

Topic 01

University Medicine

What this course is, and how to use it

This is the first course in University Medicine: **cell and molecular biology for medicine**. It is the preclinical base every later course stands on: biochemistry fills the cell, physiology makes it work, pathology explains when it breaks.

Scope limit, stated once: this is a study course, not a medical degree – no patients, no clinical placements; every case below is fictional and nothing here licenses or guides treatment.

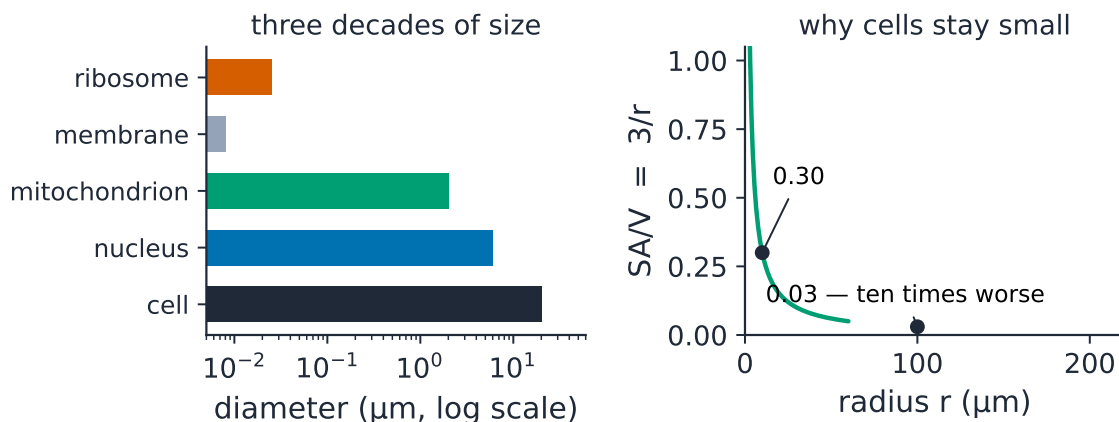
- Units 1.1–1.2 are the cell and its membranes: organisation, transport, and the electrical language of ions.
- Units 1.3–1.5 are the information line: replication, gene expression, and its regulation.
- Units 1.6–1.7 are decision-making: signalling between cells, and the cell's own life-and-death controls.

The medicine habit: **mechanism before fact**. Every number in this course is something you recompute – a Nernst potential, a compaction ratio, a mutation probability – and every clinical link names the mechanism it teaches.

Cell organisation and organelles

Why it matters. A eukaryotic cell is a factory of compartments: each organelle is a reaction vessel with the right pH, the right enzymes and the right surface. Medicine reads a cell the way an engineer reads a plant diagram – what lives where, and why.

The idea, with a picture. Organelles live on a size ladder spanning three orders of magnitude: a **ribosome** is 25 nm, a **mitochondrion** 2 μm , the **nucleus** 6 μm , the whole **cell** 20 μm . Function follows size: anything the whole cell must do quickly needs a membrane surface to do it on.



Big cells have proportionally little membrane – the surface-to-volume argument caps cell size

Definitions with numbers. The **nucleus** stores 2.2 m of DNA (6.4×10^9 bp $\times$ 0.34 nm/bp) inside a 6 μm envelope – a length about 4×10^5 times the nucleus's width. The **rough endoplasmic reticulum** is ribosome-studded protein export; the **Golgi** sorts and glycosylates; **lysosomes** keep pH $\approx$ 4.8 to dismantle cargo; **mitochondria** make ATP. **Ribosomes** translate mRNA (unit 1.4).

The derivation. Transport in and out scales with membrane **surface area** ($4\pi r^2$); demand – protein synthesis, ATP turnover – scales with **volume** ($\frac{4}{3}\pi r^3$). The ratio $SA/V = 3/r$: double the radius and every unit of cytoplasm gets half the membrane service. That single line is why cells stay small, why the ER folds so much surface into the cytoplasm, and why the intestinal lining is villi rather than flat.

