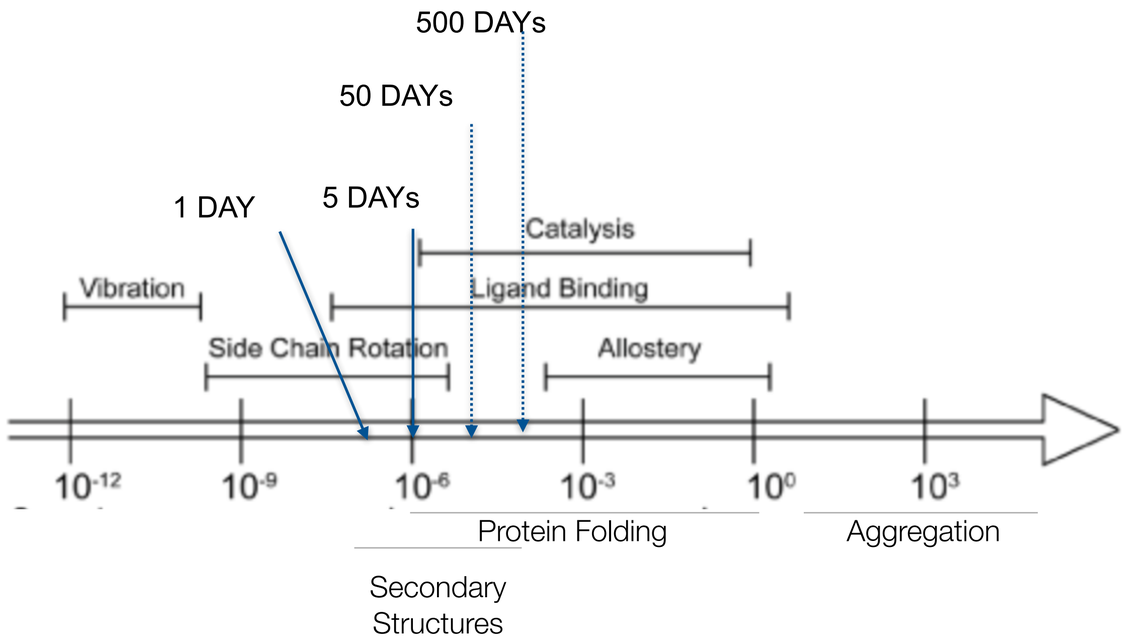
Structural Bioinformatics

Enhanced Sampling

MD simulations: time step, time scales and probabilities



time-step ~10^{-15} s

an MD simulation for a 500 AA protein can run at ~200 ns/day

the probability of observing an event with a rate of 1 ms with a simulation of 10 us is ~1% in the case of a two state kinetics

How many simulations can we run in parallel?

Free Energy Methods

How can we tweak probabilities to make more likely the observation of important configurations?

Boltzmann distribution:

$$pdf(x) \equiv \rho(x) \propto \exp\left[\frac{-U(x)}{k_B T}\right]$$

Force field:

$$V(r) = \sum_{bonds} k_b (b - b_0)^2 + \sum_{angles} k_\theta (\theta - \theta_0)^2 + \sum_{torsions} k_\phi [\cos(n\phi + \delta) + 1]$$
$$+ \sum_{\substack{nonbond \\ pairs}} \left[\frac{q_i q_j}{r_{ij}} + \frac{A_{ij}}{r_{ij}^{12}} - \frac{C_{ij}}{r_{ij}^{6}}\right]$$

Free Energy profile (potential of mean force):

$$PMF(f) \propto -k_B T \ln\left[\int \delta(f - f(x)) e^{-U(x)/k_B T} dx\right] = -k_B T \ln P(f)$$

Free Energy Methods

How can we tweak probabilities to make more likely the observation of important configurations?

Boltzmann distribution: $$pdf(x) \equiv \rho(x) \propto \exp\left[\frac{-U(x)}{k_B T}\right]$$

The probability density function depends on the **Temperature**, so by changing the temperature we change the probability. In particular we increase the probability of observing high-energy conformations.

By increasing the temperature we are actually looking at different *pdf*. The key point is that we want to use a higher temperature to learn about the *pdf* at the temperature of interest. This concept fall under the name **reweighing**.

It is immediately evident that the strength of this approach is that it doesn't need any particular input from the user. This is essentially not knowledge based. On the negative side there is not a strong control, it does not allow to focus the sampling towards a particular goal.

Free Energy Methods

How can we tweak probabilities to make more likely the observation of important configurations?

Boltzmann distribution: $$pdf(x) \equiv \rho(x) \propto \exp\left[\frac{-U(x)}{k_B T}\right]$$

The probability density function depends on the **Force Field**, so by changing the force field we change the probability. In particular we can decrease the interaction energy of some specific term.

$$V(r) = \sum_{bonds} k_b(b - b_0)^2 + \sum_{angles} k_\theta(\theta - \theta_0)^2 + \sum_{torsions} k_\phi[\cos(n\phi + \delta) + 1]$$
$$+ \sum_{\substack{nonbond \\ pairs}} \left[\frac{q_i q_j}{r_{ij}} + \frac{A_{ij}}{r_{ij}^{12}} - \frac{C_{ij}}{r_{ij}^{6}}\right]$$

By decreasing the contribution of dihedral angles for example we speed up the motion of the backbone and side chains but we are actually looking at different *pdf*. The key point is that we want to use a different force field to learn about the *pdf* for the original force field of interest. This concept fall under the name **reweighing**.

Free Energy Methods

How can we tweak probabilities to make more likely the observation of important configurations?

Boltzmann distribution: $$pdf(x) \equiv \rho(x) \propto \exp\left[\frac{-U(x)}{k_B T}\right]$$

The probability density function depends on the **Energy**, so by changing the energy we change the probability. In particular we can add a new energy term. We can add a potential to modify the *pdf* along a specific conformational parameter. For example we could try to flatten the *pdf* in some specific direction.

$$PMF(f) \propto -k_B T \ln\left[\int \delta(f - f(x)) e^{-U(x)/k_B T} dx\right] = -k_B T \ln P(f)$$

$$pdf'(x) \propto \exp\left[\frac{-(U(x) + V(f(x)))}{k_B T}\right]$$

This allows to increase the probability of observing configurations along specific reaction coordinates. Again since we will observe a new *pdf* we should make it in such a way to go back to the original one.

Free Energy Methods

How can we tweak probabilities to make more likely the observation of important configurations?

