Introduction to Biomedical Datascience

Lecture 1

From Atoms to Life: Molecular Basis of Biology

Instructor Information

Course Email:

Lecture Contents

Part 1 of 3

Atomic and Molecular Foundations

Understanding the chemical basis of life
From quantum orbitals to biological macromolecules

Atoms and Electron Orbitals

Electron Configuration

- Electrons occupy specific energy levels
- Quantum numbers define orbital characteristics
- Aufbau principle: fill lowest energy first
- Pauli exclusion principle

Valence Electrons

- Outermost shell electrons
- Determine chemical reactivity
- Participate in bond formation
- Critical for biological interactions

Orbital Shapes

s
Spherical

p
Dumbbell

d
Complex

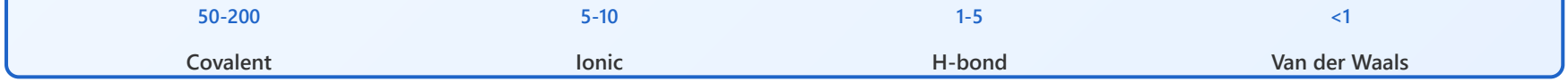
f
Very Complex

Biological Elements (CHNOPS)

C • H • N • O • P • S

Chemical Bonds in Biology

Bond Energy Comparison



Covalent Bonds

- Strong electron sharing
- Single, double, triple bonds
- Form backbone of biomolecules
- Energy: 50-200 kcal/mol

Ionic Interactions

- Electrostatic attractions
- Important in protein folding
- Salt bridges stabilize structures
- Energy: 5-10 kcal/mol

Hydrogen Bonds

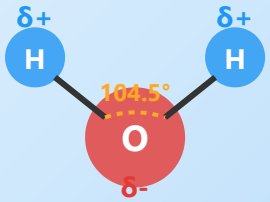
- Weak but numerous
- Critical for DNA base pairing
- Protein secondary structure
- Energy: 1-5 kcal/mol

Van der Waals Forces

- Weakest interactions
- Important in close packing
- Hydrophobic effect
- Energy: <1 kcal/mol

Water - The Solvent of Life

H_2O Molecular Structure



Unique Properties

- High heat capacity
- High heat of vaporization
- Less dense as solid (ice floats)
- Excellent solvent for polar molecules

Hydrophobic Effect

- Nonpolar molecules cluster together
- Drives protein folding
- Forms lipid bilayers
- Entropy-driven process

Molecular Structure

- Bent geometry (104.5°)
- Polar covalent O-H bonds
- Partial charges: δ- O, δ+ H
- Strong dipole moment

Solvation

- Water molecules surround ions
- Hydration shells stabilize charges
- Affects biochemical reactions
- Critical for ion transport

H-Bond Network

Hydrogen Bond Network

pH and Biological Systems

pH Scale

- pH = -log[H^+]
- Range: 0 (acidic) to 14 (basic)
- pH 7 is neutral
- Each unit = 10× concentration change

Buffer Systems

- Resist pH changes
- Blood pH: 7.35-7.45
- Bicarbonate buffer (H_2CO_3/HCO_3^-)
- Phosphate buffer in cells

Henderson-Hasselbalch Equation

- pH = pKa + log([A^-]/[HA])
- Predicts buffer behavior
- Critical for enzyme function
- Used in drug design

Enzyme pH Dependence

- Each enzyme has optimal pH
- Pepsin (stomach): pH 2
- Trypsin (intestine): pH 8
- pH affects protein charge state

Amino Acids Structure

General Structure

- Central α-carbon
- Amino group (-NH_2)
- Carboxyl group (-COOH)
- Variable R group (side chain)

Classification

- Nonpolar/hydrophobic
- Polar uncharged
- Positively charged (basic)
- Negatively charged (acidic)

Chirality

- All are L-amino acids in proteins
- D-amino acids rare in nature
- Asymmetric α-carbon
- Mirror image isomers

