Calculating protein energetics

Coulomb energy

The energy change that accompanies moving two ions from an infinite distance to distance r_{ij} is given by Coulomb's law:

$$U_{coulomb} = \alpha \frac{q_i q_j}{\varepsilon r_{ij}}.$$

- α is a constant: $\alpha \equiv 1389\ kJ \cdot mol^{-1} \cdot$ Å. This takes care of things like the Coulomb constant, unit conversion to Å and $kJ \cdot mol^{-1}$, etc.

- q_i and q_j are the partial charges on each ion (things like +1.0, +0.5, -0.2, +11, etc.)

- 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 Å.

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:

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

where κ is the *Debye Length*. This has units of Å 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:

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

- γ is a constant: $\gamma \equiv 0.01988\ M^{1/2} \cdot K^{-1/2} \cdot$ Å. Like α above, this constant takes care of vacuum permitivity, 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

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:

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

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.

Hydrogen bonding

Hydrogen bonds are challenging to (efficiently) describe. There are three basic approaches.

1. Calculate the energy using a quantum mechanics. This is the most accurate approach, but is computationally intensive. For biomolecules, quantum mechanical calculations are generally only used specific questions about enzyme active states and the like, often coupled to a classical simulation.

2. Use an empirical potential that accounts for both distance an angular information. One such function is:

$$U_{hbond} = 4.814 \times a \left[\frac{q_A q_H}{r_{HA}^2} cos(\theta) + bc^r \right]$$

   where a, b and c are adjustable parameters for eight different hydrogen bond combinations (see table below). q_A and q_H are partial charges on the acceptor and donor. r is the distance between the hydrogen and the hydrogen-bond acceptor in Å. θ is the angle between $donor - hyrogen - acceptor$ in degrees. This potential function is taken from Rezac et al. (2009) *JCTC* doi.org/10.1021/ct9000922.

   The general protocol for such functions is to calculate the angular and distance dependence for different classes of hydrogen bonds using quanum mechanics, and then fit terms like a, b, and c to reproduce those values. The sorts of functions used have been reveiwed nicely by Martin Korth (2011) *ChemPhysChem* doi.org/10.1002/cphc.201100540. These potentials are a bit hacked together: the general idea is to make them work

rather than to capture deep theory. The sorts of approximations made can be seen, for example, in O'Meara et al. (2015) *JCTC* doi.org/10.1021/ct500864r, who describe how they came up with a new potential function.

bond type	a	b	c
N (with no hydrogens) interacting with any H	14.4209	-1.3273×10^{-2}	7.2874
N (with one hydrogen) interacting with any H	73.3566	-5.3979×10^{-4}	7.0920
N (with two or three hydrogens) interacting with any H	48.7161	2.9844×10^{-4}	6.4259
O (except carbonyl) interacting with HN	29.8036	2.1262×10^{-3}	6.9768
O (carbonyl) interacting with HN	-6.4578	7.3142×10^{-3}	7.8379
O interacting with HO that is not water or carboxyl	23.1582	-4.8015×10^{-5}	6.9382
O interacting with H in water	15.3029	2.0789×10^{-3}	7.0365
O interacting with H in carboxyl	14.8668	-4.6652×10^{-3}	6.9111

3. The simplest approach is to treat them Coulomb interactions between partial charges. This approach is straightforward, but does not properly account for the angular dependence in the strength of a hydrogen bond. Because of its speed, however, many standard molecular dynamics force-fields use this approach.

van der Waal's interactions

Van der Waal's interactions are usually described as:

$$U_{vdw} = (\epsilon_i \epsilon_j)^{1/2} \left[\left(\frac{\sigma_i + \sigma_j}{2r_{ij}} \right)^{12} - 2 \left(\frac{\sigma_i + \sigma_j}{2r_{ij}} \right)^{6} \right]$$

where σ_x is the van der Waal's radius of an atom and ϵ_x measures the strength of the contribution of atom x to the interaction. The values of σ_x and ϵ_x depend on the types of atoms involved. r_{ij} is the distance between the centers of the interacting atoms.

Hydrophobic effect

Using a solvent transfer energy studies, the following empirical relationships have been established that link transfer free energy to extent of burial of surface area. Key papers along the way are:

- McAuliffe (1966) *J Phys Chem* doi.org/10.1021/j100876a049

- Baldwin (1986) *PNAS* doi.org/10.1073/pnas.83.21.8069

- Murphy and Gill (1991) *JMB* doi.org/10.1016/0022-2836(91)90506-2

- D'Aquino et al. (1996) *Proteins* doi.org/10.1002/(SICI)1097-0134(199606)25:2<143::AID-PROT1>3.0.CO;2-J

- Auton and Bolen (2005) *PNAS* doi.org/10.1073/pnas.0507053102

In the following, ΔASA is calculated in Å^2, energies are given in $J \cdot mol^{-1}$, and temperatures are in K. The thermodynamics of transferring nonpolar surface area from fully water inaccessible to fully water accessible is given by:

$$\Delta C_{p,nonpolar} = 1.883 \times \Delta ASA_{nonpolar}$$

$$\Delta H_{nonpolar} = -35.313 \times \Delta ASA_{nonpolar} + \Delta C_{p,nonpolar} \left(T - 333.15 \right)$$

$$\Delta S_{nonpolar,Tref} = \Delta C_{p,nonpolar} \times ln \left(\frac{T}{385.15} \right)$$

The thermodynamics of transferring polar surface area from fully water inaccessible to fully water accessible is given by:

$$\Delta C_{p,polar} = -1.088 \times \Delta ASA_{polar}$$

$$\Delta H_{polar} = 131.378 \times \Delta ASA_{polar} + \Delta C_{p,polar} \left(T - 333.15 \right)$$

$$\Delta S_{polar,Tref} = \Delta C_{p,polar} \times ln \left(\frac{T}{335.15} \right)$$