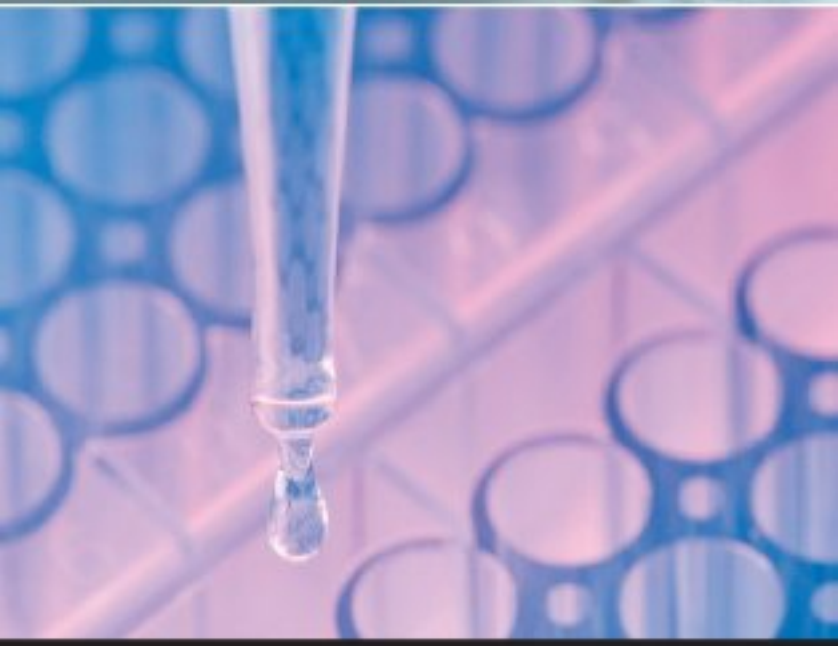

CHALLENGES IN DELIVERY OF THERAPEUTIC GENOMICS AND PROTEOMICS

Edited by AMBIKANANDAN MISRA

Challenges in Delivery of Therapeutic Genomics and Proteomics

Challenges in Delivery of Therapeutic Genomics and Proteomics

Edited by

Ambikanandan Misra

Pharmacy Department
TIFAC – Centre of Relevance and Excellence
in New Drug Delivery Systems
The Maharaja Sayajirao University of Baroda
Vadodara, India

AMSTERDAM • BOSTON • HEIDELBERG • LONDON • NEW YORK • OXFORD
PARIS • SAN DIEGO • SAN FRANCISCO • SINGAPORE • SYDNEY • TOKYO

Elsevier
32 Jamestown Road London NW1 7BY
30 Corporate Drive, Suite 400, Burlington, MA 01803, USA

First edition 2011

Notices
Knowledge and best practice in this field are constantly changing. As new research and experience broaden our understanding, changes in research methods, professional practices, or medical treatment may become necessary.

Practitioners and researchers must always rely on their own experience and knowledge in evaluating and using any information, methods, compounds, or experiments described herein. In using such information or methods they should be mindful of their own safety and the safety of others, including parties for whom they have a professional responsibility.

To the fullest extent of the law, neither the Publisher nor the authors, contributors, or editors, assume any liability for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions, or ideas contained in the material herein.

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Cataloging-in-Publication Data
A catalog record for this book is available from the Library of Congress

ISBN: 978-0-12-384964-9

Preface

All human diseases are genetic in origin. It is becoming increasingly clear that the root cause of virtually all human afflictions, from cancer to psychiatric disorders, as well as susceptibility to infection, lies in our genes. Genetic information is passed on to progeny through a process tightly controlled by regulatory factors acting in consortium. The genetic message is decoded to form functional biomolecules that are crucial effectors and regulators of biological processes. Trafficking of endobiotics and xenobiotics through transporters and cell–cell communication through various receptors is essential for the maintenance of cellular homeostasis. An increased understanding of molecular biology instigated research in the application of proteomics and genomics in the treatment of human diseases. This compilation is an effort to encompass the knowledge about the origins of human diseases, systems biology, and various aspects of personalized medicine and challenges in the field, to make application of genomics and proteomics a therapeutic reality.

The intracellular delivery of therapeutic proteomics and genomics is possible by overcoming and manipulating several uncompromising barriers. Therapeutic genomics delivery involves successful entry into the cell after overcoming extracellular diffusion barriers and enzymatic degradation, hepatic uptake and degradation, and the reticulo-endothelial system (RES) uptake barrier to finally reach the cells of the pathogenic organ for expression. Cytoplasm provides physical and metabolic barriers to further trafficking of genomics, and nuclear uptake is the most significant barrier to gene delivery. However, contribution of each barrier to the genomics delivery and its expression in different cells and ways to overcome it needs to be deeply understood in order to obtain clinically significant expression. The processes involved in each step of therapeutic genomics delivery and progress made in understanding the cellular barriers in gene transfer and expression have been explained in Chapter 2.

Nonviral gene delivery is concerned with physical and chemical methods for safe and efficient transgene expression. Chapter 3 summarizes the basic principles, mechanisms, and protocols of the various physical methods used in gene transfer. Minimum toxicity with maximum transfection efficiency and kinetics (i.e., rate and duration) and *in vivo* feasibility and reproducibility of all physical methods of transgene expression are discussed and compared. The chapter elaborates biological and cellular responses to the process of physical stimulation involving gene transfer. Chemical vectors enjoy a high profile due to the safety and stability advantages they offer as compared to viral approaches. However, these vectors exhibit comparatively low efficiency and show inability to target gene expression to the area of pathology until functionalized specifically. Targeting ligands have been used with moderate success to overcome specific barriers, while endosomal escape and nuclear targeting

peptides are some of the strategies under investigation. No major clinical toxicities of nonviral vectors have been reported in the results of clinical trials. Practically, cationic lipids, cationic polymers, and other naturally occurring compounds have proven to be extremely effective for *in vitro* gene delivery. All classes of chemical vectors are discussed in Chapter 4, and their comparative abilities in the transfection of genes.

The development of viral vectors such as adenovirus (Ad), adeno-associated virus (AAV), retrovirus, and herpes simplex viruses have advanced to clinical trials. Other viral vectors like baculovirus, lentivirus, influenza virus, human papillomavirus, and hepatitis B virus have also been successfully used as gene delivery vectors. A roadblock to the use of viral vectors in gene therapy is the potential immunogenic reaction due to the biological nature of the vector; moreover, the use of viral vectors necessitates repeated administration, which can cause a cumulative immune response. The continuous efforts of research scientists have raised the hope of convenient, safe, stable, and reproducible transduction of functional genetic material into specific cells. Because of their inherent construction, viral vectors are advantageous for targeting and efficient transport into the cells. Furthermore, they allow modification according to the required application, as discussed in Chapter 5.

Gene therapy offers an incredibly powerful tool for preventing and curing diseases because it is aimed to treat the cause rather than symptoms of diseases. A number of clinical trials are underway for human diseases with genetic causes as well as for acquired diseases. The results of these trials using both viral and nonviral vectors show the enormous potential of gene therapy for treatment of both genetic and acquired diseases, but there is still a long way to go before they can be used in regular clinical practice. Continued research in the area of gene therapy can develop the "gene as a substitute for drug" in the near future. Chapter 6 discusses the progress in applications of gene therapy in clinical practice. Gene therapy is becoming the eventual choice for treatment of various devastating genetic and acquired diseases that occur because of gene malfunction. However, it requires a multidisciplinary effort from gene therapy scientists in applying knowledge of many sciences to construct and manipulate vectors with enhanced gene expression and low immune response or toxicity.

Antisense technology presents an opportunity for utilizing antisense oligonucleotides, ribozymes, short interfering RNA, micro RNA, and aptamers to manage diseases by regulating gene expression. Highly specific and effective gene silencing can be achieved by an in-depth knowledge of the target mRNA sequence and rational design of its complementary antisense agents. Chapter 7 discusses the intricacies, delivery options, and applications of antisense technology. Extensive and appropriate pharmacological and toxicological studies are prerequisite at the initial development and preclinical stages to avoid failures at later clinical phases. The simplicity of siRNA design, its specificity, potency, the availability of human genome information, feasibility of fabrication into the required sequence, and endless applications make it a promising tool for treating otherwise difficult-to-treat diseases.

Proteomics deals with the study of proteins, their structure, and their characteristics as well as their functions in a biological system. Chapter 8 provides an insight into the conventional, classical methods of protein synthesis along with an overview of recently

developed biotechnology-based techniques for cost-effective protein production. An attempt is made to provide an outline regarding the objectives of proteomic studies, their role in diagnostics and therapeutics, the changing needs of peptide drug delivery, rate-limiting biological barriers, and factors that affect the *in vivo* performance of proteins and peptides as pharmaceuticals. We also describe briefly the sophisticated, smart drug delivery systems used for controlled protein delivery and stability and safety issues pertinent to protein drug delivery, high resolution, and throughput techniques used for protein characterization.

Although parenteral route is most efficient for proteins, it is not always suitable from the patient's perspective. The lung has been recognized as an efficient route for systemic drug delivery, particularly for macromolecules such as peptides and proteins. Also, inhalation therapy as aerosols targets the lungs through the pulmonary airways. The instant access and high concentration of the drug deposited within the lungs noninvasively are major advantages of aerosol delivery over other routes of administration. The latest findings in the research of drug development for protein products delivered via intranasal route have also shown promising results. It is very likely that in the near future more drugs intended for systemic absorption through noninvasive routes of administration will come into the market. Chapter 9 elaborates on the promise and limitations of these noninvasive techniques for therapeutic protein delivery through respiratory pathways.

Increasingly, therapeutic peptides and proteins are being explored for delivery through the oral route because of advances made in the understanding of the biological and physicochemical factors influencing absorption and proteolytic enzymatic degradation. Research in the field has focused on improving bioavailability from almost nil to at least 10–20%. Methods to increase drug flux (e.g., use of permeation enhancers) without associated toxicity, approaches to minimize proteolytic degradation, chemical modifications and innovative approaches with regard to mucoadhesive dosage forms, and targeted as well as controlled drug delivery have been employed to improve oral peptide and protein delivery. Promising results have begun to appear, and oral delivery of insulin is currently in clinical trials. These advances have been addressed in Chapter 10.

Approximately 30–40% of recent new drug approvals are in the biological category, including antibodies, proteins, and peptides. Chapter 11 elaborates on parenteral delivery of these protein-based compounds. Most of the protein-based therapeutics are normally administered by frequent injections through different parenteral routes, such as intramuscular and subcutaneous. Frequently, the product is a lyophilized form of the protein along with an excipient or any particulate or carrier system that has to be reconstituted before administration. This combination of a lyophilized protein or protein along with the carrier system and the subcutaneous route seems at present to be the most robust and successful design.

