

**Chemistry 5.07SC Biological Chemistry I**  
**Fall Semester, 2013**

**Lecture 1. What is Biochemistry?** Life at the molecular level.

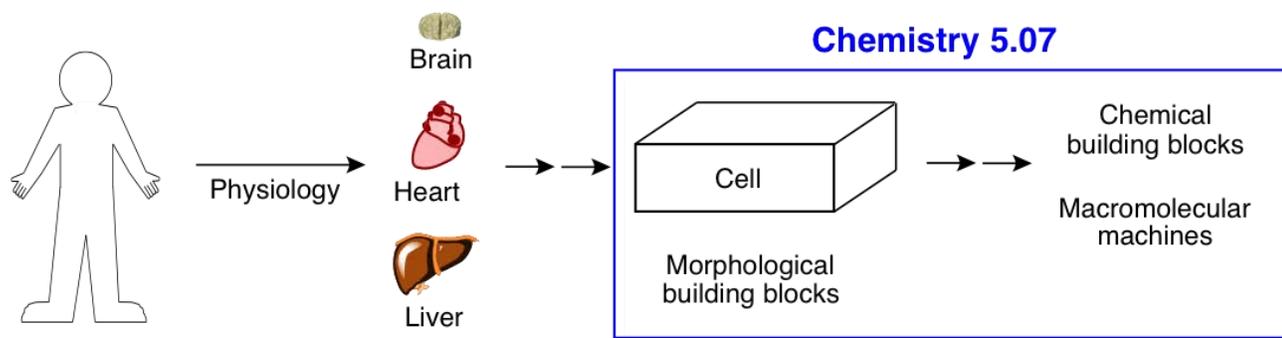


Figure 1. Chemistry 5.07 focuses on the study of life at the molecular level.

**A. What is Life?** The ability to reproduce and to make order from chaos (Chapter 1 of course textbook).

**B. All known life forms share certain properties.**

1. All organisms have the same morphological unit of life: **the cell**. Fleas and elephants have similar sized cells.

The crowded interior of the cell looks like a jungle (Figure 2). However, studies in the past decade are revealing amazing organization within what appears to be chaos. How and when you look and removing redundancy of function is essential to seeing intracellular organization. The crowdedness also requires one to think about the concept of “free” proteins or “free” small molecules. Disruption or dilution of the cell environment required for biochemical analysis, changes all the equilibria. Recent estimates suggest that macromolecules are present at 350 mg/mL and that small molecules are present at 300 mM.

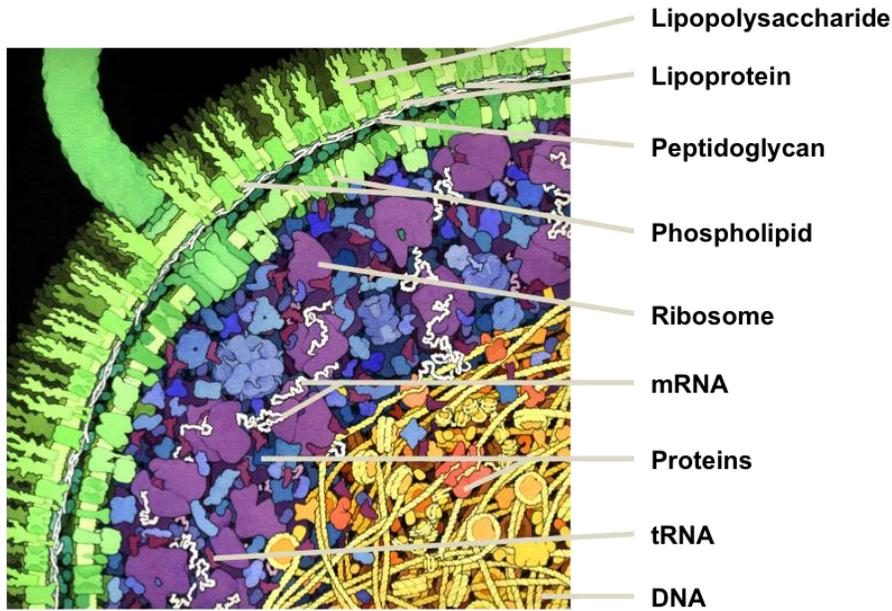


Illustration courtesy of [David S. Goodsell](#), the Scripps Research Institute.

