Introduction

1

Summary

- Molecular medicine is a highly dynamic field of life science research that uses interdisciplinary approaches to understand normal and pathological cellular processes at the molecular level. The findings of basic research have entered clinical practice, as new diagnostic assays and novel therapeutic strategies focus not only on the symptoms but also on the causes of disease.
- The development of drugs is a long-term and expensive process that starts with basic and preclinical research. A candidate drug must then successfully pass through three types of clinical trial in humans before a novel agent can be approved for therapeutic purposes.
- The eukaryotic cell is compartmentalized into several cellular organelles by intracellular membranes. The nucleus harbors the genetic material, mitochondria are the cellular power plants, and the endoplasmic reticulum and the Golgi apparatus are responsible for the glycosylation and sorting of proteins.
- Cells follow a tightly regulated cycle of four phases. These include the two gap phases G_1 and G_2, the S phase in which new DNA is synthesized, and mitosis, during which the cell divides.
- Apoptosis is the process of programmed cell death, which is important as a normal physiological mechanism and for protection against infections and cancer. Apoptosis can be triggered by extrinsic or intrinsic signals.
- Genomic DNA is amplified by DNA polymerases in a process known as replication. The synthesis occurs in a semiconservative and semidiscontinuous way.
- Expression of genes requires two steps. In the first step, the DNA is transcribed into RNA. Most primary transcripts are posttranscriptionally processed. For mRNAs, this step includes the addition of a cap at the 5′ end and of a poly(A) tail at the 3′ end. Introns are spliced out to link the exons together. Several bases are modified in various types of RNAs. The second step in gene expression is the translation of the genetic information into proteins. This process is carried out by ribosomes. Posttranslational modifications of proteins include activation by proteolytic cleavage and covalent modification of amino

Contents List

Molecular Medicine: An Introduction, First Edition. Jens Kurreck and Cy Aaron Stein.

acid side chains. This can occur, for example, by glycosylation or reversible phosphorylation.
- Sophisticated communication between cells is essential for the functioning of a multicellular organism. Neurons transmit signals at synapses. Hormones are molecules that induce physiological responses over a long distance or in adjacent cells. The extracellular signals are transmitted into the cell by cell surface receptors and induce a signaling cascade that leads to a biological response.
- The immune system protects an organism against (infectious) disease. The innate immune response recognizes general patterns of pathogens, while the adaptive immune system is directed against specific targets. The adaptive immune system involves a cellular immune response (T cells) and a humoral immune response (B cells that produce antibodies).

1.1 The Basics of Molecular Medicine

1.1.1 Topics of Molecular Medicine

Molecular medicine is a discipline dedicated to understanding normal and pathological cellular processes at the molecular level. This approach requires the use of many physical, chemical, biological, biochemical, and medical techniques (some of which are introduced in Chapter 2) to understand fundamental molecular mechanisms and how they go awry in disease. Molecular medicine combines classical disciplines such as cell and molecular biology, biochemistry, and medicine. Knowledge is often acquired via interdisciplinary investigation and can be used to develop new forms of molecular diagnosis and therapeutic intervention.

Molecular medicine can be divided into a basic research and an applied clinical discipline. The basic research component investigates molecular and genetic mechanisms of cellular function and identifies pathological processes. In many cases, this addresses a specific question with a hypothesis-driven approach, and can lead to large-scale investigations of whole genomes and proteomes (Chapter 7). The discipline known as translational research then tries to apply the findings from basic science to the clinic, where it may provide new forms of diagnosis and therapy.

A report published by Linus Pauling in 1949 laid the basis for the establishment of the field of molecular medicine. In his seminal paper, he showed that hemoglobin from patients suffering from sickle cell anemia had a different electrical charge than hemoglobin from healthy individuals. This study demonstrated that a disease could be traced to an alteration in the molecular structure of a protein. This novel perspective opened the possibility of establishing novel forms of diagnosis and therapy at the molecular level. Sickle cell anemia is not the only case in which a detailed understanding of the molecular etiology of the disease (e.g., of inherited genetic disorders, Chapter 3) has led to new diagnostic options (Chapter 8), although with only a modestly improved therapeutic outcome.

The field of oncology also illustrates the paradigm shift caused by a molecular perspective. While cancer treatment is still largely based on removal of the tumor by surgery (followed by chemotherapy and/or radiation therapy), molecular oncology (Chapter 4) tries to elucidate those pathways that lead to cellular transformation. This knowledge helps to produce a comprehensive molecular diagnosis of the disease basis in a single patient so that the treatment can be adjusted accordingly, an approach that has come to be known as "personalized medicine." Many modern anticancer drugs block specific pathways that lead to uncontrolled cellular proliferation. Similarly, elucidation of the life cycles of pathogens has helped develop new drugs for the treatment of infectious diseases (Chapters 5 and 6). For example, advancements in virus biology have led to the identification of novel targets for antiviral agents.

Most drugs belong to the class of small molecular compounds. To achieve oral bioavailability and to promote rapid diffusion across cell membranes and intracellular trafficking to their sites of action, the majority of (oral) drugs have molecular weights below 550 Da (although some antibacterial agents fall in the 700–900 Da range). A prominent example is acetylsalicylic acid (trade name Aspirin, Figure 1.1a), a drug mainly used as an analgesic.

Molecular medicine has broadened the spectrum of entities used as drugs. New medications are now often based on large molecules of biological origin (known as "biologics"). These include, for example, recombinant proteins such as monoclonal antibodies (Figure 1.1b) (Chapter 10), short pieces of DNA or RNA (Chapters 13 and 14), entire genes that can be delivered by viral vectors (Chapter 11), or even complete cells (Chapter 12). Pharmacogenetic investigations aim at discovering why the efficacy and toxic side effects of a drug at a given dose vary between individuals (Chapter 9). However, molecular medicine not only develops new diagnostic and therapeutic approaches but also poses heretofore unknown ethical issues, some of which will be introduced in Chapter 15.

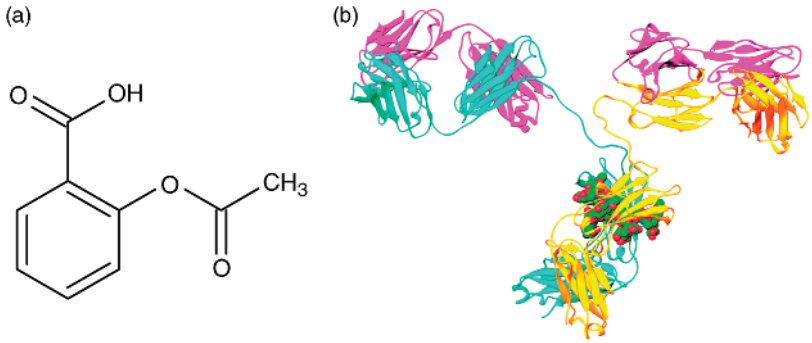

Fig. 1.1 Small molecular drugs and biologics. (a) The chemical structures of acetylsalicylic acid (Aspirin) and (b) the crystal structure of an antibody, shown for comparison. The two structures are not drawn to scale. (Part (b) adapted from Ref. [1] with kind permission from John Wiley & Sons, Inc.)

1.1.2 Stages of Drug Development

The development of a new drug is a time-consuming and expensive process (Figure 1.2) that may take 12–15 years (and in some cases even longer). The cost calculus of developing a new drug is complex and controversial, but the average cost to bring a new molecular entity (NME) to the market has been estimated to be as high as $1.8 billion. Drug development usually starts with the identification of a new target, which, for example, may be a proliferative factor that causes tumor growth. The next step in the process is to characterize the target, its location (extracellular, membrane-bound, cytosolic, and nuclear), and its function. Confirmation that the potential drug target fulfills the expected function is known as *target qualification* or *validation*. One way to identify a new compound is to perform what is known as a high-throughput screen (HTS). This approach allows testing a large number (up to millions) of compounds to identify an active molecule that modulates a particular target (e.g., inhibits a proliferative factor). In almost all cases, the primary hit must be optimized by chemical modification to obtain higher binding affinities, better solubility, and so on. The efficacy and toxicology of the substance are then investigated in *in vitro* studies and animal experiments. The process of drug development may deviate substantially from this path depending on the type of drug being developed. Biologics, for example, are usually not obtained by HTS. Any substance will only be tested in humans after having passed extensive toxicological examination.

Clinical research is usually divided into three main phases. However, these phases are sometimes preceded by an exploratory trial (frequently called as phase 0) in a small number of subjects with a very small, subtherapeutic dose designed to gather data on the agent's basic properties in humans. This trial does not produce data about safety or efficacy. The actual clinical research starts with a phase I trial, usually carried out with a small number (20–100) of subjects. The main purpose of a phase I trial is to assess the safety of the drug. The

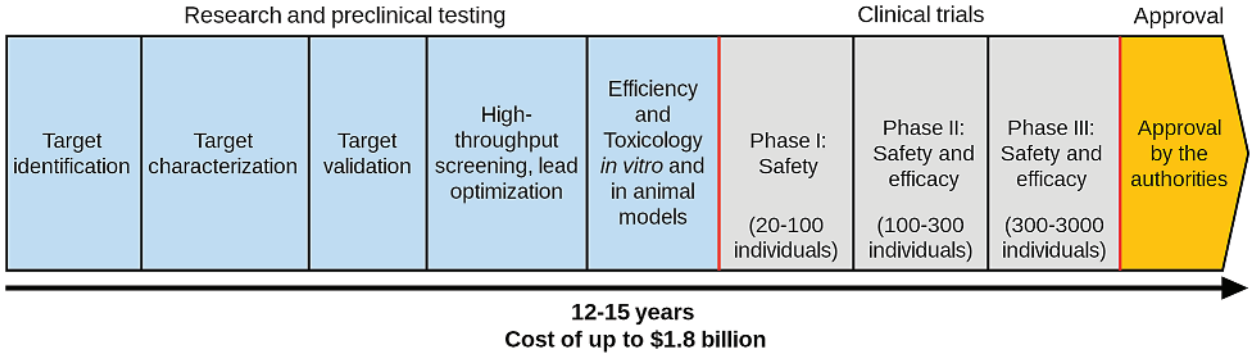

Fig. 1.2 Stages of drug development. The preclinical stages comprise the identification, characterization, and validation of a target and the identification and optimization of a compound, in addition to toxicological evaluation. The drug then undergoes three main phases of clinical testing before it is approved by the regulatory authorities.

trial is frequently designed to include a dose escalation to determine the optimal dose and the dose at which unacceptable toxicity supervenes. Although phase I trials are often carried out with healthy volunteers, under some circumstances, ill patients are enrolled. This is done most often with cancer patients, as the drugs to be evaluated are likely to make healthy individuals ill or may carry a significant risk of long-term toxicity.

Phase II trials are carried out with a larger number (100–300) of individuals. The central aim of the phase II study is to evaluate the efficacy of the drug. The trial is, therefore, usually performed in sick patients. Phase II studies are sometimes divided into phase IIA and phase IIB. While a phase IIA trial assesses the optimal dosing of the drug, a phase IIB trial is designed to study the efficacy of the drug at the prescribed dose. Another important goal of a phase II trial is to assess drug safety in a larger group of individuals.

Phase III trials are designed to assess the effectiveness and safety of a new drug in clinical practice in a large patient group (300–3000 or more individuals). These studies are carried out in randomized, controlled, multicenter trials. Phase III trials are usually designed as double blind studies, that is, the patients are randomly assigned to an experimental and a control group (in some trials, the control group may either receive a placebo or standard of care treatment). Neither the patients nor the physicians monitoring the outcome know which patient is receiving which treatment. Phase III trials aim at assessing the efficacy of a drug in comparison to placebo or the current standard of care treatment.

If drug safety and efficacy have been demonstrated in multiple phase III trial(s), approval for marketing can be applied for from the appropriate regulatory agency such as the Food and Drug Administration (FDA) in the United States or the European Medicines Agency (EMA) in the European Union. These agencies may request postmarketing monitoring, which is sometimes referred to as a phase IV trial. A phase IV trial involves safety surveillance after the drug has received permission to be marketed. In principle, it is designed to detect rare or long-term adverse effects in a much larger patient population and over a longer time period than was possible during the earlier phase of clinical trials. A phase IV trial may also identify interactions with other marketed drugs. Even after marketing, if harmful effects are discovered, any drug may be withdrawn at any time or its use restricted only to certain conditions.

The term phase V is increasingly used to describe studies that determine whether the therapeutic effect of a new drug is realized in day-to-day clinical practice. Community-based research is employed to survey whether the effects under typical (and somewhat variable) clinical contexts are similar to those that were found in the controlled efficacy studies. A phase V trial may also analyze the cost-benefit ratio of a drug or therapeutic intervention.

1.2 The Human Cell

Despite the extreme complexity of living organisms and the myriad number of functions that each constituent organ must carry on, only a surprisingly limited set of molecules are commonly employed. Typical biomolecules found in living organisms include nucleic acids, proteins, polysaccharides, and lipids. These macromolecules are composed of a relatively limited number of monomeric building blocks such as DNA and RNA nucleotides, amino acids, monosaccharides, and fatty acids. In addition, inorganic ions, organic acids, and a variety of metabolites are also important constituents of cells. While the basic features of biomolecules are extensively covered in the general textbooks of biochemistry listed at the end of this chapter, the cell as the basic functional unit of an organism and the major intracellular processes relevant to human physiology and pathology will be outlined here.

The human body consists of approximately 100 trillion (10^{14}) cells. Although all cells of a given organism carry the same genome, these cells have different functions and are highly specialized. A typical eukaryotic (animal) cell is illustrated in Figure 1.3. The most prominent characteristic that distinguishes eukaryotic cells from prokaryotes is its compartmentalization. The main membrane-bound organelles of animal cells are the nucleus, the endoplasmic reticulum (ER), the Golgi apparatus, the mitochondrion, the lysosome, and the peroxisome. Each of these organelles contains a specific set of proteins that fulfill a specific function. These organelles are embedded in a gelatinous fluid called the cytosol.

1.2.1 Organelles

1.2.1.1 The Nucleus

The central organelle of a eukaryotic cell is the nucleus, which contains the genetic material. The nucleus is surrounded by the nuclear envelope consisting of two membranes: the inner and the outer nuclear membranes (Figure 1.4a). The outer nuclear membrane is continuous with the rough endoplasmic reticulum (RER). Proteins referred to as nucleoporins form aqueous channels, called nuclear pores, through the envelope. These pores allow small water-soluble molecules to pass into and out

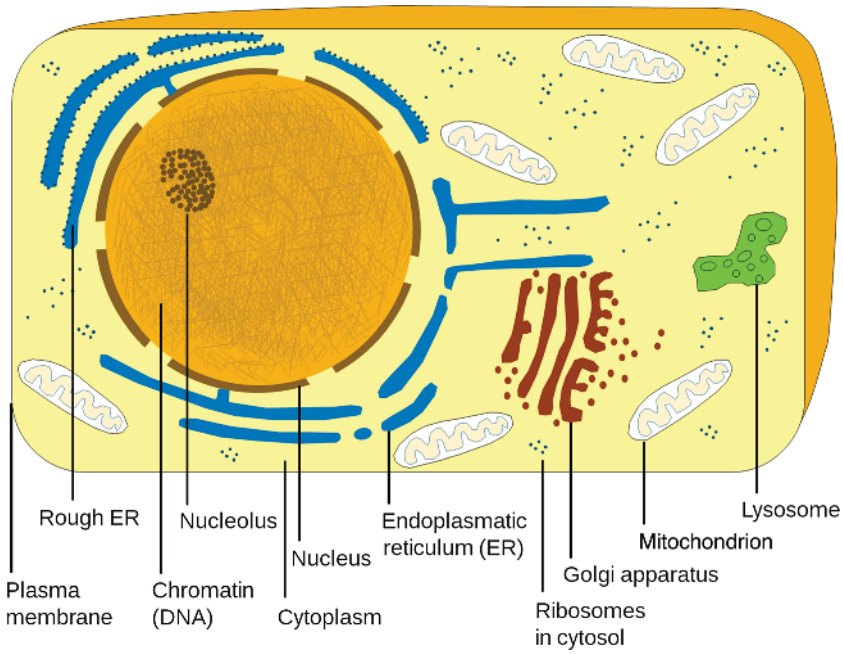

Fig. 1.3 Diagram of a typical eukaryotic cell. Membrane-bound organelles include the nucleus, endoplasmic reticulum (ER), Golgi apparatus, mitochondrion, lysosome, and peroxisome. (Reproduced with permission from Ref. [2]. Copyright 2008, Garland Science/Taylor & Francis LLC.)

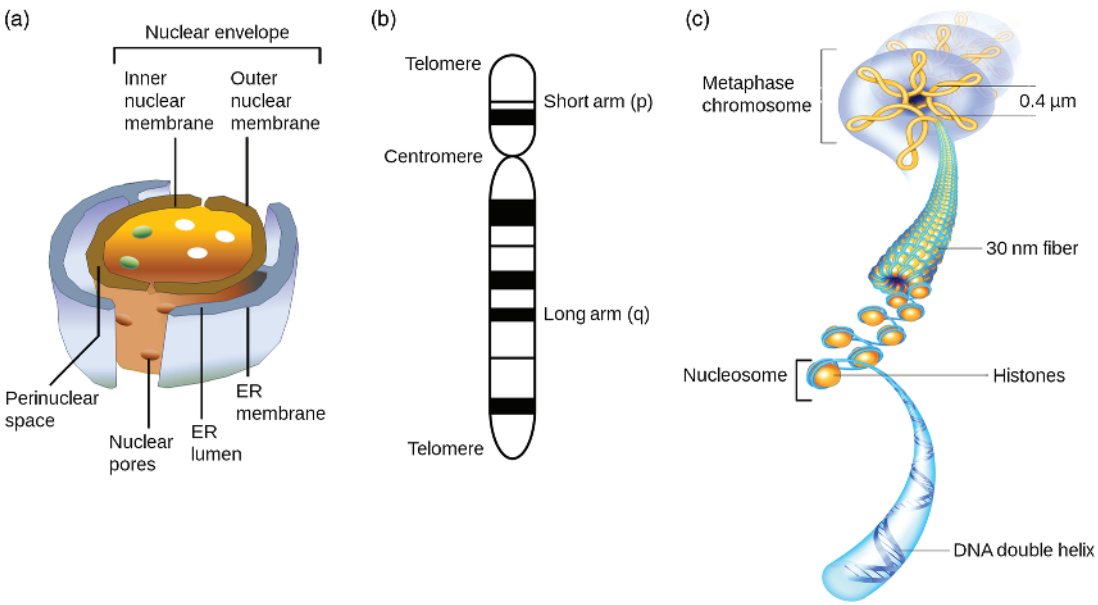

Fig. 1.4 Cell nucleus and chromatin organization. (a) The nucleus is surrounded by the nuclear envelope consisting of two membranes. Pores in the envelope allow the exchange of small water-soluble molecules. The outer nuclear membrane is continuous with the membrane of the endoplasmic reticulum (ER). (b) Schematic representation of a eukaryotic metaphase chromosome. Each chromosome has a specific banding pattern after staining. (c) DNA winds around histone proteins to form nucleosomes that fold into a 30 nm fiber. Loops of chromatin are then attached to a protein scaffold to form the metaphase chromosome. (Part (a) reproduced with permission from Ref. [2]. Copyright 2008, Garland Science/Taylor & Francis LLC. Part (c) adapted from Ref. [1] with kind permission from John Wiley & Sons, Inc.)

of the nucleus, while larger molecules must be actively transported in or out. A filamentous network of lamin proteins provides mechanical support, but is also involved in the regulation of replication and cell division. A distinct structure found in the cell nucleus is the nucleolus (indicated in Figure 1.3). The nucleolus occupies up to one fourth of the volume of the nucleus and forms around specific chromosomal regions. In the nucleolus, ribosomal RNA is transcribed and assembled with proteins to give (incomplete) ribosomes.

The DNA together with DNA-binding proteins and RNA molecules is organized in chromosomes. The human genome consists of 46 chromosomes that can be identified based on the specific banding pattern after staining. Figure 1.4b shows a schematic representation of a condensed metaphase chromosome. The ends of the chromosome are called *telomeres*. The *centromere* is the point where the two identical chromosomes touch and the microtubules attach for separation during mitosis. Most chromosomes are asymmetric; the centromere separates a short arm (p for the French word *petit*) and a long arm (q, chosen as next letter in the alphabet after p). The position of a gene or a DNA sequence in the human genome (denoted as the *locus*) is indicated by the chromosome number, the arm, and three numbers that refer to the region, the band and the subband. For example, the locus 11p15.3 indicates that a DNA sequence is located on the short arm of chromosome 11, in region 1, band 5 and subband 3.

If the DNA of every cell in the human body were lined up end to end, it would stretch from the Earth to the Sun and back 100 times. This means that a single cell must package DNA strands with a combined length of approximately 2 m into a nucleus with a volume in the cubic micrometer (μm^3) range.

