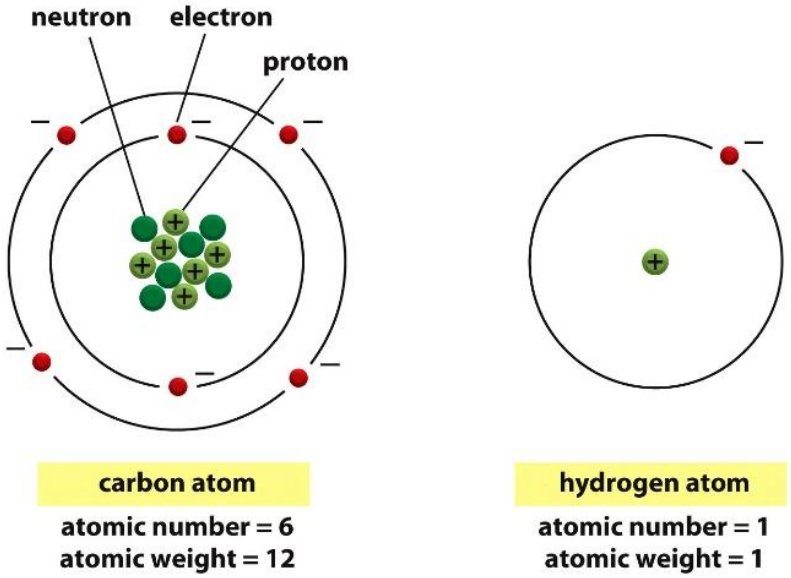
Lecture 1:
Chemistry and Biology Fundamentals

- Chemical Bonding
- Functional Groups
- Protein Structure and Stability
- Ligand Binding
- Proteins as enzymes
- Carbohydrates



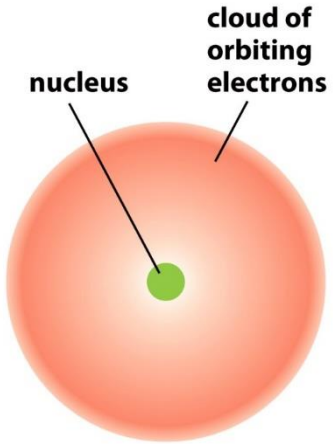
- Atoms are composed of:
  - Protons – positively charged particles
  - Neutrons – neutral particles
  - Electrons – negatively charged particles
- Protons and neutrons are located in the nucleus.
- Electrons are found in **orbitals** surrounding the nucleus.

Mass number
(number of protons + neutrons)

Atomic number
(number of protons)

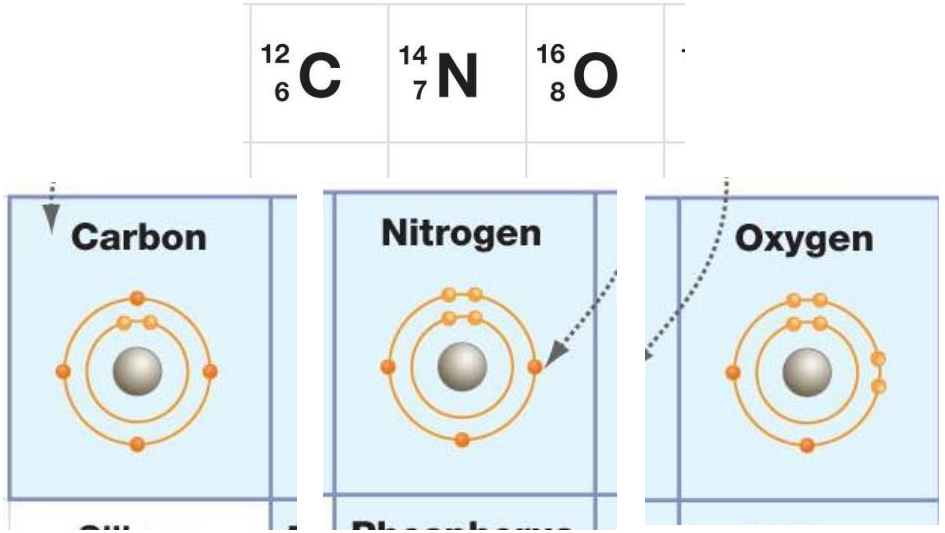
$^{1}_{1}H$							$^{4}_{2}He$
$^{7}_{3}Li$	$^{9}_{4}Be$	$^{11}_{5}B$	$^{12}_{6}C$	$^{14}_{7}N$	$^{16}_{8}O$	$^{19}_{9}F$	$^{20}_{10}Ne$
$^{23}_{11}Na$	$^{24}_{12}Mg$	$^{27}_{13}Al$	$^{28}_{14}Si$	$^{31}_{15}P$	$^{32}_{16}S$	$^{35}_{17}Cl$	$^{40}_{18}Ar$

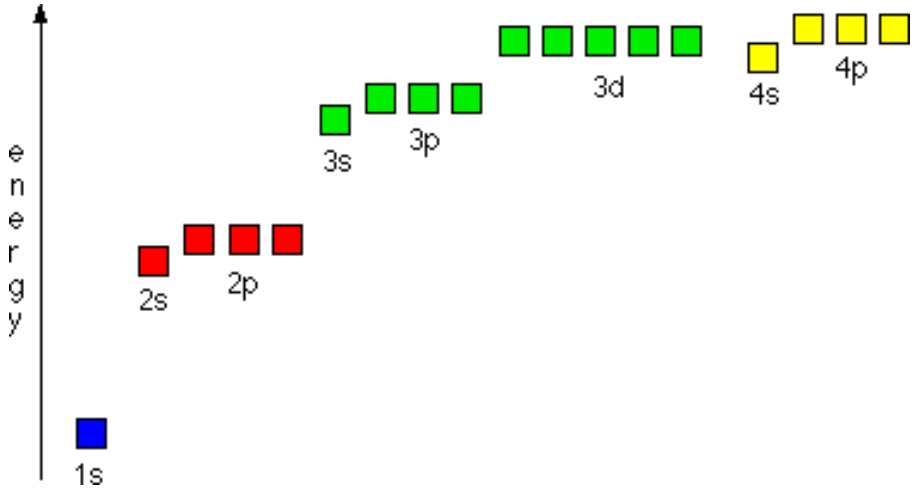
Atomic number = # of protons = # electrons in element



Electron Orbitals

- Electrons arranged around the nucleus in specific regions called orbitals.
  - Each orbital can only hold two electrons
- Orbitals are grouped into electron shells
  - Numbered 1,2,3…
  - Lower numbers = shells closer to the nucleus
  - First shell can hold a maximum of 2 electrons
  - Second shell can hold up to 8
  - Third shell can also hold 8
- Orbitals are usually filled from lowest energy (inner shell) to highest energy (outer shell)
- Outer shell is the ***valence shell*** and is used for forming bonds with other elements.
- The most stable configuration is a complete (full) outer shell.

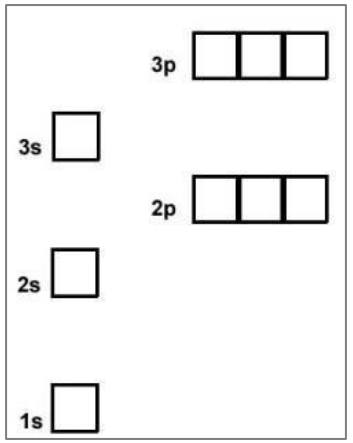




Shells: 1st = 1s, 2nd = 2s + 2p, 3rd = 3s + 3p

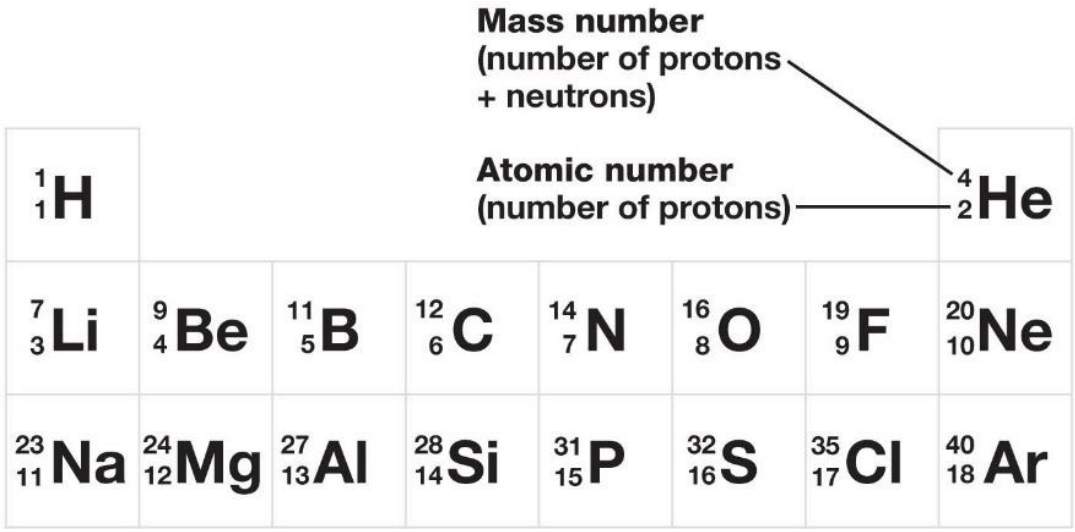
Shell is a collection of orbitals with similar energy

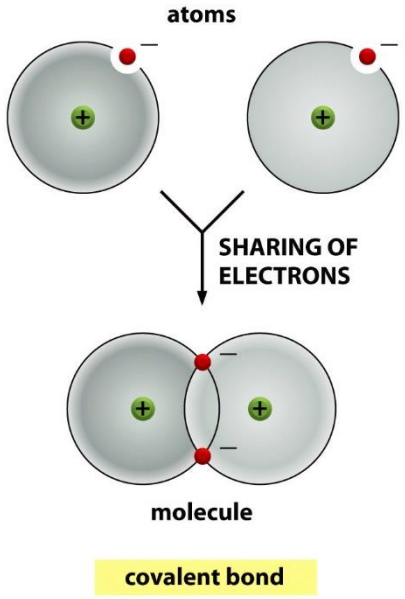
Electron Configuration of Ne – an inert gas (10e)



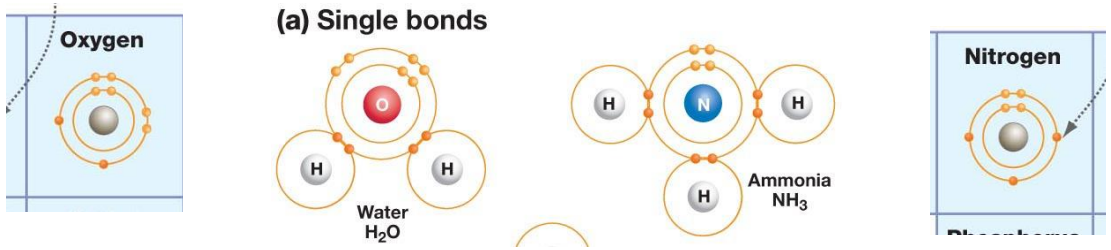
Covalent Chemical Bonds

- Elements like Li, Na, F, Cl, Mg, readily form ions to generate a complete outer shell.
- Some elements cannot form stable ions because it would involve the loss or gain of too many electrons. This includes C, N, and O – which are common in biological systems.
- Unfilled electron orbitals on elements like C, N, and O allow for the formation of ***covalent bonds***, and atoms are most stable when each electron orbital is filled.
  - Each atom's unpaired *valence* electrons are shared by both nuclei to fill their orbitals.
  - Substances held together by covalent bonds are called molecules

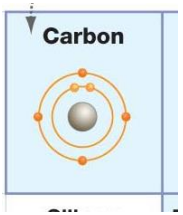




- The number of unpaired electrons (in the outer shell) determines the number of bonds an atom can make.
- Multiple bonds form when atoms share multiple electrons.



The number of covalent bonds (valence) formed by common elements.

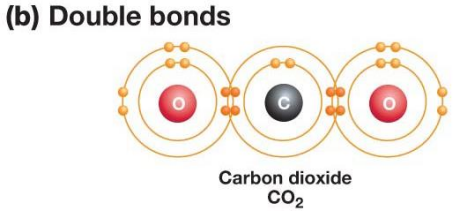


How many bonds will carbon form?

- Oxygen = 2 bonds
- Nitrogen = 3 bonds
- Carbon = _____
- Sulfur = 2 bonds (in biological systems)
- Hydrogen = 1 bond
- Phosphorous = 5 bonds in biological molecules

You must know these numbers.

O=C=O



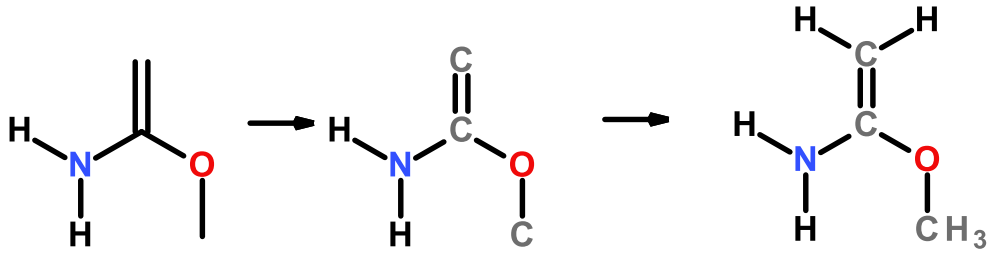
In chemical drawings:

- "C" for carbons are not drawn, but carbons are found at the ends of lines and when lines join or "kink"
- Hydrogens attached to carbon are not shown, you need to add them to complete to complete the valence of the carbon atoms.

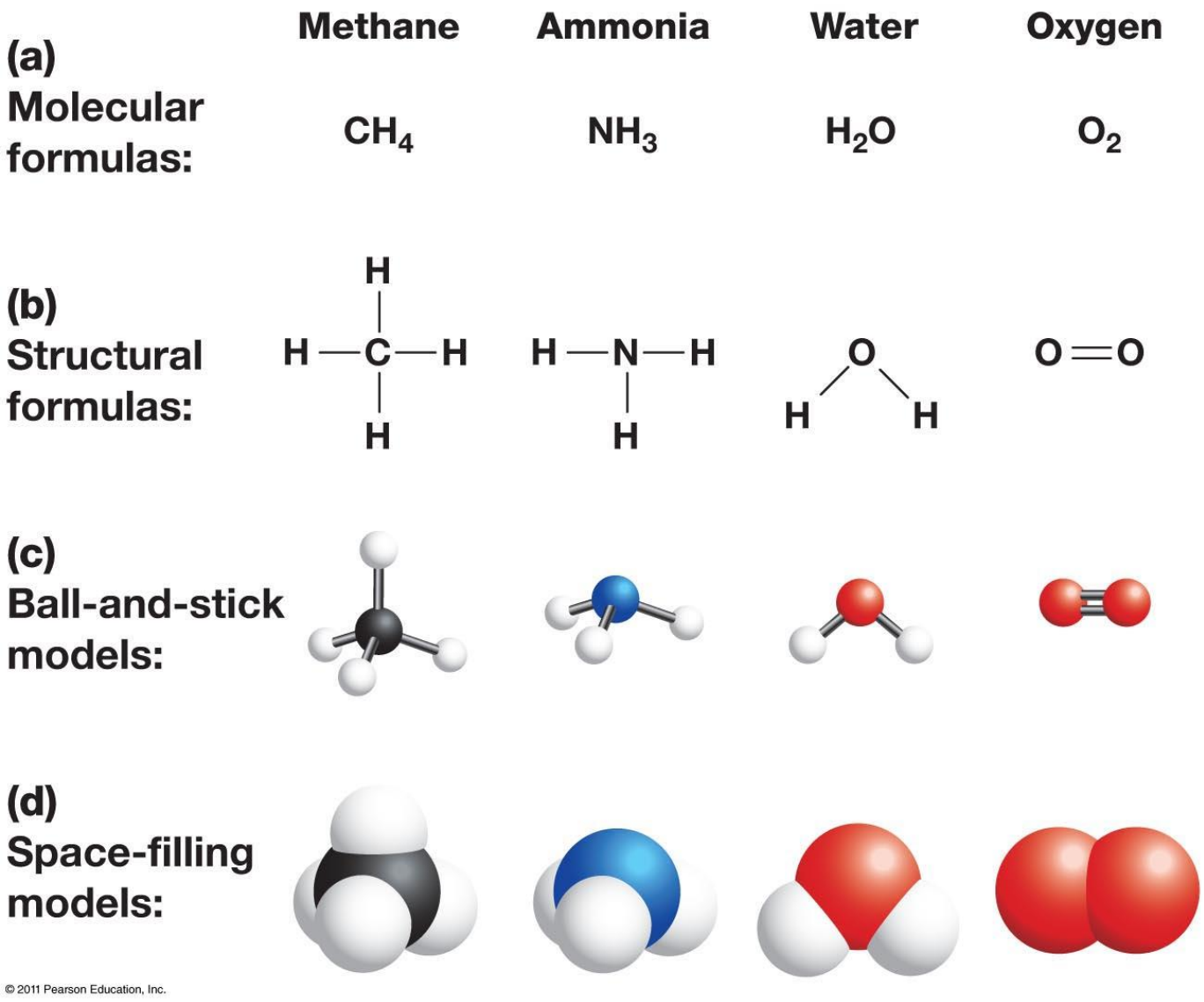
You must know how to do this.

"C" for carbons

Complete valence with H



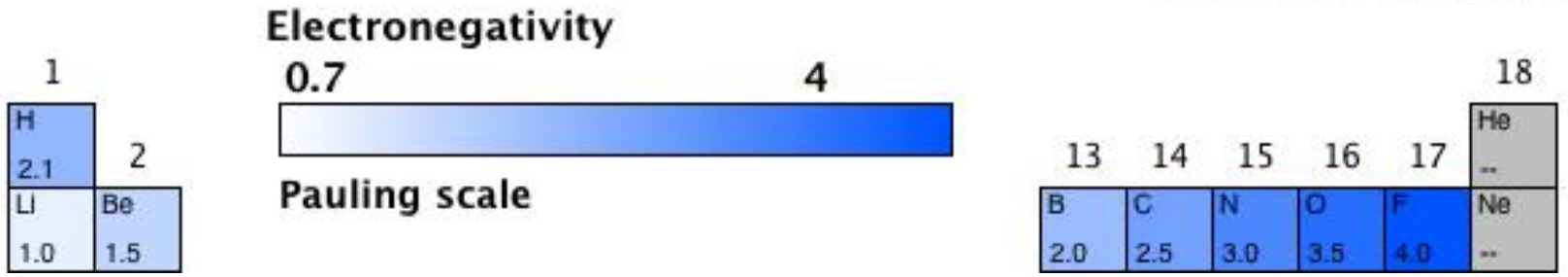
Representation of Molecules



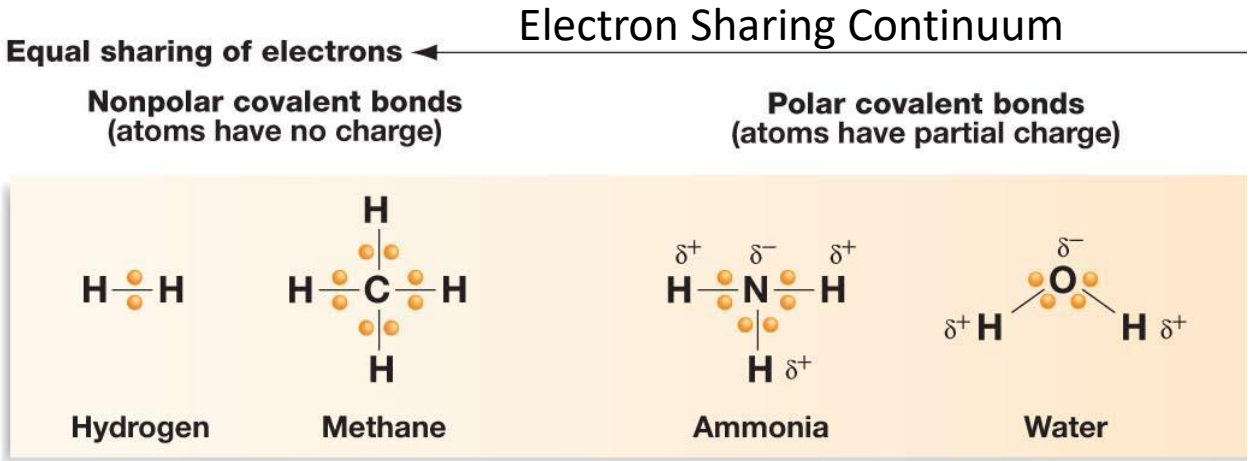
Electron Sharing and Bond Polarity

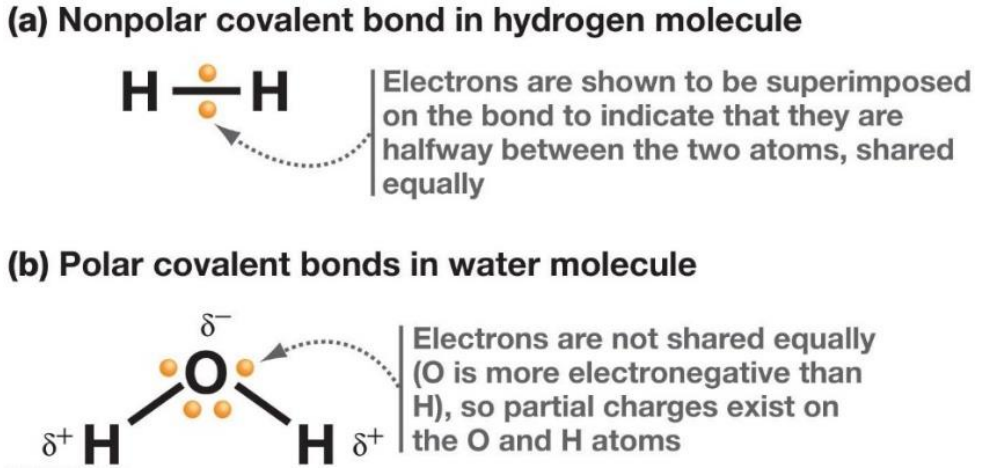
- The polarity of a bond depends on the electronegativity of the atoms.
- Electronegativity - ability of atoms to pull electrons from other atoms.
- Atoms with higher electronegativity will develop a partial negative charge, the atom they are bonded will have a partial positive charge.
- The order of electronegativity is:

  H ~ C < N < O



Increased pos. charge of nucleus.



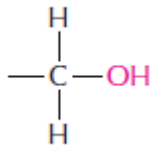


Functional Groups – You should Become Familiar with These

C–O COMPOUNDS

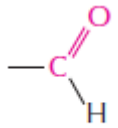
Many biological compounds contain a carbon bonded to an oxygen. For example,

alcohol

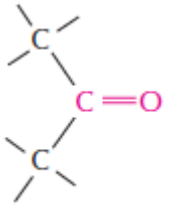


The –OH is called a hydroxyl group.

aldehyde

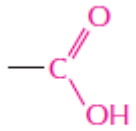


ketone



The C=O is called a carbonyl group.

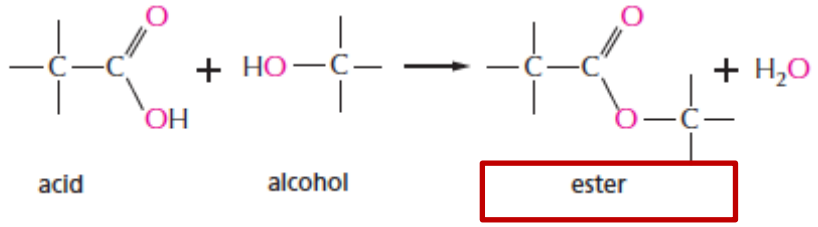
carboxylic acid



The –COOH is called a carboxyl group. In water this loses an H^+ ion to become $-COO^-$.

esters

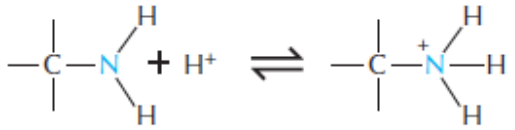
Esters are formed by combining an acid and an alcohol.



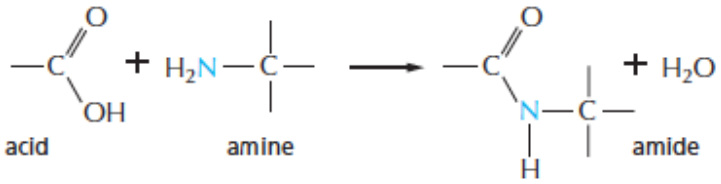
C–N COMPOUNDS

Amines and amides are two important examples of compounds containing a carbon linked to a nitrogen.

Amines in water combine with an H^+ ion to become positively charged.

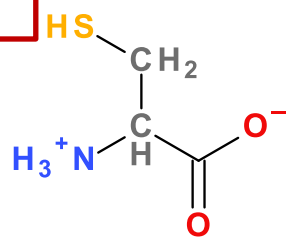


Amides are formed by combining an acid and an amine. Unlike amines, amides are uncharged in water. An example is the peptide bond that joins amino acids in a protein.



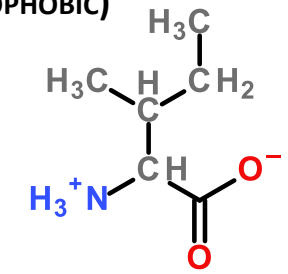
C–S COMPOUNDS

thiol



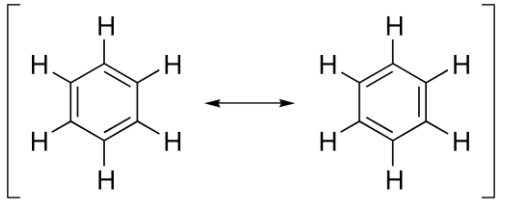
Cysteine (amino acid)

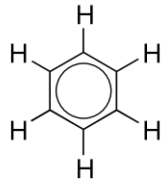
C–H GROUPS (HYDROPHOBIC)



Isoleucine (amino acid)

Aromatic – planer rings

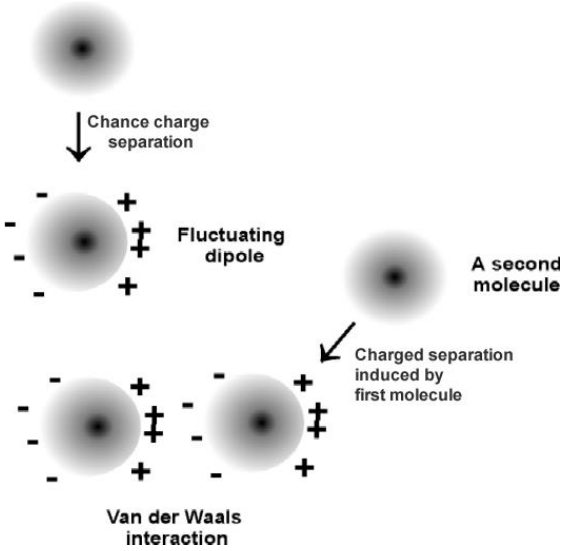


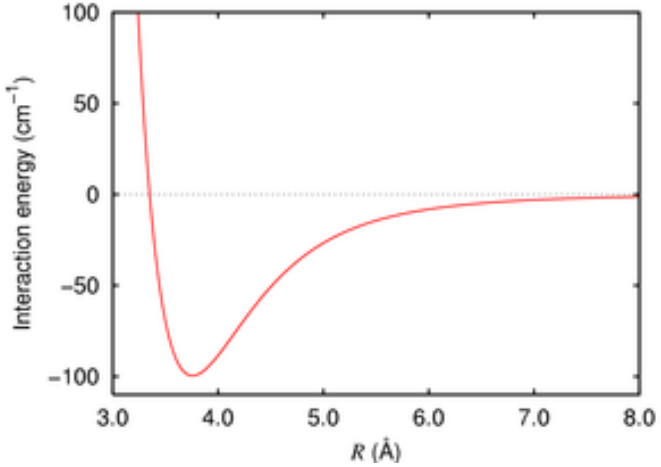


Inter-molecular Interactions in Biology

- **Electrostatic interactions** between fully charged molecules (ionic interactions)
  - Like charges repel
  - Opposite charges attract

- **van der Waals** interactions between:
  - Molecules with partial charges (polar)
  - Molecules with no charge (non-polar)





- **Hydrogen bond:**

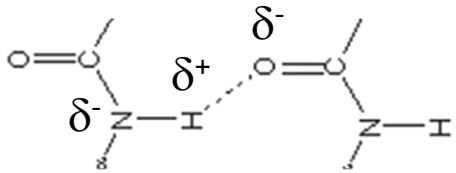
  X & Y are electronegative (N and O usually)

  X-H = *Donor* of the hydrogen bond, it provides the hydrogen (the hydrogen remains in a covalent bond to X)

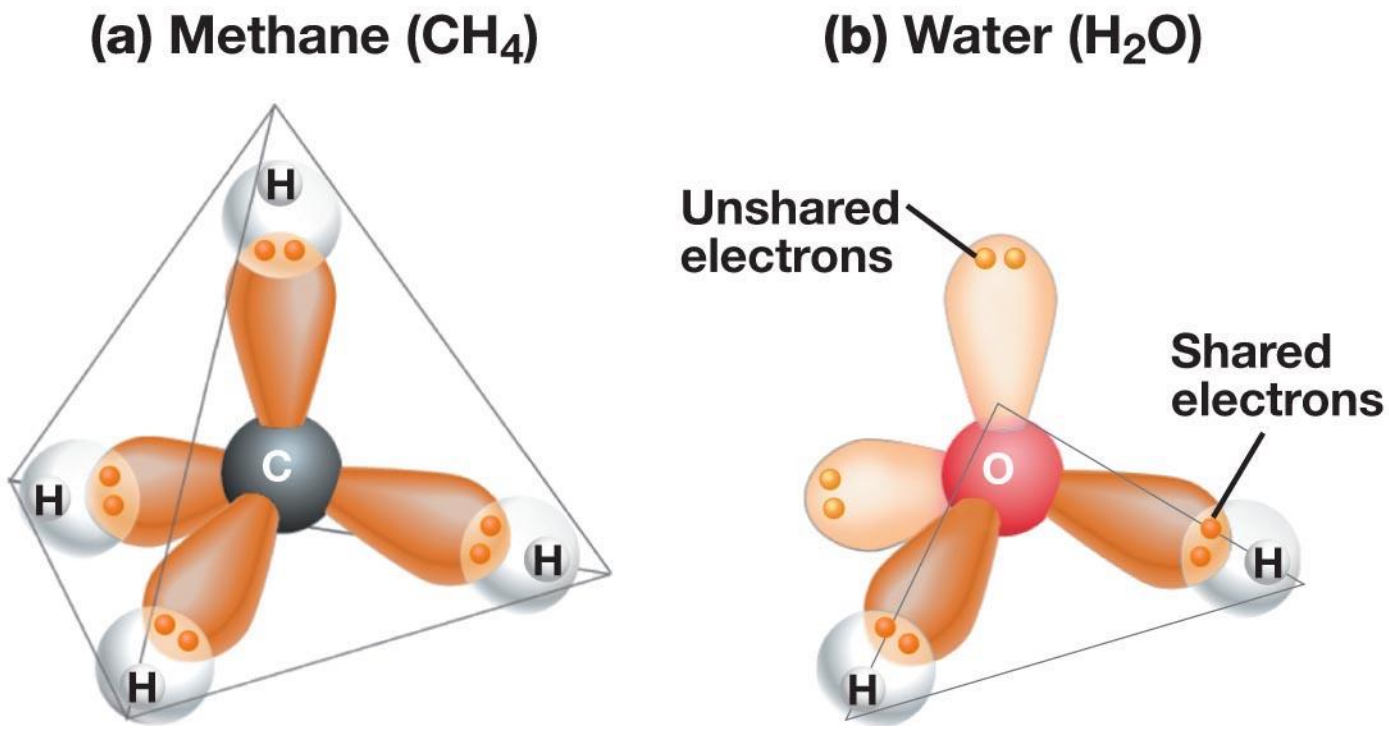
  Y = *Acceptor* of the hydrogen bond

X-H Y

Hydrogen bond between NH and O=C



The Geometry of Simple Molecules

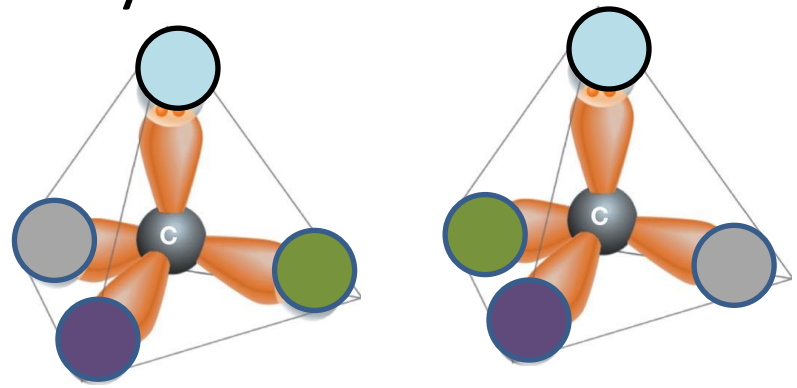


The shape of a molecule is determined by the geometry of its bonds.

