Drugs & Disease – Spring 2025

Course Overview:

1. Introductory Biochemistry
2. DNA, RNA, protein synthesis, biotechnology
3. Immunology & Immunotherapy
4. Drug Discovery – Enzyme Inhibitors
5. Genome Editing – CRISPR
6. Final presentations

Expectations:

- 5 Problem sets (***First one posted***)
- One mid-class exam
- Presentation (10 min, topic of choice)
- Short paper (Same topic as presentation)

My Story

- Born in Ottawa Canada
- Undergraduate: University of Waterloo, largely physics
- MS: Penn State University
- PhD: Carnegie Mellon
- Post-doc: Stanford University
- Faculty: University of Virginia then Carnegie Mellon

Research area:

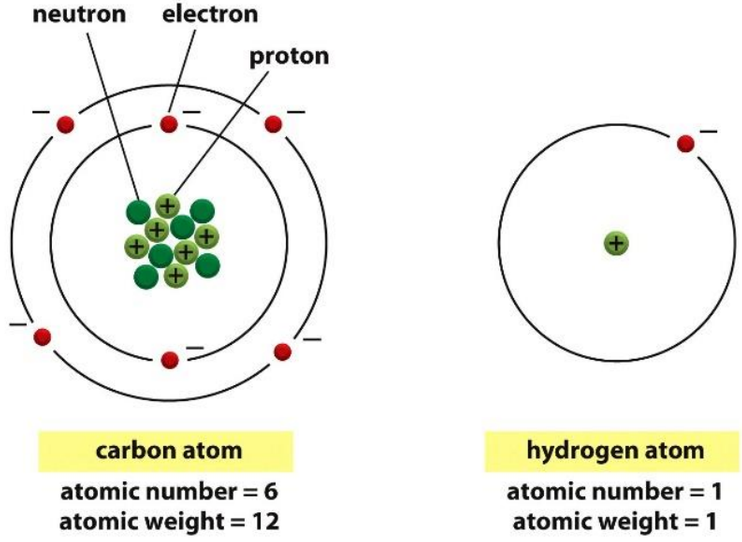
- Protein structure and dynamics
- Drug discovery

Take-home exercise:
Send me an email with a short paragraph describing why you took the course and what you hope to take away from the course.

Course materials:
https://www.andrew.cmu.edu/user/rule/Drugs_Disease/

Lecture 1
Chemistry and Biology Fundamentals

- Chemical Bonding
- Functional Groups
- Chirality of carbon, chiral drugs
- Molecular interactions
- pH and charge
- Protein Structure and Stability
- Ligand Binding
- Proteins as enzymes (PKU disease)



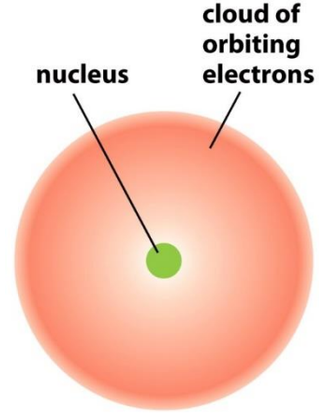
- Atoms are composed of:
  - Protons – positively charged particles
  - Neutrons – neutral particles
  - Electrons – negatively charged particles
- Protons and neutrons are located in the nucleus.
- Electrons are found in **orbitals** surrounding the nucleus.
- The overall charge on an element is neutral (#electrons = # protons).

Mass number
(number of protons + neutrons)

Atomic number
(number of protons)

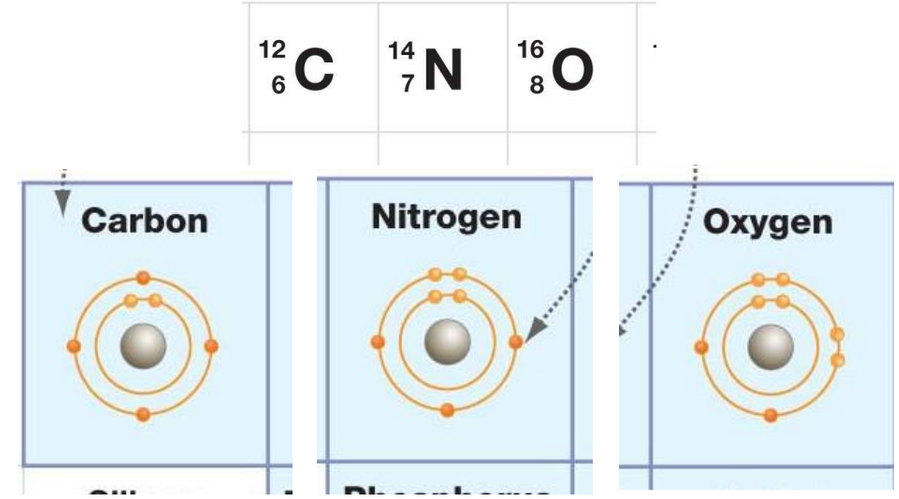
$^{1}_{1}H$							$^{4}_{2}He$
$^{7}_{3}Li$	$^{9}_{4}Be$	$^{11}_{5}B$	$^{12}_{6}C$	$^{14}_{7}N$	$^{16}_{8}O$	$^{19}_{9}F$	$^{20}_{10}Ne$
$^{23}_{11}Na$	$^{24}_{12}Mg$	$^{27}_{13}Al$	$^{28}_{14}Si$	$^{31}_{15}P$	$^{32}_{16}S$	$^{35}_{17}Cl$	$^{40}_{18}Ar$

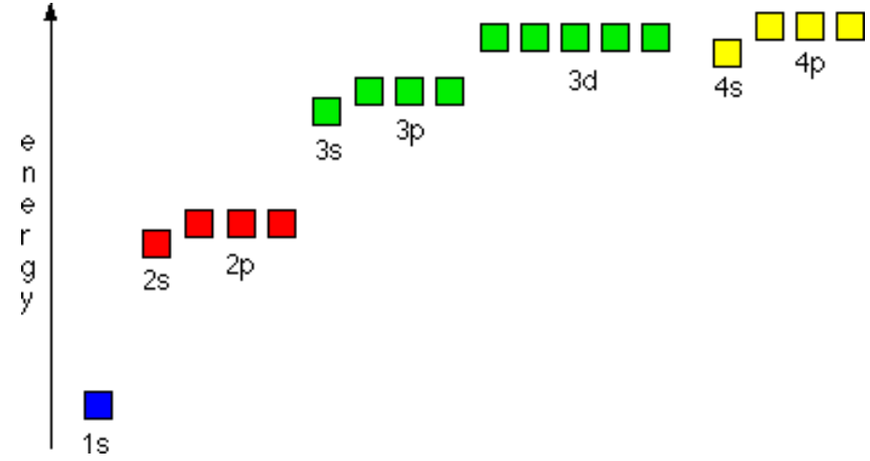
Atomic number = # of protons = # electrons in element
Isotope = different # of neutrons = same bonding capability



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 via electron sharing.
- The most stable configuration is a complete (full) outer shell.

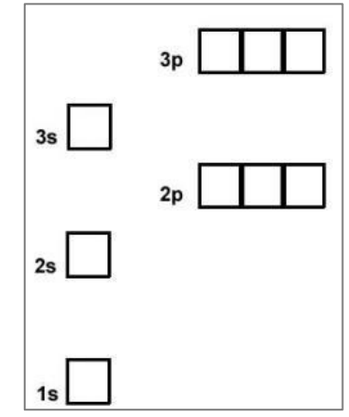




Shells: 1st = 1s, 2nd = 2s + 2p, 3rd = 3s + 3p

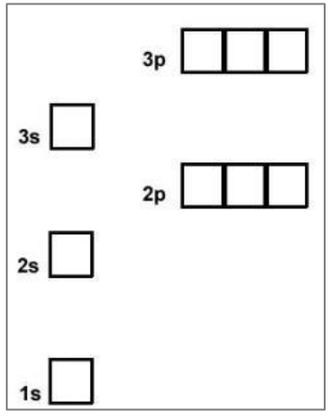
Shell is a collection of orbitals with similar energy

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



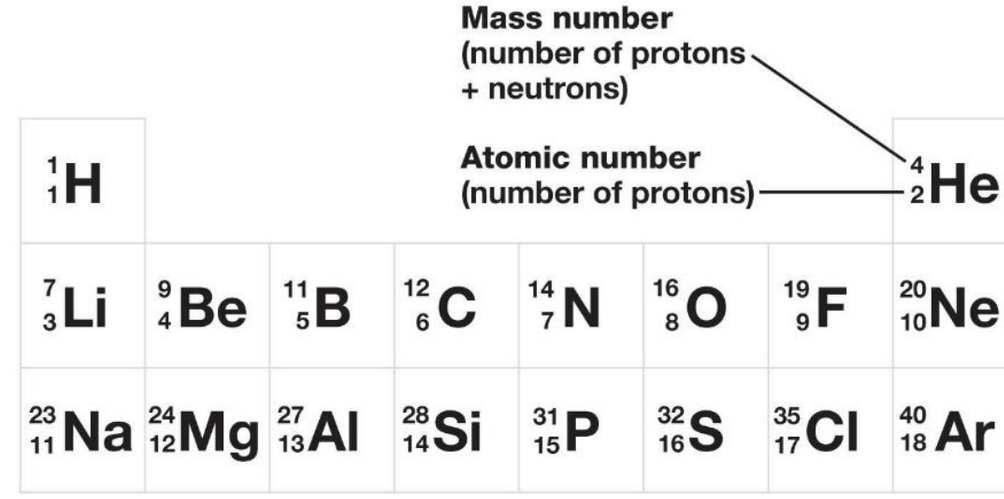
Ions or Covalent Chemical Bonds – What's an Element going to do?

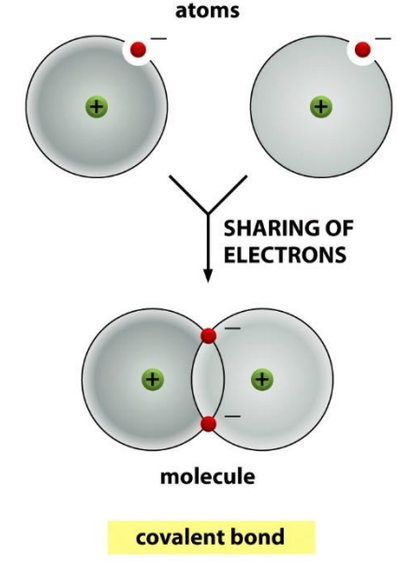
- Elements like Li, Na, F, Cl, Mg, readily form ions to generate a complete outer shell.



Example:
Li with 3 electrons

- 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





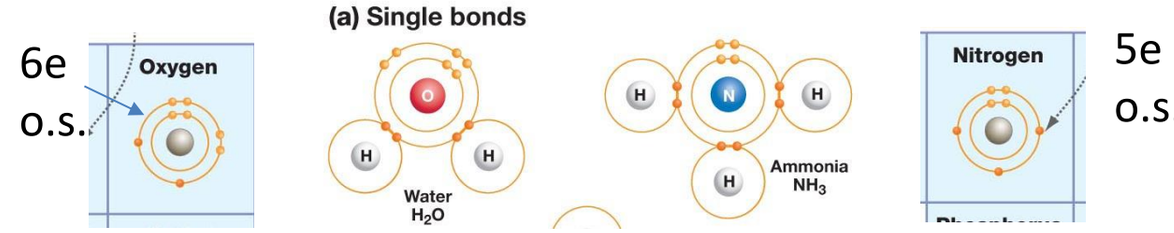
Covalent Bonds – Filling the Outer Shell by Sharing

- 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.

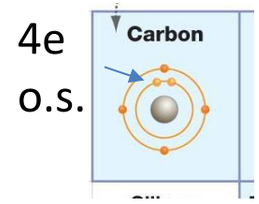
The number of covalent bonds (valence) formed by common elements.

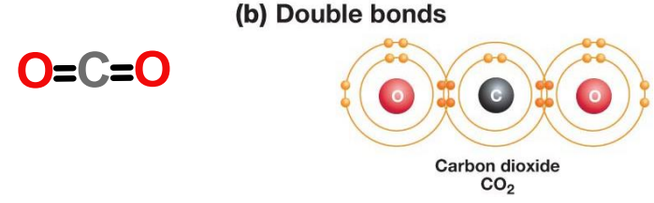
- Oxygen = 2 bonds
- Nitrogen = 3 bonds
- Carbon = _____
- Sulfur = 2 bonds (in biological systems)
- Hydrogen = 1 bond
- Phosphorous = 5 bonds in biological molecules

You must know these numbers.



How many bonds will carbon form?

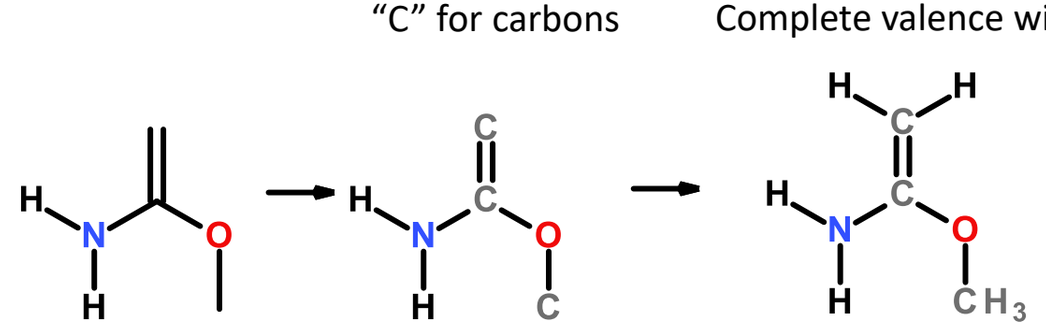




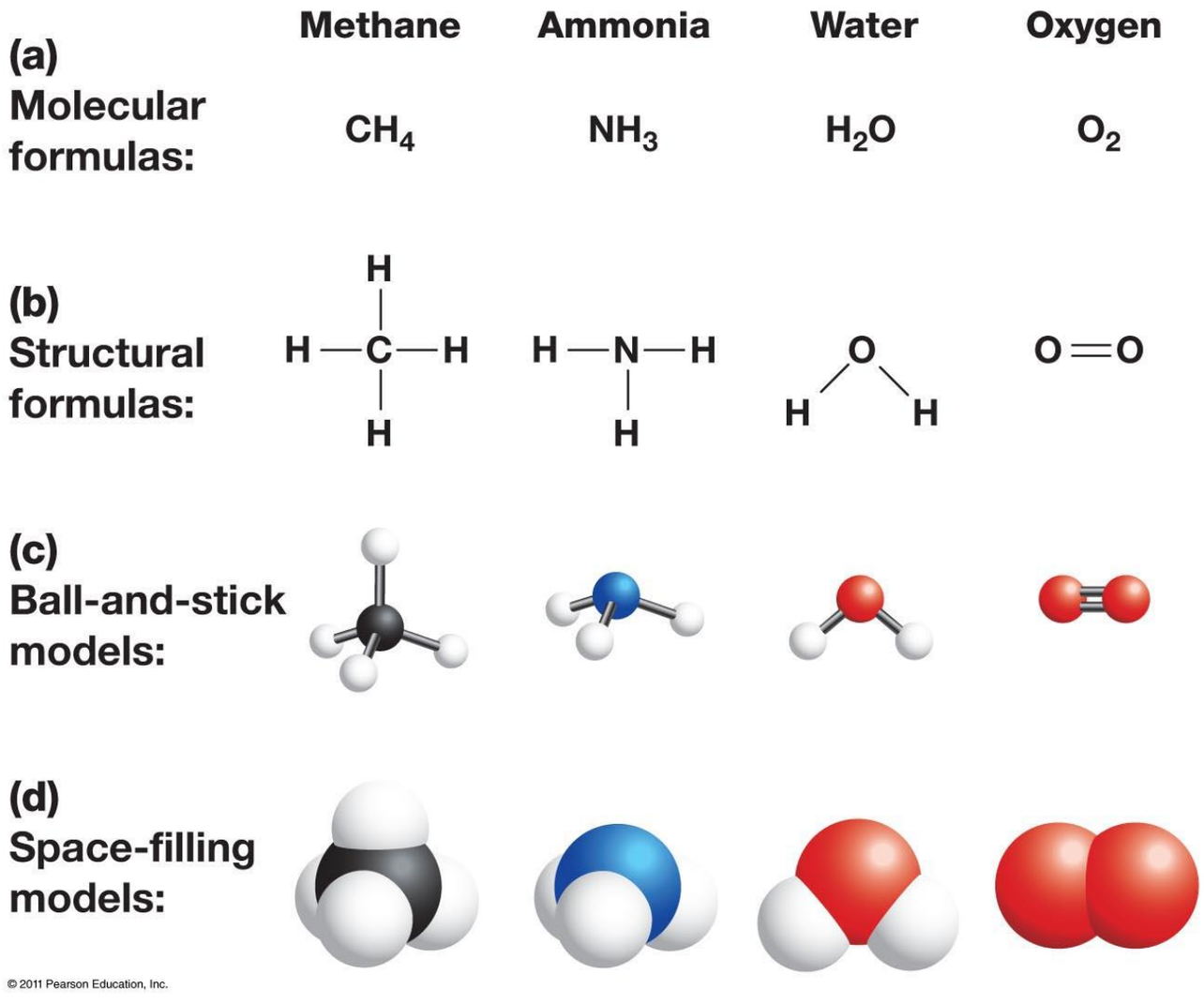
Abbreviated Chemical Drawings:

- "C" for carbon is 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.



Representation of Molecules

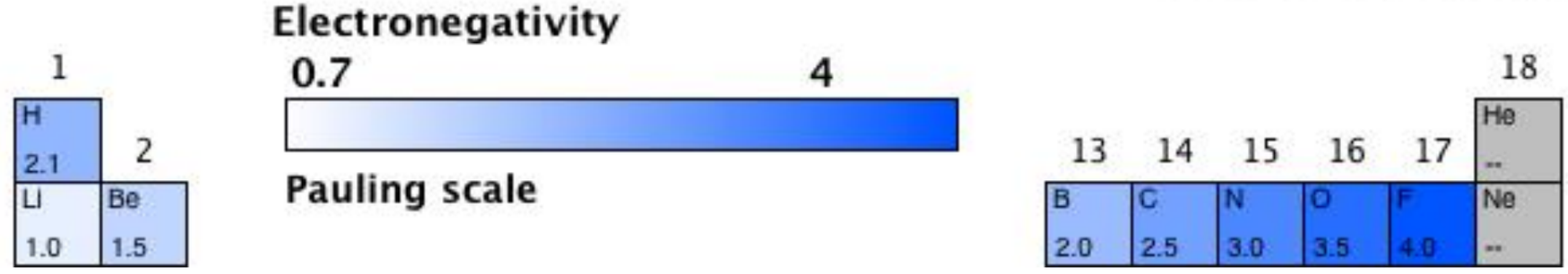


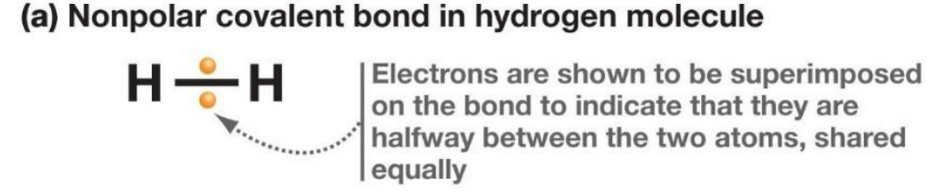
Electron Sharing and Bond Polarity – Are All Bonds Equal?

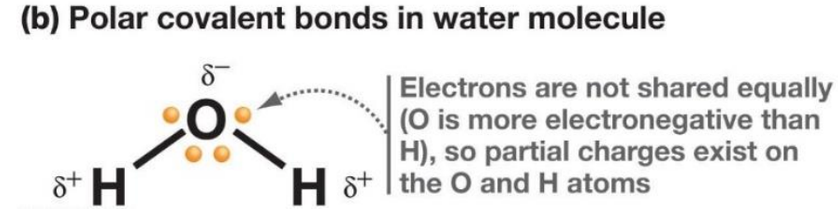
- Polar bonds = different electron density of each atom.
- The polarity of a bond depends on the electronegativity of the atoms.
- Electronegativity - ability of atoms to pull electrons from other atoms.
- Atoms with higher electronegativity will develop a partial negative charge, the atom they are bonded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

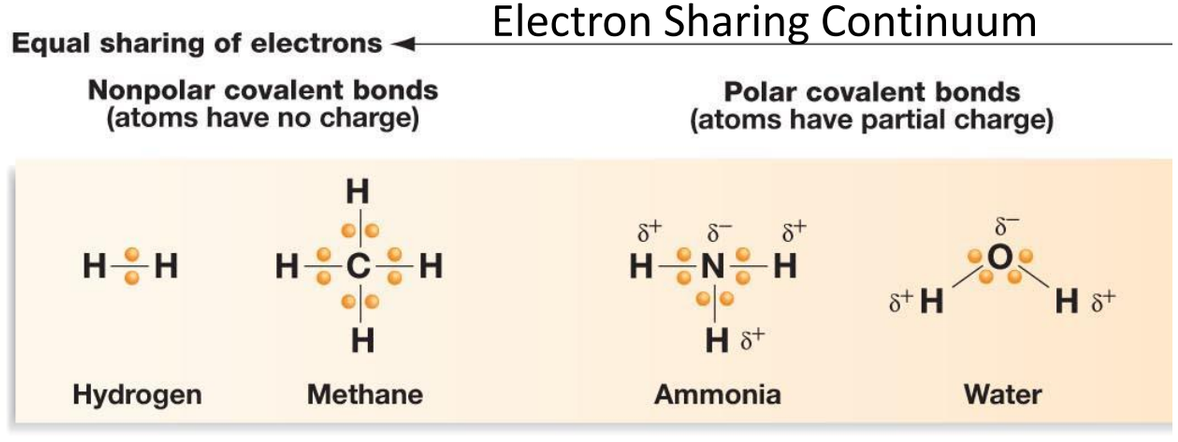
  H ~ C < N < O

Increased pos. charge of nucleus.







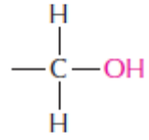


Functional Groups – You should Become Familiar with These

C–O COMPOUNDS

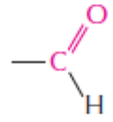
Many biological compounds contain a carbon bonded to an oxygen. For example,

alcohol

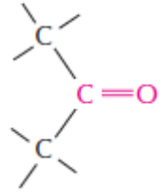


The –OH is called a hydroxyl group.

aldehyde

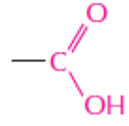


ketone



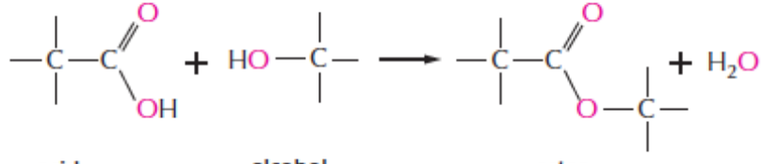
The C=O is called a carbonyl group.

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.

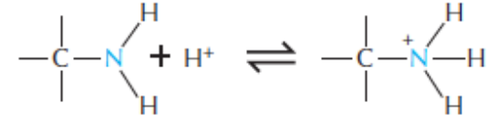


acid alcohol ester

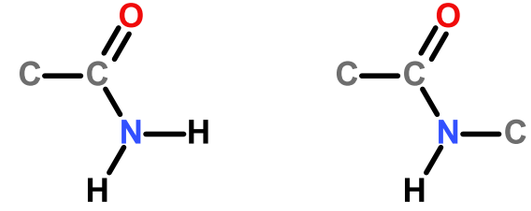
C–N COMPOUNDS

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.

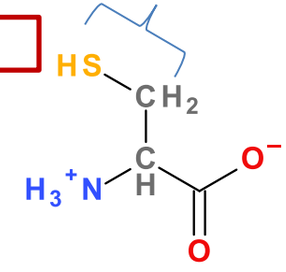


Amides are formed by combining an acid and an amine. Unlike amines, amides are uncharged in water. An example is the peptide bond that joins amino acids in a protein.



C–S COMPOUNDS

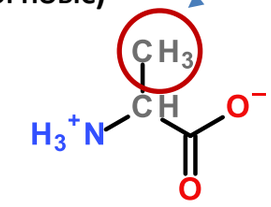
thiol



Cysteine (amino acid)

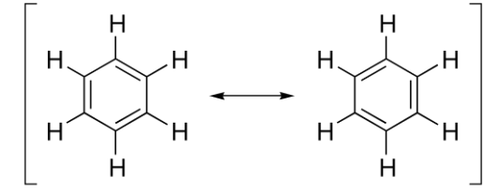
C–H GROUPS (HYDROPHOBIC)

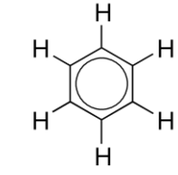
Methyl



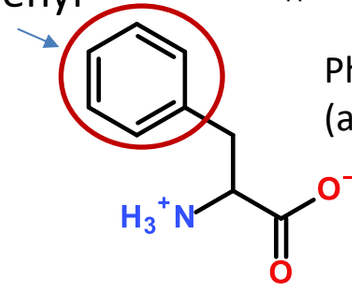
Alanine (amino acid)

Aromatic – planer rings, alt double bonds





Phenyl



Phenylalanine (amino acid)

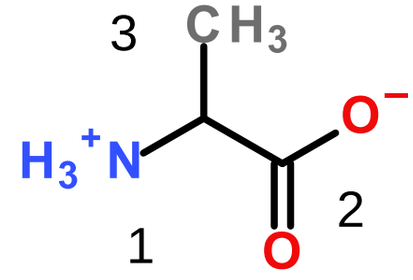
Key Points & Expectations

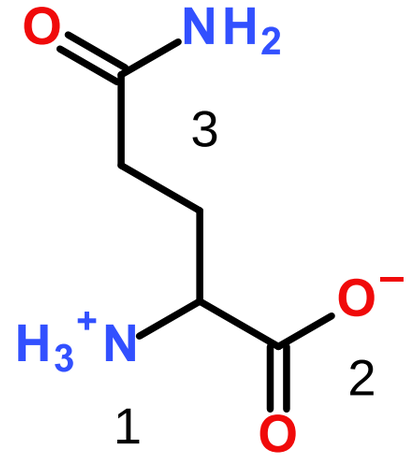
Chemistry

- 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.
- Predict degree of bond polarity based on electronegativies, N-H and O-H and C=O are polar, C-H is not.

