COMENIUS UNIVERSITY
in BRATISLAVA

JESSENIUS FACULTY of MEDICINE
in MARTIN

PETER RAČAY

Selected chapters from enzymology, membrane biochemistry and biochemistry of genetic information

Textbook

2013

REWIEVERS: Assoc. Prof. Erika Halašová, RNDr., PhD.
Assoc. Prof. Jana Muchová, RNDr., PhD.

ISBN: 978-80-89544-46-2

CONTENT

Introduction

This textbook is intended for international students of general medicine studying at Jessenius Faculty of Medicine in Martin. It follows the textbooks Medical Chemistry and Biochemistry I (written by prof. Kaplan), Medical Chemistry and Biochemistry II (edited by prof. Lehotsky) and Medical Chemistry and Biochemistry III (written by author). It covers lectures in Medical Chemistry given in summer semester of the first year of General Medicine at Jessenius Faculty of Medicine in Martin.

This text will be followed soon with regular textbook, which will include also figures that should help students to understand described topics.

This text is divided to four chapters.

1. Enzymes – provides introduction to enzymology, describing mechanism of enzyme action and nature of coenzymes
2. Biologic membranes – deals with cell membranes with a focus on structure and function of plasma membrane and membranes of intracellular organelles.
3. Membrane transport – focused on transport of ions and other substances through biologic membranes.
4. Biochemistry of genetic information – deals with the key processes involved in storage and transmission of genetic information as well as with recombinant DNA methods used in human medicine.

The content of this textbook was not edited with respect to specialized correction of terminology, language, grammar and stylistics. Author of presented textbook is fully responsible for its content.

1. **ENZYMES**

Enzymes are proteins that catalyse specific biochemical reactions. From several thousands of biochemical reactions occurring in the cells of human body, only a few reactions are not catalysed by enzymes. Likewise all catalysts, enzymes decrease the activation energy ($E_a^{\ddagger}$) of a given reaction and are not consumed by the reactions they catalyze. This significantly increases the rate of the reaction, resulting in faster formation of products and faster achievement of equilibrium state of reaction. The rates of the most of enzyme catalysed reactions are million times faster than those of comparable un-catalysed reactions. For example, the uncatalysed decarboxylation of orotidine 5´-monophosphate has a half life of 78 million years but after addition of the enzyme, *orotidine 5´-phosphate decarboxylase*, the half life of the same reaction is 25 milliseconds. However, enzymes are different from most other catalysts in that they are highly specific for their **substrates**. Substrate, reactant of enzyme catalysed reaction that is converted to specific products, binds to the enzyme's **active site**. Active site, the region of enzyme that contains around 2 - 4 amino acid residues directly involved in the catalysis, binds the substrate and then carries out the reaction. Some enzymes do not need any additional components to be fully active. However, majority of enzymes require non-protein molecules called **cofactors** to be bound for activity. Cofactors can be either inorganic (e.g. ions of heavy metals) or organic compounds (e.g., heme or biotin). Organic cofactors can be either **prosthetic groups** that are covalently bound to enzyme polypeptide or **coenzymes**, which are bound like substrate, undergo chemical change and are then released from the enzyme's active site. In contrast to substrate for which a given enzyme is specific, coenzymes are common to many different enzymes. For example, about 700 enzymes are known to use NADH as coenzyme. Some cofactors can be bound to the sites that are distinct from active sites. In enzymes containing prosthetic group, the protein part of enzyme is called **apoenzyme**, whereas completely active enzyme is referred to as **holoenzyme**.

Medical importance of enzymes

- **Diagnostics of some diseases** – specific enzymes present in human plasma or other human fluids can indicate pathologic process (e.g. appearance of cardiac isoform of

creatine kinase (CK-MB isoform) in plasma indicates necrotic death of heart muscle cells for example due to myocardial infarction)

- **Treatment of some diseases** – therapeutical enzymes (*urokinase* or *streptokinase* are used to dissolve blood clot for early treatment of myocardial infarction)
- **Drug targets** – specific enzymes are target of drugs used for treatment of diseases, (inhibition of *cyclooxygenase*, enzyme involved in production of substances mediating pain and inflammation, by aspirin)
- **Gene therapy** – delivery enzymes to the body in order to treat specific disease (e.g. delivery of *adenosine deaminase* gene for the treatment of sever combined immunodeficiency)

Nomenclature and classification of enzymes

Nomenclature of enzymes was originally based on trivial names (e.g. *α-amylase*). Latter, name of enzyme was derived from the type of catalysed reaction (e.g. *lactate dehydrogenase*). The present nomenclature is based on systematic numbering and naming according to Enzyme commission of International Union for Biochemistry and Molecular Biology (IUBMB) (e.g. EC 1.1.1.27 *(S)-lactate: NAD^+ oxidoreductase*)

According to mechanism of catalysed reaction, enzymes are classified to 6 major classes:

- **Oxidoreductases** – catalyze transfer of reducing equivalents (e.g. electrons or hydrogen atoms) between donor and acceptor substances, e.g. *lactate dehydrogenase*
- **Transferases** – catalyze transfer of functional group (e.g. amino or phosphate group) between donor and acceptor substances, e.g. *alanine transaminase*
- **Hydrolases** – catalyze hydrolytic (i.e. in the presence of water) cleavage of chemical bonds (hydrolysis), e.g. *peptidases*
- **Lyases** – can catalyze both formation of chemical bond and cleavage of different chemical bonds by means other than hydrolysis and oxidation, e.g. *glutamic acid decarboxylase*

- **Isomerases** – catalyze changes within the one molecule, e.g. *2,3-bisphosphoglycerate mutase*
- **Ligases** – join two molecules together at the expense of a ATP hydrolysis, e.g. *DNA ligase*

Mechanism of enzyme catalysis

Enzyme catalysed reaction can be divided to several steps. First step involves the binding of substrate to the active site of enzyme and formation of **substrate-enzyme complex** (ES) that can act in several ways, all of which decrease $\Delta G^{\ddagger}$:

- Creating an environment in which the transition state is stabilised and activation energy is decreased (e.g. the enzyme alters the structure of bound substrate into transition state form, thereby reducing the amount of energy required to complete the transition)
- Decreasing the energy of the transition state by creating an environment with the opposite charge distribution to that of the transition state without straining the substrate
- Providing an alternative pathway. For example, enzyme reacts with the substrate forming a covalent intermediate ES complex, which would be impossible in the absence of the enzyme
- Reducing $\Delta S^{\ddagger}$ of the reaction by arrangement of reacting substrates in the correct orientation without any effect on $\Delta H^{\ddagger}$

Finally, the complex of enzyme-substrate is decomposed to the regenerated enzyme and reaction products.

Enzyme specificity

Enzymes are usually very specific, they bind only specific substrates and catalyze specific reactions. **Absolute specificity** means that enzyme is able to bind only one specific substrate (e.g. *glucokinase* that phosphorylates only D-glucose) while **group specificity** refers to the ability of some enzymes to bind structurally related substrates (e.g. *hexokinase*

that phosphorylates almost all hexoses). Enzymes can also exhibit **stereospecifity**, it means that they are able to bind only one stereoisomer (e.g. *glucokinase* that phosphorylates D-glucose but not L-glucose) and **bond specificity** that is ability of some enzymes to cleave or create specific bond (e.g. *α-amylase* that catalyses hydrolysis of α-O-glycosidic bond in starch but not β-O-glycosidic bond in cellulose).

Enzyme specificity was explained by the fact that both the enzyme and the substrate possess specific complementary geometric shapes that fit exactly into one another. This is often referred to as the **lock and key model** suggested by organic chemist and Nobel Prize laureate Emil Fisher. However, while this model explains enzyme specificity, it fails to explain the stabilization of the transition state that enzymes achieve. Therefore Daniel Koshland suggested a modification to the lock and key model called **induced fit model**. Since enzymes are rather dynamic structures, the binding of substrate to the active site is associated with changes in enzyme conformation. As a result, the amino acid side chains which make up the active site are shifted into the precise positions that enable the enzyme to perform its catalytic function.

Isoenzymes are enzymes catalysing the same reaction, however, are different in structure and thus in physical and chemical properties. Most of the isoenzymes are different in primary, secondary and tertiary structures however some of them are different at the level of quaternary structure. For example, muscle *creatine kinase* is composed of 4 M subunits (also called CK-MM), brain *creatine kinase* consist of 4 B subunits (also called CK-BB) whereas heart *creatine kinase* is composed from 2 M and 2 B subunits (also called CK-MB). Due to differences in the structure, isoenzymes are different in kinetic parameters and often are different in sensitivity to either activators or inhibitors. Isoenzymes are expressed in tissue- or cell-dependent manner however some of them are present in different compartments of cells (e.g. cytoplasmic and mitochondrial *superoxide dismutase*).

In addition to normal biochemistry, isoenzymes are of high biomedical importance since they can be used as crucial diagnostic markers. For example, elevated level of *creatine kinase* in blood may have different meaning depending on which isoenzyme is elevated. High levels of CK-MB may indicate myocardial infarction but also inflammation of the heart muscle (myocarditis), whereas CK-BB isoenzyme may be elevated due to stroke or

brain cancer. Finally, CK-MM elevation may indicate muscle injury or muscle damage (rhabdomyolysis) due to drugs, long lasting immobilization or excessive physical activity.

Enzyme kinetics

Enzyme kinetics describes how the particular enzyme binds substrate and turns it into product. Unlike classical chemical kinetics, increase of substrate concentration is associated with linear increase or reaction velocity only in the range of very low concentration of substrate. Further increase of substrate concentration increases the velocity of reaction however the change of the reaction velocity is not linear and proportional to the change in substrate concentration. The reaction velocity becomes to be almost constant at high concentrations of substrate.

How to describe mathematically the dependence of velocity of enzyme catalysed reaction on substrate concentration. The simplest reaction catalysed by enzymes is conversion of one product to one substrate that in terms of chemical kinetics can be expressed by following reaction scheme:

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_2} E + P$$

Overall reaction consists from three particular reactions. Formation of enzyme-substrate complex ES is the first reaction characterised by rate constant k_1. This reaction is reversible since ES complex can dissociate back to free substrate and enzyme. The reverse reaction is characterised by rate constant k_{-1}. Irreversible conversion of ES complex to product and regenerated enzyme is the third reaction characterised by rate constant k_2. In steady-state, the rate of ES complex formation is equal to the rate of ES complex dissociation, thus:

$$k_1[E][S] = k_{-1}[ES] + k_2[ES]$$

Assuming that concentration of free enzyme is equal to **total enzyme concentration** ($[E]_T$) minus concentration of enzyme in complex with substrate ($[E] = [E]_T - [ES]$) and that the rate of enzyme reaction is the same as the rate of conversion of ES complex to product (v =

k_2[ES]), we will get equation describing dependence of the rate of enzyme reaction on substrate concentration, known as **Michaelis-Menten equation:**

$$v = \frac{V_{\max}[S]}{K_M + [S]}$$

Where $V_{max} = k_2[E]_T$ represents **maximal rate of reaction**, thus the rate of reaction when the enzyme is fully saturated with substrate and exists only in the form of ES complex, and $K_M = \frac{k_{-1} + k_2}{k_1}$ is **Michaelis constant** that represents the concentration of substrate at which half of V_{max} is achieved. Thus, enzyme with low value of K_M works faster at lower concentrations of substrate than enzyme with high value of K_M.

Assuming that $k_{-1} >> k_2$, K_M has character of dissociation constant of ES complex. This means that K_M describes affinity of enzyme to substrate. As lower value of K_M as higher is affinity of enzyme to substrate. For example, K_M of *hexokinase* for glucose as the substrate is 0.15 mmol/l whereas K_M of *hexokinase* for fructose is 1.5 mmol/l. Thus affinity of *hexokinase* to glucose is 10 times higher than to fructose. K_M of *glucokinase* for glucose as the substrate is 8 mmol/l.

From known values of total concentration of enzyme $[E]_T$ and maximal rate **turnover number**, $\mathbf{k_{cat}}$ (k_2) can calculated.

$$k_{cat} = \frac{V_{\max}}{[E]_T}$$

Turnover number is a measure of how efficiently an enzyme converts a substrate into product. By definition, it represents number of substrate molecules that are converted to product by one molecule of enzyme per second if the enzyme is fully saturated by substrate. Turnover number can significantly vary among the enzymes. For example, *carbonic anhydrase* is the enzyme with high k_{kat} thus each enzyme molecule can convert 600 000 molecules of H_2CO_3 to H_2O a CO_2 per 1 second while *DNA polymerase I* with low k_{kat} = 15, can incorporate only 15 molecules of nucleotides to synthesised DNA strand per second.

Estimation of V_{max} and K_M

It is difficult to estimate values of V_{max} and K_M from hyperbolic dependence of reaction rate on substrate concentration. There are several methods allowing accurate estimation of these kinetic parameters. The **Lineweaver-Burk** plot or double reciprocal plot is the common way of presenting kinetic data. This is produced by taking the reciprocal values of both reaction rate ($1/v$) and substrate concentration ($1/[S]$). This produces a straight line, a linear form of the Michaelis–Menten equation, with a *y*-intercept equivalent to $1/V_{max}$ and an *x*-intercept of the graph representing $-1/K_M$.

$$\frac{1}{v} = \frac{1}{V_{\max}} + \frac{K_M}{V_{\max}} \frac{1}{[S]}$$

Enzyme inhibition

To date, three main mechanisms of enzyme inhibition has been described.

Competitive inhibition refers to a form of enzyme inhibition where inhibitor that is structurally related to the substrate binds to the enzyme active site and prevents binding of substrate. This form of enzyme inhibition is **reversible** since it can be overcome by increase of substrate concentration. The net effect of competitive inhibitor is increase of K_M whereas maximal rate of enzyme reaction is not affected. The competitive inhibition is characterised by **inhibition constant** (K_i) that represents dissociation constant for the enzyme-inhibitor complex. Several drugs used in human medicine works as competitive inhibitors. For example, statins act as competitive inhibitors of the *3-hydroxy-3-methyl-glutaryl-CoA reductase*, a key regulatory enzyme involved in the synthesis of cholesterol. They are used in human medicine to decrease hypercholesterolemia (high concentration of cholesterol in blood) since it correlates with incidence of vascular diseases such as myocardial infarction and stroke.

Non-competitive inhibition refers to a form of enzyme inhibition where inhibitor that is not structurally related to the substrate binds to the site that is distinct from enzyme active site. Non-competitive inhibitor binds to both free enzyme and enzyme-substrate complex (ES). Unlike competitive inhibition, it cannot be overcome by increase of substrate concentration. The binding of non-competitive inhibitor is often associated with formation of covalent bond resulting in change of enzyme conformation this form of inhibition is

irreversible and long lasting. The net effect of non-competitive inhibitor is decrease of maximal rate of enzyme reaction whereas K_M is not affected. Several drugs used in human medicine works as non-competitive inhibitors. The non-competitive inhibition is characterised by K_i that represents average dissociation constant for the enzyme-inhibitor complex and ES-inhibitor complex. For example, aspirin binds to molecule of *cyclooxygenase* that is involved in production of pro-inflammatory, pain mediating and fever inducing eicosanoides. Therefore, aspirin is used as pain killer, antipyretic and anti-inflammatory drug.

Uncompetitive inhibition refers to a form of enzyme inhibition where inhibitor that is not structurally related to the substrate binds only to the enzyme-substrate complex (ES) at a site that is distinct from enzyme active site. The net effect of uncompetitive inhibitor is decrease of both K_M and maximal rate of enzyme reaction. The uncompetitive inhibition is characterised by K_i that represents dissociation constant for the ES-inhibitor complex.

Allosteric enzymes

Allosteric enzymes are enzymes that change their conformation upon binding of an **effector**, which results in an apparent change in substrate binding affinity. Allosteric enzymes play a crucial role in many physiological processes, mainly in the regulation of metabolism and cell signalling. Most of allosteric enzymes are oligomers, but many enzymes demonstrating allosteric properties are single polypeptides. Allosteric enzymes consisting of multiple coupled subunits exhibit **cooperative binding** of substrate (like binding of oxygen to haemoglobin) in contrast to **non-cooperative binding** of allosteric enzymes without coupled subunits that display normal Michaelis-Menten kinetics. Cooperative allosteric enzymes display a **sigmoidal dependence** of reaction rate on the concentration of their substrates. This allows most allosteric enzymes to greatly vary catalytic output in response to small changes in effector concentration. Instead of K_M, affinity of allosteric enzymes to substrate is characterised by $\mathbf{K_{0.5}}$, which likewise K_M refers to the concentration of substrate at which half of maximal rate is achieved. Effector molecules may be the substrate itself **(homotropic effector**) or some other small molecule (**heterotropic effector**). They cause the enzyme to become more active (**positive effector**) or less active (**negative effector**) by changing enzyme conformation to the higher affinity state or lower affinity state. Thus, positive allosteric effectors are increasing affinity of

enzyme to substrate, whereas negative allosteric effectors are decreasing affinity of enzyme to substrate. The binding site for heterotropic effectors, called **allosteric site**, is usually different from the active site.

Regulation of enzyme activity – mechanisms of short term and long term control.

Activity of enzymes is regulated at four different levels:

1. **Substrate availability**. With respect to normal physiological concentrations of majority of intermediates of metabolic pathways, most of the enzymes involved in metabolism have very low affinity to their particular substrates. Thus concentration of substrate is significantly lower than Michaelis constant of the enzyme. Considering that $[S] << K_M$, Michaelis-Menten equation can be transformed to $\boldsymbol{v = V_{max}x[S]/K_M}$. It means that for most of the reactions, the rate of reaction is proportional to substrate concentration. Thus increased substrate availability is associated with proportional increase of reaction rate. On the other hand, to maintain constant fluxes of substrates through **blood brain barrier**, transport proteins expressed on the membranes of epithelial cells of brain vessels have high affinity to transported substrates (K_M << normal blood concentration of substrate). Therefore they are almost fully saturated with substrate and $\boldsymbol{v = V_{max}}$.

2. **Presence of inhibitors or activators.** Enzymes involved in regulation of the rate of particular metabolic pathway are often regulated by the presence of inhibitors or activators. Inhibition of key regulatory enzyme of particular metabolic pathway by end product of this pathway represent the most common type of regulation of enzyme activity at this level. This type of regulation of metabolic pathway is called **feed back inhibition**. The mechanism of inhibition is either **competitive inhibitio**n or **allosteric inhibition** (since end product of metabolic pathway is often structurally unrelated to the substrate of regulatory enzyme). Some regulatory enzymes are allosterically activated by end products of parallel pathway that works as **allosteric activators**. Typical example is cross regulation of synthesis of pyridine and purine nucleotides. The rate of synthesis of purine nucleotides is stimulated by presence of pyridine nucleotides and vice versa. Finally, the rate of some metabolic pathways is

stimulated by an initial substrate or an early product of the pathway. This type of pathway regulation is called **feed forward activation**.