Chromosomes consist of the DNA complexed with proteins, together known as chromatin (Figure 1.4c). The first step in chromatin condensation is achieved by winding the DNA around so-called histone proteins. These proteins are positively charged (i.e., contain many arginine and lysine residues) and can bind negatively charged DNA by electrostatic interactions. Two copies of each of the histones H2A, H2B, H3, and H4 form a core around which the DNA is wound twice. The fifth histone protein, the linker histone H1, binds to the middle of the DNA and to its two ends, locking the DNA into place. This structural unit, called the nucleosome, allows packaging in higher ordered structures, such as the so-called structural element "30 nm fiber."

The interaction between proteins and DNA has a major impact on gene expression (Section 1.3.3). Much of a cell's DNA is not expressed, as it does not encode a gene product. In addition, the expression of a particular gene may be inappropriate for a certain cell type. Non-expressed DNA is typically highly condensed; it is called *heterochromatin*. In contrast, lightly packed DNA (*euchromatin*) facilitates access to those enzymes required for gene expression and is, therefore, transcriptionally active. Nucleosomes also undergo remodeling as proteins required for transcription must bind to their target DNA. The *chromatin remodeling complex* disrupts the interaction between histones and DNA in the nucleosomes in an ATP-dependent manner so that the DNA becomes accessible. Histones undergo extensive covalent posttranslational modification (Section 1.3.5), which, together with DNA methylation, are two of the main mechanisms of the epigenetic regulation of gene activity (Section 1.3.4).

1.2.1.2 Mitochondria

Mitochondria (Figure 1.5) are organelles with double membranes that function as "cellular power plants" since they are the site of the oxidative metabolism that leads to the production of ATP. A eukaryotic cell typically contains 800–2000 mitochondria. In the internal matrix, they contain all the enzymes required for the citric acid cycle and for fatty acid oxidation. Redox proteins involved in the electron transport chain and oxidative phosphorylation are embedded in the inner membrane, which invaginates into cristae to expand its surface area.

A special feature of mitochondria is that they carry a genome that encodes several (but not all) mitochondrial proteins. The human mitochondrial genome is a circular molecule of approximately 17 000 base pairs encoding 13 proteins and 2 ribosomal and 22 transfer RNAs. Each mitochondrion contains multiple copies of the mitochondrial genome. Additional mitochondrial proteins are encoded in the nucleus. The mitochondria and their

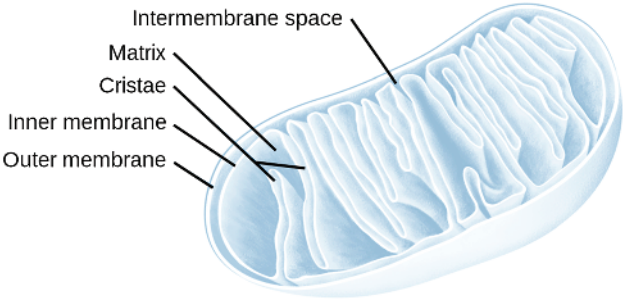

Fig. 1.5 The mitochondrion. The mitochondrion is bounded by two membranes. The inner membrane harbors the electron transport chain bearing the respiratory enzymes of the cell. It is extensively invaginated, forming the so-called cristae. The internal matrix contains high levels of enzymes involved in oxidative metabolism, and also holds the mitochondrial DNA. (Adapted from Ref. [1] with kind permission from John Wiley & Sons, Inc.)

genetic material are exclusively inherited from the mother. The *endosymbiotic hypothesis* is an attempt to explain the origin of mitochondria (and plastids such as chloroplasts in plants). According to this theory, mitochondria and plastids evolved from formerly free-living bacteria that were engulfed by another cell as an endosymbiont through endophagocytosis. The mitochondrial DNA (mtDNA) represents the remnant of the bacterial genome, while the remaining genetic material moved into the nuclear genome.

Mitochondrial dysfunction may result in mitochondrial diseases, which may be caused either by mutations in mtDNA or in nuclear genes that encode mitochondrial proteins (Section 3.1.4). Mitochondrial dysfunction may have severe consequences, due to their critical function in cellular energy supply. These effects usually vary from organ to organ and also depend on other variations in the genome. Mitochondrial diseases often lead to neuromuscular disease symptoms (myopathy), but also include diabetes mellitus, deafness, visual loss, and dementia.

1.2.1.3 Endoplasmic Reticulum and Golgi Apparatus

The endoplasmic reticulum and the Golgi apparatus play an important role in the posttranslational modification and sorting of proteins (Figure 1.6). The ER is a network of interconnected, flattened, membrane-enclosed sacs known as cisternae. The enclosed space is called the lumen (or cisternal space). The main function of the ER is to support the synthesis and export of proteins and membrane lipids. Two types of ER can be distinguished, the rough ER (RER) and the smooth ER (SER). While the SER is involved in lipid and carbohydrate metabolism and in detoxification, the RER plays an important role in posttranslational modification of proteins. The rough appearance is caused by ribosomes that bind to the cytosolic side of the ER membrane and translocate newly synthesized proteins into the lumen of the ER during the translation process. The main functions of the RER include the support and control of protein glycosylation and of correct protein folding, as will be described in more detail in Section 1.3.5. The RER carries out initial N-linked glycosylation, which is terminated in the Golgi apparatus. Proteins are shuttled from the RER to the *trans* Golgi network in membrane-bound vesicles.

The main functions of the Golgi apparatus include the termination of *N*-glycosylation and the initiation of *O*-linked protein glycosylation. Vesicles from the RER carrying proteins fuse with the membrane on the *cis* face of the Golgi apparatus and release their cargo. Proteins are modified in the lumen and leave the Golgi apparatus from the *trans Golgi network*, which sorts proteins for extracellular release or for transport to the lysosome, according to the markers they carry.

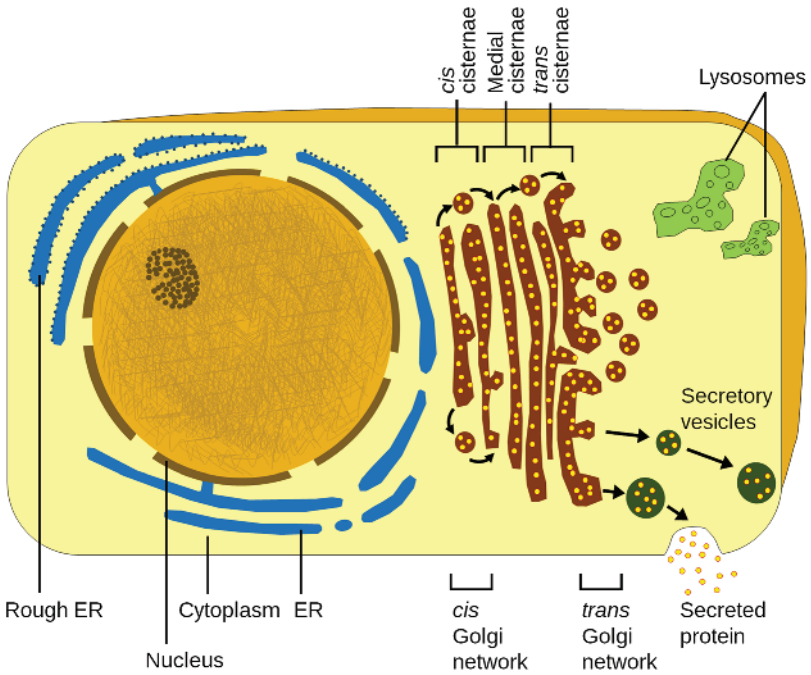

Fig. 1.6 Endoplasmic reticulum (ER) and Golgi apparatus. The endoplasmic reticulum and the Golgi apparatus are two organelles that are important for the correct folding, posttranslational modification, and sorting of proteins. *N*-Glycosylation of proteins is initiated in the ER and is terminated in the Golgi apparatus, while *O*-glycosylation exclusively occurs in the Golgi apparatus.

The importance of correct protein sorting is illustrated by the hereditary disease mucolipidosis II, which is characterized by psychomotor retardation, skeletal deformities, and death in childhood. The molecular cause of the disease is the lack of a phosphotransferase that normally phosphorylates mannose residues to mannose-6-phosphate on *N*-linked glycoproteins. Without this marker, enzymes that normally degrade glycosaminoglycans are transported from the Golgi apparatus into the extracellular space, rather than into the lysosome. This defect in protein sorting eventually leads to the accumulation of glycosaminoglycans and glycolipids in large inclusions in lysosomes that cannot be degraded. The disease is thus also known as Inclusion-cell (I-cell) disease.

1.2.1.4 Peroxisome and Lysosome

Peroxisomes and lysosomes are eukaryotic organelles with important metabolic functions, which consist mainly of degradative processes. Peroxisomes are involved in the catabolism of very long-chain and branched fatty acids and in the oxidative catabolism of amino acids. In addition, they contain enzymes of the pentose phosphate pathway. Peroxisomal disorders typically affect the nervous system, since peroxisomes also play a role in the biosynthesis of ether phospholipids, which are critical for the normal functioning of the mammalian brain.

Lysosomes are the cell's waste disposal and recycling system, as they contain enzymes that digest dispensable cellular components and ingested material. Lysosomes employ approximately 50 hydrolytic enzymes to break down cellular waste products, fats, carbohydrates, proteins, and other macromolecules into simpler compounds, which are returned to the cytoplasm as building blocks for biosynthetic processes. A variety of proteases, known as *cathepsins*, degrade proteins in a nonselective manner. Lysosomes maintain an interior pH of approximately 5, and their enzymes function optimally at acidic pH. The cytosol, with its higher pH (slightly above 7), is thus protected from the degradative activity of lysosomal enzymes that may leak into the cytosol, since they are inactive in this more alkaline environment.

Approximately 50 rare inherited metabolic disorders are known to result from malfunction of the lysosome. They are collectively referred to as *lysosomal storage diseases* (LSDs). The lack of a particular lysosomal enzyme may result in accumulation of substances destined for breakdown. LSDs have a variety of symptoms, but typically affect children, who die at a young age. Lysosomes also have a major influence on the pharmacokinetics of several drugs. Uncharged, weak bases may cross lysosomal membranes and accumulate in lysosomes. This phenomenon is called lysosomotropism and may explain the high tissue concentration and long elimination half-lives of some drugs. Accumulation in lysosomes may result in the inhibition of the activity of lysosomal enzymes.

1.2.2 Cell Cycle

The cell cycle is the sequence of events that leads to cell division into two daughter cells. It can be divided into three main periods: the interphase, the mitotic phase, and cytokinesis (Figure 1.7). Mitosis and cytokinesis together define the mitotic (M) phase of the cell cycle. A complete cell cycle in cultured cells takes approximately 16–24 h. In contrast, the doubling times of a cell in a multicellular organism can vary from 8 h to >100 days. Some terminally differentiated cells, such as muscle cells or neurons, never divide. This quiescent state is called the G_0 phase.

Interphase is characterized by cell growth, the accumulation of nutrients, preparation for mitosis, and the replication of DNA. It can be subdivided into three phases: The longest segment of interphase is the G_1 phase (G stands for gap), in which cells grow in size, have elevated biosynthetic activity, and produce enzymes required for DNA replication. DNA replication itself takes place in the synthesis phase (S phase), during which time all the chromosomes are replicated. S phase is followed by the second gap phase (G_2) during which time cells continue to grow. Several control mechanisms ensure that the cell is ready to enter mitosis.

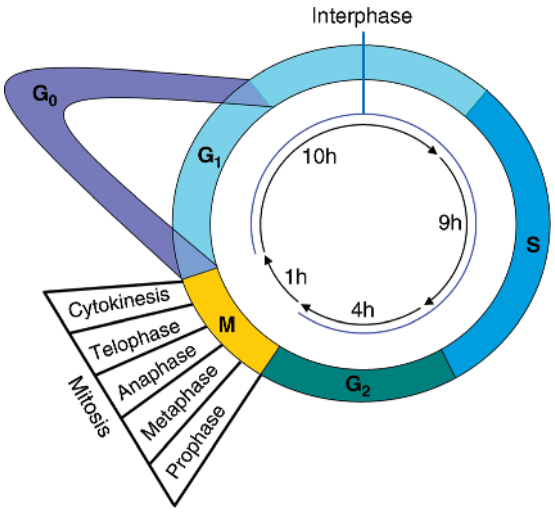

Fig. 1.7 The cell cycle. The eukaryotic cell cycle is divided into three main phases: (1) interphase consists of the G_1 and G_2 gap phases as well as the synthesis phase (S); (2) mitosis comprises pro-, meta-, ana-, and telophase; (3) in the final phase, cytokinesis, the cell divides.

Mitosis starts with *prophase*, which is characterized by spindle formation and condensation of the duplicated chromosomes. The nuclear membrane then breaks apart and the nucleus disappears (this process is sometimes called prometaphase). In the subsequent *metaphase*, the duplicated chromosomes move to positions midway between the spindle poles. The next step is the separation of the sister chromatids of the duplicated chromosomes and their movement to opposite poles of the cell. This occurs in *anaphase*. The last phase of mitosis is called *telophase* and involves decondensation of the chromosomes and restoration of the organelles and the nucleus. Although cytokinesis is not part of mitosis in the strict sense, it directly follows mitosis and involves division of the cytoplasm into two daughter cells.

As uncontrolled proliferation is a main characteristic of cancer, strict control of cell cycle is essential for the organism. The cell cycle has several checkpoints that monitor its progress and arrest the cell cycle if certain conditions have not been fulfilled. For example, the cell cycle only progresses to mitosis after the cell's DNA has been completely replicated. Another checkpoint ensures that all chromosomes are properly attached to the mitotic spindle.

Two classes of proteins, known as cyclins and cyclin-dependent kinases (Cdks), are the main regulators of progression through the cell cycle. Different types of cyclins are specific for each phase of the cell cycle. Cyclins are synthesized during one cell cycle phase and subsequently degraded during the succeeding phase. A cyclin forms a complex with its corresponding Cdk, which leads to the activation of the Cdk. Cdks are serine/threonine protein kinases that phosphorylate nuclear target proteins involved in the various processes of the cell cycle. Cdks are also regulated by cyclin-dependent kinase inhibitors, which arrest the cell cycle under certain conditions. These conditions include contact with adjacent cells, DNA damage, terminal cell differentiation, and cell senescence. Alterations of the inhibitory activity of Cdks are frequently found in cancer. Leland H. Hartwell, R. Timothy Hunt, and Paul M. Nurse were awarded the 2001 Nobel Prize in Physiology or Medicine for their discovery of the cyclin/Cdk system.

1.2.3 Apoptosis

Apoptosis (the Greek word for "falling off", for example, of leaves from a tree) is the process of programmed cell death. Apoptosis is a normal, physiological mechanism found in multicellular organisms that is important for the development and maintenance of proper physiological functioning and for the protection against cancer and other diseases. An example of the essential function of apoptosis in developmental processes is the formation of fingers in vertebrates. Initially, the digits are fully connected by webbing. During development, the web is eliminated by programmed cell death. An estimated 10^{11} of the $\sim 10^{14}$ cells of an adult human die every day through apoptosis and are replaced by new ones. Apoptosis is also important for the elimination of virus-infected or malignantly transformed cells. Abnormal apoptosis may result in neurodegenerative diseases such as Alzheimer's, Parkinson's, and Huntington's diseases, as well as in damage caused by stroke or heart attack.

The apoptotic program includes a sequence of cellular morphological changes: As a first step, the cell begins to shrink. Its chromatin becomes condensed at the nuclear periphery and the cytoskeleton collapses. Subsequently, the nuclear envelope disappears, the DNA is fragmented, and the plasma membrane forms irregular bulges (blebs). Eventually the cell disintegrates into numerous membrane-enclosed vesicles called apoptotic bodies. Apoptotic bodies are phagocytosed by neighboring cells and macrophages. Apoptosis should be differentiated from necrosis, which is a form of cell death that results from trauma, such as acute cellular injury. In this case, water rushes into necrotic cell due to the loss of cell membrane integrity. The cell swells, its organelles lose their function, and the cell eventually bursts, spilling its contents into the extracellular space.

The processes of apoptosis are mainly controlled by *caspases*, a family of *c*ysteinyl *asp*artate-specific prote*ases*. Caspases are heterotetramers, consisting of two α- and two β-subunits, which are expressed as inactive, single-chained zymogens (procaspases). Upon activation, the N-terminal prodomain is excised and the α- and β-subunits are proteolytically separated to yield the active $\alpha_2\beta_2$ caspase. Two types of apoptotic caspases can be distinguished: initiator and effector (executioner) caspases. Initiator caspases cleave inactive zymogen precursors of effector caspases, thereby activating them. The effector caspases then cleave other cellular proteins, triggering the apoptotic process.

Apoptosis may be induced in two ways (Figure 1.8): External signals trigger the *extrinsic pathway* (death by commission), while the absence of external signals that inhibit apoptosis can lead to the activation of the *intrinsic pathway* (death by omission). The extrinsic pathway is induced by binding of a so-called death ligand (e.g., the Fas ligand) on an inducing cell to the death receptor (Fas receptor) on the surface of the cell destined to undergo apoptosis. This leads to activation of the initiator caspase-8 (and possibly caspase-10), which triggers executioner caspases to cleave their substrate proteins, driving the cell into apoptosis. The internal pathway is

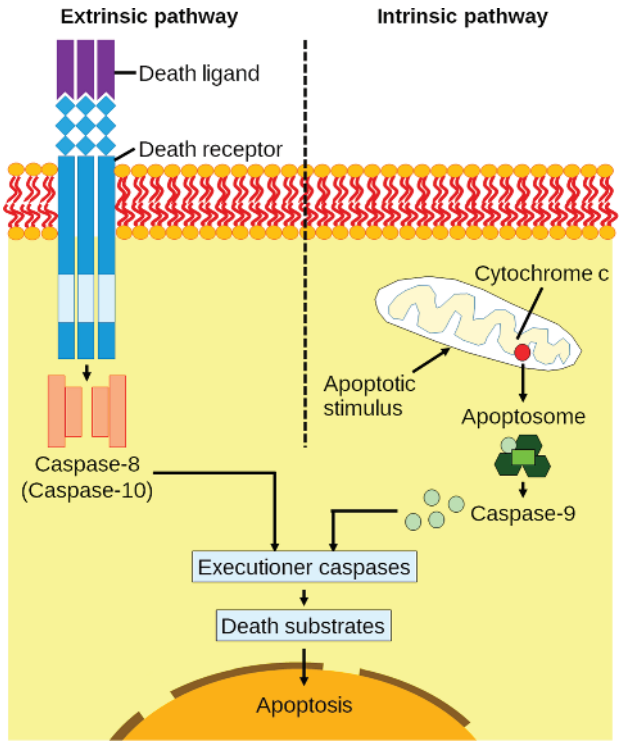

Fig. 1.8 Extrinsic and intrinsic pathways of apoptosis. Programmed cell death can be induced by extrinsic signals from a death ligand or by an apoptotic stimulus that leads to the release of cytochrome c from mitochondria. In both cases, an initiator caspase is activated that activates executioner caspases, which cleave their substrates. This program eventually leads to cell death. (Adapted from Ref. [3] with kind permission from Macmillan Publishers Ltd., Copyright 2002.)

induced in the absence of signals from the environment, for example, neighboring cells, which normally prevent the cell from committing suicide. Proapoptotic members of the Bcl-2 protein family are important mediators of these signals. They activate mitochondria to release cytochrome c into the cytosol, where it binds the apoptotic protease activating factor 1 (APAF1) to form the *apoptosome*. In the apoptosome, the initiator caspase-9 is activated, which then triggers executioner caspases as per the program described for the extrinsic pathway.

In 2002, the Nobel Prize in Medicine was awarded to Sydney Brenner, Robert Horvitz and John E. Sulston for their discoveries concerning the genetic regulation of organ development and programmed cell death.

1.3 DNA Replication and Gene Expression

(Anke Wagner, Berlin University of Technology, Berlin, Germany)

Cellular DNA must be copied before a cell can divide into two daughter cells. This occurs via a process known as *replication*. In addition, the genetic information stored in the DNA must be transferred into proteins, the actual primary effectors of cell function. This process of gene expression comprises two steps: the *transcription* of DNA into RNA and the *translation* of the RNA sequence into the corresponding amino acid sequence necessary for protein formation. This flow of genetic information, formulated in brief as "DNA makes RNA makes protein," was summarized by Francis Crick in 1958 in what is known as the *central dogma of molecular biology* (Figure 1.9). Crick himself acknowledged that the term dogma may be misunderstood as a belief that cannot be doubted. In fact, we now recognize several exceptions to the original model. The process of reverse transcription from RNA to DNA in retroviruses, or the replication of RNA molecules into new RNA molecules by other viruses (Section 5.1.1) are such examples. The dogma also does not cover processes such as epigenetic regulation of gene expression by DNA methylation (Section 1.3.4), posttranscriptional gene silencing by the RNA interference pathway (Section 13.3), or post-translational protein modification. However, the assumption that proteins cannot serve as the template for the

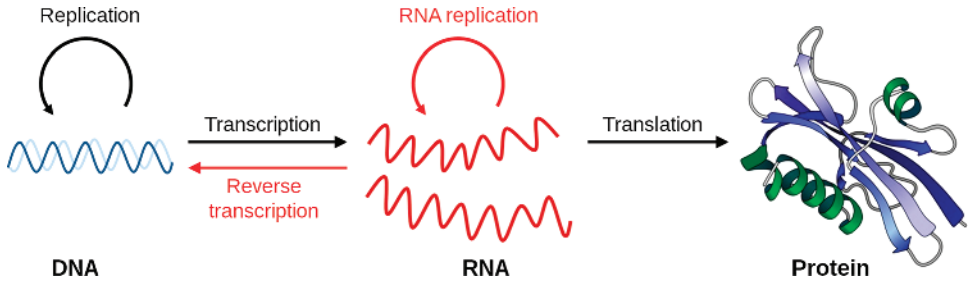

Fig. 1.9 The central dogma of molecular biology. DNA is copied in a process known as replication. The genetic information stored in the DNA is expressed by transcription into RNA, followed by translation into an encoded protein. Several exceptions to the model (marked in red) are known, including the replication of RNA into RNA and the reverse transcription of RNA into DNA.

synthesis of new proteins or for the transfer of sequence information into RNA or DNA is still valid.

