7 Nucleic acids

7.1 DNA structure and replication

The Watson and Crick model suggested semi-conservative replication

- DNA structure suggests mechanism for DNA replication
- several lines of experimental evidence play role: X-Ray diffraction patterns, Rosalind Franklin's photographs, base composition studies of Erwin Chargaff
- Watson and Crick's first model has sugar-phosphate strands wrapped around each other and nitrogen bases facing outwards; Rosalind Franklin says that nitrogen bases are relatively hydrophobic in comparison to phosphate backbone: nitrogen bases point inwards
- Franklin's X-Ray diffraction suggests DNA is tightly packed
- packing occurs if purine pairs with pyrimidine and bases are upside down to one another
- adenine has a surplus negative charge and thymine a positive one: electrically compatible
- cytosine and guanine can bond by three hydrogen bonds
- complementary base-pairing suggest mechanism for DNA replication: hypothesis of semi-conservative replication

The role of nucleosomes in DNA packing

- nucleosomes help supercoiling DNA
- eukaryotic DNA is associated with histone proteins; prokaryotic DNA is naked (=no proteins)
- histones package DNA into nucleosomes
- nucleosomes are 8 histones (octamer 2x4 different histones) in the core with DNA coiled around
- short „linker" DNA connects nucleosomes
- additional histone protein molecule (H1) binds DNA to core particle
- association of histones with DNA is supercoiling: allows great length of DNA packed into small space
- nucleosome facilitates packing of large genomes; H1 further enhances packing

The leading strand and the lagging strand

- DNA in anti-parallel fashion: synthesis of strands occurs differently
- leading strand is made continuously following the fork
- lagging strand is made in fragments moving from the fork
- lagging strands creates fragments called Okazaki fragments

Proteins involved in replication

- complex system of enzymes carry out DNA replication
- replication involves formation and movement of replication fork and synthesis of leading and lagging strands
- helicase unwinds the DNA
- topoisomerase releases strain developing ahead of helicase
- single-stranded proteins keep strands apart long enough for template stand to be copied
- primers start replication (only one in leading strand, many on lagging strand)
- DNA primase creates RNA primers which initiate DNA polymerase
- DNA polymerase covalently links deoxyribonucleotide monophosphate to 3' of growing strand
- different organisms have different DNA polymerases with different functions as proof-reading, polymerization and removal of RNA primers once no longer needed
- DNA ligase connects gaps between fragments

The direction of replication

- DNA polymerases can only add to 3' end of primer
- DNA replications starts at sites called origins of replication (prokaryotes have one, eukaryotes many)
- replication occurs in both directions from origin
- five carbons of deoxyribose sugar have a number
- phosphate group of new DNA nucleotides is added to 3' carbon of sugar of nucleotide at end of chain: replication occurs in 5' to 3' direction

Non-coding regions of DNA have important functions

- DNA is used as guide for production of polypeptides (only some DNA sequences do: coding sequences)
- many non-coding sequences are found in the genome
- functions of non-coding sequences: act as guide to produce tRNA and rRNA, play role in regulation of gene expression (enhancers and silencers), introns
- most of eukaryotic genome is non-coding
- repetitive sequences can be common in genome, especially in eukaryotes
- two types of repetitive sequences: moderately repetitive and highly repetitive (satellite DNA)
- these sequences can form from 5 to 60% of genome (humans: nearly 60%)
- one area of rep. sequences occurs on ends of eukaryotic chromosomes called telomeres
- telomeres have a protective function: during interphase enzymes cannot replicate DNA until the end of chromosome; without telomeres genes would be lost; they sacrifice repetitive sequence

7.2 Transcription and gene expression

Regulation of gene expression by proteins

- gene expression is regulated by proteins binding to specific base sequences on DNA
- some proteins are always necessary for survival and are expressed in unregulated fashion
- other proteins need to be produced at certain times in certain amounts: regulated expression
- in prokaryotes gene expression is a consequence of variations of environmental factors (e.g. lactose in *E.coli*: genes expressed in presence of lactose, not expressed absence: regulation of gene expression by negative feedback; repressor is deactivated in presence of lactose)
- eukaryotic genes are regulated by variations in environmental conditions; each cell of a multicellular eukaryotic organism expresses only a fraction of its genes
- regulation of eukaryotic gene expression is a critical part of cellular differentiation and process of development
- number of proteins whose bonding to DNA regulates transcription: enhancers, silencers, promoter-proximal elements; sequences linked to regulatory transcription are unique to gene
- enhancers increase rate of transcription when proteins bind to them
- silencers decrease rate of transcription when proteins bind to them
- enhancers and silencers can be distant from promoter, unlike promoter-proximal elements
- binding of proteins to promoter-proximal elements is necessary to initiate transcription

The impact of the environment on the gene expression

- environment of cell and organism has an impact on gene expression
- influence of environment on gene for some traits is unequivocal: environmental factors can affect gene expression such as production of skin pigmentation
- embryo contains an uneven distribution of chemicals called morphogens
- morphogens affect gene expression contributing to different patterns of gene expression