Boltzmann distribution: $$pdf(x) \equiv \rho(x) \propto \exp\left[\frac{-U(x)}{k_B T}\right]$$

The probability density function depends on the **Energy**, so by changing the energy we change the probability. In particular we can add a new time-dependent energy term. In particular we can add a potential to modify the *pdf* along a specific conformational parameter. For example we could try to flatten the *pdf* in some specific direction. One problem is that often we don't know the PMF along a coordinate *a priori* so we would like to add an adaptive potential (a time-dependent potential) that uses what we learn on-the-fly.

$$PMF(f) \propto -k_B T \ln\left[\int \delta(f - f(x)) e^{-U(x)/k_B T} dx\right] = -k_B T \ln P(f)$$

$$pdf'(x,t) \propto \exp\left[\frac{-(U(x) + V(f(x),t))}{k_B T}\right]$$

This allows to increase the probability of observing configurations along specific reaction coordinates. Again since we will observe a new *pdf* we should make it in such a way to go back to the original one.

Enhanced Sampling

Extended Ensemble Methods

1. **Parallel Tempering**
2. Hamiltonian Replica Exchange
3. Thermodynamics Integration

Collective Variables
Methods

1. **Umbrella Sampling**
2. Steering MD
3. **Metadynamics**

Parallel Tempering

At least in principle if we know all the configurations we can obtain the *pdf* at any temperature

$$pdf(x, T_1) = \frac{\exp(-U(x)/k_B T_1)}{\int \exp(-U(x)/k_B T_1) dx}$$

$$pdf(x, T_2) = \frac{\exp(-U(x)/k_B T_2)}{\int \exp(-U(x)/k_B T_2) dx}$$

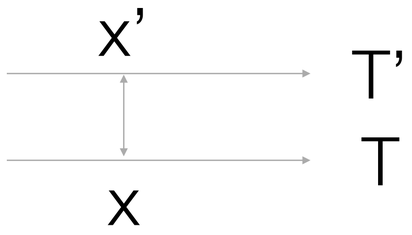
It is enough to recalculate the formula using a different temperature. But as we already discussed we are never in this condition. So this is useless.

What is instead the *pdf* for two copies of the same system at two different temperatures?

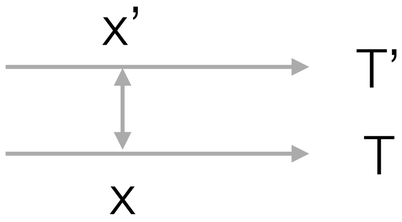
$$pdf(x, x', T, T') \propto \exp\left[\frac{-U(x)}{k_B T}\right] \exp\left[\frac{-U(x')}{k_B T'}\right]$$

This is the joint probability of observing two configurations from the same force field at two different temperatures.

In principle we can take conformations obtained from the simulation at high temperature and bring them at low temperature and vice versa. We could do it only if we are not going to change the equilibrium:



Parallel Tempering



The probability of exchange is the same as the joint probability for x'->x and x->x' at their own temperatures

$$w(x \to x', T)P(x, T) = w(x' \to x, T)P(x', T)$$
$$w(x \to x', T')P(x, T') = w(x' \to x, T')P(x', T')$$

At equilibrium the flux in one direction is the same of the flux in the opposite

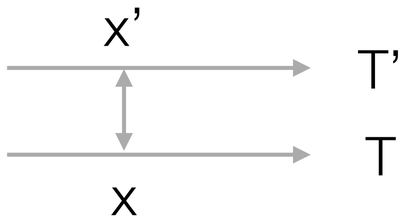
Taking the joint probability:

$$w(x \to x', T)P(x, T)w(x' \to x, T')P(x', T') = w(x' \to x, T)P(x', T)w(x \to x', T')P(x, T')$$

$$\frac{w(x \to x', T)w(x' \to x, T')}{w(x' \to x, T)w(x \to x', T')} = \frac{P(x', T)P(x, T')}{P(x, T)P(x', T')} = \frac{\exp\left[-U(x')/\,k_B T - U(x)/k_B T'\right]}{\exp\left[-U(x)/\,k_B T - U(x')/k_B T'\right]} =$$

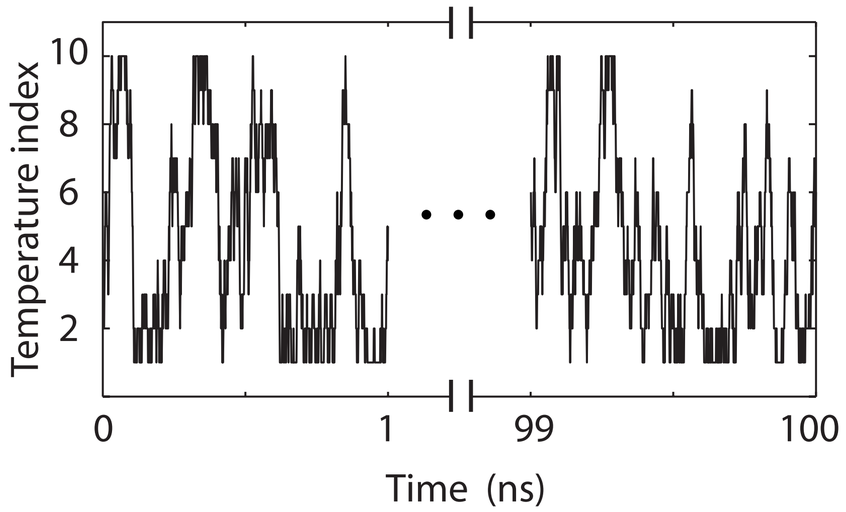
$$\exp\left[-\frac{(U(x) - U(x'))}{k_B}\left(\frac{1}{T'} - \frac{1}{T}\right)\right]$$

Parallel Tempering



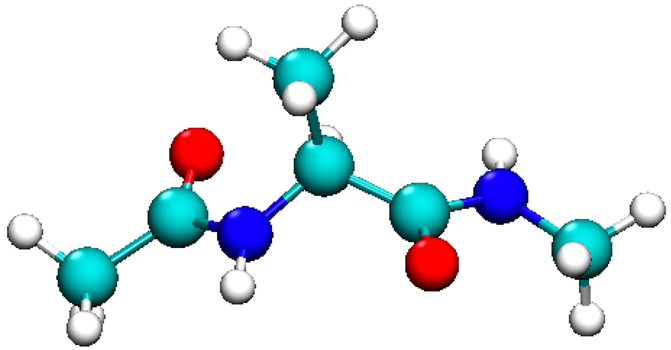
Apart from the math what happen in practice is that we have two (many) MD at many temperatures, and with some frequency an exchange is attempted between configurations at different temperature. This is done using a **Metropolis-Monte-Carlo simulation**, with a probability of acceptance given by:

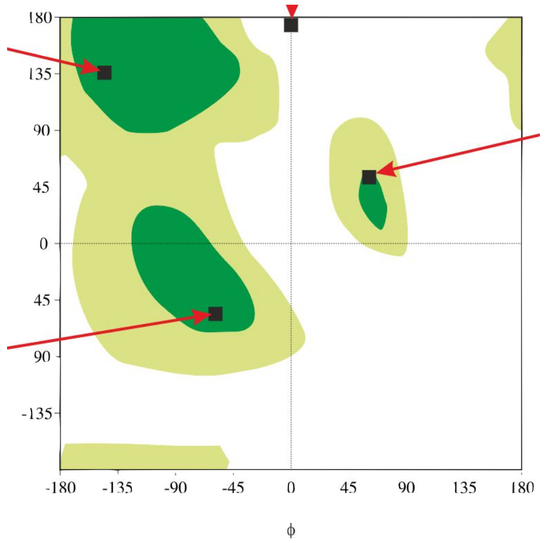
$$P_{acc}(x \leftrightarrow x') = \min\left(1, \exp\left[-\frac{(U(x) - U(x'))}{k_B}\left(\frac{1}{T'} - \frac{1}{T}\right)\right]\right)$$



So a simulation will move in time and temperature.
The result is that one obtains at once the temperature dependence of the system. The cons is that the probability of exchange decreases a lot with the system size, so for large system one could easily need hundreds of replicas. The other cons is that lack of control of the sampling, not all processes depends in the same way from temperature.

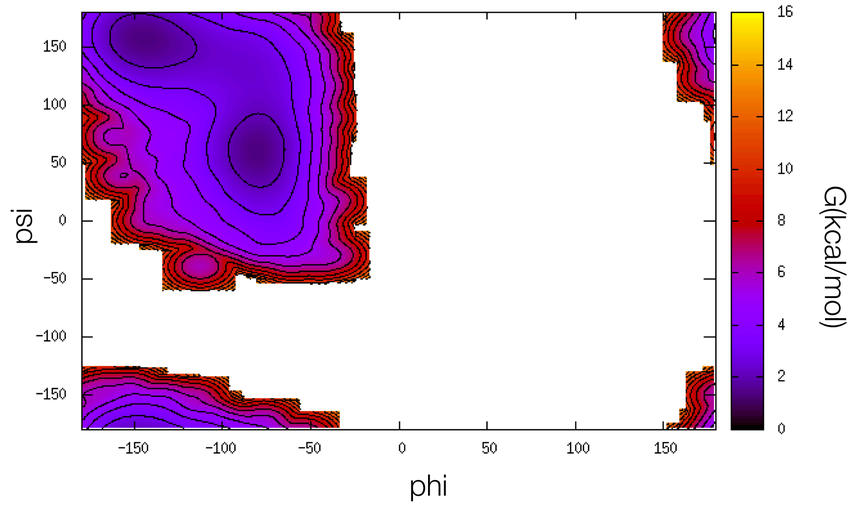
Parallel Tempering





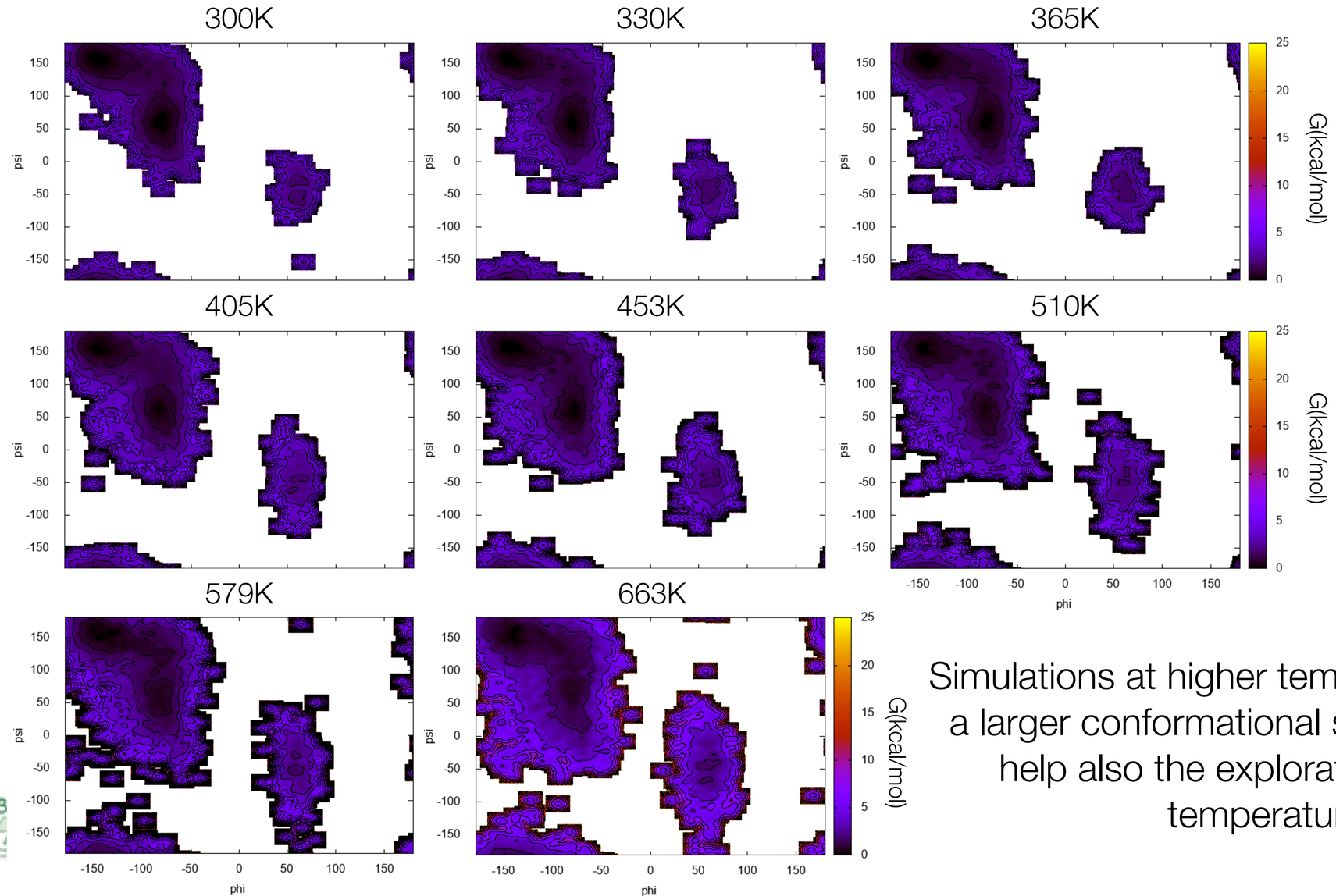
The expectation for this system is to sample three relevant conformations.

