Lecture 1

Origin of Life: from Chemical to Living

Earth is believed to be roughly **4.6 billion years old.** The evidence for this number is as follows:

- Evidence:
  - radiochemical dating of rocks (particularly the decay of ^{238}U to ^{206}P, which can be used to estimate age.
  - astrophysical sciences (planetary motion, etc.)
  - geological uniformitarianism
  -

- Microfossils
  - fossils of organic organisms dating further back than 3.5 billion years
  - reminiscent of cyanobacteria (this is indirect evidence at best, sheds no light on how life formed)
  -

Primordial Earth

Primordial earth is believed to have had a **reducing atmosphere**, void of any O_2. It did however, contain stable inorganic molecules such as H_2O, N_2, CO_2, CH_4, NH_3, and SO_2.

There was *no complex organic molecules existing anywhere on the planet* (amino acids, carbohydrates, nucleic acids, protein lipids etc).

In the 1920s, **Oparin & Haldane** made the proposal that energy present in the atmosphere (lighting, solar UV) may have converting the stable inorganic molecules into more complex, organic molecules. In 1953, Urey and Miller carried out the famous experiment, in which the aforementioned inorganic molecules were placed in the presence of discharging electrodes (lighting). Using the Urey-Miller apparatus, complex organic molecules were collected after some time, including:

Acids	Nucleic acid bases	Amino acids	Sugars
formic acid	adenine	glycine	ribose
acetic acid	guanine	isoleucine	ribulose
succinic acid	uracil	proline	

Essential, Urey and Miller proved Oparin & Haladane's hypothesis that *In a reducing environment, complex chemical products can arise from the simplest precursors, given energy and time.* Its important to note that meteorites are often found with *similar organic molecules* significant because they likely form in a similar fashion (inorganic molecule + energy + time = organic molecules).

Chemical Evolution

It is a common hypothesis that these organic molecules, or **monomers**, were able to combine to form larger, more complex molecules called **polymers** when grouped together in large concentrations. These polymers are believed to eventually become *self-replicating*.

A prevalent hypothesis is the **RNA World hypothesis**. Essentially, it is the believe that the formation of RNA kick-started life on planet earth as we know it. The following is a summary:

- RNA is able to fold
  - this contributes to its *stability and structure*. Folded RNA is able to last longer!
  - *RNA can be catalytic*, examples include ribosomes and ribozymes. These are proposed to have catalyzed its own formation.

In agreement with this "self-assembly" theory was the idea of **compartmentalization**. It is believed that the compartmentalization of organisms (from a simple bacteria, protecting its genetic information with a cell membrane, to a eukaryote, with membrane bound organelles) was able to reduce variation due to environmental change enough to allow live to thrive.

Which leads us to a definition for life (or what it means to be alive):

- Most Common Scientific Definition: Life posses the properties of replication, catalysis, and mutability.

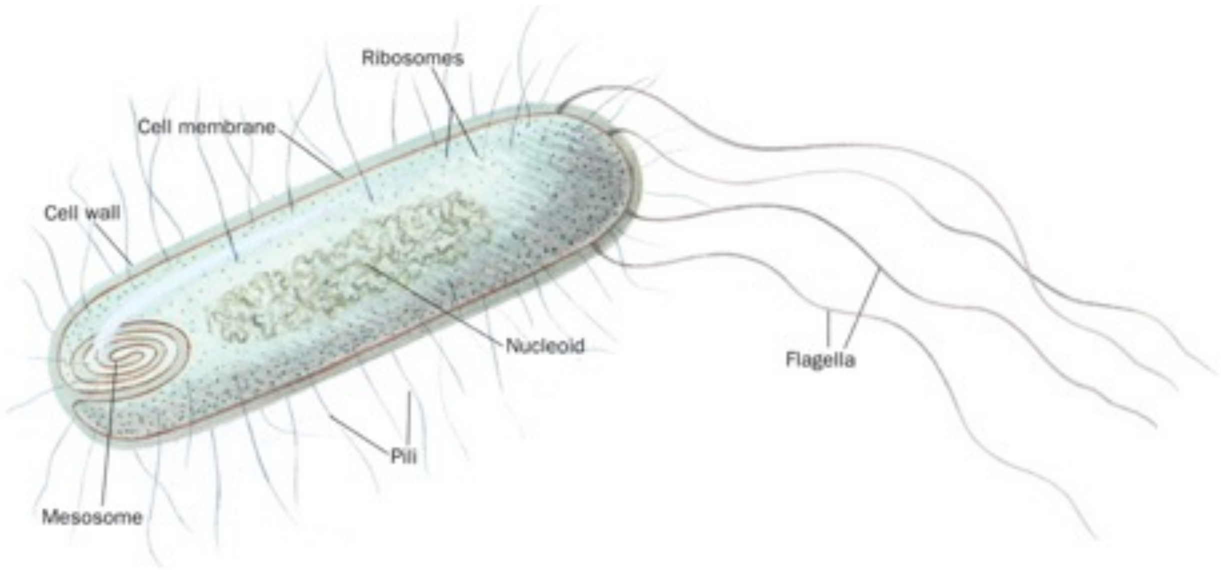
In essence, Life evolved from inanimate chemicals governed by fundamental chemical principles. Fundamental chemical principles *thus govern biochemistry.* We will be exploring these basic chemical principles in the next few chapters.

