Genetic Engineering

Principles and Practice

Second Edition

About the Author

Sandhya Mitra retired as Professor from the Department of Biological Sciences, BITS Pilani, with over 30 years of teaching experience. She graduated from Barnard College, New York, and Presidency College, Kolkata, and was the first PhD student and Research Assistant of J Herbert Taylor of Columbia University, New York City. Her eminent teachers were Leslie Dunn and Theodosius Dobzhansky, also of Columbia University. She has worked with A E Mirsky of the Rockefeller Institute of Medical Research, New York City.

Prof. Mitra has authored several books like *Genetics: The Blueprint of Life, Molecular Genetics: Concepts and Development, Vols. 1 and 2,* and *The Business of Living: An Acquaintance with Biology*.

GENETIC ENGINEERING

Principles and Practice

Second Edition

Sandhya Mitra
Retd. Professor
Department of Biological Sciences
Birla Institute of Technology and Science (BITS) Pilani
Rajasthan

McGraw Hill Education (India) Private Limited
NEW DELHI

McGraw Hill Education Offices
New Delhi New York St Louis San Francisco Auckland Bogotá Caracas
Kuala Lumpur Lisbon London Madrid Mexico City Milan Montreal
San Juan Santiago Singapore Sydney Tokyo Toronto

McGraw Hill Education (India) Private Limited

Published by McGraw Hill Education (India) Private Limited
P-24, Green Park Extension, New Delhi 110 016

Genetic Engineering: Principles and Practice, 2e

Print Edition
ISBN-13: 978-93-392-0353-5
ISBN-10: 93-392-0353-4

ebook Edition
e-ISBN-13: 978-93-392-0354-2
e-ISBN-10: 93-392-0354-2

Managing Director: *Kaushik Bellani*

Head—Higher Education (Publishing and Marketing): *Vibha Mahajan*

Senior Publishing Manager (SEM & Tech. Ed.): *Shalini Jha*
Associate Sponsoring Editor: *Smruti Snigdha*
Senior Editorial Researcher: *Amiya Mahapatra*
Senior Development Editor: *Renu Upadhyay*
Manager—Production Systems: *Satinder S Baveja*
Assistant Manager—Editorial Services: *Sohini Mukherjee*
Assistant Manager—Production: *Anjali Razdan*

Assistant General Manager (Marketing)—Higher Education: *Vijay Sarathi*
Product Specialist: *Richa Khanna*
Senior Graphic Designer—Cover: *Meenu Raghav*

General Manager—Production: *Rajender P Ghansela*
Manager—Production: *Reji Kumar*

Typeset at The Composers, 260, C.A. Apt., Paschim Vihar, New Delhi 110 063 and printed at AVON Printers, Plot No, 16, Main Loni Road, Jawahar Nagar, Industrial Area, Shahdara, Delhi - 110 094.

Cover Printer: SDR Printers Pvt. Ltd.

RZCYCRLVDACLB

Visit us at: www.mheducation.co.in

Dedicated with love

to my students
searching the bylanes for the essence of life
and

to my precious grandchildren

Abhi Ranjan Mitra
and

Ananya Mavmita Ghemawat
who may enjoy the harvest of the search.

Contents

Part III: Appendices

Preface to the Second Edition

The field of genetic engineering has expanded beyond comprehension in the years since the publication of the first edition of this book. Hence, the necessity for a revised and augmented second edition.

The sphere of genetic engineering can range from plants to cover both animal and human life. It can, for example, crossbreed the production of animals and stimulate the growth of healthier and stronger species of milk-producing animals and farm animals which can survive better in conditions of extreme weather, food shortage or tiring work. Regeneration of vital human organs to replace diseased and malformed ones is another example of how genetic engineering can prove advantageous to human beings.

The prospects of the scope of genetic engineering are boundless and the opportunities are widening with every day of research. The Human Genome Project to classify all the genes in the human species is an extraordinary effort to define the complete structure of the deoxyribonucleic acid (DNA), the human genetic material, and understand its functions. This study in human genetics targets the resemblances or differences of human beings from each other.

Molecular biologists have discovered various enzymes which can alter the DNA structure in living creatures. Using them, scientists can cut the specific genes from the DNA and build custom-made DNA. With this information, they can alter the genes of living organisms.

Genetic engineering has opened up innumerable ways of altering the genetic structure of a cell. The obvious inducement to do so is not just for scientific curiosity, which, of course, is always of primary value, but for the promise of intervening for understanding the molecular basis of diseases, particularly of the inherited ones.

The material of the first edition has been rearranged into more meaningful categories by my young friend Mr Manoj Kannan.

The present edition aims to introduce the student and the interested layperson to intricacies of manipulating biological systems for the benefit of humankind and for satisfying his/her urge to learn the fundamentals of the living world.

A later edition of the book will describe the various strategies employed to create stem cells, differentiate stem cells into specific tissue cells and ways and means employed currently to create transgenic organisms, both plants and animals.

I am grateful to Mr Manoj Kanan for re-arranging the original contents of the first edition of this book and to Dr Pankaj Hiradhar for pedagogy enhancement. I am also grateful for the help and cooperation of Amiya Mahapatra and Renu Upadhyay of McGraw Hill Education, India.

Sandhya Mitra
Gran de Gracia, 88, Apt 1
08012 Barcelona,
Spain
sandhya.mitra@gmail.com

Publisher's Note

The publisher, McGraw Hill Education (India), is pleased to present the second edition of the book with the following enhanced features:

- Excellent curricula coverage with detailed theory
- Reorganized content as per latest syllabi needs
- Techniques to culture eukaryotic systems
- Basic immunological techniques
- Pedagogy includes
  - 62 Laboratory Exercises
  - 75 Review Questions
  - 50 Objective-type Questions

At this stage, the publisher also wishes to thank the following reviewers for their comments and suggestions regarding improvement of the text:

Rohini Muthuswami	School of Life Sciences Jawaharlal Nehru University (JNU) New Delhi
Pankaj Hiradhar	Veer Narmad South Gujarat University Surat, Gujarat
P Sundaresan	Aravind Medical Research Foundation Aravind Eye Hospital No. 1 Madurai, Tamil Nadu
G Vijaya Lakshmi	Chaitanya Bharathi Institute of Technology (CBIT) Hyderabad

McGraw Hill Education (India) invites suggestions and comments from you, all of which can be sent to *info.india@mheducation.com* (kindly mention the title and author name in the subject line).

Piracy-related issues may also be reported.

Preface to the First Edition

Genetic engineering, like all other engineering, is the art of the possible based on insufficient information. Conventional engineering projects utilize the available or fabricated components to fashion products of choice. In the case of biological engineering, the components are taken or redesigned from the living system, or synthesized *in vitro,* and the know-how is that accrued from studies of the molecular biology of DNA. The products as well as the process designs are of value to research and application.

Genetic engineering is the manipulation of genetic information for achieving a predetermined objective. A new information or a modified native one is introduced into a target cell that is consequently transformed genetically. The transformed cell, whether a single one or part of a whole organism, provides a new genetic quality to the recipient. The latter is said to have been 'genetically engineered' to possess a novel or altered phenotype.

The fundamental technique that makes this feasible is embodied in Recombinant DNA Technology, which allows strips of DNA to be cut, joined, altered or otherwise manipulated *in vitro.* The hybrid or recombinant DNA is amplified to amounts large enough to be handled by conventional methods of chemistry and physics, It is this rDNA that is introduced into target cells. The products due to the imported DNA may be recovered for further use, or allowed to remain in the target cell, tissue or organism with a view to transforming them to exhibit a desired quality.

The potential of gene manipulation has revolutionized not only the strategies for investigating biological phenomena, but also those used for exploiting this potential in diverse fields of applications ranging from agriculture, medicine, fermentation-based industries to those that combat environmental pollution, that extract minerals from low-grade ores and industrial wastes and those that envisage futuristic use in areas such as bioelectronics. Potential practitioners of genetic engineering are, therefore, not limited only to those dealing directly with the life sciences. Indeed, this fact is obvious from the prominence gained by Biotechnology in recent times. Biotechnology represents all technologies based on living organisms or their products. The current usage is mainly intended for technologies that utilize recombinant DNA cells or their products. Biotechnology represents the scaling up of bench-level genetic engineering projects to that of industrial production.

The training of contemporary science and engineering undergraduates would not be complete without familiarity with genetic engineering and its scope in areas interfacing with biology, chemistry and physics. This was the rationale for

introducing a course in elementary Genetic Engineering for the first degree level, at the Birla Institute of Technology and Science, Pilani, Rajasthan. The prerequisite for this course is the common two-year foundation courses taken by all science, engineering and pharmacy majors. The foundation courses include one in elements of molecular biology and one in computational techniques.

This is sufficient preparation to assimilate the essential features of recombinant DNA technology and allied techniques and methodologies, including an ability to use the computer as an aid in sequencing and analysis of DNA.

As no inexpensive textbook was in the market, especially the one which offered a global account of this new technology, a set of lecture notes was prepared for the BITS course. These notes were progressively modified, supplemented, and reorganized as the infrastructure and expertise increased over time. This book is the final shape taken by the lecture notes.

The book is divided into two sections: Principles and Practice. The first section contains four chapters, the first one of which provides a thumbnail sketch of genetics, both molecular and classical. **Chapter 2** describes the theoretical aspects of gene manipulation, and **Chapter 3** provides a sketchy account of the diverse areas of man's endeavours that profit from the use of rDNA technology. **Chapter 4** highlights some of the aspects of use of genetically engineered organisms that have caused an unnamed dread in the minds of the general public. The new technology has advanced at a speed that is too rapid to allow a leisurely assessment of its impact on man, his society and environment. A practitioner of genetic engineering should be aware of potential hazards and influences in a realistic manner, and be able to interface with the layman to present both facts and fallacies of genetic engineering in a proper perspective.

The Practice section of the book contains five chapters. **Chapter 5** contains a cursory introduction to the organisms that are chiefly encountered in this basic course. This becomes especially necessary due to the input of students from a heterogeneous background. Knowing something about the creatures that are utilized for recruiting or introduction of genetic material makes sense. A few exercises are included that provide easy directions for handling these species. **Chapter 6** provides brief introductions, accompanied by exercises, for handling DNA, RNA and proteins from the viewpoint of an rDNA technologist. **Chapter 7** is confined to a few common procedures for introducing DNA into bacterial, yeast and plant cells. **Chapter 8** deals with background material for exercises in plant and animal tissue culture. Exercises for generating plant calluses, single protoplasts, regeneration from single cells and somatic cell hybrids are also included. This chapter also contains the rudiments of animal cell culture, with practice exercises to start a fibroblast culture, to passage cultured human cell-line cells and lymphocytes and a protocol for fusion of animal cells. **Chapter 9** introduces a few basic immunological techniques. Immune system cells and products have become essential diagnostic reagents that are integral parts of genetic engineering and application-oriented projects. Methods of immunological practice are sufficiently different from those of routine biochemical ones, to warrant inclusion in a basic course in genetic engineering.

With the limited perspective of a one-semester introductory course in view, those exercises that require the use of radioactive reagents have been omitted.

Those interested in pursuing more ambitious programmes, including preparation of radioactive probes, gene expression in cell-free systems or sequencing of DNA are advised to refer to excellent publications listed in the bibliography at the end of this book.

The scope of this book does not permit the inclusion of introductory exercises in computer-aided DNA sequencing and analysis. Such exercises have been undertaken in some semesters by interested students, as term-paper projects.

In its present reorganized form, the book addresses itself to at least four populations of users. It is primarily for undergraduate or first-degree students who need a preliminary theoretical introduction, and hands-on experience, in areas of algorithms and techniques of genetic engineering. Secondly, it can be utilized by research students, who have not been exposed to the entire spectrum of the body of knowledge underlying genetic engineering. The third category of users is intended to be college and high-school instructors who may find themselves out-of-depth in dealing intelligently with syllabi that are increasingly enriched with facts and jargon from rDNA research. Finally, there are the educated laymen, who desire an overall understanding of a technology that is constantly mentioned in the media and is being interfaced with several domains of man's endeavours. It is hoped that this book will place the reader in a position to be able to appreciate a technology that is galloping to unravel the baroque intricacies of biological development as well as to open unpredicted vistas in biotechnological applications.

The exercises in the Practice section may be undertaken by those who have access to a laboratory having the minimum additional inputs for microbiological work. They may be perused equally profitably by those who would like to know about the methodologies employed in rDNA investigations. It is well known that one who 'reads' cookbooks often gains as much, if not more, than one who actually tries out recipes in the kitchen.

I now come to the pleasant job of thanking the innumerable persons, as without their sympathy and advice, this venture would have been aborted long ago. The preparation was precipitated by the introduction, at short notice, of the course in Genetic Engineering, at BITS, Pilani, in January 1983. The prime mover, against insurmountable odds, was Dr C R Mitra, Director of BITS at the time and till 1989. I am indebted to him for the facilities extended to me and my colleagues for setting up the laboratory and the ancillary infrastructure. Dr Mitra's insistence and far-sightedness in including a basic foundation course in Molecular Biology and the elective course in Genetic Engineering in the programme of all science, pharmacy and engineering first degrees (MSc Hons., BE Hons.) has been vindicated by the success of graduates who have entered areas of higher studies and research in frontiers of biology and interdisciplinary areas, irrespective of the nature of their first degree.

It is impossible to express my gratitude to every single person to whom I am indebted. Nevertheless, I would like to mention with appreciation the names of some who came to my rescue with material, know-how and suggestions for various aspects of the laboratory exercises. They include Dr N K Notani and his colleagues; Dr S K Mahajan and Dr S Joshi of Bhabha Atomic Research Centre (BARC), Bombay; Dr H K Das and Dr Anjali Mukherjee of the Jawaharlal Nehru University

(JNU) New Delhi; Dr S K Sen and Dr S K Ghosh, Bose Institute, Calcutta; Dr Sudhanshu Das, Chittaranjan National Cancer Research Centre, Calcutta; Dr and Mrs (Dr) P K Bose, Department of Anthropology, Delhi University, and Dr Y Mohan Ram, Department of Botany, Delhi University; Dr Bharati Bhatt and Dr G P Phondke, BARC, Bombay; Dr A B Mitra, Cancer Research Centre, Maulana Azad Medical College, New Delhi; Dr S G Gangal, Tata Memorial Cancer Institute, Parel, Bombay; Dr Archana Sharma, Calcutta University; Dr R P Sharma and Dr R Majid, Nuclear Research Laboratory (NRL) at the Indian Agricultural Research Institute (IARI), New Delhi and Dr Z Lobo, Tata Institute of Fundamental Research (TIFR), Bombay. I especially thank Dr A B Joshi and Dr R H Das of the CSIR Biochemicals, New Delhi. Their troubleshooting suggestions, given generously, have been invaluable. Dr Raka Hari also graciously read certain sections and suggested useful amendments or corrections.

This is also an opportunity to express my gratitude to Dr V Allfrey, who introduced techniques of cell fractionation using the ultracentrifuge, at the Rockefeller Institute of Medical Research, New York, in the laboratory of Dr A E Mirsky and to Dr J H Taylor then at Columbia University, New York, for initiation into autoradiography and molecular biology in the early 50s.

I now thank my colleagues at BITS: Dr S Venkateswaran, the present Director and formerly the Dean, Educational Development Division; Professor I J Nagrath for resolving many ticklish bureaucratic situations, both as Dean, Instruction Division, and Administrative Dean; Dr G P Srivastava for arranging the logistics for the production of the revised form of the lecture notes; Dr R K Patnaik, Dean, Instruction Division, deserves special thanks for persuading me to collect the rambling notes and protocols into a consolidated book form. With three other textbooks just completed, I had no intention, of my own, to become embroiled in one more.

But my love and appreciation go especially to my students who not only helped me try out different protocols and standardize newly introduced ones, but also generally provided, with eagerness on their faces, a visible incentive at all times. I record my thanks to the following in particular: Rupak Banerji, Chandrasekhar Anand, Ajay Kapur, Ritu Nagrath, Suresh G, K L A Shastry, Vikas Chandhoke, S Ganesh, Ganesh Anand, Ramakrishna, S, K E Surekha, Reeta Prusty, Lakshmikiran, Sudha Sivaraman, Abraham Kavoor, Sudhir Kumar and Amitav Mukherji.

I extend my thanks to Messrs. Sardara Ram Saini, and R C Sharma and Ram Singh Saini for typing the manuscript and Mr B S Kudesia for organizing their output. I am grateful to Mr Naurang for keeping stability at the home front.

Closer to home, I am grateful to my family for sparing me from routine household obligations during the development of the course, the laboratory and the book. I thank my son, Dr Amitava Mitra, for general marshalling of the contributions of typists and drawing office personnel and for moral support throughout the venture. I express gratitude to my daughter, Dr Anuradha Mitra Ghemawat, for encouragements from across the seas; they have been as valuable as the chemicals and other tangible material for the course that she has supplied me gratis for the last six years. I record my gratitude once more to Dr C R Mitra, my husband, for boosting flagging enthusiasm at times when ill-health threatened the successful completion of the project.

This book is only a preliminary introduction to the art and craft of genetic engineering. The exercises may be preplanned and scheduled to allow completion in one semester. Long-term ones, such as tissue culture of plants and animal cells, and immunizing protocols, may be undertaken at the start of the semester, and sandwiched with exercises of culturing of *E. coli* for DNA isolation: The other exercises may be spread out according to convenience and available facilities. Assignments, based on reports of investigations in journals, and open-ended laboratory, are to be included to provide as much background and familiarity with the jargon that is possible in the period of one semester. The invaluable texts and manuals by Old and Primrose, Winnacker, and Maniatis, Fritsch and Sambrook should be available for consultation and instructions for more ambitious exercises.

I have received facts and know-how from many, but none of them are to be blamed for mistakes in representation or interpretation in this text. The errors are solely my own. I would be immensely grateful if they are brought home to me.

Sandhya Mitra

PART I: *The Principles*

Foundational Genetics 1

Introduction

Genetic engineering is not an area of knowledge that can be learnt in isolation, since its existence is only in the context of a larger and growing matrix of the science of genetics. So, before we proceed with the basis of molecular genetic technologies and recipes for achieving them, we must acquire an overall picture of the genetic material that has to be appreciated and manipulated.

1.1 ❑ CLASSICAL GENETICS IN A NUTSHELL

Gregor Mendel studied the inheritance patterns of several traits of garden varieties of sweet peas *(Pisum sativum)* and came to the conclusion that heredity was a function of discrete 'factors' that were transmitted from parents to offspring. Since the existence and behaviour of chromosomes in the cell were not known during his time, it was one of the greatest of feats to arrive at such a concrete concept from analysis of statistical data.

The above concept also postulated that genes occur in pairs in every individual (remember only eukaryotic species were well known at that period). A gene is capable, he stated, of being present in two alternate states. For instance, the gene for 'height' (a *trait*) can exist as one for 'tallness' and also as one for 'dwarfness'. Mendel proposed that one of these two alternate states or *alleles* of a gene may be *dominant* over the other. This implies that even if one of the two alleles of the same gene in an individual is a dominant one, the trait exhibited by the individual will be the dominant one. The unexpressed allele, i.e., the *recessive* one, is expressed as a trait only when present in a state where it is not overshadowed by a dominant allele. Mendel demonstrated that only individuals with both alleles of a gene recessive, will exhibit the recessive trait. The two alleles in an individual may be of the same kind (dominant or recessive), in which case the individual is said to be *homozygous* for the allele. If the pair of alleles consists of one each of the dominant and recessive ones, the individual is *heterozygous* for the gene.

1.1.1 Monohybrid and Polyhybrid Crosses

The first type of experiment performed by Mendel involved the mating (crossing) of

a pea strain homozygous for a dominant allele with one homozygous for the recessive allele of the same gene (Fig. 1.1). All the plants that grew from the seeds of this cross exhibited the dominant trait. When these Fl (Filial generation 1) plants were self-crossed and the plants from the seeds cultivated, the F2 generation exhibited two types of traits, dominant and recessive, with the former being about two-third in number of those with the recessive trait. When the plants showing the dominant trait were self-crossed, the F3 generation produced some (of the total population) that were true breeding dominants (i.e., they continued to produce only dominant trait plants when self-crossed), some (one-fourth) were true breeding recessive ones, while half the population were not true breeding. On self-crossing, the latter again exhibited a spectrum of traits (1:2:1) in the next generation. These data led Mendel to propose that one true breeding parent was homozygous for the dominant allele (written with a capital letter, such as T for tallness) and the other was a recessive homozygote (written with a lower case letter, such as t for 'dwarfness').

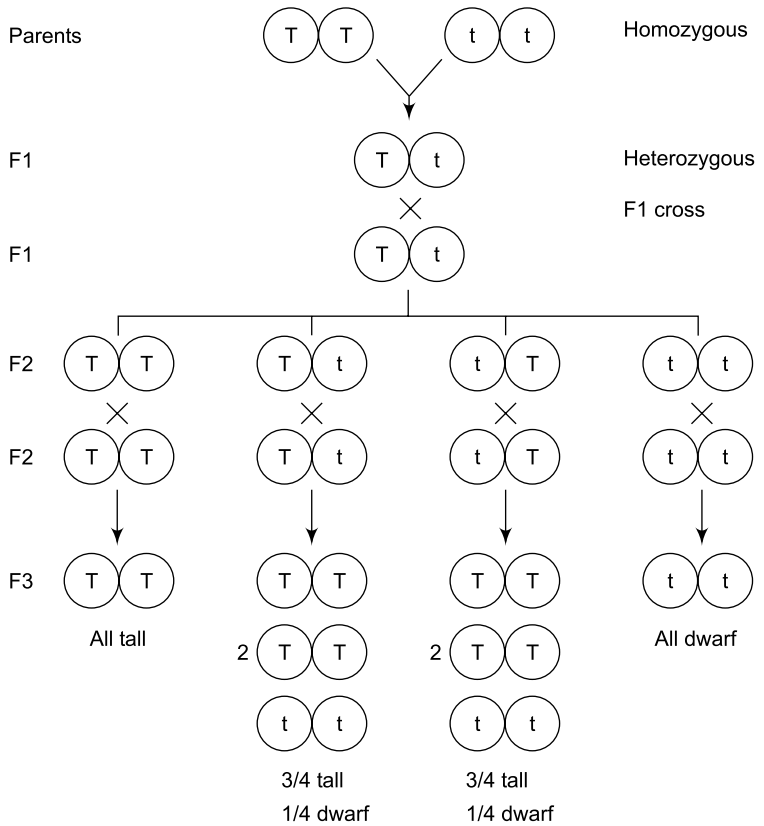


Fig. 1.1 *A monohybrid cross schematic. T and t represent alleles for the trait height. T for tallness is dominant in this example while t for dwarfness is the recessive allele; F1, F2 are filial generations 1 and 2 and P is the parental generation.*

The Fl result is explained by the suggestion that the two alleles in each parent become separated into different gamete cells, each of which is equally capable of fertilizing one from the other parent. In terms of the preceding example,

$$T\,T \times t\,t$$

results in gametes T and T in one parent and t and t in the other. The union of a gamete from each parent can then result in only Tt individuals. All these individuals will be tall, which is what was found. A distinction was made between the internal genetic constitution and the outwardly observable appearance or trait. The former is called the *genotype* and the latter the *phenotype.*

The Fl × Fl crossings, therefore, resulted in two types of gametes from each parent, T and t ones. The union of T and t with T and t in a random manner can only result in TT, Tt, tT and tt genotypes. The TT and Tt individuals (three-fourth of the population) are tall while only one-fourth of the total are dwarfs. These are the tt homozygotes.

The preceding simple experiment, using different genes, led Mendel to postulate his theory of heredity: traits are represented by factors that are discrete and do not blend in any individual. Further experiments using combinations of alleles of different genes (e.g., tall and green × dwarf and yellow) provided another dictum: the factors for different traits are inherited independent of each other.

The foregoing landmark investigation remained obscure until 1900, when it was rediscovered independently by three persons, de Vries, Correns and Tschermak. By this time the interior of the cell was known cytologically; the movements of the chromosomes during cell division suggested that Mendel's factors (now renamed as 'genes') were borne on (or were synonymous with) chromosomes (Fig. 1.2).

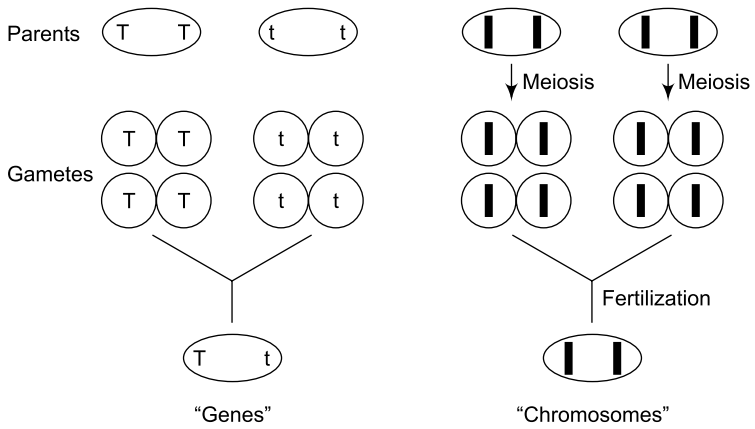


Fig. 1.2 *An obvious parallel in the behaviour of a gene (alleles), as predicted by Mendel, and chromosomes, as discovered by cytologists.*

1.1.2 Linkage and Crossing Over

Extension of Mendel's monohybrid (using alleles of one gene) and dihybrid (using crossings with two genes) crosses by others brought out exceptions to Mendelian

patterns of inheritance. It was discovered that some genes were 'linked' on the same chromosome. This linkage was also found to be breakable by crossing over. It was accepted that crossing over was brought about by a physical exchange of homologous regions of chromosomes. The molecular details were made available, however, only after the birth of molecular genetics.

1.1.3 Sex Chromosomes and Autosomes: The Chromosome Theory

Early studies in Mendelian genetics, using breeding procedures, revealed also that in animals there may be chromosomes that are different in the two sexes of species. The X and Y chromosomes of mammals were thus discovered. Females of mammals possess two X chromosomes per cell; the male of the species carries one X and one Y instead. The remaining chromosomes, called *autosomes,* are identical in both the sexes. For instance, the human female has 22 pairs of autosomes and one pair of X chromosomes, while the male possesses 22 pairs of autosomes and an XY pair. In some species, the female is the heterogametic one (with different sex chromosomes).

By the end of the second decade of the twentieth century the Chromosome Theory was enunciated by Thomas Hunt Morgan and his students. A. H. Sturtevant, one of Morgan's students, developed the method of mapping relative positions of genes on a chromosome, using recombination values between gene loci.

1.1.4 The Mutation Theory

In the 1920s H J Muller discovered that physical agents, such as X-ray radiation, cause heritable changes in the genotype. The theory of mutation was born. The allele of a gene was demonstrated to change to another one of the same gene as a result of X-ray treatment. The test organism was *Drosophila melanogaster.*

What was more interesting was the fact that the X-ray induced mutants had counterparts in the stocks of *Drosophila* collected from nature. This led Muller to state that in nature mutations are brought about by environmental agents. This suggestion immediately gave a foundation for Charles Darwin and A R Wallace's theory of natural selection in evolution, proposed almost a century ago. This theory suggested that the existing species are derivations from variants present in previous generations. At the time of proposal of this theory there was no suggestion as to how the variation was produced in the first place. Muller's theory of the origin of mutations provided this link. The mutation theory was responsible for amending Darwin's theory into what become known as 'neo-Darwinism'.

1.1.5 Extensions of Mendelism

Extensions of Mendelism include the following:

1. A gene may possess more than two types of alleles.
2. An allele may not be only dominant or recessive. The two alleles may be codominant or expressed in the same cell (e.g., the blood group alleles M and N produce both antigens in an MN individual).
3. Recombination may occur between non-homologous DNA. This is the process underlying transposition.

4. Methods of making physical maps of genes on chromosomes are now available. The exact locations and not relative distances between gene loci can now be assigned.
5. There are traits which are specified by several genes. These polygenes exhibit a quantitative nature in phenotypes.

Although molecular biology has enhanced our knowledge of genetical phenomena in an extraordinary manner in a brief period, it is still the basic tool of classical genetics that are utilized to handle problems of heredity. Matings or crossing of viruses, bacteriophages, bacteria, yeasts, fruit flies and mice, among other species with a short generation time, are still one of the chief methods in exercises linked with more modern approaches. Since the Mendelian terms are in use, it is necessary to be familiar with them.

To recapitulate, alleles are alternate states of a gene. These alleles are relics of mutations that have occurred. The rate of mutation may be enhanced by artificial mutagens, so that there are large number of mutants that may be utilized for studies. In all genetical studies, the function of a *wild-type* (synonymous with normal or most predominant in nature) is deduced initially by analysing the defect due to a mutant allele. In the diploid somatic cells of eukaryotes, a pair of chromosomes with identical gene-loci positions, may contain the same allele of a gene or contain a pair of different alleles. A gene locus may exhibit more than two alleles. The genotype is the allelic constitution of an individual, the phenotype, the expression of the genotype.

The phenotype may be influenced by the environment. This is especially true for polygenically determined traits.

The alleles of a gene may be written in different ways; *A* and *a* for dominant and recessive alleles respectively; a^+ and a^- for wild-type and mutant alleles, lac^+, lac^- for wild-type and mutant alleles or simply + for the wild-type allele and a name (e.g., *w*) for the mutant allele.

Bacteria are crossed by mixing the strains differing in the alleles for the gene in question, and plating the mixture on a growth medium that can select recombinants. Viruses and phages are crossed by infecting the same host cell with both viral strains, and again selecting for recombinants on a suitable culture medium.

1.2 ❑ MOLECULAR GENETICS IN A NUTSHELL

Genetics is the science of heredity. It began with the analysis of patterns of inheritance of observable traits, but eventually developed into studies of all aspects of the nature and activities of the hereditary material. It is the hereditary material that is utilized in the lifetime of an organism as a blueprint for the synthesis of proteins—the basic building block molecules, and catalysts, that mediate biochemical reactions and are dispensed in exact amounts to future generations.

The hereditary material was discovered to be the nucleic acid DNA (deoxyribonucleic acid) in the majority of species. In some viruses, it is represented by an allied substance RNA (ribonucleic acid).

The compositions of both DNA and RNA were known in the first half of the twentieth century, but the manner in which their component units are assembled

into the final molecule was suggested only in 1953 by J D Watson and F H C Crick. Classical genetics, that draws conclusions from results of breeding experiments, was concerned with the manner in which traits are inherited and mutated by artificial insults. With the identification of the underlying material that dictates the basis of traits, it became possible to design more objective studies for understanding heredity and, eventually, for introducing modifications by the manipulation of DNA in order to achieve the desired characteristics of an organism.

The model of DNA, proposed by Watson and Crick (Fig. 1.3), became a landmark in the history of science, for it provided an inkling of, for the first time, how the DNA is utilized in the lifetime of a cell or an organism of several cells. It is now possible to visualize how the DNA replicates into exact daughter DNA molecules—a feat that is essential for any material which has to be transmitted exactly to daughter cells. The above mentioned model suggests how the DNA molecule is the repository of genetic information. Such information is encoded in a cryptic form as units of three nucleotides, which are the basic monomers linked to form the polynucleotide DNA.

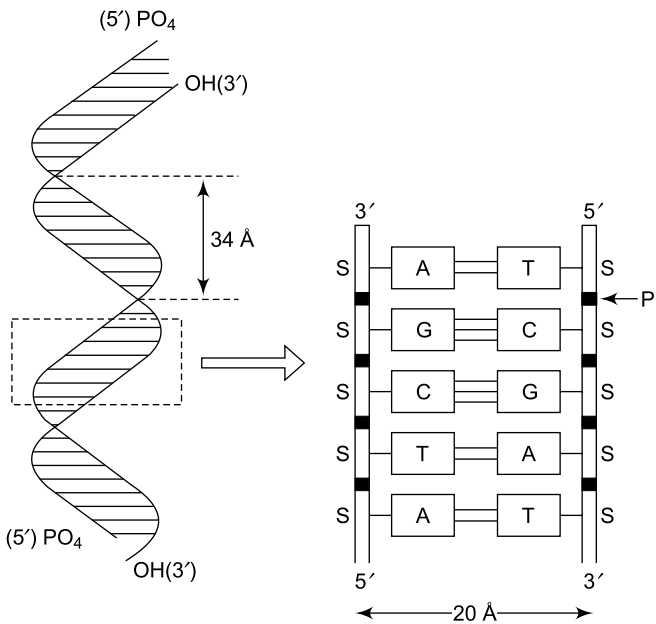


Fig. 1.3 *Diagrammatic representation of the DNA double helix model. A(adenine), T(thymidine), G(guanine) and C(cytosine) are the four bases that distinguish the nucleotide units of a DNA strand. The purines A and G base-pair complementarily with T and C, respectively. This is the B form of DNA which is coiled in a left-handed direction.*

One of the discoverers of the DNA double helix model, namely Crick, went on to decipher the coding system. A codon was found to be a set of three nucleotides, representing one of the amino acids, that are linked together to form the macromolecules called proteins (Fig. 1.4). The code was eventually cracked by Nirenberg, Mathei and Hargobind Khorana, and a chart showing the 61 codons representing the 20 known amino acids was constructed.

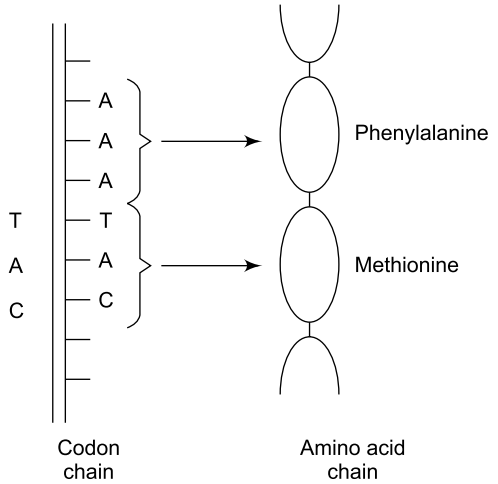


Fig. 1.4 *Each amino acid residue in a polypeptide is encoded by a triplet of nucleotides (called a codon). Of the 64 codons possible from A, T, G, C, 61 represent amino acids. The remaining three are stop codons (UAG, UAA and UGA).*

The discovery of the coding system and the use of codons in heredity provided a molecular explanation for mutations, which are heritable (permanent) alterations in a genetic trait. An alteration in one or more bases of a codon may change the specificity of the latter, resulting in an altered amino acid in the ensuing protein. Mutations had already been designated as the raw material for evolution; they provide variations in a genetic population that may be exploited differentially in diverse environmental niches. A molecular explanation of mutation suggested how such evolutionary events might take place. In addition, the manner in which mutations arise can now be appreciated with reference to the structure and activities of DNA.

The plethora of investigations that were initiated by the discovery of the DNA model soon clarified the nature and boundaries of a *gene.* The concept of the gene, the central theme of genetics, had undergone several verbal mutations until the molecular definition was presented by Seymore Benzer. A gene is the length of a DNA strand (most DNA are double-stranded, each strand with a different series of nucleotides, and hence dissimilar in terms of genetic information) that contains all the bases required to dictate the synthesis of a particular polypeptide (basic unit of a protein). This length of DNA includes a set of codons (coding sequence) that corresponds to the amino acid sequence of the polypeptide to be synthesized and other sequences that are required for the copying of the gene into RNA molecules which, in turn, are decoded into the polypeptides. The non-coding sequences in a gene are thus mainly the ones that regulate the 'expression' of a gene. A gene shorn of all non-coding sequences cannot be utilized for protein synthesis.

The manner in which information flow between the message in the gene and the final product is achieved was also revealed. The DNA is first transcribed into

a complementary RNA (Fig. 1.5). This phase is known as *transcription* and is mediated by an enzyme, namely, RNA polymerase. The second phase is *translation* and is carried out in organelles known as ribosomes, that are present abundantly in the cytoplasm. Regulatory sequences in the gene are utilized to dictate when, where and in what quantity a particular RNA may be synthesized, as well as to modulate the translational step.

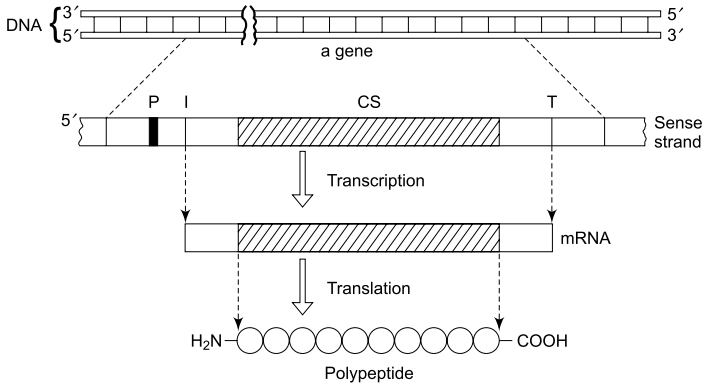


Fig. 1.5 *Information flow from a gene to a polypeptide through an intermediate RNA transcript.*

The RNA destined to be decoded into a polypeptide carries the 'message' of the gene, and hence is the *messenger* RNA or mRNA. There are two other types of RNA molecules (Fig. 1.6). These are the *ribosomal* RNAs or rRNAs and *transfer* RNAs or tRNAs, that are parts of the translating apparatus and are not decoded into proteins. A few other small RNAs (snRNAs), present in the eukaryotic nuclei, have also been discovered. They are involved in unique activities of the cell. snRNAs are also not translated into proteins.

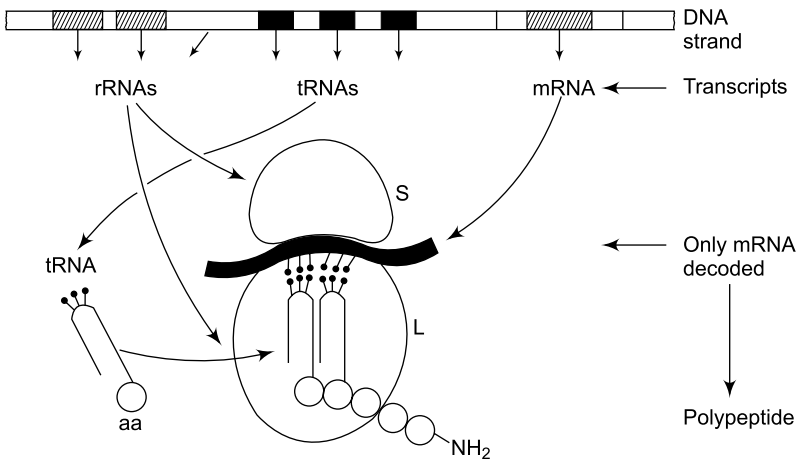


Fig. 1.6 *The three types of RNA molecules, ribosomal (rRNA), transfer (tRNA) and messenger (mRNA) and their roles in protein synthesis.*

The discovery of the two basic operations involving the DNA—that is, DNA replication (Fig. 1.7) and transcription (or DNA and RNA synthesis)—paved the way for understanding other biological phenomena, such as recombination, establishment of mutation and repair of damaged DNA. All these events use DNA synthesizing operations. Finer searches into these processes revealed that variations of the basic synthetic operations rely on variant polymerases. Thus, DNA Pols I, II and III were identified and characterized. The first one is the enzyme for routine DNA synthesis, and the last one is for synthesis of short lengths of DNA. The latter is required during the processes of DNA recombination and repair.

Recombination of traits of parents or ancestors in offspring is a known fact. In eukaryotes or higher organisms reciprocal exchange of genes is a regular event in an organism's life cycle. In prokaryotes, or primitive organisms that include bacteria and blue-green algae, there is no such regular event. However, there exist four parasexual methods of recombination that may occur in some individuals of prokaryotes. These are the only chance opportunities for a prokaryote to acquire a trait from another individual. The parasexual modes are made possible during *transformation, transduction* and *conjugation.* In transformation, naked DNA may enter a cell and recombine with the resident DNA. In transduction, a piece of DNA may be transferred from one cell to another by a parasitic organism—a bacteriophage (virus of bacterium). In conjugation, two cells become connected by a tube, through which the DNA from a donor may be transferred to a recipient cell.

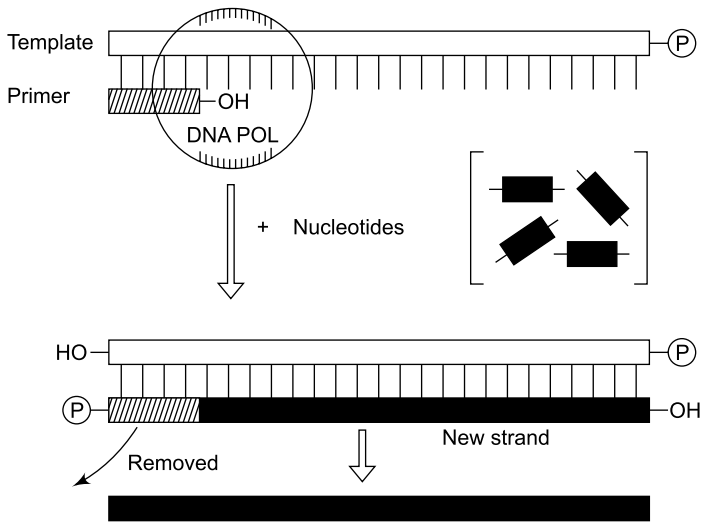


Fig. 1.7 *Schematic representation of DNA replication. The two strands separate at a region called the origin (ori) of replication. A new strand, complementary to each parental strand, is synthesized with strand elongation always in the 5′ –3′ direction. As the two strands of a ds DNA are antiparallel, the daughter strands are synthesized in opposite directions. One new strand is elongated continuously, while the other is synthesized as short lengths (Okazaki fragments) that eventually join to form the continuous new strand. Strand synthesis begins from a primer (a short length of RNA complementary to the template) which is removed and the gap replaced by an equivalent length of DNA.*

In all these instances the incoming and resident DNA recombine only in regions that are homologous (that is, carry the same gene sequence). In a fourth type of parasexually achieved recombination—called *transposition*—totally unrelated DNA molecules may be joined together. Genes on transposons, which are DNA sequences, copies of which may integrate in another DNA region, present on bacterial plasmids (extrachromosomal small circular DNAs) are transposed to other DNA by this process.

All species possess a basic minimum set of genes and non-genic DNA that serve as the fundamental genetic information for the species. A eukaryote (in which the genetic information is confined in a membrane-bound organelle in the cell, called the nucleus) possesses two of these sets in one phase of its life cycle and only one set in the other phase. These phases are, therefore, referred to as the *diploid* and *haploid* stages of a eukaryote. A prokaryote possesses only one set of DNA and is characteristically haploid. In some unusual circumstances, there may be more than one copy of a region of the prokaryotic DNA. Such a region makes the individual a *merozygote* for the genes in the region.

In prokaryotes, there is a single DNA for the haploid genetic unit. In eukaryotes, the total DNA of one set is distributed into two or more units and are associated with proteins to form what are known as chromosomes. The bacterial DNA, which is naked, is also referred to as a chromosome. The eukaryotic DNA is wrapped around small cores of histones at regular intervals to form a 'beaded' string. The 'beads' are known as nucleosomes, and the DNA-histone complex in general is called chromatin (Fig. 1.8). The chromatin is organized into compact chromosomes together with certain non-histone proteins (NHCP).

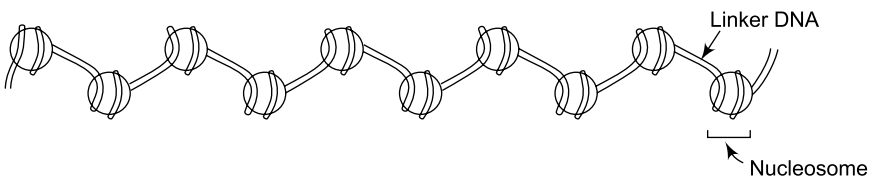


Fig. 1.8 *Schematic representation of chromatin with the bead-like nucleosomes.*

In bacteria, transmission of hereditary information is from one single celled generation to another. In eukaryotes, the haploid and the diploid phases may consist of one or more cells each. Therefore, genetic information is transmitted both *horizontally* between daughter cells in the same phase and *vertically* between individuals of different generations. Horizontal genetic transmission occurs during a process known as *mitosis*. Each eukaryotic cell has a species-specific life cycle, that consists of (i) an S or synthetic phase during which DNA and chromosomal proteins are synthesized, (ii) a G2 phase during which the daughter DNA molecules become progressively compacted from the basic straggly molecule, (iii) the M or mitotic period during which the daughter chromosomes are apportioned equally to separate cells, and (iv) a G1 phase when the new cell prepares itself for its S phase.

Mitosis results in identical chromosome sets in parent and daughter cells. There is another type of cell division that occurs before sexual reproduction in

eukaryotes. This is known as *meiosis.* It consists of two stages, in one of which there is recombination between homologous gene loci in each pair of chromosomes. In the other stage the number of chromosomes is reduced to that of the haploid set. Meiosis results in four daughter cells, with the genes recombined in each of them. These four cells are the *gametes* (egg in females and sperm or pollen in males). A male and a female gamete unite after fertilization to once more form a diploid ($2n$) cell. This first $2n$ cell is called the *zygote.* In some unicellular plant species the zygote is the only diploid cell in the life cycle. It divides by meiosis into 4 haploid (n) cells, each of which divides mitotically to generate a population of haploid stage cells. As one goes higher in the evolutionary scale, the diploid stage, brought in by mitotic proliferation of the zygote, becomes the prominent phase.

In summary, the genetic material DNA (or RNA) specifies the proteins that form structural, enzymatic, hormonal, antibody and other crucial components of a living system. Genes are arrayed on the DNA in a specific pattern in each species. Haploid species possess only one set of this genome, while diploid species carry two sets in their body (somatic) cells and one set in their reproductive (germ) cells. Regular exchange of genetic information occurs in the life of every eukaryote. Prokaryotes occasionally exploit certain parasexual modes for acquiring genes from other sources. The DNA replicates into identical daughter DNAs. The DNA is transcribed into RNAs, one species of which, the mRNAs, are translated into polypeptides.

An understanding of the manner in which DNA is expressed led to queries as to the basis for the regulation of expression of genes. Examination of life histories of viruses, bacteriophages, bacteria and higher organisms led to a confirmation of the suspicion that development is a reflection of the coordinated and/or sequential expression of specific genes (or sets of genes) in any individual, and this chronological pattern of expression is characteristic of every species (Fig. 1.9). Models of such regulation were presented. The first one, the *operon model,* by F Jacob and A E Monod, provided an insight into the molecular mechanism of one type of regulation in bacteria. It was discovered that a set of genes can be allowed to be transcribed or prevented from doing so by unblocking or blocking the region on the gene to be used by the enzyme for transcription. While this model was excellent for explaining one-shot regulation in the 20-minute life of an *Escherichia coli* bacterium faced with an environment change, it was not sufficient to suggest how a complex eukaryote (even a single celled one) manages to express its genome in a preordained and time-bound manner. Britten and Davidson presented a model for coordinate and sequential expression of proteins in response to a single stimulus. The model also suggests how different signals may result in the expression of the same set of genes. Britten and Davidson's model postulated the presence of several regulatory elements (sensor, integrator, receptor) that qualify the transcription of a 'producer' gene into one of the RNAs. Some physical correlates of this model have been discovered. They include regulatory sequences in the non-coding sequences of a gene and genes for proteins that act as regulatory molecules by binding themselves to specific regulatory sequences. There is a growing list of such transcription factors and regulatory sequences that recognize them and other molecules, such as hormones. Such discoveries are beginning to illuminate the basis of the differential expression

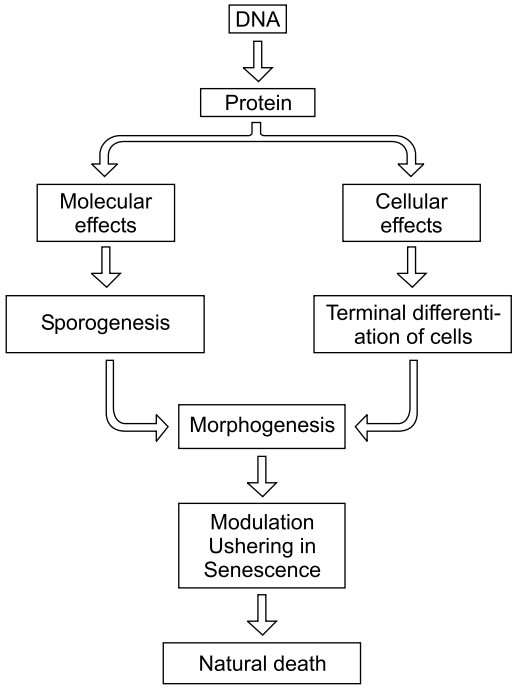


Fig. 1.9 *The DNA provides the genetic basis for developing, maintaining and perpetuating an organism.*

of genes in different tissues at different times of development, or even in the same cell at different periods of differentiation.

It has become obvious that there are certain crucial developmental genes which regulate the path of development proceeding from them, and the others which specify the exact biochemical character of what is known as the 'finished product' or 'differentiated cell'.

Enquiries into the molecular basis of development were obviously linked with those that dwelt on the causes of abnormal development. Immunogenetic, developmental, and behavioural development can go awry in several ways. The genetic basis of some of them are becoming apparent. Abnormal development includes carcinogenesis (cancer). A series of genes have been identified, the products of which mediate the biochemical pathways that link an external stimulus with expression of specific genes in the cell that result in cell division. Mutations in one or more of these cell division-regulating genes result in a 'transformed' cell that on further alteration (the nature of which is still not clear) becomes 'malignant' as well. A malignant cell has lost control over its cell division schedule and proliferates continuously.

Another area of biology that has become more tractable than before, due to the discoveries in molecular genetics, is that of evolution. It is possible now to examine in detail the pattern of amino acids in proteins and bases in DNAs by elegant sequencing techniques. Comparison of species is possible, therefore, on the

fundamental basis of their blueprints, rather than only on morphological features and fossil records. A mutation in one base may change a particular amino acid in a protein. As mutations occur in a stochastic manner, an estimate may be made of the number of mutations that are likely to show up in the gene of a particular protein in a specified evolutionary time period. A species that is advanced will be expected to have fewer substitutions of amino acids in a particular protein than a species that has emerged much earlier in the evolutionary tree. In other words, slightly variant compositions of the same polypeptide will be found in different species. This *polymorphism* of proteins has been confirmed. It was the initial enthusiastic basis for studying the evolutionary relationships between species, at a molecular level.

More recent studies, which are based on the examination of the DNA, suggest that the non-coding sequences may be more appropriate for providing the evolutionary history than the coding sequences of a protein. The argument for this suggestion is that alterations in crucial functional regions of a protein will render the latter non-functional and may result in the failure of the individual to survive and so fail to leave progeny that may be examined for the aberration. Mutations in non-coding regions, however, may be tolerated to a much greater extent, and remain as signatures of events that have occurred during the evolutionary history.

It may be mentioned here that the upsurge of studies that led to the extraordinary development of the science of biology became possible because of the tools provided by recombinant DNA techniques, the mainstay of genetic engineering.

1.2.1 DNA and RNA

DNA is a polymer of units called nucleotides (Fig. 1.10). A nucleotide is a complex of a 5-carbon sugar, a nitrogenous moiety (called base) and a molecule of a phosphate. In DNA the sugar is deoxyribose, and in RNA it is ribose.

The nitrogenous component is attached to carbon 1 of the ribose molecule. There are two classes of bases, namely, the *purines* and the *pyrimidines.* A purine is composed of two ring structures, with nitrogen as part of the structure in certain positions. There are two types of purines: guanidine (G) and adenine (A). A pyrimidine is a one-ringed molecule with nitrogen in certain positions. There are three types of pyrimidines: cytosine (C), thymidine (T), and uracil (U). The DNA is composed of G, A, C and T, while the RNA contains U in place of T.

Each nucleotide is composed of a ribose (5C) sugar that has its first carbon attached to a nitrogenous base and its fifth carbon attached to a phosphate moiety. The base may be a purine (two-ringed molecule) or a pyrimidine (one-ringed).

A nucleotide is linked to another by a phosphodiester bond between the OH group of carbon 3 of the ribose and the $–PO_4$ group of carbon 5. The backbone of the DNA polymer is thus composed of alternate sugar and phosphate molecules. The DNA strand takes on a spiral configuration as a function of the angles made by the alternate sugar-phosphate linkages.

Only some DNA viruses possess single-stranded DNA as the genetic material. All other species possess double-stranded DNA, with the two strands forming the double helical structure presented by Watson and Crick. The two strands 'face' in opposite directions; they are said to be antiparallel. However, the bases, projecting

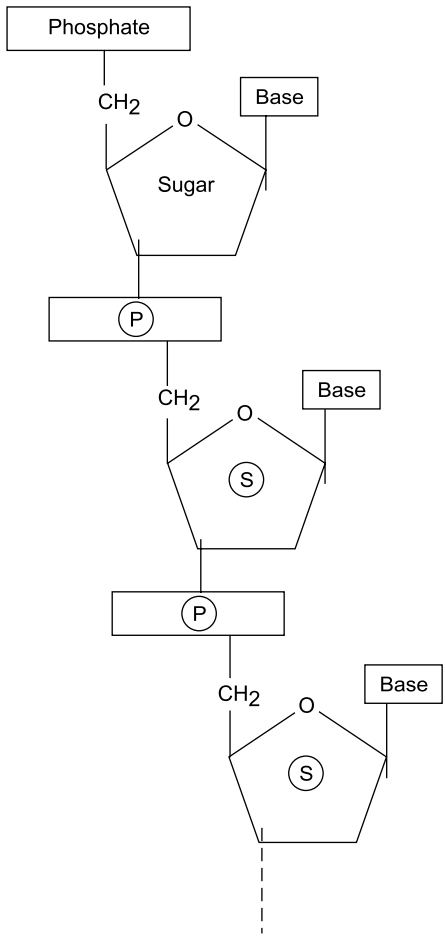


Fig. 1.10 *A single strand of DNA showing its backbone of alternate ribose sugar and phosphate molecules, and the unique base attached to each sugar.*

inwards from the two backbones, link the two molecules by hydrogen bonds. The 20 Å diameter of the DNA molecule allows only a purine to base-pair with a pyrimidine and vice versa. Specifically, only A-T and G-C pairs are formed. Although two bases are linked with only a few weak hydrogen bonds (2 between A and T and 3 between G and C), the thousands of bonds between the two strands of DNA result in a very stable molecule. Stability is also provided by an envelope of water molecules that bind themselves to the negatively charged phosphates in the backbones.

It is essential to appreciate the fact that the two strands of a DNA do not separate easily. All metabolisms of this molecule (replication, transcription), therefore, involve mechanisms that pry open the two strands in specific ways. The overall regulatory system or schedule decides which area is to be made available for use at a particular point of time.

RNA is generally a single-stranded molecule. However, certain plant viruses possess double-stranded RNA.

Single-stranded DNA and RNA are seldom stretched out fully. They base-pair in complementary stretches of the molecule to form a more compact structure (Fig. 1.11). Single-stranded nucleic acids are vulnerable to several nucleases (degrading enzymes).

Keeping much of the molecule in the double-stranded form prevents destruction by these nucleases. Even a single DNA strand of a double-stranded molecule possesses complementary regions. These become apparent when such regions become separated from the other strand of the molecule; each strand proceeds to form 'stem-and-loop' or 'hairpin' structures by base-pairing of the inverted complementary sequences (Fig. 1.12).

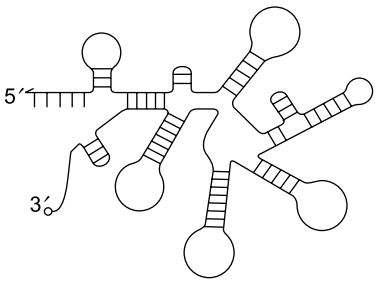


Fig. 1.11 *Secondary structure of a single-stranded RNA. Complementary regions in the RNA result in base-pairing and duplexed regions. This protects the molecule from degradation by nucleases that attack ss RNA, as well compacts the straggly chain to occupy a smaller space.*

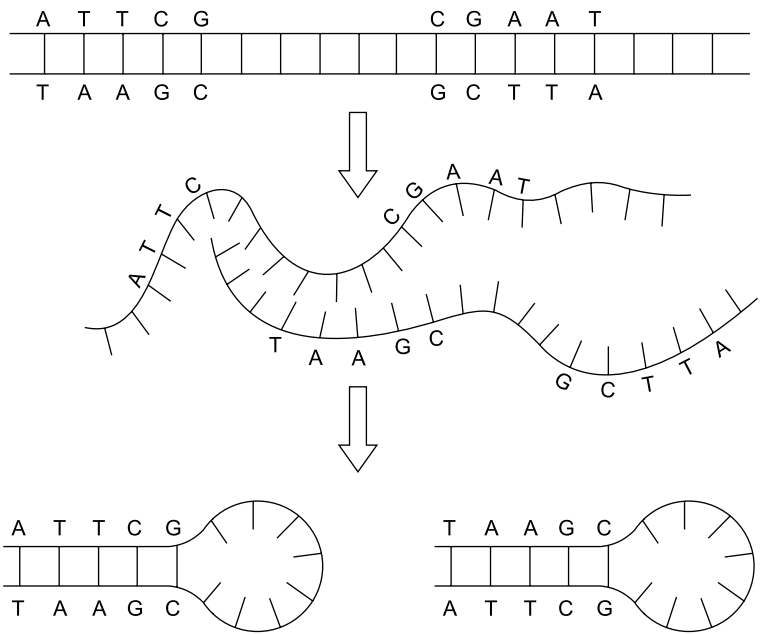


Fig. 1.12 *Palindromic sequences cause the formation of stem-and-loop or hairpin structures in single-stranded nucleic acids.*

1.2.2 Anatomy of a Gene

A gene, as mentioned earlier, is a stretch of DNA (RNA in some viruses) that contains the coding sequences for the amino acids in a polypeptide (or the complementary sequences for an rRNA or tRNA product) as well as sequence motifs recognized by

various enzymes and protein factors that are required to transcribe and translate the gene into its product. These latter are the control or regulatory sequences.

A typical simple gene consists of a coding sequence (CS) that is flanked on both sides by DNA that contains the regulatory motifs (Fig. 1.13). The most essential motif is upstream (towards the 5′-end of the DNA) of the CS and is that of the *promoter* (P); the transcribing enzyme RNA polymerase (RNA Pol) binds to the promoter before it begins transcription.

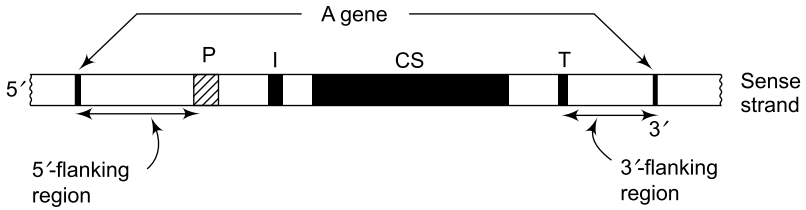


Fig. 1.13 *Anatomy of a simple gene. The coding sequence (CS) is uninterrupted, A promoter (P) and other transcription initiation/repression control sequences are present upstream (towards the 5′-end of the sense strand) of the CS. Termination (T) and other signals are present downstream (on the 3′-side) of the CS. The untranslated regions between the 5′-end and P and between T and the 3′-end of a gene are referred to as the 5′-flanking and 3′-flanking regions, respectively.*

The promoters in bacterial and viral genes usually contain a *consensus sequence* of 7 bases, named the Pribnow box, after its discoverer. Eukaryotic genes have one or more RNA Pol-binding motifs that include the TATA and the CAAT motifs. The former is proximal to the CS and the latter distal to it.

About 10 bases from the centre of the Pribnow box (ATAPyTAPy, where Py is a pyrimidine) and between the motif and the start of the CS is the initiation (I) base. The first nucleotide in an RNA is the complement of this base. The region between I and CS is called the *antileader* region and is transcribed into the *leader* region in the RNA. The leader attaches itself to the translating apparatus (small unit of the ribonucleic acid-protein complex called the ribosome) by virtue of a sequence (called the Shine-Dalgarno or SD) that is complementary to a region of the rRNA which is part of the small ribosomal unit.

The region upstream of P (5′-flanking region) usually contains one or more recognition sequences for transcription factors and other regulatory signals.

The region downstream (towards the 3′-end of the molecule) of the CS may carry one or more recognition motifs that are involved in the termination of transcription and in most eukaryotic genes for a step in the post-translational processing of the mRNA. The region downstream of the termination sequences may also possess regulatory sequences which are not clear. A gene without the 3′-flanking region (between the T or termination signal), and the 3′-end of the gene, is not transcribed. Hence, this uncharacterized region is also an integral part of the gene.

The preceding discussion has been centred on what is known as the 'simple gene'. In such a gene the CS region is one block of sequences that does not contain any intervening non-coding bases. Most eukaryotic genes appear to be what have been named as *split* genes (Fig. 1.14). In these genes, the CS consists of blocks of coding sequences (*exons*) alternating with blocks of sequences that do not specify amino

acids (*introns*). These genes are versatile, as different sets of exons may be utilized to fashion specific proteins in different developmental periods and/or tissues. The exons required for a particular polypeptide are joined together and the polypeptide translated from the tailored sequence. This operation is carried out on the RNA and not on the original DNA. The gene is initially transcribed fully into a *heterogeneous* or hnRNA, which is then 'spliced' appropriately (Fig. 1.15). The introns and exons, if any, between two exons that are to be juxtaposed are cut out and degraded. The joining is precise, so that there is no shift in the reading frame in the RNA message. Addition or loss of even one base at the exon-exon junction would result in a garbled message that would specify an abnormal sequence of amino acids.

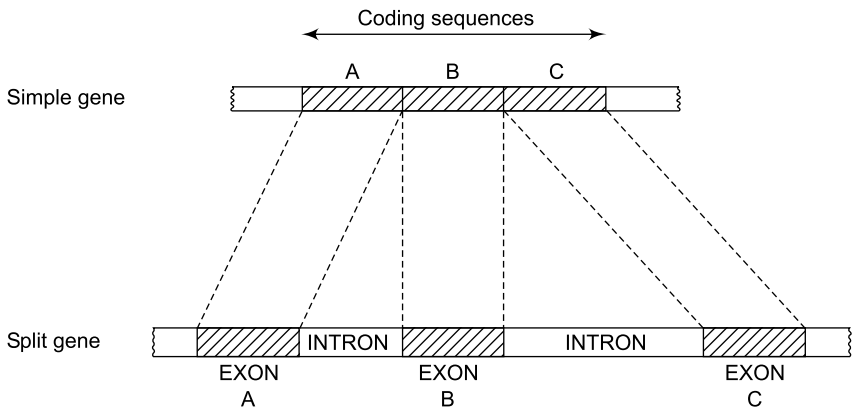


Fig. 1.14 *A split gene versus a simple gene. The former possesses modules of coding sequences (exons) separated by intervening sequences (introns). All sets or a particular set of the exons may be joined together (in the mRNA) to form a functional transcript that is translated.*

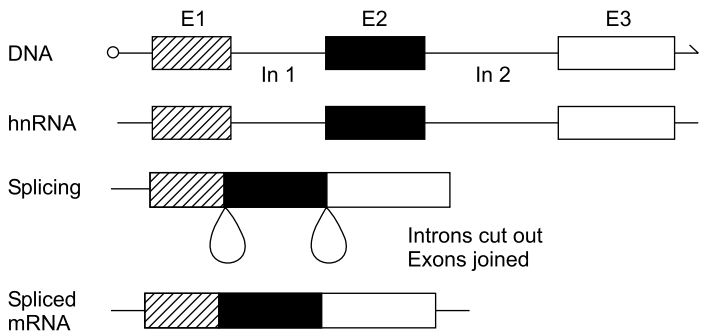


Fig. 1.15 *Splicing of the mRNA of a split gene. Different ways of splicing have been discovered. In all of them, the ends of exons and regions of the introns possess consensus sequences which are involved in the exact splicing of the exons. The introns are cleaved out and degraded.*

The coding sequence of a gene is collinear with the amino acids that they represent. This does not prevent the occupation of a common stretch of DNA by the sequences of more than one gene. Several viruses possess overlapping genes. It appears to be one way of maximizing the genetic material as a source of genetic information. Of course, the overlapping genes are not expressed simultaneously.

Developmental control regulates the sequence and time of transcription of the separate genes.

A further concession had to be made in the molecular definition of a gene, when the anatomy of immunoglobulin genes was exposed. An immunoglobulin (Ig, or antibody, in common parlance) molecule consists of one or more units. Each unit contains two identical subunits. Each subunit is composed of two polypeptides, a longer or heavy (H) chain and a shorter or light (L) chain. The H and L chains of each subunit are held together by disulphide bonds; so are the two H chains of the Ig molecule. Each subunit has one end (the aminoterminal) that binds itself to a matching antigen (a molecular moiety foreign to the animal in question). This end, called the *variable* or V region has several alternate blocks of codons of which only appropriate ones are joined together (in response to the antigen) to form a functional gene. (Here, the unwanted DNA is cleaved off, in contrast to the situation in RNA splicing.) The tailored V region is translated to form the series of amino acids that bind themselves to the antigen.

Both H and L chains have V regions (V_H and V_L). The remainder of the Ig polypeptides is called *constant* (C) regions (C_H and C_L). The V and C regions are specified by distinctly different genes, but the latter are not expressed separately. The V_H and C_H regions are linked through a *joining* (J) region on the same stretch of DNA. After formation of the functional V region, it is linked to a particular alternative subset of the J gene. The V-J joint DNA is then joined to an appropriate C region (which also possesses alternate blocks of coding sequences). The final V-J-C DNA is transcribed and translated. A similar processing occurs in the C chain gene. The Ig genes have been named as *variable* genes. In this case, different genes (each with variable components) are first linked together before producing one final polypeptide.

Two other types of genes (in terms of behaviour and not in terms of the molecular definition) have been identified. These are the mobile or transposable genes and the processed genes. The mobile gene is on a transposable element in the DNA that has characteristic repeated sequences at its two ends (Fig. 1.16). A copy of the transposable element is integrated into another region of the same DNA or a different DNA. Some transposable elements, i.e., the *transposons,* carry sequences of one or more genes. The latter may thus be transferred to a new site in the same genome or a different genome. Some of these elements appear to regulate certain developmental events by stationing a gene next to the one on the recipient DNA.

Now we come to the discovery of extrachromosomal genes and DNA. The essential blueprint for a species resides in the chromosome(s) of its cell. Some species, particularly bacteria, possess, in addition, DNA molecules in the cytoplasm. These are small circular molecules called *plasmids.* Plasmids may carry transposons containing genes for certain special functions that include drug-resistant genes, toxin-producing genes, tumour-inducing genes, nitrogen-fixing genes and others. These genes are not essential for the survival of the individuals carrying them, but do provide certain extra advantages in particular environments (Fig. 1.17).

Plasmids are the earliest and one of the most extensively used components in the hands of a genetic engineer. Pieces of DNA that are not viable on their own are spliced to plasmids, which are independently viable in a cell. The passenger DNA or added DNA is transferred, expressed, and replicated as part of the plasmid DNA (called carrier or vector DNA).

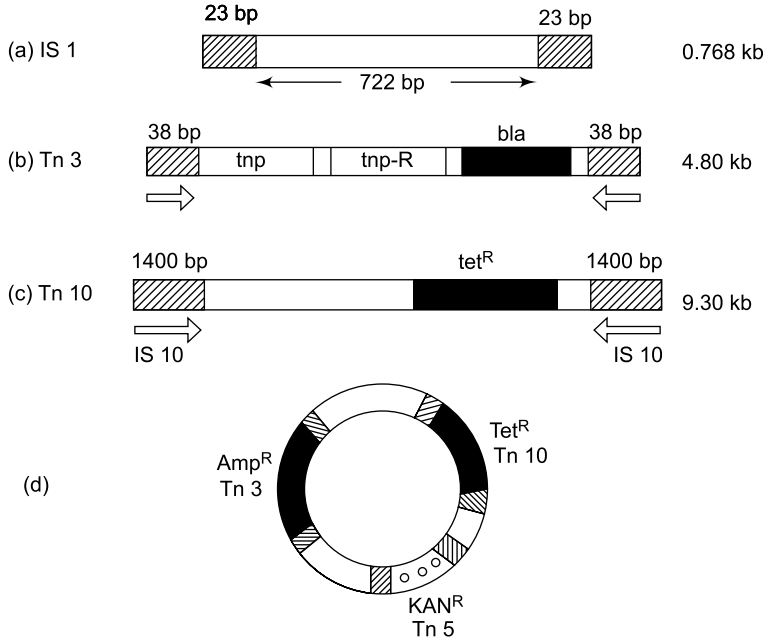


Fig. 1.16 *A transposon is a DNA unit with characteristic repeated sequences (direct, inverted, simple, compound) at the two terminals and a gene sequence. Terminal repeats are involved in the transposition of a copy of a transposon to a different site in the same or a different DNA. Many drug-resistant genes (Amp^R, Tet^R, Kan^R) are on transposons, which, in turn, are within bacterial plasmids or on chromosomes. Transposition may result in different base alterations in the host DNA at the juncture of the transposon and host DNA. Transposition may also join two non-homologous circular DNA into one molecule (cointegration).*

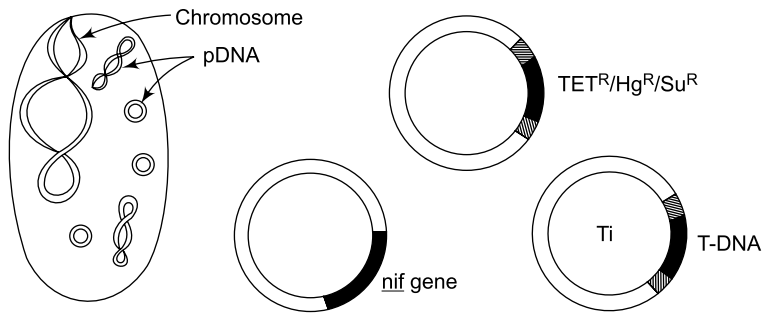


Fig. 1.17 *Plasmids are small covalently closed circular (CCC) ds DNA molecules found mainly in prokaryotic cells. Some eukaryotes have also been discovered to harbour plasmids (2 μm of yeast, Ddp of slime mold, P element of Drosophila, and possibly in maize, sorghum and Vicia faba). Plasmids usually possess one or more genes which provide special advantages to the cell carrying them. They may carry antibiotic, drug or toxic metal-resistant genes, as well as ones with specialized functions, such as the nif-genes of nitrogen-fixing organisms, and tumour-causing genes of the Ti plasmid of Agrobacterium tumefaciens. Plasmids are usually in a supercoiled state.*

Some eukaryotic species have been found to possess plasmids and elements that are also usable as vector DNA. One of them is a 2 μm plasmid in yeast (*Saccharomyces cerevisiae*) and the P element in the fruit fly (*Drosophila melanogaster*).

Other extrachromosomal DNA, found in some plant species, are not characterized clearly, but have been demonstrated to be instrumental in specifying certain genetical traits. These non-chromosomal genes do not follow the classical Mendelian pattern of inheritance. The latter results from ordered distribution, preceded by recombination, of the products of meiosis. Extrachromosomal DNA are not part of any organized equitable transfer system.

1.2.3 Expression of a Gene

As the two strands of a gene are complementary, it is obvious that the coding sequences in one strand will be different from the codons that are complementary to them. This was proved conclusively in the case of two genes in a bacterial virus, the bacteriophage T4. A gene, therefore, is only on one strand of a DNA.

RNA Pol binds itself to the P region of a gene and begins the synthesis of an RNA polymer starting with a base complementary to I (Fig. 1.18). Initiation of transcription involves a complex set of events in which one of the subunits (the sigma) of the RNA Pol is required to bind the enzyme firmly to the DNA. Soon after transcription is initiated, the *sigma* subunit gets detached from the core enzyme. The latter proceeds to elongate the RNA until it reaches the T signal in the 3′-region of the gene. At this point, the enzyme, the RNA and the DNA become separated; the core enzyme reunites with the sigma and begins another round of transcription. This is the basic process of transcribing the message of a gene into an RNA molecule.

RNAs of eukaryotes are processed before they can be translated. This involves adding a methylated 'cap' to the first few bases in the RNA, a polyadenine tail in the trailer region (between the CS and T) and splicing, if the gene is a split one.

The mature or functional mRNA is then translated. The translation apparatus consists of the mRNA, the ribosome (with a smaller and large subunit), and tRNAs, each specific for binding to a particular amino acid.

The initiation of translation is a complex process, at the end of which the 5′-SD region of the mRNA becomes bound to the 16S rRNA in the smaller subunit of the ribosome (in bacteria). The larger subunit joins the initiation complex to form the ribosome with the mRNA running through a groove-like space between the two sub-units. There is space in the groove for only two codons.

The first codon in a CS is always that after methionine (AUG). When the ribosome is fixed with the mRNA, the AUG is in the second position. The amino acid methionine is brought by the methionine-specific tRNA to the ribosome. Further, tRNAs possess three loops each in their secondary structure, the middle one of which has the *anticodon* (complementary bases) for the codon for the amino acid they represent. So, in the preceding example, the anticodon of $tRNA_{met}$ forms hydrogen bonds with the AUG. The ribosome slides backwards by one codon length. The $tRNA_{met}$ is now in the 1st position in the groove. The tRNA with the anticodon for codon 2 now binds itself to the latter; it is bound to its specific amino

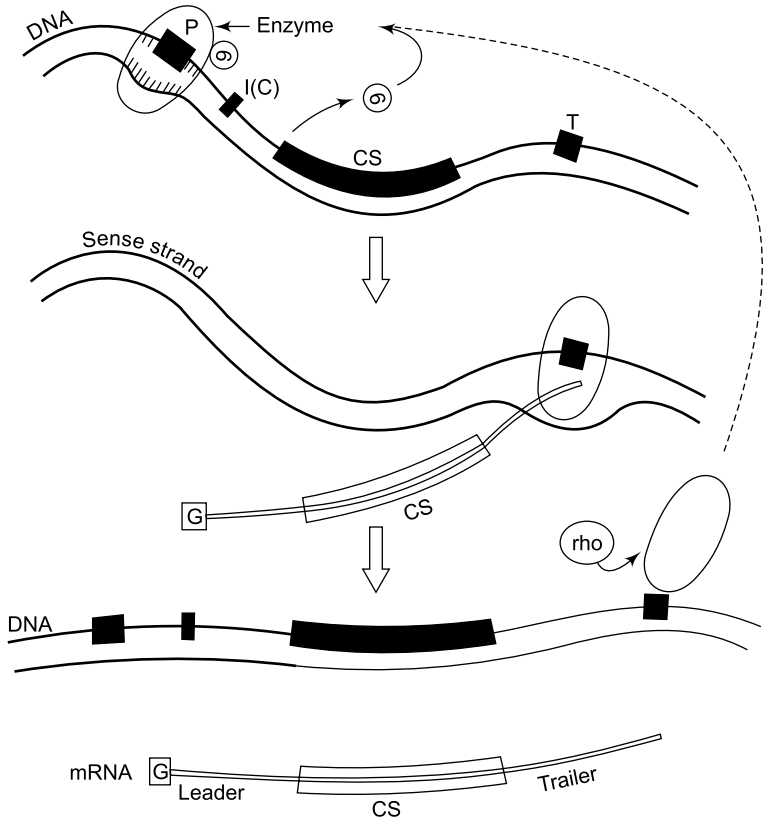


Fig. 1.18 *Transcription or copying of a DNA instruction into an RNA message.*

acid. When amino acids 1 and 2 are side by side, a peptide bond is formed between them. This releases the first tRNA from its amino acid and the ribosome moves one codon backwards. The third codon is bound to the anticodon of the tRNA for the third amino acid. The second and third amino acids are joined. In this way, the entire coding sequence is decoded into the amino acid chain, that begins with the $-NH_2$ end of the molecule.

The last codon in the CS has no representative tRNA. Hence the polypeptide chain stops elongating when it reaches this last or *stop* codon. At this point, the ribosomal subunits, the mRNA, the tRNAs and the new polypeptide are released from bondage. The apparatus is free to initiate a fresh polypeptide along the same mRNA.

The foregoing description presents in the briefest of outlines, the picture of how a gene is expressed.

1.2.4 Regulation of Gene Expression

Jacob and Monod, while studying the induction and repression of certain enzymes in the bacterium *E. coli,* discovered the molecular mechanism by which a gene is

switched 'ON' or 'OFF'. In their particular case, they were examining the induction of genes, the products of which are needed to metabolize the sugar lactose. Normally, *E. coli* thrives on glucose. When glucose is not available, the bacterium is able to use other sugars when induced with a minute quantity of the sugar in question. When glucose is available, the genes for utilization of the other sugars are in an 'OFF' position. In the absence of glucose and presence of the other sugar, the genes for the latter come into operation.

Jacob and Monod postulated that the protein decoded from a particular gene (inhibitor or regulator gene) competed with the RNA Pol for a place to bind itself on the DNA. It was discovered that the regulator protein can bind itself to a sequence that partially overlaps the promoter sequence. When the regulator is bound to this region (called the operator) the RNA Pol cannot achieve transcription. The gene(s) down-stream of the P are thus shut 'OFF'. In the case of lactose utilization there are three genes which are turned 'ON' or 'OFF' by the same regulator, in one operation.

Jacob and Monod presented the Operon Theory on the basis of the detailed study of the lactose utilization genes of *E. coli.*

The three genes in this case are parts of the lactose or *lac* operon (Fig. 1.19). It is operated by the protein from the *lacI*gene, which is not part of the operon. Binding of the *I* protein to the operator prevents transcription of the lac operon genes *lacZ, lacY* and *lacA.* Of these, *lacZ* specifies the enzyme galactosidase—that mediates the separation of the glucose and galactose components in the lactose disaccharide molecule.

Extensive investigations have since revealed other factors that regulate the expression of a gene. The basic molecular principle is, however, still the same. A regulator protein either blocks the binding of the RNA Pol by itself or by complexing with another (effector) molecule. Conversely, complexing of the regulator with an effector molecule may also make it unsuitable for binding to the operator region. In such a case, the RNA Pol is free to initiate transcription.

In higher organisms the signal for a particular gene to be expressed may be in the form of chemical messenger, or hormone, antigen, mitogen (mitoses-generating substance) or others. Hormones may be polypeptides or steroid based. The former bind themselves to the cell surface; the signal is transduced through the cell membrane to a second messenger which sets in motion a series of biochemical steps that eventually trigger a particular gene to be expressed. A steroid hormone moves into the cell and forms a complex with some receptor molecule, and it appears that it is this complex that makes a beeline for the DNA of the gene to be expressed.

1.3 ❑ EXPLOITATION OF MOLECULAR GENETICS: GENE CLONING

The most spectacular spin-off of molecular genetics is the birth of recombinant DNA technology. It is also popularly referred to as gene cloning. It is possible to obtain large quantities of a gene, or any other part of a DNA, using recombinant DNA techniques.

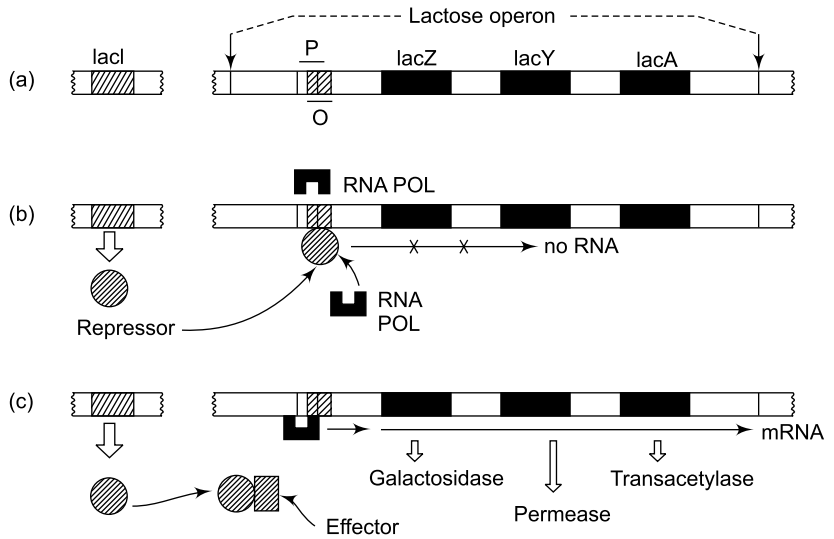


Fig. 1.19 *The lac (lactose) operon of E. coli. Three genes lacZ, lacY and lacA are under the control of cis-acting DNA elements lacP (promoter) and lacO (operator). The RNA polymerase binds itself to P, and the product of a gene lacI (repressor) binds itself to O. The lac operon may be induced to express the three genes by supplying a little lactose to the system. In practice an analogue IPTG is used as this, unlike lactose, is not degraded by β-galactosidase. lacZ encodes α- and β-subunits of the polypeptide β-galactosidase—an enzyme that mediates the conversion of the disaccharide (lactose) into the monosaccharides glucose and galactose.*

The purified clones of DNA may be utilized to further investigate the working of a biological system, or for applications that tailor products to specifications in agriculture, medicine and other fields.

Gene cloning methods have made it possible to chart the entire sequence of genetic information in any DNA. Interpretation of much of the exposed information is already available, while known phenomena are becoming understood on the basis of sequence patterns in the DNA.

Improvements in the culture of organisms, both whole or as single cells, together with sophistications in DNA engineering are rapidly enhancing the area known as genetic engineering.

REVIEW QUESTIONS

1. Describe the monohybrid and poly-hybrid crosses performed by Mendel.
2. Explain how Muller's "mutation theory" helped in amendment of Darwin's "evolutionary theory".
3. How do RNAs participate in the information flow from gene to the final product?
4. Recall the anatomy of a gene.
5. How did Jacob and Monod study regulation of gene expression?

2 Genetic Engineering Toolkit-I: Enzymes, Vectors and Hosts

Introduction

Genetic engineering is the directed manipulation of the hereditary material. It is based on a set of molecular techniques collectively called recombinant DNA (rDNA) technology. In essence, rDNA technology consists of the splicing of a piece of DNA to a suitable carrier which is then introduced in a convenient cell that allows the 'passenger' or 'foreign' DNA to express itself in the new surroundings.

The introduced foreign DNA is amplified into several thousand copies together with the carrier DNA. This has been, until recently, the only way of acquiring a measurable quantity of a DNA in the purest of forms. The latest strategy of amplifying a length of DNA is the one that uses the Polymerase Chain Reaction or the PCR technique. This technique, unlike conventional gene cloning methods, requires no cloning in host cells.

2.1 ❑ RECOMBINANT DNA TECHNOLOGY

The birth of rDNA technology is based on two sets of discoveries: the discovery of the Watson and Crick DNA model together with the studies that ensued, and the identification of enzymes that can cleave, elongate, join or otherwise modify DNA molecules. These two discoveries complemented each other, and opened up avenues of biological research in an unprecedented manner. It became possible to snip out a piece of DNA from any organism and prepare it for amplification in a conveniently handled cell.

A piece of DNA is unable to maintain or replicate itself on its own unless it possesses certain essential features. To clone such a DNA fragment it has to be first joined to another DNA that possesses these essential features. The first job in gene cloning is to find a suitable carrier or vector DNA for the passenger DNA. Now, the latter cannot be picked out effortlessly from the long straggly DNA molecule. Instead, the DNA containing the fragment of choice is first cleaved into several pieces. This collection of fragments is then mixed with the vector DNA which also has two free ends. Both types of DNA fragments join with each other forming several kinds of hybrid molecules each containing a particular fragment from the donor DNA. The linear hybrid molecules are then circularized and the cut ends sealed.

The next step is to look for the hybrids that contain the passenger DNA. To do this, the total population of hybrid and non-hybrid DNA pieces are first introduced into suitable cells that grow and replicate without undue fuss. The cell of choice to date has been that of the bacterium *E. coli.* A few out of thousands of DNA pieces manage to enter these cells. The latter are said to be 'transformed' (due to the addition of extrinsic genetic matter). Each transformed cell grows a colony of its own, in which every member is genetically alike. Out of the many colonies that surface, only some contain the DNA fragment of interest. These colonies have to be distinguished and recultured separately. This 'fishing expedition' may be undertaken in a number of ways, all of which depend on at least one tell-tale trait that is present in the chosen vector. After selecting the clones carrying the desired hybrid molecules they are cultured extensively to provide sufficient numbers of cells from which a worthwhile quantity of the hybrid DNA can be extracted. The recombinant DNA is then extracted from lysed cells, purified and used as desired.

The first gene was cloned in 1973 by Herbert Boyer and Stanley Cohen of Stanford University, California. The year before, the first hybrid DNA was produced by Janet Merty and Ron David also at Stanford. Both these successes rested on the discoveries of two unique enzymes: *ligase* in 1967 by Merty and David and *restriction enzymes* in 1970 by Hamilton O. Smith. Ligase joins DNA backbones; restriction enzymes cleave the DNA at regions that are specific for each enzyme. The specificity lies in a set of 4–7 base-pairs that are recognized by a particular enzyme which then cleaves the DNA backbone within or very near this recognition sequence.

The restriction enzymes that were first used profitably were those that could make staggered cuts on opposite strands of the DNA duplex, creating in the process a single-stranded (ss) 'tail' at each cut end. As the tails are complementary to each other, they can come together and become hydrogen-bonded once more. This 'stickiness' of DNA ends, cut by the same enzyme, is the rationale behind bringing together a donor and a vector DNA prior to sealing the backbones with ligase.

Restriction enzymes and ligase are the essential enzymes for creating hybrid DNAs. However, the utility of this technique has been very much enhanced by the discovery of other enzymes that catalyze other reactions vis-à-vis nucleic acids. These include the DNA and RNA polymerases, exonucleases and RNA and DNA nucleases that act on single- or double-stranded molecules. One of the enzymes that has proved to be invaluable in rDNA technology is *reverse transcriptase,* which can mediate the synthesis of a DNA strand along an RNA template. It is used as a tool for making copies from cell RNA molecules of genes that are not easily located or reached.

Three experimental techniques that have contributed to advances in DNA biology are *gel electrophoresis, Southern blotting* and DNA *sequencing.* A fourth one that has become possible due to these techniques and appears to be a promising tool is *site-directed mutagenesis.* The latest invaluable technique for DNA engineering is the one known as the *Polymerase Chain Reaction* or the *PCR* technique.

DNA fragments and proteins of various sizes can be fractionated elegantly by gel electrophoresis. The ability to separate fragments of DNA differing by only one nucleotide is utilized profitably in methods for DNA sequencing, which lays bare the nucleotide pattern in a DNA molecule. DNA sequencing involves generating sets

of fragments in the molecule to be read, dispersing these fragments in appropriate gels and reading the order of the nucleotides from the order of size of fragments. The utility of DNA sequencing techniques cannot be adequately appreciated. After all, one of the chief aims of molecular genetics is to read the blueprint inherent in the genetic material, and sequencing gives us this information. Of course, we have a long way to go before we can read all the sequences meaningfully. But at least, the text and verse are available; the interpretations will come in time. Southern blotting was designed by E M Southern to transfer the bands of DNA from electrophoretic gels to specialized papers. The transferred DNA bands can then be searched for wanted regions by appropriate techniques. Identification of homologous regions in DNA and RNA are invaluable tools in rDNA procedures.

A genetical technique that developed as a spin-off of DNA studies is that of site-directed mutagenesis. Traditional methods of mutagenesis are incapable of causing alterations at a particular predetermined region of the DNA. Site-directed mutagenesis does just that and provides a means to study functions of different regions of both genic and non-genic DNA.

The PCR technique can amplify any short length of DNA several million folds in a test-tube, by using only the DNA to be amplified, two primers, each matching the 5′ extremity of the single strands of the said DNA and a DNA polymerase that works at a high temperature. The mixture is heated to dissociate the strands. After a while the mixture is cooled. The primers attach to the ss DNA; the DNA Pols begin synthesizing a second strand along each separated template. When the strand has become duplex, the mixture is heated once more. Alternate heating and cooling of the reaction mixture results in an exponential increase in the number of new DNA molecules.

An experimental technique which has been devised and perfected in recent times, and owes much to recombinant DNA technology, is the synthesis of *monoclonal antibodies* in *hybridomas.* Hybridomas are tumorous growths of hybrid cells, one partner of which has the ability to proliferate, and the other to synthesize only one type of antibody. Such antibodies that are specific for their antigens are extraordinary tools for identifying minute quantities of proteins and nucleic acids in a heterogeneous population of molecules. They are used for purification of proteins and other antigens, and possess the potentiality of use as carriers of anticancer drugs directly to the target cells. Today, many pharmaceuticals provide monoclonal antibodies to antigens used routinely in investigations, as well as supply monoclonals, on order, against specific antigens.

Molecular biology, together with its potent tools of rDNA techniques, *in vitro* DNA synthesis and DNA sequencing is destined to be one of the most exciting, challenging and profitable fields of study in the next few decades.

With this brief introduction to the general principle of recombinant DNA technology and some of the developments that accompanied its use, let us proceed to discover how genes are spliced, cloned, selected, transferred or otherwise manipulated for a specific goal.

2.2 ❑ HOW TO CLONE A GENE

The operations involved in gene cloning consist of the following steps (Fig. 2.1):

1. Cut the larger DNA (donor) and the vector DNA with the same special enzyme. Join a donor and a vector DNA. A recombinant or rDNA has been made. One of these rDNAs has the fragment with a desired gene G.
2. Amplify (clone) each rDNA in *E. coli* cells.
3. Select clones with rDNA. Some of them carry G.
4. Select clones with G.
5. Increase the number of selected clones. All cells have rDNA with G.
6. Extract rDNA from cells, every rDNA has G.
7. Use the rDNA for research or application.

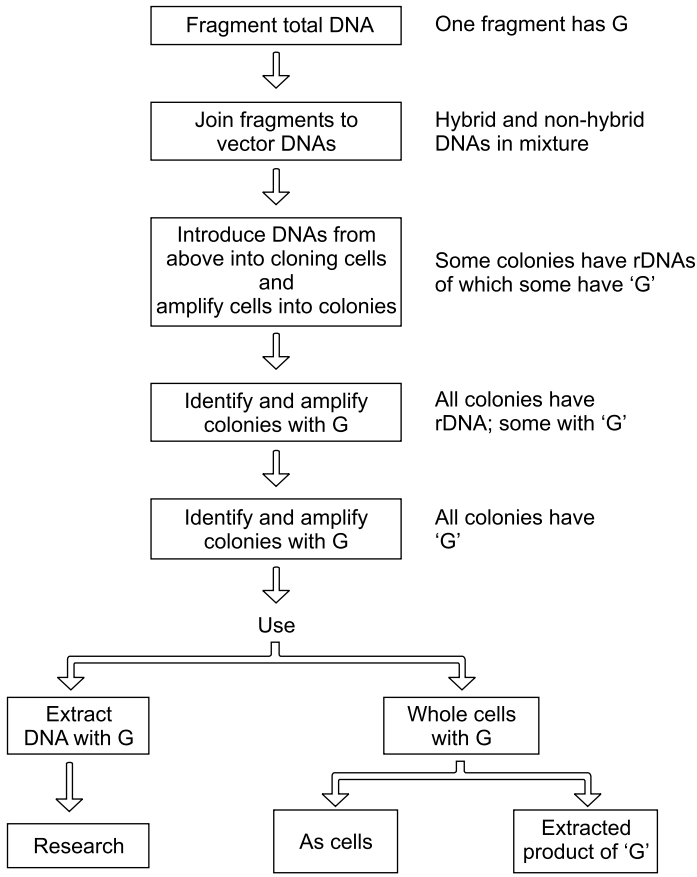


Fig. 2.1 *Cloning of a gene.*

Let us now follow the steps within each of the preceding operations (Fig. 2.2):

1. To construct the rDNAs

1.1 Select a vector (carrier) DNA and the DNA to be cloned (donor DNA).
1.2 Cut each DNA with the same restriction enzyme. The cut ends of donor and vector are complementary.
1.3 Modify the cut ends for better joining.
1.4 Mix donor DNA and vector DNA with ligase (joining enzyme).

Result Many rDNAs (hybrid, chimaeric or recombinant) are made, each having the vector DNA and one of the fragments of the donor DNA.

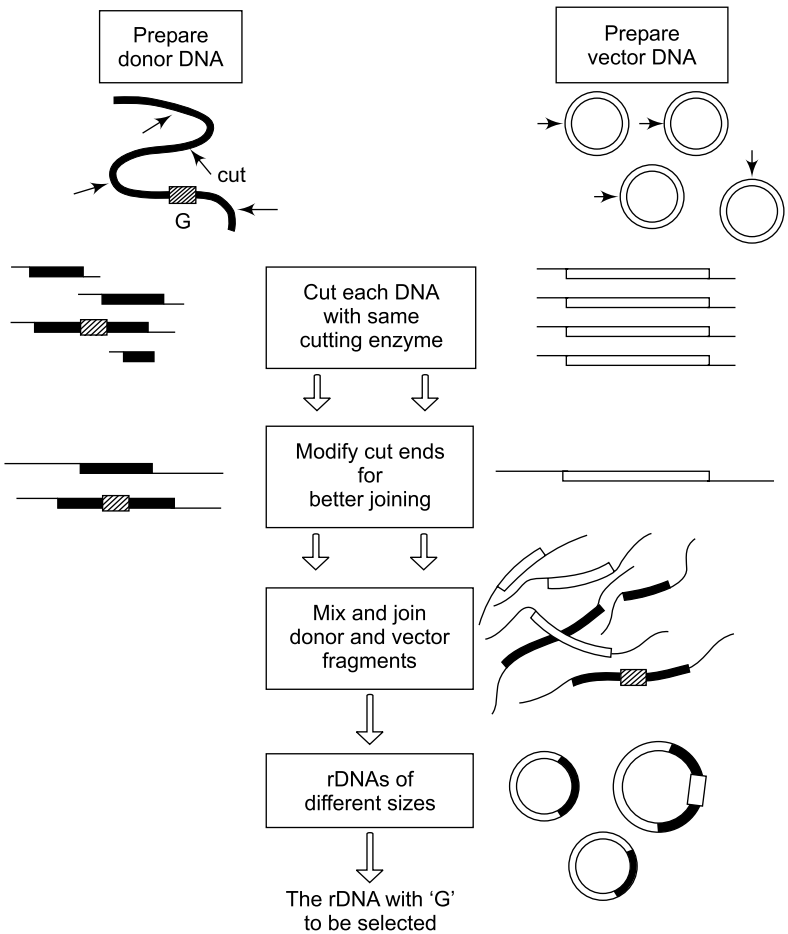


Fig. 2.2 *Construction of a recombinant DNA molecule.*

2. To amplify the rDNAs

Introduce the rDNAs into *E. coli* cells, and allow the latter to divide into large populations.

Result There will be several cells carrying the same donor fragment, and there will be as many types of transformed cells as there were donor fragments. There will also be the untransformed cells containing no vector or rDNA fragment.

3. To select the cells with an rDNA each (Fig. 2.3)

3.1 Pour the cells from Step 2.1 into culture plates containing solidified nutrient media. Each cell will divide again and again, forming a colony in which every cell is genetically identical; that is, each cell initiates the formation of a clone of cells. There will be several clones containing rDNA and several without any.

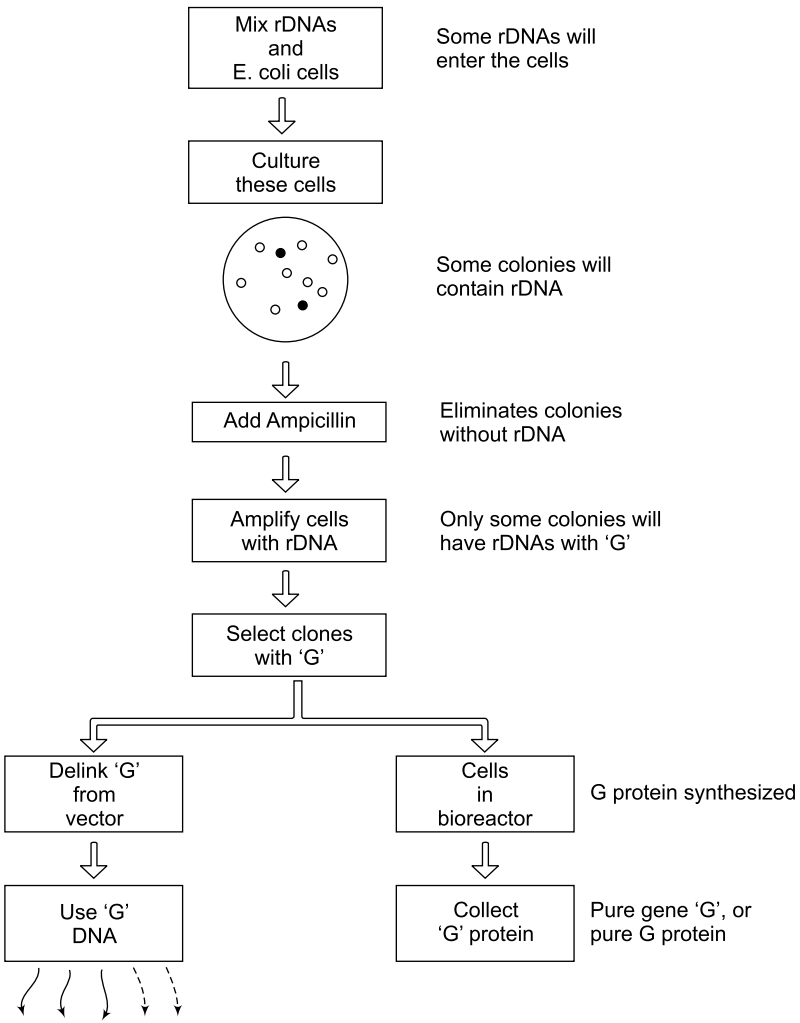


Fig. 2.3 *Selection of clones with a desired rDNA.*

3.2 Transfer each colony to a culture plate to which a selection agent is added. The nature of the selection agent depends on a gene on the vector DNA that directs the activity (or lack of it) against the agent. For instance, say, the vector carries the gene for the enzyme β-lactamase. This enzyme thwarts the activity of the antibiotic, ampicillin. The gene on the vector is said to be for ampicillin resistance (Amp^R). In the preceding general scheme, let us say ampicillin is the selection agent. Those cells containing the vector (with or without the rDNA) will be ampicillin resistant; this means that they will survive when treated with ampicillin. All untransformed cells are ampicillin sensitive, and so will not survive in the selection medium. Of the rDNA-containing colonies, only one set will contain the rDNA having G.

4. To Select Clones with G

The method for selecting the rDNA clones with G depends on various factors. Several strategies have been developed to suit different genes, vectors and cloning cells. One of them may be employed to identify the clone of interest. In screening techniques, the required clones are recognized by some easily visible feature. In the selection technique, the desired clones are the only ones that survive under certain given conditions; all other clones are eliminated.

5. To Amplify Clones with G

The screened or selected clones are transferred to fresh culture plates and grown into large populations. The rDNAs in these cells will be of an appreciable quantity.

6. To Extract the rDNA

The rDNA is extracted and purified. The cells are lysed and all material, except the low molecular weight DNAs, (these are the small rDNAs) is removed from the lysate. Several methods are available for isolating such DNAs. If desired, the rDNA can be cleaved with the enzyme used to cut the donor and vector DNA in Step 1.1. The vector DNA and the donor fragment with G can be separated by the technique of gel electrophoresis. The two DNAs, of different sizes, migrate to different distances in a gel material (porous) soaked in a buffer solution, when placed in an electric field.

7. To Use the Isolated DNA

The fragment with G may now be utilized for research or application. The opportunities for doing so are myriad in both these categories of use. The rDNA may be left intact and cloned in another set of *E. coli* cells for the purpose of collecting the protein that is synthesized on directions from the gene G. Or else, it may be utilized for gene replacement therapy in animals, plants and human beings. Agriculture, medicine and all industries based on living organisms or their products can exploit the isolated gene in specific ways. Or else, the detached fragment with G may be respliced to another vector for the purpose of studying the sequence of the gene, its mechanism of expression and/or the regulatory system that controls its expression. The uses to which the cloned gene is put will be limited mostly

by human imagination and somewhat by the not-so-clear understanding of the workings of the living system. Resources are not likely to be serious bottlenecks, as they are in most other spheres of human endeavours.

2.3 ❑ THE ENZYMES TO BE USED

Genetic engineering became possible with the discovery of two types of enzymes: the cutting enzymes called *restriction endonucleases* and the joining enzymes called *ligases.*

Restriction endonucleases, or restriction enzymes as they are called popularly, recognize unique base sequence motifs in a DNA strand and cleave the backbone of the molecule at a place within or, at some distance from the recognition site. The DNA is thus cleaved into defined and discrete fragments by a particular restriction enzyme.

Ligase is the enzyme that joins a 5′-end of a DNA with a 3′-end of the same or of another strand.

Once the means of cutting DNA into non-random fragments and of joining them became known, the stage was set for diverse DNA cloning exercises. Further embellishments have been added and are continuing to be appended to make DNA engineering a highly sophisticated technology.

Let us look at the above mentioned main mediators and the chief accessory requirements of reactions that are widely used in this technology.

2.3.1 Restriction Endonucleases

Ordinary nucleases are exonucleases or endonucleases. The former remove one nucleotide at a time, starting with the 5′- or 3′-end of a DNA strand. The latter cleave the DNA backbone between two nucleotides—one type at the bond between the 3′-end of a nucleotide and the phosphate, and the other between the 5′-end and the phosphate. In either case, the polynucleotide chain is reduced to free nucleotides. A restriction *endonuclease* cleaves only at specific regions in a particular DNA, so that discrete and defined fragments, instead of free nucleotides, are obtained at the end of total digestion. These enzymes recognize certain short motifs of base sequences and cut the DNA within it or at some specified or unspecified distance from it.

The name ‘restriction’ endonuclease originated from an observation of a system of restriction of the growth of the phage lambda in particular strains of the *E. coli* host cell. G Bertain and J J Weigle (1953) had observed that when the phage lambda is grown in strain C of *E. coli,* it does not fare well in strain K12. On the other hand, those grown in *E. coli* K12 remain unaffected when grown in a fresh batch of K12 cells. This was the first indication that *E. coli* or bacteria possess some system that preferentially restricts growth of phages as well as modifies the mechanism of restriction.

The system itself was identified by M Meselson and R Yuan in 1968. This time the DNA, and not the cells, from strains C and K12 of *E. coli* were taken and treated with an extract from K12 cells. The DNA from the C strains was fragmented, but

not that from the K12. The enzyme in the extract that cleaved the DNA was named 'restriction endonuclease'. This discovery was soon followed by another one: the protection from cleavage is due to methylated bases in the motifs or sites of recognition for the enzyme. All newly synthesized strands of DNA become methylated almost simultaneously with strand elongation, with the help of methylases. Phage DNA replicating in a host become methylated by the same enzymes. A phage DNA thus modified in one host strain is immune to the restriction enzyme of the same strain, but not to enzymes of others against which phage DNA have not been protected.

More than 600 different restriction enzymes have been identified, although all of them have not been characterized as yet. These enzymes have been classified into Types I, II and III. Type I enzymes have fairly long recognition motifs, but cleave the DNA perhaps at random points at least more than 1 kb away from the motif. This was the type of enzyme identified by Meselson and Yuan.

Type II enzymes were first discovered by Hamilton O Smith and associates, in 1968, in *Hemophilus influenzae.* Members of this class recognize motifs that are 4–7 bases long, and mostly cut at a specific place within this motif; a few are known to cleave a few bases to the 3′-end of the motif. In either case, the site of cleavage by Type II restriction enzymes is specific. These were the enzymes that launched recombinant DNA techniques.

Type III enzymes also cleave in a site-specific manner, but their other requirements for the cutting reaction are different from those of Type I and II. Type I enzymes require ATP, Mg^{++} and s-adenosyl methionine besides the catalyst. Type II enzymes need only Mg^{++}. Type III enzymes require the same additional factors as Type I enzymes. The enzymes are also different in the three cases. Type I is a small monomeric protein. The Type II enzyme is composed of two subunits, whereas the Type III is quite large with three subunits.

Restriction enzymes are obtained from bacterial species. Hence the name of the enzyme is an abbreviation of the first letter of the genus and the first two letters of the species name of the source bacterium; e.g., *Eco* RI, is from *Escherichia coli.* The last Roman numeral is the order in which the enzyme of this type was discovered. An alphabet between the first three letters and the numeral is the strain designation, where specified.

Since Type II enzymes cut the DNA at a very specific bond within the recognition motifs, they are the only ones that are useful for use in recombinant DNA exercises. Type III enzymes also cleave in a site-specific manner, but the reaction mixture needs to be supplemented with S-methionine and ATP. The discussion here will be confined mainly to the features of the Type II enzymes.

The motifs for the Type II restriction enzymes are 4, 5, 6 or 7 bases long (Fig. 2.4). There is usually a rotational symmetry within this motif; that is, the sequence in each strand of the motif is palindromic. The motif may be cleaved asymmetrically, so that each cut end has a single-stranded extension, or may be cleaved at the same place on both strands. In this case, the fragments are blunt ended. A few examples will illustrate the preceding text.

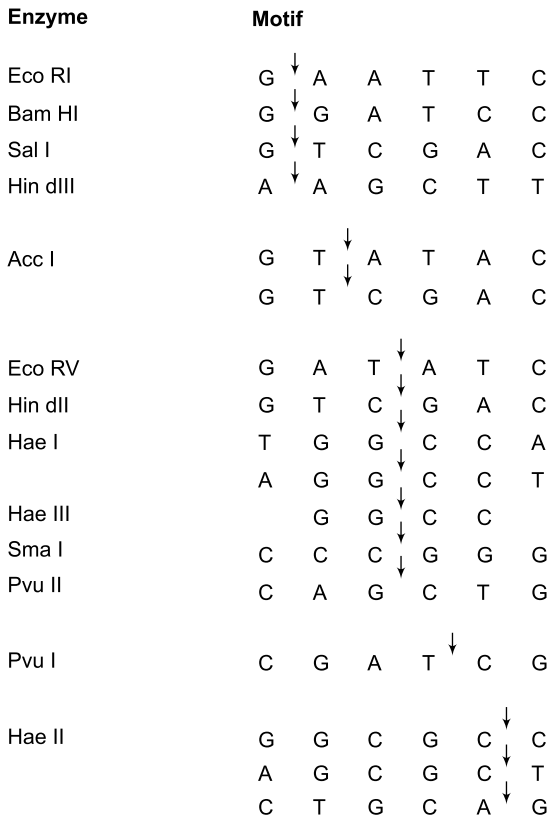


Fig. 2.4 *Some restriction enzymes and the DNA motifs that they recognize. The arrows indicate the position at which, the endonuclease cleaves the DNA backbone.*

1. The enzyme *Eco*RI and HindIII recognize motifs 5′GAATTC3′ and 5′AAGCTT3′, respectively, which are cut on the left side of the central axis, generating free 5′-ended tails (Fig. 2.5).
2. The enzyme *Pst,* on the other hand, cleaves its recognition motif to the left of the axis, generating 3′-OH ended tails.
3. The enzymes *Hin* dII and *Hae* II produce blunt ended fragments by cleaving at the axis on both strands.

You may notice in the first two of the preceding examples that the 'tails' of the two generated fragments are complementary to each other; that is, if allowed to come near each other (in a solution) the fragment ends will align themselves by hydrogen bonding between the complementary bases. This apparently unspectacular event forms the foundation of genetic engineering. Of course even if the 'tails', which are referred to as 'cohesive' or 'sticky' ends, do align themselves to give the illusion of one continuous DNA instead of two fragments, the molecule is not truly one; there are gaps at the sites of cleavage, in each strand. The gaps could be closed only when the joining enzyme ligase was discovered.

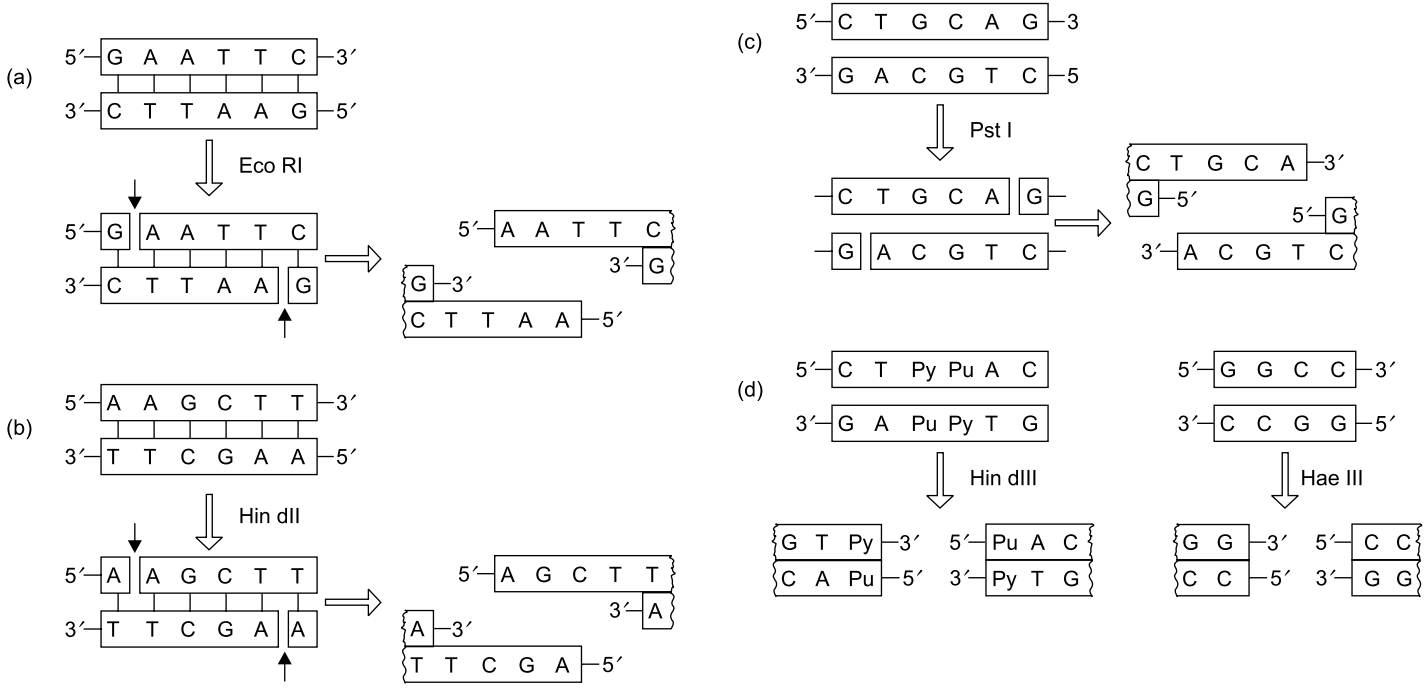


Fig. 2.5 *Some restriction enzymes generate single-stranded extensions at each end of the restriction fragment (e.g., Eco RI, HindIII and Pst I), while others create flush or blunt ended fragments (e.g., HindIII and Hae III).*

By choosing the appropriate restriction enzymes, any two DNA molecules (or the opposite ends of the same duplex) can be made to align and become covalently joined with the addition of the enzyme ligase.

A few Type II restriction enzymes have been identified that cleave the DNA strands at a specific distance downstream from the 3′-end of the recognition motif. One of them is *Hga,* which recognizes the penta-nucleotide 5′ G A C G C 3′, but cleaves 5 bases to the right in one strand (from the 3′-end) and 10 bases to the right in the other (also counted from the 3′-end) as shown in Fig. 2.6. Also we have

5′ G A C G C N N N N N N 3′

3′ G T G C T N N N N N N N N N N 5′

where N represents any nucleotide.

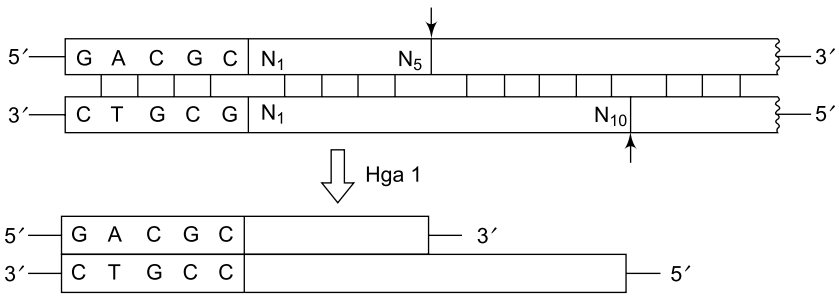


Fig. 2.6 *Cleavage pattern of Hga I. One strand is (cleaved) 5 nucleotides downstream from the 3′-end of the recognition motif, while the other is cleaved 10 nucleotides away.*

The sites of recognition of different restriction enzymes are inherent features of the DNA of a species. Hence, the exact pattern of distribution of the motifs in the total genome is theoretically identical in every member of a species. This allows one to cut different aliquotes of the same genome DNA with the same enzymes and recover identical arrays of fragments from each lot. Occasional deviations are observed in this restriction site map (hence the use of the term 'theoretically' a little earlier). These deviations are due to mutations and have special significance for certain types of investigations and, so, are exploited for the particular goals.

The presence of recognition motifs for restriction enzymes in spatially defined locations in the genome has also been exploited to prepare what are referred to as 'restriction maps'. By cutting a length of DNA with one or more restriction enzymes, one can generate a 'physical' map of the DNA in terms of specific fragment lengths and the consecutive position of each fragment. This fact has been of tremendous practical value; not only can the DNA be mapped in terms of physical distance but also specific restriction fragments may be identified with loci of particular genes. Restriction fragments are used also for charting the complete sequence of a DNA by what is known as 'chromosome walking'.

Restriction enzymes have to be extremely pure (not contaminated with any substance from the cell extract or reagents) to be able to mediate fragmenting of DNA. It is customary, therefore, to buy them from suppliers who guarantee their purity. In any event, before undertaking a 'cutting' job, the activity of the enzyme

and the proportions of ingredients in the reaction mixture, which are optimum for the purpose, have to be assessed in a trial run.

Some kinetics of the restriction enzyme action is known, but not enough to be able to control such action with confidence. The manner in which the enzyme protein binds itself to the DNA motif has been recently indicated from X-ray diffraction and other studies of *Eco* RI. For the practical genetic engineer, however, lack of sufficient knowledge of the physics of the system has not been a bottleneck so far.

The large variety of restriction enzymes, some of which alter their specificities for recognition at different temperatures, provides a very versatile tool for adding, removing, sequencing or otherwise engineering the DNA molecule.

2.3.2 Ligase

Ends of DNA strands may be joined by the enzyme polynucleotide ligase. The enzyme catalyzes the formation of a phosphodiester bond between the 3′-hydroxyl and the 5′-phosphate terminals of two nucleotides. The enzyme is thus able to join unrelated DNA, repair nicks in single strands of a DNA and join the sugar-phosphate backbones of the newly repaired and resident region of a DNA strand.

The reaction of ligation requires a source of energy. The ligase extracted from phage T4 infected cells (phage T4 *lig. gene* or *gene* 30 encoded) requires ATP, and the *E. coli* ligase utilizes NAD for the purpose. The latter joins DNA fragments with sticky ends, and the T4 ligase manages to join flush-ended fragments. The T4 enzyme is obtained commercially from the T4 *lig. gene* cloned in a phage lambda vector with which the *E. coli* K12 strain has been lysogenized.

The ligation reaction can be performed in a test-tube or inside a cell. It must be remembered that if the recombinant DNA is to be ligated *in vivo,* the ends of the DNA must possess cohesive ends. In practice *in vitro* ligation appeals to the genetic engineer, as fewer mishaps (loss of nucleotides in the ss tail due to cellular nucleases and so on) can occur outside a cell than within it. In addition, the type of ligated molecule to be produced (circularized or tandemly joined ones such as concatamers) can be controlled by *in vitro* ligation.

The ligation reaction is controlled by several factors, such as pH, temperature, concentration and kinds of sticky ends and so on. As the cohesive ends are held together only by hydrogen bonds, the temperature at which ligation is performed becomes very important. At temperatures above 15°C the single-stranded tails are liable to dissociate.

When a recombinant of, say, *A* and *B* DNA only (one each) is the goal, the fragments are made incompetent to curl up and join their own ends or form tandemly joined fragments of the same kind by removing the phosphate moiety from the ends of the *A* or *B* fragments. The *A* and *B* fragments will then get aligned at the cohesive ends and the ligase will be able to connect the 3′-end of a strand with the 5′-end of the other.

The optimum temperature for ligase action is 37°C. But in the case of joining of cohesive ended fragments, as the ends dissociate at such a high temperature, the reaction temperature is lower. This does not apply to ligating blunt ends.

In this case, the problem is to bring the blunt ends face to face long enough for the reaction to take place. This is usually achieved by having a very large number of fragments, the ends of which are to be joined. In sticky end joining, therefore, temperatures lower than 15°C (preferably 12.5°C) are practical, while for blunt end ligation, one may go to a temperature as high as around 23°C.

As ligase uses the ends of DNA molecules as substrates, rather than the entire DNA, the kinetics of joining depend on the number of ends (concentration) available for joining. DNA molecules longer than 200 bp are liable to join end to end to form circular molecules. DNA molecules longer than 20,000 bp, on the other hand, do not circularize readily.

Joining of two ds DNAs takes place in two steps: two strands are linked by a phosphodiester bond, followed by linking of the opposite strands. It takes longer for the first step to occur, as only random collisions between the free 5′ and 3′ ends bring the moieties in the proper conjunction for the enzyme ligase to operate. Once one strand is ligated, the complementary one is joined rapidly, as they are already held immobilized, by hydrogen pairing, with the other DNA strand.

A high concentration of joinable ends is required for ligation of blunt-end DNAs, as in this case too, only random encounters can clinch the issue. The initial joining of cohesive-ended molecules may be speeded up by extending the cohesive ends with homopolymer tailing. The ends to be joined are thus stabilized, or held in position by the base-paired extensions.

Many DNA engineering exercises do not require *in vitro* ligation. The hybrid, but unligated DNA, becomes ligated *in vivo,* by endogeneous ligase, in the cloning cell. Isolation procedures may introduce nicks in the DNA. It may be necessary to close these nicks with ligase before proceeding further.

The commercially available ligase may not always be upto the mark in terms of potency, due to any of a variety of environmental or handling parameters. It is wise, therefore, to test the ligating efficiency of the enzyme before utilizing it.

The optimum reaction mixture for cutting and ligating DNA has been developed empirically. It contains a reaction buffer, DTT, ATP or NAD, a restriction enzyme and ligase. One may test the efficiency of reaction mixtures in which the ratios between the above mentioned components are varied. The ligating ability is assessed by running control mixtures and ligation mixtures (after conclusion of the reaction) in agarose gels. The ratio of the ligated (single large DNA) and unligated (two or more smaller fragments of DNA) DNAs may be determined from the corresponding bands in the gel. The reaction mixture (without the ligase) containing the DNA to be fragmented and subsequently joined, is vortexed, spun briefly in an Eppendorf (microfuge) centrifuge and kept at 37°C for one hour. Reaction is stopped by holding the tube in a 65°C water-bath for about 10 minutes. This mixture is then supplemented with ligase, incubated at 14°C overnight, diluted with the reaction buffer, heated to 60°C for 10 minutes (to inactivate the ligase and to dissociate DNA ends that have base-paired). Small aliquots (50 μl) of this sample are then run on a 0.8% agarose gel.

2.3.3 Enzymes to Modify Ends of DNA Molecules

The central event in a gene-splicing project is the joining of ends of DNA molecules. The latter may be restriction fragments with blunt or sticky ends. In the latter, the protruding tail may have a 3′- or a 5′-end. The two ends to be joined may not be compatible, or may possess complementary single-stranded regions that are too short to hold the molecules together long enough for them to be ligated. It may be required to make an unevenly cut-end into a flush-ended molecule, or the latter made to acquire sticky tails. For these and other fine-tuning jobs, several enzymes are available which are widely utilized. They supplement the cutting and ligating enzymes to prepare the types of ends required for the DNA molecules participating in a splicing operation. Some of them are briefly discussed below.

1. Alkaline Phosphatase

The enzyme alkaline phosphatase removes the phosphate moiety at the 5′-end of a DNA strand, whether it is part of a blunt end single strand extension or a recessed end of a double-stranded DNA. The PO_4 at an RNA terminal is also removed by this enzyme. It is obtained commercially from two sources: bacterial and calf intestinal phosphatases (BIP and CIP, respectively). Alternatively, raising the pH to 8.3 after restriction enzyme digestion, also dephosphorylates the DNA strands.

2. Polynucleotide Kinase

Polynucleotide Kinase is a phosphorylating enzyme that transfers the γ phosphate of ATP to a dephosphorylated end of a DNA or RNA. The enzyme is encoded by a gene of phage T4 and is extracted from *E. coli* cells infected with the phage. Kinasing is utilized to rephosphorylate DNA or RNA dephosphorylation with alkaline phosphatase, or for end labelling of DNA and RNA strands. The ^{32}P label is in the γ phosphate of ATP. Mg^{++} and dithiothreitol are used in the reaction.

3. Exonuclease III

The enzyme exonuclease III is generally a phage lambda-encoded product, extracted from the phage-infected *E. coli.* Its main activity is that of removing nucleotides from a DNA strand, one at a time, in the 3′- and 5′-direction. A second activity is that of adding a phosphate to the 3′-end of an ss DNA or a ds DNA. A third function is to disrupt a DNA at places that lack a purine base. Finally it is capable of acting as RNase H. Mg ions are needed in the reaction.

An Exonuclease III is used to produce a recessed 3′-end in a ds DNA. Wu has utilized this method to degrade part of each strand in a ds DNA fragment to generate single-stranded DNA templates to be used for a sequencing procedure.

Removal of a terminal phosphate group becomes necessary for a variety of operations. One such operation is for end-labelling of a DNA strand for sequencing and another for restriction mapping. The phosphate is first removed by the phosphatase and a radio-labelled one added in its place by a phosphorylating enzyme.

The most common need for dephosphorylating DNA ends arises during the making of recombinant DNA molecules. Lack of a phosphate group at the end of

one DNA strand prevents the formation of a phosphodiester bond with another DNA as both PO_4 and OH moieties are needed for the reaction. This lacuna prevents both circularization of fragments, (donor or vector) as well as the formation of concatamers of solely vector or solely donor DNAs. The procedure, therefore, when splicing a donor fragment with a vector DNA is to dephosphorylate one of these two sets and mix it with the other sets of fragments. The two types will be aligned by their sticky ends, and joined by one of the strands (Fig. 2.7). The gaps left in the backbone due to each of the phosphates in one DNA become, repaired in the cell into which the rDNA is introduced.

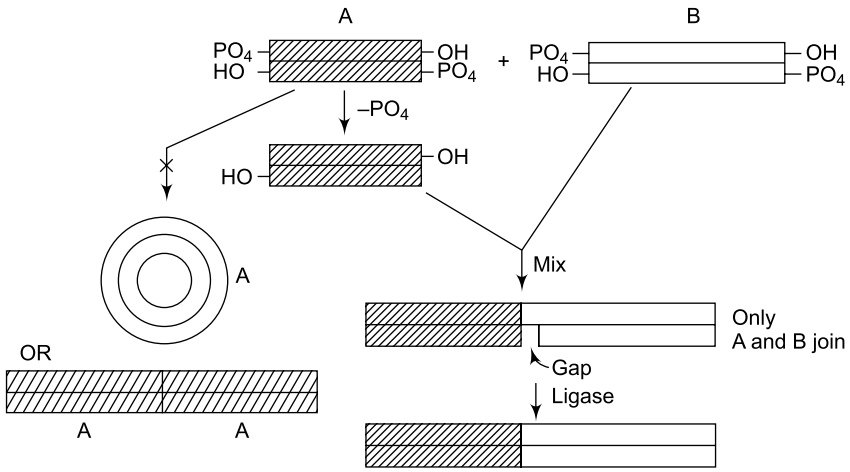


Fig. 2.7 *Dephosphorylation of ends of one DNA prevents ligation with the same DNA. This technique is utilized to ensure the joining of donor fragments only to vector ones.*

Calf intestinal phosphatase, (CIP) is preferred for most purposes, as it is easily denatured at 68°C in the presence of SDS. Bacterial intestinal phosphatase (BIP) is not affected at this temperature, and, if used, requires several phenol-chloroform extractions to remove its traces from the DNA to be eventually ligated. BIP is utilized to dephosphorylate the end of an RNA molecule, the secondary structure of which hampers the removal of phosphate by CIP. The latter may be used for all DNA, single-stranded or double-stranded with blunt, sticky tailed or recessed 5′-ends. The reaction with CIP is usually carried out at 37°C but with BIP a much higher temperature (60°C) is required.

4. DNase I

DNase I or deoxyribonuclease I is an endonuclease that generates single-stranded fragments having the phosphate at the 5′-end of the strand or of the single released mononucleotide. It can act on both ss and ds DNA.

The enzyme mostly used is derived from bovine pancreas. The strands may be cleaved at random sites or at about the same place on both strands. Addition of Mg^{++} and Mn^{++}, respectively, ensures the former and latter results.

5. Mung-Bean and S1 Nucleases

Both mung-bean and S1 nucleases cleave ss DNA into single nucleotides or into short sequences of nucleotides. Excess enzyme is capable of degrading ds DNA. The mung-bean nuclease appears to require somewhat different conditions for its activity *vis-à-vis* S1 nuclease. The latter can severe a ds DNA if there is a single nucleotide gap in one of the strands. The other enzyme operates only on larger deletions in the strand.

S1 nuclease is used particularly for DNA footprinting to remove the ss DNA not protected by the bound protein. It is also used for removing ss regions of DNA not hybridized by another nucleic acid strand. The mung-bean enzyme is derived from the legume and the S1 from the fungus *Aspergillus oryzae.*

6. DNA Polymerases and the Klenow Fragment

The DNA polymerase that is generally utilized is either the DNA Pol I from *E. coli* or the T4 DNA polymerase encoded by the phage gene.

The *E. coli* enzyme is basically a proofreading and repair enzyme. It is composed of three subunits each with a specific activity. They are 5′→3′-polymerase, 3′–5′-exonuclease, and 5′→3′-exonuclease. The enzyme is useful for synthesizing short lengths of a DNA strand, especially by the nick translation method [Sec. 2.94]. The 5′→3′ exonuclease activity may be deleted. This edited enzyme is referred to as the Klenow fragment.

The T4 DNA Pol possesses, like the Klenow fragment, only the polymerase and proofreading (3′–5′-exonuclease) functions.

The choice of DNA Pol to be used depends on the nature of the substrate. For instance, the T4 DNA Pol can elongate both ss DNA and ds DNA, the latter from both exposed and recessed 3′-OH ends. Pol I can extend a nicked DNA, starting from the exposed 3′-OH of the nucleotide on one side of the nick. Pol I and the Klenow fragment repair mismatches at the ends and fill up gaps in either strand. In addition, the Klenow fragment can elongate a primer aligned to a template strand of DNA. The latter activity is exploited in the enzymatic method of DNA sequencing developed by F Sanger and A R Coulson. The lack of 5′ → 3′-exonuclease activity makes the Klenow fragment the most suitable of the DNA polymerases for certain jobs. One of them is filling up a recessed end where the latter is a 3′-OH one.

7. Terminal Deoxynucleotidyl Transferase

Terminal transferase is a DNA polymerase that can extend a strand (from the OH-end of a primer, hydrogen paired to a complementary strand) without using a template. Any nucleotide that is provided in the reaction mixture is utilized to elongate the DNA strand, of course, with a random sequence of bases. If only one kind of nucleotide is provided a mononucleotide polymer will be produced.

The preceding strategy is employed to add 100-base-long ‘tails’ to the ends of restriction fragments. If the vector and donor fragments synthesize tails with complementary bases, only these two types of fragments join.

8. RNA Dependent DNA Polymerase

RNA dependent DNA polymerase is reverse transcriptase. This enzyme synthesizes a single strand of DNA along an RNA template. It can also synthesize a second strand along the first one to make a ds complementary or cDNA.

Reverse transcriptase is usually utilized to copy mRNAs into ss or ds cDNA, and to make short labelled probes.

9. RNases

Two RNases used in general are RNase A and RNase T1. Both cleave the phosphodiester bond between adjacent ribonucleotides. However, RNase A cleaves next to uracils and cytosines such that the phosphate remains with these pyrimidines. The nucleotide on the other side of the phosphate in the DNA backbone is dephosphorylated. RNase A is obtained from the bovine pancreas.

RNase T1, on the other hand, cleaves only next to a guanosine, again including the phosphate group at the 3′-end of the nucleotide. This enzyme is derived from *A. oryzae.*

10. RNase H

RNase H is an endoribonuclease that is useful for degrading the RNA strand from a DNA:RNA hybrid molecule. It cuts up the RNA into short fragments.

2.4 ❑ VECTORS IN VOGUE OR IN THE OFFING

Vectors for carrying the foreign DNA have been derived from naturally occurring plasmids and genomes of bacteriophages and viruses. A vector must possess certain minimum qualifications to be an efficient agent for the transfer, maintenance and amplification of the passenger DNA.

The very first vectors utilized for cloning DNA were plasmids. Next came the phages that infect *E. coli,* the cell almost exclusively used in the first days of gene cloning. These vectors were also utilized for cloning of eukaryotic DNA. It was soon apparent, however, that vectors appropriate for eukaryotic cells were needed for effective expression of eukaryotic genes. Attention was, therefore, focused on eukaryotic viruses. These were mostly animal viruses. Gene transfer into plants was a later development; plasmids and viruses that enter plants in the natural course of infections came into prominence at a later date.

As genetic manipulation at the molecular level improved, the original vectors were modified at will to acquire characteristics that made them valuable and versatile cloning agents.

The discovery of viral (including bacteriophage) vectors is very fortuitous. These vectors possess several advantages not enjoyed by plasmid vectors that enter host cells mainly by transformation. These advantages include the following:

1. Due to the natural infectivity mechanism of the phage or virus, many more cells receive the DNA than what is possible with transformation techniques using plasmid DNA, or transfection with donor DNA.

2. Viral and phage vectors can be equipped with selectable markers; this allows easy identification of clones carrying the recombinant DNA.
3. Several viral signals for transcription are more efficient than those accompanying the gene to be cloned. A gene cloned in a viral vector thus produces a higher level of products than what is possible in their natural settings. Overproduction of a gene product may also be engineered by linking the gene to strong viral promoters and certain mammalian genes that are known to amplify under specified conditions (see amplifiable vectors, Sec. 2.4.5.).
4. Many viruses have a very broad host range. Those which are restricted to only a few species may be made to extend this range by including the DNA in the shell of a broad host range virus.

Of the prokaryotic viruses, phage lambda and M13 have provided the main 'training ground' for gene splicers. Of the animal viruses, the one that was used first (and still utilized for defined projects) is the monkey virus SV40. The adenoviruses and retroviruses were felt to have greater potential for animal cell gene transfers, as these viruses infect several host species. The viruses in lytic cycles are good for studies of transient gene expression. Those viruses that become integrated 'transform' the animal cell permanently, but as only a very few—often just one—copies of the vector are integrated per cell, such vectors are not appropriate as producers of large quantities of a gene product. Recently one of the papilloma viruses, that cause warts when integrated, has been hailed as the most useful of animal viruses for use in overproduction of an inserted gene product. This is BPV-1 that exists as multicopies in the host cell and correctly expresses genes inserted into its genome.

Indeed the success with the phage and animal virus vectors has triggered off searches for other such vectors appropriate for human beings, animals and insects. The viruses which infect insects have come in for their place in the limelight as they can be exploited to undo the insect pests that are currently tackled with heavy doses of undegradable and so, environmentally undesirable chemicals. These insect viruses that may be tailored and sprayed on their hosts are referred to as 'biociders'.

Vectors are also in demand for making vaccines. Vaccines have been traditionally made from attenuated pathogens. The genetic engineer has visualized vaccines that are vectors carrying the pathogenic antigens as genes. Multistrain or multispecies vaccines may be prepared in this way. The Herpes simplex-1 and vaccinia viruses appear to be good candidates for developing such vaccines.

Let us briefly survey the vectors that are available for cloning DNA in *E. coli,* animals, insects and yeasts. The ones suitable for gene transfer in plants will be described in the section on plant genetic engineering (Sec. 5.4).

It may be mentioned here that vectors have been designed that are also 'safe' ones. These 'crippled' vectors, therefore, do not pose any biohazards if inadvertently they find themselves outside the laboratory environments.

2.4.1 Vectors from Plasmids

Plasmids are small, double-stranded, covalently closed circular DNA molecules present mostly in prokaryotic cells. A small plasmid, 2 μ, has been identified in the yeast *Saccharomyces cerevisiae.*

Every plasmid carries one or more genes, one of which is required for DNA replication, and one for a quality referred to as 'plasmid incompatibility'. Two plasmids that can survive and replicate in the same cell are said to belong to separate incompatibility groups. Those plasmids that do not coexist in the same cell belong to the same groups. There are about 30 incompatibility groups, named I, II, III,..., M, P, Q, and so on.

Some plasmids carry, in addition, a gene or genes that confer resistance to the cell to drugs, toxins or toxic metals. Others carry genes that encode bacterial toxins. Still others possess a set of genes that are required for a very special function (e.g., nitrogen-fixing genes, *nif,* of *Rhizobium* and *Klebsiella* species; tumour-causing genes of *Agrobacteria* species).

Some plasmids carry genes that allow the plasmid to self-transfer to another cell via a conjugation tube (e.g., F and other conjugative plasmids). There are some that do not possess the transfer or *tra* genes of conjugative plasmids, but a couple of mobilizing genes (*mob*), one of which encodes a protein that on complexing with the second site, causes the superhelically twisted plasmid DNA to relax and be nicked at a specific place on one strand. The free nicked end mobilizes the transfer of the plasmid to another cell. So, some plasmids are autotransmissible, while some are mobilizable. Still others can be mobilized if they possess the second *mob* element and have access to the other mob protein in *trans* from another plasmid.

A natural plasmid is large and too unwieldly for manipulations, and also may possess several cleavage sites for the same enzyme. They are not suitable as vector material.

1. Essential Features of a Plasmid Vector

To be a vector, a plasmid must possess certain minimum qualifications. They include the following:

(i) Small size Smaller DNA molecules pass through the cell membrane more easily than the larger ones. Linearized plasmids, or any linear DNA molecules pass more slowly than the circular ones. The efficiency of introduction into a cell depends, therefore, both on the size and the conformation of the molecule.

(ii) Replication *Ori,* incompatibility gene, *par* sequences Presence of a replication origin (*Ori*), and incompatibility (*Inc*) and partitioning (*par*) sequences are essential.

The *Ori* is the region where DNA replication is initiated. The *par* is required for distributing daughter plasmids stably to new cells. The *Inc* function determines the plasmid copy phenotype. Some plasmids replicate into several copies per cell, some only into 1–3 copies. The *Inc* function is somehow associated with this feature.

The reason underlying incompatibility has been partially unveiled in the case of at least one plasmid, Col E1, In Col E1, an RNA primer (0.5 kb) is initiated upstream of the *Ori.* This primer is degraded at its 3′-end if it overshoots the *Ori* region, by an enzyme RNase H. A second RNA is transcribed in the opposite strand, at the *Ori* region, ending at the initiation site of the first RNA. The second RNA is called RNA I, and the first is RNA II. Both these RNAs assume secondary structures somewhat like those of tRNA molecules; a loop in each possesses bases that are complementary to each other, like the codon-anticodon pairing between tRNA and

mRNA. RNA I prevents further primer (RNA II) formation after the first bout of DNA synthesis, by engaging the RNA II at the coupled loops. A mutation in the pairing bases of RNA I was found to have lost the power to inhibit further RNA II synthesis, so that DNA replication continued for several rounds.

Perhaps plasmids that are in the same incompatibility group possess the same or very similar RNA I and RNA II pairing regions, which compete with each other, while those that coexist are independent from the point of view of requirements of DNA replication.

When using different vectors in the same cell, one has, therefore, to make certain that they belong to separate *Inc* groups. Plasmids F, pSC101 and RP4 belong to different groups and may be used together, while pMB1 and Col E1 may not, as they are members of the same *Inc* group.

(iii) Absence of transfer or mobilizing genes Some plasmids are autotransmissible due to the presence of *tra* genes; others are mobilizable by virtue of the *mob* functions. A vector cannot afford to possess genes that allow it to dispense wantonly the cloned DNA, which has been acquired at a considerable cost of time and effort.

(iv) Marker genes One or more marker genes that allow easy detection of the recombinant clones must be present in the plasmid vector. The most commonly used markers are drug-resistance genes borne on transposons. *E. coli* vectors carry one or more of the following markers: resistance to tetracycline, ampicillin (β-lactamase gene on Tn3), kanamycin (*aph* on Tn5), streptomycin, and chloramphenicol.

(v) Unique restriction sites There must be one each unique site of cleavage for several restriction enzymes, with a few in each marker gene. Use of any of these enzymes would linearize the plasmid, by cutting only at one place. The cutting sites on a vector also allow the insertion and removal of a passenger DNA (gene or otherwise).

(vi) Multicopy nature Some plasmids occur in only a few (1–3) copies per cell. These copies replicate only once during the synthesis of the bacterial DNA. These plasmids are said to exhibit a *stringent* mode of replication. There are other plasmids, the replication of which is not linked with that of the chromosomal DNA, and are not inhibited by the presence of chloramphenicol which interferes with protein synthesis and stops chromosomal DNA synthesis. Such plasmids exhibit the *relaxed* mode of replication. Treating a cell containing relaxed plasmids with chloramphenicol stops cell division, but not plasmid replication. This allows the plasmid to form as many as 1000 copies per cell.

Since the aim of gene cloning is to recover amplified copies of the cloned DNA, the plasmid vector must be of the relaxed type.

It goes without saying that no naturally occurring plasmid is a paragon of all these virtues. Hence, for use as vectors several plasmids have been tailored by adding, deleting, and substituting specific regions of natural and tailored plasmids. Even the copy number trait of a plasmid has been altered where needed. A low copy number plasmid (stringent) is made amplifiable by joining it to a relaxed plasmid. The stringent pDF41 was spliced to the relaxed mode pMK16 to yield a larger replicon that can be increased in number by chloramphenicol treatment.

Various specialized plasmid vectors have been constructed that have insertions of sequences required for specific functions. Some of these will be described under the section on 'special plasmid vectors'.

2. Some Common Plasmid Vectors

The parents for a large number of plasmid vectors are pSC101 (stringent), Col E1 (relaxed), pMB1, R6 and p15A. The last one is a *cryptic* mini-plasmid, that has no genes.

A short list of some of the plasmid vectors in vogue is given in Table 2.1.

Table 2.1 *A Few of the Commonly Used Plasmid Vectors*

		Size	Markers	*Cutting site* (*one each*)
1.	pBR322	4.3 kb	Tet^R, Amp^R	9 enzymes
2.	pCR1	11.4 kb	Col^R, Kan^R	2 enzymes
3.	pACYC177	3.8 kb	Amp^R, Kan^R	8 enzymes
4.	pACYC184	4.0 kb	Ch_1^R, Kan^R	4 enzymes
5.	pMB9	4.4 kb	Col^R, Tet^R	6 enzymes
6.	pAT153	4.2 kb	Tet^R, Amp^R	8 enzymes
7.	pUC		Amp^R, *lacZ*	

Four of the often-utilized vectors are now described briefly.

(i) pBR322 Perhaps the most exploited vector is pBR322. It is the source of a large number of uniquely designed vectors.

pBR322 is a composite of sequences from Col E1, pMB1 and pSC101. Its geneology is an excellent example of the power of the art of cutting and pasting DNA (Fig. 2.8).

pBR322 is a 4.3 kb DNA (Fig. 2.9) carrying the β-lactamase (*bla*) gene on Tn3 (that confers ampicillin resistance) and a tetracycline resistance gene (with four open reading frames). The Amp^R was lifted from an R plasmid and the Tet^R from pSC101. The *ori* and sequences for relaxed replication mode were originally present in the Col E1 parent. It possesses the *nic* sequence, which is recognized by *mob* proteins that mediate plasmid mobilization. If there is another plasmid in the cell with pBR322 that expresses the *mob* gene, the vector could be co-mobilized into another cell. To avoid this altogether the plasmid pAT153 was developed with a deletion in the *nic* region (Twigg and Sherrat).

(ii) pACYC177 pACYC177 is a plasmid vector (Fig. 2.10) with the replicon of p15A, the cryptic small natural plasmid, that has been joined to a portion of the kanamycin resistance gene (from a transposon on plasmid R6–5) and the ampicillin resistance gene (Tn3). It is a relaxed mode plasmid, 3.45 kb in length (Change and Cohen, 1978), and pACYC177 and r vectors derived from it are compatible with Col E1 based plasmids. Both may be employed in the same cell, a strategy often required for housing two sets of genes on separate plasmids.

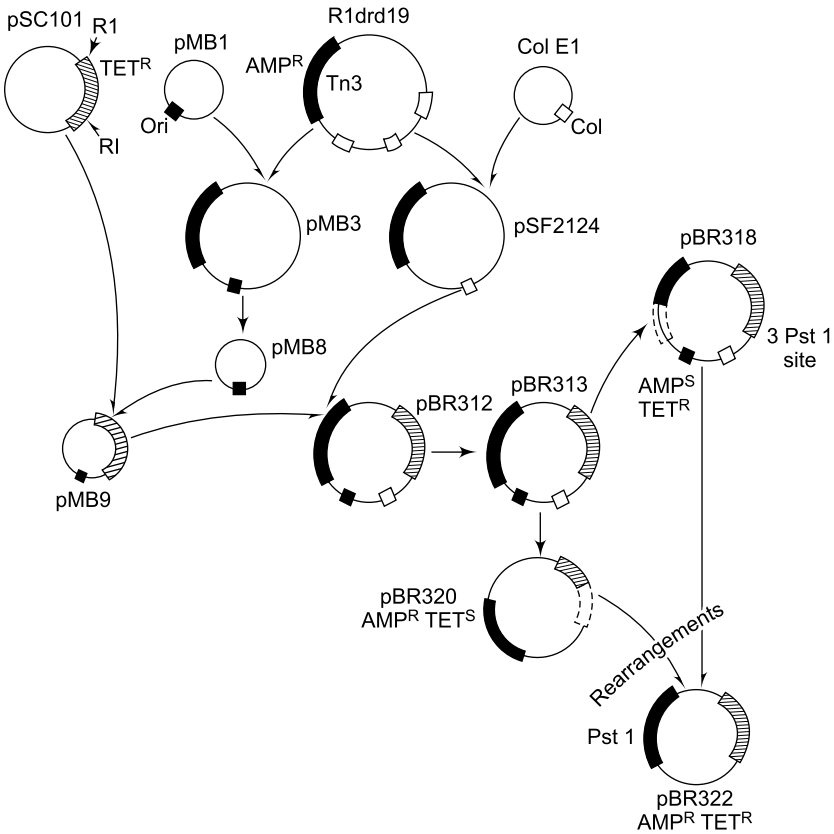


Fig. 2.8 *Ancestry of pBR322. Judicious additions, deletions and rearrangements of fragments of plasmids R1, pSC101, pMB and pColE1 generated pBR313 which, after further combinations and rearrangements with regions from pBR318 and pBR320 resulted in pBR322.*

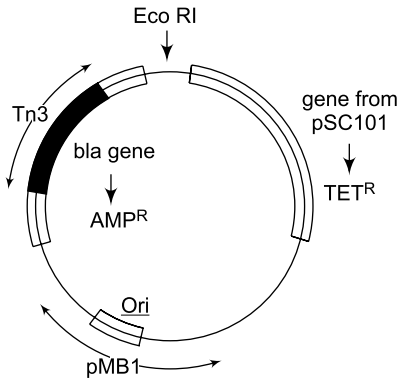


Fig. 2.9 *The origins of regions of pBR322: a schematic representation. The bla gene encodes β-lactamase and is taken from the transposon Tn3. This gene confers ampicillin resistance to the cell carrying it. The Tet^R gene confers resistance to the antibiotic tetracycline.*

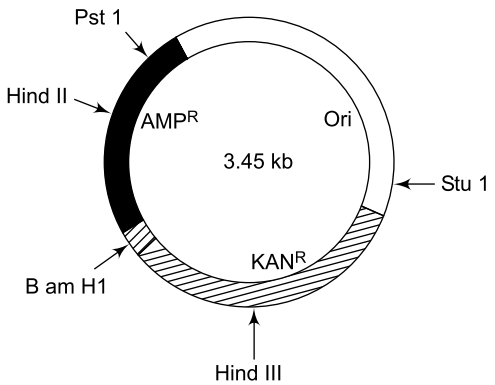


Fig. 2.10 *pACYC177 is a plasmid with a replicon of the cryptic plasmid p15A from E. coli strain 15.*

(iii) πVX This is a 902 bp Col E1 derived plasmid, into which a *supF* gene and a polylinker have been added (Fig. 2.11).

(iv) pUC pUC plasmids (Fig. 2.12) contain polylinkers, which are a set of cutting sites on a synthetic DNA, for different restriction enzymes which are used as cloning sites. Also pUC8 contains the *lacZ* and the Amp^R genes.

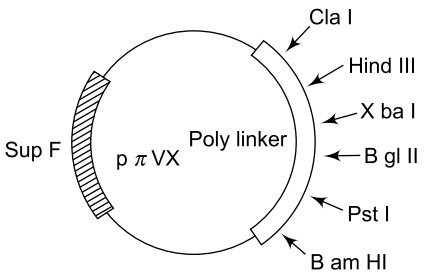


Fig. 2.11 *πVX is a microplasmid carrying a polylinker and a supF gene that suppresses amber mutation (mutation that alters the codon for an amino acid residue to a stop codon).*

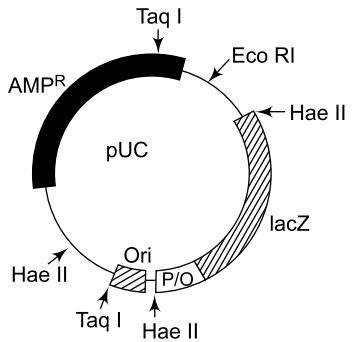


Fig. 2.12 *pUC represents a family of plasmid vectors each containing within a base pBR322 DNA the following: an Amp^R gene, an insert with the Lac promoter-operator (p/o) and a few codons of the E. coli lacZ gene, and a polylinker. Different pUCs differ in the length, composition and arrangement of restriction enzyme cutting sites.*

2.4.2 From Bacteriophages

Two bacteriophages have been tapped extensively for use as cloning vehicles. One of these is phage lambda, and the other phage M13. The former has a double-stranded and the latter a single-stranded DNA genome, respectively. Both use *E. coli* as the host species.

1. Phage Lambda

The DNA of phage lambda (ϕλ) is about 50 kb (48.6 kb) long. It has a single-stranded 12-base extension at each end. The two extensions are complementary and can align to convert the linear molecule into a circular one. They are, therefore, referred to as the 'sticky' or cohesive ends. These *cos* ends are essential signals for packaging the DNA into a mature phage particle.

(i) Biology Phage λ attaches to a receptor (product of gene *lam* of *E. coli*) on the host cell surface by a protein (product of gene *J* of the phage, in the tail region). After attachment of the particle to the cell surface, the DNA is injected into the cell, where it circularizes. The circular DNA replicates by the theta-mechanism (like the *E. coli* DNA) for about 10 minutes. Each of the many circular molecules then switches to a rolling circle method of DNA synthesis, as a result of which long concatamers of DNA become available. This becomes cleaved into 'lambda-size' DNAs by an endonuclease which requires two *cos* sites flanking a length of DNA for its activity. The *cos* sites are cut in a staggered manner generating the 12-base long sticky ends in the process (Fig. 2.13).

The transition from the theta to the rolling circle way of DNA replication is vulnerable to endonuclease V encoded by the *E. coli* genes *recB* and *recC.* However, the phage DNA remains unscathed due to the fortuitous presence of a protein that inactivates endonuclease V. The protein is a product of the phage gene *gam.*

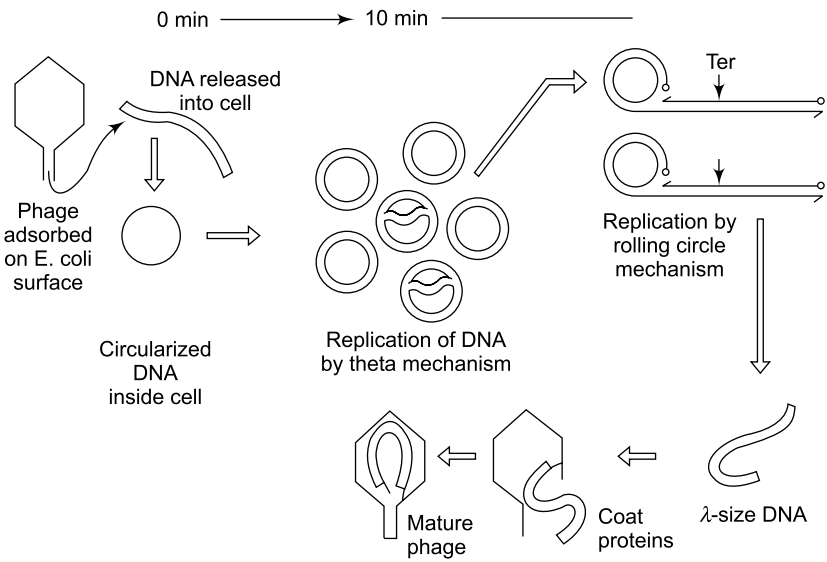


Fig. 2.13 *Biology of the DNA of bacteriophage lambda. On infection of an E. coli host cell, the injected phage DNA circularizes and replicates by the theta-mode. Later, the replication mode changes to the rolling circle one. Concatamers of lambda DNA are formed, from which lamda-sized DNA (~50 kb) are cleaved out and enclosed in a phage shell. The sites of cleavage are known as cos (cohesive end) regions; the enzyme encoded by the ter gene cleaves the cos in a staggered manner to generate 12-base long single-stranded extensions or tails.*

When 50–100 phage particles are ready, the product of phage gene *S* mediates the lysis of the host cell by destabilizing its membrane. The released phages infect fresh cells and lyse them in turn. Infection by a single phage particle, therefore, results in the degradation of host cells; this becomes evident as clear 'plaques' on an opaque 'lawn' of *E. coli.*

The lambda is a temperate phage, which means that it can take recourse to a second lifestyle (Fig. 2.14). Instead of producing a 'burst' of phages and destroying the host cell, the DNA may become integrated into the *E. coli* DNA and be replicated once per cell division as part of the cell DNA. In this case, the cell is not lysed. The cell with an integrated phage genome is called a lysogen. *E. coli* lysogens of the phage λ are written as *E. coli* (λ), if the phage was P1, it would be written as *E. coli* (P1).

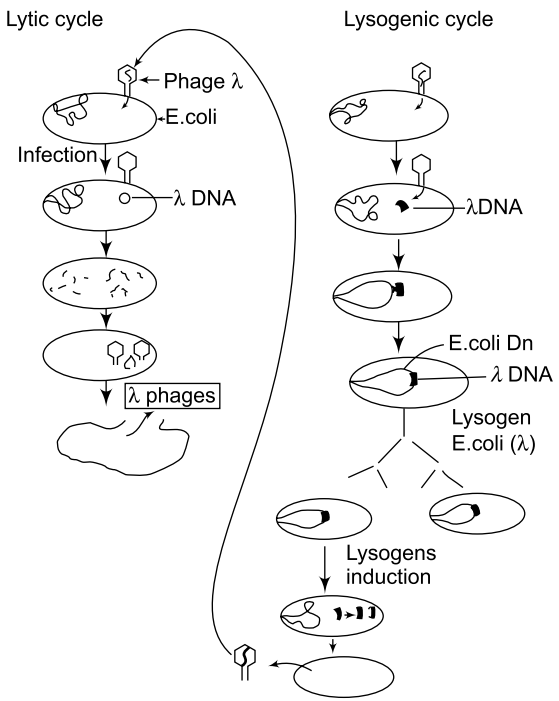


Fig. 2.14 *Lytic and lysogenic options of bacteriophage lambda. Free mature phages result from the lytic phase; the host cell is disrupted to release the phage particles. The cell with the phage DNA integrated in its own DNA is called a lysogen, and the integrated phage, the prophage.*

A very subtle play of gene regulation directs a phage λ to opt for the lytic or the lysogenic lifestyle. These are six key regulatory genes, *cro, N, Q, CI, CII* and *CIII* that control the choice of and maintain the chosen mode.

Phage lambda has genes on both strands (Fig. 2.15), that are initially transcribed from promoters P_R and P_L in the rightward and leftward directions, respectively. Both mRNAs stop at termination sites t_{R1} and t_{L1} having transcribed the genes *cro* and *N* respectively. The *N* product is an antiterminator that allows the halted

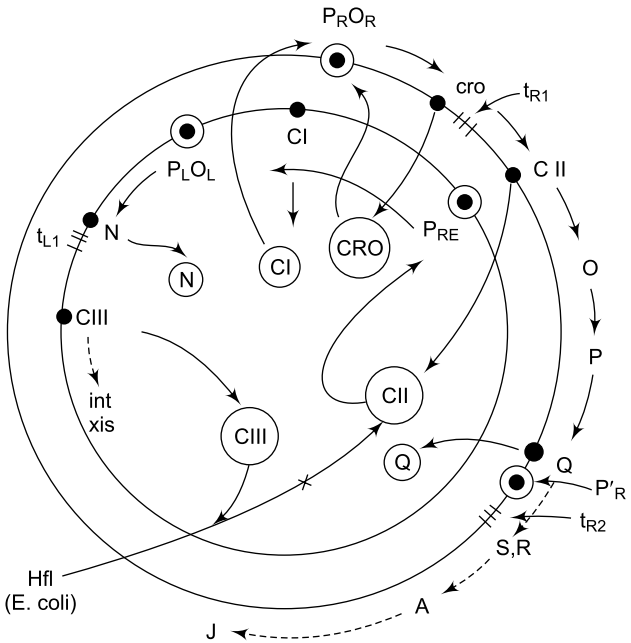


Fig. 2.15 *Gene regulation is a concerted process in the lytic and lysogenic phases of phage lambda. Transcription from P_R, P'_R and P_L results in proteins CRO, N, CII, CIII, OP, QRS, and A to Z (lytic phase). Production of CII leads to the expression of CI, the lambda-repressor protein, which stops the lytic phase and initiates and maintains the lysogenic phase.*

RNA polymerases to proceed across the t_{R1} and t_{L1}. The rightward mRNA elongates to include transcripts of *CIII* and those of *O* and *P.* The latter two are the DNA replication components. Once these genes are expressed the DNA replicates by the theta-mode. The leftward mRNA is elongated to encompass the *CII* gene. Both CII and CIII proteins are now in the cytoplasm of the cell. CIII inactivates an *E. coli* gene product, HfI, that would otherwise degrade CII. The latter is required for the initiation of transcription from another promoter (P_{RE}) upstream of the P_L region. P_{RE} establishes the expression of gene *CI,* the product of which is called the *repressor.* The mRNA initiated at P_{RE} continues beyond CI and stops before the P_L site.

In the meantime, when sufficient CRO has accumulated, a few of its molecules wrap up the P_R and P_L sites and prevent further initiation of transcription of the early genes (*cro, N*). Once CRO appears in a large quantity on the scene, it competes successfully with the RNA polymerase for the sites in P_R and P_L. This prevents expression of all the early genes and hence of all the other genes required to carry out the lytic cycle. The mRNA that transcribed *O* and *P* is extended to include gene *Q.* The Q protein initiates transcription from a fourth promoter (rightward strand) next to *Q.* This last mRNA continues until *S, R,* and *A* to *J* genes are transcribed. *S* and *R* are required for cell lysis. *A* to *J* are required for constructing the protein shell (head and tail) of the phage. When CI becomes operative these genes are not expressed. The CI also initiates transcription (leftward) from genes *int* and *xis,* the

products of which are required to cut the host DNA at a unique 15 bp-motif-containing region and also the phage DNA at an identical motif (*att* site) and promote site-specific recombination between the lambda and *E. coli* DNAs. The phage DNA thus becomes integrated into the host cell DNA. As CI is responsible for stopping the lytic cycle and maintaining the lysogenic one, it is known as a 'lambda-repressor'. Lysogens are not lysed. If cells are infected with free phages, it takes a while before the lysogens start to appear. Hence, the latter are seen as opaque spots within the already present clear plaques. As the sizes of the lysogen colonies increase, the plaques appear turbid. Turbid plaques on *E. coli* are, therefore, due to the presence of active CI, and, conversely, clear plaques are due to the absence of CI activity. This feature is exploited in various ways by genetic engineers.

The foregoing is a very sketchy outline of the life cycles of phage-lambda. The fascinating intricacies of fine balancing of gene products that determine developmental pathways in this bacteriophage, have been revealed primarily through the courtesy of recombinant DNA technology. These details have also shed light on the molecular mechanisms of gene regulation in other systems.

(ii) Vectors For the gene manipulator, the lytic cycle is useful, as this provides large numbers of phages, and is a means, therefore, to amplify any insert DNA in the vector.

The lysogenic cycle genes occupy around 20 kb of the DNA (Fig. 2.16). This region (about 40% of the genome) is not essential for propagation of the phage and is, therefore, dispensable (from the genetic engineer's point of view). An insert DNA may safely replace this region and not affect the production of free phages, provided the total length of the hybrid DNA is within the limits of packagable DNA. The phage cannot form mature particles unless the DNA is of the 'lambda' size. This may be stretched somewhat: the shell can accept 78% to about 105% of this length. In practical terms, an insert of 10–25 kb may be used, depending on the manner in which the hybrid vector is constructed.

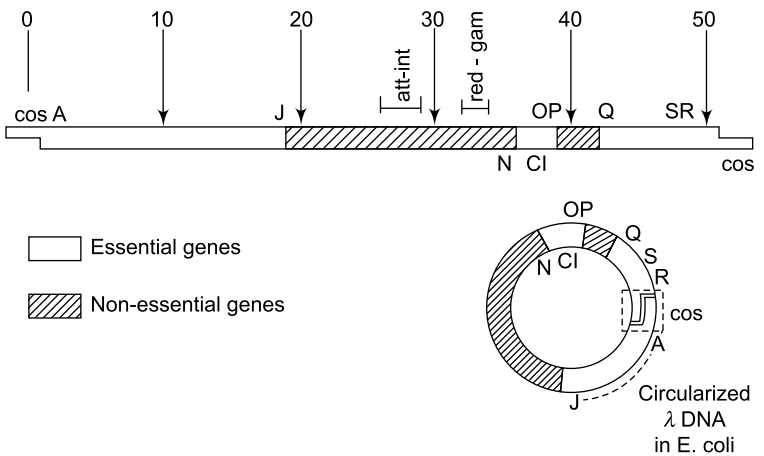


Fig. 2.16 *Regions of the phage lambda DNA that are non-essential for the lytic phase, and are, therefore, dispensable in a recombinant DNA molecule. This region, named the 'stuffer' is replaced by the DNA fragment to be cloned.*

To use the lambda DNA, the length of which has to remain within the prescribed limits, different strategies had to be devised for small and large inserts.

Small inserts (10 kb or less) are introduced by cutting the lambda DNA at one cutting site in the non-essential region. A large insert (10 kb or more) is introduced by replacing a piece from the non-essential region. The former are the *insertion* and the latter the *replacement* vectors (Fig. 2.17). A look at the map of lambda DNA will show that the replaceable DNA carries, besides the *CI* and *N* genes, genes called *gam* and *red.* The *gam* product inactivates the *E. coli recBC* encoded endonuclease V. The *red* is responsible for recombination event in the lambda DNA.

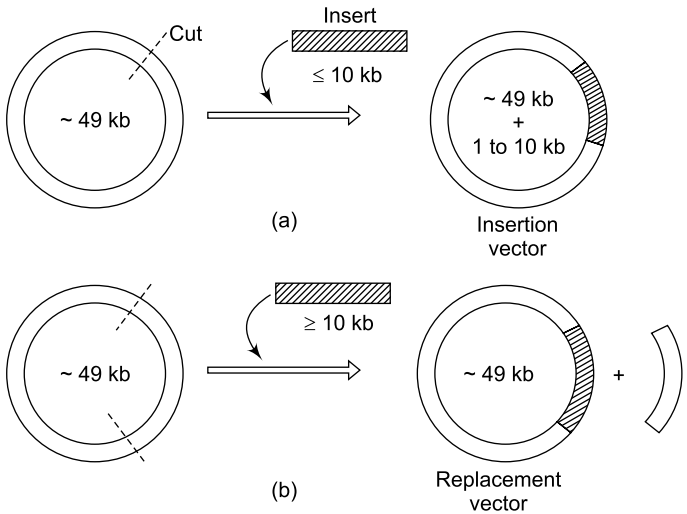


Fig. 2.17 *Insertion and replacement vectors from phage lambda DNA. Insertion vectors are linearized at one unique restriction site to ligate the DNA fragment to be cloned. A dispensable length of the phage DNA is, on the other hand, replaced with foreign DNA in a replacement vector. The central or stuffer region of the phage DNA, non-essential for the lytic phase, is the segment that may be replaced.*

A lambda lacking the *gam* gene does not grow on *E. coli recB*$^+$ *C*$^+$ unless the phage DNA carries a special 8 base-pair motif, called the *chi* (*cross over hotspot instigator*) which is a dominant mutation. The *chi* signal induces the host *recA* gene to join monomeric DNA circles into large multimeric ones (*recA* is unable to do this on its own). Of course, one could use a *recB*$^-$*C*$^-$ host with a *gam* vector to avoid the impasse. The sad part is that, in that case, no phages will be formed. The system not only requires an endonuclease to cut the concatamers but also requires concatamers for the action of the nuclease. The phage particles are formed only from concatamers. Circular concatamers may be produced by recombination of several monomeric circular DNAs. This recombination is normally mediated by the *redA* (also called *exo*) gene. Since the latter is also lost with the *gam* in replacement vectors, the job can be done with the help of the *recA* gene of *E. coli.*

The above peculiarities of the phage lambda genome and life cycle can be exploited variously for specific purposes. This has been done and is continuing to be the source of better and more versatile λ-based vectors.

One of the earliest replacement vectors was from generalized transducing (gt) phages and are referred to as 'λ-gt vectors'. The first such phage used for the purpose was a non-lethal mutant that was viable in spite of a loss of some sequences in the right arm. The lost region was that occupied in the wild-type phage DNA by the last two (rightmost) of the five *Eco* RI sites on the DNA. This length thus became available for a little longer insert than before. These mutations could be made into vectors by removing the two fragments in the non-essential region, between the first three *Eco* RI sites. These fragments are the B (at map position 21.8–26.6 kb) and C (between 26.6 and 32.14 kb) regions. A-λgt with the B replaced is called λ-gt B and that with the C region replaced as λ-gt C. As the two regions add up to about 23 kb of DNA, a vector lacking both B and C can insert a maximum length of 25–28 kb (5% more is manageable).

The λ-gt C vectors were altered suitably to make them more versatile. The λ-DNA was mapped for other restriction enzyme sites. This extended the cloning potential of these vectors.

Many λ-gt vectors also possess a small deletion (*nin*) at the t_{R2} terminator adjacent to the Q locus. t_{R2} is a stronger terminator than the t_{R1}, after *cro*, and requires the Q protein for a read-through of this site. A *nin* mutant thus transcribes the middle and late genes at a stretch, without halting, and thus ensures the production of phage particles.

Replacement of the non-essential region means a loss of function of the dispensed genes. This loss becomes a useful marker for successful insertion of the donor DNA. Only a few representative examples of such selectable vectors are discussed here, as it is impossible even to hint at the versatility of the vectors that have emerged from phage λ alone or as hybrid vectors containing useful regions of the λ genome.

(a) Immunity vectors Murray has constructed a vector with an *Eco* RI site within the *CI* gene. This gene lies within what is known as the immunity region. Such immunity insertion vectors or λ-*gt* (imm^{434} b527) can accommodate up to a maximum of around 7.6 kb of donor DNA. The imm^{434} and b527 refer to deletions in the lambda-map at sites 49.1-57.4% and 72.9-79.3% respectively.

A plasmid vector has also been prepared which utilizes inactivation of the *CI* gene as a marker. The *CI* and P_RO_R region of λ-DNA has been joined in this plasmid just upstream of a tetracycline resistance gene. The CI protein represses transcription by binding to the O_R sequences (operator sequences that overlap the three 17 bp segments of P_R). A cloning site within the *CI* gene is used to insert the foreign DNA. This inactivates the *CI* gene, an event that is reflected in the constitutive expression of the tetracycline resistance gene. Hence, cells that grow in the presence of this antibiotic are the ones harbouring the recombinant DNA.

(b) Charon phages These are a series of phages constructed by Blattner et al. (1979) to be used as replacement vectors (Fig. 2.18). The upstream promoter regions and part of the *lacZ* gene of the *lac* operon of *E. coli* are spliced to the replacement vector. Several restriction enzyme sites (called cloning sites if used to introduce a donor at these sites) are included in different regions of the promoter and in the *lacZ* sequences. A donor DNA inserted in any one of these sites inactivates *lacZ* expression, a fact that is easily assayed by a visual screening method (the *lacZ*-XGAL method, described later). These vectors were christened *'Charon'* after the Greek mythological ferryman of that name.

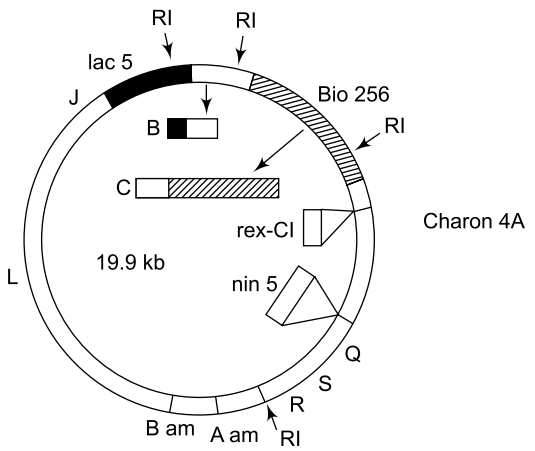


Fig. 2.18 *Charon 4A is a lambda-based vector with four Eco RI sites at which the enzyme cleaves, creating four fragments the left (L) and right (R) arms and the stuffer fragments B and C. The rex-CI region and the t_{R2} region (nin 5) are deleted. This allows a longer foreign DNA insert as well as an N-independent transcription of the late lytic phase genes; lac 5 and bio 256 are sequences from the lactose and biotin operons of E. coli DNA. Aam and Bam are amber mutations. The lac sequences are utilized for selection of hybrid DNA clones.*

(c) The λ-gt WES vectors These vectors (Fig. 2.19) are particularly safe ones. There is a stop codon (amber) mutation each in genes *W* (phage assembly), *E* (shell protein), and *S* (cell lysis), and one of two deletions (λ B or C) within the non-essential region, flanked by *Eco* RI sites. These triple mutants, λ-gt WESB and λ-gt WESC, are unable to grow on *E. coli* unless the latter has suppressors (mutations in tRNA that negate amber or other mutations in the mRNA being decoded) for the *W, E* and S genes. Strains su II and su III are such suppressor mutations containing strains of *E. coli.*

(d) The red-gam deletion identifiers: the spi phenotype A group of vectors possess a central replacement region flanked by *Bam* HI sites. The removable DNA contains the *red-gam* sites. A red^+ gam^+ phage cannot grow in *E. coli* (P2); no plaques are formed. The phenotype exhibited by the wild-type is known as Spi or sensitive to P2 interferences. A phage in which the *red* and *gam* are deleted (red^- gam^-), grows on *E. coli* (P2), i.e., are of the Spi phenotype. Clear plaques on the *E. coli* (P2) lawn, therefore, indicate successful replacement by the donor DNA.

The presence of the donor DNA, and not just the lack of the *red, gam* section, is confirmed by the production of viable infectious phage particles, as the latter cannot be formed unless the length of the DNA is within the limits of packaging. This would happen only if the insert was in place in the λ-DNA. The DNA obtained from joining of the two λ-DNA arms would be too short to give mature particles.

(iii) To construct an rDNA with a λ replacement vector To begin with, both the donor and vector DNAs have to be isolated. Care must be taken to avoid shearing of the high molecular weight donor DNA. It is important that all cut ends are due to restriction cutting, and not due to mechanical breaking at random sequences,

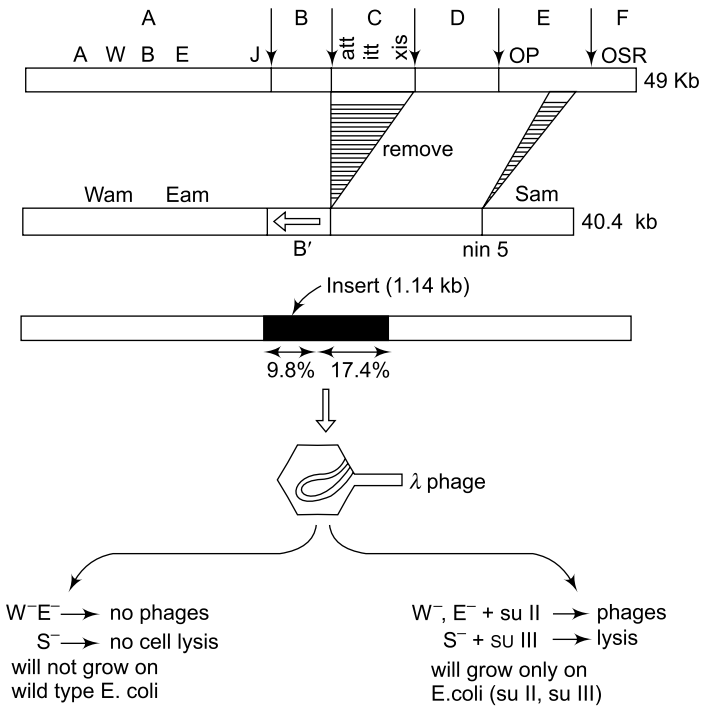


Fig. 2.19 *λ-gtWES B′/C is a lambda DNA-based replacement vector, where segments B or C may be replaced by foreign DNA. Amber mutations in the W and E genes prevent assembly of mature phage particles, while one in gene S (for lysozyme) prevents lysis of the infected host cell. These deficiencies prevent this vector from surviving outside the laboratory where a very special host, with a suppressor of amber mutations has to be provided for maintaining it. It is thus one of the 'safe' vectors. B′ represents the B segment accidentally joined in the reverse orientation. B and C together represent 27.2% of the phage DNA. The nin 5 deletion makes place for another 6.1% of insert DNA.*

otherwise many of the fragments (as well as vectors) will be wasted. Mechanical breakages are most likely to occur during the phenol extraction steps. To avoid such breakages either the mixing is done very gently or a protocol used that does not require phenol extraction. One such procedure consists of isolating nuclei from homogenized cells that are frozen in liquid nitrogen. The nuclei are lysed and the lysate spun in a tube of $CsCl_2$ at a very high speed (30,000 × g) in a Beckmann ultracentrifuge. The DNA band (visualized by the pink glow emitted when UV light is directed on ethidium bromide intercalated DNA) is recovered by puncturing the polyallomer tube with a 18-gauge syringe needle and collecting the DNA in several aliquotos. The ethidium bromide is removed by suitable treatment. This DNA is further purified by dialysis and precipitated by ethyl alcohol. The DNA is next cut with a restriction enzyme and fractionated by agarose gel electrophoresis.

The vector DNA has a left (L) and a right (R) area sandwiching a central dispensable or 'stuffer' fragment (Fig. 2.20). The stuffer possesses one site each for

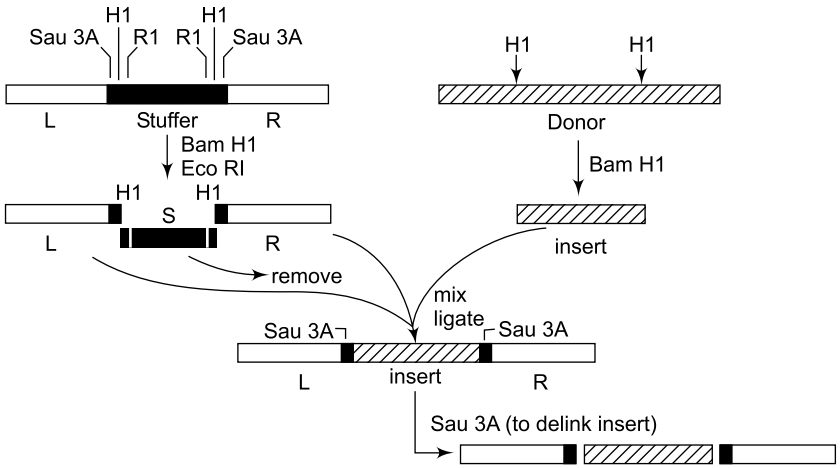


Fig. 2.20 *Construction of a recombinant DNA with a lambda replacement vector, (see text for explanation)*

three restriction enzymes one is for Bam HI one site, very often for Sau 3A, is distal to the stuffer region, and another, often Eco RI is proximal to the stuffer. The Bam HI sites are used for inserting the donor DNA also cleaved with Bam HI. The insert is now flanked by (in this example) the Sau 3A sites. The latter may be utilized to remove the insert clearly from the vector areas at a later step, if so desired.

Before isolation of the vector DNA, the phage has to be concentrated in a purified condition. This is achieved by spinning the phage suspension by $CsCl_2$ density gradient centrifugation, the phage band dialysed and the DNA extracted by phenol-chloroform-isoamyl treatment. The phenol is removed by dialysis.

The central stuffer region may be cleaved out in a number of ways. In the preceding example, Eco RI and Bam HI digestion will remove the small fragments between the sites for these enzymes; the Bam HI cut end of the stuffer will no longer join with the Eco RI end of the larger deleted stuffing. The vector areas are prevented from joining by phosphatase treatment. The sticky ends align, however, so that the vector DNA is now a joint L and R arm fragment. This is now mixed with Bam HI donor DNA cut fragment together with ligating ingredients. Some protocols include the removal of the deleted stuffer regions before the recombination step.

The donor DNA fragments are chosen from the array of restriction fragments fractionated on an electrophoretic gel. Only those in the size range of 15–25 kb are elected, cleaned by dialysis and utilized for making the hybrid DNA.

The vector areas and DNA form long concatamers. This mixture is then added to a 'packaging mix' (see next section), which contains the components needed to fashion phage particles enclosing the rDNA molecules, and an extract from λ-infected *E. coli* cells that contains the enzyme for cleaving the concatamers at the *cos* sites.

(iv) In vitro packaging of lambda DNA The linear DNA of phage lambda may be packaged *in vitro* by an ingenious method developed by Murray and Hohn (1975). Two judiciously tailored mutants of the phage are used for the purpose (Fig. 2.21). Two mutants of phage lambda, that are made to coinfect an *E. coli* host, are unable to form phage particles on their own, but together supply the wild-type gene products required to construct an infectious phage. When the coinfected cells are lysed, the lysate contains the enzymes and proteins needed to enclose the recombinant DNA in mature particles.

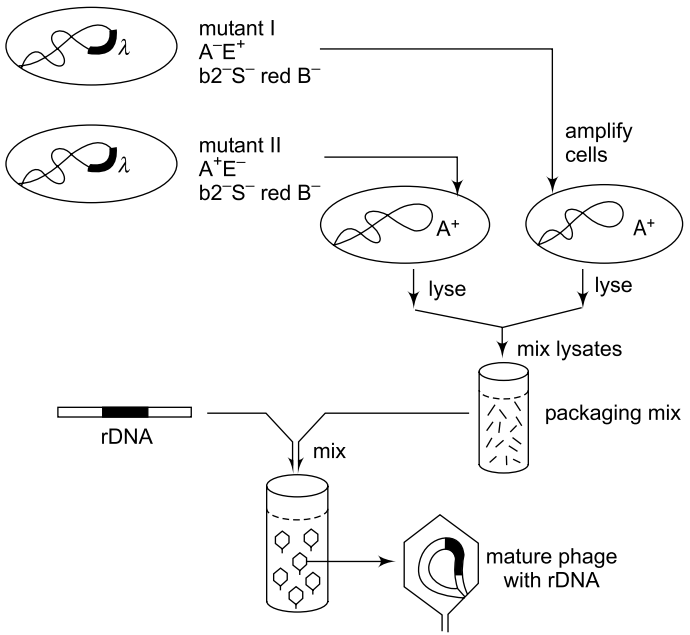


Fig. 2.21 *In vitro packaging of lambda DNA (see text for explanation)*

To understand the logic underlying the aforestated method, it would be useful to recall the functions of the genes that have been mutated in the two mutant packaging strains. Genes *A* to *E* are required for making the head proteins. Gene *S* is for cell lysis; the product of *b2* is needed for attaching the phage DNA to a capsid protein as well as for cleaving out the prophage from the *E. coli* DNA. Generalized recombination is mediated by the product of the *E. coli recA* and *redA* and *B* of the phage. One of the lambda mutants has a lesion in gene A, the other in *E.* Neither can, therefore, make complete shells. Coinfection into a cell produces a pool of wild-type proteins from which phage shells can be constructed. Both mutants possess lesions in *b2, S* and *redB.* The *b2* mutation prevents induction of the integrated DNA (both mutants are in lysogens) and also amplifies *in situ* the prophage DNA. There is a consequent accumulation of late gene products—the capsule proteins. Recombination between the DNA to be cloned and the resident DNA is prevented by the mutations in *redB.* Finally, the cells are not lysed, due to the mutant *S.* The *E. coli* cells overpacked with 'packaging material' are lysed *in vitro* and the extracts

from them mixed with the hybrid vector. The latter are efficiently packaged into mature phage particles by the enzymes and capsid proteins in the packaging extract.

2. Phage M13

M13, fd and fl belong to a group of ss DNA linear bacteriophages that infect *E. coli* carrying an F plasmid (F^+ or Hfr cells). Such cells possess hairlike extensions or pili which have receptors for an M13 surface protein.

Inside the host cell, the M13 DNA follows a peculiar pattern of replication, which has been exploited for obtaining pure collections of single strands of a passenger or carrier DNA.

(i) Biology The phage possesses a circular ss DNA, 6.4 kb in length. After entering the cell, the ss DNA becomes double-stranded (Fig. 2.22). The ds DNA circular molecule is known as the *replicative form* (RF). It replicates until there are 50–100 copies. After this, the positive (+) strands (same as infecting) are synthesized on the negative (–) template of the RFs and proteins expressed by the phage genes assemble to form the linear particles. The latter are extruded from the cell surface without damaging the cells. M13 does not, therefore, form a clear lysed plaque like lambda, but a turbid one due to the slower growth of the infected cells.

(ii) Vectors Vectors are made from the ds RF molecules. There is only one stretch of DNA (508 bp) between genes 2 and 4 (there are 10 genes) that is replaceable (Fig. 2.23). Polylinkers containing restriction sites for different endonucleases, and drug resistance markers have been inserted in this region to convert the natural genome into useful vectors. The M13 DNA can be easily extracted from any one of its growth stages. The RF is cleaved with the same enzyme as the donor DNA fragment to be joined to it, the vector and DNA fragments mixed and ligase added to the reaction mixture. The hybrid vector is introduced into *E. coli* cells that have been made receptive by a $CaCl_2$ treatment.

Messing has enhanced the versatility of the M13 vector by introducing in the RF the regulatory upstream regions and the first 145 codons of the galactosidase gene (Fig. 2.24). This region encodes one of the polypeptides of the enzyme. Such a truncated *lacZ* gene can be complemented if the cell also contains the other polypeptide portion of the gene. The latter (minus the amino terminal region) has been cloned in plasmids. *E. coli* strains JM101 and JM103 contain resident plasmids carrying this *lacZ* 15 region. The F plasmid that contains this stretch allows conjugational transfer of the plasmid to F cells. To prevent the latter, a mutation traD36 is included in the system. Hence, pili are formed, on which M13 phages can adhere, but not conjugation tubes. These vectors belong to the M13mp series.

A JM101 cell with one of Messing's M13 vectors is able to synthesize functional β-galactosidase, which converts a colourless chromogen XGAL (5-bromo-4-chloro-indolyl-β-D-galactoside) to blue non-diffusable crystals. Cells synthesizing the enzyme (natural or complemented) appear darkish blue. Cells with impaired *lacZ* genes are unable to convert XGAL and so remain colourless. Colourless plaques, therefore, indicate the presence of inactivated or missing *lacZ* genes. Messing has inserted several cloning sites within the sixth codon of the *lac* gene. If a donor DNA

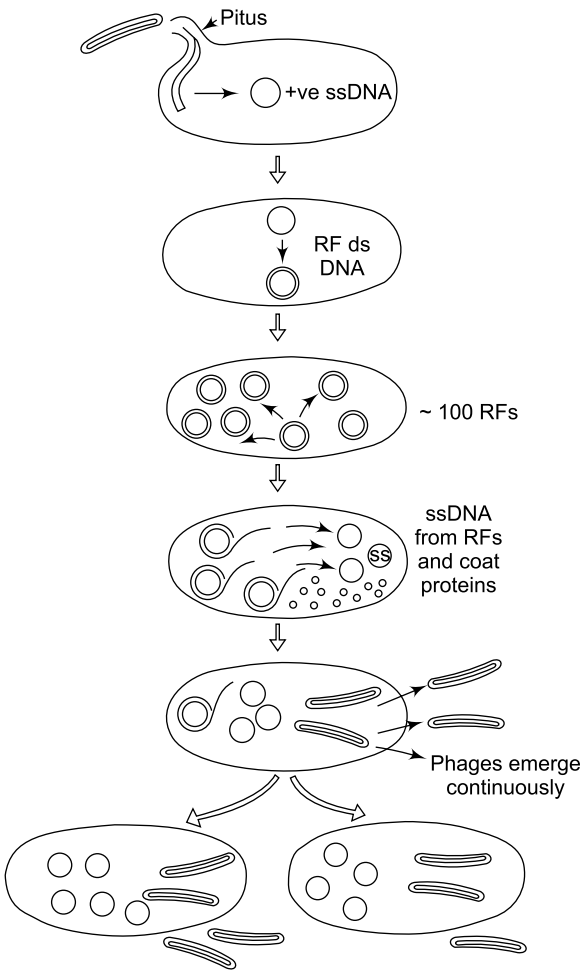


Fig. 2.22 *Biology of bacteriophage M13. This linear virus possesses a covalently closed circular (CCC) DNA which is injected into a pilus of the host E. coli cell (F^+ or Hfr). The entering +ve DNA strand is first made double by the synthesis of a complementary –ve strand. The ds DNA is known as the replicative form (RF). The latter replicates to produce 50–100 copies, each of which switches to the rolling circle mode of DNA synthesis, producing +ve ss circular DNA copies against the –ve strand template. The ss DNA is packaged in viral proteins to form mature phage particles that are extruded through the membrane without damage to the cell.*

is inserted in any one of these sites, the gene is inactivated. Recombinant plaques are colourless and are easily picked out from the background of blue coloured ones.

Two stop codon mutations have been introduced into the M13 phages to make them 'safe' ones. To allow them to grow in host cells, the latter are provided with suppressor mutations that negate the amber mutations in the phage. A cloned DNA fragment may be in the extruded phage in the correct orientation for expression from the provided signals or may be in the opposite orientation. This can be ascertained easily. Details of the above mentioned techniques will be described in later sections.

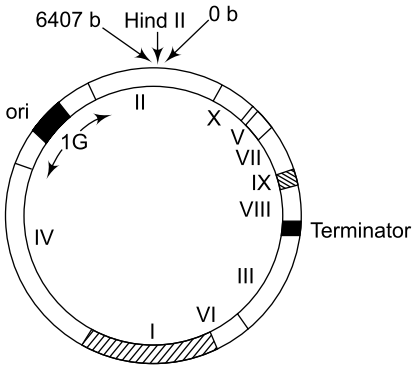


Fig. 2.23 *A schematic representation of the genetic map of the M13 DNA. There are 10 genes (I-X) and an intergenic (IG) region between genes IV and II. Vectors utilize the IG region as a cloning site.*

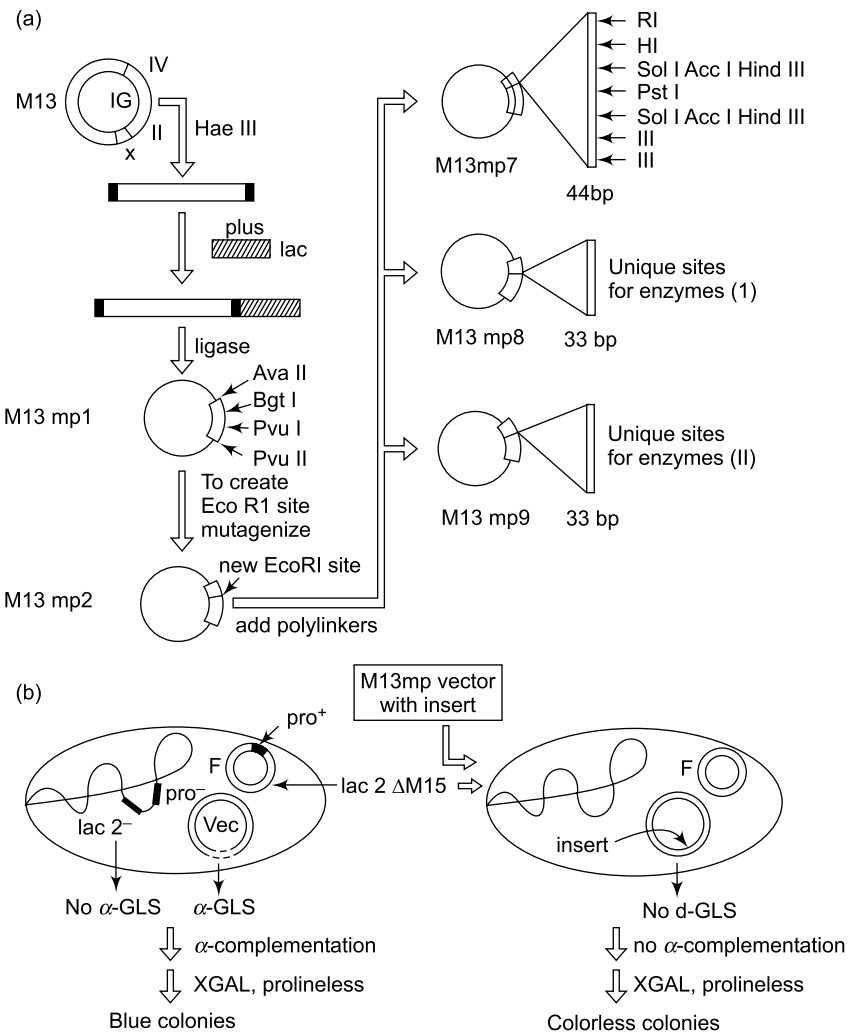


Fig. 2.24 *M13mp vectors.*

The ability of the DNA phages to switch from the RF stage to ss DNA molecules resides in certain sequences. These have been isolated from phage M13 and included in certain plasmid vectors. The latter replicate into single-stranded circular molecules. The EMBL series of plasmids are such ss DNA cloning vectors that find several specific uses in recombinant DNA technology.

Although fragments larger than 5 kb can be accommodated in the phage, such a large vector DNA is not stable. Several deletions can occur. This restricts the length of fragments clonable in M13mp vectors to 4 kb or less. This fact has been exploited extensively for situations where very small pieces of DNA need to be cloned.

Three types of activities are performed best with the small length of DNA cloned in M13 vectors. These are site-directed mutagenesis, making of small probes and DNA sequencing by the shot-gun method. All three procedures require pure collections of short, single stranded DNA. A vector tailored and packaged in the M13 phage automatically releases phage particles from an infected cell that contains only ss DNA of the same kind.

2.4.3 Cosmids

Cosmids are plasmids containing the sequences required for packaging into phage lambda particles (Fig. 2.25). The sequences include the *cos* region of lambda DNA together with about 200 bp stretches flanking both sides of the *cos.* Enclosing the cosmid in phage lambda shells allows the particles to infect *E. coli* readily and inject the cosmid into the cell. The percentage of cosmids introduced into host cells is much higher than that achieved by naked plasmids that move in through the plasma membrane. Since the particle accommodates the full lambda-sized DNA, and only the *cos* and flanking 400 bp region is required for packaging the remainder may be replaced by the donor or insert DNA.

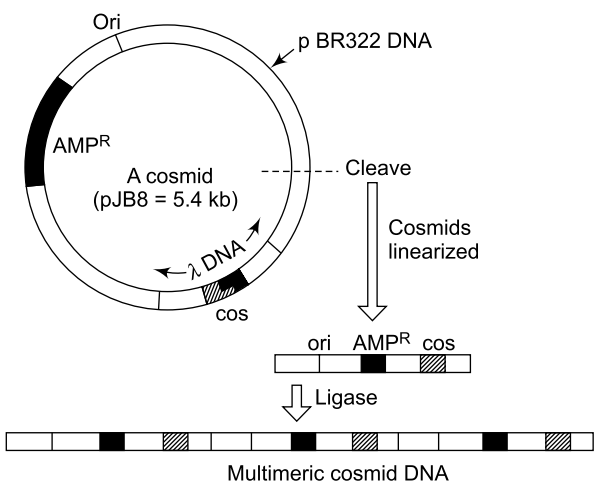


Fig. 2.25 *Cosmid pJB8. A cosmid is a plasmid (pBR322) containing the cos region from phage lambda DNA, which allows DNA to be packaged within the phage shell. A cosmid vector can thus accommodate a DNA almost as long as that of the phage. The insert (40–45 kb) is packaged into phage particles by virtue of the cos sequences. Once inside a host cell, a cosmid is a plasmid, and unlike true phage DNA, it does not go through the developmental pathways of the phage.*

Cosmids are usually derived from pBR322. Several problems associated with the replication origin of the native pBR322 have led to the introduction of the phage origin. The cosmid usually possesses one drug-resistance gene (Amp^R, Tet^R or Kan^R) and appropriate sites for restriction enzymes.

To clone a large eukaryotic DNA in a cosmid, the following steps have to be taken (the procedure was first developed by J Collins and B Hohn in 1978): (1) The donor DNA is first digested and fragments of 35–45 kb size isolated. (2) Cosmids are then linearized by cutting at one site. (3) The linear cosmid and donor DNAs are mixed and ligated. Concatamers of cosmids and donor DNAs are formed. (4) An extract from *E. coli* cells, infected with phage, is added to the above mixture, together with what is known as a packaging mix. The extract contains the *ter* (terminating) enzyme that makes staggered cuts at each *cos* site when they are separated by the usual λ length of DNA. The cleaved fragment, containing the insert DNA, becomes packaged by the shell proteins present in the mix, and infectious cosmid particles are formed. These can be used to infect *E. coli* cells, and amplify the passenger DNA by replication in the infected cells.

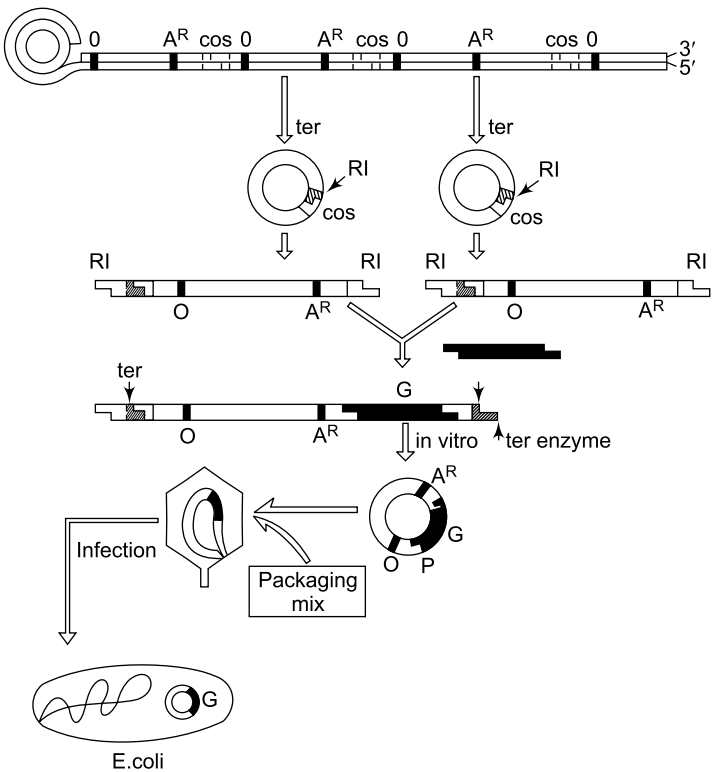


Fig. 2.26 *Cosmid cloning. Linearized cosmids are mixed with foreign DNA (35–45 kb size) and ligase added to the mixture. Concatamers of donor and cosmid DNA are formed. This is the substrate required for the cleaving activity of the lambda ter enzyme. E. coli (λ) cells that cannot be induced are lysed in vitro. The extract from these cells contains ter and the shell proteins. The concatamers are cleaved and a phage-sized DNA packed per shell. The phages containing hybrid cosmid DNA infect fresh cells and thus transform the cells with the cloned DNA.*

Several refinements have been introduced into the above basic method of Collins and Hohn. The latter created problems such as formation of separated concatamers of cosmid and DNA fragments and packaging of lambda-length DNA formed by ligation of several small non-contiguous sequences of the donor DNA. Ish-Horowicz and J F Burke published a revised and improved protocol in which the use of two restriction enzymes resulted in the elimination of much of the cosmid DNA from both ends. These methods were further modified by Maniatis, Fritsch and Sambrook.

The ability to accommodate 35–45 kb of insert DNA and greater efficiency of infecting host cells in comparison to naked plasmids make cosmids valuable cloning vectors for a variety of purposes. These include making gene libraries, analyses of large regions of the genome (particularly of eukaryotic genes that may be 20 kb or more in length) and in the method of genome analysis known as *chromosome walking* (Fig. 2.27). Briefly, the latter consists of making a cosmid library with large overlapping sequences, which are sequenced individually. By arranging the fragments with the overlaps in position, one is able to 'walk' from one sequence to the next and thus determine the sequence of a very large piece of the donor DNA; the latter may even be an entire genome. 'Walking' is aided by hybridization of the donor fragments in different cosmid clones with probes made from ends of the fragments.

Attractive as cosmids appear, there still remain several pitfalls that make their use less popular than that of phage vectors. Improvements are being reported in cosmid cloning protocols continuously, with some already quite adequate for certain special goals.

2.4.4 From Animal Viruses

Animal viruses that have been exploited the earliest for vector material are the SV40, human adenoviruses and retroviruses; others which appear to have potential as routine vectors for gene transfer into animal cells, are some parvoviruses and vaccinia viruses. In addition, vectors have been constructed which are plasmids containing cis-acting elements taken from animal viruses.