Worked example 1 – the size ladder, verified. A 20 μm hepatocyte: $SA/V = 3/10 \mu m^{-1} = 0.3$. A 200 μm sphere (an oocyte): 0.03 – ten times worse ✓, which is why large cells are rare and membrane-folded. Cross-check the ladder: 20 μm / 25 nm = 800 – a ribosome is as small against a cell as a cell is against a 1.6 cm marble ✓.

Worked example 2 – clinical link: the lysosome that fails. A (fictional) patient with progressive muscle weakness is found to lack a single lysosomal hydrolase. Mechanism, not name, is the answer: the substrate cannot be dismantled at pH 4.8, so it accumulates, the lysosome swells and the cell dies by being unable to recycle. One missing enzyme in one compartment = one storage disease: the organelle map is the diagnosis map.

Common mistakes. Treating “the cytoplasm” as one well-mixed bag – compartments differ in pH, calcium and redox state by design. Reading organelle counts as fixed (a liver cell has ~1000 mitochondria; a quiescent lymphocyte, few). And forgetting the membrane is where the cell meets the world.

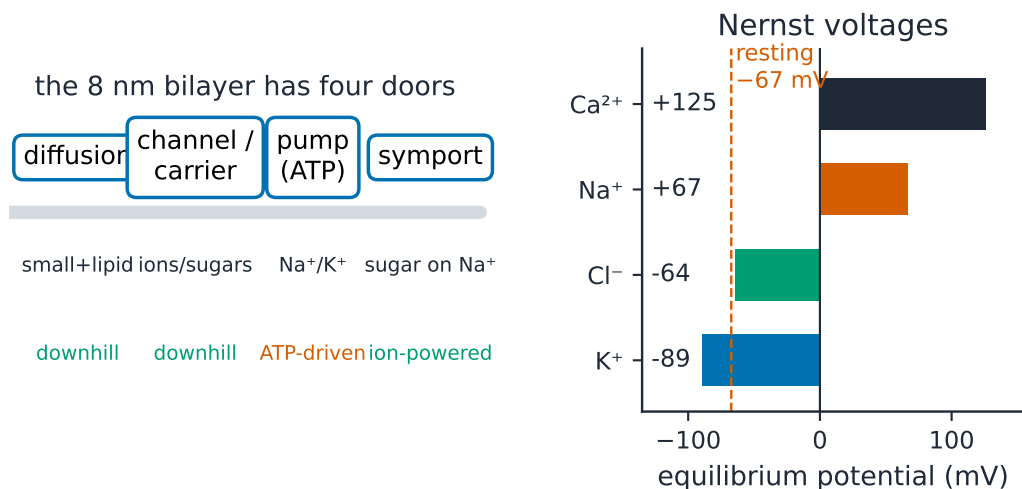
Summary. Organelles are reaction vessels on a 25 nm–20 μm ladder; $SA/V = 3/r$

caps cell size and explains folded membranes; 2.2 m of DNA packs into a 6 μm nucleus; compartment failure is disease.

Membranes and transport

Why it matters. Every drug target, every nerve impulse and every kidney's work reduces to one question: how does a molecule cross a lipid bilayer? The answer is always one of four doors.

The idea, with a picture. The **lipid bilayer** is a self-assembled 8 nm sheet impermeable to ions. Crossing is by: **simple diffusion** (small, lipophilic), **facilitated diffusion** through channels or carriers (down gradient, no ATP), **primary active transport** (ATP pumps, against gradient), and **secondary active transport** (an ion gradient drives something else uphill).



