Free energy calculations

Lauren Wickstrom

Outline

- Introduction to Free Energy
- Ensembles and Ensemble averages
- Methods
- Applications
- Major Pitfalls

What is Free Energy?

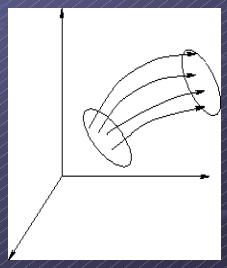
- Tells how favorable a chemical and structural change is
- Expressed as either Helmholtz free energy function A (constant NVT) or Gibbs free energy G (constant NPT)
- State function: only depends on initial and final positions

Applications of Free Energy Calculations

- Solvation of small molecules
- Molecular Association
- Ligand binding and catalysis
- Sequence dependent stabilities
- Environmental effects on reactions in solutions and in enzymes

Phase space

- N –particles (or atoms) defined by momentum (or velocity) and position
- Each point in phase space
- 3 N coordinates (x,y,z)
- 3 N momenta (p(x),p(y),p(z))



http://www.plmsc.psu.edu/~www/matsc597c-1997/introduction/intro/node5.html

Movement in Phase Space

- Governed by Hamiltonian equation

$H = p^2/2m$ (Kinetic E.) + V (Potential E.)

- System movement

(dr/dt) = (dH/dp) *(dp/dt)=(dH/dp)*

- Outcome: trajectory of points in phase space

Ensemble

- Collection of points in phase space
- Molecular dynamics – points connected in time – NVE ensemble (microcanonical)
- Monte Carlo – NVT ensemble (canonical) – no momentum components

Ensemble average

- Weighted sum over all members of the ensemble

$\langle A \rangle = \iint dp^n dr^n \, A(dp^n dr^n) \, (\exp(-E(dp^n dr^n)/kT)/Q)$

Partition Function – Q –sum of all Boltzmann weights of the all the energy levels of a system

If a system is ergodic, the ensemble average should equal the time average!!!

Free Energy Difference

$\Delta G = G_b - G_a = - RT \ln \langle e^{(-\Delta H/RT)} \rangle$

G_b = Free energy in system B
G_a = Free energy in system A
R = Gas constant
T = Temperature (K)

- Represented as an ensemble average

However, if A and B differ then other approaches must be implemented

Free Energy Perturbation

- Considers initial, intermediate and final state using the coupling parameter, λ(varying from 0 to 1)

$H^*(\lambda) = \lambda H_b + (1-\lambda)H_a$

$\Delta G = \sum - RT \ln \langle e^{(-\Delta H^*/RT)} \rangle$

$\Delta H^* = H(i+1) - H(i)$

Different methods of FEP sampling

Forward sampling : uses free energy differences from λ(i) →λ(i+1)

Backward sampling: uses free energy differences from λ(i) → λ(i-1)

Double wide sampling: uses free energy differences from both λ(i) →λ(i+1) and λ(i) → λ(i-1)

Double wide sampling is more efficient because more sampling points are obtained.

Thermodynamic integration

$\Delta G = \int(0,1) \langle \partial H/\partial\lambda \rangle$ where H_a is $\lambda = 0$ and H_b is $\lambda = 1$

- Several simulations run at discrete values of λ then ΔG is evaluated
- Most computer programs evaluate ΔH/ Δ λ instead of the partial derivative

The Slow Growth Method

- Hamiltonian is changed a very small, constant amount

From $\lambda = 0$ to 1

$\Delta G = \sum (H_{n+1} - H_n)$

Taken over an n amount of steps

H_{n+1} = next larger λ after 1 step

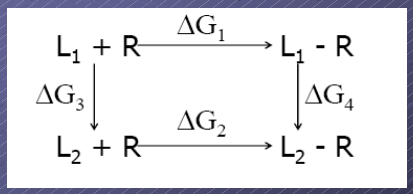
H_n = current λ

ΔG changes very slowly in this process

Simplification to the Hamiltonian

- In most situations, it is assumed the kinetic energy term cancels out.
- Potential energy used in calculation.
- Force field equation directly solves for this.

Applications – Thermodynamic cycle for binding of ligands



$\Delta G = -RT \ln K$ – Equilibrium between molecules interacting via noncovalent forces

ΔG_1 = Binding of ligand 1 to a Receptor

ΔG_2 = Binding of ligand 2 to a Receptor

ΔG_3 = Free energy difference between the two ligands in solution

ΔG_4 = Free energy difference between the two intermolecular complexes

What are the consequences of the thermodynamic cycle?

- Since ΔG = state function, this cycle must equal **zero**
- Therefore, $\Delta G_2 - \Delta G_1 = \Delta G_4 - \Delta G_3$
- Allows us to calculate values not obtainable in the laboratory - ΔG_3 & ΔG_4
- In silico, mutations to the ligand in solution and the ligand-receptor complex are possible.
- This is known as the free energy perturbation approach.

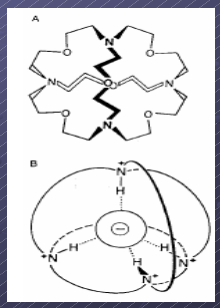
Examples - Host Guest Complexes

- Macrocycle SC24 binds halides when protonated
- SC24 binds Cl^- 4.30 kcal/mol stronger than Br^-

Two simulations were run using free energy perturbation method

1) Mutation of Cl^- to Br^- in solution = 3.35 kcal/mol
2) Mutation within the macrocycle = 7.50 kcal/mol

Overall, relative free energy of binding = 4.15 kcal/mol



Lybrand, T.P., J.A. McCammon and G. Wipff, *Theoretical calculation of relative binding affinity in host-guest systems.* Pro. Natl. Acad. Sci. USA, 1986. 83:833-35.

Accuracy of Free Energy vs Entropy/Enthalpy Calculations

- Free energy errors – 1 kcal/mol
- Entropy/Enthalpy – large errors

Why?

- ΔH involves difference of two large numbers
- Dominated by solvent-solvent interactions
- ΔG – determination of ensemble averaged difference between solvent-solute interactions

Major Pitfalls of Free Energy Calculations

- Hamiltonian

a) Bad potential energy model
b) Inaccurate force field

- Insufficient sampling of phase space

Errors can be estimated

a) Using different free energy methods
b) Perturbation can be run in forward and reverse directions

Conclusions

- Free energy calculations are applicable to many different systems.
- There are three different methods of calculating free energy.
- Accuracy of free energy calculations is far greater than enthalpy and entropy calculations.
- Accuracy of these calculations are dependent on convergence of the simulation and the force field.