Some of the vectors made from animal viruses are packaged in viral shells and infect host cells with a very high efficiency. Others are not packaged and have to be introduced into host cells by transfection. The degree of success, in this case, depends on the method of introduction. Some animal vectors are good for a short time, expressing the vector-borne gene(s) for a limited period until the host cell is killed. Others become stable components of the host system and express the insert genes constitutively (continuously). Both transient and permanent or stable expression systems are valuable for specific types of studies. For example, the small papovavirus SV40 infects mammalian cells (monkey) and produces a progeny of particles before killing it. SV40 DNA vectors allow transient expression of genes carried on them. The number of RNA molecules and proteins synthesized during this limited period are sufficient for a thorough study of their characteristics. On the other hand, vectors from some papilloma and Herpes virus families remain as uncoated plasmids within host cells, which are not lysed, and continue to be source(s) of transcript(s) and protein(s) from the insert gene(s).

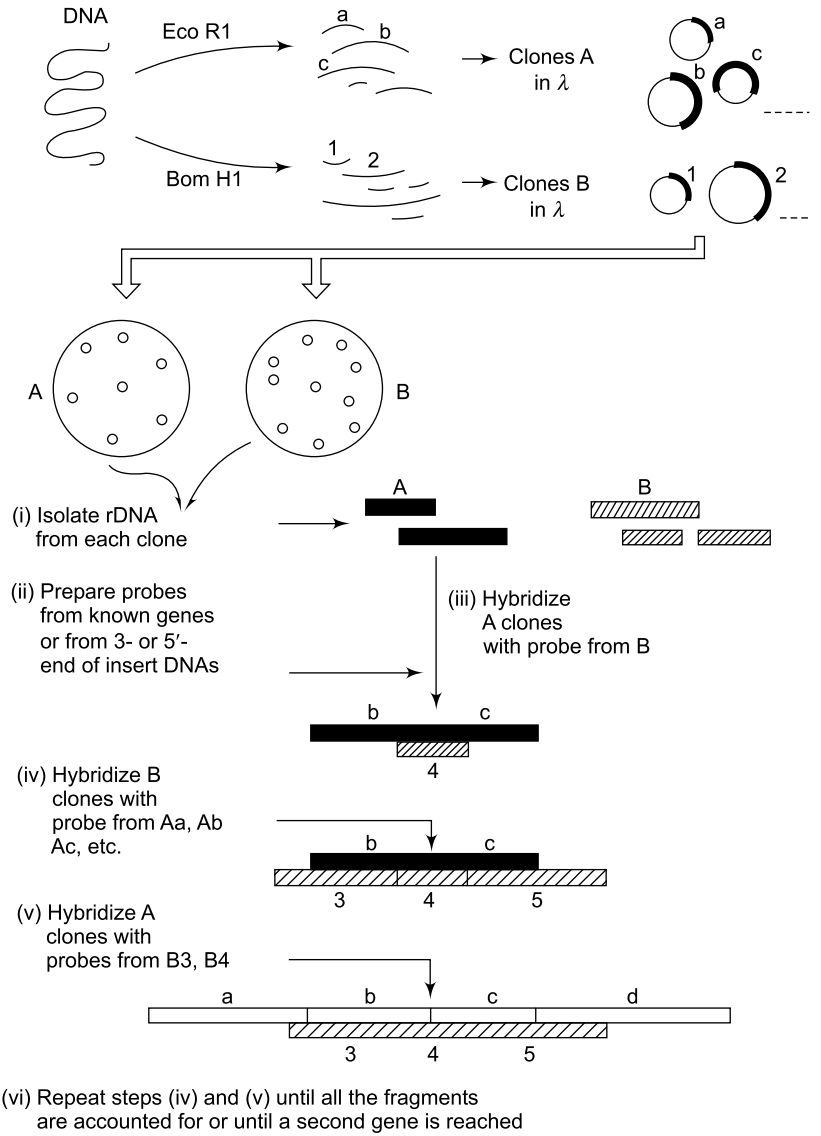


Fig. 2.27 *Chromosome walking. This is a technique for allocating sites for adjacent genes in the total DNA or a fragment of a DNA. The DNA is cleaved by two different restriction enzymes in separate lots A and B. Probes are prepared from genes known to be present in the particular DNA under scrutiny and/or from the 3′- or 5′-ends of the restriction fragments. The restriction fragments of one lot (A) are mixed with a probe from B. The fragments of A that hybridize with the B probe are used as probes for searching the A sequences in B. By repeating this process, the overlapping A and B fragments can be aigned into a restriction map and the positions of the genes indicated, without going through the longer process of DNA sequencing. One thus 'walks' from one gene locus to the other and fixes their positions on the DNA of interest.*

Most animal viruses possess only little (if any) stretch of the DNA, removal of which does not affect the activities of the viral genome. It is customary, therefore, to use vectors made from them in conjunction with helper viruses or cells.

1. Simian Virus 40

Simian virus 40 belongs to the group of papovaviruses and is capable of existing in different hosts as a lytic or productive virus or as a transforming one. Its double-stranded DNA has been exploited the most as vectors for gene transfer into animal cells.

(i) Biology SV40 leads a lytic existence in monkey cells and a transforming life-style in rodent cells. Its double-stranded DNA is about 5.3 kb long and forms a covalently closed circular molecule (Fig. 2.28). The viral particle adsorbs to the host cell surface and is endocytosed into the cytoplasm, where the viral coat is shed.

Within the permissive monkey cells two genes are expressed immediately after infection. These early genes represent two tumour causing antigens, namely a small tumour (t) and a large tumour (T) proteins (antigens). The function of the T antigen (100 kdal) is understood better than that of the t antigen (18 kdal). The T and t mRNAs are produced by differential spicing of a common initial transcript.

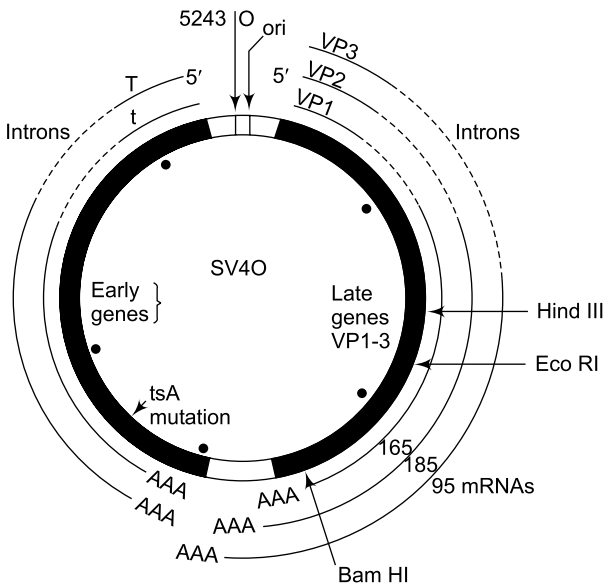


Fig. 2.28 *A schematic representation of the map of SV40 DNA. The early (E) genes for antigens T and t, and the late (L) ones for viral coat proteins VP1–3 are transcribed from different strands. The origin of replication and initiation of transcription of both E and L genes are in the region between the start of both sets of genes. On infection of a permissive host (monkey), host cell RNA polymerase begins expression of the early genes. The T antigen is required for initiation of DNA replication. Transcription of the L genes begins simultaneously with initiation of DNA synthesis. Differential splicing results in the unique T and t transcripts, and in those of the late genes, although initiation begins from a common site in both cases.*

The T antigen is required to initiate replication of the SV40 DNA. The protein binds itself around the replication origin. The T antigen is also instrumental in controlling the transcription of the early and late genes. Late transcription begins only after DNA replication is under way. The three late genes for viral coat proteins VP1, VP2 and VP3 are initiated from a common site but become distinguished into three sizes of mRNAs—16S, 19S and 18S (for VP1, VP2, and VP3, respectively), each size consisting of subsets of mRNA that differ in the regions that have been cut out by splicing. The transcripts are from overlapping genes that are decoded from separate reading frames. The polypeptides that they produce have a common carbonyl but different start or amino-ends. The 16S, 19S and 18S subsets also differ in the lengths of the leader regions. The VP3 mRNA has no spliced out region. The viral proteins, together with the enclosed DNA, form mature virus particles that are released from the lysed host cell.

The t and T antigens are expressed in non-permissive mouse cells. However, the DNA becomes integrated into the chromosomal DNA of the monkey (a permissive host) which continues to express the early genes in this integrated or *proviral* position. The provirus may be induced (released) by fusing the rodent cell with a monkey cell. A rodent cell with an SV40 provirus acquires the characteristic features of a transformed cell— spherical shape, serum independence, loss of contact inhibition and lectin agglutinability. These features are used as markers for the transformed state.

(ii) Vectors The SV40 was hailed as a much needed naturally infecting vector for transferring genes into animal cells. However, as its intricate system of gene regulation of splicing of transcripts became known, it was obvious that it could not be used as a general vector system in animals. The difficulties posed by the SV40 life-style include the following: (a) the genome is very small (5.24 kb) and no part of this is really dispensable to allow insertion of a donor DNA; (b) integration of the SV40 DNA is accompanied by unpredictable and variable sequence rearrangements in the viral DNA, host DNA or in both; and (c) the splice regions, though cleaved out in the mature transcript, appear to have some say in the successful expression of the gene; it is essential to ensure, therefore, that any replacement vector from this DNA contains the relevant splice regions.

In spite of the above mentioned deterrents, several vectors have been designed from time to time to serve specific goals. Some of these are *late replacement vectors* to be used in conjunction with a helper virus. Others are *early replacement vectors* that function only in specially designed helper or COS cells. Vectors which become packaged into viral particles are also known as infectious vectors, while those which are not packaged are the non-infectious ones.

Replacement vectors are made by deleting the early or late gene regions. Late replacement vectors are more in demand than early replacement ones, as the genes of the former are expressed in greater amounts than those of the latter. The T antigen binds itself to two sites near the promoter of the early genes after the antigen has reached a certain level in the infected cell. At the same time, an excess of T antigen in the cell promotes transcription of the late genes. Hence, late replacement vectors with an inserted gene will furnish a reasonably useful level of the product

of the latter gene before the host cell is lysed. Removal of the capsid-making genes in a late replacement SV40 vector restricts amplification of the inserted gene into viral particles. This hurdle is overcome by coinfecting the cloning cell with both the recombinant vector (carrying the gene of interest) and with what is known as a *helper virus.* This has a genome that supplies the gene functions lacking in the crippled virus.

A brief note about the helper virus: The coinfecting virus may be a ts A mutant that fails to transcribe the early genes at 42°C. The defective virus with the recombinant vector is able to do so. When vector-carrying and helper virus particles are incubated together within the host cell at 41°C, the vector supplies the T antigens, and the helper supplies the proteins. The two viral DNAs thus complement each other and a mixture of helper and recombinant particles is produced (Fig. 2.29).

(a) Late replacement vectors The first useful SV40 late replacement vectors were made in 1979 by R C Mulligan, B H Howard and P Berg. These investigators cloned the rabbit β-globin gene in cultured monkey kidney cells. The vector carried a cDNA version of the globin gene mRNA in place of the coding region of the VP1 gene. A cell with a ts A mutant helper virus was coinfected with the recombinant virus. Similar success was reported by D H Homer, K D Smith, S N Bayer and P Leder (1979). P Gruss and others associated with G Khoury were able to express the p21 (oncogene) protein of Harvey murine sarcoma virus in a similar manner.

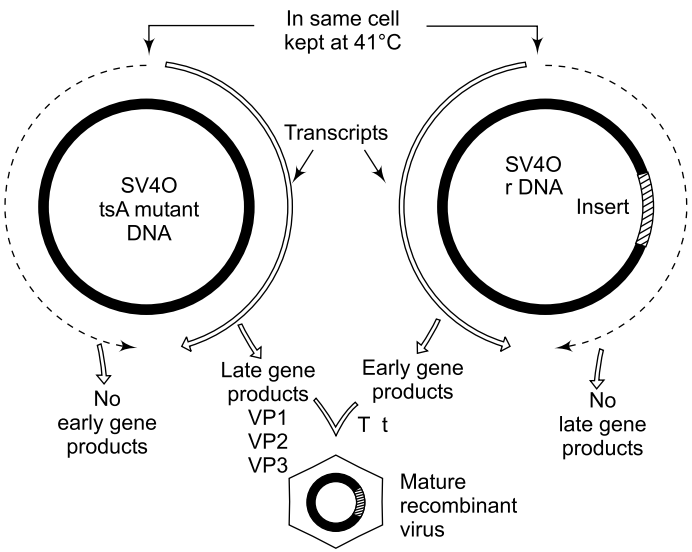


Fig. 2.29 *Late replacement vector from SV40 DNA. Foreign DNA replaces a region of the late genes, and hence inactivates the latter. No viral coat proteins are made. These are supplied in a cell coinfected with a helper SV40, with a ts A mutation. The mutation prevents expression of T antigen at 41°C. Cells containing the hybrid vector and helper DNA are kept at this higher non-permissive temperature. The coat proteins supplied by the helper and the T antigens supplied by the hybrid vector result in mature SV40 particles.*

Sometimes, the T antigen gene (A) is made ineffective by an insertion within the promoter of the early genes. There is one hitch in the use of this system; the helper virus often replicates very much more than the recombinant vector, so that only a small proportion of the particles contains the latter. Despite this shortcoming, helper viruses and replacement of late genes have been used to express several genes in monkey cells. Examples include the NS gene of influenza virus (Lamb and Lai, 1984), the EIA gene of adenovirus (Oda et al., 1983) and of the HpsAg and core antigen (Will et al., 1984).

(b) Early replacement vectors An insert DNA in place of the early genes stops expression of the T antigen. To supply this protein, necessary for SV40 DNA synthesis and expression of late genes, the vector is introduced into a cell developed by Y Gluzman, that synthesizes the missing antigen. The cell is known as a COS cell and is a cultured green monkey one (CV-1) which contains an SV40 DNA that has a UV-inactivated replication origin. The latter DNA does not replicate, but is able to synthesize the T antigen, which is utilized to replicate the vector DNA. Thus, several copies of the vector are made, each one of which expresses the inserted gene. The viral proteins are synthesized without hindrance and mature particles containing only recombinant vectors can be collected (see Figs. 2.30 and 2.31).

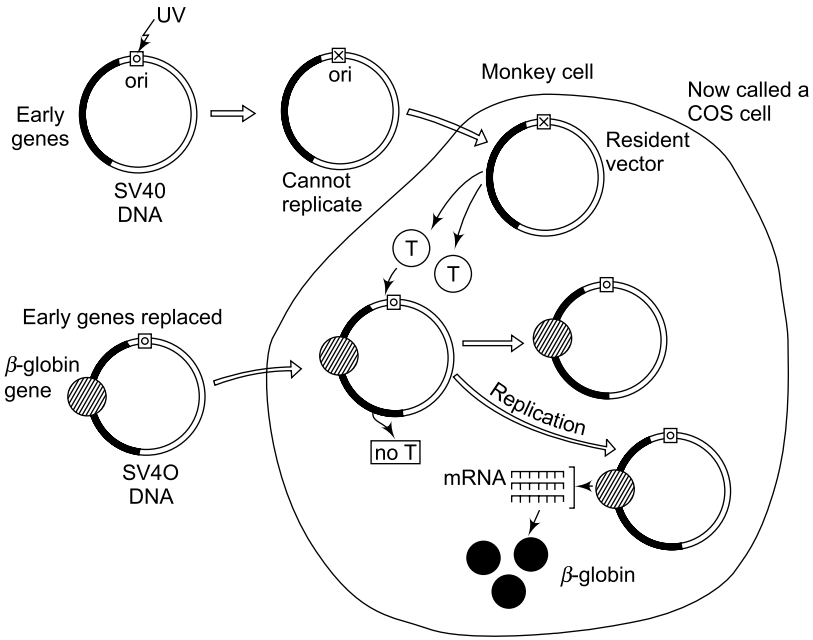


Fig. 2.30 *Early replacement vectors of SV40 are expressed in COS cells, which are cultured kidney cells of the green monkey (CV-1 cell line) possessing a resident SV40 DNA with an ori inactivated by UV irradiation. Inactivation prevents DNA replication but not T antigen expression. As the early gene region is replaced in the hybrid vector, the latter does not produce any T antigens. These are supplied by the resident vector. The vector replicates as well as expresses the proteins required to form the viral coat.*

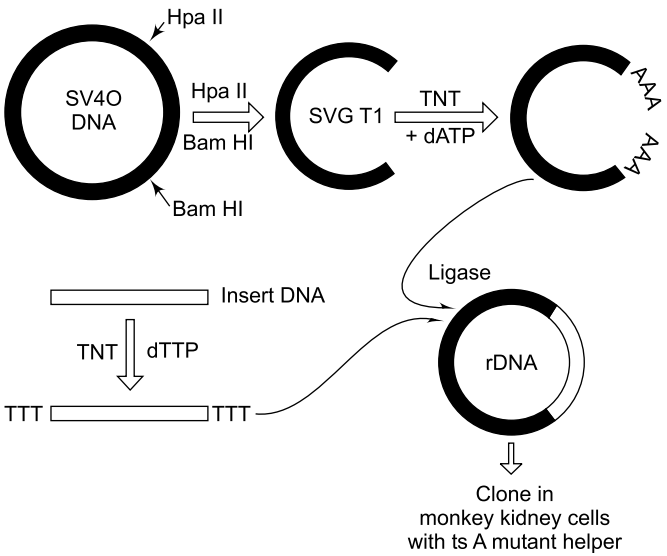


Fig. 2.31 *SVGT-1 is an SV40 vector containing an HpaII-Bam HI fragment that can be replaced by a foreign DNA. The insert and vector DNAs are joined by hompolymeric tails.*

(c) Non-infectious SV40 vectors Non-infectious SV40 vectors are plasmids (with *cis*-acting sequences of the SV40 DNA) that are not packaged into viral particles. They contain the regulatory sequences required for transcription as well as those needed for processing the transcript; splicing and polyadenylation signals are, therefore, included. A selectable marker gene is also present. Some of them are shuttle vectors (see later section) which carry replication origins and selectable markers for *E. coli* DNA. These latter vectors can then be extracted from the infected animal cell and introduced into bacterial cells, if required. Non-infectious SV40-based vectors (e.g., the pSV2 series) can be used to transfect a wide range of animal cell types (see Fig. 2.32).

Such plasmid SV40-based vectors are used fruitfully with COS cells. Boast et al. (1983) used human fibroblast cells as COS cells rather than CV-1 cells. If a replication origin of SV40 DNA is added to the vector a very large output of replicates can be recovered.

2. Adenoviruses

Adenoviruses infect human, murine, avian, bovine and simian hosts. Some of them are associated with mild respiratory diseases in man.

(i) Biology Most studies have concentrated on human adenoviruses. An adenovirus is an icosahedral particle containing a linear ds DNA that ranges in weight from 20 kdal to 30 kdal. It can lead to either a lytic or a transforming type of a life-cycle. In nature, viral particles remain within the nucleus and do not leak out until the fag end of the infection. They have to be, therefore, collected by lysing the nuclear membrane.

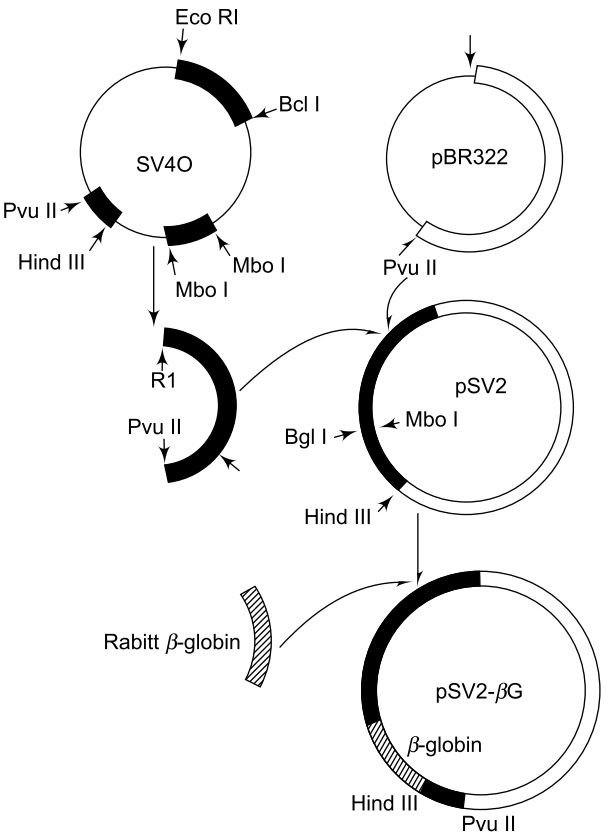


Fig. 2.32 *pSV2-β G: a non-infectious SV40 vector. A vector pSV2 was first constructed by joining portions of SV40 with an Eco RI-Pvu II fragment of pBR322. Next a β-globin gene was inserted at a Bg II site in the SV40 DNA region.*

Adenoviruses differ in surface proteins (antigens) that can be distinguished by matching antibodies. Sera containing antibodies are referred to as 'antisera', and tests utilizing them are known as 'serological assays'. Human adenoviruses of serotypes 2, 5, 7 and 12 are mostly utilized, as they are easier to cultivate than other serotypes. They grow easily in cultured epithelial cells (HeLa, KB) into free viral particles. Ad7 and Ad12 are oncogenic serotypes that cause tumour formation when injected into baby hamsters.

Ad2 has a 35 kb long DNA with a complex protein attached to the 5′-termini. There are inverted repeats at both ends of the DNA. The protein acts as a primer for viral DNA replication. The protein also improves the efficiency of transfection of Ad2 DNA into host cells. DNA replication and packaging signals are present near the inverted repeats.

Ad2 expresses four early transcripts and several sets of late ones. Both early and late mRNAs are derived by differential splicing and polyadenylation of initial transcripts that begin at a common starting-point. The early genes are on one strand of the DNA and the late ones on the other. There are splicing sites in the leader regions of

most of the mRNAs. There are three splice sites in the antileader region of most late genes. They are spliced to form a common leader for the different late transcripts. Late transcription uses a major late promoter or MLP. Adenoviruses 2 and 5 also express two late genes (VA or virus-associated genes) which are transcribed with RNA Pol III into small RNAs, VA-RNAI and VA-RNAII. The host cell is lysed when viral particles are formed, about 2–3 days after infection.

Most human adenoviruses are cultivated best on epithelial cells (e.g., human HeLa and KB cells) as monolayer or suspension cultures. Oncogenic strains do best in human embryonic kidney cells. Some human serotypes grow only on human embryo kidney cells (e.g., number 293) which have been transformed by Ad5 DNA segments.

(ii) Vectors The complexities of splicing of adenovirus mRNAs pose hurdles for the routine use of these viral DNAs as vectors. There are, however, attractive features that are expected to make adenoviruses good vector material, especially for transferring genes to human cells. The advantages include the following: (i) the blocking of cellular protein synthesis, something that does not happen with SV40 vectors; (ii) the presence of numerous control strategies for expressing the different transcripts could perhaps be exploited for achieving and/or studying mechanisms of gene control; (iii) the late proteins of the virus are synthesized in large amounts—insert genes using the *cis*-acting controls for these genes could also be made to overproduce their products.

A replacement adenovirus vector would require coinfection with a helper virus, as there are no large regions of the viral genome that may be deleted without adverse effects on the growth of the virus. The 293 cells mentioned earlier possess the E1 (early1) region of Ad5 and may be used to accept replacement vectors. The E1 product is required for adenovirus DNA replication. N Jones utilized E1-replacement vectors for cloning eukaryotic genes in 293 cells and T Shew (1979) and Grodzicker and Klessig (1980) have utilized E1-replacement vectors for cloning eukaryotic genes in 293 cells.

The pioneer adenovirus-based vector is a naturally-occurring hybrid of SV40 and adenovirus DNA packaged in viral coats. These particles grow well in monkey cells unlike pure adenoviruses. The hybrid DNA is only marginally longer than that in a regular adenovirus. The adenovirus DNA segment may be very small (about 7 kb). This would theoretically allow the remaining length (~30 kb) for the insert DNA.

Expression from the MLP is preferred. The spliced leader region appears to enhance the efficiency of transcription from this promoter. Translation of the mRNA is also aided by the presence of the spliced leader. Studies with truncated leaders preceding inserted genes show very inefficient expression.

Adenovirus vectors containing insert genes may be constructed by inserting the gene in an adenoviral DNA fragment spliced to a plasmid with the SV40 T antigen gene placed in *cis* to the adenovirus MLP. The recombinant DNA is then introduced with a helper virus into 293 cells. Mature viral particles containing the inserted gene are produced, together with others that do not have the rDNA. The latter are weeded out by passaging the particles through monkey cells, in which only the SV40 DNA containing rDNA will grow. The insert gene may be sandwiched between the MLP and the SV40 *A* gene to ensure its expression.

Recombinant adenoviral particles may also be made by replacing the E1 region with the insert DNA. The E1 represents the left part of the adenoviral DNA. The right part is added judiciously to the insert DNA to produce rDNAs with both extremities of the viral genome that are required for packaging. The rDNA is transfected into 293 cells; or, unligated left and right portions of the viral DNA (on plasmid DNA) may be cotransfected into 293 cells with the gene to be inserted. Recombination within the cell leaves adenoviral sequences containing the introduced gene. The plasmid vector DNA sequences are lost. By adding an E1(L) segment to the vector, one can recover only recombinant viruses.

The gene that has been utilized the most as an insert DNA is that for the SV40 T antigen. Success of expression was gauged by the level of T antigen synthesized. This suggests that other genes may also be expressed in the same position.

3. Retroviruses

Viruses that possess an ss RNA genome that replicates via an intermediate ds DNA copy are known as retroviruses. The ds DNA becomes integrated into the host cell DNA as a provirus that expresses its genes in a very efficient manner. This DNA is a natural vector for conveying a passenger gene into an animal cell.

There are high hopes for the use of retroviral vectors as they provide several advantages not offered by other methods of gene transfer into animals. These advantages include the following features:

(a) The insert gene becomes integrated stably into the host cell genome in a defined manner.

(b) Infectivity of cells with this virus may be as much as 100%. So, almost every cell in the culture can be made to contain the desired gene. In addition, the virus can infect all the cells in a tissue when injected *in vivo.* This offers a means for gene replacement therapy.

(c) The retroviral genome has genes, the transcripts from which are spliced in the host cell. Any split gene fitted into this genome would also be spliced; the latter can thus be recovered from a recombinant viral infection.

(d) Retroviruses can infect every cell in germ-line tissues and become part of the gametes that are produced with a high efficiency. Genes spliced to such viral genomes have been inherited stably by progeny animals and expressed developmentally in a correct manner. Creation of such transgenic animals is one of the most attractive features of using retroviral vectors.

(e) It would be obvious from the preceding enumeration that any job which requires the introduction of a foreign gene in an animal can be performed more efficiently and predictably using retroviral vectors than with other commonly utilized techniques. Genes introduced via retroviral vectors can cause insertional mutagenesis and introduce markers into particular cells or chromosomes. These are some of the approaches successfully utilized for genetical and developmental studies.

(i) Biology A retrovirus particle contains two linked copies of an ss RNA genome together with a core protein and a few molecules of the RNA-dependent DNA polymerase known as 'reverse transcriptase'. The shell is surrounded by a cellular lipoprotein membrane containing a glycoprotein encoded by a viral gene (*gag*).

On entry into a host cell the RNA is copied by the reverse transcriptase into an ss cDNA that is made duplex by a host DNA polymerase. The ds DNA becomes integrated as a provirus into the host cell DNA at, what till now appears to be, random locations (Fig. 2.33).

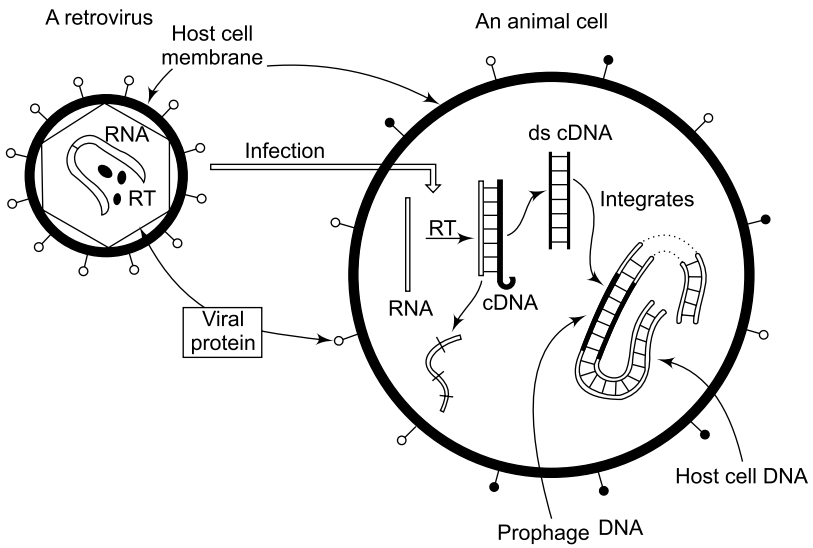


Fig. 2.33 *Biology of a retrovirus. A retrovirus possesses an ss RNA genome, which is copied, in the infected host cell, into an ss cDNA by reverse transcriptase (RT). DNA polymerase converts this into a ds cDNA, which integrates as a provirus at random sites in the host cell DNA. The cell undergoes phenotypic changes into what is referred to as a 'transformed' cell. Such a transformed cell possesses certain characteristics typical of malignant cells.*

The genome of a typical retrovirus has a characteristic set of genes and other sequences in a particular order (Fig. 2.34). The provirus has two sets of a repeated sequence, one at each end of the DNA. These are called long terminal repeats (LTRs) and have the following sequence constitutions. The left (5′-end of the genome) has a sequence named U3 followed by a sequence called R that has another sequence U5 on its 3′-end. The U3 and U5 derive their labels from the predominance of uracils within them. The right LTR is a repeat of the U3-R-U5 sequence with the U3-end following the 3′-end of the genome. The genomic RNA that is synthesized from the proviral genome lacks the extreme U units at both ends (i.e., the left repeat does not have U3 and the right has no U5). Between the LTRs are three main genes, *gag, pol* and *env,* in this same order. These genes encode the glycoprotein surface marker, the reverse transcriptase and the envelope or shell proteins, respectively.

The signals for DNA replication, transcription, shell assembly and splicing of viral mRNAs reside in the LTRs and in the genome DNA immediately adjacent to them. Transcription is initiated within the left LTR and terminated by a signal in the right LTR. Reverse transcription is a complex process that uses at least two primer attachment sites. These are immediately adjacent to the two LTRs. Both promoter and enhancer are found in the left U3 sequences, and a polyadenylation signal is present in the right LTR. The *gag* and *pol* genes are derived from a common

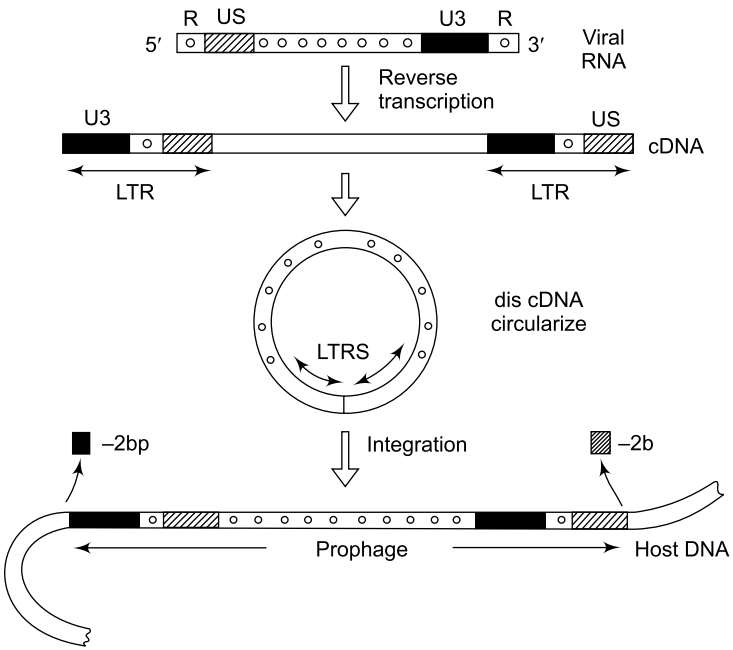


Fig. 2.34 *Schematic representations of retroviral genomes. The retroviral nucleic acid (RNA) possesses unique terminal regions composed of repeated sequences R, U3 and U5 in the pattern shown. Reverse transcription adds a U3 and a U5 sequence to the left and right repeated terminals, which are called long terminal repeats (LTRs). On integration, the viral DNA loses 2 bp at each end.*

transcript. The *env* mRNA is spliced from the initial RNA into an *env* specific mRNA. Splicing signals are present immediately adjacent to the 3′-end of the left primer binding site and 5′ to the *env* coding sequences.

The genomic RNA is of the positive kind and so is directly translated into the viral proteins. It also contains, on the 3′-side of the left splicing sequences, a signal required for packaging the genomic RNA into a viral particle. After a certain level of transcripts are made from the proviral DNA, the mRNAs are packaged into individual particles that leave the cell carrying away an envelope of the cell membrane.

Some retroviruses possess, in addition to the three common genes, a gene called an *onc* gene (oncogene) (Fig. 2.35). The presence of this gene causes cells to become transformed in an oncogenic (cancerous) manner. Retroviruses without *onc* genes also may cause cell transformation over a long period of time (several months) after infection. The mechanism underlying this is not clearly understood. The *onc* genes have been classified arbitrarily on the basis of the functions undertaken by their products. All of them appear to be involved somehow in derailing the chain of events that lead to a normal pattern of cell division. Whatever the defect caused by an oncogene, the upshot is a transformed cell that divides in an uncontrolled manner. The *onc* gene is always downstream of the last retroviral main gene.

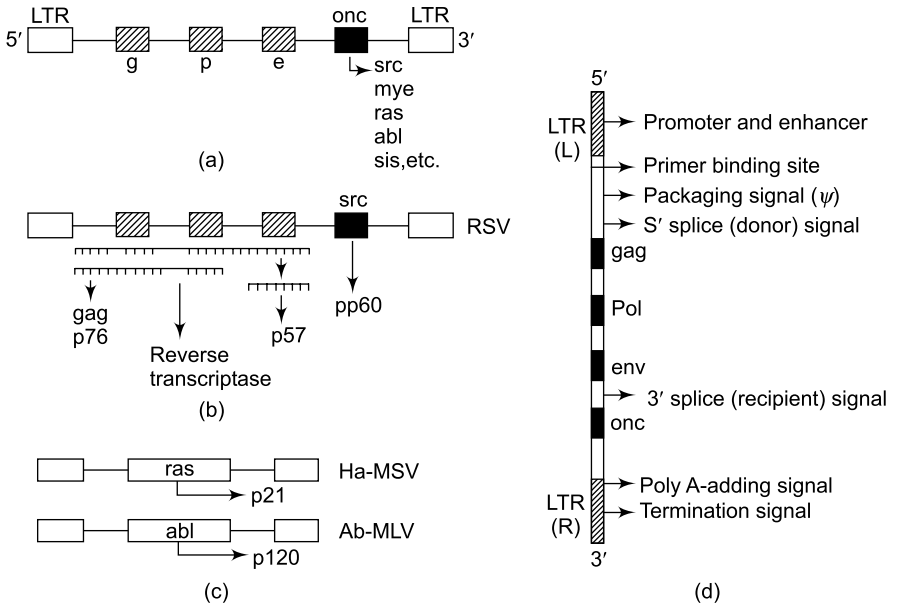


Fig. 2.35 *Highly oncogenic retroviruses carry an oncogene each, the product and function of which differ from one virus to another. Only RSV (Rous Sarcoma Virus) possesses the ideal or complete retroviral genome with a gag, pol, env and onc gene between the two LTRS. The onc gene of RSV, src, encodes a 60 kdal protein. All other oncogenic retroviruses are deficient in one or more of the first three viral genes, but each has its characteristic oncogene. The onc gene is expressed as a fusion protein in some viruses.*

(ii) Vectors The ds DNA of the provirus is the candidate vector material. The LTR and the immediately contiguous sequences, containing splicing, replication and packaging signals, are the only ones required to make infectious particles that integrate the genomic DNA into host DNA. The *gag, pol* and *env* genes are, therefore, redundant in a vector.

Fortunately, all the signals mentioned earlier are sequestered in or near the LTRs and can, therefore, be utilized for mediating the relevant functions in foreign DNA that takes the place of the dispensable genes. Unfortunately, only one oncogenic retrovirus, RSV (Rous Sarcoma Virus of the chicken) is a complete and replication-competent one. Others lack the reverse transcriptase mediated initiating sequences and operate only in the presence of genomes that supply the missing signals and/or factors.

In RSV, the *onc* gene (in this case the *src*) is replaced with the foreign gene. However, this virus only infects birds. Attempts are on to make other retroviruses replication-competent by adding the appropriate signals within the LTRs. One Maloney Murine Leukemia Viral (Mo-MLV) vector has been constructed by inserting such a signal together with a selection marker for *E. coli* cells (the supF, a tRNA suppressor). Although from a cloning point of view, RSV-based vectors appear to be attractive, they are not desirable for gene therapy as the replication competence allows them to spread indiscriminately.

One vector was constructed by deleting all the viral genes and replacing them by a DNA containing a polylinker with sites of several restriction enzymes. This hybrid DNA is latched on to pBR322 sequences containing a replication origin and an Amp^R gene. This is the pMX-1112 single gene vector from the Maloney Murine Leukemia Virus (Mo-MLV) genome. Other vectors derived from the foregoing have incorporated a selection marker between the LTRs or the promoter and initial coding sequences of the *gag* gene. A single gene vector like pMX-1112 is useful for recovering a high titre of recombinant viruses, especially if the foreign gene doubles as a marker for transformation.

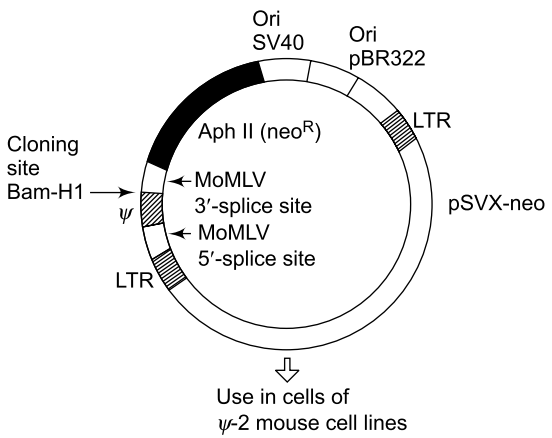


Fig. 2.36 *pSVX-neo: a retroviral shuttle vector. The signals for DNA replication and transcription initiation, termination and polyadenylation are in the appropriate LTR. Immediately adjacent to the left LTR are signals for splicing (5′-splice site), packaging (ψ or psi site) and for binding of the primer for DNA synthesis. A 3′-splice site is present further downstream. A selection marker aph II is included in the vector. This encodes the enzyme aminoglycoside 3′-phosphotransferase II that confers neomycin (or kanamycin) resistance to bacterial and G418 resistance to animal cells respectively. G418 is a 2′-deoxystreptamine derivative of the antibiotic gentamycin. The splice sites are taken from the Mo-MuLV and allow correct splicing of a split inserted gene. The signal allows the hybrid vector to be packaged into retroviral shells.*

However, as most genes of interest are not readily selectable, it is customary to add a marker in the vector DNA. The usual markers are the *tk* (thymidine kinase), *dhfr, qpt, neo* and others described in a later section (selection markers). For genes inserted in place of the distal viral genes, often a suitable promoter is included upstream of the cloning site (polylinker with restriction sites). If the donor gene transcript requires splicing, the cloning site is flanked by the 5′-and 3′-sequences of the viral genome required for splicing. The Mo-MLV derived vector pMX-1122 is such a one, having a marker gene (*neo*) in the *env* locus, preceded by the 3′-splicing signal, which, in turn, has on its upstream side cloning sites for the foreign gene (replacing *gag*). The 5′-splice signal is in its usual position next to the left LTR.

Some retroviruses produce fusion products of the *gag* or *env* gene and the *onc* gene. A foreign gene that is required to be translated efficiently and at high levels can be inserted to take the place of the *onc* gene in vectors made from such viruses.

The only snag is that often such foreign proteins fused to a viral one are not active or have other shortcomings.

Retroviral based vectors can be constructed as shuttles for cloning in both *E. coli* and animal cells such as pSVX-*neo* (Fig. 2.36). In such a case, the *E. coli ori* and usually an SV40 *ori* are included together with a marker that works in both systems. The neomycin (*neo*) marker provides resistance to kanamycin in bacterial cells and to the compound G418 (an analogue of neomycin) in animal cells.

(iii) Packaging cell lines It is not enough to design a suitable recombinant retrovirus-based vector. It should be also possible to recover stocks of mature infectious particles containing the rDNA.

Recombinant vectors can be packaged with the proteins supplied by a coinfected helper virus. However, the viral particles will be a mixture of those with helper and those having recombinant genomes. To avoid such a mixture of virions, it has become customary to use what are known as helper-free viruses, at least for two retroviral vectors (from MLV and from REV, a reticulo-endotheliosis avian virus that also infects dog cells).

A packaging cell has a provirus in its DNA that lacks the packaging signals (e.g., the 351bp psi or ψ region of Mo-MuLV). The proviral genes for shell proteins are, however, synthesized and utilized to package rDNA introduced into the cell (e.g., the γ-2 packaging cell line obtained by transforming a mouse cell line NIH3T3 with any γ-deficient vector, pMOV-γ). The rDNA contains the foreign gene and the signals for packaging. The rDNA containing particles are said to be 'helper free' viruses.

Murine viral vectors are used in BALB/C 3T3 or NIH3T3 cell line cells that have been made into helper cells by infecting them with the packaging-defective viral genomes. Various sophistications have been introduced in this basic design to develop retroviral vector-packaging cell systems with more catholic abilities than those of the initially developed ones.

Two features of retroviruses that are useful for making vectors have not been mentioned earlier. One is the fact that these viruses do not kill the host cell and continuously express products from the provirus genes. A foreign gene placed in a slot in the viral genome, therefore, becomes a source for constitutively expressed proteins, a feature much appreciated by genetic engineers interested in harvesting the product(s).

Some retroviruses have a very restricted host range. These ecotopic strains can be made to extend their host range by including in the genome genes of broad range (amphotropic) strains.

Finally, retroviral vectors do not cause rearrangements in the viral or flanking host DNA, a feature that handicaps animal vectors based on viruses such as SV40.

4. Bovine Papilloma Viruses

The bovine papilloma viruses exist as multicopy episomes in nuclei of animal cells of many different species. They are readily cultivated in cultured rodent cells. The episomal nature of the BPV genome and the ease of its extraction (rescue from the animal host) and reintroduction into *E. coli* (with minor modifications) render

vectors made from BPV valuable for establishing cell lines that overproduce the product of the cloned gone. The strain that has been utilized the most for this purpose is BPV type 1 (BPV-1).

A vector that integrates into the host genome does, of course, express the cloned gene continuously, but the result may not be very predictable or encouraging. It has been already mentioned that since only one vector or a few vectors become integrated, the total volume of the expressed product is quite little. A further disadvantage of such integrating vectors is the uncertainty in the manner of expression of a gene located in a strange genic setting (retroviral and other integrating DNA insert at random sites). The state of organization of the chromatin, or of methylation, or the actual genes or regulatory sequences flanking the inserted DNA may have unpredictable influence on its expression. Such disadvantages are avoided with the use of BPV-1 vector systems.

(i) Biology The genome of BPV-1 is a circular, 7954 bp long DNA. All its genes are on the same strand and are transcribed unidirectionally.

Two major regions may be distinguished on the BPV-1 map (Fig. 2.37). One of them is a transforming segment and the other the structural gene region. The first region constitutes about 69% of the genome; of this only the two terminal regions are required in *cis* for transformation activities. The BPV-1 DNA integrates into the host cell DNA through the mediation of these two subsegments in the 69% segment. The alternative ability to remain as free episomes within the nucleus is mediated by genes in the region between the two transforming subsegments.

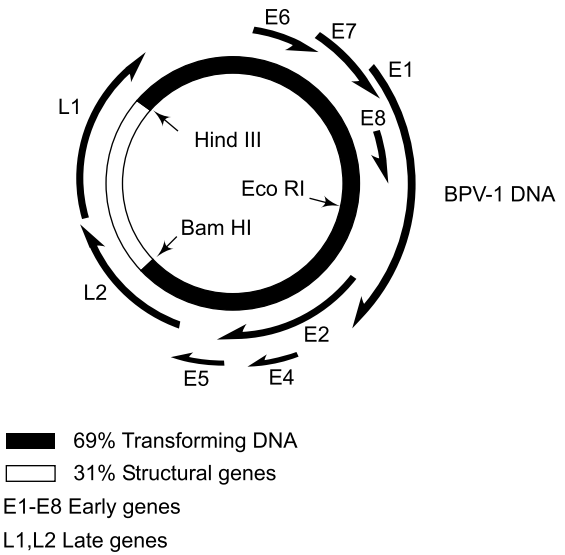


Fig. 2.37 *Schematic representation of the DNA map of bovine papilloma virus type I (BPV-1). The early genes (E1-E8) constitute 69% of the DNA which allows transformation of animal cells. The late genes L1 and L2 are part of the remaining DNA.*

There are two major early (*E1, E2*) and late (*L1, L2*) genes and several smaller early genes (*E3-E9*). The late genes are in consecutive sequences, while the early ones overlap, but have different reading frames. Several transcriptional signals (promoters etc.) are clustered in a 1 kb region between the end of the *E6* and beginning of the *E1* genes. The early mRNAs have a common spliced leader, and are polyadenylated from a common signal. About five mRNAs are seen in transformed cells of warts, while another five are associated with conditions for episome maintenance.

(ii) Vectors The biology of BPV-1 is not as well understood as that of SV40 or even as that of the adenoviruses. What little is known, however, has already proved useful in contributing to the construction of many useful vectors for both transient and long-term expressions of cloned genes.

Most vectors of the aforementioned group are plasmids with BPV-1 sequences. So far, pBR322 and its derivatives have been the base plasmids. Specialized vectors with selectable and amplifiable markers have also been developed.

Special BPV helper cell lines have also been developed which contain the viral genome in an integrated form. As the early functions are supplied by the integrated genes, the vectors can dispense with them and utilize the available space for markers and a longer length of donor DNA.

It was discovered that when the 69% segment of BPV-1 DNA was joined to the intact pBR322 DNA, the vector could no longer transform a cell. The region of the plasmid that erases the transformation function was found to be the same one, the presence of which prevents SV40 DNA synthesis in monkey cells. This region has been referred to as the 'poison sequence' of pBR322. Derivatives of this plasmid lacking this retarding sequence do not suppress transformation by the linked BPV-1 DNA; pML2 and pAT153 are two such vectors. However, this is not a general feature. The same recombinant plasmid fails to cause transformation under different experimental conditions. In fact, such variability, that is unpredictable, in several type of operations still stands in the way of making BPV-1 vectors routinely utilized ones.

The first vectors made from BPV-1 are the ones mentioned earlier. These vectors had limited host ranges and did not possess unequivocal selectable markers; cell transformation (as shown by lack of contact inhibition) was the only criterion that could be used with such vectors. This situation has been improved by designing vectors with dominant selectable markers, such as tk^+, gpt^+ and the neomycin resistance gene (amino-glycoside 3′-phosphotransferase) and the mouse and/or human metallothionin genes.

The metallothionin genes are expressed when triggered off by the excessive presence of metals. The gene product binds the excess metal and protects the cells from its toxicity effects. Vectors made by joining a metallothionin (MT) gene with BPV-1 sequences can be triggered off to express the MT gene by treating the cells with a high dose of cadmium (Cd). Two of the better known BPV-1-MT vectors are pBMTX and pBMT2X (Fig. 2.38). Both contain the BPV-1 DNA, a fragment from pBR322 with the *ori* for *E. coli* cell DNA, and the Amp^R gene, together with a mouse MT (mMT) gene. The second vector has a hMT (human MT) gene between the BPV-1 and mMT sequences. There is a unique cloning site (restriction enzyme

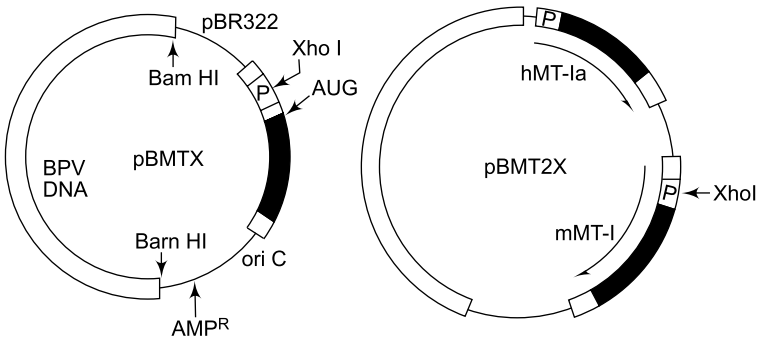


Fig. 2.38 *pBMTX and pBMT2X are two vectors based on BPV-1 DNA. The former carries one copy of MT gene from the mouse. The latter possesses one MT gene from the mouse (m) and one from the human (h) genome. Foreign DNA is cloned at the Xho-I site. An abnormally high level of metal ions in the environment is tolerated due to the overexpression of hMT-Ia.*

Xho-I) just preceding the AUG start codon for the mMT. The second vector has a hMT-Ia gene upstream of the mMT one. While the inserted gene is transcribed from the promoter of the mouse gene, the human gene mainly expresses a high level of resistance to Cd. Cells with such vectors survive and grow with prohibitory doses of the metal in the growth medium. Survivors are easily detected; they contain the inserted gene. The Cd also overexpresses from the promoter of the mMT. As the inserted gene is sandwiched between this promoter and the mMT, the inserted gene is also overexpressed. Other variations of the above BPV-MT vectors are available with subtle changes in their functions. One such series of vectors is used for expressing fused genes. Vectors have also been constructed that add the signal peptide sequences required for proteins to be secreted. In addition, vectors with *neo* markers and a cloning site downstream or upstream of the AUG initiator codon of mMT allow expression of inserted genes from the MT gene or from its own AUG. When the cloning site is within the first intron of the mMT, an inserted split gene without AUG of its own can be spliced and expressed.

A cosmid with BPV-1 DNA together with the neo^R gene promoter of the herpes simplex virus type 1 (HSV-1), tk^+ gene and some pBR322 sequences is also available. Some BVP-MT vectors also have the *neo* gene. They have, in addition, the SV40 t antigen gene and its termination signals. All these vectors are very versatile as well as specific for particular jobs.

Several genes have been cloned and their products synthesized in helper cells at a high level using BPV-1 vectors. Such genes include rat and human growth hormones, human interferon, human β-globin, rat preproinsulin, and the transcriptional activator (TA-1) of the HTLV-1 promoter in the HIV LTR.

The eukaryotic genes cloned so far in BPV-1 vectors have been mostly transcribed, translated, processed and secreted (where required) correctly. Some vectors do not give consistent results; a few become integrated and/or rearranged. The reasons for these anomalies are not clear as yet.

In spite of these lacunae in our understanding of BPV-1 biology, BPV-based vectors will certainly be utilized increasingly in the future. The value of a vector that

can remain unrearranged in a host cell and synthesize the product of an inserted gene indefinitely and continuously cannot be underestimated. Hsiung et al. have collected a gene product at the rate of 1 ng per million cells per day for an indefinite length of time.

5. Herpes Viruses

The herpes viruses consist of several members that have been classified into three groups, namely d, B and herpes virinoe. Herpes simplex viruses types 1 and 2, (HSV-1, HSV-2) and pseudorabies virus are among the varieties. The B-group contains the murine and human cytomegaloviruses. The group is represented by the Epstein-Barr virus (EBV) and Herpes virus.

Several severe disorders (such as encephalitis, meningitis and others) are caused by HSV-1 and HSV-2. These viruses have also been implicated in certain cancers.

HSV-1, EBV, HSV-2 and VZV (Varicella-Zoster Virus) are the ones that have been studied the most. VZV has been used mostly for a human vaccine.

The herpes viruses have large genomes, and were expected to be useful as vector material for animal genes. However, this expectation has not been fully justified as yet. Their potentials, at present, appear to be as candidates for vaccine making or for unravelling the biology of the herpes viruses themselves.

The HSV-1 has been inspected for a vector supplying potential. Both helper-dependent and helper-independent infectious vectors have been constructed to date; they do not appear attractive as carriers of eukaryotic genes.

Some minireplicons have been made from EBV, and hybrids with sequences from pBR322, SV40, transcriptional signals from an EBV gene and selection markers for cloning in bacterial and animal cells have been constructed and utilized for specific research goals.

6. Vaccinia Viruses

The vaccinia viruses have been in use for a long time as vaccines against smallpox. Their large size, and long history of not communicating the disease via vaccines, have prompted their examination for the making of vectors for large eukaryotic genes to be expressed in animal cells.

(i) Biology The human pox virus has been studied the most among the members of this group, and hence was the obvious choice for investigating its use as a eukaryotic vector.

The virus is quite a large (200–300 nm) box-shaped particle with a linear ds DNA genome that is about 186.5 kb in length. The virus undergoes all steps in its development in the host cell cytoplasm. The mature virus contains, besides the DNA molecule, several enzymes required to transcribe and process the viral genes. Thus, the viral RNA Pol, the methylating enzymes for capping the mRNA (5′-end) and adding a poly A tail at the 3′-end are part of the equipment carried in the viral core.

Infection by a vaccinia virus brings to a halt all DNA and protein synthesis in a cell. This is fortuitous, as foreign proteins, expressed via the viral vector, can be isolated rather easily. These viruses infect a very wide spectrum of animal species.

As the cells are killed by infection, only transient expression may be observed using vaccinia vectors. A foreign gene, inserted in a functional location on the viral DNA, is expressed and processed correctly both after transcription and translation. The hybrid DNA becomes packaged correctly and infectious vectors are produced.

(ii) Vectors To construct recombinant vaccinia vectors, the donor is first cloned with flanking vaccinia DNA, in a plasmid vector. This is transfected into cells infected with a vaccinia virus. Homologous recombination between the flanking avian sequences in the introduced plasmid and the resident homologous loci in the cell transfer the inserted gene into the vaccinia DNA (Fig. 2.39). The latter becomes packaged correctly; selection is made possible by including a suitable marker in *cis* with the foreign gene. This technique of inserting a gene into the vaccinia genome was first reported separately by Machett et al. and Panicali and Paoletti in 1982.

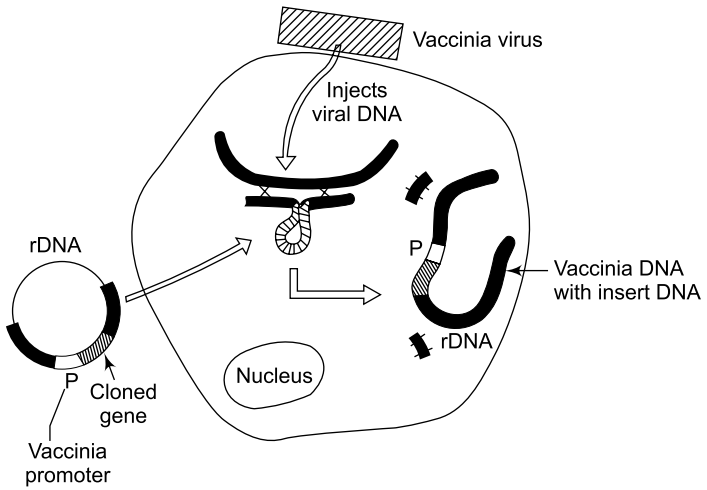


Fig 2.39 *Construction of a recombinant vaccinia virus. A vector carrying the foreign DNA sandwiched between stretches of vaccinia virus DNA is introduced in a cell infected with the virus. Homologous pairing and crossing over between the vaccinia DNA in vector and viral chromosomes insert the foreign DNA into the latter, which is packaged normally into viral particles.*

The DNA may be inserted in one of the several essential sites. These have not yet been characterized fully. Most genes have been inserted within the vaccinia *tk* locus. Successful insertion is reflected in the inability of the destroyed tk^+ gene (now tk^-) to grow in a HAT medium.

The plasmid vector containing the foreign gene is provided with transcriptional signals of vaccinia genes. Some vectors also contain the *lacZ* construct with promoters from a vaccinia gene. Recombinant cells can be visualized by the β-galactosidase-XGAL screening system.

One of the most successful vaccinia-based vector systems uses coinfection of a cell with two recombinant vectors one with the inserted gene and the other with the RNA polymerase gene of phage T7, which has been discovered to be a very efficient initiator of transcription. The target gene is flanked by the signals for the T7 RNA Pol.

The target DNA can be as long as 25 kb. Hela cells (spinner cultures preferably), chick embryo fibroblasts and TK strains (L) of mouse are equally good for growing vaccinia viruses. Infectious vectors can thus be produced by infecting cultured animal cells with recombinant viruses.

Recombinant viruses may be detected by a variety of techniques that include 'immuno dot blot' and 'DNA dot blot hybridizations' and selection of tk^+ recombinants.

Several eukaryotic genes have been studied by expressing them via vaccinia vectors. The list includes the gene for the circumsporozoite surface antigen (a surface protein of one of the various developmental stages of the malarial parasite *Plasmodium knowlesii*), the chloramphenicol acetyl transferase (CAT) gene of bacteria and the Hepatitis B surface antigen (HbsAg). Vectors with genes for surface antigens of pathogenic species are being designed to be used as vaccines instead of the conventional attenuated, but whole, pathogens.

7. Baculoviruses

Baculoviruses are viruses that infect insects. They have come into prominence for two reasons: (i) they provide a handle for controlling insect pests. (Eradicating insect pests with biological agents is cheaper, easier and without the attending pollution hazards of the currently used chemical insecticides.) (ii) They have proved to be efficient expression systems for the overproduction of animal proteins. The baculovirus-insect system has been utilized to synthesize several human and other vertebrate proteins that include fibroblast interferon, IL-2 and *c-myc.* The proteins have been processed correctly within the insect cells, in contrast to their lack of processing or indifferent processing in *E. coli* or yeasts. Extracellular secretion, signal peptide cleavage, and other post-translational modifications were correctly carried out in each case.

The baculovirus AcNPV (Autographa californica Nuclear Polyhedrosis Virus) has been the target and tool for many studies of gene expression. These viruses are easily grown in cultured insect cells. They are helper-independent.

The polyhedron gene of the baculovirus is preceded by a very powerful promoter, which has been requisitioned in cloning studies to drive expression of foreign genes. Apparently, the protein processing mechanism is identical in the insect and mammals; the foreign proteins are secreted and are functional.

Smith et al. (1985) used the insect *Spodoptera frugiperda* and a vector (pAC101 = pUC8 + AcNPV DNA-containing the promoter and gene for the polyhedron protein) that has a Bam HI site just downstream of the initiation nucleotide for RNA synthesis. An IL-2 gene was cleaved with Bam H1 upstream of the ATG (initiation codon), and spliced to the cleaved vector (Fig. 2.40). Those constructs that had the IL-2 codons in the same reading frame as the polyhedron gene, expressed the IL-2 product efficiently.

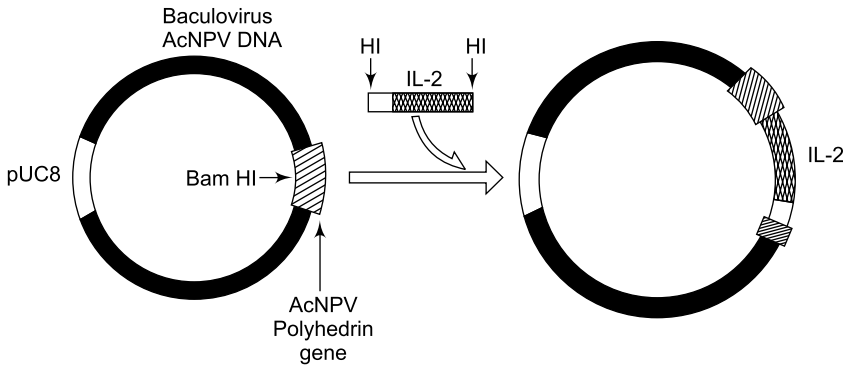


Fig. 2.40 *A baculovirus vector (derived from pAC101) carries the gene for the large nuclear protein polyhedron of the virus Autographa californica Nuclear Polyhedrosis Virus. The AcNPV gene possesses a very powerful promoter. The foreign DNA inserted within the polyhedron gene is expressed efficiently.*

Another protein that is synthesized with advantage in insect cells is the *c-myc* oncogene. Its overproduction in mammalian cloning cells proves toxic for the latter, but not for insect cells. The second and third exons with the intervening intron of human c-*myc* were joined to the promoter of the viral polyhedron gene as in the case cited earlier. The c-*myc* protein was secreted to a high level in a correctly processed manner. This protein is an important regulatory one, aberrations in which, and anomalies of regulation of which leads to malignant transformations. The animal cell systems have not been able to supply sufficient amounts of this protein for research. The insect baculovirus system is expected to fill this need.

2.4.5 Special Vectors

A plasmid, phage or viral vector with the bare minimum features of a vector has limited potentialities. Some are good only in bacterial cells. Still others do not express functional proteins or processed mRNAs. Some need to be amplified and/or tailored further in a bacterial cell for the ease of operation in these cells but expressed in an eukaryotic one. A naturally occurring plasmid or viral vector is unable to exist in such widely different hosts.

A variety of vectors have, therefore, been designed, which incorporate useful sequences from diverse sources. This is where we appreciate once more not only the unity in basic molecular processes throughout the living system but also the near-magical ingenuities and expertise of the genetic engineer who snips off a promoter from a virus, positions it next to an oncogene retrieved from a tumour tissue and adds a gene from an enteric bacterium that allows one to track the presence of this chimaeric DNA in diverse backgrounds.

Some special vectors can be used interchangeably in different species; these are the 'shuttle vectors'. Other vectors are provided with dominant selectable markers. Still others are expression vectors containing efficient transcription and/or translational signals that may be utilized for expressing the foreign gene. Sometimes, it is necessary to integrate a gene into a heterologous DNA. Special integrating

vectors that replace a section of the host DNA with the foreign one can be designed. Then again, the objective of using a foreign cloned gene may be to overproduce the gene product in a convenient cellular factory. The capabilities of the host cell for doing so may be ignored and gene inserted in the vector that is naturally amplified under specific selection pressures; with the increase in the number of copies of the amplifiable gene, the foreign gene linked to it is automatically amplified. Add to this an efficient transcription signal and you have a vector that performs better than any gene in its natural setting.

Making vectors to order is one of the foremost occupations (obsessions ?) of DNA engineers. It is, therefore, not possible to even hint at the plethora of unique vectors that are being reported continuously.

Nevertheless, certain basic model vectors have been developed, which form the basis (both physically and conceptually) for designing of other novel ones. Some of the prototypes will be enumerated here with a few examples of each type. The 'References' may be consulted for the maps and other details of the cited examples.

1. Shuttle Vectors

Shuttle vectors are plasmids that contain replication origin sequences for two different host species. The hosts belong to divergent groups such as bacteria and yeast, monkey and *E. coli,* and *E. coli* and human beings.

The shuttle vector usually also carries two selectable marker genes, one each for the two hosts. In this way not only can the plasmid replicate in the two hosts but also may be readily identified by the phenotypes imparted in each cellular background by the corresponding marker gene.

The first shuttle vector that became popular and introduced the idea of such interchangeable vectors is the one that could be used in both *E. coli* and yeast (*Saccharomyces cerevisiae*). It contains the *ori* of both species and a number of markers each for the two hosts. A suitable cloning site is of course provided. The cloned gene is first amplified in *E. coli* (Fig. 2.41) and subsequently expressed in yeast cells.

Shuttle vectors have been constructed that move between mammalian and bacterial (*E. coli*) cells. One such shuttle, designed by Essigman's group at MIT (Fig. 2.42) carries a known DNA lesion (alkylated guanidine), the replication origins for *E. coli* plasmids and SV40 and a neomycin resistance gene that confers kanamycin resistance in the bacterial cells and G418 resistance in the mammalian cells. The purpose of the study was to monitor the ways in which the lesion would become fixed mutations. The shuttle vector was introduced in CHO (see Section 2.8.7) cells where it became integrated in a host chromosome. The transformed cells are recognized by their ability to grow in spite of the neomycin in the medium. These cells are then fused with COS cells, an operation that leads to the induction of the vector with the insert DNA as circular molecules. These were introduced into *E. coli* where they became amplified. This time kanamycin was utilized to identify the neo^R containing cells. The amplified DNA was finally examined for the nature of the mutation in the target DNA. The vector pSV2-β shuttles between *E. coli* and monkey cells (Fig. 2.43).

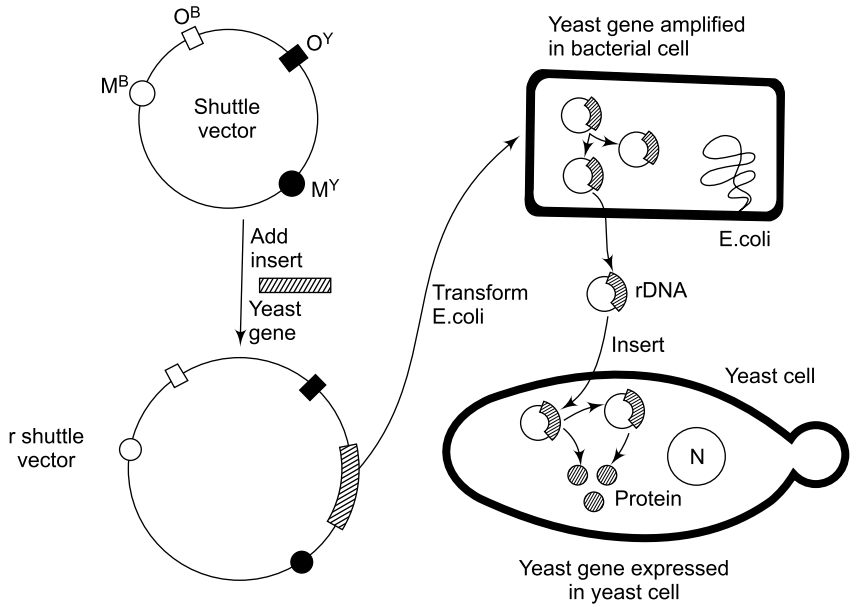


Fig. 2.41 *A shuttle vector that replicates in E. coli and yeast. The foreign gene is amplified in E. coli and expressed in yeast cells. M^B,M^Y are markers and O^B and O^Y are ori regions for bacterial and yeast cells respectively.*

Others have constructed shuttle vectors with the *ori* of EBV and that of pBR322. These vectors also carry selectable markers for *E. coli* (Amp^R) and mammalian cells (hygromycin phosphotransferase gene), and the tk^+ gene from HSV. The shuttles were utilized to study the types of mutations that can be induced in the tk^+ gene in cultured human lymphoblastoid cells. The EBNA-1 (Epstein-Barr virus nuclear antigen) gene was included in a helper cell or on the shuttle, as its product is required as a transcription factor to activate the enhancer in the EBV *ori.* (Fig. 2.44).

Several shuttle vectors have been constructed based on BPV sequences on an *E. coli* plasmid, and can be used in cultured animal cells for expression and rescued in *E. coli* cells for amplification, modification and anaylsis. Shuttle vectors are also available that move between *E. coli* and Drosophila (Fig. 2.45).

Recently, shuttle vectors that can move with ease between gram-negative and gram-positive bacteria have become available. The most successful ones, to date, are those that can be propagated in *E. coli* and *Bacillus subtilis.* The replicons are from plasmids in *Staphylococcus aureus* (with chloramphenicol resistance gene) and pBR322. This bifunctional vector is amplifiable in a particular *E. coli* mutant.

2. Expression Vectors

Expression vectors contain expression signals that best transcribe/translate the foreign gene in a heterologous system. Sometimes, these signals are promoters and terminators of *E. coli* or phage genes that are added at appropriate sites on either

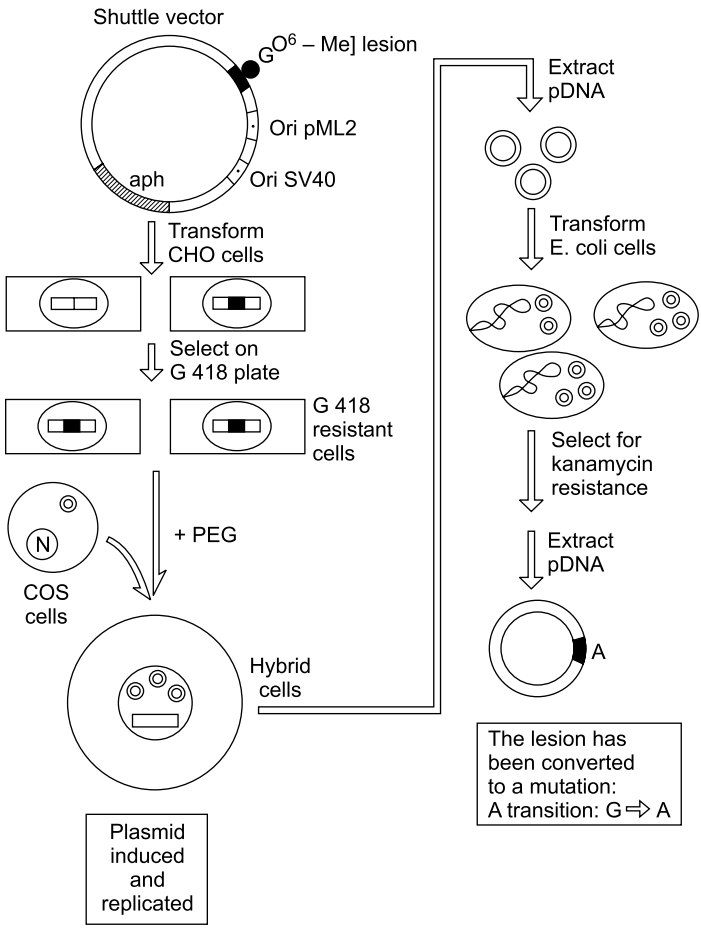


Fig. 2.42 *A shuttle vector designed by Essigman et al. that can replicate in animal and E. coli cells. This particular vector with a neo^R marker (gene aphII) was used to clone a DNA fragment with a lesion (O^6–methyl guanidine) to find out if the lesion is converted to a mutation (in this case G → A) when placed in cells possessing or lacking repair functions. The hybrid DNA was first introduced into Chinese hamster ovary (CHO) cells, and the transformed cells selected by virtue of their resistance to neomycin. The transformed cells were fused with COS (monkey) cells, an event which induced the integrated hybrid vector. The latter circularized and were replicated in E. coli cells (which became kanamycin-resistance). Some vector DNA extracted from E. coli cells were found to possess the nucleotide adenine (A) in place of the methylated guanidine adduct (the lesion). This is only one of the many ingenious ways in which shuttle vectors may be utilized.*

side of the gene to be expressed. A eukaryotic gene from an animal can be expressed in *E. coli* with promoters and terminators from *E. coli* genes. The earliest expression vectors used this strategy (Fig. 2.46).

It was soon discovered that many phage or viral promoters were more efficient as transcription initiators than the native *E. coli* signals. The promoter of phage T7 for instance, is a very powerful one; so, when the aim of cloning a gene is to overproduce its protein, an expression system is created by positioning a T7 promoter upstream of

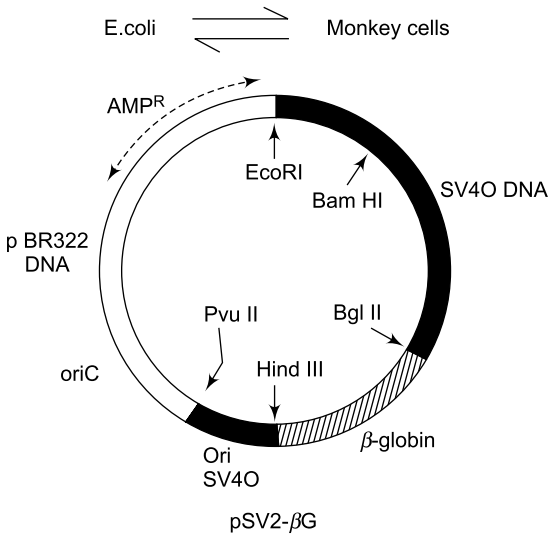


Fig. 2.43 *pSV2-β G. An E. colii-monkey cell shuttle vector. The vector was derived by including a Bgl II—Hin dIII fragment containing the globin (β) gene. The bacterial and SV40 ori sequences allow replication of the vector in E. coli and monkey cells respectively.*

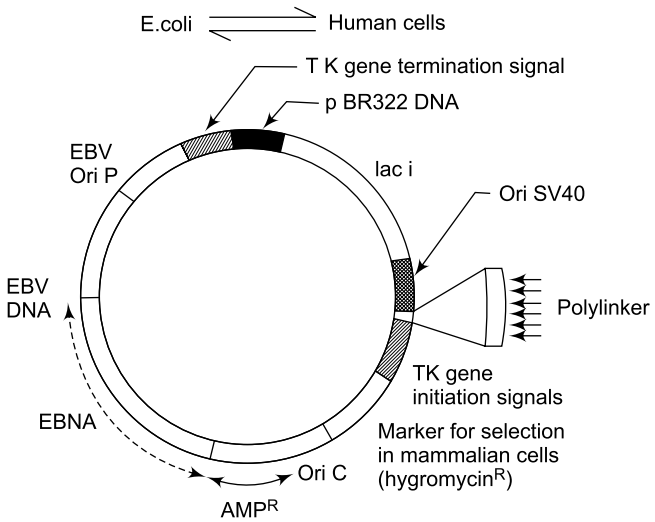


Fig. 2.44 *An E. coli-human cell shuttle vector with sequences from the EBV (Epstein-Barr virus) and SV40 DNA. Ori P is from EBV DNA; it requires the virus-encoded EBNA (E-B nuclear antigen) protein for replication of the EBV DNA. A polylinker with a variety of cloning sites is flanked by the initiation and termination signals of the TK gene. Sections of pBR322 sandwich the ori C, the Amp^R and the lacI (repressor) gene. A marker gene (hygromycin resistance) is included for selection in mammalian cells. Ori C is a BamHI-C segment of EBV DNA and is composed of two separate motifs about 1 kb apart, each with repeats of a 30 bp sequence. Vectors from EBV are found to be unrearranged in the cloning cell, unlike SV40-based vectors.*

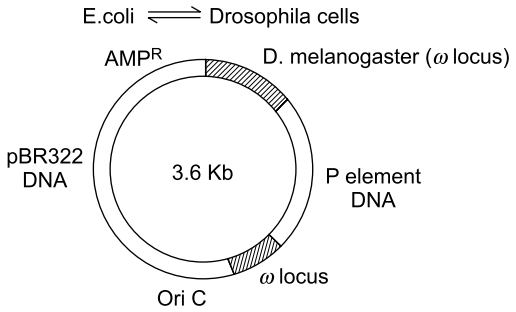


Fig. 2.45 *An E. coli-Drosophila cell shuttle vector belonging to what is referred to as the Carnegie series. A Drosophila mobile P element, with deletion of a segment that promotes autonomous transposition, was introduced into a pBR322 (pUC8). The cloning site is within the P element sequences. The latter possess typical 31 bp terminal repeats, (see Rubin and Spradling, 1992, 1993.)*

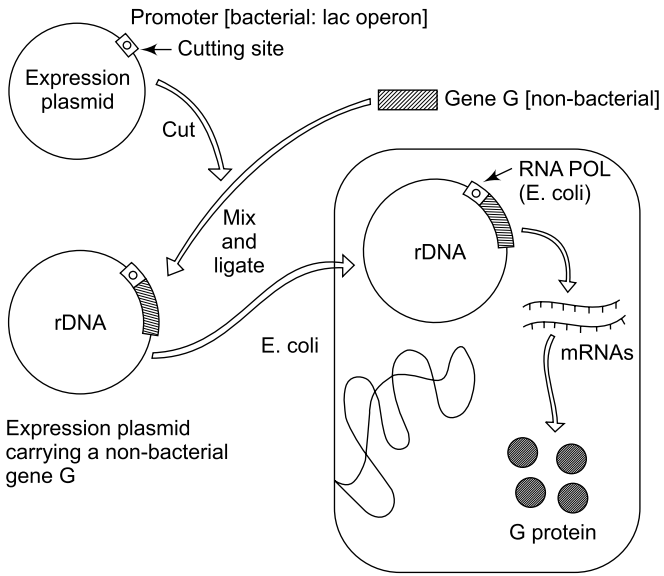


Fig. 2.46 *An expression vector with transcription signals for the bacterial system may be utilized to express eukaryotic genes in bacterial cells.*

the gene on a vector together with a gene for the T7 polymerase, often on a separate vector. The RNA Pol gene is preceded by an *E. coli lac* UV5 promoter which, in turn, is under the regulatory control of the *lacI* or repressor protein. Such a vector may be used with an *E. coli* cell with the addition of the lactose analogue, IPTG, to the glucose-deficient medium. The *lacI* product is deflected from repressing the UV5 promoter. This allows the T7 RNA Pol gene to be expressed. The phage RNA Pol then initiates high levels of transcription from the T7 specific promoter linked to the gene of interest. The latter is overexpressed. Removing the inducer stops the synthesis of the RNA Pol and consequently of the gene of interest (Fig. 2.47).

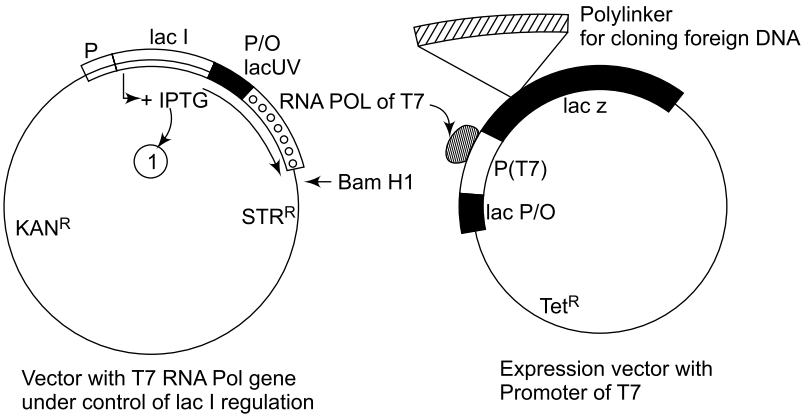


Fig. 2.47 *A two-vector expression system. An expression vector (pNP106 or pNP109) carries the phage T7 late gene promoter sandwiched between the E. coli lac promoter/operator (p/o) and the lacZ sequences. A foreign gene, cloned in a polylinker within the lacZ gene, is transcribed from the T7 promoter, which is regulated by the lacP/O. The second vector (pNP148) is a regulatory one. It carries the T7 RNA polymerase (Pol) gene controlled by the lacUV p/o and the product of the lacI gene. The inducer IPTG triggers the transcription of lacI the product of which represses the expression of the RNA Pol gene. In the absence of the repressor protein, the expressed polymerase binds itself to the T7 promoter and allows expression of the cloned gene.*

Similarly, a gene to be expressed in a mammalian cell may be placed under the control of a promoter from SV40. Usually the promoter for the early genes (T, t) is used for the purpose. When used with a COS cell (a cell line from mouse, containing a replication-deficient SV40 DNA) the gene is expressed efficiently from the SV40 promoter.

As mentioned in a previous section, the promoter for the metallothionein gene (man, mouse) is often linked to the gene to be expressed. The MT gene can be triggered into transcription initiation by a heavy dose of a metal such as Cd. In this case, that expression can be regulated. A ts mutant of a regulator gene (such as *Cl*) may likewise be utilized to control the expression of a foreign gene preceded by the P_L of phage λ.

A useful expression vector is represented by the pUC series. These vectors carry a part of the *lacZ* gene (the N-terminal with around 177 codons) and its upstream regulatory sequences (Vieira and Messing, 1982). The vector has to be used in an *E. coli* strain (JM 83) which possesses *lacZ.*

The mutations in pUC and *E. coli* chromosome complement each other and produce the functional enzyme which turns XGAL into the blue dye. A polylinker has been placed within the *lacZ* region; this additional DNA does not appear to affect expression of a functional protein. When an insert DNA is cloned at one of the sites in the linker, the *lacZ* gene is inactivated.

3. Dominant Selectable Vectors

Several genes are mentioned in a later section on methods of selection of a cloned gene. These include the genes *tk* (thymidine kniase), *dhfr* (dihydrofolate reductase),

gpt (xanthine guanine phosphoribosyl transferase), and *aph* or *neo*^{R} (aminoglycoside phosphotransferase). When such genes are included in a vector, under appropriate conditions, the presence of clones containing these tell-tale markers can be ascertained. As these selectable markers are accompanied by the gene of interest, the presence of the latter is also confirmed.

Some selectable vectors have to be used in particular cells; others are applicable in a wide variety of species.

The *aph*-1 gene occurs on Tn601, and is expressed in *E. coli* to inactivate antibiotics of the aminoglycoside kind (neomycin, kanamycin). Mammalian cells do not have the *aph* gene. The latter can be placed in a vector with the promoter of the HSV *tk*$^+$ gene and introduced into a mammalian cell. The *tk* gene is expressed correctly in such cells. If the *aph* transformed cells are grown on neomycin supplemented media, the expressed *aph* product will make the mammalian cells resistant to the antibiotic. The cells carrying the gene of choice thus can be selected by the use of such a selectable vector.

Similarly, a selectable vector carrying a CAT gene with a eukaryotic promoter can be detected in a mammalian cell which does not express it. The CAT-containing vector confers chloramphenicol resistance to the cells harbouring it (Fig. 2.48).

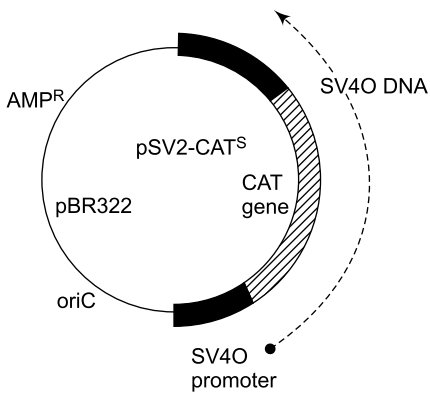


Fig. 2.48 *pSV2-CAT^s is a dominant selectable vector. A CAT (chloramphenicol acetyl transferase) gene preceded by an SV40 promoter has been inserted into pBR322. This vector can be used in an eukaryotic cell, to which it confers resistance to chloramphenicol. In bacterial cells the CAT gene cannot be transcribed due to the absence of a bacterial promoter. Bacterial cells with this vector are therefore sensitive to the drug.*

The commonest direct selection vectors for use in *E. coli* cells are the different drug-resistant marker-carrying ones. The markers are usually on transposons. The Amp^R gene is on Tn3. This transposon can be inserted within the colicin immunity gene of Col E1 DNA. When cells with such vectors are treated with mitomycin C that provokes the synthesis of colicin, lack of colicin resistance (due to the insertional mutation) forces the cells to grow very poorly into very small colonies. Hence, minute colonies under the selection pressure of mitomycin C indicate them to be recombinant ones. It may be noted that pBR322 has Amp^R and Tet^R genes. A gene cloned within one of them inactivates it and cells treated with the drug do not survive.

4. Amplifiable Vectors

It has been noticed that when cultured cells are subjected to an excessive dose of certain toxic drugs, some clones survive the treatment. On investigating these drug-resistant clones, it was discovered that their increased resistance was directly related to an increase in the number of copies of the gene that confers resistance

to that particular drug. Advantage has been taken of this knowledge in designing amplifiable vectors into which the foreign gene is inserted. On treatment with increasing doses of the relevant drug, the copies of the closely linked inserted genes increase simultaneously with those of the drug-resistant gene. Such a vector induces overproduction of the desired gene product by merely adding the drug concerned to the culture.

The first amplifiable gene was the *dhfr* which becomes amplified on treatment with methotrexate (an inhibitor of the *dhfr* product). Several other genes have turned out to be amplifiable in a similar manner. They include metallothionein-I (MT–I), glutamine synthetase (GS), aspartate transcarbamylase and adenosine deaminase.

The amplifiable vectors contain pBR322 or other plasmid sequences with a replication origin and a selectable marker for *E. coli,* another marker for mammalian cells, the amplifiable gene and a replaceable segment for the gene to be inserted.

The *dhfr*-inserted vector is usually used in a CHO mutant which is $dhfr^-$. A second strategy is to allow the high dose of the toxic drug to kill off all cells except the ones with the amplified drug-resistant gene.

Amplifiable vectors with the MT gene also carry a selection marker (usually the neo^R gene). Other genes have been reported recently that are usable as amplifiable vectors.

McConlogue has developed an amplifiable vector based on the gene for ornithine decarboxylase (ODC). DFMO (difluoromethylornithine) inhibits the action of ODC in decarboxylating ornithine to putrescine. Treatment of cells having an ODC vector with DFMO amplifies the ODC gene almost 1000 fold. The inserted gene is likewise amplified enormously. The vector has been designed to be used in ODC^- CHO cells.

5. Integrating Vectors

An integrating vector is a plasmid with an inserted gene and sequences from the DNA flanking the site where the gene is to be inserted. When such a vector is introduced into the relevant cell, recombination between the homologous sequences of the vector and host cell result in the integration of the desired gene into the chromosomal site.

The first integrating vectors were made to introduce genes into yeast chromosomes. These were the YIP (yeast integrating plasmid) (Fig. 2.49) vectors, which have a pBR322 base and a selectable marker ($URA3^+$) for yeast cells. Scherer and Davis (1979) used this method to transfer *his* 3 (a deletion mutant) into an yeast chromosome (Fig. 2.50).

Some yeast integrating vectors possess the *ARS* (autonomous replication sequences) of yeast DNA that act as replication origins. The *ARS* makes the YIP replicable. The increased number of vector copies somehow enhances the efficiency of transformation at least 10^2–10^3 times.

6. Single-Stranded Plasmid Vectors

The first ss DNA vectors in use have been those derived from phage M13 DNA. The M13 vectors cannot accommodate large pieces of inserts stably. A series of

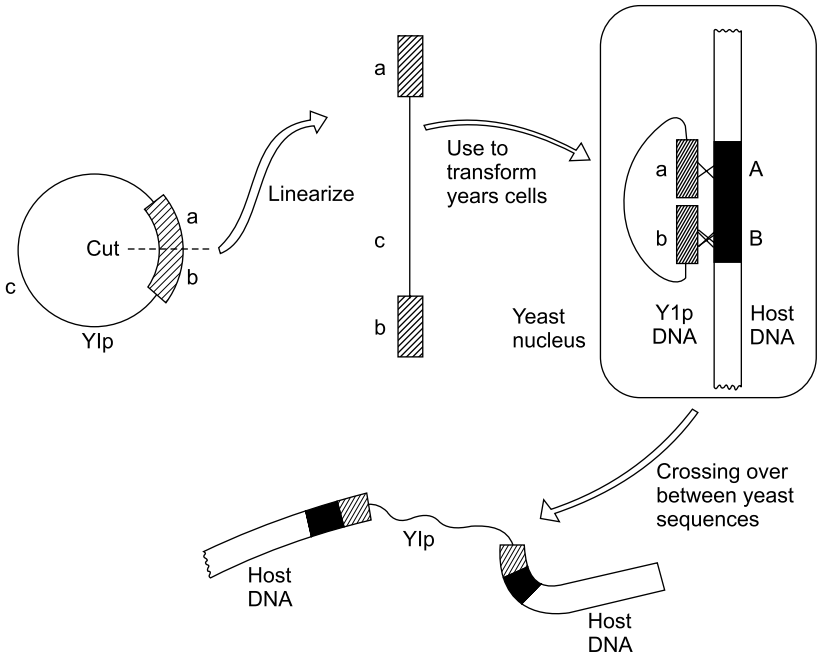


Fig. 2.49 *The YIP (yeast integrative plasmid) vector integrates into yeast DNA as a result of recombination between homologous yeast DNA sequences in the vector and chromosome.*

vectors have been developed, which possess the M13 trait of becoming sequestered as single-stranded DNA in mature phages but are more versatile in terms of length of insert and other features. These vectors belong to a series known as the EMBL vectors.

pUC is a ds DNA plasmid vector containing a polylinker with cutting sites for different restriction enzymes. This multiple site polylinker is inserted within a *lacZ* gene that is already spliced into the plasmid. A vector pD4 has been constructed by Dente et al. from pUC by inserting a 1300 bp DNA from phage *fl* into the *lacZ* gene. The *fl* segment contains the sequences required for mediating the conversion of the ds RF of the phage into the ss DNA containing mature phages. The segment has already been cloned in a pBR322. When the EMBL vectors contain, in addition to the genes on the pUC the *ori* of *fl,* the latter allows the RF to become ss plasmids on super infection with *fl* phages. Only one of the two strands (+ve or –ve) becomes encapsulated; which one depends on the orientation of the insert at the cloning site. Any plasmid can be made an EMBL type vector by incorporating the *fl ori* sequences. Dente et al. have constructed several yeast shuttle ss vectors (the EMBL family) in this manner. Yeast DNA sequences, yeast replication origin and others that are selectable markers have been introduced into a unique site in the polylinker of one of the EMBL vectors. The yeast sequences allow the vector to become integrated in yeast chromosomal DNA by homologous recombination. The auxotrophic markers (e.g., *TRP1, LEU2, URA3*) help in selecting the desired clone by complementation.

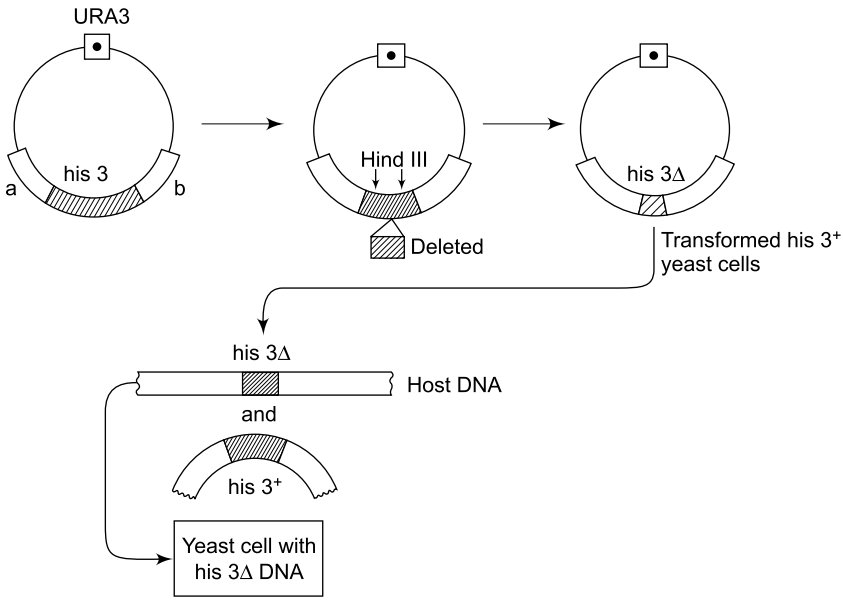


Fig. 2.50 *The YIP vector can be used to introduce a mutation into an yeast chromosome. In this illustration a his gene mutation (his 3) is carried on a YIP. Homologous recombination splices the entire vector into chromosome 15 of yeast. The two genes (his$^+$ and his$^-$) do not remain together for long. After 15–20 cell generations, one of the two alleles is lost. URA3 is a marker selectable in yeast. The desired recombinant may be selected for uracil auxotrophy by selection for resistance to 5-fluoro-orotic acid. URA3 represents orotidine phosphate decarboxylase.*

7. Artificial Minichromosomes

Artificial chromosomes are plasmid vectors that contain DNA sequences that allow them to be segregated into daughter cells like chromosomes of eukaryotes. These sequences are *ARS* and *CEN* taken from yeast DNA and *Tr* from the protozoon Tetrahymena (Fig. 2.51).

(i) ARS These are autonomous replicating sequences, that function as replication origins in the plasmid. It is not certain whether they have the same function in their original settings. The *ARS* were first cloned from yeast chromosomes and from the 2 μ yeast plasmid.

Chromosomes of other species (e.g., slime mold, maize, fruit fly, nematode, yeast) contain a common consensus sequence

A A A C / T A T A A A

which is the ARS motif.

(ii) CEN These are sequences taken from the centromere regions of yeast chromosomes numbers 3, 4 or 11. The *CEN* contains an 80–90 bp AT-rich stretch, flanked by bases that may be homologous in *CENs* from different chromosomes.

Addition of the *CEN* sequences associates the vectors with the mitotic apparatus, and the copies of the plasmid are transmitted to daughter cells. A vector without a

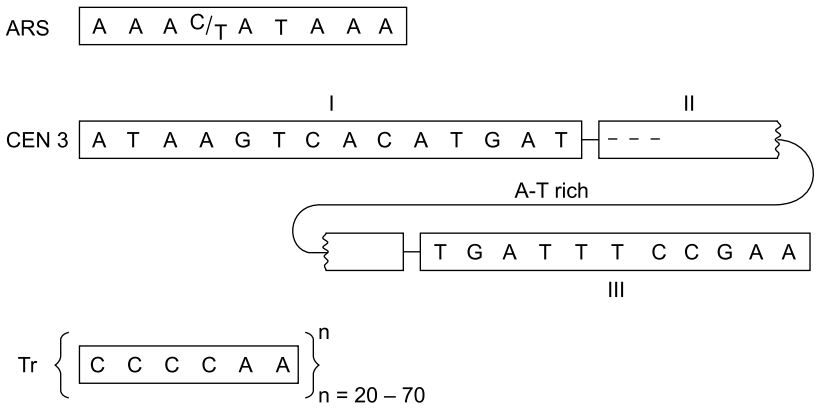


Fig. 2.51 *The ARS, CEN and Tr sequences. ARS (autonomous replicating sequences) are required for DNA replication. CEN is an A-T-rich centromeric sequence involved in an equitable distribution of daughter chromosomes to progeny cells. Tr represents tandem arrays of a specific C-A-rich motif in yeast and a C-T/A motif in other primitive eukaryotic telomeres (terminal regions of chromosomes).*

CEN is not guaranteed for equal segregation to the new cells. This is a danger faced by most plasmid vectors that are used to transfrom yeast cells. In fact, the plasmid is often lost completely in sectors of the population.

Segregation of *CEN*-containing vectors during meiosis depends on a sequence that is longer (6–10 kb) than that which is sufficient for segregation at mitosis.

The *ARS*- and *CEN*-containing pBR322 vectors are also provided with a marker selectable in yeast. One marker utilized popularly for use in yeast is LEU2 (for the enzyme β-isopropyl malate dehydrogenase) required for the biosynthesis of leucine.

(iii) Tr Ends of DNA molecules (telomeres) in yeast chromosomes, and in those of the ribosomal DNA of the ciliate Tetrahymena, have been found to be hairpin loops with a characteristic composition.

The ends of the rDNA of Tetrahymena contain 20 to 70 repeats of a sequence

```
C C C C A A
| | | | | |
G G G G T T
```

with a closed loop where the two strands end (Fig. 2.52).

Two telomeric regions cannot join. This is possibly the reason why eukaryotic chromosomes are able to remain linear. When these telomeric (*Tr*) sequences are added to an artificial chromosome, the vector becomes linear, and stays so. In fact to make the circular DNA linear, two *Tr* sequences are added, one from yeast, and the other from Tetrahymena, in a head-to-head orientation. When this vector is introduced into a cell, the plasmid breaks between the two *Trs,* and the vector becomes linear. The linear vector acts effectively like a minieukaryotic chromosome (Fig. 2.53). The advantages of such chromosomal mimics for cloning eukaryotic genes are obvious.

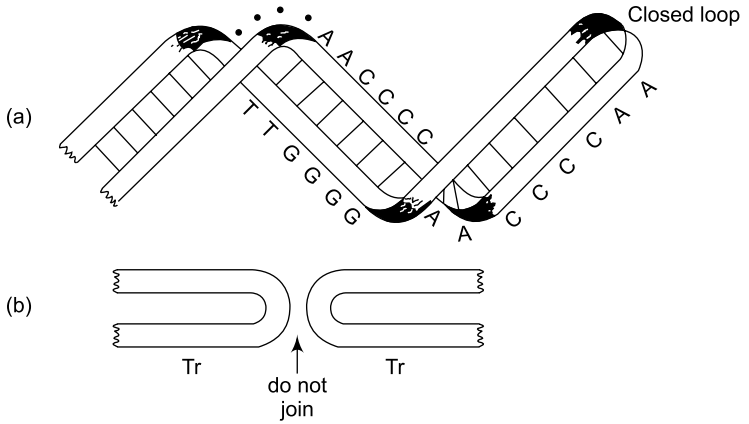


Fig. 2.52 *The Tr sequences of Tetrahymena ribosomal gene DNA (rDNA) are about 700 bp long.*

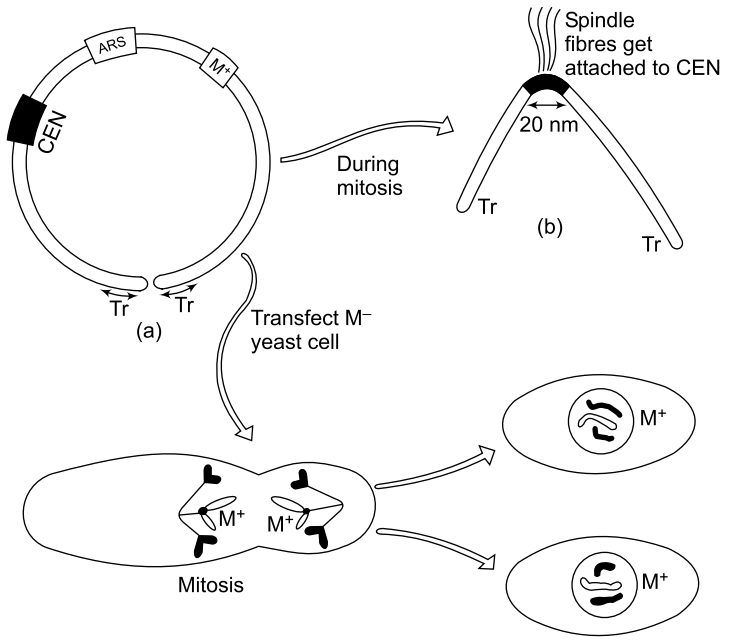


Fig. 2.53 *A minichromosome with ARS, CEN, and Tr becomes linear inside a cell and is dispensed to daughter cells in a stable manner, especially if the length of the artificial chromosome is greater than a certain minimum length.*

8. Broad Host Range Vectors

Broad host range vectors are essentially shuttle vectors that can be maintained in a wide spectrum of species. Plasmid vectors developed to date are almost exclusively based on *E. coli* plasmids. These vectors can be used only in a few related gram negative bacteria such as *Salmonella, Serratia* and *Proteus.* They are not usable in *Pseudomonas spp.,* also a gram negative bacteria.

Plasmids from incompatibility groups P and Q are more catholic in their tastes. They exist equally well in species of *Pseudomonas, Azotobacter, Rhizobium* and *Neisseria,* as well as in *Agrobacterium tumefasciens*, and *Methylophilus methylotropus.* These are all gram negative bacteria that are exploited for a number of economically profitable goals.

Most P group plasmids are very large, e.g. RP4 is 56 kb and have three drug-resistant genes, and the ones present are sequestered in the drug-marker gene regions. The replication and maintenance functions usually occur in three widely separated locations.

The Q group plasmids are much smaller in size, e.g. RSF1010 is 8.9 kb (Fig. 2.54) and are of the multicopy type. They do not autotransfer themselves, as P group plasmids do, but may be mobilized for transfer with the aid of gene products provided from another plasmid (e.g., RK2, a P group member).

Not many useful broad range vectors have been derived as yet for gram positive bacteria. *Bacillus subtilis* is one such species. Its own plasmids have not been very useful. However, certain small plasmids of *Staphylococcus aureus* transform *B. subtilis.* They have been tapped for developing vectors that may be used with several gram positive species. One shuttle vector, made with parts of *Staphylococcus* and *E. coli* replicons, can now be shuttled between these two bacteria.

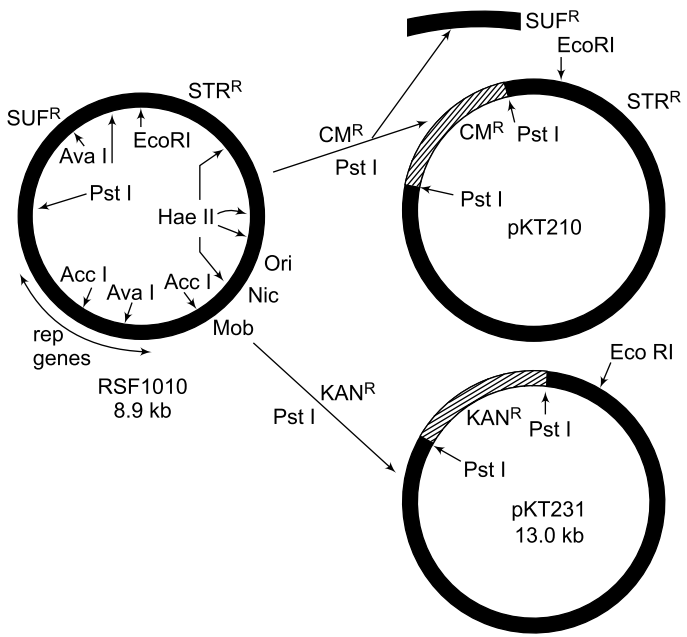


Fig. 2.54 *RSF1010 is a plasmid that replicates in Pseudomonas spp. and in E. coli. It belongs to the Inc Q/P4 group and has a low copy number. The natural plasmid bears sulphonomide (Su^R) and streptomycin resistance (Str^R) genes. Useful derivatives have been constructed, which contain other antibiotic marker genes (replacing the Su), having unique restriction sites for cloning and inactivation of marker genes: pKT210 and pKT231 possess, respectively, the chloramphenicol (Cm) and kanamycin (Kan) resistance genes. RSF1010 may be mobilized for transfer by an F or RP4 conjugative plasmid.*

As will be mentioned from time to time several special broad host range vectors are in the offing. Attempts have been successful in enclosing vectors from one strain into shells of a different strain of virus (ectopic to amphitopic strains). Similarly, agrobacterial plasmids have been spliced with others to provide vectors that will not be restricted to leguminous species.

2.4.6 From Drosophila

Drosophila melanogaster possess two classes of transposable elements that are potential cloning vectors. One class is represented by the P elements and the other by retrotransposons such as *copia.*

1. P Elements

Drosphila strains that always carry the P element (Fig. 2.55) are known as the P strains. There are others named M strains that rarely possess P elements. When a P strain male mates with an M strain female, the offspring exhibit a variety of defects, collectively referred to as the P-M dysgenesis syndrome. The aberrations include mutations, chromosomal aberrations, and germ line defects that lead to sterility. The syndrome is never exhibited in the progeny of the reciprocal cross.

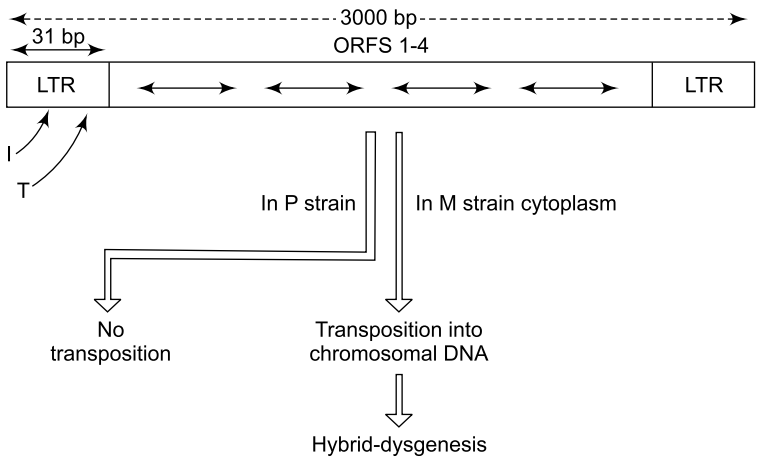


Fig. 2.55 *A schematic representation of the transposable P element (2.9 kb) from Drosophila melanogaster. Four open reading frames (ORFs) have been identified in the DNA between the two 31 bp long terminal repeated sequences. Trascription signals are within the LTRS.*

Such a one-sided affair has been diagnosed as due to the propensity of the P element to undergo transposition in the cytoplasm of the M strain, but not in that of the P strain (which perhaps harbours some repressor mechanism). In addition to this bias, transposition is also tissue-specific; it occurs only in reproductive cells or their precursors. Somatic cells are not affected.

Many naturally occurring strains are P strains, some having as many as 50 copies of the element per cell. The element varies in length from 0.5 to 2.9 kb. In fact, one class consists only of the 2.9 kb elements, and another one of the smaller ones. It

appears that the latter are derived from the former as a result of deletions. The 2.9 kb is the complete element and transposes autonomously with the aid of a transposase enzyme encoded in the element. The smaller elements are transposition defective.

The transposase gene is a spilt one with four exons. The mRNA is spliced differently in somatic and germ-line cells. Only the mRNAs from the latter are translated into the enzyme. The gene is flanked by 31 bp inverted repeats a regular characteristic of all transposable elements (IS or insertion elements without genes, Tn or transposons containing genes, and giant transposons represented by phages such as Mu). A defective element can be transposed if it is supplied in *trans* with the transposase.

The frequency of P element transposition depends on the nature of the P strain and is not quite understood. Some genes are more vulnerable to insertion of the P element than others. The locus *singed* is one of the hot spots for P insertional mutations.

Rubin and Spradling (1982) developed a strategy for the stable insertion of cloned DNA sequences into the germ line cells of *Drosophila,* using vectors based on P elements. The sequence of the *rosy* (ry^+) gene, that provides a component of the wild-type eye colour, was used for the purpose.

Two plasmids were constructed by Rubin and Spradling. One contained the inverted repeats of the P element flanking the ry^+ gene and the other had a P element lacking the terminal repeated sequences at one end. The latter was transposition defective but carried the transposase gene.

Both types of plasmids were injected into the polar plasm of the ry^- *D. melanogaster* embryos. The polar plasm becomes organized into pole cells that differentiate eventually in the adult into gametes. The ry^+ gene is transposed into a *Drosophila* chromosome and remains integrated stably in the offspring derived from the manipulated individual. The presence of the ry^+ locus in the progeny flies was demonstrated by mating adults developed from microinjected embryos to ry^- flies. Among the offspring were some with the ry^+ phenotype. The fruit fly mutant for the rosy allele would have, if untampered with, produced only a ry^- gamete that would yield only homozygous recessive offspring. But the introduction of the ry^+ gene rescued some of the progeny from the mutant state.

Rescue of the mutant phenotype by the insertion of cloned DNA in P elements has been utilized successfully for defining the actual limits of a gene. Different lengths of DNA, including the 5′-and 3′-flanking regions, are used to transform flies that carry the mutant allele for the gene. The minimum length of the inserted gene that restores the wild phenotype is the actual length of the gene. Smaller genes of *Drosophila* (43 kb to 54 kb or less) have been defined in this manner. They include the above mentioned ry^+ (encoding the enzyme xanthine dehydrogenase) and the genes for alcohol dehydrogenase and dopa decarboxylase.

P element vectors have also been useful in studies of gene regulation during development. The gene of interest is fused to the *E. coli lacZ* gene and successful insertion assessed visually *in situ* by monitoring the expression of the *E. coli* gene. Cell and tissue-specific expressions have thus been located, using the promoter for hsp 70 (the major heat shock protein) that functions in almost all tissues during development.

P element-mediated gene transfer promises to be a valuable addition to the gimmicks available to a genetic engineer.

It may be mentioned here that gene transfer into offspring via genetically engineered reproductive cells in *Drosophila* is no different from the gene transfer system that produces transgenic mammals. In the latter (see Section 4.3.8) the vector used to integrate the desired gene in the germ line DNA is predominantly based on retroviruses. The story of the 'transgenic' fruit flies is inserted here to emphasize the use of its own special vector or P element.

2. Retrotransposons

The second group of mobile *Drosophila* elements is similar to the integrated proviral sequences of vertebrate retroviruses. The elements possess long terminal repeats and open reading frames in the intervening DNA. The best characterized elements of the group are *copia* and *gypsy.*

Copia has been utilized for delineating genes by the technique of transposon tagging (see Section 3.2.2). It is not clear whether retrotransposons can be exploited as vectors as well.

2.4.7 From Yeasts

Saccharomyces cerevisiae possesses a natural plasmid in many of its strains. This is a 6.318 kb (2 μm) long sequence in a covalently closed circular form (Fig. 2.56). This plasmid, called the 2 μ. or scp (*Saccharomyces cerevisiae* plasmid), is a multicopy variety with up to 100 copies per cell (Hollenberg, 1982).

There are two 599-base-inverted regions in the plasmids that divide the DNA into two unequal regions. Recombination between these similar sequences results in plasmids that fall into two classes, A and B, differing only in the orientation of the inverted regions with respect to each other. There are several restriction enzyme recognition sites on this plasmid.

There are three *ORFs* (open reading frames, or putative gene sequences) in the plasmid; they are A, REP1 (or B) and REP2 (or C). The product of A is involved in recombination, and the others are involved in the stable but random partitioning of the plasmid copies into daughter cells.

Some strains of yeasts possess endogenous 2 μ plasmids. They are the cir^+ strains. Those without the plasmid are called the cir^- strains. It has been observed that it is more difficult to transform cir^- cells with a 2 μ vector than cir^+ cells. Vectors that contain the entire 2 μ DNA are transferred more easily than those that carry parts of the plasmid lacking the replication region.

Yeast plasmids possess *ARS* (Autonomous Replicating Sequences) and *CEN* (centromere-like) that are needed for replication of the plasmid and its partitioning to daughter cells. These sequences have been exploited to construct what are known as YEp, YRp, YCp vectors, as well as ones known as artificial minichromosomes.

YEps are yeast episomal plasmids derived from the 2 μ plasmid. They occur as multicopies of free circular DNA. Hybrid plasmids containing sequences of 2 μ and MB9 (with Tet^R) have been further modified by the addition of sequences from yeast chromosomal DNA (Beggs, 1981). Some of them are the yeast integrating

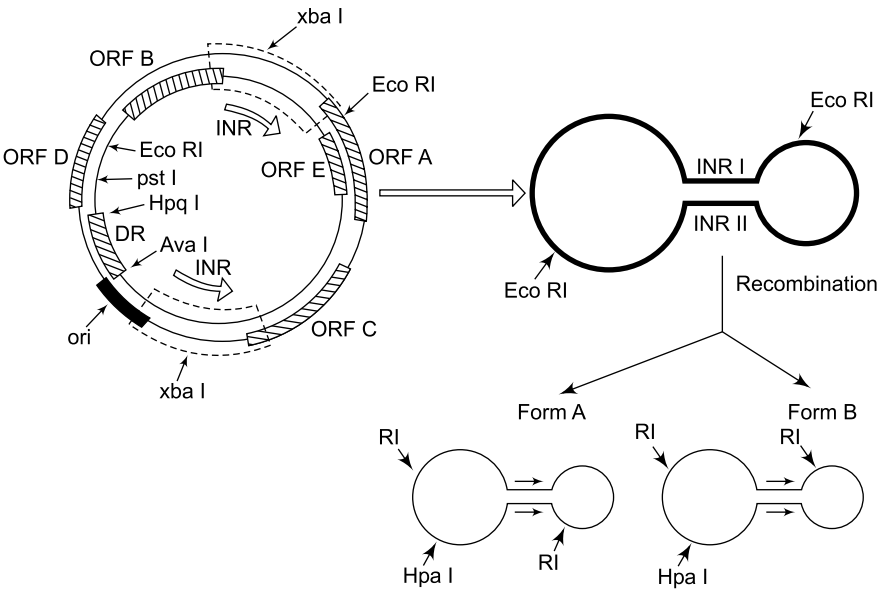


Fig. 2.56 *The 2 μm yeast plasmid. Strains of Saccharomyces cerevisiae possessing this plasmid are said to be cir⁺; those lacking it are cir⁻. There is a pair of 6000 bp long repeated sequences that are in opposite orientations (inverted repeats or INR) which, on pairing, divide the plasmid into a central stem with unequal sized loops at each end. Homologous recombination between the INRs results in types A and B plasmids that differ in the orientation of the sequences in the smaller loop. This is evident from the position of the single Eco RI site in this loop.*

plasmids (YIP) utilized for integrating a donor DNA into yeast chromosomes. LEU2 is one of the yeast markers utilized in these vectors; it allows recognition of transformed leu⁺ cells among untransformed leu⁻ cells.

A YRp contains, besides the *E. coli* plasmid sequences, some sequences from yeast DNA that include the *ARS* segment. These yeast replicating plasmids replicate in yeast cells. Struhl et al. (1979) were the first to utilize these vectors.

A YCp is a yeast vector containing the *CEN* sequences. The latter regulate, partitioning of the plasmids during cell division.

The *ARS, CEN* and *Tr* (telomere) sequences from another source have been combined in a vector known as an 'artificial minichromosome'.

2.5 ❑ THE DNA TO BE CLONED

The piece of DNA to be joined to a vector is known variously as donor, insert, passenger or foreign DNA. It can originate from any source—viral, bacterial, plant, animal or even a synthetic one.

There are basically two ways in which one obtains the donor DNA fragment. If the location of the fragment in the genome is not known, the entire DNA of the cell is fragmented and the one of choice fished out with a 'probe'. The latter is a nucleic acid (single-stranded) fragment that is complementary to the target DNA or else some selection technique may be employed to retrieve the desired fragment.

If the fragment represents a gene, the mRNA of the gene is copied into a cDNA using reverse transcriptase and the ds cDNA is used for splicing to the vector.

Both procedures are very much shortened if one begins with a gene bank or library. A set of DNA clones representing the genome is already available in such a collection and all one has to do is retrieve the clone carrying the insert DNA of choice.

A gene bank or library is a collection of fragments of DNA, each one of which is spliced to a vector DNA. The genomic library is made from fragmenting the total DNA of a cell, and hence represents, theoretically, every region of the genome. A cDNA library represents the genes in the total DNA—rather the regions of the genes that are copied into an RNA. This means that only the coding sequences and the leader and trailer regions are represented, and in the case of split genes, only the exons of the coding sequences. A cDNA library may not represent every gene, as the mRNAs of some may be very poorly represented or too transient for preparing cDNAs.

The total DNA may be fragmented mechanically by using a sonicator, or enzymatically with the help of restriction endonucleases. In the former case the approximate desired length of the fragments may be controlled, but the ends of the fragments will have random sequences. In the latter case, the fragments will be of different lengths, due to the asymmetrical locations of the restriction sites, but the cleaved ends will be defined in terms of sequences of nucleotides. Mechanically sheared DNA fragments may be made usable for cloning by modifying the ends appropriately. Modification of restriction fragment ends may also be needed to make them suitable for splicing to vectors cut with a different enzyme or for extending a single stranded tail on each side of the fragment, or for converting them to flush ends for joining to another flush end. The tricks for tailoring DNA fragments are surveyed in the section on 'DNA engineering'.

2.6 ❑ INTRODUCING GENES INTO PROKARYOTES

There are no regular sexual processes in prokaryotes that ensure imports of alleles into a cell from other members of the species. The prokaryote relies on a few parasexual channels for such acquisitions. They include the gene transfer methods referred to as *transformation, transduction, conjugation* and *transposition.* (Fig. 2.57).

Advantage has been taken of these natural routes to develop laboratory techniques for transferring exogenous DNA into prokaryotic cells. The *in vitro* techniques go one step further than the natural ones in that they insert genes that may be foreign to the recipient cell. The main gene transfer techniques for prokaryotes include transformation and transduction. As the transducing phages contain DNA foreign to the prokaryotic cell, the laboratory method is basically one of infecting the cell via transduction or transformation mechanisms. Hence these techniques are often referred to as those of *transfection.*

Most of the transfection procedures have been developed for the predominantly utilized prokaryote, the bacterium *E. coli.* Similar procedures are becoming available for other prokaryotic species such as *Bacillus subtilis, Streptomyces* and blue-green algae (*Cyanophyta*).

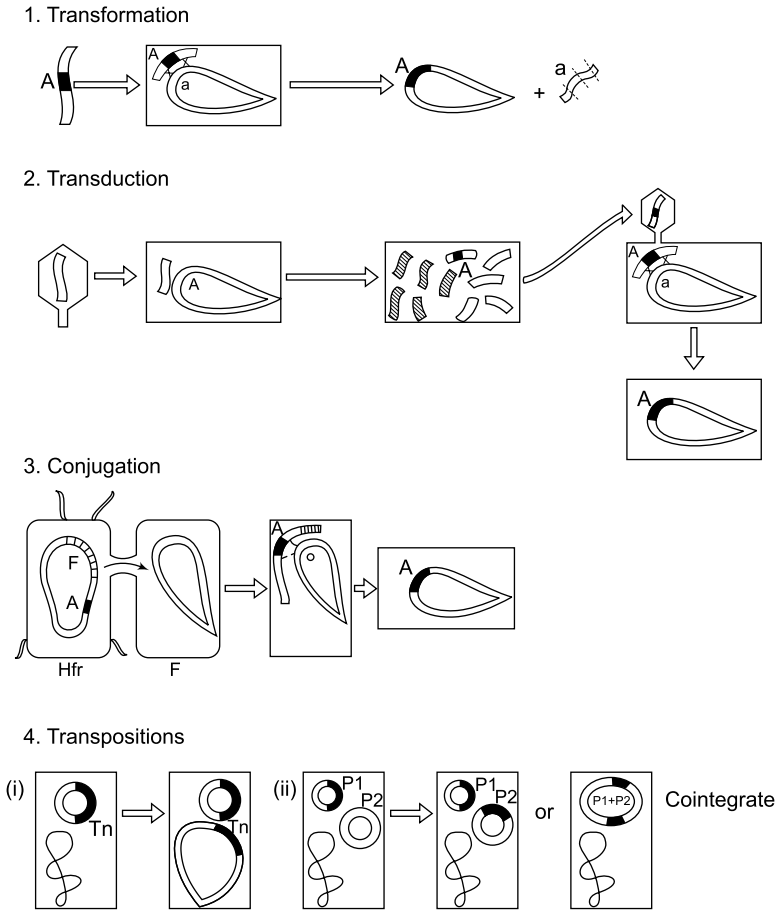


Fig. 2.57 *Parasexual modes of gene transfer in prokaryotes. (1) Transformation by naked DNA, (2) transduction by phage carried DNA, (3) conjugational transfer via a tube connecting donor and recipient cells introduces exogenous alleles via homologous recombination, and (4) the fourth mode, transposition, employs an entirely different method for joining totally unrelated DNA sequences. Gene cloning by and large uses the first two and the last modes for transforming cells with foreign DNA.*

2.6.1 Natural Parasexual Models of Gene Transfer

A brief note on the term 'transformation' would be relevant before proceeding further. The original and basic meaning is the 'genetic transformation' of the recipient cell with an imported DNA. The term is also used, in a different context, for eukaryotic cells altered genetically to acquire the tell-tale features of a tumourous or cancerous cell; so is it used freely to refer to the process of entry of free DNA (not carried by a pathogen) into a cell. In the following paragraphs the term has been utilized as a mode of entry of free DNA into the *E. coli* cell.

1. Transformation

Transformation occurs by the entry of naked *E. coli* DNA fragments through the cell membrane of a recipient *E. coli* cell. There is evidence that only ds DNA fragments take part in transformation, and that one of the strands is degraded while passing into the cytoplasm or in the cytoplasm itself. After a brief period (known as the eclipse period), the remaining strand becomes duplex and the latter aligns with homologous loci in the resident DNA. Non-reciprocal recombination between the incoming and endogenous DNA transfers the exogenous allele(s) to the recipient cell DNA. Not every cell is able to take in exogenous DNA. The cell appears to go through a physiological condition, referred to as that of 'competence' (ill-defined as yet) during which it is capable of bringing in outside DNA through the cell membrane again. Not all exogenous DNA survive in the recipient cell; they are vulnerable to the nucleases and especially restriction ones of the recipient cell, unless the restriction sites are nullified by appropriately methylated bases. These features have to be remembered while arranging for 'transformation' in the laboratory.

2. Transduction

The second parasexual means of gene transfer, or transduction, is a process by which bacterial DNA fragments are introduced via natural bacteriophage vectors that infect the bacterial strain in question. In *E. coli* there are two types of transductants—the general and the site-specific transducing phages. Phage P1 is representative of the former and phage λ of the latter classes. In *generalized transduction,* random fragments of the *E. coli* DNA are introduced into fresh cells by infection with the phages carrying the DNA fragments. This happens because the enzyme that cleaves the concatamer phage DNA into phage-size fragments also cleaves the host cell DNA into similar sized pieces. Some phage shells include the latter by mistake. These *E. coli* DNA fragments are thus transported by infection into a new host cell. In *site-specific transduction* (Fig. 2.58), the phage DNA becomes integrated at a specific site in the *E. coli* chromosome. For *E. coli* DNA this site is called the *att* (attachment) locus and it is positioned between the *gal* (galactose) operon on its left and the *bio* (biotin) operon on its right. The phage λ DNA also possesses the 15 bp motif present in the *E. coli* DNA. The phage and bacterial *att* regions get aligned, and staggered cuts are made with a phage-encoded enzyme in both these regions. Recombination integrates the smaller and larger DNAs into one large circular DNA. The *gal* and *bio* regions are now separated by a length of λ DNA. Under certain conditions (e.g., damage of the cell due to raised temperatures, infection, and so on) the integrated phage DNA is released from the bacterial DNA. This deintegration is referred to as 'induction' of the phage. The bacterial cell harbouring an integrated phage DNA is said to be a lysogen, and the phage is said to be indulging in the lysogenic mode of the life cycle.

In the majority of events, the induced phage makes a clean break from the host cell DNA; only the phage DNA is cleaved out. However, mistakes occur in which the cleavage become displaced, so that the lambda-size DNA that is released contains a portion of the adjoining *E. coli* DNA. Of course, the phage leaves behind an equivalent length of its genome in the host chromosome (Fig. 2.59). Such

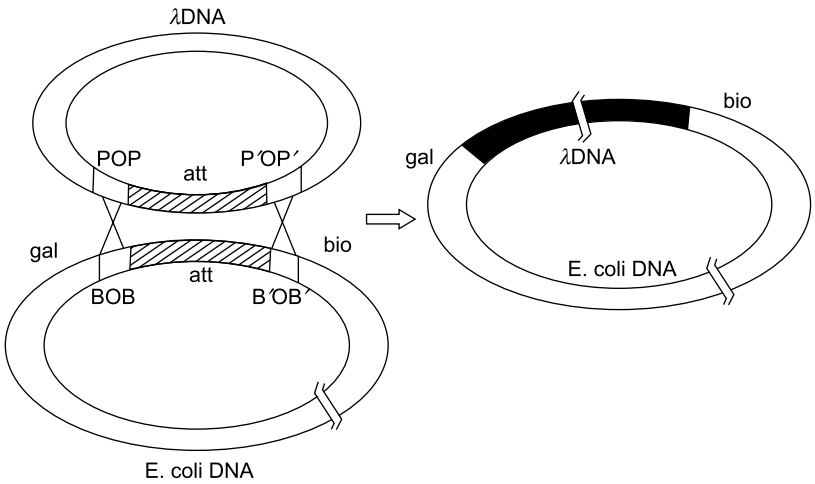


Fig. 2.58 *Site-specific recombination. Homologous recombination at a specific site (attachment or att sequences) in phage λ and E. coli DNA integrate the entire phage DNA into the E. coli chromosomal DNA. The att region in DNA is flanked by sequences POP and P′OP′, respectively. Similarly, sequences BOB and B′OB′ occur, one on each of the two sides of the bacterial att region.*

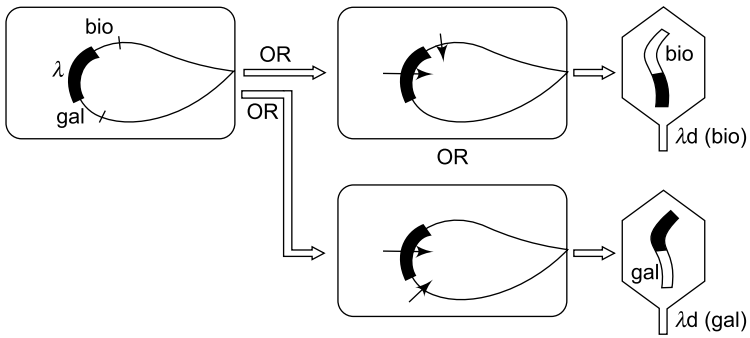


Fig. 2.59 *Defective transducing bacteriophage dλ (gal) and dλ (bio). On induction, the phage DNA is released from the chromosomal DNA in a precise manner. However, accidental and imprecise excision may include E. coli sequences on either side of the att region. The latter lies between the galactose (gal) and biotin (bio) operons of E. coli. Defective induction may thus produce phages with the gal or the bio sequences in place of an equal length of phage DNA. On infection of a new host cell, the E. coli sequences may recombine, causing replacement of a resident allele with an imported one. Thus an E. coli may come to possess two sequences of either the gal or the bio loci.*

phages are said to be *defective*. If the bias, while cleaving, is on the left side of the phage DNA, a portion or the whole of the *gal* operon may become included in the DNA of the now defective λ. Similarly, displacement to the right would include regions of the *bio* sequences. Such defective lambda phages are labelled as λ d(*gal*) or λ d(*bio*). Since phage λ preferentially integrates at a specific site in the *E. coli* DNA it is possible to exploit this fact to include genes of choice in the defective phage.

3. Conjugation

The third natural method of importing alleles by *E. coli* cells is that of conjugation. One of the natural plasmids of *E. coli*—the F plasmid—carries genes that are capable of setting up the infrastructure for autotransmission of the plasmid into an F^- cell. A cell carrying an F plasmid possesses hair-like structures on its surface; these are known as *pili* (singular *pilus*). In some cells, the F DNA is integrated into the *E. coli* DNA. Such cells also exhibit pili on their surfaces. A cell with a free F plasmid is called an F^+ cell; a cell without one is called an F^- cell, and the one having an integrated F^- is an *Hfr* (High Frequency Recombination) cell. The latter one gets its sobriquet from the fact that such cells are involved in transfers of *E. coli* alleles to F^- cells thus mediating recombination of *E. coli* alleles.

In a nutshell, the process of conjugation consists of the following steps: An F^+ or *Hfr* cell gets attached to an F^- cell through a protein at the tip of a pilus for which there is a receptor on the F^- cell surface. The membranes at the point of contact are degraded, such that the pilus becomes a hollow tube or channel connecting the two cells. Next, one of the strands in the F plasmid is cleaved at a specific site, and the free 5′-end of this strand moves into the conjugating tube. As it proceeds into the tube and out into the cytoplasm of the F^- cell, the intact circular strand of the F plasmid acquires a complementary strand. So does the strand that moves out. The linear ds F DNA in the recipient cell becomes circularized into an F plasmid. The recipient thus becomes an F^+ cell.

An integrated F DNA also is capable of transfer by conjugation. In this case, as the *E. coli* DNA is sandwiched between the region of the F DNA that has the 5′-free end after nicking of the F strand and the remainder of the nicked F DNA, the bacterial chromosome is also pulled through the conjugation tube. Under unusual conditions the *F-E. coli* DNA combine can be transferred intact; the process takes around 90 minutes. Usually, though, the conjugation tube becomes severed at varying lengths of time after the beginning of transfer. Hence, the recipient cell receives varying lengths of *E. coli* DNA—the length depending on the time of breaking of the tube relative to the starting time. Wolman and Jacob were the first to map the genes on an *E. coli* DNA by examining the genes that were transferred at different time intervals. The recipient cell used was a mutant for several markers. The 'mating' of cells was interrupted by their shaking aliquots removed at short intervals. Such aliquots with ruptured conjugation tubes were plated on several 'marker' plates. These marker plates contain the minimal medium plus a mixture of nutrients, but lacking in one of them in each different plate. If *E. coli* colonies appear on, say, a *leu* marker plate (that is, one without added leucine), one is entitled to claim that the cells possessed the leu^+ gene. Since recipients were leu^-, the leu^+ allele must come from the *Hfr* cell. The number of leu^+ revertants (due to mutation) is deducted from the number of counted leu^+ colonies on control plates where no mating mixture has been added.

This *interrupted mating technique* was used by Wolman and Jacob to assign map positions to the genes of *E. coli.* Hence, *E. coli* gene loci are usually given units of 'minutes' after the start of transfer. The circular DNA is divided into 100 and gene loci assigned on this scale.

Due to the superficial analogy of this transfer of genetic material to the process in sexually reproducing species, the F^+ and *Hfr* are referred to as 'male' cells. Phage

M13 only infects 'male' *E. coli,* as the initial attachment occurs only on a receptor on the pilus.

The process of conjugation is not exploited in genetic engineering protocols. Indeed, care is taken to see that the genes cloned after a great deal of effort do not disappear from a clone by conjugational transfer. However, certain features of the process are utilized where feasible. For instance, M13 phage vectors have to be used with F^+ cells that have pili. The transfer functions in the F in such cells are deleted by mutation of the relevant genes. In other cases, it may be required to mobilize transfer of a non-autotransmissible vector; a helper F plasmid or *tra* genes may be included in the system.

4. Transposition

Transposition involves the integration of a copy of a defined segment of a DNA from another locus, either in the same or in a different molecule of DNA. Such movable, mobile, or *transposable elements* possess a characteristic structure with repeated sequences at each end. These terminal repeat regions are essential for the transposition event. They contain the gene for the enzyme transposase that mediates the event, or the transcription signals for the gene which may be in the DNA intervening between the element ends.

Different mobile elements have been identified. They are grouped arbitrarily as insertion (IS) elements, transposons (Tn) and certain macrotransposon-like bacteriophage genomes (such as that of phage Mu of *E. coli.*). IS elements are relatively short in length and do not possess detectable genes. Transposons carry one or more genes besides the one for transposase. These genes provide additional qualities to the cell containing the transposon genes. The genes are predominantly those that confer resistance to drugs and toxins (including toxic heavy metals), encode toxins (e.g., colicin of *E. coli*) that act against susceptible microbial species or some very unique ones that specify proteins with unusual functions. To the last category belong transposons with genes for nitrogen fixation and T DNA genes on the Ti or tumour-inducing plasmids of *Agrobacteria tumefasciens* that cause infected dicotyledonous cells to become tumourous, The drug-resistant genes include Amp^R, Str^R, Tet^R, Neo^R and so on, which protect a cell from the effects of ampicillin, streptomycin, tetracycline, and neomycin, respectively.

2.6.2 Cell Transformation with Plasmids

Although bacteria are known to be transformed by exogenous DNA in nature, the efficiency of transformation in the natural habitat is very low. Mandel and Higa (1970) were the first to develop a method for improving the efficiency of uptake of free phage DNA by *E. coli* cells, that involves treating the cells in succession by chilled $CaCl_2$ and heat shock. Later, Cohen discovered that the same treatment enhanced the uptake of plasmid DNA.

Other modifications have been introduced. Of these the one by Kushner (1978) and another by Hanahan (1983) deserve mention. Kushner used rubidium chloride and $CaCl_2$ together with DMSO (dimethyl sulphoxide) followed by the heat shock treatment. Hanahan's method includes, in addition, $MgCl_2$, $MnCl_2$, and hexamine cobalt (III) chloride.

The cell transformation protocol consists of treating $CaCl_2$-treated (0–5°C for 0.5 to 24 hours) *E. coli* cells with heat shock (42°C). The cells thus made competent are mixed with the DNA to be introduced into the cells. The Ca^{++} appears to make the DNA adhere to the cell surface and the heat treatment allows the membrane to be temporarily premeable to the exogenous DNA. The mechanisms underlying these steps are not clear.

Some bacteria are readily transformed by DNA fragments from the same strains. *Hemophilus influenzae. Bacillus subtilis* and *Streptomyces* spp, belong to this group. The much utilized *E. coli,* however, is recalcitrant to transformation without treatments (such as the $CaCl_2^-$ heat shock combination described earlier) that make them 'competent'. It appears that bacteria of the first group possess receptors on their surfaces that bind them to specific sequence motifs in DNA of the same species.

Artificially induced transformation is not only restricted to DNA of a strain identical to that of the recipient cell, but it also ensures a larger proportion of transformants in a treated population. Of course the term 'efficiency' is a relative one when used in connection with bacterial transformation. The actual number of transformants after treatment is quite small. A recovery of 10 million transformants per microgram of exogenous DNA is considered to be quite good, compared to the proportion of success in untreated cells. Still, this figure represents, at most, one plasmid per 10 thousand in a case where 10^7 transformants are recovered from the addition of one μg of DNA. Hanahan's method transforms as many as 10^8 cells per microgram of DNA, especially with the χ1776 strain of *E. coli.*

2.6.3 Transfection with Phage Vectors

The phage vector-carried gene/DNA of interest may be introduced into an *E. coli* cell either by transforming with the naked vector DNA, or via infectious particles carrying the gene to be introduced. The latter particles are constructed around replicated recombinant DNA copies with the help of a λ-packaging mix that has been described earlier.

Transfection by packaged particles has a slightly higher efficiency than even the most efficient protocol using transformation protocols.

2.7 ❑ INTRODUCING GENES IN EUKARYOTES

Unlike prokaryotes, a large proportion of eukaryotes are multicellular organisms. Introduction of exogenous DNA into such organisms must, therefore, be considered, at least at two levels: transfer into single cells and transfer into whole organisms. Techniques for the former are varied; the one to be used in a project depends on the nature of gene expression expected—transient or permanent. Transfer into whole organisms includes stable integration of the exogenous gene in the genome of the species; individuals which were the original recipients of the gene should transmit it stably into the offspring, in which expression is regulated developmentally in a correct manner. Methods for creating such transgenic animals and plants have begun to be available only recently. Corrective measures for genetic defects may perhaps become possible by using such techniques. In the meantime, there is a report of

the alleviation of an individual's inherited shortcomings by the implantation of a tissue carrying the correct gene. Inder Mohan Verma and associates, at the Salk Institute, California, have successfully introduced into mouse a colony of cells with a corrected gene. The cells expressed the correct product within the live animal.

By and large, genetic engineering of eukaryotes is envisaged at present at the level of individual cells. Several ways of introducing genes into animal, plant and yeast cells have been developed that allow primarily the study of the biology of the eukaryotic genetic system. Some of the more prominent ones are briefly surveyed below. Gene transfer strategies for plants are described separately in Section 5.4.

2.7.1 The Recombinant Viral Technique

The recombinant viral technique involves the transfer of an exogenous gene into an animal cell via a recombinant DNA made with a viral vector. So far, SV40, adenovirus and papilloma viruses have been utilized.

Late replacement vectors of SV40 can be utilized to construct the recombinant DNA, and replication of the hybrid vector ensured by cotransfecting the host cell with a tsA mutation carrying helper SV40. The cloning cell for SV40 vectors is CV-1, the cultured kidney cell line of monkey. Mulligan et al. (1979) employed this strategy to clone the rabbit β-globin gene in CV-1 (see Fig. 2.30). The globin gene was taken from a clone already made by Maniatis and associates (1976). The β-globin gene was represented as a cDNA copy. The globin sequences were cleaved out from the cloning vector and spliced to a late replacement SV vector (SVGT-5), in which the signals for DNA replication, transcription, splicing and polyadenylation were intact. The transfer of the cDNA to the SVGT-5 was performed in two steps. The cleaved out globin cDNA with appropriately retailored ends was first spliced to a pBR322 and the hybrid plasmid amplified in *E. coli.* In the second step, the gene sequences were separated once more from the vector DNA and inserted in place of the late gene VPI in SVGT-5. The recombinant viral DNA was then transferred into CV-1 cells at 41°C. Correct β-globin was synthesized at a high level, indicating that it was inserted correctly in the vector. Several other genes have been expressed by this technique (see vectors from SV40 in Section 2.4.4).

2.7.2 DNA Mediated Gene Transfer Methods

There are a few techniques by which naked DNA, which is not necessarily spliced to vectors, may be introduced directly into mammalian cells. One of these DNA-mediated transfer techniques utilizes calcium phosphate and another one utilizes the compound DEAE-dextran, to facilitate the entry of the DNA into recipient cells.

1. The Calcium Phosphate Method

The calcium ion-facilitated method involves the formation of calcium phosphate, in a solution containing donor DNA, by the action of calcium chloride in a phosphate saline buffer. The DNA becomes complexed with the freshly formed calcium phosphate. The recipient cells are mixed with the DNA phosphate-containing solution and incubated for two days at 37°C. The complexed particles enter the

cells through endocytosis or phagocytosis (it is presumed); some of the DNA become integrated into the chromosomal DNA in the cells (Fig. 2.60). The 48-hour incubation allows the stably integrated gene to express itself. The transformed cells can then be identified by a suitable strategy. The procedure calls for adding the DNA first to a phosphate saline buffer, followed by addition of a $CaCl_2$ solution, and constant stirring by bubbling air through the reaction mixture. In about half-an-hour the solution turns opaque and milky in appearance. The preparation is poured on cells grown in suspension culture as suspension cultures appear to produce more transformants than monolayer ones (Chu and Sharp, 1981).

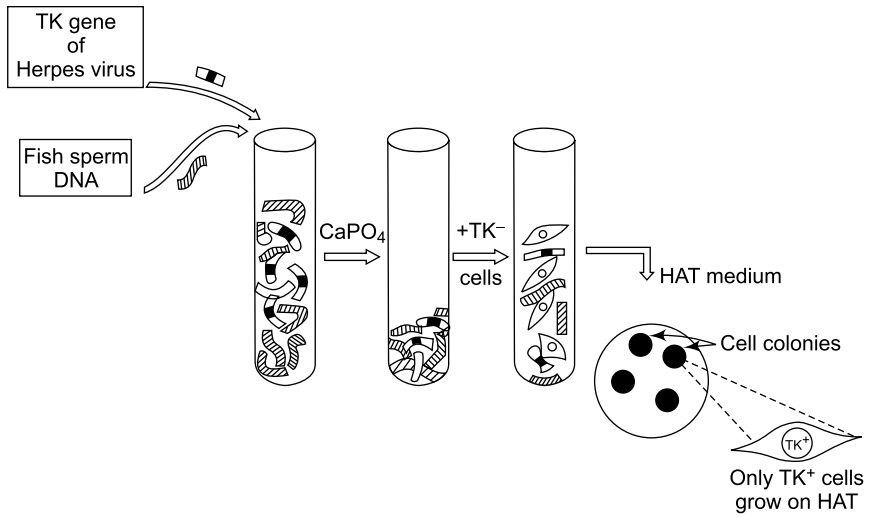


Fig. 2.60 *DNA-mediated gene transfer by the calcium phosphate precipitation method. As the gene to be transferred is usually available in very minute amounts, a carrier DNA, such as that from fish sperm, may be added to provide a reasonable bulk.*

The calcium phosphate precipitate technique was introduced in 1962 by Szybalska and Szybalski who transformed *hgprt*$^-$ human cultured cells with *hgprt* DNA. The efficiency of transformation by this method is not too impressive. Only a small proportion of the 2–3% of the cells that become transformed are able to integrate the exogenous gene into their genomes. Modified techniques have raised the value to a maximum of 15–25% of transformed cells. The success level also depends on the type of cell used as the recipient. While human cultured cells (fibroblasts) are poor transformants, various other strains are excellent importers of exogenous DNA. However, this method of DNA transfer is still more efficient than the one using calcium chloride treated *E. coli* cells. Nevertheless Axel employed this technique for transfecting mouse *tk*$^-$ cells with the Herpes simplex virus (HSV-1) *tk*$^+$ gene. The transformants were selected on a HAT medium. This DNA-mediated transfer technique requires a filler or carrier DNA to provide sufficient bulk if the level of donor DNA utilized is very low. Usually fish sperm (e.g., of salmon fish) or calf thymus DNA is used as the carrier DNA. Higher molecular DNA isolated from other eukaryotic sources may also be employed as carrier DNA.

If the donor DNA does not oblige by being a selectable marker, it may be coprecipitated in the calcium phosphate mixture with a DNA that carries a dominant marker (such as the tk^+ gene from HSV). Almost every cell that is transformed with the selectable marker containing DNA also contains the donor DNA.

2. The DEAE-dextran Method

DEAE (diethylaminoethyl)-dextran-treated viruses have been shown to increase the efficiency of infection of cultured cells of several mammalian species. Based on this fact, a technique has been evolved that transfects mammalian cells with donor DNA.

As in the case of $CaCl_2$-treated cells of *E. coli,* it is not clear as to how this treatment helps the DNA to pass into the cells. Perhaps the compound helps the DNA to come into close contact with the cell surface; alternatively, the membrane is modified in some manner to allow it to be more permeable than usual to outside DNA molecules.

The DEAE-dextran technique has a higher success rate than the calcium-phosphate mediated one; up to 25 per cent of the treated cells may be transformants.

Prolonged incubation in the treatment solution raises the efficiency of transfection to almost half the treated cell population. This was demonstrated by Sompayrac and Danna in 1981, when they incubated cells and DEAE-dextran-treated SV40 particles for about 16 hours.

2.7.3 Protoplast Fusion or Somatic Cell Hybridization

A variety of reproductive barriers prevent the formation of interspecific or intergeneric hybrids. These barriers may be bypassed by fusing two cells into a hybrid one. The somatic cell hybrids are viable long enough for studies of transient gene expression. In the case of plants, some interspecific hybrid cells have been differentiated in culture into mature and fertile plants. The majority of species, however, do not respond to the current procedures for regeneration from single cells.

Somatic cell hybridization was initially hailed as an excellent means of transferring bulk genetic material to a eukaryotic cell. However, it appeared for some time that the technique would remain a valuable one for research, but not a practical one for inserting a foreign genome into a recipient species.

There has been a renewed incentive to utilize somatic cell hybridization for gene transfer in eukaryotic cells. A bacterial cell harbouring recombinant genes can be fused to a mammalian one, thus avoiding the hazards of the inefficient transformation and transfection systems.

The technique for introducing recombinant DNA via whole *E. coli* cells was first reported by Schaffner (1980). It has been refined since then such that 100 per cent of the cells become transformed. When an SV40 sequence containing rDNA was transferred into mammalian cells in culture, successful transformation could be assessed quantitatively by the number of foci formed on the plates.

The fusion between two cells is due to a merging of the two plasma membranes in response to a fusing agent (Fig. 2.61). Polyethylene glycol is the agent of choice. However, cells with a cell wall surrounding the plasma membrane can fuse only when the wall is removed.

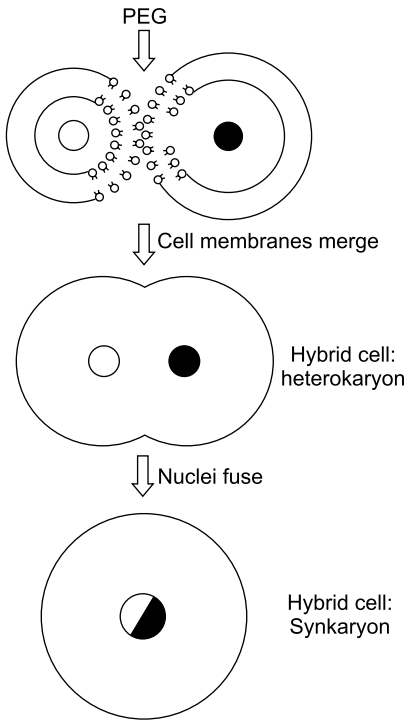


Fig. 2.61 *Gene transfer by somatic cell fusion. PEG (polyethylene glycol) is used as a fusing agent. It destabilizes cell membranes. Cells with destabilized membranes, when in contact, become enclosed within one membrane. These cell hybrids at first have separate nuclei (heterokaryons). Later, the nuclei also fuse; the hybrid cell is now known as a synkaryon. Somatic cell fusion is achieved when the partners are both eukaryotes, one a eukaryote and the other a prokaryote, as also between organelles and cells.*

The wall-denuded cells are referred to as protoplasts. In plant cells, the pectin cementing the cells in a tissue, and also the cellulose of the wall have to be removed. This is done by using the enzymes pectinase and cellulase, which are available in a very pure state from commercial suppliers. Yeasts possess a cell wall with a composition different from that of plant cells. Special yeast wall-degrading enzymes (Zymolase, Gluculase etc.) are available. Bacterial cells also possess characteristic cell walls. Enzymes for these have to be used. The protoplasts become spherical when not confined within the rigid cell walls; hence they are often known as 'spheroplasts'.

Schaffner fused *E. coli* protoplasts with mammalian cells. This method is increasingly in use for transferring rDNA between bacteria and mammalian cells, as also between members of prokaryotes.

2.7.4 Microcell Fusion Technique

Microcells are miniature structures containing one or more chromosomes and are enclosed in a plasma membrane. They are produced in the following manner: Cells

are treated with an inhibitor of mitosis, such as colecemid. Chromosomes divide, but no new cells are formed. If subjected to colecemid for an extended period, the individual chromosomes may acquire a nuclear membrane around them. If cells with micronuclei are treated with cytochalasin B (Wigler and Weinstein, 1975) that removes intact nuclei from cells, and the mixture centrifuged, free microcells (wrapped in plasma membrane) can be recovered. They may contain aberrated chromosomes or more than one chromosome each. (Fig. 2.62).

Microcells can be fused with a mammalian cell, using the fusing agent PEG. Microcells are recovered by centrifuging in a gradient of Ficoll and purified by available protocols. (Fournier and Ruddle, 1977, Zornetal, 1979, McNeil and Brown, 80).

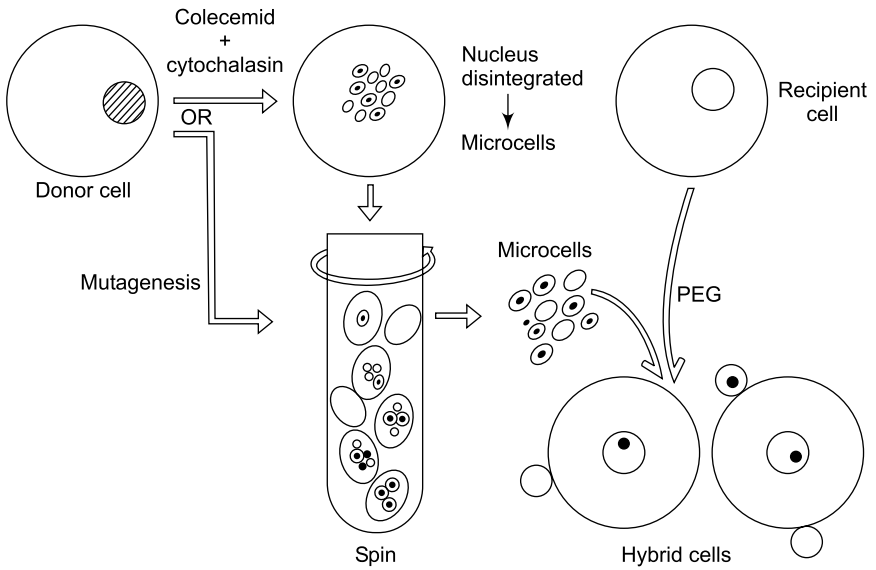


Fig. 2.62 *Microcell fusion method. Treatment with colecemid (a mitotic spindle poison) and cytochalasin causes the nucleus to become disintegrated and each fragment (one or more chromosomes) becomes enclosed in its own membrane. These micronuclei are isolated by centrifugation with added cytochalasin B. They are now called microcells. These microcells may be utilized to transfer single chromosomes by fusing to a somatic cell. It helps if the microcell carries a selectable marker. Microcell fusion has been useful for specific mammalian gene-mapping projects.*

2.7.5 Metaphase Chromosome Transfer

Genes that are arranged in the native chromosome may be transferred into a mammalian cell in the form of whole chromosomes. As the latter are condensed the most at the metaphase stage of mitosis, it is easier to isolate chromosomes at this stage.

The usual procedure (Nelson, 1941) for isolating metaphase chromosomes is to treat mitotically dividing cells with an inhibitor such as colchicine, colecemid or vinblastine, followed by lysing the blocked cells in a hypotonic buffer, with separation of the chromosomes by subsequent differential centrifugation. The

mitotic inhibitors prevent re-forming of the nuclear membrane after mitosis. The heavier metaphase chromosomes can be readily separated from the background cytoplasm.

The metaphase chromosomes are then introduced into the recipient cell by one of a few different techniques. The method of isolation was introduced by McBride and Ozer (1973) and refined further in Ruddle's laboratory (Willecke and Ruddle, 1975, Filler and Ruddle, 1978). The improved methods are significantly more efficient (10^{-5}–10^{-4}), than the earlier ones (10^{-7}–10^{-6}).

Two methods of transferring isolated chromosomes into cells have been utilized: co-precipitating with calcium phosphate and mixing the complexes with the recipient cells and fusing liposomes containing a chromosome each to the target cell.

The various studies with transferred metaphase chromosomes indicate that they are possibly taken into cells by endocytosis and are degraded in the target cell cytoplasm, except for a few which may gain entry into the nucleus. There are indications that genes on the chromosome are expressed in the target cell. There are also indications that the stably transformed cells have the foreign chromosome integrated in a host chromosome. The degree of success is, however, quite low.

One serious drawback of this technique is the absolute need of having a selectable marker gene on the chromosome and a recipient cell mutant for that gene. For experimental purposes, the standard markers such as tk^- and $hgrt^+$ have been useful. But as a practical proposition, one must have a selection system that can identify chromosomes that do not have known markers, and use a wider spectrum of target cells.

Another negative point is the lack of any technique for preserving isolated chromosomes. The transfer of a whole chromosome is highly tempting, as such a procedure would allow the desired gene to be *in situ* with its native control signals and neighbouring sequences. We do not know as yet whether development *in vivo* is dependent on signals that are identical or different from the ones that control expressions *in vitro* or from detached genes *in vivo.* Studies with transferred metaphase chromosomes should provide an answer to this question.

2.7.6 Via Liposomes

Liposomes are spheres with phospholipid membranous walls that may be filled with a drug or any other substance of choice. The latter are entrapped during the process of making the liposomes.

Liposomes with entrapped metaphase chromosomes, or vectors with donor DNA, have been fused with target cells. Wong (1980) introduced the Amp^R gene (on pBR322) into several cultured animal cell strains using this method.

Liposomes are made by sonication (Wong) or reverse phase evaporation (Szoka and Papahadjoupoulos, 1978) or by utilizing an automatic liposome maker (e.g., Liposomat). Wong's method involves sonicating for 10 seconds, in a sonicator, a mixture of the recombinant DNA in a phosphate saline buffer, and phosphatidyl serine and cholesterol, the last two in diethyl ether. Tiny spheres or liposomes are formed which are rid of the ether and adventitious donor DNA. The cleaned liposomes are introduced into target cells using the calcium phosphate precipitation technique. Some investigators use PEG to fuse the liposomes with the target cells.

2.7.7 Microinjection Technique

DNA may be injected directly into cells (nuclei, if necessary) through glass pipettes that are pulled to produce very fine bores ($\simeq$ 0.1 μ–0.5 μ) at the tip.

The cells to be injected are placed in drops on a microscope glass slide. This slide is inverted over a chamber made with glass slides and the edges sealed with grease. The cells are now in the form of hanging drops. This arrangement is placed on the stage of a phase-contrast microscope. A micromanipulator holding a micropipette is set up beside the microscope. Looking through the microscope, the DNA solution, taken up in the micropipette, is released into the nucleus of a cell (Fig. 2.63). Refinements in the apparatus allow a fairly constant amount of solution to be injected at a constant pressure.

A simpler alternative method requires the cells to be placed in a buffer solution on a glass slide that can be placed face up on the stage of an inverted microscope, and then the DNA is injected into each nucleus. One is able to recover 100 per cent transformation by this method, although it takes a while to inject each cell (maximum of around 1000 cells per hour).

Plasmids with the *tk* gene and globin gene were microinjected into mouse tk^- (L strain) cells by Kretschmer, Linda Saunders Haig and Diacumacos. Expression of the genes was assessed by Southern blotting and the autoradiography of the blots hybridized with probes for the two genes.

A variant of the microinjection method floods the cells with the solution containing the DNA, and each cell is pierced with a fine glass needle through the nucleus. Some of the DNA goes into the nucleus. However, the efficiency of transformation in this case is very low in comparison with the microinjection method.

2.7.8 Electroporation

Exogenous DNA can be introduced into a cell by subjecting the cell to electrical pulses. This technique is especially suitable for introducing DNA into plant cells for both transient gene expression and long-term integrative studies.

Short or long pulses of an electric field are used to form pores in the membrane of naked protoplasts. The short pulses use a higher voltage ($\cong$ 1400 V) than the long pulses (400 V); the former appear to be more useful for stable integration and the latter for studies of transient expression of the introduced DNA.

The protoplasts are first given a heat shock (at 45°C in a water-bath, for 5 minutes), followed by incubation in an ice-bath until the temperature of the protoplast suspension comes down to room temperature.

Several 'electroporator' instruments are available in the market, and are very efficient for transforming large populations of cells under fairly uniform conditions.

If the amount of transfecting DNA is very little, the bulk is increased to an acceptable value with a suitable carrier DNA. PEG may also be added to the DNA solution before mixing it with the protoplasts. The transformation mixture is placed in the chamber provided in the electroporator and pulses (usually at intervals of 10 seconds) of appropriate strength given. The treated cells are plated out and, in time, tested for transformation by the donor DNA.

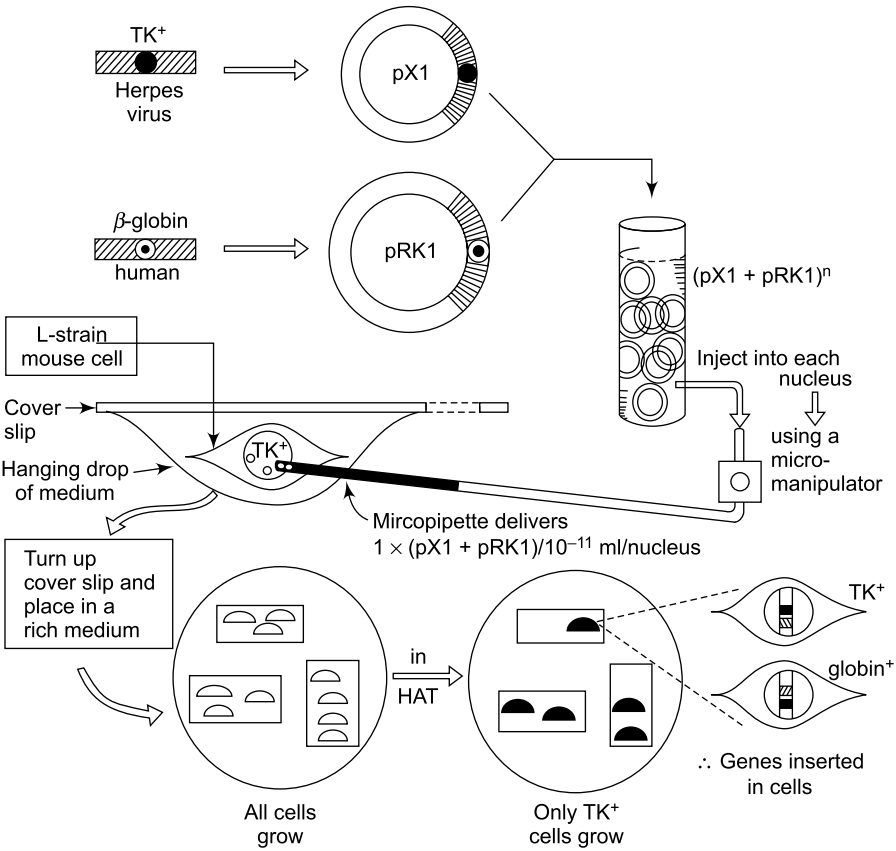


Fig 2.63 *The micrionjection technique. Genes to be microinjected are first closed in plasmid vectors. This illustration represents the introduction of the Tk+ and β-globin genes into Tk− (L strain) mouse cells.*

The electroporation technique is slightly less efficient than the DEAE-dextran method of DNA transfection. However, this technique can be especially valuable for transferring genes into cells that do not take up exogenous DNA by the conventional techniques. Falkner and Zachau (1984) achieved this transfer for lymphoid cells of mouse.

2.8 ❑ CELLS FOR CLONING

The recombinant DNA must be introduced into a cell, within which it may replicate freely. In addition, in many exercises, it is required that the introduced gene also be expressed within this cell.

Transformation of cells of all species is not equally efficient. In such cases, the manipulative exercises on the DNA are performed in *E. coli* and the ready-made recombinant DNA introduced into the more recalcitrant host cell. Several shuttle or bifunctional vectors (plasmid) have been developed with which use of two hosts, one of which is *E. coli,* can be made.

Nevertheless, there is a need to develop systems other than those of *E. coli* which may be transformed by exogenous DNA, and which will also be reliably efficient in expressing the same. A few of the more prominent alternatives are described here, as well as the stalwart standby *E. coli.*

2.8.1 *Escherichia coli*

Escherichia coli is a gram-negative bacterium with a characteristic bacterial cell wall (Fig. 2.64). This species has been the 'guinea-pig' for almost all recombinant DNA manipulations. It is not as easily transformable as *Hemophilia* species or *Salmonella typhimurium,* also gram-negative bacteria. However, today, methods are available for making the cells transformation competent, prominent among them being the calcium chloride treatment first proposed by Mandel and Higa (1970).

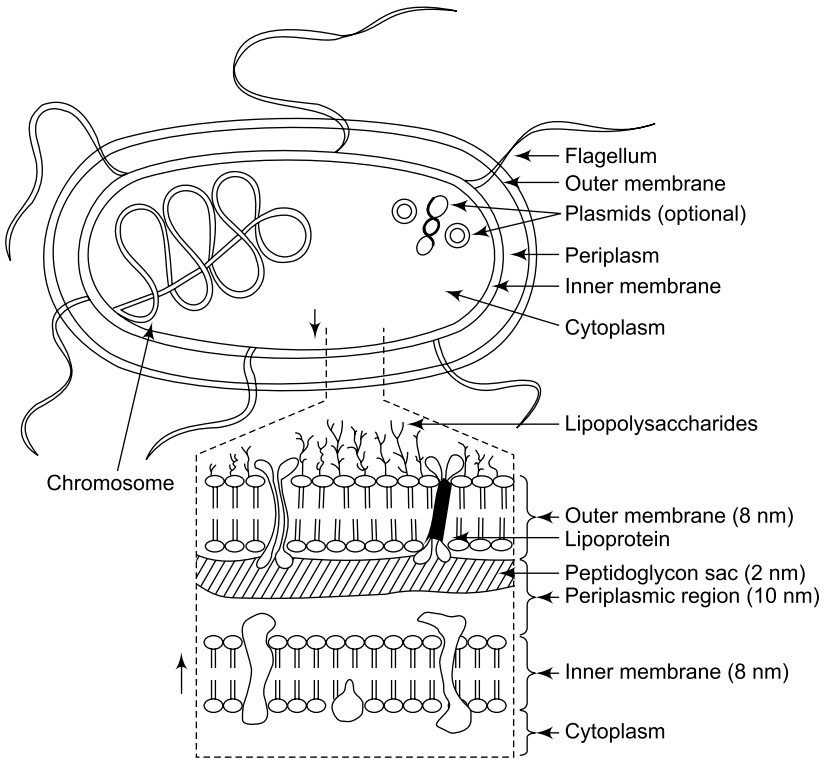


Fig. 2.64 *The E. coli cell; a diagrammatic representation. E. coli possesses two cell walls with a periplasmic region between them. A sac of peptidoglycan lines the outer wall of the periplasm. The E. coli chromosome is a highly supercoiled naked, circularly closed ds DNA that lies in the cytoplasm. Extrachromosomal DNA, as plasmids, may be present in some strains. Flagella are involved in locomotion of the cell. Cells with the conjugal F plasmid (in the free form) as F^* cells or (integrated in chromosomal DNA) as Hfr (high frequency of recombination) cells possess extensions of the plasma membrane called pili (sing. pilus) which are utilized for the process of conjugation.*

Plasmid, phage and cosmid vectors described earlier are for use in *E. coli.* Several shuttle and expression vectors are also available that allow cloning in *E. coli* and in another species (e.g., *E. coli*-yeast, *E. coli-B. subtilis; E. coli*-mammalian cell). With suitable modifications, several commercially valuable proteins have been expressed from transformed *E. coli* cells.

2.8.2 *Bacillus subtilis*

Bacillus subtilis is a gram-negative bacterium that has proved to be very difficult to transform. There has still been a continued effort to make it a suitable cloning cell, as several advantages can be perceived in its use.

The advantages of cloning in *B. subtilis* include the following: (i) *Bacillus* species are utilized in a large number of fermentation industries. (ii) Their species are absolutely non-pathogenic. (iii) Most species secrete proteins. This secretory quality, if appended to the introduced recombinant DNA, would allow the desired protein to be secreted as well. It is easier to collect the secreted protein than those which are intercalary. (iv) Many of the protein secretors are obligate aerobes, and so are easier to culture than anaerobes or facultative anaerobes. (v) Geneticists have an added interest in *Bacillus* species as these sporulate and thus offer one more opportunity to investigate the molecular basis of differentiation and development.

The difficulties of cloning in *Bacillus* species are many and have been deterrents in their use in the past. Today, with a better adeptness in rDNA techniques, a renewed effort is being made to find ways of utilizing these species for cloning.

One of the difficulties has been the highly proteolytic activities of the proteins secreted by most species of *Bacillus.* Mutants that have low levels of protease activity have been identified. Even in these mutants, the cells with intact walls are not transformable, but the cell wall-denuded protoplasts (spheroplasts) are.

Most attempts of transformation with plasmids failed, although fragments of chromosomal DNA could be introduced into *Bacillus* cells with greater ease. On investigation of the reasons underlying this failure, it was discovered that monomeric plasmids found it difficult to enter the cell, but not multimeric ones. If several monomers were ligated to form a multimeric circular DNA, *Bacillus* cells could be transformed. Further investigations revealed that if certain regions of the monomer were present in the vector as a repeated sequence (260–2000 bp), transformation became possible (Michel et al., 1980). Difficulties in transformation with plasmids led to a search for other avenues of insertion of exogenous DNA into *Bacillus* species. One such endeavour is the use of the technique of plasmid rescue described earlier. The donor DNA in this case is spliced to a plasmid which has a sequence that is homologous to another plasmid within the cloning cell. The recombinant plasmid is cleaved within the region of homology and introduced into the cell. The free ends align with the homologous sequence in the endogenous plasmid and induce recombination. This transfers the donor insert into the resident plasmid. Several obstacles to plasmid transformation are averted by this technique. Of course, when using the plasmid rescue system, the resident and incoming plasmids have to be separated at the end of the show.

One of the selectable markers that is widely employed in the *Bacillus* system is *erm* c that confers erythromycin resistance. This marker is not as useful in *E. coli* as the latter is *highly* resistant to even medium high levels of the antibiotic. Some sensitive strains have been identified (e.g., DB11) that partially solve this problem, when a *B. subtilis-E. coli* shuttle vector is employed. If the rDNA contains the insert within the *erm* c locus, the transformants (erm^-) may be recognized in *E. coli.*

Vectors presented another deterrent for *Bacillus* transformation. The native ones of the species have no detectable marker genes, while those of *E. coli* are not replicable in the *Bacillus* species, nor are their markers expressed in the latter. This obstacle has been circumvented by using vectors made by joining certain plasmids from *Staphylococcus aureus* that have a larger host range than that of the *Bacillus* plasmids. Ehrlich (1978) created a vector by joining parts of a plasmid pC194 with a fragment (*Hin* dIII) from another one pT127, that possessed two marker genes: Tet^R and Cm^R (chloramphenicol). The donor DNA may be cloned at the *Hin* dIII site. Two shuttle plasmid vectors were developed by joining the above plasmid with pBR322—both cleaved at the *Hin* dIII site (pAV14 and pAV15 differing only in orientation of the pC194 segment). Other useful vectors have been developed from these constructs. One of them carries a direct selectable marker, the thy^+ (thymidine synthesis) gene. The thy^+-carrying plasmid (also carrying the Cm^R marker) has to be in a thy^- strain of *B. subtilis* (BD 393). The recombinant DNA is constructed by inserting within the thy^+ gene sequences in the plasmid (pBD214). A thy^- *B. subtilis* when transformed with the rDNA remains thy^-. The latter are resistant to the drug trimethoprim (Tp). A recombinant cell would be Cm^R and Tp^R; the untransformed BD393 are Cm^s.

One unique difficulty raised by *Bacillus* species is their inability to express genes from other species. Is the difficulty at the stage of transcription? Moran *et al.* (1982) have sequenced *E. coli* and *B. subtilis* promoters and have not discovered any significant difference between them. Attention was, therefore, given to the translational phase. It appears that there is a ribosome-binding sequence in *B. subtilis* mRNAs that has a greater affinity for the 16S rRNA of the small ribosomal unit of *B. subtilis* than do the Shine-Dalgarno elements of *E. coli* or phage rRNAs. This is perhaps then the reason for the failure of *Bacillus subtilis* in expressing *non-Bacillus* genes.

To overcome the last-mentioned snag, some expression vectors have been specially tailored. One of them (pL603) carries a promoter-less CAT gene from *B. subtilis.* It carries a kan^R marker as well. Fragments of Bacillus species may be inserted at a cloning site (*Eco* RI) just upstream of the CAT sequence. The CAT gene will be expressed if the inserted segment contains a promoter. Further, the expression for the CAT phenotype may be determined by treating the cells with chloramphenicol. If CAT is expressed, the cells will be Cm^R; cells with unexpressed CAT will be Cm^s.

An expression system that may be controlled with an inducer has been devised for use in *B. subtilis.* The penicillinase gene of *B. subtilis* was fitted with the operator sequences of the *E. coli lac* operon, on the 3′ side of the penicillinase gene promoter. The gene for *lacI* (the repressor) was also spliced to this vector next to a *Bacillus* promoter. The penicillinase gene could now be induced by subjecting the cell to the *lac* operon inducer IPTG (analogue of lactose).

Perhaps the greatest utility of creating efficient cloning systems in *Bacillus* species lies in the production of secreted proteins. *Bacillus* proteins, like secreted proteins of other species, possess the short hydrophilic peptide sequence (signal peptide) that allows the passage of the protein through the ER membrane. The cloned gene may be provided with a signal peptide sequence from a *Bacillus* gene; this ensures efficient secretion of the product of the cloned gene. As the secreted proteins, obtained so far, have also been found to be functional, it has been presumed that post-translational modifications (such as correct protein folding) must also be carried out correctly within the *Bacillus* cell.

2.8.3 *Streptomyces* Species

Streptomyces are gram-positive bacteria that produce more than 60 per cent of the known antibiotics. There is a need to be able to manipulate their genome for improving their drug-producing ability and to engineer them to synthesize antibiotics with novel features.

To be able to engineer *Streptomyces* species genetically, it is necessary to know the biology and gene regulatory systems in these soil microorganisms. Until now improvements in performance have been brought about by traditional techniques of mutation breeding, without recourse to understanding its genetics clearly. Gene cloning techniques, however, require this knowledge for any experimental intervention to be meaningful.

We owe much of our knowledge about the biology of *Streptomyces* to the investigations in Hopwoods' laboratory for almost three decades (Hopwood et al., 1973, 1983). The first report of a cloned *Streptomyces* gene came from the laboratory of Cohen (Bibb, Schottel and Cohen, 1980).

S. coelicolor, the best studied species, possesses a double-stranded circular DNA of around 10 kb. The latter has been mapped; several loci for genes that confer resistance to drugs, for enzymes and for biosynthetic and fermentation pathways have been identified. The detailed steps in the biosynthesis of antibiotics are not clear or known. Entire sets of genes involved in a pathway have, therefore, been the target donor DNA for cloning. Malportida and Hopwood (1984) have cloned the set required for the synthesis of actinorhodin, (from *S. coelicolor*) in cells of *S. parvulus.*

One of the most difficult tasks for cloning in *Streptomyces* has been to identify strains that lack, or are weak in their, restriction systems. Almost all strains degrade phage and heterologous DNA. These are the strains preferred by biotechnologists since they are the strong producers of the pharmaceutically important products. One species that is restriction deficient is *S. ambofaciens.* The restriction activities of other strains may be reduced or nullified by heat treatment or by mutating the gene(s) for the enzyme(s).

A cloning system should not alter the recombinant DNA. In some systems, the latter is unstable as deletions and other aberrations occur due to the recombination enzymes in the cloning cell. There is still a dearth of a strain of *Streptomyces* that is recombination deficient. It may take a while to smoothen out such wrinkles in creating a useful cloning system in these bacteria. In the meantime, the complex biology of these species is becoming unveiled with each new investigation.

There are several plasmids of *Streptomyces* species that have already been 'recruited' as cloning vectors. Some are of the low and some of the high copy number varieties. DNA cloned on high copy number plasmids have not been useful, perhaps because the excess gene product has proved toxic to the cloning cell.

One plasmid recovered from *S. lividans* is pIJ101, which is autotransmissible, has a very wide host range and occurs as 40–60 copies per cell. Several vectors, smaller than the original pIJ101 (9000 bp), have been constructed.

Marker genes for vectors for *Streptomyces* include several drug resistance genes. The best one is for resistance to thiostrepton, to which most *Streptomyces* species are sensitive. Another useful marker is a tyrosinase gene, which when intact, can identify cells containing it as black in colour in a tryptone and copper-containing medium. If the donor DNA is cloned within this gene, the inactivation is reflected in colourless colonies on this special medium.

One of the bacteriophages that infect *Streptomyces* has been used as a vector. This phage C31 has a 41 kb DNA (of which 7 kb may be replaced by an insert) and, like the phage of *E. coli* can lysogenize its host cell. Hybrid vectors have been constructed by splicing the DNA of this phage with pBR322. They replicate well in *E. coli* and in *Streptomyces* species.

Only protoplasts of *Streptomyces* are transformed. The cell walls are removed by a glycine treatment followed by one with lysozyme. Protoplasts mixed with PEG and the plasmids take up the exogenous vector. If the vector has no selectable marker, a procedure known as 'pock' marking is employed to identify transformed protoplasts. The technique consists of spreading the mixture containing regenerating protoplasts, with and without plasmids, on a solidified medium. The cells with plasmids cause the formation of clear zones wherever they come into contact with untransformed cells. The bacterial lawn assumes a pock-marked appearance. Vectors with selectable markers are used in the usual manner for identifying recombinant clones.

2.8.4 *Saccharomyces cerevisiae*

Yeasts are unicellular fungi that exist both as haploids and diploids. The haploids are of mating types *a* and α. Mating between an α and an *a* cell results in a diploid cell. The latter undergoes meiosis to form two each of spore types *a* and α which germinate into the corresponding mating type strains.

Saccharomyces cerevisiae, one of the yeasts, is the best studied species (Fig. 2.65). It is of great value as it forms the basis for a very wide spectrum of technologies that depend on fermentation. This importance has led to the emphasis on investigations of the biology and genetics of *S. cerevisiae* and other species of yeasts. Recombinant DNA techniques have increased exponentially the knowledge about yeasts acquired earlier by traditional mating and mutation studies.

The nature of control of the genes in *S. cerevisiae* is of prime interest to those who would like to exploit yeasts for more efficient production of commercially valuable products. Studies of cloned genes were, therefore, initiated.

Hinnen *et al.* (1978) were the first to transform yeast with DNA using the $CaCl_2$ PEG method. The foreign DNA was found to be integrated into the yeast chromosome. This was the beginning of the construction of yeast integrating vectors.

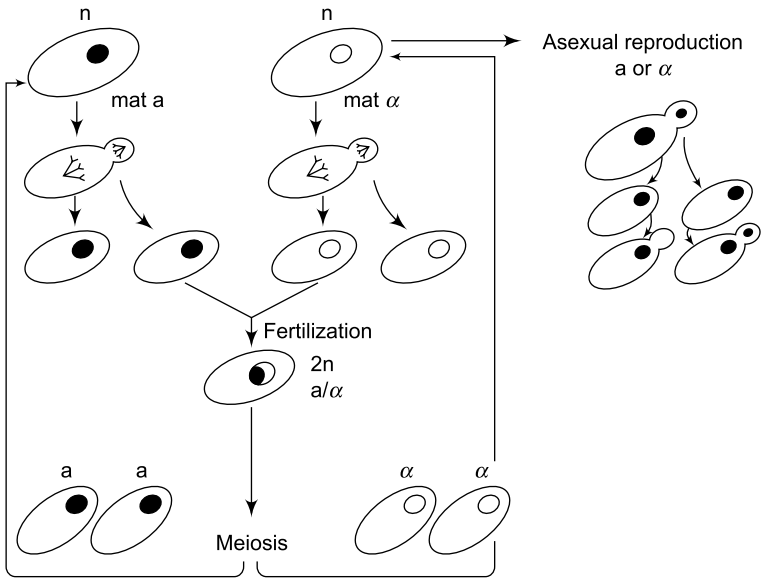


Fig. 2.65 *The yeast Saccharomyces cerevisiae. A schematic representation of the sexual and asexual reproductive cycles. S. cerevisiae exists as two haploid (n) mating strains a and α (alpha). They fuse to form diploid (2n) cells with an a α genotype. Meiosis in the zygote (2n) results in recombination between homologous alleles of the two parent strains, and the production of 4 haploid spores (environment-proof cells) that germinate into a and α type haploid vegetative cells. Yeasts are eukaryotes which can be cultivated and manipulated like bacterial cells. They are useful for studying eukaryotic molecular biology in general and yeast biology in particular. As yeast represents one of the most commercially utilized species, there is a special need to understand its genetics and developmental features in molecular terms.*

To transform yeast cells, the characteristic fungal cell wall has to be removed, and the spheroplast (protoplast) mixed with the donor DNA.

At first, yeast DNA was cloned in *E. coli,* using *E. coli* vectors. About one-third of yeast genes were found to complement mutations in *E. coli* genes. This fact was employed to identify yeast genes in *E. coli* and, in turn, to transform yeast cells. The first series of shuttle vectors thus came into existence. These vectors can be replicated in *E. coli* and yeast, and may be selected in both hosts due to the presence of *E. coli* and yeast selectable markers.

Eukaryotic genes have been introduced successfully into bacteria and the various steps required for DNA engineering have been carried out in them. But to understand the correct mode of expression and/or regulation of expression of a eukaryotic gene, the correct environment should be the interior of a eukaryotic cell. The first choice for a cloning cell among eukaryotes was that of *S. cerevisiae.* However, very soon, it became evident that many eukaryotic genes are not expressed efficiently, or not expressed at all, in yeast cells. If the cloned gene is a split one, the yeast-splicing system fails to splice the exogenous gene correctly. There also appear to be differences in transcription initiating and translation signals between those in yeast and those in other eukaryotes (mammals and human beings).

Attempts have been made to decipher why the yeast system does not do justice to other eukaryotic genes. In the meantime, yeast cells are being used extensively to fathom the nature and behaviour of the yeast system itself. This understanding is so valuable that even if the yeast cell remains appropriate only for cloning yeast DNA, it will retain its utility for a long time to come.

Other species of yeasts are being investigated for use as cloning and expression systems.

2.8.5 In *Xenopus* Oocytes

Isolated eukaryotic genes may be expressed in frog (*Xenopus* sp.) oocytes. These cells are large enough for the DNA to be microinjected directly into them.

Oocytes are extremely rich in the transcribing enzymes (the RNA polymerases) that are specific for eukaryotic genes, as well as in the factors required for post-transcriptional and post-translational modifications of the gene product. One of the main reasons for the lack of function of an eukaryotic gene in the mileu of a prokaryotic cytoplasm is the inability of the latter to put the finishing touches to the RNA transcript and/or to the translated polypeptide that make the gene products functional. Frog oocytes offer useful cloning vessels for the expression of eukaryotic genes. These cells are particularly handy for studies of expression of genes that have been mutated by design.

The frog oocyte contains a much greater quantity of the RNA polymerases than ordinary somatic cells (easily 50–100 K) times. Eukaryotic RNA polymerases are of three kinds: I, II, and III. Each of these enzymes is made of 6–10 polypeptide subunits. RNA Pol I (DNA-dependent RNA polymerase) transcribes ribosomal DNA into the 18S, and 28S 5.85S transcripts. RNA Pol II copies the structural genes into hnRNA (heterogeneous RNA, which is processed into the functional mature RNA transcripts). RNA Pol III takes care of the genes of the tRNAs (4S), of those for the 5S rRNA, and perhaps of other small RNAs such as the snRNA molecules involved in splicing of some RNA.

Genes may be microinjected into the oocyte by either of the two following methods: that of Gurdon, where the gene is pushed into the nucleus in the centre of the cell, or that of Kressman's group, which first brings the nucleus near the cell plasma membrane by mild spinning. Injecting the foreign DNA directly into the nucleus avoids degradation of the DNA by cytoplasmic DNases. Modifications have been devised for the prevention of degradation of linear DNA, which is more vulnerable to nuclease action than circularly closed DNA.

It is interesting to know that when a circular duplex DNA finds itself in the oocyte nucleus, the DNA promptly becomes organized into a chromatin, by wrapping itself characteristically around nucleosome cores, the components of which abound in the nucleus.

The genes that have been studied best by expression in frog oocytes are those of the 5S rRNAs. Brown and Gurdon demonstrated that even plasmid-cloned 5S RNA genes injected into *Xenopus* egg cells were transcribed into correct transcripts by RNA Pol III.

Most eukaryotic genes have been successfully expressed in *Xenopus* oocytes. Notable exceptions are those of human mitDNA and bacterial genes (*E. coli*). Expressions in oocytes have elucidated various facts about the process of transcription by the different polymerases. The three RNA Pols can be distinguished by their sensitivity to a fungal toxin called 'amanitin'. So, it is possible to determine which type of RNA polymerase is involved in the transcription of a particular type of gene. When the homogenates from nuclei are used for *in vitro* transcription of 5S and adenovirus associated 5.5S and 5.2S RNA genes, the proper transcripts are synthesized. If, however, only the particular RNA Pol and the DNA template are used for *in vitro* transcription (cell free), there is no synthesis of the RNA. This difference indicates the presence, in oocytes, of factors other than the enzyme and recognition signals on the template DNA that are essential for proper transcription. The *Xenopus* oocyte offers an excellent system for elucidating the finer details of the requirements for such gene expression activities.

2.8.6 In Mammalian Fertilized Egg Cell

The expressions of eukaryotic genes are under temporal and spatial control. Only a relatively small set of the entire genome is expressed in any one type of tissue and that too, not at every stage of development. Gene expression in eukaryotes is thus tissue-specific. It is also regulated in terms of the amount of gene-product synthesized in the same tissue at different times, as well as in different tissues in which it is expressed.

Introduction of a foreign gene into a random cell in an eukaryote cannot, therefore, guarantee the expression of the same. Earlier studies revealed that transfection of foreign DNA could be demonstrated in the recipient cell, but not necessarily the expression of the gene product. This difficulty may be avoided by introducing the foreign gene as early as possible in the developmental plan of the organism, hoping that the exogenous DNA would become stably integrated into the host cell DNA and be perpetuated with the latter, and eventually expressed in the proper place and time. Such a strategy has been developed using the fertilized egg of the mouse as the recipient cell. Methods of handling mice eggs and their implantation into the uterus of a surrogate mouse, to allow development of the egg to the full term, have improved considerably, and facilitated investigations of gene transfer and expression in mammalian species.

The foreign DNA, usually in the form of tandem repeats, that range in size from 0.1 kb to 30 kb, are microinjected directly into the male or larger of the two nuclei (pronuclei) in the fertilized egg, before the former fuse into a single diploid zygote nucleus. The integration of the introduced DNA has been demonstrated, and stable transmission into germ line as well as somatic chromosomes confirmed.

Does the inserted DNA cause any disturbance in the transfected animal? This question is difficult to answer. It would seem reasonable to suppose that those which derail normal functioning in the recipient individual do not come to our notice since these animals do not develop. So, animals with the foreign gene are those that have been without selection pressure and survived in the new genetic background.

The tissue-specific expression of inserted foreign genes in mice have been demonstrated by the insertion of 'fused genes' via the genome of a retrovirus. The fate of the introduced gene is assessed in all types of cells in the adult mouse by Southern hybridization of chromosomes *in situ* and by screening for the mRNA or protein product of the foreign gene.

A fused gene is hybrid DNA with the regulatory sequence and sometimes a part of a host gene preceding the foreign gene. The fused gene is carried on a plasmid vector. The regulatory sequences that have been utilized are those of a gene, the product of which complexes with certain heavy metals (Cu, Zn) and sequesters the latter for use or for scavenging (Hg). Cadmium acts as an inducer for the expression of such a *metallothionein* (*MT*) *gene.* Mouse eggs in which the foreign gene is introduced are allowed to develop in a tissue culture medium. They are treated with cadmium and evidence for the expression of the attached gene sought in them. The inserted genes that have been used first are the ones for thymidine kinase (TK) from the Herpes simplex virus (HSV) and those of the rat growth hormone (GH). The viral TK and mouse TK genes (indigenous to the recipient) can be distinguished easily. Expression of the inserted TK gene was found to be induced by the inducer (Cd) that binds itself to the upstream regulatory sequences of the MT gene.

Similarly, the growth hormone was expressed in the recipient animal in a fairly large number of subjects. Mice developed from treated eggs were larger than normal, showed up the introduced GH DNA in Southern blots made from somatic cells and also were found to synthesize more than normal amounts of the hormone. The plasmids carrying the MT (metallothionein kinase) and MGH (metallothionein-growth hormone) fused genes are usually introduced into the mouse pronuclei in a linear form, as the latter integrate into the host DNA with greater guarantee than the circular exogenous DNA.

From the foregoing experiments it became apparent that the fused gene expressed itself best, on provocation with Cd, in those tissues where the MT gene is normally expressed maximally. Tissue-specific regulation of gene expression was further confirmed by these exercises.

Insertion of the genetic information of a tumour virus (murine or mouse leukemic virus, MuLV) into the fertilized egg of mouse also demonstrated that an inserted gene is expressed only when it is in the correct chromosomal location (possibly vis-à-vis its neighbouring genes) and in the proper state of methylation. If the ds DNA copy (proviral DNA) of the retroviral ss RNA genome is introduced at a time when the cells are actively methylating new DNA, the inserted DNA also becomes methylated. However, such highly methylated DNA is not expressed. If, on the other hand, the insert DNA arrived at a developmental period or tissue where methylation was not a prime activity, such DNA escaped methylation and were found to be expressed as mRNA and/or the protein gene products.

Insertion of genes in mammalian cells was hailed briefly as an ideal method of 'genetically engineering' a commercially useful vertebrate—especially mammals of economic value. Theoretically, once the techniques of nuclear transplantation and uterine implantation are improved sufficiently to prevent damage to the developing embryo, it should be possible to make genetically identical replicas (clones) of such animals.

The introduction of DNA into mammalian species is a potent tool for investigating the mechanism(s) of gene control and expression in higher eukaryotes. Such studies need not be carried out only in fertilized egg cells which are not as easily available as other somatic cells of the body are. Research has, therefore, taken advantage of another type of mammalian cells, the genetical composition of which is largely specified and which can be manipulated in bulk under artificial tissue culture conditions. These are the various cell line cells, which are strains of 'immortal' cells derived from some primary normal or neoplastic tissue. One such cell line type that has proved to be extremely useful for studies of developmental control of gene expression in eukaryotes are the CHO or the Chinese hamster ovary cell lines (see the next section).

2.8.7 Chinese Hamster Ovary-Cultured Cells

As mentioned earlier, expression of genes in a eukaryotic setting is also studied by using cultured cells (cell line strains) of mammalian species. Rat and mouse cell line strains have mostly been in vogue. The utility of cell lines derived from the ovary of another rodent, the Chinese hamster (*Cricetulus barabensis griseus*) has been realized and exploited. One of the attractions of using Chinese hamster cells is the relatively small number ($2n = 22$) of its chromosomes, which can be readily distinguished individually (10 pairs of autosomes plus X and Y chromosomes).

CHO line cells are very useful for somatic cell genetics studies and those that involve transfer of genes, mutagenesis and DNA repair and short-term assays for genetic toxicity. CHO cells can be cultivated as single cell suspensions in standard media as well as in the form of monolayer cultures. These advantages are further supplemented by the availability of several auxotrophic mutants (mostly derived in the laboratory of Verganian and of Puck and associates by the BrdU-enrichment method). This method involves the treatment of a population of normal and mutated (very low count) cells with 5-bromodeoxyuridine (BrdU), an analogue of thymidine. The cells to be treated are kept in a minimal medium, which supports the cell division and growth of only wild-type cells. An auxotrophic mutant lacks a functional gene that allows the cells to grow on the minimal medium only when the latter is supplemented with the nutritional requirement the gene needs for the synthesis after it has mutated. So, wild-type cells grow and replicate DNA, incorporating BrdU in its place of T. When the population of cells is exposed to sunlight, the DNA of the normal cells become inactivated (BrdU instigates the photolysis of DNA), and these cells die. Only the mutated cells survive, although they do not proliferate. On transferring to a medium supplemented with the deficiency requirement, the surviving cells grow into colonies.

The CHO system has been especially adapted to serve as a selection agent for genes coinserted with HGPRT, APRT and DHFR loci, either by a procedure of whole nuclei, metaphase chromosome or other agent-mediated gene transfer.

REVIEW QUESTIONS

1. Describe how cloning of a gene is carried out.
2. Enumerate the enzymes employed in genetic engineering and briefly explain their utility.
3. Give examples of common plasmid vectors.
4. Describe the bacteriophages extensively used as cloning vehicles.
5. Write an explanatory note on cosmids.
6. How are animal viruses exploited as vectors?
7. Give an account of special vectors.
8. Describe the role of *Drosophila* transposable elements and yeast plasmid as cloning vehicles.
9. Recall natural para-sexual modes of gene transfer.
10. Describe various techniques used to introduce genes into eukaryotes.
11. Enumerate and explain cloning in representative prokaryotic cells.
12. Describe how cloning in eukaryotic cells is carried out.

3 Genetic Engineering Toolkit II: Gene Cloning, DNA Libraries and PCR

Introduction

DNA engineering involves the creation of a DNA sequence of choice. Its tools are the variety of enzymes that cleave, ligate, shorten, lengthen, methylate, phosphorylate or dephosphorylate a DNA. A short length of DNA—an oligonucleotide—may be synthesized to order. They may be utilized to join DNA molecules with incompatible ends, provide a cohesive end to a molecule without a defined end or provide a set of new restriction sites for cloning. These oligonucleotides are the linkers, adaptors and connectors of various kinds. The ends of DNAs may be engineered to specifications merely by a judicious use of polymerases and nucleases and one restriction site replaced by another or erased completely.

3.1 ❑ DNA ENGINEERING

The most significant advance in DNA engineering is, of course, the discovery of ways of causing targeted mutagenesis; single bases or short lengths of DNA may be altered deliberately in a site-specific and sequence-specific manner. It has thus become possible to introduce mutations not only into resident genes and regulatory sequences but also into specifically introduced synthesized oligonucleotides in a specific target region. The potential of this manoeuvrability is legion in terms of research and application.

3.1.1 Oligonucleotide Synthesis

Short lengths of DNA and RNA may be synthesized by very sophisticated methods of organic chemistry. The chief ones are the phosphodiester and phosphotriester methods of synthesis.

In essence, synthesizing an oligodeoxynucleotide consists of joining two nucleotides in which all reactive groups except the ones to take part in the formation of the interpolynucleotide bond are blocked or protected. This means the base and the 5′ phosphate group of one and the base and 3′ hydroxyl group of the other are prevented from undergoing chemical reactions during the duration of the synthetic

process. This leaves the 3′ hydroxyl of one and 5′ phosphate group of the second deoxynucleotide free to form a phosphodiester bond. The 3′ hydroxyl group of the last nucleotide in the elongating chain is then deblocked for the reaction with the next nucleotide. One by one the bases are added in the chosen order. When the required length of the polynucleotide is synthesized, the bases and the hydroxyl groups in the first and last nucleotides (at the free ends of the chain), as well as the bases, are deblocked.

In both methods (given in Fig. 3.1), the hydroxyl groups at the 5′- and 3′-ends of the ribose moiety are made unreactive (protected) by the addition of suitable groups R1 and R2, respectively. In the phosphotriester method, one of the OH groups in

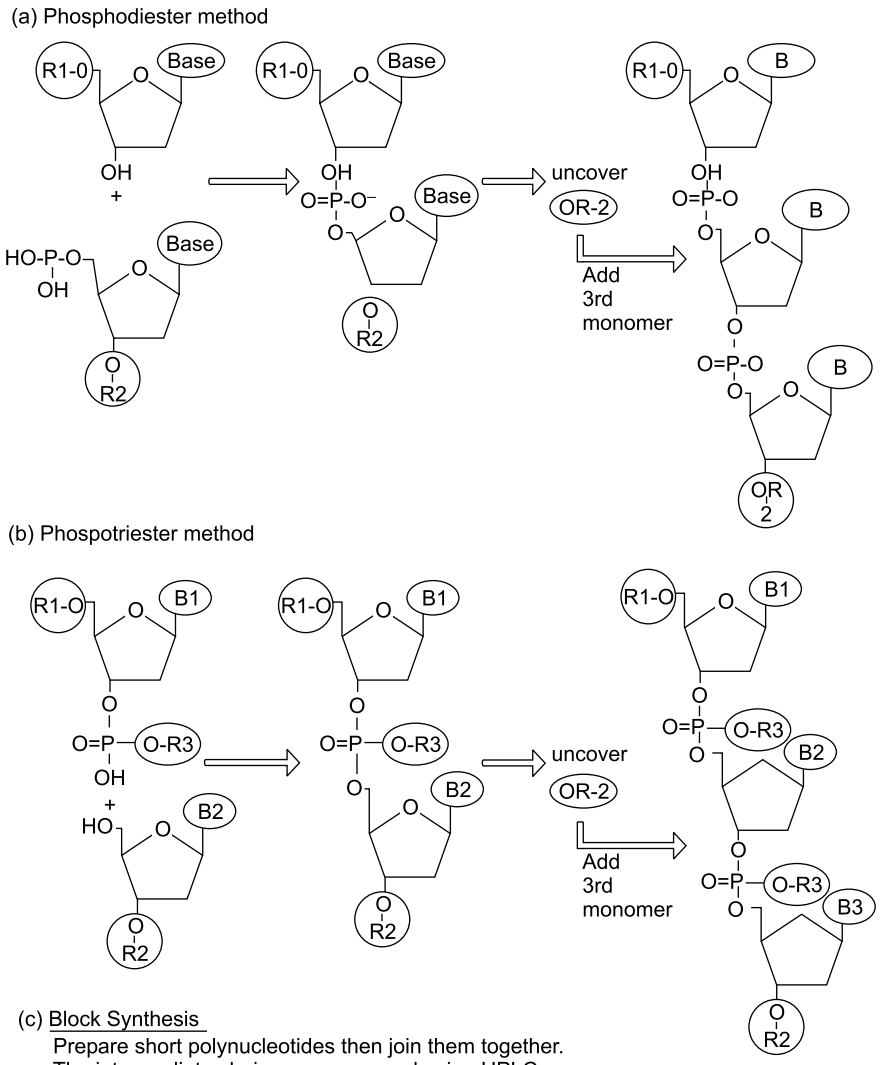


Fig. 3.1 *Oligonucleotide synthesis by the phosphodiester and the phosphotriester methods.*

the phosphate moiety is also protected (by R3). The bases are also protected. The first nucleotide of the oligonucleotide is immobilized on a solid support. The R1 of the second nucleotide is removed and the latter added to the reaction mixture; a phosphodiester bond links the first two nucleotides. The process is continued, unmasking one R1 group of the added nucleotide and the R2 of the last unit on the synthesized chain, and thus extending the chain by one nucleotide. At the end of the job, the bases and the R1 of the first and R2 of the last nucleotides are removed. In the phosphotriester method, the R3 groups are also removed at the end of the reaction.

Several sophisticated improvements have appeared in the procedures for chemical synthesis of oligonucleotides. They are essentially derived from one of the two basic protocols—the phosphodiester and phosphotriester methods. The 5′- and 3′-hydroxyls of the mononucleotides are protected in both methods. The second method also blocks a hydroxyl group in the internucleotide region. Michelson and Todd were the first to initiate chemical synthesis of DNA; they employed the phosphotriester method (1955). This method was improved upon by Khorana *et al.,* (1972, 1976), using the phosphodiester method, to synthesize entire genes, alanine and Sup F (tRNA) of yeast. The principle, simple as it appears, runs into snags in practice, especially due to various solubility and 'recovery of products' problems. Eventually, the triester method has been modified into a routine procedure.

The bases are protected more firmly than the hydroxyl ends of the sugar or phosphate group, as the latter require to be relieved of the block after the addition of each new nucleotide. Khorana *et al.,* introduced the use of N-acyl groups to tie up the nitrogenous base moiety. N-benzoyldeoxyadenosine, N-isobutyryldeoxyguanosine and N-anisoyldeoxycytidine are the usual blocks for the A, G and C bases.

The procedure for building an oligonucleotide chain is to first convert the protected bases into monophosphate esters. The second step consists of presenting the next nucleotide in the chain together with a coupling agent. Narang *et al.* (1979) and others have used a successful variation of this method. This protocol begins with a fully protected phosphotriester added to two nucleotides, to one at the 5′-end and to the other at the 3′-end, by unblocking the appropriate triester 3′- and 5′-blocks. The two esters of the nucleotides are now joined together.

This is not the forum for a detailed account of the chemical synthesis of oligonucleotides. It is sufficient to be aware that such oligonucleotides are now synthesized to order with any desired sequence of bases. They have added greatly to the 'bag of tricks' used for engineering the DNA to be cloned or causing alterations in native or introduced genes.

Khorana employed oligonucleotides to synthesize an entire gene, as mentioned earlier. Basically, the method consists of preparing four or five oligonucleotides that are joined in a given order to form a longer chain. Several long fragments are next linked to form the final chain. The fragments or oligonucleotides are prepared with overlapping sequences; in the case of one for a protein, codon usage is taken into account before deciding on the sequence that would represent the fully functional gene.

The excitement created by oligonucleotide synthesis is not limited to the use that may be made of short known sequences of nucleic acids, but can also be extended

to the possibilities of improving natural genes by synthetic versions which are calculated to possess extranatural useful functions. In short, protein engineering is acquiring an extra dimension.

One fascinating use of a laboratory-created sequence led to the discovery of a variant conformation of the double-helix. Alexander Rich and associates prepared short oligonucleotides which possessed a conformation distinctly different from the B DNA of the Watson and Crick model. Rich's Z DNA was used to raise antibodies, and the latter to probe its existence, if any, in the living system. The antibody identified Z-DNA regions in *Drosophila* species. A combination of synthetic organic chemistry and recombinant DNA technology opens up, as this example illustrates, avenues of investigations that may be unpredictably novel.

