

*Dive into
computational
physical chemistry*

*Lecture 4: Parallel
Tempering*

Glen Hocky
September 26, 2024



Recap (from lecture 2)

Real experiments are at constant temperature or pressure. What we really want in MD is to “sample” configurations with the correct probabilities. At constant temperature, this is given by the Boltzmann equation:

$$P(\mathbf{X}) \propto e^{-\left(\frac{H(\mathbf{X})}{k_B T}\right)}$$

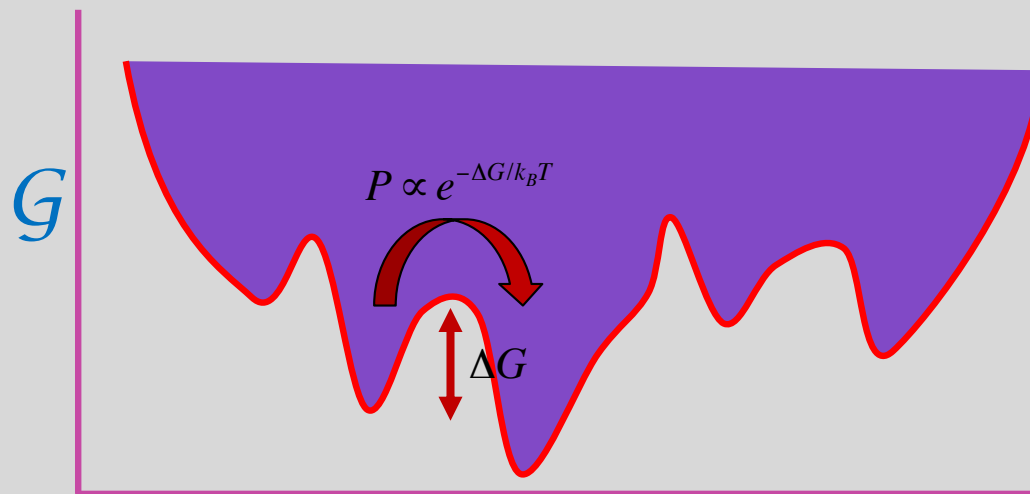
where $H(\mathbf{X}) = U(\mathbf{X}) + K.E.$

What we usually care about is

$$P(q_1, \dots, q_N) \propto e^{-\frac{U(q_1, \dots, q_N)}{k_B T}}$$

Challenge: the rare event problem

The chance of seeing some configuration depends on the (free) energy landscape. But the chance of transitioning between basins depends on barrier heights.



Reaction coordinate (Q)

To accurately sample in simulations, we need techniques that effectively (a) lower barrier or (b) raise the temperature

Running at higher temperature

Higher temperature will let us cross barriers more easily, can we just do that?

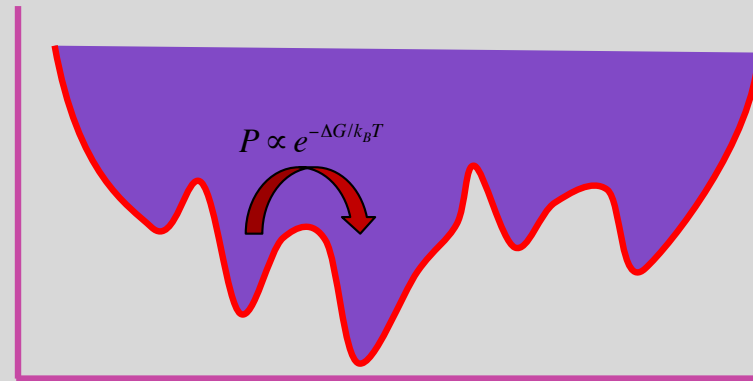
If we run a simulation at temperature T_{hot} ,

$$P_{hot}(\mathbf{X}) \propto e^{-\frac{H(\mathbf{X})}{k_B T_{hot}}}$$

If we multiply this by $1 = e^{-\frac{H(\mathbf{X})}{k_B T_{cold}}} e^{+\frac{H(\mathbf{X})}{k_B T_{cold}}}$

$$\begin{aligned} P_{hot}(\mathbf{X}) &\propto e^{-\frac{H(\mathbf{X})}{k_B T_{hot}}} e^{-\frac{H(\mathbf{X})}{k_B T_{cold}}} e^{+\frac{H(\mathbf{X})}{k_B T_{cold}}} \\ &= e^{-\frac{H(\mathbf{X})}{k_B T_{hot}}} e^{+\frac{H(\mathbf{X})}{k_B T_{cold}}} e^{-\frac{H(\mathbf{X})}{k_B T_{cold}}} \\ &\propto e^{-\frac{H(\mathbf{X})}{k_B T_{hot}}} e^{+\frac{H(\mathbf{X})}{k_B T_{cold}}} P_{cold}(\mathbf{X}) \end{aligned}$$

$$P_{cold}(\mathbf{X}) \propto P_{hot}(\mathbf{X}) e^{\left(\frac{1}{T_{hot}} - \frac{1}{T_{cold}}\right) \frac{H(\mathbf{X})}{k_B}}$$



So, can run a simulation at high temperature and calculate by “reweighting” to lower temperature (or the other way around, but this won’t help us).