Figure 2. The interior of the cell is crowded and appears chaotic as seen in this illustration of a bacterial cell magnified  $10^6$  fold. However, it is not simply a bag of enzymes and nucleic acids each occupying its own space. Cytoplasmic space is jammed full of proteins (dark yellow, pink, orange, light, medium and dark blue), tRNA (dark purple), mRNA (cream), ribosomes (light purple) and the DNA (yellow ribbons). The inner membrane is crammed with phospholipids (medium green) and various proteins and protein complexes (light green), while the periplasmic space contains the peptidoglycan (white), which is linked to the inner face of the outer membrane via lipoproteins (dark green). The outer membrane contains a variety of proteins and protein complexes interspersed with phospholipids, as well as lipopolysaccharides.

The crowded interior of the cell looks like a jungle. However, studies in the past decade are revealing amazing organization within what appears to be chaos. How and when you look and removing redundancy of function is essential to seeing intracellular organization. The crowdedness also requires one to think about the concept of “free” proteins or “free” small molecules. Disruption or dilution of the cell environment required for biochemical analysis, changes all the equilibria. Recent estimates suggest that macromolecules are present at 350 mg/mL and that small molecules are present at 300 mM.

2. The solvent of life is  $H_2O$  due to its unusual properties. [Think about H bonding and buffers, reviewed in the lexicon.]

3. All the building blocks in these organisms are the same: nucleotides, to make DNA (deoxyribonucleic acids) and RNA (ribonucleic acids); acetylCoA, to make fatty acids and triacylglycerols; sugars, to make polysaccharides; amino acids, to make proteins (polypeptides). **The chemistry of life is conserved.** There are a limited number of chemical transformations that can account for the vast majority of transformations in pathways (see the Lexicon).

4. Vitamins are conserved. Derivatives of vitamins extend the chemical repertoire of protein catalysts and are essential for life. Look at the label on your vitamin bottle. You will understand the functions of these molecules by the end of the semester.

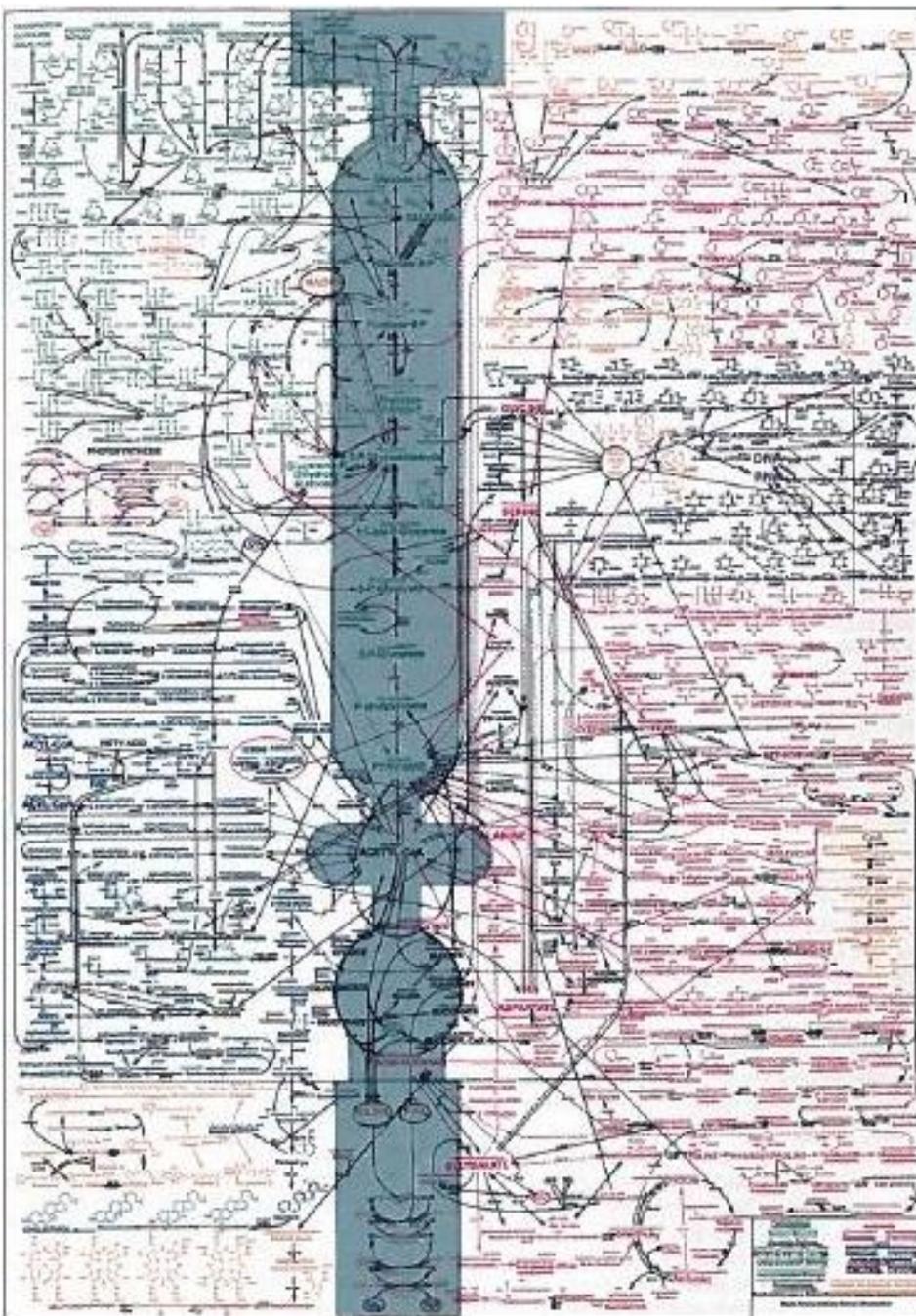
5. All organisms share the same mechanism of information transfer: DNA is transcribed into RNA, that is translated into protein. In 5.07 we will **NOT** focus on DNA replication, transcription, or translation. This information and the macromolecular machines involved in these processes are the focus, in part, of Chemistry 5.08. We will consequently not discuss nucleotide and nucleic acid structure.