3.1.2 Linkers, Adaptors and Connectors

Short lengths of DNA are available as single-stranded chains with different defined sequences. They are called linkers, adaptors or connectors, depending on the use to which they are put.

Linkers usually possess a set of restriction sites, either singly, or in pairs, that are placed symmetrically on two sides of a central axis (Fig. 3.2). A linker is introduced into a vector at a restriction site; the vector now possesses several optional sites for restriction enzymes. The pUC and pMB series vectors possess such linkers. A donor DNA cleaved with almost any restriction enzyme can be cloned into such vectors by using the appropriate cognate enzyme.

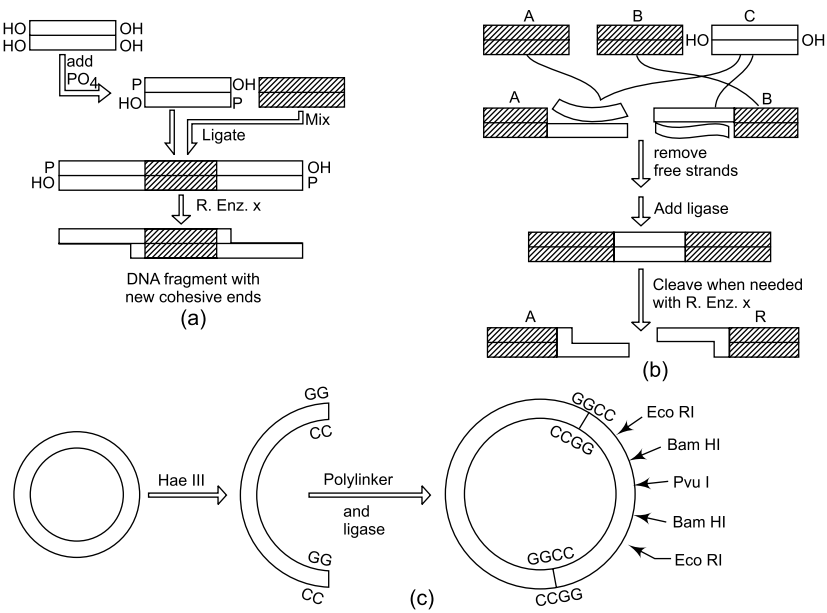


Fig. 3.2 *Linkers. The site for a restriction enzyme may be created by ligating the restriction motif (synthetic oligonucleotide) to the desired fragment and cleaving the motif with the corresponding enzyme. Polylinkers, with one or more sites for different enzymes, are utilized as cloning sites in vectors.*

Adaptors, as the term implies, serve to join DNA strands with incompatible ends. Suppose one cut end has an extended 5′-end and the other an extended 3′-end in the same orientation, an adaptor chain (at least 10 bases long) is used to first become intercalated between the recessed ends in one orientation. This leaves a gap where the two single-strand extensions fail to meet. The gap is usually filled and ligated inside the transformed cell (Fig. 3.3).

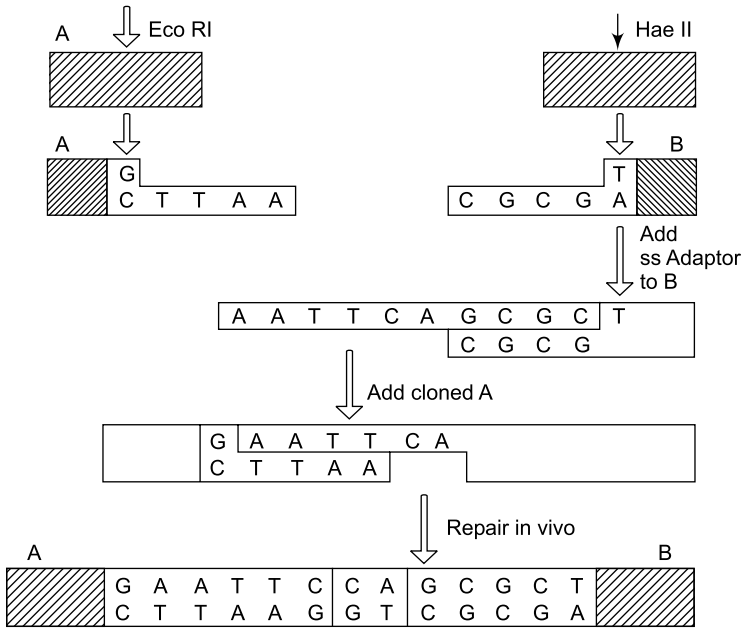


Fig. 3.3 *Adaptors. When two restriction fragments with extensions in the same strand are to be joined, use is made of an adaptor. This is a 10 bp synthetic oligonucleotide that is complementary at its terminal regions with two extensions, which on mixing with the two fragments, binds itself to the complementary tails. There is a gap in one of the strands between the joined fragments, which becomes repaired (filled and ligated) within the cloning cell by the appropriate enzymes.*

Connectors are two adaptors with an internal homology (Fig. 3.4). They are used for joining (connecting) two fragments cleaved by different restriction enzymes. One connector, on ligating to the ends of two strands (in the same orientation), completes the cleavage site at one terminus. The other ss connector restores the cleavage site at the other terminus. The two fragments are now connected with an extra length of DNA (due to the connector) between them. The new DNA may be cleaved by either of the two enzymes, the recognition sites for which have been restored, as well as by a new site that may be present within one of the connector sequences.

Linkers may be utilized for a variety of goals. If a new cohesive end is to be created in a DNA, a linker containing the site for the new enzyme is attached to the DNA. If the latter already possesses an uneven end, it is blunted before ligating to the linker (using T4 ligase). On treating the DNA with the restriction enzyme, the linker is cleaved and the DNA now possesses the new cohesive end.

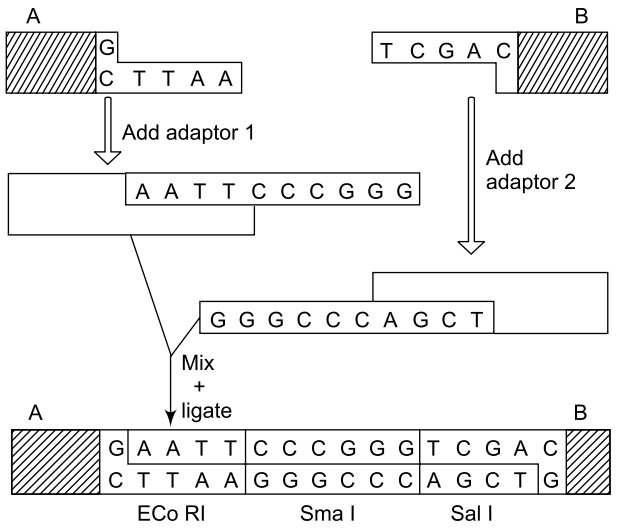


Fig. 3.4 *Connectors. Restriction fragments cut with separate enzymes (RE1 and RE2) may be joined into one DNA by using special adaptors, one end (5′) of which is complementary to the extension due to RE1, and the other contains three cyctosine nucleotides. A second adaptor possesses three G nucleotides at its 5′-end, while the remaining region is complementary to the extension in fragment 2 due to RE2. When the fragments and ss adaptors are mixed, together with ligase, they are joined into one DNA, with an extra three C-G base pairs between them. The restriction motifs for RE1 and RE2 are restored, and, as shown in the diagram, a site for a third enzyme (Sma I) is created.*

The linker may be unphosphorylated or phosphorylated. When it is unphosphorylated, one of the strands of the linker will not be ligated. The free linker strand may be removed before using the DNA for cell transformation. The gap will be filled and the new strand ligated correctly *in vivo.*

3.1.3 To Restructure Restriction Sites and Ends of DNA

If a DNA with uneven ends is to be joined to flush-ended DNA, the former may be made flush-ended and the two blunt ends ligated using T4 ligase. The blunt ending may be achieved either by repair synthesis of the recessed strand using the Klenow fragment and NTPs, or by the extending strand being removed by S1 nuclease.

If a DNA is to be shortened two strategies may be employed. (Fig. 3.5)

(i) The DNA may be treated with an exonuclease, that degrades only one strand, until the required length is reached. The remaining strand is removed with S1.

(ii) The enzyme *Bal* 31 may be used; it degrades both strands simultaneously starting from the free end of a molecule.

If it is necessary to create a nick (break in the phosphodiester bond) in only one strand, DNAse I is used in the presence of ethidium bromide. If a gap is required in one strand, first, a nick is generated. The gap is created by degrading the nicked strand using the exonuclease that will operate in the desired direction (exo III or λ exo) (Fig. 3.6).

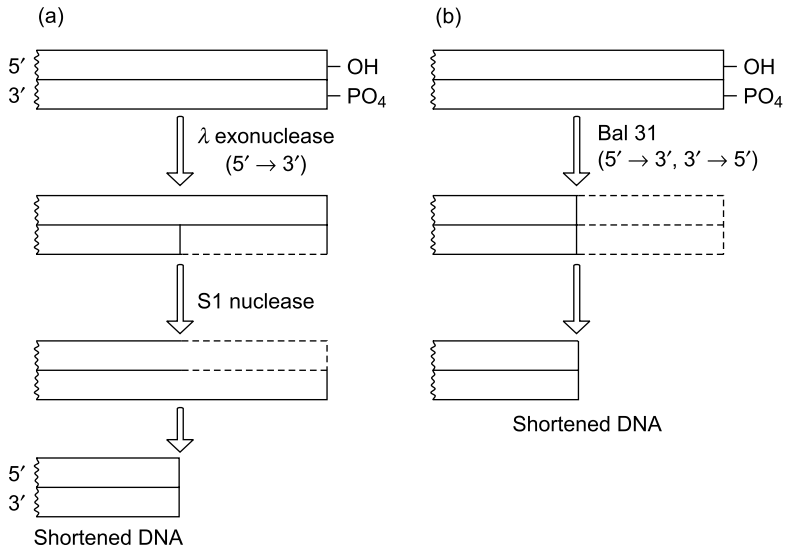


Fig. 3.5 *Shortening of a DNA. Exonucleases are available that degrade single strands from either or both ends. The λ exonuclease shortens a single strand in a ds DNA in the 5′ → 3′ direction. S1 nuclease removes a free ss DNA, while Bal 31 can remove both strands simultaneously. The reactions are continued until the desired length has been degraded.*

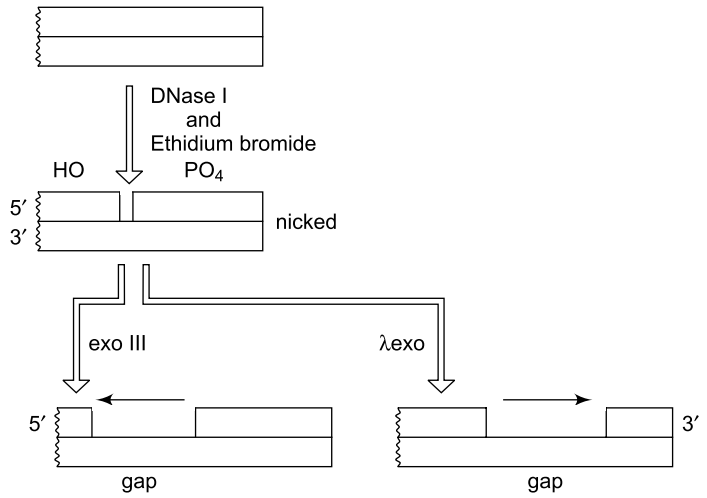


Fig. 3.6 *To create nicks and gaps in a DNA. A nick (break in the DNA backbone) can be created by using DNase I and ethidium bromide. A gap may be created starting from the 3′- or 5′-free end in the nicked region, using Exo III or λ exonuclease, respectively.*

The restriction site for one enzyme may be replaced by that for another. This will be illustrated with an example, where the site for *Xba* I is replaced by that for *Taq* I (Fig. 3.7).

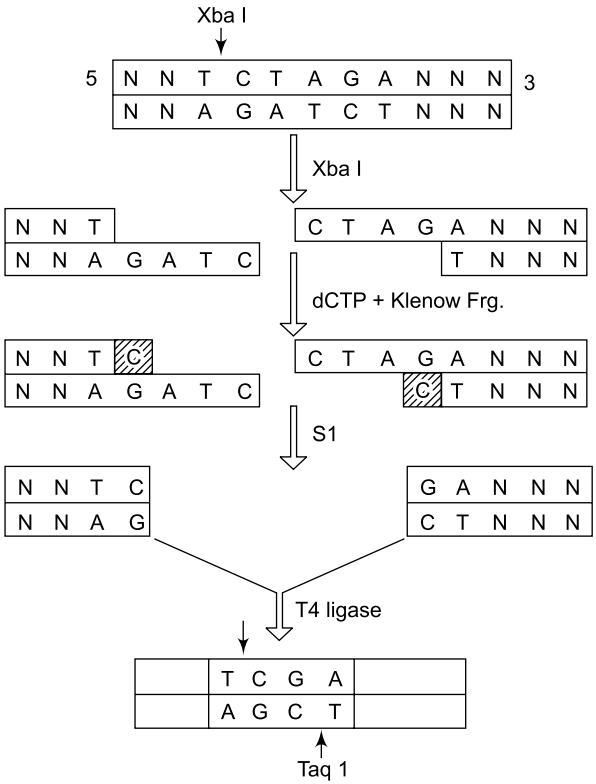


Fig. 3.7 *To replace one restriction enzyme motif by that of another. In this illustration the site for Xba I is replaced by one for Taq I. After cleaving a DNA with Xba I, a cytosine nucleotide is introduced at the 5′-cut ends, with the help of the Klenow fragment and dCTP. This operation is followed by trimming of the protruding ss DNA regions with S1 nuclease. The blunt-ended fragments are then joined with ligase. As a result, the third and fourth nucleotides of the Xba I motif (TCTAGA) are removed leaving TCGA which is the recognition motif for the enzyme Taq I.*

Restriction sites may be re-created on both sides of an inserted DNA by repair synthesis of the cleaved ends after cutting with an enzyme (say, *Eco* R1, G/AATTC), followed by ligation of the insert DNA (flush-ended, with T4 ligase). The insert can be removed by using the original enzyme; the cuts at the two ends will neatly remove the insert when needed.

Two DNAs cleaved with different restriction enzymes may be connected using homopolymer linkages (Fig. 3.8). The linker or insert DNA is provided with tails of only one type of nucleotide (say, C). The ends to be joined are provided with tails having complementary bases. The linker is sandwiched between the two DNAs by hybridization of the complementary tails. Gaps in the two strands are repaired *in vivo*.

There are endless ways in which one can manipulate the sequences within or at the ends of DNAs, and ingenuity shall discover other novel methods for the unique jobs at hand.

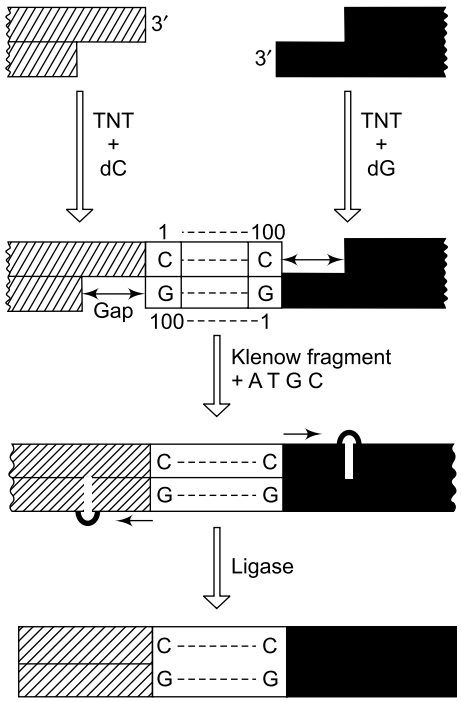


Fig. 3.8 *Restriction fragment with non-complementary ends may be joined with the help of homopolymer tails. The cleaved 3′-end of one strand may be extended into a poly C (or poly A) strand (about 100 b long) with the help of the DNA polymerase TNT (terminal nucleotidyl transferase). The 3′-end of the second strand is similarly extended into a poly G (or poly T) tail. When the two fragments are mixed together, complementary pairing between the tails joins the two fragments. The gaps in the two strands are filled, using the four nucleotides and the Klenow fragment. The nicks are repaired with ligase.*

3.2 ❑ HOW TO ISOLATE A GENE

Various techniques for isolating and purifying genes have been mentioned in the earlier sections. The two more prominent classes will be summarized here.

All techniques for isolating genes are based on gene transfer exercises. The earliest attempt to identify a portion of the genome carrying a specific gene was to transform *E. coli* cells deficient in the function of that gene with fragments of the total DNA harbouring the gene in one of them. The fragment which, on entering the mutant cell, complements the host cell mutation is judged to carry the desired gene. It remains to amplify this fragment, isolate it, check it for the presence of the gene sequences or its products by an appropriate probe.

The next advance came when pieces of yeast DNA were found to complement auxotrophic mutations in *E. coli*. Virus DNA, such as that from adenoviruses was transfected into mammalian cells and those cells that exhibited tumourous transformation were isolated and their DNA probed to identify the viral sequences

integrated in the host DNA. The gene for thymidine kinase was isolated in this way from Herpes simplex virus DNA. These results were obtained purely by traditional genetics-cum-biochemical methods.

For more complicated organisms, the simple techniques mentioned earlier were not adequate for purifying genes of interest. Techniques were developed, therefore, for the complementation of functions in mammalian cells.

The basic principle made use of in the techniques for isolating genes from mammalian cells is to tag the suspect fragments with a known DNA sequence that can be identified with a probe. The nature of the linked DNA differs in different protocols that have emerged. Some of them are described in the following subsections.

In all the techniques, the DNA to be searched for the desired gene is compartmentalized into clones (gene library) made in plasmid, phage or cosmid DNA. Each clone is used to transform recipient mammalian cells, and the transformants chosen with an appropriate selection system. These first sets or primary transformants are usually found to have taken in more than one fragment of the exogenous DNA. To reduce the number of extraneous fragments, the DNA of primary transformants is extracted, Southern blotted and probed for the presence of the carrier or tag DNA (plasmid, phage vector). The positively identified bands are matched with the original fractionated DNA in the gel and the relevant DNA eluted and utilized for a second round of transformation. Many more secondary transformants will be present than found in the primary round. The DNA from these may be probed. The identified fragments are then analysed. The fragment is generally found to contain, besides the tagged DNA, a part of or the entire length of the desired gene. The latter may be utilized to look for the gene, *in situ,* if so desired.

3.2.1 Plasmid and Marker Rescue Techniques

Plasmid and marker rescue techniques are essentially similar. In both, the DNA is cloned as a library into a plasmid, usually pBR322 (or a cosmid). The clones are used to select primary transformants. The DNA from these is used to obtain secondary transformants which may possess the entire pBR322 (if that were the vector) or only parts of it. In the plasmid rescue method, the clones that are positive for pBR322 are isolated and their DNA extracted and cleaved with an enzyme that does not cleave within the gene or between the gene and the plasmid DNA (ascertained before the protocol is followed). The cleaved fragments are circularized and used to transform *E. coli* cells. Those that are ampicillin-resistant (due to Amp^R on pBR322) are allowed to grow into large colonies, and DNA extracted from them to transform once more the mammalian cell. Perucho utilized this method to isolate the *TK* gene from chicken DNA. The cells were selected on a HAT medium for the tk^+ transformants. On sequencing the DNA of transformed cells, the entire *TK* gene was found with the pBR322 sequences.

In the marker rescue scheme, the DNA of the secondary transformants is cloned into a phage library, and each clone searched with a pBR322 probe. The ones that are positive for the probe are used to transform or complement a host defect (mutation) in the gene being searched. Pellicer *et al.* (1980) employed this strategy to isolate the

APRT gene from hamster DNA. The *aprt*$^-$ cells that became *aprt*$^+$ were selected on a special medium. The DNA of these cells contained the *APRT* gene together with a small region only of the pBR322.

A somewhat similar strategy was used to isolate oncogenes. In this case the tag for the chromosomal DNA was a plasmid carrying the *SupF* (suppressor tRNA that suppresses amber mutations) gene. Transformation of hosts with *SupF* was reflected in the ability to grow on the amber mutation-carrying hosts. The DNA of the secondary transformants was made into a phage library, with the amber mutation included in one arm of the phage. When the phages were plated on *su E. coli,* only *SupF*-carrying phages could grow. The DNA from the latter was used to transform rodent cells in culture. Those that acquired the characteristics of oncogenically transformed cells were selected. The DNA of the latter contained the oncogenes c-Ha-*ras* (Goldfarb *et al.* 1982) and N-*ras* (Shimizu *et al.,* 1983).

3.2.2 Transposon Tagging

Isolating a eukaryotic gene by transposon tagging involves inserting transposons into target DNA, fractionating fragments of the latter on Southern blots and hybridizing the transposon with a probe. Cleaving out the transposon, together with flanking DNA, has revealed adjoining sequences of a gene. Probes made from the latter are then utilized to look for the fragment having the gene.

The transposon may make its presence felt by altering a normal phenotype. The method of transposon tagging was first employed to isolate the *white* locus in *Drosophila.* The mutants with *white-apricot* eyes harboured a *copia* retrotransposon in the white locus. A probe made from *copia* identified the transposon as well as the accompanying *white* locus. Many eukaryotic genes have been purified by this method.

3.3 ❑ GENE LIBRARIES

The collection of cloned DNA fragments that represents the entire genome of a species is referred to as a 'gene library'. A gene bank or a clone bank may be a complete or an incomplete gene library.

The gene library may be a *genomic* or a *cDNA* one. A genomic library ideally contains every region of the genome represented in the collection of cloned fragments. A cDNA library consists of the genic regions of the genome. As cDNA is copied from RNAs, non-genic regions and genic stretches of sequences not represented in the transcripts are also not present in a cDNA library. A cDNA library is further liable not to be representative of all the genes, as the ratios of RNA species from different genes vary in proportion and stability and may not, therefore, be included in the RNAs that are copied into cDNAs.

3.3.1 Genomic Library

For a genomic library the total DNA is cleaved into smaller fragments (mechanically or enzymatically), each one of which is cloned in a suitable vector. The rDNA

carrying each fragment is amplified in *E. coli* cells. One can calculate the minimum number of clones that must be present to represent every region of the genome. For this calculation, one needs to know the average length of the fragments and the approximate total length of the genomic DNA. Clarke and Carbon (1976) proposed a formula that is generally utilized, keeping in mind that the values obtained are good approximations and not exact ones. One may find the value of *N,* the number of clones needed to represent every bit of the total DNA from the relationship

$$N = \frac{\ln (1 - P)}{\ln (1 - F)}$$

where F is the ratio of the length of the fragment to the entire genomic length and P is the probability of finding a specific sequence.

Fragments for a genomic library should be long enough to include an entire gene sequence. It is customary to have average fragment lengths of 15–20 kb. This can be obtained by mechanical shearing of the DNA or by the use of restriction enzymes. What enzymes would one use? The criterion for the use of the particular restriction enzyme is partly how often it is likely to come across its cleaving site in the DNA. A recognition motif with four nucleotides is likely to be present every 4^4 or 256 bases, while a hexanucleotide motif will be present in every 4^6 or 4096 bases. The latter will produce fragments too long for cloning and characterization by restriction mapping. The preferred fragment length is 15–20 kb as mentioned earlier.

The next decision to be taken is how long should the DNA be treated with the chosen enzyme? If complete digestion is allowed, every DNA will be cut up into identical sets of fragments, which will be very large in number. In practice, partial digestion is preferred, as it produces longer fragments; these are also random ones, as the enzyme does not cleave at every site, nor at the same sites in every DNA strand. Such random fragments, therefore, overlap at places, a feature that is useful for lining up the fragments in the correct original order during the melding operations in DNA sequencing. The procedure of connecting sequences of consecutive fragments by aligning their overlapping regions (ends) is popularly known as chromosome walking.

There is one other reason for the fragments to be about 15–20 kb in length. The fragments usually need to be mapped by restriction enzymes other than the one used for the initial fragment production. Restriction mapping is done best if the fragment is not too long (i.e., > 20 kb). Such a long fragment is likely to possess more than one cutting site for an enzyme. This complicates the mapping operation.

The genomic library will be screened for desired regions of the DNA. Since genes and their flanking regions are of primary interest, it would be ideal to have fragments with unmutilated genes—that is, genes without a portion cleaved out with the enzyme. Often, it is desirable to know the nature of the sequences on either side of a gene. A good collection of random fragments usually satisfies the above mentioned needs.

The cloning vector can theoretically be any one of the several available vectors. However, for a variety of reasons, λ-replacement vectors and cosmids are used for making gene libraries. The merits of these two will be discussed later.

The first step in making a genomic library is to isolate genomic DNA with as little shear as possible. The high molecular weight DNA is then digested partially with a restriction enzyme. The fragments are fractionated by gel electrophoresis or sucrose (or sodium) gradient velocity centrifugation. The sodium gradient centrifugation is preferred, as it does not require an additional step of removing the salt (as is done for the sucrose). A size standard is run in the same gel, and lengths of 18–22 kb are identified among the fragmented genomic DNA. Fragments larger than these are not packaged into phage particles, while ones much smaller may join tandemly in a random order to make up the optimum packaging size. The latter are not useful for obvious reasons.

Of course, by confining oneself to a size range, one may miss larger fragments that have no restriction sites, as well as very small ones that are lost during processing. By creating a library of a sufficient number of clones, one may hope to ensure the inclusion of every region of the genome.

The enzymes used for fragmenting the DNA are usually *Mbo* I and *Sam* 3, both of which recognize a tetranucleotide motif, GATC, which is the core of the *Bam* HI motif GGATCC. Sticky ends caused by these three enzymes are compatible. With the calculation that a 4-base motif occurs every 256 bases, a 15 kb DNA would be expected to possess about 50 such sites; to recover fragments of this size only 1/60 of the total cutting sites may be engaged by the enzyme. Controlled partial digestion would ensure this result. In the same vein, a 20 kb fragment can be recovered by using only 1/80 of the possible cleavage sites.

To prevent donor DNA fragments from joining with each other, or from circularizing, the 5′-ends of the strands are dephosphorylated with alkaline phosphatase. This DNA, when purified, is ready for ligation to the phage λ-vector. The unligated nicks at each juncture of vector and donor DNA are joined by the ligase in the *E. coli* cell in which the hybrid molecule is introduced. The donor DNA is spliced to the arms of the vector and recombinant DNAs are packaged with packaging mixes.

The hybrid DNA from each phage particle is amplified in *E. coli* either by phage infection or transfection with the naked DNA. The method of choice will depend partially on the resources of the user; packaging mixes are not inexpensive.

The library can be stored appropriately, and used to fish out a desired gene or other regions of the genome or for chromosome walking. Probes are required for both operations. The hybridization is usually carried out on nylon filters or on nitrocellulose paper to which replicas of plaques have been transferred by blotting (a modification of the Hogness-Grunestein technique). The blots are treated with a mild alkali (NaOH), which releases the DNA from intact cells and denatures the DNA into single strands. Free, unpackaged DNA present in the lysate in each plaque also becomes adsorbed on the blotting paper. The filter is then flooded with the probe, treated as usual to fix the hybridized DNA and autoradiographed for detecting the locations of the latter.

In practice, the term ‘genomic library’ is used for a library of clones made from the native DNA, as opposed to that from cDNAs.

3.3.2 cDNA Library

cDNA (complementary DNA) is synthesized against an RNA template using the enzyme reverse transcriptase. A cDNA library or clone bank contains cDNA rather than genomic DNA fragments spliced to the vector.

Although any RNA may be copied into a cDNA, in practice, mRNAs have been mostly used as templates. A cDNA clone bank, therefore, usually represents the heterogeneous population of mRNAs present in a cell.

A complete clone bank should represent every mRNA in a species. This rarely happens, as all mRNAs are not synthesized in every cell, nor are the same sets found in cells of different tissues and/or at different developmental stages. A cDNA clone bank would thus reflect the genes expressed in the cell from which the mRNA templates are taken.

Again, a cDNA clone bank even from the same cell need not contain representatives of every transcript that this cell synthesizes. Transcripts that are not very stable or long lasting, or those which are produced at negligibly low levels, are likely to be missed.

Despite these shortcomings, and in spite of the fact that it is more time consuming and expensive to prepare cDNA clone banks than genomic ones, several occasions demand the use of cDNA and not the genomic DNA banks or libraries. Let us list some of these requirements.

First of all, there are far less numbers of clones of cDNA in a bank than in a genomic library. This makes it easier to look for a desired gene. A complete genomic library may require 100–1000 k fragments, whereas a cDNA bank may contain only 10-30 k clones to choose from.

Screening a cDNA bank also provides fairly unambiguous results. The probes utilized for screening genomic fragments are made by labelling *in vitro* mRNAs or cDNAs. In practice, there is no guarantee that the isolated mRNA is not contaminated with the rRNAs that occur in profusion in every cell. Some of these rRNAs also become labelled, and result in 'false positives' when used to hybridize the genomic clones. No such mistake can occur with a cDNA clone bank. There is always a unique mRNA that will bind itself with the correct cDNA.

A cDNA bank is useful when one is particularly interested in recovering a specific gene. If the cDNA is synthesized against the mRNA from cells expressing the latter at a high level, a large percentage of the clones will represent this mRNA, and hence its gene.

There are at least four other justifications for preparing cDNA clone banks. One of them is for cloning genomes of RNA viruses; cDNA copies are the only possible inserts that can be spliced to vector DNA. Secondly, one can get an idea about the base level house-keeping genes by comparing cDNA clone banks from cells at different stages of development and from different tissues. The luxury or special mRNAs in them may be different, but there will always be one set of genes that is expressed in all of them. Again, if it were necessary to express a eukaryotic split gene in a bacterial cell, a cDNA prepared from the spliced mRNA would be translated in the prokaryote; the latter will not be able to handle splicing of the mRNA, since the infrastructure for such operations is not present in them. Further, a cDNA is

useful for delineating the exon and intron regions of a split gene. When a cDNA is hybridized with the genomic counterpart, only the exon regions pair; the introns are thrown out as loops.

As in the case of preparing a genomic library, Clarke and Carbon (1976) have provided a formula for estimating the number of clones (*N*) that will guarantee the presence of the minimum number (*n*) of clones required to represent all classes of abundance of mRNAs. There are three classes of abundance: high, medium and low. The low abundance class may have several thousand different species, but only a very few copies per species. In contrast, the other two classes may represent fewer number of genes, but the latter produce several thousands or hundreds of copies of mRNA per gene.

If one was interested in screening a cDNA clone bank for one of the high or medium abundance class genes, only a few (7000–8000) clones would be enough, as many of them would be duplicates from the same mRNA species. Low abundance clones may be recovered only from a much larger collection. If a specific probe was available for the low abundance species of interest, it would not be very cumbersome to screen even a large collection.

The cDNA bank is prepared by first extracting the total RNA from a cell and then purifying it from any contaminating DNA by caesium chloride density centrifugation. The already purified mRNAs are further purified by running them through a column containing poly-dT oligonucleotides. The poly dA tails of the mRNAs become hybridized to these oligonucleotides; the mRNAs are eluted from the column.

The mRNAs are next copied into ss cDNA using a poly dT-oligonucleotide as primer and the AMV reverse transcriptase as the DNA polymerase. There are at least two procedures for synthesizing the second-strand of the cDNA.

To enable the ds cDNA to be joined to the vector DNA, it is customary to provide suitable joining ends to the cDNA. If the ends of the latter are not 'flush', they are made so using S1 nuclease. Flush-ended cDNA may be joined to flush-ended vector ends using T4 ligase. Or else, a staggered end, specific for a particular restriction enzyme, may be created at the cDNA terminally by first joining the restriction enzyme linker (synthetic short length of DNA containing the cleavage site of an enzyme), which is subsequently cut with the restriction enzyme. The sticky tails of vector and cDNA may be extended by complementary homopolymer tails and TNT.

The vector and cDNA are ligated and *E. coli* cells transformed with them. When the transformant colonies are grown, representatives from each are stored in a storage medium (usually containing DMSO), at –20°C or –70°C. Hanahan and Meselson have developed a neat way of storing the transformants in a very small space. They grew the bacteria on millipore filters placed on the nutrient agar. The filters with the colonies were frozen and stored at –70°C. These can be utilized for screening for the desired clone. Usually, the screening is performed in two steps: the primary screening identifies the clones likely to possess the sequence in question, and the secondary screening pinpoints the ones with the sequence. The latter are then characterized.

3.3.3 Phage Lambda versus Cosmids for Gene Libraries

Cosmids can carry a maximum of 45 kb of insert DNA, whereas phage λ vectors can bear a maximum of 25 kb of the donor DNA. The larger insert DNA can encompass long genes, particularly eukaryotic split genes. This advantage of cosmids, as well as that of requiring fewer clones to include all fragments in the genome, is offset by the fact that they are less easy to handle, the longer DNA does not always become packaged, and several recombination events may reduce the size of the cloned fragment.

Phage λ-based libraries, on the other hand, do not present the above problems. In spite of the larger number of clones that have to be screened, the techniques for screening are so simple that the larger number of clones is not much of a problem. The smaller insert DNAs may cut through a gene, or very large or very small fragments may be left out during packaging. By ensuring a sufficient number of clones of random fragments, the sequences in adjacent fragments may be determined, and the intact gene characterized.

3.4 ❑ POLYMERASE CHAIN REACTION (PCR)

The most recent addition to the trend-setting techniques that have transformed the depth and precision of analysis in molecular biology is the one known as the Polymerase Chain Reaction or PCR.

PCR was invented by Kary B Mullis and first used for an application by the Department of Human Genetics of the Cetus Corporation of USA. The basic technique has been modified considerably and has become an indispensable tool in a phenomenally wide range of research and application-oriented endeavours.

The PCR technique is a simple and rapid method of amplifying a length of target DNA into thousands of copies (to the order of 10^9 fold) in a test-tube and in virtually a couple of hours time. It is, in that sense, a purely chemical technique and as such has little justification for inclusion in an account of techniques of molecular biology and genetic engineering. However, the sea-change brought about in molecular biological investigations, by the advent of this technique, has led to the equipment of most major laboratories, engaged in biological research, with PCR facilities. PCR is an invaluable tool that supplements, and often supplants the conventional techniques of recombinant DNA technology. It is, therefore, useful to have a minimal acquaintance with the fundamentals and areas of use of PCR.

The PCR technique (Fig. 3.9) involves repeated enzymatic synthesis of a segment of DNA using a specific DNA polymerase and two oligonucleotide primers that match the two extremities (one for the 5′ end of the sense and the other for the 5′ end of the antisense strand) of the DNA segment to be amplified. The method consists of repeated thermal cycles, each cycle consisting of three steps. The first step in a cycle is the dissociation of the two strands of the sample DNA, and this is achieved at 94°C. The second step is that of annealing the two primers at the appropriate locations in the separated strands; for this step the temperature is lowered to 55°C. The final step involves the extension of the primers to the 5′ ends of the templates; the temperature for this step is maintained at 72°C. Steps 1, 2 and 3 require 20, 20

and 30 seconds, respectively and each cycle time is approximately 3.75 minutes. The operation is continued for about 30 cycles. At the end of cycle 1, the original two templates of DNA acquire two complementary strands, each stretching from the primer to the end of the template strands. In the second cycle all four strands (2 new, 2 old) separate and each becomes available as a template for further DNA synthesis. At the end of cycle 2, therefore, there are 8 strands, 4 of which now stretch only between the sites of the two primers. The number of copies of DNA thus increase exponentially, until at the end of a time a little less than 2 hours there are more than 2^{30} copies of the target DNA.

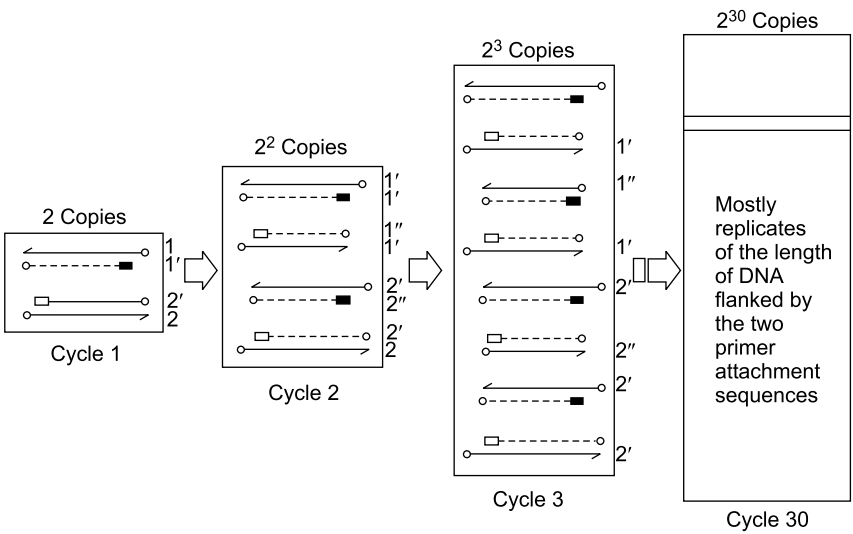


Fig. 3.9 *Principles of PCR: (a) The steps in each thermal cycle, (b) Exponential increase in the number of copies of DNA via PCR.*

The entire reaction is carried out in a test-tube containing the DNA (say, about 0.1 μg human genomic DNA, having about 10^2–10^5 copies of the target DNA), the two primers, a DNA polymerase that is not degraded by repeated subjection to high temperature, the four deoxynucleotide triphosphates and buffering material. These operations are uncomplicated and are repeatable with a satisfactory degree of success. However, there is yet no standardized reaction brew and protocol that is suitable for all projects. In general, modifications are made, to suit a particular project, in the following basic recipe. The reaction is carried out in a 50–100 μl volume of reactants. The latter include, besides the 0.1 μg of DNA, (i) the primer (0.25 μM each), (ii) a DNA polymerase (2.5 units), (iii) the 4 deoxynucleotide triphosphates (dATP, dGTP, dCTP, dTTP, 200 μm each), (iv) KC1 (50 mM), tris. HCl (10 mM, pH 8.4 at room temperature), $MgCl_2$ (1.5 mM) and (v) gelatine (100 μg/ml).

The operations may be carried out manually, using water-baths at preset temperatures, or performed in the automated DNA Thermal Cycler instrument from Perkin-Elmer Instruments, USA.

The polymerase originally employed by Mullis was the Klenow fragment from the *E. coli* DNA polymerase I. The Klenow fragment, however, becomes inactivated at the high temperature required to dissociate the strands of a duplex DNA, and hence, lead to the replenishing of the reaction mixture with fresh Klenow fragment at the end of each cycle. This botheration was overcome by using the polymerase from an organism that habitually lives in a very hot environment. The organism tapped for such an enzyme is the thermophilic eubacterical species, *Thermos aquaticus,* and the enzyme is the *Taq* polymerase. This enzyme also improved the specificity and yield of the target segment. The use of the Klenow fragment, at 37°C, results in the amplification of several non-specific (non-target) segments. The low yield of amplified target DNA have then to be identified by appropriate probes. The *Taq* polymerase works at about 75°C; at this temperature the primers also do not anneal indiscriminately to non-target sequences. In addition, as mentioned earlier, the *Taq* polymerase is not degraded by repeated exposure to high temperatures; the enzyme retains about 65 per cent of its activity even after a 50-cycle operation.

It has been observed that after a certain number of DNA strand replications, the amount of enzyme, included in the reaction mixture, is no longer sufficient to cope with the increased number of templates to be duplicated. A plateau is reached in the yield of the PCR product at this stage. When using the Klenow fragment this plateau is reached after 20 cycles. The *Taq* polymerase remains sufficient for at least 30 cycles. In addition, the *Taq* enzyme allows amplification of a much longer DNA segment (as much as 10 kb in length) than does the Klenow fragment, which can cope with a maximum of about 400 base pair lengths of DNA. These and other advantages have replaced the use of the Klenow fragment in the PCR technique with that of the *Taq* polymerase. In practice, DNA lengths with an upper limit of 3–4 kb are suitable for PCR.

One disadvantage of using the *Taq* polymerase is the propensity of induction of base-substitution mutations in the copies of the target DNA. This enzyme, unlike say the T4 polymerase, does not possess an appreciable amount of $3' \rightarrow 5'$ exonuclease activity, and is, therefore, incapable of 'proofreading' the newly synthesized strands. Hence PCR products, when using the *Taq* polymerase, are not very homogeneous in terms of the target DNA. Where greater fidelity of replacation is desired, enzymes, such as the T4 polymerase, are used instead. To obtain greater yield and homogenity of the target product in relation to the yield of non-target products, the other parameters of the reaction may have to be altered as well. Much research is focused on improving the yield and purity of the PCR-amplified DNA by modifying the parameters of the so-called basic or standard reaction.

PCR does not depend upon the construction of recombinant DNA or upon cloning the hybrid DNA inside a cell, with the attendant steps of identifying and recovering the cloned DNA. PCR thus eliminates all these laborious and time-consuming operations, characteristic of traditional gene cloning technology, and increases exponentially the copies of a target DNA in a totally *in vitro* operation.

The PCR technique can use extremly minute amounts of the DNA sample from which the target DNA is to be amplified. The ability to recover appreciable quantities of DNA starting from that present in dried blood stains, hair and other biological samples, has made PCR an important tool in forensic investigations.

PCR can also amplify the target DNA from a rather impure sample of the latter, or from a highly heterogeneous mixture of biomolecules. Indeed, adding whole cells to the reaction mixture is theoretically sufficient to provide the DNA required for the reaction; the cells lyse and release the DNA during the first step (denaturation) of the cycle. For practical reasons, however, it is preferable to prepare the target DNA more carefully. This is especially important when the number of cells containing the target DNA is a very small aliquot of the total population of background cells to be screened. Examples of such cases include the search for the few cells in a transgenic animal that carries the incorporated transgene and the screening of a large population of lymphocytes, only a very small percentage of which is infected with HIV.

Since PCR can use target DNA present in crude cell lysates, often whole cells are added to the reaction chamber together with a non-ionic detergent (e.g., Trition 100–X) that solubilizes the contents of the cells and a proteolytic enzyme (e.g., Proteinase K) that digests the proteins in the lysate. Triton 100–X and Proteinase K are common reagents in several protocols for purifying DNA from animal cells. As Proteinase K is denatured at the dissociation temperature ($\geq 75°C$) of step 1 of the PCR cycles, it is customary to add the *Taq* polymerase after the lysis-protein-degradation step is completed.

In some samples (e.g., whole blood) one or more of the cell lysis products may inhibit the action of the PCR enzyme. In the case of RBCs, hematin and heme-derived porphyrins act as inhibitory agents. Protocols are available to overcome such obstacles.

Protocols have also been developed for preparing DNA for PCR from sources such as single hairs, clinical swabs containing virus-infected epithelial cells and non-blood tissues.

RNA may also be amplified by PCR. The RNA is first used as a template to synthesize cDNA (using reverse transcriptase) and the latter used for amplification. To prepare mRNA from whole cells two precautions have to be taken; (i) the cells must be lysed keeping the nuclei intact, and (ii) the endogenous RNase must be made inactive. The former is achieved by using the non-ionic detergent NP40 to disrupt the cell wall, followed by centrifugation to pellet out the nuclei. The second goal is secured by using an inhibitor of RNAse, such as DEPC (diethylpyrocarbonate; Sigma): Additional steps are taken to ensure that only the mRNA and not the DNA is amplified.

A notable advantage of PCR is the ease with which directed mutagenesis can be achieved in the target DNA. The region to be mutated must lie next to the 5′ end of the target DNA, with the primer on the 5′ end of the former. That is, the mutated base or bases lie sandwiched between the primer and the 5′ end of the target DNA. The additional base or bases may be incorporated by adding the required segment to the 3′ end of the primers. The mismatch produced by the additional base/bases is tolerated by the replicating system in the initial cycle. After this cycle the add-ons are already part of the target DNA. The additions may represent any desired sequence—short lengths of DNA containing restriction enzyme sites—promoters or other regulatory sequences or tailor-made sequences required for a specific project.

It must be obvious from the preceding accounts that, only a DNA with a known sequence is suitable for a PCR operation, since the sequences for the two primers have to be derived from the target DNA strands. This constraint no longer poses a deterrent for using 'unknown' DNA; techniques are now available that can use DNA the primer sequences of which are not known.

In one method for using DNA of unknown sequence in a PCR, primers are added to flank the target sequences. The primers are joined to the latter by a simple ligase-mediated reaction or by the use of homopolymer tailing. In cases where the sequence of the target primer region of the DNA is to be inferred from a polypeptide with a known amino acid sequence, a pool of degenerate primers is utilized. This is necessary in view of the degeneracy of the genetic code, which does not allow one to predict the actual codes used for each amino acid during the synthesis of the polypeptide. This method has been used successfully for amplification of cDNA sequences, which are synthesized against mRNA templates, which of course, do not possess the primer-attachment region of its corresponding DNA template.

The use of PCR has been extended for the analysis of sequences flanking a target DNA. This method is known as 'inside-out PCR' or 'inverse PCR'. The first step in this technique is to cut out a length of DNA, including a known core or target DNA, with suitable restriction enzymes. The cut ends are then ligated to produce a double-stranded circular DNA. Now, two primers are constructed that are homologous to the ends of the core region, but which are oriented towards the flanking sequences (away from the core region) as shown in Fig 3.10. In one approach, when a restriction enzyme site is present within the core region, the primer extends along the flanking DNA which serve as templates; the two flanking

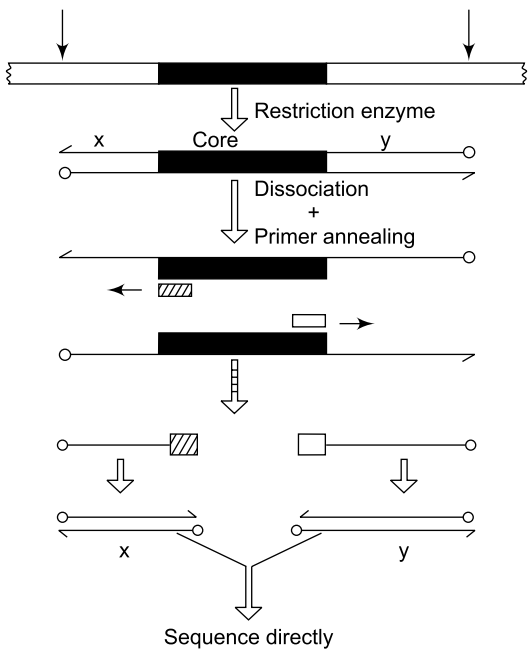


Fig 3.10 *Principles of Inverse PCR*

DNA are, therefore, amplified separately. In another approach, when the core is not cleaved by a restriction enzyme, the two flanking regions are replicated together and subsequently cleaved by the restriction enzyme used to initially circularize the DNA to be amplified.

Inverse PCR has many uses. It has been employed to analyse unknown flanking DNA both of an entire gene, and of just the 5′ and 3′ flanking regions within a gene. The technique has also been found to be useful for preparing probes for ends of sequences incorporated into various types of vectors. This extension of the straightforward PCR was developed by D L Hartl, together with his colleagues H Ochman, J W Ajioka and D Garza, of the Department of Genetics, at the Washington University School of Medicine.

PCR has made inroads into a plethora of areas of investigation and application in molecular biology. Perhaps the area of application that has benefited most by the use of PCR is that of DNA sequencing. Conventional sequencing protocols begin with the preparation of templates for the sequencing operations. The latter are purely *in vitro* affairs, and can be begun as soon as the DNA to be sequenced is present in a sufficient quantity and in a pure form. PCR eliminates the lengthy and at best messy steps of conventional DNA cloning protocols and delivers amplified samples of pure target DNA. Together with automated sequencing techniques, PCR has made DNA sequencing a simple, rapid and reliable operation.

Research goals that use PCR include creation of synthetic DNA, analysis of gene expression, detection and characterization of mutations and the induction of directed mutagenesis for functional analysis of regions of both genic and non-genic DNA. Other uses of PCR in research include the construction of genetic maps by using the DNA sequence in single meiotic products (sperm) to estimate the percentage of recombination as indicated by the presence of alternate alleles of two genes. This is a far simpler method for mapping the sites of two genes than the traditional ones based on breeding, analysis of progency, and pinpointing of the relative positions of gene loci.

PCR followed by direct DNA sequencing has become a valuable tool in the hands of population biologists, especially for evolutionary studies. Contemporary molecular approaches for investigation of evolution, by analysis of population and reconstructions of phylogenies, utilize restriction mapping of DNAs, analysis of isozymes and comparison of other proteins as well as of DNA sequences. The amino acid codons of proteins are compared, for instance, to assess the extent of evolutionary differences between the subjects furnishing the sample proteins. DNA and RNA sequences are also utilized on the same principle. All these methods examine the divergence between the samples and not the actual sequence of amino acids or codons etc. in the protein or nucleic acids respectively. PCR can furnish the actual sequence data for the DNA of the organisms to be compared.

Both nuclear and organelle DNA are material for evolutionary studies. To study nuclear DNA, using PCR and direct sequencing, convenient sequences are those of the DNA for ribosomal RNAs. T D Kocher and T J White have identified a sequence within the 18S rDNA that is conserved in *Saccharomyces cerevisiae, Dictyostelium discoideum* and *Stylonichia pustulata.* They have prepared two 'universal' primers from this sequence which may be used to amplify the 18SrDNA from a variety

of non-bacterial and non-mitochondrial DNAs. For tackling eukaryotic DNA the first choice has been the preparation of universal primers from two regions in the mitochondrial gene of cytochrome *b* that is identical in most vertebrates that have been studied so far. Sequences of *cyt b* from sample organisms may be amplified via PCR and aligned, to estimate the types and degrees of differences in them, and the data utilized for assessing evolutionary relationships.

PCR has revolutionized several aspects of analysis in the fields of diagnostic and clinical medicine. The areas that have exploited the use of PCR most include those of identification of pathogens, prenatal diagnosis of genetic disorders, analysis of X-linked mutations that result in genetic diseases, and those in oncogenes responsible for precipitating cancers. Use of PCR has made it possible to achieve the above objectives with far greater speed, precision and reliability of results than is possible by the conventional techniques.

The first genetic disorder to be identified prenatally, using PCR, is that of sickle-cell anaemia. Subsequently, other haemoglobinopathies based on single gene mutations were also diagnosed by using PCR-amplified regions of the concerned gene. An idea about the speed of operations when using PCR can be gained from the reports from the laboratory of H H Kazazian, Jr. of the Department of Pediatrics (Genetics Unit) of the Johns Hopkins Hospital in Baltimore, USA. While the usual techniques of probing Southern blots of DNA obtained from cells extracted through amniocentesis took Kazazian's team at least two weeks from the time of obtaining the test sample, use of PCR cut down the time period to 1–4 days. Similarly, instead of spending two weeks or more in indirect detection of β-thalassemias through RFLP-linked polymorphisms, Kazazian and associates accomplished the task in about one week; the latter protocol involved looking for mutations in PCR-amplified regions of the gene for β-globin.

Conventional indirect detection of genetic disorders requires painstaking family studies to determine polymorphisms, if any, that accompany the mutated version of the gene. Y W Kan and A M Dozy were the first to report the presence of a DNA polymorphism that accompanied the mutated β-globin gene in cases of sickle-cell anaemia. Point mutation markers were then (1978) searched for in the vicinity of genes, mutations in which are known to be the basis of genetic disorders. Subsequently, a better marker system was discovered. The gene responsible for the genetic disorder was often found to be associated with base polymorphisms that resulted in obliteration or alteration of cutting sites for some restriction enzymes. In any event, even this advance in technique due to the use of polymorphisms in restriction fragment lengths (RFLPs), use of which eliminated the need to search for point mutations, depended on pedigree studies. PCR-aided diagnosis of disease genes, on the other hand, does not require pedigree analysis; analysis of the parental and foetus DNA is sufficient for the prenatal diagnosis of genetic disorders, and dot-blot hybridization, restriction fragment analysis or PCR with direct sequencing of potential carriers and normal (control) subjects yields the required information in a very short time.

Other types of mutations that may be searched rapidly by PCR-based techniques are the ones borne on X chromosomes and showing up as disorders in the male. Once such a defective male is identified, it is customary to collect blood samples from

relevant family members and hybridize the DNA of cells cloned for this collection, as well as that of the affected male with probes to identify variations in the DNA of the affected person. This entire sequence of events involves many operational steps which are time-consuming, to say the least. One other disadvantage of the older method of identifying new mutations in the genes responsible for genetic diseases is that different projects have to be undertaken to decipher the type of mutation involved—point mutations, deletions or other rearrangements.

PCR-amplified DNA, followed by direct DNA sequencing, provides specific data for all types of base substitutions, deletions and other rearrangements in the genes that are involved in various genetic disorders. Unlike RFLP data or other marker data, the PCR-based techniques supply the basic information (sequence data) in the relevant DNAs. This approach has been utilized by C T Caskey and colleagues, at the Institute for Molecular Genetics at the Baylor College of Medicine, for identifying new mutations in the HPRT gene (that are responsible for the Lesch-Nyhan syndrome) and the dystrophin gene (underlying Duchenne muscular dystrophy). These researchers have succeeded in developing a strategy that identifies the entire range of DNA lesions (point mutations, deletions and other rearrangements) in one go, via PCR-based techniques. Their unified approach may be adapted for use as a standard strategy for examining the defective genes underlying other genetic disorders.

Identification of polymorphism is also a requirement for other objectives. For instance, one needs to know the allelic contents of HLA class II genes for the purpose of tissue-typing for transplantation purposes, for the presence of tell-tale signs (in the sequence) that points to susceptibility to particular autoimmune diseases, as also for unambiguous identification of an individual. H A Erlich of the Department of Human Genetics at the Cetus Corporation has PCR-amplified specific regions of the HLA class II (HLA-DRβ, HLA-DQα, HLA-DQβ and HLA-DPβ) and used them for identifying polymorphisms in humans affected or suspected to be susceptible to autoimmune diseases, and in members of different species in the evolutionary lineage to man. In the former (e.g., insulin dependent diabetes), the data was useful for exposing the sequence diversity of the HLA class II genes responsible for the disorder, as well as for charting the distribution of such allellic DNA sequences in normal and diseased members of a population. The latter case provided information that could be used to build a phylogenetic relationship between the HLA Class II genes, in terms of evolution, of a comprehensive range of primate species. The HLA-DR, HLA-DQ and HLA-DP regions of the HLA class II genes each code for two glycopeptides (α and β). The glycopeptides are anchored on the membranes of B cells, activated T cells and macrophages, with the NH_2 terminal outside the cell. It is the second exon, which encodes the NH_2-containing domain, that exhibits the polymorphisms responsible for assigning individuality to each person. Such class II polymorphism is traditionally identified by serological techniques. The PCR-based techniques, however, provide more information than possible by the serological methods. As an example, a much greater diversity in the base composition of alleles surfaced when HLA-II genes were amplified by PCR (using a primer for the second exon) and sequenced directly than realized earlier. What appeared to be distinct serologically-defined alleles were in fact found to contain considerable variations

when the PCR-based products were examined. The strategy consists mainly of the following sequence of steps The specific DNA region is PCR-amplified, and the copies immobilized on nylon membrane filters (replicate). The filters are hybridized with sequence-specific oligonucleotide probes, labelled non-radioactively with biotin or with the appropriate enzyme for a ELISA test. Each filter is then examined to find the one showing hybridization with the probe. This is the regular dot-blot method. Conversely, a reverse dot-blot method may be used, by which the probe is immobilized on the nylon membranes and the labelled PCR products used to hybridize them. The former is a more complex approach. In any case use of either or both of these methods results in rapid typing of polymorphism of HLA class II genes.

It is obvious that a technique as powerful as PCR will find use in various aspects of cancer research. J L Bos, at the University of Leiden, The Netherlands, has used PCR-amplified regions of the DNA of the *ras* oncogene from different types of human tumours that show evidence of *ras* involvement, and directly sequenced the PCR products. This provided extensive and precise data on the point mutations in positions 12, 13 and 61 that are associated with each type of tumour studied.

Similarly, PCR use has been a great boon for identification of pathogens and contaminants.

Forensic investigations have been greatly aided ever since it has been possible to identify individuals by DNA polymorphism analysis (also known as DNA-finger printing). The initial methods for detection of such genetic variation at the allelic (molecular) level rely mainly on characterization of RFLP. However, this technique requires a substantial amount of DNA ($\geq$ 50 ng), an amount not always available in practice. Furthermore, cutting up the precious meagre sample of DNA leaves nothing for any further analysis. These obstacles are avoided by using PCR-amplified DNA sequences that include specific marker loci. Two main attributes that markers for PCR should possess, are several allelic variations of the markers which are preferably situated on separate chromosomes (linkage groups), so that they do not exhibit linked inheritance. The number of such markers are still very few, but many more are expected to be identified in the human genome.

Sequence and length variations are the two types of polymorphisms present in DNA sequences. Sequence variation can be analyzed by dot-blot hybridization, with the DNA sequences immobilized on nylon membrane filters and hybridized with allele-specific oligonucleotide (ASO) probes, or vice versa. Of course, a panel of ASO probes have to be made from the DNA sequences of preferably all the alleles of a marker locus. Length plymorphism usually is due to a variable number of tandem repeat (VNTR) in the loci. The repeat regions are amplified and the products dispersed by size using electrophoresis in an analytical gel. The different lengths of the tandemly repeated regions can be characterized further.

Forensic samples have been examined most to identify specific individuals from typing of the HLA class II locus DQα. Oligonucleotide probes have been designed that can discriminate not only six of the eight alleles known for this locus, but also for subtyping one of the alleles. HLA-typing in this case is performed, as in the preceding example, via PCR-amplification and dot-blot hybridization with

appropriate probes. Other regions of the HLA class II gene segments have also been utilized for typing purposes.

When the amount of the sample is very little, it is better to PCR-amplify the mitochondrial DNA than the genomic (nuclear) DNA. The advantages of using mtDNA for PCR-based identification are many. One of them is the fact that the nuclear DNA usually carries only one copy of a structural gene, which has to be amplified, whereas a mitochondrion contains several copies of the same single-stranded circular DNA. A region of the mtDNA, known as the D-loop region shows a high order of polymorphism.

Several other tissue-typing systems based on the use of PCR products are in the offing. They would enrich further the bag of tools at the disposal of the forensic expert.

The above is only a cursory account of the exciting new developments in chemical techniques that has become the mainstay of most investigations in molecular biology. Indeed, PCR has totally altered, in many instances, the manner in which one has been approaching problems in biology. The spin-offs of PCR in the field of applications are equally spectacular. A more elaborate account with details of the challenges and pitfalls of PCR may be obtained from *PCR Technology: Principles and Applications for DNA Amplification,* edited by Henry A Erlich and published by Stockton Press, Macmillan Publishers (1989).

REVIEW QUESTIONS

1. How is gene cloning done? Mention its importance in DNA engineering.
2. What are linkers, adaptors and connectors, and how are they useful?
3. How is a genomic library created?
4. What is cDNA library and its utility?
5. Give an account of PCR technique and its various modified versions.

4 Screening, Selection and Expression of Recombinant Clones

Introduction

Why does one need to 'clone' a piece of DNA? Where does one begin to answer such a question? Recombinant DNA technology has become a tool that is so fine-honed and the sphere of its influence so staggeringly large that perhaps it would be easier to tick off areas of research where it cannot be utilized profitably. Indeed, there is hardly a question in the enigma of the living system that will not profit by being phrased in terms of cloned DNA.

4.1 ❑ IDENTIFYING THE RIGHT CLONE

The clone containing the recombinant DNA may be identified by *screening* or *selection* techniques. (These two terms are not used interchangeably.) One *screens* a clone by distinguishing it from surrounding ones by some readily visible feature. One *selects* the clone that has survived a specific treatment that discouraged the growth of clones not carrying the recombinant DNA. In other words, if a population of recombinant and non-recombinant clones are subjected to a particular selection pressure, only the desired clones will survive, and the rest will be eliminated; they will be selected (or vice versa).

There are both direct and indirect screening methods and dominant selection techniques that are appropriate for different projects.

4.1.1 Direct Screening

The most popular direct screening methods are the ones in which the desired clone can be distinguished by the inactivation of a gene, by the biochemical reaction of a cell product with a chemical in the growth medium, or by the appearance of plaques due to bacteriophages.

1. Insertional Inactivation of a Marker Gene

Cells carrying the hybrid DNA can be identified by the mutant phenotype of a marker gene. The mutation is due to the cloning of the donor DNA within the marker gene. The disrupted gene, therefore, exhibits the mutant phenotype. The most common example of this technique is the cloning of an insert in the Tet^R or

Amp^R gene of pBR332. In the former case the hybrid DNA containing cells will be Tet^S (unable to survive in tetracycline-supplemented media). Similarly, the cells with a disrupted β-lactamase gene (Amp^R) will not survive in the presence of ampicillin.

Of course, the above mentioned screening tests are not carried out on the valuable master plates. Replicas are made of the latter in several fresh petri plates, and the replicas are utilized for screening. If a clone is not evident in a replica plate, the one in the corresponding position in the master plate is identified as the one that did not grow under selection pressure. The identified colonies or plaques are then amplified in separate plates to recover the clones carrying the desired DNA.

Another 'trick' is to use vectors that carry part of a *bio* or *trp* operon of *E. coli.* These handicapped phages can only grow in *E. coli* mutants that possess mutation in some other region of these operons. When an intact *bio*-vector is introduced into the *E. coli* mutant for another part of the same operon, the two defects complement each other, so that the cells grow without added biotin in the medium. However, if a donor DNA is inserted within the *bio* sequence of the vector, it can no longer complement the mutation in the cell; the result is an inability of the cells to grow unless biotin is supplied to them. Hence, cloning in such a vector will readily allow screening of cells carrying recombinant DNA; the latter will only grow on medium with added biotin or tryptophan (in the case of the vector with the *trp* sequences).

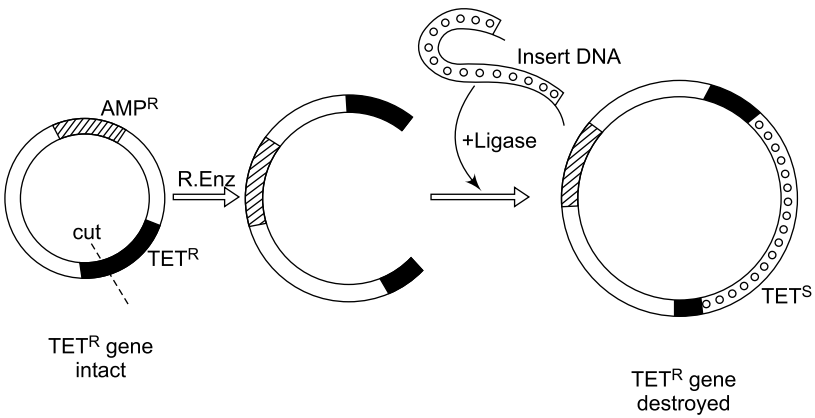


Fig. 4.1 *Insertional inactivation of the TET^R gene. Cells containing a hybrid vector with the foreign gene, inserted within the TET^R gene, do not survive in the presence of tetracycline. Cells with the intact TET^R gene (and, therefore, non-hybrid vector) grow in the presence of the antibiotic.*

Insertional inactivation of the lambda-repressor gene, *CI,* has been mentioned earlier. A DNA cloned within *CI* prevents lysogeny; hence, recombinant plaques are always clear, while non-recombinant ones are turbid in appearance (Fig. 4.2).

2. Visual Screening Methods

The commonest visual screening strategy uses the action of a gene product on a chromogenic substance to distinguish recombinant and non-recombinant clones. The gene product is the enzyme β-galactosidase that converts a colourless substance, XGAL (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) into a derivative that oxidizes into a blue non-diffusible dye (5, 5′-dibromo-4, 4′-dichloro indigo).

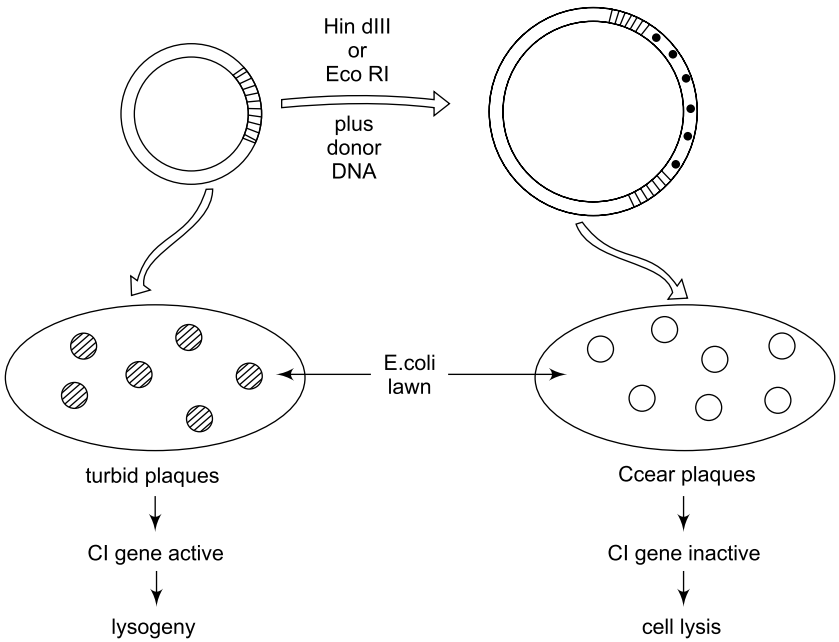


Fig. 4.2 *Insertional inactivation of the CI gene. Disruption of the CI gene, by cloning a foreign DNA within it, results in the absence of the lambda-repressor and the establishment of the lytic phase of the phage infection. Hence, clear plaques represent clones of recombinant phages.*

β-galactosidase is encoded by the *lacZ* gene of the *E. coli lac* operon. Several vectors carry the promoter and the first 145–150 codons of the *lacZ,* which represent the β-polypeptide of the enzyme. Cloning sites are available to introduce inserts into the *lacZ* sequences. A donor DNA inserted within the *lacZ* prevents enzyme synthesis. Cells grown in a medium to which XGAL has been added will be blue in colour or colourless, depending on whether or not the cell is expressing the gene for the enzyme. Recombinant sells are colourless, while non-recombinants are blue (Fig. 4.3).

Messing has exploited the *lacZ* XGAL system in several screening strategies. In one of them the β-peptide region of the *lacZ* and the promoter of an M13 vector is present. The *E. coli* in which the vector will be introduced carries a deletion in the aminoterminal region of the β-galactosidase which is complemented by the carboxyl end-deleted *lacZ* product from the vector. Together they produce functional β-galactosidase. If a donor DNA is cloned in a site within the *lacZ* region in the vector, this disrupted gene will no longer complement the truncated one in the cell. So, once again, the recombinant DNA-containing cells will be colourless.

A second vector constructed by Messing is a pUC (a derivative of pBR322) that carries the *lacZ* gene within which there is an inserted polylinker. The presence of the extraneous polylinker does not affect the expression of the enzyme. However, if the donor DNA is cloned within one of the restriction sites in the polylinker, no enzyme is synthesized.

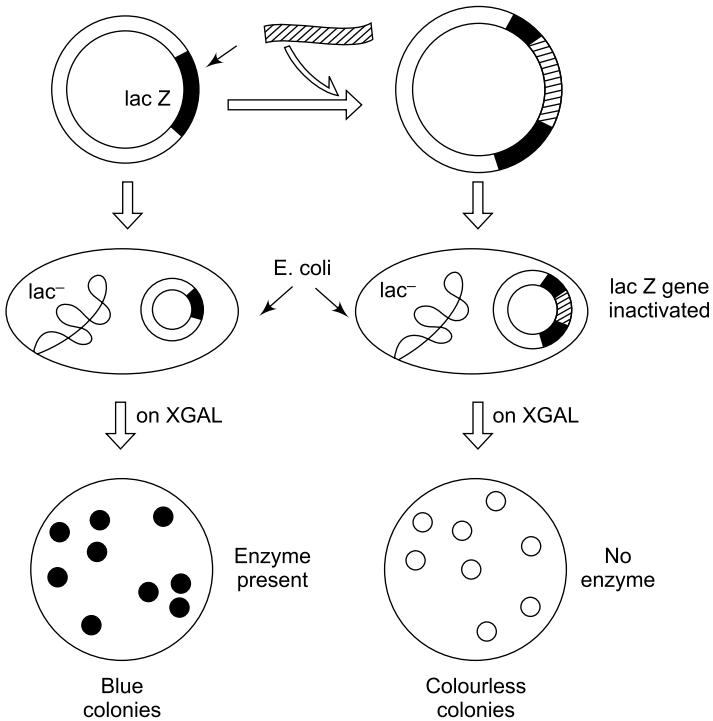


Fig. 4.3 *Insertional inactivation of the lacZ gene. The lacZ gene, one of the three genes of the E. coli lactose operon, expresses the enzyme β-galactosidase. Cloning a foreign DNA within this gene inactivates the latter; no enzyme is synthesized. The presence or absence of β-galactosidase may be detected by growing E. coli in a medium supplemented with a chromogen XGAL (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside). The enzyme converts XGAL into a product which is oxidized to a dark blue coloured product (5, 5′-dibromo-4, 4′-dichloro-indigo). Colonies synthesizing β-galoctosidase are dark blue in colour. Colonies harbouring disrupted lacZ sequences (hybrid vectors) are colourless on XGAL.*

3. Plaque Phenotype

The repressor protein, encoded by the *CI* gene establishes lysogeny and prevents the formation of free phage (λ) in *E. coli*. Vectors are available (immunity vectors) that have restriction sites within the CI (or immunity) region; imm 434 is one such vector with sites for *Eco* RI and *Hin* dIII. A DNA fragment cloned in any one of these sites destroys the *CI* gene. Hence recombinants are picked out as clear plaques.

4.1.2 Direct Selection

A dominant selectable marker is used to announce the presence of recombinant clones. The marker may be the insert DNA itself, or included in the system when the cloned DNA does not proffer a selectable phenotype. A few of the more commonly used systems are given below.

1. Complementation or Suppression of Mutation

The phenotype after transformation of the cell could be due to a complementation (Fig. 4.4) or a suppression (Fig. 4.5) of a mutant phenotype in the cloning cell. If the transfected gene is *his*$^+$ and the *E. coli* a *his*$^-$, transformation of the latter will make it a wild-type for the *his* locus. Such a cell will grow in minimal media without added histidine, unlike the *his*$^-$ cells that require histidine supplement for growth. In other words, the transfected gene complemented the lesion in the cell. K. Struhl (1976) utilized this method to isolate *E. coli* clones containing *his*$^+$ carrying fragments of yeast DNA. Fragments of yeast DNA were spliced to plasmids, that were used to transform *E. coli* cells.

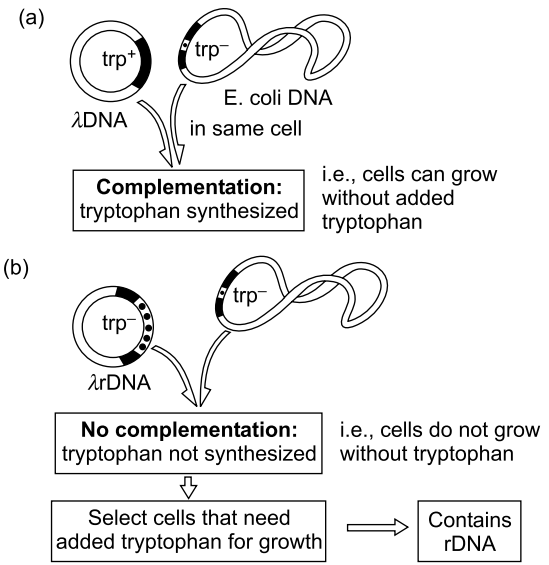


Fig. 4.4 *Selection of trp*$^+$ *by complementation*

2. Selection of Recombinant-deficient Phages

Lambda vectors, with cloning sites within the *redA,* and *redB* recombination genes, are available. Intact vectors can grow in *lig*$^-$ and *polA*$^-$ *E. coli* cells (incapable of recombining DNA) due to the presence of the phage *red* gene. If the latter is cut asunder by an inserted DNA, the phage no longer mediates recombination in the *lig*$^-$, *polA*$^-$ *E. coli.* The rDNA containing phages are recognized by plaque morphology.

3. The Spi Phenotype: Recombinant Proficiency

A phage lambda containing the recombination genes *red* and *gam* does not grow on *E. coli* (P2). The phage is said to be sensitive to P2 prophage instigation or interference (spi$^+$). A cell with a replacement vector that is *gam*$^-$, *red*$^-$, and which has the nonessential region replaced with the insert DNA, grows on *E. coli* (P2), that is, it has a *spi*$^-$ phenotype (Fig. 4.6). The latter is easily recognized, though the phenotype is expressed better if there is *chi* sequence in the cell. *Chi* induces the

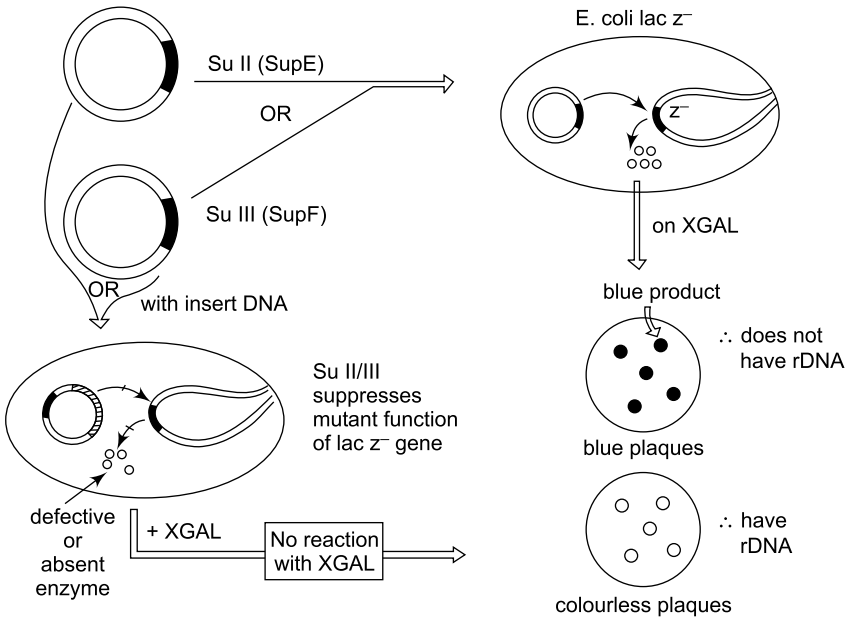


Fig. 4.5 *Selection of recombinant clones by suppression of mutations. The mutant tRNA, Su II (or SupE) and Su III (or SupF), genes suppress amber mutations (UAG) by reading the amber stop codon as an amino acid. SupF, a tyrosine-specific tRNA, reads UAG as a codon for tyrosine. Translation, therefore, continues with only one deviant amino acid residue in the polypeptide, which restores the wild-type phenotype of the mutated locus. This feature of suppression is exploited in many ways. In this illustration a foreign gene is cloned within a SupF gene inactivating the latter. That is, it is no longer able to reverse the activity (non-activity) of the mutation present in the lacZ gene. Recombinant colonies, therefore, remain colourless on an XGAL medium.*

production of concatamers of the lambda DNA with the aid of the *recA* product; only multimeric DNA can be the substrate for packaging into lambda capsids.

4. Selection for the Correct Promoter Sequence

The DNA fragment to be cloned is not necessarily only the coding region of a gene. Non-coding regions are cloned and studied for an understanding of their role, if any, in the expression of a gene.

Different regions of a promoter sequence, as well as different promoters, have been examined by using 'promoter probe' plasmids that have been derived from pBR322 (Fig. 4.7). The promoter of the Tet^R gene of this plasmid is inactivated by an 8-base insertion within it. The octanucleotide is an artificially synthesized one having an *Eco* RI site within it. A promoter sequence, from a promoter probe plasmid, may be inserted at the *Eco* RI site. The Tet^R gene thus acquires a new promoter. If the latter has the correct sequence of bases, the Tet^R gene is expressed. Degrees of efficiency of different versions of the inserted promoter may be assessed quantitatively by the activity of the Tet^R gene.

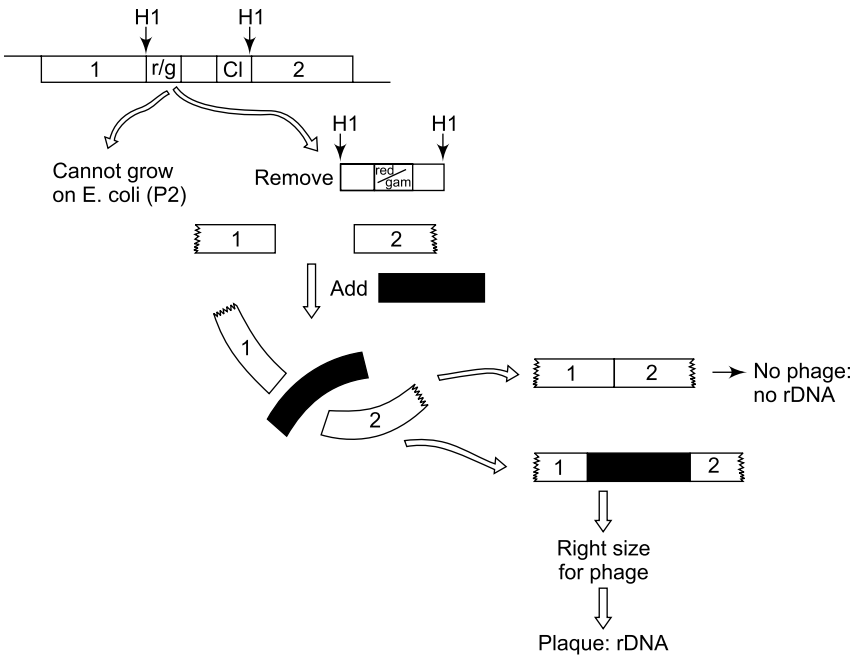


Fig. 4.6 *Spi phenotype selection. The phenotype spi⁺ is possessed by phage lambda (red⁺ gam⁺) that are unable to grow in E. coli (P2); the phage is said to exhibit 'sensitivity' to phage interaction. Spi⁻ (red⁻gam⁻) phages grow on the same lysogens. Recombinants in which the red, gam region has been replaced are, therefore, spi⁻ and can be picked up from clear plaques on lawns of E. coli (P2).*

5. Tetracycline Resistance due to Inactivation of CI

A derivative of pBR322 has been developed by Roberts and associates (1980) in which the promoter of the Tet^R gene is replaced by a piece of DNA carrying the *CI* gene and the λP_R, which are in *cis* with the Tet^R locus. There is a cloning site within the *CI* sequences. When the latter are intact the *CI* protein binds itself to the P_R region and prevents transcription of the Tet^R gene. If the *CI* gene is interrupted by an insert DNA, there is no repressor to block expression from the P_R and the Tet^R phenotype is expressed. Recombinant DNA in a cell makes the latter survive in the presence of tetracycline, but not intact vectors (Fig. 4.8).

6. Selection of the Size of an E. coli Colony

Okazaki (1980) has designed a plasmid vector with a *bla* (β-lactamase or Amp^R gene) sequence inserted in the gene for colicin resistance. The Col E1 plasmid contains both a gene for the toxin colicin and one that protects the cell from its own colicin. The colicin resistance gene is, therefore, also known as the 'immunity' (imm) region of the plasmid.

E. coli cells can be stimulated to overproduce colicin by adding mitomycin C to the medium. If such a cell possesses an Amp^R-disrupted immunity region in the vector, it cannot withstand the effect of the toxin and, so, forms a small colony. Now,

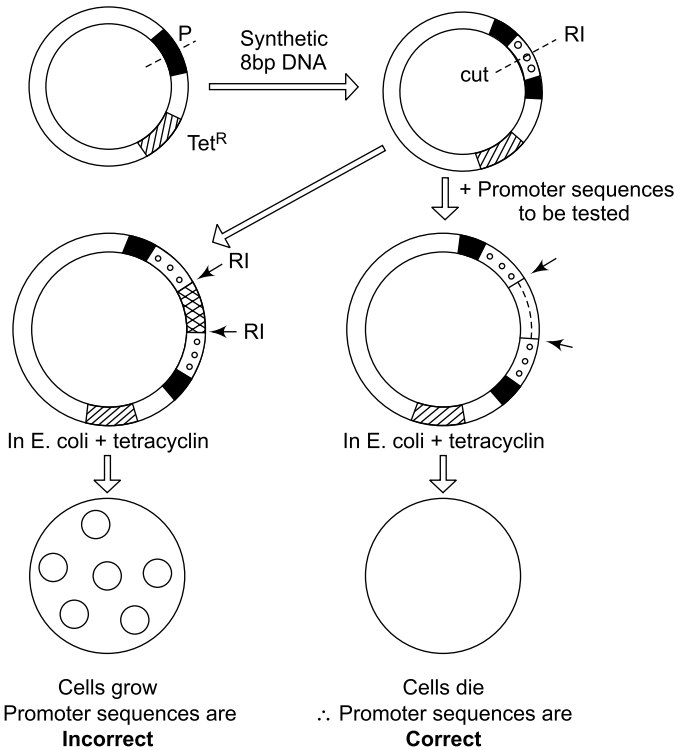


Fig. 4.7 *Selection of correct promoter sequences. Any part of the DNA, genic or non-coding may be examined by rDNA techniques. In this illustration we see a system in which different sequences of a promoter are tested to discover the correct one. Promoters to be tested are cloned in 'portable' vectors, which are inserted within the promoter of the Tet^R gene of pBR322. If the 8 bp long sequence within the promoter is a correct one the Tet^R gene will be expressed; as the test promoter is a correct one the gene will be expressed and render the cells resistant to the drug. Incorrect promoters will fail to initiate transcription of the gene, so that cells carrying them will remain tetracycline sensitive.*

if the foreign gene is cloned within the colicin E1 gene, the toxin is not synthesized even when provoked by mitomycin C. Such cells grow into normal sized colonies. Recombinants are thus picked off as normal colonies on mitomycin C plates.

7. The TK/HAT System

One of the first selectable markers for cloning in eukaryotic cells is based on the availability of the proper enzyme to synthesize thymidine monophosphate (TMP), a precursor of DNA. There are two pathways in higher eukaryotes for the synthesis of TMP (Fig. 4.9). The regular one is synthesized from dUMP (deoxyuridine monophosphate) with the mediation of the enzyme thymidylate synthetase which requires tetrahydrofolic acid as well. This enzyme can be stymied by analogues of folic acid such as aminopterin and methotrexate. However, there is a second or salvage pathway via which TMP can still be synthesized. This depends on an enzyme

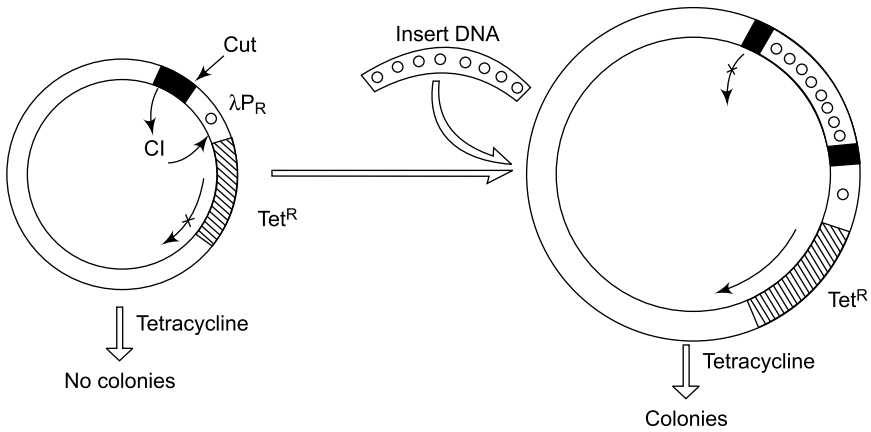


Fig. 4.8 *Tetracycline resistance due to inactivation of the CI gene. A vector has been constructed in which the Tet^R gene is under the control of the phage lambda repressor (CI) gene. The DNA to be cloned is inserted within the CI gene. An inactivated CI gene fails to repress transcription from P_R and so allows expression of the Tet^R gene. Recombinations are, therefore, colonies that grow in spite of the presence of the antibiotics.*

thymidine kinase (TK) that can utilize thymidine if supplied exogenously. If both pathways are blocked, the cell does not make any more DNA and so fails to survive. The salvage route is blocked when the cell is *tk*$^-$.

A selection system has been extensively in use for eukaryotic cells in which the above mentioned facts have been exploited (Litchfield 1964). If the media contain the inhibitor of the *de novo* pathway as well as exogenous thymidine, cells that contain the *tk*$^+$ allele manage to survive, courtesy of the salvage pathway. Such a medium is the HAT which contains hypoxanthine, aminopterin and thymidine. The hypoxanthine is added to allow purine synthesis (adenine and guanine) which is also inhibited by aminopterin. The vector carries a *tk*$^+$ gene.

The TK/HAT system is used to monitor transformation of *tk*$^-$ cells. Such a strain is the L-strain of cultured mouse cells. The recombinant and non-recombinant vectors are mixed with L-cells and first plated on a rich medium. This allows every cell to grow into substantial sized colonies. The medium is then replaced by the HAT one. Only rDNA-containing cells survive in this medium (Fig. 4.10).

Not many genes possess selectable phenotypes. Such genes are selected by cloning them in vectors carrying the *tk*$^+$ gene. The HAT positive cells are further searched for the presence of the gene of interest. Now, there are times where it becomes necessary to distinguish between the vector-specified *tk* gene product and that from the *tk* gene of the cloning cell. It is customary, therefore, to use the Herpes simplex virus-1 (HSV-1) *tk* gene as the marker. It can be discriminated from the host cell gene in a number of ways. Wilkie *et al.* (1979) have cloned the HSV-1 gene on pBR322. The gene is on a 3.4 kb Bam HI fragment (Wigler, 1977).

The TK/HAT system is very much in demand for cloning in eukaryotic cells. It suffers from a common handicap of most selection systems in eukaryotes in being limited for use only with the special L-strain cells.

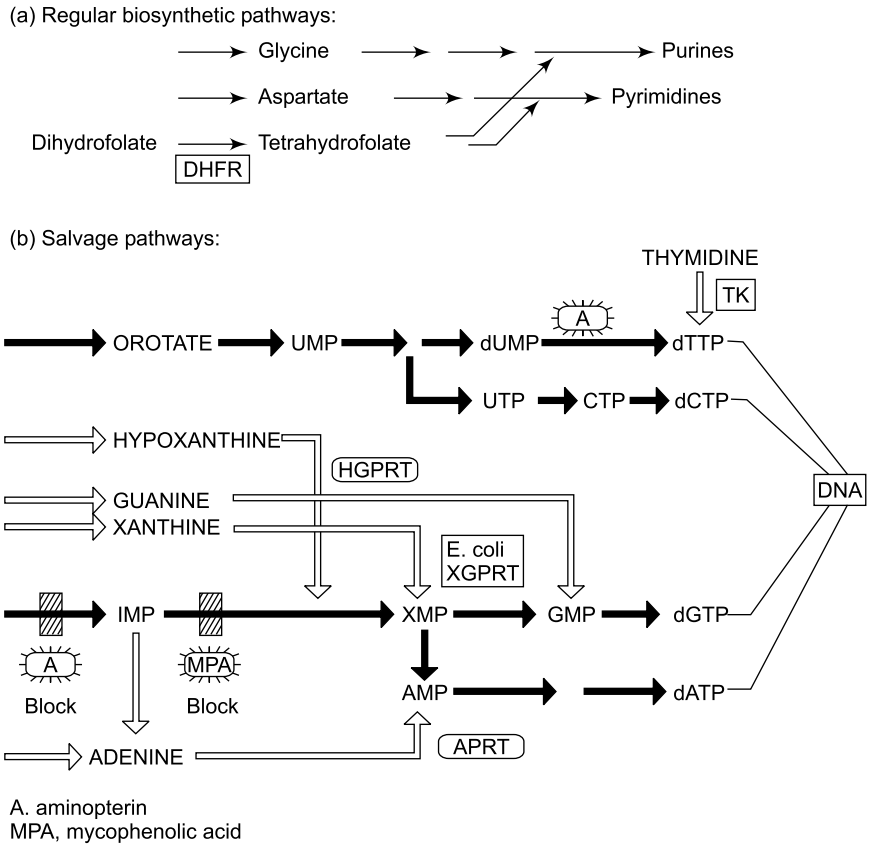


Fig. 4.9 *Regular and salvage pathways for purine and pyrimidine synthesis and sites for metabolic blocks by the inhibitors aminopterin and mycophenolic acid. There are two pathways each for the synthesis of dATP, dGTP and dTTP. If one of them is blocked, the alternate or salvage pathway is utilized. Blocking with aminopterin allows the cell to convert hypoxanthine or guanine into GMP provided the genes hgprt or xgprt are available, together with hypoxanthine. Similarly, instead of using the pathway that begins with asparate, thymidine is used to synthesize dTMP with the help of the* tk^+ *gene. The E. coli xgprt gene is analogous to the animal gpt gene. Aminopterin blocks the action of genes hgprt required for the maintenance of the dihydrofolate-tetrahydrofolate-methyltetrahydrofolate cycle and the action of the thymidylate synthetase (ts) gene that is required both for the above cycle and for the synthesis of TMP.*

The above strategy has been used successfully by Mantedi (1979) to clone rabbit β-globin genes in L cells. Perucho and associates (1980) utilized the *plasmid rescue* method to isolate the chicken tk^+ gene from total chicken DNA, in the following manner: Fragments of total chicken DNA were used to transform L-cells. The tk^+ transformants were isolated, and the DNA of the recombinant vector extracted and spliced to pBR322. These were used to transform cells to the tk^- phenotype. The DNA from the transformed cells was treated with a restriction enzyme that cleaves the chicken *tk* gene and the pBR322 on the same fragment. These restriction fragments were used to transform Amp^s or Tet^s *E. coli* cells. The tk^+ gene was finally delinked from the pBR322 DNA.

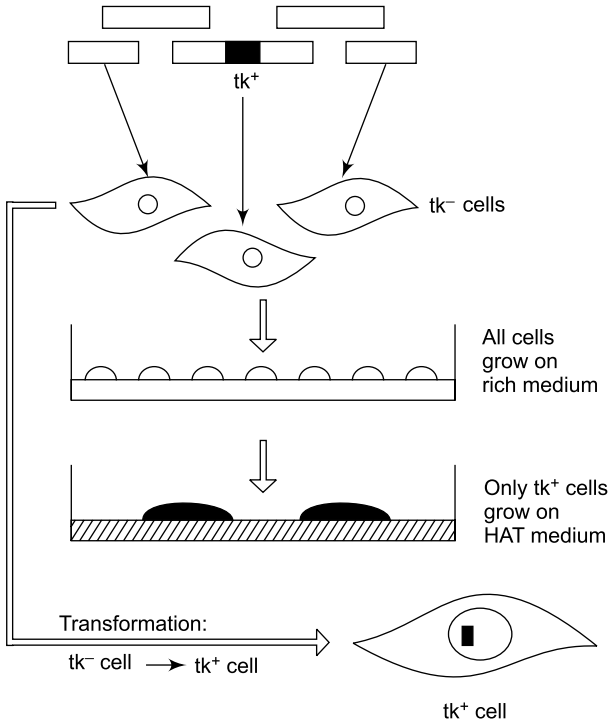


Fig. 4.10 *HAT (hypoxanthine-aminopterin-thymidine) selection of TK⁺ cells. Mouse cells of strain L (TK⁻) are transformed with DNA fragments some of which may possess the tk⁺ gene. All the treated cells are first allowed to proliferate on a rich medium into sizeable colonies. They are then transferred to a HAT medium. Aminopterin blocks the regular biosynthetie pathway for thymidine. However, the tk gene can utilize exogenously supplied thymidine in an alternate pathway. Hence, cells that are TK⁺ will grow on HAT, while TK⁻ cells will not survive on it.*

The TK/HAT system cannot be utilized for studying the expression of all eukaryotic genes in the L strain of mouse of cells, as all genes are not expressed either at the same developmental stage of the individual or in the cells of every tissue. If the *tk*⁺ marker is to be a selection criterion for a broader range of genes, a proper cloning cell that is also *tk*⁻ has to be identified for each one of the genes to be cloned. This is still a tall order. There has, therefore, been a search for other direct selection methods in which a mutation or a developmental stage does not become an essential feature of the cloning cell. One of them is described in the next subsection.

8. The gpt System

As mentioned in the previous paragraph, there is a need for a selection system that is not tied down to a special mutant cloning cell. The *gpt* system satisfies this need.

E. coli possesses a gene for XGPRT (xanthine guanine phosphoribosyl transferase) that converts xanthine to XMP (xanthine monophosphate) which is one of the alternate precursors of GMP (guanine monophosphate). XGPRT also mediates the formation IMP (inosine monophosphate) from hypoxanthine and GMP from guanosine by condensation with phosphoribosyl pyrrophosphate. Eukaryotic cells

do not possess XGPRT and hence cannot utilize xanthine in case the normal route for GMP in eukaryotic cells is blocked.

A vector with a *gpt* (XGPRT) marker can, therefore, be utilized in any cell, and not in one that needs to be a mutant for the marker gene. Cells with or without *gpt* can be distinguished easily by growing them on HAT as well as on XAT (xanthine-aminopterin-thymidine) media. Whereas gpt^+ cells will grow on both these media, gpt^- non-transformants will not (Mulligan and Berg, 1981).

The purine and pyrimidine pathways and inhibitors that block one or the other routes have been further exploited by using mycophenolic acid as the block for the conversion of IMP into XMP. A gpt^+ cell survives in a XAT or HAT medium supplemented with mycophenolic acid. A gpt^- cell will grow only if adenine and guanine are supplied exogenously (see Fig. 4.9).

9. The dhfr as a Selectable Marker

The *dhfr* (dihyrofolate reductase) gene is required for the synthesis of tetrahydrofolic acid, which is an essential precursor of purines, glycine and thymidine. Methotrexate and aminopterin inhibit the action of *dhfr*. Vectors carrying the $dhfr^+$ gene allow *dhfr* cells to be transformed and the latter may be isolated by their lack of resistance to added methotrexate. The concentration of the inhibitor added is sufficient to block *dhfr* expression in the cloning cell. The $dhfr^-$ mutated product has less affinity for the methotrexate; hence, the $dhfr^-$ transformants are resistant to the inhibitor.

Resistance to methotrexate may also be due to a greater level of the *dhfr* product brought about by amplification of the *dhfr* locus.

Subramani and associates (1981) have inserted the *dhfr* gene from mouse DNA into several SV40 vectors, and SV40-derived plasmid vectors.

10. A Phosphotransfersae Selection System

The genes for aminoglycoside 3′-phosphotransferase [*aph* (3′)I, *aph* (3′)II] confer resistance to the antibiotics kanamycin and neomycin in bacterial cells and to the antibiotic G 418 (2-dexoystreptamine) in yeast, plant, *Drosophila* and vertebrate cells. This neomycin resistance gene (*neo*) is therefore a useful selectable marker that may be shuttled between prokaryotic and eukaryotic cells. The *aph* I and *aph* II genes are derived from Tn601 (Tn903) and Tn5, respectively. They do not possess any similarities in their sequences.

11. The CAT System

The product of the gene for CAT (chloramphenicol acetyl transferase) mediates the conversion of chloramphenicol into acetylated derivatives. There is no CAT in mammalian cells, but the latter can express the gene. The gene, derived from Tn9, has been added to a plasmid vector containing the promoter of the SV40 early (T) gene.

Chloramphenicol poisons the eukaryotic mitochondrial system, so that the cells die within a very short time after treatment. A cell carrying a CAT gene can withstand the effect of this toxin, as the latter is made ineffective by the enzyme. Foreign DNA on CAT-containing vectors (e.g., pSV2-cat) can be thus utilized as a dominant selectable marker for transformed cells.

12. Alu Markers for Human Genes

The eukaryotic DNA contains many repeated sequences, many of which are similar enough to be grouped into families. One such ubiquitous highly repeated family is the *alu* (130–300 bp). The name is derived from the presence in the sequence of the recognition sequence (AGCT) for the restriction enzyme *Alu* I.

The *alu* sequences occur more than 30×10^4 times in the human genome; they are positioned at short intervals throughout the DNA, both within and outside the regions occupied by genes, as well as in exons and introns. There are indications that these sequences are results of several transposition events.

Because of their presence in almost any short length of the human DNA, *alu* sequences are useful markers for human DNA inserts in a transformed cell. With due apology to Shelley: If Alu comes, the human gene cannot be far behind.

4.1.3 Indirect Screening Techniques

A somewhat indirect—but a surer way of detecting the presence of the transferred DNA is to identify the DNA itself or the product synthesized due to the inserted DNA. The DNA that leads to the synthesis of an mRNA or protein product need not be the coding sequences alone of a gene. The insertion of a promoter or other regulatory motifs will also result in gene expression if the insert is accompanied with or placed in the proper context of expression signals and coding sequences.

Proteins expressed by the cloned fragment may be identified by very sensitive immunochemical techniques. The DNA insert may be identified by hybridization with a probe (mRNA, cDNA, etc.) or by procedures that pinpoint the presence of the desired DNA. Such procedures include techniques known as 'hybrid-arrested and hybrid-selected translations'.

1. Immunochemical Technique

Immunoglobulins (antibodies) that match the gene product in question are the basis of several tests that identify as well as isolate the protein products of the inserted DNA. Some of these tests are described in Section 4.2.3.

A crude method of identifying the clones synthesizing the desired protein is to lyse bacterial colonies *in situ* on discs of filter paper, and treat the discs with the antibody probe. The latter bind themselves only to the desired protein. In a subsequent step the antibody-antigen (protein) complexes are identified with a labelled antibody probe that matches the same protein (Skalka and Shapiro 1976; Ehrlich *et al.,* 1978; Carbon *et al.,* 1978; and Broome and Gilbert, 1978).

Broome and Gilbert's method consists of lowering an antibody-coated plastic disc on the surface of the petri dish containing bacterial colonies (lysed) or phage-caused plaques. The antibody molecules will bind themselves to the antigen only from clones that are producing it. The disc with the attached proteins is now treated with the radioactively-labelled antibody. The disc, after appropriate washes to remove non-attached probes, is screened for radioactivity. By comparing with the plate of clones, the ones that are positive for the protein are identified and if needed isolated and amplified. Young and Davies refined this method for detection of clones with a foreign gene inserted in a *lacZ* sequence in a gt vector (Figs. 4.11 and 4.12).

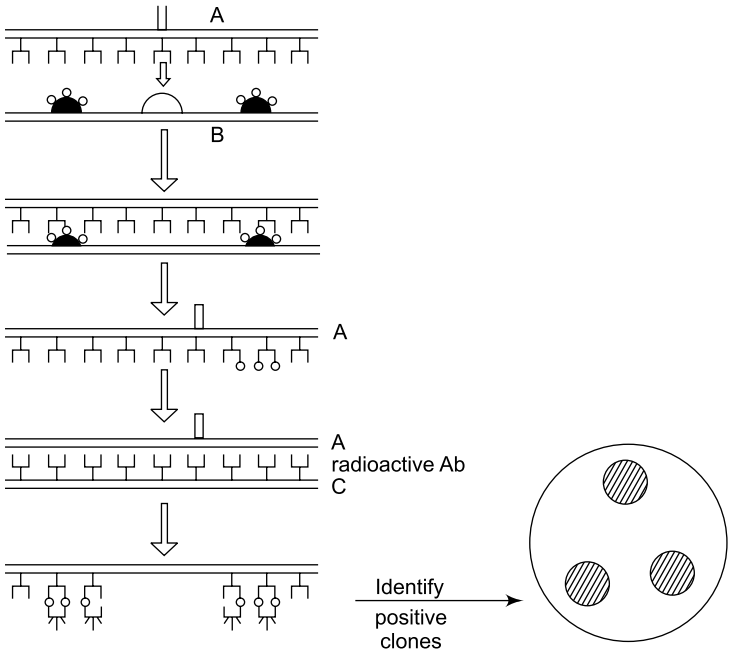


Fig. 4.11 *Immunochemical detection of a cloned gene product. A plastic disc is coated with the antibody matching the gene product to be identified in colonies growing on a petri dish. By placing the antibody-coated surface against the colonies, complexes are allowed to form between the Ab and An of interest. The disc is then placed over a surface coated with the same antibody, but one that is radioactively labelled. The antigen in the initial Ab-An complexes binds itself also to the labelled Ab. The regions of the disc that are radioactive are determined and used to identify the locations of the colonies expressing the gene product of concern.*

2. Hybrid-arrested Translation

When the mRNA transcribed from the inserted DNA forms a small proportion of the heterogeneous mRNA population in the cloning cell, a strategy is employed by which the presence of the cDNA of the gene sought is established (Fig. 4.13). The technique, designed by B. Roberts, includes the following basic steps:

(i) The insert is a cDNA copy of the mRNA of the gene in question. This rDNA is isolated and purified from the selected clones.
(ii) The rDNA is dissociated by heating.
(iii) The total mRNA of the cloning cell is extracted and the dissociated rDNA added to it. The cDNA will hybridize to its matching mRNA.
(iv) The above mixture is added to a 'cell-free translation system' that contains radioactively labelled amino acids and the paraphernalia needed for protein synthesis *in vitro.* Two results can be expected:
  (a) The cDNA has tied up the desired mRNA. In this case the latter cannot be translated.
  (b) If the mRNA is not hybridized, it will be translated.

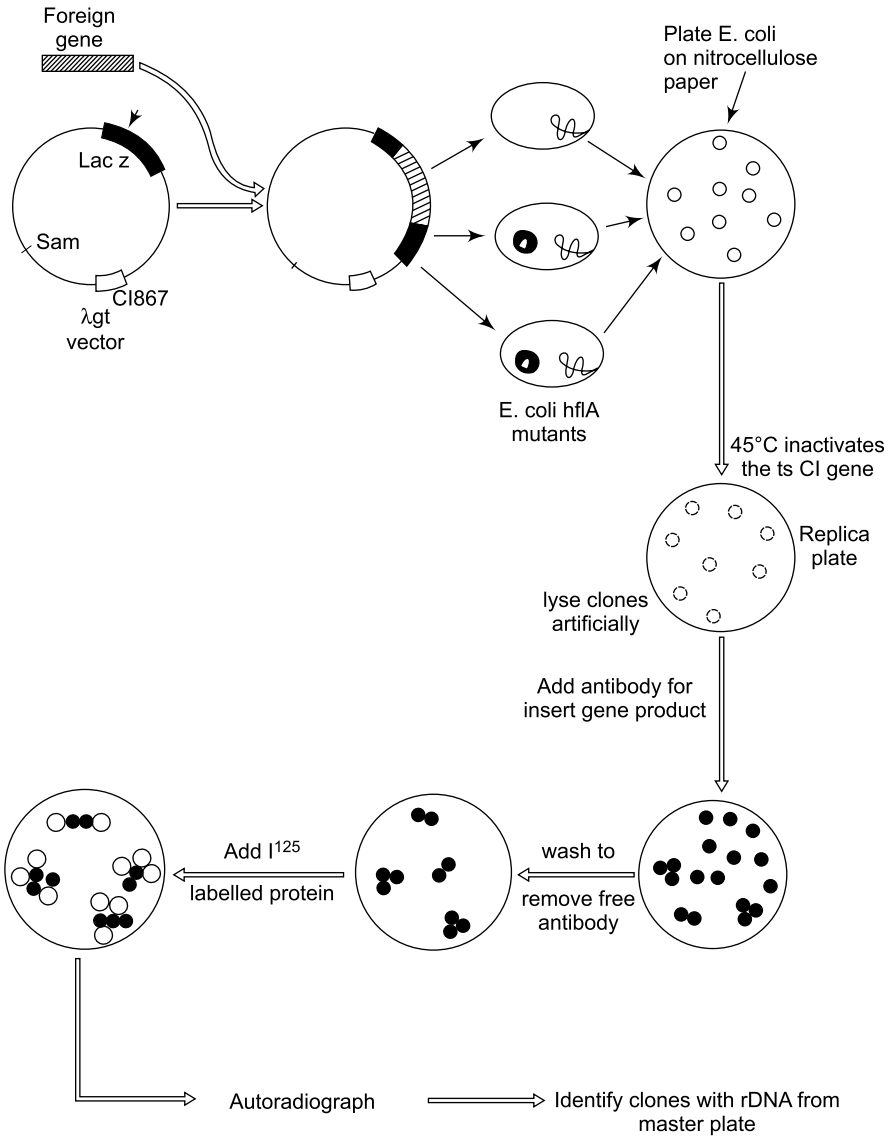


Fig. 4.12 *An improved immunochemical detection technique (Young and Davies). A lambda gt vector with a ts CI gene (CI857) and Sam is used to clone the foreign gene. E. coli cells transformed by this hybrid vector are plated either (i) directly on agar-solidified medium or (ii) on nitrocellulose paper placed over the medium. In the former case, a nitrocellulose paper (disc) is utilized to 'blot' loose cells on to it. Several replicas may be made in this way. In the latter event, the paper with the colonies is treated further. The colonies on nitrocellulose paper are lysed. They are treated to a temperature of 45°C and the colonies lysed. The clones producing the desired proteins are detected with antibody probes.*

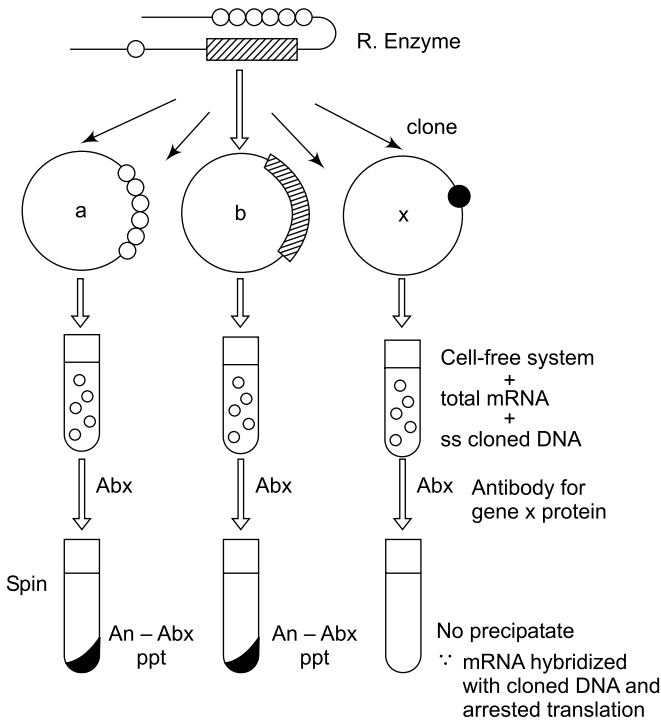


Fig. 4.13 *Hybrid-arrested translation (HART). The mRNA of the gene of interest hybridizes with the sense strand of dissociated cloned DNA fragments and prevents (or arrests) its translation into a polypeptide. Antibodies are used as probes for the product of interest. Lack of precipitates in this system is the basis of selection.*

(v) Antibody matching the desired protein is added to the reaction mixtures.
   (a) The mixtures in which proteins have not been synthesized will have no antibody-antigen precipitates. This indicates that the cDNA has found the complementary mRNA.
   (b) Antibody-antigen precipitates, on the other hand, indicate that the mRNA are free and are not complementary to the cDNA. The clones from which the reaction mixtures were made, and which showed no precipitates, are the ones containing the insert DNA being searched. In short, the arrest of translation by hybrid RNA-cDNA formation indicates the clones that have recombinant DNA.

3. Hybrid-selected Translation

This technique differs marginally from the previous one (Fig. 4.14). The rDNA from recombinant clones are hybridized with the total mRNA from the cells. The hybridized mRNAs are isolated and purified, and the mRNA released from the hybrid molecule. The cleaned mRNAs are now used for translation. The reaction mixtures contain the correct proteins translated from the mRNA transcribed from the desired gene. The proteins are identified by immunochemical or other methods.

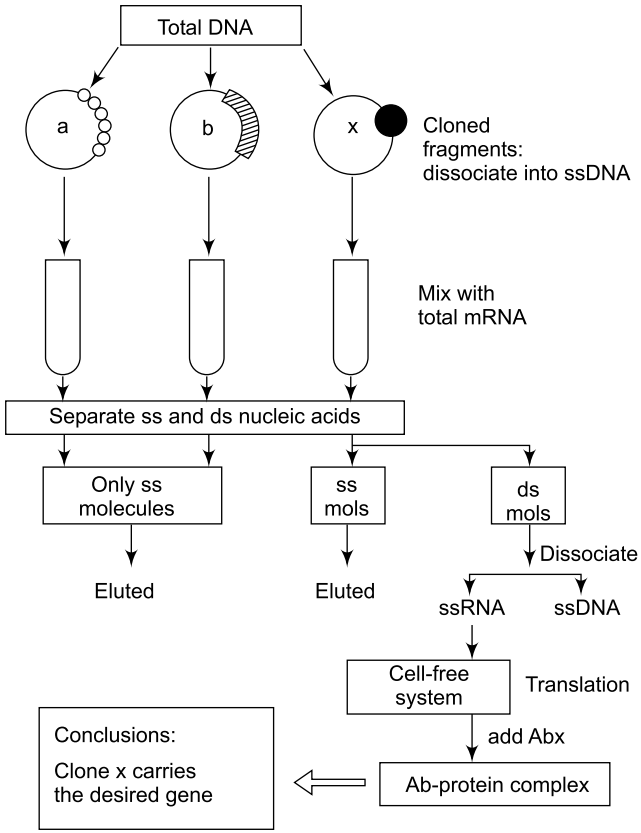


Fig. 4.14 *Hybrid-selected translation (HST). Cloned fragments of DNA to be searched for the gene of interest are dissociated and mixed with the total RNA of the cell type from which the DNA has been isolated. The DNA of concern will hybridize into DNA-RNA duplexes. The reaction mixture is passed through a column which will bind only the duplex molecules. The ss RNAs and unhybridized DNAs will be eluted. After washing the column, the ds molecules are released and dissociated. The mRNA is translated in a cell-free system. If it is the correct one, the antibody matching the protein product will precipitate the latter. The presence of Ab-An precipitates is the basis for HST selection.*

Sobel and associates (1978) and Nagata *et al.* (1980) immobilized the DNA to be hybridized, while Woolford and Rosbash (1979) allowed the hybridization to occur in solution. This is a very sensitive method and can be employed to pick out DNA sequences that are expressed at very low levels (0.1 per cent) in the cell.

4. Rescue Techniques: Plasmid and Others

Rescue techniques both identify and isolate the gene of interest. The techniques involve splicing fragments of genomic DNA to a DNA handle that can be used to advantage either for recombination by hybridization or to supply some other advantage for detecting the DNA fragments joined to the handle.

The first of these methods is the plasmid rescue technique, which was employed by Perucho *et al.* (1980) to isolate the chicken *tk*$^+$ gene from the whole chicken DNA. The technique can be described best by following the steps utilized in this project (Fig. 4.15).

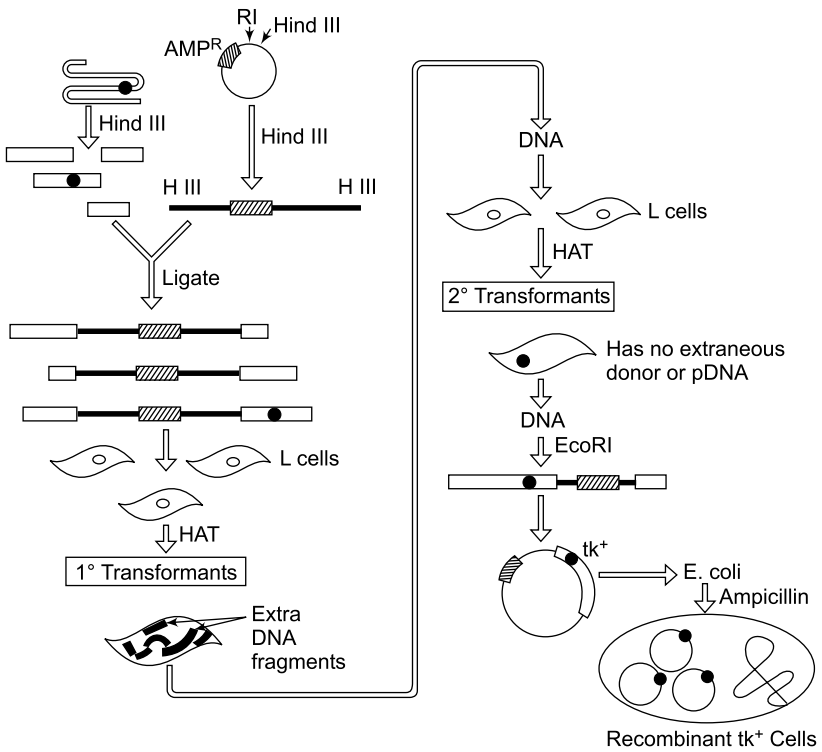


Fig. 4.15 *Plasmid rescue technique. Fragments of DNA, one of which carries the gene of interest (in this case the tk$^+$ gene) are ligated to a plasmid (pBR322), and the hybrid molecules mixed with L-strain cells (tk$^-$) for transformation. Transformed tk$^+$ cells are selected on HAT medium. These primary transformants may possess more than one fragment. It is, therefore, customary to obtain secondary transformants using the DNA from the primary ones. DNA from the last transformant is introduced into E. coli cells. Those which are Amp^R contain both the pBR322 and the required sequences.*

The gene to be isolated, in this case, was the *tk*$^+$ gene, which is itself a selectable marker. Several restriction enzymes were used to find some that did not cut through the *tk* gene but could linearize the plasmid pBR322. *Hin* dIII was one of them. Total DNA chicken was isolated and treated with *Hin* dIII. The fragments were spliced to pBR322 linearized with *Hin* dIII. The hybrid molecules were transfected into mouse L-cells. Those which grew on the HAT medium were amplified into mass colonies. The DNA from these colonies were extracted and Southern blotted. The blots were hybridized with pBR322 probes. The autoradiograms of the probed bands revealed the ones which had both the pBR322 and *tk* gene sequences (the latter because all of these grew on HAT). The DNA from these bands (from the

clones from which the band was derived) was used once more to transform a second batch of L-cells. The DNA from the secondary transformants were cut with an enzyme that does not cut through the resistance genes in pBR322 (*Eco* RI) and the *tk* gene (*Bam* H1). These fragments were used to transform *E. coli* cells and cells that grew on ampicillin-supplemented plates were identified. These were then selected on the HAT medium. Cells that grew in HAT possess the *tk* gene as well as the pBR322 sequences. The DNA from these last transformants was analysed. The *tk* gene together with flanking chicken DNA was found linked to the plasmid DNA.

The preceding technique has been refined by others, especially to increase the capacity of the vector for the donor DNA. It may not always be possible to splice a full gene to a plasmid. Cosmids are better, as up to 45 kb insert DNA may be accommodated. Also, cosmids may be packaged into infectious particles. Lan and Kan (1984) used a cosmid library to look for the human *tk* gene. The DNA from the tranformants was packaged with the packaging mix and the particles tested for their ability to transform L cells of mouse into tk^+ phenotypes.

A modification of the above strategy is the marker rescue technique employed by Pellicer *et al.* (1980) and Low *et al.* (1980) to isolate the *aprt* (adenine phosphoribosyl transferase) gene from the hamster (Fig. 4.16). Also *aprt* cells do not survive on a selection medium containing adenine and azaserine; *aprt* transformants are, therefore, selected as colonies on this medium. The scheme for this technique is similar to the previous one up to the second round of transformation. The DNA from the secondary transformants was cut up into restriction fragments and a bank of fragments was made in Charon4 vectors. These fragments were screened with a pBR322 probe. The clones that were positive were seen to contain part of the plasmid and some of the hamster DNA. This DNA was used to transform $aprt^-$ cells. Those cells which grew on the selection medium had DNA that complemented the defective $aprt^-$ in the cell. The intact *aprt* gene was discovered in the transforming DNA.

A third variation on the rescue theme is the tRNA rescue technique, in which the recognizable nucleic acid spliced to the genomic DNA fragments is a tRNA gene with a *suppressor* (sup F) for amber mutations (Fig. 4.17). In the last round of transformation the DNA from the short listed clones is used to transform an su^- strain of *E. coli.* Fragments that suppress the amber mutation in the *E. coli* are the ones carrying the desired gene. Goldfarb *et al.* (1982) and Chimizn *et al.* (1983) in Wigler's team followed this method to isolate fragments that contained the oncogenes c-Ha-*ras* and N-*ras,* respectively.

The principles underlying the aforementioned techniques have been utilized in various modifications to isolate and delineate regulatory and other functional regions of a gene.

5. Nucleic Acid Hybridization: Southern Blotting

DNA or RNA may be fragmented, fractionated on an agarose gel by ejectrophoresis and the bands transferred (blotted) to nitrocellulose filter paper to which nucleic acids adhere firmly. This blot (called the Southern blot, after the pioneer for this technique) is then flooded with a radioactive nucleic acid probe for the gene to be searched for in the bands. After appropriate steps the washed and dried paper is

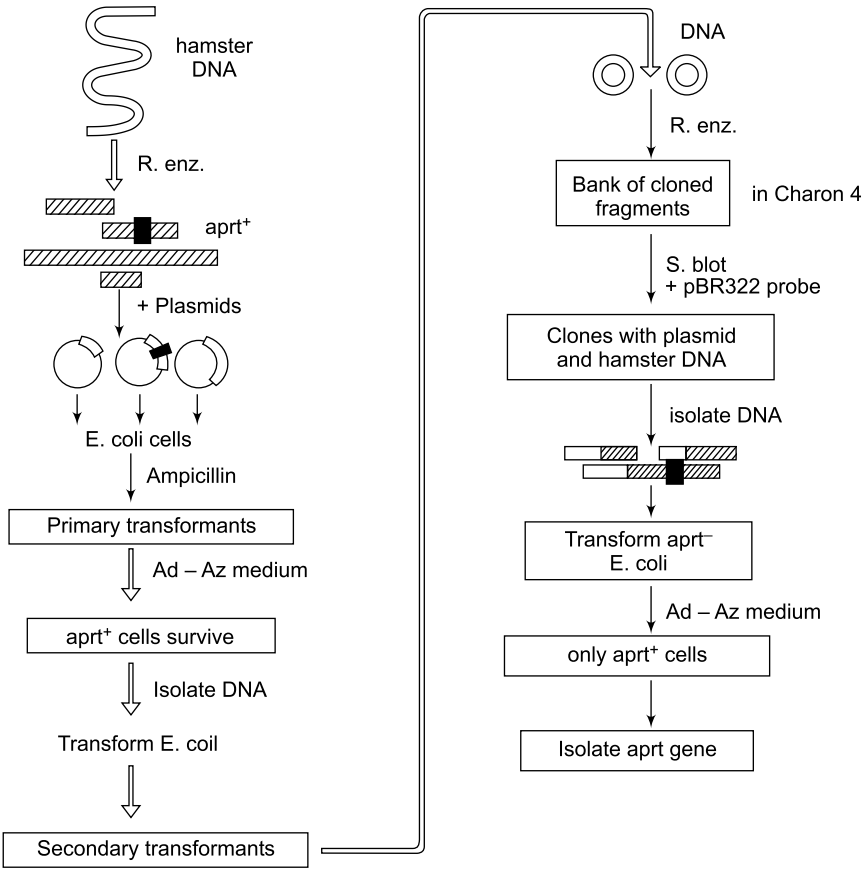


Fig. 4.16 *Marker rescue technique. This technique isolates DNA by virtue of a particular marker gene present on it. In this illustration the marker aprt gene is rescued from DNA fragments, some of which only possess this marker gene. Also, aprt encodes the enzyme adenine phosphoribosyl transferase which mediates the synthesis of AMP from adenine, by a salvage pathway of biosynthesis. The regular pathway for the biosynthesis of purines may be blocked by aminopterin (see caption of Fig. 2.78) or azaserine. A medium containing azaserine and adenine will support cell growth if the gene aprt is present as well. Cells that are aprt⁻ cannot, conversely, grow in the presence of adenine and azaserine. In this technique, cells transformed with DNA fragments ligated to a plasmid are selected on an azaserine-adenine medium. Secondary transformants are screened for the presence of the plasmid vector with an appropriate probe. The clones thus identified contain the aprt⁺ gene.*

autoradiographed. Bands on the print that show the exposure of the photographic film are the ones containing the hybridized probe (Fig. 4.18). The band in a refractionated nucleic acid is eluted and the desired gene or mRNA characterized.

The above procedure was modified to suit the hybridization of RNA and protein molecules. These techniques are described elsewhere as Northern and Western blotting, respectively.

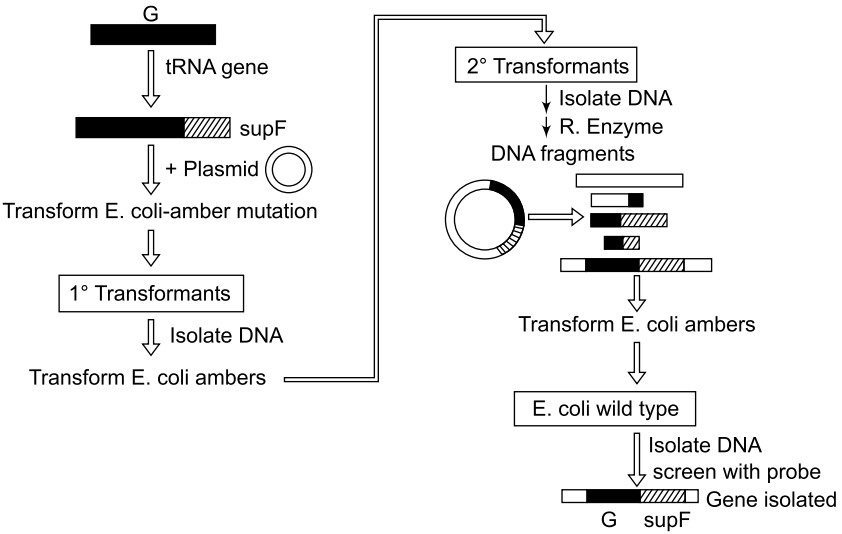


Fig. 4.17 *The tRNA rescue technique. This technique of isolating a particular DNA fragment exploits the ability of suppressor tRNAs to ignore amber mutations. Here supF is the gene for a tyrosine-specific tRNA with a mutation in its anticodon that incorporates a tyrosine residue at an amber (UAG) stop codon. Translation of the mutated mRNA is not prematurely terminated and the phenotype of the cell becomes that of the wild-type allele of the mutated gene. Fragments of DNA, from which a desired sequence is to be fished out, are spliced to plasmids bearing the supF gene. These hybrid plasmids are replicated in an E. coli strain with an amber mutation in, say, a nutritional (auxotrophic) gene. Cells with the hybrid plasmid will behave like wild-type cells and grow in minimal medium. DNA is extracted from secondary transformants and cleaved with a restriction enzyme. The restriction fragments are used to transform a third batch of the same mutated strain of E. coli. Transformed cells should possess the desired DNA fragment. The presence of the latter is confirmed by a suitable radioactive screening technique.*

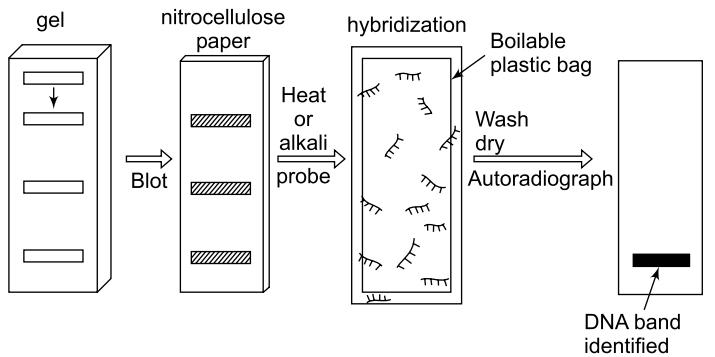


Fig. 4.18 *Southern blotting principle. DNA from an electrophoretic gel is transferred without displacement in position to the nitrocellulose paper. This 'blot' is first treated to denature the ds DNA. The paper is then treated with an RNA (or DNA) probe for the DNA to be detected. Unhybridized probes are removed, and the 'blot' autoradiographed. Only the bands containing DNA hybridized to the radioactive probe will expose the photographic film. The latter is developed. The exposed band will indicate the band in the gel that contains the desired DNA. This technique, developed by E M Southern, introduced a new dimension in investigations based on recombinant DNA techniques.*

It is sufficient to emphasize here that hybridization with a complementary nucleic acid or antibody (in the case of proteins) probe is a very sensitive technique that pinpoints even a small amount of the desired molecule in a given sample.

6. Colony Hybridization: Hogness and Grunestein

In principle, this method is similar to the Southern blotting technique. In the colony hybridization method (Fig. 4.19) first developed by Hogness and Grunestein, (1975) the colonies are allowed to grow on nitrocellulose paper placed in contact with the nutrient agar. When the colonies are about 1 mm in diameter, replicas are made in other paper-covered dishes. The nitrocellulose papers are removed, washed, treated with a detergent to lyse the cells, and heated to 80°C. The exposed DNA bound to the paper is denatured with 50 per cent formamide, treated with the probe and autoradiographed. Colonies that possess the hybridized DNA expose the films. The identified colonies are amplified separately from the master plate and the DNA utilized as required.

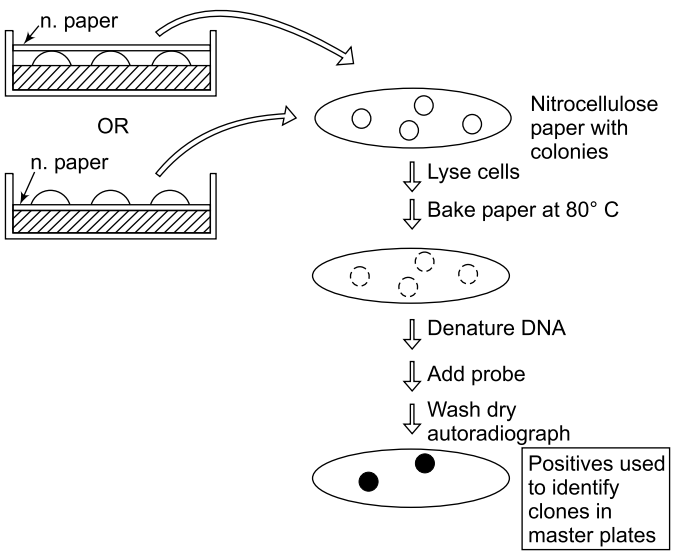


Fig. 4.19 *Hogness and Grunestein: colony hybridization technique. This is a variation of the Southern blotting technique that blots DNA from cells in situ, and identifies colonies harbouring a desired gene. Cells are plated on an agar-solidified medium and 'blotted' with a disc of nitrocellulose paper. The paper is now treated, first to lyse the cells, next to dissociate the DNA to be hybridized with the probe and finally to be autoradiographed. The exposed regions on the developed film indicate the colonies in the master plate possessing the DNA matching the probe.*

7. Dot Blot Hybridization

Small samples of RNA may be assayed by the 'dot blot' hybridization technique. This technique is useful for both tittering the concentration of RNA viruses and for screening a sample for particular species of RNAs.

The Dot Blot Hybridization technique consists of placing drops of the RNA sample (viral genomic or transcripts) on nitrocellulose or diazotized paper, fixing

the RNAs on it by baking, denaturing the secondary structure of the RNAs and hybridizing the 'dots' with given probes. The desired spots are identified as usual by autoradiography. This is a very rapid method for screening RNAs. Quantitative estimates of the wanted species may be made by comparing degrees of hybridization with known or control samples.

Samples as little as 1–5 µl may be used per 'dot'. Several commercial suppliers now carry apparatus that can be used to blot a larger amount (100 µl) of the sample uniformly and simultaneously.

4.2 ❑ PROBES AND TESTS

The exact pinpointing of a particular nucleic acid or protein is one of the important objectives in gene cloning ventures. This is achieved by employing appropriate 'probes' that match the sought-for molecule, and by a radioactive or other label declare the successful hybridization or complexing of the probe and the nucleic acid or protein.

Procedures have been also developed that have revolutionized the manner in which these probes may be utilized, particularly when one is looking for a needle in a haystack. Prominent among these procedures are the Southern blotting technique for probing fractionated DNA and allied ones for RNA and protein, and sophisticated immunodiagnostic tests that are also valuable for the recovery of minute quantities of a unique protein from a hetrogeneous mixture.

4.2.1 Nucleic Acid Probes

A nucleic acid probe is a short or long length of single-stranded RNA or DNA that is complementary to a portion of the nucleic acid to be identified. There are thus RNA, cDNA, DNA and nick-translated ss DNA probes.

A general procedure for preparing a nucleic acid probe is to have a cell-free reaction system containing the template, the polymerizing enzyme, the four NTPs, one or more of which are radioactively labelled, and other factors needed for the reaction to occur. Aliquots of the reaction mixture are removed at regular intervals and assayed for the length of the synthesized strand. This check involves a TCA assay followed by measurement of the degree of radioactivity in the sample. The amount of radioactivity counts are plotted against time. A linear relationship exists between them. The end point is reached when there is no further increase in radioactivity in an aliquot with an increase in the time of reaction. The reaction is stopped and the probe isolated. A TCA (trichloroacetic acid) assay consists of mixing the TCA with a reaction mixture chilled at 0°C, and allowing the mixture to pass through a special filter that retains the new nucleic acid strands. The filter papers are counted in a Geiger-Müller counter.

1. RNA Probes

The first RNA probes were mRNAs of a gene that are abundantly expressed in a cell; these were end-labelled with $^{32}PO_4$.

Cells such as erythrocytes (RBCs) are primarily filled with globin mRNAs. The total mRNA may be extracted by gel electrophoresis. The mRNA of the size expected for the DNA (gene) to be probed is isolated and then end-labelled, by first removing the 5′-terminal phosphate using alkaline phosphatase, and adding a ^{32}P-labelled phosphate with the help of a kinase.

If the proportion of the mRNA to be made into a probe is very low, other strategies are utilized to identify and isolate them for end-labelling. One strategy relies on a differential hybridization preliminary step. This consists of preparing a cDNA bank (see next section) from the total RNA from one type of cell. These cDNAs are then used to hybridize the RNAs in a cell from another tissue or developmental stage, which differs in some unique gene product from the first. The hybridized mixture is run through a hydroxyapatite column which binds itself to double-stranded nucleic acids (here the RNA:cDNA hybrids). The single-stranded unhybridized RNA is eluted. This is, hopefully, the mRNA that is wanted. This strategy was successfully used to isolate a protein that is present in the liver cells of male rats but not in those of female rats.

In both the above strategies, there is the possibility of false identifications. In the former method, although the predominant mRNA is expected to form the bulk of the RNA that is isolated for labelling, there is little guarantee that it does not include other RNAs that may be present in the heterogeneous pool of cellular RNAs.

A far more efficient way of preparing RNA probes involves transcription from a template that is cloned in a plasmid. This transcription is initiated from a promoter that is specifically recognized by an RNA polymerase. Now, RNA polymerases of *E. coli* are rather unspecific. They may initiate transcription from different unrelated promoters. On the other hand, a promoter of a bacteriophage is very specific for the phage polymerase. Green, Maniatis and Melton (1983) decided to utilize one of the phage promoters and the corresponding enzyme to transcribe a DNA cloned in a plasmid vector. The gene for the phage polymerase and the phage promoter were both cloned in a vector, and a cloning site for the template for transcription introduced downstream of the promoter.

The first promoter-polymerase system was taken from phage SP6 that infects *Salmonella typhimurium.* Later, use of the phage T7 (of *E. coli*) promoter-enzyme system came into vogue. The T7 promoter is very efficient and is included in many expression vectors, together with the gene for the T7 RNA polymerase (gene 1). Initially the SP6-containing vector made use of a single cloning site downstream of the promoter site. Later modifications inserted polylinkers very close to the initiation nucleotide. The T7 system uses the gene for the T7 polymerase and the T7 promoter and terminator signals flanking the DNA to be transcribed (Fig. 4.20).

The SP6 and T7 systems can be and have been utilized to express whole RNAs. However, for making a probe, only a short labelled RNA is sufficient. To enable such probes to be transcribed into uniform lengths, it is a practice to linearize the plasmid by cleaving it with a restriction enzyme. The vector thus carries the input in the following order: phage promoter-Enz. l-Enz.2. The template DNA is inserted at the number 1 site, and the composite treated with restriction enzyme 2. We now have a linear DNA with a promoter, the template DNA and an automatic termination site at the end cleaved with enzyme 2. The mRNAs fall off when they reach this end of

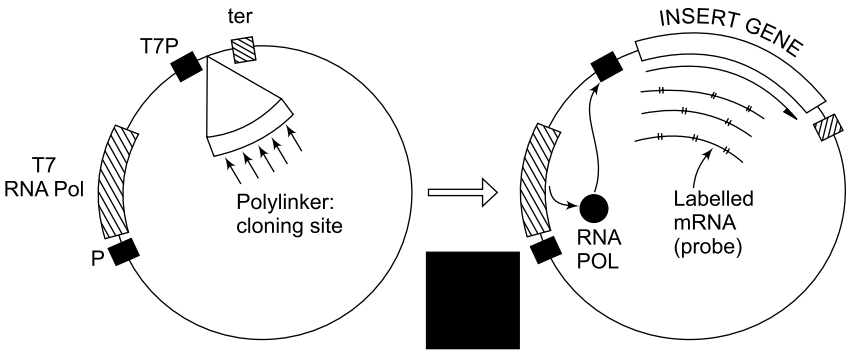


Fig. 4.20 *Synthesis of an RNA probe using the T7 promoter/RNA Pol system. A vector containing the promoter and gene for the phage T7 RNA polymerase, and a cloning site flanked by the promoter and terminator sequences of phage T7 genes, also carries a DNA insert from which the probe is to be synthesized. Phage promoters, in general, and the T7 promoter for late genes, in particular, are more powerful than those of bacterial genes. Cells transformed with rDNA are supplied with RNA precursor nucleotides, of which at least one is radioactively labelled. RNA transcripts from the insert gene are, therefore, labelled, and may be used as probes. The required RNA is isolated after fractionation by size on an agarose gel.*

the vector. Such templates, referred to as *run-off linear templates* (Fig. 4.21) are of uniform size and are easier to isolate from the reaction mixture.

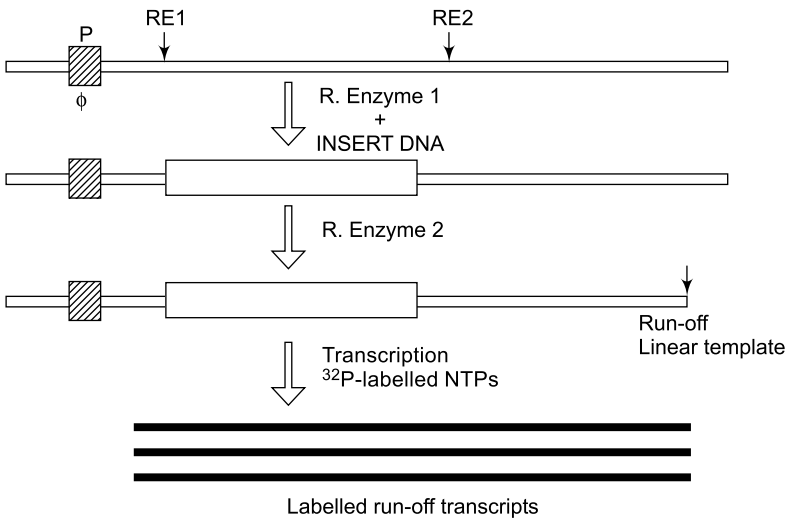


Fig. 4.21 *Run-off linear templates and labelled run-off transcripts.*

Transcription can also be controlled by placing the phage polymerase gene under the influence of an inducible promoter. A T7 expression system has been developed in which T7 polymerase gene is preceded by the *E. coli lacUV5* promoter. The latter is repressed by the *lad* repressor and is induced by lactose or its analogue, IPTG (isopropylthioglactoside). The T7 polymerase is induced by IPTG, and the enzyme

transcribes the cloned template from the T7 promoter. Krieg and Melton reported the synthesis of around 20 mg of probe RNA from 1 mg of vector DNA in one hour at 40°C.

Probes to be utilized for Southern or Northern blotting (described later in this chapter) for *in situ* hybridization, or genomic sequencing (Church and Gilbert) need to be heavily labelled. Run-off transcripts can be made heavily radioactive by including ^{32}P-labelled nucleotides in the reaction mixture. End-labelled probes are far less labelled than the transcribed ones with labels at several nucleotides in the strand.

2. cDNA Probes

If a suitable mRNA is available, it can be copied into a cDNA (complementary) with the help of reverse transcriptase (RNA-dependant-RNA polymerase). The enzyme used routinely is from the AMV (Avian Mycloblastosis Virus), which does not possess the exonuclease activity.

The polymerase requires a primer. In a eukaryotic mRNA, the polyadenylated tail may be utilized as one. A poly dT oligonucleotide primer (about 12 Ts) is synthesized chemically. It aligns with the poly dA region and transcribes the entire mRNA. Specific primers may also be synthesized chemically by deducing the base sequence of the probable primer from the amino acid sequences of the gene product (polypeptide). As the genetic code is degenerate, it is not possible to arrive at the exact sequence. In practice a mixture of the alternative sequences are made, one of which is likely to be the correct one. Methods of preparing oligonucleotides have improved and it is now possible to come very close to the actual sequence in the gene.

3. DNA Probes

If the gene to be searched *in situ,* or from clones of a gene library, is already cloned, the probe may be prepared from the vector carrying the gene. The vector DNA is dissociated and the cloned region cleaved out. The latter is then end-labelled with ^{32}P at the 5′-terminals. On dissociation of this fragment, a probe for the DNA and one for its mRNA will be available.

4. Nick-translated Probe

One way to obtain short lengths of DNA probes is to nick a ds DNA with DNase I, and initiate strand synthesis at each 3′-end in the nicked single strands. The *E. coli* DNA Pol 1, that has the 5′-3′-exonuclease activity, is generally used as the enzyme. The latter has the ability to add a new nucleotide at the 3′-cut-end while removing the nucleotide ahead of it on the other side of the nick (Fig. 4.22). In this way one of the strands between two nicks becomes replaced by a new strand. If radioactively labelled nucleotides are provided for the synthesis, the new strand becomes labelled. This strand is said to be 'nick-translated'. This term should not be confused with the term 'translation' used for protein synthesis. To obtain a uniformly labelled probe it is better that all four NTPs be radioactively labelled. Otherwise, depending on the frequency of a particular nucleotide in a short stretch of DNA, a label in only one NTP will provide very non-uniformly labelled probes.

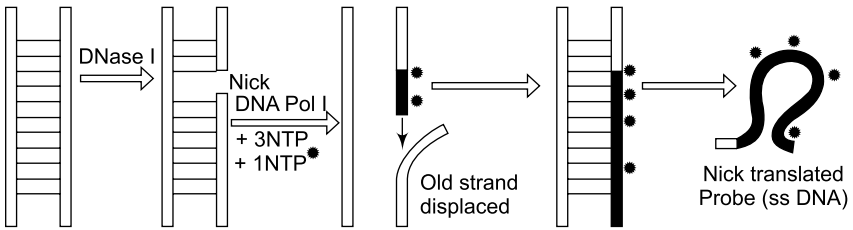


Fig. 4.22 *Nick-translated probes. Short ss DNA probe molecules may be synthesized by exploiting the capability of the DNA polymerase I (essentially a DNA repair enzyme) to initiate DNA strand elongation from a free—OH end at a nicked region in a ds DNA. As the chain gets elongated, the strand ahead of it is degraded by the DNA Pol 1, while removing the nucleotide with the free 5′-end. The DNA synthesis is carried out in the presence of at least one radioactively labelled deoxyribonucleotide, so that the nick-translated DNA are labelled. The term 'translation' refers to the movements of the discarded and newly synthesized strands, and not to the usual term for decoding of the mRNA into a protein.*

4.2.2 Hybridization Techniques

The most accurate way of recognizing a specific piece of DNA or a clone is to search for it with an appropriate probe. The probe (mRNA, cDNA or nick-translated or other DNA probes) will be hybridized only with a complementary strand.

The technique for hybridization has been streamlined by E M Southern. The Southern blotting technique was originally developed for hybridizing DNA with probes. Later, it was modified to suit hybridization of RNA and proteins. These modifications acquired the names Northern and Western blotting techniques, respectively. These are extremely sensitive techniques for instituting searches for even very rare target molecules.

1. Southern Blotting

The Southern blotting technique consists of transferring DNA fragments from a gel slab to nitrocellulose (cellulose nitrate) filter paper (0.45 mm) hybridizing them with a radioactive probe, and identifying the hybridized DNA by autoradiography.

The technique is based on the strong affinity of DNA for nitrocellulose paper. The DNA to be searched is fractionated by agarose or polyacrylamide gel electrophoresis. The bands of DNA in the gel slab are eluted into a closely apposed nitrocellulose paper by a moving front of buffer solution. The adsorbed DNA becomes immobilized on the paper in the positions in which they are transferred.

Hybridization is possible only if the DNA signal on the paper is single-stranded. The gel is, therefore, treated before the transfer step, with a mild alkali (5M NaCl and 0.5M NaOH), and then neutralized using an NaCl solution in a tris-HCL buffer (pH 7.4).

The gel slab is then placed in an assembly having the following order of components (Fig. 4.23), commencing from the bottom: buffer in a suitable tray-like container, a glass support a little larger than the size of the gel, a Whatman's 3 mm filter paper, the gel slab, a nitrocellulose paper of the same size as the gel surface, a second Whatman's 3 mm paper and a stack of absorbent paper.

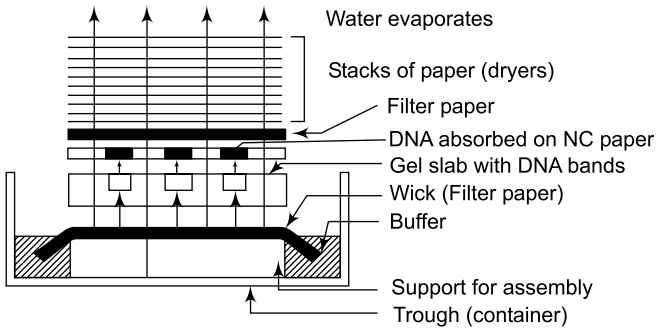


Fig. 4.23 *Southern blotting assembly. The upward capillary movement of the buffer through the gel slab pushes the DNA of each band in the gel into the overlying nitrocellulose paper. The DNA binds itself to the paper, while the buffer solution moves upwards and evaporates. In this way, the DNA bands are blotted exactly on to the nitrocellulose paper.*

The buffer moves up through the gel towards the dry filter and absorbent papers. In its passage through the gel, the DNA fragments become eluted. The latter pass into the nitrocellulose paper. The buffer solution moves upward but the DNA remains fixed to the blot paper.

In case the DNA fragments to be blotted are longer than 10 kb, it is customary to cleave them partially for more efficient transfer into the blot paper. The partial cleavage may be achieved by depurinating the DNA. This is brought about by soaking the gel in HCl prior to the alkali treatment.

To fix the dissociated DNA fragments firmly, the blotted paper is baked at 65–80°C in a vacuum oven, for about 2–3 hours. The Southern blot is now ready to be screened with a probe.

Hybridization with the probe is performed by placing the blotted paper in a boilable plastic bag, containing a measured volume of the solution, with the probe (pre-hybridizing and hybridizing solutions). The prehybridizing solution contains Denhardt's reagent and some denatured heterologous DNA. The hybridizing solution contains the probe in SSC (standard sodium citrate) or SSPE (standard sodium phosphate-EDTA) buffer. Denhardt's reagent, which contains Ficoll polyvinyl pyrollidone and BSA (bovine serum albumin), prevents ss DNA from binding itself to the filter paper without hampering hybridization of the bound ss DNA with the probe.

The filter paper is washed and dried by appropriate methods and autoradiographed. As the patterns of bands on the nitrocellulose paper are identical to those which were on the agarose gel, the location of the DNA containing the wanted DNA sequences is identified easily by looking for the band that has exposed the film.

Hogness and Grunestein's colony hybridization technique is identical to the Southern blotting one, except that colonies are grown directly on nitrocellulose paper. Lysing of the cells in the colony is followed by an alkali treatment to dissociate the cellular DNA, and the paper is further treated as for Southern blots.

Plaques are also hybridized using a modification of the colony hybridization technique. The filter paper, when placed on a petri dish with the plaques to be

screened, picks up phages and naked DNA from the plaque regions. Several replicas may be made in filter paper discs from the same master plate.

At first only DNA could be Southern-blotted effectively. RNA was found not to adhere to the paper as well as DNA. Later modifications allow RNA also to be blotted efficiently on to nitrocellulose paper.

2. Northern Blotting

RNA bands can also be transferred from electrophoresed gels to a solid substrate without displacement of the positions of the bands. Initially, nitrocellulose paper was used, and the technique acquired the modified name of Northern blotting.

RNA molecules are mostly single-stranded, but due to their secondary structure, they possess several regions of ds RNA.

This requires straightening out, and can be achieved by using one of the following strong denaturing agents: (i) glyoxal, alone, or in combination with DMSO (dimethyl sulphoxide); (ii) formaldehyde and formamide, (iii) methyl mercuric hydroxide; or (iv) urea.

RNA denatured by glyoxal can be directly blotted on to nitrocellulose paper. The use of the other agents requires a few additional steps before effective transfer can be made.

Of the denaturing agents mentioned above, methyl mercuric hydroxide reacts with free radicals of polyacrylamide gels; hence, only agarose gels may be used for use with this denaturing agent.

Nitrocellulose filter paper binds itself strongly to denatured RNA, but not to those with a secondary structure.

Another solid support for transfers of nucleic acids from gels has been developed that binds equally well both denatured and non-denatured nucleic acids. It was initially employed for blotting RNA, but subsequently found to be equally useful for DNA transfer. The advantage of this technique is that the blotted paper may be reused several times for different probes, without appreciable loss of the nucleic acid signal, unlike nitrocellulose paper which is not reusable. The alternate filter paper is DBM or diazo-benzyloxymethyl paper.

DBM can be converted from ABM (amino benzyloxymethyl) paper just before use. ABM is available commercially, or may be prepared in the laboratory by the following procedure. Whatman's paper is treated sequentially (with appropriate washings and dryings between steps) with n-nitrobenzyl-oxymethylpyridinium chloride, benzene and acetic acid. The paper is finally desiccated and stored at 4°C. DBM paper is prepared by treating ABM paper with $NaNO_2$ dissolved in HCl, and later washed with a chilled sodium borate buffer. The paper is usually referred to as the diazotized paper.

Whatever the paper utilized for transfer of the RNA, the ethidium bromide stain for nucleic acids is added not to the nucleic acid, but to the latter. Acridine orange may be used to stain the glyoxal and formamide-treated gels. If methyl mercuric hydroxide treatment is employed, the reagent is removed before the blotting step; a treatment of the gel with a 2-mercaptoethanol and NaOH-containing reagent accomplishes this goal.

3. Western Blotting

The technique of transferring protein fractionated on a gel to a solid support inevitably became dubbed as 'Western Blotting'.

In the case of protein blotting the paper to be used is generally pretreated in a manner that is appropriate for identifying or probing for a particular protein. The analytical methods used for such identification include: (i) immunodetection, (ii) binding of proteins to ligands immobilized on the filter paper, and (iii) binding of tagged molecules to proteins immobilized on the filter paper.

Proteins are usually fractionated on polyacrylamide gels, and transferred to nitrocellulose or DBM paper that has been appropriately pretreated. One precaution has to be taken if DBM paper is used. As glycine reacts with the diazonium groups in DBM paper, the latter has to be washed free of glycine.

For immunodetection of a protein on the paper, the DBM is treated with the antibody for the protein (antiserum), which may be labelled or unlabelled. If unlabelled, another labelled antibody that matches the immobilized one may be employed after the proteins are transferred to the paper. The labelled antibody probe will form complexes with the unlabelled ones. The complexes may be identified by a variety of ways.

The second method of detection involves the interaction of a ligand immobilized on the DBM paper with the protein to be detected. The ligand may be a lectin, an antibody or an antigen. The transferred protein has an affinity for its matching ligand, and so forms a complex with it. The latter may be detected by autoradiography if the protein is radioactively labelled. This immunodetection technique is a variant of the filter paper affinity one, where immunodiagnostic methods are employed to identify the wanted molecules.

The third technique reverses the positions of the protein and the binding molecule with respect to the blotting paper. In this case, the protein is immobilized on the paper, and treated with the labelled molecules (^{32}P-tagged nucleic acids or ^{125}I-tagged proteins). Protein-protein binding may be detected by labelled antisera, and protein-nucleic acid associations by autoradiography.

4.2.3 Immunodiagnostic Probes

Antibodies or immunoglobulins constitute the main ammunition of the defence system of the animal body. Antibodies are also excellent biomolecular reagents that may be used to identify and isolate even very small quantities of antigens from a heterogeneous mixture. The ability of the antibody to bind itself to a matching antigen is exploited in a variety of techniques employed in immunochemistry and immunodiagnostic tests.

Immunodiagnosis has become an essential tool in molecular biological investigations. It is almost inseparable from most DNA cloning exercises. The methodologies and approaches of immunology are so different from those of other experimental biological systems that a little introduction to the basis of the immunological tools and familiarity with the paraphernalia of immunological exercises are necessary components of a course in genetic engineering. Let us take a quick look at some of the relevant terms and components utilized in immunodiagnostic probes and tests.

The probe may be the protein itself, or the antibody it provokes in a foreign body. The protein or the other molecular complex which provokes the generation of an antibody is usually referred to as the antigen. The serum containing the relevant antibody is usually called the antiserum (plural antisera).

1. The Antigen

The immunoglobulin that matches an antigen does not necessarily cover the entire molecule or complex. It is usually a smaller moiety on the larger antigenic molecule or complex. The immune-provoking smaller moieties are the antigenic determinants. A large antigen may possess more than one type of antigenic determinant. A cell or virus or an entire organism, such as protozoan and fungal parasites and pathogens possess several antigenic complexes on their surfaces. They are, therefore, sometimes referred to as 'immunogens'.

Each antigenic determinant provokes the synthesis of an immunoglobulin molecule that binds itself only to this determinant, and not to others unless the others are very similar to the provoking antigenic determinant. In the latter case, the antibody is said to cross-react with two or more types of antigens; conversely, such an antibody identifies antigens with close homologies in their make-up.

When a poly-determinant-studded antigen is introduced into the bloodstream of an animal, a heterogeneous mixture of antibodies is synthesized in the latter, each immunoglobulin matching one of the determinants in the immunogen. The proportions of the different antibodies varies from animal to animal, and from condition to condition, in the same animal. For general diagnostic purposes, antisera with such *polyclonal antibodies* are adequate as there is always at least one population that overwhelmingly represents a particular determinant and so reacts with the antigen in question.

If, however, a very specific probe is required, one has to make use of a *monoclonal antibody* (Mab), that has been raised specifically against a single determinant. Mabs are synthesized from tumours developed from a hybrid cell with special parentage; such a tumour has been named a *hybridoma* and the body of knowledge in synthesizing Mabs *in vitro* is known as *hybridoma technology* (Fig. 4.24).

To be effective, the antigen used must be reasonably pure. Even biochemically pure proteins and other antigens (complex carbohydrates, nucleic acids etc.) may contain traces of impurities that possess antigenic qualities.

There are several ways of purifying antigens. One of the most efficient ways of selecting a particular antigen from a mixture (such as membrane proteins or subunits of an enzyme) is that of fractionating the proteins in SDS-PAGE (sodium dodecyl sulphate polyacrylamide gel electrophoresis). The band corresponding to the expected molecular weight of the desired protein is cut out from the gel, the protein eluted from it, and used as the antigen. If the band contains more than one protein in the same range of molecular weight, two-dimensional gel electrophoresis may be employed to disperse the proteins in a band by a fractionation by electrophoresis in a second dimension (at right angles to the first).

The antigen is injected into the animal for raising antibodies. In practice, the antigen is introduced concurrently with a material that enhances the production of antibody and/or prolongs the life of the antigen in the body. This additional

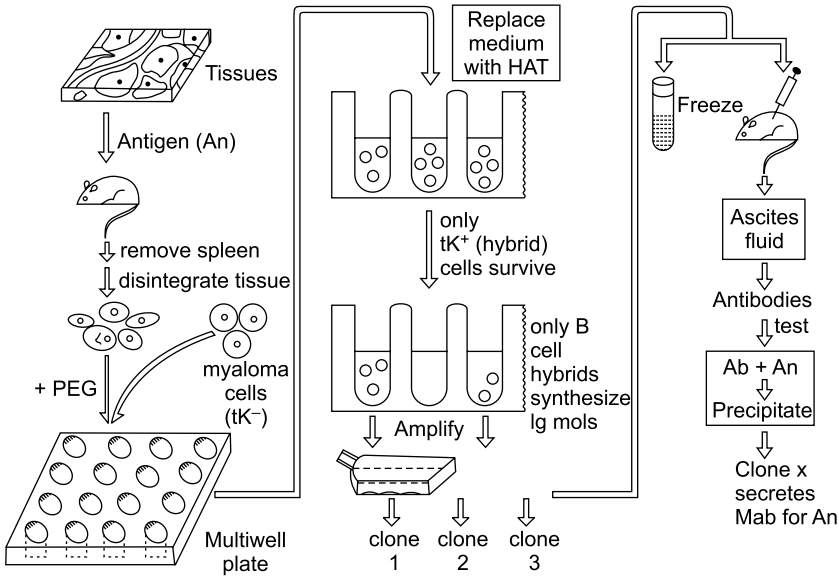


Fig. 4.24 *Hybridoma technology: the technique for the production of monoclonal antibodies (Mabs). Hybrid cells possessing the characteristics of continuous growth of cancerous cells and expression of a specific antibody of a B cell form tumours (hybridomas) that secrete antibodies made by a single clone of B cells. The parents for the somatic hybrid cells are a myeloma cell line and B cells stored in the spleen of a rabbit injected with a specific antigen. The fused cells are cultured in small volumes of media in the wells of a multiwell plate, each well containing one fused cell. The medium when replaced with HAT selects for TK cells. These are amplified in T flasks (bottles with larger sides flat, on the inner surface of which normal animal cells get attached and divide until the surface is covered with a single layer of cells). The cells are stored frozen or are introduced into peritonial cavities of rabbits. The ascites fluid that develops contains clusters of cells or small tumours. The ascites fluid is withdrawn and tested with the antigen of interest and the ones secreting the matching Mab are identified.*

material is called an 'adjuvant'. It may be a non-antigenic material such as mineral (paraffin) oil or one that itself elicits antibody formation, such as the tubercle bacilli (Mycobacteria). The adjuvant may be prepared in the laboratory although it is more usual to employ readymade ones. The most frequently utilized adjuvants are Freund's complete and incomplete adjuvants (FCA, FIA). Both contain an oil and an emulsifier. The FCA contains, in addition, killed mycobacteria. The first of the weekly series of injections with antigen usually uses FCA while subsequent immunizations employ FIA.

Very little antigen is required to raise appreciable amounts of antibodies. A rabbit may be injected with 0.1 mg of a protein in each of four weekly doses, to obtain enough antibodies for use. Even as little as 1 mg of a protein attached to a solid support can induce the synthesis of sufficient antibodies. Very small molecular groups that do not elicit antibody response, can be made antigenic by conjugating them to a protein base. There is sufficient evidence (Medawar) that it is the small group, called the *hapten,* and not the supporting protein that provokes antibody synthesis.

2. The Antibody

Antibodies are immunoglobulins, which are proteins with a very unique structure and function. In human beings, there are five types of immunoglobulin (Ig) molecules: IgM, IgG, IgA, IgE and IgD. The bulk of the Ig molecules in the body is due to the first two.

The different Ig molecules differ in the number of components and other variations but all of them possess a common prototype unit. IgG has only one of these units, and will be used here to describe an Ig molecule.

The molecule of IgG (Fig. 4.25) consists of a pair of identical subunits, each consisting of a long (heavy or H) and a shorter (light or L) polypeptide chain. The H and L chains are held together by S-S bonds; so are the the two H chains. The L and H chains of each subunit are aligned from the amino ends of the polypeptides. Each L and H chain can be divided into two distinct regions. The amino-terminal region is known as the 'variable or V region', and the remaining portion as the 'constant or C region'. The V regions of the H and L chains of one unit bind themselves to the antigen. The amino acids in them possess a special affinity for the antigenic determinant that provoked antibody synthesis. The C regions have functions other than engaging the antigen; one of them is to get attached to surfaces of different cells or to certain other molecules of the immune system (e.g., complement CIq).

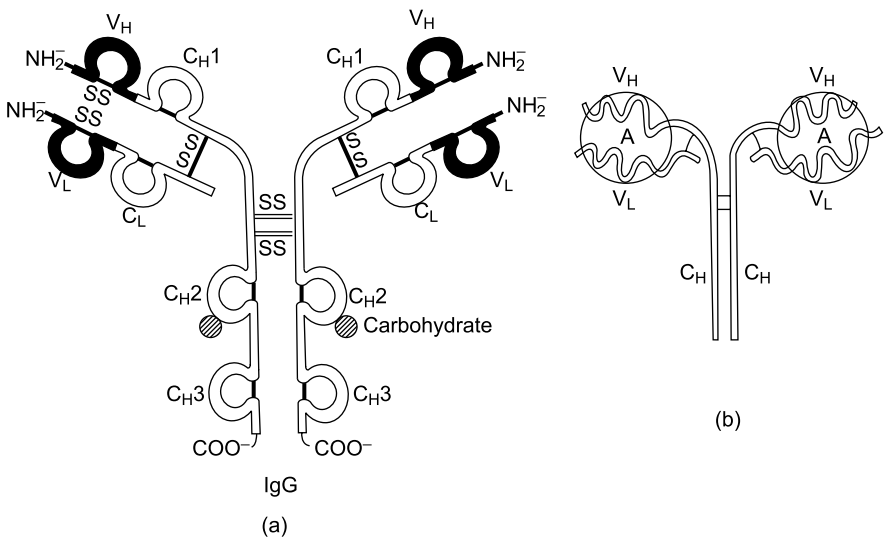


Fig. 4.25 *Diagrammatic representations of a molecule of immunoglobulin G (IgG), and of antibody–antigen binding. IgG is one of the five human immunoglobulins. It possesses the prototype unit consisting of two identical copies of a pair of light (L) and heavy (H) chains (polypeptides). The amino-terminals of the H and L chains are the antigen-binding sites.*

The genetics of antibody response and diversity is fascinating and extremely complex. Much of the knowledge in current immunology has been acquired by the use of recombinant DNA techniques. This, however, is not the forum for delving into these techniques.

It is sufficient to know that the antibody elicited by an antigen is highly specific for the latter, and either the antibody or the antigen may be recruited as a probe to search for the other. The probe (antigen or antibody) may be labelled suitably if the test requires such tagging.

4.2.4 Immunological Tests

Immunochemical methods for detecting antigen-antibody complexes, and for estimating the amount of the antigen (or antibody), in the system make use of either the biological or the physical properties of the immune system components. Biochemical techniques are used for quantizing antibody-antigen complexes: the complexes are precipitated in a 2 per cent solution of PBS. A second technique uses serum protein (complement) Clq immobilized on a solid substrate (plastic tube), to bind to Ig molecules that have complexed with the antigen. A radioactively tagged anti-Ig molecule is then used to identify and estimate the IgG of the complexes.

The antigens (protein products of cloned genes) to be detected in recombinant DNA essays are very little in quantity, and require rather sensitive tests that can be performed on minute amounts of a substrate. The principles underlying some of them are described below.

1. The Precipitin Test

When a polyvalent antigen solution is mixed with its matching antiserum (antibodies in serum), the two proteins combine to form a net-like spongy aggregate that precipitates. When there is more antigen than antibody in the mixture, the antigens usually form complexes that remain in solution. When the antibody is in excess, the complexes formed are insoluble. Some human sera form soluble complexes even in the presence of excess antibody. When the amounts of antibody and antigen are equivalent, there is a rapid aggregation and precipitation of the complexes.

Thus, by adding an equivalent volume of an antigen to that of its antisera, one obtains a preliminary idea about the ratios of the matching proteins in the mixture. If one solution is layered carefully over the other, within one day, a thin cloudy line will be seen at the junction of the two, if the antiserum matches the antigen. A quantitative estimate of the amount of antibody in the serum is obtained by adding different dilutions of the antigen to the latter at 37°C and then at 4°C. The precipitated complexes may be centrifuged out, washed in saline and tested for the quantity of antibody in the complexes.

The above is known as the 'precipitin test' (Fig. 4.26).

2. Immunodiffusion Technique

Titres of antisera may be obtained by immunodiffusion techniques. As in the case of the precipitin test, the antigen and antibody are allowed to interact and form complexes at the interface of the two reactants. However, in immunodiffusion tests, the interfacing occurs between antigen and antibody that diffuse through a solid medium, such as an agar or agarose gel.

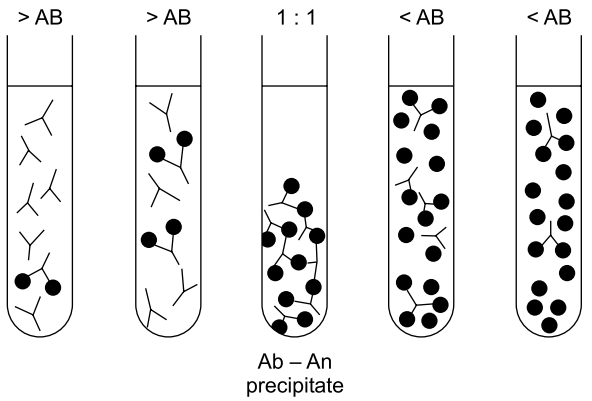


Fig. 4.26 *The precipitin test. This is the simplest immunological test to determine the presence and proportion of an antibody in a mixture of the antibody and its matching antigen. Precipitates of An-Ab complexes are formed when the two components are in equal proportions. Excess or paucity of antibodies in relation to antigen concentration results in soluble complexes. If a series of antigen solutions of graded dilution are used for the test, an estimate may be made of the antibody concentration in the test sample.*

In the earliest of such tests, the antibody was dispersed in the gel and the antigen placed in wells cut out from the gel. The antigen diffuses out from the well and interacts with the antibody to form complexes. The complexes appear as a ring of precipitates around the well. The width of the ring may be used to calculate the amount of antibody relative to the antigen present in the gel. The ring appears as an opaque band in a gel illuminated against a dark background. The ring may also be visualized by staining the precipitates with a protein stain, such as Coomassie brilliant blue.

The above-mentioned single diffusion technique has been superseded by a double diffusion one developed by Ouchterlony (Fig. 4.27). In this case, the antigen is placed in a well with the antisera in wells surrounding it or vice versa. The radially placed wells may contain different antisera, or different dilutions of the same antibody. The advancing wave fronts of the diffusing antigen and a antisera form precipitated complexes at the zones where they meet. The precipitate is soluble if there is more antigen than antibody. Hence, a very sharp precipitin line is made at the zone where amounts of antigen and antibody are equivalent. The relative distance of the precipitin line from the two reactants indicates their relative strengths.

3. The Immunoelectrophoretic Technique

The Immunodiffusion and electrophoresing techniques may be combined to assess antigen antibody complex formation. The polyclonal antibodies in a serum may be fractionated into the component antibodies with different specificities by immunoelectrophoresis.

The technique consists of placing the serum in a well at one end of a thin agarose slab prepared on a microscope slide. When the slide is placed in an electrophoresis apparatus and the current turned on, the heterogeneous antibodies migrate in the gel

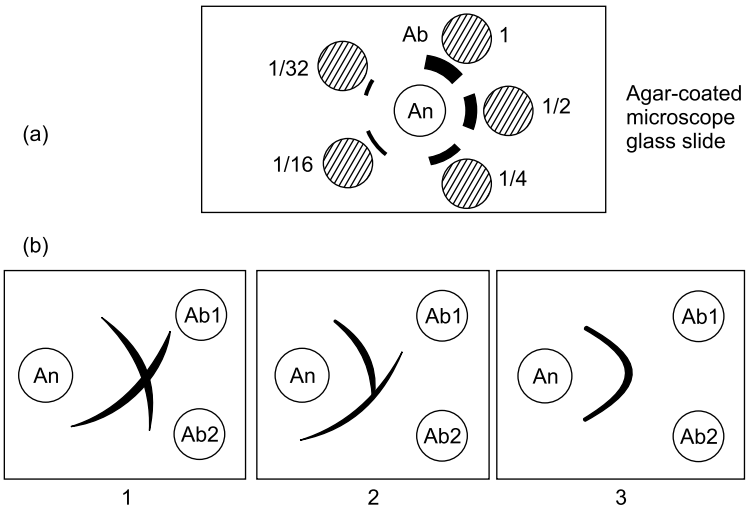


Fig. 4.27 *Ouchterlony's double diffusion technique. This is a precipitin test carried out on an agarose (solid) medium. A glass slide is coated thinly with fine quality agarose and wells made in various patterns as desired. (a) The antigen solution is placed in the centre well and dilutions of the antibody are placed in the wells surrounding it. Antigen and antibodies diffuse radially from their respective wells into the agarose gel. A line of precipitates is formed where they meet in equal concentrations. The precipitates can be visualized as 'spurs' after staining the proteins with Coomassie blue. (b) Two antibodies may be compared with respect to their affinities for a particular antigen by placing them in wells equidistant from the antigen-containing well. (1) If the spurs cross each other, the antibodies are unrelated. (2) If one spur merges with the other, and the latter extends beyond the first spur; the first antibody has some homology with the second one. (3) If both spurs combine to form one common curved spur, the two antibodies are identical.*

at rates proportional to their size. A well may be cut along the longitudinal edge of the agarose slab and the antigen to be matched placed in this. The antigen will diffuse into the surrounding gel and precipitate the matching antibody. The precipitates are seen after staining with Coomassie blue. Other immunoelectrophoretic techniques have also been developed, one of them is the rocket electrophoresis technique (Fig. 4.28). In this case the antibody is mixed uniformly in the gel. It is kept immobile by choosing a pH that prevents its diffusion. The antigens are placed in small wells near one edge of the slide or on a piece of square glass coated uniformly with agarose gel. When the current is on, the antigens move out and the reaction with the antibody is visualized after staining the gel. Rocket-shaped precipitates are seen, the heights of the cone- or rocket-shaped stained region being proportional to antigen concentration. With the help of standard markers, the concentration of the unknown antigen may be determined.

4. Radioimmunoassay (RIA)

Radioimmunoassay (*RIA*) is another technique that is used for precise and sensitive estimation of immune complexes. This technique involves the use of a known quantity of ^{125}I-labelled antigen that is allowed to compete with the antigen to

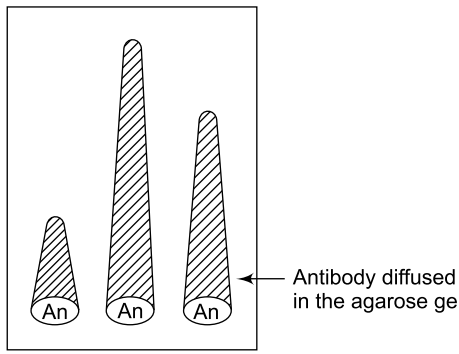


Fig. 4.28 *The rocket electrophoresis technique. Antigens are placed in a straight line at one edge of a rectangular gel, and the antibody is diffused in the latter. On electrophoresis, the antigens move out and complex with the antigen wherever they are in equal proportions. This results in cone- or rocket-shaped regions of An-Ab precipitates. The affinities of different antigen samples are indicated by the height and intensity of staining of the 'rockets' that they generate.*

be estimated for a measured amount of the matching antibody; the degree of competition can be computed. A non-radioactive biotinylated label may be used instead of the radioactive iodine tag.

The RIA procedure is generally as follows (Fig. 4.29). The antigen in a saline solution is allowed to become adsorbed on the inner surface of a plastic tube (or microwell in a multiwelled plate). Non-adsorbed antigen is removed by washing. To avoid non-specific adsorption of unmatching antibodies, a non-specific protein solution is poured into the tube (well). This is again washed out. The antisera to be tested are poured in. The matching antibody alone will complex with the immobilized antigen. Again, excess antisera and unattached antibodies are washed out. A radioactively labelled ligand is included with the antiserum; the former binds itself to the antibody. After the reaction is over, and excess unused antibodies and ligands removed, the test-tube (well) walls are tested for radioactivity. Sophisticated variation of this basic procedure are employed to even distinguish the type and subtype of Ig molecules that complex with a given antigen.

5. Enzyme-linked Immunoabsorbent Assay (ELISA)

The ELISA technique is (Fig. 4.30) similar to RIA up to the addition of the test antiserum or antigen. The antibody or antigen is immobilized (conjugated) on a solid base (test-tube or multi-well microtitre plate surfaces, or plastic beads). The matching protein, labelled with an enzyme (e.g., horseradish peroxidase) is added to the bound protein and the excess enzyme-labelled protein removed. The amount of complexes formed (immobilized) can be estimated by adding the substrate which is a chromogen for the enzyme and measuring the products (one of which is coloured) of the reaction with a spectrophotometer.

6. Immunoaffinity Chromatography

Once a clone that is expressing the protein from the recombinant DNA is identified, it may be necessary to collect sufficient amounts of the antigen in a pure form. One

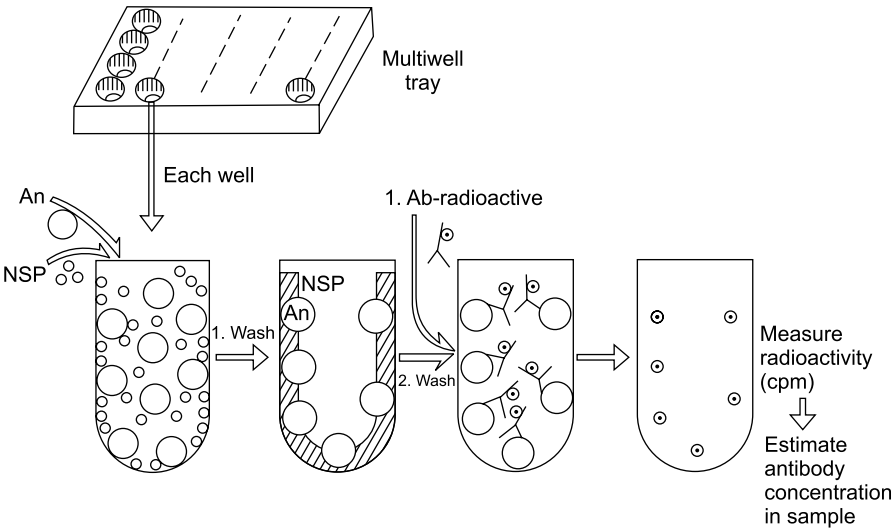


Fig. 4.29 *Radioimmunoassay (RIA). The antigen to be tested is mixed with an excess of a non-specific protein (NSP) and poured into wells in a multiwell culture plate. The antigen is adsorbed on the plastic surface of the well. Surface areas not covered by antigen are lined with NSP. The excess mixture is pipetted out, the wells washed free of proteins and a solution of radioactively labelled antibody introduced into each well. The excess antibody solution is removed and the wells washed to remove free labelled antibody. Each well is cut out from the plate and monitored for radioactivity (cpm or counts per minute) in a scintillation counter. The amount of antibody bound to antigen is estimated from the cpm in each well. The amount of antigen is calculated from this value. The technique may be utilized to estimate the amount of antibody by reversing the roles of An and Ab in the above procedure.*

of the best techniques for doing so is to use the biochemical method of affinity chromatography (Fig. 4.31), with the matching antibody in the column gel. The unpurified mixture of antigen and impurities is passed through this column; the required antigen complexes with the immobilized antibody; everything else is eluted. The antigen can then be released from the complexes. The reverse may be followed to obtain an antibody in a pure form; the antigen is immobilized in this case.

4.3 ❑ EXPRESSION OF CLONED GENES: ISSUES AND PROCEDURES

The splicing of a passenger DNA to a vector and introducing the hybrid DNA into a cell constitute just half the battle. The proof of the pudding lies eventually in the synthesis of the correct mature transcript and functional protein.

The expression of bacterial genes is far less complicated than that of eukaryotic ones. To begin with, bacterial genes are expressed by only one type of RNA polymerase, whereas eukaryotic genes are of three classes: I, II and III, transcribed, respectively, by the corresponding RNA Pols I, II and III. Prokaryotic transcripts do not require the types of processing that are mandatory for the different eukaryotic classes of genes. Control of gene expression in prokaryotes is achieved with far

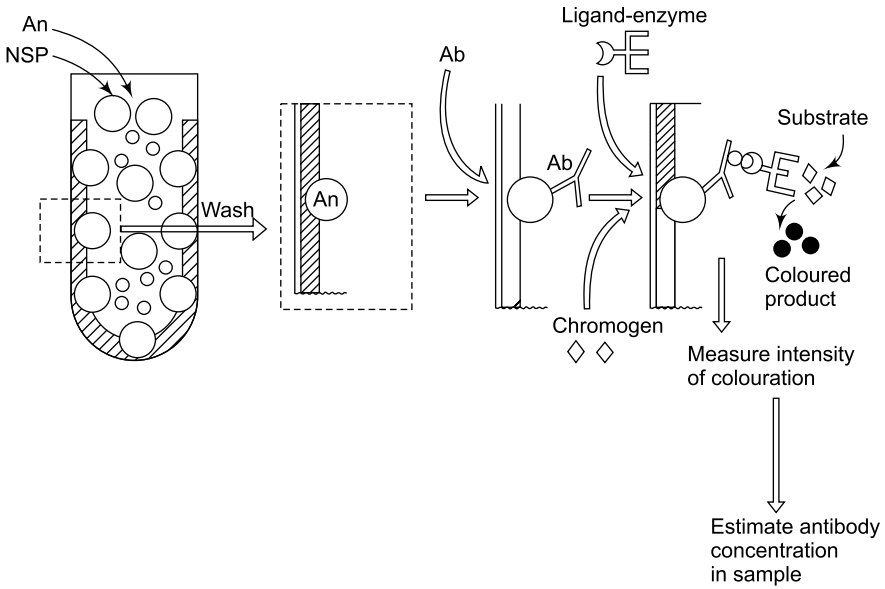


Fig. 4.30 *Enzyme-linked immunoabsorbent assay (ELISA). This technique estimates the amount of antigen (or antibody) using a non-radioactive probe. As in RIA (Fig. 4.29) antigen and NSP are first allowed to line the inner surface of the wells of the culture plate. Now, a non-radioactive antibody (Ab) is added in the following sequence: the Ab is added to each well and free Ab removed. The Ab molecules bound to the An are conjugated to an enzyme (such as horseradish peroxidase) which catalyses a reaction that results in a coloured product. A substrate for the enzyme is added. The enzyme-catalysed reaction colours the solution in the well. The intensity of colouration (measured by a colorimeter) is utilized to estimate the concentration of the antibody, and eventually that of the antigen in the test sample.*

fewer, though subtle, mechanisms than in eukaryotes. The list of *cis*-acting control elements and *trans*-acting proteins that modulate tissue-specific, developmental period-specific and signal-induced transcription of eukaryotic genes is growing longer very fast. In addition, there are several post-translational modifications that are required to make a polypeptide functional.

Any gene expression system must take into account the characteristic minimal features of prokaryotic and eukaryotic systems before attempting to study or induce gene expression in heterologous systems. Two main approaches are used for this purpose. The gene is expressed in *in vivo* or *in vitro* systems. For prokaryotic and viral genes this involves mainly expression in *E. coli* cells, or in cell-free extracts. Both methods possess intrinsic advantages. *In vivo* expression allows one to study the products in a natural setting; cell-free systems are good for manipulating factors and components to discover how the process works. Data from both methods are pooled to obtain what one hopes is a realistic picture of what happens in a real life situation.

Eukaryotic gene expression may be studied in bacterial cells, in cell-free extracts, in isolated nuclei or in eukaryotic cloning cells. Those in bacteria have to be placed on appropriate expression vectors which possess signals that can be utilized in

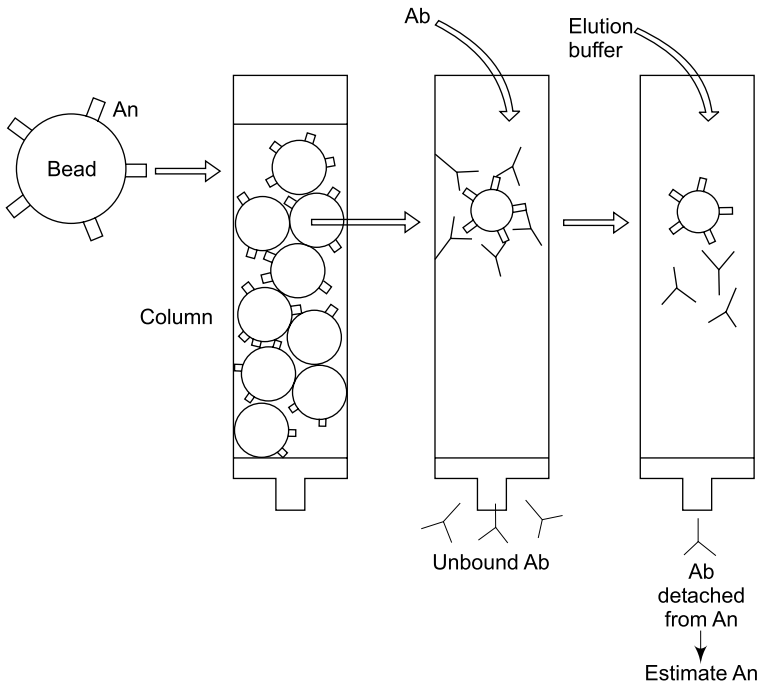


Fig. 4.31 *Immunoaffinity chromatography. Polystyrene beads coated with the antigen (An) are used to pack a column through which the solution of a matching antibody (Ab) is passed. Unbound Ab is removed by repeated elutions. Finally, the An-Ab complexes are released from the beads and eluted. Pure antigen or antibody (by reversing the roles of An and Ab in the above procedure) may be collected using this technique.*

bacterial cells. Cell-free extracts may be used for manipulative exercises as in the case of prokaryotes. As transcription and translation are coupled in bacteria but insulated in separate compartments in eukaryotes, studies have zeroed in on transcription and processing of transcripts in isolated nuclei, and translations in intact cells or cell-free extracts. Certain animal cells are especially useful for studying the expression of eukaryotic exogenous genes. They include the large egg of the African toad, *Xenopus laevis,* and the fertilized eggs of mammals, particularly of rodents. A cultured cell line, CHO (Chinese Hamster Ovary) is a favourite for specific types of studies of gene expression.

Yeast genes are expressed best in yeast cells, although with appropriate bacterial gene signals, they are expressed correctly in *E. coli* . As yeasts are commercially very valuable, it has been judged worthwhile to concentrate on developing yeast-cloning systems exclusively for studying yeast gene expression. The temptation for utilizing yeast cells for heterologous eukaryotic genes, however, persists as yeasts possess the advantages of microbial culture.

One overt advantage has accrued from the attempts to overcome hurdles encountered in expressing genes in heterologous systems; they have led to the information explosion of recent times about the mechanisms involved in the control of gene expression.

One of the main goals of gene cloning is to harvest the products of its expression, both for examining them and for utilizing them for research or application. Some studies aim to recover intermediate RNA transcripts, while others seek the final protein molecules. In the case of the structural transcripts, rRNA and tRNA, the quest stops at the transcript stage, at various stages of processing.

Some of the transcriptional and translational systems in use are described below.

4.3.1 Prokaryotic Gene Expression *in vivo*

It is easiest to study gene expression in a system where genes other than the one of interest are not expressed. In *E. coli* cells this may be achieved in three ways: (1) in maxicells, (2) in minicells, and (3) in *E. coli* cells treated with a dose of UV irradiation that is heavy enough to inhibit expression of the chromosomal genes. These methods are described briefly.

1. The Maxicell System

This system uses an *E. coli* cell that has been treated with a level of UV irradiation that extensively damages the chromosomal DNA but leaves the small vectors unscathed. Such cells are called *maxicells.*

The *E. coli* utilized to make maxicells are repair-deficient mutants, with lesions in the *uvrA* and *recA* genes of the excision-repair systems and in the *phr* gene required for repair by the light repair system. When such a cell is treated with UV irradiation the pyrimidine dimers that are formed are not repaired or removed. The mismatched single strand regions become vulnerable to degradation by nucleases. The much smaller vector DNA molecules will mostly escape destruction. The genes expressed from the vectors are labelled by supplying ^{35}S-labelled amino acids during incubation of the cells. Both small and multimeric large plasmids (say, a tetramer of pBR322) can continue transcription and translation for an appreciable period of time.

Plasmids commonly in use in maxicells are derivatives of Col El such as pBR322. The *E. coli* mutants that work best are the strains CSH603 and CSH26, the former with mutations in *uvrA6, phrI* and *recAI* and the latter in *recA.* Other auxotrophic markers are also present.

Transcription products are examined in the following manner. The maxicells containing hybrid vectors are pelleted and resuspended in nutrient media supplemented with ^{35}S-methionine. After incubating at 37°C the radioactive medium is diluted with L-methionine and the incubation continued for another 5 minutes. The maxicells are pelleted out and placed in a suitable buffer containing SDS and kept on a water-bath at boiling temperature. The lysate is run on an SDS-PAGE to fractionate the transcripts. The newly synthesized ones are detected by Northern blotting followed by autoradiography. The mRNA species are eluted from the appropriate bands in the gel and characterized.

2. The Minicell System

E. coli mutants with min A$^-$ and min B$^-$ lesions are prone to divide into unequal sized daughter cells, the smaller of the two having no bacterial chromosome.

Recombinant vectors already introduced into the cell may find some copies in these smaller minicells. Genes on the vectors in the minicells are expressed into RNAs and polypeptides. The *E. coli* strain that is used generally to make minicells is DS410, a derivative of strain 925.

The procedure consists of growing the cells on nutrient media, and pelleting them out to be treated with chilled $CaCl_2$ (0.1 M soln.) to make them competent (physiologically prepared to take in exogenous DNA). They are then mixed with the recombinant plasmid and the cells heat-shocked at 42°C. These cells are then incubated at 37°C and the minicells separated from nucleated cells by sucrose density centrifugation. The purified minicells are frozen in aliquots at –70°C after examining them under a microscope for the degree of their purity. The aliquots are thawed overnight at 4°C before use. The minicells are then suspended in a medium containing ^{35}S-labelled methionine. The minicells are lysed, and the lysate fractionated by SDS-PAGE. The labelled proteins are identified by autoradiography.

3. Expression of Phage Genes in UV-damaged E. coli Cells

The *E. coli* strain used most for monitoring expression of phage genes is the *uvrA* mutant, *E. coli* 159. Phage λ gets attached to *E. coli* cell surfaces via a receptor which also facilitates the entry of maltose. Cells grown in maltose possess an increased number of such receptors. Hence, *E. coli* into which phage λ is to be introduced is grown in a medium that possesses maltose in place of glucose.

Recombinant phage λ containing the gene that is to be studied, or the product of which is to be collected is added to *E. coli* 159 that has been grown on a maltose-M9 medium, exposed to UV (254 nm), washed in a fresh medium and resuspended in a prewarmed medium. The bacteria and phage particles are incubated with ^{35}S-methionine for about 10 minutes. Unlabelled L-methionine is added next, the cells washed and resuspended in SDS buffer and incubated at 100°C (in a water-bath) for a few minutes. Aliquots of this mixture are fractionated on SDS-PAGE and the labelled proteins identified by autoradiography.

4. Requirements for Efficient Expression

Efficient expression of a cloned gene—bacterial, phage or unsplit eukaryotic—in a bacterial cell or extract depends on at least three features: (i) the strength of the promoter, (ii) the SD (Shine-Dalgarno) sequence in the leader of the bacterial mRNA that binds itself to a complementary region in the 16S rRNA of the smaller ribosomal subunit, and (iii) the distance of the SD from the AUG or start codon also appears to be critical for correct expression.

(i) Promoters Most *E. coli* promoters are rather weak. It is a practice, therefore, to provide the cloned gene that is to be expressed with a promoter. As eukaryotic promoter sequences are not effective with bacterial RNA polymerase, an eukaryotic gene has to be furnished also with a prokaryotic promoter.

'Portable promoters' are available that can be positioned at the correct distance from the starting codon of the inserted gene. These promoters are either taken from bacterial or viral systems or are synthetically prepared oligonucleotides with restriction sites at both ends, which allow them to be fitted into the same restriction

sites in the vector. Sometimes, the blunt ended 'promoter' oligonucleotide is joined to *Eco* RI cut vector ends after converting the latter into flush ends.

The promoters in general use include the following: (i) The *E. coli lac* operon promoter. (ii) The *E. coli trp* operon promoter. (iii) The promoter of an *E. coli* membrane protein gene, *ompF,* which is under the control of the gene product of an *ompR* gene. The vector with the *ompF* is cloned in an *ompR ts* mutant cell that allows enhanced transcription at a temperature higher than 31°C. Hall and Silhavy designed pMH 621 that carries the promoter and the upstream portion of the *ompF* gene. To facilitate the insertion of DNA cleaved by different restriction enzymes, linkers with cutting sites for several enzymes are added at the Bgl II cloning site (iv) Two *lac* promoters in tandem.(v) A hybrid promoter containing the upstream region (–35) of the *trp* promoter and the –10 region of a *lac* promoter. The fused *trp-lac* or *tac* promoter mediates a high level of transcription, (vi) The promoter of the β-lactamase gene (*bla*) present on the transposon Tn3 (and in pBR322) which provides ampicillin resistance to the cell, (vii) A *lacUV5* promoter, (viii) The P_L promoter of phage.

An expression vector with P_L sequences is normally used in a lysogen having a *ts* mutation of the *CI* (λ repressor) gene. At 31°C, the repressor blocks transcription from P_L. Raising the temperature denatures the protein, and P_L is usable for transcription. When expression is not required the lysogens are maintained at 31°C. The *E. coli* strain CI_{857} possesses such a *ts* mutation in the *CI* gene.

Use of a promoter, whose activity can be controlled experimentally, is a preferred strategy. The use of *ts* mutants of a regulator molecule, as described in the preceding example, is one strategy. Another is to regulate expression with a suitable inducer. An example of the latter is the use of IPTG (isopropyl-β-thiogalactoside), an analogue of lactose, to induce transcription from a *lac* promoter.

(ii) The SD region and its distance from AUG It has been demonstrated that the length of RNA separating the SD region from the first codon (AUG) of a bacterial gene affects the efficiency of transcription. However, no hard-and-fast rules have been discovered that suggest the optimum composition of the SD region or its distance from AUG. A semi-consensus sequence for the SD region, approximately 10–15 bp distance from the AUG, is common in many of the bacterial genes that have been examined.

4.3.2 Eukaryotic Gene Expression in *E. coli*

As mentioned earlier, the transcriptional and translational signals are different in eukaryotic and *E. coli* systems. Also the bacterial cell does not have the wherewithal for splicing, capping and A-tailing the eukaryotic mRNA. A eukaryotic mRNA in a bacterial cell, of course, need not have these processed 5′ and 3′-ends, as the latter are respectively a translational signal and one for allowing the mRNA to cross the nuclear membrane barrier in the eukaryotic cell. In addition, the pattern of codon usage in the bacterial cell is often different from that in the eukaryotic system. You may recall that most amino acids are represented by two or more codons. All the codons for an amino acid are not utilized with equal frequency. The codons

employed depend on the abundance of the tRNA species with the complementary anticodon.

The above difficulties may be overcome by several methods. The following are a few of the available strategies:

(i) As eukaryotic mRNAs are not preceded by the SD region, any one of the following tactics may be used to make a eukaryotic gene translatable in *E. coli:* (a) the gene is positioned immediately after the initiation codon of a bacterial gene, and (b) a gene is synthesized with an accompanying SD sequence and spliced to the vector downstream of the bacterial promoter.

(ii) The codon usage pattern may be bypassed by chemically synthesizing the eukaryotic gene, thereby substituting the bacterially favoured codons for the corresponding eukaryotic ones.

(iii) Splicing may be avoided by cloning the cDNA copy of the processed (spliced) mRNA, or by using a synthesized gene without the introns.

(iv) Bacterial cells generally degrade small foreign proteins. To prevent this fate, the eukaryotic gene may be cloned in the same reading frame as that of a bacterial gene; the longer fused protein is usually safe from proteolytic attacks. The eukaryotic protein is subsequently cleaved out from the fused protein.

Eukaryotic genes may be expressed in *E. coli* or other bacteria in one out of two ways. One strategy is to aim for fused proteins as mentioned earlier and the other is to express it directly as an unfused gene. Both methods have their plus and minus points.

1. Fused Genes

One trick of expressing a eukaryotic gene in *E. coli* involves the fusion of the former with the promoter and upstream codons of an expressable bacterial or viral gene. The bacterial cell is hoodwinked into continuing transcription begun at the prokaryotic promoter, across the truncated prokaryotic gene and then across the adjacent eukaryotic one. The fusion protein has eventually to be cleaved to recover the pure cloned protein.

Cleavage of the eukaryotic protein from the fusion product may be carried out by chemical or enzymatic methods. The cloned fusion gene is, therefore, equipped with regions that are vulnerable to breakage by these means.

Chemical cleavage is achieved by including a pair of codons for asparatic acid and proline at the juncture of the bacterial and eukaryotic gene sequences. The asp-pro bond is cleaved when the protein is incubated in a 70 per cent formic acid, 6 M guanidine hydrochloride reagent. However, this technique leaves a proline residue at the start of the polypeptide. This slight deviation has been found to lower the functional efficiency of the expressed protein. A second chemical strategy is based on cleaving at a methionine residue with cyanogen bromide.

Enzymatic cleavage is possible when appropriate cleavage sites are included in the fused gene. In some cases copies of the gene of the required protein are joined in tandem, with the first one fused to a prokaryotic initiating sequence. Cleavage signals (e.g., arginine-arginine bridge) are provided between each copy. The long multimeric polypeptide is first cut into monomers with trypsin and then the arginine residues are removed by appropriate treatment (Fig. 4.32)

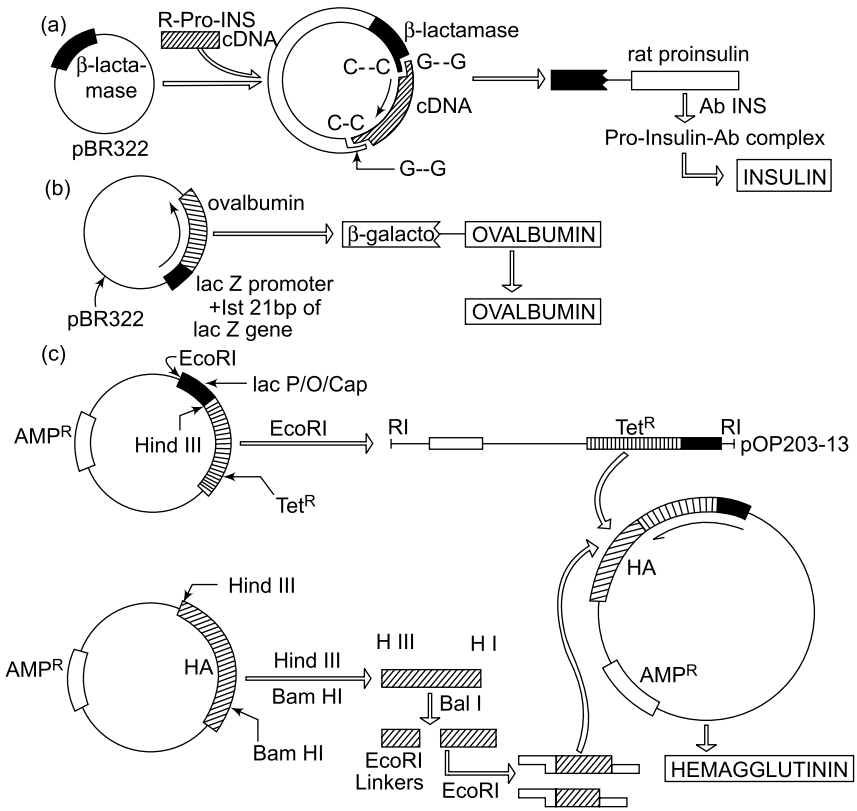


Fig. 4.32 *Fused genes and the production of hybrid polypeptides.*

In Fig. 4.32, the gene to be expressed is latched on to a known gene (or the 5′-region of its coding and regulatory sequences) and transcribed as a common mRNA-containing sequences from both genes. If the gene to be expressed as a protein is placed with its codons in the same reading frame as that of the preceding gene, functional proteins (correct ones) are synthesized. This principle has been utilized to express (a) rat insulin, (b) chicken ovalbumin, and (c) the haemagglutinin (HA) surface protein of the influenza virus.

Homopolymeric tails (C : G) were used to join the rat proinsulin gene sequences to the pBR322 at the Pst I site within the Amp gene. Insertions in the correct reading frame as that of the *bla* (Amp^R) gene produced insulin, which was detected immunologically. The hybrid protein was found in the periplasmic region. The lactamase normally is secreted into the periplasm. The fact that the insulin-producing cells continued to be resistant to ampicillin indicates that the segment of the Amp^R gene left upstream of the insertion site was functional.

The chicken ovalbumin cDNA was cleaved from the plasmid in which it was cloned, and the ends of the fragment flush ended and ligated to a vector linearized by cleaving at a site after the first few codons of the *lacZ* gene. The fused gene was transcribed from me *lacP/O* and translated into a hybrid polypeptide that could be identified using a matching antibody.

The cDNA of the HA (haemagglutinin) gene of the influenza virus capsid protein was fused in continuation with the control region of the *lac* operon which was part of a plasmid vector pOR 203–13 (the region is 203 bp long). The control region (an *Eco* RI-*Hin* dIII segment) consists of the UV5 mutant version of the *lac* promoter and a mutation L8 that releases the promoter from catabolite (CAP) control, so that transcription may be initiated at *lacP,* independent of the cyclic AMP (cAMP) level in the cell. The HA gene was cleaved out first from the plasmid containing it, as a *Hin* dIII-*Bam* HI fragment. The latter fragment was cleaved further with Bal into two unequal sized subfragments. *Eco* RI linkers were added to the larger subfragment and the latter treated with *Eco* RI to generate *Eco* RI joining ends. This fragment was inserted at the *Eco* RI site in pOP203–13. Those constructs with the HA codons in the same reading frame as the *lacZ* codons were translated.