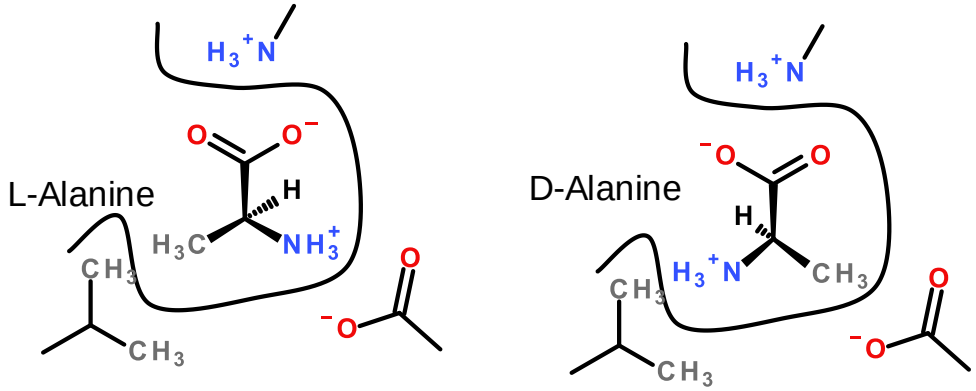
Carbon, oxygen, and nitrogen often form bonds with a tetrahedral geometry

Unique Feature of Tetrahedral Carbon - Chirality

- A single tetrahedral carbon atom can have four groups attached (group = collection of atoms)
- If the four groups are different, then two forms of the molecule are possible, they are **mirror images** of each other.
- The carbon that has four different *groups* is called a **chiral carbon.**
- The two different mirror-image molecules are called **enantiomers**
- These two *cannot be superimposed* on each other (superimposed = rotated so that the same atoms overlap)
- A mixture of both enantiomers is called a **racemic mixture**.
- One naming system to distinguish enantiomers is D & L



Two Different enantiomers (L and D alanine) binding to the same receptor

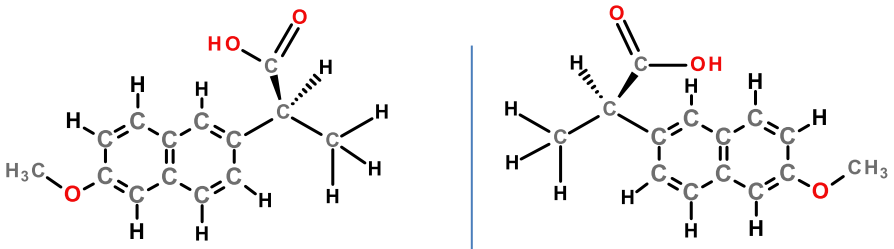


L-Alanine binds better because of more favorable _____ interactions.

- The biological properties of each enantiomer may be quite different, because they may interact with different receptors in the cell.

Naproxen (Aleve)

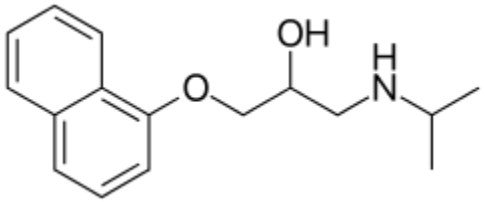
Chiral carbon



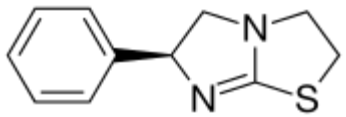
Liver toxin

Anti-inflammatory

Propranolol
1.

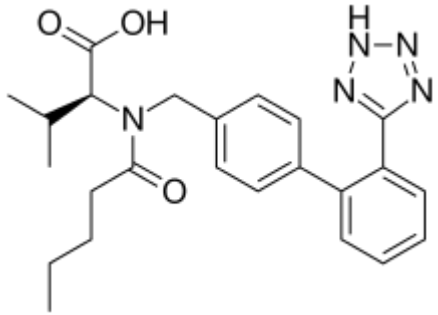


levamisole
4.

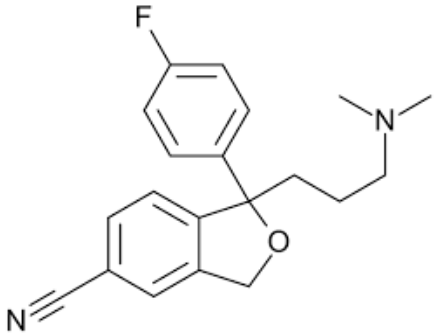


Instructions: Go to the google slide with the same number as your breakout room

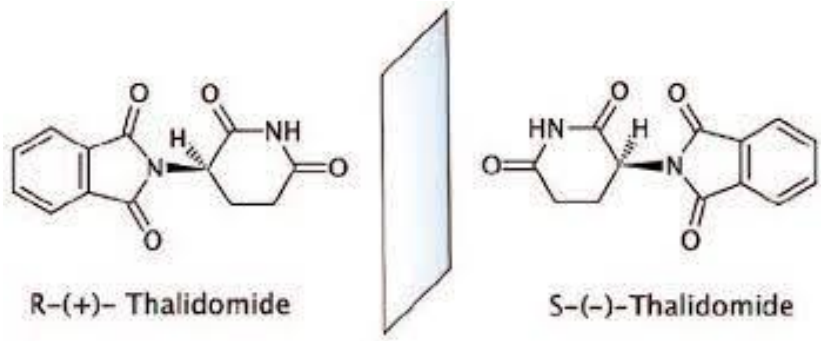
Valsartan
2.



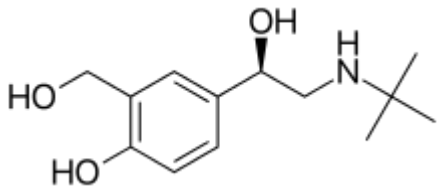
citalopram
5.

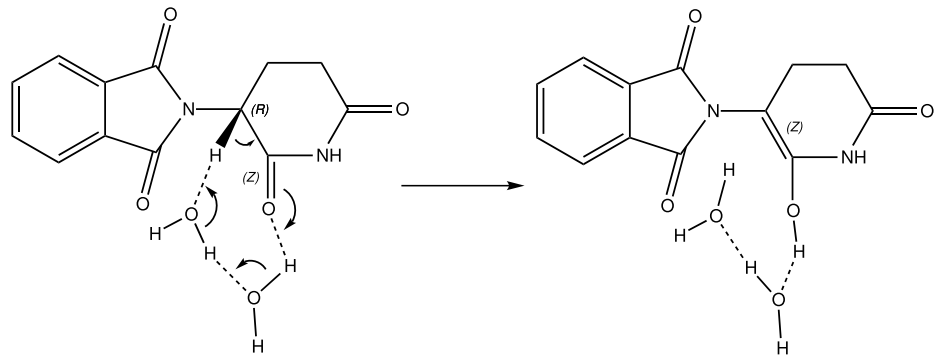


Thalidomide
6.



levalbuterol
3.





https://docs.google.com/presentation/d/1PJ33ZCi55w4Bdjg4KVQ0qf2HrCfwVHKLcdkvIbUbdfw/edit?usp=sharing

pH, Strong Acids & Bases

- pH = –log [H^+]
- The pH of a solution tells us how acidic the solution is.
- The pH scale is used to transform the large range of possible [H+] values to more manageable numbers.
- **Note a low pH is a high [H+].**

The pH is a property of the solvent (water) and can be changed by the addition of a strong acid or base, such as HCl or NaOH.

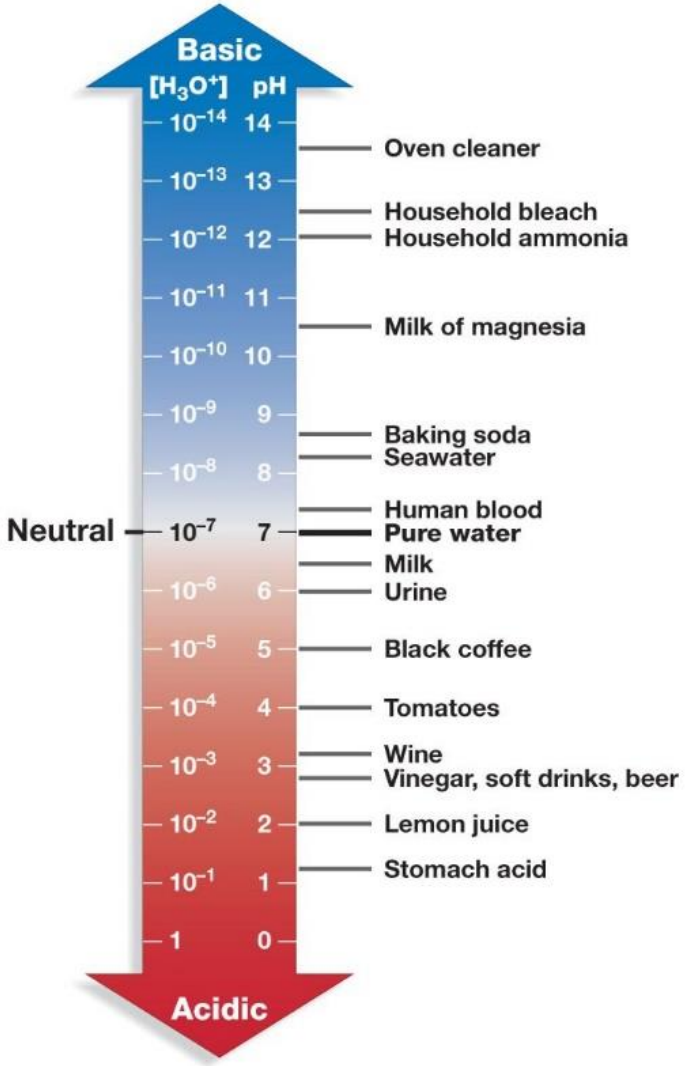
- Acids release protons and will lower the pH of the solution, e.g.

  $HCl + H_2O \rightleftharpoons H_3O^+ + Cl^-$

- Bases (e.g. ammonia, sodium hydroxide) will absorb protons and lower the hydrogen ion concentration. These increase the pH.

  $NaOH \Rightarrow Na^+ + OH^-$

  $OH^- + H_3O^+ \rightleftharpoons 2\ H_2O$



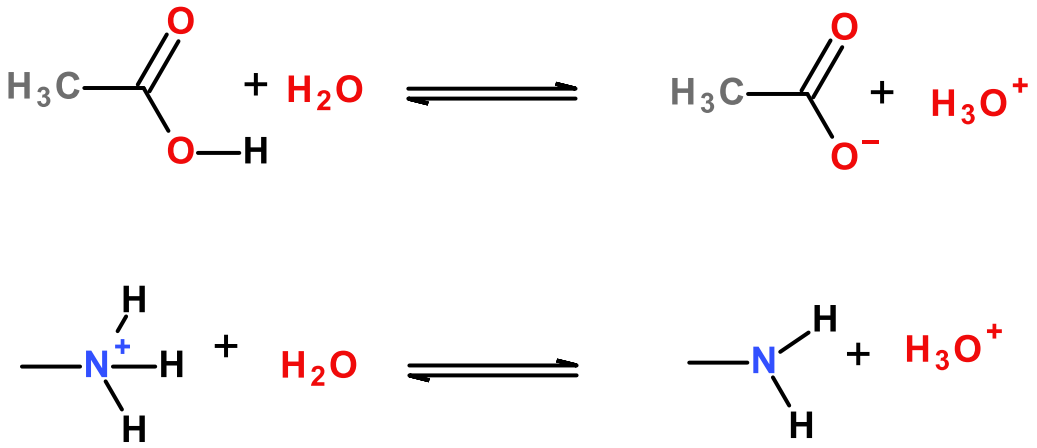
1. Which solution has a higher H^+ concentration, pH=3 or pH 4.
2. How large is the difference?

Acids and Bases.

Strong acid – complete ionization in solution. e.g.

$HCl \longrightarrow H^+ + Cl^-$

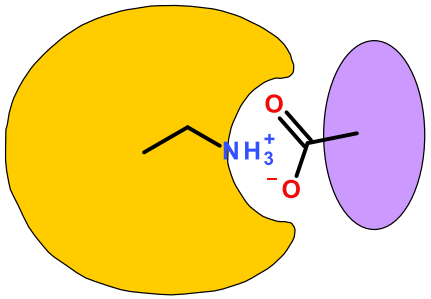
Weak Acid – incomplete ionization in solution.

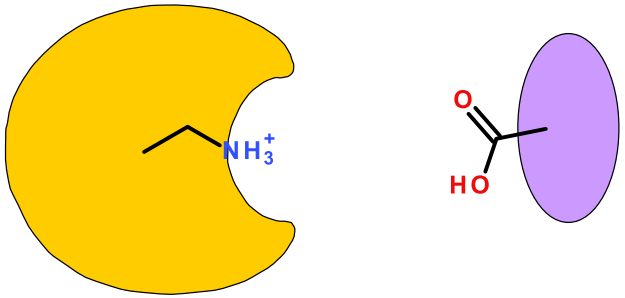


"HA"=protonated form

"A"=deprotonated form (conjugate base)

Why this is important: protonation/deprotonation changes the ***charge*** on species, either creating or destroying strong electrostatic interactions!





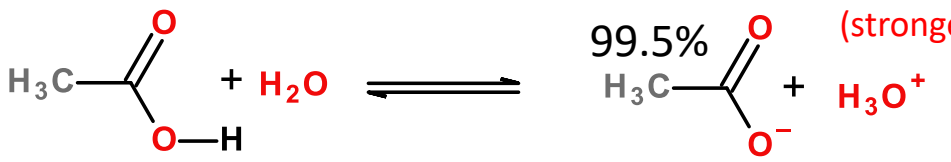
What Affects the Degree of Protonation?

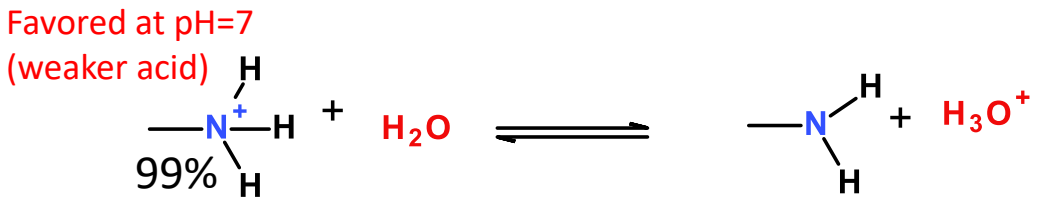
1. The extent of protonation/deprotonation depends on the pH of the solution:
   - Low pH values will favor protonation of acids since there are many protons that will collide with (A) to make (HA).
   - High pH values will favor deprotonation of acids since there are fewer protons to protonate the acid.

*What would you expect to happen to the fraction of the acid that is protonated (f_{HA}) as the pH of the solution is **decreased**?*

$$f_{HA} = \frac{[HA]}{[HA] + [A]}$$

2. The amount of protonated/deprotonated species *also* depends on the chemical properties of the acid. Comparing acetic acid to a protonated amine. At neutral pH (7) most of the acetic acid will be deprotonated while most of the amine will be protonated.





The pKa of an acid is the pH where equal amounts of protonated and deprotonated species are found.

Key Points & Expectations

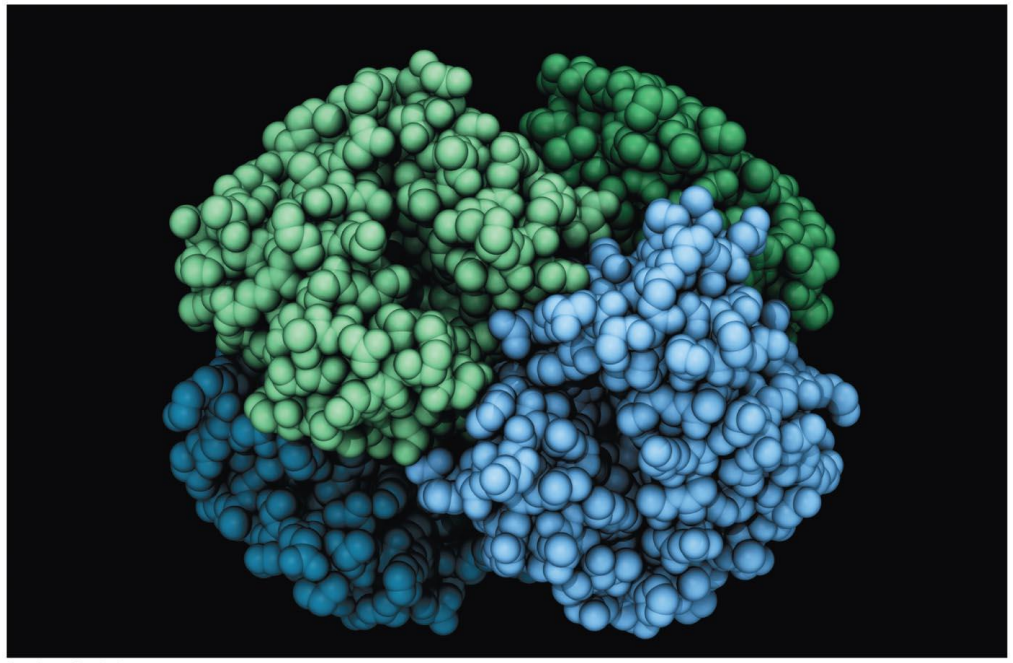
Chemistry

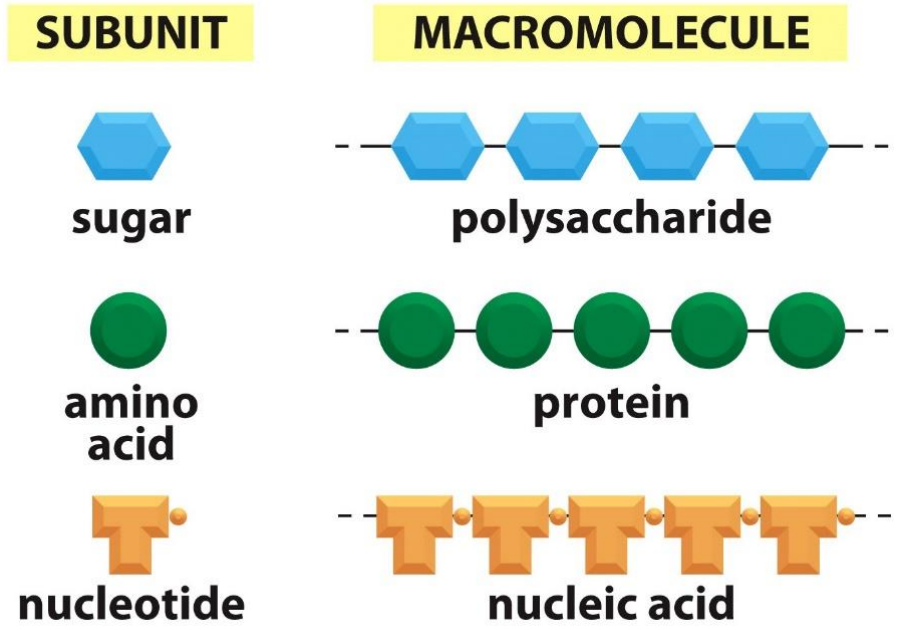
- Number of bonds formed by common elements:

  (N=3, C=4, O=2, S=2, H=1).

- You should be able to complete chemical structures by adding hydrogens to carbons.
- Chiral carbon and enantiomers - different enantiomers can have different properties. You need to identify chiral carbons.
- Polar (unequal charge distribution, e.g. N-H) versus non-polar bonds (e.g. C-H). You need to be able to identify polar and non-polar bonds.
- H-bond - Partial charges due to X-H interacting with Y (X & Y electronegative)
- H-bond - Identify donors and acceptors, partial charges
- pH – be able to predict the charge on a group, given the pH of the solution and the pKa of the acid.

Proteins and Amino Acids





MOLECULAR MACHINERY: A Tour of the Protein Data Bank

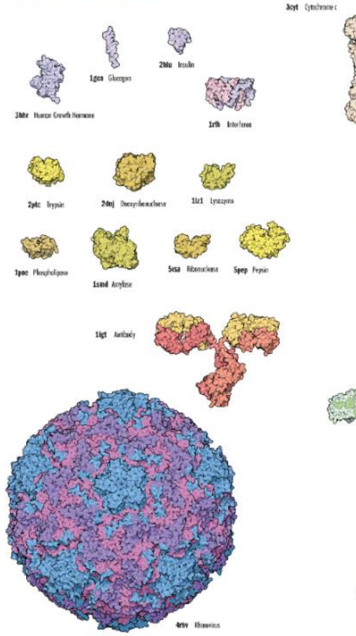
Living cells are filled with complex molecular machinery, a million times smaller than familiar machines like computers or automobiles. Cells use these tiny molecular machines to perform all of the jobs needed for life. Some are molecular scissors that cut food into cell-sized pieces. Some build new molecules when cells grow or when damaged tissues are repaired. Some are molecular bones and muscles that support cells and help them move and crawl. Some fight off attackers, defending against infection.

Researchers around the world are studying these molecules and determining their precise atomic structures. These structures are available on the internet through the Protein Data Bank (http://www.pdb.org), the central storehouse of biomolecular structures. A few of the thousands of structures held in the Protein Data Bank are shown here. In these pictures, the molecules are all drawn at a magnification of 3,000,000 times, and each atom is shown as a small sphere. Many of these structures are composed of several subunits, which are indicated by different colors. An enormous range of sizes is shown here: the water molecule at the left has only three atoms and the rhinovirus shown below has hundreds of thousands.

Water
Glucose
Adenosine triphosphate (ATP)

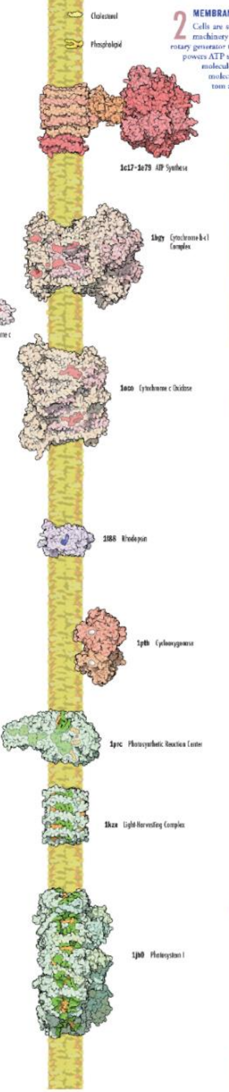
By David S. Goodsell, The Scripps Research Institute, La Jolla, California, USA

Graphic design by Gail W. Bamber, San Diego Supercomputer Center



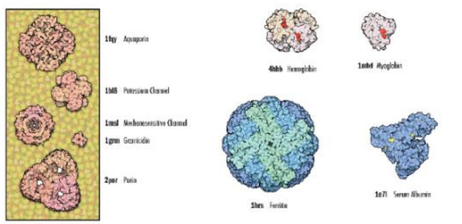
1 OUTSIDE THE CELL

Some molecular machines perform their jobs outside of cells. Many are compact, so that they can diffuse quickly to their site of action. This is true of the four hormones shown at the top: insulin and glucagon, which together regulate blood sugar levels, interferon, which carries signals in the immune system, and human growth hormone. The seven digestive enzymes (in yellow) are also small and very stable, so that they can survive the hostile environment in the digestive tract. Each of these enzymes has a small groove (oriented towards the top in each) that binds to a different target molecule and digests it. At the bottom is rhinovirus, the virus that causes the common cold, and an antibody, our major defense against viruses. Antibodies bind to viruses and prevent them from binding to cell surfaces, thus blocking infection.



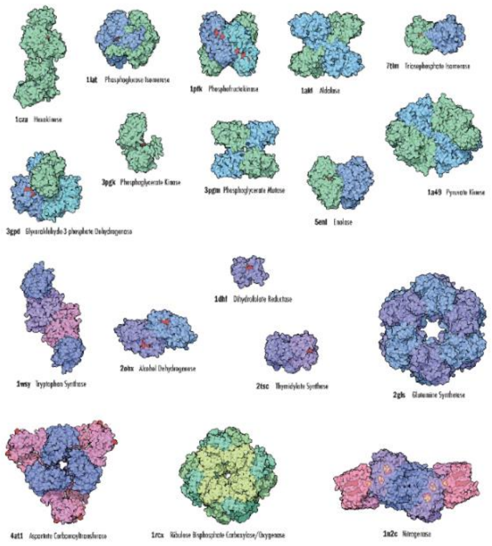
2 MEMBRANES

Cells are surrounded by a membrane made of lipids, like the phospholipid and cholesterol molecules shown at the top. Membranes keep the cellular machinery inside and unwanted material out. Many proteins are embedded in this membrane, performing a variety of essential tasks. ATP synthase is a rotary generator that produces ATP (adenosine triphosphate), the small molecule used for powering cells. The two large complexes below it charge a battery that powers ATP synthase, and the tiny protein cytochrome c shuttles electrons between them. Rhodopsin is found in membranes in the retina. The small retinol molecule inside of it changes shape when illuminated, causing the surrounding protein to send a signal to the brain. Cyclooxygenase builds one of the molecules used to signal pain—this cyclooxygenase molecule here, however, is blocked by two molecules of aspirin, shown inside in white. At the bottom are three molecules involved in photosynthesis, which capture energy from light and use it to power the synthesis of sugar in plant cells.