6. The energy currency, required to make order out of chaos, is the same in all organisms: ATP and GTP. A 70 Kg person makes 45 Kg of ATP every 24 h. The machinery to make ATP is conserved.

7. Primary metabolic pathways, both biosynthetic (anabolic) and degradative (catabolic) pathways are conserved. This semester we will investigate glycolysis/gluconeogenesis; fatty acid biosynthesis/fatty acid degradation; the TCA cycle; the respiratory chain and how the energy provided by reducing equivalents is used to make ATP; the pentose phosphate pathway.

Environmental changes alter fluxes through the metabolic pathways in the chart below from the course textbook. What are the general principles that control the changes required to meet and maintain the organism's needs? We will return to this chart over and over again. We hope to convince you that metabolism is NOT the incredible jungle that it first appears. We will focus on central metabolism, which is essential to almost all organisms: glycolysis, TCA, ET, and O/P to make

ATP. You are expected to know the structures of the central metabolites (glucose, acetyl coenzyme A, pyruvate, etc).



© John Wiley & Sons. All rights reserved. This content is excluded from our Creative Commons license. For more information, see <https://ocw.mit.edu/help/faq-fair-use/>.

Figure 3. Cellular metabolism is extremely complex and is the sum of its many substrates, products, macromolecular machines, biosynthetic and catabolic pathway, which provide life.

**C. Pathways are tightly regulated.** How do we control if fatty acids are synthesized or degraded or whether sugars are synthesized or degraded? Regulatory mechanisms have also been conserved but, in general, the more complex the organism, the more complex the regulatory mechanisms. You will be introduced to the general mechanisms of regulation.

1. Allosteric regulation: small molecules bind outside the site where chemistry occurs in an enzyme and affect the rate at which the enzyme converts a substrate to product.

2. Feedback inhibition: an end product of a pathway, may bind to the first enzyme in the pathway to turn the pathway off when too much of the end product is present.

3. Post-translational modification: many activities of enzymes are modified by covalently modifying the protein after it is made by the translational process. Proteins can be phosphorylated, acetylated, hydroxylated, ubiquitinated, methylated, glycosylated etc.

