CHEMISTRY 456 / INTERNATIONAL WORKSHOP

Fundamentals of Biological Macromolecules

Andy Howard
Associate Professor of Biology & Physics,
Illinois Tech

Monday 28 June 2021

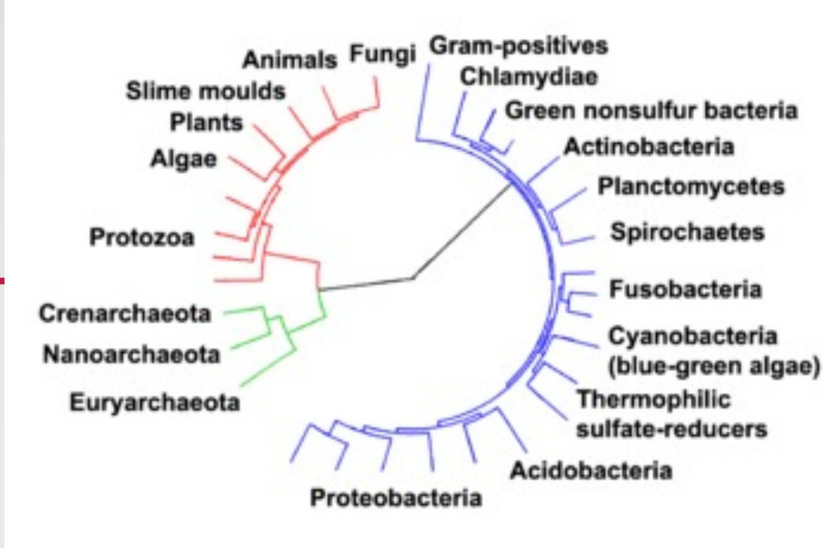
WHY BIOCHEMISTRY MATTERS

- Biochemists point out how chemical processes occur in biological systems, inside and between cells.
- These remarks are intended to introduce you to the chemistry of biopolymers and their biological significance.

WHAT WE'LL DISCUSS

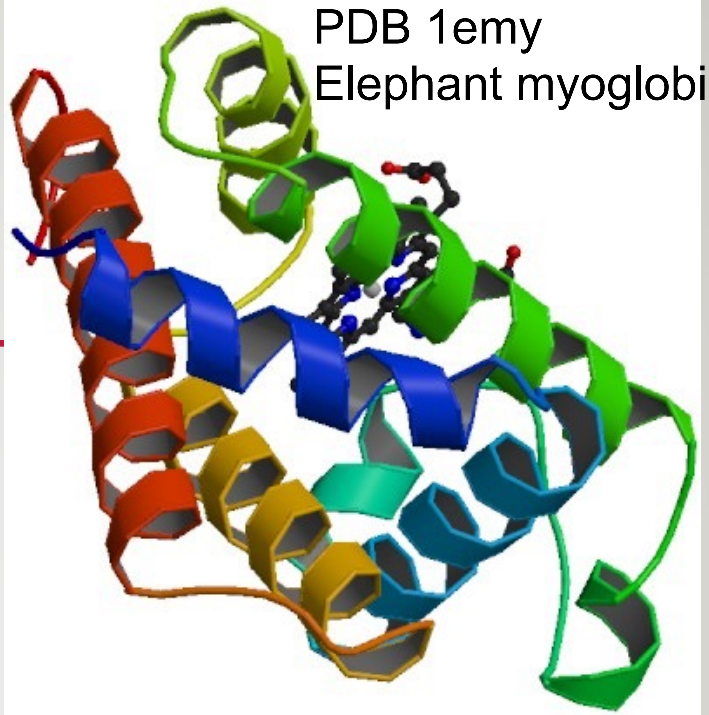
- Unity and diversity
- Thermodynamics
- Organic chemistry
- Small biomolecules
- Polymers
- Proteins
  - Components
  - Structure
  - Functions
- Enzymes
- Nucleic acids
  - Components
  - Polymers
  - Structure
  - Function
- Carbohydrates

BIOCHEMICAL UNITY



- There is a remarkable unity to biochemical processes all across eubacteria, archaea, and eukaryota
- The genetic code is almost uniform throughout all 3 super-kingdoms
- Specific proteins look the same across wide gaps of evolutionary history

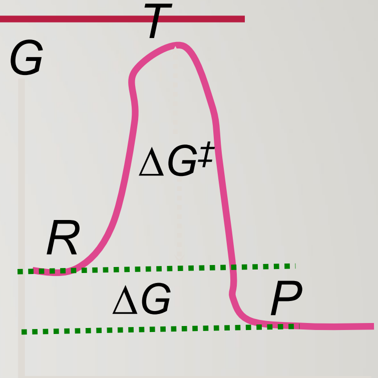
BIOCHEMICAL DIVERSITY



- The morphological and functional differences between you and an elephant, or between you and a flowering plant, or between you and a bacterium, arise from differences in biochemical phenomena
  - Some classes of proteins are present only in certain kinds of organisms
  - Structural components in vertebrates differ significantly from those in arthropods

