# *Concepts students should know for Medicinal Chemistry CHEM 515*

*you may not have all these memorized, but we will expect some familiarity to allow more efficient discussion. If you are not familiar with these concepts, please review them or talk to Dr. Mal.

## 1. Organic functional groups

Miscellaneous

a) alkyl halides: R-X, where X=F, Cl, Br, or I
**Note R is used as a generic abbreviation for alkyl groups
-fluoro- (F), chloro- (Cl), bromo- (Br) and iodo- (I)

b) ethers: R-O-R'
$CH_3CH_2OCH_2CH_3$

c) alcohols: R-OH
$CH_3CH(CH_3)OH$= isopropyl alcohol

d) amines: R-$NH_2$, $R_2NH$, $R_3N$
$(CH_3CH_2)_3N$ = triethylamine

C=O functional groups

e) aldehyde = RCHO

f) ketone = RC(=O)R'

g) carboxylic acid = $RCO_2H$ or RCOOH

h) carboxylic ester = $RCO_2R'$ or RCOOR'

i) carboxylic amide = $RC(=O)NH_2$, RC(=O)NHR' or $RC(=O)NR'_2$

Unsaturated hydrocarbon functional groups

-contain pi bonds

j) alkene = $R_2C=CR_2$

k) alkyne = R-C≡C-R

l) benzene/aromatic = $C_6H_6$ et al.

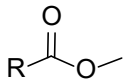
-a.k.a. phenyl (Ph) ring
**Note Ar is used as a generic abbreviation for aromatic or aryl groups

m) nitrile = R-C≡N

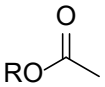
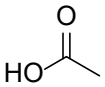
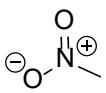
n) nitro, = R-$NO_2$

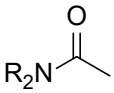
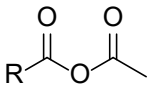
## 2. Electron releasing groups (ERG's) vs. electron withdrawing groups (EWG's)

ERG's:

R- (alkyl) RO- (ether)  (ester) $R_2N$-, RHN-, $H_2N$- (amines)

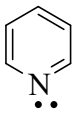
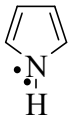
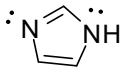
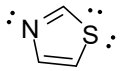
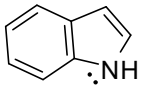
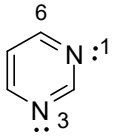
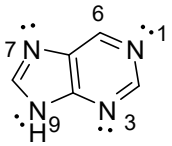
EWG's:

(ketone)  (ester)  (acid)  (nitro)

(aldehyde)  (amide)  (acid anhydride) (cyano)

## 3. Heteroaromatic rings

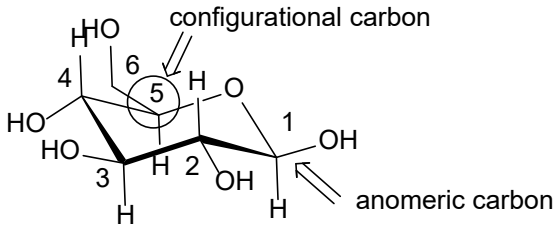
**I will review this at the start of the course if students don't have experience with heteroaromatic structures.*

| | | | |
|---|---|---|---|
| *pyridine*<br>  <br>-found in pyridoxal, a.k.a. vitamin B6 | *pyrrole*<br>  | *imidazole*<br>  <br>-note: 2 N's<br>-found in important amino acid histidine | *thiazole*<br>  <br>-found in thiamine, a.k.a. vitamin B1 |
| *indole*<br>  <br>-note: two rings with benzene fused with pyrrole | *pyrimidine*<br>  <br>-note: pyridine with a second N<br>-found in DNA and RNA | *purine*<br>  <br>=pyrimidine + imidazole<br>-found in DNA and RNA | *furan*<br> |

## 4. Carbohydrate structures

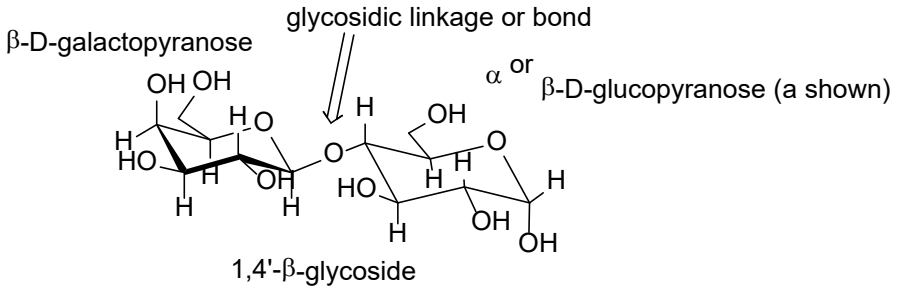
a) Monosaccharides:

e.g. D-glucopyranose (D-glucose in a cyclized form)




b) Disaccharides

e.g. lactose=β-D-Gal-(1→4)-α/β-D-Glc




-names:

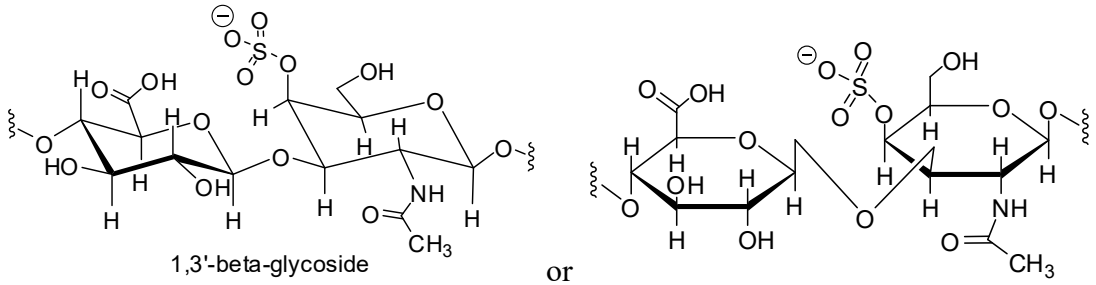
official name: 4-O-(β-D-galactopyranosyl)-α-D-glucopyranose

or O–β-D-galactopyranosyl-(1--> 4)-α-D-glucopyranose

or β-D-Galp-(1-->4)-α-D-Glcp (w/ 3 letter sugar abbreviations)

c) Polysaccharides:

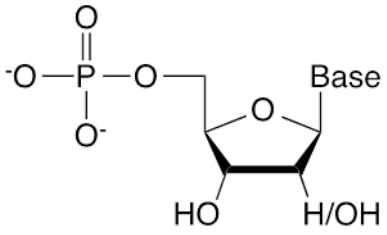
e.g. chondroitin-4-sulfate= [D-Glucuronate-beta-(1→3)-beta-N-Ac-D-Galactosamine-4-sulfate beta-(1→4)]$_n$- polymer; or a.k.a. [β-D-GlcA-(1→3)-β-D-GalNAc4S-(1→4)]$_n$ (GalN=galactosamine); n=~20-60




## 5. Nucleic acids

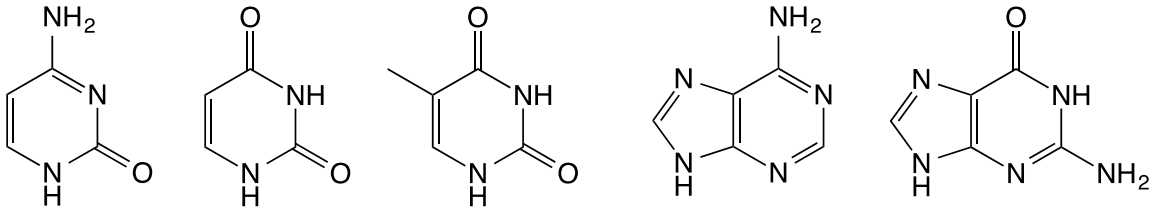
a) Nucleotide and oligonucleotide structure: nucleoside + phosphate(s)

- Ribose sugar
  - purine/pyrimidine base
  - phosphate groups
  - directionality, antiparallel




b) DNA vs. RNA

- bases: cytosine, uracil/thymine, adenine, guanine



**C** **U** **T** **A** **G**

- 2′ -OH group

- double/single stranded

- differences in stability

c) Intermolecular forces

- hydrogen bonding

- aromatic base-stacking

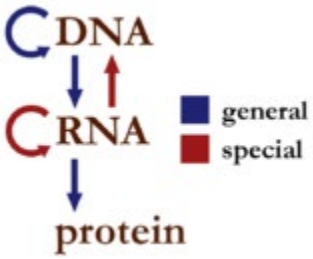
d) Forms of DNA, plus overall structural similarities/differences

- A-form: right-handed, major/minor grooves

- B-form: right-handed, major/minor grooves, most common

- Z-form: left-handed, zig-zag backbone

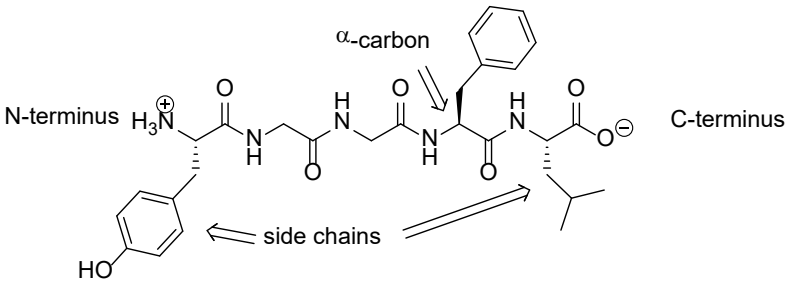
e) Central dogma




## 6. Protein structures

a) Peptides

e.g. Leu-enkephalin



-Know all 20 natural amino acid side chains

-peptide nomenclature for Leu-enkephalin:

-full name: tyrosine-glycine-glycine-phenylalanine-leucine

-3 letter abbreviation: Tyr-Gly-Gly-Phe-Leu

-1 letter abbreviation: Y-G-G-F-L

b) Proteins

4 levels: 1°, 2°, 3° and 4°.

