Chapter 7 —> Nucleic acids

7.1 DNA structure and replication

Hershey and Chase:

- Conducted a series of experiments to prove that DNA was the genetic material

1) Viruses (T2 bacteriophage) were grown in one of two isotopic mediums in order to radioactively label a specific viral component
2) Viruses grown in radioactive sulphur (^{35}S) had radio-labelled proteins (sulfur present in proteins but not DNA)
3) Viruses grown in radioactive phosphorus (^{32}P) had radio-labeled DNA (phosphorus present in DNA but not proteins)
4) Viruses were allowed to infect a bacterium (E. Coli)
5) the virus and bacteria were separated via centrifugation

- The bacterial pellet was found to be radioactive when infected by the ^{32}P–viruses (DNA) but not the ^{35}S–viruses (protein) —> pellet is the DNA

Structure of DNA:

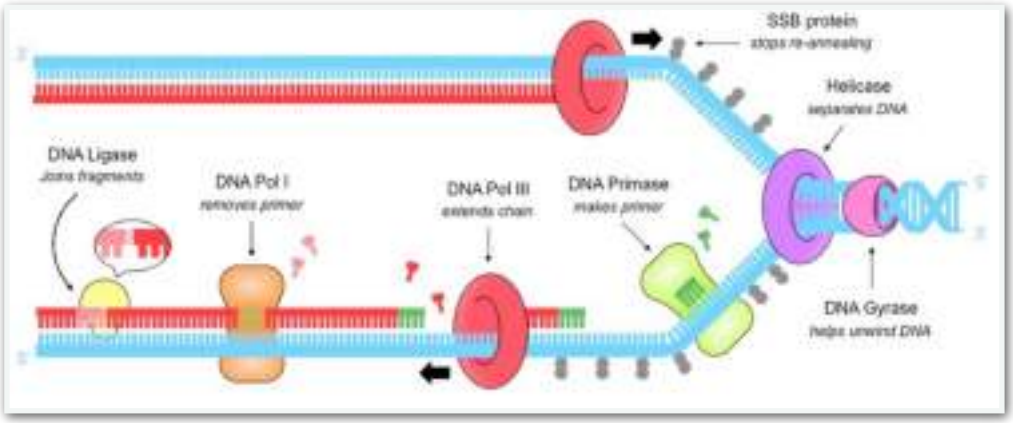
1) DNA was purified and then fibres were stretched in a thin glass tube
2) The DNA was targeted by a X-ray beam, which was diffracted when it contacted an atom
3) The scattering pattern was recorded on a film and used to see details of molecular structure

- Franklin —> Phosphates (and sugars) form an outer backbone and nitrogenous bases are packaged within the interior
- Chargaff —> nitrogenous bases are paired (purine + pyrimidine) within the double helix
  —> the two strands must run in antiparallel directions
- Watson and Crick —> Adenine and thymine (2 hydrogen bonds), guanine and cytosine (3Hy. b)

DNA replication:

- A semi-conservative process that is carried out by a complex system of enzymes
- Helicase —> Unwinds and separates the double-stranded DNA (breaks Hyd bonds of pairs)
  —> occurs at origins of replication
- DNA gyrase —> reduces the torsional strain created by the unwinding of DNA by helicase
  —> it relaxes positive supercoils that would form during the unwinding of DNA
- SSB proteins —> Single stranded. Binding proteins
  —> bind to the DNA strands after being separated and prevents re-annealing
  —> help to prevent the single stranded DNA from being digested by nucleases
  —> will be dislodged from the strand when a new complementary strand is added
- DNA primase —> generates a short RNA primer (10 / 15 nucleotides) on each templated strands
  —> provides an initiation point for DNA polymerase III
- DNA polymerase III —> can extend a nucleotide chain but not start one (5’ to 3’ direction)
  —> free nucleotides align opposite to their complementary base partners
  —> moves in opposite directions on the two strands
  —> leading strand —> moves towards the replication fork (synth. always)
  —> lagging strand —> moves away the replication fork (synth. in pieces)
- DNA polymerase I —> the lagging strand has multiple RNA primers along its length
  —> removes the RNA primers and replaces them with DNA nucleotides

- DNA ligase —> joins the Okazaki fragments together to form a continuous strand
  —> joins the sugar-phosphate backbones together with a phosphodiester bond



Okazaki fragments:

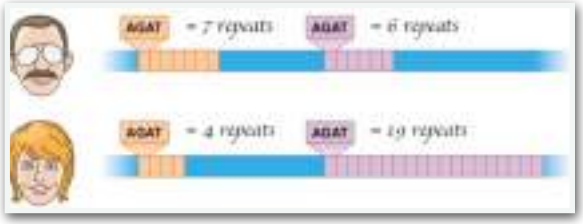
- Needed because DNA polymerase cannot initiate replication
- The lagging stand is copied as a series of short fragments (Okazaki), each preceded by a primer
- The primers are replaced with DNA bases and the fragments are joined together by a combination of DNA pol I and DNA ligase

DNA sequencing:

- The process by which the base order of a nucleotide sequence is elucidated
- Involves the use of chain-terminating dideoxynucleotides
- Dideoxynucleotides —> lack the 3'-hydroxyl group necessary for forming phosphodiester bond
  —> prevent further elongation of a nucleotide chain and terminate replicat.
  —> length will reflect the specific nucleotide position of the the ddNTP
  —> can be used to determine DNA sequence using the Sanger method
- Sanger method —> Four PCR mixes are set up —> (ddA, ddT, ddC, ddG)
  —> fragments are separated using gel electrophoresis
  —> the base sequence is determined by ordering fragments according to length
  —> if a fluorescently labelled primer is present in each mix, the sequence can be detected by using an automated sequencing machine

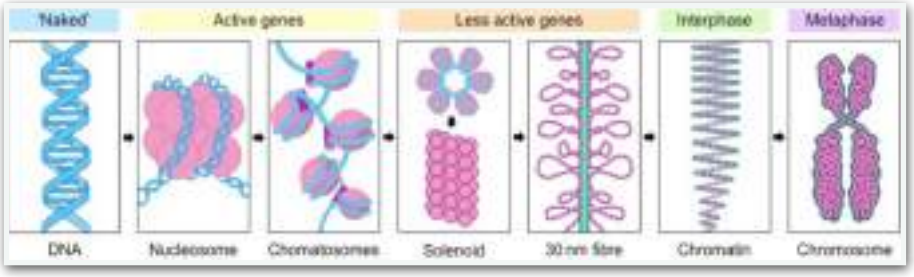
Non-coding DNA:

- Genes only account for 1.5% of the total sequence —> the rest is non-coded DNA
- DNA profiling —> a technique by which individuals can be identified and compared via their DNA profiles
- Individuals will likely have different numbers of repeats at a given satellite DNA locus, so will generate a unique DNA profile



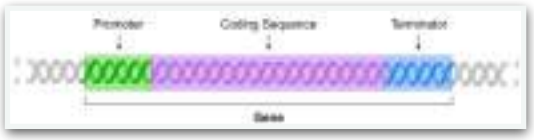
Nucleosomes:

- Help to supercoil the DNA —> better compacted structure that allows for efficient storage
- Helps protecting the DNA from damage and allows chromosomes to move during reproduction
- Eukaryotic DNA —> complexed with eight histone proteins (forms nucleosome)



- Nucleosomes —> a molecule of DNA wrapped around a core of eight histone proteins
  —> the negatively charged DNA associates with the positively charged amino acids on the surface of the histone proteins
  —> the histone proteins have N-terminal tails extruding outwards of nucleosome
  —> tails link up and draw the nucleosomes closer during chromosomal condens.

7.2 Transcription



Sections of a gene:

- Gene —> a sequence of DNA which is transcribed into RNA and contains three main parts
- Promoter —> the non-coding sequence responsible for the initiation of transcription
  —> typically located immediately upstream of the gene's coding sequence
  —> functions as a binding site for RNA polymerase (transcription RNA)
  —> binding is mediated and controlled by some transcription factors in eukaryotes
  —> transcription factors bind to the proximal control element or distal control elements
- Coding sequence —> the region of DNA that is transcribed by RNA polymerase
  —> after RNA polymerase has bound to the promoter
  —> it causes the DNA strands to unwind and separate
- Terminator —> RNA polymerase will continue to transcribe the DNA until a terminator sequence
  —> the mechanism differs between prokaryotes and eukaryotes

