3/24/2020 Week 10 Module 1
Free Energy Calculations

- This module will consist of
  - a mini-lecture describing free energies, various applications of free energy calculations, and how they are calculated based on molecular simulations
  - an exercise on setting up an alchemical binding free energy calculation with YANK
- At the end of this module, you should be able to answer the following questions:
  - What are free energies?
  - How are free energy calculations useful?
  - How can you calculate a free energy difference from a molecular simulation?
- You should also be able to run a binding free energy calculation with YANK

What are free energies?

What is ΔG?

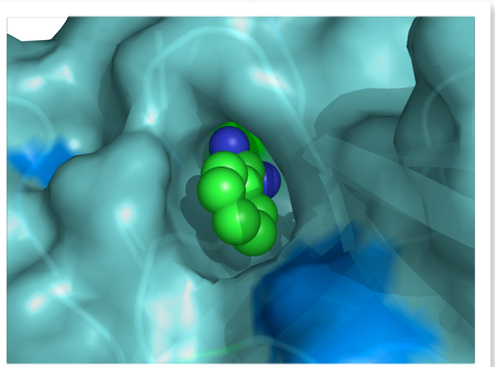
- ΔU is the change in average internal energy.
  - internal energy can be computed for individual structures
  - in biomolecular simulations, internal energy is modeled by the molecular mechanics force field
- ΔH = ΔU + Δ(pV) is the change in enthalpy
  - in biomolecular simulations, change in pV is 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
  - ΔG = ΔH + TΔS, 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



$$P + L \xrightarrow{\Delta G^0} PL$$

(from [1])

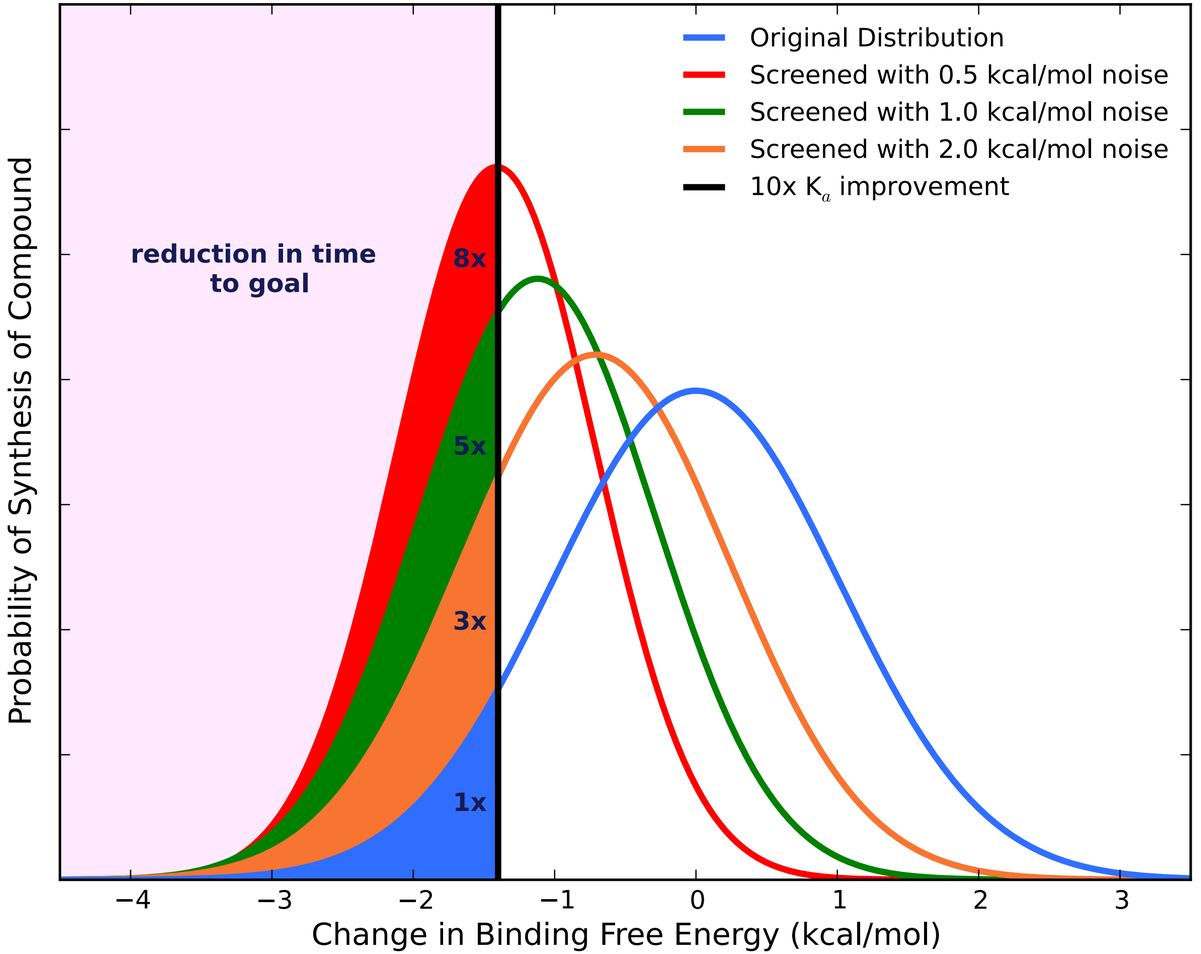
Even modest accuracies in calculated binding free energies can have significant benefits