THERMODYNAMICS AND KINETICS

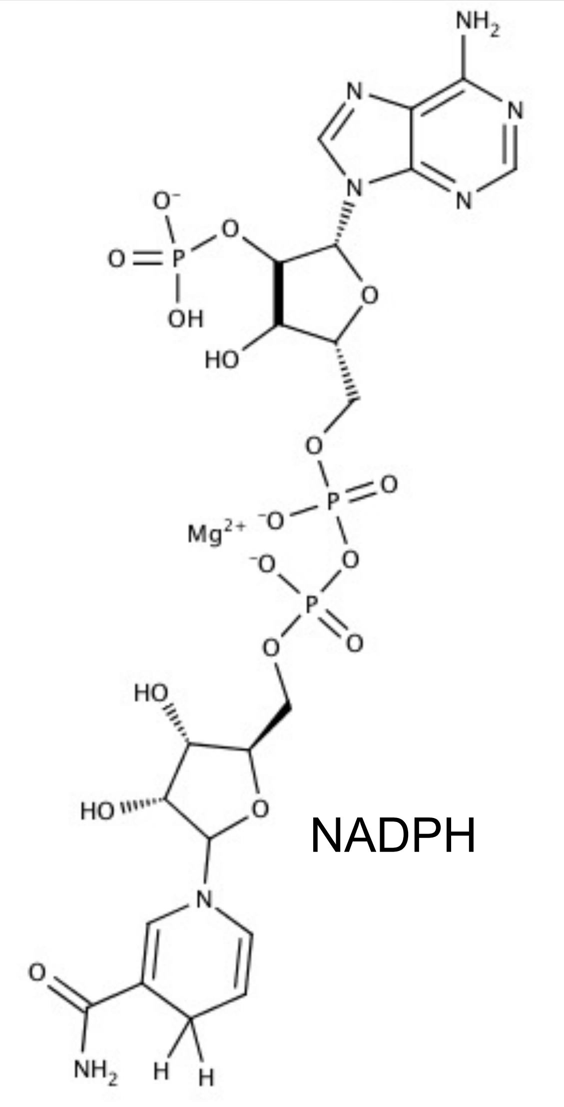
- Systems of molecules tend toward an equilibrium state in which the minimum free energy is obtained
- In order to go from an initial state *R* to a final, lower-energy state *P*, a higher-energy transition state *T* may have to be overcome
- Free energy has both enthalpic (heat) and entropic (organizational) components (Gibbs & Boltzmann): $\Delta G = \Delta H - T\Delta S$



Reaction Coordinate

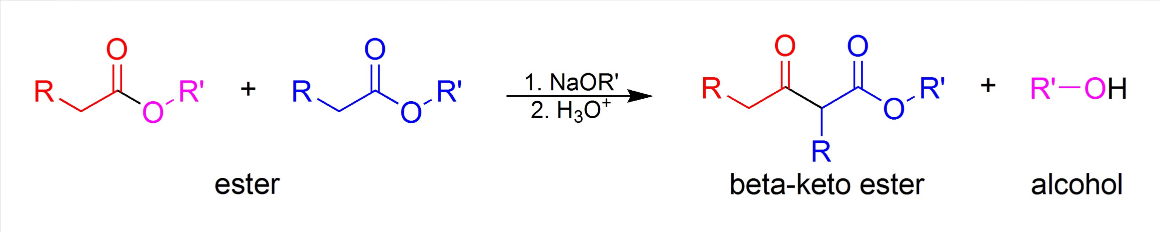
BIOORGANIC CHEMISTRY

- Most biological reactions involve carbon-containing compounds
- Therefore organic chemistry (the chemistry of compounds that contain C-C or C-H bonds) explains a large fraction of what's happening below the surface in biochemical systems



WE CAN IGNORE A LOT OF ORGANIC REACTION MECHANISMS!

- Only a subset of all possible organic reactions occur in biochemistry:
  - They're almost all aqueous
  - Generally occur at temperatures between 3°C and 50°C
  - Most occur between pH 5 and pH 9 (exception in humans: the stomach, which is highly acidic)
  - Entire classes of organic reactions are unrepresented

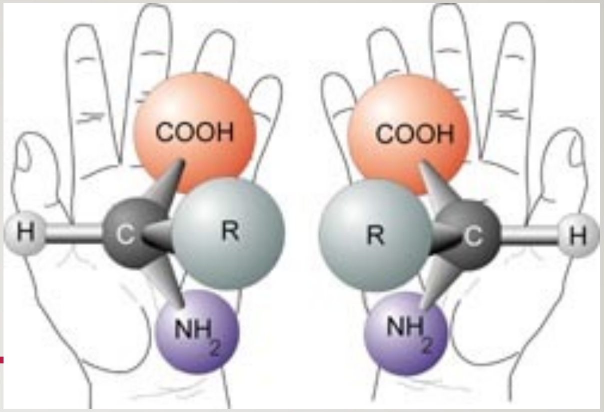


Claisen condensation

WHAT REACTIONS *DO* OCCUR?