- Be able to draw the following functional groups & identify them on larger molecules.
  - Non-polar:
    - Methyl
    - Phenyl
  - Polar:
    - Alcohol (C-OH)
    - Thiol (C-SH)
    - Carboxylate (ketone, aldehyde) (C=O)
    - Ester
    - Carboxylic acid
    - Amide
    - Amino

A. Give the names of the functional groups on these two amino acids.
B. Which functional groups are common?

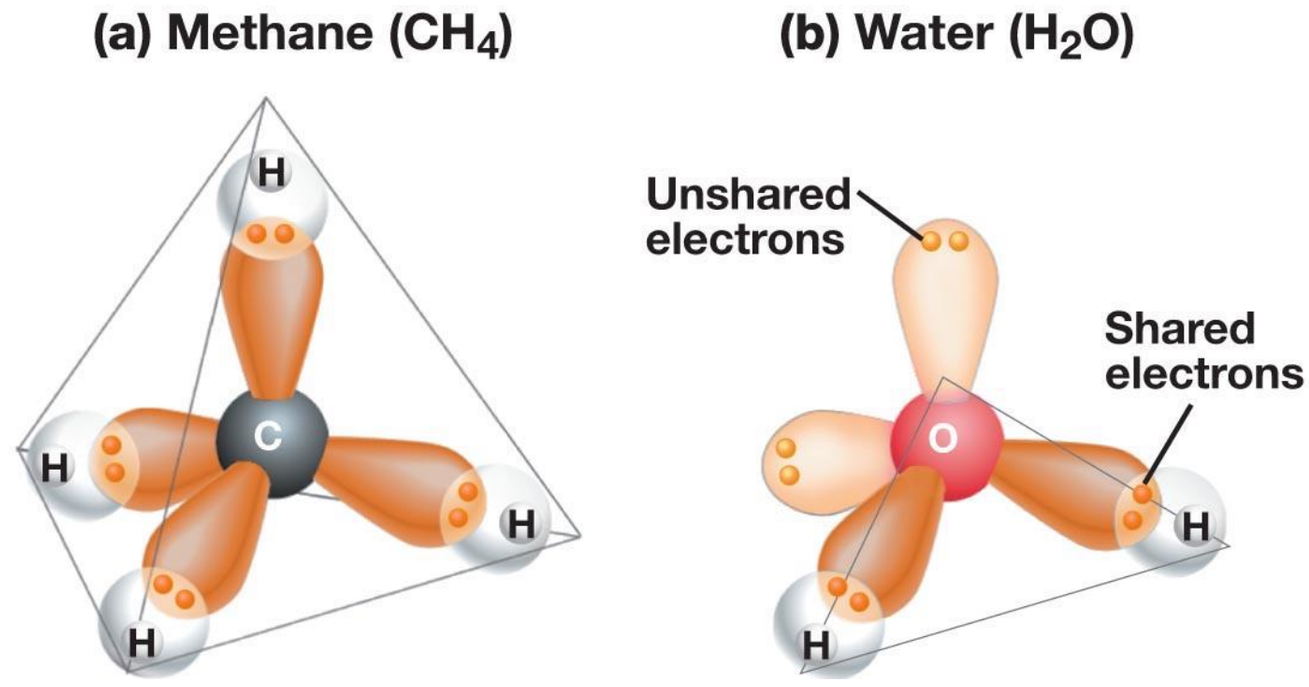




1.
2.
3.

1.
2.
3.

The Geometry of Simple Molecules

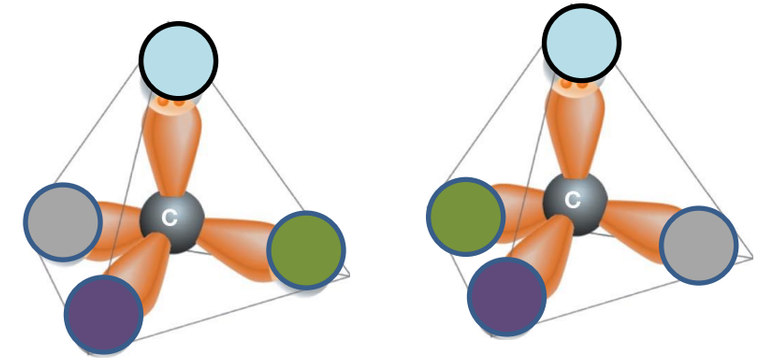


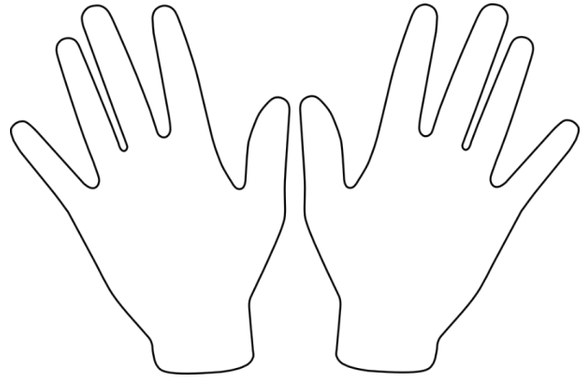
The shape of a molecule is determined by the geometry of its bonds.

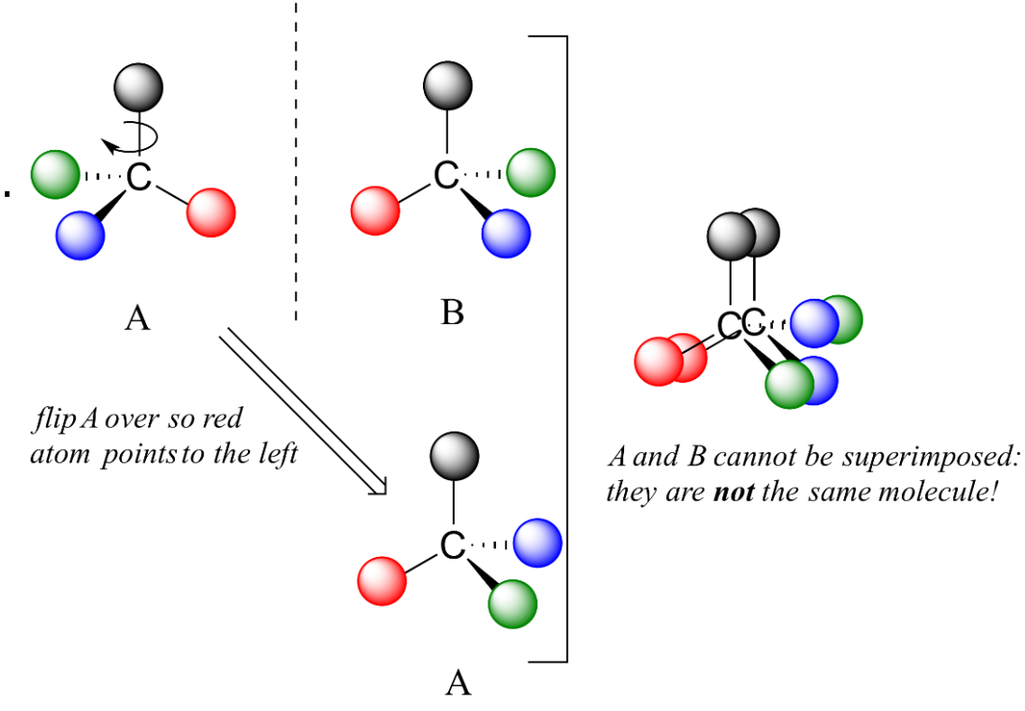
Carbon, oxygen, and nitrogen often form bonds with a tetrahedral geometry

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

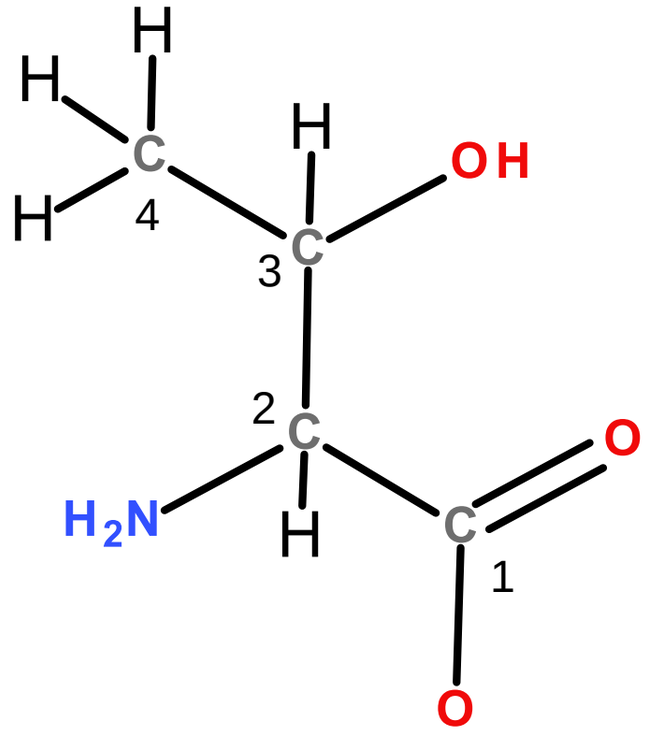






Identify the Chiral Centers on Threonine

Can you identify chiral centers?



Which carbon is chiral?

1 Yes or No?

2 Yes or No?

3 Yes or No?

4 Yes or No?

Drugs with Chiral Centers

Nobel Prize for Chiral Synthesis 2001

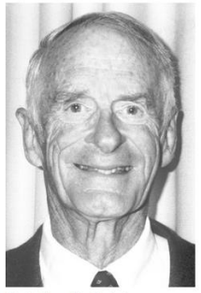


Photo from the Nobel Foundation archive.
William S. Knowles

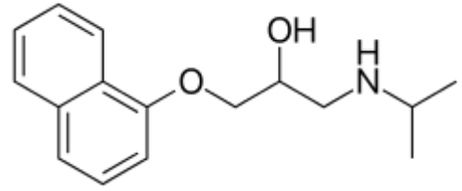

Photo from the Nobel Foundation archive.
Ryoji Noyori


Photo from the Nobel Foundation archive.
K. Barry Sharpless

and a fun game:

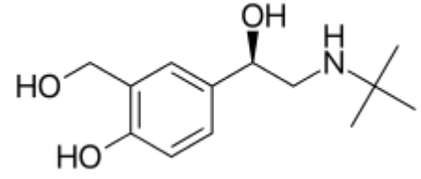
https://educationalgames.nobelprize.org/educational/chemistry/chiral/game/game.html

Propranolol



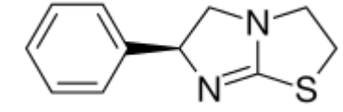
1. Racemic mixture is used to treat high blood pressure.

levalbuterol



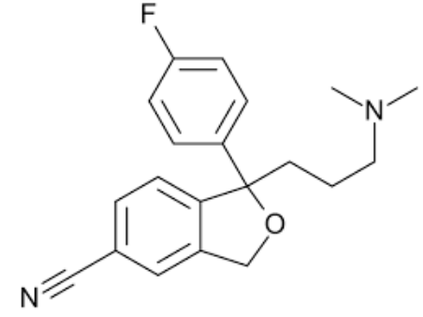
2. R-enantiomer used to treat asthma

levamisole



3. L-form used to treat parasitic worm infections

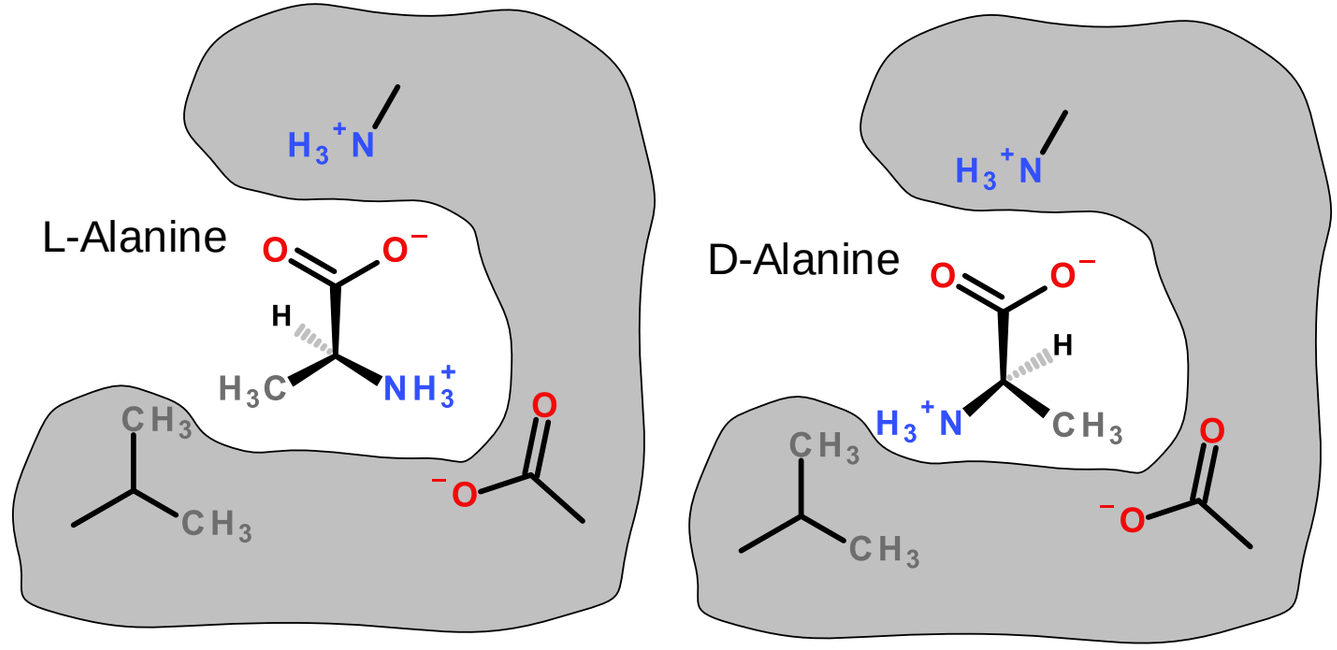
citalopram



4. Antidepressent (escitalopram is L)

Why Chirality Matters

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



L-Alanine binds better because of more

favorable _____ interactions.

pH, Strong Acids & Bases

- pH = –log $[H^+]$ = -log$[H_3O^+]$
- 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.

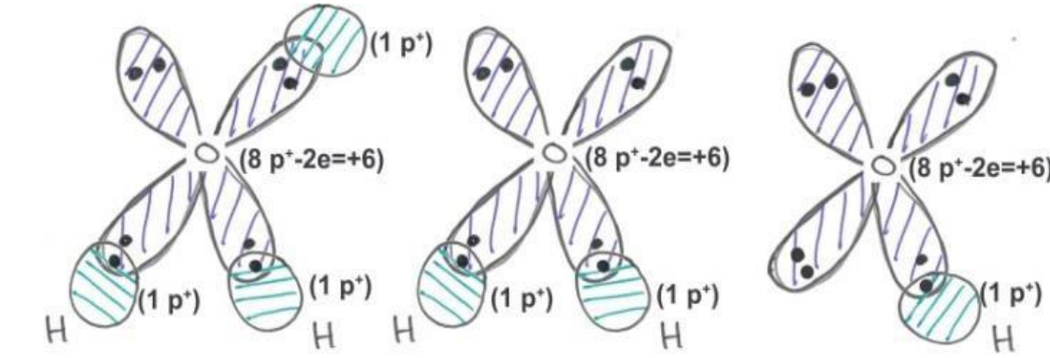
- Acids release protons and will lower the pH of the solution, e.g.

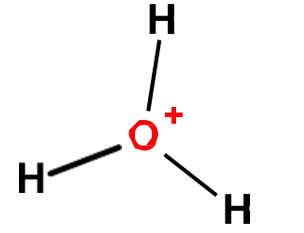
  $HCl + H_2O \rightleftharpoons H_3O^+ + Cl^-$

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

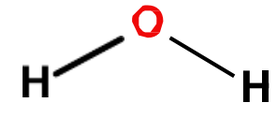
  $NaOH \Rightarrow Na^+ + OH^-$

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

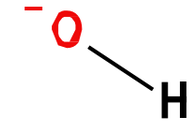




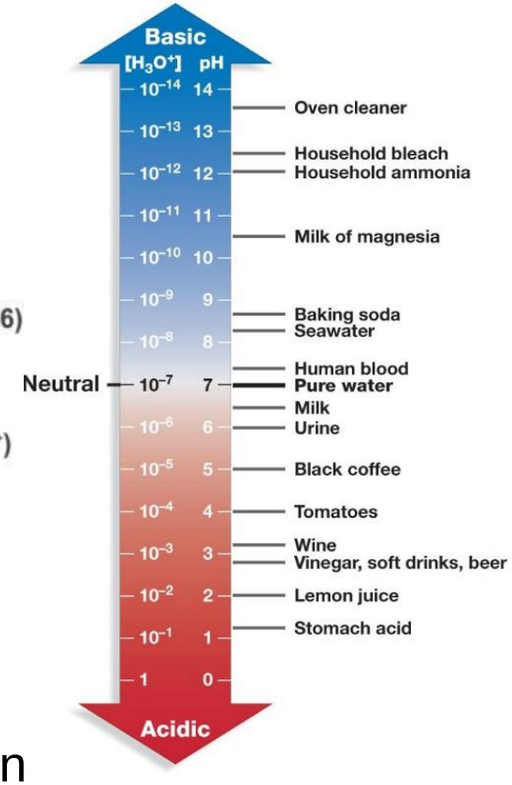
Hydronium ion



Water



Hydroxide ion



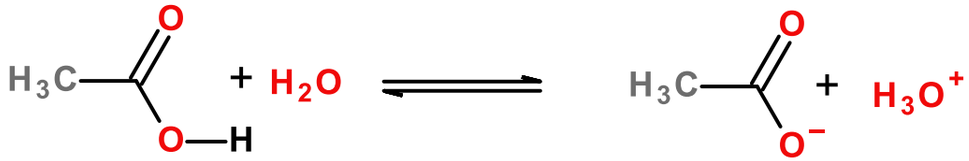
1. Which solution has a higher H^+ concentration, pH=3 or pH 4.
2. How large is the difference?

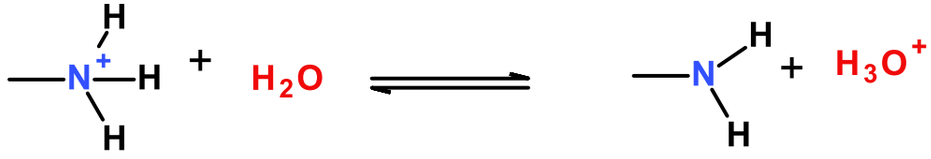
Acids and Bases.

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

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

Weak Acid – incomplete ionization in solution.



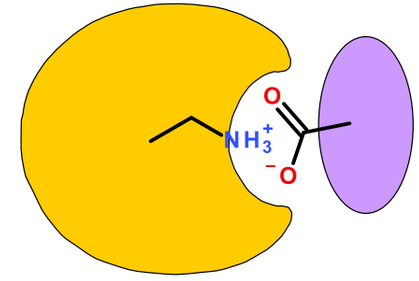


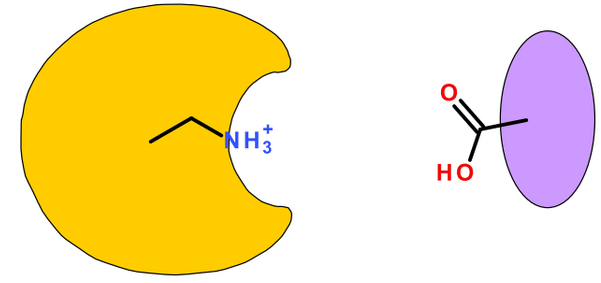
“HA”=protonated form

“A”=deprotonated form (conjugate base)

Why this is important:

protonation/deprotonation changes the ***charge*** on species, either creating or destroying strong electrostatic interactions!



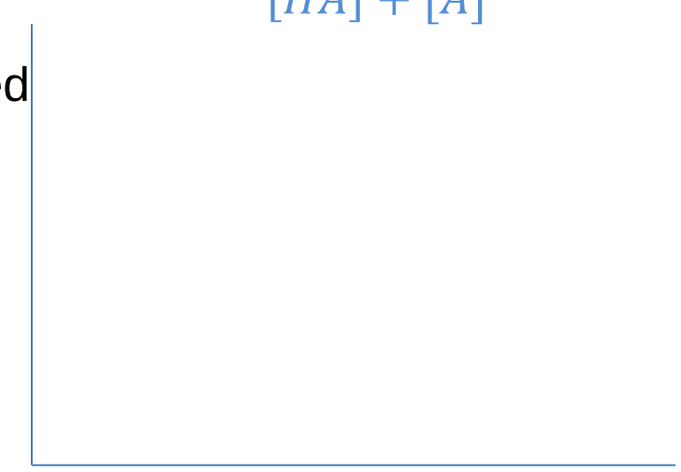


What Affects the Degree of Protonation?

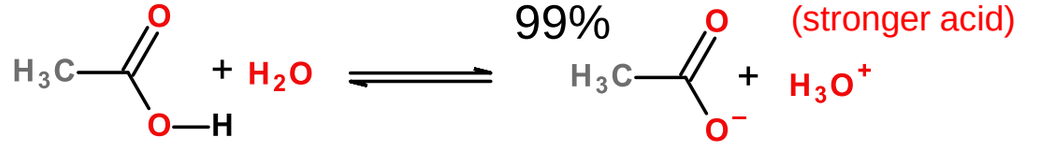
1. The extent of protonation/deprotonation depends on the pH of the solution:
   - Low pH values will favor protonation of acids since there are many protons that will collide with (A) to make (HA).
   - High pH values will favor deprotonation of acids since there are fewer protons to protonate the acid.

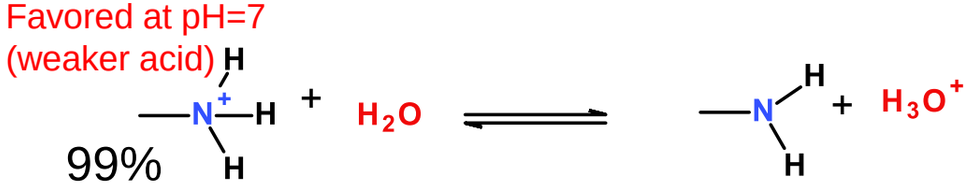
*What would you expect to happen to the fraction of the acid that is protonated (f_{HA}) as the pH of the solution is **decreased**?*

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



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.

Molecular Interactions

$$\Delta E = E_{AB} - (E_A + E_B) \sim E_{A \cup B}$$

The energy change when two things come together can be approximated to be due to ***new*** inter-molecular interactions: $E_{A \cup B}$.

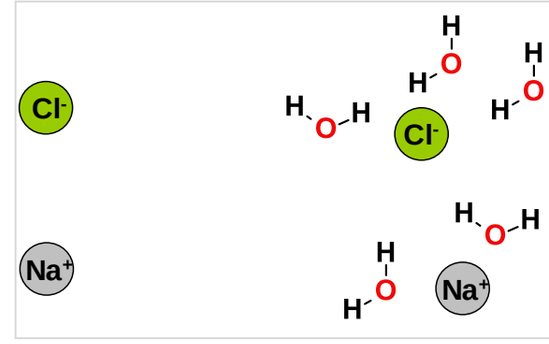
Interaction	Interaction	Energy (kJ/mol)
Electrostatic interactions (in water)	Full charges	~5 kJ/mol/single interaction
Van der Waals: Dipole-Dipole	Perm. partial charges	~0.05 kJ/A^2 x 100 A^2 = 5 kJ/mol for 100 A^2
Van der Waals: Induced-dipole	Induced partial charges	~0.02 kJ/A^2 x 100 A^2 = **2 kJ/mol for 100 A^2**
H-Bonds	Electrostatic + e sharing	~20 kJ/mol gross, **~5 kJ/mol net**

i) Electrostatics: The interaction energy between two charged particles is:

$$E = \frac{1}{D}\frac{1}{4\pi\varepsilon_o}\frac{q_1 q_2}{r}$$

The energy depends on the charges of the particles (q_1, q_2), distance (r) between the two charges, and the dielectric constant (D) of the media.

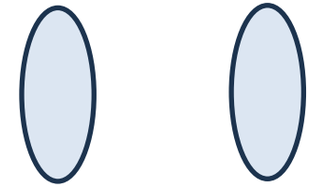
How strong are electrostatic interactions?
$Na^+ Cl^-$ = ~ -700 kJ/mol *in vacuum (D=1) when r = 2A*



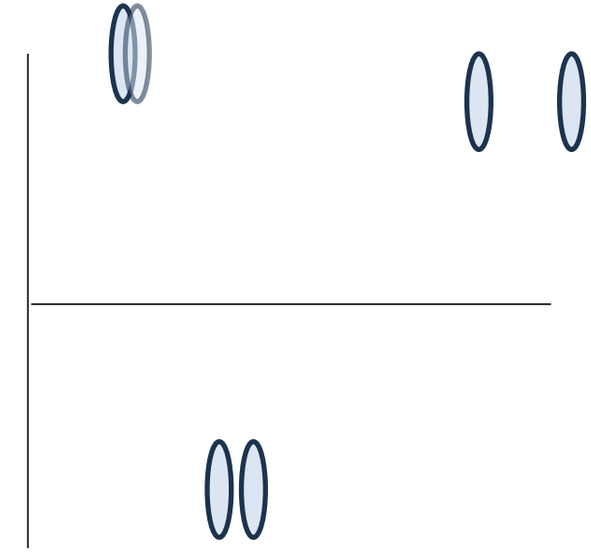
Water has a high dielectric constant of 80 due to its polar nature. How does this affect the energy of interaction?

