4.1.2 Statistical estimation of thermodynamic properties

- This module will be a mini-lecture describing statistical estimators for thermodynamic properties
- At the end of this module, you should be able to calculate
  - expectation values
  - free energy differences between thermodynamic states

- TODO: More detail on BAR/MBAR?

Thermodynamic properties

- Molecular simulations are used to calculate thermodynamic and kinetic properties
- In general, the thermodynamic properties are
  - expectation values of an observable, including
    - probability of the observable having a certain range of values
    - potential of mean force with respect to the observable
  - free energy differences between thermodynamic states
   - in biomolecular systems, $\Delta G \sim \Delta A$

What is ΔG?

- ΔU is the change in average internal energy.
  - internal energy can be computed for individual structures
  - in biomolecular simulations, internal energy is modeled by the molecular mechanics force field
- $\Delta H = \Delta U + \Delta(pV)$ is the change in enthalpy
  - in biomolecular simulations, change in pV is usually negligible
- ΔG is the Gibbs free energy
  - at constant pressure and temperature, dictates
    - spontaneity and
    - equilibrium constant of process
  - in biomolecular simulation, interest in free energy differences between
    - conformations of a macromolecule
    - thermodynamic states with different temperature, pressure, volume, or other parameters
  - $\Delta G = \Delta H + T\Delta S$, but ΔS is very challenging to compute

What is ΔA?

- ΔA is the Helmholtz free energy
  - at constant volume and temperature, dictates
    - spontaneity and
    - equilibrium constant of process
  - in biomolecular simulation, ΔA and ΔG are usually assumed to be equal

How are free energy calculations useful?

- Noncovalent binding between molecules (see [1])
  - Design molecules to manipulate protein function
  - Recognize toxins
  - Identify enzyme functions
  - Protein design: design binders to target molecule
  - Aid medicinal chemistry, guide synthesis
- Hydration free energies
  - Part of binding free energy & solubility
- Conformational free energies relevant to
  - biological mechanism
  - binding free energy

$$P + L \xrightarrow{\triangle G^0} PL$$
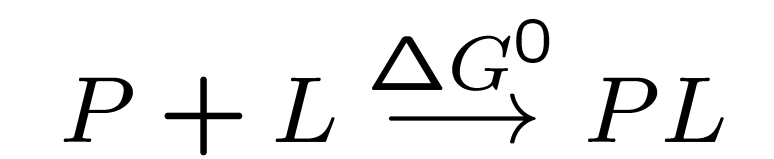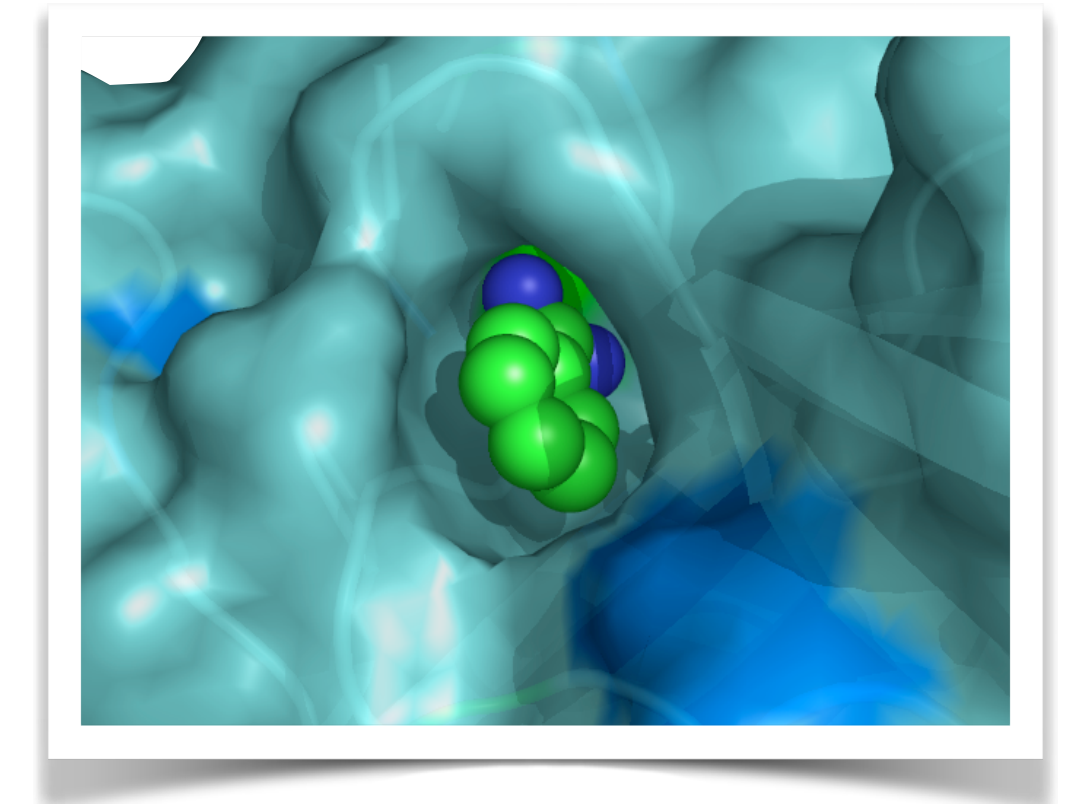
(from [1])