The fused gene method has been used to synthesize peptides that can be employed as antigens. Kleid has synthesized peptides of the foot-and-mouth virus coat protein to be utilized as antigens (vaccines in this case) that provoke synthesis of antibodies against them. Goeddel and associates demonstrated the synthesis of insulin from a fused gene. Frazer and Bruce, and Heiland and Gething, have respectively fused the *lacZ* gene (first 7 codons and promoter) with genes of chicken avalbumin and HA (haemagglutinin) of the coat protein of the influenza virus. The expression vector, pOP 203–13, constructed by Fuller contains an *Eco* RI site at the downstream edge of the seventh *lacZ* gene codon. The P_L of phage λ and a portion of the phage MS2 polymerase gene are included in another vector, PLC24, that has been utilized by Derynk *et al.* to synthesize a human fibroblast interferon fused to the truncated MS2 protein.

The Human Growth Hormone (HGH) has also been expressed from a fused gene (Fig. 4.33). An HGH gene was fused to *E. coli* genes, of the tryptophan operon, which was already cloned in pBR322. The trp promoter and the above two genes were cloned as a *Hin* fl-*Hin* dIII fragment in pBR322. The *Hin* dIII site is at the 92nd codon of the *trpD* gene. The cDNA version of the HGH gene also possesses a *Hin* dIII site in its 5′ flanking region. But the joining vector and cDNA, using the *Hin* dIII cut extension, disturb the correct reading frame by one base. To avoid this defect the *Hin* dIII end of the *trp* fragment is first filled in by DNA polymerase, ligated to a *Hin* dIII linker and treated with *Hin* dIII to generate the characteristic *Hin* dIII joining end. Ligation of this tailored fragment to the the end of the HGH fragment restores the correct reading frame of the HGH codons. As a result of such manipulations, a few extra codons are generated at the linker region. However, they do not disturb the correctness of the expressed hormone.

In Fig. 4.34, the 14-codon long somatostatin gene was prepared in segments by oligonucleotide synthesis and the segments were joined. The synthetic gene was provided with a start (AUG) codon (for methionine) and two stop codons and possessed an *Eco RI* extension at the 5′ end, and a *Bam* HI one at the 3′ end. This gene was inserted into a pBR322 vector carrying a *lac* control and some codons of the *lacZ* gene. An *Eco* RI site was present at the 3′ end of the truncated *lacZ* gene, and a *Bam* H1 site further downstream. Hybrid β-galactosidase-somatostatin expressed from the fused gene was treated with cyanogen bromide that breaks the peptide bond at the carboxyl end of methionine residue. The somatostatin polypeptide (without

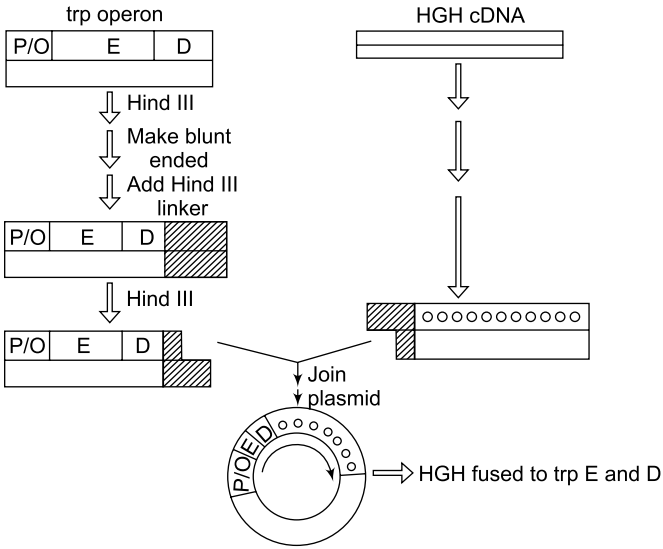


Fig. 4.33 *Human Growth Hormone (HGH) expressed from a gene fused to trpE and trpD genes of the E. coli tryptophan operon.*

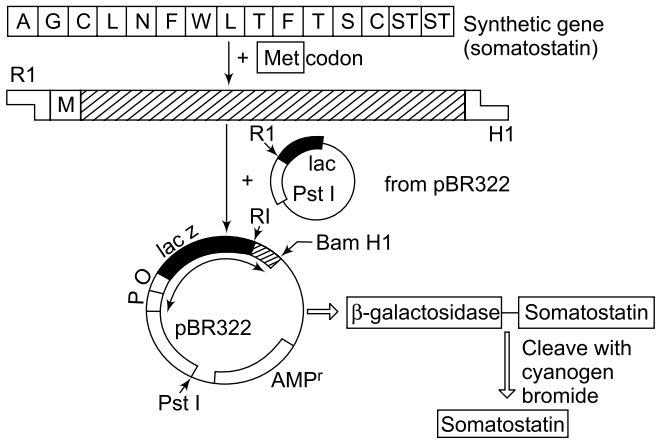


Fig 4.34 *Expression of a somatostatin as a fused gene.*

the added methionine residue) is obtained as a result. The actual investigation failed to collect the somatostatin, possibly due to the degradation of polypeptides smaller than a critical size. Later investigations used a longer fragment of the *E. coli* gene and the gene product of interest detected by immunological techniques.

2. Unfused Genes

The eukaryotic gene may be cloned downstream of a bacterial SD sequence with a restriction site on the upstream side of the initiating (ATG) codon. The human

growth hormone (HGH) was the first eukaryotic gene that was cloned and expressed in bacterial cells by this method. The strategy employed is described in brief.

The HGH is a secreted protein and so is preceded by a signal peptide sequence that enables the protein to pass out of the cell through the membrane. Such genes normally do not possess an ATG-initiating codon. The HGH gene consists of 191 codons. The first 24 codons were synthesized chemically with an *Eco* Rl recognition sequence and an ATG at its 5′ end. The natural gene possesses a restriction site between the 24th and 25th codons. A cDNA copy was made from the HGH mRNA and cleaved between the 24th and 25th codons with the restriction enzyme. The synthetic and abbreviated cDNA stretches were ligated and this tailored gene was introduced into a plasmid vector that carried the *lac* promoter, the SD sequences and 11 more base pairs of the vector (Fig. 4.35). The mRNA was transcribed correctly across the unfused gene.

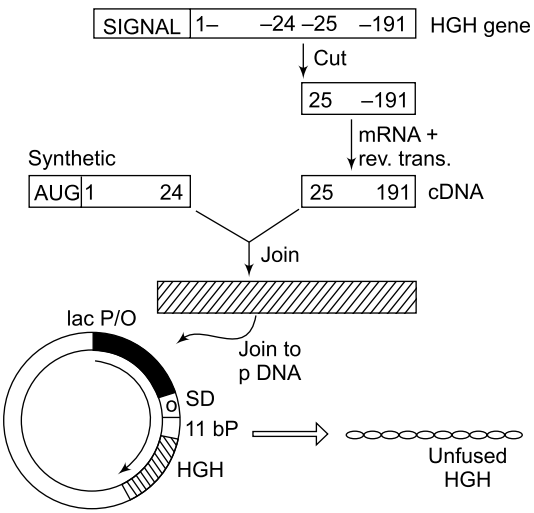


Fig. 4.35 *Human Growth Hormone (HGH) has also been expressed from an unfused gene. An Hae III site between the 24th and 25th codons of the HGH gene allowed cleavage of the HGH cDNA into two fragments (1–24) codons, (25–191) codons. Only the larger fragment was used. The smaller fragment was synthesized chemically together with an added AUG (methionine) codon preceding codon 1. The synthetic (met 1–24) and cDNA (25–191) DNA fragments were ligated and the gene inserted 11 bp downstream of a Shine-Dalgarno (SD) or ribosome-binding region between the P/O and AUG. Transcription initiated at the bacterial promoter resulted in mRNAs that were translated into functional HGH proteins. The starting methionine residue was absent in the final product; it was presumably removed by the E. coli system.*

It must be remembered that directly expressed eukaryotic proteins possess a methionine protein at the N-terminal position, unlike natural eukaryotic proteins. Such trivial difficulties will be overcome in time. What is exciting is the fact that proteins that are today procured from animal sources at a great expense, and which possess other shortcomings, may be produced by genetic engineering in easily cultivated cloning cells with assured purity, correct antigenicity and in a far cheaper manner than by traditional methods.

3. Secreted Proteins

Proteins that have to be secreted are first expressed as a pre-protein that possesses a short hydrophilic length of a polypeptide at the N-terminal of the main protein. This extra peptide, called the signal peptide (Fig. 4.36), is instrumental in guiding the rest of the polypeptide through the endoplasmic membrane into the periplasmic region. The signal peptide is removed by an enzyme on the outer surface of the membrane. Proteins to be synthesized in cloning bacterial cells may be provided with a signal peptide by attaching the consensus sequence for the same to the start of the coding sequences of the protein of interest.

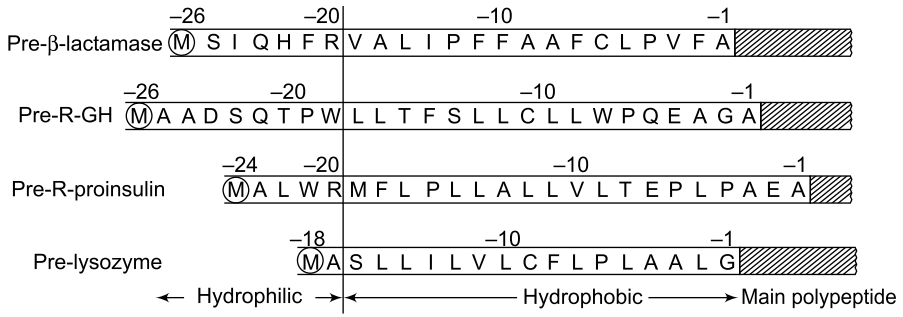


Fig. 4.36 *Sequences of a few signal peptides showing regions with hydrophilic and hydrophobic amino acid residues.*

Secreted foreign proteins are safe from cellular proteolytic enzymes and are more easily recovered from the growth medium than those that remain within the cell. In addition, continuous fermentation processes may be utilized for production rather than discrete batch lots of fermentation.

4.3.3 Prokaryotic Gene Expression *in vitro*

Cell-free systems are extracts of cells containing all the enzymes and other factors required for the biological processes to be studied. As transcription and translation are coupled in bacteria, both events can be monitored in the same cell-free extracts.

Two protocols are available and they are used widely. They are the Zubay method and the Gold and Schweiger method. As RNase is abundantly present on the palms of hands, all surfaces and apparatus to be utilized for RNA isolation projects must be made RNase-free. There are different ways of achieving this: the gloves are sterilized and the glassware thoroughly baked (dry heat) or all glassware is washed in DEPC (diethyl pyrocarbonate) solution and all reagents made from water containing DEPC.

The cells which will provide the extract are pelleted, washed in a buffer containing 2-mercaptoethanol and kept frozen at –70°C until use. Amino acids (some of which are labelled), ATP and requisite co-factors are added to the extract obtained by lysing the thawed cells, and labelled proteins are identified after the incubation is over.

The Zubay method utilizes the whole extract, while the second technique fractionates the components deemed necessary from the extract, freezes them separately and reconstitutes the mixture before use. Fractionation procedures reduce the efficiency of gene expression, but have the advantage of providing a means of identifying the exact conditions required for correct expression.

4.3.4 Transcription of Endogenous Genes in Isolated Nuclei

RNA synthesis in eukaryotes has been examined in two ways. In the first one, the *in vivo* conditions are preserved and transcription allowed to occur in conditions which are as close an approximation of natural ones as possible. This approach uses intact nuclei that have been cleaned of cytoplasmic and RNase adherents. The second approach is to allow transcription *in vitro* of exogenously supplied DNA in a defined reaction mixture.

Both the preceding methods have their advantages and shortcomings. While it is possible to know the developments perhaps as they take place in intact cells using isolated nuclei, it is not possible to intercede experimentally to discover details about the conditions and components required for the normal events. There is a greater chance of unveiling the process in cell-free expression systems. We will discuss the latter strategy in the next section. We turn our attention here to expression in isolated nuclei.

To ensure that transcription is as little deviated from that in intact cells, the nuclei must be isolated very carefully, so that the internal organization, especially of the chromatin, is not disturbed. Techniques to achieve this goal are available.

Nuclei of most animal cells are recovered from cells lysed in iso-osmotic solutions, containing a non-ionic detergent. Contaminating cytoplasm and cellular nucleases are removed by passing the released nuclei through concentrated sucrose solutions. If the cells are cultured ones, the above mentioned procedure is sufficient. If nuclei have to be extracted from cells in tissues, the latter is first homogenized in a suitable buffer. The coarse material of the tissue pieces is removed from the homogenate by straining through layers of coarse cheesecloth. The homogenate contains lysed cells. Cultured cells are also homogenized. The homogenate from either source is layered over a sucrose magnesium acetate buffer and centrifuged at about 30,000 g at 4°C for about an hour. The pellet of cleaned nuclei is resuspended in a storage buffer and stored in liquid nitrogen. The nuclei are usable for several months.

These nuclei are then utilized to study RNA synthesis from endogenous genes. The nuclei are incubated in a buffer supplemented with amino acids, one of which is labelled with ^{32}P. In practice, GTP is preferred, as the other three nucleotides often transfer the label to other transcripts; dATP takes part also in polyadenylation of the mRNA.

At the end of incubation, the total RNA is extracted and fractionated. Aliquots of the reaction mixture (with the intact nuclei) are also removed at intervals to monitor the progress of transcript formation. The lysate from the nuclei is spotted on discs of Whatman no. 1 filter paper and, after a TCA (trichloroacetic acid) treatment, the latter are counted for radioactivity in a liquid scintillation counter. TCA precipitates proteins.

To distinguish newly made transcripts from those already present in the nuclei, a specific label is given to the initiating base. Usually this is a γ-thioribonucleoside, the 5′ phosphate of which is less vulnerable to removal by phosphatases of the nuclei than those of γ-32 labelled NTPs. The labelled RNAs are easily isolated by running the reaction mixture (nuclear lysate) through an affinity column containing mercury agarose.

All three types of RNA Pols are present in the nuclei. The bulk of the transcripts are rRNAs and tRNAs. The transcript one needs to study may be recovered by treating the system with the cyclic peptide from mushrooms, α-amanitin. It does not inhibit Pol I, inhibits Pol II at a low concentration and Pol III at a very high concentration.

4.3.5 Transcription of Eukaryotic Genes in Cell-free Systems

Separate systems have been developed for exclusive transcription of class I, II or III eukaryotic genes. A general procedure is to add exogenous DNA to an extract of the whole cell. Of course, it becomes difficult to fish out the RNA of interest from the heterogeneous RNA population of a cell. Techniques to tackle such a situation have been developed.

Transcription of eukaryotic genes is accompanied by unique post-transcriptional processing operations. Extracts made from Hela cells (human cultured cell line) grown in a suspension culture, possess all the enzymes and factors to accomplish the post-transcriptional modifications. Further, mRNAs have been capped, tailed and spliced as necessary; rRNAs and tRNAs have also been processed correctly. This fact has been sufficient to encourage the use of this *in vitro* system for studying the mechanisms of synthesising mature RNAs. Conditions in such a cell-free system may be altered to test for their roles, if any, in the process.

A majority of transcription studies are concerned with mRNA synthesis and processing. Hela cells are lysed by homogenization. The thick lysate is rid of chromosomal DNA which is protected from being sheared during the earlier step. After suitable treatment, the DNA-free extract is stored at 70°C. The activity of the extract is retained for about a year.

The frozen extract is thawed before use and a transcription-reaction mixture prepared. One of the 4NTPs is labelled with ^{32}P. The reaction is terminated after an hour or so with urea-SDS-EDTA in Tris-HCl solution. The mixture is centrifuged and the supernatant phenolized. The aqueous upper layer is treated with chloroform. The RNA molecules are precipitated and lyophilised, dissolved in a solution containing sarkosyl and EDTA and stored at –20°C.

Termination in cell extracts does not stop at the natural termination signals. To overcome this difficulty, the template DNA is cut with a restriction enzyme at a desired site downstream of the termination codon. The mRNA and the enzyme fall off automatically when they reach this cut end. This gimmick also results in a collection of RNA molecules of a given size that can be readily identified after fractionation of the transcripts. Such 'run-off' transcription is particularly useful for determining the site of initiation of transcription. In this case, the DNA template is cut slightly downstream of the expected start of the coding sequences. After the

reaction, the template is removed and the RNAs synthesized and hybridized with a DNA probe. The single-stranded regions of the unhybridized DNA are removed with S1 nuclease. The RNA fragments are then fractionated.

Polyadenylation is not accomplished in *in vitro* systems, although capping is normal. The poly-A tail is added *in vitro* by incubating the RNAs in a fresh extract supplemented with components of the tailing reaction. The A-tailed transcripts are isolated by electrophoresis or by passing the mixture through an affinity chromatographic column containing oligo (T) cellulose.

Transcription of class I and class III genes is carried out more efficiently in cell-free extracts than of mRNAs (class II genes).

4.3.6 Eukaryotic Protein Synthesis in Cell-free Systems

Cell-free systems utilized for the translation of eukaryotic mRNA are extracts from suitable eukaryotic cells supplemented with specific components as required. Systems that are mostly in vogue utilize lysates of rabbit red blood cells (reticulocytes) and wheat embryo (germ). Any other type of cell that can be grown in bulk in suspension cultures may be used for making cell extract-based translation systems. Mouse ascites tumour cells, mouse L strain (TK^-) cells, Chinese hamster ovary (CHO) cells and Hela (human cell line) cells are usable for this purpose.

At first reticulocyte extracts were employed to study the mechanism of protein synthesis itself. These cells are stocked with the mRNAs for the α and β polypeptides for globin. Their translation into the corresponding globins was followed, with monitoring, in these lysates. In 1976, Pelham and Jackson discovered that if the endogenous mRNAs in the extract were removed, the system could be utilized for the translation of exogenously added transcripts. The resident mRNAs are destroyed by micrococcal nuclease that in a given condition degrades the mRNA but not the rRNAs and tRNAs. This is very fortunate as the latter are integral components of the translating machinery. The nuclease protein is inactivated after the reaction is over by adding EGTA or ethyleneglycobis (β-aminoethyl) ether tetraacetic acid that chelates calcium ions required for nuclease activity. No nuclease, therefore, remains to degrade the exogenous mRNA that is added to the system.

An essential precaution, as with any work involving the integrity of RNAs, is to have all apparatus and surfaces free of RNase. RNase is ubiquitous and is rich in the palms of the hands. Sterilized gloves must, therefore, be worn and all glassware baked thoroughly before use.

The particular mRNA to be translated is added to an incubation mixture consisting of the cell extract, amino acids, one of which is labelled (^{35}S-methionine, ^{14}C-leucine), a mixture of KCl (swells the cells) Mg-acetate and salts, and a source of energy (ATP, GTP and creatine phosphate with creatine phosphokinase). Aliquots of the mixture, which are incubated at 30°C, may be taken out to monitor the progress of translation. After the reaction is over, an aliquot of the mixture is treated with NaOH and H_2O_2 (hydrogen peroxide) to release amino acids bound to tRNAs. The proteins are precipitated with TCA (trichloroacetic acid) treatment and its radioactivity assessed. The monitoring may be also done by spotting aliquots on discs of Whatman 1 filter paper which are treated with TCA and H_2O_2 and counted

in a liquid scintillation counter. The amount of newly synthesized polypeptides is calculated, taking into account the proteins that were present in the extract *in situ.*

Rabbit reticulocyte lysate is prepared from the blood of animals that have been made anaemic experimentally about 9–10 days before bleeding. Acetyl phenyl hydrazine injected daily for about four days ensures anaemia, which, in turn, promotes an increased rate of red blood cell production in the animal. The blood is collected in heparin solution (to prevent clotting) and the cells spun out of the fluid. The cells are then lysed, centrifuged again, and the lysate stored at –70°C after freezing in liquid nitrogen. The lysate is then treated with micrococcal nuclease (from *Staphylococcus aureus*). The lysate is mixed with a haemin-containing (prevents haeme-directed inhibition of translation) buffer that also has $CaCl_2$. This mixture is incubated at 20°C with the nuclease. Reaction is stopped after 15–20 minutes by adding EGTA and the mixture kept on ice. This lysate can be stored in liquid nitrogen or used right away.

The enzyme-treated lysate is excellent for translating exogenous mRNAs, even when the latter is present in very small quantities. The efficiency of the translation may, however, be lowered by the continued presence of fragments of endogenous mRNAs containing the ribosome attachment (and, therefore, translation initiation) site.

It is far easier to prepare the wheat germ lysate translation system. Wheat germ (obtained commercially) is ground into a paste in a pestle-and-mortar using an extraction buffer containing spermidine—a polyamine that stimulates polypeptide chain elongation. In the absence of spermine and/or spermidine, there is a tendency for polypeptides translated in wheat germ lysates to be terminated prematurely. The paste is fractionated by centrifuging at 30,000 g and the lipid-free portion of the supernatant (middle region) is dialysed or purified in a biogel column (e.g., Sephadex G-25). The lysate is equilibriated with a given buffer and stored at –70°C or lower temperature (liquid nitrogen).

The wheat germ system has very little endogenous protein synthesis. Hence, the nuclease treatment may not be required. Some, however, prefer to use the nuclease-treated lysate. This lysate is incubated, as in the case of the reticulocyte lysate system, and synthesized polypeptides estimated both at intervals during incubation and at the end of the show.

Cell lysates prepared from cells other than RBCs do not require the addition of haemin. However, only the RBCs exhibit a very high level of protein synthesis. In non-RBC mammalian cell lysates, translation is correct though the chain elongation is very slow, and prematurely terminated proteins may be produced.

Cell-free systems for the translation of eukaryotic transcripts have yielded much information on the mechanism of protein synthesis. For instance, the existence and mechanism of action of the 20–30 amino acid long signal peptide, that precedes a secreted protein while passing through the membrane, was discovered by such studies. In these studies, shreds of the endothelial reticulum membrane (microsomal fraction) had to be included in the incubation mixture.

4.3.7 Expression of Genes in Oocytes of *Xenopus*

Oocytes of *Xenopus* may be used to study transcription and translation of genes from a wide spectrum of species that include viruses, animals and plants. The large egg cell is well stocked with the enzymes and other factors essential for the synthesis of the transcript and its eventual translation. Studies have demonstrated that both operations are performed correctly; that is, the start and termination sites, and post-transcriptional and post-translational modifications appear to be identical to the ones observed in the natural setting of the exogenous gene.

Of course, the finer details of the mechanisms of both processes can be monitored and manipulated more realistically in cell-free systems. However, *Xenopus* egg cells are still in demand for the study of post-transcriptional and post-translational processings of the gene products. The *Xenopus* cell has the infrastructure not only for expressing the polypeptides, but also of glycosylating them, folding them and eventually allocating them to their correct locations in the cell, or secreting them where relevant.

The exogenous DNA or RNA is microinjected using a glass pipette, the tip of which has been pulled to give an external diameter of about 200 μm. It is possible to deliver around 50 ml of material in each dose of injection. All operations are carried out at a temperature below 30°C, as above this temperature, only the heat shock protein hsp 70 is synthesized.

Let us examine the use of the preceding system for transcription and translation of exogenous nucleic acids.

1. Transcription Studies

Gurdon and associates were the first to study transcription of exogenous DNA in *Xenopus* oocytes. Since then the system has been utilized for expressions of genes by all three RNA polymerases. It has been noticed, however, that transcriptions of rRNAs and tRNAs by Pol I and Pol III are more efficient than those of mRNAs. Split genes, with the intron removed, are transcribed and sent into the cytoplasm. Complete split genes do not always achieve this success. The trouble seems to be the stage of transcription initiation. Significant research in the laboratories of Birnsteil *et al.* and McKnight and Kingsbury has identified some of the 5′-flanking signals required for correct initiation as well as those that ensure normal terminations. The conditions necessary for splicing interrupted mRNAs are also studied in this system.

The exogenous DNA works best if it is in a circular form. Where the gene is already on a vector, it is useful to delink the vector and gene sequences and circularize the insert DNA. This ensures transcription of only the inserted DNA and not of the vector sequences.

The DNA may be introduced into the nucleus very specifically by bringing the latter near the surface of the cell by centrifugation; or, the micropipette may be introduced at the general region of the nucleus after some proficiency has been attained by practising with a dye solution. The synthesis of the RNA is monitored by the incorporation of radioactively labelled transcripts. The ^{32}P-CTP or ^{32}P-GTP generally used for the purpose is injected into the nuclei. The injection can be made at the same time as that of the exogenous DNA or about one day later. The latter procedure gives better results for mRNAs that acquire poly A-tails.

The synthesis of the transcripts and their progressive processing may be monitored by examining aliquots by electrophoresis or by a procedure referred to as that of hybrid release. The latter method involves fixing the template DNA on small discs of nitrocellulose filters and flooding them with the isolated labelled RNAs. The complementary template and RNA strands hybridize. Unwanted free RNAs are washed off and the hybridized RNAs recovered by elution from the filter papers. The RNA is collected by precipitation in ethanol in the usual manner. This RNA is then fractionated by PAGE and characterized (length, 5′- and 3′-ends and so on). The ends are identified by S1 mapping. DNA probes, labelled or unlabelled, are used to hybridize with unlabelled or labelled transcripts, respectively. The probe contains the length of the gene from a region slightly upstream of the expected or surmised promoter region to a little beyond the initiating codon (into the first exon in a split gene). The length of ss DNA extending 5′ to the initiating base of the RNA will be degraded by the S1 nuclease. The two nucleic acids are then dissociated and the lengths (and sequences) of the synthesized RNAs determined. Similarly, the 3′-termination regions can be identified with a DNA probe for that region. Studies such as those mentioned earlier have indicated that the transcripts of exogenous DNA are correct with respect to their various primary features.

Post-transcriptional modifications are followed by injecting exogenous RNA into the nucleus and the latter is isolated for looking at the transcripts before they are exported into the cytoplasm. The proteins required for post-transcriptional processing have been identified in some laboratories by coinjecting proteins of interest with the exogenous RNA or DNA into the *Xenopus* nuclei.

2. Protein Synthesis

Again, it was Gurdon's team that first injected mRNAs into *Xenopus* oocytes to study protein synthesis. The injected 9S mRNA of rabbit globin was correctly expressed into the protein. In fact, its efficiency of translation was more than 1000 fold of that achieved in a cell-free lysate system. The chief advantage of using the *Xenopus* system is in the authentic expression, modification and translocation of the protein.

Translation in the system is monitored by growing oocytes in a buffered saline medium supplemented with ^{35}S-methionine (preferably). The amino acid is readily taken in by the oocytes. Secreted proteins and cellular ones can be distinguished in the fractionated aliquots of the homogenized cell lysates. Care is taken not to break the sulphur bonds in the mature polypeptides.

Xenopus oocytes do not translate organelle or prokaryotic mRNAs.

4.3.8 Expression in Transgenic Animals and Plants

The systems described so far in which gene expression may be observed involve isolated cells and *in vitro* conditions. Although such studies provide much detail about the molecular mechanisms involved, they do not tell us how the engineered gene acts in the context of the whole plant or animal. Is it possible to introduce a foreign gene in such a way that it becomes part of the gene pool for the species? In other words, will the individual transmit the gene to its offspring? Again, will the gene be expressed correctly in space and time? That is, will it be expressed in the proper tissues at its normal developmental stage?

The only way to answer these questions is to insert foreign genes into germ line tissues and look for evidence of their presence/activity in the offspring of the treated individuals. Such exercises have been performed with mice and the answer has been in the affirmative to both the earlier questions. The offspring from the treated parents, containing and expressing the introduced gene, are referred to as transgenic individuals. The technique of producing transgenic mice has become fairly well established. Other species, including those of plants, are also being converted into transgenic ones.

Initially, transgenic mice were created by infecting germ line cells with retroviral vectors carrying the foreign gene (Fig. 4.37). The technique has been augmented by that of direct microinjection into the male pronucleus of the fertilized egg. The first transgenic mice were developed in the laboratory of Beatrice Mintz (Jaenisch and Mintz, 1974); SV40 DNA-carried genes were introduced in the blastocoel cavity of the early stages of the mouse embryo. Today, there are other ways of achieving the same goal. Retroviruses have been employed to generate transgenic mice; preimplantation embryos are infected with the virus or are kept in contact with infected cells that produce free retroviruses. The infected embryo (morula) is transferred to pseudo-pregnant female mice, where the embryos develop. Evans *et al.* (1986, 1987) have infected embryonic stem cells (ES) in culture and injected them into the blastocyst.

The retroviral genome becomes integrated at random sites in the chromosomes of the recipient cell. One way of ensuring that a specific chromosome is to carry the 'transgene' into the offspring is to use Evans' method of infecting cultured cells. The site of integration may be determined in these cells by somatic cell hybridization and the desired construct chosen to inject into the embryo.

Direct microinjection into the pronucleus of the zygote results in transgenic animals that are not genetic mosaics, unlike the animals generated by retroviral infection. One consequence of this difference is that in the case of microinjected embryos, almost half the offspring will be transgenic, while a lesser proportion of the retroviral infected ones will carry the foreign genes.

On the one hand, DNA introduced directly is seen to integrate into the recipient chromosome as several copies in tandem at the site of integration. On the other hand, retroviral insertion introduces one gene copy per integration, although the retrovirus may be integrated at several sites in the same recipient.

One difficulty that became evident quite early in the days of producing transgenic mice was that constructs with plasmid or phage vector DNA did not express the gene in the transgenic animals. The reason for this failure is being investigated. In the meantime, techniques are being perfected to introduce DNA without vector material.

The most significant aspect of this method of gene transfer into individuals is that the transferred gene is expressed at the appropriate times in appropriate tissues. This offers a route for gene replacement therapy, a method of charting the lineage of cells during development, and a means to alter regulatory signals and factors, the effect(s) of which can be monitored in the proper developmental and physiological contexts.

Of the different vertebrate proteins that have been used as transgenes, the globin appears to be sensitive to the site of integration in the host DNA. The others so far have not been significantly affected in expression by the location of integration. The reason for this difference is not known. Perhaps the globin gene possesses regulatory sequences that do not become included in the limited fragment of the cloned gene. Answers to questions such as this one and other finer nuances in the hierarchy of regulation systems in the whole animal may be revealed by investigations using transgenic individuals.

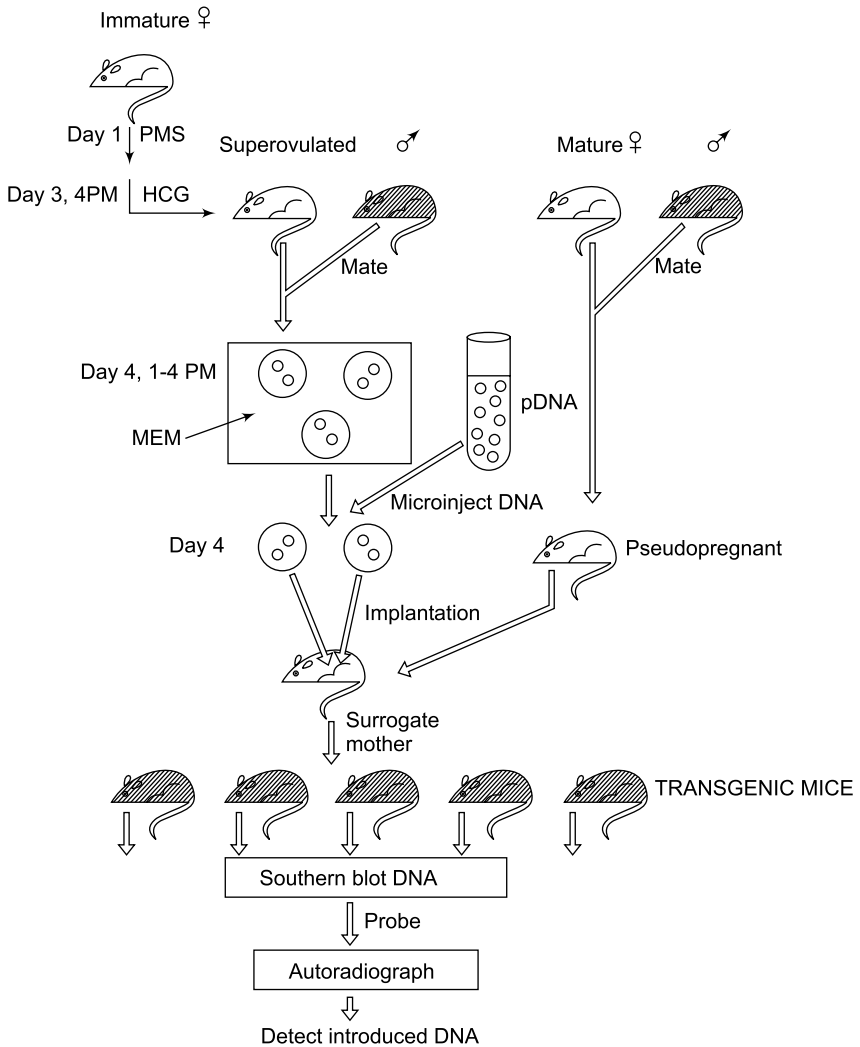


Fig. 4.37 *Algorithm for the production of transgenic mice. Fertilized eggs of mouse were transformed with the foreign gene and allowed to develop in surrogate mothers, which had been treated hormonally to accept this role. The litter included transgenic mice. The presence of the foreign gene in them was confirmed.*

REVIEW QUESTIONS

1. Describe direct selection techniques with reference to clone identification.
2. Enumerate and expand on indirect screening techniques of clone identification.
3. How do nucleic acid probes help in pinpointing exact molecules in gene-cloning ventures?
4. Describe Southern blotting techniques and development of other techniques based on it.
5. Enumerate the immunodiagnostic probes and briefly account for their utility.
6. Review the issues involved in expression of cloned genes, and consider the procedures employed.
7. How do we study prokaryotic gene expression *in vivo* and *in vitro*?
8. What are the experimental strategies employed to overcome difficulties encountered during eukaryotic gene expression in *E. coli*?
9. How are the transcription and translation of eukaryotic genes in cell-free systems achieved?
10. Assess the utility of *Xenopus* oocytes for expression of exogenous genes.
11. How is the expression of foreign genes through transgenic animals and plants studied?

5 Applications and Advances in Genetic Engineering

Introduction

The many-splendoured phenomenon of molecular biology burst on the scene almost a century after the birth of the concept of a gene. The double helix model of the genetic material provided the insight to Mendel's factors and the molecular antics of the genetic material began to be comprehensible. New facts were requisitioned into applications and became the basis for more incisive tools; recombinant DNA technology had emerged.

5.1 ❑ MUTAGENESIS

Mutagenesis is the process by which a lesion in the DNA becomes converted in a subsequent replication into a heritable genetic alteration. The alteration includes, substitution, addition or deletion of a single nucleotide or a set of nucleotides.

Agents that cause mutations are known as *mutagens*. Mutagens include physical, chemical and biological ones. Physical mutagens are the various ionizing and nonionizing radiations and heat. Chemical mutagens cannot be so readily classified but include predominantly alkylating agents and others that induce modifications in the nucleotide, and acridine dyes that bring about frame shift mutations due to the deletion or addition of bases. Substituting mutagens include nitrous acid, hydroxylamine, sodium bisulphite and alkylating agents such as ethyl- and methylmethane sulphonate (EMS, MMS) and N-methyl-N′-nitroso-N-nitroguanidine (MNNG).

A substitution may be a transition or a transversion. The former is the replacement of a purine by a purine (A $\rightleftharpoons$ G), pyrimidine by a pyrimidine (T $\rightleftharpoons$ C), or a purine-pyrimidine (or vice versa) pair by another purine-pyrimidine (or vice versa) pair. A transversion is the replacement of a purine by a pyrimidine or vice versa, or a purine-pyrimidine pair by a pyrimidine-purine one (or vice versa). Biological mutagens are transposons and integrating viruses, particularly retroviruses, that become inserted within functional sequences (coding or regulatory) in the DNA and disrupt the normal activity of the altered site.

The lesions caused by mutations are converted into permanent changes or mutations after a subsequent DNA replication. The strand opposite the one with the altered site may base-pair with the altered base in the lesion in an atypical manner;

the latter acquires its legitimate partner in the following replication. This ds DNA is now the parent for a clone of mutated cells.

Mutants have perhaps been the most useful tools for various projects of geneticists that attempt to understand the nature and function of a gene in terms of local expression and global development. And mutagens became available with H G Muller's use of X-rays on *Drosophila,* Karl Sax's on plants, and Charlotte Auerbach's pioneering use of chemicals. However, these earlier methods could not generate lesions/mutations at specific regions, leave alone at a particular single base location, in the genome. The randomly produced mutations were bred in various ways with other genotypes, and their nature and relative map positions ascertained.

Recombinant DNA techniques provided, for the first time, ways of causing mutations in very specific sites, and also of a predetermined nature.

Perhaps the most spectacular success in refashioning a DNA has been through the techniques of directed mutagenesis. By instituting mutations of a known nature in a target site, it has become possible not only to produce a DNA with novel attributes but to test the role of different regulatory and coding sequences in specific functions. Indeed, the role of every base in the sequence of interest can be tested for its contribution to the relevant function. Additions (or insertions) of single or of several bases, as well as deletions, may be achieved in a target-specific manner by several gimmicks that make use of the available data bank on the biology of phage, virus and eukaryotic genetic systems. The most specific alteration is achieved with the use of synthetic oligonucleotides. Some of these directed ways of mutagenesis are described briefly. Finer refinements of these techniques and others abound in almost every important publication in this field.

1. Oligonucleotide-directed Mutagenesis

Oligonucleotide directed mutagenesis (Fig. 5.1) can install a desired alteration in any site in a stretch of DNA. The technique is based on the use of a synthetic oligonucleotide (7-12 bases) that is identical to (complementary) the site to be mutated, except in one or more bases. This synthetic short length of DNA is hybridized with a ss DNA vector (usually from ϕ M13), and acts as the primer which elongates to make the M13 template duplex. The two strands of this circular DNA are separated and the new strand used as a template to synthesize its partner. The mismatched bases in the original oligonucleotide are now matched with the correct complementary base(s). A mutation becomes established in this manner at the site of the bases that were different in the oligonucleotide. Almost all techniques based on this pioneering one (Hutchinson and associates, 1978) utilize a ss DNA as the template. Vectors from phage M13 and fd as well as plasmids (ss) of the pEMBL series are suitable for the purpose. All of these may be packaged as ss DNA-containing particles.

While the technique appeared to be quite simple in principle, the strategy was found to encounter certain snags during practice. The chief among these obstacles was the propensity of cells to repair mismatched regions of a DNA. The second difficulty was the pushing out of the oligonucleotide by the advancing 3′-end of the new DNA strand which had reached the 5′-end of the oligonucleotide. Thirdly, the primer often hybridized to unintended regions of the template with which it had partial homology.

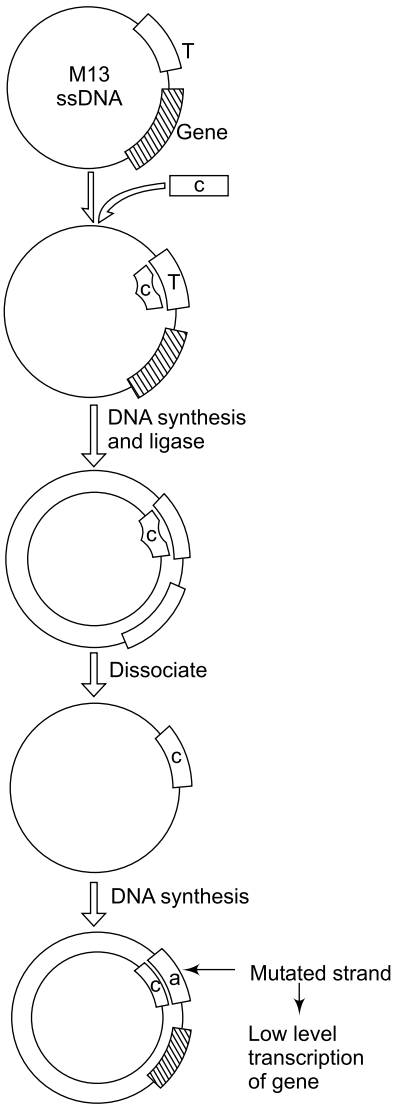


Fig. 5.1 *Site-directed mutagenesis: the principle of the earlier technique. A synthetic oligonucleotide complimentary to a region of the ss DNA vector (M13) but differing in one non-complementary nucleotide (C in place of A) is allowed to anneal to the vector. DNA synthesis extends the oligonucleotide until it makes the M13 DNA a duplex molecule. Ligase joins the free ends of the new strand and the inserted oligonucleotide. The ds circular DNA is dissociated and the new strand used as a template to synthesize its complementary strand. The new strand will possess a G residue opposite the C (lesion) in the synthetic oligonucleotide. The TA pair in the wild-type DNA has thus been mutated to GC pair.*

The first difficulty (of removal by repair or proofreading by DNA Pol I) is avoided by ensuring that the mutated strand (i.e., the one with the oligonucleotide) is methylated at appropriate bases. Endonucleases do not attack fully methylated or

even hemimethylated (one strand) DNA, except for a very few known restriction enzymes (e.g., *Dbn* I). The mutated strand is maintained by ligating the end of the new strand with the oligonucleotide. Hybridization at incorrect locations is avoided by allowing only a limited region of the template for attaching the oligonucleotide primer.

These trouble-shooting jobs led to various improvements in the technique, so that it is now a reliable and routinely used one for site-specific mutagenesis. Kramer *et al.* (1984) developed the following elegantly simple protocol (Fig. 5.2). The template was a ss M13 DNA containing a short insert at the region to be mutated. This rDNA is an RF molecule synthesized while in a methylation-deficient (*dam*$^-$) *E. coli* host. A non-recombinant M13 DNA (ss) is obtained by infection of a methylation-competent strain (*dam*$^+$). The non-methylated RF recombinant DNA is dissociated and mixed with the methylated M13 DNA, which has been linearized. Double-stranded hybrid DNA is formed between the non-methylated ss RF strand and the methylated ss M13 DNA. There is a gap in the latter opposite the insert in the former. The oligonucleotide is introduced to the left end of this gap and allowed to synthesize a new strand to fill the gap. The nicks between the new strand and the existing Ml3 DNA are ligated. By these tactics,

(i) the mutated strand is not vulnerable to attack by endonucleases (particularly the four GATC nucleosides in M13 DNA, which have been methylated in the N^6 of the adenine),
(ii) the primer does not hybridize at unwanted regions, and
(iii) the length of new strand to be synthesized is quite short.

This last feature prevents possible non-completion of strand synthesis due to incorrect primer attachment etc., and saves on time and material.

The preceding 'gapped duplex' approach has been further refined and modified both by Kramer and associates and others. Instead of having an insert in the template DNA, a gap may be generated by removing the short *Eco* R1-*Hin*d fragment present in the M13 DNA. When this shortened DNA is hybridized with a full length DNA cleaved elsewhere, a gap occurs in one strand and in a single strand region opposite the gap in the other strand. Dalbadic-McFarland *et al.* (1982) produced a similar gap in one strand by nicking the latter at a site and degrading the nicked strand using exonuclease III, that had a $3' \rightarrow 5'$ activity.

The approaches mentioned above also reduced the number of steps in producing the ds mutant DNA. After the filling up and ligating in the gap region, it may not be necessary to dissociate the two strands and use the mutated one to acquire a new complementary strand. As the template strand is unmethylated, the DNA repair enzymes favour correcting the mismatch in this strand rather than in the one with the introduced oligonucleotide. Such repair would result in a ds DNA that was mutant and without a heteroduplex region. However, all mismatches are not corrected with equal frequency; in which event, it is better to rely on other means of selection.

One such strategy is to use selection by complementation. The strand with the oligonucleotide could carry a *lacZ* gene, while its partner carries the same but with an amber mutation. So, the template DNA will be able to grow only on *E. coli* with a suppressor for the amber mutation. The mutated DNA will grow on a wild-type (non-suppressor) *E. coli* and will transform a *lac* host strain (Kramer *et al.,* 1984).

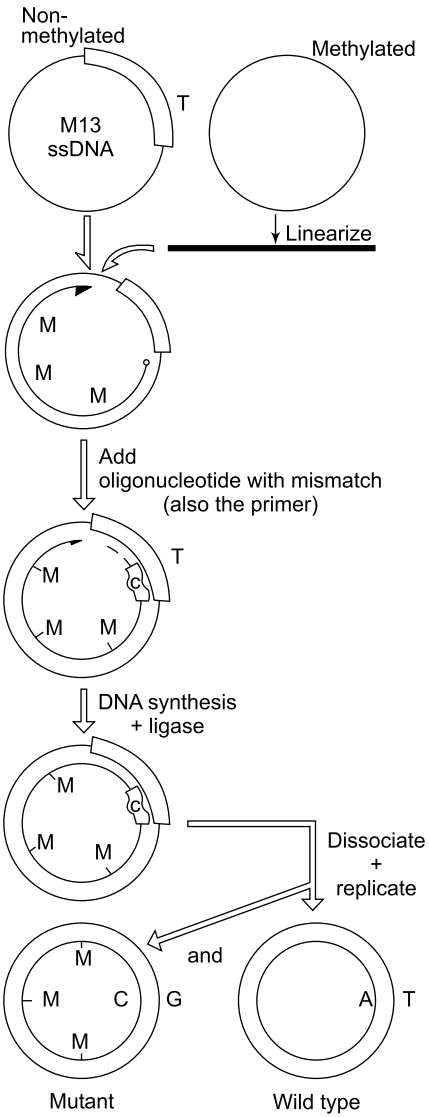


Fig. 5.2 *The site-directed mutagenesis: An improved technique. DNA to be mutagenized is cloned in an M13 vector. The ss DNA of the vector, without the insert, and which has been methylated at the C residues, is linearized and allowed to hybridize with the vector containing the insert. There will be a gap in the linear DNA opposite the insert DNA. An oligonucleotide with the desired mismatched base and a primer for DNA synthesis are added to the incomplete M13 duplex. DNA synthesis from the 3′ end of the methylated strand generates the complementary strand for the insert DNA (except for the mismatch, C, in place of A). Ligase circularizes the linear DNA. The ds DNA is dissociated and each strand allowed to replicate. Half the recovered circular DNAs will be of the wild-type, while the other half will be the mutants with the GC pair in place of the wild-type AT pair. This method limits DNA synthesis to a defined short region, a feature which makes it less vulnerable to various defects including non-specific annealing of the primer.*

2. Deletions

Deletions may be introduced in a DNA by a variety of methods, as follows:

(i) One way would be to remove a short segment flanked by restriction sites, the shortened DNA ligated to form a closed circle, and the latter used to transform *E. coli* cells.

(ii) A second method is the one mentioned earlier (Dalbadic-McFarland *et al.*), where a single nick with an endonuclease is followed by the use of exonuclease III ($3' \rightarrow 5'$), or the λ exonuclease ($5' \rightarrow 3'$).

(iii) Another strategy uses the removal of the unhybridized single-stranded loop in a hybrid between a full length DNA and one with a deletion. The S1 nuclease chews up the single-stranded loop, leaving the intended DNA with a deletion size of the loop, or a few bases longer (20–30 b) than the loop (Shenk, 1977).

(iv) The exo III may be used to create a deletion in the following manner: After an initial nick, the use of exo III degrades the nicked DNA starting from the 3′-end. This generates an extended 5′-tail. The latter is degraded with S1 nuclease; the result is a blunt-ended and shortened DNA. These ends are ligated with T4 ligase. *Bal* 3′ may also be used, as this removes both strands at one end simultaneously, leaving a blunt-ended fragment. However *Bal* 3′ operates on the distal end of the nicked DNA, shortening it from the other end. This would not matter if there were no important genes or regulatory signals in the vicinity of the cut. After shortening the DNA insert, the latter may be snipped off from the vector and relegated to an intact vector and then amplified by *E. coli* transformation. (Fig. 5.3).

(v) A more general method for creating a deletion is to generate a D-loop in a ds circular DNA (Fig. 5.3). This is achieved when an ss DNA, complementary to a short region in one of the strands, is added to superhelical ds DNA. The strands of the latter dissociate at the region of homology and the ss strand becomes hybridized in the bubble. The non-hybridized strand bulges out to form what is known as a D-loop. Adding the *RecA* protein also induces the formation of a D-loop in mixtures of superhelical DNA and a ss DNA piece complementary to a region of the ds DNA. The moment the ss DNA in the loop becomes available, the S1 nuclease attaches it at either end and degrades it. Ligation of the shortened DNA results in short deletions.

Several other ingenious methods have been reported that delete a single or several nucleotides at specific regions of a DNA.

3. Insertional Mutagenesis: Retrovirus Mediated

Insertion of a retroviral sequence within a gene or its regulatory sequences cause mutations in the functions at the site of insertion. Insertional mutagenesis has been utilized to create developmental mutants in rodents, by infecting germ line cells with retroviruses.

Insertional mutagenesis with phage Mu has become a common practice for recovering mutants from *E. coli.* The idea of using integrating viruses for the same purpose in the case of higher organisms emerged when it was discovered that many cases of oncogenesis were due to the integration of retroviral sequences within an

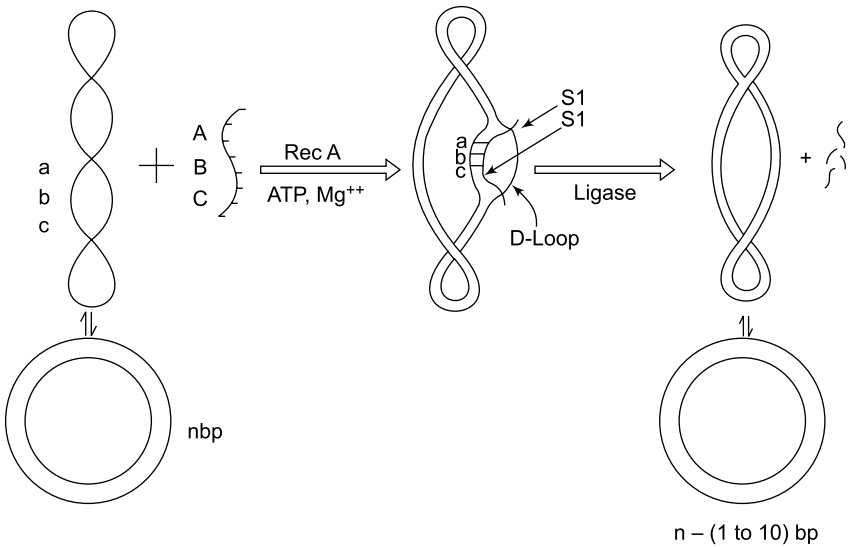


Fig. 5.3 *Deletion by the D-loop method. This technique is based on the phenomenon that on addition of a strand of DNA complementary to a region of a circular ds DNA, the strands in the complementary region separate to form a D-loop (deletion loop), and the single-stranded fragment hybridizes with the complementary region in the loop. All ss DNA regions are removed by the action of S1 nuclease. Ligation of the shortened ends automatically generates short deletions in the region occupied by the loop. This technique is useful for removing short lengths of DNA not only next to a restriction enzyme-cleaved region but also from any other region of a DNA.*

oncogene (the mutated version of a developmental gene, which causes cancerous transformation of a cell).

Jaenisch *et al.* have utilized several strategies to introduce retroviruses into germ line cells. The proviral sequence showed up in several members of the litter. When the latter were mated, the proviral sequence was found to be transmitted in a typical Mendelian fashion.

The site of insertion of the provirus was random; both sex chromosomes and autosomes were affected. However, only one of a pair of homologues showed the presence of the proviral sequence. Some offspring exhibited the phenotype as a spread of virus infection in all the tissues. Eventually, these affected animals developed the condition of leukemia. No other phenotype due to the viral integration could be discerned.

4. Insertional Mutagenesis: Transposon Mediated

Like retroviruses, transposons insert their copies into random locations in the chromosome(s), and generate mutations in the disrupted loci.

If the transposon carries a selectable marker gene, the affected clones may be identified. The clones are then tested for aberrations in functions that have been knocked out or modified.

'Transposon mutagenesis', as this mode of creating mutations is known, has been valuable in locating genes in several systems. The *nif* genes of *Rhizobium meliloti* (Ruvkin and Ausubel, 1983; Ditta *et al.,* 1980) and the tumour-causing ones in the T DNA of *Agrobacterium tumefaciens* have been pinpointed using this technique.

The transposon that is mostly utilized is Tn5, which carries the *neo* (neomycin, kanamycin, G-418 resistance) gene. In practice, the source of the Tn5 is a vector constructed by Berg (1974) called the λ *kan hoppers.* These are λ phage carrying Tn5. When such a vector is introduced into a cell, the Tn5 'hops' or transposes a copy into any other DNA in its vicinity.

Ditta *et al.* (1980) mutagenized a 3.9 kb *nif* cluster by Tn5 insertion. The *nif* was on pRMR2 (pYCAC 184 *nif DNA*). The mutated *nif* was transferred to pRK290 which has a Tet^R marker. Finally, the altered *nif* was introduced into the *Rhizobium meliloti* cell where it replaced the wild-type *nif* sequences by recombination.

5. Directed Single Base Mutations

To circumvent the difficulty of knowing which base is mutated by treatment, systems have been evolved that produce predictable alterations in single bases, within a defined region of the target DNA.

Single bases may be altered by the incorporation of analogues of nucleotides or, specifically, by the use of sodium bisulphite.

An analogue of a nucleotide can be incorporated during DNA replication. If this analogue prevents the proofreading activity of DNA Pol I, the mistake remains in the newly synthesized strand. When the latter replicates, a nucleotide that pairs favourably with the analogue is incorporated. The progeny from the last DNA strand will carry a base-pair that is different from the one at that locus in the original DNA template. Some of the most convenient analogues that fit the bill are the α-thiophosphates of the four nucleotides. They allow chain elongation by the polymerase activity but not degradation by the $3' \rightarrow 5'$ exonuclease of DNA Pol I (Shortle *et al,* 1982).

Weissman developed a method for the misincorporation of nucleotide analogues in reasonably defined regions. A DNA is cut with a restriction enzyme in the presence of ethidium bromide, which nicks only one of the two strands at the restriction site. Nick translation is initiated with one of the four NTPs as an analogue. Weissman used the enzyme *Eco* R1 and N^4-deoxycytidine triphosphate (N^4-hydroxy dCTP) as the analogue. The analogue becomes incorporated in place of thymidine. On replication, the DNA with the analogue includes the incorporation of a purine. If the purine is C, the original A-T pair will have been replaced by a G-C one. If it is an A, no alteration will be visible. Hence, half the population after mutagenesis should be mutants. As the region affected in Weissmann's study was in the *Eco* R1 sequences (G/AATTC), mutants were recognized by their inability to be cleaved at the altered recognition sequence.

The bisulphite mutagenesis method was developed by Shortle and Nathans (1978). Sodium bisulphite alters a cytosine in a single strand of a DNA into uracil (by the removal of the amine group). After synthesis, the double strand will possess a U-A pair, which on subsequent replication will be converted to A-T. The technique consists of first inducing a region of a ds DNA to become single stranded; this step

is followed by one of bisulphite treatment. The first objective is achieved by causing a nick in one strand using a restriction enzyme in the presence of ethidium bromide. A C-rich recognition site is preferable; hence *Hpa* II is generally utilized (CCGG). After nicking, the nicked strand is shortened by using an exonuclease, which preferably degrades in a $5' \rightarrow 3'$ direction; this would expose the recognition site in the remaining single strand. DNA Pol I from *Micrococcus luteus* is appropriate for this purpose; it degrades the DNA if only one NTP is available in the reaction mixture. When all four NTPs are supplied, this polymerase uses its polymerizing capability. Single-stranded regions may be induced by the D-loop creating technique (see Fig. 5.3). Of course, the mutagenesis may also be carried out on ss DNA cloned in M13 vectors or pEMBL plasmids. Several other tricks have been employed to generate single-stranded regions that will serve as substrates for bisulphite mutagenesis. Usually, the uracil-containing DNA is transfected into cells that lack the ability to repair the U-region (i.e., Ung^6 cells with a mutation in the gene for uracil-N-glycosylase).

Other sophisticated techniques that make use of the ability of an electrophoretic gel to discriminate between DNA strands that mismatch are also in use (Myers *et al.,* 1985).

5.2 ❑ MAPPING THE DNA

Genes have been mapped traditionally into relative positions by genetical breeding methods. Today, it is possible to map the DNA in physical terms.

Gross landmarks, in the form of sites for restriction cleavage, are assigned by techniques of restriction mapping. A small specific segment of a DNA strand may be identified and isolated by the method of DNA footprinting. Large regions of a DNA may be encompassed by strategies of chromosome walking or jumping, while genes may be assigned to specific bands (defined regions) of the human chromosome by a combination of somatic cell and *in situ* hybridization techniques. Of course, the finest mapping is achieved by DNA sequencing, the techniques for which are described in Section 5.3.

5.2.1 Restriction Mapping

Restriction mapping sketches a gross picture of a DNA on the basis of sites for restriction enzymes. These sites are constant in a DNA from the same strain or species, with allelic variations (polymorphism) as in the case of gene sequences. Any DNA from the same genetic background will, therefore, be cleaved into identical sets of fragments by a restriction enzyme. Making a restriction map is the first step in characterizing a genome or a cloned DNA.

The procedure for restriction mapping involves the treatment of DNA with a restriction enzyme, fractionating the fragments according to size, and analysing these fragments to ascertain the position of each in the intact DNA.

There are two methods for restriction mapping; direct and indirect. In direct methods, the DNA is digested to completion, and in the indirect ones only partial digestion is allowed.

One variation of restriction mapping that has found new applications is the determination of restriction fragment length polymorphism (RFLP). The essential features of direct and indirect approaches in restriction mapping, and one of the cases of RFLP determination, are described briefly.

As the size of fragments is the only criterion for restriction mapping, the quality of mapping depends on the accuracy of determined sizes. These are calculated from their mobilities in the gel, with reference to those of standard size markers.

Southern has developed a useful algorithm for calculating sizes from mobilities of fragments. In practice two maps are made, starting with the two ends of the DNA. The fragments that are of identical size in both gels are noted, and ambiguities resolved by different means. Both maps should come out to be identical in the end. This operation can be carried out manually, but it is time-consuming and error-prone.

1. Direct Mapping

Direct mapping is based on the analysis of partially digested DNA.

In one protocol, Smith and Birnstiel (1976) end-labelled one end of the DNA selectively and after partial digestion (so that only one cleavage per DNA was possible), the fragments with the label were examined. They formed a nested set of fragments ranging in size from the smallest to the largest. The length of each fragment represented the distance, from the labelled end of the intact DNA, occupied by a cleavage site for the enzyme employed.

Another protocol end-labels ends of the DNA and cleaves the latter once to separate the two labelled fragments. Further, partial digestion of both sets of labelled fragments with another enzyme generates fragments, the analysis of which leads to the required map.

2. Indirect Mapping

Indirect mapping procedures are based on complete digestion of the DNA with two or more enzymes, separately or in double digestion reactions.

The second digestion is often carried out in the gel itself, in a second dimension. The sub-fragments (if any) of each initial fragment are identified. Mapping becomes complicated as the length of DNA increases. To cope with the complications recourse is taken to a systematic approach, such as the use of the 'branch and bound' technique (Fitch and associates, 1983). This involves using certain defined rules to assign sub-fragments to the initial parent fragments; the technique simultaneously eliminates the alternate (but incorrect) fragments that could also add up to give the size of the fragment in question. Pairs of single digestion fragments that share one double digestion product are identified, and the ones that appear to be ambiguous are taken care of by the protocol of the technique.

This procedure is very much shortened by using DNA labelled at both ends and incorporating any other known data for any part of the DNA under scrutiny. In some cases, there may be too many complications to allow an unambiguous solution based only on the data of the fragment size.

5.2.2 DNA Footprinting

DNA footprinting refers to identifying sequences in the DNA that are bound to specific proteins (e.g., enzymes, transcription factors, growth promoting factors). The technique involves the removal of the ss region of the DNA not bound by the protein, usually with the help of S1 nuclease. The protein is degraded with a proteolytic enzyme, and the released piece of ss DNA cloned and sequenced.

5.2.3 Chromosome Walking

Chromosome walking is a technique by which mapping begun in one DNA fragment of an overlapping genomic library can be continued into the adjacent ones. Starting with any fragment one can 'walk' to the preceding or succeeding sequences of the overlapping fragments. If the process is continued, the entire map is established by 'walking' across consecutive fragments.

The walk is begun with a probe for one end of a fragment. The probe is then used to screen the library of overlapping fragments. The fragment to which the probe hybridizes is placed in tandem with the initial fragment with the probed regions overlapping each other. The two fragments thus unite to form a larger fragment. Probes from both ends of the latter screen other fragments from the collection and meld the positive ones as before. One, therefore, 'walks' forward or backwards starting from any one fragment.

Probes for the ends of the fragments are prepared in the following manner. The fragment is cloned in a vector that carries a specific strong promoter (T3, T7, SP6) just upstream of the cloning site. Replication begun at the promoter, and in the presence of two or more radioactively labelled nucleotides, yields specific probes. A more arduous method is to identify ends of restriction fragments and then isolate the fragments for preparing the probes.

Chromosome walks are useful not only for melding the sequences of large fragments of DNA into a single sequence but also for identifying the sequences on either side of a sequence of interest. Quite often, the latter is a gene that is too large to have been included in the phage or even cosmid vector. The entire gene, including the 5′ and 3′ flanking regions may be mapped by identifying clones that represent overlapping regions of the DNA. Clusters of genes have also been mapped in this manner. Gitschier *et al.* (1984) mapped the gene for the human clotting factor VIII using this technique.

5.2.4 Chromosome Jumping

Chromosome jumping is an offshoot of the technique of chromosome walking, where one can skip segments of the DNA represented by a λ-clone to reach a fragment of interest.

Chromosome jumping is carried out basically in two ways. In one approach, 'junction' sequences are cloned and used as probes to identify sequences several kilobases (up to 100 kb) away from the region hybridized by the probe. The technique involves the following steps:

(i) The DNA is cleaved into large fragments (100 kb or more).
(ii) The fragments are spliced to plasmid vectors and the hybrid DNA circularized.
(iii) The hybrid DNA are cleaved with an enzyme that does not cut within the vector. Usually, a large insert has more than one cleavage site for an enzyme. The insert region after cleaving may circularize into a smaller molecule, due to the loss of one or more internal fragments. The vector now possesses a shortened insert representing the two ends of the latter. A probe is made from this 'junction' fragment and used to screen the fragment library. A probe from one end of the junction is thus sufficient to identify a very large segment of the DNA. One thus 'jumps' over long distances of the DNA and maps the two extremities of such a large fragment (Collins and Weissman, 1989).

A second technique of chromosome jumping utilizes the presence of chromosomal rearrangements (such as an inversion or a translocation). The approach is identical to straight chromosome walking until one end of the rearranged region is reached. The probe for the overlap between the terminal in the last fragment before the rearrangement now skips a fragment or two before it hybridizes with another clone. In some cases, this is due to an inversion in the region; in others, due to a translocation of the segment next to another chromosome.

5.2.5 Mapping Genes by Somatic Cell Hybridization

Somatic cell hybrids between different mammalian species usually exhibit the phenomenon of loss of chromosomes of one species over several rounds of mitoses. This feature has been exploited to assign genes to chromosomes of the human genome.

Human cells (primary or cultured, normal or diseased), are fused to rodent cells (usually mouse or hamster that is mutant for an enzyme or other phenotype that is easily selectable). The culture media prevent growth of unfused human and rodent cells. The hybrid cells are transferred to a limiting medium in which only those hybrids survive and proliferate which possess the gene that allows overcoming of the mutation in the rodent partner. Let us say in a man-mouse hybrid, the mouse cell is tk^-. In a HAT medium, only hybrids with the tk^+ gene would grow. This gene must be from a human chromosome. All hybrids are scanned for a common chromosome. It is a safe bet that the *TK* gene is on this chromosome.

This fundamental approach is combined with other data and genes assigned not just vaguely to a chromosome but also specifically to regions in either arm of a chromosome. The human genome has been mapped at an exponential rate in this manner. Retention of a chromosome is correlated with the presence of a particular gene. It is confirmed by the loss of function of that gene in cells that lack the particular chromosome (Creagan and Ruddle, 1975; Shows *et al,* 1982.)

Chromosomal abnormalities are also pressed into service for assigning genes to human chromosomes. The procedure is the same; the function is now correlated with the presence or absence of the chromosomal segment in the hybrid cell. The genes *hprt, g6pd, pgk* (phosphoglycerate kinase) and the β-galactosidase gene have been mapped to the human X chromosome using this technique.

It is particularly easy to distinguish the activity of human genes in a man–mouse hybrid. In some cases, the mouse marker gene is experimentally inactivated. In others, distinct differences between the two allow identification of expression of a human marker gene. Many of the markers are isozymes that are distinguishable by their electrophoretic mobility, antigenecity or reassociation kinetics.

The site of a gene on the human chromosome in a man–mouse somatic hybrid cell can be assigned unambiguously by using a probe for the wanted gene. The probe may be from another species, provided there is sufficient homology between it and the wanted gene. The probe is used on Southern blots of restriction fragments of human genes. These can be distinguished from the fragments of the mouse DNA generated by the same enzyme.

Single copy gene sequences are currently mapped by *in situ* hybridization of metaphase chromosomes with an appropriate probe. The chromosomes are banded by a high-resolution banding technique (produces more than 1000 bands per chromosome). This allows assigning the probed gene to exact locations in the chromosome (Zabel *et al.* 1985). Morton *et al.* (1986) have employed high-resolution banding with *in situ* hybridization of human chromosomes in man–mouse cell hybrids to assign the genes for insulin-like growth factor (IGFF) I and II to chromosomes 12 and 11, respectively.

5.3 ❑ DNA SEQUENCING

Sequencing a DNA means mapping the DNA in terms of its nucleotide sequences. Sequencing became possible when methods were developed for generating DNA fragments with a defined nucleotide terminus. Two methods, chemical and enzymatic, have been developed that are used routinely, with or without modification of the initial protocols.

Both the preceding methods are based on fragmenting the DNA into four aliquots, with fragments in each aliquot ending at one end with one of the four nucleotides. The other ends of the fragments are identical. This results, in each aliquot, in lengths of DNA that form a nested collection. The fragments in each aliquot are separated on a sequencing gel (8 per cent polyacrylamide/8M urea, 400 mm long vertical gel). The four gels are aligned side by side and the order of fragments, from the smallest (at the bottom of the gel) to the largest (next to the intact whole length DNA fragment), is read off in terms of the nucleotide at the end of the fragment. In practice, all of the DNA fragments in a reaction mixture are not placed in the well of the gel in one go. Aliquots are removed at intervals from the start of the reaction and placed in the well. The earliest aliquots are at the bottom of the gel and the last one at the top. Within each aliquot the sequence is read from the bottom upwards. This procedure is preferred as short tracks are not distorted on the gel, whereas there can be considerable distortion at the bottom of a long track.

Figure 5.4 represents the distribution of DNA fragments in four gels. The fragments obtained by the *chemical method* of Maxam and Gilbert, terminate in a particular base at the cut end. Those generated by the enzymatic method also end in a specific base in each reaction set. The fragments are radioactively labelled at the 5′-end. Sanger and Coulson's enzymatic method creates a set of nested fragments

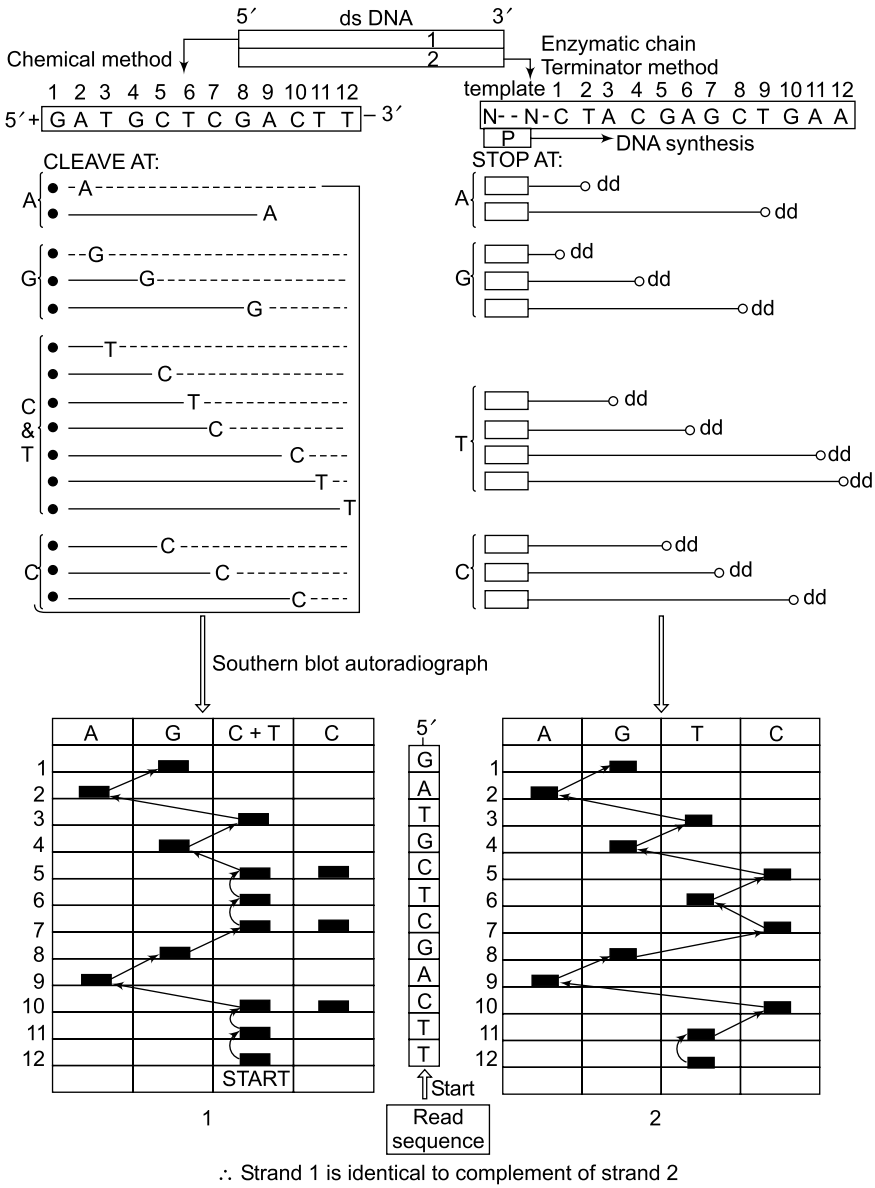


Fig. 5.4 *DNA sequencing: chemical and enzymatic chain termination methods.*

beginning from a defined region (a primer), with consecutive fragments differing in lengths by a single nucleotide. Both methods employ four reaction mixtures; in the chemical method, fragments in one mixture are cleaved after the same nucleotide. In Sanger's method, each reaction mixture contains a dideoxynucleotide of one of the four kinds. Incorporation of a ddNTP in place of a dNTP stops further DNA strand elongation. Any reaction mixture, therefore, contains samples of new strands that

have stopped at least one of the sites occupied by its complementary nucleotide in the template strand. The fragments from each mixture are fractionated on a special vertical polyacrylamide electrophoretic gel that can separate DNA strands differing in length by even one nucleotide. The sequence is read off from the four gels which are aligned next to each other.

The gels are Southern blotted and autoradiographed. The sequence is read from the autoradiogram. The process of reading the bands has been streamlined and made practically error-proof by using an electronic device to scan the intensity of each band in the gel (due to fluorescence of ethidium bromide) and record it in a computer. This digitizer allows rapid and correct recording of every band in the four gels.

The original chemical method devised by Maxam and Gilbert and the enzymatic method by Sanger and Coulson have been replaced by modifications that are more rapid in execution and less arduous in operations. The chemical method has also been automated by the use of a sequencer instrument. The enzymatic method depends on the availability of a pure strand of DNA to serve as the template. The initial time-consuming methods of isolating pure collections of ss DNA are replaced with Messings' and Wu's modifications. In the former, the RF vector from the ss DNA phage M13 is exploited for carrying the inserted fragment that is to be sequenced. In Wu' method, a linear DNA is degraded from the 3′-end to the 5′-end with exonuclease III until only a central region of ds DNA remains. This region is cleaved to yield asymmetric fragments, each of which is used for sequencing.

The nested fragments show up on the autoradiogram due to the labels on the DNA fragments. In Maxam and Gibert's method, the labels are at one specific end of the fragments. In the Sanger and Coulson's method, the label is distributed within the fragment due to the presence of the label in one or more of the nucleotides used to synthesize the fragment. End-labelling of the 5′-terminus is achieved by replacing the resident phosphate (using alkaline phosphatase) with γ^{32}P ATP (using polynucleotide kinase). The 3′-end may be labelled by using the enzyme TNT and the four α^{32}P-labelled ribonucleotides; a subsequent alkali treatment removes all labelled ribonucleotides except the one at the original 3′ end. A 3′ recessed end may be labelled by using the Klenow fragment and the four labelled deoxynucleotides to make the end a flush one.

Reliable and rapid sequencing is presently achieved by either of the two methods. Some modifications of the chemical method are preferred for DNA lengths less than 0.5 kb, and those of the enzymatic method for DNA longer than 5 kb. The latter procedure is also more cost-effective than the chemical one.

5.3.1 Maxam and Gilbert's Chemical Method

Maxam and Gilbert's chemical method depends on the production of four sets of fragments, each set with one identical end and the other ending in one particular nucleotide. That is, in one aliquot, all fragments end in G, in another in A and so on. The identical end is labelled. These fragments are created by chemical reactions that cleave the DNA preferentially after a particular type of nucleotide.

The cleavage reaction consists of three steps: (i) modification of a base next to the base of interest (the one that is to terminate a fragment in one reaction mixture);

(ii) removal of the modified base, thus creating a gap in the base sequence; and (iii) cleavage of the backbone at the site of the gap.

The DNA fragment to be sequenced is usually a restriction fragment, one end of which is labelled at only the 5′ or 3′ nucleotide or at both. The labelled DNA is dissociated and the single strands fractionated on a denaturing gel. If the DNA had labels on ends of each strand, the ds DNA is first cleaved into asymmetric sized fragments and each of the latter sequenced separately. The two strands of each fragment are sequenced individually. This is a check for the validity of the derived sequence; the results from the two strands should be complementary to each other. The ds strands are dissociated by mild alkali or heat.

The cleavage is carried out in four reaction mixtures, with conditions in each aliquot for cleaving after one particular kind of nucleotide. The reaction conditions (i.e., reaction temperature, reaction time and concentrations of reagents) are chosen to allow only one base to be modified per DNA strand, or a very few in the entire strand. Since modification occurs at random in the various likely sites, the reaction mixture contains the end fragments that possess the labelled end and terminate in one of the required positions. Say, the reaction is set for cleaving the DNA only after a G. The final collection of fragments will have every length of the DNA that ends in a G, and begins with the labelled nucleotide.

The fragments from each reaction mixture are fractionated and read as described earlier.

The four reactions that cut the DNA at G, G + A, T + C and C (and also if needed at A + C) are based on the following principles. In each case the reaction weakens the N-glycosidic bond between the ribose and the base moieties:

(i) To cleave after a G, the DNA is methylated with dimethyl sulphate and treated with alkaline piperidine that removes the modified G (N^7 -methyldeoxy-guanidine) and cleaves the backbone at the site of the gap.

(ii) To cleave after an A, the DNA is methylated as done earlier, but treated subsequently with acid and piperidine. Only the backbone at the site of the N-3 methylated adenine is cleaved. Thus in one reaction mixture, fragments end in G, and in the other in A, depending on the treatment after methylation.

(iii) Both T and C are modified by hydrazine and removed by piperidine; the latter also breaks the DNA backbone at the site of the gap as in (i) and (ii).

(iv) If hydrazine treatment is carried out in a background of 1M NaCl, only C is affected, and the DNA cleaved at the gap left by removal of a C.

A fifth reaction, involving NaOH treatment followed by piperidine, is sometimes set up; the data from this merely confirming the positions of bases derived from the use of the first four reactions. This fifth reaction cleaves at both A and C sites, but more at A than at C under equivalent conditions.

You will notice that there is no reaction that specifically terminates a DNA at a T site. The positions of T's are determined by comparing the lanes in the gels for fragments cleaved at T and C and only at C. The bands that are not common to both lanes are due to cuts made at Ts.

A minimum of 1 picomole of DNA is sufficient for sequencing by this method. Although factors such as the incubation time, the concentration of reagents and the temperature of reaction are the parameters that control how many bases are to

be modified by a reaction, in practice, incubation time is used as the controlling variable.

5.3.2 Sanger and Coulson's Enzymatic/Chain Terminator Method

The enzymatic method of sequencing DNA also results in an array of fragments, each of which differs from the ones larger and smaller than it by one nucleotide each. In this case the fragments are synthesized in four reaction mixtures, up to a particular nucleotide.

The method is based on synthesizing a labelled strand of DNA starting from a primer annealed to a template and terminating the synthesis at one of the four nucleotides. The terminator is a dideoxynucleotide that has no 3-OH to which another nucleotide may be added. The terminator is mixed with the corresponding deoxynucleotide in a proportion that allows the terminator to be incorporated randomly in at least one of each of the sites for the particular nucleotide. The reaction mixture thus contains, at completion time, fragments that represent lengths of DNA from one identical end to each of the sites for a particular base. The gels are run, autoradiographed and interpreted as already described for the chemical sequencing method.

The enzymatic synthesis (or dideoxynucleotide chain terminator) method requires the following material: (i) a short primer; (ii) a pure ss DNA template; (iii) the four terminators (ddATP, ddGTP, ddCTP and ddGTP); (iv) the four dNTPs of which some or one or two are ^{32}P labelled in each reaction mixture; and (v) a DNA polymerase without $5' \rightarrow 3'$ exonuclease activity.

The reactions are carried out in four aliquots in small volumes (0.5–10.0 ml) either in micropipettes or microfuge tubes. The polymerase is usually the Klenow fragment. Avian reverse transcriptase may be utilized with appropriate alterations in the buffer and proportions of dNTPs.

1. The Primer

The length of the primer should be between 30 and 80 kb. The primer may be a nick-translated oligonucleotide or a short restriction fragment. The primer DNA is usually purified by fractionating on a polyacrylamide gel, eluting from the latter by the electroelution technique, and passing the eluted DNA through a DEAE cellulose column. The primer fragments may be concentrated by cold ethanol precipitation and stored a –20°C in an appropriate solution.

2. The Template

The DNA to be sequenced is usually cloned in a λ-vector. The λ-DNA has regions rich in cytosine. By adding uracil and guanidine, the strand with more Cs is made heavier than the other. The two strands are dissociated and separated by caesium chloride density gradient centrifugation. The two strands are pooled in separate tubes and dialysed.

Messing's M13 phage method simplifies the procedure for obtaining ss DNA templates for enzymatic synthesis. This will be described later.

3. The Dideoxynucleotide Terminators and Deoxynucleotides

In each reaction four deoxynucleotides and one dideoxynucleotide are mixed in a given proportion. One of the deoxynucleotides is $\alpha^{32}P$-labelled. The dideoxynucleotide that is to compete with its deoxynucleotide version is kept at a higher concentration with respect to the cognate deoxynucleotide. The chains to be stopped at As have dideoxyadenine as the terminator. Similarly, ddGTP, and ddTTP terminate chains in nucleotides with the corresponding bases.

4. The Polymerase

Klenow fragment or avian reverse transcriptase are used, as both lack the 5′- to 3′-activity. Its absence prevents degradation of the primer from the 5′-end.

The primer is annealed to the template before the synthetic reaction is begun. The primer (ds DNA) and the ss DNA template are boiled in a mixture, and the latter transferred to an oven at 60°C. (Boiling dissociates the primer strands, one of which hybridizes with the template at 60°C). The operation is performed in small capillary tubes that are sealed after filling with the primer and template mixtures.

Four reaction mixtures are prepared with a unique terminator in each. The primer-template mixture is added to this together with the DNA polymerase and incubated at 18–24°C for 15 minutes. The reaction is stopped and the mixture loaded into a sequencing gel. The primer may be removed before placing the mixture in the gel by cleaving with a restriction enzyme (the site for which is restored at the 3′-end of the primer as strand synthesis begins).

5.3.3 Messing's Shot-Gun Method

Messing developed a rapid sequencing variation of the enzymatic method that bypasses the error-prone and laborious steps of isolating pure ss template DNA and primers specific for one of its ends. The method is said to be a 'shot-gun' one as the DNA to be sequenced is broken or cleaved into random fragments. Each fragment is cloned into an M13 RF vector and the latter hybrid DNA used to transform JM101 strains of *E. coli* (Fig. 5.5). Each clone of cells will represent one of the fragments. The vectors used are of the M13p series, of which M13p2 is the most widely used for sequencing.

The M13p2 contains a part of the *lac* operon of *E. coli,* including the *lacI* gene (repressor), the promoter and the first section of the *lacZ* gene that encodes the α-region of β-galactosidase. A polylinker at an *Eco* R1 site between the 5th and 6th codons of *lacZ* is present in M13p2. The insert DNA is cloned by cutting the polylinker with the appropriate restriction enzyme.

The RF molecule of the M13 phage is cleaved and the insert DNA ligated at this site. The hybrid RFs are used to transform the *E. coli* cells. After a few rounds of replication of the ds DNA, ss circular DNA (positive strand) are synthesized. These DNA become enclosed in the phage proteins and are extruded from the host cell into the growth medium. The DNA is released from the phage particles and used for sequencing. The template is already in an ss DNA state, and all phages from the same plaque have template DNA in the same orientation. Other plaques may possess the same DNA fragment, but in the opposite orientation.

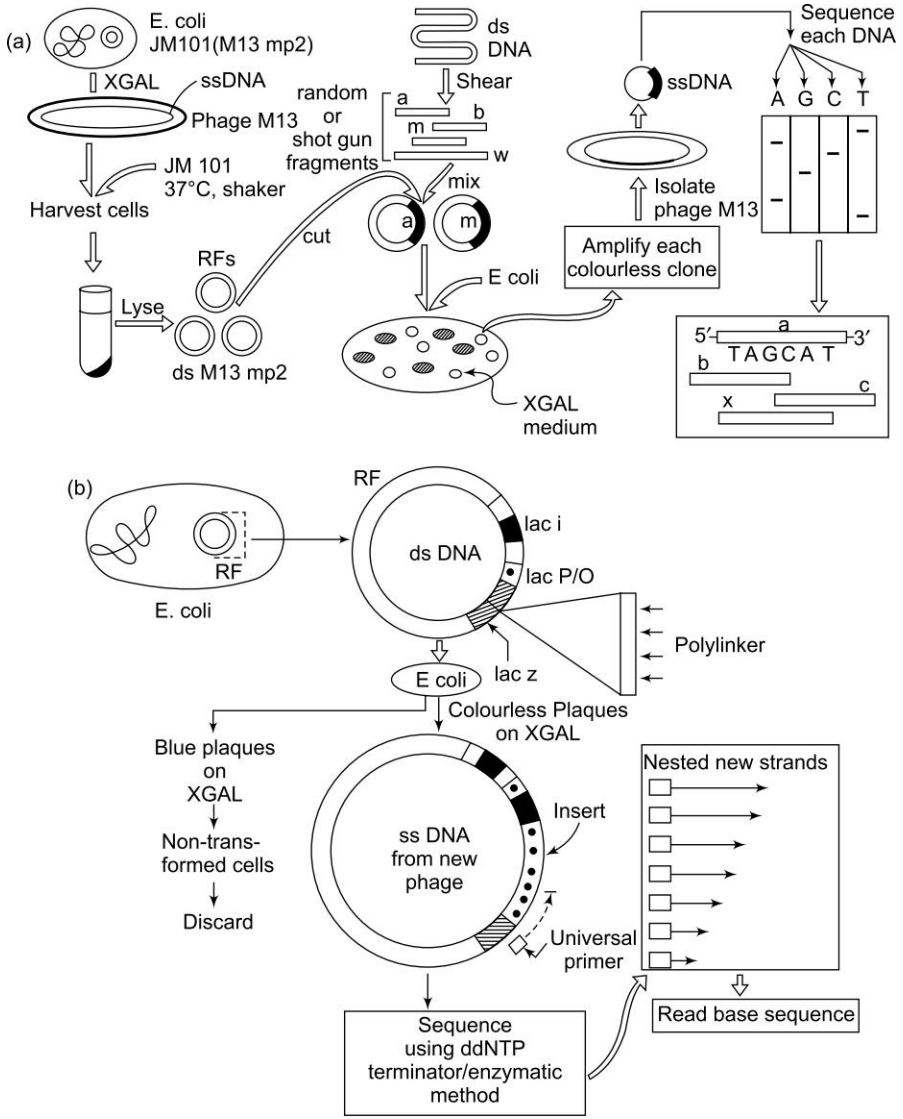


Fig. 5.5 *Algorithm for Messing's shot-gun method of DNA sequencing. RF (replicative form) molecules of vector M13mp2 are extracted from the E. coli strain JM101 which harbours the vector, and are used to ligate randomly (shot-gun) sheared DNA molecules of the DNA to be sequenced. Cells transformed with the mixture of hybrid and unhybridized vectors are identified on XGAL media. The recombinant DNA-containing cells produce colourless plaques. Each colourless plaque is amplified, its hybrid ss DNA extracted from the phage M13 particles and each hybrid DNA sequenced by the enzymatic dideoxynucleotide termination method. Use is made of computer programs for melding the sequenced fragments into the integrated sequence.*

The defined nature of the M13p2 makes it easy to have a single primer for every sequencing project that uses this vector. The primer is complementary to a region on the vector, having its 3-OH end at the start of the cloning site. The common primer, available commercially, is known as the 'universal primer'. The latest one is a 15-base oligonucleotide that hybridizes just to the right of the polylinker (Norrander *et al.,* 1983 Messing *et al,* 1981). A vector with the insert DNA disrupts the *lacZ* gene, so that recombinant cells show up as colourless colonies in an XGAL supplemented medium. The colourless colonies are picked up, grown to amplify the number of cells, and the DNA from them extracted. This DNA (circular, single-stranded) is utilized for sequencing, which is performed using the Sanger's chain terminator method.

The template DNA may be obtained by DNase I cleavage or by mechanical shearing, and the ends suitably processed (made blunt or provided with restriction enzyme ends). Totally random fragments are generated in this way. They do represent every part of the DNA; but due to amplification, several clones may represent the same fragment. This creates a problem as sequencing proceeds, because one comes up again and again with the same clone. To avoid this slowing down of the sequencing job, half the DNA is sequenced by the Messing's method, and the remainder concluded by using the Sanger's original chain terminator method.

The following is an outline of the steps followed in Messing's method of sequencing DNA:

(i) Pure colonies of *E. coli,* JM101 and blue plaques of M13p2 phage streaked on JM101 in an XGAL plate, are isolated.
(ii) The phage is mixed with the host cells in suspension and incubated with shaking (at 37°C). Mature phages will be extruded in the surrounding fluid.
(iii) This mixture is added to a suspension of JM101 cells grown from a single colony innoculum. The mixture thus obtained is incubated with shaking at 37°C for 4 to 6 hours.
(iv) The cells are harvested, lysed and the ds RF DNA extracted from them. This DNA is purified by CsCl gradient centrifugation at 15°C (in three steps at 136 k, 90 k and 90 k × g respectively) for 24, 36 and 36 hours, respectively.
(v) The DNA to be cloned and sequenced is cleaved or sheared, and the ends modified suitably.
(vi) Each fragment is ligated to an RF molecule and used to transform *E. coli.* The treated cells are plated with XGAL. Colourless plaques are picked out and transferred to individual plates to amplify the clones of each fragment. DNA isolated from each clone (phage) and used for sequencing. The ss DNA within each phage contains the vector DNA, the *lacZ* sequences, the insert DNA and the appropriate sequence at the beginning of the insert for annealing to the universal primer.
(vii) The DNA from each clone is sequenced using the chain terminator method.

5.3.4 Using Computers for Sequencing and Analysing DNA Sequences

DNA sequencing is performed on fragments that are 200–400 bases in length. The larger DNA to be sequenced is converted to these sequenceable lengths. How

does one order the different fragments that have been sequenced into the correct sequence of the larger DNA?

In essence, each fragment is compared with the sequences of all other fragments in order to look for regions that overlap or are common to each other. By placing these fragments together so that the common regions overlap, one can produce larger sequences. The process is continued until the entire sequence is reconstructed and no unmelded fragments remain. Computers have been recruited to carry out these comparison and melding operations; manually, such operations take an inordinate amount of time and are prone to incorporation of many errors. Several computer programs have been developed to do this job quickly and efficiently. The pioneers in this venture include R Staden (1980) of UK and L J Korn and C L Queen (1980), and T R Gingeras and R J Roberts of the USA.

The methods consist of recording the sequence of each fragment in a memory bank. The sequence from one set of fragments is then compared with those from every fragment of a second set. The moment two fragments have some sequences in common, they are 'melded' into one longer sequence (called a CONTIG in one program, from *contiguous sequences*). This sequence is matched with others, and so on, until the CONTIG becomes one long sequence incorporating all the fragments. Of course, this is the base outline of the protocol. Several sophisticated improvements have been introduced to make sure that a long stretch of the same nucleotide is not mistaken for overlapping regions, to name just one of the pitfalls of melding sequences.

Computer programs are also becoming indispensable for analysing the DNA sequences for the genic and non-genic regions of interest. Comparisons of sequences are fast becoming the strategy for discovering evolutionary relationships between species and for a host of other objectives. Data for particular motifs in the nucleotide sequences, such as Pribnow boxes, SD regions, TATA and CAAT boxes, motifs recognized by transcription factors and restriction enzymes, to name a few, are stored in the computer. An unknown sequence may be scanned by one of the several SEARCH programs with the speed and ease afforded by the use of computers for such operations. Nucleotide sequences of genes are also stored. These sequences are utilized for a variety of objectives that include finding homologies with other sequences of interest. Some software packages are already available commercially. Others may be obtained from various laboratories, on request.

5.4 ❑ PLANT GENETIC ENGINEERING

Attention has been focused on plant genetic engineering only recently. It appears at present that in spite of this tardy and modest beginning, plant genetic manipulation may become a routine application in the near future, a goal that does not appear to be as readily attainable in animal genetic engineering.

Plants possess one major advantage over animals that allows the engineering of novel traits in whole populations of a strain beginning with a single genetically tailored cell. Plant cells are differentiated readily into mature, fertile plants by supplying only a few hormones to the culture media. This ability to regenerate from single cells not only allows studies of expression of exogenous genes introduced into

single cells or to recover products of such genes in bulk but also provides genetic engineers with a handle for introducing improvements into agriculturally and other commercially profitable plant species.

As in the case of non-plant species, rDNA applications are based, in plant systems, on modes of gene transfer into single cells and whole plant (transgenic) analysis of native and foreign gene expression within them and examination of the gene structure in plant systems. A better understanding of the molecular biology of plants and their natural vectors will provide the foundation for meaningful manipulations of plants of interest.

The rigid cellulosic wall of plant cells is an effective barrier to the easy entry of naked DNA. Hence, transformation exercises are carried out on cells whose walls are removed enzymatically. A second strategy for entry into plant cells takes advantage of natural entrants such as bacterial and viral pathogens. Additional mechanical methods of transferring genes into plant cells include making the cell membrane more permeable to exogenous DNA by electroporation and 'shooting' fine globules of metal coated with the DNA of interest into plant cells. Neither of these last two techniques causes adverse effects in the cell.

Studies of plant gene expression have been responsible for the development of a variety of assays, both qualitative and quantitative, that are used for gauging the expression of the gene of interest. Some of them have been employed in animal and bacterial molecular genetical studies and have been adapted for use with plant material. These include isolation of DNA species, run-off transcription in isolated nuclei, preparation of probes to be used for hybridization both on Western blots and *in situ* hybridization in cells and tissues.

Studies of expression of foreign genes are carried out in single cells (transient expression) and in whole plants regenerated from genetically engineered single cells. The latter technique is not yet universally applicable, as only a very few species have been found to be amenable to regeneration by current technologies. A considerable volume of research is concerned, therefore, with finding ways of achieving regeneration of transformed cells of different plant species. Monocotyledonous plants are particularly recalcitrant to regeneration by current gene transfer systems. However, as cereal crop plants are monocots, there is a need to understand their molecular biology on the one hand and to search for natural or engineered vectors for gene transfer on the other.

Studies of plant gene structure involve the use of techniques for DNA isolation and analysis. These techniques have been adapted successfully for plants and for the genomes present in their organelles (mitochondria and chloroplasts). Productivity in plants is dependent on their cell respiratory and photosynthetic capacities. An enquiry into the molecular biology of these organelles is, therefore, warranted. Their genomes would be vulnerable to manipulation; the goal would be increased efficiency of performance by the tailored plant.

As the basic techniques of recombinant DNA and ancillary technology are the same as those described earlier, the following discussion will highlight only those features which are relevant to working with the plant system.

5.4.1 Methods of Gene Transfer in Plants

Gene transfer may be effected in plants by methods that (i) use naked cells (protoplasts) as recipients of exogenous DNA, (ii) are based on the use of recombinant vectors, (iii) involve mechanical introduction of the DNA into cells. Some of these approaches are briefly introduced below.

1. Protoplast Fusion

Plant protoplasts may be fused to generate inter-specific or intra-specific cell hybrids. The hybrid protoplast regenerates a cell wall within 24–48 hours. Somatic cell fusion is one way of introducing bulk DNA into a recipient cell.

A single metaphase chromosome may be transferred to a protoplast via a liposome or a microcell. The fusion is achieved as, described earlier, with the agent PEG.

Protoplasts of haploid cells may be fused to generate hybrids that may be made viable by converting them to allopolyploids. This objective is achieved by blocking cell division after mitosis with colchicine; the resultant cell thus possesses double the number of chromosomes of each parent—a feature that makes segregation of chromosomes possible in subsequent cell divisions.

Microbial cells have been fused with plant protoplasts to transfer recombinant plasmids from the former to the latter. This is a more efficient method of introducing exogenous DNA into a plant cell than by the conventional technique of transformation with naked DNA. There is evidence for this from fusions of *Agrobacterium tumefaciens* with plant protoplasts. *A. tumefaciens* possesses a very large plasmid (Ti) that is more readily introduced by cell fusion than by DNA-mediated transformation.

2. Organelle Engineering

Heterozygosity of extranuclear genetic information may be attained by somatic cell fusion. This information may be useful for, say, improving the photosynthetic quality of one variety by importing chloroplasts from another. However, this route has not had much success. The main problem is the non-retention of the chloroplasts from either of the parental cells. Until more is known about why and how organelles are maintained and segregated, meaningful applications would be difficult to achieve.

In the majority of studies no recombination has been found between the DNA from chloroplasts of different parental stocks. In one report, *Nicotiana tabacum* and *N. plumbaginifolia* protoplasts were fused and some progeny cells were observed to be recombinants of the two chl DNAs (Fig. 5.6). *N. tabacum* had genes for streptomycin resistance, a suppressor for the str^R gene, and lincomycin resistance. The second species was str^S and $linc^S$. Some of the progeny cells were str^R and $linc^S$. *N. tabacum* was sensitive to streptomycin due to the presence of the suppressor gene. The str^R gene from this parent segregated and contributed to the resistance phenotype in progeny hybrid cells.

3. Recombinant Vector Techniques

A foreign gene may be introduced into a plant cell via a vector derived from plasmids or plant viruses. Some vector DNA remain independent of the host cell DNA, while

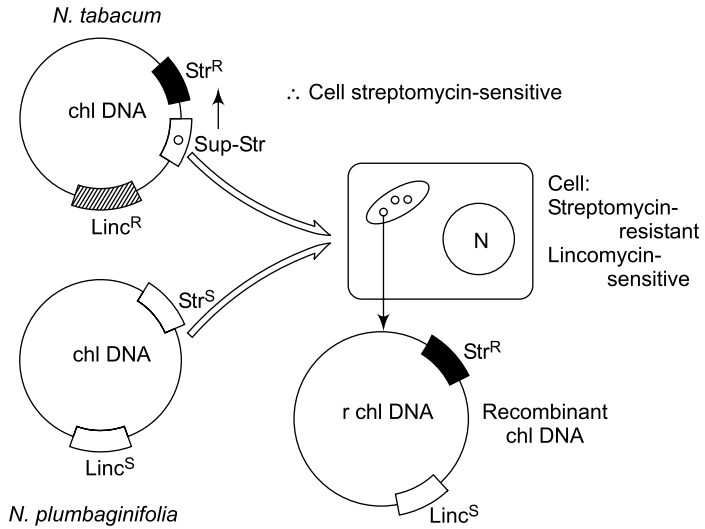


Fig. 5.6 *Organelle engineering. Transforming cells of tobacco species (Nicotiana tabacum) with chl DNA from another species (N. plumbaginifolia). The chl DNA of N. tabacum carries genes for lincomycin resistance, streptomycin resistance, and a suppressor gene that nullifies the action of the latter. Co-transfection of a Nicotiana cell with both chl DNA results in recombination between them which is evident from the recovery of cells that are both* str^R *and* $linc^R$ *.*

others become integrated into the latter. Both integrative and non-integrative DNA transfer approaches are being investigated as methods of gene delivery into plant systems.

I. Non-integrative DNA transfer Non-integrative vectors are based mostly on the nucleic acid of plant viruses. The ones that have been studied most are the caulimo viruses that infect *cruciferous* species (mustard, cauliflower, radish, etc.). Gemini viruses with their ss DNA genome also appear to be potential candidates for supplying plant cell vectors. Viruses with RNA genomes have also been considered for the same purpose. There is one main disadvantage of using viral vectors; they are pathogenic. This obstacle may be overcome when we know the molecular mechanisms underlying pathogenic symptoms. Plasmids of the bacterial species belonging to *corynebacteria,* that infect cereal crop plants, have recently come into prominence as yet another source of vectors for use in plant genetic engineering.

(i) Caulimoviruses These are small viruses with a ds DNA genome. The best studied member of such a virus is CaMV (cauliflower mosaic virus) that is transmitted by aphids (a kind of insect).

The DNA of the CaMV may be introduced into plant cells by rubbing leaf surfaces with the extracted DNA. It rapidly infects other cells neighbouring the initial recipients and is amplified to about 10^5 particles/cell. As the DNA remains non-integrated, its propagation into subsequent generations of plants cannot be guaranteed. The DNA remains in the nucleus.

The CaMV DNA is 8 kb long and may assume a linear or open circular, or pretzel-like, knotted form. There are three or four discontinuities in the DNA strands. One is in the A strand with 8 genes (or ORFs) and two or three (depending on the variety) in the other strand. There is no gap at these discontinuous regions as the free end of one strand overlaps for a short distance the free end of the other strand after the discontinuity (Fig. 5.7).

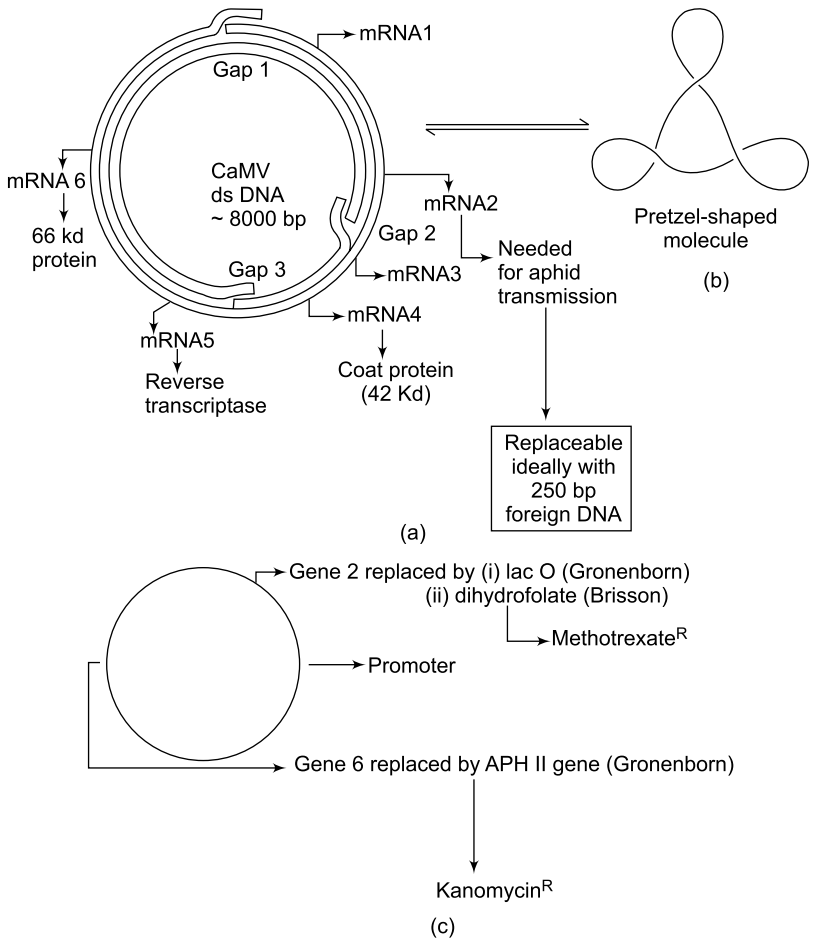


Fig. 5.7 *Schematic map of the ds DNA of cauliflower mosaic virus (CaMV). Six ORFs have been identified, of which ORF II appears to be hardly essential for infection and development traits of the virus. A lacO/P, inserted in this gene, does not seem to affect viral properties. It has also been replaced by the dhfr gene, which is functional in cells transformed with such a vector. Although a DNA virus replication uses a 35S RNA as the template for synthesizing the DNA genome, the reverse transcriptase for this operation is encoded by ORF V.*

The overlaps are digested and the nicked ends of each strand joined before transcription. The most abundant transcripts are a 35S species and a 19S species that possess a common starting region. All the ORF transcripts are included in the 35s species.

There are certain characteristic viral-encoded protein-based structures in the cytoplasm known as inclusion bodies. Viral DNA replication and particle construction are assumed to occur within them. RNA transcripts synthesized in the nucleus are translated in the cytoplasm using host cell RNA Pol II. A 35S RNA serves as the template for synthesis of a ss DNA with the help of reverse transcriptase. The ss DNA acquires a complementary partner after the 35S RNA and primer are degraded.

CaMV DNA has one region (~1 kb) within the 8 gene cluster that includes several sites for *Eco* R1. ORF II encodes a gene that is necessary for transmission by aphids but not for infecting cells. This region of the DNA is dispensable and may be replaced by the gene to be introduced. The *dhfr* gene has been inserted at this site in the laboratory of Hohn (Brisson *et al,* 1984). The *dhfr* gene was correctly expressed in the transformed cells.

The difficulties raised by the size constraint and lack of sufficient clonable sites in the CaMV DNA have been partially overcome by construction of vectors with parts from CaMV and another vector system (Ti plasmid). This feature will be described later.

The promoter for the 35S transcript is a very strong one, and may be used without additional virus-encoded or host-encoded factors, especially tissue-specific ones. There is a 59 b stretch upstream of the promoter that contains a CAAT box, a palindromic sequence and an enhancer sequence, which is possibly responsible for the efficiency of the 35S promoter. The last one is used to prefix foreign genes to be expressed in transformed plant cells.

(ii) Gemini viruses Gemini viruses are ss DNA plant viruses usually transmitted by leafhoppers. The genome consists, of one or two circular molecules which are 2.5 kb to 3 kb in length. After entering the plant, the virus moves into the bast tissue (phloem or food channel) and into the cell nuclei. The ss DNA replicates in special proteinaceous bodies within the nucleus via a ds DNA intermediate.

Gemini viruses infect several plant species. Their host ranges may be extended by the technique of 'agroinfection' (see following Section III), which involves insertion of the viral DNA in the Ti plasmid DNA. BGMV (bean gold mosaic virus) is the best characterized Gemini virus; it infects leguminous plants. Its DNA may be in both linear and circular forms.

(iii) Plant RNA viruses Although recombinant molecules cannot yet be made with the RNA of RNA viruses, the problem may be side stepped by using the cDNA made from the RNAs. The hybrid cDNA may be transcribed back into the complementary RNA, and the latter used to infect host plants. Alquist and associates have used this strategy to transfer a drug-resistant gene into barley plants, which are monocots.

Koziel and Siegel (1981) have developed a system using TRV (tobacco rattle virus) RNA. The TRV is fragmented asymmetrically and each RNA enclosed in coat proteins. The larger particle lacks genes for viral coat proteins. The smaller particle lacks all genetic information except for the genes for the coat protein. The two complement each other when cotransferred into a cell, to produce complete viral

particles. Koziel and Siegel aim to utilize the smaller RNA as a vector for transferring genes into plants.

(iv) Corynebacterial plasmids Corynebacteria infect a wide spectrum of cereal grasses. These bacteria contain small plasmids that may be exploited as vectors for plant genetic enginerring.

II. Integrative DNA transfer Integrative DNA transfer into plant cells may be achieved with plasmids of the *Agrobacteria* species, transposable elements and homologous DNA vectors. None of these systems, as they occur in nature, is ideal for use in research or application. Hence, mixed vectors have been developed that possess parts from a variety of vectors from plasmid, viral and cDNA sources.

(i) Plasmids of Agrobacteria The most promising of the integrative vectors are the Ti (tumour-inducing) and Ri (root-inducing) plasmids of *Agrobacterium tumefaciens* and *A. rhizogenes* respectively. Most studies are based on the Ti plasmid of the former species.

(a) Biology of A. tumefaciens *Agrobacteria* are soil organisms that infect dicotyledonous species. The bacteria enter the plant through damaged regions close to the soil-air interface. Wounded plant tissues exude certain phenolic compounds, two of which appear to act as chemoattractants for these bacteria. They are acetosyringone and an α-hydroxy derivative of the same. These compounds enhance the expression of two plant genes *chvA* and *chvB* (chromosomal virulence A and B), that are normally expressed constitutively at a low level. The attracted *Agrobacteria* attach themselves to the wound regions with the help of cellulose fibrils (Fig. 5.8).

The acetosyringone stimulus also triggers off a chain reaction that begins with the transduction of signal(s) by the phenolic compound across the bacterial cell membrane to a protein on the cytoplasmic surface of the membrane, encoded by the plasmid bourne *virA* (virulence) gene. The latter is one of seven *vir* loci present as a cluster in the agrobacterial plasmid. *virG* regulates the expression of the remaining *vir* genes, which are involved in the transfer of a copy (transposition) of a region of the Ti plasmid known as the T DNA; the latter integrates into the plant cell DNA.

Two other developmental pathways are known in which proteins initiate a unique phenotype. These are (i) the triggering of the *tra* genes of the *E. coli* F. plasmid by products of the *E. coli* gene *SfrA,* and (ii) the activation of the *nif* (nitrogen fixing) genes of *Rhizobium meliloti* by the products of the bacterial genes *ntrB* and *ntrC.* These gene pairs show extensive homologies with the *virA* and *virG* sequences.

The integrated T DNA is transcribed in the nucleus into at least four different transcripts. One of them is for an enzyme (synthetase) that mediates the synthesis of a sugar-amino acid complex called an 'opine' (Fig. 5.9). There are different opines, of which two, octopine and nopaline, are the most abundant ones. Each Ti plasmid carries a gene for the synthetase for one type of opine. They are, therefore, classified according to the opine synthetase that they specify (e.g., octopine and nopaline Ti plasmids).

The second transcript is due to the gene *tmr* that specifies an enzyme needed for the synthesis of a cytokinin, isopentenyladenine. The third and fourth transcripts

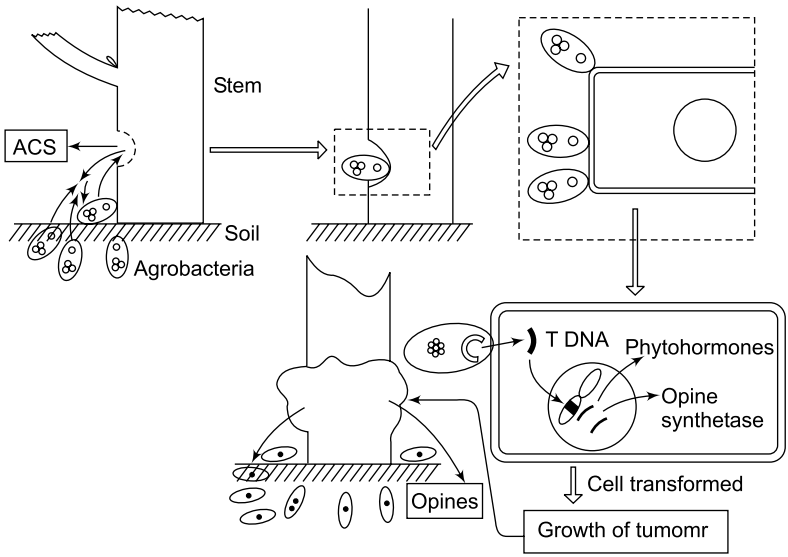


Fig. 5.8 *Biology of Agrobacterium tumefaciens. This soil bacterium enters dicotyledonous plants through wounds near the base of the plant. Wound tissue exudates chemoattractants for the bacteria. The Ti plasmid of the bacterium contains a transposon T DNA which is introduced into host cell DNA after infection. Phytohormones expressed from genes in the T DNA trigger off plant cells to divide and form a tumour or gall. T DNA genes express nitrogenous compounds called opines that are catabolized by enzymes, also encoded by the Ti plasmid. The plant cells provide the setting for the production and use of opines, which can be used as sources of C and N, only by agrobacterial species.*

are from the genes *tms* I and *tms* II that encode an enzyme for the synthesis of the auxin, indoleacetic acid (IAA). The overproduction of the auxin and cytokinin results in the tumorous proliferation of plant cells containing integrated T DNA. A mutation in *tmr* results in the development of roots (rooty mutations), and in the *tms* loci to differentiation of shoots (shooty mutations) from the undifferentiated tumour tissue.

To summarize A Ti plasmid is a very large circular ds DNA containing a transposable T DNA region and a non-T DNA region (Fig. 5.10). The latter carries the replication origin of the plasmid and the *vir* cluster. It also contains a gene, the product of which is required to catabolize the opine synthesized by the plant cell on specifications of the opine synthetase gene. The *Agrobacteria* utilize the opines as the chief source for N and C. No other bacterial species can make use of opines. The T DNA contains four genes, three of which are required for synthesis of phytohormones (auxin and cytokinin) and one for opine synthetase.

The T DNA has two 25 bP repeats at its two ends. These border sequences (RB and LB) are essential for transfer and integration of a copy of the T DNA into the plant DNA. The *vir* loci products are essential for the transfer of the T DNA.

The *tms* and *tmr* genes alter the regulatory pattern of the host cell, causing tumorous replication of the T DNA integrated cells. By amplifying the cells that produce the opine, the *Agrobacteria* manipulate an ever-widening environment

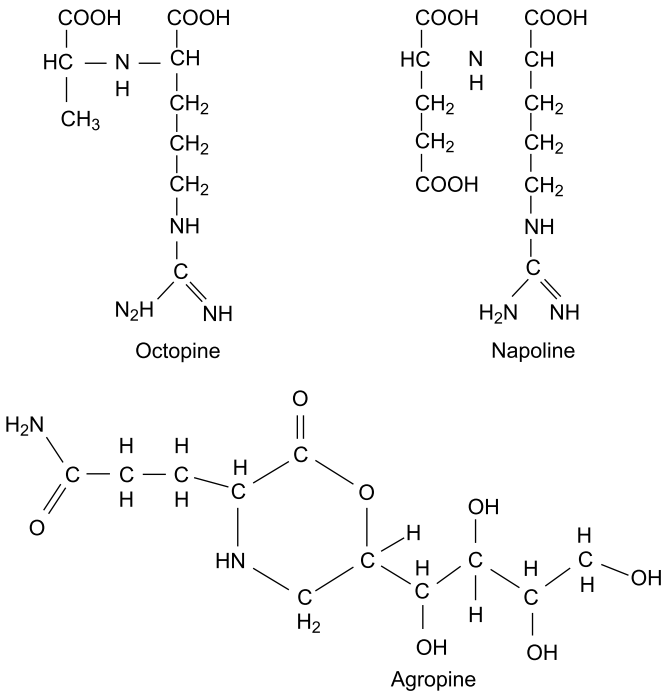


Fig. 5.9 *Structures of three types of opines. Opines are unique derivatives of arginine.*

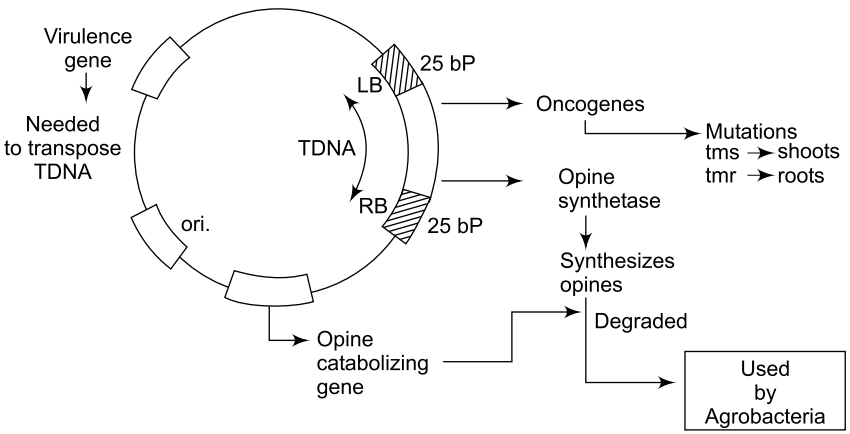


Fig. 5.10 *A schematic representation of a generalized map of Ti plasmids. The T DNA is subtended by repeated sequences (25 bp) characteristic of transposons. The transposition ability resides in these terminal repeats (left and right borders or LB and RB). Plant gall tissue is undifferentiated. However, a mutation tms may upset the balance between root and shoot-differentiating hormones and differentiate the callus into shoots. Similarly, a mutation tmr induces differentiation of roots. These genes have also been called Shi (shoot-inducing) and Roi (root-inducing) loci.*

that may maintain bacterial proliferation. As the transformed cells produce large amounts of hormones, they can be grown on media lacking hormones. In fact, one may distinguish *Agrobacteria*-caused tumours from explant calluses by the fact that the former can grow on a plain agar medium, whereas the latter require a hormone-supplemented growth medium. The wild-type alleles of *tms* and *tmr* are believed to ensure suppression of shoot and root formation in the tumour cells. Skoog's studies indicated that the development of a root or a shoot is a function of a ratio between the shoot-suppressing (but root-promoting) and root-suppressing (but shoot-promoting) phytohormones. A high cytokinin-to-auxin ratio favours shoot differentiation, while a high auxin-to-cytokinin ratio promotes root formation. If both are present in a balanced ratio, the tissue remains undifferentiated.

The wild-type alleles of *tms* and *tmr* have been renamed as genes 1 and 2 or Shi (shoot inhibition) and gene 4 or Roi (root inhibition) (Fig. 5.11). Genes 1 and 2 encode respectively, tryptophan-2-monooxygenase and indole-3-acetamide hydrolase that are used to convert L-tryptophan into AA. Gene 4 represents an enzyme, i.e., dimethylallyl pyrophosphate transferase, that catalyses one of the steps in the biosynthetic pathway for cytokinin.

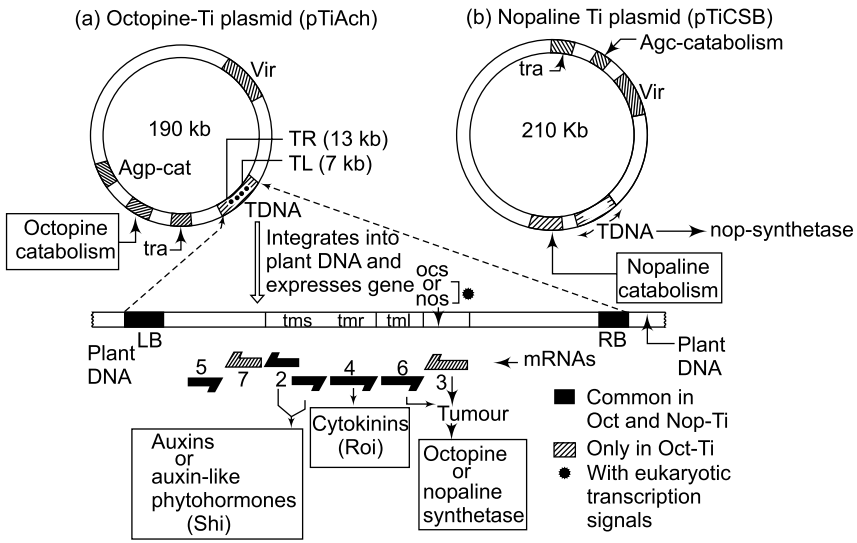


Fig. 5.11 *Schematic map of octopine and nopaline Ti plasmids, pTiAch and pTiC58 respectively. The non-T DNA regions of the plasmids carry the vir (virulence, required for infection), tra (transfer) genes and the gene for the appropriate opine catabolizing enzyme. Here, ocs and nos are the octopine and nopaline synthetase genes, respectively. LB, RB are the left and right borders of the T DNA. Agn and Agc refer to genes of other opines, agropine and agropinocine, respectively. In one plasmid the T DNA is separated into a left (TL) and a right (TR) segment.*

The T DNA of the nopaline Ti plasmid occurs as one uninterrupted block of sequences, whereas that in the octopine Ti plasmid occupies two separate sites, each flanked by the border repeat sequences, like all transposons. The left or TL block contains the Shi and Roi sequences.

The opines that find their way into the soil at the base of the infected plant trigger off the transfer of non-virulent *A. tumefaciens* (without Ti) into the host plant. The original Ti-containing bacterium thus sets up a situation that not only ensures a continuous supply of opines for its metabolic needs, but also inducts other *Agrobacteria* into the same job. An ecological niche, especially favourable to these bacteria, is perhaps engineered by this process.

(b) Vectors from Ti plasmids The T DNA may be employed as a vector for ferrying genes into plant cells. However, the large size, lack of unique cloning sites and the difficulty of regenerating normal plants from Ti-induced calluses make the use of natural Ti plasmids an impractical proposition. Vectors are, therefore, being continuously developed that make the most of the usable features of the Ti plasmids.

To avoid formation of tumours, Ti-based plasmids lack the *onc* (tumour-causing *tms* and *tmr*) regions (Fig. 5.12). They contain the LB and RB repeat sequences, and a selectable marker. The region between the LB and RB can house the gene to be transferred. The *vir* functions may be supplied *in trans* or *cis* with the gene to be transferred.

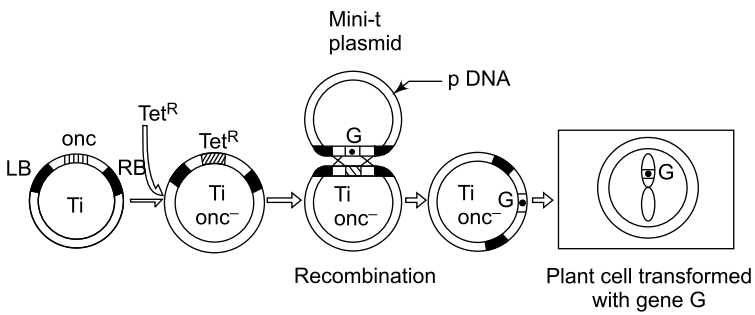


Fig. 5.12 *Introducing genes in plant cells via Ti plasmids: the homologous recombination method. A Ti plasmid vector may be constructed in which the onc gene of the T DNA has been replaced by a Tet^R gene. This vector is coinfected into plant cells with a mini-plasmid containing essentially only the terminal border sequences and the gene (G) to be transferred to the plant cell. Homologous recombination between T DNA sequences in the two plasmids results in an onc$^-$, G$^+$ Ti vector. The latter is used to transform plant cells*

There are basically two ways in which the exogenous gene may be integrated into the plant DNA via Ti-based vectors. In one strategy, the desired gene is introduced into a Ti-vector by recombination with a plasmid carrying the gene. Recombination is possible as in one plasmid the RB and LB flank the gene of interest cloned in pBR322, while a *onc*$^-$ Ti plasmid contains the same pBR322 sequences sandwiched between an LB and RB sequences. Homologous recombination between the pBR322 sequences transfers the cloned gene between the LB and RB in the Ti vector. The latter is then utilized to transform plant cells. The cloned gene becomes stably integrated into the plant cell DNA.

The above strategy was first employed by Ausubel's group at Harvard University. The *onc* gene between LB and RB was replaced by a Tet^R gene in one Ti plasmid and a donor DNA inserted within a T DNA that is spliced to a plasmid vector.

Both these vectors were introduced into the same plant cell. Reciprocal recombination between the homologus T DNA regions flanking the cloned gene and the Tet^R gene, transferred the former in place of the latter in the Ti vector.

The second strategy is based on the use of *E. coli-A. tumefaciens* shuttle vectors (Fig. 5.13). The vector carries the LB and RB regions and an intercalated one with several cloning sites, into one of which the gene of interest is inserted. This vector is introduced into an *A. tumefaciens* cell that possesses an onc^- vir^+ (deleted) Ti plasmid. The *vir* functions mobilize the transfer of the gene of interest between the LB and RB into the plant cell DNA.

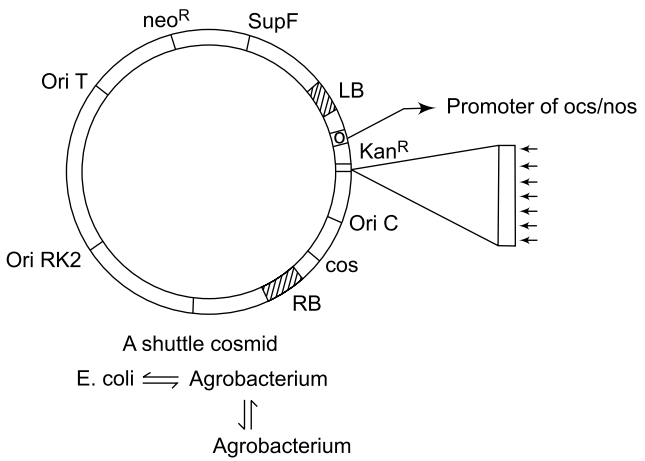


Fig. 5.13 *A Ti cosmid shuttle that can replicate in E. coli, A. tumefaciens and dicotyledonous plant cell (Ausubel), due to the presence of Ori C, Ori RK2 and Ori T, respectively. The promoter of the octopine synthetase (ocs) or nopaline synthetase (nos) gene is a eukaryotic one: a polylinker is provided as sites for cloning. The cos sequence allows packaging of the hybrid cosmid into phage lambda particles. The transposition specific LB and RB of the T DNA flanking the promoter and cloning sites allow integration of the foreign DNA into the plant cell DNA. The neo^R marker may be used to select kanamycin- or neomycin-resistant bacterial cells and G418-resistant plant cells.*

Ausubel's group has developed several cosmids that are better vectors than the earlier versions. These contain the following useful sequences: (i) the *oriC* of Col E1, that allows plasmid replication in *E. coli;* (ii) the *ori* of plasmid RK2 that allows vector replication in *Agrobacteria;* (iii) the *oriT* of the T DNA that is required for the start of the T DNA transfer; (iv) a Tet^R gene for selection in bacterial cells; (v) a kanamycin-resistant or neo^R (neomycin phosphotransferare II) gene as a selectable marker in plant cells; (vi) a Sup^F (tRNA suppressor) gene to be used in *E. coli;* (vii) a polylinker with unique cloning sites; (viii) a lambda *cos* region for insertion of plant DNA fragments cloned in cosmid banks; and (ix) the LB and RB sequences of T DNA flanking the kan^R gene, the polylinker, the Co1 E1 *ori* and the *cos* sequences.

An additional modification had to be introduced into the Ti-based cosmids before they were fully utilizable. The kan^R and methotrexate-resistant genes (which are selectable in plant cells) are not expressed in plant cells unless a promoter for one of the T DNA genes is prefixed to the drug-resistant gene. Promoters for octopine

synthetase (*ocs*) and nopaline synthetase (*nos*) genes are generally used for the purpose. The appropriate termination signals have also to be provided downstream of the drug-resistant gene. The revised vector transfers the desired gene to the plant DNA and also expresses the selectable marker in the transformed cells. Herrera-Estrella *et al.* (1983) constructed a vector with the promoter for a gene for the small subunit of Rubisco (Ribulose biphosphate carboxylase) prefixed to the CAT gene. When the promoter is induced by light, the CAT expression could be detected in transformed tobacco plant cells.

(c) Vectors from Ri plasmids Ri plasmids of *A. rhizogenes* are also classified on the basis of the opine synthetase gene that they carry. Agropine and mannopine Ri plasmids have been studied better than the others that have been reported.

The Ri plasmids also possess LB and RB (15–20 bp repeats each) sequences that flank approximately 15 kb of T DNA. Auxin synthesis is mediated by products of *tms* genes (1 and 2) as in A *tumefaciens.* Root differentiation specific genes (Ro1A and Ro1D) are present in the LB region and *tms* ones in the RB region.

Vectors have been constructed where the *vir* (*tra*) gene products are supplied in *trans* by a helper vector while the foreign gene is sandwiched between LB and RB regions of a hybrid vector. Hairy roots differentiated from Ri-transformed cells can be regenerated into plants.

These regenerated plants exhibit a variety of morphological and reproductive abnormalities that include crinkled leaves, dwarf plants, much branched stems and reduction in pollen fertility. Moreover, these abnormalities differ from one species to another, as well as at different stages of development in the same plant. The reasons underlying these anomalies are not known. Ri-based vectors may be used appropriately only when these aberrations are controlled; for that to happen more has to be known about the molecular biology of Ri plasmid T DNA-directed differentiations.

The Ri plasmids, however, provide fewer problems than Ti plasmids in terms of regeneration of transformed cells.

To summarize, the following is a short list of the strategies that have been reported, using agrobacterial plasmid-based vectors for plant-cell transformation:

- TDNA crippled at the *tms* or *tmr* sites.
- Insertion of hybird Ti plasmids directly into cells by precipitation with Ca^{++} and PEG.
- Direct insertion of hybrid Ti plasmids via liposomal vessels.
- Tentative use of Ri T DNA vectors.
- Somatic hybridization of a Ti-transformed cell and a normal cell followed by the regeneration of the whole plant.

(ii) Homologous DNA as vector Usually, ribosomal sequences occur in tandem repeats of several units. A passenger gene inserted in such a repeated sequence has a chance of being introduced into plant chromosomal DNA as a consequence of recombination between the resident and the imported rDNA sequences.

Such a strategy has been attempted using the extrachromosomal plasmid-like rDNA present in the lower eukaryote, *Physarum polycephalum.* Potrykus and

associates (Friedrich Miescher Institute, Switzerland) extracted the circular rDNA from the nucleolus region and used them as vectors. The efficiency of transformation was at least 20 per cent when the recipient cells were yeast. Tobacco plant cells showed a greater success, and the transformed cells regenerated into fertile plants.

(iii) Transposons as vectors Transposons were first identified in *Zea mays* (maize) by Barbara McClintock. Further characterization was possible only after the development of molecular biology and recombinant DNA techniques. The maize transpons that are known best are the AC, Mu and Spm elements. Others have been discovered in other species (e.g., Antirrhinums and soyabeans).

Transposons have been useful to date for gene tagging. Their sites of integration appear to be random. Until the latter are regulated, transposon-based vectors will not become a routine reality. Insertion within an important gene may cause mutations that may jeopardize the goal of the gene transfer project.

Use of transposons to transfer genes has been made for one insect, namely *Bacillus thuringensis,* which produces an insecticide that is lethal for certain insect pests of maize plants. The gene for the biosynthesis of the insecticide was introduced via a transposon into *Pseudomonas fluorescence* that colonise the root system of maize plants. Perhaps these bacteria would transfer the insecticide gene into the plant DNA. However, before such methods can be put into practice, much has to be learnt about the effect of the insecticide within the plant. Also, will they affect the consumers of maize grains?

III. Agroinfection: use of mixed vectors One clever strategy to expand the host range of vectors is to use the technique of agroinfection, which consists of constructing hybrid vectors that possess the infectivity due to viral sequences and other properties due to other vector sources.

Both CaMV and *Agrobacteria* sp. have rather limited host ranges. However, a judicious combination of useful parts of the viral and Ti-encoded genes can generate vectors that find entry into species impenetrable alone by the two parental species. The first of such vectors carried the foreign DNA in CaMV sequences that were flanked by LB and RB sequences of T DNA. The vector has to be coinfected with a helper virus that will compensate for the loss of function of the crippled recombinant vector.

Shewmaker and associates reported integration of a CaMV DNA into plants that are not its natural host by agroinfection. However, the rates of expression of the 35S and 19S transcripts varied greatly. These experiments repeatedly highlight our present inadequate knowledge about molecular mechanisms in plant systems. If one could find out why the CaMV are transcribed correctly only in crucifers but not in other dicot species, we would be slightly better informed about the strategies of gene expression and regulation in both plants and viruses.

5.4.2 Expression of Foreign Genes in Plants

Plant pathogens possess only a limited number of genes; they depend on the infrastructure of the host for many of their functions. They also possess replication and transcriptional signals especially geared for use with enzymes encoded by their

genes. In addition, a viral infection (or a bacterial one) is amplified to 105 to 107 pathogens per cell. These features make vectors based on the DNA or RNA of these pathogens potential candidates for both carrying foreign genes into plant cells and facilitating their expression in a heterologous system.

A few assays have become available that may be used to assess the expression of the transferred gene in the plant cell background. These include assays for the opine synthetase activities, assays for neomycin phosphotransferase (npt II) DHFR assay and CAT assay.

Octopine and nopaline synthetase assays have been developed by Schilperoot and associates (Otten and Schilperoot, 1978). For the nopaline synthetase assay, accumulated nopaline is first removed from the plant extract by passing it through a Sephadex G100 column. The presence of the enzymes (NOS and OCS) is determined by running spots of the reaction mixture (plant extract + arginine + NADH + pyruvate for OCS and the same, but betaglutarate in place of pyruvate for NOS by paper chromatography. The paper after staining with phenanthroquinone (very toxic) is seen under UV light. The npt II assay is based on the ability to phosphorylate kanamycin. The test mixture is fractionated by electrophoresis on an agarose gel containing labelled ATP and kanamycin. The label will be transferred to kanamycin if the enzyme npt II is present in the test mixture.

The reaction products are transferred to P81 phosphocellulose paper (ion exchange) and the latter autoradiographed. Radiolabeled kanamycin will show up in the autoradiogram, confirming the presence of even very small amounts of npt II.

The DHFR assay monitors the incorporation of labelled phosphate into DNA in the presence of methotrexate. The test plant tissue is incubated with labelled orthophosphate, and the DNA extracted, fractionated on an agarose (1 per cent) gel and autoradiographed. Labelling of the DNA with ^{32}P confirms the presence of *dhfr* in the system. Brisson *et al.* (1984) used this assay to check the expression of a *dhfr* gene that they had transferred to turnip plants using a CaMV vector.

The CAT assay is based on the acetylation of chloramphenicol derivatives by the chloramphenicol acetyl transferase (CAT) gene. The test plant extract is mixed with ^{14}C-labelled chloramphenicol, and acetyl-Co-A in an appropriate reaction mixture. The mixture is extracted with ethyl acetate to collect all chloramphenicol and its products. This fraction is fractionated by ascending paper chromatography and the paper autoradiographed. The amount of chloramphenicol converted to its monoacetate form (labelled) is estimated.

5.5 ❑ RESEARCH

All genetic material is not composed of genes or plants for synthesizing polypeptides. The genic regions are interspersed among those that do not code for proteins but contain different sequence stretches that are concerned with the regulation of gene expression at different hierarchical levels. Recombinant DNA technology has to be credited for the identification, isolation and characterization of many of these sequences, and for the subsequent increase in understanding of the biology of that piece of DNA.

Recombinant DNA technology allows the most precise way for the molecular dissection of the genetic material. Combined with accessory techniques such as Southern blotting, DNA sequencing and immunodiagnostic methodologies, these techniques of molecular biology have been the sole tools for analysing and characterizing the genetic material and its products in the past two decades.

It would be impossible and also pointless to attempt to list every area of research that has been or can be benefited by the use of recombinant DNA technology. Only a few domains of investigation will be briefly recounted here to emphasize the power of this technology for unravelling the fundamental mechanisms that underlie the living state.

5.5.1 The Biology of DNA

Delineating the physical map of a DNA, pinpointing operational subsets within it and discovering the anatomy of genes have been possible only with the use of recombinant DNA techniques.

1. Physical Mapping

Actual physical addresses of regions of a gene can be pinpointed only by using the recombinant DNA techniques. Restriction enzymes can be used to map gross regions in the whole DNA, and this can be followed up by sequencing the restriction fragments or cloned subsets of the same.

2. Role of a DNA Region

The functional role of any nucleotide in a sequence, or of a particular sequence, can be discovered by site-directed mutagenesis of the concerned region, followed by a search for deviations in function(s) or loss of the same in the expression vehicle. The functional roles of the genes in T DNA of Ti and Ri plasmids of *Agrobacteria* and those of the nitrogen fixation (*nif*) genes of nitrogen-fixing organisms (*Klebsiella, Rhizhobia, Azolla* etc.) have been discovered with the aid of site-directed mutagenesis.

3. Protein-Binding Sequences

The DNA sequences to which proteins bind themselves in the course of different activities are identified and isolated by using the strategy known as 'DNA footprinting' (Fig. 5.14). Promoters, enhancers, terminators, attenuators and various upstream and downstream regulatory sequences have been characterized by this and other methods of molecular biology.

4. Inserted DNA Sequences

The location and types of repeated sequences, transposons and other transposable elements, and integrated viral genomes have been discovered by gene cloning methods combined with Southern blotting and autoradiography.

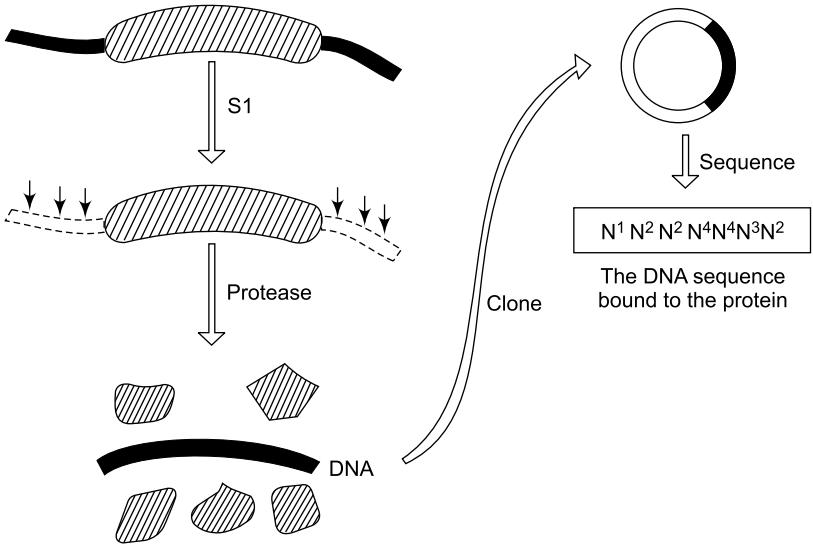


Fig. 5.14 *DNA footprinting. The sequences of a DNA that are bound to a particular protein (such as the RNA polymerase) may be determined by degrading the DNA not bound to the protein with the S1 nuclease. A proteolytic enzyme is used to degrade the bound protein. The released DNA fragment is cloned and sequenced.*

5. Gene Banks

The entire DNA of a species can be stored in the form of banks of cloned fragments. Each fragment is spliced to a vector, and the hybrid vector amplified. Search for specific regions of the entire genome can be made by examining the gene bank by appropriate means.

6. Sequencing

Constructing recombinant DNA is prerequisite for sequencing protocols. Improved modifications of sequencing strategies have been based on skilful and ingenious use of cloned fragments of DNA. The particular fragment to be sequenced or the total DNA of a cell can be sequenced by using the appropriate technology. Sequencing has revealed the presence of unsuspected genes, the products of which are too transient to have been recognized and captured by conventional biochemical methods. Several hormones and neuropeptides have been identified by isolating the pre-protein unprocessed mRNA. This heterogeneous (hn) **RNA** or the initial transcript from a DNA region was found to contain the coding sequences for several peptide products, the relationships between which were not clear from earlier studies, although they were known to possess certain identical regions. A look at the initial transcript or the sequenced complementary DNA has also revealed the presence of other proteins, some of which have not yet been isolated. The initial transcript is translated into separate proteins, which in some cases are further split into smaller peptides by post-translational proteolysis.

Sequencing has also elucidated the organization of split, mobile, overlapping and other types of genes.

7. Types of Genes

Recombinant DNA studies have revised earlier notions about the way in which a gene is organized on a DNA, *vis*-à-*vis* other genes. There are split, overlapping, mobile, variable and processed genes. Recently, there is a report of a protein that is encoded in two parts, one on one strand of DNA, and the other on the complementary strand, the final product translated from the transcript spliced from the two. Recombinant DNA together with Southern blotting and autoradiography are routinely used for locating genes, or portions of DNA *in situ* in a cell or used to search for the same in different tissues and/or different species.

8. Making Useful Vectors

A part of the research in molecular biology is engaged in constructing useful vectors, tailored for specific jobs. As the complexities of plant and animal biology are becoming apparent, the need for vehicles suitable for successful gene transfer-cum-expression is becoming imperative.

5.5.2 Development Studies

Regulation of gene expression underlines the phenomenon of development. There has been a consistent search for checkposts in the path of development in terms of (i) agents that instigate determination or differentiation, (ii) genes that become programmed to create stem cells, and (iii) mechanisms that are utilized as biological switches for the concerted management of genes to be turned on or off. It is possible now to address incisive questions framed in terms of recombinant DNA that may reveal answers for some or all of these issues.

Earlier development studies relied on observations and genetical manipulations of spontaneous or artificially induced genetic mosaics, homeotic mutants and on ingenious surgical manipulations of developing embryos. Molecular biology methods have already unravelled much of the fine-tuned control of sequential gene expressions in a host of phages and viruses. The power of rDNA techniques for the understanding of development is illustrated by the following selected areas of research.

1. Homeotic Genes and Development

Following the developmental programme at the molecular level is much more difficult in higher organisms than in viruses and prokaryotes. It is customary, therefore, to choose systems that allow one to take a closer look at least at the transition point from one developmental phase to another. One such system is represented by the developmental aberrations found generally in dipteran flies and known as 'homeotic mutations'. In a homeotic mutant an atypical appendage of the body is found in place of the normal one; an antenna is replaced by a leg (*antennapedia*), the third part of the thorax is formed by a duplication of the second segment of the thorax

(*bithorax*) and so on. The genes, which, on mutation, cause these developmental defects are called 'homeotic genes'. They are the first key developmental genes that have been identified as those which determine the pattern or architectural plan of the body. Although these mutations had been observed and described in the second decade of the twentieth century, the mechanism by which they possibly operate became amenable to analysis only after the development of rDNA techniques. Analysis of the DNA sequence of the *antennapedia* homeotic gene disclosed the presence of a stretch of DNA about 180 bp long. Using this as a probe, similar sequences (called 'homeo box') were traced in other homeotic genes.

It was tempting to know whether the conserved homeo box sequences were merely curiosities found only in *Drosophila* type insects or were routine components of the DNA of other animal species. It was exciting to discover that homeo boxes occur in vertebrates and other animal species. The next obvious step was to use probes made from known homeo boxes to search for evidences of expression of developmental genes in more complex species, including human beings. The homeo box appears to encode the information for a particular domain in a homeotic protein. Some idea about its role in the protein has come from the discovery that another kind of developmental gene (found in yeasts) that is concerned in deciding the differentiation of two phenotypic patterns (mating types *a* and α), also contains a domain that is closely homologous to the homeo box. The *MAT a*1 and *MAT* α2 yeast gene products are DNA-binding proteins that bind themselves to control sequences upstream of the mating type genes. These studies have elaborated the actual sequences that are involved in binding to the DNA and have shown in detail the similarities and dissimilarities between the sequences of a wide spectrum of species from primitive metazoa to human being. These findings suggest that homeotic genes play a crucial role in regulating development in most organisms, including prokaryotic species. It is sufficient to mention here that rDNA techniques were responsible for unearthing a group of genes, the products of which play a role in the control of development, especially of the form of the body. Armed with cloned homeotic genes, the search has now proliferated into finding their role in mammalian development. There is a dearth of homeotic mutations in higher organisms; rather they have not been recognized as yet, even though their existence has been noticed. In mouse there is a defect (the mutation *rachiterata*) in which a first vertebra of the thorax is found to occupy the position of the seventh one of the cervical region. Perhaps this is a candidate homeotic mutation. E B Lewis, one of the pioneers in the molecular study of homeotic mutations, suggests a combinatorial model for the development of body pattern (segmentation) in which a combination of a set of homeotic genes in a particular segment regulates the unique differentiation of the same. Concerted expression of sets of such developmental genes is, therefore, suspected to be the basis for the chronological and spatial development of the body of an organism.

2. Control of Temporal and Tissue-specific Gene Expression

That there are genes which take part in the regulation of expression of other genes (or gene) has been known for some time. Mutations in such regulatory genes do not alter the structural or functional traits of the regulated gene, but alter its place

(tissues), time (in development) and/or rate of expression. The only way to study these genes had been to look for (or introduce) mutations in the regulatory genes or sequences and assess the deviation in the expression of the controlled gene. The ability to clone specific regions of the DNA has revealed much information about the proteins that get attached to the regulatory sequences and induce their expressions spatially, and temporally. Promoters, operators, enhancers, attenuators, termination and splicing signals and many other regulation sequences have been revealed by DNA cloning.

Regulation of gene expressions is achieved by the binding of regulatory proteins and/or transcription factors (also proteins) to the regulatory sequences.

Details about the nature of these factors and their mode of action are becoming known from studies that are based on the use of cloned DNA.

3. Regulation of Development by Anti-sense RNA

Recombinant DNA techniques have been instrumental in the discovery of a still another unique device for the fine-tuning of regulation of development. This device uses what has become known as anti-sense RNA. Anti-sense RNA is usually transcribed from the strand region complementary to the sense-strand of a gene. The mRNA and anti-sense RNA are, therefore, complementary to each other, but in the opposite direction [Fig. 5.15(a)]. Base-pairing of the two RNA species [often only a part of the region, as shown in Fig. 5.15(b)] blocks further transcription or translation of the mRNA.

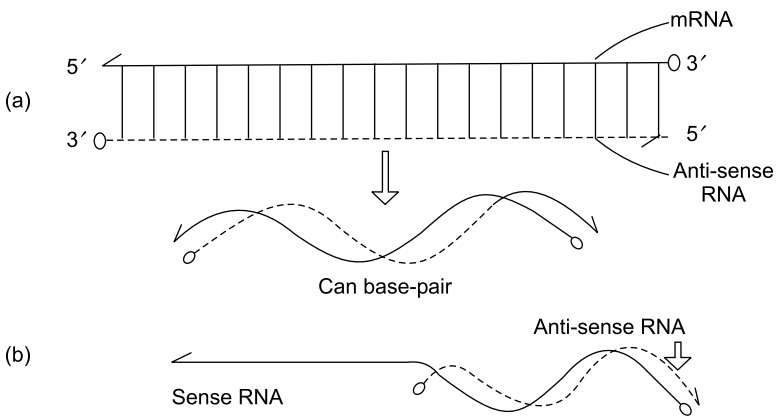


Fig. 5.15 *Anti-sense RNA. (a) The RNAs complementary to the two strands of a ds DNA are complementary to each other. (b) Base pairing of a region of the sense RNA with its anti-sense RNA. This device is used in nature and in applications to control development.*

Expression of the particular gene does not culminate in the polypeptide product, when translation is thwarted. This is a subtle mechanism for modulating diverse developmental pathways. One of the first and best studied example of developmental control via anti-sense RNA is the control of transposition of Tn10. Tn10 is a mobile unit of DNA that possesses short IS 10 (Insertion element) at the two ends. The transposon carries a gene for the enzyme transposase which is essential for the

transposition activity. The promoter for the transposase is within the right IS 10. The mRNA initiated from this promoter has been named as RNA-IN, and its corresponding promoter as P-IN. A promoter (P-OUT) in the anti-sense strand of the DNA initiates an RNA, known as RNA-OUT. The RNA-OUT base-pairs for a length of 35 bases with RNA-IN. This duplex-formation stops translation of the mRNA into the enzyme. As a consequence of an increase in the copies of Tn10, therefore, the level of transposition decreases.

Anti-sense RNA, and subsequently anti-sense DNA are being exploited for a bizarre variety of investigations and applications. To mention one, tomatoes have been made transportable, without any damage due to overripening, by preventing the translation of the gene for the enzyme required for the softening process. The engineered component is temperature sensitive, so that the fruits can be induced to ripen on demand, by altering the storage temperature from the unsuitable one during transportation to the one needed for ripening.

4. Transfection into Nuclei

Another way of ascertaining how development is managed is to introduce DNA sequences of choice into the nucleus of an undifferentiated cell, such as the fertilized egg cell of a frog or mouse. Methods of microinjecting genes into foreign nuclei have been perfected for both large cells such as those mentioned earlier and for normal animal cells that are much smaller. Again, the injected gene is sent in via a cloned vector.

5. Cell-Cell Recognition and Interaction

Development in animals is a function also of both cell-cell recognition and cell migration to appropriate regions in the growing embryo. Almost all investigations that have led to the discovery of cell adhesion molecules (CAM) and substrate adhesion molecules (SAM) and to the receptors to which they bind themselves for aggregation and migration, respectively, have been based on the use of cloned hybrid DNA.

It is difficult, in most cases, to study the metabolisms of a gene in its natural setting, since the products of the gene may become eclipsed by those of several others in the same milieu. This obstacle can be avoided by studying a copy of the gene in a different setting. The copy may be obtained by synthesizing a DNA from the mRNA of the gene to be cloned.

6. Haematopoietic Cell Differentiation

Another popular system for the molecular dissection of development is the differentiation of haematopoietic cells (Fig. 5.16). Stem cells in bone marrow become differentiated into the different blood cells, including the red blood cells, lymphocytes, macrophages, platelets and so on. The factors involved in differentiating and proliferating different types of cells from presumably one or a few kinds of stem cells are searched best by recombinant DNA techniques.

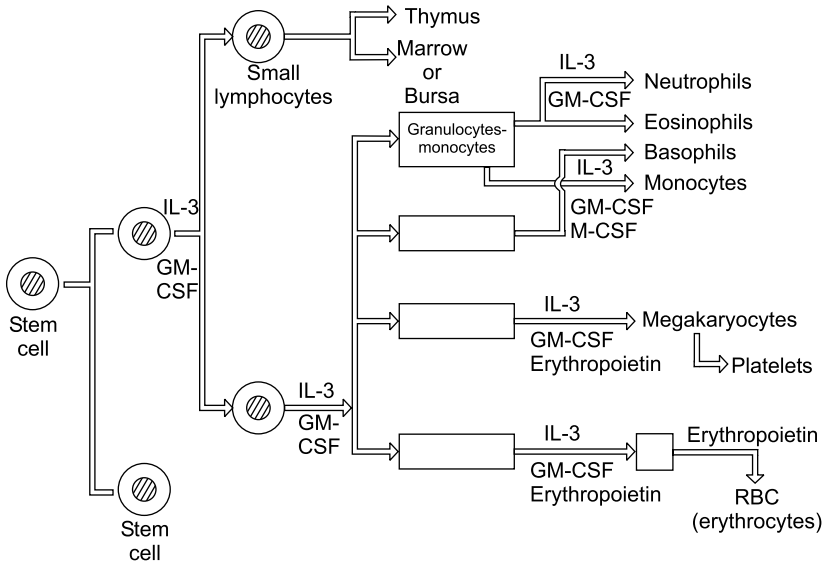


Fig. 5.16 *Hematopoietic cell differentiation. Erythrocytes and leucocytes originate from stem cells in the bone marrow (in foetal liver in embryos). They differentiate into the various specific cell types of the blood tissue system when triggered off by specific stimulating factors.*

Studies in the last two decades have revealed that the bone marrow stem cell differentiates into different types under the influence of highly specific regulator proteins. These protein factors have been named, in general, as the colony-stimulating factors (CSFs). Several CSFs have been identified in both mouse and human beings. Particularly important are four CSFs: the G-CSF (granulocyte-CSF), the M-CSF (macrophage CSF), the GM-CSF (granulocyte-macrophage CSF) and interleukin-3 (IL-3), a multiCSF. These haematopoietic factors are synthesized in too minute a quantity for extraction in a pure form by conventional biochemical techniques. Gene-cloning methods, however, have identified the genes for the above four growth factors, and cloned them. Expression of the cloned genes is now the source of sufficient quantities of these proteins for investigative studies.

In the case of the M-CSF and G-CSF, the proteins were purified and their amino acid sequences determined. The likely DNA sequences for the polypeptide chains were predicted from these data. Oligonucleotide probes were synthesized chemically corresponding to several regions of the DNA sequences. These probes were then utilized to locate the sequences of the genes for M- and G-CSFs. In the case of the other two CSFs (GM-CSF and IL-3) the genes were cloned and tested for the correct expression in COS-cells. The IL-3 and GM-CSF genes occur very close together in the same stained band in chromosome 5 (long arm), a region at which the genes of many other growth-stimulating factors are clustered. These genes include those for M-CSF and its cell surface receptor, the protooncogene *c-fms.* Several myeloid disorders can be traced to deletions in this 5qr band. Cloned genes of CSFs are being used fruitfully to decipher the mechanism of regulation of differentiation into the different types of blood-tissue cells. Probes made from the genes have been used to

locate the expression of the CSF genes in the progenitors of these different clones of cells. The biological activities of the hybrid CSF DNAs were equivalent to those of naturally occurring haematoproteins.

Cloned CSF genes have not only provided a means of obtaining large amounts of pure protein products, that can be used for research, but also have shown promise for use in clinical applications.

7. Architectural Plan of the Body

Pattern formation, another feature of development, is scrutinized at present in dipteran insects (e.g., *Drosophila*) that develop segmentally into symmetrical regions. Can the information gained from these studies be extrapolated to explain the development of form and pattern in higher animals, such as mammals? Perhaps it can be, as suggested by certain reports about abnormal patterning of parts in certain developmental mutants of mouse.

The *Drosophila* develops from an elongated egg cell (fertilized) into a mobile embryo (larva) that grows in size, acquires the internal anatomy of the adult fly, but retains certain packets of cells (imaginal discs) in a semi-differentiated status. These cells, on stimulation by a steroid hormone (ecdysone) complete their differentiation and form the adult appendages, such as the legs, antennae, eyes, halters, wings and genitals. The body of this insect develops symmetrically; it is composed of several segments, each of which carries specific adult appendages. Even mutants possess symmetrically placed defects. Geneticists have long grappled with the available developmental mutants to find out by traditional methods what went wrong and conversely to determine the normal course of events. It was only with gene-cloning methods, however, that facts became known about the possible manner in which the development of the body pattern is regulated at the molecular level.

By using cloned genes (normal and mutant) and reverse genetics—as well as all the weapons in the artillery of the molecular biologist—it has become possible to trace the path of development, at least in *Drosophila,* in terms of the sequential expression(s) of a series of developmental genes.

The initial determinants in the fertilized egg are products of maternal (egg cell) genes that are distributed in specific gradients in the egg cell cytoplasm. After fertilization, and several initial bouts of nuclear division, the nuclei move to the periphery of the egg cell (now the zygote), and carve out a little of the cytoplasm to form a layer of cells. The gradients (*combinatorially*) of the *maternal effect* gene products in the vicinity of a particular nucleus decide which of the gene(s) in the diploid nucleus is (are) to be expressed. There is thus a subtle distinction between the cells that line the outer boundary of the larva. The process continues, with products expressed by the first set of zygotic genes acting (singly and in combination) as stimuli for expression of the next set and so on, until the adult insect is constructed.

The *maternal effect* genes are followed by the expression of *segmentation genes* which were discovered to be of three types: the *gap, pair-rule* and *polarity* genes. Reporter genes spectacularly showed up the segmentation process as zebra-type stripes. The genes have been individually cloned and their products characterized. Some of them turned out to be transcription factors (*fushi tarazu* or *ftz, engrailed* or *eng, paired, evenskipped*, etc.).

One of the genes, dissected by rDNA methods, revealed a coding sequence preceded by a very long 5′-flanking upstream region containing several regulatory sequences. One of the latter was found to be activated in neural tissues, leading to neural differentiation. Another sequence regulates the striped or zebra patterning of the embryo into segments.

The beauty and utility of gene-cloning techniques lie in the possibility of framing unusual experimental questions. For instance, to obtain an answer to the question 'what does a gene do in a particular case?', one method would be to inactivate the gene and look at the consequences. Another would be to introduce short stretches of DNA at putative functional regions in the gene and assess the outcome. The striping pattern dictated by the *ftz* gene regulator was revealed by one such study, where specific DNA pieces were introduced into the germ line cells with the help of a vector derived from a naturally occurring plasmid (P-element) in *Drosophila.* The story, however, is not as simplistic as it appears from this brief report of the action of the segmental *ftz* gene. The proper expression of the gene has been found to rely also on that of at least one other gene (*hairy*). Mutants of this gene (*hairy*) possess broader and often fused stripes, indicating that in normal embryo development, the *hairy* gene represses the action of the *ftz* one in the interstripe regions.

Is the regulation of tissue-specific developmental genes, due to differences in rates of expression of the genes, different in different cellular environments? It appears to be so, at least in the case of the *Antennapedia* gene in *Drosophila.* When a hybrid DNA containing a cDNA copy of the *Antp* gene and the promoter of the gene of the heat shock protein 70 was introduced into germ line cells of wild-type flies, cells which were transformed by the hybrid or fused gene, gave rise to larvae that on heat shock treatment developed into *Antennapedia* homeotic mutants. The inference is that the hsp 70 promoter induced the expression the *Antp* gene out of turn in the larvae, and this excess of the *Antp* protein at the wrong time caused the development of a leg in place of the antenna. Tampering or interfering with the normal pattern of expression of developmental genes in still other ways is expected to provide answers about the yet-unknown mechanism(s) underlying patterning during development.

These are but a few areas in development that are engaging the attention of molecular biologists. Frankly, it is difficult to categorize the type of work being carried out using recombinant DNA technology, since the answers obtained from any one type of research are directly relevant to almost all other areas of biology. A study of regulatory sequences exposes the activities of proteins and factors, that in turn control the rate and spatial and temporal expression of genes that may be key developmental switches or ones that become active under special provocation. The result of such a study belongs as much to the domain of deciphering the plan in the genome as to that of development. Further, abnormal development is only the other side of the coin. Disorders and diseases are but expressions of abnormal genes in right place and time, abnormal expression of the correct gene in the proper tissues and developmental period, or the out-of-turn expression of normal genes in other times and places in the body. All of these aberrations lead to a lack of ease or disease of the body.

5.5.3 Immunogenetics

Recombinant DNA techniques have been responsible for the rapid unfolding of the molecular mechanisms underlying the phenomenon of immunity in animals. These studies not only have revealed the intricacies of the immune response but also have disclosed the basis for cell-cell recognition that is an essential strategy in development. Commercially useful spin-offs include the development of techniques of synthesizing various immunodiagnostic molecules, which have become the mainstay for reaching specific macromolecules or for identifying them. The knowledge of immunogentics has recently been successfully utilized to synthesize designer proteins and catalysts.

One of the enigmas of biology concerns the basis for antibody diversity. Several hypotheses or speculations have been presented earlier, but only molecular dissections by current techniques revealed that antibody genes possess several alternative sequences for the region of the immunoglobulin that binds itself to the antigen. It was possible to piece together the story as it emerged in several laboratories—prominently in those of Susumu Tonegawa and Leroy Hood—into our current understanding about the structure and expression of antibody genes.

Antibodies are the main weapons for attacking non-cellular antigens. Cells and microbial parasites are engaged by another subset of the immune system dominated by T cells and macrophages. Recombinant DNA techniques revealed that recognition and attachment of T cells and macrophages to foreign bodies were due to characteristic surface molecules. Such molecules were discovered not only on T, B and other immune system cells, but also on every kind of cell in the body. Indeed, it was soon realized that all interactions between cells and other cells, tissues or substances occurred via cell surface proteins that match with each other.

The entire Pandora's box of genes that specify surface proteins was opened by gene-cloning techniques. The histocompatibility genes, their location on the chromosomal maps, their different classes of protein products, and the significance of the latter in discriminating self from 'non-self' were exposed by rDNA techniques.

Immunity is a function also of a variety of mediator substances. Genetic manipulation disclosed the nature of many of the mediators of immune response and clarified ambiguities about some of them that originated before the immune system was studied by contemporary strategies. The most prominent ones are interferons, interleukins, a variety of growth stimulating factors, and proteins that control movements of immune system cells to and away from sites of action. Many of the genes for these mediators have been cloned and their products collected from transformed cells. Interferon, which is produced in too little a quantity to be sufficient for any research, is now available from genetically manipulated cells in an unlimited quantity.

5.5.4 Cancer

Certain cells become genetically altered in a particular manner, after a viral infection or as a consequence of treatment with a mutagenic agent. These transformed cells display a close resemblance to spontaneously occurring cancer cells. Transformed cells

(induced in the laboratory), on introduction into animal hosts, cause tumour formation in the latter. Transformation of cells is, therefore, studied to obtain insights into what happens during oncogenesis or initiation of cancer (Fig. 5.17).

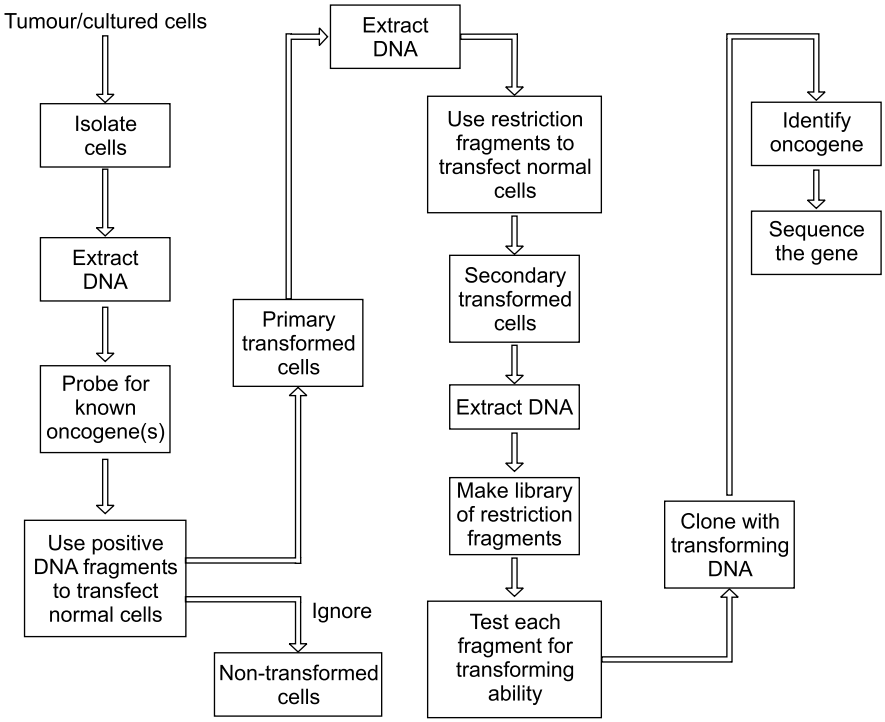


Fig. 5.17 *Isolation of an oncogene (see text for details).*

Cancer and transformed cells share certain traits that are different from those of normal cells. It had been customary earlier to investigate the basis of cancer from the starting premise of these phenotypic differences. Studies in the last decade have revealed that a cancerous transformation is based on alteration(s) in the genetic quality or/and programming of certain genes called oncogenes. Recombinant DNA technology has had a major role to play in the molecular dissection of oncogenesis.

Some viruses, particularly the retroviruses (ss RNA), were found to carry an extra gene that transformed cells on integration of the viral genome (cDNA) into the recipient cell DNA. Using probes made from oncogenes, identical sequences were detected in every cell of a normal individual. It was realized from further studies that oncogenes originated in the eukaryotic DNA and were picked up by viruses sometimes in the past, and that they play normal developmental roles in a normal individual. Why then does not every individual suffer from cancer? Gene-cloning techniques revealed that resident protooncogenes may become activated to behave abnormally if provoked in any one of several ways. These include integration of a strong viral promoter in its upstream regulatory region, and mutation of the gene sequences. Weinberg and Barbacid have shown by their fabulously meticulous

rDNA work that certain cancers are triggered off by as much as one base change in a particular codon of a protooncogene.

Recent excitement in cancer research concerns the discovery of anti-cancer or suppressor genes that control normal gene expression. Mutations in these controllers (including loss due to deletion) precipitate cancer by removing the control; this leads to overproduction of the protooncogene product. This was first discovered in cases of retinoblastoma and later in cases of lung cancers.

All the foregoing factors and much more was inconceivable even a few years ago. It may not be too long before even this dreaded malady is brought under control. As this book goes to the press, there is a report from John Hopkins University, Baltimore, that one of the products of one oncogene (p53) is able to inhibit proliferation of colon cancer cells, in culture. This report awaits confirmation.

A general method for isolating oncogene sequences consists of the following operations:

1. Isolate tumour cells from
   (i) an animal, or
   (ii) cell-line culture.
2. Extract high molecular weight DNA.
   (Test the DNA for known genes by hybridizing Southern blots).
3. Transfect normal cells in culture.
4. Pick out colonies of transformed cells (using morphological criteria).
5. Isolate DNA from these cells.
6. Make
   (i) a library from this DNA, and
   (ii) a restriction fragment profile.
7. Use each restriction fragment to identify cells that become transformed.
8. Extract DNA from transformed cells.
9. Identify oncogenic sequences (using a probe if available, and by sequencing the DNA).

The oncogenes for which viral and cellular counterparts have been identified include *src, myc, sis, nos, fos, ras, myb, abl, erb, fes, ets, mil* and *fms.* The animal hosts in which the c-*onc* have been discovered are mostly avian (fowls), murine (mouse), feline (cat) and rat.

Recombinant DNA studies revealed the differences between a v-*onc* and its counter part c-*onc* (e.g., v-*src* and c-*src*). The viral oncogene has no introns, while the cellular one possesses exons and introns.

The use of appropriate probes has shown that cellular oncogenes are very well conserved in widely divergent species (e.g., yeast, *Drosophila,* mouse and human beings) suggesting that, they play essential roles in cellular or developmental events. The products of these genes are mediators of events from the transduction of signals to the preparation of the cell for mitosis. Realization of this fact has led to a molecular analysis of events triggered off by a variety of mutagenic and carcinogenic agents (e.g., lectins, growth hormones, phorbol esters).

5.5.5 Inherited Disorders and their Timely Diagnosis

Perhaps the most sensitive method for discovering fine differences between two otherwise homologous DNA is that which uses recombinant DNA. These techniques have not only increased our understanding of diseases due to anomalies in the DNA but have also provided a means for precisely identifying such defects in both adult and prenatal stages of development.

There are more than 2500–3000 inherited disorders, of which the genetic basis is known in a greater or lesser degree for the majority. Prenatal diagnosis is helpful for those diseases that can be managed by palliative management measures (e.g., haemophilia by the administration of the clotting factor and phenylketonuria by providing a special diet to the newborn). Such diagnosis is also useful for taking decisions for performance of therapeutic abortions.

DNA techniques have gained importance for detecting inherited or heritable genetic diseases as well as those acquired due to viral, oncogenic or mutagen/carcinogen induced transformations.

1. Diseases Due to Defective Globins

The first disease that was tackled by DNA techniques was sickle-cell anaemia, where a single base substitution makes all the difference between the oxygen-adsorbing abilities and other properties of normal and diseased (sickle-cell) haemoglobin.

Several other globin-based diseases—collectively called *haemoglobinopathies*—are now investigated by using DNA techniques. Haemoglobinopathies are based on three types of defects, which are as follows:

1. Mutation in the structural gene of one or both types of globins of a haemoglobin molecule, e.g., HbS.
2. Mutation in regulatory genes that interferes with the synthesis of the globin chains. These defects cause disorders known as *thalassaemia,* in which one or the other type of globin chain is lacking or inadequately synthesized. This results in precipitated inclusions in the RBC that lead to severe *anaemias.*
3. Mutations in developmental control genes, resulting in the out-of-turn expression of globin genes at wrong periods of the developmental time table, (e.g., persistence of foetal γ-globin chains in adult haemoglobin).

Haemoglobinopathies are diagnosed traditionally by prenatal and antenatal analysis of blood. This analysis does not always detect defects that are not expressed in cells used for the analysis. Haemoglobin consists of four polypeptides (a pair each of two types of polypeptide chains). The chains are of the following types: ξ, ε, γ, α, β and δ (a variant of β). Different combinations of these chains (pairs) are expressed at different stages of development in human beings:

foetuses < 3 mths : 2 ξ and 2 ε chains
> 3 mths : 2 α and 2 γ chains
infants of 6 mths to adults : 2 α and 2 β chains

The foregoing defects may be detected prenatally by examining the DNA for mutations in the structural genes for these chains, using suitable probes. More than 300 mutants have thus been identified in α, β, γ and δ-globin chain genes.

Defects in the non-gene regions on the DNA, which control the expression of the gene in question, can be shown up as polymorphisms in the DNA, especially with respect to cutting/recognition sites for restriction enzymes.

The first globin-based disease that has been investigated by rDNA methods is sickle-cell anaemia. The sickle cell β-globin gene differs from the normal β-globin gene by one base substitution:

Normal or Hb^A gene: — CCTG [A] GGAG _____
Sickle or Hb^s gene: — CCTG [T] GGAG _____

2. Restriction Fragment Length Polymorphism (RFLP)

The exact genetic lesions for all inherited diseases are not known. However, the loci for genes involved in such diseases have been mapped in most cases. Analyses of the DNA in the neighbourhood of the disease gene in members of a family, and their immediate filial relatives, have disclosed an association of many of the mutant loci with polymorphism in the flanking DNAs. The polymorphic variation in the non-gene sequences associated with a gene may alter the normal pattern of restriction enzyme recognition and/or cutting site(s).

DNA probes are utilized to identify genetic markers in the neighbourhood of a gene involved in a disease. Here a marker is a sequence of nucleotides that accompanies the disease gene. The latter may be attended with several such markers. If a marker is found consistently with a gene, detection of the marker with a probe will not only tell us about the presence of the gene but also pinpoint the gene in a specific DNA region.

The foregoing method is particularly useful for the diagnosis in persons predisposed towards diseases that normally may surface at a later stage of development. The normal gene or DNA upstream of the gene has fixed restriction enzyme-cutting sites. So, if a normal gene DNA and its flanking DNA are cut with a particular RE and the fragments fractionated by gel electrophoresis, certain standard fragments will be obtained. If, however, there is polymorphism in the flanking area with respect to the RE-cutting site(s), an altered fragment profile will result after treatment with the same RE.

It has been found that there is a 1:1 correlation between restriction fragment length polymorphism (RFLP) and certain inherited diseases. For example, it was discovered that US American populations that possessed normal (non-sickle cell) β-globin were associated with a particular RFLP only in 3 per cent cases, whereas those with HbS were accompanied in 87 per cent of the cases with the same aberrant RFLP. These findings have opened the way for diagnosing inherited disease alleles with the help of the RFLP associated with them.

The stretch of nucleotides containing the site of the A → T lesion in the Hb^S gene is associated with a 1.35 kb fragment cut by the RE *Mst* II. The normal counterpart (of Hb^A) is associated with a 1.15 kb fragment. Diagnosis was made easier by the synthesis of probes that were oligonucleotides (19 bp long) complementary to the normal and sickle-cell β-globin stretch of DNA and including the lesion for each mutant allele. Modifications of this basic strategy are now utilized for detecting mutations that indicate various inherited diseases.

Phenylketonuria (PKU), one of the first discovered 'inborn errors of metabolism' disorders', has been traditionally diagnosed by enzymatic assays of amniotic cells. RFLP and probes for the mutant allele are replacing the earlier methods. The recombinant DNA methods not only diagnose the diseased foetus but also the carriers of the mutant gene. Early foetal diagnosis of PKU and certain other diseases allows management of the diseases at the earliest opportunity.

Several defects that are associated with X-linked genes undergo new mutational lesions. Recombinant DNA methods are the only procedures available for diagnosis and prenatal therapeutic intervention. These diseases include deficiency of HPRT (hypoxanthine-guanine phosphoribosyl transferase) causing the Lesch-Nyhan syndrome, the Duchenne muscular dystrophy and deficiency of the enzyme ornithine transcarbamylase causing a disorder of the urea cycle system. Some of these mutations can be identified by allowing hybridization between the mRNA of the concerned gene and an RNA probe and using RNase to nick the probe on one side of the mismatch (Fig. 5.18).

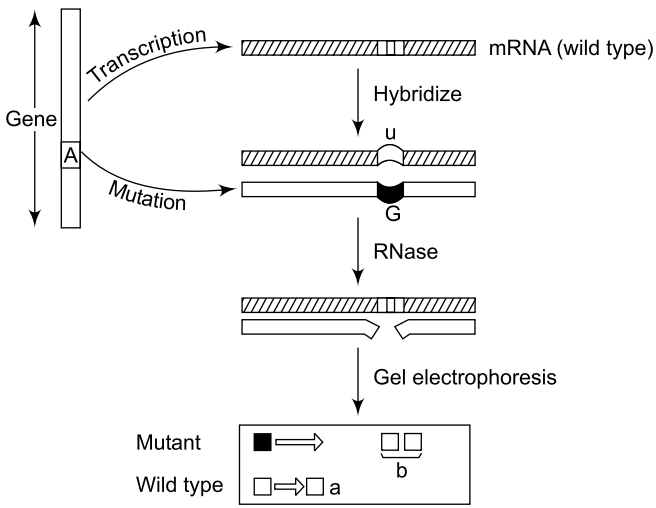


Fig. 5.18 *Detection of the gene of an inborn error of metabolism by hybridizing its nucleic acid with that of the normal or wild-type cell.*

It should be noted that RFLPs are of two types:

(i) With a single base-change lesion.
(ii) With lesions in certain tandemly repeated sequences in the RFLP.

Pooling of information from linkage studies and RFLP-associated disease genes has become a diagnostic strategy for several inherited disorders, the nature of mutations of which is not known precisely.

Diseases such as Huntington's chorea, cystic fibrosis, Duchenne muscular dystrophy, and adult polycystic kidney disease may be diagnosed by examining the RFLP accompanying their genes. Patients of the last named disease occur 1 in 1000 (approximately) in the USA population. A probe made from a variable number of tandem nucleic acid repeat sequences (VNTRs) associated with the gene (in

chromosome 16) picks out the β-globin locus as well as the disease gene. Prenatal diagnosis of this disease is useful particularly since a way of management is known, given sufficient notice by early detection of the defect.

3. Neoplasias

Recombinant DNA methods have also been utilized to diagnose DNA changes associated with certain neoplasia. Certain lymphomas are associated with translocations of a portion of an oncogene containing chromosome (c-*myc* on chromosome 8) and an Ig gene-containing chromosome (21 or 22). It appears that the translocation occurs as a result of aberrant joining of the V and J regions of the Ig gene. Probes made from the regions at which the chromosome breaks occur may be used for the detection of such lymphomas and leukemias. Of course, we must first be armed with a collection of probes made from the translocated region and from DNA lengths on either side of the potential break points.

4. Virus-caused Diseases

Recombinant DNA methods have potentialities for use in the diagnosis of viral infection. Hybridization of fresh tissues or cultured cells with probes made from viruses such as Herpes simplex, papilloma virus, human immunodeficiency virus and cytomegavirus have already been reported.

The use of recombinant DNA methods certainly provides precise means of detecting various genetic diseases and propensities or a vulnerability to them. These diagnostic methods are, however, still quite laborious and require the utilization of radioactive material that cannot be used generally in all diagnostic centres. There is a search, therefore, for non-radioactive probes and simplified protocols for diagnosis. One such method is the use of *biotinylated probes.* Analogues of dTTP and dUTP are first linked with *biotin,* and this nucleotide is then inserted into a DNA probe in place of the dTTP (in DNA) or dUTP (in RNA) (Fig. 5.19).

The biotin is conjugated to avidin, which is linked to an antibody linked to a fluorescent molecule. The labelled nucleic acids can be detected by observing fluorescence under a fluorescent microscope. There are other methods of detecting the presence of the probe. These methods do not compete well with the standard methods using radioactive tracers, since single sequences in the DNA may escape identification by the cruder non-radioactive probe methods. However, certain improvements have been reported that may make the use of biotinylated probes feasible for routine diagnosis of nucleic acid-based disorders.

5.6 ❑ APPLICATIONS OF GENE-CLONING TECHNIQUES

Gene-cloning techniques have unlimited scope for application, which includes technologies that cater to human health and prosperity.

Agriculture, animal husbandry, microbe-based industries and clinical intervention in human beings in terms of gene replacement therapy are loosely included under the global term 'biotechnology'. In a more circumscribed sense, biotechnology includes those enterprises that depend on the fermentation processes. Currently, it

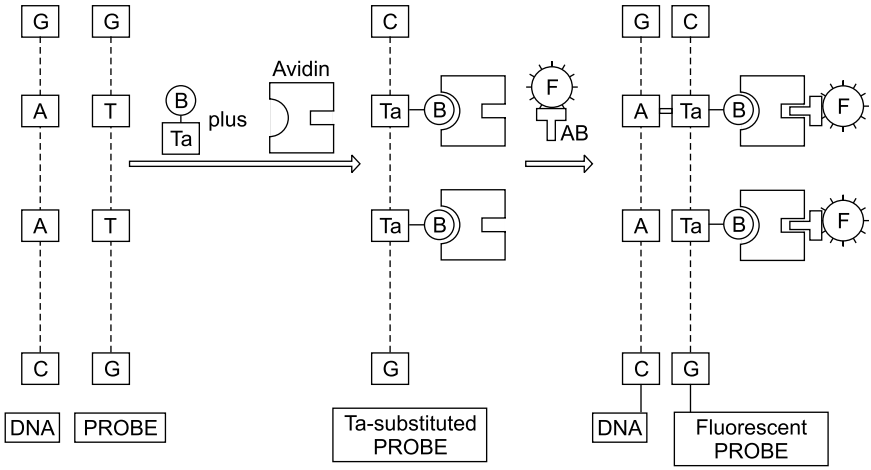


Fig. 5.19 *Non-radioactive labelling using a biotinylated probe. Biotin has a strong affinity for a protein such as avidin. An anti-avidin antibody, conjugated to a fluorescent dye, readily indicates the locations of avidin-bound biotin. The biotin is bound to the probe DNA in the form of biotinylated dUTP. The latter may be used in place of dTTP during preparation of a DNA probe by nick translation. The uracil-biotin-avidin-antibody-fluorescent dye complex may be visualized using a fluorescent microscope.*

is in vogue to restrict the definition further to fermentation-based technologies in which recombinant DNA plays a role.

Whatever the terms of reference in using the label of biotechnology, applications of genetic engineering in plants, animals, microbes, and single eukaryotic cells utilize characteristically unique ancillary techniques, and, of course, cater to different objectives. Hence these areas will be treated under separate headings.

Applications of gene-cloning techniques are being discovered in less traditional fields as well as in certain novel areas. These areas include the management of the environment, the supplementing of mining capabilities, and the engineering of designer proteins.

The unambiguous goal of plant, animal and microbe breeders has been to improve the quality and yield of products. This has been achieved by the traditional methods of breeding and selection, practised since the dawn of human civilization. The discovery of artificial mutagenesis augmented these methods by providing sources of newer alleles in a short time. The other alternative has been to screen naturally occurring populations for desired genes. Both strategies depend on random discoveries of wanted genes. This risk of uncertainty is compounded by the randomness inherent in the products of genetic recombination. In addition, the genetic alterations brought about by mutagenesis are also of a hit-and-miss (or shotgun) nature, and cannot be dictated, either in terms of the locus and/or those of the particular region of the locus that is to be manipulated. In short, artificial general mutagenesis causes an unpredictable set of alleles to appear in which the site of mutagenesis is also unpredictable. Traditional methods of gene introduction suffer from a further liability; the search for alleles is confined to the gene pool of

the variety in question. In other words, only members of a species can serve as the hunting ground for a desired allele. This is analogous to creating a new combination of cards by random shuffling of a finite deck. In addition, traditionally we have depended on the organism—plant, animal or microbe—to provide products that are already within the scope of their own repertoires. An *E. coli,* for instance, can be selected for better performance of the activity of one of its natural genes; it is not expected to provide products of genes uniquely present in other species.

Several of these disadvantages of traditional methods of strain improvement are overcome by gene-cloning strategies. First of all, they provide the means to transcend the gene pool barrier and coopt genes traditionally outside it. These genes may include man-made ones synthesized *in vitro.* Secondly, site-directed mutagenesis allows precise alterations at specific locations in or outside a gene sequence. Finally, a cell may be utilized as a mini-chemical factory for churning out a product made from an extrinsically supplied blueprint. In other words, the *E. coli* cell may take on, when supplied with the gene, the added chore of synthesizing and secreting insulin, a polypeptide characteristically produced by a very special type of animal cell.

5.6.1 Agriculture

Successful genetic manipulation of bacteria has been the incentive for similar interventions in plants and animals. The greater complexities of these higher eukaryotes have, however, been the major deterrent in the development of successful genetic engineering protocols for their improvement.

Mere alteration of the genetics of a single cell is not sufficient to result in the betterment of an existing individual or for the recovery of genetically altered progeny. Ancillary techniques are required for scaling up of manipulations performed at the single cell stage to that of the multicellular individual. Such back-up services are different for plant and animal systems. Plant and animal genetic engineering applications are, therefore, limited by the effectiveness of these methods.

Plants and animals differ in several ways; the difference is especially evident in their developmental strategies. Plant cells are totipotent; any not-too-old plant cell can be regenerated into an adult plant on a synthetic medium. Although all cells in an animal body also possess the same genetic information, the cells themselves are totipotent for a very limited early period of embryogenesis. Both these features are exploited in applications of gene cloning in plants and animals.

A corollary of plant and animal development is the tissue-specific and development-period-specific expression of most genes. Introduction of a cloned gene into any plant or animal cell does not guarantee the expression of the foreign gene in that cell. It requires a particular cellular environment for its transcription. This milieu is provided at the proper time of development and in the appropriate tissue. So, in order to express an inserted gene in a multicellular plant or animal, it has to be incorporated in the chromosomal DNA.

Now, in plants, there is no early separation of cells that will contribute to the formation of gametes in the future. On receiving the correct developmental signals, plant somatic cells become differentiated into precursors of reproductive or germ line cells. In plants, therefore, a gene installed stably in the chromosomal DNA is

likely to be included in the DNA that is transmitted to the progeny. In animals, however, very early in the process of development, certain cells or nuclei become earmarked for germ line cell formation. Hence, to insert a gene permanently in an animal stock, the gene must find its way into the DNA of cells that will form the gametes. Transgenic animals and plants have been obtained by using appropriate methods. These results indicate that it should not be impossible to introduce a gene permanently in the gene pool of an animal species, including that of human beings.

Not all goals of plant and animal genetic engineering involve such a permanent insertion of foreign genes. Sometimes, transient expression of genes may be the only objective. In such cases involving single free cells, the inserted gene is borne on a vector and is expressed in the cloning cell. Applications of transient expression have mostly been in research, where idiosyncrasies of gene expression are studied, and in the new biotechnological industries which aim to synthesize eukaryotic proteins in eukaryotic cells.

The scope of single cell cultures in different industries will be mentioned under the subsection 'New biotechnologies'.

Let us first examine the scope of applications and the issues that surface in the realm of plant genetic engineering.

1. Plants

As mentioned earlier, to make the new technology viable in agriculture, several lacunae in our understanding of plant molecular biology must be filled. This would perhaps allow us to (i) induce meiosis in somatic cells, (ii) regenerate cells of all economically important species into viable whole plants, (iii) select calli resistant to environmental stresses and regenerate them, (iv) improve techniques of protoplast fusion, (v) engineer successfully organelle genomes, and (vi) grow fertile plants from somatic hybrids in a routine manner.

Although plant cells are totipotent, the correct media for regeneration are not available for most species. Indeed, only a very few of the economically valuable families have responded to existing media. These include *Solanacese* (family to which belong potatoes, tomatoes, eggplants, tobacco and datura), and *Cruciferae* (the family of mustards, cabbages, cauliflowers and radishes). There is a demand for recipes of media for the regeneration of members of the pea family (*Leguminosae*) and of the cereal families (*Gramineae*) that supply the bulk of staple food crops. Successes with some commercially produced vegetables and fruit trees have been reported.

The best way of introducing a gene into a plant stock is to insert it first in a single cell. Separation of rigid plant tissues into naked protoplasts is not difficult. However, single protoplasts of only a very few species have been successfully regenerated into whole plants. Presently, efforts are being concentrated in various laboratories to develop methods for routine regeneration of several plant species.

(i) Single gene transfer The introduction of single genes permanently in the plant genome is banking on future developments in the use of Ti- and Ri-based vectors and on mixed vectors with useful genes from plasmids and plant viral genomes. Once again transformation of single cells and calluses from these cells have been

realized, but regeneration into normal plants is still not quite on the horizon for most species.

(ii) Mass gene transfer Somatic cell hybridization and regeneration of hybrid plants have been successful also mostly in *Solanaceous* species. The technique involved, including the regeneration of plants from hybrids, and from cells with introduced and genetically engineered organelles, offers exciting challenges to the applied plant scientist. As plants are not hampered, in general, by the sort of immune reactions typical in higher animals, regeneration of plants from somatic cell hybrids will provide a means for introducing bulk genetic material into an existing plant stock. This will be of immense utility in, say, incorporating nodule-specific genes from leguminous plants into non-leguminous crop species. Leguminous plants associate symbiotically with soil nitrogen-fixing bacteria (the *Rhizobia* species). Each species of the legume family has its own pet bacterial one; the root cells possess receptors for specific *Rhizhobium* species. If the gene for these receptors (plus perhaps a set of associated genes) can be introduced *en masse* into, say, a rice or a wheat plant stock, the need for externally supplied nitrogen fertilizers will be greatly, if not altogether, reduced.

(iii) Somaclones Callus cells, to recollect, have several types of anomalies in their genomes. Mutations and abnormalities of the chromosome number (aneuploidy) are present in many callus cells and are inherited by plants regenerated from them Such varieties derived from mutations in somatic cells are known as 'somaclonals' and are exploited as sources of variants for agricultural needs. Potato somaclones have resulted in some useful commercial varieties. Somaclonals have been obtained from a few other important crop species. DNA Plant Technology in New Jersey is using somaclonals to select strains of popcorn that have a butter-like flavour. These strains would not require the added butter with the accompanying calories and cholesterols.

(iv) Plant cell-bacterium somatic hybrids Somatic hybrids are also formed when plant protoplasts are mixed with lysozyme-treated bacteria and a fusing agent. This is a less arduous and more effective way of gene transfer into a plant system. The *Agrobacterium* vector, Ti plasmid, has been used experimentally to transform plant cells, but the percentage of success is not very impressive. Greater success is achieved when the agrobacterial cell is fused with the plant protoplast. The latter method also dispenses with the additional chore of extracting and purifying the vector DNA.

Somatic hybrids of nitrogen-fixing bacteria and crop plant cells may be envisaged that make the regenerated hybrid plants self-reliant in terms of nitrogen fertilizer requirements. Nitrogen fixers include the symbiotic *Rhizobium* species that associate with roots of leguminous plants and non-symbiotic free-living bacteria and blue-green algae.

At present one of the ways of studying the biology of nitrogen fixation by these microorganisms is to fuse plant cells in suspension, obtained from calli, with nitrogen-fixing bacteria. Since only legumes appear to possess the genetic information required by *Rhizobia* for forming root nodules, there are attempts also to obtain somatic hybrids of leguminous and non-leguminous species that may be amenable to infection and nodulation by *Rhizobia.*

(v) Pathogen resistance Yields of crops depend not only on the genetic heritage and conditions of cultivation but also on the susceptibility to their natural pests. The manner in which host-pathogen relationships have evolved together ensures that most hosts are infected by specific pathogens; or conversely, most pathogens have a fixed host range. As in the case of legume root cells and nitrogen fixers, hosts possess cell surface receptors that match antigens on the pathogen. Host-pathogen studies at the molecular level are expected to tell us the mutual requirements for a successful infection. It would then be possible to thwart the host-pathogen combine by gene-cloning methods. One method could be to mutate the host receptor sufficiently for the pathogen to fail to get a foothold on its host; or else, single cells could be selected against pathogens or toxins released by them, and the selected cells regenerated into plants that will be immune to attack by the same pathogens. Such successes have been reported from at least two laboratories: (i) A variety of maize resistant to the toxin present in its pest, *Dreschslera maydis,* has been regenerated from single cells selected against the toxin, and (ii) a tobacco cultivar resistant to methionine sulphoximine has been regenerated in a similar manner. Resistance to a pathogen may be induced in a plant by incorporating the gene that makes it resistant to the pathogen, and also the gene in the pathogen that makes it virulent. It has been observed that the virulence genes and resistance genes usually act in concert to prevent infection. If both genes become part of the host—plant genomes, it is likely that the plant will always be resistant to attacks by the pathogen.

(vi) Biociders Plant insecticides may be commercially prepared—chemical ones or naturally occurring microbial pests of the insects that destroy crops. Several bacteria are known to synthesize toxins that act against their insect hosts. One such bacteria is *Bacillus thuringiensis.* It synthesizes a crystalline protein during sporulation that is highly toxic to larvae of several ($\simeq$ 200) moth and butterfly species. Commercial preparations of *B. thuringiensis* toxin are widely used. Such preparations do not harm animals that may ingest them together with the plant product. The bacterial cells do not survive long in sunlight. The gene for the toxin has been cloned in *E. coli.* Other *Bacillus* species synthesize products toxic to beetles, causing 'milk' disease of their larvae. These biological insecticide producers—biociders in short—help tremendously in controlling pernicious pests like the Japanese beetle. Most of these natural biociders have a limited host range. Fungal insecticides have a broader host range and kill the insects on contact. Genetic engineering methods are being employed to improve the use of biociders by making them more effective under varying environmental conditions and against a much wider spectrum of host species.

(vii) Resistance to herbicides Crop plants may be made indifferent also to herbicides sprayed to protect them from competition from fast growing herbaceous weeds. Often, these herbicides cause undesirable metabolic effects on the main plants. There are reports of petunia and tobacco plants that have been made resistant to a much-used herbicide, namely glyphosphate. Some plants possess natural resistance to particular toxins. Maize, for instance, can detoxify triazine herbicides. It would pay to insert the gene for the enzyme that is responsible for the detoxification (one which substitutes a hydroxyl group for an atom of chlorine) into other plants.

(viii) Homozygous cultivars One of the demands of cultivators is to obtain plants that are homozygous in most, if not all, loci. Such totally homozygous plants can be obtained by diploidizing haploid callus, embryoid or plantlet cells with colchicine treatment. This would be a less time-consuming job than using the traditional practice, in agronomy, of performing several bouts of inbreeding and back-crossing.

(ix) Enrichment of storage proteins At present, there appears to be little hope of improving the sort of traits cultivators cherish in a crop plant, since most of these traits are suspected to be due to multiple genes about which frankly, not much is known. However, it should be possible to transform single cells into producers of high levels of certain metabolites for which single genes have been identified. These cells may be regenerated into plants. One such project would be to enrich foodgrains and legumes with genes that will deliver balanced amino acids in the digestive systems of human beings and livestock. Seeds of cereals—that is, rice, wheat, barley, maize, millet grains—contain carbohydrates and storage proteins, which, on digestion, release amino acids. These do not include proline and tryptophan. Similarly, leguminous seeds fail to provide sufficient methionine. To make both types of seeds self-sufficient in terms of balanced amino acids required for the proper health of the consumers, it is first important to know something about the regulation mechanisms that control the expression of the storage protein genes only in the seed stage.

Soybeans are very rich in storage proteins. Indeed, if only about 1 per cent more of methionine could be guaranteed on the degradation of these proteins, soybean flour would be comparable in nutritive value to meat. An enzyme, namely urease, degrades storage proteins, with methionine as one of the end products. Jack-bean urease action causes 3.7 per cent to 9 per cent methionines while soybean urease is responsible for only 1.2 per cent to 1.3 per cent methionine, under the same circumstances. It would be useful to know enough about the metabolism of the urease gene in leguminous species to be able to supplement the soybean urease gene with that from the jack-bean.

Considerable information about storage proteins of maize is available at present. In maize, as in other cereals, these proteins may be classified into three categories, on the basis of their solubilities at different pH and salt concentrations. These are albumins, globulins and a third group rich in proline and a nitrogen amide, such as glutamine. In maize (*Zea mays*) this third group is known as 'zein'; in barley (*Hordeum vulgaris*) it is called 'hordein'. Zeins occur as several groups or families of related proteins. The genes for these proteins are usually in several copies that occur in tandem clusters. The genes are intronless, unlike most major eukaryotic genes, and their expression is regulated during the development of the endosperm (the tissue in the seed that contains the storage macromolecules). Much research is focused on the details of the gene control system, including the characteristics of regulatory DNA sequences and requirements of specific transcription factors. At present, these studies are extending our knowledge not only about the mechanism of storage protein synthesis but also about control systems of plant genes. It is not unreasonable to expect that enrichment of seed storage proteins by genetic manipulation may be achieved earlier than other targets set up by plant genetic engineers.

Once the constraints and requirements of expressions of specific genes in plants are better understood, it should not be impossible to allow gene expressions in tissues other than the ones in which they normally are expressed. For instance, the genes for storage proteins are over active only in the seed formation stage. If these genes could be equipped with the correct transcriptional signals and factors in, say, the leaves, we would have plants with protein-rich leaves.

Jaynes (Lousiana State University) and Dodds (International Potato Centre, Lima, Peru, South America) are collaborating to introduce genes for proteins into carbobydrate-rich staple foods, such as potato, cassava and rice, that will make them as rich as meat in proteins.

(x) Improvement of photosynthetic capacity Another application in agriculture should be to make plants more efficient trappers of solar energy. The energy of the sun is converted into usable chemical energy in the form of carbohydrates by the process of photosynthesis within the green organelles, chloroplasts, present in plant cells. The proteins used for the process are encoded partly by chloroplast and partly by nuclear genes. Two of these nuclear genes-specified proteins are (i) the one binding chlorophyll *a* and *b,* and (ii) one of the subunits of the rate-limiting photosynthetic enzyme, ribulose-1, 5-biphosphate (RUBP) carboxylase. Plants differ in their efficiencies with which they photosynthesize. Perhaps these efficiencies can be enhanced by genetically engineered chl DNA and/or nuclear DNA concerned with photosynthesis.

(xi) Stress resistance Improvement of plants by gene manipulation techniques may be brought about in terms of making them withstand chronic stress conditions prevalent in their habitat or places of cultivation. The stresses most commonly encountered are brackish or saline soil, drought, excessive heat or freezing conditions. Perhaps, genes of plants that grow habitually in the same chronically stressed conditions may be introduced into the useful plants that normally succumb under these conditions. Of course, before this becomes possible in practice, much more has to be known about the genes or combinations of genes that allow the native plants in these unusual habitats to thrive.

Studies have been reported which indicate that tolerance to salt may be correlated with that of calli in several species. Calli grown from naturally salt-sensitive (*Lycopersicum esculentum*) and salt-tolerant (*L. peruvianum*) species of the tomato were found to be sensitive and tolerant to abnormal levels of salt in the medium. Tobacco single cell cultures grown in strong brine (6.4 g NaCl/1) resulted in a few cells that grew into calli and regenerated into plants that were also salt tolerant. Studies such as these suggest that plant tissue culture methods may be utilized to select for cultivars against desired selection pressures. Indeed cells of *Lycopersicum esculentum,* tolerant to inhibitory levels of metals such as aluminium, were recovered by Meredith in 1978 by using such a selection method.

(xii) Post-harvest preservation Post-harvest deterioration of agricultural produce is another source of tremendous economic loss. An understanding of the molecular genetics of the biochemical reactions participating in these destructive metabolic processes will pave the way for crippling them by appropriate gene manipulations.

(xiii) Selection of auxotrophs Single cell cultures may be utilized for biochemical, physiological and genetic studies. For such investigations it is convenient to possess a large spectrum of mutant marker genes. Attempts have been successful in developing auxotrophic mutants by treatment of single cells in culture. Auxotrophic mutants, as in the case of microorganisms are the easiest to select, and are, therefore, particularly useful for research.

The first auxotrophic mutants were developed in tobacco by Carlson (1970). There were six mutants: defective in biotin, P-aminobenzyoic acid, arginine, lysine, proline and hypoxanthine synthesis. They were identified from clones of cells that survived a 5-bromouridine treatment. As mentioned elsewhere (enrichment method of mutations), wild-type cells that grow on a minimal medium incorporate 5-BrdU (a toxic analogue of thymidine) during DNA synthesis. These cells subsequently die, as the analogue breaks up the DNA in the presence of light. The mutants that do not grow in the minimal medium are not affected, since no DNA replication takes place in static or non-dividing cells. The surviving cells are then tested for the mutations that they carry.

Another 'suicide' agent used to recover plant single cell auxotrophs is chlorate. Muller and Grape (1978) used chlorate to recover cells that cannot synthesize nitrate reductase. In this case, the chlorate is converted, in the cell, into chlorite which is the toxic substance. The conversion is mediated by nitrate reductase. Hence, nitrate reductase-deficient cells will be immune to chlorates, a substance which is not toxic by itself.

Mutagen treatment of protoplasts and single cells of plants, followed by growth of colonies in minimal and enriched media, has also been used to select several types of auxotrophic mutants.

The list of what may be expected by genetic engineering of plants is too open-ended to be given even as a tentative one. Almost any process that contributes to yield, quality, resistance to pests and pesticides, ability to thrive in inimical environments and to synthesize metabolites on demand should be amenable to modification by gene-cloning methods.

Monsanto, a multinational company, has already developed a tomato that resists infection by natural viral and other pests, and it is also immune to the commonly used herbicides. A particular soup company is bent on developing more solid tomatoes, fewer of which will be required to make ketchups. Still others are trying to develop tomatoes that can be kept longer on the plants before harvesting. Such tomatoes will have better flavours than the ones that are today harvested green and ripened artificially for the market.

(xiv) Secondary metabolite production An ever-widening range of plant secondary metabolites that may be utilized profitably is being discovered. It is reasonable to expect single cell cultures to supply the metabolites that are now being extracted from intact plants. These include steroids, flavonoids, alkaloids and terpenoids that are useful as drugs, flavourings, perfumes and insecticides to name just a few areas of interest.

Natural vanilla, containing more than 150 flavour-imparting components is very expensive, as it is obtained mostly from one small country. Artificial vanilla has

only one of these components and so lacks the natural rich flavour. Genetically engineered cells are expected soon to provide vanilla with almost natural qualities in a cheaper manner.