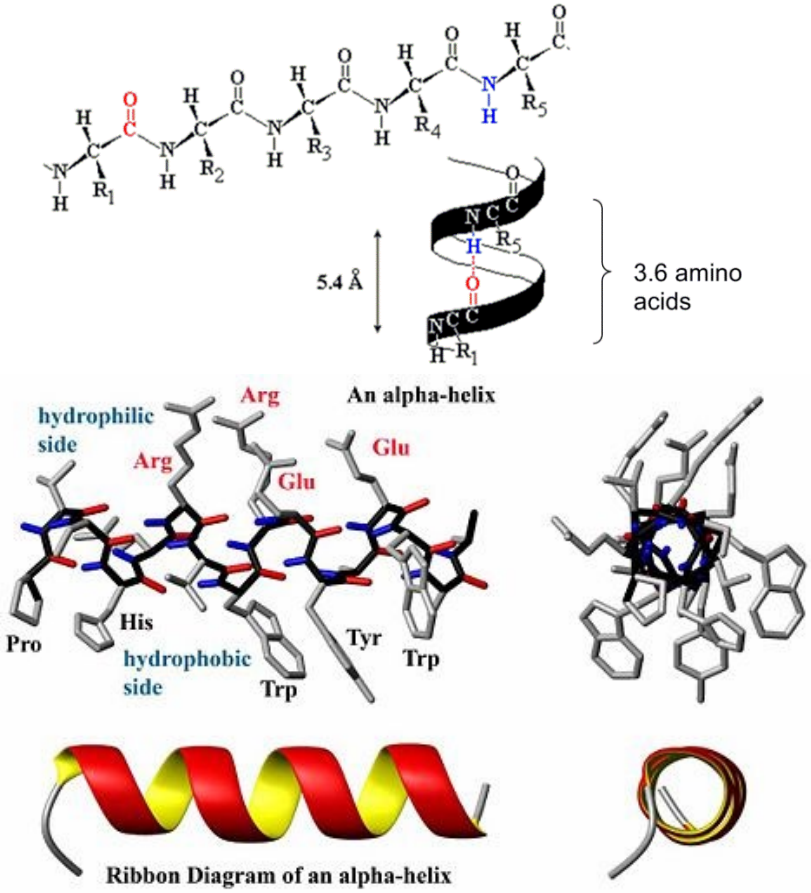
(1) Primary, 1°

=linear amino acid sequence; unique for each protein

(2) Secondary, 2°=local spatial arrangement of amino acids

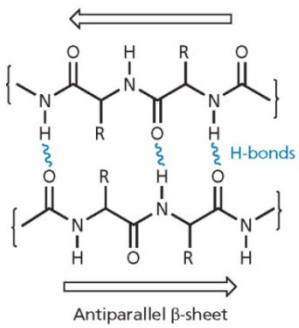
-4 basic common forms for proteins:

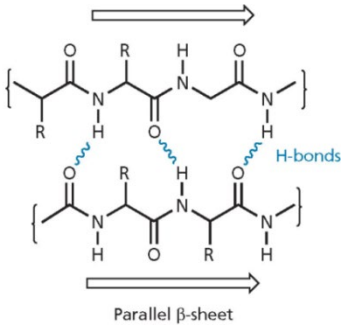
a) alpha-helix

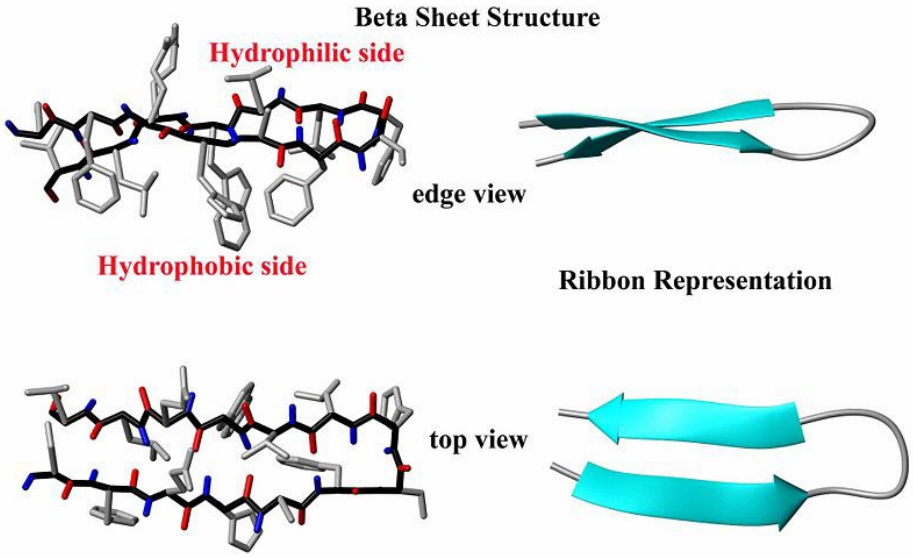


b) beta-sheets

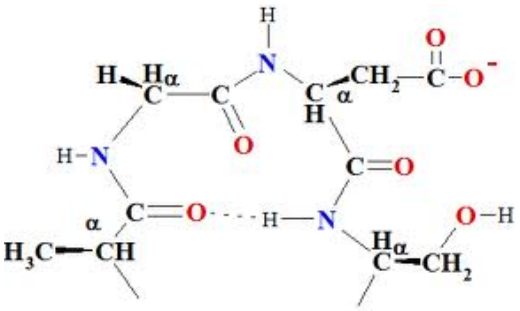
two flavors: parallel and anti-parallel

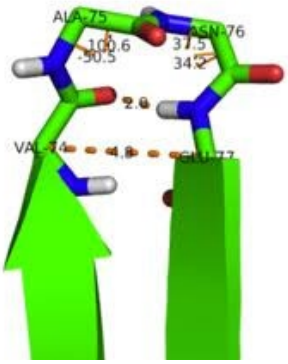


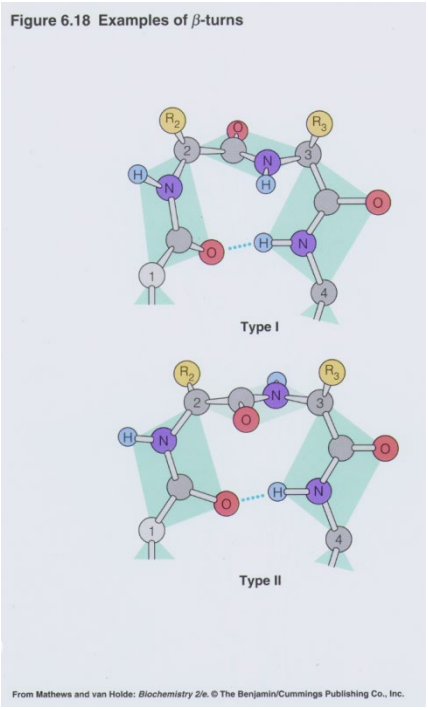




c) reverse turns or beta-bends



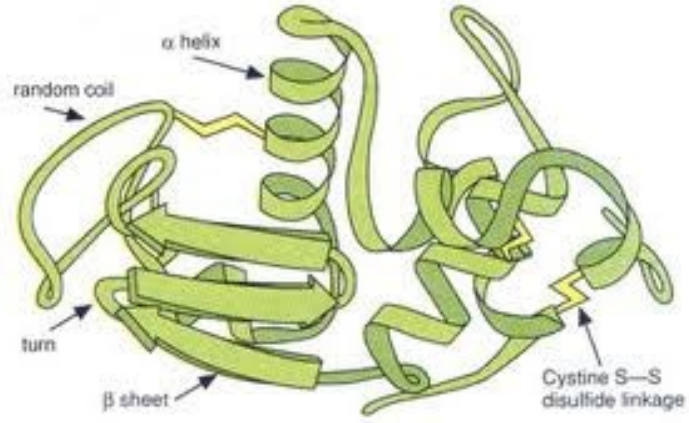


*numbered i, i+1, i+2, i+3 in the N→C direction



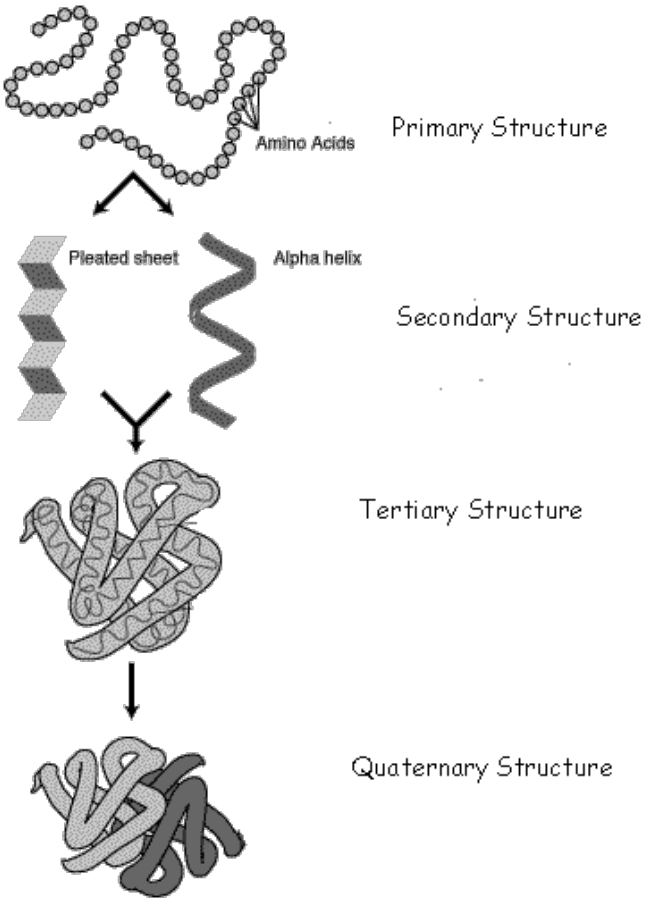

d) random coil (not an organized secondary structure)