- Mostly nucleophilic substitutions (SN1 and SN2)
  - These are two-electron transformations, as we've discussed in previous lectures
  - These take place under enzymatic control
  - Therefore they can proceed at measurable rates even though the nucleophiles aren't very powerful by organic lab standards
- Some free radical reactions, as we've said
  - There are enzymatic free-radical reactions
  - But there are non-enzymatic ones too

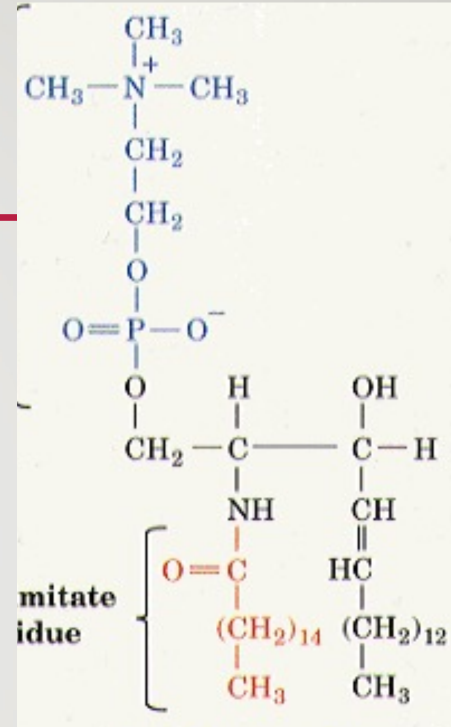
CHIRALITY



- Many biomolecules are chiral:
  - At least one carbon atom has 4 distinct substituents
  - Therefore the mirror image of the molecule is different from the original molecule
- Specific classes of biomolecules are chiral
- Standard amino acids are L- and so are their polymers
- Most sugars are D- and so are their polymers
- Lipids are usually not chiral
- Nucleic acid bases: achiral; but ribose is chiral

SMALL BIOMOLECULES

- Most common small molecule in biological systems is water
- There are other important small molecules and ions:
  - CO_2, O_2, glyceraldehyde, glucose
  - HCO_3^-, PO_4^{3-}, Na^+, Cl^-, K^+, Ca^{2+}, Fe^{2+}, Mg^{2+}
  - Palmitate, oleate, stearate, choline, sphingomyelin
  - Amino acids (20 ribosomal + others)
  - Nucleic acids (A,C,G,U,dA,dC,dG,dT, …)



A sphingomyelin

BIOPOLYMERS

- Most complex activities in biology involve polymers
- Biology achieves catalysis, reproduction, specificity, and many other things by making polymers, not by building complex monomeric molecules
- Biological polymers are built up by formal (and actual) elimination of water ($M_1 + M_2 \rightarrow M_{1\text{-}2} + H_2O$)
    - Protein: polymer of amino acids
    - RNA: polymer of ribonucleotides
    - DNA: polymer of deoxyribonucleotides
    - Polysaccharides: polymers of sugars
- Except for polysaccharides, they're unbranched

MAJOR CLASSES OF BIOPOLYMERS

Category	Types of monomers	<mol wt/ monomer >	# mono-mers per polymer	Branched?
Protein	20	110	65-5000	no
RNA	4-10	150-300	10-15K	no
DNA	4	130-270	$50\text{-}10^6$	no
Polysac-charide	~10	162	$2\text{-}10^5$	Some-times

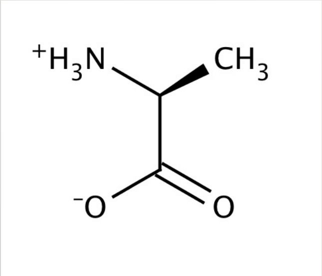
PROTEINS

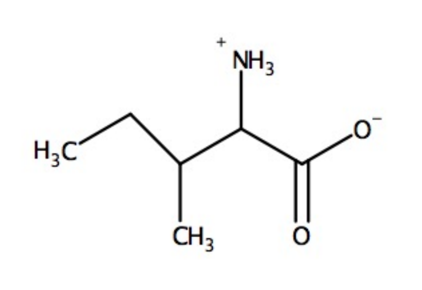
- Proteins are polymers of L-amino acids
- There are 20 amino acids in proteins from 99% of all organisms; a few use 2 others