Four doors, four rules – and the equilibrium voltage each ion would need to stop moving

The Nernst equation. For ion X at 37 °C the **Nernst equation** gives the voltage at which the ion stops moving:

$$E_X = \frac{RT}{zF} \ln \frac{[X]_{out}}{[X]_{in}} = \frac{61.5 \text{ mV}}{z} \cdot \log_{10} \frac{[X]_{out}}{[X]_{in}}$$

Typical mammalian values – K⁺ (140 in / 5 out): $E_K = 61.5 \cdot \log_{10}(5/140) = -89.0$ mV ✓; Na⁺ (12/145): +66.6 mV; Cl⁻ (10/110, $z = -1$): -64.0 mV; Ca²⁺ (100 nM/1.2 mM, $z = 2$): +125.4 mV.

The derivation. At equilibrium the electrical force balances the chemical: $zFE = RT \ln([out]/[in])$ – that IS the Nernst line. With several ions permeable at once, replace the single ratio with a permeability-weighted sum (Goldman–Hodgkin–Katz):

with $P_K : P_{Na} : P_{Cl} = 1 : 0.04 : 0.45$,

$$V_m = 61.5 \log_{10} \frac{[K]_o + 0.04[Na]_o + 0.45[Cl]_i}{[K]_i + 0.04[Na]_i + 0.45[Cl]_o} = -67.3 \text{ mV}$$

✓ the **resting potential** – about 20 mV above E_K , because Na^+ leaks in and the Na^+/K^+ -ATPase (3 Na^+ out, 2 K^+ in, 1 ATP) holds the gradients it needs.

Worked example 1 – hyperkalaemia arithmetic. A (fictional) patient's extracellular K^+ doubles from 5 to 10 mM. New Nernst: E_K shifts by $61.5 \log_{10}(10/5) = +18.5$ mV, from -89 to -71 mV ✓ – the depolarised E_K flattens the resting voltage, which is why raised potassium is a cardiac emergency: the mechanism is one logarithm.

Worked example 2 – clinical link: a broken chloride door. The most common cystic-fibrosis variant deletes one phenylalanine ($\Delta F508$) from a Cl^- channel, so the channel misfolds and never reaches the membrane. Follow the mechanism: Cl^- cannot leave $\rightarrow$ less salt reaches the airway surface, so less water follows it by osmosis $\rightarrow$ the surface liquid dehydrates $\rightarrow$ mucus cannot clear. One folding defect, one door, one disease – and small molecules that hold the fold open are the therapy rationale.

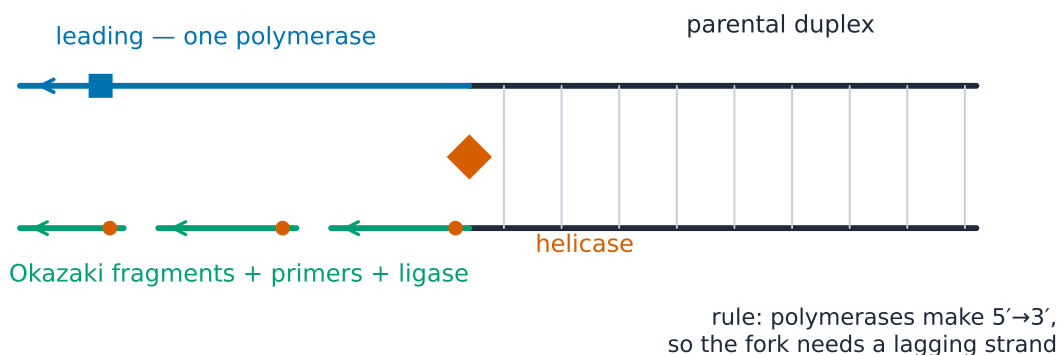
Common mistakes. Using the Nernst equation on non-equilibrium gradients (it is an equilibrium statement; GHK is the honest steady approximation). Forgetting z in the denominator (Ca^{2+} halves the slope). Calling channels “active” – only pumps burn ATP.