Smallest unit of life: the cell

The major cell types are **eukaryotes** and **prokaryotes.**

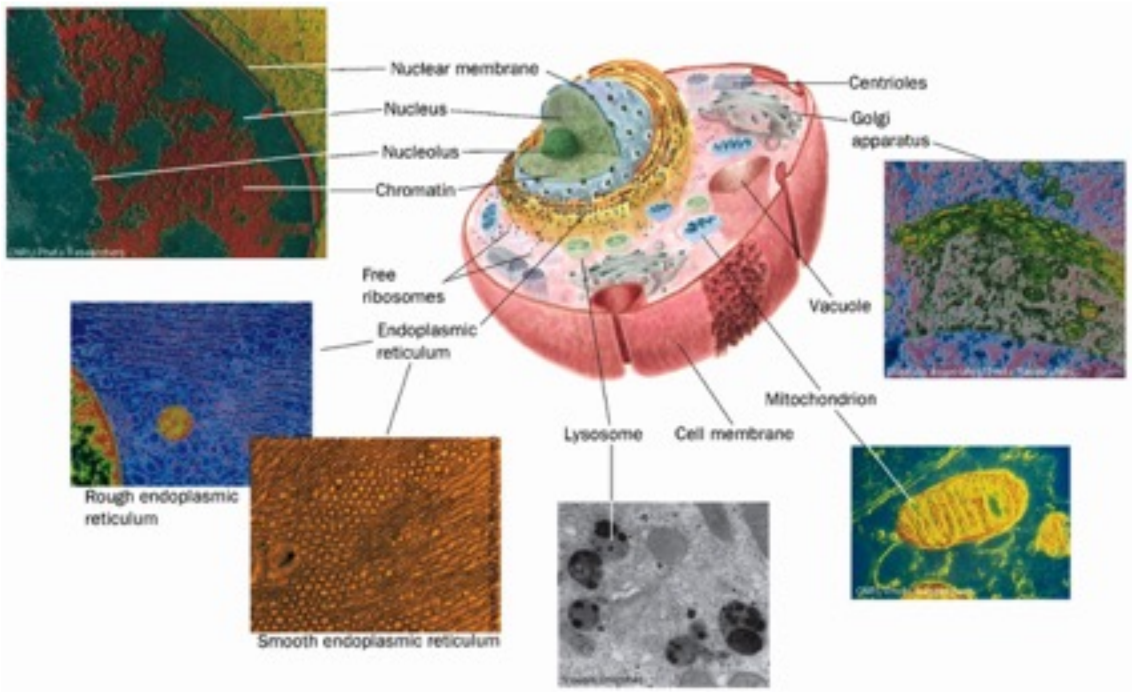
Prokaryotes (1-10um)	Eukaryotes (10-100um)
Bacteria	Fungi
Archaea	Protists
	Animals
	Plants

General Scheme of a prokaryotic cell



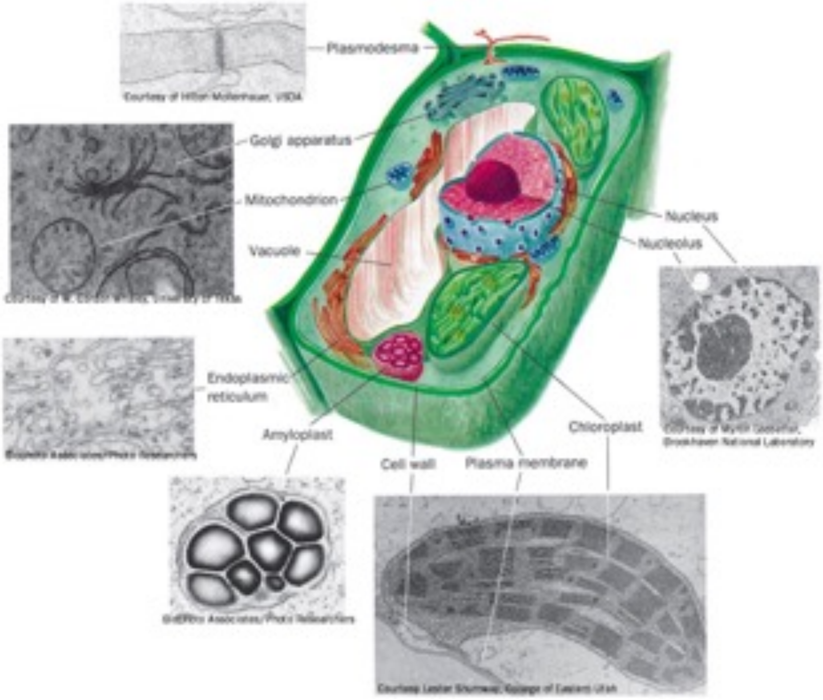
- **~6000-10,000** different biomolecules present.
- genetic material is grouped in the centre of the cell (nuceloid) but is NOT compartmentalized from the cytosol.
- The **cytosol** (cytoplasm analog) is densely packed with *proteins, nucelic acids, and metabolites.*

General Scheme of a animal cell



- **~100,000** different biomolecules present (far more complex than bacteria)
- **compartmentalization** is observed (all organelles are membrane bound)

General Scheme of a plant cell

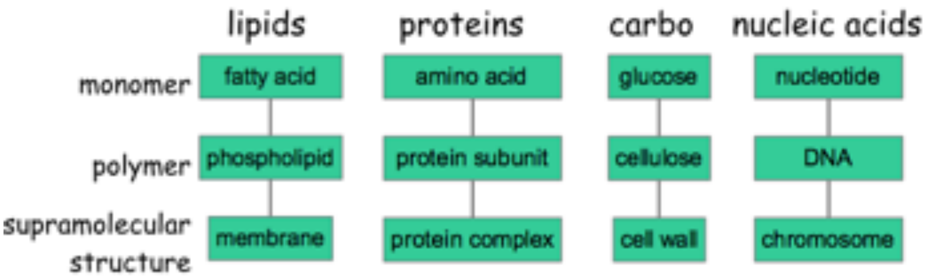


- **~100,000** different biomolecules present (far more complex than bacteria)
- **compartmentalization** is observed (all organelles are membrane bound)