3. **Covalent modifications.** Majority of regulatory enzymes are regulated by covalent modification of their polypeptide chain. The term covalent modification reflects attachment to or removal of some groups from polypeptide chain. Several different modifications (e.g. hydroxylation, methylation, adenylation, ADP-ribosylation…) have important impact on enzyme activity. With respect to regulation of enzyme activity and metabolism, the most important covalent modifications are phosphorylation and dephosphorylation of polypeptide chain. There is not general rule describing effect of these modifications on enzyme activity, some enzymes are more active when phosphorylated while others are more active dephosphorylated. However, if particular enzyme is more active phosphorylated, dephosphorylation of its polypeptide chain leads to decrease of enzyme activity and vice versa. For example, *glycogen synthase,* the primary enzyme of glycogen synthesis, is inhibited by phosphorylation and activated by dephosphorylation whereas *glycogen phosphorylase*, the primary enzyme of glycogen breakdown, is activated by phosphorylation and inhibited by dephosphorylation. Phosphorylation of regulatory enzymes is catalysed by specific group of enzymes called ***protein kinases*** that catalyse transfer of phosphate group from ATP to hydroxyl group containing amino acids (serine, threonine and tyrosine) present in polypeptide chain of regulated enzyme. In general, majority enzymes that are regulated by phosphorylation are modified at one amino acid residue despite presence of several hydroxyl containing amino acids present in their primary structure. Dephosphorylation of enzyme polypeptide chain is catalysed by specific enzymes called ***protein phosphatases***. Both *protein kinases* and *phosphatases* are regulated at several levels. Since enzymes can be phosphorylated and dephosphorylated later or vice versa, this type of covalent modifications is **reversible *(***as well as majority of covalent modifications). However, some enzymes are synthesised from inactive precursor called **zymogen** or **proenzyme**. Conversion of proenzyme to active enzyme requires hydrolysis of peptide bond at certain position by specific protease (e.g. conversion of nonactive tryspsinogen to active trypsin). Unlike phosphorylation/dephosphorylation, such covalent modification is **irreversible**.

4. **Enzyme concentration.** Maximal rate of enzyme reaction is proportional to total enzyme concentration ($V_{max} = k_2 x [E]_T$) that can significantly fluctuate depending on physiologic demands. Increase of enzyme concentration due to **increased expression** of enzyme is associated with increase of reaction rate. In addition, level of some enzymes might be decreased due to **repression of gene expression** or **proteolytic degradation** of the enzyme (e.g. by proteasomal 26S complex, see further). This is associated with decrease of reaction rate. Most of enzymes are **constitutive enzymes** that are produced constitutively by the cell during all physiological conditions and are not controlled by induction or repression of gene expression. The level of some enzymes (mainly regulatory enzymes) called **inductive enzymes** is influenced by induction or repression of gene expression in response to changes in physiologic demands. For example, *hexokinase* that catalyses glucose phosphorylation in extrahepatic tissues is constitutive with high affinity for glucose. Its function is to ensure supply of glucose for the tissues even in the presence of low blood glucose concentration. In contrast, *glucokinase* with low affinity for glucose is expressed in the cells of liver parenchyma as the response to increased blood glucose concentration. Expression of *glucokinase* gene is induced by insulin in order to remove glucose from blood and convert it into glycogen and TAGs.

Regulation of enzymes by substrate availability and presence of activators or inhibitors represent **short term regulation** since the response to changes of substrate concentration or concentration of regulatory compounds is fast. On the other hand it is short lasting.

Regulation of enzymes by covalent modifications and changes of enzyme concentration, called **long term regulation**, is not so fast but is long lasting. The modification of polypeptide chain requires some time due to chemical reactions. The process of enzyme expression is even longer involving transcription of particular gene to mRNA and its consequent translation of mRNA to enzyme polypeptide. However, both changes take longer.

In addition, enzyme activity depends on pH and temperature however these two parameters should rather be constant during normal physiologic conditions and therefore are not listed among factors contributing to regulation of enzyme activity.

COENZYMES

Coenzymes are small non-peptide organic compounds which are essential for some enzymatic reactions. Since coenzymes are chemically changed as a consequence of enzyme catalysed reaction, they can be considered as a special class of substrates, or second substrates, which are common to many different enzymes (for example about 700 different enzymes are using NADH as coenzyme). Coenzymes altered due to enzyme reaction are continuously regenerated and their concentrations are maintained at a certain steady intracellular level. For example, NAD^+ is reduced in glycolysis, β-oxidation of fatty acids and citric acid cycle to NADH that is oxidised in mitochondria by respiratory chain back to NAD^+. This continuous regeneration means that even small amounts of coenzymes are used very intensively. For example, the human brain turns over around 3.3 kg of NADH each day while it contains approximately only 75 mg of NADH. Some coenzymes or parts of some coenzyme molecules are **vitamins**, important compounds that are not synthesised in human body but are essential and therefore must be present in diet. Deficiency of some vitamins is associated with serious health problems that are indicated in paragraph dealing with particular coenzyme. If any problem is indicated, this in majority coenzymes means that deficiency of particular vitamin is rare.

Coenzymes are classified according to reaction they are involved in to several different classes.

Redox coenzymes

Nicotinamide adenine dinucleotide (phosphate), $NAD(P)^+$

The purine nucleotide containing NAD^+ and $NADP^+$ are coenzymes of *dehydrogenases*, subclass of *oxidoreductases*. In addition, they contain redox active **nicotinamide** that is **vitamin B_3**. While NADH transfers protons and electrons from catabolic pathways to the respiratory chain, NADPH, reduced form of $NADP^+$, is the most important proton donor involved in biosynthesis.

Nicotinamide can be synthesised in human body from tryptophan, however with very low efficiency. Therefore, vitamin B_3 deficiency only occurs when both nicotinamide and tryptophan are lacking in human diet and is manifested in the form of skin damage (**pellagra**), digestive disturbances and depression.

Flavin mononucleotide, FMN and flavin adenine dinucleotide, FAD

The flavin containing FMN and FAD are coenzymes of *dehydrogenases*, *oxidases* and *monooxygenases.* In addition to redox active flavin, they contain phosphorylated ribitol attached to the flavin ring. **Riboflavin** is known as **vitamin B_2**. FAD is produced from FMN by reaction with AMP. In contrast to $NAD(P)^+$, both FMN and FAD are covalently bound to the enzymes as prosthetic groups.

Ubiquinone

Ubiquinone (**coenzyme Q**) transfers electrons in respiratory chain. In addition to quinone ring, it contains isoprenoid chain that holds ubiquinone in inner mitochondrial membrane where it is freely mobile. Ubiquinone is synthesised by cells of human body, but its synthesis decreases significantly with age. There are some indications about association of poor availability of ubiquinone with major age-related diseases (e.g. neurodegenerative and cardiovascular diseases, metabolic syndrome etc.), however, consistent evidence is still lacking.

L-Ascorbic acid

L-Ascorbic acid, **vitamin C**, is coenzyme of various *monooxygenases* and *dioxygenases*. Majority of mammals, except of higher primates including humans, can synthesize L-ascorbic acid from glucose. L-Ascorbic acid is involved in hydroxylation of proline and lysine residues during collagen synthesis, in the synthesis of catecholamines and bile acids. Due to involvement of ascorbic acid in proline and lysine hydroxylation, chronic low intake of vitamin C is associated with symptoms of **scurvy** (abnormal bleeding, gum disease, loosening of teeth and poor wound healing) that are attributed to synthesis of defective collagen and consequent loose connective tissue formation.

During reactions, L-ascorbic acid is oxidised to **dehydroascorbic acid** that can be reduced back to L-ascorbic acid by glutathione and other thiols. However, dehydroascorbic acid can also be further oxidised to **oxalic acid** that is toxic and tends to precipitate with calcium ions. Therefore, to prevent formation of **kidney stones**, people suffering from kidney malfunction should not take high doses of L-ascorbic acid, as it is recommended during common cold or flu.

Tetrahydrobiopterin

Tetrahydrobiopterin containing hydrogenated **pteridine** ring that is also part of **tetrahydrofolic acid** (see further) is coenzyme of some *monooxygenases* (e.g. *tyrosine hydroxylase* catalyzing hydroxylation of tyrosine to DOPA).

Coenzymes transferring one carbon units

Tetrahydrofolic acid

Tetrahydrofolic acid (THF) consists of pteridine ring attached to amino group of 4-aminobenzoic acid that is attached via carboxyl group to amino group of glutamic acid. It is coenzyme that transfers one carbon residues in different oxidation states like formyl (oxidation state of formic acid), methylene (oxidation state of formaldehyde) and methyl (oxidation state methanol). THF is produced from vitamin, **folic acid** (**vitamin B_9**), by hydrogenation of heterocyclic pteridine ring. Transfer of one carbon unit by THF plays role in synthesis of purine nucleotides and dTMP.

Deficiency of folic acid is relatively common and is associated with disturbances in nucleotide biosynthesis and therefore cell proliferation. Since the precursors of blood cells divide particularly rapidly, deficiency of folic acid can lead to the changes in blood cell count with elevated amounts of large immature and dysfunctional red blood cells (megaloblasts) causing **megaloblastic anemia**. In contrast to animals, bacteria are able to synthesize folic acid. Therefore, the growth of bacteria can be inhibited by **sulfonamides**, which competitively inhibit synthesis of folic acid (attachment of 4-aminobenzoic acid to pteridine ring catalysed by enzyme *dihydropteroate synthetase*). Sulfonamides have minimal effect on human metabolism, since humans acquire folic acid through the diet. Therefore, sulfonamides are used for treatment of different bacterial infections.

S-adenosyl methionine

S-adenosyl methionine that is synthesised from adenosine triphosphate and methionine is donor of methyl group involved in many methylation reactions, like synthesis of epinephrine from norepinephrine or in methylation of DNA.

Methylcobalamin

Methylcobalamin that is methyl derivative of cobalamin, **vitamin B_{12}**, is coenzyme of different *methyltransferases* (e.g. synthesis of methionine from homocysteine). Methylcobalamin is used in the treatment of **peripheral neurophaty** (damage to the nerves of peripheral nervous system, PNS) or **diabetic neurophaty** (disorder of nervous system associated with diabetes mellitus).

Biotin

Biotin is the coenzyme of *carboxylases*. Using ATP, it reacts with bicarbonate (HCO_3^-) to form **N-carboxy biotin**, activated form of CO_2 that is then transferred to other molecules (e.g. formation of oxaloacetic acid from pyruvic acid).

Vitamin K

Vitamin K represents group of structurally related compounds that work as coenzymes of ***γ-glutamyl carboxylase***. Carboxylation of glutamic acid residues to γ-carboxyglutamic acid that is catalysed by *γ-glutamyl carboxylase* occurs in proteins involved in blood clotting cascade, factors II, VII, IX and X, and some bone proteins. This modification increases affinity of these proteins to calcium that serves as initiator of blood clotting and is essential component of bones. Inhibitors of the enzymatic conversion of inactive vitamin K to its active form (e.g. warfarin) are used to prevent blood coagulation in prophylactic treatment against **thrombosis** (excessive formation of a blood clot, **thrombus**, inside a blood vessels, obstructing the blood flow through the circulatory system) for example in people with high risk of brain stroke.

Coenzymes transferring two and more carbon units

Coenzyme A

Coenzyme A (CoA) in chemical terms is thiol that can react with carboxylic acids to form **thioesters**, thus functioning as an **acyl group carrier**. Activation of carboxylic acids to their acyl-CoA thioesters is essential metabolic reaction. **Acetyl-CoA** is central metabolic molecule allowing integration of saccharide, lipid and amino acid metabolism. In addition to cysteamine and 3´-phosphoadenosine diphosphate, CoA contains **pantothenic acid**, **vitamin B_5**.

Thiamine diphosphate

Thiamine diphosphate (TPP), phosphoderivative of thiamin, **vitamin B_1**, is coenzyme of enzymes that are able to activate aldehydes and ketones as hydroxyalkyl group and transfer them on the other molecule. It is involved either in **oxidative decarboxylation of α-keto acids** (e.g. conversion of pyruvic acid to acetyl-CoA) or in **transketolase reactions** (e.g. in the pentose phosphate pathway).

Due to involvement of thiamine diphosphate in reaction that are essential for energetic metabolism, vitamin B_1 deficiency leads to **beriberi**, disease with different symptoms including neurologic disturbances, cardiac insufficiency and muscular atrophy.

Lipoic acid

Lipoic acid is involved as coenzyme in oxidative decarboxylation of α-keto acids that is catalysed by enzyme complexes like *pyruvate dehydrogenase*, *α-ketoglutarate dehydrogenase* and *branched chain α-keto acid dehydrogenase*. It transfers acyl residue to CoA leading to the formation of acyl-CoA.

Coenzymes transferring other groups

Pyridoxal phosphate

Pyridoxal phosphate, **vitamin B_6**, is the most important coenzyme in amino acid metabolism. It works as the coenzyme in transamination, decarboxylation and dehydration reactions of amino acids.

Deoxyadenosylcobalamin

Deoxyadenosylcobalamin is coenzyme of various *isomerases*, which catalyse rearrangement of different molecules. The most important reaction of this type is conversion of methylmalonyl-CoA to succinyl-CoA that is involved in catabolism of branched chain amino acids, valine and isoleucine.

3´-phosphoadenosine -5´-phosphosulfate

3´-phosphoadenosine-5´-phosphosulfate (PAPS) serves as coenzyme in *sulfotransferase* reactions that are important for sulfatation of proteoglycans and glycsoaminoglycans as well as for synthesis of sulfocerebrosides.

2. BIOLOGIC MEMBRANES

Biologic membranes are dynamic cellular structures composed of phospholipids and embedded proteins. Salts of higher fatty acids are forming **micelles** in water solutions. However, amphiphilic compounds with hydrodrophobic hydrocarbon chains and polar groups, as glycerol- and sfingophospholipids, are forming higher structures of micelles with **bilayer of lipids**. Such lipid bilayer, impermeable for ions and polar compounds, is the main component of biologic membranes. In addition to lipids, proteins are components of biologic membranes.

Different types of biologic membranes are present in eukaryotic cells. All cells have got **plasma membrane**, separating them from environment and determining the shape and mechanical resistance of the cell. In addition to plasma membrane, intracellular membranes form intracellular organelles – mitochondria, nucleus, sarco/endoplasmic reticulum, lysosomes, Golgi apparatus, and secretory granules.

Biologic membranes perform several important functions:

1. separation of cell from environment or separation of subcellular organelles from cytoplasm
2. controlled and selective transport of ions and metabolites
3. signal transduction
4. site of some enzymatic reactions
5. contact and interaction with other cells
6. anchor for cytoskeletal proteins

Membrane structure

Phospholipids and proteins are the main components of biologic membranes. The ratio between amount of proteins and lipids in particular membranes is different in different types of membranes and depends on function of particular membrane. The high content of proteins in membranes of intracellular organelles is determined by presence of different enzymes and other membrane proteins (Table 1).

Table 1. Protein/lipid ratio in different types of cell membrane

Membrane	protein/lipid
Myelin	0.2
Erythrocyte	1.1
outer mitochondrial membrane	1.2
sarcoplasmatic reticulum	2.0
inner mitochondrial membrane	3.2

Membrane lipids

The main lipid components of biologic membranes are phospholipids, glycolipids and cholesterol. All of them are of **amphiphilic nature**. They spontaneously arrange so that the hydrocarbon chains due to their mutual hydrophobic interactions are shielded from the surrounding polar environment, causing that the hydrophilic regions, cytoplasm and extracellular space, associate with the cytosolic and extracellular faces of the resulting bilayer. Thus the formation of lipid bilayer is driven by **hydrophobic interactions** of hydrocarbon chains of fatty acids as well as by **increase of water entropy** (since water has partial liquid crystal structure, formation of membrane decreases order of water structure). **Electrostatic interactions** between polar groups of phospholipids and hydrophilic components of cytoplasm, extracellular space or organelles also contribute to stabilisation of membranes.

The amount of particular lipid depends upon the cell and organelle type (Table 2), but in the majority of cases phospholipids, especially glycerolphospholipids as phosphatidylcholines and phosphatidylethanolamines, are the most abundant lipid components of biologic membranes. The amount of other glycerolphospholipids (phosphatidyserines, phosphatidylinositols and plasmalogenes) is significantly lower. Cholesterol is abundant in plasma membrane and myelin while low concentration of cholesterol is in inner mitochondrial membrane. **Cardiolipin** is almost exclusively present in **inner mitochondrial membrane**, being the third most abundant lipid of inner mitochondrial membrane. Sfingolipids are present mainly in plasma membrane. The most abundant is sfingomyelin. In central nervous system, glycosfingolipids,

galactocerebrosides and gangliosides, exert very important function. Their saccharide residues, extending surface of membranes, are involved in cell-cell interactions and binding of important substances (e.g. hormones).

Table 2. Lipid composition of some membranes

Lipid	Relative amount (%)			
	Human erythrocyte	**Human myelin**	**ER**	**Beef heart mitochondria**
Phosphatidic acid	1.5	0.5	0	0
Phosphatidylcholines	19	10	40	39
Phosphatidylethanolamines	18	20	17	27
Phosphatidylinositols	1	1		7
Phosphatidylserines	8.5	8.5	5	0.5
Cardiolipin	0	0	0	22.5
Sfingomyelin	17.5	8.5	5	0
Glycolipids	10	26	0	0
Cholesterol	25	26	6	3

Asymmetric distribution of lipids in plasma membrane

Composition of lipid layers of membrane is not equal and significant difference between outer and inner layer exists with respect to content of particular lipid type. Outer (extracellular) part of membrane is rich in choline containing phospholipids (phosphatidylcholines and sphingomyelines). In addition, glycolipids involved in cell-cell interactions are present in outer layer**.** Inner (intracellular or cytoplasmic) part of membrane contains mainly phosphatidylethanolamines and phosphatidylserines. Apoptosis is associated with appearance of phosphatidylserines at the surface of apoptotic cells. *In vivo*, functional consequence of the presence of phosphatidylserines at the outer leaflet of

plasma membrane is removal of such cell by phagocytes – specific cells of immune system.

Membrane proteins

Membrane proteins perform several important functions:

- transport of ions and molecules
- some of them are enzymes
- some membrane proteins are receptors or other components of signal transduction
- structure and mechanical stability of cells
- cell-cell or cell-extracellular matrix interactions

Membrane proteins are classified according to the nature of their interaction with membrane to:

- integral (transmembrane) proteins
- peripheral proteins
- anchored proteins

Integral proteins span the membrane and have a hydrophilic cytosolic domain which interacts with intracellular molecules, a hydrophobic membrane-spanning domain that anchors protein within the membrane, and a hydrophilic extracellular domain that interacts with molecules of extracellular matrix. The transmembrane domains are predominantly α-helices composed mainly from hydrophobic amino acids (valine, leucine, isoleucine, and tyrosine).