Ionization States

- pKa values determine charge
- Zwitterion at physiological pH
- Affects protein interactions
- Important for enzyme catalysis

Protein Structure Levels

Primary (1°)

Amino acid sequence connected by peptide bonds

Secondary (2°)

Local folding patterns: α-helix and β-sheet

Tertiary (3°)

Overall 3D structure of single polypeptide chain

Quaternary (4°)

Assembly of multiple polypeptide subunits

Nucleotides and DNA Structure

Nucleotide Components

- Nitrogenous base (A, T, G, C)
- 5-carbon sugar (deoxyribose)
- Phosphate group
- Connected by glycosidic bond

Bases

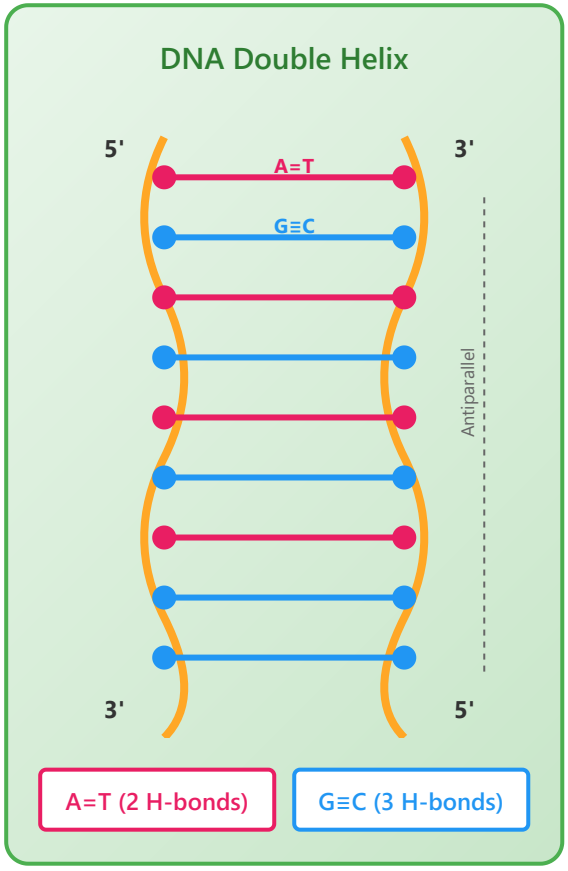
- Purines: Adenine (A), Guanine (G)
- Pyrimidines: Thymine (T), Cytosine (C)
- Watson-Crick base pairing
- A=T (2 H-bonds), G≡C (3 H-bonds)

DNA Double Helix

- Antiparallel strands
- Sugar-phosphate backbone
- Major and minor grooves
- Right-handed B-form (most common)

Structural Parameters

- Diameter: ~2 nm
- Rise per base pair: 0.34 nm
- 10.5 base pairs per turn
- Pitch: 3.57 nm



Part 2 of 3

Central Dogma

DNA → RNA → Protein

The flow of genetic information

DNA Replication Mechanism

Semiconservative Replication

- Each strand serves as template
- Two identical daughter DNA molecules
- Proven by Meselson-Stahl experiment

Key Enzymes

- **Helicase:**Unwinds DNA helix
- **Primase:**Synthesizes RNA primers
- **DNA Pol III:**Main replication (5'→3')
- **DNA Pol I:**Removes primers
- **Ligase:**Joins Okazaki fragments