Biomolecules

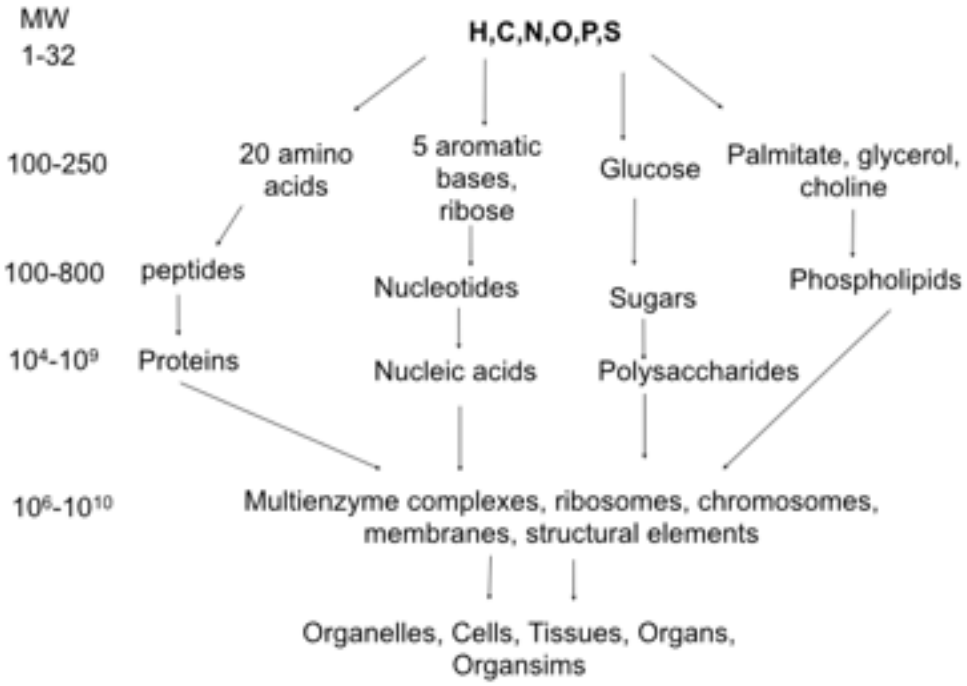
Biomolecules are **polymers**, and are comprised of a *limited number of building blocks.* Their complexity and diversity arises due to the order and type of monomers that make up their composition.

Biomolecule map:



*omitting any chemical explanation of the major biomolecules, as these will be discussed in much greater detail later on in the course.

Organizational chart, from the simplest organic elements to complex, living structures:



Lecture 2

Properties of Water

Much on this topic will be omitted, as it is review. Some key properties of water:

- Solvent Ability: Water is an amazing solvent, due to its net dipole or “polar nature”. Water is able to form a salvation shell (or *hydration* shell) around the ionics of dissolved ionic compounds, or polar regions of organic molecules.
- H-bond donor and acceptor: Water is able to form H-bonds both by accepting hydrogens via oxygens two L.Ps, and by donating its existing hydrogens H’s. Water is thus able to form up to 4 H-bonds (i.e. the crystal lattice of ice).
- High B.P, M.P, heat of vaporization and surface tension: Essentially, this H-bonding potential drastically increases the stability of water molecules, leading to high melting points, boiling points, and a incredible surface tension.

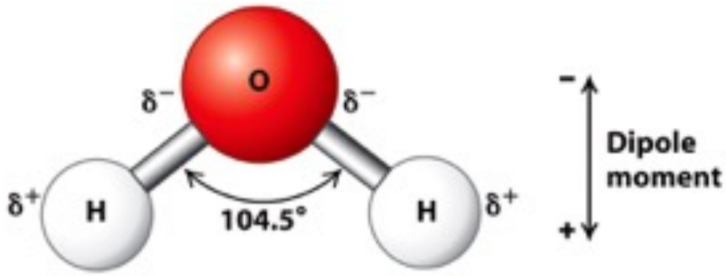


Figure 2.5
Molecular Cell Biology, Sixth Edition
© 2008 W.H. Freeman and Company

Biomolecule shapes and interactions are mediated by 4 types of non-covalent bonds.

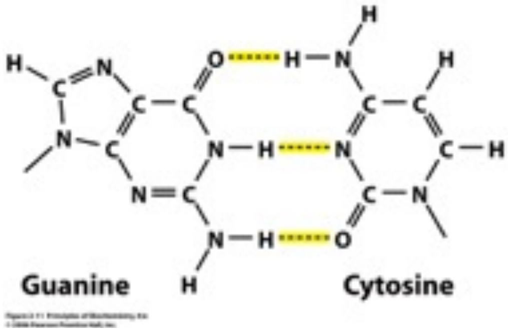
There are four types of bonds responsible for the overall shape and interaction among biomolecules and can be modified by thermal energy.