Various alternative routes investigated for protein and peptide delivery like transdermal, topical, uterine, and rectal are discussed in Chapter 12. The transdermal route includes chemical approaches, that is, the use of prodrug, permeation enhancers, and protease inhibitors, and physical approaches like sonophoresis, microporation, and electrical methods (electroporation and iontophoresis), and combinations thereof.

Topical delivery of peptide and protein drugs is important for treating dermatological ailments. The rectal route provides systemic bioavailability by avoiding first-pass elimination. Site-specific deliveries are logical choices for reducing the total dose to be delivered, gaining access to specific organs, and for delivering therapeutic dose for site-specific pharmacological action.

The value of this book depends on understanding and applying the gathered knowledge in resolving the limitations of existing technologies and in creating new opportunities with due consideration given to regulatory aspects right from the development stage. We must adopt these technologies, but keeping in mind the ultimate aim of actual therapeutic benefit of genomics and proteomics to the patients.

It gives me great pleasure to extend my thanks to all contributors who brought together their collective experience, knowledge, skills, and wisdom to bestow upon this book its present shape. Thanks are also due to the authors and publishers whose contributions have been used by us through permission.

Ambikanandan Misra

List of Contributors

Himanshu Agrawal Pharmacy Department, The Maharaja Sayajirao University of Baroda, PO Box 51, Kalabhavan, Vadodara 390 001, Gujarat, India

Tamishraha Bagchi Department of Microbiology and Biotechnology, Faculty of Science, The Maharaja Sayajirao University of Baroda, Vadodara 390 001, Gujarat, India

Kiruba Florence Pharmacy Department, The Maharaja Sayajirao University of Baroda, PO Box 51, Kalabhavan, Vadodara 390 001, Gujarat, India

Kaustubh A. Jinturkar Pharmacy Department, The Maharaja Sayajirao University of Baroda, PO Box 51, Kalabhavan, Vadodara 390 001, Gujarat, India

Gitanjali Kher Pharmacy Department, The Maharaja Sayajirao University of Baroda, PO Box 51, Kalabhavan, Vadodara 390 001, Gujarat, India

Mukesh Kumar Pharmacy Department, The Maharaja Sayajirao University of Baroda, PO Box 51, Kalabhavan, Vadodara 390 001, Gujarat, India

Manisha Lalan Pharmacy Department, The Maharaja Sayajirao University of Baroda, PO Box 51, Kalabhavan, Vadodara 390 001, Gujarat, India

Jigar Lalani Pharmacy Department, The Maharaja Sayajirao University of Baroda, PO Box 51, Kalabhavan, Vadodara 390 001, Gujarat, India

Ambikanandan Misra Pharmacy Department, TIFAC – Centre of Relevance and Excellence in New Drug Delivery Systems, The Maharaja Sayajirao University of Baroda, PO Box 51, Kalabhavan, Vadodara 390 001, Gujarat, India

Deepa H. Patel Pharmacy Department, The Maharaja Sayajirao University of Baroda, PO Box 51, Kalabhavan, Vadodara 390 001, Gujarat, India

Gaurang Patel Pharmacy Department, The Maharaja Sayajirao University of Baroda, PO Box 51, Kalabhavan, Vadodara 390 001, Gujarat, India

Mohan N. Rathi Pharmacy Department, The Maharaja Sayajirao University of Baroda, PO Box 51, Kalabhavan, Vadodara 390 001, Gujarat, India

Bhavik Shah Pharmacy Department, The Maharaja Sayajirao University of Baroda, PO Box 51, Kalabhavan, Vadodara 390 001, Gujarat, India

Tapan R. Shah Post Graduate Department of Pharmaceutical Sciences, Sardar Patel University, Vallabh Vidyanagar 388 120, Gujarat, India

Naazneen Surti Baroda College of Pharmacy, Parul Arogya Seva Mandal, Limda, Waghodia 391 760, Gujarat, India

Hemal Tandel Pharmacy Department, The Maharaja Sayajirao University of Baroda, PO Box 51, Kalabhavan, Vadodara 390 001, Gujarat, India

Sonia Trehan Pharmacy Department, The Maharaja Sayajirao University of Baroda, PO Box 51, Kalabhavan, Vadodara 390 001, Gujarat, India

Nipa Thacker Pharmacy Department, The Maharaja Sayajirao University of Baroda, PO Box 51, Kalabhavan, Vadodara 390 001, Gujarat, India

1 The Cell

Manisha Lalan[1], *Tamishraha Bagchi*[2], *Ambikanandan Misra*[1]

[1]Pharmacy Department, TIFAC – Centre of Relevance and Excellence in New Drug Delivery Systems, The Maharaja Sayajirao University of Baroda, PO Box 51, Kalabhavan, Vadodara 390 001, Gujarat, India
[2]Department of Microbiology and Biotechnology, Faculty of Science, The Maharaja Sayajirao University of Baroda, Vadodara 390 001, Gujarat, India

1.1 Introduction

The cell is the most fundamental and primary unit of life. These membrane-enclosed units are composed of an aqueous sol and endowed with a unique capability of replication. Lower organisms are unicellular, whereas higher organisms like us are communities of a wide variety of cells working in consortium to perform the essential functions. Although all living cells exhibit a similar basic cellular chemistry, there is enormous variance in the appearance and function of cells. Although the solitary cell of unicellular organisms can perform all the vital functions, the cell groups in multicellular organisms are differentiated to perform specific functions.

Cells have been classified as prokaryote, or primitive, and eukaryote, according to a very basic fundamental feature: the absence or presence of a well-defined nucleus that contains genetic information. Prokaryotes have been synonymous with bacteria. Most prokaryotes live as unicellular organisms, but they also exhibit association in the form of chains, clusters, and so on. These have been known to survive on almost anything, from wood to petroleum, in the most extreme environmental conditions. Oxygen is essential to some prokaryotes, but fatal for others. Although some can perform photosynthesis, other prokaryotes are totally dependent on others for survival. *Escherichia coli* is one of the most widely studied bacterium, by scientists seeking to understand the basic principles of life; *E. coli* is also used as a biotechnological tool.

Eukaryotic cells are larger and more complex than prokaryotes (Fig. 1.1), but they may also be unicellular or multicellular. They are characterized by a well-defined nucleus. Apart from the nucleus, there are other organelles that are key to their survival and homeostasis. In general, the biochemical makeup and the organelles of the cells are similar in fundamental ways. In this chapter, only major cellular organelles that perform vital functions are discussed.

Challenges in Delivery of Therapeutic Genomics and Proteomics. DOI: 10.1016/B978-0-12-384964-9.00001-3

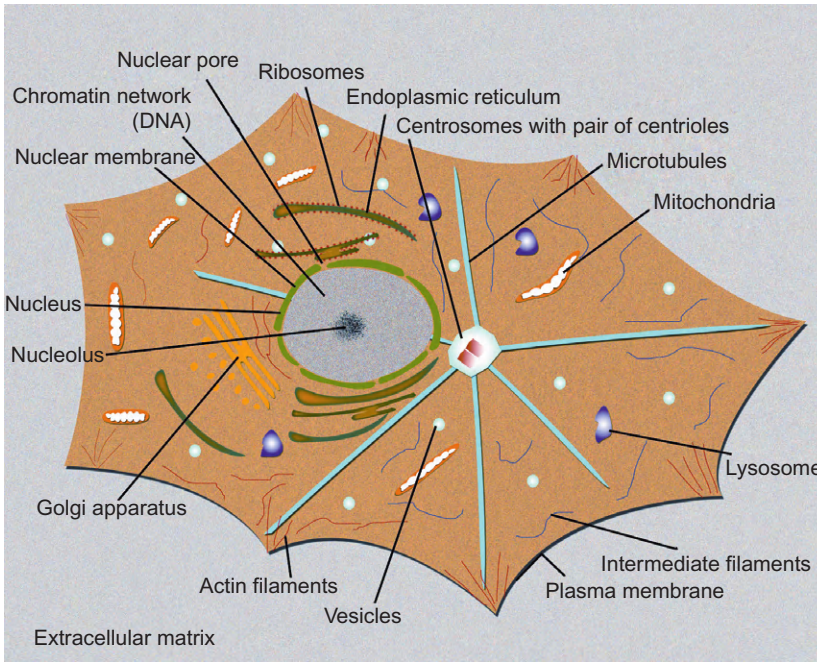

Figure 1.1 Animal cell and its internal organization.

Plasma Membrane

The plasma membrane forms the outermost boundary of the cell, enclosing the cytosol and organelles. The membrane is made up of a phospholipid bilayer with embedded proteins. It is a quasi-fluid structure where the membrane glycoproteins, in addition to serving as channels for transport of specific molecules, are also involved in physical interaction with other cells, as well as in signal transduction. The carbohydrate moieties in the cell membrane are mostly linked with the proteins and lipids, giving rise to glycocalyx on the outer side of the cell, which imparts the characteristic negative surface charge to the cell. Although all biomembranes have the same basic phospholipid bilayer structure and certain common functions, each type of cellular membrane also has certain distinctive activities determined largely by the unique set of proteins associated with that membrane.

Cytoplasm

Enclosed inside the plasma membrane is the cytoplasm—which can further be categorized as cytosol and the organelles. The organelles include the nucleus, mitochondria, peroxisomes, endoplasmic reticulum (ER), Golgi body, endosomes, and lysosomes. The cytosol is the liquid phase in which the organelles are suspended,

and it is the site for protein synthesis, metabolism, and various other reactions. Although diffusion governs the movement of soluble material through the cytosol, larger molecules, complexes, and organelles move around by active transport.

Nucleus

The nucleus is the control center of the cell. A bilayer membrane with pores encloses the nuclear material and forms a formidable barrier to any unwanted molecular penetration. Material transport is permitted through the nuclear pores, about 60–100 nm in dimension, and is regulated by more than 50 different proteins called nucleoporins. These allow easy unregulated passage of molecules of 5000 Da while progressively restricting the larger molecules and regulating their transport. The nucleus contains the genetic information of the cells in the form of deoxyribonucleic acid (DNA), which in turn is packaged as chromosomes and associated with proteins called histones. Within the nucleus is the nucleolus, a spherical body composed of DNA, proteins, and ribonucleic acid (RNA); the nucleolus is the site of RNA synthesis.