A short MD (80ns) at 300K samples only the left region:



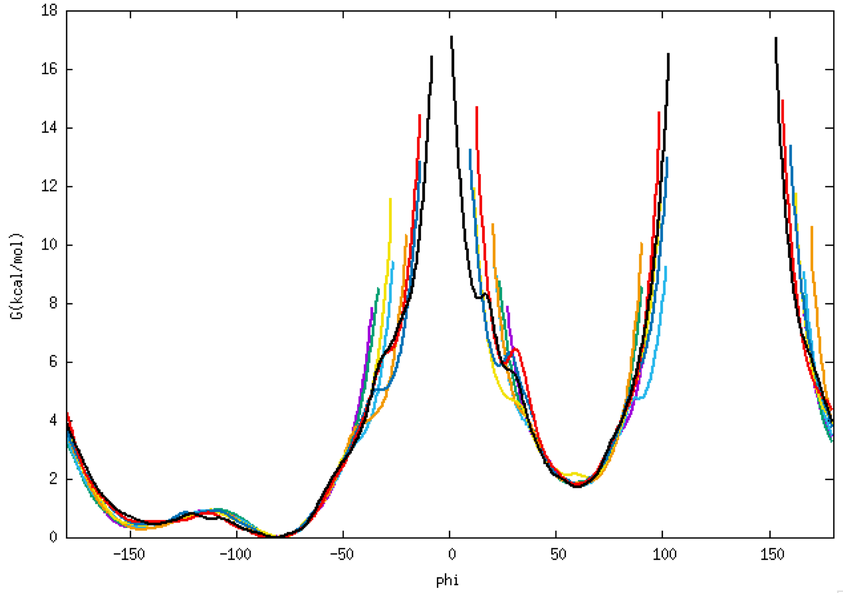
A longer simulation could sample the right region as well, but how longer? Alternatively here we run 8 simulations at 8 different temperature for 10ns each (the same total simulation time)

Parallel Tempering



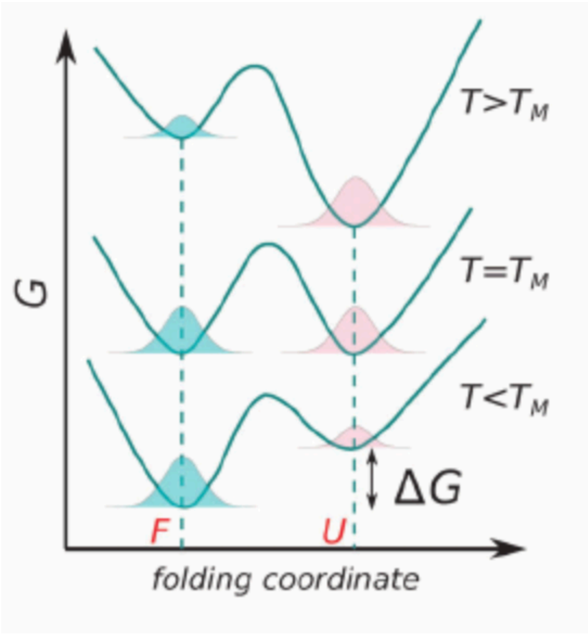
Simulations at higher temperature explore a larger conformational space and thus help also the exploration at lower temperature

Parallel Tempering

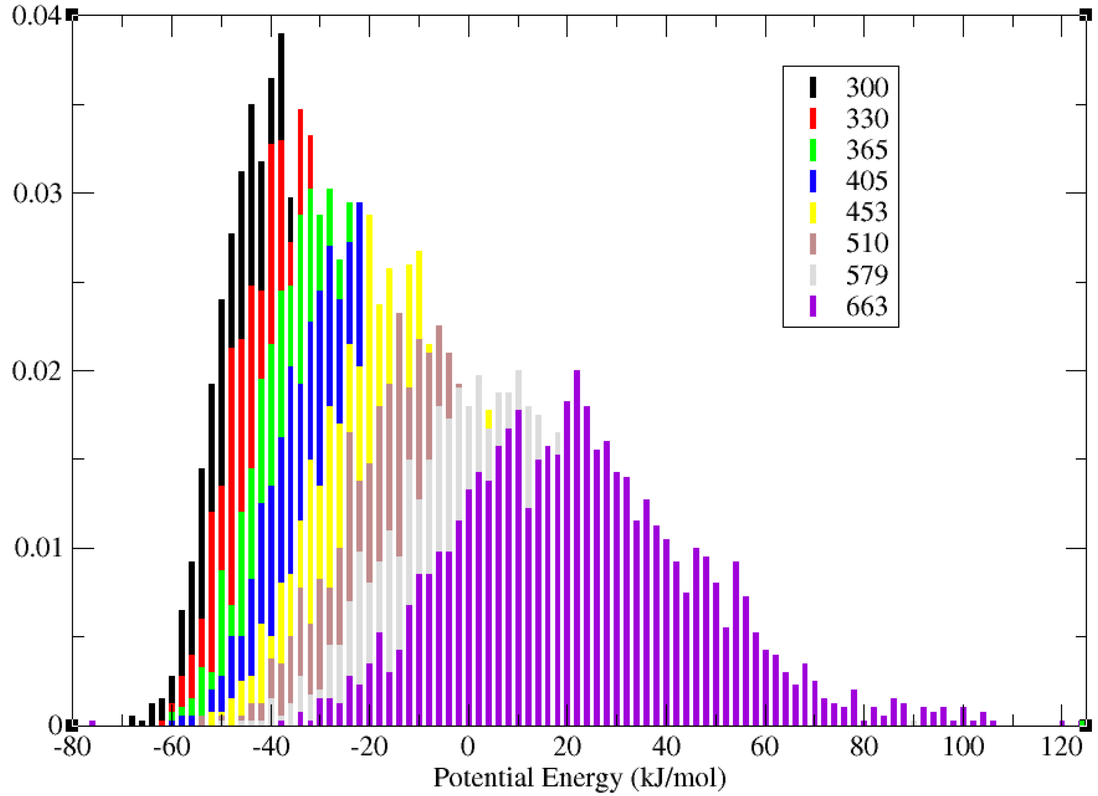


By looking a the free energy projected only on phi it is clear that the free energy in this case does not change with the temperature, so using higher temperature here is particularly efficient. Furthermore form the barrier estimate of ~10kcal/mol we could say that we would have needed a 1us simulation to sample this conformational change! We got the result with 80ns so we have speed up the simulation by a factor >10

This is an extreme case, usually the free energy will change with the temperature, think at proteins, the higher the temperature the more is gonna be populated the unfolded state. In this case the high temperature will not help too much.