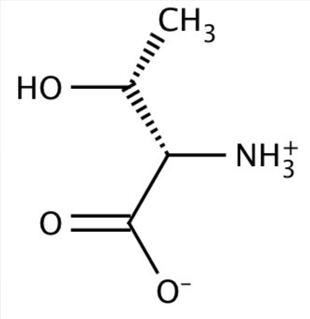
- Variety of functions:
  - Enzymes (catalysts)
  - Structural proteins
  - Storage and transport proteins
  - Hormones
  - Receptors
  - Nucleic-acid binding proteins
  - Molecular machines
  - Specialized functions (e.g. sweetness)

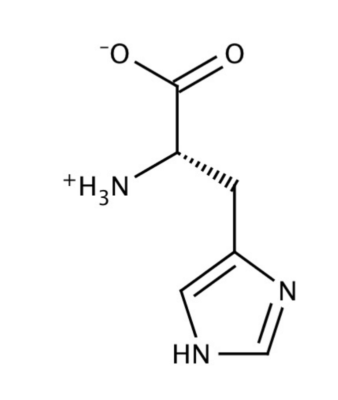
ALPHA-AMINO ACIDS

- These are the building blocks from which proteins are made
- Alpha carbon covalently bonded to a carboxylate and an amine
- Polymerization makes peptide bonds between the α-amino acids









PROTEIN STRUCTURE

Human DHFR
PDB 1KMV, 1.05Å

- Most (>80%) proteins fold into a specific 3-D conformation at room temperature
- This isn't intuitively obvious: many polymers are much floppier than that
- A few proteins are "intrinsically unstructured"
- Held together by ionic interactions (H-bonds, salt bridges) and nonpolar interactions (van der Waals and solvent exclusion)

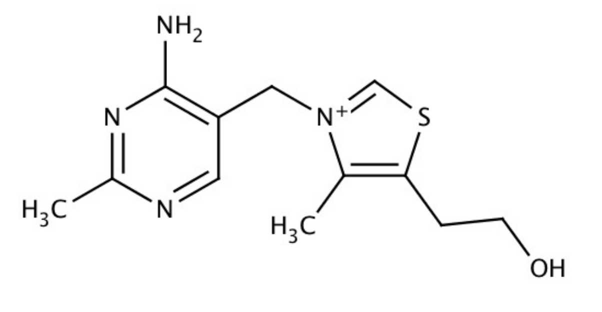
LEVELS OF PROTEIN STRUCTURE

- Primary structure: linear sequence of amino acids, e.g. val-gly-ser-leu-asn-cys-ile-val-ala-val-ser-gln-asn-met-gly …
- Secondary structure: short-range ordering provided by main-chain hydrogen bonds between amine nitrogens and carbonyl oxygens
- Tertiary structure: three-dimensional arrangements of all atoms in a single polypeptide
- Quaternary structure: arrangement of individual polypeptides into a complete multimeric structure

HOW DO PROTEINS PERFORM SO MANY FUNCTIONS?

- Amino acid side chains have a wide variety of reactivities
- Scaffolding surrounding active or binding site can significantly change reactivity of specific moieties within the protein
- Side chains that would normally be hard to ionize become easy to ionize because the ion is stabilized by interactions with neighboring groups
- Regions that would be hydrophilic if exposed become hydrophobic when buried within protein

IS THAT ALL?



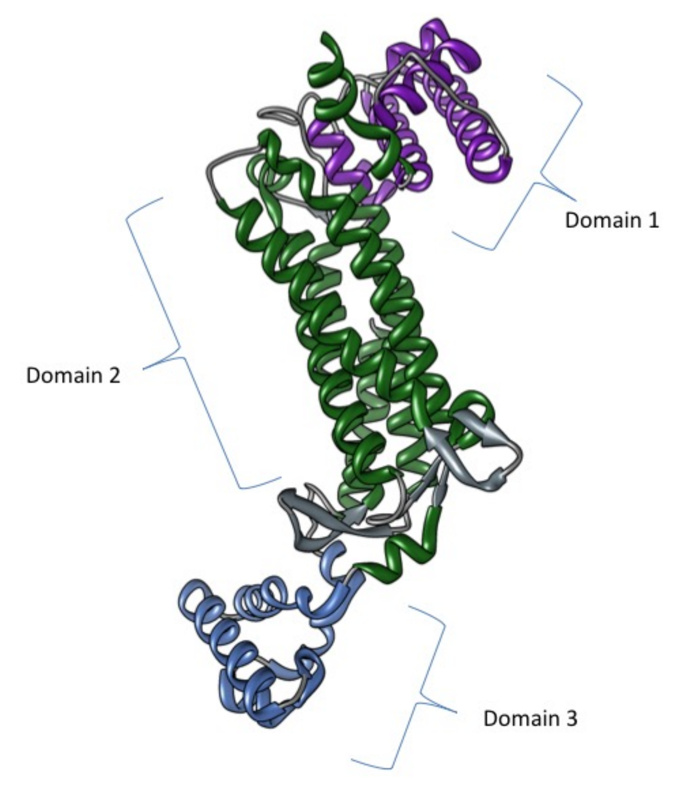
- No; Some proteins have non-amino-acid components bound, loosely or tightly, to the amino acid chain(s); these entities are *cofactors*
- Cofactors can be inorganic (e.g. Mg^{2+} ions) or organic (coenzymes)
- These provide functionalities that would otherwise be unavailable to the polypeptide
- Many coenzymes are derived from vitamins, particularly the B-complex vitamins

WHY ARE PROTEINS SO BIG?