Peripheral proteins are either attached to integral membrane proteins or associated with polar region of the lipid bilayer due to electrostatic interactions. These proteins interact only temporary with biologic membranes. After reaction, the molecule dissociates from membrane to perform its function in the cytoplasm.

Anchored proteins are covalently-bound, via ester, thioether and amide bonds, to single or multiple lipid molecules; hydrophobically inserted into the cell membrane and anchor

the protein. The protein itself is not in contact with the membrane. The types of anchors are:

Palmitate anchor consists of palmitic acid attached via ester bond to hydroxyl group of serine present in anchored protein.

Myristate anchor consists of myristic acid attached via amide bond to amino group of amino acid at the N-end of anchored protein.

Farnesyl anchor consists of farnesyl attached via thioether bond to thiol group of cysteine present in anchored protein.

Glycosylphosphatidylinositol (**GPI) anchor** is composed of a phosphatidylinositol group linked through an oligosaccharide linker (glucosamine and mannose bound to the inositol residue via glycosidic bond) and via an ethanolamine phosphate bridge to the C-terminal amino acid of anchored protein. The two fatty acids within the hydrophobic phosphatidyl-inositol group anchor the protein to the membrane.

Fluid mosaic model

According to the fluid mosaic model, the biological membranes can be considered as a two-dimensional liquid where all molecules of membrane lipids and proteins are not fixed but diffuse more or less easily. This diffusion occurs within the space scale of 10 nm. However, the plasma membranes contain different structures or domains that can be classified as: protein-protein complexes, lipid rafts and pickets and fences formed by the actin-based cytoskeleton

Lipid rafts are relative stable micro-domains that are characterised by specific lipid and protein composition. Lipid rafts are more ordered and tightly packed than the surrounding bilayer, but can move freely in the membrane bilayer. Lipid rafts are enriched in sphingomyelin, which is typically elevated by 50% compared to the rest of plasma membrane, and cholesterol, 3 to 5-fold the amount of cholesterol present in the surrounding bilayer.

Movement of lipids in membranes

Membrane phospholipids can perform three types of movement:

Lateral diffusion, including bending of fatty acid chains, is very fast and spontaneous process that refers to the lateral movement of lipids in the membrane. Membrane lipids are generally free to move laterally if they are not restricted by certain interactions.

Rotation of hydrocarbon chains of fatty acids around C-C bonds is the simplest movement of lipids in membrane. This type of motion is greatest at the hydrophobic ends (methyl end) i.e. in the middle of lipid bilayer.

Flip-flop, or transverse diffusion, involves the movement of a lipid from one membrane layer to the other. Unlike lateral diffusion, it is a fairly slow process due to the fact that a relatively high amount of energy is required for flip-flopping to occur. It can be catalysed by specific enzymes called ***flippases***.

Movement of membrane proteins, membrane polarity

Membrane proteins (e.g. ion channels or pumps) can freely move throughout whole membrane by lateral diffusion. Thus, distribution of membrane proteins within membrane should be even and therefore identical in composition and activity without significant differences between different parts of cell membrane. However, some cells, in general called **polarized cells**, exhibit significant differences with respect to membrane proteins distribution. This is particularly evident in epithelial and endothelial cells, but also describes other polarized cells, such as neurons. According to orientation, membrane of polarized cells can be classified to **apical** and **basolateral** membrane.

The apical membrane of a polarized cell is the surface of the plasma membrane that faces the lumen whereas the basolateral membrane of a polarized cell is the surface of the plasma membrane that forms its basal and lateral surfaces. Basal surface faces towards the interstitium, and away from the lumen. Lateral surface faces towards lateral surface of neighboring cell.

The uneven distribution of membrane proteins is caused by different sorting of newly synthesised membrane proteins (in all types of polar cells) as well as by presence of specific complexes of membrane protein called **tight junctions** (in epithelial and endothelial cells). Tight junctions that join epithelial cells near their apical surface prevent

the movement of proteins from the basolateral membrane to the apical membrane. The basal and lateral surfaces thus remain roughly equivalent to one another, yet distinct from the apical surface. In addition to cell polarity, tight junctions represents important barrier to prevent entry of some substances into our body or to some specific tissues like CNS. Tight junctions between enterocytes (specialised epithelial cells of the small intestinal mucosa) prevent absorption of undigested oligosaccharides and peptides from GIT to the blood as well as large hydrophilic molecules that are not of physiologic importance or that are toxic. They also occur between endothelial cells of brain capillaries restricting diffusion of solid particles (like bacteria and viruses) and large hydrophilic molecules into cerebrospinal fluid (special extracellular fluid of CNS). The barrier function of tight junctions can be disrupted by inflammation, this might cause to entry of substances that do not normally occur in blood and CNS or bacterial infection of CNS.

Likewise membrane lipids, membrane proteins can also perform rotational movement.

Membrane fluidity

Membrane fluidity that refers to the viscosity of the lipid bilayer is consequence of the movement of lipids in the membrane. It depends on temperature and lipid composition of membrane. At low temperature, fatty acid chains of membrane lipids are arranged tightly next to each other creating ordered system like **gel-crystalline state**. As temperature is increased, there is a phase transition into non-ordered system with higher fluidity that resembles **liquid-crystalline state**. **Transition temperature**, T_m, refers to temperature of the transition of membrane from gel- into liquid-crystalline state. Due to heterogeneity in membrane composition, T_m is not precise parameter for biologic membrane.

Membranes with glycerophospholipids containing short –chain and unsaturated fatty acyl groups have higher fluidity. Longer and saturated chains produce stronger interactions that are increasing T_m, whereas cis-configurations of double bonds lead to the kinks in hydrocarbon chains, which prevent tight packing of the chains. Flat rigid structure of cholesterol reduces the coiling of the fatty acid chains thus cholesterol decreases fluidity of membrane. Membrane fluidity is also reduced by Ca^{2+} due to its interaction with negatively charged phospholipid head groups. This reduces repulsion between head groups and consequently increases tight packing of lipid molecules. In addition to temperature,

there are other exogenous factors that affect membrane fluidity. Membrane fluidity can be altered by alcohol, drugs and anaesthetics.

Membrane fluidity has important impact on the functions of biologic membrane. Higher fluidity is associated with higher membrane permeability for water and other hydrophilic substances. Membrane fluidity affects the activity of membrane-bound receptors and enzymes and can also control processes as phagocytosis and cell growth and death. One of the intoxicating effects of ethanol is attributed to modification of membrane fluidity, altering functioning of membrane receptors and ion channels.

Changes in membrane structure at pathological conditions

Acanthocytes are a form of erythrocytes that are coarse and spiked, or possess various abnormal thorny projections resembling many-pointed stars. Altered membranes may contain decreased phosphatidylcholine levels but increased levels of cholesterol and sphingomyelin. **Acanthocytosis**, spur cell anaemia, is the condition with acanthocyte-like erythrocytes. They are seen on blood films in abetalipoproteinemia, inherited disorder associated with pathologically low blood cholesterol level. In addition, presence of acanthocytes may indicate liver diseases and several inherited neurological disorders, such as neuroacanthocytosis.

The term "**echinocyte**" (or "**burr cell**") is similar to "acanthocyte", but implies more moderate speculation of erythrocytes. Appearance of burr cells in blood usually indicates uremia (high concentration of blood urea).

Liposomes

Liposomes are synthetic membranes that can be used as experimental models to study functions of membrane proteins. In human medicine, they can be used as carriers of drugs and therapeutical biomolecules like enzymes or DNA. Liposome prevents destruction of biomolecules by proteases or nucleases and allows translocation of biomolecules through cellular membrane. Modification of liposome surface (e. b. by specific antibodies that recognize surface proteins of the targeted cells) enables targeted scavenging of liposomes by tissues or tumours.

Plasma membrane

Plasma membrane interacts with both extracellular and intracellular space. At the extracellular space, the plasma membrane is coated with **glycocalyx**, the polysaccharide matrix protecting cell membrane from chemical and other injury and that is also involved in cell-cell interactions. At intracellular space, the plasma membrane interacts via membrane proteins with cytoskeletal proteins. This interaction is important for maintaining the cell shape and mechanical stability and resilience of cells.

For example, mechanical stability and resilience of the erythrocyte membrane is determined by interactions among membrane integral proteins and cytoskeletal network. The major transmembrane proteins erythrocyte membrane are glycoproteins, **glygophorin** and **band 3**, that is a multi-spanning anion channel (permeable for Cl^- and HCO_3^-) existing in a dimer/tetramer equilibrium. It is structurally important because band 3 tetramers tether the lipid bilayer to the skeleton via an interaction between its cytoplasmic domain and **ankyrin** which is associated with **spectrin**. The erythrocyte cytoskeleton is an irregular hexagonal lattice of polymeric spectrin molecules which are tied together by actin, **protein 4.1** and other proteins at nodes called junctional complexes. The skeleton makes a two dimensional network which is very flexible and compressible. The flexibility and compressibility as well as mechanical stability are very important since erythrocyte width is on average about 25% larger than diameter of capillaries.

Membranes of intracellular organelles

Nuclear envelope is a double membrane intracellular organelle that contains the genetic material in eukaryotic cells. It also serves as the physical barrier, separating the contents of the nucleus (genomic DNA) from the cytoplasm. The nuclear envelope contains many **nuclear pores** that facilitate and regulate the exchange of important biomolecules (e.g. proteins such as transcription factors or RNA) between the nucleus and the cytoplasm.

Endoplasmic reticulum, or **sarcoplasmic reticulum** in muscle cells, is one membrane organelle of almost all eukaryotic cells that forms an interconnected network of tubules, vesicles, and cisternae. **Rough endoplasmic reticulum** is involved in synthesis of extracellular and plasma membrane proteins as well as proteins of endoplasmic reticulum and Golgi apparatus, while **smooth endoplasmic reticulum** plays role in synthesis of

lipids and steroids, and saccharide and steroid metabolism. Endoplasmic reticulum is important intracellular store of calcium. For example, release of calcium from with sarcoplasmic reticulum is essential with respect to muscle contraction (both skeletal and heart muscle). Endoplasmic reticulum of hepatocytes is also involved in drug catabolic metabolism.

Golgi apparatus is composed of stacks of membrane-bound structures known as cisternae. A mammalian cell typically contains 40 to 100 stacks. Each cisterna comprises a flat, membrane enclosed disc that contains special Golgi enzymes which are involved in glycosylation of proteins and distribution of proteins to plasma membrane or extracellular space. Proteins delivered to plasma membrane or secreted to the extracellular space are engulfed in specific vesicles which are generated from Golgi apparatus membrane.

Lysosomes are one membrane containing intracellular organelles that are involved in elimination of modified or exogenous proteins as well as degradation of other biomacromolecules.

Peroxisomes are one membrane organelles involved in the catabolism of very long and branched fatty acids, D-amino acids and polyamines as well as biosynthesis of plasmalogens.

Mitochondria composed from two membranes are organelles found in almost all eukaryotic cells. **Outer mitochondrial membrane** separates mitochondria from cytoplasm. **Inner mitochondrial membrane** that separates **intermembrane space** form **mitochondrial matrix** contains enzyme complex of **oxidative phosphorylation**, the major ATP producing pathway. The inner mitochondrial membrane is compartmentalized into numerous **cristae**, which expand the surface of the inner mitochondrial membrane, enhancing its ability to produce ATP. Mitochondria are the most important source of ATP; almost all cellular ATP is produced in mitochondria. In addition, they are involved in modulation of intracellular calcium concentration, fatty acid and amino nitrogen metabolism and initiation of programmed cell death, apoptosis.

3. MEMBRANE TRANSPORT

The main function of plasma membrane is to separate cells from their environment. However, cells depend on some ions and essential substances, physiologically important substances which are not produced by particular cell. On the other hand, some substances produced by particular cells have to be delivered to another cell or can be potentially harmful and have to be eliminated from cells. In addition, intracellular concentration of some ions is not constant and can significantly fluctuate within the time (e.g. intracellular concentration of calcium). Finally, the membrane potential of excitable cells is given by regulated permeability of plasma membrane to main extracellular cation, sodium cation, and main intracellular cation, potassium cation. Thus, there must be mechanisms allowing both entry of ions and essential substances into the cells as well as extrusion of ions and intracellularly produced substances out of the cells. The controlled and selective transport of ions and metabolites is performed by several membrane transport mechanism.

Membrane transport can be classified according to **morphology**, **mechanism** and **energy demands**.

According to morphology membrane transport is classified to three classes:

- **homocellular transport** – transport of substances into or from the cells, e.g. entry of glucose into cells via glucose transporters
- **intracellular transport** – transport of substances through membranes of intracellular organelles, release of calcium from sarcoplasmic reticulum via intracellular calcium channels
- **transcellular transport** – transport of substances through cells, transepithelial transport of glucose from lumen of small intestine through epithelial cells into blood

According to number and direction of transported substances membrane transport is classified to:

- **uniport** – transport of one substance through membrane, transport of glucose into cells via glucose transporters.

- **symport** – transport of two substance through membrane in the same direction, sodium-dependent uptake of amino acids
- **antiport** – transport of two substances through membrane in opposite direction, ADP/ATP translocation through inner mitochondrial membrane

According to energy demands membrane transport is classified to:

- **passive transport** – transport of substance does not require energy
- **active transport** – transport of substance requires energy

In addition, transport of ions can be **electroneutral**, if the net difference of charge of transported ions is zero (e.g. antiport of equal amount of equally charged ions). If the difference of charge of transported ions is not zero, the transport is associated with changes of charge on both sides of membrane and it is therefore **electrogenic**.

Passive transport – passive and facilitated diffusion, channels

Passive transport does not require energy and substances are transported down of the concentration or electrochemical gradient of transported substance via two different mechanisms, **passive diffusion** or **facilitated diffusion**.

Passive diffusion is the simplest membrane transport. The substance is transported from the area with its higher concentration to area with lower concentration without aid of specific carrier. With respect to thermodynamics, this process is exergonic with a tendency to reach equilibrium, equal concentration of substance. Transport of uncharged molecule from the area with concentration c1 to area with concentration c2 is associated with change of Gibbs energy that is given:

$$\Delta G = RT \ln \frac{c_2}{c_1}$$

Transport of charged particle is function of electrochemical potential:

$$\Delta G = RT \ln \frac{c_2}{c_1} + nF\Delta\Psi$$

where n is ion charge, F is Faraday constant and $\Delta\psi$ is electric potential.

The rate of flow of uncharged molecules is given:

$$J = -P(c_2 - c_1)$$

where P [cm/s] is **coefficient of permeability** that characterizes ability of molecule to cross the membrane, For example, permeability coefficient of Na^+ in artificial phospholipid membrane is $<1.10^{-12}$, glucose 4.10^{-10} and water 5.10^{-3} cm/s.

Gases (like O_2, CO_2, N_2 and NO) and small non polar substances, including drugs, can cross the biological membrane by simple diffusion. The cell membranes are also partially permeable for water. Water crosses the membrane due to osmosis through leaks in membrane that are spontaneously formed during random movement of hydrocarbon chains of fatty acids. The membrane is impermeable for large, polar or charged substances

The rate of passive diffusion of some important substances through membrane is not high enough to cover physiological demands of cells. In addition, the rate of transport of some substance cannot be constant and therefore is strictly controlled depending on physiologic demands of cell (e.g. increased entry of glucose to muscles or_regulated permeability of plasma membrane to ions). Thus, some substances are transported into cells by mechanism called facilitated diffusion that involves specific transport proteins. Unlike simple diffusion, facilitated diffusion is:

- **Specific** – only specific substrate can be transported by particular transporter.
- **Saturable** – increased concentration of transported substance increase rate of transport, however, above certain concentration of transported substrate the rate transport is constant since transport protein is fully saturated by transported substance.
- **Inhibitable** – facilitated diffusion can be inhibited either by mechanism of competitive or non competitive inhibition.

Glucose transporters

Family of glucose transporters (GLUT) represents isoforms membrane proteins that are involved in facilitated diffusion of glucose from extracellular space to the cells e.g.

transport of glucose from intestinal lumen into enterocytes. Other D-hexoses as well as glycerol are translocated by some isoforms of GLUT. In addition to insulin, expression of GLUT4 isoform in muscles is stimulated by exercise and hypoxia, reflecting a need for increased glucose utilization by the muscled during these conditions.

Table 3. Localisation and function of different isoforms of glucose transporter

Type	Localisation	K_M (mmol/l)	Function	Characteristic
GLUT1	Majority of tissues	1	Basal uptake of glucose	
GLUT2	Pancreas and liver	10-20	Pancreas – regulation of insulin release Liver – removal of glucose excess	
GLUT3	Majority of tissues	1	Basal uptake of glucose	
GLUT4	Heart and striated muscles, adipocytes	5	Regulated uptake of glucose	Inducible by insulin
GLUT5	Small intestine	?	Absorption of glucose	

Channels

Ion channels are important proteins allowing controlled entry of ions into the cells via facilitated diffusion. Majority of ion channels are present on plasma membrane of the cells, however there are some **intracellular channels** localised at the membranes of intracellular organelles (e.g. **ryanodine receptor** – calcium channel mediating release of Ca^{2+} from muscle sarcoplasmic reticulum, **IP_3 receptor** – calcium channel that after binding of IP_3 mediates release of Ca^{2+} from endoplasmic reticulum of some cells (e.g. neurones), **ATP-dependent potassium channel** of inner mitochondrial membrane). Since ion channels are open only as the response to specific stimuli, they are classified according to different mechanisms of their gating to:

- **voltage dependent** ion channels – opening of channel depends on membrane potential
- **ligand dependent** ion channels – opening of channel depends on binding of specific ligand to channel (ionotropic receptors), e.g. neuronal glutamate receptors permeable for calcium
- **mechanically dependent** ion channels – opening of channel depends on the influence of stretch, pressure, shear, vibration and displacement.

In addition to the main three types of ion channels, gating of some ion channel responds to **light**, **heat** and **cold**.

Nicotinic acetylcholine receptor

Nicotinic acetylcholine receptor is ligand dependent sodium channel that is permeable for sodium ions after binding of acetylcholine. It is partially sensitive to **nicotine**, therefore the name nicotinic (in contrast, muscarinic acetylcholine receptor that is partially sensitive to muscarine is not sodium channel but it is coupled to G proteins). It consist from five subunits each of them contains four transmembrane domains. On transmembrane α-helix of each subunit is involved in gating of acetylcholine receptor. It is expressed mainly on **neuromuscular junctions** (specific connections between peripheral nerves and muscles) where it is involved in transmission electric impulses from nervous system to muscles. This process is essential for initiation of **muscle contraction**. Therefore, blockers of nicotinic acetylcholine receptors are causing muscle paralysis (like curare used as a paralyzing poison by South American indigenous people).