Note: $\frac{1}{T_{hot}} < \frac{1}{T_{cold}}$, so the prefactor is negative. This has the result that lower energy structures are more likely at low T and higher at high T (e.g. consider $H +$ vs $-$)

Using intermediate temperatures

Challenge is that this doesn't actually work for large temperature differences, because the configurations explored at higher temperatures may not be those that are relevant at low temperatures.

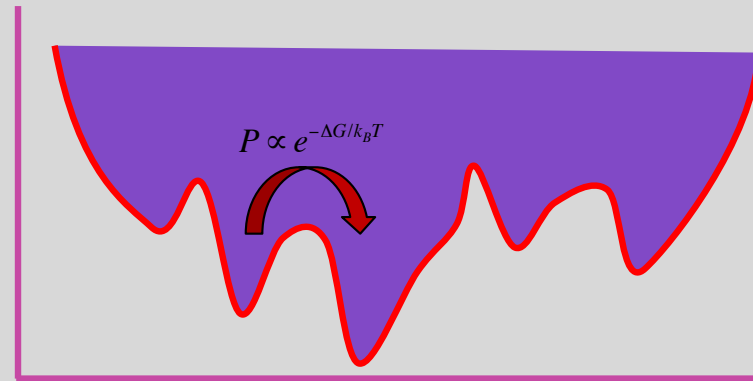
There can also be numerical problem due to the small exponents, since $H(X) \propto N$

$$P_{cold}(\mathbf{X}) \propto P_{hot}(\mathbf{X}) e^{\left(\frac{1}{T_{hot}} - \frac{1}{T_{cold}}\right) \frac{H(\mathbf{X})}{k_B}}$$

Idea: use in-between temperatures where this exponential factor is smaller

J-walking: Frantz, Freeman, Doll; JCP 1990 (<https://doi.org/10.1063/1.458863>), Simulated Tempering: Marinari and Parisi, EPL 1992 (<https://iopscience.iop.org/article/10.1209/0295-5075/19/6/002>)

But how to ensure equilibrium sampling when doing this?



Monte Carlo: another way to generate samples from $P(X)$

Before molecular dynamics, Monte Carlo simulations were invented at Los Alamos during the Manhattan Project

Like MD, want to generate a chain of configurations where each X would arise with proportion $P(X)$

This can be done if we satisfy *detailed balance*

$$P(X)P(X \rightarrow Y) = P(Y)P(Y \rightarrow X)$$

Metropolis Monte Carlo

How do we enforce detailed balance?

$$P(\mathbf{X})P(\mathbf{X} \rightarrow \mathbf{Y}) = P(\mathbf{Y})P(\mathbf{Y} \rightarrow \mathbf{X})$$

Figure out what move probabilities would have to be:

$$P(\mathbf{X} \rightarrow \mathbf{Y}) = \frac{P(\mathbf{Y})}{P(\mathbf{X})}P(\mathbf{Y} \rightarrow \mathbf{X})$$

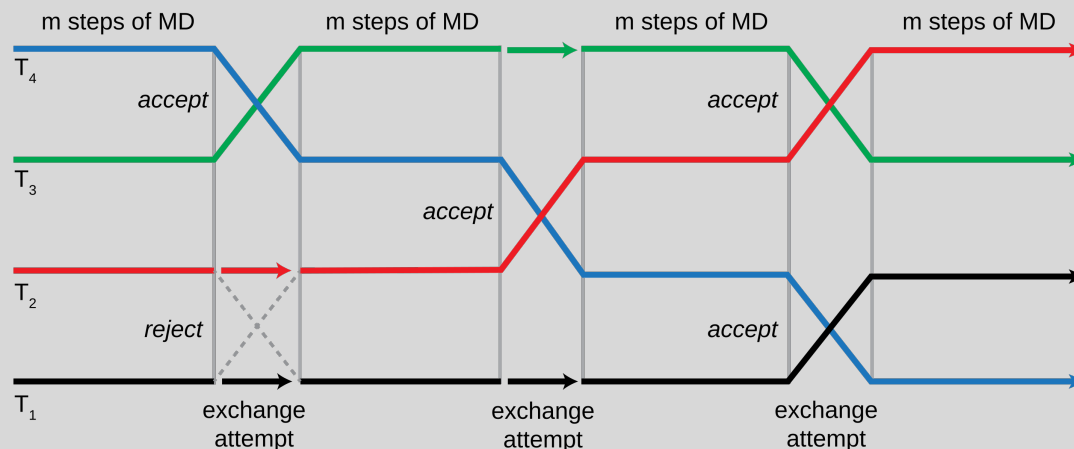
Metropolis, Rosenbluth, Rosenbluth, Teller, Teller solution