Some of the better known plant-produced drugs include quinine (an alkaloid from *Cinchona ledgeriana*) digoxin (a glycoside used for cardiac trouble, from *Digitalis lanata*), codeine (an alkaloid analgesic from *Papaver somniferum* or poppy), diosgenin (a steroid from *Dioscorea deltoidea* which is an anti-fertility agent) and vincristine [an alkaloid from *Vinca* (*Catharanthus*) *roseus* which is anti-leukemic]. Various plant-essential oils, such as jasmine, rose, and lemon grass oil, are in demand in the highly profitable perfume industry. Thaumatin, a substance that is several times sweeter than sugars is used as a non-nutritive sweetener in products such as toothpaste. It is at present extracted from the plant *Thaumatococcus danielli.* An insecticide, pyrethrin, is obtained from one of the chrysanthemums.

Traditional methods require a very large quantity of starting plant material for the recovery of milligram amounts of the product. Cultivation of single cells in biofermenters should provide a more effective way of producing the same plant metabolites at a lower cost.

Biotransformation of plant products is also visualized as a commercially profitable venture. Immobilized cells of *Digitalis lanata* have been used to convert β-methyl-digitoxin to β-methylidigoxin. Although both products are used for cardiac treatment, the latter is preferred. Maximum drug recovery is attained at present from *D. lanata,* which synthesizes both forms.

2. Animal

Improvement in animals of economic value has been accomplished traditionally by breeding and selection methods. Considerable improvement can be, and has already been, realized by employing a few novel or improved management techniques that are based on a better knowledge of the reproductive biology of animals.

Direct manipulation of the genome, with the alterations transmitable to the progeny, is of course, a certain way of introducing genetic quality. As mentioned earlier, the unique developmental pattern of animals, with reproductive cells ensconced into a definite tissue early in development, makes it rather difficult to create such transgenic animals in a routine manner. However, transgenic mice with genes introduced in the germ line DNA have been reported and serve as a milestone for future interventions in other animals, including human beings.

(i) Cloning Cloning of an important strain or plant cultivar is a prime objective of both plant and animal applied scientists. It has been achieved in mice by allowing the separated cells of an 8-cell embryonic mass to develop in surrogate mothers. Introduction of a desired DNA into these cells, before implantation in the uterus, should be one of the avenues for certain animals with targetted genetic alterations. In human beings, one can visualize this method for correcting inherited genetic disorders once the ancillary techniques are developed and mastered.

(ii) Twinning and superovulation Twinning and superovulation are two techniques that offer ways of producing extra offspring from the same maternal parent.

Hormone treatments can induce an embryo to split into two viable ones, as well as induce the release of extra ova for fertilization. These ova may be fertilized *in vitro* or by artificial insemination with sperms stored by freezing. These methods, however, will become practical realities only when we develop better techniques for (i) storing sperms and embryos, (ii) for transferring embryos without damage, (iii) for identifying the sex chromosome of the sperm, and (iv) for synchronous estrus in large numbers of animals. Gene-cloning methods can be visualized for introducing known genetic capabilities in the ova, in single cells of early embryos, or by fusion of somatic cells with fertilized ova. At the bench level, some of these projects have been successful. At present, these pioneering exercises are adding to our meagre knowledge about the reproductive biology of animals. Improvements in genetic engineering of animal cells welded with those in ancillary techniques would result in usable methods of animal stock improvement by design and not by chance.

(iii) Veterinary medicine Other standard methods for the improvement of animal stocks include prevention, diagnosis and management of pathogen-caused diseases, improvement in nutrition and in growth potentials. Vaccines produced through genetic engineering, against foot-and-mouth diseases of cattle, rabies affecting cattle, pigs, and dogs, and common poultry viral diseases such as Newcastle's disease, that are already in the market, have taken the edge off most losses due to recurrent epidemics.

(iv) Improving quality by genetic engineering The success of transgenesis in mouse has provided sufficient confidence of repeating this performance with animals, fowls and fishes having economic importance. Palmiter and Brinster from the Universities of Washington and Pennsylvania, respectively, injected rat growth hormone genes into fertilized eggs of mouse. These genes were found to have been transmitted and expressed in the progeny, which were larger in size than normal mouse. Similar success has attended transgenesis in pigs, sheep and rabbits. The transgenic animals were normal, expressing the product(s) of the additional gene(s). At least 12 strains of transgenic pigs and over 1000 transgenic mice have been produced in less than ten years. The transgenic pigs gain greater weight on less food and are ready to be sold at least six to seven weeks earlier than normal strains. These animals are still undergoing testing (some are in the seventh generation or more) and will be ready for release to farmers in Australia in the near future.

Another example of the potential power of genetic engineering is the development of a strain of carp (*rohu* fish) with a growth hormone gene from a trout. This was developed at the Chesapeake Bay Institute in Shady Side, Maryland, USA. The Chinese performed a similar 'magic' with goldfish. Armed with a human growth hormone these small fish grow to four times their normal size. Indeed, gene transfer for improvement of fish is one of the most busy areas in the domain of animal improvement.

The above and other success stories impress upon us the fact that an entirely new approach has come to stay in matters of improving animals of interest. Almost any particular trait may be altered in an animal, by injecting the gene responsible for it, provided we know what the gene is and how it is regulated in the animal. There

is no danger of creating weird hybrids, as the genome of an individual tolerates just so much imported genes and not more. In fact, one does not need to worry about a foreign gene or two being introduced by the genetic engineer. Viruses and transposons have been natural genetic engineers of longstanding. They have peppered the genome of most of their host species with part or whole of their own genomes, with very little appreciable difference, if any.

5.6.2 Biotechnological Industries

The pharmaceutical and chemical industries were the first to exploit the new biotechnology. The chemical industries are broadly of two types: those that produce commodity (or very cheap) chemicals, and those which manufacture speciality chemicals, that represent products used in the food-processing industries.

Other industries that manufacture products pertinent to the management of the environment have also realized the potential of using the new biotechnological strategies. Environmental applications include control of pollution due to (i) toxic wastes, (ii) various technological advances, and (iii) insect pests. In addition, unexploited environmental riches such as metals and minerals, present in minute quantities in low grade ores, may be made available by using specially engineered biological organisms and/or their products.

A futuristic application is envisaged in the field named 'bioelectronics' or 'bionics'. The potential of biological material to serve as extra-sensitive conductors, sensors, and support matrices for semiconductor molecules is already being explored.

The different biotechnological industries differ in the standards expected for the purity and lack of contamination of their products. Pharmaceuticals have to be strictly contamination-free and also usually of absolute purity. Such strict purity and asepsis are not demanded for commodity chemicals. Speciality chemicals, again, must adhere to strict standards of purity and non-contamination. These differences in the quality of the product, representing degrees of rigour employed to produce them, are reflected in the price of the finished products. Drugs are highly expensive, but are used in minute doses. Commodity chemicals (amino acids, industrial alcohols) that are used in bulk and need not be too pure, are very cheap (₹20–30 per kilogram).

Let us look at a few of the products that are manufactured or are in the pipeline, using biotechnological approaches.

1. Pharmaceuticals

Pharmaceutical industries have been the pioneers as well as the most successful in the industrial applications of recombinant DNA-based product improvement.

Traditionally, pharmaceuticals have been derived by fermentation via microorganisms, extracted and purified from natural sources, or synthesized chemically. They are neither easy nor inexpensive to prepare chemically and are scarce in the natural sources. Recombinant DNA techniques provide opportunities to synthesize large quantities of pure products under controlled conditions in both homologous and heterologous systems.

Earlier improvements in strains producing a fungal metabolite depended on genetic information present only in the native cell. Molecular genetic methods permit the synthesis of substances from genes taken from other species.

Drugs derived from animal or other sources and used for human therapy often cause immune reactions in the patient. This shortcoming is bypassed if the proteins are expressed from homologous—that is, human—genes. Insulin, as an example, synthesized in *E. coli* cells, from the human insulin gene, is already in the market.

The drugs and pharmaceuticals may be synthesized in heterologous cells or in animal cell lines that produce them. For instance, human chorionic gonadotropin (HCG), erythropoietin and calcitonin synthesizing cell lines have been reported. Of these, only HCG is successfully used for commercial production. The others did not survive, as optimum culture conditions under industrial constraints have not been developed for these cells. This is an obvious area that demands further development.

The drugs and pharmaceuticals that are candidates for the new biotechnology belong broadly to the following categories, although, strictly speaking, many of them are pleiotropic in action: (i) hormones, (ii) growth factors, (iii) blood-derived factors, (iv) enzymes, (v) antibiotics, (vi) immunoproteins, (vii) immunodiagnostics, and (viii) DNA probes for molecular diagnosis.

(i) Hormones Hormones are proteins, glycoproteins and steroids, usually made in one part of the body and transported by body fluids to target tissues where they trigger off specific gene expressions. Hormones, or hormone-like substances, may also be synthesized locally. This is suggested by the presence, in brain cells, of receptors for insulin-like substances that are known not to enter the brain from the blood stream. Both peptide and steroid hormones may be manufactured by rDNA technology.

(a) Protein hormones Insulin, the polypeptide hormone used against diabetes, was the first rDNA-produced drug that was released for consumption. The type of problem encountered in raising a laboratory success to a mass-scale production venture is given by the example of successful recovery of functional insulin by the appropriate manufacturing process. Insulin consists of two peptides, A and B. It is rather difficult to join them correctly by chemical methods. In the native cell, insulin is derived from a proinsulin, which contains a small segment between the peptides A and B. The intervening extra peptide is removed and functional insulin created in the beta cells of the pancreas. One of the companies marketing rDNA insulin has developed a method to synthesize A and B and also to achieve their joining within *E. coli* cells. Others are trying to develop bioreactors with immobilized enzymes that will convert nascent secreted A and B peptides into a functional hormone.

The human growth hormone (HGH) is another polypeptide hormone that has been produced by rDNA technology. Other polypeptide hormones that are candidates for production by similar methods are parathyroid hormone, erythropoietin and the nerve growth factor. Smaller peptide hormones are prepared easily and inexpensively by chemical synthesis. These include somatostatin (the first hormone obtained in the laboratory by synthesis in *E. coli*), calcitonin, cholecystokinin, bombesin and the adrenocorticotropic hormone (ACTH), all of which are important pharmaceutical

products. Whether rDNA technology is used for their production in future will depend on the relative costs of production of the old and new technologies.

Certain small polypeptides have been discovered in animal systems that behave like opiates from plant sources. These neuroactive peptides or endogenous opiates are hormone-like and are classed into two groups: enkalphins (small, 5 amino acid lengths) and endorphins (longer, 31 amino acid chains). There is a great deal of interest in finding out more about the nature and action mechanism of these substances. Analogues have been developed that mimic the natural product, but which are less degradable by enzymes in tissues than the endogenous proteins. One such analogue, dynorphin, has been found to be an extremely powerful analgesic, more than a 1000 times as effective as the standard standby, morphine.

One other neuroactive peptide of interest is the melanocyte-stimulating hormone (MSH) which appears to improve nervous mechanisms involved in memory and also the attention span.

Gene sequences for such short amino acid chains may be synthesized and expressed in bacterial cells. Since these products are synthesized in only trace amounts in the body, it may be profitable to use genetic technology for obtaining them in bulk. Sufficient quantities are required immediately for at least research purposes.

(b) Steroid hormones Steroid hormones are extracted from natural biological sources, or converted by microbial fermentation from precursors extracted from plants. The latter process is several orders less expensive than the one using only chemical synthesis. For example, cortisone may be made by chemical synthesis, or be converted by microbial enzymes from progesterone. The latter process yields the same product quality at a fraction of the cost of the chemical one.

There are, therefore, searches for microbes with enzymes that ferment an inexpensive intermediate substrate into non-protein drugs. The conversion need not be carried out in the microorganism. The gene(s) for the latter may be introduced into a bacterium such as *E. coli* or in a single celled eukaryote like yeast.

A very large proportion of drugs for cardiovascular and hypertension conditions are obtained from plants. The morphine group of drugs is synthesized by biochemical pathways from one or a few precursors. If the genes for mediating these conversions are identified and expressed in heterologous industrially cultivable cells, the dependence on natural sources, and hence costs, would be greatly reduced. The special genius of recombinant DNA practitioners would also be challenged with engineering of enzymes more efficient than the ones in the native plants.

(ii) Growth factors Several factors that are required for promoting the growth of specific types of cells have been identified. Most of them have been isolated from serum and are responsible for the indispensability of serum in current methods of animal cell cultures. Animal cell culture can be used in mass-scale ventures only when culture conditions are standardized. A serum-free medium would eliminate much of the variability in the nutrient medium, provided of course it is replaced with the appropriate growth factors. There is, therefore, a need for easily produced large quantities of these growth factors, for industrially used animal cell cultures, apart from their potential clinical applications.

The factors already identified include the epidermal growth factor (EGF), the nerve growth factor (NGF), the platelet-derived growth factor (PDGF), the fibroblast growth factor (FGF), the colony-stimulating factor (CSF), the macrophage-derived growth factor (MDGF), the tumour angiogenesis factor (TAF), the skeletal growth factor (SGF), and the extracellular matrix protein fibronectin (FN), that promotes adhesion of fibroblast cells to a substrate and also acts as a mitogen.

(iii) Blood-derived factors The largest volume of pharmaceutical products on the market is obtained from blood. These products include human serum albumin (HSA), the anti-clotting (anti-haemophilic) factor and gamma globulin. Of these, the last is an immunoprotein and will be discussed elsewhere.

HSA is used in considerable quantity in surgery and in the treatment of physical traumas resulting from burns, shocks, and so on. HSA is obtained from blood collections; however, the novel threat of AIDS as well as of other identified (hepatitis virus) and non-identified pathogenic viruses have made the use of untested blood products of questionable value. Hence, the need to synthesize contamination-free HSA has been felt. HSA is a 585 amino acid-long polypeptide. It was first synthesized by genetically engineered yeast and bacterial cells in the R & D laboratories of Genentech. Upjohn Pharmaceuticals developed their own methods for doing the same.

The second largest blood-derived proteins that are in demand are those that permit clotting of blood in haemophilics. Haemophilia is an inherited disorder, where blood fails to clot for the lack of a specific protein. In the A type of hemophilia, this factor is known as factor VIII; in B haemophilias it is factor IX. At present both factors are extracted from blood and are extremely expensive (profitable for the industry, of course). Biosynthetically produced clotting factors would lower the cost and also increase the sales. Cloning of both factors are underway. They are very large glycoproteins. Hence, a great deal of R & D will be required before they are ready for marketing.

(iv) Enzymes The pharmaceutically important human enzymes are those that dissolve blood (thrombolytic) and those that allow blood to clot (fibrinolytic). The antihemolytic factors VIII and IX are actually administered as precursor proenzymes, which get converted in the body to the required factors.

Two thrombolytic enzymes in maximum use are streptokinase and urokinase. Streptokinase is synthesized by the soil bacteria *Streptomyces.* Urokinase is the only industrially produced drug obtained from mammalian cell cultures. Efficient recovery of the product using monoclonal antibodies has given a boost to the production of these thrombolytic enzymes by rDNA technology. Costs are still prohibitive. Also, clinical tests do not show complete satisfaction. There are investigations, therefore, to understand the structure and function of these enzymes, to discover other such enzymes, and to develop better ways of producing usable enzymes via genetically engineered cells.

(v) Antibiotics The horizon of pharmaceutical industries was widened spectacularly by the discovery of the antibiotic penicillin in the 1940s. Antibiotics are antimicrobial agents obtained from soil organisms. They are non-protein secondary

metabolites that result from a complex network of biochemical reactions within the microorganism.

The usual method of improving the quality and yield of an antibiotic has been the path of mutation breeding and selection. One of the best examples of success using this method is that of the improvement of penicillin-yielding *Penicillium chrysogenum.* Fleming had discovered penicillin in *P. notatum,* which is not a great producer. A strain of *P. chrysogenum* was discovered in Illinois (USA) that had greater potential for synthesizing the antibiotic. An original NRRL-1951 strain passed through the laboratories of the Carnegie Institute of Washington, the University of Wisconsin and E. Lilly Industries before emerging as strain E.15.1 that produces 55 times more penicillin than the initial one. Wisconsin employed a sequential series of mutation and selection procedures using mainly X-rays and UV rays as the physical mutagens. The Lilly concern employed mainly chemical mutagens.

It should be possible to improve the production of the long list of antibiotics using rDNA techniques once more is known about the gene/genes involved in the process and the mechanisms regulating its/their expression(s). The secondary metabolite appears to be synthesized—courtesy of several genes. In *Streptomyces,* the other great producer of a variety of antibiotics, the genes for at least some products have been found to be clustered. In many cases these genes are on plasmids. There are attempts, therefore, to clone the cluster *in toto,* until more is known about the pathway of synthesis and contributions of individual genes in the venture.

Such a report was made in 1984 from the John Innes Institute in Norwich, England. A continuous region of *Streptomyces coelicolor* DNA was cloned into *S. parvulus,* where the antibiotic actinorhodin was synthesized correctly. The use of vectors which can comfortably carry large DNA inserts (such as the cluster of genes in an antibiotic synthesis pathway) that also are easily introduced into *Streptomyces* should be explored for the production of a single antibiotic in a cloning cell. Normally, the native antibiotic producer synthesizes a heterogeneous mixture of different antibiotics, and also various species of each type.

Incorporation of 'useful' genes by sexual recombination has not been successful due to a variety of reproductive barriers. Far better promise for recombination is offered by somatic cell hybridization procedures. Recently, as many as four different strains of *Streptomyces* have been combined into somatic hybrid strains that segregate recombinants after meiosis.

Similar somatic hybrids of *Cephalosporium acremonium,* that synthesize cephalosporins, have been found to be better producers than those derived from matings. Interspecific and intergeneric somatic hybrids can also be visualized. *Aspergillus nidulans* and *A. rugulosus* have produced viable somatic hybrids. Intergeneric hybrids are rarer; the only report is that between two yeasts: *Saccharomycopsis fibuligera* and *Candida tropicalics.*

There is another way in which recombinant DNA techniques may influence antibiotics that are to be produced industrially. Very often only a partly finished product is retrieved from the synthesizing organisms. The finishing touches to make the active product is given by chemists. Such semisynthetic penicillins have been made from fermentation-derived penicillin by further fermentation to 6-aminopenicillinic acid (6-APA) with a final conversion to α-aminobenzyl penicillin (ampicillin) (Fig. 5.20).

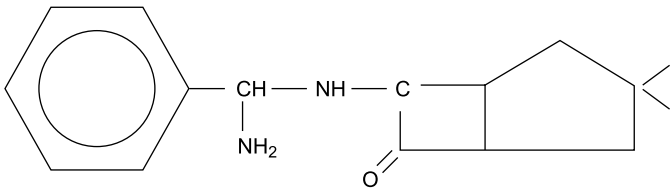


Fig. 5.20 *Ampicillin or α-aminobenzyl penicillin.*

Many of the semisynthetic penicillins are resistant to degradation in the digestive tract due to the penicillinase producing bacilli and, therefore, may be administered orally and not by injection. Several such broad-range penicillinase-resistant semisynthetic antibiotics are on the market.

Streptomyces griseus produces the valuable anti-gram-negative bacterial antibiotic, namely streptomycin. Between penicillin (which acts against gram-positive bacteria) and streptomycin, the majority of common infections may be controlled.

(vi) Immunoproteins and immuno-diagnostics The immune system consists of varieties of proteins, which cooperate in different combinations to ensure both short-term hypersensitivity-immune responses as well as long-term cell-mediated and humoral immune reactions. The major components are antibodies, antigens, lymphokines and cytokines.

(a) Antibodies Antibodies are the immunoglobulin molecules (IgM, IgG, IgA, IgD and IgE) which mount immune reactions in the body against incompatible substances generally called 'antigens'. Hundreds of uses have been discovered for antibodies for therapy, diagnosis, and analytical and drug delivery procedures. Usually, an antigenic challenge triggers off the synthesis of antibodies that match different antigenic determinators on the antigen. It is difficult to fish out an antibody for a specific antigen from such a polyclonal mixture. Today, it is possible to synthesize monoclonal antibodies against practically any given antigen by the hybridoma technology. Monoclonal antibodies (MABs) are one of the finest immunodiagnostic reagents available that can accomplish a staggeringly large spectrum of jobs.

Hybridoma technology has developed to a stage where a limitless supply of MABs is made available by mass culture of hybridoma cells. Their most varied use is in diagnosis and assays. They can be used to search for any antigen in whole organisms or *in vitro* systems. They can be used to deliver tumour-specific drugs, and to find tumour foci by tagging the MAB with a fluorescent dye. They are used to identify microbial pathogens, look for deleterious proteins in prenatal conditions and also used as immobilized reagents for identification and purification of proteins and nucleic acids. Hybridoma technology, one of the spin-off ancillary techniques of molecular biology, has an increasing potential for use, and, hence, serves as an incentive for industrial production.

(b) Antigens/vaccines Antigens are substances that provoke synthesis of matching antibodies. The body retaliates against non-cellular substances as well as antigens on surfaces of invading organisms—bacterial, protozoal, fungal and viral—under natural circumstances, and against man-made circumstances of tissue transplantation

(skin, organ, blood). Vaccines are antigens of infecting organisms or non-self cells, which have been prepared to elicit antibody formation but not the disease syndrome of the pathogen in the body.

The normal practice for preparing vaccines is to introduce a small dose of a crippled pathogen into the body. The body responds by a primary reaction where a number of β-cells are primed to differentiate into antibody-secreting plasma cells while a certain population remains static at a partially differentiated stage. The latter (memory cells) respond immediately on a second presentation of the antigen (e.g., during an epidemic in the neighbourhood) and amplify in a short time the number of B cells that synthesize the antibody.

The antigens used as vaccines are the killed or attenuated pathogens. There is always the danger of contamination by recoverable live pathogens. Recombinant DNA methods may be used to synthesize specific non-pathogenic vaccines by incorporating the gene sequence for a particular external antigenic region of the pathogen in a DNA vector. These hybrid DNAs are expressed in bacterial cells, and the products are extracted for use as vaccines. These vaccines elicit the antibody response characteristically produced by the whole pathogen, though they make weaker vaccines than the natural pathogens.

A vaccine for the foot-and-mouth disease, a great destroyer of livestock property, has been developed and now commercially produced by using such a method (Fig. 5.21).

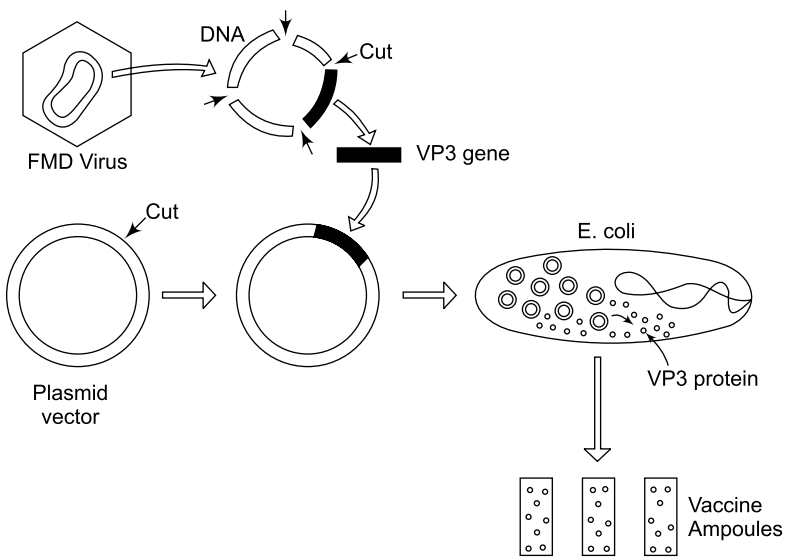


Fig. 5.21 *Synthetic peptide vaccine for the viral foot-and-mouth disease.*

Two strategies are being employed to prepare such subunit vaccines. In one method, portions of the pathogen DNA that codes for a surface protein are expressed in bacterial cells, as in the case of the FMDV. In the other, regions of the antigen that interface with the environment (that is, regions on the surface of the antigenic group or protein) are scanned for short peptides, which are then synthesized chemically

or in genetically engineered bacteria. These synthetic peptide vaccines are just as effective in invoking antibody formation in the body as the intact antigen. Creation of a synthetic peptide vaccine has been facilitated by developments of computer programmes that graphically depict the chemical groups on a given antigen that may be copied into synthetic peptides.

Vaccines for certain developmental stages of the malarial parasite are in the offing. Development of antiviral vaccines is proving to be more difficult. One problem with some virus pathogens—such as the influenza virus—is the rapidly altered surface antigens in successive epidemics. There are attempts to splice the characteristic surface antigen gene from several such variants or races of a virus to produce a 'consensus' antigen gene, which should express a hybrid antigen to be used as a vaccine against the different alternate coat proteins of the pathogen.

Development of vaccines by rDNA technology will benefit both man and animals. The number of pathogenic viruses, bacteria, nematodes, fungi and protozoa that cause animal and human diseases is very large and varied. They are responsible for enormous losses in terms of health and productivity.

(c) Interferons The antiviral, antitumour and other activities of interferons make them one of the most sought-after compounds for both research and application. The rDNA technology has been instrumental in providing methods of interferon recovery that are superior in terms of purity and yield of product to extraction procedures from natural sources. However, purification methods had to be improved further to allow production at even a pilot plant level. Rubinstein *et al.* (1978) reported an efficient method of human interferon recovery. Human peripheral leucocytes were cultured together with the Newcastle disease virus, an affliction which affects fowls. By HPLC (involving reverse phase, arid ordinary affinity chromatography), the product was bound to silica gel and ricin particles and eluted by the solvent *n*-butanol. High yields of very high activity interferon were obtained by this procedure. MABs have also been inducted for use in the purification of interferons. Gel electrophoretic fractionation of the different species of interferons is an additional refinement that is utilized effectively.

There is a great demand for mass-produced interferons, as their involvement in diverse immune activities may be exploited in therapy. At present the interferons are being studied to discover the speciality features of these multifunctional immunoproteins.

2. Commodity Chemicals

Commodity chemicals are very cheap compounds produced by the chemical industries. They are represented by industrially used solvents, enzymes, amino acids, organic acids, fertilizers and pesticides. They cost typically not more than about ₹24–30 per kilogram.

Most commodity chemicals are at present derived or synthesized chemically from petroleum, or coal and natural gas. The inevitably dwindling non-renewable energy sources as well as the potentials of genetic engineering are expected to bring about a shift towards the use of renewable sources, such as biomass, which includes agricultural and industrial wastes. A few examples of such possibilities are now given.

(i) Solvents One of the larger classes of industrial chemicals comprises solvents and organic acids. Glycerol, ethanol, *n*-butanol and acetone are the most important solvents produced in very large quantities.

Yeast is the traditional fermenter that converts sugars to ethanol. Other microorganisms have been discovered that are more efficient producers than yeasts. One of them is *Zymomonas mobilis,* a bacterium that ferments palm syrup into wine. For convenience of the chemical engineer, it would be worthwhile to look for fermenters among thermophilic bacteria. Several operations would be simplified and energy costs reduced if the fermenting organisms can operate at high temperatures in bioreactors. Heat-transfer devices would become redundant and discharges at higher temperatures will cut down on the energy requirements for heating them again for distillation.

The starting material for the solvent produced in the largest bulk, namely ethanol, is sugarcane or beet-derived sugar or molasses left behind after the sugar has been crystallized. Starches from wheat, rye, maize and starchy underground roots have also been used. There is a need to discover alternative sources—at least for industrial quality alcohol—since the traditional substrates are also used as food. The most likely substrates are cellulose from agricultural and timber wastes, and methanol and various other organic wastes. Wood and plant wastes are the sources for cellulose. However, wood also contains the substance lignin, and hemicelluloses. There are some organisms that can be employed to convert cellulose into its carbohydrate monomers, which, in turn, may be fermented into alcohol. It would be useful to combine the enzymes for both operations within one organism—say, an yeast, a *Zymomonas* sp. or a thermophilic bacterium, *Thermoanaerobacter ethanolicus.*

For converting methanol, which is available by simple synthesis from coal, use may be made of methanogenic bacteria, which possess the rare ability to use a simple one-carbon molecule in its metabolism. Imperial Chemical Industries (ICI) of UK were the first to develop a *Methylophilus methylotrophus* that had a better conversion ability than the native strains. The bacterium converts the substrate into a protein which is marketed as a single cell protein (SCP) used as animal feed.

The hemicellulose in wood contains one species in large quantities. This is a polysaccharide xylan, the xylose monomer of which can be converted to xylulose, a sugar that can be fermented into ethanol. Attempts to introduce the gene for conversion to xylulose and then into yeasts is one more direction which will provide a way of utilizing the great biomass represented by wood.

Futher, *n*-butanol is a valuable solvent for a variety of materials such as plasticizers, resins and brake fluids. It had, at one time, been derived via fermentation of starches and sugars by *Clostridium* species. This route of production has almost disappeared as it is cheaper to make *n*-butanol from petroleum products. With the current need to use non-petroleum substrates, this older method may be reconsidered and, with the help of genetic engineering may increase the tolerance of the bacterium (*Clostridium acetobutylicum*) to its own product. Ethanol producers are not hampered by this sort of handicap. More research to understand and overcome this problem is required.

(ii) Enzymes The industrially used enzymes that have large markets are glucamylase, alpha-amylase, glucose is omerase and protease. Alpha-amylase is synthesized by

Bacillus subtilis. Recombinant DNA techniques have developed *B. subtilis* that synthesize an enzyme which is stable at higher temperatures. Fructose is used as a sweetener in the soft drink industries. The sugar sucrose, used mostly for sweetening has been almost replaced by corn syrup with a high fructose content, obtained by the conversion of corn starch by alpha-amylase, glucamylase and glucoseisomerase.

The synthetic polymer industries require polymerizing agents such as alkene oxides. These agents can now be made by the enzymatic conversion of alkanes. The enzymes are taken from certain bacteria and fungi. Once again, the natural synthetic programmes may be tailored genetically to provide better yields and more efficient catalytic activities. Some of the polymer industries that may use microbial conversion are those manufacturing thermoplastic and thermosetting resins (polyethylene, polyvinyl alcohol, polystyrene, epoxy resins, polyesters, urea, and so on), synthetic rubbers (polybutadiene, styrene-butadiene, polyisoprene and others) and the synthetic fibres (nylon polyester, acrylic, glass fibre, rayon, etc.) The Cetus Corporation is engaged in exploiting many of the above potentialities, especially as the microbiological production method offers several advantages over those employed by the chemical industries.

Another class of enzymes that has created a market of its own is that which is essential in the armoury of recombinant DNA technology. Restriction enzymes are produced in nature by bacteria partly to protect themselves from foreign DNA abounding in their ecological niches. Polymerases are of course utilized for synthesizing the nucleic acids, each enzyme being structured appropriately for the type of DNA or RNA synthesis that it has to mediate. Similarly, ligases join up nicked or new ends of DNA so that the sugar-phosphate backbone becomes continuous or uninterrupted. Already, natural producers of some of these enzymes have been genetically engineered to step up their production figures. Introduction of multiple copies of the phage T4 ligase gene has resulted in cells with enhanced ligase synthesis.

Restriction enzymes are not only used for cutting DNA that are to be spliced but also used as diagnostic reagents. It has been mentioned elsewhere how RFLPs (restriction fragment length polymorphisms) are used as powerful tools for the diagnosis of inherited genetic disorders.

Indeed, the enzymes on which recombinant DNA technology is based are the ones that have revolutionized biological investigations and applications. This use of rDNA to reveal more about the DNA itself reminds the author of a common Bengali saying: 'to use the sacred Ganges water for offerings to the Goddess of the Ganges herself!'

(iii) Organic acids Organic acids now produced by chemical industries that may profitably be manufactured by biotechnological methods include acetic acid, lactic acid and citric acid. These can be derived via microbiological fermentations of glucose, cellulose and molasses. Future considerations of scarcity of present raw materials for such organic acids may make it obligatory to use biotechnological routes of production.

(iv) Amino acids All 20 amino acids are in demand as they are used both in pharmaceutical products as food supplements and, of course, in research. Lysine and methionine, not made available by degradation in the digestive tract of herbivores, are supplied as supplements to animal and poultry feed. Glutamic acid is the raw material for the Japanese food flavouring agent, monosodium glutamate (Accent, Ajinomoto, MSG etc.). Today, the production of L-lysine by microbiological synthesis is competing with that produced by chemical industries.

(v) Fertilizers The largest class of fertilizers comprises nitrogen fertilizers. With the exception of a few species all plants derive their nitrogen supply in the form of reduced nitrogen. The exceptional species are prokaryotes that can convert gaseous nitrogen into ammonia, nitrites and nitrates. While agricultural scientists are trying to understand the mechanism of nitrogen fixation by these microorganisms, with the intention of making the plants nitrogen fixers themselves, there are others who have made genetically engineered yeasts become nitrogen fixers. In such a case, production of nitrogen salts by fermentation would bypass many of the disadvantages of present chemical production methods—provided the method can successfully compete with fertilizer plants—to supply the fertilizer needs of the world.

(vi) Pesticides Fungi, insects, weeds and rodents are the recognized pests that destroy agricultural produce during production and post-harvest storage periods. Chemical pesticides have stemmed the onslaughts on man's production efforts very effectively. However, continuous use has resulted in accumulations of non-degradable pesticides, that ultimately find their way into the food chains. The deleterious effects of the rise in the level of pesticides in food and, eventually, in the tissues of human beings have led to frantic efforts to both degrade unused chemicals as well as identify biological means for controlling the pests.

Biological pest controllers have 'enjoyed' anonymity for eons. They have now become interesting to human beings as non-chemical killers of insect pests.

Almost all pests are harassed by pests of their own. The strategy for curbing the former consists of encouraging the potency for killing of the latter. Luckily, most of these insect killers or *biociders* are harmless to human beings and animals.

The baculoviruses which infect lepidopteran insects may be genetically engineered to deliver toxins that kill the host insect. Baculovirus DNA may also be enlisted to synthesize pheromones of insects. Pheromones have been utilized for enticing insects into traps where they are killed. Insect-repelling hormones may be synthesized microbiologically, once the complex synthetic strategies of these substances are better understood. In fact, the plants to be protected may be fortified by the genes for these pest repellents and not depend for their safety on externally applied biociders.

Three bacterial biociders have been developed and marketed. These are three species of *Bacillus—B. thuringiensis, B. popillae* and *B. moritai*—that synthesize substances toxic to insects. The bacteria are cultivated and their spores used for formulations that are applied to the plants in the field in an appropriate manner. *B. thuringiensis* and *B. moritai* can be cultured by standard microbial fermentation. *B. popillae* requires larvae of the Japanese beetle for propagation.

Already the capabilities of *B. thuringiensis* have been exploited for protection of maize crops. The gene for the insecticidal toxin is present on a plasmid in the bacterium. This plasmid has been successfully transferred to *Pseudomonas fluorescence*—a bacterial species that infests and destroys roots of the maize plant.

There is a very wide scope for genetically engineering such biociders not only for increasing their insecticidal potency but also for widening the spectrum of host species that they can infect. This would require the successful introduction of genes for several toxins, their correct expression in the bacterial cell, or selective expression of a particular one of this mixture of genes in specific host plants.

3. Speciality Chemicals

Speciality chemical industries produce chemicals which are used usually in food and food-processing operations. These chemicals do not include agricultural produce, that is consumed directly, but food that has been obtained from conventional and unconventional sources or processed by microorganisms or their products. Such chemicals include food additives, flavours, quality enhancers, retarders of deterioration, as well as protein-rich whole cells that may be used as food and feed supplement. The last item is usually referred to as single cell protein or SCP. Speciality chemicals demand a certain degree of purity and non-contamination which is more than that required for commodity chemicals but far less than the norms for drugs and pharmaceuticals. Their prices, therefore, lie a little above those of the cheaper commodity chemicals.

The additives and process mediators generally used are fatty acids, amino acids, steroids, lipids, enzymes and vitamins.

Microorganisms have perhaps been exploited the earliest in the preservation of perishable food. The most important of these is milk. The excess milk produce of agricultural communities has been saved from being spoilt by different fermentation methods in various countries.

(i) Milk products In the Middle East countries and India, the chief mode of milk-conversion has been the preparation of yogurt (*dahi*), sour cream and several fermented skimmed milk beverages. *Lactobacillus bulgaricus* and *Streptomyces thermophilius* are used to convert the milk sugar lactose to lactic acid. In India, butter is churned out from yogurt and the buttermilk utilized as a beverage.

Western societies convert milk to cheese that has a longer shelf life than yogurt and sour creams. The milk is first separated with rennet or microbial enzymes, and the soft casein processed further to give soft or hard cheese. The empirical choice of fermenting fungal species has resulted in the creation of cheese varieties that are prized internationally and are the pride of the nations producing them. Hard cheeses may be innoculated with different organisms and treated initially to lower the water content. Gruyere type cheeses require processing by *Propionbacterium shermanii,* which synthesizes propionic acid that gives a special aroma and texture to the cheese. Blue cheeses, Roqueforts, Gorgonzolas, and Stiltons are made with additions of *Penicillium roquefortii* to the curdled milk. Swiss cheese uses *Lactobacillus helveticus.* Individual protocols of processing are responsible for the large variety of cheeses made in Europe.

(ii) Vegetable products Europe has also preserved vegetables by processing them through a sequential series of concentrations of common salt solution (brine), each of which favours the growth of a particular type of fermenter (lactic bacteria and yeasts, in general).

In the Far East countries, soybeans have been fermented to provide both protein-rich products and flavourings. These countries have also developed ways of enhancing protein contents of vegetables by allowing them to be overgrown, under controlled conditions, with edible fungi, such as *Aspergillus niger,* and using them as fried foods.

(iii) Baked foods Another prominent use of microorganisms with respect to food processing has been in bread making. Baker's yeast, the spores of which are ubiquitous in the environment, is added to dough made usually with wheat flour, sugar, salt and water. The water induces the breakdown of the starch with the amylase enzymes. These breakdown products (sugars) are fermented by the yeast, *Saccharomyces cerevisiae,* to alcohol and carbon dioxide. The gas 'raises' the dough, by getting collected in pockets within it, and the alcohol is eliminated during baking. The bread-making industries are highly successful ventures, in which the strain of yeast used is very important. The reasons for success rest on improvements both in baking technologies and in the production of correct yeast strains by separate organizations to meet the quality and quantity demands of industrialized bread-making.

(iv) Alcoholic beverages Yeasts are also the mainstay of alcoholic beverage industries. Starches, sugars, syrups from palm trees and sugarcane have been fermented into characteristic brews the world over. The closely guarded formulae or protocols for the fermentation make beers, wines, hard liquors, *sakes* (the Japanese drink) and so on, the prized products of different nations.

In all the above traditional food-processing ventures, improvements and alterations by the usual genetic methods may be augmented by recombinant DNA techniques. Indeed, modern nutrition-conscious nations demand novel qualities in the products, such as non-fattening beer. The production of dietetic beer by suitable genetic manipulations has already been demonstrated. The new variety, however, is yet not marketable since it has acquired an undesirable taste during the process of improvement. Speciality chemicals, therefore, require an extra rigour in production; they must conform to conservative taste and new demands for flavours.

(v) Nutritional additives Amino acids (lysine, methionine, aspartic acid, tryptophan and phenylalanine) and vitamins (especially the B2, B12, C and E) are speciality chemicals used mainly as nutritional supplements for both animal and human consumption.

(vi) Enzymes Enzymes are used to process raw materials or intermediate products in various food industries. Mention has already been made of the maize-degrading enzymes, the alpha-amylase and glucoamylase and glucose isomerase, which process maize starch , in a stepwise manner, to corn syrup that is very rich in fructose. High-fructose corn syrup has practically supplanted the use of other sweeteners in the soft drink industries. The success of continuous processing of glucose by immobilized

glucose isomerase has made corn syrup manufacturing the largest industry that uses immobilized enzymes.

Proteases are also utilized in a variety of industries. Alkaline protease is added routinely to detergents. Any or all of these enzymes are potentially alterable by design.

Lipids and, especially lipid breakdown products, are used in several industries. In the food-processing ones, they are used as polymer emulsifiers, as reagents for the floatation process of product separation, as well as directly as food additives. They are also in demand for the plastic, detergent, cosmetics, paints, varnishes, and other industries. Lipids that are saponifiable are the ones that contain fatty acids and glycerol. Fatty acids are now obtained by the chemical hydrolysis of the backbone of fatty acids, that is, a triglyceride. Enzymatic methods for this reaction would be preferable, for several reasons. Lipases that split a triglyceride into free fatty acids and glycerol are in use in certain Japanese firms, which use plant oils and other natural fats as the raw material. Conversion of fatty acids to fatty acid alcohols is also of commercial value, as the latter are used extensively as plasticizers in polymer industries.

(vii) Single cell proteins SCPs are whole cells that can be consumed as protein-rich food additives. These are generally single celled organisms (mostly yeast) that have been grown on otherwise non-usable material or biological wastes. These cells thus retrieve non-usable material to synthesize their own macromolecules, which are edible for both human beings and animals. In short, SCPs are microbial food.

SCPs were first exploited by the Germans during the two world wars to replenish their fast-dwindling food resources. Yeasts, grown on molasses, that were discards from the sugar industry, were rich in proteins, carbohydrates and vitamins. The yeasts were added to foods and feeds.

The next boost to the production of SCPs came with the discovery of bacteria that could convert petroleum fractions and products into the materials of the cell. Imperial Chemical Industries (ICI), UK, came out with 'Pruteen', an SCP product from the growth of *Methylophilus methylotrophus* on methanol, obtained by chemical conversion from methane. The sensation created by the possibility of converting hydrocarbon wastes into SCPs, however, decreased as there are today more competitive sources of cheap proteins (fish meal, soybeans). Production of SCPs using agricultural wastes and discards from paper pulp and timber industries, however, remains an attractive supplementary method, where other raw material and energy resources are scarce.

5.6.3 Protein Engineering

Protein engineering, one of the major offshoots of genetic engineering, deals with the designing of proteins with desired functions. The potentials for applications in biotechnology are obvious. Proteins may be modified from existing ones or designed *de novo* using the methods of (i) site-directed mutagenesis, (ii) gene fusion with useful parts of different proteins merged into one, or (iii) providing novel sequences for synthetic proteins. The latter may be produced biologically in cells, or

constructed chemically in the laboratory.

To achieve such goals, first of all it is necessary to know more about the biology of proteins themselves. We have to know, in particular, what makes a portion of the molecule competent to undertake a particular activity. The bulk of protein engineering research is engaged at present in probing the structural–functional rapport of the macromolecules.

1. Protein Folding

What do we already know about proteins? We do know that a protein consists of one or more polypeptide chains, the patterns of amino acids of which are determined by the sequences of codons in the corresponding DNA. Secondly, we know that a protein usually assumes secondary, tertiary and often quarternary structures. The secondary structure consists of α-helix coils, β-antiparallel sheets, and random coils or loops between and beyond these two structures.

The tertiary structure is due to folding of a polypeptide in a characteristic fashion for each kind, with the folds held in place by disulphide bridges that connect sulphur-containing amino acids (e.g., cysteine). The quarternary structure is the association of two or more polypeptides (of the same kind or different kinds) to form a functional protein (e.g., haemoglobin consists of two each of two types of polypeptides, $\alpha_2\beta_2$ or $\alpha_2\gamma_2$).

It has been further discovered that a folded protein consists of a hydrophobic core and a hydrophilic outer surface that interacts with the environment. To be able to interevene in a directed manner in the synthesis of a protein, more information is required about what makes different regions of a polypeptide assume a helical, sheetlike or loop configuration and about the properties of each of these structures that are essential for different functional roles of the protein.

It is common knowledge that a stretched out (denatured) polypeptide is meaningless in terms of activity. The protein is only active when folded into the proper tertiary configuration. We must, therefore, also know how the active sites of a protein are ensured positions on the outer surface of the folded molecule. In short, it is essential to know why a protein folds up the way it does in nature.

The cell, at least, has guidelines for folding a protein correctly, and it follows the dictates meticulously every time the same protein is synthesized. Unless we are also let in on this secret, there is no hope of designing a protein rationally. How does one proceed to extract this biological 'trade secret'?

Obviously, we have to glean the relevant information from proteins themselves. The DNA sequence of the gene encoding the protein provides data on the primary structure of the polypeptide. X-ray analysis of the protein crystals tells us about the distribution of atoms in the folded molecule. Circular dichroism analysis permits pinpointing of secondary structures of proteins in solution. Relevant information for protein folding lies hidden in these data.

Let us first consider the DNA-bound information. Thanks to the already discovered genetic code, we can decipher the amino acid pattern in the polypeptide. Does this sequence also provide hints about which regions should curl up into helices, become stacked as sheets or enjoy a random or loop configuration? Is there, in addition, a clue as to which residues should be joined by S-S bridges to allow

the protein to assume its tertiary structure? Studies to answer these questions are of paramount importance.

2. Amino Acid Sequence Folding

One way to find answers to the above questions is to determine how critical it is for a particular secondary structure to possess the amino acid sequence that it enjoys. What would be the effect, if any, on the integrity of the function if the amino acids are replaced, singly, or in sets of more than one? Site-directed single base or oligonucleotide-directed, site-specific mutagenesis could be used to answer this question.

A second method would be to study differences in homologous proteins, and in proteins with targeted modifications, in terms of the efficiency of performance of a particular function.

Whichever approach one favours, one has to have access to data on a large number of wild-type and mutant proteins, DNA sequences, X-ray crystallography and to data from other physical and biochemical studies that provide such information.

Comparison of protein sequences have revealed that (i) there are homologous proteins with the same activities, although they differ in details of their amino acid sequences, and (ii) unrelated amino acid sequences may form identical secondary structures. An example of the first statement is the family of hemoglobin proteins. In the large number of haemoglobins found in nature only a very few amino acids have been conserved at key places in the molecules. And yet, all haemoglobins perform by and large the same jobs. The second conclusion has been indicated by various studies. In one of them, the greater part of a natural toxin (mellitin of the honey bee venom) was replaced by synthetic stretches of amino acid sequences. Yet both natural and doctored sequences coil into identical α-helices. So, although the types of amino acids do determine the shape taken by the polypeptide, it is not obligatory to use the same residue sequence for that shape.

3. Folding is Context-dependent

The above evidence suggests that unlike the genetic code system, which appears to be context-independent, the folding of proteins into secondary and tertiary structures is context-dependent. This means, that while, say, AUG or UUU always specify, respectively, methionine and phenylalanine, irrespective of its place of occurrence in the mRNA, a set of amino acids found in a particular α-helix may not form one in another region of the protein. It is suspected that there are landmarks in the coding region that influence the formation of a secondary structure in its neighbourhood.

Indeed, Jonathan King, of the Massachusetts Institute of Technology, USA, has discovered mutants of a bacteriophage coat protein that do not fold correctly unless certain sequence motifs are reinstated. The *Salmonella* phage P22 has a tail-spike protein that is involved in the attachment of the virus to the host cell surface. This protein is a trimmer, with three identical polypeptides associated into the mature protein. Each polypeptide assumes a partially folded conformation as it is translated. The polypeptides associate three at a time and complete the remaining folding operations. King has found evidence that the tail-spike folding (TSF) mutant

polypeptides do not fold correctly during the intermediate steps of maturation of the protein. This inability appears to be due to mutation in certain consensus sequence motifs in the coding region that possibly signal folding in its neighbourhood. Mutations in bases outside these motifs do not result in TSF mutants. It would appear that these motifs mark the places in the protein that undergo folding or are involved in protein-protein associations. The question surfaces: Is there another yet-to-be discovered code that underlies protein folding?

4. Model Making and Computer Graphics

Let us now examine what information may be gleaned from physical studies of protein structures. High resolution X-ray crystallographic data tell us about the distribution of the atoms in space. With much painstaking labour and perseverance, the three-dimensional pictures of a few proteins have been deciphered from X-ray analysis. This is hardly the quickest way to visualize the shapes of proteins in a routine manner from sub-molecular coordinates.

How then would one construct the shape of an unknown protein? The most feasible way at present is to look for a homologous protein, for which such a structure has been worked out. Since homologous proteins have been shown to possess roughly the same shape and pattern of packing of coils and sheets in the core, a tentative structure may be initially produced for an unknown protein from this known structure.

Model making is a very useful (sometimes the only) way of proceeding to understand a problem. Model making in protein research has been revolutionized by computer graphics. Given the proper parameters—even when insufficient—the programme displays the likely shape of the molecule and the arrangement of the secondary structures in the mature protein. Brown et al. (1979) were the first to produce a computer model for α-lactalbumin, on the basis of an existing model for a homologous protein, i.e., the enzyme lysozyme.

Similarity of motifs has helped in computer modelling of insulin-like growth factors (ILGFs), on the basis of a consensus model of insulins from mammals and invertebrates. ILGFs and insulins possess identical cores. To build the model for a sample ILGF, this core is first used as a scaffold to which peripheral sections are added. Even in proteins with non-identical cores, but with homologous activities, there is a great similarity in the structure of the outer loop regions—or seats of activity.

5. Intermolecular Relationships

Since the activity of a protein involves intermolecular relationships with a matching substrate, studies of structure–function correlations include the examination of complementarity at molecular interfaces. In other words, if the function of, say, an enzyme is to be determined, the interface between the test protein and its substrate has to be scrutinized.

Clues for the rules of natural protein engineering are, therefore, also sought among different cases of protein–substrate interaction. These clues include nucleotide–DNA associations, and those with hormones and other biologically active peptides,

antibody-antigen reactions, membrane proteins that transduce external signals into the cell interior, or those that act as pumping devices for entry and exit of ions, protons and even of polypeptides, signal peptides that allow the export of proteins from the cytoplasm, toxins that penetrate plasma membranes, and other surface-acting proteins.

6. Stability and Structure

Another type of protein is studied to understand the features that make a protein stable under conditions of unusual temperature and pH. Natural proteins of this type are found in members of a primitive bacterial group, i.e., the *Archebacteria.* Some of these are tolerant of high salt concentrations in the medium; some thrive in hot spring waters. Both halophilic (salt-loving) and thermophilic bacteria were once mere curiosities to the biologist. Today, they are being coaxed to yield their knowhow for synthesizing proteins that can withstand these abnormal environmental conditions. One of the hurdles in the way of successful proliferation of biotechnological industries is the inability of bacteria, microorganisms and other cells to survive and perform well outside a very limited range of temperature. There would be several advantages in having cells and/or enzymes that can be used at higher temperatures preferred for the chemical engineering operations in a biotechnology. It is hoped that pointers from thermophilic bacterial proteins will provide the expertise to design 'thermophilic' enzymes for industrial use.

7. Toxins and Mechanism of Action

The structure of a toxin and the mechanism of its action are of interest to protein engineers for two reasons. First, these are small proteins showing spectacular alterations in conformation, in tune with the activities to be undertaken. Secondly, there is the possibility of using toxins conjugated to antibody proteins to deliver drugs to specific sites in the body. Such immunotoxins have therapeutic applications in cancer management. To make such drug delivery feasible *in vivo,* much has to be learnt about how the toxin operates, how it may be dissociated from the carrier protein at the proper target and not react with healthy tissues and so on. A great favourite for such studies is the diphtheria toxin, that consists of two polypeptides (A and B), only one of which is internalized in the cell by endocytosis. The internalized vesicle containing the toxin fuses with an endosome with an acidic interior. Under this low condition of pH, the toxin is separated from its receptor and becomes anchored on the membrane bounding the endosome. The details of this process at the molecular level are intriguing, an understanding of which will provide us with enough wherewithal to design toxins for therapy use as well as enable us to optimize the antibody linker and toxin for maximum *in vivo* activity.

Another area of protein research that has burgeoned over the past decade and a half is the study of antigen–antibody interactions. To a protein engineer it is worthwhile to clarify the principle of immune recognition of proteins. With this aim in view, M Z Atassi of the Baylor College of Medicine, Houston, USA, developed a unique method to study the intermolecular surface interactions. This method, named 'surface-simulation-synthesis', consists of simulating antigenic regions (or

any other active region) on a protein, and using this as a mimic for the protein surface structure to be studied. Atassi used lysozyme, which has the three outer loops clustered on the protein surface that act as the antigenic sites. This cluster was mimicked by joining synthetic peptides representing the loops with appropriate lengths of peptides. These peptides (SP) were used as antigens to raise antibodies (A-SP). The region of the antibody that was bound to SP was also mimicked in a synthetic peptide (Ab-SP). The latter was used to raise antibodies. When these last mentioned antibodies (Ig-Ab-SP) were mixed with antibodies to the synthetic antigenic sites (A-SP), complexes, which were similar to the natural lysozyme, antigen antibody were formed. This indicates that both A-SP and Ig-Ab-SP were good mimics of the antigenic sites of lysozyme and of the antilysozyme antibody variable regions bound to them. This technique may be used profitably to study other protein–protein interacting surfaces (see Fig. 5.22).

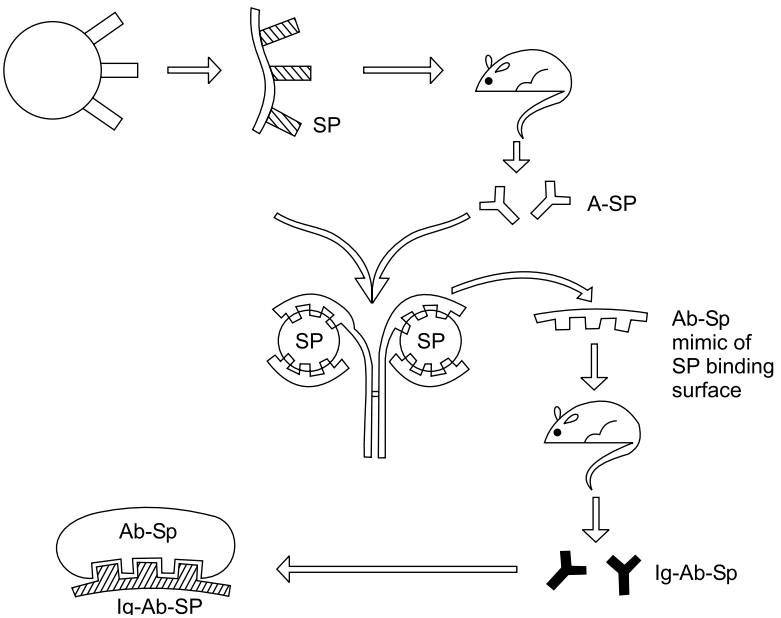


Fig. 5.22 *Synthesis of a protein that mimics the activity of the enzyme lysozyme.*

8. Site-directed Mutagenesis Used for Protein Studies

Once the structure of the protein is determined, it is necessary to find out the basis for structural–functional relationships. The main tool for such analyses is site-directed mutagenesis. This technique can replace single bases in a DNA, as well as blocks of bases. The latter technique is known as 'oligonucleotide-directed site-specific mutagenesis', and has been developed by Hargovind Khorana at MIT. Khorana has studied the protein bacteriorhodopsin (BR) that performs certain intriguing functions in the presence of light. Synthetic polypeptides of BR were made with provisions for restriction enzyme cutting at specific distances. The section between two such sites was replaced by synthetic oligonucleotides of the same length but

with different base alterations. The doctored genes were expressed and the activities characteristic for BR were tested. Thus, piece by piece, a protein may be scanned for its activity, and the roles of the nucleotides in a piece showing activity may be ascertained by the replacement of each base in the oligonucleotide.

9. Use of Mutated Primers

Another technique for producing random mutations in a protein sequence uses mutants of a primer, rather than site-directed, or oligonucleotide-directed mutations. Vershon et al., also at MIT, constructed a vector with a repressor gene (*Arc*) of phage P22, fused to a hybrid *trp-lac* promoter (P_{tac}); mutated copies of this DNA were recovered by providing mutated primers for DNA replication. A variety of transition and transversion mutations thus obtained were then tested for *Arc* repressor activity. The *Arc* protein represses expression of the chloramphenicol-resistant gene (*cat*). Mutants of the *Arc* gene can, therefore, be selected on media supplemented with chloramphenicol. This rapid method for isolating and sequencing defective versions of a gene would be useful in studies that enquire whether there are potential sites for mutations of protein-folding manoeuveres.

So much for the types of investigations that are aimed primarily to elucidate the structure–function principle that makes proteins fold correctly and perform in a characteristic manner. Let us now take a brief look at a few areas where applications of protein engineering are expected to be concentrated in the near future.

There is no doubt that for rational designing of proteins, maximum data about them would be ideal. But like all aspects of engineering, which has been called 'the art of the probable in the face of insufficient data', genetic engineering has not waited for scientists to harvest all possible information before launching on ventures to produce useful designer proteins.

10. Designer Enzymes

Designer enzymes will perhaps be the largest and most immediate targets for use in industries. At present, only a few enzymes out of several hundred characterized ones are used as catalysts in industry. The chief reasons for this lacuna are the instability and lack of consistent behaviour of biological catalysts in unusual non-biological environments common in industrial reactors and down-processing equipment. Investigations of structure and function of proteins of thermophilic bacteria should help in this area. Tairo Oshima, of the Tokyo Institute of Technology, Japan, has analysed the molecular basis of thermophily in *Thermophilus thermophilus,* that thrives in very hot water. Certain intrinsic and biochemical post-translational modifications that promote heat stability of proteins are indicated by these studies, and may be further probed and eventually exploited to modify existing heat-labile protein to make them stable at higher temperatures preferred for chemical engineering operations.

(i) Semisynthetic enzymes Semisynthetic methods for modification of enzymes for industrial use have been attempted by E T Kaiser, of the Rockefeller University, New York City, USA. Kaiser's group has been engaged mainly in the chemical modification of existing enzymes. They have substituted functional groups with

analogues of co-enzymes and have been successful in altering the catalytic behaviour as well as the stereospecificity of the enzyme. For instance, the hydrolytic enzyme papain was converted to an oxidoreductase. This approach together with the use of site-directed mutagenesis would allow tailoring of enzymes in a desired manner.

(ii) Enzymes to be used in organic solvents Industry would also welcome enzymes that can be used in a non-aqueous mileu, since many industrial operations use organic solvents. Since enzymes act in aqueous solutions, it had been tacitly assumed that the entire medium should also be water. In fact, recent work shows that only a film of water molecules is required to maintain the proper undenatured form of the protein molecule. Studies were, therefore, undertaken to see what happened when minimally water-coated enzymes were mixed with the substrates in a water-immiscible solvent. A M Klibanov's group at MIT, USA, was the first group to attempt such studies. They demonstrated that a large number of enzymes were fully active in the non-conventional environment provided they were coated with water. Indeed, these exercises led to the discovery of several unusual properties of enzymes working in organic solvents in an almost anhydrous state. To mention one such trait: the catalysts become very stable at high temperatures, a quality, as mentioned earlier, that will endear such enzymes to industrial engineers. In addition, several enzymatic activities occur readily in organic solvents. For instance, Klibanov demonstrated that pancreatic lipase catalysed a reaction between butyric acid and heptanol with close to cent per cent conversion in organic solvents, and less than 0.1 per cent in aqueous media.

(iii) Selectivity in synthesis Enzymes may be exploited for industrial use for their quality of selectivity. Very often chemical methods are incapable of producing only one type of isomer of a required product. Enzymes are meticulously choosy, and can use one isomer in a racemic mixture as a substrate and ignore the other. The utility of such enzymes for application on an industrial scale is obvious. Two examples of the use of enzymes, rather than the chemical method of synthesis, that have already entered the repertoire of industry are (1) the synthesis of a peptide sweetener, aspartame, from phenylalanine, and (2) aspartic acid with the help of a protease thermolysine. Normally, proteases are associated in our minds with a proteolytic activity. However, many enzymes, such as the one mentioned earlier, also mediate a reverse reaction; in this case, that of peptide synthesis. An understanding of how these separate activities are performed, by the same protein, will permit redesigning of enzymes to primarily catalyse synthetic reactions.

(iv) Abenzymes and other antibody-protein conjugates Both enzymes and antibodies bind themselves specifically to chemical groups in substrates. The enzyme, in addition, causes catalysis of a particular reaction involving the substrate(s). The question that arises is: would it be possible to raise antibodies that also function as enzymes?

The answer to the preceding question is in the positive. The basis for the answer is the proposal of Linus Pauling that enzymes bind themselves, not to the primary substrate, but to an activated intermediate or transition state. Jenks proposed that it should be possible to use the transition-state molecule as an antigen to raise a

matching antibody. The latter should be identical or quite similar to the binding site of the enzyme that mediates the catalysis of the transition state to the next product. Massey and associates at Igen, Inc. (Rockville, Maryland, USA) demonstrated the feasibility of this strategy. Pollack, Jacobs and Shultz in one laboratory; Tramontano, Janda and Larner in a second; and Jacobs, Shultz, Sugasawara and Powell in a third, among others, have been engaged in raising antibodies that simulate the structure of the binding site of certain hydrolases. These antibody-cum-enzymes, or abenzymes, enhance rates of reactions as much as 1000 fold of those of uncatalysed reactions. They are, however less efficient than naturally occurring enzymes (e.g., chymotrypsin increases the rate of a reaction 3000 fold).

Abenzymes possess the other basic features of enzymes. Powell's group raised abenzymes against an analogue of a transition state molecule in the hydrolysis of methyl 4-nitrophenyl carbonate. The analogue—a tetrahedral molecule of 4-nitrophenyl phosphonate—induces the production of antibodies that mediate the hydrolysis of a carbonate substrate.

Abenzymes or catalytic antibodies offer an easy route for preparing enzymes, the nature of whose active sites is not known. By using the transition-state product as the hapten of an immunogen, the test animal effortlessly produces the required matching proteins. The potential for such designer proteins is unlimited.

One profitable use would be to create abenzymes that synthesize and hydrolyze peptide bonds. It would then be possible to develop abenzymes which like restriction enzymes, will cleave and ligate polypeptide chains as desired. The scope for such protein engineering is obvious.

Attempts are being made to produce abenzymes that mediate synthesis of antibiotics, one of the most important of pharmaceutical products. The pathways of synthesis of antibiotics are mostly unknown, especially with reference to the enzymes mediating the reactions in the pathways. It may be easier to isolate transition state intermediates to be used as haptens than deciphering the biochemistry of the synthetic pathway. An active abenzyme has been prepared by Neuberger et al. of Medical Research Council (MRC) laboratory at Cambridge, UK, by fusing the Fab region of an antibody to that of a nuclease of *Streptomyces aureus.*

(v) Novel antibiotics An area in pharmaceutical industries which will benefit by protein engineering is that of the production of novel antibiotics. To date, antibiotics are produced microbiologically, and further modified chemically or enzymatically where required. An increase in production yields and the introduction of new antibiotics may be brought about by protein engineering. The goal of research in antibiotics has been to identify new compounds with clinical properties superior to the existing ones. Post-synthetic, modifications to improve the features of natural products are fraught with difficulties, which may perhaps be alleviated or removed by protein engineering. There is a need for such novel products, especially for catalysing the biosynthesis of penicillin and cephalosporins, which belong to the lactam group of antibiotics. Usually, antibiotics are produced as secondary metabolites at the end of a complex network of biosynthetic reactions. Several genes are involved for the conversion of substrates to the end product. Only recently, has it been possible to clone the cluster of genes producing an antibiotic from *Streptomyces.* Individual genes, such as that for isopenicillin N-synthetase or cyclase (that converts a precursor

of cephalosporin C to isopenicillin N) have been cloned successfully. Elucidation of the genetics of antibiotics synthesis together with gene manipulation techniques will provide the basis for protein engineering in this field.

5.6.4 Clinical Applications

Genetic engineering techniques have greatly improved the possibilities of applications in the construction of drugs and other products used in medicine. Here we will restrict ourselves to direct use of rDNA-transformed cells in clinical practice. Two areas have received more attention than others. These are the gene replacement therapy and the transformation of whole animals. Theoretically, it is not impossible to replace a defective gene that is responsible for a genetic disease. However, several problems have to be solved before such replacement may be practised as a clinical intervention. Perhaps, it may be easier in the near future to implant genetically altered or healthy foreign cells on a suitable matrix within the body. It is expected that the gene products from the introduced cells will compensate for the native deficiency or will counteract its defect. Such whole animal transformation has been reasonably successful on animal models. Extensive research has still to be done about the manner in which this can be made a routine procedure that will not harm the body in any manner.

1. Gene Replacement Therapy

Gene replacement therapy is the correction of an inherited disorder by insertion of the correct version of the involved gene. Successful transfer and expression of exogenous genes in mouse embryos and whole mice have offered the possibility of intervention in diseases due to single genes or a cluster of genes that are also expressed in a single tissue. At least these diseases appear to be the first candidates for attempts to replace defective genes.

Several clinical and surgical constraints prevent gene replacement therapy in human beings. Such constraints are expected to be overcome in time. The disease chosen for intervention must first be cured successfully in an animal model. There are searches, therefore, for disorders of human beings that find a close similarity with a similar disorder in an animal.

One of the obvious difficulties is our present lack of control of the site(s) of integration of the introduced gene. First, insertion in inappropriate regions, including within the coding or regulatory sequences of another gene, is known to result in adverse affects which may be obvious immediately or become evident at a later period. Secondly, the expression of the defective gene must be limited to readily transplantable tissue. The basic procedure would be to transform cells from such a tissue e*x vivo* and reintroduce (transplant) it into the diseased body. The only two tissues that appear to be suitable for this purpose, at present, are blood and liver tissues. The disease chosen for therapy must also be such that one gene transfer transplantation will be enough to favourably influence the effect of the defective gene throughout the body.

The above constraints delineate the domains that may be explored for this type of therapy. Disorders due to defective expression of stem cells in bone marrow and in

liver cells are the ones that have, therefore, received much attention. Several anemias and thalassemias are included in the list of disorders of blood cells and liver cells. There are already sufficient indications that bone marrow and liver transplantations may become practical in the near future.

In some diseases, the expression of the defective gene occurs in one tissue and the pathological alterations in others. In other diseases, the sites of expression and disorder are the same. An example of the first type is the deficiency of LDC (low density lipoprotein) receptors in liver cells—a condition that underlies atherosclerosis and several coronary diseases. A genetically treated liver cell with the correct receptor gene, should remove the disorder. An example of the second type is the deficiency of a mucopolysaccharide, aryl sulphate B (ASB), that results in alterations in skeletal and connective tissues. Krivit et al. (1984) have stemmed the disorder in an animal model by transplanting bone marrow stem cells carrying the correct ASB gene.

Success has been recorded in a limited number of bone marrow transplants with genetically doctored cells. These include immunological and hematopoietical disorders that originate in defective bone-marrow cells. The failures are due to technicalities that include insufficient HLA-matching (tissue-typing) and incomplete elimination of mature T cells from the transplanted cells.

The use of cyclosporin A (Iwatsuki et al., 1985) has contributed to some success in liver tissue transplants. Better success is expected with improvements in surgical, clinical and post-operative management strategies.

Exploration of gene replacement therapy is taking place at several levels. One of them is concerned with the best vector for introducing the gene into the maximum number of cells. Retroviral vectors and BPV-1 episomal vectors appear to be the first two choices, with the first claimed to be a better candidate. Injection of retrovirus-carried genes into the marrow cavity is expected to automatically introduce the genes into almost every cell in that region, and the infection would tend to be a continuing one. This route is rather attractive as it dispenses with troublesome tissue-matching and *ex vivo* transformation steps. The disadvantage is the unpredictability of the site(s) of insertion in the DNA. Transgenic mice, though expressing the particular introduced gene, are known to harbour several mutations—an outcome that has to be avoided in clinical interventions.

The genetic disorder which is likely to be the first candidate for successful gene therapy is ADA (adenosine deaminase) deficiency. Deficiency of ADA leads to SCID (severe combined immunodeficiency), a disorder due to the accumulation of toxic metabolites that include dATP. This disorder affects both T and B cells which are rapidly eliminated. The gene for ADA has been cloned from both mouse and human beings and its structure is under investigation. This gene has been successfully introduced into target mouse cells. Only a low level of expression of the correct gene appears to be sufficient for correcting the defect, and bone marrow transplants have shown their corrective nature in the animal. As mentioned earlier, gene therapy with defective retrovirus-carried human/mouse ADA gene in bone marrow precursor cell transplants require no immunosuppressive strategies and do not precipitate graft-versus-host diseases in the treated animal. Again, the disadvantages are generation of insertional mutations and of active retroviruses through recombination of the vector with endogenous viruses. In addition, the effect of the introduced gene wears out in time. A more permanent expression must be ensured.

These crinkles and others have to be smoothened out before the gene replacement therapy becomes a reality. The stage is set with cloned genes available for a wide spectrum of blood- and liver-based diseases. These diseases include haemoglobinopathies, complement disorders, thalassemias, diseases due to deficiency of lysosomal enzymes, phenylketonuria, familial hypercholesterolemia and haemophilia. Several of the foregoing and other genetic defects can be recognized in carriers by the analysis of DNA sequences and restriction fragment length polymorphisms (RFLFs). If the ADA model works out, the way will be clear for attacking other genetic disorders.

It must be noted at present that genetic intervention is visualized only for curing affected individuals by transplantation of somatic cells. Transgenesis may not be contemplated until a foolproof method is developed to allow genes to be targeted without supplementary unfavourable consequences. The ethical and legal aspects of such intervention are also under consideration.

2. Whole Animal Transformation

Gene transfer into whole animals will effectively supply the gene product throughout the life of the recipient only if the gene is introduced into cells that divide throughout the life of the animal. Initial attempts using retroviral vectors to infect stem cells in bone marrow have been successful, but the percentage of correct expression or any expression of the gene is quite low.

To overcome the preceding difficulty, a novel strategy has been introduced by Bell and associates (1979) and developed further in the laboratory of Inder M. Verma (1985). Retroviral vectors carrying the gene, to be established in the animal, were introduced into fibroblast cells; the marker for the vector was a neomycin gene. The fibroblasts were attached to a matrix of collagen and the patch of tissue thus formed implanted under the skin of the recipient (mouse). The gene of interest was that for the human clotting factor IX. This was provided with a promoter of a mouse metallothionin gene. Successful expression of factor IX was demonstrated with the help of a monoclonal antibody for the factor. The level of expression suddenly shot up in the second and third weeks. This corresponded with an increase in growth of the collagen implant (Fig. 5.23).

The above method appears to be a promising one for introducing desired genes into a whole animal. It is definitely a simpler procedure than the introduction of transformed bone marrow cells. Studies are continuing to workout a usuable alternative that will be first tested on an animal model. At least a way has been indicated for treating individuals in their lifetime with transformed cells that will continue their growth in the recipient body.

5.6.5 Control of Environmental Pollution

Natural wastes from animals, plants, humans and petroleum products have been present in the soil and water for ages. To these wastes has been added toxic and carcinogenic chemicals that are by-products of industrial development. These chemicals have posed a serious problem of environmental pollution in developed as well as developing countries.

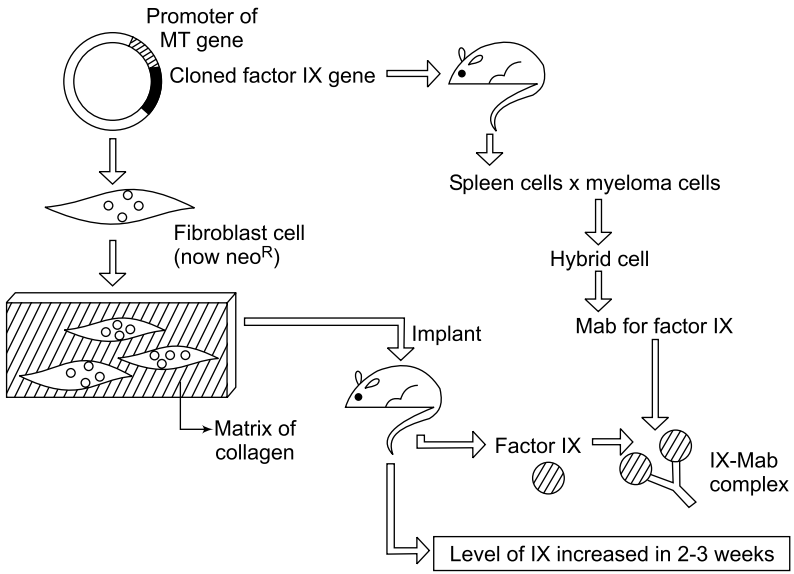


Fig. 5.23 *Whole animal transformation for gene-mediated therapy. Inder M Verma and associates implanted recombinant cells, containing the cloned gene for the clotting factor IX, on a collagen matrix, under the skin of mouse. The level of the factor was found to increase significantly in the recipient (IX) animal.*

Natural wastes and hydrocarbons have been degraded naturally—courtesy of a plethora of microorganisms—at a slow but steady rate. However, in areas steeped in petroleum products due to seepage from stored material or on the waters at sea and river ports blanketed with oil spills, the natural process is not fast enough to ensure total or even significant partial removal of pollutants. Man-made chemicals that are released into the environment include pesticides, herbicides, fire retardants, and dielectric fluids from capacitors and transformers. These substances have been designed in the first place to make them long-lasting, and hence are not readily destroyed. This is particularly true of pesticides and herbicides that are mostly halogenated aromatic compounds. Accumulated deposits of these chemicals eventually lead to their entry into the food chain and they get lodged in the fatty tissues of animals and human beings. Very often, storage and dumping also result in seepage into the ground waters present in the soil or/and into the rivers and lakes.

While there are bacteria that are capable of degrading one or more components of petroleum, there is no single species or strain that can tackle all these different fractions. Anondo Chakravarty, while at General Electric Corporation, was the first to attempt the development of such a 'paragon'. There are separate strains of *Pseudomonas putida* which carry genes for enzymes that mediate the degradation of xylenes, toluenes, octanes and camphors into carbon dioxide and water. These genes are borne on plasmids. Chakravarty was able, by general microbiological and genetic engineering techniques, to create a 'superbug' strain that contained all four sets of genes in its cells. These engineered bacteria rapidly degrade oil spills both in laboratory and field tests (see Fig. 5.24).

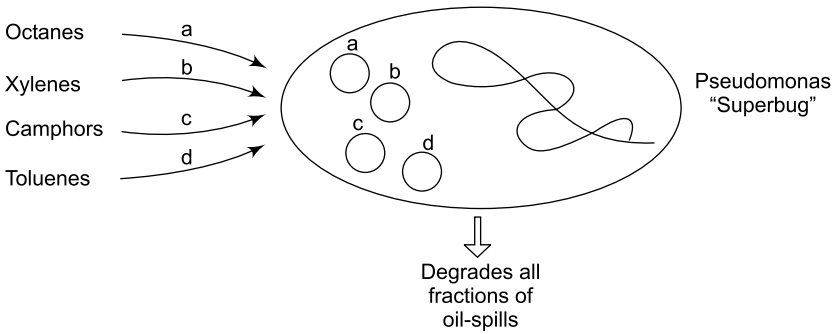


Fig. 5.24 *Anondo Chakravarty created a Pseudomonas 'superbug' strain that degrades all fractions of hydrocarbons in oil spills at harbours and river ports. The genes for enzymes that degrade camphors, xylenes, toluenes and octanes, CAM, XYL, TOL, and OCT respectively, were taken from strains that possess only one of these activities.*

Such success led Chakravarty and his associates to tackle the halogenated compounds that include DDT, 2,4-D and 2,4-T. A rigorous search for microflora that attack these compounds was made (and is continuing to be made). As there are very few chlorinated compounds in nature, it was not surprising to find only very few bacteria that matched them. Nevertheless, some bacteria were discovered that degraded simple chlorinated compounds completely, and others that mediated partial conversion to intermediates which required further dechlorination. These bacteria included strains of *Acinetobacter,* and *Alcaligenes, Arthrobacter* and *Pseudomonas* species.

Traditionally, biological waste treatment strategies involve allowing the natural microbial population to adjust to the presence of the toxic material. The strains that can metabolize it, even partially, eventually predominate. This is a sort of a chemostat system. A second strategy is to mix cultures of strains that degrade different intermediates in the degradation pathway. In most cases, this ability resides as genes on plasmids.

In a few cases, single strains have emerged from such mixed cultures that possess the plasmids from the different strains, and, like the petroleum-devouring 'superbug'; accomplish complete degradation by itself. This system is referred to as 'plasmid-assisted molecular breeding' (PAMB).

Chakravarty's laboratory at Illinois isolated such a strain, named AC1100 from a mixed culture that degraded 2, 4, 5-T (2, 4, 5-trichlorophenoxy acetic acid), one of the components of 'agent orange' that defoliates broad-leaved plants.

Simple chlorinated compounds, such as chlorobenzoic acids, are degraded by single strains. For example, 4-chlorophenoxy acetic acid is degraded fully by a species of *Pseudomonas.* Another species of the same bacterium degrades 3-chlorobenzoic and 4-chlorobenzoic acids. Polychlorinated biphenyls (PCBs) have their own degraders: *Alcaligenes* Y42 and *Acinetobacter* P6. The end products are chlorobenzoic acids. Also, 2, 4, 5-trichlorobiphenyl (2, 4, 5-T) is degraded to chlorobenzoic acids by the latter species.

A species of *Alicaligenes* degrades 2, 4-D (2, 4-dichlorophenoxy acetic acid) to an intermediate product 2, 4-dichlorophenol. The gene for this event is carried by

a plasmid pJPl of *Alicaligenes paradoxus.* The intermediate product is a substrate for certain microbial species. Similarly, p-chlorobiphenyl (4CB) is converted to intermediates that are still chlorinated. Dechlorination may be completed by certain mutant strains of a *Pseudomonas* species; the genes for this activity are on the plasmid pAC25. Chakravarty used a mixed culture of *Acinetobacter* P6 with a 4C base (p-chlorobenzoic acid, a product of 4CB breakdown) strain of *Pseudomonas putida* to totally degrade the compound 3, 5-dichlorobiphenyl (3, 5-DCB). Combinations of other species have been identified that also act in a similar manner.

The single strain AC1100, mentioned earlier, metabolizes 2,4,5-T very rapidly. While natural degradation of this compound is very slow (0.1 per cent in a period of 15 minutes to 96 hours), incubation with AC100 removes 6.43 per cent, 68.75 per cent and 89.57 per cent of the pollutant in 15 minutes, 48 hours and 96 hours, respectively.

Pollution by toxic chemicals may be controlled by either encouraging the growth of known naturally occurring degraders, or by increasing the degrading capacity in genetically engineered species. An example of the first approach is the use of microbes together with nitrogen and phosphate-containing nutrients, undertaken to remove petroleum seepages in Pennsylvania. The easily removable oil is floated on surfaces of water pumped through bores drilled into the soil. The nutrients, in solution, are then pumped in. The Sun Oil Company used this method to remove enormous quantities of petroleum that would otherwise be virtually irremovable under natural conditions. Several other USA based firms market microbes for removing pollutants from soil and water.

For smaller operations, such as cleaning up drains in restaurants and oil on aircraft carriers, dry bacterial mixtures are available commercially that destroy harmful material without releasing the obnoxious smelling hydrogen sulphide, a by-product of natural degradation or putrefaction of lipid compounds. Companies such as Polybac, Sybron Biochemicals and others regularly sell microbial degraders to several oil companies.

In this context, first, it may be mentioned that the introduction of microbes—especially engineered ones—is not universally accepted without qualms. Secondly, there is no absolute proof that the added genetically manipulated bacteria are able to compete successfully with the naturally occurring ones under field conditions. Nevertheless, the search for 'more efficient performers' and attempts to develop suitably tailored ones with enhanced degrading qualities are continuing in several laboratories. These researches also include investigations that will reveal the molecular pathways of degradation. One such research in the Universities of Gottingen (Germany) and Geneva (Switzerland) has resulted in the cloning of the gene for one of the enzymes in the degradative pathway for 2, 4, 5-T. An enhanced understanding of the processes involved will be valuable for designing better performers in the near future, it is hoped.

5.6.6 Recovery of Minerals

Certain bacteria, usually found in mining areas, are able to release minerals from the ores. *Thiobacillus ferrooxidans* is one such species that has been employed to leach

copper and uranium from low-grade ores. This bacterium is also able to leach nickel, lead, cobalt and zinc.

Microbial leaching is gaining prominence with the rapid depletion of high-grade ores and the rising cost of energy. The leaching bacteria utilize CO_2 from the atmosphere and generate energy by oxidizing inorganic salts; thus, no additional nutrients are required for the growth of the bacteria. An added advantage of biological leaching is its non-polluting operational mode. The time taken by this method is, of course, much longer than that taken by chemical or physical processes. However, the ease of leaching operations and the value of the minerals warrant additional methods for making the most of poor natural resources.

The leaching operation depends on two events: a transformation of a ferrous iron to a ferric iron and the use of the latter (strong oxidizing agent) to convert metal sulphides into salts that are soluble in acids. Some bacteria directly convert metal sulphides into sulphates as well as ferrous iron to the ferric state; again acid-soluble metal sulphates are obtained, as a result.

Keeping the potentialities of recombinant DNA technology in mind, attempts to improve the efficiency of these bacteria in carrying out the leaching operations are being made.

Wastes from mining operations usually contain trapped sulphur in the form of salts. Suitable bacteria are available for the recovery of pure sulphur from such waste. Coal usually contains an appreciable amount of sulphur, which is released as a pollutant (sulphur dioxide) during combustion. Leaching of sulphur from coal may reduce the polluting effect of burning coal—an operation that is being revived due to the declining resources of petroleum fuels.

Metals are recovered from the acid-soluble salts by standard physical and chemical processes. Genetic engineering of such bacteria may improve their capabilities in aspects that include

(a) an ability to withstand higher acidic conditions, as well as higher temperatures and pressures (for deep mining situations),
(b) an increase in their ability to convert ferrous to ferric iron, and
(c) development of insensitivity (or less sensitivity) to toxic metals such as cadmium, mercury, silver and thorium.

Recovery of petroleum from soils which have already yielded the major proportion of their oil is another area that attracts genetic engineers. The oil remaining after regular drilling and recovery may be too viscous to be pumped out or may be bound within rock formations. Microbial enhanced oil recovery (MEOR) has invited a search for appropriate microorganisms. Microbial species may be able to grow on the less valuable components of the oil and produce emulsifiers that decrease the viscosity of the trapped oil and thus facilitate its recovery by conventional means. There is especially a requirement, for microorganisms that can tolerate the higher temperatures, salinity and pressure in the deep-seated oil deposits. Some species that have been isolated from such sites, especially those that are anaerobic, and so do not require an additional carbon source for growth, are candidates for the enhancement of efficiencies by genetic manipulation.

Two microbially produced substances that reduce the viscosity of crude oil and emulsify it, respectively, are a glycolipid from bacterium H-13 and emulsan. The

latter is being marketed by Pfizer and has been used for cleaning oil tankers and runways of aircraft carriers. Although MEOR appears to be too speculative at present for R&D, several smaller American companies have initiated research in this area.

5.6.7 Novel Products: Biosensors and Biochips

Recombinant DNA technology has made it feasible to think in terms of entirely novel applications, and not merely in improvements in existing ones. Two of these applications are in the areas of biosensors and biochips.

Biosensors are biologically based sensing devices. Enzymes have been used for a long time as indicators of the presence and quantity of diverse organic substances. The reaction of the enzyme with the substrate is measured directly from the ensuing colour or temperature changes, with the help of sensitive electrodes. These devices have remained as esoteric ones chiefly due to the high cost of enzymes and non-availability of enzymes for many situations.

At present, enzymes may be produced in the required quantity (1) by cloning cells with naturally occurring genes, (2) from laboratory-designed ones, and (3) from performers improved by rDNA techniques. In addition, monoclonal antibodies that are highly specific for their substrates are also available for sensitive assays. There is, therefore, a renewed enthusiasm for developing cost-effective biosensors.

The enzyme-substrate reaction is recorded by a handy transistor in the form of electronic signals. Biosensors find use in medical diagnostics, in monitoring of bio-processing in biotechnological industries and in assessing the level of pollution in the environment, particularly in air and water. The routine application of biosensors will depend on lowering their cost, on making them portable and especially on rendering them tolerant of high temperatures. *Thermophilus* bacteria are being searched for enzymes that can withstand high temperatures, and also for the understanding of the molecular features that make proteins thermolabile.

A biochip is a futuristic application which uses a semiconductor in the form of an organic molecule that is stabilized in a jacket of proteins. Biochips are potential competitors of silicon chips. In fact, if successful, they will be cheaper, more energy-efficient and will act more rapidly than the conventional microchips of today. Enthusiasm to develop biochips has been generated with the possibility of obtaining rDNA-generated proteins that are tailored from computer-aided designs.

One of the areas of use of biochips that is obvious is in medical practice. Implants of biochips may be used to monitor blood flow, heart beat and brain functions, to name a few applications.

One of the main advantages of biochips over silicon ones would be in lowering the limit of the width of the circuit, which, in silicon chips, is governed by the wavelength of light required for indenting the circuit on the silicon wafer. As molecules of proteins are themselves the conductors, the width of the circuit will be that of these molecules. Secondly, molecular circuits can be packed closer to each other without what is known as the 'tunnelling' effect that occurs when silicon circuits are closer than a certain minimum distance. Lastly, conducting currents

through molecules would not generate heat, as happens in the case of silicon circuits. However, unique methods of fabrication have to be developed before biochips can become significant products in the market. Monoclonal antibodies are also hopeful candidates as material for making biochips. Extensive research is required also for understanding the relationship between the form, the structure and the function that makes a protein active in a particular manner.

5.7 ❑ SAFETY, ETHICS AND SOCIETAL ISSUES IN GENETIC ENGINEERING

Genetic engineering involves redesigning the basic blueprint of life in a deliberate manner. This has understandably given rise to grave concerns about indiscriminate release of recombinant DNA-containing organisms into the environment. The risks include harmful biological impact on gene pools and ecosystems, as well as that on global economy, social values and legal structures of our existing systems.

Most advances in science and technologies that have the potential of disturbing the integrity of health or the survival (or the social fabric) of the human society have spawned controversies regarding the extent to which society should tolerate further research that could have deleterious effects on the society itself. In no case have these fears been more turbulently expressed than in those which were awakened by recombinant DNA technology.

Understandably, the immediate worry revolved around biohazards due to the presence of chimaeric or synthetically reorganized organisms in the environment at large. It was soon apparent, though, that much more was at stake than immediate health and ecological imbalances. A technology that possesses the power to genetically redesign even human beings could be expected to have significant effects on ethical, moral, social, economic and legal values and structures.

Traditionally, the scientist has remained insulated from his/her society, pursuing 'truth' or 'knowledge' without concern about the effect of his/her discoveries on the human community. This insular attitude was first shaken when nuclear power became available for use. On the one hand were the fabulous benefits of harnessing this tremendous energy; while, on the other, were the nightmares evoked by the catastrophes of Hiroshima and Nagasaki. It was a group of genuinely concerned atomic scientists who initiated debates about the advisability of using nuclear power and desirability of setting up of regulations and institutions to ensure and oversee the safety measures imperative for handling nuclear power.

Recombinant DNA technology triggered off a far greater concern due to its unique characteristics of being able to pollute the biosphere in an irrevocable manner. This anxiety led to a substantial erosion of the traditional aloofness of the scientist ensconced in an ivory tower. The scientist could no longer remain unconcerned about the impact of his/her findings on the environment in general and mankind in particular. A group of socially conscious and eminent biologists was responsible for inducing a brainstorming session at Asilomar, USA, in 1974. The risks and ways of assessing them, if any, were discussed among scientists who were predominantly molecular biologists. The immediate agenda of the meeting pertained

to the hazards of inadvertent or deliberate release of chimaeric organisms into the environment. What were the types of dangers that could be anticipated? How could this eventuality be prevented? Opinions varied from one extreme to the other. The alarmists perceived dire irreversible effects and recommended a complete shutdown of all rDNA research, even if tantalizing projected benefits were in the offing. The opposite camp was more conservative and did not believe in impending doom. The members belonging to the second camp, cited especially the known fact that most redesigned organisms usually did not survive outside laboratory environments when pitted against the more hardy wild-type strains. According to this group, the benefits to society that could accrue from rDNA research far outweighed the speculative dangers that could be ushered in.

The Asilomar conference included a few non-experts who were also non-scientists. This was in response to a need felt by the experts for opinions of the lay public, who, after all, would be sharing the hazards, if any. Subsequently, committees in different cities of the USA and in certain campuses were organized which contained a proportion of representative lay public.

The immediate sequel to the Asilomar meeting was the creation of a set of regulations to be followed in rDNA research laboratories under the auspices of the National Institute of Health (NIH), Maryland, USA. The NIH safety measures were called guidelines, which could be mandatory for only those institutions funded by the NIH. Other countries evolved their own versions of NIH-like guidelines. Laws that would apply to all researchers could not be formulated as once again it was realized that, in doing so, one would have to seriously assess the tenets of our current legal values. Till such time as and when appropriate laws could be designed, guidelines would have to suffice for safety in ongoing research.

The debates and controversies have not as yet resolved themselves into unanimous or even consensus opinions regarding the future of rDNA research. However, the deliberations were not without value. They made it abundantly clear that we, as a society, have not learnt how to cope with a novel advance in knowledge that has potentialities of leading to both benefits and harm to our health and cultural norms. These debates, in addition, were responsible for the induction of a larger lay public in the deliberations to decide the fate or extent of rDNA research to be allowed by society. This new relationship between experts, who apparently know what they are talking about, and the lay public, which may or may not understand fully the implications of the new technology, has to be fostered and meaningful ways designed to make similar participation, in the future, effective and democratic. Such regulatory bodies will have to find algorithms that are flexible enough to keep pace with a rapidly changing scenario, not only for the case of recombinant DNA technology but also for other technologies in the future that may be expected to drastically influence mankind and the environment. A failure to do so, would mean a repetition of the exercises in an ad hoc manner for every such revolutionary advance in science and technology.

5.7.1 Biological Risks

Genetically engineered organisms, unlike toxic chemicals and other noxious pollutants, but like all other living creatures, can increase their populations and spread

far and wide both in space and time. This is the first characteristic of deleterious chimaeric microorganisms that causes consternation. It is not unreasonable to assume that a bacterial plasmid spliced to a viral genome, will be disseminated in nature at a faster rate than the naturally occurring bacterial species. If these hybrid DNA molecules carry potentially harmful genes (such as resistance to a miracle drug), tremendous dangers could be anticipated. The use of viral genomes that cause cell transformation and tumor formation was looked upon with added trepidation. Would an SV40 DNA accidentally integrated in that of a human lead to malignancy in the latter non-permissive host?

Biological risks may also accompany the release of engineered microorganisms that are judged on the basis of available data to be safe, but actually turn out to be hazardous. It is argued that splicing of DNA from species that do not exchange genetic matter in nature may trigger off abnormal expression or repression of genes in the heterologous host. We do not know, for instance, what a viral promoter may initiate, *in vivo* in an unnatural genetic setting. It is this unknown nature of biological risks which has taxed the minds of experts and the lay public alike.

The proponents of the new technology argue that most engineered microorganisms are not likely to survive in competition with their wild-type counterparts in the environment outside the laboratory or reactor. The trouble is, in the case of recombinant DNA technology, no method is available for making risk-benefit assessments.

A second type of biological hazard that is envisaged concerns the release of genetically engineered strains of agriculturally important species. The development of a paragon superproducer may lead to the exclusive use of this particular strain. Such a development would result in the loss of several other naturally occurring strains. Improvement in strains has depended till now on the availability of a variety of alternate alleles in nature. Diminishing this gene pool, by disuse, will leave behind only engineered genes, that may be wiped out at a future date by some unforeseen circumstances. The naturally occurring varieties, if present, would perhaps have included one or more that could have coped with the new environmental milieu. To counteract this hazard, banks of germ plasms of different species are being maintained or are planned to be maintained. In the long run, germ plasm banks cannot be the practical answer to dwindling strains. After all, maintenance of such banks is also vulnerable to a variety of accidents. To name one, a severe power failure may destroy an entire germ plasm bank in a short period of time.

A third area of biological risks involves the population of workers in industries utilizing or creating rDNA organisms. Inadvertent contamination of workers with a lowered resistance (such as those undergoing antibiotic treatment) may lead to unforeseen complications. Major organisations that belong to this category have begun to keep regular records of workers exposed to rDNA sources. All illnesses (even one-day absences) are checked before these workers are allowed to resume work. Some nations make such medical surveillance of rDNA-exposed workers mandatory, and supervise permission for research in university and industrial laboratories. This is not a novel strategy; similar precautions have been prescribed and ensured for work with toxic and radioactive material, as also with pathogenic organisms.

A biological hazard that one cannot afford to overlook is that of using engineered organisms for biological warfare. Undoubtedly, microbes are far more cost-effective as exterminators than even the deadliest of explosives. A biological weapon is one that will not be restricted, like nuclear ones, to nations that are sufficiently affluent and also possess the esoteric know how. Any nation can create an arsenal of microbial killers, or order, with a minimum of expertise and resources. One of the main deterrents, however, for using genetically engineered microorganisms for biological warfare is the current lack of methods required to confine the spread of the agents only to defined populations (the enemies). Besides, it is still cheaper and easier to use naturally available deadly organisms such as the food-poisoning bacterium (C. *botulinum*), the tetanus bacterium (C. *tetani*) or the encephalitis viruses, than undertake the expense and labour of creating chimaeric monsters.

The biological risk which is speculated to have a profounder influence is that on the ecosystem. The biosphere has evolved, by natural selection, species that live harmoniously and share equally profitably in the resources of the earth. There is worry that DNA created by splicing of genes from different species may disturb the biosphere in ways that are not apparent today. This may result in upsetting the delicate balance maintained in the natural ecosystem.

One argument against the view that deliberate splicing of unrelated DNA may affect the natural evolutionary rate and/or processes is the recognition of the fact that similar gene-splicing is a regular event in the lives of viruses and bacteria. Similarly, the DNA of higher organisms, including that of human beings, reveals the presence of sequences that have originated elsewhere (e.g., an integrated cDNA version of a retroviral genome); not all such integrations result in abnormal physiology or development.

The hazard that has been of immediate worry is that of accidental release of potentially harmful chimeras due to errors of handling. As the types of harm that can ensue cannot be specified as yet, measures have been concentrated in devising safe laboratory practice. The NIH regulations, mentioned earlier, provide comprehensive guidelines for safe handling of engineered organisms. It recommends categories of containment laboratories on the basis of the risk involved in a project, as also the use of crippled organisms that cannot survive in the human body or even in the environment outside the special one provided in the laboratory.

The initial reaction of a section of the community of molecular biologists was to stop all rDNA research. Those who were not in favour of a total moratorium on rDNA research recommended that caution and discretion must dictate the type and extent of research to be pursued.

Media reporting, necessarily sketchy and sensation-mongering, incited the untutored public to imagine the worst, from deliberate productions of Frankensteins and robots to cripples according to the goal of the controller of the project! The NIH guidelines, at least, stemmed somewhat this escalating hysteria. The two years it took to formulate the guidelines further quenched this fear and also slowed down rDNA research until even scientists began to have a more realistic perspective.

Further understanding of molecular biology and a lack of any obvious catastrophe even after two decades of rDNA research has lessened the intense alarmist attitude in the majority of the scientific and lay communities. Containment of chimaeric organisms and discretion in the types and goals of a research were deemed to be sufficient retardants for gene cloning to go out of hand, both physically and metaphorically.

5.7.2 Ethical Issues

The controversy surrounding rDNA research was fanned by the growing realization that its results could have grave ethical repercussions. Would success with lower organisms encourage genetic manipulation of mammals and eventually of human beings? Questions arise as to whether society should allow experimentation on its own members, and further, who would be the chosen few, or many, who would be competent to be effective and can democratically grant permission? Indeed, one may go one step higher and question who would be the ones to decide on the body of permission tenderers.

Debates on the 'ethics' of 'tampering' with the social values that we cherish today were as heated as the ones on the effective logistics of controlling research deleterious to the society of the present and of the future.

One example of how certain advances in the use of rDNA techniques may immediately affect society can be cited. Today, there are several rDNA-based strategies for diagnosing incipient hereditary diseases, symptoms of which may not appear until middle age. By this time the subject has already passed on his/her defective genomic contribution to the offspring. The question would then arise: What would be the correct (read 'ethical') procedure to follow: to suppress this information from the general public, or to include it in a person's personal data sheet? Such a person, healthy till middle age, would be deprived of the advantages of both a family and a career in anticipation of the time when he/she will be ill and/or incapacitated, if the fact of his/her disorder was known to society. Would the employer, on the other hand, be guilty of discrimination against a member of the society by not employing him/her? Could the employer take the risk of appointing someone who is doomed to be of substandard health and is likely to become unfit for work in the future? Would it be unethical of him to reject such a person in favour of one who will continue to give returns to the employer's investment in him/her over the years? Such questions were raised several times, but especially after DNA sequencing became a facile project. Some advocated that this most accurate and unique 'data sheet' should form part of every person's list of personal qualifications. Does society possess the right to invade its member's privacy to this most intimate of levels?

Ethical questions are raised when the misuse of some techniques is already evident. Two types of misuse of the new biological techniques are here with us already. Amniocentesis, hailed as one of the most useful clinical techniques for assessing genetic aberrations at the early fetal stage, has acquired the status of a 'menace' in India, where it is being used surreptitiously for aborting female fetuses. The social and moral impact of such a behaviour is obvious.

Another misuse that may be prominent in the future is that of using surrogate mothers to develop embryos of women reluctant to undergo the rigours and inconveniences of pregnancy. Obviously, poorer members of the society will be the candidate surrogate mothers, and like professional blood donors, they will use this stratagem for supplementing their meagre incomes. This would lead to ill-health and shortened life expectancies and at the same time stretch further the gap between the more affluent and the less endowed members of society. Similarly, cases of subversions of other useful genetic engineering techniques may materialize in the future.

Conversely, the society has to develop strategies to ensure that useful spin-offs of rDNA technology do not languish untapped in archives, because commercial promoters do not consider them to be potential profit earners. 'Good' ethics demand that the nation provides for the development of underprivileged candidates, i.e., for the benefit of even a few in a society, which has funded most rDNA research to date.

A case-by-case assessment may be made of every potential application of genetic engineering, and the ethics of sanctioning, fully or partially (or not at all), and rDNA research considered item by item. However, there are other primary ethical issues, the analysis of which is of greater fundamental importance.

The primary ethical issues have engendered heated controversies first among molecular biologists themselves, and later among the general lay public. Basically, the debate is concerned with the ethics of allowing research that may usher in harmful consequences of a profound nature. Would 'stopping' or curbing of such research be an infringement of the traditional freedom of the scientist to extend the horizon of knowledge? Furthermore, who or what body would decide which band of assessors would be appropriate to give or refuse permission, impose restrictions and oversee their compliance by the practitioners?

The basic ethical conundrum is represented by the controversy to continue research, that may harm individuals while bringing in advantages to many, versus maximum social benefits for the price of disadvantages to a few. The option chosen depends on the value one places on the dignity and freedom of every individual. Adherents of the first option place a premium on an individual's right to his (or her) own welfare: those of the second, while not opposing this view, believe that a pragmatic approach would be to make compromises which would ensure maximum benefits for most. The former, essentially upholding the views of the philosopher Kant, insist that all rDNA research should be stopped, if there is the slightest chance of its leading to deleterious effects or to misuse. Projected benefits may be sacrificed according to them, in order to safeguard society from even speculative disasters.

Social and ethical obligations also play a part in one other avenue of use of rDNA research. It was a recognized practice in the past to export to less developed countries, technologies or their products which had lost favour or had even been banned in the developed countries of their origin. It has been speculated that such exports are not only of financial value to developed countries; but also represent experimentation on the non-affluent, and hence non-influential, members of the world community (expendable as well?) that will be abhorrent in the parent countries. The result of their use—deleterious or beneficial—will supply data not available by other strategies. The world is no longer a collection of disparate nations

scattered far and wide, and such experimentation, in the long run, cannot fail to affect the world community at large.

5.7.3 Economic Issues

Several effects on the economic stability of the society may be envisioned as a result of the choice of priorities in the use of recombinant DNA products. The more prominent effect represents the fear of an inequitable distribution of the benefits of this technology.

It is argued, for instance, that due to the essentially expensive production of genetically engineered products, only large organizations, mainly multinationals, will dominate as producers and acquire monopolies, one of the ways known to upset economic stability. High-yielding crops usually require extra inputs of fertilizers and proper water management routines that all add to the cost of production. A failure of a harvest may drive a smaller producer to rely heavily on loans. This can be visualized to generate a viscious cycle of dependency that would progressively impoverish the poorer farmer and prevent him from providing the optimum conditions for taking advantage of the specially bred seeds.

Other economic instabilities may result from the manufacture, by rDNA technology, of products which at present form the main money-earners in certain smaller nations. Then again, loans given for fertilizers, pesticides and crop management may be diverted to the purchase of labour-saving machinery. This is likely to deprive large populations of labourers of their only means of earning a livelihood.

The use of rDNA processes and products cannot, therefore, be made without taking into account the various facts of a society's state of economy.

The misuse of human genetic engineering, such as subversive use of surrogate mothers (mentioned in the earlier section), may also lead to widening the breach between the haves and have-nots—which surely is not the ambition of the society in a global sense. No one questions the efficacy of several projected human engineering schemes. Indeed, it is ridiculous to believe that such a powerful tool, for intervention in health and hygiene, should not be exploited maximally. It is the anxiety of the power going to the 'wrong' hands (i.e., in the control of those whose objectives are contrary to that of mankind in general) that has triggered off thinking about the role of society in safeguarding the use of a technology that may eventually have adverse repercussions on the economic stability of the society as a whole.

The 'wrong' hands may indulge in sophisticated biological warfare, by creating microorganisms redesigned to carry deadly toxins or other lethal agents. One section of society, comprising mainly the experts, will insist that in order to know the extent of damage by such agents, as also to develop antidotes for them, rDNA research becomes a necessity. Society has to decide whether or not this goal-oriented rDNA research should be allowed. The basic issue is whatever the motivation (defence or aggression), the research will still be attended by hazards due to accidental contamination of the environment. Besides, how will society guarantee that what started out as 'defence research' will not be used for attack? Military research, by its very nature, is secretive. How would a society keep abreast or monitor the type of research that is pursued in military and other restricted laboratories?

A further source of fear of rDNA products is the unpredictability of the impact that may result from secondary causes. Technological forecasting, always a difficult venture, is particularly inadequate at present, for assessing the secondary impacts of the use of rDNA products. This is one more item in the agenda of deliberations of socially conscious bodies.

Monopoly of production in a few developed nations may cause additional financial burden on nations that import their products. Say, one product carries a shortcoming that can be remedied by another product marketed by the same multinational only. The poor consumer will have to buy both. This situation can be visualized in terms of the use of pesticides containing a non-degradable component. If the biodegrader is also patented by the same company, the customer would have no choice but to purchase both the inherent poison and its antidote and that too from the same company.

The above discussion is meant to highlight only a few of the avenues by which rDNA research and products may affect economic and social stability of a nation. The moral of the story is that serious thinking should evolve appropriate rules that will protect the society from the evils visualized today and future ones inherent in the use of a revolutionary technology. Equitable distribution of the benefits of the same has also to be ensured.

It will not be an easy task to design democratic methods of sanctioning (or not) and of monitoring the progressive use of such technologies, as the interests of several lobbies are at stake. The scientist primarily perceives the career ahead of him; the funding agent visualizes the profits (social and economic) to be gleaned out of the ventures; commercial exploiters have their own obvious axes to grind. There have been few examples of public efforts rewarded with easy success when pitted against such powerful lobbies with common interests. Nevertheless, avenues are being pursued for designing workable, democratic measures of sanction and control of recombinant DNA research and products.

5.7.4 Legal Issues

Legal issues arising from the use of rDNA products or experimentation fall broadly into two categories: (i) individual issues that arise on a case-to-case basis, and (ii) global issues that demand decisions from the community at large.

The issues belonging to the first category were initially, almost predominantly, about laws for protecting intellectual property. Patents have been conventionally granted for products and processes dealing with non-living matter. Anondo Chakravarty's claim to patent oil-spill antagonists, namely the *Pseudomonas* superbug strain created by him, brought the issue to the forefront. A long and intense legal battle in the USA ended in success for the claimant: the superbug became patented. This single victory has not closed the chapter on patent rights for recombinant organisms. What would be the rights of a commercial group that creates, say, a paragon cereal crop plant? Will the strain be exclusively its own property, or should the general public have claims on its use without strings attached? Other complex legal problems may be perceived if one persists in this line of questioning. These issues need, therefore, to be identified, faced and tackled.

Legal knots would be encountered if human genetic engineering developed to the stage when cloning humans becomes a reality. This is not an improbable expectation. Cloning of animals of economic value and of genetically redesigned ones (such as animals with supplemented growth hormone genes) have already been approved. Indeed, cloning is recognized as an effective technique of amplifying (mass-producing) desirable strains. If human cloning becomes possible, what will be the legal status of the copies in terms of inheritance laws and of social norms in terms of spouses and parents? Again, if tissue culture is utilized to perpetuate an individual, where would the line be drawn for each generation? These questions, apparently weird today, will have to be faced in the future, if advances in synthetic biology keep pace with contemporary expectations. Considering the manner in which science advances by spurts and surprises, one cannot avoid speculating about situations that have a way of turning science fiction into reality.

It is, however, the second category of legal issues with which the concerned bodies are grappling at present. These issues include regulations or guidelines for recombinant DNA based investigations and ways and means of controlling the type and extent of rDNA applications.

The Asilomar conference had included a few non-biologists; among them were four lawyers. This fact indicates that even in those early days it was felt that legal minds would have to be engaged in the problems thrown up by this unusual technology.

The NIH guidelines that followed provided a measure to chalk out norms for safety in the laboratory, that would discourage environmental contamination with escaped man-made genetic chimeras. Work environments were categorized on the basis of the degree of hazard inherent in the project, and the use of 'safe' or 'crippled' strains was recommended. No directions were available, however, for dealing with chimeras that are bound to escape, given the fallibility of humans and the unpredictability of natural accidents.

A more fundamental issue is: should there be laws to restrict rDNA research? Would the passing of such laws jeopardize a scientist's freedom to advance along his/her own path of science? Does anyone possess a right to follow a path that may spell trouble to other members of a society? What is so sacrosanct about a desire to follow a certain route of scientific investigation? Should there be laws to protect society from a line of research expected to bring damage in its wake? If so, who would be the law-makers? Who will decide the composition and mandates of a legal body formed for the above purpose? What proportion of the bodies should consist of the lay public? Can the lay public, given its lack of understanding of the subject matter, do justice to the responsibility given to it as representative of a law-making body? Conversely, can one guarantee that the experts, who do know the subject matter in depth, will be free from bias? Would the scientists, sometimes in league with commercial lobbies, insist on railroading research in a direction which honestly is 'not right'? Questions such as these, and others, need to be tackled and reasonable answers found for them.

The immediate task is to decide how society can control an activity that may not be totally in the best interest of all its members, and not infringe on those activities that aim to increase knowledge and offer future benefits.

Debates have ranged from the nature and tightness of control that will be acceptable to all sections of a society to the nature of the composition of bodies that will undertake such a task.

The status of the regulatory body is very important, as all individuals and organizations engaged in rDNA research and applications should come within its purview. It should also be decided whether or not rDNA products and processes of production ought to be subjected to the scrutiny of environmental protection agencies. If so, ways must be devised to shorten the period spent in gaining approval from such agencies, and prevent unnecessary bureaucratic bottlenecks in the path of the advance of an unusually rapidly developing technology.

These issues and many other crucial and finer points need to be thrashed out once and for all, so that identical exercises are not repeated for every technological advance that moves at a rate far greater than that at which rational assessments may be made of its impact on society.

Recombinant DNA technology has created a novel situation where we have to search for new norms, scrutinize conventional values, and design user-friendly social behaviours to deal with issues that may vitally affect the well-being of the human community. Effective participation in decision making, without unnecessary trauma due to ill-informed members or parties with vested interests, would be the major objective. Only when there is some measure of understanding as to how this objective may be achieved will society be educated or matured enough to face this technology as well as other unconventional technologies of the future.

This chapter attempts neither to side with the voices of doom and cry 'do we have the right to play God?', nor to uncritically extol the miraculous benefits that are sure to emerge as a result of rDNA technology. It aims to stimulate the novice genetic engineer to think beyond the unique techniques and their immediate applications, and thus begin the process of becoming a discerning and responsible member of society, ready to participate in decision making in the future.

REVIEW QUESTIONS

1. What is mutagenesis? Enumerate the alterations encountered, and mutagens that bring about such changes.
2. Recall the principal technique in directed mutagenesis.
3. How can a gross picture of a DNA be sketched with restriction mapping techniques?
4. What do DNA footprinting, chromosome walking and chromosome jumping involve?
5. Explain the importance of somatic cell hybridization in mapping of human genes.
6. Recall the early chemical/enzymatic methods of DNA sequencing.
7. What is shotgun sequencing?
8. How are computers commissioned to analyze sequences?
9. Review the methods of gene transfer in plants.
10. Recall the assays employed to assess expression of a transferred gene into the plant cell.

11. Explain how recombinant DNA techniques make it possible to delineate the physical map of a DNA?
12. Elaborate on various studies in development vis-à-vis regulation of gene expression.
13. How do the r-DNA techniques help in unfolding of the molecular mechanism underlying the phenomenon of immunity in animals?
14. Recall the general method for isolation of oncogene sequences.
15. How can anomalies identified in DNA aid in timely diagnosis of inherited diseases?
16. Briefly recall the genetic engineering strategies vis-à-vis plants.
17. How can animal genetic engineering be employed for human welfare?
18. Enumerate and elaborate on how pharmaceutical and chemical industries have gainfully exploited biotechnological interventions.
19. What are the ways in which structural-functional rapport of macromolecules can be unveiled? What is the fall-out in terms of exploiting the knowledge thus gained?
20. Explain the role of genetic engineering in clinical applications.
21. What is the contribution of genetic engineering in control of environmental pollution?
22. How can the potential of r-DNA technology be exploited in mineral recovery from mines?
23. What is the importance of biosensors and biochips developed with help of r-DNA technology?

PART II: *The Practice*

Working with Bacteria and Phages

6

Introduction

Genetic engineering may be undertaken for any species of prokaryotes, eukaryotes or viruses. The initial and most voluminous studies have been performed in bacteria and bacteriophages using the *Escherichia coli* as the cloning cell. Construction of hybrid DNA, modification of vectors and other initial exercises of DNA engineering are still carried out best in this species.

With progress in technical aptitude, other species of prokaryotes and eukaryotes were explored for use as cloning and/or expression vessels. Species of *Salmonella, Bacillus,* and *Streptomyces* are also in use by now. In eukaryotes, yeasts came into use early in the history of recombinant DNA technology. Because of the commercial interest in yeast, exploitation of the latter has become the basis for research and application in a particular population of investigators. Next came the tentative use of more complex eukaryotic systems, including those of plants.

It is necessary, therefore, to be familiar with the most common systems that are in use and acquire a proficiency in handling them. After all, the recombinant DNA is to be constructed and utilized in the milieu of these systems in the majority of projects.

A few of the more frequently utilized species are introduced in the following pages, together with a few representative exercises for handling them. A familiarity with basic microbiological techniques is taken for granted.

6.1 ❑ WORKING WITH BACTERIA

Bacteria are members of prokaryotes and are the most primitive of cellular organisms. A bacterial cell contains the DNA as a single double-stranded covalently closed (circular) molecule. Since only one DNA molecule is present, the cell is a haploid one, with only one copy of each allele. In unusual cases, due to transduction via a bacteriophage, regions of the DNA may contain two alleles of a gene. Such a cell is said to be a merozygote for the loci in question.

The bacterium that has been in maximum use for recombinant DNA studies and applications is *Escherichia coli.* Others, such as *Bacillus subtilis, Salmonella typhimurium* and *Streptomyces* spp. are in the process of being exploited for use in gene-cloning projects.

A brief introduction to *E. coli,* that will be utilized in the following exercises, is given in the next page.

6.1.1 *Escherichia Coli*

E. coli is a single-celled bacterium with a single covalently closed double-stranded DNA molecule as its main chromosome. It may also harbour one or more plasmids, which are very small circular ds DNA molecules. The chromosomal DNA is attached to the inner surface of the plasma membrane at a particular point.

The *E. coli* cell divides once every 20 minutes at 37°C. This event is preceded by the replication of its DNA. The cell membrane invaginates between the two new DNA molecules and separates the parental cell into two independent cells.

The *E. coli,* like most bacteria, follows a standard pattern of growth. Cultures are usually stored at 4°C, as stab or slant cultures. To obtain a growing culture, a very small aliquot (loopful) of the stock culture is added to 10 ml of a liquid culture medium (such as LB broth) and incubated at 37°C. When a new culture is thus started it may take a long period for it to begin growing vigorously. This is known as the *lag* phase; it is followed by an exponential growth phase which slows down due to lack of oxygen in the limited medium. In the laboratory, the exponential rate is maintained at a steady pace by making available extra oxygen (by shaking the culture in a temperature-controlled water-bath).

The overnight culture is introduced into a larger volume of the same liquid medium and incubated with shaking for 2–3 hours. The concentration of cells at the end of this period is adequate for further use. The flask or container in which the cells are cultured should not be more than 1/3rd full, in order to ensure a sufficient volume of oxygen for the cells.

For most experimental work, a minimum density of 10^7–10^{10} cells per millilitre is adequate. The density is checked with a spectrophotometer (Spectronic 20 in the author's laboratory). A reading of A_{600} = 0.7 coincides generally with the required density. One may, however, work out the relation between the A_{600} reading and the cell density in the following manner.

Take two equivalent samples from an *E. coli* culture at one hour intervals from the end of the lag phase to the stationary phase. Check A_{600} for one of the two samples for each interval; use the other sample to find out the actual cell count in the suspension. To do this, aliquots from the second sample are plated on LB-agar and incubated at 37°C for 1–2 days and the number of colonies counted. The number of cells/ml may be calculated from these data, from the volume of suspension plated and from the dilution of the sample. The A_{600} and cell count values are plotted against each other and a graph drawn. Provided the culture conditions remain the same, this graph may be used to read off the cell count (density) corresponding to the A_{600} value of a given sample. (It must be remembered that while the reading of a coloured solution in a spectrophotometer is an indication of the amount of light absorbed by the solution, the reading from a cell suspension is a measure of the light not reaching the photomultiplier in the instrument.)

The strains of *E. coli* to be used carry known marker genes (e.g., *thr,* str^s). The markers are liable to mutate in a small percentage of the cells in the stock culture (i.e., $thr^- \rightarrow thr^+$ $str^5 \rightarrow str^r$). It is essential, therefore, to check the integrity of the strain before use. One can save much time and avoid futile work by ensuring that the markers are as indicated on the label or the stock record book.

E. coli cells may carry the F plasmid, which is autotransmissible to an F^- cell. The former is called an 'F^+ cell'. Sometimes, the F plasmid DNA is integrated into the *E. coli* chromosomal DNA. In this case, the cell is said to be an Hfr (high frequency of recombination) one. Both F^+ and Hfr cells possess fine hair-like outgrowths on the cell surfaces. These are the pili (singular: pilus). When the tip of a pilus contacts a receptor protein on the surface of an F^- cell, the membranes between the two cells, at the point of contact, become degraded. As a result, the two cells become connected by a thin channel. The DNA of the F plasmid (free or in the Hfr cell) is nicked at a specific region on one of the two strands. The free 5′-end at the nicked region moves out and passes through the connecting tube into the F^- cell. As the strand is pulled out from the ds DNA, a new strand takes its place (by addition of a nucleotide complementary to the free base in the intact strand). Similarly, the moving ss DNA acquires a complementary strand. In the case of an Hfr cell, the moving region of the F strand is separated from the remaining F region by the strand of the chromosomal DNA. Transfer of the F DNA is completed only after the entire *E. coli* DNA strand has been transferred.

In reality, the entire Hfr chromosome rarely becomes transferred, as the fragile pilus breaks with the slightest movement. Whatever is passed on to the F^- cell possess alleles of *E. coli* that recombine with homologous alleles in the F^- cell. The recombination is non-reciprocal, that is, only the exchanged allele in the main chromosome remains, while the fragment of DNA (reciprocal recombinant) is degraded. This parasexual mode of gene exchange in bacteria is known as conjugation. Haploid *E. coli* cells do not recombine with genes on a regular basis. However, a parasexual event introduces alleles from a donor cell and promotes recombination. This is why the donor cells that transfer *E. coli* alleles and thus show a high frequency of recombination are known as 'Hfr cells'. The Hfr cell is often referred to as the male cell, and the minimal medium plate in which it grows is called the 'male plate'.

As is obvious from the foregoing information, the F^- cell acquires varying lengths of the chromosomal DNA from an Hfr cell. In an ideal condition, the entire chromosomal DNA may be transferred; it takes 90 minutes for the complete transfer. The length of *E. coli* DNA transferred is assessed from the percentage of the loci which exhibit recombination in the F^- colonies. To identify the recombinants, the mixture of Hfr and F^- cells is plated on a series of plates, called 'marker plates', each of which contains the minimal medium minus one of the essential metabolites (amino acid, sugar etc.), the biosynthetic genes of which are to be monitored for recombination. In other words, if one needs to find out the recombination value for a locus (*leu* for leucine synthesis), the donor Hfr is a leu^+ and the F^- a leu^- strain. The recombination mixture will contain some cells in which the leu^- has become exchanged for a leu^+ allele. Such a cell will be able to grow on a medium that lacks leucine. In short, recombinants for the *leu* locus are scored as the number of colonies that appear on a leu^- marker plate.

Wolman and Jacobs used the above principle to find out the sequence of alleles that are transferred during the conjugation of *E. coli.* Since the pilus breaks easily, there are always more of the shorter fragments of transferred DNA than the longer ones. This is reflected in a larger percentage of recombinations in the alleles that are

transferred earlier. By arranging the recombination values in a descending order, one can map the sequence of gene loci in the circular DNA. This is what Wolman and Jacob did; their method of mapping *E. coli* genes is, therefore, known as the 'interrupted mating method'. In the laboratory, Hfr and F^- strains of known genetic constitution are allowed to mate, and an aliquot removed, shaken and plated on marker plates. It may be necessary to plate each aliquot in different dilutions on each plate, if the cell density of the recombinants is expected to be high in the mixture. The mating is interrupted every minute (or as required) from 0 to 120 minutes. After every 90 minutes, the values of recombinations will increase per locus. This fact led to the conclusion that the *E. coli* chromosome is a circular DNA, that requires 90 minutes for one round of transfer to an F^- cell. The *E. coli* map is divided into 100 minutes and the gene loci are marked on this in the sequence in which they are transferred, the values on the map indicating the time of transfer.

To obtain a purified *E. coli* (or other microbial) strain, the culture should be started with a single colony, where every cell possesses the same genetic constitution. To achieve this, streak a loopful of the stock culture across the surface of the medium in a petri dish. Some cells will be scattered individually outside the main mass of colonies that arise along the streaks. These separated cells will grow into distinct colonies that may be picked up with sterile toothpicks and placed in separate, sterile 10 ml LB broth (in test-tubes). These colonies are incubated at 37°C overnight, and plated on petri dishes. The latter are incubated at 37°C until the colonies are 1–4 mm in diameter. These colonies are checked for the integrity of the marker genes before proceeding with an exercise using the *E. coli.*

6.1.2 *E. coli* Carrying pBR322

The popular plasmid pBR322 possesses the marker genes AMP^R and TET^R It is customary to write the names of drug-resistant genes on plasmids in capitals using the first two or three letters of the drug in question (e.g., AP^R or AMP^R for ampicillin resistance). The names of drug-resistant genes on chromosomal DNA are written in italics (e.g., amp^R).

E. coli (pBR322) is grown on a LB medium supplemented with ampicillin (50 μg/ ml) and tetracycline (10 μg/ml). This prevents 'curing' of the cells of the plasmid. That is, the selection pressure due to the drugs, prevents loss of pBR322 from the cells—an event that may occur in the absence of selection pressure (the antibiotics in this case).

It should be noted that pBR322 is a relaxed mode plasmid. It may be amplified by adding the protein synthesis inhibitor, chloramphenicol, to the growing cells. The cells may eventually contain 1000 or more copies of the plasmid.

6.1.3 *Agrobacterium tumefaciens*

Agrobacteria occur as four known species. *A. tumefaciens* causes *crown gall* disease in *dicotyledonous* plants. *A. rhizogenes* is responsible for the *hairy root* disease; *A. rubi* for *cane gall* disease, and *A. radiobacter* is a virulent species that kills bacteria with a toxin called *bacteriocin.*

Crown gall disease was reported as early as 1901 by Tswett, and a tumour-inducing principle was suspected but not verified. In 1965 Lippincott and associates felt that the DNA was involved in some manner in the origin of the disease. Hamilton and his team (1971) reported that a virulent *Agrobacterium* became avirulent at 37°C, but grew comfortably at 27–30°C. In 1974 it was discovered that the conversion to an avirulent form is accompanied by the loss of a plasmid from the infecting bacterial cell. This plasmid was named the *tumour-inducing* or Ti plasmid.

The Ri plasmid causes the formation of bunches of fine hairy roots from the infected region of the plant, and is a *root-inducing* plasmid.

The *Agrobacteria* enter the plant only through a wound. Certain receptors in the inner tissues recognize, and bind themselves to, specific cell surface molecules on the surface of the *Agrobacteria.* The bacterium then synthesizes cellulose fibrils which get attached to cell walls and lyse the cell.

Three characteristic changes occur in plant cells infected with *A. tumefaciens:*

1. A tumour is formed at the site of the wound and eventually stops further growth of the stem.
2. Certain unusual amino acids called *opines* are synthesized in the affected cells. These are utilized by the bacteria as a source of C and N. Indeed, excess opines find themselves in the soil at the base of the plant and foster the growth of *Agrobacteria* in the soil. Since only these species can utilize the opines, very soon an ecological niche is carved out, that excludes the presence of other bacterial species.
3. The infected cell relinquishes its normal need of phytohormones. Animal cancer or 'transformed' cells also do not require serum proteins (in contrast to normal cells).

Different strains of *A. tumefaciens* are committed to the synthesis of different opines. These are classified into the following four types:

1. (a) Octopine : pyruvate + arginine
   (b) Octopinic acid : pyruvate + ornithine
   (c) Lysopine : pyruvate + lysine
   (d) Histopine : pyruvate + histidine
2. (a) Nopaline : α-ketoglutaric acid + arginine
   (b) Nopalinic acid : α-ketoglutaric acid + ornithine
3. Agropine : mannose + glutamine
4. Agrocinopine : pyruvate + sugar derivative

The functions of opine synthesis and usage and the consequent formation of the plant tumour are due to the presence of genes on the Ti plasmid. This is a large plasmid (150–200 kb), which contains an inset of a 15–23 kb region called the T DNA (tumour DNA). When a Ti plasmid enters a plant cell (via the bacterium), a copy of the T DNA (a transposon), becomes integrated into a plant cell chromosomal DNA. After this integration, the bacterium is no longer needed for tumour formation. The *Agrobacterial* plasmids are the only known examples where a prokaryotic gene is integrated into a eukaryotic genome and *expressed.*

The T DNA region carries genes for opine synthesis and degradation, and for morphology differentiation: *tms* (Tumour Morphology Shoot), *tmr* (Tumour

Morphology Root) and *tml* (tumour morphology). There is a gene for virulence in the Ti outside the T DNA region. The gene ensures infection by the bacteria.

The normal undifferentiated cells of *Nicotiana* (as also other plant species) possess a specific ratio of auxin (root-forming hormone) and cytokinin (shoot-forming hormone). Imbalance causes both the tumour and the undifferentiated tissue to differentiate into a shoot or a root. Shoot formation is induced by a reduced proportion of the cytokinin transribosyl-zeatin (by inducing a mutation with a 1.1 kb transposon in the *tms*). If this mutation is corrected, callus formation is restored. Similarly if a mutated *tmr,* induced with a 1.6 kb transposon, causes the level of the aforementioned cytokinin to increase abnormally, the callus differentiates into roots (the ratio of auxin is decreased).

The *Agrobacterial* pDNA is of interest to gene manipulators, since it can be utilized as a vector for introducing genes into plant species. The tumours caused by Ti plasmids have not always been successfully differentiated into whole plants. Some laboratories have designed methods of transformation and differentiation for certain plant species, but they are usually not 100 per cent successful. Ri plasmid-induced adventitious roots can be on the other hand, differentiated into complete, tumourless plants. So, interest is veering towards the use of Ri plasmids for genetic engineering in plants. However, most of the investigations have, to date, been concerned with *A. tumefaciens* and its Ti plasmid; hence, we have much more data on Ti than on Ri plasmids. There is some homology in the non-T DNA regions of Ti and Ri plasmids. The Ri plasmids also cause the synthesis of agropines.

The main problem of using Ti as a vector is its large size which contains recognition sites for several restriction enzymes. In spite of this problem, attempts have been made to transform plant cells with T DNA carrying a *passenger gene.* The transposon Tn7, carrying a gene for resistance to the drug methotrexate, has been introduced into a plant cell via the T DNA. The gene has been expressed in the foreign recipient cells, so that the cells as well as Fl progenies of them were found to be methotrexate-resistance. In still other attempts, transformation and gene insertion took place, but the gene was not expressed in the heterologous cell (e.g., genes for interferon and phaseolin, a protein of beans). The mRNAS were produced, but they were not translated.

If a method could be found to express foreign genes in plant cells, via the Ti plasmid, and, at the same time, prevent the formation of the tumour that retards and stops the growth and development of the plant to be genetically engineered, it would be an ideal way of introducing useful genes in plants of commercial importance (timber trees, resin forming trees and others giving fibre, oil, etc.).

Exercise 1: To Check the Genetical Constitution of a given Strain of E. coli

Have ready

1. *E. coli* MD1767 (F^-)
2. *E. coli* AB1157 (Hfr)
3. M9 stock solution (see procedure, given later, for preparation)
4. $CaCl_2$ (1 M) solution (sterilized)
5. $MgSO_4$ (1 M) solution (sterilized)

6. Glucose (20%) solution (sterilized)
7. Vit Bl (thiamine) (0.1%) solution (sterilized)
8. Sterile dist. water (SDW) 1000 ml
9. Agar (3% in water) solution (autoclaved) 500 ml
10. Petri dishes 40 (standard 90 mm diameter)
11. Amino acid stock solutions (filter sterilized):

Threonine	71 mg/10 ml
Leucine	79 mg/10 ml
Proline	460 mg/10 ml
Histidine	31 mg/10 ml
Arginine	253 mg/10 ml

12. Streptomycin (10% in water) solution (filter sterilized)
13. Saline (9% in water); sterilized
14. Glass rod spreader (L-shaped rod)

Summary of steps

1. Grow Hfr and F^- strains in overnight cultures.
2. Make dilutions of Hfr and F^- cultures.
3. 'Spot' dilutions of Hfr and F^- on the following:
   (i) Five marker plates (*leu*$^-$, *his*$^-$, *pro*$^-$, *arg*$^-$, *thr*$^-$)
   (ii) Three minimal medium (MM or *male*) plates
   (iii) One control plate (MM + 5 amino acids)
   (iv) One control plate (MM + streptomycin)
   (v) One control plate with no cell
4. Incubate the plates at 37°C overnight.
5. The next day, observe and note the number of colonies, if any, in the 'spots' and controls.
6. Draw conclusions about the presence of the alleles for *leu, pro, his, arg* and *thr* in the Hfr and F^- strains.

Procedure

Day I

1. Prepare amino acid stock solutions:
   (i) Dissolve each amino acid in SDW and filter sterilize.
   (ii) Store in sterile bottles at 4°C.
2. Prepare a 10% stock solution in SDW.
3. Prepare M9 stock for minimal medium (MM).
   (i) Dissolve the following in 1000 ml dist. water (DW).

Na_2HPO_4	5.8 g
KH_2PO_4	3.0 g
NaCl	0.5 g
NH_4Cl	1.0 g

   (ii) Autoclave.
   (iii) Keep at 4°C until use.

Day 2

4. Prepare MM (just before use):
   (i) In a 500 ml (sterile) measuring cylinder, pour 100–200 ml SDW.
   (ii) Add the following in the given order:

1 M $CaCl_2$	0.1 ml
1 M $MgSO_4$	1.0 ml

   (iii) Stir until dissolved; then add

M9 stock	40.0 ml
Vit B stock	1.0 ml
Glucose stock	20.0 ml

   (iv) Add SDW up to 500 ml. (The salts do not dissolve in larger quantities of water.)
   (v) Make in 500 ml DW 3% agar solution. Autoclave to melt the agar and sterilize.
   (vi) Heat the stored agar solution to 45–50°C (in a water-bath).
   (vii) Add the agar to the M9 solution (500 ml + 500 ml).
   (viii) Distribute this mixture, while hot, into sterile petri dishes. You may do this by using a 20 ml pipette with the tip sawed off to provide a wider passage.
   (ix) Add two drops of streptomycin to plates 2 and 4 and spread evenly with the bent glass spreader. *Open lid of dish as little as possible to avoid contamination.*
   (x) Add two drops of each of the five amino acids to (3) and (4) above. There will be about 25–50 mg/ml or 1 mg/plate of each amino acid. Spread the amino acids on the plates.
   (xi) Place all the plates in the incubator (37°C) with the lid slightly open, for 30 minutes.
   (xii) Close the lids and leave plates (inverted) in the incubator (37°C) until use.

Day 3

5. Take out eight MM and F plates and five marker plates and label them (at bottom of each plate) as follows:
   1. MM + Hfr
   2. MM + Strep + Hfr
   3. MM + F^-
   4. MM + Strep + F^-
   5. F + F^-
   6. F + Strep + F^-
   7. F + Strep no cells
   8. F – no Strep, no cells
   9. –(13) Marker plates *leu, thr, pro, arg, his*
6. Dilute the overnight Hfr and F^- cultures in saline to have dilutions of 10^{-2} and 10^{-3} of each culture.
7. Take a small drop of each Hfr and F^- dilution and 'spot' it on each plate, which has been labelled on the bottom (of the plate) for each strain and dilution.
8. Incubate inverted plates at 37°C overnight.

Day 4 or 5

1. Observe growth/no growth of each spotted culture.
2. Record results.
3. Give explanations for your observations.

6.2 ❑ WORKING WITH BACTERIOPHAGES

The bacteriophages that have been utilized mostly are the ones that infect *E. coli* strains. These include the ds DNA-containing virulent bacteriophages (T2, T4 etc. and the T_{ODD} series), the temperate bacteriophages (lambda and P1) and the ss DNA-containing phage M13.

The T phages possess a 'head', a tail with additional covering proteins and a basal region with fibres that protrude from the region where the tail meets the base. The phage gets attached to the *E. coli* surface by means of the tips of these fibres. After attachment, the phage DNA (linear) is shot into the *E. coli* cell, where it is first transcribed (early genes) by the *E. coli* RNA polymerase. One of the products of the *early* genes shuts off transcription from the *early* promoter, while a phage-encoded gene makes possible the replication of the T DNA and transcription of the *middle* genes, many of which specify phage proteins. At the end of the replication of a sufficient number of copies of the phage DNA, individual phages are formed, by including one DNA each within a protein shell. The *late* genes are mostly for shell proteins. One of them is for the enzyme lysozyme that lyses the *E. coli* membrane. The mature phage particles are released. The number of phages produced per cell is referred to as the 'burst' size.

Phage lambda is a temperate phage. Its DNA has been exploited mostly as vector material. It can undergo a *lytic* cycle as in the case of a T phage, or a *lysogenic* cycle in which the lambda DNA becomes integrated in the *E. coli* DNA. The integrated DNA may be released under certain conditions. If the deintegration is faulty, a portion of the *E. coli* may be carried away by a phage. This hybrid DNA will cause recombination on transduction to another *E. coli* cell.

Phage P1 is also a temperate bacteriophage. However, P1 causes generalized transduction, while lambda brings about specialized transduction. This means that the piece of *E. coli* carried away by mistake can be from any region of the *E. coli* in the first case, but only from a specific region in the second. This will be elaborated in subsequent sections 6.2.1 and 6.2.2.

Phage M13 has a linear ss DNA in an elongated shell. It infects *E. coli* strains carrying F^+ or Hfr DNA; these cells possess pili to which the M13 is attached.

Phage M13 has a unique manner of producing the ss DNA in the mature phage. The infecting DNA first becomes circular and then double, with the synthesis of a complementary strand within the host cell. This replicating factor or RF replicates several times before only one of the strands becomes replicated. The new ss DNA is the one that is enclosed in the mature phage. This fact has been exploited brilliantly by J Messing for obtaining pure aliquots of one of the strands of DNA—a requirement for methods of DNA sequencing.

6.2.1 Phage Lambda

Phage lambda is a small bacterial virus that has a lytic and a lysogenic option in its life cycle.

In the lytic cycle, the infecting ds DNA becomes covalently closed into a circular molecule by complementary pairing of 12 bp long regions that protrude as single-stranded 'tails', one from each end of the linear DNA.

The circular DNA first replicates into several copies after which concatamers of the phage DNA are made by the rolling circle mechanism of DNA synthesis. The long multimer concatamer is then cleaved into phage sized lengths by a *ter* (terminating) enzyme acting on *cos* (cohesive end) sites in the DNA. The 12 bp single strand tails are due to staggered cleavage of the *cos* region. Each lambda-sized DNA becomes included in the mature phage particle. One of the late lytic genes expresses lysozyme that lyses the host cell, whereby free phage particles are released.

The lysogenic cycle is initiated by a regulator protein, CI, which in one stroke stops the expression of lytic cycle genes and maintains the lysogenic cycle. The latter is also initiated by other gene products that cleave the *E. coli* DNA at a specific site (a 15 bp region) between a *gal* and a *bio* operon and also join (integrate) the phage DNA at this site. During deintegration the lambda DNA is released. In faulty deintegration, either some or all of the *gal* or all of the *bio* operon is carried away by the phage. Such phages are referred to as gt (*gal*) and gt (*bio*), respectively. A very elegant control system, comprising genes *cro, N, Q, CI, CII* and *CIII* regulates the transition from one type of cycle to another, and also maintains either of the cycles.

Lambda DNA is about 50 kb in length. The phage is rather fussy about the length of DNA that may be packed in the head in a functional phage. It can dispense with about 30 kb of DNA, provided the length is replaced by some other DNA. Use is made of this feature to make appropriate vectors from lambda DNA.

It is customary to maintain phage lambda as stocks of lysogens (the host cell is a *lysogen* when the phage DNA is integrated in the host DNA). The phage may be induced when necessary by using one of several strategies.

A convenient procedure is to use a *ts* mutant of CI that induces the phage DNA from the lysogen at a specific non-normal temperature, and leads to the formation of free phage particles.

Pure stocks of phage suspensions may be stored in a buffer to which a drop or two of chloroform has been added. This prevents bacterial contamination. The chloroform has to be eliminated (by evaporation) before the phage suspension is used.

A lysate of a phage is a suspension of the phage in lysed *E. coli* cell material. The phage suspension may be concentrated and purified, if need be before isolating its DNA. One way of releasing the phage DNA is to crack the phage shell by treatment in alternate extreme temperatures (liquid air or nitrogen and ambient temperature). Another way is to dialyse the phage in the buffer.

The most used ts mutant of phage lambda is CI_{857}. This phage can be induced by keeping cells at a temperature of about 42°C for 5 minutes. Mutations in the *S* gene (for lysozyme) prevent synthesis of the enzyme and the host cell is not lysed. So, lambda strains with *S* mutations can form an amplified number of phages (>1000) in a host cell. These cells are lysed artificially.

Lambda phages are very useful organisms for two types of workers: (i) for those interested in the mechanism of controlled and coordinated regulation (i.e., development) and the associated biology of the phage, and (ii) for genetic engineers, who utilize phage lambda DNA as convenient tailored vectors.

When making a lysate, the mixture should not be carried overnight, since lysogens—*E. coli* (λ) — are formed easily and these latter are not lysed. The lysogens increase in number and give a turbid look to the suspension.

Another precaution has to be taken when phage is used to make an *E. coli* lysate. Although the *E. coli* (λ)— do not support the growth of the phage, many phage particles become adsorbed on the host cell wall and the phages inject their DNA into the host cells, where the phage DNA is degraded (due to *immunity to superinfection*). In this way much of the phage is put out of commission. If you have to use λ, use the strains which are resistant to adsorption and yet lysogenize the *E. coli* cells.

6.2.2 Phage PIKC

PIKC is a phage that transduces random regions of the host cell (*E. coli* K12) DNA. After infection of a cell by PIKC DNA, the latter replicates in the form of long concatamers. These are cleaved into phage size DNA by a phage-encoded enzyme. The latter also fragments the host cell DNA. Occasionally, a fragment of the chromosomal DNA becomes included in the phage shell and is transducted to a new host. The imported DNA may recombine with its homologous stretch in the recipient *E. coli* DNA. Since random regions of the host DNA are transduced, the phenomenon is referred to as *generalized transduction* to distinguish it from *specialized transduction,* the event that occurs via phage lambda.

Phage P1 carries Tn9 with its chloramphenicol resistance gene. The PIKC strain is inducible at 42°C, as the repressor protein is temperature-sensitive; strain PI is not temperature-sensitive.

Exercise 2: To Make a Lysate of Phage Lambda—I

A ts *E. coli* lysogen will be used as the source for the phage lambda.

Have ready

1. LB broth (120 ml)
2. NaCl
3. TE buffer:

Tris base	60.55 g
EDTA	14.61 g
Dist. water (DW)	800 ml
Adjust pH	8.2 (with HCl)
Add DW to make	1000 ml buffer

4. Polyethylene glycol (PEG) 6000
5. Water-baths at 37°C, 43°C, and 70°C
6. *E. coli* (λ) strain (ts CI_{857})
7. A Sorvall RC5 (centrifuge) with SS34 rotor

Procedure

1. Start an overnight culture of *E. coli* (λ) at 37°C in 10 ml LB broth.
2. Dilute the overnight culture by adding it to 100 ml LB broth (kept warm at 37°C).
3. Shake this culture for 2–3 hours in a water-bath at 37°C.

Check turbidity

4. Take a 10 ml turbid culture in a 50 ml conical flask (sterile).
5. Place this flask in a 43°C water-bath for 10–15 minutes.
6. Shake the above at 37°C for 2–3 hours.

The suspension should appear clear. This is your phage suspension (*lysate*).

7. Check m.o.i. (multiplicity of infection). It should be 0.1 (1 phage/10 cells).

You may now concentrate the phage suspension as follows:

1. To a 10 ml phage suspension in a test-tube, add solid NaCl to make a 0.5 M NaCl solution (29.5 g NaCl/1*l* water).
2. At 4°C (cold room), add PEG 6000 to make a 2% solution of PEG. Shake gently by hand and later with a magnetic stirrer to dissolve the PEG. This process takes 30–40 minutes.
3. Centrifuge (spin) at 13,000 g (7,000 rpm in a Sorvall SS34 rotor) for 10 minutes.
4. Add to the supernatant in the same tube 4% PEG 6000 solution to make a total of 6% PEG in the supernatant.
5. Spin at 20,000 g (13,000 rpm; rotor as above) for 30 minutes. Discard the supernatant.
6. Resuspend the pellet in TE buffer.

This is your concentrated phage suspension.

Exercise 3: To Make a Lysate of Phage Lambda-II

Have ready

1. *E. coli* host strain
2. Phage suspension from Exercise 2 or stock
3. LB broth 1.25 *l*
4. Ca-Mg solution:

$CaCl_2$ (10 mM)	1 ml
SDW	99 ml

5. $MgCl_2$ (10 mM) 1 ml
   SDW 99 ml
6. Chloroform 5 ml
7. NaCl 35 g
8. PEG 6000 80 g
9. PEG buffer:

NaCl (10 mM)	2 ml
Tris-HCl, (10 mM), pH 7.4	1 ml
$MgCl_2$ (10 mM)	1 ml
DW	96 ml

Procedure

1. Grow overnight a culture of *E. coli* in 10 ml LB broth.
2. Pour culture into a Sorvall plastic centrifuge (50 ml) tube.
3. Add 0.45 ml $CaCl_2$ and 0.45 ml $MgCl_2$ solutions to 0.1 ml *E. coli* culture.
4. Add phage suspension (1 phage/10 *E. coli*) or m.o.i. = 0.1
5. Keep tube in a water-bath at 37°C for 20 minutes.
6. Add 18 ml LB (kept at 37°C).
7. Take two 2-litre flasks- each containing 500 ml LB + 10 mM $MgCl_2$.
8. Add to each flask half of the phage –*E. coli* mixture.
9. Incubate with shaking (fast speed) at 37°C for 8–12 hours.
10. Add 2.5 ml chloroform to each flask and continue to shake for another 20 minutes.
11. Spin the contents of each flask in the cold room (4°C) at 5,000 g for 30 minutes. The pellet consists of cell debris.
12. Transfer the supernatant to a large flask (2-litre capacity).
13. Add 35 g NAC1 and shake to mix and dissolve.
14. Add and transfer to the cold room 80 g PEG 6000.
15. Mix gently by hand. If PEG has not dissolved use a magnetic stirrer to do so.
16. Let the flasks rest at 4°C overnight. This should precipitate all the phage particles.
17. Spin the flasks in the cold room at 5,000 g for 35–45 minutes. Discard the supernatant carefully with a pipette.
18. Add 10 ml PEG buffer to the pellet kept on an ice-bath and allow to rest for 30 minutes.
19. Resuspend the pellet gently with a Gilson pipette.
20. Pool resuspensions in a 50 ml centrifuge tube. Add washings with PEG buffer to this tube.
21. Allow the above to remain in the cold room (4°C) overnight.

This is your concentrated phage lambda ready for purification.

Exercise 4: To Assay Phage Lambda

The *E. coli* strain 600 (ts CI_{857}) or 594 (ts CI_{857}), Su^- will be used in this exercise. A suspension of the bacteria will be mixed with phage lysate and plated in soft agar over a hard agar medium, and incubated. The number of phages per ml of lysate is estimated from the number of plaques that appear on the petri dishes. Each plaque originates from a single phage particle.

Have ready

1. Lysate of phage lambda (from Exercise 2 or 3), kept with 2–3 drops of chloroform, at 4°C (cold room)
2. *E. coli* strain 600 or 594; plating bacteria stock
3. Tryptone agar plates:
   Hard agar:

Bactotryptone	1.0%
NaCl	0.5%
Bactoagar	1.2–1.5%

Autoclave. Pour into 80 mm sterile petri plates. Keep at 37°C in the incubator overnight. This ensures that there is no contamination before use. Contaminated material will appear as colonies on the plates.

4. *Soft agar:*

Bactotryptone	1.0%
NaCl	0.5%
$MgCl_2$ 0.01M	0.2%
Bactomaltose	0.2%
Bactoagar	0.7 0.8%

Autoclave. Maintain the soft agar in a melted condition at 45°C (in a water-bath), till use.

5. Tryptone bactomaltos medium (TBM):

Tryptone	1.0%
NaCl	0.5%
$MgCl_2$	0.2%
Bactomaltose	0.2%

6. Phage dilution medium (PDM):

Tryptone	0.10%
NaCl	0.85%
$MgCl_2$	0.20%

Autoclave.

7. Tryptone broth (TB):

Tryptone	1.0%
NaCl	0.5%

8. Plating bacteria grown overnight at 32°C to OD at A_{600} = 0.3

Procedure

1. Remove all chloroform (preservative) from 1 ml of the lysate by keeping the latter in an open tube or vial at 37°C for about 30 minutes (or till there is no smell of chloroform).
2. Shake the lysate (for mixing) for 1 minute.
3. (a) Make dilutions of the lysate (10^{-2} to 10^{9}) using PDM.
   (b) Start by making the 10^{-2} dilution:

PDM	5.00 ml
lysate	0.05 ml

   (c) Prepare a 10^{-8} dilution:

PDM	0.9 ml
dilution 10^{-6}	0.1 ml

   (d) Use for assay dilutions. 10^{-6} , 10^{-8} , 10^{-9}
4. Have the soft agar ready at 45°C.
5. Take out 0.1 ml of the following dilutions in a tube each: 10^{-6}, 10^{-8}, 10^{-9}.
6. To each of the above tubes, add 0.1 ml (2 drops) of *E.. coli* in TB.
7. Mix and keep at room temperature (25°C) for 15 minutes.
8. Dispense into each of the four sterile test-tubes (one for control) 2.5 ml of soft agar using a Pasteur pipette.

9. (a) Add to three tubes (labelled 10^{-6}, 10^{-8}, 10^{-9}) 0.1 ml of the mixture from Step 7.
   (b) Mix by tapping the tubes with fingertips.
10. Pour immediately on the surface of the correspondingly labelled hard agar plate. Spread evenly (quickly) by tilting the covered plate.
11. Keep at room temperature for 5–10 minutes (or until the soft agar is set completely).
12. Place plates (inverted) at 32°C (incubator) for 16–20 hours or overnight.
13. Count the number of plaques in each plate, including the control plate.
14. Calculate the amount of the phage (in litres) in the sample lysate.

Exercise 5: To Make a Lysate of Phage T2

Have ready

1. LBT broth: 5 ml
2. *E. coli* strain MD1767 (sensitive to T2)
3. Stock of phage T2
4. Chloroform
5. Shaker-bath at 37°C
6. Incubator at 37°C.
7. Centrifuge 8,000–10,000 rpm (Sorvall RC5) with SS34 rotor

Procedure

1. Make an overnight culture of MD1767 in a 10 ml LBT broth at 37°C.
2. (a) Take 4 ml of overnight culture and 10 ml of LBT (37°C) in a small conical flask. Grow to the exponential phase by shaking at 37°C for 2–4 hours (2×10^8 cells/ml).
   (b) Make a control flask with the overnight culture and LBT and grow it to the exponential phase.
3. Take phage T2 (of m.o.i. 0.2–0.4) lysate and remove the chloroform by keeping at 37°C, i.e., until there is no smell of chloroform.
4. Add to (a) only, a drop or two of the lysate.
5. Incubate the above and control (b) at 37°C overnight.
6. Observe the cultures (a) and (b):
   (a) Infected culture: Should be clear due to cell lysis by T2.
   (b) Control culture: Should be turbid due to the growth of unlysed bacteria.
7. Centrifuge (a) at 8,000–10,000 rpm for 10 minutes.
8. Take out supernatant and add to it 1–2 drops of chloroform. Keep in a screw-capped vial.
9. Store the lysate in a refrigerator. This lysate can be kept for many months with no loss of phage activity.

Procedure

1. (a) Seed plate *E. coli* in TBM and incubate at 37°C overnight.
   (b) Seed the *E. coli* strain into a 10 ml tryptone broth and incubate at 37°C overnight to reach the stationary phase.
2. (a) Pellet cells at 5,000 g for 5 minutes. Resuspend the pellet in 5 ml $MgSO_4$ (0.01 M).

(b) Place the *E. coli* culture from (b) in a water-bath (43°C) and shake for 15–20 minutes.

3. Transfer to a shaker-bath at 37°C and shake vigorously for 2–3 hours. Cells would have lysed by then. The broth should no longer be turbid.
4. Chill the tube with lysed cells on ice (kept in a small ice bucket).
5. Add 2–3 drops of chloroform.
6. Shake, to mix, for about 1 minute.
7. Dilute the phage suspension in a phage dilution medium. Start with a 5 ml medium (yeast-tryptone) + 0.05 ml of phage suspension. Make dilutions: 10^2–10^9.

Exercise 6: To Assay a Phage T2 Lysate

Have ready

1. Phage T2 lysate (stored at 4°C)
2. *E. coli* recipient strain MD1767 (Hfr) made competent
3. LB Broth 125 ml (for plates) + 50 ml (for dilutions)
4. LB hard agar plates 8
5. LB soft agar: 50 ml (with 0.7% agar)
6. Incubator (at 37°C)
7. Vortex mixer

Procedure

1. Take out the lysate from the refrigerator.
2. In a test-tube, take 0.3 ml of the lysate, and keep at 37°C for 30 minutes to remove the chloroform (preservative in the lysate).
3. Make lysate dilutions in LB broth:
   10^{-2}, 10^{-4}, 10^{-6}, 10^{-7} in the following manner:
   (i) Place, in each of 4 test-tubes, 4.95 ml LB broth.
   (ii) Add to tube 1: 0.05 ml of lysate (undiluted). This makes the suspension in tube 1 a 10^{-2} dilution.
   (iii) Transfer from tube 1 to tube 2 the 0.05 ml suspension to make a 10^{-4} dilution.
   (iv) Similarly make a 10^{-6} dilution.
   (v) To make 10^{-7} dilution , add 3.06 ml of LB and 0.04 ml of 10^{-6} dilution.
   (vi) Take 5.00 ml undiluted lysate in a tube.
4. Take four Eppendorf tubes and place in each 0.1 ml (2 drops) of lysate. One of the dilutions plus the undiluted lysate is in the 5th tube.
5. Add to each tube 0.1 ml of recipient culture. Now, we have
   Tube 1 = 10^{-2} + bacteria
   Tube 2 = 10^{-4} + bacteria
   Tube 3 = 10^{-6} + bacteria
   Tube 4 = 10^{-7} + bacteria
   Tube 5 = 10^{0} + bacteria
6. Mix the lysate and bacteria in each tube gently by rotating the tube between the fingers.

7. Pour 3–4 ml of contents of each tube on to two LB hard agar plates and pour out the excess mixture. This leaves about 1 ml per plate.
8. Add and spread over this in each plate 2.5 ml of soft agar (melted and cooled to 45°C).
9. Incubate plates after 30 minutes at room temperature at 37°C overnight (or 5–6 hours).
10. Observe and count the plaques per plate:
    10^0 should have a very large number of plaques.
    10^{-2} should have around 1000 plaques.
    10^{-4} should have 20–30 plaques.

N.B. The same procedure may be followed to assay litres of T_{even} (T4, T6, etc.) and T_{odd} (T1, T3, etc.) phages. The plaques of most T_{even} phages are small. The plaques of most T_{odd} phages are large. The plaques of most T_{rapid} lysis or *r* mutants of T2 are large, but not as large as those of T_{odd} phages.

6.3 ❑ WORKING WITH EUKARYOTES

The eukaryotic species used frequently are those of yeast, the alga *Chlamydomonas,* the various protozoa, including *Tetrahymena,* the nematode *Cenhorrhabditis elegans,* the slime mold, *Dictyostelium discoideum,* the fruit fly *Drosophila melanogaster,* the amphibians including *Xenopus* spp., birds, mainly, chicken, and the mammals including rodents (mouse, rat, hamsters), apes (monkey, chimpanzee, gorilla) human beings (*Homo sapiens*), as well as, lately, plant species.

Eukaryotic species normally possess diploid somatic cells. Plant cells possess a complex cellulose-based wall surrounding the cell membrane; plant cells are also cemented to each other with a substance called 'pectin'.

Progress in eukaryotic genetic engineering has been boosted by improved methods of culturing tissues and single cells and also by better methods of extracting unsheared DNA and of gene transfer into recipient cells.

We now present a brief introduction to the main species to be used in the following exercises.

6.3.1 *Saccharomyces cerevisiae*

Yeasts are unicellular fungi (eukaryotes), which are used commercially for baking and brewing. In the laboratory yeasts are very useful organisms that can be grown and manipulated with the same ease as encountered while cultivating and experimenting widi bacterial cells. Since yeasts are very much more complex than prokaryotes, they offer a convenient system for studying the DNA biology of eukaryotes. Yeasts are valuable to gene manipulators since they can be used as a cloning vessel for eukaryotic genes. Yeasts are, to date, easily transformed with cloned yeast DNA, and with that of bacteria. Attempts are afoot to successfully clone a larger spectrum of eukaryotic DNA in this species.

Yeasts are known to possess 2 μm diameter plasmids (called 2 μm plasmids) that can be used for cloning in yeast. Other heterogeneous vectors are also used successfully in yeast.

Recently, 'mini-chromosomes' with the principal features of eukaryotic chromosomes [centromere (CEN) and autoreplication sequences (ARS) and telomeric ends (Tr)] have been constructed by introducing CEN and ARS from yeast chromosomes and Tr from *Tetrahymena* DNA into plasmid vectors. These 'artificial chromosomes' behave like normal eukaryotic chromosomes during mitosis and meiosis, and thus ensure the equitable and stable distribution of genes cloned in them in eukaryotic-recipient cells. Refinements in using artificial chromosomes are promising for cloning of genes in eukaryotic species.

The yeast we will use in our exercises is baker's yeast or *Saccharomyces cerevisiae,* which thrives on carbohydrate-rich substrates. *S. cerevisiae* reproduces asexually by budding. Sexual reproduction involves the fusion of two haploid mating type cells (+ and –), which are formed after meiosis in the single celled zygote. Only the mating type *plus* and the mating type *minus* can unite to form the diploid zygote.

The *S. cerevisiae* DNA possesses a row of three genes in chromosome number 3. These are the mating type genes, namely HML, MAT and HMR. These genes have common sequences; HML and HMR possess an inserted sequence (transposon) each, $\underline{a}$ and $\underline{\alpha}$—MAT contains either $\underline{a}$ or $\underline{\alpha}$ The mating type of a haploid cell depends on the nature of the transposon it contains ($\underline{a}$ or $\underline{\alpha}$). An $\underline{a}$ or $\underline{\alpha}$ strain of this yeast divides mitotically into the parental mating type except occasionally when a small population of the opposite mating type is found in the population derived from a pure $\underline{\alpha}$ or $\underline{a}$ strain. This is due to an event of transposition by which the $\underline{\alpha}$ or $\underline{a}$ component of the MAT locus is replaced by $\underline{a}$ or $\underline{\alpha}$ sequences from the appropriate adjacent mating locuses (HML/HMR). This event is a very interesting one to geneticists and developmental biologists who try to track down initial moves during key steps in the developmental programme of eukaryotes.

Yeasts, like other fungi, possess a complex cell wall that is different from the cellulose walls of plants. To extract macromolecules (DNA, RNA, proteins) from yeast, and for introducing genes into yeast, this tough cell wall has to be removed first. This can be done by subjecting the wall to extreme temperature shock (–70°C to 37–40°C) or to fungal cell wall-digesting enzymes (such as zymolase). The naked protoplast becomes spherical and is called a 'spheroplast'.

Yeast cells, like all eukaryotic cells, possess mitochondrial DNA, which carries genes for a limited but specific number of mitochondrial proteins. Mutations in proteins involved in respiratory events cause disturbances in respiration that surface as an inability to grow well. The respiratory mutants of yeast form smaller colonies than normal yeasts. The small mutant colonies are called 'petite' mutants. There are three types of petite mutants:

1. Those due to mutation in chromosomal genes for respiration.
2. Those due to mutation in mit DNA.
3. Those due to lack of mit DNA.

The first type segregate according to Mendelian laws, and are called *segregational petites.* The other two mutants have the characteristic random segration features of extrachromosomal genes. We will study the respiratory capacities (resp. quotients) of normal and petite yeasts, as well as try to mutagenize yeast cells using UV and a chemical mutagen.

6.3.2 *Drosophila melanogaster*

Drosophila melanogaster is an insect of the Dipteran group. It has a life span of about 15 days at a temperature of 20° ± 2°C, and lives on carbohydrates (overripe fruits in nature and a pudding made with cereal, molasses, yeast and an antimicrobial agent in the laboratory).

Drosophila is the organism first utilized to work out the intricacies of heredity in higher organisms. Thomas Morgan of Columbia University, New York City, together with his students Sturtevant and Bridges, and colleagues Muller, and Dobzhansky, were the pioneers who elucidated the basis of Mendelian inheritance using this fly for their studies. The concepts of the *chromosome theory* (genes are located in a linear array on chromosomes), of sex-linked inheritance (the presence of X and Y type of chromosomes and the biased inheritance of genes carried by them) and of the *mutation theory* (genes can be altered permanently by treatment with physical agents) were established by Morgan and his well-known 'fly squad', and have been found to be universally applicable to all species of organisms.

Drosophila is a very convenient organism for genetical research, and as it turns out today, also for studies of the molecular mechanics of development. Its short life span of about two weeks, small size and negligible bother and cost of rearing, the ease with which mutations can be induced, observed and analysed have made this organism perhaps the most versatile and valuable for biological studies.

The *Drosophila* begins life as a fertilized egg which is covered by a membrane called the *chorion* with hexagonal cells. Under the chorion is a chitin-containing tough membrane, the *vitelline* membrane. The egg is fertilized by a sperm which may be stored for several days in the uterus. Hence, in genetical studies we have to make sure that the female fly is being fertilized by the male of one's choice, by ensuring that the female is still virgin. This is done by collecting female flies within two hours of appearance in a flyless bottle.

The egg cell nucleus undergoes meiosis to form haploid gamete nuclei only after the sperm enters the cell through an opening in the anterior part of the egg. The female and male haploid nuclei unite and the diploid zygote cell is formed. This divides mitotically to form the *embryo* or *larva.*

The larva is an independent, mobile unit that feeds, grows in size and develops internally until it reaches the next stage—that of the *pupa.* The mature larva becomes sluggish and eventually stops moving and attaches itself to a dry spot (on the side of the bottle). A hard coat covers the *pupa.* Within the pupa an exciting transformation takes place. This is known as metamorphosis, during which the larval structure is altered to that of the *adult* mature fly. The life cycle of *Drosophila,* therefore, consists of four stages: eggs, larva, pupa and adult.

The adult flies can mate almost immediately after emerging from the pupa. The female fly begins to lay eggs from the third day of emergence and continues to do so throughout its life. Hence, in the laboratory, when you want to study the effect of a particular mating, you have to make sure that the adult flies (parents) are removed before they come in contact with the offspring, with whom they can mate again.

The larval period is divided into three stages, called *instars.* The embryo starts to develop within the original chorionic covering. Since this covering does not stretch,

it is replaced periodically to accommodate the larger larval body. The period of exchanging the new for the old skin is known as that of *molting*. Molting occurs between each larval instar.

The three larval stages can be distinguished by the additional developed features at each stage. But in general, the three larval stages can be picked up in the following crude manner:

First instar This is found within 24 hours of placing a pair of mating flies in a bottle or a vial. These larvae are not seen outside the medium; they can be strained out of the medium.

Second instar This is found in the second 24 hours after mating (i.e., 48 hours after mating, or 24 hours after the 1st instar). These larvae crawl in and out of the medium but do not climb out on to the glass walls or on the crumpled paper inserted in the bottle.

Third instar These larvae appear about the third 24 hours after the initial mating time. The third instar larvae crawl all over the wall of the bottle and the paper and feed for about 2 days until they become fat and much longer in size than the earlier instar larvae.

The satiated third instar larvae creep up to a dry place in the bottle and become the pupa. The colour of the pupal skin changes to a deep brown and, after around 4–5 days, the adult fly emerges. The newly emerged flies are light coloured and their wings are not expanded. Soon they become darker in colour and the wings open out.

The female and male flies may be recognized by the features shown in the following table:

		Female	Male
1.	Abdominal segments that can be seen easily	7	5
2.	Markings on abdomen	Separate	Last few fused
3.	Sex combs	None	Present
4.	Shape of abdomen	Larger, pointed end	Rounded end; ten thick bristles on the first tarsal joint of the first legs.
5.	Size	Larger	Smaller

Recently, there has been a renewed interest in the developmental features of *Drosophila*. It is known that the development of an organism involves or represents a sequential expression of sets of genes. All along the developmental pathway there are key points during which the genetic information in the chromosomes of particular cell(s) become reprogrammed so that a different combination of genes, from that in the previous cells, is now expressed. These newly programmed cells (stem cells) divide mitotically until they reach another developmental landmark, when one of the cells becomes the *stem* cell (reprogrammed cell) for the next set of cell divisions.

Eventually, at the fully mature stage, several new tissues are established, each of which has originated from a stem cell somewhere in the developmental pathway.

Drosophila offers unique advantages for investigating the molecular bases for the aforementioned feature in the developmental pathway. These advantages include a way of identifying the agents that control the switch from one cell type to the next one, the molecular picture of how these switches operate, and identification of the crucial molecules or events that actually trigger off developmental switches. To understand how this fly is useful for such studies you have to become familiar with a few more gross peculiarities of the *Drosophila.*

First of all, it is easy to locate the region of the chromosome that is being currently expressed by looking for structures called 'puffs' in certain special chromosomes in glandular cells (salivary gland, pancreas, and so on). These chromosomes are called *polytene* or giant chromosomes, since they are enormously larger in diameter than the usual chromosomes. The chromosomes of the salivary gland are the largest and are, therefore, used for most studies of puff formation. The polytene chromosome results from the repeated mitosis of chromosomes that remain paired, and the mitotic products (daughter chromosomes) which do not move apart from each other. Now, when a gene is being expressed (transcribed into RNA, as the first step in expression) the two DNA strands separate, and each strand becomes decondensed (so, very long).

In a single chromosome this 'bubble'-like structure can be seen only under the high-powered electron microscope. Since the individual strands in polytene chromosomes are identical, bubble formation occurs in the same locations in all the strands. The effect is a 'puffed' appearance at these locations. So, if you see a puff in a salivary gland chromosome you can rest assured that a transcriptional activity is occurring at that region.

Another (and earlier discovered) feature of *Drosophila* salivary gland chromosomes is that the idiosyncracies in the physical appearance of each single chromosome (like centromeres, darkly and lightly stained bands, secondary constrictions and so on) become magnified and can be recognized easily in suitably stained polytene chromosomes. These band regions are unique for each of the four chromosomes of the fly genome. Morgan's laboratory had correlated each band in detail with the loci of genes studied by them. Chromosome maps are, therefore, available, which number each band in each chromosome and also indicate the gene loci represented by them. By consulting these maps you can more or less pinpoint the gene loci that are involved in the puff.

It was discovered, further, that the pattern of puffing is different at different periods of development of the *Drosophila* larva, and is specific for any one period. This observation corroborates the speculation that different sets of genes are utilized at each stage of development. U Clever had found that a switch in puffing pattern occurred at the two molting junctures: at the end of the second and third instar periods. These periods are also determined to see the advent of new molecules, without which the switch cannot occur. These are the juvenile and ecdysone hormones for transition from the second to the third instar and from the third to the pupal stage, respectively. Clever discovered that the set of puffs induced in the second instar stage is different from that of the third instar. Not only do some new

puffs emerge in the latter, but also others disappear. Clever simulated this natural event by treating young larvae, not ready to move into the pupa stage, with ecdysone (steroid). Ecdysone-specific puffs promptly appeared after the treatment and the larvae began to pupate. At least one chemical was thus identified that served as a trigger for reprogramming stem cells to express the next set of genes.

Culturing of *Drosophila* salivary gland chromosomes, as well as those of other glands and other insect species opened up a new avenue of research. One can dissect out intact salivary glands, maintain them under culture conditions and experiment on the various parameters involved in this developmental event.

A second feature of *Drosophila* development that proved useful is the occurrence of developmental freaks known as *homeotic mutants.* These have mutations by which the development of an entire appendage is replaced by another normal one. When such replacements of normal appendages by ones not meant for that position in the body are encountered in the laboratory they are called cases of *transdetermination.* It appears that homeotic mutations and transdetermination occur by similar, if not identical mechanisms. There are several well-known homeotic mutations, such as *bithorax,* where one of the abdominal segments is replaced by another thorax, and *Antennapedia,* where a leg takes the place of an antenna.

Homeotic mutations indicate at least two conclusions about development, at least in insects: (1) There are junctures in the developmental pathway where a mutation in a gene causes an entire organ to be altered, rather than just a single trait or a few traits in the expected normal organ. (2) Development occurs in 'compartments' with a key stem cell launching a particular compartment. This is Garcia-Bellido's (another pioneer in *Drosophila* developmental studies) *polyclonal hypothesis* for development.

In order to understand the basis of homeotic mutations, investigators have studied another special feature in Dipteran development. During the development of the embryo (larva), certain packages of cells on the epidermal layer are retained in a determined, but undifferentiated, state, until the signal for metamorphosis. These packages are called *imaginal discs.* These discs become embodied in the body of the larva and remain there in a particular convoluted arrangement, while the main body cells of the larva proliferate. The imaginal discs are attached to the epidermal surface by a thin tube-like connection. At the end of the second instar, the ecdysone hormone triggers off these dormant disc cells to begin differentiation. During the pupal stage this differentiation is completed. At this stage, each imaginal disc emerges to the outside of the body much as a glove is turned inside out. Each imaginal disc is differentiated into a specific cuticular (skin)-based appendage, such as the eyes, antennae, legs, halteres, wings and genital organs. There are pairs of imaginal discs, one on each side of the median line running from the anterior to the posterior end of the body.

The imaginal discs are already determined for the external appendages that they form. Even parts of an imaginal disc are committed to the programmes that they have to follow in order to develop into specific parts of an appendage.

Imaginal discs are useful for studies in developmental genetics, since here you find cells kept in an undeveloped state in a bed of developing body cells. The imaginal discs can be kept in this state as long as they have no access to ecdysone. This can be

easily achieved in the laboratory by picking out a disc from a larva and introducing it into another larva that is not about to synthesize ecdysone, or from the abdomen of one adult fly to that of another. You may transplant the disc from larva to larva for unlimited generations without causing them to differentiate. That the determined state is quite stable can be proved by transplanting a much transferred imaginal disc into a third instar larva which has ecdysone. The introduced disc will be found to start developing into its destined appendage in practically every case. Occasionally (and that is why the word practically has been used), the imaginal disc will be found to be developing into an entirely different appendage (e.g., an eye into an antenna or an antenna into a leg). Such events, when found under laboratory conditions, are called *transdeterminations.* It is as if the cue for the proper development has been replaced by a trigger for that of another appendage. This entire area of biology is extremely exciting since there is a reasonable possibility of deciphering several general developmental strategies during the delayed (and in the laboratory, manipulated) development of imaginal discs.

Perhaps the most promising of the information that has been acquired from studies of imaginal disc development is the identity of DNA regions that house the sets of sequences for each type of imaginal disc elaboration. These are the *homeotic genes,* mutations in which presumably result in homeotic mutations and transdeterminations. Homeotic gene complexes of *bithorax* and *antennapedia* mutations have been sequenced. It was found that a particular 180 bp sequence was common to both the complexes. On searching sequences of other homeotic genes, a similar sequence with more or less homology was found. These sequences have been named homeo box sequences and are believed to be responsible for regulating the pattern of expression of developmental genes for each appendage and its component parts. The idea took root that perhaps similar development-regulating key genes may be found in other animal species. This suspicion has been confirmed.

Another type of gene that became labelled as a regular developmental gene is the *protooncogene.* This type of gene was originally discovered in cancerous cells or in cells transformed by viral DNA integrated in their own DNA. Normal DNA—or rather DNA from non-cancerous cells—were probed for the presence of oncogenes. It is an astounding discovery to note that oncogenes are part of the normal repertoire of all cells and are used as developmental agents. Oncogenes have been found even in the primitive slime molds. They also carry homeo boxes. The results from different areas of research are thus converging. This will perhaps enlighten us about the step-by-step mechanics of the highly, well-managed control system that operates to develop a normal individual. This information is of interest to experts in management, servo-mechanisms, and automatic control systems, since they expect to take a leaf out of Nature's book and apply it to advantage in man-made devices and organizations.

Drosophila has been an excellent choice for studies of development, as we already know a great deal about an extraordinary number of mutations that are found in this species.

A remarkable synthesis of genetical, developmental and psysiological studies has taken off from the ground in the form of neurogenetical studies of the development of the nervous system of *Drosophila.* Once again, there are several

known neurological mutants, the genetics of which have been worked out fairly well. Combining molecular studies of development of nerve cells, their connections, excitator and inhibitory activities and the overall developmental pathway for each component of the nervous system has been a great achievement that is expected to flourish further in the not-so-distant future. These studies hope to decipher the basis of behaviour—both normal and abnormal—an information that will hopefully be profitably utilized eventually for explaining and managing aberrant behaviour in human beings.

Drosophila has been the mainstay of research in both population genetics and studies of evolution. Sewall Wright, Haldane, and Dobzhansky have been some of the key personalities who have opened up areas of population and evolutionary genetics by studying *Drosophila.* Factors such as the: large populations, small generation time and impressive spectrum of strains adapted to special ecological niches have, of course, aided such investigations.

In this book, you will be introduced to the rudiments of handling *Drosophila melanogaster* using the wild-type strain Oregon-K (OR-K). You would learn to make crosses between flies with different alleles, and chart the types and ratios of these traits in the ensuing generations. You will be introduced to culturing salivary glands, preparing polytene chromosomes for microscopic observation and isolate DNA from embryos and adult flies. To launch any investigation that requires the use of cloned DNA in *Drosophila,* you should be knowledgeable about these minimum qualifications and modes of confronting this tiny, but magnificient example in the biological world.

6.3.3 Plant Species

Green plants range in the evolutionary scale from the most primitive species of the group of algae to the most advanced ones of flowering plants.

The algae that have been used so far for genetic analysis include *Chlamydomonas reinhardii,* which are single celled chlorophyll-bearing organisms. The cells here are haploid and motile, with a pair of flagella at one end of the pear-shaped cell. They occur as two reproductive haploid strains (+ and –) which fuse to form a diploid zygote. The latter divides meiotically into four haploid cells, two each of the *plus* and *minus* strains.

The highest evolved groups of plants are the Angiosperms which bear flowers and seeds. There are two main groups in Angiosperms—the monocotyledonous and the dicotyledonous species. A cotyledon is the nutrition-packed region of a seed which is expended to nourish the embryo (plantlet) to grow until it becomes self-reliant (by synthesizing glucose via photosynthesis). Monocots possess one cotyledon per seed and dicots two per seed. Grasses, lilies, orchid and plantains are monocots, while legumes (pea or bean species), tobacco, tomato, potato, *datura* (all of the same family Solanaceae) belong to the group of dicots. Monocots and dicots differ characteristically in morphological and other features.

Plants are grown from seeds that are soaked in water at appropriate temperatures to make them begin germination. The embryo grows rapidly and the precursor of the root system (the radicle) emerges. This is followed by the development of the

precursor of the shoot system, i.e., the plumule. By the time the first or the first two or three leaves appear, the cotyledon(s) is (are) lost, or has (have) shrivelled up.

Plant cells are cemented together. The growing or undifferentiated tissues are at the tips of stems and just behind the tips of roots and their branchlets. The mature cells become differentiated into the main tissues of the stem and root systems. A cross-section of a dicot stem reveals an outermost layer of epidermal cells, followed inwards by a softer tissue called the phloem, that consists of elongated cells that distribute prepared food from leaves and other green cells to all parts of the plant. The ring of phloem tissue encloses a ring of harder tissue that forms passages for transport of raw minerals and water from the root to every part of the plant. This is the xylem tissue. In mature plants, the phloem forms the bast fibres (used commercially, e.g., jute, hemp) and the xylem the wood. In the centre of the stem is a tissue made of soft-walled cells. This is the pith region. In older trees, the pith may be eliminated by the inward growth of the xylem tissue during each growing season.

In monocots, the xylem and phloem form bundles that are arranged randomly in the interior of the stem. They do not form solid wood-like dicotyledonous trees.

Dicotyledonous species have been found to be more amenable to tissue culture than monocots, although ways of dealing with the latter are also being developed.

Plants are different from animals in several ways. Apart from the ability of plants to synthesize basic food from CO_2, water and sunlight, their mode of development is such that individual regions may be removed and used to propagate species. A wound is immediately covered by a mass of freshly divided cells that may form a tumour or scab tissue. The ability to form such scabs is exploited for developing tissue cultures from portions of a plant and from cells which have been rid of their cellulose wall and intercellular cementing pectin. One other difference from animals is that an entire adult, mature plant may be regenerated from a callus tissue or a single cell.

Such regeneration has been possible in several dicot species. Attempts are continuing to develop ways of tissue culturing monocot species, since many of the economically important plants (cereals, palms) belong to this group.

Cultures of animal and plant tissues for genetic manipulation are described at greater length in Chapter 8.

REVIEW QUESTIONS

1. Evaluate *E. coli* as a preferred cloning cell enumerating its salient features.
2. Assess the candidature of *Agrobacterium* spp. in plant genetic engineering.
3. Describe various baceriophages utilized in cloning exercises.
4. Discuss the utility of yeast as a eukaryotic cloning candidate.
5. Recapitulate the importance of *Drosophila* in hereditary studies.

7 DNA Isolation, Analysis and Cloning

Introduction

DNA isolation involves extracting the DNA molecule from the interior of the cell in a clean form. This molecule is then purified by different methods.

7.1 ❑ DNA ISOLATION

In prokaryotes, the DNA is a naked (protein-free) molecule that lies in a tightly supercoiled mass in the cytoplasm and is not surrounded by a membrane. Extraction can be effected easily by bursting the cell and eliminating all non-DNA matter from the lysed cell. Prokaryotes possess a special carbohydrate-cum-protein cell wall around each cell. This has to be first degraded by an enzyme. Lysozyme (present in egg white) is used for this purpose.

In eukaryotes the main (or genomic) DNA is in the form of chromosomes within a nucleus. To isolate nuclear DNA, it is advisable to first isolate intact nuclei and then lyse them. This prevents degradation of the DNA by cytoplasmic nucleic acid-degrading enzymes (nucleases) and by those present within the nuclei that have become active due to the rupture and resulting chaos in the organelle.

Plants, which are also eukaryotes, have cellulose cell walls in addition to the plasma membrane surrounding the cell contents. This cellulose has to be removed with a combination of carbohydrate-degrading enzymes: cellulase and pectinase. The latter degrades the substance pectin which binds plant cells together.

Yeast cells (eukaryotic single-celled organisms), classified within the plant kingdom such as chlorophyll-less thallophytes or fungi, have a complex cell wall containing chitin, like the exoskeleton of some primitive animal species. The yeast cell wall has to be removed with the enzyme zymolase (brand names: Glusulase, Driselase).

Removal of cell walls from plant and yeast cells leaves behind spherical cells that are called free or naked protoplasts or spheroplasts.

7.1.1 Bacterial Chromosomal DNA

To isolate DNA from a cell, the contents of the latter have to be spilled out of broken cells and the DNA extracted from the complex mixture of proteins, and

other macromolecules of the cell. This crude DNA is further purified and stored as a desiccated powder, or in a buffered solvent. Figure 7.1 shows a flow diagram for the isolation of high molecular weight DNA (chromosomal) from *E. coli* cells. The procedure of isolation is described in Exercise 7.

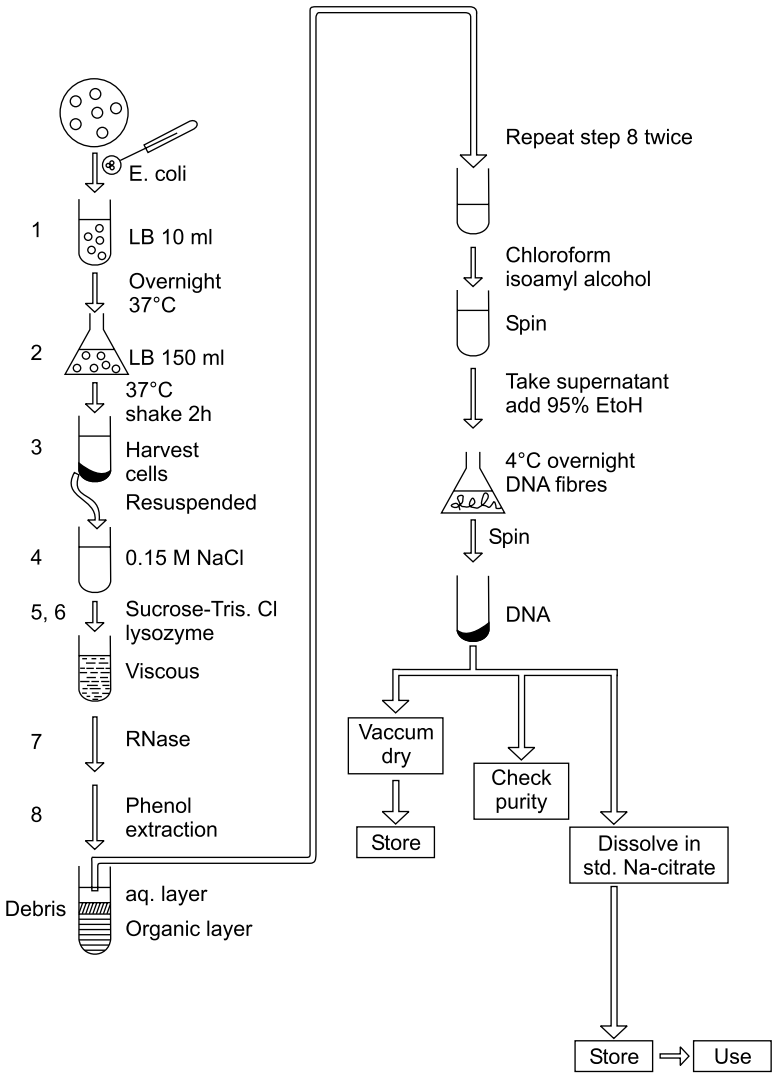


Fig. 7.1 *Flow diagram for the isolation of high molecular weight DNA (chromosomal) from E. coli.*

E. coli cells possess a cell wall surrounding the plasma membrane of the cell proper. To burst an *E. coli* both the cell wall and the plasma membrane have to be damaged. The enzyme lysozyme digests the cell wall. A combination of substances destabilizes the plasma membrane. These substances include a high molar solution of sucrose,

EDTA (chetaling agent) and SDS (sodium dodecyl sulphate, a detergent). Proteins are then removed by the addition of phenol. RNA is removed before or after protein extraction by using an RNase. This enzyme often contains DNase as an impurity, although the label may proclaim it to be pure RNase. We should take no chances and destroy the DNase activity before using the RNase. The latter is heated to 80°C in a water-bath, and allowed to cool down to room temperature by itself. If the RNase is used after the phenol extraction step, another round of phenol extraction has to be performed to remove the added protein (the enzyme) from the solution. The phenol extraction is carried out by separating the DNA from the proteins and the remaining cell contents in a 1:1 mixture of equilibriated, distilled phenol and chloroform-isoamyl alcohol (24:1, v/v). The proteins remain in the organic phenol phase, and the DNA partitions into the aqueous phase. The aqueous solution is pipetted out and the DNA precipitated from this by adding two volumes of 95 per cent ethyl alcohol. The DNA precipitates as a sluggish and thick fibre. The precipitation of DNA from the aqueous solution is a thrilling sight, particularly if the flask is held up against light!

The DNA fibres can be removed by winding them around a glass rod, or pelleted by centrifugation. The isolated DNA is then dissolved in a buffer solution or desiccated before storage or use. This DNA has, of course, to be purified to get rid of all other contaminating smaller molecules. Such purification can be achieved by a variety of methods:

1. By gel electrophoresis, followed by elution and dialysis
2. By column purification
3. By caesium chloride or sucrose density centrifugation

The third method is usually utilized when ultrapure DNA is required (as for restriction enzyme cutting).

The purity of the DNA is checked with a spectrophotometer. Double-stranded DNA, and RNA give a purity value of 1 for a wavelength of 260 nm for a solution containing 50 μg nucleic acid/ml of aqueous solution.

E. coli cells possess several kinds of natural plasmids. The ones carrying free F plasmids are F^+ cells. Those with the F DNA integrated into the chromosomal DNA are the Hfr cells (high frequency of recombination). The presence of F in *E. coli* is advertised by small hair-like projections called pili (sing: pilus). F^- cells have no pili.

Several phage parasites of *E. coli* enter the cell by first attaching the viral shell to a matching receptor protein on the surface of the pilus. Phage M13 aggregates along pili of *E. coli* cells. One may distinguish F^+ and Hfr cells from F^- cells by mixing the bacteria and M13 particles, and looking for the culture which shows clustering of M13 along the edges of the *E. coli* cells.

E. coli cells are readily infected by the phage T1. The *E. coli* strain used (C600) is immune to this phage. To be sure that the gene introduced in a clone is not transferred autonomously to another cell, it is customary to use an F^- strain, such as HB101.

7.1.2 Plasmid DNA

Isolation of plasmid DNA is essentially similar to that of chromosomal or high molecular DNA, with minor modifications that ensure good and pure yields of the very low molecular weight pDNA molecules. Figure 7.2 shows a flow diagram for the isolation of low molecular weight DNA (plasmid) from *E. coli* (pBR322). The procedure for isolation is described in Exercise 8/9.

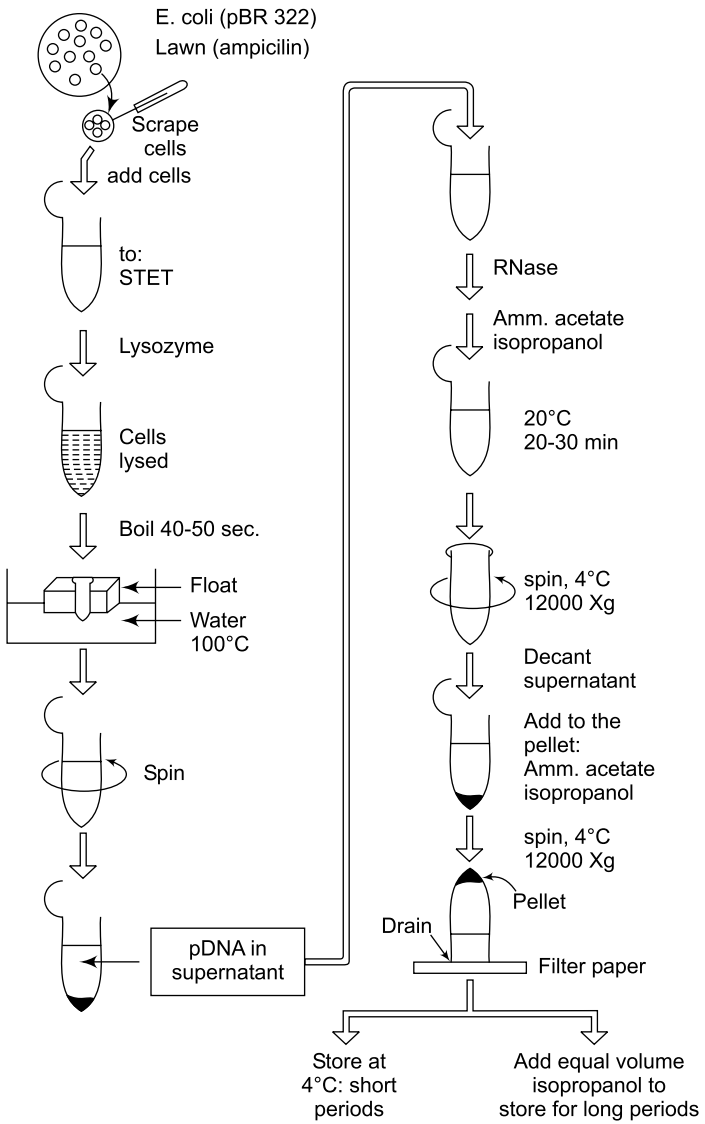


Fig. 7.2 *Flow diagram for the isolation of low molecular weight DNA (plasmid) from E. coli. pBR322 is a small plasmid vector carrying ampicillin and tetracycline resistance genes. It is isolated from the cell lysate in which the chromosomal DNA has been made ineffective for isolation by boiling. Hence this procedure is often referred to as the 'boiling method' of plasmid DNA isolation.*

The circular pDNAs are usually supercoiled. The *E. coli* plasmid pBR322 is amplifiable. Chloramphenicol is, therefore, added after the bacteria have been grown in the shaker water-bath for 3 hours and then incubated with shaking for another 20 hours.

The procedure also includes heating the cell extract to 45°C in order to denature the DNA. The longer chromosomal DNA become entangled and do not get reassociated fast enough when the temperature is lowered. The small plasmid DNA circles become easily reassociated.

The next step consists of centrifugation at a very high speed. This pellets out all high molecular DNA and other cell matter (cell membrane, wall, etc.) The lighter plasmid DNA remain in the supernatant. These are centrifuged in a CsCl gradient and extracted from the tube by puncturing the polyallomer centrifuge tube at the base of the pDNA band. The DNA band is identified by its reddish-orange glow in long wave UV light. The glow is due to an added reagent, namely ethidium bromide, that intercalates between the base pairs of DNA. On exposure to light, ethidium bromide-associated DNA promotes the breaking of the DNA backbone; hence, it has to be kept away from light until the ethidium bromide is removed from the DNA.

7.1.3 Phage DNA

We will isolate DNA from the phage lambda in an *E. coli* lysogen. Stocks of lambda can be preserved for long periods of time by keeping them in the lysogenized form. Figure 7.3 shows a flow diagram for the isolation of bacteriophage lambda DNA from *E. coli* (λ) lysogens. The procedure for isolation is described in Exercise 10.

Sufficient phage particles have to be present in order to extract a reasonable quantity of DNA. A lysogen contains only one lambda DNA per *E. coli* cell. To increase the number of lambda DNA molecules, we have to *induce* the phage DNA (allow it to deintegrate from the *E. coli* DNA) and replicate into many copies, each one of which becomes part of an individual infectious particle. A large number of free phages are harvested from *E. coli* cells, then purified and their density/ml assayed before cracking the viral protein shell to release the DNA. The phage shell can be broken by alternate exposure to a very low temperature (liquid N^2 or liquid air temperature of –180°C) and room temperature. The shell and DNA are separated out by centrifugation and the purity of the phage DNA assessed.

7.1.4 Slime Mold DNA

We shall use the slime mold *Dictyostelium discoideum* for our exercise. Slime molds are eukaryotes that live as independent amoeboid cells on damp surfaces of trees and fallen plant material that contain bacteria. When the surface dries up and/or the bacteria are used up, the cells aggregate into masses that move as sloppy slugs (called grex). After moving about for some time, if starvation conditions still persist, the grex reorganizes its cells into a highly structured stalked body with a bulbous spore container at one end and an anchoring base at the other. The lifecycle of *Dictyostelium discoideum* is shown in Fig. 7.4.

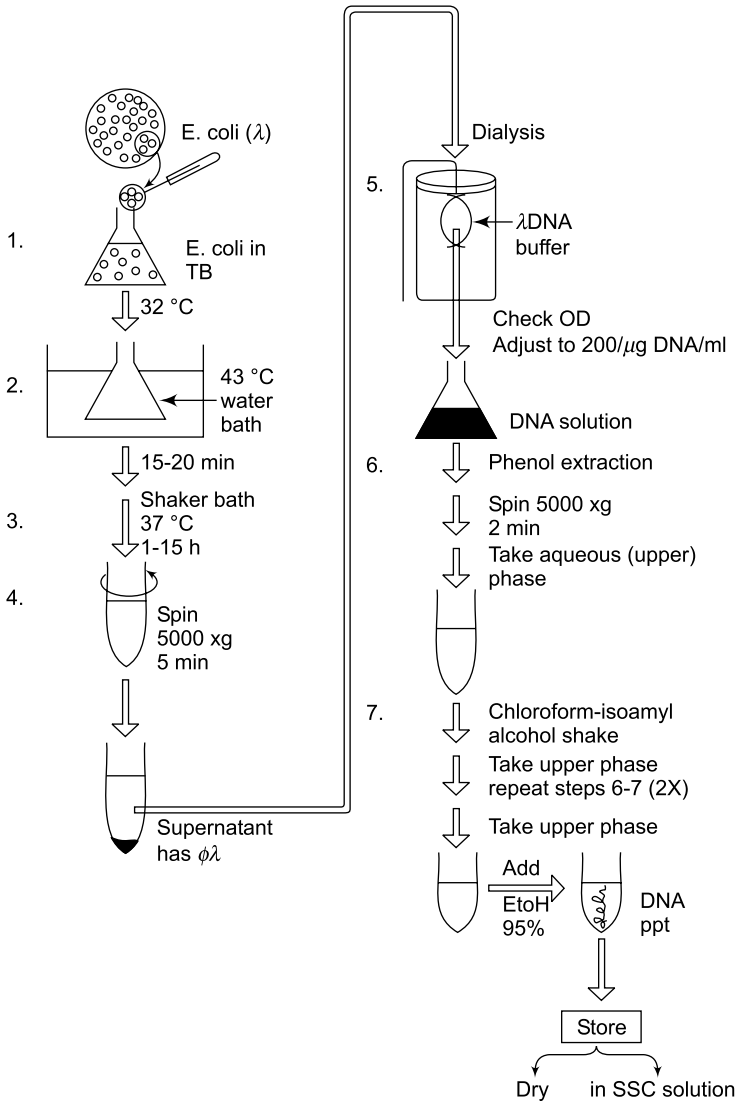


Fig. 7.3 *Flow diagram for the isolation of bacteriophage lambda DNA from E. coli (λ) lysogens. The prophage is induced at 43°C and allowed to amplify as free phages at 37°C. The phage shells are disrupted by dialysis (osmotic shock) and the DNA isolated by phenol/chloroform-isoamyl alcohol treatment and precipitated by 95% chilled ethanol.*

The best studied species of slime molds is *Dictyostelium discoideum.* The biology of slime molds is particularly interesting to developmental biologists as they undergo the basic steps of development of higher animal species (aggregation, migration, and differentiation), but with fewer frills. This organism thus serves as a good model for elucidating the molecular basis of cell-cell recognition, cell migration, determination

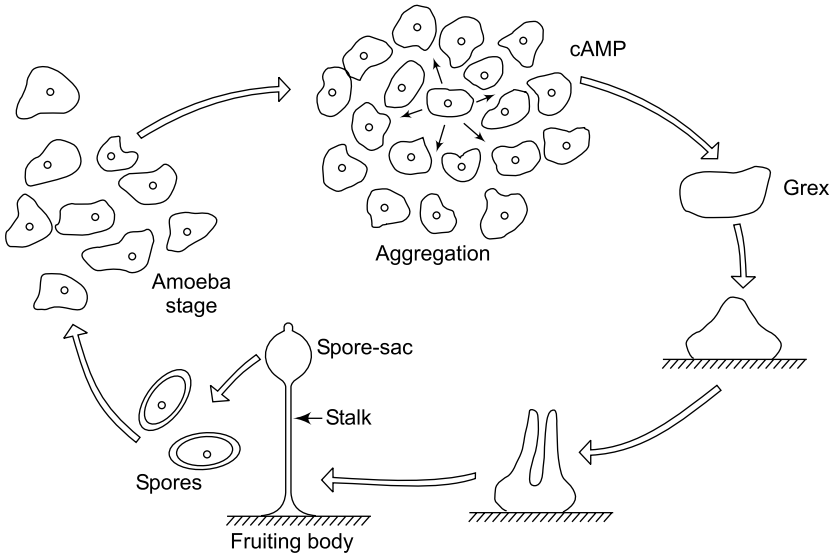


Fig. 7.4 *Lifecycle of Dictyostelium discoideum. The single-celled amoebae aggregate around a cell that begins to release cAMP when triggered by conditions of drought. The aggregated cells (grex) behave like one entity and move around like the slug of a snail. A series of developmental steps converts it into a stalked-spore sac that is anchored to the substrate. The spores, released from the sac, germinate when conditions are appropriate.*

and differentiation. Attention is especially focused on the signals and switches that determine the progress from one stage of development to another.

7.1.5 Yeast DNA

Yeast is a single-celled eukaryotic fungal cell which propagates asexually by budding, with the buds separating to form free cells. It reproduces sexually by forming four gametes per diploid cell. The fertilizing strains are called mating types *a* and α.

The yeast cell has a tough, complex carbohydrate wall that can be removed enzymatically. So DNA isolation from this species involves an initial removal of the cell wall with an enzyme known as zymolase (brand names Glusulase, Driselase) (Fig. 7.5).

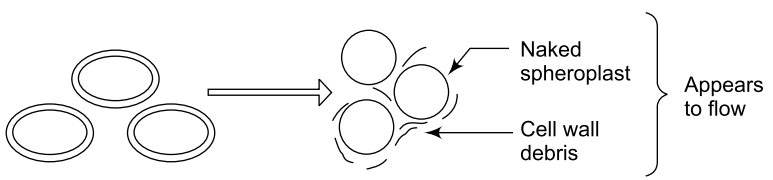


Fig. 7.5 *'Spheroplasting' of yeast cells. The enzyme zymolase degrades the characteristic wall of the yeast cell, leaving a naked cell or protoplast (also referred to as spheroplast, as the wall-less cells assume a spherical form).*

7.1.6 *Drosophila* DNA

Isolation of DNA from *Drosophila melanogaster* chromosomes is similar to that from any other animal species. In this case, embryonic cells (larval stage) are utilized as the source of DNA.

Animal cells contain low molecular weight DNA in mitochondria. To isolate chromosomal DNA from such cells it is better to restrict isolation from the nucleus. This procedure not only bars contamination of nuclear DNA with cytoplasmic DNA but an intact nucleus also prevents the degradation of the DNA by cytoplasmic and nuclear nucleases.

The first job then is to carefully isolate intact pure nuclei from the larval cells. The animal cells do not possess cell walls like bacteria and plant cells. The isolated nuclei are suspended in a special buffer and centrifuged. The nuclei are then lysed with a detergent buffer. The nuclear proteins are degraded by the protein digesting enzyme pronase. The DNA (with added ethidium bromide) is isolated by CsCl centrifugation and eluted from the band in the CsCl tube. The DNA is rid of the ethidium bromide and dialysed against a TE (tris-EDTA) buffer.

Exercise 7: Isolation of E. coli Chromosomal DNA (I)

1. Seed *E. coli* from single colonies from a streaked plate: 1 loopful of cells in 10 ml LB. Incubate at 37°C overnight.
2. Add overnight culture to 150 ml LB (37°C) in a 500 ml flask. Incubate at 37°C for 2 hours.
3. Add the above culture to 3 litre LB (37°C) in 6 × 1000 ml flasks. Incubate with shaking in a water-bath at 37°C for 4–6 hours.
4. Keep the flasks at 37°C in the incubator overnight.
5. Check OD_{600} next morning. The suspension should be very turbid with A_{600} = 0.7 or more (1×10^8 cells/ml OD 0.7).
6. Harvest cells in the Sorval centrifuge (4°C) at 8,000 rpm for 10 minutes.
7. Resuspend the pellets in each tube and pool the suspensions in one centrifuge tube. Spin once more to pellet the cells.
8. Resuspend the pellet with 0.15 M NaCl. Spin to pellet.
9. Resuspend the pellet with sucrose-tris HCl and pour into a beaker.

In cold room (4°C)

10. Add to the above 60 ml lysozyme. Incubate for 5 minutes.
11. Add and mix gently 120 ml EDTA. Incubate for 5 minutes.
12. Add and mix gently 12 ml SDS soln. You will now have a clear viscous solution.
13. Add and mix gently 48 ml RNase solution. Incubate at 37°C for 20 minutes.
14. Transfer the liquid to glass centrifuge tubes.
15. Add to each tube equal volumes of phenol.
16. Add also chloroform-isoamyl alcohol.
17. Shake by hand at room temperature for 20 minutes.

18. Spin in the table-top centrifuge (Remi R4C) at 4,000 rpm for 30 minutes. There will be three layers (see Fig. 7.6).
19. Remove the supernatant (aqueous layer) carefully with a Pasteur pipette and place in a sterile flask (500 ml).
20. Add chloroform-isoamyl alcohol. Layer and spin and collect the supernatant in the flask.

