LN's

Endless

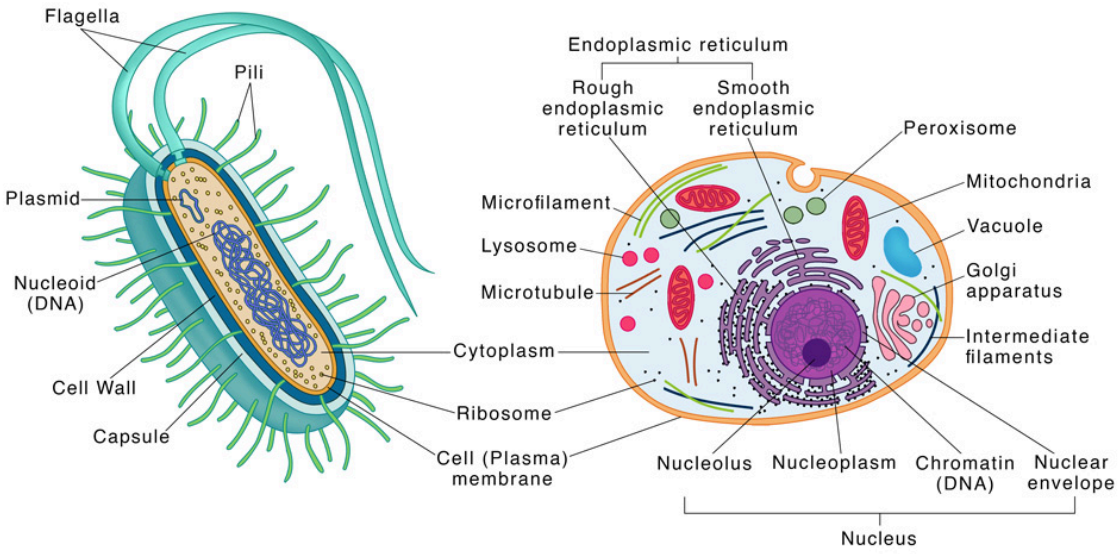
Biology Notes

Art:
Kira, by @Ekkoberry

B17. MOLECULAR BIOLOGY

17.1. Cellular Biology and Metabolism

17.1.1. Generalised Structure of Cells



Prokaryotic Cell

(Bacteria, Archaea)

Eukaryotic Cell

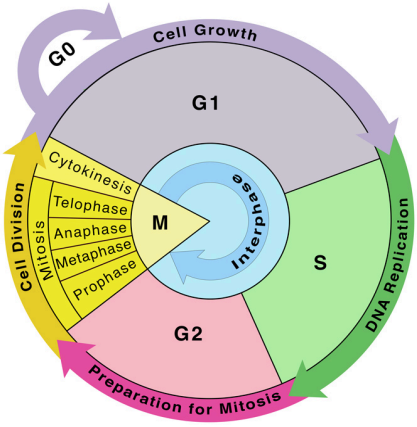
(Protists, Fungi, Animals, Plants)

- **Plant cells additionally include a cellulose cell wall and chloroplasts** (not shown).
- Gram-positive bacteria (shown left) have a thick peptidoglycan cell wall and a thin inner cell membrane.
- Gram-negative bacteria (e.g. *E. Coli*) have a thin cell wall and two cell membranes inside and outside.
- The capsule (slime layer) on a bacteria (shown above) is not always present.
- Hypertonic (concentrated) media lead to plasmolysis (shrinking). Hypotonic (dilute) media lead to cytolysis (swelling).

Structure and Functions of the Main Components and Organelles:

- **Cell membrane:** a phospholipid bilayer bounding the cytoplasm.
- **Cell wall:** a cellulose barrier giving plant and algal cells structure.
- **Cytoplasm:** the interior of the cell, composed of the organelles in an aqueous solution (cytosol).
- **Ribosome:** contains the cell machinery to synthesise proteins from mRNA (translation).
- **Mitochondria:** where respiration occurs to release energy for the cell.
- **Vacuole:** stores water and nutrients as cell sap to maintain the cell's turgor pressure.
- **Nucleus:** contains chromosomes (bundled DNA as chromatin). The nucleolus synthesises ribosomes.
- **Plasmid:** bacterial circular DNA, which can be exchanged with other bacterial cells to pass along genes.
- **Endoplasmic reticulum:** the RER is a ribosome-dense region where proteins are modified for their function.

17.1.2. The Cell Cycle



Healthy, non-stem, somatic cells can undergo ~50 cell divisions before entering cell senescence (the Hayflick limit) due to telomere attrition (Section 17.2.3). Stem cells and gamete cells have much longer replication limits, and cancer cells can divide forever (immortality).

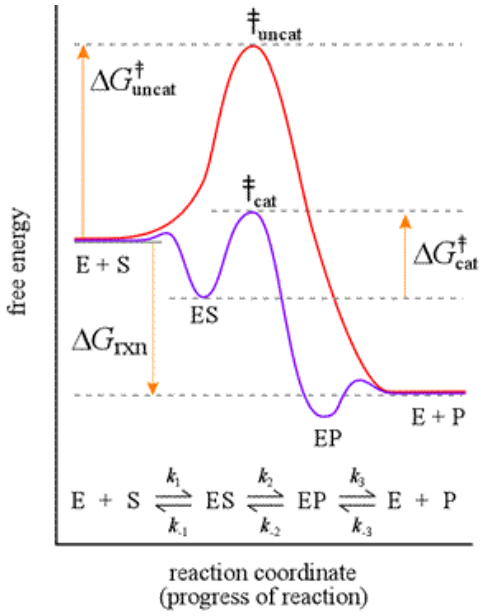
17.1.3. Basic Functions of Common Enzymes

Enzymes are biologically-derived proteins with the ability to act as enantioselective, stereospecific, homogeneous catalysts.

Enzyme name	Biotransformation typically catalysed *in vivo*
Invertase (sucrase)	sucrose → glucose + fructose
Maltase	maltose → 2 glucose
Lactase	lactose → glucose + galactose
Amylase	starch → n glucose
Emulsin	cellulose → n glucose
Lipase	triglyceride (lipids) + H_2O → triol + carboxylic acid
Urease	$NH_2CONH_2 \rightarrow CO_2 + 2\ NH_3$
Carbonic anhydrase	$H_2CO_3 \rightarrow CO_2 + H_2O$
Pepsin	proteins → α amino acids (acidic conditions)
Trypsin	proteins → α amino acids (basic conditions)
Nucleases	DNA or RNA → nucleotides
DNA or RNA polymerase	nucleotide phosphates → DNA or RNA
Reverse transcriptase	RNA → DNA

For more on enzymes, including their use in biotechnology, see Section 17.1.11-15.
For the chemical structures of enzymes and other biomolecules, see Section 16.5.

17.1.4. Thermodynamics of Enzymatic Metabolic Reactions



Thermodynamics of the Enzyme-Substrate Binding Model

The binding step E + S → ES is entropically unfavourable but exothermic due to the electrostatic stabilisation in the active site, so is exergonic.

The reaction ES → EP is then performed in the active site, with lower activation energy and so at much higher rate:

$$\Delta G^{\ddagger}_{cat} < \Delta G^{\ddagger}_{uncat} \Rightarrow k_2 >> k_{\text{uncat}}$$

The product then dissociates, which may be either exergonic or endergonic (entropy and enthalpy effects compete).

Catabolic processes (bond breaking) are net **exergonic**. **Anabolic** processes (bond forming) are net **endergonic**.

Non-Spontaneous Enzymatic Reactions: when $\Delta G_{rxn} > 0$

If the reaction to be catalysed is endergonic, free energy must be supplied to the system in order to render it spontaneous. Most bond-forming reactions (e.g. DNA replication, protein translation, photosynthesis) are endergonic. The energy may be provided from various sources depending on the reaction and context, such as light, a proton gradient or chemical energy (e.g. ATP). In DNA synthesis, DNA polymerase reacts deoxyribonucleotide triphosphates (dNTPs) with the DNA chain, forming the bond and releasing inorganic pyrophosphate (PP_i).

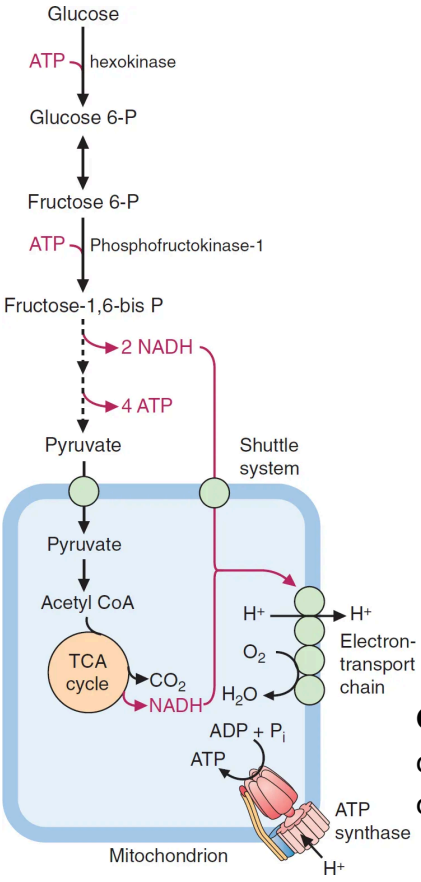
Cells as Open Systems: non-equilibrium allows for spontaneous local structural organisation

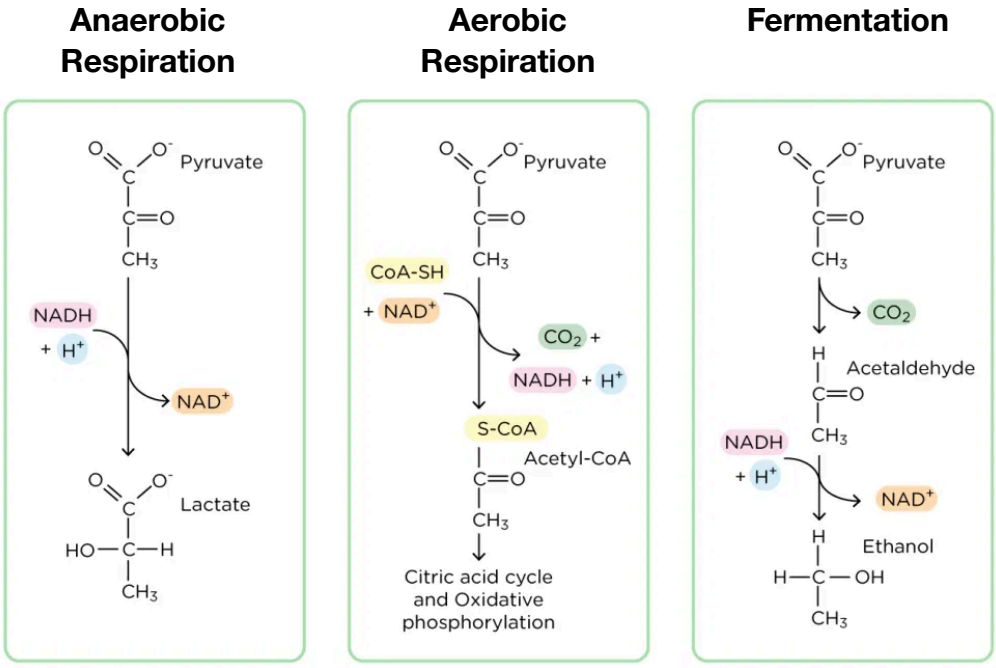
- A cell maintains low-entropy homeostasis by increasing the entropy of its surroundings by release of heat to the environment.
- A cell uses available energy sources to drive endergonic reactions away from equilibrium.
- Any cell or organism in equilibrium with its environment is dead.

Thermodynamics of Photosynthesis: $Q < 0$: heat absorbed by system, $W > 0$: work done by system

The reaction 6 CO_2 + 6 H_2O → $C_6H_{12}O_6$ + 6 O_2 has $\Delta G > \Delta H > 0$, $\Delta S < 0$. The available non-pV work supplied by the sunlight is $W_{\max} = \Delta G$, and $\Delta H = Q + W$ ($W \leq W_{\max}$). The reaction releases heat $Q \geq T\,\Delta S$ to the environment, increasing its entropy as $\Delta S_{\text{env}} = -Q / T \geq -\Delta S$, so overall entropy $\Delta S + \Delta S_{\text{env}} \geq 0$. Real photon absorption is irreversible, leading to faster entropy generation and lower exergy efficiency ($W = \eta W_{\max} \rightarrow -\Delta S / \Delta S_{\text{env}} = \eta$ where $\eta \approx 0.35$).

17.1.5. Respiration





Glycolysis is the conversion of glucose (sugar from foods) to pyruvate. It is the common first part of respiration in all life. It is a 10-step enzymatic process with overall reaction:

glucose + 2 NAD^+ + 2 ADP + 2 P_i → 2 **pyruvate** + 2 NADH + 2 H^+ + 2 **ATP** + 2 H_2O

The pyruvate can then be metabolised in different ways: ($P_i = PO_4^{3-}$: inorganic phosphate)

Aerobic Respiration: in mitochondria, with oxygen present, pyruvate is converted to acetyl-CoA:

pyruvate + CoA-SH + NAD^+ → **acetyl-CoA** + NADH + H^+ + CO_2

The acetyl-CoA is then used in the citric acid cycle (Section 17.1.4) and then to the electron transport chain (Section 17.1.5), which produces a total of up to 38 ATP per glucose.

Anaerobic Respiration: in the cytoplasm, pyruvate is converted to lactate:

pyruvate + NADH + H^+ → **lactate** + NAD^+

Lactate is typically oxidised to lactic acid, yielding no further ATP, but can also be used to regenerate glucose by gluconeogenesis in the liver (reverse glycolysis, consuming 6 ATP).

Fermentation: in both prokaryotes and eukaryotes, pyruvate is converted to ethanol:

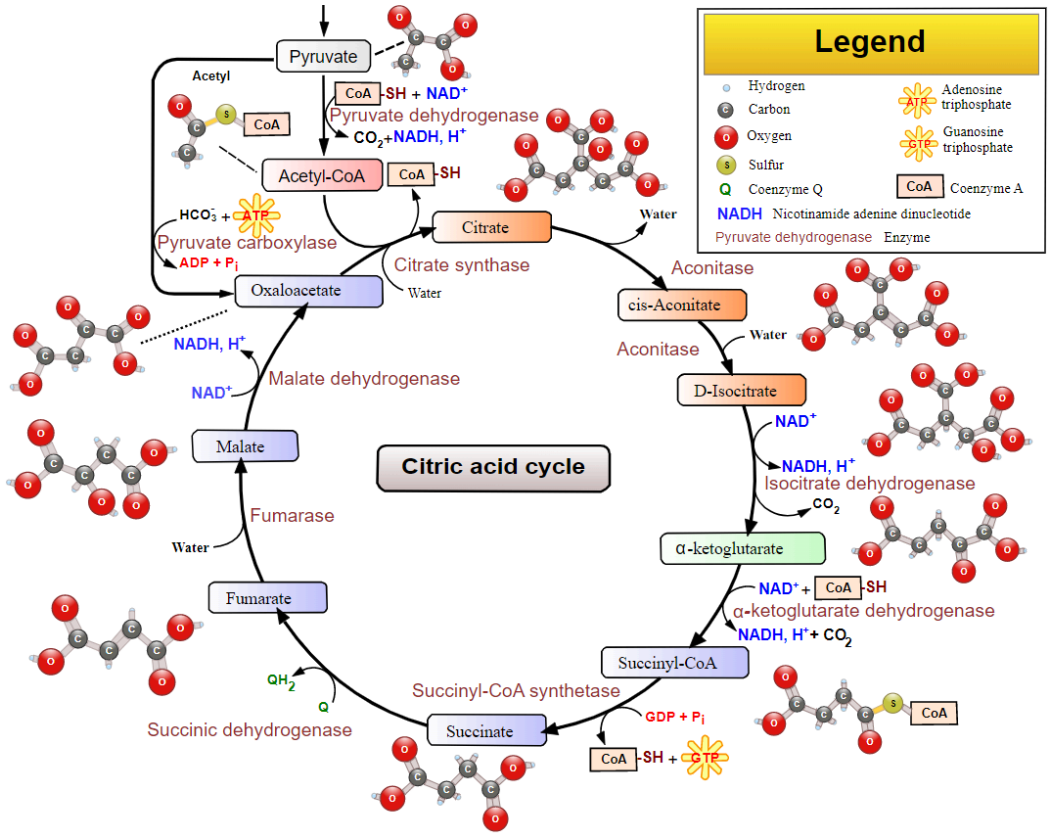
pyruvate + NADH + H^+ → **ethanol** + NAD^+ + CO_2

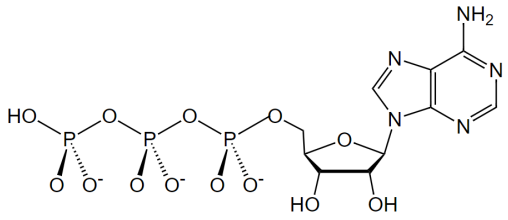
17.1.6. Citric Acid Cycle (Krebs Cycle; TCA Cycle) for Aerobic Respiration

The citric acid cycle phosphorylates GDP to GTP (guanosine triphosphate) from pyruvate (from glycolysis). On a per-cycle basis, the overall reaction is:

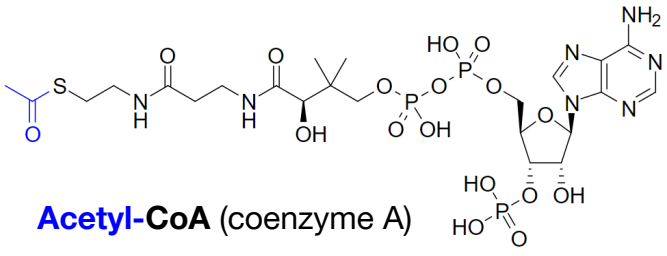
acetyl-CoA + 3 NAD^+ + FAD + **GDP** + P_i → 2 CO_2 + 3 NADH + 3 H^+ + $FADH_2$ + **GTP** + CoA-SH

GTP can then be converted to ATP via nucleoside-diphosphate kinase: GTP + ADP → ATP + GDP.

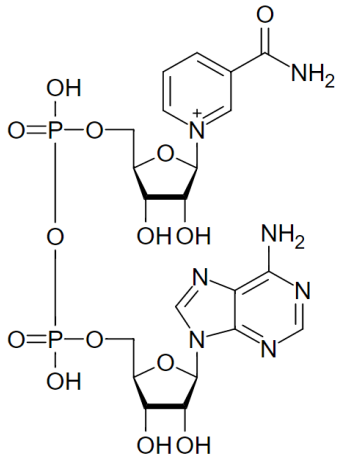




ATP (adenosine triphosphate)



Acetyl-CoA (coenzyme A)



NAD (nicotinamide adenine dinucleotide)

17.1.7. Oxidative Phosphorylation and the Electron Transport Chain

Mitochondria convert ADP to ATP, which provides the necessary chemical energy to carry out a wide range of catabolic reactions.

Electron Transport Chain: sets up and maintains a proton-motive force gradient across the inner mitochondrial membrane (or the plasma membrane in prokaryotes).

1. **Delivery of electrons by NADH and $FADH_2$.** Reduced electron carrier cofactors (NADH and $FADH_2$) from glycolysis/aerobic respiration/citric acid cycle transfer their electrons to molecules near the beginning of the transport chain. In the process, they turn back into NAD^+ and FAD, which can be reused in other steps of cellular respiration.
2. **Electron transfer and proton pumping.** As electrons are passed down the chain, they move from a higher to a lower energy level, releasing energy. Some of the energy is used to pump protons, moving them out of the matrix and into the intermembrane space. This pumping establishes an electrochemical proton gradient across the membrane (outside: high H^+, inside: low H^+).
3. **Splitting of oxygen to form water.** At the end of the electron transport chain, molecular oxygen is the electron acceptor, which splits and takes up H^+ to form water: $½\ O_2 + 2\ e^- + 2\ H^+ \rightarrow H_2O$.

Chemiosmosis-driven synthesis of ATP: as protons flow down their gradient from the intermembrane space back into the mitochondrial matrix, they pass through the molecular machine enzyme ATP synthase, which harnesses the proton flux to synthesise ATP.

$$\textbf{ADP} + P_i + 2\ H^+ \text{(out)} \rightleftharpoons \textbf{ATP} + H_2O + 2\ H^+ \text{(in)}$$

ATP synthase consists of two main parts: the F_O proton channel embedded in the membrane and the F_1 catalytic domain extending inside the matrix. F_O consists of the *c*-ring ('rotor/turbine') and γ subunit ('axle'), while F_1 ('stator') is an assembly of alternating nucleotide (ADP)-binding proteins (α and β) whose conformations are changed between 'closed' and 'opened' by the continuously-rotating γ subunit, permitting reaction with ADP and dissociation of ATP. Two of the proteins in the *c*-ring are encoded in the mtDNA genome (Section 17.2.4).

The same process occurs in chloroplasts through the thylakoid membrane, with the electron transport chain being initiated by photochemical oxidation (photosynthesis) rather than NADH. The resulting ATP is used by plants in the light-independent reactions of photosynthesis (Calvin cycle, Section 17.3.4). In vacuoles, this system works in reverse, acting as 'V-ATPase', creating a proton gradient by consuming ATP (active transport).

In prokaryotes, these processes all occur across the cell membrane and are used for chemotaxis (Section 17.3.6) via the molecular motor at the base of their flagella.

17.1.8. Glucose Transport and Metabolism in Animals

When a meal is eaten, a series of events allows for the energy to be released in the body:

1. **Digestion:** salivary amylase enzymes in the mouth partially breaks down starch (more complete for soluble starch) into oligosaccharides (e.g. maltotriose, dextrins) and disaccharides (e.g. maltose). On entering the small intestine, pancreatic amylase, as well as maltase, lactase and sucrase break these down completely into simple sugars (e.g. glucose).
2. **Transport:** Glucose undergoes active transport into enterocytes lining the small intestines, through a sodium-ion symport transporter protein SGLT1. Glucose then leaves the cells into the hepatic portal vein through GLUT2 (a glucose transporter protein) by facilitated diffusion, **entering the bloodstream.** This occurs within ~1 hour of eating.
3. **Glucose sensing:** In the islets of Langerhans of the pancreas, glucose enters beta cells via GLUT2, undergoes metabolism to produce ATP. An ion channel cascade leads to exocytosis of vesicles containing insulin, **releasing insulin hormone** into the bloodstream.
4. **Insulin action:** Insulin binds to its receptor (IR) on muscle cells, fat cells and liver cells, where they induce production of vesicles containing GLUT4, which anchor within the cell membrane (through the cascade: phosphodiesterase activation → cAMP levels drop → activation of protein kinase B (Akt) → translocation of GLUT4). The increased number of cells expressing GLUT4 transporters allow for **more glucose uptake** by muscle and fat cells.
5. **Respiration:** Intracellular glucose rapidly undergoes glycolysis to pyruvate in the cytosol, then is transported into the mitochondria to undergo the citric acid cycle, which drives the electron transport chain, which in turn drives ATP synthase to generate ATP (Section 17.1.5-7). The ATP is used as chemical energy in a wide variety of metabolic processes.

Glycogenesis: glucose-rich blood enters hepatocytes in the liver via GLUT2. Excess glucose can then be stored as glycogen (glycogenesis), which requires signalling from insulin.

When blood glucose is low, pancreatic alpha cells release **glucagon** hormone. Glucagon acts to raise blood glucose by 1) stimulating **glycogenolysis**, 2) promoting **gluconeogenesis**, 3) inhibiting glycolysis, 4) promoting **lipolysis**. Insulin and glucagon together form a negative feedback control system by which dietary variations in glucose are smoothed out over time, maintaining cellular homeostasis.

Type 1 Diabetes: an autoimmune disease in which pancreatic beta cells are attacked, blocking insulin production, leading to insufficient glucose uptake by cells in the body. Glucose therefore remains at high levels in the bloodstream (hyperglycemia). This promotes lipolysis, from which ketone metabolites accumulate leading to metabolic acidosis. Type 1 diabetes requires regular insulin injection.

Type 2 Diabetes: cells become resistant to the presence of insulin, rendering it ineffective at releasing GLUT4 transporters, leading to insufficient glucose uptake by cells in the body. The mechanism by which GLUT4 production is impaired in type 2 diabetes is not fully understood. Type 2 diabetes is generally associated with obesity; losing visceral fat and eating less sugar often resolves type 2 diabetes by increasing insulin sensitivity.

17.1.9. Gas Exchange in Animals

Nonpolar gas molecules such as O_2 and CO_2 are soluble in lipids and can freely cross cell membranes by diffusion. Oxygen gas is inhaled from the atmosphere into the lungs, through the bronchioles and diffuses through alveoli into capillaries. In the bloodstream, oxygen binds to haemoglobin proteins in erythrocytes (red blood cells) for transport, forming oxyhaemoglobin.

In **oxyhaemoglobin**, the haem B cofactor complexes with an iron ion centre, which is also bound to oxygen and a proximal histidine residue on the polypeptide. A distal His also stabilises the oxygen by hydrogen bonding. There is a resonance structure between Fe^{2+}/O_2 (ferrous, oxygen) and Fe^{3+}/O_2^- (ferric, superoxide). The electronic properties of the iron ion and the porphyrin ring contribute to blood's red colour. The iron ion may also occur purely in the Fe^{3+} oxidation state (methemoglobin), which has reduced oxygen-carrying capacity.

In **deoxyhemoglobin**, the haemoglobin complex dissociates, and the Fe^{2+} ion lies just outside of the conjugated porphyrin plane, broadening the band structure and making deoxygenated blood a slightly darker red than oxygenated blood. This releases free O_2 into the blood plasma, diffusing through the interstitial fluid, across cell membranes and into mitochondria, where it is consumed in aerobic respiration (the electron transport chain, powers ATP synthase (Section 17.1.7)). Oxygen is also used in gluconeogenesis (rev. glycolysis) in the liver to convert lactate (product of anaerobic respiration) to glucose.

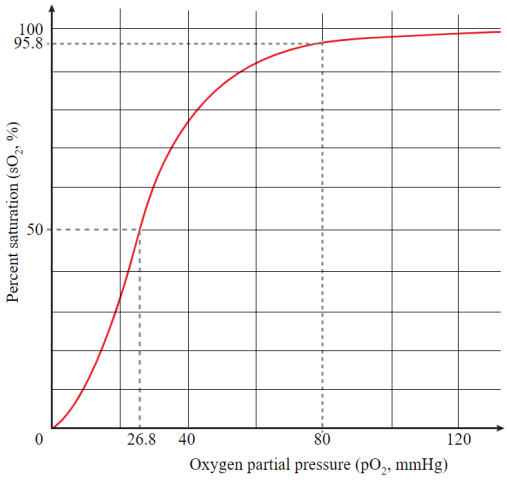
Aerobic respiration generates CO_2, which is dissolved in the bloodstream by the carbonate buffer, with some gaseous CO_2 remaining dissolved in the plasma. It is transported through alveoli to the lungs where it is exhaled.

Blood contents: ~55% plasma (low-density pale yellow fluid containing water, proteins, nutrients and hormones), ~44% red blood cells (erythrocytes: most dense) and ~1% white blood cells (basophils, neutrophils, eosinophils, monocytes, lymphocytes) and platelets (thrombocytes).

Blood pH buffers: the pH of the blood plasma must be 7.40 ± 0.05 to prevent acidemia/alkalemia.

- Carbonate buffer: $CO_2 + H_2O \rightleftharpoons H_2CO_3 \rightleftharpoons HCO_3^- + H^+ \rightleftharpoons CO_3^{2-} + 2\ H^+$
- Phosphate buffer: $H_2PO_4^- \rightleftharpoons HPO_4^{2-} + H^+$

Oxygen-haemoglobin dissociation curve (1 mmHg = 133.322 Pa)



The pH of blood is controlled by the carbonic anhydrase enzyme, which catalyses the reaction

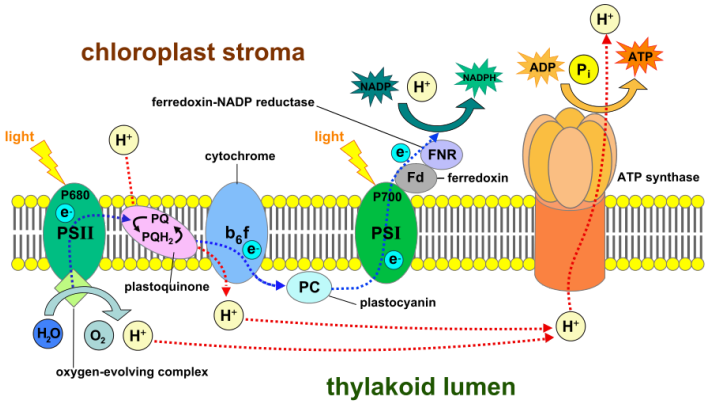
$$CO_2 + H_2O \rightarrow H_2CO_3 \rightarrow HCO_3^- + H^+$$

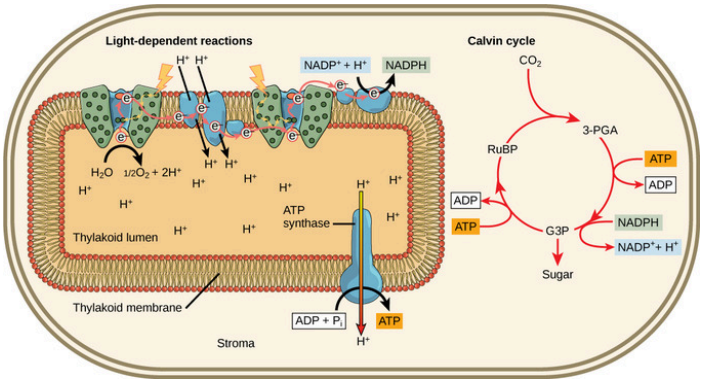
Bohr effect: an increase in blood CO_2 causes a decrease in pH, which induces a conformational change in haemoglobin proteins (relaxed; high-affinity → taut; low-affinity) which makes it more difficult to bind O_2.

Carbon monoxide (CO) severely inhibits O_2 transport by binding ~250 times more efficiently. CO also triggers the Bohr effect, further enhancing CO uptake by other haem sites.

17.1.10. The Calvin Cycle and Photosynthesis in Plants

Photosynthesis occurs in chloroplasts (and cyanobacteria). It consists of two reaction sequences (light dependent set in the thylakoid lumen and light independent set in the stroma) which overall generate glucose ($C_6H_{12}O_6$) and oxygen from CO_2, water, and light.





Light-dependent reactions **Light-independent reactions** (Calvin cycle; shown on right)

1. Photosystem II: photons incident on the fully-conjugated porphyrin ring of a chlorophyll molecule form polarons (excited resonance electrons), which via carotenoids conjugate to two chlorophyll molecules (P680). The electron is passed on to pheophytin, leaving a hole that induces oxidation of water at the oxygen-evolving complex (a Mn_4CaO_5 complex).

2. Electron transport chain: the passed electron is shuttled through a series of other membrane-bound proteins (plastoquinone, cytochrome b_6f, plastocyanin). Photosystem I (P700) re-energises the electrons. The energy of the electron is used to form a proton gradient which converts NADP to NADPH.

3. ATP production: the proton gradient is used to drive ATP synthase, forming ATP in the stroma, used to power the Calvin cycle.

1. CO_2 fixation: 3 CO_2 + 3 RuBP → 3 3-PGA, via RuBisCO.

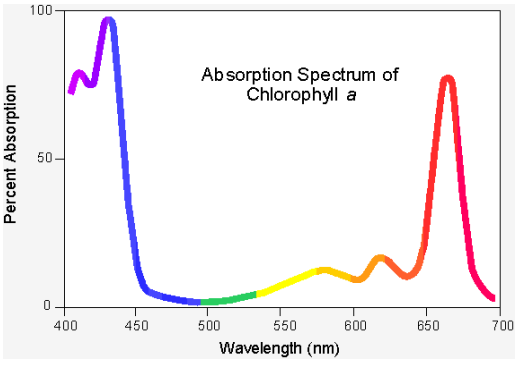
2. Reduction: 6 3-PGA + 6 ATP + 6 NADPH → 6 G3P + 6 ADP + 6 $NADP^+$.

3. Regeneration: 5 G3P + 3 ATP → 3 RuBP + 3 ADP + 2 P_i, via several sugar intermediates.

Glucose formation: G3P is converted to fructose-1,6-bisphosphate via aldolase, which is used to make glucose via gluconeogenesis.

(RuBP: ribulose-1,5-bisphosphate,
3-PGA: 3-phosphoglycolic acid,
G3P: glyceraldehyde 3-phosphate.)

The Calvin cycle is the only mechanism of photosynthesis in C3 plants. In **C4 plants**, CO_2 is fixed into malic acid in the mesophyll as an intermediate before being converted back into CO_2 to enter the Calvin cycle in the bundle sheath, as a defence against photorespiration. In **CAM plants**, the malic acid intermediate step occurs only in the dark. In cyanobacteria, photosynthesis occurs directly on free chlorophyll A molecules in the cytoplasm.



17.1.11. Chemosynthesis, Nitrogen Fixation and Anaerobic Digestion in Prokaryotes

A variety of primitive metabolic processes evolved early in life, as a method of deriving energy from the few consistent chemical sources available on early Earth (see Section 17.5).

Chemosynthesis: conversion of inorganic chemicals to organic chemicals without light

Some extremophile deep-sea microorganisms (chemolithotrophs), including bacteria and archaea, can derive energy from chemicals released from hot hydrothermal vents and cold seeps (seafloor fissures). There are a range of possible pathways, including sulfur, iron pyrites and methane, such as:

$$H_2S + 2\ O_2 \rightarrow SO_4^{2-} + 2\ H^+; \qquad CH_4 + 2\ O_2 \rightarrow CH_3OH + 2\ H_2O$$

The carbon products can then be converted to glucose by carbon fixation processes. Sulfides and sulfates can be reduced to elemental sulfur. Some chemosynthetic organisms are hosted as internal endosymbionts within deep-sea tubeworms, clams and mussels.

Nitrogen Fixation: conversion of nitrogen into ammonia

Some bacteria, including cyanobacteria and green sulfur bacteria, can metabolise nitrogen into ammonia with a high ATP cost via the nitrogenase enzyme complex with overall reaction:

$$\mathbf{N_2} + 16\ ATP + 16\ H_2O + 8\ e^- + 8\ H^+ \rightarrow 2\ \mathbf{NH_3} + H_2 + 16\ ADP + 16\ P_i$$

Nitrogenase is composed of an iron protein (Fe) reductase and an iron-molybdenum (FeMo) protein. The Fe protein binds ATP, pumping electrons and protons into the FeMo cofactor (Fe_7MoS_9C) over a P-cluster between the two to form the 'Janus intermediate'. With further proton addition, this can progressively reduce a molecule of nitrogen (N_2) bound to the FeMo cofactor into hydrazine (N_2H_4) then ammonia (2 NH_3). The precise mechanism varies by circumstances and is not fully understood.

Molecular O_2 inhibits nitrogenase, so cyanobacteria must isolate photosynthesis from nitrogen fixation, either by only nitrogen-fixing at night, or in communities of cyanobacteria, by nitrogen-fixing only in specialised cells (heterocysts) that are impermeable to O_2 and that degrade photosystem II prevent local oxygen production.

The resulting ammonia is used in various reactions, such as amino acid synthesis, nucleotide synthesis, glutamate metabolism and the urea cycle. Nitrogen-fixing bacteria in soils are a key part of the nitrogen cycle in plants (Section 19.2.7), where they can release ammonium into the soil for uptake as nitrates by plants.

Anaerobic Digestion: used by microorganism communities to break down organic material.

1. **Hydrolysis:** insoluble organic polymers are hydrolysed to monomers e.g. carbohydrates → sugars.
2. **Acidogenesis:** bacteria convert sugars/amino acids into CO_2, H_2, NH_3 and organic acids.
3. **Acetogenesis:** bacteria convert organic acids into acetic acid. This process can also occur starting with CO_2 and H_2 (the Wood-Ljungdahl pathway).
4. **Methanogenesis:** methanogenic archaea release energy from carbon sources, via reactions such as $CO_2 + 4\ H_2 \rightarrow CH_4 + 2\ H_2O$ (via the Wolfe cycle) and $CH_3COOH \rightarrow CH_4 + CO_2$. This forms 'biogas'.

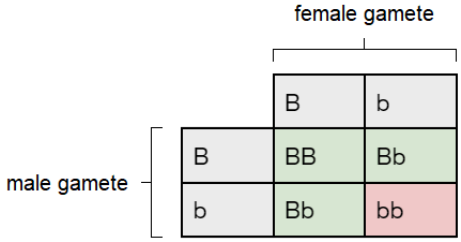
17.2. Genetics and Evolutionary Developmental Biology

17.2.1. Mendelian Genetics and Inheritance

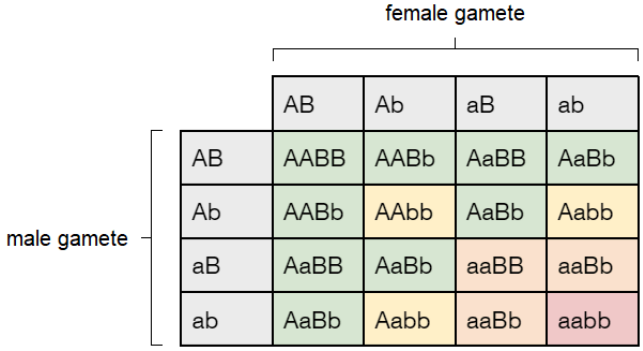
Mendelian genetics provided a biological mechanism for Darwin's theory of evolution by natural selection, with their combination forming the **'modern synthesis'** (**'neo-Darwinism'**). Evolution is therefore defined as the variation of allele frequency in a given population. By studying pea plants, in 1866, Mendel proposed that traits are controlled by alleles (copies of genes), described by:

- **Law of segregation:** one of the parent's alleles is distributed **randomly** in each gamete.
- **Law of independent assortment:** each allele is distributed **independently** in each gamete.

Alleles: different sequences of a gene at a given locus on a chromosome. A pair of chromosomes may have the same (homozygous) or different (heterozygous) alleles of a given gene: in the simplified model, the genotype is written e.g. "BB / Bb / bb". The phenotype (presenting trait) for the dominant "B" is expressed unless both gametes provide the recessive "b". Heterozygous alleles are 'carriers' of the recessive trait despite not showing it in their phenotype.



Monohybrid Cross: one allele pair (3:1)



Dihybrid Cross: two allele pairs (9:3:3:1)

Most traits are not controlled by a single gene. Simplifications are often made to allow for probabilistic analysis of Punnett squares, such as from inherited disorders (e.g. polydactyly, cystic fibrosis, alopecia). The sex chromosomes X and Y can be treated as codominant 'alleles'.

- **Gametes (sex cells)** are **haploid** (n chromosomes; n = 23 in humans = 22 autosomes + X or Y)
- **Somatic cells** are **diploid** ($2n$ chromosomes; $2n$ = 46 in humans = 22 autosome pairs + XX or XY)

Sexual reproduction involves the joining (fusion) of male and female gametes (sperm and egg cells in animals; pollen and egg cells in flowering plants). There is mixing of genetic information (gamete formation uses meiosis: spermatocyte → 4 spermatozoa, oocyte → 1 ovum + 3 nonfunctional eggs, then gamete fusion at fertilisation to form the zygote) which leads to variety in the offspring.

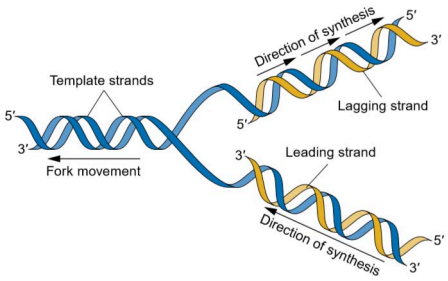
Asexual reproduction involves only one parent and no fusion of gametes. There is no mixing of genetic information (mitosis), producing genetically identical offspring (clones).

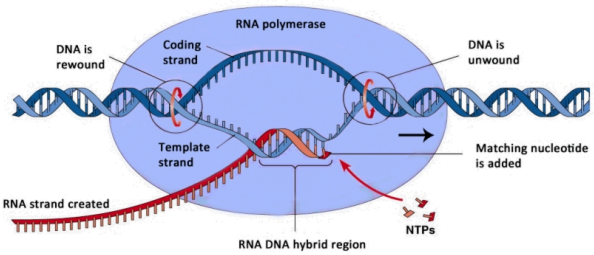
Both types may also occur, e.g. malarial parasites (asexually in human host, sexually in mosquito), many fungi (asexually by spores, sexually to give variation), many plants (sexually by seeds, asexually by runners such as strawberry plants, or bulb division such as daffodils).

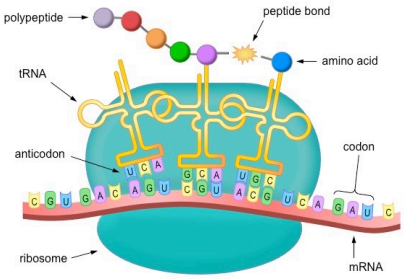
17.2.2. DNA: Replication, Transcription and Translation

DNA semi-conservative replication occurs in the S phase of the cell cycle, in the nucleus (in eukaryotes).
DNA transcription and RNA translation occur in the G_1 and G_2 phases.
For the structures of DNA and RNA, see Section 16.5.7.

The 'central dogma of molecular biology': **DNA → RNA → Protein.** (exception: reverse transcription)







DNA Replication

Initiation: helicase unwinds the helix by disrupting the hydrogen bonds. This creates a replication fork.
Topoisomerase relieves torsional strain ahead of the replication fork by breaking, untwisting, and reconnecting the DNA. SSB proteins stabilise the separated strands.

Leading Replication: primase anneals a short RNA sequence (primer) to each half of the DNA. DNA can only be synthesised in the 5' → 3' direction, so only one strand can anneal free dNTPs (deoxynucleotide triphosphates) continuously (the leading strand), using DNA polymerase III in prokaryotes (different in eukaryotes).

Lagging Replication: short 'Okazaki fragments' on the 5' to 3' strand (lagging strand) are formed, one at a time, requiring a primer for each. DNA polymerase I (in prokaryotes) replaces the RNA primers with DNA nucleotides In eukaryotes, several enzymes fulfil this role together. Ligase joins fragments of DNA together on the lagging strand.

DNA Transcription

Initiation: RNA polymerase binds with various transcription factors to a promoter sequence (e.g. TATA box in eukaryotes) of DNA near the beginning of a gene. The two strands (coding, template) of DNA are separated, forming a replication bubble.

Elongation: RNA polymerase adds complementary ribonucleotides of the template strand, building in the RNA 5' → 3' direction, with the synthesised strand of RNA being released continuously.

Termination: a termination sequence (AAUAAA) is recognised near the 3' end of the pre-mRNA once the full gene has been transcribed. The DNA is returned to its original state.

Post-Transcriptional Modification: A 5' cap (single modified G nucleotide) and a 3' tail (poly-A sequence) are added to each end of the pre-mRNA. A spliceosome enzyme cuts out non-coding intron sections. The remaining coding exon sections are combined. The 5' cap is recognised by a nuclear pore complex, allowing the mRNA to leave the nucleus.

RNA Translation

The small ribosomal subunit (**rRNA**) binds to the start of the mRNA.

Aminoacyl-tRNA synthetases (ARSs) attach amino acids to one end of tRNA molecules. A tRNA molecule carrying methionine binds to the start codon of the mRNA.

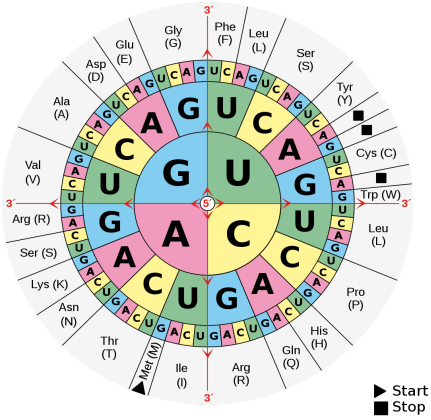
The large ribosomal subunit binds to form the complete initiation complex.

During the **elongation stage**, new tRNA molecules corresponding to the next codon triplet enter the ribosome, binding, and resulting in the carried amino acids becoming linked by a peptide bond, forming a polypeptide.

When a stop codon is encountered, there exists no corresponding tRNA carrier, so **termination** occurs. The complex dissociates and the amino acid sequence detaches.

The polypeptide enters the **rough endoplasmic reticulum** for folding and the **Golgi apparatus** for further modification (e.g. to form a proteoglycan), assisted by chaperone proteins, to form the final protein.

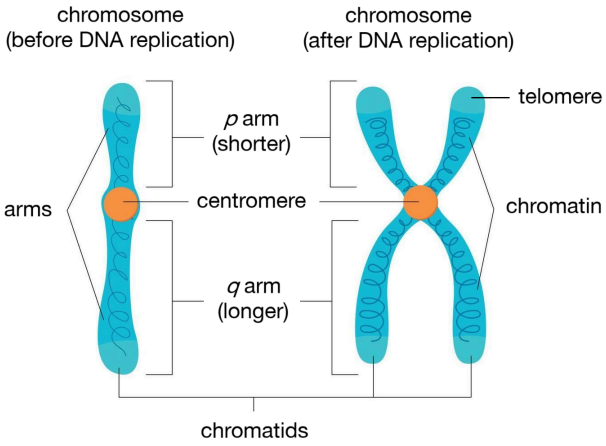
17.2.3. RNA to Amino Acid Inverse Codon Mappings



The genetic code conversion process is highly conserved throughout evolution, in most organisms.

In a coding section of a gene, each consecutive triplet of mRNA codons (directed from 5' to 3') forms a single amino acid, as determined from the chart. The **start codon** is methionine (in eukaryotes) or formyl methionine (in prokaryotes). The **stop codons** are sometimes called amber (UAG), opal/umber (UGA), ochre (UAA). In some cases, these serve other functions, and can incorporate special amino acids instead e.g. UGA → selenocysteine.

17.2.4. Nuclear DNA (Chromosomes)



In eukaryotes, the main component of the genome is the chromosomes, which are found inside the nucleus of most cells. The nucleus is surrounded by the nuclear envelope (with an inner and outer membrane) which has **nuclear pores** to protect the chromosomes from chemical damage while allowing controlled access across the membrane. Diploid cells have 2*n* chromosomes, with chromosomes coming in pairs (one maternal, one paternal). The two chromosomes are homologous (contain the same genes) but are not identical due to natural variation in the population. During the interphase of a cell cycle, each chromosome duplicates into two identical sister chromatids, conjoined at the centromere.

Chromosomal Origins of Mendelian Genetics:

- The **'law of segregation'** is accounted for by the fact that during meiosis I, the maternal and paternal chromosomes are separated into each new cell, so only one allele is carried forward;
- The **'law of independent assortment'** is accounted for by the fact that different traits arise from genes on different chromosomes, which are arranged randomly when meiosis occurs.

Centromere: the site where the two sister chromatids bind together. The centromere DNA sequence contains groups of long repeating sequences (such as α-satellite DNA; alphoid repeats). The identity of the centromere sequence varies between species.

Telomeres: the telomere sequence 'TTAGGG' is repeated several hundreds of times on the ends of the arms of each chromosome. The RNA primers for the last Okazaki fragment at the end of a DNA strand cannot be removed during DNA replication, resulting in slight chromosome shortening at each cell cycle. In gametes, telomerase is responsible for gradually regenerating the telomere sequences, but it is marginally too slow to keep up with the rate of cell replication, leading to gradual information loss (telomere attrition; replicative senescence), which is the main biomarker for cell ageing. Telomerase is activated in cancerous somatic cells, contributing to their immortality.

17.2.5. Non-Nuclear DNA and Ribosomal RNA

Mitochondria and chloroplasts inside eukaryotic cells originate as endosymbionts and share many similarities with prokaryotes.

Mitochondrial DNA (mtDNA)

- **Structure:** animal mtDNA is a singular circular chromosome, with a 'heavy' (H) strand and a 'light' (L) strand, no introns, of size 11-28 kbp. Arrangements differ in other kingdoms.
- **Transcription:** mitochondria undergo their own cell cycle, independently of the host cell, using binary fission for replication. Many components required for mtDNA replication, including subunits of DNA polymerase and helicase, must be sourced from the host cell as they are encoded in nuclear DNA. mtDNA uses single gene promoters, not operons.
- **Translation:** The mt-mRNA is translated in mitochondrial ribosomes. Vertebrate mtDNA uses different start/stop codons than nuclear DNA, and also has some differences in the codon table.
- **Mitoepigenetics:** mtDNA methylation is observed as a mode of mitochondrial gene regulation.
- **Human mtDNA genes:** human mtDNA is 16,569 bp long, with genes primarily responsible for respiration, including for subunits of ATP synthase, subunits for cytochrome c oxidase, NADH dehydrogenase, humanin micropeptide, as well as 22 tRNAs and 2 rRNAs. Transpositions from mtDNA into nuclear DNA (NUMTs, nuclear-mitochondrial segments) happen often. Mutations in either the main nuclear DNA or mtDNA can cause a variety of mitochondrial diseases, limiting the energy availability from respiration for the host cell/organism.
- **Human Haplogroups:** mtDNA is maternally conserved (matrilineal), since the mitochondria-containing midpieces of sperm fall away when they fertilise the egg. It is subject to a slow and steady mutation rate (the human mtDNA molecular clock). Variations in mtDNA can be grouped by single nucleotide polymorphisms (SNPs) and microsatellite DNAs (STRs) and are indicative of broad human communities, useful for tracing migration patterns in recent evolutionary history (paleogenomics). The patrilineal equivalent is obtained by similar analysis of the Y-chromosomal DNA (Y-DNA).

Chloroplast DNA (cpDNA)

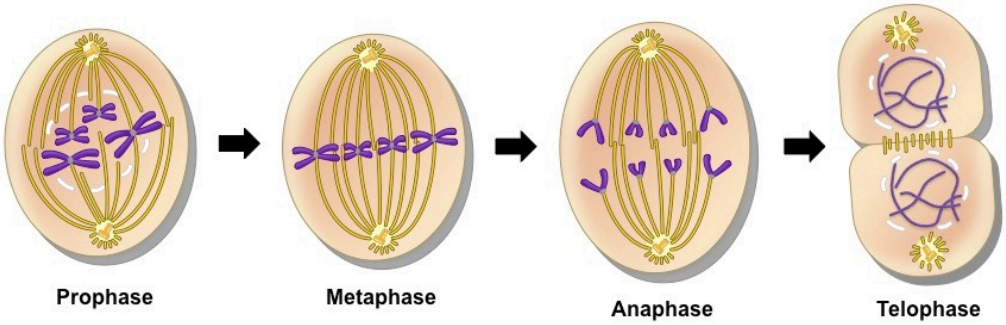
- Plant cpDNA is a singular circular chromosome of size 120-170 kbp.
- Some cpDNA genes have introns and are typically organised into operons.
- cpDNA genes are primarily responsible for photosynthesis, including RuBisCo, photosystem I and II, and ATP synthase, as well as components for replication, including subunits of RNA polymerase, ~30 tRNAs and 4 rRNAs.

Ribosomal RNA (rRNA)

- Ribosome synthesis: nuclear genes for the rRNA precursors are transcribed and ligated to form the small and large ribosomal subunits in the nucleolus.
- Small and large subunits bind with ribosomal proteins to form the ribosome with stem-loop motifs.
- Most of the rRNA genes are highly conserved across all three domains of life.

17.2.6. Mitosis (Somatic Cell Division)

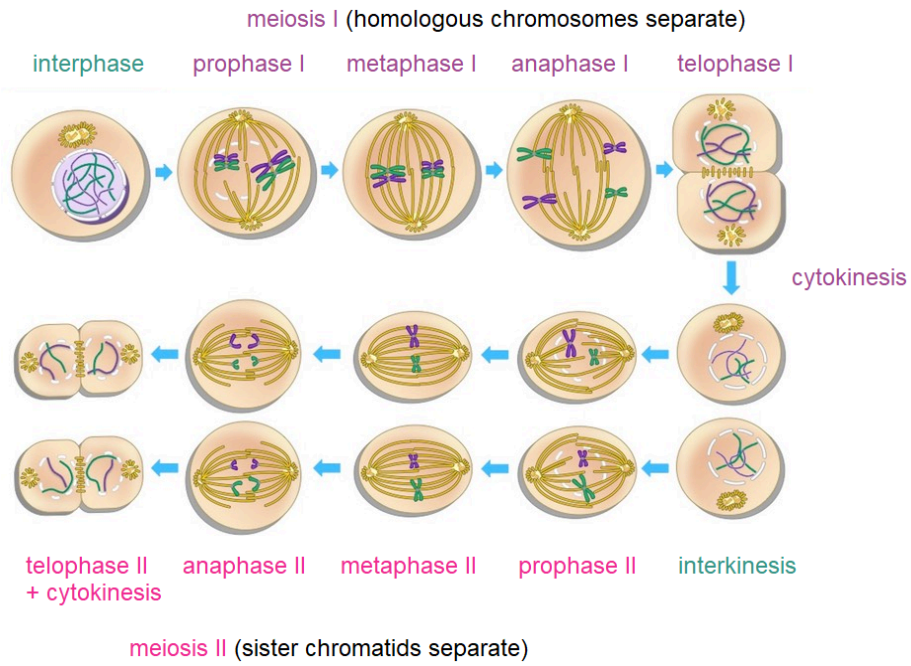
Mitosis involves the copying of genetic material to produce identical cells (barring mutations).



1) **Prophase:** chromosomes assemble, centrosome duplicates, spindle fibres form. The microtubules between the centrosomes extend in length.
2) **Prometaphase:** nucleus breaks apart, kinetochores attach at the centromeres of the pairs.
3) **Metaphase:** chromosomes aligned in the centre, the two centrosomes are at the poles of the cell.
4) **Anaphase:** separase enzyme cleaves the cohesins of the centrioles, then the kinetochores pull the chromatids apart by motor proteins, cell elongates.
5) **Telophase:** two new nuclei form around each set of chromosomes.
6) **Cytokinesis:** the cytoplasm will continue to divide, pinching off in two.

17.2.7. Meiosis (Gametic Division)

Meiosis involves mixing of genetic material in a diploid cell to produce four identical haploid cells.



1) Meiosis I, interphase: DNA replication to form a diploid cell.

2) Meiosis I, prophase I: pairs of chromosomes may exchange fragments of themselves by **'crossing over'**.

3) The remainder of meiosis I is similar to mitosis, with each chromosome being pulled randomly to a pole.

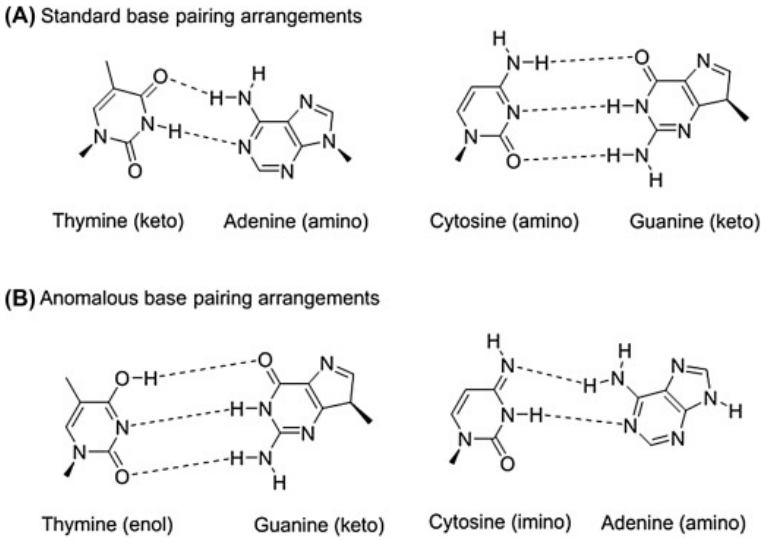
4) Meiosis II: functionally identical to mitosis in each cell.

Anisogamy: the difference in size and form between the two sexes of gametes.

17.2.8. Genetic Mutations

For the structure of DNA and nucleobases, see Section 16.5.9.

Sources of DNA Mutations



- **Tautomerism** of nucleobases (shown right), where ketone/amine groups on bases are in chemical equilibrium with their rarer enol/imine groups respectively (Section 16.1.4), which have different hydrogen bonding environments, prompting DNA polymerase to insert the wrong base. Fixed by other DNA polymerases (proofreading / DNA repair).
- **Deamination** of 5-methylcytosine (epigenetically marked C) to thymine, and of cytosine to uracil. These hydrolysis reactions occur spontaneously. Fixed by nucleases and glycosylases (base excision repair).
- **High energy radiation (e.g. UVB, X-rays)** can cause bonding between adjacent T or C bases on the same strand (pyrimidine dimer) to form a cyclobutane ring between them, causing a kink in the DNA helix. Fixed by photolyases (nucleotide excision repair).
- **Chemical mutagens** such as oxidising agents can oxidise nucleobases, changing their structure and identity of the corresponding base. Fixed by nucleases and glycosylases (base excision repair).