Aquaporins

Aquaporins are cell membrane proteins that regulate the flow of water through plasma membrane. They are permeable for water but impermeable for H_3O^+. There are essential for re-absorption of water in kidney.

Ionophores

Ionophores are substances produced by some microorganisms that are able to transport ions through cell membrane disrupting concentration gradient of certain ions. Due to these properties, some of them are used as antibiotics. According to the mechanism of ion transport, ionophores are classified to:

- **Mobile carriers of ions** (valinomycin)
- **Channel-like ionophores** (gramicidin A)

Gap junctions

Gap junctions are produced by transmembrane proteins that cross plasma membrane of the cells creating specialised intercellular connection between the neighboring cells which allows various molecules smaller than 1kDa and ions to pass freely between the cells. Gap junctions allow for direct electrical and chemical communication between cells ensuring that molecules and current transferred between two cells do not leak into the intercellular space.

One gap junction channel is composed of two **hemichannels** each of them penetrating through the membrane of particular cell and which connect across the intercellular space. Thus, two hemichannels, joined together across a cell membrane comprise a gap junction channel. One hemichannel consists of six subunits, called **connexins** that can be identical (homomeric) or different (heteromeric hemichannel). One connexin protein has four transmembrane domains.

Gap junctions are particularly important in cardiac muscle: the calcium ions causing muscle contraction are passed efficiently through gap junctions, allowing the heart muscle cells to contract in tandem. Gap junctions are expressed in virtually all tissues of human body, except of adult skeletal muscles and mobile cells such as sperm and erythrocytes.

Active transport

Unlike passive transport, active transport requires energy. According to source of energy, active transport is classified to two groups:

- **Primary active transport**
- **Secondary active transport**

Primary active transport

Primary active transport requires energy that comes directly from hydrolysis of ATP. Therefore proteins transporting substances at the expense of ATP hydrolysis are also called ATPases. ATPases are almost exclusively involved in transport of ions though plasma membrane and membranes of some intracellular organelles.

Na^+/K^+-ATPase

Na^+/K^+-ATPase (or **sodium-potassium pump**) is present on the plasma membrane of cells. It pumps three sodium cations out of the cells and two potassium cations from extracellular space into the cells at the expenses of hydrolysis of one ATP molecule. Due to unequal stoichiometry of ion transport ($3Na^+$ versus $2K^+$), Na^+/K^+-ATPase is electrogenic – it leads to the polarisation of cell membrane. Na^+/K^+-ATPase is involved in following cellular functions:

- **Maintaining of the cell membrane potential** - keeping a low concentration of sodium ions and high levels of potassium ions within the cell. In addition, selective permeability of the cell's plasma membrane for the different ions plays an important role in generation of resting membrane potential.
- **Transport of nutrients** - sodium gradient generated by pumping of sodium from the cell provides the driving force for several secondary active symport transporters, which import glucose, amino acids and other nutrients into the cells.
- **Regulation of cell volume** – the cell action of pump is associated with net loss of ions out of the cell, since three Na^+ are transported out and two K^+ into the cell. In addition, the cell membrane that is almost impermeable to sodium ions has higher

permeability for K^+ due to potassium leak channels embedded in the membrane. Therefore K^+ ions leak back out of the cell down their concentration gradient and the higher osmotic pressure of extracellular fluid drives the water out of the cells. Furthermore, the cell swelling activates the Na^+/K^+ -ATPase, which drives still more ions to the extracellular space.

- **Signal transduction –** it has been demonstrated that Na^+/K^+-ATPase can also relay extracellular signalling into the cell through regulation of protein tyrosine phosphorylation.

Inhibitors of Na^+/K^+-ATPase (e.g. digoxin) decrease sodium gradient across plasma membrane of cardiomyocytes (heart muscle cells). Since rate of calcium extrusion from cardiomyocytes via Na^+/Ca^{2+}-exchanger is proportional to sodium gradient, inhibition of Na^+/K^+-ATPase is associated with delayed extrusion of intracellular Ca^{2+} from cardiomyocytes and thus with increased contraction of heart muscle and performance of heart. Therefore, digoxin is used for treatment of chronic heart insufficiency.

H^+/K^+-ATPase

H^+/K^+-ATPase is present on apical membrane of parietal cells of gastric mucosa. It transports one proton out of the cells in exchange for one potassium ion retrieved from the gastric lumen. Due to equal stoichiometry, H^+/K^+-ATPase is electroneutral. **Inhibitors of H^+/K^+-ATPase** (e.g. omeprazol), decreasing proton concentration in gastric juice, are used for treatment of inflammation of gastric mucosa cells (gastritis) and gastric ulcers. Secretion of gastric acid can be also blocked by inhibitors of specific histamine receptor since the binding of histamine to its receptor localised on basolateral membrane of parietal cells triggers production and release of HCl.

Ca^{2+}-ATPases

Ca^{2+}-ATPases are involved in active transport of calcium ions across different membranes. According to localisation they are classified to three different groups:

- **Plasma membrane Ca^{2+}-ATPases**
- **Sarcoendoplasmic reticulum Ca^{2+}-ATPases**

- **Secretory pathway Ca^{2+}-ATPases**

Plasma **m**embrane **C**a^{2+}-**A**TPase (**PMCA**) present on plasma membrane of virtually all cells removes calcium from cells. It transports one calcium ion at expense of hydrolysis of molecule of ATP to ADP. It binds calcium with high affinity (K_M is around 100 nmol/l) but does not remove Ca^{2+} at a very fast rate. Thus it works in cooperation with Na^+/Ca^{2+} exchanger (see further) that is faster but does not bind calcium as effective as PMCA.

Sarco/**e**ndoplasmic **r**eticulum **C**a^{2+}-**A**TPase (**SERCA**) is localised on the membranes of sarco/endoplasmic reticulum of cells. SERCA is the most abundant membrane protein of sarcoplasmic reticulum of heart and muscle cells. It transports 2 ions of calcium through membrane at the expenses of one molecule of ATP.

Secretory **p**athway **C**a^{2+}-**A**TPase (**SPCA**) is expressed at the membrane of Golgi apparatus. Following a calcium spike, SPCA is responsible for transport of Ca^{2+} ions from the cytosol to the lumen of the Golgi apparatus, thus decreasing the cytoplasmic concentrations of Ca^{2+} to resting levels. So, it works in cooperation of SERCA.

ATP-binding cassette transporters

ATP-binding cassette transporters (ABC-transporters) are members of a large superfamily of transmembrane proteins that catalyze active transport of a great variety of endogenous and exogenous substrates across biological membranes. Although they are coded by several genes producing 48 different isoforms, a prominent characteristic of these proteins is that they share a highly conserved domain, the **ATP binding cassette**, which binds and hydrolyzes ATP. There are distinct differences in mechanism of transport among different isoforms. With respect to normal physiology, they transport a wide variety of endogenous substrates across extra- and intracellular membranes, including products of metabolism, phospholipids, cholesterol and other sterols. In addition, they are involved in extrusion of exogenous substances, drugs and toxins, from cells. Some ABC transporters are important from medical point of view. They are often involved in **multidrug resistance** of tumour cell to cytostatics and other antineoplastic agents as well as multidrug resistance of bacteria to antibiotics and chemotherapeutics, cystic fibrosis and a range of other inherited human diseases. Multidrug resistance is important phenomenon associated with extrusion of a

variety of structurally unrelated compounds from the cells what may hinder the treatment of malignant diseases and microbial infections. The most prominent representatives of ABC transporters responsible for multidrug resistance are **P-glycoprotein** and **multi resistance protein** (MRP). Inhibitors of P-glycoprotein are tested in order to overcome resistance of malignant cells to cytostatics to improve efficiency of chemotherapy used for cancer treatment.

Cystic fribrosis transmembrane conductance receptor (CFTR) is member of the family of ABC transporters that after binding of two molecules of ATP transports chloride anions. Some mutations of gene coding for CFTR are cause of autosomal recessive genetic disorder called **cystic fibrosis** that is characterised by production of abnormally thick mucus by epithelial cells of respiratory system and GIT associated with obstruction and bacterial infections of pulmonary airways.

Secondary active transport

Proteins involved in secondary active transport utilize energy of chemical gradient established in majority of cases by the action of ATPases. Transport of some ions and proteins through inner mitochondrial membrane is driven by electrochemical proton gradient established by complexes of mitochondrial respiratory chain.

Na^+/Ca^{2+} exchanger

Na^+/Ca^{2+} exchanger is an antiporter transport protein expressed mainly on cell plasma membrane that removes calcium from cells. Transport of calcium ions out of the cell is driven by sodium gradient established across plasma membrane by the action of Na^+/K^+-ATPase. The Na^+/Ca^{2+} exchanger removes a single calcium ion in exchange for the import of three sodium ions. Due to higher capacity to transport calcium, it represents principal mechanism of extrusion of calcium ions from cardiomyocytes although it has lower affinity for calcium than PMCA. There are some isoforms of Na^+/Ca^{2+} exchanger expressed on inner mitochondrial membrane that is involved in maintenance of low Ca^{2+} concentration in the mitochondria.

Na^+/H^+ exchanger

Na^+/H^+ exchanger is an antiporter transport protein expressed on plasma membrane of many cells (especially in specific cells of kidney) that is involved in maintaining the balance of sodium as well as volume- and pH-regulation.

Increased concentration of protons in blood, **acidosis**, can be associated with reversal of the mode of action of Na^+/H^+ exchanger and consequent decrease of Na^+/K^+-ATPase activity leading to leak of potassium ions from the cells. This leads to elevated blood potassium concentration (**hyperkalemia**) which can be potentially harmful due to depolarisation of plasma membrane of cardiomyocytes which can be associated with **disturbances of heart rhythm** and can lead eventually to **heart arrest**.

Other Na^+-dependent secondary active transporters

Sodium gradient established across plasma membrane by the action of Na^+/K^+-ATPase drives sodium-dependent symport of glucose (by **sodium/glucose cotransporters**) and amino acids (by **sodium/amino acid cotransporters**) through plasma membrane. Both transport systems are expressed mainly in the plasma membrane of intestinal epithelial cells.

Mitochondrial Ca^{2+} uptake

Mitochondrial Ca^{2+} uptake is driven by electrochemical gradient of protons established on inner mitochondrial membrane by mitochondrial electron transport chain. It is involved in transport of calcium from cytoplasm to the mitochondrial matrix.

Transport of amino acids by *γ-glutamyl transferase* system

The γ-glutamyl transferase is an important system for uptake of amino acids by some mammalian tissues (e.g. liver, kidney, heart and brain). It is localised in the plasma membrane and catalyzes attachment of extracellular amino acid to glutamic acid of intracellular glutathione. The resulting dipeptide, **γ-gluatamylamino** acid, is transported into the cell where it is hydrolyzed to amino acid and **5-oxoproline**. Cysteinylglycine produced during reaction is hydrolyzed to cysteine and glycine. Glutathione is regenerated

from 5-oxoproline in the three consecutive ATP-requiring reactions. Thus, the γ-glutamyl transferase consumes **three molecules of ATP** per one transported amino acid. It is more energy consuming system than other mechanism involved in amino acid transport (e.g. sodium-dependent symport) but is faster and has higher capacity. The *γ-glutamyl transferase* has several applications as a **diagnostic marker** in medicine. Elevated serum activity of *γ-glutamyl transferase* can be found in diseases of the liver, in particular in **alcoholic liver diseases**.

Endocytosis

Endocytosis represents important mechanism of transport of some substances through plasma membrane from extracellular space into the cells. It is used by all cells of the body because most substances important to them are large polar molecules that cannot pass through the plasma membrane. Endocytosis is either nonspecific or specific – receptor mediated.

Nonspecific endocytosis is associated with random invagination of transported particle by part of plasma membrane producing endocytotic vesicle.

Specific endocytosis requires binding of transported substance to specific receptor. In majority cases, the next step involves formation of clathrin coated vesicle that is taken up by cell (e.g. transport of LDL to the liver or transport of iron bound to the transferin to cells).

Phagocytosis is a specific form of endocytosis involving the vesicular internalization of solid particles such as bacteria by phagocytes that are white blood cells, protecting human body from foreign particles.

The final step of endocytosis involves translocation of transported molecule or particle to **lysosomes**. The main function of lysosomes is to break down cellular waste products, lipids, saccharides, proteins, and other macromolecules into simple compounds that are then returned to the cytoplasm as new cell-building materials.

Exocytosis

Exocytosis represents important mechanism of transport of some compounds to the plasma membrane (receptors, transporters, surface antigens) or through the plasma membrane out

of cells (hormones, secreted proteins, neurotransmitters). The transported substances are engulfed in **secretory vesicles** that are produced by Golgi apparatus. These membrane-bound vesicles contain soluble proteins or peptides to be secreted to the extracellular environment, as well as membrane proteins and lipids that are sent to become components of the cell membrane.

In multicellular organisms there are two mechanisms of exocytosis:

- **Ca^{2+} triggered non-constitutive (regulated)**
- **non Ca^{2+} triggered constitutive**.

Regulated exocytosis is regulated process that is response to some specific stimuli. It requires an external signal (e.g. signal causing excitation of presynaptic neuron) that is associated with an increase in intracellular Ca^{2+}. Regulated exocytosis is involved in release of hormones from the producing cells to blood or neurotransmitters from presynaptic neurons to synaptic cleft during transmission of signal from presynaptic to postsynaptic neuron.

Constitutive exocytosis is performed by all cells and serves the release of components of the extracellular matrix, or just delivery of newly-synthesised membrane proteins and lipids that are incorporated in the plasma membrane.

Mechanism of exocytosis involves at least three steps:

- **Vesicle trafficking** – delivery of vesicle to the plasma membrane
- **Vesicle docking** – attachment of vesicle to the plasma membrane
- **Vesicle fusion** – fusion of vesicle lipids with membrane lipids followed with incorporation of proteins or lipids into membrane or release of proteins or peptides to extracellular space.

Transepithelial transport

Transepithelial transport is transport of substances or ions through epithelial cells. It depends on asymmetric distribution of transporters in apical and basolateral membrane of epithelial cells. Transepithelial transport is mainly involved in the absorption of nutrients from GIT, like absorption of glucose or amino acids. It is also important for re-absorption

of glucose and amino acids from primary urine back to blood. Production of gastric hydrochloric acid represents another example of transepithelial transport

Blood brain barrier

The **blood brain barrier** is a selective barrier that separates circulating blood from the brain extracellular fluid in the CNS. It consists of tight junctions between endothelial cells around all brain capillaries that restrict the diffusion of solid particles (e.g. bacteria) and large or hydrophilic molecules into the cerebrospinal fluid, while allowing the diffusion of small molecules (O_2, CO_2, hormones). Cells of the barrier allow active transport of nutrients such as glucose and amino acids across the barrier through specific transport proteins (mainly Na^+-dependent symport). Despite its protective role, existence of blood brain barrier might be serious medical problem since it prevents delivery of drugs to CNS. For example, the treatment of brain infections or tumours requires either specific molecules that can cross this barrier or artificial opening of blood brain barrier (using manitol).

4. BIOCHEMISTRY of GENETIC INFORMATION

Transfer of information that is essential for existence, survival and reproduction of all living organisms is one of the most characteristic features of life. Heritable information that is transferred from generation to generation is called **genetic information**. It has been unequivocally documented that nucleic acids are carriers of genetic information and that sequence of nucleotides in some types of nucleic acids represents particular form of record of certain primary genetic information. The term primary expresses the fact that genetic information is not recorded in any other type of biologic macromolecules.

Sequences of nucleotides may contain following information:

- Information about primary structure of all cellular proteins that can be present in DNA and mRNA.
- Information about primary structure of all cellular functional RNAs (tRNA, rRNA, miRNA and snRNA) that is present in DNA.
- Some sequences of DNA and RNA can bind specific proteins. Binding of regulatory proteins to specific sequences can be associated with initiation of replication and with initiation or termination of transcription. In addition binding of regulatory proteins to specific sequences can modulate rate of transcription or translation
- In some types of viruses (retroviruses), sequence of RNA carries information about primary structure of DNA.

Transfer of genetic information is formulated in **central dogma,** which says that genetic information can be transferred from nucleic acid into either nucleic acid or protein but reverse transfer from protein into nucleic acid in not possible. To date, three ways of transfer of genetic information were identified:

1. **Replication** – transfer of genetic information from DNA into DNA or from RNA into RNA.
2. **Transcription** – transfer of genetic information from DNA into RNA. In some retroviruses, genetic information is stored in RNA thus incorporation of viral genetic information into human genomic DNA requires transcription of genetic information from RNA into DNA that is called **reverse transcription**.

3. **Translation** – transfer of genetic information from mRNA into protein primary structure.

Structure and organization of prokaryotic genome

A **genophore** is the DNA of a prokaryote. It is commonly referred to as a prokaryotic chromosome; however, the term "chromosome" does not reflect reality since the genophore lacks chromatin. The compact structure of genophore is maintained through a mechanism known as supercoiling whereas a eukaryotic chromosome is additionally compacted via chromatin (see further).

All or most of the genetic material of prokaryotes is present in the **nucleoid** (meaning nucleus-*like*) that is an irregularly-shaped region within the prokaryotic cell. In contrast to the eukaryotic cell nucleus, it is not surrounded by a nuclear membrane. Generally, the genome of prokaryotic organisms is a circular, double-stranded piece of DNA that may exist in multiple copies at any time. The circular structure of genophore in most prokaryotes allows replication to occur without telomeres (see further). The length of a genophore widely varies. Genophores are generally of a much smaller size than eukaryotic chromosomes (e.g. 580,073 base pairs in *Mycoplasma genitalium*) but size of genophore can be a few million base pairs. It has been evidenced experimentally that the nucleoid is largely composed of DNA, about 60%, with a small amount of RNA and protein. The latter two constituents are likely to be mainly mRNAs and **the transcription factors** (see further). The supercoiled structure of the prokaryotic genophore is maintained by proteins known as **nucleoid proteins** or **nucleoid-associated proteins** that are distinct from histones of eukaryotic nuclei. In contrast to histones, the DNA-binding proteins of the nucleoid do not form nucleosome, in which DNA is wrapped around a protein core, but these proteins often use other mechanisms, such as DNA looping, to promote compact structure of genophore.