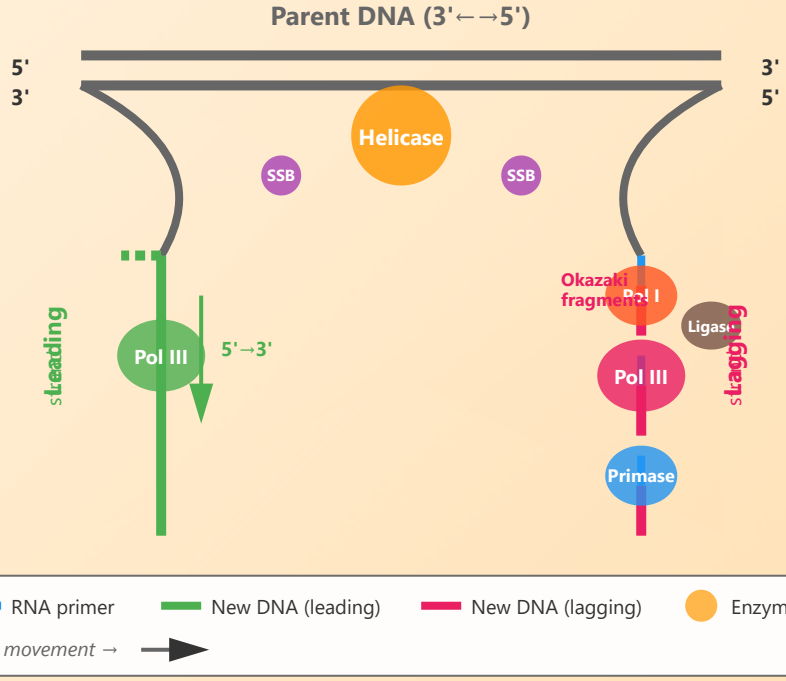
Leading vs Lagging

- **Leading:**Continuous synthesis
- **Lagging:**Discontinuous (Okazaki)
- Fragment size: 1000-2000 nt

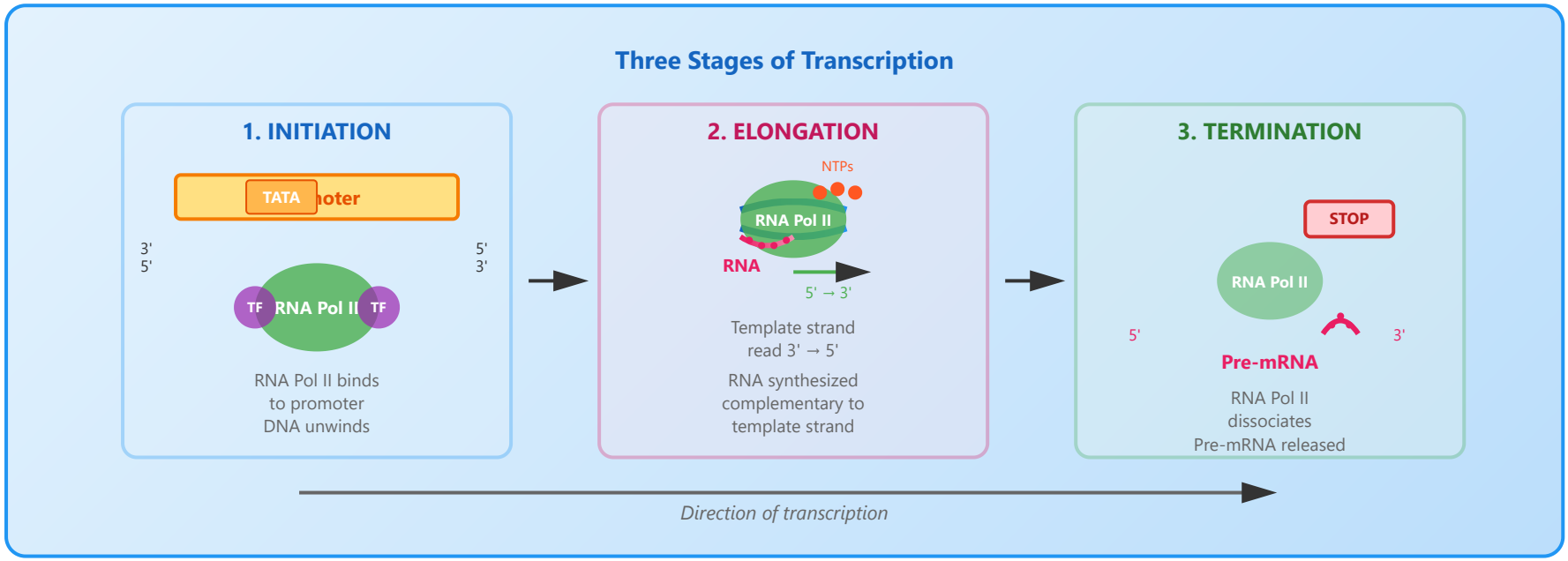
✓ Proofreading & Fidelity

- 3'→5' exonuclease activity
- Error rate: ~1 in 10^7 bases
- Mismatch repair systems

