

Instructions:

- Turn in your results in some kind of clear form (handwritten, typed, pdf).
- Feel free to work in groups.
- I have tried to be consistent in my language in my prompts if I want something specific.
 - *Sketch*: Hand draw a plot. Axes should be labeled conceptually, with key features indicated and explained.
 - *Generate a plot*: Plot using software. Label axes, use appropriate significant figures, etc.
 - *Calculate*: Actually calculate numbers using math. Report units and error bars, as appropriate.
 - *Describe/Argue*: Use any combination of writing, sketching, plotting, and calculation to argue for an interpretation.
 - Finally, I will sometimes specify the sort of explanation I want.
 - * *Molecular*: Describe what the atoms and molecules are doing in space and time. Depending on the context of the question, this might also involve explaining the result in terms of atomic properties like hydrophobicity.
 - * *Energetic*: Describe in energetic terms (entropy, free energy, statistics).
 - * *Mathematical*: Answer in terms of how the functions behave. For example, if I asked “mathematically, why does $Kx/(1 + Kx)$ saturate with increasing x ” the answer would be: “because as $x \rightarrow \infty$, $Kx \gg 1$ and the function tends to 1.”
- Some of the questions may require you to play with math and/or ideas we did not explicitly discuss in class. This is intentional. Many times in science you will be faced with a paper that uses an approach you are not familiar with. Learning how to gather enough information to critically evaluate their findings, as well as understand any mathematical models employed, is important.
- Some questions are listed as **GRAD STUDENT** questions. Undergrads are free to do those questions, but it is not required.

1 Protein structure and energetic reasoning

1.1 Warm up and review

- Consider a unimolecular reaction, $A \xrightleftharpoons{K_{eq}} B$, in which ΔG° favors the formation of B . You start the reaction with $[A] = A_0$ and $[B] = 0$.
 - Sketch a curve of $[B]$ vs. time (t) for the un-catalyzed reaction. Explain why the curve has this shape in terms of the behavior of the molecules in space and time.
 - As $t \rightarrow \infty$, what value does $[B]$ approach, and why?
 - You add an enzyme which catalyzes the reaction. Sketch a new curve for this reaction on the same graph as before. As $t \rightarrow \infty$, do the curves reach the same value? Why or why not?
- You are studying spontaneous unimolecular reactions at 25°C .
 - The reaction passes through a single transition state whose energy is $100\text{ kJ} \cdot \text{mol}^{-1}$ greater than the reactant. What is the fold speed up ($10\times$, $20\times$, etc.) if a catalyst is added that lowers the transition state energy by $25\text{ kJ} \cdot \text{mol}^{-1}$?
 - Now consider a reaction with a more complicated energy barrier. It passes through two transition states: the first is $100\text{ kJ} \cdot \text{mol}^{-1}$ higher in energy than the reactant, the second is $25\text{ kJ} \cdot \text{mol}^{-1}$ higher in energy than the reactant. The catalyst stabilizes the first transition state by $25\text{ kJ} \cdot \text{mol}^{-1}$ but has no effect on the second transition state. How much will this speed up the reaction?
 - Could you tell the difference between the reaction/enzymatic mechanisms on parts *a* and *b* from their reaction kinetics? Why or why not?
- Now consider a different unimolecular reaction, $C \xrightleftharpoons{K_{eq,C \rightarrow D}} D$, with $\Delta G_{C \rightarrow D}^\circ = 10\text{ kJ} \cdot \text{mol}$ (that is, it mildly *disfavors* the formation of D).
 - You start with $[C] = 10\text{ }\mu\text{M}$, $[D] = 0\text{ }\mu\text{M}$. As $t \rightarrow \infty$, what is the concentration of C and D in solution?
 - Would the addition of an enzyme that catalyzes this reaction change the outcome? If not, what would addition of the enzyme change about this process?
 - Now, imagine that you add a second enzyme that catalyzes the conversion of $D \xrightleftharpoons{K_{eq,D \rightarrow E}} E$, where $\Delta G_{D \rightarrow E}^\circ = -100\text{ kJ} \cdot \text{mol}^{-1}$ (it strongly favors formation of E).
 - As $t \rightarrow \infty$, what (qualitatively) happens to the concentrations of C and D ?
 - Calculate the $[D]_{t \rightarrow \infty}$.
 - Draw a free energy reaction coordinate for this pair of reactions. How does this reaction coordinate compare to the reaction coordinate for a single enzymatically catalyzed reaction?
- What is kT and what does it measure?
- The fastest enzymes end up being *diffusion limited*. What does this mean, and why can't diffusion limited enzymes go any faster? (Put another way, why doesn't it help to lower the activation energy after some point?).