Summary. Four doors cross the bilayer; $E_X = (61.5/z) \log_{10}([out]/[in])$ mV at $37^\circ C$; GHK with $P_K:P_{Na}:P_{Cl} = 1:0.04:0.45$ gives -67 mV; 3 Na^+ out / 2 K^+ in per ATP keeps the battery charged.

DNA replication and repair

Why it matters. A cell must copy 6.4 billion letters with almost no mistakes, once per division, forever. The machinery that manages this – and the repair systems behind it – is where hereditary disease, mutagenesis and most oncology begin.

The idea, with a picture. The **replication fork** : helicase unzips, polymerases extend $5' \rightarrow 3'$ only, so one strand is copied continuously (**leading**) and the other backwards in **Okazaki fragments** (**lagging**), each restarted by a primer.



One antiparallel rule forces the fork into a semiconservative machine with a lagging strand

Definitions with numbers. Rate: bacterial polymerase ~ 1000 nt/s; human ~ 50 nt/s per fork. Fidelity: raw polymerase error $\sim 10^{-5}$ per bp; **proofreading** buys $\times 100 \rightarrow 10^{-7}$; **mismatch repair** another $\times 100 \rightarrow 10^{-9}$ per bp ✓ – three systems deep, one billionth.

The derivation (the timing argument). Human S phase copies 6.4×10^9 bp in 8 h. A single origin with two forks at 50 bp/s would need $6.4 \times 10^9 / 100 = 6.4 \times 10^7$ s $\approx$ **2 years**. Replication works because $\sim 30\,000$ origins fire in timed clusters: $6.4 \times 10^9 / (100 \times 3 \times 10^4) \approx 2100$ s $\approx$ 35 min of pure synthesis ✓, wrapped in chromatin time to fill S phase.

Worked example 1 – the fidelity budget. After proofreading and mismatch repair the residual rate is $\sim 10^{-9}$ /bp/division. For a 6.4×10^9 bp diploid genome: $6.4 \times 10^9 \times 10^{-9} \approx 6$ new mutations per daughter cell ✓ – the number oncology (unit 1.7) and genetics both lean on. Skip mismatch repair and the rate rises $\times 100$: hundreds of mutations per division, which is hereditary non-polyposis colon cancer's mechanism.

Worked example 2 – clinical link: sunlit skin without repair. (Fictional patient) severe sunburn blisters after minutes outdoors, skin cancers by twenty. Mechanism: ultraviolet light fuses adjacent thymines into dimers; **nucleotide excision repair** is the system that cuts them out; without it every sunbeam is a mutagen with no auditor. The lesson generalises: each repair pathway owns one lesion class, and its absence names one syndrome.

Common mistakes. Reading polymerases as reading $3' \rightarrow 5'$ – they MOVE $3' \rightarrow 5'$ along the template while MAKING $5' \rightarrow 3'$. Calling S phase "copying once" without the origins argument. Quoting 10^{-9} as the raw polymerase rate (it is the post-repair rate).

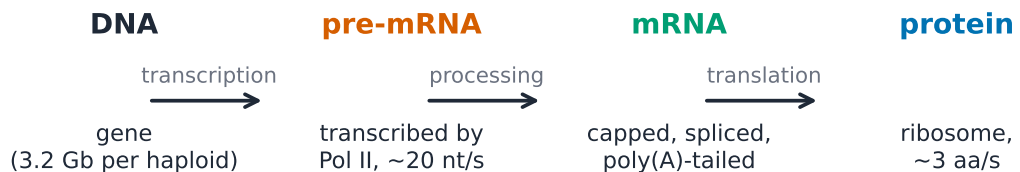
Summary. Semiconservative, $5' \rightarrow 3'$, one leading + one lagging strand; $10^{-5} \rightarrow 10^{-7} \rightarrow 10^{-9}$ fidelity ladder; $\sim 30\,000$ origins turn 2 years into 8 hours; ~ 6 mutations

survive per division; each repair pathway owns one lesion class.