1) Electrostatic Interactions

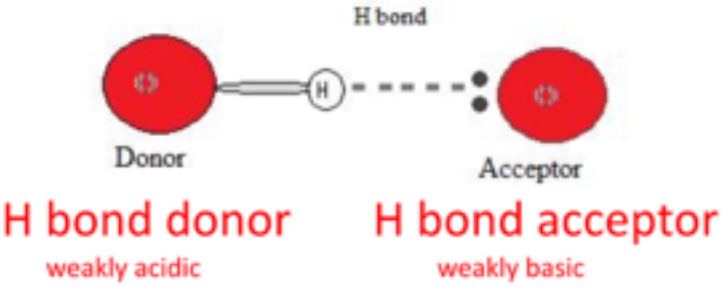
Electrostatic interactions are between and among cations and anions, which are species with formal charge of ...-2,-1,+1,+2... Electrostatic interactions can be either attractive or repulsive, depending on the signs of the charges. The electrostatic interactions within a sodium chloride crystal are called *ionic bonds* (covalent). But when a single cation and a single anion are close together, within a protein, or within a folded RNA, those interactions are considered to be *non-covalent electrostatic interactions*. Electrostatic interactions can be very strong, and fall off gradually with distance (1/r).

2) Hydrogen bonds

In biological systems, it is common place of a Hydrogen atom to form non-coavalent bonds with two atoms simultaneously. This requires that the X-H (X for some electronegative atom, not necessarily a halogen) be acidic, and each of X-H-A X and A are able to donate a lone pair to the H-bond. *Special note:* C-H bonds are NOT Hydrogen bonds. C cannot form a hydrogen bond (not e^- enough)



Typically, H-bond acceptors are weakly basic, and H-bond donators are weakly acidic.



3) Van der Waals attraction

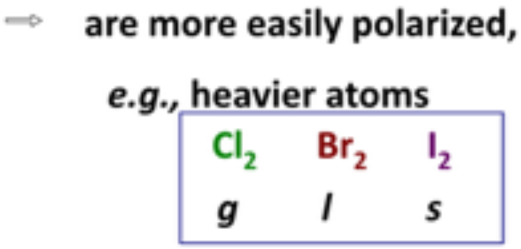
Also known as London Forces (or dispersion force), Van der Waals attraction is a weak force between any pair of molecules - even ones which have no permanent dipole movement.

Atoms behave as though they are surrounded by constantly swirling and shifting delocalized clouds of electrons. Even if a molecule has no permanent dipole moment (resulting from the unequal sharing of electrons in covalent bonds), it will still have a *transient* dipole moment, because, at any instant, the distribution of the electron clouds is not symmetrical. And a transient dipole moment in one molecule will exert a electrical force on any other molecule that is near, inducing a transient dipole moment in it. You might think that all these shifting dipoles would cancel out, but they don't; all the electron motions *correlate* and add up to a net attractive force. We call these forces "London forces" or "dispersion" forces, since they result from the "dispersion" of the electron clouds in molecules. Most people consider London forces to be a special sub-class of "van der Waals interactions".

London forces are a bit like gravity. Gravity is actually a very weak force (compared to electrical forces) but it's always attractive. So, when you have lots of mass, gravity adds up to a big force.

London forces are also very weak - individually, they are weaker than other van der Waals interactions. But ***every atom attracts every other atom*** by London forces, at least when the two atoms are fitting snugly together How tightly they can pack depends on their shapes. You can pack books into a carton very tightly, but you can't pack coffee mugs into a carton very tightly, because the shapes don't fit together very well.

Look at the examples below. None of these molecules has any charge or any significant permanent dipole moment, so their intermolecular attractions are mediated mainly by London forces. Heavier atoms have "looser" (more easily polarizable) clouds of electrons, so they experience bigger London attractions. That's why, for example, iodine is a solid, and bromine is a liquid, and chlorine is a gas, at room temperature.



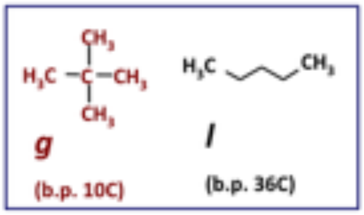
→ **are more easily polarized,**

***e.g.,* heavier atoms**

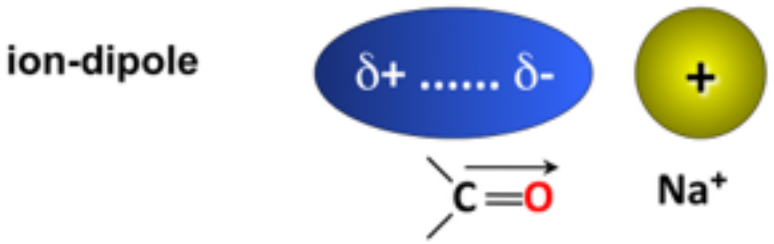
Cl_2	Br_2	I_2
g	*l*	*s*

Look at the hydrocarbons on the bottom. *n*-Pentane molecules pack together like books, so the substance is more "cohesive" and it is a liquid at room temperature (it boils at 36ºC). But *neo*-pentane (2,2-dimethylpropane) molecules pack together like coffee mugs, so it is a gas (it boils at 9ºC).

→ **can fit together snugly**



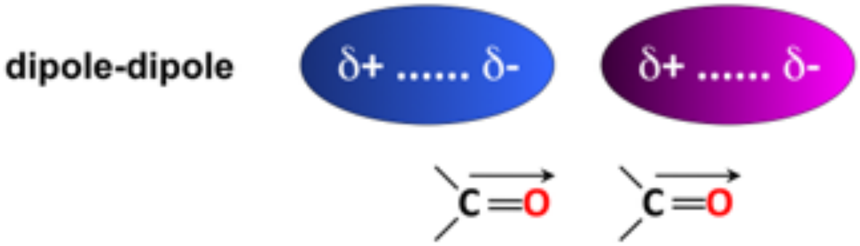
The following are two examples of Van der Waals forces:



An ion can attract a dipole, swinging it around so that the attractive interaction is closer and stronger than the repulsive interaction. And two dipoles can do the same thing.