- Typical enzyme active site or carrier-protein binding site involves < 5% of the amino acids in the protein
- So what are all those other amino acids doing?
  - Creating environment perfectly suited to the electrostatics and hydrophobicity or hydrophilicity required for proper function
  - Guarantees that the active site or binding site *keeps* its shape and electrostatic properties over time
  - Enables binding of cofactors and other effectors
  - Allows for appropriate interactions with other macromolecules

DOES EVERY PROTEIN HAVE EXACTLY ONE FUNCTION?

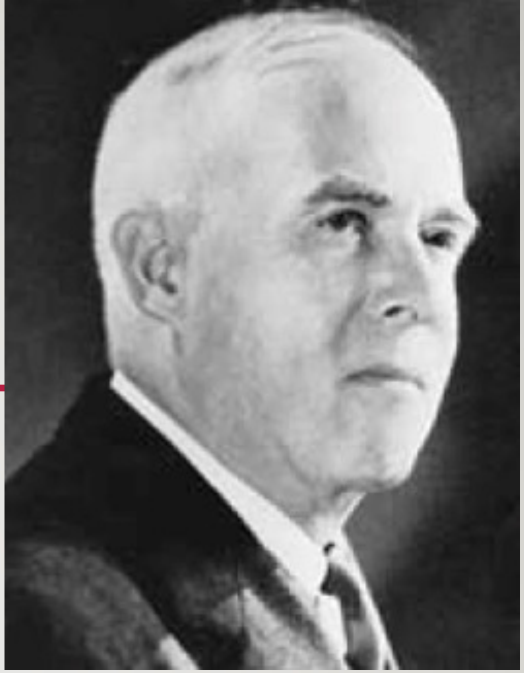
- No.
- Some proteins that have evolved for one purpose become pressed into service in a different guise
- Example: some of the proteins that make up the eye lens are actually enzymes, reused as structural proteins



Adapted from: Sampaleanu, L. M; Vallee, F.; Thompson, G. D.; Howell, P. L. Biochemistry. 2001, 40, 15570-15580.

ENZYMES

- Enzymes are defined as biological catalysts
- Biochemists figured out in the 1920's that all enzymes were proteins, although it took until the 1930's for that notion to be definitively accepted
- In the 1970's this assertion had to be modified: some enzymes (biocatalysts) are actually RNA molecules
- 1990's: Recognition that the catalytic step in protein synthesis is performed by a specific adenine base in ribosomal RNA



James Sumner

WHAT PROPERTIES MUST AN ENZYME HAVE?

- **Catalysis:**
  must lower the activation energy of a reaction
- **Specificity:**
  must act preferentially on certain substrates
- **Regulation:**
  must be subject to some sort of regulation

HOW IS CATALYSIS ACCOMPLISHED?

- Multiple mechanisms, often acting in concert:
  - Acid-base catalysis (specific amino acid side-chains participate temporarily in reactions by acting as nucleophiles or electrophiles)
  - Proximity effect (enzyme contains channels with appropriate shape and electrostatic properties to cause substrates to travel down them toward one another)
  - Induced fit (enzyme reshapes itself to bind substrate, but substrate gets reshaped as well)
  - Transition state stabilization

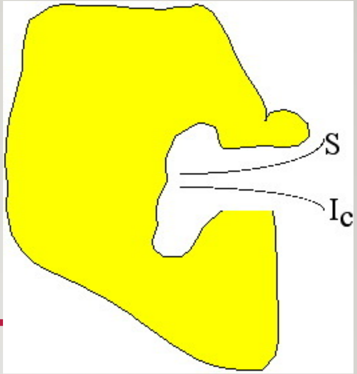
HOW EFFECTIVE ARE ENZYMES, AND HOW ARE THEY REGULATED?