***Picture with all four possible secondary structures displayed:***



(3) Tertiary, 3° (structure)
-combination of 2° elements into a more complex protein 3-D structure

(4) Quaternary, 4°
-combination of multiple proteins together
-multimeric proteins
multimer= multi-subunit protein
-each unit of a multimer is a protomer
oligomer= a few subunits making protein



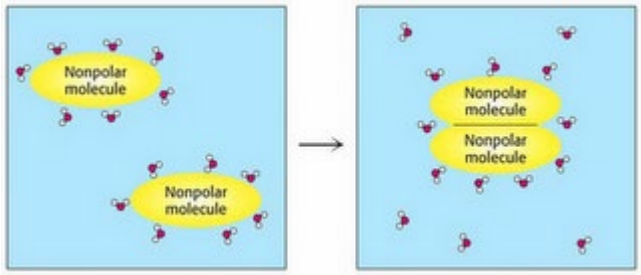

<u>One additional extremely important concept: *HYDROPHOBIC EFFECT*</u>

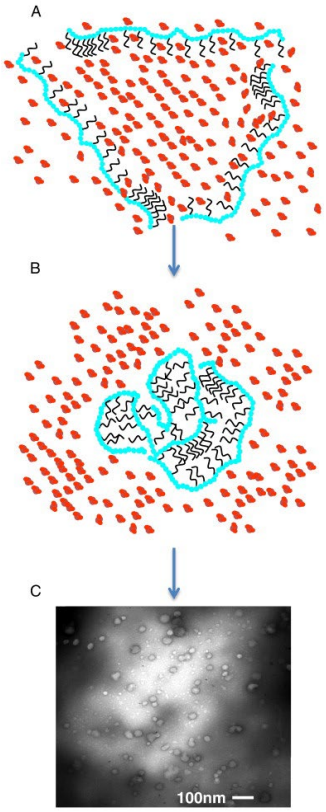
-combining greasy, non-polar regions together→ reduces the surface area of greasy, hydrophobic structures→gives water molecules more freedom to move around→ increased entropy and thermodynamically favorable!

-normally, water is wallowing around in a sea of H-bonds like a pig in mud

-hydrophobic hydrocarbons intercept H bonds and force water into a more ordered arrangement around the hydrocarbon surface: Fewer H-bond options available → more $H_2O$ organization → lower entropy→ less thermodynamically favorable

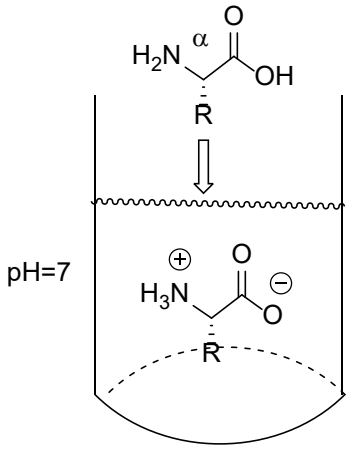
<u>Pictorially:</u>




## 7. pKa/pH

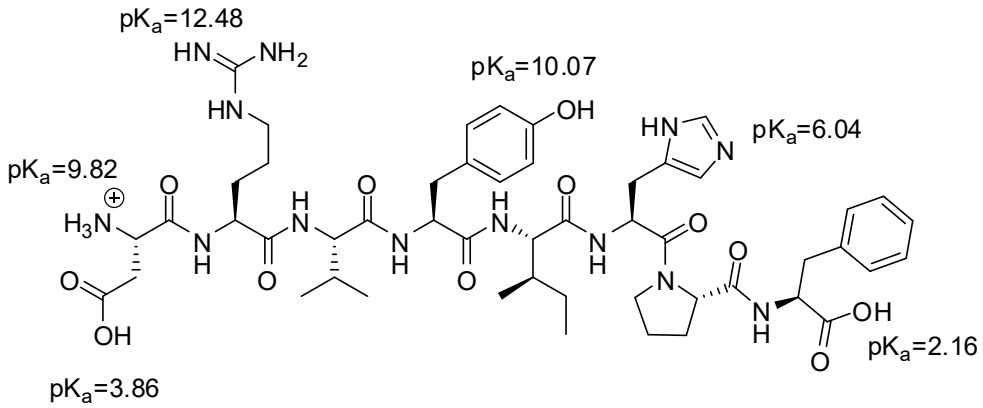
*protonation state of amine and acid change based on pH of solution…remember the comments on amine pH/$pK_a$ interactions:




*<u>note:</u> we are talking about R-$CO_2H$ and R-$NH_3^+$ acid dissociation equilibrium!!