In addition to genomic DNA, many prokaryotes contain **plasmid DNA** that is separated from and can replicate independently of the genomic DNA. Plasmids are double-stranded and, in many cases, circular. Plasmids provide a mechanism for horizontal gene transfer within a population of microbes and typically provide a selective advantage under a given environmental state. They may carry genes that provide resistance to naturally occurring

antibiotics in a competitive environmental niche, or the plasmid proteins may act as toxins under similar circumstances. Due to ability of plasmids to replicate independently of genomic DNA and to confer the resistance of bacteria to antibiotics, they are often used as vectors for gene transfer and expression in biotechnology (see further).

Structure and organization of eukaryotic genome

In eukaryotic cells, genomic DNA is localised in nucleus where it is closely associated with proteins and RNA forming nucleoprotein complexes that contain approximately one third of DNA and that are known as **chromatin**. During interphase of cell cycle, most of the chromatin is loose and can be morphologically characterised as tightly packed **heterochromatin** and less dense **euchromatin** that is site of active transcription (see further). However, cell division is associated with condensation of chromatin into **chromosomes** that are visible under light microscope.

Chromatin proteins are classified as either **histones** or **non-histone proteins**. Histones that are directly associated with DNA are small and strongly basic proteins due to high content of basic amino acids (histidine, lysine). They are important component of chromatin structure since positive charge of histones compensates for negative charge of DNA phosphate groups. This compensation allows dense packing of DNA in nucleus. In human diploid cells, 46 molecules of genomic DNA containing **3 x 10^9 base pairs** (bp) of total length 2 meters have to be packed in nucleus with diameter about 10 μm. Histones are also involved in regulation of transcription. Two molecules of histones **H2A**, **H2B**, **H3** and **H4** form an octameric complex around which 146 bp of genomic DNA are wrapped in 1.8 turns. These particles with a diameter 7nm are called **nucleosomes** and connect together by linker DNA (approximately 20-80 bp in length) that interacts with monomer of histone **H1**. Nucleosomes with linker DNA form spirally coiled superstructures with diameters of 30 nm that are known as **solenoids**. During condensation of chromatin to chromosomes, solenoids form 200 nm long **loops**, containing 80 000 bp, that are bound to protein framework, the **nuclear scaffolding**. Twenty loops are organised by framework protein into **minibands**. A large number of stacked minibands forms chromosomes that contain densely packed genomic DNA. The smallest human chromosome contains more than 50 million bp. Histones present in octamer contain N-terminal stretches of 20 amino acids that project out of nucleosomes and are site of covalent modifications involved in regulation of

chromatic structure and gene transcription. The condensation of chromatin into chromosomes is associated with phosphorylation of histones whereas acetylation of free amino group of lysine residues causes relaxation of chromatin and is associated with gene transcription.

Internucleosomal cleavage of genomic DNA by specific nucleases produces DNA fragments of approximate sizes nx180 bp (where n is natural number from 1 to approximately 10) that are one of the characteristic features of **apoptosis**.

The non-histone proteins perform different functions like structural proteins of nucleus, enzymes and transcription factors.

In addition to nuclear DNA, eukaryotes carry in intracellular organelles such as mitochondria (in plants and animals) and chloroplasts (in plants) genophore that is structurally very similar to genophore of true prokaryotes. Some eukaryotes (e.g. *Saccharomyces cerevisiae*) also contain circular plasmid DNA in cytoplasm.

It has been estimated that only 10 % of human DNA is coding for some product. Based on DNA analysis, the sequences of human genome were classified into three groups:

- **Nonrepetitive sequences** that comprise 60 % of total genomic DNA that occur in one or a few copies. Majority of sequences coding for proteins are present in two copies (alleles).
- **Medium repetitive sequences** that comprise around 30 % of DNA sequences and that occur in 10-100 copies throughout genomic DNA.
- **Highly repetitive sequences** that comprise 10 % of total genomic DNA that occur in 100 – 100 000 copies.

Repetitive sequences are further classified into following groups:

- **Tandem repeats** are stretches of identical sequences separated by short sequences called **spacer sequences**. **Long tandem repeats** are usually genes. Human DNA contain 250 tandem ordered copies coding for precursor of 45S rRNA, 2 000 copies for 5S rRNA and 1 300 copies for tRNA. **Short tandem repeats** are constituted of 5 - 200 bases in total length 100 000 bases. These sequences are transcriptionally inactive. Since they are mainly localised in **centromere** and **telomere**, it is supposed that they

have structural function in chromosomes. Many of short tandem repeats are present in less number of copies than those located in centromeres. They are interspersed throughout whole genome and their function is unclear. These sequences are also called **hypervariable regions**, since number of repeats varies significantly among individual people. This is the base of specific test of identification of persons that is called **DNA fingerprinting** (see further).

- **Interspersed repeats** are interspersed throughout whole genome without significant arrangement. Short interspersed repeats consist of hundreds of bp whereas long may contain 5 000 – 7 000 bp. Some of them were produced from RNA that was reversely transcribed into DNA and then incorporated to genome. These sequences are called **pseudogenes**. They contain sequence of original gene but they lack promoter and other regions that are essential for expression.
- **Inverse repeats** are sequences (around 100 - 1000 nucleotides in length) that are into reverse complementary to another sequence further downstream (e.g. 5´-GCTCATG……CATGAGC-3´). Specific form of inverse repeats are **palindromes** with lengths of 4 – 8 nucleotides, which correspond to sequences between complementary strands that are identical when read from the 5' to 3' direction (e.g. 5´-GCTCGAGC-3´ and sequence of complementary strand 3´-CGAGCTCG-5´). Inverse repeats can produce hairpins or loops but their exact role is not known.

REPLICATION – DNA SYNTHESIS

During S phase of cell cycle, genomic DNA must be duplicated producing a complete copy of the genome since after mitosis both mother and daughter cells have to contain complete genome. The process of intracellular duplication of DNA is called **replication**. Given the complementary structure of double stranded DNA, it is relatively easy to understand how DNA as a molecule is well structured for replication. Each strand serves as a template for a synthesis of new strand. It has been documented experimentally that after DNA replication, each new DNA double helix will have one strand from the original DNA molecule, and one newly synthesised molecule. This is referred to as **semiconservative replication**.

Mechanism of DNA replication, enzymes and other factors needed for replication

Prokaryotic replication starts with formation of **replication fork**. Initiation of replication is strictly regulated process. In *E. coli*, only one **initiation site**, called **oriC**, exists on genomic DNA. OriC consists of 245 bp that bind replication protein, **DnaA**. This is followed by unwinding of three AT-rich regions by interruption of hydrogen bonds between bases. The next step involves ATP-dependent binding of **DnaB**, **DnaC** and **DnaT** proteins. DnaB protein exhibits ***helicase*** activity that further unwinds DNA and leads to the formation of replication fork. The supercoiling caused by helicase is relaxed by ***gyrase*** (form of bacterial ***topoisomerase II***) by cutting the DNA strands, allowing it to rotate and release the supercoil, and then rejoining the strands. *Gyrase* is most commonly found upstream of the replication fork, where the supercoils form. Replication fork is region of two separated strands of DNA that is site of replication. Each separated strand of DNA binds s**ingle strand binding proteins** (SSB) that prevents renaturation of DNA (reassociation of both DNA strands)**.** This structure is called **pre-primosome complex**. Finally, **DnaG** will bind to the pre-primosome forming a complete **primosome**. DnaG that exhibits ***primase*** activity attaches 1-10 nucleotide RNA primer to the single stranded DNA creating a DNA-RNA hybrid. This sequence of RNA is used as a primer to initiate binding of ***DNA polymerase III*** and **elongation** of DNA strands. The primosome is utilised once on the leading strand of DNA and repeatedly, initiating each Okazaki fragment, on the lagging DNA strand.

Elongation of synthesised DNA occurs simultaneously on both antiparallel parental DNA strands and in most prokaryotic organisms in both directions. **Leading strand** is elongated continuously in the direction 5'→ 3' by attachment of complementary nucleotides according to sequence of parental strand that is read in the direction 3´→ 5´, which is the same with movement of replication fork proteins. Synthesis on antiparallel strand is complicated since replication fork proteins moves in direction 5'→ 3' with respect to this strand. Therefore elongation of this strand called **lagging strand** is **discontinuous** and occurs in the form of short fragments called **Okazaki fragments**. Synthesis of each Okazaki fragment starts with synthesis of RNA primers by *primase*. Primers are further elongated by ***DNA polymerase III***. RNA primers are removed by the ***flap endonuclease*** (FEN1) or ***RNase H*** and so produced gaps between Okazaki fragments are filled *by* ***DNA***

polymerase I. Finally, Okazaki fragment are connected by ***DNA ligase***. Since prokaryotic genomic DNA is circular and replication occurs in both directions there are any problems associated with termination of replication.

The DNA **replication in eukaryotes** is similar to prokaryotic replication but is more complex. The genome of humans contains about **10 000 origins of replication** distributed throughout the chromosomes at 30 000 to 300 000 bp intervals. The rate of eukaryotic replication is about 10 times slower than the rate of prokaryotic replication however the multiple origin of replication ensures that whole genome is replicated within a reasonable time period (about 8 hours). Eukaryotes contain at least **three different nuclear *DNA polymerases*** and one **mitochondrial *DNA polymerase*** responsible for replication of mitochondrial DNA (Table 4).

Table 4. Properties of eukaryotic *DNA polymerases*

	DNA pol α	***DNA pol β***	***DNA pol δ***	***DNA pol γ***
Location	nucleus	nucleus	nucleus	Mitochondria
5'→ 3' exonuclease	-	-	-	-
3′→ 5′ exonuclease	-	-	+	+
Primase	+	-	-	-
Function	Replication of lagging strand	DNA repair	Replication of leading strand	Replication of mtDNA

Important difference between prokaryotic and eukaryotic replication is associated with **replication of telomeres. Telomeres** are stretches (approximately 1000) of six nucleotide tandem repeats (human telomere sequence: TTAGGG) located at the end of linear eukaryotic chromosomes. They protect end of the chromosomes from degradation and fusion with neighbouring chromosomes by **homologous recombination** (see further) or **non-homologous end joining** (see further) due to the essential "capping" role of telomeres that distinguishes them from DNA double-strand breaks. In addition, telomeres compensate for incomplete semi-conservative DNA replication at chromosomal ends. Otherwise, the cell without telomeres would lose the ends of their chromosomes, and possibly the necessary information they contain.

Telomeres are replicated by ***telomerase*** that is a "ribonucleoprotein complex" composed of a protein component and an RNA primer that is complementary to telomere sequence and is essential for initiation of telomere replication. The *telomerase* binds to the 3'-end of the DNA strand and the RNA template of *telomerase* (sequence CCCUAACCC), pairs with GGG triplet. A six-nucleotide repeat TTAGGG is synthesised by means of ***reverse transcriptase*** activity of *telomerase*. Then the *telomerase* dissociate from binding to DNA and reassociate with synthesised GGG triplet to add another TTAGGG hexamer.

In most multicellular eukaryotic organisms, *telomerase* is active only in germ cells, stem cells, and some white blood cells while telomerase activity in human somatic cells is either undetectable or inadequate for telomere maintenance. Since *DNA polymerase* is not capable of complete replication of telomeres, 50 nucleotides are lost during each cell cycle, which results in gradual **telomere length shortening**. In human somatic cells, proliferation potential is strictly limited to approximately 50-70 cell divisions. Critically short telomeres cause **senescence** associated with **cell cycle arrest** that is followed by **cell death**. Multiple studies have shown that telomere length in peripheral white blood cells inversely correlates with increasing age however it is not yet clear whether shorter telomeres are just a sign of **ageing** or actually contribute to ageing. On the other hand, approximately 85%–90% of human **cancer** cells have detectable activity of *telomerase* that is responsible for telomere length maintenance in these cells thus making them immortal with unlimited potential of replication and proliferation. The critical role of telomerase in facilitating and enabling unlimited proliferation of cancer cells is also documented by the fact that inhibition of telomerase activity in *telomerase*-positive cancer cells results in cell death and tumour growth inhibition.

Drugs that affect replication

Some bacterial and antiviral drugs and many chemotherapeutic drugs affect replication. They are classified according to the mechanism by which they affect replication.

Antimetabolites

An **antimetabolite** is a chemical compound that inhibits the use of a metabolite, which is product of normal metabolism and is often similar in structure to the metabolite that it interferes with. Antimetabolites affecting DNA replication inhibit production of substrate

for replication (dNTPs). Decreased substrate availability is associated with decreased rate of replication. Typical examples of therapeutically important antimetabolites, **5-fluorouracil** and **mercaptopurine**, are direct inhibitors of dNTPs synthesis. In addition, **methotrexate** inhibits *difydrofolate reductase* that is essential for synthesis of thymidine monophosphate.

Substrate analogues

Many analogues of dNTPs can be incorporated into DNA by DNA polymerases and then inhibit further replication. **Azidothymidine** and **cytosine arabinoside** are typical examples of therapeutically important substrate analogues.

Drugs interacting directly with DNA

Drugs interacting directly with DNA include **intercalating agents** or **drugs that damage DNA**.

Intercalating agents are drugs that insert between DNA strands. They disrupt or change DNA conformation that is associated with inhibition of replication. Many of them are mutagenic. **Anthracycline glycosides**, **daunorubicin** and **doxorubicin**, have wide application in human medicine. Daunorubicin is used for treatment of acute myeloblastic leukaemia while doxorubicin is used to treat different types of solid tumours. Cardiotoxicity of anthracyclines represents serious side effect of anthracyclines. Another intercalating agent, **actinomycin D**, is an antibiotic used for treatment of different types of cancer.

Drugs that damage DNA are **alkylating agents, platinum complexes** and **bleomycins.**

Alkylating agents attach the alkyl group to the nitrogen atom number 7 of guanine base of DNA. Alkylation of guanine is associated with mispairing during replication, breakage of DNA and crosslinking of the double helix. The later processes are associated with initiation of apoptotic cell death. Several different alkylating agents (e.g. cyclophosphamine, chlorambucil, streptozocin, busulphan) of different chemical nature are used for treatment of wide variety of cancers.

Platinum complexes (e.g. cisplatin, carboplatin, oxaliplatin) that cause crosslinking of the DNA double helix are used for treatment of solid tumours (e.g. colorectal or breast cancer).

Bleomycins bind to DNA and together with oxygen and Fe^{2+} cause DNA breakage. They are used for treatment of wide variety of cancers.

Inhibitors of replicative enzymes

Inhibitors of *topoisomerase*

Inhibitors of human *topoisomerase I* (e.g. irinotecan, topotecan) and *topoisomerase II* (e.g. etoposide) are used for treatment of different types of cancer whereas inhibitors of bacterial *topoisomerase II* (e.g. ciprofloxacin) are used for treatment of bacterial infections, mainly infections of urinary tract.

Factors that affect DNA structure

Although DNA is essential molecule, it is not resistant to different endogenous and exogenous conditions. Factors that affect DNA structure can be classified as **physical**, **biologic** and **chemical**. **Ionizing radiation** and **UV ligh**t are the most important physical factors, causing double strand brakes of DNA. In addition to chemical factors causing hydrolysis or denaturation of DNA, **chemical mutagens** are the most important factors affecting DNA structure. Chemical mutagen is each compound that increases the probability of DNA mutations upon exposure. Chemical mutagens affect the DNA structure in multiple ways. Some of them must be incorporated in DNA during replication (e.g. base analogues) while others change directly structure of nitrous bases (for example nitrous acid causing deamination of cytosine to uracil or alkylating agent causing alkylation of bases). Finally **intercalating agents** such as ethidium bromide or acridine dyes are molecules that are inserted between DNA strands, causing frameshift mutations during replication. Some chemical compounds became to be mutagenic after their chemical transformation in the human body. For example aflatoxins naturally occurring mycotoxins that are produced by many species of *Aspergillus* (e.g. *Aspergillus flavus*) are metabolised by the liver to a reactive epoxide intermediate that is strong mutagen causing liver carcinoma. Some chemical mutagens such as base analogues and alkylating agents are used for treatment of different types of cancer. Biologic factors affecting structure of DNA are either **bacteria** or **viruses**. For example, bacteria *Helicobacter pylori* causes inflammation of gastric parietal cells during which oxidative species are produced that reduce efficiency of DNA repair systems thereby increasing mutations and probability of

gastric cancer. On the other hand, mutations caused by viruses are associated with direct incorporation of viral genetic material into genomic DNA.

Mechanisms of DNA repair

Since DNA is under permanent attack of different factors that can alter structure and thus function of DNA, each cell is provided with several systems responsible for repair of DNA. According to the mechanism of their action, DNA repair systems are classified to three groups; **direct repair**, **excision repair** and **recombinant repair**.

Direct repair

Mechanism of direct repair involves direct reverse of modified base to original base. Dimers of thymine and cytosine bases that can significantly alter DNA structure and that are produced mainly by UV light are converted back to monomers by the action of specific enzyme ***photolyase***.

Some chemical mutagens called alkylation agents cause alkylation of guanine to O^6-alkylguanine that is associated with incorrect pairing with thymine instead of cytosine. Alkylated bases can be directly repaired by group of enzymes called ***alkyltransferases*** that transfer alkyl from bases to their own molecule. However, some alkylating agents (e.g. temozolomid) are used for treatment of some types of cancer (e.g. brain tumours) and high level of *alkyltransferases* in tumours is associated with poor response to treatment.

Excision repair

Removal of modified nucleotide and its substitution by nucleotide complementary to unmodified strand is a common mechanism of excision repair that can involve three different processes.

Base excision repair (BER) is characterised by removal of modified bases (oxidised, alkylated and deaminated) by the group of enzymes called ***DNA glycosidases***. The next step involves removal of remained deoxyribose by ***AP endonuclease***. Then, *exonuclease* removes short oligonucleotide. The resulted gap is filled by the action of *DNA polymerase*

and the synthesised oligonucleotide that is complementary to unmodified strand is joined to the rest of repaired strand by *DNA ligase*.

Nucleotide excision repair (NER) involves removal of short oligonucleotide encompassing modified base (mainly pyrimidine dimers) by ***excision endonuclease***, synthesis new oligonucleotide by *DNA polymerase* and the final connection by *DNA ligase*.

Mismatch repair is a system that recognizes and repairs erroneous insertion, deletion and mis-incorporation of nucleotides that can arise during replication and recombination. Mechanism of mismatch repair involves binding of specific **Mut protein** to mismatched base, excision of short oligonucleotide encompassing mismatched base by specific *enxonucleases*, synthesis of new oligonucleotide by *DNA polymerase* and the final connection by *DNA ligase*.

Repair of double strand breaks

Double strand breaks of DNA that are result of exposition of DNA to ionizing radiation or UV light are repaired by two different mechanisms, **homologous recombination** or **nonhomologous end joining**.