DNA Replication Fork - Detailed Mechanism



Transcription Process



Initiation

- RNA polymerase binds promoter
- TATA box recognition
- Transcription factors assist
- DNA unwinds at start site

Elongation

- RNA synthesized 5' → 3'
- Template strand read 3' → 5'
- Ribonucleotides added
- Transcription bubble moves

Termination

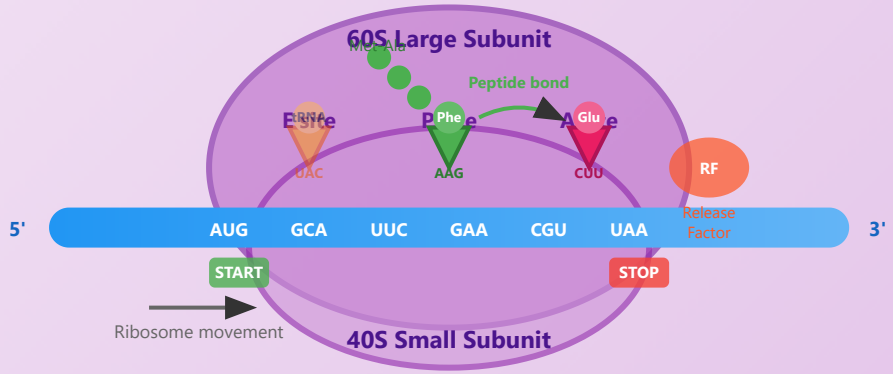
- Rho-dependent or independent
- Hairpin structure formation
- RNA polymerase dissociates
- Pre-mRNA released

RNA Processing

- 5' cap (7-methylguanosine)
- 3' poly(A) tail
- Splicing removes introns
- Mature mRNA produced

Translation and Genetic Code

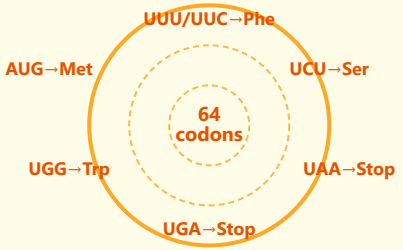
Translation Process



Steps:

1. Initiation: Ribosome assembles at AUG
2. Elongation: tRNAs bring amino acids
3. Peptide bonds form in P site
4. Ribosome translocates 3 nucleotides
5. Termination: Release at stop codon

Genetic Code



- 61 amino acid codons
- 3 stop codons
- Degenerate code

Ribosome Structure

- Large subunit (60S in eukaryotes)
- Small subunit (40S in eukaryotes)
- rRNA and ribosomal proteins
- Three tRNA sites: A, P, E

tRNA Function

- Anticodon pairs with codon
- Carries specific amino acid
- Wobble base pairing
- Aminoacyl-tRNA synthetases

Translation Steps

- **Initiation:**AUG start codon

- **Elongation:**peptide formation
- **Termination:**UAA, UAG, UGA
- Energy: 2 GTP per amino acid

Gene Regulation Overview

Transcriptional Control

- Promoter accessibility
- Transcription factor binding
- RNA polymerase recruitment
- Primary regulation point

Enhancers and Silencers

- Regulatory DNA sequences
- Can be far from gene
- Increase or decrease transcription
- Bind transcription factors

Chromatin Remodeling

- ATP-dependent complexes
- Alter nucleosome positioning
- Expose or hide DNA
- Control gene accessibility