- The best can speed up a reaction by a factor of 10^{12} or more
- They don't change equilibrium: they just make the reaction go faster.
- Regulation by:
  - Control of transcription
  - Control of translation
  - Interference with activity via inhibitors

INHERENT PROPERTIES OF ENZYMES

- K_m is an inherent property of an enzyme
  - Can be measured for a specific substrate, independent of enzyme and substrate concentration
  - Measures binding affinity of enzyme for substrate
- V_{max} clearly depends on how much enzyme we put in: the more we put in, the faster the reaction
- But $k_{cat} = V_{max}/[E]_{tot}$ is an inherent property
  - Measures turnover rate
  - k_{cat} = number of substrate molecules converted to product per unit time by a single enzyme molecule

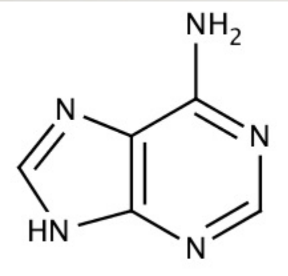
ENZYME INHIBITION

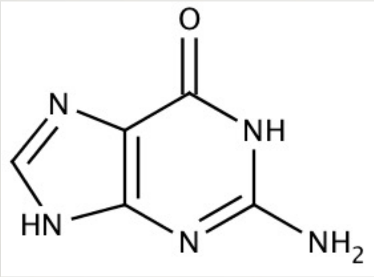


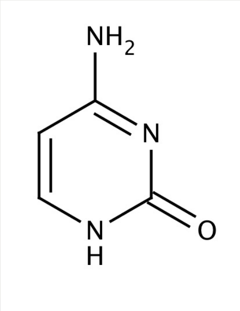
- Various small molecules can bind to an enzyme and slow it down
- Competitive inhibitors compete with the substrate for the active site and raise K_m
- Noncompetitive inhibitors interfere with catalysis and decrease V_{max}
- Uncompetitive inhibitors decrease both V_{max} and K_m, but do so in a way that slows the reaction

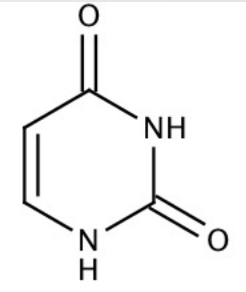
NUCLEIC ACID BUILDING-BLOCKS

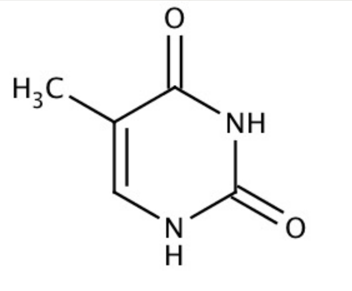
- All nucleotides are based on *nucleobases*, which are either purines (9 ring-atoms, double-ring structures) or pyrimidines (6 ring-atoms)
- The nucleobase is attached to a ribose or deoxyribose sugar by an N-C bond
- A phosphate is attached to the 5'-position of the sugar





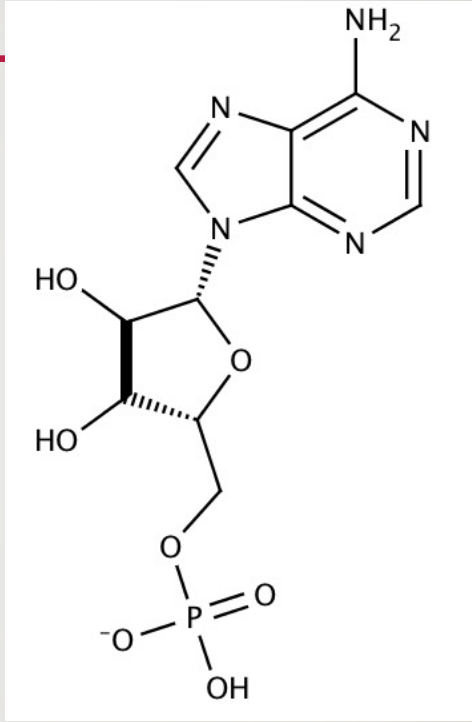






NUCLEOSIDES AND NUCLEOTIDES

- Attach a nitrogen on a nucleobase to carbon 1 of ribose to make a nucleoside
- Attach a phosphate to the 5' position on the nucleoside to make a nucleotide

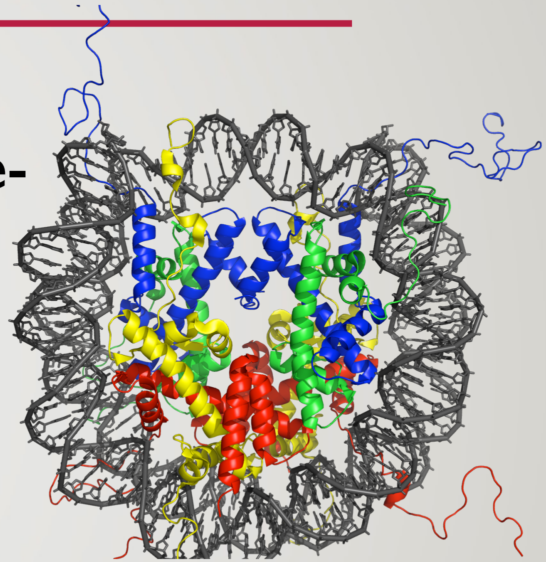


POLYNUCLEOTIDES

- Connect the 5’ phosphate of one nucleotide to the 3’ position on the previous nucleotide and continue
- A DNA molecule can comprise millions of nucleotides
- RNA is somewhat smaller but still substantial