1.2 Structure/energy relationships

- Conceptual goals:** Understand that the energetic properties of proteins can be estimated using a physical model and the 3D structure of the protein. Specifically, understand some of the considerations that go into electrostatics calculations.
- Skill goal:** Reason effectively about protein energetics. Manipulate quantitative descriptions of protein energetics. Reason about links between experiments and physical models.

Introduction

Coulomb energy

According to Coulomb's law, the maximum free energy that can be extracted by moving two ions from an infinite distance to distance r_{ij} is given by:

$$\Delta G_{coulomb} = \alpha \frac{q_i q_j}{\epsilon r_{ij}}.$$

- α is a constant: $\alpha \equiv 1389 \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{\AA}$. This takes care of things like the Coulomb constant, unit conversion to angstroms/kJ/ etc.
- q_i and q_j are the unit charges on each ion (things like -1, +1, or +11)
- ϵ is the dielectric constant, a unitless real number $\in [1, \infty)$ that scales how much the medium separating the charges attenuates the interaction. In biological systems, the relevant range is from 1 (vacuum) to water (78.5).
- r_{ij} is the distance between the ion centers in \AA .

Salt screening

In most biological contexts, there is salt in the solution that forms a mobile cloud which screens Coulomb interactions. This effect depends on distance (with greater attenuation over longer distances) and concentration (with greater attenuation with higher salt concentrations). One common way to model this is using the Debye-Hückel approximation:

$$\Delta G_{coulomb}([ions]) = \Delta G_{coulomb} e^{-r_{ij}/\kappa}$$

where κ is the *Debye Length*. This has units of \AA and measures how quickly interaction strength decays. A shorter κ means that the Coulomb interaction is felt at longer distances. At $r_{ij} = \kappa$, the Coulomb interaction is $\approx 37\%$ of what it would be without salt present. κ is given by:

$$\kappa \equiv \gamma \sqrt{\frac{\epsilon T}{I}}.$$

- γ is a constant: $\gamma \equiv 0.01988 \text{ M}^{1/2} \cdot \text{K}^{-1/2} \cdot \text{\AA}$. Like α above, this constant takes care of vacuum permittivity, unit conversion, and the like.
- ϵ is the dielectric constant.
- T is the temperature in K .
- I is the ionic strength in M (https://en.wikipedia.org/wiki/Ionic_strength).

Solvation energy

Finally, ions possess a “self energy”: the energy required to carve out a charged sphere in a given environment. If an ion maintains its charge and environment throughout a reaction, we can generally ignore this term. If, however, the magnitude of the charge and/or its environment changes, we must take this term into account. For example, if an aspartic acid deprotonates in the course of the reaction, we must account for the cost to create the charge in any energetic calculation. The Born solvation model describes this as the energy to bring the charge from being evenly distributed across the the surface of an a sphere with infinite radius to being distributed across a sphere with radius r_{ion} . It is given by:

$$\Delta G_{born} = \alpha \frac{q_{ion}^2}{\epsilon r_{ion}}.$$

This can be derived from Coulomb's law by breaking q_{ion} into infinitesimal pieces and integrating over the process of bringing these charge pieces together onto the surface of a sphere.