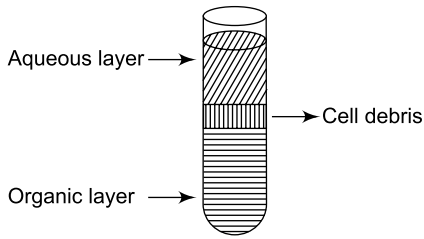


Fig. 7.6 *Three layers after spinning the table-top centrifuge*

21. Repeat Step 20 once more (total three times).
22. To the aqueous solution in the flask, add 2 volumes of 95% ethanol.
23. Aspirate the mixture gently with a Pasteur pipette. The DNA fibres are precipitated out from the solution.

Follow Step A for one-half of the isolated DNA and Step B for the other half.

Step A

1. Wind the DNA fibres around the end of a glass rod and transfer these into a centrifuge tube having a little of 95% ethanol or spin in Remi R4C at 4,000 rpm for 10–15 minutes. Discard the supernatant.
2. Pool the pellets using 95% ethanol.
3. Add some more ethanol and spin.
4. Vacuum dry the DNA pellet in a desiccator for 10 minutes or till it is quite dry.
5. Cover the mouth of the tube with parafilm and aluminium foil and store at 4°C. Label the tube with your name, date of isolation, and type of DNA [e.g., *E. coli* (C600) chromosomal DNA]; or dialyse the solution against TE buffer and check purity of DNA in the dialysis bag.

Step B

1. Wind the DNA fibres onto a glass rod and transfer these with a little dil. Na-citrate solution to a glass tube.
2. Add dil. Na-citrate twice the volume of DNA soln. Allow the fibres to dissolve completely.
3. Add to the above cone. Na-citrate soln. until the concentration of standard Na-citrate solution is attained. (Std. Na-citrate = 0.15 N NaCl, 0.015 M trisodium citrate, pH 7.0).
4. Repeat the preceding Steps 14–22, and Steps 1–3 of B *twice.*
5. Repeat the preceding Steps 14–22 and Steps 1–2 of B. The DNA is now in 9 ml of dil. Na-citrate.
6. Add 1 ml Na-acetate-EDTA.
7. While stirring with a glass rod, add ice cold isopropanol to 2/3 of the volume of liquid after the above step. DNA fibres will spool out on the rod.
8. Pass the spooled DNA fibres successively through 75%, 80% and 95% ethyl alcohol.
9. Allow the DNA to go into solution by keeping the spooled DNA for 24 hours in a test-tube containing dil. Na-citrate solution.

Exercise 8: Isolation of pBR322 from E. coli

Have ready

1. *E. coli* strain with pBR322—3 litre culture at the exponential phase
2. Chloramphenicol solution (for 3 litre culture): 250 mg chloramphenicol and 1 ml ethanol
3. TE buffer 1–100 ml: 50 mM Tris (pH 8); 15 mM EDTA; and 100 mM NaCl
4. Lysozyme soln. 50 ml (5 mg/ml): 250 mg lysozyme and 50 ml sterile water
5. Triton X-100 (detergent) solution:
   (i) Make stock solutions:
   Buffer; 50 mM Tris
   15 mM EDTA
   pH 8
   Detergent: 1 g Triton X–100
   100 ml sterile water
   (ii) Make 0.4% solution of Triton X–100 using the above two stock solutions.
   RNase (A) soln.: 1 ml
   Stock: 5 mg/ml (kept at –20°C)
   To use RNase solution:
   First heat RNAse stock to 80°C in a water-bath and let it cool down to room temperature on its own (*do not cool rapidly* under tap water).
7. Equilibriated phenol: see Appendix.
8. Choloroform-Isoamyl alcohol (24:1 v/v)
9. Na-acetate: 50 ml of 3 M soln.
10. Isopropanol (undiluted)
11. Sorvall RC-5B, Refrigerated Superspeed Centrifuge

Procedure

1. Grow *E. coli* (pBR322 strain)—3 litre culture to exponential stage. Check OD_{600}.
2. Amplify pBR322 with chloramphenicol treatment, as follows:
   (i) Add 1 ml chloramphenicol to the culture.
   (ii) Shake at 37°C for 17 hours.
   (iii) Check OD_{600}.
3. Harvest the cells in a table-top centrifuge (Remi R4C) at 5,000–6,000 rpm for 10–15 minutes.
4. Add to the pellet 60 ml of TE buffer.
5. Add: 50 ml lysozyme soln.
6. Shake at room temperature (30–37°C) for 20 minutes.
7. Add: Triton X-100 soln. Keep at room temperature for 10–15 minutes. The suspension will become thickish due to cell lysis.
8. Spin at 17,000 rpm (Sorvall SS34 rotor, 30,000 g) for 30 minutes.
9. Transfer the supernatant to a glass centrifuge tube.
10. Add: 1 ml RNase A (2 mg/ml). Keep at room temperature for 30 minutes.
11. Add to the above an equal volume of phenol. Shake carefully by hand.
12. Spin at 10,000 rpm (HB4 rotor) or at 15–20,000 g. Two layers separate out in the tube.

13. (a) Transfer the upper aqueous layer into a fresh tube.
    (b) Add: an equal volume of chloroform-isoamyl alcohol.
    (c) Shake by hand.
14. Spin at 15–20,000 g. Transfer the upper aqueous layer to a fresh tube.
15. Add: 3 M Na-acetate (1/10 of upper layer taken)
16. Add: an equal volume of isopropanol. This precipitates out DNA only. (Ethanol precipitates out both RNA and DNA.)
17. Keep at –20°C overnight. The DNA precipitates out by this time.
18. Spin in Sorvall (HB4 rotor) at 10,000 rpm for 30 minutes (at 5–6°C). The pellet contains pDNA.
19. Discard the supernatant by decanting.
20. Dry the pellet in a vacuum desiccator for 10–15 minutes.
21. Resuspend the pellet and dissolve it in 0.1 ml TE buffer 2.
22. To store the pDNA:
    (i) For a short period: Cover the vial with parafilm and aluminium foil and keep it in a cold room (4°C).
    (ii) For a longer period: Add an equal volume of isopropanol and keep in a screw-capped vial in the refrigerator (not freezer).

    Expected yield:
    Ideally, 1 litre culture should give 1 mg DNA. Actually 1 litre culture gives about 0.1 mg DNA.

Exercise 9: Rapid Isolation of Plasmid DNA

This exercise will be used for isolating pBR322 from *E. coli* cells. It may also be utilized for isolating other plasmids from *E. coli* and other bacterial species. This pDNA may be used for transformation or restriction analysis.

Have ready

1. LB plate with streaked *E. coli* (pBR322)
2. STET:
   (a) Sucrose 8%
   (b) Triton X–100 5%
   (c) Tris 50 mM pH 8.0
   (d) EDTA 50 mM
3. Lysozyme (egg white) 5 mg/ml (freshly made)
4. RNase 5 mg/ml (heated to 80°C and cooled)
5. Ammonium acetate 5 M and 0.3 M
6. Isopropanol 100% and 70%
7. TE:
   (a) Tris 10 mM
   (b) EDTA 1 mM pH 7.4
8. Tracking dye: Bromophenol blue

Procedure

1. Scrape 1 loopful of cells from the plate.
2. Mix the loopful of cells in 0.3 ml of STET in an 1.5 ml Eppendorf tube.

3. Add about 0.3 ml of lysozyme solution and allow the tube to rest for 5–10 minutes at room temperature. Cells should be lysed by now (the mixture should be quite viscous).
4. Treat the lysate at 100°C for 40–50 seconds (float the tubes in boiling water by fixing in slots made in a slab of styrofoam).
5. Spin in the Eppendorf microfuge for 10 minutes so that the liquid settles down and the chromosomal DNA and cell debris are deposited at the bottom of the tube. The pDNA is left in the supernatant.
6. Remove the supernatant with a pipette and place it in a fresh tube. (You may check the presence of the DNA by running the solution in a minigel.)
7. Add RNase; about 0.5 mg/ml, final concentration.
8. Add to the above:
   Ammonium acetate (5 M) 2 ml
   Isopropanol (100%) 1 ml
9. Keep at –20°C for 20–30 minutes.
10. Spin in the microfuge at 4°C (in a cold room) for 10 minutes.
11. Decant off the supernatant.
12. Add to the pellet:
    Ammonium acetate (0.3 M) 2 ml
    Isopropanol (70%) 1 ml
13. Spin in the microfuge at 4°C for 5 minutes.
14. Decant the supernatant and invert the tube over a clean filter paper, to drain off all liquid. Dry the pellet.

Exercise 10: Isolation of Phage Lambda DNA (I)

We will use an *E. coli* lysogenic strain for phage λ.

Have ready

1. *E. coli* (λ) strain, overnight culture grown in TB at 32°C
2. TB (tryptone broth): 100 ml
   (a) 1% tryptone 1% soln.
   (b) 0.5% NaCl 0.5% soln.
   Add dist. water to make a 100 ml broth.
3. Water-bath at 43°C and shaker-bath at 37°C
4. Dialysis tubing no. 20
5. Lambda DNA buffer:
   (a) 0.01 M Tris-HCl pH 8.0
   (b) 0.1 M KCL
   (c) 10^{-4} M EDTA
6. Phenol, redistilled and equilibriated with 0.2 volume of 1 M Tris-HCl, pH 8
7. Chloroform-isoamyl alcohol (24:1 v/v)
8. 95% ethanol
9. Vacuum desiccator
10. Dil. Na-citrate solution for dissolving extracted DNA (see Appendix).

Procedure

1. Grow *E. coli* (λ) at 32°C to OD = 0.3 in TB (in a flask)—100 ml.
2. Place the culture flask in a water-bath, with shaking, at 43°C, for 15–20 minutes.
3. Transfer the flask to a water-bath, with vigorous shaking, at 37°C for 1–1$\frac{1}{2}$ hours.
4. Fill the dry centrifuge tubes with the culture from Step 3.
5. Spin in Sorval GSA or SS34 rotor at 5,000 g for 5 minutes.
6. Decant the supernatant into a sterile flask. (This would contain the phage DNA.)
7. Fill dialysis tubing with the supernatant and immerse in a lambda DNA buffer at 4°C for 1 hour. The solution should be viscous due to the bursting of phage shells.
8. Check the concentration of DNA in Spectronic-20 (spectrophotometer) 50 µg/ml DNA = OD at A_{260} = 1.0.
9. Adjust the concentration of DNA to 200 µg/ml with lambda DNA buffer.
10. Pour DNA solution from the dialysis bag into a sterile conical flask.
11. Add to the DNA in the cold room (4°C) an equal volume of the phenol (equilibriated).
12. Mix DNA and phenol gently by shaking for about 25 minutes at 4°C.
13. Pour the mixture into glass centrifuge tubes and spin at 5,000 g for 2 minutes.
14. Remove the upper phase, with a bent Pasteur pipette, into another tube (Fig. 7.7).
15. Add to the upper phase an equal volume of the phenol.
16. Spin at 5,000 g for 2 minutes. Transfer the upper phase into a fresh tube.
17. Add to the above an equal volume of chloroform-isoamyl alcohol. Shake and allow separation into distinct phases.
18. Remove the upper (aqueous) phase to a fresh tube.
19. Repeat Steps 17–18 [two times more].
20. Transfer the upper phase to a small flask or tube.
21. Add 95% ethanol (twice the volume of the upper phase fluid).
22. Shake to precipitate out the DNA with minimum breakage.
23. Store the DNA in a dry form or in a dil. Na-citrate solution.

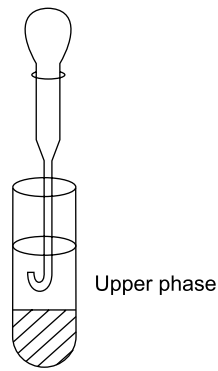


Fig. 7.7 *Removing the upper phase with a bent Pasteur pipette*

Exercise 11: Isolation of Agrobacterium Tumefaciens Ti Plasmid

Have ready

1. *Agrobactrium tumefaciens strain*
2. Medium 523:

Yeast extract	4.0 g
K_2HPO_4	2.0 g

$MgSO_4$	0.15 g
Casein hydrolysate	3.0 g
Sucrose	10.0 g
Twice dist. water	1000 ml
pH	7.1

3. Buffer (stock 20 X):

Tris base	0.8 M
Na-acetate	0.1 M
Na_2 EDTA	0.02 M
pH	7.9 (use gl. acetic acid to adjust)

For use: 1 X solution (1 ml stock/19 ml water).

4. Solution for lysing cells (SLC):

Tris base	3.6 g (50 mM)
NaOH soln.	24.6 ml (2 N). (Freshly made before use).
SLS or SDS	18.0 g (3%)

Twice dist. water to make 500 ml

5. Tracking dye:

Bromophenol blue	0.25%
Glycerol	50.0%
Tris-acetate	50 mM (pH 7.9)

6. Phenol:chloroform mixture (1:1, v/v)

Distilled Phenol	50 ml	(Stored at 4°C.)
Chloroform	50 ml	[Kept at room temperature (25°C).]

7. Eppendorf (microfuge: speed 12 or 13 krpm) centrifuge and tubes.

Procedure

1. Make an overnight culture at 28–30°C in 10 ml of Medium 523.
2. Innoculate 100 ml Medium 523 (29°C) with the overnight culture and shake at 29°C for 5 hours.
   Check cell density: should be 10^8/ml.
3. Distribute into each of eight Eppendorf tubes: 0.5 ml of culture.
4. Spin in the Eppendorf for 1 minute.
   Discard the supernatant.
5. Add to each pellet: 1 ml of buffer solution.
   Mix by a vortex mixer.
6. Add to each tube: 2–3 ml of the SLC soln.
   Mix by inverting the tube with fingers 10–15 times.
7. Keep at 55°C for 90 minutes.
8. Add to each tube: 4 ml phenol:chloroform.
   Mix by inverting about 100 times.
9. Spin at 25°C for 3 minutes.
10. If lysis is not complete (i.e., if there is no clear layer between the upper and lower phase), there are too many cells in the culture. Therefore, if not fully lysed, begin all over again from step 3 after diluting the culture.
    If lysed well, proceed to Step 11.

11. Transfer the upper aqueous phase (containing pDNA) to a vial (screw-capped) and store at 4°C until use.
12. Check the quantity and quality of the pDNA (≃ 300,000 bp) by gel electrophoresis.

Gel electrophoresis to screen extracted large-size plasmid DNA

1. Use 0.7% agarose gel.
2. Use the buffer prepared for this exercise for making the gel and for filling the buffer tanks in the apparatus.
3. Combine with each on a parafilm square:
   0.5 ml of pDNA
   0.05 ml of tracking dye.
4. Load wells in agarose gel with the DNA/dye mixture.
5. Run the gel at 12 volt/cm of gel slab length for 3–3$\frac{1}{2}$ hours.
6. Add to the gel: ethidium bromide soln.
7. Remove excess stain by washing many times with twice distilled water.
8. Record DNA bands by ultraviolet photography.

Exercise 12: Isolation of DNA from D. Discoideum

The biology of slime molds is particularly interesting to developmental biologists since it undergoes the basic steps of development of more complicated higher animal species, but with far fewer frills. Since basic bilogical principles of all species are essentially similar, if not identical, *Dictyostelium* offers a good system for understanding the molecular basis of cell–cell recognition, cell migration, determination and differentiation. It is expected that molecular studies of this species will identify the developmental infrastructure (signals and switches) that are essential for progressing from one stage of development to another.

Outline of procedure

A. Culturing *E. coli* on LBA; *E. coli* is the food for the slime mold cells.
B. Culturing *D. discoideum* (slime mold).
C. Reducing *E. coli* contamination. This reduction has to be done before isolating DNA in order to avoid impurity of bacterial DNA in the extracted slime mold DNA.
D. Harvesting slime mold cells.
E. Isolation of DNA.

Have ready

1. *E. coli* stock on LB plates or stock
2. Slime mold stock
3. LB broth
4. PBS
5. Hemocytometer
6. SMA plates (SM broth + agar soln.)
7. Glass-spreader (L-shaped glass rod)
8. Brown wrapping paper

9. Streptomycin
10. BOD at 22°C
11. TEN buffer

Procedure

A. Culturing E. coli on LBA

1. Add *E. coli* scraped from streaked LB agar (LBA) plates or tubes (stored stock);
   Scraped cells 1 loopful
   LB broth 10 ml
2. Incubate at 37°C overnight.
3. Add the following material:
   Overnight culture 2 loopfuls
   LB broth 100 ml
   Incubate with shaking at 37°C overnight.
4. Divide the suspension equally into two centrifuge tubes (Remi table-top centrifuge tubes).
   Spin at room temperature at 2,500 rpm for half an hour.
5. Resuspend the pellet (cells) in PBS 5 ml (per tube).

The cell count should be about 10^{10}/ml.
These cells may be stored at 4°C for 2–3 weeks.

B. Culturing D. discoideum Slime mold stocks are stored at 4°C as (a) spores or (b) fruiting bodies on plates (sealed). They can be stored for about a month.

1. Spores (stock) or pick from fruiting bodies using a platinum loop.
2. Add from stock to PBS spores or fruiting body, head regions carefully picked with a platinum loop.
   Count cells/ml with a hemocytometer: 2×10 spores/ml.
3. Add in a 1:1 ratio:
   E. coli cells (in PBS)
   Spore suspension (in PBS)
4. Vortex the mixture thoroughly.
5. Pour on each SMA plate (dry agar surface) 0.25 ml of the mixture. Spread cells with a glass-spreader.
6. Cover plates with brown wrapping paper.
7. Incubate at 22°C (BOD):
   Cells aggregate about 48 hours after seeding.
   Fruiting bodies appear 24 hours after aggregation.

*C. Reducing E. coli contamination**

1. Pour on each plate, just as cells begin to aggregate, 5 ml PBS broth + streptomycin (0.5 µg/ml broth).
2. Keep for 3 minutes.
   Scrape off cells into the PBS with a spreader.
3. Pipette out the cell-PBS suspension from every 2 plates into one test-tube containing 5 ml PBS-streptomycin.
4. Incubate tubes at 22°C (BOD) for 2 hours.
5. Spin to pellet cells.

The antibiotic will prevent bacterial proliferation but not affect that of *Dictyostelium,* nor prevent the latter from ingesting the *E. coli.* After this incubation, there should be a considerable reduction of *E. coli* cells in the suspension. Further reduction will occur during the repeated washings of the harvested cells.

* Try the following after Step 4 of Section D:

D–5 • Resuspend the pellet in 10 ml PBS-streptomycin.
• Incubate at 22°C for 2 hours.

D–6 • Spin the discarded supernatant.
• Resuspend the pellet in 3 ml PBS (ice-cold).

D–7 • Repeat D–6 until the supernatant is very clear.

D–8 • Resuspend the final pellet in TEN buffer.

** Also repeat both the above alternatives using SM-streptomycin (0.5 μg/ml broth) instead of PBS-streptomycin.

D. Harvesting Dictyostelium cells Harvest cells when aggregation just begins.

1. Add to each plate 5 ml PBS (ice-chilled). Let it remain so for 3 minutes.
2. Scrape off cells gently from the agar surface with a glass-spreader. The cells will remain suspended in the PBS added in Step 1.
3. Pipette out the cell-PBS suspension into Remi centrifuge tubes (table-top Remi), and add 5 ml suspension/tube.
4. Spin in the table-top centrifuge (Remi) at position 3 at 4°C (in cold room) for 3 minutes.
5. Discard the supernatant. Resuspend the pellet (cells) in 3 ml PBS (ice-cold).
6. Repeat Steps 4 and 5 until the supernatant is very clear (1–3 times).
7. Resuspend the final pellet in TEN buffer.

E. Isolation of DNA All operations should be performed using glass (Remi) centrifuge tubes.

1. Suspend the pellet in TEN buffer.
   Count cells: should be 2.5×10^7 cells/ml.
2. Add 0.5 M EDTA to the above to give a final concentration of 25 raM EDTA.
3. Add 10% SDS stock to the above to give 1% SDS in the liquid. Mix gently by shaking the tube with hand and fingers.
4. Add to the above an equal volume of equilibriated phenol.
5. Spin at position 4 (Remi table-top) at 4°C (cold room) for 10 minutes.
6. Transfer the aqueous phase (upper) with a Pasteur pipette (or decant carefully) into a fresh Remi tube.
7. Repeat Steps 4, 5 and 6 once more.
8. Add to the upper phase of the last step 5 M ammonium acetate to give a final concentration of 2 M ammonium acetate.
9. Add to above 2 volumes of absolute ethanol (ice-cold).
10. Keep the above at –20°C (deep freeze) overnight.
11. Spin at position 6 (Remi) at 4°C for 20 minutes.
12. Decant the alcohol (supernatant). Add fresh absolute alcohol, and decant again. If the pellet breaks up, spin, to pellet, again, in fresh alcohol.
13. Resuspend the washed pellet in 2–3 ml of TE buffer (or less). Keep at 4°C until the DNA dissolves.

14. RNase treatment: Add RNase to the above to give a final concentration of 25 μg/ml sample. Incubate the mixture at 37°C for 1 hour.
15. Add immediately (to stop enzyme action) an equal volume of equilibriated phenol. Shake for 10 minutes (by hand). Spin (position 4, 4°C, 10 minutes) in a table-top centrifuge.
16. Repeat Step 8 (ammonium acetate).
17. Repeat Steps 9–12 (alcohol-precipitaton).
18. Resuspend the final pellet in 2–3 ml of TE buffer (or less).
19. Check purity of DNA with UV spectrophotometer (Spectronic 20). Measure OD of sample at 260 nm (for nucleic acids) and 280 nm (for protein). A_{260}: A_{280} = 1.8 or > 1.8 is considered good for most studies.
20. Store DNA in a small vial or Eppendorf tube at 4°C.

Exercise 13: Isolation of Saccharomyces cerevisiae (Yeast) DNA

Outline of strategy

A. Culture yeast cells.
B. Harvest cells.
C. Make spheroplasts.
D. Isolate DNA.

Have ready

1. YPD medium for yeast:
   (i) Mix Dextrose 209 g
   2 X dist. sterile water 100 ml
   (ii) Mix Yeast extract 10 g
   Peptone 20 g
   2 X dist. water to make 900 ml
   Autoclave
   (iii) Mix (i) and (ii)
2. Pretreatment solution: Make 50 ml

β-mercaptoethanol	0.1 M
Tris-HCl	0.2 M
EDTA	0.1 M
Sorbitol	1.0 M

3. TEN buffer:

Tris-HCl	10 mM pH 7.6
EDTA	1 mM
NaCl	10 mM

4. Enzyme solution:

Zymolase	2.5 mg
TEN buffer	5.0 ml

(Zymolase solution = 0.5 mg zymolase/ml and 5000 units/g zymolase)
5. SDS or SLS 25% in dist. water
6. Chloroform:iso- (or n-) amyl alcohol mixture
7. Ethanol 95%

8. RNase Working solution: Make 1 ml
   (i) Na-acetate 0.1 M soln. in dist. water
   RNase 10 mg/ml of Na-acetate soln.
   (ii) Add to the above EDTA 0.3 mM.
   (iii) Adjust to pH 4.8.
   (iv) Heat to 80°C and cool slowly to room temperature (25°C).
9. –70°C freezer or dry ice
10. Pronase (proteolytic enzyme): Make 1 ml of solution
    Pronase 2 mg/ml of TEN buffer
11. Phenol—redistilled and equilibriated
12. Chloroform

A. Culturing yeast cells

1. Innoculate 10 ml YPD with 1 loopful yeast from stock.
2. Incubate at 30°C overnight.
3. Add:
   10 ml overnight culture to 1000 ml YPD (liquid) medium
   (or 2.5 ml/250 ml YPD in four 1-litre flasks).
4. Incubate the yeast culture with shaking at 30°C for 4–5 hours
   (or till stationary phase is reached).

B. Harvesting cells

1. Transfer the yeast culture into 50 ml polypropylene centrifuge tubes and weigh each tube.
2. Spin at 7,000 rpm (Sorvall RC5) at 4°C for 8 minutes.
3. Decant the supernatant.
4. Weigh each tube and pellet (wet weight).

There should be about 10 g of cells from the 1-litre culture.
The following procedures are based on 10 g yeast cells.
Adjust proportions of materials for amounts other than 10 g cells.

5. Suspend the pellet (pooled in one tube) in 60 ml of 2 × dist. water.
6. Spin in Refrigerated Sorvall at 7,000 rpm (6,000 g) at 4°C for 5 minutes. Discard the supernatant.

C. Spheroplasting

1. Resuspend the pellet in 45 ml of the pretreatment soln.
2. Add to the above 2 ml of zymolase soln.
3. Incubate with shaking at 30°C for $1\frac{1}{2}$ hours.
4. Spin at 4–5,000 rpm (4000 g) at 4°C (Refrigerated Sorvall) for 5 minutes.

The pellet can be stored at –20°C until use. If being used immediately, continue with the following steps:

5. Resuspend the pellet in 40 ml of pretreatment soln. Spin as before for 5 minutes.
6. Check the extent of spheroplast formation under a microscope.

Place a drop of the soln. on a glass microscope slide and observe under low and medium high power.

Most cells should be naked by now.

D. Isolation of DNA You may lyse the spheroplasts by spinning 10 ml of pellet suspension in TEN + glass beads (0.45–0.50 mm diam. acid-washed beads: "Glass Perlen" B. Braun Melsungen) and continue directly from Step 6 to 7 or begin with Step 13. The pellet should be almost pure chromatin.

With glass beads

- Add enough glass beads to make a slurry of the solution.
- Shake (in cold room) for 10–15 minutes.
- Check every few minutes to observe progress of spheroplasting (one drop of soln. covered with coverslip on a microscope slide).
- Spin in a Remi Research Centrifuge at 7,000 rpm (4,000 g).
- Pipette out the supernatant with a large pipette and transfer to a small conical flask.
- Distribute the supernatant into ultracentrifuge tubes. Weigh and accurately balance pairs of tubes.
- Spin in Beckmann's ultracentrifuge (model L5–65, L5–50BH) at 30,000 g for 30 minutes, allowing vacuum to reach > 100 μb during the first 15 minutes.
- Keep at –20°C for 30 minutes.
- Load suspension over 2.6 M sorbitol.
- Spin at 80,000 g for 1 hour.

7. Add to spheroplast suspension:
   TEN equal vol. suspension
   SDS 2 ml (final concen. 1%)
   RNase soln. 0.5 ml
8. Mix carefully by tilting the tube by hand.
9. Incubate at 37°C for 2 hours.
10. Add Pronase (1 ml).
    Incubate at 37°C for 2 hours.
    Take out and rotate the container by hand to mix reagents.
11. Heat the above in a water-bath at 65°C for 30 minutes. Let it cool to 25°C (room temperature)
12. Add to above an equal vol. of phenol: chloroform (1:1). Mix gently (to avoid shearing of DNA) for 20 minutes.
13. Distribute the mixture into glass centrifuge tubes and spin at 7,000 rpm (Remi table-top) at 4°C (cold room) for 35 minutes.
14. Transfer top layer (aqueous phase) to fresh glass tubes.
    Add an equal vol. of the chloroform:n-amyl alcohol (24:1) (or isoamyl).
    Mix for 15 minutes by hand.
    Spin to separate phases.
15. Add to the above 5 M NaCl – 1/25 vol. of the above soln.
16. Transfer all aqueous phases to one 250 ml flask.
17. Add ice-cold 95% ethanol to twice the vol. of the above soln.
    Keep the flask in an ice-bath (in cold room) for 5 minutes mixing by hand intermittently.

The DNA will be precipitated now.

18. Spool out the DNA with a glass rod.
19. Transfer the DNA into a flask (100 ml) with TEN 50 ml.
20. Add RNase solution (0.1 ml). Keep at 37°C for 1 hour.
21. Add an equal vol. of phenol:chloroform (1:1).
    Mix by rotation by hand for 15 minutes.
    Spin to separate phases (as in Step 13).
22. Add to the aqueous phase 5 M NaCl — 1/25 vol. of aqueous phase.
23. Add ice-cold 95% ethanol to twice the volume of the above soln.
    Keep in ice-bath for 5 minutes as in Step 17.
24. Spool out the DNA with a glass rod.
25. Dissolve in 3 ml TEN.
26. Add to a small screw-cap vial:
    Chloroform 0.1 ml
    DNA in TEN solution
27. Check the purity of DNA. A_{260}/A_{280} should be 1.8 or more.

Exercise 14: Isolation of Drosophila melanogaster (Fruit Fly) DNA

Drosophila chromosomal DNA is confined within the nucleus, as in all eukaryotic species. To isolate this DNA, care has to be taken to first isolate the clean intact cell nuclei. If the nucleus is lysed within the cytoplasm, the DNases present in the latter would destroy the DNA to be extracted. There are also DNases in the nucleus, but in the intact condition they appear to be prevented from acting on the nuclear DNA by perhaps the combined effects of other nuclear molecules. The DNA in the nucleus is closely associated with proteins to form the nucleoprotein fibres known as chromatin. Five histones are involved in the formation of chromatin, four of which (in 2 molecules each) form a spherical body around which the ds DNA winds 1.75 times to form structures called nucleosomes.

This thin (30 Å diameter) fibre is organized by repeated hierarchies of coiling into compacted structures called chromosomes. Chromosomes also contain a large and variable population of non-histone proteins. All these associated proteins have to be removed to obtain pure DNA.

The method to be used here will attempt to isolate intact nuclei which will then be lysed and cleaned out of all material except the DNA. The operations have to be carried out very carefully so that the high molecular weight DNA is not sheared (fragmented) by this process of isolation.

Before isolating the nuclei, the chorion of the embryo has to be removed.

Have ready

1. *Drosophila melanogaster* mating pairs (several) per milk bottle-culture
2. *Drosophila melanogaster,* wild-type strain (OR–K). Have several bottles in which fresh eggs (less than 22 hours) have been laid.
3. Fine nylon net cloth to strain embryos from fly-food
4. Mortar and pestle (kept in the cold room)
5. Triton X-100
6. Saline 0.7%—make 1.5 litres.

7. Sodium hypochlorite stock, 5.25% in water. For use: Dilute 4 fold.
8. $CaCl_2$ 10 mM
9. Hexylene glycol
10. Cyclohexyl aminopropane sulphonic acid (CPS)
11. Sucrose
12. Tris-HCl pH 8
13. EDTA 1 mM
14. Enzyme solution: with Pronase
    Mix Tris-HCl 10 mM
    EDTA 1 mM
    ] T–H–E
    Add Pronase 5 mg/ml
    Keep at 37°C (incubator) for 60 minutes.
    Do not store this solution beyond 1 day.
15. Buffer for lysing (BL):
    Detergent (SDS) 5% solution
    EDTA 0.5% pH 9.2
16. Buffer for dechorinating embryos (BDE)
    To make 10 X stock soln.
    (i) Add to a sterile flask (1 litre) in the following order:
    CPS (to make 200 mM)
    $CaCl_2$ (to make 10 mM)
    Hexylene glycol (to make 2 M)
    (ii) Add dist. water to make 1000 ml.
    Store at room temp.
    Solution for use: 1 X pH 9.2. Make 1 litre.
    Store in the uppermost shelf of the refrigerator.

Procedure

1. Prepare 20 culture bottles with $\frac{1''}{2}$ depth of food.
2. Into each bottle let loose
   10 virgin females, and
   10 males.
   Keep at 24 ± 1°C for 18–20 hours.
3. Remove parent flies into another set of 20 bottles (for additional embryos if needed).
4. Strain out the food from each bottle through the nylon cloth, using lots of water (4–6 litres), and stirring.
5. Wash the embryos on the cloth with the following mixture (T-N):
   Triton X–100 0.01%
   NaCl 0.7%
6. Weigh the embryos. Take 50 g of embryos to continue.
7. Wash the embryos with sodium hypochlorite (soln. for use) keeping them in the solution for 2–3 minutes.
8. Drain out the above solution and now wash with T N.

9. Wash with BDE to remove chorionic membrane using about 100–150 ml BDE.
10. Drain out as much water as possible by placing the nylon cloth with embryos on a double layer of blotting paper.

In the cold room Have the following in the cold room in ice-baths for keeping them ice chilled, before beginning the next steps. (A deep tray filled with crushed ice is ideal.)

1. BDE
2. Beaker (250 ml)
3. Sucrose 50% in BDE
4. BL
5. Enzyme solution
6. CsCl
7. Ethidium bromide

N.B. Work rapidly to prevent shearing of DNA.

11. Transfer the embryos to the mortar and pestle and grind vigorously. At intervals, look at samples of the ground material under the microscope to assess the number of embryos dechorionated. (Broken chorions give the embryos a transparent appearance). Dechorionate all embryos (90–95%).
12. Add to the ground embryos 100 ml BDE. Stir to allow the membranes and other debris to become suspended in the fluid.
13. Strain the suspension through a nylon cloth into the beaker. Lift the nylon cloth to form a bag of suspension and press with fingers to allow all the fluid to be filtered through nylon into the beaker.
    Wash out the mortar and pestle with another aliquot ($\simeq$ 100 ml) of BDE and filter this also, with the filtrate in the beaker.
14. Transfer the filtrate to suitable tubes for the Sorvall centrifuge. Spin at 13,000 rpm at 2°C for 20 minutes.
15. Discard the supernatant. Suspend the pellet in 10 ml BDE (ice-cold).
16. Make BDE-sucrose solution: 150 ml (BDE–S)
    Add to the suspension BDE–S 135 ml. Mix gently by hand.
17. Spin at 13,000 rpm at 2°C for 60 minutes. Discard the supernatant.
18. Cool the pellets using ice-cold BDE (4 ml/pellet).
19. Add to the cooled suspension BL 6 ml (ice-cold).
20. Add to the above enzyme soln. 1.6 ml and keep this at 50°C for 1 hour.

The suspension now contains DNA and degraded proteins. The DNA can be isolated from this by the usual phenol extraction method. But that much handling breaks up the long DNA molecules into small fragments. To avoid this, it is better to remove the DNA by CsCl gradient centrifugation.

For CsCl gradient centrifugation of the above lysate:

21. Add to the enzyme-degraded lysate:
    CsCl 22 g
    Ethidium bromide 1.6 ml (15 mg/ml)
    Refractive index should be 1.403 at 20–25°C.
    Spin in the Beckmann apparatus for 45 hours in an angle rotor (42 krpm).
22. Dialyse DNA against lots of TE (pH 8).

Exercise 15: Isolation of Plant-Cell DNA

We will use wheat (or gram) embryo for this exercise.

Have ready

1. Wheat 500 grains (4–5 g) or gram: 178 seeds (20 g)
2. Saline-EDTA (Sal-E):
   NaCl 0.15 M
   EDTA 0.10 M
   pH 8.0
3. SDS 20% in water
4. Sodium perchlorate 5 M stock soln.
5. Chloroform:Isoamyl alcohol (24:1, v/v)
6. Absolute ethanol
7. *Saline-citrate solution:
   NaCl 0.15 M
   Na_3 citrate 0.01 M
   pH 7 ± 0.2
   * maintains ionic strength of dissolved DNA and chelates divalent ions.
8. Mercuric chloride soln. in water (0.1%)

Procedure

1. Soak about 500 grains of wheat (gram: 20 g) at 37°C overnight (in a large beaker, covered with aluminium foil).
2. Switch on the UV lamp in the Laminar Flow Cabinet for 30 minutes, before use.
3. Wash the grains in the following solutions:
   (i) tap water + Teepol
   (ii) tap water, several washes
   (iii) Sterile dist. water (SDW)
4. In the Laminar Flow Cabinet (LFC), surface sterilize the seeds with mercuric chloride (0.1%) for 2 minutes. Make sure that all the seeds are submerged in the solution.
5. Wash in SDW (several changes).
6. With a knife and forceps, cut out the embryo from each seed.
7. Place the embryos in a petri dish, lined with a wet filter paper.
8. Weigh the embryos, transferring them to a clean petri dish.
9. Transfer the seeds (with the water) to a 50 ml conical flask (sterile). Decant the water.

10. Add to the seeds Sal-E. Swirl the flask to wash seeds well. Decant the Sal-E.
11. Add 9.8 ml SDS soln.
12. Add sodium perchlorate soln. 2.70 ml.
    Swirl the flask.
13. Add 12 ml chloroform:isoamyl alcohol (24:1, v/v).
14. Shake at room temperature for 30 minutes.
15. Spin in the cold room (4°C) in a Remi clinical centrifuge at 4,500 rpm for 15 minutes.
16. Pipette (gently) the supernatant into a tube.
17. Layer twice the volume of supernatant gently along the side of the tube of absolute alcohol (chilled).
18. Leave the tube at 4°C overnight.
19. Spool out the DNA fibres with a chilled glass rod.
20. Transfer the DNA-wrapped end of the rod into a small conical flask containing Na-citrate soln.
21. Check the purity of DNA.

7.2 ❑ RNA ISOLATION

RNA molecules occur as different species in several distinct classes. The most abundant class consists of (i) ribosomal or rRNAs. In some eukaryotic cells they may represent more than 80 per cent of the RNA species. The remainder including typically smaller molecules consist of (ii) transfer or tRNAs, (iii) several types of very small RNAs (snRNAs, 7S RNAs) that take part in the process of regulation of gene expression, and (iv) the messenger or mRNAs that are the only ones which are translated into polypeptide products. The kind and size of mRNAs vary from a few hundred to several thousand bases. In any one cell, normally only up to 5 per cent of RNA molecules are mRNAs, as only a particular small set of genes are expressed at a time in a cell. In eukaryotes, certain cells specialize in the overproduction of one specific protein, and hence are found to be packed with large amounts of mRNA expressed from the genes of those proteins.

From the above description of the heterogeneity of RNA molecules in an organism, it is obvious that methods had to be developed to isolate and purify the RNA molecules of choice. One may desire to extract the total RNA in a cell (prokaryotic or eukaryotic), only the cytoplasmic ones in an eukaryotic cell, or the ones confined within the nuclear membrane. One may specifically require to purify only one particular species of RNA whatever its class. Different techniques are growing in number which make such fishing expeditions feasible. Figure 7.8 shows the flow diagram of one method for isolation of total RNA from cells.

No matter what the technique for the isolation of RNA molecules, there is one aspect that has to be heeded, and that is to protect the RNA to be isolated from degradation by nucleases that act on them. In the intact cell, the RNA species are kept stable by a variety of gimmicks, including complexing of base-paired secondary structures with proteins. During the process of isolation these protective assemblies are destroyed, and the naked RNA molecules become vulnerable to degradation. *In vitro* isolation methods, therefore, ensure the absence of all such ribonucleases.

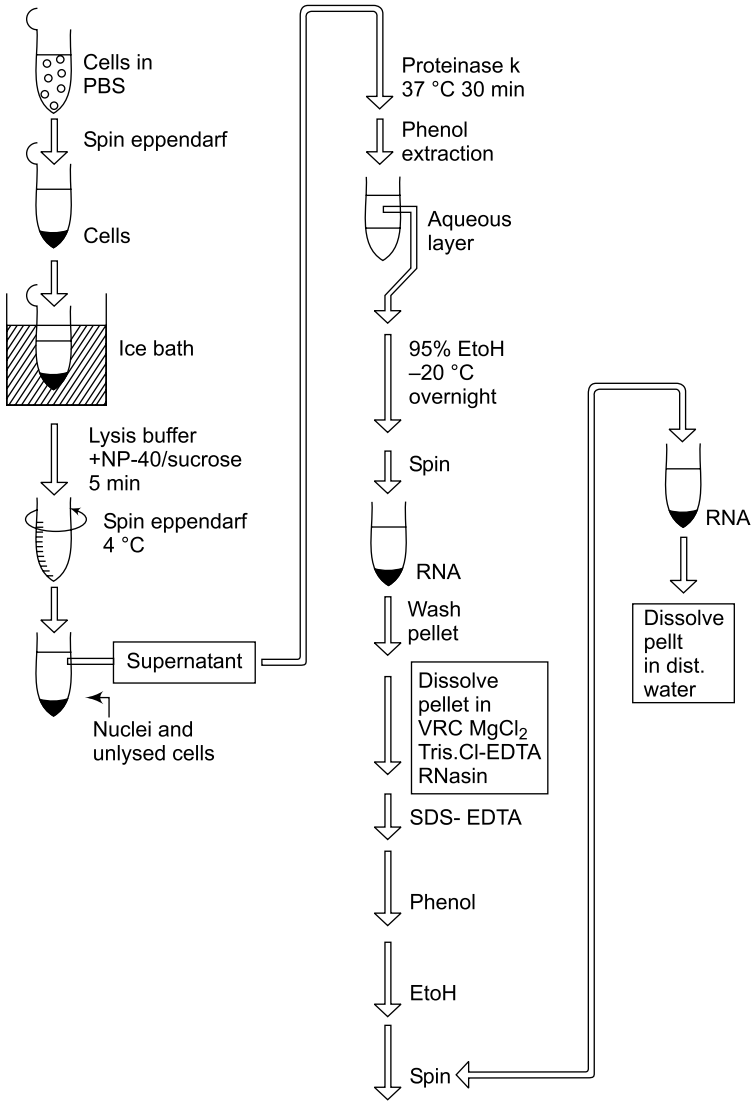


Fig. 7.8 *Flow diagram for the isolation of total RNA from cells.*

A basic precaution, therefore, is to have all apparatus, including fingers, free of ribonucleases. Glassware are baked (250°C for 4 hours) and solutions are autoclaved and touched only with hands covered with sterile gloves. Those solutions that cannot be sterilized by autoclaving are made ribonuclease-free by adding special proteins. Some protocols for RNA isolation utilize ribonuclease inhibitors, such as ribonucleoside-vanadium complexes and RNasin. The former are transient analogues of ribonucleosides on which the enzymes start to act. So, if complexes of vanadium ion and the four different ribonucleosides are added during the RNA isolation procedure, the RNA molecules are recovered unharmed and are capable

of undergoing translation. RNasin, a 40 kilo-dalton protein present in rat liver and other mammalian tissues also protects RNA from degradation and inactivation. RNA extracted in the presence of RNasin can be utilized for translation purposes in the test-tube (cell-free system).

RNase action may be thwarted by another method. There are compounds that facilitate proteins to go into solution. When cells are treated with these compounds the proteinaceous structures become degraded and proteins bound to nucleic acids liberated and destroyed. Cox developed a protocol that uses two such compounds: guanidinium isothiocyanate and guanidinium chloride. Chirgwin and associates have modified this procedure by adding a reducing agent (β-mercaptoethanol) to a 4 M solution of the isothiocyanate compound. Both these protocols and the variations that have been further introduced, ensure die extraction of clean and active mRNA.

The task of isolating mRNAs from eukaryotic cells is somewhat simplified by the presence of poly-dA tails of mRNAs destined to be translated. An oligo (dT) cellulose column may be used to trap the poly-(dA) tails of the mRNAs. The latter are eluted from the column, through which the extracted RNA solution has been passed several times.

The type of RNA molecule to be isolated is determined by the objective of the project. One may require the mRNA of a particular gene for analysis, translation or what have you. It may be desired to analyze the detailed sequence and secondary structure of the ribosomal, transfer or snRNAs for the understanding of their involvement in a particular function. RNA molecules may serve as sources of templates for synthesizing cDNAs. Parts of an RNA or the entire molecule may be utilized as a probe for searching sequences in the DNA involved in coding and control sequences. Spliced eukaryotic mRNAs derived from split or interrupted genes are useful to map the sites and extents of introns in the complementary DNA strand. The radioactively tagged spliced mRNA is hybridized *in situ* to genomic DNA, and viewed after autoradiography with the electron microscope. The presence of introns in eukaryotic genes was revealed by using such a procedure. Processed mRNAs of globin (mouse) and ovalbumin (chicken) genes, aligned against the complementary DNA strand showed up one and seven introns respectively.

Methods for the isolation of RNAs vary with the source and abundance of that particular species. Some cells are specialized to overproduce a characteristic protein product (e.g., globins, immunoglobulins, myosin, keratin, fibroin). These cells naturally abound in the mRNA for the relevant product. Isolation of total RNA from such cells followed by fractionation methods, leads to a very high yield of only one kind of mRNA.

Another method of RNA extraction relies on immunological techniques. During the period of active translation of mRNA, it is studded with ribosomes (the entire complex known as a polysome), with graded lengths of newly synthesized polypeptides. Antibodies against these protein chains, (or better still anti-antibodies against these nascent polypeptides) may be used to complex with the antigen (or antigen-antibody complex) and precipitated. The precipitated complexes are freed of extraneous matter (ribosomes, tRNAs, polypeptides) to yield a pure population of one type of mRNA.

Of course, separation of RNA species by electrophoretic fractionation is also utilized wherever a profile of RNA species in a cell or tissue or detection of a particular molecular weight species is desired.

The amounts of mRNA recovered may be determined in several ways. One method uses the mRNA to obtain translation products. The translation is performed with known variables *in vitro,* using cell-free systems, or in a living test-tube. Wheat germ and reticulocyte lysates have been found to provide the necessary machinery for cell-free translations of mRNA. The best used living test-tube is the oocyte of the African toad, *Xenopus laevis.* The egg cell of this amphibian has been found capable of tackling the translation of all exogenous mRNA used so far.

Reassociation kinetics of RNA-DNA hybrids is also exploited to assess the amount of RNA that has been purified from a tissue. The RNA is allowed to hybridize with dissociated DNA and the extent of hybridization is estimated from RNA-reassociation values, which are determined by different methods. One method is similar to that of obtaining $C_o t$ values for DNA reassociation, and consists of determination of optical values of solutions containing single-stranded and double-stranded nucleic acids. The hybridized region may also be isolated by degrading the single-stranded unhybridized regions of the DNA, using the S1 nuclease. A further device is to trap the hybrid regions to hydroxyapatite columns, that have affinity for double-stranded nucleic acids. The ss DNA and RNA pass out of the column. The RNA-DNA hybrids are eluted from the column and the dissociated strands cleaned of the DNA by treatment with a DNase. The RNA is purified by phenol extraction and alcohol precipitation.

Exercise 16: Isolation of Slime Mold Total Cellular RNA (I)

The slime mold, *Dictyostelium discoideum* will be used for this exercise. RNA will be extracted from cells in the three stages of development of the slime mold—spores, stalks of fruiting bodies, and slugs.

Outline of procedure

I. Grow *D. discoideum* to spore-sac bearing stage.
II. 1. Collect samples of the following for RNA isolation.
   (a) spores
   (b) stalks
   (c) slugs.
   2. Prepare each sample for RNA isolation. The preparation will be in suspension in a lysis buffer.
III. Isolate RNA from the above suspension.

I. To grow *D. discoideum*

Have ready

1. *D. discoideum* stock
   (a) Spores (preserved at 4°C in 0.5 ml PBS in a 3 ml vial)
   (b) Plates containing fruiting bodies
2. *E. coli* culture (in LB broth after 4 hours of shaking at 37°C)
3. Prepare SMA medium.

4. SMA plates (SM broth + Agar)
   (i) Heat sterilize the paper-wrapped petri dishes in the oven at 90°–120°C for 2 hours.

IN THE LAMINAR FLOW CABINET (LFC)

   (ii) Pour the melted SMA into plates and leave the UV light on while the medium solidifies (about 30 minutes).
   (iii) Wrap the plates with brown paper and keep in the LFC. The plates may be kept at room temperature (22–25°C) for a week. The plates can also be kept for a longer period, after wrapping in aluminium foil, at 4°C.
5. Glass T-spreader
6. Eppendorf tubes 2.5 ml
7. A 15 ml vial containing PBS (sterile)
8. Eppendorf tubes with PBS 1 X (sterile)
9. Gilman's 100 µl pipettes (P200)
10. Innoculation loop (Platinum)

Procedure

I (a) Growing from spores

1. Vortex the spore stock to resuspend the spores in the PBS in the stock vial.
2. Remove 100 µl of the above suspension and place it in a 2.5 ml Eppendorf tube.
3. Add to the above 200 µl of the *E. coli* suspension.
4. Transfer the above mixture (with a Gilson pipette) to the centre of an SMA plate and spread with a glass-spreader.
5. Cover the plates and keep in the BOD at 22°C. (This temperature is critical for the first 24 hours).

At the end of the first 24 hours the cells will be at the aggregation stage.
At the end of 48 hours fruiting bodies will be present.

You are now ready for Step I.

I (b) Growing from cultures on SMA plates

1. Flame the loop of the innoculation needle and dip it in the PBS in the vial. A transparent film of PBS will form across the loop.
2. Touch the film to about 15 spore-sacs. Avoid touching the medium to prevent picking up *E. coli* cells. The spore-laden film will appear clouded and yellowish.
3. Dip the spore-laden PBS film into 200 µl of PBS (1 X) kept in an Eppendorf tube.
4. Add to the above 200 µl of the *E. coli* suspension.
5. Transfer the mixture (with a pipette) to the centre of an SMA plate. Spread the mixture with a glass-spreader.
6. Cover the plate and incubate in the BOD at 22°C for 48 hours, (or until, fruiting bodies fill the plate).

You are now ready for Step II.

II. To collect samples of slugs, stalks and spores

Have ready

A.

1. Nylon cloth (fine mesh, < 25 μm size)
2. Pasteur pipettes
3. PBS (sterile)
4. Eppendorf tubes

B.

1. Innoculation loop
2. Eppendorf tubes each with 1 ml PBS (1 X)

C.

1. PI000 Gilson pipette
2. Extra PBS (I ×) in a flask
3. Nylon filter (25 mm size)
4. Beaker (50 ml)
5. Rubber band.

For A, B and C.

1. Ice-bath
2. TE lysis buffer, containing Vanadyl Ribonucleoside Complex (VRC) and the detergent NP40 (Nanodet P40)*

* Amount of NP40 (5%) to be used per 2.5 μl of TE lysis buffer:

A. 5 μl (slug)
B. 20 μl (spore)
C. 15 μl (stalk)

Procedure

A.

1. Squirt PBS (1 X) with a pipette on the slugs in the plate (after about 24–30 hours from the time of plating) to dislodge the slugs from the surface of the plate.
2. Pour the suspension through a filtering device made by tieing a fine nylon fabric (double sheet, weight of fine lawn) over a beaker.
3. Scrape the slugs from the filter into an Eppendorf tube containing PBS (1 X).
4. Spin at 3,000 g to pellet.

This is now ready for Step III.

B.

1. Dip innoculation loop in a vial of PBS (1 X) to form a film.
2. Pick up 20–25 spore-sacs (these usually contain around 3×10^8 cells/ml).
3. Spin to pellet the spores.

This is now ready for Step III.

C. Use the culture plates from which spore-sacs have been removed.

1. Use a P1000 pipette to squirt PBS (1 X) on the culture plate. This action dislodges the stalks, which then float in the PBS.

2. Prepare a filtering device as in the case of A, but by using a nylon filter (25 μm size) instead of nylon cloth.
3. Depress the surface of the filter and wet it with some PBS (1 X).
4. Pour the stalk-suspension through the filter. The stalks will remain on the filter.
5. Using an innoculation loop, scrape off the stalks and by holding the loop over the mouth of an Eppendorf tube dislodge the stalks into the tube by squirting PBS (1 X) with a Pasteur pipette.
6. Spin the suspension to pellet the stalks.

This is also now ready for Step III.

III. To isolate total cellular RNA

N.B. All apparatus and material used for RNA isolation should be strictly free of RNase contamination. Gloves should be worn, and test tubes, pipettes, pipette tips and so on should be handled with flamed forceps.

1. *To make glassware RNase-free*
   Bake clean glassware at 180–250°C for at least 4 hours.
2. *To make test-tubes, pipettte-tips RNase-free*
   Place test-tubes and pipette-tips in beakers and cover them with aluminium foil or use plastic containers having tight-fitting covers and autoclave for 20 minutes at 15 psi.

Have ready

1. Ice-bath
2. TE lysis buffer (kept standing in the ice-bath)
3. Eppendorf tube fixed in a adaptor for spinning in a refrigerated centrifuge (Sorvall RC5)
4. Protein kinase buffer (kept at 22–25°C)
5. Proteinase K
6. Phenol, equilibiriated as for DNA isolation
   or
   Phenol:chloroform (1:1, v/v).
7. Eppendorf tubes
8. Phenol:chloroform:isoamyl alcohol (25:24:1)
9. 95% ethyl alcohol (4°C)
10. Vacuum-Desiccator
11. Pasteur and Gilman pipettes

Procedure

1. Stick tubes from A, B, and C in the ice-bath.
2. Add to each tube 200 μl of lysis buffer and resuspend the pellet by vortexing.
3. Gently introduce into each tube, along the inner wall, 200 μl of lysis buffer with NP40 1% sucrose 24% (w/v).
4. Leave on ice for 5–6 minutes.
5. Spin at 4°C in the Sorvall at 10,000 g.

The pellet consists of nuclei and unlysed cells. The supernatant contains RNA bound to VRC, cytoplasm and some sheared membranes.

6. Transfer the supernatant (quite slimy) to another Eppendorf tube and add an equal amount of Protein kinase buffer (2 X).
7. Add to the above about 1/2 mg of Proteinase K and mix well (makes about 200 μg/ml).
8. Keep at 37°C for 30 minutes.
9. *Now follow one of the following procedures:*
   No. 1
   1. Add equilibriated phenol.
   2. Vortex for 2–3 seconds.
   3. Spin for 3–5 minutes.

   No. 2
   1. Add phenol:chloroform (1:1).
   2. Vortex for 2–3 seconds.
   3. Spin for 3–5 minutes.

   No. 3
   1. Add phenol:chloroform:isoamyl alcohol (25:24:1).
   2. Vortex for 2–3 seconds.
   3. Spin for 3–5 minutes.
10. Transfer the aqueous phase to a fresh centrifuge tube and add chloroform:isoamyl alcohol. Let the layers separate out.
11. Transfer the aqueous phase from the above to a fresh tube and add $2\frac{1}{2}$ times its volume of 95% ethyl alcohol (chilled).
12. Keep the tube at –20°C overnight. The RNA will be precipitated out.
13. Spin the tube for 10 minutes at 5,000 g, at 0°C.
14. Pour off the supernatant completely.
15. Add a little of the following:
    Ethyl alcohol 75%
    Sodium acetate 0.1 M, pH 5.2
    Slosh around in the tube and decant the fluid. This washes the pellet.
16. Add to dissolve the pellet:

Tris. Cl	50 mM, pH 7.5
EDTA	1 mM

17. Add to the above solution to give the following final concentration:

$MgCl_2$	10 mM
VRC	2 mM

18. Add to the above solution to give the following final concentrations:

EDTA	10 mM
SDS	0.2 mM

19. Add an equal volume of phenol:chloroform (1:1).
20. Transfer the aqueous phase to a fresh tube.
21. Add the following to the above solution to give a final concentration of 0.3 M Na-acetate, pH 5.2
22. Now add 2 volumes of ethyl alcohol and keep at 4°C overnight.
23. Spin briefly to pellet the precipitated RNA.

24. The RNA may be dissolved in 100 µl sterile distilled water.
25. Test the purity of the RNA in a spectrophotometer at 260 nm.
26. Run the RNA solution in a 1% agarose gel in a horizontal electrophoresis apparatus. Ethidium bromide (3 drops/100 µl solution) is added to the agarose solution before the gel is set. Bromophenol blue (1 drop) is mixed with each 20–30 µl sample of the RNA solution, and electric current (12 volts/cm of gel length) is turned on for 35–40 minutes. The RNA bands can be seen in long wavelength UV light.

Exercise 17: Isolation of D. Discoideum Total Cellular RNA (II)

Have ready

1. Extraction Buffer:

NaCl 2.5 mM	4.6 g
Na-acetate 2.5 mM	2.9 g
Double dist. water	1000 ml
pH	7.5

   Autoclave.
2. Water-saturated Phenol:
   Redistill commercial phenol. Collect distillate under water (chilled). To 550 ml of this distillate add 70 ml of m-cresol and 0.5 g of 8-hydrozy quinoline. Shake well, add some more water and store at 4°C. This can be kept for a month or two.
3. SDS Solution 20%:

SDS	20 g

   Double dist. water to make 100 ml
4. Polypropylene (opaque) screw-capped centrifuge tubes 50 ml
5. Sorvall, with SS34 rotor
6. –70°C freezer
7. Vortex mixer
8. Pasteur pipettes
9. Absolute ethyl alcohol
10. Chloroform
11. Sterile double dist. water

Procedure Use RNase-free tubes, gloves, etc., as in Exercise 16.

1. Harvest cells ($5–10 \times 10^7$/ml) in the centrifuge tubes (8,000 rpm in Sorvall).
2. Keep the tubes in the –70°C freezer (1 day to several days).
3. Remove the tubes from the freezer, wipe off the external moisture.
4. Add 10 ml extraction buffer.
5. Vortex 5 seconds each, 2–3 times.
6. Add 500 ml SDS. Mix first by gentle rocking, then by vortexing.
7. Add 10 ml phenol. Shake by inverting the tube several times, until the solution is turbid and yellow.
8. Spin at 1,000 rpm for 10 minutes at 4°C
9. Transfer the aqueous phase to a fresh tube, using a Pasteur pipette with the delivery end sawn-off.

10. Add 5 ml extraction buffer to re-extract the organic phase. Shake well.
11. Spin at 1,000 rpm for 10 minutes at 4°C.
12. Combine both the aqueous phases. Add 15 ml phenol and re-extract as above.
13. Collect the aqueous phase (~4 ml) in a fresh tube.
14. Extract twice with 15 ml chloroform.
15. Spin as in Step 8 but for 2–3 minutes to separate the phases clearly.
16. Collect the aqueous phase in a new tube.
17. Add about two and a half times the volume of absolute ethyl alcohol. Keep for more than 12 hours at 20°C. The RNA will precipitate.
18. Spin at 1,000 rpm for 45 minutes at 4°C. The RNA pellets out.
19. Decant off the alcohol. Drain off the remaining alcohol by inverting the tube over a thick filter paper (2 layers).
20. Add 500 μ*l* of sterile double dist. water and keep on ice (to dissolve) for 30 minutes.
21. Read OD at 260 nm.
    OD at 260 nm ≈ 42 μg RNA/ml
22. Store the RNA at –20°C.

7.3 ❑ PROTEIN ISOLATION

Proteins are the most important 'enabler' molecules in the living system. They may be enzymes, hormones, antibodies, various factors required in DNA and cell metabolism, as well as material required to build the structure of an organism.

One class of proteins that interests molecular biologists is that which includes various non-histones binding to the DNA as factors which influence the expression or regulation of expression of genes, especially in eukaryotes. Such proteins are extracted from intact nuclei. The cells are first lysed in an appropriate lysis buffer, which preferably contains a non-ionic detergent such as NP40.

The nuclei are spun out of lysed cells and treated with a buffer that extracts nuclear proteins, including histone, non-histone chromosomal and nuclear sap proteins. The proteins are extracted from this after a 1 M NaCl treatment at 0°C followed by spinning at high speed in a Beckmann centrifuge at 4°C.

The supernatant from the above is mixed with NaCl (to 1 M) and PEG. The latter separates out the DNA from the proteins in the solution. The DNA is pelleted out by spinning. The supernatant is stored at 4°C in glycerol and PMSF (phenyl-methyl-sulphonyl-fluoride).

The proteins may then be eluted by passing the above supernatant through an appropriate DNA cellulose affinity column. The proteins are then fractionated by polyacrylamide gel electrophoresis (PAGE) into bands that may be visualized after staining with Coomassie Brilliant Blue. The proteins from each band may then be eluted and characterized by standard biochemical procedures.

If, on the other hand, it is required to isolate a particular protein (antigen) from the surface of a cell membrane, the cell is disrupted mechanically (by sonication) and the heavier cell debris and unbroken cells pelleted out. The lighter membrane fractions are in the supernatant. The latter is spun at a much higher speed (to about 225,000 g) to pellet out the membrane fraction.

The pellet is resuspended in a buffer and fractionated by SDS-PAGE, which denatures the proteins. The protein bands, indentified by staining in Coomassie Blue, are eluted and characterized. If the molecular weight of the wanted protein is known, the band corresponding to or close to a molecular weight marker band of about the same weight is identified and eluted.

Exercise 18: Isolation of DNA-Binding Proteins from Slime Mold

Outline of procedure

I. Grow *D. discoideum* on *E. coli* on SMA plates (8–10 plates).
II. Isolate intact nuclei.
III. Isolate nucleoproteins from nuclei.
IV. Isolate DNA bound non-histone proteins.

I. To grow *D. discoideum*

Have ready

1. 8–10 plates of SMA
2. *E. coli* culture (after 4 hours of shaking)
3. *D. discoideum* stock

Procedure

1. Mix *E. coli* and slime mold suspensions and spread out on the SMA plates (see Exercise 16).
2. Incubate for 48 hours at 22°C.

II. To isolate intact nuclei

Have ready

1. Lysis buffer
2. Lysis buffer with 5% NP40 (detergent)
3. Centrifuge tubes (12 ml or so) for refrigerated centrifuge
4. Sucrose 1 M
5. HMS lysis buffer
6. Nuclear storage buffer (kept at 4°C)
7. Parafilm

Procedure

1. Remove slime mold matter (stalks, slugs and spores) from 8–10 plates and place it in approximately 5 ml of lysis buffer + 1 ml NP40 (5%).
2. Distribute the above matter into two 12 ml centrifuge tubes.
3. Spin at 2,000 g for 5 minutes at 4°C.
   The pellet consists of nuclei and some intact cells.
   Discard the supernatant.
4. Resuspend the pellets in lysis buffer without NP40.
5. Spin at 700 g (4°C) for 2 minutes.

The pellet consists of whole cells. The suspension contains intact cells.

6. Place 10 ml of 1 M sucrose (in HMS lysis buffer) in a centrifuge tube (12 ml). Layer over this the suspension from the above step.
7. Spin at 5,000 g for 10 minutes. The pellet consists of nuclei.

8. Resuspend the nuclei in 5 ml HMS lysis buffer.
9. Layer the suspension over 5 ml of HMS buffer (with 1 M sucrose) in a 12 ml tube.
10. Spin at 5,000 g for 10 minutes. Discard the supernatant.
11. Add nuclei storage buffer (4°C) to the pellet in the tube with parafilm.
12. Store nuclei at 4°C (refrigerator)

or

Continue with Step III

III. To isolate nuceleo proteins from slime mold nuclei

Have ready

1. Nculear pellet from Step II
2. Nucleoprotein extraction buffer (NPEB)
3. NaCl
4. Ice-bath
5. Sterile dist. water (SDW)
6. Ultra centrifuge tubes (7ml)
7. Beckmann's ultra centrifuge
8. PEG 8000
9. Glycerol
10. PMSF (phenyl-methyl-sulphonyl-fluoride)
11. Polyacrylamide gel electrophoresis set-up (PAGE)
12. DNA cellulose affinity column

Procedure

1. Resuspend the nuclear pellet in approximately 4–5 ml of NPEB.
2. Spin at 1,000 g (4°C).
3. Resuspend the pellet in NPEB (4–5 ml) and spin.
4. Repeat Steps 2 and 3 once more (i.e. wash twice)
5. Transfer the NPEB suspension to a 7 ml ultra centrifuge tube.
6. Add enough NaCl to make a 1 M concentration.
7. Keep the tube in an ice-bath for 30 minutes.
8. Add sterile dist. water to fill the 7 ml tubes completely. Seal the tubes well with sealing wax.
9. Spin in the Beckmann ultra centrifuge at approximately 29,000 rpm (4°C) for 90 minutes.
10. Decant the supernatant ($\simeq$ 7 ml) into a 12 ml centrifuge tube.
11. Add enough NaCl to make a 1 M soln.
12. Add PEG 8000* to make a 30% (w/v) solution in the 7 ml fluid. Shake the tube a little to mix the PEG.
    * PEG dissociates most of the DNA from proteins.
13. Spin at 11,000 rpm (4°C) for 20 minutes. The DNA precipitates out.
14. Decant the supernatant into a glass, screw-capped vial containing the following mixture.
    glycerol $\simeq$ 500 µl *
    PMSF $\simeq$ 3–4 mg*
    * These concentrations are not critical.

15. Store at 4°C.
    This contains the total nuclear proteins (Histones, NHCP and nuclear sap proteins).

You may proceed to step IV.

IV. To isolate the DNA-building proteins

1. Run the supernatant through a DNA cellulose affinity column. The eluted solution contains the proteins.
2. Run the elute on 12.5% PAGE. The proteins form separate bands.
3. Cut out each protein-containing band and elute the protein from the gel.
4. Run each elute on a fresh PAGE.
5. Characterize the proteins.

Exercise 19: Isolation of Proteins (Antigens) From E. coli Cell Membrane

This project will be divided into five parts:

I. To make a preparation of *E. coli* cell membrane in 2 mM Tris HCl (pH 7.8). This may be stored in a screw-capped vial in the deep freeze (–20°C). It can be kept for about 15 days.
II. To set up and run the apparatus for SDS-polyacrylamide gel electrophoresis (vertical slab or disc cylinders) with 25 µl of the membrane sample solution in each well (4–5 wells).
III. To identify the protein bands in the gel by staining with Coomassie Blue.
IV. To run the remaining membrane solution in the same electrophoretic gel, using a single large well. Stain the gel and elute the proteins from the bands.
V. Purify the proteins and characterize them by standard biochemical protocols.

For Parts II-IV, see Exercises under Section 7.4

Have ready (for Part I).

1. A growing culture of *E. coli* in glycerol minimal medium
2. Tris HC1 50 mM pH 8.5
   EDTA 2 mM
   in 25 ml solution.
3. Tris-HCl 2 mM pH 7.8 (may be stored at –20°C)

Procedure (for Part I)

1. Spin the cells in the medium at 6,000 rpm (Sorvall) and 2°C for 10 minutes.
2. Resuspend the cell pellet in 25 ml of 50 mM Tris HCl pH 8.5 and 2 mM EDTA.
3. Sonicate cells for 30 seconds, four times, with 20 seconds gaps for cooling. Place the tube with cells in crushed ice.
4. Spin at 6,000 g for 20 minutes. Discard the pellet (cell debris and unbroken cells).
5. Spin the supernatent in the Beckmann centrifuge at 225,000 g (2°C) for 10 minutes.
6. Resuspend the membrane pellet in 2 mM Tris-HCl pH 7.8. Store this in a screw-capped vial in the deep freeze (–20°C). It can be kept for about 15 days.

An SDS-PAGE (Polyacrylamide Gel Electrophoresis) is run with the above sample. The sample is adequately prepared by first denaturing the proteins in a sample buffer.

This denatured sample is then loaded (25 µl per well) in the electrophoretic gel.

7.4 ❑ GEL ELECTROPHORESIS

Electrophoresis is an analytical technique used for separation of molecules according to their size and shape. The technique is particularly useful for dispersing molecular species that are chemically too similar for separation by other techniques. The dispersion is based on the variable speeds of movement of the components to be separated in a given electric field. A charged molecule, when placed in an electric field, migrates towards the appropriate electrode. The speed of movement, or electrophoretic mobility, of the molecule is a function of its size (or molecular weight) and shape. The type of material (paper, gel) through which the sample is electrophoresed depends on the nature of the material to be fractioned and the aim of the exercise.

Gel electrophoresis is utilized for quantitizing and characterizing proteins and nucleic acids. The sample travels through a porous material or gel that is non-reactive and neutral and does not move in the electric field. The molecules move via a buffer solution with which the gel is innundated. When the current is turned on the sample migrates from a cavity (well) at one edge of the gel block towards the opposite edge. When the current is stopped the dispersed sample remains immobilized in the gel at varying distances from the starting well. The distance travelled by a molecular species depends on the nature of the gel (its pore size), the strength of the electric field and on the size and shape of the sample molecules. Since the electric field and gel are constant features in an exercise, the distances travelled by different sample molecules are measures of their size and shape. The electrophoretic mobility or velocity of a molecule moving through such a gel, is inversely proportional to the molecular weight, net charge and shape of the molecule. The smaller the molecule the greater the distance travelled in the same time span. In the case of nucleic acids, the size of the molecule is the length of the polymer. The shape depends on the type and degree of compaction caused by folding of the linear molecule. An RNA molecule may be compacted due to complementary duplex formations in its different parts. A DNA molecule, single- or double-stranded, may exist as a covalently closed circular structure. If two DNA molecules, one linear and one circular, of the same size are electrophoresed simultaneously, the circular molecule will move faster than the linear one. In the case of proteins a better estimate of the molecule size is obtained when the folded molecule is denatured before electrophoresis.

The extent of dispersion of molecules of similar sizes depends on the viscosity of the gel materials, which again is proportional to the pore size of the latter. The pore size may be varied by the concentration or/and length of components in the material of the gel. Two types of materials are in vogue for use in gel electrophoresis: agarose and polyacrylamide.

Agarose is a linear polymer of the carbohydrate galactose and one of its derivatives. Polyacrylamide gels are built from the monomers acrylamide and bisacrylamide.

The polymer chains in an agarose gel are held together by hydrogen bonds; those in the polacrylamide ones are cross-linked covalently. The agarose gel is made by boiling the carbohydrate in water and cooling to room temperature. Crosslinking of the polyacrylamide chains is catalysed by ammonium persulphate. The reaction, however, has to be triggered by an amine, N, N, N′, N′-tetramethylene-ethylenediamine or TEMED.

The choice of material for the gel depends on the sizes of the molecules to be dispersed. The smaller the size of the molecules of the sample, the smaller are the pores desired in the gel. Usually polyacrylamide gels are used for separating proteins and nucleic acids that are less than 2 kb in length. For nucleic acids that range in size between about 0.15 kb to about 7 kb, the most suitable pore sizes are offered by agarose gels in the concentration range of 0.5 to 1.8 per cent. Nucleic acid molecules that are 10 kb are electrophoresed best in 0.5 to 0.3 per cent agarose gels. Polyacrylamide gels used for separation of restriction fragments usually have a concentration of about 4 to 6 per cent. For sequencing nucleic acids, where there is a graded series of fragments each differing in length from the next one by only one nucleotide, 8–12 per cent polyacrylamide gels are very effective in resolving the closely graded molecular species.

There are two basic designs of apparatus for gel electrophoresis — the horizontal and the vertical. Agarose gels are used in the horizontal position (Fig. 7.9). Vertical gels usually use polyacrylamide. Horizontal agarose gels, though quite delicate or

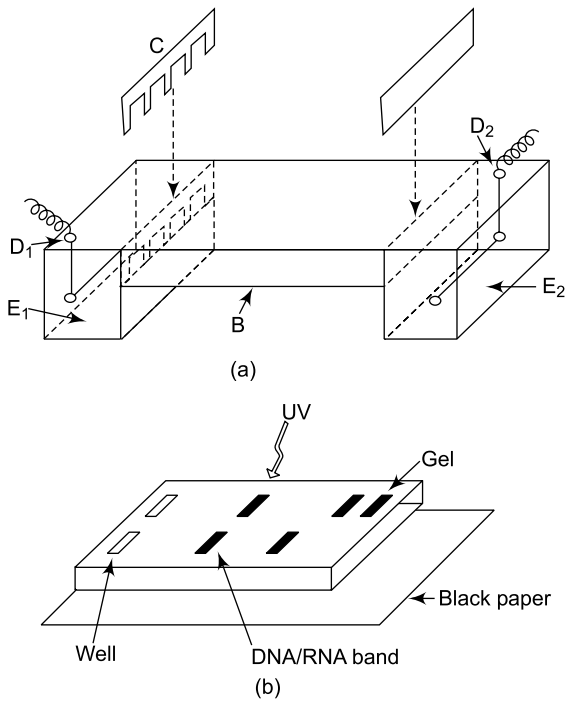


Fig. 7.9 *Prototype of horizontal agarose gel electrophoresis apparatus.*

fragile, provide a less distorted view of the separated bands and a better estimate of the relative positions of the bands than those in vertical gels. Horizontal gel electrophoresis has several other advantages over the vertical type. Both are utilized depending upon the aim of the job. The concentrations of the gels, as mentioned earlier, can be varied and the range most suitable for the expected molecular sizes in the sample is utilized.

Vertical gels can be in the form of a single rectangular slab or that of several thin cylinders in the same apparatus. The choice of the vertical slab or DISC (discontinuous) gel type of apparatus depends on the nature of the answer required in an exercise. Polyacrylamide is usually the material in the vertical gel apparatus. As this polymer oxidizes in air, the gel is protected by keeping it sandwiched between two rectangles of the plastic material of the (Fig. 7.10) apparatus or in the narrow tubes of the DISC gel apparatus (Fig. 7.11).

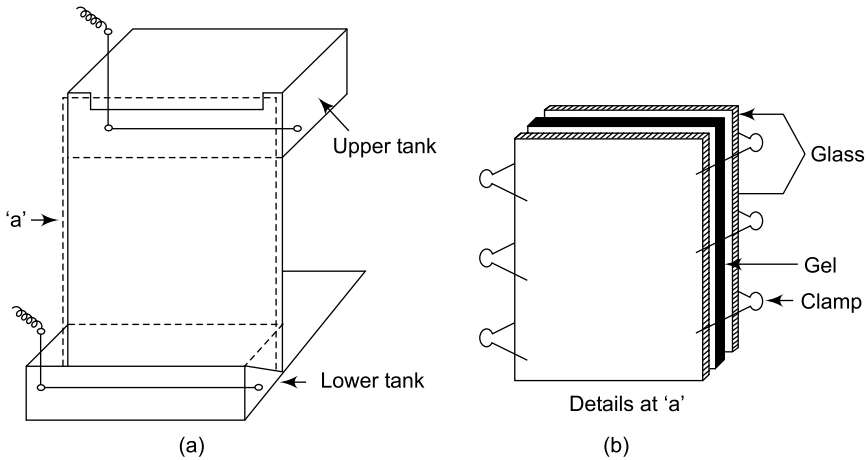


Fig. 7.10 *Prototype of vertical polyacrylamide gel electrophoresis apparatus.*

The electrophoresed nucleic acids or proteins are visualized by staining the sample appropriately. The acridine dye ethidium bromide (EtBr) intercalates between the bases of both RNA and DNA. EtBr fluoresces in UV light. EtBr is, therefore, added to the sample and/or gel or buffer. The migrating nucleic acid bands can be seen in a dark room by flashing long-range UV light over the gel. The bands glow with a reddish-orange colour. The intensity of fluorescence is directly proportional to the length of the DNA chains. This property is useful for identifying those DNA fragments that migrate through the same distances in the gel as well as restriction fragments of different lengths based on comparison with molecular standards.

EtBr is a highly carcinogenic substance and must be handled with gloved hands.

Nucleic acids stained by EtBr are vulnerable to breaks in the sugar-phosphate backbones when exposed to visible light. Hence operations of gel electrophoresis must be carried out in a darkened room. When electrophoresis is performed to purify or isolate a particular sized nucleic acid for recombinant DNA operations, the dye has to be removed from the DNA; procedures are available for doing this.

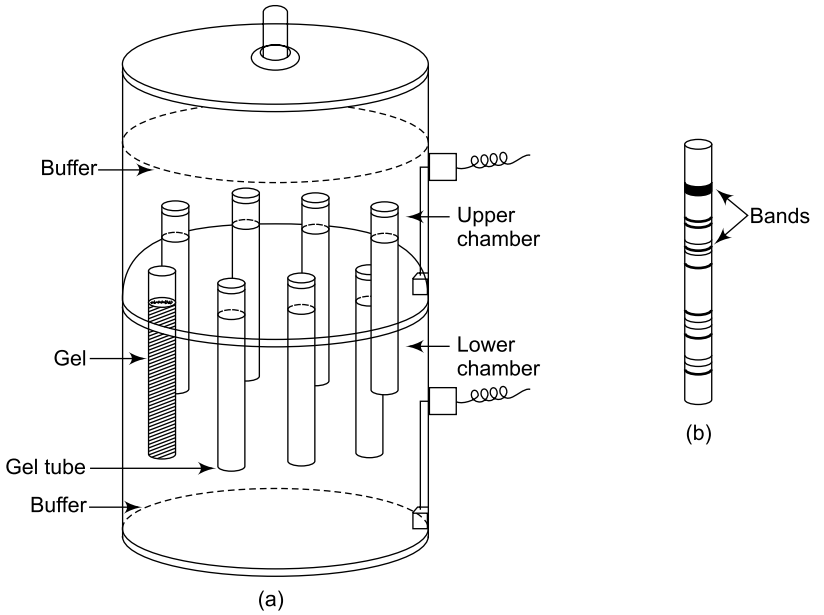


Fig. 7.11 *Prototype of DISC (discontinuous) polyacrylamide gel electrophoresis apparatus.*

Protein bands in the gel are stained after the completion of electrophoresis with the dye Coomassie Brilliant blue R (Sigma). The gel is carefully removed from the vertical sandwich or from each DISC-gel tube and stained in a flat covered dish or a large petri plate. The excess stain from the non-band regions of the gel is then removed by destaining in a mixture of acetic acid and methanol. The protein regions stand out as bright blue bands.

The samples of protein or nucleic acids electrophoresed are run simultaneously with a set of molecular weight standards. The DNA standard is often the DNA of bacteriophage lambda, whole or cleaved by known restriction enzymes. That for proteins depends on the size range of the sample proteins. A set of protein standards is available commercially. Or one may use proteins of known molecular weight such as albumin.

Several different buffers are utilized based on the requirement of a particular exercise. The pH of the buffer influences the extent and speed of separation of the electrophoresing molecules.

Gel electrophoresis may be utilized for various types of separations. SDS-electrophoresis of proteins is performed after denaturing the protein with the detergent SDS (sodium dodecyl sulphate) before electrophoresis. One may place this denatured protein directly in the wells of the gel to be used, or in a gel with larger pores (the stacking gel) placed above the separation gel. The latter procedure allows the SDS denatured proteins to percolate down to a narrow band before entering the separating gel. This ensures that the molecules leave the well more or less simultaneously; it results in clearly separated bands. Isoelectric-focusing gels distinguish proteins on the basis of their isoelectric points. A very useful innovation

in gel electrophoresis is the one known as *two-dimensional gel electrophoresis* (Fig. 7.12). In the first part of this procedure, the proteins are dispersed on the basis of their isoelectric points. In the second part, the gel is turned 90° and the dispersed proteins are now separated according to their molecular weights. Urea is often used for gels which partially denature proteins, a result desired in certain studies.

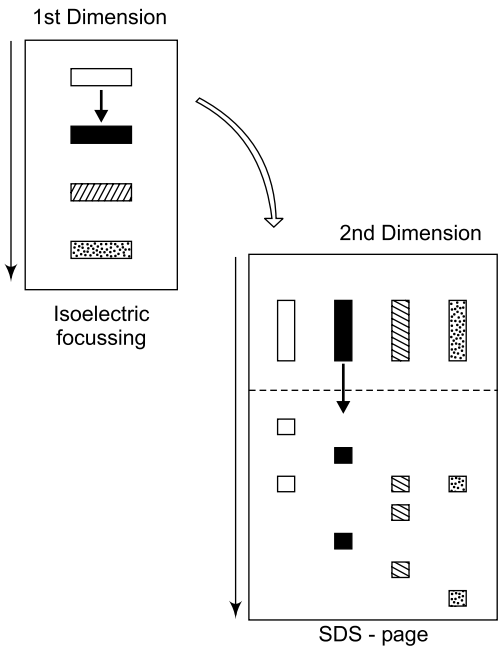


Fig. 7.12 *Two-dimensional gel electrophoresis of proteins. The technique of isoelectric focusing is employed to fractionate molecules of same weight (but not necessarily identical) in one direction. The proteins in this gel are layered on top of an SDS-polyacrylamide gel, and the proteins in each band are dispersed according to size by one-dimensional electrophoresis. Thus, quantitative (size) and qualitative (charge) differences in proteins expressed in different tissues or/and different developmental phases may be determined with greater accuracy by this technique than by simple fractionation by size.*

For electrophoresis of DNA, the sample in the well contains, besides the DNA, a dye bromophenol blue (with or without another dye xylene cyanol) and either glycerol or a concentrated solution of sucrose. The dye(s) is/are used for tracking the progress of the nucleic acid migration front in the gel. The bromophenol blue moves just a little faster than the very small DNA molecules (30 bp DNA). Xylene cyanol co-migrates with the DNA that is about 4 kb in length. Electrophoresis is stopped when the blue band of the tracking dye is about 1 cm or more from the edge of the gel opposite to the edge adjacent to the wells. The glycerol or sucrose increases the viscosity of the sample and keeps the DNA well-below the upper surface of the well. This ensures the movement of the DNA through the body of the gel. The necessity for this is obvious.

The interpretation of the gel data can be distorted due to certain standard artifacts. In most protocols the optimum size of the gel apparatus and the current per cm length of the gel are specified. These specifications ensure that under the given conditions the gel does not become over heated. This would happen if there was no proper balance between the capacity for dissipating heat and the heat generated during electrophoresis. Again, in a slab gel, the longitudinal edges can lose more heat per unit time than the regions in the interior of the gel. Since the electrophoretic mobility is directly proportional to the rise in temperature, samples at the edges and those in the interior wells of a slab gel will move at different velocities; the ones in the interior will move over longer distances than those at the sides, although molecules in both wells are of the same size.

Another wrong interpretation of the gel data is made when the ionic strength of the sample is increased. At higher concentrations of the salt in the sample, the mobility of the smaller molecules or fragments is decreased. Hence the spread of the fragment sizes will not give the correct values of the small fragment sizes.

Also, one has to be alert when running gels of plasmid DNAs. Supercoiled pDNA being the most compact form of pDNA moves the fastest. The runner-up is a linear DNA of the same size. A little slower than the linear DNA is a relaxed circular pDNA, again of the same size as the other two. Addition of EtBr induces supercoiling in relaxed circular DNA. So, an unstained supercoiled DNA and an EtBr-stained-originally relaxed DNA will both move with the same velocity in the gel.

The above pitfalls must, therefore, be kept in mind before a final judgement is passed on the bands seen in the gel. Several sophistications have been introduced to reduce the causes of the usual artifacts, and to standardize the velocity of molecules in gels made with the same apparatus.

One modification that is utilized abundantly is that of the 'minigel' system. The larger the length of the gel through which the current passes, the greater the resistance. This impedes the mobility of the molecules in the gel. It, therefore, takes a long period ($1\frac{1}{2}$ hour to half a day) for the sample to be dispersed properly in a standard large-sized gel. For a quick assessment of the DNA in a sample (presence/absence and gross size characteristics) it is customary to run an abbreviated or 'minigel'. It takes less than half an hour to see the results in such a gel (usually 2.5 × 1.5 × 0.05 inches).

7.4.1 Agarose Gels

Agarose is a polysaccharide, with galactose and a derivative of galactose, with the fibres crosslinked by H-bonds. Different lengths of DNA fibre can be fractionated by using a variety of agarose concentrations:

Use 1.4% agarose for < 0.5 kb
Use 1.0% agarose for 0.2–30 kb
Use 0.7% agarose
Use 0.3% agarose
for 10.0 kb

7.4.2 Polyacrylamide Gels and DISC Gel Electrophoresis

The polyacrylamide gel is a polymer of the monomers acrylamide and bisacrylamide.

acrylamide, $CH_2 = CH.CO.NH_2$
bisacrylamide, $(CH.CO.NH_2)_2.CH_2$

The gel is run in a vertical position either in the form of cylinders of narrow bore or a rectangular slab. The latter is more useful when comparisons have to be made between the separating bands of several samples. For separation only of the component molecular species in a sample, the thin gel cylinders are adequate.

The buffers used for running polyacrylamide gels contain chloride and glycinate ions. The chloride ions move faster than the glycinate ions. The tracking dye, bromophenol blue marks the boundary between the chloride and glycinate ions. The sharp boundary between these ions is maintained in the electric field at the pH 8.9 of the separation gel, in which the glycinate remains just ahead of the sample proteins to be separated. As the buffer thus forms a *discontinuous* system, this form of electrophoresis is also known as *DISC gel electrophoresis.*

The sample may be separated by loading directly over the separation gel. For better resolution or sharpness of the separating bands, it is more usual to allow the sample to become concentrated before it enters the separation gel. This is achieved by loading the sample over a *stacking gel* which has larger pores than the separating gel. The sample proteins percolate through the pores of the first gel and form a tight band at the entry of the second one.

Polyacrylamide gels fractionate DNA fragments shorter than 1 kb in length. They are also appropriate for sequencing of DNA. 5 per cent gels give good separation of restriction enzyme fragments.

Vertical slab gels are prepared by polymerizing an acrylamide-bisacrylamide mixture in a chamber in a gel-forming assembly consisting of two glass sheets that are kept separated by rubber or perspex spacers. The assembly is held together with wide clamps and the polymerizing mixture poured into the enclosed space. Enclosing the latter between the glass plates prevents entry of oxygen which affects polymerization.

Exercise 20: Agarose Gel Electrophoresis of DNA

Have ready

1. DNA samples
2. Molecular weight standards (usually restriction fragments of phage lambda DNA)
3. Horizontal gel electrophoresis apparatus
4. Agarose (Electrophoretic quality)
5. Electrophoretic buffer (10 X):

Tris	48.4 g
Glacial Acetic Acid	11.42 g
Na_2-EDTA	7.44 g
Dist. water to make	1000 ml
pH	8.0

(pH should be adjusted using 1 N HCl or 1 N NaOH)

N.B. The Tris-acetate buffer is used most as it tolerates high gradients of current without overheating. It should, however be replaced after about 3 hours, as it does not buffer well beyond this period.

6. Ethidium bromide*:
   Stock solution: 10 mg/ml in water
   Working solution: 1 mg/ml
   * This is a carcinogen and should be handled carefully, with rubber-gloved hands. The solution should be stored in a dark bottle covered with aluminium foil.
7. Tracking dye: Use one of the following:
   1. Bromophenol blue 0.25 g
      Xylene cyanol 0.25 g
      SDS 0.005 g
      Glass-dist. water 50.00 ml
   Mix the above. Just before use dilute the above with 50.00 ml glycerol.
   2. Bromophenol blue 50.0 g
      Glass-dist. water 50.0 ml
      Sucrose 5.0 g
8. UV transilluminator (short wavelength, 350 nm)

N.B. Wear thick, large plexiglass eye-shields or special goggles when working with UV, as the latter damages the eyes and skin tissues.

9. Parafilm
10. Gilson pipette, with tips held in position in a slab of styrofoam.

Outline of procedure

A. Prepare dil. buffer.
B. Prepare agarose solutions.
C. Set up the apparatus.
D. Load DNA samples and molecular weight marker DNAs.
E. Run the gel (i.e., switch on the current).
F. Stop the current when the tracking dye is 1–2 cm away from the farthest end of the slab (i.e., the end opposite to the edge near the wells).
G. Note the positions of the bands in the sample and marker lanes. Measure the distances of the bands from the well.

Procedure

A. Prepare dilute buffer

To 10 X Tris-acetate buffer 20 ml
Add dist. water 180 ml

B. Prepare agarose solutions

1. Prepare 1% agarose for gel slab:
   Agarose 300 mg
   Glass-dist. water 27 ml
   } Heat to dissolve.

2. Prepare 2% agarose for sealing:
   Agarose 200 mg
   Glass-dist. water 10 ml
   Heat to melt and keep warm (45–50°C) in a water-bath (beaker).

C. Set up the electrophoretic (horizontal) apparatus (use the 'comb' appropriate for the required volume of the wells to be made in the gel)

1. Wash and dry the apparatus and place it on a glass sheet on a level surface, with a sheet of thick black paper under the glass sheet. This helps the viewing of the bands during electrophoresis.
2. Place the gel-trough side supports in position.
3. Seal all edges of the gel-trough with the warm 2% agarose solution, in a warm pipette.
4. Allow the seal to solidify.
5. Add:
   27 ml of the 1% agarose soln.
   and 3 ml of buffer (10 X)
   to make 30 ml of 1% agarose soln.
6. Add to the 1% agarose soln., 1–2 drops of ethidium bromide (working solution).
7. Pour the warm 1% agarose solution into the gel-trough.
8. Fix the 'comb' in the gel about 1 cm from one short end and 0.5 mm–1 mm from the bottom surface of the gel. (The size of the comb will automatically ensure the distance from the bottom.)
9. Allow the gel to solidify. This takes about 30 minutes.
10. Prepare the working dil. buffer solution.
11. Add to the above solution 2–3 drops of ethidium bromide.
    (It should give a slight pinkish tinge to the buffer.)
12. Pour the dil. buffer into both the troughs and on the solidified gel. Remove the side supports, so that the buffers in the troughs and over the gel are able to mingle.
13. Remove the comb carefully without piercing the bottom of the gel or otherwise damaging the wells.
14. Add more buffer until there is about 1 mm buffer above the gel surface.

The apparatus is ready for use.

D. Load the sample and marker DNAs in wells

1. Place on a 2″ × 2″ square of parafilm sheet as many drops of tracking dye as there are DNA samples (including the marker), using a Gilson pipette and a disposable tip.
2. With a fresh 'tip' and a 10 lambda holder, take 10 lambda of a DNA sample and mix it with one drop of EtBr working solution.
3.* Take up the sample + dye mixture with a fresh tip of the Gilson pipette and gently release the fluid into one of the wells. Make a note of the samples and the wells into which they have been placed.
4. Repeat Steps 2 and 3 above for each sample and the marker DNA. Place the latter in the first or the last well.

- Do not load more than 0.2–0.5/g DNA per 5 mm well.
- Be careful not to allow the sample to fill more than 2/3rd of the well. This would force passage of the DNA through the gel and prevent it from flowing over the gel. The sucrose increases the density of the test solutions and thus helps to keep the DNA lower in the well.

E. Run the gel

1. Connect wires from the electrodes in the two buffer troughs to the power pack, with the 'sample end' of the gel connected to the –ve terminal and the distal end to the +ve terminal.
2. Current to be passed at 12 V/cm of gel (30 ma).
3. Run the gel for at least $2\frac{1}{2}$ hours or until the tracking dye band is about 1 cm from the distal edge of the gel.

F. Switch off the current.

G. Observe the bands

1. Shine the UV light on the gel in the dark.
2. The DNA bands will fluoresce with a reddish-orange colour.
3. Note the positions of the bands relative to those in the marker lanes. A rough estimate of the size of the DNA in each band is made in this way. Measure the distance travelled by each band from the wells.
4. Photograph the gel using a UV light source and a red filter.*
   *see the Appendix.

N.B. You may preserve the gel for long periods submerged in dil. buffer solution, in a covered large petri dish. Seal the edges of the cover with parafilm.

Exercise 21: Polyacrylamide DISC Gel Electrophoresis of Proteins

Have ready

1. DISC gel apparatus
2. 12 tubes for cylindrical gels
   (a) Blood (b) Other protein mixtures
3. Gel buffer solution:

NHC1	48 ml	
Tris	36.6 g	
*TEMED	0.2 ml	*catalyst for polymerization
Dist. water to make	100.0 ml	pH 8.9

4. Stock solutions for making gel:

*1. Acrylamide 60 g in 100 ml H_2O. Make 100 ml

*2. Bisacrylamide 2.6 g in 100 ml H_2O. Make 150 ml.

3. Ammonium persulphate 0.2 gm in 100 ml. Make 150 ml.
4. Bromophenol blue 0.05%. Make 10 ml.

The (1) and (2) solutions can be kept at 4°C for about 6 months. The (4) solution may be kept at room temperature.

*These are neurotoxins; so handle with gloved hands and *never* pipette by mouth.

5. Electrode buffer: Make 2 litres
   Add to a volumetric flask (1 litre)

Dist. water	500 ml
Glycine	2.8 g
Tris	0.6 g
Add dist. water to make	1000 ml pH 8.3

6. Tracking dye–10 ml, Bromophenol blue 0.1 g/*l*
7. 7% acetic acid
8. Stain: Amido black 1 g/*l* in 7% acetic acid

Procedure

I. Make the gel cylinders.
II. Prepare the sample.
III. Assemble the apparatus.
IV. Run the gel.
V. Remove the gel cylinders.
VI. Stain gel and observe the bands.
VII. Observe the positions and the widths of the bands.

I. To make gels

1. Mix in a 100 ml beaker:

Acrylamide stock soln.	4 ml
Bisacrylamide stock	2 ml
Dist. water	4 ml
	10 ml

2. To another 100 ml beaker:

Transfer	6 ml	of the above solution
Add	3 ml	gel buffer (pH 8.9) and
	15 ml	ammonium persulphate stock
Total	24 ml	of gel solution

   Mix well.
3. Using a Pasteur pipette fill the tubes in the following manner (Fig. 7.13):
   Fix each tube in a rubber bung. Place a little gel solution in the tube. Tap the assembly on the bench-top to remove the air bubbles. Now fill the tube with the gel solution leaving 1 cm from the top edge.
4. Place a drop of dist. water on top of the gel solution and leave the assembly undisturbed for about 30 minutes.

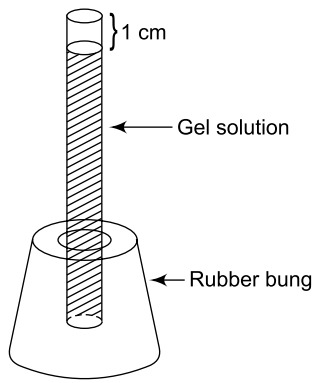


Fig. 7.13 *Set-up for preparing the gel in tubes of the DISC electrophoresis apparatus. The gel solution is poured to within 1 cm from the top of each tube, which is held upright (by the rubber bung or some other holder) until the gel has set.*

II. To prepare the sample Drain 2 drops of blood, from an alcohol-sterilized finger tip punctured with a sterile needle, to a small tube and add 1 ml distilled water to it. The osmotic shock due to the pure water bursts the red blood cells, resulting in hemolysis. This hemolysate is used for fractionation by DISC gel electrophoresis.

III. To assemble the apparatus

1. Shake out the water from the top of the solidified gel (separation gel).
2. Fit each gel-tube in a hole in the upper electrode chamber.
3. Pour Electrode buffer (800 ml) into the lower chamber to a depth of about 3 cm.
4. Fix the upper chamber over the lower one.
5. Fill the upper chamber with 500 ml of Electrode buffer (i.e., to a level just below the top edge of the tubes).
6. Add the sample solution, with a Pasteur pipette, to the top of the separation gel in each tube.
7. Leave undisturbed for 5 minutes.
8. Add another 500 ml (approx) Electrode buffer to which 2–3 drops of Bromophenol dye have been added. The buffer should submerge at least 1 cm of the carbon electrode.

IV. To run the gel.

1. Fit the electricity supply unit over the upper electrode chamber.
2. Connect leads to a power pack with the upper chamber connected to the negative electrode (cathode) and the lower to the positive one (anode).
3. Switch on the current*. Note the time.
   *The apparatus used by the author supplies 2–3 mA per gel tube, automatically.
4. Switch off the current when the blue band, due to the dye, is almost at the bottom of the tube.

V. To remove the gel cylinders

1. Take out the tubes and place them horizontally in a dish, filled with water.
2. The gel is removed from the tube by first injecting a stream of water between the gel and the inner wall of the tube, using a syringe and an 18 gauge needle. The gel then slips out easily.
3. Transfer each gel carefully into a large test-tube.
4. Fill the test-tube with stain solution. Keep the gel in stain for 5–10 minutes.
5. Decant off the staining solution.
6. Add 7% acetic acid, keep for a few minutes and replace it with fresh 7% acetic acid several times, until only the bands in the gel remain coloured blue. (Destaging procedure).

VI. Observations

1. Mark the positions and widths of each band on a sketch.
2. Photograph the gels, if required.
3. Estimate the range of molecular weight of each band in the protein mixture samples by comparing with the molecular standards of the proteins* that separate from the haemolysed blood*.
   *Haemolysed blood contains the following major protein species:
   Haemoglobin
   Albumin
   Fibrinogen
   Globulins: These are the immunoglobulins that include 5 types (M, G, A, D, E), one or more of which are predominant in each blood sample.

7.5 ❑ PURIFICATION OF DNA

The isolated DNA, stored in solution or in a desiccated form, is usually not pure enough for use in various operations, especially those involving restriction enzymes and ligase.

The DNA may be purified in a number of ways. The preliminary separation of DNA of a particular molecular length is achieved by *gel electrophoresis.* This DNA can be removed from the gel by *electroelution.* The two most common methods of electroelution are (1) Dialysis in an electric field, and (2) Elution of the DNA into a well in the gel itself. The eluted DNA may be purified further by passing the sample through a column packed with a material that preferentially adsorbs the impurities in the DNA solution. The elutant is DNA that may be purified further by *density gradient centrifugation.* The purest DNA is obtained by the last method, which uses either a sucrose or a caesium chloride (CsCl) density gradient for spinning out the sample DNA. The DNA in a CsCl gradient is identified by the fluorescence in UV light of ethidium bromide that is used to stain the DNA. The EtBr is removed subsequently. In some protocols this last step is avoided by a method that does not employ EtBr.

In summary, the DNA may be purified by the following procedures. For very pure DNA, the procedures (1) and (3) should be followed in succession.

1. Electroelution from the gel
2. Biogel column purification
3. Density gradient centrifugation

7.5.1 Electroelution from an Agarose Gel

The DNA bands in the gel are dialysed into a buffer in an electric field by this procedure. The DNA moves out of the gel into the buffer in a dialysis bag. The DNA is then precipitated out from the buffer solution in the bag.

7.5.2 Biogel Column Purification

A DEAE-cellulose (Sephadex) column will be used in this exercise for cleaning the dialysed DNA. Sephadex is a product of Pharmacia and consists of cross-linked dextrans. The porosity of the gel depends on the degree of cross-linking. A wide range of Sephadex products are available which are appropriate for fractionating molecules within small size ranges. These are named as Sephadex G-25, G-50 and so on.

Sephadex remains stable for long periods (unless contaminated with micro-organisms) since it is not affected by weak acids and bases. The column can be reused if washed thoroughly after use and stored at 4°C with a preservative such as chloroform (added to the last wash water).

7.5.3 Density Gradient Centrifugation of DNA

Very pure DNA may be obtained by centrifuging it in a density gradient made in a centrifuge tube with a suitable material. Caesium chloride (CsCl) solutions can be

spun at very high speeds in a fixed angle rotor to obtain a layering of the molecules according to their densities, with the heaviest ones at the bottom of the tube and the lightest at the top. The density of each layer of CsCl can be determined with reference to a standard. When the DNA is spun in a CsCl density-wise layered tube, the nucleic acid settles in the layer having the same density as its own.

This method has been employed for determination of the densities of prokaryotic and eukaryotic DNA, and has led to the facts that (i) all of the DNA in a prokaryote possesses one particular density, which is characteristic of the species; (ii) eukaryotic cell DNA is, on the other hand, heterogeneous. All eukaryotic cell DNA separates out into two bands—the bulkier one representing the main chromosomal DNA, and the smaller band(s) representing the mitochondrial and or chloroplast DNA depending on the inclusions in the cell. The main chromosomal DNA may be further fractionated (if spun after fragmentation) density-wise, into one or more additional bands. These latter DNA are called satellite DNA and represent the repeated sequence regions in the genome.

To obtain pure DNA from the samples purified and cleaned by one or more of the preliminary procedures (electrodialysis, column purification) it is mixed with the CsCl solution and spun at a very high speed for long periods (40–60 hours). The DNA is stained with ethidium bromide in most protocols. The DNA band, visualized in UV light is transferred from the polyallomer centrifuge tube and the ethidium bromide is removed from it.

For purifying the plasmid DNA, a sucrose density gradient is often utilized, spinning at about 33 krpm for about $2\frac{1}{2} - 3\frac{1}{2}$ hours at 10°C.

If the concentration of the DNA is very low in the initial solution, the DNA is concentrated before centrifuging it in the density gradient.

The DNA may also be purified by spinning at high speeds in a sucrose density gradient. In this case a sucrose density gradient is created with the help of a gradient maker. This consists of two tubes connected at the bottom by a horizontal tube that can be opened or closed with a valve. Sucrose solutions of required concentrations are placed in the two tubes and the contents of both allowed to mix gradually and then eluted from the bottom of the gradient maker, via a thin bored tube into 5 mm centrifuge tubes that fit the rotor of a Beckmann centrifuge. The filled tubes are kept overnight at 4°C. The DNA sample is next added to this and spun at a high speed. The DNA settles down in that sucrose layer which has the same density as the DNA.

Exercise 22: Electroelution of DNA from Agarose Gels (I)

Have ready

1. Agarose gel with DNA bands stained with ethidium bromide
2. Gel Electrophoresis Apparatus (Horizontal)
3. Electrophoresis buffer:

   Stock 10 X:

Tris	48.40 g	
Glacial acetic acid	11.42 g	
Na_2-EDTA	7.44 g	
Dist. water to make	1000 ml	pH 8.0

   (Adjust pH using 1 N HCL or 1 N NaOH)

Working solution:

10 X Stock	20 ml
Dist. water	180 ml
Total	200 ml

4. Dialysis tubes: Size 20, 13 cm lengths that have been cleaned according to the following schedule. Boil in the following solutions in a large beaker, taking care that the tubes remain submerged in the solution at all times.

	Double dist. water	0.2 M EDTA	Time
(i)	750 ml	5 ml	10 min
(ii)	Repeat (i)		
(iii)	500 ml	2.5 ml	15 min
(iv)	500 ml	1.0 ml + 10 g Na_2CO_3	20 min
(v)	500 ml	1.0 ml	15 min

Store at 4°C in double dist. water in a wide-mouthed, screw-capped bottle.

5. Parafilm pieces 5 cm × 5 cm
6. Petri dishes (one for each parafilm piece)
7. Equilibriated phenol
8. Chloroform:Isoamyl alcohol (24:1)
9. Sodium acetate 3 M solution
10. Ethanol (chilled)
11. Eppendorf tubes
12. Eppendorf centrifuge microfuge.

Procedure

1. Carefully transfer the 'run' gel on the glass window of the UV transilluminator or UV-box.
2. Switch on the UV lamp and cut out the DNA bands to be eluted, using a sharp pointed blade.
   Pare the gel as close to the band as possible without damaging the latter.
3. Place a square of parafilm on a petri dish and transfer the cut out gel-DNA piece onto this parafilm. Cover the dish until use.

4. Preparation of the dialysis bag for electrodution of DNA

1. Twist a tubing length to loosen the sides that stick to each other.
2. Make a knot at one end of the tube.
3. Introduce a little of the working buffer (~1 ml) into the knotted tubing. Squeeze several times to wash the inside of the tube with the buffer.
4. Carefully transfer 2 gel-DNA blocks (replicas of the same band) into the tubing. Add a little more of the working buffer (total of about 2 ml) into the tubing so that the gel-pieces are totally submerged in it.
5. Knot the open end of the tubing. The gel-DNA blocks should be immersed in the buffer and there should be no air bubbles.
6. Place the dialysis bag in the gel-trough of the electrophoresis apparatus, with the DNA bands in the same orientation as in the intact gel slab.
7. Pour enough working buffer into the side trough to allow the dialysis bag to be completely immersed in it.

8. Connect wires to the power pack and apply a current of 30 mA for $1\frac{1}{2}$–2 hours.
9. Stop the current, reverse the electrode connections and apply current in the opposite direction for 2–5 minutes.
10. Switch off the current.

5. Precipitation of DNA from the dialysed solution

11. Remove each dialysis bag and hold upright in one hand.
    Cut off the bag at one knot.
    Pour out the contents of the bag (dialysed DNA in buffer) into a 15 ml graduated centrifuge tube. Let volume of this solution be *x* ml.
12. Add 1/*x* ml of equilibriated phenol and 1/*x* ml of choloroform:isoamyl alcohol (24:1)
    Mix well.
13. Spin in the clinical centrifuge (4,000 rpm) for 10 minutes.
14. Transfer the upper (aqueous) phase into a fresh centrifuge tube.
    Add equal volumes of chloroform:isoamyl alcohol.
    Mix well.
    Spin as above for 10 minutes.
15. Transfer the upper (aqueons) phase to a glass centrifuge tube.
    Add Na-acetate (to make 0.15 M solution).
    Add twice the volume of chilled ethanol.
16. Leave overnight in the deep freezer (–20°C).
17. Spin at 10,000 rpm in the Sorvall (SS34) rotor for 20 minutes.
18. Decant off the alochol and allow the inside of the tube to dry at room temperature. (Keep it inverted over the blotting paper.)
19. Add 0.2 ml (or 200 µl) of distilled water (sterile) and suspend the DNA pellet.
20. Transfer the resuspended DNA into an Eppendorf tube. Rinse out the glass tube with another 200 µl of water and add this to the Eppendorf tube (total = 0.4 ml)
21. To the above, add
    0.02 ml 3 M Na-acetate
    1.00 ml chilled ethanol
    (total soln. = 1.42 ml)
22. Keep the above soln. at –20°C overnight.
23. Spin in the Eppendorf for 10 minutes. (12,000 g)
24. Decant the liquid.
25. Add to the pellet 1 ml chilled ethanol. Spin for 5 minutes.
26. Dry the pellet under vacuum in a desiccator.
27. Add sterile dist. water and resuspend the DNA for use.
28. Check the purity of the DNA with the UV-spectrophotometer.

Exercise 23: Elution of DNA from a Polyacrylamide Gel

Have ready

1. Polyacrylamide gel with DNA bands
2. UV lamp (360 nm)

3. Ammonium acetate 0.5 M:
   (i) Dissolve in 800 ml dist. water, 385 gm ammonium acetate.
   (ii) Make up volume to 1000 ml with dist. water.
4. EDTA 1 mM, pH 8.0:
   (i) Dissolve 372.2 gm EDTA in 800 ml H_2O.
   (ii) Make up volume to 1000 ml.
   (iii) Adjust pH to 8.0 with about 20 g of NaOH pellets.
5. Elution buffer:
   (i) Adjust pH to 8.0 with 20 g NaOH pellets.
   (ii) Dilute stock (5 M) to 0.5 M (10 ml).
   (iii) Distribute into 100 ml aliquots and add 0.5 M ammonium acetate and EDTA 1 mM, pH 8.
   (iv) Filter sterilize 1.0 mM, pH 8.
6. Parafilm
7. Pasteur pipette with narrow tip
8. Centrifuge tubes
9. Vortex
10. Ethyl alcohol 95% and 70%
11. TE buffer:
    10 mM Tris-HCl, pH 7.9
    1 mM EDTA, pH 8.0
12. 3 M sodium acetate pH 5.2 (Dissolve 408.1 g of Na-acetate, $3H_2O$ in 800 ml dist. water. Adjust pH to 5.2. Distribute into 100 ml aliquots; autoclave.)

Procedure

1. Locate the DNA bands in the gel by flashing UV light onto them.
2. Remove the band to be eluted using a sharp blade.
3. Place the piece of gel on a glass plate and trim away, as much as possible, the gel surrounding the DNA band.
4. Now, cut the gel into very fine pieces with the help of a blade.
5. Push the pieces into a test-tube with the help of a blade or a scalpel and cover them with the elution buffer (same volume as the gel-pieces).
6. Cover the tube with aluminium foil or parafilm and incubate at 37°C overnight, with very slow shaking.
7. Spin the contents of the tube at 10,000 g (20°C) for 10 minutes. Transfer the supernatant to a test-tube.
8. Add some more (half of the earlier volume) elution buffer to the pelleted gel-pieces. Vortex and spin.
9. Add the supernatant to the supernatant obtained after Step 7.
10. Add 95% ethyl alcohol. The DNA precipitates out.
11. Pour off the ethyl alcohol.
12. Add 200 µl of TE buffer to the pellet and allow the DNA to dissolve in it.
13. Add 25 µl of sodium acetate and allow the DNA to precipitate out.
14. Discard the supernatant, resuspend the pellet in 70% ethyl alcohol and spin briefly to pellet out the DNA.

15. Dry the pellet in a vacuum desiccator.
16. Resuspend the DNA in TE buffer.

Exercise 24: Column Purification of DNA

The DNA isolated from the cell debris is still not pure enough for diverse operations. Impurities in DNA prevent most restriction enzymes from cutting the DNA strands. The DNA may even be degraded by some contaminants, which may be microorganisms or DNA-degrading nucleases. These require certain divalent cations for the enzymatic reactions. Adding a chelating agent, such as EDTA (1 mM in the solution), removes these essential cations and prevents enzyme action.

One of the best ways of purifying the DNA eluted from gels is to pass it in solution through a column packed with a biogel, such as DEAE-cellulose. Only the DNA is adsorbed on the column material, while the unbound impurities pass through it. The smaller lengths of DNA ($\simeq$ 1 kb long) that bind to the biogel can be eluted easily by a suitable buffer. The longer DNA fragments that are more difficult to extricate from the gel particles, require a modified procedure for elution.

Have ready

1. DNA solution to be cleaned
2. DEAE cellulose (Whatman, DE 52)
3. A small Pasteur pipette (to be used as the column) 7 cm column length.
4. Glass wool (siliconized, if possible)
5. Column buffer:
   Tris-HCl 10 mM pH 7.5
   NaCl 0.3 mM
6. Elution buffer:
   NaCl 1 M 2.5 ml
   Tris-HCl 10 mM pH 7.4
7. 95% ethyl alcohol
8. Clamp-stand to hold the column in a vertical position
9. Beakers for eluted fluids
10. Eppendorf tube to collect the eluted DNA

Procedure

1. Set up the column (Pasteur pipette) using the clamp-stand.
2. Pack a little glass wool at the lower end of the column.
3. Mix the DEAE-cellulose (about 1/2 teaspoon) in the column buffer in a beaker and allow the gel to swell (takes about 5–8 minutes).
4. Using a pipette, transfer about 1 ml of the swollen gel slurry into the column. Allow it to settle for 2–5 minutes.
5. Add about 5 ml of column buffer to the column, and collect the eluted liquid in a beaker. Discard this fluid.
6. Repeat Step 5 about 4–5 times.
7. Now, transfer the DNA solution to the top of the slurry in the column.
8. Collect the eluted fluid and discard it.

9. Add about 2 ml of column buffer and discard the elutant. Repeat 5 times.
10. Add to the column 2.5 ml of elution buffer *Collect the eluted solution, which will now contain the DNA.
11. Add water to the DNA solution so that the strength of NaCl becomes 0.2 M (from the 1 M of the elution buffer).
12. Add 2 volumes of ethyl alcohol. The DNA will precipitate out.

***N.B.** If it is suspected that some of the biogel particles have been eluted with the DNA solution, it is advisable to spin the latter in an Eppendorf for a few seconds and collect the supernatant containing the DNA.

Exercise 25: Sucrose Density Gradient Centrifugation of DNA

Have ready

1. DNA samples kept in solution at 4°C
2. Sterile 5% and 20% sucrose solutions in water (w/v)
3. Gradient maker assembly
4. Gradient-tube holder assembly with a 18 gauge syringe needle having a small plastic funnel fitted at the broader end of the needle
5. Slanting holder for sucrose tubes
6. Straight holder for sucrose tubes
7. Straight holder for test-tubes
8. Thin 10 ml test tubes (12 per DNA sample)
9. Chloroform:isoamyl alcohol (24:1)
10. Chilled ethanol

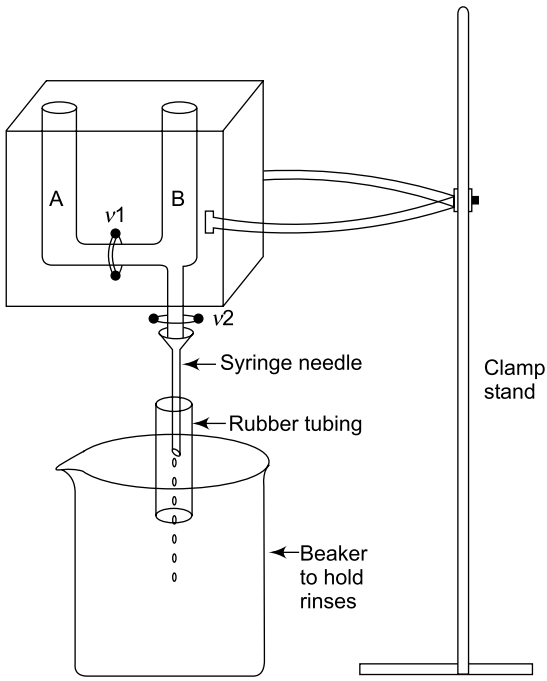


Fig. 7.14 *A sucrose gradient maker*

11. Standard SSC solution:
    NaCl 0.15 M
    Na_3 citrate 0.01 M
12. Beckmann's ultracentrifuge:
    L 5 65B, class H, with Rotor SW 65 (kept at 4°C in the refrigerator or cold room)
13. Vacuum grease
14. Eppendorf centrifuge
15. Eppendorf tubes
16. Standard 5 ml nitrocellulose centrifuge tubes for Beckmann's Rotor SW 65.

Procedure

I. Prepare sucrose gradients using the gradient maker.
II. Load the DNA sample on gradient tubes and spin.
III. Collect the purified DNA.
IV. Concentrate the DNA if needed.

I. To prepare sucrose gradient tubes

1. Prepare and autoclave 5% and 20% sucrose in water (w/v)
2. Hold the gradient maker about 12″ above the bench-top, using a clamp-stand.
3. Rinse the gradient maker with distilled water, use a Pasteur pipette to dispense water in the wells (A and B) of the apparatus (see Fig. 7.14).
4. Close the valves, numbered 1 and 2.
5. Open valve number 1, and rinse the channel between the wells. Remove the 5% solution with a pipette, from the side of well A. This rinses the well A and the channel with the 5% solution.
6. Close valve 1, open valve 2. Fill well B with 20% sucrose solution and rinse, letting the 20% solution pass out into a beaker held below the rubber tubing. Close valve 2.
7. Now fill:
   Well A with 2.6 ml of 5% solution.
   Well B with 2.4 ml of 20% solution.
8. Stir the 20% solution in well B with the stirring rod of an electric stirrer (washed first by stirring with dist. water in a test tube).
9. Place a rack (lucite) with slanting holders for 5 ml centrifuge tubes that fit Beckmann SW 65 rotor under the gradient maker, with the delivery tube (fixed to a syringe needle) of the gradient barely touching the side of the centrifuge tube.
10. Fill the centrifuge tubes with the sucrose solutions by opening both valves simultaneously.
11. Rinse out wells A and B with 5% and 20% solutions respectively and fill with the appropriate solutions. Empty this into the next tube.
12. Repeat Step 11 until all the tubes are filled.
13. Keep the tubes, in a *straight* (vertical) position in the refrigerator for about 30 minutes. This stabilizes the gradient in each tube.

II. To load and spin the DNA sample

1. Take out the sample DNA kept at 4°C in screw-cap vials.
2. Take out the Rotor SW 65 (Beckmann's) from the cold room (or refrigerator).
3. Apply vacuum grease to insides of tube holders (buckets) and screws of the rotor.
4. Remove 'sucrose' tubes from the refrigerator and load them with the DNA samples in the following manner:
   (a) Take 0.3 ml of the sample with a pipette and gently trickle it down the side of a 'sucrose' tube, without disturbing the sucrose gradient.
   (b) Place this tube carefully in a 'bucket' by lowering it with a pair of thin forceps, and touching the rim of the tube to the inside of the bucket.
   (c) Place a nut-cap on the bucket and tighten the nut with the tool provided for this purpose.
   (d) Insert the pin and screw tight.
5. Place the loaded rotor in the centrifuge (Beckmann L5 65B ultracentrifuge, class H).
6. Close the top of the centrifuge and spin at 33 krpm at 10°C for 3 hours.
7. Take out the rotor. Take out the buckets.
   Open the 'cap' of each bucket with the special tool provided.

III. To collect the purified DNA

1. Place a sucrose tube in the channel in a gradient tube holder (Fig. 7.15). Cover the top of the holder and screw it tightly. Arrange this holder with a clamp-stand so that its lower edge is at least 12 inches from the bench top.
2. Remove cotton plugs from 12 thin sterile 10 ml test tubes, kept in vertical positions in a tube-holder.
3. Puncture the bottom of the sucrose tube with a needle having a small plastic funnel attached to its broader end.
4. Hold the tube under the needle and loosen the screws of the holder. Drops will emerge through the needle. Collect 6 drops of liquid in each one of the 12 tubes, per sucrose tube.
5. Remove the screws and cap of the gradient tube-holder, clean the rubber pad. The holder is now ready for the insertion of another sucrose tube.

N.B. When using pDNA, the 4th to 7th fractions of the liquid are expected to contain the pDNA, with fraction 7th having the peak amount.

These purified pDNAs are suitable for transformation of recipient cells.

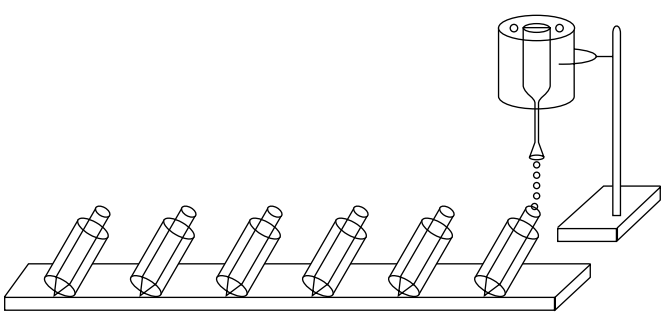


Fig. 7.15 *Delivering sucrose solution gradients in 5 ml centrifuge tubes.*

6. Run a minigel to estimate the quantity of DNA per sample. If the bands are very faint, concentrate the 6th, 7th, 8th fractions of the samples in the manner described below.

IV. To concentrate the DNA

1. To sample DNA (kept at 4°C) add equal volumes of Chloroform:isomyl alcohol (24:1). Shake by hand to mix well.
2. Spin in an Eppendorf (12 krpm) for 15 minutes.
3. Transfer the aqueous (upper) layer into a fresh Eppendorf tube and add to it 2 volumes of ice-chilled ethanol. Leave undisturbed on an ice-bath for 30 minutes to 12 hours. (A longer period is better.)
4. Spin at 12 krpm for 10 minutes.
5. Decant the supernatant.
6. Dry the pellet under vacuum in a desiccator.
7. Resuspend the pellet in a minimum volume (10–20) of SSC buffer. (Std. sodium citrate soln. = NaCl 0.15 M, Na_3 citrate 0.015 M).

Exercise 26: Caesium Chloride Density Gradient Centrifugation

Have ready

1. DNA solution for further purification
2. CsCl
3. A saturated solution of CsCl in isopropanol
4. Ethidium bromide
5. TE buffer
6. Liquid paraffin
7. Sterile disposable syringes (20 gauge needles)
8. Isopropanol
9. Dialysis tube (10 cm diameter or less)
10. Magnetic stirrer
11. Beaker
12. Polyallomer centrifuge tube
13. Beckmann's ultracentrifuge
14. UV transilluminator (long wavelength)

Procedure

1. Spin me DNA sample (in solution) at 13,000 rpm for 30 minutes in a Sorvall centrifuge at 4°C.
2. Add to above 0.98 g of CsCl and 200 mg of ethidium bromide per ml of the solution.
3. Spin in the Beckmann's centrifuge at 35–40,000 rpm for 48–60 hours at 10–15°C.

N.B. There should be no air bubbles in the tube before spinning. If there is space above the liquid in the tube fill it up with liquid paraffin and close the cap tightly; the paraffin should overflow while the cap is being screwed on.

4. Take out the tubes immediately and observe DNA bands with a UV transilluminator.

5. Fix a tube with a clamp on a vertical stand. Stick a cellophane tape on the tube over the band to be eluted. Using a syringe, puncture the tube through the tape at the lower surface of a DNA band and run off the fluorescent band material (DNA + ethidium bromide) into a small tube.
6. Extract the DNA from the material with isoproponal in the following manner:
   (i) Add equal volume of 'saturated CsCl solution in isopropanol' to each tube, and shake gently by rocking. The ethidium bromide will come out into the isopropanol (which will turn pink).
   (ii) With a Pasteur pipette remove the isopropanol. Add more isopropanol-CsCl, shake and remove the pink solution. Repeat this step until the isopropanol is no longer stained pink.
7. Remove CsCl by dialysis in the following manner:
   (i) Transfer the DNA with a Pasteur pipette into a prepared dialysis tube (washed and cut into 3–4 inch length and knotted at one end). Knot the open end of the bag.
   (ii) Place the bag in a beaker and pour TE buffer to cover the bag.
   (iii) Place the magnetic needle of the 'stirrer' in the beaker.
   (iv) Place the beaker on the 'stirrer' and keep it on overnight, (although 2–3 hours may be enough). All the CsCl should come out at this stage.
8. Store the DNA in a screw-cap bottle at 4°C.

7.6 ❑ MANIPULATION OF DNA

The DNA to be cloned has to be first manipulated in a variety of ways to make it suitable for the exercise to be carried out. These manipulations may include one or more of the following:

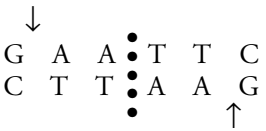
1. Cleaving the DNA into fragments with restriction endonucleases
2. Removing the $-PO_4$ moiety from both ends of the molecule
3. Digesting each DNA strand with endonuclease
4. Adding poly-tails to appropriate ends
5. Ligating DNA fragments

7.6.1 Cleavage with Restriction Enzymes

Restriction enzymes are endonucleases that cleave double-stranded DNA in specific sequences of 4 to 7 bases in length. Populations of identical regions of a DNA can, therefore, be isolated by using a particular restriction enzyme. It thus becomes possible to isolate specific subsets of a genome for utilization for a variety of goals. Perhaps the most important factor (if one is permitted to make a distinction between more than one essential factor) in the birth of genetic engineering is the discovery of restriction enzymes.

There are two main types of restriction enzymes—Types I and II. The first type, discovered initially, is a large protein with many subunits, and which requires ATP, Mg^{++} and a co-enzyme called S-adenosylmethionine for its activity. Type I enzymes bind to DNA in characteristic sequences, but actually cleave outside these regions at points that are not found to be specific. Type II enzymes are monomeric proteins

that cleave the DNA within unique sequences. These sequences exhibit a dyad symmetry. This means that both strands of the DNA have the same sequences when read from the 5′ to the 3′ end. The enzyme delinks the sugar-phosphates in the DNA backbones at the sites of cleavage.



The dotted line is the central axis of this *recognition sequence* for a restriction enzyme called *Eco* RI. The enzymes are named after the organisms from which they have been isolated. The first alphabet is the initial letter of the genus name (*Escherichia*), while the next two are the first two letters in the species name (*coli*). Another alphabet may signify a particular strain of the species, while a Roman number (I, II ...) gives the order of discovery of the enzyme.

Type II endonucleases cut the DNA in one of three ways within the recognition sequence. In the first case (a) each fragment produced by cleavage of the DNA possesses a single-stranded extension ending with a phosphate (5′ end) group. In the second case (b), the single-stranded extensions end with an OH (3′ end) group. In the third type (c), the two strands are cleaved at the central axis of the recognition sequence. The fragments in this case are 'blunt-ended'.

Each restriction fragment has its own characteristic cleavage point, yielding tailed- or blunt-ended fragments. The above facts are exploited for the manipulation of DNA ends. For instance, a DNA polymerizing enzyme (DNA Pol) can lengthen the polynucleotide chain only at the 3′ end of a primer that is bonded complementarily to a template DNA. Again, if it is necessary to label a DNA strand with a radioactive phosphate moiety, this is done readily if the PO_4 of the strand to be tagged is at the end of the single-stranded tail. In view of these features one uses the particular restriction enzyme that will provide the type of fragment that is desired.

Sometimes it becomes necessary to provide a cohesive end to a blunt-ended fragment. This is achieved by adding short chains (oligonucleotide lengths) of DNA to one strand at a blunt end of the ds DNA. These additions are popularly known as 'linkers'. They are the recognition sequences of specific enzymes. Linkers are synthetic and are available commercially. If a linker attached to a blunt-ended ds DNA fragment is allowed to react with the corresponding enzyme, it is cleaved and a sticky end generated in the process.

A sticky or cohesive end may be generated in a blunt-ended fragment by another method. This involves selective degrading of one of the strands using an appropriate ($5' \rightarrow 3'$ or $3' \rightarrow 5'$) exonuclease. Conversely, a fragment with a recessed 3′ end can be blunt-ended by extending the complementary recessed strand with the help of a polymerase that does not possess exonuclease activity, such as the Klenow fragment of DNA Pol I. Use of the Klenow fragment also avoids the formation of hair-pin shaped branches, often found after artificial synthesis of a DNA strand along a given template. In practice, the Klenow fragment of the DNA Pol encoded by the phage T4 DNA Pol gene is used to fill up the single-stranded tail with a complementary strand.

One can insert any number of useful cutting sites within a stretch of a given DNA; a linker DNA, containing one or more recognition sequences, may be inserted in the recipient DNA at the required site. Other desired modifications of DNA, in terms of sequences required for cleaving with restriction enzymes, may be made by judicious use of appropriate restriction enzymes, polymerases and other nucleases.

Some restriction enzymes cleave at the stated sites only under specific conditions. These cutting sites and/or recognition sequences may become different if reaction conditions are altered. One example is provided by *Eco* RI which cuts at the G A A T T C sequence only when the pH is 8 and the salt concentration is 50 mM. If the pH is higher than 8 and the salt concentration is above 50 mM, and if, in addition, glycerol is also present in the mixture, the *Eco* RI cleaves the sequence G A A T T C either between G and A or T and C. In the natural scheme of events, base substitution mutations can occur anywhere in the DNA. If either G or C is substituted with A or G or AATTC with AGTTC, new sticky ends will be generated by the *Eco* RI, under altered conditions of reaction.

One of the requirements for restriction enzyme activity is an absolutely pure DNA substrate. Hence, unlike the case for transformation by exogenous DNA, restriction enzyme cutting demands impurity-free DNA. Special measures are taken, therefore, to rid the isolated DNA of molecular contaminants.

Restriction enzymes (store bought or home-purified) require special buffers for the cleavage reactions. The latter may be stopped either by heating (5 min) at 65°–70°C or by the addition of a 'stop mixture', which contains one or more agents that denature and thus inactivate the enzyme.

The restriction enzyme cleavage reaction is carried out using small quantities of DNA in solution within glass test tubes or microfuge tubes held in ice-baths. The activities of these enzymes are affected by various factors that include pH, ionic strength, presence of organic solvents or divalent ions, as well as inappropriate temperature. Needless to say, all operations are to be performed under aseptic conditions. Exact amounts of substrate and reagents are a must for successful digestion. This is ensured by using fixed volume micropipetters fitted with sterile disposable pipette tips. As in other cases of handling DNA, the tips of the disposable pipette end should be cut back sufficiently to allow the DNA to pass through without being sheared.

The time required to completely cut up 1 μg of DNA or pBR322 is taken to be the standard unit of activity of a restriction enzyme. Controlling the length of time of the reaction allows the DNA to be digested fully (that is, at all cutting sites) or only partially. Both results are useful for specific purposes.

A rough calculation may be made of the number of sites that may exist in a DNA for a particular restriction enzyme. The smaller the number of bases in a sequence, the larger the possibility for its recognition site in a DNA. The number of sites may be estimated by using the equation

$$\text{Frequency of sequence} = \left(\frac{1}{4}\right)^{N}$$

where N represents the number of bases in the recognition sequence. However, this equation is not necessarily true for every enzyme and every sample of DNA.

Some DNA, though expected to possess cutting sites for a number of different enzymes, are actually found not to possess them. In general, if large fragments are desired from an unmapped DNA it is safe to use an enzyme that cuts at a 6 or 7 base sequence. 4-base sequences occur with greater frequency and with shorter distances between them in the same DNA. In using a 4-base site, therefore, there is always the danger of cutting through a gene. Restriction enzyme sequences occur both within and outside the boundaries of a gene sequence, with a crowding usually noticed near genes that code for polypeptides. Strictly speaking a 'gene' is defined as all the sequences needed to express a polypeptide, but occasionally regulatory or control sequences have been called 'genes' instead of the more appropriate term 'elements'.

The restriction enzyme solution usually contains 5 per cent or less of glycerol. As mentioned earlier, larger proportions of glycerol may affect the activity of the enzyme. The enzyme solution is stored at –20°C. The glycerol acts as an antifreeze.

All work with restriction enzymes should be performed in the cold room, with the reagents ice-chilled. The enzyme should be allowed to be exposed to 4°C only for short periods of time, to avoid inactivity.

If a DNA is to be cleaved by more than one restriction enzyme, all traces of the earlier enzyme should be removed before setting up the reaction mixture for the subsequent enzyme. This is achieved by extracting the protein (enzyme) from the solution with a 1:1 mixture of phenol:chloroform/isoamyl alcohol (24:1).

Cleavage reactions are carried out in small quantities (usually 15–25 µl for which Eppendorf tubes are suitable).

Each restriction enzyme works best in a particular buffer solution. However some common buffers are now in use that are suitable for several enzymes. Some enzymes require a higher salt concentration and others a low one. If more than one enzyme is to be used for the same sample, it is customary to use the buffer with the low salt concentration first, and then to add (i) the appropriate amount of the salt to this cocktail to give the higher percentage of salt, and (ii) the second enzyme.

The reaction is stopped either by heating at 70°C for about 5 minutes or by the addition of a *restriction* endonuclease *stop-solution.*

Exercise 27: Cutting DNA with Eco RI and Bam HI

Have ready

1. DNA to be cleaved, in a solution in dist. water
2. DNA, as a standard of molecular weight
3. *Eco* RI (Restriction enzyme)
   The enzyme *Eco* RI comes in a lyophilized form and is kept stored at –70°C. It is revived for use according to the instructions of the supplier. The enzyme can be stored at –20°C in a buffer containing 50% glycerol.
4. *Eco* RI buffer 10 X (High salt):

1 M Tris HCl	1.0 ml	pH 7.5
50 mM $MgCl_2$	0.02 ml	
20 mM β-mercaptoethanol	0.16 ml	
500 mM NaCl	0.12 ml	

5. Sterile dist. water
6. A slanting stand to hold Eppendorf tubes

7. Eppendorf tubes 1.5 ml
8. Eppendorf centrifuge
9. Phenol
10. Chloroform:Isoamyl alcohol (24:1)
11. Ethanol.

Procedure

I. To prepare reaction mixture

1. Add to an Eppendorf tube the following in the order given:

(i) Sterile dist. water	60 µl
Eco RI buffer 10 X	30 µl
DNA soln. (1 µg/ml)	200 µl
Eco RI	10 µl
Total: Volume*	300 µl

*Usually not more than 200–300 µl mixture is taken in a tube.

2. Mix the components by quickly turning the tube upside-down once, or by tapping the tube with fingertips.
3. Spin the tube for 2 seconds in the Eppendorf (*optional*) for settling all the liquid at the bottom of the tube.

II. Enzymatic reaction

1. Incubate the mixture at 37°C by holding* the tube in water at 37°C or in an incubator.
   *This is done by inserting the tube in a hole in a stryrofoam slab that is floated in a water-bath (37°C).
2. Withdraw 10 µl aliquots of the mixture at intervals of 20 minutes, with the last aliquot allowed to incubate a total of 70 minutes from the start.
3. Stop reaction in each withdrawn aliquot, by heating the tube at 65°C (in a water-bath) for 10 minutes. This inactivates the enzyme.

*You may now directly run a gel with the stopped reaction mixture**.

or

Extract digested DNA *and then run a gel**

*Add Bromophenol blue to the mixture to be run on a gel.

7.6.2 Joining with Ligase

Ligase is an enzyme that joins a free 5′ end of a deoxynucleotide with a free 3 end of another by a phosphodiester bond using an energy source such as ATP or NAD. Ligase may be extracted from either *E. coli* or *E. coli* infected with phage T4. The T4 phage ligase is more useful than the *E. coli* ligase as it is capable of ligating both staggered and blunt-ended DNA. The T4 ligase is expressed from *gene 30* of the phage. This phage is of the virulent type, which liberates 50–100 phage particles per cell per infection particle. Extraction of T4 DNA ligase, therefore, involves repeated infectious cycles of the phage and intermittent extraction of the enzyme. However, a better method is in use. The T4 *gene 30* is introduced into the DNA of a lysogen. At 42°C the T4 ligase gene is very efficient, so that excess gene product becomes available in the host cell. The enzyme is rapidly extracted from the induced cells.

Ligating two DNA fragments to make a hybrid molecule may be achieved both inside a cell or in a test-tube. Construction of recombinant molecules in a test-tube has several advantages over those joined inside the cell. Within a cell the ends of fragments to be joined may become altered by the action of different cellular nucleases. The nucleotides at the cohesive or blunt ends may be removed before ligation occurs. This would alter the recognition sequence of the restriction enzyme used to cleave the DNA, and prevent the hybrid DNA from being cleaved again by the same enzyme for recovering the vector and inserted fragments in intact forms. In a test-tube, where the main ligating reaction is not hampered by undesirable enzymes, the hybrid DNA is formed with unchanged sticky or blunt ends, and can, therefore, be separated into the joining components by using the same enzyme. The immediate advantage is an increase in the efficiency of transformation. In hybrids made inside the cell using the resident ligase, an unknown quantity of vector and insert DNA pieces become crippled and hence are unable to form the desired hybrid molecules. This lowers the efficiency of transformation.

Ligation, like other enzymatic reactions, depends on several factors that include concentration of free ends for joining, the lengths of the DNA pieces, the temperature, ionic concentration of the reaction mixture, the base composition of the ends to be joined, ratios of vectors and inserts, level of ATP and so on. The optimum temperature for hydrogen pairing of sticky tails is around 5°C to 15°C. That for ligation is closer to 37°C. A compromise is usually made and 12.5°C often used as the temperature for annealing and joining operations which are carried out simultaneously. A low concentration of ATP (0.5 mM) is favourable for T4 ligase joining of blunt and staggered ends. But an excess of ATP (2–5 M) prevents blunt-end ligation.

The ligase does not join all the ends simultaneously, but in a progressive manner. Formulae are available for calculating the optimum conditions for the joint or separate operations of annealing and ligation.

DNA ligation occurs during repair of nicks in the DNA strands, and during recombination and replication of DNA. Restriction enzyme cutting of DNA and joining of cut DNA backbones by ligase are the two basic operations on which rests the entire superstructure of recombinant DNA technology.

Exercise 28: Joining DNA Fragments with T4 DNA Ligase

T4 DNA ligase is added to a reaction mixture containing appropriate buffers, the DNA fragments to be ligated, and some additional ingredients. These ingredients include, in some protocols, spermidine (spermine tetrahydrochloride), ATP, DTT (dithiothrietol) and BSA (bovine serum albumin). Protocols for DNA ligation vary in details. The addition of spermidine cleans the DNA fragments by removing a variety of impurities and condensing the DNA. This results in a clean DNA precipitate. The DNA from which spermidine has not been removed may be run on an electrophoretic gel, as the mobilities of the DNA fragments are not altered for all practical purposes.

The activity of the T4 DNA ligase is expressed in Weiss units. One Weiss unit of T4 DNA ligase is equivalent to the amount of the enzyme required to mediate the catalysis of 50% ligation of *Hin* dIII fragments of λ DNA in 30 minutes at 16°C in

20 µl of the ligation mixture, having a concentration of 0.12 µM (~ 300 µg/ml) of 5′ DNA termini. The amount of the enzyme required varies according to the nature of the DNA ends to be ligated. Blunt-ended joining requires 0.1 units per 15 µl of the reaction mixture, while 0.01 units suffice for ligation of sticky-tailed ends.

Some protocols prefer mixing the fragments to be joined, followed by dialysis of the mixture against TEN buffer (1 litre at 22–25°C for 3 hours) before ligation.

Have ready

1. T4 ligase (0.01 units per 15 µl reaction mixture)
2. Ligation solutions:

Solution I	Tris	100 mM	pH 7.5	
	KCl	60 mM		
	$MgCl_2$	100 mM		
	DTT	10 mM		
Solution II	Spermidine	10 mM		
	DTT	100 mM		
	BSA	0.1 M		
	ATP*	10 mM	pH 7	*ATP is stored frozen

3. Ligation buffer (10 X):

Solution I	10 ml
Solution II	10 ml
Dist. water	80 ml

   Store at 20°C in 10 ml aliquots. Thaw before use.
4. Stop mix soln. (to terminate the ligation reaction):

Urea	5 M
Glycerol	10%
SDS	0.5%

   Bromophenol blue 0.025%
5. pBR322 (vector) cut with *Bam* HI, 1 µg
6. *E. coli* chromosomal DNA (insert) cut with *Bam* HI, 3 µg
7. Eppendorf microfuge tubes 1.5 ml
8. Gilman pipettes with disposable tips

Procedure

1. Prepare the following reaction mixture in a 1.5 ml Eppendorf tube.

Solution I	4 µl
Solution II	4 µl
Cleaved vector DNA	5 µl
Cleaved *E. coli* DNA	5 µl
Distilled water	22 µl
Total volume	40 µl

2. Withdraw 5 µl from the above mixture for using as a control.
3. Add to the remaining 35 µl of the reaction mixture—Ligase 1 µl (about 4 units).
4. Incubate the above at 15°C overnight or at 4°C for about 2 days.
5. Transfer 2 µl from the above to a fresh Eppendorf tube. Store the remainder at –20°C.

6. To the 2 μl ligation mixture (Step 5) add 8 μl stop mix soln.
7. Keep the above at 65°C for 5 minutes.
8. Run a minigel electrophoresis to assess the extent of ligation, with one well containing the control (unligated) DNA fragments (Step 2). Ligation is incomplete if the pBR322 band is still evident.
9. (a) If ligation is complete, thaw the stored ligation mixture and fractionate it by electrophoresis. Use the ligated DNA eluted from the different bands.
   (b) If ligation is incomplete, thaw the stored ligation mixture and add to it
   T4 DNA ligase 1 unit
   ATP 5 μl
   and keep at 4°C for another 2 days.

7.6.3 Removal of the Terminal Phosphate

The enzyme that can cleave off the phosphate moiety from a 5′ end of DNA is known as alkaline phosphatase. Both *E. coli* and calf intestine alkaline phosphatase are used (Fig. 7.16).

It is necessary to remove phosphate moieties from the 5′ position of a terminal nucleotide for the following reasons:

1. To prevent a DNA fragment from becoming circularized by the joining of its terminal 5′ phosphate and 3′–OH groups.
2. To replace a PO_4 group with a radioactively labelled one.

7.6.4 Addition of Labelled Phosphate

There are occasions where it is necessary to add a phosphate to the terminal 5′ end of a polynucleotide. This reaction is catalysed by an enzyme polynucleotide kinase (Fig. 7.17). Kinases are enzymes that phosphorylate molecules. There are differences in kinases that depend on the substrates on which they operate.

One need to phosphorylate a terminal nucleotide of a DNA or RNA arises when the end is to be ligated to a 3′–OH of the same or a different polynucleotide. Certain projects require end-labelled polynucleotides (e.g., DNA sequencing). In these cases, the kinase is used to replace the phosphate from a 5′ end with a ^{32}P labelled moiety. The radioactively tagged P is derived from the ^{32}P-labelled γ-phosphate of ATP.

The polynucleotide kinase used generally is expressed in *E. coli* cells by a kinase gene of phage T4. The T4 kinase is also capable of phosphorylating the 3′–OH end of a single nucleotide.

After the dephosphorylation reaction, the polynucleotides are first purified by gel electrophoresis to ensure that small nucleic acid fragments are removed. Otherwise the kinase used for the next reaction will add ^{32}P to these unwanted fragments, and much of the labelled phosphates will be wasted.

An important precaution to be taken while using polynucleotide kinase is to ensure that no ammonium salts are present in the reaction mixture, as ammonium ions inhibit the kinase-catalysed reaction.

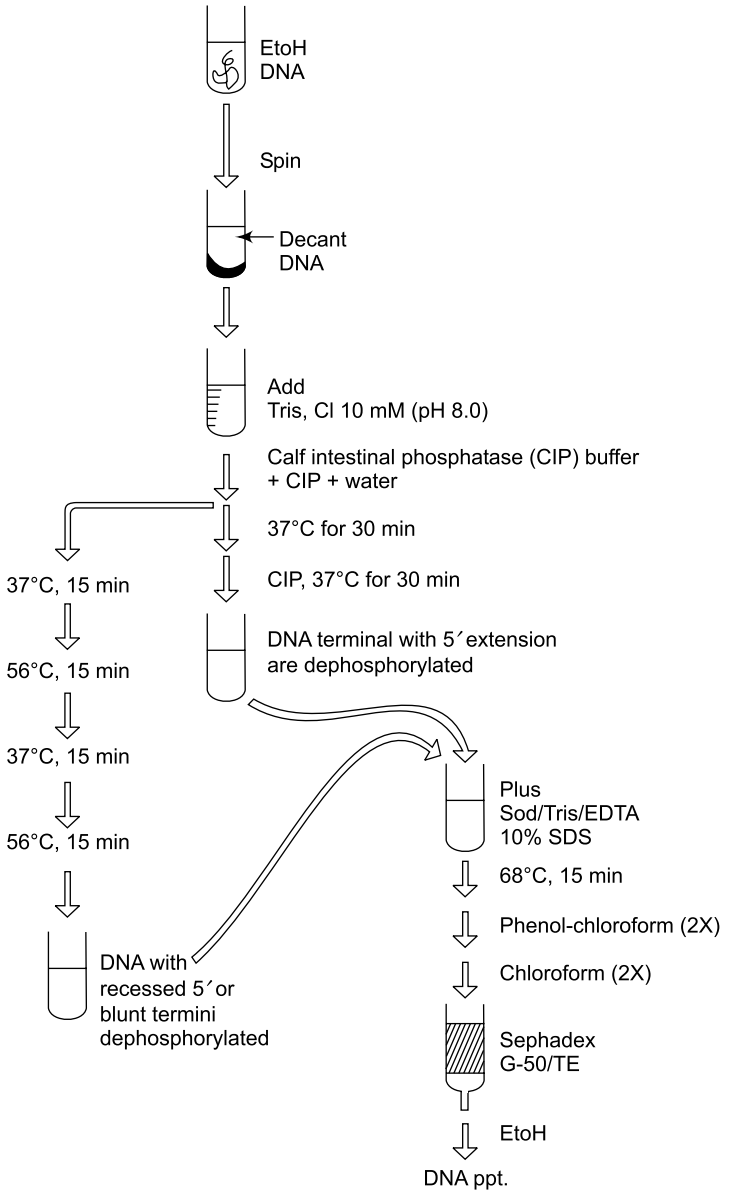


Fig. 7.16 *Dephosphorylation of ends of DNA strands using CIP (calf intestinal phosphatase)*

7.6.5 Adding Linkers

Linkers are synthetic pieces of DNA that are 8- or 10-base long. Each includes the recognition site of a particular restriction enzyme. If it is necessary to splice a mechanically sheared (broken) or blunt-ended DNA(x) to a restriction fragment (y), a linker with the cohesive end of the restriction fragment is attached to the end of the x DNA using T4 DNA ligase. The linkers available commercially provide a

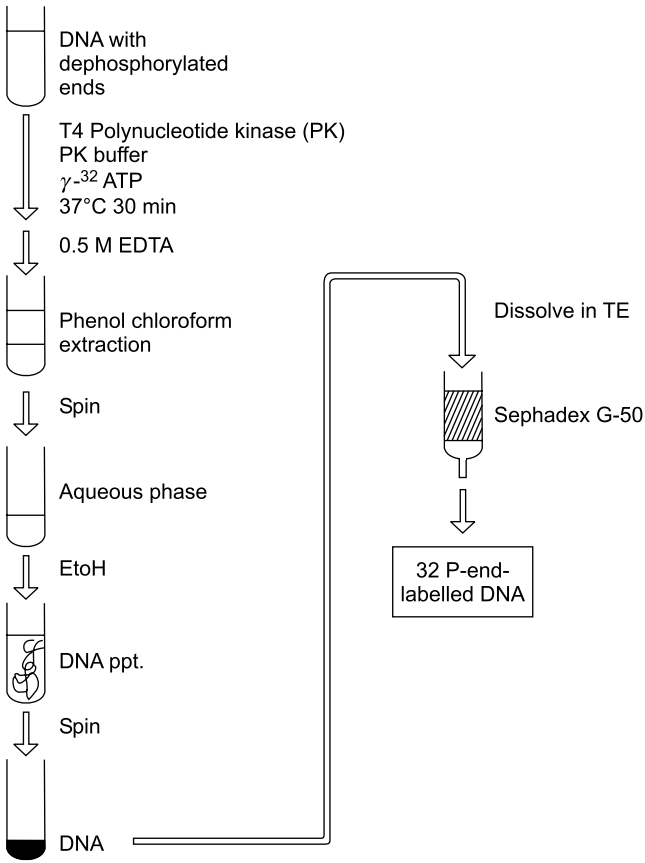


Fig. 7.17 *End-labelling of DNA with ^{32}P using T4 polynucleotide kinase (PK). PK buffer (10 X): Tris.C1 0.5 M (pH 7.6); $MgCl_2$ 0.1 M, Spermidine 1 mM, EDTA 1 mM,TE: Tris-EDTA.*

tremendous versatility for tailoring the ends of DNA strands, such as (1) for inserting known short lengths of DNA between a donor DNA and a vector, (2) for making cohesive ends for insert DNA that will anneal with those of the vector DNA, and (3) for creating new sites for cleavage for a restriction enzyme and adding known restriction sites for cloning in a given DNA. Linkers of different lengths may be made by ligating 3 or 4 linkers containing sites for the same enzyme and recleaving the joined molecules again with the same enzyme. (4) 'Adaptor' or readymade cohesive end is made for a blunt-ended DNA by anealing two different enzyme linkers of unequal size that have 3–4 consecutive bases of one complementary to those of the other. The 5′ end of the longer piece has a hydroxyl group, which prevents joining of this end to the blunt end of the insert DNA.

7.6.6 Adding Poly(d)Nucleotide Tails

A DNA polymerase, terminal nucleotidlyl transferase or TNT, possesses, besides the properties common to other DNA Pols, the property of extending a polynucleotide

strand without referring to a template. That is, it requires a free 3′–OH end of a nucleotide, base-paired to another strand to begin the chain elongation, but it can extend the chain beyond a minimum of three bases in the complementary strand if nucleotides are available. The cell uses TNT to repair long gaps in damaged DNA by creating a filler piece containing random nucleotides. *In vitro,* this enzyme is useful for extending a short single-stranded tail of a double-stranded DNA molecule. When vector and insert DNA are cut with the same enzyme, both have mutually complementary tails. However, these are very short; they are not long enough to keep the hybrid molecule components aligned stably. In practice, therefore, the tails of the vector are lengthened by adding about 100 identical nucleotides and those of the insert DNA by a similar number of nucleotides complementary to those used for the vector tails. TNT that mediates the synthesis of a polyadenine strand is known as poly(A) polymerase.

Such homopolymer tails in vectors and complementary homopolymers in the insert DNAs ensure that: (i) there is no illegitimate joining of vectors and inserts in their own subsets, (ii) very long complementary tails prevent the vector and insert regions of the hybrid DNA from becoming separated. A hundred or so hydrogen bonds make the aligned DNA regions definitely more stable than only 4–6 base containing tails.

TNT is extracted usually from plant sources and thymus tissues.

7.6.7 To Make Restriction Maps

When very little information is available about the patterns of sequences in a DNA sample, the first step to characterize the DNA would be to cut it up with one or more restriction enzymes and then to arrange the fragments in the correct order. The sizes of the fragments may be estimated by a variety of methods, the most commonly used one, as a first step, being a comparison with the molecular weight standards co-run in the same electrophoretic gel.

A restriction enzyme (Type II) physical map gives us the number of cutting sites (as well as their locations) for the DNA in terms of the enzymes utilized. The DNA is thus characterized in terms of specific fragment lengths. Each fragment may be mapped in further detail either by the same method or by other methods. The exact order of bases in each fragment or sub-fragment may then be determined (sequencing).

A restriction enzyme map is made by using, in the first phase of operations, an enzyme that cuts the DNA in not too many places (usually one with more than 4 bases in the recognition sequence). After cutting, the ends of the fragments are dephosphorylated to prevent religation of the cut ends. The fragments are correctly fractionated on a gel, and the correct order of the fragments in the intact DNA is assessed by different methods.

One method of physical mapping is to digest the DNA only partially so that the DNA is cut only once in a specified period of time. This method may be combined with the labelling of one end of the DNA with ^{32}P. When the fragments are separated on the gel and an autoradiograph is made of the DNA bands, only the labelled fragments will be registered on the film. The different sites for cutting may then be

established by arranging the labelled fragments in order of size, since all of the labels are at the same end of the DNA (Fig. 7.18).

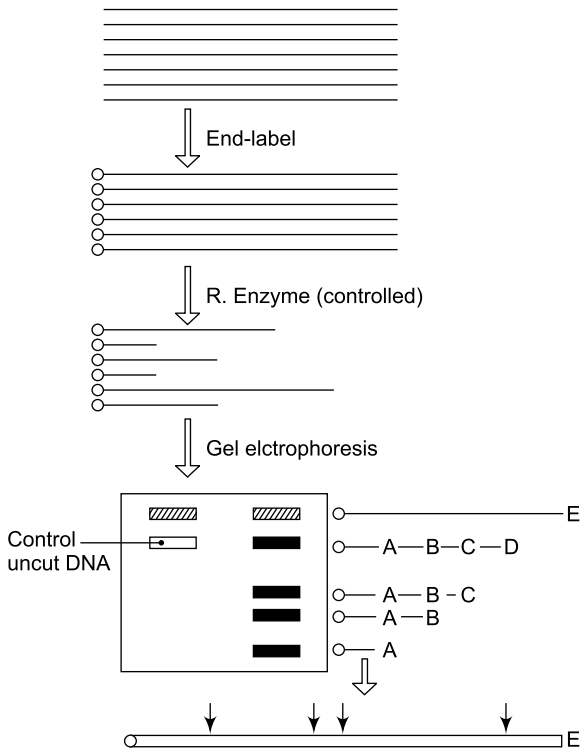


Fig. 7.18 *Restriction mapping: Partial cleavage of a DNA with a restriction enzyme results in fragments due to cleavage at one or only a few of the available cutting sites. As the sample DNA is end-labelled, only fragments with the label (therefore, beginning from the same terminal) show up in the autoradiogram. The fragments are arranged in the order in which they occur in the intact DNA using the data thus obtained. Two or more enzymes may also be used for restriction mapping, strategies have to be employed to distinguish fragments of the same size but from different regions in the sample DNA.*

A second modification of the restriction enzyme mapping technique is to cut the DNA, first with one enzyme (X) then with another (Y) in separate reactions, and further cut the fragments in each set with the enzyme of the other. The correct order and positions of the cutting sites for both enzymes may be constructed from the data, like putting together the pieces of a jigsaw puzzle.

Apart from its use in general physical mapping of a DNA, restriction enzyme mapping is also useful for locating certain genes which appear to be associated with one or more restriction enzyme recognition sites preceding the gene sequence. Particularly, polymorphism of these sites often exists. Differences in restriction fragment lengths or RFLP (RFL polymorphism) then become cues for locating certain genes.

The art of restriction mapping has become extremely refined in several laboratories. Clones of restriction fragments are kept in banks of phage λ vectors or in cosmids and are used for both mapping the entire genome and laying hands on a particular segment of the genome for further study or application.

7.7 ❑ TRANSFER OF NUCLEIC ACIDS AND PROTEINS FROM GELS TO NITROCELLULOSE PAPER

One of the key advances in recombinant DNA technology is the development of techniques for transferring the fractionated bands of DNA, RNA and proteins from impermanent (and often fragile) electrophoresed gel blocks to a stable solid support. The technique consists of 'blotting' the bands from the gel to a suitable paper. The paper first utilized was nitrocellulose filter paper. This paper is still widely utilized; however, the later improvements include specially treated papers that retain RNA (also DNA) and protein and/or its ligands, and allow reuse of the paper for different probes.

The dried 'blots' on the paper can be searched for particular nucleic acids or proteins with the help of suitable probes. For the nucleic acids, the strands in the paper must be single-stranded to allow hybridization with a suitable probe. For proteins, tagged antibodies or antigens serve as the tracker.

7.7.1 Southern Blotting

E M Southern has developed a method of transferring DNA fragments from the bands in an electrophoretic gel to a more permanent support, where they can be manipulated further. The technique is based on the strong affinity of DNA for nitrocellulose paper. The DNA fragments in the gel bands are eluted with a buffer solution into a closely apposed nitrocellulose paper. The adsorbed fragments become immobilized on the paper in the positions in which they are transferred (see Fig. 4.23). The DNA in the 'Southern blotted' paper is then available for further manipulation, the most prevalent one being screening of the DNA for a particular sequence with a radioactively tagged probe.

Hybridization of the DNA on the Southern blot with a probe is possible only if the DNA signal is single-stranded. Therefore, the gel with the fractionated DNA fragments is first treated with a weak alkaline solution (5 M NaCl and 0.5 M NaOH) and then neutralized using a NaCl solution in a Tris-HCl buffer (pH 7.4).

The blotted filter paper is baked at 65.8°C in a vacuum oven for about 2–3 hours in order to fix the DNA strands firmly on the paper. The Southern blot is now ready for screening with a radioactive probe. The filter paper is placed inside a boilable plastic bag together with a pre-hybridizing and a hybridizing solution. The former contains Denhardt's reagent, and some denatured heterologous DNA and the latter solution has the probe that has been denatured by heating at 65°C for 5 minutes—all in a SSC (standard sodium citrate) or SSPE (standard sodium phosphate EDTA) buffer solution. Denhardt's reagent, which contains Ficoll polyvinyl pyrollidone and bovine serum allumin (BSA), prevents single-stranded DNA from binding to the filter paper without hampering hybridization of the bound ss DNA with the probe.

In case the DNA fragments to be transferred are 10 kbp in size, it is customary to partially cleave the DNA in the gel for more efficient transfer of the fragments to the filter paper. The partial cleavage can be achieved by depurinating the DNA by soaking the gel in HCl followed by denaturation and neutralization treatments.

7.7.2 Northern Blotting

RNA bands can also be transferred from electrophoresed gels to a solid substrate without displacement of the relative positions of the bands. Initially nitrocellulose paper was used for blotting RNA from gels, and the technique became known as Northern blotting.

Northern blotting is more efficient if the RNA molecules are denatured before tansferring to the blotting paper. Most RNA molecules are single-stranded and possess a secondary structure on account of formation of stems and loops due to pairing of complementary sequences within the molecule. The RNA is straightened out using one of the following strong denaturing agents:

(i) Glyoxal alone, or in combination with DMSO (dimethylsulphoxide),
(ii) Formaldehyde and formamide,
(iii) Methyl mercuric hydroxide, and
(iv) Urea.

The RNA denatured by (i) can be directly blotted on to nitrocellulose paper. Use of the other two agents requires a few additional steps before effective transfer can be made. Of these latter agents, methyl mercuric hydroxide reacts with free radicals of polyacrylamide gels and hence only agarose gels should be used if this agent is employed for denaturation of RNA.

Nitrocellulose filter paper (0.45 μm) strongly binds to denatured RNA. The RNA with a secondary structure does not adhere to this paper very well. Another solid support for transfers has been developed which binds equally well both denatured and non-denatured nucleic acids. This is called DBM (diazobenzyloxymethyl) paper, and is converted just before use from ABM (aminobenzyloxymethyl) paper which is made from Whatman's paper in advance and stored at 4°C. ABM can also be bought. ABM paper is made by treating Whatman's paper sequentially (with appropriate washings, and dryings between steps) with n-nitrobenzyloxymethylpyridinium chloride, benzene and acetic acid. The paper is finally desiccated and stored at 4°C. DBM paper is made by treating ABM paper with $NaNO_2$ dissolved in HCl and later washed with a chilled sodium borate buffer. The diazotized paper may be reused several times without appreciable loss of the nucleic acid (signal) unlike nitrocellulose paper which is not reusable. In any case, ethidium bromide is added later to stain the nucleic acid, and not mixed with the sample. Acridine orange may be used to stain the glyoxal and formamide-treated gels. If methyl mercuric hydroxide treatment is used, it is removed before transferring the RNA to the blotting paper, with a treatment that includes NaOH and β-mercaptoethanol.

7.7.3 Western Blotting

The technique of transferring proteins fractionated on an electrophoretic gel to a solid support is known as Western blotting, extending the nomenclature system begun by using Southern's name for DNA transfer to nitrocellulose paper.

Usually, in the case of protein transfer, the paper to be used is pretreated in an appropriate manner to also identify, or probe for, a particular protein. The analytical methods used for such identification include (i) immunodetection, (ii) binding of proteins to ligands immobilized on the filter, and (iii) binding of tagged molecules to proteins immobilized on the filter.

Proteins are usually fractionated on polyacrylamide gels, and transferred to nitrocellulose or DBM paper that has been similarly pre-treated. Before transfer, however, the DBM paper has to be washed free of molecules such as glycine that react with diazonium groups in the DBM.

For immunodetection of a specific protein in the gel, the DBM paper is treated with the appropriate antisera. The latter (antibody to the searched protein) may be unlabelled or labelled. If unlabelled, another labelled antibody that matches the immobilized antibody may be used after the proteins have been transferred to the paper. The proteins matching the immobilized antibodies will bind with the latter, and these complexes can be identified by a variety of ways.

The second method of analysis involves the interaction of a ligand, immobilized on the DBM paper, with the protein to be detected. The ligand may be a lectin, an antibody or an antigen. The transferred protein has affinity for its matching ligand and, therefore, complexes with it. The complexes can be detected by autoradiography if the transferred protein is radioactively labelled. The immunodetection method is a variant of this filter affinity technique, where immunodiagnostic methods are employed to identify the transferred protein.

The third method reverses the positions of the protein and the binding molecule with respect to the blotting paper. In this case the protein is immobilized on the paper, and the latter treated with the appropriate labelled molecules such as DNA tagged with ^{32}P or protein labelled with ^{125}I. Protein-protein binding may be detected by labelled antisera and protein-nucleic acid associations with autoradiography.

The choice of nitrocellulose or DBM paper depends on a variety of factors. DBM paper requires prior preparation and binds the proteins for a limited time period. However, DBM paper can be reused after immunodetection. Nitrocellulose paper, on the other hand, may be used as is. It binds proteins for longer periods but is not reusable after immunodetection.

Exercise 29: Southern Blotting and Hybridization of DNA

Have ready

1. The gel with the DNA to be transferred
2. A glass dish ($33 \times 22 \times 4.5$ cm). A baking dish will also do.
3. NaOH 0.5 M solution
4. NaCl 1.5 M solution; 3 M solution
5. Tris-HCl 0.5 M solution (pH 7.0)
6. Plastic gloves
7. Nitrocellulose paper, same size as that of gel
8. Whatman's filter paper No. 3 MM—3 sheets 26×16 cm each.
9. Whatman's filter paper No. 2 MM—10 sheets—same size as that of gel
10. SSC 10 X; 2 X
11. Glass rod

12. About 90 sheets of newspaper cut to the size of the nitrocellulose paper
13. A pair of forceps to handle filter papers
14. Glass plates that are slightly larger than the size of the gel.
    A stack of plates should be at least 1 cm high.

Outline of procedure

A. Separate DNA fragments by electrophoresis on an agarose or polyacrylamide slab gel.
B. Denature the DNA in the gel.
C. Blot DNA onto the nitrocellulose paper.
D. Dry the Southern blotted paper.

Procedure

A. To separate the DNA fragments

1. Load the wells in the gel with the DNA samples.
2. Run the gel, until the DNA fragments are separated into bands.

B. To denature the DNA

1. Transfer the gel to a glass baking dish, wearing plastic gloves.
2. Flood the gel for 15 minutes with:
   NaOH 0.5 M
   NaCl 1.5 M
   Remove the excess liquid with a Pasteur pipette.
3. Repeat Step 2. This denatures the DNA in the bands.
4. Soak the gel for 15 minutes in NaCl 3 M and Tris-HCl 0.5 M (pH 7.0). Remove the excess liquid as before.
5. Repeat Step 4.

C. To blot the DNA

1. Stack the plates in the baking dish (washed and dried) (see Fig. 6.16).
2. Place 3 sheets of Whatman's No. 3 MM paper on the topmost plate.
3. Place on this 10 sheets of 2 MM paper.
4. Pour about 750 ml SSC 10 X solution to a level just below the top of the stack of 3 MM paper.
5. Press out air bubbles from the stack of papers by rolling a glass rod back and forth over the top paper.
6. Place the gel on this stack of papers.
7. Float the nitrocellulose paper, on SSC 2 X, so that the paper is wetted thoroughly. Again all air bubbles must be removed.
8. Immerse the paper for a few minutes in the SSC 2 X solution.
9. Wet the upper surface of the gel with SSC 2 X and place the soaked nitrocellulose paper over it. Remove the trapped air bubbles.
10. Place a stack of 2–3 Whatman's paper No. 3 MM over the nitrocellulose paper. Over this place the stack of newspapers.
11. Place a glass plate on top of the stack of papers and place a heavy weight (about 1 kg) over the assembly. A bottle containing a litre of water would do.
12. Leave the assembly for 12–18 hours.

13. Remove the top stacks of paper and then the nitrocellulose paper, using a pair of forceps.
14. Soak the nitrocellulose paper for 15 minutes in SSC 2 X.

D. To dry the Southern blotted paper

15. Place this paper between two layers of 3 MM paper and leave at 65°C overnight. The blotted DNA will be firmly immobilized on the nitrocellulose paper. This is the Southern blot. It is next treated with a probe (radioactively labelled) that hybridizes with complementary sequences in the blotted ss DNA fragments.

Hybridization

Have ready

1. A boilable plastic lunch bag (Sears and Roebuck Co., USA)
2. Hybridization solution (Formamide 50% SSC 4 X)
3. Radioactive probe ^{32}P RNA or DNA denatured for 10 minutes at 100°C in SSC 0.1 X
4. Water-bath at 65°C
5. Container for radioactive waste
6. SSC 0.5 X
7. RNase A (10 µg/ml) in SSC 2 X

This is boiled first at 80°C for 30 minutes and then cooled slowly.

8. Tris base pH 9
9. Arrangement for autoradiography

Procedure

1. Cut one end of the plastic bag and insert the Southern blotted paper.
2. Pour 15 ml of hybridization solution and the probe solution into the bag.
3. Force out all air bubbles from the surface of the filter paper.
4. Heat-seal 3 sides of the bag, close to the edges (1 cm) of the filter paper. Snip off one corner of the paper to indicate orientation of the paper.
5. Heat-seal all sides of the bag.
6. Incubate the bag in a 65°C water-bath for 18–48 hours.
7. Remove the bag from the bath, cut open one side, pour out the solution into the radioactive waste container. This probe may be reused.
8. Wash the filter paper at room temperature in 200 ml of SSC 0.5 X.
9. If the probe used was an RNA, remove unhybridized RNA with RNase A.
10. Wash the hybridized filter paper in 200 ml of SSC 0.5 X for 15 minutes with gentle shaking.
11. Repeat Step 10.
12. Wash in Tris base (pH 9) 2–3 times, for 5 minutes each.
13. Dry the paper.
14. Autoradiograph the dried hybridized Southern blot.

The hybridized bands of DNA will show up as exposed regions in the autoradiographic film (X-ray film).

Exercise 30: Colony Hybridization on Nitrocellulose Paper

The colonies to be screened for a particular DNA sequence are first transferred by blotting to nitrocellulose filter paper. These colonies are then treated *in situ* with solutions that lyse the cells exposing the DNA and denaturing the latter. The probe hybridizes the ss DNA, which are then identified by autoradiography. The corresponding colonies on the master culture plate are pinpointed as the ones carrying the desired DNA.

Have ready

1. The material required for pre-hybridization, hybridization and autoradiography as in Exercise 29.
2. Petri plates (90 mm) with colonies (*E. coli*) to be screened
3. Nitrocellulose papers cut to fit inside the petri plates (90 mm). These are autoclaved and dried.
4. Whatman 3 MM filter papers cut to fit ≃ 150 mm petri plates.
5. 3 large petri plates (diameter 150 mm)
6. NaOH 0.5 M solution
7. Tris 1 M pH 7.4.
8. ST solution:
   NaCl 1.5 M
   Tris 0.5 M pH 7.0
9. Fine forceps with bent tips (sterile)

Procedure

1. Place a dry nitrocellulose paper (NP) disc with a pair of forceps over the colonies in the petri plate.
   Press without displacing the paper, making sure that there are no air bubbles trapped between the plate surface and the disc of paper.
2. Incubate the paper-covered colonies (covered plates) in the incubator (37°C) for 2–2.5 hours.
3. Pour a little NaOH into a large petri plate. Place one large Whatman paper disc in the solution until the paper is wetted thoroughly. Decant off the excess solution.
4. Lift the nitrocellulose paper, with the pair of forceps, from the surface of the colonies. Turn the colony side up and place on the NaOH soaked paper. Pour a little NaOH around the edge of the NP. Keep for 5–10 minutes. The cells in the colonies are lysed by this treatment, and the DNA denatured by the mild alkali.
5. Repeat Step 4 with a Whatman 3 MM disc soaked in Tris 1 M. Pour 5–10 ml of Tris around the NP and keep for about 2 minutes.
6. Repeat Step 4 with the ST solution. Keep for 5 minutes.
7. Dry the NP in air for 2 hours.
8. Bake the NP in a vacuum oven at 80°C for 2–4 hours.
9. Process the NP for hybridization with a given radioactive probe and autoradiograph as in Exercise 29.

The positive clones may be picked up from the master plate and amplified further. These clones carry the desired DNA sequence.

7.8 ❑ MUTAGENESIS

DNA lesions may be caused by a variety of agents: physical, chemical and biological. A lesion is made permanent in the genome as a defect or mutation only after replication(s) of the initial affected DNA.

Mutagens may be naturally occurring ones, or products of man's technologically advanced civilization. Awareness of the long-term genomic damage that may be wrought by the mutagens has led to the development of several tests that try to assess the mutagenic potential of (i) chemicals introduced into the environment by various industries, and (ii) electromagnetic radiations used for both military and peaceful operations. The concern has appreciably increased since most carcinogens also happen to be mutagens.

7.8.1 Sources and Types of DNA Damage

The type of initial DNA damage or lesion depends on the mechanism of action of the concerned agent vis-a-vis the DNA molecule. Broadly speaking, the lesion may be (i) at a single nucleotide site, (ii) an extended region in one strand, or (iii) a gross perturbation in the genetic map of a chromosome.

Physical mutagens include electromagnetic radiations and heat. Ionizing radiations (X-rays, γ-rays) break the sugar-phosphate backbone(s) of a DNA, often indirectly via hydrolytic action of radiation-generated free radicals in the cell sap. X-rays break one or both chromatids in a chromosome. The broken ends may rejoin illegitimately forming aberrations that can be usually distinguished readily in a cytological preparation. Nonionizing radiations, such as short-wavelength UV (254 nm) cause the formation of pyrimidine dimers between bases on the same or on the opposite strands. Such dimer formation causes distortion of the DNA structure and prevents replication of the molecule. Heat, a potent physical mutagen, mimics the effects of some chemical mutagens.

Similarly, alkylation favours the altered base to mispair in the subsequent DNA synthesis. Adducts may range from simple alkyl groups (methyl, ethyl) to much larger multi-ring compounds such as aromatic amines, and polycyclic aromatic hydrocarbons (PAH or PCAH), and fungal products such as aflatoxin. Guanine with its many sites for adduct formation (O^6 , C^8 , N^7 , or the N^2 outside the rings) is more prone to alteration than the other bases. The potential for causing mutation, however, depends not only on the nature of the adducts but also on the location of the lesion. For instance, adding a methyl group to a G at N^7 is less mutagenic than the one at O^6 of the same base. On the other hand, aflatoxin complexed at the N^7 of G causes gross distortion of the DNA, that is reflected in subsequent mutagenesis.

Additions and deletions are usually due to distortions in the DNA structure followed by faulty lengths of DNA after replication. Certain chemicals, like acridine dyes, intercalate between base-pairs in the DNA and distort the backbones abnormally. According to one suggestion, when repair is initiated by nicking one strand at the distorted region and completed by replacing it with a newly synthesized one, the nicked or intact strand may buckle, so that more or less bases may be inserted at the corrected region. On replication the product will contain one normal

and one lengthened or shortened DNA. There are other mechanisms also by which a region of the chromosome may be duplicated, the initial reasons for which are not always clear.

Additions and deletions generally cause frame-shift mutations, in which an altered pattern of amino acid residues replaces the original one in a polypeptide.

Biological mutagens include DNA elements (transposons or viral DNA or cDNA that become integrated into the DNA of the infected or transfected cell. Such insertions may cause (i) disruption of a unique genetic message that needs to remain intact for correct expression, and/or (ii) cause various rearrangements of the base sequences of the recipient DNA on either or both sides of the insertion.

7.8.2 Some Mutagenecity Tests

Several tests are available for testing agents for both short-term and long-term effects. Short-term effects are due to direct toxic reactions of the agent in the organism. The testing of potential toxicity of chemicals involves at least the following operations:

1. Estimation of the LD_{50}, a measure of the dose required to kill (lethal dose) half the population of treated organisms.
2. Assessment of the types of epidemiological changes wrought by the treatment in the whole organism, as well as in specific organs, and
3. Investigation of the mode of action of the agent at the molecular level.

The conclusions from the above tests are used to establish confidence parameters or safety doses for chemicals used in foods, fads, drugs and pesticides.

In addition to the short-term effects, or even in the absence of obvious immediate effects, the mutagen may generate long-term genetic alterations, even with doses of the agent that are not considered toxic. These long-term consequences include carcinogenesis and teratogenesis. The later refers to abnormal development of the embryo in response to an agent, which is then referred to as teratogenic. All three effects—mutagenic, carcinogenic, teratogenic—are essentially the outcome of a lesion in the DNA that failed to be corrected and became perpetuated.

In view of the above, it is customary to test a sample first for its mutagenic potential. Those showing positive results are then subjected to more rigorous tests for carcinogenecity and teratogenecity. The tests for the latter are more time consuming and expensive in terms of materials. The test animals for these latter screens are mostly rodents; in some cases simian species are employed.

Mutagenic tests aim to be simple and quick. Preliminary tests use bacterial or yeast cells as test material. Bacteria and haploid yeasts are preferable, as recessive mutations become readily identifiable in haploid cells. Damage is scored as the percentage of mutations that appear in the test cells after treatment with the test chemical. It is assumed that the effect of a chemical on the DNA of these simple cells will not be basically different from that of on the DNA of humans. It is, however, not firmly established that this is so.

The complexities in the human body may thwart the primary effect of a chemical or convert it after metabolism to a different mutagen that has effects different from those seen in the bacterial cells. There are tests, therefore, that try to assess the damage in terms of chromosomal aberrations. In plants this is done by treating the growing root-tips with the chemical, making slides of the cells undergoing mitosis

and scoring aberrations under the microscope. In humans, peripheral lencocytes may be treated in test-tubes with the chemical, and the cultured cells observed later during mitosis arrested at metaphase.

The tests employing bacteria score the percentage of mutations caused in a given locus. One set of tests is based on the induction of reverse mutations and another on forward mutations. The first type, developed by B N Ames at the University of California at Berkeley, is the *reverse* mutation test, and scores the reversion of *his* strains of *Salmonella typhimurium* to *his*$^+$ strains. The second type of test was developed in the laboratory of W G Thilly, at MIT, Massachussetts and also uses special *S. typhimurium* strains that change from a wild-type to a mutant at a particular locus; this is a *forward* mutation test. Both techniques have their plus points and are employed accordingly.

1. The Ames Test

A carefully selected set of *his*$^-$ strains of *Salmonella typhimurium* is employed for the Ames test. The *S. typhimurium* operon for synthesizing histidine contains the genes for at least ten enzymes that mediate as many reactions in the biosynthetic pathway. Each *his*$^-$ strain owes its defect to a specific type of base-substitution or frameshift event in the gene for one of the ten enzymes. A test-chemical may be specific in the manner in which it generates a mutation. By using a collection of strains with different *his*$^-$ loci, it is possible to make sure that whatever the mutagenic specialty of the chemical, it will be identified by the conversion of one or more *his*$^-$ strain cells to the *his*$^+$ counterparts.

The strains used for the test carry, besides the *his* mutation, at least two other mutations that give credibility to the results obtained by these assays. An assay may indicate a very low, if not nil, percentage of mutations, merely due to the fact that the test-chemical was unable to penetrate the cell well. Again the results may be on account of missing lesions due to their subsequent correction by repair enzymes. The mutation *rfa* allows penetration of foreign matter into the test bacteria, while the *uvrB*$^-$ provides a strain deficient in excision-synthesis repair. So any lesion caused by an agent remains uncorrected and a mutation ensues following DNA replication. The *his* cells that revert (mutate) to *his*$^+$ ones will be now able to proliferate in the histidineless medium.

The basic Ames test is carried out by plating a *his* strain in a soft agar containing a very little quantity of histidine in the medium, over a petri plate of hard agar containing no histidine. The bacteria use up the histidine in a couple of hours and are unable to proliferate further. The test-chemical is placed in the centre of the plate in a solution that either wets a small disc of filter paper or is in a well cut out in the centre of the agar plate. In either case the chemical diffuses radially into the surrounding agar and comes into contact with the *his* cells. If the chemical is mutagenic it will induce mutations in some of these cells. It will, in fact, cause mutations randomly at any loci, which of course includes the different *his*$^-$ loci. The *his*$^+$ (revertant) cells are obvious as colonies, as they are able to proliferate in the histidine-deficient medium, while the *his* cells are unable to do so (Fig. 7.19). The hist colonies will appear in a ring around the source of the mutagen. The density of the colonies in a ring, the width of the ring or band as well as the radius of the ring provides data about the strength of mutagenic potential of the test-chemical.

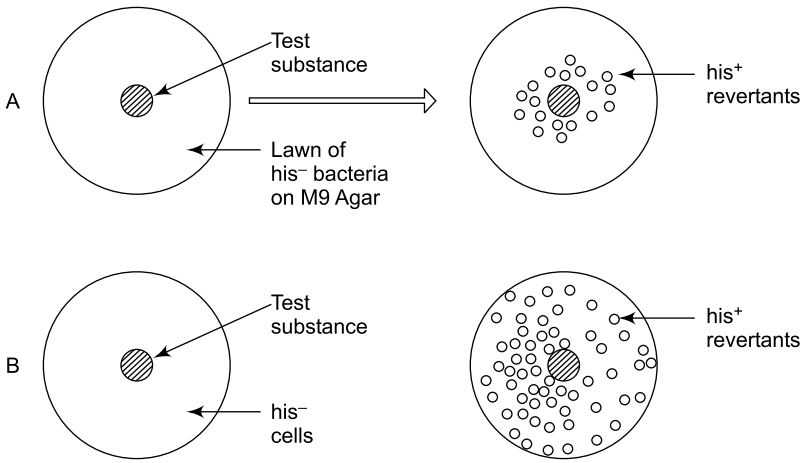


Fig. 7.19 *Ames Test for mutagenecity of a substance, his cells (mutants) of Salmonella typhimurium are allowed to come into contact with the test substance (placed in the centre of a petri dish lawn of the bacterium). If the sample is mutagenic, some of the bacteria will mutate (back-mutation) to the wild-type (his⁺) version of the gene locus. The degree of mutagenecity is calculated from the size of the back-mutated population, as well as the dilution of the sample sufficient to be mutagenic. Of the two samples of A and B in this figure, the former is less mutagenic than the latter; only a strong concentration of A mutates the his locus, whereas even dilute concentrations of B (far from the central well or sample-soaked disc of filter paper) are mutagenic.*

The basic test is, however, unable to provide the correct answers about the mutagenecity of chemicals that do not directly cause DNA lesions. There are chemicals which only after metabolism within an animal provide the products that cause DNA damage. Such chemicals thus induce mutations indirectly. To ensure that even indirect chemical mutagens are identified by this test, the sample is presented to the *his* bacteria after incubation with enzymes that convert the primary chemical into metabolic products. The enzymes are mostly oxygenases and are present in about all tissues, including those of plants. In practice, they are extracted from the livers of rodents which have been pre-treated with substances that induce overproduction of these oxygenases in liver cells. The cells are homogenized and the extract centrifuged. The S9 fraction contains microsomes rich in the required enzymes. This microsomal or S9 fraction is added to the test-chemical before launching an Ames test. The substances used for inducing overproduction of the microsomal enzymes are polychlorinated biphenyls or PCBs. Arochlor 1254 (Analabs, USA) is often used for this purpose. Livers from PCB-untreated mice or rats may also be used to extract the S9 fraction.

The strains of *S. typhimurium* used for the Ames test (TA98, TA100, TA1535, TA1537, TA1538, etc.) are indicators of GC substitutions or of frame-shift mutations. To make sure that mutations due to AT pair substitutions are also picked up, other strains have to be used. *E. coli* strains have been developed that pinpoint such a substitution. A *S. typhimurium* strain that also indicates alternations at AT pairs and is sensitive to a large spectrum of mutagens has been developed in the laboratory of S. Venitt, at the Institute of Cancer Research in Sutton, UK. The

E. coli strains sensitive to AT substitutions are auxotrophic mutants for tryptophan synthesis; the *trp E* reverts to the wild-type *trp* E^+ condition and can be picked up as colonies on a tryptophanless medium.

The actual test is carried out simultaneously with all the above strains. The chemical and the S9 fractions are used in a range of doses (concentrations). In one set the chemical is used without the S9 treatment, in another together with the S9 fraction. Appropriate controls are maintained for each set. Since there is always a low base-level incidence of spontaneous mutations in any cell population, plates without chemical (with and without S9) are scored as well, and the value of revertant obtained is deducted from that found in the other test plates.

2. The Forward Mutation Test

Thilly and associates have developed a forward mutagenecity test using *rfa* and *uvrB* carrying *Salmonella typhimurium* that promises to be a quicker and less elaborate toxicity test than that provided by the Ames reverse mutation assay. Only one tester strain is sufficient to screen for the mutagenic potential of a test chemical, when using Thilly's test.

The forward mutation assay uses the inability of a mutant to convert a chemical into toxic metabolites inside the cell. An analog of guanine called 8-azaguanine is transported into a cell by an enzyme bound to the plasma membrane and also converted to a phosphorylated version inside the cell. If a chemical is mutagenic, it knocks out one or more genic or regulatory regions of this enzyme. The outcome, whatever the nature of the lesion, is an inactivated enzyme. If the enzyme is in form, the phosphorylated 8-AG interacts with the DNA and precipitates the actual toxic activity and kills the cell. A mutated enzyme fails to induce such toxicity. Cells that are resistant to the 8-AG are, therefore, mutants for the enzyme (xanthine phosphoribosyl transferase). The strain first used by Thilly (TM35) was a revertant (*his*$^+$) of the Ames strain TA1535. A derivative of TM35 that contains a plasmid with a drug resistance gene (pKM101) was found to be a better candidate as a tester strain. This strain, TM677, showed the same degree of sensitivity to identifying a lesion-causing chemical as five of the TA strains used by the Ames test.

The forward mutation test may also be used with the addition of the S9 fraction.

7.8.3 Some Useful Marker Mutations

An abnormality in gene expression is best indicated by a defective function. A mutation of the coding region of a gene results in an abnormal gene product or in an absent one. Mutations in control elements show up as irregularities in the time, place and rate of expression of the gene product. Mutations have been used in genetics to identify the locus of genes as well as to elucidate the mechanism of their functioning.

Haploid cells are particularly useful for detecting mutations, as even one recessive allele is expressed in the absence of a dominant one. In bacteria, the most commonly used mutations are the auxotrophs that are deficient in their ability to synthesize basic macro-molecules from inorganic and small organic precursors. If it is necessary to find out about the biosynthetic pathway for a particular amino acid, a set of mutants

have to be collected that prevent the following of the pathway due to a defect in one of the enzymes that mediate the reactions in the series. If a recombinant cell is to be identified, the system uses a readily identifiable drug-resistance gene. Thus, to map the circular DNA of *E. coli,* conjugation is allowed between a cell containing several auxotrophic alleles (leu , pro , thr , his , etc.) and an Hfr strain with wild-types of these alleles (leu^+, pro^+, etc.). The cells which have received DNA regions from the donor Hfr will become prototrophic (leu^+, pro^+) and be able to grow on minimal media.

There are other metabolic activities which cannot be detected by auxotrophic or drug-resistance type of mutations. Activities such as DNA replication, repair or recombination rely on the efficiency of the concerned enzymes at the temperature that is optimal for the growth of the organism. Mutations in the genes of such enzymes may make it impossible for their products to function in the optimal range of temperature or other parameters, rather than knock out the products altogether or make them defective. It may, however, function at a different (lower or higher) temperature (or under other conditions). By observing the activity of the mutated gene at the altered temperature (or other conditions), it is possible in many cases to decipher the actual function of the gene product. Such mutations are known as conditional mutants; the ones that function in an altered range of temperature are known as temperature sensitive (*ts*) mutants.

Auxotrophic, drug-resistant and *ts* mutants are essential tools in genetic analysis. Eukaryotic cells are also vulnerable to mutations in their extra-chromosomal DNA. All eukaryotes possess the organelles called mitochondria (sing., mitochondrion) in which the major activities of cell respiration are carried out. Mitochondria possess a few or several hundred copies of a small plasmid like circular DNA. The mtDNA (or mit DNA) carries genes for the rRNA and tRNA molecules used within the organelles, as well as genes for parts of some of the respiratory enzymes, such as cytochrome c. Mutations in mitochondrial genes show up as abnormalities in respiration. The cells are not as efficient in extracting energy from energy rich molecules and grow very slowly compared to wild-type cells. This phenomenon is seen dramatically in the yeast, *Saccharomyces cerevisiae* which produces respiratory deficient cells that form small or 'petite' colonies among the larger normally metabolizing ones.

Petite mutants are, therefore, due to alterations in genes of proteins required for cell-respiration. As indicated above, all the respiratory proteins are not encoded in the mtDNA. Some of them are from genes in the nuclear DNA. This sharing of mitochondrial protein genes between the chromosomal and organelle DNA results, on mutation, in three types of defects. They are the segregational, neutral and suppressive petites. The first is due to nuclear genes and is subject to the laws of Mendelian inheritance. The latter two are mutations in a gene called *rho* (ρ) in the mtDNA and show non-Mendelian inheritance, characteristic of cytoplasmic or extra-chromosomal genes. Neutral mutations result from a loss of the *rho* or mtDNA (ρ^0) and the suppressive ones are due to a mutation in the mtDNA (ρ^-). The three types of petites may be easily isolated by suitable recombination exercises.

Mutations in the other prominent extrachromosomal DNA, the ch1DNA, found in the green organelles called chloroplasts in plant cells, also result in defective photosynthetic functions. The ch1DNA is also small, circular and occurs in several

single, interlocked or concatennated copies of the same genetic map. They have been especially useful for studying the roles of nuclear and chloroplast genes in the photosynthetic set-up.

It is relatively more difficult to collect strains of higher eukaryotes with readily identifiable marker genes. The predominant diploid phase of the life cycle of an eukaryote masks the presence of a single dose of a recessive allele in a genome. Studies using single cell cultures of plant, animal and human cells, however, would not have been successful without the use of at least a few such marker genes The ones employed widely in animal and human cell lines are the mutations for the *tk, dhfr, gpt* and *hgprt* loci. Auxotrophic and drug-resistant mutations are also being continuously searched in lines of single cell cultures of plant species.

Exercise 31: Screening UV-induced Mutations in E. coli

A short wavelength UV source will be used for this exercise. The tube may be covered with a cardboard jacket with an opening cut out for the UV rays. Goggles should be worn and the hands should receive minimum exposure to radiation during the exercise. Mutagenesis is undertaken only after the dose required to kill 90 per cent of the cells is determined.

Have ready

1. A wild-type strain of *E. coli*
2. LB broth and LB agar plates
3. PBS (pH 7.0)
4. 15 W UV lamp or tube (short wavelength, 254 nm) with a window in a cardboard jacket
5. Sterile tubes for dilution
6. Pasteur pipettes, alcohol lamp etc.
7. A 100 ml conical flask
8. A timer

Procedure

1. Make an overnight culture of *E. coli* in 10 ml of LB broth.
2. Add 0.4 ml of the overnight culture to 20.0 ml of fresh LB in a sterile conical flask. Shake in a 37°C water-bath incubator for 3 hours. Check OD. The density of the cells should be about 10^8/ml.
3. Arrange a series of sterile test-tubes with 9 ml of sterile PBS in each. Add 1 ml of the above culture to tube No. 1. Make dilutions in the remaining tubes by transferring 1 ml of fluid from tube No. 1 to tube No. 2, then 1 ml from tube No. 2 to tube No. 3 and so on until the last tube also contains 10 ml of PBS culture. This gives a series of tubes with the culture diluted from 10^1 to 10^9 strength of the original culture.
4. Turn on the UV lamp now (in the Laminar Flow Cabinet).
5. Keep controls by plating cells from dilutions 10^5–10^9 on LB agar plates. Keep in the incubator at 37°C.
6. Pour the diluted culture from each tube (10^{-1} to 10^{-9}) into a similarly labeled sterile petri dish (10^{-1} to 10^{-9}). Cover the dishes.

7. Place two covered petri dishes on a rectangular piece of plywood (or a small tray) about 20 cm below the 'window' in the UV source. Remove the covers of the dishes.
8. Turn on the timer. Swirl the contents of the dishes intermittently, using both hands.
9. Remove the tray with the dishes after a period of 30 seconds (or some other length of time, depending on data from a preliminary exercise that indicates the dosage that kills about 90 per cent of the cells). Swirl the suspension with your hands.
10. Expose the remaining dilutions (2 plates at a time) for the required period of time (same for all dilutions).
11. Pour the mutagenized cells on LB agar plates and incubate at 37°C for 18–24 hours.
12. Count the number of colonies in each plates.

N.B. The ultraviolet light causes the formation of pyrimidine dimers which are repaired by both light and dark repair systems of repairing damaged DNA. When not repaired, the distorted DNA strands obstruct DNA replication and cause cell death. To assess the degree of repair, aliquots of the mutagenized cells may be kept in the dark (by covering the tube with Al foil) at 4°C and under a table lamp overnight, before plating out on LB agar plates. The plates after incubation at 37°C for 24 hours will indicate the differences in survival in the three sets of treated cells and, therefore, in the degree of repair in the light and dark post treated cells.

Exercise 32: Screening for 'Petite' Mutants of Saccharomyces cerevisiae

A fairly high percentage (0.5–2%) of yeast (*S. cerevisiae*) cells mutate spontaneously into 'petites'. Higher frequencies of petites may be generated by deliberate mutagenesis. Ethidium bromide will be the mutagen used in the following exercise, as it is known to cause lesions predominantly in mtDNA. Chromosomal genes for respiration are more vulnerable to mutagens that induce adduct formation (e.g., alkaylating agents such as MNNG, EMS etc.) and to the non-ionizing UV radiations.

Petite mutants are observed as very minute colonies that fail to grow on non fermentable carbon sources (e.g., glycerol). To gauge the deficient respiratory function quantitatively, the respiratory quotient of petite and normal cells may be measured with the Warburg's apparatus.

Genetic crosses may be made with normal yeast to ascertain whether a mtDNA petite is neutral (rho°) or suppressive (rho$^-$). The progeny of neutral and normal yeast cells will all be normal, while that of suppressive and normal will give unpredictable ratios of petites and normals—a behaviour typical of extrachromosomal traits.

S. cerevisiae comes in two haploid mating strains *a* and α (alpha), referred to as Mat *a* and Mat α respectively.

In the following exercise, one set of plates will be scored for spontaneous petites and one set each for petites generated by ethidium bromide and UV irradiation. The petite colonies and controls will be amplified and their respiratory indices estimated.

Have ready

1. A haploid strain of *S. cerevisiae* (wild-type).
2. YPD medium with agar:

Bacto-yeast extract	10 g
Bacto peptone	20 g
Dextrose	20 g
Bacto agar	20 g
Dist. water to make	1000 ml

   Autoclave
3. YPD Plates:
   Pour 20–25 ml of YPD agar into 10 plates. Allow to set and keep at 30°C until use (at least 24 hours).
4. YPD Slants:
   Pour 250 ml of YPD agar into 20–35 sterile test-tubes and make slants. Keep at 30°C until use (at least 24 hours).
5. YPDG with agar (YPD + Glycerol). Make on Day 4:

Bacto-yeast extract	10 g
Bacto peptone	20 g
Dextrose	1 g
Glycerol	30 ml
Bacto agar	20 g
Dist. water to make	1000 ml

   Autoclave
6. YPDG Plates:
   Pour YPDG agar medium into sterile dishes and store at 30°C until use.
7. Ethidium bromide:
   This is carcinogenic and should be handled with gloved hands and spills are to be avoided.
   The solution is made and stored in a dark-coloured bottle that is wrapped with Al foil.
   Dissolve:

Ethidium bromide	10 mg
in sterile dist. water	1 ml

8. Sterile test-tubes with 9 ml of sterile dist. water in each for dilutions of yeast culture.
9. 1 ml Pasteur pipettes
10. Sterile petri dishes
11. Sterile dist. water
12. Glass spreader (L-shaped glass rod)
13. Rotating disc on tripod for spreading cells on dishes (optional)

Procedure All operations are to be carried out in the Laminar Flow Cabinet.

1. *Day one* Make the YPD plates and slants. Keep at 30°C until use.
2. *Day two* Streak one loopful of cells from stock tube on each of 6 YPD plates. Incubate at 30°C for 48 hours. (Keep the petri dishes in the top-up position.)
3. *Day four* Streak a single colony in each of the YPD slants. Incubate at 30°C for 48 hours.

Day six

4. Add to each YPD slant 1.5 ml sterile dist. water. Shake to get cells into suspension.
5. Arrange 6 dilution tubes, each containing 9 ml of sterile dist. water and label them 10^{1}, 10^{2}... 10^{6}.
6. Make dilutions, beginning with 1 ml of cell suspension in 9 ml of sterile dist. water (tube 10^{1}).
   Make 2 tubes per dilution and label the sets I, II. Set I will be used to score spontaneous mutations, set II for UV induced mutations. The undiluted cell suspensions will be used in set III for mutagenesis with ethidium bromide.

Set I

7. Plate 0.2 ml of each of dilutions 10^{4}, 10^{5}, 10^{6} per plate on 3 each of YPD, and YPDG plates (18 plates). Incubate at 30°C for 48 hours.

Set II

8. Pour each of dilutions 10^{4}, 10^{5}, 10^{6} into separate petri dishes and expose to UV irradiation (at 25 cm from lamp for 15 minutes as in Exercise 31).
9. Distribute the irradiated cells from each dilution into three each of YPD and YPDG plates (0.2 ml cell suspension/plate). Incubate at 30°C for 48 hours.

Set III

10. Prepare cell suspension (1.5 ml sterile water/YPD slant) and transfer the suspension to a sterile centrifuge tube. Spin at 1000 g for 5 minutes. Resuspend in sterile dist. water (check cell count: should be 10^{7}/ml).
11. To 5 ml YPD liquid medium in PBS (pH 6.24): add 0.05 ml of the cell suspension and 0.01 ml of the ethidium bromide stock solution. Cover the vial with Al foil and incubate, with shaking, at 30°C for 24 hours.
12. Spin the cells and wash the pellet with sterile dist. water. Adjust cell count with sterile dist. water to 10^{3}/ml.
13. Plate 0.1 ml of above per YPD and YPDG plate (3 each) and incubate at 30°C for 48 hours.

Day 8

14. Examine the incubated plates of all three sets and count the number of small colonies in each. Calculate the number of petites per number of cells treated in each case.
15. Use the Warburg apparatus to measure the respiratory indices of wild-type and petite cells from each set. Plot the data.

Exercise 33: The Ames-Microsomal Fraction Test

The *Salmonella typhimurium* strains TA98 and TA100 will be used for the following tests. See Appendix for sources of tester strains for mutagenecity tests.

A known mutagen, benz[α] pyrene (BP) will be used in the following exercise. The chemical is dissolved in dimethylsulphoxide (DMSO), a solvent for both polar and non-polar substances.

The S9 fraction is prepared and stored at –80°C until use. It is then allowed to thaw to 0°C, mixed with a buffer containing NADP and Glucose-6-phosphate and filter-sterilized. This generates NADPH in the system. The working S9 mixture may be kept below 4°C until used (a couple of hours). Unused S9 mixture is discarded.

All operations are carried out using ice-cold reagents and under aseptic conditions. S9 is inactivated above 0–4°C.

Procedure

Day 1 to Day 5

I. Preparation of S9 fraction.

Day 2

II. Preparation of Bottom Agar Plates.
III. Preparation of Top Agar Tubes.

Day 4

IV. Start overnight cultures of testers.

Day 5

V. Preparation of S9-mixture (working solution).
VI. Preparation of BP solution.
VII. Melting of Top Agar and maintaining at 45°C.
VIII. Carrying out of the Assay.

Day 6 or 7

IX. Count revertant colonies.
X. Calculate mutagenecity of BP.

Have ready

I. For S9-mixture:
   1. Male rats, each about 200 or 250 g in weight (preferably inbred strain)
   2. PVC gloves
   3. Aroclor (Analabs, or Monsanto Co., USA)
   4. DMSO
   5. Screw-capped vials or Universal bottles (cylindrical with conical bottom, that can stand on their own)
   6. Disposable sterile syringes (1 ml) with 21-gauge (1 inch) needles.
   7. Dissecting board, covered with plastic and absorbent paper
   8. 70% isopropanol in water (in a plastic wash bottle)
   9. At least three sets of instruments, each with a pair each of scissors and blunt forceps
   10. NaCl 0.9% solution, 6 × 500 ml aliquots
   11. KCL 0.15 M solution 6 × 200 ml aliquots. aliquots. Solutions (10) and (11) are kept in screw-capped bottles at 4°C.
   12. Stainless steel sieve (tea strainer)
   13. Sterile screw-capped vials
   14. Sterile Pasteur pipettes

15. Freezing mixture (dry ice + acetone)
16. Fume-hood
17. Homogenizer with teflon pestle (Potter-Elvehjem)
18. Refrigerated Centrifuge (Sorvall RC 5, rotor SW 34).
19. Low temperature freezer (80°C)

II. Bottom Agar with M9 medium:

1. Bacto agar	10.0 g
Dist. Water	440 ml
2. **Solution A:**	
Na_2HPO_4, $2H_2O$	7.0 g
KH_2PO_4	3.0 g
NH_4Cl	1.0 g
NaCl	0.5 g
Dist. water to make	100.0 ml
Solution B:	
$MgSO_4 \cdot 7H_2O$	20.0 g
Dist. water to make	100.0 ml
Solution C:	
$CaCl_2$ $2H_2O$	200 mg
Dist. water to make	100.0 ml
Solution D:	
Bacto agar	10.0 g
Dist. water to make	440 ml

III. Top Agar (10 X):

Solution a (10 X):	
Histidine	10.048 mg (or 0.5 mM)
Biotin	12.22 mg (or 0.5 mM)
Dist. water to make	100.0 ml
Solution b:	
Glucose	40 g
Dist. water to make	100 ml
Solution c:	
NaCl	0.5 g
Bacto agar	0.6 g
Dist. water to make	80.0 ml

IV. Tester strains:
S. Typhimuruim, TA98, TA100

V. Stock Solutions for S9 mixture:

1. Buffer (0.2 M $NaHPO_4$)	
$NaH_2PO_4.2H_2O$	15.6 g
Dist. water	375 ml
pH adjusted to	7.4 (using 1 N NaOH)
Dist. water to make up to	500 ml
Autoclave and store at	25°C

2.	$MgCl_2 \cdot 6H_2O$ (0.25 M)	1.271 g
	Dist. water to make	15.0 ml
	Autoclave and store at	25°C
3.	KCl (1 M)	1.864 g
	Dist. water to make	25.0 ml
	Autoclave and store at	25°C
*4.	Glucose-6-phosphate (0.2 M)	0.282 g
	(Mol. wt. 282.1)	
	Dist. water to make	5.0 ml
*5.	NADP (0.04 M)	0.153 g
	(Mol. wt. 765.4)	
	Dist. water to make	5.0 ml
	*Filter-sterilize using	0.22 μM filter paper

VI. Chemical to be tested and Solvent:
Benzo (α) pyrene (BP)
DMSO
Weighing bottle
Aluminium (Al) foil
Screw-capped glass vials.