Parallel Tempering: How to set the temperatures



In order to have a good probability of exchange

$$P_{acc}(x \leftrightarrow x') = \min\left(1, \exp\left[-\frac{(U(x) - U(x'))}{k_B}\left(\frac{1}{T'} - \frac{1}{T}\right)\right]\right)$$

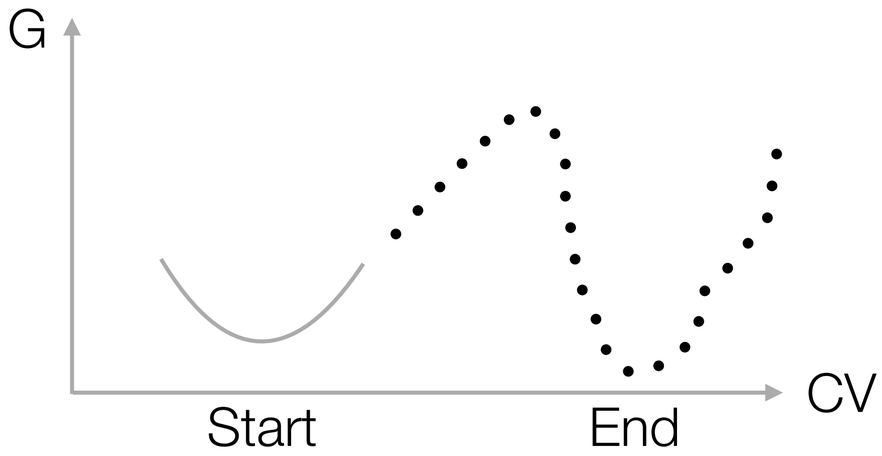
The histograms of the potential energy should overlap significantly (~10-30%), here it means that we could have used even less replicas, the overlap is very high, but for real system in water (here we are in vacuum) the temperature difference needed can become of the order ~1 K.

Umbrella Sampling

PT allow speeding up the sampling without the need of specific knowledge of the system.
Oftentimes we have ideas on how a process should work.

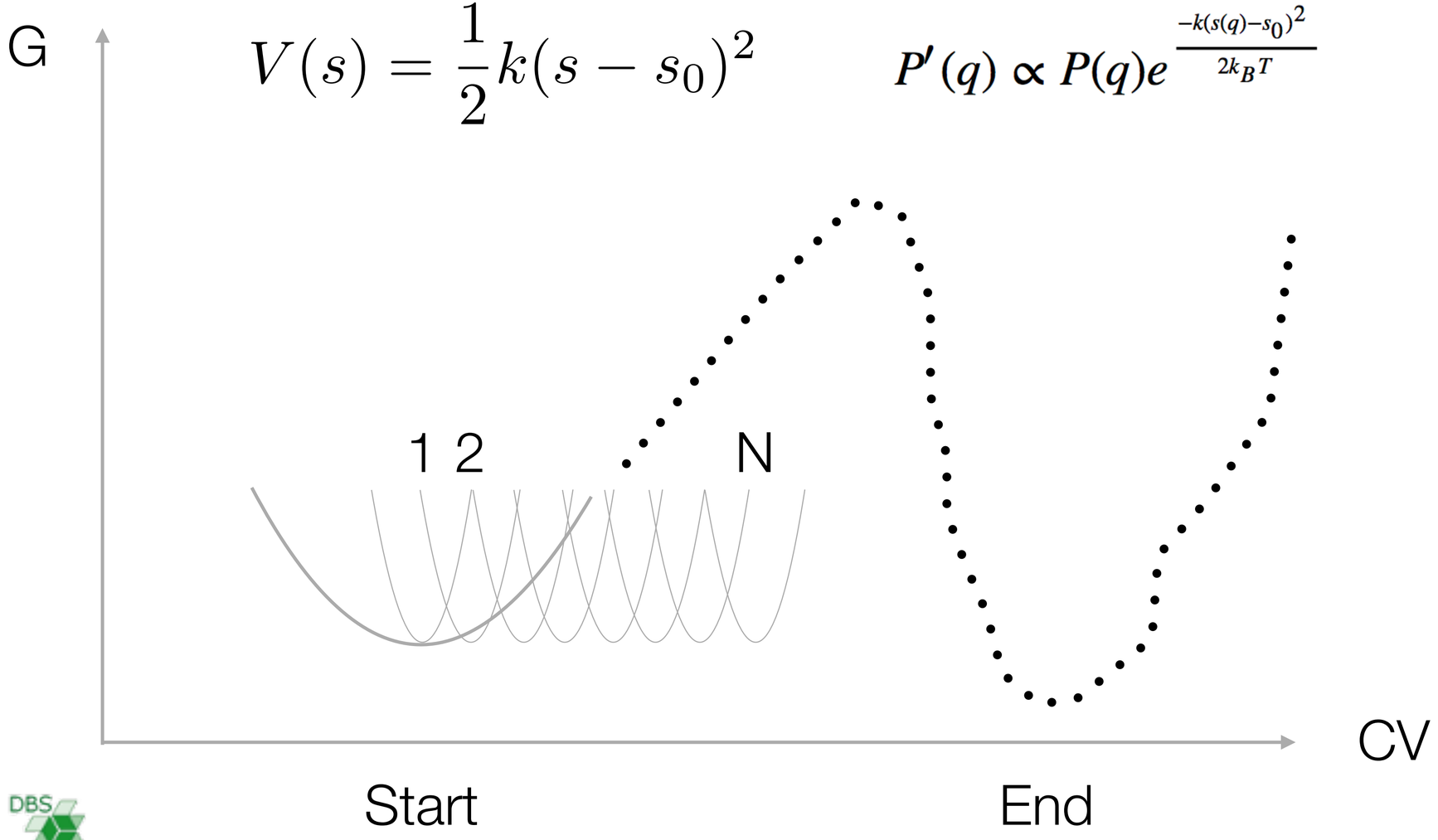
For example:

1. alanine dipeptide: the conformational change is related to phi-psi;
2. ligand-binding: the binding requires a decrease in the distance ligand-binding-site
3. maybe one knows the initial and final structure of a conformational change
4. …