$$P(\mathbf{X} \rightarrow \mathbf{Y}) = \min\left(1, \frac{P(\mathbf{Y})}{P(\mathbf{X})}\right)$$

For regular Boltzmann, $P(\mathbf{X} \rightarrow \mathbf{Y}) = \min\left(1, e^{-\frac{H(\mathbf{Y})-H(\mathbf{X})}{k_B T}}\right)$

Parallel Tempering/Replica exchange

An easy and good method: combine MD and MC to sample accurately at low temperatures, using information from high temperatures



$$P(\mathbf{X}, T_1 \rightarrow \mathbf{Y}, T_2) = \min \left(1, \frac{P(\mathbf{Y}, T_1)P(\mathbf{X}, T_2)}{P(\mathbf{X}, T_1)P(\mathbf{Y}, T_2)} \right) = \min \left(1, e^{\left(\frac{1}{T_1} - \frac{1}{T_2}\right)(H(\mathbf{X}) - H(\mathbf{Y}))/k_B} \right)$$

Challenge: overlap between replicas

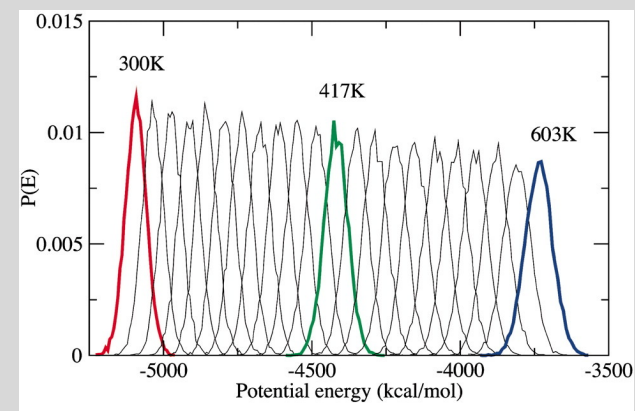
Exchanges happen when energy distributions overlap so that there is some chance $H(\mathbf{X}) - H(\mathbf{Y})$ is negative, where \mathbf{X} is from low T and \mathbf{Y} is from high T

$$\min \left(1, e^{\left(\frac{1}{T_1} - \frac{1}{T_2}\right)(H(\mathbf{X}) - H(\mathbf{Y}))/k_B} \right)$$

Having a large system requires many replicas

Note: temperatures not evenly spread out, how should one choose? (see gromacs manual)

Alternative: Scale part of energy function, effectively making solute hotter (REST)



Analyzing PT simulations

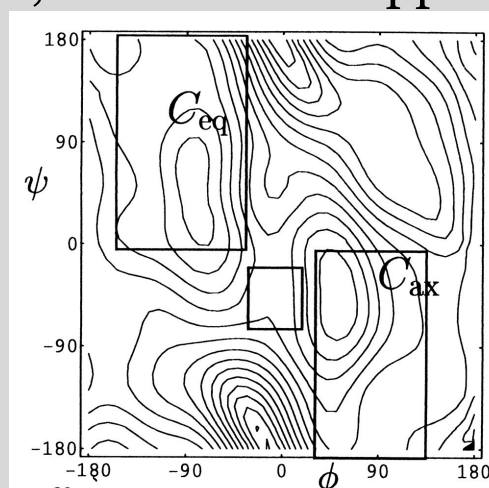
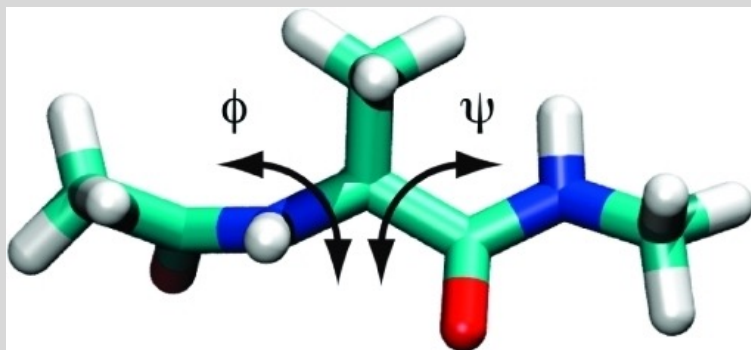
Parallel tempering is still used because it is so simple to analyze. The samples from “Replica i ” are already equilibrated at temperature T_i . To get energy about higher energy states, we must combine information from all simulations using fancier techniques (WHAM, MBAR, etc)