Nucleosomes regulate transcription

- chemical modification of histone tails as important factor in determining expression of a gene
- modifications can be addition of acetyl, methyl or phosphate group
- example: residues of amino acids lysine on histone tails add or remove acetyl groups, lysine residues bear positive charge binding to negative DNA forming a condensed structure inhibiting transcription, acetylation neutralizes positive charges causing higher transcription levels
- chemical modification activates or deactivates genes by decreasing or increasing accessibility of gene to transcription factors

The direction of transcription

- transcription occurs in a 5’ to 3’ direction
- synthesis of mRNA in three stages: initiation, elongation, termination
- transcription begins at promoter; once RNA polymerase binds, DNA is unwound by RNA polymerase, RNA polymerase slides along DNA synthesizing single strand of RNA

Post-transcriptional modification

- mRNA is modified after transcription in eukaryotes

- Regulation of gene expression can occur at several points: transcription, translation, post-translational regulation in eu- and prokaryotes
- in prokaryotes mostly at transcription
- eukaryotes additionally use post-transcriptional modification of RNA
- prokaryotes lack a nuclear membrane: transcription and translation can be coupled
- separate locations for transcription and translation in eukaryotes allows significant post-transcriptional modification before mature transcript exits nucleus (e.g. removal of intervening sequences (introns) from RNA transcript)
- prokaryotic DNA does not contain introns
- immediate product of mRNA transcription is called pre-mRNA (several stages until it's mature)
- one stage is RNA splicing: RNA contains sequences not contributing to formation of polypeptide: introns: must be removed: remaining parts are called exons
- exons will be spliced together to form mature mRNA
- post-transcriptional modification includes addition of 5'-cap (occurs before transcription) and poly-A tail (occurs after transcript has been made)

mRNA splicing

- splicing mRNA increases number of different proteins an organism can produce
- alternative splicing: process when a single gene codes from multiple proteins (occurs in genes with multiple exons)
- proteins translated from alternatively spliced mRNAs will differ in their amino acid sequence and possibly in their biological functions
- example in mammals: Tropomyosin is encoded by gene that has 11 exons
- example in fruit flies: potential 38'000 different mRNAs possible based on number of introns

7.3 Translation

Initiation of translation

- involves assembly of components that carry out the process
- beginning of process is the binding of mRNA to small ribosomal subunit at mRNA binding site
- initiator tRNA molecule carrying methionine binds to start codon AUG
- large ribosomal subunit then binds to small one
- Initiator tRNA is in P site, next codon signals binding to A site, peptide bond is formed between amino acids in P and A site

Elongation of the polypeptide

- synthesis of polypeptide involves a repeated cycle of events
- elongation occurs after initiation
- ribosome translocates three bases along mRNA, moving tRNA to E site: this frees the A site and allows the new tRNA with appropriate anticodon to bind

Termination of translation

- process of elongation continues until a stop codon is reached
- movement along the mRNA is from 5' end to 3' end

Free ribosomes

- in eukaryotes, proteins are synthesized in cytoplasm or at endoplasmic reticulum, depending on final destination of protein
- proteins used in cell (cytoplasm, mitochondria, chloroplasts) are synthesized by free ribosomes

Bound ribosomes

- proteins perform a function within specific compartment of the cell or are secreted
- proteins must be sorted to end up in correct location
- proteins destined for use in endoplasmic reticulum, Golgi apparatus, lysosomes, plasma membrane or outside the cell are synthesized by ribosomes bound to endoplasmic reticulum
- whether a ribosome is bound or free depends on presence of signal sequence on polypeptide; in the first part of sequence: if signal sequence is created it binds to signal recognition protein stopping translation until bound to receptor on surface of endoplasmic reticulum
- translation begins again and polypeptide is created into lumen of endoplasmic reticulum

The coupling of transcription and translation in prokaryotes

- eukaryotes have compartmentalized cellular functions while prokaryotes do not
- as soon as mRNA is transcribed, translation begins

Primary structure

- chain of amino acids is a polypeptide; 20 different common amino acids can be combined in any sequence: huge diversity of proteins
- sequence of amino acids in a polypeptide is the primary structure

Secondary structure

- chain of amino acids in polypeptide has polar covalent bonds within its backbone: chain folds so hydrogen bonds form between carboxyl group (C=O) and amino group (N-H)
- folding creates α-helix and β-pleated sheet and are the secondary structure

Tertiary structure

- overall three-dimensional shape of protein: happens through interaction of R-groups between one another and surrounding water medium
- different types of interaction: positively and negatively charged R-groups, hydrophobic amino acids orientate inwards while hydrophilic turn outwards, polar R-groups form hydrogen bonds between themselves, R-group of amino acid cystine can form covalent bond with another cystine (bond is called disulphide bridge)

Quartenary structure

- proteins can be formed from single polypeptide chain or multiple chains
- quartenary structure is the way polypeptides fit together when there is more than one chain
- biological activity of protein is related to primary, secondary, tertiary and quartenary structure
- treatments as high temperatures, changes in pH can cause alterations in structure of protein and therefore its biological activity; permanent loss of structure is called denaturation