Gene expression: transcription and translation

Why it matters. The gene is a recipe, not a machine: between DNA and protein stand RNA polymerase, processing, splicing and the ribosome – and most human disease mutations, it turns out, break the in-between steps, not the recipe itself.

The idea, with a picture. The **central dogma** with its eukaryotic detours: DNA → **pre-mRNA** mRNA (**transcription** by Pol II at ~20 nt/s) → 5' cap + poly(A) + **splicing** (introns out, exons in) → mature mRNA → ribosome (**translation**) → protein (~3 amino acids/s). One gene, alternative exons, several proteins.



splicing: one gene → several proteins (alternative exons)

the code: $4^3 = 64$ codons = 61 sense + 3 stop, for 20 amino acids

Transcription, processing, translation – three stages, three chances for regulation

Definitions with numbers. The genetic code: 64 **codons** = 61 sense + 3 stop; 20 amino acids from 4 bases in threes ($4^3 = 64$ ✓ – the arithmetic OF the code). Degeneracy averages ~3 codons per amino acid; AUG starts (methionine), UAA/UAG/UGA stop.

The derivation (the counting argument). Why triplets? Two bases give $4^2 = 16 < 20$: too few. Three give $4^3 = 64 > 20$: enough, with surplus spent on degeneracy – mostly at the third position, which is why third-base substitutions are so often silent. The code's redundancy is not waste; it is damage tolerance.

Worked example 1 – timing a protein. Preproinsulin is 110 amino acids: translation at ~3 aa/s takes ~37 s; its 0.5 kb mRNA took ~25 s to transcribe at 20 nt/s. The numbers set expectations: making one insulin molecule costs about a minute of machinery, which is why secretion systems pre-build and store (unit 1.6 signalling triggers release, not synthesis).

Worked example 2 – clinical link: the splice that breaks. (Fictional pedigree) anaemia, and sequencing shows no coding mutation – the defect is a single base that creates a new splice site inside a β -globin intron. Mechanism: the spliceosome obeys consensus sequences; a new site lures it, a stretch of the intron is kept in the mRNA, the reading frame shifts and a premature stop follows. β -thalassaemia without a single amino-acid change: the dogma's middle step is the disease.

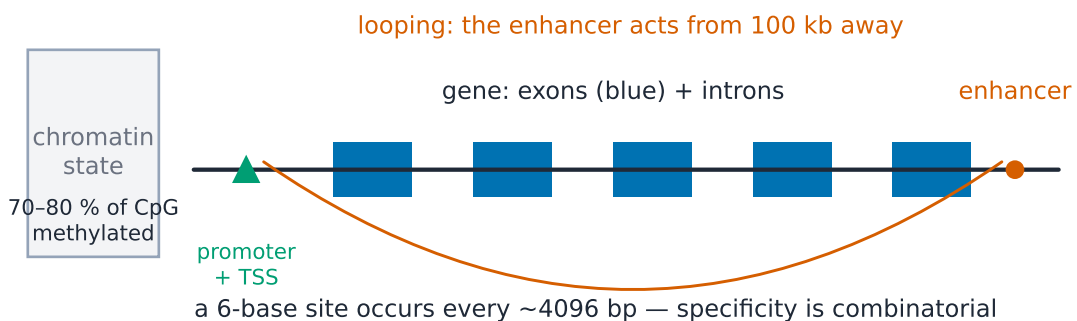
Common mistakes. Calling introns “junk” (regulatory and splicing-adjacent function, unit 1.5). Confusing the template strand with the coding sequence (the mRNA matches the coding strand, T→U). Expecting one gene = one protein (alternative splicing).

Summary. Pol II ~ 20 nt/s $\rightarrow$ processing (cap, poly(A), splice) $\rightarrow$ ribosome ~ 3 aa/s; $4^3 = 64$ codons for 20 amino acids with third-base degeneracy; one insulin molecule $\approx$ one minute; splice-site mutations break genes without touching their protein code.