Post-transcriptional

- mRNA stability regulation
- Alternative splicing
- MicroRNA regulation
- Translation control

Epigenetic Modifications

DNA Methylation

- Addition of methyl groups to cytosine
- CpG islands near promoters
- Gene silencing mechanism
- Maintained through cell division

Histone Modifications

- Acetylation: gene activation
- Methylation: activation or repression
- Phosphorylation: chromatin structure
- Histone code hypothesis

Chromatin States

- Euchromatin: transcriptionally active
- Heterochromatin: transcriptionally silent
- Dynamic transitions
- Cell type-specific patterns

Disease Implications

- Cancer: aberrant methylation
- Imprinting disorders
- X-chromosome inactivation
- Environmental influences

RNA Types and Functions

Messenger RNA (mRNA)

- Encodes protein information
- Short-lived in cells
- 5' cap and poly(A) tail
- Template for translation

Transfer RNA (tRNA)

- Adapter molecule
- Brings amino acids to ribosome
- ~75-90 nucleotides
- Post-transcriptional modifications

Small Regulatory RNAs

- miRNA: post-transcriptional silencing
- siRNA: gene knockdown
- ~20-25 nucleotides
- Therapeutic potential

Long Non-coding RNAs

- >200 nucleotides
- Chromatin remodeling
- Transcription regulation
- Emerging therapeutic targets

Protein Folding and Misfolding

Anfinsen's Principle

- Sequence determines structure
- Spontaneous folding possible
- Minimum free energy state
- Reversible denaturation

Chaperone Proteins

- Assist protein folding
- Prevent aggregation
- HSP70, HSP90 families
- ATP-dependent mechanisms

Folding Funnels

- Energy landscape model
- Multiple pathways to native state
- Local minima can trap
- Kinetic vs thermodynamic control

Misfolding Diseases

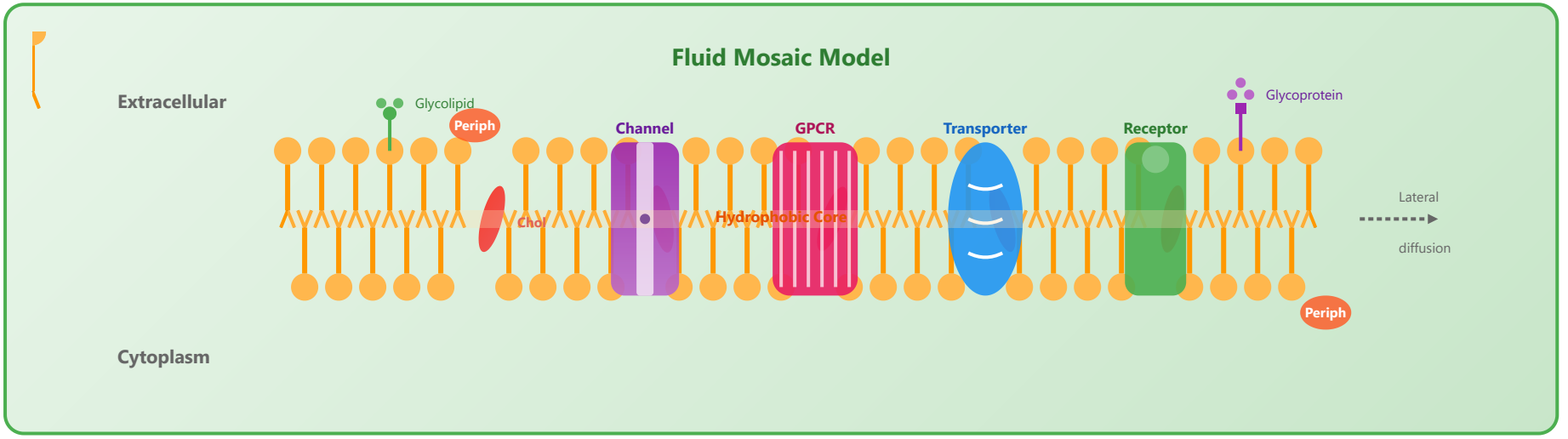
- Alzheimer's: Aβ plaques
- Parkinson's: α-synuclein
- Prion diseases: PrP
- Therapeutic targets