VII. Others:
45°C water-bath
37°C incubator
Vortex Mixer
Al foil
Pipettes, test-tubes, petri plates, screw-capped vials
Colony Counter (manual or automatic)
Marking pen
$Na_2 HPO_4 \cdot 2H_2O$

Procedure

I. Procedure for S9 preparation

1. Prepare Aroclor solution in DMSO in the fume hood.
   - Weigh a Universal bottle.
   - Add some Aroclor to it.
   - Calculate the amount of DMSO to be added to make a 200 mg/ml solution.
   - Add the required amount of DMSO. Cap the bottle.
   - Shake the contents until the mixture appears homogeneous.
   - Warm to 37°C before use.
2. Inject each animal intraperitoneally with Aroclor (500 mg Aroclor/kg body weight, i.e., about 100 mg/200 g animal).
3. Maintain the animals with normal food and water for four days, followed by about 18 hours without food and water.
4. The animals are to be sacrificed on Day 5 and the livers dissected out aseptically. Have material and instruments (2, 7–17) ready. The pestle of (17) should be pre-chilled in the freezer.
   - Sacrifice the animal (cervical dislocation).

- Place, ventral side up, over the dissecting board lined with a layer of 2–3 absorbent paper and fixed to the board with dissecting pins.
- Squirt 70°C isopropanol over the abdomen to wet it thoroughly.
- Make an incision on the skin and extend it to expose the muscle layer, using one set of instruments.

5. With the second set of instruments cut through the muscle layer to expose the liver.
6. With the third set of instruments remove the liver and transfer it to a beaker of ice-chilled NaCl solution. Steps (4), (5) and (6) prevent contamination of the dissected levers.
7. After all the livers are collected, transfer them to the sieve placed over a deep petri dish. Wash the livers with NaCl solution until the washings are no longer red.
8. Weigh the livers in a sterile pre-weighed beaker and place the beaker (covered) on crushed ice.
9. Mince the livers with fine scissors in a petri dish containing ice-chilled KCl solution (3 ml/gm of liver).
10. Homogenize the chopped livers in ice-chilled KCl in the cold room with the chilled pestle.
11. Centrifuge at 9000 g (2°C) for 15–20 minutes.
12. Decant the supernatant (S9 fraction) into a sterile conical flask (200 ml) on crushed ice.
13. Distribute the S9 fraction into sterile screw-capped tubes held ready in a bucket containing the freezing mixture. Store at –80°C.

II. To make bottom (hard) agar plates Prepare solutions A, B, C, and autoclave them separately and keep at 25°C.

Dissolve the agar for solution D in 440 ml water, autoclave and cool to 65°C.

Mix the following:

Solution D	440 ml
Solution A	50 ml
Solution B	5 ml
Solution C	5 ml

Pour the above mixture into sterile petri plates (20–25 ml/dish). Store the plates for 24–48 hours at 25°C.

III. To make top (soft) agar tubes Prepare solutions a and b, autoclave and store at 25°C until use.

On day of use, prepare solution c. Autoclave and cool to 65°C.

Mix the following:

Solution c	80 ml
Solution a	10 ml
Solution b	10 ml

Dispense 2 ml of this mixture in each of 20–25 sterile vials, and maintain at 45°C in a water-bath.

IV. To prepare the Salmonella cultures

1. Make an overnight culture (1 loopful from a slant into 10 ml of M9 medium (liquid).
2. Next morning plate the 10 ml on 2–3 bottom (hard) agar plates. Incubate at 37°C for 24 hours. There should be a uniform lawn of cells on the agar surface (Master Plates).
3. Scrape 1 loopful of cells from the master plate and add to it 10 ml of liquid M9 medium. Incubate by shaking at 37°C (overnight). Cell density required: $2–3 \times 10^{8 \text{ or } 9}$/ml.
4. Remove cultures and hold at 37°C in the dark (cover with Al foil).

VI. To prepare the working S9 mixture

1. Take out the number of frozen S9 vials required for the day's assays. Allow to thaw and keep on crushed ice.
2. To make 15 ml of a 4% S9 mixture (v/v): Mix the following:
   (a) Buffer 7.50 ml
   (b) $MgCl_2$ solution 0.48 ml
   (c) KCl solution 0.50 ml
   (d) Glu-6-phos. solution 0.38 ml
   (e) NADP solution 1.50 ml
   (f) Dist. water (deionized) 4.05 ml
3. Filter-sterilize the above through a syringe filter assembly directly into a 50 ml sterile flask kept on crushed ice.
4. Add to the above S9 (ice-cold) 0.60 ml
5. Mix the S9 mixture, shaking by hand, and keep on crushed ice until use.

VII. To prepare the test-chemical solution Work in a fume-hood, with gloved hands:

1. Weigh a 5 ml weighing bottle with the stopper.
2. Add a few crystals of BP.
3. Weigh the bottle, BP and the stopper.
4. Add the amount of DMSO needed to give a concentration of 500 μg BP/ml DMSO. Close the stopper.
5. Swirl the bottle to make a homogeneous mixture.
6. Prepare dilutions of the above stock solution in DMSO in screw-capped glass tubes (5 μg, 10 μg and 30 μg/ml). Cover each tube with Al foil.

VIII. Melt soft agar and maintain at 45°C in a water-bath.

IX. To perform the assay

Have ready

1. Bottom agar plates
2. Top agar tubes at 45°C
3. Cultures at 37°C
4. S9 mixture (kept on crushed ice)
5. DMSO

Procedure

1. To a 2.0 ml top agar in tube
   Add 0.1 ml culture (TA-98)
   0.1 ml BP solution
   0.5 ml S9 mixture (4%)

*2. Immediately vortex briefly and pour on a bottom agar plate. Swirl to spread the soft agar mixture evenly.

*3. Repeat with 2 more samples of the TA-98 strain.

*4. Repeat using the TA-100 strain (X3).

5. Cover the plates with Al foil or thick brown paper while the soft agar sets.
6. Transfer the plates to an incubator (37°C). Leave for 25–48 hours.
   * It may be necessary to use other percentages of the S9 mixture when testing an unknown chemical.

Controls Incubate the following plates (37°C)

	Hard agar	Soft agar	Bacteria	S9 mixture	BP	DMSO
1.	√	—	—	—	—	—
2.	√	√	—	—	—	—
3.	√	√	—	√	—	—
4.	√	—	—	—	—	√
5.	√	√	—	√	√	√

X.

1. Check plates and discard the contaminated ones.
2. Count the number of revertant colonies (appearing as discrete ones) that are visible to the naked eye in each set.

XI.

3. Prepare dose-response curves.

7.9 ❑ INTRODUCTION TO DNA CLONING

The term *DNA cloning* strictly applies to the production of colonies of cells each one of which harbours an identical genotype. Asexual cell division ensures this uniformity in the clone. If a piece of DNA is introduced into one of the cells, and the latter starts a colony, each member of the colony will possess a copy of the added DNA. This is the rationale of introducing a recombinant DNA into a cloning cell.

The introduced DNA may be amplified more than a 1000 fold from a few nanograms to several micrograms. If a single clone containing the inserted DNA is used to inoculate a vigorously aerated liquid culture, the micrograms become milligram quantities within a relatively short period of time.

Amplification of DNA is the primary purpose of cloning it. However, other aims may be served as well. One of them is the purification of a particular species of DNA. Purification becomes necessary as, besides the DNA of interest, the cloning cell may

contain other unwanted species of the molecule. The latter include unhybridized vectors, unligated vectors or donor DNA fragments and DNA fragments other than the ones to be cloned.

In general, pieces of DNA with free ends (that is, linear DNA) are degraded by cellular nucleases. But circular vector molecules and circularized donor DNA, ones that include at least one replication origin each, survive and replicate as easily as the hybrid molecules of interest. We have seen, in Chapter 2, how the cells with the desired DNA may be fished out or identified by using the appropriate screening or selection techniques. Once identified, cells from them are grown to large populations that are genetically identical.

In a less specific sense, the term *DNA cloning* refers to the entire operation of creating and amplifying the hybrid DNA. In this chapter, the term will be used to signify the introduction of a DNA into a cell that amplifies the imported molecule by replication within each cell, as well as by dividing it into a population of identical cells, each one of which amplifies the foreign DNA.

A cell that accepts an exogenous DNA stably is said to be transformed genetically. The introduced DNA, if a gene, adds a new quality to the genotype as well as to the phenotype of the recipient cell. If the import is a regulatory sequence, it may influence the pattern of expression of a resident gene. In either case, the transformed cell is different genetically—even by a single allele—from the parent, undisturbed cell. The term 'transformation' is also used, in context, for those genetically altered conditions that are similar or identical to a pre-cancerous or cancerous type of cell. In such cases the genetic difference lies in the acquisition by the recipient cell of a cancer-causing gene, known as an oncogene, or DNA sequences which due to positioning in the host DNA lead to the activation of a protooncogene in the host cell DNA. The term 'transformation' will be used in this chapter for the general meaning of altering the genotype of a cell by an imported DNA.

By tradition, and due to convenience, the first choice of a cloning cell is that of *E. coli* (See Exercise 34). Other cloning cells, as mentioned in Chapter 2, are also coming into vogue. As DNA replication, unlike gene expression, is carried out equally well in prokaryotes and eukaryotes, it is customary to amplify engineered DNA in *E. coli,* even if the ultimate objective is to express it elsewhere. Shuttle vectors are used for this purpose. Shuttles are available that move with ease between *E. coli* on the one hand, and yeast, *Drosophila,* mammalian cells and *Bacillus subtilis* on the other.

E. coli may be transformed by plasmid, cosmid (see Exercise 36) or phage DNA-borne foreign DNA. The term 'foreign' applies to any DNA taken up by the cell, be it from another cell of the same strain or species or from an entirely different species. Plasmid-carried DNA enters through the cell membrane and remains as an episome in the cell or may become integrated into the host cell DNA. The latter may happen if the imported DNA is homologous to a region of the host DNA or is flanked by sequences that are homologous to host DNA or possess flanking sequences that 'transpose' the DNA into the cell DNA.

Efficiency of transformation by naked DNA is enhanced if the cells are first altered to a state, which for lack of proper characterization is known as that of 'competence'. Cells are made competent by treating them first with an ice-cold

calcium chloride solution followed by heat-shock at 42–43°C. The calcium chloride is adsorbed on the outer surface of cells and in some manner deposits the extraneous DNA on the surface as well. It is not known whether the salt also makes it possible for the DNA to cross the membrane barrier. In any case, the DNA is not taken up unless the cells are treated with the pulse of raised temperature.

Recombinant DNA carried in phage particles transforms many more cells than naked DNA, as the particles adsorb specifically to receptors on the cell surface. Thus every cell becomes 'infected' with the phage, and the DNA in the latter, natural or chimaeric, is injected into the cell. If the vector DNA is a phage genome, the rDNA replicates within the cell, as would that of the natural virus, and becomes amplified eventually into several mature 'infectious' phage particles. The size of the insert DNA, however, has to be fairly small when the vector is a phage DNA, as the latter can spare only a relatively small proportion of its genome for replacement with a foreign DNA. The situation improves with the use of cosmid DNA (see Exercise 36), which is basically a plasmid, the length of a phage lambda DNA, but only with the *cos* region of the phage that is needed for packaging the DNA within the viral shell. Cosmid carried DNA is thus much longer than what is possible in phage DNA vectors. Cosmids packaged in lambda shells have an added advantage. Once the rDNA is released into the cell, the DNA is not repackaged into phage particles—a consequence inevitable for phage vector carried DNA—unless of course the phage vector carries mutations in genes that prevent shell formation and/or lysis of the host cell membrane.

Entry into plant and yeast cells is facilitated by removing the specialized carbohydrate cell wall. The enzyme zymolase degrades the yeast cell walls, releasing the naked cells that round up to form free protoplasts or spheroplasts (see Exercise 37). Plant cell walls are removed with cellulase. Cells of plant tissues are bound together with a cementing material called pectin. The latter may be degraded by pectinase. A combination of cellulase and pectinase removes the pectin and cellulose walls of plant tissues, resulting in free protoplasts. Protoplasts are as vulnerable to transformation by naked DNA as bacterial cells. The transformation process can be enhanced by adding PEG (polyethylene glycol), a fusion agent, to the mixture of protoplasts and the DNA.

The wall-less transformed cells regenerate new cell walls in about twenty four hours. Theoretically, a protoplast can regenerate into a whole plant. In reality, regeneration has been possible in only a very limited number of plant species. Nevertheless, transformation of free plant protoplasts is useful both for studies on the mechanism of expression of plant genes and for the production of plant metabolites of commercial value.

Entry into plant cells is enhanced if the DNA of interest rides on a naturally entering DNA. Species of the bacteria, *Agrobacterium,* infect plant cells and cause proliferation of the infected cells into 'galls' or tumours. The bacteria achieve this courtesy of a plasmid (Ti or tumour inducing) which carries a transposable unit, the T DNA region. A copy of the T DNA becomes integrated in the plant nuclear DNA and its further metabolism leads to a genetical tumourous transformation of the cell. A foreign DNA spliced to the T DNA, or a T DNA containing plasmid vector becomes installed in the plant cell chromosomal DNA and is amplified as the

'tumourous' cell proliferates (see Exercise 38). Efforts are continuing to develop Ti based vectors that will allow the inserted DNA to function as desired, but suppress those of the tumour causing genes.

Animal cells may be transformed by utilizing any one of the methods described in the section on gene transfer into animal cells in Chapter 2. The calcium phosphate precipitation technique is used more often than the others chiefly due to its ease of operation. The micro-injection and electroporation strategies are definitely more efficient in terms of the number of cells transformed per nanogram of DNA. These are preferred in certain laboratories in spite of the added expense and paraphernalia accompanying these operations.

The remainder of this chapter will provide protocols for transforming *E. coli* and yeast cells with plasmid and cosmid borne DNA and for transforming plant protoplasts with the Ti plasmid of *Agrobacterium tumefaciens.* Before launching into the exercises, let us briefly recapitulate the steps involved in DNA cloning, including those for constructing the rDNA.

7.10 ❑ ALGORITHM FOR CLONING DNA

To recapitulate, DNA cloning consists of the following steps (see Exercise 38):

1. Isolation of vector and donor DNA in a fairly pure form.
2. Cleaving of the vector and donor DNAs with one restriction enzyme. The enzymes that cleave a vector, within or outside a marker gene, are known. These enzymes are first tried out on the donor DNA to make sure that there is no cleavage site for them within the stretch of the DNA to be cloned. The ones which pass muster may be utilized.
3. Modification of the vector and donor fragment ends as required.
4. Mixing of vector and donor fragments together with ligase for ligation of vector and donor DNA.
5. The ligation mixture containing hybrid and non-hybrid DNAs is then transferred into a cloning cell, by one of the various gene transfer techniques.
6. Where the donor DNA consists of different sheared or restriction fragments, there may be as many kinds of rDNA as that of the fragments. The rDNAs carrying the DNA to be cloned have to be screened or selected from this heterogeneous population of transformed cells.
   When pBR322 is the vector, the marker genes for AMP^R and TET^R are utilized for selection of the transformed cells. This step is followed by a more specialized selection/screening step that identifies the desired rDNA carrying clones.
7. The identified clones are amplified.
8. A further step of Southern blotting and hybridization with a probe, may be taken to identify the rDNA carrying cells in an unambiguous manner. In some cases, this may be the only strategy that will do the job.
9. The selected clones may be utilized to express a gene on the cloned DNA or the rDNA extracted and used in some other manner.

7.11 ❑ TO TRANSFORM *E. coli* WITH A PLASMID-BORNE DNA

The first step in introducing the DNA into *E. coli* cells consists of making the latter 'competent'.

To make *E. coli* cells competent, they are first soaked in ice-cold 50 mM solution of calcium chloride, then chilled for some time in an ice-bath and finally placed in a water-bath at a temperature of 42–43°C for a brief exposure to this higher temperature. Some investigators find rubidium chloride to be equally effective. Most protocols use $MgCl_2$ with the $CaCl_2$.

The competent cells are next mixed in suspension with the DNA (on a pBR322). The DNA treated cells are plated out on selective media and the recombinant cells identified.

7.12 ❑ TO TRANSFORM *E. coli* CELLS WITH PHAGE LAMBDA CARRIED DNA

Foreign DNA spliced to a phage DNA vector may be introduced into *E. coli* cells by transfection or by packaging *in vitro* into phage shells.

Transfection is the infection of *E. coli* with phage DNA carried foreign DNA that is not enclosed in a phage shell. The method is essentially identical to that used for transforming cells with pDNA. The cells are made competent with calcium chloride and heat-shock treatment before allowing them to mix with the DNA to be transferred.

In vitro packaging involves the enclosing of rDNA (phage vector or cosmid) into phage shells. This operation is carried out in a test-tube, using home-made or ready-made 'packaging mixes', that are available commercially. The mix consists of two mutant strains of phage lambda, each defective for the ability to synthesize a unique shell protein. When an *E. coli* culture is infected with both these mutants, the phage DNA of each is replicated and all save one shell protein synthesized. When these cells are lysed, the lysate contains all the shell proteins in abundance as well as the enzymes and factors needed to construct mature phage particles (see Exercise 36).

When the mix-lysate is allowed to mingle with the rDNA, each one of the latter becomes encapsulated, by the shell proteins in the mix, to form 'infectious' particles. These become adsorbed on recipient *E. coli* cells and inject the rDNA into them.

7.13 ❑ THE DNA TO BE CLONED: DONOR AND cDNA

Recombinant DNA molecules are made with the vector DNA and the DNA to be cloned. The latter may be a fragment from the total or a specific limited region of a donor DNA. Or, it may be the cDNA version of an mRNA or of the RNA genome of a phage or virus.

Donor DNA is cleaved with one or two different specific restriction enzymes to make it fit the cloning site in the vector. A donor DNA fragment may be spliced

in either orientation at the cloning site, as DNA strands are ligated as the result of formation of a phosphodiester bond between the ends of two strands. The inserted DNA will be functional in only one of these orientations. The functionally correct inserted DNA clones will be subsequently screened or selected. However, it may be necessary to ensure the insertion in a specific orientation. In such a case, the vector and donor DNA are cleaved by two different restriction enzymes. Complementarity of ends produced by different enzymes will ensure joining of the DNA in the desired orientation.

When cDNA represents the donor molecule, ends of vector and cDNA may be created that are complementary to each other. This is achieved by adding a poly(d) T tail to the ends of the linearized vector, and poly(d)A tails to the ds cDNA. On the other hand, the cDNA may be synthesized right on one of the ends of the vector DNA (Fig. 7.20). Most eukaryotic mRNAs possess a poly(d)A tail. This tail region may be hybridized to a poly(d)T tail on one free end of the linearized vector. Addition of reverse transcriptase produces a ss cDNA copy, complementary to the attached mRNA. A poly(d)G linker is fixed to the remaining free end of the vector. The hybridization between the G and C 'tails' circularizes the molecule. We are left with a hybrid circular DNA where one of the strands is an RNA. The latter is removed by degradation with RNase H and a new DNA strand takes its place after the addition to the reaction mixture of DNA Pol I, the four NTPs and other factors required for DNA synthesis.

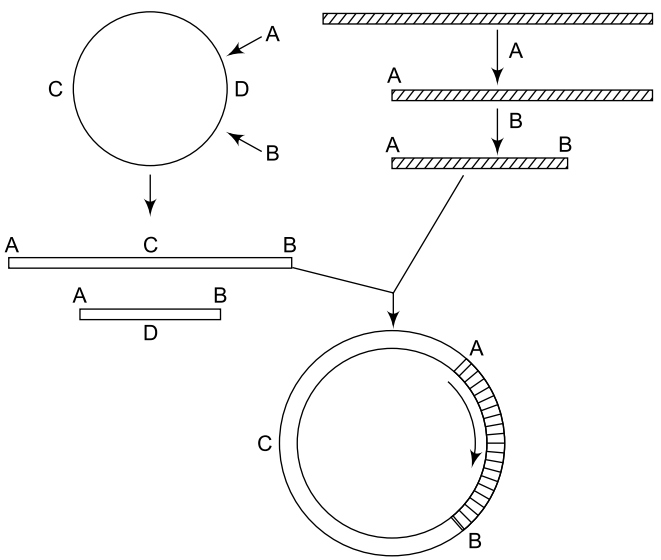


Fig. 7.20 *Joining a DNA fragment to a vector in a desired orientation. The donor and vector DNA are shown as hatched and solid line (or open box) respectively. Donor DNA cleaved with enzymes A and B when mixed with vector cleaved with the same enzymes, becomes inserted in only one orientation. If the fragments were prepared with only one enzyme, they would join the vector in either orientation.*

7.14 ❑ TO IDENTIFY THE mRNA TO BE COPIED INTO A cDNA

If the gene to be cloned is expressed abundantly in a cell, it is relatively easy to isolate its mRNA and use this as a template for synthesizing a cDNA. Of course, there is no guarantee that every cDNA synthesized from the extracted abundant mRNA is the transcript of the gene in question. The sample is likely to include other species of mRNA.

If the mRNA occurs only sparsely in the cell, other strategies have to be employed to detect and isolate it and then use it as a template for cDNA synthesis. The best method to achieve this is to fractionate the total mRNA of a cell according to size (gel electrophoresis or gradient centrifugation), denaturing the mRNA before fractionation. The bands may be Northern blotted and searched by a probe, if available. The identified band of mRNA is then purified and utilized for cDNA generation.

A second approach is available for identifying a cDNA of an mRNA that is present at a low abundance or occurs transiently during a particular phase of a developmental period. The total mRNA of two cells that differ in this low abundance mRNA or a few mRNAs may be converted to cDNA copies and the latter hybridized with the total mRNA from both cells. The odd-man-out will be detected as an extra band in the Southern blot of cDNAs from one of the cells. This non-common cDNA is likely to represent the mRNA that is different.

Another method is to artificially synthesize a sequence from the gene using the amino acid sequence of the gene product. As the genetic code is degenerate, certain practices have to be followed to narrow down the sequence that possibly is nearest to the sequence of the required gene. The synthetic oligonucleotide is then employed as a probe to identify the mRNA or cDNA of the gene to be cloned.

A further strategy is to isolate the required mRNA from polysomes by using monoclonal antibodies that match the polypeptide that is being translated from the mRNA. One of the polypeptide chains of a human MHC gene product (HLA-DR) was trapped in this manner by Korman et al. (1982) and the mRNA isolated from the polysome-polypeptide-antibody complexes. The purified mRNA obtained in this manner then serves as a template for synthesizing die required ss or ds cDNA.

Exercise 34: Transformation of E. coli with pBR322

Strain HB101 of *E. coli* is utilized for transformation with pBR322. HB101 is an F^- strain.

Have ready

1.	*E. coli*	HB101	
2.	$CaCl_2$	50 mM	50 ml (autoclaved)
3.	$MgCl_2$	10 mM	50 ml (autoclaved)
4.	pBR322	2 X	10 µl
5.	LB broth:		
	Bactotryptone		10 g
	Yeast Extract		5 g
	NaCl		10 g
	Dist. water		100 ml
	Autoclave the above.		

6. LB agar plates with 50 μg/ml of ampicillin
7. Saline solution:
   NaCl 9% in dist. water (autoclaved)
8. Ice-bath (use an ice bucket filled with crushed ice cubes)
9. Selection plates:
   (i) Ampicillin-LB
   Add 50 μg/ml ampicillin (filler-sterilized) to LB broth.
   (ii) Tetracycline-LB
   Add 10 μg/ml of tetracycline (filler-sterilized) to LB broth.
10. Sample (transformation tubes)

Tris-HCl	10 mM pH 7.0
$CaCl_2$	10 mM
$MgCl_2$	10 mM
pBR322	1 μg

Procedure

Summary of steps

I. Grow *E. coli* recipient cells.
II. Make cells competent by $CaCl_2$, $MgCl_2$ and heat-shock treatment.
III. Mix competent cells with plasmid.
IV. Add LB broth to the above mixture and spread on selection plates.
V. Score the number of transformants (colonies) on plates.

I. To grow E. coli recipient cells

1. Innoculate 10 ml LB with a loopful of HB101 from the streaked plate.
2. Incubate at 37°C (incubator) overnight.
3. Add the overnight culture to fresh LB broth (190 ml) and mix well.
4. Transfer 50 ml of the above to a 250 ml flask.
5. Incubate at 37°C with vigorous shaking until cell density is about 5×10^8 /ml.
6. Transfer 20 ml of the culture to each of two 50 ml centrifuge tubes and spin in a Sorvall (SS34 rotor) at 4°C, 8,000 rpm for 5 minutes.
7. Decant off the supernatant (or remove with a Pasteur pipette).

II. To make cells competent.

8. Add to the pellet 20 ml of chilled $CaCl_2$ (50 mM).
9. Resuspend the pellet by vortexing 2 ml $CaCl_2$ (50 mM) at 0°C.
10. Stand the tube for 15 minutes in crushed ice (ice-bath).
11. Centrifuge as before (4°C, 8,000 rpm) for 5 minutes.
12. Repeat Steps 8 and 9, using this time 4 ml $CaCl_2$.

III. To mix competent cells with plasmid

13. Add to each sample (transformation tube) standing in the ice-bath 0.2 ml of $CaCl_2$ treated cell suspension.
14. Leave on ice (in the refrigerator shelf) for 25 minutes.
15. Transfer to a 45°C water-bath (heat-shock treatment) for 2 minutes.
16. Remove from the water-bath and keep at room temperature (22°C) for 10 minutes.

IV. To add LB broth

17. Add to each tube 1 ml LB broth (warmed to 37°C).
18. Incubate at 37°C for 2–3 hours.
19. Spread 0.2 ml of the cells on each selection plate (ampicillin and/or tetracycline).
20. Incubate at 37°C (incubator) for 2–3 days. The transformed cells alone will form colonies.

V. To score the number of transformanets (colonies) on plates

21. Score the number of colonies per plate for transformation samples and controls (without heat-shocked cells, and with only heat-shocked cells but no transformation mixture).
22. Calculate the number of transformants.
    (i) per number of cells on plate or per ml
    (ii) per ng of pBR322

N.B. Try a third set of transformations with linearized pBR322. Do you find any difference in the efficiencies of transformation with circular and linear plasmids?

Exercise 35: Enrichment of Tetracycline-Sensitive (Tet^s) Transformants using Cycloserine.

The drug cycloserine is used to kill off those cells that are resistant to antibiotics such as penicillin, ampicillin, and tetracycline. The cells transformed by hybrid pBR322 in which the insert DNA is within the Tet^R gene are selected easily on AMP-TET LB plates. Sometimes the number of transformants are so few in number that they may be overlooked by spot plating on a TET-LB plate. In such cases, addition of cycloserine eliminates the Tet^R cells, leaving the Tet^s cells which contain the vector with the insert DNA.

Have ready

As in Exercise 34 for *E. coli* cell transformation with pDNA, along with the following additions:

1. Transformation tubes containing pDNA and rDNA mixture (i.e., pDNA that has been made hybrid by the insertion of a donor DNA)
2. Cycloserine 150 mg/ml

Procedure

Summary of steps

I. Grow *E. coli* recipient cells.
II. Make cells competent by $CaCl_2$, $MgCl_2$ and heat-shock treatment.
III. Mix competent cells with plasmid.
IV. Add LB broth to the above mixture and spread on selection plates.
V. Score the number of transformants (colonies) on plates.

I. To grow E. coli recipient cells

1. Innoculate 10 ml LB with a loopful of HB101 from the streaked plate.
2. Incubate at 37°C (incubator) overnight.

3. Add the overnight culture to fresh LB broth (190 ml) and mix well.
4. Transfer 50 ml of the above to a 250 ml flask.
5. Incubate at 37°C with vigorous shaking until cell density is about 5×10^8 /ml.
6. Transfer 20 ml of the culture to each of two 50 ml centrifuge tubes and spin in a Sorvall (SS34 rotor) at 4°C, 8,000 rpm for 5 minutes.
7. Decant off the supernatant (or remove with a Pasteur pipette).

II. To make cells competent.

8. Add to the pellet 20 ml of chilled $CaCl_2$ (50 mM).
9. Resuspend the pellet by vortexing 2 ml $CaCl_2$ (50 mM) at 0°C.
10. Stand the tube for 15 minutes in crushed ice (ice-bath).
11. Centrifuge as before (4°C, 8,000 rpm) for 5 minutes.
12. Repeat steps 8 and 9, using this time 4 ml $CaCl_2$.

III. To mix competent cells with plasmid

13. Add to each sample (transformation tube) standing in the ice-bath 0.2 ml of $CaCl_2$ treated cell suspension.
14. Leave on ice (in the refrigerator shelf) for 25 minutes.
15. Transfer to a 45°C water-bath (heat-shock treatment) for 2 minutes.
16. Remove from the water-bath and keep at room temperature (22°C) for 10 minutes.

IV. To add LB broth

17. Add 1.0 ml of LB broth (warmed to 37°C) to each tube and incubate the latter at 37°C for 30 minutes.
18. Add to each tube 1 ml ampicillin (50 μg/ml LB) and incubate at 37°C for 2.5 to 3 hours or overnight.
19. Dilute the contents of each tube with LB broth (1:2).
20. And to each tube 2 ml of LB.
21. Add 10 g/ml of tetracycline and incubate with shaking at 37°C for about 45 minutes.
22. Add to the above cycloserine 150 g/ml. Incubate with shaking at 37°C for 2–4 hours.
23. Spin the tubes at 8,000 rpm (Sorvall SS34 rotor) for 5 minutes.
24. Resuspend in 5 ml of NaCl soln. (0.9%).
25. Spin and discard the supernatant.
26. Resuspend the pellet in 1.5 ml AMP-LB broth and incubate at 37°C from 30 minutes to overnight.
27. Add to each tube LB-agar (7%) 2.5 ml (kept warm at 37°C) and swirl to mix.
28. Pour from each tube onto AMP-LB and TET-LB plates. Incubate the plates at 37°C overnight.

V. To score the number of transformants (colonies) per plate

29. Score the surviving cells. These are AMP^S and Tet^S transformed cells.

N.B. You may streak some plates, with one transformant colony per plate, and incubate them at 37°C to enlarge the colonies. pDNA may be isolated from them, cleaved with restriction enzymes (the ones used to make the hybrid DNAs) and the sizes of fragments (plasmid and insert DNA) checked in a minigel.

Exercise 36: Transformation of E. coli with Recombinant Cosmids

The *E. coli* strain HB101 (pJC74) carries a cosmid with a gene for ampicillin resistance. The cosmid pJC74 will be isolated, spliced to *E. coli* chromosomal DNA fragment and packaged in a packaging mix. The latter will be prepared from two strains of *E. coli* (λ); one lysogen harbours a phage lambda with an amber mutation in the D shell protein (*Dam*), and the other a phage λ with *Earn.* The recombinant phages will be used to transduce the hybrid cosmids into a recipient strain of *E. coli* (HB101). Phages from the plaques on the *E. coli* lawn will be amplified, and DNA extracted from them. The extracted DNA from these phages and the unhybridized cosmid DNA may be fractionated on a horizontal agarose gel. The hybrid DNA will be larger in size than the original cosmid. Figure 7.21 gives the flow diagram for transformation of *E. coli* with recombinant cosmids. The procedure for the same is described here:

Summary of steps

I. Isolate cosmid from *E. coli.*
II. Purify the cosmid DNA.
III. Construct recombinant cosmids.
IV. Prepare the packaging mix.
V. Transfer the DNA into *E. coli* by transduction.
VI. Check clones for insert DNA.

Have ready

1. Strains of *E. coli,* K12
   (i) Wild-type, r_k^+ m_k^+
   (ii) HB101F$^-$, $r_k^-m_k^-$
   (iii) HB101 (pJC74), in which the cosmid pJC74 has the AMP^R gene.
   (iv) BHB2690, with Dam_{15},CI^{ts} r_k^+ m_k^+
   (v) BHB2680, with Eam_4, CI^{ts}, r_k^+ m_k^+

I. To isolate the cosmid

2. LB-Ampicillin agar plates:
   (i) LB 1000 ml
   (ii) Agar 15 g
   (iii) Ampicillin 50 μg/ml
   Autoclave.
3. LB:
   Bactotryptone 10 g
   Yeast Extract 5 g
   NaCl 10 g
   Dist. water 1000 ml
   Autoclave.
4. LB-Ampicillin broth:
   LB 1000 ml
   Ampicillin 50 g/ml
5. Chloramphenicol (see Exercise 8)

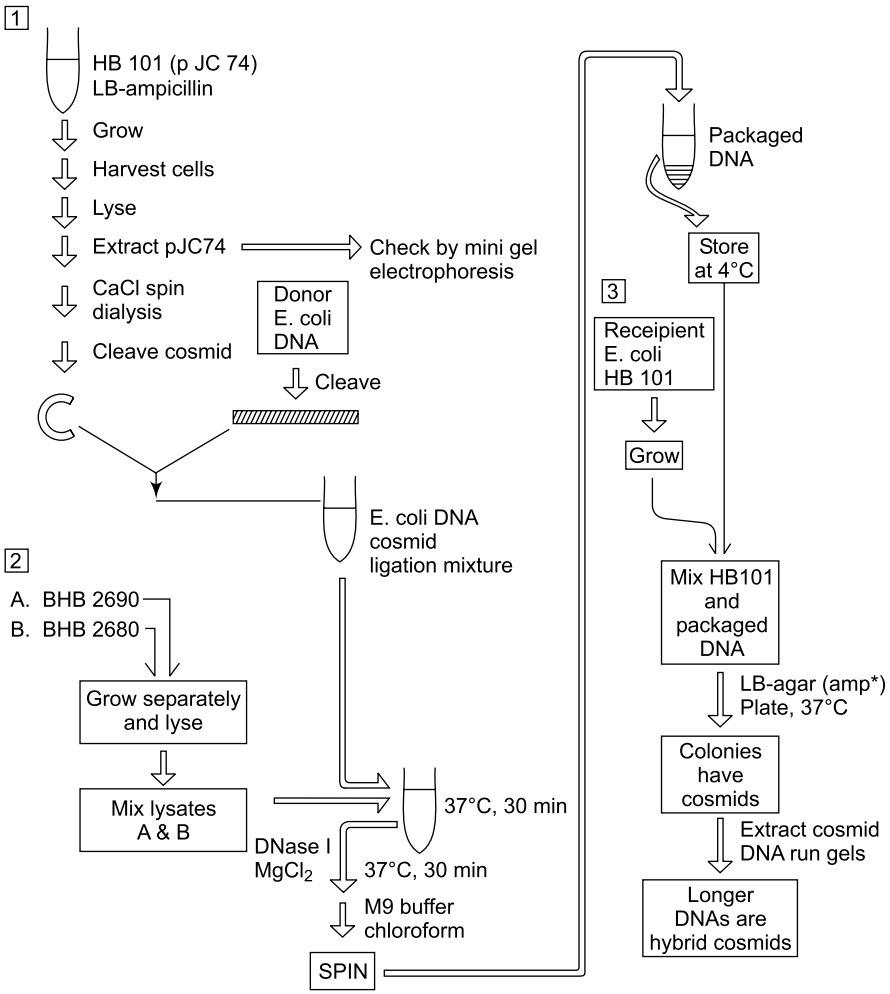


Fig. 7.21 *Flow diagram for transformation of E. coli with recombinant cosmids. The cosmid pJC74 is isolated from E. coli strain HB101 (pJC74), purified and cleaved with a restriction enzyme. The cleaved cosmids are mixed and ligated with E. coli DNA fragments cut by the same enzyme. The packaging mix is prepared from E. coli strains BHB2690 and BHB2688. The hybrid cosmids are added to the mixture of lysates from these cells. The recombinant phage particles that result from the reaction are used to transform E. coli HB101. Cosmids extracted from them are run on agarose gel together with a control (intact cosmid). Cosmids larger than the control are hybrid.*

6. Sucrose-EDTA-Tris HCl (SET):
   Tris HCl 0.1 M pH 8.0
   Sucrose 15 %
   EDTA 10 mM
7. NaCl 5 M soln.
8. Lysozyme (to be added fresh)
9. SDS 10%

II. To purify the cosmid

10. CsCl
11. Ethidium bromide
12. Isopropanol saturated with CsCl

III. To construct recombinant DNA

13. *E. coli* wild-type chromosomal DNA
14. pBR322
15. *Eco* RI
16. *Eco* RI 10 X buffer

(15–16) See Exercise 27

17. T4 ligase
18. Ligation solutions

(17–18) See Exercise 28

IV. To prepare the packaging mix

19. LB-Agar plates (Agar 1.5%)
20. LB broth
21. Phage mix buffer:

NaN_3	10 mM
Tris HCl	40 mM pH 8.0
Putrescine	10 mM
Spermidine	10 mM
ATP	10 mM
β-mercaptoethanol	0.1%
DMSO	7.0%

22. $MgCl_2$ 100 mM
23. DNase I
24. Mg buffer:

Potassium dihydrogen phosphate	3.0 g
Disodium hydrogen phosphate	7.0 g
NaCl	0.5 g
Ammonium chloride	1.0 g
*Magnesium sulphate	1.0 mM
*$CaCl_2$	0.1 mM
Dist. water	1000 ml

Autoclave all except*
Autoclave *–marked salts separately
Mix before use.

25. Chloroform

V. For transduction

26. LB-Maltose broth:

LB broth	1000 ml
Maltose	4 g

27. Soft agar-LB:
    LB 1000 ml
    Agar 7 g
28. LB-Ampicillin agar plates

Procedure

I. Cosmid extraction

1. Make an overnight culture (37°C) of *E. coli* BH101 (pJC74) in 10 ml LB-AMP broth (see Exercise 33).
2. Mix 10 ml overnight culture with LB-AMP broth 1000 ml.
3. Incubate in a shaker water-bath at 37°C for 2–3 hours (OD should be A_{600} = 0.5).
4. Add to this culture chloramphenicol 170 mg/ml. Incubate in a shaker water-bath at 37°C overnight.
5. Centrifuge for 10 minutes at 8,000 rpm in a Sorvall. The pellet consists of cells.
6. Add to the pellet 10 ml SET and resuspend the cells using a Pasteur pipette.
7. Add to the suspension lysozyme (100 mg/ml). Keep on ice for 60 minutes
8. Add 5 M NaCl, such that the final concentration of NaCl is 1 M.
9. Add SDS (10%) to a final concentration of 1%. Store overnight in the cold room (4°C) or on a refrigerator shelf. *Steps 6–9 lyse the cells.*
10. Centrifuge the lysate for 30 minutes at 15,000 rpm in a polyallomer tube. The supernatant contains the smaller cosmid DNA.
11. Pipette the supernatant into a graduated centrifuge tube. Note the volume.

II. Purification of cosmid

12. Add CsCl (0.98 g/ml of supernatant).
    Add ethidium bromide (100 mg/ml) adjusted refractive index to 1.385.
13. Centrifuge in the Beckmann ultracentrifuge for 48 hours at 40,000 rpm.
14. Remove the tube; fix it vertically to a stand with a clamp in a dark room. Turn on the UV transilluminator on the tube. The reddish bands represent DNA. There may be more than one band due to small fragments of chromosomal DNA.
15. Remove the lowermost band (which contains the cosmids) into a glass test tube, by piercing the centrifuged tube at the lower margin of the band with a No. 18 syringe needle. The material of this band, eluted through the needle, will be collected in a test-tube.
16. Add to the DNA the isopropanol saturated with CsCl. The ethiduim bromide is removed by this step.
17. Transfer the above carefully into a washed dialysis tube. Knot the open end of the tube and hang the bag in a large beaker (1 litre) containing TE. Change the TE every 8 hours. This removes the CsCl from the cosmid preparation.
18. Transfer the dialysed DNA into a graduated glass centrifuge tube.
19. Add cold ethanol (twice the volume of DNA solution). Keep at –20°C for 2 hours. The cosmid DNA will precipitate out.
20. Pellet the precipitate by spinning.
21. Decant the supernatant. Resuspend the pellet in TE, check OD, and adjust the concentration of DNA to 0.5 mg/ml of TE.

III. Construction of recombinant cosmids

22. Take in an Eppendorf tube, using Gilson pipettes:

Cosmid DNA	1.0 ml
Eco RI 10 X buffer	1.5 µl
Dist. water	2.5 µl
Eco RI	1.0 µl (5 units)

Gently mix by tapping with finger tips.

23. Incubate at 37°C for 60 minutes. The cosmids should be linearized by this time. You may run a minigel to check if digestion is complete. Uncleaved cosmids will form a separate band (faster moving than the linearized cosmids).
24. Similarly, fragment the *E. coli* DNA in the following manner:
    Take in an Eppendorf tube.

E. coli DNA	25 ml (0.5 mg/ml)
Eco RI 10 X buffer	5 ml
Dist. water	1 ml
Eco RI	1 ml (5 units)

Mix gently, incubate at 37°C for 15 minutes. Digestion should be checked on a minigel.

N.B. Steps 22 and 23 may be taken simultaneously, and minigel checks run on the same agarose gel. Join the *E. coli* DNA fragments and cosmids in the following manner:

25. Take in an Eppendorf tube:

E. coli DNA	5 µl [(from Step (24)]
Cosmid DNA	10 µl
LS 5 X	5 µl
LA 5 X	5 µl
T4 ligase	1 µl

Keep at 12°C overnight.

26. Check ligation by running an aliquot of the above ligation mixture in a minigel; also run a linearized cosmid as a standard. There should be several bands of DNA larger than the one with the linearized cosmid band.

IV. Preparation of packaging mix

27. Grow 10 ml overnight cultures of the mutants BHB2690(A) and BHB-2688(B) in LB broth (at 32°C).
28. Add overnight cultures to 250 ml LB broths. Adjust ratio of culture to broth until OD at A_{600} = 0.1
29. Incubate in a shaker water-bath at 32°C until OD at A_{600} = 0.3
30. Transfer the cultures to a slowly shaking water-bath at 42°C. Keep shaking for 15 minutes.
31. Transfer the cultures to a regular shaker water-bath at 37°C for 2.5–3 hours. The liquid should be quite turbid due to cell lysis. Take a few ml of the culture in a test-tube and add a drop of chloroform. The culture becomes clear if the phage has been induced completely.
32. Mix in a large test-tube equal volumes of cultures A and B.
33. Harvest the cells by centrifugation. Discard the supernatant.

34. Add to the pellet 1.0 ml of the phage mix buffer, and resuspend the pellet.

This is the packaging mix.

35. Distribute 20 ml of the above packaging mix into each of several Eppendorf tubes (1.5 ml). Freeze in liquid nitrogen and store at –70°C.

To package rDNA in phage shells

36. Transfer the packaging mix tubes to a crushed ice-bath (0°C). This thaws the mix.
37. To each tube add:
    Ligated DNA 5 ml (0.01–0.2 mg)
    Mix by tapping the tube.
38. Keep at 37°C for 30 minutes.
39. Add to the mixture and mix:

$MgCl_2$	2 ml
DNase I	2 ml (100 mg/ml)

40. Incubate for 30 minutes at 37°C.
41. Add to the mixture and mix:

M9 buffer	0.5 ml
Chloroform	0.2 ml (1 drop).

42. Spin at 8,000 rpm for 2–3 minutes. The phage particles will be in the supernatant.
43. Store the supernatant at 4°C.

V. Transfer of rDNA into E. coli by transduction

44. Innoculate 1 litre of LB-Maltose broth with a 10 ml overnight culture of the recipient strain of *E. coli* HB101.
45. Incubate in a shaker water-bath at 37°C until OD at A_{600} = 2.0
46. In an Eppendorf tube, take 1.0 ml of HB101 culture and 0.4 ml of phage suspension. Mix by tapping the tube.
47. Keep at 30°C for 10 minutes. The phage particles become adsorbed on the *E. coli* cells.
48. Spin the mixture in the Eppendorf centrifuge and discard the supernatant.
49. Resuspend the pellet in 1.0 ml LB-AMP broth. Keep at 30°C for 2 hours.
50. Add to the above 1.0 ml LB soft agar (kept at 45°C).
51. Immediately pour on a LB-AMP agar plate and spread by swirling the covered plate.
52. Incubate (in the incubator) at 37°C for 24 hours.
53. Observe and count the colonies that appear. These colonies are ampicillin resistant, which indicates that they contain cosmid DNA. The latter could be either the hybrid or the non-hybrid cosmid.
54. To discover the clones that contain insert *E. coli* DNA, add one colony each to at least 15 tubes containing 5 ml LB-AMP broth, incubate these at 37°C overnight, and proceed as given below.

VI. To check clones for insert DNA Isolate low molecular weight DNA from each overnight culture or from a more amplified culture, and run minigels together with a linearized non-hybrid cosmid as a standard marker.

Only inserts that are within a particular size range will allow the hybrid cosmid to be packaged into viable particles. This means that very small inserts as well as very large fragments will be missed by this technique.

The bands of DNA that are due to hybrid cosmids (larger than the hybrid cosmid) contain fragments of *E. coli* DNA.

If required, these clones may be subjected to further selection pressures to specifically identify a particular region of the *E. coli* DNA.

Exercise 37: Transformation of Saccharomyces cerevisiae

Have ready

1. YEPD medium:
   Yeast Extract 10 g
   Peptone 20 g
   Dextrose 20 g
   Dist. water 1000 ml
2. SET medium:
   Sorbitol 1 M
   Na_2 EDTA 25 mM pH 8.0
   DTT 6.7 mg/ml
3. SCE:
   Sorbitol 1 M
   Na-citrate 0.1 M
   Na_2 EDTA 10 mM pH 5.8 (use HCl to adjust pH)
4. Zymolase 60,000. Prepare fresh 1 mg/ml in TEN buffer.
5. PEG 4000
6. SOS:
   Sorbitol 1 M
   YEPD 33% (v/v)
   $CaCl_2$ 6.7 mM
   Leucine 14 g/ml
7. Sorbitol 1 M
8. Calcium-Sorbitol soln.:
   Tris 10 mM pH 7.5
   Sorbitol 1 M
   $CaCl_2$ 10 mM
9. Leucine marker: (2.5% top agar with sorbitol (LSAI))
   Sorbitol 1 M
   Difco Yeast Nitrogen base 0.67% (without amino acids)
   Glucose 2.0 %

Amino acids:	
Adenine Methionine	10 g/ml each
Arginine Tryptophan Histidine	20 g/ml each
Lysine Uracil	40 g/ml each
Tyrosine Threonine	50 g/ml each
Isoleucine Phenylalanine	60 g/ml each
Agar	2.5%

10. Leucine marker: 2% agar plate with sorbitol (LSAII). Same as above, except for 2% agar in place of the 2.5%.
11. Leucine marker: agar (2%) (LA). Same as serial 9, but without sorbitol.
12. Yeast strain to be transformed.
13. pDNA (purified by CsCl gradient centrifugation and stored at 4°C).

Procedure

1. Start an overnight culture with 5 ml of YEPD. Incubate at 30°C (preferably in a rotating incubater).
2. Next morning, dilute the overnight culture by adding it to 100 ml of YEPD.
3. Incubate the diluted culture at 30°C by shaking for about 5 hours (about 2×10^7 cells/ml or OD at A_{600} = 1.5–3)
4. Harvest the cells by centrifuging at about 5,000 g for 5 minutes. Discard the supernatant.
5. Resuspend in DW. Spin, and discard the supernatant. Repeat this step to wash the cells free of medium,
6. Resuspend the cells in 10 ml SET. Incubate at 30°C for 10 minutes.
7. Spin; discard the supernatant.
8. Resuspend the pellet in 1 M sorbitol. Spin. Discard the supernatant.
9. Resuspend the pellet in 10 ml of SCE.
10. Add 0.1 ml of Zymolase (0.5 mg/ml).
11. Incubate at 30°C for 20 minutes or more, shaking occasionally. Monitor spheroplasting, by observing under the microscope. At least 50% cells should have 'loose' cell walls.
12. Stop the action of the enzyme by spinning the spheroplasts in 1–2 ml of the calcium-sorbitol solution in Tris buffer. Resuspend in 0.5 ml of the same solution.
13. Distribute the spheroplasts (0.1 ml each) into Eppendorf tubes (1.5 ml).
14. Add to each tube ~1 µg of pDNA. Keep a control tube without any pDNA. Leave at room temperature (25°C) for about 20 minutes.

15. Add to each tube PEG (10 ml for each 0.1 ml of cell suspension).
16. Spin and resuspend cells in SOS. Keep at 30°C for 30 minutes.
17. Add to each tube 6 ml of LSAI (top agar) (at 45°C); swirl the tube to mix and then pour out on a LSAII plate. Let die top agar solidify.
18. Incubate the LSAII plates at 30°C for 2–3 days or until colonies appear. The colonies are transformants that can grow on the media lacking leucine (LSAI & LSAII).

Exercise 38: Plant Cell Transformation using Agrobacterium tumefaciens

Have ready

1. Minimal Medium A (MA) — 1 litre stock.

$(NH_4)SO_4$	1.0 g
K_2HPO_4	10.5 g
KH_2PO_4	4.5 g
Na-citrate · $2H_2O$	0.5 g
$MgSO_4 \cdot 7H_2O$	0.2 g
Glucose	2.0 g
Dist. water to make	1000 ml

2. MA Medium with the following added per litre:
   (MA-SIX)

Sucrose	0.2 M
Inositol	100 mg
Xylose	250 mg
pH	5.6 (Adjust with KOH)

   Autoclave.
3. Antibiotic to kill agrobacteria (e.g., the Hoechst product 'Cefotaxime' or Kanamycin, 75–100 mg/ml)
4. Shoot-Inducing MS medium (MS-SI)
   MS powder (for 1 litre)

Nicotinic acid	0.5 ng
Glycine	2.0 mg
Pyridoxine-HCl	0.5 mg
Myo-inositol	100.0 mg
Thiamine	0.4 mg
Sucrose	10.0 mg
Agar	8.0 mg
pH	5.7 (Adjust with KOH).

5. *Agrobacterium tumefaciens* stock.
6. Petunia leaves (young)
7. 1 pair of sharp scissors (sterile)
8. Parafilm
9. $HgCl_2$ (Mercuric chloride) soln. (0.1%)

Procedure

1. Grow an overnight culture of *A. tumefaciens* at 28°C (4 ml MA with one loopful of bacteria scraped from stock culture).
2. Sterilize washed petunia leaves in 0.1% mercuric chloride.
3. Pour about 15–20 ml of MA-SIX in a sterile 90 mm petri dish.
4. Float a leaf with its upper side touching the solution in the dish.
5. Carefully remove the edges of the leaf with a pair of scissors, and cut the leaf blade into square pieces (10 mm × 10 mm).
6. Pour MA-SIX (4 ml) into each of 5 very small petri dishes (1″ diameter).
7. Place a few leaf squares (lower side facing up) in each dish.
8. Pipette (using lambda-pipettes with sterile tips) about 50 µl of *Agrobacterium* overnight culture into the medium in each dish.
9. Seal edges of petri dishes with parafilm.
10. Incubate at 25°C for 2–4 days.
11. Add about 18.7 g of antibiotic, such as Claforan (cefotaxime) of Hoechst, to 125 ml of MA-SIX. Filter sterlize. (MA-SIX-A).
12. Pick up each leaf square with a pair of bent forceps and gently wash its upper surface by gently swishing it in MA-SIX-A kept in a petri dish.
13. Pick up each leaf square with a pair of bent forceps and pull it along the surface of a petri dish of MA-SIX-A. This washes the wet surface (dorsal surface of leaf) and kills any bacteria adhering to it.
14. Place the leaf pieces on the surface of MS-SI medium in small petri dishes. Seal the dishes with parafilm.
15. Incubate at 25°C on MS plates containing 0.1 mg NAA and 1 mg BA per litre of medium. (Tobacco cells fare well with 2 mg NAA and 0.5 mg BA.) Calluses appear after about a week. Let them grow for about 2–3 weeks. Shoots should appear by this time; roots appear simultaneously or later.
16. Gently remove the plantlets with a pair of fine forceps and transfer them to fresh dishes containing the same MS-SI (plus hormones) medium.
17. The plantlets may be placed individually in tissue-culture test-tubes containing MS-SI (plus hormones) medium (slants).
18. When the plantlets are stronger they are gradually transferred to progressively larger containers, until they are acclimatized enough to grow outdoors.

The exercises in this chapter provide a few basic procedures for introducing foreign DNA into host cells. Needless to say, the basic original protocols have been improved and made very sophisticated in a variety of ways. Perhaps the worker using these exercises will develop modifications to suit very specific requirements. The basic exercises are enough, though, to introduce a novice to the laboratory operations routinely utilized in most DNA cloning procedures. For more specific information, the reader may consult the list of references that includes excellent books on gene cloning as well as several articles which provide detailed protocols of various benchmark DNA cloning ventures.

REVIEW QUESTIONS

1. Giving examples, briefly recall how DNA can be isolated from prokaryotic and eukaryotic cells.
2. What precautions and care must be taken in extraction of RNA molecules?
3. How is cellular protein isolation carried out?
4. Assess the utility of gel electrophoresis in separation of molecules.
5. How is isolated DNA purified?
6. Enumerate the techniques of DNA manipulation.
7. Recall various blotting techniques.
8. Recapitulate steps in DNA cloning.
9. How is transformation of *E. coli* effected with a plasmid-borne DNA and with a DNA carried by phage lambda?
10. Describe how r-DNA molecules are made.
11. How is mRNA, which is to be copied in cDNA, identified?

8 Eukaryotic Cell Culture Systems

Introduction

Bacteria and other prokaryotes are haploid and usually single-celled species with a very elementary pattern of development. Genetic manipulation of prokaryotes is, therefore, fairly easy, as well as gratifying since the effects of the interventions become apparent as morphological or biochemical variants.

The single-celled eukaryotes, such as the green alga, *Chlamydomonas reinhardii* and the yeast (fungus) *Saccharomyces cerevisiae,* are less elaborate than the multicellular eukaryotes and genetic tinkering can be achieved in them with some degree of success and satisfaction.

The trouble comes when one deals with the multicellular creatures, be they animals or plants. First of all, such species possess complex schedules of development, which rest on strict regulation of expression of genes in terms of kind of gene expressed in time, place and quantity of gene product. This is the basis for the following two difficulties of gene-cloning in eukaryotes:

1. Since many genes are developmentally regulated and are, therefore, tissue-specific, introduction of a gene into a random cell in an organism does not ensure its expression in the cloning cell.
2. In order to incorporate a gene into the organism as part of the active genome, the introduction must occur as early as possible in the time-table of development. Even if this was achieved, in the case of an animal (say a mammal), we do not know the environmental, cellular and other requirements for the successive developmental phases to culture a viable, fertile adult from the manipulated early embryo. At present, the difficulty of culturing animals to the adult stage in strictly *in vitro* conditions is side-tracked by implanting the treated embryo in a surrogate mother. However, these obstacles have not prevented attempts to engineer animal cells with defined DNA, if only to observe and learn about the biology of the phenomenon under study. Techniques are being continuously upgraded for culturing single animal cells in nutrient media. The medium has to be supplemented with serum (calf, horse etc.) for the cells to be able to survive and differentiate. The serum presumably contains the mixture of factors essential for the growth of animal cells *in vivo.* More about this will be described later.

We find a much more satisfying situation in the case of plant tissue culture. Almost any healthy, young plant cell can be regenerated into a fertile adult plant. The requirements for development of what are obviously totipotent cells are met with a variety of defined synthetic media.

In any event, attempts to genetic engineer both animals and plants must be backed with adequate techniques for culturing cells, tissues and developing organisms. Due to certain basic differences between the molecular biology of plants and animals, methods of culturing tissues of these two kingdoms are also divergent. A brief introduction to the problems, methods and algorithms and efficiencies of plant and animal tissue culture will, therefore, be presented below.

8.1 ❑ PLANT TISSUE CULTURE

Despite the fact that plant molecular biology has been delved into seriously only in the very recent past, application of genetic engineering resulting in altered whole plants is likely to outstrip earlier similar efforts by animal scientists. The prediction is based on the two special features of plants mentioned earlier. These are the totipotency of plant cells, and the possession of two autonomous organelle genomes that may be engineered independently.

Plant tissue culture aims to induce plant organs, or portions of them, to proliferate cells, each one of which can be regenerated into whole plants. These plants would be virtually replicas of the parent plant. The usefulness of such replica-making or cloning in agriculture is obvious (Fig. 8.1).

When a portion of a plant tissue or organ is placed on a culture medium and kept under optimum growth conditions, the meristematic cells de-differentiate and proliferate to form a mass of cells known as a callus. Under the influence of proper phytohormones (auxins and cytokinins), the undifferentiated cells can be induced to undergo morphogenesis into roots, shoots or both. The individual plantlets can be nurtured to grow into mature plants.

The starting material for callus production may be a small section of any part of the plant—generically called the explant. Root, shoot, leaf, fruit tissue, petals, hypocotyls of germinating seeds and so on may be used to obtain explants. These of course will result in calluses of diploid cells (see Exercise 39).

If haploid calluses are desired, the starting material should be ovules, immature pollen or whole young anthers (See Exercise 40). The calluses formed from these cells may be unorganized ones or may form miniature embryo-like structures that regenerate directly into complete plantlets. These latter structures are known as embryoids.

Instead of using multicellular tissue as the starting point, it may be dissociated into single cells, which may then be cultured in suspension in liquid medium or immobilized on the surface of an agar-solidified medium (see Exercises 41–44).

To regenerate cells that are single or are part of a callus, appropriate auxin and cytokinin have to be added to the growth medium. It may be necessary to try combinations of different ratios of sets of these phytohormones to determine the optimum mixture for the particular plant material in use.

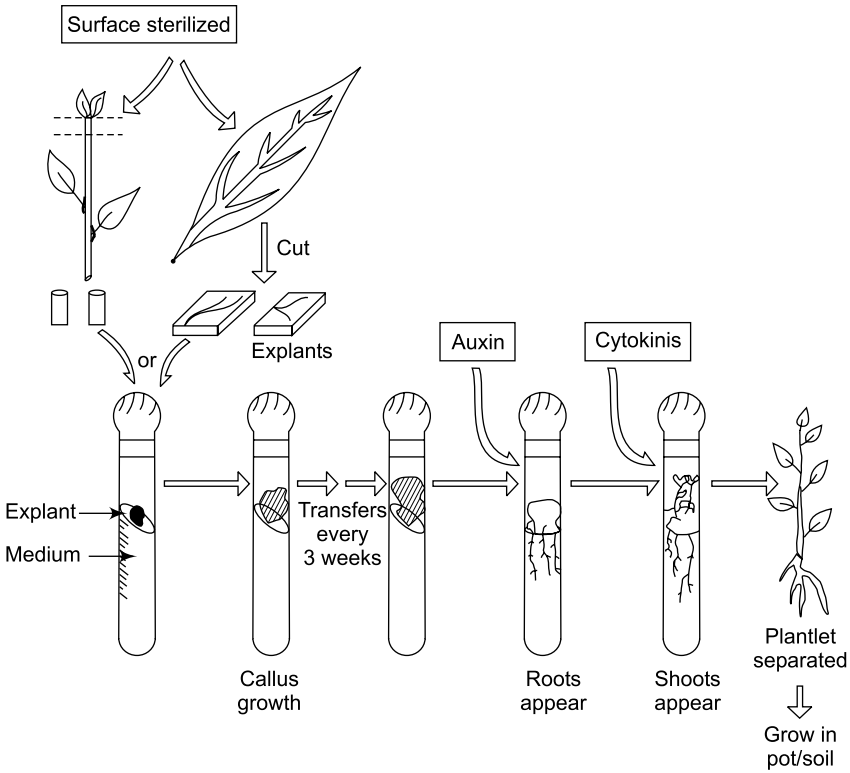


Fig. 8.1 *Explant tissue culture. A piece of surface sterilized plant tissue from any organ (stem, leaf, root etc.) when placed on a solidified synthetic nutrient medium develops calluses (aggregations of newly divided cells) which on transfer to growth hormone-supplemented media regenerate (from each callus cell) into whole, viable plants. Auxins and cytokinins are root and shoot promoting hormones, respectively*

Culture conditions include: (i) a suitable medium, (ii) optimum light intensity, (iii) appropriate humidity, and (iv) a suitable temperature. The culture room must be as contamination free as possible and all operations are to be carried out under utmost aseptic conditions. A special cabinet, called the Laminar Flow Cabinet, is available that allows flow of air through a filter that keeps out all microorganisms including viruses. The cabinet is also fitted with a germicidal UV lamp, that is turned on for about half-an-hour preceding the use of the cabinet.

The tissues and cells are usually cultured in large rimless test-tubes plugged by cheesecloth covered nonabsorbent cotton. Petri dishes, conical flasks or other suitable sterilizable or disposable sterile containers may be used for the purpose.

Light is usually provided by fluorescent tubes placed at measured distances from the cultured material, so that the latter receives the optimum quantity of light (measured in 'lux'). The requirements of the duration of light exposure varies with different plant species.

The growth medium comes in several formulations, bearing the name of the initial reporter of the concoction. All of them contain a carbon source, major and minor elements and vitamins. The temperature for most plant tissue culture is around 26 ± 1°C.

Not all species of commercial or other importance have responded well to tissue culture and regeneration methods. Attempts are in progress to overcome this difficulty. To date, members of the tomato family (Solanaceae) have responded most successfully. Tobacco, *daturas,* eggplants, and potatoes belong to this family. The most important cereal crops, and the important food crops, belonging to the beans and lentils family (Leguminoseae) are proving to be rather recalcitrant to existing tissue culture protocols. A few useful protocols for culturing some species of cereals have been reported.

In some cases single cells or calluses have been cultured successfully, but it has not been possible to regenerate them. These failures stem mainly from the lacunae in our knowledge about the workings of plant systems at the molecular and cellular levels.

Calluses are actually wound tissues that develop at the injured regions of a plant organ. Calluses developed under laboratory conditions show significant alterations in the proliferated cells. Chromosome number anomalies, chromosome aberrations, as well as mutations, are featured in the callus cells and plants regenerated from them. While this is a disadvantage when replicas of a given strain are desired, this mutability of callus cells is used to advantage to isolate new varieties or clones of an existing one. The mutated cells maintain the mutation stably in progenies derived from the regenerated plants. The clones obtained from such somatic tissues are often referred to as somaclones. Chromosome counts and inspection should always be made of plants regenerated from calluses.

8.1.1 Calluses from Explants

Explants may be taken from almost any region of the plant that is young and healthy—rather that which is not too mature or senescent. Shoots roots, leaves, stem tips, soaked and germinating seeds, their endosperms, plumules, radicles, hypocotyls or epicotyls may be used as the starting material. Explants develop calluses at the cut surfaces (see Exercise 39).

Callus cultures may be utilized for generation of clones or may be dissociated into individual cells that are then cultured as single protoplasts or cells (see Exercises 41–44).

In vitro differentiation of callus cells is induced by plant growth hormones (phytohormones) called cytokinins and auxins. Cytokinins trigger the differentiation of shoots, while auxins promote that of roots. The cytokinins generally in use are:

1. Kinetin (6-furfurylamino purine)
2. BA (6-benzyladenine)
3. BAP (6-benzoadeno purine)
4. Zeatin, and
5. 2, i-P (N-isopentenylamino purine)

The commonly used auxins are:

1. IAA (Indole-3-acetic acid)
2. NAA (1-Naphthaleneacetic acid), and
3. 2, 4-Dichlorophenoxyacetic acid

Development of the shoot or root is a function of a particular ratio between the auxin and the cytokinin in the medium. Increasing the relative proportion of cytokinin(s) or of auxin(s) results in the formation of shoots and roots respectively. Equal amounts of both usually result in undifferentiated masses of callus cells. The types of phytohormones in each group and the proportion of each needed for successful regeneration of plants differ from species to species, and are, to date, determined by trial and error, and the best combination used as the standardized recipe.

The callus usually consists of the variety of cell types present in the original explant. Plant regeneration may occur directly from cells of the explant without passing through a callus stage of development. Cell masses that directly give rise to separate plants, and therefore, mimic embryos, are known as embryoids. It is more usual for a callus to differentiate into root or shoot, and then be induced to complete the regeneration process.

Cells in the calluses and the plants developed from them are normally not identical to the cells of the primary explants in terms of their genetic endowment. Callus cells and derivatives usually exhibit wide ranging anomalies in chromosome numbers. Also, calluses that have been subcultured serially several times, lose their ability to regenerate. Perhaps, besides the visually distinguishable chromosomal defects, there are also accumulated mutations in the repeatedly subcultured cells, that prevent normal morphogenesis. This genetic instability of callus-derived cells make them unsuitable for commercial cloning. However, this very handicap is exploited as a source of genetic variants of commercially important plant stocks.

8.1.2 Haploid Culture

Techniques are available for generating plants from single micro- or mega-spores (Fig. 8.2). The most common source for haploid cells is the anther, the male reproductive cell-forming region of a flowering plant. A whole anther placed on a culture-medium, fortified by (see Exercise 40) appropriate growth factors, develops embryoids from single pollen cells that occur in the interior of the organ. The embryoids emerge from the interior of the anther. The pollen, separated from the surrounding diploid parental tissue of the anther, may also be cultured *in vitro.*

Embryoids developed from intact anthers or from isolated pollen cells may be regenerated into plants, which of course are haploid. There is far less genetic abnormality of haploid culture cells and plants than found in explant callus-derived material. More embryoids are obtained from intact anthers than from isolated pollen cells.

The stage of development of the pollen cells is important for the successful yield of a large number of haploid cells per anther. In most cases, the micro-spore with the undivided haploid nucleus gives the best result. In some species better results have been obtained, when pollen cells with the nuclei having undergone one mitosis

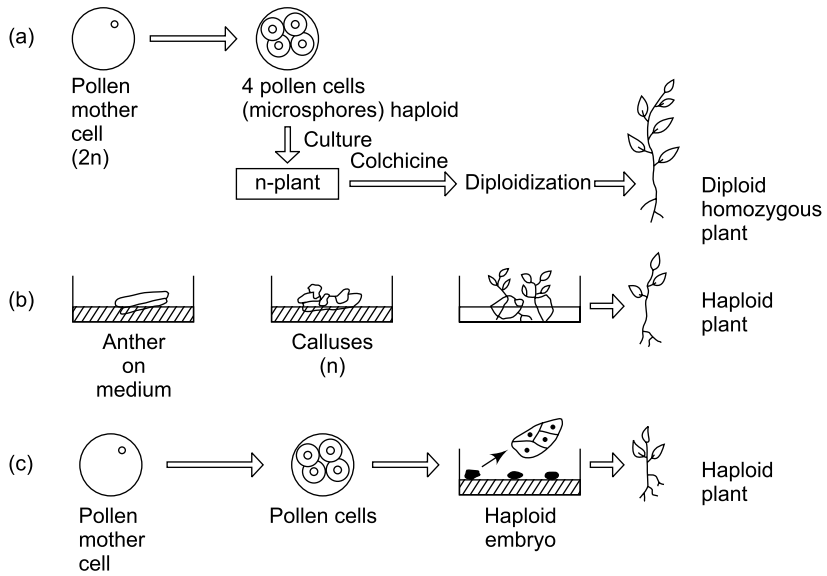


Fig. 8.2 *Haploid tissue culture. Haploid cells of plants (pollen and ovule) may be cultured and regenerated into haploid whole plants, (a) A haploid plant may be made diploid by treatment with the mitotic poison colchicine, (b) Anther culture produces calluses which may regenerate into plantlets. (c) When single pollen cells (n) obtained after meiosis of pollen mother cells (2n), are cultured, an embryo-like aggregate of cells, referred to as an embryoid, may result, which grows directly into a plantlet.*

were used. Pretreatment of buds (or the infloresence or stalk carrying the buds) with high or low temperatures have been found to be efficacious in some species: the pretreatment enhances the number of cells that are regenerated into plants. For some species of *Brassica* (mustard), low temperatures (4–5°C) and a period of darkness are known to enhance embryoid formation. In another species of *Brassica* the treatment consisted of the following consecutive steps: 45°C for one hour, several steps at progressively lower temperatures until room temperature (25°C) is reached. This protocol promoted increase in embryoid development.

There is no clue as yet, as to the conditions which alter the direction of normal development of pollen to a path that results in mimicking the seed embryo.

Pollen or haploid cultures from anthers would be logical systems for investigating the requirements of a switch from a gametophytic (pollen germination and formation of two generative nuclei) to a sporophytic mode of development, and other phenomena associated with the onset of the process of meiosis. The ability to induce meiotic divisions in non-spore-mother cells will provide a variety of ingenious uses both for research and application.

Haploid cultures also possess other advantages not enjoyed by cultures of diploid cells. As mentioned earlier, haploid cultures exhibit far less (if at all) chromosomal and genetic abnormalities, than their diploid explant derived counterparts. Secondly, a mutation induced or introduced into a haploid cell becomes apparent as the mutant phenotype, a feat not always possible in diploid cells where heterozygosity

may cloud the issue. Diplodization of haploid cultures further provides homozygous stocks for breeding purposes.

Perhaps one of the unique uses of haploid cultures is in the participation of their cells in the formation of somatic cell hybrids. Haploids of two parent stocks can be fused to form a *heterokaryon* (with the parent nuclei unfused), a *synkaryon* (with a fused nucleus) or a *cybrid* (where the genome of one parent has been eliminated from a hybrid somatic cell). More about this will be discussed in the next section.

It would be fair to confess that despite the potentialities of haploid cultures, only a very narrow spectrum of species has responded well enough to serve as the basis for commercial exploitation. Several cereals, including rice, have yielded calluses from haploid cells, but very few, if any, have been regenerated into whole plants.

8.1.3 Protoplast Culture

The term protoplast is used for plant cells from which the cellulose cell wall has been removed. The naked cell has the plasma membrane as its outermost boundary, a feature that is usual in animal cells.

Plant cells, unlike animal ones, are cemented together into tissues with a carbohydrate called pectin. To release naked protoplasts from plant tissues, therefore, both the cementing pectin and boundary cellulose wall have to be removed. This is currently achieved by digestion of pectin and cellulose with the enzymes pectinase and cellulase respectively. Pectinase preparations are also sold under the name of macerozyme (since this enzyme helps to macerate tissues). Initially, protoplasts were isolated by mechanical abrasion of cells that were first plasmolyzed. The yield of protoplasts is very much poorer by this mechanical method than by the enzymatic one. However, the former technique has one advantage over the latter: there is no danger of enzymes destablizing the exposed plasma membrane.

Protoplasts may be isolated from stem, root, or mesophyll tissues of the leaf, as well as from calluses grown *in vitro.*

The procedure consists of the following basic steps (Fig. 8.3):

1. The explant or leaf to be used is first washed with a detergent followed by several changes of clean water (see Exercises 41–44).
2. The washed tissue is surface sterilized using a very dilute solution of mercuric chloride ($HgCl_2$) or sodium or calcium hypochlorite, or in the commercial solution named Chlorox (hypochlorite).
3. The material is washed in several changes of sterile distilled water to remove traces of the surface sterilizer.
4. The material is then treated with a mixture of cellulase and macerozyme which attack the lipids, proteins, cellulose and hemicelluloses in the plant cell wall and the intercellular pectin.
5. The liberated naked cells are isolated from the digestion mixture by filtering the latter through a sieve with a pore size of around 45 μm. The released protoplasts are spherical in shape. The progress of protoplast formation is observed under a microscope at regular intervals from the time of addition of

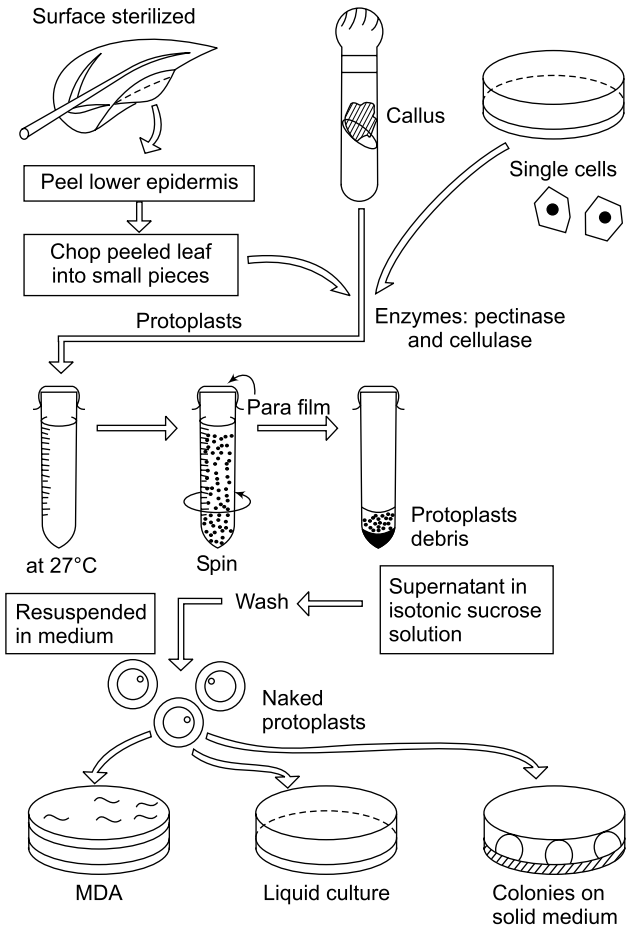


Fig. 8.3 *Single cell or protoplast culture. Intercellular pectin and cell wall cellulose are removed enzymatically to yield naked cells or protoplasts. These acquire new cell walls, usually within 24 hours. The protoplasts may be cultured by the MDA (microdrop array) procedure (in hanging drops) or in a liquid or solidified medium. Colonies from liquid culture may be transferred to solidified media for further development.*

the enzymes. The treatment is stopped when the majority of the cells are seen to be spherical.

6. The crude suspension of protoplasts obtained in the previous step is then filtered through a nylon material (45 µm mesh size) and centrifuged for 5 minutes in a clinical table top centrifuge (75 g). The protoplasts pellet out.
7. The protoplasts are now ready to be used.

As the naked cells are sensitive to the osmolality of the surrounding fluid, all the above operations are carried out with a fluid base in which the cell does not burst nor become plasmolyzed. Usually a mannitol solution is used that has the same osmolality as that of the interior of the cells.

The isolated protoplasts are then cultured by one or more of the following methods:

(i) Liquid culture The protoplasts are suspended in a liquid medium, usually in very small petri dishes or in 'hanging drops'. Very often, the development of a single cell proceeds better in a confined environment such as a drop of the medium. The cells are placed within a drop of medium that is placed on a hanging drop microscope slide. Development of the cells can be then visualized under a microscope.

(ii) MDA culture MDA or microdrop (or multidrop) array culture is a hanging drop culture where an array of drops is provided on the inner surface of the lid of a petri dish. This method is utilized particularly when different culture media are to be tested for their efficiencies of culturing the protoplasts under study.

(iii) Solid surface culture In this case the liquid medium is solidified with agar. Protoplasts plated on the surface of this soft agar divide and form colonies that are not dispersed in the medium. This method is, therefore, useful for identifying clones of specific protoplasts.

The cell wall that is removed when releasing the naked protoplasts is regenerated, usually within 24 hours. Any manipulations of the protoplasts, such as introduction of foreign genes, treatment with mutagens, or fusion of cells, have to be done before the new cell wall is synthesized. The cell with the regenerated wall then divides mitotically and forms a callus. With modification of the culture medium, it has been possible to develop embryoids directly from the calluses of certain species. The embryoids can be regenerated into whole plants.

Single plant cells, particularly naked protoplasts, have several uses. They can internalize, by endocytosis, a variety of substances as well as organized structures such as synthetic beads, viruses, bacteria, vectors, and intact cell organelles. The single cells are used for transient gene expression studies or as minifactories for the synthesis of the product of an introduced gene. On the other hand, if it is required to make the foreign gene a stable member of the plant genome, the single cells are regenerated into plants.

One of the obstacles to a rapid development of successful genetic engineering in plants is the dirth of mutations that may be utilized as genetic markers. Single plant cells are ideal candidates for the induction of such mutations. It has to be remembered that plant somatic cells are diploid, a condition that is unlikely to advertise the presence of a single dose of a newly acquired recessive mutant allele. Haploid single cells would be better candidates for acquiring cell mutants. Strains of tobacco have been developed in this manner that are resistant to streptomycin (the SR strains).

8.1.4 Induction of *Agrobacteria* caused Calluses

Ti-induced tumours can be either on whole plants or on excised plant organs. In the case of the former, regions on the stem surface may be wounded by sandpapering and by smearing a loopful of *A. tumefaciens* or *A. rhizogenes* overnight culture over the abrasion. Calluses appear in 3 to 4 weeks or more.

Calluses may be induced on root (such as carrot) or stem (such as potato) pieces by smearing the bacteria on small discs of surface sterilized explants. Leaves may also be used. Either large sections of surface-sterilized leaves should be slashed at several places with a blade and placed on the medium and smeared with the bacterial innoculum, or discs should be punched out of the leaf surfaces and treated as above. Calluses will be induced at the damaged regions—along the edges of the leaf pieces and along the slashed leaf surfaces.

Since *Agrobacteria* caused calluses are neoplastic and do not require exogenously supplied phytohormones, the medium used is a simple agar one. When the calluses are large enough they can be transferred to fresh hormone-free media.

The same procedure is followed if the infecting bacterium is *A. rhizogenes.* In this case, instead of tumourous growths, a profusion of fine hairy roots will emerge from the cut surfaces of the explants.

As an exercise you may transfer a well-grown tumour (say, of the size of a pea or smaller) to a fresh hormone-less medium every three weeks or so. If you find tumours in which roots or shoots have differentiated, you may isolate them as possible mutants of the tumour morphology loci.

8.1.5 Somatic Cell Hybridization

The usefulness of protoplast culture of plants is realized when it is needed to combine the genomes of two parent stocks, especially when they belong to different species.

Somatic cell hybrids are made by mixing freshly isolated protoplasts of the contributing parents in the presence of PEG (see Exercise 45). At first the two nuclei remain distinct in the combined cytoplasm. The cell at this stage is called a heterokaryon. Later, the nuclei fuse and the cell becomes a synkaryon. Cybrids can be made by destroying the nucleus of the parent cell, only the cytoplasmic qualities of which are required.

Usually a new cell wall appears within 24 hours, and cell division may be observed within another day. Several attempts are in progress to discover the culture conditions that will allow successful cell hybrids to be made and regenerated into plants. This has been possible, to date, in only a very few plant species.

For purposes of demonstration, somatic cell hybrids may be made from protoplasts isolated from carrot root and tomato leaves. The presence of orange carotenoids and green chloroplasts in the hybrid cells will easily distinguish the latter cells from the non-hybridized ones.

8.1.6 Culture Media

A general defined culture medium that induces the formation of calluses from both explants and single cells, as well as maintains them, consists of the following minimum ingredients:

1. Major Inorganic Elements
2. Minor Inorganic Elements
3. Organic Components
4. Iron Compounds

5. Trace Elements
6. Carbon Source
7. Growth Regulators
8. Undefined Substances

Solidified media contain, in addition, tissue-culture quality agar. Special culture techniques include other ingredients found suitable for specific objectives.

Several plant tissue culture media have been prescribed from time to time, one of the earliest being that of P R White. The more popular of these media include the ones named MS (Murashige and Skoog) SH (Schenk and Hildebrandt) and B5 (Gamborg). The MS and SH have higher concentrations of salts than the earlier White's medium and are the more effective of the commonly used formulations. Modifications of these media have been found to be suitable for different species and are, therefore, patronized in different laboratories. In some cases, the addition of an extra component, such as coconut milk, has been found to be especially efficacious. One of the features of the much-used media (SH and MS) is the presence of an agent that chelates the iron salt and releases it gradually without disturbing the pH of the medium during the long period of culture.

For making a solid medium, high quality (bacteriological grade) agar is added to the liquid medium, the pH of which is adjusted to around 5.5–5.9 and the final mixture autoclaved in the culture vessel to be utilized. This may be a conical flask, a jam bottle, a petri dish or a tissue culture test-tube stoppered with a cotton wool plug, screw cap or lid, depending on the nature of the container.

It is customary to start the culture of a sample with an MS, SH or B5 medium. Controlled ratios of a cytokinin and an auxin may be added to the medium and samples grown in each combination. The one providing the best growth is then selected for further modification, if necessary.

Monocotyledonous species, especially those of cereals, have been difficult to culture in unmodified standard media. The addition of higher amounts of growth regulators (e.g., 9.0–45.2 μm of 2, 4-D for rice, and 22.6–67.8 μm for sorghum) have given good results. However, it has not been possible to date to differentiate these cereal calluses into shoots and roots.

The formulae for the standard MS, SH and B5 media are given in the Appendix. Several firms are now supplying these media in a powdered form. If expense is not of any concern, these media may be used for most routine cultures. If the medium is to be prepared in the laboratory, the following steps have to be undertaken:

1. Prepare stock solutions of the first five groups of ingredients mentioned earlier.
2. Take a 1 litre flask (with a magnetic stirrer) and add to it the volume of stock solution of each ingredient as given in the table of constituents (see Appendix).
3. Add the carbon source (sucrose) and the solid myo-inositol
4. Add distilled water until the volume is 950 ml.
5. Stir the ingredients in the water with the magnetic stirrer.
6. Check the pH; adjust it to 5.8–5.9 using 0.5 M NaOH.
7. Make up the volume accurately to 1000 ml (using a volumetric flask) with distilled water.
8. Pour medium into flasks (≃ 75 ml per 250 ml Erlenmeyer flask) or tissue culture tubes (≃ 25 ml per 10 inch tube).

9. Plug the containers with non-absorbent cotton covered tightly with cheese-cloth.
10. Cover the plugs individually or in groups with aluminium foil.
11. Autoclave for 15 minutes at 120°C (15 psi pressure).
12. Place the tubes in a slanting position, while cooling, in order to form a slanted surface of the solidified medium.

When a thermolabile growth regulator has to be added to the medium, it is filter-sterilized and then added to the autoclaved medium when it has cooled down to 45–40°C (in the Laminar Flow Cabinet). The growth regulator is filter-sterilized by straining the stock solution to be used through a 0.45 μm pore size filter in a Millipore filter assembly.

It is a good practice to rest the media contained in tubes or flasks for 1 to 2 days at 25°C to check for contamination.

In the case of single cell cultures, the osmolarity of the culture medium is of critical value. It should possess the same value as that within the cell, otherwise the latter will burst or be plasmolyzed. Usually an osmoticum such as mannitol is utilized for this purpose. The proper osmolarity is determined by trial for each new material before launching of mass isolation of cells or protoplasts. The enzymes used for removing pectin and celluloses are dissolved in the same osmoticum to be used later for culturing the cells.

8.1.7 Surface Sterilization

Explants obtained from sources outside the tissue culture room have to be cleaned of gross soils and surface-sterilized before plating on culture media.

The particles of solid and other extraneous fine particles of matter are first washed off using a detergent solution (household products such as DET or Teepol), followed by several rinses of tap water. After a final rinse or two in sterile distilled water, the explant is placed in a beaker or a test-tube for surface sterilization.

The sterilizing agents commonly used are mercuric chloride ($HgCl_2$, 0.1–0.2%), sodium or calcium hypochlorite (common household bleach diluted for use) or chlorine water (prepared in a set-up in the laboratory). The last set-up consists of a thistle funnel fitted to a glass jar having an outlet near its bottom. The outlet is connected with a rubber tubing to a glass tube that enters into another glass bottle. The open end of this glass tube is submerged in distilled water present in the receiving glass bottle. The first glass bottle contains $KMnO_4$ solution at a level below the outlet. When HCl is added to the thistle funnel, its drops fall on the $KMnO_4$ solution. Chlorine is released under water in the receiving glass bottle and dissolves in it. The water becomes straw-coloured. The strength of the chlorine water may be standardized by testing density with a densitometer. This will ensure a standard concentration of the sterlizing solution for routine use.

The time required to surface sterlize explants or seeds depends on the type of organ and the size of the explant. Usually 1–5 minutes are sufficient for carrots and stem explants, while leaves, and hypocotyles require less time. Usually leaves and seeds are especially washed well with Teepol to remove microorganisms clinging to dry crevices in the material. Teepol may also be added to the sterilizing solution.

A sterilizing agent, such as $HgCl_2$ kills the cells at the cut surfaces on prolonged exposure. The cut edges should be, therefore, sliced off before culturing the explants. Dead explants usually turn brown, and should be discarded.

When it is required to avoid all contaminants, plants are grown in the tissue-culture room from surface-sterilized seeds. For instance, surface-sterilized and sterile-water washed tobacco seeds may be planted on a solid medium which contains the following ingredients:

1. Inorganic salts of SH; use $\frac{1}{4}$ of the stock
2. Fe-EDTA of MS; use $\frac{1}{4}$ of the stock
3. Agar to make 0.9–1% solution
4. Sucrose to make 1% solution

   pH 5.9

The plants germinated from the seeds are then micro-propagated by transferring the stem tips (with 3 leaves) to a fresh medium surface (another flask). If growth is not satisfactory, gibberelic acid (0.1 mg/litre) may be added to the culture medium. Stems and leaves from these plantlets may be utilized for isolating the uncontaminated single cells or naked protoplasts.

8.1.8 General Directions for Plant Tissue Culture

Callus, and single cell culture, strain amplification or making of somatic cell hybrids involve the right choice of culturing vessels, method of sterilization and maintenance of aseptic work environment and the preparation of culture media.

1. Culture Vessels

For standard work, large size rimless glass test-tubes of good quality are used. 100 ml conical flasks, and autoclavable screw-cap bottles (like jam bottles) and petri dishes may be used depending on the objective of the exercise. Pre-sterilized disposable tissue culture tubes and petri dishes are also available; they are too expensive for regular use by the developing countries.

The culture vessels must be cleaned with detergent, rinsed several times in distilled water and oven-dried before filling with the prepared medium. The mouth of a tube or flask is stoppered with non-absorbent cotton plugs wrapped in cheesecloth. These plugs may be reused after sterilization.

2. Sterilization

The medium is autoclaved at 15 psi for 20 minutes with the mouths of the vessels plugged with the cotton-cheesecloth stoppers and covered with aluminium foil. A pressure-cooker type or larger autoclave is used for sterilization.

Hormones added to the medium are autoclavable. But for macerating tissues and obtaining protoplasts, the enzymes must be filter-sterilized using a Millipore filter assembly. If a small amount is required, a small syringe-filter-assembly may be used (see Appendix). For larger quantities, use the Millipore Buchner funnel-vacuum-pump assembly that filters about 250 ml at a time.

Explants are surface-sterilized with any of the following:

(i)	$HgCl_2$	0.5–0.1%	2–8 minutes
(ii)	Ca-hypochlorite	0.1%	5–10 minutes
(iii)	Na-hypochlorite	0.1%	5–10 minutes
(iv)	Chlorine water	10%	5–10 minutes

A nonionic detergent, Teepol, may be added (0.05%) to the sterilisant.

3. Preparation of Medium

The MS (Murashige and Skoog), NN (Nitsch and Nitsch), SH (Schenk and Hildebrandt) and B5 (Gamborg) are some of the media used for most routine work, with some species responding better to one or more of them. MS medium is used for carrots in the author's laboratory. NN medium is particularly good for Solanaceous species. (*Datura innoxia,* tomato, potato, tobacco). If you are not sure of which medium to use for your experiment, you should try MS, NN and SH as starters. As you gather experience, you may modify the standard medium if necessary, to suit the species with which you are working.

Directions are given for preparing the MS medium from (i) the basic ingredients and from (ii) a prepared commercial mix. The same general directions may be followed for the preparation of the other media.

(i) Examine the list of ingredients given in the Appendix, and collect the items on the work bench.
(ii) Have ready several bottles of sterilized distilled water (SDW).
(iii) Have ready the appropriately sized conical flasks for each class of items (e.g., Major Elements A, B, C, D, E, F and G) and label them. After solutions are made, 'stock' solutions C, D, E and F may be stored in the freezer in polyethylene bottles (screw-capped), that are labelled with water-proof ink.
(iv) Have ready pieces of butter paper (wax paper) (8 cm diameter) in a covered petri dish, for weighing the chemicals.
(v) Now prepare the medium. Weigh out and add in the following order:

Solution A

1. Weigh out the following major inorganic salts and place them in a 1 litre glass beaker, containing about 250 ml of SDW. Place the beaker on a magnetic stirrer.

(i)	Ammonium nitrate	1.65 g
(ii)	Potassium nitrate	1.90 g
(iii)	Potassium dihydrogen phosphate	0.17 g
(iv)	Calcium chloride	0.44 g

Solution B

2. Weigh out the following and mix them:

(i)	$MgSO_4$, $7H_2O$	0.37 g
(ii)	$MnSO_4$, $4H_2O$	0.223 g

3. Add the above mixture to the beaker.

Solution C

4. Make the following stock solution:
   (i) Boric acid 62 mg
   (ii) $ZnSO_4$, $4H_2O$ 86 mg
   (iii) KI 8.3 mg
   (iv) Sodium molybdate 2.5 mg
   (v) Sterile dist. water 100 ml
5. Add 10 ml of stock C to the beaker.

Solution D

6. Make the following stock solution:
   (i) $CuSO_4$ 2.5 mg
   (ii) $COCl_2$ 2.5 mg
   (iii) SDW 100 ml
7. Add 1 ml of stock D to the beaker.

Solution E

8. Make separately the following solutions 'a' and 'b' and mix them.

Solution a $FeSO_4 \cdot 7H_2O$ 55.1 mg
SDW (hot) 100 ml

Solution b Na_2EDTA 0.745 mg
SDW (hot) 100 ml

9. Add 5 ml of the above mixture to the beaker.

Solution F

10. Prepare a stock solution in 1000 ml:
    (i) Thiamine HCl 0.1 mg
    (ii) Pyridoxine HCl 0.5 mg
    (iii) Nicotinic HCl 0.5 mg
11. Prepare* a 10 ml of the stock **solution a** containing.
    10 mg of (i)
    50 mg of (ii)
    50 mg of (iii)
    10 ml of SDW
    * Use 0.1 ml of this solution in the 1 litre of medium.

Solution G

12. Weigh the following and add to the beaker:
    (i) Glycine 2.0 mg
    (ii) Myo-inositol 100 mg

Carbon Source

13. Add to the above: sucrose 30.0 g
14. Add more water to dissolve the above, if necessary.
15. Make up the volume to about 950 ml with dist. water.
16. Adjust pH to 5.8, using 0.5 M NaOH.

17. Pour the mixture into a 1 litre volumetric flask, and add dist. water until the 1 litre mark is reached.
18. Pour the medium into a large beaker and stir till the ingredients are mixed thoroughly. Place on a hot plate and warm to about 45°C.
19. Add to the beaker:
    Tissue Culture Grade Agar 8.0 g
20. When agar melts and mixes with the mixture, dispense the latter into (50–60 ml) test-tubes (20 ml/tube) Plug mouths of tubes with cotton plugs.
21. Autoclave at 120°C for 20 minutes.
22. Place 12 empty tissue culture test-tubes horizontally, in tandem on a work bench and place each autoclaved tube in a slanting position, with the mouth resting an inch beyond the empty tubes. Leave overnight, or till the medium has solidified.
23. Store at 4°C if not required to be used within the next week. Otherwise, keep in the tissue culture room (25 ± 2°C). The medium should be used within the next 7 days.

4. To Reconstitute MS Medium

1. Add 200 ml dist. water in a one litre volumetric flask.
2. Add to this the MS medium (without sugar, agar and hormones) and allow to dissolve.
3. Add 20 gm sucrose and shake to dissolve, adding more water if required.
4. Add to the flask: 440 mg $CaCl_2$, $2H_2O$.
5. Add hormones: (if required)
   IAA 1–30 mg
   Kinetin 0.04–10 mg
6. (a) If the medium is to remain liquid, the volume is made up to 1 litre at this stage. Check pH (5.8).
   (b) If the medium is to be solidified, do not let the volume of the liquid mixture exceed 700 ml.
   Take 100 ml dist. water in a beaker.
   Add 8 g agar (tiss. cult. grade).
   Place the beaker on a hot plate and let the agar dissolve in the warm water. Add this agar solution to the liquid medium, and make the volume up to 1 litre. Check pH (5.8).
7. Distribute the 1 litre medium into 50 large (60 ml) culture tubes (about 20 ml/tube); plug the tubes and place them on the test-tube holders.
8. Autoclave tubes at 15 psi (120°C) for 20 minutes.
9. Keep the autoclaved tubes in an inclined position for making slants.

5. Laboratory Requirements

1. **A Laminar Flow Cabinet** This is a work bench enclosed by a plastic roof, and walls on three sides. The back of the cabinet is covered by a filter system through which air flows into the cabinet (Fig. 8.4). Coarser particles in the

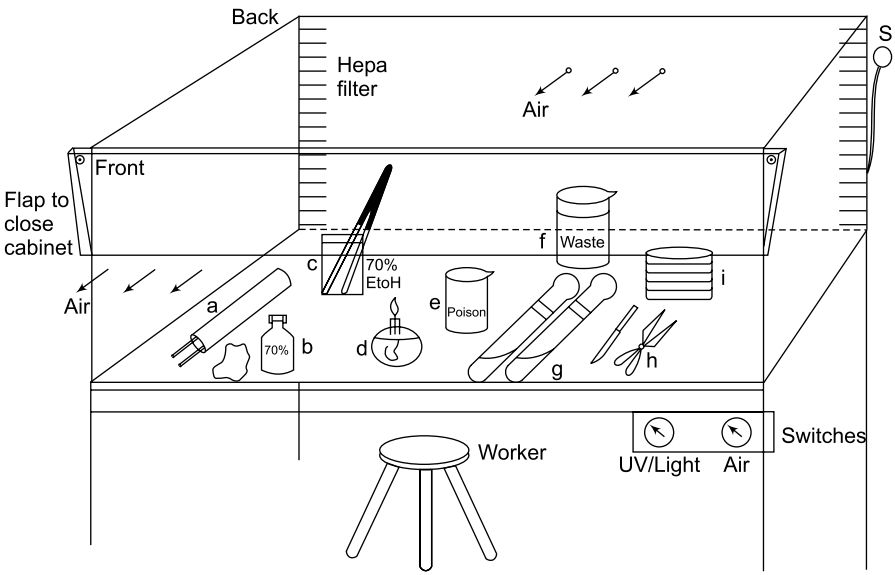


Fig. 8.4 *Laminar Flow Cabinet for working in an aseptic atmosphere. In this horizontal flow cabinet, air enters through a HEPA filter which removes the finest dust particles. An alcohol lamp or Bunsen burner is used for sterilization of instruments, mouths of tubes and bottles and so on. The layout in this figure includes (a) pipette container with sterile pipettes, (b) cotton wool and 70% ethyl alcohol, (c) a pair of large forceps, kept dipped in 70% ethanol when not in use, (d) spirit lamp, (e) $HgCl_2$ solution (labelled POISON), (f) beaker for waste water and solutions, (g) sterile culture tubes with medium, (h) dissecting instruments, (i) stack of sterile petri dishes. S, the main switch. The UV switch is turned on half an hour before work, to sterilize the inside of the cabinet. During work, the light switch is turned on, and the UV switch turned off.*

air are sieved out by a filter having larger pores than a second one, the HEPA (high efficiency particulate air) filter which keeps out particles as small as 0.2–0.3 μm in diameter. The air that constantly flows through the cabinet is thus free of viruses, bacteria and microbial spores, as well as dust particles.

2. **A Plant Tissue Culture Room** It is a room with temperature maintained at 25°C ± 2°C, humidity 50–60%, and light intensity of 1,000–12,000 lux from 40 watt fluorescent tubes.
3. A thick, black cloth to cover material that is to be kept in the dark.
4. Butter (wax) paper to cover fluorescent tubes, for material required to be exposed to diffuse light.
5. Equipment for rotating suspension cultures, the type depending on the particular objective and culture vessels utilized.
6. Millipore syringe type filter sterilization kit for sterilization of small quantities of thermolabile material.
7. Millipore Buchner funnel type-vacuum pump-sterilization assembly for filter sterilization of larger quantities of thermolabile substances.

6. Requirements for any Plant Tissue Culture Work

1. 20 test-tube slants of desired medium. We will mostly use MS, with and without hormones.
   1 litre MS makes 40 slants.
   or
   Conical flasks, petri dishes or jam bottles
2. $HgCl_2$, 0.1% (or Ca-hypochlorite 9–12% or Na-hypochlorite, 2.5% or chlorine water).
3. A screw-capped large-mouthed bottle (jam bottle) for surface sterilization of explants.
4. A large bottle or beaker labelled 'waste' for rinse waters.
5. A bottle labelled 'waste $HgCl_2$'.
6. 2–3 500 ml screw-capped bottles containing sterile dist. water.
7. A tall heavy glass cylinder with 70% ethyl alcohol for keeping forceps and scalpels dipped, when not in use (a bottle with the neck sawed off and the edge flame polished will be more stable).
8. One 500 ml bottle of 70% ethyl alcohol.
9. The following items should be wrapped individually in brown paper and dry heat sterilized at 180°C for 2 hours (in a dry heat oven).
   (a) A pair of sharp pointed small forceps
   (b) A pair of 10 inch forceps with rounded ends.
   (c) 2 scalpels (with detachable blade)
   (d) 2 dissecting needles.
10. Alcohol lamp, and a $6'' \times 7''$ tin partition to keep the flame from blowing out inside the laminar flow cabinet.
11. 2 sterile petri dishes
12. A sterile ceramic tile ($6'' \times 6''$).
13. Cotton wool, marker pen, box of matches.
14. A bucket for waste plugs, paper etc.

Exercise 39: To Culture Calluses from Explants

Have ready

1. Leaves, petioles, carrot root, young stem pieces, hypocotyls or soaked seeds (ideal material for initial attempts are petunia, tomato, *Datura* spp., carrot, and potato tuber)
2. $HgCl_2$, 0.1% (in a bottle marked POISON)
3. Dist. water
4. Sterile dist. water (SDW)
5. Teepol
6. Sterile forceps (8–10″)
7. Sterile large scalpels, and smaller forceps
8. A few sterile petri dishes
9. Tubes with slants of agar-solidified MS medium, without growth regulators
10. A large heavy bottle containing 70% ethyl alcohol, for dipping long forceps when not in use.

11. A bottle of 70% ethyl alcohol for surface sterilizing laminar flow cabinet working surface, hands, bottles, and other containers to be placed in the cabinet.
12. Two large beakers for discarding water from rinses. Mark them POISON, as the rinses will contain the washed off $HgCl_2$.
13. A Laminar Flow Cabinet (LFC); Turn on the switch for the UV lamp at least 30 minutes before start of work. Then turn off UV lamp and turn on the tube light switch for work.

N.B. Do not stay in the room when the UV lamp of LFC is switched on.

Procedure

1. Bring material for explants and keep them in the laboratory with cut ends dipped in water to prevent desiccation.
2. Cut off the unwanted parts (leaves, buds, rootlets etc., from stems and roots).
3. Cut the material into suitable lengths and wash with Teepol and water. Use a brush to scrub carrots, potatoes and roots.
4. Rinse several times to remove the detergent.
5. Place the explants in a beaker containing enough 0.1% $HgCl_2$ in order to submerge them; or place one explant per test-tube containing 10 ml of the sterilizing solution. Cover the mouth of the beaker or tube with Al foil and swirl the material to allow the surfaces to interact with the solution. $HgCl_2$ treatment varies from 1–2 minutes (for leaves) to 4–5 (for stems) and 7–10 (for seeds).
6. Pour out the $HgCl_2$ into the waste beaker, and rinse the explant with several changes of SDW.
7. Place the explant on a sterile glass plate or petri dish and cut it into 1 cm long or 1 cm^3 pieces, after first removing the browned off ends of the explant.
8. Innoculate one explant per tube, using the long forceps, flamed after removing from the ethyl alcohol bottle.

N.B.
- Be careful not to let the ignited ethyl alcohol, from the flamed forceps, reach your hand.
- First cool the tip of the flamed forceps by letting it touch the inner surface of the sterile petri dish, before picking up an explant.

9. Keep the tubes in the tissue culture room in the dark (cover the tubes in racks with black cloth) for 24 hours, and then expose to light for 16 hours or more (according to specific requirements). The temperature is maintained at 25 ± 2°C, the humidity at 50–60%, and intensity of light at 1000–1200 lux.

Expected results After 4–6 days you may observe initiation of callus at the cut edges of the explant.

10. Subculture into fresh containers after every 3–4 weeks, cutting up the callus into smaller pieces each time. The minimum size that will subculture differs from species to species. For instance, carrot can be subcultured from much smaller pieces of the initial callus than the potato.

Exercise 40: Anther (Haploid Culture)

Have ready

1. Flower buds of different sizes from the same species (use *Datura* sp., petunia, tomato, tobacco or *Brassica* sp)
2. Acetocarmine stain (1% in 45% acetic acid)
3. Microscope slides and coverslips
4. Teepol
5. $HgCl_2$ 0.1% solution.
6. Sterile dist. water (SDW)
7. Innoculating needles (sterile)
8. Forceps (sterile)
9. Large beaker labelled: WASTE-POISON
10. 70% ethyl alcohol
11. Tissue culture Medium* in culture tubes NN (Nitsch and Nitsch) or MS (Murashige and Skoog)**
    * The medium is solidified with agarose instead of agar, since the latter often contains impurities that kill pollen cells.
    ** The medium may be supplemented with coconut milk instead of defined growth regulators. The milk may be used fresh or obtained from material kept stored in sterile screw-capped vials in the deep freeze (–20°C). 15 ml coconut milk/1000 ml medium to be added before adjusting the pH of the medium to 5.5–5.7 (using NaOH/HCl).
12. • Tissue culture medium (same as Step 11) but in 250 ml conical flasks (50 ml per flask)—to be used 25 days or so after start of anther culture.
    • If there is no shoot or root differentiation in the first 25 days, the above subculturing medium should be supplemented with the usual growth regulators.
    • *For shoots:* BAP, Kinetin or Zeatin
    • *For roots:* IAA. IBA, NAA, or 2, 4-D

Outline of procedure

I. Select buds with anthers at the correct stage (tetrad or uninucleate microspores)
II. Dissect anthers from surface-sterilized buds.
III. Place anthers on the surface of solid tissue culture medium in test-tube slants.
IV. Incubate the tubes in tissue culture room at 25 ± 1°C, 2,500–16,000 lux (16 hour light 8 hour dark periods) or 6,000–7,000 lux (continuous light).
V. Transfer calluses/embryoids for subculturing to 250 ml flasks with medium supplemented with coconut milk and/or growth regulators. Plantlets will develop.

These plantlets or calluses may be used for diploidization of haploid cultures

Procedure

1. Select buds with microspores at the uninucleate stage.
2. Place the selected buds in a beaker of Teepol solution. Swirl to wet the buds thoroughly
3. Rinse the buds several times (3–5 times) with dist. water, ending with a rinse in SDW.

4. Place the washed buds in a 50 ml beaker filled 1/3 with 0.1% $HgCl_2$ soln. Keep for 20 minutes, swirling the beaker every few minutes.
5. Decant the $HgCl_2$ soln. into the 'Waste-Poison' beaker.
6. Rinse the buds 4–6 times with SDW, discarding the rinse water in the waste beaker.
7. Dissect each bud using flamed forceps and a dissecting needle, removing first the bracts, if any, and then the sepals and petals.
   Check the stage of microspore using one anther per bud.
   If the stage is correct, place the remaining anthers on the slant surface of the medium in tubes, pressing the anthers firmly into the medium (large anthers, 1/tube; small anthers 2–4/tube).
   Replug the tubes after flaming the plugs and mouths of the tubes.
8. Incubate the tubes in the culture room at 26 ± 1°C at a high light intensity (2,500 to 8,000 lux depending on the species).

N.B.
- Plants kept at a lower temperature (12–15°C) for several days often yield more embryoids when their anthers are cultured at 26 ± 1°C or 35°C (some Brassicas).
- After innoculating the anthers on the medium, the tubes may be kept in the cold room (4°C) for a few days (the period varies with different materials), before incubating at 26 ± 1°C in the tissue culture room.
- Instead of continuous exposure to light, the cultures may be subjected to 16 hour light and 8 hour dark periods by covering the tubes with black cloth during the working day (8 hours).

9. After 10–15 days, calluses or embryoids appear, and in about 25 days from the start, distinct plantlets may be seen.
10. After 25 days, transfer calluses or embryoids to 250 ml flasks containing 50 ml each of the medium used earlier.
    * If no root or shoots are seen by the 25th day, add growth regulators to the medium in the flasks to be used for subculturing. (See page 498 Step 5 for amounts of hormones to be used.)
    - *For shoots use:* BAP, Kinetin or Zeatin
    - *For roots use:* IAA, NAA, IBA, or 2, 4-D
11. Transfer the plantlets to small pots, and later to larger pots containing sterile soil.
12. Check the chromosome numbers of cells from calluses, shoots and roots. Are the numbers the same as the haploid number characteristic of the species? If not, what are the differences, and what percentages of the plants show abnormal chromosome numbers?

Exercise 41: Protoplast Culture from Friable Callus

A friable callus is a fast-growing one in which the small aggregates of cells disintegrate readily into fragmented calli. Naked protoplasts are obtained by treating the calli with cellulase. Optimum conditions for the isolation and culture of protoplasts have been arrived at empirically, and, therefore, are subject to modification.

The naked protoplasts are extremely sensitive to the osmotic value of the medium in which they are bathed. An appropriate osmoticum is, therefore, utilized in the

enzyme and washing solutions. If the osmolarity of the solution is equivalent to that within the cell, the latter will neither burst nor be plasmolyzed; it would be perfectly spherical in shape. Usually, carbohydrates mannitol and sorbital are used as the osmotic stabilizer (500–800 mmol).

The commercial enzyme preparations that are relatively free of nucleases and proteases are Onozuka R–10 cellulase and macerozyme from Yakult Co., Japan.

After removal of the cell walls, the debris is discarded by one of the following methods:

1. By passing through a stainless steel sieve (of 0.45 μm pore size) or through a double layer of fine nylon or cheesecloth.
2. By loading on 21% sucrose solution in a centrifuge tube and spinning at 200 g (600 rpm in a clinical centrifuge) for 10 minutes. The protoplasts float on top, while the debris settles at the bottom.
3. The enzyme solution with the protoplasts and debris is centrifuged at 100–200 g for 3–5 minutes, the pellet resuspended in the washing medium and spun at the same speed. The pellet may be washed 2–4 times by resuspension in the washing medium and spinning for 3–5 minutes each time. Use the screw-capped glass centrifuge tubes (15 ml) or use tubes capped with Al foil.

The density of cells should be about 2×10^5 per ml of the liquid. A glass slide with hatchings, called the hemocytometer is used to count cells under the microscope. The volume of fluid under the coverslip in the hemocytometer is known; the number of cells per ml is calculated from this. The viability of isolated protoplasts may be assessed by treating them with a solution of Evan's blue (0.025%). The viable cells will not stain, while the damaged ones turn blue.

The naked protoplasts acquire new cell walls within 1 to 3 days after the enzyme treatment. It is preferable to keep them in the dark or in diffuse light until cell walls are formed.

At first, cells divide within 1 to 6 days; sometimes it takes longer for the first mitosis to occur. Cells from younger tissues also divide earlier than those from older tissues. After about 5 weeks, clusters of cells become visible. These colonies may now be subcultured in a medium in which no osmoticum has been added. The colonies grow into calluses, which may be regenerated with a combination of NAA or IAA and Kinetin or only the auxin 2, 4-D.

Have ready

1. 40 sterile 55 mm petri dishes
2. 40 sterile 85–90 mm petri dishes
3. Sterile screw-capped or Al foil-capped centrifuge tubes (15 ml)
4. Sterile Pasteur pipettes with cotton plugs
5. Gilson's Pipette (50 μ) with sterile tips
6. Fine mesh cheesecloth (sterile)

   or

   50 μ and 100 μ size stainless steel sieve
7. The standard glassware, lamp, forceps, alcohol etc., for tissue culture work
8. Sucrose: 20% solution
9. Mannitol 0.7 M solution pH 5.8 (sterilized) — 1 litre

10. MS medium (liquid) 500 ml
11. Enzyme solution: filter sterilized
    Macerozyme 1.5%
    Cellulase 2.0%
    made in 0.7 M Mannitol.
12. Soft agar (0.4%) kept melted at 45°C (in a water-bath)

Outline of procedure

I. Grow friable callus from explants.
II. Isolate protoplasts from calluses
III. Purify the viable protoplasts and prepare suspensions of required cell density.
IV. Culture the protoplasts in each of the following:
    (a) Liquid Culture Medium
    (b) Semi-solid Culture Medium
    (c) Micro-Droplet Arrays (MDA).

Procedure

I. Follow the instructions of Exercise 39 to obtain calluses (allow 2 months for this) or use calluses already available from the earlier cultures.

II. Isolation of protoplasts

WORK IN THE LAMINAR FLOW CABINET

1. With the help of flamed forceps, transfer the calluses to a sterile petri dish.
2. Using flamed scalpels, chop the calluses into small fragments.
3. Add to the petri dish 10 ml enzyme soln./1 gm (for a callus of ~2 mm diameter).
4. Incubate at 25 to 28°C for ~2 hours, swirling the dish occasionally.
   * Take a drop with a sterile pipette and place it on a microscope slide. Cover the drop with a coverslip and observe, under high power (40 × objective), the degree of protoplast formation.
5. Stop incubation when the majority of the cells have become spherical (spheroplast, protoplast).
6. Pour the suspension through a sterile gauze fixed over a beaker or consecutively through a 50 μ and 100 μ stainless steel sieve fixed to a holder. The debris will remain on the sieve, and protoplasts collect in the beaker.
7. Pour the protoplast suspension into sterile centrifuge tubes (≃ 5 ml per tube). Close the mouth of the tube with a cap or use Al foil and spin at 100 rpm in a clinical table top centrifuge for 3–4 minutes.
8. Decant the supernatant and discard it.
9. Add 0.7 M mannitol to the pellet, and resuspend the cells using a sterile Pasteur pipette. Check the cells under the microscope for viability, using Evan's blue dye.

III. Preparation of protoplast suspensions

10. Take 5 ml of sterile sucrose solution in a sterile centrifuge-tube. Using a fine Pasteur pipette, carefully layer (add drop wise along the wall of the tube) 1.5 ml of protoplast suspension over the sucrose solution.

11. Centrifuge the above at about 200 rpm for 10 minutes. The debris will collect at the bottom.
12. With a sterile Pasteur pipette, remove a little suspension from each layer in the centrifuge tube. Place one drop on a microscope slide for examination. Transfer the remaining suspension from the pipette to a sterile centrifuge tube containing 0.7 M mannitol ($\simeq$ 5 ml).
13. The tube with the maximum protoplasts is spun at 100 rpm for 3–4 minutes. Discard the supernatant.
14. Resuspend the pellet in 5 ml MS medium containing 0.7 M mannitol. Check protoplast density with a hemocytometer; it should be 10^4–10^7/ml. Adjust the volume with MS medium to give the required density.

IV. Culture methods This protoplast suspension is now to be cultured by methods (i), (ii) and (iii), mentioned in the 'outline of procedure'. Start each suspension with step 15 (a, b, c), and then follow the individual steps for each type of culture.

IV(a). Liquid Culture

15a. Place 2 ml of suspension in a sterile small petri dish. Seal the sides of the dish with parafilm. Leave in the Tissue Culture Room (TCR) (25 ± 2°C) in the dark.
16a. Check for mitosis after 48 hours. In most cases, there will be cell-division by this time. Add to the culture, that shows mitosis, 1 ml MS + 2/3 of 0.7 M mannitol (= 0.47 M). Leave in the TCR in the dark.
17a. There should be clumps of cells after 10–12 days. Transfer the clumps with a pipette to 2 ml MS + 1/3 of 0.7 M mannitol (= 0.23 M). Leave in the TCR in the dark.
18a. After 2 weeks transfer the clumps to 2 ml MS (without mannitol) in a test-tube.
19. Add to the tube 2 ml melted agar (45°C). Swirl the mixture and pour into a sterile small petri dish. Seal the sides of the dish with parafilm and keep in the TCR in diffuse light.

IV(b). Semi-solid Culture

15b. To 1 ml protoplast suspension in MS, add 1 ml melted 0.8% agar. Swirl the mixture and pour into a small (5 mm) sterile petri dish. Seal the edges of the dish with parafilm.

N.B. The protoplasts should not sink.

16b. Leave the culture in the TCR (25 ± 2°C) in the dark for 24 hours, then expose to diffuse light.
17b. After 10–12 days, there should be growing colonies of cells. Cut out blocks of agar containing colonies and transfer to fresh petri dishes containing liquid MS medium + hormones.

N.B. If the growth of the callus is not good, increase the proportions of the hormones in the liquid MS of the next transfer.

18b. Subculture the calluses on MS + hormone slants.

IV(c). MDA Culture

15c. Mark the outer surface of the lid of a sterile 90 mm petri dish into squares with a marker pen.

16c. Take the lid off; place it with its inner side facing up. Place 1 drop of the protoplast suspension (using a 50 µm Gilson or pushbutton pipette) in each square.

17c. Turn the lid right side up, without disturbing the drops, in the following way: Hold the lid in the palm of your hand and swing your arm swiftly through 360° by moving down, back, up, and front like a discuss thrower. The drops will be hanging from the roof of the lid. Practice first with drops of water placed, as stated above, on the inner surface of a petri dish lid.
Fit the lid over the petri dish bottom.
To prevent the drops from drying up, you may use one of the following methods:
(i) Seal the edges of the petri dish with parafilm.
(ii) Add a little MS or water to the bottom half of the dish and seal the edges with parafilm.

or

(iii) Place the dishes in a large humid chamber.

18c. Leave the MDA dishes in the TCR in the dark for 24 hours and then in diffuse light.

19c. Using a Pasteur pipette, transfer each drop with dividing cells on to the surface of a semi-solid medium (2 ml MS + 2 ml 0.8% agar) in a 55 mm petri dish.

20c. Cut out blocks of agar containing growing clumps of cells and place them in a fresh petri dish containing 2 ml liquid MS + hormones.

21c. Incubate in diffuse light in the TCR (25 + 2°C) and transfer, every 2 weeks, to fresh liquid MS, until the calluses are large.

22c. Transfer the calluses to MS + hormone slants in culture tubes.

N.B. • You may try different media (SH, NN, B5) in the same MDA lid. Be sure to label the columns correctly.

Exercise 42: Protoplast Culture of Leaf Cells (I)

The cells from the spongy mesophyll tissues of leaves are easily converted into naked protoplasts. These may be cultured into single-cell derived calluses, regenerated into plants, and used for somatic cell hybridization or for induction of mutations.

Have ready

1. Young opened out leaves of *Vinca rosea* (periwinkle) or *Nicotiana tabacum* (tobacco)
2. Teepol
3. 70% ethyl alcohol
4. Na-hypochlorite (0.5% w/v) soln. in water
5. Sterile double dist. water
6. CPW (cell-protoplast-washing) solution:

$CaCl_2 \cdot 2H_2O$	1480 mg
$MgSO_4 \cdot 7H_2O$	246 mg

$KH_2 \cdot PO_4$	27.2 mg
KNO_3	101.0 mg
$CuSO_4 \cdot 5H_2O$	0.025 mg
KI	0.16 mg
Double dist. Water	1000 ml
pH	5.8

Autoclave at 15 psi (120°C) for 20 minutes.

7. Mannitol 12% in CPW solution. Make 1 litre of this solution (12 M CPW):
8. Sucrose 20% in CPW soln. (20 S CPW)
9. Enzyme solution:

Macerozyme	0.5%
Cellulase	3.0%
12 M CPW	100 ml

Place powders in a measuring cylinder and add 12% mannitol to dissolve and make up volume to 100 ml.

10. One extra large (140 mm) petri dish (sterile)
    Two 90 mm petri dishes (sterile)
11. One 100 ml sterile flask
12. One wide-mouthed, screw-capped bottle (jam bottle); sterilized
13. Sterile screw-capped or Al foil-capped centrifuge tubes
14. Sterile forceps (small, sharp, and with bent ends; one each)
15. Sterile Pasteur pipettes
16. Waste beaker for discards

Procedure

1. Take about 25 *Vinca* leaves and wash them in a beaker containing Teepol solution. Rinse several times in tap water, and finally in sterile distilled water 2–3 times.

WORK IN THE LF CABINET

2. Place the leaves in a wide-mouthed bottle or beaker and submerge them in 70% ethyl alcohol for half a minute.
3. Decant the alcohol and add 0.5% Na-hypochlorite solution to submerge the leaves. Keep for 30 seconds, swirling gently so that every leaf is bathed in the sterilizing solution.
4. Decant the Na-hypochlorite solution and rinse 3–5 times with sterile distilled water.
5. Have ready: A sterile petri dish (140 mm or two 90 mm) with about 25–30 ml of 12 M CPW. (Place half the volume of 12 M CPW in each 90 mm petri dish)
6. Peel off the lower epidermis of each leaf in the following manner:
   (i) Place a leaf, with the lower surface facing up, on a sterile petri dish bottom.
   (ii) Hold a vein of a leaf with a pair of fine forceps in one hand and peel off the epidermis with a bent fine forceps held in the other hand pulling from a vein outwards.
   (iii) Place the peeled leaves, with exposed surfaces touching the solution, in the petri dish kept ready as per Step 5.

7. When the surface of one petri dish is covered with peeled leaf pieces, move on to the next dish, until all mannitol-CPW surfaces are covered.
8. Using a fine bore Pasteur pipette, remove the mannitol CPW from each dish and discard it into a 'waste' beaker.
9. Perform substep (i) or (ii) below:
   (i) Now, add about 25–30 ml of the sterilized enzyme mixture to the 140 mm dish or 12–15 ml to the 90 mm ones. Seal the edges of the petri dishes with parafilm and leave in the dark in the TCR (25 ± 2°C) overnight or for about 18 hours.
   (ii) Place the leaf pieces in the enzyme solution in a sterile wide-mouthed, screw-capped bottle. Keep in the dark 25–27°C for about 18 hours or overnight.
10. (i) Remove the enzyme solution from beneath the leaf pieces. Add 25 ml of 20 S CPW soln. below the leaf pieces. Tap the latter with the tip of a Pasteur pipette to release the protoplasts. The fluid will be green due to a suspension of protoplasts. Tilt the petri dish and pipette off the suspension into a flask.

    or

    Sieve the material in the petri dish after enzyme digestion, through a fine mesh cheesecloth or stainless-steel (100p) gauze. Collect the filtrate in a beaker or flask.
    (ii) Tap the sides of the bottle several times. The protoplasts will be released into the enzyme solution. Transfer the suspension to a sterile flask.
11. Transfer the suspension to sterile centrifuge tubes and spin at 1000 rpm (~100 g) for 10 minutes. Discard the supernatant.
12. Add 10 ml 12 M CPW and resuspend the protoplasts.
13. Spin at 1000 rpm for 10 minutes. Discard the supernatant.
14. Resuspend in 10 ml 12 M CPW soln. and spin at 1000 rpm for 10 minutes.
15. Repeat Step 14 once more (total 3 times)
16. Resuspend the protoplasts in 12 ml of 12 M CPW solution:
17. Take out one drop of the above to make a count of the viable cells. Adjust the cell density to 2×10^5 or 5×10^4 cells/ml.
18. Pipette out 4 ml of melted agar into a test-tube. Add to this quickly 4 ml of the suspension mix.
19. Pour the above (medium + agar + protoplasts) into a 90 mm petri dish. Allow the medium to set (30–40 minutes). Seal the edges of the petri dish with parafilm.
20. Incubate the petri dishes (inverted) in the TCR in the dark for 24–48 hours, and subsequently in light (25 ± 2°C).

Exercise 43: Protoplasts from Leaves (II)

In this exercise we will use the leaves from asceptically grown pea seedlings. The pea leaves are too small for the epidermis to be peeled off; hence the leaves are shred into small pieces for enzyme digestion.

Have ready

1. Pea seeds
2. Na-hypochlorite solution (12%) in water
3. $HgCl_2$ 0.1% soln. in water
4. Sterile dist. water
5. A large petri dish (140 mm)
6. Sterile Al foil to cover both halves of the petri dish
7. Sterile CPW solution
8. Sterile Mannitol 0.6 M solution in CPW (0.6 M CPW)
9. Enzyme solution:
   Macerozyme 0.5%
   Cellulase 2.5%
   } in 0.6 M CPW
   Filter-sterilize these enzymes.
10. Sterile bent forceps, blunt-ended forceps, and scalpels with sharp blades.
11. Sterile Pasteur pipettes
12. Push-button (Gilson) pipette with sterile tips
13. Sterile 100 ml conical flask
14. Waste beaker

Procedure

A. To grow seedlings under aseptic conditions

1. Wash the seeds, soak for 10 minutes in a Na-hypochlorite solution, wash 4–5 times with sterile dist. water. Soak in sterile dist. water at 26°C for 24 hours.
2. Half fill each half of the petri dish (top and bottom pieces) with sterile dist. water. Cover with Al foil. Make holes in the foil in a grid fashion.
3. Place the soaked seeds on the Al foil, with each seed next to a hole.
4. Keep in the TCR. The seeds will germinate.
5. When 2–4 leaves have appeared on each seedling, snip them off with a flamed sterile forceps and collect them, with a petri dish cover, in a pre-weighed small beaker (25 ml). You need about 4 gm of leaves.

B. To isolate protoplasts from leaves

1. Add $HgCl_2$ soln. to cover the leaves in the beaker. Swirl for about 2 minutes.
2. Decant off $HgCl_2$.
3. Rinse 2–3 times with sterile dist. water.
4. Add sterile 0.6 M CPW.
5. Lift out the leaves with a flamed pair of forceps and place in a petri dish containing 0.6 M CPW.
6. Shred the leaves with sharp blades in this petri dish and swirl with forceps.
7. Pour the enzyme solution into two 90 mm petri dishes.
8. With a Gilson pipette, transfer the 'green soup' from Step 5 into the waste beaker.
9. Now transfer the leaf debris into the enzyme solution using the pipette tip to spread out the leaf pieces. Cover the plates; seal the sides with parafilm and keep at 25°C for 24 hours (TCR).

10. Filter the material in Step 9 through two layers of fine stainless-steel net.
11. Distribute the filtrate into 4 centrifuge tubes and spin at 400–500 g for 3 minutes.
12. Discard the supernatant. Resuspend the protoplasts in 0.6 M CPW and spin.
13. Repeat Step 11 2–3 times. Pool the pellets.
14. Suspend the final pellet in 10 ml 0.6 M CPW.
15. Transfer the 10 ml suspension into a screw-capped tube (15 ml).
16. Check the tube directly under a stereo-dissecting microscope for protoplasts.
17. You may store the protoplasts in 0.6 M CPW for 24–36 hours or use them immediately for culturing, using the liquid, semi-solid or MDA method, for somatic cell hybridization or for the induction of mutations.

Exercise 44: Protoplasts from Stems, Roots, Petioles, or Hypocotyls

Have ready

1. Stem: Young stem tips that have not become woody, potato tubers.
   Root: Carrot
   Petioles: Petunia, Hollyhock
   Hypocotyl: Seeds germinated aseptically from surface-sterilized and washed seeds, soaked in sterile water.
2. CPW soln.
3. 12 M CPW
4. 20 S CPW
5. Enzyme solution:
   Cellulase 1%
   Macerozyme 0.1%
   } in 12 M CPW
   pH 5.6
6. Na-hypochlorite (0.1%) or $HgCl_2$ (0.1%) for surface sterilization of stem, roots and petioles and 10% Na-hypochlorite for seeds.
7. Sterile dist. water
8. A stainless steel or nylon sieve (50 μm size)
9. Sterile forceps, scalpels, beakers, flasks, petri dishes etc.

Procedure

1. (a) Surface sterilize in 0.1% $HgCl_2$:
   (i) Washed stems, petioles — for 2–3 minutes.
   (ii) Washed and scrubbed carrot roots, potato tubers — for 2–3 minutes.
   (b) Surface sterilize in 10% Na-hypochlorite:
   (iii) Washed seeds of
   - mustard (in a gauze pouch)
   - *Cucurbita pepo* (pumpkin)
   - mung bean, gram, kidney bean, groundnut (legumes) — for 10 minutes.
2. Wash several times in sterile tap water.

WORK IN THE L.F. CABINET

3. (a) Cut 1.0 mm cubes from stems and petioles, 0.5 mm cubes from carrots and potatoes and place in 12 M CPW for 1 hour.
   (b) Germinate the seeds, soaked for 24 hours in sterile water, in petri dishes lined with sterile wet filter papers
      (i) Cut 1 mm sections from the hypocotyles of at least 100 large or 200 small (mustard) seeds; place them in 12 M CPW.
      (ii) Remove 1 cm from the root apex and cut the remaining radicles into 0.5 mm sections. Use 100 large or 200 small seeds for one isolation; place in 12 M CPW for 1 hour.
4. Transfer the foregoing material to enzyme solution and shake slowly in the dark at 25°C for 18 hours or overnight (use a rotary shaker, if you have one, at 20–25 cycles/min).
5. Pour the enzyme-cell suspension mixture through the sieve. Add more 12 M CPW to wash off the protoplasts adhering to the debris on the sieve.
6. Transfer the filtrate to sterile centrifuge tubes and spin at 1000 rpm (100 g) for 10 minutes. Discard the supernatant.
7. Add to the pellet 10 ml of 20 S CPW and spin at 1200 rpm for 10 minutes. The protoplasts will collect above the 20 S CPW. Skim off the protoplasts with a Pasteur pipette and add to 12 M MS medium for culturing. Make a cell count.

Exercise 45: Plant Somatic Cell Hybridization

Somatic cell hybrids of plants are totipotent like non-hybrid cells. However, the actual culture conditions for regeneration of somatic hybrid cells (as for non-hybrid ones) are not known for the majority of species. Individual modifications for different inter- and intra-specific plant cell hybrids will be attempted to discover specific or general requirements for regeneraton.

The steps for creating a somatic cell hybrid in plants consist of:

1. Isolation of protoplasts, from parents, to be hybridized.
2. Fusion of cells using polyethylene glycol (PEG).
3. Isolation and maintenance of the fused cells.
4. Regeneration from fused cells.

Have ready

1. Material for fusion:
   Carrot root
   Tobacco or eggplant (brinjal) leaf
   10 ml of protoplasts (2×10^5 cells/ml) of each of the above are to be prepared in a protoplast culture medium.
2. Culture Medium for protoplasts:

Basal MS	1000 ml
Sucrose	20 gm
Mannitol	0.65 M
2, 4-D	2 mg
Kinetin	0.5 mg
pH	5.8

   Autoclaved at 15 psi (120°C) for 20 minutes.

3. Teepol, $HgCl_2$ 0.1% for surface sterilization
4. Sterile dist. water 6 × 500 ml (in bottles)
5. CPW soln. 1000 ml
6. Mannitol 0.7 M solution in CPW
7. Enzyme solution:
   Cellulase 2 gm
   Macerozyme 0.03 gm
   0.7 M Mannitol CPW 100 ml
8. 20% sucrose in 0.7 M Mannitol CPW (20 S CPW)
9. Fusion Agent:
   Polyethylene glycol (PEG)
   (a) *For Procedure A:*
   PEG 6000 30 gm
   0.7 M Mannitol CPW 100 ml
   (b) *For Procedure B:*
   PEG 4000 25 gm
   0.7 M Mannitol CPW 100 ml
   (c) Autoclave the PEG solutions.
10. Sterile centrifuge tubes with caps.
11. Sterile 90 mm petri dishes:
    8 for interspecific fusion
    4 for intraspecific (carrot) fusion
    4 for intraspecific (leaf) fusion
12. Sterile 50 mm petri dishes — 20
13. Sterile 140 mm petri dishes — 2
14. Sterile pasteur pipettes with cotton plugs
15. Sterile nylon fine sieves 100 μ
16. Sterile forceps, scalpels etc.
17. Sterile coverslips and microscope slides
18. Parafilm

N.B. You may follow procedures *A* or *B* or both.

Procedure A

1. Isolate the carrot protoplasts from roots, or from calluses from explants. Have ready 10 ml of protoplasts in the given MS medium.
2. Isolate the leaf protoplasts by the method given in Exercise 42 or 43. Have ready 10 ml of protoplasts in the given MS medium.
3. Arrange 8 sterile centrifuge tubes in a test-tube rack and label them as follows:
   I — Carrot
   II — Leaf
   III — Carrot × Carrot
   IV — Leaf × Leaf
   V — Carrot × Leaf
   VI — Carrot × Leaf
   VII — Carrot × Leaf
   VIII — Carrot × Leaf

4. Add 4 ml of protoplast suspension to the following tubes.
   I — Carrot protoplasts
   II — Leaf protoplasts
   III — Carrot protoplasts
   IV — Leaf protoplasts
   V-VIII — Carrot protoplasts
5. Add 4 ml of the given protoplast suspension to the following tubes:
   III — Carrot
   IV — Leaf
   V-VIII — Leaf
6. Mix the suspensions in each tube, spin the tubes at 1000 rpm or less for 10 minutes to pellet out the protoplasts.
7. Pipette out the supernatant leaving 0.5 ml with the pellet.
8. Add 2 ml PEG solution (6000) to the tubes III-VIII and let them rest at room temperature for 15 minutes.
9. Centrifuge at 1000 rpm for 10 minutes.
10. Resuspend the protoplasts in the given MS medium.
11. Spin at 1000 rpm for 10 minutes.
12. Pool pellets of the same kind into one centrifuge tube and make up the volume to 12 ml with the given MS medium.
    So, you have III 1 tube — Carrot × Carrot
    IV 1 tube — Leaf × Leaf
    V + VI 1 tube — Carrot × Leaf
    VII + VIII 1 tube — Carrot × Leaf
    The tubes I and II with carrot and leaf protoplasts respectively are to be used to assess the percentage of cell survival for each parental strain.
13. Arrange sixteen 50 mm sterile petri dishes each containing 4 ml of the given MS medium.
14. Add 4 ml to each petri dish as follows:

Tube No.	*Petri dish No.*
I	1
II	2
III	3 and 4
IV	5 and 6
V	7 and 8
VI	9 and 10
VII	11 and 12
VIII	13 and 14

15. Seal the edges of all petri dishes with parafilm and keep them in the TCR in diffuse light.

Procedure B

1. Have ready 10 ml each of carrot and petunia leaf protoplast suspensions in the given MS medium.
2. Place a sterile 90 mm petri dish in a sterile 140 mm petri dish, with drops of sterile water in the latter.

3. Place 3 sterile coverslips in the centre of the 90 mm petri dish and place drops of the culture medium on the petri dish around the coverslips.
4. Mix 5 ml of carrot and 5 ml of leaf protoplast suspensions.
5. Take a 50 µl Gilson pipette and place a drop of the protoplast mixture in the centre of each coverslip. Let the cells settle down (15 minutes).
6. Add 450 µl of the PEG (4000) solution medium around the protoplast mixture. The PEG and protoplast medium will coalesce.
7. Close the cover of the 90 mm petri dish, and then that of the larger one.
8. After 20 minutes, wash out the PEG by adding several drops of the culture medium to the protoplasts. (The excess medium will run off the coverslip.)
9. Close the cover of the 90 mm petri dish and seal its edges with parafilm and keep in diffuse light in the TCR (25 ± 2°C).
10. Observe the fusion of cells by placing the sealed petri dish under a microscope. Heterokaryons will have green chloroplasts from leaves and orange chromoplasts from carrot cells. Mitosis will be seen after cell wall regeneration (after 24 hours.).
11. Transfer the coverslips to 50 mm petri dishes containing 4 ml of the MS a medium + hormones + 0.7 M mannitol.
12. After 2 weeks, transfer the cell clumps on coverslips to fresh MS + 2/3rd of 0.7 M mannitol + hormones (NAA or IAA 1% and Kinetin 0.5%).
13. After 10 days transfer to the same MS-hormone mixture but with l/3rd of 0.7 M mannitol.
14. After 2 weeks transfer again but leave out the mannitol in the culture medium.
15. Transfer calluses to the agar solidified MS medium + hormones.
16. Try different combinations of auxins and kinetins (as also different ratios of the two) to find out the most effective ratio for plant regeneration.
17. You may examine the cells of calluses/regenerated shoot or root cytologically to determine whether they are interspecific somatic hybrids.

8.2 ❑ ANIMAL CELL CULTURE

Animal cells are more difficult to culture *in vitro* than plant cells. The former require special conditions that cannot be duplicated as yet *in vitro.* For instance, animal cells will grow *in vitro* only if some serum is added to the basic synthetic tissue culture medium. The serum represents a collection of components obtained *in vivo.*

Advances in the methods of animal tissue culture have followed, and also contributed to progress in our knowledge of animal cell biology. In the earliest days of animal cell culture, the cells or pieces of tissues were grown on some fluid or the tissues were obtained from a live animal. Lymph and discs of clotted blood were the favourite culture media. Such non-standard or black-box conditions were gradually replaced by a mixture of salts and organic compounds, plus, of course, a dose of the 'unknown' in the form of serum.

Earle (1948) and Eagle (1955) provided for the first time fairly standard recipes for culture media. Most animal tissue culture media in use are modifications of or derivations from 'Earles' and 'Eagles' media.

Single cell cultures could be made only when it was suspected that there had to be an optimum ratio between the volume of a cell and the media, for the latter to support the former. It was speculated that the cell had to modify the medium sufficiently before it could remain active in it. This suspicion led to the use of hanging drop cultures, cultures in narrow glass tubes and flask cultures. The optimum cell density in each case was determined empirically.

It was 'Eagles' group at NIH, USA, that discovered that most animal cells need to have a solid surface for growth and development. In the body, this is supplied by the extra-cellular matrix comprising several proteins and other macromolecules. In the laboratory a smooth glass or plastic surface serves the same purpose.

Several cell types, such as lymphocytes, do not undergo cell division *in vitro.* It was Newell who discovered by chance that plant glucosides called lectins, can trigger lymphocytes to begin mitosis. Newell had been using a lectin called phytohaemagglutinin (PHA) for agglutinating red blood cells in blood samples, in order to separate the leucocytes. In 1960, he reported that the PHA left in the supernatant, after clumping of the RBC, stimulated some of the lymphocytes to begin mitosis. This accidental discovery provided an important tool for studying lymphocytes and other leucocytes *in vitro.* PHA-induced lymphocyte cultures have provided unique opportunities to study cell-biology, and differentiation in both normal and neoplastic human cells. Lymphocyte cultures also opened up avenues for characterizing human chromosomes.

The next major landmark in the development of animal cell culture came with the discovery of what is known as the Hayflick limit. It was common knowledge that tissues or cells can be cultured and transferred serially for only a limited period, after which the culture died out, in spite of no further alterations in the culture conditions. Hayflick and Moorhead observed that each species exhibits a limit in the number of cell divisions that their cells undergo from the start point. In man, this number, called the Hayflick limit, is around 50. Cells taken from an older person reach the limit of extinction earlier than the cells from younger individuals. Hayflick and Moorhead demonstrated this in a very dramatic manner. They also demonstrated that the Hayflick limit is species-specific.

The Hayflick limit prevents indefinite subculturing of animal cells, unlike those of plants. In practice, a primary culture (begun from a tissue from an individual) of human cells rarely survives beyond two dozen transfers. This is a serious difficulty for any study which requires replicas of the original, possibly unique, tissue.

It was found, however, that occasionally a few cells become released from whatever control that determines a regulated cell-division schedule and become what are known as 'immortal' cells. These cells do not appear to observe any Hayflick limit and can be subcultured indefinitely. Clones of such cells have been derived by different investigators from a variety of animal and human sources. Each of these clones constitutes a 'cell line' and have been assigned labels to distinguish them. Cell line cells are analogous to stocks of bacteria and microorganisms having identical genetic qualities. Cell line cells made it possible to undertake long term studies, and provided a means of comparing data from different studies on a common basis. Hayflick and Moorhead were responsible for giving us a method for obtaining 'immortal' cell lines from a primary culture.

Initially, cell lines were obtained from easily grown tumour tissue. Later, run-away 'immortal' cells were collected from primary normal tissue cultures. These latter cells, like tumour cells, exhibited the capacity for uncontrolled growth and/or a dedifferentiation to an embryonic/tumour-type mass of cells. Methods were developed to select immortal cells in which cell-replication was somewhat under control, and in which the cells continued to remain differentiated. That is, the cells still synthesized the proteins they did in the primary tissue. These 'almost normal' cells became models for molecular and other studies of normal animal cells. These developments came in the wake of knowledge about the roles of various hormones in growth and differentiation.

The cycle of events between two mitoses was studied using incorporation by cells, of radioactively labelled (^{3}H or tritium) thymidine (one of the components of DNA). Such studies required large populations of cells at the same stage of the cell cycle—that is of synchronously growing cells. It became possible to induce synchrony by first blocking the cell cycle with a rate-limiting component for DNA synthesis (usually a nucleotide) or by a drastic change in temperature and then releasing the block after a while. Cells at different stages of the cell cycle proceed to develop until they halt at a certain stage when further progress is prevented by the 'block'. The release of the block results in a large population of cells, proceeding in unison through the remainder of the cell cycle.

The study of the cytology of animal cells, particularly of human cells, was facilitated by the technique of chromosome preparation described in 1956 by Tjio and Levan. They used a mitotic poison, an alkaloid called colchicine, to arrest cell mitosis at metaphase. Large populations of cells were thus made available at the metaphase stage, when the chromosomes are maximally condensed and, therefore, easily distinguished and characterized. The colchicine treated cells were killed and fixed, and then placed in a hypotonic KCl solution in which they became bloated. The swelling helps to separate the metaphase chromosomes, so that when the cell is dropped from a height on the microscope slide, the entire chromosome complement spreads out. Examination of the chromosomes was first limited to gross morphology. The chromosomes (actually their photographs or camera lucida made drawings) are arranged with their centromeres on one line, and with the shorter arm above the line. The members of each homologous pair are placed side by side. This 'typing' of the contents of the 'karyon' (nucleus) is known as karyotyping. This was the first step in the analysis of animal chromosomes from cells grown *in vitro.* Finer analysis of each arm of each chromosome became possible when better methods of staining were developed, which showed up bands of deeply stained regions and interbands of unstained ones. The pattern of banding was found to be unique for each chromosome in any one species. It was now possible to detect anomalies even in small regions of a chromosome.

McKusick's group became involved in characterizing the bands on human chromosomes and mapping known genes along them. A system of nomenclature was adopted for specifying the geographical location of parts of the chromosome. The small arm above the centromeric line of a karyotype is called the 'p' and the arm below the line, the 'q' arm. Each arm is divided and subdivided into regions that are numbered. Using this nomenclature, a defect in any part of the chromosome complement can be indicated unambiguously.

Further characterization of each chromosome involves pinpointing of a wanted region in the chromosome with the help of a suitable probe. Usually, radioactively labelled mRNA or cDNA is used as a probe. The cell is fixed on a microscope slide *in situ,* the DNA allowed to dissociate by a mild alkali, and the probe added to the preparation. The probe hybridizes with the complementary DNA; the superfluous probe is removed by enzyme digestion and washing, and an autoradiograph made of the preparation. Exposed grains over the chromosome indicate the region that was being searched.

Cell culture and allied techniques were the foundations for studies of animal cell genetics. A need in all genetic studies, especially for single-cell ones, is the availability of suitable genetic markers. Mutations of cloned animal cells were, therefore, identified that could serve as markers. One such mutation belongs to the class known as conditional mutants. These are the *temperature sensitive*(*ts*) mutants. The *ts* mutants are functional at one temperature but not at another. They are thus different from mutants where the defect results in a defective protein due to the loss or alteration of a gene. In *ts* mutants one can find out about the activities represented by the mutant locus by merely shifting the cells to a temperature at which the activity is able to occur. Most information about the functions of crucial enzymes (such as polymerases) has been derived from studies of *ts* mutants.

Another useful series of cell lines that have come into existence are those of the nutritional mutants of Chinese Hamster Ovary (CHO) cells. These were developed in the laboratory of Kao and Puck and reported in 1968. These, and the discoveries mentioned above, laid the foundation for somatic mammalian cell genetics.

Inevitably, the development of cell culture and other techniques were exploited for investigations of virus-caused and oncogene-mediated transformation.

8.2.1 Requirements for Cultivation of Animal Cells

Most animal cells require a rich growth medium, a solid support on which to anchor and a pH maintained at 7.0–7.3. Some cells can grow as suspensions in liquid media. Certain others that normally require a solid substrate can now be cultured as suspensions.

The type of culture and the aims of the study or project dictate the type of vessel to be used for culturing the cells. The containers range in variety from small tubes, hollow fibres, petri dishes, bottles, arrays of miniature culture wells to bioreactors, suitable for batch or continuous cultures, used by the industries.

The pH of the medium is maintained best by incubation of the cultures in an atmosphere of 5 per cent CO_2 in air. This is best achieved by sophisticated CO_2 incubators. Small batches of cultures may, however, be maintained in improvised CO_2 chambers made from desiccators. In addition, the medium is buffered.

Another important need for *in vitro* culturing of animal cells is a high degree of aseptic conditions. Animal cells are particularly susceptible to viral, yeast and mycoplasma infections which may spread like wild fire throughout the laboratory. Once infected, such cultures can rarely be salvaged. Stringent steps are, therefore, taken to ensure prevention of such contamination.

1. Media and Growth Factors

The culture medium should possess at least the following components:

(i) An energy source, such as glucose
(ii) A balanced mixture of inorganic salts
(iii) Essential amino acids
(iv) Inositol, choline and the B group vitamins
(v) Growth factors, usually in the form of serum
(vi) A buffering agent, and a pH indicator
(vii) Distilled and deionized water.

Let us first consider the last item in the above list, i.e., water. Most laboratory distilled waters contain trace metals that are highly toxic to animal cells. It is, therefore, imperative that deionized water is used to make the media, and also for washing of the labware to be used for the culturing operations. Glassware cleaned with a non-ionic detergent (e.g., Teepol) or with nitric or hydrochloric acid, should be rinsed thoroughly with deionized water, and then heat sterilized in ovens at 180°C.

A good buffering device is required as the pH (of about 7) for the medium is critical for animal cell culture. The commonest buffering agent is sodium bicarbonate. Its buffering efficiency, however, depends on varying factors, and is, therefore, not suitable for all types of cells and/or media. Organic buffers, one of which is the commercially available HEPES (N-2-hydroxyethylpiperazine-N′2-ethanesulphonic acid) are better, though far more costly than sodium bicarbonate. More about this will be discussed in Section 4 of this chapter.

Several formulae for basal medium incorporating the minimal requirements have been developed by Eagle and Hanks. Modifications have been introduced by others including McCoy and Dulbecco. The media may be concocted in the laboratory from the primary constituents or reconstituted from commercially available liquid concentrates or dry powders. Eagle's Minimum Essential Medium (MEM) and Basal Medium Eagle (BME) and modifications such as Dulbecco's Modification of MEM (DME) are standard supplies. In most of these prepared media, the sodium bicarbonate is added when the media are made. The pH is monitored by the colour imparted to the solution by a pH indicator, such as phenol red. Phenol red is already present in the medium powder.

The minerals and amino acids are sufficient to keep the cells alive for long periods of time. To make it possible for the cells to grow, proliferate and/or differentiate, certain additional factors are required. These include amino acids, vitamins (especially of the B group), choline, inositol and growth factors.

However, the growth factor requirements have not been fully defined. This lacuna in knowledge is counteracted by adding serum to the medium. Serum is the fluid remaining after blood cells and other inclusions are strained away from clotted whole blood. The blood carries all the physiological components to all cells of the body. So the serum from any animal is likely to possess a cocktail of the growth factors. Perhaps all the components of growth present in the serum are not required for every type of cell. However, until the serum is characterized completely and the individual requirements of each type of cell identified, animal tissue culture protocols would continue to depend on serum to supply the unknown (to us) but essential growth factors.

Although it has been apparent for long that animal cells cultured *in vitro* undergo mitoses only when serum is present in the medium, it is only recently that some of the reasons for this have become evident. Blood consists of a basic interstitial fluid which contains soluble matter and cellular inclusions. The fluid that passes out of capillaries to supply nutrients to cells (the interstitial one) does not possess factors for stimulating cell division. At points of injury, where blood vessels have ruptured, the flow of blood is stopped by the formation of a clot. A clot consists of a network of proteins converted from precursers in whole blood, the conversion mediated by certain factors. Hemopoietic cells, including the fragmented non-nucleated ones known as platelets are caught in the protein mesh. In the process, platelets provide certain mitogens that stimulate the otherwise non-dividing cells at the site of the wound to proliferate and repair the injured region. It was discovered that these proteins are the growth factors PDGF (platelet derived growth factor), EGF (epidermal growth factor) and TGF (transforming growth factor).

Now TGF is a factor that makes normal cells anchorage independent. This is also a feature of transformed animal cells—hence the term 'transforming' growth factor. Similarly, PDGF has been identified with the product of a virus-borne oncogene —the v-*sis* (Simian sarcoma) protein, which transforms animal host cells. EGF has also been found to have homologies with another viral gene product involved in cell transformation. In all these instances, these growth factors instigate cell division *in vivo.* It is no wonder then that serum has been empirically discovered to be an essential growth and proliferation inducer in cells cultured *in vitro.*

Most cells *in vitro* can be stimulated to divide by adding the above three growth factors. Some cells that rarely divide in adult tissues are unaffected by these factors; they have other specific growth promoters. Nerve cells, for instance, also require NGF (nerve growth factor); endothelial cells, found only in linings of blood vessels, are not renewed once formed. They proliferate only where new blood vessels are developing (as in tumours, placenta and so on). The specific factors that trigger endothelial cell proliferation are ECGFs (endothelial cell growth factors), FGFs (fibroblast growth factors) and a ribonuclease ressembling, but non enzymatic, protein called angiogenin (because it stimulates angiogenesis or formation of blood vessels). This brief discussion of the growth factors will highlight the importance of serum in animal cell cultures. As more specific requirements are discovered, serum free media will come into routine use. This will greatly reduce the non-standard conditions which one faces with serum from different lots, animals, strains and so on. Further, serum may harbour viruses that escape detection.

It is customary to use calf, bovine, porcine, horse or goat serum for different projects. For culturing human lymphocytes, autologous AB serum may be used. The serum from calf fetus is preferable, as it does not contain γ-globulins (as carried by adult animal serum), which inhibits the growth of many animal cell cultures. Fetal calf serum is, however, extremely expensive. Where fetal calf serum is found to be essential, protocols are modified, so that microquantities of media are utilized in small culture vessels. In other cases, serum from adult animals is used after ridding it of the γ-globulins.

Of course, the best solution for removing the unknown quantity represented by serum is to supply the exact growth promoter(s) needed by each type of cell.

This has been attempted in several laboratories. There are reports of the successful use of EGF, NGF, FGF in special circumstances. Various hormones, separately and in combinations, have also been used. There are a few reports of completely defined media that show some promise.

Animal cells also appear to modify their surrounding medium by contributing substances synthesized by them. This 'conditioning' of the media has to be remembered when replacing the old, used medium with a fresh one. Usually the cells transferred to the new medium start to proliferate only after a slight lag period. It is assumed that the medium has to be conditioned by substances secreted by the transferred cells, before further differentiation can occur.

Cells *in vivo* rest on extra cellular matrices, the composition of which is often determined by products secreted from the cells themselves. It has been observed that unless receptors on cell-membranes become engaged to one or more of these self-supplied substances, the cells do not differentiate. Another observation is that serum contains substances that inhibit action of protein-degrading enzymes (proteases). Developers of serum-free media, therefore, try to incorporate protease inhibitors and proteins and polyamines known to be present in basal membrane molecules—such as Substrate Adhesion Molecules (SAMs) represented by fibronectin, transferrin and laminins.

Proliferation in basal media, even those with added serum, or serum substitutes, is not enough to induce cultured animal cells to undergo mitoses. Usually a mitogen has to be supplied, especially for cells that grow in suspension, such as lymphocytes. Mitogens include antigens, growth factors, certain hormones and complex carbohydrates called lectins. The lectin PHA triggers human T cells in culture to proliferate. Pokeweed mitogen induces mitoses in cultured human B cells.

Animal cell culture media have to be sterilized by filteration and not by autoclaving, since many of its components are heat-labile. A Millipore-vacuum pump-filtering assembly is best for this purpose. Prepared media usually come in quantities that are to be reconstituted into one litre of liquid medium. This is filter-sterilized in batches of 250 ml aliquots. The first aliquot may be used immediately. The remaining aliquots may be stored in dark-coloured bottles in the freezer. Just before use, the medium is thawed, allowed to attain room temperature (25°C) and re-filter-sterilized. As a further precaution, antibiotic solutions (a mixture of penicillin and streptomycin) are added to the freshly prepared culture mix (medium + serum + cells).

2. Substrate

Animal cells, with a few exceptions, thrive best and differentiate only when they are anchored to a solid support. Perhaps the support takes the place of some of the functions provided by the extra cellular matrix or the basal membrane present *in vivo.* The exceptions include blood tissue cells and ascites tumours that develop in rodents; these grow best in suspensions in liquid media.

There are three categories of solid supports that are used for *in vitro* culture of anchorage-dependent animal cells. These are (i) glass or plastic surfaces, both treated and untreated for maximum cell adhesion, (ii) semi-solid surfaces provided by agar or agarose gels, and (iii) solid surfaces with a monolayer of feeder cells.

(i) Rigid substrates Glass surfaces are best for animal cell culture for the following reasons: (a) glass surfaces are naturally charged, a quality required for cell-anchorage. (b) they may be reused several times without loss of optical properties; animal cells spread out on the substrate and have to be viewed through the glass container using an inverted microscope, (c) they are cheaper in the long run, if facilities for cleaning and sterilization are not excessively expensive.

Plastic surfaces are also utilized—especially factory-sterilized disposable ones. Some synthetics are hydrophobic by nature: they are treated to make them hydrophilic before use. The cheapest synthetic in use is polystyrene. Others, including teflon, are also usable. Some plastics used for flasks and petri dishes, are permeable to gases, and may be useful for particular projects. Plastic materials, however, lose their optical properties after repeated use. In projects where large volumes of culture vessels with guaranteed sterility are required, synthetic, disposable, pre-sterilized culture vessels are the best choice.

A glass or plastic surface treated with one or more compounds, found in the extracellular matrix, *in vivo,* usually provides better anchorge to cells than the untreated surfaces. In certain cases, the cells do not proliferate unless such a treated surface is provided for them. Denatured collagen, in solution, has been used for this purpose. A culture flask or petri dish is rinsed with the solution, air-dried and sterilized using a UV lamp. Epithelial cells have been found to do well on such collagen-treated surfaces, especially if another SAM, laminin, is added to the medium. Similarly, gelatine treatment is beneficial for culture of muscle cells.

Instead of coating the surface with the above type of proteins, the surface may be made ready for cell adherence and growth by first growing a layer of some other cell that secretes extra cellular matrix proteins. In this case, the conditioning cells are removed by the action of a detergent. The cells together with the detergent solution are removed, leaving traces of cell-secreted proteins on the culture surface.

(ii) Semi-solid surface Cells can be grown on semi-solid media, made by the addition of a gelling substance such as agar or agarose. As daughter cells tend to remain adjacent to the parent ones, separate clones of cells are best identified by culturing on such a surface.

Semi-solid substrates are also useful for screening of virally—or otherwise transformed cells from the non-transformed ones. The latter usually fail to adhere and form colonies, unlike the former. Clones of the transformed cells can thus be easily recognized. This is especially useful where only very few cells in a large population become transformed.

(iii) Substrates coated with live feeder cells Some types of cells appear to differentiate well only when grown over a layer of some other type of cell. The latter not only provide anchorage material, but also supplement in some way the growth-promoting qualities of the prepared media. Hence, these help-supplying cells are referred to as feeder cells.

Feeder cells from various tissues are in use: Rodent embryo cells, intestine epithelial cells and glial cells of the nervous system are utilized as feeders for a variety of cultured cells. Some types of cells, such as neurones differentiate well only when overlaid on a feeder layer of glial cells.

3. Types of Culture Vessels

Cells grown in suspension may be cultured in small, screw-capped vials for short term projects or in special spinner bottles where the medium is stirred at a very low speed by an in-built magnetic stirrer (Fig. 8.5).

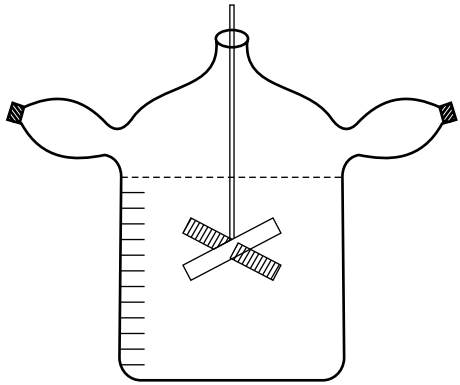


Fig. 8.5 *A spinner bottle.*

For cells to be grown on soft-surfaces or over-feeder layers, petri dishes are ideal. Sometimes glass slides or coverslips are used for attaching the cells, and the former kept submerged in medium in a petri dish. The coverslips may be removed and viewed under an upright microscope.

The commonest vessel used for culture of anchorage dependent cells is the T-flask, which is a flat-sided, prescription bottle-type of container with gently sloping shoulders. The cells spread on the inner surface of the flat side that is at the bottom. Several such flasks can be stacked over each other in the incubator. The area of the surface for cell attachment ranges from 25 cm^2 to 175 cm^2 and is able to contain 5 to 100 ml of culture media respectively. Small volumes of media and small cell populations are cultured in special small tubes with one side flattened and extended like a split level floor. Coverslip cultures may be inserted in this sunken chamber, and removed for examination. Such special tubes are generally called Leighton tubes. They can accommodate cells on an area of about 4 cm requiring only 1.0 ml of the medium (Fig. 8.6).

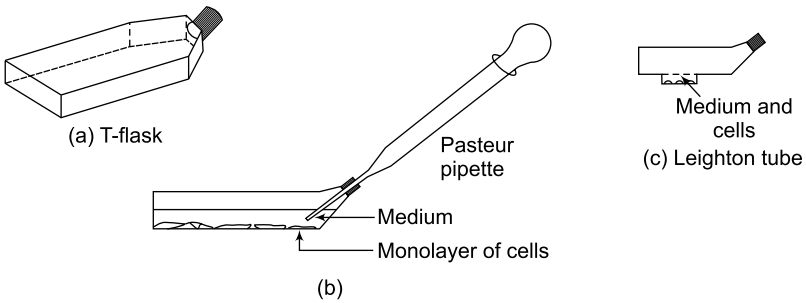


Fig. 8.6 *(a) T-flask, (b) Pasteur pipette, and (c) Leightor tubes used for animal cell culture. Petri dishes may also be used.*

To maximize the surface on which cells can attach, some cultures are grown in bottles with the usual round cross-sections. These are, however, rotated steadily at a slow speed, so that cells can spread on the entire inner surface. The rotation also bathes the cells constantly in the medium. A rotator, the speed of rotation of which can be set at desired values, is used to rotate several bottles at a time.

Sometimes it is required to culture several replicates for a variety of manoeuvres and observations. For this, multiwell plates and a range of microtitration plates are available. (Fig. 8.7). These are like miniature egg holders, with a few large or several smaller wells or depressions, each of which is used for a culture. Very small wells, or rings fixed on a flat plate are usually used for microtitration and other microtests. Some cell cultures grow best in a small volume of conditioned medium. Somatic cell fusions are also carried out best in small volumes of media. Multiwell plates with 72 or 96 wells are available. These are also used when replicates have to be analyzed in different ways.

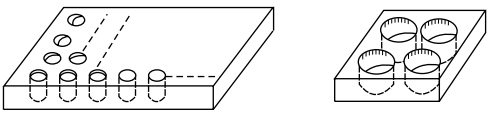


Fig. 8.7 *Multiwell plastic culture plates come in different sizes and with different numbers of wells per plate.*

Ordinary soda glass prescription bottles may be used for routine work as they are cheap and as they do not usually contain lead that is found in the more expensive heatproof pyrex glasses.

Commercial production of gene-products by animal cells makes use of bioreactors with spherical synthetic beads as surfaces for cell anchorage. In this way, both ease of factory level operations and a biological culture requirement (anchorage) of the cells are satisfied.

4. CO_2/O_2 — or Gas Phase

Unlike plant cells, animal cells have to be provided with a medium with a pH that strictly stays within bounds. Cells in culture release acids and CO_2, the latter forming weak carbonic acid in solution. The culture medium thus faces the danger of becoming progressively acidic. To offset this, a sodium bicarbonate buffer is used generally. The CO_2 and hydroxyl ions, formed by dissociation of the salt, buffers the solution against the increasing acid inputs. In the event that the cell metabolises and secretes acids at a rate lower than the CO_2 released by the sodium bicarbonate, the medium would turn alkaline. Excess CO_2 in the air interfacing with the medium (referred to as the 'gas phase') would prevent too rapid a dissociation of the buffer salt. For small volumes of short term cultures, a closed tube or a petri dish accumulates enough CO_2 from the cells to provide the 5% CO_2 blanket needed to control the dissociation of the buffer.

However for larger volumes and long-term cultures that require a standard composition of the gas phase, a CO_2-incubator is to be employed. Such an incubator may be (i) improvised out of a desiccator containing a vial with sodium bicarbonate and HCl that releases CO_2 within the vessel, (ii) a small perspex-walled box, fabricated or store-brought, that can be purged with CO_2 from a cylinder, or (iii) a sophisticated equipment, such as a Hereus CO_2 incubator, in which all the culture parameters can be met steadily. The levels of CO_2, O_2, air, humidity and temperature can be set as desired. The incubator can be self-sterilized. The latest

models allow the use of small compartments within the large incubator, without disturbing the environmental conditions in the remaining regions of the incubator.

Almost all types of animal cells do well in normal atmospheric levels of O_2. Some types, such as certain embryonic cells, may require more than 90 per cent of O_2 in the gaseous phase. Such special and constant conditions can again be provided by the sophisticated CO_2 incubator describe above.

5. Avoiding Contamination

Animal cells in culture are vulnerable to viral, bacterial, yeast, fungal, protist and mycoplasmal infections. All of these, except the last one, can be detected by the naked eye and characterized under the light microscope, and standard steps taken immediately to eradicate the contamination from the laboratory. Mycoplasma infections are not readily detected and much damage can incur before there is any suspicion of contamination.

Mycoplasmas are too small to be seen even under very high magnification. Scanning electron microscope photographs show them up as tiny beads on the surfaces of infected cells. Mycoplasmas are prokaryotes without any cell wall surrounding the plasma membrane and they pass through filters having 22 μm pore size.

Mycoplasmal infections do not appear to alter the morphological features of cultured cells. But they create havoc in studies based on cultures of animal cells. For instance, if antibodies are prepared against particular fractions of a cell, there is always the danger of the antibodies being raised against the mycoplasmal components and not against those of the cell. Similarly, if cell studies require the use of tracking tags, the labelled nucleotides may be incorporated in mycoplasmal DNA, which metabolises faster than the cell DNA.

The best way to keep out the mycoplasmal infections is to make periodic checks of culture materials in the laboratory. The DNA of mycoplasma may be stained with the usual DNA stains, Feulgen or Giemsa, or with a fluorescent dye. There are stains that show up the nuclear DNA of the host cell and specks of DNA in the cytoplasm due to staining of the mycoplasmal DNA.

8.2.2 Strategy for Culture

The cells to be cultured are added to the buffered medium in the appropriate vessel to which the antibiotics penicillin and streptomycin are also added. The cells are incubated at about 37°C for the required length of time. For suspension cultures, the period of incubation is determined by the density of cells in the medium, as assessed by cell counts and spectro-photometric measurements of turbidity of the suspension. In the case of cells that attach and spread as monolayer on a solid support, the period of incubation is the time taken to form a confluent single layer of cells on the available surface. At the end of this period, aliquots of the suspension cultures are removed, centrifuged to pellet the cells, fresh medium added, and the latter innoculated into fresh culture vials or bottles.

For cells grown on solid glass or plastic surfaces, at the end of the period of formation of confluent monolayer covering the entire available growth surface, the

medium is removed using a pipette. A solution of trypsin in buffered salt solution is added to the culture vessel, and allowed to incubate for a short time. Trypsinization removes the proteinaceous material that sticks the cells to the solid surface. The trypsin is replaced by a fresh medium. This stops further proteolysis by trypsin. The cells become suspended in the medium. A small amount of the cell suspension is pipetted out and inoculated into a fresh culture vessel. The latter is again incubated at 37°C until the surface is covered with one layer of cells. In this manner, cells are transferred at regular intervals.

Cells that grow anchored on glass or flask surfaces have to be viewed with a special microscope capable of viewing the cells through the bottom surface of the flask. The light from the material passes through the glass or plastic wall of the flask and into objectives placed below the microscope stage. The image can be pictured through eyepieces in the normal position. Such inverted microscopes are a must for routine work involving animal cell culture. T-flasks are expensive since the glass or plastic has to be of microscope slide quality, through which the cells may be viewed. Both simple and sophisticated inverted microscopes are available in the market. The latter allow culture dishes, plates and flasks of varying sizes to be accommodated on the stage, and also possess attachments for phase-contrast, fluorescent microscopy and still or cinematic photography.

In the absence of an inverted microscope, a quick look can be had of the attached cells by inverting the T-flask fleetingly, and placing it under the objective of a standard upright microscope. Keeping the culture vessel for a prolonged period in this position will dry out the cells, and so should be avoided.

The number of times cells may be subcultured depends on the nature of the cultures. Primary and secondary cultures, derived from an explant or directly from an organ, degenerate after a limited number of subculturings. Some cells in primary cultures, however, overcome some 'crisis', as yet not understood, and fail to observe the specific Hayflick unit. These cells form clones that may be subcultured indefinitely. Such 'immortal' cells can be typed on the basis of selectable or screenable markers and maintained as stocks of continuous cell lines.

8.2.3 Primary Cultures

Cell cultures originating from explants and organs are known as primary cultures. Primary cultures may be started in either of the following two ways:

(i) The material is chopped into small pieces, the latter mixed with a serum-enriched medium and seeded on the surface of a coverslip or a small T-flask (Exercise 46).

(ii) The material is disintegrated by removing the proteinaceous and mucopolysaccharide intercellular binding material before mixing with a medium and seeding on coverslips or flasks.

Most animal explants are heterogeneous mixtures of several types of cells. Each of these, or only some of them, may initiate primary cultures. When a chopped material is the source of the culture, the cells migrate along the glass (or plastic) surface and generally form aggregates of each type of cell. When the disintegration method is employed, not all cells remain equally capable of adhering to the support

surface. There is thus a selection of particular cell types that form primary cultures, irrespective of the procedure followed for obtaining them. Special selection methods are available for choosing a particular type of cell for subsequent culturing.

The disintegration method yields more cells for culture than the chopping one; hence one has to use a larger explant if the latter procedure is to be utilized.

The explant or organ is dissected out and first washed in a standard balanced salt solution (BSS). The material is then transferred to a sterile petri dish in which it is chopped (finely or not so finely, depending on the procedure to be followed). The finely chopped material is strained through a fine-sieve and the BSS with the sieved material is centrifuged to concentrate the latter. After a few washings with the BSS, the material is mixed with a very small amount of the medium to be used. Aliquots of the tissue-medium mixture are placed on coverslips or small (25 cm^2) T-flasks and incubated for about a day at 37°C. More medium is added gradually to the flasks or petri dishes (with coverslips) and incubation continued for a few (4–5) days. By this time about half the surface of the 25 cm^2 flask should be covered with cells. At this point a subculture should be started. Transferring cells from one culture to start another one involves certain standard steps, that are collectively known as 'passaging' of cells.

If, on the other hand, disintegration of the explant is desired, a proteolytic enzyme or a chelator such as EDTA is utilized. The most common enzyme used for this purpose is trypsin. Trypsin, however, affects cells adversely if the latter are treated for too long; so methods are available for minimizing the undesirable effects of the enzyme. The explant is washed in BSS and chopped into small pieces and trypsin added to the mass. The treatment should not exceed 30 minutes at 37°C. After trypsin-treatment the cells are centrifuged out into a pellet and the action of trypsin stopped by adding the medium containing fetal bovine serum and keeping the tube at 0°C (on ice). The cells are pelleted out again and treated a second time with trypsin. This procedure is repeated a few times, until the material appears to have completely disintegrated. The cells are finally pelleted out and added to a fresh serum-enriched medium and cultured in small T-flasks. After a few days when the flask surface appears to be fairly covered with cells, the latter are 'passaged'.

Passaging of cells consists of the following operations (see Exercise 47):

(i) The medium in the flask is pipetted out and discarded.
(ii) A trypsin solution is added to the monolayer of cells.
(iii) After about 30 seconds the trypsin is pipetted out without disturbing the cells, which are then allowed to remain at 37°C for a short time ($\simeq$ 15 minutes). The cells can be seen to detach themselves from the surface and become spherical. Fresh medium is added. The serum in the medium stops the action of trypsin.
(iv) The cells are suspended in the medium and disaggregated with a pipette into single cells.
(v) A calculated aliquot of the cell suspension is then introduced into a fresh flask containing a measured amount of the fresh medium. Passaging is completed.
(vi) The flask is then incubated at 37°C as before.

Primary cultures of different cell types may be selected and maintained as stocks, by using the appropriate selection methods. Epidermal, melanocyte, embryonic and fibroblast cells are some of the more frequently used cell types for animal cell studies.

1. Fibroblast Cells

It is generally observed that out of the mixed population of cells present in an explant, one particular type becomes the most abundant one in the ensuing primary cultures. These cells are spindle-shaped and strongly resemble the fibrous cells of the connective tissues. Hence these cultured cells are referred to as *fibroblasts.* Majority of general animal cell studies use fibroblasts as the test material.

Although there are certain similarities between fibroblast and connective tissue cells, it is not certain that they are identical. Fibroblasts are known to differentiate into other cell types and fibroblasts from explants of different species may not be quite from the same origin. In any case, these cells are the most hardy of cultured cells and can tolerate experimentally caused insults more than the other cultured cell types.

As the origin of fibroblast cells is not always clear, long, spindly cells in culture are generally dubbed as fibroblasts. Similarly, polygonal type of cells, unless their origin is clear, are referred to as epithelial. If you are a stickler, you may call them fibroblast-like or epithelial-like cells.

2. Epithelial Cells

The next most prolific types of cells emerging from explants are epithelial cells. They can be isolated or selected from the overwhelming population of fibroblasts by manipulating the culture conditions. Serum promotes a greater growth of fibroblasts than of the epithelial cells. Hence isolation of the latter is best carried out in a serum-free media to which growth factor supplements have been added.

In addition, instead of trypsin, another proteolytic enzyme, collagenase, may be used with better success. This enzyme degrades the collagen fibres in the stroma adjacent to the epithelial cells.

Another way of selecting epithelial cells against fibroblasts is to add d-valine to the medium. Fibroblast cells cannot use d-valine and die out, leaving the epithelial cells which exploit this amino acid.

There is much interest in culturing epithelial cells, as they play crucial roles in the function of most organs. Cancers of various types also originate from epithelial tissues. Furthermore, these cells serve as good models for the study of cell differentiation.

Epithelial cells can be isolated from different tissues including the epidermis, milk, liver, cervix, pancreas, bronchi, and gastrointestinal walls. The epithelial cells grow well on feeder layers of irradiated NIH3T3 (a mouse cell line) or fetal human intestine cells, on collagen-coated flasks or on floating sheets of collagen.

By appropriate modification of the general explant culture method, cultured epithelial cells may be observed undergoing differentiation. Differentiation of keratinocytes from epidermal tissue has been a favourite research problem.

3. Nervous Tissue Cells

Neurones are more difficult to grow in culture than the glial cells. The fastidiousness of neurones for culture conditions has prevented culturing of neuronal cells that undergo mitosis. Hence most nerve-cell cultures are derived from neoplastic tissues, such as neuroblasts.

Neurones grow only on treated surfaces; they thrive on collagen covered or poly-L-lysine treated substrates. Differentiation of neurones, such as dendrite formation, is dependent on nerve growth factors (NGF).

Glial cells are easier to culture, which may be started by, mechanical, trypsin or collagenase-based maceration of embryonic avian or mammalian brains. Cholera toxin is a suitable mitogen for inducing proliferation of glial cells.

4. Embryonic Cells

Embryos are the easiest starting material for obtaining primary cultures. They can be macerated with ease, and they double the cell populations in a shorter time than the cells from neonatal or adult tissues.

Rodent and chicken embryos have been in use most for preparing primary cultures of cells. The hen's egg contains the growing embryo and the yolk paraphernelia needed for the growth of the embryo. One has to take an egg of the proper age, remove the shell from the blunt end, peel away the underlying opaque shell membrane and pick up the visible embryo with a pair of forceps. This is then chopped, trypsinized and seeded in the appropriate medium. Mouse embryos are dissected out when they are large enough (usually 12th or 13th day of embryogenesis), and used for deriving primary cultures.

If specific cells are to be cultured, the parent tissue is dissected from the embryo and treated for obtaining primary cultures.

5. Bone Marrow Cells

Bone marrow contains the stem cells of the hemopoietic system (blood and lymph). As bone marrow precursor cells are usually undergoing mitoses, additional mitogen may not be required for their proliferation in short term cultures.

It has not been easy to develop methods for culturing hemopoietic cells. Usually the mature cells have a finite life-span (such as erythrocytes) and do not undergo mitosis after maturation. Culturing became possible with the availability of information about the biology of some of these cells. It was discovered that there are specific growth stimulating factors that induce cell-division/differentiation of stem cells in culture, and also of small lymphocytes that *in vivo* differentiate into T or B cells. These factors include the colony-stimulating factors (CSFs), the B-cell stimulating factors, those that trigger T-cell precursors to differentiate and activate the subset families of T-cells such as interleukins (e.g., IL-2), antigens, and mitogens.

Bone marrow cells are usually cultivated in suspension; there are also methods available to grow them on semi-solid media. To obtain bone marrow cells for culture, they are flushed from femur bones, into the medium in a small T-flask by aspirating with a large-bore needle. The method consists of cutting off the ends of a femur (mouse), placing one end inside the medium, and fitting a needle-syringe assembly at the other. Aspirating with the syringe forces the marrow cells to pass out of the bone into the medium. The cells attach themselves to the surface of the T-flask (placed on one flat side). These cells release new cells at varying stages of maturity into the liquid medium. The suspended cells are then subcultured or studied. The differentiation of lymphocytes, macrophages, RBCs and other granulocytes may be observed using the appropriate culture strategies.

Hemopoietic cells and their differentiation are of great interest to almost all the molecular biologists, developmental biologists and applied biologists, as they provide useful models for investigating the parameters that regulate gene expression and eventually development.

Study of hemopoietic cells was accelerated after several continuous cell lines became available. In some cases, it is possible to induce the formation of such cell lines by infection with a transforming virus such as the Epstein-Barr virus and Friends' virus. EBV transforms both T and B cells. Friend's virus (with an RNA gene) induce erythroid cell precursors to become neoplastic (erythroleukemic). In addition, cancers of blood cells have yielded myeloid cell lines, some of which synthesize complete H and L chains of IgA and IgG. These are exploited for the development of hybridoma cells—cells that produce monoclonal antibodies, and proliferate due to their cancerous characteristics. The above cell lines are very useful for a variety of studies, especially those concerned with differentiation.

Macrophages may be selected from bone marrow cells by trypsin treatment. Proteolytic enzymes do not affect macrophages, which, therefore, remain attached to the substrate, while other cells become detached from the substrate. Another way to obtain large numbers of macrophages is to wash out mouse peritoneum that has previously been subjected (injected with) to mineral oil or some other irritant.

In any case, conditions have not yet been discovered that allow normal macrophages to undergo mitoses. Most long-term studies are, therefore, based on cell lines derived initially from cancerous mouse tissues.

6. Lymphocytes

Lymphocytes are also members of the hemopoietic group of cells. There are two main classes of lymphocytes, the T and B cells, the stem cells for which originate in the bone marrow.

Small lymphocytes from bone marrow migrate to the thymus, where differentiation furnishes different T-cell precursors, with surface antigens and other characteristics, that convert them to T_H, T_C, T_S cells, and perhaps also to the NK cells. Similarly, B-cell precursors are differentiated into B lymphocytes in an, as yet, unknown site in mammals. In birds this event occurs in the Bursa of Fabricus, a gland attached to the intestine and hence the name B cell.

Both T and B are small lymphocytes that undergo dramatic changes due to differentiation, when challenged by certain factors. These include foreign antigens and mitogens.

Lymphocytes are usually cultured for a short term, during which they are provoked into cell division by an added mitogen. The cells in mitoses are arrested at the metaphase stage with a spindle-poison. This prevents the mitosis to be completed and so results in a cell in which the metaphase chromosomes are dispersed in the cytoplasm. The chromosomes can then be studied for anomalies and aberrations. Metaphase arrest also pools a large number of cells at this stage. All the cells in a population are not in synchronization in terms of their cell cycles; some reach metaphase earlier than others. By arresting the cells at metaphase, the ones that have reached the stage are prevented from proceeding into anaphase, telophase and into formation of new daughter cells. Cells in earlier stages of mitosis proceed until

stopped at metaphase. If cells are allowed to proceed to complete mitosis, a very small population will be available at any time for scoring metaphases.

The spindle-poison generally used is colchicine or a derivative, colcemid. The mitogen used in lymphocyte cultures is PHA (for T cell proliferation) and PWM (pokeweed mitogen, for B cell mitoses).

Lymphocytes for culture are generally derived from peripheral blood (see Exercise 48). Blood taken from a vein and collected in a vial, containing an inhibitor to prevent clotting of blood (heparin), is mixed with culture medium, serum, mitogen and antibiotics, and dispersed in several screw-capped vials. The latter are incubated at 37°C for 24, 36, 48, and 72 hours. About three hours before termination of the incubation period, colcemid is added to each vial. The cells are harvested, treated with a hypotonic KCl solution and spread on microscope slides. The swollen cells burst and the dispersed and condensed metaphase chromosomes stick to the glass slide. They are air-dried and stored. The slides are stained by a flat chromosome stain such as Feulgen or acetocarmine, or treated for banding, using the appropriate protocols (see Exercise 49).

Lymphocyte cultures are useful for assessing damage to human cells by environmental mutagens. Persons exposed to unusual atmospheric or other pollutions may be tested for significant aberrations in their chromosomes, by making metaphase preparations from only 1 ml of blood.

Lymphocytes are also used to assess the immunocompetence of a subject. The percentage of cells that can be induced to respond to an antigen or mitogen gives an idea about the ability of the subject to withstand environmental insults.

8.2.4 Culture of Transformed Cells and Established Cell Lines

When we use the term 'transformed' in connection with the mammalian cells, we refer to an alteration in the genetic qualities that modify the phenotype of the cell in a specific manner. The alterations may be induced by a variety of insults—chemical, physical or viral—but the new phenotype has several traits in common irrespective of the mode of induction of transformation.

In the most general usage of the term transformation (say, in bacteria) we mean a genetic change brought about by an exogenous genetic material, usually DNA. In such transformations, the type of change varies with the introduced DNA. In contrast, the type of transformation mentioned in the beginning results in, more or less, a particular phenotype, irrespective of the provocation that induced the change. This 'transformed' phenotype has at least two prominent traits. Firstly, a propensity of the cell to continue cell divisions *ad infinitum*, and secondly, several biochemical alterations in the cell that make it distinctly different from a normal or nontransformed one. The newly acquired qualities of the transformed cell include differences with regard to anchorage ability for growth, need for added growth factors, and reaction on contact with adjacent cells. Most normal cells need a solid support for anchorage. They cannot grow if serum (or growth factors) is not added to the medium. Also, the normal cells stop proliferating after they come into contact with each other. In other words, once a confluent monolayer is formed by normal cells, or the cell density reaches a certain high value in a suspension culture, there is very little, if any, further increase in cell numbers.

In contrast, the 'transformed' animal cells are anchorage-independent, serum-independent (almost), and are not inhibited by contact or high density of cells in a suspension culture. When grown on a glass slide they can pile up on top of each other and form a 'focus'. They can also be shaken off the slide. Some transformed cells can be grown on agar-media or on feeder layers.

It is now apparent that the road to malignancy is via transformed cells. Several steps are required before a cell can become tumourigenic (malignant). Tumourigenic cells have the transformed phenotype as well as the ability to proliferate indefinitely and at a faster rate than normal cells. These two components—immortalization and transformation—may occur in different sequences *in vivo,* and *in vitro.* That these are independent events can be seen from the fact that there would be cells that become immortal—that is, can be subcultured continuously—but are not tumourigenic.

It is obvious from the above brief introduction to transformation in animal cells, that the events involved are of prime interest not only to cancer specialists and therapists but also to those who would like to unravel the global features that are involved in the process of gene-regulation-cum-differentiation.

Much of the recent advances in animal cell studies should be credited to the availability of several 'immortal' cell stocks. These are known as continuous cell lines. Many such continuous cell lines have been isolated, subcultured and maintained as established cell lines. Whatever genetic change converted normal cells to a particular immortal type, is retained in the latter and propagated. So they are definite genetic variants of the normal cells.

Special 'strains' have also been delineated from continuous cell lines that possess particular 'markers', which give them a selectable or screenable identity. The following is a list of some of the established strains which are used universally as standard laboratory test materials.

	Name	*Type of cells*
1.	Klein	Human fibroblast
2.	NIH3T3	Mouse fibroblast (from embryo)
3.	L cells	Mouse fibroblast (from embryo)
4.	BHK	Baby Hamster Kidney
5.	CHO	Chinese Hamster Ovary
6.	HeLa	Human cervical carcinoma
7.	NB41A	Mouse neuroblastoma

The human and murine (mouse) fibroblast cell lines (1–3) also possess the following useful selectable markers.

Klein	TK^+ $HPRT^-$
3T3	TK^+ $HPRT^+$
L	TK^- $HPRT^+$

The Klein and 3T3 survive on HAT (hypoxanthine aminopterin-thymidine) media, while the L cells do not. So gene transfer studies in which TK^+ containing foreign DNA is introduced into L cell, select the transformed cells on HAT but not the untransformed L cells.

Most of the cell lines have originated from malignant tissues. They are presumed to be quasi-transformed. At least they are immortal and can, therefore, be propagated in culture, indefinitely. After a few or several subcultures, some of these cells may become tumourigenic. This can be ascertained by injecting cells from continuous cell lines in culture into a test animal (mouse, hamster). Tumours develop in the injected animal if the immortal cells are also transformed to be tumourigenic. Tumourigenic cells can invade the underlying tissues and metastasize (migrate to other regions of the body), causing tumour formation at new focal sites.

Tumourigenic cells may be selected from primary cultures by a variety of methods. One of them is based on cloning cells in suspension. Tumour cells form aggregates more often than normal cells. A more efficient method is to grow cells on a suitable feeder layer. Both mouse 3T3 embryonic fibroblasts and human fetal intestinal fibroblasts have been used successfully as feeder layers for the isolation of tumour cells from various carcinomas. Usually the feeder layer inhibits the growth of fibroblasts, which otherwise would overwhelm the smaller tumour cell population, if present. Another procedure is to employ a medium that is selective for tumour cells. A medium developed by Carney et. al., (1981), known as HITES, which is a modified version of the standard RPMI 1640 medium, selects for small lung cancer cells. Some other modifications of RPMI 1640 are suitable for selection of adenocarcinomas and other tumours. One medium inhibits fibroblast domination by removing them with fibroblast-specific monoclonal antibodies.

It is usually difficult to disaggregate tumour tissue for preparing primary cultures. If trypsin is used, the treatment has to be for a considerably long period of time, which does not suit the health of the cells concerned. To avoid the toxicity due to lengthy trypsinization, collagenase may be utilized.

One occasion where culturing of tumour cells is utilized routinely is in the case of characterization of biopsies from human subjects. The tissue taken out is too small for use in diagnostic or analytic protocols; it needs to be amplified. This is sometimes achieved by injecting the test tissue into athymic mice (that lack T cells and NK cells and so do not mount cell-mediated immune responses nor remove nascent tumour cells) known as 'nude' mice (they are hairless).

Usually cell line cells possess anomalous numbers of chromosomes as compared to those in normal cells. For instance, 3T3 fibroblasts possess around 75 chromosomes in place of the normal diploid number, 40, for mouse, while HeLa cells may carry as many as 76 chromosomes instead of the 46 of normal diploid human cells.

It must be, therefore, kept in mind that immortal cells of established cell line cultures are not quite equivalent to normal cells. However, they are sufficiently similar to primary culture cells to be used as representatives of the latter for specific objectives. The most useful feature of established cell lines is that they provide standard stocks of cells which allow meaningful comparison and extrapolation of conclusions from results obtained by diverse investigations. Indeed, the study of human molecular and cellular genetics and biology have advanced extensively with the availability of cell line stocks.

8.2.5 Somatic Cell Fusion

The principle as well as the method of fusing two somatic cells to form a single hybrid heterokaryon or synkaryon are the same as that described earlier for creating plant somatic cell hybrids.

Animal somatic cell hybrids are useful for a variety of research objectives. They allow studies of the effect of a particular genome on another cytoplasm and vice versa. Harris had demonstrated that when a chicken erythrocyte (with a condensed, non-active nucleus) is hybridized with a mouse fibroblast cell (that possesses an active nucleus), the chick nucleus (in the heterokaryon) shows signs of activity. It becomes decondensed, and chick specific gene products make their appearance in the cytoplasm of the hybrid cell.

Interspecific mammalian somatic cell hybrids do not retain both genomes stably as do plant cell hybrids. Normally one of the parental genomes is lost over a few or several cell generations. In man-mouse cell-hybrids, the human chromosomes are lost. By judicious figuring out of the gene-products contributed by the retained chromosome(s), it is possible to assign genes to specific chromosomes. Human genes have been mapped rather rapidly in the past few years using this technique.

Some somatic cell hybrids are exploited for the gene products they synthesize. The most glamourous example of this are the hybridomas—tumours formed by the fusion of a cell committed to synthesize a monoclonal antibody and a cancerous cell that proliferates and thus amplifies the antibody-producing hybrid cells.

The somatic cell hybrids are selected in special selective media that kill off the non-hybrid cells. The commonest of such media that is in use is the HAT medium described in Chapter 2. After selection of the hybrid clones, their hybrid nature has to be confirmed by other techniques (hybridization with probes, immunodetection, enzyme analysis and so on).

PEG (1000, 4000 or 6000), in a 20–60 per cent solution is used for fusing animal cells. The PEG is autoclaved (this also melts it) before adding it to a mixture of cells to be fused that are already attached to coverslips or small petri dishes. After treating the cells with the PEG for 1 minute, it is removed and the cells washed a few times with serum-free medium. Finally, serum-enriched medium is added to the cells and the latter incubated overnight. The selective HAT or other medium is added after removing the used rich medium. The cells that survive after a period of incubation are expected to be the required somatic cell hybrids (see Exercise 51).

How would you obtain TK^- and $HGPRT^-$ cells for use in gene transfer or cell fusion exercises? The simplest way is to request for a sample from someone you know who uses it. Failing that, you may get them from suppliers. You may also isolate them in your laboratory by growing cells in (i) 8-azaguanidine, and (ii) in bromouridine. If cells survive in (i) above they are TK^+ and $HGPRT^-$. Those that survive in spite of (ii) above are TK^- but $HGPRT^+$.

Exercise 46: To Make a Primary Culture of Chicken Fibroblast

The chicken egg has a brittle outer shell, and is more pointed at one end than at the other. It takes about 21 days for a hen's egg to hatch. Fertilized eggs, 8–10 days old, are ideal for making primary cultures of cells that are still not quite organized into

committed tissues. The 8–10 day embryo is removed from the egg, chopped finely and the cells dispersed in a suitable tissue culture medium. The medium containing the cells is introduced into a small $(25\ cm)^2$ T-flask. Cells migrate from the explant pieces and adhere to the inner surface of the flask. Although the explant may contain a variety of types of cells, usually fibroblasts dominate in the culture. Usually in 6–7 days half the flask surface will be covered with a single layer of cells. These cells are now ready to be 'passaged' (transferred) into fresh medium in a new flask. The cells are incubated in the CO_2 incubator in an atmosphere of 5 per cent CO_2, at 37°C and 90–100 per cent humidity.

For passaging cells, they have to be dislodged from the flask surface with trypsin treatment. Trypsin action is terminated by the serum present in the added fresh culture medium. Long trypsin treatments damage the cells. Hence, the period of treatment has to be controlled strictly (about 30 minutes or less at 37°C). The cells may be treated with trypsin for a longer period (overnight) at 4°C and then exposed to warm trypsin for a shorter period (15 minutes).

Have ready

1. A fertilized hen's egg, 8–12 days old, preferably kept at 38.5°C
2. Medium TC 199 100 ml
3. Medium TC 199 + 10% FCS 10 ml
4. 10 T-flasks ($25\ cm^3$)
5. 2 sterile centrifuge tubes (15 ml) with Al foil caps
6. A pair of fine surgical scissors (sterile)
7. A pair each of straight and curved forceps (sterile)
8. A scalpel (sterile)
9. 2 sterile petri dishes (90 mm)
10. Sterile beaker (25 ml)
11. 70% ethyl alcohol
12. Cotton wool
13. Sterile Pasteur pipettes with cotton plugs.

Required for passaging after 1 week

14. PBS soln. (see Appendix)
15. Trypsin soln. (0.25% in PBS). Make 100 ml of this stock solution and dispense in 1–2 aliquots in small screw-capped vials and store at –20°C.

Outline of procedure (Fig. 8.8)

I. Removal of chicken embryo from the egg
II. Initiation of primary culture
III. Passaging of cultured cells

WORK IN THE LAMINAR FLOW CABINET

I. Removal of chicken embryo from the egg

1. Wipe the egg with cotton soaked in 70% alcohol.
2. Fix the egg on the mouth of the beaker with the blunt end facing up (Fig. 8.8).
3. With sharp forceps, crack the blunt end and carefully remove pieces of the shell to expose about $1\frac{1}{2}$ cm of the white opaque membrane below the air pocket.

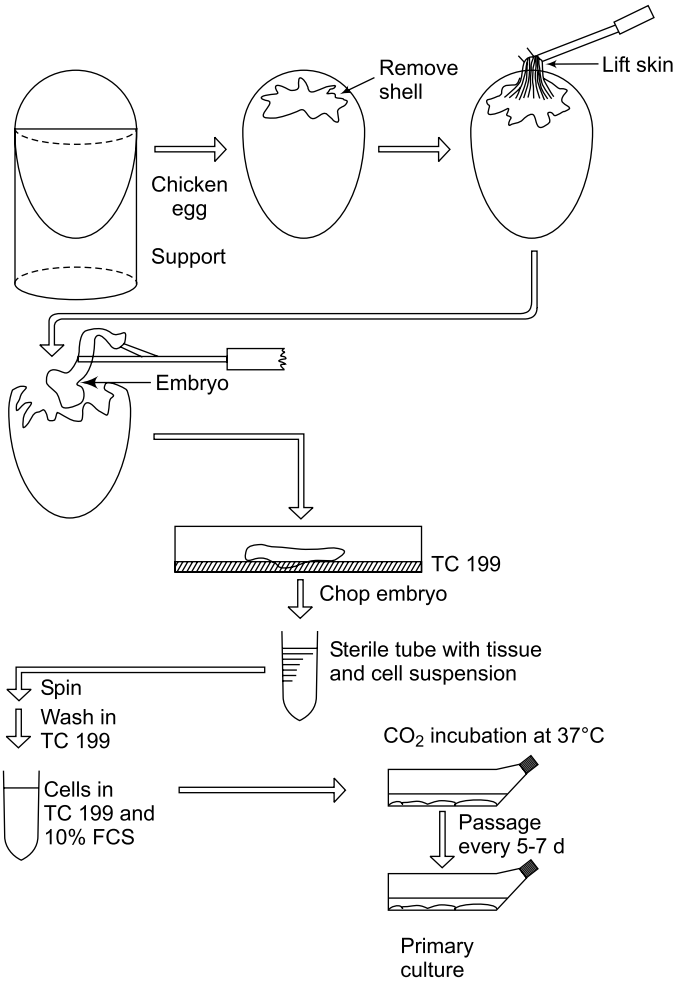


Fig. 8.8 *Steps for generating a primary chicken fibroblast culture. TC 199 is one of the animal tissue culture media.*

4. Flame the alcohol-dipped forceps and remove the exposed white membrane. A transparent membrane with blood vessels of the embryo will now be visible.
5. Now, flame the curved forceps, and gently lift out the embryo.
6. Place the embryo in a petri dish containing sterile medium TC 199 (20 ml).

II. Initiation of primary culture

7. Using the scalpel and fine scissors remove the fatty regions of the embryo, and chop up the remaining tissues into fine pieces. The fluid in the petri dish should become turbid.
8. Using a flamed sterile Pasteur pipette, transfer the turbid and chopped tissues into a sterile centrifuge tube.
9. Spin at 2000 rpm (clinical centrifuge) for 10 minutes. Discard the supernatent.

10. Resuspend the pellet in TC 199 (≃10 ml).
11. Repeat steps 9 and 10 twice.
12. Resuspend the final pellet in TC 199 + 10% FCS (10 ml)
13. Pull up some TC 199 in the sterile tip of a Pasteur pipette and let it out. (This wets the pipette tip; otherwise cells would stick to the dry pipette surface.)
14. Use the wet pipette to withdraw 5 ml of cell suspension and explant pieces from the centrifuge tube and introduce 1.5 ml of it into a T-flask, held with one large flat side facing down. Place 1.5 ml in each of 4 other T-flasks. Repeat step 14 using a fresh pipette to introduce 1.5 ml of fluid + explants into 5 other T-flasks (total 10 T-flasks) Cork the flasks after flaming.
15. Swirl the flasks carefully to spread out the material on the bottom surfaces.
16. Keep in the CO_2 incubator at 37°C overnight.
17. Add 4.5 ml of sterile TC 199 + 10% FCS to each flask after flaming pipettes and flask mouth.
18. Use the inverted microscope to observe the cells adhering to the flask surface.

N.B. • The culture medium should remain lightish pink (pH 6.8–7.0). A darker pink colour indicates alkalinity and a yellowish colour acidity. Change the medium if the colour changes either way.

19. Observe growth once a day. Usually in 5–7 days, more than half the surface should become covered with growing cells. The cells are then said to have formed a confluent monolayer.

N.B. • It is time now to passage these cells.

III. Passaging of a confluent monolayer of cells

20. Remove the medium from each T-flask with a Pasteur pipette. Discard it.
21. Introduce with a fresh pipette 5 ml of PBS into the flask. Unhealthy and loose cells will become suspended in this fluid.
22. Remove the PBS with a pipette.
23. Add 2 ml trypsin (thawed, if kept stored at –20°C) to each flask.
24. Incubate at 37°C (incubator) for 5–10 minutes.
25. Pour out the trypsin (this will contain the detached cells) into a sterile centrifuge tube. You may tap the flask to release the trypsinized cells.
26. Add to the tube 2–3 ml of TC 199 + 10% FCS.
27. Spin at 2000 rpm for 10 minutes. Discard the supernatant.
28. Resuspend the pellet in 5 ml TC 199 + 10% FCS in the following manner:
    (a) Distribute 1 ml of TC 199 + 10% FCS in each tube.
    (b) Add more of the medium to make up the volume to 5 ml.
29. Plate the material in each tube in a T-flask or a sterile petri dish (sealed with parafilm after innoculation).
30. Incubate at 37°C until a confluent monolayer is formed. The cells may be passaged every 5–7 days.

N.B. • Note the number of days required to form a confluent monolayer after each passaging. Does it take the same number of days to form equivalent confluent layers after each passaging? If not, why?

Exercise 47: Passaging of a Human Cultured Cell Line

You will be provided with a T-flask confluent culture of a human cell line, originating from a neoplastic tissue. You will passage this once every week, or when the cells form a confluent monolayer.

Note the number of times you transfer the cells in one semester. How does this compare with the number of passaging in Exercise 46? Note the extent of coverage of the surface every fifth day and on the day of passaging. Compare these cells with those of the chicken fibroblast in terms of their morphology. Add 0.2 ml of PHA-P (Difco) diluted stock soln. to 1 ml of the cell suspension in TC 199. Treat 1 ml of chicken fibroblast suspension in the same way. What do you observe?

Have ready

1. A T-flask with the given cultured cell line.
2. PBS (phosphate buffered saline):

KCl	0.20 g
KH_2PO_4	0.20 g
NaCl	8.00 g
$Na_2 HPO_4 \cdot 7H_2O$	2.16 g
Dist. water	1000 ml
pH	7.3

3. Trypsin soln. 0.25% in PBS (kept in 2 ml aliquots at –20°C)
4. MEM with 10% FCS 100 ml
5. Sterile 25 cm^3 T-flasks for subculturing
6. A sterile screw-capped vial (15 ml)
7. Sterile Pasteur pipettes 1 ml, 2 ml, 5 ml sizes
8. A 'waste' flask.
9. 70% ethyl alcohol; cotton swabs
10. Hemocytometer
11. Incubator at 37°C
12. CO_2 incubator at 37°C, 5% CO_2, 90–100% humidity
13. Inverted microscope

WORK IN THE LAMINAR FLOW CABINET

Procedure

1. Using a cotton swab soaked in 70% ethyl alcohol, wipe off the given T-flask, especially around the mouth and stopper.
2. Pour off the used culture solution from the T-flask into the 'waste' flask (WF).
3. Wash the cells with PBS to remove all serum in the following way:
   (i) With a Pasteur pipette, introduce 5 ml of PBS without disturbing the cell surface.
   (ii) Roll the flask, so that the PBS covers the cells.
   (iii) Keep for 1 minute at room temperature.
   (iv) Pipette out the PBS into the WF.
4. (i) Add to the T-flask 2.0 ml of trypsin soln. (thawed). Close the mouth with the stopper (wiped with 70% alcohol).
   (ii) Place in the incubator (37°C) for 2 minutes. The separated cells should round up by this time.