4. Hormonal regulation: binding of hormones (extracellular) to receptors in membranes that transmit information to the cellular interior. Examples encountered will be epinephrine, insulin, and glucagon.

5. Transcriptional and translational regulation (not discussed in 5.07)

**D. Biological relevance.** Why should we care about biochemistry, which encompasses all of the above topics? Understanding regulation of the “Fed” state and the “Fasting” state, and how an organism switches its metabolism to accommodate these states is central to understanding chronic disease mechanisms. Most diseases are associated with metabolic dysregulation. Thus, understanding biochemistry plays a central role in dealing with and reversing diseased states.

### **E. Explicit GOALS of Chemistry 5.07**

1. To introduce you to all of the chemical players of life: their structures and chemistry and, consequently, function. Without this background, you cannot understand biochemistry.

2. To introduce you to the central pathways of metabolism. At first glance, metabolism can be daunting, but we shall show you that all of biochemistry is made up of a surprisingly small number of chemical transformations, which are used again and again. Understanding these transformations is

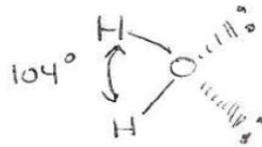
sufficient to allow you to predict most metabolic interconversions. See the Lexicon of reactions for the course.

3. Once the pathways have been introduced, one needs to understand the regulation of the pathways and their integration under different environmental conditions. This is the future.

**F. The path ahead.** ‘The more we learn about these complex systems, the more we realize that our knowledge is dwarfed by our ignorance.’ This is a fascinating discipline with deep roots in both basic and applied science ... and there is still much to do.

### Cellular components H<sub>2</sub>O and amino acids:

**I. Water** is essential for life and its properties are central to the function of all biological molecules.

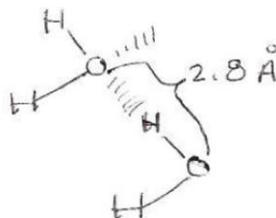


H<sub>2</sub>O has two O-H bonds (0.95 Å) and has two pairs of non-bonded electrons that play a key role in forming weak, non-covalent H (hydrogen) bonds.

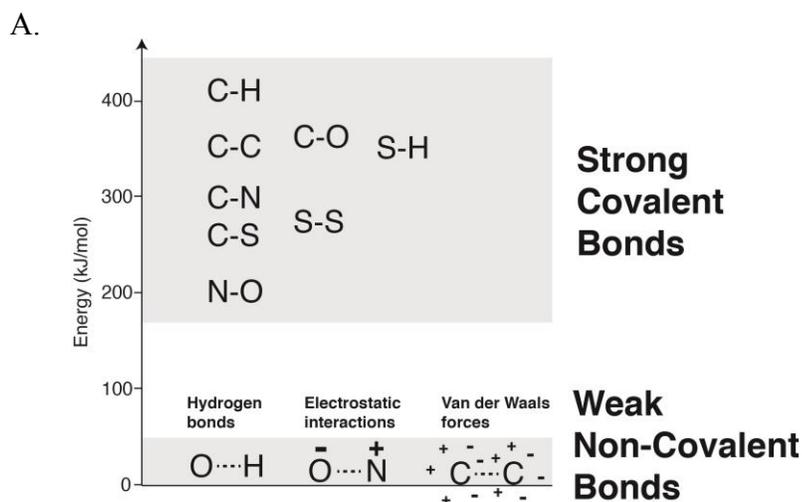
There are four unique, physical properties of H<sub>2</sub>O:

- A. H<sub>2</sub>O can form H bonds with itself and with many biological molecules.
- B. H<sub>2</sub>O has dipolar character and plays an important role in the solvation of many molecules.
- C. H<sub>2</sub>O minimizes interactions with non-polar species.
- D. H<sub>2</sub>O is a neutral molecule with the ability to ionize. **You have/will discuss pH and pKa.**

Definition of an H bond: if an H is bonded to a very electronegative atom (in biology this atom is usually O or N), it can form an additional weak bond with another electronegative atom that has a pair of non-bonded electrons.