Part 3 of 3

Cellular Systems

Integration of molecular processes
The cell as a functional unit

Cell Membrane Structure



Lipid Bilayer

- Phospholipids: hydrophobic core
- Cholesterol: membrane fluidity
- Glycolipids: cell recognition
- Asymmetric distribution

Membrane Proteins

- Integral: span membrane
- Peripheral: surface attachment
- Channels and transporters
- Receptors and enzymes

Fluid Mosaic Model

- Dynamic structure
- Lateral diffusion of components
- Restricted rotation
- Temperature-dependent fluidity

Transport Mechanisms

- Passive: down concentration gradient
- Active: against gradient (ATP)
- Facilitated diffusion
- Endocytosis and exocytosis

Organelles and Functions

Nucleus

- Houses genetic material
- Nuclear envelope with pores
- Nucleolus: rRNA synthesis
- Chromatin organization

Endoplasmic Reticulum

- Rough ER: protein synthesis
- Smooth ER: lipid synthesis
- Calcium storage
- Detoxification

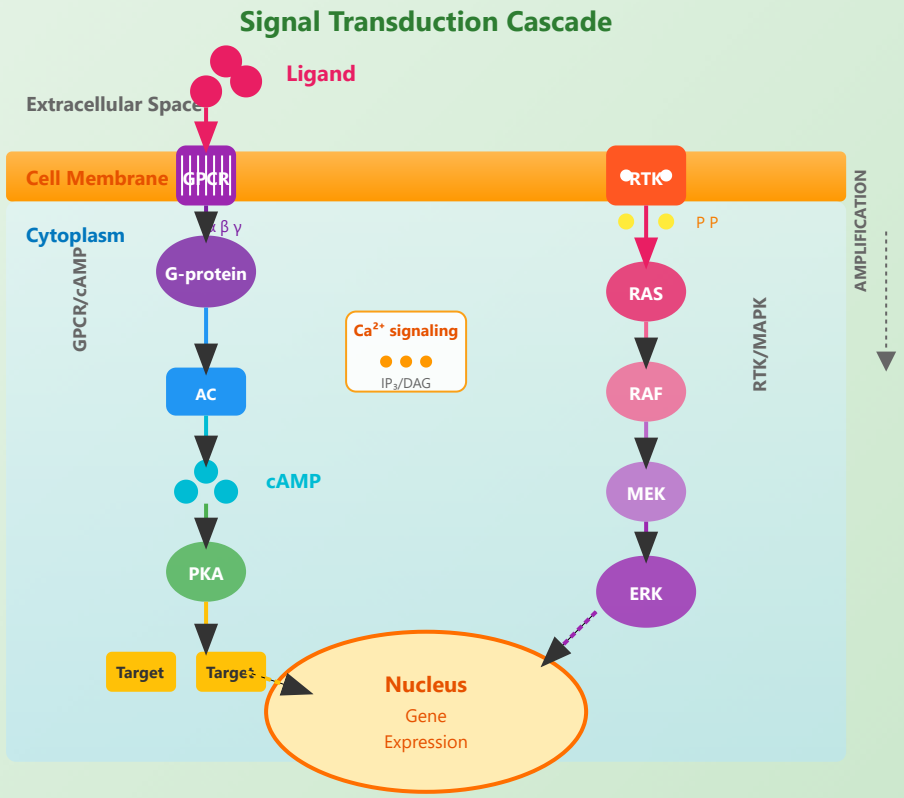
Mitochondria

- ATP production (powerhouse)
- Double membrane
- Own DNA and ribosomes
- Apoptosis regulation

Golgi Apparatus

- Protein modification
- Glycosylation
- Protein sorting and packaging
- Vesicle formation