Van Der Waals Forces:

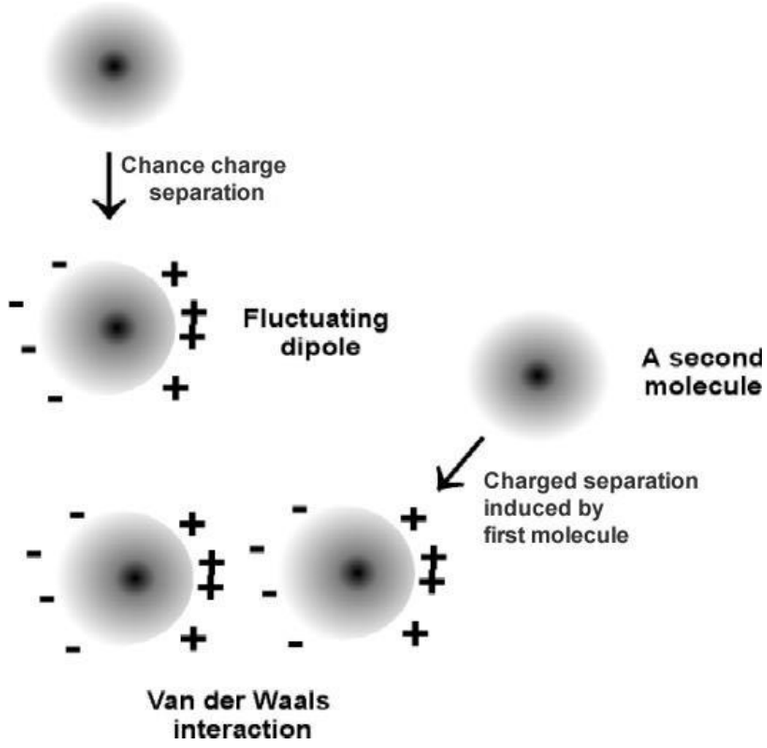
ii) Dipole-dipole – an electrostatic interaction that involves permanent ***partial*** charges (these are sometimes called Keesom forces).

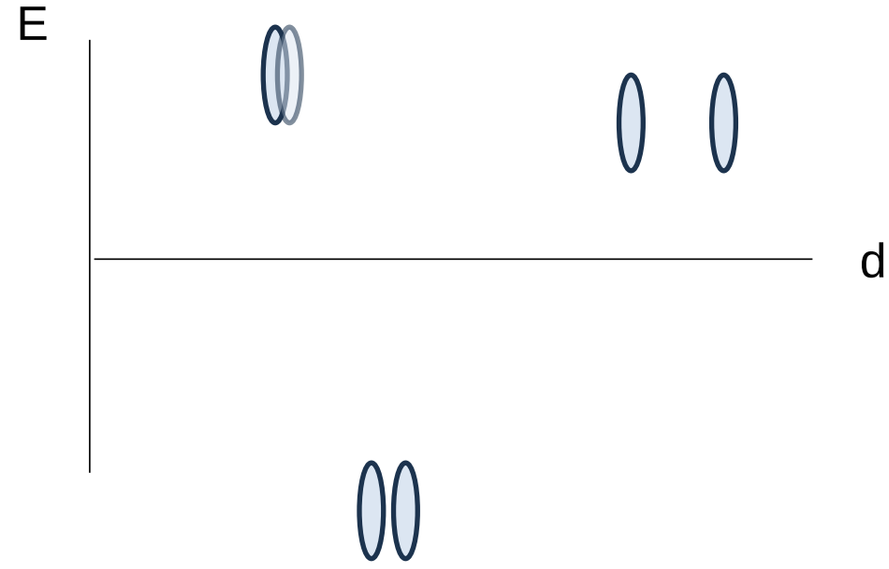


Electrostatic energy of *point* charges.



iii) Induced dipole (often referred to as London Dispersion)

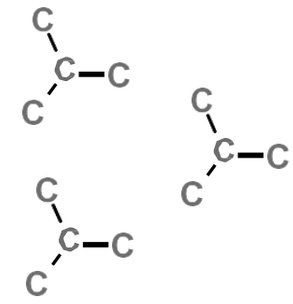




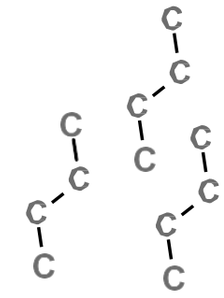
iii) Induced dipole (often referred to as London Dispersion)

Although weak, the effects of van der Waals are easily observed: Boiling points of two hydrocarbons:

isobutane: 261 K

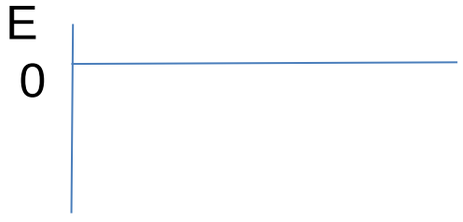


butane: 272 K



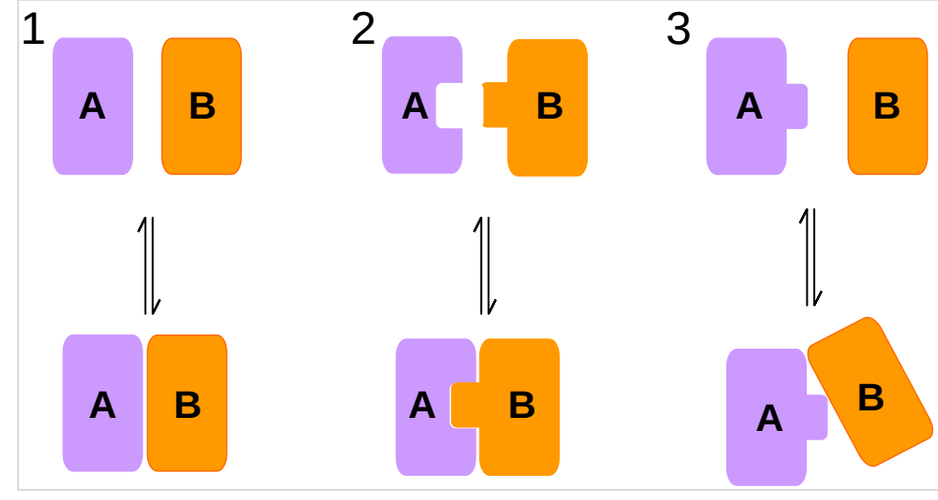
1. Same number of carbons, why the difference in boiling points?

2. How will van der Waals interaction energies scale with contact area?



3. Which of these will have the most favorable vdw interaction:
1 2 3

4. Which of these will have the least favorable vdw interaction:
1 2 3



https://www.youtube.com/watch?v=uhfXbSSrabw

iv) Hydrogen Bonds

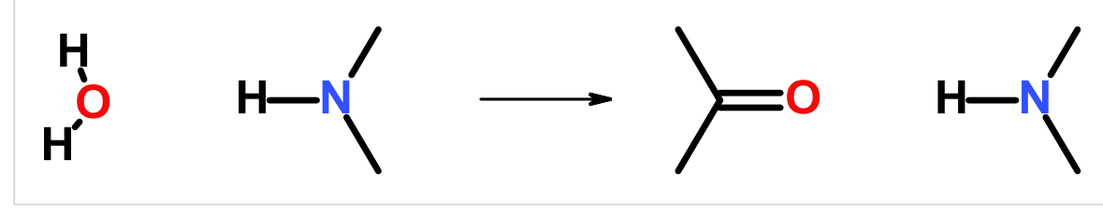
- H-bonds are primarily (90%) an electrostatic attraction between:
  - Electropositive hydrogen, attached to an electronegative atom is the hydrogen bond donor (i.e. NH).
  - Electronegative hydrogen bond acceptor (e.g. the lone pairs of oxygen, or C=O group of an amide).

	$\lvert\Delta e\rvert$
N——**H** **3.0** **2.1**	0.9
O——**H** **3.5** **2.1**	1.4

A "bond" implies electron sharing – about 10% of the electron is shared from one molecule to the next in the case of H-bonds

*Note that the proton is **NOT** transferred to the acceptor, it remains covalently bonded to the donor atom. The Hydrogen Bond is the **interaction** between the X-H donor and electronegative acceptor.*

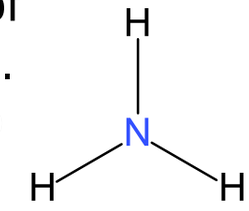
- The energy released when an H-bond forms depends on the distance and angle of the bond.
- Usually hydrogen bonds are exchanged, resulting in small ***net*** energy differences:

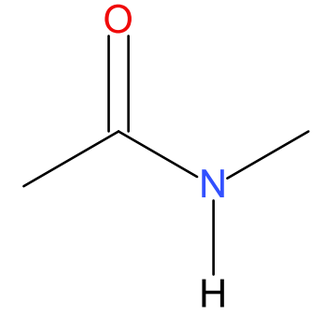


How to Identify Hydrogen Bond Donor and Acceptors

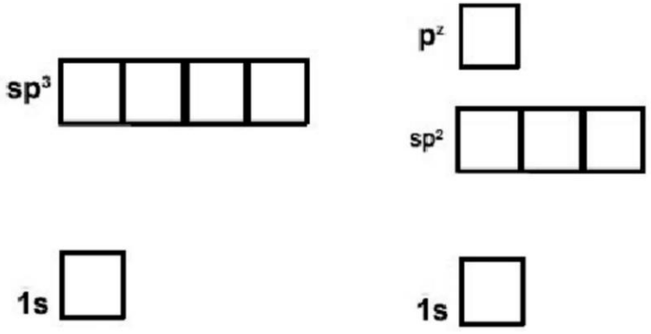
Exceptions, N in a delocalized system:

- Will not accept from above or below the plane of the system, because the lonepair is delocalized.
- Can accept in the plane of the ring if there is no attached hydrogen, via lone pair in sp2 orbital



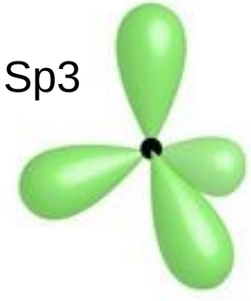


- Sp2 is used in amides, allowing favorable overlap of the full pz orbital with the pz on C and O
- The lone pair in the pz is shared with the pz electrons on carbon and oxygen.
- Due to electron sharing, there is only a slight neg. charge and the group does not accept an H-bond.

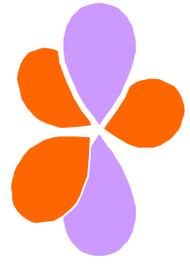


N – 7 electrons

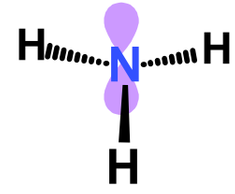
Sp3

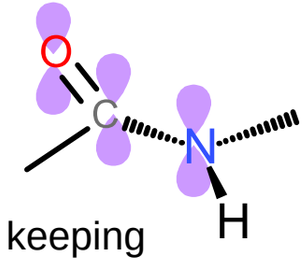


Sp2

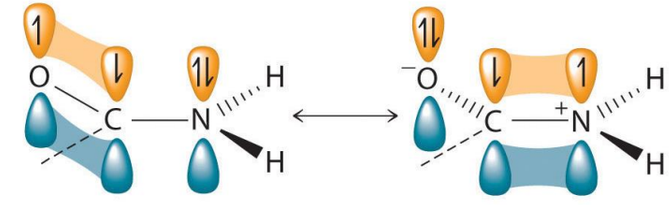


Nitrogen can form two types of hybrid orbitals, sp3 (tetrahedral geometry) or sp2 (planer) + pz



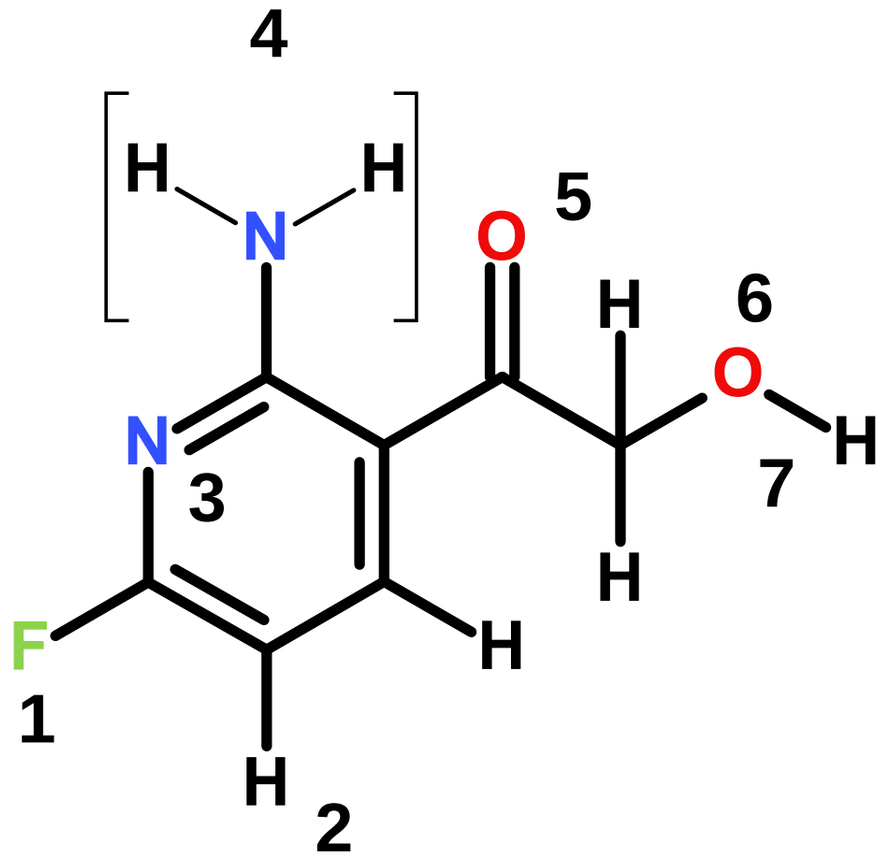


- Sp3 is used in ammonia, keeping the three hydrogen atoms as far from each other as possible. The fourth sp3 orbital is full with two electrons (lone pair).
- The lone pair is an excellent acceptor.



(a)

1. *Indicate which atoms could donate an H-bond and which could accept and H-bond*



ATOM	Donor?	Acceptor?	Neither
1 (F)			
2 (C_{aro}-H)			
3 (N)			
4 (-NH_2)			
5 (C=O)			
6 (O)			
7 (O-H)			

Can you?

- Identify groups that can donate or accept hydrogen bonds?

Relative Energy of Interactions

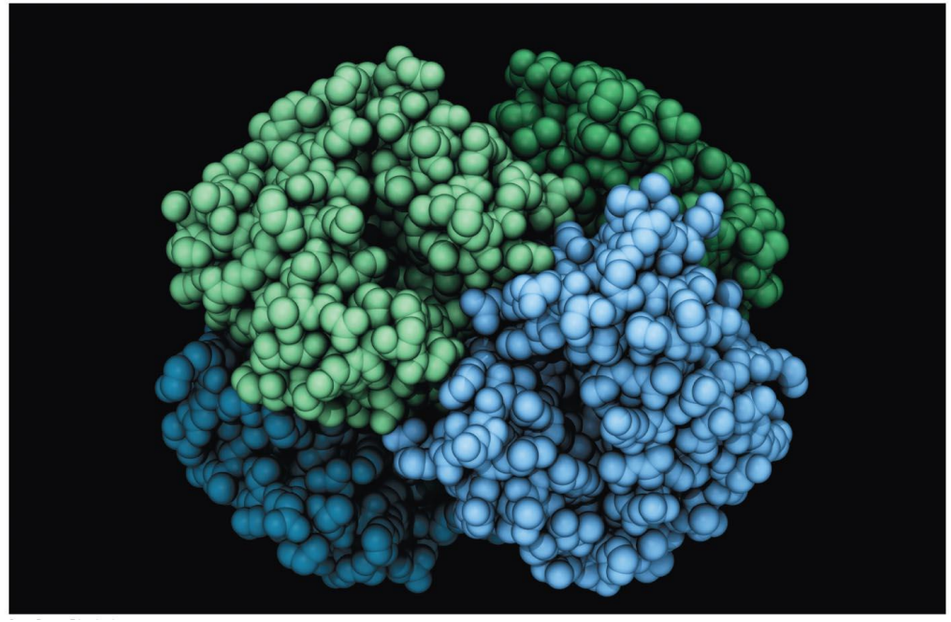
Interaction	Interaction	Energy (kJ/mol)
Covalent Bond	Electron sharing	200-400 kJ/mol
Electrostatic interactions (in water)	Full charges	~5 kJ/mol/single interaction
VdW - Dipole-dipole (Keesom)	Perm. partial charges	~0.05 kJ/A^2 x 100 A^2 = 5 kJ/mol for 100 A^2
VdW – Induced dipole (London)	Induced partial charges	~0.02 kJ/A^2 x 100 A^2 = **2 kJ/mol for 100 A^2**
H-Bonds	Electrostatic + e sharing	~20 kJ/mol gross, **~5 kJ/mol net**

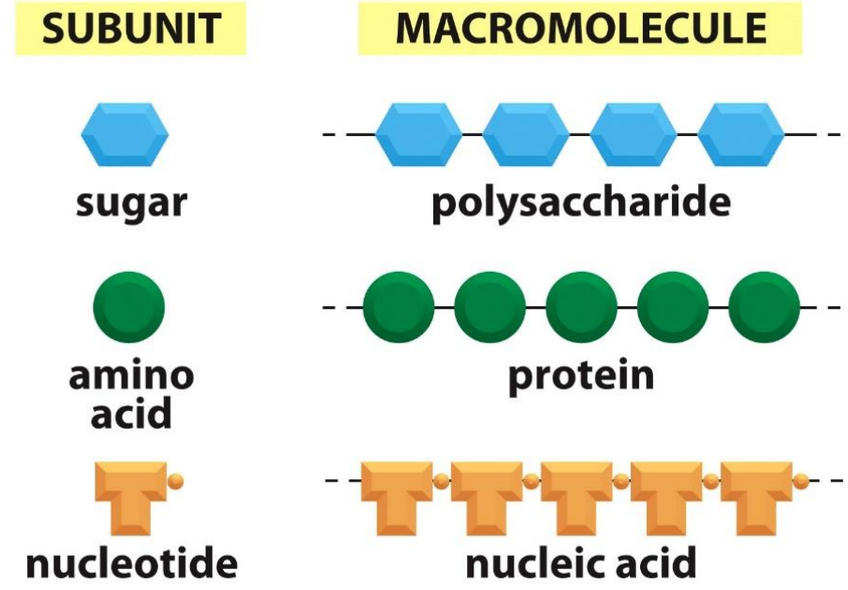
1. How does the energy of the last four interactions compare to covalent bonds?
1. Stronger 2. Weaker 3. The Same

2. Which of these are closer to thermal energy, kT = 2.5 kJ/mol @ room temp.

3. What is the advantage of a weak interaction in biology?

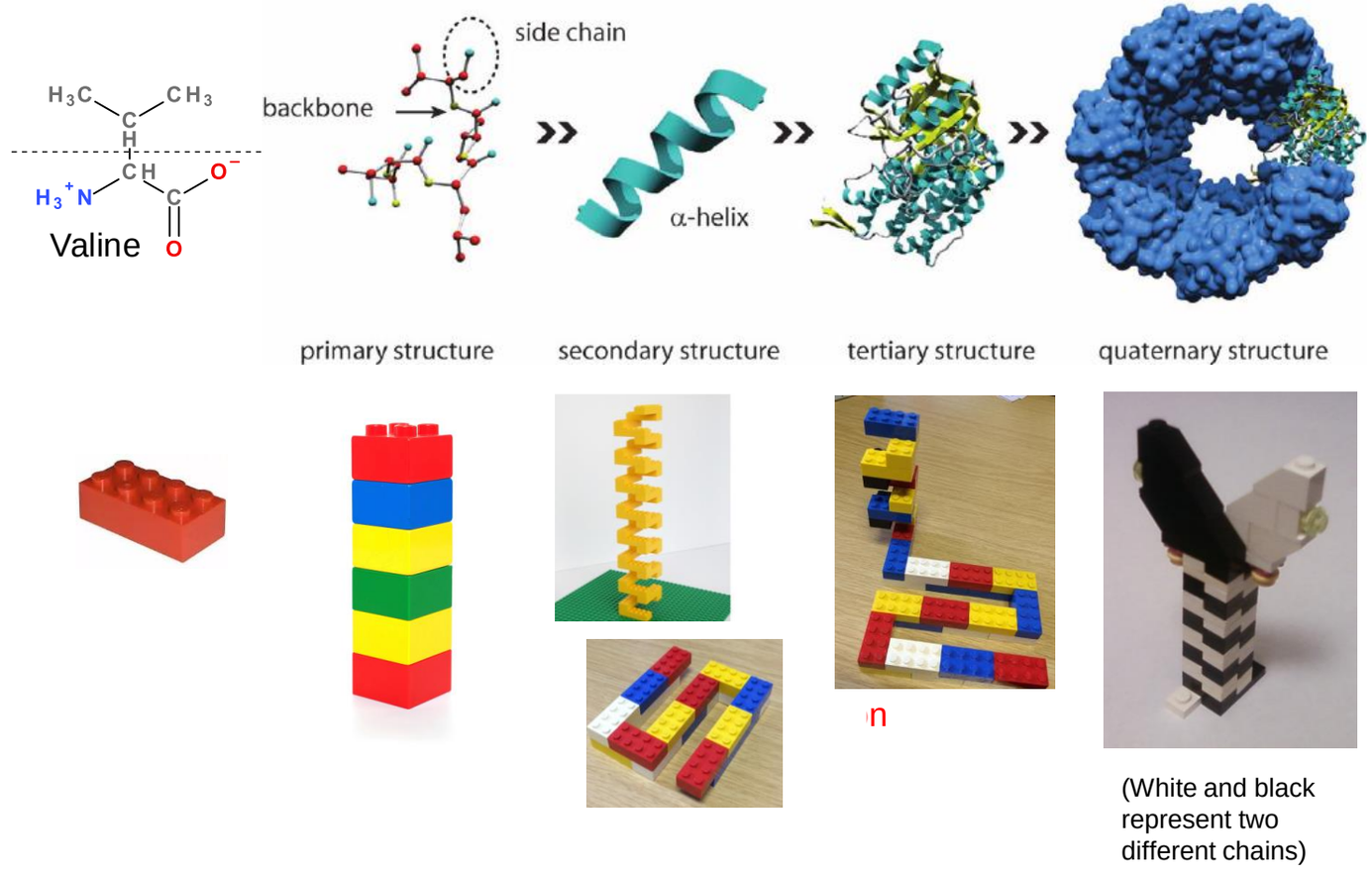
Proteins and Amino Acids



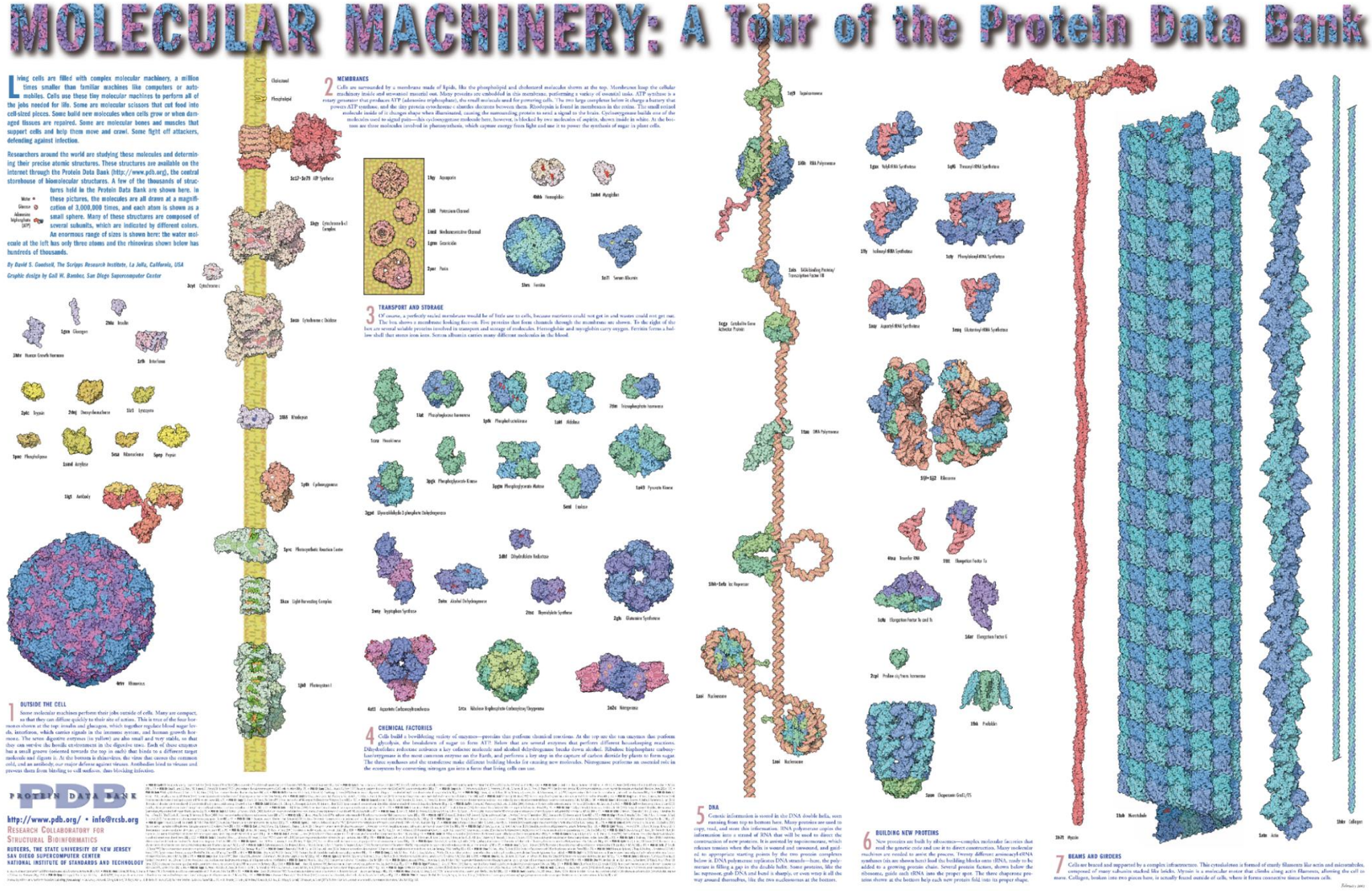


Structural Hierarchy of Proteins

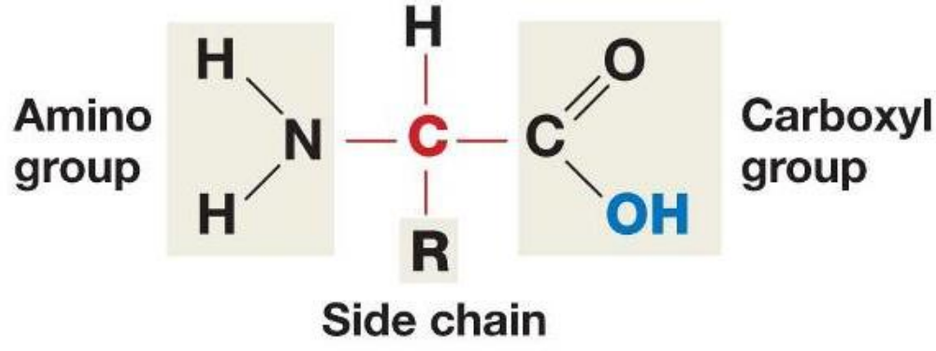
- Primary - sequence of amino acids, no 3D structural information
- Secondary - local structural elements, only mainchain atoms involved
- Tertiary - 3D position of ***all*** atoms, functional form of many proteins.
- Quaternary - multiple chains – multiple chains often required for function.