In homologous recombination, the sequences of DNA around the 5´-ends of the break are cut away by specific endonucleases. In the next step called **strand invasion**, an overhanging 3´-end of the broken DNA molecule combines with a similar or identical DNA molecule that is not broken and provides a template for repair. The invading 3' strand is extended along the recipient DNA duplex by *DNA polymerase*. The newly synthesised 3'-end of the invading strand is then able to anneal to the second 3'-end overhang in the damaged DNA through complementary base pairing. This is followed by addition of missing nucleotide sequences and ligation of repaired strands.

In nonhomologous end joining the break ends are directly ligated by *DNA ligase IV* without the need for a homologous template.

Hereditary DNA repair disorders

Several different hereditary diseases are associated with defects in DNA repair mechanisms. Defects in the NER mechanism are responsible for several genetic disorders. The most known disorder called **Xeroderma pigmentosum** is associated with hypersensitivity to sunlight or UV light, resulting in increased skin cancer incidence and premature ageing.

Other disorders of NER include **Cockayne syndrome** (hypersensitivity to UV light and chemical agents) and **trichothiodystrophy** (sensitive skin, brittle hair and nails) that are often accompanied by mental retardation, suggesting increased vulnerability of brain in development neurons.

Repair of mitochondrial DNA is less effective then repair of genomic DNA, therefore progressive age-dependent accumulation of mtDNA mutations has been observed and is considered as one mechanism responsible for mitochondrial dysfunction during ageing.

TRANSCRIPTION – RNA SYNTHESIS

Transfer of genetic information stored in DNA into protein primary structure requires **transcription** of DNA sequences into sequences of functional RNAs.

Mechanism of transcription

In prokaryotes, transcription begins with the binding of specific protein, called **σ factor**, to the promoter sequences on DNA located -35 and -10 downstream of promoter sequences. Promoters are regions of DNA that promote transcription. The σ factor initially recognizes and binds to -35 sequence TTGACA called **recognition sequence**. At the start of initiation, the core enzyme of *RNA polymerase* is associated with σ factor. Combination of the σ factor with *RNA polymerase* core enzyme (consisting of five subunits: 2 α subunits, 1 β subunit, 1 β' subunit, and 1 ω subunit) forms a holoenzyme. Sequence -10 downstream of promoter sequences called **Pribnow box** (TATAAT) facilitates the localised DNA unwinding that is essential prerequisite to synthesis a new RNA.

In eukaryotes, the initiation of transcription requires binding of specific proteins called **basal transcription factors** to a core promoter sequence in the DNA that are found in eukaryotes at -30, -80, and -90 base pairs upstream from the **transcription start site** (TSS). Core promoters are sequences within the promoter that are essential for transcription initiation. Different promoters and transcription factors are utilised by the *RNA polymerases I, II* and *III*. ***RNA polymerase I*** synthesizes a pre-rRNA 45S. ***RNA polymerase II*** synthesizes precursors of mRNAs and most snRNA and microRNA. ***RNA polymerase III*** synthesizes precursors of tRNAs, rRNA 5S and other small nuclear and cytosolic RNAs.

In this text book we will focus on the synthesis of pre-mRNA by RNA polymerase II that is the most studied type, and transcribes majority of genes. The most characterised type of core promoter in eukaryotes is a short DNA sequence known as a **TATA box** that is located 25-30 bp upstream from the TSS and has consensus sequence TATAAAA. The second important element called **CAAT box** is located near the position -80 and has consensus sequence GGCCAATCT. The transcription starts with binding of a transcription factor known as **TATA-binding protein** (TBP), which is itself a subunit of another transcription factor called **transcription factor II D** (TFIID), to the TATA box. After TFIID binds to the TATA box via the TBP, five more transcription factors (**TFIIA**, **TFIIB**, **TFIIF**, **TFIIH** and **TFIIJ**) and *RNA polymerase II* combine around the TATA box forming a preinitiation complex. Transcription factor II H (TFIIH) consists of two subunits with helicase activity and so is involved in the separation of opposing strands of double-stranded DNA to form the initial transcription bubble. However, only a low, or basal, rate of transcription is driven by basal transcription factors. Other specific transcription factors (see further) along with associated coactivators (called **enhancers**) or corepressors (called **silencers**) are responsible for modulation of transcription rate.

Elongation is practically the same comparing prokaryotes and eukaryotes. One strand of the DNA, the template strand (or non-coding strand), is used as a template for RNA synthesis. As transcription proceeds, *RNA polymerases* traverse the template strand and uses base pairing complementarity with the DNA template to create a new RNA. Although *RNA polymerases* read the template strand in the 3' → 5' direction, the synthesis of newly-formed RNA occurs in the 5' → 3' direction with the addition of ribonucleotides to the 3´-hydroxyl group of ribose at the end of the chain. This produces an RNA molecule from 5'

→ 3' that is an exact copy of the coding strand except that thymines are replaced with uracils. Unlike DNA replication, mRNA transcription involves often multiple *RNA polymerases II* on a single DNA template and multiple rounds of transcription, so many mRNA molecules can be rapidly produced from a single copy of a gene (amplification of particular mRNA). In eukaryotes, early during elongation, the 5´-ends of pre-mRNA are modified by the addition of **7-methyl guanosine caps** that are added when the growing RNA chain is long only about 30 nucleotides. The caps that contain unusual **5´-5´ triphosphate** bond are recognised by specific proteins (initiation factors) involved in initiation of translation (see further) and also protect the growing chain from degradation by nucleases.

The incorrectly incorporated bases during elongation can be replaced by a proofreading mechanism. In eukaryotes, this may correspond with short pauses during transcription that allow appropriate RNA editing factors to bind. These pauses may be intrinsic to the *RNA polymerase* or due to chromatin structure.

Prokaryotes use two different strategies for transcription termination. In the first type, RNA transcription stops when the newly synthesised RNA molecule forms a G-C-rich **hairpin loop** followed by a stretch of uracils. Formation of hairpin is associated with mechanical break of the weak rU-dA bonds of the DNA-RNA hybrid. This pulls the poly-U transcript out of the active site of the *RNA polymerase* causing termination of transcription. In the second type of termination, a binding of specific protein called ρ-factor to specific sequences destabilizes the interaction between the template and the mRNA, thus releasing the newly synthesised mRNA from the elongation complex.

Transcription termination in eukaryotes is less understood but involves cleavage of the new transcript followed by template-independent addition of stretch of adenosine at its new 3'-end, in a process called **polyadenylation**.

Post-transcriptional RNA processing

All types of cellular RNAs are synthesised as non-functional precursors that have to be further processed to yield functional RNA. Post-transcriptional RNA processing depends on the type of RNA.

Production of functional tRNA

The tRNAs are synthesised as large precursors, both in prokaryotes and eukaryotes. These precursors often contain the sequences for more than one tRNA, which are cut out and reduced in size by ***specific ribonucleases***. The genes of some tRNA molecules contain single intron that is 10-40 nucleotides long and that is located near the sequences of the anticodon loop. These **tRNA introns** are transcribed. Therefore, the processing of tRNA precursors must include removal of the introns by **splicing** to generate active tRNA molecule for protein synthesis.

The further modifications of tRNA molecules include nucleotide alkylation and the attachment of the canonic **CCA sequence** at the 3´-end of tRNA molecule. The alkylation of tRNA precursors occurs in nucleus whereas splicing of intron and attachment of CCA sequence is located in cytoplasm.

Production of functional rRNA

The rRNA is synthesised as 45S pre-rRNA, which is cleaved into mature 28S, 18S and 5.8S rRNAs that are the major RNA components of the ribosomes.

Maturation of mRNA

The mRNA is synthesised as large precursor pre-mRNA that has to undergo three major processing events to produce mature mRNA that is transported from nucleus to cytoplasm. The steps of mRNA maturation comprise **capping** of 5´-end (), **splicing** out of introns from within the body of the pre-mRNA, and the **polyadenylation** of a 3´-end.

Processing of mikroRNA (miRNA)

Specific genes coding for different miRNAs are transcribed to specific stem loop containing **primary miRNA** (**pri-miRNA)** that consists of 500 – 3000 nucleotides and that is further processed by specific enzyme complex called ***Drosha*** to **pre-miRNA** (containing 70 nucleotides). After transport from nucleus, pre-miRNA is processed by specific enzyme complex called ***Dicer*** to mature miRNA that contains approximately 22 nucleotides.

RNA as an enzyme

Some RNA molecules with a well defined tertiary structure that are able to catalyze a chemical reaction are called **ribozymes (ribo**nucleic acid en**zymes)**. It contains an active

site that consists entirely of ribonucleotides. Many natural ribozymes cleave one of their own phosphodiester bonds (self-cleaving ribozymes), or phosphodiester bonds in other RNAs. Some have been found to catalyze different reactions, for example the aminotransferase activity of the ribosome.

Inhibitors of transcription

Transcription inhibitors can be used as antibiotics against pathogenic bacteria (antibacterials) and fungi (antifungals). For example, antibacterial **rifampicin** inhibits prokaryotic DNA transcription into mRNA by binding to β-subunit of *RNA polymerase* and **8-hydroxyquinoline** is an antifungal transcription inhibitor.

Amanitin that is potent inhibitor of *RNA polymerase II* is deadly toxin found in several species of the *Amanita* genus of mushrooms (e.g. *Amanita phalloides*). Around 15% of persons poisoned with *Amanita* mushrooms will die within 10 days. Amanitin can also be used to determine type of *RNA polymerase* by testing the sensitivity of the polymerase to α-amanitin. *RNA polymerase I* is insensitive, *RNA polymerase II* is highly sensitive (inhibited at 1μg/ml), and *RNA polymerase III* is moderately sensitive (inhibited at 10μg/ml).

TRANSLATION – PROTEIN SYNTHESIS

Characteristics of genetic code

Protein synthesis represents the final step of transfer of genetic information stored in DNA into protein primary structure. In fact, protein synthesis can be considered as the **translation** of genetic information from the language of nucleotides in to language of amino acids. The set of rules by which information encoded in nucleotide sequences of DNA or mRNA is translated into protein primary structures is called **genetic code.** The basic unit of genetic code is **codon** that is triplet of nucleotides coding for one amino acid. Due to this fact, DNA and proteins exhibit **colinearity** that means dependence of protein primary structure on sequence of coding regions of DNA.

Translation starts from **start codon** (**AUG**) that is bifunctional since it is important for initiation of translation and specifies for amino acid, methionine. The **termination codons** (**UAA**, **UAG**, and **UGA)** do not specify for any amino acid and are important for termination of translation. In some instances **UGA** codon is bifunctional, since it can specify a specific amino acid, selenocysteine.

General properties of genetic code:

Genetic code of all organisms exhibit three common general properties and is:

- **Universal** – genetic code is valid for all organisms (except of mitochondrial genetic code).
- **Degenerated** – majority of amino acids are coded by more codons (only methionine and tryptophan are coded by one codon).
- **Non-overlapping** – one nucleotide of coding sequence is present only in one codon.

Despite the fact that genetic code is non-overlapping, there are three possible combinations of reading of genetic information. However, the correct reading of coding sequence starts with AUG codon that is located close to the specific sequences (**Kozak** sequence in eukaryotes and **Shine-Dalgarno** sequence in prokaryotes) important for proper position of mRNA on ribosome during translation. Reading of sequence can lead to two different reading frames.

Open reading frame is determined by AUG (ATG when DNA is analysed) and stop codon giving a polypeptide of proper size.

Closed reading frame is determined by AUG and stop codon, giving a short non-functional polypeptide.

The concept of open reading frame (ORF) is important in terms of computer analysis of DNA sequences. For example, analysis of human DNA sequences revealed approximately 26 000 ORFs in human genome (that is much less than anticipated, 55 000 in 1990s or 35 000 the first estimation from complete human genome sequence in 2001).

Interaction codon (mRNA) – anticodon (tRNA), wobble pairing

If each tRNA molecule paired with its complementary mRNA codon is using canonical **Watson-Crick base pairing** (complementary pairing), then 61 types of tRNA molecule would be required (since three of 64 codons are stop codons). So canonical pairing would require 61 species of tRNA, however, most organisms have fewer than 45 species of tRNA. Therefore, some tRNA must pair with more than one codon. This fact was explained by Francis Crick who proposed the **wobble pairing** hypothesis. He postulated that the 5´-base (first base) on the anticodon, which makes pair with the 3´-base (third base) of codon on the mRNA could have non-standard base pairing (Wobble pairing) that is shown in Table 5.

Table 5. Wobble pairing

First base of anticodon	**Prokaryotes** **Third base of codon**	**Eukaryotes** **Third base of codon**
C	G	G
A	U	U
U	A or G	A
G	U or C	U or C
I	U, C or A	U or C

Classification of mutations according to their impact on protein primary structure

Mutation is each sudden and spontaneous change of DNA nucleotide sequence that is permanent and that can be transmitted to next generation. According to their impact on protein primary structure mutations occurring in coding sequence of mRNA can be classified to following groups.

Silent mutation is the one in which a base was changed but because of the degeneration of the genetic code the same amino acid is coded and therefore primary structure of particular protein is not affected.

Conservative missense mutation is the one in which a base was changed in the way that mutated codon is coding for amino acid that belongs to the same group of amino acid as original amino acid (e.g. serine substituted by threonine). Primary structure of particular protein is changed but because of similar chemical properties of changed amino acid such type of mutation has often minimal impact of protein function.

Nonconservative missense mutation is the one in which a base was changed in the way that mutated codon is coding for amino acid that belongs to the completely different group of amino acid as original amino acid (e.g. serine substituted by proline). Primary structure of particular protein is changed and because of significantly different chemical properties of changed amino acid such type of mutation has often strong impact of protein function. For example, **sickle cell anaemia** is caused by nonconservative missense mutation in β-globin gene. Codon GAG coding for glutamic acid is changed to GTG that is coding for valine. Such subtle change has strong impact not only on haemoglobin function (decreased capacity to bind oxygen) but is also associated with a dramatic change of erythrocyte shape. Instead normal oval and flexible biconcave disk, affected erythrocytes are of sickle shape and rigid. Therefore the sickle erythrocytes have tendency of log jamming, sticking and accumulating at the branching point in a vein.

Nonsense mutation is the one in which a base was changed in the way that mutated codon is not coding for any amino acid (stop codon). Such type of mutation is causing synthesis of truncated protein that may often exhibit significantly different biologic activity. For example, nonsense mutation within the gene coding structural protein dystrophin is cause of severe inherited disease called **Duchenne muscular dystrophy**.

Frame-shift mutation is the one in which some base or bases were added or removed from coding sequence causing a change in reading of genetic information (shift of reading frame). Deletion of UUU codon in the cystic fibrosis transmembrane conductance receptor gene is a cause of an autosomal recessive genetic disorder, **cystic fibrosis**.

There are some other types of mutations affecting primary structure of proteins like **gene translocation** associated with production of aberrant fusion proteins (Bcr-Abl fusion *protein kinase* associated with development of chronic lymphoblastic leukaemia) or mutations in introns that can affect splicing of mRNA primary transcript.

In addition to mutations that have impact on protein primary structure, some other mutations are important from medical point of view. These include mutations in promoter sequences or chromosome rearrangement associated with up- or down-regulation of expression of particular gene (e.g. t(14;18) chromosomal translocation associated with over-expression of proto-oncogene Bcl-2 identified in 70% of human follicular B-cell lymphomas).

Mechanism of translation

As it was already mentioned, mRNA serves as template for translation that occurs at translation apparatus, ribosomes. The **70S ribosomes** are operating in prokaryotes while eukaryotes are using **80S ribosomes**. The rate of translation is 10-20 amino acids/s that is quite low however the rate of translation is significantly increased by the fact that usually 10-15 ribosomes associate with one mRNA. The fidelity of translation is high, 1 error per 2000 incorporated amino acids, and is increased by fact that incorrectly synthesised protein is submitted to degradation.

In prokaryotes, all proteins are synthesised in cytoplasm. In eukaryotes, translation of all proteins begins in cytoplasm. Cytoplasmic, nuclear, mitochondrial and peroxisomal proteins are completely synthesised in cytoplasm, while synthesis of secretory, plasmalemal and lysosomal proteins continues at **rough endoplasmic reticulum** and **Golgi apparatus**. Some mitochondrial proteins (13 polypeptides in human beings) coded by mtDNA are synthesised in mitochondria on 70S ribosomes that are similar to prokaryotic ribosomes. In plants, protein synthesis occurs also in chloroplasts on 70S ribosomes.

Translation consists of five independent steps.

- **Activation of amino acids** – synthesis of aminoacyl-tRNA
- **Initiation** – formation of initiation complex
- **Elongation** – connection of amino acids according to sequence of codons in mRNA
- **Termination** – end of protein synthesis

- **Modifications of synthesised polypeptide** – additional chemical modifications of polypeptide to produce functional polypeptide

Since initiation of translation in eukaryotes is complicated process we will focus only on prokaryotic translation in this text book.

Activation of amino acids

Activation of amino acids is associated with attachment of amino acid to tRNA that bears corresponding anticodon. The reaction is catalysed by the group of enzymes called ***aminoacyl-tRNA synthase***; each enzyme is specific for one amino acid. Activation occurs in two steps. In first step, amino acid reacts with ATP producing **aminoacyl adenylate**. In the second step, aminoacyl residue is attached to 3´-hydroxy group of the terminal ribose of tRNA producing **aminoacyl-tRNA**.

Initiation

Initiation of translation involves formation of **initiation complex** that consists of ribosomes, mRNA, formylmethionine-tRNA and requires specific proteins called **initiation factors**. Initiation in prokaryotes involves several steps. In the first step, initiation factors IF-1 and IF-3 bind to the 30S subunit of ribosome. During the next step, initiation factor IF-2 is first charged by GTP and this complex binds to 30S subunit of ribosome. In the third step, the subunit associates with mRNA and releases IF-3. The right position of mRNA on 30S subunit is ensured by the interaction of **Shine-Dalgarno** sequence (AGGAGG) located in the 5´-UTR of mRNA (8 nucleotides upstream from start codon) with complementary sequence on rRNA. Finally, formylmethionine-tRNA binds to the start codon producing preinitiation complex and functional initiation complex is formed after binding of 50S subunit of ribosome to preinitiation complex a release of IF-1, IF-2 and GDP.

In the initiation complex, formylmethionine-tRNA is bound to the binding site known as **peptidyl site** (P-site). The second binding site located at ribosome, the **acceptor site** (A-site), in not yet occupied during initiation. The third active site, the **exit site** (E-site), is the site from which uncharged tRNA leaves ribosome.

Elongation

Once initiation complex is formed, the peptide chain is elongated by the addition of further amino acids according to sequence of codons on mRNA until ribosome reaches a stop codon.

