**Expectations for Biochemistry Majors**

Biochemistry, the study of how molecules - small metabolites and large macromolecules - carry out the complex functions of living cells, lies at the intersection of three basic sciences: Biology, Chemistry, and Physics. It draws upon general principles operating in these broad areas of knowledge. For this reason, students undertaking training in biochemistry must first obtain mastery of several more fundamental subjects and must be proficient in certain quantitative skills; typically, students obtain this mastery by the end of sophomore year, so as to be ready for BCHM 100 in junior year. This handout aims to alert students to what they will encounter in the Brandeis Biochemistry curriculum, and to the already-mastered skills that will be expected of them in the core courses for the Biochemistry major. These fall into the areas of **mathematics**, first-year **physics,** general **chemistry** concepts, and **organic chemistry**. The specific elements within these areas are detailed below, followed by the particular skills that students will use frequently in each of the core courses for Biochemistry majors.

**1. Mathematical skills expected for Biochemistry majors** (after taking MATH 10a,b or equivalent)
> All majors will have taken algebra, geometry, and pre-calculus in high school, and many will have taken calculus. Some may have placed out of calculus with high AP scores. But it's easy to get rusty in math if you have not used it for a year or two. In that case, you will have to sharpen your math skills so that you can easily solve problems such as the sample problems below.
>
> **Algebra and pre-calculus mechanics:**
> * Solution of algebraic equations, including quadratic equations
> * Understanding intuitively the 'shapes' of functions in graphical representation
> * Solving 'word problems' by translating them into algebraic language
> **Univariate calculus:**
> * Familiarity with basic analytic functions: polynomials, logarithms, exponentials, trig functions
> * Methods of differentiation, including chain, product, and quotient rules
> * Determining limits, slopes, and curvatures of functions
> * Finding maxima and minima of functions
> * Methods of integration, including substitution, parts, and partial fractions
> * Calculating numerical or algebraic values of definite integrals

**2. Knowledge of introductory physics expected for Biochemistry majors** (after taking PHYS 11)
* Handling problems of position, velocity, acceleration with calculus
* Kinetic and potential energy and relation to force
* Conservation of energy and momentum
* Coulomb's law and basic electrostatics

**3. General chemical knowledge expected for Biochemistry majors** (after taking CHEM 11 or 15)
* Nature of chemical bonds, electron orbitals
* Equilibrium constants and their practical use in problems
* Gibbs free energy and its relation to equilibrium constants
* pH and its relation to proton-binding equilibrium constants
* Fundamentals of chemical kinetics, rate laws
* The difference between thermodynamics and kinetics

**4. Elements of organic chemistry** (after taking CHEM 25)
* Redox states of carbon compounds
* Carbonyl chemistry, proton abstraction, and reaction mechanisms

**Math self-test exercises**

Here are some problems (along with answers for some of them) for you to check your familiarity with math tools that you will be expected to know from your previous math courses. If you find that most of these problems stump you, you will need to sharpen your tools, or to consider a different major.

1. SKETCH roughly the functions on the interval (-1,4):

   a. $f(x) = (x-1)^2$ $\qquad$ $f'(x) = 2x - 2$ $\quad$ min at x=1
   b. $f(x) = e^{-x}$

   c. $f(x) = e^{-3x}$ $\qquad$ $f'(x) = -3\ e^{-3x}$ $\quad$ no extremum
   d. $f(x) = 1 - e^{-2x}$

   e. $f(x) = e^{1/x}$ $\qquad$ $f'(x) = (-1/x^2)e^{1/x}$ no extremum
   f. $f(x) = x^2 e^{-2x}$ $\qquad$ $f'(x) = 2xe^{-2x}(1-x)$ max at x=1, min at x=0
   g. $f(x) = \ln((rc-1)/x^2)$

   h. $f(x) = 1/(1+x^2)$ $\qquad$ $f'(x) = -2x/(1+x^2)^2$ $\quad$ max at x=0
   i. $f(x) = \exp(-x^2)$
   j. $f(x) = \sin(2x + rc/2)$ $f'(x) = 2\cos(2x+rc/2)$ max and min at $2x+rc/2 = nrc/2$ where n is odd !!
   k. $f(x) = e^{-2x} - e^{-3x}$ $\qquad$ $f'(x) = -2e^{-2x} + 3e^{-3x}$ $\quad$ max at x = $\ln(3/2)$
   l. $f(x) = x\ln x$
   m. $f(x) = 2x/(1+2x)$

   n. limit x=>0 of $x\ln(x)$ [Hint: L'Hopital's rule!]