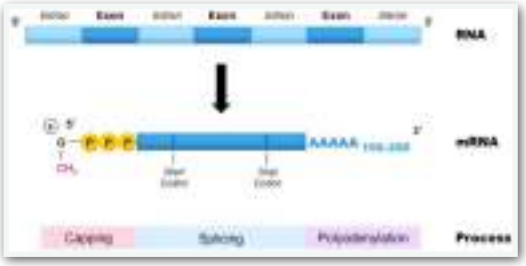
Antisense vs Sense:

- A gene consist of two polynucleotide strands (just one is transcribed into RNA)
- Antisense strand —> the strand transcribed into RNA (template strand)
  —> the DNA version of the tRNA anticodon sequence
- Sense strand —> the strand not transcribed into RNA
  —> the DNA version of the RNA sequence
  —> the coding strand

Transcription:

- The process by which a DNA sequence is copied into a complementary RNA sequence by RNA polymerase
- Nucleotides triphosphates —> free nucleotides inside a cell (NSPs)
- RNA polymerase covalently binds them together in a reaction involving two phosphate release
- The 5'-phosphate is linked to the 3'-end of the growing mRNA strand (5' to 3' direction)

Messenger RNA:



-In eukaryotes three post-transcriptional events must occur in order to form an mRNA

-Capping —> involves the addition of a methyl group to the 5'-end of the transcribed RNA

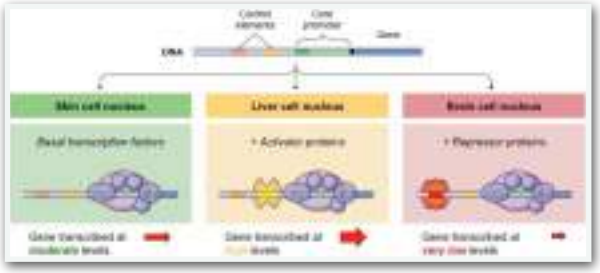
—> provide protection against degradation by exonuclease

—> it allows the transcript to be recognised by the cell's translation machinery

- Polydenylation —> the addition of a long chain of adenine nucleotides to the 3'end of transcript
  —> improves stability of the RNA transcript and facilitates export from nucleus
- Splicing —> introns —> non-coding sequences which must be removed prior to form mRNA
  —> exons —> coding regions that are fused together when introns are removed
  —> introns are intruding sequences while exons are expressing sequences
  —> can also result in the removal of exons —> alternative splicing —> will result in the formation of different polypeptides from a single gene sequence

Gene expression:

- Transcriptional activity is regulate by two groups of proteins
- Transcriptional factors form a complex with RNA polymerase at the promoter
- RNA polymerase cannot initiate transcription without these factors



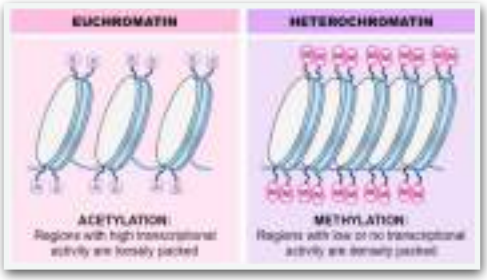
- Regulatory proteins bind to DNA sequences outside of the promoter and interact with factors
  —> activator proteins bind to enhancer sites and increase the rate of transcription
  —> repressor proteins bind to silencer sequences and decrease the rate of transcription
  —> the presence of certain transcription factors may be tissue-specific
  —> chemical signals can also moderate protein levels
- Control elements are the DNA sequences to which regulatory proteins bind to
  —> some are proximal elements (near to promoter) and others are distal elements
  —> regulatory proteins typically bind to distal control elements while transcription factors usually bind to proximal elements
  —> most genes have multiple control elements —> gene expression is tightly controlled
- Changes in the external or internal environment can trigger changes to gene expression patterns
  —> chemical signals within the cell can change the levels of regulatory proteins or factors in response to stimuli
  —> this allows gene expression to change in response to alterations in the environments
  —> ex. humans produce different amounts of melanin depending on light exposure

Modification of Histone tails:

- Histone tails have a positive charge and hence associate tightly with the negatively charged DNA
- By adding an acetyl group, the tail neutralises —> DNA il less tightly coiled and < transcription
- By adding a methyl group, the tail remains positive —> DNA more coiled and > transcription

Types of chromatin:

- Heterochromatin —> DNA is supercoiled/condensed and so is not accessible for transcription
- Euchromatin —> DNA is loosely packed and is accessible to the transcription machinery
- Different cell types will have varying segments of DNA packaged as hetero/euchromatin
- Some segments of DNA may be permanently supercoiled



DNA methylation:

- Can also affect gene expression patterns (prevents the binding of transcription factors)
- Increased methylation of DNA decreases gene expression —> genes that are not transcribed tend to exhibit more DNA methylation than genes that are actively transcribed
- Ex. Maternal diet, exposure to microbes, environmental, diet, lifestyle and age-related changes

Epigenetics:

- The study of changes in phenotype as a result of variations in gene expression levels
- Shows that DNA methylation patterns may change over the course of a lifetime
- It is influence by heritability but is not genetically pre-determined
- Different cell types in the same organism may have totally different DNA methylation patterns
- Environmental factors may influence the level of DNA methylation within cells

---

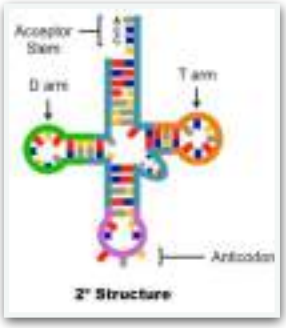
7.3 Translation

Ribosomes:

- are made of protein (for stability) and ribosomal RNA (for catalytic activity)
- Two subunits —> small subunit contains an mRNA binding site
  —> large subunit contains three tRNA binding sites (aminoacyl, peptidyl, exit)
- Can be found freely floating in the cytosol or bound to the rER for eukaryotes
- Prokaryotes —> 70s Eukaryotes —> 80s

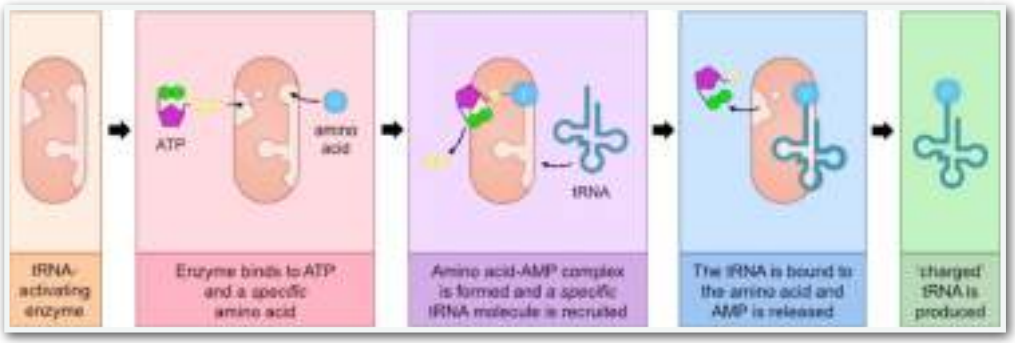
Transfer RNA:

- tRNA molecules fold into a cloverleaf structure with four key regions:
  —> the acceptor stem (3'-CCA) —> carries an amino acid
  —> the anticodon —> associates with the mRNA codon
  —> the T arm —> associates with the ribosome (E, P, A sites)
  —> the D arm —> associates with the tRNA activating enzyme



tRNA activation:

- Each tRNA molecule bind with a specific amino acid in the cytoplasm —> each amino acid is recognised by a specific enzyme
- The binding of an amino acid to the tRNA acceptor stem occurs as:
  —> enzyme binds ATP to the amino acid to form an amino acid —> AMP complex linked
  —> the amino acid is then coupled to tRNA and the AMP is released —> now charged
- The function of the ATP is to create a high energy bond that is transferred to the tRNA molecule
  —> this energy will provide the energy required for peptide bond formation in translation



Translation:

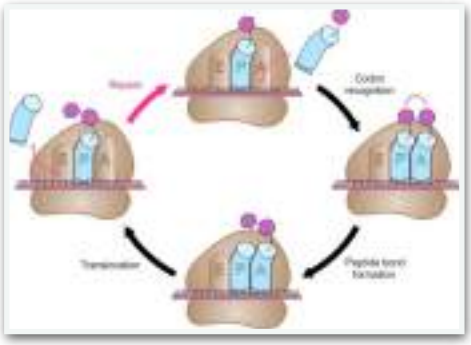
- Initiation —> the first stage and involves the assembly of the three components that carry out the process (mRNA, tRNA, ribosome)
- The small subunit binds to the 5'-end of the mRNA and moves along it until it reaches AUG
- Then the appropriate tRNA molecule binds to the codon via its anticodon (comp. base pairing)
- Lastly the large subunit aligns itself to the tRNA molecule at the P site and forms a complex