*<u>zwitterion</u>=molecule with both a + and – charge.

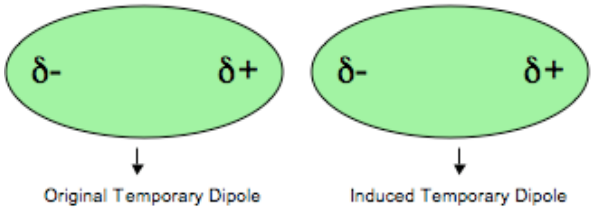
-more complex example: (angiotensin II)




## 8. Intermolecular forces

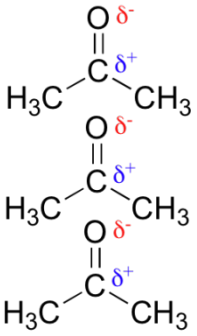
*Listed in order of increasing strength:

a) induced dipole-induced dipole
-due to transient areas of high and low electron densities cause temporary dipoles



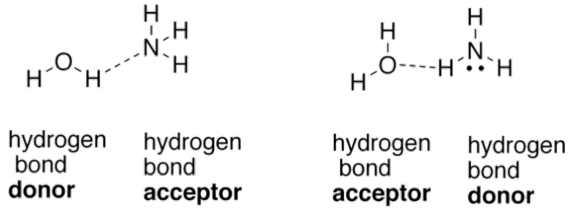

-a.k.a. van der Waals forces or London forces
-weakest of all intermolecular forces: 2-4 kJ/mol
-longer, thinner alkanes (less branching) will have stronger induced dipole-induced dipole interactions (polarizability, the ability to distort an electron cloud, is a characteristic that contributes to stronger intermolecular forces)
-only intermolecular force with greasy alkanes; will be present in interactions of all alkyl chains as well.

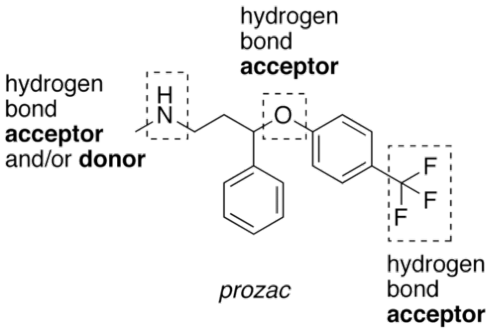
b) dipole-dipole interactions
-slightly stronger at 3-4 kJ/mol
-need bonds with dipole moments, e.g. C-Cl, C-O, C-N (not C-H)
-dipoles align themselves to hold molecules together
-e.g. just like pic of induced dipole-induced dipole, but greater in magnitude

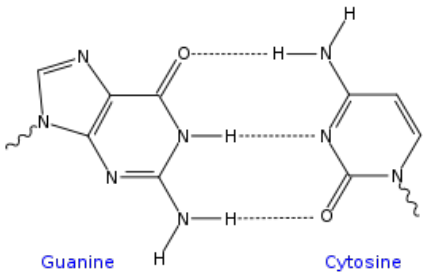


c) hydrogen bonding (H-bonding)

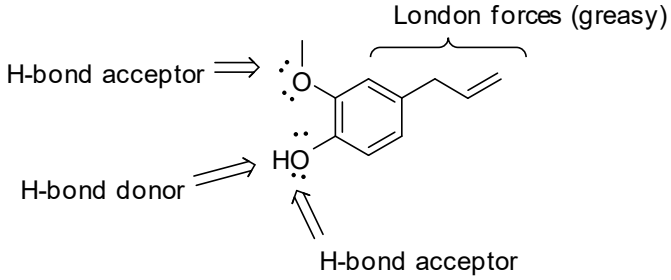
-stronger still at ~10-40 kJ/mol
-a special kind of dipole-dipole interaction between an electron deficient hydrogen and an electron rich heteroatom (N or O; F also)
-H-bond acceptors: O's and N's with lone pairs; also F
-H-bond donors: primarily N-H, O-H
(H-bond donor) -N-H----:N- (H-bond acceptor)
also, (H-bond donor) –O-H----:O-R(H) (H-bond acceptor)




Important application:
-Analyzing molecules for intermolecular forces:



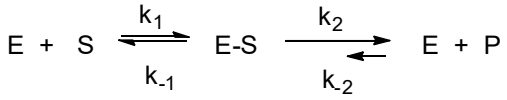

*General characterization of structures based on intermolecular forces and "like dissolves like" application:*

-hydrophobic/lipophilic = lots of C's and H's → doesn't dissolve in water ("oil and water")