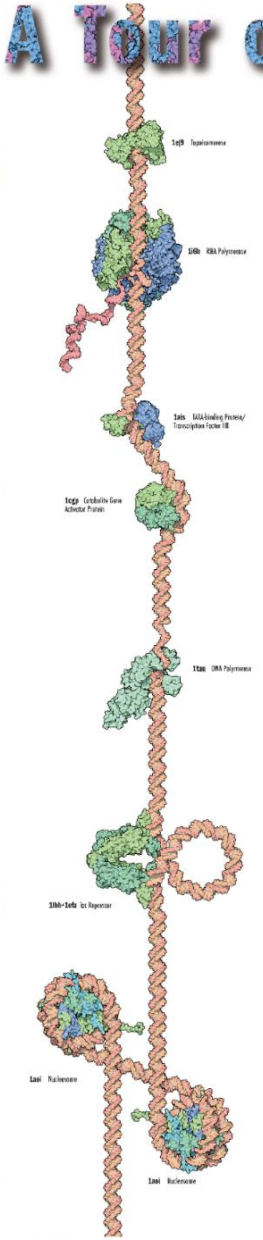
3 TRANSPORT AND STORAGE

Of course, a perfectly sealed membrane would be of little use to cells, because nutrients could not get in and wastes could not get out. The box shows a membrane looking face-on. Five proteins that form channels through the membrane are shown. To the right of the box are several soluble proteins involved in transport and storage of molecules. Hemoglobin and myoglobin carry oxygen. Ferritin forms a hollow shell that stores iron ions. Serum albumin carries many different molecules in the blood.


1cza Hexokinase
1iat Phosphoglucose Isomerase
1pfk Phosphofructokinase
1ald Aldolase
7tim Triosephosphate Isomerase
3gpd Glyceraldehyde-3-phosphate Dehydrogenase
3pgk Phosphoglycerate Kinase
3pgm Phosphoglycerate Mutase
5enl Enolase
1a49 Pyruvate Kinase
1dhf Dihydrofolate Reductase
2wsy Tryptophan Synthase
2ohx Alcohol Dehydrogenase
2tsc Thymidylate Synthase
2gls Glutamine Synthetase
4at1 Aspartate Carbamoyltransferase
1rcx Ribulose Bisphosphate Carboxylase/Oxygenase
1n2c Nitrogenase

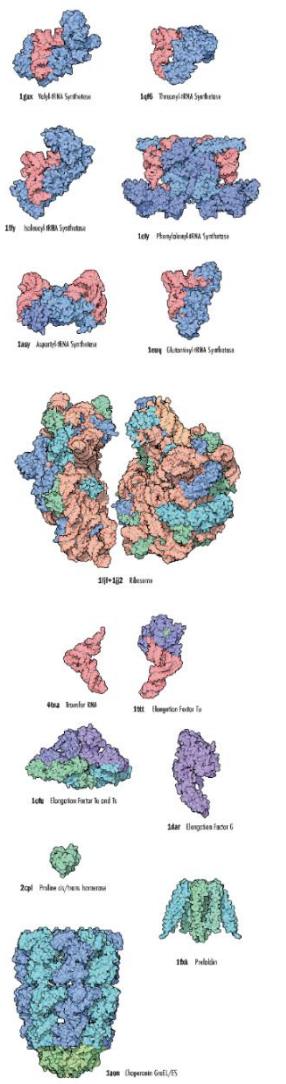
4 CHEMICAL FACTORIES

Cells build a bewildering variety of enzymes—proteins that perform chemical reactions. At the top are the ten enzymes that perform glycolysis, the breakdown of sugar to form ATP. Below that are several enzymes that perform different housekeeping reactions. Dihydrofolate reductase activates a key cofactor molecule and alcohol dehydrogenase breaks down alcohol. Ribulose bisphosphate carboxylase/oxygenase is the most common enzyme on the Earth, and performs a key step in the capture of carbon dioxide by plants to form sugar. The three synthases and the transferase make different building blocks for creating new molecules. Nitrogenase performs an essential role in the ecosystem by converting nitrogen gas into a form that living cells can use.



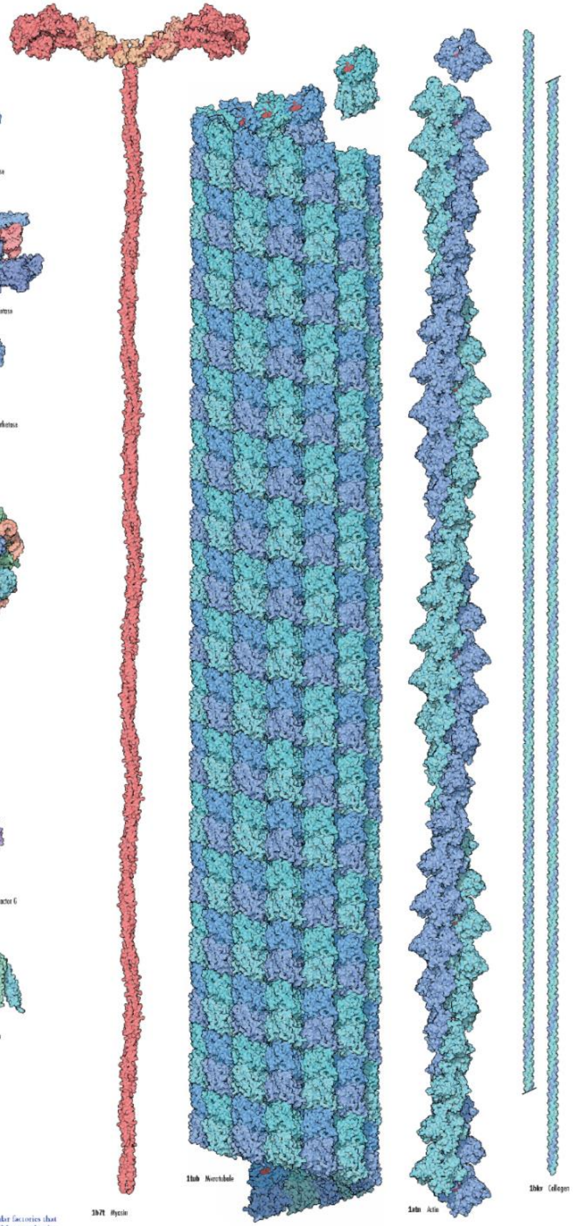
5 DNA

Genetic information is stored in the DNA double helix, seen running from top to bottom here. Many proteins are used to copy, read, and store this information. RNA polymerase copies the information into a strand of RNA that will be used to direct the construction of new proteins. It is assisted by topoisomerase, which releases tension when the helix is wound and unwound, and guided to appropriate starting points by the two protein complexes below it. DNA polymerase replicates DNA strands—here, the polymerase is filling a gap in the double helix. Some proteins, like the lac repressor, grab DNA and bend is sharply, or even wrap it all the way around themselves, like the two nucleosomes at the bottom.


1gax Valyl-tRNA Synthetase
1qf6 Threonyl-tRNA Synthetase
1ffy Isoleucyl-tRNA Synthetase
1eiy Phenylalanyl-tRNA Synthetase
1asy Aspartyl-tRNA Synthetase
1euq Glutaminyl-tRNA Synthetase
1fjf+1jj2 Ribosome
4tna Transfer RNA
1ttt Elongation Factor Tu
1efu Elongation Factor Tu and Ts
1dar Elongation Factor G
2cpl Proline cis/trans Isomerase
1fxk Prefoldin
1aon Chaperonin GroEL/ES

6 BUILDING NEW PROTEINS

New proteins are built by ribosomes—complex molecular factories that read the genetic code and use it to direct construction. Many molecular machines are needed to assist the process. Twenty different aminoacyl-tRNA synthases (six are shown here) load the building blocks onto tRNA, ready to be added to a growing protein chain. Several protein factors, shown below the ribosome, guide each tRNA into the proper spot. The three chaperone proteins shown at the bottom help each new protein fold into its proper shape.



7 BEAMS AND GIRDERS

Cells are braced and supported by a complex infrastructure. This cytoskeleton is formed of sturdy filaments like actin and microtubules, composed of many subunits stacked like bricks. Myosin is a molecular motor that climbs along actin filaments, allowing the cell to move. Collagen, broken into two pieces here, is actually found outside of cells, where it forms connective tissue between cells.

February 2003

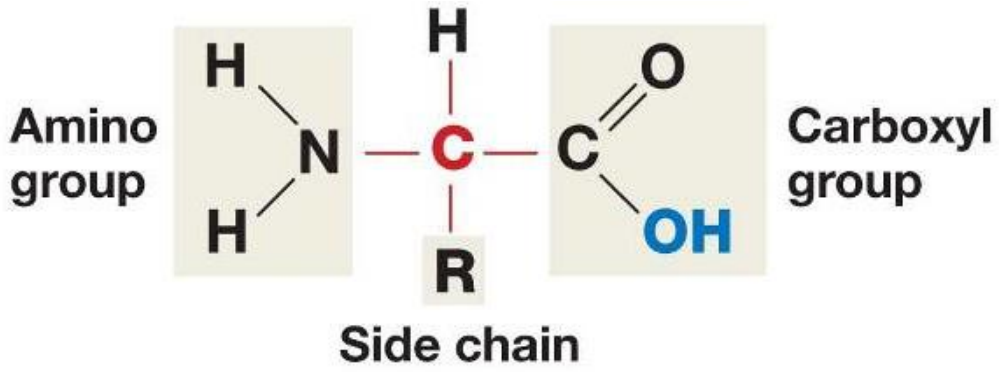
PDB PROTEIN DATA BANK

http://www.pdb.org/ • info@rcsb.org

RESEARCH COLLABORATORY FOR STRUCTURAL BIOINFORMATICS

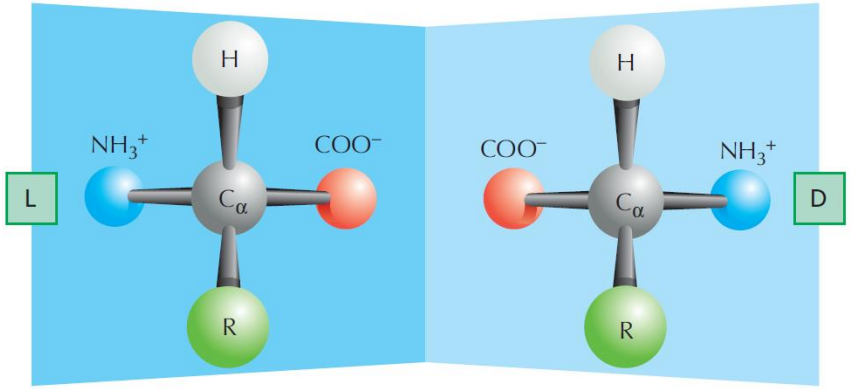
RUTGERS, THE STATE UNIVERSITY OF NEW JERSEY
SAN DIEGO SUPERCOMPUTER CENTER
NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY

The Structure of Amino Acids and Proteins

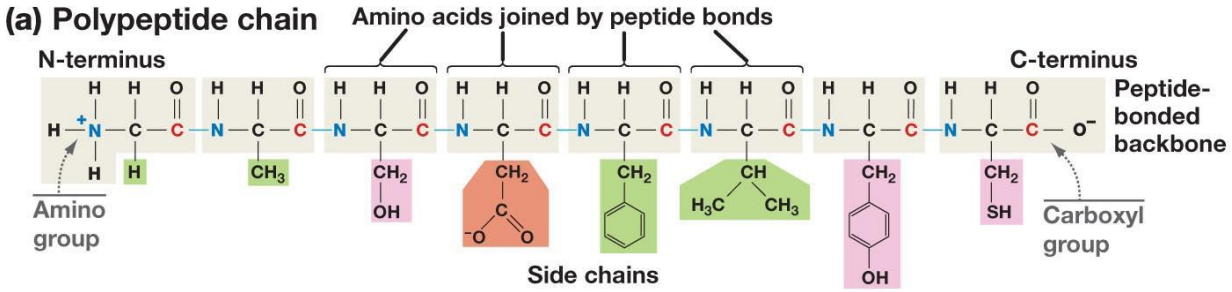


Is there a chiral carbon on amino acids?

- *The amino group, Cα (and one hydrogen), and the carbonyl group are common to all amino acids*
- *The N-Cα-C=O are the mainchain of the protein polymer.*
- *The R groups are different –there are 20 common R groups they are the sidechain of the protein polymer – their **sequence** defines the properties of the protein.*

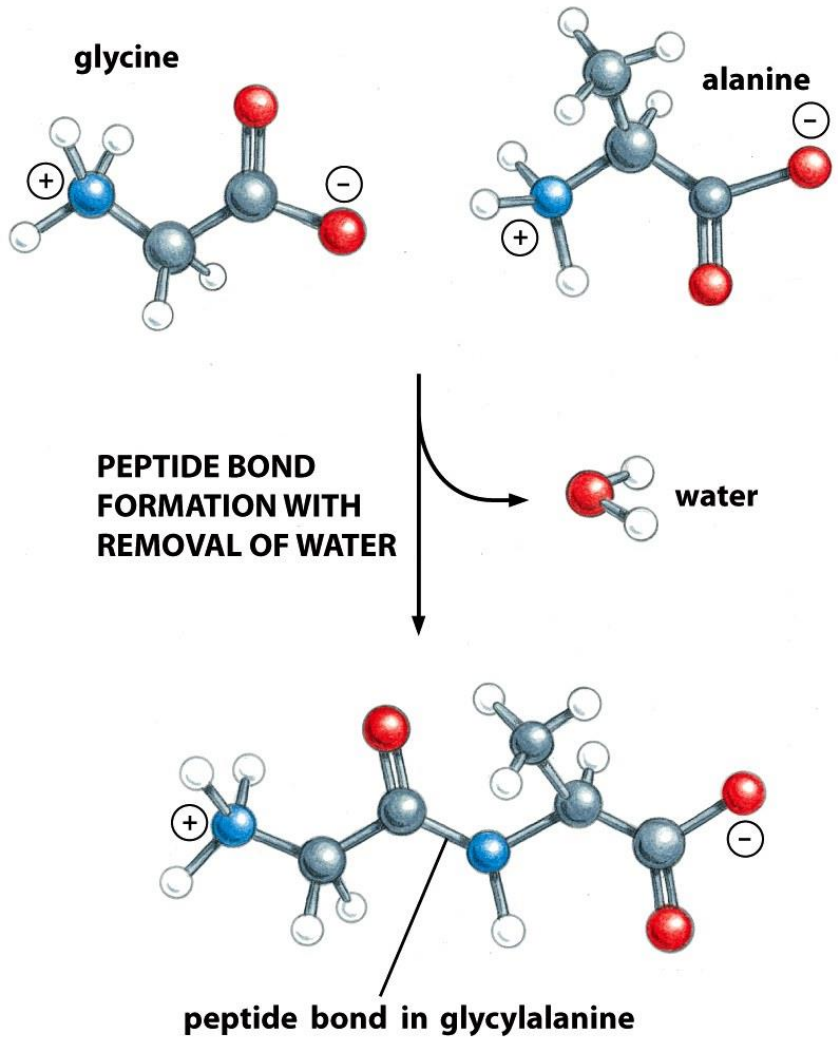


Proteins consist exclusively of L-amino acids.
(as does the ribosome that make them)

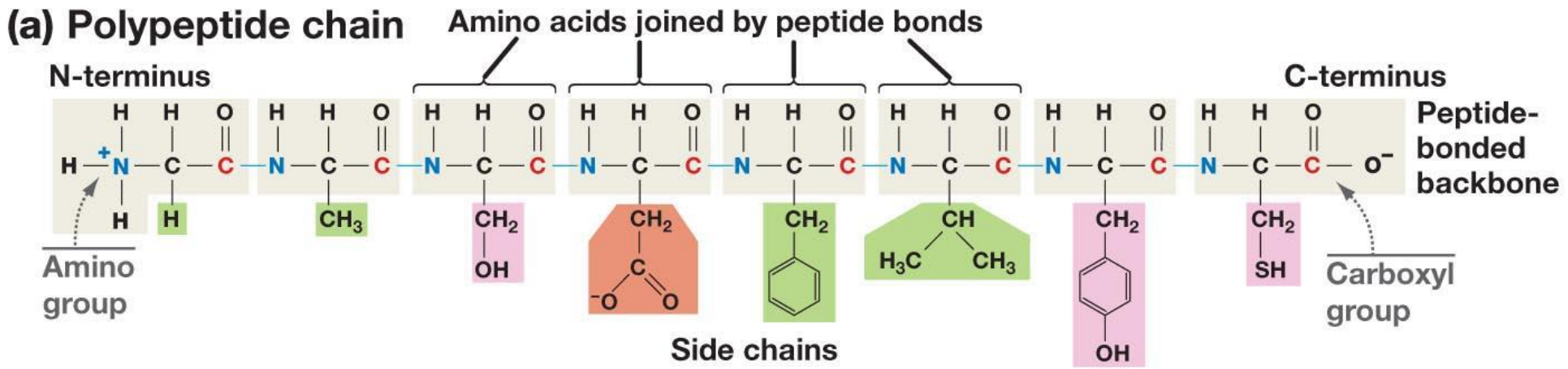


Proteins are generated by formation of Peptide Bonds

- Amino acids are linked to form ***linear*** polymers:
- Dehydration reactions bond the carboxyl group of one amino acid to the amino group of another to form a **peptide bond**.
- A chain of amino acids linked by peptide bonds is called a **polypeptide**.
  - Polypeptides containing fewer than 50 amino acids are called **oligopeptides** (**peptides**).
  - Polypeptides containing more than 50 amino acids are called **proteins.**
  - The four atoms involved in the peptide bond all lie on the same plane.
  - The C=O is usually across from the N-H (trans conformation)



Sidechain *Functional* Groups Affect Behavior (and the order is important)

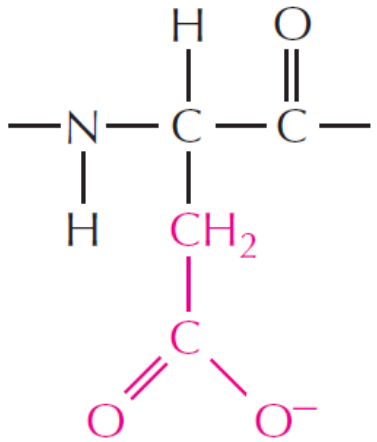


- Sidechains (R-groups) differ in their size, shape, reactivity, and interactions with water.
  1. Nonpolar Sidechains: **hydrophobic**; do not form hydrogen bonds; coalesce in water - typically form the core of folded proteins.
  2. Polar Sidechains: **hydrophilic**; form hydrogen bonds; readily dissolve in water
  3. **Ionizable** Sidechains: Can be charged at certain pH values. Interact strongly with water.

ACIDIC SIDE CHAINS

aspartic acid

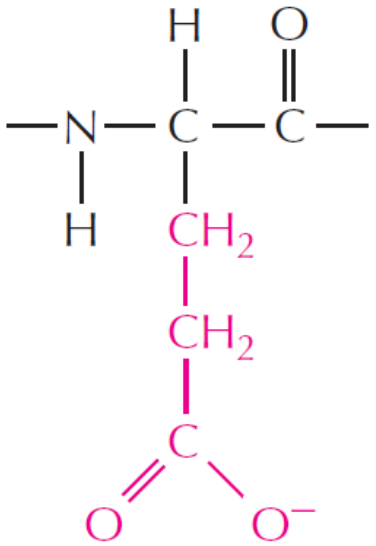
(Asp, or D)



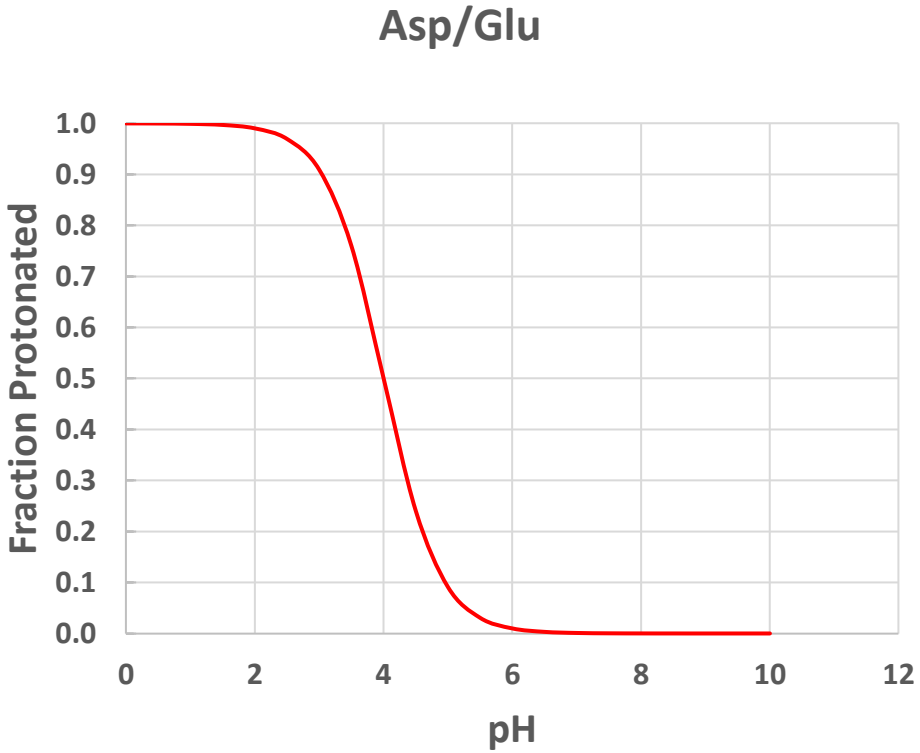
pK_a of sidechain ~ 4

glutamic acid

(Glu, or E)



pK_a of sidechain ~ 4

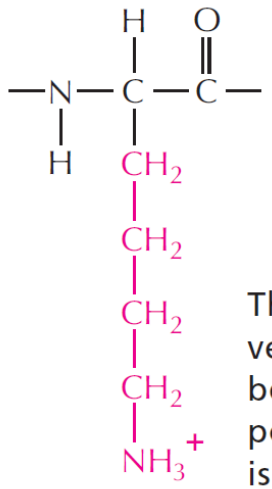


Have a net negative charge at pH 7.0

BASIC SIDE CHAINS

lysine

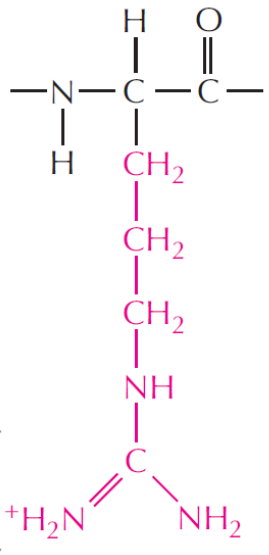
(Lys, or K)



This group is very basic because its positive charge is stabilized by resonance.

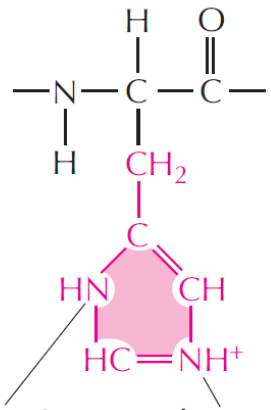
arginine

(Arg, or R)

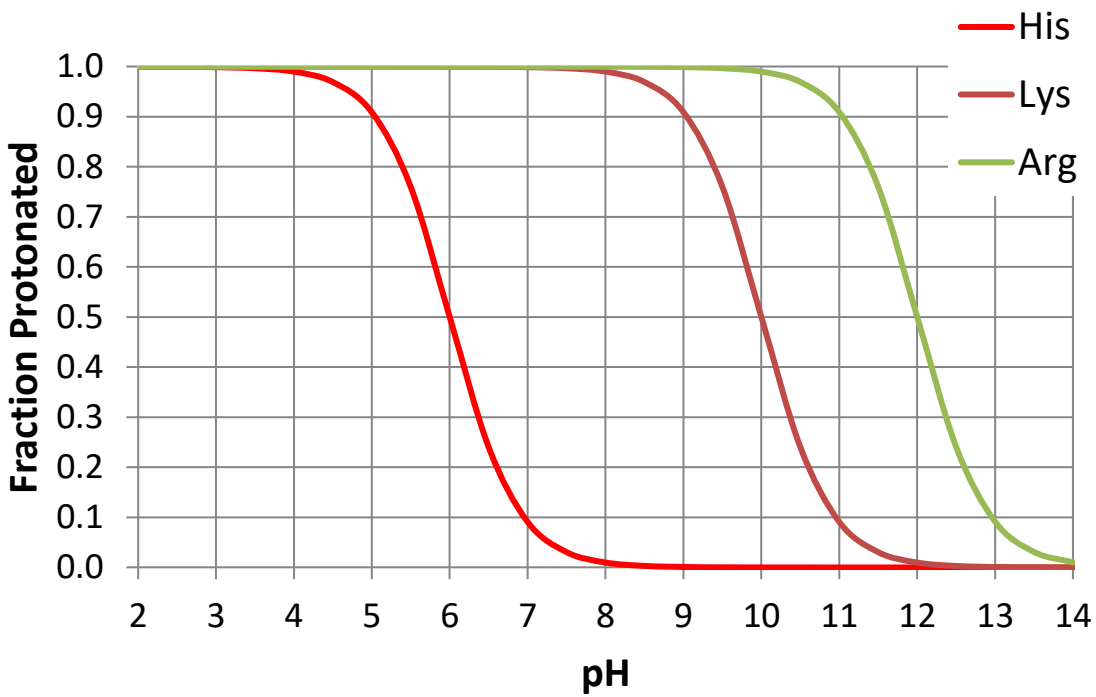


histidine

(His, or H)



These nitrogens have a relatively weak affinity for an H^+ and are only partly positive at neutral pH.



pK_a ~ 10

pK_a ~ 12

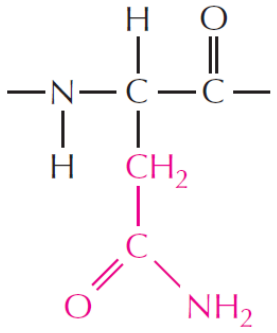
Have a net positive charge at pH 7.0

pK_a ~ 6

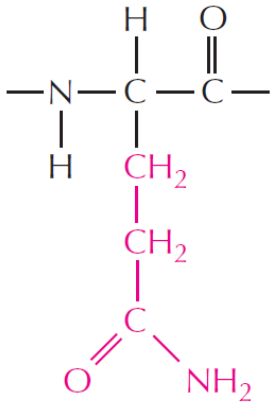
Positive charge when protonated
10% Protonated at pH 7.0

UNCHARGED POLAR SIDE CHAINS

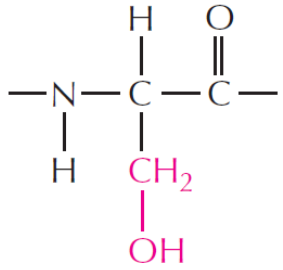
asparagine (Asn, or N)



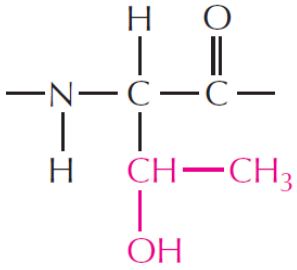
glutamine (Gln, or Q)



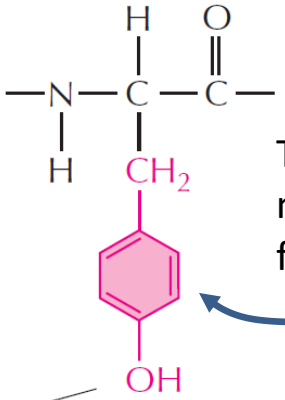
serine (Ser, or S)



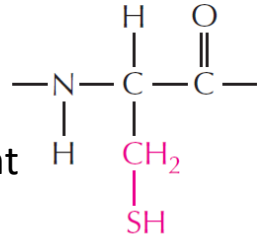
threonine (Thr, or T)



tyrosine (Tyr, or Y)



cysteine (Cys, or C)



This is a significant non-polar functionality

Although the amide N is not charged at neutral pH, it is polar.

The –OH group is polar.

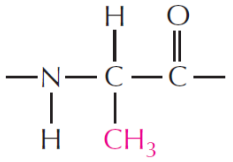
These can form what type of bond with water?