Regulation of gene expression

Why it matters. Every one of your $\sim 20\,000$ genes exists in nearly every cell; what makes a neuron a neuron is which genes are ON, when, and how much. Regulation is the actual genome – and the most common site of both evolution and disease.

The idea, with a picture. Control stacks: **transcription factors** bind promoters and **enhancers** (which can act from 100 kb away by DNA looping); **chromatin** state – histone marks and CpG **methylation** – decides access; **epigenetics** is the heritable part of that access without changing sequence.



Enhancer, promoter, transcription start, exons – and the chromatin gate in front of them

Definitions with numbers. Only $\sim 1.5\%$ of the genome codes protein; regulatory sequence dwarfs it. 70–80 % of CpG dinucleotides in a differentiated cell are methylated – the switched-off baseline. The 2.2 m of DNA (unit 1.1) wraps nucleosomes (~ 200 bp each, about 7-fold shorter) folded into a ~ 30 nm fibre (about 40-fold); loops

and further folding bring a mitotic chromosome near 10^4 -fold.

The derivation (the combinatorics of specificity). A transcription factor recognises ~6–8 bases. Random 6-mers occur every ~4096 bp – about 800 000 times along each strand of the 3.2 Gb haploid genome: far too often for clean regulation. Specificity comes from COMBINATIONS: several factors binding adjacent sites cut false hits multiplicatively (4^{-6} to the power of cooperating sites) – the same logic that makes passwords long. Enhancers are cooperative logic gates, which is why regulation is network-shaped.

Worked example 1 – lactase persistence. The enzyme LCT falls after weaning in most of the world's population; northern Europeans carry a single base change ~14 kb UPSTREAM – inside an intron of a neighbouring gene, where it creates new enhancer binding. Mechanism chain: one regulatory base → one enhancer strengthened → LCT stays on into adulthood → milk digestible. A zero-protein-change trait that swept a continent: regulation evolves fastest.

Worked example 2 – clinical link: silencing gone wrong. (Fictional twin pair) identical genotypes, different severity: one carries methylation of the disease gene's promoter, the other does not. Same letters, different access – an **epigenetic** modifier of severity. The mechanism explains why “identical genes” never promises identical disease, and why methylation assays sit next to sequencers in the diagnostic laboratory.

Common mistakes. Equating epigenetic with permanent (marks divide with the cell but can be reprogrammed). Treating enhancers as promoters-with-a-different-name (orientation- and distance-independent vs the fixed start site). Reading “gene ON/OFF” as binary – expression is graded and tissue-timed.

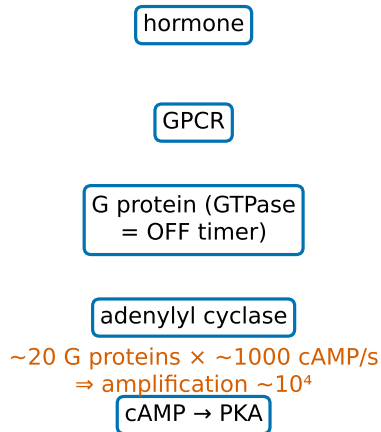
Summary. ~20 000 genes, ~1.5 % coding; enhancers are cooperative combinatorial switches acting by looping; methylation (70–80 % of CpG) and histone marks gate access heritably; one upstream base can rewire a phenotype.

Cell signalling

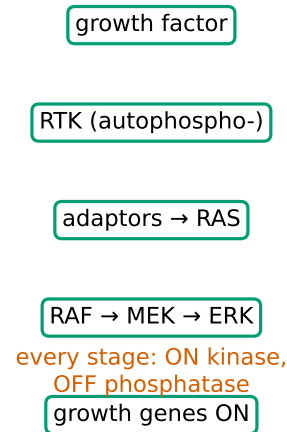
Why it matters. Cells are not self-sufficient: hormones, neurotransmitters and growth factors are instructions arriving at the membrane. Signal transduction is how a message at the surface becomes biochemistry inside – and roughly half of all drugs act at this layer.