DNA Repair: DNA damage can be fixed, but occasionally a mistake will occur (mutation).

- **3' (to 5') exonuclease activity (proofreading):** nucleotides at 3'-OH termini can be removed, allowing for in-place correction during replication before continuing. Exhibited by DNA polymerase I and III.
- **5' (to 3') exonuclease activity (DNA repair):** nucleotides at 5'-P termini can be removed, allowing for removal of existing DNA nucleotides. Exhibited by DNA polymerase I.
- **Nucleotide excision repair (NER):** replacement of a short single strand of nucleotides.
- **Base excision repair (BER):** replacement of oxidised, alkylated or deaminated bases.
- **Double-stranded break repair (DSBR):** includes homologous recombination (copying from undamaged sister chromatid) and non-homologous end-joining (NHEJ, deliberate indel formation).

Point mutation: a change in a single nucleotide in the template strand of DNA (a missense mutation). Can be benign if no change in the resulting amino acid, or if in an intron section. Point mutations can be written in terms of the amino acids: e.g. Glu346Lys (glutamic acid at protein position 346 becomes lysine). If the new mRNA triplet is a stop codon, the resulting incomplete protein is usually non-functional (a nonsense mutation).

RNA mutations can also occur (post-transcriptional modification of mRNA). The main mechanism is by adenosine-to-inosine RNA editing, in which an adenine nucleoside (adenosine) is deaminated by an ADAR enzyme to form inosine, which mimics a guanine nucleobase instead (a point mutation).

This mode of mutation does not affect the genome, but can increase the diversity of the 'transcriptome'. It occurs at high rates in cephalopods (a class of molluscs: octopuses, squid, cuttlefish), potentially as a method of rapid environmental adaptation.

Other types of mutations: often having larger influence than point mutations

- **Frame shift mutation:** single insertions or deletions (indels) in DNA, resulting in a change in the reading frame (large number of mutations). This can significantly change the tertiary structure and size of the encoded protein, making the protein non-functional or of alternative function.
- **Transposons:** movement of a section of DNA. Can lead to **gene/chromosomal duplication**, **inversion** or **translocation**, which allow for significant diversification of the genome (neofunctionalisation). If the transposon contained *cis*-regulatory elements e.g. a promoter region, an adjacent non-coding section may become coding (*de novo* genes).

If a mutation occurs in gametes or their stem cell progenitors, it can be passed to the next generation, as the alleles have the chance to form the offspring's genome at fertilisation.

Gene flow and horizontal gene transfer: exchange of genetic material between cells or populations e.g. plasmid conjugation: bacterial plasmids are transferred from one bacterium to another, conferring e.g. antibiotic resistance. This forms one of the mechanisms of evolution.

17.2.9. Gene Regulation in Eukaryotes (Epigenetics)

The majority of the genetic information in a eukaryotic cell is found in the chromosomes, in the nucleus of eukaryotic cells. Histones are octamer protein particles around which DNA coils to form nucleosomes, which in turn supercoils to form chromatin which comprises the chromosomes. DNA may be accessible (euchromatin) or inaccessible (heterochromatin) to transcription factor enzymes, altering gene activity.

Mechanisms of Regulation

- **DNA methylation:** DNA methyltransferases (DNMT) add a CH_3 group to cytosine bases to form 5-methylcytosine. Results in **gene silencing** due to inhibition of transcription proteins and recruitment of methyl-cytosine-binding proteins (MeCPs) and is associated with increased cell specialisation. Demethylation results in a reversal towards the pluripotent stem cell state. During DNA replication, methylation can be produced (by e.g. DNMT1) on the new strand at CpG dinucleotides (CG islands) to produce methylated bases at adjacent positions on each strand. 5-methylcytosine can spontaneously deaminate to thymine, causing a point mutation (C → T).

- **Histone acetylation and deacetylation:** histone acetyltransferases (HAT) adds and histone deacetylases (HDAC) remove $COCH_3$ groups to and from N-terminal tails of core histone subunit proteins. Acetylation results in **gene expression** due to the loose packing of histones. This can also induce histone remodelling (euchromatin formation) due to the conformational changes, which can further increase DNA availability. This can also be caused by ATP-driven complexes which eject nucleosomes from the histones.

- **Other histone modifications:** histone methyltransferases add a CH_3 group to lysine side chains (**histone methylation**); protein kinases add a phosphate group to serine side chains (**histone phosphorylation**); core histone subunits can bind to ubiquitin or SUMO proteins, which in turn bind lysine side chains (**histone ubiquitination** / **SUMOylation**). Depending on the site targeted, these can result in **either silencing or expression** by inducing chromatin remodelling.

- **Non-Coding RNA:** micro-RNA (miRNA) and short-interfering-RNA (siRNA) can bind or cleave newly-transcribed mRNA to prevent its translation, resulting in **gene silencing**.

- **Positive Gene Regulation:** a signalling molecule (e.g. cAMP) can bind to an activator protein, which can bind to DNA near the promoter sequence, enhancing the affinity of RNA polymerase for the promoter, resulting in **gene expression.** This can also occur in operons (Section 17.2.5).

- ***Cis*-Regulatory Elements (Promoters, Enhancers, Silencers, Operators):** transcription factors can bind to specific intron (non-coding) regions of DNA, which then recruit other activator proteins associated with the promoter region (e.g. TATA box) of a given gene (a protein-DNA interaction). This results in the DNA in a 'topologically associated domain' bending and looping around to bring the two regions close together, which recruits RNA polymerase to the developing transcription-initiation complex.

Example: Human lactose tolerance. The *LCT* gene codes for lactase (breaks lactose into glucose and galactose). A regulatory region on the intron of the *MCM6* gene, which typically codes for a helicase, functions as an enhancer for an *LCT* promoter. Recently in human evolution (~10 kYA), for Western populations, a mutation in *MCM6* has increased its regulatory activity further, enabling 'lactase persistence', in which *LCT* is expressed at high levels throughout adult life rather than only in the first few years of life, allowing lactose to be metabolised, naturally selected for as these societies farmed dairy. Those without the mutation (~75% of humans worldwide) remain 'lactose intolerant': the absence of lactase in the small intestine leaves the lactose available for the bacteria in the large intestine, which ferment it, leading to fatty acid and gas production, causing symptoms.

Epigenetic Inheritance

In some organisms, some regulatory signals can be inherited from one generation to another, if they occur in gametes (epimutations). This can lead to a change in phenotype in the offspring.

Epigenetic inheritance is not 'Lamarckism', which was the pre-Darwinian idea that organisms inherit traits based on the parents' behaviours during their life.

Intergenerational epigenetic inheritance: epigenetic marks are inherited up to one generation. This is common for most epigenetic markers expressed in gametes at fertilisation.

Transgenerational epigenetic inheritance: epigenetic marks are inherited through multiple generations, potentially forever. TEI is well studied in plants, occurring alongside polyploidy to provide significant capacity for rapid adaptation, but may also happen in animals to a lesser degree.

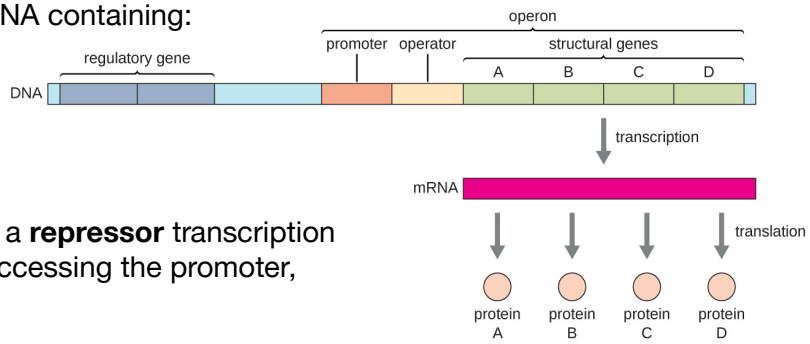
There is disputed evidence of TEI in humans, from 'foetal programming' after environmental stressors induced by famines, in which babies born to malnourished mothers in the famine exhibit phenotypic changes.

17.2.10. Gene Regulation in Prokaryotes (Operons)

Modulation of plasmid gene expression is based on negative feedback from the surrounding chemical environment: if there is a high concentration of a particular metabolite in the cell, this will act as an inhibitor for its own production, helping to reduce its level and prevent accumulation of toxins.

An **operon** is a sequence of prokaryotic DNA containing:

1. **Genes** coding for proteins.
2. **Promoter** sequence, signalling the start region of a gene to be transcribed.
3. **Operator** sequence, which can bind to a **repressor** transcription factor to block RNA polymerase from accessing the promoter, effectively turning the gene off.



In operons synthesising amino acids Trp, His, Phe, Thr only, a **leader** sequence is found between operator and first gene, which encodes for a **leader peptide** and an **attenuator**. The leader peptide contains several of the target amino acid residues. The attenuator mRNA has self-complementary regions which form hairpins. When RNA polymerase transcribes the operon, ribosomes can begin translating the growing mRNA. Negative feedback (**attenuation**) is established as the ribosome can stall if the target amino acid tRNAs are absent. If present, translation occurs quickly and the ribosome detaches, forming the terminator which blocks operon transcription.

- The repressor protein is synthesised by a repressor gene elsewhere (typically upstream).
- The repressor may be either **active** or **inactive** when it is newly transcribed - presence of a molecule can switch the repressor's state (**metabolite**: to active, **inducer**: to inactive), switching the operator's state, turning the gene on or off.
- If the gene codes for the metabolite, then a negative (inhibitory) feedback loop is established.
- The genes on the operon are all adjacent to a single promoter, so are all transcribed at the same time.
- Operons can be monocistronic (1 gene), bicistronic (2 genes), tricistronic (3 genes), polycistronic (more).
- Operons are rare in eukaryotes. Some are known in model organisms *D. melanogaster* and *C. elegans*.

Tryptophan synthesis: the *trp* operon, found in *E. coli*. Negative feedback regulation of tryptophan (Trp).

- *trp* has five genes: *trpE, trpD, trpC, trpB, trpA*.
- When Trp is **absent**, the *trp* repressor (encoded by *trpR*) is inactive (operon **on**).
- When Trp is **present**, Trp acts as a metabolite for the *trp* repressor, making it active (operon **off**).
- Attenuation mechanism: when Trp is **absent**, leader translation is slow, antiterminator hairpin forms, transcription continues (operon **on**). When Trp is **present**, leader translation is fast, terminator hairpin forms, transcription ends (operon **off**).

Lactose metabolism: the *lac* operon, found in *E. coli*. Digest lactose (Lac) when glucose (Glu) is absent.

- *lac* has three genes: *lacZ* (β-galactosidase), *lacY* (permease), *lacA* (acetyltransferase).
- When Lac is **absent**, the *lac* repressor (encoded by *lacR*) is active (operon **off**).
- When Lac is **present**, (allo)lactose is the metabolite for the *lac* repressor, making it inactive (operon **on**).
- Positive gene regulation mechanism: when Glu is **absent**, cAMP is present, attaches to CAP, allowing it to bind the upstream CAP site of DNA. Bound CAP helps RNA polymerase bind to the promoter (operon **strongly on**). When Glu is **present**, cAMP is absent, CAP is unbound and cannot help RNA polymerase (operon is **weakly on**).

17.2.11. Stochastic Modelling of Genetic Regulatory Networks (Systems Biology)

Interactions within genetic circuits yield time-varying concentrations of molecular species (mRNAs, proteins, signalling molecules etc).

Master Equation: models the joint probability density function of a given state

(n_i: number of molecules of species i (mRNAs, proteins etc), k_j: rate constant for reaction j, $p(n_1, \ldots, n_N)$: joint PDF of number of molecules.) p varies over time according to a Poissonian discrete state space Markov chain model (Section 4.4.8): the current state $\mathbf{x} = [n_1, \ldots, n_N]^\mathsf{T}$ can evolve over time due to reactions changing each n_i.

Rate equations: e.g. if $\mathrm{A} \rightarrow \mathrm{A} + \mathrm{B}$ (rate k), then $\mathrm{d}p/\mathrm{d}t = -kn_\mathrm{A}p(n_\mathrm{A}, n_\mathrm{B}) + kn_\mathrm{A}p(n_\mathrm{A}, n_\mathrm{B} - 1)$ (first term: transition from $[n_\mathrm{A}, n_\mathrm{B}]^\mathsf{T} \rightarrow [n_\mathrm{A}, n_\mathrm{B} + 1]^\mathsf{T}$, second term: transition from $[n_\mathrm{A}, n_\mathrm{B} - 1]^\mathsf{T} \rightarrow [n_\mathrm{A}, n_\mathrm{B}]^\mathsf{T}$.

Moment generating function (MGF, Section 5.2.8) of p: $F(z_1, \ldots, z_N) = \sum_{n_i} (z_1^{n_1} \ldots z_N^{n_N})\, p(n_1, \ldots, n_N)$.

Reaction type	Reaction	Time Evolution of MGF, $\mathrm{d}F/\mathrm{d}t$
Synthesis from a template (e.g. transcription, translation)	$A \xrightarrow{k} A + B$	$\dot{F} = kz_A(z_B - 1)\dfrac{\partial F}{\partial z_A}$
Degradation (e.g. removal of mRNA, protein)	$B \xrightarrow{\gamma} 0$	$\dot{F} = -\gamma(z_B - 1)\dfrac{\partial F}{\partial z_B}$
Forward irreversible reaction ($n_\mathrm{A} + n_\mathrm{B} = n$, a constant)	$A \xrightarrow{k} B$	$\dot{F} = kn(z_B - 1)F - kz_B(z_B - 1)\dfrac{\partial F}{\partial z_B}$
Dynamic equilibrium (e.g. enzyme or receptor process)	$A \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} B$	$\dot{F} = k_1(z_B - z_A)\dfrac{\partial F}{\partial z_A} + k_{-1}(z_A - z_B)\dfrac{\partial F}{\partial z_B}$

Combinations of these terms yield more complete models of genetic cascades. Statistical moments (mean, variance) can be easily found from the MGF at **steady state:** $\mathrm{d}F/\mathrm{d}t = 0$. Under the condition $\{z_1, \ldots, z_N\} = \{1, \ldots, 1\}$, at which $F(1, \ldots, 1) = 1$, important statistics are:

$$\mathrm{E}[n_i] = \frac{\partial F}{\partial z_i}, \quad \mathrm{E}[n_i(n_i - 1)] = \frac{\partial^2 F}{\partial z_i^{\,2}}, \quad \mathrm{E}[n_i n_j] = \frac{\partial^2 F}{\partial z_i \partial z_j}\ (i \neq j), \quad \mathrm{Var}[n_i] = \frac{\partial^2 F}{\partial z_i^{\,2}} - \left(\frac{\partial F}{\partial z_i}\right)^2 + \frac{\partial F}{\partial z_i}.$$

For systems of interactions, the equations are typically nonlinear, and a matrix formulation can be written, linearising about the steady state i.e. $\mathrm{d}\mathbf{x}/\mathrm{d}t = (\mathbf{K} - \mathbf{\Gamma})\mathbf{x}$ (can be extended with control inputs $\mathbf{u}$ and outputs $\mathbf{y}$, see Section 5.4.11), with $\mathbf{E}[\mathbf{x}] = \nabla F(\mathbf{1})$ etc. The 'Langevin approach' is another method, starting with the deterministic rate equations and adding a stochastic noise term, solving to find the resulting output disturbance by Fourier transforms (Section 3.6.4).

17.2.12. Composition of Genomes

Only a small proportion of the genome is for protein-coding genes. Organisms with larger genome sizes tend to have smaller protein-coding fractions:

Organism	Genome size (bp)	Proportion protein coding
Humans (*Homo sapiens*)	3.2 billion	1.5%
Fruit fly (*Drosophila melanogaster*)	180 million	20%
Nematode (*Caenorhabditis elegans*)	100 million	25%
Yeast (*Saccharomyces cerevisiae*)	12 million	70%

The non-coding section of the genome contains mobile genetic elements, introns, repetitive/duplicate elements and entirely non-functional DNA ('junk DNA').

Mobile genetic elements: variable loci in the genome; forms ~50% of the human genome.

- **Transposons:**
    - Retrotransposons (Class I): 'copy and paste': transcribed to mRNA, then reverse transcribed back into DNA and re-integrated into the genome at another position, using the encoded reverse transcriptase (RT) and integrase enzymes.
        - LTR retrotransposons: commonly encodes the genes *gag* and *pol*. The *gag* protein assembles a virus-like particle (VLP) to protect the mRNA from degradation. The *pol* gene has four proteins: protease, reverse transcriptase, RNase H and integrase.
        - Endogenous retroviruses (ERVs): benign retroviruses, permanently incorporated.
        - Short/long interspersed retrotransposable elements (SINEs and LINEs): common are the *Alu* and SVA SINEs and the LINE1 gene, highly conserved in primate genomes.
    - DNA transposons (Class II): 'cut and paste' by transposase enzymes.
- **Viral agents:**
    - Viruses and viroids: free viral DNA or RNA transcribed within the cell during a viral infection.
    - Endogenous viral elements: proviral sequences in the genome during a retroviral infection.
- **Plasmids:** circular chromosomes found in prokaryotes and some fungi (e.g. yeast).

17.2.13. Comparative Genomics and Bioinformatics

Metrics of Distance: compute a difference score between sequences

- **Hamming distance:** number of positions at which the corresponding symbols in two strings of **equal length** are different.
- **Edit distance (Levenshtein distance):** minimum number of single-character edits (insertions, deletions, or substitutions) required to change one sequence into the other.

Smith-Waterman Local Alignment Algorithm: compute a similarity score between sequences

- Initialise a $(m+1) \times (n+1)$ matrix of zeros. (m, n: sequence lengths)
- Define the scoring scheme e.g. match: +1, mismatch: -1, gap penalty: -2.
- For each cell (i, j) in the matrix:
    - Calculate the score for extending the alignment ending here by considering **1)** extending from the diagonal (match/mismatch), **2)** extending from the left (gap in sequence 1), **3)** extending from above (gap in sequence 2).
    - Let this cell in the matrix be the maximum score among these (or zero: local alignment).
- Traceback: starting with the maximum cell in the matrix, follow the decreasing values to reach the cell with score zero. This forms the optimal alignment path.

This is an $O(mn)$ algorithm (quadratic), so heuristics are necessary for practical applications, where seed strings are used to find diagonal banded matches. FASTA and BLAST are popular heuristic-assisted algorithms for efficient comparison. Repeats in vertebrate genomes can make alignments easier by providing conserved anchors. However, when the repeats are variable (e.g. transposons, microsatellites/STRs), alignment is complicated and gaps (indels) must be accounted for suitably (such as in whole-genome comparisons).

Phylogenetic Analysis: infer evolutionary relationships among taxa from genetic similarity.

While evolution is not necessarily parsimonious (due to e.g. variable mutation rates, selection pressures, genetic drift, gene flow, horizontal gene transfer), it is often a strong indicator of relatedness, especially at and above the taxonomic genus level.

- **Fitch algorithm:** minimise the number of character state changes.
- **Sankoff algorithm:** minimise the costs associated with different types of evolutionary events.

Clustering methods can be used to generate a phylogenetic tree from the distance matrix, by iteratively starting with a star topology (all equally related) and merging nearest neighbours.

17.2.14. Experimental Analysis of Gene Expression

***In vitro* microarray analysis techniques** are used in interpreting the data generated from experiments on DNA (gene chip), RNA, and protein microarrays. After isolation:

1. Solvent separation: move mRNA into aqueous phase, protein and DNA into phenol phase.
2. Reverse transcription (RT) is used to convert mRNAs into cDNA.
3. Coupling: fluorescent cyanine (Cy) dyes used to label the cDNA targets (single channel). For two-colour arrays, Cy3 and Cy5 are used to distinguish control and test groups.
4. Hybridisation: the labelled cDNAs bind with complementary probes in the array wells.
5. Scanning: fluorescence microscopy is used to record which array wells responded.
6. Analysis: plotting to calculate RNA abundance across the array, normalising the intensity against control spots.

Statistical rank products analysis: used to identify differentially expressed genes. The significance is assessed by bootstrapping (comparing with a null RP population from a large random set of permutations) to estimate the false discovery rate. Hierarchical clustering groups genes on their similarity of expression pattern (either by gene or by sample) by constructing a dendrogram (tree structure) and applying clustering algorithms.

Serial analysis of gene expression (SAGE): a high-throughput assay for quantifying gene expression, used to analyse whole transcriptomes with SAGE libraries of tags. Robust variants include RL-SAGE and SuperSAGE. It has been expanded with **massively parallel signature sequencing** (MPSS), which forms part of modern **next-generation sequencing** (NGS), a much higher throughput method than the traditional Sanger sequencing and shotgun assembly methods, making rapid whole genome accession viable.

Chromatin immunoprecipitation (ChIP): used to investigate protein-DNA interactions. Cells are crosslinked with formaldehyde to preserve protein-DNA bonds. After lysis, and shearing of the chromatin, an antibody for the protein of interest is added to precipitate selected protein-DNA complexes, isolated with protein A/G beads (binds IgA/IgG) or magnetic beads. The DNA can then be sequenced (ChIP-Seq: ‘ChIP on chip’).

***In vivo* morpholino knockdown tests** can be done with morpholino oligomers (synthetic RNA analogs) which can be designed to bind specific mRNA sequences to form a morpholino-RNA heteroduplex, preventing their expression. Analysis of the resulting phenotype in animal models informs on the function of the gene.

17.2.15. Homeotic Genes and Body Plan Development

Homeotic genes regulate morphogenesis (changing the shape of an organism), which are active during embryonic development of multicellular eukaryotes. They encode the body plan in plants, animals, fungi and some protists. Homeotic mutations can cause serious developmental defects and are often fatal. Homeotic genes can exhibit:

- Heterochrony: time-dependent expression during development.
- Heterotopy: space-dependent expression during development.
- Heterometry: changes in the extent of gene expression.
- Heterotypy: changes in developmental outcome due to regulatory mutations.

Homeobox genes are homeotic genes that contain a conserved 180 bp sequence (homeobox) within them. The homeobox gene translates into a transcription factor protein, with the homeobox region translating into a constant 60 amino acid sequence (homeodomain: three α-helices folded around a hydrophobic core and a flexible N-terminal arm). The helix-turn-helix motif can bind DNA, epigenetically up/downregulating transcription. The homeodomain can recruit histone-modifying enzymes and interact with chromatin-remodelling complexes and microRNAs.

Hox genes: the main class of homeobox genes in bilaterian animals, originally discovered in *Drosophila* (fruit flies). They are found arranged in clusters on chromosomes. In some lineages, including vertebrates, Hox genes have been duplicated, giving four clusters across multiple chromosomes. The ordering of the genes along the chromosome corresponds to the order of phenotypic expression along an anterior-posterior axis (spatial collinearity).

Pax genes (paired box genes): another class of homeobox genes.

Gap genes demarcate the boundaries between body segments, and are activated through interactions between the protein products of MEGs (maternal effect genes, Section 17.2.11), and they also regulate each other. This allows for space-dependent control of mitosis, apoptosis and cell differentiation, forming different segments of the body.

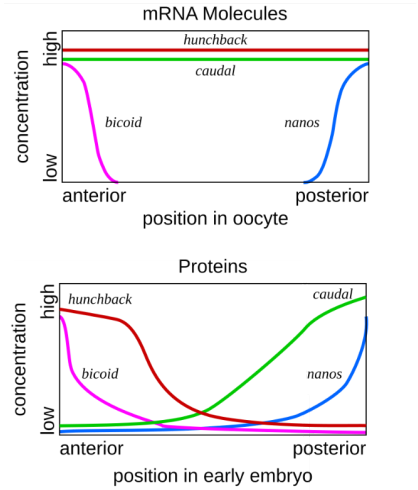
17.2.16. Embryonic Development

Morphogenesis: the progressive developmental organisation in a growing embryo

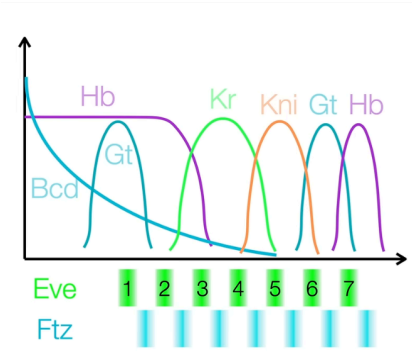
- Haeckel: development follows evolutionary history ('ontogeny recapitulates phylogeny').
- Von Baer: development diversification reflects evolutionary diversification (no recapping).

Von Baer's concept (while not perfect), is the more accurate summary for evo-devo biology.

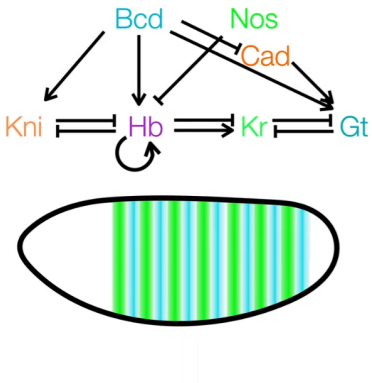
Maternal Localisation: At fertilisation, maternal effect genes (MEGs) encode factors (e.g. mRNAs) that are present in the oocyte (egg cell), and are bound to the poles of the cell (maternal localisation), which initialises the head-tail axis. The cascade following MEG expression in *Drosophila* are shown:



Initial mRNA concentration and resulting protein distribution due to diffusion



Second cascade layer forming segments (*Eve*, *Ftz*: pair-rule genes)



Regulatory network for cascade layers 1 and 2, and resulting embryo segments

The third cascade layer sets up the polarity within each segment using 'segment polarity genes'. The fourth cascade is the spatially dependent activation of the homeotic selector genes (Section 17.2.10). These cascades can be modelled as bistable oscillators using systems of ODEs (Section 17.2.7).

17.2.17. Sexual Dimorphism and X-linked Genes

During embryonic development of biologically female organisms (XX chromosomes), one of the X-chromosomes in each diploid cell is randomly permanently inactivated (lyonisation), in order for X-linked gene expression to equalise in males and females. For heterozygous females, neighbouring cell clusters may express different alleles of X-linked genes (mosaicism). X-linked recessive inheritance causes various disorders in homozygous XX females.

17.2.18. Organogenesis and Stem Cell Differentiation

The Hedgehog Signalling Pathway

Embryonic stem cell differentiation is controlled by spatial expression of hedgehog (*hh*) genes. Mammals have three hedgehog homologues, ‘desert’ (DHH), ‘indian’ (IHH), and ‘sonic’ (SHH).

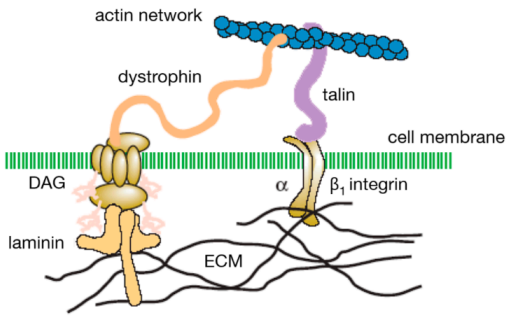
When SHH is translated, the protein is processed by forming an N-terminal domain and O-linkage with a cholesterol molecule for lipid transport out of the cell membrane. By paracrine signalling, a DISP protein

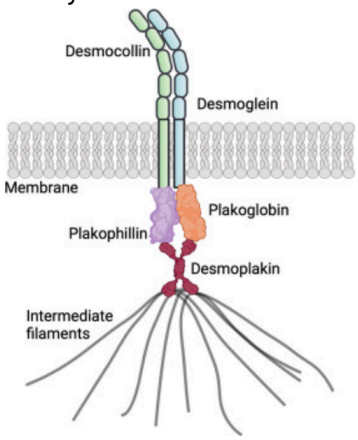
17.3. Biomechanics and Biophysics

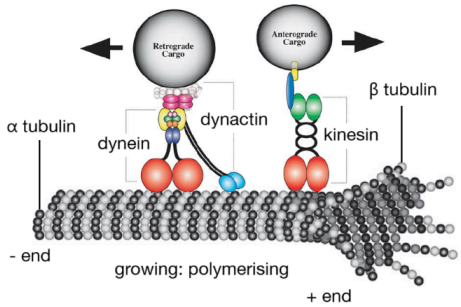
17.3.1. Cytoskeleton and Cellular Biomechanics

The cytoskeleton is a protein scaffold. The overall mechanical rigidity of a eukaryotic cell is dominated by the cytoskeleton, since the lipid membrane is compliant and the cytoplasm is primarily viscous only.

The **persistence length** of a rubber-like polymeric fibre is the length scale over which the curvature of the fibre remains correlated before the effect of random thermal Brownian motion 'scrambles' the structure. Persistence length of a fibre: $l_p = EI / k_B T$ (EI: flexural rigidity, $k_B T$: thermal energy). Sections of polymer shorter than l_p can be modelled as elastic beams, while for sections longer than l_p, an 'entropic spring' statistical model (like natural rubber) is more appropriate. Flexible polymer chains can migrate by a 'slithering' mechanism (reptation). The cytoskeleton has three main parts at different scales of l_p:







Actin Filament

thin strands linked to cell membrane

t = 7-9 nm, l_p = 15 μm, E = 1.3-2.5 GPa

- In **'actin networks'**, filaments form loosely crosslinked 3D meshes near the cell membrane, bound to integrins/laminins by dystrophin and talin.
- In **'actin bundles'**, filaments form tightly-bound parallel arrays, cross-linked by actin-bundling proteins with Ca^{2+}-binding domains.

Intermediate Filament

bundles of twisted-coil chains

t = 12 nm, l_p = 1-3 μm, E = 2-5 MPa

- Made of fibrous proteins (in the cytoplasm: keratin, vimentin, neurofilament).
- **Cytoplasmic IFs** are **passive** elements, providing elasticity.
- **Nuclear lamins** form square lattices and support the inner nuclear membrane. It also helps anchor heterochromatin to the nuclear membrane and can indirectly affect gene expression via epigenetics.
- **Spectrin** (α and β) forms a triangulated lattice on inner cell membranes with flexible hinge proteins, providing high compliance, high ultimate tensile strength and high lock-up strain (until filaments straighten out).

Microtubule

rigid hollow cylindrical tubes

d_o = 25 nm, d_i = 14 μm, l_p = 5200 μm, E = 1.9 GPa

- Made of helical globular tubulin heterodimer proteins (α and β). These blocks can have post-translational modifications (e.g. phosphorylation) to influence microtubule functionality (the tubulin code)
- Microtubules are used as a **'molecular highway'** by motor proteins (dynein and kinesin).
- The mitotic spindle organises the chromatids during mitosis.
- The '-' end is anchored at the cell centrosomes, adjacent to the Golgi apparatus, where modified proteins in vesicles are collected by kinesin. The '+' protrudes freely into the cytoplasm and has a **'dynamic instability'**.

17.3.2. Mechanobiology and Mechanotransduction

Cells respond to mechanical and osmotic forces in their environment, as well as use enzyme-catalysed reactions to perform mechanical work at the molecular level.

Mammalian cells are ~70% water by mass. Their dry mass (without water) is composed of 60% protein, 13% lipids, 4% RNA (ribosomes), 1% DNA (remaining 22% polysaccharides, metabolites, ions etc). The overall cell mass is approximately 10^{-12} g = 1 pg.

Experiments find that various gene expression pathways are epigenetically inhibited in microgravity (μg), suggesting that mechanotransduction is fundamentally an adaptation to the presence of gravity on life.

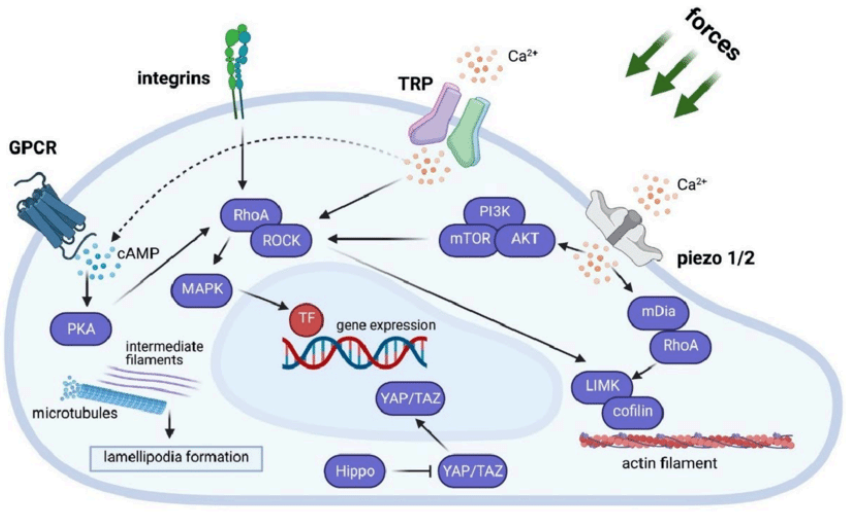
Cell migration can be influenced by molecules (chemotaxis) or stressors (mechanotaxis):

- Durotaxis: migration up gradients in ECM substrate mechanical rigidity
- Plithotaxis: cells migrate in-plane along axes of principal stresses (minimal shear stress)
- Haptotaxis: directional outgrowth up gradients of adhesion sites and chemoattractants
- Gravitropism: plant cell growth in the directions aligned with self-weight (roots: with, shoots: against)

Mechanisms of Mechanotransduction

- Stretch-activated ion channels (SACs)
- Contact with fibronectin in the ECM promotes focal adhesions via vinculin, talin, integrins
- Shear forces cause glycocalyx deformation with coupling to the cytoskeleton

Signalling Cascades Associated following Mechanical Stimulation



Ca^{2+} is transported through mechanosensitive ion channels such as Piezo1 and Piezo2 (biologically analogous to the piezoelectric effect).

Integrins bind cells to the ECM signalled by RGD sequences, and the focal adhesions are matured by vinculin, promoting stability.

Mechanical forces can influence gene expression via pathways to mechanotransducers (e.g. YAP/TAZ) and transcription factors (e.g. NF-κB).

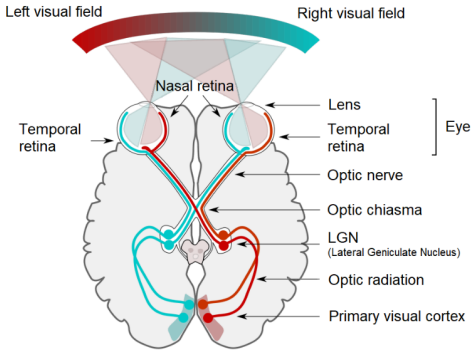
(GPCR: G-protein coupled receptor, PKA: protein kinase A, TRP: transient receptor potential)

Together with chemical signals, mechanical signals also control cell growth, differentiation and fate switching in developing embryos, giving rise to ‘evo-devo mechanobiology’.

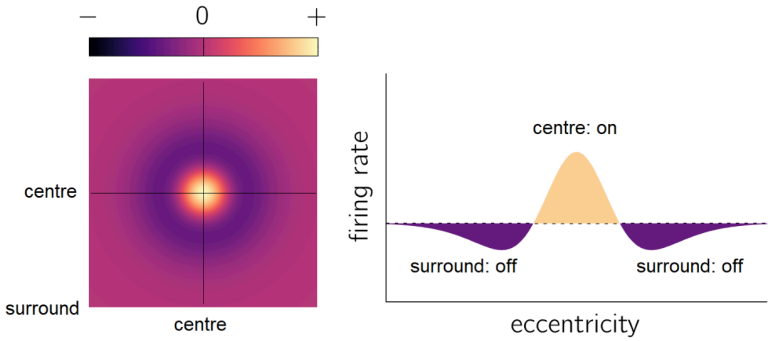
17.3.3. Photoreception (The Eye)

Neuroscience of Vision and Visual Processing in Primates

The Visual Pathway



On-Centre Ganglion Cell Receptive Field: Difference of Gaussians



Receptive Fields: the variables (spatial position etc) which determine when a neuron will fire.

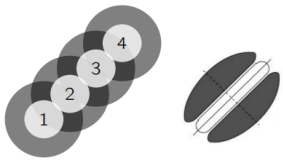
- Photoreceptor: the small angular position on the retina.
- Retinal ganglion cell (GC) and LGN neurons: the angular positions of nearby photoreceptors, with feedback from the centre of the field and opposing feedback from the surroundings. This approximates a difference of Gaussians (DoG) filter which acts as a contrast (edge) detector:

  ($G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right)$ is the 2D Gaussian blob detector kernel.)

  $$G(x, y, k\sigma) - G(x, y, \sigma) \approx (k-1)\sigma^2\nabla^2 G(x, y, \sigma)$$

- 'Simple' neurons: oriented, edge/line-like stimulus. Hubel-Wiesel's feedforward model: the V1 neurons are fed by a line of GCs. Found in Brodmann area 17 (V1: primary visual cortex). Approximates a Gabor filter:

  $$f(x,\ y)\ =\ exp\left(-\frac{x^2+(yy)^2}{2\sigma^2}\right) cos(\omega x\ +\ \phi)\ \ (y\text{: line axis})$$

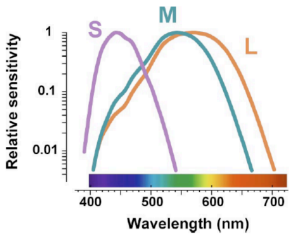


- 'Complex' neurons: nonlinear summation, motion sensitive along an axis, less specific to retinal location. Found in areas 17 (V1) and 18 (V2). Note: 'hypercomplex' neurons have end-stopped receptive fields, inhibiting after a certain length. Both simple and complex neurons can have this property. Found in areas 18 (V2) and 19 (V3) and **not** in V1.

Spatial Organisation: V1 is highly nonlinear, as shown by cortical maps (retinotopy). Cortical columns are block regions with similar stimuli (but have a range of spatial frequencies). Cortical hypercolumns combine nearby columns to include all orientations, with long-range connections encompassing multiple blobs.

Channel theory: different features of the visual scene (colour, orientation, spatial frequency, direction of motion) are represented in independent channels (merged in later processing stages). Evidence: contrast sensitivity function, neural fatigue/after-effects on over-exposure.

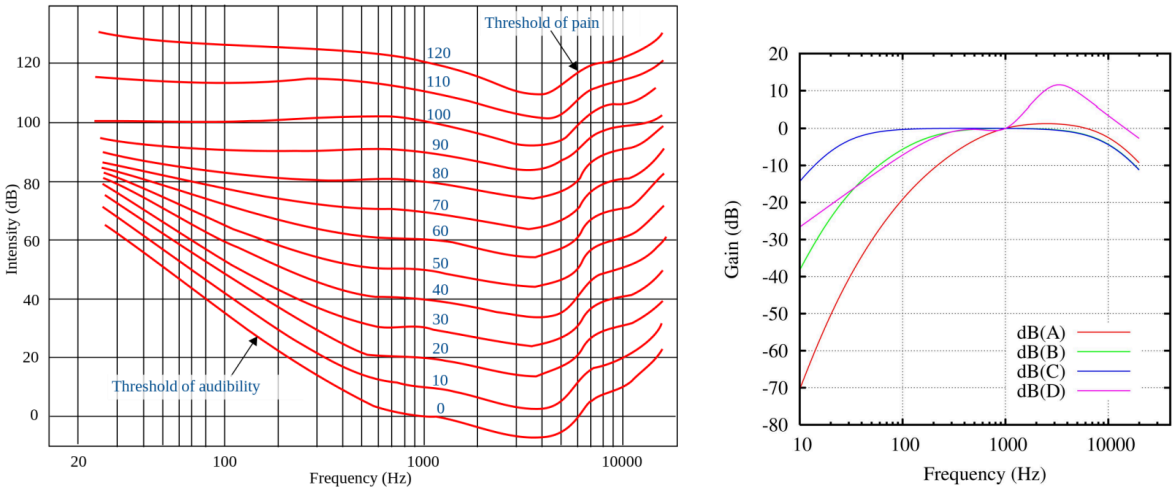
Trichromatic vision: three cones (S, M, L) in the fovea respond to wavelength bands. The colour-opponent channels are L+M+εS (achromatic), L-M (red-green) and L+M-S (yellow-blue). It is thought that these evolved to be the optimal way of coding (i.e. principal components) of the reflectance spectra for typical natural materials.



17.3.4. Phonoreception (The Ear)

Audiogram and Weighting Filter for Human Hearing Sensitivity

Audible range: approximately 20 - 20 kHz.



An **A-weighting filter** (red curve on the right) is applied to loudspeakers to compensate for the equal-loudness curves of human hearing (left).

Damage to hearing: hearing loss due to age raises the equal-loudness curves across all frequencies. Hearing loss due to an injury raises the curves more around one (over-exposed) frequency.

17.3.5. Bioacoustics

Phonation (Vocalisation)

Echolocation

Echolocation is used by bats and toothed whales.

17.3.6. Olfactory Reception (The Nose)

17.3.7. Mechanoreception and Thermoreception

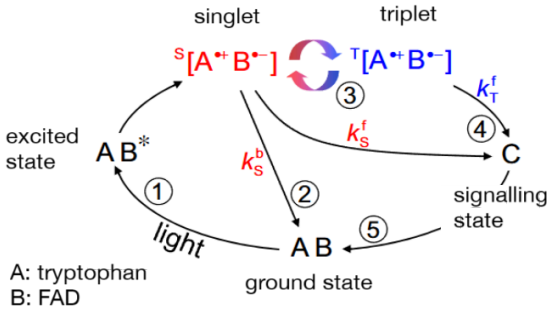
17.3.8. Electroreception and Electrogenesis

17.3.9. Magnetoreception (Response to Magnetic Fields)

Many animals (some arthropods, molluscs, and vertebrates) use magnetoreception for navigation by orienting with Earth's magnetic field. Several independently evolved mechanisms exist.

Cryptochrome: a flavoprotein in retinal rod cells in the eyes of some animals.

In plants, cryptochromes mediate phototropism. In animals, type 1 cryptochromes regulate the circadian rhythm while type 2 cryptochromes can sometimes be used for magnetoreception via the quantum mechanical 'radical pair mechanism', modelled as follows:



① Cryptochrome has a FAD chromophore cofactor absorbing at ~465 nm (blue light), forming an exciton FAD* with ~50% quantum yield.

② FAD* dissociates by charge transfer with a tryptophan residue in cryptochrome. The radical spins are correlated in the singlet state (↑↓).

③ Intersystem crossing to the triplet state (↑↑). When a magnetic field aligns with the spins, the triplet state is stabilised (Zeeman effect).

④ H^+ transfer to $FAD^{\bullet-}$ induces conformational change in the signalling state of the complex, with field-dependent formation kinetics ($k_T^f << k_S^f$). A downstream neurotransmitter cascade is therefore field-dependent.

In birds, the magnetoreception signal activates 'cluster N' in the forebrain of the bird brain. Low-intensity stray radio-frequency electromagnetic fields can disorient migratory birds by interfering with the geomagnetic field at the cryptochrome complexes. This biological compass is likely noisy and requires signal integration over several seconds to improve the fidelity.

Cryptochrome has high DNA sequence homology with the DNA repair protein (6-4) photolyase, sharing high-energy photon sensitivity due to the FAD cofactor. Molecular clock phylogenies indicate that animal and plant cryptochromes both evolved in the late Precambrian, ~1000-541 MYA.

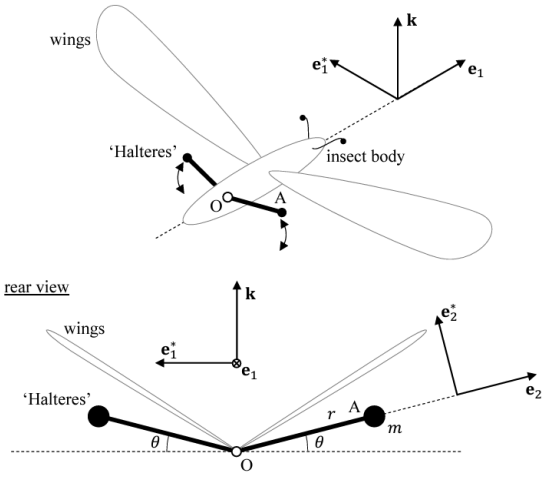
Magnetosome: liposomal superparamagnetic ~100 nm particles (magnetite (Fe_3O_4) / greigite (Fe_3S_4)).

Magnetosomes are found in the polyphyletic group of magnetotactic bacteria (MTB), most being in phylum *Pseudomonadota*, with inter-phylum horizontal gene transfer. MTBs are found at the boundary between oxic and anoxic waters and sediment and use their passive magnetic alignment to travel up and down the water column. The magnetosome crystals are produced by biomineralisation, arranged in a linear chain inside the cell membrane as their total magnetic moment produces the least demagnetizing field when aligned parallel to the field. The chains are also bound to the cytoskeleton filaments (via the actin homolog mamK) and so orient the cell.

The core operons for magnetosomes are conserved across all MTBs on the 'magnetosome island' (MAI). The mamAB operon initiates formation, while the mm6, mamGFDC and mamXY operons control crystal growth. Magnetosomes are likely very ancient, occurring around the time of the Great Oxidation Event or earlier, with an unknown selective pressure. Eukaryotic magnetosomes are rare, only occurring with functionality in some aquatic unicellular protists (euglenids and algae). Free magnetite particles are known in many animals but their purpose is unclear, and they may be serving as stores of iron. Sea turtles migrate with an unknown mode of magnetoreception, hypothesised to be due to symbiotic MTBs in the lacrimal glands of the eyes, potentially transported by macrophages, with connections to the ophthalmic nerve.

17.3.3. Proprioception

Body Rotation Sensing: halteres in insects.



Insects in orders *Diptera* (flies) have structures called **halteres** (instead of hindwings) that give feedback on the body's rotation during flight. Halteres are rods that protrude from the body and vibrate at high frequency $\theta(t) = \theta_0 \sin \dot{\theta} t$ perpendicular to the insect's heading $\mathbf{e}_1$. When the insect's heading is changing with angular velocity $\Omega\mathbf{k}$ ($\dot{\theta} >> \Omega$), the halteres act as vibrating structure gyroscopes, and mechanoreceptors (campaniform sensilla) detect the Coriolis torque ($Q = 4m\Omega r^2\dot{\theta} \cos \dot{\theta} t$) at the root of the halteres.

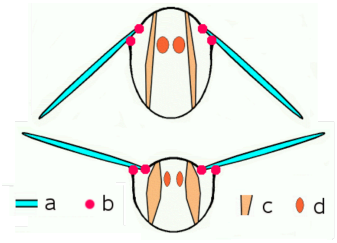
Halteres also convergently evolved in order *Strepsiptera* (twisted-wing endoparasites), and similar flight stabilisation apparatus is seen in order *Lepidoptera* (moths and butterflies) as their antenna.

To control movement, the nervous system must integrate multimodal sensory information (both from the external world as well as proprioception) and generate the necessary signals to send to muscles to actuate the body. Motor control is typically handled in the cerebellum of the brain.

17.3.4. Aerial Locomotion (Flight)

Flight in animals may be **powered** (stationary body, moving wings) or **unpowered** (stationary wings, body moving relative to air).

Powered Flight in Insects: both thrust and lift generated continuously by flapping wings



Insect wings **(a)** are aerofoil-shaped. During the 'downstroke', the wing moves down and forward, and the high angle of attack forms a low-pressure leading-edge vortex, increasing the lift produced. The wing is then inverted quickly before the upstroke, moving up and backward, to ensure lift continues to act upwards, and then inverted again to complete the cycle. The movements are actuated via a complex hinge **(b)**, using stretch-activated muscles **(c)** in the thorax.

Powered Flight in Birds: thrust and lift mainly generated on the downstroke only

The signals to actuate wings are generated by the cerebellum in the brain as part of motor control.

17.3.5. Aquatic Locomotion (Swimming)

17.3.10. Structural Analysis of Biomaterials and Cellular Tissue

Periodic Stiff Lattices (Rigid Biopolymer Foam Model)

Structural biopolymer networks can be modelled as 3D frameworks of elastic-perfectly plastic 'bars' with plastic hinge joints to model frictional torques at cross-links. The mechanical properties of a rigid-jointed truss are similar to the equivalent pin-jointed truss, but:

- The rigid-jointed truss will have a weak **bending** mode if the pin-jointed truss is mechanistic.
- The rigidity of the truss depends on the lattice connectivity, Z (average number of bars per joint).

Connectivity Criteria for Rigidity of Periodic Lattices

	Stiff Stretching dominated	**Compliant** Bending dominated
2D	$Z \geq 4$	$Z < 4$
3D	$Z \geq 6$	$Z < 6$

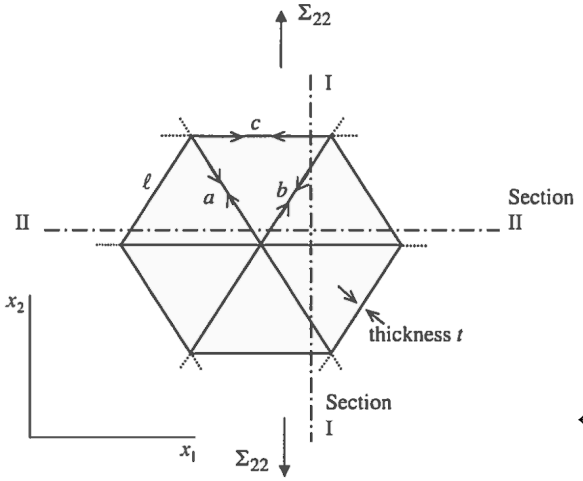
Effective Properties of Anisotropic Lattices in 2D (Prismatic) and 3D (Foams)

Direction	**Modulus** reduction factor		**Strength** reduction factor	
Transverse (weak)	2D	$\overline{E}/E_s \sim \overline{\rho}^3$	2D	$\overline{\sigma_y}/\sigma_y \sim \overline{\rho}^2$
	3D	$\overline{E}/E_s \sim \overline{\rho}^2$	3D	$\overline{\sigma_y}/\sigma_y \sim \overline{\rho}^{3/2}$
Axial (strong)	$\overline{E}/E_s \sim \overline{\rho}$		$\overline{\sigma_y}/\sigma_y \sim \overline{\rho}$	

Notable points:

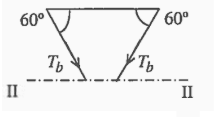
- Any 2D lattice with 120° rotational (C_3) symmetry has an isotropic in-plane modulus.
- Irregular/random polygons in the lattice do not significantly affect the structural properties (~3%).
- For a 3D foam topology to be rigid, its unit cell must tessellate and be rigid ($Z \geq 6$). None of the platonic solids achieve this, so combinations are necessary (e.g. FCC octet truss).

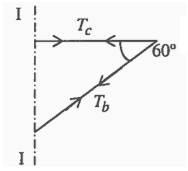
Example: for the stiff 2D triangulated lattice shown, what is the effective Young's modulus and yield stress?



(Σ_{22}: macroscopic applied axial stress, ε_{22}: axial strain, E_s: link elastic modulus l: cross-linker length, $\{a, b, c\}$: link tensions. Unit cell: equilateral triangle with thickness $t/2$.)

Elastic Analysis: consider **free-body cuts** of the unit cell





$$\overline{\rho} = \frac{\text{unit cell solid area}}{\text{unit cell total area}} = 2\sqrt{3}\,\frac{t}{l}$$

$$A = \frac{\sqrt{3}}{4}l^2$$

By symmetry, $T_a = T_b$

$\leftrightarrow: (T_a + T_b)\cos 30° = \Sigma_{22}\, l$

$T_a = T_b = \Sigma_{22}\, l/\sqrt{3}$

$\leftrightarrow: T_c + T_b \cos 30° = 0$

$T_c = -½\, T_b = -\Sigma_{22}\, l/2\sqrt{3}$

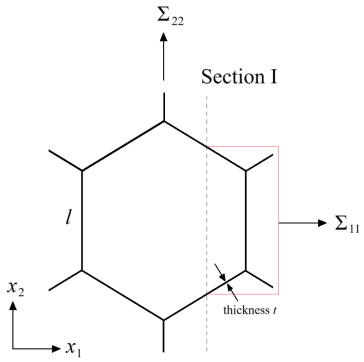
Virtual work (per unit depth): $\Sigma_{22}\,\varepsilon_{22}\,A = = \frac{1}{2}(T_a e_a + T_b e_b + T_c e_c)$; (factor $\frac{1}{2}$ as bar forces are shared per cell)

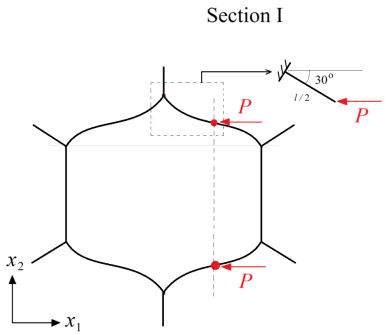
Elastic law: $e = \frac{Tl}{E_s t} \rightarrow \frac{\Sigma_{22}}{\varepsilon_{22}} = \overline{E} = \frac{1}{3}\overline{\rho}E_s$.

Plastic Analysis: assume bars a, b attain yield (perfectly plastic) while c remains rigid. Let $T_a = T_b = \sigma_y t$. The FBD cut at Section II-II shows that $\sigma_y t = \Sigma_{22}\, l/\sqrt{3} \rightarrow$ therefore $\Sigma_{22} = \overline{\sigma_y} = \sqrt{3}\sigma_y\, t/l = \frac{1}{2}\overline{\rho}\sigma_y$.

Periodic Compliant Lattices (Honeycomb Biopolymer Foam Model)

Example: for the compliant 2D hexagonal foam shown, what is the effective modulus and yield stress?





Imagine pulling the lattice in each direction. Place Section I at the inflection point of a beam.

Relative density: $\bar{\rho} = \frac{2}{\sqrt{3}} \frac{t}{l}$

Elastic analysis: at Section I, curvature $\kappa = 0$ $\rightarrow$ bending moment $M = 0 \rightarrow$ only a force exists.

By force balance on Section I, per unit depth, $P = \Sigma_{11}\, l\,(1 + \sin 30°) \rightarrow P = \frac{3}{2}\,\Sigma_{11}\, l.$

Consider a half-section of the beam anchored at the joint, which can be modelled as a cantilever.

- Force P has transverse component $\frac{1}{2}P$, which causes transverse tip deflection by $u = \frac{Pl^3}{48\, E_s\, I}$.
- The strain $\varepsilon_{11} = \frac{2(u\ sin\ 30)}{l\ cos\ 30^o} = \frac{2\sqrt{3}}{3}\frac{u}{l}$. ($I = \frac{1}{12}t^3$: second moment of area per unit depth)
- The effective modulus is then $E_{11} = \frac{\Sigma_{11}}{\varepsilon_{11}} = \frac{3}{2}E_s\,\bar{\rho}^3$. This modulus is isotropic.

Plastic analysis: assume plastic hinge formation (Section 6.3.22) at the cantilever roots (top, bottom joints).

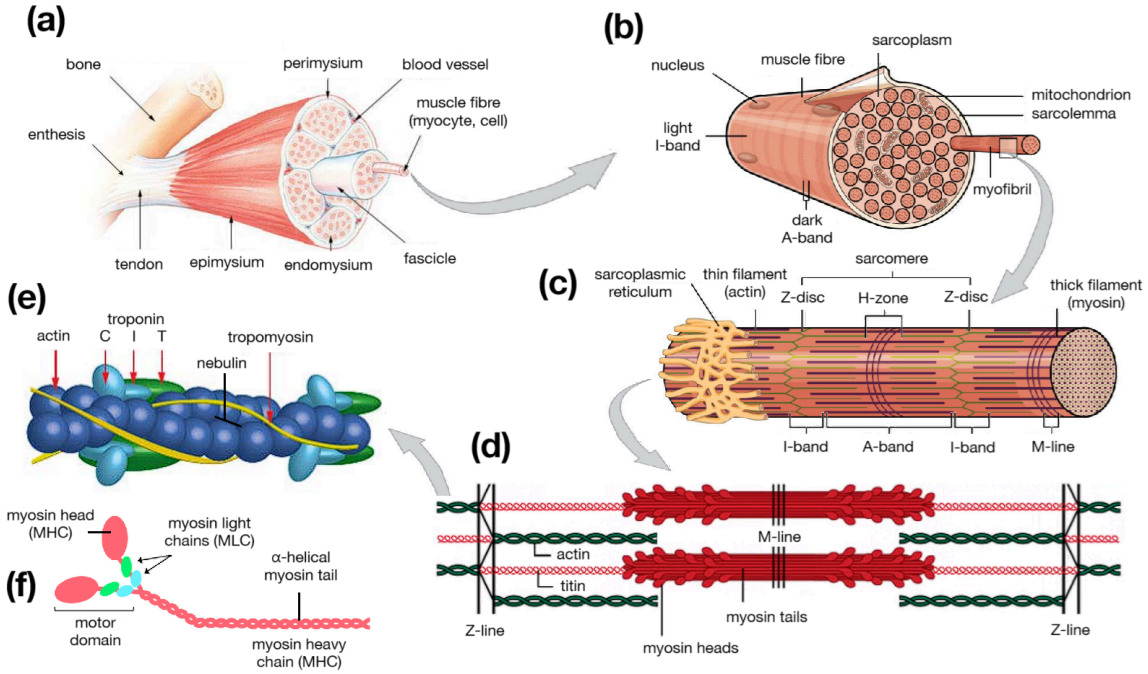
- External work = internal dissipation : $\Sigma_{11}\,\varepsilon_{11} \times \frac{3\sqrt{3}}{2}l^2 = \frac{1}{2}M_p \times (4\theta \times 2)$ (θ: hinge rotation)
- Plastic modulus per unit depth: $Z_p = \frac{1}{4}t^2 \Rightarrow$ plastic moment capacity: $M_p = \frac{1}{4}t^2\,\sigma_y$ (Section 6.3.21)
- Deformation geometry: $\varepsilon_{11} = \frac{l\theta\ sin\ 30^o}{l\ sin\ 60^o} = \frac{\theta}{\sqrt{3}}$. At yield, let $\Sigma_{11} = \overline{\sigma}_y$. Then $\overline{\sigma}_y = \frac{2}{3}\,\sigma_y\,\frac{t^2}{l^2} = \frac{1}{2}\,\sigma_y\,\bar{\rho}^2$.