5. (i) Add to the flask, using a 5 ml pipette, 2.5 ml of MEM + 10% FCS.
   (ii) Aspirate the fluid gently to dislodge the cells and break up the aggregates.
6. (i) Pipette out the fluid (with cells) into a sterile 15 ml screw-cap vial.
   (ii) Make a cell count with a hemocytometer using one drop of the cell suspension.
7. Place 5 ml MEM + 10% FCS in 10 fresh T-flasks.
8. Pipette into each of the 10 flasks, 0.2 ml of the cell suspension from Step 6(i). Stopper the flasks.
9. Incubate horizontally in the CO_2 incubator.

N.B.
- Watch the colour of the medium. If it turns dark pink, add a drop or so of 1 N HCl until the proper pink colour (pH 7.3) is restored.
- If the colour of the pH changes after one day, replace the medium with fresh MEM + 10% FCS or replace only half the medium with fresh medium.

8.3 ❑ HUMAN PERIPHERAL LYMPHOCYTE CULTURE

Free lymphocyte culture has proved to be a valuable asset for the study of human cells. This is especially true for the karyotyping of the human chromosomes and for assigning genes to them. A by-product of karyotyping is the opportunity to detect aberrations and anomalies of chromosomes. This facility has allowed the detection of chromosome defects associated with several inherited human diseases, as well as the assessment of gross effects of mutagens, both of environmental and experimental origin (see Exercises 48–50).

Although conclusions drawn from responses of single cells may not be extrapolated quite correctly in multitissue composite organisms, leucocyte cultures, especially those of small lymphocytes, have contributed to a gross understanding of the biology of the human cells at the genetic and molecular levels.

T cells can be stimulated to divide with the mitogen PHA and B cells with PWM. Cell differentiation is easily studied in cultured lymphocytes. The series of alterations that differentiate a stem cell small lymphocyte into a specialized B or T cell may be followed and monitored experimentally in cultures of lymphocytes stimulated by an antigen and/or mitogen. A new or different antigen on a cell surface can be located in mixed lymphocyte cultures (MLC). MLC is a mixture of two lymphocyte populations differing in surface antigens. The difference may be due to their origins from separate (non-twin) individuals, or to alterations in the surface antigens in populations of lymphocytes in the same individual. Again, in the latter case the difference may be a reflection of changes due to progressive normal differentiation (as the acquisition and loss of surface antigens during the development of the members of the T-cell family), or due to genetic disturbances caused by viral infection or/and oncogenic transformation. In any event, when two differing lymphocyte populations are mixed, each triggers the other to undergo cell-division; each cell surface acts as an antigenic stimulation for the other. Undirectional mitotic induction can be achieved by treating one set of cells with mitomycin C. Now, why this elaborate exercise? What purpose does an MLC serve? The MLC system is a sensitive technique for detecting subtle differences in surface antigens of cells, and have been used to locate particular ones on different T-cell types. The cells that undergo mitosis are called *responders.*

8.3.1 Preparation of Metaphase Slides

A small volume of peripheral blood is collected in a heparinized sterile vial. A small aliquot of this is added to sterile culture vials containing the tissue culture medium, serum, PHA, and antibiotics. The culture is incubated in an atmosphere of 5% CO_2 at 37°C for 48 or 72 hours. Three hours before harvesting time, colchicine is added to the culture vial. Incubation is continued for the remaining 3 hours. The cells are then harvested by low speed centrifugation. The pellet of cells is allowed to swell in a hypotonic KCl solution and then the cells are killed and fixed by die fixative. Slides of lymphocytes are made with the fixed cells. The dried slides are stained by one of the given methods and mounted in Euparol or DPX (mounting fluid). The preparations may be now viewed under a compound microscope. The details of the metaphase plates show up best under oil immersion.

8.3.2 Chromosome Staining and Banding

Chromosomes may be stained with a uniform tint, or banded by alternate regions of stained and unstained chromatin.

1. Staining Methods

Three methods of flat chromosome staining are used for human chromosome preparations. They are also suitable for chromosomes of other eukaryotes. These are: (i) Aceto-carmine or aceto-orcein, (ii) Feulgen, and (iii) Giemsa staining techniques. Feulgen staining imparts a salmon or fuschia-pink colour to chromosomes, aceto-carmine a deep reddish-pink tone and Giemsa a bluish purple colouration.

Aceto-carmine staining is the quickest of the three methods and is usually used for examination of temporary preparations. The material on the slide is covered by a drop or so of a solution of carmine in 45 per cent acetic acid, die excess stain removed with a blotting paper and the preparation treated with 45 per cent acetic acid. This leaves the background clear and the chromosomes stained. The slides may be made permanent by dehydrating in progressively higher percentages of ethanol in water, with a final passage through absolute alcohol, followed in sequence by passage through xylol and die mounting medium (Euparol, DPX).

The Feulgen method consists of hydrolyzing acetic acid-alcohol fixed chromosome material in 0.1 N HCl at 60°C for 8–10 minutes and subsequently treating it with a decolourized stain solution (a Schiff s reagent). The hydrolysis makes available the aldehyde moiety in the deoxyribose units, which combine with the bleached dye (basic fuchsin) to give a salmon-pink colour (Schiff s reaction) to the chromosomes. Either of the above methods of staining leaves the chromosomes in a uniformly stained condition, which allows observation of the gross morphologies of the chromosome arms.

Another technique of staining, that is useful especially for the study of animal and human chromosomes, uses a dye called Giemsa. Giemsa staining can produce either uniformly coloured chromosomes, or ones that have a 'banded' appearance.

Let us first take a quick look at the principle of the Giemsa staining technique. Giemsa is a complex mixture of dyes that include thionin, methylane blue and the

Azure dyes A, B, and C, as well as the red dye eosin. All of die latter dyes, except eosin, consist of a molecule of thiazine, with methyl groups at specific locations in each case. These dyes are positively charged molecules that aggregate along the negatively charged phosphate groups of the DNA backbones. In an unfixed chromosome, these phosphate group charges are associated with the histones of chromatin, most of which are lost when the chromosome is treated with an aceticacid-alcohol fixative and 0.2 N HCl. The freed negatively charged phosphate moieties can now combine with more dye molecules and thus stain the chromosome deeply.

2. Banding Methods

Variations in chromosome treatment, followed by Giemsa staining form the bases for five banding techniques. These delineate what are referred to as C—,G—,R—,Q— and Hoechst 33258 band patterns. Hoechst 33258 is a DNA binding fluorochrome that has been found to increase the photosensitivity of replicating cells in which 5-Bromodeoxy-uridine (5-BrdU) is incorporated in place of thymidine.

Each chromosome in a species exhibits specific band patterns with each of the first four banding techniques. It is thus possible to physically or visually identify homologues and demarcate every region of the linear chromosomes. Comparison of banding in normal and aberrated chromosomes, not only enables one to pinpoint the aberrated region but also indicates the type of abberration. Chromosome banding is, therefore, utilized for locating aberrations in animal cell cultures, particularly in cultures of peripheral human lymphocytes.

Bands and inter-bands are suspected of mirroring some morphological (and underlying molecular?) features of a chromosome. However, such correlation is still far from understood.

In order to discuss the possible significance of the banding patterns in terms of chromosomal features, it would be useful to mention certain cytological observations that are widely acknowledged. Usually, the stained mitotic chromosomes possess certain regions that are more intensely coloured than the others. While the faintly tinted (or unstained) areas are named as euchromatin, the deeply stained heterochromatin is believed to represent highly condensed regions of the chromosome. Some heterochromatin remain deeply stained throughout the cell cycle; these are said to be *constitutive heterochromatin.* Probes made from satellite DNA (that fractionates out from bulk eukaryotic DNA during CsCl density gradient centrifugation) have been found to hybridize with heterochromatic regions. In mouse, these regions are next to the centromeres.

C-bands usually represent heterochromatic satellite DNA. They may also occur, besides next to the centromere, in other regions of chromosomes. *G-bands* are believed to represent the intercalary or facultative heterochromatic regions. *Q-bands* occupy the same regions as the G-bands. *R-banding* is the reverse of Q-banding; areas that are interbands in Q-banding are the fluorescing R-bands, and the bands of the former are the interbands of the latter.

To produce C-bands, one-day old slides are treated with 0.1 N HCl and incubated in 4–5 per cent barium hydroxide solution at 50°C. The washed slides are then kept for an hour at 60°C in a saline citrate solution, after which they are stained with Giemsa.

G-banding is achieved by treating five-day old slides with hydrogen peroxide solution for five minutes, washing in normal saline solution and incubating with the proteolytic enzyme, trypsin. The slides are finally stained with Giemsa.

Q-bands are seen when fixed chromosomes, that are allowed to age for a few days, are immersed in a solution of quinacrine mustard or quinacrine hydrochloride for 5–10 minutes, in a phosphate buffer. The slides, washed in the same buffer, are then mounted in a mixture of the buffer and glycerol. When observed with UV light in a fluorescent microscope, the Q-bands appear as prominent fluorescent areas on the chromosomes.

R-bands are obtained by treating Q-banded chromosomes with acridine orange. The Q-interband regions glow as bright green bands. These are the same as the R-bands which can also be visualized when fixed chromosomes treated with a phosphate buffer at 80°C are stained with Giemsa.

Replication of eukaryotic chromosomes is asynchronous. This means that different regions of the chromosomes in a set synthesize new DNA at differing—though characteristic—times within the allotted S-phase of the cell cycle. The early and late replicating DNA have been mostly identified as GC and AT rich regions respectively. The R-bands coincide with the early replicating GC rich regions and the G-bands with the late-replicating AT rich regions. The Hoechst dye staining method (to be described later) reveals the AT and GC rich stretches in replicating DNA.

Chromosome banding methods, together with molecular biology techniques, are being utilized to answer another nagging question about chromosome morphology. It has been observed that extended meiotic chromosomes possess a beaded appearance, with a characteristic pattern of beads of different sizes on each chromosome. Homologous chromosomes possess identically beaded regions. These beads have been called chromomeres. They are not to be confused with nucleosomal beads which can be observed only with the resolution power of the electron microscope. In the early days of cytogenetics the beads were even believed to represent the 'sites' of genes on the chromosome string. Today, it is suggested that a chromomere is a region of the DNA with a cluster of loops, having AT rich regions at the bases of the loops and GC regions constituting the loops themselves. When several chromomeres lie next to each other, they give the appearance of a larger bead of chromomere. G-banding is seen to intensify the appearance of the chromomeres. The reason for this is not clear.

8.3.3 Differential Staining of Sister Chromatids

Old and new chromatids in a replicated chromosome, that were still attached at the centromere, were first differentiated by J H Taylor, using autoradiography of cells that had incorporated radioactive thymidine ^{3}H-T during one S-phase. Another way of distinguishing newly synthesized chromatids from their parental chromatid partner is to use the Hoechst 33582 dye with incorporation of a halogenated nucleotide, rather than a radioactive one. Cells are allowed to incorporate 5-bromodeoxyuridine (5-BrdU or BU) instead of thymidine during one cell cycle (or two consecutive ones). The newly synthesized chromatids containing the BU can be distinguished

by their enhanced fluorescence from the parental template. If the preparations are further stained with Giemsa, the old chromatids appear intensely stained. This type of staining is especially useful for the detection of chromosome abnormalities where regions of sister chromatids are exchanged reciprocally (SCE). Such exchanges do not show up in material stained with Feulgen, aceto-carmine or Giemsa, since the exchanges do not alter the overall shape or size of the chromatids (see Exercise 50).

8.3.4 Scoring Aberrations in Metaphase Chromosomes

Abnormalities in the chromosome complement may be due to aberrations caused by breakage and illegitimate rejoining of broken ends, or due to anomalies in numbers of single or whole sets of-chromosomes. All aberrations, with the exception of SCEs, can be scored in material stained with Giemsa or by employing modified Giemsa techniques.

Aberrations that show up at mitotic metaphases are usually deletions, translocations, inversions, and dicentric chromosomes.

1. *Deletions* may be terminal or intercalary. A deletion may be small—appearing as a 'dot' or a linear fragment. Often the dot is a minute ring formed by the joining of the broken ends of an intercalary deletion.
2. *Translocations* occur when two non-homologous chromosomes are broken, and rejoined illegitimately.
3. *Inversions* occur when a chromosome breaks in two places and rejoins the centre piece in an inverted position. Inversions can be detected in mitotic chromosomes by comparing the band patterns of homologous chromosomes.
4. *Dicentric* chromosomes are usually accompanied by one or a pair of homologous fragments, depending on whether the break had occurred in (i) a pre S-phase chromosome, or (ii) one that had already replicated. Dicentrics are formed when two centric fragments join. They are accompanied by one or two acentric fragments.

Sister chromatid exchanges are scored in BrdU-Hoechst stained material. The harlequin patterns due to reciprocal exchanges of homologous regions of sister chromatids give away the number of places where such damage has occurred.

Exercise 48: Culture of Human Peripheral T-Lymphocytes

Body fluids, (blood and lymph) contain red and white blood cells, one class of the latter being the circulating T-small lymphocytes. These cells do not undergo mitosis *in vivo* unless challenged by a non-self antigen. Mitosis is triggered, *in vitro,* with a mitogen. T-cell division is stimulated *in vitro* by the lectin phytohemagglutimin (PHA).

Mitosis is induced in T cells so that condensed metaphase chromosomes are obtained for cytological observations. Various gross aberrations and anomalies in the chromosome set are indicative of specific diseases or disorders. Short-term cultures of peripheral T cells, arrested at metaphase are examined cytologically after appropriate staining procedures.

Like other mammalian cells, lymphocytes have to be cultured in a medium supplemented with FCS. However, unlike most animal cell types, lymphocytes

can be cultured in suspensions. Short-term cultures range from 2 to 3 days; CO_2 liberated by the cells into the gas phase of the culture vials is sufficient to maintain the culture fluid in the proper buffered condition.

The method of short term peripheral lymphocyte culture (Fig. 8.9) consists of obtaining blood from a vein, keeping it unclotted with the help of an anticoagulant (heparin) and adding adequate aliquots to sterile medium + FCS + PHA already kept ready in a screw-capped vial. The vials are tightly capped and incubated in a humid incubator at 37°C. Cell division starts after about 24 hours. Colchicine is added to arrest cells at mitosis, three to five hours before the time of harvesting

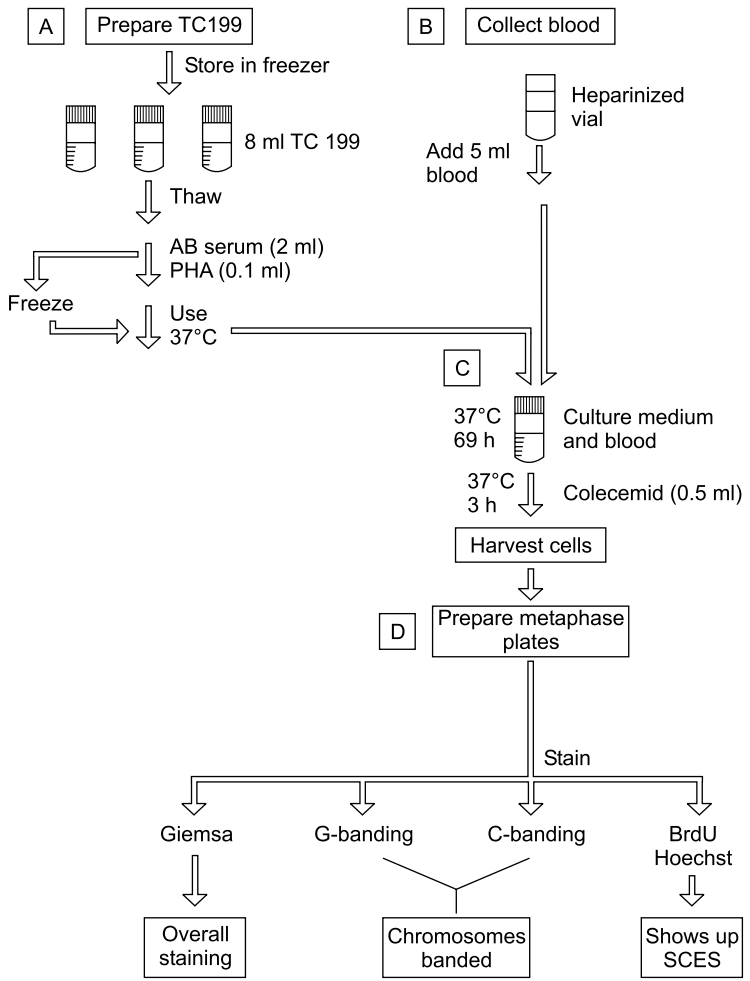


Fig. 8.9 *Steps for short-term culturing of peripheral lymphocytes and staining metaphase chromosomes. PHA, phytohemagglutinin is a lectin that triggers T cells to divide. Mitosis is stopped at the metaphase stage with colchicine (colecemide) treatment. Cells arrested at metaphase are harvested, smeared on microscope slides and stained by an overall or a banding staining procedure. SCE, (sister chromatic exchanges) can be visualized with BrdU treatment, followed by staining with Hoechst 33258. See text for details.*

the cells. The harvested cells are killed and fixed and spread out to dry on clean microscope slides. The slides are stained by the Giemsa trypsin technique or by the G- and/or C-banding method. Gross chromosomal aberrations and anomalies become apparent, if present, in the stained material. One type of aberration—the sister chromatid exchange—is not visible by these staining techniques; another technique—the BrdU-Hoechst staining method is, therefore, employed. Exchanges between the chromatids of the same chromosome become apparent as 'harlequin' or alternately stained and unstained regions. A fourth technique—the Q-band one—requires a fluorescent microscope for observing the fluorescing quinacrine mustard-stained chromosome bands. We will not attempt this last method in this volume.

Now, a word about some of the materials to be used:

1. Heparin This comes in 5 ml tubes with a coating of heparin on its inner walls sufficient to prevent clotting of 5 ml of collected blood, or as a solution (5000 IU/5 ml) in sealed sterile vials. In the latter case, withdraw the solution using a sterile syringe (21 gauge needle) and place 2–3 drops in a sterile screw-capped 15 ml vial. Roll the vial to coat the bottom and sides of the vial with heparin. Close the cap tightly and use immediately or store at 4°C for 2–3 weeks.

2. Phytohemagglutinin PHA-P and PHA-M are sold commercially. PHA-P is a very concentrated solution of the lectin and must be diluted as follows for use.

A. Diluted stock Add to 5 ml of PHA-P (in vial) 5 ml of sterile dist. water. Mix, and distribute 1 ml each in 5 ml screw-capped sterile vials. Store at 4°C.

B. Working solution Add to 1 ml of dil. stock 9 ml of sterile dist. water and mix.

C. Solution for use Use 0.1 ml of solution B for each 5 ml culture. PHA-M is less concentrated. Follow the suppliers' directions for dilution before use. Use 5 μg/ml concentration in the final 10 ml culture fluid.

3. Colecemid We will use the Sigma product Demecolcine. Make the following dilution.

Stock A. 1 mg/ml (1000 μg/ml)
Stock B. 01 mg/ml (100 μg/ml)
Add 1 ml stock A to 9 ml dist. water.
Stock C. 200 μg/ml
Add 2 ml stock B to 98 ml dist. water.
Use. 0.1 μg/ml (or 1 μg/10 ml culture)
Add 0.5 ml of stock C to 10.0 ml of culture medium in vial.
The lyophilized Gibco colecemid may be used as follows:

Dissolve the colecemid in 10 ml sterile dist. water. Use 1 drop/5 ml or 2 drops/10 ml culture.

4. Hypotonic KCl soln. Dissolve 0.56 g KCl in 100 ml sterile dist. water (0.075 M). This solution swells the fixed lymphocytes, so that the metaphase chromosomes are well dispersed and do not lie over each other in the cytological preparation.

5. Fixative For each culture you require 30 ml fixative. In a glass-stoppered bottle, mix 150 ml of methanol and 50 ml of glacial acetic acid. Keep in the refrigerator to chill. Use within 2 hours.

6. Autologous serum Human AB serum is collected, heat-inactivated and stored in the deep-freeze in aliquots of 10 ml in sterile screw-capped vials. It is thawed at room temperature and allowed to warm up to 37°C in the incubator before use.

To prepare the AB serum

(i) Collect 200–250 ml of AB blood in a sterilized saline bottle. The blood collecting kit may be attached to the rubber-stopper of the bottle (to be done by the hospital staff).
(ii) Do not disturb the bottle for 2–3 hours to minimize hemolysis.
(iii) Transport the bottle to the laboratory and keep it at at 4°C overnight. The blood will clot and a yellowish fluid (serum) will separate out.
(iv) Pipette out or decant the serum into sterile centrifuge tubes.
(v) Spin at 600 rpm for 15 minutes.
(vi) Pool the supernatants in a sterile flask. Cap the latter with a sterile Al foil.
(vii) Place the flask in a 55–60°C water-bath for 1 hour.
(viii) With sterile pipettes, dispense 5 ml of the serum into each of several sterile 10 ml screw-capped vials.
(ix) Store the vials in the deep-freeze.
(x) Remove a vial at least two hours before use to thaw the serum; then place it in the 37°C incubator until use.

7. Antibiotics (i) The antibiotics may be used separately: Add 5 ml of sterile dist. water to the powder in the ampoule and filter sterilize (0.22 μ size filter) into a sterile screw-capped vial, using the syringe filtering assembly. For a 1 litre medium use 0.5 ml of banzylpenicillin (200 units/ml) and 1.0 ml of streptomycin.
(ii) We will use Discristicine which is a mixture of penicillin and streptomycin. 5 ml of sterile distilled water is added to an ampoule of Discristicine and shaken till the powder dissolves. Use 5 ml of this in 1 litre of medium.

8. Na-bicarbonate solution Mix 4.5 gm Na-bicarbonate and 100 ml dist. water. Adjust pH 7.0–7.2 (using 1 N HCl). Distribute the solution in 20 ml aliquots in screw-capped milk dilution bottles and autoclave.

9. The Medium TC 199 (with Hank's base), F-10, MEM and RPMI are suitable for lymphocyte and other leucocyte cultures. We will use TC 199 prepared powder. The dry prepared powder remains active for at least 2 years. The liquid medium may be made. It can be stored for about one month without losing its activity.

10. Water As even minute traces of minerals affect animal cell culture, deionized triple-distilled water is preferred.

Outline of procedure

I. To prepare the TC 199 medium.
II. To prepare the culture tubes.
III. To collect the blood for culture.
IV. To culture the lymphocytes for 48 or 72 hours.
V. To prepare the slides.
VI. To stain the slides.

Schedule of work

A. Prepare the medium, store in vials in the freezer.
B. One day before blood collection, prepare the culture tubes (TC 199 + AB serum + PHA). Keep in the deep-freeze.
C. 2 hours before blood collection:
   (i) Have ready in the laboratory—culture vials (thawed) and warmed to 37°C.
   (ii) Go to the hospital for blood collection.
D. In the laboratory: Add blood to each culture tube and keep in the incubators.
E. (i) After 45 or 69 hours, add to each culture, the required amount of colecemid, and incubate for 3 more hours.
   (ii) During this time, have ready:
   - Clean microscope slides in 90% alcohol.
   - Make the fixative and keep in the refrigerator to chill.
   - Make fresh KCl soln.
   - Make working solution of Giemsa stain.
   - Arrange copling jars with stain, dist. water rinses and xylene.

F. Harvest cells after 48 or 72 hours and proceed to fix cells and make slides.
G. Slides may be stained immediately [by technique (A) or after 5 days by technique (B) or after 7 days by technique (C) (Exercise 49)].

I. To prepare the TC 199 medium This may be prepared in advance and stored in several bottles at 4°C.

Materials

(i) 10 gm powder of TC 199 with Hank's base (Gibco)
(ii) Triple distilled (glass), deionized water 1000 ml
(iii) Sterile 4.5% Na-bicarbonate soln.: 20 ml
(iv) Antibiotics (Driscristicine) soln.
(v) One 2-litre flask
(vi) Millipore vacuum pump filter kit with 0.22 μ pore filter (use +ve pressure as the pH alters under –ve pressure).

Method

(i) Take in a 2 1 flask 980 ml water.
(ii) Add to it 10 gm of TC 199 powder. The colour of the medium will be yellow.
(iii) Add slowly, from the 20 ml of the Na-bicarbonate soln., drop by drop, and shake to mix after every few drops, until the colour of the medium changes to a watermelon red.

(iv) Adjust pH to 7.0 to 7.2, using 1 N HCl.
(v) Add to the medium 5 ml of the Driscristicine soln.
(vi) Use the Millipore vacuum pump filter-assembly to filter sterilize the prepared medium. 250 ml lots are filtered at a time.
(vii) Working in the LFC and using aseptic methods, dispense the filtered medium into 30 ml screw-capped vials. Place vials in a wire-basket and store in the deep freeze.

N.B. • Filter sterilize the stored media before use.

II. To prepare the culture medium This may be prepared in advance and stored in the deep freeze.

In each 30 ml screw-capped vial add (in the LFC):

TC 199	8.0 ml
AB Serum	2.0 ml
PHA-M	0.1 ml

III. To collect the blood

A. Have ready before collecting blood (in LFC):
   1. Vials with culture medium mixture (37°C)
   2. Sterile Pasteur pipettes and bulbs
   3. 70% ethyl alcohol
   4. Alcohol lamp.

B. Take to the hospital:
   1. Blood donors
   2. Heparinized collection vials (1 per donor)
   3. Sterile disposable syringes and gauge 20 needles (one set per donor)
   4. 70% ethyl alcohol
   5. Cotton swabs

C. At the hospital (by the medical staff):
   1. Light alcohol lamp
   2. Wipe hands and forearms with ethyl alcohol
   3. Open a packet of syringe-needles and assemble parts handling them with sterile tissue paper
   4. Withdraw 5 ml blood (per donor) and inject into a heparinized vial.
   5. Shake continuously to mix the blood and heparin to prevent clotting.
   6. Return to the laboratory.

IV. To set up the culture Work in the LFC using aseptic methods:

1. Assemble parts of sterilized syringe and needle.
2. Using the syringe and needle, remove heparinized blood and add to each culture vial about 0.5 ml of the blood (about 12 drops). Close the screw-caps tightly.
3. Incubate the culture vials in a humid incubator at 37°C.
4. Mix the culture material, twice a day, by gently turning the vials upside down a couple of times.
5. Three hours before harvesting the cells (i.e., 45 hours or 69 hours after the start), take out the vials and set up in the LFC.

6. Add to each vial and mix 0.5 ml colecemid, stock C.
7. Return the vials to 37°C for 3 hours.
   The cells are ready to be harvested.

V. To prepare the slides

8. Transfer the contents of each culture tube into a 15 ml centrifuge tube (graduated).
9. Spin at 1000 rpm for 5 minutes. Discard the supernatant.
10. Add to the pellet 10 ml of KCl soln. (warmed to 37°C). Keep in the incubator (37°C) for 1.5 minutes.
11. Spin at 1000 rpm for 5 minutes. Discard the supernatant.
12. Add to the pellet, 1 ml of ice-cold fixative, drop by drop, shaking gently after each addition. Add more fixative until the volume is 10 ml. Mix well. The colour of the cells will become almost black.
13. Keep the tubes in the refrigerator (4°C) for 30 minutes, (or in an ice-bucket containing crushed ice).
14. Spin at 1000 rpm (100 g) for 5 minutes. Discard the supernatant.
15. Hold the tube in a vortex mixer while adding 5 ml of fresh, chilled fixative. Keep on ice for 5 minutes.
16. Repeat Steps 14 and 15 until the cells are colourless.
17. To the final pellet add 1.5 ml of the fixative.
18. Take the cells in the fixative in a Pasteur pipette and place 1 drop per microscope slide in either of the following two ways:
    (i) Hold a clean slide in one hand inclined at an angle to the desk top and raise the pipette with the other hand, and let the drop fall by 12–18 inches upon the raised end of the slide.
    (ii) Set up an arrangement for making slides so that the pipette remains fixed (held in a ring stand) and slides can be placed inclined on the bench-top. In either case, the drop of fixative with the cells will run down the slide. The impact of fall will burst the fragile swollen cells, and the metaphase chromosomes will stick to the glass surface. Observe one slide under the microscope to check the density of the cells. If the density is too high add a little more fixative to adjust the cell density, if necessary. If the density is very low, pool the cells from replicate vials and suspend them in 0.5 ml of fixative and check the density again.

VI. To stain the slides

19. Air-dry the slides, or blow over them from one end.
20. Keep the slides in a dry, dust-proof slide box until required for staining.
    The slides will be stained by the Giemsa-trypsin, modified Giemsa-trypsin, G-banding, C-banding and the BrdU-Hoechst methods.

Exercise 49: Staining of Peripheral Lymphocyte Chromosomes

The techniques to be used for staining the prepared slides of lymphocytes from Exercise 48 are:

A. Giemsa staining: Conventional

B. G-Banding, and
C. C-Banding.

These techniques may be also employed for staining of other mammalian cells.

Technique (A) tints the chromosomes uniformly with a bluish mauve colour. Technique (B) stains the chromosomes in bands, which alternate with unstained interbands. The banding pattern is characteristic of each chromosome in a complement. The trypsin treatment in technique (B) removes chromosomal proteins. C-banding reveals regions of constitutive heterochromatin—highly condensed regions with a high density of repeated DNA sequences that do not stain well. Some regions of a chromosome remain heterochromatic throughout the cell cycle. These are the constitutive heterochromatic zones. Other regions become heterochromatic at certain times and under certain conditions. These regions are said to be facultatively heterochromatic.

In mouse chromosomes, constitutively heterochromatic regions occur on either side of the centromere.

Technique (A) is useful for karyotyping and for scoring aberrations such as gaps, breaks, and rings. Technique (B) shows up specific regions of each chromosome. Hence deletions, duplications, inversions and translocations may be identified by comparing with a normal karyotype. Technique (C) of C-banding identifies the centromeric regions in mammalian chromosomes.

(A) Giemsa staining

Have ready

Materials

1. 5% Giemsa stain in phosphate buffer (pH 6.8):
   (a) You may use readymade liquid Giemsa stain and dilute it for use. Mix:

Liquid stain	5 ml
and	
Phosphate buffer	95 ml

   or

   (b) Prepare Giemsa stock (GS) from powdered stain as follows:
      (i) *Have ready:*

Giemsa Powder	500 mg
Glycerol	50 ml
Methanol	50 ml

      (ii) Dissolve the powder first in 5 ml glycerol and 5 ml methanol. Shake to mix.
      (iii) Add the remaining glycerol and mix.
      (iv) Add the methanol and mix. Allow the solution to rest overnight at 30–50°C.
      (v) Filter the solution and keep in a brown, screw-capped bottle, labelled Giemsa Stock Soln.
      (vi) To prepare Giemsa working solution (GWS) from GS:
      For 10 ml : use 10 ml buffer and 0.5 ml GS
      For 100 ml : use 100 ml buffer and 5 ml GS
      If only a few slides are to be stained singly, make 10 ml of GWS.
      If many slides are to be processed, make 100 ml of GWS. This is enough to fill 1 coplin jar.

2. Sorensen's Phosphate Buffer pH 6.8:
   Prepare solutions A and B separately and mix before use.

Soln. A.	KH_2PO_4	9.08 g	or	4.54 g
	Dist. water	1000 ml		500 ml
Soln. B.	$Na_2HPO_4 \cdot 2H_2O$	11.88 g	or	5.94 g
	Dist. Water	1000 ml		500 ml

3. Distilled water
4. Xylene 100 ml
5. A mounting fluid, such as DPX or Euparol
6. Clean coverslips
7. Immersion oil for use with the oil-immersion lens.
8. A binocular high power microscope with high dry (× 100) and oil immersion lenses and 10 × or 15 × eyepieces.
9. Camera to fit the monocular eyepiece attachment of the above mentioned microscope.
10. Seven coplin jars with covers.

A. Procedure

1. Arrange 6 coplin jars, with labels as follows, and fill them with the appropriate material.
   Giemsa, DWI, DWII, DWIII, DWIV, Xylene.
2. (i) Place the slides horizontally in a petri dish and cover the cells with one or two drops of the Giemsa stain for 2 minutes.
   or
   (ii) Place the slides, back-to-back in a Giemsa coplin jar (holds 8 slides). Keep the slides in this jar for 2 minutes.
3. (i) Place the slides in a coplin jar of tap water, under a running tap. Let the slides be washed in running water until excess and precipitated stain granules are washed away.
   or
   (ii) Transfer the slides from the Giemsa jar to a coplin jar of tap water and wash under running water as in (i) above.
4. Pass the slides through the four dist. water jars, i.e., dip, allow the drip to drain in the same jar and place the slide in the next dist. water jar.
5. Finally transfer the slides to the xylene jar for 10 seconds. Remove the slides and allow the drip to drain into the jar.
6. Place 1 drop of DPX on the cells and cover with a fine (00) coverslip (rectangular).
7. Allow the DPX to set in a dust proof, undisturbed place (between blotting paper sheets inside a chest of drawers).
8. Place under a microscope and scan for lymphocytes arrested at metaphase.

9. Take photographs of good metaphase spreads, using a 35 mm camera green filter 100 ASA black and white film and 20–25 second exposure.
   This staining is good for making karyotypes and for scoring aberrations such as gaps, breaks and rings.

(B) G-Banding

Have ready

1. Use slides that are at least 5 days old.
2. Normal saline solution:

NaCl	0.89 gm
Dist. water	100 ml

3. H_2O_2—15% soln. in water.
4. Phosphate buffer pH 6.8:

NaCl	4 gm
KCl	0.1 gm
Na_2HPO_4	0.825 gm
KH_2PO_4	0.0425 gm
Glucose	0.5 gm
Dist. water	50 ml

5. Giemsa 5% solution [as in technique (A)].

   or

   (i) Dissolve 1 gm Giemsa powder in 66 ml glycerol (AR), stirring with a magnetic stirrer.
   (ii) Keep the solution (in a flask) at 56°C (water-bath) for about 2 hours.
   (iii) Let the solution cool to room temperature.
   (iv) Add 66 ml of methanol and stir to mix with a magnetic stirrer.
   (v) Keep this Giemsa stock (GS) in a dark bottle at 4°C.
   (vi) To use: Add to 100 ml phosphate buffer (pH 6.8) 5 ml GS. This gives a 5% Giemsa Working Soln. (GWS).
6. Trypsin solution — Stock:

Trypsin	0.2 gm
Phosphate buffer	10.0 ml
Dist. water	90.0 ml
Trypsin Working Soln. (TWS):	
Trypsin stock soln.	2 ml
Dist. water	48 ml

7. Ten coplin jars

B. Procedure

1. Label the coplin jars as follows and arrange them in the given order. Fill each jar with the appropriate material.
   Peroxide, Nor. Sal I., TWS, Nor. Sal. II, Giemsa, DWI DWIII, Xylene.
2. Keep each slide for 5 minutes in the peroxide jar.
3. Transfer each slide for 10–15 seconds in Nor. Sal I.
4. Keep each slide for about 100 seconds in TWS at 20°C.
5. Transfer each slide, for 10–15 seconds to Nor. Sal II.
6. Pull out each slide, allow the drip to drain into the jar, and air dry the slide.

7. Place each slide for 2–5 minutes in Giemsa.
8. Place each slide for 1–2 minutes in Phosphate buffer.
9. Air-dry each slide and observe under the microscope:
   - If the chromosomes are indistinctly stained, restain in Giemsa, wash in buffer and check again.
   - If banding is not distinct, return to Step 3 and repeat the trypsin treatment for a longer period and continue with the remaining steps.
10. If the staining and banding are satisfactory, pass each slide for 5 seconds each through DWI, DWII and DWIII.
11. Dip each slide in xylene; drain the drip.
12. Place 1–2 drops of DPX over the stained cells and cover with a coverslip.
13. Allow the DPX to set in a dust proof, undisturbed place (takes about 1 day).
14. Score the slides for metaphases.

(C) C-Banding There are certain regions of the chromosomes that are always highly condensed in the metaphase state. These constitute heterochromatin regions. In mouse, they occur on either side of the centromere.

Have ready

1. Slides that are at least 7–10 days old
2. 0.2 N HC1 in a coplin jar
3. Dist. water in 5 coplin jars (DWI-DWV)
4. Barium hydroxide — 5% solution in dist. water — in a coplin jar

   or

   0.07 NaOH soln. in dist. water.
5. SSC 2 X solution:

Sodium citrate	0.09 M soln. in water
NaCl	0.9 M soln. in water
pH	6.9–7.0

6. Giemsa working solution (5% in phosphate buffer, pH 6.8) — as in (A).
7. Xylene
8. DPX
9. Oven set at 50°C
10. Oven set at 60°C

Procedure

1. Place each slide in coplin jar of 0.2 N HCl for 1 hour (room temperature).
2. Pass each slide for 30 seconds each through DWI, DWII and DWIII.
3. Transfer each slide to Barium hydroxide solution and keep the coplin jar (covered) in the 50°C oven for about 10 minutes.

N.B. • Find the optimum period of time for satisfactory staining, by trying out aliquots with 5–15 minutes treatments of $Ba(OH)_2$

4. Air-dry each slide and place in SSC 2 X solution. Keep the coplin jar in the 60°C oven for 1 hour.
5. Wash slides, by passing them through DWIV and DWV. (You may rinse once more in dist. water.)

6. Place each slide in Giemsa working solution at room temperature for 30 minutes.
7. Wash off the excess stain, pass through dist. water rinses and xylene and mount cells in DPX. Allow to dry as in (A).

Exercise 50: Staining of Chromosomes to Reveal Sister Chromatid Exchanges (SCEs)

Exchanges between homologous chromatids cannot be detected by the conventional staining methods. J H Taylor was the first to distinguish the new from old strands of replicated DNA in autoradiographic pictures of chromosomes that had undergone one or more replications after an initial exposure to a labelled nucleotide (tritiated thymidine) in the growth medium. The newly replicated strands had incorporated the radioactive dTs and could be distinguished from the unlabelled parental strands.

Another method has been in use which reveals the new and the conserved DNA strands in chromosomes that have passed through one or two replications in the presence of a non-radioactive analogue of thymidine. The analogue used frequently is 5-Bromodeoxyuridine or 5 BrdU. Fluorodeoxyuridine (FdU) has also been used for the same purpose.

To find out if SCEs have occurred in response to an environmental insult, the test cells (most conveniently, peripheral T lymphocytes) are treated with the test material together with 5 BrdU, and cultured for 72 hours by the method described in Exercise 48. The slides prepared from these cells are treated with a DNA binding fluorochrome, the bisbenzimidazole dye Hoechst 33258. Chromosomes with bromodeoxyuridine are sensitive to light; the polynucleotides break at the sites of BrdU incorporation. Hoechst 33258 enhances this photosensitivity. The Hoechst-treated preparations, when viewed under a fluorescent microscope, show brightly fluorescing and dull looking chromatids. In the absence of a fluorescent microscope, the BrdU-Hoechst treated chromosomes may be further treated with Giemsa stain. Incorporation of BrdU causes alterations in the binding capacity of dyes like Giemsa mixtures. These alterations are particularly noticeable with the enhancement caused by the Hoechst dye.

So, if SCEs occur during DNA replication, the regions of exchange will show up as a harlequin pattern of Giemsa-stained and unstained (or faintly stained) regions after the BrdU-Hoechst treatment. BrdU itself is a mutagenic agent that induces sister chromatid exchanges. To determine the SCEs caused by a test agent, lymphocytes may be cultured in three sets: one control set without any agent, another with the test material and BrdU, and a third set only with BrdU. The base level SCEs due to spontaneous events, and those due to BrdU may thus be detracted from the data in the second set.

We will culture, human peripheral T lymphocytes with 5 BrdU in the dark at 37°C for 72 hours, arrest dividing cells at metaphase, prepare slides, age them for a few days and treat them with Hoechst 33258, exposing the slides to sunlight during this treatment. The slides will be stained with Giemsa (A) of Exercise 49. After you become proficient in this technique, you may try to assess the damage caused by a chemical mutagen (e.g., EMS) as indicated by sister chromatid exchanges.

Have ready Items 1–9 as in Exercise 48. In addition you will require the following:

10. SSC 2 X:
    Dissolve NaCl 17.55 g, Na_3 Citrate 17.53 g in 1000 ml dist. water.
11. (i) 5 BrdU stock solution:
    Dissolve 100 mg 5-BrdU (Sigma)
    in 50 ml sterile, double-dist. water.
    And filter-sterilize using the Millipore vacuum pump-filter assembly. Store the stock soln. in a dark bottle wrapped in Al foil at 0°C (Freezer of the refrigerator).
    (ii) Working solution:
    Add 2 ml stock BrdU soln. to 8 ml sterile dist. water.
    (iii) Solution for use:
    Add 0.25 ml of (ii) above to each culture vial of 10 ml medium + blood.
12. Hoechst 33258 (Sigma):
    Stock solution (50 μg/ml water)
    2.5 mg Hoechst 33258.
    50 ml dist. water
    Store the solution in a light proof (Al foil wrapped) bottle at 4°C.
    - Bis-benzimide, 2′-(4-hydroxyphenyl)-5-(4-methyl-l-piperazinyl)-2, 5′-bi-H-benzimidazole trichloride pentahydrate).
13. A shallow tray or casserole dish lined with wet cotton wool and filter paper.
14. Microscope slides and rectangular coverslips.
15. Microscope with oil immersion lens, and an Interference band filter.
16. Photographic equipment.

Procedure

1. Collect blood and set up culture vials as in Exercise 48.
2. Keep aside 2 vials to serve as CONTROL material. To each of the remaining vials add 0.25 ml of BrdU solution (10 ml/culture vial).
3. Culture cells and prepare slides as in Exercise 48.
4. Allow the slides to remain in the slide box for at least 5 days at room temperature, then proceed as given below.
5. Wash the slides in running water.
6. Place each slide on the lower half of a petri dish.
7. With a Pasteur pipette, cover the cells with 2–3 drops of the Hoechst solution. Cover with a coverslip.
8. Let the dye remain for 10 minutes.
9. Remove the coverslip and wash each slide individually in running water.
10. Add SSC 2 X soln. to the slide (to prevent drying of cells). Add coverslip.
11. Place the slides (with SSC 2 X) on the surface of a wet filter paper placed on wet cotton-wool lining the bottom of a tray.
12. Place the tray with slides in direct sunlight (strong) for a minimum of 2 hours.
13. Remove coverslips.
14. Wash the slides in running water.
15. Wash the slides in dist. water.
16. Stain in 5% buffered Giemsa (pH 6.8) as in Procedure (B) Exercise 49. Mount in DPX.

17. Examine the slides under the oil-immersion objective for 'harlequin' chromosomes.
18. Score the number of such SCEs and make sketches of the different types of SCEs.
19. Score the number of SCEs in the CONTROL CELLS.
20. Score the number of cells per 100 cells examined, that contain SCEs, in both sets 19 and 20.
21. Take photographs.

Exercise 51: Somatic Fusion of Animal Cells

Human peripheral and mouse peripheral lymphocytes will be used for the somatic cell hybridization. PEG will be the fusing agent. Equal numbers of cells of men and mouse will be mixed in suspension cultures with PEG 1000. To select the hybrids, a selection method based on mutant genes in one of the parents (e.g., TK^-, $HGPRT^-$) is required. You may, on the other hand, as in this exercise, take aliquots of non-mutant cell mixtures after the PEG treatment and make metaphase preparations. These may be treated with Hoechst 33258 and stained with 5% buffered Giemsa. The mouse chromosomes stain deeper with Giemsa under these conditions than the human chromosomes.

We will be using lymphocytes that grow in suspension. Somatic fusion may also be achieved by using cells that grow in monolayer cultures. These latter cells can be released from the growth surface by trypsinization and the trypsin replaced by TC 199 or MEM with FCS. The parent cells may then be mixed in suspension and plated out in T-flasks, petri dishes or multiwell dishes, each containing a few cells. The latter may be examined under the microscope for evidences of fusion. If TK^- cells are available (e.g., L strain of mouse cell line), then after growing the mixed cells first in a complete medium with FCS, they may be transferred to a HAT (hypoxanthine-aminopterin-thymidine) medium. The HAT medium will support the growth of fused cells better than that of the TK^+-containing parent cells. If the second parent is $HGPRT^-$, then neither parent, but only the hybrid cells will survive in the medium.

The HAT medium has the following composition:

Hypoxanthine	10 mM
Aminopterin	0.6 μM
Thymidine	1.6 μM

Have ready

1. Heparinized human and mouse peripheral blood
2. TC 199 medium with Hanks base (without serum)
3. TC 199 medium with Hank's base and 20% FCS
4. 1 M NaOH
5. PEG 1000 solution in TC 199

Take a 25 ml flask and place about 1 teaspoonful of PEG 1000 in it. Plug the flask and cover the plug with Al foil and autoclave. Let the melted PEG cool down to 37°C. Mix equal volumes of melted PEG and TC 199 without FCS (10 ml each).

Use 1 ml per 15 ml cell suspension.

6. Sterile Pasteur pipettes
7. 5% Giemsa in phosphate buffer (pH 6.8)
8. Hoechst 33258 stock solution (see Exercise 50)
9. SSC 2 X soln. (see Exercise 50).
10. Dist. water
11. Xylene
12. DPX (mounting medium)
13. Microscope slides and rectangular coverslips
14. Photographic equipment.
15. PHA (see Exercise 48 on culturing of lymphocytes)

Procedure

1. Pellet the human and mouse cells, separately, after 48 hours in TC 199 + 20% FCS + PHA + antibiotics
2. Check cell densities. Adjust to about 2×10^5/ml in TC 199 without serum. You should have at least 5 ml cell suspensions of each kind.
3. Mix 5 ml of human cells culture and 5 ml of mouse cells culture in a sterile 30 ml centrifuge tube.
4. Spin at 1200 rpm (150 g) in a clinical centrifuge. Discard the supernatant.
5. Wash by spinning in 15 ml of TC 199 without FCS.
6. Remove the supernatant with a Pasteur pipette.
7. Add to the pellet, little at a time and gently, 1.0 ml of the prepared PEG and let the tube rest for 1 minute.
8. Add to the suspension 9 ml of TC 199 (without FCS).
9. Distribute this 10 ml into two centrifuge tubes and add 15 ml TC 199 to each tube (without FCS).
10. Spin and discard the supernatant.
11. Resuspend the cells in 15 ml TC 199 (with FCS).
12. Incubate at 37°C for 20 hours.
13. Spin to pellet cells.
14. Fix in 3:1 (methanol; glacial acetic acid) fixative.
15. Treat with Hoechst 33258 as in Exercise 50 (without the BrdU treatment).
16. Stain with 5% Giemsa in phosphate buffer (pH 6.8) as in Exercise 49.
17. Screen metaphases containing both darkly stained and lightly stained chromosomes.
18. Karyotype mouse chromosomes and human chromosomes from the mouse and human cells.
19. Karyotype the chromosomes in the hybrid cells.

REVIEW QUESTIONS

1. How are calluses produced from explants?
2. Describe haploid culture technique for plant generation.
3. Review methods of culturing isolated protoplasts.
4. Enumerate the ingredients of a culture medium that induces the formation of calluses from explants and single cells.
5. Recall landmark investigations in animal-cell culture.
6. Review the requirements for animal cell culture.
7. How is culturing of transformed cells and established cell lines carried out?
8. Describe the chromosome staining and banding techniques.
9. How is the differential staining of sister chromatids performed?
10. Enumerate the aberrations encountered in metaphase chromosomes that are scored.

Basic Immunological Techniques

9

Introduction

Antibodies or immunoglobulins are the main ammunition of the defence system of the animal body. The ability of an antibody to bind to a matching antigen is exploited in a variety of techniques used in immunochemistry and immunodiagnosis. Antibodies are, thus, excellent chemical reagents that are used to identify and isolate even small amounts of an antigen from a heterogeneous mixture.

Immunodiagnosis has become an essential tool in molecular biological investigations. It is almost inseparable from most DNA cloning projects. The methodologies and approaches of immunology are so different from those of other experimental biological systems that an introduction to the basis of use of the immunological tools and a familiarity with the paraphernalia of immunological techniques become necessary components of even an elementary introduction to genetic engineering.

The ensuing pages will include simple procedures for the following: (i) raising polyclonal antibodies against a single protein (BSA) and against a whole cell with several antigenic determinants on its surface; (ii) isolating leucocytes from rat spleens, and T and B cells from lymph nodes of mice and from peripheral blood of man; (iii) assay methods for antibody-antigen response; (iv) purification of an antibody from mouse serum; (v) determination of the number of cells secreting immunoglobulin molecules; and (vi) the raising of monoclonal antibodies.

A brief introduction to the principles underlying the above techniques will be provided first.

9.1 ❑ RAISING OF POLYCLONAL ANTIBODIES AGAINST A SINGLE PROTEIN OR CELL

Antibodies are raised in animals for two general objectives. One of them is that of immunizing the animal (or man) against a specific pathogenic or allergenic antigen. The second goal is to procure antibodies to be used as reagents. An antigen often possesses several antigenic sites on the outer surface, each one of which may provoke the synthesis of a matching antibody. The different antibodies are generated by separate clones of B cells. Hence the blood or lymph of an animal injected with a multi-determinant antigen (immunogen) contains a mixture of what are referred

to as polyclonal antibodies. Immunoglobulins also come in many sub-types and varieties. The sub-types of Ig as well as the ratio between their number vary from animal to animal and under different conditions of immunization.

To raise antibodies, a suitable animal is injected with the desired antigen according to a particular schedule. The antigen may be available in a purified form, or may have to be purified from a mixture of antigens. In either case, the antigen is rarely injected alone into the animal. It is necessary to add a reagent, called an adjuvant to the antigen, an emulsion made of the two, and this emulsion introduced into the animal. The antigen may be injected intramuscularly, subcutaneously, intravenously or interperitonially, depending on the antigen-adjuvant usage. If they are utilized as a mixture, the first two methods are preferred. If used separately, the antigen may be injected into footpads (in the case of an animal like a rabbit) or intravenously, while the adjuvant is delivered subcutaneously.

9.1.1 The Antigen to be Used

To be effective, the antigen used must be reasonably pure. Even biochemically pure proteins and other antigens, may contain traces of impurities that have antigenic qualities. There are several ways of purifying antigens. One of the most efficient ways of selecting a particular antigen from a heterogeneous mixture (such as membrane proteins or sub-units of an enzyme) is that of fractionating the proteins by sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE). The band corresponding to the expected molecular weight of the desired protein is cut out from the gel, the protein eluted from this, and used as the antigen. There is one snag in this method if more than one protein in the mixture, has similar or identical molecular weights, they migrate together to the same band in the gel. Two-dimensional gel electrophoresis may be then utilized to separate the different molecular species.

If polyclonals are to be raised against a protein found only in a mixture, the latter is subjected to SDS-PAGE having one large well for the prepared protein sample. After electrophoresis, a vertical strip may be removed from the gel slab and stained for proteins with Coomassie Brilliant blue. The bands of proteins in the remaining gel can then be identified by placing the stained gel strip, alongside. If it is not known which of the bands is the antigen to be isolated, it will be economical in terms of animals to pool the material from 3 or 4 consecutive bands, and use these mixtures for immunization. Only the one containing the desired antigen will give a positive response in the treated animal. In the next step of the same exercise, the proteins of the separate bands in the positive mixture are used as antigens and injected into separate animals. The band that induces the production of the required antibody, represents the antigen to be isolated for utilization in further projects.

The protein from a polyacrylamide gel may be prepared for injection in the following manner:

The band is cut out from the gel, and excess gel material is shaved off from this block with a sharp knife or blade. The acrylamide is toxic to animals and should, therefore, be introduced only very sparingly. The gel-DNA block is broken up

by aspirating several times with a syringe having a needle with a large bore. This macerated material is then mixed with the adjuvant in a ratio (by volume) of 4 of the former and 5 of the latter.

Very little antigen is required to raise appreciable amounts of antibodies. This fact is exploited for vaccination against pathogens. A small quantity of a dead or disabled (attenuated) pathogen (bacteria, virus, protozoa, others) is injected into the animal (or man) to be immunized. The body of the latter responds within less than a week (4–6 days) by producing antibodies against the epitopes (provoking chemical groups) of the antigen. The antibodies populate the blood, lymph systems, spleen and bone marrow regions. When the individual is assailed in future with live and active pathogens of the kind used for immunization, the body is already ready to attack them with its pre-prepared store of Ig molecules and with cells (memory) which are already primed to differentiate further for synthesizing immunoglobulins identical to the ones provoked initially (primary).

In practice, a rabbit is injected with 0.1 mg of a protein in four weekly doses to obtain enough antibodies for use. Even as little as 1 μg of a protein, attached to a solid support, can induce the synthesis of sufficient antibodies. This is very fortunate, as in many instances only micro-quantities of the antigen are recoverable or available.

The antigenecity of a protein depends also on the presence of proteins in the immunized animals that are polymorphic versions of the introduced antigen, or are otherwise its close evolutionary relatives. In such cases, either an animal not likely to have related proteins is chosen for immunization, or the antigen is coupled to another protein which is immobilized on a solid support such as Sepharose.

Before beginning the schedule of immunization, a sample of blood should be removed (from the ear vein of a rabbit, when the latter is used) to serve as a control system. At the end of the immunization period, or at weekly intervals, blood may be removed and the antiserum (serum in this blood containing antibodies) may be extracted and tested for the titre of synthesized immunoglobulins.

9.1.2 Use of Adjuvants

An adjuvant is a material that enhances the production of antibodies and/or prolongs the life of an antigen in the body. An adjuvant may be a non-antigenic material such as mineral oil or alum, or one that itself elicits antibody formation, such as tubercle bacilli (mycobacteria). An animal is injected once a week or every 10 days 4 or 6 times with the antigen-adjuvant mixture to raise sufficient antibodies to be used *in vitro.*

The adjuvant may be prepared in the laboratory. It is more usual to use the readymade ones. The most frequently employed store-bought adjuvants are Freund's Complete and Incomplete Adjuvants (FCA and FIA). Both contain an oil and an emulsifier. The FCA possesses, in addition, killed mycobacteria. The first of the weekly (or 10 day) series of injections usually uses FCA, while subsequent immunizations (sensitizations) employ the Incomplete version.

9.2 ❑ DETECTION OF ANTIBODIES

The techniques for testing antibodies range from simple qualitative assays to sophisticated methods that allow specific identification of subclasses of Ig molecules and quantitative estimation of the reacting molecules.

The techniques of one set of tests are based on the fact that populations of an antigen and its matching antibody, when present in equivalent concentrations, form a network or spongy mass that precipitates from the solution. These techniques include: (1) the precipitin, (2) the immunodiffusion, and (3) immunoelectrophoretic tests. (4) Antibodies also precipitate larger units, such as bacterial and other cells, by cross-linking the antigenic determinants on their surfaces. Red blood cells are used in one type of tests with antibodies against their surface antigens—this is the basis for the hemagglutination assays. A second set of tests that aims to detect particular antigens in cell preparations or on live cells are represented by the (5) immunofluorescent techniques. A third type of assay involves antibodies labelled by radioactive or non-radioactive tags. (6) Radio-immunoassay (RIA) and (7) enzyme-linked immunoadsorbant assay (ELISA) belong to this category, and are useful for both detection and quantitative estimation of antibodies that are present in only microquantities in the sample. In addition, particular subtypes of antibodies may be purified by adsorbing them on antigens conjugated to a support material. This is a modified version of (8) affinity chromatography.

Needless to say, the above techniques may be used conversely to assay the presence and quantity of antigens, by providing antibodies as the known components in the tests.

There are a few other special tests that exploit the activities of members of the immune system other than the Ig molecules. One of them is (9) the complement fixation test.

In addition to the above techniques for antibody/antigen assays there are others for estimating hypersensitivity reactions. These reactions are allergenic (short-term) responses to antigens that challenge cells which occur in the lining of passages of the body exposed to the environment (i.e., mouth, nose, windpipe etc.). Hypersensitivity reactions involve the activities of the IgE which are found on the outer surface of mast cells and some leucocytes in the linings of the above mentioned passages. The IgE present needs to be detected and estimated for certain purposes; (10) the radioallergosorbent test (RAST) is one such test that may be used on such occasions.

9.2.1 The Precipitin Test

The simplest precipitin reaction is carried out in solution in test-tubes. It is usual to carry out a precipitin test for a quantitative estimation of the amount of antibody in a serum, by mixing different dilutions of the antigen to the same volume of the antiserum, incubating the mixtures first at 37°C for about an hour, and then at 4°C for a longer period (half an hour to overnight for stronger concentrations of antibody and for 10 to 14 days for poorer concentrations). The material in each tube in the series is then centrifuged and the supernatant tested for excess antigen or antibody. The pellet, if formed, is washed with the buffered saline (PBS) and

assayed for the amount of antibody in the complex. The estimate is usually made by dissolving the precipitate in NaOH (0.1 M), measuring the optical density of the solution at 280 nm (extinction wavelength for proteins) and calculating, from the OD data, the amount of antigen in the solution giving the precipitate.

The precipitin test may be used just for ascertaining whether a particular antibody (or antigen) is present in a serum. The thin cloudy line of precipitate at the juncture of the antibody and antigen solution, that appears after the incubation period, is an evidence for the presence of the molecule under consideration.

9.2.2 Immunodiffusion Tests

Immunodiffusion tests are precipitin reactions made possible in an agar slab. The antigen and antibody are placed in a thin slab of agar or agarose and allowed to diffuse towards each other through the gel material. Here too, aggregates of antibody-antigen complexes form only at a distance between the reacting molecules where the latter are at equivalence value. The aggregates precipitate in a sharp line that appears opaque against light, and which may be made prominent by staining. The immunodiffusion techniques are more useful in most circumstances than the precipitin test as (i) very little of either of the reactants is required, and (ii) responses to individual antigens in a mixture or vice versa may be seen as well as measured.

The single-diffusion technique described above has been superseded by a double-diffusion method developed by Ouchterlony. In this case, the antigen is placed in a well with antisera in wells surrounding the antigen or vice versa. The radially placed wells may contain different dilutions of the same antiserum. In the Ouchterlony method, precipitin lines can be seen to form between the wells of antiserum and antibody at the point of equivalence. The immunogenic relationships of two antibody samples may be qualitatively determined by placing them in adjacent wells in the ring around the central antigen-containing well.

If two adjacent wells contain identical antibodies, the opaque line of precipitates will be a curved one (a). (b) If the antibodies are partially related (that is if they possess one or more common antigenic determinants)—a curve with a spur at a tangent to it forms between the two antibody wells and the antigen (c). If the adjacent antibodies are totally unrelated, the lines due to each antibody cross each other and do not form any joint curved region. The curved line is generally concave to whichever of the reactants that has the higher molecular weight.

The agar slab is a thin one (2–4 mm) that is allowed to solidify on a clean microscope slide or petri dish bottom. The slides and petri plate bottoms must be allowed to sit on perfectly levelled surfaces both during setting of the gel and the double diffusion operation. The latter is carried out in a moist chamber such as a large petri dish with a wet filter paper lining the bottom plate. Holders are available, or may be home-made, that carry six or more slides each and allow them to be transported without movement of the slides as well as facilitate easy removal of each slide.

9.2.3 Immunoelectrophoretic Tests

Normally antigens are mixtures of proteins or determinants on the same antigen. If such a mixture is used to immunize an animal, the antibodies formed against each antigenic challenge will result in a heterogeneous collection of antibodies of different types and subtypes in unknown proportions to each other. These immunoglobulins may be resolved precisely by the technique of immunoelectrophoresis. When a complex mixture of antigens is placed in a well in an agar slab on a glass slide and the latter placed in an electric field, the differences in electrophoretic mobility of each antibody move them to different distances away from the starting well. The fractionated antibodies may then be challenged with the antigen.

This is a good method to separate the proteins in serum. The main proteins in all sera are albumin, and globulins μ, α_1, α_2, β and γ. The γ-globulin fraction migrates towards the cathode (–ve) while the other proteins move towards the anode (+ve).

One modification of immunoelectrophoresis utilizes this last fact in a highly sensitive technique—that of *countercurrent electrophoresis.* The pH of the agar is such that the antibody and antigen used are positively and negatively charged respectively. When current is applied, the materials from the two wells move towards each other, and form a precipitin reaction at a line where the antigen and antibody are equivalent. In contrast to non-electrophoretic double diffusion, the precipitin line is straight and not a curved one. This is due to the fact that the antigen and antibody molecules are moved at a far higher speed in the electric field than in a non-electric one.

9.2.4 Hemagglutination Test

When bacteria are to be identified or red blood cells to be typed, use is made of the hemagglutination test. Cells possess characteristic antigens on their surfaces. If sufficient antibodies against a surface antigenic determinant is added to a cell, the cells are cross-linked by the immunoglobulins via the surface antigens. This principle has been extended for the detection of antigens in solution. The antigen may be attached to red blood cell (RBC) surfaces and antibodies presented to these cells. The cells will be precipitated as a spongy mat if there are sufficient antibodies in the reaction mixture to crosslink most of the RBCs. If the antibody concentration is poor, smaller aggregates form which precipitate as a loose pellet at the bottom of the reaction tube. The latter is customarily a small v-shaped well in a plastic multiwell tray. Synthetic (polystyrene latex) beads may be used in place of the antigen-carrying cells for such a test. The surfaces of the beads can be coated with a desired antigen, and a solution of matching antibody added to them. The beads become cross-linked by the antibodies and are precipitated. The antigen very often is itself an immunoglobulin, an antigenic determinant of which is engaged by the serum (added) antibody.

Agglutination tests are useful for assaying the amount of antibody present in a sample that can cause cross-linking and aggregation. Blood group typing is also a hemagglutination test, that may be carried out rapidly on a microscope slide or in a tube. A drop each of A-group and B-group antigen carrying RBCs is placed on a

slide and the sample blood to be tested added to each. If there is agglutination of cells the blood on the slide and that from the test sample are not compatible. It is possible in this way to characterize the test sample as type A, B, AB or O according to the A-B-O blood group system. The other blood group that requires typing in certain cases is the RH system.

Surfaces of RBCs to be coated with an extraneous antigen require special pre-treatment with tannic acid or chromium chloride in the case of most protein antigens. Polysaccharide antigens may be adsorbed non-covalently merely by incubation of RBCs in a solution of the carbohydrate (polysaccharide).

The hemagglutination test is particularly useful for isolating antibodies (or antigens) that have a very low titre in the antiserum (or sample).

The hemagglutination test may thus be employed to test for (i) antibodies to antigens present naturally on red blood cells, (ii) other antibodies deliberately bound to RBC surfaces, or conversely to assay the amount of antigen present in a sample by challenging the latter with known amounts of the antibody.

9.2.5 Immunofluorescent Techniques

Antigens may be identified by the immunofluorescent technique with the help of antibodies conjugated to a fluorochrome which fluoresces in UV light. The fluorescent dyes generally used are fluorescine isothiocyanate and tetramethyl rhodamine isothiocyanate.

The dye may be bound to an antibody that directly binds to the antigen to be located or indirectly via an antibody to the antigen. The presence of an antigen may also be detected by attaching the fluorochrome to an antibody to one of the serum proteins (complement 3b) that binds to the immunoglobulin Fc region in a cell-bound antigen-antibody complex. Fluorescein gives a greenish colour in UV light.

Immunofluorescent techniques are useful for locating specific antigens (or antibodies) in a cell or tissue. The sample cells or tissues to be tested are frozen to prevent antigens and antibodies from being altered by the subsequent killing and fixing of cells. Live cells may be also distinguished by this technique, on the basis of their surface antigens. The ratio of fluorochrome-conjugated antibody and the protein antigen varies depending upon use with either fixed or live cells; in the former the proportion of dye to protein is much less than when the antigen to be detected is on live cells.

Live cells of different surface antigen specificity may be separated into lots by staining them with fluorescent dye and then sorting them out in a fluorescent activated cell sorter (FACS). The intensity of fluorescence varies with the types of the cells and the type of fluorescent dye used to label them. The FACS separates them on the basis of differences in intensity of fluorescence. This is a fast and efficient method for collecting particular types of cells from a heterogeneous population.

The indirect immunofluorescence test is more sensitive than the direct one. In the direct test only one labelled antibody can bind to an antigen, whereas in the indirect test several labelled antibodies can attach to the same antibody molecule that first binds to the antigen. The indirect test also provides other advantages that makes it a preferred one in many investigations.

A further modification in the protocol of the immunofluorescence test allows the detection of antibodies that are being synthesized in a cell against a particular antigen. The cells are treated with the antigen and the cellular antibody allowed to bind to it. The preparation is then treated with a fluorochrome labelled antibody to the same antigen, and viewed in UV light. In this case the antigen is bound by the cellular antibody as well as the labelled reagent antibody. Needless to say, the protocol involves the removal of all unbound antigens before addition of the anti-antibody to the antigen.

9.2.6 Radioimmunoassay (RIA)

The RIA technique detects and estimates the titre of antigens (or antibodies) with the use of radioactively labelled antibodies (or antigens). The RIA is based on the principle that as both labelled and unlabelled antibodies bind with equal avidity to matching antigens, the amount of radioactive complexes formed will depend on the rates of the labelled and unlabelled component in the mixture presented to the partner. Say an antigen is presented with only labelled antibody (Ab^{l} : Ab^{ul} = 1:0 where 1, ul are labelled and unlabelled Abs), the complexes will register as much radioactivity as the number of antibody molecules bound to antigens. If the challenging antibody solution contains labelled and unlabelled antibodies in equal amounts, only half the complexes will be seen to be radioactive, as the antigens are shared with equal chance by tagged and untagged antibodies. A correlation may be made between the amount of antigen and bound antibody by measuring the inhibition of radioactivity due to the presence of unlabelled antibodies. A standard inhibition curve of radioactivity versus the ratio of labelled and unlabelled antibody in the challenging mixture is made. The RIA is then carried out with the unknown sample. Comparison of the measured radioactivity with values in the inhibition curve provides the titre of the antigen in question. The same procedure may be used to determine the titre of an antibody to a given antigen. In this case use is made of a mixture of labelled and unlabelled antigen.

In either case, one of the two partners for the immune complex is immobilized on a plastic support, such as PVC, polystyrene or dextran, which is usually the inner surface of a tube or the outer one of a bead. The tubes may be small wells in a multiwell dish or tray. Proteins adsorb with ease on these synthetic surfaces.

The RIA is especially useful for estimation of very small amounts of both antigen and antibody. The actual design of the test depends on the goal of the exercise. If the antigen in a sample is to be estimated, it is allowed to incubate in a tube and then discarded. The tube is washed well to remove all unbound antigen. It is customary to add a protein that does not react with the antibody in this first antigen solution. The irrelevant protein covers those patches in the tube surface that are not bound to the given antigen. The tube is next filled with a mixture of unlabelled and radioactively labelled antibody. After incubation for a period that allows the antibodies to bind to the available antigens, the antibody solution is removed and the tube washed thoroughly so that no free antibody remains in the tube. The tube is then tested for radioactivity using the appropriate techniques and the amount of antigen is deduced by comparing the experimental data with the previously obtained standard inhibition curve.

If the antibody is to be estimated, the same procedure is followed with a slight modification. The known antigen is first bound to the surface of the tube, and the antibody solution with unknown titre allowed to react with the antigen as before. An antibody (radioactively labelled) to the Fc region of the first employed antibody is then added to the tube containing the antigen-antibody complexes. The radioactivity of the tube is measured and compared with that of known values in a standard inhibition curve made with labelled and unlabelled anti-Fc Ig molecules.

The titre of an antigen may also be estimated by allowing it to complex with antibody bound to Sepharose (dextran) or polystyrene beads. The latter are first added to labelled antigen solution, removed by centrifugation after complexes are allowed to form, and washed thoroughly before their radioactivity is measured. Again, comparison with an inhibition curve is used to calculate the amount of antigen in the sample.

The RIA is a very sensitive test that can detect and estimate minute concentrations of an antigen or antibody in a very small sample. The antibody used must be very pure; the antigen may be a pure one, or one that is available in a complex mixture.

9.2.7 The Enzyme-Linked Immunoabsorbent Assay (ELISA)

The principle underlying the ELISA and RIA tests is identical. In the RIA, an antigen is recognized and estimated with the help of a tagged antibody or anti-antibody to the antigen. In ELISA, in place of the radioactive label of RIA, an enzyme is used which makes its presence known by catalyzing the conversion of a colourless substrate (chromogen) to a coloured end product. The intensity of colouration after the reaction indicates the concentration of enzymes and in turn the antigens to which the enzyme-linked immunoglobulins have bound. The depth of colour is read in a spectrophotometer at the appropriate extinction wavelength for the coloured product.

The enzyme-antibody conjugate may be prepared in the laboratory, and stored without loss of activity for a year. This feature, together with the lack of any radioactivity hazard, makes the ELISA a technique of choice in many laboratories.

The ELISA test is carried out in the following manner:

1. The antigen is poured into polystyrene tubes (wells) in a multiwell microtitre tray as a buffered solution, and the tubes incubated at 4°C for 2–3 hours. This allows the antigens to become absorbed on the inner surfaces of the tubes (wells). At the same time an irrelevant protein is added.
2. The antigen solution is pipetted out and the wells washed thoroughly with a detergent containing buffer (washing buffer) to remove all unbound antigen: 0.05 per cent Tween in 100 ml of PBS may be used.
3. The wells are then filled with the serum to be tested for the presence of antibodies matching the adsorbed antigen. The test-serum is diluted with the washing buffer. The wells are incubated at room temperature to allow the antigens to be bound to the antibodies in the serum.
4. The test-serum is removed and the wells washed thoroughly to ensure the absence of unbound antibodies in the wells.

5. Now, the enzyme-conjugated anti-serum is added to the wells. The anti-antibody binds to the Fc region of the earlier (bound) Ig molecule (usually IgG), while the trays are kept overnight in the refrigerator (4°C).
6. The solution of the substrate for the enzyme action is prepared and introduced into the wells, which are then kept in the dark for half an hour or so at about 25°C. The reaction is stopped with an appropriate stop-solution.
7. The tubes are measured for absorbance in a spectrophotometer, at the appropriate wavelength for extinction of the coloured end product. The correlation between absorbance and the amount of enzyme is obtained from a previously prepared standard curve using dilutions of a known serum.

The enzyme to be conjugated may be horseradish peroxidase. The substrate for the enzyme is o-phenylene diamine together with hydrogen peroxide in a citrate phosphate buffer (pH 5.0); 12.5 per cent sulphuric acid is used as the stopping reagent. The readings are taken at 492 nm in this case.

9.2.8 Purification by Affinity Chromatography

Specific antibodies may be isolated from a heterogeneous mixture in a serum by adsorbing them on matching antigens, that are bound to solid substrates. The antigen bound substrate is packed in a column and the test-serum poured through it. The wanted Ig molecules bind to the antigen and the unadsorbed ones flow out of the column. The column is washed to clear it of all unbound antibodies and the bound ones, released from the column by an eluting buffer, are collected as a sample of pure antibody. The technique, of course, depends on the extreme specificity or affinity of an antibody for the matching antigen.

When an antigen is allowed to respond to an antibody in solution, the antigen-antibody cross-linked complexes are precipitated. These precipitates may be isolated, and in principle the two proteins (if the antigen also is one) separated by a fairly harsh treatment (e.g., extreme lowering of pH) that distorts their configurations. If the liberated protein(s) are kept exposed to this drastic treatment for long, the reactive configuration of the antibody molecule becomes permanently distorted. Affinity chromatography on the other hand allows the antibody to be released from the bound or insoluble (called immunoadsorbent) antigen in a condition that is still reactive.

The non-reactive support for the antigen may be Sepharose (dextran) beads treated with cyanogen bromide to allow binding of the antigen to the synthetic surface. The cyanogen bromide is a very toxic substance; hence all operations with it are carried out in a fume-cupbard or hood. The elution buffers generally used include glycine-HCL (ph.2.5) diethylamine (pH 11.5), urea, acetic acid and 5 M guanidine HCl. The released antibodies are separated from the elution chemicals by filtration through a biogel or by dialysis.

Affinity chromatography is one of the best techniques for isolating, in a pure form, specific subclasses of Ig molecules, even when present in a very poor concentration in the test sample.

9.2.9 The Complement Fixation Test (CFT)

Antibodies (or antigens) may be detected and estimated in a sample by exploiting a characteristic response of complement to complexes formed by antigens and antibodies. Complement is a set of serum proteins that take part sequentially in a series of reactions which begin with the binding of one (C1) of the set to the Fc region of IgM or IgG. If the antigen is one on a cell surface, the antigen complexed with C1 triggers the formation of a cascade of C3b which bind all over the surface of the cell. Subsequent events lead to the formation of 'holes' or perforations in the plasma membrane and lysis of the cell. If the latter happens to be an erythrocyte (RBC), lysis is detected by the release of haemoglobin by the disintegrated RBC.

The Complement Fixation Test makes use of the above observations to detect the presence of an antibody (or antigen) in a test sample. As complement becomes bound (fixed) to the antigen-antibody complex, it is no longer available for reacting with fresh antigen-antibody complexes. This decrease in the activity of a given concentration of complement is indicated by the number of RBCs that can be still lysed by the unattached or non-fixed complement. The test, therefore, consists of two parts. In the first part the antigen and antibody solutions are allowed to mix. If they match, Ab-An complexes are formed. Complement (in serum preparation) is added to this mixture. Complement will become fixed to the complexes. How much complement has become inactivated may then be deduced from the second part of the test when erythrocytes (preferably especially coated with an antibody) are added to the An-Ab-complex-complement mixture. Any free complement in the mixture will become fixed by binding to the Ig molecules coating the RBCs. These cells will be lysed, and the degree of lysis calculated by comparison with a known or standard solution of the antigen, and complement.

Sheep red blood cells (SRBC) are generally used for the test. They are coated with antibody to their surface proteins, obtained by sensitizing rabbits with SRBC. The anti-SRBC serum is first rid of rabbit blood complement by heating the serum for half an hour at 56°C. These antibody coated erythrocytes or EA are readily attacked by free complement.

The CFT is useful for detecting the presence of pathogens in sample serums (or antibodies matching them). If it is a virus, the antibody to it is added to the sample together with a known amount of complement obtained from, say, the guinea pig. If the virus is present, it will complex with the antibody and induce the complement to become fixed, the amount depending on the titre of the virus in the sample. When Ab-coated SRBC is now added to the mixture, any decrease from the maximum lysis value of SRBCs will reflect the presence and concentration of the virus in the sample. If no virus is present, none of the complement will be fixed and all the SRBCs will be lysed (maximum lysis). Conversely, one can detect the presence of the antiserum to the virus, by using the same technique. Again the extent of lysis of SRBC will indicate the concentration of the antibody in the test serum.

The CFT is used routinely for diagnosis of certain diseases (e.g., the Wasserman test for syphilis).

9.3 ❑ T AND B CELLS

T and B cells are the mediators of the two main branches of the immune system—the *cell mediated* and *humoral.* B cells secrete immunoglobulins synthesized in response to a foreign antigen. T cells respond by synthesizing very similar proteins that remain bound to the cell surface. The antibody can complex with a free non-cellular or cellular antigen. The T cell interacts with the antigen on a cell or body surface in cooperation with the phagocytic white blood cells or macrophages. All three interact in diverse ways to mount characteristic immune responses.

Several types of investigations employ T and B cells. One of them assesses mutagenic damage in these lymphocytes, as an indication of the mutagenic potentiality of an untested substance. Some of the exercises are devoted to separation of these cells from rodent and human blood.

T cells and B cells circulate in the blood and the lymphatic system. The latter consists of a network of lymph channels connected in places with glandular organs. Lymph nodes, lymph nodules without an outer sheath and lymph glands belong to the latter category. The most important lymph glands are the thymus and the spleen. Both T and B lymphocytes originate from precursor small lymphocytes in the bone marrow, in adults and from fetal liver in the embryo stage. The T cell precursors migrate to the thymus where they acquire unique surface antigens, and where they become differentiated into subpopulations with unique functions. The latter include the T_H (helper), T_C (cytotoxic) and T_S (suppressor) cells. The B cell precursor differentiates into immature B cells, which begin differentiation, in response to an antigen challenge, and organize the immunoglobin genes into the blue-print for the antibody that will match the antigen. An Ig molecule is composed of two types of polypeptides (Heavy and Light), the gene for each of which contains several units, each having a number of alternate subunits. During differentiation of a B cell, on antigen provocation, the parts or modules of these variable genes are selected (to match the antigen determinant) and spliced into the sequences that translate into the two antibody chains matching the antigen.

T cells and B cells may be isolated from any of the following sources: peripheral blood, lymph nodes, and the spleen. To obtain an appreciable titre (concentration) of the lymphocytes, it is customary to stimulate their production by immunizing the animal over a period of 4–6 weeks. The antigen is injected at weekly intervals.

T and B cells possess distinctive surface proteins. These are exploited in some protocols to separate the two types of cells from a common pool of lymphocytes isolated from blood or lymphatic glands. B cells have surface receptors which are very similar to the IgM and IgD immunoglobulins. These membrane Ig molecules act as receptors for specific antigens. B cells possess receptors for the Fc region of circulating immunoglobulins. These free Ig molecules in turn possess sites in their Fc that bind to the complement serum protein C3b. One convenient way of separating T and B cells is to mix them with sheep red blood cells or SRBC. It so happens, that there are receptors for SRBC surface antigens on the surfaces of T cells. The SRBCs become attached to the T-cell and form aggregations referred to as 'rosettes'. When SRBC is coated with anti-SRBC immunoglobulins and mixed with T and B cells, the Fc-receptors of the latter bind the anti-SRBC Fc region. You now have a B cell

surrounded by Ig molecules which in turn are attached to SRBC. Such a rosette is called an EA (erythrocyte-antibody) rosette. A third type of rosette is obtained when a receptor binds to the C3b serum protein, which binds to the Fc region of an Ig molecule bound to SRBC. These are the EAC (erythrocyte antibody-complement) rosettes.

Exercise 52: To Raise Antibodies Against a Soluble Antigen

Have ready

1. Sterile Normal Saline+ or sterile PBS*
   *See Appendix
2. Bovine Serum Albumin (Sigma):
   BSA 10 mg
   PBS or Saline 5 ml
3. Freund's Complete Adjuvant — Difco
4. Freund's Incomplete Adjuvant — Difco
5. Na-azide. 0.02% (w/v) in dist. water (stored in deep freeze)
6. A rabbit for immunization
7. Disposable sterile syringes and sterile no. 20 needles
8. Screw-capped sterile vials (15 ml) for storing the anti-serum

Procedure

1. In the LF cabinet,
   (i) Take in a sterile test-tube 1 ml of BSA stock.
   (ii) Add to this, using a no. 18 syringe, drop by drop, 1 ml of Freund's Complete Adjuvant.
   (iii) Aspirate the mixture with a syringe until a very fine emulsion of water-in-oil is formed. If one or two drops of this creamed mixture do not disperse in a test-tube half-filled with saline, the emulsification is complete.
2. Hold the rabbit in a wooden box or wrapped in a blanket and inject 1 ml of the above emulsion intraperitoneally or subcutaneously at 2–3 places on the thigh and back.
3. Repeat Step 2, using an emulsion of 1 ml BSA stock and 1 ml of Freund's Incomplete Adjuvant on the 8th, 15th and 22nd day from Day 0.
4. Rest for 7–10 days.
5. (i) Draw 10 ml blood from the central artery of one ear and transfer to an open centrifuge tube.
   (ii) Place the tube in a slanting position at 37°C for 1 hour, or at 25°C (room temperature for 4 hours).
   (iii) Keep the tube at 4°C overnight.
6. (i) Spin the tube at 1000 rpm (100 g) for 10 minutes.
   (ii) Transfer the straw-coloured supernatant (serum) with a Pasteur pipette into a sterile test-tube.
   (iii) Incubate the test-tube at 56°C (water-bath) for 30 minutes.
   (iv) Add the 0.2% Na-azide solution and mix.
   (v) Distribute into smaller aliquots in sterile screw-capped vials.
   (vi) Store in the deep freeze.
7. Use this rabbit anti-BSA serum in Exercises 55, 56 and 57.

Exercise 53: To Raise Antibodies Against a Cellular Immunogen

Have ready

1. Sterile saline phosphate buffer, pH 7.4 (SPB) (See Appendix)
2. A growing culture of *E. coli* in LB
3. Haemocytometer
4. Freund's Complete Adjuvant (FCA) — Difco
5. Freund's Incomplete Adjuvant (FIA) — Difco
6. Na-azide 20% (w/v) solution in dist. water (stored in deep freeze)
7. Disposable sterile syringes and needles (no. 20)
8. Screw-capped sterile vials (15 ml) for storing the antiserum.

Procedure

1. Harvest *E. coli* cells by spinning in the refrigerated Sorvall at 1000 rpm for 10 minutes.
2. Suspend the pellet in SPB. Spin once more and discard the supernatant.
3. Repeat Step 2 two times more (total 3 washings in SPB)
4. (i) Using the haemocytometer, count cells per ml.
   (ii) Adjust cell count to 10^7/ml.
5. Mix in a sterile test-tube 0.5 ml of cell suspension in SPB, 0.5 ml of FCA into a stable emulsion (see Exercise 52, Step 1).
6. Inject 0.5 ml of the above cells in FCA into 2–3 places in the neck and thigh of the rabbit (intermuscularly).
7. Inject once a week 3 times more, using 0.5 ml of a mixture of 0.5 ml cell suspension in SPB and 0.5 ml of FIA
8. Rest for 7–10 days.
9. (i) Withdraw 10 ml of blood from the central artery of one ear.
   (ii) Place the blood in a 15 ml open centrifuge tube.
   (iii) Place the tube at a slant at 37°C for 1 hour or at 25°C (room temperature) for 4 hours.
   (iv) Leave the tube at 4°C overnight.
10. (i) Spin the tube at 1000 rpm (100 g) for 10 minutes.
    (ii) Transfer the straw coloured supernatant (serum) with a Pasteur pipette into a sterile test-tube.
    (iii) Incubate the tube at 56°C (water-bath) for 30 minutes.
    (iv) Add the 0.2% sodium azide
    (v) Distribute, if necessary, into smaller vials (screw-capped).
    (vi) Store in the deep freeze.
    (vii) Use this rabbit anti-*E. coli* serum to separate different immune complexes formed with *E. coli* membrane proteins.

Exercise 54: To Isolate Leucocytes from Rat Spleens

Outline of procedure

I. Sensitize the animal with antigen 8 days before cell isolation.
II. Kill the animal and dissect out lymph glands.
III. Macerate the glands, to release cells, in a basal salt solution (BSS)
IV. Isolate leucocytes from the cell suspension.

Have ready

1. Three rats (Wistar strain)
2. Antigen (use BSA in FCA as in Exercise 52)
3. A covered bin for discarded animals
4. A large funnel with a cotton plug
5. Ether for anaesthesizing the animals
6. Dissecting instruments: forceps, scalpels, fine scissors
7. A quarto size stiff white paper
8. Hank's Basal Salt Solution (HBSS) in a conical flask kept immersed in crushed ice
9. Crushed or small ice-cubes in an ice-bucket
10. A 90 mm petri dish
11. A fine mesh nylon tea strainer
12. A pestle having a rubber stopper at the pounding end
13. Ammonium chloride, 0.83% in water
14. Four centrifuge tubes (15 ml)
15. Pasteur pipettes
16. Sterile screw-capped vials
17. Haemocytometer
18. High power microscope
19. A timer
20. 70% ethyl alcohol
21. Cotton swabs
22. Sterile Nylon Wool (Leukopac)
23. 5 ml plastic syringe

Procedure

I.

1. Sensitize the rats seven days before the exercise for isolating leucocytes.
   (i) Apply alcohol with a cotton swab at the place to be injected on the abdomen, near the left hind leg.
   (ii) Hold the animal in an inclined position with the head lowered and the ventral side facing up.
   (iii) Lift up the skin with a pair of forceps and insert the needle into the lower skin.
   (iv) Inject the antigen emulsion (1 ml).
   (v) After 7 days contact-sensitize the animals with the same antigen emulsion by smearing (painting) the latter on one ear and the solvent on the other (control).
   (vi) Test for hypersensitivity after 24 hours. Measure the thickness of the ears with a pair of calipers. The sensitized ear would have swollen.

II.

2. Place the rats under the large funnel with a cotton-plug soaked in petroleum ether. This will anaesthesize the animals.
3. Kill the animals by cervical dislocation (pull head and tail in opposite directions with a jerk).

4. Place one animal on a sheet of white paper with the ventral side upwards. Wipe the left side of the ventral surface with 70% ethanol to surface sterilize the skin.
5. Lift the skin with a pair of forceps, and with the scissors cut the top skin, and then the lower flesh, lifting them high, to avoid puncturing the inner organs.
6. The spleen is just below the cut out flesh. It is dark maroon, like the liver.
7. Lift out and severe the spleen with the scissors, and place it on the paper. Snip off connective tissues, blood vessels etc.
8. Partially fill a petri dish (without the cover) with ice-cold HBSS. Wedge a rubber eraser under the petri dish on one side, in order to incline the latter. Place the mesh of the strainer in the liquid in the inclined petri dish.
9. Place the dissected spleen in the strainer so that it is submerged in the HBSS. Similarly remove the spleens from the remaining two rats.
10. Discard the dead animals and their tissues in the waste bin.

III.

11. Wipe the forceps, scissors, scalpels and nylon tea strainer with 70% ethyl alcohol before proceeding further.
12. Cut up each spleen into very small pieces, using scissors and forceps.
13. Using the rubber stopper end of the pestle, macerate the spleen fragments against the nylon net of the strainer. The cells pass into the HBSS. More HBSS may be added from the flask kept on ice. No spleen fragments should be visible after complete maceration. This cell suspension may contain aggregates. These are broken up into single cells by filtering through a nylon-wool column as per Step 14 below.
14. To prepare the column:
    (i) Pack a 5 ml sterile plastic syringe loosely, with sterile nylon wool (300 mg/syringe) with the help of a glass rod and dist. water.
    (ii) Load 5 ml of HBSS on this column (held vertical by a clamp-stand) and allow the fluid to drain out.
15. Add the cell suspension, and allow it to pass through the column. Collect the eluted fluid.
16. Chase the cell suspension with 2 ml of HBSS. Collect this fluid. Pool the fluids from Steps 15 and. 16.
17. Transfer the cell suspension to four centrifuge tubes (15 ml).
18. Spin the tubes at 1250 rpm (150 g) in the clinical centrifuge (kept at 4°C in the cold room) for 70 minutes. Discard the supernatant.
19. Add to each tube 0.83% ammonium chloride to fill 3/4th of the tube.
20. Treat the cells in this ammonium chloride solution for 7 minutes, resuspending them during this period with a Pasteur pipette. (Use the timer.)
21. Spin at 150 g at 4°C for 5 minutes.
22. Resuspend the pellet in HBSS.
23. Repeat Steps 21 and 22 two times more.

24. Pool the pellets in one tube and resuspend the cells in 6–7 ml HBSS.
25. Make a cell count, using a haemocytometer. Adjust the number of cells in the tube to 5×10^7 cells/ml.

Exercise 55: To Isolate T and B Cells from Lymph Nodes of Mice

B lymphocytes adsorb avidly to nylon wool at 37°C in the presence of serum. Advantage is taken of this property to separate T and B cells from the pool of lymphocytes isolated from lymph nodes of mice. Lymphocytes from spleens may also be separated into T and B cells using this method.

Outline of Procedure

I. Sensitize the animals with antigen 8 days before commencement of the exercise for isolation of cells.
II. Kill the animals, dissect the lymph nodes and suspend the lymph node cells in medium.
III. Pellet the cells and suspend in medium with fetal calf serum.
IV. Set up the nylon mesh column.
V. Separate T and B cells.

Have ready

1. Four mice (strains: C57B1, C3H or CBA)
2. Antigen in Freund's CA (see Exercise 52) and syringes for injection
3. A beaker with cotton wool soaked in ether (to anaesthetize the animals), and covered with a glass plate
4. A pair of calipers
5. 70% ethanol and cotton swabs
6. A thermocole (foam) dissecting board: A 9″ × 12″ × 1″ slab of synthetic foam enclosed in a plastic bag
7. A stiff white paper for placing on the dissecting board
8. Sterile dissection instruments (wiped after each use with 70% ethanol): scalpels, scissors, forceps and pins
9. Two 90 mm petri dishes
10. Two plastic tea strainers with nylon sieves: One large and one small
11. A plastic pestle having a rubber pounding end (a rubber cork)
12. Ammonium chloride 0.83% soln. in water
13. MEM medium — (will be called BM, for buffered medium, in this protocol)
14. Pasteur pipettes (sterile)
15. Centrifuge tubes (7 ml)
16. Clinical centrifuge
17. Five plastic syringes (5 ml)
    Four 21 gauge needles
    Six stoppers for open ends of syringe needles, use teflon tubing sealed at one end, or rubber bungs (e.g., erasers from one end of pencils provided with erasers)
18. An arrangement to hold 4 separation columns (syringes) vertical, 12″ from the surface of the table (one control)
19. Sterile Nylon Wool (Leukopac) — 300 mg/syringe

20. Parafilm
21. MEM with 10% FCS (will be referred to as BMFCS) — 20 ml
22. Small beakers with ice and water for holding eluted T cells
23. 25 gauge needles — Four to fit 4 syringes
24. 15 ml centrifuge tubes — six.
25. Haemocytometer.

Procedure

I.

1. Inject antigen on Day 0.
2. On Day 7
   (i) mildly anaesthestize the animals in the beaker containing ether-soaked cotton wool.
   (ii) Measure the thickness of the ear-lobes with a pair of calipers, and
   (iii) Contact-sensitize (paint) one ear lobe per animal with the antigen solution and the other ear lobe (control) with the solvent.
   (iv) Allow the animals to recover.
3. On Day 8, measure the thicknesses of the treated and control ear lobes.
4. Leave the mice in the ether chamber.

II.

5. Place an animal (ventral side up) on a white paper placed over the dissection board and pin the four limbs with pins.
6. Open up the 'skin' and pin it away from the body. Look for the lymph nodes. These will look enlarged due to antigen-treatment, and can thus be easily identified.
7. With a pair of forceps and scissors, snip off at least 6 lymph nodes per animal: (i) inguinal, axillary (ii) neck-cervical, submaxillary, (iii) popliteals.
8. Place each node on a blotting paper square, and remove the bloody and non-bloody connective tissues.
9. Have ready:
   (i) ice-cold BM
   (ii) Petri dish with a large tea strainer and containing ice-cold BM (A).
   (iii) Petri dish with a small tea strainer and containing ice-cold BM (B).
10. Place the lymph nodes of treated animals in (A) and those of control animals in (B).
11. (i) Crush the lymph nodes and macerate tissues against the nylon sieves using the rubber tipped pestle.
    (ii) Filter the cell suspension through a nylon column as in Exercise 54, Steps 14–16 to break up the cell aggregates into a single cell suspension.

III.

12. Transfer the cell suspension in BM to 4 centrifuge tubes (7 ml).
13. Spin in a clinical centrifuge at 2500 rpm for 5 minutes. Discard the supernatant.
14. Add ammonium chloride to fill 3/4th of each tube; resuspend the cells in this solution, keeping the cells in it for a total of 7 minutes.

15. Spin at 2500 rpm for 5 minutes. Discard the supernatant.
16. Resuspend in BM.
17. Repeat Steps 15 and 16 two times more.
18. Check the cell count with a haemocytometer. There should be 4.5 ml of suspension with 5×10^7 cells/ml.

IV.

19. (i) Pack each of 4 plastic syringes (5 ml) with 300 mg of nylon wool. This gives a column volume of about 1.5 ml. Pack the wool under dist. water with a glass rod. One column is for a control.
    (ii) Fix the 4 syringes in the holes in the thermocole slab.
    (iii) Using another plastic syringe, rinse four of the 21 gauge needles with BM.
    (iv) Fix a BM-rinsed needle to each syringe. Make sure that the column and needle are filled with BM and that there are no air bubbles.
    (v) Wash nylon-packed columns (syringes) with 15 ml BM to drive out the dist. water in the following way:
        (a) Let column fill with 15–20 ml of BM and let the BM drip out. Repeat two more times.
        (b) Wash with BMFCS (8 ml). Allow to drip out.
        (c) Add BMFCS (3 ml or twice the column volume) once more.
        (d) Seal the end of the needles with the teflon-tubing caps.
        (e) Take out each 'column' from the holder. Cover the open top of each with parafilm, and place in a vertical position at 37°C (incubator). Incubate for 40–45 minutes.
        (f) Fix the columns to the holder and allow the BMFCS to drain out.

These columns are now ready for use.

V.

20. Place a beaker below each column.
21. Add to each column 1.5 ml of the cell suspension in BM. Remove the teflon cap from the needle.
22. Allow the liquid to drip until the level of the BM is 1 ml above the nylon in the column. Seal the needles with the teflon caps.
23. Keep columns at 37°C for 30 minutes.
24. Take out columns and fix again on the holder.
25. Replace the 21 gauge needles with 25 gauge ones (washed in BM).
26. Add 4.5 ml of the medium (kept at 37°C) to each column (a little at a time); and collect the eluted liquid in a small centrifuge tube held in the beaker containing ice and water. This eluted liquid is Fraction I and contains the T cells, that do not adsorb on the nylon. Label these tubes 'T'.
27. Remove needles, and wash each column with several column volumes of BM. Collect the eluted liquid. This is Fraction II and is to be discarded, as it contains some T and B cells.
28. Remove all of the Fraction II in the following manner:
    (i) Place the column inside a larger test-tube and with the help of a syringe piston, squeeze the nylon packing.

(ii) Pull up more BM from a petri dish and squeeze out again.
(iii) Repeat (ii) above a few times.

29. Now squeeze out BM inside a fresh test-tube. There will be about 1–1.5 ml of the liquid. Pull up a little more BM and squeeze out into the above fresh tube. Repeat this a few times. The liquid in the tube is Fraction III and contains the B cells; label these tubes 'B'.
30. Spin both T and B suspensions at 2500 rpm for 5 minutes.
31. Resuspend the pellets in BM.
32. Make cell counts of:
    T cells in Fraction I and in control Fraction I
    B cells in Fraction III and in control Fraction III.

Exercise 56: To Isolate Human T and B Cells

The different types of cells in blood differ in their densities and can be separated by isopycnic sedimentation. This means that when the cell mixture is centrifuged in a density gradient, the different types of cells collect at the layers that have the same density as those of the cells. Leucocytes and erythrocytes were at one time separated by centrifuging in albumin solutions which had densities intermediate between those of the red and white blood cells. The technique has improved with the use of other high molecular weight substances, such as methyl cellulose and products with trade names that include Dextran, metrizamide (Nygaard), Ficoll (Pharmacia) and Percoll (Pharmacia). Percoll is a preparation of colloidal silica. Both Ficoll and Percoll can be autoclaved and hence are used frequently. Very rapid sedimentation of leucocytes is achieved by density gradient centrifugation in mixtures of Ficoll and a radio-opaque substance (metrizoate-or its derivative). Combinations of Ficoll and Hypaque (sodium metrizoate) are sold under different trade names. Ficoll-Hypaque (Pharmacia, Flow Laboratories, Sigma) has a density of 1.114 g/ml and separates the lymphocytes as a distinct layer between the densities of the two isopycnic substances. The erythrocytes pellet at the bottom of the centrifuge tube. The 'buffy coat' that forms a distinct layer midway in the tube consists almost exclusively of lymphocytes.

We will use Ficoll-Hypaque isopycnic density centrifugation to isolate human T and B cells from peripheral blood.

Have ready

1. SRBC in Alsevier's solution (2×10 T cells/ml)
   (a) Alsevier's solution from suppliers. Keeps well for about 4 weeks at 4°C

   or

   (b) Prepare:
      (i) Alsevier's solution:

Dextrose	2.05 g
Sodium citrate	0.80 g
Sodium chloride	0.48 g
Citric acid	0.05 g
Glass dist. water	100 ml

(ii) Collect 10 ml sheep's blood (jugular vein) in a heparinized syringe and add this to 10 ml Alsevier's solution in a conical flask, store at 4°C. Use within 7 days.

2. Human peripheral venous blood collected in heparin (see Exercise 48).
3. Sterile normal saline:

NaCl	8.5 g
Dist. water	1000 ml

4. (i) Ficoll-Hypaque (Sigma) comes as a sterile liquid in 100 ml bottle packs. Its density is adjusted for optimal separation of lymphocytes and erythrocytes.
   (ii) Ficoll-Hypaque gradient solution may be prepared in the laboratory:
      (i) Prepare a 9.2% aqueous soln. of Ficoll.
      (ii) Mix Ficoll soln. 43.3 ml and Hypaque 6.6 ml
      (iii) Use the Ficoll-Hypaque after filter sterilization through a 0.22 μ size filter (Millipore).
5. Culture Medium: TC 199
   (other media that may be used: RPH1, 1640, Dulbecco's or BME 4 X).
   See Appendix for formula of the tissue culture medium, TC 199
6. Fetal Calf serum (FCS)
7. PBS pH 7.2 (Phosphate buffered saline):

Soln. A: 0.15 M Na_2HPO_4, $2H_2O$

$Na_2HPO_4 \cdot 2H_2O$	26.724 g
NaCl	8.000 g
Glass dist. water	1000 ml

Soln. B: 0.15 M KH_2PO_4

KH_2PO_4	20.413 g
NaCl	8.000 g
Glass dist. water	1000 ml

Mix to make PBS pH 7.2 (fresh).

Soln. A	72.0 ml
Soln. B	28.0 ml

8. Anti-SRBC antibody (Haffkine Institute, Bombay)
9. Methylene blue stain
10. Table-top, swing-out type centrifuge
11. Haemocytometer
12. Complement (use guinea-pig serum as source of complement) See below (A) for preparation of serum for use of its complement, and (B) for standardization of complement for use.

(A) To Prepare Complement (Serum) for Use

1. Use human or guinea-pig blood 20 ml (non-heparinized).
2. Clot the blood at 4°C for 4 hours or about 1 hour at room temperature.
3. Spin and transfer the supernatant (serum) to a small screw-capped vial. It may be stored in the deep-freeze and thawed before use.
4. Add 0.1 ml of SRBC pellet to 4.0 ml of the serum and keep at 0°C for 15 minutes.

5. Spin at 150 g for 10 minutes at 0°C.
6. Take the supernatant and repeat Steps 4 and 5 two times more.
7. Use this serum which is now free from natural hemolytic antibodies.

(B) To Standardize Complement on the basis of SRBC Lysis

I. Prepare complement dilution medium (CDM)
II. Prepare SRBC in CDM (5% suspension)
III. Add different dilutions of complement (serum) to SRBC suspension and measure the amount of RBC lysing by a spectrophotometer. Use the serum that gives 50% lysis (Complement Hemolysis or CH_{50}).

I. To prepare CDM (5 X)

(a)	NaCl	85	g
	Na-5 diethylbarbiturate	5.75	g
	Triple dist. water	500	ml
(b)	5-diethylbarbituric acid	5.75	g
	Triple dist. water	1400	ml
(c)	$CaCl_2$	0.3	M
	$MgCl_2$	1.0	M

(c) stock solution

(d) Mix (a) and (b) and add 5 ml of (c)
(e) Make up the volume to 2000 ml with triple dist. water.
(f) For use: Dilute the 5 X stock to a 1 X soln. with water.

II. To Prepare SRBC for Use

1. Spin SRBC in Alsevier's soln. at 3000 rpm for 5 minutes.
2. Remove the top layer plasma and the buffy coat (leucocytes), and the bottom layer. Leave the pellet.
3. Resuspend the pellet in CDM solution, spin, and discard the supernatant. Repeat this Step 3 times. The SRBC are now washed.
4. Resuspend the SRBC in the CDM to give a 5% suspension.
5. Place a loose plug (small, of sterile cotton) in a glass funnel and filter the SRBC suspension. Use the filtrate.
6. Mix 1 ml of this suspension and 14 ml of triple dist. water.
7. Measure the OD of the above at 541 nm. Adjust OD to 0.7 using sterile dist. water. This suspension has 1×10^9 SRBC/μl.

III.

1. Clear the serum of naturally occurring hemolysing antibodies (as given in A).
2. Dilute the serum with CDM (1:50).
3. To each of 6 tubes add 1 ml of antibody coated SRBC. (See Steps 1 and 2 in the procedure for making EAC Rosettes given later in this exercise).
4. To no. 1 tube add 5 ml of serum
   no. 2 tube add 2.5 ml of serum
   no. 3 tube add 1.5 ml of serum
   no. 4 tube add 1.0 ml of serum
   no. 5 tube add 6.5 ml triple dist. water
   no. 6 tube add 6.5 ml CDM

tube no. 7 is a control with 5 ml of CDM and 2.5 ml of serum. This has a dilution of 1:20.

5. Measure OD of tube no. 7 (blank).
6. Calculate OD of serum in each tube.

$$\text{OD of serum dilution} = \frac{20}{\text{reciprocal of ser. dil.}} - \times \text{OD of blank}$$

7. Now correct this value:
   Corrected OD = (OD of sample) – (OD of no. 6 + OD of ser. dil.)
8. Calculate lysis of SRBC (%)

$$\text{Lysis \%} = \frac{\text{corrected OD}}{\text{OD of no. 7} - \text{OD of no. 6}}$$

9. Use the sample that just reaches 50% lysis. (Plot a graph on a log-log paper with lysis on the *x*-axis and logarithm of reciprocal complement dilution on the *y*-axis.)

Procedure (see end of the protocol for modifications)

1. Collect peripheral blood (10 ml) in a heparinized tube. Tilt the tube gently to mix the blood with the heparin (anti-coagulant).
2. Mix the heparinized blood with 5–8 ml of normal saline in a conical flask (50 ml).
3. Take 2.5–3 ml of Ficoll-Hypaque (density adjusted) and distribute it in two centrifuge tubes.
4. Tilt each centrifuge tube and introduce half of the heparnized diluted blood from the flask, with a Pasteur pipette, gently down the side of the tube. The blood is now loaded on the Ficoll-Hypaque solution.
5. Spin at 1520 rpm (400 g) for 20 minutes. The lymphocytes form a layer at the interface of the plasma and hypaque layers.
6. Using a Pasteur pipette, carefully remove the dilute plasma layer and discard it.
7. With the pipette, gently remove the 'ring' of lymphocytes and transfer to a centrifuge tube. (15 ml).
8. Add 2/3 of tube of normal saline and spin at 1000 rpm for 5 minutes. Discard the supernatant. Resuspend the pellet in saline.
9. Repeat Step 8 two times more.
10. Resuspend the cells in medium (TC 199), supplemented with 10% FCS. Cell count should be 4 × 10 /ml of medium.
11. Prepare SRBC for use:
    (i) Spin in Alsevier's solution (10 ml)
    (ii) Resuspend the pellet in normal saline (10 ml) or PBS pH 7.4. Spin, wash the pellet once more in HBSS (10 ml). This is a 1% suspension of SRBC.

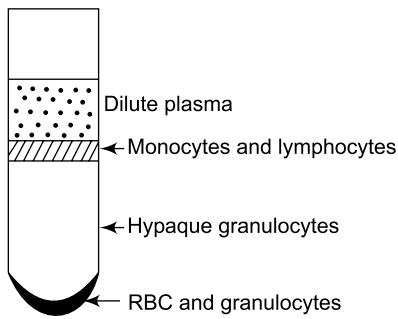


Fig. 9.1 *Lymphocytes may be isolated by centrifuging heparinized dilute blood in a Ficoll-Hypaque Solution.*

12. Mix gently in a sterile Al foil capped centrifuge tube: 0.1 ml of lymphocyte suspension and 0.1 ml of 1% SRBC.
13. Incubate at 37°C (incubator) for 10 minutes.
14. Add a drop of methyline blue to stain the cells.
15. Spin at 500 rpm for 5 minutes.
16. Keep the tube at 4°C overnight.
17. Pick up a small amount of the pellet with a pipette and place it on a haemocytometer or microscope slide. Cover with coverslip.
18. Observe under the microscope and identify the rosettes. Count at least 200 lymphocytes and the number of rosetted and of non-rosetted cells. These are the T rosettes (EA). The cell suspension now contains T rosettes and B cells.
19. Layer this suspension carefully over Ficoll-Hypaque as before.
20. Spin at 1,520 rpm (400 g) for 20 minutes. The T rosettes precipitate out, and the B cells remain at the gradient-medium interface.
21. Transfer the supernatant to another tube. This contains the B cells.
22. The T cells can be freed of the SRBC by lysis of the latter:
    Treat T rosettes for 8 minutes at 37°C in NH_4Cl 0.83% and in Tris buffer 0.01 M, pH 7.2.
23. Spin at 500 rpm. Discard the supernatant. Resuspend the T cells in TC 199, make a cell count.

To Make EAC Rosettes with B cells

1. Coat SRBC with anti-SRBC serum in the following manner:
   (i) Spin the SRBC in normal saline at 300 g for 5–8 minutes.
   (ii) Discard the supernatant and resuspend the SRBC in normal saline. Repeat (i) and (ii) two times more. The SRBCs are now washed in saline.
   (iii) Resuspend the cells in TC 199 with a cell density of 2×108 ml.
      (a) Mix 0.5 ml SRBC in TC 199 with 0.08 ml anti-SRBC serum.
      (b) Mix 0.5 ml SRBC in TC 199.
   (iv) Add the following amounts of anti-SRBC serum to each of 0.5 ml SRBC in TC 199.

I	SRBC 0.5 ml	20% antiserum
II	0.5 ml	10% antiserum
III	0.5 ml	2.5% antiserum

2. Spin each tube at 1250 rpm (120 g) for 5 minutes. Discard the supernatant. Resuspend in TC 199 and wash again two times more. Finally suspend the cells in 1 ml TC 199.
3. Check a drop of suspension from each titration tube (I, II, III) and disregard those tubes which show agglutination (clumping of SRBC). We require Ig coated SRBC and not agglutinated ones.
4. Mix with aliquots of the correct Ig coated SRBCs, different dilutions of the complement solution. Observe the treated material under the microscope. The dilution which does not give cell lysis is the correct one to use. The SRBCs are now coated with Ig molecules to which complement has been bound (C3b). This is the EA complex for EAC suspension.
5. Mix 0.1 ml of B cell suspension in TC 199 and 0.1 ml of EAC complex suspension.

6. Add a drop of methyline blue stain and incubate at 37°C (incubator) for 15 minutes.
7. Layer this mixture on Ficoll-Hypaque and spin at 400 g for 20 minutes. The B-EAC rosettes precipitate out.
8. Count the number of rosetted and non-rosetted B cells.
9. You may lyse the SRBC with 0.83% ammonium chloride in TC 199 to release the B cells.
10. Wash the cells 3 times in TC 199.

Modification of the above Method of Isolation of T cells T cells form more stable rosettes with SRBC coated with neuraminidase, which is available from Sigma Chemical Co.

1. (i) Prepare a 5% SRBC in TC 199.
   (ii) Add to 1 ml of 5% SRBC 0.1 ml Neuraminidase (25 units/ml).
   (iii) Keep at 37°C (water-bath) for 1 hour.
   (iv) Add PBS (5 ml) and spin at 150 g for 5 minutes. Discard the PBS.
   (v) Repeat Step (iv) two times more.
   (vi) Resuspend the pellet in TC 199 to give 5% SRBC in TC 199.

You now have a 0.5% neuraminidase treated SRBC.

2. Mix 1 ml of SRBC and 1 ml of lymphocytes.
3. Incubate at 37°C (incubator) for 10 minutes.
4. Spin at 200 g for 5 minutes.
5. Keep the vial in an ice-bath for 15 minutes.
6. Layer the mixture on Ficoll-Hypaque gradient and spin at 400 g for 20 minutes.
7. Remove the supernatant carefully and discard the same.
8. Remove the cells at the interface of Ficoll-Hypaque and supernatant, and transfer to a tube containing 5 ml TC 199.
9. Remove the lower layer (Ficoll-Hypaque). Retain the pellet.
10. Add a few drops of TC 199 to the pellet and examine a drop of the pellet suspension under the microscope. There will be T rosettes.

Exercise 57: The Antigen-Antibody Response (I) The Precipitin Test

When a polyvalent antigen solution is mixed with its matching antiserum (serum with antibodies), the two proteins combine to form a net-like spongy aggregate that precipitates out. When there is more antigen than antibody in the mixture, the antigens usually form complexes that remain in solution. When the antibody is in excess, the complexes formed are insoluble. Some human sera form soluble complexes even in the presence of excess antibody. When the amounts of antibody and antigen are equivalent, there is a rapid aggregation and precipitation of the complexes.

Thus by adding an equivalent volume of an antigen to that of its antisera, one obtains a preliminary idea about the ratios of the matching proteins in the mixture. If one solution is layered carefully over the other, within one day a thin cloudy line will be seen at the junction of the two, if the antiserum matches the antigen. A quantitative estimate of the amount of antibody in the serum is obtained by adding

different dilutions of the antigen to the same volume of serum in a number of test-tubes, incubating the latter at 37°C and then at 4°C. The precipitated complexes may be centrifuged out, washed in saline and tested for the quantity of antibody in the complexes.

In the following exercise we will observe (i) the formation of the antigen-antibody-complex-ring at the interface of the two protein solutions, and (ii) discover the antigen quantity that precipitates all of the antibody in one of several mixtures of the two proteins.

Have ready

1. Blood (5 ml) from ear artery of a pre-immunized rabbit
2. BSA (500 μg/ml of normal saline) to immunize the rabbit over 4 weeks (as in Exercise 52)
3. BSA (20 mg/ml) in normal saline, as the challenging antigen
4. Sterile fine tipped Pasteur pipettes
5. Small, rimless test-tubes, 40 × 3 × 1.5 mm (Durham tubes)
6. Sterile syringes, needles etc.

I. To Test for the Presence of Antibody in a Serum

Day 1

1. Collect 5 ml blood from the pre-immunized rabbit in a screw-capped vial. Close the cap.
2. Keep the blood at 37°C (incubator) for 40–80 minutes.
3. Transfer the blood to 4°C. Keep for 16–18 hours.
4. Remove the serum into a fresh vial and heat it at 56°C for 30 minutes, to inactivate the complement. Otherwise, these will adhere to the complexes and precipitate with them.
5. Store the serum in the deep freeze. Thaw the serum before use.

Day 3

6. Collect 5 ml blood from the immunized animal and separate the serum as in Steps 2 to 4.
7. Introduce into two Durham tubes 1 ml of the antiserum (from Step 6) and into two other tubes, 1 ml of serum from Step 5.
8. Now carefully overlay on each serum 1 ml of a solution of BSA in saline.
9. Leave the tubes in the stand at 25°C (room temperature) for 1 hour and at 4°C for 15–18 hours.
10. Observe if any 'ring' has formed in any of the four tubes. What is your explanation of the results you see?

II. To Make a Crude Estimate of the Amount of Antibody Present in the Antiserum

1. Arrange in a row (in a stand) 7 Durham tubes, labelled normal serum, control, 15 μg BSA, 30 μg BSA, 60 μg BSA, 90 μg BSA and 120 μg BSA.
2. Add to the first tube, serum taken from the preimmunized animal.
3. Add 0 μg BSA to the control tube and 15 μg ... 120 μg of BSA (challenging antigen) in saline to the tubes labelled correspondingly.

4. Add normal saline to the tubes to make the volumes equal.
5. Incubate the tubes at 37°C for 1 hour.
6. Keep the tubes at 4°C (cold room) for 1–7 days, mixing the material in the tubes twice or thrice a day (by inverting the tubes a couple of times).
7. Observe the tubes every day, and note the formation of precipitates. (Record notes on the date, tube, comments about the quantity of the precipitate.)
8. Spin the contents of each tube at 2500 rpm at 4°C for 30 minutes (clinical centrifuge in the cold room). Do not discard the supernatant. Transfer the supernatant to the labelled tubes (see after Step 10).
9. Resuspend the pellet in 0.5 ml of chilled normal saline. Wash down the sides of tube with more chilled saline (1.5–2 ml).
10. Spin again and wash the pellet with chilled saline two times more. Now check the supernatants for the presence of antibody or antigen.
11. Divide each supernatant into two tubes. To one add the antigen solution, and to the other an equal volume of serum. Incubate at 37°C for 1 hour and at 4°C for a few days to see if any precipitate is formed.
12. The tubes with no precipitate represent the aliquots in which the antigen and the antibody amounts were equivalent.
    The amount of antibody in the precipitate may be determined by several methods. A simple one is to:
    (i) Check the OD of the BSA at 280 nm.
    (ii) Dissolve the precipitate in about 2 ml of 0.1 N NaOH, and check the OD of this solution at 280 nm.
    (iii) Calculate the OD due to the antibody

    $$OD(Ab + Ag) - OD(Ag) = OD(Ab)$$

    given that 1 mg of antibody is equivalent to an OD_{280} value of 1.46.

Exercise 58: Estimation of Antigen-Antibody Response by the Immunodiffusion Technique

The titre of antisera may be obtained by immunodiffusion techniques. As in the case of the precipitin test, the antigen and antibody are allowed to interact and form complexes at the interface of the two reactants. However, in immunodiffusion tests, the interfacing occurs between antigen and antibody that diffuse through a solid medium, such as an agar or agarose gel.

In the earliest forms of such tests the antibody was dispersed in the gel and the antigen placed in wells cut out from the gel. The antigen diffuses out from the well and interacts with the antibody to form complexes. The latter appear as a ring of precipitates around the well. The width of the ring may be used to calculate the amount of antibody relative to the antigen present in the gel. The ring appears as an opaque band in a gel illuminated against a dark background. The ring may be also visualized by staining the precipitates with a protein stain, such as Coomasie Brilliant Blue.

The above mentioned single diffusion technique has been superseded by a double diffusion one developed by Ouchterlony. In this case, the antigen is placed in a well with the antisera in wells surrounding it or vice versa. The radially placed wells may

contain different antisera, or different dilutions of the same antibody. The advancing wave-fronts of the diffusing antigen and antisera form precipitated complexes wherever they meet. The precipitate is soluble if there is more antigen than antibody. Hence a very sharp precipitin line is made at the zone where the amounts of antigen and antibody are equivalent. The relative distance of the precipitin line from the two reactants indicates their relative strengths.

Have ready

1. Microscope slides
2. A frame for holding 6 or 12 slides, and a device for levelling the frame
3. Glass or metal tube to punch holes in the agar (diameter 2–3 mm)
4. Agar (Difco or Noble agar, a special quality agar used for immunoelectrophoresis)
5. Normal saline (0.85% NaCl in water)
6. Antigen (BSA in PBS). Make the following dilutions: 1.0 mg, 0.5 mg. 0.25 mg, 0.125 mg per ml.
7. Antisera (raised in rabbit):
   (i) Add to 1.9 ml warm PBS (56°C) 0.075 ml antisera.
   (ii) Dilute the above into the following: 50, 100, 150, and 200 μg/ml
8. Staining Solution:

Coomassie Brilliant blue 250 (BDH)	1.25	g
Glacial acetic acid	50	ml
Dist. water	185	ml

9. Destaining Solution:

Glacial acetic acid	100	ml
Dist. water	370	ml

10. Whatman's filter paper

Procedure

I. Preparation of Slides

1. Prepare 50 ml each of a 1% and a 1.5% agar solution in normal saline. Heat to 100°C and then maintain it at 56°C until use.
2. Arrange 6 slides per frame, and level the latter.
3. Brush a fine film of 1% agar on each slide and allow it to set.
4. Take 1.5% agar in a Pasteur pipette and dispense 2–4 ml on each pre-coated slide. Allow to set.
5. Make holes in the solid agar as per the following pattern:

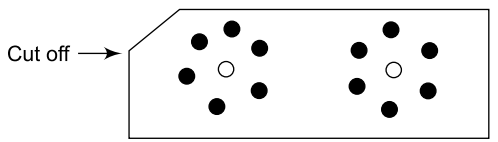


Fig. 9.2 *Patterns of wells cut out in the agar of an agar-coated microscope slide.*

   Cut off one corner for identification of the correct orientation.
6. Remove the plugs using a pair of dissecting needles or a Pasteur pipette. Be sure to keep the sides of the wells vertical.

II. Setting up the Test

7. Place 10 µl of the reactant per well in the following patterns:

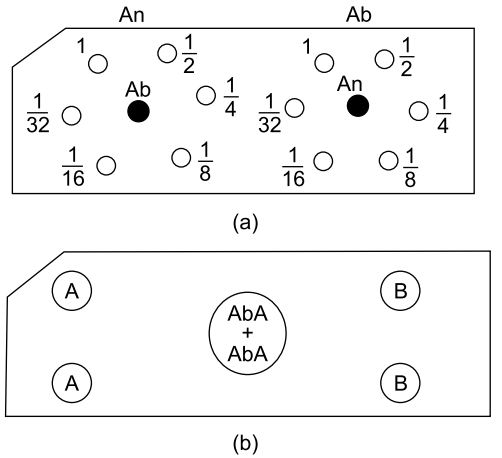


Fig. 9.3 *Patterns (a) and (b) for placing antigen (An) and antibody (Ab) dilutions in wells for immunodiffusion-based tests.*

8. Lift up the frame and place it in a moist chamber. The chamber can be a large petri dish or a plastic covered box, lined at the bottom with moist paper towel or tissue.
9. Leave undisturbed overnight or for longer periods at 25°C (room temperature) or for 2–4 hours at 37°C (Incubator).

III. Staining the Protein Complexes

10. Wash each slide in several changes of PBS spread over one day.
11. Place a stack of 3 or 4 rectangles of filter paper over each slide and leave overnight. This removes the excess solutions. Remove the filter papers.
12. Place the slide in a petri dish containing the staining solution. Stain for about 5 minutes.
13. Decant the staining solution into a beaker. (This stain may be reused.)
14. Wash the slide with changes of the destaining solution until only the reaction lines remain.

IV. Observations

15. Record your observations:
    (i) Is there any reaction between the central well and the outer wells?
    (ii) If so, what is the nature of the stained zone? Is it (a) a single line, (b) a curved line, (c) a curved line with spurs, or (d) lines which cross each other? (see Fig. 9.4).

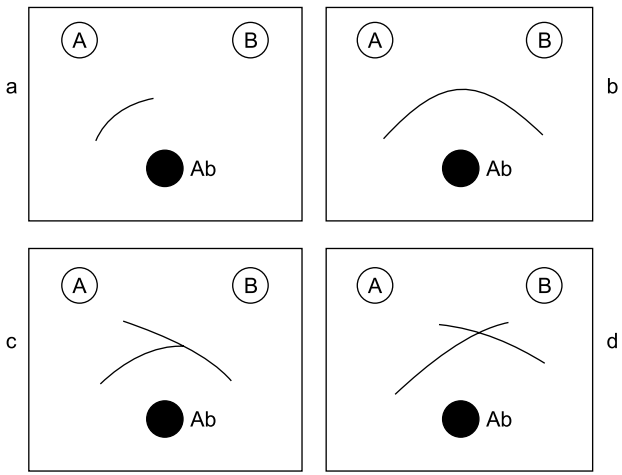


Fig. 9.4 *Degrees of identity between two antigens, A and B, may be established by allowing them to react with an antibody to one of them placed equidistant from the two antigens. See text for explanations of the curved and spurred reaction lines as shown in this figure.*

The following interpretations may be made (refer Fig. 9.4).

(a) Only antigen A reacts with the antiserum.
(b) Both antigens A and B have common antigenic determinants that react with the antiserum.
(c) Besides the determinants common with A, there is another determinant in B that has a matching antibody in the antiserum.
(d) A and B possess nonidentical antigenic determinants, the antibodies against which are present in the antiserum.

If there is no blue/opaque line between the antiserum and an antigen, the former has no antibody corresponding to the latter. By judicious combinations of different antisera and antigens, it is possible to establish the identity and non-identity between different antigens as recognized by an antiserum.

Exercise 59: Estimation of Antigen-Antibody Response by Immunoelectrophoresis

Immunodiffusion and electrophoresis may be combined to assess antibody-antigen complex formation. Normally, antigens are mixtures of proteins or determinants. If such a mixture is used to immunize an animal, the antibodies formed against each antigenic component (and present in the serum of the immunized animal) can be resolved precisely by the technique of immunoelectrophoresis.

This is a good technique to separate the proteins in sera. The main proteins in human serum are albumin, and globulins α_1, α_2, β, μ and γ. When serum is placed in a well in a gel on a microscope slide, and the latter is placed in an electrophoretic apparatus, these proteins, except IgG (γ), migrate away from the well in the order given above. The IgG migrates towards the cathode (–ve); the others move towards the anode (Fig. 9.5). If a little bromophenol blue dye is added to the serum, it binds to the albumin, which moves the fastest. When the blue band almost reaches the end of the slide, the current is discontinued.

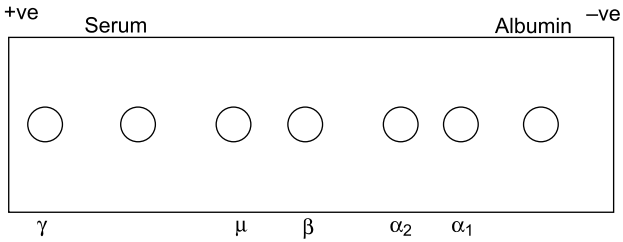


Fig. 9.5 *Fractionation of the main proteins, albumin and globulins (α, β, γ, μ) of serum, by the immunoelectrophoretic technique.*

Some white blood cell cancers known as myelomas produce only monoclonal antibodies. To find out which immunoglobulin is present in the serum of a myeloma patient, serum from both a normal human and the patient (NHS and PS) are fractionated on the same slide. The band of PS which corresponds with that in the NHS channel identifies the monoclonal globulin in the PS.

Have ready

1. Agar 1.5% made in barbitone buffer
2. Barbitone buffer: Make 500 ml
   (i) Dissolve:
   Barbital* 4.40 g
   in dist. water 150 ml at 95°C
   (ii) Dissolve:
   Na-barbital** 12.0 g
   in dist. water 800 ml
   (iii) Mix (i) and (ii).
   (iv) Adjust pH to 8.2.
   (v) Make up the total volume to 1000 ml with dist. water.
   *5, 5-diethylbarbituric acid
   **Sodium salt of the above acid
3. Microscope slides
4. Normal Human Serum (NHS)
5. Serum from a multiple myeloma patient (PS)
6. Coomassie Brilliant blue stain (see Exercise 58)
7. Glacial acetic acid 10% (v/v) in water
8. Device for punching holes in the gel (see Exercises 58)
9. Horizontal Electrophoresis apparatus
10. Two 8 cm × 2 cm strips of Whatman's filter paper

Procedure

1. Prepare a 1.5% agar solution in the buffer.
2. Prepare a 1% agar solution in dist. water
3. Brush each slide with 1% agar. Allow it to set.
4. Dispense 2–3 ml of the 1.5% agar over the pre-coated slide. Allow it to set.
5. Punch two small holes (diameter 1 mm) in the gel (NHS, PS = normal human serum and patient serum respectively).

6. Place two slides in the gel trough of the electrophoresis apparatus.
7. Fill the two tanks with the buffer.
8. Connect each short end of a slide with the buffer trough on its side with a filter paper strip (wick).
9. Add a drop of bromophenol blue solution to each serum.
10. Fill the two wells with the NHS and PS as shown in Fig. 9.6.

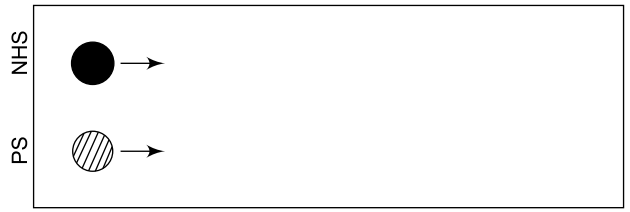


Fig. 9.6 *A comparison of the serum proteins in normal human serum (NHS) and patient serum (PS) by using the technique of immunoelectrophoresis.*

11. Turn on the current (16 amps) for about 1 hour. The blue-albumin band should be near the end of the gel on the anode side, by this time.
12. Stop the current and cut out a long trough with a sharp blade as shown in Fig. 9.7. Remove the gel from this rectangular trough.

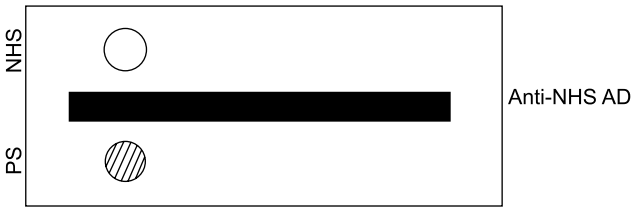


Fig. 9.7 *Immunoelectrophoresis is used to identify the similar and dissimilar serum proteins in NHS (normal human serum) and PS (patient serum) by allowing the electrophoresed serums to react with the antiserum to the NHS.*

13. Fill the trough with anti-NHS (raised in rabbit).
14. Place the slide in a humid chamber (see Exercise 58) for 12 to 18 hours at 25°C or about 4 hours at 37°C (incubator).
15. Stain the slides with Coomassie Blue and destain as in Exercise 58.
16. Record your observations about the serum proteins of NHS and PS that have matching antibodies in the anti-NHS serum.

Exercise 60: To Purify Immunoglobulin Molecules of Mouse

Most proteins dissolve in pure water or salt solutions. The proteins of serum include albumin and the immunoglobulins. The latter may be precipitated in an ammonium sulphate solution.

IgG of mouse may be obtained from mouse serum by the following exercise.

Have ready

1. Solid ammonium sulphate
2. Dil. ammonia
3. Ether to anaesthetize mice
4. Rabbit
5. Centrifuge tubes
6. Clinical centrifuge; Sorvall centrifuge
7. Normal saline (0.85% NaCl in water)
8. PBS (Phosphate buffer) 0.01 M, pH 8.0
9. Dialysis tube
10. 18 gauge needle and syringe
11. Screw-capped bottle for blood collection
12. DEAE (diethyl aminoethyl) Cellulose (DE 52, Whatman) 100 g
13. Whatman's Filter paper, No. l
14. Buchner Funnel
15. PBS: NaH_2PO_4, pH 1.0 + 0.15 M NaCl
16. Spectronic 21 (UV spectrophotometer).

Procedure

I. To Prepare a Crude IgG Fraction

1. Collect in the bottle about 40–50 ml blood from the central artery of the ear using the syringe and 18 gauge needle.
2. Let the blood clot in the bottle at room temperature (25°C) overnight.
3. Decant the serum into a conical flask. Keep at 0°C.
4. Add solid ammonium sulphate to the serum so that there is a concentration of 75% of the salt in the fluid. Stir continuously while adding the salt.
5. Stir for 60 minutes at 25°C.
6. Pour into centrifuge tubes and spin at 5,000 g for 20 minutes.
7. Resuspend the pellet in 1.75 M ammonium sulphate solution and spin again.
8. Discard the supernatant and wash the pellet two more times with 1.75 M ammonium sulphate solution.
9. Resuspend the pellet in 10 mM NaH_2PO_4 (pH 7.0).
10. Pour into a washed dialysis bag and hang overnight in distilled water, in the cold room (4°C).
11. Pour the liquid from the dialysis tube into the centrifuge tube and spin at 5000 g for 20 minutes.
12. Pour the supernatant into a fresh dialysis tube and hang overnight in 10 mM NaH_2PO_4 (pH 8.0) at 4°C.
13. Purify the IgG by passing the dialyzed fluid through DEAE cellulose (an anionic exchanger).
    (i) Equilibriate the DEAE cellulose by mixing in a flask 100 g DE 52 (Whatman) in 550 ml 0.01 M phosphate buffer (PB) pH 8.0.
    (ii) Adjust pH of mixture to 8.0 (use 1 N HCl)
    (iii) Carefully pipette off the supernatant (This will remove very fine particles of the DE).

(iv) Fill the flask with PB. Allow the cellulose to settle. Then remove the supernatant.
(v) Repeat Step 13 (iv) once more.
(vi) Weigh the equilibriated cellulose.

14. For every 10 ml of starting serum you may use 50 g (wet weight) of the equilibriated DE slurry.
15. Mix the dialyzed fluid with the slurry (chilled at 4°C) and stir intermittently for 1 hour in the cold room (4°C).
16. Filter the mixture through a Buchner funnel (lined with 2 layers of Whatman's filter paper No. 1). Collect the eluted fluid in a flask.
17. Add 20 ml of 0.01 M PBS (pH 8.0). Collect the effluent.
18. Repeat Step 17 two more times. The pooled effluents contain IgG. The other Ig molecules remain bound to the cellulose.
19. Add ammonium sulphate to make a 75% solution in the effluent (at 0°C).
20. Pour into centrifuge tubes and spin at 10,000 g for 10 minutes. The pellet is pure IgG.
21. Dissolve the pellet in PBS.
22. Check protein content by using the UV spectrophotometer (Spectronic 21) OD_{280} = 1 is equivalent to 0.69 mg/ml of IgG

Exercise 61: To Assay the Number of Cells Secreting Ig Molecules

The number of B cells in a sample secreting immunoglobulins may be estimated in a quantitative manner by a method originally developed by Jerne, Nordin and Henry (1963). The following exercise will quantitize the number of B cells from the spleen of an animal sensitized with SRBC, that are producing anti-SRBC immunoglobulins. It requires one molecule of IgM but two of IgG to start the cell-lysing cascade activities of complement. There are thus two ways in which these predominantly occurring Ig molecule secreting cells may be detected; these are the direct and the indirect methods. Both IgM and IgG secreting cells may also be counted by using a third technique that combines the logic of the first two. We will assay the antibody secreting cells, (with respect to IgM and IgG) using the direct and the indirect methods.

Have ready

1. Mice for immunization
2. Sheep blood in Alsevier solution
3. Hank's saline solution:

NaCl	8.09 g
KCl	0.40 g
$CaCl_2$	0.20 g
$MgSO_4$	0.20 g
KH_2PO_4	0.10 g
$NaHCO_3$	1.27 g
Glucose	2.00 g
Dist. water	1000 ml

4. Rabbit anti-mouse serum (Haffkine Institute, Bombay).
5. Complement (guinea pig serum)
6. Bacto Agar (Difco)
7. DEAE (diethylamine ethyl) cellulose
8. Petri dishes (50 mm)
9. Pasteur pipettes

Procedure

1. On Day 1, immunize 4 mice with SRBC (2×10^8 cells/ml) injected intraperitoneally.
2. On Day 5 prepare:
   (i) *Underlay Agar*
   Hank's saline soln. 100 ml
   Agar 1.4 gm
   Warm the above to melt agar. Add in each of the several petri dishes enough of this agar to form a thin film, making sure that the dishes are on a level surface. Allow the agar to solidify.
   (ii) *Overlay Agar*
   Add to 80 ml dist. Water
   0.7 gm Agar
   50.0 mg DEAE-cellulose.
   (iii) (a) Spin SRBC in Alsevier's solution at 250 g for 10 minutes. Discard the supernatant.
   (b) Resuspend the cells in Hank's saline and spin as before. Discard the supernatant.
   (c) Repeat the last step two more times. You now have Hank's SS-washed SRBC ready for use.
3. (i) Prepare spleen cell suspension (as in Exercise 54) in 2.5 ml of Hank's SS.
   (ii) Make dilutions 10^{-1} and 10^{-2} using 1 ml of the cell suspension (in Hank's SS).
4. Arrange 2 sets of 6 petri dishes each, with underlaid agar (A and B sets) and label them as follows:
   10^0 10^{-1} 10^{-2} 10^0 10^{-1} 10^{-2}
5. With a Pasteur pipette dispense 0.25 ml of spleen cell suspensions (original, and dilutions 10^{-1} and 10^{-2}) into all the above petri dishes.
6. Using a warmed Pasteur pipette dispense 0.8 ml overlay agar (kept melted at 45°C in a water-bath) in 12 small tubes. Three pairs of duplicate tubes will be used for the direct (A) method and three pairs for the indirect (B) method.
7. To each of these tubes add 0.15 ml of SRBC suspension. Tap the bottom of the tube to mix.
8. Quickly pour the agar-SRBC mixture on to the underlaid agar plates and swirl to mix evenly with the spleen cells. Let the agar set.
9. (a) To the set A dishes add 1 ml of guinea pig serum diluted to 10^{-1} with Hank's SS.
   (b) To the set B dishes add 1 ml of rabbit anti-mouse Ig diluted to 10^{-1} with Hank's SS.

10. Keep the plates of set A in the incubator (37°C) for 50–60 minutes.
11. Set A: Allow the plates to stand at room temperature for about 30–40 minutes.
12. Set B:
    (i) Add excess Hank's SS and pipette off the solution.
    (ii) Repeat once.
    (iii) To the plates of set B, add 1 ml of guinea pig serum diluted to 10^{-1}.
13. Keep at 37°C for 40–45 minutes.
14. Leave the plates at room temperature for about 30–40 minutes.
15. Observe and count the plaques produced by lysis of SRBCs under a low power binocular microscope.
    Set A: Plaques are due to IgM, and
    Set B: Plaques are due to IgM and IgG.
    No. of IgG secreting cells = No. of plaques in B minus the No. of plaques in A.

9.4 ❑ MONOCLONAL ANTIBODIES

As early as 1958 Nossal and Lederberg reported that any one B cell secreted only one kind of antibody—that is, an immunoglobulin that matched one particular antigenic provocation. Since antigens trigger B cell precursors to multiply and differentiate into mature plasma or B cells, all cells that originate from a single precursor secrete identical antibody molecules. In other words, each clone of B cells is committed to the synthesis of the same antibody molecule. At about the same time as the above discovery, Okada observed that an infection by the Sendai virus causes animal cells to fuse their plasma membranes. In 1973, Cotton and Milstein demonstrated that a hybrid of rat and mouse B cells continued to synthesize the antibodies characteristic of the parental rodent cells.

Two years later Köhler and Milstein reported the fusion of a normal and a tumourous B cell of mouse. The hybrid cell proliferated, with each cell synthesizing the Ig molecules of the normal and tumourous plasma cells. The tumourous plasma cells are known as plasmacytomas or myelomas. The latter may be induced by a course of intraperitoneal injections of mineral oil. Several murine and rat myeloma continuous cell lines are available. In the majority of myelomas there is a loss of heavy (H) chains or/and an excess of light (L) chains. The normal plasma cells are obtained from the spleen. The spleen contains both the B and T types of lymphocytes as well as macrophages. The spleen is macerated and aggregates of cells are broken up to give suspensions of single cells. The latter are fused with the myeloma cells. The myeloma cells may fuse with all three types of cells in the spleen cell suspension but only the B cell hybrids will produce antibodies, which may be detected by the appropriate screening methods. Myeloma cell lines are available that do not synthesize any immunoglobulins. Fusion of spleen cells with such cell lines results in the production of only the desired antibody.

When a hybrid cell is injected into a mouse (or rat) the cell proliferates to form a tumour. The latter tumour has been named as a hybridoma. The hybridoma continues to secrete the monoclonal antibody with defined specificity. This antibody is the predominant immunoglobulin in the serum of the animal harbouring the hybridoma.

The first hybridomas were made by Köhler and Milstein by fusing a mouse myeloma cell line (MOPC-21)—that was $HGPRT^-$—with the cells taken from the spleen of a mouse sensitized with sheep red blood cells (SRBCs). The latter were, of course, $HGPRT^+$. When the spleen and myeloma cell mixture treated with the fusion agent, was transferred to a semi-solid selection medium (HAT), only hybrid cells continued to survive and proliferate. The original spleen cells do not live long on the culture medium, while the MOPC-21 cells fail to multiply on the HAT due to the mutant HGPRT gene. The hybrid cell clones secreted only anti-SRBC antibodies. The use of Sendai virus as a fusion agent has long been replaced by polyethylene glycol.

The clones of hybrid cells that secreted the desired antibody were selected in the manner described below. The cells were grown in micro-wells on a semi-solid agar medium, with as few (ideally, only one cell) cells per well as possible. The clones were then overlaid with SRBC and haemolysing complement. Those wells which contained hybrid cells secreting anti-SRBC antibodies, showed hemolysis of the added SRBC. The anti-SRBC Ig molecule binds to the surface of the SRBC while the complement serum binds to the Fc region of the Ig molecule and mediates the lysis of the red blood cells.

Preparation of monoclonals has become standardized enough for reproducible production of homogeneous antibodies with defined antigen binding properties. Many commercial suppliers provide tailor-made monoclonals. The rudiments of the process of production of monoclonals will be indicated in this chapter.

BALB/C strain of inbred mice are the best recipients for the formation of hybridomas. Where antibodies against a mouse antigen are required, the recipient animal is rat. There are some reports of human monoclonal production. When a hybrid of two species is to be injected for hybridoma formation, an appropriate histocompatible recipient (often a hybrid of two strains of mice) is chosen.

Exercise 62: Preparation of Monoclonal Antibodies

Have ready

1. BALB/C mouse
2. MOPC-21 myeloma cell line, or the NS-1 line which does not produce any immunoglobulins
3. Apparatus and equipment used for animal cell culture
4. Dulbecco's Modified Eagle's Medium (DME)
5. Fetal Calf Serum (FCS)
6. Antigen (Bovine Serum Albumin or BSA)
7. PBS
8. PEG 1000 — BDH
9. Hypoxanthine (6-hydroxypurine) — Sigma
10. Aminopterin (4-aminofolic acid) — Sigma
11. Thymidine — Sigma
12. HAT Medium: (i) Prepare HT (hypoxanthine and thymidine) and A (aminopterin) stock solutions (100 X)

HT Stock:

- Dissolve in 100 ml deionized dist. water
  136 mg hypoxanthine
  38 mg thymidine.
  Heat to 60–75°C to dissolve.
- Filter sterilize using 0.2 μm millipore filter.
- Store in the deep freeze (–20°C).

Outline of procedure

1. Immunize mice with antigen.
2. Prepare suspension of spleen cells in DME + HT + L-glutamine.
3. Grow an NS-1 (or MOPC-21) culture in DME + HT + L-glutamine.
4. Add PEG to cell pellet. Keep at room temperature for 6 minutes.
5. Stop the action of PEG by flooding with medium.
6. Replace medium with HAT medium.
7. Dispense cell suspension into micro-wells of a multiwell plate, with very few cells in each well.
8. Incubate in the CO_2 incubator, replacing half of the HAT medium once per day.
9. After 7–9 days screen the hybrid clones.
10. Replace HAT with HT medium.
11. Change the HT medium 2–3 times in 2–3 days.
12. Replace HT with DME + FCS medium or continue culturing in the HT medium.
13. Screen the hybrid clones for presence of the desired antibody (anti-SRBC) using SRBC and complement. If present, the anti-SRBC Ig will bind to the SRBC and complement C3b. Binding to the Ig-SRBC complex will bring about the lysis of the red blood cells. This will be indicated by a zone of hemolysis in the well.

Procedure

1. (i) Immunize BALB/C mice with SRBC (collected in Alsevier's solution) washed 3–4 times with PBS, injected intraperitoneally, using about 3×10^7 cells per mouse.
   (ii) Give two booster doses at intervals of 3 weeks.
   (iii) Wait for 2 weeks before use.
2. Prepare the HT-DME growth medium.
   (100 ml with FCS, 100 ml without FCS)

DME	100 ml + antibiotics
FCS	10 ml (inactivated at 56°C)
Glucose	0.45 g
HT	1.0 ml
L-glutamine (200 mM)	1.0 ml

3. Kill mice by cervical dislocation and dissect out the spleens. Place the latter in the HT-DME medium (~5 ml) in a petri dish. Macerate spleens and prepare a suspension of spleen cells in the medium.

4. Grow in the same medium a culture of MOPC-21 or NS-1 cells.
5. (i) Pellet the spleen cells and resuspend in 50 ml DME (without FCS).
   (ii) Cell density should be about 10^8/ml.
6. (i) Trypsinize myeloma monolayers and prepare pellet of cells in FCS-free DME. Resuspend in 50 ml DME (without FCS). Cell density should be about 10^7/ml.
7. Mix equal volumes of spleen and myeloma cell suspensions in a conical bottomed centrifuge tube.
8. Spin for 5 minutes at 400 g.
9. Remove all supernatant with a Pasteur pipette. Turn the tube upside down on a filter paper to drain away all the medium from the tube. (This is essential for ensuring that the PEG to be added is not diluted by the medium.)
10. Disintegrate the pellet by tapping the bottom of the tube.
11. Prepare the PEG solution:
    (i) Take 10 gm of PEG in each of the two tubes and autoclave them to melt the PEG.
    (ii) Allow the PEG to cool. While it is still liquid, remove the liquid from one tube with a warm graduated pipette. Note the volume of the PEG. Discard this PEG.
    (iii) Take DME (without FCS) that has been kept at 50°C (water-bath or incubator), and add it to the second vial of PEG to make a 35% (w/v) solution.
    (iv) Mix the medium and PEG quickly.
12. Add to the disaggregated pellet 5 ml of the PEG-DME solution.
13. Spin for 2 minutes at 200 g.
14. Rest at room temperature (25°C) for about 6 minutes.
15. Now stop the PEG action, by very gradual dilution with the serum-free DME in the following manner: First add 1 ml of DME spread over a period of 1 minute. Then add 20 ml of DME over a period of 5 minutes. The cells treated with PEG are very fragile and should be subjected to as little mechanical injury as possible.
16. Spin for 10 minutes at 200 g.
17. Resuspend the pellet in 10 ml DME with FCS (15%).
18. Spin for 5 minutes at 75 g.
19. (i) Dispense 50 µl of the cell suspension in each well of a 96-well culture plate.
    (ii) Mix the remaining suspension with twice its volume of DME + FCS, and dispense 50 µl per well in a second 96-well plate.
    (iii) Repeat the above, so that a third 96-well plate is filled with twice the diluted cell suspension.
20. Incubate the plates in the CO_2 incubator at 37% and 8% CO_2 for 24 hours.
21. (i) Prepare the HAT medium: Mix
    2 ml HT stock soln.
    2 ml A stock soln.
    100 ml DME + Glucose + Glutamine + FCS

(ii) Add to each well of the above three plates 50 µl of the HAT medium.
(iii) Incubate at 37°C in the CO_2 incubator.
(iv) Replace 25 µl of the medium in each well with 25 µl of fresh HAT.

22. Wells in which hybrid cells are growing show a change of colour (to yellow) in the growth medium.
    (i) Place each plate on the stage of the inverted microscope and examine each well for clones.
    (ii) Note wells with 1–3 clones of cells.
    (iii) Transfer cells from these wells into fresh 96-well plates (use HT-DME with FCS instead of HAT).
23. Transfer a few cells from each clone to a 25 cm^3 T-flask in the HT-DME + FCS medium and allow the clone to expand.
24. The wells should be tested for antibody (anti-SRBC) production within 3–4 days of seeding.
25. Positive clones should be expanded in T-flasks, and an aliquot from each clone should be preserved in ampules by freezing in liquid nitrogen.
26. The positive clones may be expanded in suspension culture and the antibody purified from the supernate.
27. The hybrid positive cells may be injected into the peritoneal cavity of histo-compatible mouse, which has already received a dose of 0.5 ml Pristane one week earlier. The cells grow into ascites tumours in 1–2 weeks.

The abdomen is swollen with ascites fluid, in which the tumours (aggregates of cells) are suspended. This fluid has a very high titre of the required antibody. Of course, it also contains antibodies native to the particular animal, but they are much less in titre than the desired monoclonal. The latter may be isolated from the ascites fluid by a variety of techniques, including gel electrophoresis.

The art of genetic engineering has flourished spectacularly in the past couple of decades. Sophistications and innovations, beyond the crude initial protocols and strategies, now allow operations that are not only extremely precise but also immensely elegant. Most laboratories have found improved ways of cloning and manipulating the DNA segments of their concern. Several commercial organizations have on the market products and processes that remove most of the tedium, uncertainty and impreciseness inherent in the basic laboratory practices.

One of the techniques that has revolutionized recombinant DNA technology and investigations in molecular biology is that based on the Polymerase Chain Reaction (PCR). This is an *in vitro* chemical technique that amplifies minute quantities of DNA several thousand folds in a very short period of time. The original PCR technique has undergone various modifications to suit the objectives of different laboratories. Other techniques, allied to the original PCR—such as Inverse PCR and Ligase Chain Reaction—have extended the usefulness of this novel strategy of DNA amplification *in vitro.* Basically, PCR and the associated techniques have not only reduced the time and the number of operations required to amplify a piece of DNA by the standard cloning techniques, but also opened up an entirely new approach to designing strategies in biological research and applications.

Greater advances are to be expected in the future in the practice of genetic engineering, molecular biotechnology and ancillary fields of cell culture and whole organism cloning. Those superstructures will but have taken off from the foundation laid by the basic principles and practice of genetic engineering that ushered in a sea-change in attitudes and methods in investigations of the living world at the molecular level.

REVIEW QUESTIONS

1. How are polyclonal antibodies raised against a single protein or cell? What are the precautions taken?
2. Enumerate and elaborate on various tests and techniques used for detection of antibodies.
3. Recapitulate the steps followed in ELISA test.
4. What is Complement Fixation Test?
5. Give a brief account of T and B cells.

PART III: *Appendices*

Appendix I: *Media, Buffers and Other Reagents*
Appendix II: *Some Useful Facts*
Appendix III: *Common Lab Procedures*
Appendix IV: *Multiple-Choice Questions*

Appendix I

Media, Buffers and Other Reagents

AI.1 ❑ MEDIA

1. Luria Bertani (LB) Broth:
   - Mix — Bactotryptone 10 g; Yeast Extract 5 g; NaCl 10 g; DW 900 ml
   - Adjust pH to 7.5, using I N NaOH
   - Make up volume with DW to 1000 ml
   - Autoclave.
2. LB-Agar:
   (i) For Plates: Hard
   LB 1000 ml
   Glucose 0.2% (w/v)
   Agar 20.0 g
   Autoclave.
   To add antibiotics: Cool LBA to 45°C.
   (ii) For Stabs: Soft
   LB 1000 ml
   Agar 10 g
   Boil to melt agar. Distribute into small vials (2 ml in each). Autoclave.
   (iii) Top Agar: Soft. [see (ii) above]
3. YPD (or YEPD) Medium:
   - Mix — Yeast Extract (DIFCO) 10 g; Peptone (DIFCO) 20 g; 2x DW to make 900 ml; Autoclave; Add filter sterilized; Dextrose (20% in DW) 100 ml
4. YPD-Agar:
   - Mix — Bacto-Yeast Extract 10 g; Bacto-Peptone 20 g; Dextrose 20 g; Bacto-Agar 20 g; DW to make 1000 ml
   - Autoclave

5. YPDG-Agar:
   - Mix

Bacto-Yeast Extract	10	g
Bacto-Peptone	20	g
Dextrose	1	g
Glycerol	30	ml
Bacto-Agar	20	g
DW to make	1000	ml

   - Autoclave
6. SM (Standard Medium):
   - Mix

Yeast Extract	1	g
Bacto-Peptone	10	g
Glucose	10	g
$MgSO_4, 7H_2O$	1	g

   - Dissolve in PBS to make 1000 ml
   - Autoclave
7. SMA (SM-Agar):

SM Broth	1000	ml
Agar	20	g

   - Autoclave
8. TB (Tryptone Broth):
   - Mix

Tryptone	1.9	g
NaCl	0.5	g
DW to make	100	ml

   - Autoclave
9. TCC (Ready-made powder to be reconstituted in 1000 ml DW): Composition

A. *Inorganic Salts*

KH_2PO_4	0.06	g
$MgSO_4, 7H_2O$	0.20	g
NaCl	8.00	g
$NaHCO_3$	0.35	g
Na_2HPO_4 (anhyd.)	0.475	g
$CaCl_2 \cdot 2H_2O$	0.186	g
$Fe(NO_3)_3 \cdot 9H_2O$	0.10	g
KCl	0.40	g

B. *Amino Acids*

L-Alanine	0.025	g
L-Arginine · HCl	0.07	g
L-Aspartic Acid	0.03	g
L-Cystine, 2Na	0.0237	g
L-Cysteine · HCl	0.0987	mg
L-Glutamic Acid	0.0668	g
L-Glutamine	0.001	g
Glycine	0.05	g
L-Histidine · HCl · H_2O	0.0219	g
L-Hydroxyproline	0.01	g
L-Isoleucine	0.02	g

L-Leucine	0.06	g
L-Lysine · HCl	0.07	g
L-Methionine	0.015	g
L-Phenylalanine	0.025	g
L-Proline	0.040	g
L-Serine	0.025	g
L-Threonine	0.030	g
L-Tryptophan	0.01	g
L-Tyrosine, 2Na	0.0497	g
L-Valine	0.025	g
C. *Vitamins*		
L-ascorbic acid	0.050	mg
Biotin	0.010	mg
D-Ca panthotenate	0.010	mg
Calciferol	0.100	mg
Choline Chloride	0.500	mg
Folic acid	0.010	mg
i-Inositol	0.050	mg
Nicotinamide	0.025	mg
Pyridoxel HCl	0.025	mg
Riboflavin	0.010	mg
Thiamin HCl	0.10	mg
Pyridoxine HCl	0.025	mg
Para-aminobenzoic acid	0.050	mg
Nicotinic acid	0.025	mg
Menaphthone sodium bisulphite $3H_2O$	0.019	mg
Dl-tocopherol	0.010	mg
Vit. A acetate	0.115	mg
D. *Others*		
D-glucose	1.00	g
Phenol Red	0.017	g
Hypoxanthine	0.30	mg
(reduced) Glutathione	0.05	mg
Tween 80	0.005	mg
Adenine SO_4	0.01	g
5′ AMP	0.20	mg
ATP · 2Na	0.01	g
Cholesterol	0.20	mg
2-Deoxyribose	0.50	mg
Guanine HCl	0.30	mg
D-Ribose	0.50	mg
Na-acetate	0.0367	g
Thymine	0.30	mg
Uracil	0.30	mg
Xanthine	0.30	mg

10. MS Medium (Murashige and Skoog):
Add the ingredients in the given order.

A. *Major Elements* — 1000 ml

Ammonium nitrate	1.65	g
Potassium nitrate	1.90	g
Pot.-dihydrogen phosphate	0.17	g
$CaCl_2$	0.44	g

B. *Mix Separately* — 1000 ml

$MgSO_4 \cdot 7H_2O$	0.37	g
$MnSO_4 \cdot 4H_2O$	0.0223	g

C. * *Stock Solution* — 1000 ml

Boric acid	6.2	mg
$ZnSO_4 \cdot 4H_2O$	8.6	mg
KI (potassium iodide)	0.83	mg
Sodium molybdate	0.25	mg

**To make this:*
Add the following amounts to 100 ml DW

Boric acid	62	mg
Zn-sulphate	86	mg
K-iodide	8.3	mg
Sodium molybdate	2.5	mg

N.B. Use 10 ml/litre.

D. ** *Stock Solution* — 1000 ml

$CuSO_4$	0.025	mg
$CoCl_2 \cdot 6H_2O$	0.025	mg

** To make this:

Dissolve $CuSO_4$	2.5	mg
$CoCl_2 \cdot 6H_2O$	2.5	mg
In DW	100.0	ml

N.B. Use 1 ml/litre.

E. *Stock Solution* (Fe Salts)

$FeSO_4 \cdot 7H_2O$	27.8	mg
Na_2EDTA	39.3	mg

- Dissolve 55.1 mg of Fe-sulphate in 100 ml of *hot* DW.
- Dissolve 74.5 mg of Na_2EDTA in another 100 ml of *hot* DW.
- Mix both to get 200 ml.

N.B. Use 5 ml/litre.

F. *Stock Solution in* 1000 ml

1. Thiamine HCl	0.1	mg
2. Nicotinic HCl	0.5	mg
3. Pyridoxine HCl	0.5	mg

- Prepare 10 ml of the Stock solution, containing 10 mg of (1)
  50 mg of (2)
  50 mg of (3)

N.B. Use 0.1 ml/litre.

G. For	1000	ml
glycine	2.0	mg
myo-inositol	100.0	mg

N.B. Do not prepare stock solution.

- Add fresh to the medium.

H. *Agar* 1.8% (18 g/1)

The solutions are combined as follows:

1. Weigh A and place in a large flask with some measured volume (< 1000 ml) of DW.
2. Mix the items in B, and add to A.
3. Take 10 ml of stock C, add to the above.
4. Take 1 ml of stock D, add to the above.
5. Take 5 ml of stock E, add to the above.
6. Take 0.1 ml of stock F and add to the above.
7. Add G (freshly weighed).

To Prepare the Solid Medium:

You may prepare the above stock solutions and keep them in plastic bottles in the deep freeze for use and re freeze them for further use.

To prepare 1 litre of the medium (enough for 40 tubes):

1. Add agar to the flask and heat to suspend the agar uniformly.
2. Pour medium into tissue culture glass tubes (held in test-tube stands), using a small beaker for pouring.
3. Plug each tube with cottonwool covered with cheesecloth.
4. Autoclave, after covering tops of tubes with Al foil, for 20 minutes.
5. Remove from autoclave.
6. When temperature comes down sufficiently to be able to handle with a towel, place each tube carefully in a slanting position. This is done in a ready make stant maker or by placing the tubes with the 'head' resting on an empty tube, at right angles to the media-filled tubes.
7. When the medium has solidified, store the tubes upright in the cold room or Plant Tissue culture room as required.

N.B. See 10a and 10b below.

10a. Prepared MS Medium: The MS medium is sold in packs for each 1 litre reconstituted medium. If agar is not included, add agar and calcium chloride 400 mg/litre. Autoclave and make slants.

10b. To add auxin/kinetin to MS:

After the agar is suspended well in 1 litre of the medium, add the powders of auxin or kinetin directly (separately or together in the same medium) as required.

11. Medium for *Agrobacterium:*

Nutrient Broth	8	g
Yeast Extract	5	g
NaCl	5	g
DW	1000	ml
*Agar	1.5–1.8%	

*Agar is added after adjusting the pH.

12. Drosophila Food:
    1.5 litre medium fills about 30 bottles
    - You require:

Agar	20	g
Yeast Powder	22	g
Maize Flour or *Suji	150	g
*Cream of wheat		
Unbleachdd sugar	130	g
DW	1500	ml
Propionic Acid	5	ml

    1. Prepare (i) Agar, (ii) sugar and suji, and (iii) yeast separately as follows.
       (i) Add agar to warmed dist. water (200 ml) in a 1000 ml beaker and boil while stirring — (A).
       (ii) Add sugar and suji to 250 ml DW and boil the mixture, while stirring — (B).
       (iii) Mix yeast in 100 ml DW and boil — (C).
    2. Boil in a 2 litre beaker for 10–15 minutes.

DW	400	ml
(A) (B) (C)	550	ml

       Keep stirring constantly.
    3. Add propionic acid; boil for 4–5 minutes.
    4. Dispense into sterile bottles (to a depth of about 5 mm).
    5. Keep bottles on a tray, covered with a clean towel, in the Laminar Flow Cabinet (overnight), to allow the steam to condense.
    6. Close mouths of bottles with sterile plugs (cotton wool or synthetic foam).
    7. Keep in the cold room till use.

    N.B. This food should be used within 4–6 days.

13. Drosophila Culture Medium (Schneider's):
    A. *Inorganic Salts*

	mg/litre
$CaCl_2$, anhydrous	600
KCl	1600
KH_2PO_4	450
$MgSO_4 \cdot 7H_2O$	3700
NaCl	2100
$NaHCO_3$	400
$Na_2HPO_4 \cdot 7H_2O$	1321

    B. *Amino Acids*

Alanine	500.0
L – R*	400.0
L – D	400.0
L – C	60.0
L – Cystine	100.0

L – E	800.0
L – Q	1800.00
L – G	250.0
L – H	400.0

*See symbols for Amino acids in Appendix 11.10

L – I	150.0
L – L	150.0
L – K-HC1	1650.0
L – M	800.0
L – F	150.0

C. *Others*

L-Ketoglutaric acid	200.0
Fumaric acid	100.0
D-Glucose	2000.0
Malic acid	100.0
Succinic acid	100.0
Trehalose	2000.0
Yeastolate	2000.0

14. HAT Medium:
    - Dissolve (i) Hypoxanthine 136 mg
      In 0.05 N HCl 100 mg
      (ii) Aminopterin 1.76 mg
      In 0.1 N NaOH 100 ml
      (iii) Thymidine 38.7 mg
      In BSS 100 ml
    - Filter sterilize solutions (i), (ii) and (iii).
    - Mix equal volumes of (i), (ii) and (iii).
    - Store aminopterin soln. at –20°C, hypoxanthine soln. and thymidine soln. at 4°C.
    - Concentration in medium 3%.
15. Maintenance Media:
    - Maintain stabs of *E. coli* strains, individually, by yearly transfers into stabs of the following composition.

Bacto Tryptone	5 g
Bacto peptone	8 g
NaCl	5 g
Agar	20 g
DW	1000 ml
pH	7.2 – 7.4
Add thiamine	50 µg/ml

    - Dissolve the above and dispense appropriate volumes in small screw-capped vials.
    - Autoclave 15 psi, 20 minutes.
    - Store the stabs, with or without cultures at ROOM TEMPERATURE, (*not* at 4°C or deep freeze).

16a. Maintenance of Liquid Stock Cultures:
- Since the frequent opening of vials leads to contamination, it is preferable to store culture stocks for a month or two at 4°C. While remaking the stock, it is essential to check the markers of each strain and prepare fresh stocks from the checked clones.
- While performing experiments, only fresh overnight cultures must be used. (1 loopful from stock culture in 10 ml liquid broth, incubated at appropriate temperature overnight).

AI.2 ❑ BUFFERS

1. PBS (Phosphate Buffered Saline):
   - Dissolve NaCl 10.0 g
     KCl 0.25 g
     Na_2HPO_4 1.43 g
     KH_2PO_4 0.25 g
   - In DW 100 ml
   - Add DW to make up volume to 1000 ml.
   - Adjust pH to 7.2.
   - Autoclave.
2. PBS (for RNA Isolation):
   - Dissolve 5 M NaCl 20 ml
     1 M Tris.Cl 20 ml pH 7.4
     0.5 M EDTA 2 ml pH 7.5
   - In DW 958 ml.
3. TE*:
   Tris base is Tris (hydroxymethyl) Aminomethane, M.W. 121.14
   - Dissolve Tris base 0.12 g 10 mM
     Na_2EDTA 20.3 mg 1 mM
   - In DW 80.0 ml
   - Adjust pH to 8.0 using conc. HCl.
   - Add DW to make up volume to 100 ml.
4. Tris.Cl (or Tris-HCl): 1 M
   Tris hydrochloride, M.W. 157.64
   - Dissolve Tris base 12.1 g
   - In DW 80.0 ml
   - Let the solution cool to room temperature.
   - Adjust pH as required (7.4, 8.0 or 9.5), using conc. HCl.
   - Add DW to make up volume to 100 ml.
   - Autoclave.
5. Tris.Cl: 10 mM
   - Dissolve Tris Base 0.12 g
   - In DW 80.0 ml
   - Adjust pH as required (7.5 or 8.0) using conc. HCl.

6. SET:

Sucrose	20	%	
EDTA	50	mM	
Tris.Cl	50	mM	pH 7.6

7. STET:

Sucrose	8	%	
Triton X-100	5	%	
EDTA	50	mM	
Tris.Cl	50	mM	pH 8.0

8. TEN (for DNA):
   - Mix

1 M Tris.Cl	10	ml	(10 mM)
0.1 M EDTA	10	ml	(1 mM)
5 M NaCl	2	ml	(10 mM)

   - Add DW to make up volume to 1000 ml.

9. RNase Buffer:
   - Mix

Na-acetate	0.1	M	
EDTA	0.3	M	pH 4.8

   - Heat the solution to 80°C for 10 minutes.
   - Add to the above.

RNase A	10	mg/ml of solution.

10. TM:

Tris.Cl	50	mM	pH 7.5
$MgSO_4$	10	mM	

11. Lambda DNA Buffer:

Tris.Cl	0.01	M	pH 8.0
KCl	0.1	M	
EDTA	10^{-4}	M	

12. NPEB (Nucleoprotein Extraction Buffer):

Tris.Cl	50	mM	pH 7.0
EDTA	5	mM	
β-mercaptoethanol	10	mM	

13. High Salt NPEB:

NPEB + NaCl	1	M

14. Lysis Buffer (for slime mold cells):

Hepes	50	mM	pH 7.5
$MgAc_2$	5	mM	
Sucrose	10	%	
KCl	40	mM	
PMSF	1	mM	

15. Storage Buffer (For slime mold nuclear pellet):

Tris.Cl	40	mM	pH 7.9
$MgCl_2$	10	mM	
Na_2EDTA	0.1	mM	
DTT	1.0	mM	
Glycerol	50	%	

16. Buffers for Reactions of Some Restriction Enzymes:

Eco RI	Tris.Cl	100	mM	pH 7.5
	$NaCl_2$	50	mM	
	$MgCl_2$	10	mM	
5′G/AATTC3′			37°C	
Bam HI	Tris.Cl	20	mM	pH 8.00
	NaCl	100	mM	
	$MgCl_2$	7	mM	
5′G/GATC3′	β-mercaptoethanol	2	mM	
Alu I	Tris.Cl	50	mM	pH 8.00
	NaCl	50	mM	
	$MgCl_2$	5	mM	
	DTT	1	mM	
5′AG/CT3′				37°C
Hin dIII[a], *Mbo* I[b]				
	Tris.Cl	50	mM	pH 8.0
	NaCl	50	mM	
	$MgCl_2$	10	mM	
a 5′A/AGCT3′				37°C
b 5′/GATC3′				
Sal I	Tris.Cl	8	mM	pH 7.6
	NaCl	150	mM	
	$MgCl_2$	6	mM	
	Na_2EDTA	0.2	mM	
5′G/TCGAC3′			38°C	
Sau 3AI	Tris.Cl	6	mM	pH 7.5
	NaCl	50	mM	
	$MgCl_2$	6	mM	
	BSA	100	µg/ml	
5′GATC3′				37°C

AI.3 ❑ OTHER REAGENTS

1. Ammonium Acetate 5 M STOCK:
   - Dissolve Ammonium acetate 385 g
   - In DW 800 ml
   - Add DW to make volume up to 1000 ml
   - Filter sterilize.
2. β-mercaptoethanol (BME):
   - It comes in a 14.4 M solution.
   - To be stored at 4°C in the dark.
   - Not to be heat-sterilized.
3. Dithiothreitol (DTT):
   - Add to DW 20 ml
     3 M Na-acetate 67 µl

- Dissolve in the above solution

DTT	1.55	g

- Filter sterilize (0.2 µm membrane).
- Dispense into 0.5 ml aliquots in small vials.
- Store at –20°C.

4. EDTA (Ethylene-diamine-tetra-acetate, disodium salt) 0.5 M, pH 7.0, 7.5 or 8.0
   - Dissolve *EDTA 18.6 g
   - In DW 80 ml
   - Adjust pH as desired, using NaOH pellets.
   - Add DW to make up volume to 100 ml.
   - Autoclave.

   *Dissolves completely only after pH is around 7.0.
5. Ethidium Bromide *Stock 10 mg/ml
   - Add Ethidium bromide 0.5 g
   - To DW 50 ml
   - Dissolve by stirring on a magnetic stirrer.
   - Cover bottle with aluminium foil to protect the EtBr from light.

   *This is a strong mutagen/carcinogen and should therefore be handled with care using gloved hands.
6. Lysing solution (for slime mold cells):

TE	42.5	µl
VRC*	5.0	µl
5% NP40**	5.0	µl

   *VRC, Vanadyl Ribonucleoside Complex (available commercially, or may be prepared in the laboratory as given in Maniatis, T. et al., *Molecular Cloning: A Laboratory Manual*).

   **This is a detergent Nonadet P40. (nonionic)
7. Sodium Chloride (NaCl) 5 M
   - Dissolve NaCl 29.0 g
   - In DW 100 ml
   - Autoclave.
8. Sodium acetate 3 M pH 5.2
   - Dissolve anhydrous Na-acetate 24.6 g
   - In DW 80 ml
   - Adjust pH to 5.2 using glacial acetic acid.
   - Add DW to make up volume to 100 ml.
9. Sodium perchlorate 5 M
   - Dissolve $NaClO_4$ 612.5 g
   - In DW 800 ml
   - Add dDW to make up to 1000 ml.
10. Spermidine 100 mM
    - Dissolve Spermidine trihydrochloride 255 mg
    - In DW to a volume of 10 ml
    - Filter sterilize (0.2 µm membrane).
    - Store at –20°C.

11. SSC (Standard Sodium Citrate):

	NaCl	0.15	M
	Na_3 citrate	0.01	M

- Adjust pH to 7 ± 0.2

12. Tracking dye I:

• Dissolve	Bromophenol blue	50.0	g
	Sucrose	5.0	g
• In	Glass DW	50.0	ml

13. Tracking Dye II:

• Mix	Bromophenol blue	0.25	g
	Xylene-Cynol	0.25	g
	SDS	0.005	g
	Glass DW	50.0	ml

- Just before use:
- Add to the above solution

	Glycerol	50.0	ml

14. Triton X-100 5%

• Dissolve	Triton-X 100	5	ml
• In	DW	95	ml

15. Earles' Balanced Salt Solution (Earle's BSS):

• Dissolve	$CaCl_2$ (anhydrous).	0.02	g
	KCl	0.04	g
	$MgSO_4 \cdot 7H_2O$	0.20	g
	NaCl	6.68	g
	$NaHCO_3$	2.20	g
	$NaH_2PO_4 \cdot H_2O$	0.14	g
	D-Glucose	1.00	g
	Phenol Red	0.01	g
• In deionized 3x DW		1000	ml

- Filter sterilize.

16. HBSS (Hank's Balanced Salt Solution)

	KCl	0.40	g
	KH_2PO_4	0.06	g
	$MgCl_2 \cdot 6H_2O$	0.10	g (optional)
	$MgSO_4 \cdot 7H_2O$	0.10	g
	NaCl	8.00	g
	$NaHCO_3$	0.35	g
	Na_2HPO_4	0.06	g
	D-Glucose	1.00	g
	Phenol Red	0.02	g
• In deionized 3x DW		1000	ml

- Adjust pH to 6.5 with sterile 0.1 N NaOH.
- Autoclave.

17. Hoechst 33258*

• Mix	BSS	10	ml (without phenol red)
	Hoechst 33258	10	mg

- Store at –20°C.

For Use

Dilute stock : BSS : : 1.0 µl/20 ml

*Handle carefully. The reagent is carcinogenic.

18. Trypsin* Stock:
    - Dissolve Trypsin 2.5 gm
    - In 0.85% NaCl 100 ml
    - Stir for 1 hour at room temperature to dissolve completely.
    - Filter sterilize and store in aliquots at –20°C.

    *Crude Trypsin (Difco) may be used at the above 2.5% concentration. If cells are damaged by above, use purified trypsin (Sigma) at a concentration of 0.05%.

AI.4 ❑ ANTIBIOTICS

1. Ampicillin (Sodium Salt):
    - Dissolve Ampicillin 25 mg
    - In DW 10 ml
    - Filter sterilize and dispense into Eppendorf tubes.
    - Store at –20°C.

    To Add to Media:
    - Autoclave media; allow to cool to 55°C.
    - Add ampicillin solution about 40 µg/ml (which is about 0.8 mg/20 ml media in a petri dish)
2. Streptomycin:
    - Dissolve Streptomycin sulphate 200 mg
    - In DW 10 ml
    - Filter sterilize and store in aliquots at –20°C.
3. Tetracycline:
    - Dissolve Tetracycline hydrochloride 125 mg
    - In Ethyl alcohol: water (50% v/v) 10 ml
    - Filter sterilize. Store in small aliquots in Eppendorf tubes, covered with aluminium foil at –20°C.
4. Chloramphenicol:
    - Dissolve Chloramphenicol 340 mg
    - In Absolute ethyl alcohol 10 mg
    - Store at –20°C.

AI.5 ❑ ENZYMES

1. RNase A (pancreatic):
    - Prepare a solution of
      Tris.Cl 10 mM pH 7.5
      NaCl 15 mM
    - Add RNase 10 mg/ml of above solution

N.B. Heat to 80–100°C for 10–15 minutes. Allow to cool gradually to room temperature. Distribute into Eppendorf tubes and store at –20°C.

- This destroys the usual DNase contaminant.

2. Proteinase K (make fresh):
   - Dissolve Proteinase K (Boehringer) 1 mg/ml in DW.

   N.B. This is usually used in a buffer. (See buffers).
3. Lysozyme I (made fresh for use):
   - Dissolve Lysozyme 2 mg/ml
   - In a solution of 0.15M NaCl
     0.1 M EDTA pH 8.0
4. Lysozyme II:
   - Add solid lysozyme to a final concn. of 5 mg/ml.
   - Mix to dissolve.

AI.6 ❑ HORMONES, MITOGENS AND MITOTIC POISONS

- *Cytokinins*

Adenine sulphate	
Kinetin	6-Furfurylamino purine
Auxins	
BAP or BA	6-Benzylamino purine or N^6 Benzyl Adenine
CPA	4-chlorophenoxyacetic Acid
2,4-D.	2,4-Dichlorophenoxyacetic Acid
IAA	Indole-3-Acetic Acid
IBA	Indole-3-Butyric Acid
NAN	α-Naphthaleneacetic Acid
Mitogens	
CON. A	Concanavaline A (Jack Beans)
PWM	Pokeweed Mitogen (*Phytolacca americana*) leaves and roots; mitogen for both B and T Cells
PHA	Phytohemagglutinun, for T cells
WGA	Wheat Germ Agglutinin (*Triticum vulgaris*)
Colchicine	
Colecemid/ demecolchicine	N-diacetyl-N-methyl colchicine

AI.7 ❑ STRENGTHS OF SOME REAGENTS AS SUPPLIED

HCl	11.6 M (36% HCl)
H_3PO_4	14.7 M (ortho; 85%)
H_2SO_4	17.8 M
HNO_3	16.4 M
Glacial Acetic Acid	17.4 M
β-mercaptoethanol	15.6 M

Appendix II

Some Useful Facts

AII.1 ❑ SOME UNITS AND B DNA STATISTICS

- To Recapitulate:
  1 mg = 10^{-3} g
  1 µg = 10^{-3} mg = 10^{-6} g
  1 ng = 10^{-9} g
  1 pg = 10^{-12} g
- Some B DNA statistics:

Length of DNA/nucleotide	3.4 Å
M.W. of 1 bp (average)	650 daltons
270 bp DNA	10 kdal
810 bp DNA	30 kdal
1035 bp DNA	50 kdal
2700 bp DNA	100 kdal
1 bp (dNMP pair)	600 g/mole (average)
1 base (dNMP)	300 g/mole (average)

M.W. of a DNA fragment

$$= \text{no. of bp} \times \frac{6000 \text{ g/mole}}{\text{bp}}$$

= g/mole DNA.

Moles of ends (2/fragment)* of a DNA fragment (5′ or 3′ ends)

$$= \frac{\text{gDNA}}{\text{M.W.DNA}} = \text{moles DNA} \times 2^*$$

1 p mole of 1 kb DNA	= 0.66 µg
1 µg of DNA	= 1.52 p mole
1 µg pBR322	= 0.36 p mole DNA
1 p mole pBR322	= 2.8 µg
1 kb ds DNA (Na salt)	= 6.6×10^5 dal
1 kb ss DNA (Na salt)	= 3.3×10^5 dal
1 kb ss RNA (Na salt)	= 3.4×10^5 dal
ds DNA A_{260nm} OD = 1	= 50 µg/ml
ss DNA A_{260nm} OD = 1	= 33 µg/ml
ss RNA A_{260nm} OD = 1	= 40 µg/ml

AII.2 ❑ SIZES OF SOME VECTORS AND CHROMOSOMAL DNA

Chromosomal DNA		
SV40		5 kb
Adenovirus		21 kb
Bacteriophage Lambda		50 kb
Bacteriophage T2		200 kb
Escherichia coli		4600 kb
Drosophila melanogaster		$1–6 \times 10^5$ kb
Mouse		2.3×10^5 kb
Man		2.8×10^6 kb
Plasmids		
pACYC177	2.5 Mdal	
pBR322	2.5 Mdal	4.363 kb
pBR329		4.150 kb
pMB9	3.6 Mdal	
Col E1	4.2 Mdal	
pSC101	5.8 Mdal	
RSF2124	7.4 Mdal	
RP4	30.0 Mdal	
Phage Vectors		
M13 mp7		6.4 kb
Charon4		45.0 kb
λ gt WES λ B		49.0 kb

AII.3 ❑ SIZE MARKERS FOR DNA, RNA, PROTEINS

A. DNA Size Markers*

1. λ DNA. Strain CIts857, 48,540 bp

   With *Eco* RI:
   21.240, 7.420, 5.810, 5.650, 4.880, 3.540

   With *Hin* dIII:
   23.150, 9.420, 6.560, 4.380, 2.320, 2.020, 0.560, 0.125

   With *Bam* HI:
   16.840, 7.230, 6.785, 6.530, 5.620, 5.530

2. pBR322, 4362 bp

   With *Acc* I
   2.767, 1.595

   With *Taq* I
   1.444, 1.307, 0.475, 0.368, 0.315, 0.312, 0.141

   With *Hpa* II
   0.622, 0.527, 0.404, 0.309, 0.242, 0.238
   0.217, 0.201, 0.190, 0.180, 0.160, 0.160
   0.147, 0.147, 0.122, 0.110, 0.090, 0.076

0.067, 0.034, 0.034, 0.026, 0.026, 0.015
0.009, 0.009.

Bromophenol blue, the tracking dye co-migrates with 300 bp ds DNA, in a 1% agarose gel.

B. RNA Size Markers*	kb	Daltons
E. coli 5S rRNA	0.120	3.72×10^4
Mouse α-globin mRNA	0.696	0.22×10^6
Mouse β-globin mRNA	0.783	0.24×10^6
E. coli 16S rRNA	1.776	0.53×10^6
Hela 18S rRNA	2.366	0.71×10^6
E. coli 23S rRNA	3.566	1.07×10^6
Chicken Myosin H chain mRNA	6.500	2.02×10^6

(Xylene cyanol runs just behind 18S rRNA, and can be used as a marker)

C. Protein Size Markers**	kdal
Cytochrome C	12.4
Lysozyme (egg white)	14.3
Trypsin inhibitor (soybean)	20.1
Lactate dehydrogenase	36.0
RNA Pol, α-subunit (*E. coli*)	39.0
Ovalbumin	43.0
Gamma Globulin	53.0
Bovine serum albumin	68.0
β-galactosidase (*E. coli*)	130.0
Myosin H chain (rabbit)	212.0

*From: Minter, S. and Sealey, P. (1987)
**From: Fasman, G.D. (1976)

	Restriction Enzymes	R.E. Motifs	No. of Sites of Cleavage in pBR322	Source
1.	*Alu* I	AG/CT	16	Arthrobacter luteus
2.	*Ava* I	C/PyCGPuG	1	Anabaena variabilis
3.	*Bal* I	TGG/CCA	1	Brevibacterium albidum
4.	*Bam* HI	G/GATCC	1	Bacillus amyloliquifaciens H
5.	*Bgl* I	$GCC(N)_4/NGGC$	3	Bacillus globigii
6.	*Eco* RI	G/AATTC	1	Escherichia coli RY13
7.	*Hae* I	(A) GG/CC (A) T T	7	Haemophilus aegypticus
8.	*Hin* cII	GTPy/PuAC	2	Haemophilus influenzae Rc
9.	*Hin* dIII	A/AGCTT	1	Haemophilus influenzae Rd
10.	*Hpa* I	GTT/AAC	0	Haemophilus parainfluenzae

(*Contd.*)

(*Contd.*)

11.	*Mbo* I	/GATC	22	Moraxella bovis
12.	*Pst* I	CTGCA/G	1	Providencia stuartii 164
13.	*Sau* 3A	/GATC	22	Staphylococcus aureus 3A
14.	*Sal* I	G/TCGAC	1	Streptomyces albus G
15.	*Sma* I	CCC/GGG	0	Serratia marcescens Sb
16.	*Taq* I	T/CGA	7	Thermus aquaticus YTI
17.	*Xba* I	T/CTAGA	0	Xanthomonas badrii
18.	*Xho* I	C/TCGAC	0	Xanthomonas holcicola
19.	*Xma* I	C/CCGGG	0	Xanthomonas malvacearum
20.	*Xpa* I	C/TCGAG	0	Xanthomonas papavericola

AII.4 ❑ RANGE OF SIZE RESOLUTION OF DNA IN GELS

Size of DNA (kb)	% Agarose gel	%Polyacrylamide gel
5–60	0.3	
0.8–10	0.7	
0.4–6	1.2	
0.2–4	1.5	
0.1–3	2.0	
0.1–1.0		3.5
0.08–0.5		5.0
0.06–0.4		8.0
0.04–0.2		12.0
0.01–0.1		20.0

AII.5 ❑ COMPOSITION OF POLYACRYLAMIDE GELS

For 9″ length of gel slab (*1.5 mm thick*)

	Reagents	3% gel	4.5% gel	7.5% gel
1.	Tris-borate 10X (ml)	5.0	5.0	5.0
2.	D. Water (ml)	40.3	37.8	33.0
3.	Acrylamide 29.2% bis-acrylamide 0.8%	5.00	7.6	12.6
4.	Ammonium persulphate (ml)	0.25	0.25	0.25
5.	TEMED (µl)	25	25	25
TOTAL VOLUME (ml)		50	50	50

AII.6 ❑ TEMPERATURES MAINTAINED IN DIFFERENT CULTURE AND WORK ROOMS

- General Laboratory 25°C
- Genetic Engineering Lab 20–25°C.
- Cold Room 2–4°C
- *E. coli* Culture Room 37°C (with shakers)
- Animal Cell Culture Room 25°C–30°C.
  The CO_2 Incubator is maintained at 37°C.
  This means that the ambient temperature must be 7–10°C lower.
- Plant Tissue Culture Room 25 ± 2°C
- Drosophila Room 22 ± 2°C.

AII.7 ❑ KEY TO SOME MARKER SYMBOLS IN *E. coli*

- Wild-type *E. coli* represents strains with no mutations.
- Strain C600, which is used often as the wild-type strain, however, possesses a few mutant loci (see next section).
- Mutants are denoted by the mutant loci they carry. Some of the more common markers for these loci are given below:
  end, endonuclease I
  hsdM, host-specific modification
  hsdR, host-specific restriction
  rpsL, ribosomal protein L12
  StrA, Streptomycin resistance
  RecA, general genetic recombination
  lac utilization of lactose
  λ^- sensitive to phage λ infection

  SupE or su II, SupF or su III — mutations in tRNAs that suppress amber mutations

  $r_k^- m_k^-$ are K12 strains deficient for restriction and methylating (modifying) activities. Restriction is encoded by hsdR, methylation by hsdM and recognition of the sites by the enzymes is due to the activity of hsdS.

AII.8 ❑ STRAINS GENERALLY USED

1. *E. coli* K12 strains

• HB101	F^-, leu, thi, thr, lacY, tonA,
• C600	supE 44.
• BHB2688	recA, (λ imm 434CIts b2 red3, Eam, Sam7) λ^-: r_k^+. m_k^+.
• BNB2690	Same as BHB2688, but with Dam instead of Eam.
• BHB2600	$SuII^+$ $SuIII^+$, $r_k^- m_k^-$ (SupE) (SupF)
• JM101	(F′ traD36, $proA^+B^+$, lac, IqZAM13)/thi, Δ (lac, $proA^+B^+$), SupE. $r_k^+ m_k^+$; rec a^+ host, also for M13mp vectors.

- JM103 — This is an F^+ strain with pro A^+B^+ on the F compensating for pro, lac deletion (Δ) in the *E. coli* DNA. end A, thi, Δ (lac, pro) hsd R4, SupE, λ^-, str A (FtraD, proA^+B^+ lacI Z Δ M15)
- DH1 — F^-, end A, hsd R17, $r_k^- m_k^-$, recA1, supE44, Good for plasmid and cosmid cloning.
- JM83 — ara, Δ (lac, pro A^+B^+), ϕ80, lac2Δ M15, r_k^+ m_k^+ may be used as host for pUC.
- Q359 — hsdR (r_k^-), hsdM (m_k^-), supF, ϕ80, P2 Suppressor host used for screening Spi^- recombinants of Charon28.
- HB101 (pJC74) — This is HB101 carrying the cosmid pJC74 (Amp^R).
- x1776 — A much crippled 'safe' strain that transforms readily, Cm^s Kan^s and Tet^s.

2. *Drosophila melanogaster*
   Strain Oregon K.
   The wild-type is designated as OR – K.
   The mutants used are
   w (white eyes)
   y (yellow body)
   e (ebony body)
   vg (vestigeal wings)
   The gene loci w and y are in chromosome 1, and for vg on chromosome 2.

3. *Dictyostelium discoideum*
   The wild-type strain NC4.

4. *Animal Cell-Line Strains*

	Source	*Looks like*	*Special feature*
HeLa	Human	Epithelial	GGPD
3T3-L1	Swiss mouse	Fibroblast	
3T3-A31	Mouse BALB/C	Fibroblast	
L	Mouse	Fibroblast	TK^- Grows well in suspension cultures.
BHK21	Syrian Hamster kidney	Fibroblast	polyoma virus infection transformation

   You may prepare primary cultures from any tissue/animal.

5. *Datura inoxii* for haploid culture
   - *Nicotiana tabacum* (tobacco plant)
   - *Daucus carota* (carrot) for single cell and explant cultures.

N.B. You may also try tomato, petunia, potato and mustard plants for tissue culture and somatic cell hybridization.

AII.9 ❑ GENETIC CODE AND LETTER SYMBOLS FOR AMINO ACIDS

20 amino acids are represented by 61 of the triplet codons. The remaining 3 of the total 64 codons are stop codons. The codons are in the RNA. They are matched by amino acid-bearing tRNAs during protein synthesis.

Ist	Middle Base								3rd Base
5′	U		C		A		G		3′
U	UUU	F	UCU	S	UAU	V	UG	C	U
	UUC		UCC		UAC		UG		C
	UUA	L	UCA		UAA	STOP	UGA	STOP	A
	UUC		UCG		UAG	STOP	UGG	W	G
C	CUU	L	UCU	P	CAU	H Q	CG	R	U
	CUC		CCC		CAC		CG		C
	CUA		CCA		CAA		CG		A
	CUC		CCG		CAC		CG		G
T	AUU	I M	ACU	1st Base	AAU	D E	AG	S R	U
	AUC		ACC		AAC		AG		C
	AUA		ACA		AAA		AG		A
	AUC		ACG		AAC		AG		G
G	GUU	V	GCU	A	GAU		GG	G	U
	GUC		CCC		GAC		GG		C
	GUA		GCA		GAA		GG		A
	GUC		GCG		GAG		GG		G

STOP CODONS: UAA ochre
UAG amber
UGA opal

Key to Letter Symbols for Amino Acids

A	alanine	L	leucine
R	arginine	K	lysine
N	asparagine	M	methionine
D	aspartic acid	F	phenylalanine
C	cysteine	P	proline
G	glycine	S	serine
E	glutamic acid	T	theonine
Q	glutamine	W	tryptophan
H	histidine	Y	tyrosine
I	isoleucine	V	valine

Appendix III

Common Lab Procedures

AIII.1 ❑ TO CHECK THE PURITY OF DNA

1. Use the Spectronic 20 to find the OD of the DNA solution:
   If A_{260} is 1, there is 1 μg/ml
   If A_{260}/A_{280} is 1.8–2.9, the DNA is quite pure.
   If A_{260}/A_{280} is < 1.8, the DNA is impure*.
   * It may contain traces of proteins and/or phenol.
2. To check the physical state of DNA (amount of shear or fragmentation), fractionate a sample on a mini-agarose gel (0.8%) with 0.5 μg/ml ethidium bromide.
   If the DNA is fairly intact there will be one major band. Shearing is indicated by one or more bands further away from the main band.

AIII.2 ❑ TO PREPARE CLEAN DNA FOR RESTRICTION ENZYME OR LIGASE ACTION

Conventional isolation of DNA using phenol, CsCl density gradient centrifugation, agarose gel column chromatography and even ethanol precipitation leave heavy metal and other matter that inhibit the action of restriction enzymes and ligase. To avoid this, the DNA may be precipitated with ethanol and *ammonium acetate* in the following manner:

- To the DNA solution add 7.5 M NH_4 – acetate (pH 7.5) and 95% ethyl alcohol.
- To the DNA solution add $\frac{1}{2} \times 7.5$ M NH^{-}_{4} acetate, pH 7.5.
- Add twice the above volume of 95% ethanol.
- Keep the mixture overnight at –20°C.
  The DNA precipitates.

AIII.3 ❑ TO REDISTIL COMMERCIAL PHENOL

Commercial phenol is generally yellowish in colour. It is redistilled at 160°C through a distillation column and the phenol collected under chilled water in warm weather. In winter a blower-heater is focused on the column to prevent solidification of the phenol in the column.

The distillation is carried out in a chemical hood with the phenol in a round-bottomed flask heated in a mantle-heater.

The redistilled phenol is stored in aliquots of 250 ml at –20°C.

AIII.4 ❑ TO EQUILIBRIATE DISTILLED PHENOL

1. Take out distilled phenol from storage (–20°C) and allow it to warm to room temperature.
2. Melt the phenol at 68°C.
3. The melted phenol is now neutralized and water saturated by shaking with 2–3 changes of a Tris.Cl buffer in the following manner:
   (i) In a large beaker mix the phenol and an equal volume of 1 M Tris.Cl buffer (pH 8.0).
   (ii) Fill half the bulb (chamber) in a 250 ml separation funnel; holding the bulb with two hands shake the mixture (BE VERY CAREFUL*).
   (iii) Fix the funnel on a ring stand. Open the lower stop-cock and let out the bottom phenol phase. Discard the upper aqueous phase.
   (iv) Repeat (i) – (iii) once more.
   (v) Repeat (i) – (iv) using Tris.Cl (pH 8.00) with EDTA (1 mM).
   (vi) Finally add to the organic/phenol phase an equal volume of Tris.Cl (pH 8.0).
   (vii) Store in aliquots at 4°C in a dark bottle.
   * Phenol causes severe burns, and nausea if its fumes are inhaled. Distillation should be performed in a fumehood. Phenol should be washed off using plenty of water (*not ethanol*)

AIII.5 ❑ TO PURIFY IgG BY AFFINITY CHROMATOGRAPHY

You may use Protein A Agarose (BRL Cat. No. 5918SA), which binds to the Fc region of IgG. Protein A is bound to the surface of the 4% agarose beads (40–150 μm size) by cyanogen bromide.

18 mg of IgG binds to 1 ml of the agarose gel.

Appendix IV

Multiple-Choice Questions

1. Mutation theory was proposed by
   (a) Alfred H Wallace (b) H J Muller
   (c) Gregor Mendel (d) Charles Darwin
2. The genetic code was cracked by
   (a) Watson and Crick (b) Jacob and Monod
   (c) Nirenberg, Mathaei and Khorana (d) Benzer
3. The first codon in the coding sequence is AUG which codes for
   (a) arginine (b) lysine (c) methionine (d) threonine
4. In 1973, the first gene was cloned by
   (a) Merty and David (b) Boyer and Cohen
   (c) Bertain and Weigel (d) Hamilton O Smith
5. Synthesis of DNA strand along an RNA template is mediated by
   (a) restriction endonuclease (b) polynucleotide ligase
   (c) DNA polymerase (d) reverse transcriptase
6. Plasmid-carrying nif gene is encountered in
   (a) *Agrobacterium* (b) *Saccharomyces*
   (c) *Escherichia coli* (d) Rhizobium
7. SV40 is a vector from
   (a) animal virus (b) bacteriophage (c) *Klebsiella* (d) *E. coli*
8. Cloning vectors from *Drosophila* are
   (a) autonomous replacing sequences (b) transposable elements
   (c) putative gene sequences (d) transposable sequences
9. The process of introduction of random fragments of the *E. coli* DNA into fresh cells by infection with phages carrying the DNA fragment, is called
   (a) transfection (b) transformation
   (c) transduction (d) transposition
10. Transferring RNA bands from electrophoretic gels to solid substrate, is called
    (a) Southern blotting (b) Northern blotting
    (c) Western blotting (d) dot-blot hybridization
11. Short lengths of DNA which serve to join DNA strands with incompatible ends are called
    (a) ligators (b) connectors (c) linkers (d) adaptors
12. Use of chemicals as mutagens was pioneered by
    (a) H G Muller (b) Karl Sax
    (c) Auerbach (d) Alexander Rich

13. A clone bank contains
    (a) genomic DNA fragments (b) cDNA
    (c) transposons (d) connector sequences
14. Identifying sequences that are bound to specific proteins is called
    (a) chromosome walking (b) DNA footprinting
    (c) chromosome jumping (d) restriction mapping
15. A rapid sequencing variation of enzymatic method was proposed by
    (a) Sanger (b) Maxam (c) Gilbert (d) Messing
16. In PCR, dissociation of the two strands of the sample DNA is achieved at
    (a) 55°C (b) 94°C (c) 72°C (d) 37°C
17. For study of specific types of gene expression, the most favorite cultured cell line is
    (a) mouse cell line (b) hamster cell line
    (c) monkey cell line (d) fish cell line
18. Viral oncogene lacks
    (a) exons (b) introns
    (c) both exons and introns (d) neither exons nor introns
19. An anti-leukaemic alkaloid vincristine is obtained from
    (a) *Catharanthus roseus* (b) *Cinchona legerians*
    (c) *Papaver comniferum* (d) *Digitalis lanata*
20. The first hormone obtained in the laboratory by synthesis in *E. coli* is
    (a) calcitonin (b) somatostatin
    (c) cholecystokinin (d) ACTH
21. Successful tissue transplantation often depends on administration of antibiotics, such as,
    (a) penicillin (b) cyclosporin
    (c) chloramphenicol (d) streptomycin
22. Degradation of oil spills can be achieved through engineered bacteria utilizing genes of the strains of
    (a) *Bacillus thuringiensis* (b) *Bacillus subtlis*
    (c) *Thiobacillus ferroodoxin* (d) *Pseudomonas putida*
23. The term *bioinformatics* was coined by
    (a) Margaret Dayhoff (b) Rita-Levi Montalcini
    (c) Paulien Hogeweg (d) Dorothy Hodgekin
24. The genome sequence of the first multicellular organism deciphered in 1998 was that of
    (a) *Caenorhabditis elegans* (b) *Arabodopsis thaliana*
    (c) *Brachidanio rerio* (d) *Takefugu rubripus*
25. Thomas Rodericke introduced the term
    (a) genomics (b) proteomics (c) stammzelle (d) apoptosis
26. The technique in which marker locations are mapped by hybridizing a probe containing the marker to intact chromosomes is called
    (a) STS mapping (b) FISH technique
    (c) restriction mapping (d) genetic linkage mapping

27. In 1980, the Nobel Prize for contribution to genome sequencing was shared by Sanger and Gilbert with
    (a) Thomas Rodericke (b) Paul Berg
    (c) Allan Maxam (d) Margaret Dayhoff
28. Shotgun sequencing is designed for sequences of base pairs
    (a) less than 500 (b) between 100 and 500
    (c) between 500 and 1000 (d) more than 1000
29. Post-translationally modified proteins can be identified through a technique developed in 2008, known as
    (a) MALDI (b) mass spectrometry
    (c) SDS-PAGE (d) PROTOMAP
30. The ENCODE consortium member from Europe is
    (a) France (b) Spain (c) Germany (d) Belgium
31. The National Human Genome Research Institute (NHGRI) is located in
    (a) Washington DC (b) Chicago
    (c) Atlanta (d) Bethesda
32. When adult stem-cell types differentiate into cell types seen in organs or tissues other than those expected from the cells' predicted lineage, it is called
    (a) apoptosis (b) transdetermination
    (c) transdifferentiation (d) senescence
33. Hfr cells of *Escherichia coli*
    (a) lack the F plasmid (b) have free F plasmid
    (c) have integrated plasmid (d) are rare
34. The *Agrobacterium* species that causes hairy root disease is
    (a) *A. rubi* (b) *A. rhizogenes*
    (c) *A. tumifaciens* (d) *A. radiobacter*
35. The basis of Mendelian inheritance in *Drosophila* was elucidated at Columbia University in the laboratory of
    (a) Dobzhansky (b) Muller (c) Sturtevant (d) Morgan
36. An example of a monocotyledonous plant is
    (a) tomato (b) tobacco (c) pea (d) plantain
37. Chitin is a major component in the cell wall of
    (a) *Agrobacteria* (b) *E. coli* (c) *Dyscostelium* (d) yeast
38. The proteins fractionated by PAGE into bands may be visualized after staining with
    (a) Ninhydrin (b) Victoria blue
    (c) Congo red (d) Coomassie Brilliant Blue
39. Agarose used in gel electrophoresis is a linear polymer of
    (a) galactose (b) glucose (c) fructose (d) xylulose
40. Plant cells are cemented into tissues by
    (a) cellulose (b) chitin (c) lecithin (d) pectin
41. Long spindly cells used in majority of animal-cell culture studies are
    (a) epithelial cells (b) fibroblasts
    (c) lymphocytes (d) embryonic cells

42. Chromosomes processed for banding reveal intercalary or facultative heterochromatin regions that are represented by
    (a) C-bands (b) G-bands (c) Q-bands (d) R-bands
43. Mitotic metaphases with aberrations in one or a pair of homologous fragments indicate
    (a) inversion (b) translocation
    (c) dicentric (d) deletion
44. A material that enhances the production of antibodies and/or prolongs the life of an antigen in the body is called
    (a) epitope (b) adjuvant (c) promoter (d) facilitator
45. The class of antibodies involved in the hypersensitivity reaction is
    (a) IgE (b) IgGα (c) IgA (d) IgM
46. In immuno-electrophoretic tests, the protein that moves towards the cathode (–ve) is
    (a) α_1 globulin (b) α_2 globulin (c) β-globulin (d) ✓ globulin
47. One of the most sensitive tests to detect and estimate minute concentration of an antigen or antibody in a very small sample is
    (a) immuno-diffusion test (b) radio-immuno assay
    (c) immune-electrophoresis (d) haemagglutinin test
48. The test useful in detecting presence of pathogens or in a sample serum is
    (a) complement fixation test (b) immuno-diffusion test
    (c) immuno-electrophoresis (d) haemagglutinin test
49. In an embryo, T and B lymphocytes originate from the
    (a) spleen (b) liver (c) bone marrow (d) thymus
50. Differentiation of T-cell precursors into subpopulations takes place in the
    (a) bone marrow (b) thymus (c) liver (d) spleen

Answers

1. (b); 2. (c); 3. (c); 4. (b); 5. (d); 6. (d); 7. (a);
8. (b); 9. (c); 10. (b); 11. (d); 12. (c); 13. (b); 14. (b);
15. (a); 16. (b); 17. (b); 18. (b); 19. (a); 20. (b); 21. (b);
22. (d); 23. (c); 24. (a); 25. (a); 26. (b); 27. (b); 28. (d);
29. (d); 30. (b); 31. (d); 32. (c); 33. (c); 34. (b); 35. (d);
36. (d); 37. (d); 38. (d); 39. (a); 40. (d); 41. (b); 42. (b);
43. (c); 44. (b); 45. (a); 46. (d); 47. (b); 48. (a); 49. (b);
50. (b)

Index

C

D

E

F

G

H

I

J

K

L

M

N

O

P

Q

R

S

T

U

V

W

X

Y

Z