(White and black represent two different chains)

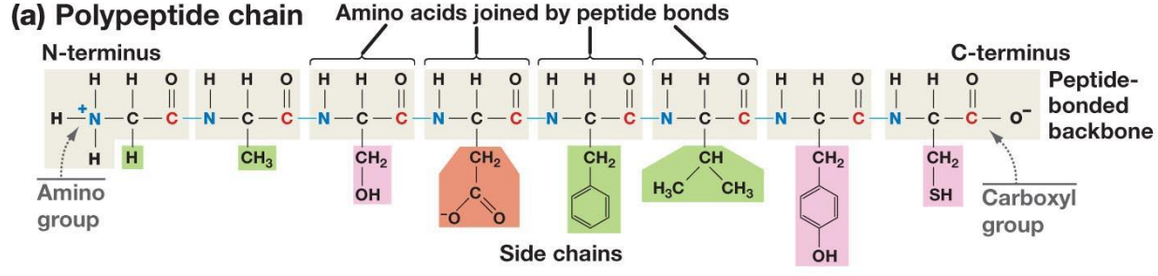

MOLECULAR MACHINERY: A Tour of the Protein Data Bank
Living cells are filled with complex molecular machinery, a million times smaller than familiar machines like computers or automobiles. Cells use these tiny molecular machines to perform all of the jobs needed for life. Some are molecular scissors 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 those 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. Bamba, 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 retinal 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.
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 catalyzes 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 it sharply, or even wrap it all the way around themselves, like the two nucleosomes at the bottom.
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.
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

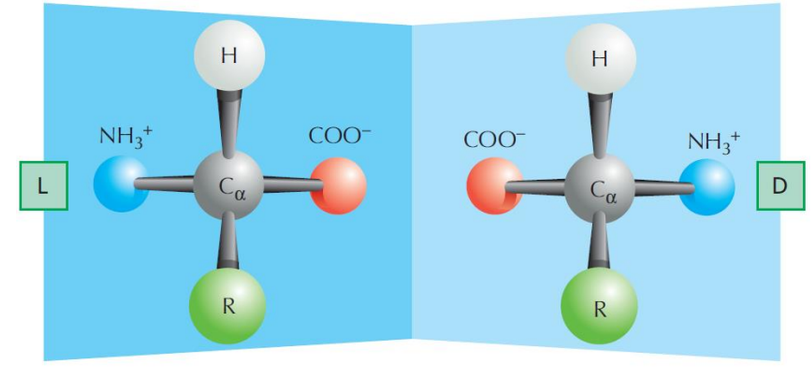
The Structure of Amino Acids and Proteins



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)

Primary Structure

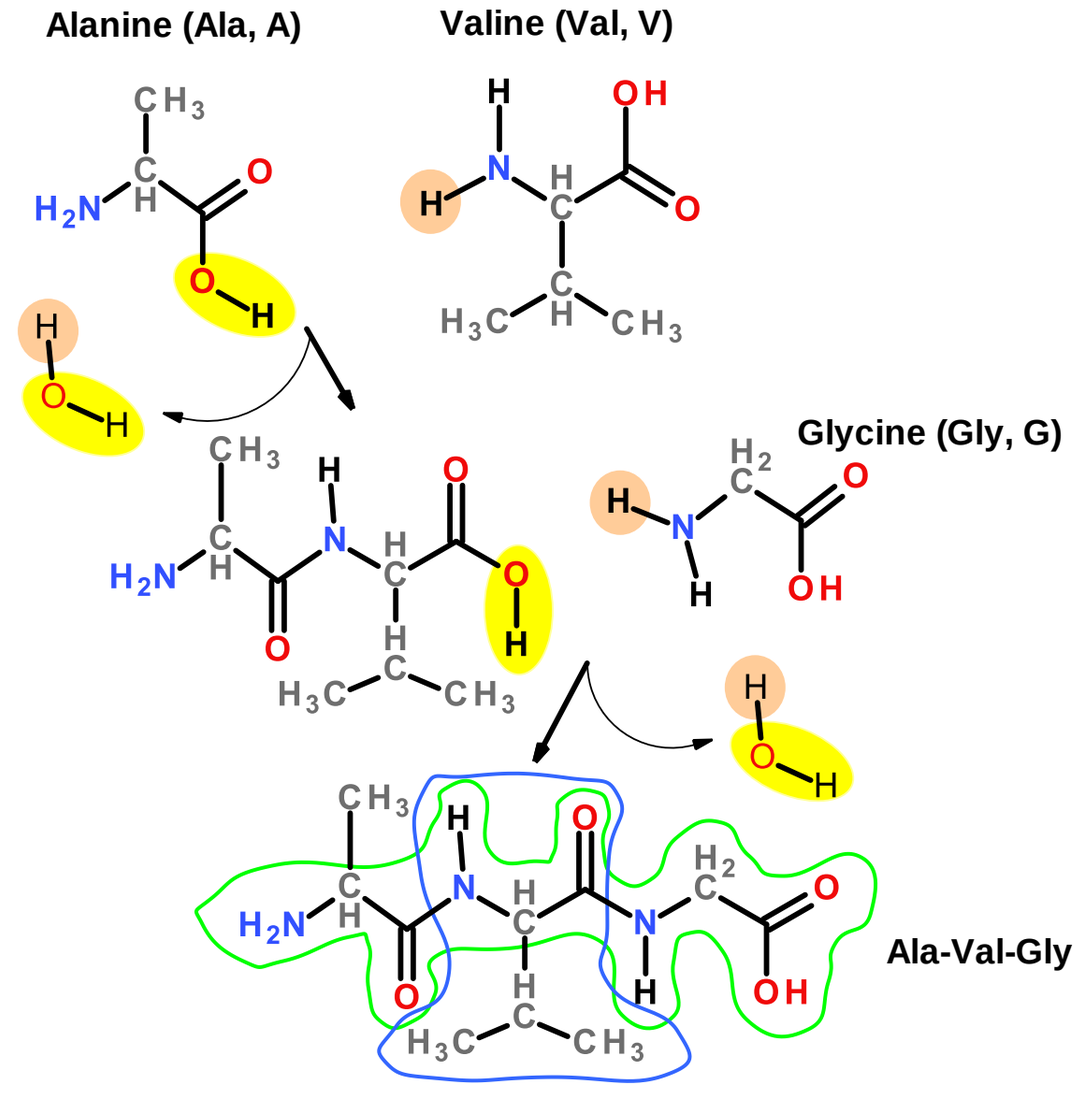
- Amino acids are joined together to form linear polymers by the formation of a **peptide bond** between the carboxyl of one amino acids and the amino group of the next.
- This reaction releases water: a **dehydration** reaction.
- The peptide bond can be broken (*lysis*) by the addition of water = **hydrolysis**.

Incorporated amino acid = *residue* (atoms are lost when the peptide bond is formed).

Polarity of chain direction – amino (N) terminus to carboxy(C) terminus = order of amino acids = *sequence* = *primary structure*

Mainchain (or backbone) – linear atoms of the polymer

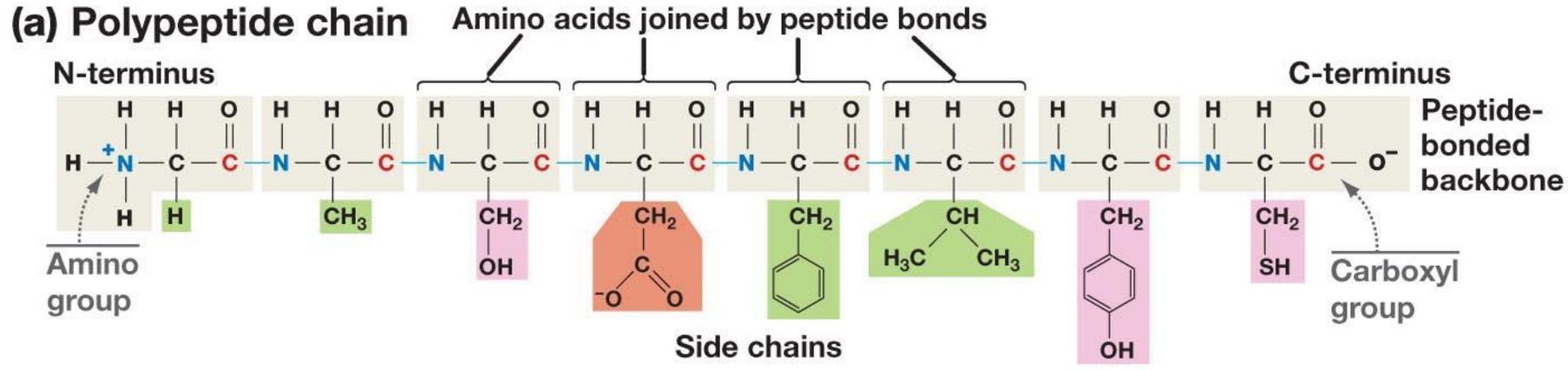
Sidechain – atoms off the Ca carbon



Primary Structure – Expectations

- Draw chemical structure given the sequence.
- Determine the seq. from chemical structure.
- Distinguish/identify:
  - **Mainchain** & **Sidechain** atoms,
  - **Residue** = aa in polymer,
  - N & C terminus,
  - Peptide bond(s).

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

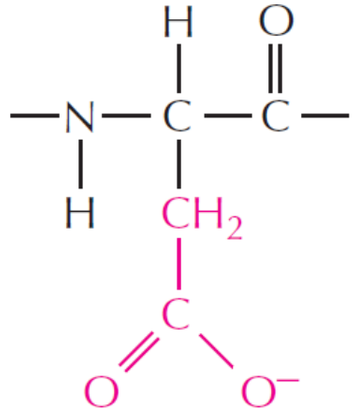


- Sidechains (R-groups) differ in their size, shape, reactivity, and interactions with water.
  1. Nonpolar Sidechains: **hydrophobic**; do not form hydrogen bonds; coalesce in water - typically form the core of folded proteins.
  2. Polar Sidechains: **hydrophilic**; form hydrogen bonds; readily dissolve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

ACIDIC SIDE CHAINS

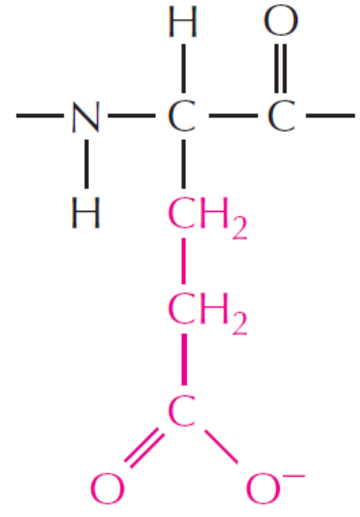
aspartic acid

(Asp, or D)



glutamic acid

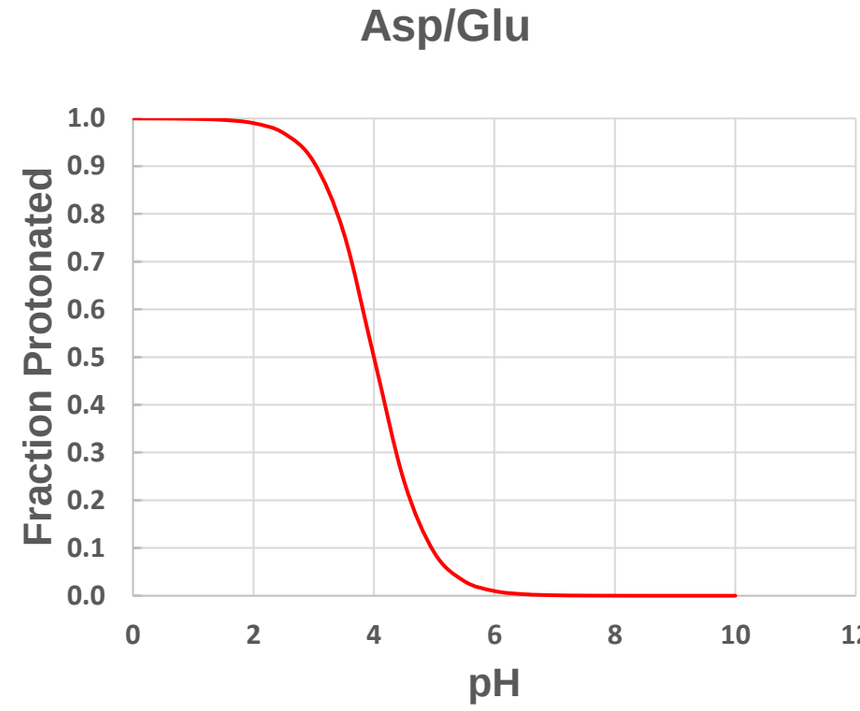
(Glu, or E)



pK_a of sidechain ~ 4

pK_a of sidechain ~ 4

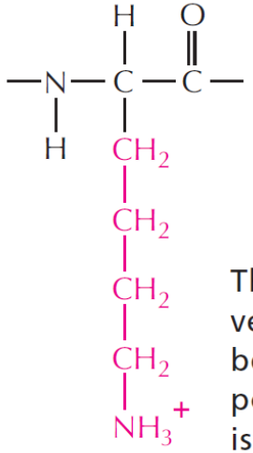
Have a net negative charge at pH 7.0



BASIC SIDE CHAINS

lysine

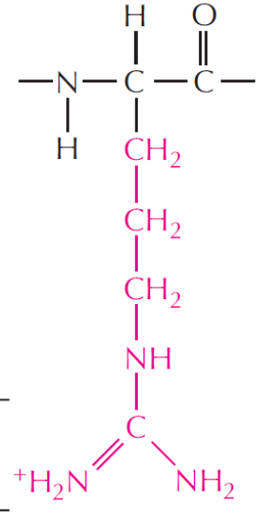
(Lys, or K)



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

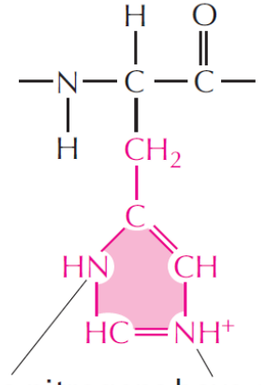
arginine

(Arg, or R)

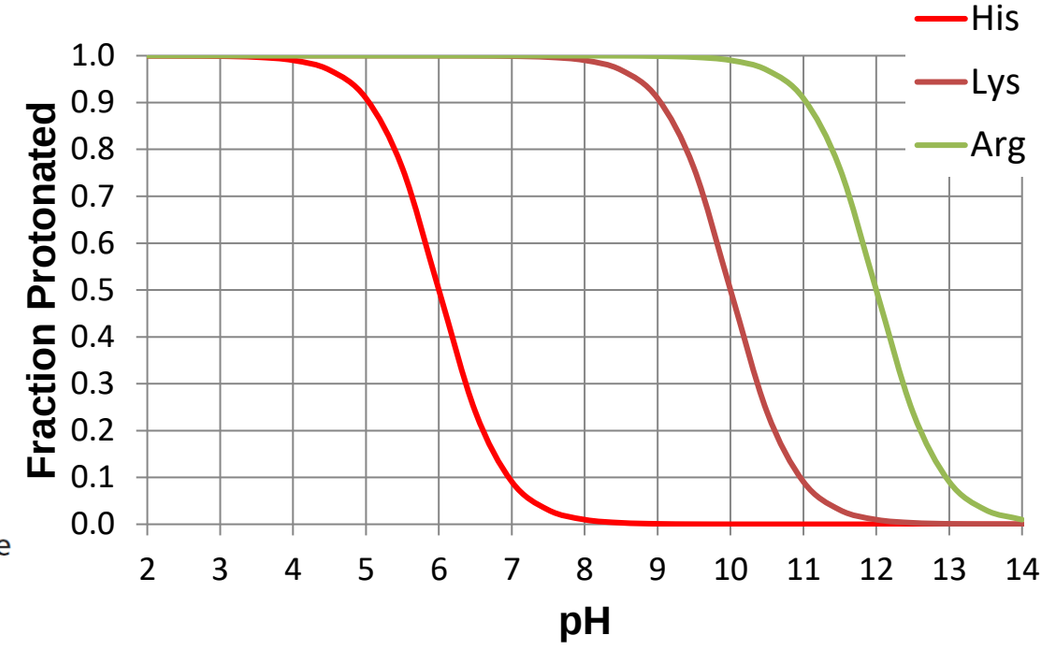


histidine

(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)

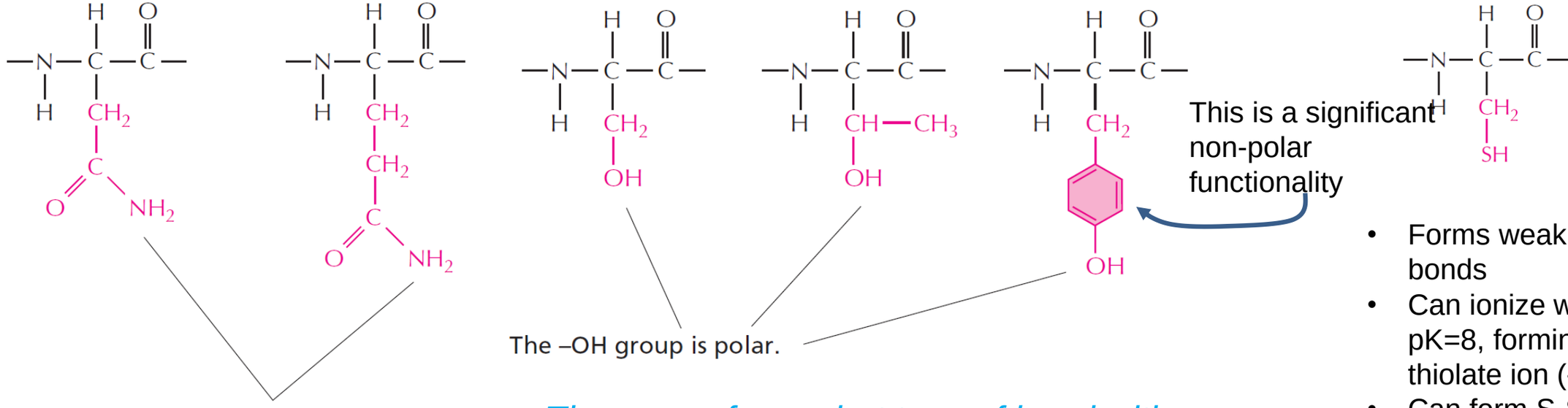
glutamine (Gln, or Q)

serine (Ser, or S)

threonine (Thr, or T)

tyrosine (Tyr, or Y)

cysteine (Cys, or C)



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

The –OH group is polar.

This is a significant non-polar functionality

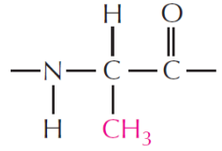
These can form what type of bond with water?

- 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

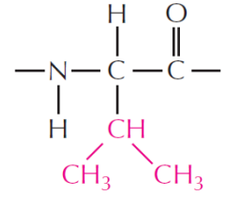
alanine

(Ala, or A)



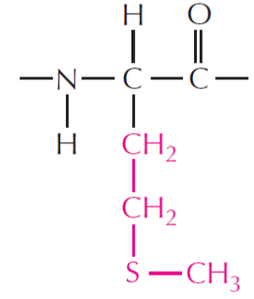
valine

(Val, or V)



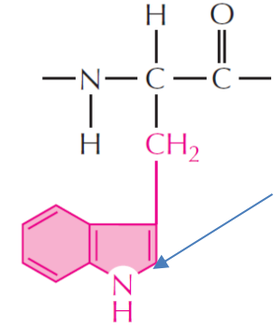
methionine

(Met, or M)



tryptophan

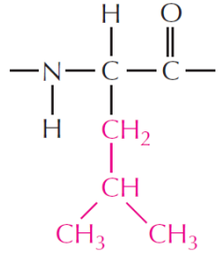
(Trp, or W)



H-bond donor
(polar functionality)

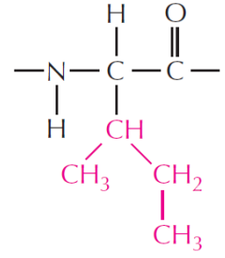
leucine

(Leu, or L)



isoleucine

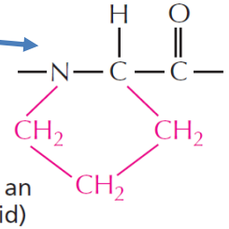
(Ile, or I)



Ring results in no NH group for H-bonding

proline

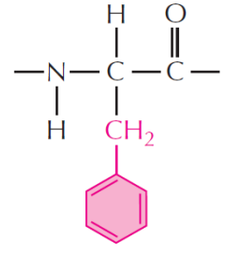
(Pro, or P)



(actually an imino acid)

phenylalanine

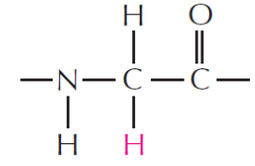
(Phe, or F)



AND LASTLY THERE IS GLYCINE:

glycine

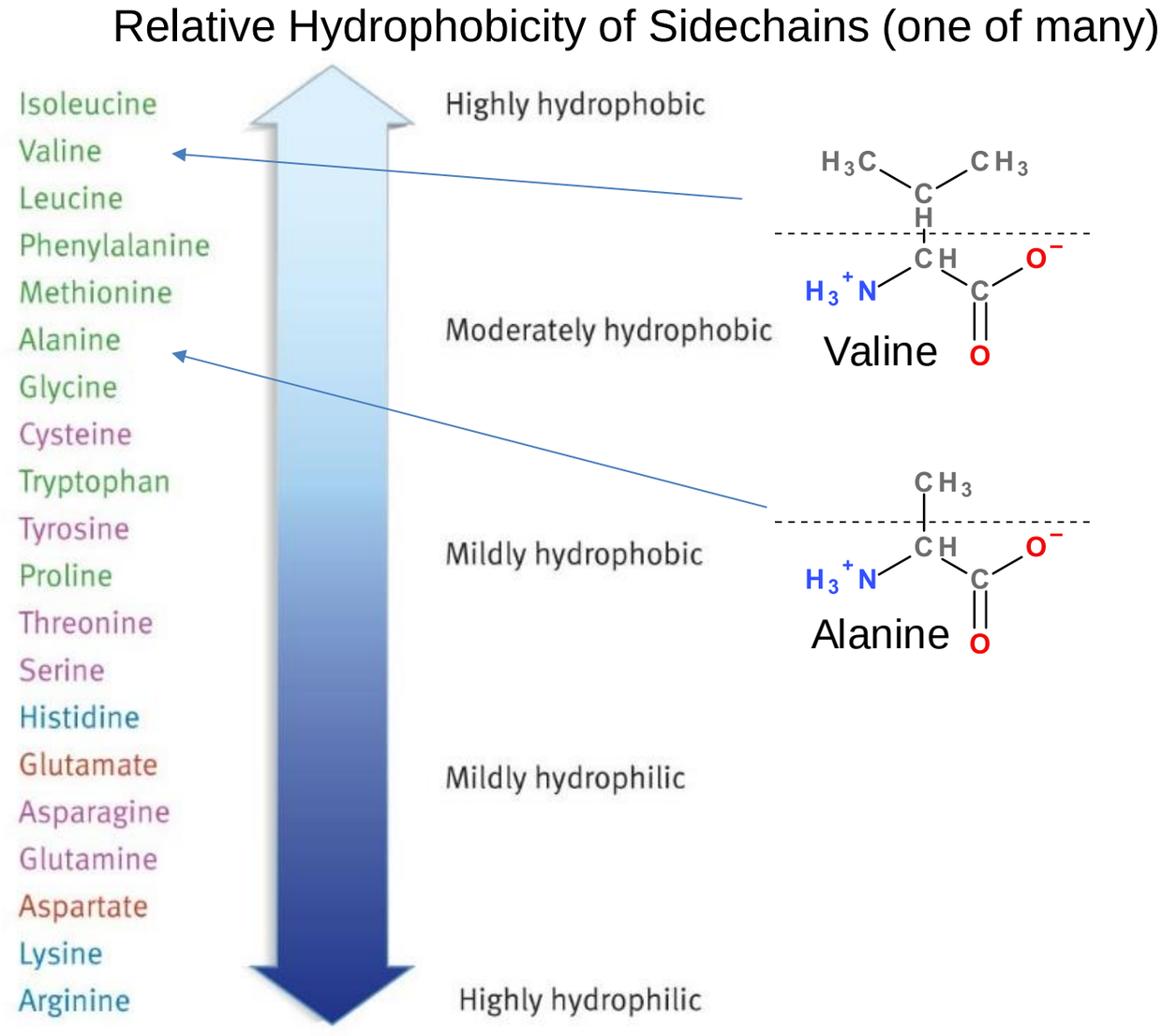
(Gly, or G)