- Forms weak H-bonds
- Can ionize with a pK=8, forming thiolate ion ($-S^-$)
- Can form S-S disulfide bonds with other Cys residues

NONPOLAR SIDE CHAINS

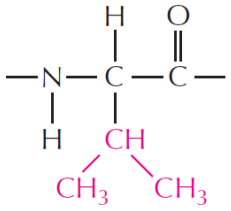
alanine

(Ala, or A)



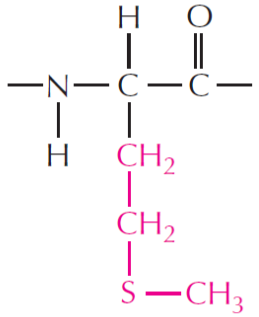
valine

(Val, or V)



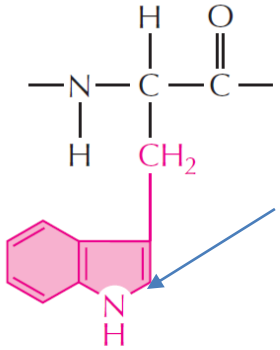
methionine

(Met, or M)



tryptophan

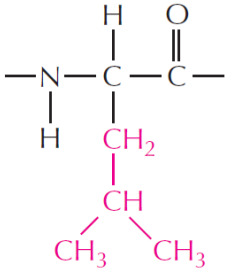
(Trp, or W)



H-bond donor
(polar functionality)

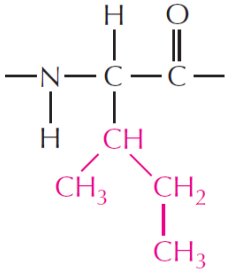
leucine

(Leu, or L)



isoleucine

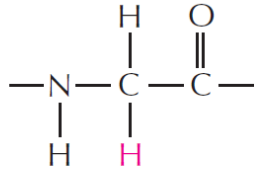
(Ile, or I)



AND AT LAST THERE IS GLYCINE:

glycine

(Gly, or G)

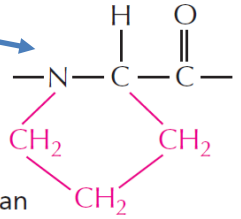


No real functionality for its R group (H)
Only AA that is achiral

Ring results in no NH group for H-bonding

proline

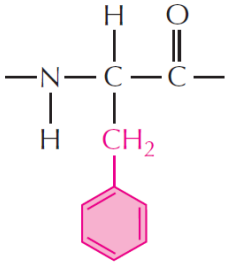
(Pro, or P)



(actually an imino acid)

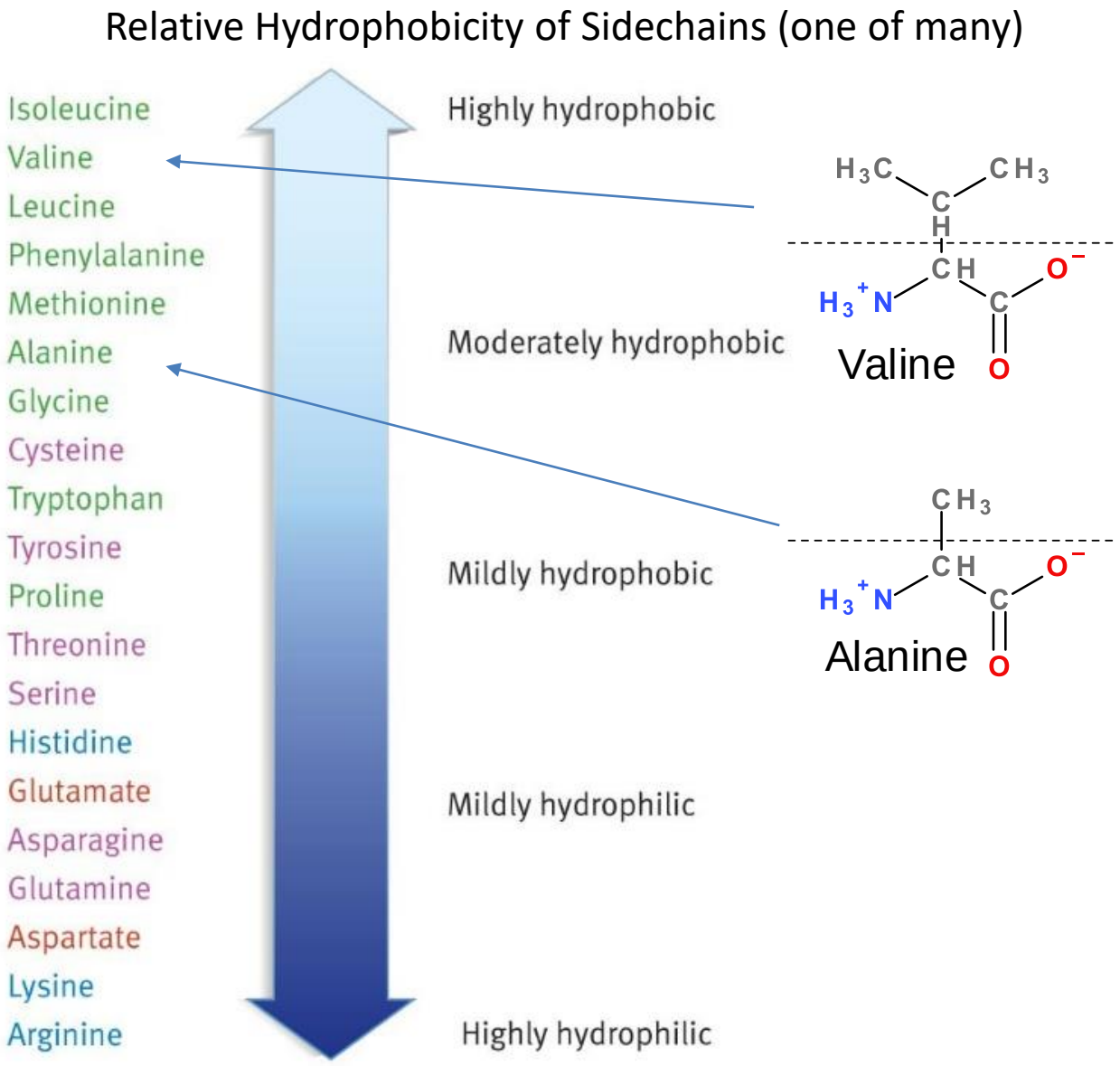
phenylalanine

(Phe, or F)

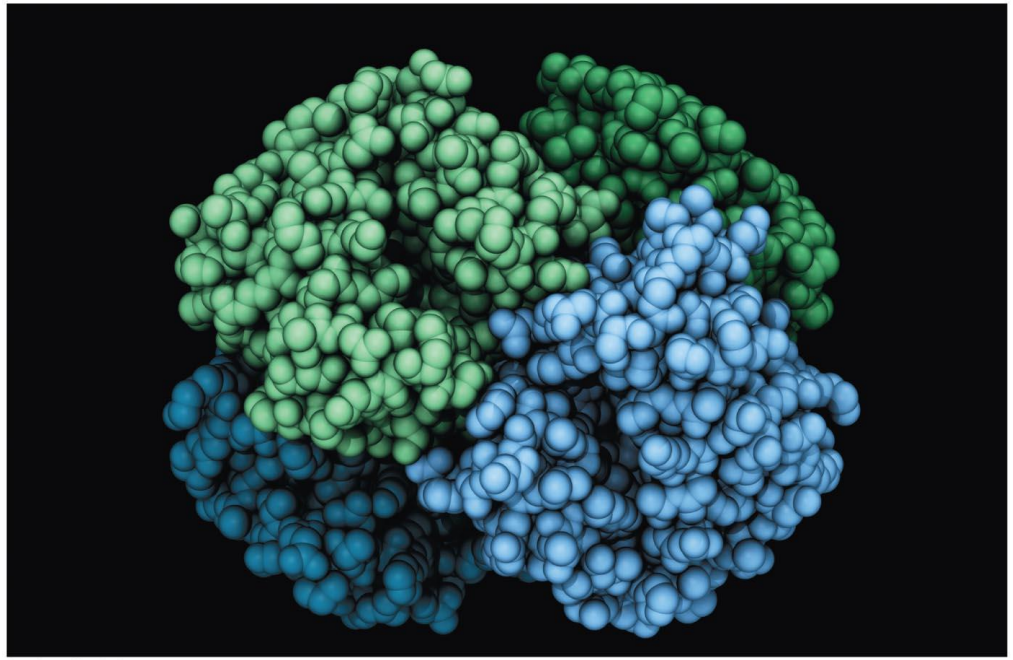


Summary of what you should know about amino acids:

- All amino acids have a carbon atom bonded to an amino group, a hydrogen atom and a carboxyl group. What makes each amino acid unique is its **sidechain**.
- The common atoms will form the **mainchain** of the protein.
- Most amino acids have at least one chiral center - the alpha carbon, exception is glycine, which is achiral.
- You should be able to look at the functional groups on the ***side-chain*** and determine how they will interact with water:
  - Polar
  - Charged
  - Non-polar (hydrophobic). You should be able to justify large differences in hydrophobicity, e.g. Val versus Ala

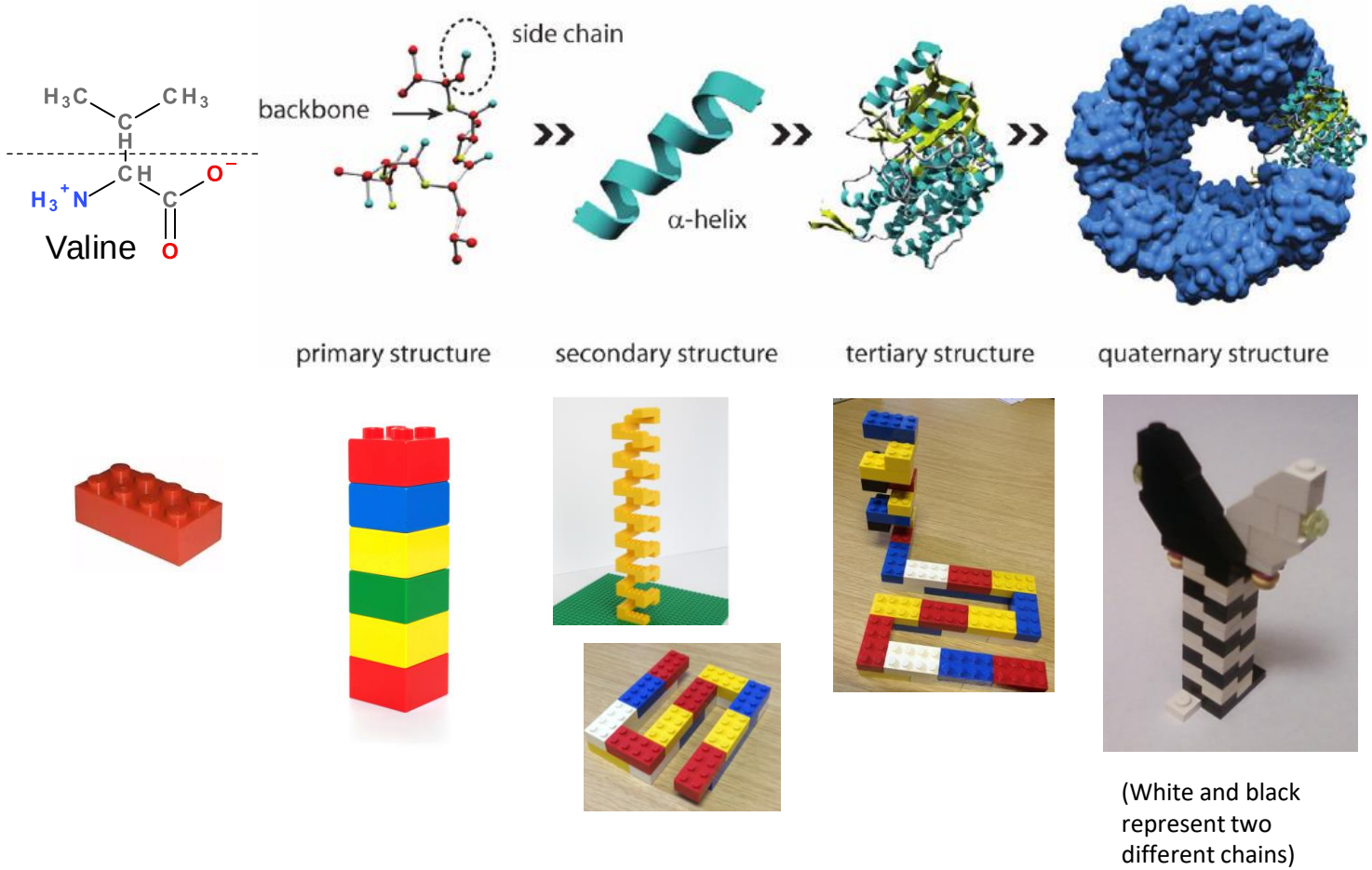


Protein Structure and Stability



Structural Hierarchy of Proteins

- Primary - sequence of amino acids, no 3D structural information
- Secondary - local structural elements, only mainchain atoms involved
- Tertiary - 3D position of ***all*** atoms, functional form of many proteins.
- Quaternary - multiple chains – multiple chains often required for function.



(White and black represent two different chains)

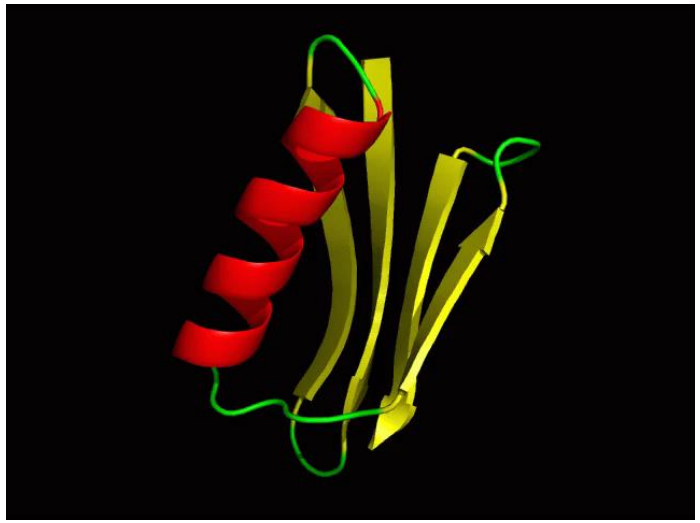
Secondary Structure

"Building blocks of proteins"

- **Hydrogen bonds** between the *mainchain* carbonyl group of one amino acid and the *mainchain* amino group of another form a protein's **secondary structure**.
  - A polypeptide must bend to allow this hydrogen bonding, forming:
    - **α-helices**
    - **β-pleated sheets**

- The large number of hydrogen bonds in a protein's secondary structure increases its stability - each hydrogen bond that is formed releases some energy.
- All amino acids can be incorporated into either secondary structure

  (However, some are found more frequently in one structure)



General Rule for Hydrogen Bonds:

X-H Y

X & Y are electronegative (N and O usually)

X-H = *Donor* of the hydrogen bond

Y = *Acceptor* of the hydrogen bond

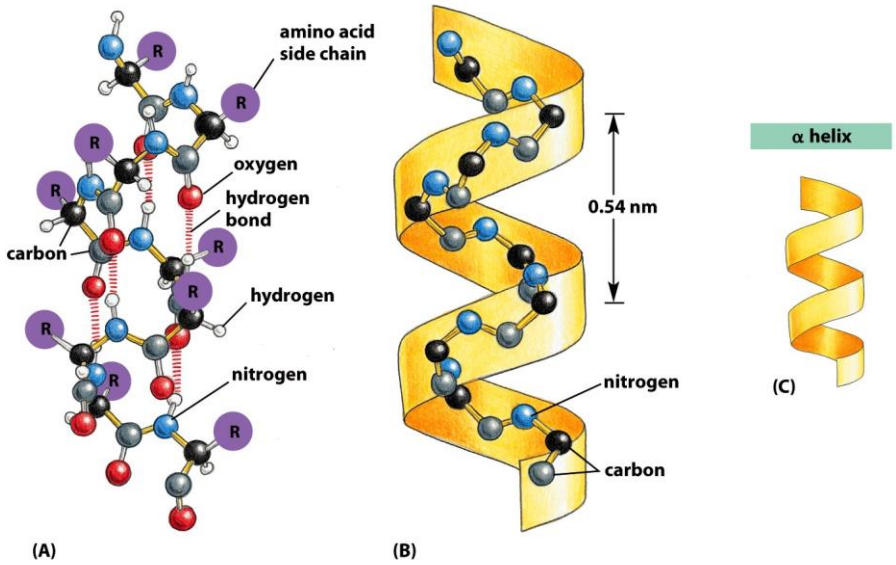
Mainchain hydrogen bonds

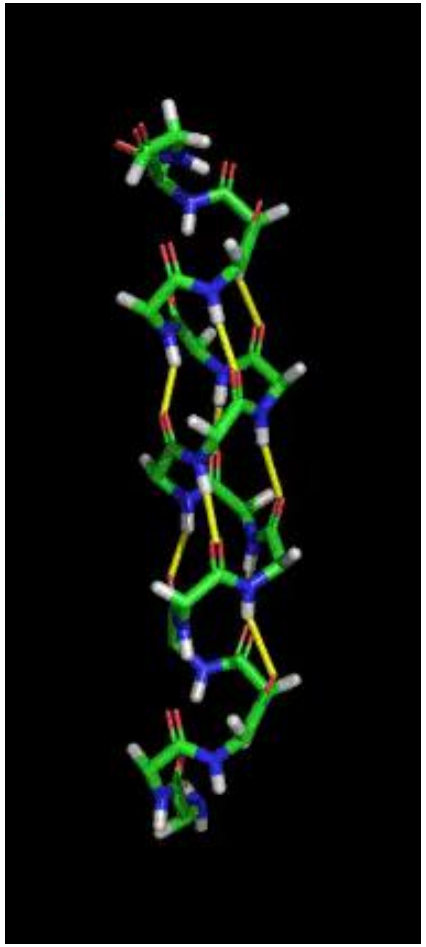
N-H O=C

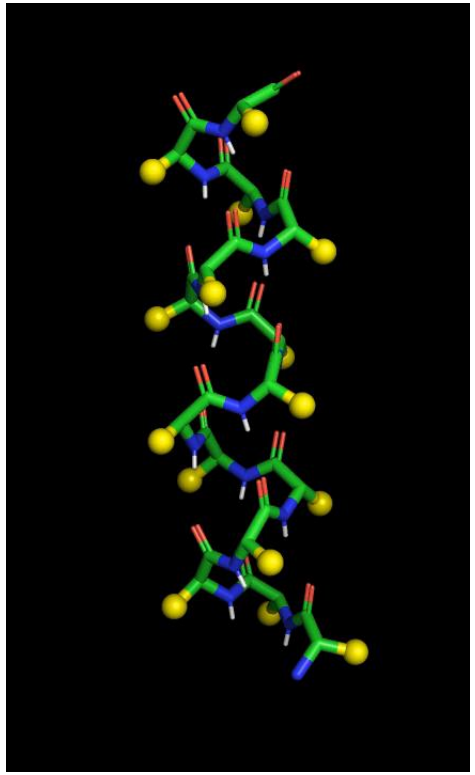
The NH is the hydrogen bond_____.

The C=O is the hydrogen bond _____.

Alpha Helix





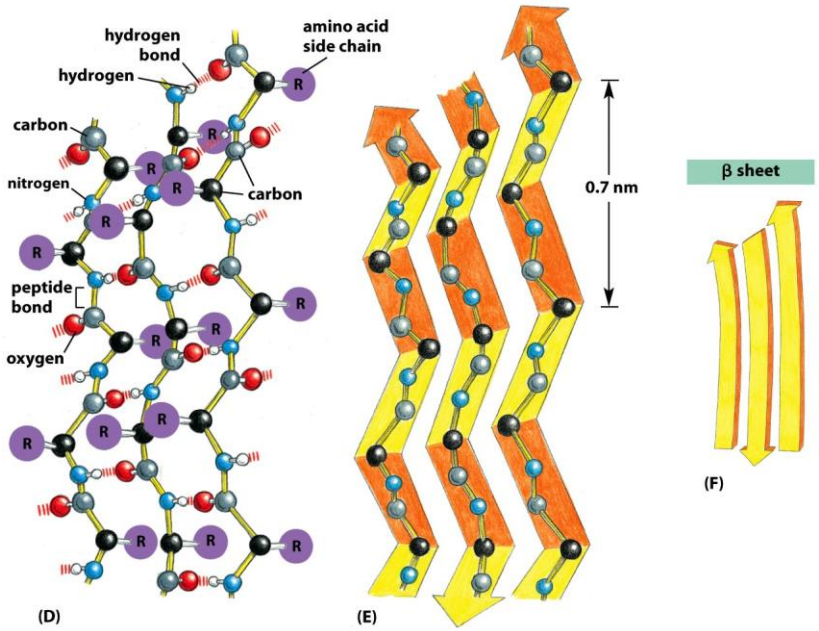


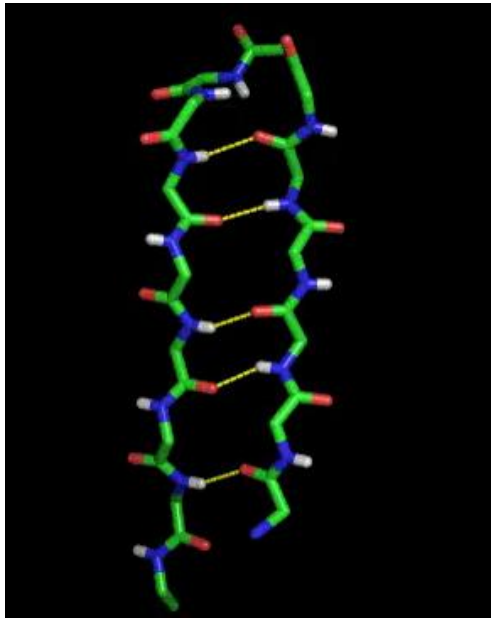
Spiral conformation (*helix*) in which every backbone N-H group donates a hydrogen bond to the backbone C=O group of the amino acid four residues earlier:

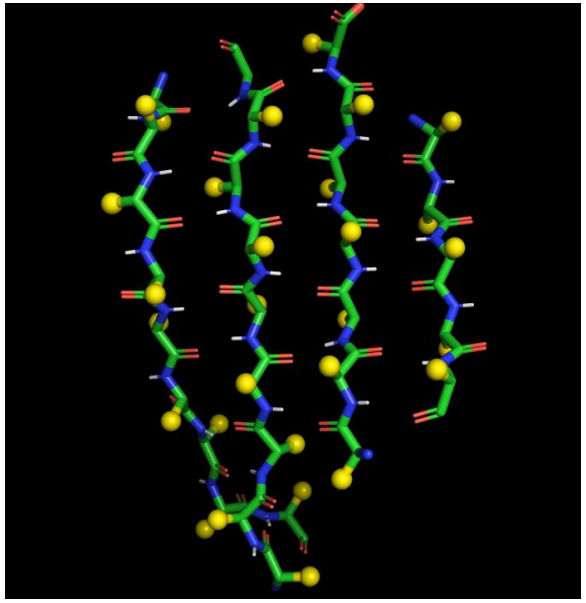
Intra-strand H-bonds, parallel to helix axis.

Side-chains project outwards.

Beta Sheet

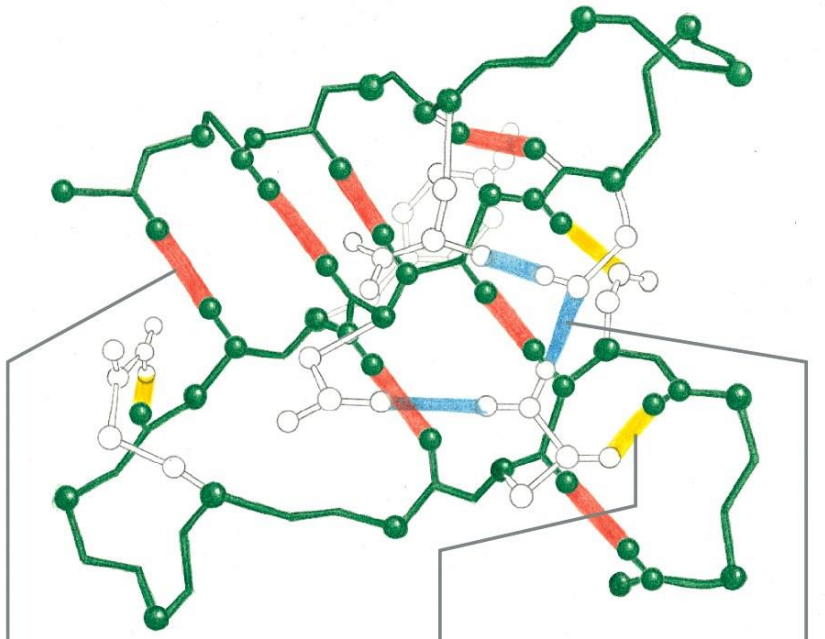


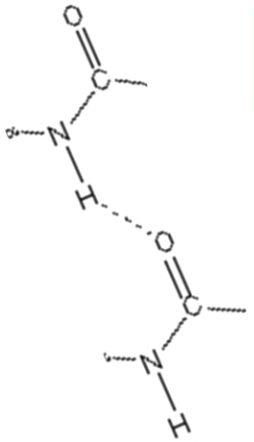


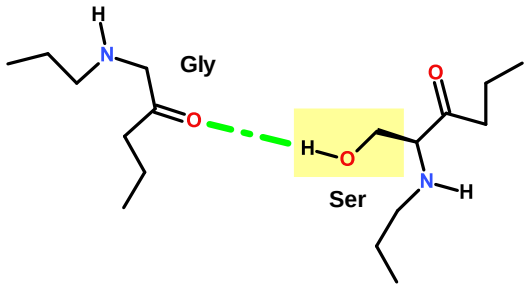


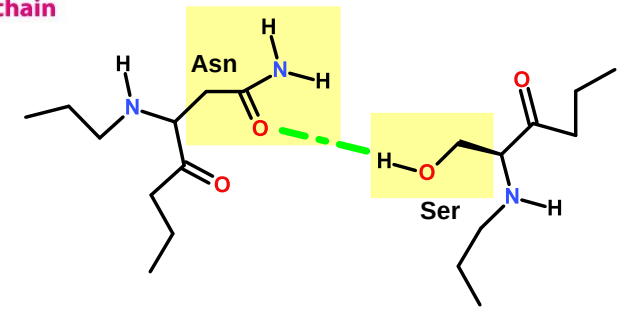
- Beta-Strands connected laterally by backbone hydrogen bonds that are perpendicular to the strand, forming a generally twisted, pleated sheet.
- Sheets can have two or more strands
- Side-chains:
  - project up and down along a strand.
  - project in the same direction going from strand to strand across the sheet.

Hydrogen Bonding Stabilizes the Tertiary Structure





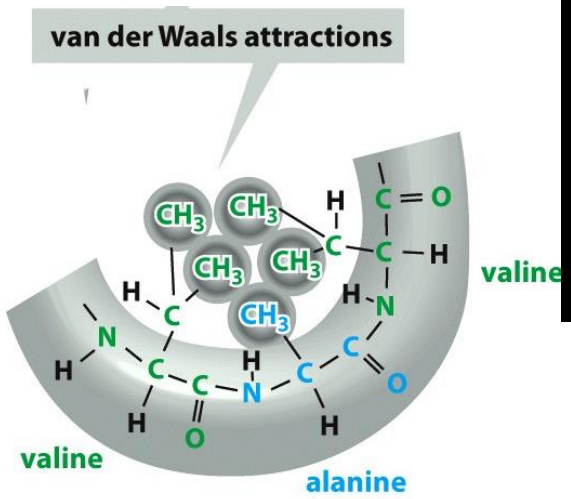


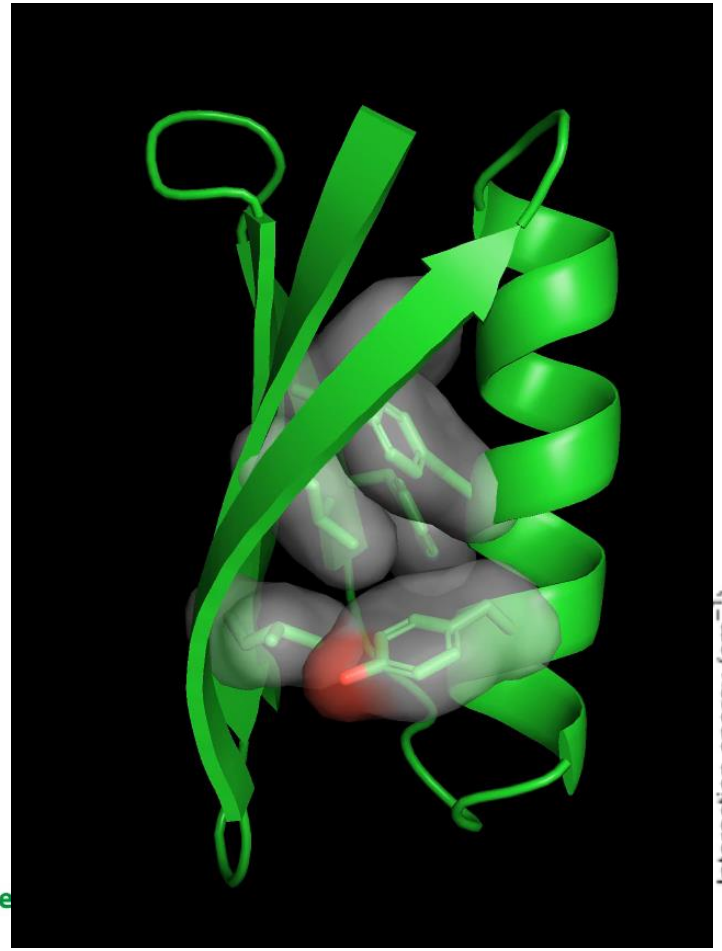


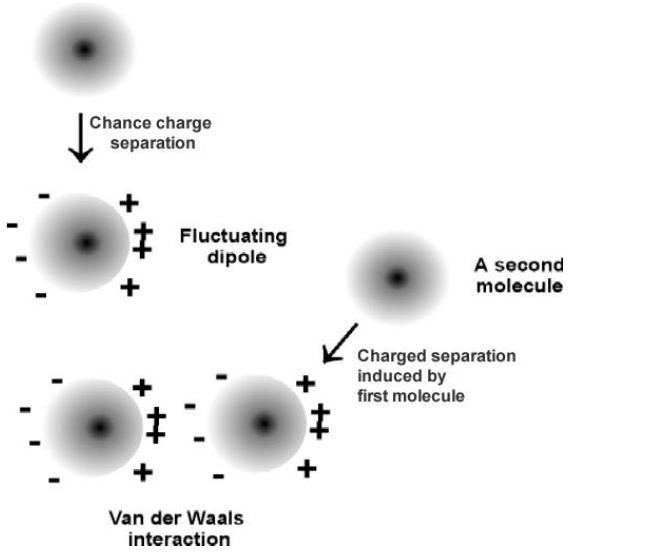
- **Hydrogen bonds** form between hydrogen atoms and the carbonyl group in the peptide-bonded backbone – secondary structure
- Hydrogen bonds are also found between hydrogen and electronegative atoms in side chains (sidechain-sidechain)
- Sidechains can form hydrogen bonds to the mainchain too.