1.3.1 DNA Replication

Replication is the process of copying the genetic information contained in cellular DNA. The double-helix model of DNA first proposed by Watson and Crick suggested a molecular mechanism for the transmission of hereditary information: Each strand of the double-stranded DNA molecule can serve as a template for the synthesis of a complementary daughter strand. Since the newly synthesized DNA double strands consist of one parental strand and one daughter strand, this mode of replication is termed *semiconservative*.

The process of prokaryotic and eukaryotic replication is similar, although it is more complex in the eukaryotes. Altogether, replication involves the concerted action of more than 25 proteins. DNA synthesis occurs at the so-called replication fork, a simplified scheme of which is depicted in Figure 1.10.

In eukaryotic cells, the initial step of replication is the separation of the two strands of the DNA by a heterohexameric helicase called minichromosome maintenance (MCM). The unwinding process, which is driven by the hydrolysis of ATP (or another NTP), causes supercoiling of the DNA ahead of the replication fork, which would eventually halt replication. To overcome these topological problems, the enzyme topoisomerase catalyzes the relaxation of DNA by cutting one or both strands of the DNA, rotating it, and reannealing the cleaved strand(s).

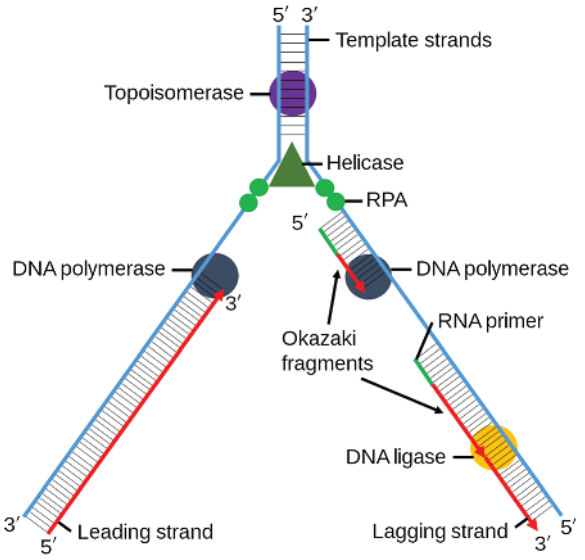

Fig. 1.10 Replication fork. Two DNA polymerase enzymes synthesize the new DNA strands, which are called the leading and lagging strands, respectively. A helicase unwinds the double-stranded template molecule and topoisomerases prevent supercoiling of the DNA. The replication protein A (RPA) stabilizes the unwound DNA single strands. DNA polymerase requires RNA primers to initiate synthesis of the Okazaki fragments, which are finally joined together by DNA ligase.

The two strands separated by the helicase have a strong tendency to reanneal to form a DNA duplex again. To prevent reannealing, the replication protein A (RPA) coats the single-stranded DNA. The central enzymes of DNA replication are the DNA-dependent DNA polymerases, simply known as DNA polymerases. Animal cells contain at least 13 distinct types of DNA polymerases. Like bacterial DNA polymerases, they can extend DNA strands only in the 5′ → 3′ direction. The two strands of parental DNA, therefore, have to be replicated in different ways. While the *leading strand* is synthesized continuously in the 5′ → 3′ direction of the replication fork movement, the other strand, known as the *lagging* strand, is also synthesized in the 5′ → 3′ direction, but can only be made in small segments known as Okazaki fragments. The overall process is, therefore, described as a *semidiscontinuous* mode of replication.

Another common feature shared by all known DNA polymerases is that they cannot synthesize a new strand *de novo*, but can only add nucleotides to the free 3′-OH group of a base-paired polynucleotide. The replication of DNA is, therefore, initiated by the synthesis of a short RNA primer. In eukaryotes, a primase synthesizes a 7- to 10-nt RNA primer. DNA polymerase α is associated with the primase and extends the primer by approximately 15 DNA nucleotides. The chain is then further extended by DNA polymerases δ and ε. Both polymerases are highly processive, meaning they can polymerize long stretches of the DNA without dissociating from the strand. DNA polymerase δ needs a sliding clamp protein named proliferating cell nuclear antigen (PCNA) to maintain its processivity. PCNA is also required for the synthesis of the lagging strand. In contrast, DNA polymerase ε is highly processive in the absence of PCNA and is involved in leading strand synthesis.

Both polymerases δ and ε carry out an additional *proofreading* step, for which they possess 3′ → 5′ exonuclease activity: In case the DNA polymerase incorporates a mismatched nucleotide into the newly synthesized strand, its polymerase activity is halted and the mismatched nucleotide is hydrolytically excised by the exonuclease activity. The enzyme can then incorporate the correct nucleotide and resume its normal polymerase activity. DNA polymerases with proofreading activity are characterized by high fidelity, that is, their error rates are only about one in every 10^6–10^{11} nucleotides, depending on the type of polymerase.

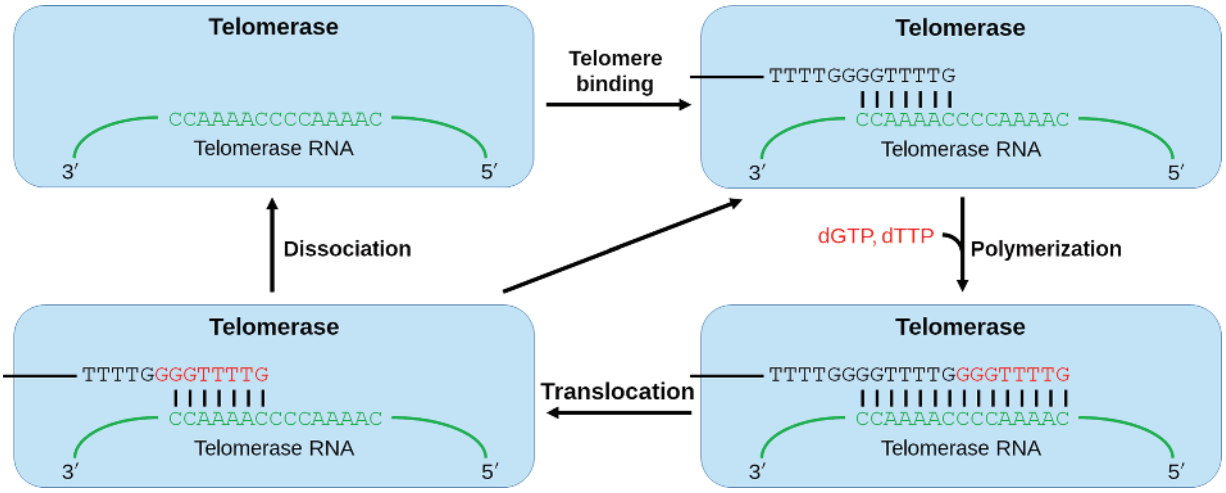

Fig. 1.11 Telomerase. Telomerase is a cellular reverse transcriptase that extends the ends of linear chromosomes (telomeres). It contains an RNA component to which the single-stranded 3′ end of the chromosome hybridizes. Telomerase then adds several nucleotides. The telomere can translocate and undergo several rounds of extension before it dissociates from telomerase.

Eventually, the RNA primers must be removed from the DNA and the Okazaki fragments must be joined together. In eukaryotes, this step is initiated by RNase H1 (H stands for hybrid, since the enzyme recognizes a DNA–RNA hybrid and hydrolyzes the RNA; this catalytic activity is also used in antisense technologies, as will be described in Section 13.1). The RNase removes most of the RNA and leaves only one RNA nucleotide adjacent to the DNA, which is then removed by the flap endonuclease-1. The excised nucleotides are replaced by DNA polymerase δ and the nick is eventually sealed by DNA ligase.

Eukaryotic chromosomes are large: The longest human chromosomes consist of more than 200 million base pairs. To achieve replication in a few hours, eukaryotic chromosomes contain multiple origins of replication, one every 3–300 kb depending on the species and tissue. Clusters of 20–80 adjacent replicating units (*replicons*) are activated simultaneously. Several sets of these clusters must be activated so that the entire chromosome can be replicated.

The linear ends of eukaryotic chromosomes, called telomeres, present a particular problem for the replication machinery. As already outlined, an RNA primer is required for the initiation of DNA replication. This primer can be removed but not replaced by the DNA polymerase. To prevent shortening of the linear chromosomes with each round of replication, eukaryotes have developed a mechanism to maintain the telomeres. The enzyme telomerase is a ribonucleoprotein that consists of a protein and an RNA component (of 451 nucleotides in humans). Since it functions as a reverse transcriptase, as illustrated in Figure 1.11, the protein component is called telomerase reverse transcriptase (TERT). The RNA component of telomerase is a complementary sequence to the repeating telomeric sequence and directs the addition of nucleotides to the 3′ end of the DNA. This process of polymerization followed by translocation can be repeated several times before the telomere dissociates from the enzyme. The complementary strand can then be synthesized by the normal cellular machinery for lagging strand synthesis.

Somatic cells of multicellular organisms lack telomerase activity; thus, the telomeres shorten with each round of replication. Cells in culture can only undergo a certain number of cell divisions before they reach senescence and die. The absence of telomerase in somatic cells is at least part of the basis for aging in multicellular organisms. The presence of shortened telomeres in somatic cells used for the cloning of animals is also believed to be one of the reasons for the health problems found in these clones (Section 10.1.5, Box 10.1). Telomerase is active mainly in two types of cells: germ cells that must maintain intact telomeres and cancer cells that divide rapidly and would stop growing without a mechanism to prevent shortening of the telomeres. The senescence of somatic cells and the process of aging due to lack of telomerase activity can be viewed as mechanisms that protect multicellular organisms from cancer. Telomerase is an attractive target for new anticancer drugs, since its inhibition should prevent the uncontrolled growth of tumor cells.

1.3.2 Mutations

Mutations are changes in the genetic information. They may cause cancer or genetic disorders. A large number of human cancers (estimates reach up to 80%) arise from substances referred to as *carcinogens*, which induce mutations. Over 70 000 man-made chemicals are currently of commercial importance and 1000 new ones are introduced every year.

Germinal mutations occur in germline cells and are transmitted through the gametes to the progeny. In contrast, *somatic mutations* occur in somatic cells; thus, the mutant phenotype will occur only in the descendants of that cell and will not be transmitted to the progeny. While *spontaneous mutations* occur without any known cause, perhaps due to inherent metabolic errors or to unknown agents in the environment, *induced mutations* result from exposure of organisms to mutagens. The degeneracy and order in the genetic code outlined below help prevent many mutations from affecting the phenotype of the organism. These changes are called *neutral* or *silent mutations*.

Several types of mutations are known: These include substitutions of one base pair for another and deletions and insertions of one or more base pairs (Figure 1.12). The most common type of mutation is the substitution of one base pair for another. A *transition* replaces a pyrimidine with another pyrimidine or a purine with another purine. A *transversion* replaces a pyrimidine with a purine or a purine with a pyrimidine. Altogether, four different types of transitions and eight different types of transversions are possible in DNA.

An example of a germline mutation is a single base substitution in the ß-globin gene that results in sickle-cell anemia in homozygous individuals (Sections 3.1.2 and 6.2). This disorder is caused by an A to T substitution, causing the hydrophilic amino acid glutamic acid to be replaced with the hydrophobic amino acid valine at the sixth position of the ß-globin protein. The exchange promotes the noncovalent polymerization (aggregation) of hemoglobin, which distorts red blood cells into a sickle shape and decreases their elasticity.

Base pair insertions or deletions within the coding regions of genes usually lead to so-called *frameshift mutations* (except if three nucleotides or multiples thereof are added or removed), because they alter the reading frame of all the base pair triplets. Since the triplets specify codons in mRNA and amino acids in the encoded protein (Figure 1.12), frameshift mutations mostly result in the synthesis of nonfunctional protein products.

Mutations can be induced physically (e.g., by UV radiation), chemically, or biologically (e.g., by transposons or viruses). Chemical mutagens may be mutagenic for replicating DNA only (e.g., acridine dyes) or mutagenic for both replicating and nonreplicating DNAs (e.g., alkylating agents and nitrous acid). For example, treatment of DNA with nitrous acid causes point mutations by oxidative deamination. Thus, cytosine and adenine can be converted to uracil and hypoxanthine, respectively.

Insertions and deletions can arise from the treatment of DNA with intercalating compounds such as acridine orange. Intercalation of such an agent into DNA almost doubles the distance between two consecutive base pairs. As a consequence, the replication of the distorted DNA occasionally results in the insertion or deletion of one or more nucleotides.

Environmental factors, such as UV or ionizing radiation, and some chemical compounds can also cause physical damage to the DNA. For example, UV light promotes the formation of a cyclobutyl ring between adjacent thymine residues that leads to the formation of

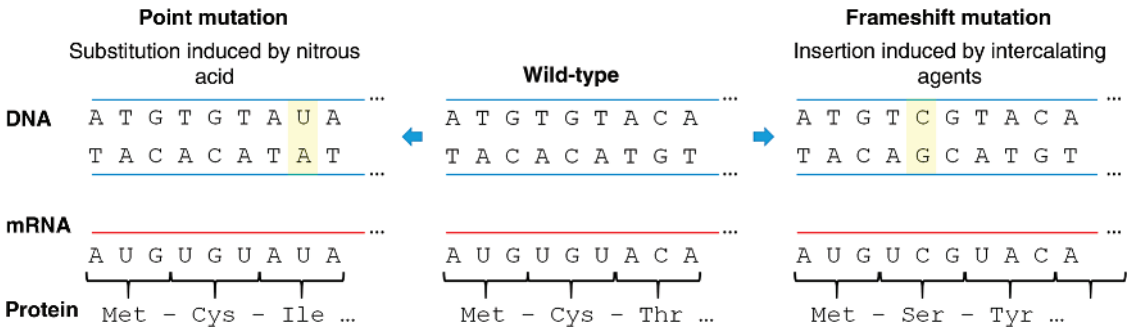

Fig. 1.12 Types of mutations. Substances such as nitrous acid can induce point mutations, that is, the substitution of one base by another. This may lead to a change of the codon and the incorporation of an incorrect amino acid into the protein. Frameshift mutations, induced, for example, by intercalating agents, occur by the introduction (or deletion) of one or multiple nucleotide(s). A reading frameshift changes all the codons and results in the synthesis of a nonsense protein.

intrastrand thymine dimers. Cytosine or cytosine–thymine dimers can also be formed, but at lower rate. These pyrimidine dimers all result in errors during transcription and replication. Treatment of DNA with ionizing radiation may lead to strand breakage. DNA damage may either be caused directly by radiation or by the formation of free radicals in the surrounding aqueous medium. Cells have developed several very effective mechanisms for the repair of DNA damage. Clinically, individuals with inherited disorders in the repair systems can develop disease such as *xeroderma pigmentosum*, which can lead to the early development of cancer through exposure to sunlight (Section 3.1.2).

The *Ames test* is a widely employed method to investigate the mutagenic potential of a substance. It is a quick and convenient bacterial test that serves as an alternative to time-consuming carcinogenesis tests on mice and rats; such evaluation in animals may take up to 3 years. In the assay, a special tester strain of *Salmonella typhimurium* that cannot synthesize histidine (denoted as his^-) is used. These bacteria are unable to grow in the absence of the amino acid. Bacteria of the tester strain are spread on a culture plate lacking histidine. In the presence of a mutagen, some bacteria revert to the his^+ phenotype, that is, they regain the ability to synthesize histidine and can grow and form visible colonies on the histidine-deficient medium. When this happens unusually frequently, it is an indication that the tested material is mutagenic.

1.3.3 Transcription

Transcription is the process of copying of DNA into RNA by the enzyme RNA polymerase (RNAP). Similar to replication, transcription relies on the transmission of genetic information by the pairing of complementary bases; however, transcription usually involves only copying of one of the two DNA strands, while both strands are doubled during replication. The DNA strand that serves as a template for the synthesis of RNA is called the noncoding or antisense strand, while the strand with the same sequence as the RNA is known as the coding or sense strand. As will be outlined in various chapters of this book, it has become apparent only recently that a great diversity of RNA molecules exists (Sections 7.2.4 and 13.4.2), all of which are generated by RNAPs. Animals contain three distinct types of RNAPs in the cell nucleus (as well as a separate mitochondrial RNAP), which transcribe different types of RNAs (Table 1.1).

Eukaryotic RNAPs are large complexes consisting of up to 14 subunits. They possess DNA unwinding activity, synthesize the new RNA molecule in the 5′ → 3′ direction, and carry out a proofreading step similar to the one described for DNA polymerases. The mode of action of the three RNAPs is somewhat different, and the following paragraphs will focus on the synthesis of mRNA precursors by RNAP II.

Table 1.1 Eukaryotic RNA polymerases.

Polymerase	*Location*	*Transcribed RNAs*
RNA polymerase I	Nucleolus	rRNA precursors
RNA polymerase II	Nucleoplasm	mRNA precursors; U1-, U2-, U4-, and U5-snRNAs
RNA polymerase III	Nucleoplasm	5S-rRNA, tRNAs, U6-snRNA, most miRNAs, and various other small RNAs

The table summarizes the location of the three eukaryotic RNA polymerases in the nucleus and the types of RNAs they transcribe.

Transcription is initiated at specific sites of the genome known as promoters. RNAP II promoters contain certain sequence motifs, sometimes referred to as core promoter elements (Figure 1.13), such as a GC-rich region or the so-called TATA box (thymidine–adenosine–thymidine–adenosine) located 25–31 nucleotides upstream of the transcription start site (it is often overlooked that approximately two thirds of protein-encoding genes lack the TATA box!). The promoters of housekeeping genes that are expressed in all tissues have a GC box upstream of their transcription start site. However, this GC box is missing in structural genes selectively expressed in certain cell types. About half of the protein-coding genes have a conserved 7-nucleotide initiator element (Inr) that includes the first nucleotide of the transcribed RNA (+1). None of the core promoter elements are present in all promoters. Further elements, such as the CCAAT box located in the region −50 to −110 to the transcriptional start site, are indicated in Figure 1.13.

Additional elements, called enhancers, are required for the full activity of their cognate promoters. These control elements do not have fixed positions and orientations relative to their corresponding transcribed sequences. Enhancers are recognized by specific transcription factors that stimulate RNAP II to bind to the corresponding, but often distant, promoter. This requires that the DNA between the enhancer and the promoter forms a loop that permits simultaneous interaction of the transcription factors with the enhancer and the RNAP at the promoter site.