Here, we are showing carbonyls (C=O) as examples of dipoles. The O is more electronegative than the C, of course.

Notice that van der Waals interactions of dipoles, like hydrogen bonds, are



4) Hydrophobic interactions

Hydrophobic interactions are a function of the tendency of polar molecules to attract polar molecules, resulting in a cluster of non-polar molecules that appears separate from the polar grouping. This is not a force, but an *effect* or *consequence* of Van der Walls interactions. To be clear, there is *no actual repulsion between polar and non-polar molecules* but a very large attraction between water molecules (cohesion) and thus a tendency to group.

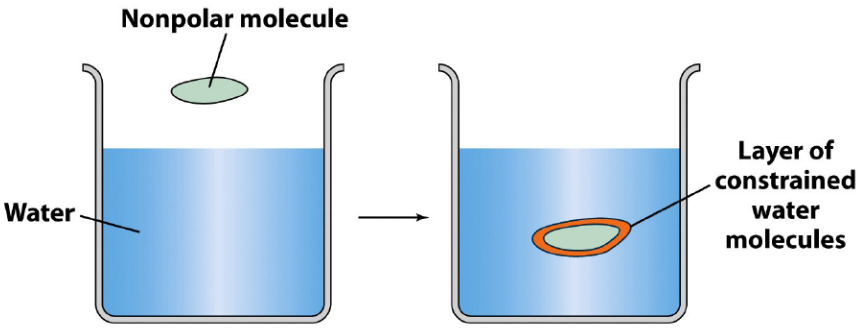
The "hydrophobic effect" drives formation of lipid bilayer membranes, the folding of proteins into globular molecules, the winding of DNA into the double helix, and many other aspects of biochemistry.

Etymology:

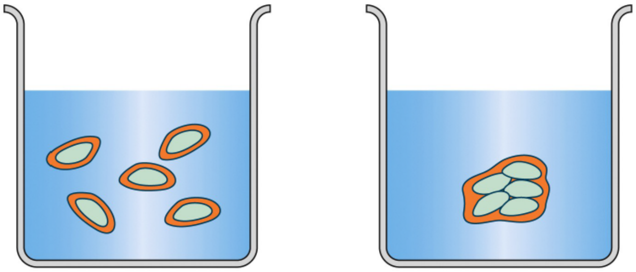
Hydrophilic **Hydrophobic**

hydros = water	*hydrology*
philos = love	*philosophy*
phobos = hate	*agoraphobia*

Example:



© John Wiley & Sons, Inc. All rights reserved.



Aggregated hydrophobic reduces volume of constrained water molecules

Lecture 3

Acid-base chemistry

Acid-base chemistry is extremely important in biological systems. The maintenance of pH plays a huge role in homeostasis - and is critical for the structure and function of biomolecules. As such, control of pH is one of the most critical parts of any biochemical experiment.

This lecture will be a review of **acid-base chemistry** and **buffers**.

Brønsted Acids and Bases

This course will use the Brønsted definition of acids and bases, namely:

Acids: a substance that can *donate* a proton
Base: a substance that can *accept* a proton

$$\underset{A}{HA} + \underset{B}{:B} \rightleftharpoons \underset{CB}{A^-} + \underset{CA}{HB^+}$$

Above is a generic A/B equillibirum reaction, one which follows *Le Chatelier's Principle*.

$$K_{eq} = \frac{[HO^-][H^+]}{[H_2O]}$$

$$K_w = K_{eq}\,[H_2O] = [HO^-][H^+] = 1 \times 10^{-14}$$

Water itself is *both an acid and a base*. It dissociates, as shown above. The amount of dissociation is tiny - about one water molecule in a billion. (This amount of dissociation is so small that it makes no discernible change in the concentration of the water itself.)

But even this little bit of dissociation matters, because the products (H^+ and OH^+) have very different properties from H_2O. The calculation shows you that neutral water has pH = 7, more or less. But this is only an approximate value, because the dissociation equilibrium depends on temperature.

The fundamental definition of neutrality is that $[H^+] = [OH^-]$.

Thus, [H3O+] = [OH-] = 1 x 10-7 M and pH = -log[H30+]

The pH scale:

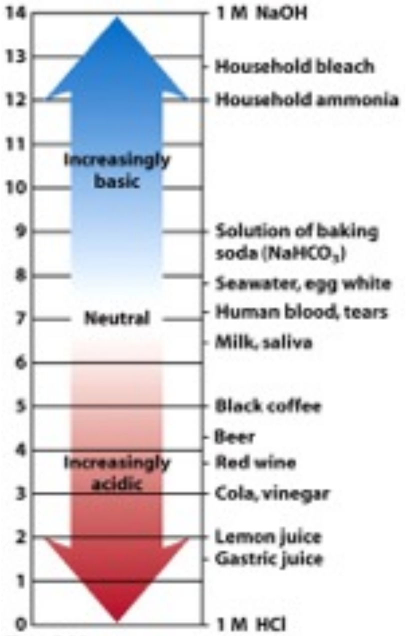


Figure 2-14
Lehninger Principles of Biochemistry, Fifth Edition
© 2008 W. H. Freeman and Company

We can rank acidities of various solutions on the pH scale. Never forget that it is a logarithmic scale, so decreasing the pH by one unit means increasing the H^+ concentration by a factor of 10.