Endoplasmic Reticulum

Endoplasmic Reticulum (ER) is the meshlike network of flattened tubular sacs found throughout the cytosol and also extending to the nuclear membrane. The structure at places may be bound with ribosomes (rough ER); in other regions it may be smooth in texture (smooth ER). Smooth ER is the site of lipid synthesis, whereas the rough ER is the site of protein synthesis. Smooth ER also transports proteins and peptides by forming transport vesicles. Endoplasmic reticulum provides the machinery for metabolic functions of the cells.

Lysosomes

These membrane-enclosed organelles, termed the "suicide bags of the cell," perform the function of breaking down macromolecules and ingested material. The intracompartmental pH of around 5 aids in hydrolytic enzymatic activity. The intracellular digestive activity of lysosomes recycles some of the macromolecules by breaking them down into constituting monomers, and lysosomes also help rid the cell of toxic molecules and damaged cellular structures.

Mitochondria

Mitochondria are the powerhouses of the cell. They are double-membrane-bound structures having their own DNA and ribosomes, and hence are self-replicative. They are one of the most abundant organelles in the cytosol. Mitochondria are the residence site for many oxidative enzymes, which extract energy from sugars and fats to generate adenosine triphosphate (ATP).

Endosomes

Endosomes are membrane-bound vesicles that have a very important role in the endocytic pathway. They are involved with the movement of material internalized from the plasma membrane to lysosomes for degradation, or they may recycle the material back to the plasma membrane itself. Endosomes also transport material from the Golgi body to lysosomes, and they provide an environment for material to be sorted before it reaches the degradative lysosome. Endosomes are characterized as early, late, and recycling endosomes differentiated by their role and specific markers present on them.

Peroxisomes

Peromisomes are single-membrane-bound vesicles harboring various oxidase enzymes. Primarily, they regulate fatty acid metabolism and remove toxic peroxides from the body.

Golgi Body

The Golgi body is composed of membrane-bound cisternae, important organelles of the cell, that deal with processing and secretion of proteins and lipids in vesicles.

Cytoskeleton and Intracellular Movement

The cell is crisscrossed with a network of cytoskeletal structures that provide it necessary structural support while imparting resilience and capacity to bear external stress. The three major types of cytoskeletal structures include microtubules, actin filaments, and intermediate filaments, which fulfill different functions of the cell. Microtubules are comparatively firm structures composed of tubulin, originating from the centrosomes and extending into the cytosol. They exhibit dynamism because they can elongate or shrink by the addition or loss of tubulin subunits. The nucleation process is a guanosine triphosphate (GTP)-aided mechanism, and the growing ends of the microtubules are stabilized by a few cytosolic structures. This stability-enhancing attachment determines the cellular location of the microtubules. These are further stabilized by specific microtubule-associated proteins (MAPs). The microtubule motor proteins, kinesins, and dynesins constitute MAP and mediate the movement along microtubules. Actin filaments are also similar to microtubules in the dynamism exhibited; they too are rapidly formed and lost. Actin filaments are present as bundles or networks rather than as filaments. They form a cortical ring just underneath the plasma membrane and are key regulators of cellular shape and movement, as well as playing a crucial role in cytokinesis and phagocytosis. The movement along actin filaments is effected through myosin protein. Intermediate filaments, long polymers of fibrous polypeptide providing the resilience and stress-bearing property to the cell, exhibit different chemical composition in different tissues. For example, they are present as keratin filaments in epithelial cells and as

desmin filaments in muscle cells. Both ends of these coiled structures are capable of extending. All three types of filaments act in consortium and mediate the intracellular and cellular movement [1–6].

1.2 Genes, Chromosomes, DNA, and RNA

Nearly all cells of an organism contain similar genetic material called its genome. Scientists have defined the gene as that part of the chromosome that encodes the necessary information for the synthesis of a functional biological product, which could be either a protein or RNA. The units carrying genetic information are made up of DNA. Besides DNA, the cells contain another type of nucleic acid, RNA, which is produced by transcription of DNA, as discussed later in this chapter. The nucleic acid molecule is made up of nucleotides, which in turn have three basic units: a nitrogenous base, a pentose sugar, and a phosphate group. A phosphate-less molecule is termed a nucleoside. The nitrogenous bases found in the nucleic acids are of five different types—adenine (A), guanine (G), cytosine (C), thymine (T), and uracil (U)—derived from purine and pyrimidine parent structures. DNA is known to be made up of A, T, G, and C, while RNA has A, U, G, and C. The pentose sugar is also different in the two nucleic acids. DNA contains deoxyribose, and RNA has ribose present in its structure.

Apart from the regular purines and pyrimidines, other bases, usually in methylated forms, may also be present. The nucleotides of DNA and RNA are linked to each other by phosphodiester linkage, wherein the 5′ phosphate group of one is attached to the 3′ hydroxyl end of the other. So the backbone structure has the sugar and phosphate group linked with the nitrogen base protruding from the backbone at regular intervals (Fig. 1.2). These phosphodiester linkages have the same orientation, giving a specific 3′ and 5′ end to the nucleic acid and thus a specific polarity.

A short linear polymer of nucleotide is termed an oligonucleotide, whereas longer chains containing 50 or more nucleotides are called polynucleotides. The DNA molecule is made up of two polynucleotide chains, or DNA strands, held together by the hydrogen bonding that exists between their nitrogen bases. The nitrogen-base pairing is very specific; A always pairs with T with the help of two hydrogen bonds, and G pairs with C with the help of three hydrogen bonds. In essence, a purine always pairs only with a pyrimidine. This specific or complementary base pairing allows the two strands to be held together in a thermodynamically favorable antiparallel double helix. The sugar phosphate backbones twist around one another to give shape to this double helical structure. The hydrophilic backbone of sugar and phosphate groups is on the outer side, whereas the hydrophobic nitrogen bases are on the inner side; this creates a major and minor groove on the surface of the duplex. Crystal analysis of DNA revealed that the bases stacked in the helix are 3.4 Å apart, and it takes a complete turn for every 10 bases, in other words, at 34 Å.

Erwin Chargaff and coworkers showed that the DNA in an organism does not change because of external or internal factors like age, nutrition, and so forth. The

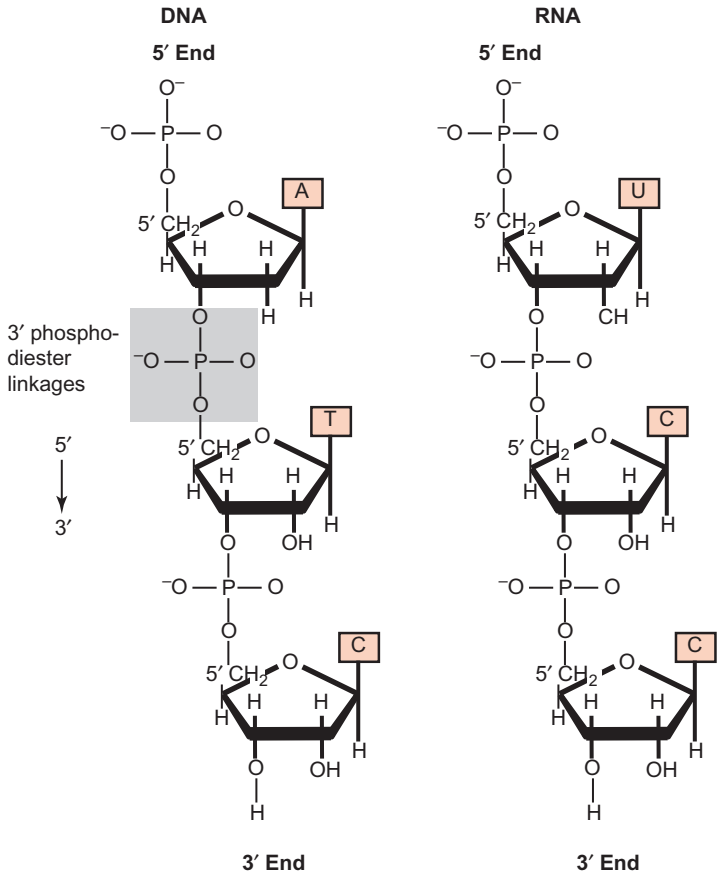

Figure 1.2 Nucleotide chain seen in DNA and RNA.

base composition remains identical in different tissues of the same species, and quantitatively the number of purine residues and pyrimidine residues are the same.

The DNA is too large compared to the cell in which it is packaged. The organization and structure of DNA is comparatively simple, both structurally as well as functionally in prokaryotes. The bacterial chromosome is usually a long, single circular DNA molecule, whereas some organisms may have more than one unit. Apart from this large, single, circular DNA, some extrachromosomal DNA called plasmids are also present in numbers ranging from a few thousand to several thousand base pairs. Plasmids are self-propagating in nature, and daughter plasmids are passed on to progeny. These plasmids generally carry genes useful to the host; for example, some confer the host with resistance to antibiotics.

The organization of eukaryotes, however, is complex. Eukaryotic genes are known to carry nucleotide sequences that are not coding for any biological product interspersed between the coding regions. Such intervening sequences are termed introns,

and the part of the gene that is translated is called an exon. In fact, it is claimed that only 1.5% of the total human genome consists of exon. Also the presence of transposable elements is very common—short, repeated sequences of DNA that keep moving in the genome; these are also called molecular parasites. About 3% of the genome consists of simple sequence repeats that are a few base pair long sequences known as "satellite" bands. There are specific areas in chromosomes that, by virtue of their unique sequences or base composition, help the chromosomes in particular functions; for example, the centromere is the site where the chromosome attaches to spindle fibers with the aid of proteins. These sequences are usually rich in A=T pairs. The sequences, called telomeres, help to stabilize the chromosome. The sequence-specific changes in chromosomes become very crucial in certain diseases and can be used for other applications, hence should be characterized. For example, the simple sequence repeats of the chromosome are individual specific because of unequal crossing over during meiosis. The difference can be traced by Southern blotting, polymerase chain reaction and forms the basis for DNA fingerprinting [7–11].