Elongation involves several steps. In the first step, aminoacyl-tRNA corresponding to the next codon binds to **elongation factor Tu** (EF-Tu) that is charged by GTP. The next step is associated with binding of aminoacyl-tRNA to the codon that is located on the A-site of ribosome. The binding is accompanied with hydrolysis of GTP to GDP and further dissociation of complex GDP-EF-Tu from ribosome. EF-Tu is regenerated after exchange of GDP for GTP that is catalysed by **elongation factor Ts** (EF-Ts). The next step involves formation of peptide bond. Ribosomal ***peptidyl transferase*** catalyzes transfer of the peptide chain from the tRNA at the P-site to the amino group of the aminoacyl-tRNA at the A-site. After the transfer of elongating peptide to A-site, the free tRNA at the P-site dissociates and another **elongation factor G** (EF-G) charged with GTP binds to the ribosome. Hydrolysis of GTP provides energy for translocation of ribosome three bases along mRNA in the direction of 3´-end. The translocation causes appearance of the next codon at the A-site while elongating polypeptide bound to tRNA reaches P-site. The free tRNA dissociates from the E-site of ribosome.

Termination

End of protein synthesis occurs as soon as one of the three stop codons appears at the A-site of ribosome. Since there are not corresponding tRNA for stop codons, two **releasing factors** bind to the ribosome. One of these factors (RF-1) catalyses hydrolysis of ester bond between tRNA and C-end of polypeptide. The synthesised polypeptide is released from ribosome complex that completely dissociates to the components. The energy for dissociation is supplied by the factor RF-3 that hydrolyzes another molecule of GTP.

The **energy requirements** of translation are high since two macroergic bonds are consumed during amino acid activation and two GTP molecules are hydrolyzed during elongation. In addition, both initiation and termination require one GTP.

Differences in translation between prokaryotes and eukaryotes

- Ribosomes – eukaryotes 80S – prokaryotes 70S
- Initiation – first amino acid in eukaryotes is methionine (instead of **formylmethionine** in prokaryotes) and eukaryotic initiation is controlled by 11 initiation factors (3 IF in prokaryotes)
- Elongation is the same just using eukaryotic elongation factors **eEF-1α** (analogue of EF-Tu), **eEF-1βγ** (analogue of EF-Ts) and **eEF-2** (analogue of EF-G)
- Termination of translation in eukaryotes requires only one release factor (2 RF in prokaryotes)

Synthesis of secretory and membrane proteins

Synthesis of secretory and membrane proteins starts in cytoplasm however as soon as **signal peptide** is synthesised, **signal recognition particle** binds to signal peptide. This is followed by binding of complete translational complex to specific receptor localised on membrane of rough endoplasmic reticulum. Translation continues in the way that elongated polypeptide penetrates into lumen of endoplasmic reticulum. After synthesis of part of polypeptide, signal peptide is cut out by means of specific enzyme called ***signal peptidase*** and translation continues till complete polypeptide is synthesised and translation is terminated. Termination of translation is associated with ribosome dissociation. Synthesised polypeptide is further processed in endoplasmic reticulum and Golgi apparatus. Finally, mature polypeptide is coated into specific **vesicles** and either secreted out of the cell or delivered to plasma membrane. Some proteins synthesised on rough endoplasmic reticulum are delivered to lysosomes.

Targeting of synthesised proteins

In eukaryotes, proteins synthesised in cytoplasm have to be delivered to the site of their function. This is enabled by specific amino acid sequences at N-end of polypeptide that serves as a signal for delivery of protein to particular cell compartment.

Modifications of synthesised polypeptide

Depending on timing, modifications of polypeptide chains are classified to two groups.

- **Cotranslational modifications** are the modifications that occur simultaneously with elongation of incomplete polypeptide.
- **Posttranslational modifications** are the modifications of completely synthesised polypeptide

The most important cotranslational modifications are:

- Deformylation of initial methionine that occurs in prokaryotes, mitochondria and chloroplasts
- Removal of N-terminal amino acids (initial methionine, signal peptide)
- Formation of disulfide bonds
- Chemical modifications of amino acids (phosphorylation, methylation, acylation, hydroxylation, glycosylation, iodation)
- Formation of secondary and tertiary structure – folding.

The most important posttranslational modifications are:

- Cut out of peptides (activation, e.g. proinsulin – insulin, pepsinogen-pepsin) and proteolytic degradation of proteins
- Attachment of prosthetic group (e.g. heme to globin)
- Formation of quaternary structure
- Folding of proteins and assembly of protein complexes by means of chaperones
- Chemical modifications affecting activity of proteins (phosphorylation, hydroxylation, γ-carboxylation).

Disturbances of translation

Disturbances of translation are rare. Clinically most important are disturbances of **mitochondrial translation** (mainly due to mutations of genes coding for mitochondrial tRNA). The most affected organs are that containing high amount of mitochondria like heart (decreased cardiac output), muscles (weakness, atrophy), CNS (different neurologic symptoms) therefore the common name of such diseases is **mitochondrial cardio-, myo-, encephalopathies**

Inhibitors of translation

The most common inhibitors of translation are antibiotics and antimetabolites. They affect different processes on different sites (Table 6). Due to differences between prokaryotic and eukaryotic translation, specific inhibitors of prokaryotic translation can be used in human medicine as powerful antibiotics for treatment of bacterial infections.

Table 6. Inhibitors of translation

	Affected process	Site of effect
Streptomycin	Initiation, elongation	30S subunit
Neomycin	Nonspecific	Nonspecific
Tetracycline	Binding of AA-tRNA	30S subunit
Puromycin	Premature termination	70S a 80S ribosomes
Erythromycin	Translocation	50S subunit
Cycloheximid	Nonspecific	80S ribosome
Ricin	Nonspecific	60S subunit
Diphtheria toxin	Elongation	Inactivation of eEF-2

REGULATION OF GENE EXPRESSION IN EUKARYOTES

In multicellular organisms, complete genomic DNA is present in all cells (except of erythrocytes). Differences in gene transcription and expression of proteins depend on type and function of cells and tissues. In addition, many genes are expressed differentially during development, maturation and ageing of organism. To assure such comprehensive expression of proteins, gene expression in eukaryotes is controlled at six different and independent levels.

Six steps of regulation of gene expression in eukaryotes.

1. Level of DNA and transcription
2. Maturation of mRNA and RNA editing
3. Transport of mRNA from nucleus and localisation in cytoplasm
4. Stability of mRNA
5. Regulation of translation
6. Control of protein activity and post-translational modifications

Regulation at the level of DNA and transcription

Regulation at the level of DNA

At the level of DNA, gene expression is regulated by several different processes. The level of chromatin condensation is important factor affecting transcription. Chromatin is usually present in two forms as **euchromatin** that is relaxed and transcriptionally active or **heterochromatin** that is condensed and transcriptionally repressed. The level of chromatin condensation is regulated by **histone modification**. The most important chemical modification of histones is associated with **acetylation** of lysine residues producing ε-N-acetyl lysine that leads to the relaxation of chromatin and consequent transcriptional activation.

In addition to histone modification, some modifications of DNA represent important mechanism of gene expression regulation. The most important DNA modification is associated with **methylation** of **CpG islands** leading to production of 5-methyl cytosine. Since CpG islands are very frequent in promoter regions of genes, methylation of CpG islands is associated with repression of transcription what is called **imprinting**.

In some conditions genes can be either **amplified** or **deleted**. These mechanisms are pathologic and often associated with development of cancer (e.g. amplification of HER2 gene associated with development of breast cancer).

Regulation of transcription

Initiation of transcription absolutely depends on **basal transcription factors** however the level of transcription is rather low. The level of transcription is mainly modulated by **specific transcription factors** (e.g. CREB – cAMP responding element binding protein) that bind to specific promoter sequences and may respond to specific extracellular and intracellular signals. Specific transcription factors are often regulated by posttranslational modifications (phosphorylation) or ligand binding (steroid hormones). They bind to the specific regulatory DNA sequences called **responding elements**.

In addition to specific transcription factors, the rate of transcription is regulated by specific DNA binding proteins, **enhancers** that increase rate of transcription or **silencers** that decrease rate of transcription.

Combination of all components is associated with comprehensive regulation of transcription in space and time.

Maturation of mRNA and RNA editing

With respect to regulation of gene expression, **tissue specific alternative splicing** represents important mechanism of tissue or cell specific expression of isoforms of particular protein depending on physiological demands of given cell populations or tissues. Alternative splicing is the process by which the exons of the RNA produced by transcription of a gene (a primary gene transcript or pre-mRNA) are joined in multiple ways during RNA splicing. The resulting different mRNAs are usually translated into

different protein isoforms; thus, a single gene may code for multiple proteins (in humans, ~95% of multiexonic genes are alternatively spliced). There are numerous modes of alternative splicing observed, of which the most common is **exon skipping**. In this mode, a particular exon may be included in mRNAs under some conditions or in particular tissues, and omitted from the mRNA in others.

RNA editing represents processes in which the information stored in an mRNA is altered through chemical modifications of mRNA bases or whole mRNA molecule. The mechanisms of mRNA editing involve insertions or deletions of uracil as well as deamination of cytidine into a uridine or adenine to inosine. For example, tissue specific expression of Apo B proteins in humans is regulated by conversion cytidine into a uridine in Apo B pre-mRNA that changes CCA codon coding for glutamine to STOP codon.

Transport of mRNA from nucleus and localisation of mRNA in cytoplasm

Transport of mRNA from nucleus and localisation of mRNA in cytoplasm represent important type of gene expression regulation during development since uneven distribution of synthesised proteins is base for differentiation of cells. To date three main mechanisms of targeted mRNA localisation are supposed:

- directed transport of mRNA on cytoskeleton
- random diffusion of mRNA and trapping
- degradation of mRNA combined with local protection.

Stability of mRNA

Binding of specific proteins on 3´-UTR or polyadenylated tail of mRNA significantly increases mRNA stability. Increased mRNA stability is normally associated with higher expression of particular protein whereas decreased mRNA stability is associated with lower expression of particular protein.

Regulation of translation

Eukaryotic translation is mainly regulated by four different mechanisms:

– Phosphorylation of eIF2α

– Binding of initiation factors eIF-4E and eIF-4G on CAP and internal ribosomal entry sites (IRES)

– Binding of regulatory proteins on 5´-UTR or 3´-UTR

– Regulation by miRNA

Phosphorylation of eIF2α

Phosphorylation of eIF2α by different *protein kinases* is associated with inability of phosphorylated eIF2α to bind GTP and thus with inhibition of translation initiation.

Binding of initiation factors eIF-4E and eIF-4G on CAP or internal ribosomal entry sites (IRES)

Binding of initiation factors eIF-4E and eIF-4G on CAP is essential step for initiation of translation. Both initiation factors can be modified by posttranslational modifications that is associated with changes of their level or binding properties and thus with changes in the rate of translation initiation. In addition to CAP, both initiation factors are involved in interaction with IRES that is sequence in 5´-UTR of mRNA that allows translation of the mRNAs in a CAP-independent manner.

Binding of regulatory proteins on 5´-UTR or 3´-UTR

Although many features of an mRNA can contribute to its translation, most control elements are located within the untranslated regions (UTRs). Both 5′-UTRs and 3′-UTRs can contain sequences that function as binding sites for regulatory proteins. Binding of such proteins to their binding sequence in 5′-UTRs is often associated with inhibition of translation whereas binding of regulatory proteins to 3′-UTRs increases efficiency of translation mainly due to increased stability of mRNA. Binding of the regulatory proteins to their binding sequences in UTRs is often strongly modulated by binding of specific ligand to particular regulatory proteins. Thus the rate of synthesis of particular protein might be quickly modulated by changes in metabolism or physiologic conditions.

Regulation by miRNA

Regulation of translation by miRNA involves binding of particular miRNA on complementary sequences of its target mRNA. Such mRNA is either quickly eliminated by specific *ribonucleases* or is not translated to the particular polypeptide due to bound miRNA that blocks translation elongation. One miRNA can block translation of several different polypeptides whereas translation of one polypeptide can be regulated by combination of several different miRNAs (so called **miRNA clusters**).

Control of protein activity and post-translational modifications

There are several post-translational modifications (also called **covalent modification** since they are associated with production or extinction of covalent bonds) with significant impact on function of modified protein (see also regulation of enzyme activity at the level of covalent modifications). Post-translational modifications involve either **proteolytic activation** or **degradation** of proteins as well as **chemical modifications** of polypeptide chain (e.g. phosphorylation, hydroxylation, methylation, acetylation, adenylation, ADP-ribosylation…)

Proteolytic activation of proteins

Many proteins are synthesised as inactive polypeptides that are converted to active protein by proteolytic cut of part of polypeptide. Typical examples are pepsinogen-pepsin, fibrinogen-fibrin or prothrombin-thrombin.

Blood clotting cascade

Proteolytic degradation of proteins

„Aged“ and damaged proteins, proteins of cell cycle, transcription factors, apoptotic and signal transduction proteins are removed from the cells by proteolytic degradation via ATP–dependent proteasome 26S complex. Proteins destined to be eliminated by proteasome 26S complex are in first step labelled with small peptide called **ubiquitin**. The attachment of multiple ubiquitines to target protein is associated with ATP consumption and is catalysed by group specific enzymes called *ubiquitin ligases*.

Inadequate ability of proteasomal 26S complex to cope with aberrant proteins is considered to be a cause of several neurodegenerative diseases including Alzheimer and Parkinson disease. On the other hand, specific inhibitor of proteasomal 26S complex (bortezomib) is used for treatment of some hematologic malignancies (e.g. multiple myeloma).

Chemical modifications of polypeptide chain

Phosphorylation/dephosphorylation of proteins

Phosphorylation of tyrosine, serine or threonine residues by ***protein kinases*** using ATP as phosphate group donor can significantly change protein function. ***Protein kinases*** refer to a family of enzymes whose activity is dependent on cellular levels of different molecules involved in signal transduction. For example *protein kinase A* (**PKA**) is activated by cAMP, *protein kinase C* (**PKC**) is activated by diacyl glycerol or tyrosine kinase receptors (e.g. insulin receptor, EGFR) phosphorylate tyrosine residues of substrate proteins after binding of particular ligand (e.g. insulin, EGF, respectively).

The opposite reaction (dephosphorylation), removal of phosphate group from proteins, is catalysed by group of enzymes generally called ***protein phosphatases*** that often respond to signal molecules.

AS it was already mentioned (see page 12), there are not general rules describing effect of phosphorylation/dephosphorylation on protein activity.

Biologic function of some proteins is regulated at the level of posttranslation modifications in very complex manner. For example, **hypoxia induced factor** (HIF) is transcription factor regulating expression of proteins involved in **energetic metabolism** and **tissue vascularisation** (formation of new vessels and capillaries delivering blood to the tissues). During **normoxia** (normal tissue concentration of oxygen), HIF is hydroxylated by specific *hydroxylase* and then proteolytically degraded. However, during **hypoxia** (low tissue concentration of oxygen), hydroxylation of HIF does not occur HIF is stable and triggers transcription of target genes.

Examples of pathological post-translational modifications of proteins

- Nonenzymatic glycation of proteins (see text book III)

- Modification of proteins by homocysteine
- Oxidative modifications of proteins during ageing
- Prion diseases

Modification of proteins by homocysteine

Homocysteine is intermediate of methionine metabolism. Increased concentration of homocysteine in blood (**homocysteinemia**), from genetic or metabolic reasons, is associated with higher risk of atherosclerosis and thrombosis that are further responsible for cardiovascular diseases as myocardial infarction and stroke. Pathologic effects of homocysteine are attributed to homocysteine lactone that is formed in blood from homocysteine and that is able to react with free amino groups of plasma proteins (altering functions of blood clotting factors associated with thrombosis) or endothelial membrane proteins (causing dysfunction of endothelial cell and consequent formation of atherosclerotic plaques).

Oxidative modifications of proteins during ageing

Ageing is associated with slow accumulation of oxidised proteins mainly in extracellular matrix, since intracellular proteins that are modified by oxidation are normally efficiently eliminated by proteasomal 26S complex or degraded in lysosomes. The most important oxidative modifications are **formation of dityrosines**, **oxidation of tryptophane** or reaction of free amino groups of lysine residues with aldehydes that are **products of lipid peroxidation** (see Medical Chemistry and Biochemistry III). Such modifications might be responsible for muscle rigidity (by changing collagen flexibility) and decreased reactions (due to changes in neuronal extracellular matrix) that are often observed in senescence.

Prion diseases

Prion diseases are neurodegenerative disorders that are associated with unusual posttranslational modification of priones. **Priones** are integral plasma membrane proteins with unknown physiologic function that are expressed in CNS and muscles. Normal proteins (**PrP^{C}** – C comes from cellular) with dominant α-helix secondary structure are not

inducing any pathologic process. However, aberrant priones (**PrP^{SC}** – SC comes from scrapie that is one of the forms of prion diseases) with dominant β-sheet secondary structure are inducing spongiform degeneration of CNS. PrP^{SC} are working as chaperone proteins that are solely able to change tertiary structure of PrP^{C} into PrP^{SC}. In addition, accumulation of PrP^{SC} is associated with formation of protein aggregates in form of fibrils that are not degraded by proteasomal 26S complex and that are responsible for induction of neuronal death and consequent neurodegeneration. Thus intoxication with PrP^{SC} is able to initiate pathology process (infection like origin of prion diseases). Interestingly, muscle PrP^{C} are resistant to PrP^{SC}.

GENE MANIPULATION

Recombinant DNA technology

By definition, **recombinant DNA** is any DNA that is composed from at least two parts of DNA coming from different organisms and that can be replicated in host organisms. The population of host cells containing identical recombinant DNA is called **clone**. **Recombinant DNA techniques** represent all methods used for production and analysis of recombinant DNA. Finally, the term **biotechnology** represents all technologies based on recombinant DNA to produce different bioproducts (either proteins or small molecules).

Methods used in recombinant DNA technology

Construction of functional recombinant DNA is fundamental method of recombinant DNA technology. Formation of recombinant DNA requires a **cloning vector**, a DNA molecule that is generally derived from plasmids or viruses that represents relatively small segments of DNA containing all signals necessary for replication (replication origin). Molecule of DNA vector has to be attached to the fragment of foreign DNA by means of *DNA ligase*.

Restriction endonucleases

Restriction endonucleases are DNA cutting enzymes that recognise short palindrome sequences (4 - 6 nucleotides) of DNA (Table 7). Unlike *endonucleases* and *exonucleases*,

they cut DNA within recognised palindromic sequence. With respect to recognised sequence, they are highly specific, change of one nucleotide within recognised sequence is associated with lost of *endonuclease* activity.

