

**68/ The structure of DNA, the structure of chromosome, euchromatin, heterochromatin, regulation of transcription, enhancer, silencer, the difference between prokaryotic and eukaryotic gene expression.**

The structure of DNA: "primary str." the order of nucleotides, having the potential to code for information. The nucleotides are purin or pirimidine bases connected to ribose, and the ribose molecules are connected by phosphate groups covalently (sugar phosphate backbone). Between the single strands, A and T form two hydrogen bonds, the C and G form three ones. The two single strand form double strand in a helix "secondary structure". The double helix has major grooves (the transcription factors bind to the major groove, because from there the G C from C G and A:T from T:A can be distinguished without separating the two DNA strand therefore the specific sequences can be recognised) and minor grooves.

The structure of chromosome: the elementary units are the nucleosomes, in which the DNA double helix is coiled around the histone octamers (2x H2A,H2B,H3,H4). The octamers are connected by the H1 protein. This condenses the genomic DNA about 40-fold. The DNA is 100-1000-fold condensed in the chromatin and about 10,000-fold in the metaphase chromosome (the 180 cm haploid human genome becomes 180  $\mu$ m). The euchromatin is less condensed, actively transcribed region, while heterochromatin is more condensed inactive, non-transcribed region.

The regulation of transcription: The RNA polymerase binding to DNA is facilitated by a number of proteins called transcription factors. These protein show helix-loop-helix, Zn-finger motives to bind to the DNA. The Leu-zipper motives help to attach two proteins. Transcriptional factors can be hormone receptors, tumor suppressor proteins.

Enhancer is a DNA sequence that binds a transcriptional activator, silencer is a DNA sequence that binds repressor of the transcriptional machinery. They are mostly located before the promoter, upstream of the start of the gene. The activators connected with co-activators and the co-activators with basal factors, the later ones are bound to the RNA-polymerase.

The prokaryotic gene expression (transcription ) is not separated from the translation by the nuclear membrane, therefore the first part of the RNA can be translated by the ribosomes while the rest of it is still transcribed or have not even been transcribed yet. In prokaryotes the splicing is less common than in eukaryotes where one mRNA codes for one type of protein.

## **69/Histones and their function, nonhistone proteins, regulation of replication of the eukaryotic cell: exit from G0 phase, protooncogenes, cyclins, tumor suppressor genes.**

The DNA is condensed into chromatin with the contribution of histones. 2-2 molecules of H2A, H2B, H3 and H4 are connected into histone octamer. The double stranded DNA is coiled around the histone octamer forming a nucleosome (the elementary unit of the chromatin). The nucleosomes are connected by H1 proteins. H1, when highly phosphorylated, connects nucleosomes condensing the chromatin. H1 phosphorylation is a subject of regulation at the G2/M border of the cell cycle, when the cyclin B and the cyclin dependent kinase 2 (CDK1) complex induces phosphorylation of H1, therefore leads to the condensation of chromatin and then to that of mitotic chromosomes.

Nonhistone proteins binds to specific sequences between the nucleosomes, many of them are involved in regulation of gene expression.

The cell cycle consist of interphase (G1,S,G2) and mitotic phase (M). The DNA is replicated in the S phase. Transition from the G1 to S phase is regulated by so called tumor suppressor genes e.g. the gene of the retinoblastoma protein (Rb) or the p53 protein. Rb inhibits the start of replication when it is highly phosphorylated, while if it is dephosphorylated or mutated it can not prevent DNA-replication and subsequent mitosis. The Rb-defect is inherited by the daughter cells that are likely to become tumorous because their proliferation is uncontrolled. Heterozygote carriers of Rb mutations are more likely to develop tumor, because the normal copy of the RB gene is susceptible to mutagenic effects. p53 is a transcriptional factor inducing genes involved in the DNA repair. p53 also prevents the cell from starting the replication (G1/S border) before the repair is completed (inducing p21 that inhibits the formation of the cyclinD-CDK2 complex). If the p53 is mutated the daughter cells will carry a vast amount of mutated DNA finally leading to apoptosis of the cell.

Cells in the G0 phase are latent or quiescent, but they can exit from this phase to the G1 by the effect of growth factors. The genes of several growth factors are protooncogenes that can be mutated into oncogenes. An oncogene stimulates tumorigenesis by its increased gene activity not like by its inactivation as the tumor suppressor gene.

Differentiation and cell proliferation are often mutually exclusive events. Cells in the terminal differentiation are in a sort of G1 phase. Tumorous cells never differentiate because they can not stop in the G1 phase.