RNA

The second major type of nucleic acid present in cells is RNA, of which there are three types: ribosomal RNA (rRNA), messenger RNA (mRNA), and transfer RNA (tRNA). All three have specific intermediary functionality to ensure that the genetic message encoded in the DNA is expressed. Because DNA is present mostly in the nuclear region in the cell, whereas the site for protein synthesis is in cytoplasm, a mechanism for transfer of the genetic message is essential. RNA is the molecule that fills this lacuna. This is evident because RNA is present both in the nuclear region and in the cytoplasm, and increased protein synthesis is seen with increased levels of cellular RNA. Single-stranded RNA results from the transcription of DNA, which also assumes a right-handed helical conformation. Weak base stacking interactions stabilize its structure, which is unique like proteins and unlike DNA. Among the three types of RNA, rRNA is associated with ribosomes, the machinery for protein synthesis. mRNA are messenger molecules that transfer the genetic information from DNA to the ribosomes to allow protein synthesis. mRNA serves as the template that codes for the amino acid chain (polypeptide). If mRNA codes for one polypeptide, it is referred to as monocistronic, and it is called polycistronic when several polypeptides are encoded in its message. Most of the mRNA is monocistronic, and its length varies with the polypeptides it codes for. Apart from the coding region, there are additional bases present that may serve to control protein synthesis. tRNA are molecules that transfer the information of mRNA into the amino acid chain by serving as adaptors. tRNA carries a specific amino acid on its end and attaches with the mRNA to transfer the amino acid to the growing amino acid chain [12,13].

Properties of Nucleic Acid

Isolated DNA solutions exhibit a marked decrease in viscosity on heating. The thermal change causes it to denature, or melt down, because of the breaking of hydrogen

bonds between base pairs and consequent strand separation. This separation may be partial or complete, and leads to an increase in the ultraviolet (UV) absorption properties called hyperchromicity. The DNA reverts back to double helical structure or anneals when temperature or pH is favorable. Complete separation requires a longer time to anneal compared to partial separation. First, a small region anneals by collisions, and then the whole strand zippers together very fast. DNA double-helix formation can therefore be confirmed from its hypochromic effect on UV absorption. The meltdown temperature is species specific and determined by base composition of DNA in that species. Species having higher G–C content show higher meltdown temperature on account of tighter bonding between G–C compared to weaker bonding between the A–T base pair. RNA duplexes and RNA–DNA hybrids can also be denatured in a similar way. If DNA from two species are denatured together in a solution and allowed to anneal, some hybrid duplexes also form, indicating common evolution. Hybridization is an important biotechnological tool for detecting a specific RNA or DNA.

Mutations are the permanent alterations in the DNA sequence structure and affect the genetic message and thus its expression. There are several reasons for the mutative changes such as oxidative damage, presence of alkylating agents, exposure to near UV radiations, spontaneous deamination of bases, and so on. All such reasons bring about permanent and inheritable changes in the genome, resulting in faulty expression of the genetic message [14–16].

1.2.1 DNA Supercoiling

There are several stages in the life of a cell, and the structure of chromosomes also varies with the stage. During interphase, they exist as a network of long, entangled, thin strands of DNA, and as the cell approaches the division stage, they become condensed and compacted.

The long-stranded DNA molecules are condensed into the chromosome to allow packaging into the eukaryotic chromosome and also to aid in their allocation into daughter cells during cell division. A number of proteins help the DNA to coil and wind, organizing into loops, rosettes, and so forth, with an increasing complexity that enables it to condense into manageable size without entanglement (Fig. 1.3). However, this organization is so perfectly managed that essential activities like replication, repair, transcription, and so on, are not hindered. Each chromosome is a long linear molecule that links with several proteins to form a compacted structure. This DNA and protein network is called a chromatin. Each somatic cell contains pairs of similar chromosomes called homologous chromosomes.

The proteins that bind to the DNA are known as histone proteins and nonhistone proteins. The first level of compaction is a nucleosome, which is a “beads-on-a-string” structure where the core is a group of eight histones, (two of each H2A, H2B, H3, and H4) and the DNA is the string. They are further condensed into a 30-nm chromatin fiber with the aid of the fifth histone, H1. The nucleosomes become organized into regular repeating arrays providing around 100-fold condensations. Further compaction is less understood but involves attachment of a 30-nm fiber to

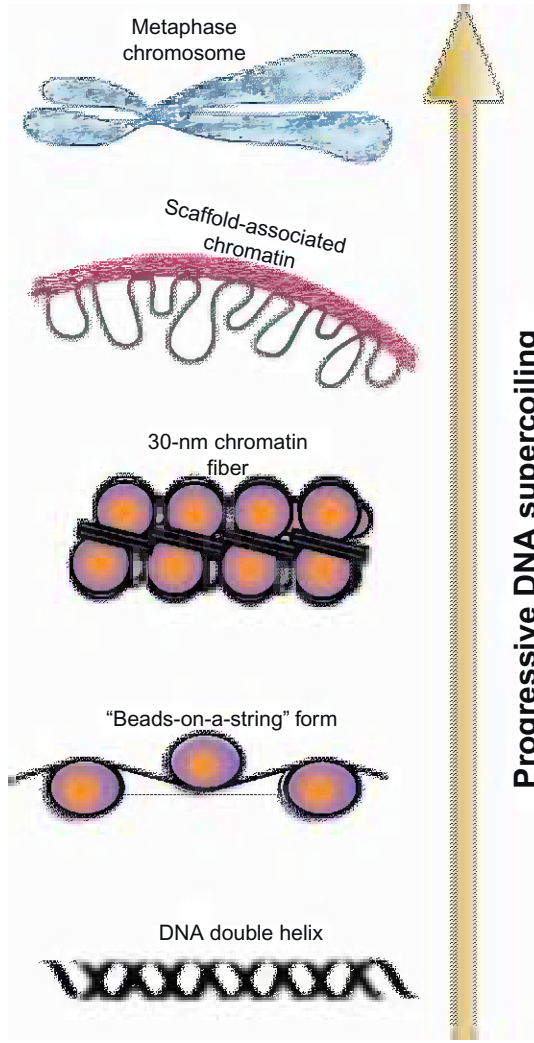

Figure 1.3 Progressive packing and condensation in the chromosome.

a nonhistone chromosome scaffold. Another class of chromatin proteins function as structure maintenance proteins called cohesins and condensins. These help in chromosome condensation and keep the sister chromatids attached during replication. Enzymes such as helicases and topoisomerases present in the cell are responsible for unwinding, or relaxation, of the DNA.

The size, number, and shapes of chromosomes traced at the metaphase form its karyotype. During the metaphase stage of cell cycle, the chromosomes are condensed to the maximum. They can be visualized by dye staining using Giemsa dye. A newly

developed method that involves visualizing the chromosome in bright colors is called chromosome painting, wherein a fluorescent probe specific for sites on the chromosomes is used [17–23].

1.2.2 DNA Replication, Repair, and Recombination

It has been seen that the two strands of DNA have complementary base pairing, thus enabling each to serve as a template for the synthesis of a new complementary strand. This is the basis of DNA replication and recombination. In eukaryotes, the replication machinery, a group of proteins, ensures correct and efficient replication of the DNA strand. The DNA replication is semiconservative in nature, with each double helix generated by replication containing one parent strand. DNA replication begins with separation of the two strands at the specific replication origins—regions rich in A=T base pairs—brought about by origin recognition complex and some other proteins like CDC6, CDT1, and MCMs. The DNA double helix is unwound near the replication origin, and in this place a Y-shaped structure called a replication fork is generated by new strand synthesis along the parent strand. A protein called single-strand binding protein or replication protein A attaches to the unwound DNA strand and prevents reformation of the base pairs by hydrogen bonding. The enzyme topoisomerase relieves the stress created from the unwinding of DNA strands. DNA polymerase, the most important enzyme of the replication machine, is involved in the synthesis of new strands by catalyzing the addition of nucleotides at the 3′ end by esterification. However, the polymerase requires a primer to which it can add nucleotides. An oligonucleotide RNA/DNA usually serves as a primer that is synthesized by enzyme DNA polymerase α. Subsequently, enzyme DNA polymerase δ moves along the DNA strand as it attaches more and more nucleotides. The direction of elongation is always in 5′→3′ direction. Because the two strands are antiparallel, the nucleotide addition cannot proceed continuously in the same direction on both strands.

The continuous strand is called the leading strand. On the other strand, DNA is synthesized in small fragments (Okazaki fragments) in 5′→3′ direction discontinuously and then ligated together (Fig. 1.4). The fidelity of the process is well assured. The replication mechanism is precise and accurate because DNA polymerase also exhibits a proofreading activity by 3′→5′ exonuclease activity, whereby it removes any wrong nucleotide added. Another protein known as sliding clamp protein/replication factor C prevents detachment of DNA polymerase from the DNA strand but releases it once an Okazaki fragment ends. The ends of the chromosomes have special sequences called telomeres. The enzyme telomerase binds to the telomeres and helps to end the replication mechanism. Finally, the primers of RNA are removed and replaced with DNA and then ligated to the newly formed strand by DNA ligases.

When a cell's replication and repair process fails, a rare event, a permanent change called mutation results. DNA also continuously incurs chemical changes like insertion or deletion, substitution, depurination, deamination, dimer formation, and so forth. The cell has several repair mechanisms in place. In the mismatch repair system, a group of repair proteins recognizes these mistakes, cuts the damaged DNA strand, and adds nucleotides at the 3′ end complementary to the undamaged strand.

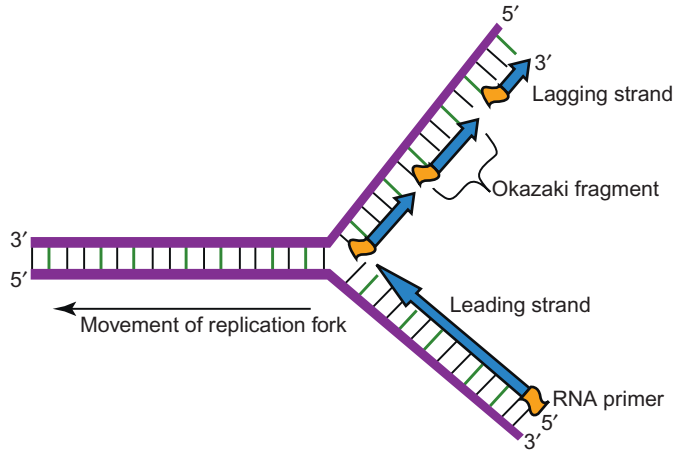

Figure 1.4 DNA replication—replication fork.

Finally, the broken ends are sealed by DNA ligase. Base excision repair mechanism requires DNA glycolsylases that recognize specific deamination and remove the base by cleaving the *N*-glycosyl bond, then repairing it using endonucleases. In some cases where the whole nucleotide is excised, the distortion is repaired by making cuts at a distance of a few base pairs from the lesion site and then resynthesizing the DNA strand using DNA polymerase ε and ligating the two DNA pieces. In some other cases like dimer formation, the lesion is corrected without excision. Other cases may involve damage to the complementary strand itself, like double strand breaks, crosslinks, or damages encountered on the replication fork. In a few such cases, a recombinational repair system involving homologous chromosomes helps in repair; otherwise, error-prone translesional DNA synthesis takes up the repair responsibility. But the chances of mutation are very high in such cases as repair is not very accurate [24–31].

