Free energy calculations

Jochen Hub & David van der Spoel



Overview

- Free energies and Probabilities
- Thermodynamic cycles
- (Free energy perturbation (FEP))
- Thermodynamic integration (TI)
- (Jarzynski equality
- Crooks fluctuation theorem)
- Potential of mean force (PMF)

Free Energies and Probabilities

Free energies are related to probabilities of states

Partition function $\mathcal{Z}(N,V,T) = \int d^{3N}x \, d^{3N}p \, e^{-\beta E(\mathbf{x},\mathbf{p})}$ Probability to be in state s $P_s = \frac{1}{Z} e^{-\beta E_s}$

Normalization with $Z \rightarrow$ probability to be in any state = one

• Helmholtz free energy

$$A(N, V, T) := U - TS = -k_B T \ln \mathcal{Z}(N, V, T)$$

Thermodynamic Cycles

• Which ligand binds better? What you often want to know: $\Delta\Delta G = \Delta G_1 - \Delta G_2$

$$A + B \xrightarrow{\Delta G_1} AB$$
$$\Delta G_3 \parallel \qquad \qquad \parallel \Delta G_4$$
$$A' + B \xrightarrow{\Delta G_2} A'B$$

- Instead of measuring ΔG_1 and ΔG_2 , it is often easier to computationally "measure" ΔG_3 and ΔG_4
- Alchemical transformation in bulk and in binding site
- Other important example: How does a mutation destabilize a folded compared to the unfolded protein?

Thermodynamic Integration

- Instead of FEP, better do a TI !
- Idea: Parameter λ smoothly switches from state A to B

$$H(\lambda) = (1 - \lambda)H_A + \lambda H_B$$
$$G(A \to B) = \int_0^1 d\lambda \left\langle -\frac{\partial H(\lambda)}{\partial \lambda} \right\rangle_{\lambda}$$

- Switch slowly from $\lambda=0$ to $\lambda=1$ and integrate the dgdl.xvg output of mdrun. Problem: non-equilibrium/hysteresis effects
- Choose 20 to 50 discrete λ-steps (init-lambda), sample each λ-step at constant lambda (delta-lambda=0), integrate the averages (discrete TI)

Example TI: Solvation of Ethanol



Example TI: Solvation of Ethanol



Constrained free energies

• "Constrained free energies". Some observable f constrained to ξ

Partition function
$$Z(\xi) = \int d^3x \, d^{3N}p \, \delta(f(\mathbf{x}, \mathbf{p}) - \xi) \, e^{-\beta E(\mathbf{x}, \mathbf{p})}$$

 $Z(x_1) = \int dy_1 \, dz_1 d\mathbf{x}_2 \cdots d\mathbf{x}_N \, d^N \mathbf{p} \, e^{-\beta E(\mathbf{x}, \mathbf{p})}$
Helmholtz Free Energy $A(\xi) = -k_B T \ln Z(\xi)$

Probability to be in state with f= ξ $P(\xi) = \frac{Z(\xi)}{Z} = \frac{1}{Z} e^{-\beta G(\xi)}$

Relative probabilities of states ξ_1 and ξ_2

$$P(\xi_2) = P(\xi_1) e^{-\beta [A(\xi_2) - A(\xi_1)]}$$

Typical reaction coordinates ξ

Continuous reaction coordinates

- Distance between two ions/molecules
- Ion position along an ion channel
- Tilting angle of helix in membrane

Discrete reaction coordinates

- ligand bound or unbound
- alchemical transformation (mutation)

Potential of mean force (PMF)

The negative derivative of a constrained free energy is the mean force along the reaction coordinate

$$-\frac{\partial G(a)}{\partial a} = \langle F(a) \rangle$$
$$G(b) = G(a) - \int_{a}^{b} \mathrm{d}\xi \, \langle F(\xi) \rangle$$

 \rightarrow Constrain system at various positions $\xi_1,...,\xi_n$ along reaction coordinate, measure the average force, and integrate the force

A potential of mean force **is** a free energy profile !

Potential of mean force (PMF)

→ Alternative: Umbrella sampling



From Justin Lemkul's umbrella sampling tutorial (gmx website)

- → Restrain system with harmonic potential w_i(ξ) along ξ Check pull_XXX mdp-options in GROMACS
- \rightarrow generate trajectories (probabilities) referring to V(x)+ $w_i(\xi)$
- → Use Weighted Histogram Analysis Method to compute G(ξ) (g_wham)

Barriers are related to rates

- Experimentalists measure rates
- But: many biological processes hardly ever occur in simulation under equilibrium



Computing PMFs from simulation

How do I get a PMF (or free energy profile) from a simulation?