More generically, an H bond exists between a donor (D) and an acceptor (A) (that is, D-H---A), when the distance between D and A is 2.7 to 3.1 Å. D and A are electronegative atoms O, N and sometimes in Biology S. H bonds are directional, but not as directional as covalent bonds. A good H bond does not need to be linear.



B. **H bonds found in nature**

Example	Donor $\cdots$ Acceptor	Bond length (Å)
Formed in water	$\text{—O—H} \cdots \text{O} \begin{array}{l} / \\ \text{H} \end{array}$	$2.8 \pm 0.1$
Often formed between water and other molecules	$\text{—O—H} \cdots \text{O}=\text{C} \begin{array}{l} / \\ \backslash \end{array}$	$2.8 \pm 0.1$
Often formed between water and other molecules	$\begin{array}{l} / \\ \backslash \end{array} \text{N—H} \cdots \text{O} \begin{array}{l} / \\ \text{H} \end{array}$	$2.9 \pm 0.1$
Proteins & nucleic acids	$\begin{array}{l} / \\ \backslash \end{array} \text{N—H} \cdots \text{O}=\text{C} \begin{array}{l} / \\ \backslash \end{array}$	$2.9 \pm 0.1$
Proteins & nucleic acids	$\begin{array}{l} / \\ \backslash \end{array} \text{N—H} \cdots \text{N} \begin{array}{l} // \\ \backslash \end{array}$	$3.1 \pm 0.1$
Rare and weaker than above H bonds	$\begin{array}{l} / \\ \backslash \end{array} \text{N—H} \cdots \text{S} \begin{array}{l} / \\ \backslash \end{array}$	3.7

Figure 6. Types of bonds and bond energies. A. Bond energies for strong vs. weak non-covalent bonds (previous page). B. Examples of H bonds found in nature, including schematics of the donor and acceptor, as well as their bond lengths.

H bonds are on the order of 2 to 7 kcal/mol and fit into the category of weak non-covalent bonds. These weak interactions (electrostatic, van der Waals,  $\pi$  cation, dipole-dipole, dipole-induced dipole, etc) play an essential role in the structure and function of all biological macromolecules in aggregate.

## II. Amino acids, the building blocks of proteins.

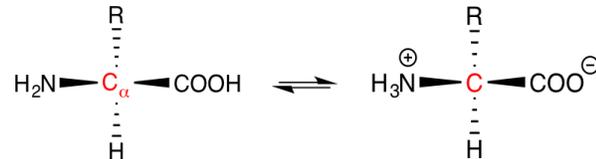


Figure 7. The structures above are Fischer projections of amino acids. The C is tetrahedral and the R and H substituents go into the plane of the paper, while the amino and carboxylate groups, come out of the plane of the paper. In aqueous solution, amino acids are **not neutral** (left), but are zwitterionic (right) where the amino group is protonated and the carboxylic acid group is deprotonated.

All amino acids with the exception of glycine ( $R = H$ ) are chiral. Their carbon possesses four different substituents. Note that the R group may have additional chiral centers. Refresh your understanding of enantiomers (two molecules that are non-superimposable mirror images and indistinguishable except for their ability to rotate plane polarized light) and diastereomers. **Chirality is central to biology.** Most amino acids are in the S configuration and are L amino acids (based on a comparison with L-glyceraldehyde).

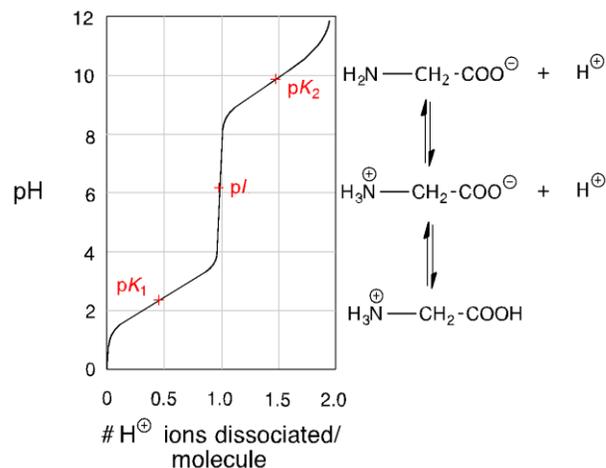
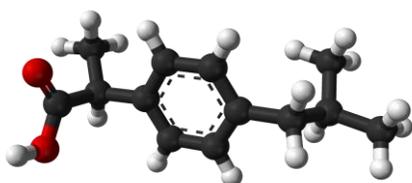


Figure 8. Titration experiment of an amino acid as a function of pH. The pKa of the acid group is around 2.5 and that of the ammonium group is around 10. Thus at pH 7 (interior of the cell), free amino acids exist in the zwitterionic form.