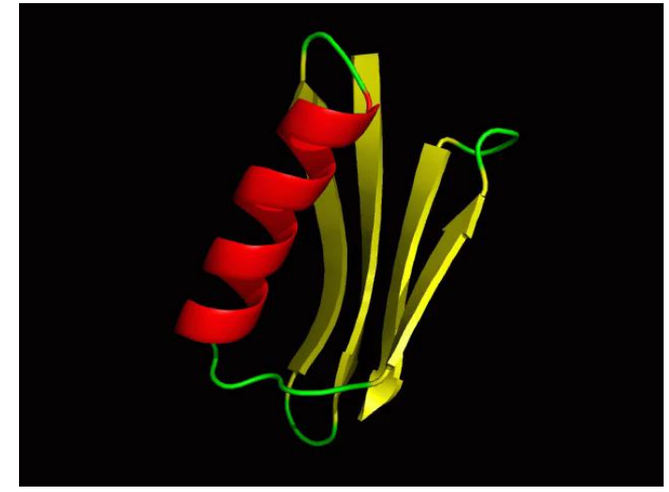
No real functionality for its R group (H)
Only AA that is achiral

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



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

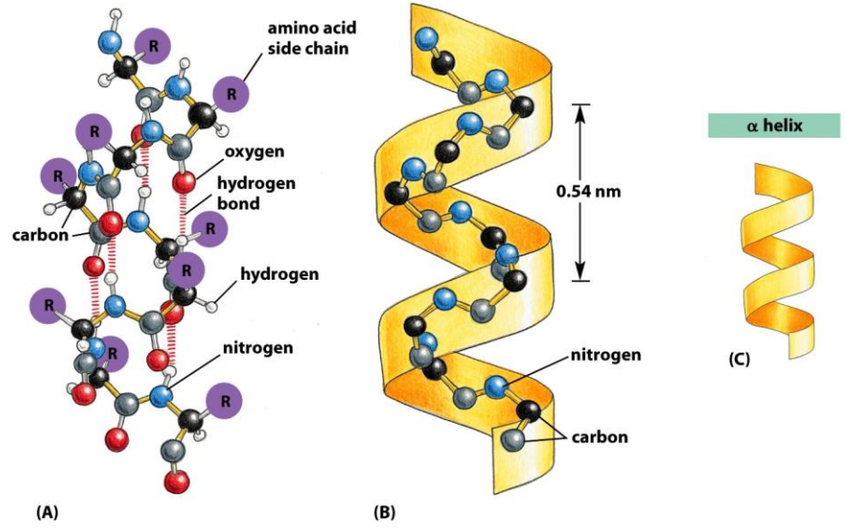
Mainchain hydrogen bonds

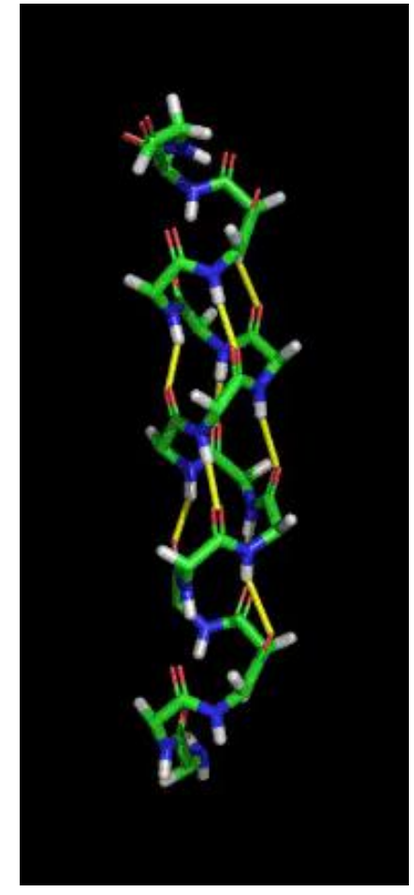
N-H O=C

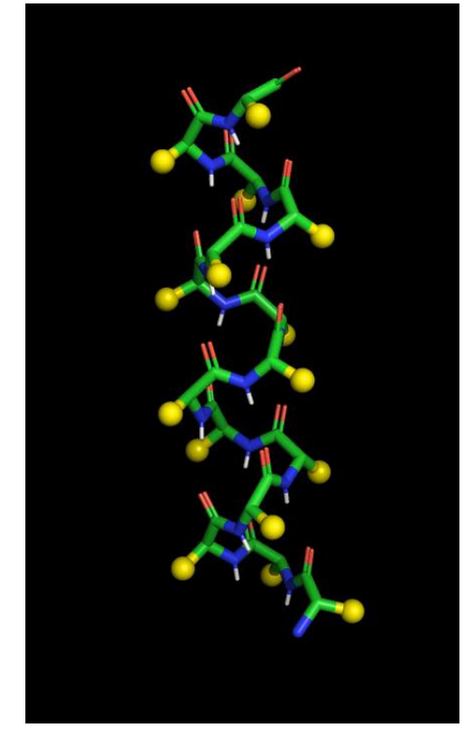
The NH is the hydrogen bond_____.

The C=O is the hydrogen bond _____.

Alpha Helix





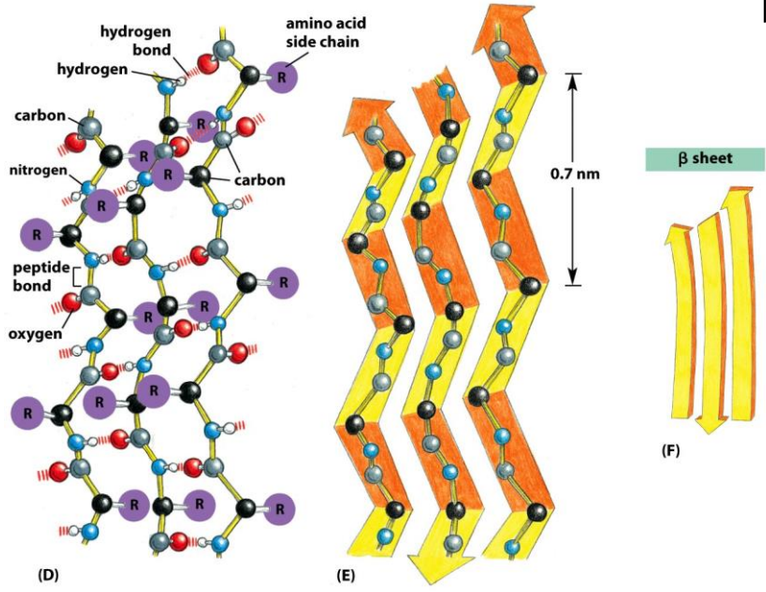


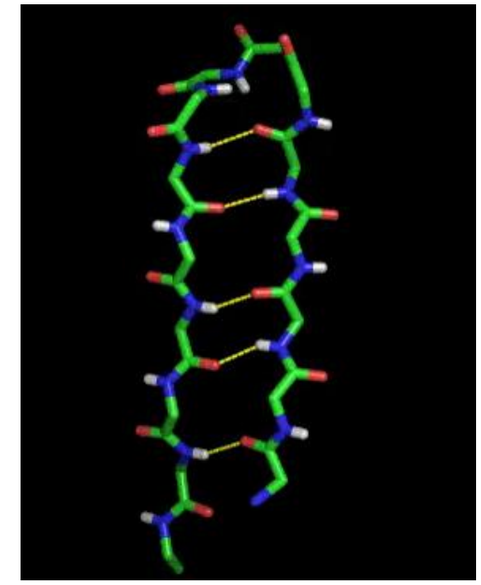
Spiral conformation (*helix*) in which every backbone N-H group donates a hydrogen bond to the backbone C=O group of the amino acid four residues earlier:

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

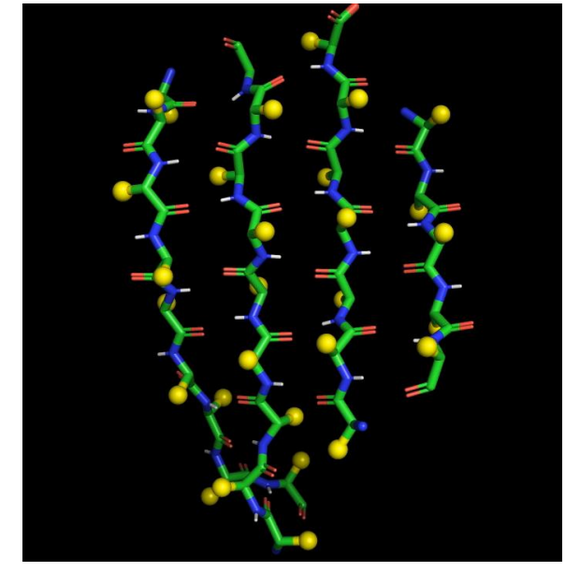
Side-chains project outwards.

Beta Sheet

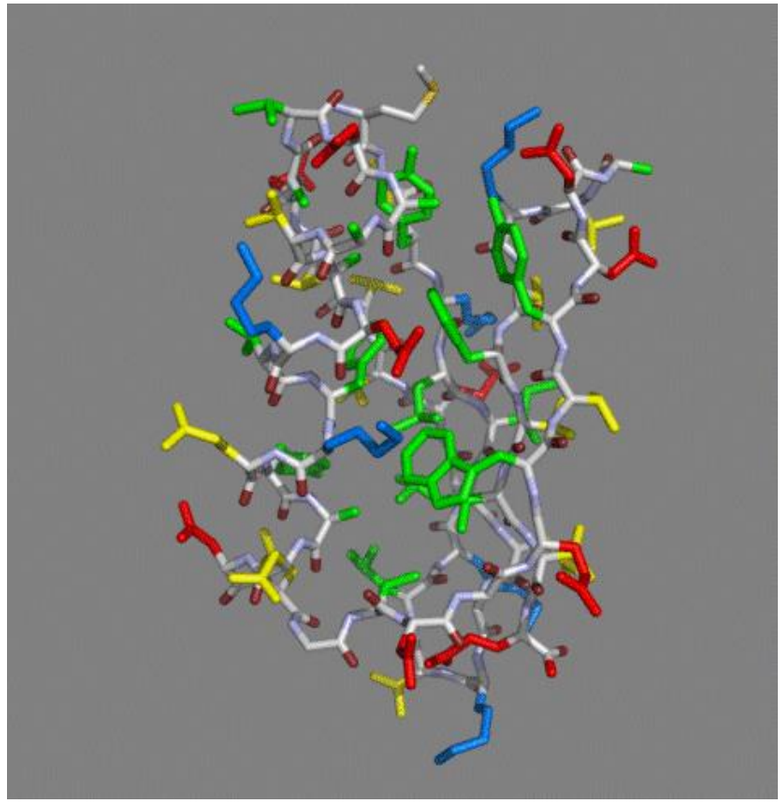




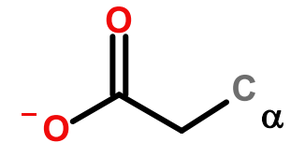
- Beta-Strands connected laterally by backbone hydrogen bonds that are perpendicular to the strand, forming a generally twisted, pleated sheet.
- Sheets can have two or more strands
- Side-chains:
  - project up and down along a strand.
  - project in the same direction going from strand to strand across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



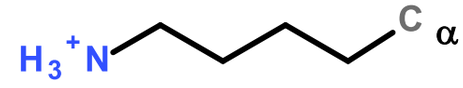
Tertiary Structure - Location of Residues in Globular Proteins



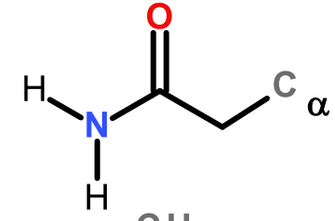
Red - amino acids with neg. sidechains (e.g. Asp)



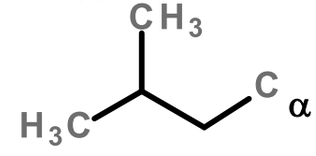
Blue - amino acids with pos. sidechains (e.g. Lys)

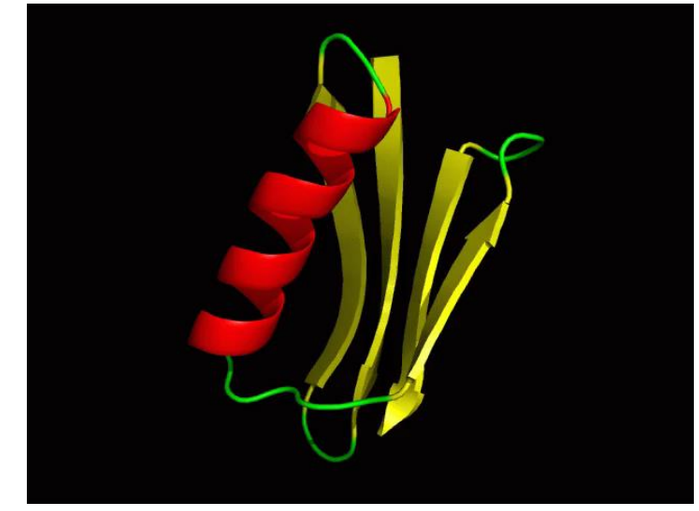


Yellow – amino acids with polar sidechain (e.g. Asn)

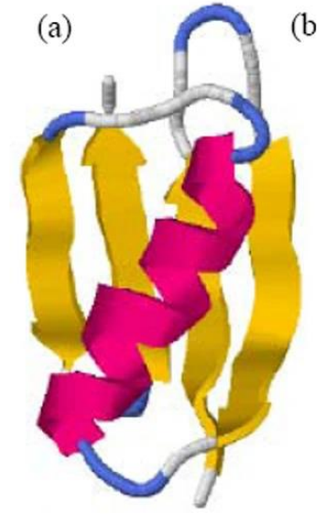


Green - amino acids with hydrophobic side chains (e.g. Leu)





Amino Acid Type		
	Inside (I)	Surface (S)
Charged		
Polar		
Non-polar		



Protein Stability:

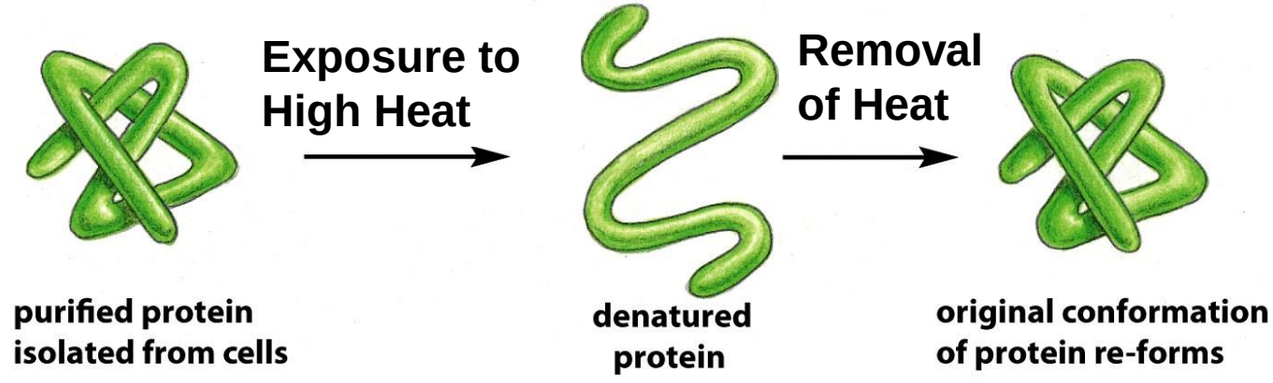
H-bonds
van der Waals
Hydrophobic effect

Native

Chain disorder

Unfolded

Protein Denaturation

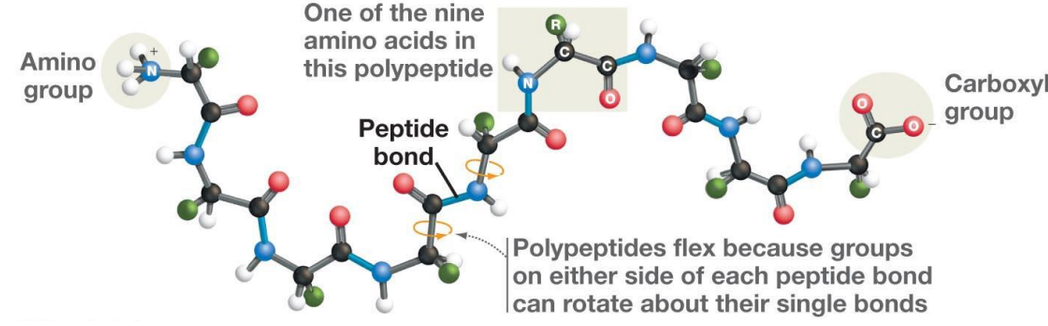


- Often, unfolded protein aggregate, which prevents refolding.

Unfolded Polypeptides Are Flexible – High Entropy stabilizes the Unfolded state

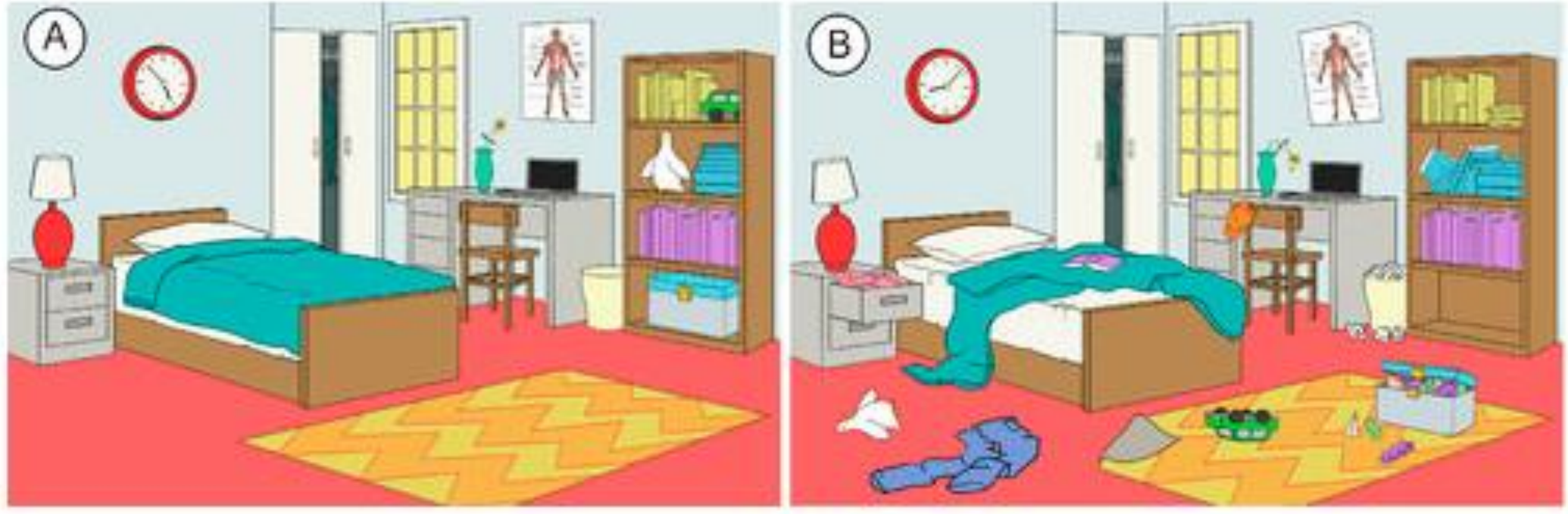
purified protein
isolated from cells

denatured
protein

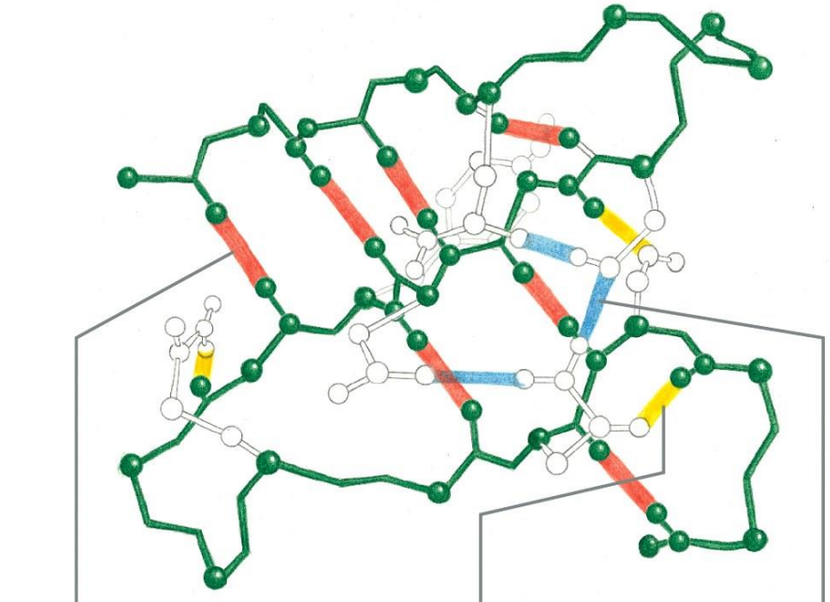


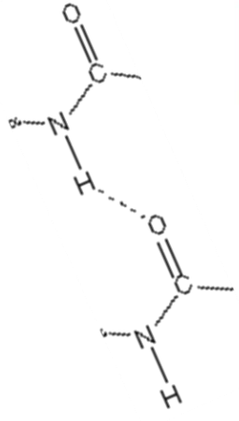
Entropy

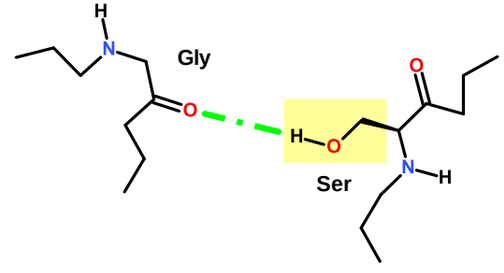
Energy and Entropy

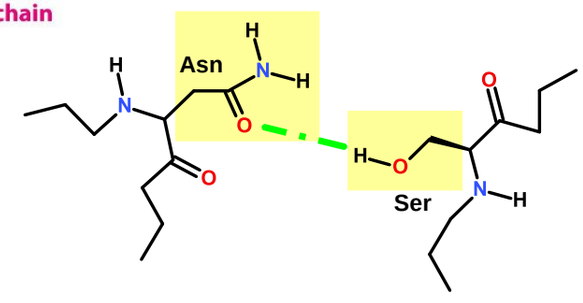


Hydrogen Bonding Stabilizes the Tertiary Structure





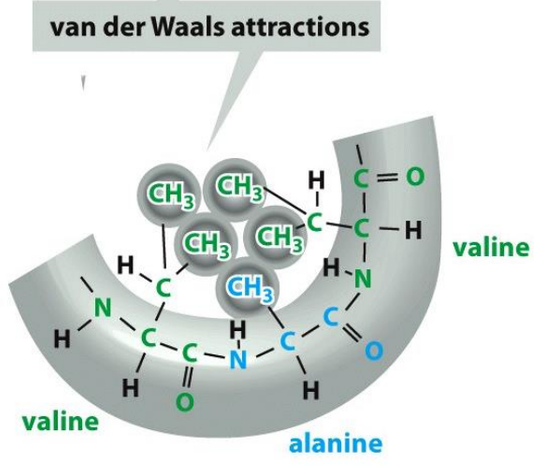


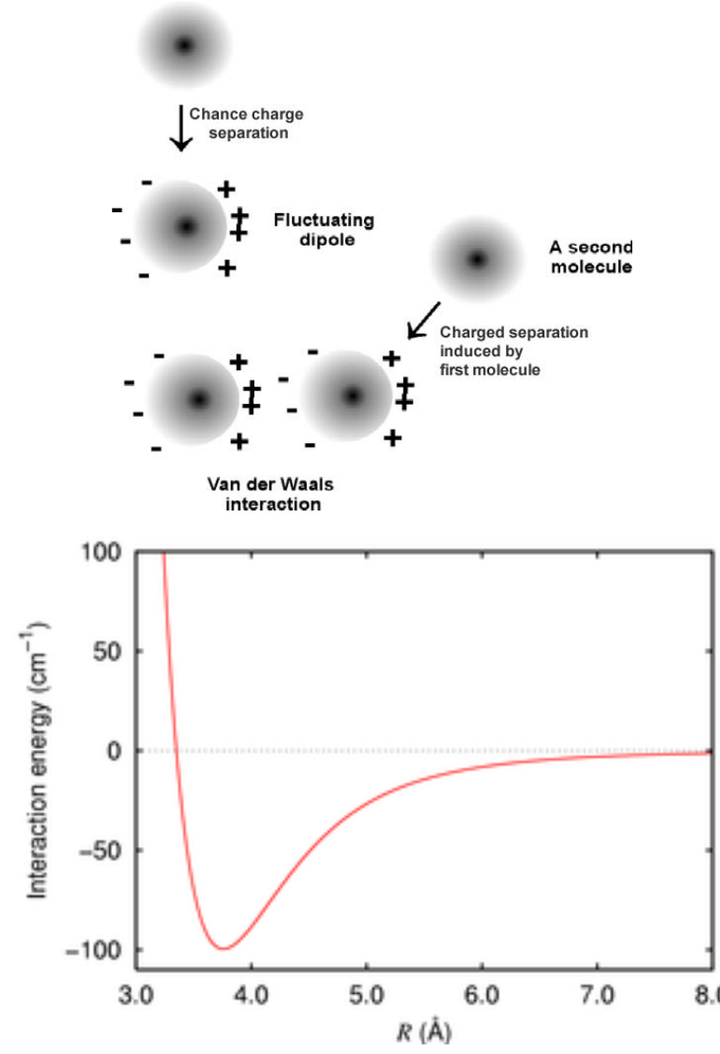


- **Hydrogen bonds** form between hydrogen atoms and the carbonyl group in the peptide-bonded backbone – secondary structure
- Hydrogen bonds are also found between hydrogen and electronegative atoms in side chains (sidechain-sidechain)
- Sidechains can form hydrogen bonds to the mainchain too.