While a large number of promoters are ubiquitously active in all cell types, others are tissue specific. For example, α_1-antitrypsin is a protease inhibitor that is expressed only in the liver. Its promoter, abbreviated

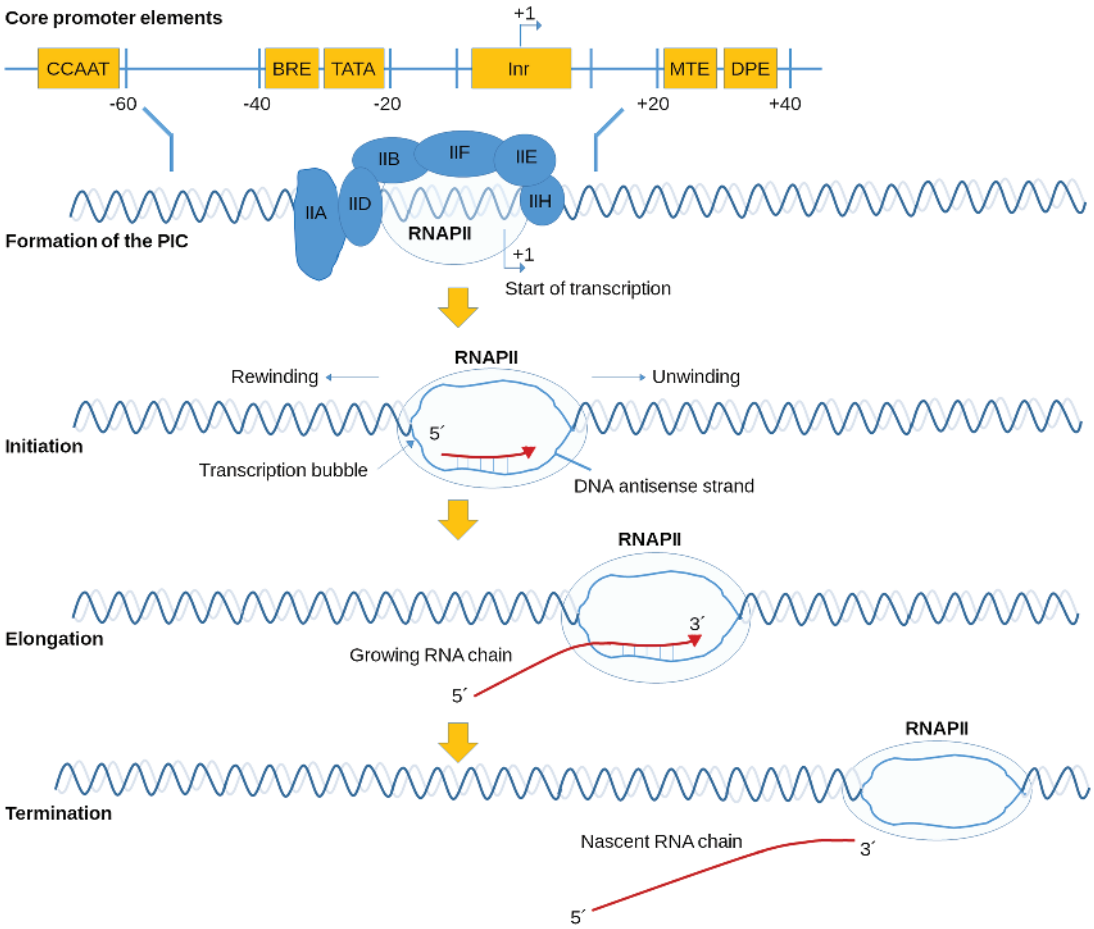

Fig. 1.13 Model of transcription. The DNA is shown in blue and the growing RNA strand in red. A preinitiation complex (PIC) is formed by binding of RNA polymerase II (RNAP II) and general transcription factors (blue circles; IIA, B, D-F, H) to the core promoter sequences. The initiation step includes unwinding of the double-stranded DNA to form an "open transcription complex." Elongation of the RNA chain occurs in the 5′ → 3′ orientation. During termination, the nascent RNA chain is released. CCAAT, CCAAT box; BRE, TFIIB recognition element; TATA, TATA box; Inr, initiator element; MTE, motif ten element; DPE, downstream promoter element.

as hAAT (human alpha antitrypsin), is active only in the liver and can be used for the tissue-specific expression of a transgene in gene therapeutic applications (Section 11.3).

The process of transcription can be divided into three stages: initiation of the synthesis of a new RNA chain, elongation of the chain, and termination of the transcription and release of the newly synthesized RNA molecule (Figure 1.13). During initiation, RNAP associates with the promoter sequences near the transcription start site. A so-called closed promoter complex is formed as the DNA at the transcription start site is still double stranded. Initiation of transcription is mediated by several proteins called *transcription factors*. Accurate transcriptional initiation of most genes requires the presence of six general transcription factors (GTFs). The designation of GTFs starts with the initials TF (for transcription factor) followed by a Roman numeral (e.g., II indicates involvement in transcription catalyzed by RNAP II). The properties and functions of these multiprotein complexes are summarized in Table 1.2. Together with the RNAP, they form the so-called preinitiation complex (PIC).

Transcription factors are proteins that bind to specific DNA sequences, thereby controlling transcription. They perform this function alone or with other proteins in a complex by promoting (as an activator) or blocking (as a repressor) the recruitment of RNAP to specific genes. All transcription factors contain one or more DNA-binding domains (DBDs), which attach to specific sequences of DNA adjacent to the genes that they regulate. Approximately 2600 proteins (~10% of all protein-coding genes) encoded in the human genome contain DNA-binding domains, and most are presumed to

Table 1.2 Functions of the eukaryotic general transcription factors.

Factor	*Number of unique subunits in humans*	*Function*
TFIIA	3	Stabilizes TBP and TAF binding
TFIIB	1	Stabilizes TBP binding; recruits RNAP II; influences start site selection
TFIID:		Recognizes TATA box; recruits TFIIA and TFIIB
TBP	1	
TAFs	16	
TFIIE	2	Recruits TFIIH and stimulates its helicase activity; enhances promoter melting
TFIIF	2	Facilitates promoter targeting; stimulates elongation
TFIIH	10	Contains an ATP-dependent helicase that functions in promoter melting and clearance

TF: transcription factor; TBP: TATA-binding protein; TAF: TBP-associated factor.

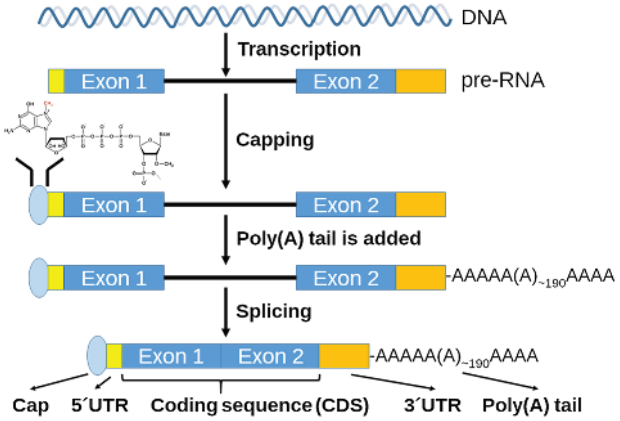

Fig. 1.14 Posttranscriptional processing of pre-mRNA. The products of transcription, the *primary transcripts*, can be altered by appending nucleotides to the 5′ end (capping) or the 3′ end (poly(A) tailing). Introns are removed during the splicing process. Posttranscriptional modifications lead to the formation of typical human mRNAs, including the untranslated regions (UTR).

function as transcription factors. Transcription factors are involved in numerous important biological roles, such as basal transcription regulation (GTFs TFIIA, B, D, E, F, H), the differential enhancement of transcription, developmental processes, response to intercellular signals (e.g., estrogen receptor transcription factors) or environmental factors (e.g., hypoxia inducible factor), and cell cycle control (oncogenic factors (e.g., myc) or tumor suppressor proteins).

RNA synthesis, such as DNA synthesis, occurs in the 5′ → 3′ direction (Figure 1.13). In the region being transcribed, the DNA double helix is unwound by RNAP so that the template strand can be transcribed into its complementary RNA strand. This complex is referred to as an "open promoter complex." During transcription, the DNA antisense strand transiently forms a short RNA–DNA hybrid with the nascent RNA chain. The transcription bubble travels along the DNA.

After initiation of transcription and the production of a short transcript, RNAP II changes into an elongation mode and leaves (clears) the promoter. A complex composed of six proteins known as "elongators" binds to the polymerase and assists with the synthesis of the mRNA. These proteins are dispensable, but transcription is accelerated in their presence.

In eukaryotes, a termination sequence for transcription has not been identified. A defined termination site is not required, since the transcript undergoes posttranscriptional processing (Figure 1.14): The major steps of maturation of the *primary transcript* are the addition of a "cap" at the 5′ end, appending a poly(A) tail to the 3′ end, and the removal of noncoding introns by a process denoted as splicing.

Eukaryotic mRNAs have a peculiar 7-methylguanosine at the 5′ end that is linked to the transcript via a 5′-5′ triphosphate bridge. This structure is called the "cap." It is added to the growing transcript before it is 30 nucleotides long. After formation of the 5′-5′ bridge, the guanine is methylated to produce 7-methylguanosine. Further nucleotides may be methylated at the 2′-OH-group of the ribose, generating Cap1 (first nucleotide of the transcript is methylated) or Cap2 (first and second nucleotide of the transcript are methylated).

Almost all mammalian mRNAs have 3′-poly(A) tails consisting of ~250 nucleotides. The poly(A) tail is generated from ATP by poly(A) polymerase, a template-independent RNA polymerase. The main function of the poly(A) tail is protection against nucleases. Poly(A) tails are added in a two-step reaction: In the first step, the transcript is cleaved 15–25 nucleotides downstream of an AAUAAA sequence, and less than 50 nucleotides before a U- or U- and G-rich sequence. In a subsequent reaction, the poly(A) tail is generated from ATP in a stepwise action by poly(A) polymerase. The cleaving enzyme and poly(A) polymerase are located within a large protein complex. Hence, the cleaved mRNA cannot dissociate before polyadenylation. In the cytoplasm, the poly(A) tail enhances the stability of the mRNA by interacting with the poly(A)-binding protein (PABP).

In eukaryotic genes, coding sequences (exons) are interspersed with noncoding sequences (introns). The primary transcripts initially transcribed are heterogeneous in length (*heterogeneous nuclear RNA* (*hnRNA*)). After capping and polyadenylation, the introns are excised and the exons are spliced together to form the mature mRNA (Figure 1.15). On average in humans, exons are 150 nucleotides long. Most of them are <300 nucleotides in length, but the largest exon (from the muscle protein titin) is 17 106 nucleotides in length.

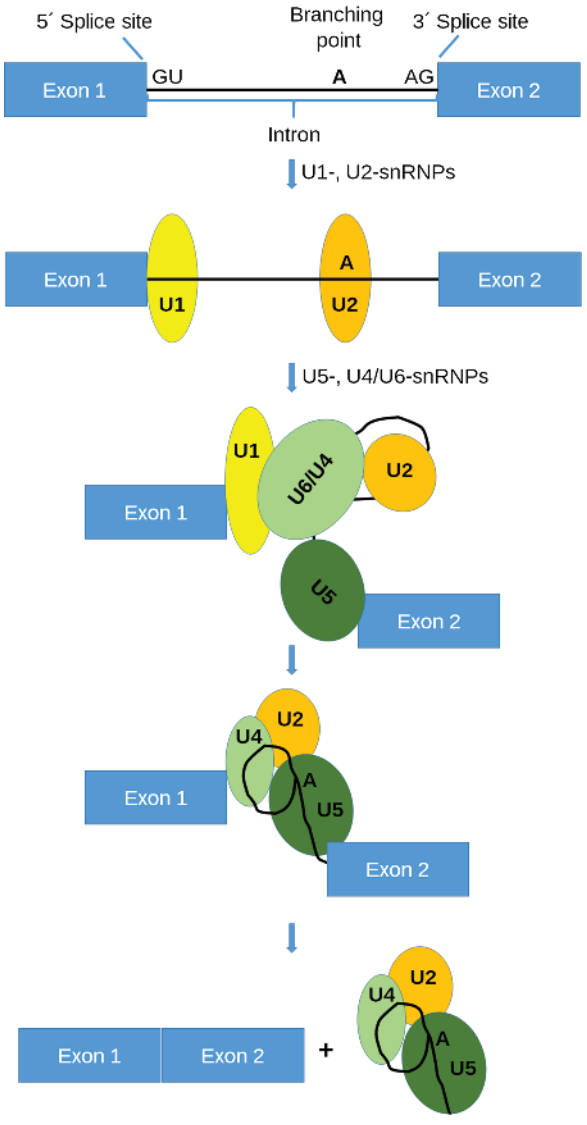

Fig. 1.15 The splicing reaction. In the first step, U1-snRNP binds to the 5′ splice site and U2-snRNP attaches to the branch site. Afterward the complete spliceosome assembles and cleaves the transcript at the 5′ splice site. The 5′ end of the intron is joined to the adenine in the branch position to form a lariat. U1- and U4-snRNP leave the complex. Finally, the 3′ splice site is cleaved. The 5′ end of exon 2 is joined to the 3′ end of exon 1 and the lariat structure is released.

Introns are usually much longer. On average, they are 3500 nucleotides in length, but they may be as long as 2.4 million nucleotides. The number of introns in a gene in the human genome averages ~8 and can vary from none to more than 200.

Splicing occurs via two transesterification reactions. The first one yields a 2′,5′-phosphodiester bond between an adenosine at the branch point and the 5′ splice site. This process results in the formation of a lariat structure, as schematically shown in Figure 1.15. The second transesterification reaction splices the two exons together and releases the intron.

The splicing reactions are mediated by the *spliceosome*, which consists of the pre-mRNA, small nuclear ribonucleoprotein complexes (snRNPs) (commonly pronounced "snurps") and a number of pre-mRNA-binding proteins. In total, the spliceosome contains more than 100 proteins and five RNA molecules, called small nuclear RNAs (snRNAs). The snRNP complexes are named U1, U2, U4, U5, and U6, where U indicates that the RNAs are uracil-rich.

The composition of a protein encoded by a gene can be modulated by the selection of alternative splice sites. The most common mode of alternative splicing in mammalian cells is exon skipping. This is where an exon may either be spliced out of the primary transcript or preferentially retained (Figure 1.16a). Further modes of alternative splicing are the mutual exclusion of exons, the use of alternative splice donor or acceptor sites, and the retention of an intron. For the muscle protein α-tropomyosin, alternative splicing leads to the generation of seven tissue-specific variants encoded by a single gene (Figure 1.16b).

Alternative splicing is found in all multicellular organisms and is especially prevalent in vertebrates. It has been estimated that up to 95% of all human genes undergo alternative splicing. On average, each pre-mRNA is spliced to give three different variants. Splice variants may control the subcellular localization of a protein or whether it is soluble or hydrophobic. The selection of splice sites is tissue- and developmental stage-specific and, therefore, tightly regulated. Several studies have stated that 30–60% of the genetic diseases in humans are pre-mRNA splicing defects caused by point mutations. In addition, tumor progression is correlated with changes in the protein levels implicated in alternative splice site selection.

After completion of the various steps of post-transcriptional processing, a typical, mature eukaryotic mRNAs consists of I: the cap at the 5′ end; II: a short 5′ untranslated region (UTR); III: the coding sequence (CDS); IV: the 3′ UTR; and V: the poly(A) tail (Figure 1.14).

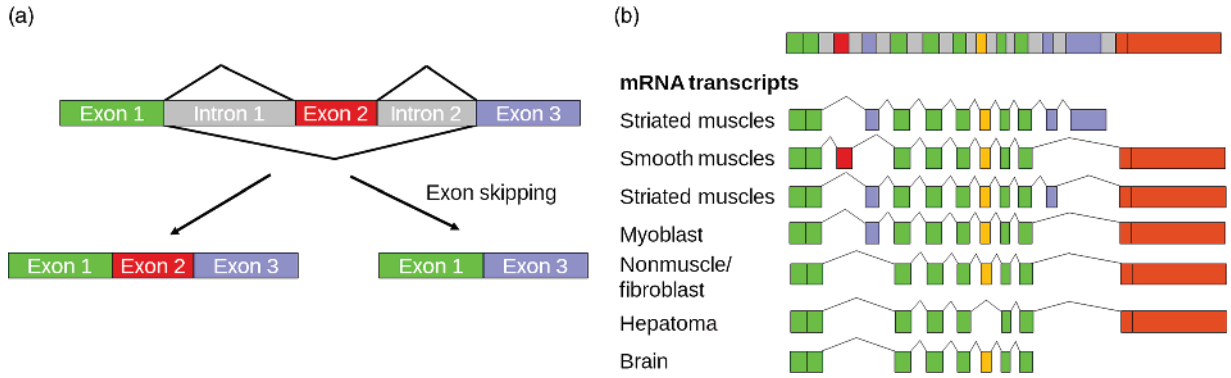

Fig. 1.16 Alternative splicing. (a) Exon skipping. Exon skipping is the most common mode of alternative splicing in mammals. In this case, an exon may be spliced out of the primary transcript or retained. (b) Alternative splicing of rat α-tropomyosin. Seven α-tropomyosin variants are produced in a tissue-specific manner by alternative splicing. The thin lines indicate the positions of introns before splicing. Tissue-specific exons are indicated with a color code: Green boxes represent exons expressed in all tissues; the exon indicated by a brown box is expressed in smooth muscles (SM); purple boxes depict exons expressed in striated muscles; variably expressed exons are depicted as yellow boxes. (Part (b) adapted from Ref. [1] with kind permission from John Wiley & Sons, Inc.)

Similar to mRNAs, rRNAs and tRNAs undergo post-transcriptional processing. The eukaryotic genome typically has several hundred tandemly repeated copies of rRNA genes that are contained in small, dark-staining nuclear bodies, the nucleoli (see nucleolus above). The nucleolus is a nonmembrane-bound substructure, in which rRNA transcription and processing and ribosomal subunit assembly take place. The primary rRNA transcript is about 7500 nucleotides in length and contains three rRNAs separated by spacer sequences.

In the first stage of processing rRNAs, the 45S RNA precursor is specifically methylated at numerous sites (106 sites in humans). In addition, many pre-rRNA uridines (95 in humans) are converted to an isomeric nucleoside called pseudouridine, which is characterized by a C-glycosidic bond instead of the normal *N*-glycosidic bond. The subsequent cleavage and trimming of the 45S RNA is catalyzed by enzymes possessing RNase activities. Some bases in rRNAs are methylated, a process that is guided by members of the large family of small nucleolar RNAs (snoRNAs). The length of these snoRNAs varies from 70 to 100 nucleotides. They contain segments of 10–21 nucleotides that are complementary to fragments of the mature rRNAs that contain the nucleotides to be methylated at the 2′ OH position. Methylation is mediated by a complex of at least six nucleolar proteins, including *fibrillarin*, the likely methyltransferase. Eukaryotic ribosomes contain four different rRNAs. While the 28S, 18S, and 5.8S rRNAs are processed as described, the fourth type, the 5S rRNA, is separately processed in a manner similar to tRNAs.

tRNAs deliver amino acids to the ribosome, the protein synthetic machinery. They consist of ~80 nucleotides and are commonly represented as a two-dimensional *cloverleaf structure*, although their three-dimensional structure was determined to be L-shaped. tRNAs contain a substantial number of modified bases (Figure 1.17). Mature tRNAs have the highest content of covalent modification diversity of all classes of RNA molecules. Up to 25% of the nucleotides of a tRNA are covalently modified. For accurate translation, the correct amino acid needs to be attached to its tRNA; the amino acid is linked to the 3′-terminal CCA sequence. This reaction is catalyzed by *aminoacyl-tRNA synthetases* that specifically link amino acids to the 3′-terminal ribose residue of their respective tRNAs to form aminoacyl-tRNAs. The *anticodon* that is complementary to the mRNA codon specifying the tRNA's amino acid is located in the loop of the cloverleaf structure opposite to the stem containing the terminal nucleotides.

Primary tRNA transcripts contain one to five tRNAs. The 5′ end of the tRNA is processed by RNase P, which consists of an RNA and a protein component. An important feature of RNase P is that the RNA component is the enzyme's catalytic subunit, that is, it functions as a ribozyme (Section 13.2). Endonucleases cleave the tRNA precursors close to the 3′ end. Subsequently, exonucleases such as RNase D remove nucleotides stepwise until they reach the 3′ end of the tRNA. Finally, CCA ends are posttranscriptionally added to the tRNA by the enzyme *tRNA nucleotidyltransferase*, which sequentially adds two Cs and an A using CTP and ATP, respectively, as substrates. Many eukaryotic primary tRNA transcripts contain introns adjacent to their anticodons that must be removed during the processing steps.

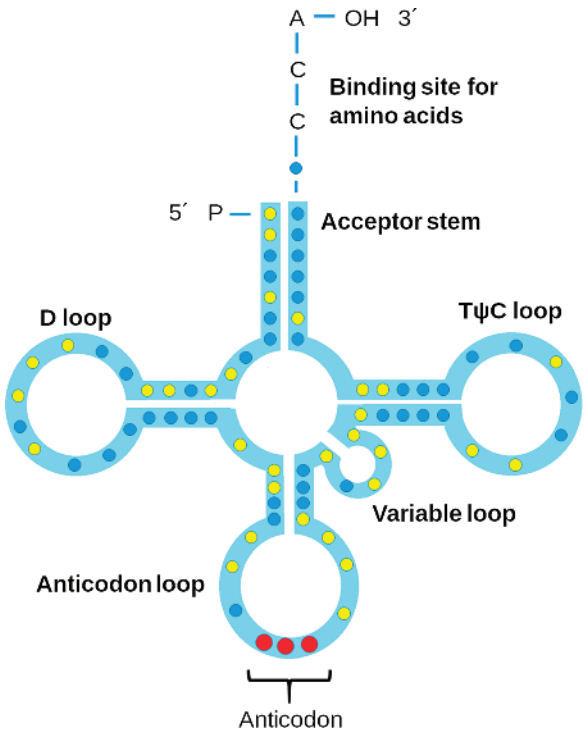

Fig. 1.17 Cloverleaf structure of a tRNA. Blue circles show nonmodified RNA nucleotides. Red circles reflect nucleotides forming the anticodon sequence. The average tRNA has 13 modified residues. The putative positions of modified nucleotides are highlighted by yellow circles. These modifications are highly conserved in eukaryotes, including humans.