2. Calculate the first derivatives of the functions in (1). For each one that has an extremum, find it.
See derivatives and extrema above

3. Calculate and sketch on a crude graph the definite integral:

$$\int_0^1 e^{-2x}\,dx = -e^{-2x}/2\Big|_0^1 = \frac{1-e^{-2}}{2}$$

4. Integrate by parts:

$$\int xe^{3x}\,dx = (e^{3x}/3)(x-1/3) \qquad \int \ln(2/x)\,dx = x(\ln(2/x) + 1)$$

5. For the following functions f(x), find the special value of x, x*, where f(x) is maximum. Also find the value of the function at this point, f(x*).
   $f(x) = xe^{-x}$ $\qquad$ x*=1 $\quad$ f(1) = 1/e
   $f(x) = x^2 e^{-x}$ $\qquad$ x*=2 $\quad$ f(2) = 4/e²
   $f(x) = x^3 e^{-x}$ $\qquad$ x*=3 $\quad$ f(3) = 27/e³

2

Roughly sketch these functions on the same graph to compare them.

Generalize the above results – what happens to $x^n e^{-x}$ as n varies?

Make a "*Table of useful integrals*" section in your notebook where you can quickly refer as you do problems. ***You*** should derive the following integrals and write down the results in this section. You will be using these in various problems, and it will be very useful for you to know

(1) that there is a place you can easily look them up to save time, and

(2) that you are the mathematician who evaluated them (rather than some anonymous geek).

1. $\int e^{-kx} dx$

2. $\int_0^\infty e^{-kx} dx$

3. $\int xe^{-x} dx$

4. $\int_0^\infty xe^{-x} dx$

5. $\int_0^\infty xe^{-kx} dx$

YOU do these - Check your answer by differentiating!

(Evaluate integral #5 without actually doing the integration, but rather by taking the result of integral #4 and using substitution. This is a useful trick that will save you lots of time.)

6. $\int x^2 e^{-x} dx$

7. $\int_0^\infty x^2 e^{-kx} dx$

8. $\int \frac{1}{1+kx} dx$

9. $\int \frac{x}{1+kx} dx$    (Use the substitution trick for #7)

for 8, let u=1+kx then, du=kdx, and the integrand becomes (1/ku)du

10.    Given that $\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}$ , evaluate the Gaussian integral:

$$\frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{(2\sigma^2)}} dx$$    , where **a** is a constant.

Use substitution here: Let $u^2 = x^2/(2a^2)$    $2udu = 2xdx/(2a^2) \Rightarrow dx = (2a^2)^{1/2} du$

Then the integral becomes $(rc)^{1/2}/(rc)^{1/2} = 1$ !!

10.  A hormone L binds reversibly to a receptor R to form a complex C, with equilibrium constant K. Derive an equation for the concentration of C as a function of the ***total*** concentrations of ligand and receptor, $L_T$ and $R_T$, added to the system. (You will have to come up with and then solve a quadratic equation. The concept of an equilibrium constant is required for this problem.)

11.  The $pK_a$ of acetic acid is ~5. What is the pH of a solution of 0.1 M Na-acetate in water? Answer: pH = 9. Make reasonable approximations, and don't forget to take into account *all* ions in the system! This is an algebra problem that requires some background in general chemistry.