## The importance of chirality:

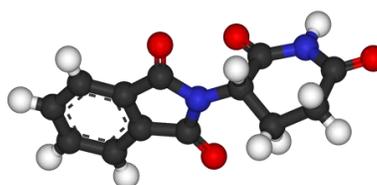
The importance of chirality can be seen using two drugs (shown below), Advil and Thalidomide, as examples. For Advil, though only one enantiomer is effective as a painkiller, the other enantiomer is not harmful. In the case of thalidomide, one enantiomer was helpful for pregnant women suffering from morning sickness and was prescribed in the late 1950s and early 1960s. However, one of the enantiomers caused terrible teratogenic-fetal defects. Though the exact mechanism of action has not been elucidated, thalidomide is now used to treat multiple myeloma cancers and a condition called erythema nodosum leprosum associated with leprosy and is prescribed only under extremely careful medical supervision.

A.



Credit line to be added: Public domain on Wikipedia.

B.



Credit line to be added: Public domain on Wikipedia.

Figure 9. Drugs with medically important chirality. A. Advil (Ibuprofen, Motrin, Nuprin), a painkiller. B. Thalidomide.

There are 21 amino acids (actually 22, a new amino acid pyrrolysine was discovered several years ago and its biosynthesis worked out in 2011) that are distinguished by their R groups. **You are required to know the structures of the amino acids and the one and three letter codes for each of them.** Below is a table with the structure, residue mass, and pKa for each of the common amino acids. The R groups of amino acids play a central role in defining protein structure and function and thus in addition to knowing the structures of each amino acid, you also need to **know a ball park pKa of the side chains** that can change protonation state under physiological conditions.

## Important points about amino acids to remember:

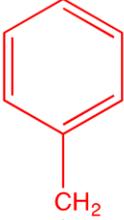
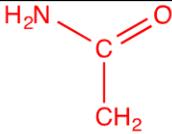
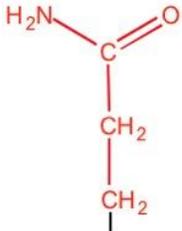
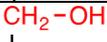
1. G (R = H) is achiral and is **flexible**. This is the only achiral amino acid. This amino acid, due to its flexibility, is often conserved in protein sequences. G plays an important role in the active sites of enzymes that catalyze phosphoryl transfer reactions and that use phosphorylated substrates (ATP, NADH). You will see that  $Mg^{2+} \cdot ATP$ ,  $Mg^{2+} \cdot GTP$  and many

of the vitamins (riboflavin, niacin, etc) have binding sites in the protein for phosphate that are enriched in G residues.

2. P (proline) is the only amino acid that has a secondary amine. When proline is part of a protein, it has thus lost its ability to form a H bond. Thus prolines are often secondary structure breakers (secondary structure of proteins will be defined in the next lecture).

3. H, C, E, D, K, S, T, Y All of these amino acids have a pKa of their side chain in the physiological range and thus can play an important role in catalysis by enzymes. You will learn, however, that the **pKa of these side chains can be perturbed by the protein environment by five or six orders of magnitude!**

Amino acid (3-letter name, 1-letter name)	Side chain R	Molecular weight/ Residue mass (Da)	-NH <sub>2</sub> pKa	-COOH pKa	Side chain pKa	pI
Glycine (Gly, G)	H 	57.05	9.60	2.34	-	5.97
Alanine (Ala, A)	CH <sub>3</sub> 	71.09	9.69	2.34	-	6.00
Leucine (Leu, L)	$  \begin{array}{c}  \text{H}_3\text{C} \quad \text{CH}_3 \\  \quad \diagdown \quad / \\  \quad \text{CH} \\    \\  \text{CH}_2 \\    \\  \text{CH}_3  \end{array}  $	113.16	9.60	2.36	-	5.98
Isoleucine (Ile, I)	$  \begin{array}{c}  \text{CH}_3 \\    \\  \text{CH}_2 \\    \\  \text{H}_3\text{C} - \text{CH} \\    \\  \text{CH}_3  \end{array}  $	113.16	9.60	2.36	-	6.02
Valine (Val, V)	$  \begin{array}{c}  \text{H}_3\text{C} \quad \text{CH}_3 \\  \quad \diagdown \quad / \\  \quad \text{CH} \\    \\  \text{CH}_3  \end{array}  $	99.14	9.62	2.32	-	5.96