1.3.4 Epigenetic Regulation of Gene Expression

The cell uses several mechanisms to regulate gene expression, for example, by the actions of transcription factors. A recent focus of intensive research is the *epigenetic* modulation of gene activity. The term "epigenetics" describes heritable changes in the level of gene expression that are not caused by changes in the DNA sequence. In a broader sense, epigenetics also refers to stable, long-term alterations in transcriptional levels that are not necessarily heritable. The term epigenetics (*epi*, Greek: outside of, around) was coined in 1942 by C.H. Waddington and describes phenomena other than those determined by the DNA sequence. Epigenetic processes are crucial for cellular differentiation. All cells of an organism contain the identical genetic information, but the selective activation and inhibition of genes determines their differentiation into a specific cell type, for example, a liver cell, a muscle cell, or a neuron.

Two main mechanisms produce epigenetic alterations: DNA methylation and histone modification (Figure 1.18), each of which influences gene expression without altering the DNA sequence. In addition, positioning of the nucleosomes relative to the transcription start sites and additional regulation by small noncoding RNAs such as miRNAs (Section 13.4) are sometimes considered mechanisms of epigenetic regulation.

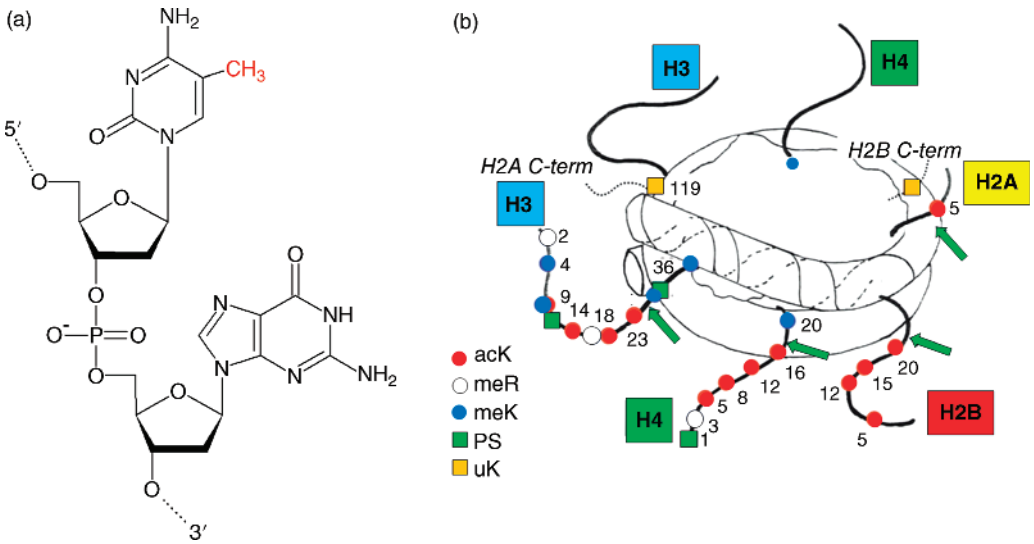

Fig. 1.18 DNA methylation and histone modification are the main mechanisms of epigenetic modification. (a) The chemical structure of a methylated CpG dinucleotide. The methyl group is marked in red. (b) The properties of histones are altered by posttranslational modifications. The figure shows the N- and C-terminal domains of histone proteins in the nucleosome core particle. Sites marked by green arrows are susceptible to cutting by trypsin in intact nucleosomes. Modifications are indicated by colored symbols: acK: acetyl lysine; meR: methyl arginine; meK: methyl lysine; PS: phosphoryl serine; uK: ubiquitinated lysine. (Part (b) adapted from Ref. [4] with kind permission from Elsevier.)

DNA methylation refers to the well-characterized chemical modification of nucleotides that influences the expression levels of certain genes. The only methylated base found in eukaryotic DNA is cytosine. This form of methylation occurs largely in dinucleotides of cytosine and guanosine (abbreviated as CpG dinucleotides, where p indicates the phosphate group connecting the nucleosides (Figure 1.18a)). CpG dinucleotides tend to cluster in regions called CpG islands. DNA methyltransferase (DNMT) enzymes transfer a methyl group to the C5 position of cytosines. The methyl group can be removed by demethylases.

CpG islands are frequently found in the 5′ region of genes, where promoters are situated. Methylation at cytosine switches off gene expression. In contrast, genes that are transcriptionally active will be hypomethylated at the CpG islands in their upstream regions. Methylation represses gene expression by recruiting chromatin corepressors and inhibiting the binding of transcription factors.

DNA methylation is stably maintained during cell division: The parental DNA strand directs the generation of the identical methylation pattern in its daughter strand. This process, also known as *maintenance methylation*, is mediated mainly by the DNA methyltransferase DNMT1 and results in the stable inheritance of methylation patterns in a cell line. Thus, all the cells will have the identical differentiated phenotype.

While methylation patterns are maintained during somatic cell division, they vary during the stages of embryologic development. Initially, the level of DNA methylation is high in mature gametes, but it is nearly eliminated in the blastocyst stage of development (i.e., approximately 200–300 cells, when the embryo embeds itself into the endometrical layer of the uterine wall). The level of DNA methylation then rises again. This *de novo* methylation is mediated by two specific DNA methyltransferases (named DNMT3a and DNMT3b). At the gastrula developmental stage (when the three germ layers ectoderm, mesoderm, and endoderm are formed), the level of DNA methylation rises to the adult level, where it remains for the lifetime of the individual. However, the CpG islands of germline cells are an exception, as they remain largely unmethylated throughout life.

There is a clear difference in the levels of DNA methylation between embryonic and somatic cells. This has consequences for patterns of gene expression. These differences probably explain (at least in part) the high failure rates of cloning experiments in mammals by somatic cell nuclear transfer (SCNT) (Section 10.1.5, Box 10.1). There, a nucleus of a differentiated somatic cell is injected into a denucleated egg. However, the epigenetic pattern of the transferred genomic material differs from that present naturally at this early stage of embryological development. As a consequence, only a small fraction of cloned embryos develop normally and even those that survive to birth usually have a variety of abnormalities. For example, the live newborn may be unusually large, a condition known as the large offspring syndrome.

DNA methylation has also been found to play an important role in cancer development, because of the induction of classic genetic mutations. Methylated cytosines mutate to thymidines at a comparatively high frequency, a mutational change frequently found in human cancers. These mutations may convert proto-oncogenes to oncogenes or inactivate tumor suppressor genes. In addition, many cancer cells also have a distorted epigenetic landscape. Typical changes include a profound global loss of DNA methylation, which occurs concomitantly with specific patterns of extended methylation at CpG islands of certain gene promoters. Hypermethylation at the CpG islands of tumor suppressor genes further favors tumor initiation and progression, whereas hypomethylation can induce aberrant expression of oncogenes and induce loss of imprinting (LOI) in some loci (see below).

Drugs are available that affect the epigenetic modulation of CpG motifs. An example is decitabine (5-aza-2′-deoxycytidine, trade name Dacogen) that hypomethylates DNA by inhibiting the DNA methyltransferase. This drug is indicated for myelodysplastic syndrome (MDS) and acute myeloid leukemia (AML). However, since the unspecific inhibition of DNA methylation has widespread consequences, the drug has many toxic side effects in treated patients.

After DNA methylation, posttranslational histone modification is the second major mechanism of the epigenetic fine-tuning of gene expression (Figure 1.18b). Histones are positively charged proteins that package and order genomic DNA. Modifications at the N-terminal tails of the histones influence the ability of genes to transcribe DNA into RNA and also affect replication, DNA repair, and chromosomal condensation. Posttranslational modifications include methylation, acetylation, phosphorylation, and ubiquitination. These modifications (except for methylation) produce negative charge, thus weakening the interaction of the histone with the negatively charged DNA. As a consequence, the chromatin becomes decondensed. In contrast, methyl groups increase the hydrophobicity of the side chain and tend to stabilize chromatin structure. The histone tails also interact with specific proteins, which further change the transcriptional accessibility of the associated genes. The different combinations of modifications are sometimes regarded as a *histone code*, where a certain modification pattern evokes a specific biological effect.

Posttranslational modification of histone proteins is carried out by specific enzymes, including histone acetyltransferases (HATs) and histone methyltransferases. These processes are reversible as histone deacetylases (HDACs), and the more recently discovered demethylases, may remove the substituents.

Aberrations of the histone-modifying proteins may have severe clinical consequences. An example is the Rubinstein–Taybi syndrome, which is characterized by short stature, distinctive facial features, and mental retardation. This condition is caused by a mutation in the CREB-binding protein (CBP), which binds to the cAMP-response element-binding protein (CREB). CBP has histone acetyltransferase activity and opens the chromatin structure. It also interacts with additional proteins of the transcription complex and controls the activity of many genes. CBP has an important role in regulating cell growth and division and is essential for normal fetal development.

Alterations in histone modification have also been found to promote tumor growth. A common alteration is the global reduction of an acetylated lysine residue in histone H4. The loss of acetylation is mediated by histone deacetylases that are overexpressed or mutated in different tumor types. Vorinostat (trade name Zolinza), used for the treatment of cutaneous T-cell lymphoma, was the first HDAC inhibitor that was approved by the FDA. In some cases, chromosomal translocations modulate the activity of histone acetylases and affect the global balance of histone acetylation. Altered methylation of histones is primarily found in hormone-dependent types of cancer, such as breast and prostate cancers.

Another important function of epigenetic regulation is *genomic imprinting*, a process by which certain genes are expressed in a manner dependent on the parent from which they are inherited. Since mammalian embryos carrying two copies of either only female or only male origin do not develop into viable offspring, the diploid genome always consists of genetic material obtained from both a male and a female parent. The vast majority of genes are expressed from both alleles simultaneously. However, a small proportion of genes are imprinted, that is, their expression originates from only one allele. In humans, only roughly 200 genes are known or predicted to be expressed in a manner dependent upon their parental origin. An example is insulin-like growth factor 2, which is only expressed from the paternal allele. Imprinted genes lie within clustered regions of the genome; the two largest regions are located on chromosomes 11 and 15 (11p15 and 15q11). Genomic imprinting employs DNA methylation and histone modulation to achieve monoallelic gene expression without altering the genetic sequence. Several imprinted genetic disorders have been described. For example, the Prader–Willi syndrome (PWS) and the Angelman syndrome (AS) are both associated with genetic defects in the 15q11-13 chromosomal region (Section 3.1, Box 3.1). However, the phenotypic consequences depend on whether the mutation is inherited from the father or from the mother. Paternal inheritance of the alteration is associated with the PWS, while inheritance of the same mutation from the mother leads to the development of AS.

Epigenetic modifications are also important for a process called X-inactivation (or lyonization after the discoverer of this phenomenon, Mary Lyon). As females carry two X chromosomes, one of them must be inactivated to prevent them from having twice as many X chromosomal gene products as males, who only posses a single copy of the X chromosome. This process is also known as *dosage compensation*. Although most of the X chromosome is inactivated, some segments escape this process because there are comparable genes on the Y chromosome.

X-inactivation is initiated from a specific site on the chromosome, which expresses a long noncoding RNA denoted as the X-inactive specific transcript (Xist). The inactive X chromosome is coated by this RNA, whereas the active X chromosome is not. It becomes highly condensed and visible under a microscope in the nucleus as the so-called Barr body. Once an X chromosome is inactivated, it will remain silenced throughout the lifetime of the cell and in all of its descendants in the organism. The inactivation occurs as a random process in the human embryo approximately 16 days after fertilization. Therefore, female cells are mosaics; some cells have an inactive maternal X chromosome, while others have an inactive paternal chromosome. The process of X-inactivation is important for the symptomatic occurrence of X-linked recessive disorders in females (Section 3.1.3).

In addition to the inherited disorders described above and in the development of cancer, epigenetic alterations also play an important role in some neurological and autoimmune diseases. The findings that epigenetics are highly relevant to health and disease have triggered large-scale studies, sometimes referred to as epigenomics. These initiatives have succeeded in characterizing genome-wide DNA methylation patterns at single-nucleotide resolution (known as the human DNA methylome), and have also found new variants and modifications of histones.

1.3.5 Translation

The genetic code is a set of rules that defines how the information encoded within genetic material is translated

into proteins. As only four bases must code for 20 amino acids, multiple bases (*codons*) are required to specify a single amino acid. Cells use a three-base code that can theoretically code for $4^3 = 64$ amino acids. Therefore, most of the amino acids are specified by more than one codon, that is, the code is *degenerate*: While most of the amino acids are specified by two, three, or four codons, the amino acids arginine, leucine, and serine are each specified by six codons. Only methionine and tryptophan are represented by a single codon. Codons that specify the same amino acid are termed *synonyms*.

Table 1.3 summarizes the standard genetic code. While 61 triplets correspond to particular amino acids, 3 triplets indicate chain termination. The stop codons UAG, UAA, and UGA are referred to as amber, ochre, and opal, respectively. The codon AUG (and less frequently GUG) serves as the initiation codon of translation. However, the codons also specify the amino acid residues methionine and valine, respectively, at internal positions in the polypeptide chains.

The standard genetic code is (almost) universal, that is, it is used by bacteria as well as eukaryotes. This phenomenon forms the basis of genetic engineering. The genetic code not only facilitates research but also permits the production of human proteins in bacteria (Chapter 10). It came as a surprise when research in the early 1980s revealed that certain mitochondria and ciliated protozoa employ a slightly different genetic code.

Table 1.3 The genetic code.

First position (5′ end)	*Second position*				*Third position (3′ end)*
	U	*C*	*A*	*G*	
U	Phe	Ser	Tyr	Cys	U
	Phe	Ser	Tyr	Cys	C
	Leu	Ser	Stop	Stop	A
	Leu	Ser	Stop	Trp	G
C	Leu	Pro	His	Arg	U
	Leu	Pro	His	Arg	C
	Leu	Pro	Gln	Arg	A
	Leu	Pro	Gln	Arg	G
A	Ile	Thr	Asn	Ser	U
	Ile	Thr	Asn	Ser	C
	Ile	Thr	Lys	Arg	A
	Met	Thr	Lys	Arg	G
G	Val	Ala	Asp	Gly	U
	Val	Ala	Asp	Gly	C
	Val	Ala	Glu	Gly	A
	Val	Ala	Glu	Gly	G

The amino acids are given by their three-letter codes. The stop codons and the start codons encoding methionine and valine, respectively, are marked.

The code as depicted in the table is nonrandom. Changes in the first position tend to specify similar amino acids. Most synonyms differ only in their third nucleotide, which minimizes the deleterious effects of mutations. Many point mutations at the third codon position do not change the encoded amino acid.

While transcription and RNA processing take place in the nucleus, translation occurs in the cytoplasm (Figure 1.19). The primary sites of biological protein synthesis are ribosomes, which decode an mRNA and catalyze the formation of peptide bonds between amino acids that are delivered by tRNAs. Ribosomes are large and complex particles consisting of two subunits composed of protein and RNA molecules. The RNAs account for approximately two thirds of the total mass. The large subunit is mainly involved in mediating biochemical tasks such as catalyzing the peptidyl transferase reaction. The small subunit is responsible for recognition processes such as mRNA and tRNA binding. Messenger RNAs are read in the $5' \rightarrow 3'$ direction and encode the polypeptide, which is synthesized from the N-terminus to the C-terminus. For protein synthesis, the proper tRNA is selected only through codon–anticodon interactions, while the aminoacyl group does not participate in the process.

In 2000, the X-ray structure of the 50S ribosomal subunit at atomic resolution was published. Soon thereafter, the X-ray structure of the 30S ribosomal subunit was solved. For elucidating the structure of the ribosome, Ada Yonath, Thomas Steitz, and Venkatraman Ramakrishnan were awarded the Nobel Prize in Chemistry in 2009.

Their structure analysis revealed that the distribution of the proteins in the two ribosomal subunits is not uniform. The vast majority of the proteins are located on the surfaces of their subunits. The interface between the two subunits, particularly those regions that bind the tRNAs and mRNAs and form the new peptide bonds, is largely devoid of proteins. Thus, rRNA rather than the ribosomal proteins has the major functional role in ribosomal processes. Ribosomes have three functionally distinct tRNA-binding sites, known as A (aminoacyl) site, P (peptidyl) site, and E (exit) site (Figure 1.20). Ribosomes are tandemly arranged on mRNAs. The individual ribosomes in these *polyribosomes* (*polysomes*) are separated by gaps of 50–150 Å, producing a maximum density of ~1 ribosome per 80 mRNA nucleotides.

Protein biosynthesis requires a number of protein factors in addition to the ribosomal proteins. These soluble proteins can be grouped into three classes: initiation factors, elongation factors, and release factors (RFs).

Eukaryotic initiation is a complex process that involves at least 11 initiation factors (abbreviated as eIF*n*; e for

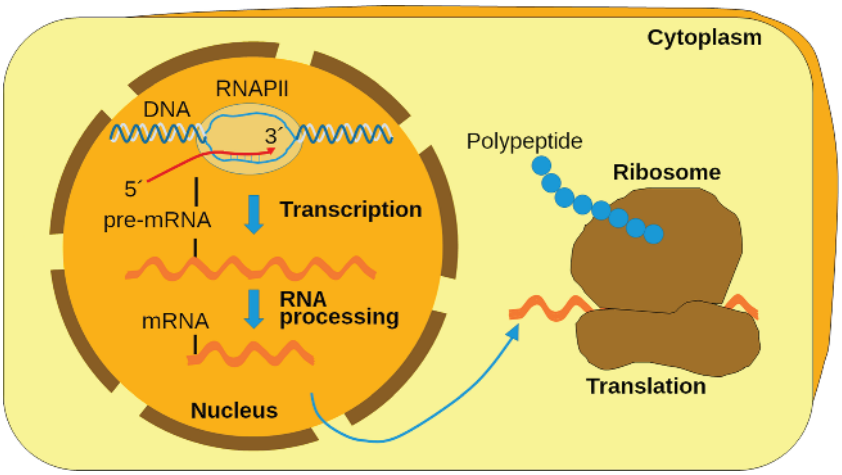

Fig. 1.19 Subcellular localization of transcription, RNA processing, and translation in a eukaryotic cell. RNAs are transcribed and processed in the nucleus. After their export to the cytoplasm, ribosomes translate the genetic information into proteins.

eukaryotic), including the Met-tRNA$_i^{Met}$ (initiator tRNA) that delivers a methionine that is never formylated in eukaryotes (as it is in the case of prokaryotes). Eukaryotic mRNAs also lack the Shine–Dalgarno sequence that is present in prokaryotic mRNAs, but nearly all eukaryotic mRNAs have the m^{7}G cap described above. Furthermore, they are almost always monocistronic, that is, they contain the genetic information for a single protein only, so that they initiate translation at their leading AUG. The AUG is embedded in the consensus sequence GCCRCCAUGG, known as the Kozak sequence (R denotes one of the purines adenine or guanine). Initiation factor eIF4E binds the cap by intercalating between two tryptophan residues. In addition, eIF4G interacts with the poly(A)-binding protein bound to the mRNA's poly(A) tail, circularizing the mRNA. Additional initiation factors subsequently bind to the eIF4F–mRNA complex.

The resulting complex then associates with the ribosomal 43S preinitiation complex. This complex translocates along the mRNA in an ATP-dependent process called *scanning* until it encounters the AUG initiation codon, thereby yielding the 48S initiation complex. Hydrolysis of GTP results in release of all the initiation factors. Finally, the large 60S subunit joins the mRNA-bound Met-tRNA$_i^{Met}$–40S subunit complex in a GTPase-driven reaction to yield the 80S ribosomal initiation complex.

Ribosomes elongate polypeptide chains in a three-stage reaction cycle that adds amino acid residues to a growing polypeptide's C-terminus: These three stages

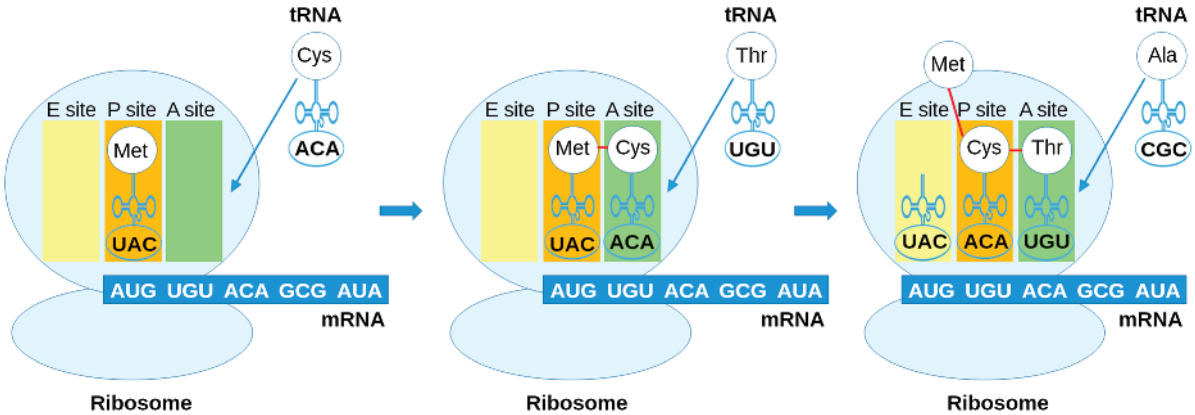

Fig. 1.20 Protein biosynthesis by ribosomes. Ribosomes elongate polypeptides in the process of translation. The ribosome harbors three functional sites, the E(xit) site, the P(eptidyl) site, and the A(minoacyl) site. tRNAs deliver the amino acids. The interaction between the codon of the mRNA and the anticodon of the tRNA determines the incorporation of the correct amino acid. A peptide bond is formed between the growing polypeptide chain and the incoming amino acid.

are decoding, transpeptidation (peptide bond formation), and translocation of tRNAs. The process occurs at a rate of 10–20 residues/s and requires several elongation factors.