Examples of Biological Structures as Lattice Models

- Square lattice (2D, $Z = 4$): nuclear lamins supporting the inner nuclear membrane of eukaryotic cells.
- Hexagonal lattice (2D, $Z = 3$): cellulose and lignin honeycomb cell walls in xylem tubes and wood.
- Triangulated lattice (2D, $Z = 6$): spectrin fibres supporting the cell membrane of erythrocytes.
- Amorphous network (2D): wavy collagen fibres in skin giving very low elastic modulus.

17.3.11. Muscle Anatomy and Physiology

Muscle Anatomy: a hierarchical arrangement of parallel fibrous structures



(a) Macroscopic skeletal muscle anatomy. **(b)** A single muscle fibre (cell). **(c)** A single myofibril. **(d)** A single sarcomere lengthwise unit of myofibril. **(e)** A single actin filament. **(f)** A single myosin unit.

Muscle Signalling: pathway for a conscious striated muscle contraction

1. **Motor cortex** (part of the cerebral cortex) of the brain generates a signal from conscious inputs.
2. **Motor neurons** in the CNS (spinal cord) propagate the action potential to the PNS.
3. Exocytosis of **acetylcholine (ACh)** into synaptic cleft of motor unit neuromuscular junction.
4. Binding of ACh to receptors on sarcolemma, inducing **Na^+ uptake** into the muscle cell.
5. An **action potential** propagates along the **sarcolemma** and into T-tubules.
6. Voltage-gated dihydropyridine receptors are activated, mechanically linked to ryanodine receptors, which induce **Ca^{2+} ion release** from the sarcoplasmic reticulum into the sarcoplasm.

Muscle Function: Lymn-Taylor crossbridge cycle / sliding filament theory for **striated muscle** contractions

1. **Calcium binding:** Ca^{2+} binds to troponin C, displacing tropomyosin, exposing actin binding sites.
2. **Hydrolysis:** a myosin head containing ATP is hydrolysed to ADP and P_i, retaining them and changing the conformation of the myosin head to the 'cocked' state.
3. **Cross-bridge formation:** cocked myosin head (holding ADP + P_i) binds to actin sites.
4. **Power stroke:** the P_i is released, causing the myosin head to pull the whole actin filament towards the centre of the sarcomere (power stroke). The ADP is then also released.
5. **Dissociation:** a new ATP binds to myosin, breaking the cross-bridges and returning the myosin to the 'cocked' state, ready to release the conformational strain energy again.

This is the basis of the **sliding filament theory**: myosin thick fibres are fixed in place, and their heads continuously bind actin filaments and pull it towards the M-line (sarcomere centre). The sarcomere length (between Z-discs) decreases, so the muscle experiences contraction.

Muscle **relaxation** is the reverse process, with Ca^{2+} being actively transported **out** of the sarcoplasm back into the sarcoplasmic reticulum by calcium pump ATPases.

All skeletal muscles act in **antagonistic pairs** (one shortens, one lengthens). The striated muscles are attached to bones in the skeleton by **tendons** (made of collagen and proteoglycans), so the contractions exert forces and torques on the rigid skeleton, which induce translational and rotational motion.

Muscle fibres can be **slow twitch or fast twitch**. These use different isoforms of myosin heavy chain (slow: MHC I, fast: MHC IIa) to achieve differences in cross-bridge cycling rate. Fast twitch muscle fibres are suited for short bursts of force with high maximum mechanical power output, while slow twitch muscle fibres are suited for long-lasting low force production.

Smooth Muscles

In **smooth muscles**, Ca^{2+} instead acts via binding to calmodulin and phosphorylation of myosin light-chain (MLC) starts the cross-bridge cycle. Smooth muscles are involuntary muscles (stimulated by unconscious nerve impulses), found near soft organ tissue (except the heart cardiac muscle) and do not actuate the skeleton.

Muscles in Other Animals (comparative biomechanics)

Muscles are highly evolutionarily conserved from around ~600 MYA (Precambrian; Ediacaran), appearing in all but the most basal animal phyla (e.g. *Porifera*: sea sponges) as smooth muscle cells. Phylum *Cnidaria* (e.g. jellyfish) have the most primitive of muscle apparatus, lacking the embryonic mesoderm layer that all other phyla develop muscles (and the heart and circulatory system) from but having striated cells in its reproductive stage.

In **small animals** (e.g. insects, arthropods and some amphibians), very rapid movements can be generated despite tiny muscle masses, due to 'latch-mediated spring actuation' (LaMSA). The muscle fibres contract and store elastic energy in tendons or specialised structures, and a latch mechanically holds the tendon in place without expending further energy. On release of the latch, the stored energy is released as kinetic energy. This exploits passive elasticity to amplify the mechanical power significantly.

In **apes**, muscularity varies due to evolutionary adaptation. Humans have on average ~70% slow and ~30% fast twitch muscle fibres, giving us relatively high endurance and manual dexterity but low overall muscle strength. The other great apes (e.g. gorillas, chimpanzees) possess higher fast-twitch muscle fibre content as well as increased overall muscle mass due to decreased myostatin expression.

Muscle Energy Sources

Striated muscles are specialised for respiration. Oxygen gas binds to myoglobin for facilitated diffusion, ensuring a plentiful supply for mitochondrial respiration. Slow-twitch muscle fibres have many mitochondria for aerobic respiration to generate ATP. Fast-twitch muscle fibres have few mitochondria, performing glycolysis and anaerobic respiration to form lactic acid, which must be cleared by transport to the liver, and the oxygen must be replenished by panting (oxygen debt).

The ATP required for the myosin crossbridge cycle is sourced from existing ATP stores, as well as the on-demand **Lohmann reaction** (phosphocreatine + ADP $\rightleftharpoons$ creatine + ATP, via creatine phosphokinase) in the sarcoplasm.

Muscle Growth and Repair

Muscles grow primarily by hypertrophy (myocyte size) rather than hyperplasia (myocyte count). During **strenuous exercise**, minor damage to muscles can result from increased metabolic reactive oxygen species (ROS) production, leading to inflammation which triggers a local innate immune response. Satellite cells (muscle stem cells) undergo **myogenesis** to produce new muscle myofibrils. Microtears occurring from mechanical stress trigger nuclear migration to the site to deliver mRNA for **protein synthesis** to fill in the microtears. The mTOR pathway also promotes muscle growth by upregulating protein synthesis in myocytes.

Muscle injuries and soreness is often caused by eccentric contractions (muscle lengthening) since high tensile forces are developed even for small extensions. A careful balance in resistance training intensity and protein intake is required to promote muscle growth without damage.

Abnormal Muscle Function: disruptions at any point in the muscle signalling cascade. Examples:

When an organism **dies**, the calcium pumps are shut down and Ca^{2+} diffuses into the sarcoplasm, causing contraction and the diminished ATP stores cannot break the myosin crossbridges, so the muscle remains firmly contracted for ~12 hours (*rigor mortis*), only breaking when the proteins themselves begin to degrade by action of free proteases and microorganisms.

Nerve agents inhibit acetylcholinesterase in the synaptic cleft, leading to ACh accumulation and uncontrolled muscle contractions.

Botulinum toxin (botox) from the bacteria *Clostridium botulinum* blocks ACh exocytosis, promoting muscle relaxation (paralysis / botulism), used by cosmetic surgeons to relax facial muscles, reducing wrinkles. The toxin from *Clostridium tetani* has the opposite effect, causing tetanus disease (forced contractions).

17.3.12. Experimental Muscle Mechanics

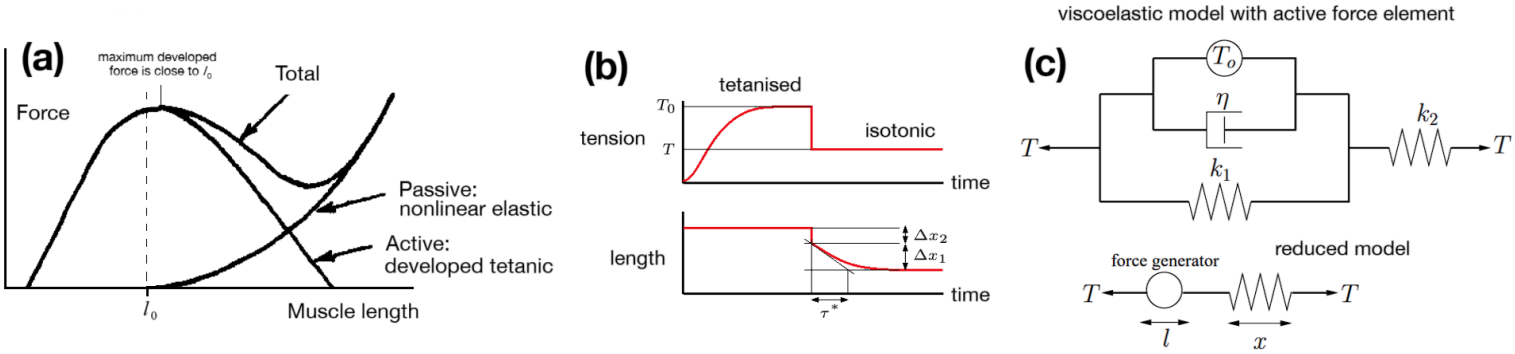
Isometric Stimulation of Muscles: hold at constant length L, stimulate, measure force $T(x)$.

A muscle removed from an animal and kept in a cold sterile isosmotic solution can be studied *ex vivo* for its response to stimuli. On electrical stimulus (applied voltage), a muscle held isometrically (constant length) responds with a tensile force depending on the applied stimulus:

- **Twitch**: single impulse stimuli → transient decaying response. The tension rises and falls back to zero within about one second.
- **Unfused tetanus** (clonus): low-frequency impulse train → superposition of twitches. The tension reaches a moderate steady-state value with small oscillations.
- **Fused tetanus**: high-frequency (>50 in mammalian muscles) impulse train → tension rises to maximum steady-state value with no oscillations.

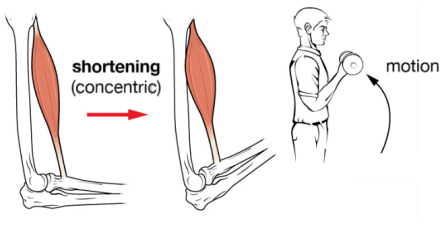
Total muscle force $T(x, t)$ = Tetanic (Active) force $T_a(x, t)$ + Passive Elastic force $P(x)$.

Isotonic Stimulation of Muscles: tetanise to tension T_0, stop and apply step tension $T < T_0$, measure length response $L(t)$. The negative creep response resembles a viscoelastic standard linear solid model (Section 6.2.9) with an additional active force component. The initial gradient of the relaxation curve $x(t)$ is the 'shortening velocity', $v = \Delta x_1 / \tau^*$.

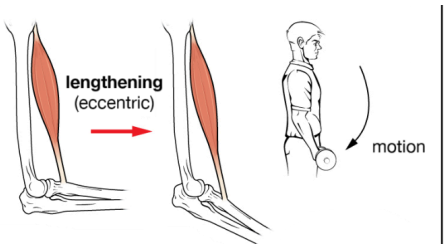


(a) Isometric force-length characteristic. **(b)** Hill's step drop experiment. **(c)** Active standard linear solid model (Section 6.2.9) and reduced model to explain the isotonic experimental data.

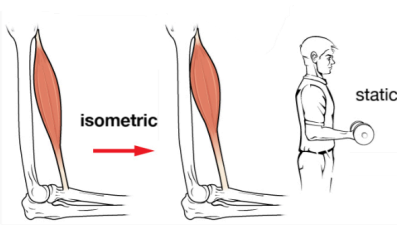
Types of Active Muscle Contractions in Exercise: muscle is usually in unfused tetanus state



Isotonic Shortening (concentric)
bicep shortens, tricep lengthens



Isotonic Lengthening (eccentric)
bicep lengthens, tricep shortens

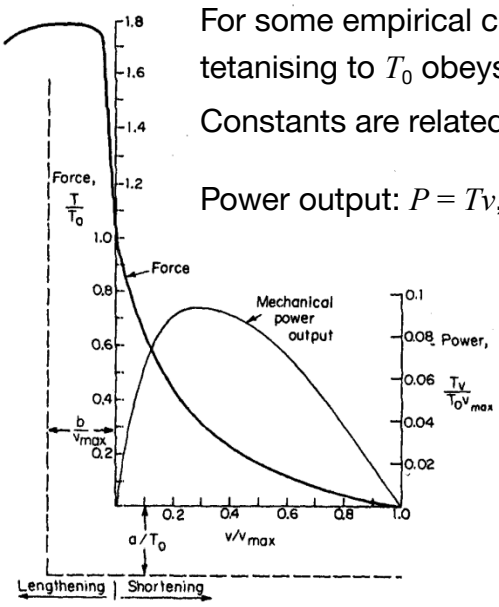


Isometric Contraction
neither changes length

Note: bicep curls as shown are not truly isotonic due to time-varying torque demand ('auxotonic').

17.3.13. Modelling of Muscle Mechanics

Hill's Active State Model: relates T and v from **isotonic** tetanising experiments.



For some empirical constants (a, b), the tension T and shortening velocity v after tetanising to T_0 obeys the **Hill equation**, $(T + a)v = b(T_0 - T)$.

Constants are related by $v_{\max} = \frac{b}{k}$ where $k = \frac{a}{T_0}$. Typically, $4 < \frac{1}{k} < 6$.

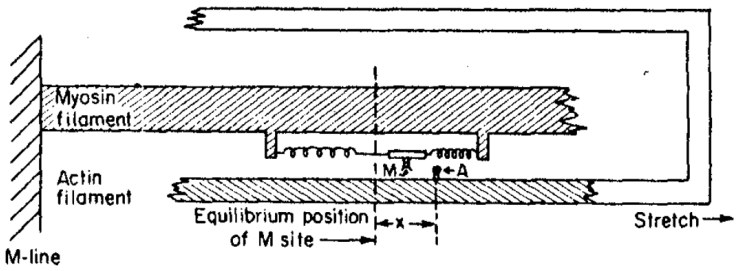
Power output: $P = Tv$, normalised as $\widehat{P} = \widehat{T}\,\widehat{v}$ where $\widehat{T} = \frac{T}{T_0}$, $\widehat{v} = \frac{v}{v_{max}}$, $\widehat{P} = \frac{P}{T_0 v_{max}}$

Maximum power output occurs at $\widehat{v} = k(\sqrt{1 + 1/k} - 1) \sim 0.3$.

Fenn's energy balance: the total heat developed by a muscle is $Q = -(T + \alpha)\,\Delta x$ (work done + additional heat). Experiments find strong agreement with the Hill model, with $a = \alpha$ and $\mathrm{d}/\mathrm{d}t\ \Delta x = -v$.

Forces above isometric in lengthening can cause yielding (failure), experienced as 'pulling' the muscle.

Huxley's Crossbridge Dynamics Model: explains the T-v characteristic considering a sarcomere.



Assume 1) only active contractile components (neglect passive parallel elastic elements), 2) peak of the tension- length curve (all available cross-bridges will attach), 3) full tetanisation, 4) velocity and tension are constant, 5) each attached cross-bridge undergoes the full ATP reaction cycle.

(x: instantaneous displacement of an actin active site A from the equilibrium position of a myosin head M. $n(x, t)$: proportion (i.e. probability) of attached cross-bridges at position x and time t.)

- Isotonic cross-bridge kinetics: $\mathrm{d}n/\mathrm{d}t = -V\,\mathrm{d}n/\mathrm{d}x = (1 - n(x))f(x) - n(x)\,g(x) \rightarrow -V\,\mathrm{d}n/\mathrm{d}x + (f + g)n = f(x)$.
  ($V = -\mathrm{d}x/\mathrm{d}t$, f: attaching rate const: 'desire to contract', g: detaching rate const: 'desire to relax').
  Let $f(x) = \{x < 0: 0;\quad 0 < x < h: (f_1/h)x;\quad x > h: 0\}$ and $g(x) = \{x < 0: g_2;\quad x > 0: (g_1/h)x\}$ to fit data well.
  (h: bridge maximum reach)
  For $x > 0$, attached bridges act to shorten. For $x < 0$, attached bridges act to lengthen.

- Boundary conditions:
  if isotonic shortening ($V > 0$), $n(x) = 0$ for $x \geq h$; if isotonic lengthening ($V < 0$), $n(x) = 0$ for $x \leq 0$.
  Solve the DE for a piecewise continuous pdf $n(x)$.

To find the tension-velocity relationship $T(V)$, consider an energy balance on a half-sarcomere:

- Total work done: $T \times l = \frac{mAs}{2} \int_{-\infty}^{\infty} n(x, V)\, kx\, dx$ ($l >> h$: contraction distance, m: crossbridges per unit volume, A: area of muscle, s: sarcomere length, k: myosin filament spring constant).

Unlike Hill's model, Huxley's model additionally predicts accurate lengthening behaviour:

- Discontinuous slope: $\frac{\lim_{V \to 0^-} dT/dV}{\lim_{V \to 0^+} dT/dV} = \frac{f_1}{g_1} > 1$.
- Yielding: $\lim_{V \to -\infty} T(V) = T_0 \frac{f_1 + g_1}{g_1}$ where $T_0 = T(0)$ (isometric tension).
- Total rate of energy release (work + heat): $E' = \frac{mAse}{2l} \int_{-\infty}^{\infty} (1 - n(x))\, f(x)\, dx$ (e: energy per contraction site per cycle), matching Fenn's experiment with suitable choice of constants.
- Overall muscle efficiency: $\eta = \frac{TV}{E'}$ (mechanical power output / total energy)

At $V = V_{max}$, equal numbers of bridges pull in opposite directions, generating no overall force but consuming energy nonetheless. This corroborates Hill's model, which predicts zero efficiency at both $V = 0$ (isometric) and $V = V_{max}$, and by choosing the constants suitably, very close agreement can be found with Hill's experimental T-v data. By fitting a value for the energy released per ATP cycle, strong agreement with Fenn's heat experiments is found as well. Huxley's model additionally predicts the yielding phenomenon for moderate negative v (lengthening).

- At isometric conditions ($V = 0$), $E'_0 = \frac{mAsh}{4l}\left(\frac{f_1 g_1}{f_1 + g_1}\right)$ and $\eta = 0$.

Huxley-Simmons Kinetic Model: explains the transient effects of Hill's experiments.

The force response of a step-strain applied to Hill's T-v model does not fit experimental data, which sees a 'fast' partial development of force followed by a 'slow' continuation to the required force.

The **change in configuration** of the myosin motors represents the **'fast'** process.
The **crossbridge cycling** of the myosin motors represents the **'slow'** process.

The 'fast' rate constant (reciprocal time constant) is $r = ½\, r_0\, (1 + e^{-\alpha y})$ (y: step change in length). This can be derived by considering Arrhenius equations / Boltzmann statistics for the equilibrium between the (assumed) two starting and ending crossbridge states. The myosin head 'rocks' around, changing the force suddenly.

Further adjustments include modifying the myosin head potential energy wells.

17.3.14. Modelling Diffusion-Controlled Homeostasis

Mechanisms by which a cell can achieve chemical homeostasis include:

- **Passive carrier diffusion** of small hydrophilic molecules through porins (water-filled channel proteins in the cell membrane) or ion channels, or hydrophobic molecules through the cell membrane directly.
- Passive osmosis of water through aquaporins (water-selective channel proteins).
- **Facilitated diffusion** of sugars and amino acids by mobile protein mediators (carriers; ionophores) which pass through the cell membrane. This is also passive.
- **Active transport** via ATP-powered pumps e.g. Na^+/K^+-ATPase and Ca^{2+}-ATPase (Section 17.3.7).

Modelling Diffusion Across Membranes (L: membrane width)

1D Reaction-diffusion equation: $\frac{\partial c}{\partial t} = D\frac{\partial^2 c}{\partial x^2} + r(c)$ (c: concentration, r: reaction rate to form c)

At steady state, $\frac{\partial c}{\partial t} = 0$; boundary conditions $[c_0, c_L]$ with solution $c(x) = c_0 + (c_L - c_0)\frac{x}{L}$

By Fick's law, flux $J = -D\frac{\partial c}{\partial x} = \frac{D}{L}(c_0 - c_L)$, so the 'diffusion resistance' is $\frac{L}{D}$ (per unit area).

Passive Carrier Diffusion (Slab Reactor): substrate S, enzyme carrier E, bound complex ES

- Equilibrium: $E + S \underset{k_-}{\overset{k_+}{\rightleftharpoons}} ES$ with kinetics $r = k_+[\mathrm{E}][\mathrm{S}] - k_-[\mathrm{ES}]$.
- Conservation of enzyme amount: $[\mathrm{E}] + [\mathrm{ES}] = [\mathrm{E}]_{\text{total}} = \text{constant} \rightarrow \frac{\partial [E]}{\partial t} + \frac{\partial [ES]}{\partial t} = 0$.
- Diffusion: $\frac{\partial [S]}{\partial t} = D_S\frac{\partial^2 [S]}{\partial x^2} - r,\ \frac{\partial [E]}{\partial t} = D_E\frac{\partial^2 [E]}{\partial x^2} - r,\ \frac{\partial [ES]}{\partial t} = D_{ES}\frac{\partial^2 [ES]}{\partial x^2} + r.$

  with boundary conditions $\{x = 0\colon [\mathrm{S}] = s_0;\ x = L\colon [\mathrm{S}] = s_L\}$ and $\{(x = 0, x = L)\colon \frac{\partial [E]}{\partial x} = 0, \frac{\partial [ES]}{\partial x} = 0\}$.
- Total substrate flux: $J = J_S + J_{ES} = -\left(D_S\frac{\partial [S]}{\partial x} + D_{ES}\frac{\partial [ES]}{\partial x}\right)$ at steady state ($\frac{\partial [S]}{\partial t} = \frac{\partial [ES]}{\partial t} = \frac{\partial [E]}{\partial t} = 0$)
- Solution: $J = \frac{D_S}{L}(1 + \mu\rho)(s_0 - s_L)$ where $\rho = \frac{D_{ES}}{D_S}\frac{[E]_{total}}{K}$ and $\mu = \frac{K^2}{(s_0 + K)(s_L + K)}$ ($K = k_+ / k_-$)

  assuming $r \approx 0$, which is larger than unaided diffusion by an increase of $\mu\rho > 0$.

Facilitated Transport: carrier E conformation switches S between 'in' to 'out'. x is irrelevant.

- Equilibria: $E_{in} + S_{in} \underset{k_-}{\overset{k_+}{\rightleftharpoons}} ES_{in} \underset{k}{\overset{k}{\rightleftharpoons}} ES_{out} \underset{k_+}{\overset{k_-}{\rightleftharpoons}} E_{out} + S_{out};\quad E_{in} \underset{k}{\overset{k}{\rightleftharpoons}} E_{out}$ constant influx of S = J

  with kinetics $\frac{d[S_{in}]}{dt} = k_-[\mathrm{ES_{in}}] - k_+[\mathrm{E_{in}}][\mathrm{S_{in}}] - J$ and $\frac{d[E_{in}]}{dt} = k_-[\mathrm{ES_{in}}] + k[\mathrm{E_{out}}] - k_+[\mathrm{E_{in}}][\mathrm{S_{in}}] - k[\mathrm{E_{in}}]$, etc.
- Conservation of S and E: $[\mathrm{S_{in}}] + [\mathrm{S_{out}}] + [\mathrm{ES_{in}}] + [\mathrm{ES_{out}}] = [\mathrm{S}]_{\text{total}}$ and $[\mathrm{E_{in}}] + [\mathrm{E_{out}}] + [\mathrm{ES_{in}}] + [\mathrm{ES_{out}}] = [\mathrm{E}]_{\text{total}}$
- At steady state, $\frac{d[E_{in}]}{dt} = \frac{d[E_{out}]}{dt} = \frac{d[ES_{in}]}{dt} = \frac{d[ES_{out}]}{dt} = 0$
- Solution: $J = \frac{1}{2}K_d K_m k_+ [E]_{total}\frac{s_{out} - s_{in}}{(s_{in} + K_m + K_d)(s_{out} + K_m + K_d) - K_d^2}$ where $K_m = \frac{k_-}{k_+},\ K_d = \frac{k}{k_+}$.
- Symport vs antiport ionophore: two different substrates move in the same vs opposite directions.

17.3.15. Modelling Active Transport-Controlled Homeostasis

ATP-powered pumps use an ATPase to pump chemical species across a membrane against their concentration gradient by hydrolysing ATP into ADP and P_i. They are responsible for maintaining the difference in conditions between the cytosol (inside) and extracellular fluid (ECF, outside).

- P-type pump: cell membrane active transport of **ions**: H^+ (bacteria), Na^+/K^+ (animals) or Ca^{2+} (sarcoplasm)
- ABC transporter: cell membrane active transport of **small molecules** (amino acids, sugars, cholesterol etc)
- V-type pump: organelle (vacuole, lysosome) membrane active transport of **protons**: H^+ (eukaryotes)
- F-type ATP synthase: mitochondrial and chloroplast membrane, **passive proton-powered ATP synthase.**

Cytosolic conditions must be maintained independent of the ECF conditions. This includes pH 7.2, **high [K^+], low [Na^+], low [Ca^{2+}] and low [Cl^-]** (relative to the ECF e.g. blood).

Calcium Pumps: removes Ca^{2+} from the muscle cell cytosol to enable muscle relaxation

In muscles, Ca^{2+} is stored in the sarcoplasmic reticulum. When a muscle relaxes, P-type Ca^{2+} pumps (SERCAs) actively transport Ca^{2+} from the muscle cell cytosol (deactivating the contraction due to tropomyosin) back into storage.

The action of the Ca^{2+} pump involves 1) binding of cytosolic Ca^{2+} and ATP, 2) Mg^{2+}-facilitated ATP hydrolysis to ADP and P_i, 3) P_i transfer by phosphorylation of a distal Asp residue, 4) conformational change, 5) dissociation of Ca^{2+} due to lower affinity, now on the sarcoplasmic reticulum side, 6) hydrolysis of Asp-P_i bond and returning to original conformation.

Sodium-Potassium Pumps: removes 3 Na^+ from the cell and brings 2 K^+ into the cell, maintaining the membrane resting potential in neurons after an action potential floods the neuron with Na^+ (high V)

The Na^+/K^+ pump is a tetramer ($\alpha_2\beta_2$), whose action is similar to the calcium pump. It involves 1) binding of cytosolic Na^+ and ATP, 2) Mg^{2+}-facilitated ATP hydrolysis to ADP and P_i, 3) P_i transfer by phosphorylation of a distal Asp residue, 4) dissociation of Na^+ due to lower affinity, now on the outside, and binding of outside K^+, 5) hydrolysis of the Asp-P_i bond and returning to original conformation, 6) dissociation of K^+ due to lower affinity, now on the inside.

Na^+/K^+ pump equilibria (using simplified 1:1 Na:K binding model instead of 3:2): (E: pump enzyme)

$$\underbrace{Na^+_{in} + E \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ENa_{in} \underset{k_{-P}}{\overset{k_P}{\rightleftharpoons}} EPNa \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} Na^+_{out} + EP;}_{\text{Steps 1-4: pumping of } Na^+} \quad \underbrace{K^+_{out} + EP \underset{k_{-3}}{\overset{k_3}{\rightleftharpoons}} EKP_{out} \underset{k_{-4}}{\overset{k_4}{\rightleftharpoons}} P + EK_{in}; \quad EK_{in} \underset{k_{-5}}{\overset{k_5}{\rightleftharpoons}} K^+_{in} + E}_{\text{Steps 4-6: pumping of } K^+}$$

Steady-state Na^+/K^+ flux: $J = [E]_{total} \dfrac{[Na^+_{in}][K^+_{out}]K_1K_2 - [Na^+_{out}][K^+_{in}]K_{-1}K_{-2}[P_i]}{([K^+_{out}]K_2 + [K^+_{in}]K_{-2})K_n + ([Na^+_{in}]K_1 + [Na^+_{out}]K_{-1})K_k}$

($K_1 = k_1k_2k_P$; $K_{-1} = k_{-1}k_{-2}k_{-P}$; $K_2 = k_3k_4k_5$; $K_{-2} = k_{-3}k_{-4}k_{-5}$; $K_n = k_{-1}k_{-P} + k_2k_{-1} + k_2k_P$; $K_k = k_3k_4[\mathrm{P}_i] + k_{-3}k_5 + k_4k_5$.
k_P and k_{-P} are the forward and reverse rate constants for Mg^{2+}-facilitated ATP ⇌ ADP + P_i.)

Goldman-Hodgkin-Katz (GHK) equation for membrane potential: $V = \dfrac{RT}{F} \ln \dfrac{\sum_{\text{cations } i} p_i[i^+_{out}] + \sum_{\text{anions } j} p_j[i^-_{in}]}{\sum_{\text{cations } i} p_i[i^+_{in}] + \sum_{\text{anions } j} p_j[i^-_{out}]}$

17.3.16. Biofluid Mechanics

Biofluids (e.g. blood, urine, synovial fluid, sweat, cerebrospinal fluid (CSF), saliva, cytosol, interstitial fluid) are a chemically diverse range of water-based solutions. They can transport ions, dissolved gases, organic molecules and whole cells by advection in internal flows. For the theory of fluid mechanics and thermofluid dynamics, see Sections 7.1 and 7.2. Biofluids may be:

- Non-Newtonian: shear and/or time-dependent viscosity (Section 14.3.8)
- Viscoelastic or Bingham plastic: interstitial fluids provide biopolymers in tissues with viscosity in addition to elasticity (Section 6.2.9).
- Water-based colloids: proteins agglomerate in water (Section 14.3), providing oncotic pressure (colloid-osmotic pressure). Starling equation: $Q = L_p S(p_c - p_i - \sigma(\pi_p - \pi_g))$ (Q: volumetric flow rate, L_p: hydraulic conductivity, S: surface area, $p_{c/i}$: capillary/interstitial hydrostatic pressure, $\pi_{p/g}$: plasma protein/subglycocalyx oncotic pressure, σ: Staverman's reflection coefficient)
- Cell suspensions: transport properties affected when cell size is comparable to pipe diameter, e.g. blood viscosity increases for $d \sim 10$ μm (Fåhræus-Lindqvist effect / sigma effect)

Haemodynamics and Pulsatile blood flow: oscillating pressure gradients due to heart contractions
Consider laminar axisymmetric homogeneous incompressible Newtonian pipe flow.

If the pressure gradient $\frac{\partial p}{\partial x}$ oscillates as described by its complex Fourier series $\frac{\partial p}{\partial x}(t) = \sum_{n=0}^{\infty} P'_n e^{in\omega t}$:

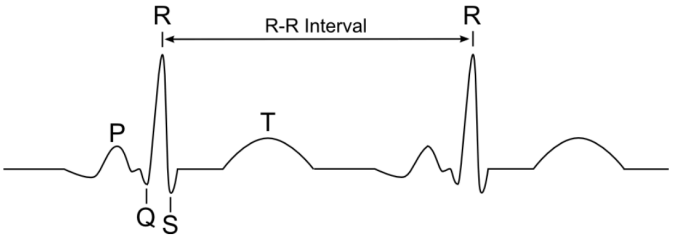
Analytical solution (velocity field): $u(r, t) = Re\left\{\sum_{n=0}^{\infty} \frac{i P'_n}{\rho n\omega}\left(1 - \frac{J_0(\alpha n^{1/2} i^{3/2} r/R)}{J_0(\alpha n^{1/2} i^{3/2})}\right) e^{in\omega t}\right\}$, for $0 \le r \le R$

(J_0: Bessel function (Section 1.7.9), $\alpha = R\sqrt{\frac{\omega\rho}{\mu}}$: Womersley number.)
The zeroth harmonic in the Fourier series of $u(t)$ corresponds to Poiseuille flow (Section 7.1.9).

The walls of the blood vessels are subject to stresses (like pressure vessels, Section

17.3.17. ECG (Electrocardiogram) Waveform

The contractions of the heart's atria and ventricles are controlled by electrical signals produced by the in the sinoatrial (S-A) node, which firstly move across the atria, then after a short delay of the signal in the atrioventricular (A-V) node, the signal moves across the ventricles causing them to contract. These electrical signals can be measured using an electrocardiograph, which forms an ECG (electrocardiogram: voltage-time plot).



- P wave - this is when the electrical signal is first produced in the sinoatrial (S-A) node and causes the atria to contract (depolarisation).
- QRS wave - this occurs around 0.2 s later and this is where the electrical signal leaves the atrioventricular (A-V) node and causes the ventricles to contract (depolarisation) (The relaxation/ repolarisation of this atria is masked by the large spike caused by the contraction of the ventricles).

T wave - this occurs around 0.2 s later again and is where the ventricles relax and are repolarised in preparation for a second heartbeat.

17.4. Biotechnology

17.4.1. Enzyme Classification

Enzyme database available here.

EC 1: oxidoreductases; $A^- + B \rightarrow A + B^-$; $AH + B \rightarrow A + BH$ or $A + O \rightarrow AO$; uses NAD(P)H / FAD cofactor

EC 1.1. dehydrogenases; acting on alcohols
EC 1.2. acting on aldehydes or oxo groups
EC 1.3. acting on CH-CH bonds
EC 1.4. acting on 1° amines
EC 1.5. acting on 2° amines
EC 1.6. acting on NADH or NADPH
EC 1.7. acting on other nitrogen-containing groups
EC 1.8. acting on sulfur-containing groups
EC 1.9. acting on a haem group
EC 1.10. acting on diphenols or quinones
EC 1.11. peroxidases; acting on peroxides
EC 1.12. hydrogenases; acting on molecular hydrogen
EC 1.13. dioxygenases; acting on two oxygen atoms
EC 1.14. monooxygenases; acting on paired donors, with O_2 addition
EC 1.15. acting on superoxides as acceptors
EC 1.16. oxidising metal ions
EC 1.17. acting on CH or CH_2 groups
EC 1.18. acting on iron-sulfur proteins as donors

EC 2: transferases; $AB + C \rightarrow AC + B$; can use pyridoxal 5'-phosphate (P5P) cofactor

EC 2.1. methyltransferases; moving 1 C groups
EC 2.2. transketolases/transaldolases; moving C=O
EC 2.3. acyltransferases; moving COO groups
EC 2.4. glycosyltransferases; moving sugars
EC 2.5. moving alkyl or aryl groups other than methyl
EC 2.6. transaminases; moving nitrogen groups
EC 2.7. phosphotransferases; moving P_i groups
EC 2.8. sulfur/sulfo/CoA transferases; moving sulfur groups
EC 2.9. selenotransferases; moving selenium groups
EC 2.10. molybdenum/tungstentransferases; moving Mo or W

EC 3: hydrolases; $AB + H_2O \rightarrow AOH + BH$; no cofactor required

EC 3.1. esterases/lipases; acting on ester bonds
EC 3.2. glycosylases; acting on sugars
EC 3.3. ether hydrolases; acting on ether bonds
EC 3.4. proteases; acting on peptide bonds
EC 3.5. acting on other C-N bonds
EC 3.6. acting on acid anhydrides
EC 3.7. acting on C-C bonds
EC 3.8. acting on halide bonds
EC 3.9. acting on P-N bonds
EC 3.10. acting on S-N bonds
EC 3.11. acting on C-P bonds
EC 3.12. acting on S-S bonds
EC 3.13. acting on C-S bonds

EC 4: lyases; X-A-B-Y → A=B + X-Y;

EC 4.1. carboxylase/aldolases; acting on C-C bonds
EC 4.2. hydratases; acting on C-O bonds
EC 4.3. acting on C-N bonds
EC 4.4. acting on C-S bonds
EC 4.5. acting on C-X (halide) bonds
EC 4.6. acting on P-O bonds
EC 4.7. acting on C-P bonds
EC 4.8. acting on N-O bonds

EC 5: isomerases; ABC → BCA

EC 5.1. epimerases/racemases; affecting R/S
EC 5.2. *cis/trans*-isomerases; affecting E/Z
EC 5.3. intramolecular oxidoreductases
EC 5.4. intramolecular transferases
EC 5.5. intramolecular lyases
EC 5.6. affecting macromolecular conformation

EC 6: ligases; $X + Y + ATP \rightarrow XY + ADP + P_i$

EC 6.1. forming C-O bonds
EC 6.2. forming C-S bonds
EC 6.3. forming C-N bonds
EC 6.4. forming C-C bonds
EC 6.5. forming phosphoric esters
EC 6.6. forming N-metal bonds
EC 6.7. forming N-N bonds

EC 7: translocases; $AX + B_{side\ 1} \| \rightarrow A + X + \| B_{side\ 2}$

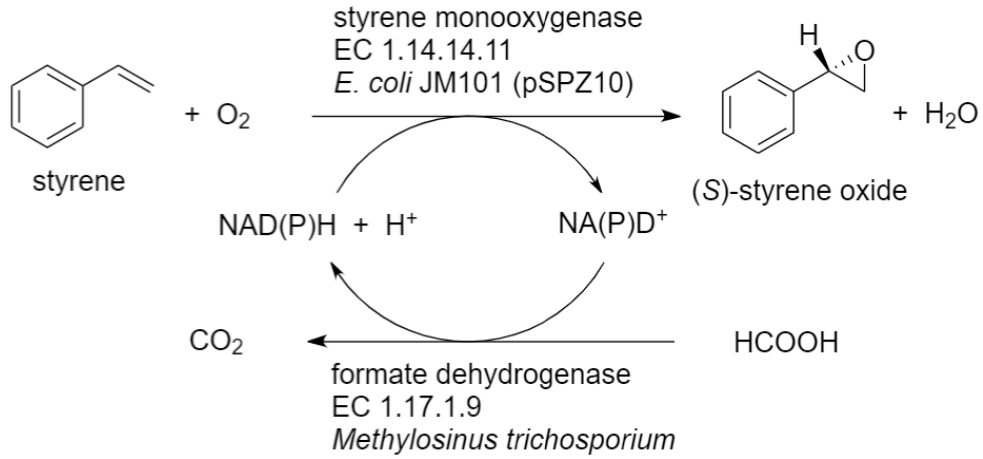
EC 7.1. acting on hydrons
EC 7.2. acting on inorganic cations
EC 7.3. acting on inorganic anions
EC 7.4. acting on amino acids/peptides
EC 7.5. acting on carbohydrates
EC 7.6. acting on nucleosides

17.4.2. Enzymatic Biotransformations

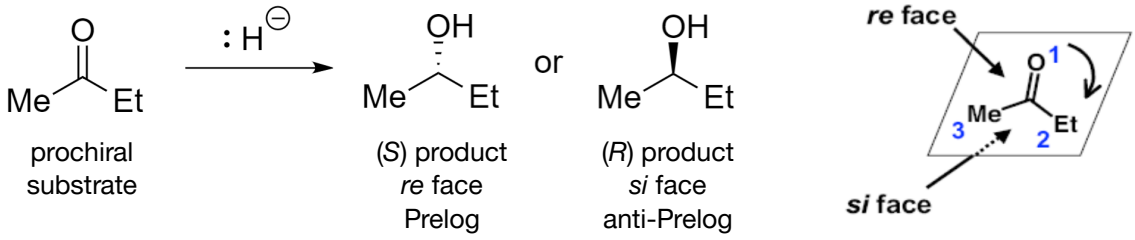
Common synthetic enzymatic reactions applied in biotechnology / biotransformations:

Enzyme class	Reactions (**not** complete; small molecules such as H_2O, O_2, H^+, NH_3 omitted)	Cofactors (for forward reaction)	Notes
Mono-oxygenase (MO) (uses O_2) EC 1	alkane ⇌ alcohol alkane ⇌ 1,2-diol (dioxygenase) arene ⇌ phenol sulfide ⇌ thioketone ketone ⇌ ester alkene ⇌ epoxide alkene ⇌ peroxide (dioxygenase, no NADH)	NAD(P)H ⇌ $NAD(P)^+$ (balance with $H^+ + O_2 \rightleftharpoons H_2O$)	High enantioselectivity and regioselectivity. Equilibrium favours the alcohol In cytochrome P450 monooxygenases, O_2 is activated by Fe^{2+} in haem. Iron must be added if using cell-free systems. FAD / FMN is an alternative cofactor which activates O_2 into O.
Dehydro-genase (DH) EC 1	alcohol ⇌ ketone aldehyde ⇌ carboxylic acid amine ⇌ ketone α-amino acid ⇌ α-oxocarboxylic acid alkane ⇌ alkene	$NAD(P)^+$ ⇌ NAD(P)H (balance with ~ ⇌ H^+)	Equilibrium favours the **alcohol** or **α-amino acid**. Can also use FAD ⇌ $FADH_2$. Equilibrium favours the carboxylic acid. Prelog's rule: the *S* enantiomer is favoured (e.g. R_1-C(···OH)-R_2 where $R_2 > R_1$)
Oxidase EC 1	amine ⇌ ketone α-amino acid ⇌ α-oxocarboxylic acid alcohol ⇌ ketone	None	Often $FADH_2$ dependent. Equilibrium favours carbonyl compounds, no net consumption of cofactors (easier to use in cell-free systems).
Reductase EC 1	alkene ⇌ alkane	None	Best enantioselectivity when EWGs present on C=C. Enantiomer depends on E / Z.
Transaminase EC 2	R_1-CH(NH_2)-COOH + R_2-COCOOH ⇌ R_2-CH(NH_2)-COOH + R_1-COCOOH (amine + ketone ⇌ ketone + amine)	None	Often dependent on pyridoxal 5'-phosphate (P5P).
Hydrolase (uses H_2O) EC 3	ester ⇌ acid + alcohol amide ⇌ acid + amine epoxide ⇌ 1,2-diol	None (balance with H_2O ⇌ ~)	Esterase, lipase: in organic solvent for reverse Amidase, protease Epoxide hydrolase
Lyase EC 4	alkene ⇌ amine alkene ⇌ alcohol α-oxocarboxylic acid ⇌ aldehyde nitrile ⇌ amide aldehyde ⇌ α-hydroxynitrile	None	Ammonia lyase Hydratase Decarboxylase: uses thiamine pyrophosphate Nitrile hydratase Hydroxynitrile lyase
Isomerase EC 5	glucose ⇌ fructose (aldohexose ⇌ ketohexose) *cis*-alkene ⇌ *trans*-alkene	None	Important in resolutions
Ligase EC 6	glutamate ⇌ glutamine (carboxylic acid ⇌ amide)	ATP or GTP	Glutamine synthetase: uses ATP and NH_4^+

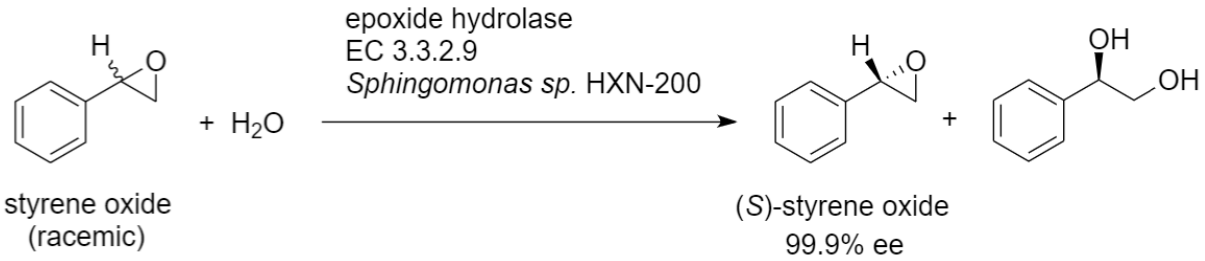
Example Biotransformation with Cofactor Recycling: enzymes which use opposing cofactors can be coupled together to exchange cofactors (catalytic cycle).



Prelog's Rule for Enantioselectivity in Asymmetric Catalysis: nucleophilic additions are carried out by delivering nucleophiles (e.g. hydride) onto the *re* face of planar prochiral substrates. (The *re* face is the front when there is more steric hindrance on the right-hand side i.e. Et > Me.) Enzymes can be classified as being enantioselective towards Prelog or anti-Prelog.



Kinetic Resolution: enzymes only react with one enantiomer, purifying the other.



Curves can be plotted of ee against % conversion for different enantiomeric ratios E.
An enzyme is said to be 'enantiospecific' if it leads to 100% ee (perfect enantioselectivity).

Enzyme Cascades: a reaction scheme with multiple enzymatic reactions.

- **Rate-limiting step:** the step setting the overall steady-state rate i.e. maximum flux control coefficient $C = \frac{dJ}{J} / \frac{d[E]}{[E]} = \frac{d(\ln J)}{d(\ln [E])}$. ($J$: product flux, $[E]$: concentration or activity)
- **Committed step:** a near-irreversible step which, once occurred, ensures the target will form.

17.4.3. Purification and Identification of Enzymes

Separation: remove the solution containing the enzyme from the cells.

1. **Cell lysis** (not required if enzyme is extracellular): osmotic shock / homogeniser / blender / grinding / French pressure cell press / ball mill / ultrasonication
2. **Separation:** filtration / centrifugation / two-phase aqueous liquid-liquid extraction, to produce the cell-free extract (CFE) in the supernatant (contains the proteins of interest).

Concentration: starting with the enzyme solution, increase its concentration.

1. **Precipitation:** salting out with $(NH_4)_2SO_4$ / adding organic solvent / precipitate at pH = pI / heat treatment / add polyvalent cations / add hydrophilic polymer e.g. PEG
2. **Stabilisation:** use buffer at optimal pH / work at low temperature / exclude O_2 to avoid free radical formation / avoid foaming to limit mechanical stress / use DTT or 2-mercaptoethanol to stabilise labile thiol groups / add BSA as a dilution standard / add protease inhibitors to prevent proteolysis

Chromatography: starting with a concentrated enzyme solution, separate the desired one while retaining biological activity.

A variety of selective chromatography methods are available (Section 16.4.7), based on isoelectric point (IEC), hydrophobicity (HIC), fusion tag affinity (IMAC) and size (SEC).

Automated UV/Vis spectrophotometry at a wavelength calibrated to the target enzyme can be used to record elution concentration in real time.

SEC / gel filtration is often used as a final polishing or desalting step, purifying the enzyme.

Quality Control and Characterisation: various assays can be used to investigate the enzyme:

- SDS-PAGE (Section 16.4.6): check for molecular mass and purity
- Electrospray mass spectrometry: determine precise molecular mass in kilodaltons (kDa)
- Isoelectric focussing (IEF): determine isoelectric point.
- Edman degradation: obtain the amino acid sequence of the protein.
- UV/Vis spectrophotometry: cofactor determination (holoenzyme or apoenzyme).
- X-ray crystallography / cryo-electron microscopy / 3D NMR: 3D structure determination
- Kinetic studies: calculate enzyme K_m, k_{cat}, K_i, proficiency, productivity (IU), etc.
- Chiral column chromatography with supercritical fluid: record enzyme enantioselectivity.
- HPLC with circular dichroism (CD) / optical rotation (OR) / refractive index (RIU) detection.
- Peptide mass fingerprinting and identification by e.g. MOWSE score.

17.4.4. Quantitative Analysis of Enzymatic Reactions

Enzyme Kinetics: rates of reaction of enzymatic biotransformations, accounting for saturation.

- **Typical enzyme-substrate (induced fit model) kinetics:** $E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_{\text{cat}}} E + P$
- **Michaelis-Menten equation:** $r = \frac{r_{max}[S]}{K_M + [S]}$ ($r_{max} = V_{max}$ = maximum reaction velocity)

  ($r = k_{\text{cat}}[\text{ES}]$; $r_{\text{max}} = k_{\text{cat}}[\text{E}]_{\text{tot}}$; $[\text{E}]_{\text{tot}} = [\text{E}] + [\text{ES}]$; $K_M = \frac{loss\ rate\ of\ ES}{gain\ rate\ of\ ES} = \frac{k_{-1} + k_{cat}}{k_1}$)
- **Lineweaver-Burk plot:** plot $\frac{1}{r}$ against $\frac{1}{[S]}$. x-intercept is $\frac{-1}{K_M}$; y-intercept is $\frac{1}{r_{max}}$; gradient is $\frac{K_M}{r_{max}}$.
- **Eadie-Hofstee plot:** plot r against $\frac{r}{[S]}$. x-intercept is $\frac{r_{max}}{K_M}$; y-intercept is r_{max}; gradient is $-K_M$.
- **Second-order rate constant (specificity constant):** $\frac{k_{cat}}{K_M}$
- **Enzyme proficiency** = $\frac{k_{cat}/K_M}{k_{non-cat}}$.

Enzyme productivity: measures the quantity of product formed.

- **Enzyme activity** [symbol: U or IU = μmol min^{-1}] = $\frac{quantity\ of\ product\ formed\ [\mu mol]}{time\ duration\ [min]}$
- **Specific activity** may be measured in IU g^{-1} (per unit mass of enzyme **or** the cell dry weight (g_{cdw}).)
- **Space-time yield:** tons of product per operating year per m^3 of bioreactor volume
  (1 ton yr^{-1} m^{-3} ≈ 2 mg min^{-1} L^{-1}; 1 operating year ≈ 500,000 min).

Enzyme Inhibition: there may be additional E + I ⇌ EI (competitive, equilibrium constant K_i) or ES + I ⇌ ESI (uncompetitive, equilibrium constant K_i') reactions.

Let $\alpha = 1 + \frac{[I]}{K_i}$ and $\alpha' = 1 + \frac{[I]}{K_i'}$. Then the kinetics of the formation of the product are:

	Type of inhibition	Apparent K_M	Apparent r_{max}
K_i only ($\alpha' = 1$)	**Competitive**	$K_M\,\alpha$	r_{max}
K_i' only ($\alpha = 1$)	**Uncompetitive**	K_M / α'	r_{max} / α'
$K_i = K_i'$ ($\alpha = \alpha'$)	**Non-competitive**	K_M	r_{max} / α'
$K_i \neq K_i'$ ($\alpha \neq \alpha'$)	**Mixed**	$K_M\,(\alpha / \alpha')$	r_{max} / α'

Enzyme Enantioselectivity: extent of specificity towards chiral substrates and/or products.

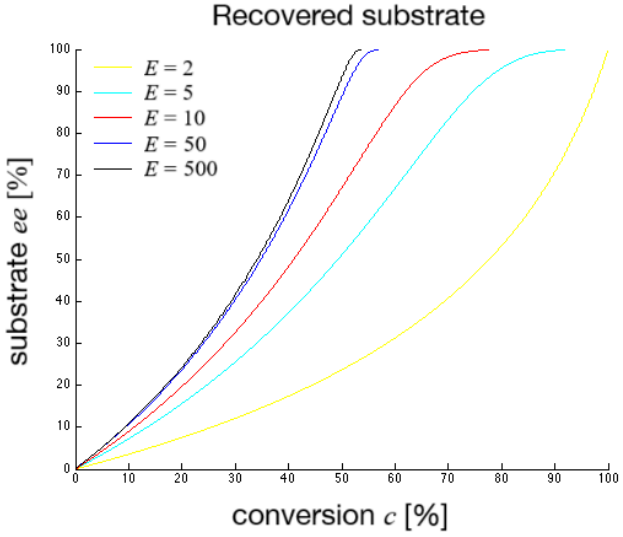
Kinetic resolution is used to convert a chiral racemic substrate A (A_R : A_S = 1 : 1) into a chiral product B ($ee_{product}$, yield c), leaving enantioenriched substrate A ($ee_{substrate}$, yield 1 - c).

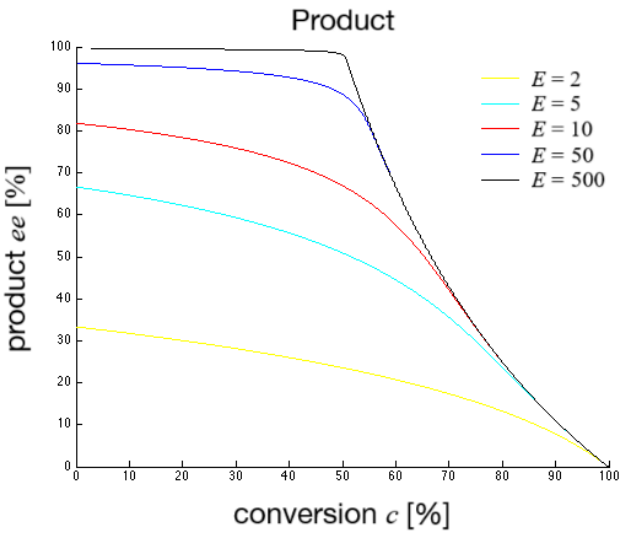
Enzyme E catalyses $A_R \rightarrow B_R$ with rate constant k_R and $A_S \rightarrow B_S$ with rate constant k_S.
If E is enantiospecific then one of k_R or k_S is zero and the corresponding [B] is zero for all t.

- **Enantiomeric excess (for R):** if the concentration ratio (or molar ratio) is $R : S$ then $ee = \frac{R-S}{R+S}$, as %.
- **Conversion (percentage yield of product):** $c = 1 - \frac{[A_R]+[A_S]}{[A_R]_{t=0}+[A_S]_{t=0}} = \frac{ee_{substrate}}{ee_{substrate}+ee_{product}}$.
- **Mole fractions:** $A_R = \frac{(1+ee_{substrate})(1-c)}{2}$, $A_S = \frac{(1-ee_{substrate})(1-c)}{2}$, $B_R = \frac{(1+ee_{product})c}{2}$, $B_S = \frac{(1+ee_{product})c}{2}$ such that $A_R + A_S = 1 - c$, $B_R + B_S = c$ and $A_R + A_S + B_R + B_S = 1$.
- **Enantiomeric ratio** E (for R) (aka selectivity factor s): $E = s = \frac{k_R}{k_S} = \frac{[k_{cat}/K_M]_R}{[k_{cat}/K_M]_S} = \frac{ln[[A_R]/[A_R]_{t=0}]}{ln[[A_S]/[A_S]_{t=0}]}$.

  In terms of ee and c, $E = \frac{ln[1-c(1+ee_{products})]}{ln[1-c(1-ee_{products})]} = \frac{ln[(1-c)(1-ee_{substrates})]}{ln[(1-c)(1+ee_{substrates})]}$.
- **Prochiral substrates:** for asymmetric transformations (A_R, A_S = A), $E = \frac{1+ee_{product}}{1-ee_{product}}$.

Kinetic resolutions typically have a maximum of $c < 50\%$, since one enantiomer of A cannot react. However, if the substrates can interconvert and racemise (epimerise) spontaneously ($A_R \rightleftharpoons A_S$), then greater product yields (up to 100%) can be attained (dynamic kinetic resolution), though there is a trade-off due to lower ee in the product: likewise, there is higher ee in the substrate but less of it in total. Characteristic curves for attainable ee against conversion (yield) for different enantiomeric ratios E are shown.





17.4.5. Protein Engineering (Protein Design and Directed Evolution)

Protein engineering is improving the properties of enzymes for practical applications. This includes enhancing thermostability, enantioselectivity, yield, proficiency or altering chemical reactivity to catalyse novel biotransformations.

Rational protein engineering: using the 3D structure of an enzyme (obtained by X-ray crystallography, or prediction from sequence with machine learning e.g. AlphaFold 3) to inform on its activity (structure-function relationships).

Directed evolution: starting with the cloned gene for a wild-type enzyme performing its natural function, perform random mutagenesis (amplifying with mutagenic PCR and testing for reactivity) by either:

- **Sequential mutagenesis:** select the single best enzyme and repeat.
- **Combinatorial mutagenesis:** select the top K best enzymes and cross them over with DNase I to get recombinant genes.
- **Saturation mutagenesis:** at each generation, choose a single site and generate all possible variants.

Number of possible variants by M simultaneous mutations over N amino acids = $19^M \times {}^N C_M$.

The methodologies can be combined by using **site-directed mutagenesis**, using the 3D structure to choose or guide the allowable sites of mutations.

Random Mutagenesis: introducing mutations into the genes coding for the enzyme

Error-prone PCR is used with conditions tailored to achieve only 1-2 bp changes per gene. These conditions include:

- Presence of UV light.
- Mutagens (e.g. MNNG (methylnitronitrosoguanidine), MMS (methyl methanesulfonate), 2-AP (2-aminopurine), AFB_1 (aflatoxin B_1))
- Host cell mutations in genes for DNA repair (e.g. uvrABC for UV lesion repair, phr for pyrimidine dimers, dnaQ for proofreading).
- Low $[Mg^{2+}]$ and high $[Mn^{2+}]$.