## 2. Basic elements of chemical thermodynamics (after taking CHEM 11 or 15)

The Biochemistry curriculum frequently uses common thermodynamic results that you have seen before - and have probably forgotten. These are relations that all students get to know and hate in Gen-Chem (and later in P-chem). The relations below are stated without explanation as to where they come from. (The best source for derivations that are both clear and rigorous is K. Denbigh, *The Principles of Chemical Equilibrium*.)

### A. Equilibrium Constant

Every chemical reaction has associated with it a numerical value called the equilibrium constant that is given in terms of the concentrations of reactants and products existing at equilibrium.

If, for instance, the reaction is described by: **aA + bB => c C**
**(where lower case letters represent stoichiometric coefficients), then**

$$K_{eq} = [C^*]^c / ([A^*]^a [B^*]^b),$$

where [A*], [B*], [C*] represent the concentrations of reagents *observed at equilibrium*. The constancy of this value at equilibrium is the basis for the empirical "laws" of mass action and LeChatelier's principle. It describes the magical tendency of a reaction to adjust its equilibrium composition so that this ratio of concentrations remains constant. (We are so familiar with the concept of an equilibrium constant that we tend to forget what an amazing thing it is: how does the reaction "know" to maintain this value constant as you dump reactants into the system?)

### B. Gibbs free energy.

For reactions that occur at constant temperature and pressure, as in nearly all biochemical applications, the condition of equilibrium may be expressed in terms of the Gibbs free energy, G, which is defined in terms of enthalpy, H, and entropy, S:

$$G = H - TS.$$

With this definition, a reaction is at equilibrium if and only if the free energy of the whole system is at a minimum. This can be stated mathematically in terms of an imagined change in the composition of the reaction near equilibrium:

$$\Delta G = 0$$

## C. Gibbs energy change when a reaction occurs.

If a reaction occurs, such as A => B, at some ambient concentrations [A], [B], then the free energy change peer mol of A converting to B associated with this reaction is:

$$\Delta G = \Delta G^o + RT \ln([B]/[A]).$$ This familiar equation is actually tricky to think about. **It's got 2 subtleties in it that you must understand *inside out*.**

**Subtlety #1**. We are talking about a reaction occurring, where a mole of A turns into a mole of B, *and yet the* <u>concentrations</u> *of A and B stay constant during the reaction*. How can such a thing be possible? The way to think about this is to imagine the reaction occurring in a gigantic volume, such as the Atlantic Ocean made up to 0.1M of A and 0.05 M of B. There are gazillions of moles of A and B in there, and if a mole of A reacts to form a mole of B, the *concentrations* at which this occurs remain essentially unchanged. Alternatively (a less weird and more "biochemical" way to look at it), think of a test tube of 1mL of A and B at some ambient concentrations, and then sprinkle in an enzyme so that a tiny bit of A reacts to form a tiny bit of B, say 1 nanomole. Since this amount of reaction is so small, the concentrations won't change much; there will be only a tiny amount of free energy change, too, because only a nanomole has reacted. However, you can still imagine the free energy change *per mole of A*, since this will be just $10^9$-fold bigger than the actual nanomole-scale $\Delta G$ of the reaction. This is obviously the situation we've got in the cell: ambient concentrations of metabolites maintained roughly constant, and far away from equilibrium, by cellular homeostasis mechanisms. Thus, we can speak sensibly about the "free energy of hydrolysis of ATP" as being, say, -11 kcal/mol under ambient cellular conditions, when [ATP]=5 mM, [ADP]=1mM, [$P_i$]=1mM, all maintained constant by the hardworking cell in a steady-state, even as the ATP is actually being hydrolyzed.