Cell Signaling Pathways



Receptor Types

- GPCR: G-protein coupled
- RTK: receptor tyrosine kinase
- Ion channel receptors
- Nuclear receptors

Second Messengers

- cAMP: activates PKA
- Ca^{2+}: multiple targets
- IP_3 and DAG
- Amplify signal

Kinase Cascades

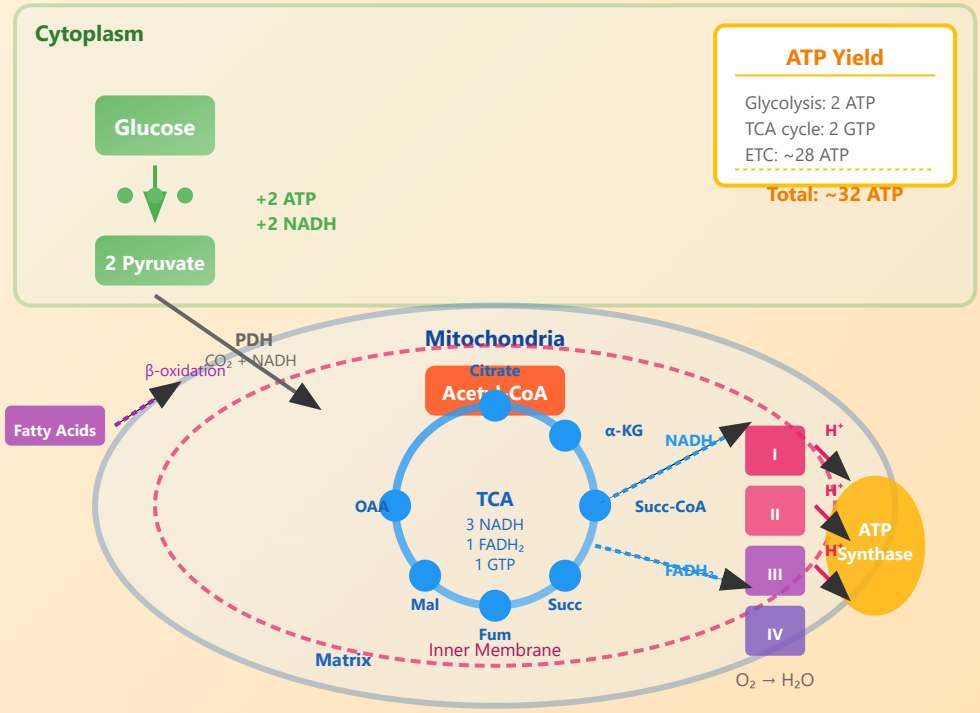
- MAPK pathway
- Sequential phosphorylation
- Signal amplification
- Specificity and crosstalk

Feedback Regulation

- Negative feedback: stability

- Positive feedback: switches
- Desensitization
- Temporal dynamics

Metabolic Pathways Overview



Glycolysis

- Glucose → 2 Pyruvate
- Net: 2 ATP, 2 NADH
- Cytoplasmic pathway
- Aerobic and anaerobic

TCA Cycle

- Acetyl-CoA oxidation
- Produces NADH, $FADH_2$
- Mitochondrial matrix
- Central metabolic hub

Oxidative Phosphorylation

- Electron transport chain
- Proton gradient formation
- ATP synthase
- ~30-32 ATP per glucose

Pathway Integration

- Metabolic flux control
- Allosteric regulation
- Hormonal control

- Compartmentalization

ATP and Energy Transfer

ATP Structure

- Adenosine + 3 phosphates
- High-energy phosphate bonds
- Hydrolysis: ATP → ADP + Pi
- $\Delta G^{\circ\prime} = -7.3$ kcal/mol

Energy Coupling

- Links exergonic to endergonic
- Common intermediate strategy
- Enzyme catalyzed
- Metabolic efficiency

Other Energy Carriers

- GTP: protein synthesis
- NADH: reduction reactions
- $FADH_2$: electron transport
- Creatine phosphate: muscle

Cellular Energy Budget

- Daily ATP turnover: ~body weight
- Majority for biosynthesis
- Transport and signaling
- Mechanical work

🔄 Cell Cycle and Division

G1 Phase (Gap 1)

Cell growth and normal metabolism. Decision point for division at G1/S checkpoint.