Table 7 Examples of *restriction endonucleases*

Restriction endonuclease	**Source**	**Cutting site**
*Eco*RI	*Escherichia coli RY13*	↓ 5′-GAATTC-3′ 3′-CTTAAG-5′ ↑
*Eco*RV	*Escherichia coli*	↓ 5′-CCCGGG-3′ 3′-GGGCCC-5′ ↑
*Sma*I	*Serratia marcescens* S_b	↓ 5′-GAATTC-3′ 3′-CTTAAG-5′ ↑
*Hae*III	*Haemophilus aegyptius*	↓ 5′-GGCC-3′ 3′-CCGG-5′ ↑

Restriction endonucleases are important molecular tools used for cutting of DNA to discrete fragments. In addition, specificity of *restriction endonucleases* is used for fast and accurate mapping of single nucleotide mutations (see further).

DNA library

Since human genome contains approximately 26 000 open reading frames each having own regulatory sequences, the identification or isolation of gene of interest would be the

needle in the haystack dilemma. In order to prepare catalogue of DNA sequences, a **DNA library** approach is used. This helps to identify genes or find investigated sequence. DNA library is a collection of DNA fragments from one organism that have been cloned into vectors so that DNA fragments of interest can be identified and isolated for further study or analysis. There are basically three kinds of libraries: genomic, chromosome and cDNA libraries. **Genomic DNA library** is a comprehensive collection of cloned DNA fragments from a whole genome of a cell, tissue, or organism. **Chromosome DNA library** is a collection of cloned DNA fragments from one chromosome and **cDNA library** is a collection of cloned DNA sequences that are transcribed into mRNA in particular cell or tissue.

In order to prepare DNA library, the investigated DNA has to be first cut to fragments by proper *restriction endonuclease*. These fragments are then cloned to the vector and the mixture of recombinant DNA clones is transferred to the host organism, for example enterobacteria *E. coli*. The transformed bacteria are then seeded on solid growth media. They grow in colonies each colony contains bacteria producing only one identical clone of recombinant DNA. Sequence or gene of interest has to be further indentified by some method of identification of DNA sequences (either by means of hybridization probes or polymerase chain reaction, see further).

Methods of identification of DNA sequences

Analysis of investigated DNA is often associated with identification of DNA sequences of interest. The DNA sequences can be identified by:

- Hybridization probes
- Southern blotting
- DNA sequencing.

Hybridization probe is labelled single stranded DNA or RNA that can produce base pairs (hybridize) with complementary sequences of investigated DNA or RNA. Hybridization probes were widely used for identification of DNA library clones containing the sequence

of interest. In order to identify colony containing DNA library clone with sequence of interest, culture dishes containing the bacterial colonies are blotted with a piece of specific membrane, to which some bacteria of each colony adhere. The adhering colonies are treated with weak basic solution to disrupt the cells and to separate the strands of the cloned DNA. The membrane is then incubated with either radioactive or chemically labelled DNA probe containing part of the complementary sequence of the gene of interest. After hybridization and washing out of unbound probes, the positive colonies are visualised depending on the nature of label (either by exposition to X-ray film or chemically). Position of the signal on the membrane allows the identification of colony of bacteria that contain cloned DNA with sequence or gene of interest. The positive colony is used for inoculation of growth media where bacteria with positive clone will grow. This will allow production of sufficient amount of recombinant DNA for further analysis. The method is relative fast and if necessary thousands of bacterial clones can be screened in this way to find the one clone that hybridizes with the probe.

Southern blotting is method of detection of DNA sequences that is based on separation of DNA fragments by electrophoresis, transfer of separated DNA fragments to specific membrane and final hybridization with hybridization probe.

Steps of Southern blotting include:

1. Cutting of genomic DNA by particular *restriction endonuclease* to discrete fragments

2. Separation of DNA fragments by gel electrophoresis

3. Transfer of DNA on nitrocellulose membrane – blotting

4. Hybridization with labelled hybridization probe and detection of sequences of interest.

Northern blotting that is similar to Southern blotting is used for detection of RNA while **Western blotting** is method of detection of proteins. Western blotting includes separation of proteins by electrophoresis, transfer of separated proteins on specific membrane and final visualisation of protein of interest by labelled antibodies raised against detected protein.

DNA sequencing enables to perform a thorough analysis of DNA because it provides the most basic information of all: the sequence of nucleotides. DNA sequencing can be used for the localisation of regulatory and gene sequences, comparisons between homologous genes across species or identification of mutations. In 1974, two methods were independently developed by an American team, led by Maxam and Gilbert, and an English team, led by Sanger. The Maxam and Gilbert method is based on chemical cleavage of DNA, while the **Sanger's method** is using a procedure similar to the natural process of DNA replication. At the present, the Sanger's method became the standard procedure because of its practicality and ability to be fully automated. The sequences of several thousands genes of different species, including human being, are recorded in several bioinformatics databases (e.g. database of National Center for Biotechnology Information, http://www.ncbi.nlm.nih.gov/nuccore).

Sanger's method of DNA sequencing is based on synthesis of DNA strand in the presence of labelled **dideoxynucleotides** (ddNTP) in addition to the normal deoxynucleotides (dNTP). Dideoxynucleotides are essentially the same as deoxynucleotides except they contain a hydrogen on the 3' carbon of 2´-deoxyribose instead of a hydroxyl group. During synthesis of DNA strand, ddNTPs integrated into a sequence of synthesised strand prevent the addition of further dNTPs. This occurs because of missing 3´-OH group. Therefore phosphodiester bond cannot form between the ddNTP and the next incoming dNTP, and thus the synthesis of DNA chain is terminated.

During sequencing of DNA by Sanger's method, the sequenced DNA has to be first thermally denatured into single strands and then a primer is annealed to the template strands. This primer is specifically designed so that its 3' end is located next to the DNA sequence of interest. After annealing of primer, reaction mixture is supplemented by *DNA polymerase*, all four dNTPs and all four ddNTPs, each at about one-hundredth the concentration of the normal dNTP. Each ddNTP is labelled by different fluorescent dye. As the DNA strand is synthesised, nucleotides are added on to the growing chain by the DNA polymerase depending on sequence of template strand. However, incorporation of dideoxynucleotide into the chain in place of a normal deoxynucleotide results in a termination of chain synthesis. In this way, DNA fragments of all different lengths and terminated with a specific labelled base are produced. At the end of reaction, DNA is denaturated and the synthesised fragments are separated by capillary electrophoresis. Since

the fragments are separated according to their size and specific detector is able to recognize fluorescent dye and thus the base of terminal dideoxynucleotide, the flow of labelled fragments through detector is translated to the sequence of synthesised strand that is complementary and has opposite orientation as sequenced region.

Methods of DNA amplification

In the case that the amount of analysed DNA is not sufficient, DNA has to bee amplified before analysis. Two different methods, **cloning** and **polymerase chain reaction** (PCR) are used for amplification of DNA.

Cloning is based on transfer of recombinant DNA into host cells that are then cultivated in growth media. Growth of host cells is associated with DNA replication, including recombinant DNA that is further isolated from host cells at the end of cultivation. Depending on the type of host cells, three different types of transfer are recognised.

Transformation means transfer of recombinant DNA into bacteria, whereas **transduction** refers to transfer of recombinant DNA into bacteria using bacteriophages (bacterial viruses) and **transfection** is transfer of recombinant DNA into eukaryotic cells. In addition, **genetic manipulations** that transfer exogenous gene into animals or plants, producing genetically modified organisms, are associated with DNA amplifications however they are not used for DNA amplification.

Polymerase chain reaction (PCR) is based on repetitive DNA synthesis using thermostable *DNA polymerase* in the presence of specific primers that specify target amplified sequence. One cycle of polymerase chain reaction involves three steps:

1. **Thermal denaturation** – the reaction mixture, containing thermostable *DNA polymerase*, primers, all four types of dNTP and DNA that has to be amplified, is heated in specific thermal cycler to 95 °C that causes denaturation of double stranded DNA to two strands of single stranded DNA.
2. **Hybridization of primers** – the reaction mixture is cooled to °60 C that causes hybridization of primers with complementary sequences.

3. **DNA synthesis** – the reaction mixture is heated to 72 °C that is associated with elongation of DNA strands from primers by thermostable DNA polymerase.

In theory, each cycle of PCR should produce double amount of DNA. Due to limited stability of thermostable *DNA polymerase* and limited amount of primers, maximal number of cycles in one run is 40 that should amplify DNA by factor 2^{40} (approx. 1.1×10^{12}). This is just theoretical yield of PCR since during so called lag phase of reaction, the factor of amplification is significantly less than 2. The second limiting step is related to exhaustion of primers and DNA polymerase. Despite this, PCR can amplify DNA amount by factor 10^6 (from picograms to micrograms of DNA).

Reverse transcription PCR (RT-PCR) refers to amplification of cDNA or mRNA detection and is preceded by synthesis of cDNA from RNA isolated from tissue of interest using *reverse transcriptase*.

In human medicine, PCR can be used for:

- detection of mutations in prenatal diagnostics and oncology,
- molecular classification of tumours (RT-PCR detection of mRNA of oncogenes)
- identification of bacteria and viruses in human body
- microbiological screening of food
- DNA amplification in forensic medicine

Recombinant DNA technology in diagnostics of human diseases

DNA array

DNA array method is high throughput method that can be used to determine expression levels of large numbers of genes or mutations of several thousands of genes simultaneously.

A **DNA array** (also known as DNA chip) is a collection of hybridization probes attached to a solid surface containing approximately picomoles (10^{-12} moles) of a specific DNA sequence that are used to hybridize a cDNA or cRNA sample (called target) under conditions ensuring specific hybridization to target sequences (so called **high-stringency conditions**). Probe-target hybridization is usually detected and quantified by detection of

fluorophore-labelled targets to determine abundance of investigated nucleic acid sequences in the sample relative to the level of the same sequence in control.

Restriction fragment length polymorphism is a method of detection of point mutations that are associated with production or extinction of site recognised by particular *restriction endonuclease*. The method is based on amplification of region with suspected mutation by PCR using specific primers. After amplification, the product of PCR reaction is treated with particular *restriction endonuclease*. The consequence of treatment is appearance of different length restriction fragments depending on the presence of mutation. This method is fast and reproducible but the use is significantly limited by fact that not all important mutations are localised within the restriction sites of to date known *restriction endonucleases*. For example, diagnostics of **sickle cell anaemia** can be easy performed by RFLP using *restriction endonuclease* Mst II. Mst II recognizes the sequence CCTNAGG (where N is any nucleotide). Single point mutation in the β globin gene associated with sickle cell anaemia changes CCTGAGG to CCTGTGG. Thus, the A to T mutation that causes sickle cell anaemia also causes the loss of the recognition site for the restriction enzyme Mst II. PCR amplification of DNA sequence that include affected region of the β globin gene gives PCR product that is further treated with *restriction endonuclease* Mst II. In the case that both alleles of β-globin gene are normal, the PCR product will be cut by MstII giving two smaller DNA fragments. If both alleles of β-globin gene are affected, the product is not cut. Since heterozygote contains one normal and one mutated allele of β-globin gene all three fragments will be present in reaction mixture after treatment with Mst II. The resulted fragments are easy analysed by gel electrophoresis.

Southern blotting can be used for detection of DNA insertions or deletions since their presence within investigated region is associated with a change of the profile after Southern blotting. In the case of short insertions or deletions, profile is not changed significantly and the detection of insertion/deletion requires the use of both wild type and mutant probe. Diagnostics of **cystic fibrosis** that is caused by deletion of UUU codon in CFCTR gene is based on two separate Southern blots each analysed with specific probe, one blot is analysed with probe specific for wild type allele while the second is analysed with probe specific for mutated allele.

PCR can be used for detection of DNA **insertions** or **deletions** as well as **gene amplification**. The consequence of insertion or deletion is reflected by change of PCR

product size what can be easy detected by gel electrophoresis. For example, diagnostics of Huntington disease by PCR is based on fact that mutated gene contain unusual number of CAG codons (coding for glutamine). While normal number of CAG codons is 18, the disease is manifested after expansion of number of CAG codons to 48. The presence of multiplicated codons will result in larger PCR product using primers encompassing region with mutation.

Due to higher input, assuming that the same amount of DNA contains more copies of investigated gene, **gene amplification** is associated with higher amount of PCR product after certain number of PCR cycles using primers targeting investigated gene.

In situ hybridization is used for detection of DNA or RNA sequences directly in investigated tissue. The issue of interest has to be first fixed to avoid chemical changes. Then it is cut to slices, mounted on microscope glass and hybridised with specific labelled hybridization probe targeted on sequence of interest. After hybridization, the tissue is examined under microscope. If the sequence of interest is present in the examined tissue, the labelled probe will be visible under microscope. The method can be used in prenatal diagnostics and oncology to detect specific mutations. In the case of RNA detection, in situ hybridization can be used in oncology for morphologic and molecular classification of tumours.

DNA fingerprinting is method used mainly in forensic medicine for identification of crime culprits, victims of murders or catastrophes as well as paternity test. The method is based on the presence of **variable number of tandem repeats** (VNTR) in hypervariable regions of genomic DNA at several chromosomal locations. Since genomic DNA of each person is given by combination of two unique alleles from each parent the result is again unique structure of hypervariable regions that can be analysed by DNA fingerprinting. DNA isolated from white blood cells, semen, hair bulbs or other cells is amplified by PCR and further analysed by the Southern blot using carefully chosen hybridization probes. Such procedure gives unique profile after hybridization for each person. Only profiles of one egg twins are equal. Approximately half of the profile of child's DNA will match with half of bands of mother's DNA profile and the second half will match with half of bands of father's DNA profile. This is principle of paternity test by DNA finger printing. In the case of personal identification, the DNA profile of identified person is compared with the DNA profiles of suspected persons or with DNA profiles of close relatives of identified person.

Recombinant DNA technology in treatment of human diseases

Production of therapeutical proteins by recombinant DNA technology

Cloning of DNA and consequent expression of cloned gene in host cells is currently used by biomedical companies to produce hormones (e.g. insulin), therapeutical proteins (e.g. antibodies, streptokinase) and vaccines.

The gene coding for protein of interest is inserted into **expression vector** that in addition to replication origin has to contain promoter sequences to initiate transcription and sequences important for translation of produced mRNA into particular polypeptide. Such vector allows over-expression of functional protein in host organism. Thus, cultivation and growth of host cells is associated with production of cloned protein that is finally purified from host cell lysates.

Antisense oligonucleotides and RNA interference

Both methods of disease treatment are based on inhibition of synthesis of target protein by reducing the level of particular mRNA. **Antisense oligonucleotides** are short sequences of single-stranded synthetic deoxyribonucleotides that bind to specific complementary coding regions on mRNA. Formed RNA-DNA heteroduplex is a substrate for *RNase-H*, which cleave the target mRNA and prevent its processing and subsequent protein synthesis. The specificity of method is determined by the fact that each DNA sequence longer than 17 nucleotides occurs only once within human genome. **RNA interference** (RNAi) based on knowledge of translation regulation at the level of miRNA (see above) requires delivery of particular miRNA or its precursor (e.g. **short hairpin RNA** (shRNA)) to target cells. Several protocols of treatment of different diseases using either antisense oligonucleotides of RNAi are currently in the different phases of clinical testing.

Gene therapy

Only a few, from over 4000 to date known, inherited human diseases can be currently treated with successful efficiency. For majority of these diseases, the missing or defective gene product cannot be delivered exogenously as for example insulin in the case of

diabetes treatment or factor VIII used for treatment of haemophilia A. Most of proteins are unstable and cannot be delivered in functional form to their site of action in the body. In addition, the cell membranes are impermeable for large molecules as proteins, therefore missing or defective proteins have to be synthesised in the affected cells. To date, the treatment of inherited diseases is significantly restricted to the diseases where missing metabolite is a small molecule that can be delivered to the affected cells by blood circulation. For many inherited diseases, the gene therapy represents the most promising approach to efficient treatment. Gene therapy involves delivery of a normal (wild type) copy of defective gene to the genome of affected person. The normal gene has to be inserted in specific vector that will allow incorporation of replacing gene into genomic DNA of recipient cells and further expression of functional protein.

A number of **viruses** have been used for human gene therapy, including retrovirus, adenovirus, lentivirus, and herpes simplex virus. All viruses bind to their hosts and introduce their genetic material into the host cell as part of their replication cycle. Therefore this can be used as a feasible strategy for gene therapy, by modification of the viral DNA and using the virus as a vehicle to deliver the therapeutic DNA.

Non-viral methods can present certain advantages over viral methods, such as large scale production and low host immunogenicity. On the other hand, low levels of transfection and expression of the gene was the major disadvantage non-viral methods. Recent advances in vector technology have discovered molecules and techniques that reach the transfection efficiencies of viruses. There are several methods for non-viral gene delivery, including the injection of naked DNA, electroporation, oligonucleotides, and inorganic nanoparticles.

Although gene therapy is promising approach for treatment of inherited diseases, it is associated with serious problems. Major problems of gene therapy include:

- Unstable integration of therapeutic DNA into the genome and the rapidly dividing nature of many cells prevent gene therapy from achieving long-term benefits. Patients will have to undergo multiple rounds of gene therapy.
- Introduction of foreign object into human tissues is always associated with the risk of stimulating the immune system in a way that reduces gene therapy effectiveness or may be potentially harmful for treated person.

- Viruses used as the carriers in most gene therapy studies, present a variety of potential problems to the patient: toxicity, immune and inflammatory responses, and gene control and targeting issues.
- Recently, gene therapy is significantly restricted to disorders that arise from mutations in a single gene. However, the most commonly occurring disorders are caused by the combined defects in many genes. Multigene disorders are difficult to treat effectively using gene therapy.
- Chance of inducing a mutation due to nonspecific incorporation of therapeutic DNA to the genome of treated person. One possible solution for this is to add a functional gene by replacing defective gene on the base of homologous recombination.
- Modification of genome of germ-line cells is associated with ethic problems and is restricted to be used only for experimental animals.

REFERENCES

Davidson, V. L. and Sittman D. B.: Biochemistry. Third Edition. Harwal Publishing. 1994.

Devlin, T. M., et al.: Textbook of Biochemistry with Clinical Correlations. Sixth Edition. Wiley-Liss. Hoboken NJ, 2006.

Kaplán, P.: Medical Chemistry and Biochemistry I. Comenius University Bratislava, 1997.

Kierszenbaum, A. L.: Histology and Cell Biology: an Introduction to Pathology. Mosby Inc., 2002.

Koolman, J. and Roehm, K. H.: Color Atlas of Biochemistry. Second Edition. Georg Thieme Verlag Stuttgart, 2005.

Lehotský, J., et al.: Medical Chemistry and Biochemistry II. Comenius University Bratislava, 2012.

Murray, R. K., et al.: Harper's Illustrated Biochemistry. 26th Edition. McGrawe-Hill, 2003.

Račay, P. Medical Chemistry and Biochemistry III. Comenius University Bratislava, 2012.

Snustad, P., and Simmons M. J.: Principles of Genetics. Third Edition. John Wiley & Sons, Inc., 2003.


9 788089 544462