Phenylalanine (Phe, F)		147.18	9.21	1.83	-	5.48
Tryptophan (Trp, W)		186.21	9.39	2.83	-	5.89
Tyrosine (Tyr, Y)		163.18	9.11	2.20	10.07	5.66
Asparagine (Asn, N)		114.11	8.80	2.02	-	5.41
Cysteine (Cys, C)		103.15	10.28	1.96	8.18	5.07
Glutamine (Gln, Q)		128.14	9.13	2.17	-	5.65
Methionine (Met, M)		131.19	8.95	2.28	-	5.74
Serine		87.08	9.15	2.21	-	5.68

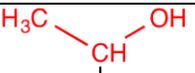
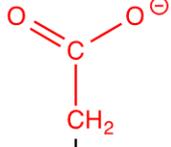
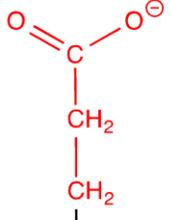
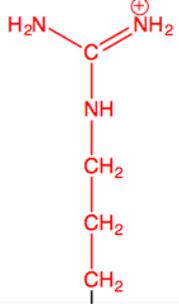
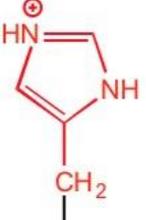
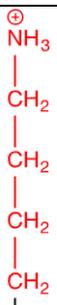
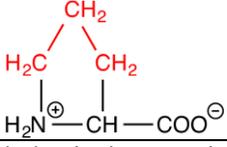
(Ser, S)						
<b>Threonine</b> (Thr, T)		101.11	9.10	2.09	-	5.60
<b>Aspartic Acid</b> (Asp, D)		115.09	9.60	1.88	3.65	2.77
<b>Glutamic Acid</b> (Glu, E)		129.12	9.67	2.19	4.25	3.22
<b>Arginine</b> (Arg, R)		156.19	9.04	2.17	12.48	10.76
<b>Histidine</b> (His, H)		137.14	9.17	1.82	6.00	7.59
<b>Lysine</b> (Lys, K)		128.17	8.95	2.18	10.53	9.74
<b>Proline</b> (Pro, P)		97.12	10.60	1.99	-	6.30

Table 1. The structure and physical properties of the common 20 amino acids. The full, 3-letter, 1-letter name, as well as the side chain and physical properties of each amino acid are given.

4. Amino acids are the building blocks of proteins. To study proteins one needs to know their concentrations ( $[ ]$ ). The  $[ ]$  can be determined from their extinction coefficient ( $\epsilon$ ). Proteins have  $\lambda_{\max}$  at 280 nm. This absorption feature is distinct from nucleic acids (DNA, RNA) that have  $\lambda_{\max}$  at 260 nm. The amino acid side chains that are largely responsible for the  $A_{280\text{nm}}$  are W and Y. From Beer's law  $[A = \epsilon lc]$  where  $A$  is the absorbance,  $l$  is the path length of light and  $c$  is the concentration of your protein, one can calculate the protein concentration from the  $A_{280\text{nm}}$  and the  $\epsilon$ . There is a website, called [Expasy](#), where one can calculate the  $\epsilon$  of any protein in the denatured (unfolded) state.

As discussed above, with the exception of  $R = H$ , all amino acids are chiral. Think about the significance of the amino acid stereochemistry and how chirality affects protein conformation.

Proteins in the cell function in structural, catalytic, or receptor capacities and are built from amino acids. In Chemistry 5.07, we will discuss proteins in each of these roles:

A. Collagen, which is the most abundant protein in vertebrates is extracellular and organized into insoluble fibers of great tensile strength. Collagen is a major stress bearing component of connective tissue.