DNA Recombination

The sequence of genes present often is altered or changed due to continuous rearrangement mechanisms. These rearrangement mechanisms aid in the evolution/adaptation in response to changes in external environments. They also constitute a type of specialized DNA repair system. The precise mechanism by which such rearrangements take place is termed genetic recombination.

The basic recombination is homologous recombination, where the chromosomes having regions with similar base pair sequences align near each other, and then crossover takes place between chromatids. The double strands of DNA are broken, and some of the nucleotides are digested by exonuclease activity. The cut ends seek a similar region in the opposite strand and join there. This brings about branch migration and creates a crossover structure called Holliday intermediates. These Holliday intermediates undergo cleavage to produce two recombination products (Fig. 1.5)

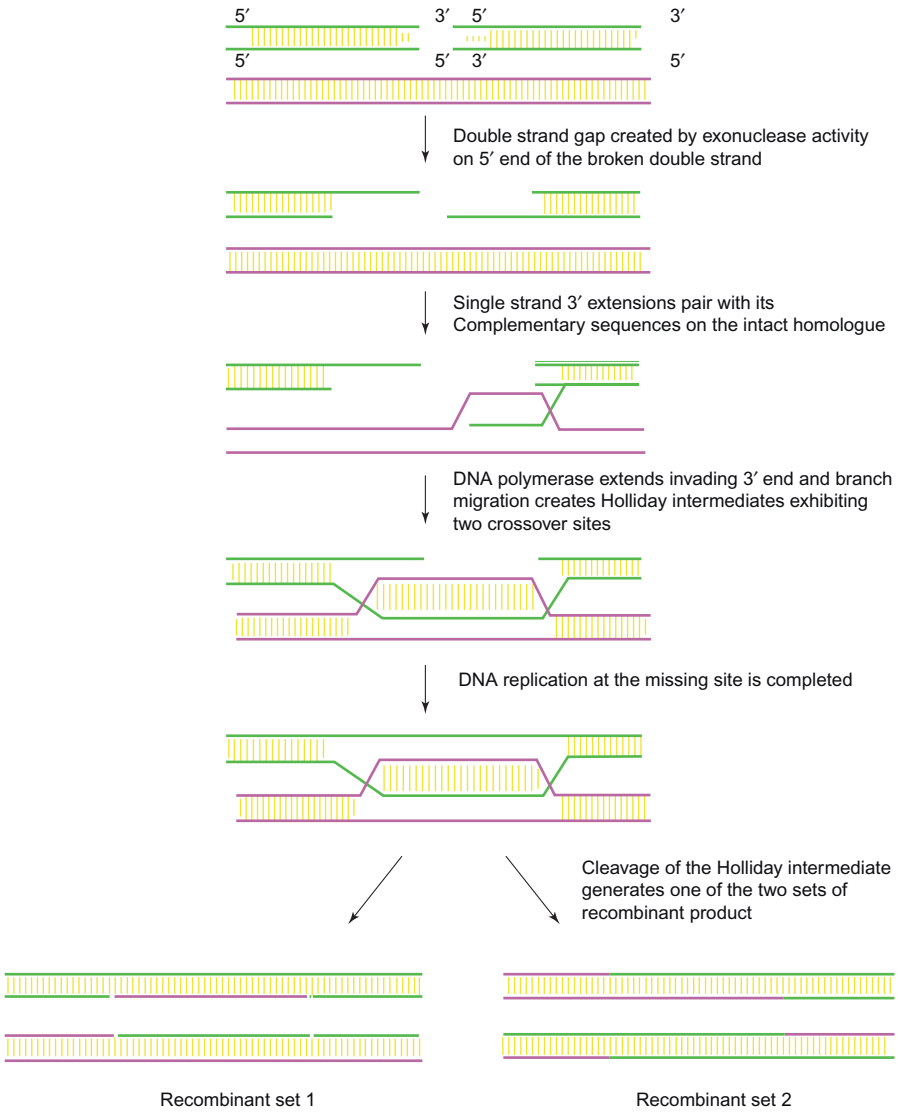

Figure 1.5 DNA recombination.

through a very precise mechanism, without any addition or deletion of nucleotides. This recombination is not site specific and can occur anywhere in the region of similarity, but regions of high probability do exist. Homologous recombination is also a DNA repair mechanism and is important for maintenance of genetic diversity. Homologous recombination uses a host of enzymes during the process.

Although homologous recombination is not site specific, recombination can be site specific and also involves recombination between nonhomologous regions. Enzyme

recombinase and a short specific DNA sequence are the key necessities apart from several other auxiliary proteins and enzymes. One recombinase enzyme binds at each of the two sites for recombination, which may be on the same or different DNA strands. Then the DNA strands are cleaved and in some cases rejoined to form Holliday intermediates at both the sites. In some systems, both DNA strands are cut and rejoined simultaneously without the formation of Holliday intermediates.

Transposition is another type of recombination process. This occurs mostly to allow for movement of mobile genetic elements or transposons. Mobile genetic elements have accumulated over ages by mutations and make up a sizeable part of the genome. These elements usually code for the proteins or enzymes that aid in their movement around the genome. The movement can be of a cut-and-paste type, where the whole piece of DNA is cut from one chromosome and inserted as such in the other. Movement can also be replicative, where a new copy of transposon is inserted in the target DNA and the original remains with the donor DNA. Similar to DNA-only transposons, as in the case of prokaryotes, eukaryotes have DNA-only transposons as well as retrotransposons. The latter elements use RNA as an intermediate molecule. The DNA mobile element is first transcribed and the RNA so formed is acted upon by reverse transcriptase to yield a DNA replica element that is inserted into the opposite DNA strand [32–40].

1.3 Transcription

The information is carried in genes as DNA, but to transduce this information into functional biomolecules requires the dual process of transcription and translation. The genetic information is copied in the form of RNA by transcription and translated to produce proteins. As with DNA replication, the DNA double strand must separate to allow for growth of the new polynucleotide chain by base pairing. The genes that carry the information to code for proteins are copied to RNA, and such RNA copies are termed mRNA molecules. The four-base language of DNA is converted to the four-base language of RNA by transcription. The DNA strand acts as a template strand for polynucleotide chain formation catalyzed by RNA polymerase (Fig. 1.6). The RNA polymerase identifies and associates with a region in the DNA called the promoter region. Transcription factors aid RNA polymerase in this promoter site recognition. The site of transcription initiation is called +1 and the direction in which it is transcribed is downstream and the opposite direction is upstream. RNA polymerase melts around 14 base pairs of DNA to create a transcription bubble. Transcription progresses by the joining of ribonucleotides by a phosphodiester bond formation at the 3′ end of the growing chain. About half of the bases are paired to ribonucleotides in the transcription bubble as the elongation complex comprising RNA polymerase, template DNA, and the growing RNA moves ahead. In the final stages, specific sequences in the template DNA direct RNA polymerase to terminate transcription, release the RNA transcribed, and dissociate from the template DNA strand, ready to transcribe again.

The transcription in prokaryotes differs from the eukaryotic process. The genes in prokaryotes have a common metabolic goal, and such a group of genes is called

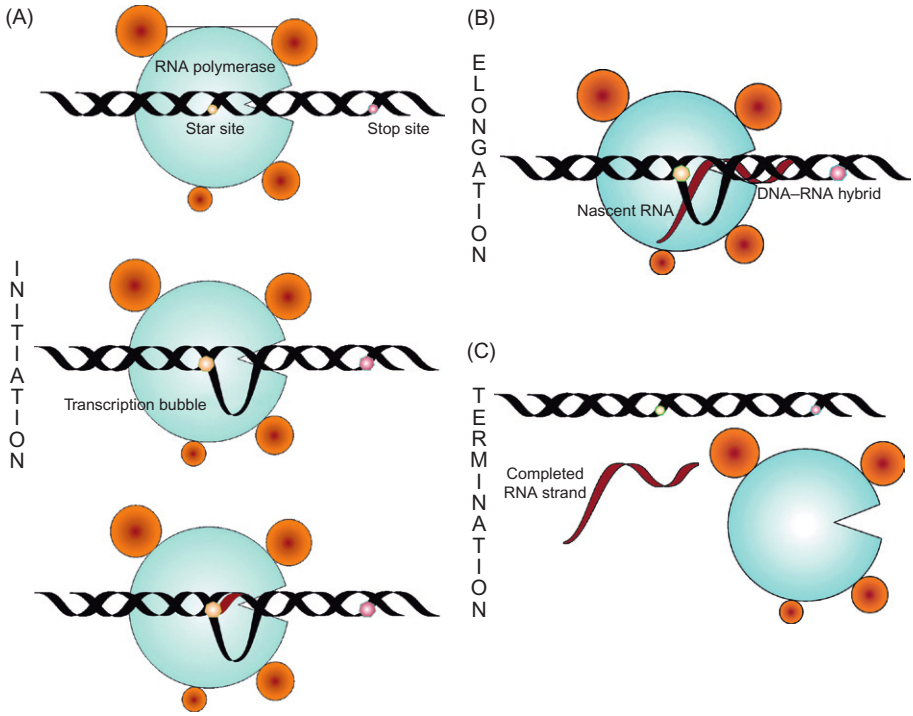

Figure 1.6 Steps in transcription: (A) initiation; (B) elongation; and (C) termination.

an operon. The transcription of operon produces polycistronic mRNA that codes for several functionally related proteins. Such clustering of genes is not seen in eukaryotes, and in fact the genes for related proteins are often present on different chromosomes and transcribed separately. But the eukaryotic genes have parts of coding sequences (exons) intervened by noncoding sequences (introns), requiring that the long mRNA strand transcribed be spliced or clipped to remove the noncoding parts and then ligated back together. Following the above modifications, the pre-mRNA transcript containing exons and introns is spliced to form functional mRNA. This is accomplished by spliceosome, wherein a group of proteins, along with small nuclear RNA, splice together the exons by forming loops of the introns and excising them.