The idea, with a picture. Two canonical cascades: a **GPCR** G (the hormone never enters; a G-protein switch relays to **second messengers** like cAMP) and an **RTK** (receptor phosphorylates itself; adaptor proteins pass the signal to RAS→MAPK growth commands). Both **AMPLIFY**: one surface event becomes thousands of intracellular events.

GPCR — second messenger



RTK — kinase cascade



One molecule at the membrane becomes tens of thousands inside – amplification is the design

Definitions with numbers. Resting cAMP $\sim 10^{-8}$ M; a maximal hormone pulse drives it to $\sim 10^{-6}$ M – a hundredfold switch in seconds. One receptor can activate tens of G-proteins; each adenylyl cyclase makes ~ 1000 cAMP/s: amplification $\sim 10^4$ from hormone to second messenger ✓. GTPase activity of the $G\alpha$ subunit is the built-in OFF timer (seconds).

The derivation (the amplification product). Signal output = $\prod$ (stage gains): hormone $\rightarrow$ receptor (1 : ~ 20 G-proteins) $\rightarrow$ cyclase (each ~ 1000 cAMP/s) $\rightarrow$ PKA (each phosphorylates many targets). The gains MULTIPLY, so the cascade is exponential in gain and, crucially, every stage is a control point: phosphodiesterase destroys cAMP, GTPase switches off $G\alpha$, phosphatases strip phosphates. Amplification and reset are the same machinery seen from two sides.

Worked example 1 – the cAMP ledger. Basal 10 nM in a 2×10^{-12} L hepatocyte is 10^{-8} mol/L $\times 2 \times 10^{-12}$ L $\times N_A \approx 12\,000$ molecules ✓. One second of 10 activated cyclases at 1000/s would produce 10 000 cAMP – the entire resting pool is replaced in about a second, so the OFF enzymes matter exactly as much as the ON ones.

Worked example 2 – clinical link: the switch that cannot turn off. (Fictional outbreak vignette) profuse secretory diarrhoea within hours. Mechanism: cholera toxin ADP-ribosylates the $G\alpha_s$ subunit, killing its GTPase – the OFF timer is disabled, adenylyl cyclase runs constitutively, cAMP stays at maximum, and the intestinal cell pumps Cl^- (unit 1.2's door) and water into the lumen. Therapy is fluid replacement while the toxin-modified proteins decay: you cannot inhibit a switch that is already welded on.

Common mistakes. Saying the hormone “enters the cell” (GPCR ligands do not; steroid hormones do – different mechanism entirely). Forgetting second messengers

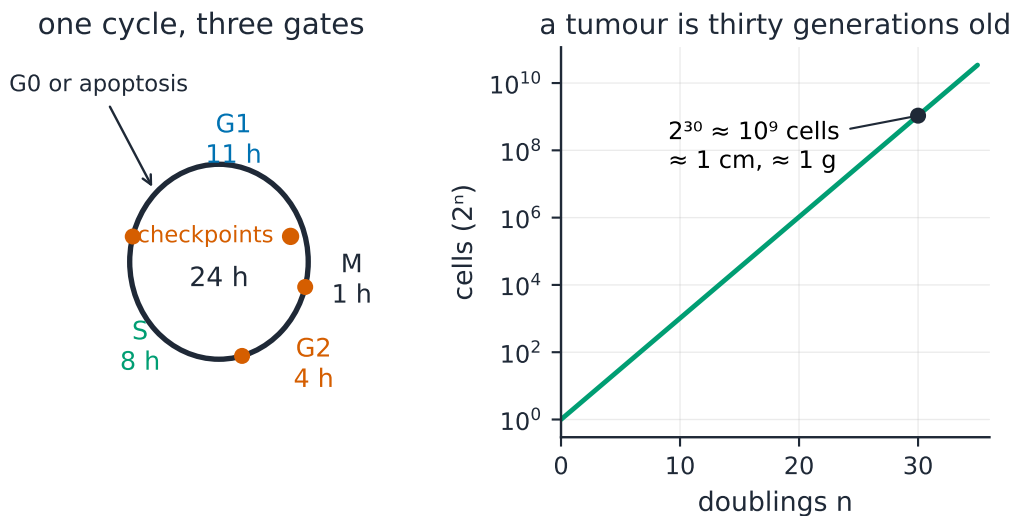
are destroyed continuously (PDE vs caffeine and other inhibitors). Reading kinase cascades as linear (feedback everywhere).