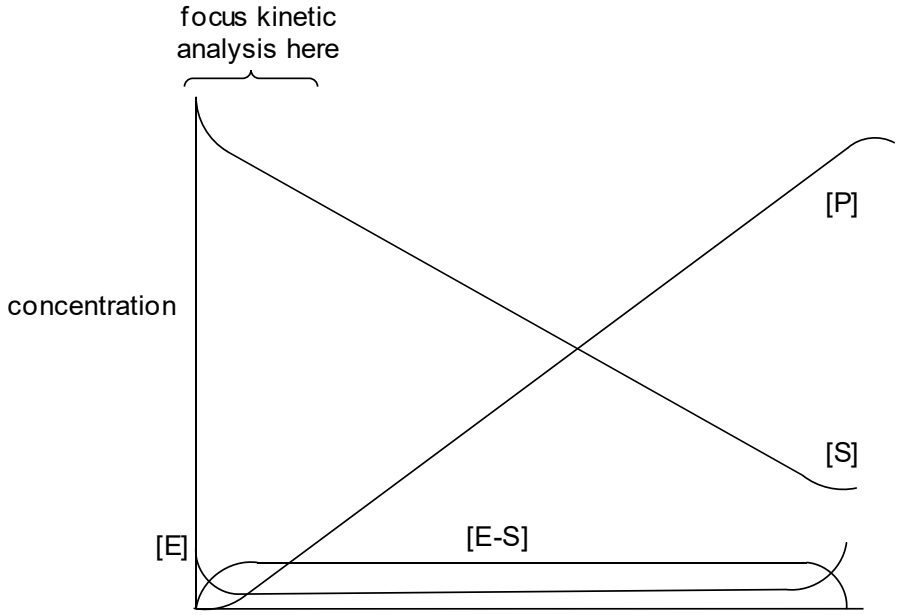
-hydrophilic (lipophobic) = more O's and N's → does dissolve in water (H-bonds)

## 9. Enzyme catalysis

Rxn:




Graphical look at the process:




a). Key assumptions in analyzing enzyme kinetics:

(1) study initial rate, $v_o$
- -before 10% of S consumed (usually first minute of rxn)
- -eliminates P→S ($k_{-2}$), P inhibition and E inactivation

(2) rate limiting step is breakdown of E-S (Michaelis and Menten, 1913), $v_o=k_2[E\text{-}S]$ (focus on E-S!)
- -equilibirum established : E + S → E-S and amount of E-S determines rxn rate
- -"E is saturated with S"
- -otherwise stated: $k_{-1} >> k_2$

(3) steady-state assumption (Briggs and Haldane, 1925)=rxn quickly achieves a steady state defined by stable/constant [E-S]

b). Important constants and formulas

**(1) Michaelis-Menten constant, $K_m=(k_2 + k_{-1})/k_1 = k_{off}/k_{on}$**

=rate constants of rxns. breaking down E-S over rate constants forming E-S

***$K_m$ becomes one important constant for measuring enzyme activity

*varies with temp and pH

-units of M

*_sometimes_ $k_{-1}>>k_2$, then $K_m=(k_2 + k_{-1})/k_1$ becomes $K_m=k_{-1}/k_1$

**So....$K_m$ is an apparent dissociation constant (~processes decomposing E-S/processes forming E-S) and therefore a measure of how well enzyme binds substrate!

*the smaller the number, the stronger the E binds to the S

THIS IS A GROSS SIMPLIFICATION THAT DOES NOT ALWAYS APPLY!!

### (2) $k_{cat}$ or turnover number:

…from the basic equation for enzyme catalytic activity…

$v_o=k_2[E\text{-}S]$ and $V_{max}=k_2[E_{total}]$

$k_2=k_{cat}$=turnover number, i.e. how fast enzyme can make product after substrate is bound
*the higher the number, the faster the enzyme makes product
*this works even with multi-step enzyme rxns b/c $k_{cat}$ is the rate of rxn *after the rate limiting step*. It can become a conglomeration of several rate constants.

-units of $s^{-1}$

### (3) Michaelis-Menten equation:

$$v_o = \frac{k_2[E_{total}][S]}{K_m + [S]} = \frac{V_{max}[S]}{K_m + [S]}$$

**These are all values that can be experimentally determined.

**This is a characterization of the quantitative relationship between initial velocity, $v_o$, the maximum velocity, $V_{max}$, and the [S].