In contrast, prokaryotes can exhibit transcription and translation simultaneously, as generally noncoding sequences are nonexistent in such cells. Moreover, there is no separate nuclear and cytoplasmic region to separate the two processes. However, in eukaryotes, mRNA primary transcripts must undergo processing in the nucleus before they are functional in the cytoplasm. As the primary RNA transcript dissociates from the RNA polymerase, a 5′ cap is formed by the addition of methyl guanylate at the 5′ end. This cap allows the easy export of RNA to cytoplasm and is the site for a specific protein association that is required to initiate translation. At the

3′ end, a chain of adenylic acid residues is added. The most important step in RNA processing is RNA splicing to remove the introns and ligate back the broken coding sequences. The fact that eukaryotic genes exhibit several introns allows for coding of several related proteins by alternative splicing and is the reason for the presence of isoforms of several proteins [41–43].

1.4 Translation

Transcription is usually followed by translation, the process by which the blueprint for protein production now copied into mRNA is used to form a functional protein by the addition of amino acids in the specified sequence. Protein synthesis is carried out in the cytoplasm in conjunction with all three types of RNA. The mRNA carries the information in a codon, the triplet-base language wherein each triplet of bases, or codon, can code for an amino acid or act as a stop codon. Of the existing 64 codons from combinations of four bases, 61 code for amino acids and 3 are stop codons. In general, with the exception of methionine and tryptophan, all other amino acids are coded by multiple codons.

The initiation of synthesis of the polypeptide chain in both prokaryotes and eukaryotes begins with the amino acid methionine. The start codon specifying it is AUG, and the sequence of codons from start to stop codon is called a reading frame. An mRNA molecule can be read in different reading frames but is generally read in a single reading frame, where a functional protein is formed before termination.

The conversion of the four-base language of nucleic acids into the 20-amino acid language of proteins requires an adaptor, and tRNA fits the role of an adaptor molecule aptly. To fulfill this requirement, it is necessary that an amino acid link to tRNA to form aminoacyl tRNA, and tRNA must have an anticodon to pair with mRNA. tRNAs are 70–80 nucleotides long and have a well-defined clover leaf-type three-dimensional structure. The four stems are stabilized by base pairing. The 3′ end of the unlooped stem is the acceptor stem. The base pairing rule demands the presence of 61 tRNAs, each specific for one codon coding for an amino acid, although cells may contain a lesser number of tRNA. This is because the first position of the anticodon on tRNA is the wobble position. It allows base pairing with more than one type of bases and accounts for the fewer number of tRNA molecules.

The attachment of tRNA with the appropriate amino acid is catalyzed by specific aminoacyl tRNA synthetase. The 20 different enzymes recognize one of the 20 amino acids and link them to their appropriate tRNA at the 3′ terminus. This reaction requires energy, and the aminoacyl tRNA is said to be activated. The fidelity of the attachment is ensured by the proofreading activity of the synthetase enzymes.

The next step in protein synthesis involves the ribosomes. Ribosomes are rich in RNA, containing about 60% RNA. They are composed of three to four different rRNA molecules and around 80 proteins, and form a large subunit and a small subunit. The ribosomal subunits and rRNA are differentiated by their sedimentation

rate, which is quantified in terms of Svedberg (s) units. The prokaryotic RNA is 70s, whereas generally the eukaryotic RNA is 80s.

An initiation complex is formed by the assembly of one of the ribosomal subunits, mRNA, methionine-charged tRNA, and some associated proteins. The initiation complex moves along the mRNA to identify the start codon. Helicase activity of the associated proteins aids in the unwinding of the mRNA. The protein synthesis always starts at the AUG, the start codon for methionine. This ensures that mRNA is read in the correct reading frame. The tRNA that charges the first methionine is unique in both prokaryotes and eukaryotes and differs from the tRNA that brings methionine at positions other than the starting one in the growing polypeptide chain. Hence, only initiator methionine tRNA ($tRNA^{i}_{met}$) is capable of binding at the P site on the ribosome, whereas others bind at the A site. As $tRNA^{i}_{met}$ recognizes the start codon, the movement of the complex is halted, and the larger subunit of the ribosome also links to form the 80s ribosome.

Several proteins called elongation factors now come into play to carry out the process of polypeptide chain elongation. As discussed earlier, the $tRNA^{i}_{met}$ is at the P site, which is the site of polypeptide elongation. The next aminoacyl tRNA is brought at the acceptor (A) site, with the appropriate base pairing between the codon on mRNA and anticodon on tRNA. GTP hydrolysis and some conformational changes in the ribosome cause tight binding of charged tRNA at the A site and bring the amino acid close to the tRNA at the P site. The α amino group of the incoming amino acid reacts with the carboxylic group of the charged methionine to form a peptide bond. This peptidyl transferase activity is catalyzed by larger rRNA. Subsequent to bond formation, ribosome moves along or translocates by one codon. After this translocation, the $tRNA^{i}_{met}$ without its methionine is positioned at the exit (E) site of ribosome and the second tRNA with a dipeptide attached is now at the P site, leaving the A site empty for the incoming aminoacyl tRNA. The peptide chain grows in similar fashion (Fig. 1.7) until it comes across one of the stop codons. Thereafter, a battery of protein factors bring about the release of the completed polypeptide chain. The polypeptide finally assumes its native three-dimensional conformation on release [44–50].

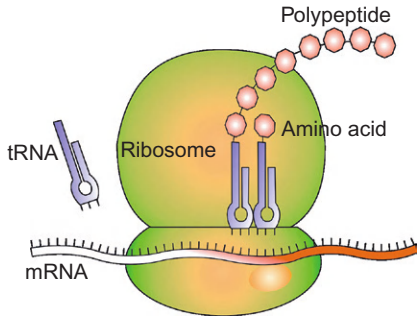

Figure 1.7 Translation of mRNA.

1.4.1 Protein Translocation

As evident from earlier discussions, proteins and enzymes are essential to every metabolic, synthetic, structural, or regulatory functioning in the cell. They are synthesized on the ribosomes but are present at all nooks and corners of the cell. This is possible because of the existence of a system for protein translocation. Protein targeting to various cellular destinations is made possible by the presence of a short sequence of amino acids called a signal sequence. The signal sequence may or may not be cleaved, and if cleaved, it is cleaved either during transport or after reaching the target organelles/location. The proteins that need to be translocated to ER, mitochondria, or chloroplasts have an amino terminus-attached signal sequence. The signal sequences are 13–36 amino acids long, with around half of the amino acids in the chain being hydrophobic. The protein synthesis initiates on the free cytosolic ribosome, and the signal sequence is synthesized initially. The large signal recognition particle (SRP) binds with the ribosome and directs the ribosome bound to mRNA to the SRP receptor on the ER. The length of the polypeptide is around 70 amino acids at this stage. The polypeptides interact with the peptide translocation complex at ER and release the SRP. The polypeptide synthesis resumes again, and the growing chain is directed inside the ER lumen until complete synthesis. The ribosome becomes dissociated, the signal sequence attached to the polypeptide is cleaved, and the nascent protein achieves its three-dimensional conformation. The protein so formed may be further modified by glycosylation and post-translational modifications and is packaged into transport vesicles and moved to the Golgi complex, where it undergoes further sorting. Similarly, the proteins intended for mitochondria are also targeted using specific amino acid sequences. But the process differs because the complete synthesis of protein takes place on the ribosomal assembly, and once released, they become associated with chaperone proteins that deliver the protein to the mitochondrial membrane. The proteins are then internalized using specific mechanisms not yet fully understood, and subsequently the signal sequences are removed. The nuclear targeting of proteins follows yet another process in which the nuclear localization sequence (NLS) targets the protein to nucleus. The peculiarity of NLS is that it may be present anywhere in the polypeptide sequence and is not cleaved at destination or post-polypeptide synthesis. The process is mediated by proteins called importins and a GTPase known as Ran. Similar to eukaryotes, bacteria also target their proteins to either the cell membrane or extracellular matrix through signal sequences at the amino end. In eukaryotes, apart from intracellular localization of synthesized proteins sometimes the proteins are imported from extracellular media. These proteins bind to special areas called coated pits. The coated pits are the areas where there is a high presence of endocytic receptors. The cytosolic side of the pits is lined with clathrin protein, which helps the membrane to invaginate, and an endocytic vesicle containing proteins is formed. The clathrin then becomes dissociated from the vesicle, and the endocytic vesicle fuses with endosome to ultimately release the imported protein [51].

1.5 Transcriptional Control of Gene Expression

Gene expression is regulated by regulatory factors at various levels so that a differential synthesis of protein is observed. This is important if proteins are being produced in anatomically and physiologically different cells or in different development stages or in response to various external stimuli. The regulation of gene expression in prokaryotes is with the basic aim to help the organism meet the challenges of survival in accordance to the external conditions and to optimize its growth. Gene expression regulation in eukaryotes has several different objectives to meet. These include differentiation during development, immune response development, and so on. In spite of the difference in purpose of regulation, there are some basic features common to both the gene regulatory sequences and control elements associated with genes. Gene regulatory proteins bind to these elements to enhance or suppress gene expression. Activator proteins promote binding of RNA polymerase to the promoter region, whereas repressor proteins inhibit this association. These protein binding sites are often located close to or at distant locations from the site of initiation. There are three types of RNA polymerases in eukaryotes. RNA polymerase I is involved in the synthesis of pre-rRNA, RNA polymerase II synthesizes mRNA, and RNA polymerase III synthesizes tRNA and some other small RNA molecules. Promoter regions are important in regulating binding of RNA polymerase II to influence the site of initiation and rate of transcription. Three types of important promoter regions have been identified in eukaryotes. These include TATA boxes, initiators, and CpG islands. Further, additional cell-specific elements like enhancers and promoter proximal elements located upstream or downstream play a role in gene expression regulation. Corresponding elements called repressors are functionally opposite to activators. Transcription activators or repressors have a unique single-DNA binding site and one or more activating or repressing regions in their three-dimensional conformation. There are some common structural conformations in the DNA-binding domains of transcription factors like C_2H_2 zinc finger, homeodomain, basic zipper, and basic loop–helix–loop. These structures have alpha helices that interact with the major groove in the DNA. Similarly, activation and repression domains have a diverse three-dimensional conformation and bind with coactivators and repressors to eventually modulate the gene expression. Thus, a very complex process involving several proteins and genetic elements is important to the regulation of gene expression [52–56].