**Subtlety #2**. What do we mean by the bizarre symbol $\Delta G^o$? This is the *standard-state free energy* of the reaction, and its meaning generates endless confusion. For now, suffice it to say that it is an experimentally measurable quantity that allows you to actually *do* the calculation of free energy change above. For its *meaning*, wait until the next page. You can *measure* $\Delta G^o$ from the experimentally determined equilibrium constant - or you can look it up in a reference book (or Google it):

$$\Delta G^o = - RT \ln K_{eq}$$

In other words, $K_{eq}$ and $\Delta G^o$ are exactly equivalent things that express, in different "words," how far over to the "right" or the "left" a chemical reaction sits at equilibrium. If you tell me something about the $\Delta G^o$ of a reaction, you're not saying anything different, or being any cleverer, than if you tell me about the $K_{eq}$ of that reaction (but your friends will think that you're cleverer if you use $\Delta G^o$ language...)

## D. Decomposition of $\Delta G^o$ into $\Delta H^o$ and $\Delta S^o$.

$$d\Delta G^o/dT = - \Delta S^o \qquad -- OR -- \qquad d\ln K_{eq}/d(1/T) = - \Delta H^o/R$$

### E. Standard States

You hate, 'em, but you can't get away from 'em, so we have to confront them head-on. Whenever you want to measure a free energy change of a reaction, every time you want to interpret an equilibrium constant, you crash right into the standard-state free energy:

$$\Delta G_{actual} = \Delta G^o + RT \ln \Pi[products]/\Pi[reactants] \qquad \text{and} \qquad \Delta G^o = -RT\ln K_{eq}$$

### E1. **What's the meaning of $\Delta G^o$?**

<div align="center">Consider a reaction A => B + C</div>

The standard-state free energy change for this reaction, $\Delta G^o$, is the free energy change that would occur in an *imaginary reaction*, a reaction that *never actually takes place* but is very easy and useful to think about. The imaginary reaction is:

> **Start with pure A at 1M concentration and convert it *completely* to pure B and pure C, each at 1M concentrations.** (Ask yourself: why is this reaction imaginary? Why does such a thing never actually take place?) $\Delta G^o$ is the free energy change that *would* occur if the imaginary reaction actually happened.

This is a useful thing to think about because you can look at $\Delta G^o$ in a slightly different way than that stated above. It is the free energy difference between a "final state" - a bucket of B and C (each at 1M) - and an "initial state" - a bucket of A (at 1M). It's the free energy change resulting *just from interconverting these molecular species - the intrinsic difference in free energy just because they are different molecules*. If, for instance, we want to try to imagine what forces lead to the stability of a protein, we can focus on the imaginary process of starting with fully unfolded protein and ending up with it completely folded, asking about the molecular contributions to the free energy of this imaginary change - the intrinsic energy and entropy differences between these two forms of the protein. (Remember, to actually *measure* $\Delta G^o$ for folding, you measure the equilibrium constant of the folding reaction by measuring the fractions of folded and unfolded protein at equilibrium; these fractions are *not* 0 and 1, but somewhere in between.)

<span style="color:red">A rule never to forget in doing problems with equilibrium constants or $\Delta G$:</span>

Always use **MOLAR** for the concentration unit, even if the problem is stated in other units (mM, µM, etc). Here's a common mistake - a quiz asks: what's $\Delta G$ for ATP hydrolysis under cellular conditions, where [ATP]= 5 mM, [ADP]= [P_i] = 1 mM, and $\Delta G^o = -7$ kcal/mol? You answer:

$\Delta G = -7 + RT\ln(1)(1)/5 = -7 - 0.6 \ln(5) = -8.2$ kcal/mol    **WRONG!!**

$\Delta G = -7 + RT\ln(10^{-3})(10^{-3}) / (5\times10^{-3}) = -7 + 0.6 \ln(2\times10^{-4}) = $ **- 11 kcal/mol - CORRECT!**
**Don't forget this! Always use Molar units!**

## F.  The four faces of acid-base behavior

Acid-base behavior is just a special case of a small molecule ($H^+$) binding to a receptor (the buffer), but tradition has established a different 'language' to handle this. You must master this language. Consider the **dissociation** of a weak acid buffer:

$$BH => B^- + H^+$$

*1.*  **Equilibrium constant** defined: $K_a = (B^-)(H^+) / (BH)$ 'acid dissociation constant', M units.
  *Note: Don't worry about the electric charge on BH or $B^-$ ; that depends on the type of protonatable group . Two important ones in biochemistry are:*