Elongation:

- A second tRNA molecule pairs with the next codon in the ribosomal A site
- The amino acid in the P site is covalently attached via a peptide bond to the amino acid in A
- The tRNA in the P site is deacylated (no amino acid)

Translocation:

- The ribosome moves along the mRNA strand by one codon position (5' to 3')
- The deacylated tRNA moves into the E site and is released while the tRNA on A goes to P
- Another tRNA molecule attaches to the now unoccupied A site
- The process is repeated until the stop codon



Termination:

- Final stage of translation —> involves the disassembly of the components and the release of the polypeptide chain
- Elongation and translocation continue until a stop codon is reached
- These codons do not recruit a tRNA molecule, but instead a release factor
- The polypeptide is released and the ribosome disassembles back in its two independent subunits

Polysomes:

- In eukaryotes, ribosomes are separated from the genetic material by the nucleus
- After transcription the mRNA must be transported from the nucleus (pores) prior to translation
- This transport requires modification to the RNA construct

- Prokaryotes lack a compartmentalised structure, so transcription and translation are not separated
- Ribosomes begin translating the mRNA molecule while it is still being transcribed from the DNA
- Possible because both transcription and translation occur in a 5' to 3' direction

- A polysome is a group of two or more ribosomes translating an mRNA sequence simultaneously
- Will appear as a bead (ribosomes) on a string (mRNA)
- Polysomes in prokaryotes may form while the mRNA is still being transcribed from the DNA

- Ribosomes located at the 3'-end of the polysome cluster will have a longer polypeptide chain

Protein destinations:

- If the protein is targeted for intracellular use within the cytosol —> ribosome free and unattached
- If protein is targeted for secretion, membrane fixation or lysosomes —> ribosome binds to ER

- Destination is determined by the presence or absence of an initial signal sequence on a chain
- The presence of a signal sequence results in recruitment of a signal recognition particle (SRP)

1) SRP halt translation in ribosomes
2) The SRP-ribosome complex is docked to a receptor located on the ER membrane
3) Translation is re-initiated and the polypeptide chain continues to grow into a transport channel to finish inside the ER lumen
4) The synthesised protein will then be transported via a vesicle to the Golgi complex or lysosome
5) Proteins targeted for membrane fixation get embedded into the ER membrane
6) The signal sequence is cleaved and the SRP recycled once the chain is fully synthesised

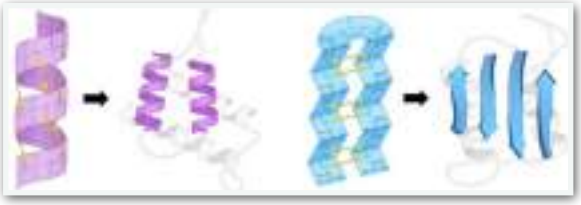
Protein structure:

Primary structure:

- The order of amino acids from which the polypeptide chain is comprised
- Formed by covalent peptide bonds between the amine and carboxyl groups of adjacent amino.
- It controls all subsequent levels of protein organisation —> determines the interactions between R groups of different amino acids

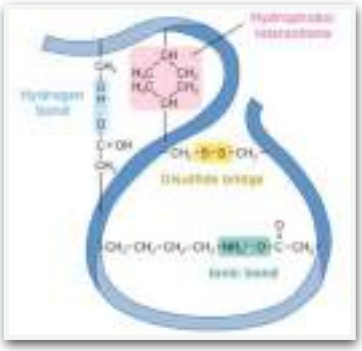
Secondary structure:

- Way a polypeptide folds in a repeating arrangement, forms Alpha helices or Beta-pleated sheets
- Results of hydrogen bonding between the amine and carboxyl groups of non-adjacent amino.
- Random coil —> sequences with neither alpha or beta arrangement
- Provides the polypeptide chain with a level of mechanical stability



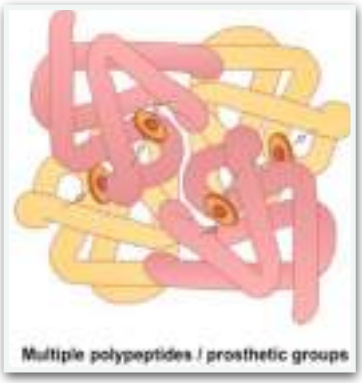
Tertiary structure:

- The way the polypeptide chain coils and turns to form a complex molecular shape
- Caused by interactions between R groups (disulphide bridges, ionic bonds, hydrophobic interactions)
- Relative amino acid positions are important
- It may be important for the function of the protein



Quaternary structure:

- Multiple polypeptides or prosthetic groups, interact to form a single, larger, biologically active protein
- Prosthetic group —> an inorganic compound involved in protein structure or function
- A protein containing a prosthetic group is called a conjugated protein (haemoglobin)
- May be held together by a variety of bonds



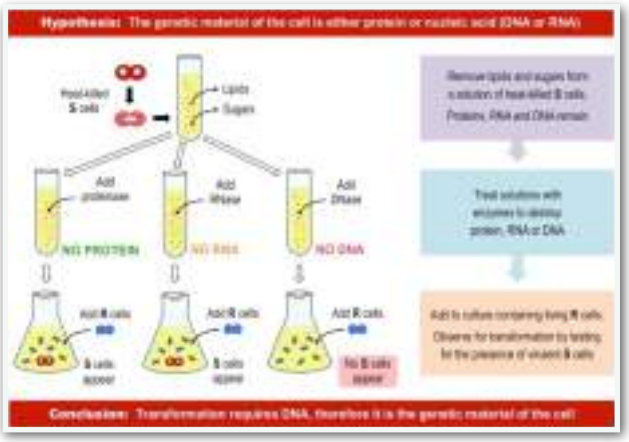
Extra:

DNA experiments:

Griffith's experiment:

- 1928, Frederick Griffith —> one of first experiments to show that cells posses genetic material
- Involved the use of two strains of pneumococcus (deadly virus strain and non-virulent strain)
- When Griffith infected mice with the non-virulent bacteria the mice survived
- When Griffith infected mice with the virulent bacteria the mice died
- When Griffith infected mice with a mix of heat-killed strain virulent bacteria and non-virulent living strain the mice were found dead —> there were some form of genetic material transfer

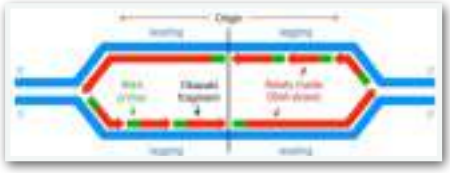
Avery-Maclord-McCarty experiment:



DNA structure elucidation:

- Rosalind Franklin and Maurice Wilkins used X-ray diffraction to identify key properties of DNA
- Wilkins shared this data with Watson and Crick
- W and C used this data to help construct an accurate model of DNA structure (double helix)
- Watson, Crick and Wilkins were awarded the Nobel prize, but not Franklin

Origins of replication:



- Sequences where DNA replication is initiated in a genome
- DNA synthesis may occur bi-directionally from an origin of replication —> greatly limits the time required for the process —> when it happens the two replication forks move in opposite directions to create a replication bubble
- Replication bubbles expand in both directions —> will fuse together as intervening regions copy

Supercoiling:

- Refers to the additional twisting of a DNA strand and is an expression of the strain on that strand
- Can be overwound —> positive supercoiling
- Can be underwound —> negative supercoiling —> most DNA is like this
- Supercoiling functions to reduce the space required for DNA packaging
- DNA will form positive supercoils when unwound by helicase and requires DNA gyrase to reduce the strain

Genes versus repetitive DNA:

	Simple copy Gene	Repetitive DNA
Proportion	Small (~1.5%)	High (5 – 45%)
Rate of mutation	Low	Higher
Occurrence	Once in genome	Occurs many times
Function	Makes protein	Not translated
Identification	Similar b/w individuals Not used for profiling	Varies greatly Used for profiling
Length	Long unique sequence	Short repeating sequence
Example	Exons	Introns

Telomerase:

-Regions of repetitive DNA located at each end of chromatid (prevent chromosomal deterioration)
-The extreme end of the telomere cannot be copied —> gets marginally shorter —> the terminal RNA primer on the lagging strand cannot be replaced
-The progressive shortening of telomeres is ageing
-Cells have limited capacity for cellular division (Hayflick limit —> 40/60x)
-Telomeres can be lengthened by enzyme telomerase
-Permanent activation of telomerase can cause cells to become immortal and leads to cancer