In the decoding stage, the ribosome selects and binds an aminoacyl-tRNA whose anticodon is complementary to the mRNA codon at the A site (Figure 1.20). A complex of GTP and the elongation factor EF-Tu is associated with an aminoacyl-tRNA. The resulting ternary complex binds to the ribosome. GTP is hydrolyzed and while the aminoacyl-tRNA is bound to the A site, EF-Tu-GDP and P_i are released. The GDP in EF-Tu is replaced by GTP in a reaction mediated by the elongation factor EF-Ts, which functions as a guanosine exchange factor (GEF).

The next step of the elongation cycle is the formation of the peptide bond through the nucleophilic displacement of the P-site tRNA by the amino group of the 3′-linked aminoacyl-tRNA in the A site (Figure 1.20). The nascent polypeptide chain is thus lengthened by one residue at its C-terminus and transferred to the A-site tRNA.

In the final translocation stage of the elongation cycle, the now deacylated, uncharged P-site tRNA is transferred to the E site and expelled. Simultaneously, the peptidyl-tRNA in the A site, together with its bound mRNA, is moved to the P site. The translocation process requires the participation of the elongation factor EF-G, which is bound to the ribosome along with GTP, and is released only after the hydrolysis of GTP, which generates the free enthalpy for the translocation of the tRNA.

The three stop codons UAA, UGA, and UAG do not normally have corresponding tRNAs. Thus, they indicate that the newly synthesized polypeptide chain should be released. In eukaryotes, all three stop codons are recognized by a single *release factor*, the eukaryotic release factor 1 (eRF1). This protein functions as a tRNA mimic. Through its binding to the A site of the ribosome, eRF1 triggers the hydrolysis of the peptidyl-tRNA, resulting in the release of the peptide chain. eRF3, a second eukaryotic release factor, accelerates this process in a GTP-dependent manner. Finally, ribosome recycling occurs. This process is mediated by components of the initiation complex (eIF3, eIF1, and eIF1A).

Antibiotics are substances produced by bacteria or fungi that inhibit the growth of other organisms. They have many metabolic targets, but most antibiotics, including many used as drugs, block translation. Antibiotics are specific for bacterial ribosomes and bind to different regions of the ribosome, and block different ribosomal processes such as peptidyl transferase activity, translocation, and elongation. Antibiotic mechanisms of action are extensively discussed in Section 6.1.2.

Virtually all proteins are modified after their synthesis by the ribosome. Posttranslational protein processing may be reversible or irreversible. A common irreversible modification is the proteolytic cleavage of peptide bonds. The first step, occurring shortly after release of the newly synthesized protein from the ribosome, is the removal of the leading methionine. In addition, many proteins are synthesized as inactive precursors that are activated by limited proteolysis (examples are digestive enzymes such as trypsinogen, factors of the blood coagulation pathway, and insulin). The inactive proteins are called *proproteins* or *zymogens*.

In addition to activation by limited proteolysis, many proteins are processed by covalent modification of their amino acid residues. An example of this is the glycosylation of proteins; this type of modification promotes several processes, such as correct protein folding, enhancement of protein stability, and cell–cell adhesion. Glycosylation occurs in transmembrane proteins, lysosomal proteins, and proteins that are destined to be secreted. Proteins are initially synthesized with an N-terminal *signal peptide*, which is recognized by the signal recognition particle (SRP), a ribonucleoprotein. SRP mediates translocation of the nascent protein into the ER. Proteins bearing a signal peptide are called *preproteins* (or, if they also contain propeptides, as preproproteins). The signal peptide is cleaved off once the peptide chain is translocated through a protein pore known as a *translocon*.

Proteins may be modified either by *O*-glycosylation at serine or threonine residues or by *N*-glycosylation at asparagine residues. The latter is initiated in the ER and continued in the Golgi apparatus. In contrast, *O*-glycosylation occurs only in the Golgi apparatus. Based on the glycosylation pattern, the proteins are sorted for their final destination, which may be the plasma membrane, secretory vesicles, or lysosomes.

In addition to glycosylation, many other types of posttranslational modifications have been described. Among them are the acetylation of the N-terminus of the protein or of lysine or serine residues, hydroxylation, methylation, or the addition of myristic acid, a common saturated fatty acid. Another common posttranslational modification is the reversible phosphorylation of serine, threonine, and tyrosine residues. Phosphorylation status modulates the activity of many enzymes.

1.3.6 Protein Degradation

Proteins are continuously synthesized and degraded. This turnover eliminates abnormal proteins whose accumulation would be harmful to the cell; it also permits

the regulation of cellular metabolism by eliminating superfluous enzymes. The half-lives of enzymes vary greatly. While short-lived enzymes (e.g., ornithine decarboxylase and phosphoenolpyruvate (PEP) carboxylase) have half-lives of 0.2–5 h, long-lived enzymes (e.g., aldolase and lactate dehydrogenase) can have a half-life of more than 100 h.

Eukaryotic cells have two systems for protein degradation: a lysosomal mechanism and a cytosolic ATP-dependent system. As already described, lysosomes are membrane-encapsulated organelles that maintain an acidic internal pH and contain more than 50 hydrolytic enzymes, including a number of proteases collectively known as *cathepsins*. Lysosomal protein degradation is normally nonselective.

In addition to the lysosomal system, an ATP-dependent, targeted proteolytic system exists in the cytosol. The protein *ubiquitin*, consisting of 76 amino acid residues, is the marker for those proteins to be degraded. Ubiquitin, as its name suggests, is a ubiquitous, highly conserved abundant protein in eukaryotes. Proteins that are selected for degradation are initially marked by covalent linkage to ubiquitin. This process occurs in a three-step pathway (Figure 1.21a). At first, ubiquitin's terminal carboxyl group is conjugated to the *ubiquitin-activating enzyme* (E_1) in an ATP-depending reaction. Humans have only one type of E_1 protein. The ubiquitin is then transferred to one of the *ubiquitin-conjugating enzymes* (E_2). Mammals have about 20 E_2 proteins. Afterward, *ubiquitin-protein ligase* (E_3) transfers the activated ubiquitin from E_2 to an ε-amino group of a lysine residue of a target protein, forming an isopeptide bond. Cells contain many different E3 proteins, each of which mediates the ubiquitination of a specific set of proteins.

For efficient degradation, the target protein must be covalently linked to more than one ubiquitin. Typically, a chain of at least four tandemly linked ubiquitin molecules is attached to the target protein. Lys-48 of each ubiquitin peptide forms an isopeptide bond with the C-terminal carboxyl group of the next ubiquitin. The polyubiquitin chains can consist of 50 or more ubiquitin molecules.

A ubiquitinated protein is proteolytically degraded to short peptides in an ATP-dependent process. This process is mediated by a large multisubunit protein complex called the *26S proteasome* (Figure 1.21b). The 26S proteasome is composed of the 20S proteasome, which forms a hollow barrel and carries out the proteolysis of condemned proteins, and the 19S caps. The 19S cap consists of approximately 18 subunits. These caps may be attached to one or both ends of the 20S proteasome. The caps recognize ubiquitinated proteins, unfold them, and feed them into the 20S proteasome in an ATP-dependent manner.

1.4 Biological Communication

Intercellular communication is an essential prerequisite for the functioning of multicellular organisms. The

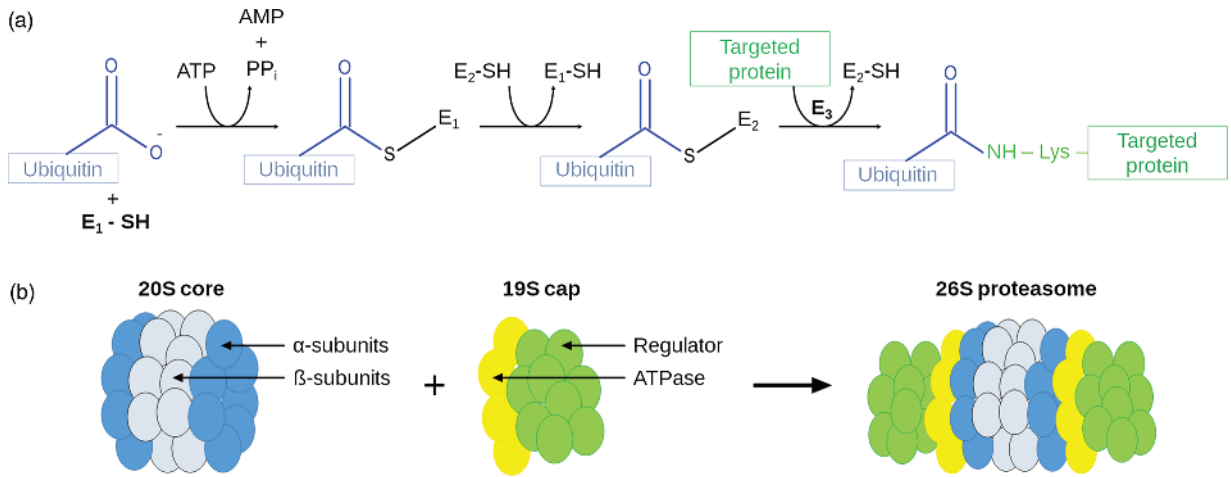

Fig. 1.21 Protein degradation by the proteasome. (a) Attachment of ubiquitin to a protein targeted for degradation. In the first step of the process, ubiquitin is covalently linked to the E_1 protein in a reaction driven by ATP hydrolysis. Subsequently, the activated ubiquitin is transferred to the sulfhydryl group of the E_2 protein. The final step, catalyzed by E_3, is the transfer of ubiquitin from E_2 to the target protein, which is then marked for degradation by the proteasome. (b) Structure of the proteasome. The proteasome consists of a central 20S barrel and 19S caps. Together, these components form the complete 26S proteasome. E_1: ubiquitin-activating enzyme; E_2: ubiquitin-conjugating enzyme; E_3: ubiquitin ligase.

following section will introduce the concepts of signal transmission between neurons, long-range and short-range intercellular signaling by hormones, and transduction of an extracellular signal across the cell membrane.

1.4.1 Neurotransmitters

The nervous system is characterized by a special mechanism that permits passage of a signal from a neuron to another cell, which may be another neuron, or an altogether different cell type. This occurs at a specific structure called a *chemical synapse* (Figure 1.22). When a nerve impulse in a presynaptic cell reaches the synapse, it triggers the fusion of synaptic vesicles, which contain neurotransmitters, with the presynaptic membrane. The neurotransmitter is then released via exocytosis into the synaptic cleft, and diffuses across it in less than 0.1 ms. The neurotransmitter then binds to specific receptors on the postsynaptic membrane, which trigger the continuation of the impulse in the postsynaptic cell. Due to the complexity of receptor signal transduction, chemical synapses can have many effects on the postsynaptic cell.

The known neurotransmitters are mainly classified into three groups (Table 1.4). The common transmitters glutamate, gamma aminobutyric acid (GABA), and glycine belong to the amino acid class. Serotonin, dopamine, or acetylcholine are biogenic amines. The third group includes neuropeptides, of which more than 50 are known today. In addition, nitrous oxide and ATP are involved in neuronal communication, but are not a member of any of the three groups.

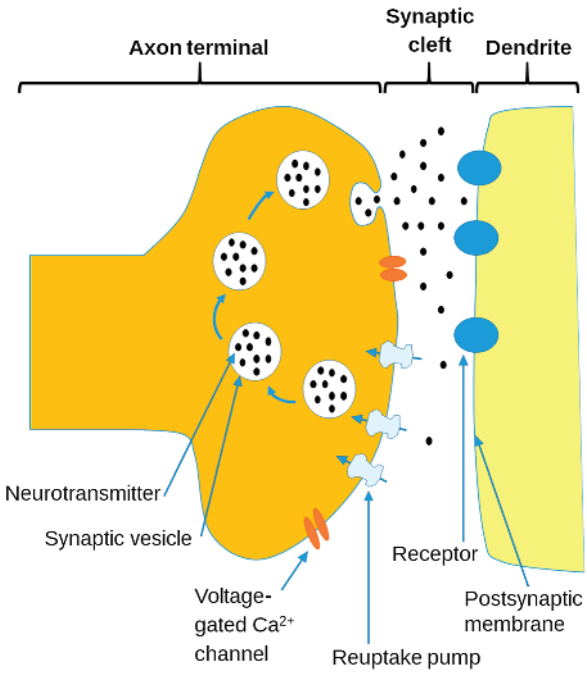

Fig. 1.22 Structure of a chemical synapse. At the synapse, the presynaptic neuron releases a neurotransmitter into the synaptic cleft. Binding of the neurotransmitter to receptors in the membrane of the postsynaptic neuron leads to its excitation (or inhibition).

Table 1.4 Characteristics of major neurotransmitters.

Class	*Neurotransmitter*	*Function*	*Secretion sites*
Biogenic amines	Acetylcholine	Excitatory for skeletal muscles; excitatory or inhibitory at other sites	CNS, PNS, neuromuscular junction
	Norepinephrine	Excitatory or inhibitory	CNS, PNS
	Dopamine	Generally excitatory	CNS, PNS
	Serotonin	Generally inhibitory	CNS
Amino acids	GABA	Inhibitory	CNS
	Glycine	Inhibitory	CNS
	Glutamate	Excitatory	CNS
	Aspartate	Excitatory	CNS
Neuropeptides	Substance P	Excitatory	CNS, PNS
	Met-enkephalin	Generally inhibitory	CNS

CNS: central nervous system; PNS: peripheral nervous system; GABA: gamma aminobutyric acid.

The duration of action of neurotransmitters is usually a few milliseconds; their effects are strictly restricted to the synaptic area. However, cells that are far away from the source of the neurotransmitter can be affected by diffusion. Neurotransmitter activity is terminated either by enzymatic cleavage or by uptake of the neurotransmitter into the presynaptic cell. Neurotransmitters can act in an inhibitory or excitatory manner depending on the postsynaptic receptor and the intracellular signal transduction pathway. Multiple types of neurotransmitters can coexist in a single neuron.

Physiologically important and well-studied examples of neurotransmitters are acetylcholine, serotonin, and dopamine. A special feature of these three transmitters is that they have a relatively limited origin, that is, they are only produced by certain defined groups of neurons. However, their influence extends to over 100 000 synapses per neuron in many different sites in the brain. In addition, acetylcholine, serotonin, and dopamine have a longer lasting effect than, for example, glutamate.

Therefore, they play an important role in the regulation of conditions such as sleep or mood and in the development of disease.

Because it plays a crucial role in the vegetative nervous system and is also active at the interface between motor neurons and skeletal muscles, acetylcholine was the first neurotransmitter discovered. Neurons that mainly use acetylcholine as neurotransmitter are referred to as cholinergic neurons. They are present in the brain and participate in the control of attention and in the excitability of the brain during the sleeping and waking circadian rhythm. In addition, they are believed to play a crucial role in plasticity and learning. The hippocampus, the neocortex, and the olfactory bulb are innervated from the cerebrum. Cells of these areas are among the first to die in Alzheimer's disease. Among the approved drugs that aim to delay the progression of Alzheimer's disease are those that slow the cleavage of acetylcholine in the brain.

In the brain, the neurotransmitter serotonin is detectable only in neurons whose cell bodies are located in the *raphe nuclei* in the brain stem. These neurons innervate virtually all regions of the brain and affect pain, sleep and wake rhythms, and mood. The raphe nuclei are very active in a state of heightened vigilance and have the lowest activity during sleep. Studies have shown that agitation and hallucinations may occur with excessive brain levels of serotonin. Serotonin deficiency can lead to depression, anxiety, and aggression.

Dopamine-containing cells are found in many areas in the central nervous system. One group of dopaminergic neurons in the *substantia nigra* of the midbrain is of particular medical relevance. These neurons send their axons into the striatum and are essential for the control of volitional movement. The degeneration of dopaminergic cells in the substantia nigra triggers the motor dysfunction seen in Parkinson's disease.

1.4.2 Hormones

Hormones mediate another important mechanism of intercellular signaling. They are typically secreted by one cell type and act on another, inducing physiological responses such as cell growth or the activation of metabolic processes. Hormones can be classified in different ways. One way is to group them according to the distance over which they act (Figure 1.23a). *Endocrine* hormones act over a long distance, that is, on cells far away from the site of their release. They are produced in

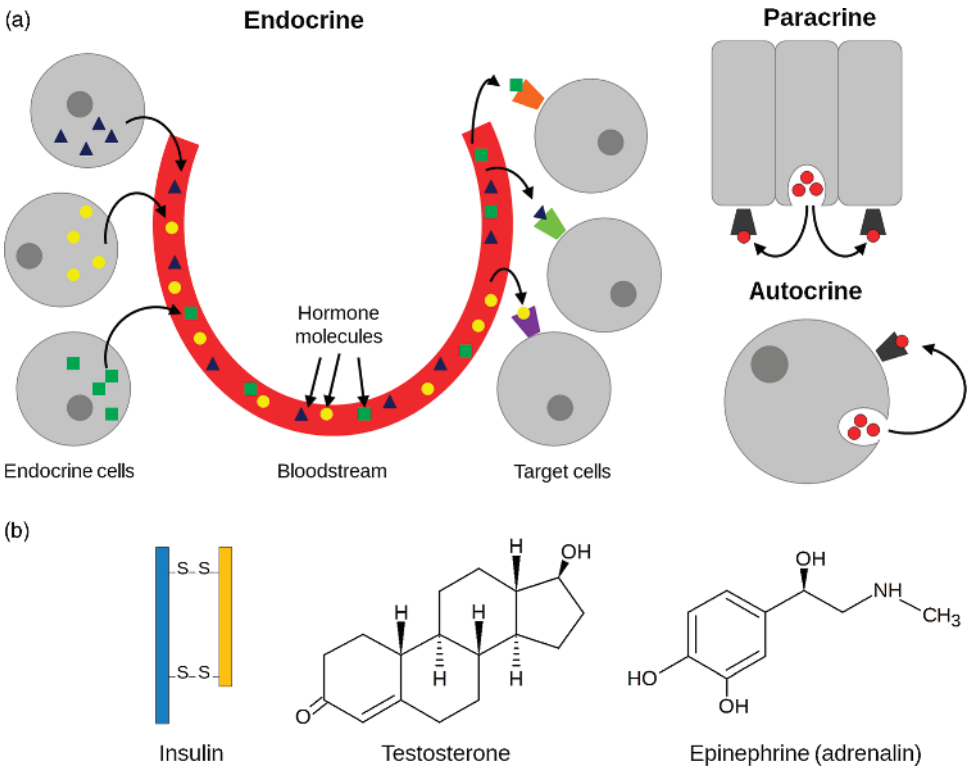

Fig. 1.23 Classification of hormones. (a) Hormones can be grouped into endocrine, paracrine, and autocrine hormones according to the distance over which they act. (b) Hormones belong to three different major chemical classes: peptides, steroid hormones (derived from cholesterol), and derivatives of aromatic amino acids. (Part (a) adapted from Ref. [5] with kind permission from John Wiley & Sons, Inc.)

specialized endocrine glands (e.g., the pancreas, adrenal glands, and the gonads) and released into the bloodstream. The endocrine system regulates many processes in the body. These include homeostasis (e.g., blood glucose levels, which are controlled by the hormones insulin and glucagon), the response to external stimuli (e.g., preparation for the fight-or-flight-response, regulated by epinephrine and norepinephrine), and various cyclic and developmental programs. *Paracrine* hormones act only on cells close to the cell that releases them. They are secreted into the immediate extracellular environment and diffuse over relatively short distances. An example is the release of interleukin-12 by macrophages, which stimulates other local immune cells. In some cases, hormones bind to receptors on the cells from which they were released. This is called *autocrine* signaling. Examples of autocrine signaling are the autoactivation of immune cells and the development of tumors through the autocrine production of growth and survival factors.

Alternatively, hormones can be classified by their basic chemical structure (Figure 1.23b). The major groups are peptide hormones (e.g., insulin), steroid hormones (testosterone), and amino acid derivatives (epinephrine and norepinephrine belong to the catecholamine subclass). Only cells with a specific receptor will respond to a given hormone.