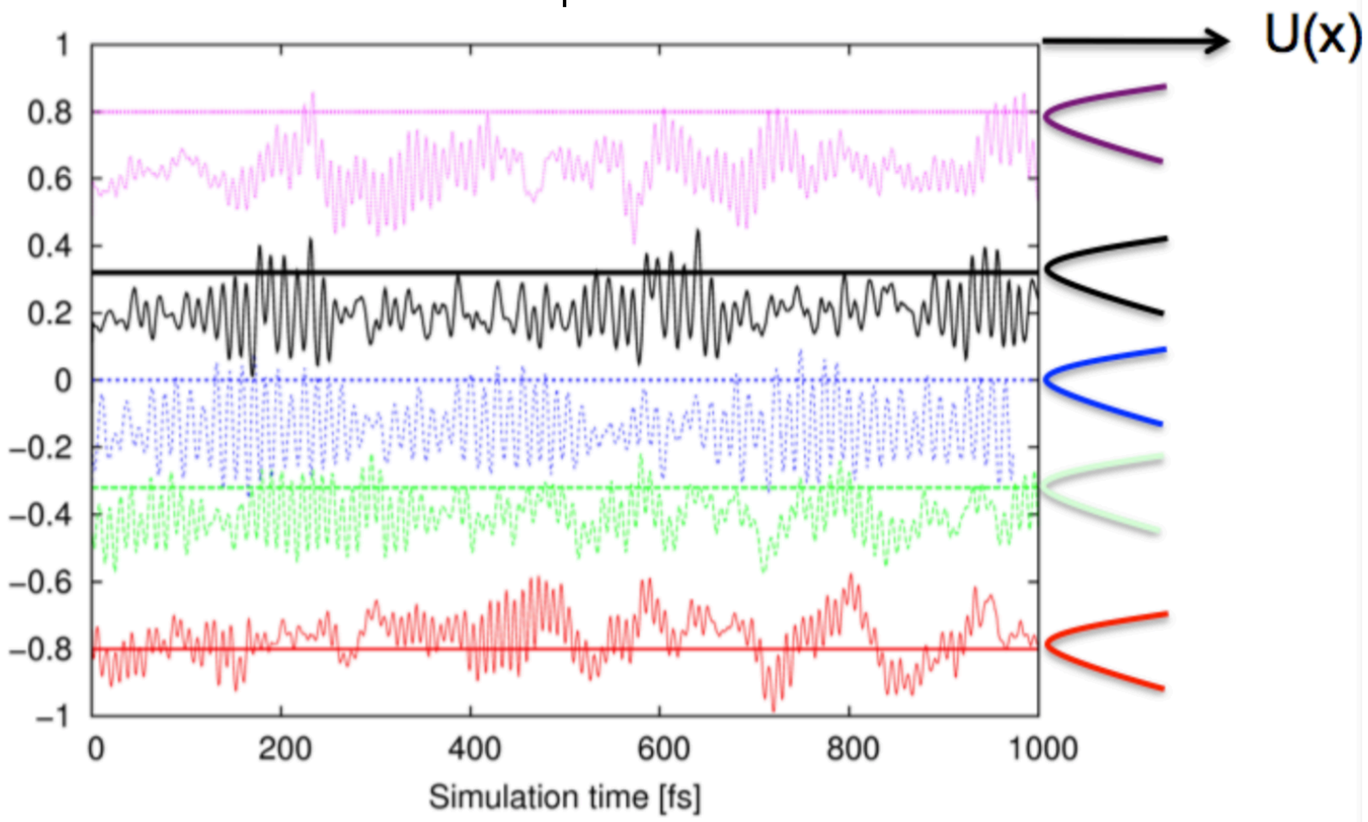
How can we get the free energy profile over one (few) conformational parameters?

Umbrella Sampling



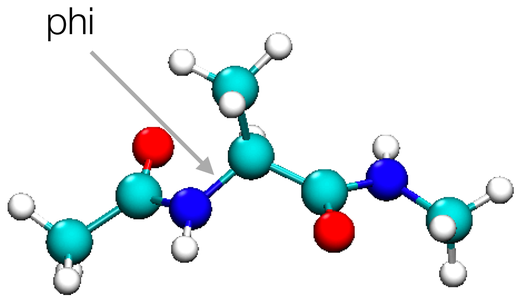
Umbrella Sampling

Many simulations are performed each one centred around a specific value of a conformational parameter

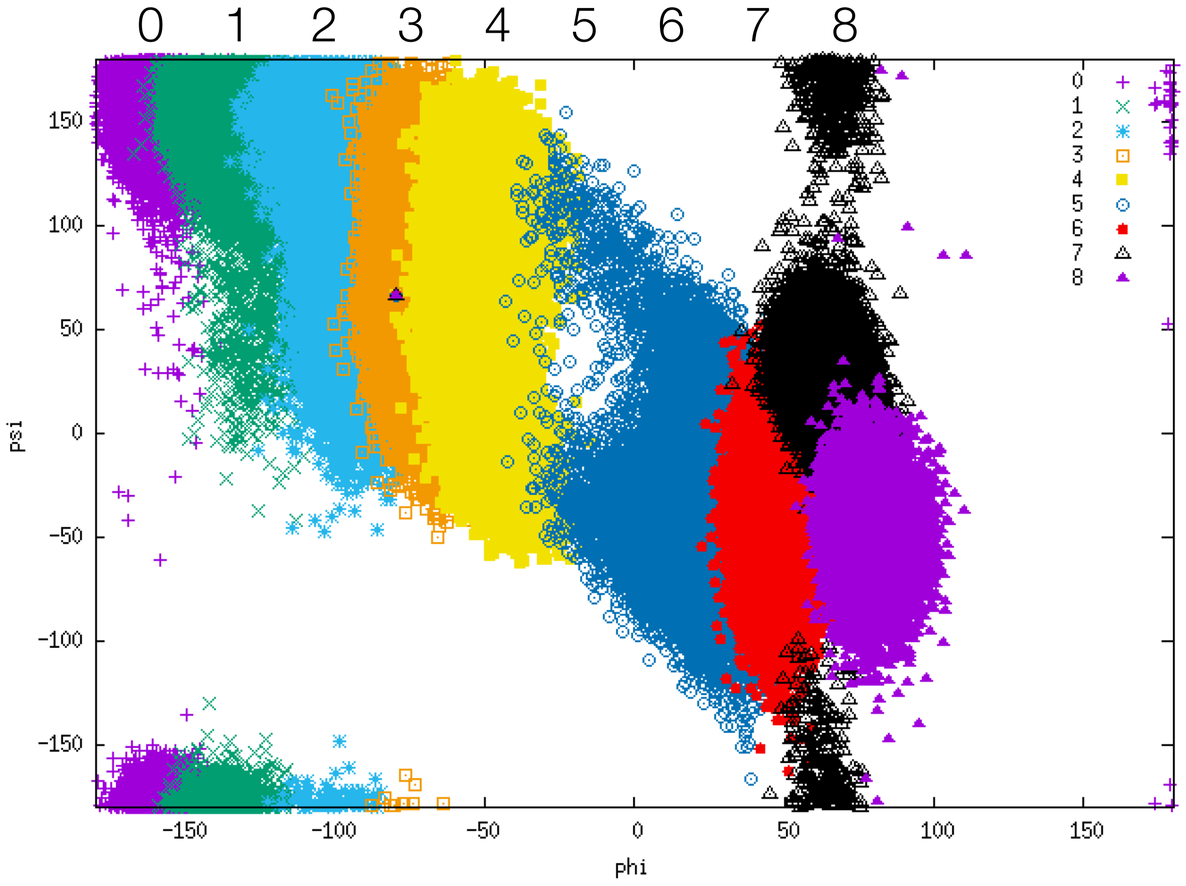


Umbrella Sampling

Many simulations are performed each one centred around a specific value of a conformational parameter