How are ΔG/ΔA calculated from molecular simulations?

Basic Statistical Mechanics

- In the Boltzmann distribution, the probability of a configuration r^N with energy $U_s(r^N)$ is,

$$\pi_s(r^N) \propto \exp\left[-\beta U_s(r^N)\right] \text{ (unnormalized)}$$

$$\rho_s(r^N) = \exp\left[-\beta U_s(r^N)\right]/Q_s \text{ (normalized)}$$

- A partition function is the normalizing constant of the Boltzmann distribution

$$Q_s = \int \pi_s(r^N)dr^N$$

- The free energy difference is related to a ratio of partition functions

$$\beta(A_1 - A_0) = -\ln\left(\frac{Q_0}{Q_1}\right)$$

The Zwanzig Relation: Derivation

- From before, $\beta(A_1 - A_0) = -\ln\left(\dfrac{Q_0}{Q_1}\right)$.

- Substituting in partition functions, $\beta(A_1 - A_0) = -\ln\left(\dfrac{\int e^{-\beta U_1(r^N)} dr^N}{\int e^{-\beta U_0(r^N)} dr^N}\right)$.

- Multiplying by one, $\beta(A_1 - A_0) = -\ln\left(\dfrac{\int e^{-\beta U_1(r^N) + \beta U_0(r^N) - \beta U_0(r^N)} dr^N}{\int e^{-\beta U_0(r^N)} dr^N}\right)$.

- Defining the potential energy difference $\Delta U(r^N) = U_1(r^N) - U_0(r^N)$,

$$\beta(A_1 - A_0) = -\ln\left(\dfrac{\int e^{-\beta \Delta U(r^N)} e^{-\beta U_0(r^N)} dr^N}{\int e^{-\beta U_0(r^N)} dr^N}\right)$$

The Zwanzig Relation: In Practice

- Using the definition of $\rho_s(r^N)$, $\beta(A_1 - A_0) = -\ln \int \rho_0(r^N) e^{-\beta \Delta U(r^N)} dr^N$.

- The Zwanzig relation [2] is
  - $\beta(A_1 - A_0) = -\ln \left\langle e^{-\beta \Delta U} \right\rangle_0$ in a simpler notation.
  - $\beta(A_1 - A_0) = -\ln \left\langle e^{\beta \Delta U} \right\rangle_1$ can be derived with similar steps

- This shows us that
  - The free energy difference can be computed based on an average over configurations taken from one of the states of interest
  - We can generate these configurations with MC or MD
  - The free energy comes from evaluating the energies of these configurations in both potentials U_0 and U_1, and taking an appropriate average of the energy difference

The Zwanzig Relation: Limitations

- In terms of an integral over the distribution of ΔU (instead of over $\rho_o(r^N)$) the Zwanzig relation is,

$$\beta(A_1 - A_0) = -\ln \int e^{-\beta \Delta U} \rho_0(\Delta U) d\Delta U.$$

- Sampling is from the red curve
- Accurate estimation requires the purple curve
- The calculation will not be accurate if U_0 and U_1 are very different!
- Potential energies will be different if the states access different parts of configuration space

(from [1])