**Carboxyl**   $RCOOH => R\text{-}COO^- + H^+$
  Amino       $R\text{-}NH_3^+ => R\text{-}NH_2 + H^+$
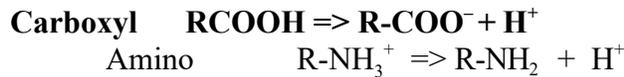
A **strong base** is a $B^-$ that wants to hold onto its proton very tightly, so the equilibrium above tends far to the left. Values of $K_a$ for all biological acids and bases can be found in reference books - or, of course, in Google and Wikipedia.

2.  **Ratio** of $BH/B^-$ increases with ($H^+$):   $(BH)/(B^-) = (H^+)/K_a$   (all in units of moles/L)
**Note: This means that when ($H^+$) = K , 50% of the buffer is in BH and 50% in $B^-$**

3. **Definition of pH and $pK_a$.** Long ago, a pair of meat-head professors - Henderson and Hasselbach - decided that students should suffer more than they already do. They thought that it would be fun to torture them by using the $\log_{10}$ **of ($H^+$)** as the metric for proton concentration, and worse - the *negative* log! (Don't ask why.) So they rewrote the ratio equation above as:

$$-\log_{10}(BH)/(B^-) = -\log(H^+) + \log(K) \text{ or: } \mathbf{\log(B^-)/(BH) = pH - pK_a}$$

  *Note: For some strange reason, textbooks rewarded these two bozos by assigning both their names to this simple equation, despite the fact that it says exactly the same thing as the simpler ratio relation above. Go figure - **but make sure that you can derive the 'Henderson-Hasselbach equation' in your sleep!***
Exercise: What is the ratio BH / $B^-$ for acetic acid ($pK_a$= 4.7) at pH 4 or at pH 6?

4. The **proton titration** curve. Actually, the Henderson-Hasselbach equation is pretty useless because we don't often care about the *ratio* of $B^-$ to BH. Typically, we'd rather have a picture of how the ***fraction (or percentage)*** of the buffer in the $B^-$ or BH form varies as you vary proton concentration. That's called a titration curve. Do the two exercises below:
*Using the ratio equation, derive the **fraction** of buffer in the $B^-$, BH forms as a function of ($H^+$):

$$f_{B-} = \frac{1}{1 + (H^+)/K_a} = \frac{1}{1 + 10^{pKa-pH}} \qquad f_{BH} = 1 - f_{B-}$$

*Plot $f_{B-}$ *vs* ($H^+$) and *vs* pH for acetic acid. See how different they look. Which is easier to interpret?

### 3. Expectations for using carbonyl chemistry (after taking CHEM 25)

**Elements of organic chemistry**
- Represent carbon-based molecules with appropriate bonding, charge, stereochemistry
- Identify functional groups: alcohol, carbonyl (ketone, aldehyde), carboxyl, ester, amide, amine, imine, thiol (sulfhydryl)
- Understand resonance and its relation to reactivity/stability
- Draw mechanisms using electron pushing arrows
- Be familiar with nucleophile and electrophile; substitution and elimination reactions

**Functional groups**: Biomolecules are carbon based. To understand biological structures and reactions, you must be familiar with the identity and reactivity of the following functional groups:

Name each functional group: (L—+ R: amine, aldehyde, ester; alcohol, thiol or sulfhydryl, ketone, thioester; enolate, carboxylate, amide)
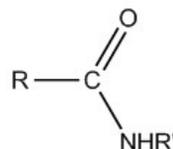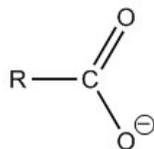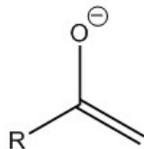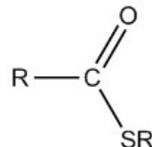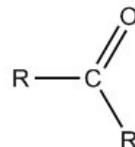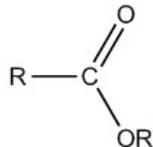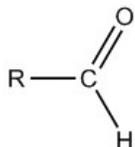
**Resonance**: To predict reactivity, we need to know where a molecule's electrons are….or are likely to be. For some molecules, we cannot use a single structure to represent its electronic configuration. Instead, we need to use resonance structures which have identical atom placement but different electron locations, as indicated by double bonds and lone pairs. These resonance structures are \*not\* in equilibrium with each other; to show this we use the double headed arrow, ↔ , rather than the equilibrium arrow, "'. The "true" structure is some hybrid of the possible resonance structures.
Draw all possible resonance structures for these molecules:
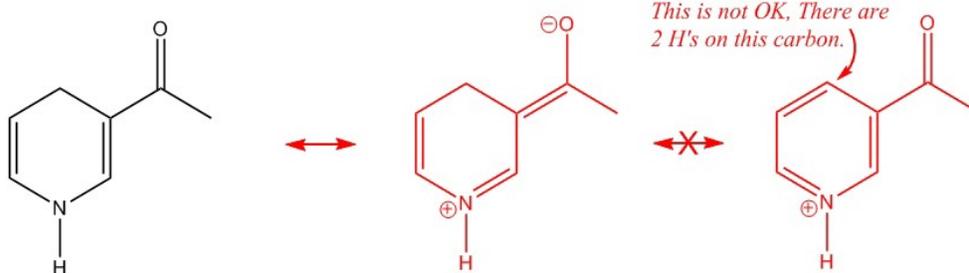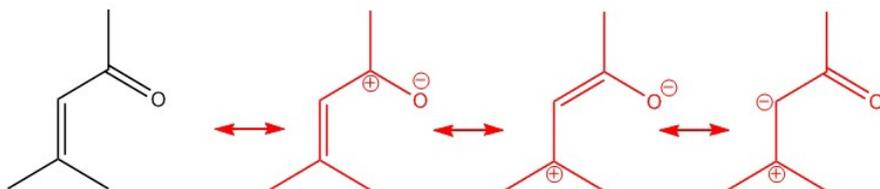
**Electron pushing**: Chemical reactions in biological systems obey the principles of organic chemistry. We use electron pushing arrows to describe reaction mechanisms as a way to show the changes in electron sharing that occur. You will need to remember a few key rules for electron pushing mechanisms:

1. Arrows must start at an electron-rich site (bond or lone pair) and point toward an area of electron-deficiency.
2. Obey the octet rule (or 2e- for H's)
3. Always show any formal charges and maintain conservation of charge.