Special case:
-when $v_o = 1/2(V_{max})$ and solve…

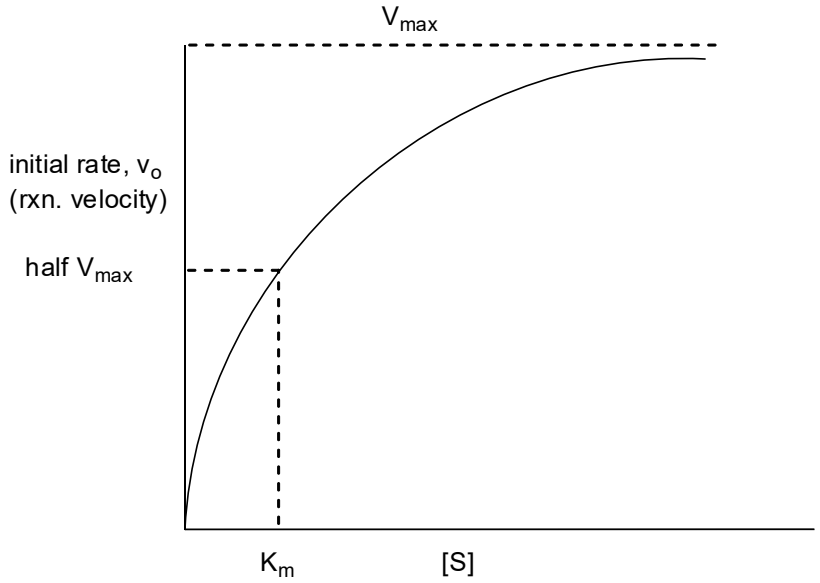
$$v_o = \frac{V_{max}[S]}{K_m + [S]} = \frac{1}{2} V_{max}$$

cancel $V_{max}$

$$\frac{[S]}{K_m + [S]} = \frac{1}{2}$$

$$K_m=[S]$$

then $K_m$=[S]! So,…
-** $K_m$ measures the [S] at which half the enzyme active sites are filled.




Special note:
-$K_m$ ~ natural [S].

## (4) Specificity constant: $k_{cat}/K_m$

$v_o = k_2[E_{total}][S]/(K_m + [S])$
IF, $[S] << K_m$, then there is very little E-S formed (due to small amount of S) and…

$v_o = k_2[E_{total}][S]/(K_m + [S])$ becomes… $v_o = k_2/K_m ([E_{total}][S])$

$k_2/K_m = k_{cat}/K_m$ is a 2nd order rate constant = measure of catalytic efficiency

**Interpretation: $k_{cat}/K_m$ is a measure of how well enzyme works when there is very little [S]. When E is starved for S, how fast can it process S once it gets it.

**Limited only by the rate that a S molecule encounters E means the best E's will have a $k_{cat}/K_m$ that approaches the diffusion limit of molecules in solution, i.e. how fast can a S molecule travel through soln to get to an E molecule.

The diffusion limit is $10^8$-$10^9$ $M^{-1}$ $s^{-1}$ (note: second order rate constant units!)

Examples of parameters for a range of enzymes:

| **Table: Values of $K_m$, $k_{cat}$, and $k_{cat}/K_m$ for Some Enzymes and Substrates** | | | | |
|---|---|---|---|---|
| **Enzyme** | **Substrate** | **$K_m$ (M)** | **$k_{cat}$ ($s^{-1}$)** | **$k_{cat}/K_m$ ($M^{-1}$ $s^{-1}$)** |
| Acetylcholinesterase | Acetylcholine | $9.5 \times 10^{-5}$ | $1.4 \times 10^4$ | $1.5 \times 10^8$ |
| Carbonic anhydrase | $CO_2$ | $1.2 \times 10^{-2}$ | $1.0 \times 10^6$ | $8.3 \times 10^7$ |
| | $HCO_3$- | $2.6 \times 10^{-2}$ | $4.0 \times 10^5$ | $1.5 \times 10^7$ |
| Catalase | $H_2O_2$ | $2.5 \times 10^{-2}$ | $1.0 \times 10^7$ | $4.0 \times 10^8$ |

| | | | | |
|---|---|---|---|---|
| Chymotrypsin | N-Acetyl-glycine ethyl ester | $4.4 \times 10^{-1}$ | $5.1 \times 10^{-2}$ | $1.2 \times 10^{-1}$ |
| | N-Acetyl-valine ethyl ester | $8.8 \times 10^{-2}$ | $1.7 \times 10^{-1}$ | 1.9 |
| | N-Acetyl-tyrosine ethyl ester | $6.6 \times 10^{-2}$ | $1.9 \times 10^{2}$ | $2.9 \times 10^{5}$ |
| Fumarase | Fumarate | $5.0 \times 10^{-6}$ | $8.0 \times 10^{2}$ | $1.6 \times 10^{8}$ |
| | Malate | $2.5 \times 10^{-5}$ | $9.0 \times 10^{2}$ | $3.6 \times 10^{7}$ |
| Superoxide dismutase | Superoxide ($O_2^{-\cdot}$) | $3.6 \times 10^{-4}$ | $1.0 \times 10^{6}$ | $2.8 \times 10^{9}$ |
| Urease | Urea | $2.5 \times 10^{-2}$ | $1.0 \times 10^{4}$ | $4.0 \times 10^{5}$ |

*note: acetylcholinesterase, catalase and superoxide dismutase are nearly perfect enzymes since they can process S as soon as they encounter a S molecule! Carbonic anhydrase is also really good.

*note: efficiency and variability of constants for different enzymes