1.4.3 Signal Transduction

The message of hormones and other extracellular signal molecules must be transmitted into the cell. With the exception of some nonpolar steroid hormones that can easily pass through the cell membrane and bind to their receptor in the cytoplasm, most signaling molecules are hydrophilic and cannot easily reach the interior of the cell. Instead, they bind to a receptor protein on the cell surface that is specific for that hormone (or other type of ligand). The binding event is transmitted to the cell interior and a series of intracellular signaling events is induced. These responses usually involve the generation of a *second messenger* and protein phosphorylation/dephosphorylation events. Signal intensity is often amplified by a signal cascade.

Signaling by G protein-coupled receptors (GPCRs) resulting in activation of adenylyl cyclase (AC) will be outlined as a paradigm for the various signal transduction pathways (Figure 1.24). A common characteristic of the more than 800 GPCRs encoded in the human genome is that they contain seven transmembrane helices. Binding of a specific ligand results in a conformational change of the GPCR, which transmits the extracellular signaling event across the plasma membrane to the cytoplasmic face. The intracellular domain of the GPCR is coupled to a heterotrimeric G protein, which consists of a complex of one α-, β-, and γ-subunit each. In the nonactivated state, the α-subunit binds GDP. Activation of the G protein by the GPCR in complex with its ligand induces the exchange of bound GDP for GTP. When GTP is bound, the α-subunit dissociates from the β- and γ-subunits and stimulates adenylyl cyclase. α-Subunit activation is temporary, since it functions as a GTPase and hydrolyzes GTP into GDP + P_i. The inactivated α-subunit carrying GDP then reassembles with the β- and γ-subunits. Inhibition of GTPase activity is the way cholera toxin leads to an overactivation of the signaling cascade. In the intestine, this causes the failure of water resorption and voluminous diarrhea (Section 6.1.1, Box 6.1).

The activated α-subunit carrying GTP stimulates AC to convert ATP into cyclic AMP (cAMP) (Figure 1.25). cAMP acts as a second messenger, that is, as a molecule that transmits the signal intracellularly to the target molecule. cAMP can freely diffuse through the cytoplasm and bind to its target protein kinase A (PKA). In the absence of cAMP, PKA forms an inactive heterotetramer consisting of two catalytic and two regulatory subunits (R_2C_2 in Figure 1.24). Binding of cAMP causes dissociation of the complex so that the catalytic subunit can phosphorylate its substrate proteins to produce a cellular response.

To terminate the signal, the signaling molecule must eventually be eliminated. In the case of cAMP and the related second messenger cGMP, signaling activity is terminated through the action of phosphodiesterases (PDEs) that hydrolyze the cyclic phosphodiesters to AMP and GMP, respectively. In mammals, phosphodiesterases are grouped into 11 families, called PDE1–PDE11. Some PDEs are specific for either cAMP or cGMP, while others can hydrolyze both second messengers. Due to their unique tissue distribution and functional properties, PDEs are attractive drug targets. The drug sidenafil (Viagra, Box 1.1) is an example of a PDE inhibitor. The stimulatory activity of caffeine is (in part) caused by nonselective inhibition of PDEs. In addition, caffeine antagonizes adenosine receptors that act through inhibitory G proteins. Both effects contribute to an increased concentration of stimulatory cAMP in the cell.

An alternative signaling pathway starts with the binding of a ligand, for example, a growth hormone, to a receptor tyrosine kinase (RTK). In contrast to GPCR, RTKs contain only a single-transmembrane segment. Binding of the ligand induces dimerization of the

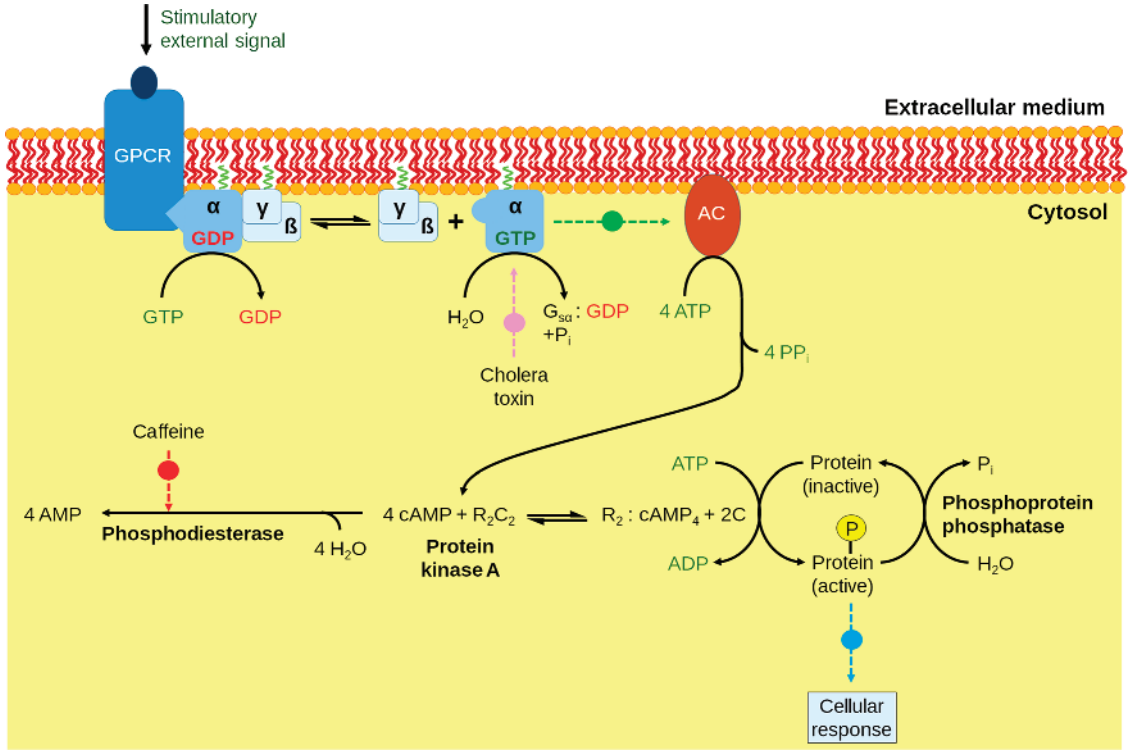

Fig. 1.24 G protein-coupled receptors and the adenylyl cyclase (AC) signaling system. Binding of the ligand to the GPCR causes exchange of GDP for GTP in the α-subunit of the protein, which then dissociates from the β- and γ-subunits and activates adenylyl cyclase. AC generates cAMP from ATP; cAMP activates protein kinase A to phosphorylate its substrate proteins, causing the cellular response. The signaling system depicted here results in activation of AC, but other hormones can bind to inhibitory GPCRs that trigger an almost identical pathway, except that the α-subunits inhibit AC. R_2C_2: the two regulatory and two catalytic subunits of protein kinase A. (Adapted from Ref. [1] with kind permission from John Wiley & Sons, Inc.)

receptor proteins, which brings the intracellular domains close together so that they can phosphorylate each other. This event activates a pathway that is known as the Ras–Raf signaling cascade. A central step in this pathway is the exchange of GDP for GTP by the G protein named Ras. This activates Ras, which then binds to Raf and induces a cascade of several protein phosphorylation events that finally activates various transcription factors. Mutations in Ras are among the most frequent changes in malignant transformation and can lead to uncontrolled cell growth.

The third signaling pathway, one that mediates the effects of a variety of hormones, is the phosphoinositide (PI) pathway. This pathway is initiated by binding of a ligand to a cell surface receptor, followed by activation of a heterotrimeric G protein, as described above. In the PI pathway, the α-subunit of the G protein in complex with GTP activates phospholipase C

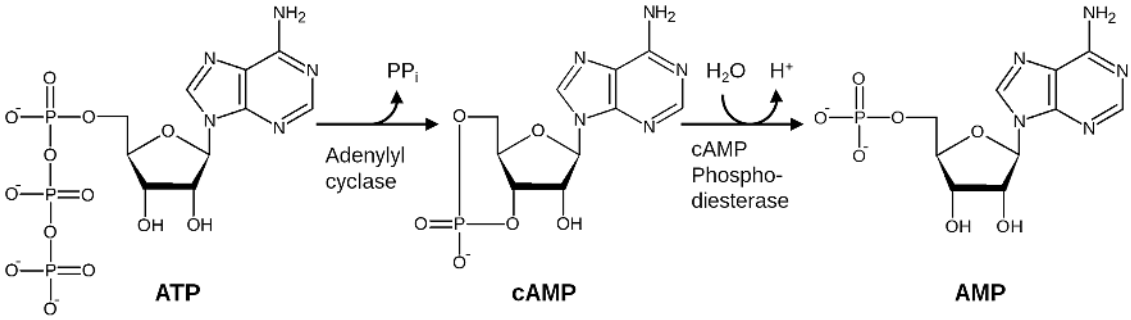

Fig. 1.25 Formation and hydrolysis of cyclic AMP (cAMP). Adenylate cyclase converts ATP into the second messenger cAMP. To terminate the cAMP activity, phosphodiesterases hydrolyze it into AMP.

Box 1.1. The Discovery of Viagra™

The discovery of Viagra as a drug to treat erectile dysfunction was accidental. Its active pharmaceutical ingredient, sidenafil, selectively inhibits the cGMP phosphodiesterase PDE5. Sidenafil was intended as a treatment for *angina pectoris* as it relaxes the arterial wall, which leads to decreased blood pressure. However, the outcome of early clinical trials was disappointing and the trials were terminated. According to the myth, the male participants in the trials refused to return the remaining pills because they claimed they were having firmer and longer lasting erections. In fact, sexual stimulation in males causes penile nerves to release nitric oxide, which activates guanylate cyclase to produce cGMP from GTP. cGMP induces relaxation of the vascular smooth muscle (SM) and blood inflow increases, leading to an erection. Hydrolysis of cGMP by PDE5 interferes with this process; thus, inhibition of PDE5 by sidenafil improves erectile function. Due to this entirely unintended consequence, Viagra has become one of the most frequently prescribed drugs in the world.

(PLC), which then hydrolyzes phosphatidylinositol-4,5-bisphosphate (PIP_2), a phospholipid found on the inner face of the plasma membrane, into the second messengers inositol-1,4,5-trisphosphate (IP_3) and 1,2-diacylglycerol (DAG). While the latter activates protein kinase C, IP_3 stimulates the release of calcium ions from the ER. Ca^{2+} then acts as another second messenger and modulates the activity of a number of proteins to induce a cellular response.

1.5 The Immune System

Animals (as well as virtually all other living organisms) are continually subject to attack by disease-causing pathogens. As a countermeasure, animals have evolved an immune system to protect themselves. The immune system must detect and destroy a wide variety of infectious agents, including viruses, microorganisms, and parasitic worms (Chapters 5 and 6), distinguishing them from the organism's own cells. More recently, it has become evident that the immune system not only fights invading pathogens but also provides an important defense against the development of cancer. It does this by recognizing and destroying transformed cells.

The cells of the immune system are distributed throughout the body and in the bloodstream. Blood cells fall into three different categories: (i) red blood cells (erythrocytes) primarily carry oxygen and carbon dioxide; (ii) platelets (thrombocytes) are involved in the formation of blood clots; and (iii) white blood cells (leukocytes) are the cells of the immune system. The formation of different blood cells, a process called hematopoiesis, is illustrated in Figure 1.26.

Immune cells can be further categorized as being part of the *innate* or *adaptive immune system* (Figure 1.27). The innate immune system is the first line of defense. It recognizes general characteristics of pathogens and provides a rapid, nonspecific response. In addition, it activates a second layer of response, the adaptive immune system, which provides a specific defense against a given target. Two types of adaptive immunity can be distinguished: cellular and the humoral immunity. This is outlined in Section 1.5.2.

1.5.1 The Innate Immune System

The innate immune system detects macromolecules that represent a danger to the organism and generates the means to destroy them. It provides an immediate defense against infection, but does not confer specific, long-lasting, protective host immunity. The innate immune system does not identify a specific microbe. Rather, it recognizes general patterns that are known as pathogen-associated molecular patterns (PAMPs). The characteristics of invading pathogens that can distinguish them from host cells include components of the outer bacterial membrane (lipopolysaccharides) and viral double-stranded RNA molecules. PAMPs act as agonists for pattern recognition receptors (PRRs) on innate immune cells.

One of the first responses of the innate immune system is the stimulation of inflammation. Acute inflammation is initiated by cells present in the tissue; these cells include macrophages, dendritic cells, and mastocytes. Upon activation, these cells release inflammatory mediators (e.g., histamine, bradykinin, and/or cytokines) that establish a physical barrier against the spread of infection and trigger additional immune system effectors.

Different types of specialized phagocytes destroy harmful particles, such as bacteria, small parasites, fungi, and viruses, by internalization (phagocytosis,

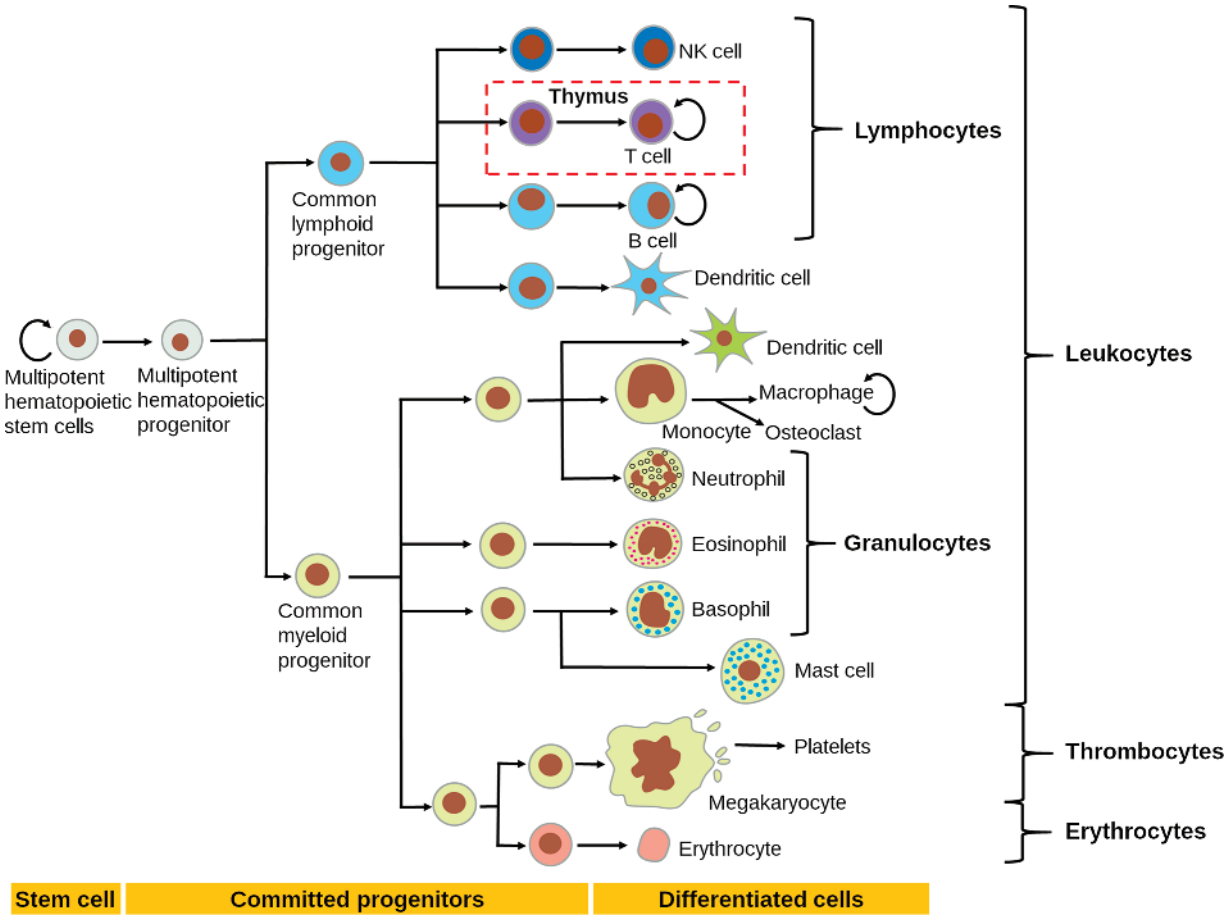

Fig. 1.26 Hematopoiesis. Multipotent hematopoietic stem cells differentiate into committed progenitors and then develop into the terminally differentiated cells. In adult mammals, the cells shown develop primarily in the bone marrow. Exceptions are T lymphocytes, which differentiate in the thymus, and macrophages and osteoclasts that develop from blood monocytes. Dendritic cells may also derive from monocytes. NK cell: natural killer cell. (Reproduced with permission from Ref. [2]. Copyright 2008, Garland Science/Taylor & Francis LLC).

Figure 1.28). Macrophages and neutrophils are the most important phagocytes in the defense against infection. The process of phagocytosis starts with the recognition of the microorganism by PRRs. The invading microorganism is then internalized into the so-called phagosome, a membrane-bound vacuole, which fuses with lysosomes. Lysosomes contain digestive enzymes in an acidic environment that degrade and kill the microorganism.

Antigen presentation is an important process of the adaptive immune system. Some phagocytes (macrophages and dendritic cells) move molecules from engulfed pathogens that have been degraded by the proteasomes (described in Section 1.3.6) back to the surface of their cells. These molecules are then presented to other cells of the immune system. Such cells are known as antigen-presenting cells (APCs). APCs break down foreign proteins to peptides in the proteasome. The peptides formed are then bound to the so-called major histocompatibility complex (MHC), which carries them to the surface and presents them to lymphocytes. In humans, MHC is also called human leukocyte antigen (HLA). HLA is an important determinant of the compatibility of donors for organ transplants. Cells express different types of MHC molecules: MHC class 1 is expressed on almost all nucleated cells. The expression of MHC class II is restricted to special cells, including thymic epithelial cells, dendritic cells, B cells, and some macrophages. By activating lymphocytes, APCs connect the innate and the adaptive immune systems.

1.5.1.1 The Complement System

The complement system is another part of the immune system that helps to clear pathogens from an infected organism. Since it does not adapt to a specific pathogen

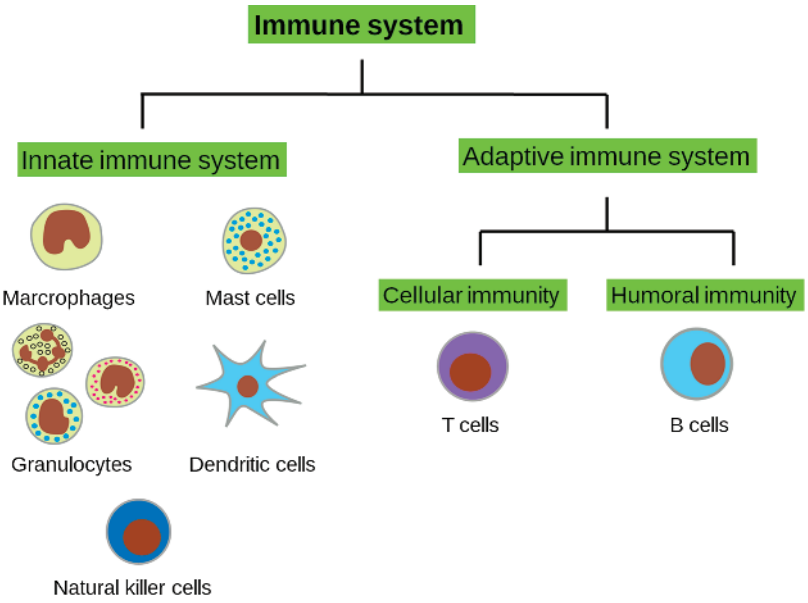

Fig. 1.27 Cells of the immune system. The cells of the innate immune system include macrophages that digest pathogens, mast cells that release histamine and other substances that induce inflammation, granulocytes (neutrophils, eosinophils, and basophils) that release toxic substances and promote inflammation, dendritic cells that present antigens and induce the adaptive immune system, and natural killer cells that destroy infected cells (as well as cancer cells). The key components of the adaptive immune system are T lymphocytes (or T cells) that mediate cellular immunity and B lymphocytes (or B cells) that constitute humoral immunity.

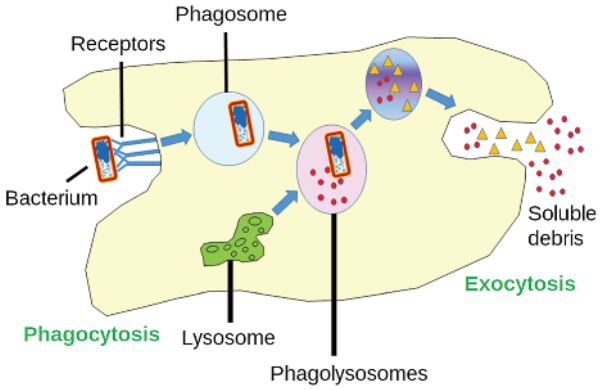

Fig. 1.28 Phagocytosis. Bacteria are recognized by special pattern recognition receptors (PRRs) on the surface of phagocytes. The bacterium is taken up by a process called phagocytosis, and is then trapped in a compartment known as phagosome, which fuses with a lysosome. Several enzymes then destroy the pathogen; the soluble debris is released by exocytosis.

and does not change over the lifetime of an individual, the complement system is considered part of the innate immune system. However, it can be activated by the adaptive immune system. The complement system consists of more than 25 proteins, most of which circulate in the blood as inactive precursors. The major components of the complement cascade are named with the letter "C" followed by a number. Cleavage products are assigned as "a" and "b," with "b" being the *b*igger fragment.