Van der Waals (VdW) interactions Stabilize the Folded State

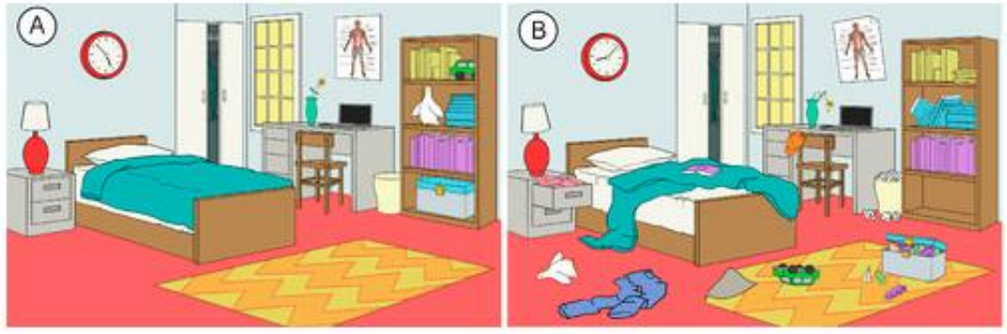
- VdW are weak electrostatic interactions between side chains due to temporary (fluctuating) charges.
- Attractive from long distance
- Distance at lowest energy is at the van der Waals radii of the atoms.
- Optimized in the core of folded proteins by “knobs fitting into holes”
- Strength proportional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



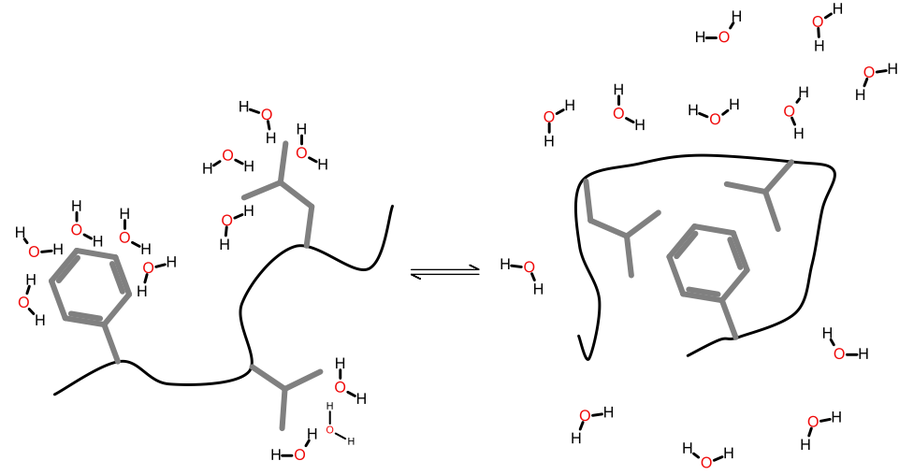


Hydrophobic Interactions are Critical for Stabilizing the Folded Structure

Energy and Entropy



Ordered water hydrating a non-polar group



Hydrophobic interactions within a folded protein increase stability of the folded protein by releasing the ordered water that surrounded exposed non-polar groups in the unfolded protein. *Folding increases the entropy of the water – favorable.*

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.**

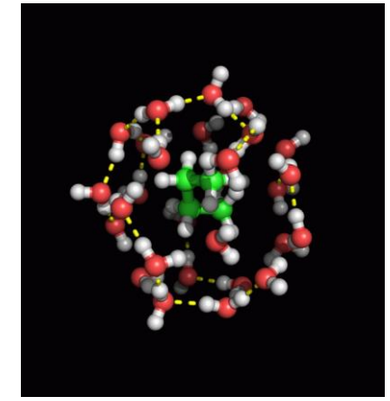
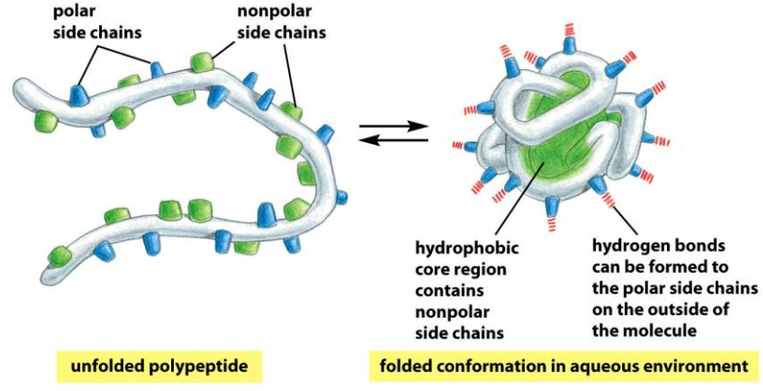


TABLE 3.1 **How Amino Acids Interact with Water**

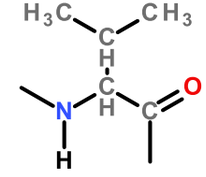
20 amino acids are ranked according to how likely they are to interact with water. Color codes are based on Figure 3.3.

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	Highly hydrophilic
Lysine	
Arginine	

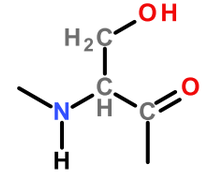


Which amino acid is most likely to be found in the core of a folded protein:

Valine

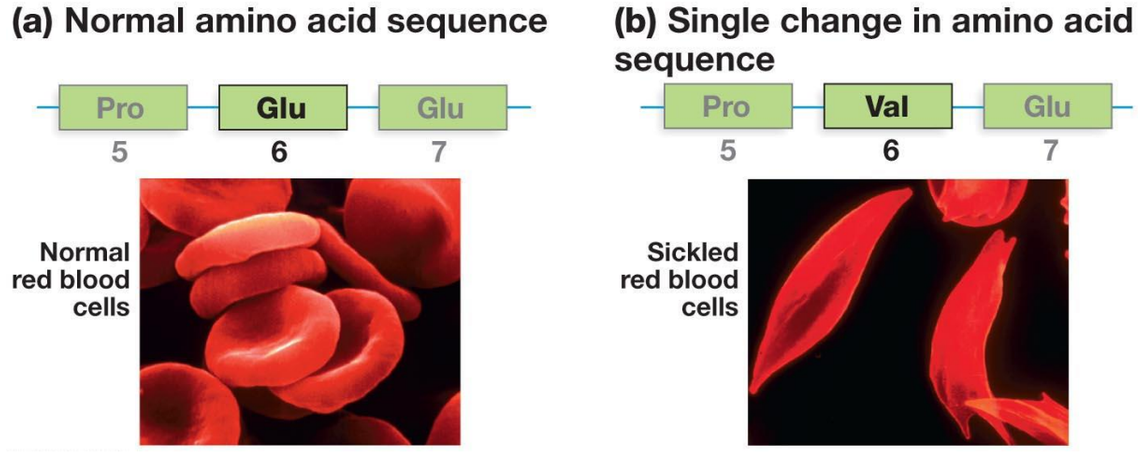


Serine



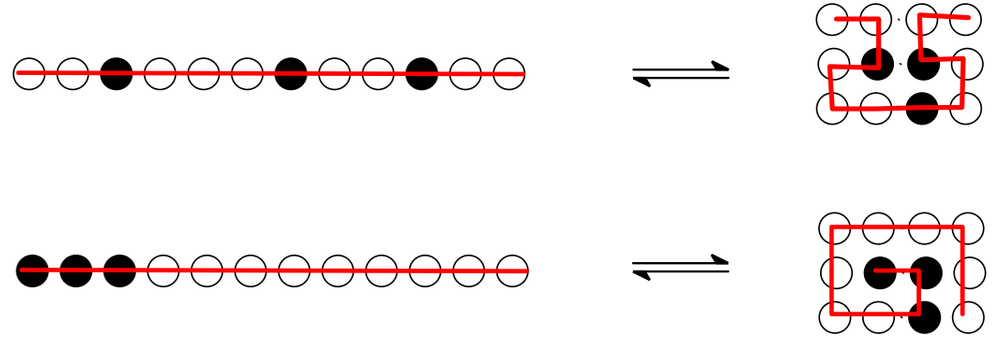
Fold Depends on Amino Acid Sequence

Effect of mutations on protein folding – sickle cell anemia

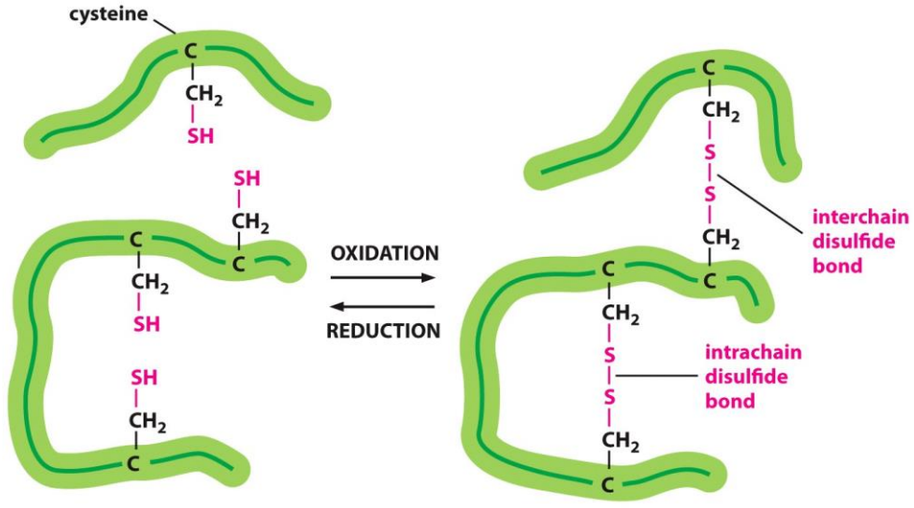


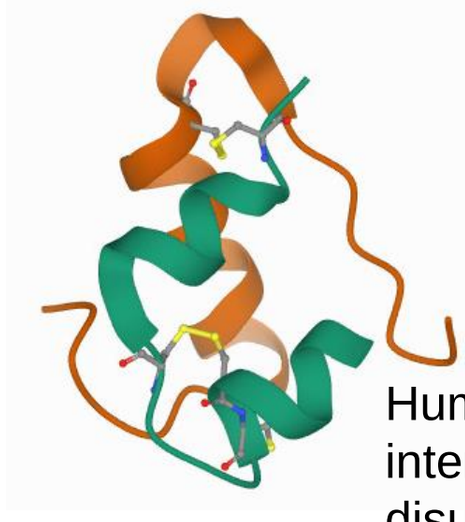
A single change in the amino acid sequence can change the function of a protein, and often affecting how it folds.

The *position* of non-polar residues (filled circles) mostly affects the final fold:

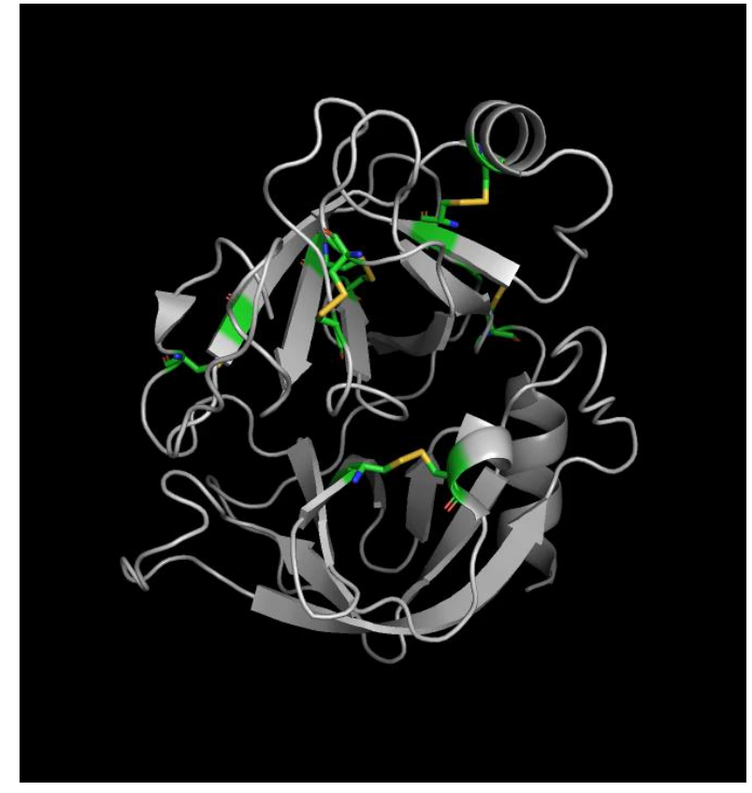


Disulfide Bonds Stabilize Some Proteins Outside the Cell (and body)





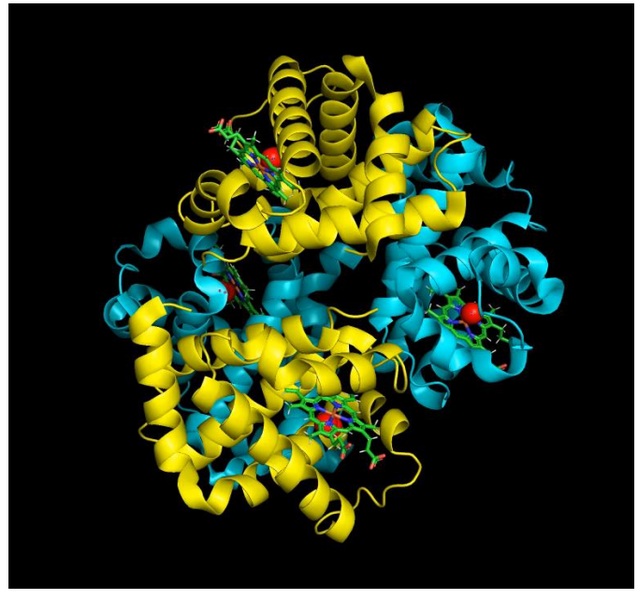
Human Insulin – interchain disulfide bonds



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

Quaternary Structure

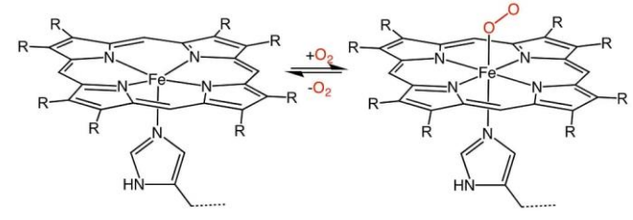
- Combinations of polypeptide subunits (combinations of tertiary structures).
- May be held together by covalent bonds (disulfide), but usually non-covalent interactions between amino acids on the different chains.
- Proteins can be a dimer, a tetramer, etc.
- If the chains are the same, called homo_____.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



Quaternary structure of hemoglobin (oxygen transport protein):

- two α chains
- two β chains

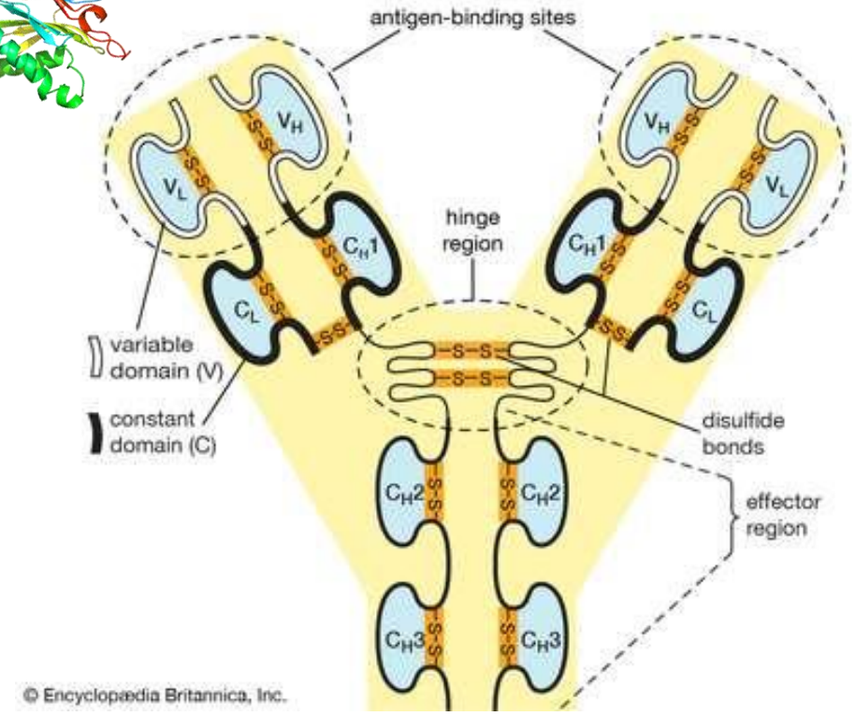
Oxygen is carried on Fe^{2+} within heme groups:

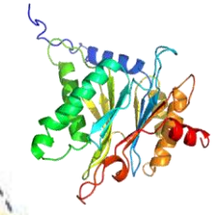


Antibodies – Produced by the Adaptive Immune system to Fight Pathogens.

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.

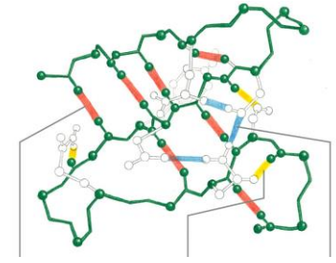
Summary - Interactions that Stabilize Folded Proteins.

- **Hydrogen bonds** form between hydrogen atoms (NH) and the carbonyl group in the peptide backbone (mainchain), and between and donors and acceptors on sidechains. *Mainchain-mainchain H-bonds are responsible for secondary structures.*
- **Hydrophobic interactions** within a protein increase stability of the folded state by *increasing entropy due to the release of water that was ordered by the exposed non-polar groups in the unfolded protein.*
- **van der Waals interactions** are *optimized in the well packed core of the protein*.

- **Covalent disulfide bonds** form between sulfur-containing cysteine residues stabilizing them (usually only exported, secreted proteins).

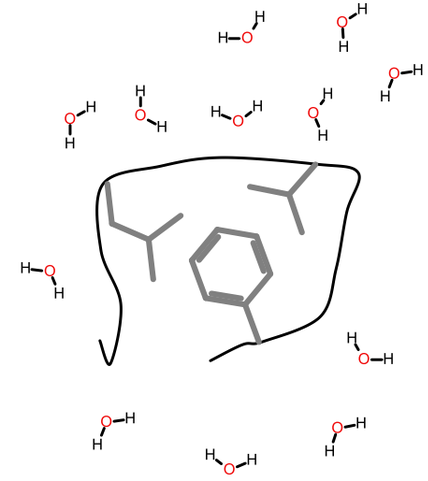
H-bonds
van der Waals
Hydrophobic effect

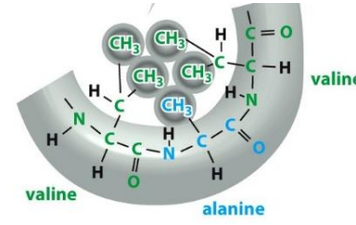
Chain disorder

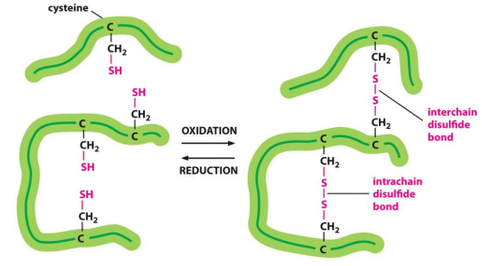


Native

Unfolded

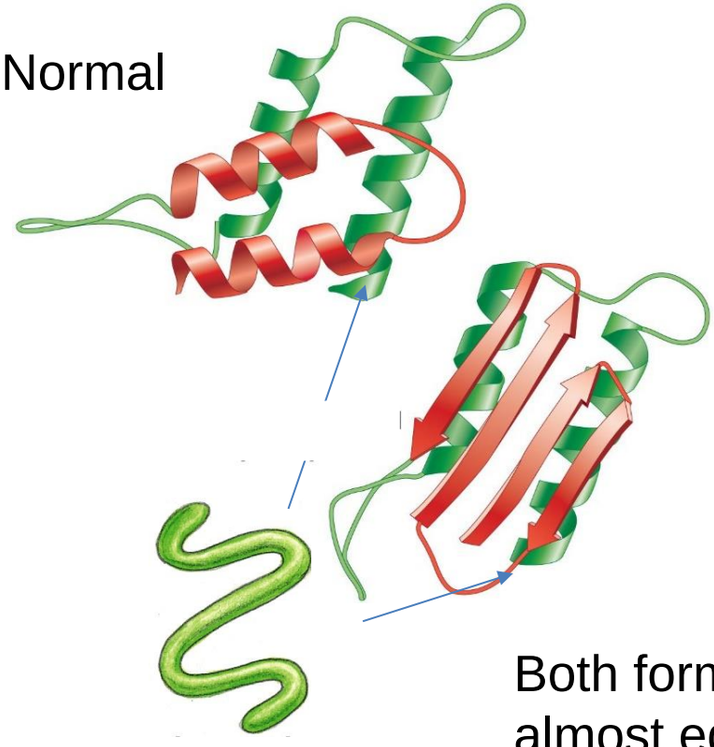




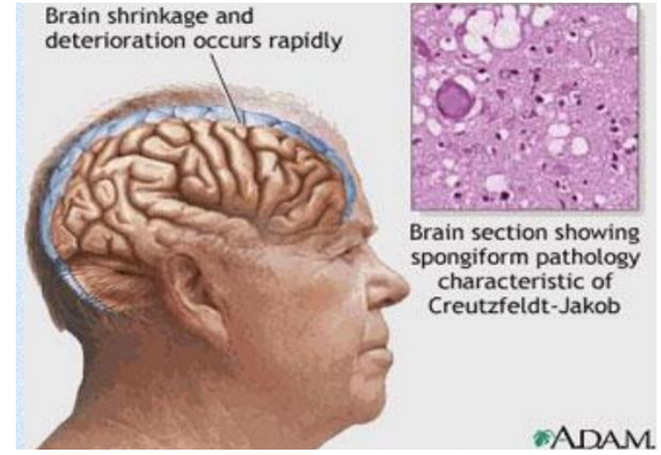


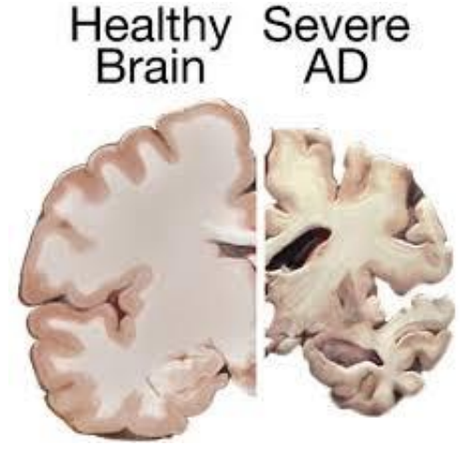
What happens when proteins don’t fold properly?

Prions are improperly folded proteins that cause neurodegenerative diseases



Both forms are almost equally stable.





What is the effect on the brain?

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

Why do the brain cells die?