Most biological systems are within a pH of 6 to 8.

Strength of an acid: Qualitative measure

Strong acids and bases completely (at least theoretically) dissociate within water. Therefore, all HA is converted in H^+ and A^-. Because of this stipulation, the pH can simply be calculated as follows pH = -log[H^+] = -log[HA].

Weak acids and bases are slightly more complicated. The acid (or base) dissociation constant (K_a or K_b respectively) is calculated from the weak acid/base equilibrium equation, and from this the pK_a can be found - one of the most common measures of acid strength in a biological system.

ex:)

Acid dissociation reaction

$$\mathrm{HA} \rightleftharpoons \mathrm{H}^+ + \mathrm{A}^-$$

acid conjugate base

$$K_a = \frac{[\mathrm{H}^+][\mathrm{A}^-]}{[\mathrm{HA}]} \qquad \mathrm{p}K_a = -\log K_a$$

Within biological systems, the **main rage of acidity and basicity is between -2 and 16 pK_a.**

The **Henderson-Hasselbalch equation** is used to calculate the pH of a weak acid/base equilibrium (i.e. a buffer!)

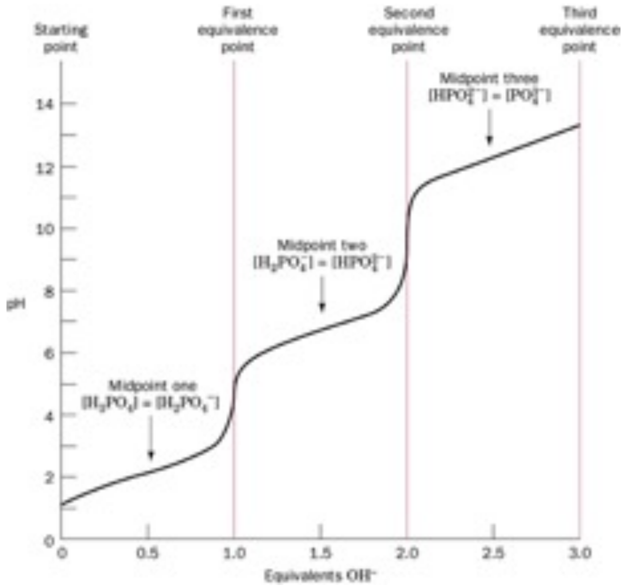
$$pH = pK_a + \log \frac{[A^-]}{[HA]}$$

One important result of this equation arrises when you set the $[HA] = [A^-]$. At which point, **$pK_a = pH$** *therefore the* pK_a of an acid is simply its pH when *half of the acid has dissociated*.

A stronger acid has a lower value of pK_a. That is, it dissociates at a higher concentration of H^+. So, we say, for example, that glycine ($pK_a = 2.3$) is a *stronger weak acid* than acetic acid ($pK_a = 4.7$).

Polyprotic acids

Polyprotic acids have more than one denotable hydrogen, and thus demonstrate a titration curve. Even closely related functional groups on a molecule will have differing pK_a's, mainly because the removal of one H^+ changes the molecule and affect affinity for the remain H^+'s.



A titration curve is calculable using the henderson hassel-balch equation.

if we let x = equivalents of OH^- added, and C_0 = equivalents of HA initially present, the pH of each curve simply becomes:

$$\mathrm{pH} = \mathrm{p}K_a + \log\left(\frac{x}{c_o - x}\right)$$

Buffers

Buffers are solutions that resist changes in their pH - and are extremely important in biological systems and biological experimentation. Typically, buffers are composed of a weak acid and its conjugate base. Acids and their conjugate bases are in equilibrium. Equilibria are related to the properties of the reactants and products, so for weak acids, the tendency to give up its proton determines its buffering property. The tendency to ionize can be put in an equilibrium equation - remember, when pH approaches pK_a, the equilibrium contains equal parts HA and A^-.

Solving Acid-Base Buffer Problems - Special Section

Case 1:

A case 1 problem is when only the “acid form” of HA is initially present. H^+ ions arise from the partial dissociation of HA and from the dissociation of H_2O. Because water is a very weak acid, we will assume that most of the $[H^+]$ ions come from the weak acid, HA.

- Set up an ICE table for the equilibrium reaction of the dissociation of the weak acid,
- Set the K_a up appropriately, and solve for X, which is your $[H^+]$.
- pH is then simple to calculate.

Remember, the -X in the denominator of the K_a equation can be ignored if $[HA_{init}] > 100 \times K_a$

required formulae:

- $pH = -\log[H_3O^+]$ or
- $pOH = -\log[OH^-]$ if base
- pH + pOH = 14 if base

- $k_a = 10^{-pKa}$

Case 2:

"*What is the pH of a solution of weak acid NaA?"* In case two, only the base form of the weak acid HA is initially present. The A^- will therefore tend to combine with any H^+ present - which in this case comes sole from H_2O. The dissociation of water and the reaction of H^+ and A- occurs simultaneously. A net reaction can be written, and K_{eq} determined simply by multiplying the K_{eq} of each reaction.