Method I: Simulate very long, get probability along reaction coordinate, compute $G(\xi)$ via

 $G(\xi) = -k_B T \ln P(\xi)$

E.g. the RDF of

water



Method 3: Restrain (bias) system at various positions ξ₁,...,ξ_n along reaction coordinate using additional harmonic potentials ("umbrealla sampling")

$$V_i(\xi) = \frac{1}{2} (\xi - \xi_i)^2$$

Compute unbiased probabilities using WHAM.

Umbrella sampling - simple example

Unknown PMF

Histograms from simulations with restrained system

(a) 2 G(E) / KT 0 ∇ (b) 1000 count 500 n 오 (c) 1000 count 500 0 오 (d) 2 G(ξ) / KT 0 -1 0 reaction coordinate ξ

PMF constructed from histograms

NH₃ permeation across lipid membrane



Permeation of different molecules through different membranes



Pull multiple molecules through in one simulation

Permeation of different molecules through different membranes



Permeation of different molecules through different membranes



Results correlate somewhat with Khex



Neopentane in membrane



Cholesterol makes membrane rigid

Neopentane disrupts favorable Van der Waals interactions in DMPC/Chol

Summary Membrane Permeability

- PMF calculations give Gibbs energy as a function of a coordinate
- Often these coordinates correspond to a physical variable e.g. the membrane orthogonal
- Very useful to extract thermodynamic parameters from MD simulations
- Very complex relation between phospholid, cholesterol concentration, solute molecule and energetics.

Solvation of lons and Organic Molecules in Water Droplets

David van der Spoel





Atmospheric chemistry:

Surface solubilities influence reactions and processes in air or vacuum.

Why do some ions prefer to be on the surface?

- Alkali ions are always completely solvated
- F- is solvated
- Cl-, Br-, I- prefer the surface

Simulations

- Polarizable models from the Roux group for water and ions
- GROMACS 4.0.7
- Potential of mean force (PMF) = free energy profile $\Delta G(r)$
- $\Delta G = \Delta H T\Delta S$

ΔH is related to Etot

- The total energy in the system, or in case of constant temperature simulations the potential energy is directly related to the enthalpy
- For a profile we are interested in the relative enthalpy and entropy

Partial desolvation of a potassium ion from a water droplet





Potential of mean force





There is no simple structural explanation.
Energetics does explain the surface preference.



Caleman, Hub, Van Maaren, Van der Spoel, Proc. Natl. Acad. Sci. 108 (2011) p. 6838

Organic molecules in droplets:



Surface orientation





(E) n-butylamine





(F) neopentane





PMF

- All molecules have minimum at the surface
- Enthalpic barrier to "droplet entry"

Hub, Caleman, Van der Spoel Phys. Chem. Chem. Phys. 14 (2012) pp. 9537

tisdag 18 december 12

Summary Droplets

- Small to medium droplet and water systems studied by polarizable models
- It is possible to get detailed energetics (< 4 kJ/mol resolution)
- New results in atmospheric chemistry
- Sampling is bottleneck



Embrace Complexity

- Simple methods often give incorrect answers - even trends
- Simulation can give well-founded results within an inherent error margin

Acknowledgments









Jochen Hub Carl Caleman Paul van Maaren Maarten Wolf







Further reading

- Umbrella sampling introduction
 B. Roux. The calculation of the potential of mean force using computer simulations. Comp. Phys. Comm., 91:275–282, 1995.
- Free energies from non-equilibrium pulling, excellent summary

 Kosztin, B. Barz, and L. Janosi. Calculating potentials of mean force and
 diffusion coefficients from nonequilibrium processes without jarzynski's equality. J
 Chem Phys, 124 pp. 64106 (2006).
- Error analysis when using Crooks
 Götte and Grubmüller, J Comp Chem 30:447-456 (2009)
- WHAM with Bayesian error estimates: Hub et al. J. Chem. Theor. Comput. 6 pp. 3713-3720 (2010)
- Wennberg et al. Large Effect of Cholesterol on Membrane Permeability, J. Amer. Chem. Soc. 134 (2012) 5351
- Hub et al. Organic molecules on the surface of water droplets An energetic perspective Phys. Chem. Chem. Phys. 14 pp. 9537-9545 (2012)