Point mutation notation: e.g. ‘D256A’ means the amino acid D (aspartic acid) at position 256 counting from the N-terminus was replaced with the amino acid A (alanine).

Thermostability: resilience to denaturation (melting) at high temperatures. Increased by having fewer heat-sensitive amino acids (Asn, Gln: deamination; Cys: disulfide bond cleavage; Asp: peptide bond hydrolysis), and a more densely packed hydrophobic interior.

Applications of Protein Engineering:

- **Subtilisin:** a protease used in bio-detergents, engineered to remove Ca^{2+} dependence.
- **Glucose isomerase:** used to make high-fructose corn syrup (HFCS) by the Maillard reaction, engineered to have a more suitable optimum pH and increased thermostability.
- Enhanced performance in biosensors, biofuel cells and antigen/antibody applications.

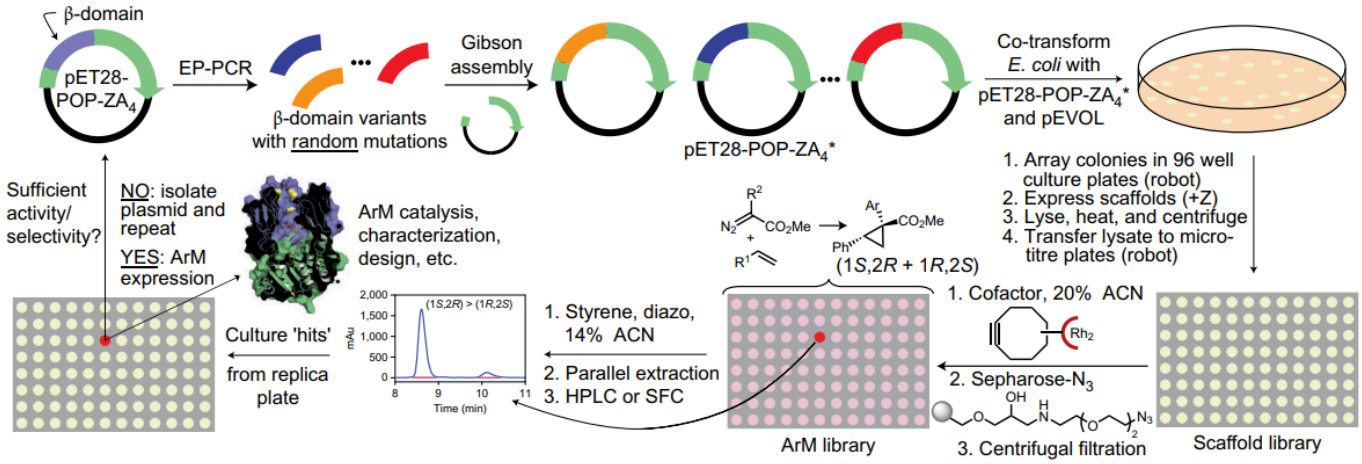
Application to Novel Biotransformations and Computational Drug Design

Depending on the active site geometry, slight modifications to the substrate may still allow a reaction to proceed, allowing for another degree of freedom to investigate when designing an enzyme for the purpose of carrying out a new reaction. This is known as **'enzyme promiscuity'**.

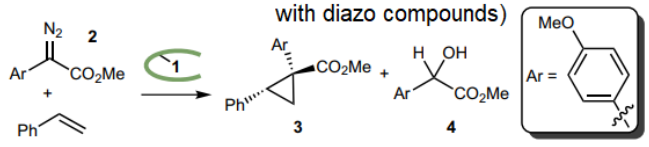
Example: Designing Artificial Metalloenzymes and Evolving by Random Mutagenesis

An artificial metalloenzyme (ArM) based on the POP (prolyl oligopeptidase, a natural serine protease) enzyme chain is designed. An azide group was genetically encoded into the protein chain by mutating an active site Ser residue, incorporated using bioorthogonal tRNA/synthetase pair (reassigns the amber stop codon), which then reacts with a metal-containing linker (**c**) to form the metalloenzyme (**b**) by a click reaction. The goal is to improve the enantioselectivity of this enzyme by referring to a model reaction (**a**) and using random mutagenesis (error-prone PCR) to vary the composition near the active site (the β-domain).

Procedure for carrying out mutagenesis:

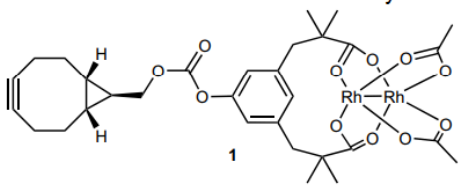


a: reaction to be carried out (olefin cyclopropanation with diazo compounds)

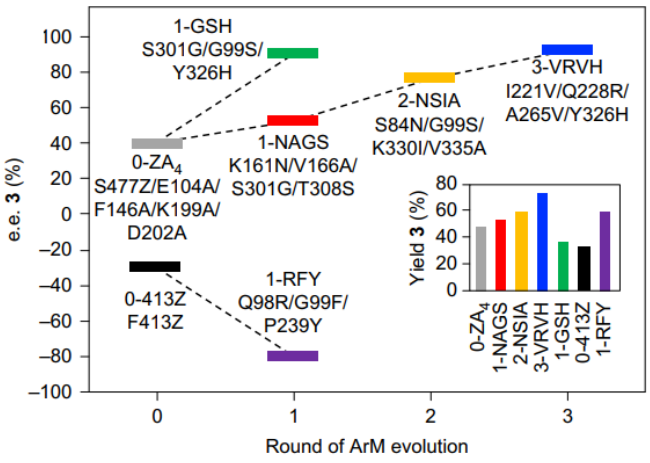


b: formation of the artificial metalloenzyme (ArM) by click chemistry

c: full structure of metal cofactor with linker to alkyne



Results: e.e. of product **3** after several rounds of ArM evolution



(Source: Yang *et al*, 2018, *Evolving artificial metalloenzymes via random mutagenesis*)

17.4.6. Industrial Biotransformations

The main classes of bioprocesses are:

- Fermentation of bulk chemicals (e.g. ethanol, citric acid, lactate)
- Bioproduction (e.g. indigo, penicillin G, vitamin B_{12}, mABs, enzymes)
- Bioconversion with whole cells (free or immobilised, for e.g. corticosteroids, acrylamide)
- Bioconversion with isolated enzymes (free or immobilised, for e.g. aspartate, HFCS, 6-APA. Materials used for immobilisation include Sepharose beads and magnetic nanobiocatalysts.)

Cofactor Recycling: cofactors are expensive, so recycling significantly helps reduce material costs
Common enzymes used for cofactor regeneration (coupled enzymes):

NAD^+ regeneration	**$NADP^+$ regeneration**	**NADH regeneration**	**NADPH regeneration**
Lactate dehydrogenase ($NADH + pyruvate + H^+ \rightarrow NAD^+ + lactate$)	Glutamic dehydrogenase ($NADPH + 2\text{-}oxoglutarate + NH_3 + H^+ \rightarrow NADP^+ + glutamate + H_2O$)	Alcohol dehydrogenase ($NAD^+ + ethanol \rightarrow NADH + acetaldehyde + H^+$)	Glucose dehydrogenase ($NADP^+ + glucose \rightarrow NADPH + gluconolactone + H^+$)
Glutamic dehydrogenase ($NADH + 2\text{-}oxoglutarate + NH_3 + H^+ \rightarrow NAD^+ + glutamate + H_2O$)	Alcohol dehydrogenase ($NADPH + acetone + H^+ \rightarrow NADP^+ + isopropanol$)	Formate dehydrogenase ($NAD^+ + formic\ acid \rightarrow NADH + CO_2 + H^+$)	
Alcohol dehydrogenase ($NADH + acetaldehyde + H^+ \rightarrow NAD^+ + ethanol$)		Glucose dehydrogenase ($NAD^+ + glucose \rightarrow NADH + gluconolactone + H^+$)	
		Hydrogenase ($NAD^+ + H_2 \rightarrow NADH + H^+$)	

Alternatively, the same enzyme can be used to carry out two reversible reactions **(coupled substrates)**.

Use of Organic Solvents

Hydrophilic solvents can be used to solubilise lipophilic substrates without deactivating the enzymes. Partially miscible solvents ($0 < \log P < 4$) are typically not used as they cause enzyme distortion, decreasing activity, while lipophilic solvents cause no distortion and retain high activity. For a table of partition coefficients (miscibility), see Section 13.4.7.

Klibanov rules: requirements for an enzyme to perform well in an organic solvent.

- Specific enzyme conformations can be attained by adding a specific ligand and lyophilising (freeze, remove water under vacuum, anhydrous extraction of the ligand). Adding water reforms the natural conformation; adding anhydrous solvent retains the conformation.
- Enzyme conformations are more ‘rigid’ in organic solvents and more flexible in water, so activities of optimal catalytic conformations are diminished in water and improved in solvent.
- Lyoprotectants (glycerol, sugars, salts) can be used to block these changes if needed.

Organic solvents are typically used to sterilise the biotransformation environment, although some strains of bacteria can survive in miscible organic solvents, by releasing emulsifier (LPS) exosomes from their cell membrane to dissolve the organic phase into the aqueous phase, away from the bacterium.

Industrial Biotransformations in Microbial Cells

Bacteria can be gene edited to express specific enzymes by splicing in a plasmid coding for the required protein, or by deactivating proteins which inhibit the desired activity. Bacteria benefit from having 'built-in' cofactor recycling as part of their natural metabolism. However, product extraction is complicated by additional purification/separation steps.

Bacteria can also be immobilised on an adhesive surface for easier re-use, but with lower activity.

Bacterial growth phases: lag (adjustment to conditions) → exponential growth → growth plateau (nutrients are depleted) → death (accumulation of toxic byproducts).

Growth in the presence of a limiting substrate of concentration [*S*]:

$$\mu = \mu_{max} \frac{[S]}{K_S + [S]}$$ (proportional growth rate: such that $\frac{dN}{dt} = \mu N$) (Monod equation)

Enzyme synthesis rate, r_e [IU L^{-1} hr^{-1}] = Specific synthesis rate, q_e [IU g_{cell}^{-1} hr^{-1}] × Biomass, c_x [g_{cell} L^{-1}]

Enzyme productivity, P_e [IU L^{-1} hr^{-1}] = $\frac{1}{T}\int_0^{t_e} q_e(t)\, c_x(t)\, dt$ (T = fermentation time t_e + turnaround time)

Product Recovery

The stages of recovering product from a bioreactor with microbial cells include:

Disruption of microbe cell wall: release the cytoplasm, generating cell debris.

- Chemical methods (can inactivate/denature the enzyme).
  Can use alkali, detergents or organic solvents.
- Enzymatic methods (specific, mild). Can use lysozymes (for Gram-positive bacteria), lysozymes with EDTA (for Gram-negative bacteria) or β-glucanases/mannanase/chitinase (for yeast and fungi).
- Physical methods. Can use freeze-thaw, sonicator, pressure bomb, wet milling, French press/homogeniser.

Removal of cells or other particles: cake filtration / cross-flow filtration / centrifugation.

Primary isolation of dissolved compounds: solvent extraction / sorption with ion exchanger or hydrophobic resin / adsorption chromatography / precipitation / membrane filtration (ultrafiltration, composite membranes) / reverse osmosis / pervaporation (permeation then evaporation) / perstraction (permeation then solvent extraction (dialysis))

Purification and final product isolation: chromatography, crystallisation, concentration, centrifugation and drying, distillation.

Bioprocess Economics

Costs can be broken down as being associated with 1) running the fermentor, 2) product recovery, 3) raw materials for cell growth, 4) substrate for reaction.

Approximate correlations for startup and operating costs are:

- Fermentor costs [US $] $\approx 40,000 \times V^{0.6}$ (includes cell recovery) (V: fermentor volume [m^3])
- Cost per unit product [US $ kg^{-1}] = $\frac{31250}{M_w\,[g\,mol^{-1}] \times P_e\,[IU\,L^{-1}\,min^{-1}]}$ where $P_e = \frac{t_{cycle}}{t_{turnaround}} \times q_e\, c_x$

Raw materials: carbon sources (e.g. whey: $ 0.1 kg^{-1}, molasse: $ 0.2 kg^{-1}, glucose: $ 0.38 kg^{-1}), nitrogen sources (e.g. ammonium sulfate: $ 1.8 kg^{-1}), phosphate/magnesium/trace sources (e.g. as salts: approx $ 0.5 kg_{cdw}^{-1})

- Source costs [$ kg_{cdw}^{-1}] = $\frac{cost\ of\ source\ [\$\,kg_{source}^{-1}]}{yield\ [kg_{cdw}\,kg_{source}^{-1}]}$, where yield = $\frac{mass\ fraction\ of\ atoms\ in\ cells}{mass\ fraction\ of\ atoms\ in\ source}$
- Cost per unit product [US $ kg^{-1}] = $\frac{source\ costs\ [US\,\$\,kg_{cdw}^{-1}]}{productivity\ [kg\,kg_{cdw}^{-1}]} = \frac{source\ costs\ [US\,\$\,kg_{cdw}^{-1}] \times 10^6}{activity\ [IU\,g_{cdw}^{-1}] \times M_w\,[g\,mol^{-1}] \times t_{turnaround}\,[min]}$
- Substrate costs [$ kg^{-1}] = $\frac{substrate\ cost\ [\$\,kg_{substrate}^{-1}] \times M_w\,[g\,mol^{-1}]_{substrate}}{yield\ [mol\,mol_{substrate}^{-1}] \times M_w\,[g\,mol^{-1}]_{product}}$
- Product recovery costs [$ kg^{-1}] $\approx k \times (product\ concentration\ [M])^{-0.75}$ (k: method dependent)

17.4.7. Bioremediation and Genetic Engineering of Bacteria

Bioremediation involves the use of microorganisms (usually bacteria or fungi) to remove pollutants from, or provide nutrients to, a region of a biome in the natural environment.

Genetic engineering can be used to deploy GMOs with the genes coding for proteins which can degrade harmful compounds. However, there are concerns about the possibility of horizontal gene transfer with other organisms.

Bacteria can accept exogenous plasmids. If a plasmid containing a gene for a desired protein is synthesised (a recombinant plasmid), it will be expressed and secreted by the bacteria. A wide range of synthetically useful proteins can be produced directly in this way (e.g. enzymes: Section 17.4.6, but also peptide hormones e.g. insulin to be used as medicine for diabetes). Enzymes can also be designed and optimised to carry out a new biotransformation on native metabolites into a useful product (Section 17.4.5) e.g. monomers for bioplastics.

17.4.8. Reproductive Cloning of Organisms

Somatic Cell Nuclear Transfer (SCNT) can be used to clone a living organism. A somatic cell is extracted and its nucleus is isolated. An egg cell is taken and its nucleus is removed and replaced with the somatic nucleus. With an electric shock, the egg now behaves as if it were newly fertilised, and develops into an embryo. The embryo is transplanted into a surrogate mother and carried to term, where it is born as a clone of the original organism.

In 1996, 'Dolly' became the first mammal (a sheep) to be created as a cloned organism. Dolly lived for 7 years, appearing to age faster than usual, as well as having various health complications. A variety of animals, including cattle and primates, have been cloned with SCNT since, with increasingly successful results.

This cloning procedure could hypothetically be applied to humans, but is currently far too uncertain and unethical for such an experiment to be performed in the foreseeable future.

17.4.9. Polymerase Chain Reaction (PCR)

PCR is a synthetic technique to amplify a given sequence of DNA. It can be considered *in vitro* DNA replication, using a polymerase enzyme (Taq polymerase, a thermostable polymerase I), 3' end complementary primer sequences, and a source of nucleotides (supplied as deoxynucleoside triphosphates, dNTPs).

Thermal cycling is used to carry out the following steps repeatedly:

1. Denaturation at 95 °C: the DNA is unravelled by melting the hydrogen bonds.
2. Annealing at 55 °C: primers anneal to the 3' ends of each open strand.
3. Extension at 70 °C: Taq polymerase builds the complementary strand from the template.

A pH buffer is used for optimal activity. A solution of Mg^{2+} (or other divalent cations) is used to provide the metal cofactors for the polymerase enzymes. If mutations are desired (PCR-mediated DNA mutagenesis), Mn^{2+} can be used instead, which displaces Mg^{2+} polymerase cofactors and are less specific to nucleotide pairing, increasing the rate of errors.

PCR can be used to test for the presence of specific viruses. Once the genome of the virus has been sequenced, its primers can be developed. For RNA viruses, the RNA must be converted to cDNA (with reverse transcriptase) first. Fluorescent DNA probes (nucleic acid stains e.g. SYBR green) are used to indicate the presence of the viral cDNA. This is the basis for the COVID-19 PCR test.

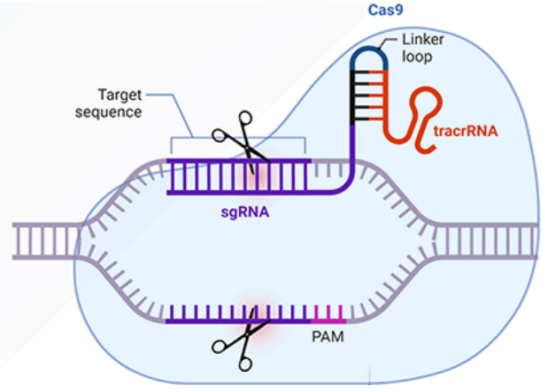
17.4.10. Gene Editing with CRISPR

CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) is a naturally-occurring component of the adaptive immune systems of prokaryotes (bacteria and archaea / extremophiles) to defend against bacteriophages and plasmids. It can be employed to edit genomes with high precision.

Natural CRISPR Mechanism (*in vivo*: in a living organism)

When viral DNA is injected into the bacteria, it can be inserted between palindromic sequences in the CRISPR locus to form a CRISPR array. This DNA can be transcribed into pre-crRNA. Effector complexes are Cas9 (nuclease) proteins bound to tracrRNA strands. If the complex encounters viral DNA complementary to its crRNA, it will cleave the DNA at the end of the protospacer adjacent motif (PAM) sequence, preventing its transcription and therefore preventing viral replication.

Artificially Engineered CRISPR Mechanism (initially *in vitro*: in a laboratory)



crRNA and tracrRNA can be linked to form a single molecule, sgRNA. When combined with Cas9, this system will read DNA until the pre-specified complementary strand is found. It will cleave the DNA after the PAM sequence.

For **gene knock-out**, the non-homologous end-joining (NHEJ) repair pathway rejoins the two sections of DNA. For **gene knock-in**, the homology-directed repair (HDR) repair pathway fills the gap by replicating from a template DNA. This can be used for transfection.

Gene therapy: the gene-edited cells can then be used appropriately (*in vitro*: e.g. CAR T-cell therapy for anticancer antibody production, viral vector vaccine production, etc).

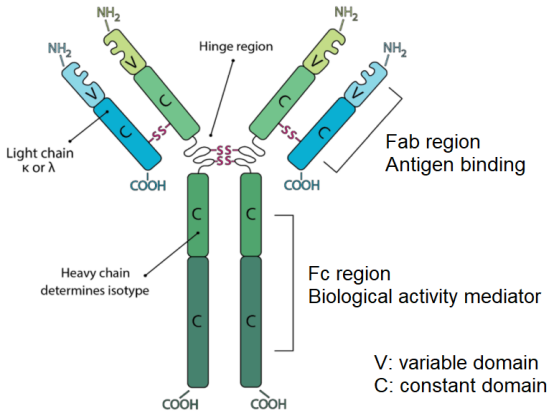
In vivo usage of CRISPR/Cas9 is currently underway (as of 2021) which could edit the genomes of live adult cells in an organism. Such a usage is limited to adult somatic cells: extension to gene editing of human embryos remains extremely controversial and unethical. In 2019, a Chinese scientist was imprisoned for conducting such an experiment in which two human twins were equipped with a gene intended to protect against HIV. As of 2023, they are alive and remain healthy, although long-term effects are highly uncertain as their natural development continues.

17.4.11. Monoclonal Antibodies

Antibodies are immunoglobulin proteins which are freely circulating in the blood plasma. They bind to matching antigens on cells, which are then recognised by macrophages. Polyclonal antibodies (pAbs) can bind to multiple epitopes (regions; antigenic determinants) of the same type of antigen. Monoclonal antibodies (mAbs) are specific to a single epitope of a single antigen.

IgG: monomer, activates phagocytosis; IgM: pentamer, activates complement

Manufacture of Murine Antibodies



- An animal (e.g. mouse) is vaccinated with the target antigen
- The immune response generates pAbs for this antigen
- B cell antiserum is extracted from the spleen
- Myelomas (cancerous white blood cells) are grown in 8-azaguanine
- The B cells are fused with the myelomas in a PEG medium
- The resulting hybridomas are immortal, divide fast, and generate pAbs.

The cells are cultured in a HAT medium. Aminopterin (A) inhibits the folate pathway of new DNA synthesis while hypoxanthine (H) and thymidine (T) provide raw materials for the salvage pathway (in order to do this, they must have the HGPRT gene, which B cells have but myelomas do not). The result is purification of only hybridomas (B cells die naturally over time).

To obtain mAbs from the pAb product, isolate individual cells by 'infinite dilution' (1 cell per well). Screen using ELISA (Section 16.6.17) with target antigen, then scale up for mass production.

Second Generation Antibodies: Murine (mouse-derived, suffix: -momab) antibodies are typically not used when immunocompatibility in humans is important. By making the antibody structure more similar to those found in humans, antibodies can be administered more safely.

- Chimeric: the constant domain of the antibody is human. Suffix: -ximab.
- Humanised: the constant domain, and most of the variable domain, is human. Suffix: -zumab.
- Human: the whole antibody is human. Suffix: -umab.

Production: *in vitro* immunisation / SCID mice / phage display / transgenic mice.

Common Applications of mAbs:

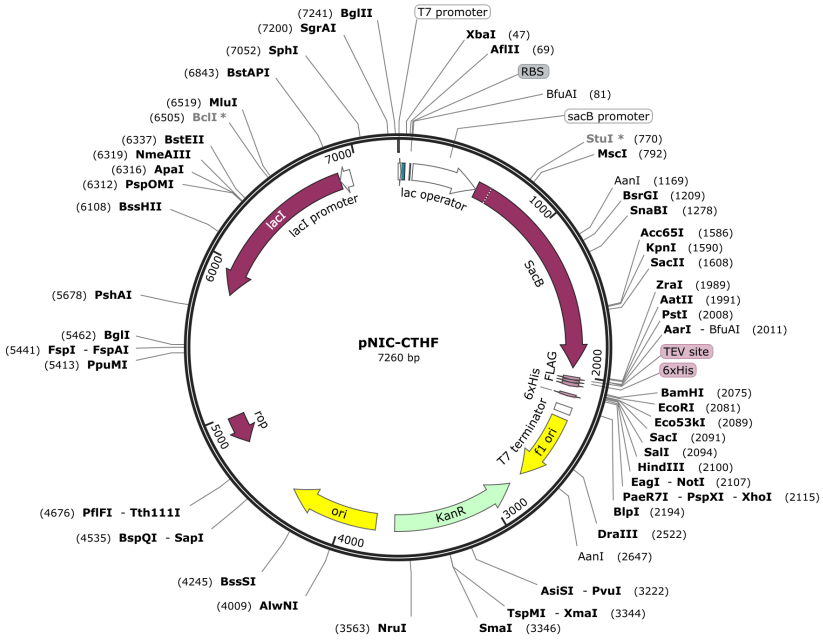
- Pregnancy test kits: based on an ELISA test strip for hCG hormone antigen with fluorescent markers.
- Lateral flow tests: same principle, testing for presence of viral antigens
- Allergen detection in foods: tests for the presence of common allergens e.g. gluten protein.
- Radiopharmaceuticals: mAbs for cancer cell antigens bonded by a linker to a chelated radioactive atom (e.g. ^{131}I, ^{90}Y, ^{177}Lu), which kills the targeted cell by emitting β radiation
- Bioassays: various methods of identifying, separating and/or characterising proteins e.g. ELISA, Western blotting, flow cytometry, immunohistochemistry, affinity chromatography

17.4.12. DNA Sequencing, Cloning and Synthesis Technology

DNA Synthesis

DNA Cloning: plasmids are used to carry genes, amplify and express in bacteria

Artificial plasmids are available to molecular biologists for cloning purposes. For example, pNIC-CTHF is a bacterial plasmid (expression vector) designed to inducibly express genes in *E. coli*. It contains:



***T7* promoter:** the main driver of gene expression.
***Lac* operator:** typically binds to the repressor *lacI* (encoded by *lacI* regulator gene) preventing transcription, but when IPTG (isopropyl β-D-1-thiogalactopyranoside), which mimics allolactose, is added to the culture, it deactivates the repressor and permits transcription (inductive gene expression).
RBS: ribosome binding site, for protein production.
***SacB* gene:** confers sensitivity to sucrose, such that expression in 5% sucrose prevents growth (negative selection marker).
***XbaI*, *PstI* and *EcoRI*:** restriction enzyme sites.
***TEV* site:** protease recognition site.
***6xHis* and *FLAG*:** protein tags useful for purification and identification.
***KanR*:** gene for kanamycin antibiotic resistance.

The plasmid pNIC-CTHF can be linearised with the restriction enzymes *XbaI* and *PstI*, resulting in excision of *SacB*, to be replaced with a suitable DNA fragment containing the new gene. The new DNA fragment must contain a ribosome-binding site downstream of *XbaI*, a start codon for the open reading frame, and ending with a *PstI* site. It is common to obtain the DNA fragments via PCR amplification. The primers used in the PCR must be designed carefully to ensure that the resulting DNA is in the same phase (reading frame) as the adjacent sequences in the plasmid, using additional sequences if needed.

Example Workflow using Plasmid Engineering: testing of a modern protein subunit vaccine

Novel gene design and synthesis:

- Obtain the DNA sequence of a spike protein on the virus using next-generation sequencing.
- To the designed sequence, add a His-tag, RBS and TEV protease recognition site. Optimise the codons for the spike protein if necessary.
- Synthesise the designed DNA sequence using e.g. phosphoramidite method, Klenow/Gibson assembly.
- Design and synthesise PCR primers for the DNA sequence, containing *XbaI* and *PstI* restriction sites.
- Carry out PCR on the template DNA using the designed PCR primers, amplifying the DNA. Purify the DNA.

Cloning process:

- Obtain bacterial plasmid pNIC-CTHF, suitable for inductive gene expression, and linearise using restriction enzymes *XbaI* and *PstI*, resulting in excision of the *SacB* gene from the plasmid.
- Use DNA ligase to splice the PCR-amplified DNA into the plasmid, forming the recombinant plasmid.
- Transform *E. Coli* with the plasmids and culture on a multiwell microtiter plate containing kanamycin antibiotic and 5% sucrose to kill cells which did not take up the plasmid, or took up an unmodified/self-ligated plasmid (no spike protein gene), respectively.
- Inoculate a culture of the transformed cells in growth medium (e.g. lysogeny broth) and culture to the mid-log phase of cell growth to produce recombinant clones.
- Induce protein expression by adding IPTG and continue to culture for some time period.

Product extraction:

- Extract the protein by cell lysis and centrifugation, and isolate by affinity chromatography (IMAC: nickel-agarose column to bind with the His-tagged protein, then elution with imidazole).
- Digest the His-tags with TEV protease solution to remove the tags, yielding the free protein.
- Purify the protein product by gel filtration chromatography.
- Analyse the protein for purity by 2D gel electrophoresis (using isoelectric focussing and SDS-PAGE), and Western blotting with anti-FLAG antibodies.
- For the purposes of protein subunit vaccine development, suspend the proteins (either freely or in liposomes) in a suitable buffer and formulate with adjuvants.
- Store the solution, typically at -80 °C for long-term storage or distribution.

17.4.13. Vaccines

Vaccines can ‘simulate’ a viral infection without a real infection, priming the body’s immune system and allowing it to respond quickly, significantly increasing the chance of recovery. They work by causing the characteristic biomarkers belonging to the pathogen responsible for the disease to be presented on host cells, allowing for the immune system to recognise them and proceed with the innate immune response. Once immunity is achieved, subsequent exposure with the same pathogen will immediately stimulate a strong immune reaction, preventing infection.

Types of vaccines: different ways of inducing antigen presentation in the host cells

- **Live attenuated virus:** a live virus which has acquired mutations that severely reduce its toxicity.
- **Inactivated virus:** a virus that has been ‘killed’ using chemicals/heat to break the virus into pieces.
- **Toxoid:** used for anti-bacterial vaccines, containing inactivated toxins released by the bacteria.
- **Protein subunit:** a component of the virus (antigen; spike protein) or bacteria (cell wall polysaccharide)
- **Viral vector vaccine:** a replication-deficient adenovirus with DNA coding for the target antigen.
- **mRNA vaccine:** lipid nanoparticle containing mRNA coding for the antigen.
- **Dendritic cell vaccine:** personalised dendritic cells already presenting neoantigens, for cancers.

Adjuvants: compounds that stimulate the immune response

For non-live vaccines, adjuvants accompany the vaccine load in small quantities to increase the sensitivity of the immune system. These work by (directly or indirectly) increasing the number of pro-inflammatory cytokines (interleukins) in the blood plasma, which attract immune cells. Common adjuvants include aluminium hydroxide nanoparticles, squalene emulsion, emulsifiers (e.g. polysorbate 80) or Toll-like receptor (TLR) agonists. Some adjuvants also promote the initial endocytosis of the vaccine payload into host cells by increasing the affinity for cell receptors.

Vaccines may be monovalent (targets only one serotype) or multivalent (targeting multiple serotypes). The serotype targeted is dependent on the identity of the surface antigens.

17.4.15. Gene Transfection

HEK-293 cells (human embryonic kidney cells) are an immortalised cell line commonly used as a common eukaryotic target for gene transfection.

17.4.16. Stem Cell Biotechnology

Stem cell potential to undergo further differentiation: Totipotent (e.g. zygote) > pluripotent (e.g. embryonic stem cells (ESCs)) > multipotent (e.g. adult stem cells: hematopoietic stem cells (HSCs), mesenchymal stem cells (MSCs)).

Induced pluripotent cells (iPSCs): treatment with the Yamanaka reprogramming factors (Oct4, Sox2, Klf4, C-Myc) leads to reversion of cells to the pluripotent state. Conditions can then be tuned to favour differentiation to a desired cell type.

Tissue engineering (Section 17.4.17) with stem cells has several applications in regenerative and personalised medicine, as well as in sustainable agriculture:

- Osteoblasts (bone stem cells) → bone regeneration.
- Myoblasts (muscle stem cells) → cultured meat (“lab-grown meat”).

Stem Cell Culture Media

A culture medium must maintain the self-renewal and pluripotency of stem cells, as well as the growth of the cells themselves. Modern growth media are xeno-free (no animal components; human-derived only) chemically-defined medium with recombinant growth factors and small molecules.

17.5. Astrobiology and Abiogenesis

17.5.1. Obsolete Hypotheses for the Origin of Life

Creationism (religious/mythological ideas) has historically been a popular explanation for the origin of life, but the methodological naturalism required of the scientific method meant that a scientific explanation should exist. **Abiogenesis** is the general term for the natural process(es) by which life developed from non-living matter. Obsolete concepts relating to abiogenesis include:

Spontaneous generation: small organisms such as larvae of insects and worms could be produced from decaying flesh of larger organisms.

Spontaneous generation (in some form) had remained a prevailing explanation for life since ancient times, but was eventually disproven in the 19th century following experiments:

- **Francesco Redi's experiment (1665):** meat was left to rot in a gauze-sealed jar. Redi observed flies would be attracted to the meat, but laid their eggs to hatch into maggots on the gauze, showing they were not generated by the meat itself.
- **Louis Pasteur's experiment (1859):** a meat broth was boiled in a swan-neck flask, separating the broth from the outside, and the broth remained sterile. When the neck from the flask was removed, allowing direct air access, microbial growth occurred in the broth.
- In the 1860s, Pasteur proposed the **law of biogenesis** (life can only arise from pre-existing life), though this does not consider the origin of life itself.

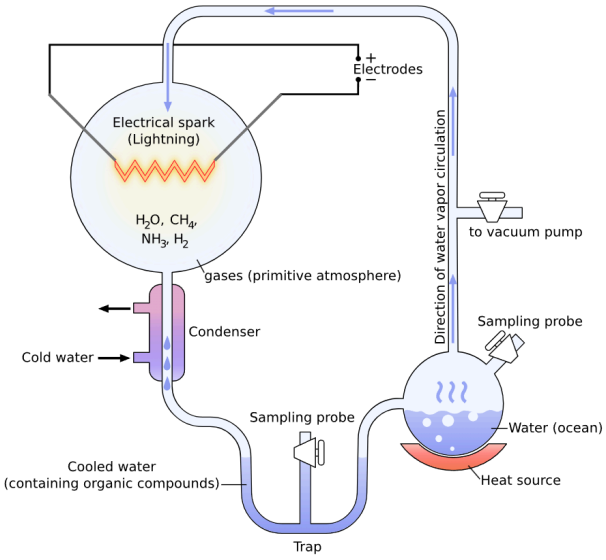
Vitalism: the concept that life is fundamentally distinct from non-life, since life has a special 'vital force', which is the only way to convert 'inorganic' chemicals into 'organic' chemicals.

- **Wöhler's synthesis of urea (1828):** a synthesis of urea ($CO(NH_2)_2$) from ammonium cyanate (NH_4OCN), showing that organic chemistry can be accessed from inorganic materials.
- **Kolbe's reaction (1845):** a four-step synthesis of acetic acid (CH_3COOH) from carbon disulfide (CS_2), which was effective in disproving vitalism.

17.5.2. Oparin-Haldane Hypothesis

The leading scientific concepts for abiogenesis involve progressive formation of simple cells from a suitably energised 'prebiotic soup' of chemicals, over a very long period of time for complexification. This possibility was also mentioned by Darwin in 1871 (the 'warm little pond').

Miller-Urey experiment (1952): attempts to simulate a prebiotic atmosphere to generate amino acids



A 2:2:1 mix of methane, ammonia and hydrogen was subjected to an electrical arc (representing lightning), leading to a range of chemical reactions in the water.

The resulting 'prebiotic soup' in the water contained intermediates such as HCN, CH_3OH (methanol), CH_2O (formaldehyde), C_2H_2 (acetylene), HCCN (cyanoacetylene), among others. Important organic products included α-hydroxy acids, formed by cyanohydrin reactions, and several racemic amino acids (e.g. glycine, α-alanine and β-alanine), formed by Strecker syntheses.

The initial conditions of the Miller-Urey experiment have been varied in subsequent experiments to more faithfully capture the composition of the prebiotic atmosphere, which was weakly reducing (more CO_2 and H_2S, less CH_4 and H_2), but likely had short periods of strongly reducing conditions following impact events. The results show a similar range of organic products.

17.5.3. Astrochemistry

Many organic molecules have been observed in space, including alcohols, aldehydes, cyanides, amino acids, ribose, glucose, nucleobases, lipids and simple polypeptides (e.g. hemoglycin) on meteorites (most commonly carbonaceous chondrites, e.g. Murchison meteorite, Allende meteorite) as well as many asteroids, dwarf planets and moons (e.g. Ceres, Titan, Ryugu) and Kuiper belt objects (e.g. Arrokoth). Some of these materials, as well as water were likely delivered to the early Earth by asteroid impacts.

Miller-Urey-type astrochemical reactions occur regularly throughout space, using solar and galactic cosmic rays as the external energy source rather than electricity.

Irradiation of simple carbon-containing molecules (e.g. CO_2, CH_4, C_2H_6) by cosmic rays, in combination with N_2, NH_3 and H_2O, can produce amorphous polymers called thiolins. These are responsible for the reddish hue of many icy bodies.

17.5.4. Driving Forces for Complex Chemistry on the Prebiotic Earth

The early Earth was highly energetic, with a variety of sources allowing the possibility of complex chemical dynamics.

Potential setting	Energy sources	Chemical sources	Concentration mechanisms	Reaction enhancement mechanisms
Surface ocean, lakes, ponds	Lightning, UV radiation, redox gradients	Meteorite deposition, Miller-Urey-type processes	Mineral surfaces, evaporation/ rehydration	Clays at the bottom of lakes
Sea ice	Lightning, UV radiation, redox gradients	Uptake from seawater, atmospheric deposition	Brine pockets	Dehydration
Hydrothermal vents	Redox gradients (between H_2 and CO_2)	Serpentinisation in peridotite-hosted vents	Carbonate or sulfide mineral pores	Metal sulfides, mineral surfaces, heat
Beaches	Lightning, redox gradients, radioactive radiation from accumulated uraninite grains, UV radiation	Meteorite input, Miller-Urey-type processes, *in situ* production by nuclear fission	Evaporation/ rehydration or pore spaces	Dehydration, nuclear fission by radioactive uraninite

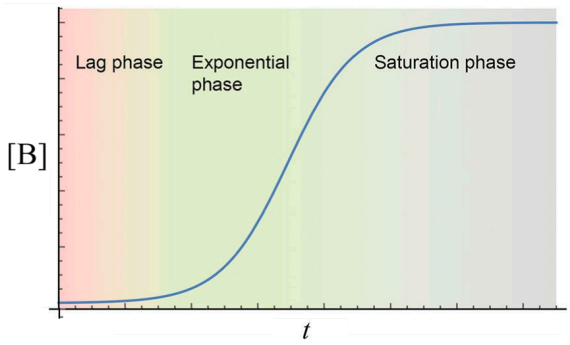
In general, factors contributing to accelerated chemical processing include:

- **High temperatures:** the Earth's surface was ~70 °C throughout the late Hadean, allowing hot liquid water oceans across the surface. Chemical and diffusion kinetics were therefore much faster.
- **Lunar tides:** the moon formed 4.527 BYA, and orbited ~16 times closer in the Hadean than today, amplifying the strength of lunar tides, allowing cyclic forcing of lava flows and generating heat due to tidal friction.
- **Radiation:** the Sun was more intense in the high energy UV and X-ray wavelengths, although was ~30% less intense overall. A hot greenhouse layer of CO_2 insulated the Earth. Earth's magnetic field was only beginning to start in the late Hadean, allowing a larger flux of cosmic rays and ions from the solar wind to reach the Earth.
- **Impact events:** a large influx of meteorites led to regular redistribution of land mass, including ejection to space. The meteorites delivered highly reduced minerals to the surface (e.g. metallic iron).
- **Weakly reducing atmosphere:** the Hadean's thick, hydride-rich, oxygen-poor atmosphere resembling the solar nebula slowly gave way to the 'prebiotic atmosphere', containing methane (CH_4), ammonia (NH_3), hydrogen (H_2), carbon dioxide (CO_2) as well as nitrogen (N_2) and some noble gases (He, Ne, Ar…).

17.5.5. Autocatalysis in Systems Chemistry

Autocatalytic reaction: a reaction in which the catalyst is also the product. The simplest form of an autocatalytic reaction is A + B ⇌ 2 B. The rate equation and sigmoid-like solution curves are: (k_+: forward rate constant, k_-: reverse rate constant, $[\mathrm{A}]_0$, $[\mathrm{B}]_0$: initial concentrations)

$$r = -\frac{d[A]}{dt} = \frac{d[B]}{dt} = k_+[A][B] - k_-[B]^2$$



Initially, no catalyst B is present and the reaction can only proceed at the minimal uncatalysed rate (lag phase).

As more product B is formed, it catalyses its own formation (exponential phase).

As the substrate A becomes limiting the rate falls again (saturation phase).

A common model is for [A] to be held constant (a buffered supply), with mechanism 1) A + B ⇌ AB, 2) AB → AA, 3) AA ⇌ 2 A. In this case, the model $\frac{d[B]}{dt} = k[B]^p$ is more appropriate, where $p < 1$ is the order of reaction. Often $p \approx ½$ (the **square root law of autocatalysis**).

In biological (enzyme-catalysed) reactions, the range of the rate constants k for catalysed reactions spans a much smaller range (~4 orders of magnitude) than the uncatalysed reactions (~15 orders of magnitude), helping to synchronise biochemical cascades and prevent rate limiting bottlenecks.

Autocatalytic Cycle:

Autocatalytic Set: a set of reactions in which every reaction is catalysed by one of the products of the reactions in the set (mutual catalysis).

Hypercycles:

17.5.5. Homochirality and Symmetry Breaking

All extant biochemistry is enantiospecific, using only L-amino acids and D-sugars to form functional homochiral macromolecules (proteins, carbohydrates and RNA/DNA). Since most ordinary chemical processes tend to produce racemic mixtures, a prebiotically plausible mechanism for breaking homochirality is required.

Enantioselective Recrystallisation: some supersaturated racemic amino acid solutions can nucleate as conglomerate solids (homochiral crystal grains) which selectively precipitate enantiopure crystals, increasing the *ee* of the supernatant solution. Enantioenrichment can also be induced in mixtures of other (racemate-forming) amino acids. Partially enantioenriched solutions of racemate-forming amino acids can also re-crystallise again, forming racemate crystals and entering the supernatant at the eutectic composition, which can have much higher *ee*, partitioning the chiral purity in each phase. This process can also occur exploiting differential volatility using sublimation/deposition (solid ↔ gas) instead of dissolution (solid ↔ aqueous).

Suitable environments include warm pools of water containing amino acids (solid ↔ aqueous), with cycles of rainfall and evaporation, or in sublimating organic space dust brought to Earth (solid ↔ gas).

Chiral-Induced Spin Selectivity (CISS): helical (spin-polarised) electrons interact with chiral molecules differently, such as asymmetric kinetics of electrochemical redox reactions due to differing activation energies, or selective degradation of helical molecules (like RNA). Sources of helical radiation include:

1) Photoelectron emission from UV irradiation of ferromagnetic materials (e.g. magnetite, Fe_3O_4).
2) Circularly polarised light, such as from the Sun or from rapidly-rotating neutron stars in space.
3) Surface muon flux from cosmic rays, polarised by the parity-violating weak nuclear force.

Suitable environments for (1) include shallow evaporative lakes with magnetite deposits utilising cyanosulfide chemistry (forming simple sugars and hydroxy-aldehydes). Mechanism (2) may be responsible for the slight enantioenrichment of the biomolecules found on meteorites.

Chiral Mineral Surfaces: the chiral surfaces of enantiomorphic inorganic crystals (e.g. quartz, calcite, cristobalite, many clays minerals) selectively adsorb corresponding enantiomers of chiral molecules. Some of these also act as catalysts, facilitating asymmetric catalysis for chemical reactions.

Asymmetric Autocatalysis: an unusual type of reaction in which racemic starting materials react in the presence of a small amount of low *ee* to amplify both the yield and *ee* of the product. A well-established example is the Soai reaction, although this is not prebiotic.

Kinetic Resolution at the Polymer Level: polymers with self-replicating capabilities e.g. ribozymes (RNA) can perform autocatalysis faster when they are homochiral, allowing them to outcompete other polymers over time with many cycles of replication.

17.1.6. Phosphate Chemistry

Phosphorus is often the limiting resource for biological systems on Earth, as it is incorporated into nucleotides, phospholipids and many metabolic molecules (e.g. ATP) in all known life. However, the geochemical cycling of phosphate is slow, and most phosphate minerals are insoluble. Reliable sources of reactive and soluble phosphorus are therefore required for prebiotic chemistry.

A variety of prebiotically plausible pathways for phosphorylation exist, such as via pyrophosphate ($PP_i = P_2O_7^{4-}$, a possible chemical ancestor of ATP), trimetaphosphate ($P_3m = P_3O_9^{3-}$), phosphorylated amino acids, diamidophosphate ($DAP = P(NH_2)_2O_2^-$), hydroxyapatite ($Ca_5(PO_4)_3OH$) and phosphide minerals (Fe_3P).

Phosphorylation of Nucleosides: the regioselectivity at the 5’ position is promoted in the presence of borate minerals, which can stabilise the ribose by forming 2’-3’ cyclic dimers. Nucleoside phosphorimidazolides ($-OPO_2(C_3N_2H_3)$) are an alternative.

17.1.7. Sugar Synthesis

Sugars are components of nucleosides (as ribose) and precursors to carbohydrates (e.g. glucose).

Simple sugars can be made by the formose reaction (Butlerov reaction), in which small molecules such as formaldehyde and glyceraldehyde (which are known to form from CO_2 and H_2O) can react under alkaline conditions in an autocatalytic cycle to form a range of monosaccharides. By varying the conditions, such as with borate minerals or chiral amino acids (e.g. proline), product specificity (e.g. forming ribose) can be increased significantly, as well as the yield. Under strongly alkaline conditions, the competing Cannizaro reaction may disrupt this process, so the milder pathways are considered more prebiotically relevant.

17.1.8 Lipid Synthesis

Lipids are precursors to cell membranes and other compartmentalised systems.

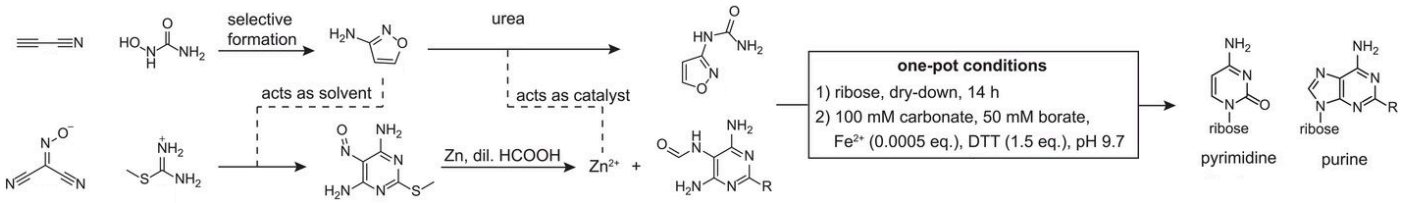
The acylation and phosphorylation of glycerol are both condensation reactions, with possible condensing agents being urea, cyanamide and imidazole. Lipid mixtures (mono/di/triglycerides) can be formed from ammonium salts of fatty acid and glycerol on hot kaolinite clay. Glycerol can also be phosphorylated by inorganic phosphate sources which react with fatty acid salts, and then further with cyanamide and difunctional hydroxy-amines to form phospholipid esters.

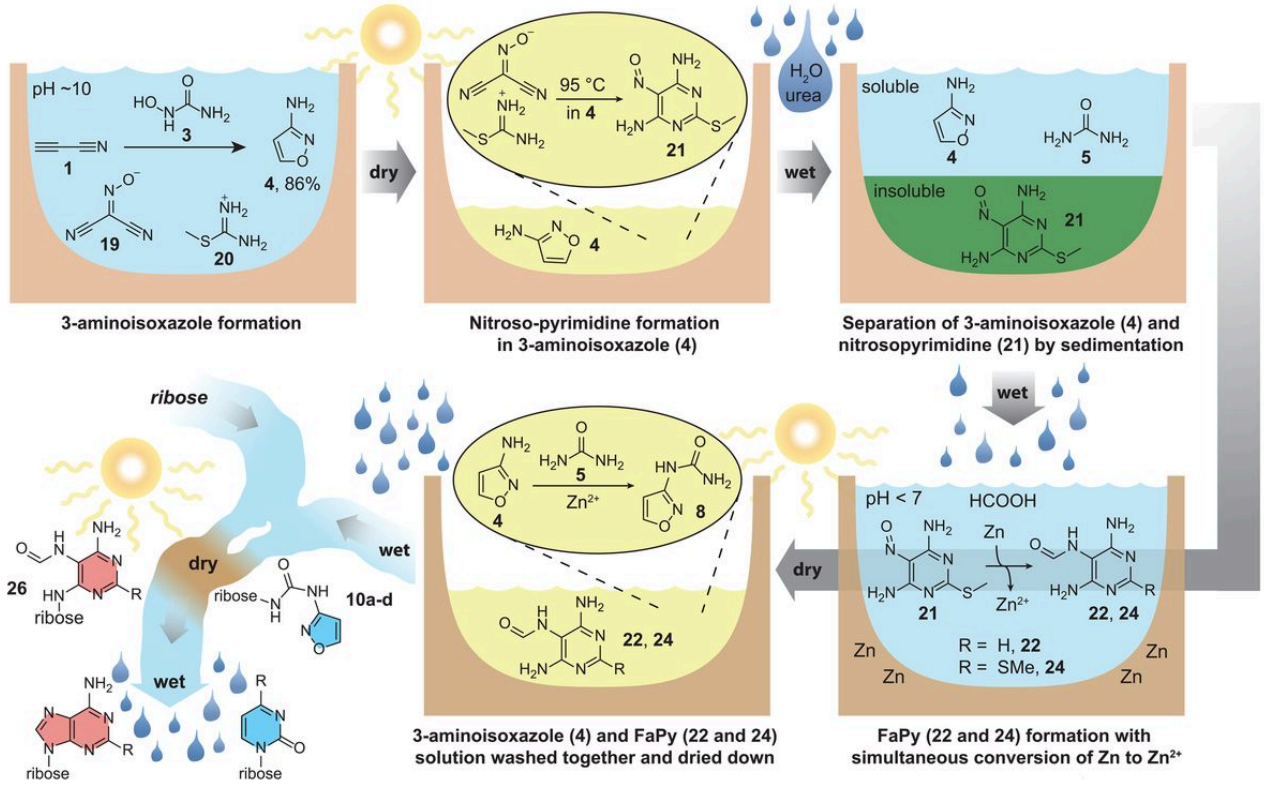
17.1.9. Nucleotide Synthesis

Nucleobases, Nucleosides and Nucleotides: precursors to genetic material (RNA, DNA).

The purine and pyrimidine nucleobases are known to be able to form from oxazole or iso-oxazole derivative precursors and glyceraldehyde. The products are chiral in high e.e. when small amounts of near-racemic (low e.e.) proline catalyst is present. Cyanoacetylene (C_2HCN) can be formed from nitrogen and methane. Hydroxyurea ($HONHCONH_2$) can be formed from hydroxylamine (NH_2OH) and isocyanic acid (HNCO). Malononitrile ($CH_2(CN)_2$) can also be generated from cyanoacetylene.

Once these are made, a prebiotically plausible synthesis of the four nucleosides is as follows, using wet-dry cycling, in which consecutive periods of hot and cold cause repeated precipitation and solution:





17.5.10. Protein Synthesis

17.5.11. RNA Synthesis

17.5.12. Carbohydrate Synthesis

17.5.13. Lipid Membrane Formation

Amphiphilic molecules such as fatty acids and phospholipids form micelles spontaneously in aqueous solution. The rate of micelle formation has been shown to increase in the presence of mineral clays (e.g. montmorillonite).

17.5.14. RNA World Hypothesis

17.5.15. Metabolism First Hypothesis

17.5.16. Amyloid World Hypothesis

17.5.17. Iron-Sulfur World and Hydrothermal Vent Hypothesis

17.5.18. Other Hypotheses of Abiogenesis

PAH (Polycyclic Aromatic Hydrocarbon) World:

Lipid World:

Panspermia and the Lukewarm Universe Hypothesis:

Panspermia is the hypothesis that life may have originated elsewhere throughout the universe, and was delivered to Earth via meteorites and space dust.

A related hypothesis is that life began very early in the universe's history, about 10 - 17 million years after the Big Bang, during which the temperature across the universe underwent gradual cooling from 100 ºC to 0 ºC (the 'habitable epoch').

These hypotheses are usually not considered scientifically valid.

17.5.19. Ribonucleoproteins

RNP world

Protoribosomes

17.5.20. First Universal Common Ancestor (FUCA)

Progenotes (ribocells)

17.5.21. Protocells

17.5.22. Origin of Viruses

Giant Viruses

Reduction hypothesis

17.5.23. Last Universal Common Ancestor (LUCA)

LUCA is expected to be a unicellular prokaryote that lived around 4 billion years ago, very simple compared to life today but nonetheless complex in chemical structure. It was not the first life, but is the earliest life from which all extant life is derived. It must have had an irreducible set of several hundred DNA-based genes which took over 100 million years to form. Synthetically developing a LUCA-like organism is infeasible on experimental time-scales, and is outside the scope of origin of life research (creating life *in vitro* is the study of synthetic biology, which primarily uses material derived from existing life).

LUCA is the root of the evolutionary tree of life, marking a reference point in the transition from chemical to biological evolution. LUCA represents an individual cell within a community of primitive cells, which exchange genetic material with each other by horizontal gene transfer. At this stage, heredity occurred both horizontally and vertically, so the tree of life more resembles an interrelated web, which stabilises into two domains: bacteria and archaea.

17.5.24. Origin of Bacteria and Archaea (4 - 3 BYA)

During the Archean era (4 - 2.5 billion years ago (Ga)), while surviving geologic formations from this era are rare, rocks from this time contain some of the oldest evidence of life. Carbonaceous matter has been identified in rocks as old as 3.95 Ga, but the oldest forms may be abiogenic (not produced by living organisms). Metamorphism has altered the molecular composition of Archean organic matter, making it challenging to distinguish between biological and non-biological carbonaceous compounds, such as those that could have formed in seafloor hydrothermal systems.

Benthic microbial mats, dating back to 3.47 Ga, are supported by the presence of organic laminae in stromatolitic carbonates, siliceous sinters, and siliciclastic sediments. However, the organic matter in these deposits rarely preserves fossil cellular structures, making it difficult to attribute the simple textures to microfossils or mineral templates. Filamentous-sheath microfossils are found in rocks from 2.52 Ga, and there are indications of altered counterparts as old as 3.47 Ga. Spheres and complex organic lenses in rocks as old as 3.22 Ga and ~3.4 Ga, respectively, are considered strong candidates for the oldest microfossils.

Titaniferous microtubes previously thought to be microbial trace fossils have been reevaluated as metamorphic or magmatic textures. Microbially-induced mineralization is supported by $CaCO_3$ nanostructures in 2.72 Ga stromatolites. Sulfides from 3.48 Ga and younger rocks bear S-isotope ratios indicative of microbial sulfate reduction. Ferruginous conditions, suggested by Fe-isotope ratios, may have fueled primary production via anoxygenic photosynthesis, potentially as early as 3.77 Ga. Microbial methanogenesis and methane oxidation, both likely anaerobic processes, are indicated by C-isotope ratios as early as 3.0 Ga and ~2.72 Ga, respectively.

Photosynthetic production of O_2 is inferred to have started between 3.2 and 2.8 Ga, well before the Great Oxidation Event (2.45–2.31 Ga). This is supported by various inorganic tracers of oxidation reactions, the morphology of benthic deposits, and evidence for aerobic nitrogen metabolism in N-isotope ratios at ~2.7 Ga.

17.5.25. Origin of Eukaryotes and Protists (2.6 BYA)

Today, eukaryotes are much more complex than prokaryotes, but they are thought to have split off from archaea via recombination with bacterial DNA when they were still relatively similar. The microfossil evidence suggests that early eukarya was as simple as other life at the time, and only later complexified through a series of one-off events.

The endosymbiotic theory states that about around this time, one archean prokaryote engulfed a small bacterial prokaryote of the phylum *Alphaproteobacteria* without destroying it, which would specialise to become the eukaryotic cell's mitochondria. Extant eukaryotic cell's mitochondria have numerous features indicative of this theory: having their own highly maternally-conserved mtDNA in the form of plasmids, a cell cycle based on binary fission that operates independently of the host cell's, and are specialised into a specific niche of energy production.

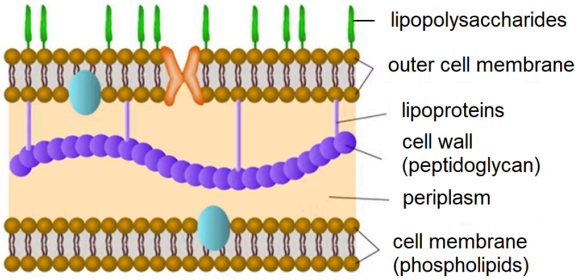
Eocyte hypothesis.

B18. MICROBIOLOGY

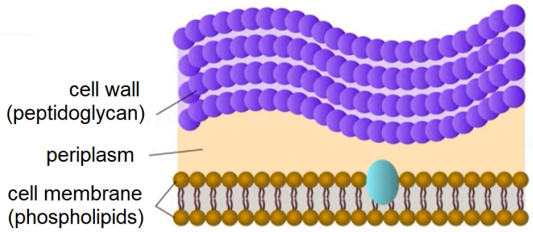
18.1. Bacteria and Archaea

18.1.1. Bacterial Structure and Organelles

Bacterial Cell Walls: contains different amounts of peptidoglycan.