You learned the keto-enol tautomerization reaction in Organic Chemistry; this reaction is used in cellular metabolism. Write the electron pushing arrows for this reaction under either acidic or basic conditions.
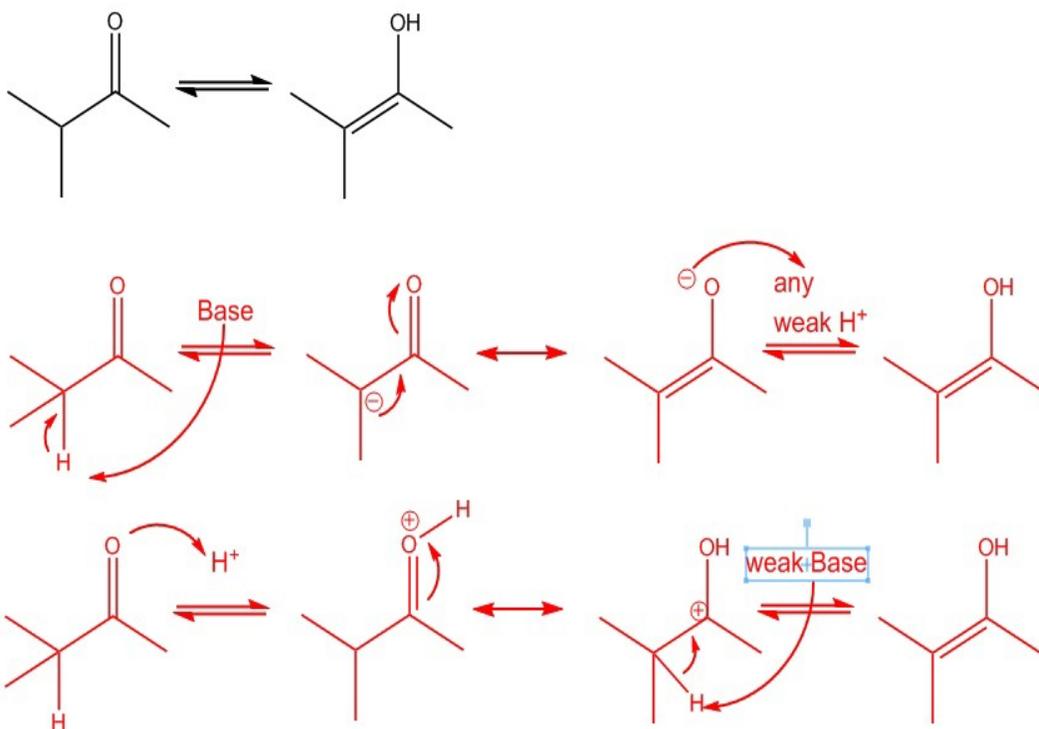
<p align="center">**Required core courses in the Biochemistry curriculum**</p>

### BCHM 100: Advanced Introductory Biochemistry

This course introduces protein structure and treats intermediary metabolism and how enzymes apply principles of organic chemistry to accelerate the rates of biochemically important chemical reactions. The class depends heavily on the rules of carbonyl chemistry, as encountered previously in your Organic Chemistry course. These principles are reviewed in the first few classes, as a "chemical toolbox" of a small number of organic reactions that apply to a very large number of biochemical reactions. In addition, familiarity with equilibrium constants, free energy, and concepts of pH will be assumed - and briefly reviewed in the first few classes.
Principles of chemical kinetics will be introduced, and the mathematics needed to apply these principles to enzyme catalysis will be taught and used extensively throughout the course.

### BCHM 103: Information Transfer in Biochemistry

This course presents a chemistry-based treatment of the central dogma of molecular biology: DNA => RNA => protein. Its subjects include the enzymes bringing about DNA replication, RNA transcription, and protein translation by the ribosome. Organic reaction mechanisms, especially phosphate chemistry, will be emphasized. Primary literature readings will be introduced to instruct students in the "art" of reading the biochemical literature.

### BCHM 104a: Physical Chemistry of Macromolecules, 1.

This course covers classical thermodynamics and some elements of statistical mechanics, leading to basic concepts of energy, entropy, free energy, and chemical equilibrium. The course is heavily mathematical and assumes proficiency in univariate calculus. Some multivariate calculus and probability & statistics, including probability density functions, gaussian integrals, and gamma functions, will also be used; knowledge of these more advanced subjects is not assumed and will be explicitly taught in the course.

### BCHM 104b: Physical Chemistry of Macromolecules, 2.

This second-semester P-chem course applies principles of chemical equilibrium emerging from BCHM 104a to macromolecules: proteins, nucleic acids, membrane assemblies. It relies heavily on proficiency in the application of equilibrium constants and Gibbs free energy to problems arising in behavior of macromolecules - molecular forces determining protein and nucleic acid structure, lipid bilayer formation, etc. The course uses calculus, much less than in BCHM 104a, and it extensively uses algebra.

### BCHM 101: Advanced Biochemistry

This is a course in mechanistic enzymology, the details by which enzymes apply the principles of organic reaction mechanisms to produce catalysis. It relies heavily on knowledge of carbonyl chemistry and on quantitative understanding of chemical kinetics. The use of differential equations to derive rates of chemical reactions will be reviewed at the beginning of the course, and will be used throughout.