NUCLEIC ACID STRUCTURE

- DNA is generally double-helical
- Many RNA molecules are partially double-stranded too
- DNA is unstructured above the double-helix level during G1 but becomes structured during S within spool-like structures called nucleosomes
- Higher levels of organizations occur too

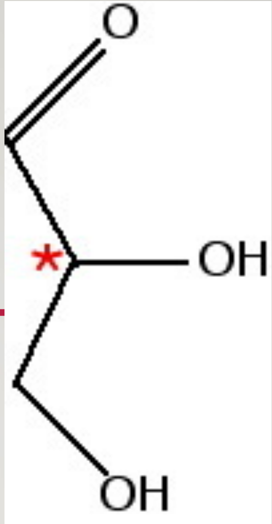


TYPES OF RNA

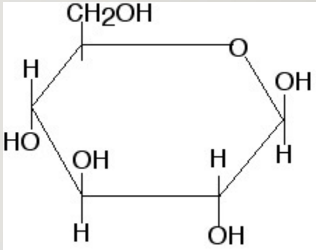
- We generally recognize four major classes of RNA, differing in size, cellular role, and replacement rate

Type	*% RNA*	*%turn-over*	*bases*	*Size, DS?*	*Partly Role*
mRNA	3	25	50-10^4	no	protein template
tRNA	15	21	55-90	yes	aa activation
rRNA	80	50	10^2-10^4	yes	transl. catalysis & scaffolding
sRNA	2	4	15-10^3	yes	various

CARBOHYDRATES



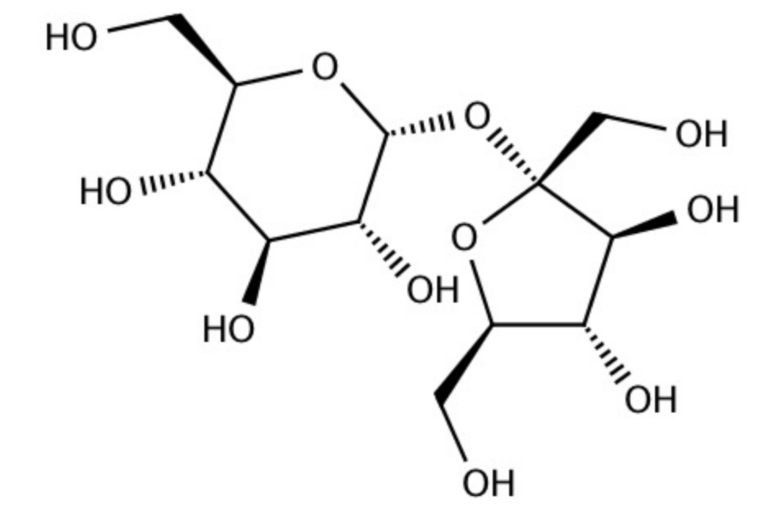
- Monomers are three-carbon to seven-carbon polyalcohols with one carbonyl group at either C1 or C2. They're very soluble.
- Simplest: glyceraldehyde & dihydroxyacetone - C_3
- Most important other ones:
  - Ribose, xylose (C5)
  - Glucose, fructose, galactose (C6)
- Monomers are fuel, signals, intermediaries (especially when phosphorylated)



SUGAR DERIVATIVES

- Sugar monomers that have been modified
  - Deoxyribose
  - Phosphorylated sugars (especially at C1 and C5 or C6)
  - Aminated sugars (mostly C6 sugars)
  - N-acetyl sugars
  - Sugar acids
  - Sugar alcohols (carbonyl converted to alcohol)
- These derivatives are often the active agents in intermediary metabolism