B. Enzymes are proteins that catalyze all of the transformations in the metabolic pathways. They can accelerate rates of reactions relative to non-enzyme catalyzed reactions by rates of  $10^6$  to  $10^{15}$ . We will discuss mechanisms of catalysis and you will be given a lexicon of catalytic transformations that cover most reactions in the metabolic pathways we will examine.

C. Receptors, which are proteins (found in membranes) play a key role in signal transduction and you will discuss G-protein receptors in insulin/glucagon (fed/starved states) regulation and in adrenaline (fear) response.

In addition to the 22 amino acids, many amino acids are modified subsequent to their incorporation into proteins (the translational process) as shown in Figure 10. This process is called post-translational modification or PTM. These amino acids derivatives have unique properties that allow them to play a central role in the chemistry of enzyme-catalyzed reactions or in the regulation of the function of a protein. Mass spectrometric technology has allowed us to identify these modified

amino acids and some of these experimentally verified amino acids are shown in Figure 10. A decade ago we had no idea of their prevalence. Below are a number of examples of amino acids generated by PTM. These modifications include acetylation, methylation, hydroxylation, phosphorylation, attachment of sugars, proteins, etc.

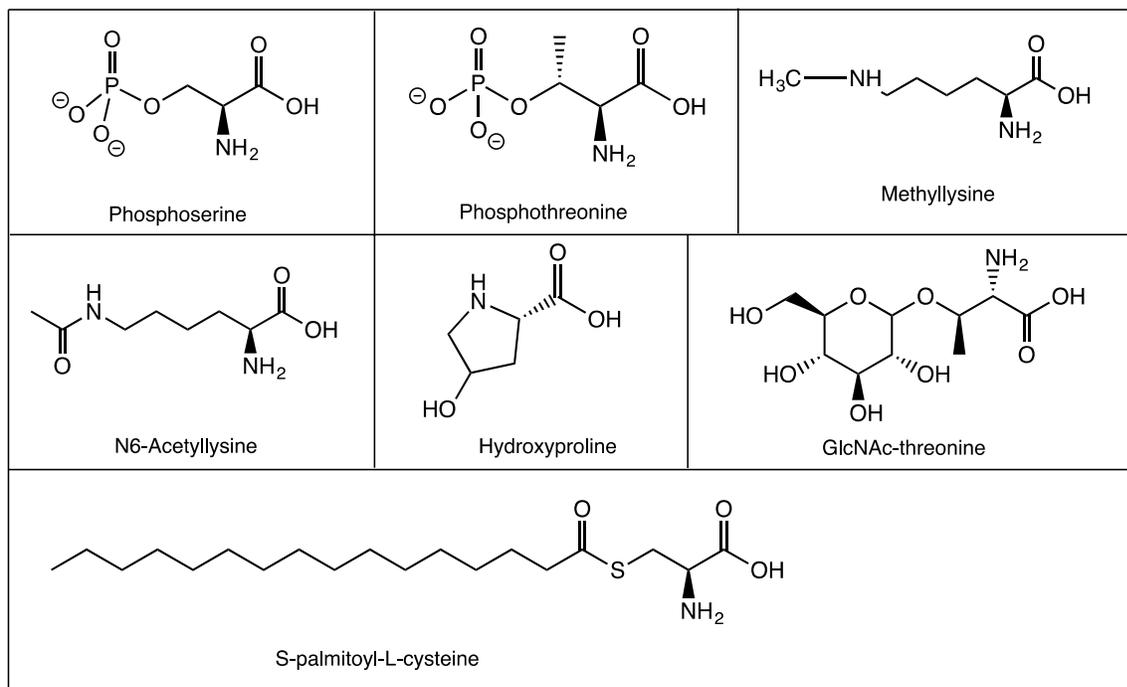


Figure 10. Sample of experimentally-verified post-translationally modified amino acids. These modifications include phosphorylation, methylation, acetylation, attachment of a lipid moiety, hydroxylation, and attachment of a sugar moiety.

MIT OpenCourseWare  
<https://ocw.mit.edu>

5.07SC Biological Chemistry I  
Fall 2013

For information about citing these materials or our Terms of Use, visit: <https://ocw.mit.edu/terms>.