Questions

- Both $\Delta G_{coulomb}$ and ΔG_{born} use dielectric constants.
 - What does a dielectric constant measure?
 - Why, molecularly, is the dielectric constant of oil ($\epsilon_{oil} \approx 2$) lower than water ($\epsilon_{water} = 78.5$)?
 - If you lyophilize a water-soluble protein (e.g. dry it out almost completely), it has a dielectric constant of $\epsilon_p \approx 4$. What does this tell you about the chemical features of protein molecules?
 - Why might it be difficult to map this “bulk” dielectric constant back to individual protein molecules?
- Consider a protein in which the carboxylic acid of an aspartate sidechain forms an ion pair with the amine of a lysine sidechain. The interaction is favorable by $-15 \text{ kJ} \cdot \text{mol}^{-1}$. If the pK_a of the aspartate carboxylic acid in water is 4.0, what is its pK_a value in this protein?
- Baran et al. set out to measure the effective dielectric constant for Coulomb interactions between charged amino acids on the surface of a protein (Baran et al. *JMB* 379(5):1045-1062). To do so, they measured changes in the pK_a values of histidine residues when they mutated acidic residues throughout the protein. They then measured the distance between the acidic residue and histidine in the crystal structure of the protein. The experiments were done at 298 K using $I = 0.1 \text{ M}$.
 - Their results are a .csv file that should have come with this homework. Use this information to determine the apparent dielectric constant for these Coulomb interactions.
 - Does this result match what you expect? How might you explain this result, mechanistically?
- Proteins often bury ion pairs in active sites. We are going to explore whether this is favorable or not.
 - Derive an equation that expresses the electrostatic component of the energy to transfer an ion pair from water to the protein interior ($\Delta G_{transfer}^\circ$). This will be a function of the system dielectric constants (ϵ_{water} and $\epsilon_{protein}$), the charge on each ion (q_i and q_j), the radius of each ion (r_i and r_j) and the distance between the centers of mass for each ion (r_{ij}). You may ignore the effects of salt in this analysis.
 - Create three plots, all as a function of the protein dielectric constant ϵ_p : ΔG_{born}° , $\Delta G_{coulomb}^\circ$, and $\Delta G_{transfer}^\circ$. You may assume that $\epsilon_{water} = 78.5$, $q_i = 1$, $q_j = -1$, $r_i = r_j = 2 \text{ \AA}$, and that $r_{ij} = 3 \text{ \AA}$.
 - Does the Coulomb interaction *ever* fully offset the desolvation penalty? Whatever your answer: can you justify this result molecularly?
- Molecular dynamics simulations are powerful tools for understanding protein function, but they can also be black boxes. In the following, you will develop a very simple forcefield and use it to calculate the dynamics of two particles. Consider two ions in water at 25 °C with no other ions present. They have the same charge (q) and their centers of mass are separated by r . They begin with zero relative velocity. Write expressions for:
 - ...the *work* required to bring the two ions from infinite distance to a separation r_{ij} .
 - ...the instantaneous *force* experienced by each ion when at r_{ij} (remember, $E = \int F dx$)
 - ...the instantaneous *acceleration* of each ion.
 - ...the *velocity* of each ion after some short time dt .
 - ...the *position* of each ion after dt .
 - From these expressions, what information do you need to calculate their position and velocity?
 - Assume these are sodium atoms separated by 4 Å. Find the information you discovered you needed in (f) and calculate their positions and velocities after 10^{-15} s .
 - From this analysis, what are basic the “ingredients” to any molecular dynamics simulation?

6. **GRAD STUDENTS:** Write a program that can read a PDB file and calculate the sum of the Coulomb interactions between charged residues in the protein, accounting for the effect of salt concentration using the Debye-Hückel approximation. You may assume that all Lys, Arg, His, Asp, and Glu are charged. Place the charge on the following atoms:

amino acid	atom	charge
ASP	CG	-1
GLU	CD	-1
ARG	CZ	+1
HIS	NE2	+1
LYS	NZ	+1

Calculate $\Delta G_{coulomb}$ as a function of ε for the pdb file 1STN at $I = 0.1\text{ M}$, $T = 298\text{ K}$.

- (a) How much does total energy change between $\varepsilon = 1$ and $\varepsilon = 78.5$?
- (b) Which dielectric constant is best (and why do you think so)?