1.6 Cell Communication or Biosignaling

For the cell to maintain homeostasis, it must respond to changes in its environment. This prepares the cell for a defensive, proactive reaction or helps it in acquiring nutrition and so on. This process requires signal transduction, which is the conversion of signals like pH, osmotic strength, oxygen, light, and so on, into a chemical change. The basis of cellular communication, or biosignaling, is that a specific signal is identified by the receptor proteins in the target cell, and in response to the signal an appropriate response is elicited. Signal transmission may be widespread, as

in the case of hormones secreted into the blood or plant sap by endocrine cells. The other method is paracrine transmission, which is the local distribution of the signal in the extracellular matrix. Signal transmission through neurons is still another type of long-distance dissipation of information; in this case, however, the signal is delivered quickly and to specific cells.

There are a few fundamental characteristics seen during cell communications, like sensitivity of signal transducers. The cell has number of receptors, but they interact with specific signals only. Secondly, a single signal itself can produce a wide range of effects in the same cell. Receptors recognize the signal and relay it to the site of action through a signaling cascade. During this course, the signal undergoes amplification through secondary messenger systems and so forth and simultaneously modulation too. Modulation is observed because, at any time, in each cell a number of different receptors are being activated to produce varying responses. It is important for a cell to integrate these signals to produce a unified reaction in the best interest of the cell. The presence of a persistent signal desensitizes the receptor, and the receptor returns to its normal state slowly.

There are several types of receptors involved in the signaling process. The major ones include ion-gated channels or ion channel-linked receptors, where a flow of ions produces an ultimate response; G protein-coupled receptors that activate G protein, starting a cascade of events; and receptor enzymes that act as enzymes on activation or activate further enzymes to produce the response and steroid receptors [57–60].

1.6.1 Ion-Linked Receptors

These are the major signal transducers at the synaptic junctions, neuromuscular junctions, and myocytes. These signal transducers are working comparatively in a more simple way and act directly to produce the ultimate action. The signals for such receptors are neurotransmitters that bring about a conformational change in the receptor. This opens/closes the specific ion channels for Na^+, K^+, Ca^{2+}, and Cl^- (Fig. 1.8). The driving force for the movement of ions is electrochemical gradient. The flow of ions in and out of the cell causes a redistribution of charges, and hence a change in the membrane potential of the plasma membrane is observed. Nicotinic acetylcholine receptor is a very classical example of a ligand-gated ion channel. Calcium ions often act as second messengers in these cases. The calcium ion channels are opened by depolarization waves in neurons and increase calcium ions concentration intracellular. This rise initiates the release of neurotransmitter acetylcholine at the synapse, and the released acetylcholine acts at the next neuron and brings about a similar chain of events to pass on the signal to the target cell. A neuron responds to several neurotransmitters, and each neurotransmitter opens up a particular ion channel. The resultant depolarization or hyperpolarization is the integrated input from the several signals received [61,62].

1.6.2 G protein-Coupled Receptors

These are the largest family of cell surface receptors, with around 1000 G protein-coupled receptors having already been identified in humans. These receptors mediate

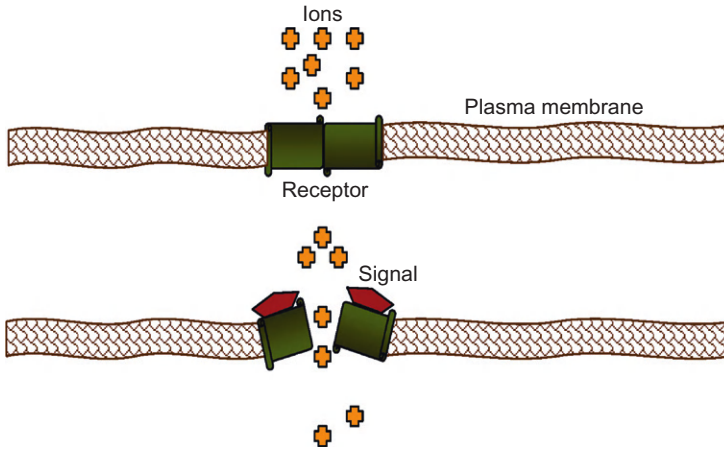

Figure 1.8 Signal transduction via ion channel-linked receptor.

responses to a wide variety of signals and produce diverse response like the sensations of taste, smell, perception of light, and so on. Although there are several subclasses of this large family of receptors, they have a certain structural similarity. All these receptors are made up of a protein that traverses the membrane seven times, that is, has seven transmembrane passes.

In response to an extracellular signal, the receptor undergoes a conformational change to activate a membrane-attached trimeric G protein located on the cytosolic side of the receptor. The G proteins are also a large family, and each responds to a specific set of signals. The structure of G proteins is composed of three subunits—α, β, and γ. The α-subunit has a guanosine diphosphate (GDP) bound to it in the unstimulated state. The G protein activation causes the α unit to exchange GDP for GTP; simultaneously, it detaches from the β–γ complex. These two separate parts of the G protein move along the plasma membrane; they can bind with target proteins to relay the signal, and their intensity of binding decides the strength and length of the signal relay. The α unit slowly hydrolyzes back its GTP to GDP and returns back to relink with the other two units, stopping the signal transduction. The target proteins for G proteins may be enzymes, ion-gated channels, and so forth.

The interactions through an ion channel bring immediate change in the response of the cell. For example, the slowing down of heart muscles in response to the release of acetylcholine is mediated through a G protein-coupled receptor. The acetylcholine activates a G protein, causing it to break apart into parts, and the β–γ complex opens the K^+ channel, changing its electrical properties (Fig. 1.9A).

The G protein-coupled receptors also interact with other enzymes; these often include adenylyl cyclase and phospholipase C. The adenylyl cyclase pathway uses cyclic adenosinemonophosphate (cyclic AMP) as the intracellular-signaling molecule, and the phospholipase C pathway involves inositol trisphosphate and diacylglycerol. These small molecules are termed the second messengers. The G protein's detached α unit affects the adenylyl cyclase activity and increases the intracellular levels of

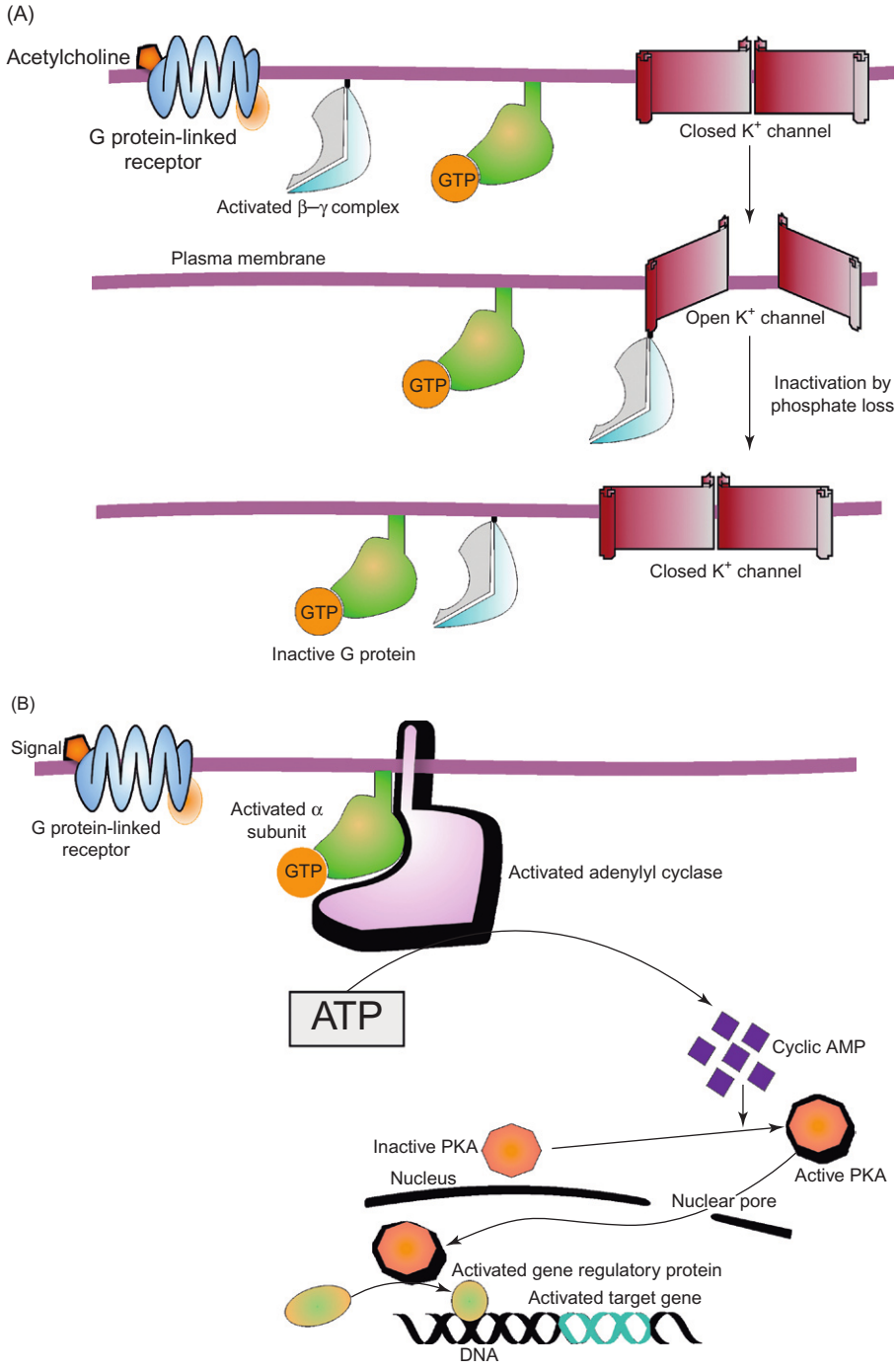

Figure 1.9 Signal transduction via G protein-coupled receptor.

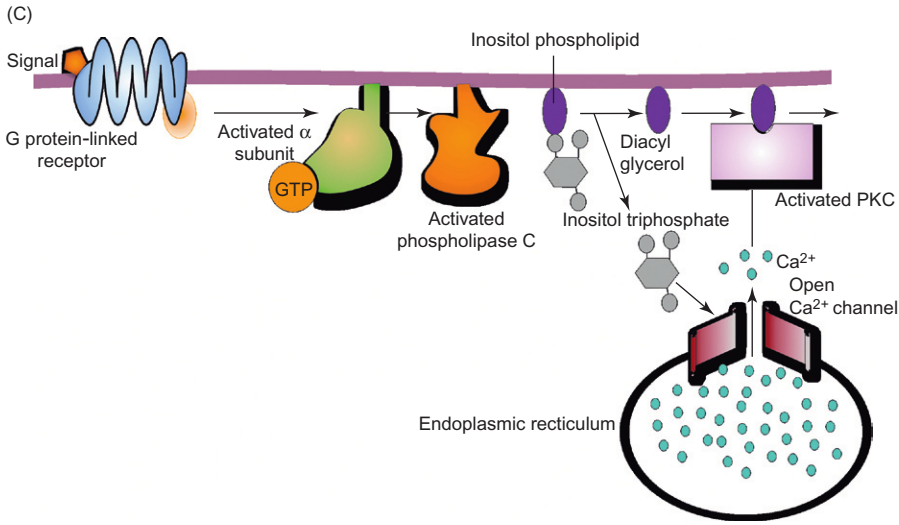

Figure 1.9 (*Continued*).

cyclic AMP tremendously by synthesis from ATP. Cyclic AMP in turn produces its effects by activation of a protein called protein kinase A. The protein kinase is activated by binding with cyclic AMP, and the activated kinase then phosphorylates specific proteins at the serine or threonine residues. This alters their activity and in different cells a wide variety of responses can be seen (Fig. 1.9B) [63–66].