S Phase (Synthesis)

DNA replication occurs. Chromosomes are duplicated. Histone synthesis.

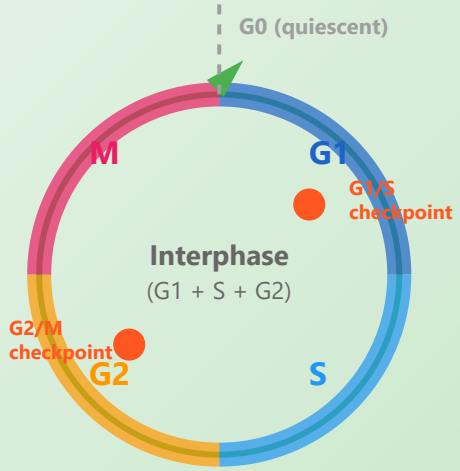
G2 Phase (Gap 2)

Preparation for mitosis. Protein synthesis and organelle duplication. G2/M checkpoint.

M Phase (Mitosis)

Nuclear division: Prophase → Metaphase → Anaphase → Telophase → Cytokinesis

Cell Cycle



Typical Duration:
G1: 5-6h | S: 10-12h
G2: 3-4h | M: 1-2h

Apoptosis and Cell Death

Intrinsic Pathway

- Mitochondrial pathway
- Cytochrome c release
- Apoptosome formation
- Triggered by DNA damage, stress

Extrinsic Pathway

- Death receptor activation
- FAS, TNF receptors
- DISC complex formation
- Immune-mediated

Caspase Cascade

- Initiator caspases (8, 9)
- Executioner caspases (3, 7)
- Proteolytic cleavage
- Irreversible commitment

Regulation

- Bcl-2 family: pro and anti-apoptotic
- IAPs: caspase inhibitors
- p53: apoptosis inducer
- Cancer dysregulation

Stem Cells and Differentiation

Stem Cell Types

- Totipotent: can form organism
- Pluripotent: all cell types
- Multipotent: limited lineages
- Unipotent: single cell type

Differentiation Signals

- Growth factors
- Cell-cell interactions
- Extracellular matrix
- Mechanical cues

Epigenetic Changes

- Progressive restriction
- DNA methylation patterns
- Chromatin remodeling
- Transcription factor networks

Regenerative Medicine

- iPSCs: induced pluripotent
- Tissue engineering
- Disease modeling
- Drug screening

Hands-on: PyMOL Molecular Visualization

Getting Started

- Install PyMOL (open source available)
- Load PDB files: fetch 1AKE
- Basic navigation: mouse controls
- Command line interface

Visualization Options

- Cartoon: secondary structure
- Sticks: detailed bonds
- Surface: molecular surface
- Ribbon: protein backbone

Analysis Tools

- Distance measurements
- Hydrogen bond identification
- Surface area calculations
- Electrostatic potentials

Creating Figures

- Ray tracing for publication
- Color schemes
- Label atoms/residues
- Export high-resolution images

Hands-on: PDB Database Exploration

Search Strategies

- Keyword search: protein name
- Advanced search: filters
- Sequence similarity
- Structure similarity

Structure Quality

- Resolution: <2Å is high quality
- R-factor: fit to data
- Ramachandran plot
- Missing residues

Functional Analysis

- Active site identification
- Ligand binding
- Protein-protein interfaces
- Conformational changes

Integration with AlphaFold

- Predicted structures available
- Confidence scores (pLDDT)
- Complement experimental data
- AlphaFold database

Thank You!

Continue exploring the molecular basis of life

Next Lecture: Advanced Bioinformatics Tools

Assignment: PDB Structure Analysis

Office Hours: By appointment