Gram-negative Cell Wall
(stains pink)



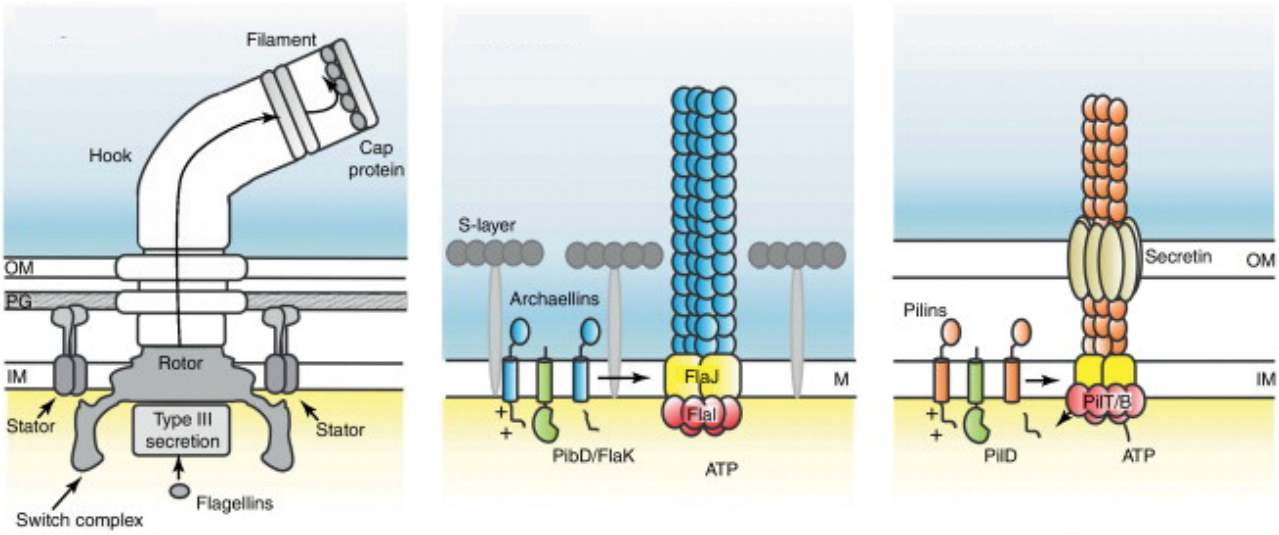
Gram-positive Cell Wall
(stains purple)

The Gram stain test (crystal violet stain to turn all cells purple → add iodine to trap crystal violet in the peptidoglycan layer of the bacteria → add decolouriser and counterstain (e.g. basic fuchsin)) can be used to discriminate between the types of cell wall.

18.1.2. Motility in Prokaryotes

Taxis: directional movement in response to a stimulus (as opposed to kinesis; no stimulus). Prokaryotes primarily use a flagellum for chemotaxis, and pili for inter-cell communication.

Bacterial Flagellum | Archaeal Flagellum | Pilus (type IV)



Abbreviations: OM: outer membrane; IM: inner membrane; PG: peptidoglycan; M: membrane.

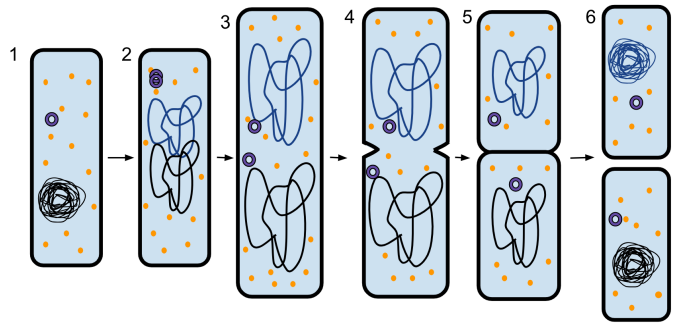
Other features found protruding from the cell membrane include the cilia and injectisome, many of which comprise similar protein subunits but with the overall organelle serving different functions.

18.1.3. Morphological Classification of Bacteria

- **Coccus (spherical):** unicellular. Plural: cocci.
  - Diplococci: two cocci.
  - Tetracocci: four cocci, in a square configuration.
  - Streptococci: a long chain of cocci.
  - Staphylococci: a cluster of cocci.
  - Sarcina: a 3D compact cluster of cocci.
- **Bacillus (rod-shaped):** unicellular. Plural: bacilli.
  - Diplobacilli: two bacilli.
  - Streptobacilli: a long chain of bacilli.
- **Spiral-shaped bacteria:**
  - Spirillum: rigid, external flagella.
  - Spirochete: internal flagella.

18.1.4. Bacterial Replication: Binary Fission

Prokaryotes replicate exclusively by binary fission, a form of asexual reproduction. Mitochondria and chloroplasts within eukaryotes also use this method, and their cell cycle is timed independently of the host, supporting the endosymbiotic theory for their evolutionary origin from ancient prokaryotes. Many protists can also undergo fission, sometimes to multiple (more than 2) daughter cells.



1: Before - tightly coiled DNA.

2: DNA duplication.

3: DNA separation to the cell poles and elongation of the cell wall.

4: New cell wall forms in centre.

5: New cell wall (septum) complete.

6: Separation to whole cells.

The resulting cells are identical (with random mutations). Bacterial binary fission occurs quickly (doubling time ~30 minutes at 37 °C), allowing populations to proliferate rapidly.

18.1.5. Infectious Bacteria

Bacteria evolved the ability to infect host organisms as a side effect of exploiting their resources. Bacteria can enter a host's body through inhalation, ingestion, open wounds, by sexual transmission or by being bitten from a vector. Once inside, bacteria can cause disease by:

- Adhering to cells in the body and colonising (multiplying), occurring in the incubation period.
- Biofilm formation (sticky webs of polysaccharides supporting a community of microorganisms) by quorum sensing (population-controlled gene expression), when a critical number of bacteria are present in one area. Biofilms commonly form on dental plaque or on implanted foreign bodies (microplastics, catheters, stents, pacemakers etc) and can impede the immune response.
- In Gram-positive bacteria, the peptidoglycan cell wall can stimulate fever/inflammation.
- In Gram-negative bacteria, the surface lipopolysaccharides can be recognised as an endotoxin by the immune system.
- Exotoxin release, which can interfere with a wide variety of host bodily functions, in addition to causing cytolysis. These exotoxins can be encoded either on the bacterial plasmid or present in a phage in the bacteria.
- Superantigen release, which trigger a strong immune response (toxic shock syndrome).

Bacteria can also evade host defences by changing their surface proteins, hiding within host cells, inactivating host defences (e.g. complement). Some bacteria have a capsule (slime layer) which mimics host cells, and some can form physical barriers using enzymes.

18.1.6. Archaea

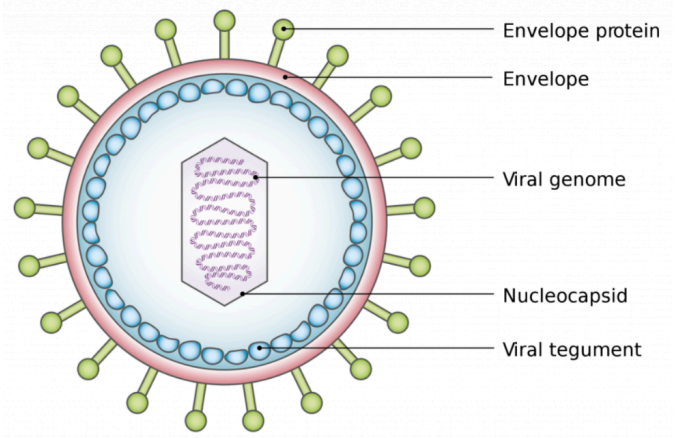
Archaea are single-celled prokaryotes. They are visually similar to bacteria: their genetics is based on plasmids with operon-based gene transcription and have no nucleus or membrane-bound organelles, but in fact are a completely different kingdom of life. They have cell walls and cell membranes (phospholipid monolayers based on ether-linked L-glycerol and terpenoid chains) with different compositions to bacteria, and can sometimes incorporate hard particles of inorganic minerals into their structure. They also use biomolecules not common in other organisms (e.g. heptose sugars). Archaea are often extremophiles, and were more prevalent on early Earth when life was far more primitive and conditions were harsher.

Unlike bacteria, archaea do not cause any known diseases in humans. In the large intestines of mammals, methanogenic archaea constitute about 1% of the gut microbiome, where they remove the hydrogen gas byproduct of early digestion and convert it to methane.

18.2. Virology and Parasitic Agents

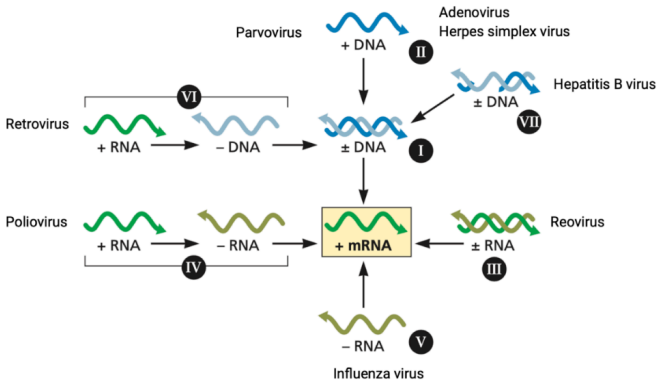
18.2.1. Structure of Viruses

Viruses: obligate intracellular parasites containing genetic material to be injected into and expressed by a cell.



Virus structure:

capsid around genome, in envelope with spike proteins



Viral genome types:

I: dsDNA, II: ssDNA, III: dsRNA, IV: ss(+) RNA, V: ss(-) RNA, VI: SS(+) RNA with DNA intermediate, VII: gapped dsDNA. Only +mRNA is 'ribosome ready'.

Virus Life Cycle: after attachment to a cell and release:

- Lytic cycle: synthesis of protein → reassembly → lysis
- Lysogenic cycle: incorporation of DNA (prophage) → cell replication → lysis on trigger

The genetic material inside a virus contains genes coding for most of the proteins contained within it (*gag* gene: capsid, *env* gene: envelope, as well as genes for various enzymes e.g. polymerase for replication). The envelope may also contain some of these enzymes already.

Viruses may exit cells via viroporins (protein ion channels) created by the viral genome.

18.2.2. Retroviruses

Retroviruses are ssRNA viruses. Within their envelope, they also carry a reverse transcriptase (RT) enzyme and an integrase enzyme, which are also encoded by its genome. They act by:

1. Fusion with the host cell membrane, release of viral RNA and enzymes into host cytoplasm.
2. A host cell tRNA primer binds to a specific sequence at the 5' end of the viral ssRNA.
3. Using host cell dNTPs, RT builds cDNA from the viral RNA while degrading the RNA template. RT then synthesises a complementary DNA strand to form dsDNA.
4. The pre-integration complex (PIC), including viral integrase with nuclear localisation signals (NLS), transports the dsDNA through a nuclear pore into the host nucleus.
5. Integrase interacts with host nuclear proteins to cut host DNA and insert the viral DNA. The integration site is somewhat random but often targets euchromatin.
6. The viral DNA has long terminal repeat (LTR) sequences at each end, which promote its transcription into new viral RNA and proteins. New viruses are assembled and can leave the cell to infect more cells.

Types of Retroviruses

- **Oncoviruses:** viruses that directly or indirectly contribute to cancer development.
- **Lentiviruses:** slow-acting retroviruses, which go 'dormant' (chronic; lysogenic cycle) for a long period of time once integrated by suppressing their transcription by various mechanisms (e.g. epigenetics). Includes HIV which causes AIDS by infecting $CD4^+$ T-lymphocytes. When the host cell undergoes mitosis, the new cells retain the proviral DNA.
- **Endogenous retroviruses (ERVs):** retroviruses infecting gametes can be incorporated into the organism's whole line of surviving descendants. These are used to study evolutionary relatedness (paleovirology). Due to mutations over time, they are now harmless and are unable to reform as viruses anymore due to being epigenetically marked as non-coding DNA. A small proportion of ERVs have been turned into useful genes by this process.

Some viruses replicate with similar characteristics as retroviruses (e.g. SV40, a DNA virus using large T antigen instead of integrase). The SV40 virus has a strong promoter sequence making it a useful vector in gene transfection.

Antiretroviral drugs e.g. azidothymidine can inhibit reverse transcriptase. Azidothymidine is a nucleoside with a substituted 3' azide group, which (after conversion to the dNTP form) blocks polymerisation on the DNA chain, terminating the reverse transcription process. The incomplete DNA fragments cannot be incorporated into the host genome. This can also hinder host cell DNA (including mtDNA) replication to some extent, causing side effects.

18.2.3. Prions

Prions are mutated endogenous lone proteins which, due to having a different folding pattern (isoform), are infectious agents. Prion proteins replicate by coming into contact with normal proteins, and can induce their misfolded conformation onto these proteins. The misfolded proteins do not perform any useful functions, leading to the progressive destruction of tissues within the host. The change in shape can also cause proteins to aggregate, causing e.g. brain shrinkage. The immune system is tolerant to PrP^{Sc} and does not destroy them.

The major prion protein (PrP) is encoded by the gene PRNP (CD230) in humans. PrP is highly conserved among mammals, despite its capacity for prion disease. PrP^{C} undergoes post-translational modification with a C-terminal glycosylphosphatidylinositol (GPI) anchors it to cell membranes. Prion proteins (PrP^{Sc}) typically have many β-sheets while the healthy proteins (PrP^{C}) have many α-helices. The β-sheets in prions can polymerize on contact like a template to form aggregated amyloid fibril structures that are stable, insoluble, and protease-resistant. The structures responsible for propagation can be studied using high-intensity X-rays to mutate the structures in various ways and test for replicative ability.

Prions are known to be responsible for some high-fatality-rate neurodegenerative diseases, such as:

- Bovine spongiform encephalopathy (BSE, ‘mad cow disease’), in cattle.
- Variant Creutzfeldt-Jakob disease (human variant of BSE), in humans.
- Ovine scrapie, in sheep.
- Chronic wasting disease (CWD), in goats.
- Kuru, in humans, due to cannibalism of prion-infected individuals.
- Alzeimers’ disease, indirectly caused by two slow-acting prions.

In the UK in the 1980s and 90s, an outbreak of BSE and vCJD originating from prion-contaminated meat-and-bone meal (MBM) used to feed cattle, transmitted to other species including pet cats and humans. Human susceptibility to this outbreak of vCJD was found to be genetic, occurring in those who are methionine-homozygous at PRNP codon 129, where there is a methionine/valine polymorphism. Those who are heterozygous or valine-homozygous may also be susceptible with an asymptomatic period of several decades, and the first heterozygous case of vCJD was found in 2016 with no mutations to the prion agent.

18.2.4. Viroids

18.2.5. Bacteriophages

18.2.6. Protists

Protists are a diverse set of single-celled eukaryotes with a wide range of characteristics. They include algae, flagellates, amoebas and slime moulds. There are also fungi-like protists. Protists primarily reproduce by binary fission, but some can also use meiosis (sexual).

Some common phyla of protists:

- **Euglenids** have flagella, can perform photosynthesis, and have photoreceptors (eyespots).
- **Diatoms** are unicellular algae, with a hard silica cell wall for surviving high pressures.
- **Dinoflagellates** have hard cellulose cell walls and flagella that cause rotational motion only.
- **Ciliates** have many cilia, an oral groove for eating, as well as trichocysts (spikes) for defence.
- **Radiolarians** have silicate internal skeletons and have a high degree of symmetry.
- **Amoebas** have psuedopod protrusions that can capture unicellular prey.

18.3. Mycology

18.3.1. Characteristics of Fungi

Fungi are small but larger-than-microscopic organisms such as mushrooms, yeasts and moulds. They reproduce both sexually and asexually, involving germination of spores (similar to how plants grow from seeds). Many fungi form large networks (hyphae) of mycelium fibres (like roots) for nutrient acquisition. Often, these filaments attach to roots of plants and trees in symbiotic relationships to form mycorrhizae, but may also be parasitic depending on the fungal phylum.

B19. ANIMALS AND PLANTS

19.1. Ecology and Evolutionary Biology

19.1.1. Foundations of Evolution

Observations leading to the formation of evolutionary theory (Darwin in *Origin of Species*, 1859)

- Organisms are well suited (adapted) for their environment.
- Life shares many characteristics (traits) despite rich variation in form and function.
- Populations can only expand to sizes sustainable by the available resources.
- Populations are naturally generally stable in size (resources are the limiting factor).
- Variation between individuals affects access to resources, influencing their reproduction rates.
- Variation is heritable from parent to offspring.
- The environment (and therefore what is beneficial to organisms) is constantly changing.

Darwin's Theory of Evolution: common descent with modification.
Selective pressures lead to competition creating differential reproductive success (**natural selection**).
The 'fitness' (conceptualised as a fitness landscape) is context-dependent for every population.

Neo-Darwinian Evolution: the 'modern synthesis', defined as variation in the allele frequency in a population over generations. Mendelian inheritance states **mutations** are the mechanism of variation that drives Darwinian natural selection. For Mendelian genetics, see Section 17.2.1.

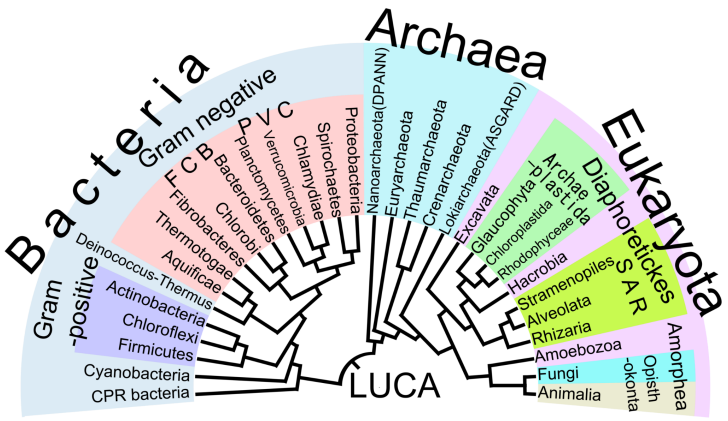
Additional mechanisms in the propagation of genetic variation:

- **Gene flow (migration):** movement of individuals between populations.
  - **Sexual selection:** preferential mating of individuals with particular heritable traits, possibly without an intrinsic survival advantage.
- **Genetic drift:** variation in allele frequency due random chance.
  - **Population bottleneck:** extinction events leave much smaller populations, whose new allele distribution may not represent the previous distribution.
  - **Founder effect:** the first individuals to occupy a new niche or habitat will set the allele frequency in the resulting population if they remain isolated.

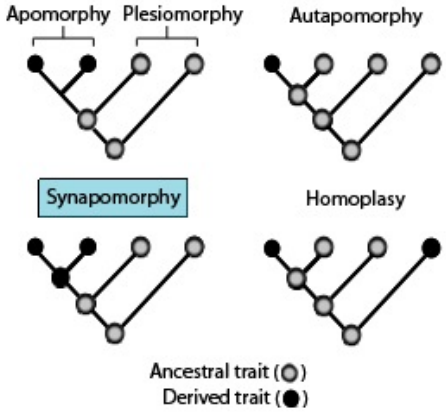
Neo-Darwinian evolution was the basis of the **extended evolutionary synthesis** (EES), including:

19.1.2. Phylogeny and Cladistics

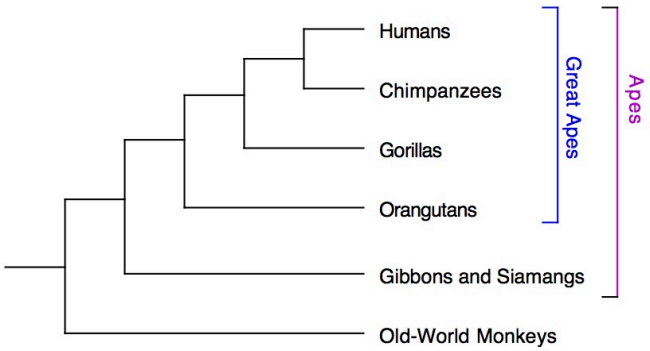
Phylogenetic tree: shows evolutionary diversification of life scaled in time, anywhere between the Last Universal Common Ancestor (LUCA) and extant life.



Cladograms: shows diversifications without considering the time taken.



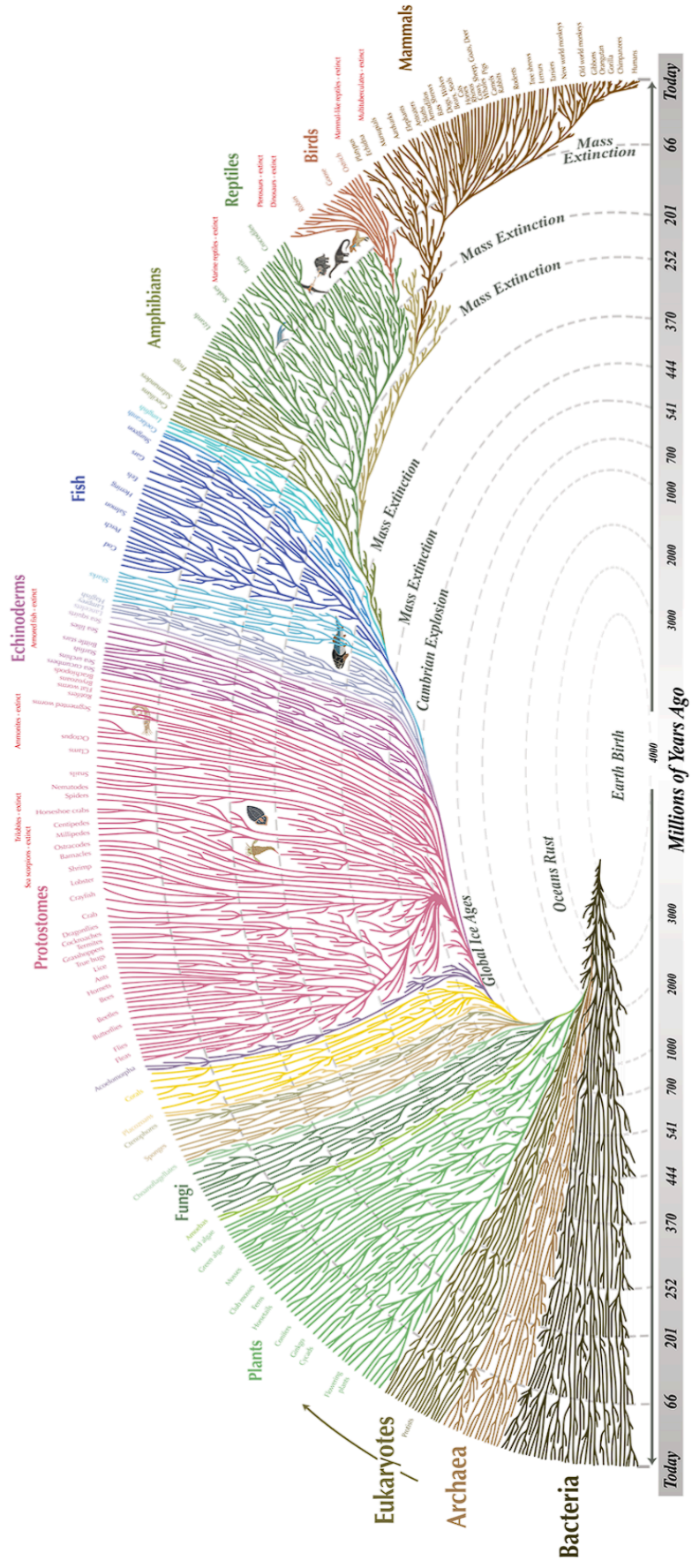
Variation in traits represented on a cladogram



Cladogram of extant New World monkeys showing the clades of Great Apes and Apes

Apomorphy:	a specialised (derived) trait unique to a clade.
(Sym)plesiomorphy:	a primitive (ancestral) trait inherited by a common ancestor.
Autapomorphy:	a trait **unique** to a single species or taxon within a larger group.
Synapomorphy:	a specialised (derived) trait **inherited from** a common ancestor.
Homoplasy:	a common trait of unrelated species e.g. due to convergent evolution.

19.1.3. Tree of Life



19.1.4. Taxonomy

The Linnaean classification system defined hierarchical tiers of organisms with similar characteristics. The precise tier system can change over time.

Life ⊃ Domain ⊃ Kingdom ⊃ Phylum ⊃ Class ⊃ Order ⊃ Family ⊃ Genus ⊃ Species

A common system used today, listed to the phylum level, is: Binomial name: *Genus species*

Domain *Bacteria*: unicellular prokaryotes (no nucleus or membrane-bound organelles), rigid peptidoglycan cell wall, circular chromosome.

Kingdom *Eubacteria*:

Phyla: *Acidobacteriota, Actinomycetota, Aquificota, Armatimonadota, Atribacterota, Bacillota, Bacteroidota, Balneolota, Bdellovibrionota, Caldisericota, Calditrichota, Campylobacterota, Chlamydiota, Chlorobiota, Chloroflexota, Chrysiogenota, Coprothermobacterota, Deferribacterota, Deinococcota, Dictyoglomota, Elusimicrobiota, Fibrobacterota, Fusobacteriota, Gemmatimonadota, Ignavibacteriota, Kiritimatiellota, Lentisphaerota, Mycoplasmatota, Myxococcota, Nitrospinota, Nitrospirota, Planctomycetota, Pseudomonadota, Rhodothermota, Spirochaetota, Synergistota, Thermodesulfobacteriota, Thermomicrobiota, Thermotogota, Verrucomicrobiota.*

Domain *Archaea*: unicellular prokaryotes, rigid non-peptidoglycan cell walls, unique rRNA genes/polymerases.

Kingdom *Archaea*:

Phyla: *Nitrososphaera*, *Thermoproteota, Euryarchaeota, Thorarchaeota, Methanobacteriota,* as well as many other candidate phyla

Domain *Eukarya*: eukaryotes (nucleus, membrane-bound organelles)

Kingdom *Protista*: a polyphyletic clade for any eukaryote not fitting into the below three (may be unicellular or multicellular, may or may not have cell walls).

Phyla: *Amoebozoa, Apusozoa, Bigyra, Cercozoa, Choanozoa, Ciliophora, Cryptistanas, Eolouka, Euglenozoa, Gyrista, Haptista, Malawimonada, Metamonada, Miozoa, Opisthosporidia, Percolozoa, Retaria.*

Kingdom *Fungi*: multicellular, chitin cell walls, heterotrophic

Phyla: *Ascomycota, Basidiomycota, Blastocladiomycota, Chytridiomycota, Glomeromycota, Microsporidia, Neocallimastigomycota, Zygomycota.*

Kingdom *Plantae*: multicellular, cellulose cell walls, autotrophic

Phyla: *Anthocerotophyta, Bryophyta, Charophyta, Chlorophyta, Cycadophyta, Ginkgophyta, Glaucophyta, Gnetophyta, Lycopodiophyta, Magnoliophyta, Marchantiophyta, Polypodiophyta, Coniferophyta, Rhodophyta.*

Kingdom *Animalia*: multicellular, no cell walls, heterotrophic

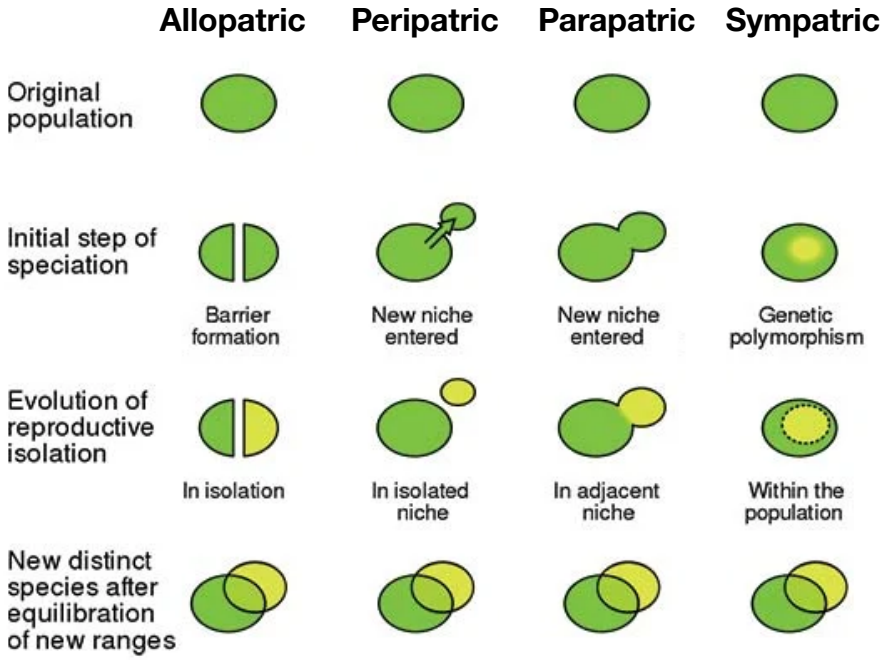
Phyla: *Annelida, Agmata, Archaeocyatha, Arthropoda, Brachiopoda, Bryozoa (Ectoprocta), Chaetognatha, Chordata, Cnidaria, Ctenophora, Cycliophora, Echinodermata, Entoprocta, Gastrotricha, Gnathostomulida, Hemichordata, Kinorhyncha, Loricifera, Micrognathozoa, Mollusca, Nematoda, Nematomorpha, Nemertea, Onychophora, Petalonamae, Phoronida, Placozoa, Platyhelminthes, Porifera, Priapulida, Proarticulata, Rhombozoa (Dicyemida), Rotifera, Saccorhytida, Tardigrada, Trilobozoa, Vetulicolia, Xenacoelomorpha.*

19.1.5. Evolution

Causes of Speciations

- **Natural Selection:** the ongoing process by which genes which produce useful adaptations to an organism's current environment can proliferate, while the others die off. The small changes caused by natural selection can lead to speciation over time, and are amplified/accelerated by other mechanisms.
- **Genetic Drift:** change in the allele frequencies in a population as a result of "sampling error" while selecting the alleles for the next generation from the gene pool of the current population. It is debated whether genetic drift contributes to speciation directly or only to evolution in general.

Modes of Speciation



19.1.6. Principles of Evolution

Species: a set of similar individual organisms capable of interbreeding or exchanging DNA (the 'biological species concept').

Adaptation: a beneficial characteristic to a species in its environment.

Variation: the diversity in the genes carried by a single species, caused by natural mutations.

Natural selection: the ongoing process by which genes which produce useful adaptations to an organism's current environment can proliferate, while the others die off.

Selective pressure: an external factor that influences the survivability of a species, potentially accelerating the process of natural selection towards adapting to the pressure.

Speciation: the diversification of a species into multiple species, often due to geographic separation into different environments (allopatric speciation).

Fitness landscape: the suitability of a gene as a function of its many possible mutations, which varies with time and environment. Natural selection can be thought of as moving towards the peaks of the fitness landscape at all times.

Rate of evolution: phyletic gradualism is the idea that evolution occurs slowly in a population, while punctuated equilibrium claims that populations are relatively stagnant most of the time, but abruptly diversify over short periods occasionally. Both models contribute to evolution, as the rate of evolution is indicative of the strength of the selective pressures acting in a given context.

19.1.7. Evidence for Evolution

Direct observation: the basic principles of evolution by natural selection are well-established scientific facts and are easily observed in organisms with short life cycles e.g. antibiotic resistance, seasonal viral infections, domestication of animals, as well as larger animals (including humans) over historical time scales.

Homology and embryology: similarity of a structure or function of body parts with different origins, implying descent from a common evolutionary ancestor. For very distant species, these similarities are only present in the embryo stage of development, before significant differentiation takes place. This is studied from the genetics perspective in evolutionary developmental biology.

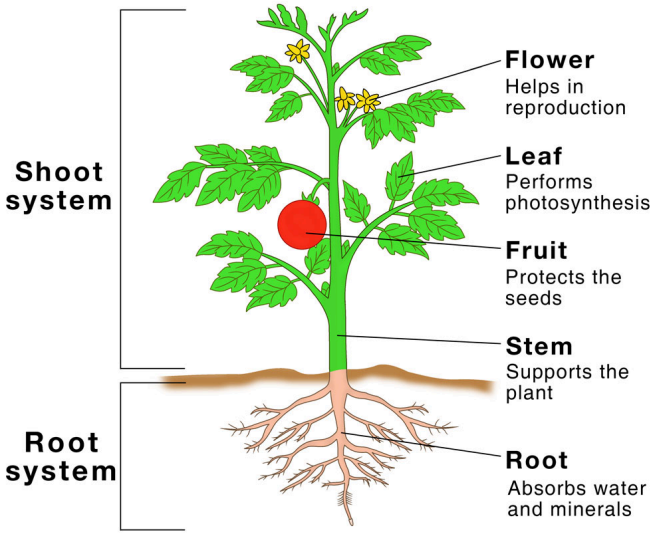
Genetics: DNA sequencing of extant life reveals a hierarchical similarity, implying common descent. Endogenous retroviruses (ERVs) are especially strong indicators of relatedness.

Transitional fossils: the fossil record contains various examples of organisms with clearly intermediate morphologies (e.g. *Archaeopteryx*, *Ichthyostega*) between distinct extant species, including those on and around the human lineage.

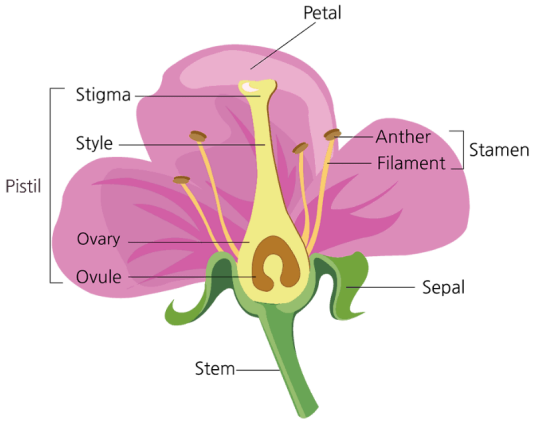
Stratigraphy and the geologic column: there is a strong correlation between the depth at which fossils are found and the age of the fossils, allowing trends over geologic timescales to be identified.

19.2. Botany and Agriculture

19.2.1. Structure of Flowering Plants (Angiosperms)



Parts and functions of a plant



Parts of a flower

Functions of the parts of a plant:

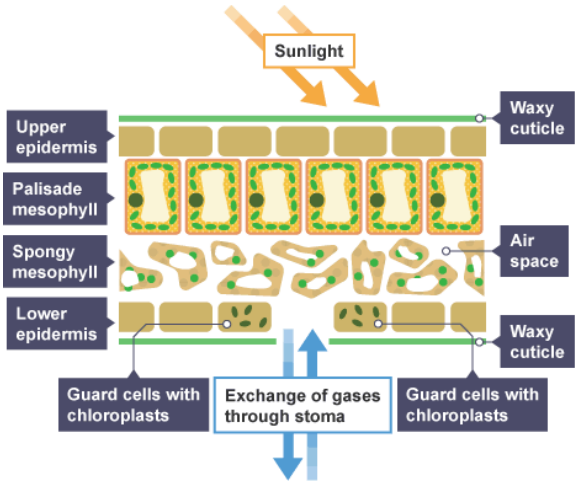
- **Stem:** provides structural support and allows for food, water and nutrient exchange.
- **Roots:** ingrains into soil. Made of root hair cells which absorb water by osmosis and minerals by active transport.
- **Leaves:** captures oxygen and performs photosynthesis (in the chloroplasts of palisade cells).
- **Sex Organs:** forms the reproductive system of the plant (pollen + ovule → seeds → plant).
    - **Sepal:** protects a developing flower bud.
    - **Stamen:** male reproductive organ. The anther produces and holds pollen grains (contains male gametes). The filament supports the anther and permits pollen dispersal.
    - **Petal:** brightly coloured to attract pollinators (bees, insects).
    - **Stigma:** female reproductive organ. Captures pollen grains (pollination).
    - **Ovule:** matures into a seed (fertilisation). Contains female gametes.
    - **Ovary:** matures into a fruit (which protects the seeds) after fertilisation. The seeds are to be dispersed and grow into new plants (germination).

Most plants are monoecious (shown above: contain both male and female reproductive organs). Some are diecious, where individual plants have separate sexes.

Life cycle: pollination → fertilisation → seed dispersal → germination.

19.2.2. Plant Organisation and Tissues

Layers of a Leaf: permits transport of gases.



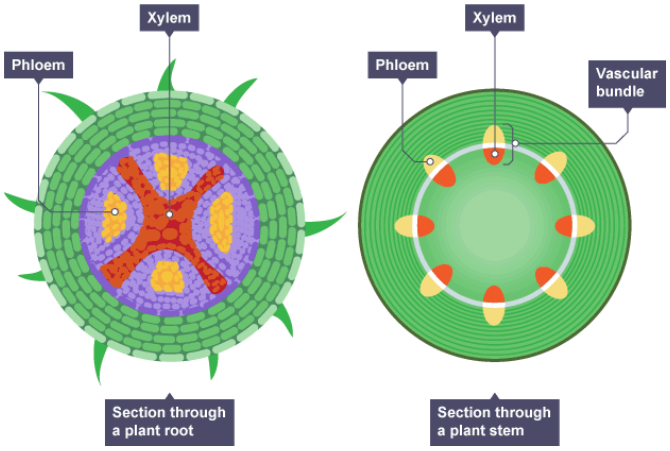
Waxy cuticle: a transparent layer which prevents water loss from the leaf in dry conditions.

Upper/Lower epidermis: protects from damage and infection. Contains stomata surrounded by guard cells to allow gas exchange (oxygen out, CO_2 in, water vapour in/out).

Palisade mesophyll: palisade cells contain lots of chloroplasts to generate glucose by photosynthesis.

Spongy mesophyll: loose-packed surrounded by water in which gases can dissolve.

Cross-Section of a Stem: permits transport of food and water.



Xylem: hollow tube of dead cells strengthened by lignin (primary component of wood). One-way (roots → leaves) transport of liquid water and minerals (the transpiration stream) by capillary action (cohesion-tension theory).

Phloem: Two-way (roots ↔ leaves) transport of cell sap (dissolved sugars and amino acids) (translocation), with porous sieve tube plates at regular lengthwise intervals.

The xylem and phloem are paired in vascular bundles running along the stem.

Rate of transpiration (volumetric flow rate of water) is set by the rate of evaporation: variations in temperature (hotter → faster), humidity (humid → slower: higher vapour pressure gradient), air movement (windy → faster: humid air is removed) and light intensity (brighter → faster: stomata open to increase oxygen uptake for photosynthesis).

Meristem tissue: exists at roots and shoots. Contains undifferentiated (totipotent; unspecialised) cells which serve to grow into the various parts of the plant.

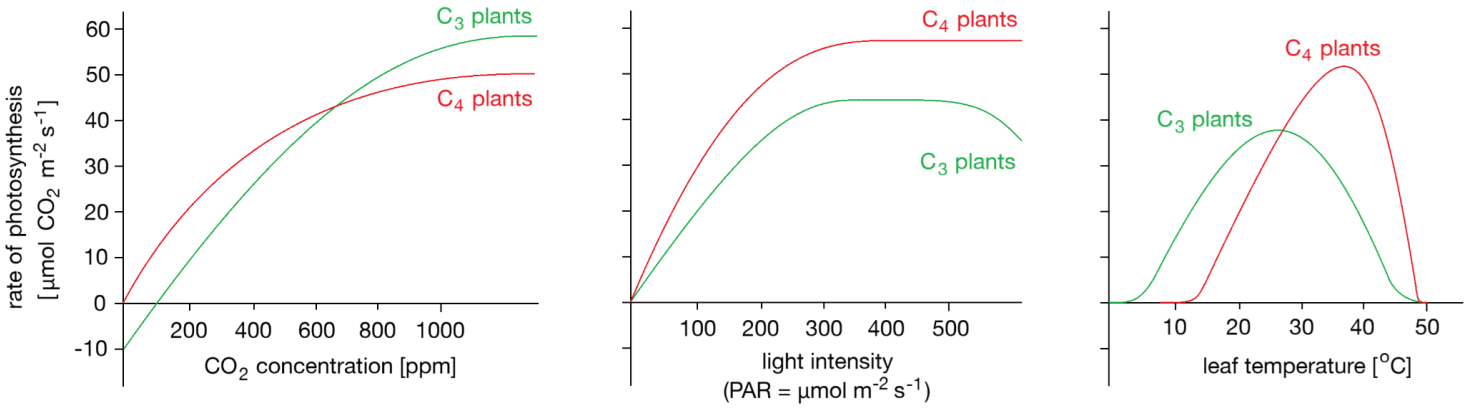
19.2.3. Respiration and Photosynthesis

For the detailed molecular biology of respiration and photosynthesis, see Section 17.3.1-4.

Aerobic Respiration:	$C_6H_{12}O_6 + 6\ O_2 \rightarrow 6\ CO_2 + 6\ H_2O$	(releases energy as ATP)
Anaerobic Respiration:	$C_6H_{12}O_6 \rightarrow 2\ C_2H_5OH + 2\ CO_2$	(releases energy as ATP)
Photosynthesis:	$6\ CO_2 + 6\ H_2O \rightarrow C_6H_{12}O_6 + 6\ O_2$	(absorbs energy as light)

Plants are a key step in the carbon cycle: atmospheric carbon dioxide is fixed into glucose, which is metabolised in animals when they eat the plants, and mineralises the soil when they/the plants die. Fossil fuel usage puts the carbon back into the atmosphere.

Rate of photosynthesis: usually measured by volumetric flow rate of CO_2 [μmol m^{-2} s^{-1}]



- **CO_2 concentration:** rate increases with concentration up to saturation at around 0.1% (1000 ppm). At [CO_2] < 100 ppm, photosynthesis may be reversed (photorespiration) in C_3 plants. Under atmospheric conditions, CO_2 is typically the limiting factor (diffusion controlled).
- **Light intensity:** rate increases with light intensity up to saturation (depending on chlorophyll content). At very high intensity, photoinhibition can reduce the rate in C_3 plants. The light intensity can be measured in lux or 'photosynthetically active radiation' (PAR = μmol m^{-2} s^{-1} of photons for wavelengths 400 nm < λ < 700 nm). For sunlight, 1 μmol m^{-2} s^{-1} = 45 lux.
- **Temperature:** enzyme kinetics favours an optimal temperature. The optimal temperature increases with CO_2 concentration: 25 °C at 325 ppm to 35 °C at 1900 ppm for C_3 plants. C_4 plants have a narrower tolerance than C_3 plants, with a higher optimal temperature.

The glucose produced in photosynthesis may be used for respiration, converted into insoluble **starch** for storage, used to produce **fat or oil** for storage, used to produce **cellulose** (strengthens the cell wall), and used to produce **amino acids** for protein synthesis. To produce proteins, plants also use nitrate ions that are absorbed from the soil.

19.2.4. Plant Hormones and Homeostasis

Plant Hormones: natural signalling molecules, also used in agriculture / horticulture.

- **Gibberellins:** initiates seed germination. Artificially used to end seed dormancy, promote flowering and increase fruit size.
- **Auxin:** promotes cell growth/division. Artificially used as weed killer, rooting powder and promoting growth in tissue culture.
- **Ethene:** limits cell growth/division and promotes ripening of fruits. Artificially used to control ripening of fruit during storage and transport. The mechanism involves π-bonding of ethene to a Cu^{+} centre in a transcription factor.

Overuse of artificially-distributed plant hormones can reduce biodiversity.

Plants produce hormones to coordinate and control growth and responses to light (phototropism) and gravity (gravitropism / geotropism). Unequal distributions of auxin cause unequal growth rates in plant roots and shoots.

19.2.5. Phytopathology and Plant Diseases

Plant diseases can be identified by stunted growth, spots on leaves, areas of decay (rot), abnormal growths, malformed stems or leaves, discolouration, and the presence of pests. Identification can be made by reference to a gardening manual or website, taking infected plants to a laboratory to identify the pathogen, and using monoclonal antibody testing kits.

Ion deficiency conditions: fix by using e.g. NPK fertiliser: nitrate, phosphate, potassium.

- Nitrate (NO_3^{-}) deficiency: stunted growth due to decreased amino acid synthesis.
- Magnesium (Mg^{2+}) deficiency: chlorosis due to decreased chlorophyll synthesis.
- Phosphate (PO_4^{3-}) deficiency: dark green/purple spots on leaf undersides.
- Potassium (K^{+}) deficiency: browning at the edges and yellowing between veins.

Common plant diseases: plants are often infected by viruses, fungi and pest insects.

- Tobacco mosaic virus (TMV) causes yellow mottling on leaves.
- *Diplocarpon rosae* (a fungus) causes rose black spot.
- Aphids (insects) eat through leaves.

Plant defences: adaptations include

- Physical (cellulose cell walls, tough waxy cuticle on leaves, layers of dead cells around stems (bark on trees) which fall off)
- Chemical (antibacterial chemicals, poisons to deter herbivores)
- Mechanical (thorns and hairs deter animals, sometimes with warning colouration (aposematism), leaves which droop or curl when touched, mimicry to trick animals)

19.2.6. Horticulture

Annual plant: lives for one year
Biennial plant: lives for two years
Perennial plant: lives for a longer period

Deciduous plants: leaves fall off once mature
Evergreen plants: leaves remain throughout life

Gardening

Soil Properties

The soil texture affects the levels of moisture, aeration and biological activity (microorganisms and earthworms), as well as the soil's mechanical workability (ease of digging and tilling).

The soil chemistry affects the ease of nutrient absorption, with most plants being optimally grown in soils of pH 6-7. Simple colourimetric soil pH test kits can be used to determine pH.

← Soil Texture Triangle

Greenhouses

19.2.7. Symbiotic and Synergistic Agriculture Techniques

Permaculture: a design philosophy that seeks to mimic natural ecosystems to create sustainable, self-sufficient, and regenerative agricultural systems, maximising biodiversity, and minimising waste.

Agroforestry (Afforestation, Rewilding): combines agricultural crops with trees and shrubs in the same farming area. Trees provide shade, windbreaks, nutrient cycling, and soil improvement. The combination of crops and trees enhances biodiversity, soil fertility, ecosystem resilience, and promotes environmental conservation. A traditional regenerative form of agroforestry is 'syntropic agriculture', in which whole communities live among the ecosystem, in which the natural progression of vegetation is followed: bare rocks → lichens → small annuals, grasses and perennials (pioneer species) → shrubs and shade intolerant species → shade tolerant trees (climax community).

Hydroponics: a soilless farming technique where plants are grown in nutrient-rich water solutions. This method allows for precise control of nutrients and water, leading to faster growth and higher yields compared to traditional soil-based farming. A variation is 'sandponics', where sand is used as a growth medium and filter instead of soil. It provides better root support and water retention, making it suitable for certain plant species, especially those in desert biomes.

Aquaponics: combines hydroponics with aquaculture (fish farming). Fish waste (ammonia) can be metabolised to nitrate by nitrifying bacteria in soil for plants, and the plants' roots help to filter the water, creating a closed-loop system. Variations include sandponics and aeroponics, suitable for different types of plants.

Agrivoltaics (dual-use farming): co-location of solar panels and agricultural crops on the same land. The panels provide shade for crops, reducing water evaporation and improving crop yields. Most suited for crops requiring low sunlight e.g. leafy greens, herbs.

Vertical Farming: stacking of plants in vertical layers, often indoors or in controlled environments. It minimises land use, optimises light exposure, and conserves water. LED lights can be used to provide optimal conditions, with different wavelengths for different stages of growth, useful for space-limited urban areas.

19.2.9. Issues in the Agriculture Industry

Food and Water Security: ensuring a stable and consistent supply of nutritious food and safe freshwater to all people is a global challenge. Population growth, climate change, and resource limitations can impact availability and affordability.

Food Miles: the distance travelled by a food product from where it is grown to where it is sold can be several thousand miles. This limits freshness and requires additional processing to maintain the appeal of the food, as well as contributing to greenhouse gas emissions in transport, and takes away jobs from local farmers. Government subsidies to local farmers can help encourage consumers purchasing locally and in season, helping to reduce food miles.

Genetically Modified Organisms (GMOs): use of organisms with altered genetic material for specific traits. They can enhance crop yields, nutritional content, and resistance to pests and diseases. While some people view GMOs as a way to address food security, others have concerns about potential environmental effects (e.g. horizontal gene transfer), human health impacts, and corporate control over the food supply.

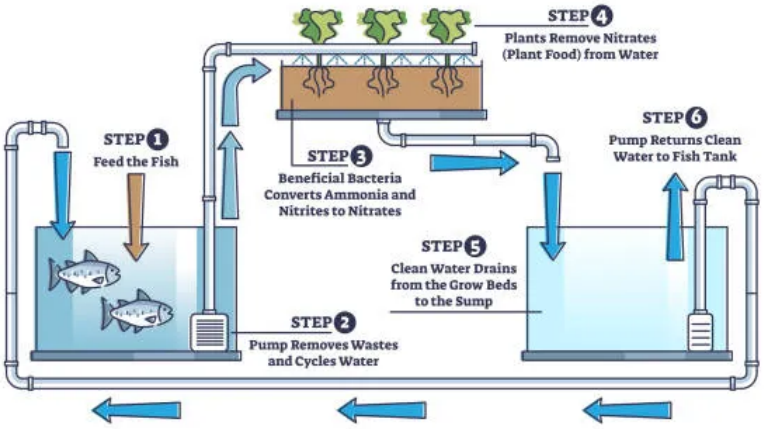
Pesticides: used to control pests and increase agricultural yields. However, they can also have negative impacts on ecosystems, non-target organisms, and human health. Integrated Pest Management (IPM) strategies aim to minimise pesticide use by combining various techniques such as biological control and cultural practices.

Lab-Grown Meat (Cellular Agriculture): involves producing meat from animal stem cells and an animal-free growth medium without traditional animal farming, providing a sustainable protein source with fewer ethical concerns. It is currently thought to be the best approach for widespread adoption of vegetarian or vegan diets among the general population, which could alleviate food security concerns significantly. However, challenges include scaling up production, addressing regulatory concerns, and public acceptance.

Sustainable Fisheries: overfishing and illegal, unreported, and unregulated (IUU) fishing threaten marine ecosystems and fish populations. Sustainable fisheries management involves setting catch limits, protecting vulnerable species, and implementing measures to prevent overfishing. Marine Protected Areas (MPAs) also contribute to ecosystem preservation.

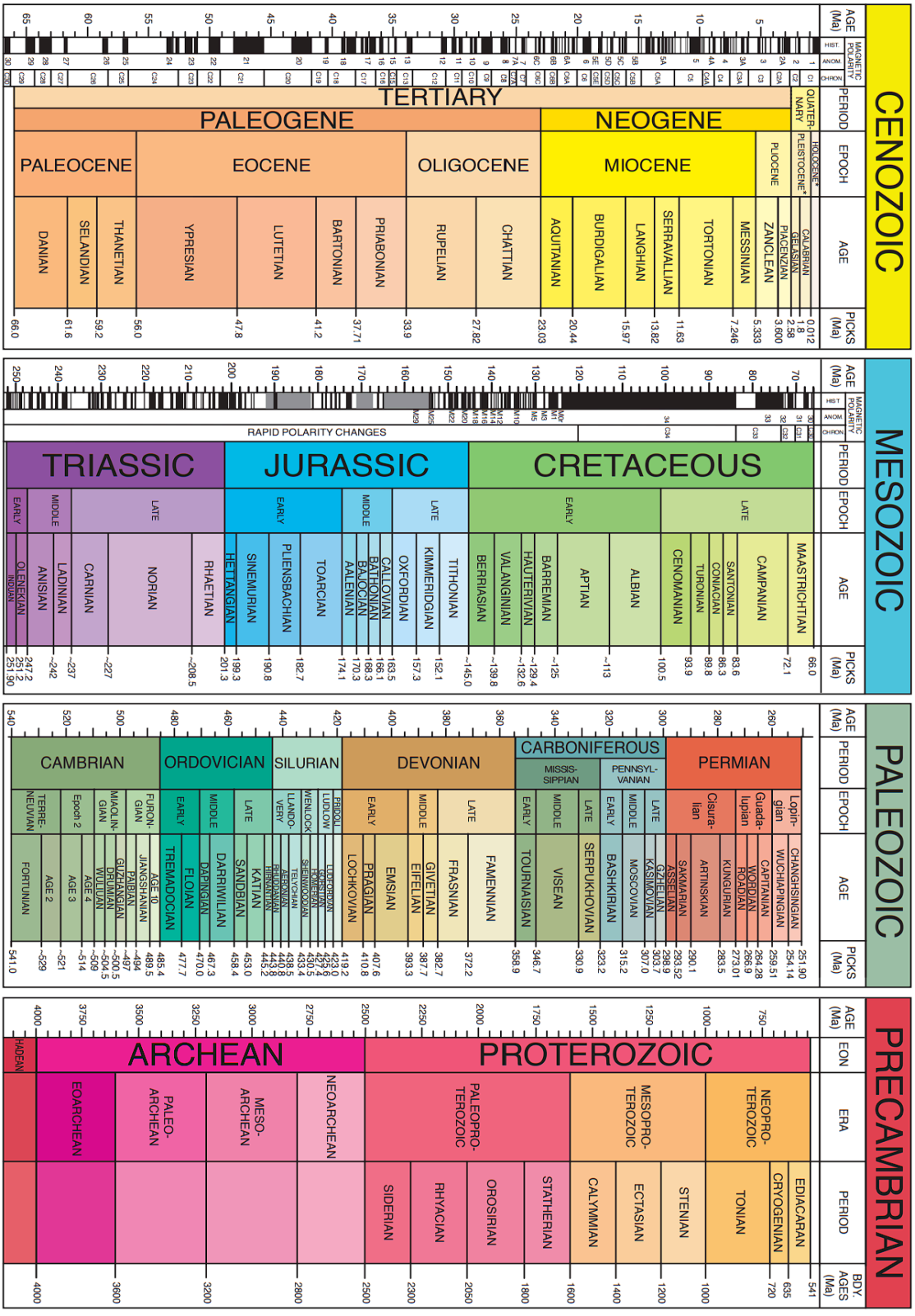
Deforestation and Desertification: degradation of natural environments by overfarming (subsistence, ‘slash and burn’ or commercial) and overgrazing leads to infertile land and soil erosion. Widespread logging contributes to global warming due to the reduction in CO_2 storage.

19.2.10. Aquaponics



19.3. Zoology and Paleontology

19.3.1. Geologic History of Earth



19.3.2. Significant Events in Evolutionary History

(ka: kilo-annum (thousand years ago), Ma: mega-annum (million years ago),
Ga: giga-annum (billion years ago))

Mass Extinction Events: the ‘Big Five’.

- **Ordovician mass extinction** (445 Ma).
- **Devonian extinctions** (372-359 Ma). A combination of the Kellwasser event and the Hangenberg event.
- **Permian extinction event** (The Great Dying) (252 Ma).
- **Triassic extinction event** (201 Ma)
- **Cretaceous (K-Pg) extinction event** (66 Ma). Caused the extinction of all non-avian dinosaurs, and the resulting open niches gave rise to the ‘age of mammals’, from which all primates arose.

Other Extinction Events:

- Great Oxidation Event (2.3 Ga).
- End-Ediacaran extinction event (550 Ma - 539 Ma).
- Capitanian mass extinction event (259 Ma).
- Pleistocene glaciation (2.58 Ma - 11.7 kA): the most recent ice age, out of five in history.
- Anthropocene extinction (100 ya - present): the ongoing rapid endangerment and extinction of many species due to human activity post-industrial revolution.

Biodiversification Events

- Avalon explosion (565 Ma):
- Cambrian explosion (540-515 Ma):
- Great Ordovician Biodiversification Event (GOBE) (497-467 Ma): It has been suggested that the Cambrian explosion and GOBE may have in fact been one long continuous radiation.

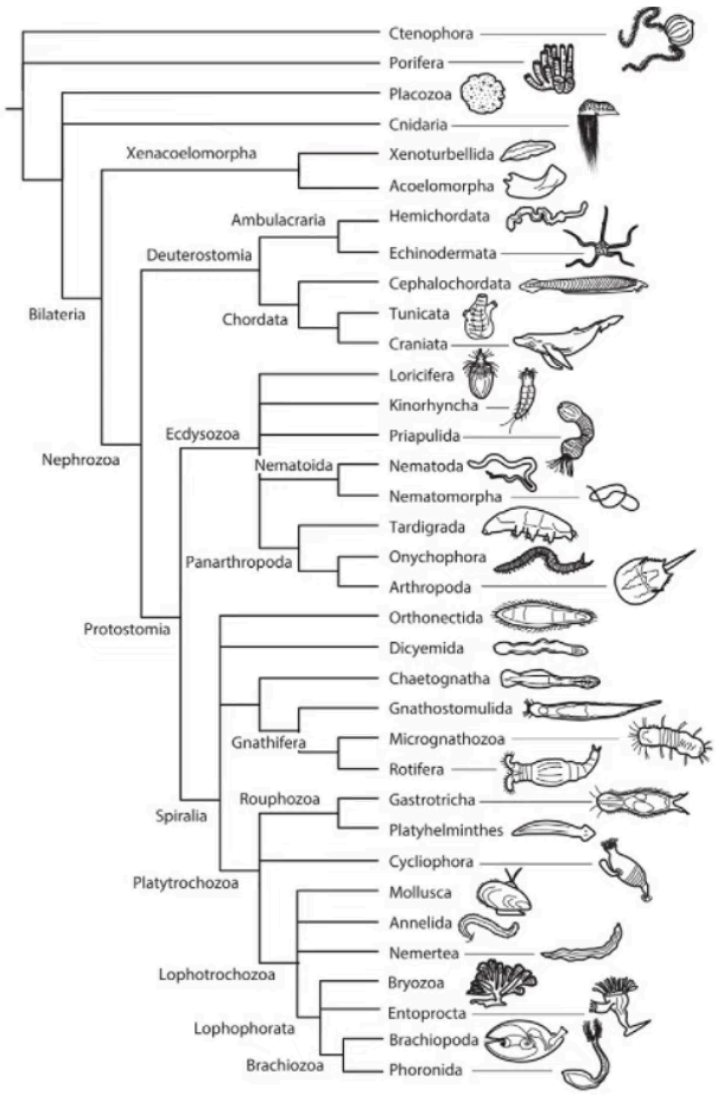
19.3.3. Taxonomic Classification of Animals

Appearance of Animals

Domain ***Eukarya*:**
- Clade *Parabasalia* / *Fornicata* / *Preaxostyla* [protists]
- Eukaryote host + *α-proteobacterium* [mitochrondrial endosymbiont]
    - Clade *Discoba*
        - Clade *Jakobida* / *Heterolobosea* / *Euglenozoa* [protists]
    - Clade *Neokaryotes*
        - Clade *Harosa*
        - Clade *Archaeplastida*
            - **Kingdom *Plantae***
        - Clade *Amorphea*
            - Clade *Amoebozoa*
            - Clade *Obazoa*
                - Clade *Apusomonadida*
                - Clade *Opisthokonta*
                    - Clade *Holomycota*
                        - Clade *Nucleariae*
                        - **Kingdom *Fungi***
                    - Clade *Holozoa*
                        - Clade *Ichthyosporea*
                        - Clade *Pluriformea*
                        - Clade *Filozoa*
                            - Clade *Filasterea*
                            - Clade *Choanozoa*
                                - Clade *Choanoflagellatea*
                                - **Kingdom *Animalia***

Diversification of Animals

There is some debate as to relationships between the most basal clades *Ctenophora*, *Porifera*, *Placozoa* and *Cnidaria*, with other arrangements proposed. The phylogeny shown is due to Giribet and Edgecombe (2020).