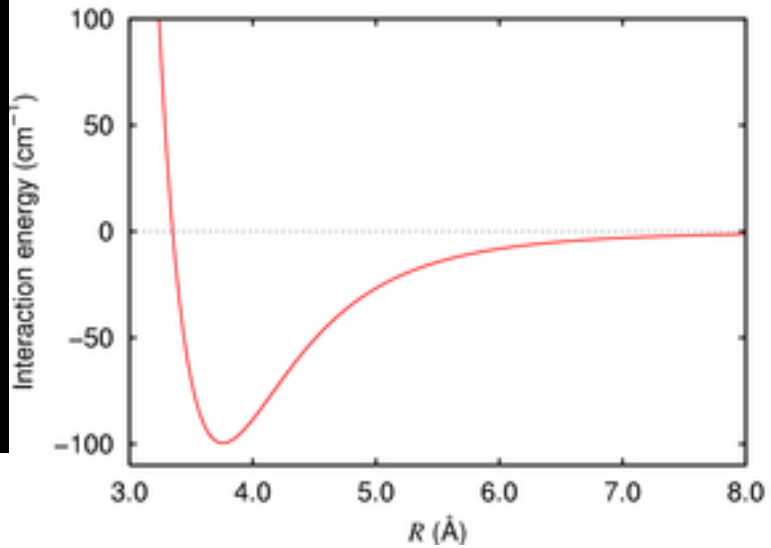
Van der Waals (VdW) interactions Stabilize the Folded State

- VdW are weak electrostatic interactions between side chains due to temporary (fluctuating) charges.
- Attractive from long distance
- Distance at lowest energy is at the van der Waals radii of the atoms.
- Optimized in the core of folded proteins by “knobs fitting into holes”
- Strength proportional to contact area.



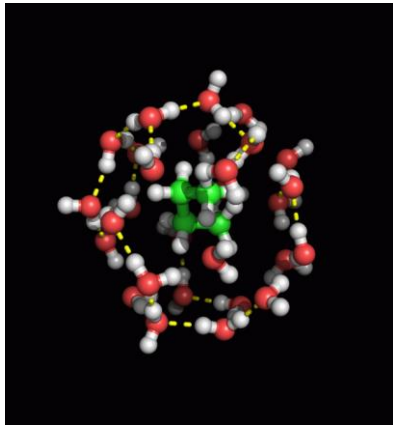




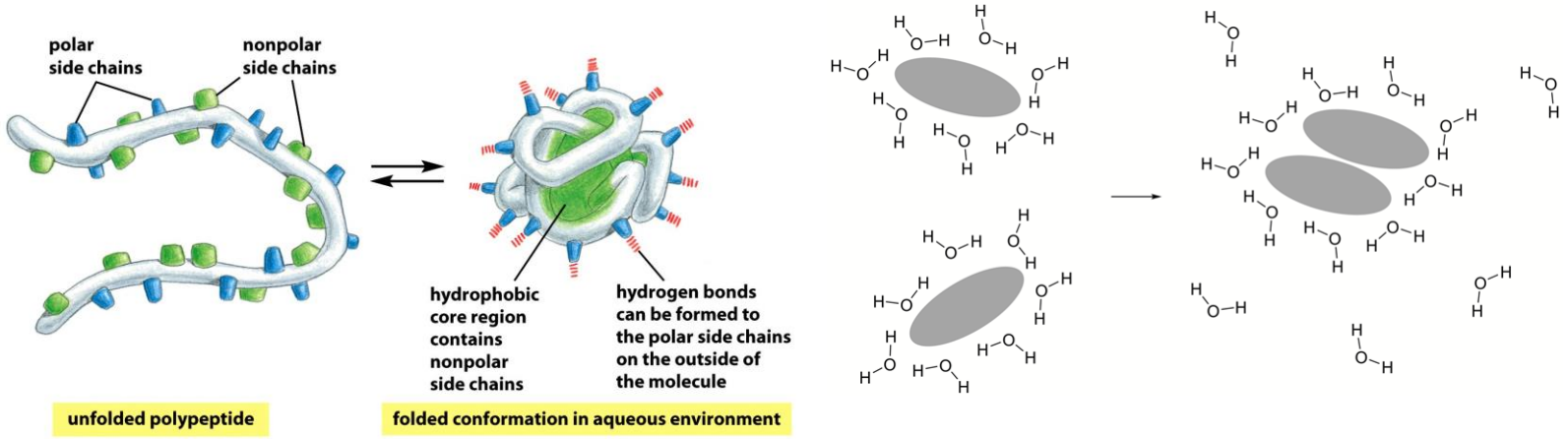


Strength of Van der Waals Depends on the Surface Area

Hydrophobic Interactions are Critical for Stabilizing Folded Proteins

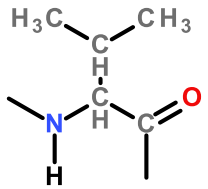


Hydrophobic interactions within a folded protein increase stability of surrounding water molecules by releasing the ordered water that surrounded exposed non-polar groups when the protein is unfolded, **increasing the entropy of the water – disorder is favorable.**



Which amino acid is most likely to be found in the core of a folded protein:

Valine



Serine

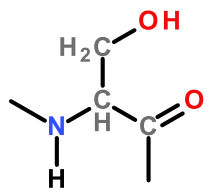
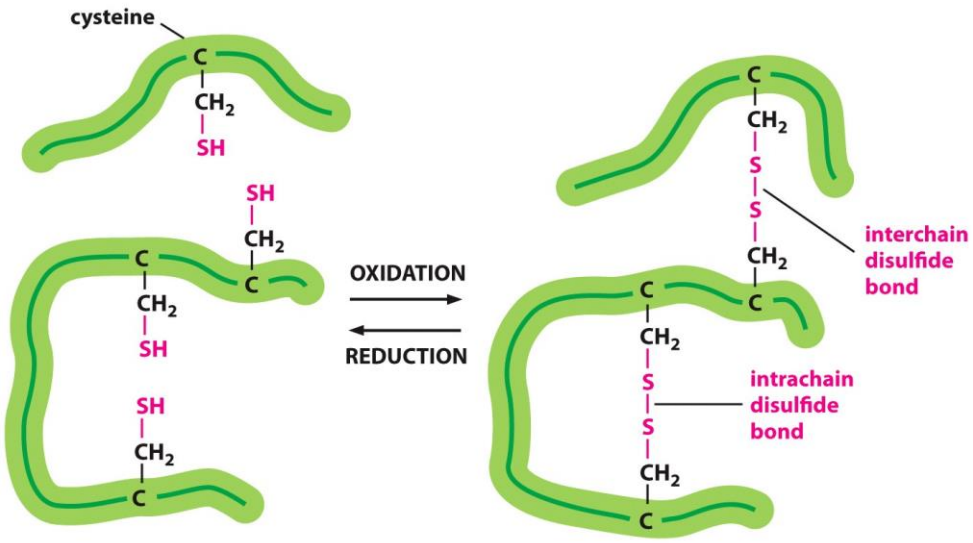


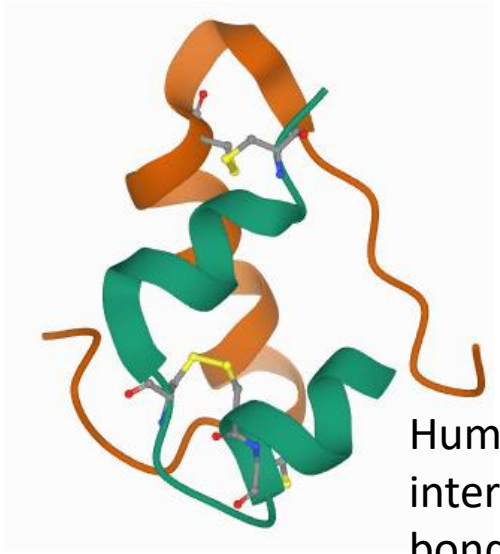
TABLE 3.1 **How Amino Acids Interact with Water**

20 amino acids are ranked according to how likely they are to interact with water. Color codes are based on Figure 3.3.

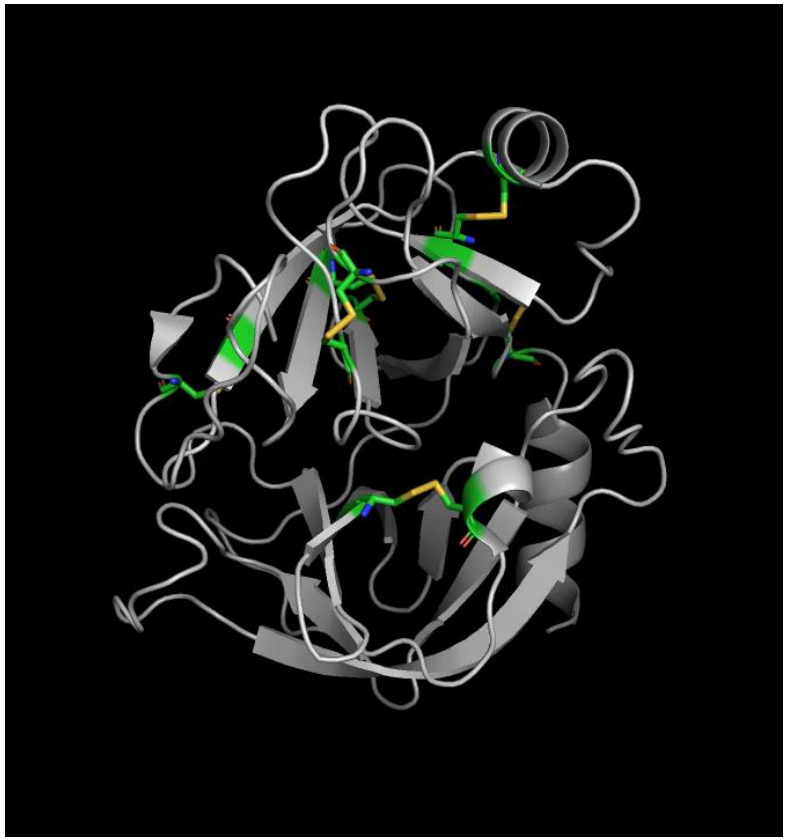
Isoleucine	Highly hydrophobic
Valine	
Leucine	
Phenylalanine	
Methionine	Moderately hydrophobic
Alanine	
Glycine	
Cysteine	
Tryptophan	
Tyrosine	Mildly hydrophobic
Proline	
Threonine	
Serine	
Histidine	Mildly hydrophilic
Glutamate	
Asparagine	
Glutamine	
Aspartate	
Lysine	
Arginine	Highly hydrophilic

Disulfide Bonds Stabilize Some Proteins Outside the Cell (and body)





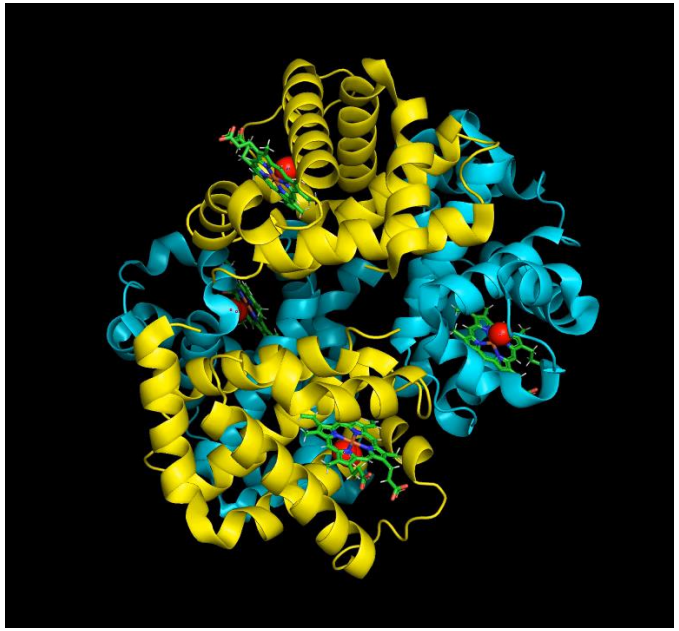
Human Insulin – interchain disulfide bonds



Trypsin – a digestive enzyme produced in the pancreas, exported to the small intestine – disulfide bonds within a single chain.

Quaternary Structure

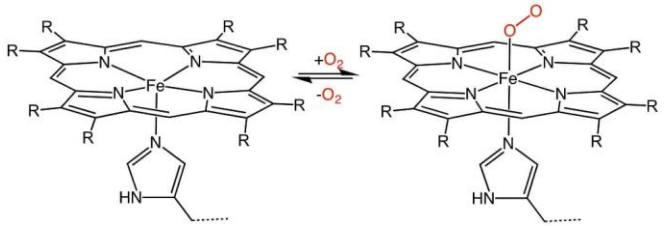
- Combinations of polypeptide subunits (combinations of tertiary structures).
- May be held together by covalent bonds, but usually non-covalent interactions between amino acids on the different chains.
- Proteins can be a dimer, a tetramer, etc.
- If the chains are the same, called homo_____. If chains are different, hetero_____



Quaternary structure of hemoglobin (oxygen transport protein):

- two α chains
- two β chains

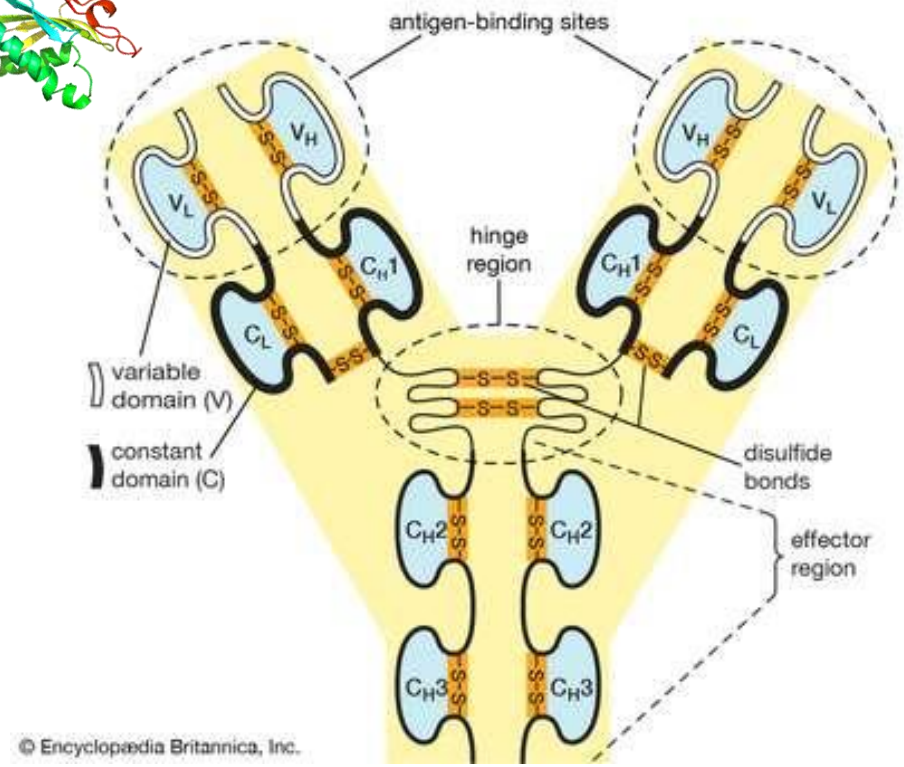
Oxygen is carried on Fe^{2+} within heme groups:

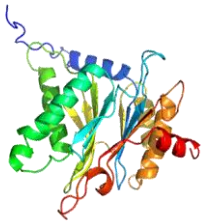


Antibodies – Produced by the Adaptive Immune system to Fight Pathogens.

Antigen

Antigen = ligand that binds to an antibody.





Antigen

Properties of Antibodies:

- 4 chains – two identical light (200 aa), two identical heavy (400 aa).
- Bind two identical antigens (pathogens, toxins)
- Chains crosslinked with disulfide bonds, increasing stability.

Protein Stability:

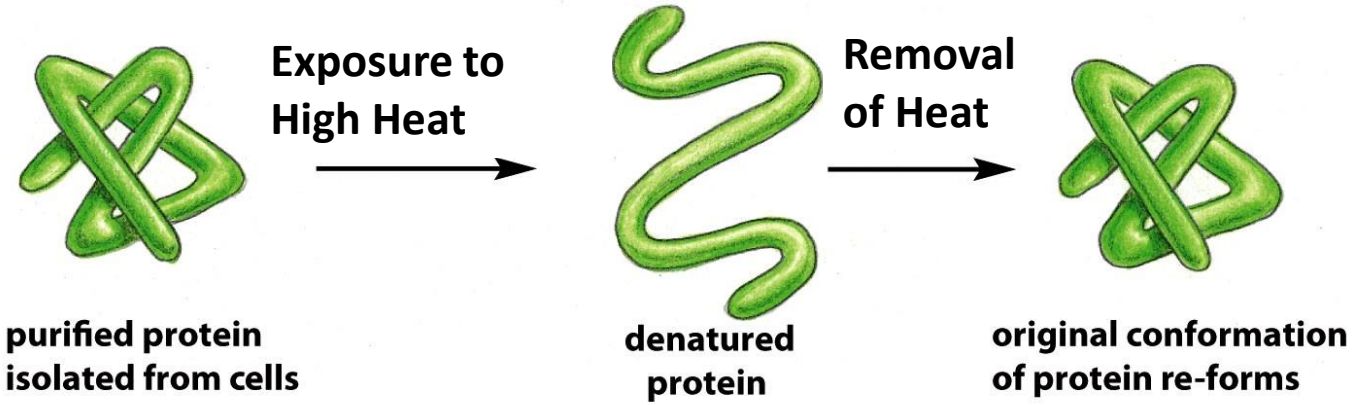
H-bonds
van der Waals
Hydrophobic effect

Native

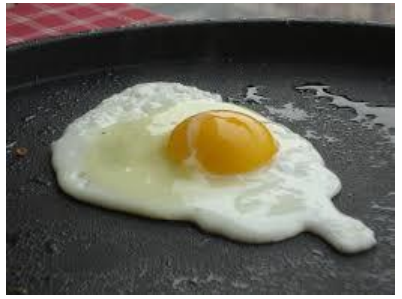
Unfolded

Chain disorder

Protein Denaturation

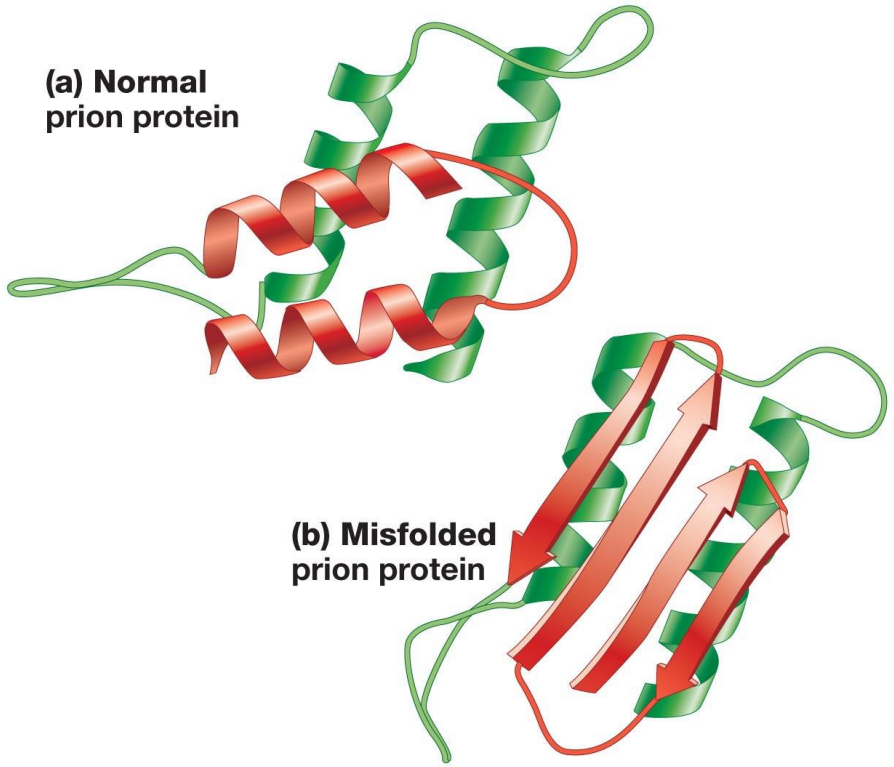


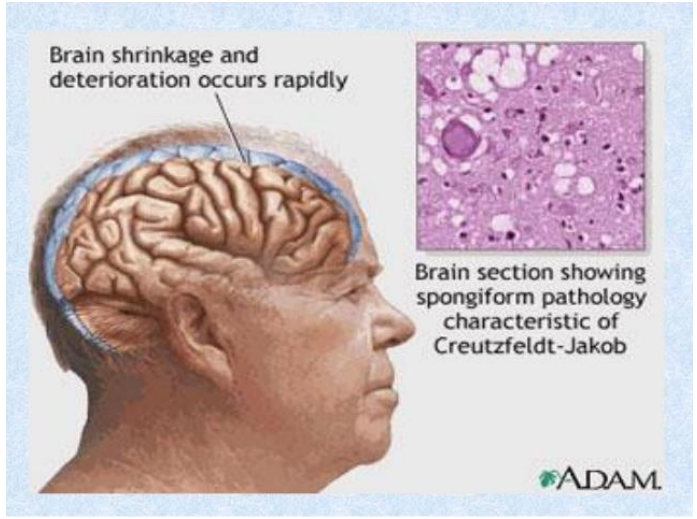
- Often, unfolded protein aggregate, which prevents refolding.

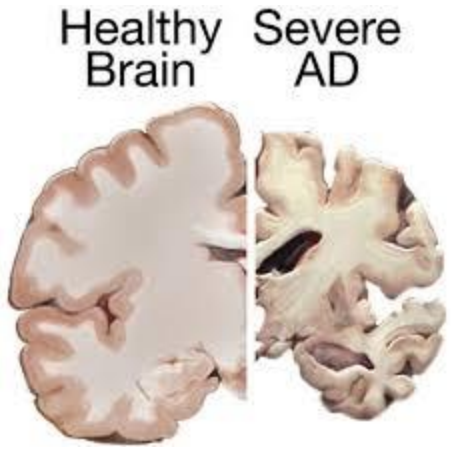


What happens when proteins don’t fold properly?

Prions are improperly folded proteins that cause neurodegenerative diseases







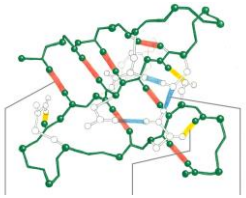
Unfolded protein response (UPR): The presence of unfolded proteins can trigger the unfolded protein response which can turn off protein synthesis in the cell, leading to cell death.

Summary of Interactions that stabilize folded proteins.

- **Hydrogen bonds** form between hydrogen atoms (NH) and the carbonyl group in the peptide-bonded backbone (mainchain), and between and donors and acceptors on sidechains and between sidechains and mainchain. *These are the most directional of all forces in biochemistry. Mainchain-mainchain H-bonds are responsible for secondary structures.*
- **van der Waals interactions** are weak interactions between side chains, *optimized in the **well packed core** of the protein*. They involve temporary *partial* charges.
- **Hydrophobic interactions** within a protein increase stability of the folded state by *increasing entropy due to the release of ordered water from the previously exposed non-polar groups. They are responsible for a **non-polar core***
- **Covalent disulfide bonds** form between sulfur-containing cysteine residues (usually only exported, secreted proteins), stabilizing them.

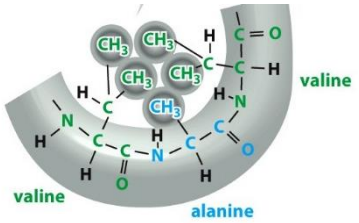
H-bonds
van der Waals
Hydrophobic effect

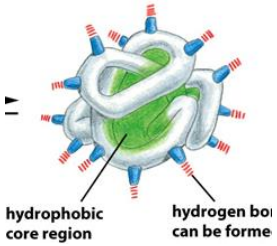
Chain disorder

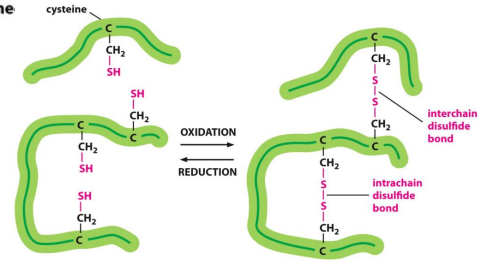


Native

Unfolded







Protein Structure - Summary and Expectations

Primary Structure:

- Can you describe the mechanism of peptide bond formation
- Can you draw structure of peptides.
- Can you identify amino terminus and give the sequence of amino acids, N -> C

Secondary structure:

- Identify helical and sheet secondary structures,
- know that they are stabilized by **mainchain** hydrogen bonds between N-H and O=C.
- Location of H-bonds and sidechains

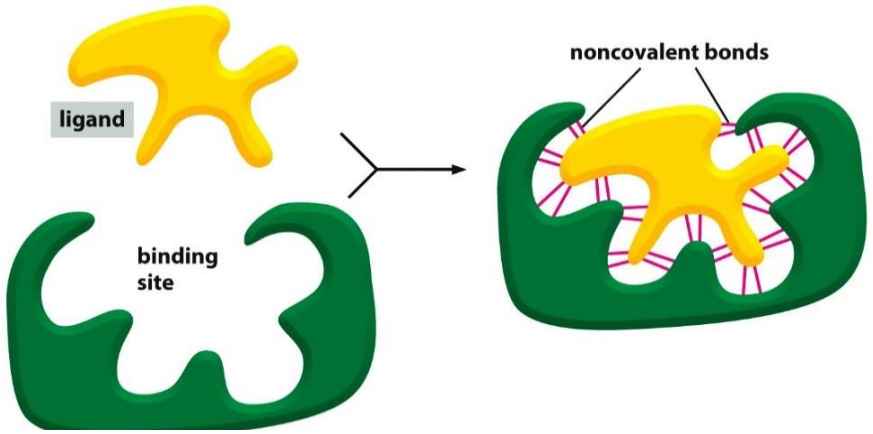
Tertiary Structure:

- Can you describe and identify role of the following in stabilizing the folded state.
  - H-bonds,
  - van der Waals,
  - hydrophobic effect
- Can you predict, based on sidechain, which amino acids are found in the core of the protein and which are found on the surface.

Quaternary Structure:

- Multiple chains, stabilized by non-covalent and covalent (disulfide bonds) interactions.

Ligand Binding: Most Proteins Bind to Other Molecules In Biological Important Interactions:



Ligand: Something that binds to a protein, usually small molecules (e.g. cyclicAMP, cAMP).

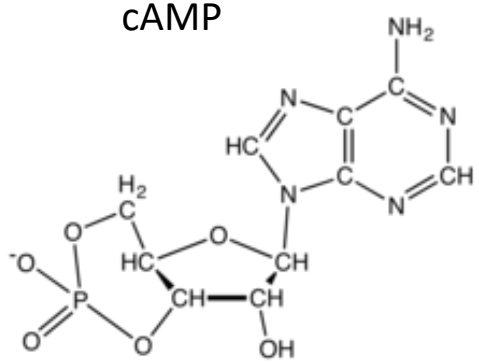
Binding site allow a protein to interact with specific **ligands**

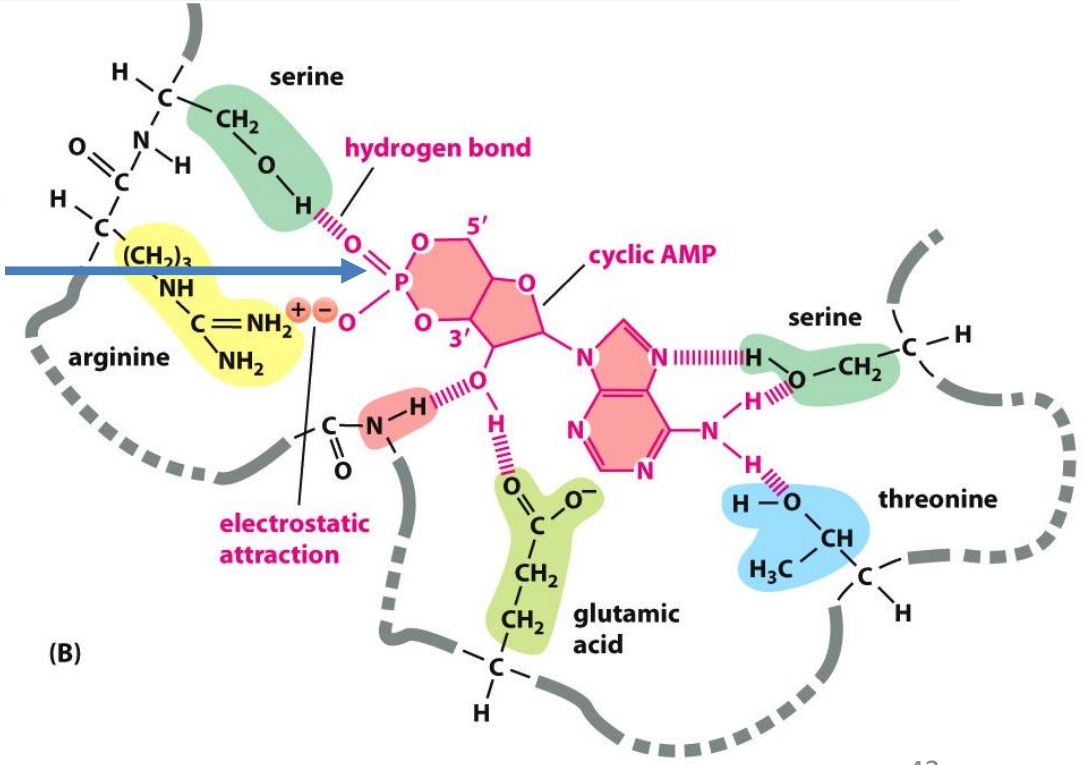
Binding site is generated by the **folded** form of the protein.