The complement cascade can be activated by three pathways: the classical, lectin, and alternative pathways (Figure 1.29a). The classical pathway is triggered by activation of the C1 complex, which consists of the C1q protein and the proteases C1r and C1s. C1q binds to antibody–antigen complexes on the surface of pathogens, but can also be activated by direct binding to the surface of bacteria. The C1 complex then splits C2 and C4, and the cleavage products C4b and C2a form the C3 convertase that promotes cleavage of C3 into C3a and C3b.

The lectin pathway functions in an analogous manner, except that the initial step consists of activation of mannose-binding lectin (MBL) by mannose residues on the pathogen surface. This activates MBL-associated serine proteases (MASPs) followed by cleavage of C2 and C4 and formation of the C3 convertase, as in the classical pathway.

The alternative pathway is continuously activated at a low level due to spontaneous hydrolysis of C3. The cleavage product C3b is rapidly inactivated by the two factors H and I. However, binding of C3b to the surface of a pathogen protects the C3b from inactivation and leads to the activation of factors B and D, resulting in the formation of an alternative C3 convertase.

Production of the cleavage product C3b by any of the three pathways mediates the main functions of the complement system. These include coating of microbes and

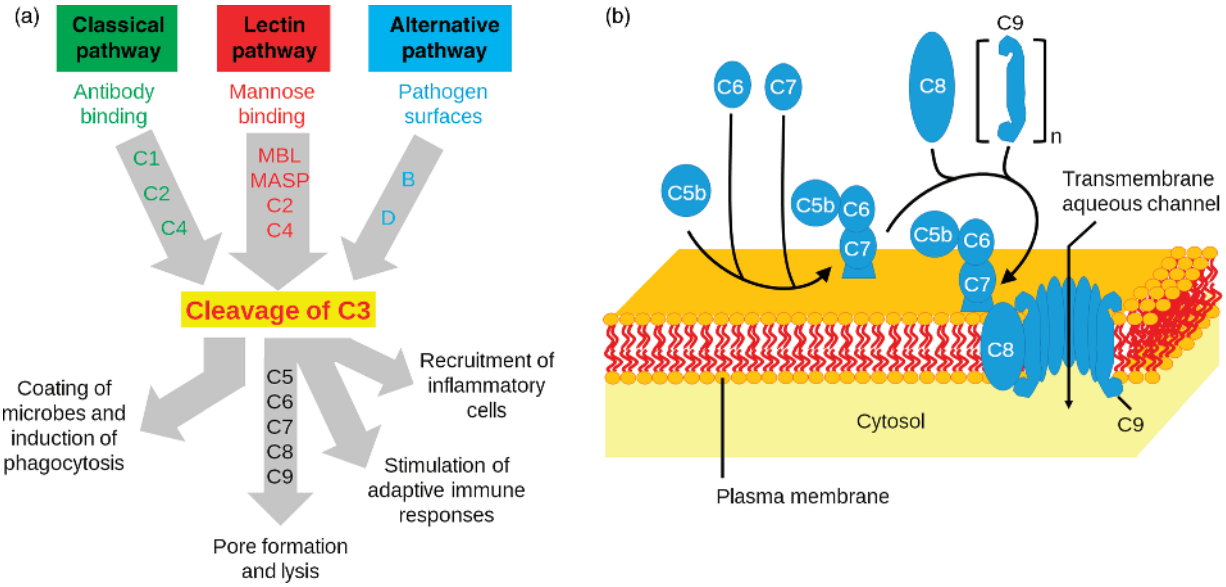

Fig. 1.29 Complement system. (a) The complement system can be activated by the classical, lectin, and alternative pathways. A central step is the activation of C3. Activation of the complement system results in the killing of a foreign cell, stimulation of phagocytosis, triggering of inflammation, and adaptive immune responses. (b) The late components C5b–C9 form a membrane attack complex that leads to osmotic lysis of the microbe or infected host cell. MBL: mannose-binding lectin; MASP: MBL-associated serine protease. (Reproduced with permission from Ref. [2]. Copyright 2008, Garland Science/Taylor & Francis LLC.)

induction of phagocytosis, recruitment of inflammatory cells, and stimulation of the adaptive immune response. In addition, the late components C5–C9 form a pore in the target cell's plasma membrane, the so-called membrane attack complex (MAC, Figure 1.29b). This channel allows free ingress of water and egress of electrolytes, which causes osmotic lysis of the target cell.

1.5.2 The Adaptive Immune System

The adaptive immune response is primarily based on two types of lymphocytes: B cells confer humoral immunity (*humor* is an archaic term for fluid) by producing antibodies (known as immunoglobulins), whereas T cells mediate cellular immunity. The adaptive immune response is triggered by antigens (*anti*body *gen*erators), macromolecules (in most cases proteins or carbohydrates) that are recognized as foreign.

1.5.2.1 Cellular Immunity

The main function of the cellular immune system is to prevent the spread of viral infection by killing virus-infected cells. In addition, the cellular immune system is effective against intracellular bacteria and parasites, where they are protected from attack by antibodies. The cellular immune system is also effective against certain types of cancer. The cells of the cellular immune system mature in the thymus and are referred to as T cells. The major populations of T cells are helper T cells (T_H), cytotoxic T cells (T_C), memory T cells, and natural killer T cells (not to be confused with the natural killer cells of the innate immune system). A common characteristic of T cells is the presence of a T-cell receptor (TCR) on the cell surface. This receptor is composed of two chains that form the antigen recognition site. The best understood populations are the αβ T cells, in which the TCR is composed of an α and a β chain. TCRs possess unique antigen specificity, which is determined by the structure of the antigen-binding site formed by the two chains. The diversity of TCRs is based mainly on somatic recombination, as described for immunoglobulins in the following section. The T-cell population is further categorized by the additional proteins found on their surface. The *cluster of differentiation* (CD) system is used for immunophenotyping cells according to cell surface markers. Conventional αβ T cells develop into two classes expressing either the CD4 or the CD8 glycoprotein.

TCRs, together with the respective CD proteins, bind to the MHC proteins displaying antigenic fragments on the surface of APCs, as described in the previous section. T cells expressing CD4 (abbreviated as $CD4^+$ T cells) interact with MHC class II molecules, while $CD8^+$ cells interact with MHC class I molecules (Figure 1.30). Binding of a T cell to an APC displaying an antigen–MHC complex causes it to reproduce, a process called

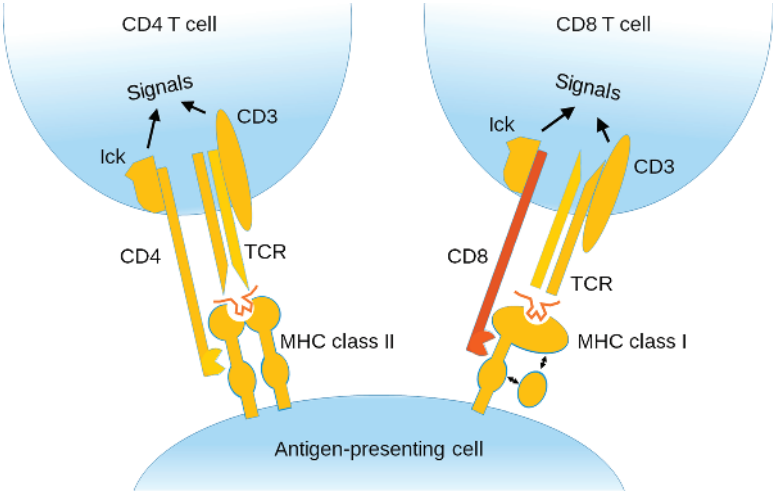

Fig. 1.30 Interaction between T cells and APCs. T-cell receptors (TCRs) and CD4 or CD8 molecules on the surface of T cells interact with proteins of the major histocompatibility complex (MHC) that present a peptide processed from an invading pathogen. CD4 interacts with MHC class II; CD8 binds to MHC class I. CD4 and CD8 are associated with additional factors that mediate signaling to activate the T cell. (Adapted from Ref. [6] with kind permission from Wiley-VCH Verlag GmbH.)

clonal selection. As a consequence, the T cells are produced in large quantity after recognizing an invading pathogen.

CD4 is expressed on T helper cells (T_h cells), which assist other lymphocytes by releasing cytokines. These cytokines induce the maturation of B cells into plasma cells, which release antibodies and activate memory B cells, cytotoxic T cells, and macrophages (Figure 1.31). Upon activation, T_h cells proliferate and differentiate into several subtypes, the two most important of which are known as T_h1 and T_h2 cells. These two types of helper cells are activated by different partners and release different kinds of cytokines that have specific effects. T_h1 cells are important for the immune reaction against intracellular bacteria and protozoa, while T_h2 cells mediate the immune reaction against helminths (parasitic worms). As will be outlined in Section 5.1.1, the human immunodeficiency virus (HIV) uses CD4 as a cellular entry receptor. At a later stage of the infection, massive destruction of $CD4^+$ cells leads to the development of the acquired immunodeficiency syndrome (AIDS).

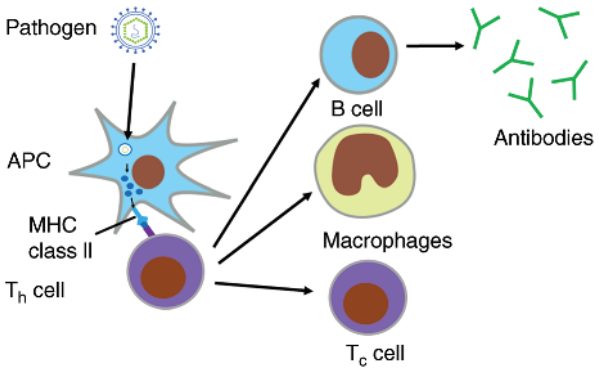

Fig. 1.31 The function of T helper cells (T_h cells). T_h cells are activated by binding to antigen–MHC class II complexes on the surface of APCs. By releasing stimulatory signals (cytokines), they induce maturation of B cells and activation of macrophages and cytotoxic T cells.

Cytotoxic T cells (T_C cells) express the CD8 glycoprotein. They destroy cells infected with a virus (or other pathogens) and tumor cells. After recognizing a specific antigen–MHC class I complex, the activated T cell releases cytotoxins, such as perforin, which form pores in the target cell's plasma membrane. This allows water and another toxin called granulysin (a protease) to enter the target cell and kill it.

Some T cells, known as memory T cells, persist for long term after an infection has resolved. They constitute an immune system "memory" of past infections and quickly expand into large numbers of effector T cells upon re-exposure to their cognate antigen. They induce a faster, stronger immune response than the first time the immune system encountered the pathogen.

1.5.2.2 Humoral Immunity

The humoral immune response is largely mediated by B cells, which mature in the bone marrow. The principal function of the humoral immune system is the production of antibodies against soluble antigens. B-cell precursors display immunoglobulins on their surface (Figure 1.32). They can be activated in a T cell-dependent or T cell-independent manner. The latter process is

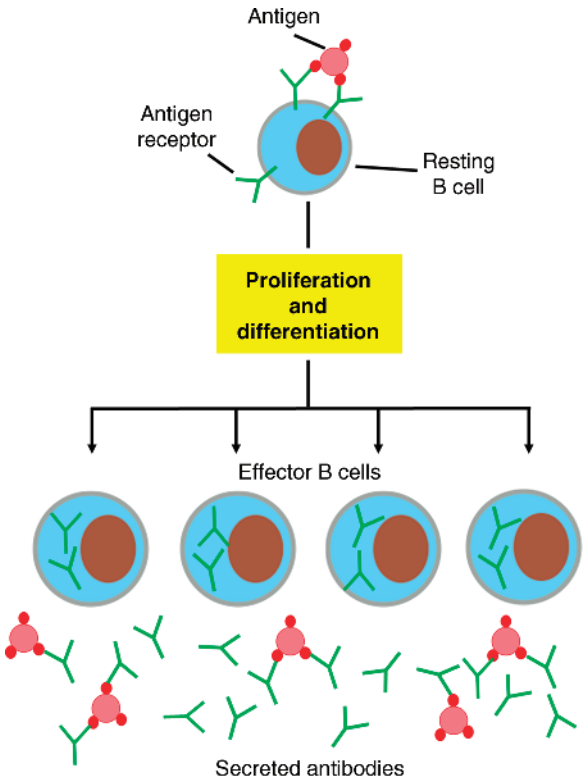

Fig. 1.32 B-cell activation. Activation of a B cell includes recognition of the antigen by the antigen receptor on the surface of the B cell. The activation may be T cell-dependent or T cell-independent. The activated B cell proliferates and differentiates to a plasma cell that secretes antibodies. (Reproduced with permission from Ref. [2]. Copyright 2008, Garland Science/Taylor & Francis LLC.)

triggered when a B cell encounters its matching antigen. The antigen–antibody complex is taken up by the B cell and digested into peptides, which are bound by a MHC class II molecule and then displayed on the cell surface. In doing so, B cells perform the role of APCs. The antigen–MHC complex is recognized by a T cell, which releases cytokines that activate the B cell. B cell activation includes a combination of clonal proliferation and differentiation into effector B cells known as plasma cells. Plasma cells produce and secrete millions of copies of the antibody that recognized the antigen. Although most B-cell progeny are plasma cells, others develop into memory B cells that have a similar function as memory T cells.

Antibodies are large, Y-shaped glycoproteins. In mammals, there are five different classes of antibodies: IgA, IgD, IgE, IgG, and IgM (Ig stands for immunoglobulin). While the monomeric structures of all the isotypes are similar, two classes can form multimers: IgMs are pentamers and IgAs are dimers consisting of linked monomeric units. The five isotypes differ in their biological properties, functional locations, and ability to deal with different antigens. The antibody isotype of a B cell changes during the differentiation process (*immunoglobulin class switching* or *isotype switching*). This switch occurs by exchanging the constant region of the heavy chain while retaining the variable region that determines antigen specificity. Immature B cells express IgM in a cell surface-bound form. At a later stage, these cells express surface IgD in addition to IgM. When activated to plasma cells, B cells produce antibodies in a secreted rather than membrane-bound form. Secreted IgMs are important for the rapid elimination of pathogens in an early stage of the B-cell response, before there is sufficient IgG. In the final step of isotype switching, cells change from producing IgD or IgM to the production of the IgE, IgA, or IgG classes of immunoglobulin.

IgG is the most abundant class of antibody in the blood, providing most of the antibody-based immunity against invading pathogens. Its general structure is depicted in Figure 1.33. IgGs protect the body from infection by several mechanisms. The antibody may bind to the pathogen and cause its immobilization and binding together via agglutination, preventing its entry into

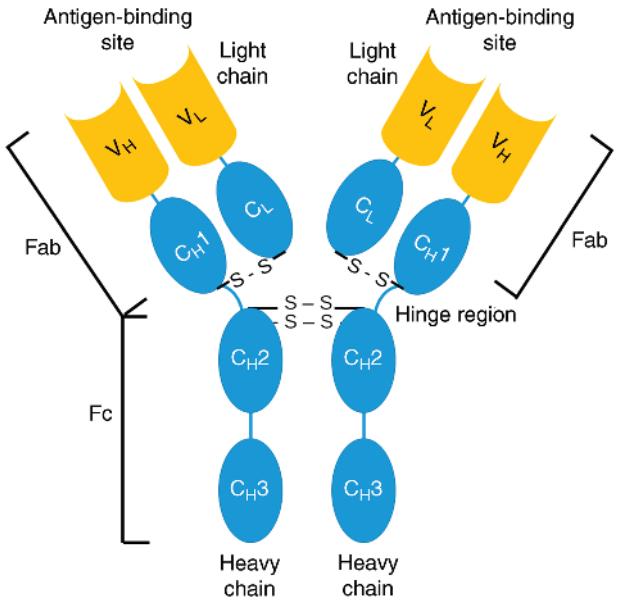

Fig. 1.33 Structure of immunoglobulin G (IgG). An antibody is made up of two identical heavy and two identical light chains connected by disulfide bonds. Each chain is composed of several immunoglobulin domains. The heavy chain has one variable (V_H) and three constant regions (C_H1–C_H3). The light chain is only composed of one variable (V_L) and one constant chain (C_L). The arms of the Y are known as the Fab antigen-binding fragment and are composed of one constant and one variable domain from each heavy and light chain of the antibody. The two variable domains form the site for antigen binding. The base of the Y is called the F_c for crystallizable fragment. It plays an important role in modulating immune cell activity.

host cells. This is also the way IgG neutralizes toxins. An additional defense mechanism mediated by IgG is *opsonization*; here, the IgG coats the surface of the pathogen, permitting its recognition and ingestion by phagocytic immune cells. IgGs also activate the classical pathway of the complement system (Section 1.5.1) and induce antibody-dependent cellular cytotoxicity (ADCC). In this process, the binding of antibodies to surface antigens marks the target cell for lysis by an effector cell of the immune system (e.g., an NK cell).

The human immune system can produce antibodies against virtually any antigen it encounters. Estimates suggest that humans produce more than 10^7 different antibodies (some estimates reach an order of 10^{12}). As the human genome encodes less than 25 000 genes (Chapter 7), each antibody cannot be encoded by a specific gene. Instead, antibody diversity is achieved by two mechanisms: *somatic recombination* and *hypermutation*. Immunoglobulin genes exist as discrete groups of gene segments. Somatic rearrangements are also known as V(D)J recombination. The variable region of each heavy or light chain is encoded by several gene segments. Heavy chains are composed of variable (V), diversity (D), and joining (J) segments; light chains consist of only V and J segments. Multiple copies of each segment in a tandem arrangement exist in the genome. During B-cell differentiation, intrachromosomal recombination occurs, and an immunoglobulin variable region is assembled by randomly combining one V, one D, and one J segment for the heavy chain, or one V and one J segment for the light chain. As the multiple copies of each type of gene segment can be combined in different ways, this process allows the generation of a large number of different antibodies.

An additional layer of diversity is achieved by two types of somatic mutations: (1) During the recombinant joining of the different segments, a terminal deoxynucleotidyl transferase randomly adds up to 30 base pairs that increase variability. (2) The variable regions mutate at rates that are at least a million fold higher than the rates of spontaneous mutation in other genes. This somatic hypermutability is presumably mediated by error-prone DNA polymerases.

The enormous diversity of antibodies has attracted the interest of researchers, who have used them for technical and medical purposes. As will be outlined in detail in Section 10.2.1, monoclonal antibodies are among the most successful classes of modern biomolecular drugs.

This brief introduction does not adequately reflect the enormous complexity of the immune system, which would be beyond the scope of any introduction to molecular medicine. The interested reader is referred to a more in-depth introduction given in the textbook cited in the reference list at the end of this chapter.

References

1. Voet, D., Voet, J.G., and Pratt, C.W. (2013) *Principles of Biochemistry – International Student Version*, 4th edn, John Wiley & Sons, Inc., Hoboken, NJ.
2. Alberts, B., Johnson, A., Walter, P. *et al.* (2007) *Molecular Biology of the Cell*, 5th edn, Taylor & Francis, New York.
3. Igney, F.H. and Krammer, P.H. (2002) Death and anti-death: tumour resistance to apoptosis. *Nat. Rev. Cancer*, **2**, 277–288.
4. Turner, B.M. (2002). Cellular memory and the histone code. *Cell*, **111**, 285–291.
5. Voet, D. and Voet, J.G. (2004). *Biochemistry*, 3rd edn, John Wiley & Sons, Inc., Hoboken, NJ.
6. MacPherson, G. and Austyn, J. (2012). *Exploring Immunology – Concepts and Evidence*, Wiley-VCH Verlag GmbH, Weinheim.

Further Reading

Biochemistry, Molecular, and Cell Biology

Alberts, B., Johnson, A., Walter, P. *et al.* (2007) *Molecular Biology of the Cell*, 5th edn, Taylor & Francis, New York.

Portela, A. and Esteller, M. (2010) Epigenetic modifications and human disease. *Nat. Biotechnol.*, **28**, 1057–1068.

Snustad, D.P. (2011) *Genetics – 6th International Student Edition*, John Wiley & Sons, Inc., Hoboken.

Voet, D. and Voet, J.G. (2011) *Biochemistry*, 4th edn, John Wiley & Sons, Inc., Hoboken, NJ.

Immune System

MacPherson, G. and Austyn, J. (2012) *Exploring Immunology – Concepts and Evidence*, Wiley-VCH Verlag GmbH, Weinheim.