Hypothetical pipeline:

- Medicinal chemist suggests 100 derivatives or compounds per week
- Your job is to pick the top 10 to carry forward

Question: How many molecules do we have to make to gain a factor of 10 in affinity?

- 0.5 kcal/mol noise: Decreases # required by 8x
- 1.0 kcal/mol noise: Decreases by 5x
- 2 kcal/mol noise: Decreases by 3x

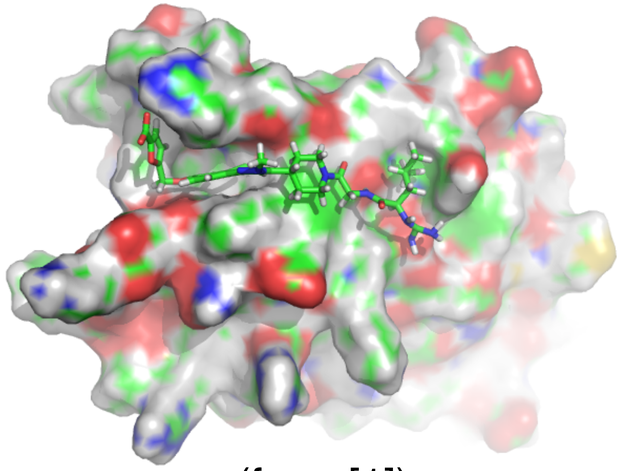


Shirts, Mobley, and Brown, “Free Energy Calculations in Structure-based Drug Design”, in “Structure Based Drug Design” by Merz et al., Cambridge University Press, 2010. See also Mobley & Klimovich, JCP 2013

Docking approximates binding ΔG

- $\Delta G = \Delta H + T\Delta S$

- Docking score ~ ΔH
- It sometimes involves
  - ad hoc ΔS based on the number of rotatable bonds
  - ΔG_{solv}
- Docking is based on “optimal” orientations



(from [1])

How is 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,

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

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

- 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, $\beta(A_1 - A_0) = -\ln\left(\frac{Q_0}{Q_1}\right)$.
- Substituting in partition functions, $\beta(A_1 - A_0) = -\ln\left(\frac{\int e^{-\beta U_1(r^N)} dr^N}{\int e^{-\beta U_0(r^N)} dr^N}\right)$.
- Multiplying by one, $\beta(A_1 - A_0) = -\ln\left(\frac{\int e^{-\beta U_1(r^N)+\beta U_0(r^N)-\beta U_0(r^N)} dr^N}{\int e^{-\beta U_0(r^N)} dr^N}\right)$.
- Defining the potential energy difference $\Delta U(r^N) = U_1(r^N) - U_0(r^N)$,