10
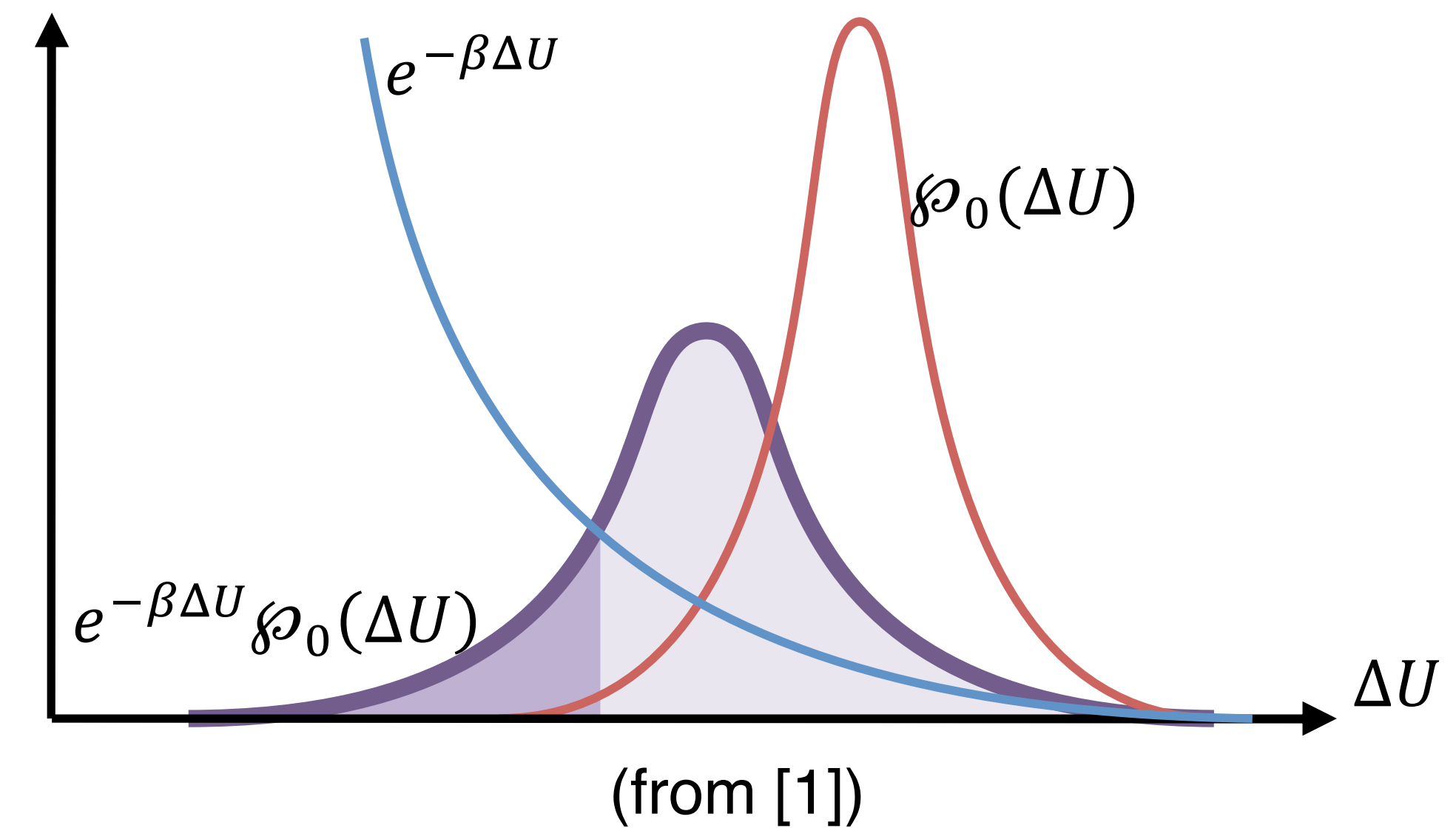
Other ways to calculate ΔG

- Multistate Bennett Acceptance Ratio (MBAR) [4]
  - estimates free energies and thermodynamic expectations from a series of states
  - extension of Bennett Acceptance Ratio (BAR) [3], which uses data from two states
  - Proven to be statistically optimal
- Thermodynamic integration is based on the fundamental theorem of calculus, integrating one the derivative of the free energy with respect to a parameter
- All of the methods require thermodynamic states with configuration space overlap, meaning that
  - similar configurations have similar energies
  - the most relevant configuration space is similar

References

- [1] Many parts of this module were adapted from a lecture by David Mobley (https://github.com/MobleyLab/drug-computing/tree/master/uci-pharmsci/lectures/free_energy_basics) under the CC BY 4.0 license. The lecture is part of the Drug Discovery Computing Techniques course (PharmSci 175/275) at UC Irvine.
- [2] Zwanzig, R. High-Temperature Equation of State by a Perturbation Method. I. Nonpolar Gases. Journal of Chemical Physics 1954, 22 (8), 1420.
- [3] Bennett, C. H. Efficient Estimation of Free-Energy Differences from Monte Carlo Data. Journal of Computational Physics 1976, 22 (2), 245–268.
- [4] Shirts, M. R.; Chodera, J. D. Statistically Optimal Analysis of Samples from Multiple Equilibrium States. Journal of Chemical Physics 2008, 129 (12), 124105.