19.4. Primatology, Paleoanthropology and Archaeology

19.4.1. Taxonomic Classification of Extant Primates

Primates are a taxonomic order of class *Mammalia*, phylum *Chordata*, characterised by grasping hands, forward facing eyes (binocular vision), nails instead of claws, enlarged brains, diverse diets, reduced olfactory sense, social behaviours and flexible locomotion.

Order *Primates*:

- Suborder *Strepsirrhini* (lemurs)
  - Infraorder *Chiromyiformes*
    - Family *Daubentoniidae* (aye-ayes)
  - Infraorder *Lemuriformes*
    - Family *Cheirogaleidae*
      - Subfamily *Cheirogaleinae* (dwarf and mouse lemurs)
      - Subfamily *Phanerinae* (fork-crowned lemurs)
    - Family *Lemuridae* (true lemurs)
    - Family *Megaladapidae* (sportive and koala lemurs)
    - Family *Indridae* (indris, sifakas, and avahis)
    - Family *Palaeopropithecidae* (sloth lemurs)
    - Family *Archaeolemuridae* (baboon lemurs)
  - Infraorder *Lorisiformes*
    - Family *Lorisidae*
      - Subfamily *Lorisinae* (lorises)
      - Subfamily *Perodicticinae* (pottos and angwantibos)
    - Family *Galagidae* (bush babies, or galagos)
- Suborder ***Haplorhini*** (‘haplorhine’)
  - Infraorder *Tarsiiformes*
    - Family *Tarsiidae* (tarsiers)
  - Infraorder ***Simiiformes*** (‘simians’; ‘anthropoids’)
    - Parvorder *Platyrrhini* (‘platyrhines’; ‘New World monkeys’)
      - Family *Callitrichidae* (marmosets)
      - Family *Cebidae* (capuchins and squirrel monkeys)
      - Family *Aotidae* (durukulis, or night monkeys)
      - Family *Pitheciidae* (sakis, uakaris, and titis)
      - Family *Atelidae* (spider, woolly, and howler monkeys)
    - Parvorder ***Catarrhini*** (‘catarrhine monkeys’; ‘Old World monkeys’)
      - Superfamily *Cercopithecoidea*
        - Family *Cercopithecidae*
          - Subfamily *Cercopithecinae* (macaques, baboons, geladas…)
          - Subfamily *Colobinae* (‘leaf-eating monkeys’; langurs…)
      - Superfamily ***Hominoidea*** (‘hominoids’; apes)
        - Family *Hylobatidae* (gibbons and siamangs)
        - Family ***Hominidae*** (‘hominids’; great apes)
          - Subfamily *Ponginae* (orangutans)
          - Subfamily ***Homininae*** (‘hominines’; African apes and humans)
            - Tribe *Gorillini* (gorillas)
            - Tribe ***Hominini*** (‘hominins’)
              - Genus *Pan*
                - Species *Pan troglodytes* (chimpanzees)
                - Species *Pan paniscus* (bonobos)
              - Genus ***Homo*** (humans)
                - Species ***Homo sapiens*** (extant humans)

19.4.3. Evolutionary History of Primates

Genera and Species of Extinct Primates: selected species are mostly (but not all) on the human lineage.

Southern Africa Eastern Africa Northern/Western Africa Europe Asia America Oceania

Paleogene (66 - 23.03 MYA)

- *Purgatorius* (66-63 MYA)
- *Altiatlasius* (58-55 MYA)
- *Cantius* (56-47.8 MYA)
- *Smilodectes* (50.3-46.2 MYA)
- *Notharctus* (54-38 MYA)
- *Parapithecus* (40-33 MYA)
- *Biretia fayumensis* (~37 MYA)
- *Aegyptopithecus zeuxis* (38-29 MYA)
- *Proteopithecus* (34-28 MYA)
- *Catopithecus* (32-31 MYA)
- *Apidium phiomense* (30-28 MYA)
- *Kamoyapithecus hamiltoni* (27.5-24.2 MYA)
- *Rukwapithecus fleaglei* (25.2 MYA)

Miocene (23.03 - 5.33 MYA)

- *Proconsul africanus* (23-14 MYA)
- *Morotopithecus* (21-20.6 MYA)
- *Rangwapithecus* (~20 MYA)
- *Ekembo nyanzae* (20-17 MYA)
- *Dendropithecus* (20-15 MYA)
- *Limnopithecus* (20-11.6 MYA)
- *Micropithecus clarki* (19-15 MYA)
- *Prohylobates* (18 MYA)
- *Afropithecus turkanensis* (18-16 MYA)
- *Heliopithecus* (~17 MYA)
- *Victoriapithecus macinnesi* (17-15 MYA)
- *Nyanzapithecus* (16-11.6 MYA)
- *Otavipithecus* (16-11.6 MYA)
- *Equatorius* (15.58-15.36 MYA)
- *Nacholapithecus* (15-14 MYA)
- *Kenyapithecus* (14 MYA)
- *Griphopithecus* (13.65-11.1 MYA)
- *Pierolapithecus* (13 MYA)
- *Dryopithecus* (12.5 MYA)
- *Sivapithecus* (12.5 MYA)
- *Danuvius* (11.6 MYA)
- *Hispanopithecus* (11.1-9.5 MYA)
- *Rudapithecus* (10 MYA)
- *Ankarapithecus* (10-9.8 MYA)
- *Nakalipithecus nakayamai* (9.9-9.8 MYA)
- *Khoratpithecus* (9-7 MYA)
- *Samburupithecus kiptalami* (9.5 MYA)
- *Ouranopithecus* (9.5-7.4 MYA)
- *Graecopithecus* (9-8 MYA)
- *Chororapithecus abyssinicus* (~8 MYA)
- *Yuanmoupithecus xiaoyuan* (8.2-7.1 MYA)
- *Oreopithecus* (7 MYA)
- *Sahelanthropus tchadensis* (7 MYA)
- *Lufengpithecus* (8-6 MYA)
- *Orrorin tugenensis* (6 MYA)
- *Ardipithecus kadabba* (5.8-5.2 MYA)

Pliocene (5.33 - 2.6 MYA)

- *Indopithecus giganteus* (5-4 MYA)
- *Ardipithecus ramidus* (4.4 MYA)
- *Australopithecus anamensis* (4.2-3.8 MYA)
- *Kenyanthropus platyops* (3.5 MYA)
- *Australopithecus afarensis* (3.6-2.9 MYA)
- *Australopithecus africanus* (3.3-2.1 MYA)

Pleistocene (2.58 MYA - 11.7 kYA)

- *Paranthropus aethiopicus* (2.7-2.3 MYA)
- *Homo habilis* (2.3-1.65 MYA)
- *Homo ergaster* (1.8-1.3 MYA)
- *Paranthropus boisei* (2.5-1.15 MYA)
- *Australopithecus sediba* (1.98 MYA)
- *Homo rudolfensis* (2 MYA)
- *Paranthropus robustus* (2.87-0.87 MYA)
- *Homo erectus** (1.89-0.11 MYA)
- *Homo antecessor* (1.2-0.8 MYA)
- *Gigantopithecus blacki* (1-0.1 MYA)
- *Bunopithecus sericus* (~700-130 kYA)
- *Homo heidelbergensis*** (700-200 kYA)
- *Homo floresiensis* (100-50 kYA)
- *Homo neanderthalensis* (400-40 kYA)
- *Homo naledi* (336-235 kYA)
- Denisovans / *Homo longi**** (285-51.6 kYA)

Holocene (11.7 kYA - present)

- *Homo sapiens* (300 kYA - present: extant

* *H. erectus sensu lato* = {*H. georgicus, H. ergaster, H. erectus sensu stricto*}, a 'lumped' taxon.
* *H. habilis* = {*H. rudolfensis*, *H. gautengensis*}
Genus *Australopithecus* = {***A. africanus, A. afarensis, A. sediba, A. anamensis***, *A. deyiremeda, A. bahrelghazali, A. garhi, A. prometheus*} (**bold:** most commonly accepted taxa)

** wastebasket taxon (‘muddle in the middle’). ***unofficial; *H. longi* is suggested by some to be a Denisovan.

Characteristics found in species **most closely related to humans** include bipedality, an anterior foramen magnum, parabolic palate of teeth, smaller canines and larger molars, larger brain size, opposable thumbs, grasping hands, nails instead of claws, shorter snout/nose, reduced olfactory sense, stereoscopic trichromatic vision, extended parental periods, complex social structures with advanced languages and behavioural traits.

Representative Skulls of Some Known Hominins (arranged by general morphologic similarity)

Binomial classification **(Specimen identifier), brain case size in cc, time period**

A *Sahelanthropus tchadensis* (TM 266), 350 cc, 7-8 MYA
B *Ardipithecus ramidus* (ARA-VP-6), 300-350 cc, 4.4 MYA
C *Australopithecus afarensis* (multi-specimen cast), 350-500 cc, 2.95-3.24 MYA
D *Australopithecus africanus* (STS 5), 400-555 cc, 2-3 MYA
E *Australopithecus sediba* (MH1), 420-440 cc, 1.95-2 MYA

F *Homo habilis* (KNM ER 1813), 509-687 cc, 1.65-2.3 MYA
G *Homo rudolfensis* (KNM ER 1470), 750-825 cc, 2 MYA
H *Homo gautengensis* (STW 53), est. 500-700 cc,1.5-1.8 MYA
I *Homo georgicus* (D2700), 600 cc, 1.7-1.8 MYA
J *Homo ergaster* (KNM ER 3733), 850 cc, 1.4-2.27 MYA

K *Homo erectus* (Tattersall-Sawyer cast), 600-1200 cc, 0.1-2 MYA
L *Homo heidelbergensis* (Rhodesian man), 1220-1300 cc, 0.2-0.6 MYA
M *Homo neanderthalensis* (La Chapelle), 1500-1700 cc, 0.044-0.4 MYA
N Archaic *Homo sapiens* (Cro Magnon 1), 1200-1600 cc, 0.3 MYA-present
O *Homo sapiens* (modern cast), 1200-1400 cc, 0.3 MYA-present

Compiled by Erika (Gutsick Gibbon)

Locomotor Styles: main modes of movement across the ground or through the trees

- Bipedal (habitual / obligate): two limbs support the body weight, used primarily.
- Quadrupedal: movement with all four limbs supporting the body weight.
- Brachiator: swinging through trees by hanging from the arms.
- Pronograde: quadrupedal knuckle-walking with the body approximately horizontal.
- Palmigrade: walking with flat hands (open palms) or feet on the ground.
- Plantigrade: walking with the toes and metatarsals flat on the ground (done by humans).

Suite of Characteristics Indicative of Bipedalism: originated in late Miocene hominids

Morphology and biomechanics are linked by causal morphogenesis (Wolff's law).

- Anterior foramen magnum*: allows the skull to rest on the top of the spine.
- Sagittally-oriented iliac blades*: allows the pelvis to rest upright
- Valgus knee (bicondylar angle)*: the femur is angled to keep the knees in line.
- In-line hallux: the big toe is aligned with the other toes, aiding in walking.
- Bowl-shaped pelvis: supports the visceral organs around the abdomen.
- Lumbar lordosis (S-shaped vertebral column): supports an upright posture.
- Arched foot: three arches (medial, lateral, transverse) in the feet act as shock absorbers in walking.

* strongest indicators, since these biomechanically prevent quadrupedalism.

Autapomorphies of Hominoids (apes) include having no tail, having flexible shoulder joints, a Y5 lower molar pattern, and complex social structures.

Autapomorphies of *Homo neanderthalensis* include retreating cheekbones (zygomatics), the occipital bun, large nasal aperture, enhanced prognathism, enhanced brow ridges (supraorbital torus), platycephalic skull, angled squamosal suture, retromolar gap, elliptical foramen magnum.

Sites of Interest to Paleoanthropology: significant hominin fossil finds and associated artefacts

- Lomekwi (Kenya): oldest stone tools, attributed to *Australopithecus* or *Kenyanthropus*.
- Koobi Fora (Kenya): *A. anamensis*, *P. boisei* and *H. habilis* discovered.
- Olduvai Gorge (Tanzania): many stone artefacts discovered with hominin remains around ~2 MYA.
- Hadar (Ethiopia): Lucy (*A. afarensis*) discovered.
- Laetoli (Tanzania): footprints of *Australopithecus* discovered, most likely *A. afarensis*.
- Cradle of Humankind (South Africa): Taung child, Little Foot, Mrs Ples and other australopithecines.
- Rising Star Cave (South Africa): *Homo naledi* discovered within.
- Liang Bua (Indonesia): *Homo floresiensis* discovered within.
- Jebel Irhoud (Morocco): archaic *Homo sapiens* discovered.
- Fertile Crescent (Middle East): neolithic farming, ancient cities and civilisations from ~12 kYA.

19.4.4. Genetic Mutations in Human Evolution

The full genomes of *Homo sapiens*, *Homo neanderthalensis* and Denisovans are available, as well as all extant primates.

Neurological development:

- **ARHGAP11:** the basal form, ARHGAP11A, encodes the protein RhoGAP with nuclear localisation, found in all extant non-human mammals. A partial duplication ~5 MYA seen in *Homo sapiens*, Neanderthals and Denisovans led to them additionally acquiring ARHGAP11B, which shows mitochondrial localisation instead. It promotes basal progenitor cells (BP cells) and increases the neocortex size significantly.
- **TKTL1 (transketolase-like 1):** modern *Homo sapiens* has an arginine point mutation (K261R) while Neanderthals, Denisovans, archaic *Homo sapiens* and other extant primates have the lysine form. The human gene promotes production of basal radial glial cells (bRG cells, neural stem cells), significantly increasing upper-layer cortical neuron production and the size of the brain's gyri (ridges) in the frontal lobe.
- **NOTCH2NL**

Musculature and biomechanics:

- **PPARGC1A and promoters for MHY7:** promotes a higher proportion of slow-twitch muscle fibres rather than fast-twitch.
- **GDF8 (myostatin):** negatively regulates skeletal muscle growth.

Metabolism:

- **GULO (L-gulonolactone oxidase):** in mammals, encodes an enzyme for producing vitamin C, but is non-functional in haplorhines, occurring as the GULOP pseudogene.

19.4.4. Analytical Methods in Paleontology, Geology and Archaeology

Dating:

- Radiometric dating:
  - Carbon dating (^{14}C)
  - Potassium-argon dating (^{40}K-^{40}K)
  - Argon-argon dating (^{39}Ar-^{40}Ar)
  - Uranium series dating (^{230}Th-^{234}U, ^{234}U-^{238}U, etc)
- Dendrochronology (tree ring dating)
- Aspartic acid racemisation
- Seafloor spreading (marine magnetic anomalies, Section 15.2.12)

Climate Analysis:

- Oxygen isotope ratio cycle ($\delta^{18}O/\delta^{16}O$)
- Ice cores
- Sediment analysis
- Pollen grains

Genetic Analysis:

Y chromosomal DNA has a faster mutation rate than other chromosomes, with significant degradation in humans. However, in recent times the mutations occurred primarily in non-coding regions and no genes have been lost for 7 MA, and only one gene lost for 25 MA.

19.4.5. Paleolithic Stone Age

The ‘old’ (paleolithic) stone age ranges from the earliest known stone tools used by many hominins, including those prior to genus *Homo*. It spans from ~3.3 MYA (middle Pliocene) to 11.65 kYA (end of the Pleistocene).

The domestication of dogs from wolves (*Canis lupus*) began approximately ~40 kYA.

19.4.6. Mesolithic Stone Age

The ‘middle’ (mesolithic) stone age ranges from ~15 kYA to 5 kYA in Europe, and from ~20 kYA to 10 kYA in the Middle East.

The domestication of cats from the African wildcat (*Felis silvestris*) began ~12 kYA.

‘Cheddar Man’ is a skeleton dated to 11 kYA, found in Britain. DNA analysis found that he was likely a hunter-gatherer with blue-green eyes and black skin, with no lactase persistence.

19.4.7. Neolithic Stone Age

The ‘new’ (neolithic) stone age ranges from ~12 kYA to ~4 kYA (10,000 BC to 2,000 BC).

The Neolithic Revolution (agricultural revolution) introduced subsistence farming and settlements developing in the Fertile Crescent.

Evolutionary traits unique to humans in the Neolithic period include white skin colour, lactose tolerance, blue eyes, malaria resistance, wisdom tooth reduction, increased height, the *palmaris longus* muscle, the median artery split, amylase production, decreased bone density, slight brain size decrease.

19.4.8. Copper Age (Chalcolithic)

The Copper Age ranges from ~5,000 BC to ~3,000 BC, and it overlaps with the Neolithic stone age, as both copper and stone tools were used in this period.

The beginning of documented history, coinciding with the invention of writing, began around 4,000 BC.

'Ötzi the Iceman' is a naturally frozen mummy carbon dated to ~3,200 BC, found in the Alps.

19.4.9. Bronze Age

The Bronze Age ranges from 3,300 BC to 1,200 BC.

The Ancient Egyptians lived from ~3,100 BC to 332 BC.

19.4.10. Iron Age

The Iron Age ranges from ~1,300 BC to ~300 AD.

The Ancient Greeks lived from 1,200 BC to 323 BC.

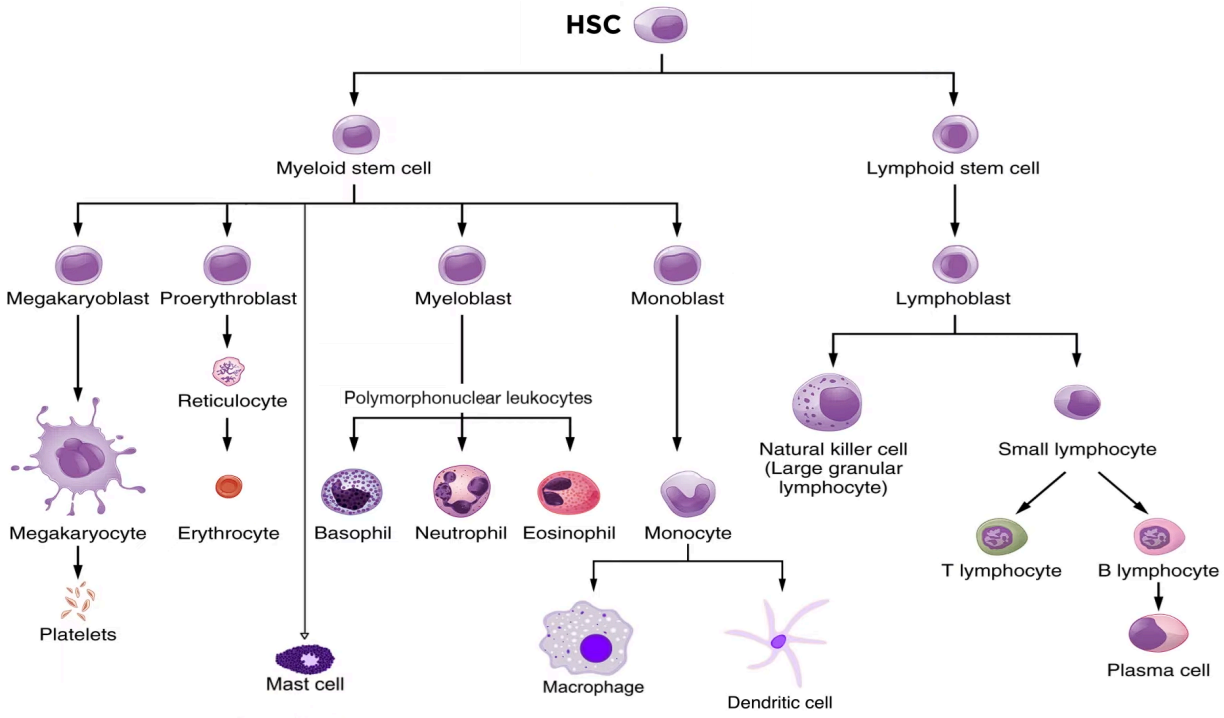
The Ancient Romans lived from 753 BC to 476 AD.

B20. MEDICAL BIOLOGY

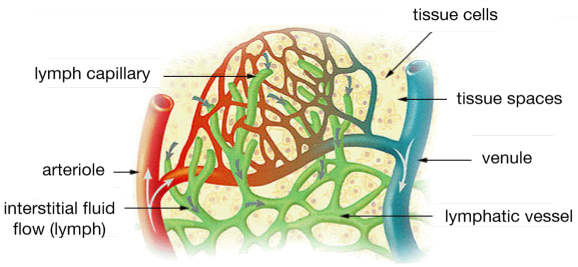
20.1. Immunology

20.1.2. Production of Immune Cells

Haematopoiesis: differentiation from multipotent hematopoietic stem cells (HSCs) in bone marrow. HSCs form two separate lineages (myeloid and lymphoid) of stem cell progenitors.



20.1.3. The Lymphatic System



Interstitial fluid containing cellular waste products and dead cell debris from tissues around the body is pulled into the lumen of lymphatic capillaries by capillary exchange. The lymphatic vessels form a network between lymph nodes. There are ~600 lymph nodes in an adult human, distributed across the body, mostly around vital organs.

Immune cells congregate at lymph nodes to help organise an immune response. Lymph node sinuses are lined with macrophages and dendritic cells that sample lymph fluid for pathogens and debris that could trigger an immune reaction. Chemokines demarcate B cell zones (follicles) and T cell zones within the lymph nodes, housing each cell type. Once a naive lymphocyte has found its cognate antigen, it undergoes costimulation and activation, followed by cell division and migration to the site of infection.

- T-cell zone: antigen-presenting dendritic cells can find a T cell with the cognate antigen
- B-T zone margin: helper T cells find B cells with the cognate antigen, forming germinal centres
- Naive B/T cells circulate through their respective zones to try finding their cognate antigen.

The spleen acts as a single giant lymph node. Healthy red blood cells are able to re-enter the venous sinus by flexing through the inter-endothelial slit in the splenic cords / cords of Billroth in the ‘red pulp’ region. Old and damaged red blood cells are unable to pass and are phagocytosed. The iron from senescent red blood cell haemoglobin is also recycled into new haem cofactors. Other filtration processes of pathogens occur in the ‘white pulp’ region, in the PALS (periarteriolar lymphoid sheath) region, serving as the spleen’s T cell zone, inside which is the marginal zone.

The thymus is a region of early T cell development after progenitor migration from the bone marrow, most active in infancy. Naive T cells leave the thymus to enter circulation through lymph nodes.

The mucosa-associated lymphatic tissues (MALTs) line the digestive tract, airways etc equip thin tissues with lymph-like structures , supplying them with antigens from microfold cells (M cells).

20.1.4. Signal Transduction in Immune Cells

Lymphocyte cell receptors are typically transmembrane proteins with an extracellular and intracellular domain. Ligand binding to the extracellular domain initiates a signalling cascade on the intracellular side.

Kinase enzymes (e.g. Lck: lymphocyte-specific protein tyrosine kinase) catalyse the phosphorylation of amino acid residues in proteins. Phosphatase enzymes reverse this process. The presence or absence of a phosphate group on a protein can promote or inhibit its activity due to conformational change, acting as protein-level regulation.

Another pathway starts via G-protein coupled receptors in the cell membrane; on ligand binding, a heterotrimeric complex (αβγ) binds to the receptor. The inactive complex is bound to GDP, but on conversion to GTP (through e.g. guanine-nucleotide exchange factors (GEFs)), the active state forms, which can phosphorylate downstream. Large/small GTPase enzymes, which convert GTP to GDP (analogous to ATP/ADP), regulate this pathway. Ras, Rac and Rho are common small GTPases.

Types of Immune Cell Receptors

- **Antigen Receptors:** pattern recognition receptors (PRRs) (e.g. Toll-like receptors (TLRs), nucleotide-binding oligomerisation domain-like receptors (NOD-like, NLRs), C-type lectin receptors (CLRs), Rig-1-like receptors (RLRs)): PRRs recognise pathogen-associated molecular patterns (PAMPs) and damage-associated molecular patterns (DAMPs). Antigen-specific lymphocyte receptors are expressed in only one unique type per lymphocyte.
- **Costimulatory Receptors:** T-cells express CD28 which is costimulatory for CD80 and CD86 on antigen-presenting cells. T-cells also express ICOS (inducible costimulator), which binds to ICOS-L on antigen-presenting cells, which helps prepare $CD4^+$ T cells. Antigen-presenting cells express CD40, which binds CD40L on T-cells, increasing CD80 expression on dendritic cells, sustaining a positive feedback loop.
- **Inhibitory Receptors:** receptor CTLA-4 (cytotoxic T-lymphocyte-associated protein 4) on T-cells binds and competitively inhibits CD80/CD86 on antigen-presenting cells, limiting CD28 effects. PD-1 binds PD-L1/L2, promoting T-cell apoptosis. Natural killer cells inhibitory receptors bind MHC class I molecules for healthy self cells, protecting them from attack.
- **Cytokine Receptors:** cytokines are soluble protein signalling molecules, determining immune cell development and inflammation vs repair.
- **Chemokine Receptors:** guides immune cells to the site of damage by chemotaxis. Most chemokine receptors are large G-protein coupled receptors, inducing changes in cell adhesion and motility (chemotaxis).

20.1.5. Cytokines and Chemokines

Cytokine families: interferons (IFN-), tumour necrosis factors (TNF-), transforming growth factors (TGF-), colony stimulating factors (-CSF), interleukins (IL-), lymphotoxins (LT-).

Types of Adaptive Immune Response: characterised by different **cytokine profiles** for fighting different types of pathogens. The immune response types are named after the type of helper T cell involved.

- **Type 1 Response (Th1 Cytokines, Th1 Helper T Cells):** promotes cellular **immunity** against intracellular pathogens (e.g. **viruses**, some bacteria) by activating $CD8^+$ **T cells and NK cells**. Th1 cytokines include **IL-2** (critical for T cell growth and differentiation), **IL-12** (activates NK cells, polarises T cells to be more type 1-like), **TNF-α** (promotes apoptosis, broad pro-inflammatory effects), **LT-α** and **LT-β** (kills chronically infected cells, activates macrophages, promotes lymphoid tissue development), **IFN-γ** (antiviral by promoting cellular MHC class I expression, activates macrophages, inhibits polarisation to type 2-like T cells).

- **Type 2 Response (Th2 Cytokines, Th2 Helper T Cells):** activates humoral responses (**antibodies**) for destruction by **B cells**, with high presence of **eosinophils, basophils and mast cells**, common for fighting **parasitic worms** (helminths) and promote wound healing/tissue repair. Th2 cytokines include **IL-4** (polarises T cells to type 2-like, promotes mast cell growth in bone marrow, stimulates eosinophils, activates B cells), **IL-5** (eosinophil development in bone marrow), **IL-10** (enhances B cell activation and antibody production, suppresses Th1 cytokine expression), **IL-13** (signals B cells to make IgE antibodies), **IL-25** (amplifies response by inducing IL-4/5/13, activates type 2 innate lymphoid cells).

- **Type 17 Response (Th17 Cytokines, Th17 Helper T Cells):** fights **fungal** and large bacterial infections by recruiting **neutrophils**. Th17 cytokines include **IL-17** (increases production of chemokine IL-8 (CXCL8, a strong neutrophil chemoattractant), **IL-22** (promotes proliferation in epithelial barrier cells in the skin/lung/gut, production of antimicrobial peptides), **IL-23** (polarises helper T cell to type 17).

- **Regulatory Response (Treg Cytokines, Regulatory T Cells):** limits the effects of an immune response to prevent excessive damage and autoimmune diseases. Treg cytokines include **IL-10** and **TGF-β** (promotes tissue repair, wound healing, anti-inflammatory, promotes Treg development, inhibits B cell proliferation, inhibits activated macrophages).

Chemokines: induces chemotaxis in immune cells to guide them towards a stimulus. Chemokine families include CC (contains two adjacent Cys residues near N-terminus) and CXC (Cys residues separated by a variable residue). Names with ‘L’ indicate ligands (e.g. CCL2). Names with ‘R’ indicate receptors (e.g. CCR2 is a receptor for either a CC or CXC chemokine).

- **CC Chemokines:** induces migration in lymphocytes and monocytes. CC chemokines include **CCL2** (MCP-1) (promotes Th2 immunity and histamine release from basophils), **CCL3** (MIP-1α) (recruits monocytes/macrophages/neutrophils, promotes Th2 immunity), **CCL4** (MIP-1β) (recruits monocytes and NK cells), **CCL5** (RANTES) (recruits eosinophils/T cells/basophils, activates NK cells), **CCL18/19** (secreted by dendritic cells, recruit T/B cells to lymph node), **CCL21** (secreted by stromal cells in lymph node to recruit dendritic cells).

- **CXC Chemokines:** induces migration in neutrophils. CXC chemokines include **CXCL1/2/3** (released by endothelial cells/fibroblasts/monocytes, promotes angiogenesis (growth of new blood vessels), activates neutrophils, stimulates fibroblast proliferation), **CXCL7** (released from activated platelets, activates neutrophils, promotes angiogenesis , **CXCL8** (IL-8) (recruits neutrophils to infected tissues), **CXCL13** (secreted by follicular dendritic cells to recruit B cells to the B cell zone (binds CXCR5)).

20.1.6. Barrier Defences of the Innate Immune System

The innate immune system is the fast-acting, nonspecific, most evolutionarily conserved (~700 MYA: plants, invertebrates, single celled organisms) part of immunity, utilising Toll-like receptors (TLRs), the complement system, phagocytosis and antimicrobial peptides.

Barrier Defences: structural and biochemical obstructions to pathogens at the surface of tissues.

Barrier surfaces are lined with epithelial cells connected by tight junctions. Barriers also possess a microbiome with commensal microorganisms (prevents colonisation, helps train the immune system).

Surfaces produce antimicrobial enzymes such as lysozymes (lyses peptidoglycan cell walls), secretory phospholipase A_2 (hydrolyses membrane phospholipids), antimicrobial peptides (AMPs) (directly kills pathogens: defensins (hydrophobic insertion into cell membrane and charge disruption), cathelicidins (pro-peptides, C-terminus acts like a defensin), histatins (fights pathogenic fungi)).

Barrier surfaces include:

- **Skin** (~2 m^2 on adults): contains a Gram-positive bacterial microbiome, an acid mantle, and dead keratinocyte epithelial cells. Skin-resident dendritic cells (Langerhans cells) quickly help mobilise innate and adaptive immune responses.
- **Respiratory tract** (~100 m^2, exposed to ~10 m^3 air per day): uses the mucociliary escalator to push mucus away from the lungs up to the larynx for coughing or swallowing. Airway epithelial cells produce lamellar bodies containing beta-defensins released into the pulmonary surfactant of alveoli, and these cells express a wide range of PRRs. The nasal microbiota is similar to that of the skin but diversified, and prevents infections spreading to the lower respiratory tract.
- **Gastrointestinal tract (GI tract):** saliva contains defensins and antimicrobial enzymes. The stomach produces gastric acid, proteases and lipases for destroying microbes. Peristalsis in the intestines forces a constant movement towards excretion, preventing colonisation. A lubricated mucus surface layer generated by goblet cells also nourishes gut microbiome commensals. The intestinal crypt contains Paneth cells, secreting antimicrobial peptides, lysozymes and cryptdins (alpha defensins).
- **Female reproductive tract:** vaginal microbiome (mainly *Lactobacillus*) prevents yeast infections and bacterial vaginosis, metabolising glycogen to lactic acid and producing hydrogen peroxide.

20.1.7. Complement System

Complement System: circulating protein effectors activated on binding to Ag/Ab complexes, a key part of the innate immune system.

Most complement proteins are zymogens (inactive enzymes, activated by cleavage of a precursor). There are 9 main complement proteins (C1-9). Cleaved products have suffix ‘a’ (smaller, anaphylatoxin) and ‘b’ (larger, binding portion) (exception: reverse convention in C2) e.g. C3 → C3a + C3b. Fragments complex together to form enzymes, e.g. C3 convertase = C4b2a = C4b + C2a; C5 convertase = C4b2a3b = C3b + C4b2a.

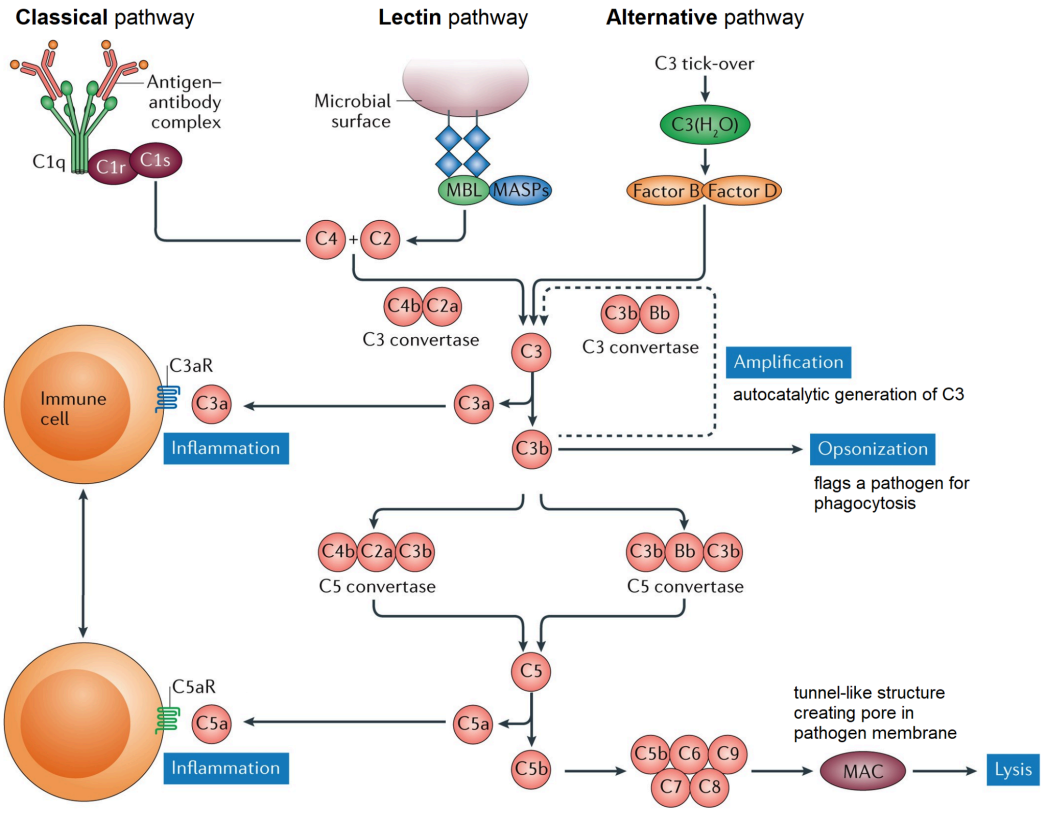
Complement is activated through classical, lectin or alternative pathways, converging at C3 cleavage:

1. **Classical:** antibody recognition, most evolutionarily recent (found in jawed vertebrates).
2. **Lectin:** acts as PRRs for binding pathogenic cell wall polysaccharides.
3. **Alternative:** uses a distinct C3 convertase (C3bBb), most evolutionarily conserved.

C3 is the most abundant complement protein (1.2 mg/ml in the serum), and is amplified autocatalytically by the alternative pathway. Downstream effects of C3 include:

1. **Inflammation:** C3a and C5a promotes inflammation and recruits neutrophils and monocytes.
2. **Opsonisation:** C3b coats pathogen surfaces and are recognised by phagocytes for phagocytosis. C3b has an exposed labile thioester for rapid attachment, but is inactivated by hydrolysis.
3. **Nucleation of membrane attack complex (MAC):** pore formation in a cell membrane.

Regulatory proteins (C1 inhibitor, decay-accelerating factor, C4 binding protein, CR1, factor H, vitronectin, protectin) limit complement overactivation and harming of host cells.



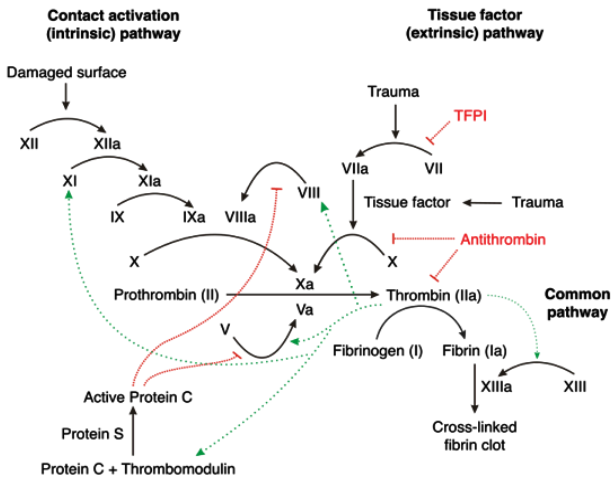
20.1.8. Inflammatory Response

Inflammation is characterised by pain, redness, swelling, heat and loss of function. Acute inflammation is an evolutionary adaptation to contain and limit early spread of infectious microbes, by destroying invading microbes, inducing local blood clotting and repairing injured tissue.

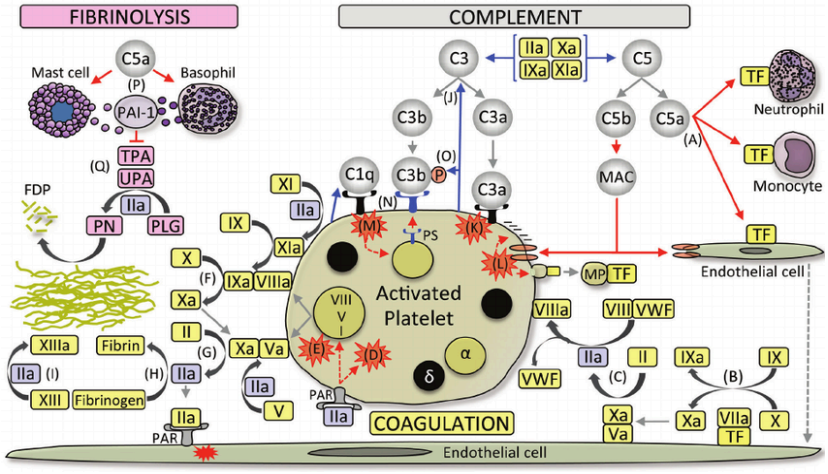
Tissue-resident immune cell PRRs sense PAMPs and DAMPs indicative of tissue damage or microbial infection. Activated macrophages release inflammatory lipid mediators (leukotrienes, prostaglandins and platelet activating factor) and nitric oxide, acting on smooth muscle cells to induce vasodilation, slowing blood flow and increased vascular permeability for neutrophil, complement and antibody migration.

In endothelial activation, cell adhesion molecule (e.g. P-selectin) expression on the lumen is increased, facilitating immune cells leaving the bloodstream. P-selectin is expressed in response to cytokine TNF-α, bacterial endotoxin, leukotriene B4, histamine or complement protein C5a. E-selectins are also expressed, which bind glycoproteins on immune cells. The immune cells then slow to a roll along the lumen wall, express adhesion molecules LFA-1 and CR3, interacting with ICAM-1/2 on endothelial cells, promoted by chemokine CXCL8, where they leave the blood vessels (extravasation) by binding of PECAM-1 (CD31).

Coagulation Cascade: forms a fibrin blood clot, trapping pathogens.



Interactions between Coagulation and Complement:



Kinin System: another protease cascade, forming vasodilator oligopeptides bradykinin and kallidin from Factor XIIa, as well as plasmin ('PN' in fibrinolysis). Plasmin also amplifies complement C3 / C5 formation.

Resolution of inflammation: controls the immune response and promotes wound healing.

The short lifespan of neutrophils (efferocytosis: phagocytosis by infiltrating macrophages after ~3 days in the tissue) help to limit inflammation. Anti-inflammatory cytokines are then released from the macrophages. Lipid mediators for resolution (derived from polyunsaturated fatty acids: resolvins, protectins, lipoxins, maresins) promote production of anti-inflammatory IL-10, stimulates efferocytosis, inhibits neutrophil chemotaxis, decreases expression of endothelial cell adhesion molecules, decreases vascular permeability. Resolution also includes angiogenesis, epithelial cell proliferation and deposition of ECM proteins (collagen). Chronic inflammation results if resolution does not occur, with many serious side effects (no return to homeostasis, rheumatoid arthritis, autoimmune disease and cancer).

20.1.9. Pattern Recognition Receptors (PRRs) and the Inflammasome

Innate immune cells contain pattern recognition receptors (PRRs), that can bind to pathogen-associated molecular patterns (PAMPs) and damage-associated molecular patterns (DAMPs). PRRs can be found facing the extracellular space on the cell membrane, free in the cytosol, or facing endosomes of ingested pathogens.

PAMPs include bacterial cell wall lipids, peptidoglycans lipopolysaccharides, bacterial flagellin, lipoteichoic acid, fungal polysaccharides, viral endosomal ssDNA/dsRNA, and unmethylated CpG-rich DNA. PAMPs are highly conserved among pathogens and tend to be essential for their survival, preventing them from evolving away from innate immune defences. DAMPs are host cell molecules that should normally not be present outside of cells (e.g. ATP), indicative of leakage from damaged cells, or ROS (reactive oxygen species), amyloid beta (prion-like) and inorganic irritants (silica, asbestos).

PAMP Detection: PRR families that can detect PAMPs include

- **Toll-like receptors (TLRs)** are homologs of the *Drosophila* Toll protein found across mammals, invertebrates and plants. TLRs are transmembrane proteins with leucine-rich repeats at the extracellular side, forming a C-shape ligand-binding domain. The intracellular side contains a TIR domain used for signal transduction. On binding to a ligand, TLRs dimerise and bring together the tails on the intracellular side, allowing docking of adaptor proteins (MyD88, TRIF, Mal, TRAM). The MyD88 and TRIF pathways lead to release of NF-κB which translocates to the nucleus, acting as a transcription factor for inflammatory cytokines (TNF-α, IL-6). The TRIF pathway also forms dimerised IRF3, a transcription factor for type 1 interferon cytokines.
- **NOD-like receptors (NLRs)** are similar to TLRs, often expressed in epithelial cells and are suited to detecting intracellular bacteria due to peptidoglycan recognition. The intracellular side contains an N-terminal CARD domain, which activates NF-κB release (similar to TRIF pathway of TLRs).
- **Rig-I-like Receptors (RLRs):** contains a helicase-like domain for binding viral RNA with two CARD termini for signal transduction. RIG-I detects viral ssRNA, while MDA5 detects cytosolic dsDNA. Activation produces transcription factors (IRF3/7, NF-κB) for type 1 interferons.

DAMP Detection using PRRs:

- **P2X7** senses extracellular ATP.
- **RAGE** senses the DNA-binding protein HMGB1 (typically found in the cytosol) and the S100 family of calcium-binding proteins.

DAMP recognition generally leads to pro-inflammatory cytokine production, like with PAMPs. Many PRRs that normally recognise PAMPs can also recognise DAMPs.

Inflammasome: a faster route of threat detection than through PRRs.

Inflammasomes are complex structures that sense threats and initiate inflammation directly. They contain sensors for a wide range of PAMPs and DAMPs (without direct binding necessary), a CARD domain, and caspase zymogens as effectors. After priming on sensing, pro-inflammatory IL-1β and IL-18 are produced, and pyroptosis is induced (inflammation-associated cell death) through N-GSDMD. Well-studied inflammasomes include NLRP1, NLRP3 (cryopyrin), NLRC4, NLRP6 and AIM2.

20.1.10. Neutrophils

Neutrophils are the main **pathogen-fighting cell** of the **innate** immune system.

HSCs undergo granulopoiesis (differentiation through myelocytes and band cells) to form neutrophils in large numbers (~10^{11} produced in the bone marrow per day). Mature neutrophils are retained in the bone marrow due to their expression of chemokine receptors CXCR2/4, which are activated by CXCL12 released by osteoblasts and stromal cells. Growth factor G-CSF promotes neutrophil exit from the bone marrow by downregulating expression of these chemokines and receptors.

Neutrophils contain many granules (secretory vesicles containing antimicrobial chemicals) which can be exocytosed, and are polynuclear (3-5 lobules). Mature neutrophils express CD11b (Mac-1), CD15, CD16/Fcγ receptor III, CD33 and CD66b. They characteristically do not express CD14. Their cytokine, chemokine and growth factor secretion profile indicate a range of roles including hematopoiesis, angiogenesis and wound healing.

Under homeostatic conditions, neutrophils enter circulation, migrate to tissues and are eliminated by macrophages (short half-life ~ 6-8 hours) to form the viscous pus fluid after an infection. Neutrophil functions in tissues include:

- Phagocytosis: clears microbes and dead cells. Occurs by either a ‘trigger’ or ‘zipper’ mechanism, using lipid remodelling and the actin cytoskeleton to shape the cell membrane into engulfing and binding the target pathogen. The pathogen is internalised into a phagosome.
- Degranulation: sequential exocytosis of chemical-filled granules. Firstly, the ‘tertiary’ granules containing myeloperoxidase/lysozyme/cathepsin/gelatinase are released to facilitate extravasation. The ‘secondary’ granules contain lactoferrin/alkaline phosphatase/NADPH oxidase for neutrophil recruitment, chemotaxis receptor upregulation, bactericide and chemoattractant generation. Lastly, the ‘primary’ granules contain elastase/MPO/defensins/cathepsin/HNP-1/2/3 / BPI, which fuse with phagosomes to deliver the contents to kill ingested pathogens.
- NETosis (formation of neutrophil extracellular traps): the chromatin in neutrophil DNA loosens and is released to form complexes with histones as well as azurophilic granules (MPO/elastase/cathepsin G). These complexes accumulate concentrations of antimicrobial agents, and pathogens entering them will be immobilised for phagocytosis.

In a tumor microenvironment, neutrophils may become polarised to tumor-associated neutrophils (TANs). They may have N1 (anti-tumor, by IFN-β) or N2 (pro-tumor, by TGF-β) phenotypes.

20.1.11. Mast Cells, Basophils and Eosinophils

Mast cells play effector roles in both innate and adaptive immunity, as well as in allergic inflammation (hypersensitivity type 1) reactions and in fighting pathogens and parasites. Together with basophils and eosinophils, they form the granulocyte white blood cells.

Mast Cells

Mast cells leave the bone marrow as progenitors and mature in peripheral tissues (not in blood), while basophils and eosinophils remain in the bone marrow until maturation. Mast cells are typically first to recognise damage from the external environment due to being localised on barrier tissues. Activated mast cells coordinate the early innate immune response, recruiting basophils and eosinophils by providing inflammatory chemotactic stimuli.

Mast cells in humans are classed as MC_T (tryptase, found in mucosa of respiratory and GI tract) or MC_{TC} (chymase, found in connective tissues including dermis and heart). Mast cells express the stem cell factor receptor CD117 (KIT), FcεR1, the IgG receptor CD32a (FcγR2a), the IFN-γ receptor CD64 (FcγR1), the receptors for complement C3a and C5a, the nerve growth factor receptor, as well as various other interleukins and chemokines. Activation of mast cells occurs by aggregation of FcεR1 when allergens bind to IgE antibodies. Activation results in exocytosis of granules containing e.g. histamine, serine proteases.

Basophils

Basophils have fewer, larger granules than mast cells and reproduce less. They express receptors for cytokines (IL-3/5R, GM-CSFR), chemokines (CCR2/3), complement (CD11b/11c/35/88), prostaglandins (CRTH2), immunoglobulins (FcεR1, FcγR2b) and TLRs. Their activation is similar to mast cells. The gp120 protein from the virus HIV is a superantigen for basophil IgE.

Eosinophils

Eosinophils have a bi-lobed nucleus with condensed chromatin. They have two types of granules (specific: cationic proteins e.g. MBP, eosinophil peroxidase, ECP, EDN and primary: Charcot-Leyden crystal protein). They also contain lipid bodies with eicosanoid synthetic enzymes, formed only after eosinophil activation. They express a diverse receptor profile.

Eosinophil degranulation is tightly regulated, and they can form DNA traps similar to neutrophils, as well as phagocytosis and respiratory bursts.

20.1.12. Macrophages

Macrophages are long-lived innate immune cells for recognition and destruction of pathogens.

Macrophages are evolutionarily conserved phagocytes found in vertebrates, exhibiting tissue-specific phenotypic plasticity with different functions under pro/anti-inflammatory conditions. Tissue-resident macrophages are present from embryonic development of hematopoietic stem cells (HSCs) in the foetal yolk sac. They can be grouped as classically-activated (M1) macrophages and alternatively-activated (M2) macrophages, polarised by Th1 or Th2 cytokines. Unique phenotypes of macrophages include tumour-associated macrophages (TAMs) and adipose tissue macrophages (ATMs), with specific expression of cytokines, chemokines, TLRs and matrix metalloproteinases (MMPs).

M1 macrophages express high levels of MHC class II, CD68/80/86.
M2 macrophages express low levels of MHC class II, high CD163/200R and MGL1/2.

Phagocytosis cascades are diverse, starting with pathogen recognition by receptors on macrophages. The pathogen is endocytosed into a phagosome, where NADPH reacts with oxygen to release H^+ and superoxide free radical ($O_2^{\bullet -}$, an ROS), while phospholipase A_2 produces free fatty acids. Together with fusion with a lysosome, this creates a broad-spectrum fatal environment for pathogens, where they are lysed, degraded and recycled.

20.1.13. Natural Killer Cells

NK cells show spontaneous cytotoxic activity towards any physiologically-stressed cell (tumour cells and virally-infected cells). As a member of the group 1 innate lymphoid cells, they secrete IFN-γ and are dependent on transcription factor Tbet. They develop from common lymphoid progenitor (CLP) cells in the bone marrow, liver and thymus. Most mature circulating NK cells are ‘dim’ for CD56 and NKp46 (low expression) and produce few cytokines/chemokines on activation, but can sometimes spontaneously lyse tumour cells. NK cells have both activating and inhibitory receptors, allowing a range of activity. The inhibitory receptors bind MHC class I molecules, present on healthy cells but can be absent in virus-infected cells, triggering activation. Tumour cells also tend to show upregulation of expression of ligands for the NK cell activating receptors.

One mechanism of cytotoxicity in active NK cells begins with the formation of an immunological synapse between the NK cell and the target cell, inducing cellular reorganisation of the NK cell’s actin cytoskeleton. The microtubule-organising centre (MTOC) is polarised and secretory lysosome is prepared for degranulation into the target cell. An alternative mechanism is ‘death receptor induced target cell apoptosis’, in which the NK cells express FasL, TNF (tumour necrosis factor) and TRAIL receptors that bind to ligands on the target cell, triggering a caspase enzyme cascade for apoptosis within the target cell.

NK cells can acquire immunological memory, unlike the rest of the innate immune system. This may occur via either antigens or cytokines.

20.1.14. Innate Lymphoid Cells

Innate lymphoid cells share similar functionality with the T cells of the adaptive immune system, but lack antigen specificity. They develop from a common lymphoid progenitor before differentiating into either NK cells (without transcription factor GATA3) or ILCs (with GATA3).

The ILC groups 1, 2, 3 and LTi are the innate counterparts of the helper ($CD4^+$) T cell groups Th1, Th2, Th17 and Treg, respectively. Similarly, NK cells mirror the killer ($CD8^+$) T cells. These phenotypes show adaptive plasticity based on the microenvironment.

ILC1 cells (including NK cells) are weakly cytotoxic, producing IFN-γ via the Tbet transcription factor, for fighting viruses and bacteria. ILC2 cells produce type 2 cytokines (IL-4/5/13) via the GATA3 transcription factor, for fighting parasites e.g. helminths. ILC3 cells produce IL-22 via RORγt transcription factor, for fighting bacteria at mucosal sites.

20.1.15. Dendritic Cells

Dendritic cells help switch from an innate immune response to an adaptive immune response.

Immature dendritic cells (DCs) are produced from differentiation of $CD34^+$ HSCs in the bone marrow. On antigen presentation to T cells, the DC may assume an immunity role (activating effector ($CD4^+$/$CD8^+$) T cells and B cell maturation) or a regulatory role (activates regulatory T cells for peripheral immune tolerance and homeostasis).

Types of DCs include: (some variations on whether human-derived or murine (mouse) derived)

- Classical / conventional / myeloid DCs (cDC): has dendrite extensions, CD11c/1a/14^+, $CD11b^{+/-}$. cDCs patrol tissues for bacterial and viral antigens, then migrate to lymph nodes for antigen presentation to T cells, producing the stimulatory mediators TNF, nitric oxide and IL-12.
- Plasmacytoid cells (PC): expresses Siglec-H and Ly6c, CD4/123^+, HLA-DR, BCDA-2, TLR-7/9. PCs also present antigens but mainly recognise viral antigens, entering the lymph nodes directly from the bloodstream. They secrete IL-12, IL-6, TNF-α and pro-inflammatory chemokines.
- Langerhans cells (LC): for immune surveillance (activation of T cells), expresses $CD207^+$ (langerin)
- Monocyte-derived cells (MC): generated upon inflammation and prime T cells in adaptive immunity.

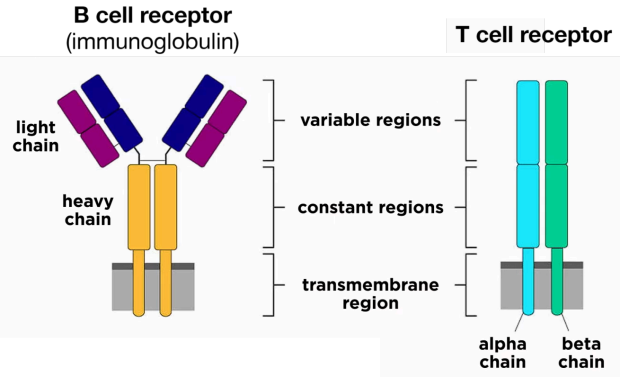
Genetic and environmental factors can alter the function of DCs, producing autoreactive killer T cells, leading to autoimmune diseases (e.g. rheumatoid arthritis, multiple sclerosis).

20.1.16. Development of Adaptive Immune Cells (T cells and B cells)

The adaptive immune system is the more advanced, evolutionarily recent (~450 MYA in jawed vertebrates) part of the immune system, responsible for antigen-specific targeting and production of long-lasting immunity against subsequent infection.

Adaptive immunity is driven by T cells (for cell-mediated immunity) and B cells (for antigenic/ humoral immunity). The B and T lymphocytes develop in the bone marrow from HSCs, with T cells migrating to the thymus, and then to the lymph nodes and spleen once matured, where they capture circulating antigens from the lymph and blood.

B-cells express both antigen receptors and antibodies on their surface, and secrete their Ag receptors into circulation where they can bind Abs. The isotypes of immunoglobulin Abs are IgM, IgD, IgA, IgE and IgG (classed by their constant domain).



V(D)J Recombination: allows complete antigen receptor diversity in lymphocytes

During the development of B and T cells, *de novo* BCR and TCR (B/T cell receptor) genes are created by recombining pre-existing gene segments, in order to translate into a huge range of combinations of receptors.

The variable (V), diversity (D) and joining (D) genes lie upstream of the constant region in the germline DNA. In humans, the loci for the heavy chain and the light chain lie on different chromosomes. Through somatic recombination, DNA is rearranged to form a VJ sequence (in light/alpha chains) or a VDJ sequence (in heavy/beta chains). A section of DNA is processed to join a random D to a random J, using RAG proteins to splice out a loop of the DJ sequence between recombination signal sequences (RSS) with 12 and 23 bp-length spacers, with the RSS orientation determining whether inversion or deletion is applied. Additional variation introduced by adding random nucleotides to the open ends. Rejoining occurs by the NHEJ pathway. The excised loops are closed and discarded on cell division. When transcribed, the mRNA introns code for a unique protein each time.