DISACCHARIDES



sucrose

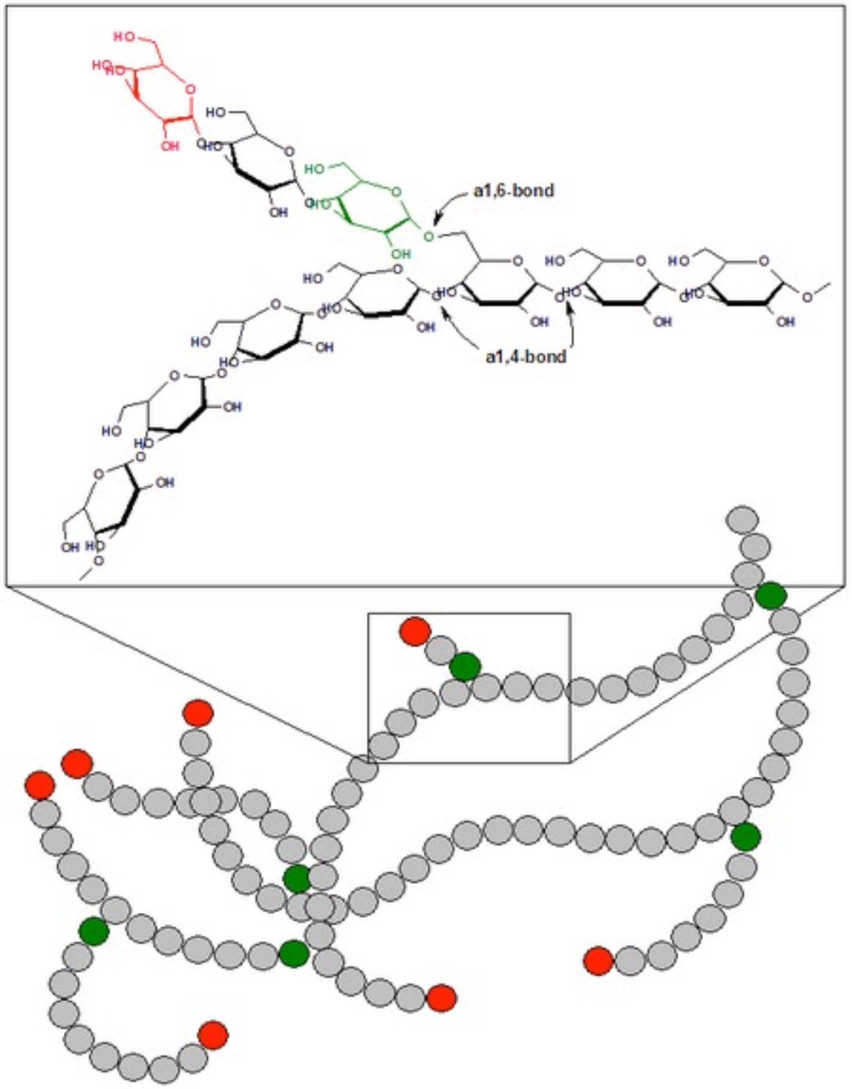
- Compounds made up of two sugar monomers
- Glucose + glucose → Maltose + H_2O
- Glucose + fructose → Sucrose + H_2O
- Glucose + galactose → Lactose + H_2O
- These are instances of *oligomers* (few rather than many building blocks)

STORAGE POLYSACCHARIDES

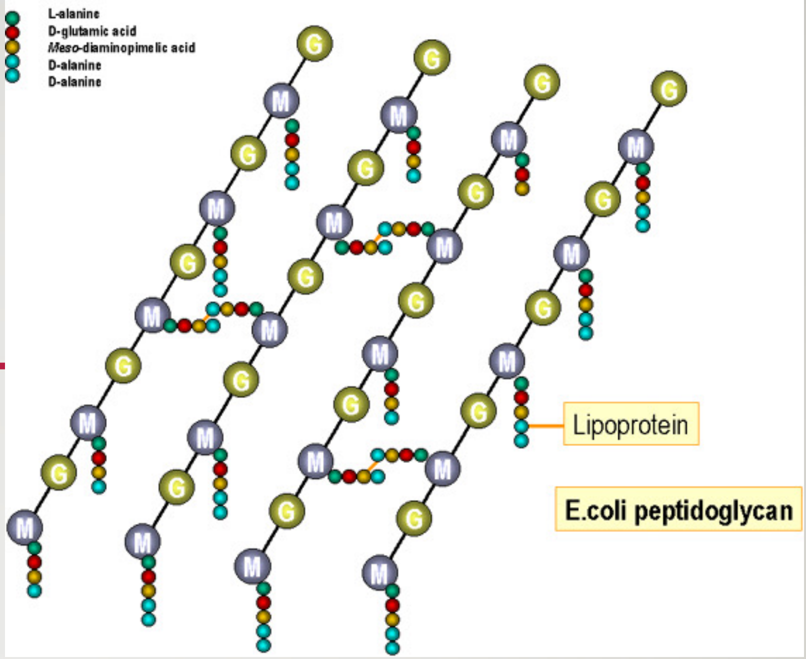
- Starch: storage form of glucose in plants
  - Amylose: 1-4 linkages only
  - Amylopectin: mostly 1-4 linkages, a few 1-6
- Glycogen: storage form of glucose in animals, some bacteria
  - mostly 1-4 linkages, some 1-6
  - Glycogen is primary short-term fuel in humans; long-term fuel is usually fats (more efficient)

GLYCOGEN STRUCTURE

- Single glucose units broken off & phosphorylated to use as fuel
- Amylases cleave the 1-4 linkages;
- Other enzymes deal with the 1-6 linkages



SUGAR-PEPTIDE COMPLEXES



- Sugars are often complexed either to peptides (oligomers of amino acids) or proteins (polymers of amino acids)
- Smaller-peptide complexes make up cell walls
- Full-size proteins that have been decorated with a few sugar molecules are ubiquitous in eukaryotes and often alter the properties of the protein or enable it to be involved in cell-cell communication