How do we check if it’s working? We want a “random walk in temperature space,” so we should plot the exchanges of temperatures and see that following a single structure would go up to high temperature and down to lower temperature many times

Some codes are temperature continuous, and some are structure continuous. Gromacs is temperature continuous, so the output is ready to analyze already. Otherwise we have to “demux”

Test system: alanine dipeptide

“Alanine dipeptide” is not actually a di-peptide, but rather a capped alanine peptide monomer.



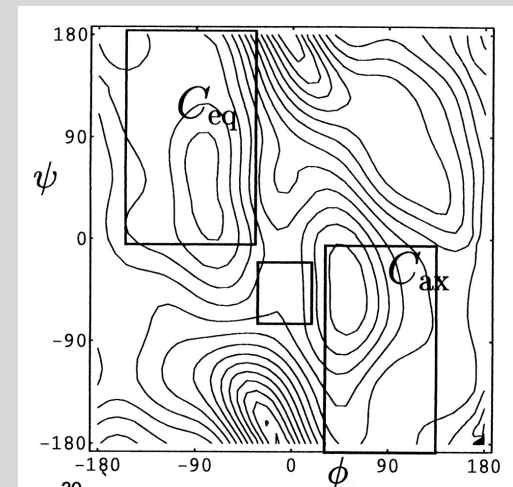
We study its free energy landscape in terms of two main dihedral coordinates that determine the configuration of a peptide.

At low temperature, only one basin (equatorial) is occupied.

Bolhuis, Dellago, Chandler. PNAS 2000. <https://doi.org/10.1073/pnas.100127697>

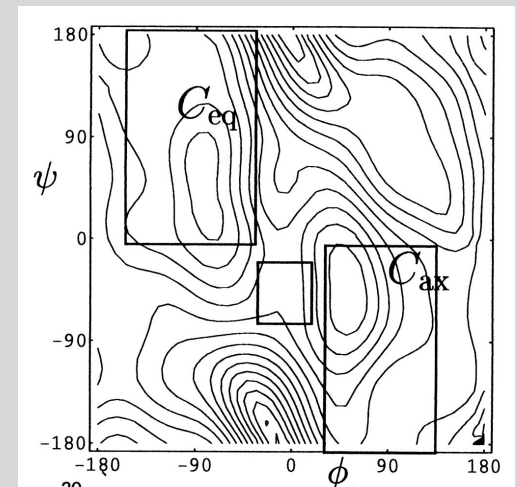
How do we compute a free-energy surface

- MD simulations and replica exchange simulations generate samples from the Boltzmann distribution, $P(X) \propto e^{-\beta U(X)}$
- The *partition function* represents essentially how many configurations the system can be in, but weighted by how likely they are: $Z = \int dX e^{-\beta U(X)}$
- The free energy of the system is given by $F = -kT \ln(Z)$
- The free energy of a 'state' is computed by restricting the integral to that state, so for A and B , $\Delta F = -kT \ln(Z_B) + kT \ln(Z_A)$, where $Z = \int dX h_A(X) e^{-\beta U(X)}$



How do we compute a free-energy surface

- The *free energy surface* or *potential of mean force* is an extension of this idea, where instead of actually defining a state *A* or *B*, we just ask the free energy at a particular grid point in some coordinates, e.g. $-89 < \phi < 91$ and $-49 < \psi < 51$ for a grid size of two degrees
- For MD or PT, this comes down to computing a simple histogram of the observed values since the samples are already Boltzmann weighted
- $F(\phi, \psi) = -kT \ln(P(\phi, \psi))$ where $P(\phi, \psi)$ is a 2d histogram over (ϕ, ψ) , which you will be able to compute w/ python from an MD trajectory



Next week

1. Pull updates into your GitHub and do Week 4 assignment, running MD simulations in gromacs with CharmmGui

Next week

1. Low temperature and replica exchange simulations of alanine dipeptide
2. Pull updates on comp-lab-class github page; see Week 5 assignment