Determining Free Energy Differences Through Alchemical Transformations

International HPC Summer School

July 2018

Martin Reinhardt

Max Planck Institute for Biophysical Chemistry

Max-Planck Institute for Biophysical Chemistry
Department of Theoretical and Computational Biophysics
Göttingen, Germany
Supervisor: Prof. Helmut Grubmüller

Background

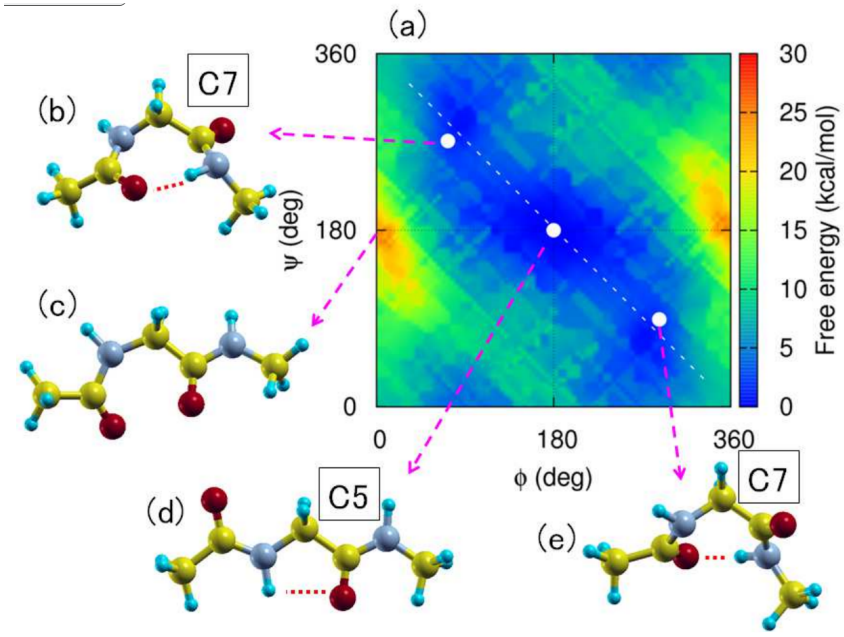
Free energy (differences): Drive (almost) everything, e.g.:

- Conformational changes
- Protein folding
- Binding processes

$$G = -k_B T \ln \int_{-\infty}^{\infty} e^{-H(\mathbf{x})/k_B T} d\mathbf{x}$$

$H(\mathbf{x})$: Hamiltonian of state $\mathbf{x}$

Image: Glycine dipeptide (Nakamura et al., 2014)

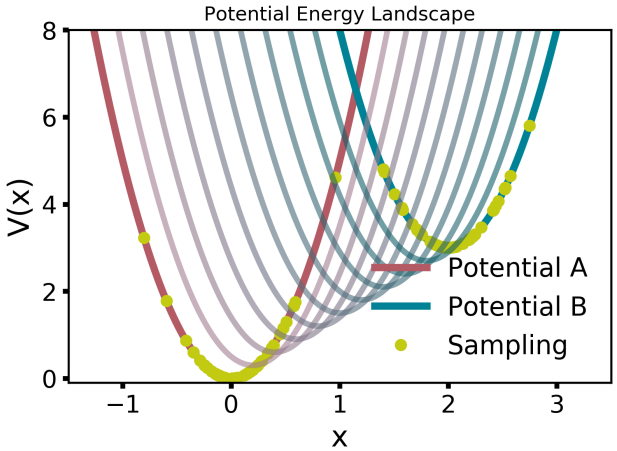


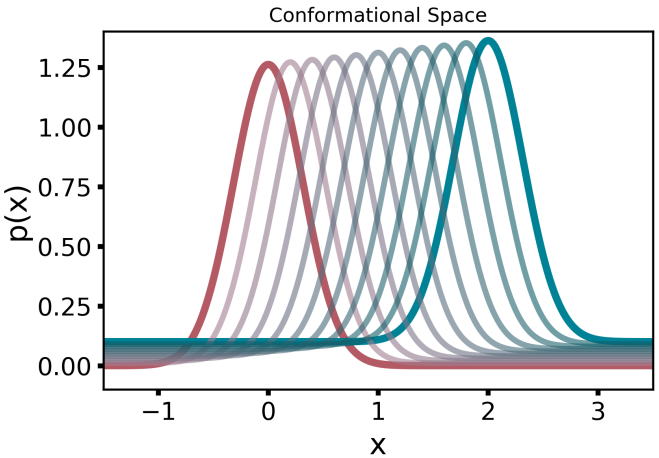
- Problem: Often **high dimensional**, cannot integrate over full phase space
- But for two systems: Often only **difference** required

Max Planck Institute for Biophysical Chemistry

Free Energy Differences

- Use Perturbation Approach: Formally correct, but **sufficient overlap** of (high probability) conformational space required
- Solution: Use set of N_{int} intermediate systems $\Delta G = \sum_{i=1}^{N_{int}+1} \Delta G_i$