T Cell Receptor Selection and Activation:

Once a T cell has developed to express a unique receptor, they are filtered (selected) for self-antigen recognition in the thymus, with self-reactive and non-reactive receptors being eliminated by induced apoptosis. Surviving naive mature T cells exit the thymus and enter the peripheral lymphoid system, where they begin their life of being exposed to foreign peptides in MHC molecules of antigen-presenting cells (e.g. macrophages, dendritic cells, B cells) during an immune challenge (infection). Once the matching ligand has been found, they undergo clonal expansion in either the $CD4^+$ (helper) or $CD8^+$ (killer) phenotype, depending on whether the ligand was held by MHC class I (infected/cancer cell → $CD8^+$) or MHC class II (macrophage → $CD4^+$). The helper T cells activate naive B cells to produce antibodies and rejuvenate macrophages, while the killer T cells directly phagocytose infected target cells.

20.1.1. Summary of the Immune Responses

Innate Immune Response: non-specific and fast. Highly conserved evolutionarily (~700 MYA).

1. Damaged cells release signals (interferons: cytokines, histamine, bradykinin, prostaglandins) which initialises the immune system. Signals also increase MHC Class I production to increase cell transparency to immune cells.
2. **Macrophages consume bacteria by phagocytosis**.
3. Macrophages release messenger proteins to make blood vessels release water (inflammation: prostaglandin A's), bringing in complement proteins.
4. Macrophages signal for neutrophils (a type of white blood cell) to move to the site.
5. **Neutrophils release toxins** which obstruct and kill nearby pathogens and cells. Neutrophils have a short life to prevent excess damage.
6. If the pathogen persists, macrophages signal to dendritic cells to begin the adaptive immune response.

Adaptive Immune Response: specific and slow. Evolutionarily more advanced (~450 MYA).

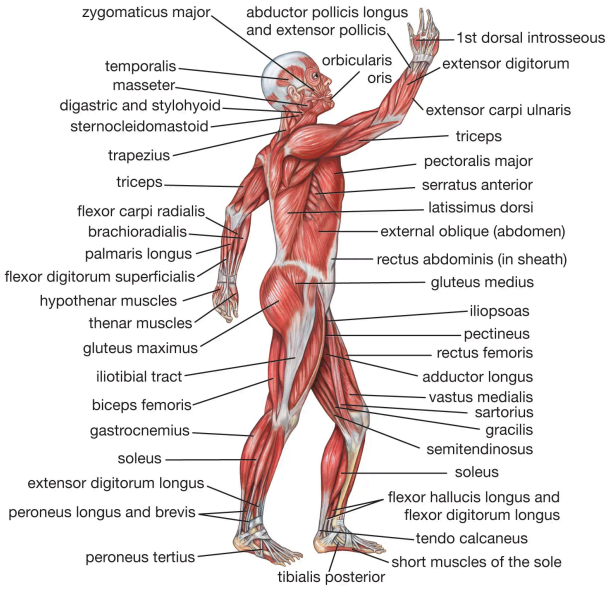
1. Dendritic cells retain the antigens of the pathogens they destroy and display them on their surface in their MHC class I windows.
2. The dendritic cells migrate to a lymph node.
3. The dendritic cell finds a (virgin / naive helper) $CD8^+$ T cell with a matching receptor.
4. The **T cell becomes activated** and rapidly divides, creating clones of (effector) helper T cells which all recognise the correct antigen. Some of the T cells become memory T cells instead and remain in the lymph node.
5. Some of the helper T cells move to the site of infection and will revitalise macrophages to kill more vigorously. The rest activate (virgin / naive) B cells.
6. The **B cells produce large amounts of matching antibodies**, stimulated to persist by the T cells.

The large influx of antibodies bind to the pathogens, which impedes their motion, making it easy for killer cells e.g. macrophages to destroy them. Once the pathogens have cleared, remaining helper cells become memory cells, the B cells continuously produce antibodies for long-lasting immunity.

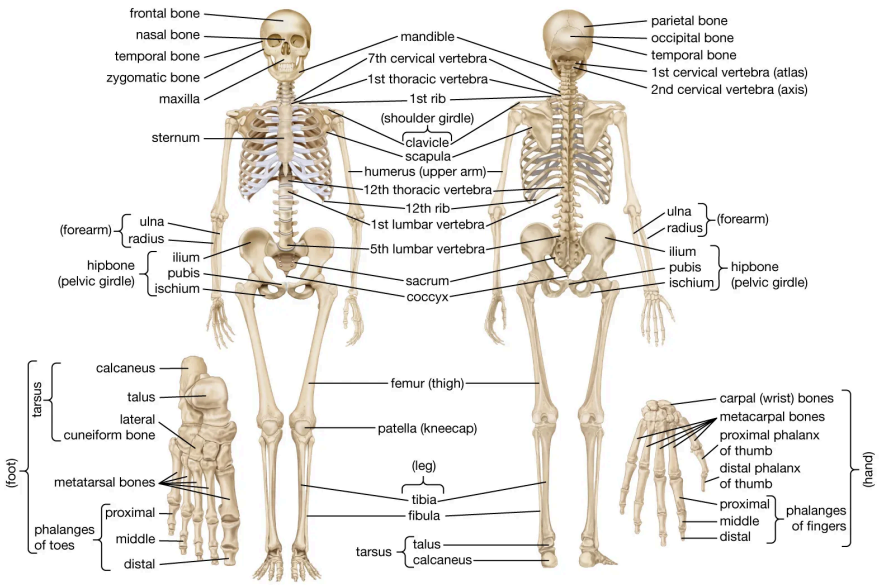
20.2. Human Anatomy and Physiology

20.2.1. Organ Systems and Their Parts

The Muscular System

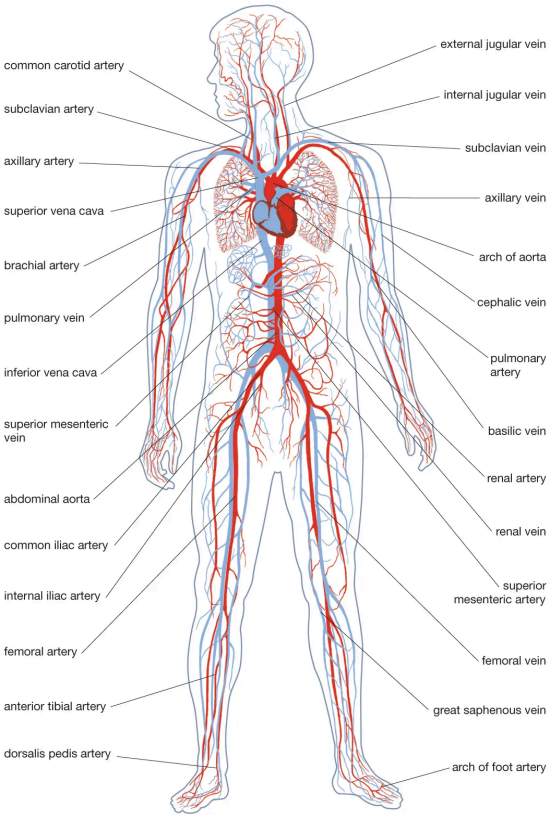


The Skeletal System



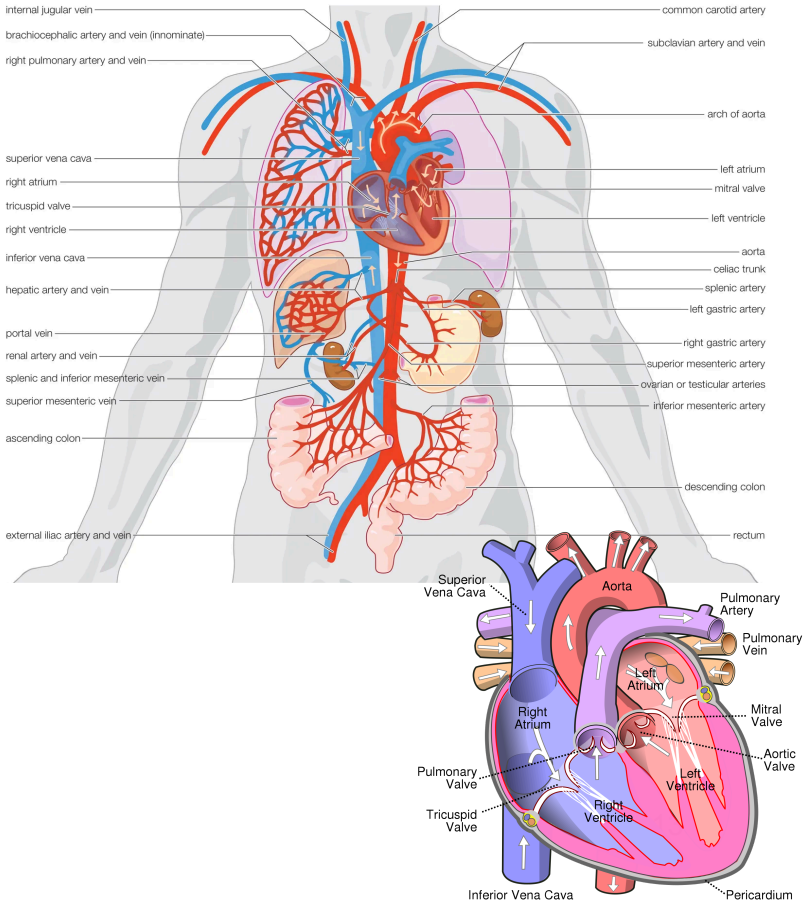
The Cardiovascular System

Arteries (carries blood **away** from the heart) and **Veins** (carries blood **into** the heart):



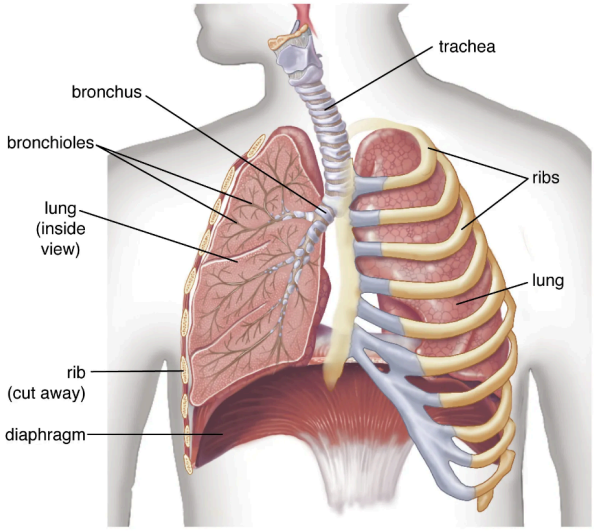
The Circulatory System and The Heart

blue: deoxygenated → lungs → **red**: oxygenated

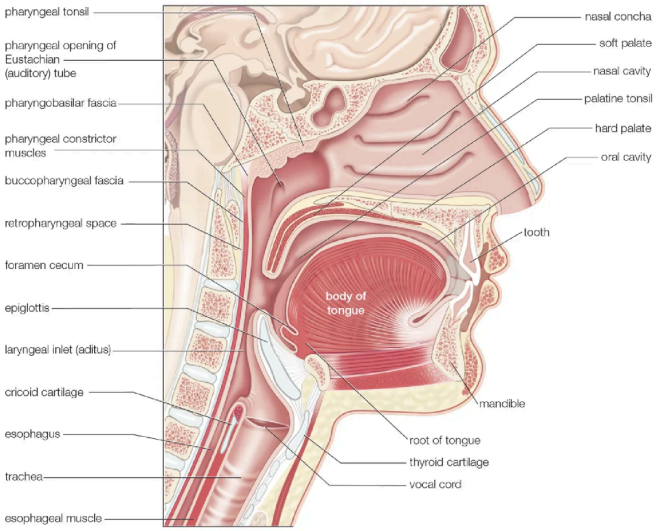


The Respiratory System

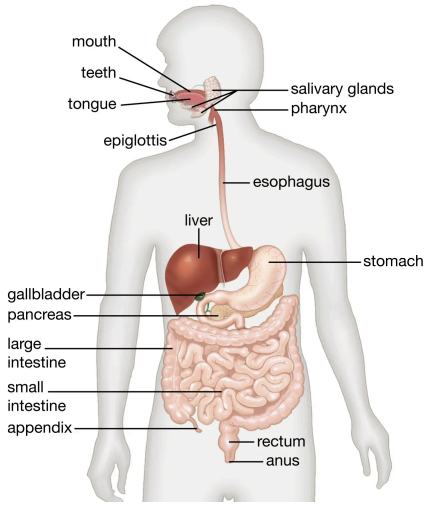
The Lungs



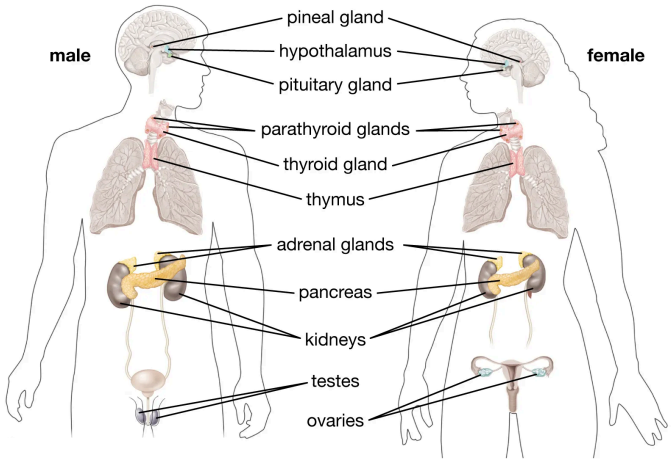
The Nose, Mouth and Throat



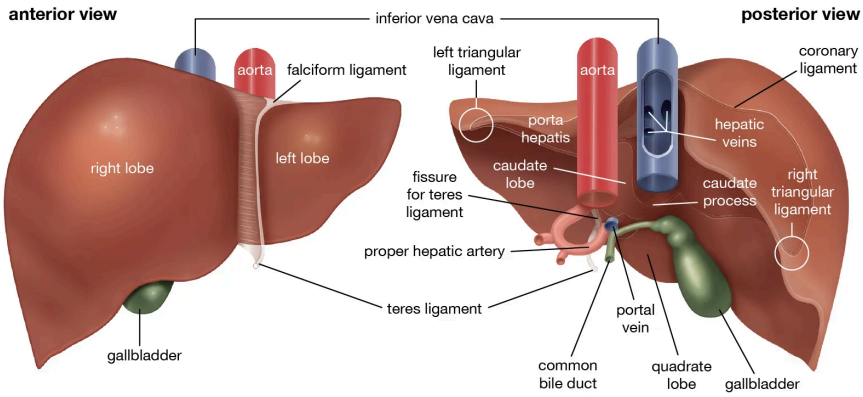
The Digestive System



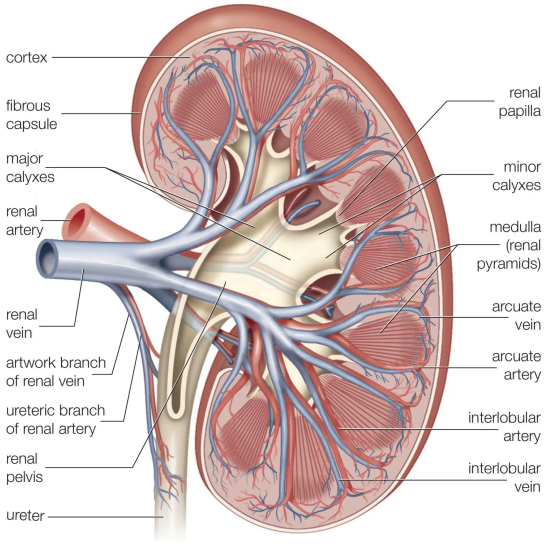
The Endocrine System



The Liver

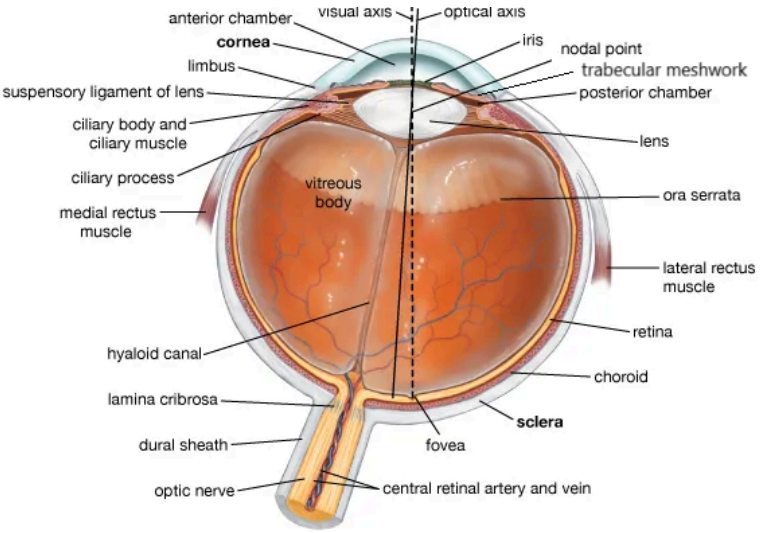


The Kidneys

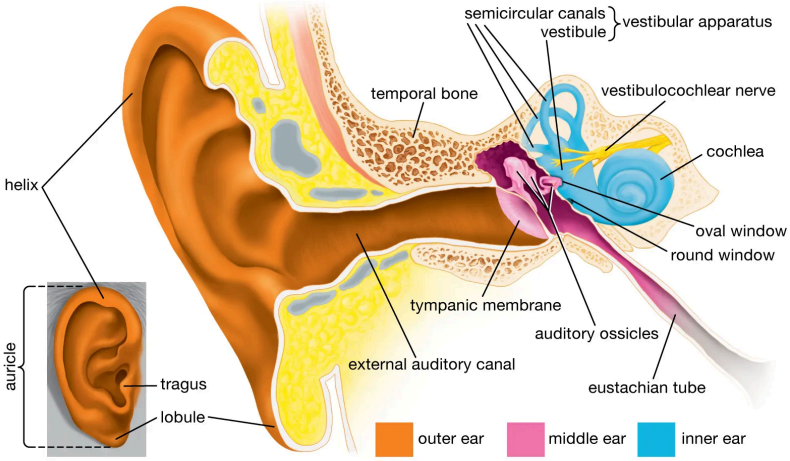


The Sensory Organs

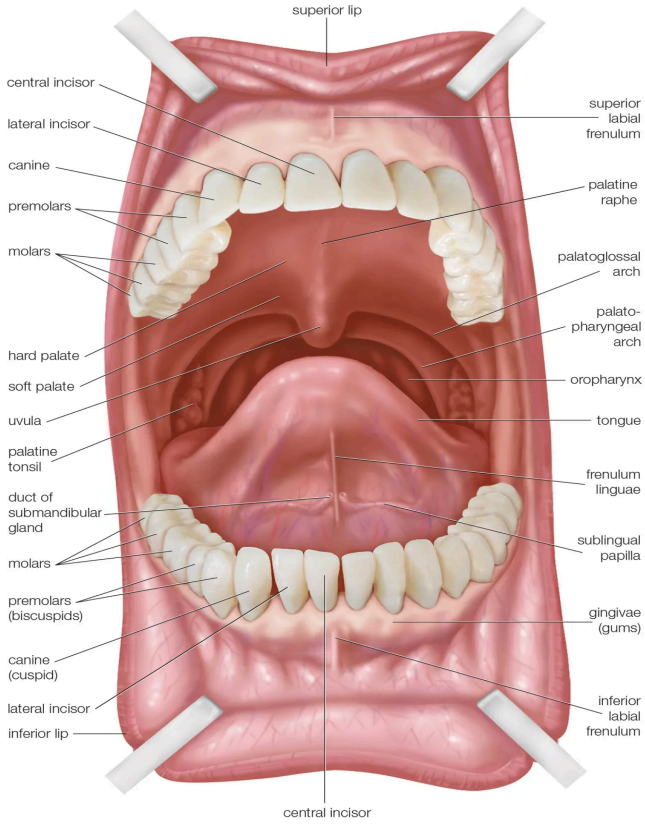
The Eye



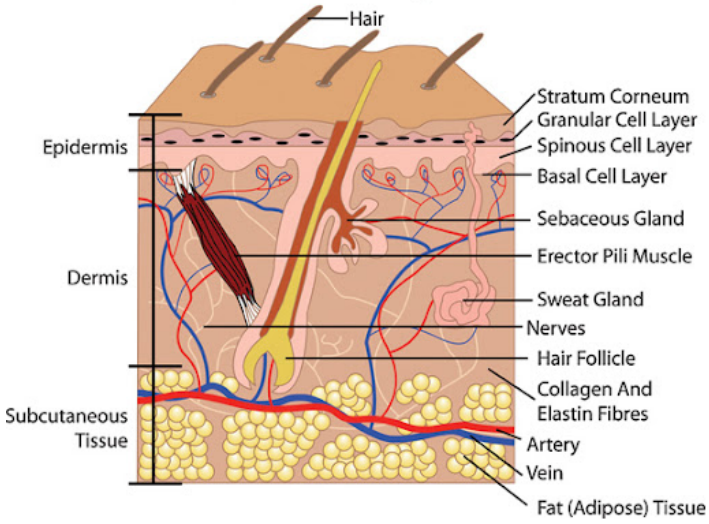
The Ear



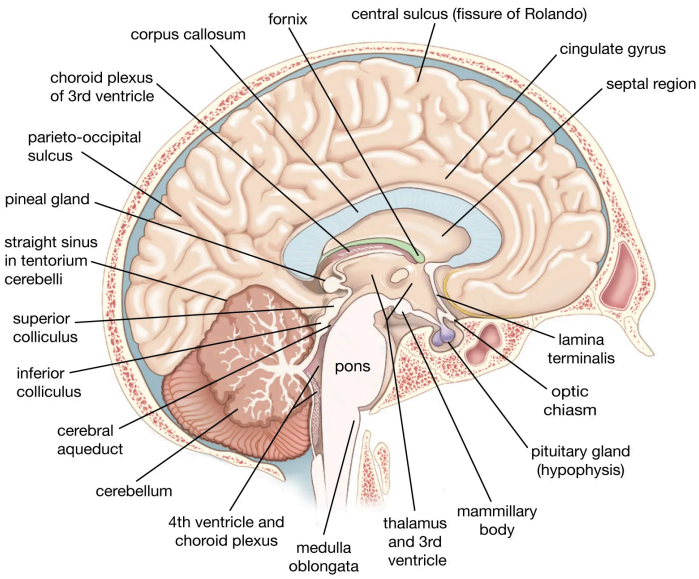
The Mouth and Teeth



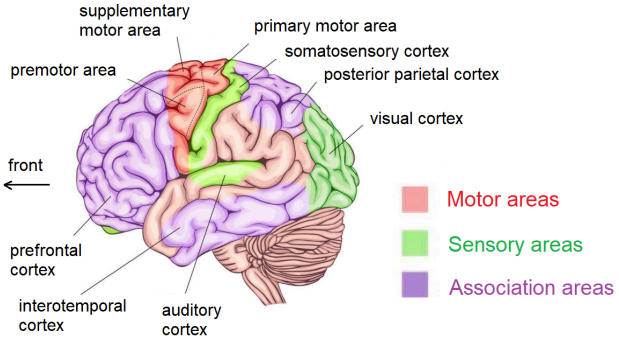
The Skin



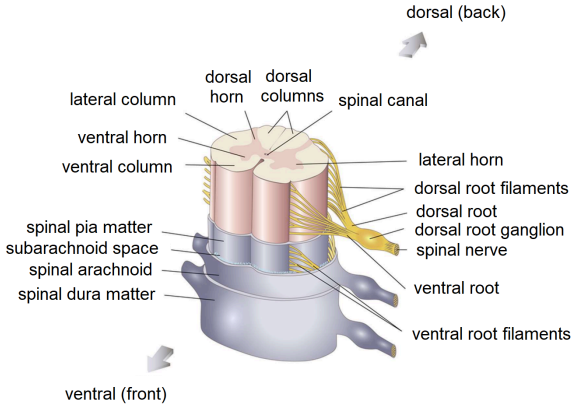
The Brain (By Structure)



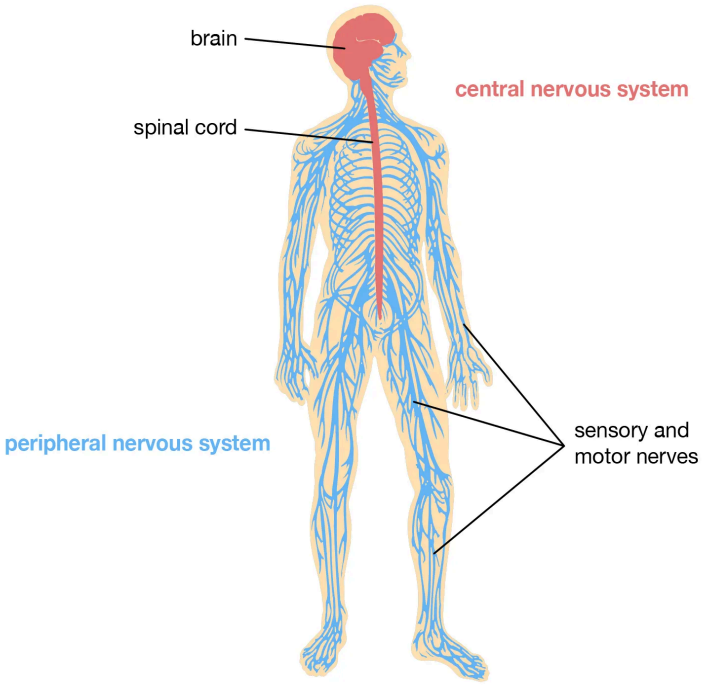
The Brain (By Function)



The Spinal Cord

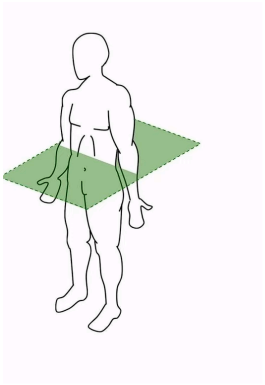


The Nervous System (CNS and PNS)

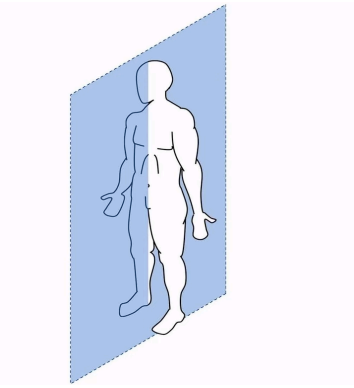


20.2.2. Terminology of Human Skeletal Anatomy

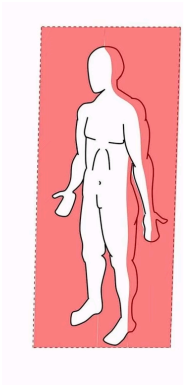
For labelled diagrams of bones in the skeleton, see Section 20.2.1.



Transverse plane
superior: above
inferior: below



Sagittal plane
medial: closer to centre
lateral: further from centre



Coronal plane
anterior / ventral: front
posterior / dorsal: back

Axial skeleton: skull, rib cage and vertebral column.
The trunk includes the central parts of the body excluding the limbs. ‘Proximal’ refers to being connected closer to the trunk. ‘Distal’ refers to being connected further away from the trunk.

The **skull** contains 8 cranial bones fused together by sutures. The spinal cord exits through the foramen magnum. The paired **facial bones** are the nasals, lacrimals, palatines, inferior nasal concha, maxilla, zygomatics (cheek bones), while the mandible (houses the teeth) and vomer are unpaired. In each quadrant of the mouth, there are 2 incisors, 1 canine, 2 premolars and 3 molars (2:1:2:3 dental formula); the most posterior molars can cause crowding (the ‘wisdom teeth’).

The **vertebral column** protects the spinal cord and consists of cervical (C1-C7), thoracic (Th1-Th12), lumbar (L1-L5) and sacral vertebrae (5 fused vertebrae in the sacrum and 3-5 fused vertebrae in the coccyx/tailbone). Cervical discs C1 (atlas) and C2 (axis) permit movement of the head.

The paired **ribs** contain 7 ‘true ribs’ (connected to the sternum), 2 ‘false ribs’ (connected via cartilage) and 2 ‘floating ribs’ (not connected to the sternum). The sternum is composed of the manubrium (most superior), body and xiphoid process.

Appendicular skeleton: limbs and pelvis. Fully symmetric about the sagittal plane.

The **pectoral girdle** (clavicle and scapula, articulated by the acromion process) connects the upper limbs (articulating with the humerus) to the axial skeleton.

The **forearms** contain two bones, the radius (distal) and the ulna (proximal). The hands, from proximal to distal, contain the carpals, metacarpals, phalanxes (proximal, middle and distal phalanges). There are eight carpals (wrist bones).

The **pelvis** consists of three fused bones, the ilium, ischium and pubis. The pelvis articulates with the femur by the acetabulum (hip socket). The leg bones are the fibula and tibia.

20.2.3. Hormones

Hormones are signallers of the mammalian endocrine system.

Class	Hormone	Tissues secreting	Effects
amino acid derivative	**adrenaline (epinephrine)**	adrenal gland (near the kidneys)	• Vasoconstriction, causing increase in blood pressure • Increases heart rate and contraction force
	melatonin	pineal gland (near the brain centre)	• Regulates the circadian rhythm (sleep-wake cycle)
	noradrenaline (norepinephrine)	adrenal gland (near the kidneys)	• Vasoconstriction, causing increase in blood pressure • Increases heart contraction force
	triiodothyronine (thyroid hormone)	thyroid gland (near the throat)	• Increases metabolism
	dopamine	substantia migra (in the midbrain)	• Regulates cellular cAMP levels
eicosanoid	**prostaglandins**	from all cells with nuclei	• Vasodilation, causing decrease in blood pressure
	leukotrienes	from white blood cells	• Increases vascular permeability
	prostacyclins	from endothelial cells	• Vasodilation • Inhibits platelet production
	thromboxane	from platelets	• Vasoconstriction • Inhibits platelet aggregation
peptide	**antidiuretic hormone (ADH, vasopressin)**	pituitary gland (near the hypothalamus)	• Water retention in kidneys • Moderate vasoconstriction
	follicle-stimulating hormone (FSH)	pituitary gland (near the hypothalamus)	• Promotes maturation in the ovary (females) • Promotes production of sperm cells (males)
	glucagon	pancreas	• Glycogenolysis, increasing blood glucose level
	human growth hormone (hGH)	pituitary gland (near the hypothalamus)	• Cell growth and division
	insulin	pancreas	• Glycogenesis and glycolysis, decreasing blood glucose level • Synthesis of triglycerides from lipids in fat cells
	luteinising hormone (LH)	pituitary gland (near the hypothalamus)	• Ovulation (females) • Production of testosterone (males)
	oxytocin	pituitary gland (near the hypothalamus)	• Lactation (release of breast milk) • Cervix contraction (orgasm)
androgen steroid	**testosterone**	testes and ovaries	• Libido and maturation of sex organs (males) • Muscle mass gain
	dihydrotestosterone (DHT)	testes, ovaries and liver	• Promotes cell growth and division in the penis • Can miniaturise hair cells (usually a side effect as a metabolite of testosterone)
oestrogen steroid	**estradiol**	testes and ovaries	• Maturation and release of an egg • Promotes growth in the female sex organs
glucocorticoid steroid	**cortisol**	adrenal gland (near the kidneys)	• Inhibits glucose uptake in muscle and fat cells
progestogen steroid	**progesterone**	ovaries, adrenal gland, placenta (when pregnant)	• Prepares the uterus lining for a fertilised egg
secosteroid (vitamin D_3)	**calcitriol**	skin, kidney tubule	• Increase absorption of calcium and phosphate

20.2.4. Neurotransmitters

Neurotransmitters can induce or inhibit action potentials in neurons.

Class	Neurotransmitter	Effects
catecholamine	**dopamine**	• Motivation, reward, pleasure, mood regulation, motor control
	epinephrine (adrenaline)	• Increases heart rate and alertness, redirects blood flow to muscles
	norepinephrine (noradrenaline)	• Constricts blood vessels, raises blood pressure • With epinephrine, initiates the 'fight or flight' response
indolamine	**serotonin** (5-HT: 5-hydroxytryptamine)	• Mood regulation, sleep, appetite
	melatonin	• Sleep-wake cycle, circadian rhythm
	histamine	• Immune response, alertness, wakefulness
trace amine	**phenethylamine and *N*-methylphenethylamine**	• Mood regulation
	tyramine	• Releases norepinephrine
	3-iodothyronamine	• Thyroid hormone metabolism, thermoregulation
	octopamine	• Arousal, movement
	tryptamine	• Precursor to serotonin
amino acid	**glutamate**	• Main excitatory neurotransmitter, all synaptic transmission
	aspartate	• Excitatory neurotransmitter
	D-serine	• NMDA glutamate receptor coagonist, synaptic plasticity, learning
	γ-aminobutyric acid (GABA)	• Main inhibitory neurotransmitter
	glycine	• Inhibitory neurotransmitter, motor/sensory functions (in spinal cord)
gastro-transmitter	**nitric oxide (NO)**	• Cell signalling, regulation, vasodilation, anti-oxidant (in the GI tract)
	carbon monoxide (CO)	
	hydrogen sulfide (H_2S)	
neuropeptide	**oxytocin**	• Love, emotional bonding
	somatostatin	• Inhibits release of various hormones e.g. GH, TSH, insulin, glucagon
	substance P	• Pain transmitter, neuroinflammation
	cocaine and amphetamine regulated transcript (CART)	• Energy homeostasis, controls appetite, found in the reward system • Mimics cocaine and amphetamine alone, but inhibits their effects when these substances are present • Cocaine is an epigenetic promoter of CART gene transcription
	opioid peptides e.g. β-endorphin	• Neuromodulators in the pain/reward system, • Binds to opioid receptors
purine	**adenosine triphosphate (ATP)**	• Pain signalling, synaptic plasticity, and neuroprotection
	adenosine	• Reduces neural activity, promotes sleep
others	**acetylcholine (ACh)**	• Released at neuromuscular junctions to effect motion • In the CNS, promotes cognition, memory and learning
	anandamide	• Binds to cannabinoid receptors to affect mood, pain perception, appetite, and neuroprotection

20.1.5. Blood Types

The ABO system: based on the presence of two common antigens, A and B.

Group	**Antigens** on erythrocytes	**Antibodies** flowing in the plasma	Can donate to:	Can receive from:
A	A	anti-B	A, AB	O, A
B	B	anti-A	B, AB	O, B
AB	none	anti-A and anti-B	AB	O, A, B, AB
O	A and B	none	O, A, B, AB	O

Blood can be donated only if the recipient does **not** have antibodies for the incoming (donor's) antigens, otherwise the antibodies will attack the corresponding blood cells.
The genes for A and B antigens are codominant (so type O is recessive).

The Rh system: based on the presence of rhesus (Rh) antigens, primarily RhD.

If RhD antigens are present, then the blood type is positive (+).
If RhD antigens are absent, then the blood type is negative (-).
Recipients with Rh- blood can **only** receive blood from Rh- donors.
The gene for RhD+ is dominant (so Rh- is recessive).

Rare Blood Types: the ABO and Rh(+/-) make up 8 different types based on antigen composition, but some people may be missing Rh antigens other than RhD, which require special attention. In the most extreme case, a person's blood may contain no Rh antigens at all (Rh-null). Compatible blood donors are much more difficult to find due to their rarity.

20.3. Pharmacology

20.3.1. Some Important Receptors and Signalling Molecules

Toll-Like Receptors (TLRs)

Found on the surface of macrophages and dendritic cells.

NF-κB (Nuclear Factor Kappa-light-chain-enhancer of Activated B Cells)

G-protein Coupled Receptors (GPCRs)

Adenylyl Cyclase, ATP and cAMP

Protein Kinases

PAMPs and DAMPs

20.3.2. Neurotransmitter Receptors

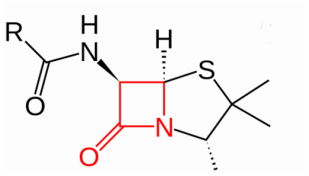
The binding affinity K_i of a substrate/ligand/drug S with respect to an enzyme/receptor E is the equilibrium constant for

$$ES \underset{k_{\text{on}}}{\overset{k_{\text{off}}}{\rightleftharpoons}} E + S \qquad K_i = \frac{k_{off}}{k_{on}} = \frac{[E][S]}{[ES]} \qquad \text{typical units: μM / nM}$$

K_i is identical to the Michaelis-Menten constant (Section 17.4.4) for the system assuming no further reaction ($k_{\text{cat}} = 0$). When binding is stronger, K_i is smaller.

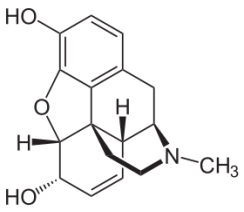
- Adrenergic receptors (α_{1A}, α_{1b}, α_{1c}, α_{1d}, α_{2a}, α_{2b}, α_{2c}, α_{2d}, β_1, β_2, β_3): targeted by catecholamines and beta-blockers.
- Muscarinic acetylcholine receptors (mAChRs) (M1, M2, M3, M4, M5).
- Dopaminergic receptors (D_1, D_2, D_3, D_4, D_5).
- GABAergic receptors ($GABA_A$, $GABA_{B1a}$, $GABA_{B1\delta}$, $GABA_{B2}$, $GABA_C$).
- Glutamatergic receptors (NMDA, AMPA, Kainate, $mGluR_1$, $mGluR_2$, $mGluR_3$, $mGluR_4$, $mGluR_5$, $mGluR_6$, $mGluR_7$)
- Glycine receptors: mediates inhibition in the spinal cord and brainstem.
- Histaminergic receptors (H_1, H_2, H_3).
- Opioidergic receptors (μ, δ_1, δ_2, κ).
- Serotonin receptors ($5\text{-}HT_{1A}$, $5\text{-}HT_{1B}$, $5\text{-}HT_{1D}$, $5\text{-}HT_{1E}$, $5\text{-}HT_{1F}$, $5\text{-}HT_{2A}$, $5\text{-}HT_{2B}$, $5\text{-}HT_{2C}$, $5\text{-}HT_3$, $5\text{-}HT_4$, $5\text{-}HT_5$, $5\text{-}HT_6$, $5\text{-}HT_7$)

20.3.1. Some Common Drug Classes



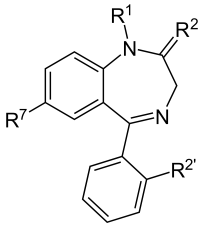
Beta-Lactams

(Antibiotics: penicillin, ampicillin…)



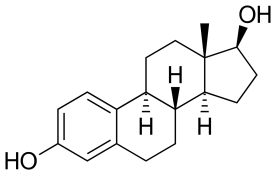
Opioids

(Potent analgesics: oxycodone, morphine, cocaine, heroin, fentanyl…)



Benzodiazepines

(Tranquilisers: diazepam, oxazepam alprazolam…)



Steroids

(Synthetic hormones: cholesterol, estradiol, cortisol, testosterone…)

Beta-lactam antibiotics bind to PBPs (penicillin binding proteins) and inhibit their usual enzymatic activity of peptidoglycan synthesis for building cell walls. The carbonyl group in the lactam ring undergoes nucleophilic attack by the hydroxyl group of the active site serine residue, forming a covalent bond. Antibiotic resistance can occur naturally when mutations produce beta-lactamase enzymes, which cleave the amide bond in e.g. penicillin, completely breaking its ring structure.

Opioids are agonists to opioid receptors, found on the surface of nerve cells. Inhibition of the release of certain neurotransmitters, such as ‘substance P’ (an 11-member neuropeptide pain signaller), occurs due to activation during binding to μ-opioid receptors. The ‘morphine rule’ is a rule of thumb for the pharmacophores (ligand motifs) to trigger this effect: (1) a benzene ring (2) attached to a 4° carbon connected by (3) an ethyl linker to (4) a tertiary amine. An exception is fentanyl (no 4° carbon) which still fits into the receptor. Opioids can produce intense psychoactive effects including euphoria and are addictive with high risk of dependency/abuse.

Benzodiazepines bind to GABA-A receptors in neural CNS synapses. This increases the activity of adjacent GABA receptors. The frequency of Cl^- ion channel opening increases while bound, increasing the hyperpolarisation potential of the cell and decreasing neuronal excitability. Some drugs also act directly on the Ca^{2+} voltage-gated ion channels. They are used to treat anxiety, insomnia and seizures by promoting a relaxed state of mind.

Steroids bind to steroid hormone receptors in various cell cytoplasms. A conformational change permits translocation to the nucleus where it binds to hormone response elements of DNA. The steroid then acts as a transcription factor for genes relating to metabolism. Steroids may be corticosteroids (anti-inflammatory) or anabolic (promoting muscle growth).

20.3.2. Naming and Regulation of Illicit Drugs

Drugs can be named using:

- Chemical name: the unique systematic IUPAC name of the drug, often long (and can induce chemophobia in the general public). e.g. 2,4-diamino-6-piperidinopyrimidine 3-oxide.
- Generic name: a trivial but unique name assigned to drugs, used as a practical substitute for the chemical name. e.g. minoxidil.
- Brand name: a trademarked name for use when sold by a pharmaceutical company, used for marketing e.g. *Regaine*. The same active ingredient can have different brand names, since it can be sold by different companies, and a brand name drug may also include multiple active ingredients (a combination drug). Often the most recognisable.
- Street name: illicit drugs may be given informal names by users to avoid disclosing their possession to law enforcement. e.g. ‘Special K’ for ketamine.

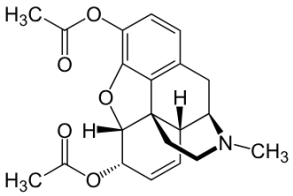
Drugs are often regulated in order to limit their availability to the general public. The methods of regulation vary by country and their legality is roughly (not always) determined by their potential to cause harm and induce dependency.

- In the USA, the DEA (Drug Enforcement Agency) sets five schedules, including Schedule I: no medical use (heroin, LSD, ecstasy/MDMA, marijuana/weed), Schedule II: limited medical use (cocaine, oxycodone, fentanyl, stimulants), down to Schedule V: negligible risk (cough medicines).
- In the UK, the Misuse of Drugs Act (1979) sets three classes. Class A (heroin, cocaine, ecstasy), Class B (cannabis/weed, amphetamine, ketamine) and Class C (steroids, benzodiazepines).

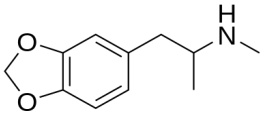
Different countries may restrict different substances, and there can be debate about the class status of certain substances (e.g. cannabis, alcohol, caffeine). Since studies on the long-term effects of drugs are often tentative and inconclusive, their interpretation and subsequent legislation is often subject to biassed politics. For example, cannabis is widely considered to be minimally harmful despite being classed severely, while alcohol (which has various known health implications) is typically not classed as a controlled drug at all. Propaganda campaigns can seek to prohibit some drugs by creating and/or targeting stereotypes about the socioeconomic or racial groups that use them, a form of scientific racism (pseudoscience).

New psychoactive substances (NPS) are often derived from common drug motifs, with slight functional group or steric modifications (e.g. DMT → 5-MeO-DMT, 4-HO-MET; MDMA → synthetic cathinones), or to sidestep prohibitive laws by passing as ‘research chemicals’. However, many countries have laws with “catch-all” clauses, outlawing drug-like substances with similar structure to known drugs, regardless of whether it is explicitly listed.

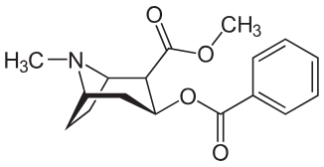
20.3.3. Illicit Recreational Drugs



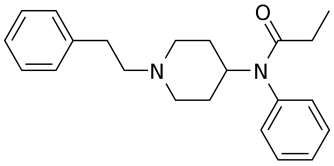
Heroin



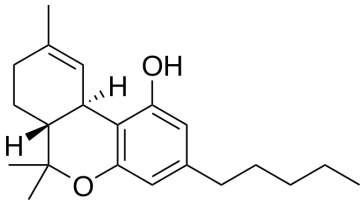
3,4-methylenedioxy-methamphetamine (MDMA)



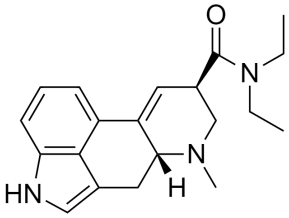
Cocaine



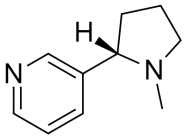
Fentanyl



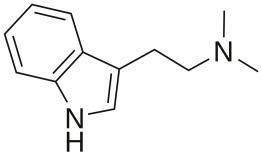
Tetrahydrocannabinol (THC)



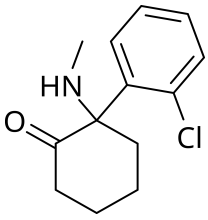
Lysergic acid diethylamide (LSD)



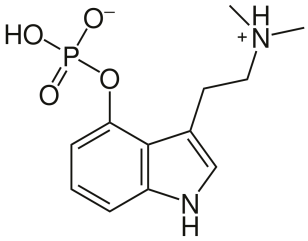
Nicotine



Dimethyltryptamine (DMT)



Ketamine



Psilocybin

20.3.3. Drug Administration, Pharmacokinetics and Pharmacodynamics

The 'Six Rights' of Drug Administration: principles to ensure safe usage of medical drugs.

1. **Right patient.**
2. **Right medication.**
3. **Right dose.**
4. **Right route of administration**, which may be:
   - enteral (into mouth/stomach e.g. sublingual (under tongue), buccal (against cheek))
   - topical (on the skin)
   - parenteral (through skin e.g. intravenous, intravenous, intradermal, subcutaneous).
5. **Right time of delivery**. May be at regular intervals or with meals.
6. **Right documentation**. Record the medication time, dosage, route and any observations.

Patients also have the 'Right to Refuse' any medication for any reason (best noted as 'declined at this time').

Pharmacokinetics: how the body impacts the drug's ability to deliver its effect.

ADME: absorption, distribution, metabolism, excretion. The kinetics of drug delivery depend on the diffusion into the bloodstream (bioavailability), distribution through the bloodstream (affinity for off-target receptors), resistance to degradation (in e.g. first pass through the liver) and ease of excretion once the drug has delivered its effect.

Lipinski's Rule of Five (RO5): a rule of thumb for the properties of most orally available drugs:

"No more than 5 H-bond donors; No more than 10 H-bond acceptors; M_w < 500 Da; log $P_{oct} < 5$."

Pharmacodynamics: how the drug impacts the body once delivered.

Drugs interact with parts of proteins (active sites on receptors or enzymes). These sites typically accept a ligand in order to become activated and propagate a signal in some way. Drugs can be agonists (mimicking the native ligand) or antagonists (binds to and inhibits the receptor). Binding may be competitive, uncompetitive or noncompetitive (see Section 17.1.3).

The binding affinity is higher when the drug fits well into the active site and makes more electrostatic interactions. Drugs can be tailored to specific receptors if their structure is known by modelling their interactions (molecular docking) or testing variations (structural-activity relationship studies (SARs)). The mechanism of action concerns the way in which the receptor-ligand complex reacts, potentially involving a chemical reaction in which the drug is metabolised to other substances.

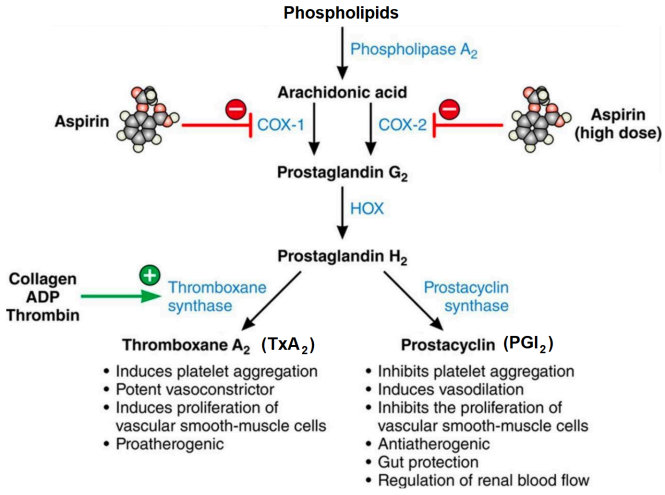
Measures of a drug's pharmacodynamic effects include:

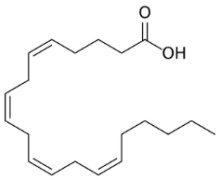
- Drug efficacy: the maximum effect of a drug as the dose approaches infinity (often qualitative). Determined by the drug's ability to act as an agonist or antagonist.
- Drug potency (median effective dose, ED_{50}): concentration for it to produce 50% of the maximum effect. Determined by the drug's affinity for the target receptor.
- Drug lethality (median lethal dose, LD_{50}): concentration that kills 50% of a test population (animals in preclinical trials) exposed to it. Determined by the types of receptors targeted.
- Therapeutic index = LD_{50} / ED_{50}.

20.3.4. Pharmacodynamics of Common Drugs

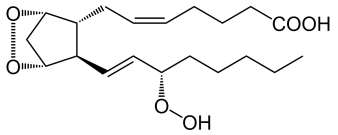
NSAIDs (Non-Steroidal Anti-Inflammatory Drugs): aspirin, acetaminophen, ibuprofen, naproxen, celecoxib

NSAIDS are antipyretic (reduces fever), analgesic (reduces pain) and anti-inflammatory. They operate by inhibiting the cyclooxygenase (COX) enzymes, preventing synthesis of inflammatory prostaglandins.

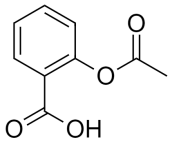




Arachidonic Acid



Prostaglandin G_2



Aspirin

Aspirin binds irreversibly, by acetylation of homologous Ser-529 in COX-1 and Ser-516 in COX-2, nullifying the nucleophilicity of the active site, while ibuprofen and naproxen bind reversibly. Celecoxib inhibits COX-2 selectively. Acetylated COX-2 retains some functionality and generates aspirin-triggered lipoxins (ATLs) such as 15-epi-lipoxin A_4 (an epimer of a natural lipoxin), a potent anti-inflammatory agent.

Antihistamines (Allergy-Alleviating Drugs)

Histamine is released by mast cells in tissues and by basophils in the blood. Its binding to the H1 receptor triggers the inflammatory cascade and then an innate immune response. Antihistamines prevent these causes of allergies from occurring.

1st generation antihistamines cross the blood-brain barrier and have psychoactive side effects (sedation and anticholinergic effects). 2nd generation antihistamines do not cause these effects as they activate protein pumps which remove them from the brainspace.

DMARDs (Disease-Modifying Antirheumatic Drugs): methotrexate, tofacitinib citrate, infliximab

DMARDs treat Rheumatoid arthritis (RA). They are classified as either small molecule (activate / deactivate proteins) or biologic (antibodies to target the proteins) DMARDs. Biologic DMARDs cause less side effects than small molecule DMARDs due to their higher specificity but are more complex to produce and require injection. They are both immunosuppressive.

Methotrexate is a prodrug, being modified *in vivo* by mimicking folic acid as a substrate to the FPGS enzyme, becoming activated to MTXPG. This inhibits the ATIC enzyme, leading to accumulation of adenosine, stimulating adenosine 2A receptors, providing anti-inflammatory effects, by reducing cytokine production and preventing activation of immune cells.

Glucocorticoids Immunosuppressants (Steroidal Drugs): hydrocortisone, dexamethasone

The glucocorticoid nuclear receptor NR3C1 is activated by glucocorticoids, and the complex is translocated into the nucleus through a nuclear pore. This complex has epigenetic activity. Gene expression of COX enzymes is decreased, reducing inflammation. Expression of transcription factor NF-κB is also decreased, reducing cytokine production and suppressing immune activity. A wide range of other signalling pathways are also affected.

Non-Glucocorticoid Immunosuppressants: cyclosporin, tacrolimus, rapamycin (cyclic oligopeptides)

- **Immunophilin binding drugs:** immunophilin typically isomerises proline and are the receptor for these drugs, found in the cytosol of T-cells. The complex then binds to calcineurin (or mTORC1 in the case of rapamycin), preventing it from acting as a transcription factor for interleukin 2 (IL-2), impeding activity, maturation and differentiation of T cells.
- **Cytostatic drugs:** these drugs inhibit dihydrofolate reductase, preventing thymidine synthesis (required for *de novo* DNA synthesis), preventing cell division in immune cells.
- **Antilymphocyte antibodies:** these are specific antibodies for CD52 receptors on mature B/T cells, triggering their immediate apoptosis, preventing their function.
- **Monoclonal antibodies:** binds even more specifically, to CD3 receptors, similarly triggering apoptosis of T cells.

Immunostimulants:

- **Colony stimulating factors:** enhance production and release from the bone marrow of white blood cells.
- **Interleukins:** recombinant forms of immune-stimulating chemokines.
- **Interferons:** binds to interferon receptors on plasma membranes of white blood cells, activating transcription factors for immune stimulating proteins such as cytokines and growth factors for natural killer cells, B cells and T cells.
- **Vaccines:** triggers a slow immune response to an unfamiliar pathogen, priming the immune system for future attack by formation of memory cells.

Bronchodilators: salbutamol, levalbuterol, triotropium bromide

Used to treat chronic breathing disorders such as asthma, emphysema, bronchitis by relaxing smooth muscle cells to increase the radius of the lumen.

- **β_2 adrenoreceptor agonism:** non-selectively activated by adrenaline/noradrenaline. The β_2 receptor is specifically activated via the Gαs pathway, in which a heterotrimer dissociates into Gαs and a GβGγ heterodimer. Gαs binds to adenylate cyclase and increases its conversion of ATP into cAMP, increasing binding to protein kinase A, activating myosin phosphatase, which phosphorylates the myosin light chain, preventing it binding to actin, reducing contractility. GβGγ also plays a role by activating K^+ channels to hyperpolarise (efflux K^+) the cells, reducing mobilisation of Ca^{2+}, reducing calmodulin activation, further inhibiting muscle cell contraction.
- **Muscarinic receptor antagonism:** non-selectively activated by acetylcholine. The M3 receptor is activated via the Gαq pathway, in which a heterotrimer dissociates into Gαq and a GβGγ heterodimer. Gαq binds to phospholipase C, which hydrolyses PIP_2 into DAG and IP_3. DAG activates protein kinase C (acts similar to the above protein kinase A). IP_3 binds to its receptor on the endoplasmic reticulum, liberating Ca^{2+} from cell stores, increasing contractility.

Antitussives: codeine, dextromethorphan, benzonatate (tessalon)

Used to block the coughing reflex. They only treat the symptoms (not the root cause).

- **μ opioid receptor agonism:** codeine is a prodrug, metabolised by enzymes UGT2B7 and CYP2D6 to form codeine 6-glucuronide and morphine respectively, which together signal the μ opioid receptor on neurons in the GI tract via the Gαi pathway. The resulting K^+ efflux due to GβGγ release reduces neurotransmission in the reflex arc of the medulla. The risks of addiction mean that codeine is only administered in extreme cases.
- **Voltage-gated sodium channel blocking:** benzonatate blocks VGSCs in the bronchi and alveoli reducing action potential frequency output from the afferent cough pathway via the vagal fibres. It also blocks VGSCs in the medulla, affecting processing of the cough reflex.

Decongestants: phenylephrine, oxymetazoline, tramazoline, ephedrine

Used to relieve nasal congestion by restricting blood flow to the mucosal membranes of the nose.

- **α_1 adrenergic receptor agonism:** activates the Gαq pathway, which increases contractility.

Cannabinoids: tetrahydrocannabinol (THC), cannabidiol (CBD)

From the *Cannabis* plant, tetrahydrocannabinolic acid (THCA) is extracted and decarboxylated to THC.

- Δ^9-tetrahydrocannabinol is a partial agonist to cannabinoid receptor CB_1 (a G-protein receptor).

20.4. Diseases

20.4.1. Bacterial Infections

Food Poisoning: *E. Coli*

Certain strains of *E. Coli* are enterohemorrhagic, releasing shiga toxin (Stx).