EQ6 $H_2O \rightleftharpoons H^+ + OH^-$ $K_{eq} = \frac{K_w}{[H_2O]} = 1.8 \times 10^{-16}$

EQ7 $H^+ + A^- \rightleftharpoons HA$ $K_{eq} = {}^1/K_a = 10^{+5}$

The overall reaction is the sum of the two:
EQ8 $H^+ + A^- + H_2O \rightleftharpoons HA + H^+ + OH^-$

$$K_{eq} = \frac{[HA][H^+][OH^-]}{[H^+][A^-][H_2O]} = \frac{1}{K_a}\frac{K_w}{[H_2O]}$$

The overall reaction is the sum of the two:
EQ8 $H^+ + A^- + H_2O \rightleftharpoons HA + H^+ + OH^-$

$$K_{eq} = \frac{[HA][H^+][OH^-]}{[H^+][A^-][H_2O]} = \frac{1}{K_a}\frac{K_w}{[H_2O]}$$

Its important to understand two caveats:

1) Since the molar concentration of water is exceptionally large, it can effectively be ignored. The x in the denominator can also be effectively ignored. Now this problem solves just like a case 1 problem.

Case 3:

"What is the pH of a solution containing HA and A^-"

With these problems, its essential to realize that the henderson-hasselbalch equation can be used to quickly find the pH of the solutions. However, when adding more acid/base to the system we must be more careful. If adding a strong base/acid, simply subtract the number of models added from the weak acid or base of the buffer, to determine how much will remain (remember, if you and x moles of strong base it will consume x moles of weak acid) this problem then becomes a case three problem once again.

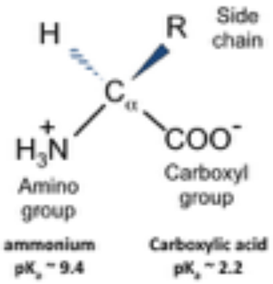
List of Formulae

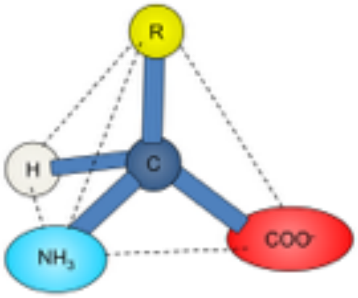
$pH = -\log[H_3O^+]$
$pOH = -\log[OH^-]$
$pH + pOH = 14$
$pH = pK_a + \log([A^-]/[HA])$; *Henderson-Hasselbalch*
$k_a = 10^{-pKa}$
$k_w = [OH^-][H^+] = 1 \times 10^{-14}$

Amino Acids

Protein is the best studied biopolymer. The building blocks or monomers of this polymer are of course amino acids. 20 of which, are used in the construction of proteins across most life forms.

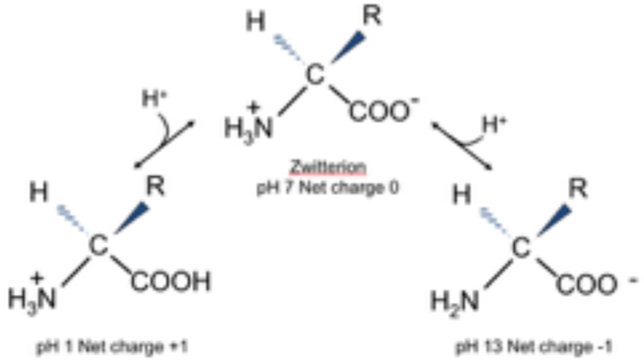
Amino acids have a general structure which is as follows (note, proline DOES NOT follow this convention):





Many biological molecules are highly sensitive to pH change, and amino acids are no exception. At neutral pH (~7.0) amino acids exist in their **Zwitterion** form. This is a really important concept to understand. *At neutral pH, all but 5 amino acids contain a neutral formal charge.* These include + charged Lysine, Arginine, Histidine, and negatively charged aspartic acid and glutamate.

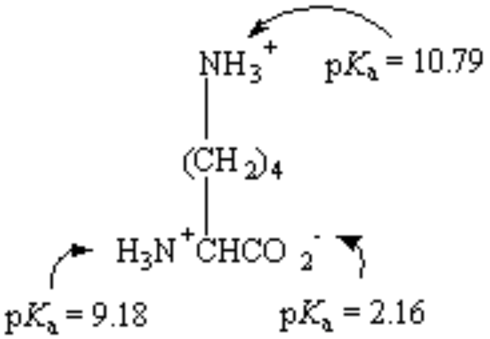
As the pH rises, there are more equivalents of A^- in solution accept H^+ ions from the amino acid, and the net charge will drop. A rise in pH of course, leads to the opposite effect.



Aside - questions involving amino acids

1.) *Being able to draw the predominant structure of an aa at a certain pH is crucial.* Begin by examining the pK_a of each functional group (COO^-, NH_3^+). If the pK_a is below the pH of solution, the functional group will be *unprotonated.* If the pK_a is above the pH, the group will be protonated.

Ex:) What is the predominate form of lysine at pH 7.0

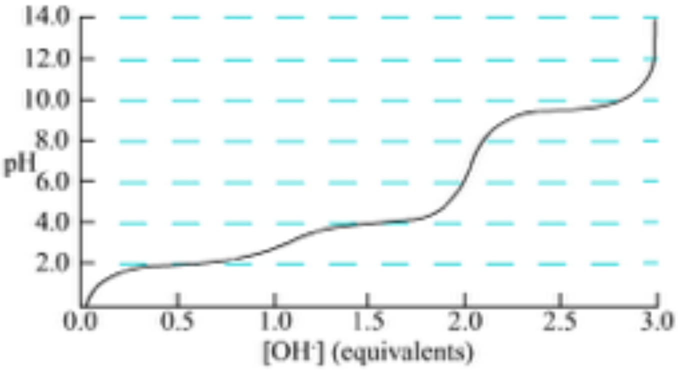


Ignore values given on the picture. Go with approximations

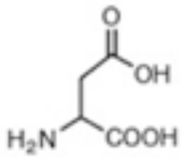
COO^- : pK_a is ~ 2.2 which is < pH unprotonated
NH_3^+ : pK_a is ~ 9.4 which is > pH protonated
R-group: pK_a is ~ 10.53 which is > pH protonated

2) Sometimes you will be asked to identify the aa associated with a titration curve plot. If so, follow these steps

1. Identify the # of titration curves. This = the number of denotable or acceptable H's of the aa
2. Draw suspected aa, and label each functional groups *approximate* pK_a's.
3. Remember, pH = pK_a when $[A^-] = [HA]$, so the pH reading on the graph of each buffer zone should equal the pK_a of the aa in one of its forms.