## **70/ Repair mechanisms, tumorous cell proliferation, apoptosis, tumor sensitivity of knock out mutants of the p53 gene.**

The repair system correcting the thymine-dimers (formed by the effect of irradiation by UV light) starts with the action of photolyase cutting the bindings between the dimers, then a specific endonuclease hydrolyzing out the affected sugar-phosphate backbones then the DNA polymerase and ligase correct the DNA.

Chemical mutagens often result in deamination of citidine (C) converting it into uracil. U is not a proper nucleotide for the DNA, therefore the glycosidase hydrolyzes away the base leaving the sugar-phosphate backbone behind. The deoxyribose and phosphate will be hydrolyzed away from the DNA by endonuclease. The DNA polymerase will attach C to the 3'OH end. The ligase will connect the 3'OH and P-5' ends.

The repair mechanisms in the cell are largely regulated by the p53 protein. p53 keeps the cell in the G1 phase until the repair is completed. If excess damage has occurred to the DNA, the amount of p53 could increase and thus inhibits cell division and stimulates genes needed for repair, allowing time for the corrections. If the p53 is inactivated by mutation, then it would not fulfill this function, making the cell more vulnerable, genetically less stable, and leaving chance for the continuous cell divisions (tumors). After accumulating mutations, in extreme cases the cell will lose its ability to function and undergo apoptosis. p53 can be called the "guardian of the genome". The knock out mutant mice of the p53 gene survive but are more sensitive to tumors than the healthy mice, because they contain only one or no copy of the normal p53 gene, therefore the suppression of tumor formation is weak or does not exist in these animals.

## **71/ The structure of eukaryotic genes, exon, intron, splicing, coding and noncoding fragments, classification of genomic sequences by repetitive character, IRE and UTR**

The majority of the genome in eukaryotes does not code for genes. Less than 5% of the human genome codes for about 30,000 genes. More than 90% of the genome is noncoding region (e.g. pseudogenes that are never expressed, regulatory sequences, and spacers that do not have known functions, but can be suspected to contribute to the chromatin structure). Most eukaryotic gene is monocistronic. Starting from the 5' end they have a regulatory region, transcription initiation site, a promoter and a structural part of the gene. The transcription is terminated at the termination site. The structural gene often has introns and exons. The introns are spliced out from the pre-mRNA (heteronuclear or hnRNA) after the splicing the mRNA has exons, only. The exons contain the protein coding region (within the open reading frame).

According to repeats the genomic sequences can be classified into three groups. 1/ Sequences that are present in a few copies e.g. mRNA genes, histone genes. 2/ Sequences that are repeated a few hundred times e.g. transcription factor binding regulatory regions, protein binding sequences. 3/ Sequences repeated a few thousand times, these are probably involved in the organization of chromosomes (mostly at the telomere and centromere).

The iron response element (IRE) is a sequence located at the 5' end of the ferritin mRNA and also at the 3' end of the transferrin receptor mRNA. The IRE binds aconitase (also known from the citrate/isocitrate conversion in the mitochondria) in case of iron starvation, this confers stability to the transferrin receptor mRNA so that to be translated but it prevents translation of the ferritin mRNA (by inhibiting ribosomal binding). As a result the iron perception and the take up will be increased but not the transport. If there is excess of iron, the aconitase will be dissociated and the ferritin mRNA can be translated while the transferrin receptor mRNA will be degraded.

Untranslated regions (UTRs) are found at the 3' of mRNA and confer stability and information for the transport of mRNAs, e. g. to the apical or basal part of an epithelial cell.

## **72/ RNA types, RNA polymerases, the transcription process, maturation of mRNA, the mechanism of splicing, tissue specific and developmental dependent splicing, thalassemias, antisense RNA.**

The ribosomal (rRNA) is 70-80%, the transfer (tRNA) is 12%, the mRNA is 3-5% of the total RNA in a cell, the remains are hnRNA, scRNA, snRNA. The rRNAs are transcribed by the polymerase I, the mRNAs by RNA polymerase II, the 5s rRNA and tRNAs by RNA polymerase III.