Cholera: *Vibrio cholerae*

Typhoid Fever: *Salmonella typhi*

Anthrax: *Bacillus anthracis*

Lyme Disease: *Borrelia burgdorferi*

The Plague: *Yersinia pestis*

Tuberculosis: *Mycobacterium tuberculosis*

Legionnaire's Disease: *Legionnella pneumophila*

Gonorrhoea: *Neisseria gonorrhoeae*

Syphilis: *Treponema pallidum*

Whooping Cough: *Bordetella pertussis*

MRSA: *Staphylococcus Aureus*

Rocky Mountain Spotted Fever: *Rickettsia rickettsii*

20.4.2. Viral Infections

The Common Cold: rhinovirus

Influenza: the flu

Herpes: herpes simplex virus

Chickenpox and Shingles: Varicella-Zoster virus

Mononucleosis: Epstein-Barr virus

Hepatitis A: hepatovirus A

Poliomyelitis: poliovirus

Smallpox: variola virus

Measles: morbillivirus

Rabies: lyssavirus

West Nile Virus: flavivirus

HPV: human papillomavirus

AIDS: human immunodeficiency virus (HIV)

Ebola: ebolavirus (EVD)

Zika Virus Disease: flavivirus (ZVD)

SARS, MERS and COVID-19: coronavirus

20.4.3. Protist Infections

Malaria: *Plasmodium falciparum*

The protist parasite *Plasmodium falciparum* is responsible for causing malaria, using a mosquito host as its vector. When the mosquito bites a human or other animal, it injects sporozoites, which transform and multiply in the liver to form merozoites, before migrating into red blood cells, destroying them in the process. Gametocytes (male and female forms of the parasite) can develop at this stage. If another mosquito bites, it draws up some blood containing the gametocytes as part of its 'blood meal', which reproduce sexually inside the mosquito, completing the life cycle of the parasite.

Dysentery: *Entamoeba histolytica*

Giardiasis: *Giardia lamblia*

Toxoplasmosis: *Toxoplasma gondii*

Primary Amoebic Meningoencephalitis: *Naegleria fowleri*

20.4.4. Fungal Infections

Onychomycosis (nail infections)

Ringworm:

Vaginal Candidiasis (Thrush):

Athlete's Foot:

Dandruff: *Malassezia*

20.4.5. Non-Communicable Diseases of the Eyes

Glaucoma: the second most common form of blindness, after cataracts. Glaucoma causes loss of peripheral vision. If the trabecular meshwork is blocked, aqueous humour can no longer drain out, causing an increase in intraocular pressure (IOP), which exerts a stress on the porous connective tissue filling the optic nerve head (lamina cribrosa), damaging nearby retinal ganglion nerve cells.

20.4.6. Non-Communicable Diseases of the Ear, Nose and Throat

Ménière's disease:

Tinnitus:

20.4.7. Non-Communicable Diseases of the Circulatory System

20.4.8. Non-Communicable Diseases of the Nervous System

Guillain-Barré syndrome:

20.5. Health, Fitness and Nutrition

20.5.1. Sex Determination

Sex: the purely biological classification based on the reproductive functions of the organism.

Binary Model of Sex: humans are ~15% sexually dimorphic by weight: the body morphology of the two sexes is fundamentally different. In **most** humans, the sex chromosomes present in the cells determine the sex of the human: only X in gametes and **XX** in somatic cells indicates **female** sex, while either X or Y in gametes and **XY** in somatic cells indicates **male** sex. At fertilisation between a male and a female, the male gamete (sperm) fertilises the female gamete (egg) to form a zygote with either XX or XY sex chromosomes, which determines the sex of the resulting baby.

The binary model is a simplified view of sex and neglects cases where the most common route to forming a viable foetus is not followed. In these cases, alterations to anatomy and/or hormone profile may occur without impacting the viability of the foetus. In normal prenatal sex differentiation, the male and female embryos are anatomically identical until week 7 of the pregnancy, when the presence or the absence of various genes influences gonad development, for example:

- The SOX9 gene (chromosome 17) forms the testes by producing Sertoli cells.
- The NR0B1 gene (chromosome X) codes for protein DAX1 which inhibits SOX9.
- The SRY gene (chromosome Y) codes for a protein which inhibits transcription of NR0B1.

Differences in Sex Development (DSD) / Variations in Sex Characteristics (VSC) / Intersex: any atypical combination of sex chromosomes, gametes and gonads.

- **Aneuploid karyotypes:** any variation in sex chromosome number in somatic (diploid) cells. Examples: 47,XXY (Klinefelter syndrome); 47,XYY (Jacob's syndrome); 47,XXX (trisomy X syndrome); 45,X (Turner's syndrome). These are usually not considered a deviation from binary sex.
- **Gonadal dysgenesis:** genotype is the binary opposite of the phenotype.
  - Swyer syndrome: 46,XY but develops female gametes and reproductive organs.
  - De la Chapelle syndrome: 46,XX but develops male gametes and reproductive organs.
- **Sex chromosomal mosaicism:** different cells around the body have opposite genotypes.
- **Congenital adrenal hyperplasia:** virilised (masculinised) hormone profile and ambiguous genitalia, caused by mutations in the 21-hydroxylase enzyme, leading to accumulation of cortisol precursors, overstimulating the adrenal glands and producing an excess of androgens.
- **Androgen insensitivity syndrome:** entirely feminised development in karyotypic males, caused by mutations in the androgen receptor, rendering androgens an ineffective hormone.
- **Persistent Mullerian duct syndrome:** partial or complete female gonads in karyotypic males.

Cultural Variations: some communities have their own unique experiences of sex. Some examples:

- **Guevedoces people:** a small Dominican community with a common mutation in 5-alpha-reductase, rendering testosterone in the foetal stage ineffective. They are karyotypic males who appear female until puberty when they develop male gonads. The study of this population in the 1970s led to the development of the DHT blocker medication finasteride.
- **Hijra people:** a community in India/Pakistan/Bangladesh who are sometimes intersex, but also have unique experiences of gender (often being transgender or a 'third gender').

20.5.2. Gender Identity

Gender encompasses various socially-constructed concepts relating to psychology and sociology, including gender roles and gender presentation (culture-dependent expectations about how people with a perceived biological sex should behave in society). Gender identity is a person's innate sense of what aspects of gender feel the most natural, and is generated by the brain.

Gender Identity: a person's internal and individual experience of gender.

Gender identity has its basis in brain structure, which in turn is influenced strongly by hormone exposure during development. All humans have varying amounts of testosterone and oestrogen, and the receptor proteins for these hormones are controlled by numerous separate gene pathways. Estradiol produces female secondary sex characteristics but male brain anatomy, and is typically blocked from crossing the blood-brain barrier in females by alpha-fetoprotein.

It is known that the human brain assumes a 'female' phenotype from conception, and differentiates to the 'male' phenotype only if exposed to the gonadal hormones, which operate by epigenetic mechanisms (inhibition of methyltransferase enzymes, reducing suppression of masculinising genes). When this process occurs, the person's gender identity matches their sex (a cisgender person) and when it does not, the gender identity may differ from the sex (a transgender person).

Sexually dimorphic regions of the brain include the SDN-POA (sexually dimorphic nucleus of the preoptic area), BsTC (bed nucleus of stria terminalis) and VIP-expressing neurons in the SCN (vasoactive intestinal polypeptide in the suprachiasmatic nucleus), which are small densely packed regions near the hypothalamus. In transgender individuals, these regions typically show structures skewed towards those resembling their gender identity rather than their sex.

- **Cis man:** assigned male at birth, has a male gender identity.
- **Cis woman:** assigned female at birth, has a female gender identity.
- **Trans man:** assigned female at birth (AFAB), has a male gender identity.
- **Trans woman:** assigned male at birth (AMAB), has a female gender identity.
- **Non-binary / Genderqueer:** any different, fluctuating and/or indeterminate gender identity.

Transgender people may (if possible and desirable) undergo 'gender affirming surgery' (top and bottom surgery) in which the external appearance of the body is changed to match those associated with their gender identity, as well as hormone therapy to adjust the endocrine profile and bring about changes in the secondary sexual characteristics. Neither the sex nor the gender identity changes in this process, but the resulting congruence between gender presentation and gender identity allows for the resolution of gender dysphoria in transgender people. Since transgender identities are rooted in brain structure, the organ that generates the sense of self, it is more practical and customary to treat transgender individuals according to their gender identity rather than their sex, unless there is good reason not to do so (e.g. physical capabilities in sports, which are segregated by sex).

20.5.3. Sexuality

Sexuality: a natural desire to be sexually or romantically attracted to members of a particular gender.

Some of the sexually dimorphic regions of the brain also show differences in homosexual and heterosexual people, e.g. the BsTC and SCN, indicating that sexuality can also be rooted in biology, like gender identity. However, the SDN does not show such changes, indicative of the difference between sexuality and gender identity. Sexuality is known to be influenced predominantly by genetic (deterministic) factors, but also by epigenetic (environmental) factors.

Sexualities can be classified into:

- Heterosexual (straight): a man or woman attracted to another of the opposite gender.
- Homosexual (gay / lesbian): a man or woman attracted to another of the same gender.
- Bisexual: being sexually attracted to both men and women, and potentially other genders.
- Asexual: not having strong sexual attractions towards anyone.
- Pansexual: being sexually attracted to all genders (gender is not a factor in determining attraction).
- Queer: usually describes any other sexuality that fluctuates or is not easily categorised.

Besides sexual attraction, individuals may also be independently ‘romantically attracted’ (can develop strong emotional bonds but may not have sexual desires) to people of certain gender(s). There are many other terms occasionally used to describe these groups, but most can be considered combinations (e.g. aro-ace = asexual + aromantic) or some point on a spectrum between those of the above.

Many aspects of gender identity and sexuality can also be considered a spectrum rather than discrete or binary classifications, owing to the complexity and potential for variation of the neurochemical profile responsible for creating these traits. This is an active area of research, and many of the sociocultural aspects fall outside the domain of ‘hard’ science.

LGBTQ+ Community: lesbian, gay, bisexual, transgender, queer, …

Due to the historically dominant ideologies of heteronormativity and cisnormativity, LGBTQ+ individuals may face discrimination (homophobia, transphobia), and form a community to improve visibility and promote understanding. The LGBTQ+ acronym is sometimes extended to include other named groups (I: intersex, A: asexual, P: pansexual…).

Sexuality in Other Animals: a range of sexual behaviours are observed in other primates

Many primates display ‘sexually liberal’ behaviour.

20.5.4. Neurodiversity

Neurodiversity refers to the range of brain functions and behavioural traits in a population.

Neurodivergence refers to deviations from the most common ('neurotypical') traits in the population, and includes the autism spectrum, ADHD, dyslexia, dyscalculia and dyspraxia.

Autistic Spectrum Disorder (ASD)

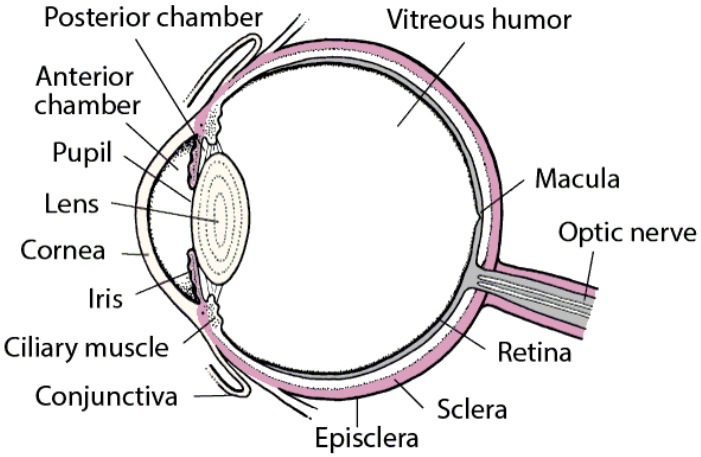
Autism is characterised by some degree of decreased social awareness and increased adherence to repetitive behaviour (DSM-5), with specifics varying by age, culture, gender, and co-occurence of other intersectional mental conditions. The variations in brain structure associated with autism are subtle, primarily concerning neural connectivity, and cannot be reliably detected even in brain scans.

Autism may be 'low support needs' or 'high support needs', reflecting the spectrum of assistance required to comfortably navigate day-to-day life. Autistic individuals may learn to use behavioural management ('masking') to try suppressing expression of some autistic traits if they feel the social environment is unsupportive or hostile to their expression, especially in school children. This may contribute to developing social anxiety.

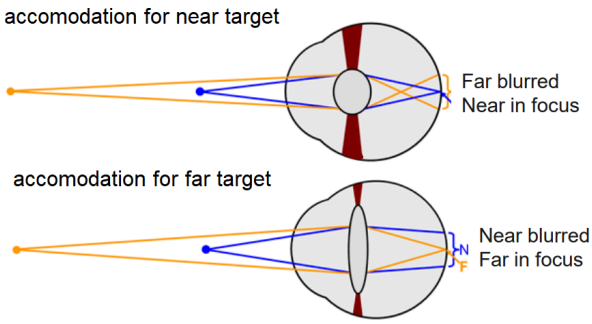
Issues with Reliable Diagnostics

Diagnostic data and guidelines for a variety of psychiatric, medical and wellness issues have historically been biassed towards studies on White males, leading to a failure to account for the variation in the symptoms, risk factors and mode of expression in other populations. Disorders such as ASD are often misdiagnosed in girls and Western ethnic minorities.

20.5.5. Eyesight



Parts of the Eye



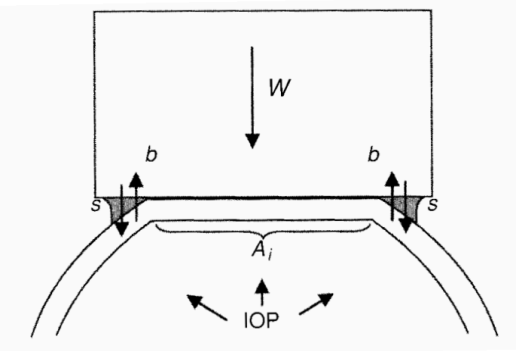
Lens Accommodation

The transparent biconvex lens structure changes shape to change focus. Lens curvature is controlled by ciliary muscles – by changing curvature, the eye focuses on objects at different distances. This process is called accommodation. The **amplitude of accommodation** is the max amount that the lens can accommodate, in diopters (m^{-1}).

- Myopia: nearsightedness, an inability to view distant objects in focus; this is corrected with **concave** lenses.
- Hyperopia and Presbyopia: farsightedness, an inability to view close-up objects in focus; this is corrected with **convex** lenses.

The lens continually grows throughout life, laying new cells over the old cells. There appears to be no protein turnover in the lens nucleus throughout lifespan. The lens is thus prone to age-related changes, in particular stiffening, making near-field accommodation impossible.

Measurements of Intraocular Pressure by Optometrists: indicates various eye defects



Goldmann tonometry: The intraocular pressure (IOP) is given by $p_{IOP} = W / A$, where A is the flattened area of the cornea, calibrated to be exactly 7.35 mm^2 (empirically found to eliminate tear film and bending stress effects), and W is the applied load. Further correction needed for abnormal corneal thickness.

The coefficient of rigidity is $\frac{dp_{IOP}}{dV} \approx \frac{2hE}{3RV(1-\nu)}$, where V is the eyeball volume, h is average cornea and sclera thickness, R is the radius of the eyeball, and (E, ν) are the Young's modulus and Poisson's ratio of the eyeball structure.

20.5.5. Bioenergetics of Macronutrients for a Healthy Diet

- Food energy is measured using bomb calorimetry to determine the total enthalpy (heat) of combustion, and is displayed on food labels.
- Metabolisable (available) energy = Energy in food − Energy in excreted substances. The digestibility efficiency is usually very high (~97%).
- The brain consumes about 20% of the total energy intake.
- Basal metabolic rate (BMR): energy expenditure rate when at rest.
  Men: $BMR\,[kcal/day] \approx 10W\,[kg] + 6.25H\,[cm] - 5A\,[yr] + 5$
  Women: $BMR\,[kcal/day] \approx 10W\,[kg] + 6.25H\,[cm] - 5A\,[yr] - 161$
  (W: weight, H: height, A: age) (Harris-Benedict equations, rev. by Mifflin *et al*, 1990)
  Individuals with 'fast' or 'slow' metabolisms have larger or smaller BMR respectively, with deviations up to ±15% from the above average value.
- Recommended daily energy intakes, accounting for background thermogenic activity, are:
  Men: 2500 kcal (10.46 MJ); Women: 2000 kcal (8.34 MJ)

	Reference daily intake (adult RDI)	Energy content	Specific dynamic action (SDA, thermic effect, energy required for metabolism)
Protein	50 g	4 kcal/g	20-30% (net: 2.8-3.2 kcal/g)
Carbohydrates	260 g of which sugar: 90 g	4 kcal/g	5-10% (net: 3.6-3.8 kcal/g)
Fats	70 g of which saturated: 20 g	9 kcal/g	0-5% (net: 8.5-9.0 kcal/g)

- Net energy balance = Food energy × (1 − SDA) − BMR − Calories lost due to exercise
  An overall SDA of approximately 10% = 0.1 is a reasonable estimate (higher if protein-rich).
- If the net energy balance is negative, the body sources the missing energy from stored tissue: 1) liver glycogen, 2) blood glucose, 3) fat, 4) muscle protein (detrimental).
- If the net energy balance is positive, the body stores the excess energy by forming glycogen/fat/protein (anabolism).
- A calculator for recommended macronutrient intake is available here.
- Body mass index: BMI = Mass [kg] / (Height [m])^2. Healthy range: 18 < BMI < 25.
  BMI is an approximate indicator of body fat (not accounting for muscle mass) for adults.
- Somatotype model: three characteristic morphologies of the body (descriptive, not fixed for life)
  - Ectomorph: thinner body, lower muscle mass, faster metabolism
  - Mesomorph: average weight and build
  - Endomorph: broader body, slower metabolism

20.5.6. Vitamins and Minerals

Essential vitamins: these nutrients are not produced endogenously and must be obtained from food. The 'recommended daily intake' often represents a range and is quoted for healthy adults. Typically, children require less than these values, men require closer to the upper value, women require closer to the lower value except when pregnant or lactating.

Vitamin	Functions and Purposes	Food Sources	Symptoms of Deficiency	Recommended Daily Intake
Vitamin A (retinol, retinal, retinoic acid, β-carotene)	• Vision: forms the rhodopsin pigment in rods on the retina • Skin: promotes collagen growth • Immune system: differentiation of B/T cells	milk, cod liver oil, carrots, butter, green leaves, tomatoes, eggs	night blindness, xerophthalmia (hard cornea)	600 - 700 µg
Vitamin B_1 (thiamine)	• Metabolism of carbohydrates • Digestion: helps produce HCl in the stomach • Nerves: production of neurotransmitters • Cardiovascularity: production of red blood cells	nuts, green vegetables, yeast, egg yolk	beriberi	1.1 - 1.2 mg
Vitamin B_2 (riboflavin)	• Metabolism of carbohydrates and proteins • Cofactor for antioxidant enzymes • Eye health: protects against cataract formation	milk, meat, green vegetables, yeast	ariboflavinosis (tongue inflammation)	1.1 - 1.3 mg
Vitamin B_3 (niacin)	• Cholesterol management • Skin health: reduces inflammation, hyperpigmentation, rosacea • DNA repair	meat, fish, nuts, legumes; from tryptophan	pellagra (diarrhea, dermatitis, dementia)	13 - 16 mg
Vitamin B_5 (pantothenic acid)	• Hormone production (testosterone, oestrogen, cortisol)	meat, fish, eggs, whole grains	very uncommon	5 mg
Vitamin B_6 (pyridoxine)	• Metabolism of proteins • Neurotransmitter production (serotonin, dopamine, GABA) • Antibody production • Hormones: production of melatonin and cortisol; regulates oestrogen and testosterone	meat, fish, nuts, legumes	various generic symptoms	1.2 - 1.4 mg
Vitamin B_7 (biotin)	• Skin, hair and nail health • Foetal development • Blood sugar regulation	egg yolks, liver, nuts, whole grains	hair loss, skin rash, brittle nails	30 - 100 µg
Vitamin B_9 / B_{11} (folate / folic acid / folacin)	• DNA synthesis • Foetal development in early stage of pregnancy • Heart health: reduces homocysteine (heart disease)	leafy green vegetables, citrus fruits, beans, fortified foods	fatigue, anaemia, birth defects	400 µg
Vitamin B_{12} (cobalamin)	• Red blood cell production • Nerves: produces myelin • DNA synthesis	meat, fish, eggs, dairy	fatigue, anaemia	2.4 µg

Essential vitamins (continued)

Vitamin	Functions and Purposes	Sources	Deficiency	Recommended Daily Intake
Vitamin C (L-ascorbic acid)	• Antioxidant: protects against free radicals • Collagen synthesis for skin and bones • Immune system: produces white blood cells • Iron absorption • Wound healing: helps to form vascular tissue	citrus fruits, strawberries, kiwi, papaya, bell peppers	scurvy (joint pains, bleeding gums)	75 - 90 mg
Vitamin D (calciferol)	• Calcium absorption • Cell growth and differentiation	sunlight, fatty fish	osteoporosis, autoimmune disease	15 µg
Vitamin E (α-tocopherol)	• Antioxidant • Skin health: protects against UV damage	nuts, seeds, vegetable oils	muscle weakness, vision problems	15 mg
Vitamin K (K_1: phylloquinone, K_2: menaquinone)	• Clotting factors in blood • Bone health: regulates calcium metabolism, bone mineralisation	leafy greens, fermented foods	abnormal bleeding, bone fracture	90 - 120 µg
Choline	• Forms the neurotransmitter acetylcholine	meat, broccoli	liver damage	425 - 550 mg
Ubiquinone (CoQ_{10})	• Energy production in mitochondria • Topically, reduces oxidative stress in skin	fatty fish	fatigue	100 - 200 mg (not well established)

Essential Minerals: ions used for osmoregulation and metabolic processes (in metalloenzymes)
Essential in substantial quantities (> ~10 mg/day); Essential in trace quantities (< ~10 mg per day)

Mineral	Purposes	Recommended Daily Intake
chloride (Cl^-)	• Production of stomach acid and pH regulation	800 - 2000 mg
phosphorus (as phosphate, PO_4^{3-})	• As phosphate, forms hydroxyapatite for bone matrix including teeth • Production of ATP for energy storage • Backbone of DNA and RNA • Phosphate buffer for pH regulation in the blood and intracellular fluid	700 mg
sulfur (as sulfate, SO_4^{2-})	• Component of cysteine, methionine, biotin, thiamine, glutathione, taurine etc. • As sulfate, forms sulfate ester bonds in glycosaminoglycans • Sulfated steroids for cell signalling	850 mg
calcium (Ca^{2+})	• Triggers release of neurotransmitters • Required for muscle contraction and relaxation, and heartbeat • Forms hydroxyapatite for bone matrix including teeth • Cofactor for amylase	700 mg
iron (Fe)	• Cofactor for catalase, cytochrome and haemoglobin (in haem), nitrogenase, hydrogenase	8.7 - 14.8 mg
magnesium (Mg^{2+})	• Cofactor for glucose 6-phosphatase, hexokinase, DNA polymerase • In plants, used in chlorophyll as the initiator of photosynthesis	270 - 300 mg
potassium (K^+)	• Regulates blood pressure via osmotic pressure • Nerve impulses: intracellular K^+ exits the neuron while the action potential propagates	2600 - 3400 mg
sodium (Na^+)	• Regulates blood pressure via osmotic pressure • Nerve impulses: extracellular Na^+ enter the neuron while the action potential propagates	1500 mg
zinc (Zn)	• Development and function of immune cells • Taste and smell receptors • Cofactor for carbonic anhydrase, alcohol dehydrogenase, DNA/RNA polymerase, carboxypeptidase and zinc finger proteins	8 - 11 mg
iodide (I^-)	• Used by the thyroid gland to make hormones for growth and metabolism	150 µg
chromium (Cr)	• Regulation of blood sugar levels by enhancing insulin activity	25 - 35 µg
cobalt (Co)	• Cofactor for cobalamin (vitamin B_{12}), nitrile hydratase and carbon monoxide dehydrogenase	5 - 8 µg
copper (Cu)	• Cofactor for cytochrome oxidase and superoxide dismutase	900 µg
manganese (Mn)	• Cofactor for arginase	1.8 - 2.3 µg
molybdenum (Mo)	• Cofactor (in molybdopterin) for nitrate reductase, sulfite/xanthine oxidase and nitrogenase • Prosthetic group (in molybdopterin) for formate dehydrogenase	45 µg
nickel (Ni)	• Cofactor in urease	23 - 35 µg
selenium (Se)	• Cofactor for glutathione peroxidase	55 µg

The following minerals are essential in some organisms, but with little or no known use in humans:

Mineral	Potential Purposes
boron (B)	• As inorganic borates, essential for plants and may be beneficial in maintaining healthy bones and joints.
silicon (Si)	• May help to synthesise collagen for connective tissues e.g. bones, cartilage.
vanadium (V)	• Cofactor for bromoperoxidase in marine algae • Can mimic insulin signalling for glucose metabolism in some mammals, but less effective in humans
nitrate (NO_3^-)	• In plants, used as an essential nitrogen source for growth and development. Not required for humans.
tungsten (W)	• In some extremophiles, replaces molybdenum in molybdopterin cofactor. Toxic in many other organisms.
lanthanides (Ln^{3+})	• In some bacteria, La, Ce, Pr, Nd are used in some PQQ-dependent methanol dehydrogenase (MDH) enzymes

20.5.7. Metabolism of Proteins

Proteins: metabolised to amino acids

- **Essential amino acids:** of the 20 proteinogenic amino acids (Section 16.5.10), 9 are essential (not produced inside the body): histidine (His), isoleucine (Ile), leucine (Leu), lysine (Lys), methionine (Met), phenylalanine (Phe), threonine (Thr), tryptophan (Trp), and valine (Val).
- **'Complete proteins'** are sourced from foods containing all amino acids: all animal proteins, including eggs and fish, are complete, and some but not all plant proteins are also complete e.g. quinoa, whey, blue-green algae, soybeans.
- Protein deficiency rarely occurs as an isolated condition. It usually accompanies a deficiency of dietary energy and other nutrients resulting from insufficient food intake.
- Proteins are broken down into amino acids by proteases in the stomach and intestine (e.g. trypsin, pepsin, chymotrypsin), then used for cell growth and repair to build endogenous proteins from aminoacyl-tRNAs, or used for gluconeogenesis if blood sugar is low, or for biosynthesis of neurotransmitters and hormones.

20.5.8. Metabolism of Lipids

Lipids are hydrophobic molecules, including fatty acids, triglycerides, phospholipids, cholesterol and fat-soluble vitamins. Most lipid metabolism occurs in the small intestine.

Bile is produced by the liver and stored in the gallbladder, and contains bile salts, formed from the condensation of bile acids (steroid acids, e.g. cholic acid) with amino acids (e.g. glycine, taurine). These are bio-surfactants in the duodenum (upper small intestine) that emulsify fats and facilitate lipolysis by pancreatic lipase enzymes, breaking down triglycerides into free fatty acids.

The fatty acids are collected into water-soluble micelles by the bile salts, and are absorbed by enterocytes (intestinal epithelial cells). The triglycerides are then reassembled and packed into chylomicrons (lipoproteins) for transport via the lymphatic system, with further processing by the liver.

Types of lipoprotein (lipid microparticles produced by the body for transporting cholesterol and triglycerides) include:

- **Chylomicrons:** produced in the small intestine villi, transporting triglycerides to the liver, skeletal muscles and adipose tissues via the lacteal lymphatic vessels. Once emptied, the chylomicron remnants are cleared from the bloodstream by being consumed in the liver, where their lipids are used to synthesise VLDL, LDL and HDL.
- **Low density lipoproteins (LDL):** circulates fat around the bloodstream to the whole body. They carry ~60% cholesterol. High genetic expression of the **ApoB** protein increases the proportion of LDL in the body, and mutations in the LDL receptor (LRL-R) contribute to their slower removal from the bloodstream. It is considered '**bad** cholesterol', as it can deposit the cholesterol in arterial walls, narrowing the blood vessels (atherosclerosis) and increasing risk of heart attack, stroke and coronary artery disease.
- **High density lipoproteins (HDL):** collects fat from cells and tissues, returns it to the liver for conversion to steroid hormones. High genetic expression of the **ApoA1** protein increases the proportion of HDL in the body. The microRNA MiR-33 epigenetically inhibits ApoA1 production, reducing HDL. It is considered '**good** cholesterol', as it removes cholesterol from circulation.

Non-genetic factors associated with increased LDL proportion include high-calorie, high-saturated fat, smoking and binge drinking. Dietary cholesterol consumption only slightly increases LDL content, and is not considered a risk factor if part of a balanced diet and in moderation. Statins are drugs for inhibiting HMG-CoA reductase for liver LDL production.

Saturated Fats: straight-chain lipids with no C=C bonds. Consumption increases LDL content.

Unsaturated Fats: straight-chain lipids with C=C bonds.

- **Trans fats:** contain a *trans* C=C bond. Sharply raises LDL and lowers HDL. Processed foods containing partially hydrogenated vegetable oils are high in trans fats.
- **Omega-3 fatty acids:** helps lower triglyceride levels, aids in brain grey matter and heart health.

20.5.9. Food Ingredients and Processed Foods

Ingredients in Foods: ingredients that often appear in dietary headlines

- **High-fructose corn syrup (HFCS):** a sweetener used in processed foods, derived from corn starch. Indirectly associated with obesity, insulin resistance and some metabolic disorders.
- **Hybridised wheat / maize:** breeding of different plant species by cross-pollination.
- **Genetically modified organisms (GMOs):** direct genetic modification of plants or small animals, typically with the aim of increasing crop yields or resilience against pest insects.
- **Lab-grown meat (cultured meat):** meat grown from cultivated animal cells (satellite stem cells). As of 2024, lab-grown meat is legal in the US (chicken), UK (pet food), Singapore (chicken) and Israel (beef).
- **Nitrites:** sodium nitrite is a preservative and inhibitor of *Clostridium botulinum* used in cured meats (bacon, ham, sausages) that can react with creatine in the meat on cooking to form nanogram quantities of nitrosamines, which are carcinogenic and hepatotoxic.
- **Fortified foods:** addition of normally-absent nutrients to foods e.g. vitamin B_9/D in bread, iodine in salt, iron in wheat/maize flours.
- **Ultra-processed foods:** edible food-like products subjected to extensive processing including non-culinary ingredients. UPFs are designed to be addictive (‘hyper-palatable’).
- **Pasteurised milk:** milk subjected to heat treatment to kill pathogens. Unpasteurised (raw) milk can cause tuberculosis, salmonella and bird flu.
- **Animal fats and vegetable fats:** animal fats are high in monounsaturated and saturated fats, while vegetable fats are high in omega/mono/polyunsaturated (oleic acid and linoleic acid).
- **Hydrogenated oils:** processing of vegetable oils forms saturated fats and trans fats in high proportions.
- **Refined oils:** volatile fatty acids removed from olive oils. Unrefined oils include ‘extra virgin’ olive oil.
- **Inverted sugar syrup:** an added sugar in honey, jam or chocolate, that crystallises less readily, consisting of hydrolysed sucrose (glucose and fructose), using citric acid or invertase enzyme catalyst.
- **Refined sugars:** processed sugar from sugarcane, found in sodas and sweets. Linked to obesity, type 2 diabetes and tooth decay.
- **Gluten:** a structural protein made of gliadin and glutenin, found in wheat grain, including bread. People with coeliac disease and other gluten-related disorders must avoid gluten.
- **Artificial sweeteners (aspartame, sucralose, saccharin, acesulfame K):** a sugar substitute in low-calorie diets. Sucralose, saccharin and acesulfame K are not metabolised in the body.
- **Artificial food colourings (Red 3/40, Yellow 5/6, Blue 1/2):** coloured compounds used in processed foods, with some associations with hyperactivity in children.
- **Monosodium glutamate (MSG):** a lower-sodium salt natural flavour enhancer, widely used in Asia.

- **Bovaer (3-nitrooxypropanol):** a non-toxic additive to cattle feed that inhibits methanogenesis in their microbiome, significantly reducing greenhouse gas (methane) emissions from cattle. This additive is not present in any milk or butter obtained from the animals who consume it.
- **Chlorinated chicken and hormone-treated beef:** in the US, slaughtered chicken is routinely washed in chlorine to kill pathogens, and cows are given steroids for growth. Importing these from the US is banned in the EU and UK, due to concerns that it is used to conceal the effects of substandard animal welfare, or potentially for protectionism as the basis of trade disputes.

Organic Food: avoids the use of synthetic fertilisers, pesticides, growth regulators (hormones) and livestock feed additives, as well as avoiding irradiation and GMOs, in growing the food. Labels include:

- **“100% Organic”:** all ingredients in a food product are organic (meeting the above definition).
- **“Organic”:** at least 95% of the ingredients are organic. Up to 5% may be non-organic.
- **“Made with Organic”:** at least 70% of the ingredients are organic. Up to 30% may be non-organic.

Diets:

- **Pescatarian:** no eating red meat. Fish, eggs and dairy are allowed.
- **Plant-based:** prioritises eating foods from plants, minimising meat consumption.
- **Vegetarian:** no eating meat. Eggs and dairy are allowed.
- **Vegan:** no eating meat and no eating any product derived from animals (eggs, dairy).
- **Paleo:** non-processed foods diet. Organic meats, fruit and vegetables, nuts and seeds.
- **Keto:** low-carb, high-fat diet. Meat, dairy, seafood, nuts and seeds.

20.5.10. Skincare

Components of the skin surface: forms the first barrier against pathogenic infection

- **Lipid layer:** the outermost physical barrier (the stratum corneum), made of ceramides, cholesterol and fatty acids.
- **Acid mantle:** the thin film of sebum and natural oils at pH 4.7-5.7, which inhibits pathogen colonisation. Skin pH influences cellular signalling pathways, skin pigmentation and epidermal stem cell differentiation.
- **Microbiome (skin flora):** the community of non-pathogenic bacteria, which competes for resources with potentially harmful bacteria. The microbiome is destroyed when the surface pH becomes alkaline.

Active ingredients in Skin Moisturisers:

- Polyelectrolytes: e.g. hyaluronic acid, chitosan, alginate
- Retinol (vitamin A), niacinamide (vitamin B_3) and ascorbic acid (vitamin C)
- Hydroxy acids: mild exfoliants (shedding of dead surface skin), with three types:
  - Alpha-hydroxy acids (AHAs): e.g. glycolic acid, lactic acid, mandelic acid
  - Beta-hydroxy acids (BHAs): e.g. salicylic acid, mild anti-inflammatory agent
  - Poly-hydroxy acids (PHAs): e.g. gluconolactone, lactobionic acid.
- Ceramides
- Probiotics

Other ingredients used in Cosmetics: used to optimise formulation properties

- Carbomer (polyacrylic acid): increases formulation viscosity
- Polyquaternium-n: range of 4^{o} ammonium polycations, holds active ingredients in place for longer

Sun Protection: the Sun emits UV radiation which can damage the skin

- **UVA (lower energy)** is responsible for tanning, hyperpigmentation and skin ageing.
- **UVB (higher energy)** is responsible for sunburn (erythema) and skin cancer (melanoma).
- Incident UV may be **direct** (in direct sunlight, most intense), **specular** (reflected from shiny surfaces) or **diffuse** (reflected from environment / scattered by clouds, least intense), so UV exposure still occurs when in the shade or on a cloudy day. UV intensity varies by time of day, day of year and latitude.
- **Physical (mineral) sunscreens** work by reflecting UV radiation. Active ingredients include ZnO or TiO_2 nanoparticles. They are non-comedogenic (they do not clog pores, which results in blackheads).
- **Chemical sunscreens** work by absorbing UV radiation and emitting at longer wavelengths. Active ingredients include organic compounds such as diethylamino hydroxybenzoyl hexyl benzoate (UVA), ethylhexyl triazone (UVB), oxybenzone (both), drometrizole trisiloxane (both), among many others. Some people may have allergies to some of these ingredients, causing e.g. rosacea (redness/rashes).
- Sunscreens marked with **'SPF50+'** indicate less than 2% of UVB radiation penetrates the skin when applied fully.
- **Broad-spectrum sunscreens** also block UVA, indicated by a PPD (persistent pigment darkening) score, with the best being 'PA++++'.
- Sunscreens only last ~6 hours before needing reapplication due to continuous absorption into skin (even if not used in the sun) and photodegradation. Physical sunscreens typically have more desirable ADME properties (absorbed less readily, less harmful metabolites) than chemical sunscreens.
- Sunscreen is typically recommended when the outdoors UV index is 3 or above.
- When the UV index is low, sunscreen is not advised to allow the body to synthesise sufficient vitamin D_3 from 7-dehydrocholesterol in response to the UVB exposure.
- Sunscreen quantity and formulation must be suitable for one's skin tone, as melanin acts like a mild sunscreen: ~SPF 5 for Black people relative to White people, blocking ~80% of UV. This means a White person with SPF 50 and a Black person with SPF 10 in the same setting would experience equal UV penetrations (2%).
- When indoors, glass windows block UVB but allows most of UVA to pass.
- Some formulations contain probiotics (fermented extracts; lysates of bacteria) which promote growth of the skin's microbiome, improving the skin's capacity as a barrier defence to pathogens.

20.5.11. Haircare

Head hair on the scalp grows in permanently from about 1 year after birth, with vellus hair (short, thin, light-coloured) occurring over the body. During puberty, vellus hair is converted to terminal hair (long, thick, melanated) by androgen hormones. In men, the extent and thickness of terminal hair is larger than in women.

The hair cell cycle: anagen (growth: 2-6 years) → catagen (pause) → telogen (shedding and rest). During the anagen phase, hair grows at 1-2 cm per month.

Hair colours and types:

- **Colours:** black, brown (more eumelanin), red (more pheomelanin), blonde (less melanin)
- **Shape:** frizzy, curly, wavy, straight. Determined genetically by asymmetry of the hair bulb.
- **Porosity:** high porosity (absorbs and releases moisture quickly), low porosity (does not absorb moisture). Determined by the openness of the cuticle surface. Intermediate porosity provides optimal moisturability, while high or low porosity hair tends to dry out easily.

Hair shampooing: cleans the scalp skin and hair

- A blend of surfactants remove contaminants in a manner similar to soap.
- Sulfate-free shampoos avoid using SLS as the surfactant, which reduces irritation and oil removal.
- Shampoo is usually applied less frequently than conditioner, as it can lead to drying and frizzing.

Hair conditioning: provides moisture to hair, increasing ease of brushing, shininess and smoothness.

- Protein-free conditioners are suitable for low porosity hair, as this tends to increase porosity.
- Leave-in conditioners with shea butter/argan oil/coconut oil are suitable for high porosity hair.
- Conditioner is usually applied after shampoo and rinsing, resupplying the lost moisture.

Chemical depilatory: used to remove hair from the skin. Thioglycolates cleave disulfide bridges in keratin and hydrolyse for easy removal. These are also used in perms for curly hair.

Hair follicle miniaturisation: in adult males, the enzyme 5α-reductase converts testosterone to dihydrotestosterone (DHT), which is a much more potent agonist of the androgen receptor (NR3C4), accelerating virilisation (masculinisation). DHT promotes pubic hair growth but also induces scalp recession (receding hairlines) and male pattern baldness (androgenic **alopecia** in men). Inhibitors of 5α-reductase such as **finasteride** or **dutasteride** can treat these conditions, as well as its use in hormone therapy as an androgen inhibitor for transgender women. The non-hormonal vasodilator **minoxidil** is also known to promote hair growth in most men by a not-fully-understood mechanism.

Greying of hair (canities): decreased melanin pigment production from the root hair cells leads to grey/white hair follicles. This occurs when the melanocyte stem cells at the root stop undergoing melanogenesis, beginning in early adulthood and progressing throughout later life, fundamentally due to a mixture of environmental (e.g. stress) and genetic factors.

20.5.12. Dental Care

Active ingredients in toothpastes:

Bacteria colonise on teeth, feeding on sugar deposits, forming dental plaque. These bacteria release acids as they metabolise the sugar, which dissolves the hydroxyapatite in the enamel, demineralising the enamel and forming caries and cavities.

- Fluoride: prevents cavities by remineralising enamel (converting hydroxyapatite to the less soluble fluorapatite), and also provides antimicrobial protection. Usually present as sodium fluoride (~1500 ppm), which is soluble and can wash away. It does not 'whiten' the teeth.
- Arginine and calcium carbonate ('Pro-Argin'): helps reduce hypersensitivity. It is the active ingredient in some Colgate toothpastes.
- NovaMin: a calcium phosphosilicate bioactive glass for steady hydroxycarbonate apatite formation in the presence of acid. It is the active ingredient in Sensodyne toothpaste.
- Nano-hydroxyapatite: nanoparticles of hydroxyapatite that deposit in cracks on teeth. Prevents cavities and also has a whitening effect, but not antimicrobial.
- BioMin F: a fluoro-calcium phosphosilicate bioactive glass for steady remineralisation by insoluble fluorapatite deposition.

Drinking water, which is fluoridated at a low concentration (~1 ppm), provides a baseline source of fluoride between meals to help when teeth brushing is missed. Fluoride operates primarily by the topical mechanism (fluorapatite formation) but also the systemic mechanism (incorporation into children's teeth through the bloodstream prior to eruption). Excessive fluoride can result in fluorosis, a discolouration of the teeth due to CaF_2 formation.

Active ingredients in mouthwashes:

- Cetylpyridinium chloride: a quaternary ammonium surfactant, anti-plaque and anti-tartar.
- Alcohol: ~20% ethanol solution. People with dry mouths should not use alcoholic mouthwashes.
- Hydrogen peroxides: ~3% peroxide solution. A bleaching agent with a teeth whitening effect.
- Salt water: a simple solution for rinsing, relieving some minor irritations (e.g. canker sores). Limited effectiveness in killing pathogens and reducing plaque.
- Chlorhexidine: a disinfectant and antiseptic, preventing gingivitis (gum disease) and plaque formation. Only recommended for short-term use (e.g. after dental surgery) due to binding with dietary tannins (polyphenols) to stain teeth.

20.5.13. Other Cosmetics

Deodorants: masks body odour

- Sold in roll-on (balm stick) or spray (liquid formulation in a pressurised canister) forms.
- May contain aerosolised fragrances, odour absorbers (e.g. zinc ricinoleate), antimicrobial agents (e.g. chlorhexidine gluconate), carrier fluid (propellants e.g. liquefied butane).
- Natural ingredients include baking soda (increases skin pH), coconut oil, shea butter, cocoa butter, essential oils (tea tree oil, lavender oil). Some of these are also moisturisers.

Anti-perspirants: inhibits sweating

- Sold in roll-on or spray forms.
- Some contain aluminium (as aluminium chlorohydrate), with some evidence of toxicity.

20.6. Neuroscience and Biopsychology

20.6.1.

20.7. Bioengineering, Bioelectronics and Biointerfaces

20.7.1. Implantable Devices

Cardiac Pacemaker: to maintain a regular heart rate.

Fitted using local anaesthesia, for bradycardia, tachycardia, heart block and cardiac arrest, by delivering electrical signals to the heart muscles to modulate contraction and pumping. They can be single-chamber, dual-chamber, or biventricular. Most models pace only when needed and are rate responsive. Implantable cardioverter defibrillators (ICDs) can 'reboot' the heart.

Auditory Prostheses: to allow hearing in hearing-impaired or deaf people

A **hearing aid** is a simple amplifier with background noise filtering, fitted in the ear canal. It requires all parts of the ear to remain functional despite being reduced in sensitivity. A **cochlear implant** features a microphone, processor, transmitter (located on the scalp) and a receiver (located under the skin). Electrical signals from the receiver are conducted by an electrode threaded into the cochlea and stimulate the auditory nerve. Current spreading in the cochlea can result in decreased sound clarity, so configuration of electrodes (multipolar arrangement) and stimulation protocols (encoding of their signals e.g. multiple phases) are important design parameters. If the auditory nerve is damaged, an auditory brainstem implant (ABI) can be used to directly send signals to the brain.

Visual Prostheses: to allow sight in visually-impaired or blind people

Depending on which stage of the visual pathway is damaged, different levels of intervention are required. Stimulation can occur at the retina or the visual cortex, with varying degrees of information pre-processing and encoding of the input required. Bio-inspired photovoltaic devices are currently under research, which use semiconductive polymers to excite neurons in the retina.

CNS and PNS Implants: to stimulate particular areas of the nervous system.

There is a trade-off between invasiveness and signal resolution. The dura of the brain acts as a low-pass filter, removing important information, so subdural electrodes are required for clear signals.

- Electrocorticography (ECog): long electrodes for deep brain stimulation (DBS). The incoming electrical signals interrupt signals that cause tremors/rigidity/stiffness, although the neuromodulatory mechanism is not understood.
- Brain-computer interface: activate motor neurons directly from signals in the brain.
- Spinal cord stimulation: can relieve chronic pain in the back, neck, legs or arms.
- Single nerve stimulation: e.g. phrenic nerve for assisted breathing, sacral nerve for incontinence, PNS for drop foot, oesophagus for reflux.
- Electroceuticals: stimulation of the vagus nerve using a specific neural biomarker (code) to induce specific changes e.g. reducing inflammation. Finding such markers is an active area of research for a variety of disorders.

Implants can induce the foreign body response (broken capillaries on entry → protein release and adhesion → summoning of macrophages → macrophage fusion into giant cell → ECM deposition), which creates a deadzone around the electrode. It is difficult to prevent entirely, but can be reduced using e.g. hydrogels to match mechanical properties with surface nanostructures and immunomodulatory coatings.

20.7.2. Cutaneous Devices

Cutaneous (on the skin) electrophysiology uses electrodes to perform:

- Recording: brain (electroencephalography, EEG), heart (electrocardiography, ECG), muscles (electromyography, EMG)
- Stimulation: transcranial direct current stimulation (tDCS), transcutaneous electrical nerve stimulation, electroconvulsive therapy (ECT), defibrillation

Simple metal electrodes include Ag/AgCl (Ohmic, non-polarisable). Modern developments include inks or textiles with conductive polymers.

EEGs measure 'brain waves', coordinated synchronised nerve impulses, by an array of cutaneous electrodes. The frequency of the EEG signal can indicate the state of the brain: delta (0.5-4 Hz) → deep sleep, theta (4-7 Hz) → light sleep, alpha (8-13 Hz) → relaxed, beta (13-30 Hz) → alert/focussed, gamma (>30 Hz) → high-level cognition/memorising/problem-solving.

Transcranial direct current stimulation (tDCS) uses low currents between two electrodes; ECT uses much higher currents to induce seizure. These treatments can often be effective despite uncertainty on their mechanism of action: it is thought to act as a 'soft reset', inducing changes in neural chemistry connectivity, breaking or deactivating neural circuits associated with various mental disorders. Temporally-interfering fields can localise excitation to specific parts of the brain, especially targeting deep brain stimulation rather than neocortical regions.

Electrical impedance tomography (EIT) can be used as an imaging technique in which the impedance is reconstructed from fast-switching AC pulses through a set of electrodes around the perimeter of the target. It has also been applied to touch-sensitive biomimetic skins (tactile sensors).

Glucose meters are biosensors that use a small filament-like needle as an electrode to measure glucose concentration in the interstitial fluid. In an electrochemical biosensor (Section 13.5.12), an immobilised enzyme (e.g. glucose oxidase with FAD) reacts with glucose and produces $FADH_2$, which undergoes a redox reaction to generate a current signal indicative of glucose concentration. There is a short time lag between blood glucose and measured glucose. Glucose meters are useful for patients with diabetes, and can be coupled with insulin autoinjectors as part of a control system.

EMGs can be used as data sources for machine-learning training of personalised prosthetic limbs, in which muscle signals of the stump are mapped to target robot pose movements, although surface electrodes (sEMG) produce noisier and less area-precise data than subcutaneous (needle) EMG.

Iontophoresis involves inducing the release of small molecules using a small current between two electrodes, as a form of drug delivery. Can be used in combination with biosensors for signalled release.

20.7.3. *In-vitro* Medical Devices

In-vitro Electrophysiology: recording electrical activity on cell tissues only

Useful for drug discovery and toxicology, and helps to minimise the use of animals. In 2022, the FDA (USA regulatory body) approved the sale of new drugs without requiring any animal testing data.

Extracellular IVE involves growing cell culture on a sensor, which can detect signals of order ~100 μV. Intracellular IVE involves use of thin electrodes being pushed into individual cells (patch clamp technique with a micromanipulator), which is high-effort, low-throughout and the cells die within ~24 hours. Using a gentle pick-up, a piece of the cell membrane can be extracted, which will contain ion channels, which can then be recorded individually. The signals are much clearer (~50 mV).

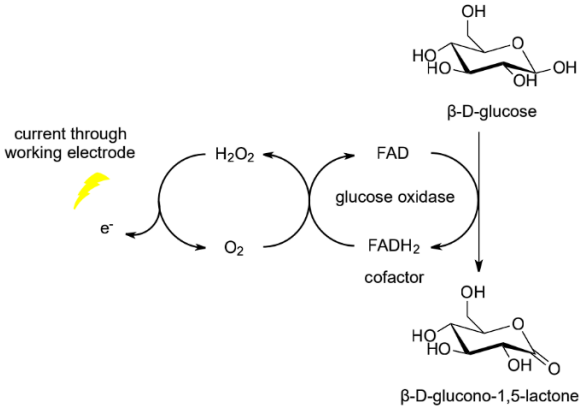
Organoids: physical models of organs to test e.g. drug pharmacokinetics.

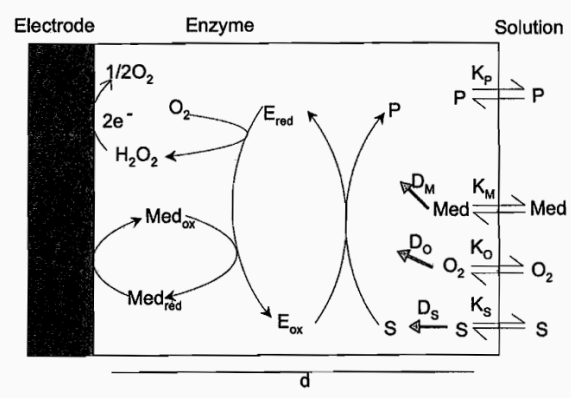
Organ on a chip: simulates tissue activity in vitro. Use of microfluidics allows for fluid flow through micro-scale channels (e.g. blood through vessels, air through bronchioles).

20.7.5. Electrochemical Biosensors and Biofuel Cells

Electrochemical Biosensors: current signal output indicates substrate concentration

An enzyme catalyses a redox reaction with a substrate using a cofactor, and the electrons are transferred to an electrode. The mechanism may involve electron transfer through a mediator molecule acting as a redox couple. Using chronoamperometry (Section 13.5.6), the substrate concentration can be estimated. Example: **glucose biosensor** (below).





For an oxidoreductase enzyme and cofactor (oxidised form: E_{ox}, reduced form: E_{red}) converting analyte/substrate S to product P via a simple oxygen mediator, reactions are:

$$E_{ox} + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_2} E_{red} + P; \qquad E_{red} + O_2 \xrightarrow{k_3} E_{ox} + H_2O_2$$

Rate of oxidoreductase enzyme reaction: r [mol dm^{-3} s^{-1}] $= -\frac{d[S]}{dt} = k_2[E_{\text{total}}]\frac{[S]}{K_M + [S]}$.

Enzyme current: I_{enz} [A cm^{-2}] $= nFdk_2[E_{\text{total}}]\frac{[S]}{K_M + [S]}$ (d: enzyme layer thickness, n: electrons)

- For $[\text{S}] >> K_M$, I_{enz} levels off to a maximum: $I_{\max} = nFdk_2[E_{\text{total}}]$
- For $[\text{S}] < K_M$, $I_{\text{enz}}([\text{S}])$ is approximately linear: $I_{\text{enz}}([\text{S}]) \approx \frac{I_{max}}{K_M}[\text{S}]$ (else: $\frac{I_{enz}}{I_{max}} = \frac{[S]}{K_M + [S]}$)

Under steady-state diffusion-limited conditions, $I_d = \frac{nFD[S]}{d}$ (linear).

For best calibration, set $\frac{D}{d^2} = \frac{k_2[E_{total}]}{K_M}$ so that I_d is linear with gradient equal to $\frac{dI_{enz}}{d[S]}$ at $[\text{S}] = 0$.

Under this constraint, $\frac{I_d}{I_{max}} = \frac{D}{d^2 k_2[E_{total}]}[\text{S}]$ has a linear response.

The sensor is then linear for $[\text{S}] < K_M$. Too far outside this range would underestimate the concentration. The effective K_M can be tuned by diffusion-limiting membranes.

When a current is produced during operation in the working electrode, the biosensor circuitry detects the signal (transimpedance amplifier and ADC, Section 8.7.8), and after digital filtering (DSPs) the analyte concentration is computed (microcontroller).

Biosensors are used to detect e.g. glucose, lactate, cholesterol, urea and HbA1c (glycated haemoglobin) in biomedical diagnostics and monitoring. They can be coupled with control systems (e.g. insulin autoinjectors to treat type 1 diabetes).

Biofuel Cells: produced electricity is used as a power source

Biofuel cells are based on the same principle as the biosensor, with the aim of delivering high electrical power rather than linear responsiveness.

Organic fuel (e.g. glucose) can be oxidised in the enzymatic reaction to produce current through the electrode. Alternatively, the enzymatic reaction may be chosen to form a secondary fuel as a product from a substrate (e.g. hydrogen), which can be passed to a conventional fuel cell to cogenerate electricity. Microbial cells have also been used in this arrangement.

Performance Optimisation

Protein engineering techniques (Section 17.4.5) can be used to optimise the properties of the enzyme used in the device e.g. thermostabilisation, faster kinetics, higher specificity…

Using cofactors with lower reduction potentials (e.g. ferrocene/ferrocenium) helps to prevent unwanted oxidation of interferents in biofluids e.g. ascorbic acid in blood. Measuring from the blood requires diffusion coefficient correction due to hematocrit h: $D' = D_0\,(1 - h) / (1 + 0.952h)$.

Mediators may have poor electron transfer kinetics and are often toxic, motivating the ‘third generation biosensors’, which do not use mediators, and electrons are directly transferred from the cofactor to a nanofunctionalised electrode (e.g. conducting polymer/CNT surface). These can be prepared using electrodeposition during cyclic voltammetry and can be used for continuous implantable *in vivo* blood glucose monitors.

Enzyme-free biosensors have also been developed using electro-oxidation to drive the reaction with a transition metal complex replacing the cofactor with direct electron transfer to a nanomaterial matrix. These are the ‘fourth generation biosensors’. However, they have poor selectivity and can require unsuitable chemical conditions (e.g. high pH).

20.7.6. Surface Plasmon Resonance Biosensors

20.7.7. Tissue Engineering and Bioprinting

Regenerative medicine is an emerging field in biomedical science, involving the *in vivo* formation of personalised tissues (e.g. skin, bone, organ tissue). In order to successfully integrate with the host tissue, implanted tissues must contain:

- Matrix / scaffold: polymeric 3D structure mimicking the extracellular matrix (ECM).
- Signalling molecules: various biomolecules e.g. growth factors, Yamanaka factors
- Cells: live cells, depending on the type of tissue to be regenerated e.g. stem cells.

An alternative cell-free approach to tissue engineering involves removing cells from the formulation, and instead using 'homing' of host cells into the new scaffold, promoted by the biomolecules.

Scaffold: typically made of synthetic resorbable aliphatic polyesters (e.g. PCL, PTMC). It must have a porous interconnected structure (to allow nutrient flow), be biodegradable/bioresorbable at a rate matching the cell proliferation rate, have suitable surface chemistry to permit cell attachment and proliferation, have matching mechanical properties and be easily processed into various shapes.

The scaffold geometry must have very high surface area. For this purpose, geometries using triply periodic minimal implicit surfaces (e.g. Schwarz "P" unit cell) or space-filling curves can be used. Scaffolds can also be made by decellularising tissue *ex vivo*, leaving a natural matrix.

Bioprinting: automated and controllable deposition of matrix-cell-biomolecule bioinks.

Bioink contains live cells in a hydrogel-biomolecule formulation. The bioink may contain the fully-formed scaffold, or its precursors (controlled induced cross-linking). Bioprinting can be done by various methods, adapted from and inspired by the approaches in conventional 3D printing.

- Extrusion: continuous deposition of a viscous bioink stream from a syringe.
- Inkjet: directed emission of bioink droplets from a piezoelectric nozzle.
- Projection: UV-initiated cross-linking of photosensitive hydrogels (**not** polymerisation like photolithography).

The rheological properties of the bioink (gelation kinetics, surface tension, viscoelasticity, shear-dependent and time-dependent viscosity) are important design parameters, depending on the printing method. The addition of cross-linkers into the bioink allows the solidification of the hydrogel in the scaffold once deposited. Depending on whether the hydrogel is UCST or LCST (Section 16.5.11), either the syringe or stage must be heated to form the gel before deposition.

The cells can be supported on microcarriers suspended in the bioink, which are made of either natural (cellulose, gelatin, collagen) or synthetic (e.g. dextran, plastic, glass) hard materials. This arrangement permits high cell counts to be used without compromising viability, as well as improved nutrient exchange. Some types of cell e.g. chondrocytes can help to replace the scaffold with new ECM tissue if rigidity is desired. Functional materials (e.g. graphene, magnetic fibres, metal nanoparticles) can also be supplied in the bioink with careful control of suspension rheology and biocompatibility.

Like other CAM hardware, bioprinters can be designed to print from inputs given by parametric equations for the toolpath, G-code (low-level language), CAD models or images (by segmentation).

~ LN, 2024