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

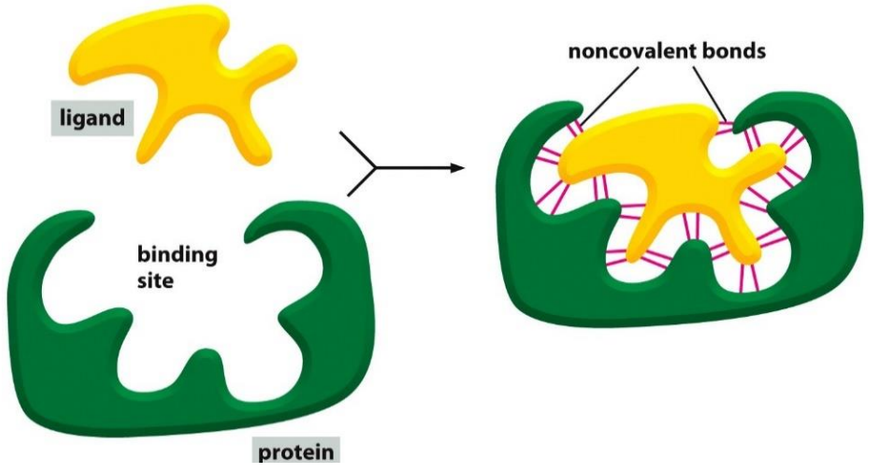
Tertiary Structure:

- 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 Interactions:



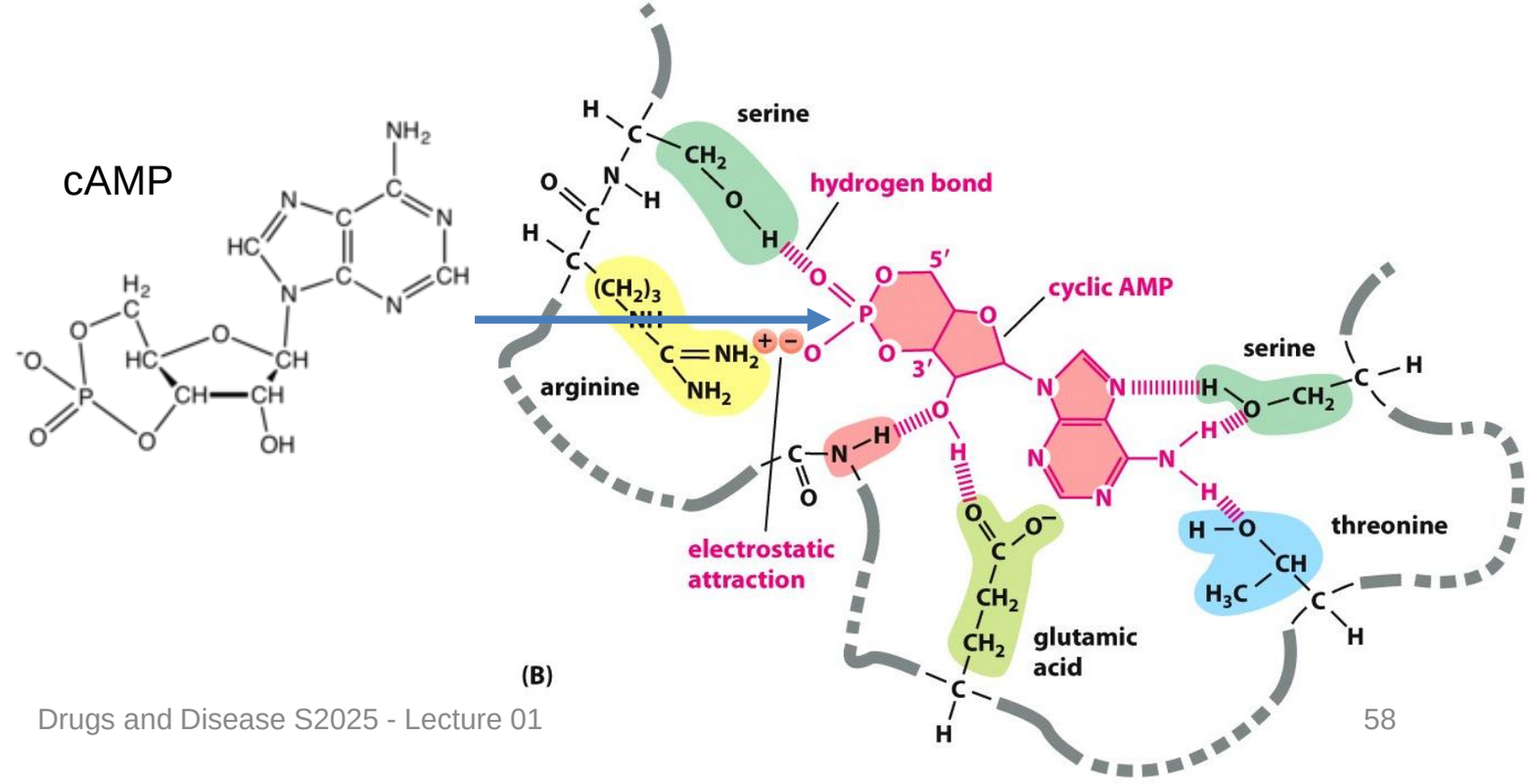
Ligand: Something that binds to a protein, usually small molecules (e.g. cyclicAMP, cAMP).

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

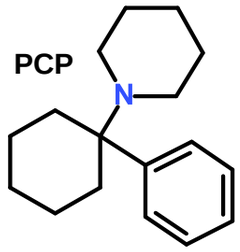
Binding site is generated by the **folded** form of the protein.

The bound ligand can be stabilized by any and all of:

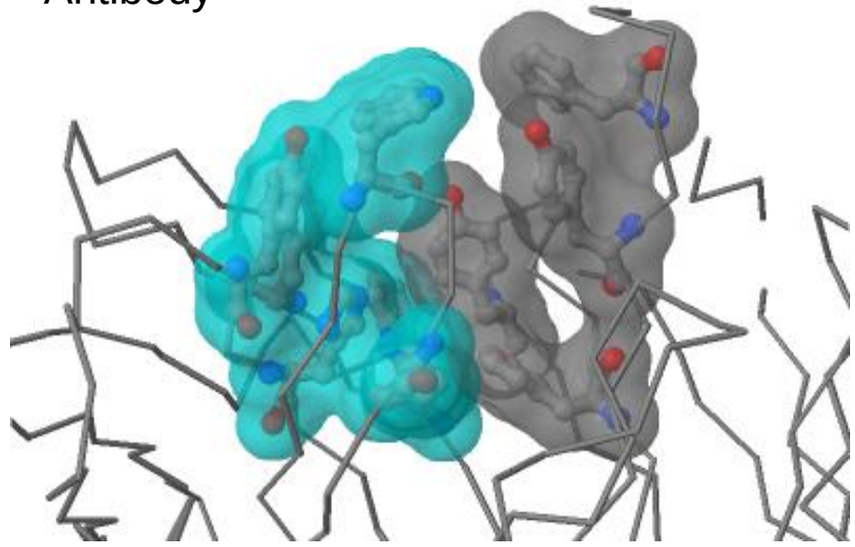
Interaction	Which stabilize cAMP Binding?
van der Waals	
H-Bonding	
Electrostatic	
Hydrophobic effect	



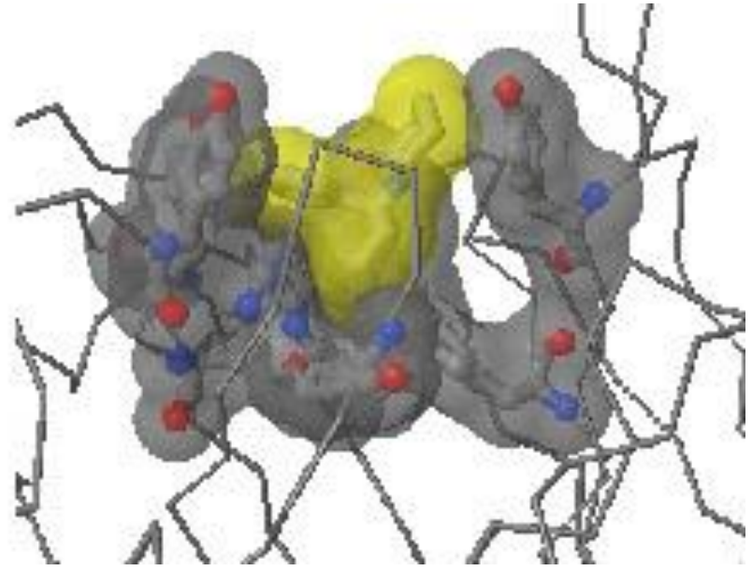
Hydrophobic Effect and Ligand Binding

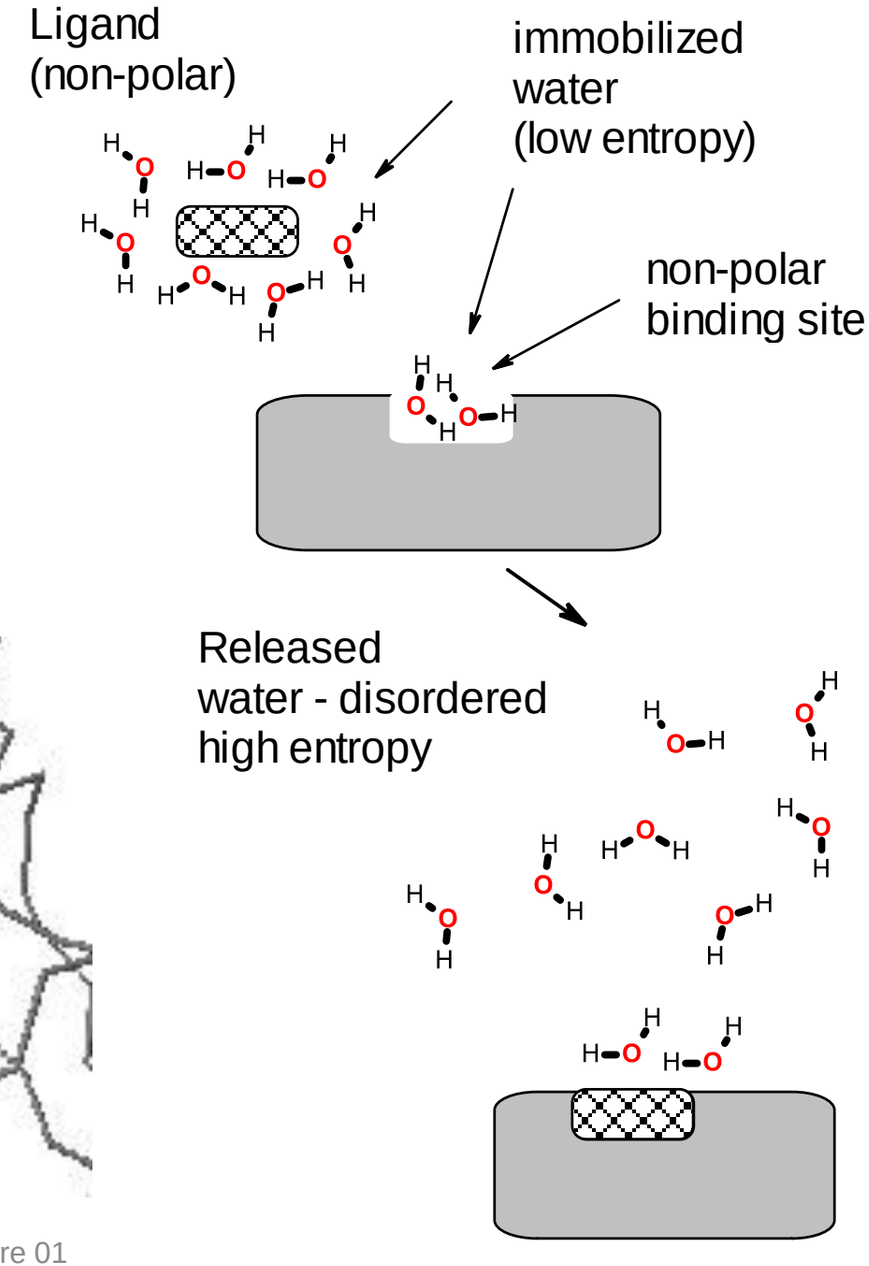


Antibody



Antibody-PCP complex

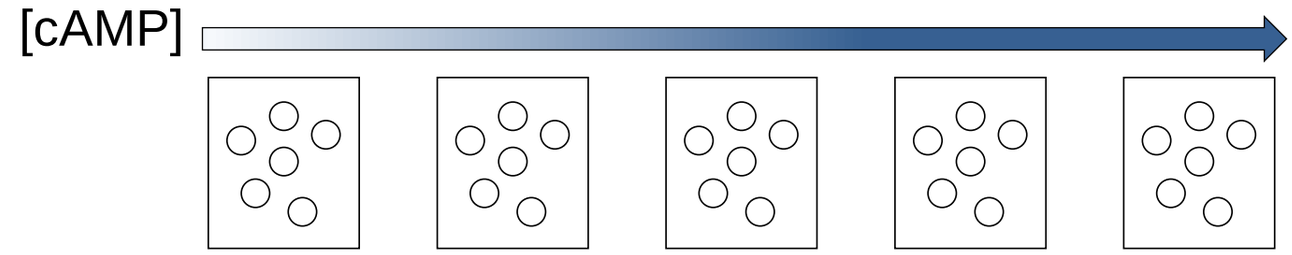


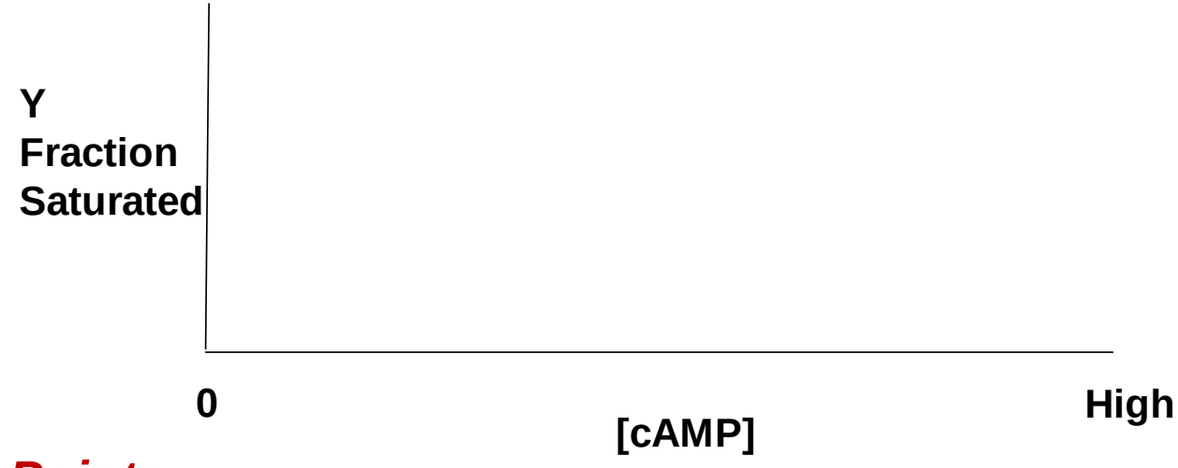


Ligand Binding & Saturation:

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. *How will the number of filled circles depend on the cAMP concentration?*

2. *Plot the location on the fraction saturated curve for each box.*

Key Points:

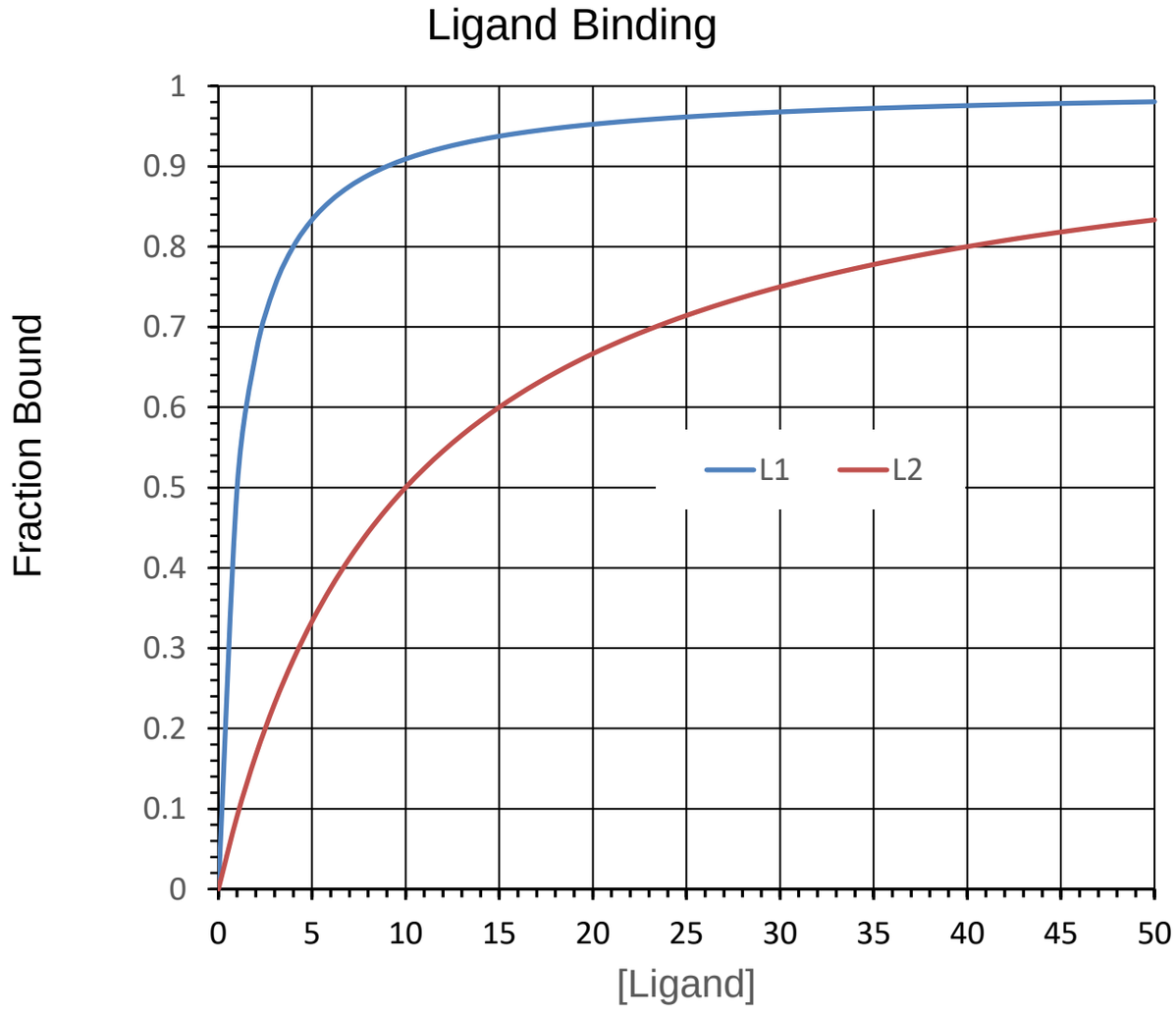
1. The binding sites saturate, when all are full no more ligand can bind.
2. There is a ligand concentration, [L], where ½ the sites are full. This [L] is K_D
3. K_D is the equilibrium constant for ligand dissociation:

$$K_{Eq} = \frac{[products]}{[reactants]}$$

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

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

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?

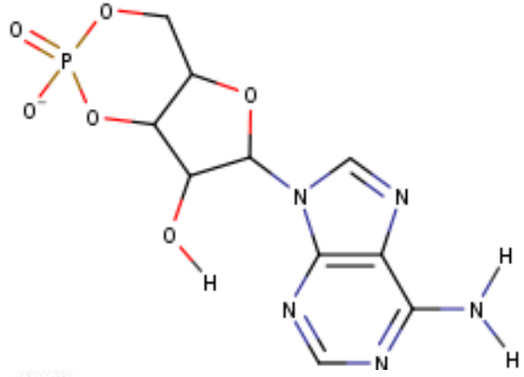
Which ligand has a K_D of 10? L1 or L2?

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

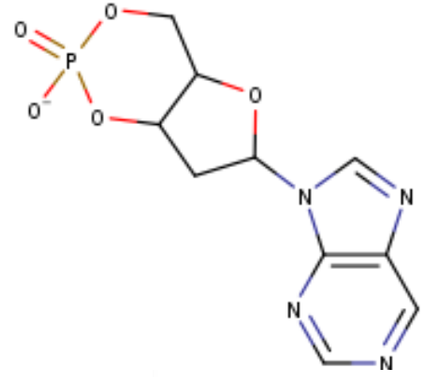
***Why** does L1 bind more tightly (higher affinity)?*

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

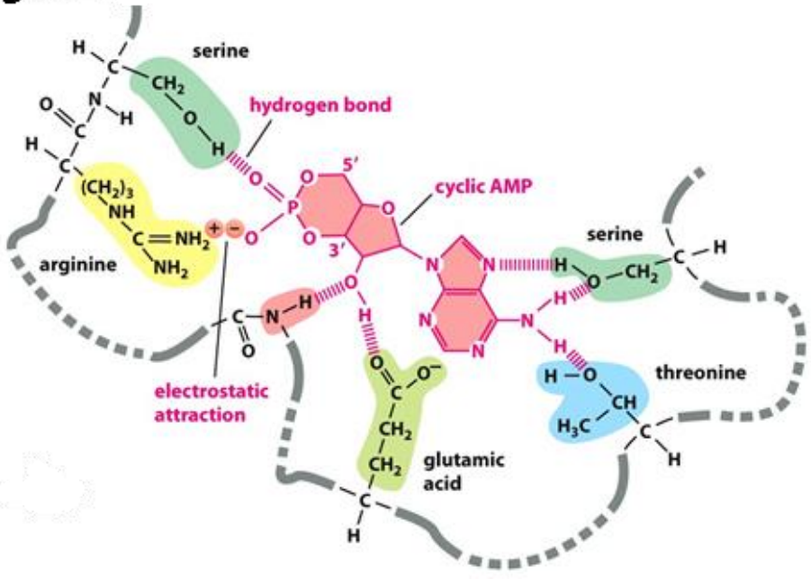
Ligand 1 (cAMP)



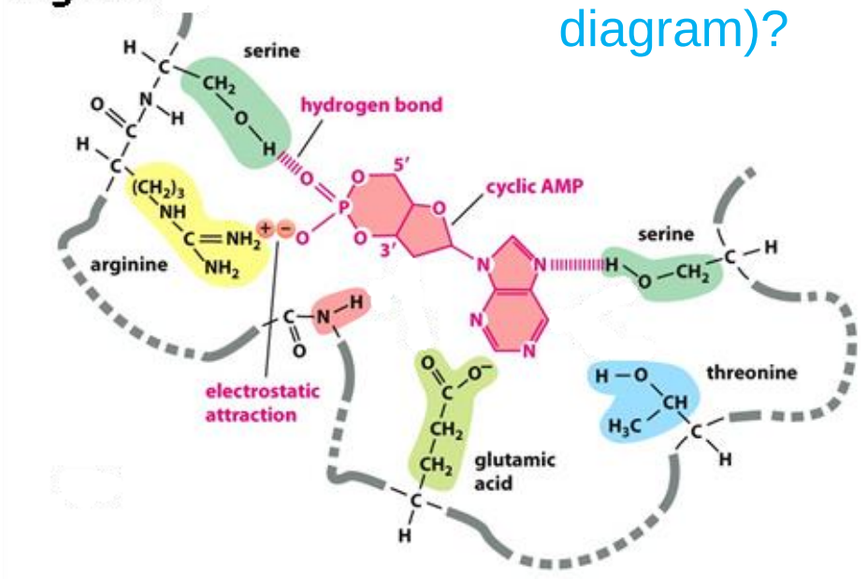
Ligand 2



Ligand 1



Ligand 2



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

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:

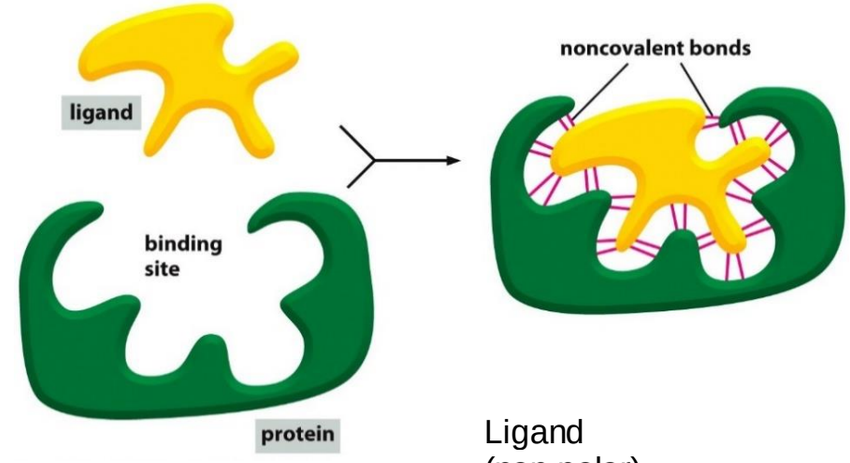
- H-bonds,
- van der Waals,
- Electrostatic,
- Non-polar interactions (hydrophobic)

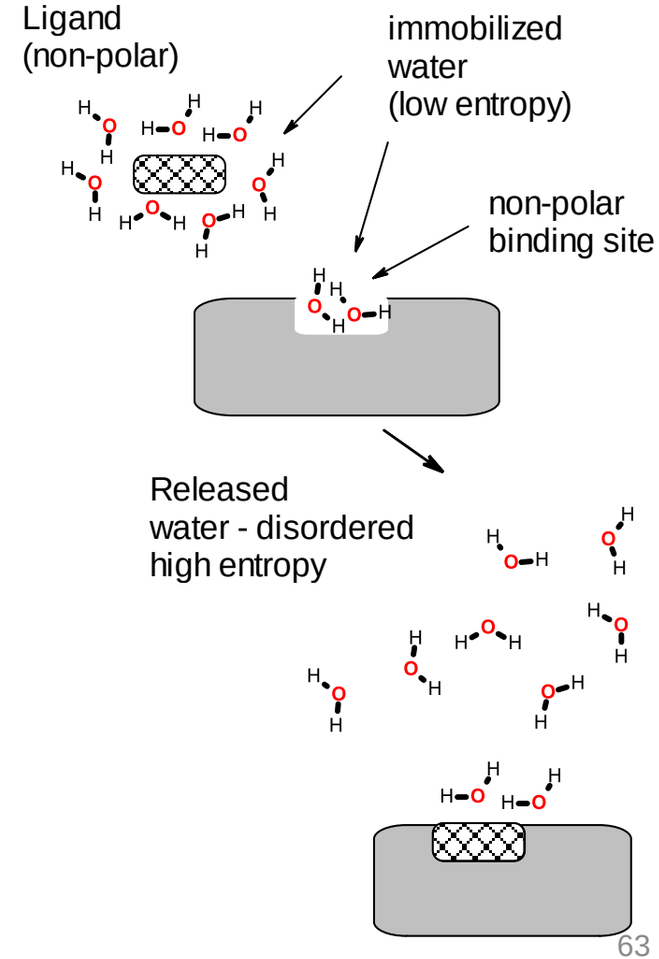
Binding is **reversible**

Binding is **saturable**

Binding ½ point (Y=0.5) occurs at K_D

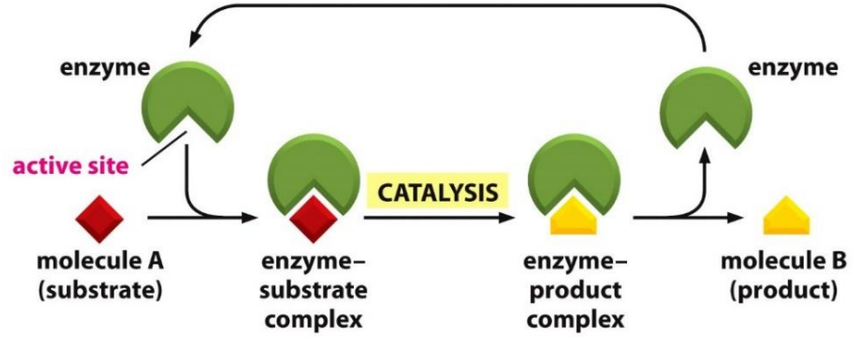
The higher the affinity (strength of interaction), the lower the K_D