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

The Zwanzig Relation: In Practice

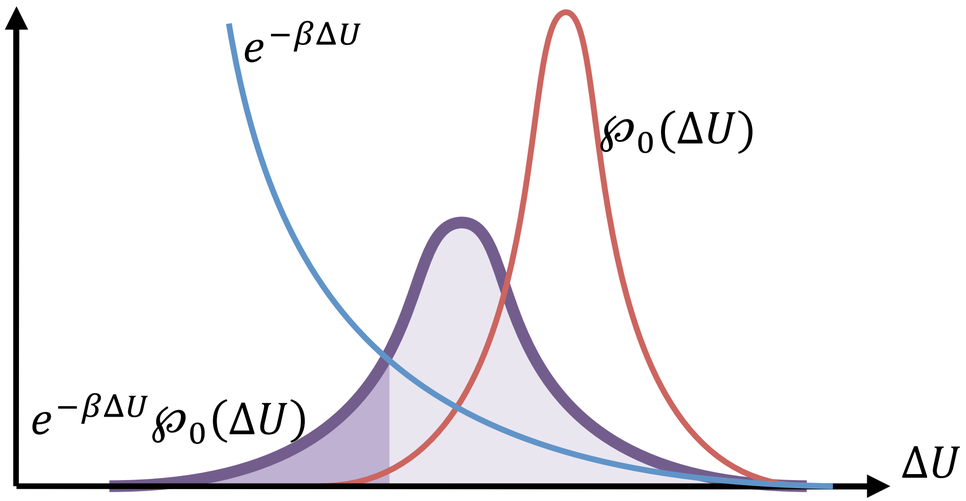
- Using the definition of $\rho_s(r^N)$, $\beta(A_1 - A_0) = -\ln \int \rho_0(r^N) e^{-\beta \Delta U(r^N)} dr^N$.
- The Zwanzig relation [2] is
  - $\beta(A_1 - A_0) = -\ln \left\langle e^{-\beta \Delta U} \right\rangle_0$ in a simpler notation.
  - $\beta(A_1 - A_0) = -\ln \left\langle e^{\beta \Delta U} \right\rangle_1$ 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!



(from [1])

Other ways to calculate ΔG

- The Bennett Acceptance Ratio (BAR) [3] uses data from two states
- The Multistate Bennett Acceptance Ratio (MBAR) [4] uses data from a series of states
- BAR/MBAR are 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