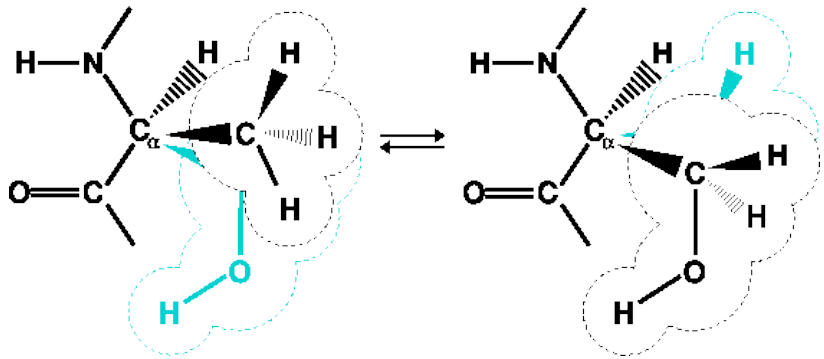




For two different molecules called

Alchemical Transformation

Reason: In the intermediate system we simulate with a parameter mix from the start and the end molecule

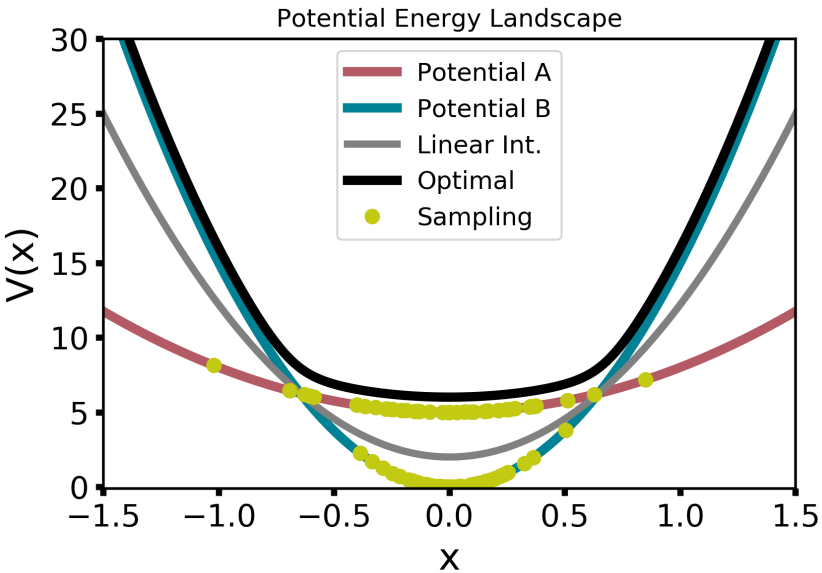


Intermediate Systems - What is the Best Choice?

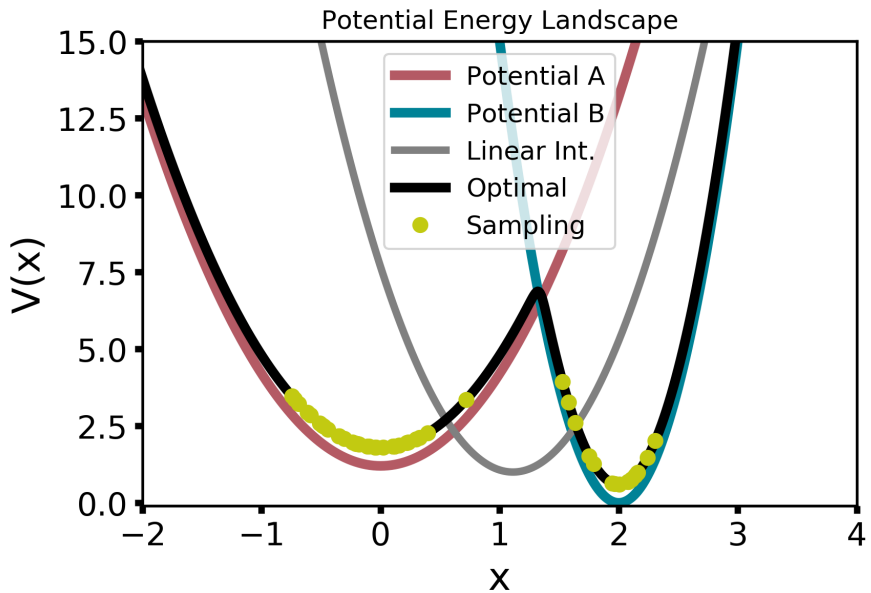
- So far: Linear interpolations of the start and end system

$$H_{int}(\mathbf{x}) = (1-\lambda)H_A(\mathbf{x}) + \lambda H_B(\mathbf{x}) \quad , \quad 0 \leq \lambda \leq 1$$

- But non-linear options also possible
- Examples for two-step case can be found through variational calculus



$$H_{int}(\mathbf{x}) = \ln\left(e^{H_A(\mathbf{x})} + e^{H_B(\mathbf{x})}\right)$$



$$H_{int}(\mathbf{x}) = -0.5\ln\left(e^{-2H_A(\mathbf{x})} + e^{-2H_B(\mathbf{x})}\right)$$

Project Points and HPC component

1. **Theoretical Analysis** :

   - Several intermediate steps

   - Correlated sampling points

2. **Implement** :

   Using GROMACS (GROningen MAchine for Chemical Simulations)

3. **Apply** :

   Free energy calculations necessary for parametrization of labels used in FRET experiments

4. **HPC Component** :

   For converged ΔG values, large sample sizes in many intermediate steps are required. New methods need to be implemented in a parallelized framework.

Max Planck Institute for Biophysical Chemistry