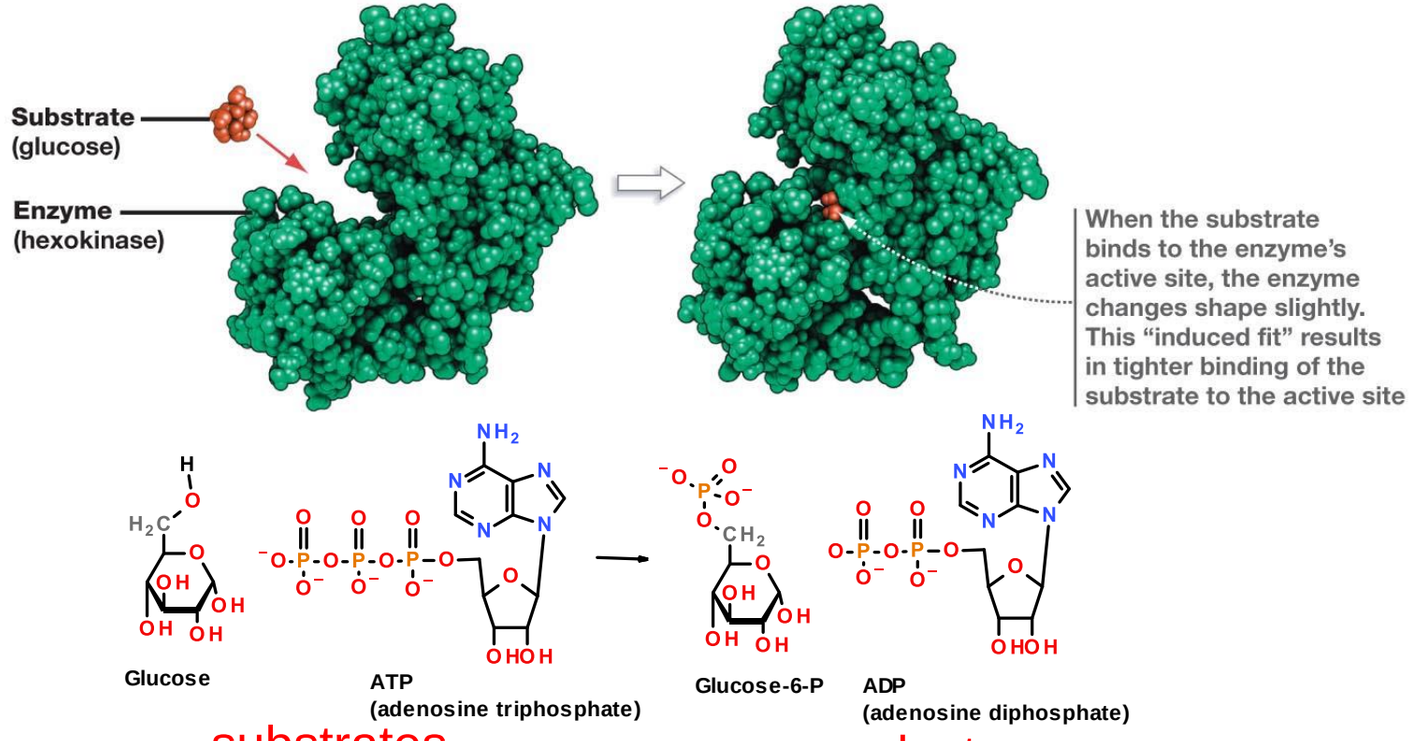




Enzymes

- **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

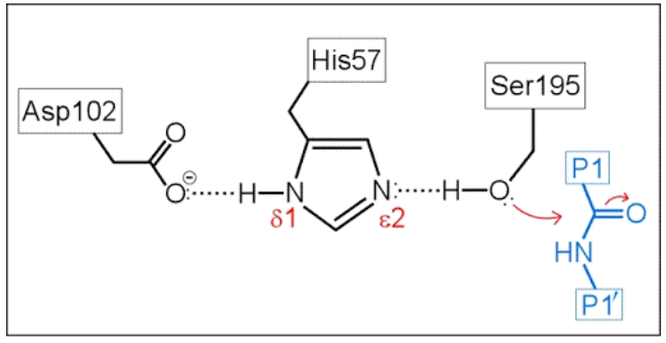
TABLE 4–1 SOME COMMON FUNCTIONAL CLASSES OF ENZYMES	
ENZYME CLASS	**BIOCHEMICAL FUNCTION**
Hydrolase	General term for enzymes that catalyze a hydrolytic cleavage reaction.
Nuclease	Breaks down nucleic acids by hydrolyzing bonds between nucleotides.
Protease	Breaks down proteins by hydrolyzing peptide 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.

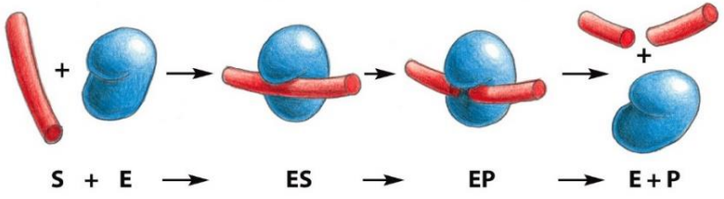
- Most enzyme names end in “-ase”
- Usually named by their substrates and the reactions they catalyse, i.e. glucose kinase

Example of Active Site Functional Groups:

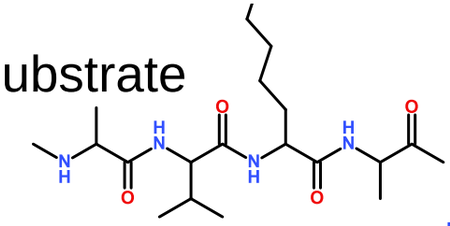
Catalytic triad (Asp, His, Ser) in Protease Trypsin cleaves after Lys Residues

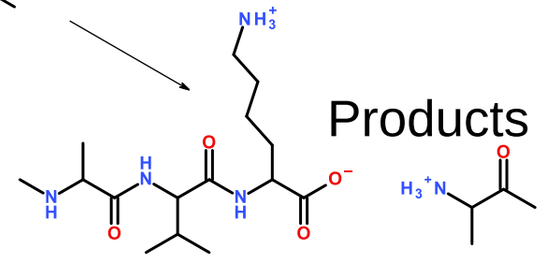
Catalytic triad



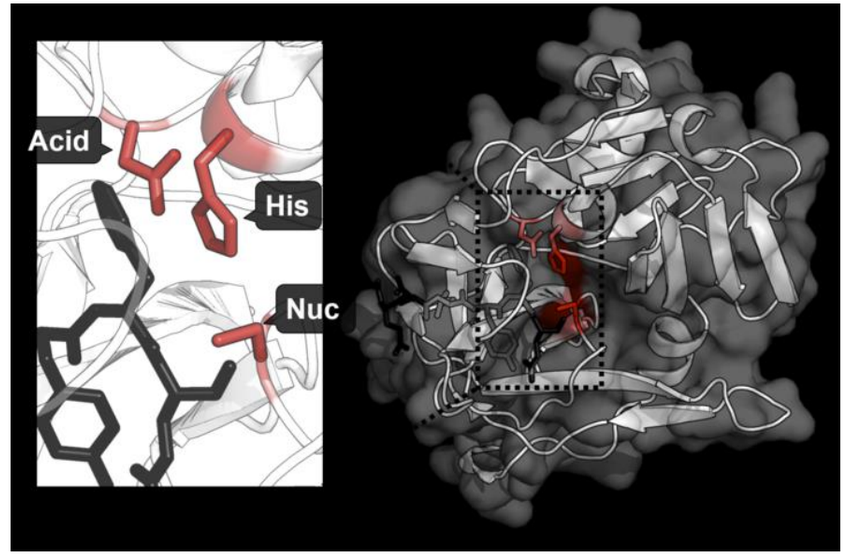


Substrate

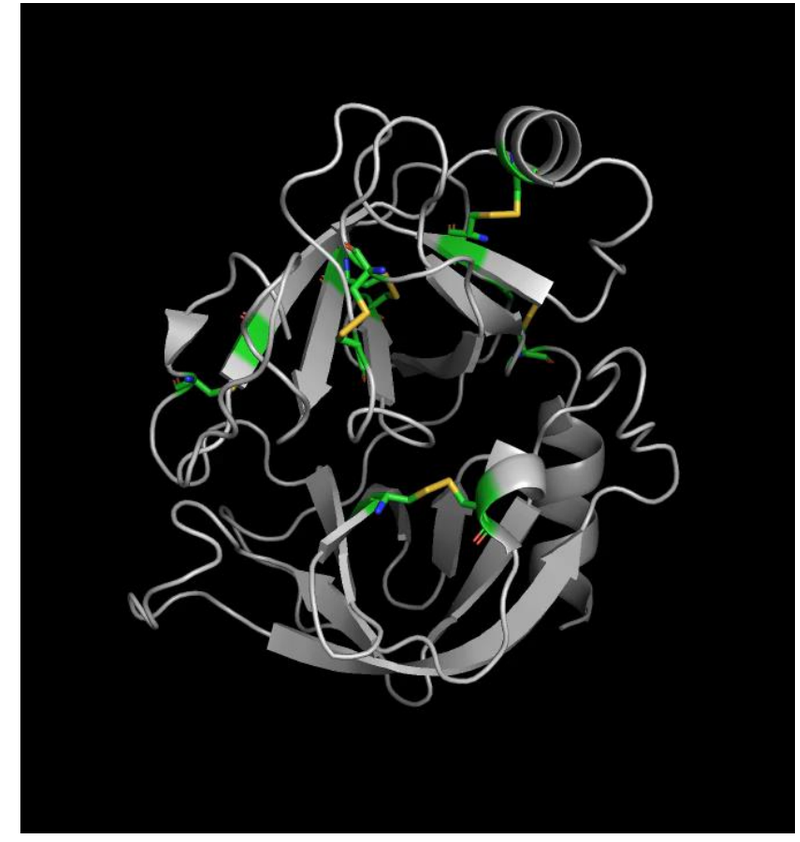




Products



https://shirleychemproject.weebly.com/



Disulfide bonds in trypsin

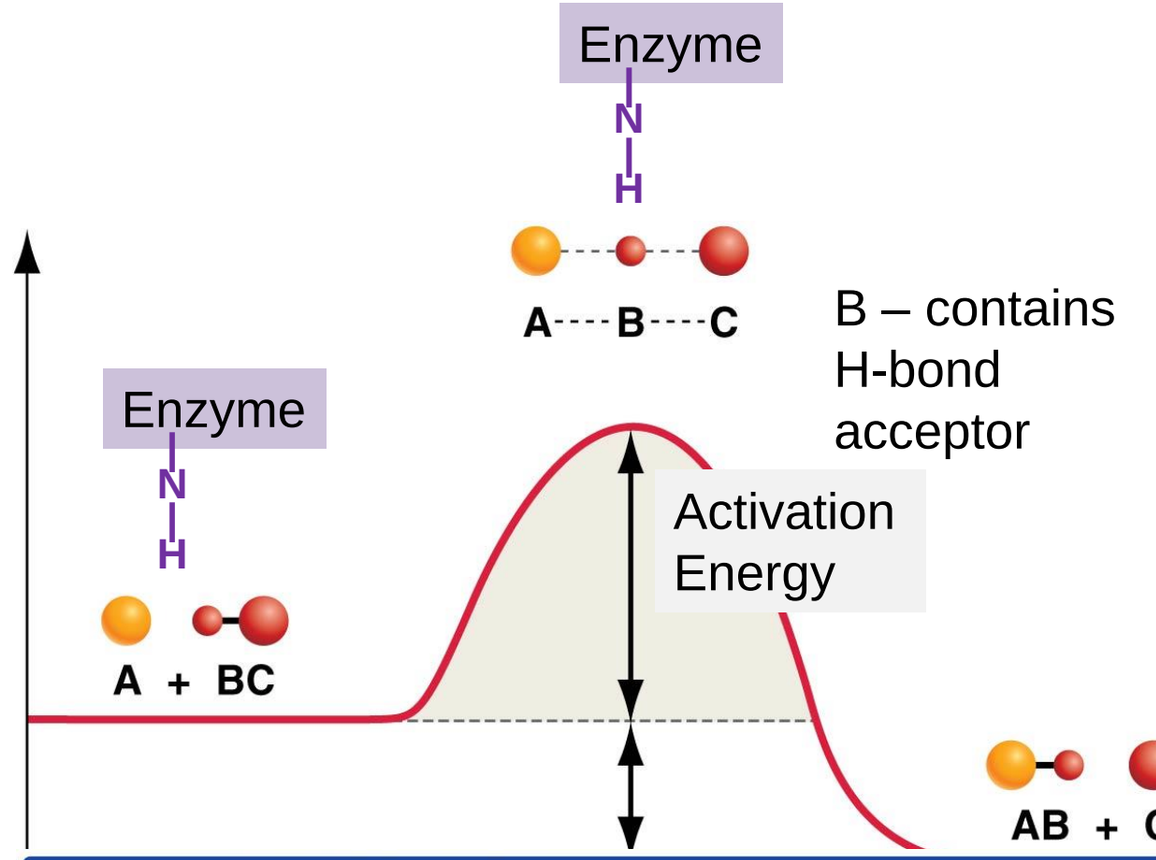
How Do Enzymes Increase Rates?

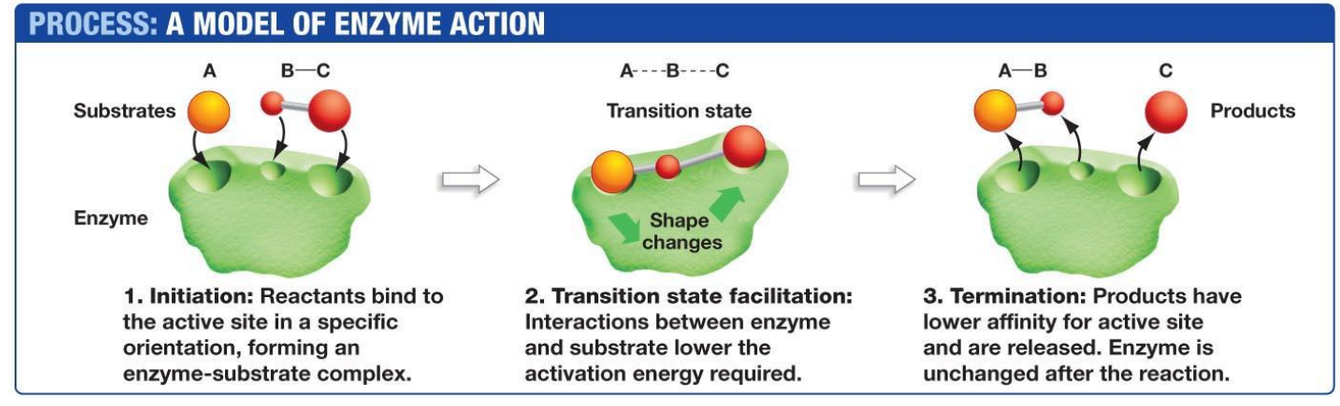
- **Transition state** = high energy intermediate that occurs during the reaction.
- Energy barrier is called the activation energy.
- Rate of product formation depends on the concentration of the transition state.

Low [X] = Slow reaction

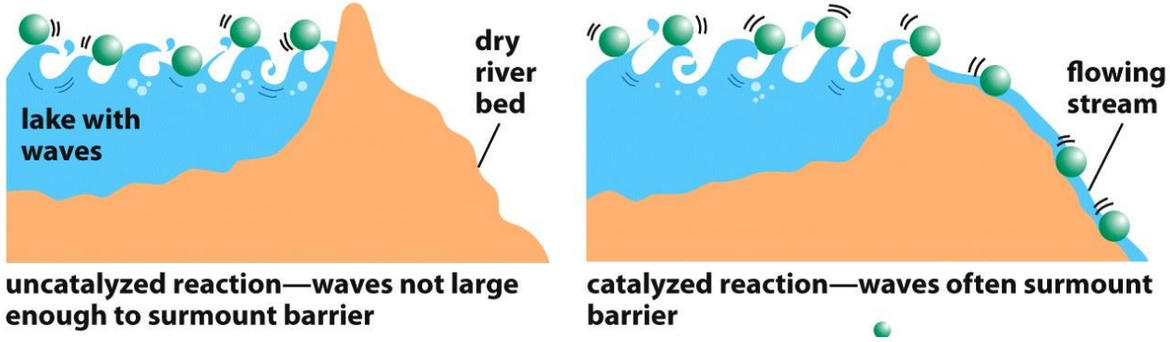
Higher [EX] = Faster reaction

- Interactions between the enzyme and the substrate stabilize the **transition state** (X) and lower the activation energy required for the reaction to proceed.
- Stabilization can include:
  - Pre- alignment of key groups in the active site, reducing entropy cost of organizing groups.
  - Direct interactions with the transition state (see diagram, N-H group interacts more favorably with the transition state)

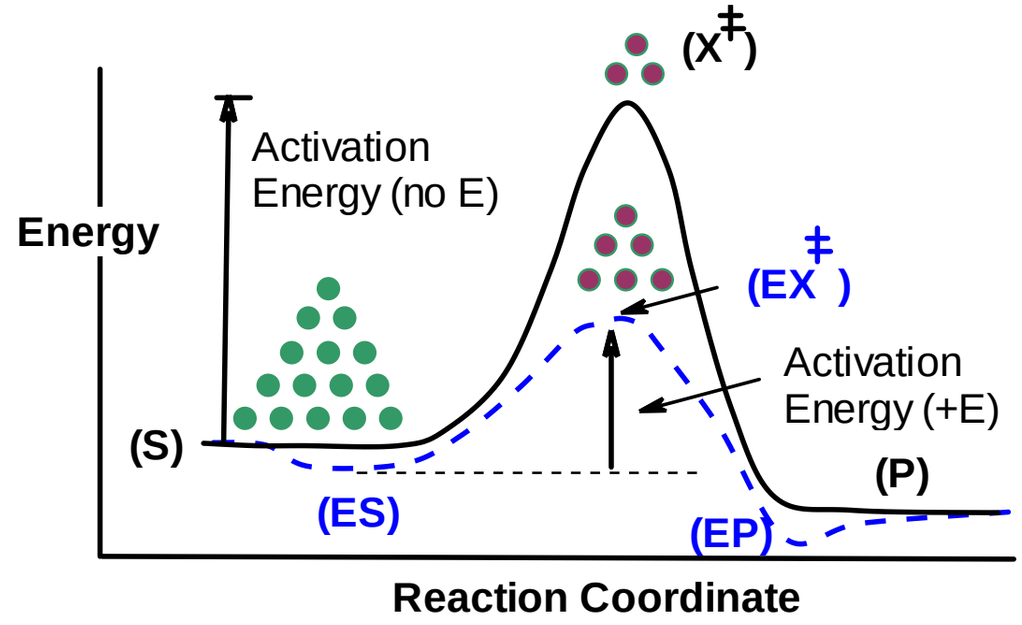




A model of transition state stabilization.



Lower energy of transition state allows more substrates to reach transition state due to their thermal energy.



[S] = 15

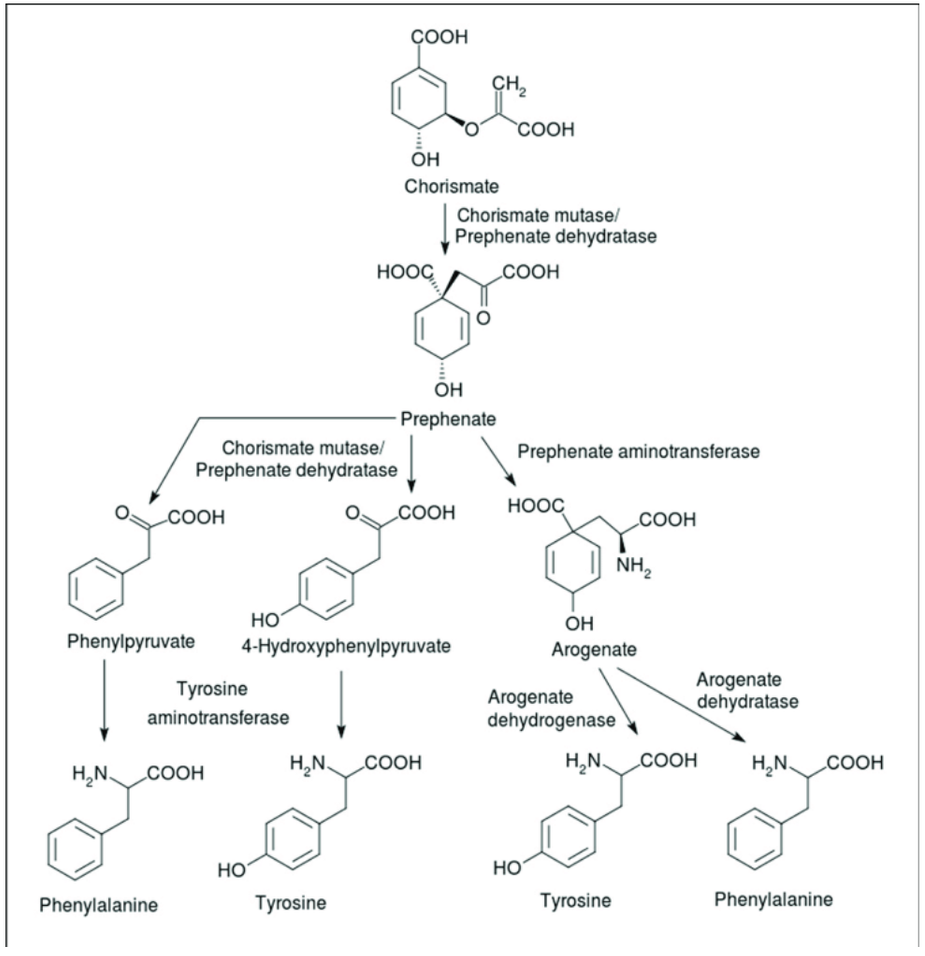
[X] = 3
[EX] = 6

How much faster will the rate be when the enzyme is present?

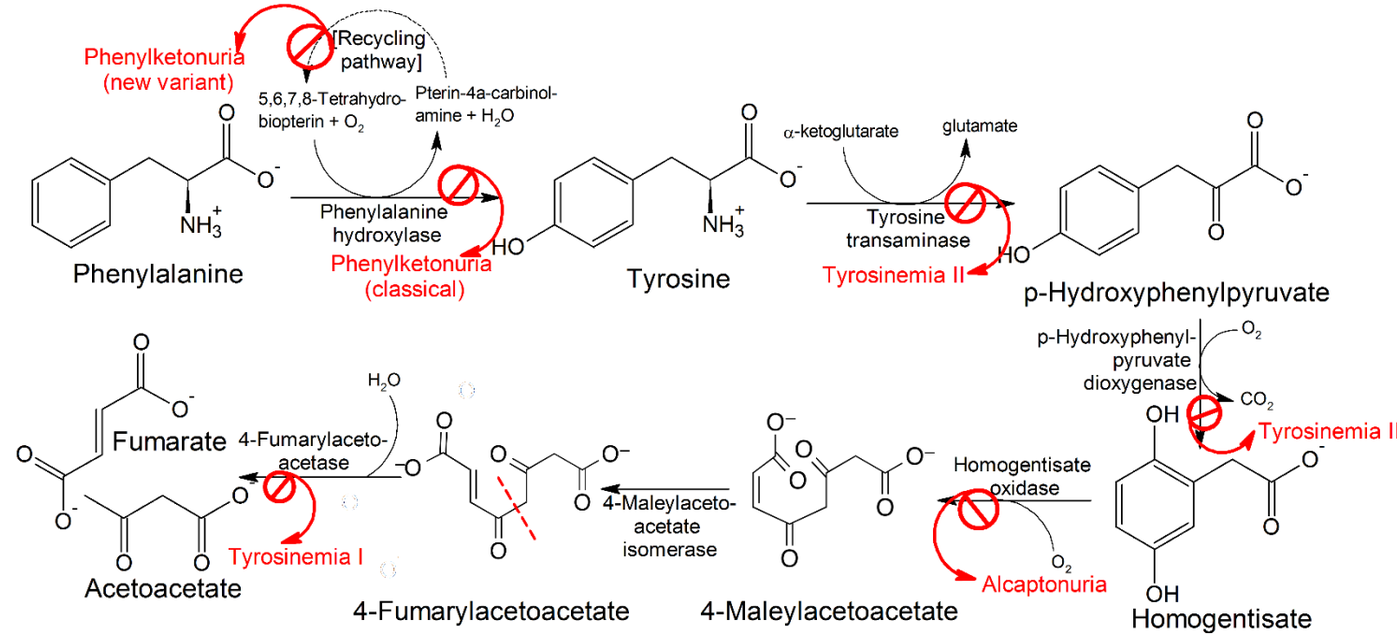
Enzymes, Metabolic Pathways, and Diseases

Synthetic Pathway for Phe, Tyr (beginning with chorismite)

- Each step catalyzed by an enzyme



Pathway for Degradation of Phenylalanine

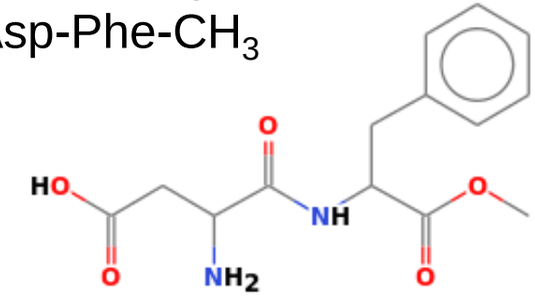


PKU Disease:

- Inactive phenylalanine hydroxylase
- Phe levels become toxic:
- Neurological problems
  - Intellectual disability
  - Developmental delays
  - Mental health disorders.

Aspartame (artificial sweetener)

Asp-Phe-CH_3



Key Points:

Enzymes:

- Enzymes bind substrates (S), forming (ES) complex in active site, converting to P, releasing P.
- Rate enhancement since 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.
- Genetic diseases that lead to inactive metabolic enzymes can cause disease.