The bound ligand can be stabilized by any of:

Interaction	Which stabilize cAMP Binding?
van der Waals	
H-Bonding	
Electrostatic	
Hydrophobic effect	

cAMP





Ligand Binding & Saturation:

Define fraction saturated: $Y = \frac{[ML]}{[M] + [ML]}$

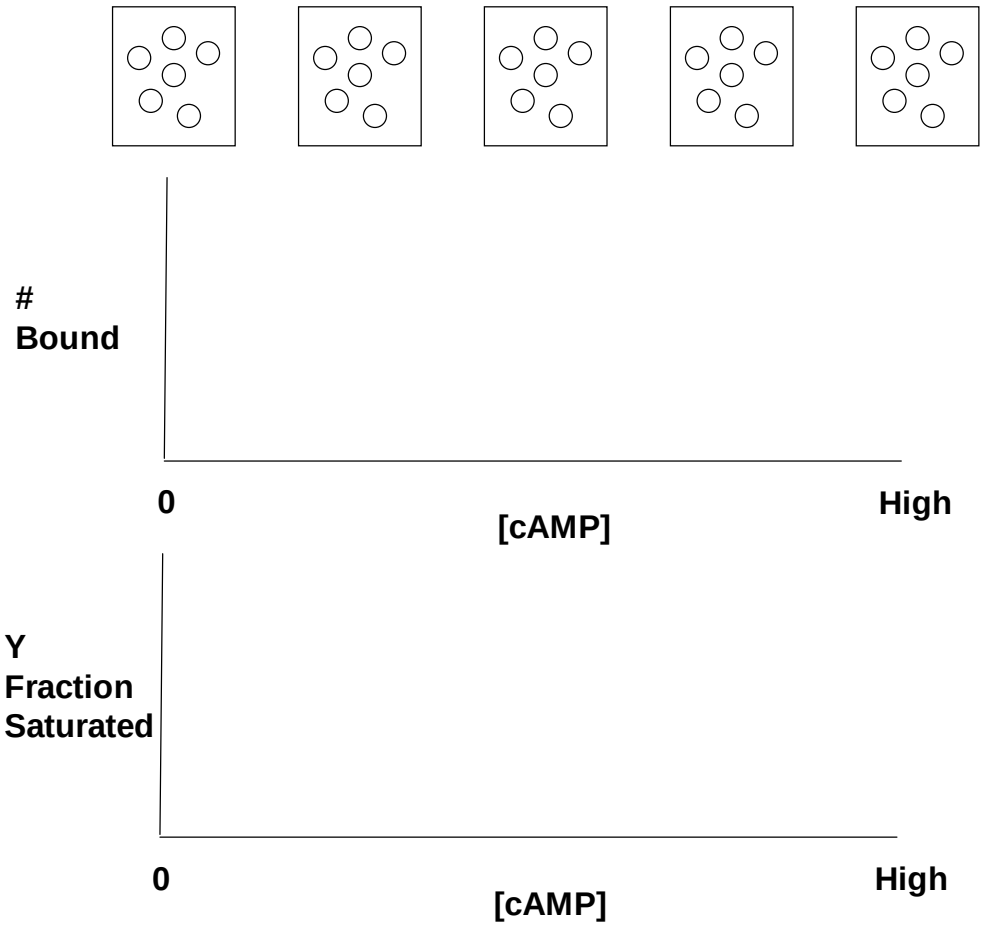
[M] = free macromolecule (e.g. antibody with no antigen).

[ML] = macromolecule with ligand bound (e.g. antibody with antigen bound).

The boxes with circles represent proteins with no cAMP bound, each box (left to right) is at a higher [cAMP]. Filled circles indicate bound ligand.

1. What would you expect to see for a graph of # bound versus [L] (complete the upper graph). Why did you draw it this way?

2. How would you modify the upper plot to obtain the lower plot?



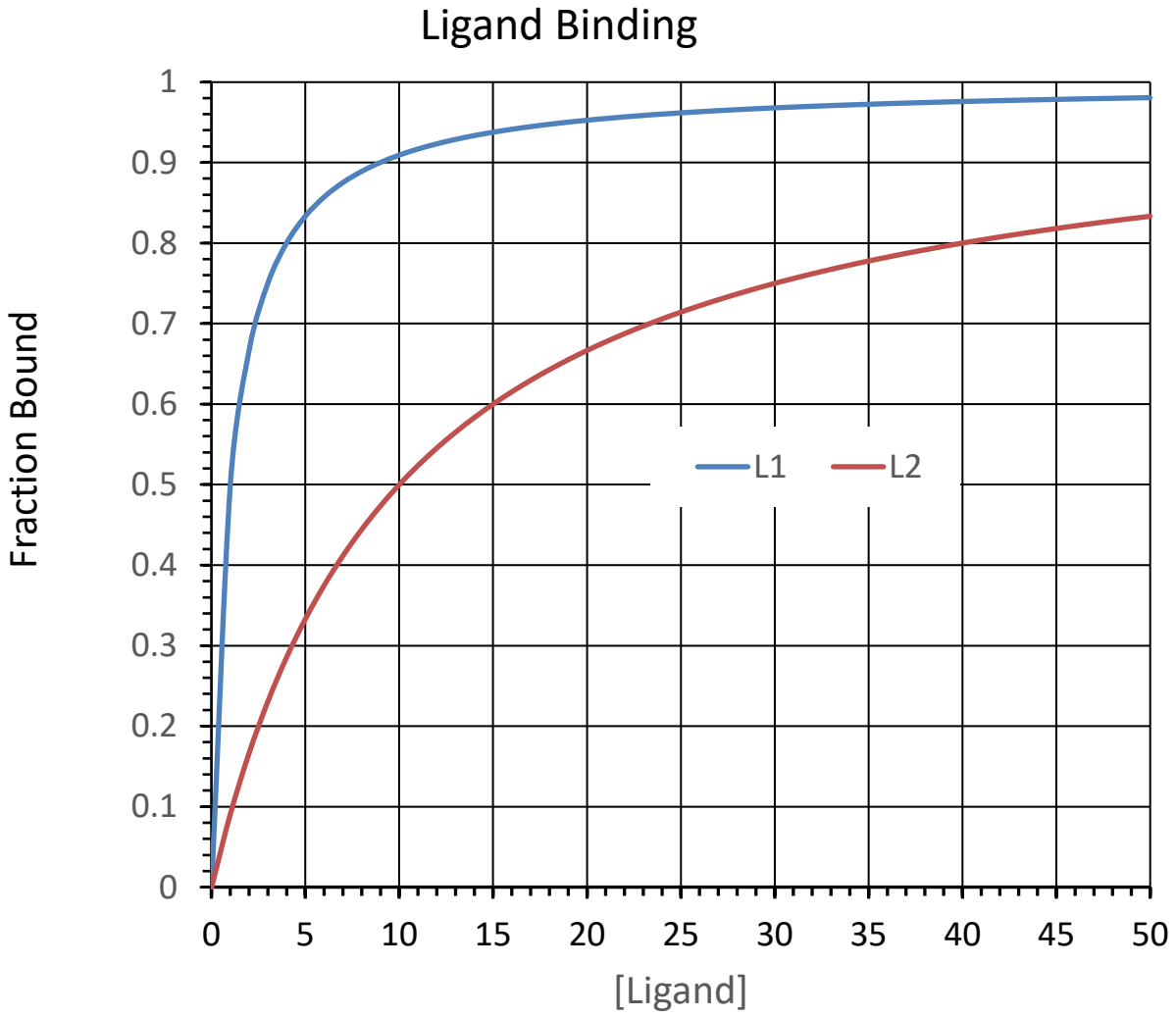
Key Points:

1. The binding sites saturate, when all are full no more ligand can bind.
2. There is a ligand concentration, [L], where ½ the sites are full. This [L] is K_D.
3. K_D is the equilibrium constant for ligand dissociation:

$$(ML) \rightleftharpoons (M) + (L)$$

$$K_D = \frac{[M][L]}{[ML]}$$

Using K_D to Compare Ligand Binding



The binding of two different molecules to the same protein was measured and the data is shown on the right. L1 is cAMP, L2 is similar to cAMP

Which ligand has a K_D of 1? L1 or L2?

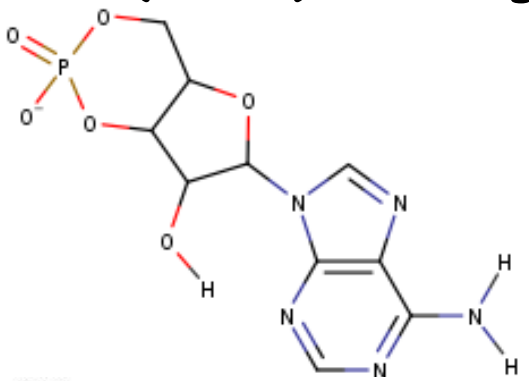
Which ligand has a K_D of 10? L1 or L2?

Which ligand binds more tightly to the protein (higher affinity)? L1 or L2?

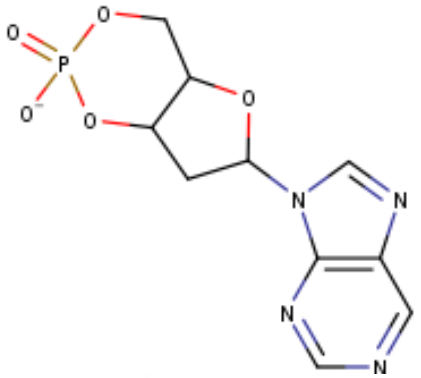
Why does L1 bind more tightly (higher affinity)?

1. What are the chemical differences between L1 and L2 (Upper diagram)

Ligand 1 (cAMP)

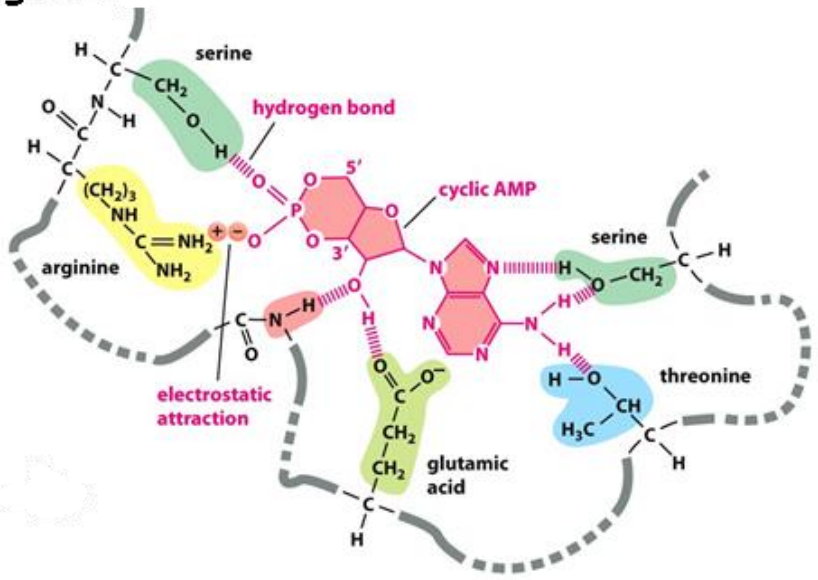


Ligand 2

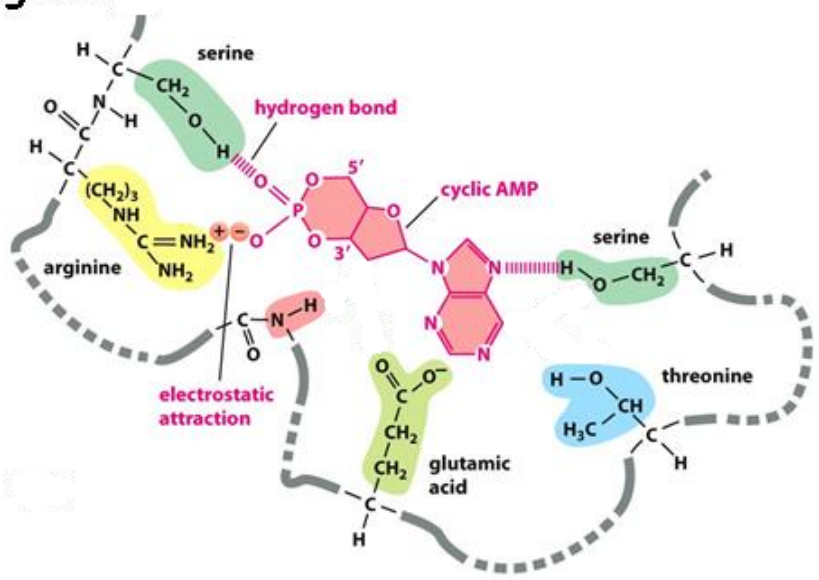


2. How do these differences affect the interactions with the protein (lower diagram)?

Ligand 1



Ligand 2



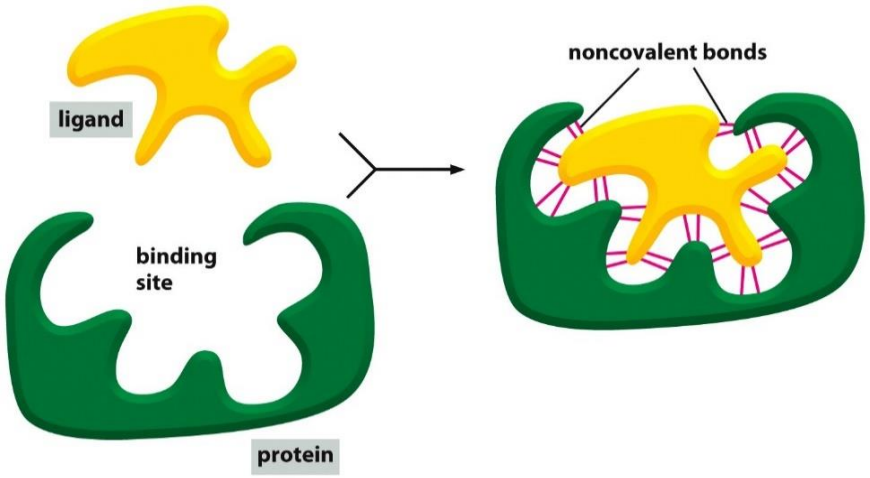
3. How do the differences affect K_D?

Key Points:

Binding:

Folded proteins have ***binding sites*** that recognize other molecules (***ligands***) using ***any and all*** of the following:

- H-bonds, van der Waals, Electrostatic, Non-polar interactions (hydrophobic)



Binding is **reversible**

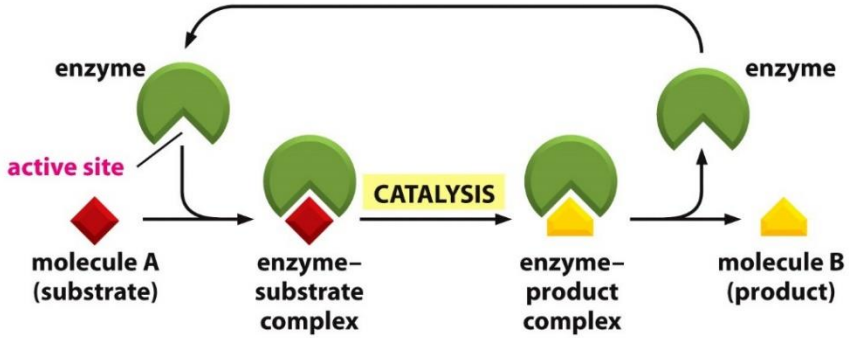
Binding is **saturable**

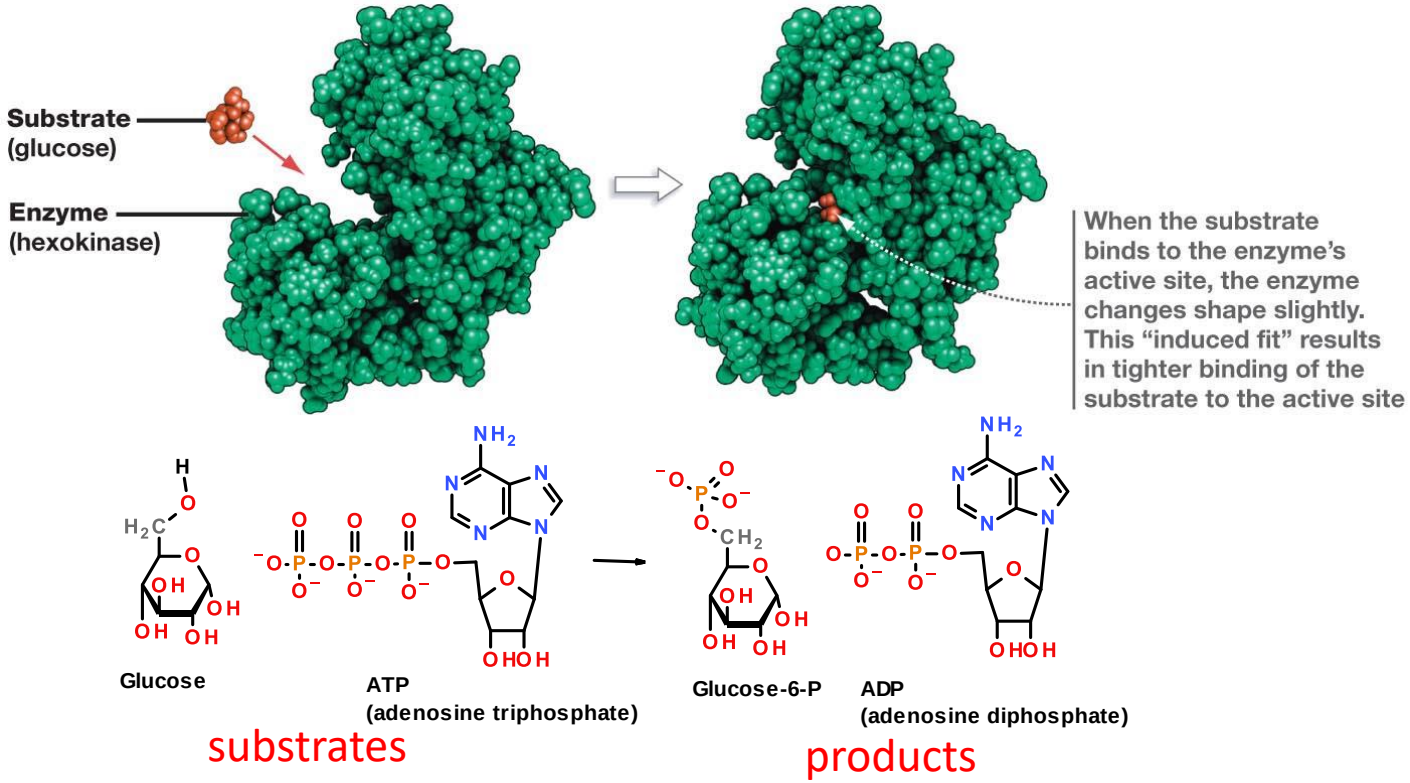
Binding ½ point (Y=0.5) occurs at K_D

The higher the affinity (strength of interaction), the lower the K_D

Enzymes

- **Enzymes** are protein or RNA catalysts. They increase the rate of the reaction.
- They bind "substrates" and convert them to "products". Usually, the substrate undergoes a chemical reaction and is changed in its structure.
- Most biological chemical reactions occur at meaningful rates only in the presence of an enzyme.
- Substrates bind specifically to the enzyme's **active site,** interacting with amino acid side chains (or RNA bases). Usually a single enzyme binds one substrate.
- The chemical change caused by the enzyme is catalyzed by additional functional groups in the active site.
- Many enzymes undergo a conformational change when the substrates are bound to the active site; this change is called an **induced fit**.





Enzyme – Chemical Diversity

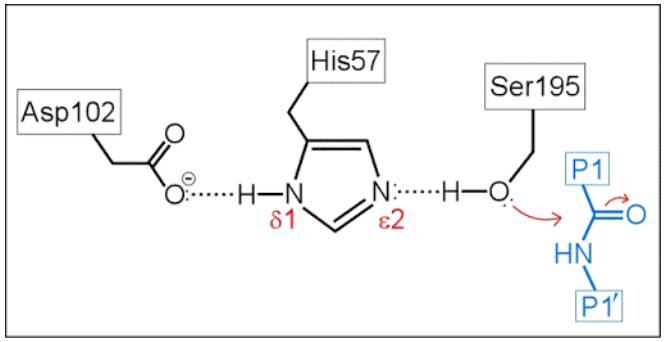
TABLE 4–1 SOME COMMON FUNCTIONAL CLASSES OF ENZYMES

ENZYME CLASS	BIOCHEMICAL FUNCTION
Hydrolase	General term for enzymes that catalyze a hydrolytic cleavage reaction.
Nuclease	Breaks down nucleic acids by hydrolyzing bonds between nucleotides.
Protease	Breaks down proteins by hydrolyzing peptide bonds between amino acids.
Synthase	General name used for enzymes that synthesize molecules in anabolic reactions by condensing two molecules together.
Isomerase	Catalyzes the rearrangement of bonds within a single molecule.
Polymerase	Catalyzes polymerization reactions such as the synthesis of DNA and RNA.
Kinase	Catalyzes the addition of phosphate groups to molecules. Protein kinases are an important group of kinases that attach phosphate groups to proteins.
Phosphatase	Catalyzes the hydrolytic removal of a phosphate group from a molecule.
Oxido-reductase	General name for enzymes that catalyze reactions in which one molecule is oxidized while the other is reduced. Enzymes of this type are often called oxidases, reductases, or dehydrogenases.
ATPase	Hydrolyzes ATP. Many proteins with a wide range of roles have an energy-harnessing ATPase activity as part of their function, including motor proteins such as myosin and membrane transport proteins such as the sodium–potassium pump.

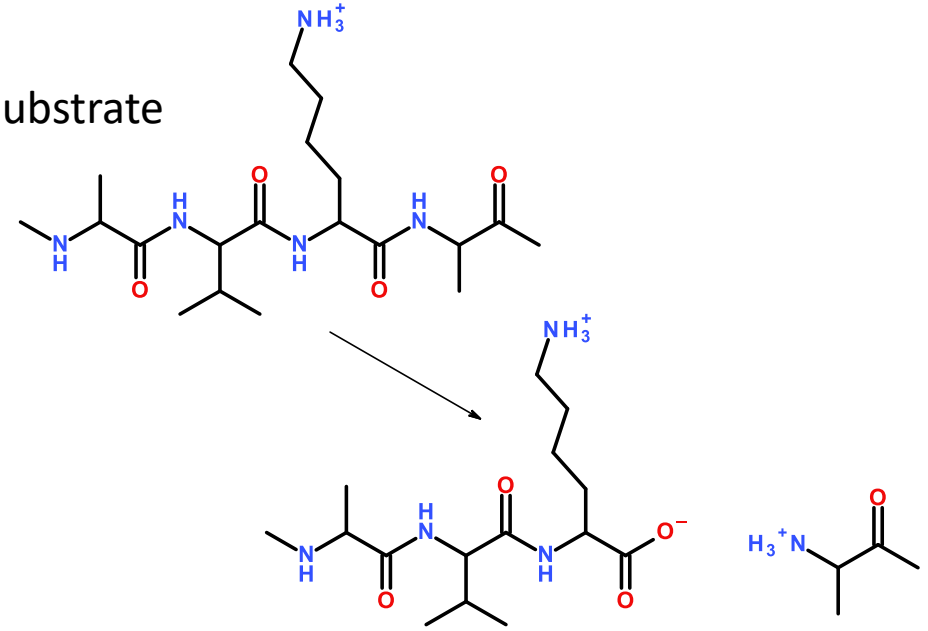
- Most enzyme names end in “-ase”
- Usually named by their substrates and the reactions they catalyse, i.e. glucose kinase

Example of Active Site Functional Groups:
Catalytic triad (Asp, His, Ser) in Protease Trypsin cleaves after Lys Residues

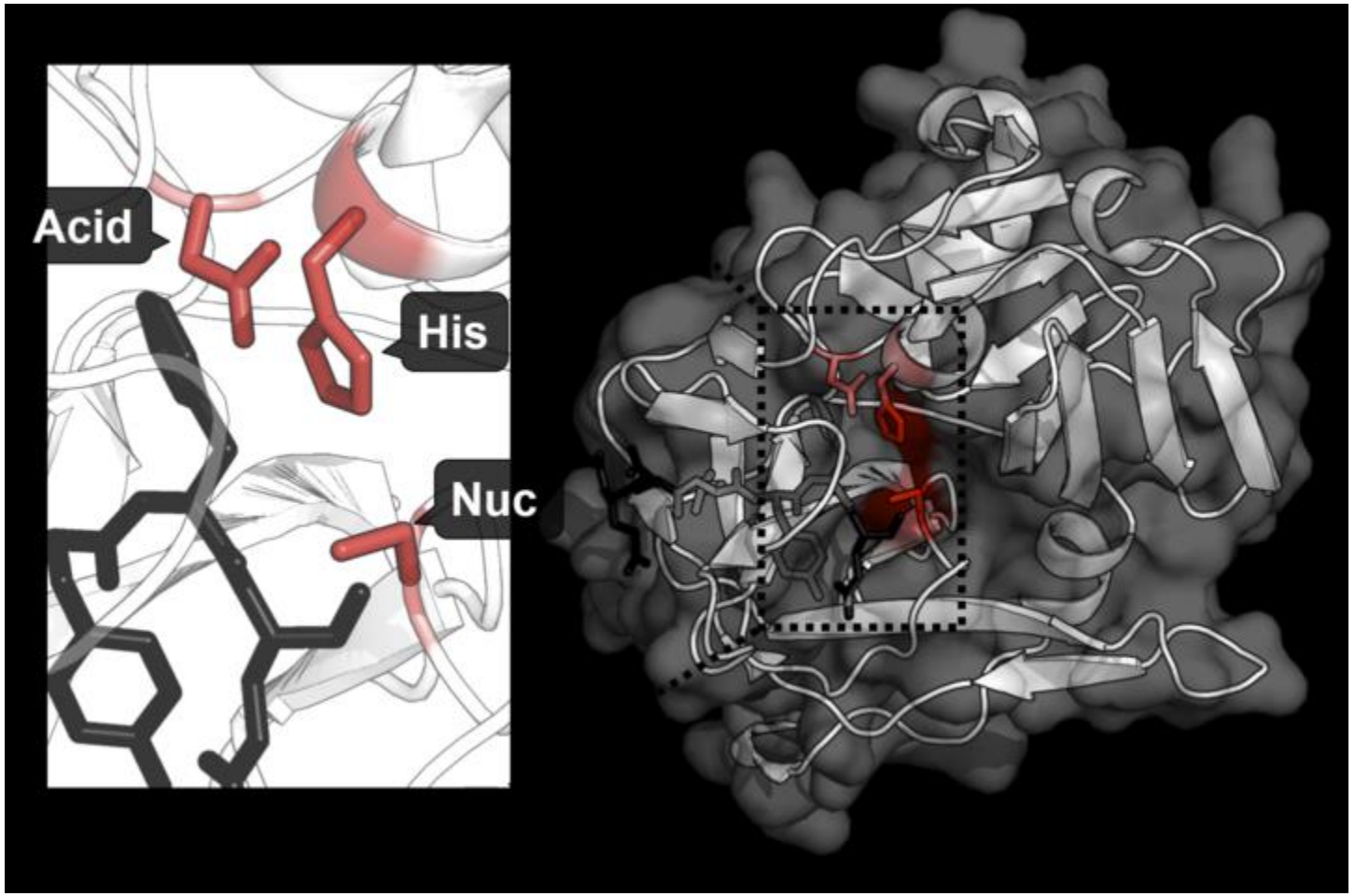
Triad



Substrate



Products

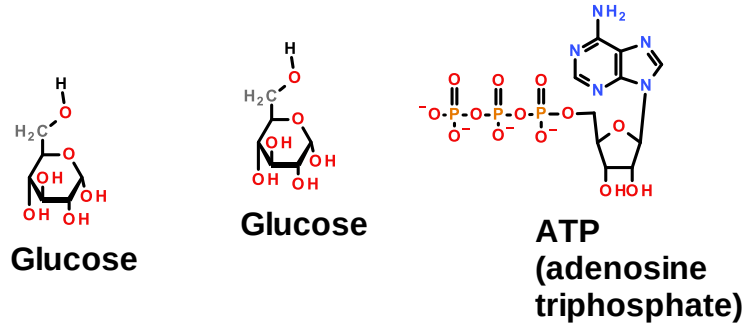


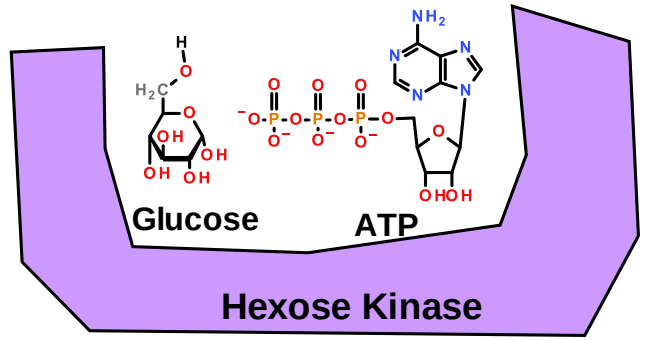
https://shirleychemproject.weebly.com/

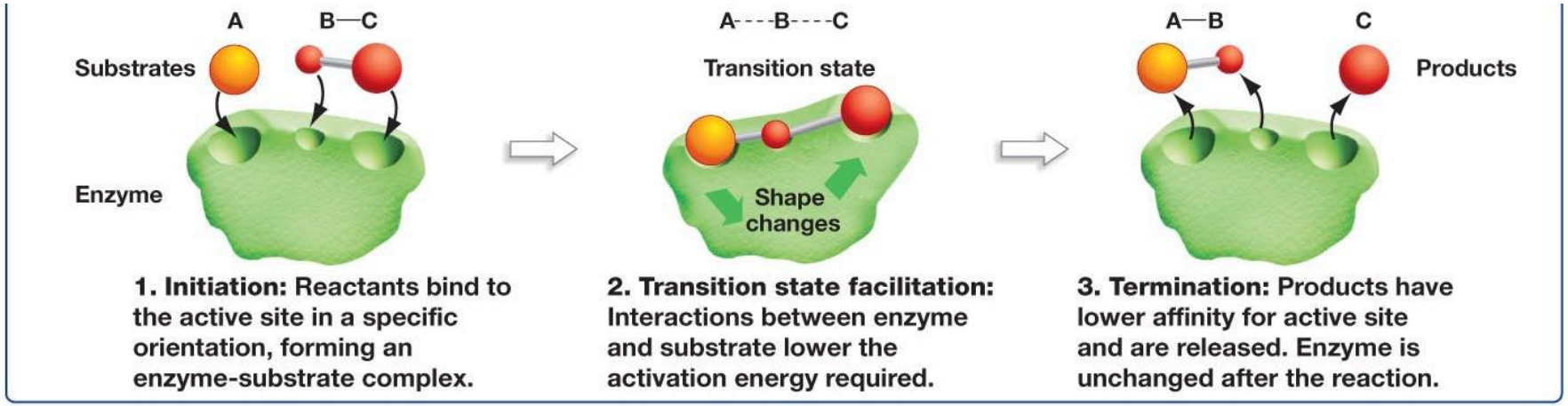
How Do Enzymes Increase Rates?