Alchemical Pathways

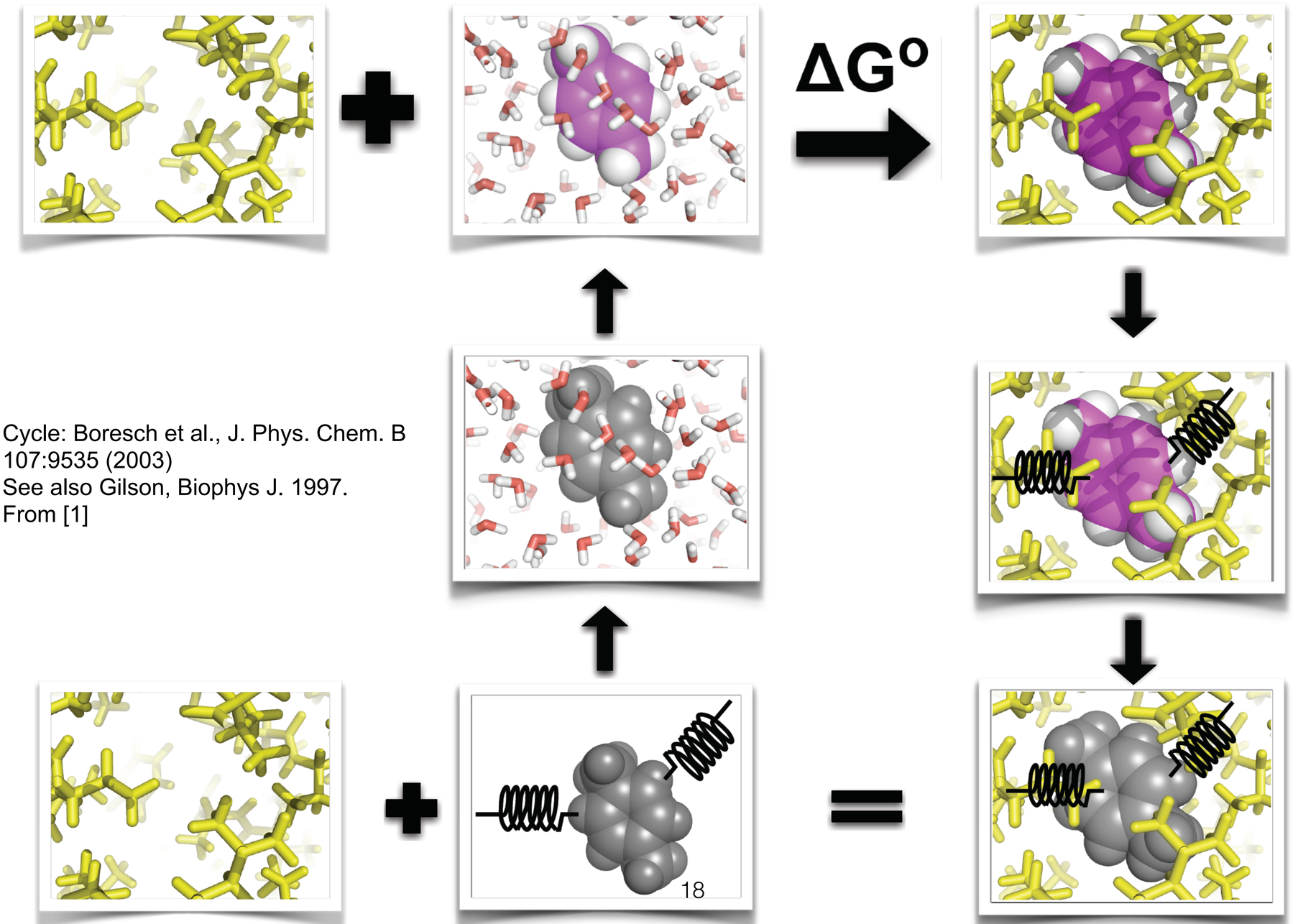
What are alchemical pathways?

- Alchemical pathways are a series of thermodynamic states where intermediate states do not necessarily model a physical system. For example,
  - a drug lead can be morphed into a similar proposed compound
  - harmonic restraints can be added to keep atoms in a certain position
  - states whose energy is a linear interpolation between states 0 and 1 can be defined as, $U_\lambda(r^N) = (1-\lambda)U_0(r^N) + \lambda U_1(r^N)$

Why do we use alchemical pathways?

- Adjacent states along an alchemical pathways have high *configuration space overlap*
- Using alchemical pathways is valid because thermodynamic functions like the Gibbs free energy are *state functions*
  - they only depend on the final values, not the path between them
    - height, weight, coordinates are other state functions
  - Hess' law is based on this property of the Gibbs free energy
- Binding free energy calculations usually involve connecting *alchemical pathways* in a *thermodynamic cycle* that joins the end states of interest

A Thermodynamic Cycle for Absolute ΔG of Binding



Cycle: Boresch et al., J. Phys. Chem. B 107:9535 (2003)
See also Gilson, Biophys J. 1997.
From [1]

References

- [1] Many parts of today's lecture 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.

Additional Resources

- Resource for alchemical binding free energy calculations (http://www.alchemistry.org/wiki/Main_Page)
- Thermodynamic cycle in YANK (http://getyank.org/latest/theory.html)