Summary. GPCR → G-protein → cAMP → PKA; RTK → RAS → MAPK; stage gains multiply to $\sim 10^4$; every ON stage has an OFF enzyme; cholera is the GTPase welded open.

Cell cycle, cell death and cancer biology

Why it matters. Division and death are the cell's most expensive decisions, gated by checkpoints. Cancer is not one disease but the accumulated failure of those gates – and every gate is a named molecular system you have already met (repair, signalling, regulation).

The idea, with a picture. The cell cycle : **G1** (growth, ~ 11 h) → **S** (DNA replication, ~ 8 h, unit 1.3) → **G2** (~ 4 h) → **M** (mitosis, ~ 1 h), with **checkpoints** at G1/S (the restriction point), G2/M and the spindle assembly check. Cells can leave to **G0** quiescence – or to **apoptosis**, programmed death, executed by caspases.



Three gates around one cycle – and thirty doublings from one cell to a visible tumour

Definitions with numbers. A typical cycling human cell: 24 h total (11 + 8 + 4 + 1) ✓. **p53** (guardian of the genome) pauses the cycle after DNA damage for repair – or commits the cell to apoptosis. Cancer's requirement: ~ 4 – 7 rate-limiting driver events, plus replicative immortality (telomerase) and escape from apoptosis.

The derivation (the doubling count). One cell doubling unchecked: after n divisions there are 2^n cells. A 1 cm³ tumour masses about 1 g $\approx 10^9$ cells: $2^{30} = 1.07 \times 10^9$ ✓ – a detectable tumour is already THIRTY generations old. The clinical implication is arithmetic: screening intervals and “why did it appear so fast” (it did not – it was silent for years) live on this exponential.

Worked example 1 – the driver-event lottery. Unit 1.3: ~6 mutations survive per division. Most hit passengers (vast non-coding space, unit 1.5); the chance any given mutation disables a specific gate is small – but 10^9 cells $\times$ decades buys many trials, each driver raising the rate further. Cancer incidence rising with age is this arithmetic: $\sim 10^{16}$ divisions in a lifetime, each holding a lottery ticket.

Worked example 2 – clinical link: a family born one gate down. (Fictional pedigree) early sarcoma, breast cancer and a brain tumour across three generations – one inherited mutated p53 allele (Li–Fraumeni pattern). Mechanism: every cell starts one gate down; a single somatic hit on the second allele removes the checkpoint entirely. That is the two-hit logic of a **tumour suppressor gene**, and it ties the course together: repair (1.3), signalling (1.6) and regulation (1.5) are the gates cancer must pass.

Common mistakes. Treating apoptosis as passive death (it is active, programmed, ATP-dependent – necrosis is the messy one). Reading oncogenes and tumour suppressors as the same (gain-of-function vs loss-of-function; one hit vs two). Equating fast growth with fast cycles (usually fewer cells dying, not faster divisions).

Summary. 24 h cycle, three checkpoints, G0 and apoptosis as exits; p53 is the damage gate; $2^{30} \approx 10^9$ cells = 1 cm³ = thirty silent doublings; 4–7 drivers plus immortality.