Lets guess aspartate, which contains three buffer regions. As usual we estimate pKa's of 2.2 and 9.4 for the carboxyl and amine functional groups. We know the R group has a pKa of ~3.65. The pH of our buffer regions are 2, 4, and 10.



Aspartic Acid (Asp, D)
MW: 115.09, pK$_a$ = 3.9

Thus, our guess seems to be correct. The donation of H's is from the carboxyl, then amide, then r group with ~ 2, 4 ,10 pH of those functional groups

Amino acids can join via peptide bonds

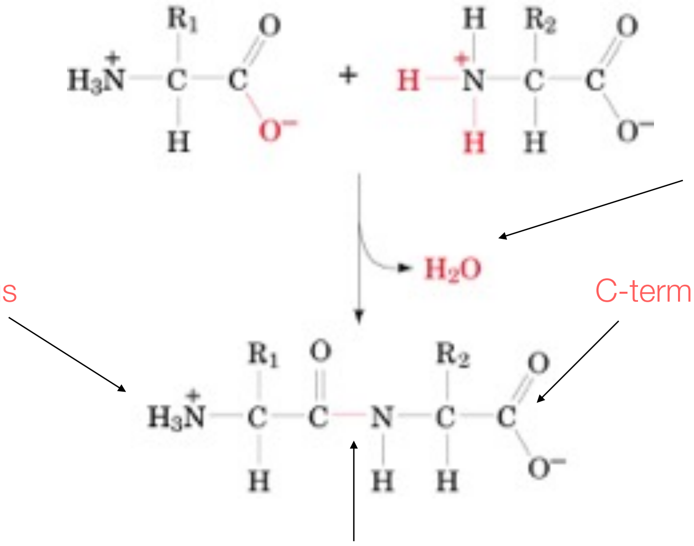
Polymerization of amino acids occurs in a *head-to-tail* fashion. The amino (-NH_3^+) and carboxyl (COO^-) groups allow AA to undergo a condensation reaction - eliminating a water molecule - and forming a covalent amide bond or **peptide bond.**

Peptides can be

- dipeptide (2 aa residues)
- tripeptide (3 aa)
- oligopeptide (a *few* aa)
- polypeptide (many aa)

Proteins themselves, are molecules that consist of one or more polypeptide chains, each ranging from 40 to 33,000 aa.

Condensation of two aa to form peptide bond:



Notice lose of water, and where this came from

N-terminus

C-terminus

peptide or amide bond

Peptides are written from *left to right*, or **N-terminus** to **C-terminus.** We will explore their formation (biosynthesis) later in the semester.

Amino acids structure and classification

There are 20 standard amino acids used to build the polypeptide chains that constitute life. Typically, there are 3 main groups based on the R-groups, or side chains, on the a backbone.

Nonpolar
- non-polar aliphatics, aromatics

Polar, uncharged
- Alcohols, Thiols, Amides

Charged polar side chains
- three bases and two acids

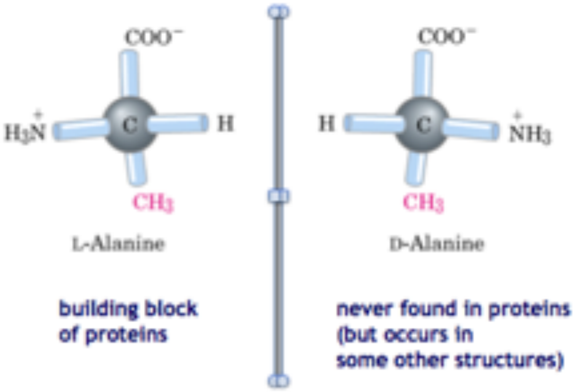
From these 20 choices we get diversity!

Lecture 4

Amino Acids - Cont'd

Most amino acids are chiral (except for glycine) (contains at least one stereo-centre, and therefore enantiomeric forms). In place of common R and S nomenclature (Cahn-Ingold-Prelog) amino acids rely on the D and L Wsterochemical descriptors, known as the **Fisher-Rosanoff** convention.

Its important to note that the L form of each amino acid are always used as the building blocks of protein. D forms are fond elsewhere in nature, but *never in proteins*. There is no chemical reason, but consistently using one form is crucial (imagine assembling a book shelf from left and right handed screws).



Identifying the D or L configuration of an amino acid

1. Place highest oxidation state chain group pointed up and *in* (note: this is usually COO^-)
2. Place longest (most carbons) chain pointing down and in
3. Left and Right substituents are horizontal and pointing *out of the plane*

If the substituent with the highest prority is on the left, this aa is in the L form. Otherwise, it is D.

ex:)

S amino aicd (for all R except Cys, R = CH_2SH)

L-amino acid

All amino acids are L including Cys

Amino acid structure, descriptions are *omitted here*. See Lecture 4 PPT if they are necessary.

Lecture 5