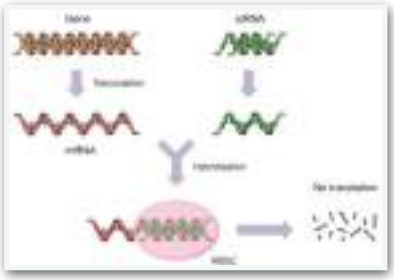
Types of RNA:

- RNA functions to transfer genetic instructions from the nucleus to the cytoplasm
- Messenger RNA (mRNA) —> a transcript copy of a gene which encodes a specific polypeptide
- Transfer RNA (tRNA) —> carries the polypeptide subunits (amino acids) to the organelle responsible for synthesis (ribosome)
- Ribosomal RNA (rRNA) —> primary component of the ribosome —> responsible for catalysis
- Cells may produce other variants of non-coding RNA to support and regulate gene expression:
    —> small nuclear RNA —> component of the spliceosome (involved in intron splicing)
    —> short interfering RNA —> moderates gene expression levels via RNA interference

Operons:

- A sequence of DNA containing a cluster of genes under the control of a single promoter
- Genes within an operon will always be expressed together or not at all
- Three basic components to an operon:
    —> Promoter —> upstream sequence to which RNA polymerase binds
    —> Operator —> Segment of DNA to which a repressor protein binds
    —> Structural genes —> genes that are collectively regulated by the operon
- Operons are related to stimulons —> set of genes under regulation from a single cell stimulus while a regulon is a set of genes under regulation from a single regulatory protein

RNA interference:



- Short interference RNA is a double-stranded RNA molecule that is roughly 20-25 base pairs
- SiRNA interferes with expression of genes —> mRNA transcripts to be broken prior translation
- RISC —> RNA induced silencing complex

Protein modification:

- Post-translational modif. —> all change in the chemical composition of proteins after translation
- These modifications may be vital to the formation of a mature functional protein
- Addition of new functional groups —> enzymes modify protein structure via the introduction of a new chemical group to specific amino acids in the molecule (phosphorylation, acetylation, …)
  —> can alter the properties of the chain and induce conformational change affecting activity
- Proteolytic cleavage of existing elements —> proteins may also be modified via the removal of specific amino acid segments from a propeptide
  —> occurs in zymogens —> active site is occluded and inactive untile proteolytic … occurs
  —> insulin requires the separation of a middle segment to form two polypeptides linked by disulphide bridges
- Racemization —> amino acids can exist as chiral enantiomers (mirror image)
  —> involves converting proteins from one enantiomeric arrangement to another
  —> different enantiomers may have distinct chemical properties

Protein expression:

1) Transcription control —> by controlling the amount of transcription (less mRNA = less protein)
   —> achieved primarily through the effects of transcription factors and regulatory proteins
2) RNA processing con. —> involves regulation the formation of mature mRNA in eukaryotes
   —> necessary to help direct the mRNA to the ribosome and prevent premature degradation
3) RNA transport con. —> RNA must be transported out of the nucleus to associate with ribosome
   —> preventing transport, the related protein cannot be synthesised
4) Translation con. —> protein expression can be regulated by controlling amount of translation
   —> may involve inhibiting ribosomal subunit or actively targeting mRNA for degradation
5) Protein activity con. —> protein expr. patterns may be affected by the rate of prot. degradation
   —> post-translational modification may target proteins for destruction (ex. phosphorylation)

Visualising proteins:

- Haemoglobin —> a globular protein responsible for oxygen transport within red blood cells
  —> it has a quaternary structure made up of four polypeptide subunits
  —> each polypeptide chain is associated with a prosthetic heme group
- Aquaporins —> are integral membrane proteins that form channels that allow water passage
  —> form tetramers in the cell membrane with each monomer allowing water move
  —> are impermeable to charged species —> prevent passage of ions or solutes
- Keratin —> is a fibrous protein that functions as a key structural material in hair, skin and nails
  —> form long twisted stands that may interconnect via disulphide bridges
- Green fluorescent protein —> is a fluorophore produced by jellyfish
  —> the fluorescing chromophore is attached to a central alpha helix surrounded by 11 beta
  —> the tightly packed beta barrel excludes solvent molecules