Other G protein-coupled receptors mediate their effects via phospholipase C present in the membrane. The activated phospholipase C produces two messengers, inositol 1,4,5 trisphosphate and diacylglycerol, by hydrolysis of membrane inositol phospholipid. The trisphosphate moiety is hydrophilic and diffuse through the cytosol, to release calcium ions from ER, whereas the diacylglycerol remains in the plasma membrane only. The increased cytosolic calcium level, along with diacylglycerol, activates protein kinase C, which subsequently phosphorylates and activates several other intracellular proteins (Fig. 1.9C).

The several signal-transducing mechanisms involve calcium at one stage or another. Calcium is a very important second messenger, and its levels are grossly different across the membranes. These gradients are maintained by active membrane pumps. The effects of calcium are mediated through calcium-binding proteins; a frequently involved example is calmodulin. The calcium binds with calmodulin, and in turn this activated complex activates calcium/calmodulin-dependent protein kinase, which subsequently phosphorylates other proteins [67–71].

1.6.3 Enzyme-Linked Receptors

Enzyme-linked receptors are also transmembrane proteins, and the extracellular ligands bind to them on the extracellular side. They comprise a very large family, and a major subclass includes receptor tyrosine kinases that phosphorylate the tyrosine

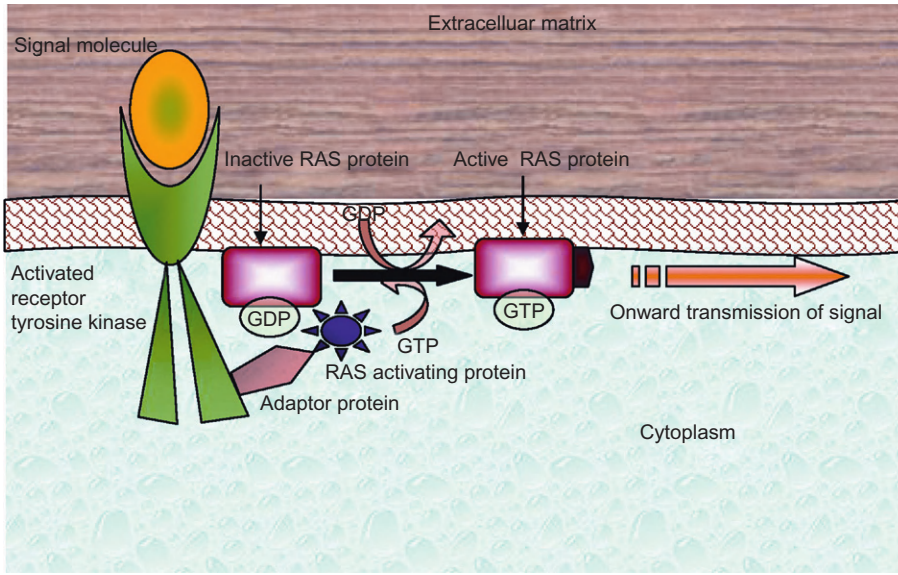

Figure 1.10 Signal transduction via enzyme-linked receptors.

residue on the cytosolic side of these proteins. The receptors are usually helical in structure, and binding with a triggering signal leads to dimer formation by the linkage of two receptors. The tail regions of the two proteins in the dimer interact with each other and each phosphorylates the other.

The activated receptors attract several intracellular-signaling proteins like protein kinase B, phosphatidylinositol 3-kinase, and others. But the receptor tyrosine kinases have a separate route of action. Receptor tyrosine kinases activate Ras, a small protein bound by a lipid tail to the cytoplasmic side of the plasma membrane. The Ras group of proteins are single-unit monomeric GTP-binding proteins and are similar to the α subunit of trimeric GTP-binding proteins. The activated Ras phosphorylates kinase enzymes, which in turn phosphorylate other kinases, passing the signal from plasma membrane to the nuclear region. The whole relay is called a MAP kinase cascade because mitogen-activated protein kinase (MAP kinase) is the last kinase in the series. MAP kinase brings about phosphorlyation of serine and threonine present on transcription-controlling proteins. The phosphorylation may activate or suppress activity of the gene (Fig. 1.10). The result may be observed as proliferation, differentiation, and so on.

Ras is a very important protein whose role has been implicated in cancers. In response to the activation of Ras, the cells respond by growth and proliferation. If Ras remains activated, then an uncontrolled growth and proliferation are observed. Some mutation in the Ras gene or other proteins in the pathway often leads to an uncontrolled outburst of growth, causing cancer [72,73].

The information dissipation through signaling cascades is one way of communication; there are hormones and local mediators like cytokines that have a differing

route of signal transduction. Their receptors do not possess any inherent enzymatic activity, but they are associated with cytoplasmic tyrosine kinases called JAKs. These are present on the plasma membrane, and once a cytokine binds to its receptor, it activates the JAKs that phosphorylate gene regulatory proteins STATs (Signal Transducer and Activator of Transcription). Activated STATs migrate to the nuclear region and control gene expression by binding to the responsive elements, resulting in modified gene expression. Another way of direct communication is adopted by serine/threonine kinases, which are similar to receptor tyrosine kinase and directly activate proteins termed SMADs (Sma and Mad related family protein(s)) by phosphorylation. SMADs are proteins capable of controlling gene expression, and this pathway is often used by the TGF-β superfamily [74].

1.6.4 Hormone Receptors

Several steroid hormones and thyroid hormones transduce the signals by a significantly different mechanism. These hormones often traverse the blood stream associated with specific carrier proteins to reach their site of action. Being lipid soluble, they pass through the plasma membrane to reach and bind with specific receptor proteins in the nuclear region. The altered conformation of receptor proteins, due to hormone binding, allows them to interact with hormone-responsive elements on the DNA. The interaction causes a changed pattern of gene expression, either enhancing or suppressing it. Apart from this classical mechanism, some other pathways involving cyclic AMP and MAP kinases are responsible for other effects exhibited by these hormones [75,76].

1.7 Cell–Cell Adhesion and Extracellular Matrix

Cells in a multicellular organism are organized into specific tissues, and the individual cells in the tissues are associated with each other as well as the extracellular matrix. The linkages between the individual cells, called *cell junctions*, are classified into three major categories based on their physiological role: occluding junctions, anchoring junctions, and communicating junctions. Occluding junctions include tight junctions that serve to create an impermeable or semipermeable barrier between the adjoining epithelial cells. They are barriers to the transportation of material and control the movement of membrane transport proteins between the apical and basal layers of epithelia. The second category, anchoring junctions, connect the cytoskeletal network of the cell to the adjoining cells and/or the extracellular matrix, helping the tissue to survive any inflicted mechanical stress. These types of junctions are quite abundant and are further classified as *adherens junctions*, *desmosomes*, and *hemidesmosomes*. The adherens junctions connect the intracellular actin filaments to other cells or extracellular matrix through cadherin and integrin proteins, respectively. Similarly, desmosomes connect the intermediate filaments from one cell to the next, and hemidesmosomes are the connection site for the intermediate filaments to the basal lamina. Gap junctions constitute the communicating type of cell junctions.

Their presence allows for selective transport between the cells of material having molecular weight less than 1000 Da. Molecularly a channel-like structure made up of proteins, they permit movement of inorganic ions and are crucial in regulating the electrical signals at neural and neuromuscular junctions [77].

1.8 Cell Cycle

In general, cell cycle is defined as a sequence of events where the cellular and nuclear contents are duplicated and divided. So we can say that cell cycle ultimately brings about cell reproduction. Cell cycle differs in different organisms, in various cells of multicellular organisms, and even at different stages of life. But certain fundamental processes remain the same: the genetic material is replicated and passed onto progeny following division. This ensures passing the entire genome to each daughter cell. The time duration of cell cycle also exhibits considerable variance. The eukaryotic cell cycle is divided into four major phases: the M phase and the other three phases, collectively termed the interphase.

The M phase is the most spectacular phase—a complex phase of cell cycle in which the replicated genome splits during mitosis, in other words, the nuclear contents are segregated into the daughter cells. Mitosis is further subdivided into various phases discussed later in this chapter. This is quickly followed by cytokinesis, the division of cytoplasmic contents into the progeny. The cytoplasmic contents are already replicated during biosynthesis in interphase, in preparation of the impending cell division. The transcription and protein synthesis is already in continuance during interphase, and hence the cytoplasmic contents are easily divided into the daughter cells. A contractile ring, made of nonmuscle myosin II and actin filaments, assembles equatorially (in the middle of the cell) at the cell cortex, constricting the cell membrane, to form a cleavage furrow eventually dividing the cell into two.

The time phase between the two M phases is the interphase, which can be further segregated into three phases: G1, S, and G2 (Fig. 1.11). The cell division must go hand in hand with cell growth so as to maintain cell size. The G1 phase is the gap phase between the M phase and the S phase. During the G1 phase, the cell responds to its external and internal stimuli and decides to commit to the S phase, or extends the gap phase to prepare for the upcoming tumultuous S phase. Sometimes the cells may switch to a G0 phase of prolonged nondividing state. The cell responds to the external stimuli to enter this nonproliferating phase.

The S phase is called the *synthetic phase* and involves the replication of the cell's nuclear contents. The replication machinery comes into action and, as described earlier, the DNA replicates to double the entire genome.

The G2 phase is the gap phase between the S phase and the next M phase. It provides additional time to grow and duplicate cytoplasmic contents, proteins, and so on. The gap phases are also important regulatory checkpoints before the cell commits itself to upcoming phases in the cell cycle. This is crucial because once committed, it is difficult for the cell to revert back or halt the further processes [78–80].