The transcription starts from the promoter region (strong and weak promoters), continues with elongation and ends by termination at a special terminating sequence which forms a hairpin like structure. The RNA (hnRNA in eukaryotes) undergoes processing; Cap addition at the 5' end, polyA tail formation at the 3' end and splicing in the intermediate sequence. The splicing is the removal of the introns and the connection of the bordering exons. Exons contain the protein coding region. The altered form of splicing can be tissue specific or developmental stage dependent, meaning that different cells are able to translate different proteins from the transcripts of the same gene by the alternative splicing. The mechanism of splicing: the 2'OH of an A nucleotide in the middle of the intron attacks the 5' of the GU sequence after exon1; UGA complex and a hairpin like structure are formed. The 3'OH attacks 5' end of the exon2 after the AG sequence and exon1 and exon 2 are connected, meanwhile the intron is removed. In

the  $\beta$ -thalassemias the deficiency or mutation of the  $\beta$ -globin gene cause alterations in the splice sites (e.g. the UG at the 3' end of exon1 or AG at the 5' end of exon2) that results in the abnormal or nonspliced RNA, reduced chains of hemoglobin and anaemia. The mRNA can be bound by its antisense complement and the RNA:RNA hybrid can not be translated. This is an interesting form of the pretranslational regulation. The antisense mRNA can be synthesised from a gene different from that of the sense mRNA.

### **73/ Transcription factors, HLH proteins, Zn-finger proteins, Leu-ziper, intracellular hormone receptors, fos and jun proteins, p53, hox proteins, the connection of transcription and the chromatin structure.**

The binding of the RNA polymerase to the promoter is helped by many proteins, among these the basal factors, the TATA binding protein, the coactivators, activators and repressors (the activators bind to enhancer sequence, the repressor to silencer sequence). A number of the transcription regulating proteins have helix-loop-helix motif, this structure fits to the major groove of DNA (e.g. the MyoD, Myf-5, Myogenin, MRF4 that are important in muscle differentiation or the proteins of the achete-scute gene complex of *Drosophila* that are important for neurogenic differentiation. The other typical motif, the Zn-finger, also fits to the major groove of DNA. The Zn-finger is found e. g. in the TFIIIA transcription factor regulating the transcription of 5S RNA and the steroid receptors. The Leucine zipper facilitates connections between proteins in a zipper-like manner e.g. the *fos*, *Jun*/Ap1 and *myc* proteins have this motif. These proteins are transcription factors and protooncogenes (can be involved in tumor formation). The *p53* protein is an important transcriptional factor in controlling the cell cycle and it is also a tumor suppressor gene. The family of *hox* proteins are transcriptional regulators involved in decisions for the cell fate (e.g. mesodermal, nonmesodermal).

Only a few transcriptional factor is suspected to be involved in chromatin regulation, but the fact that the euchromatin is better transcribed into RNA than the heterochromatin proves that there is an obvious connection between the transcription and the chromatin structure.

#### **74/ The mechanism of prokaryotic translation, the initiation complex, the three steps of elongation; termination, the signal peptide and signal recognition particle**

The major elements of translation are the mRNA, the ribosomes and the aminoacyl-tRNA. The am. ac. tRNA is synthesized by an enzyme (synthetase) in two steps. In the first step it binds the am.ac. to AMP (activated am.ac.). In the second step the AMP is hydrolyzed and the am.ac. is bound to tRNA. There is a special tRNA for each of the 20 am.ac.. The enzyme recognises them and binds the proper am.ac. to the proper tRNA, therefore controls the translation of the genetic code. The aminoacyl-tRNA synthetase is one of the most specific enzyme toward its substrates in the cell.

The initiation complex is formed from formyl-methionin-tRNA, the small ribosomal subunit and the initiation factor (IF). This complex binds the mRNA. The IF dissociates when the larger ribosomal subunit is connected.

The ribosoma has two sites (some model defines a third E site for tRNA exit), one for the growing peptide chain (P site) and one for the incoming aminoacyl-tRNAs (A-site). The peptide bond is formed between the am.ac. on the A site and the peptide chain on the P site by the peptidyltransferase. The growing peptide chain is translocated on the tRNA from the A site to the P site by the EF-2 (elongation factor 2). These steps are repeated until the termination.

Termination occurs when the nonsense codon is recognized by the releasing factor and the translational apparatus dissociates.

The signal peptide at the N-terminal of proteins helps to direct them to the lumen of the endoplasmic reticulum. The ribosome is bound to the receptor protein of the ER (remember, rough ER, with ribosomes on the surface) and the signal peptide (the polymer of hydrophobic am.acs) will bind to the signal recognition particle and enters the ER membrane. Then the signal peptide will be hydrolyzed away from the rest of the protein and it stays in the ER membrane. The proteins are further processed in the Golgi complex.