- Enzymes bring **substrates** and **functional groups on the enzyme** together in specific positions that facilitate reactions. This increases the likelihood a reaction will occur.
  - Molecules unlikely to interact in solution are more likely to react in the active site.
  - The relative orientation of the substrates is also controlled by the enzyme.

- Interactions between the enzyme and the substrate stabilize a high energy **transition state** (X), lowering the thermal energy required for the reaction to proceed.





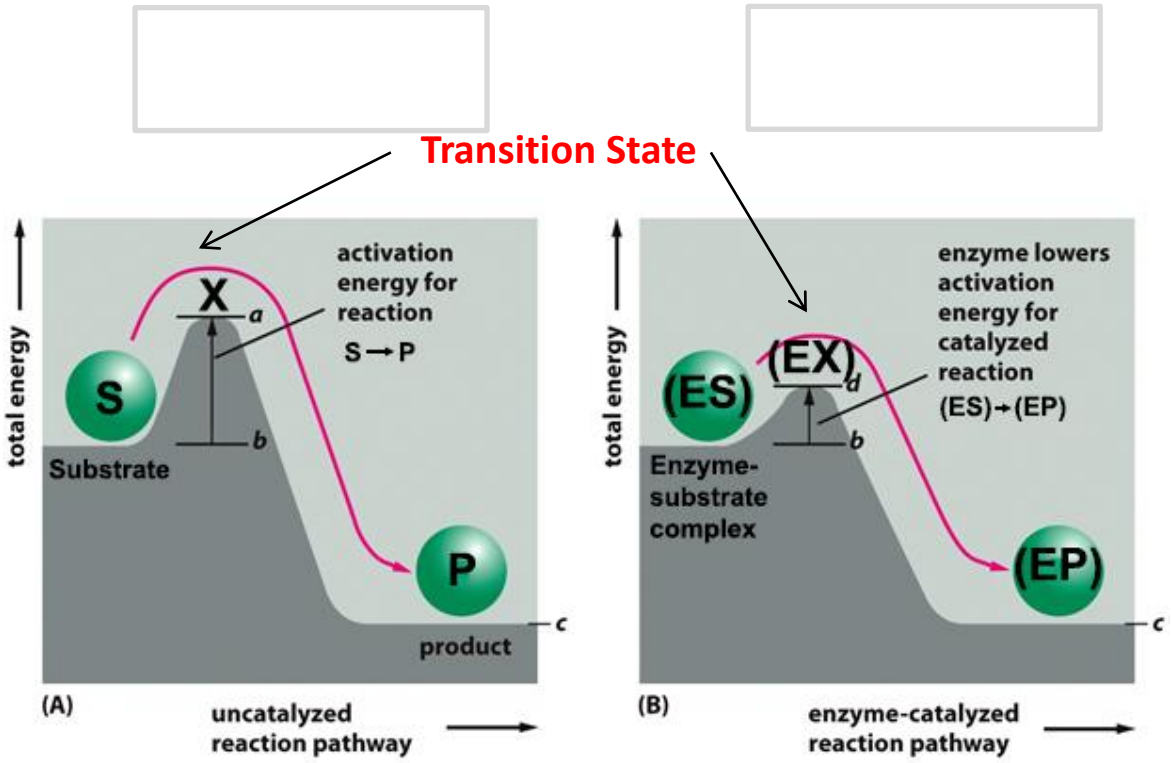


What is the transition state in the reaction on the left?

$$E + S \rightleftharpoons (ES) \longrightarrow (EX) \longrightarrow (EP) \longrightarrow E + P$$

What does the reaction rate depend on?

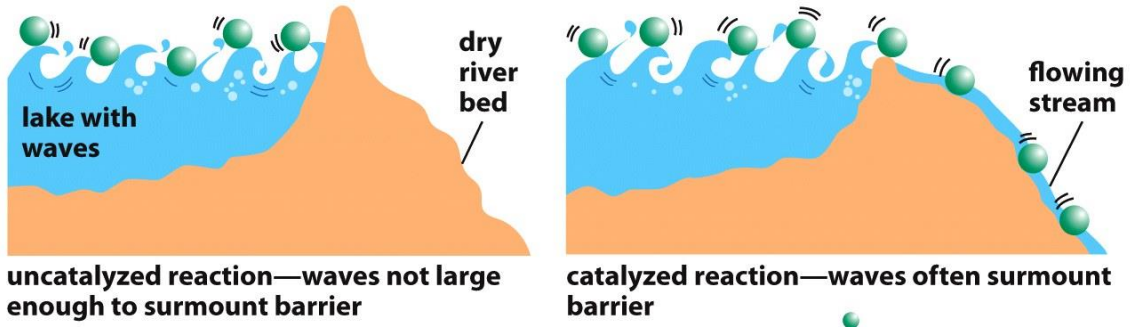
$$E + S \rightleftharpoons (ES) \rightarrow (EX) \rightarrow (EP) \rightarrow E + P$$



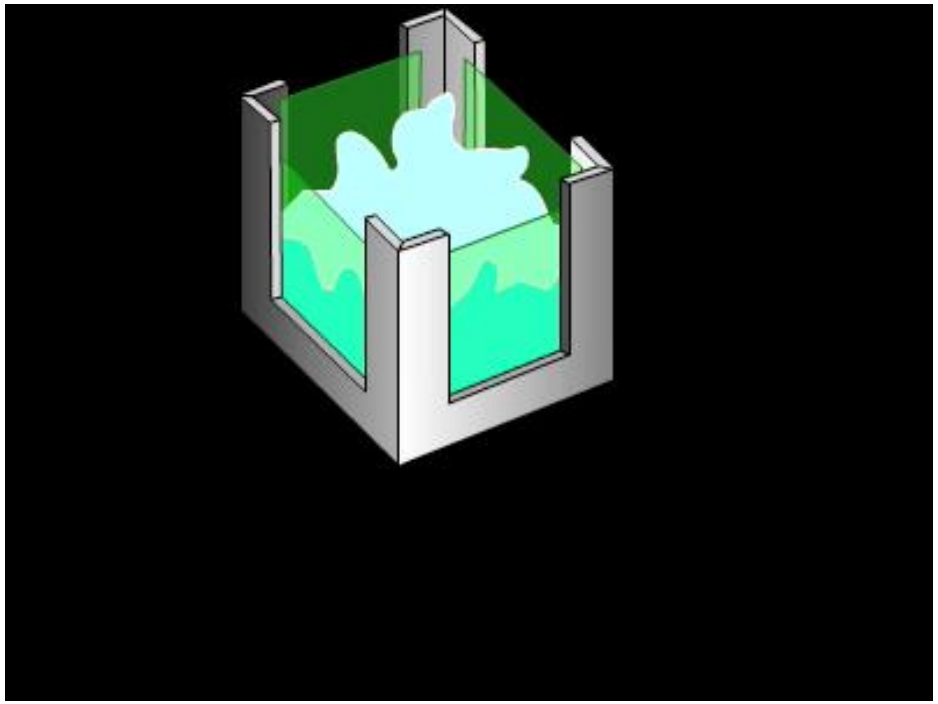
Transition State (X):

- **The transition state is high energy intermediate in the reaction going from substrate to product.** The substrate must become the transition state on the way to product.
- **The rate (how fast) the reaction goes depends on the amount of transition state.**
- In uncatalyzed reactions the energy of the transition state is high, so it is difficult to reach, [X] is low.
- **In catalyzed reactions, the energy of the transition state is low,** so it is easier to reach, and the rate is faster.

- **The energy of the transition state is lowered by favorable interactions between the active site and the transition state.**



Please watch this video and answer the following in the TA Class



Substrate = ball in upper tank
Product = ball in lower tank

How did heating make the reaction occur?

How did the enzyme make the reaction occur?

How did the enzyme control the product?

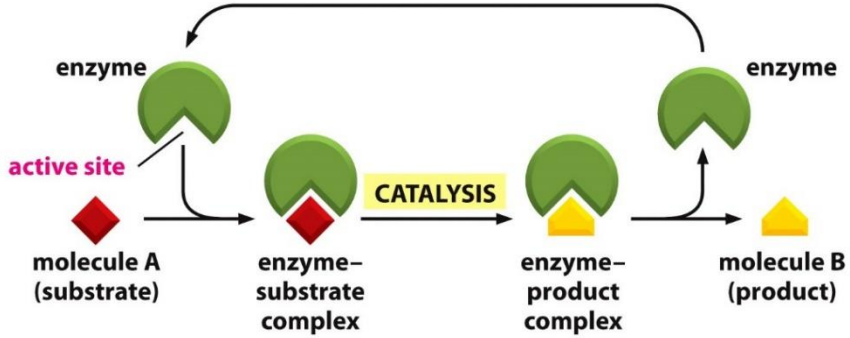
Key Points:

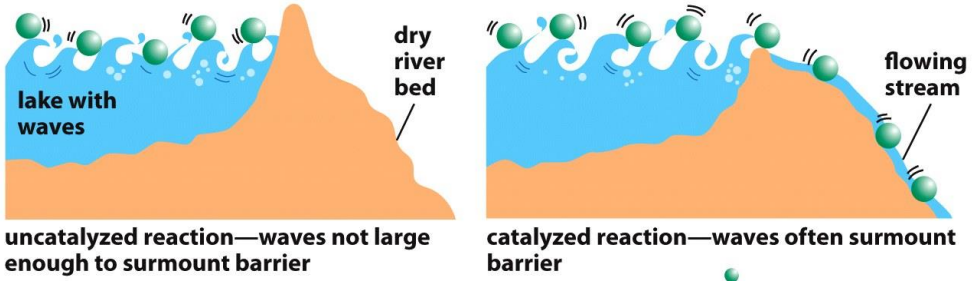
Enzymes:

Enzymes bind substrates (S), forming (ES) complex in active site, converting to P, releasing P.

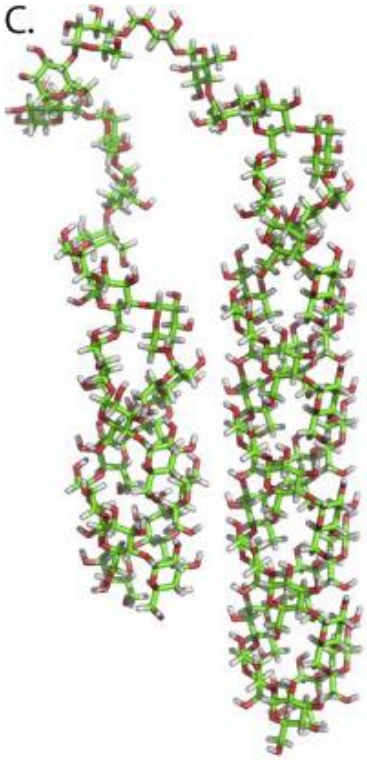
Rate enhancement due to the fact that the transition state complex (EX) forms more readily with enzymes due to:

- Bringing substrates and functional groups on the enzyme together by binding (less entropy change)
- Directly lowering energy of transition state (X) through favorable interactions that are unique to the transition state, such as forming unique hydrogen bonds.

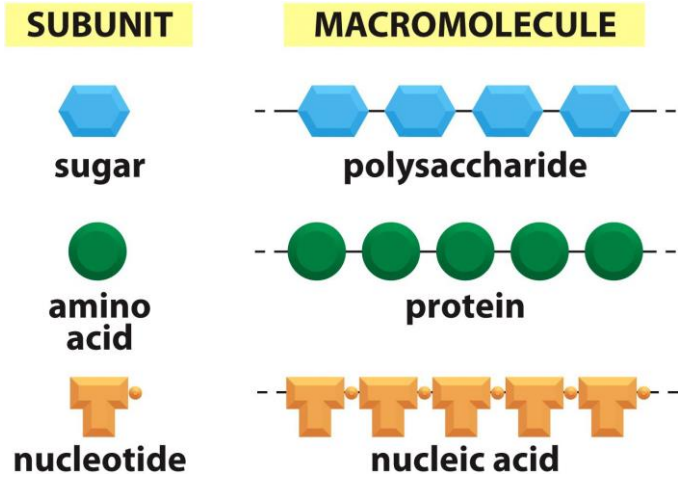


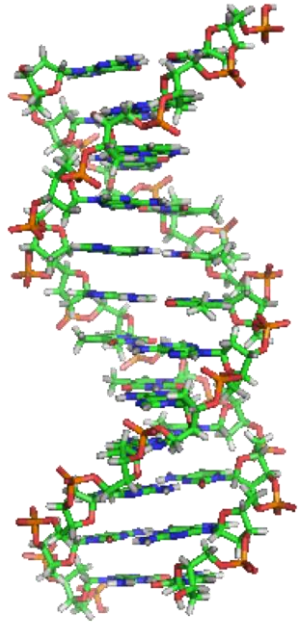


Carbohydrates



Polysaccharide



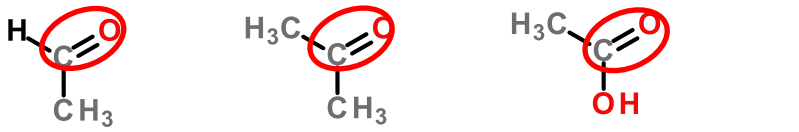


DNA (Nucleic Acid)

Carbohydrates

- Monosaccharides (one sugar),
- oligosaccharides (few sugars)
- polysaccharides (many sugars)
- Chemical formula is $(CH_2O)_n$ (e.g. hydrated carbon)
- They are molecules with:
  - one aldehyde or ketone group, on 1st or 2nd carbon
  - -OH group on *all* other carbons, leading to a chiral carbon for most carbons.

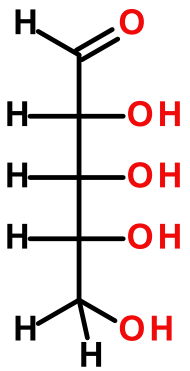
Functional groups:



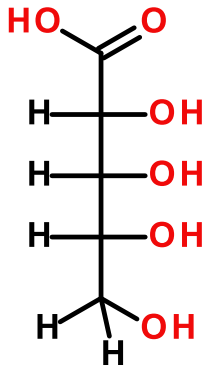
aldehyde ketone carboxylic acid

Carbonyl group –> C=O

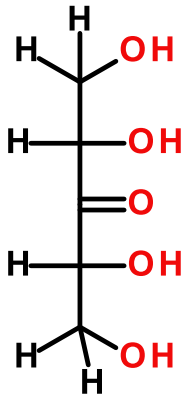
A



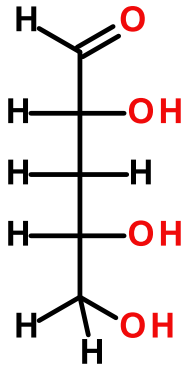
B



C



D

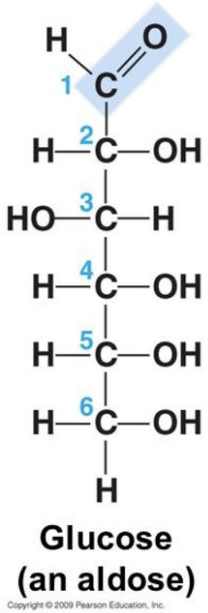


Which of the following are carbohydrates? Which are not? Why?

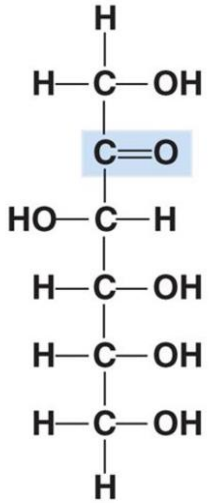
3 ways simple sugars (monosaccharides) differ from each other

1. Location of the carbonyl group
2. Number of carbons
3. Spatial arrangement of atoms (the position of the OH groups)

Carbonyl group is located on C_1



Glucose (an aldose)

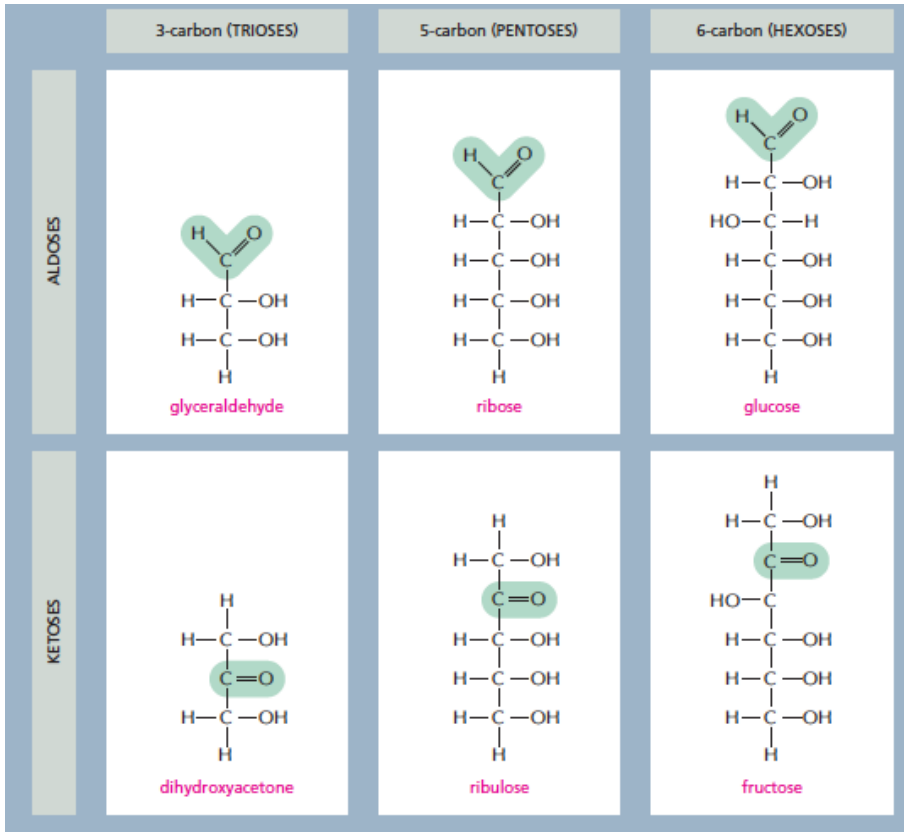


Fructose (a ketose)

Numbering carbons:
Carbon 1 is at the end closest to the C=O group.

What carbon is the carbonyl?

1. Location of the carbonyl group
2. Number of carbons
3. Spatial arrangement of atoms (the position of the OH groups)

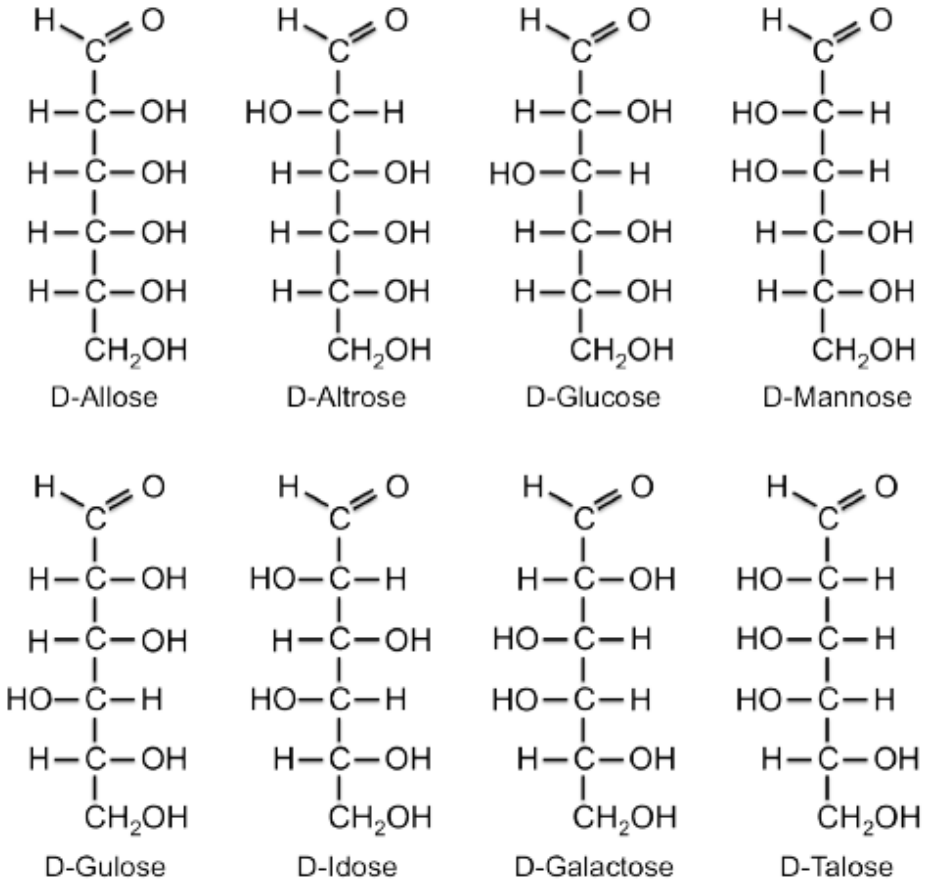


3 ways simple sugars (monosaccharides) differ from each other

1. Location of the carbonyl group
2. Number of carbons
3. **Spatial arrangement of atoms (the position of the OH groups – opposite chirality)**

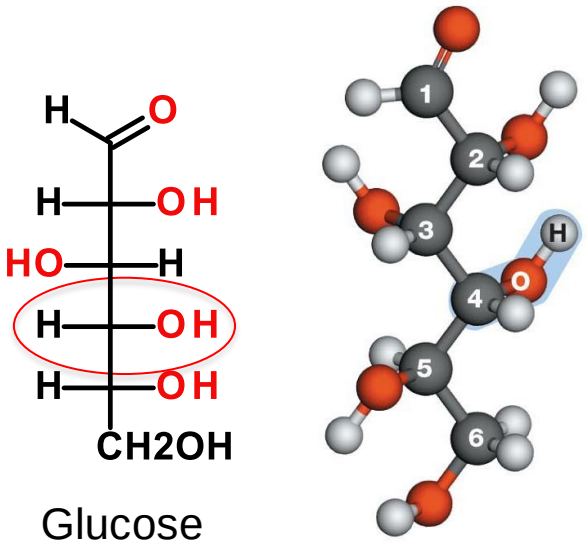
The sugars on the right are all 6-carbon aldoses.

How do they differ?

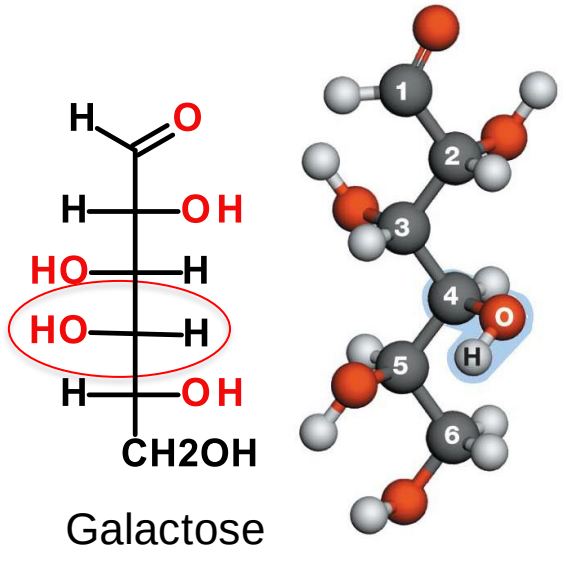


3 ways simple sugars (monosaccharides) differ from each other

1. Location of the carbonyl group
2. Number of carbons
3. **Spatial arrangement of atoms (the position of the OH groups)**



Glucose

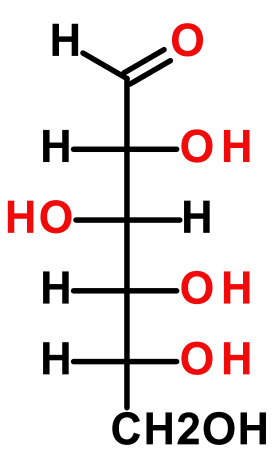


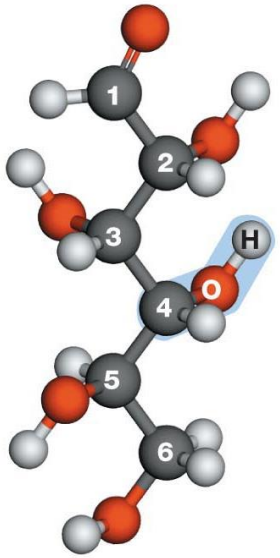
Galactose

Both have the same chemical formula $C_6H_{12}O_6$. Both are aldose sugars with 6 carbons. Yet their functions are different.

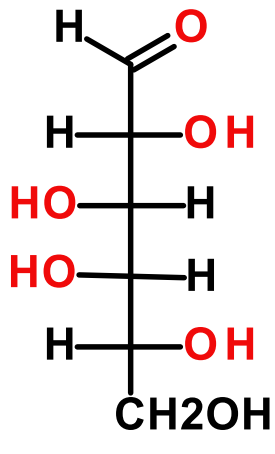
- Glucose can be used for energy immediately.
- Galactose has to be converted to glucose before it can be used for energy.

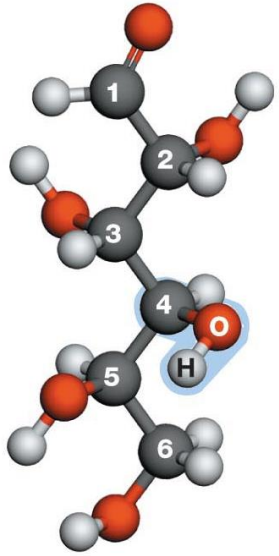
Why Glucose ≠ Galactose



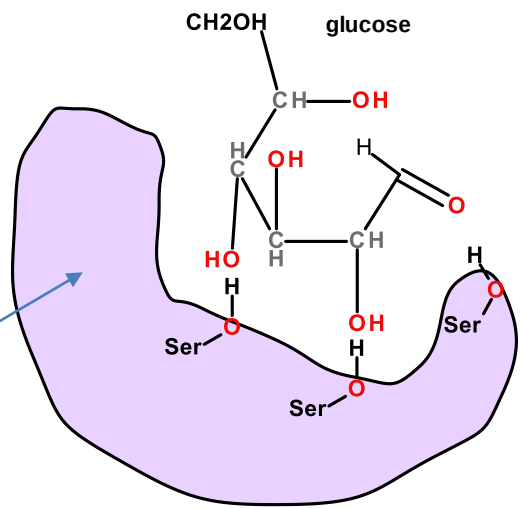


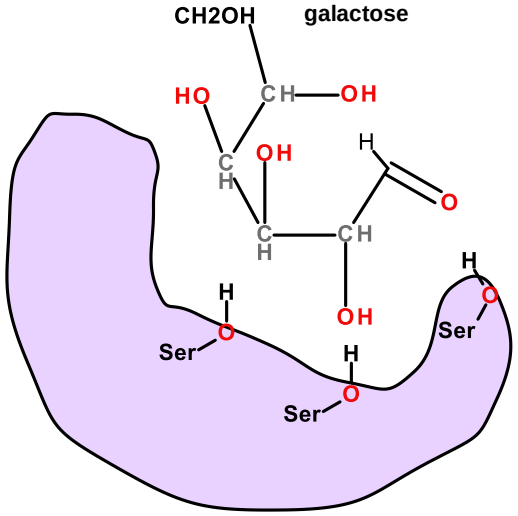
Glucose





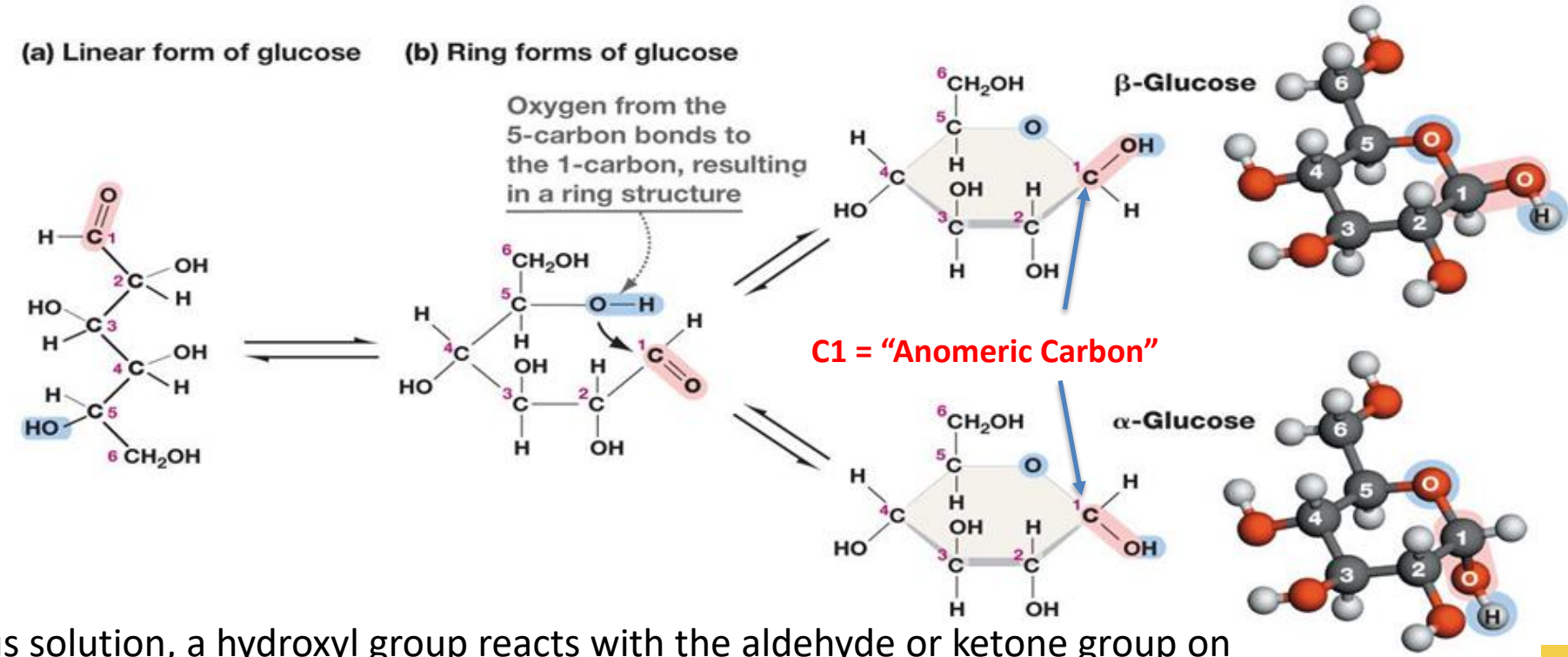
Galactose





Enzyme specific
for a-glucose

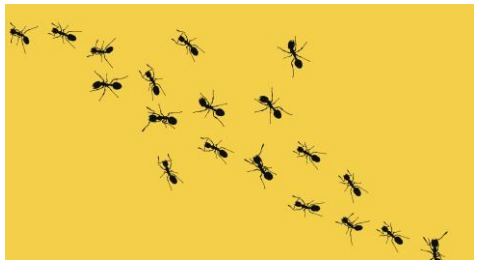
Ring formation in monosaccharides:



- In aqueous solution, a hydroxyl group reacts with the aldehyde or ketone group on the same molecule, closing the molecule into a ring, with a bridging oxygen
- Stable ring sizes are 5 atoms or 6 ato
- It is usually the 2nd to last -OH group, i.e. C5 in glucose, C4 in ribose.
- No atoms are lost or gained in this reaction.
- The carbonyl carbon becomes chiral, and is called the ***anomeric carbon*****.**
- The rings with different chirality at C1 are different:

  α (new OH is down), β (new OH is up)

  "(ants are down, birds are up)"



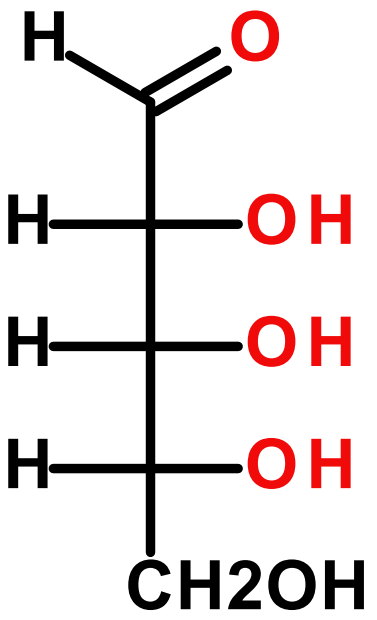
Example Problem:

The linear form of ribose, a 5 carbon aldose is shown on the right. This sugar is found in RNA (ribonucleic acid).

1. Number the carbons.

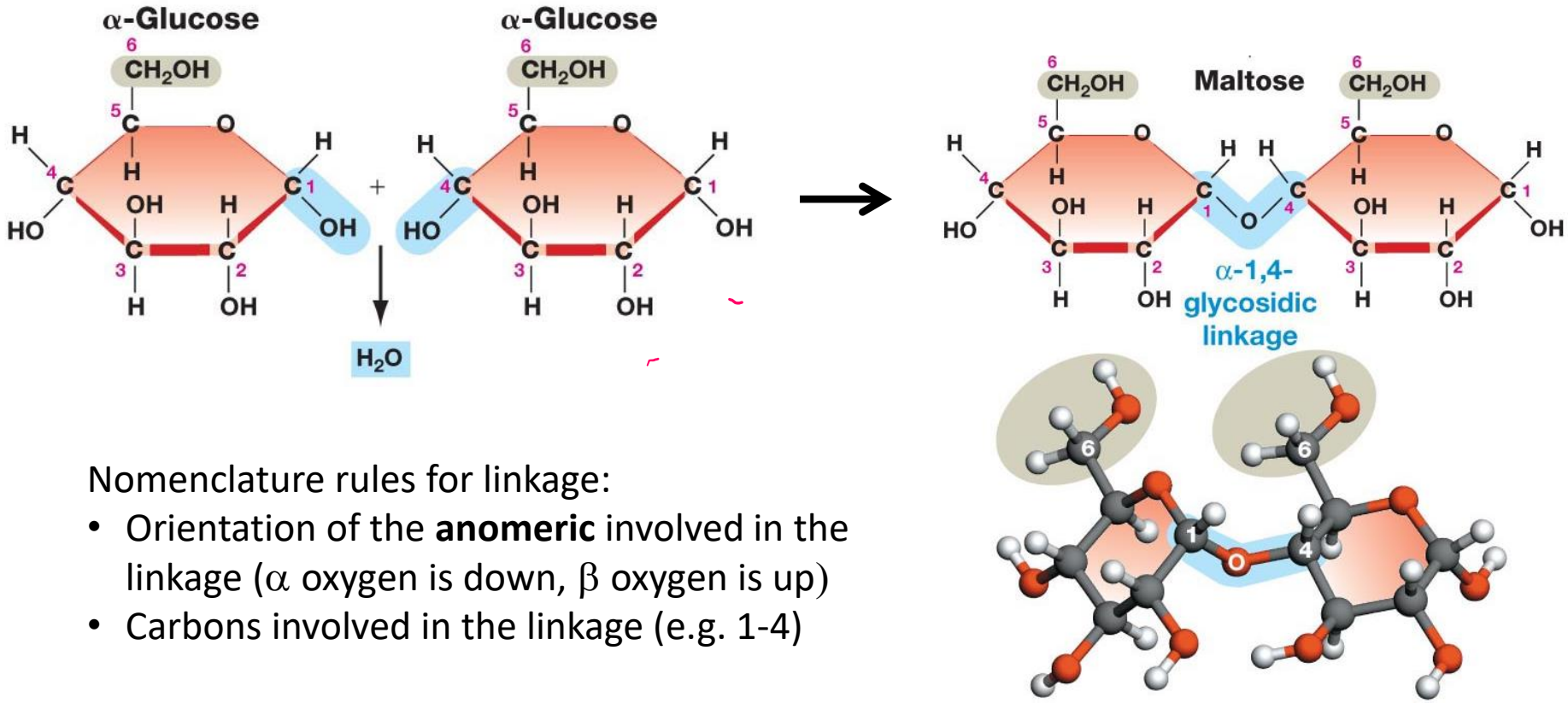
2. Which carbons are chiral? Mark them with a *.

3. Draw the cyclic form of α-ribose



Glycosidic Linkages - Disaccharides

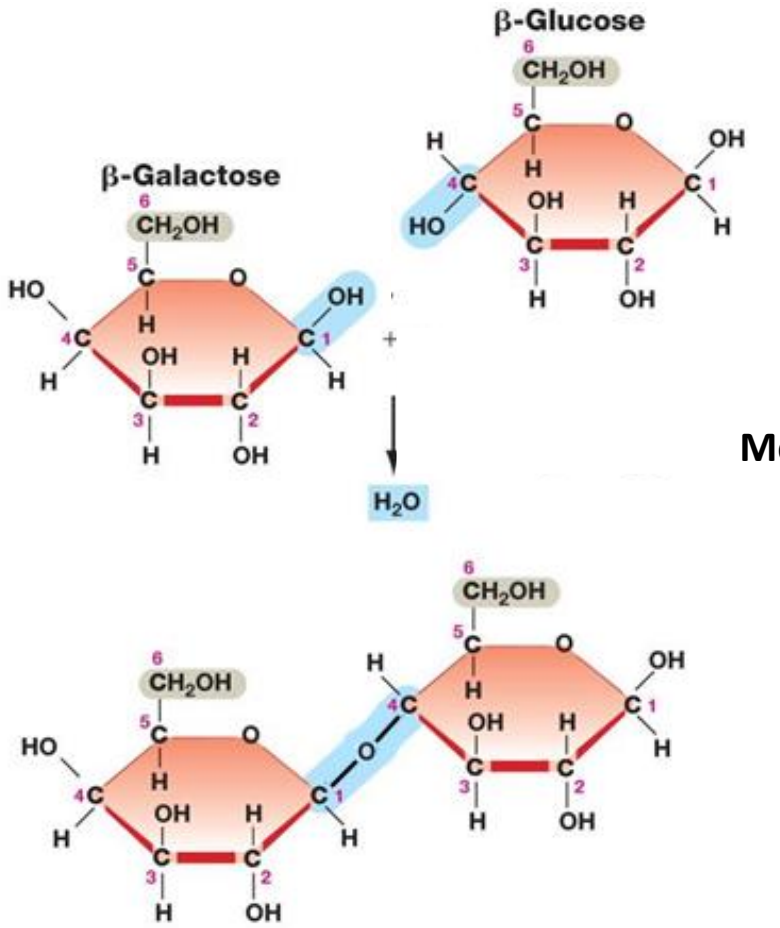
- **Glycosidic** bond formed between **any** -OH of one sugar and the anomeric carbon of another (e.g. at least one anomeric is involved.
- Water released (dehydration reaction).



Nomenclature rules for linkage:

- Orientation of the **anomeric** involved in the linkage (α oxygen is down, β oxygen is up)
- Carbons involved in the linkage (e.g. 1-4)

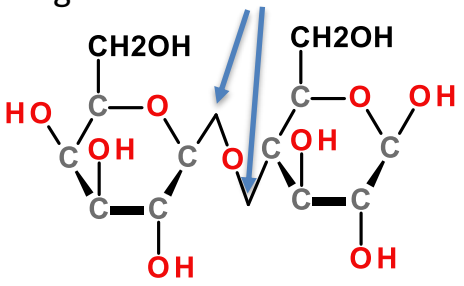
Lactose = galactose + glucose



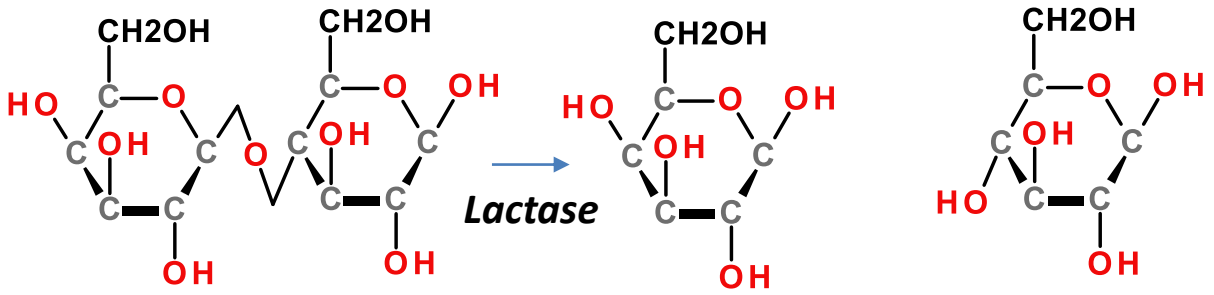
β-1,4- glycosidic linkage

Alternative drawing

These are not carbons



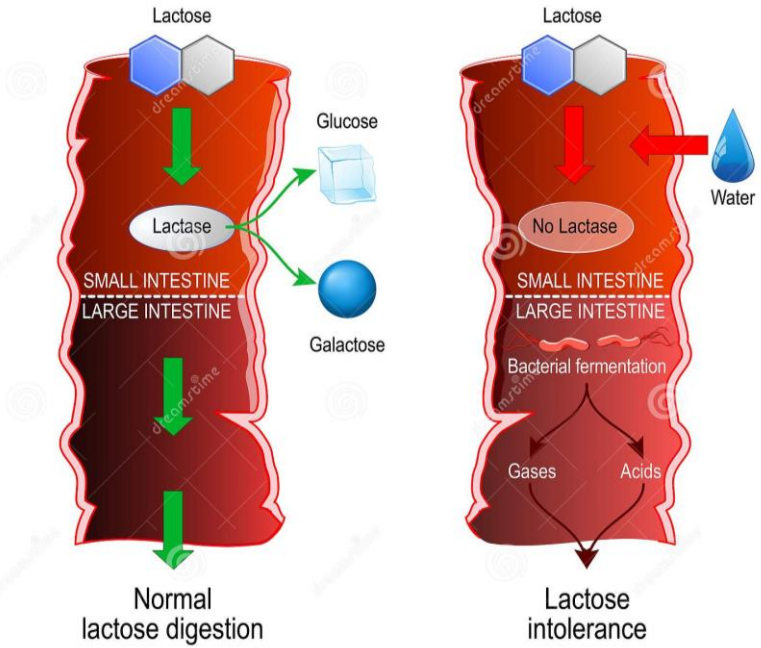
Metabolism of Lactose



Galactose

Glucose

Fun Fact - Lactose Intolerance



Most individuals with lactose maldigestion can tolerate up to 12g of lactose as a single dose with no, or minor, symptoms
The European Food Safety Authority (EFSA)

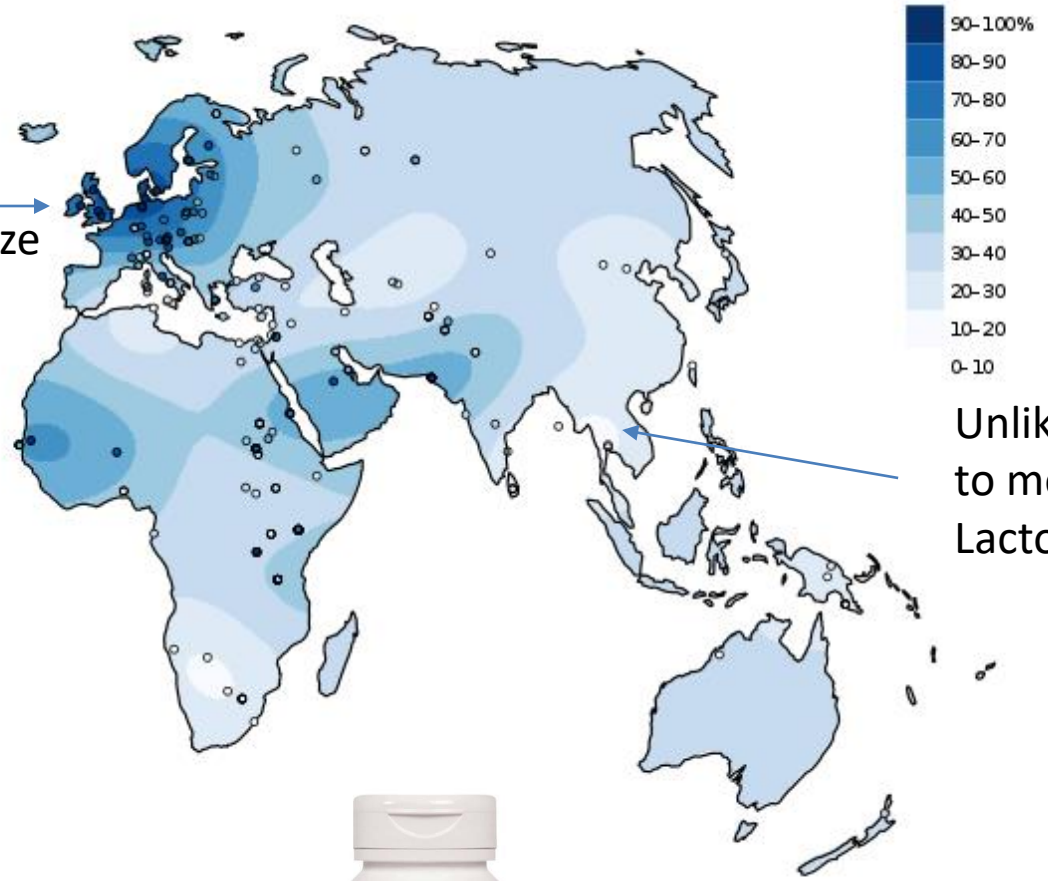
Whole/low-fat Milk
200ml
9-10g lactose

Cheddar Cheese
25g
0.03g lactose

Whole Milk Yogurt
125g
6g lactose

Lactose-free Milk
200ml
0g lactose

Lactase Persistence

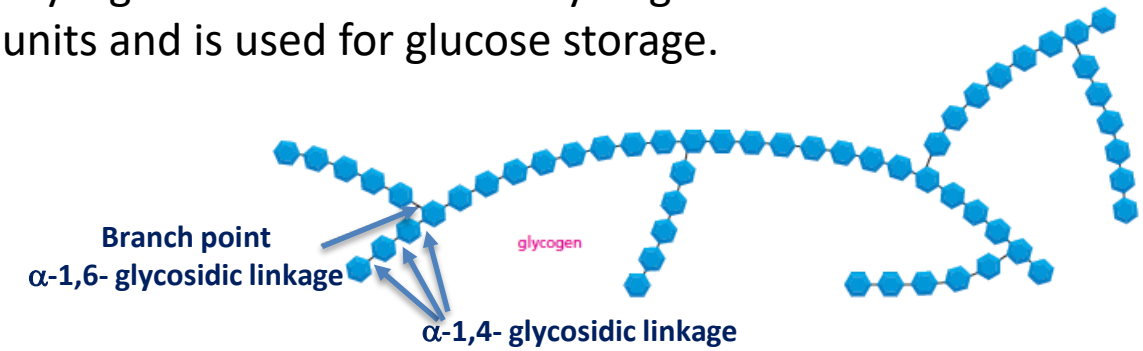


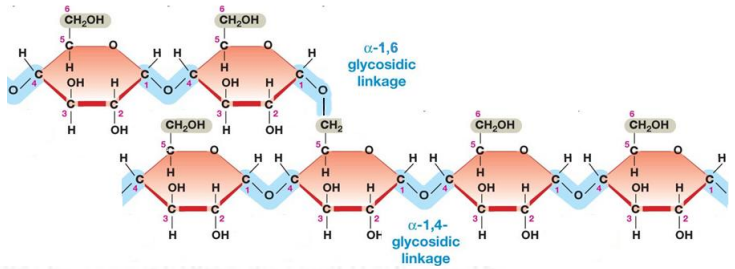
Can metabolize Lactose

Unlikely to be able to metabolize Lactose

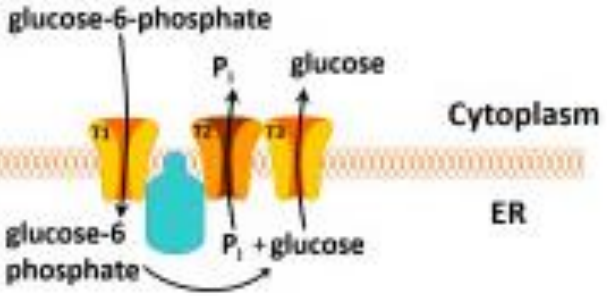
Polysaccharides as Energy Storage – Glycogen Storage Disease

Glycogen and is made entirely of glucose units and is used for glucose storage.

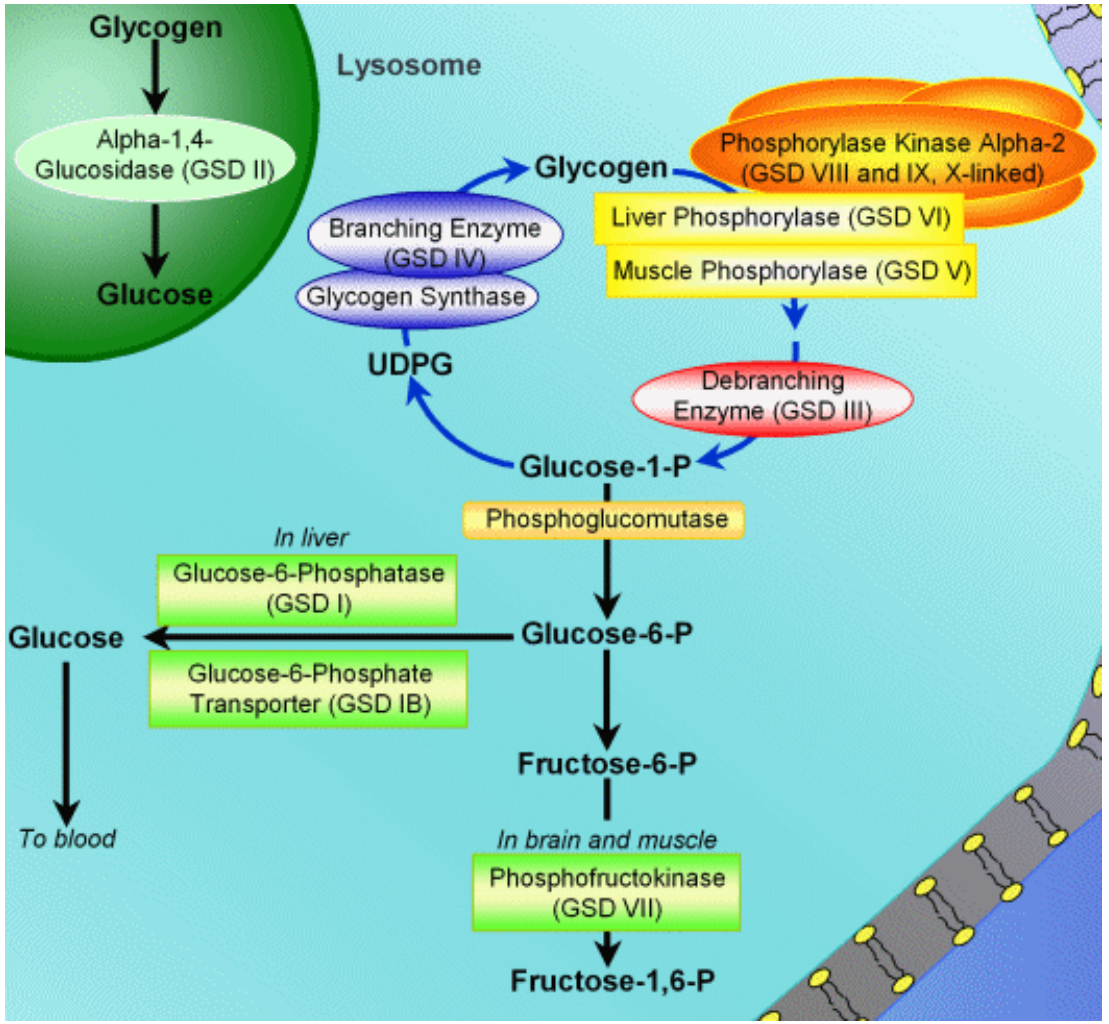




glycogen + P_i → glycogen(n-1) + glucose-1-phosphate



https://www.ncbi.nlm.nih.gov/books/NBK459277/



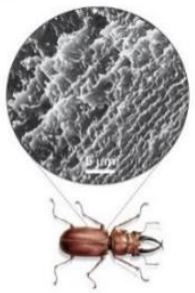
https://catalog.coriell.org/0/sections/collections/nigms/gsd_pathway.aspx?PgId=254

Polysaccharides as Structural Molecules

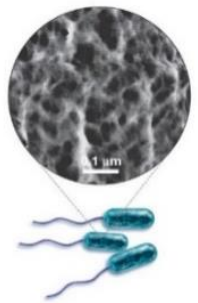
Cellulose, Chitin, Peptidoglycan Form Tough Fibers

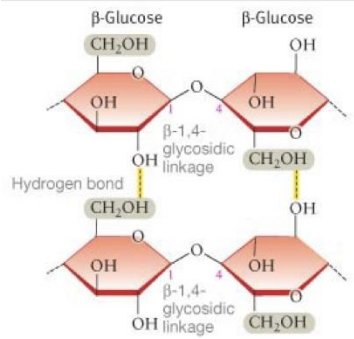
Cellulose in plant cell wall

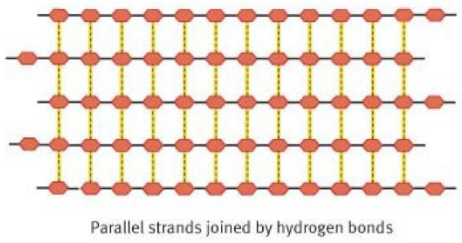
Chitin in insect exoskeleton



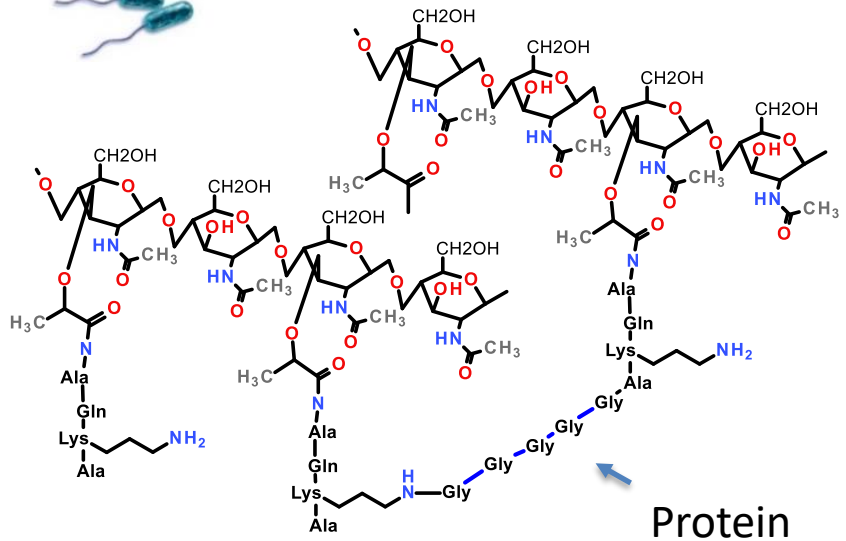
Peptidoglycan (protein + sugar) in bacterial cell wall







Cellulose



Peptidoglycan (Bacterial Cell Wall)

Summary and Expectations for Carbohydrates

Key Points:

- General structure of monosaccharides - be able to distinguish between aldose and ketose (and identify compounds that are not sugars).
- Know how to number carbons on aldoses and ketoses
- Be able to describe the linkage between two monosaccharides (configuration at the anomeric carbon, atoms linked)
- Be able to describe the linkage between glucose molecules in:
  - Glycogen (glucose storage)
- Be able to describe the overall structure of the peptidoglycan in bacterial cell walls.