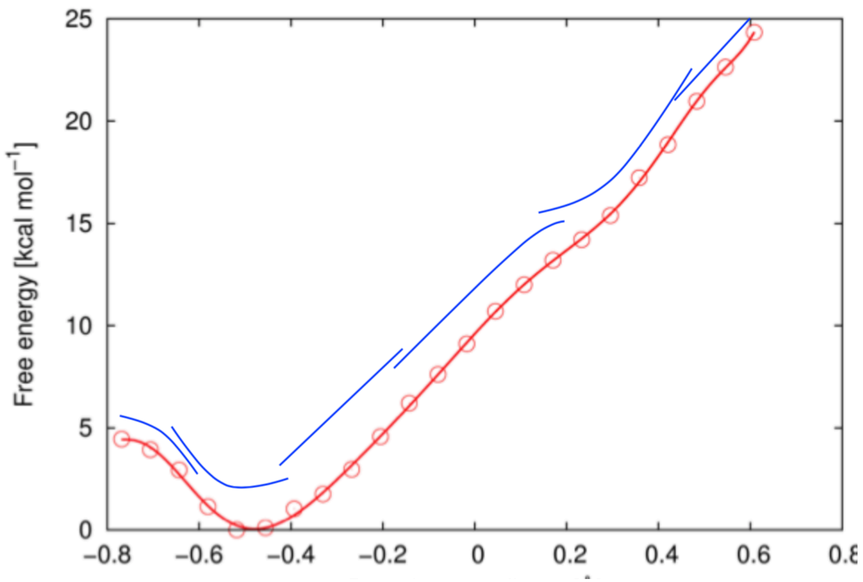
In this way we can force the simulation to sample regions that would not normally sample

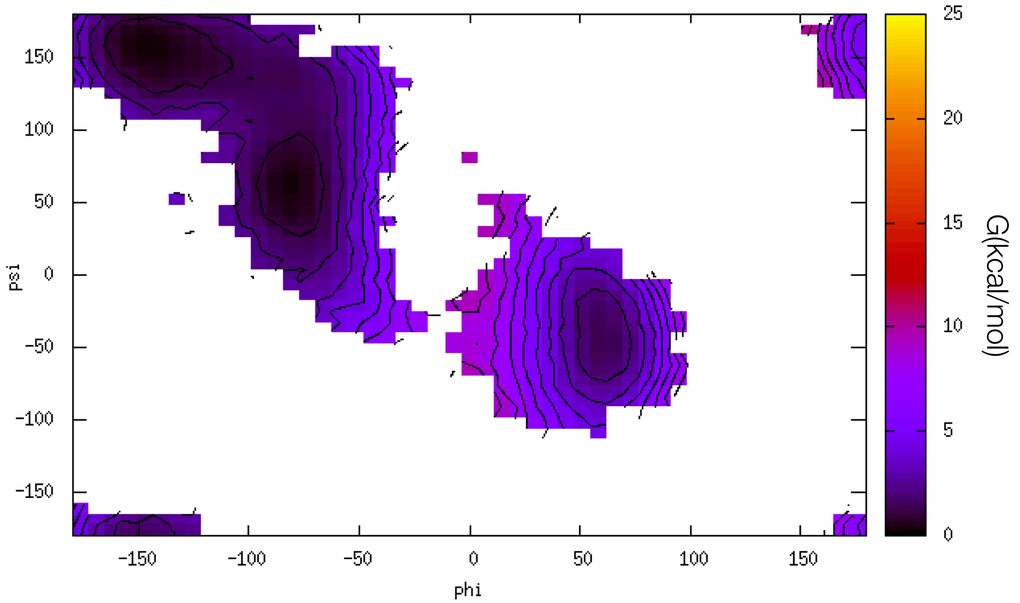


Umbrella Sampling

The problem is how to use this to go back to the original force-field behaviour?

Qualitatively the idea is that each simulation will give a local estimate and that we need to merge them together:





Metadynamics

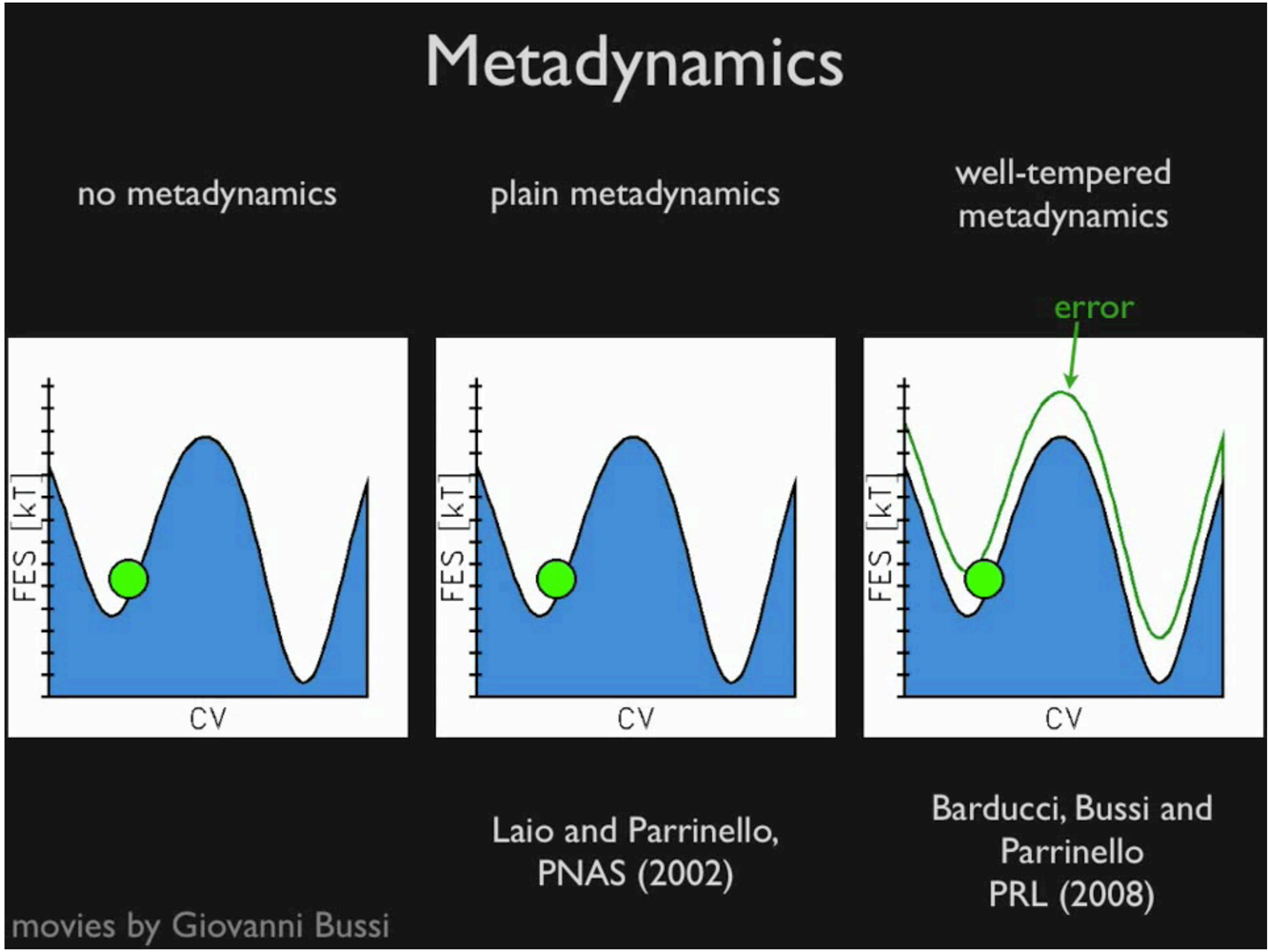
$$pdf'(x,t) \propto \exp\left[\frac{-(U(x)+V(f(x),t))}{k_B T}\right]$$

As a last case we will see how we can build a bias to speed up the simulation that learns by going.

In a standard MD the probability of visiting a conformation is constant

$$\dot{V}(s,t) = 0$$

This means that if two states are separated by a barrier it will be unlikely to cross the barrier (low probability) and visit a different state

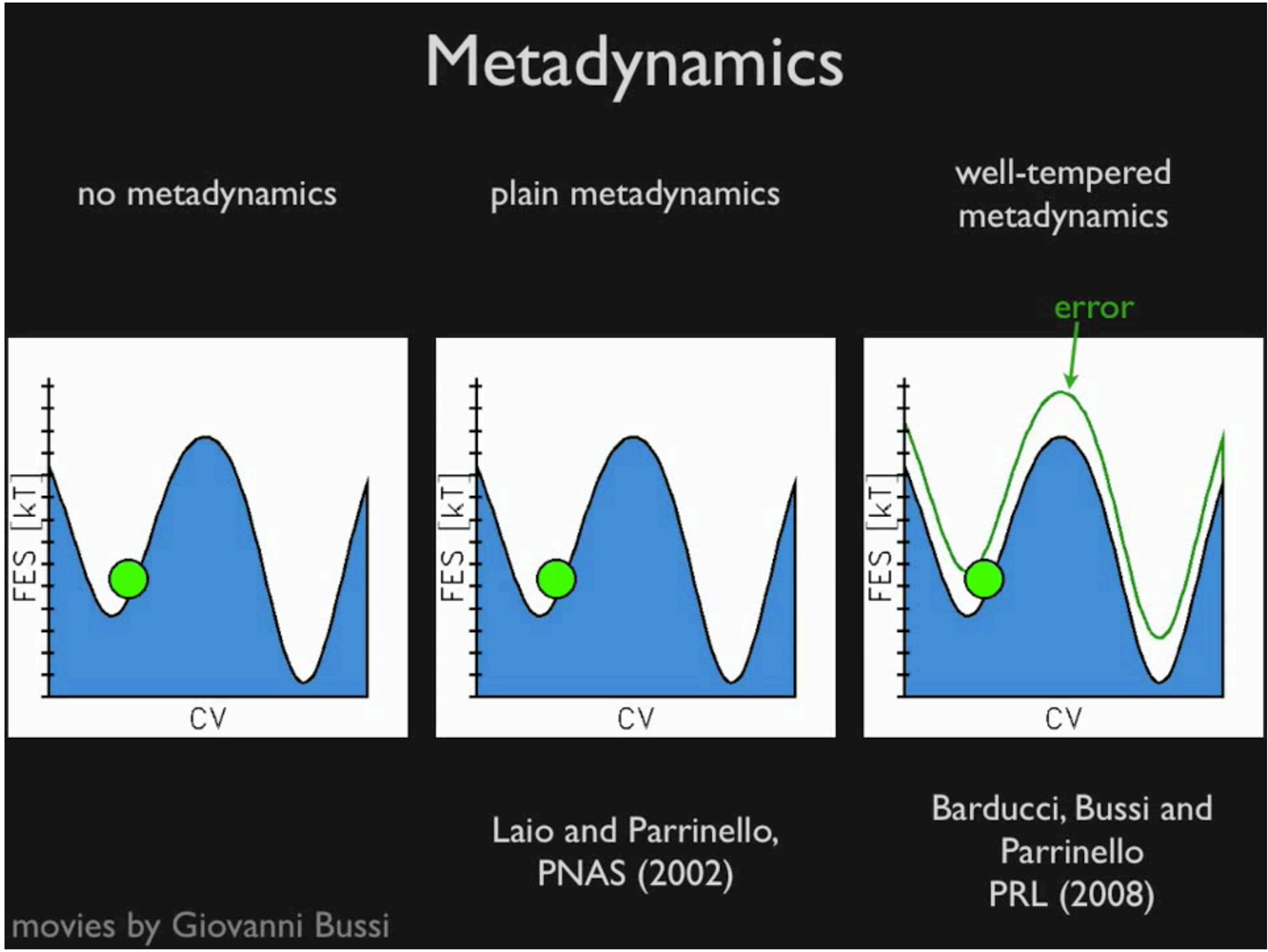


Metadynamics

We can add a bias proportional to the time spent in a particular region. The problem is this never ends.

$$\dot{V}(s,t) = \omega e^{-(s-s(t))^2/2\sigma_s^2}$$

The original idea of metadynamics was to try to make the probability of visiting any conformation equal. But this result in making likely also very uninteresting configurations.

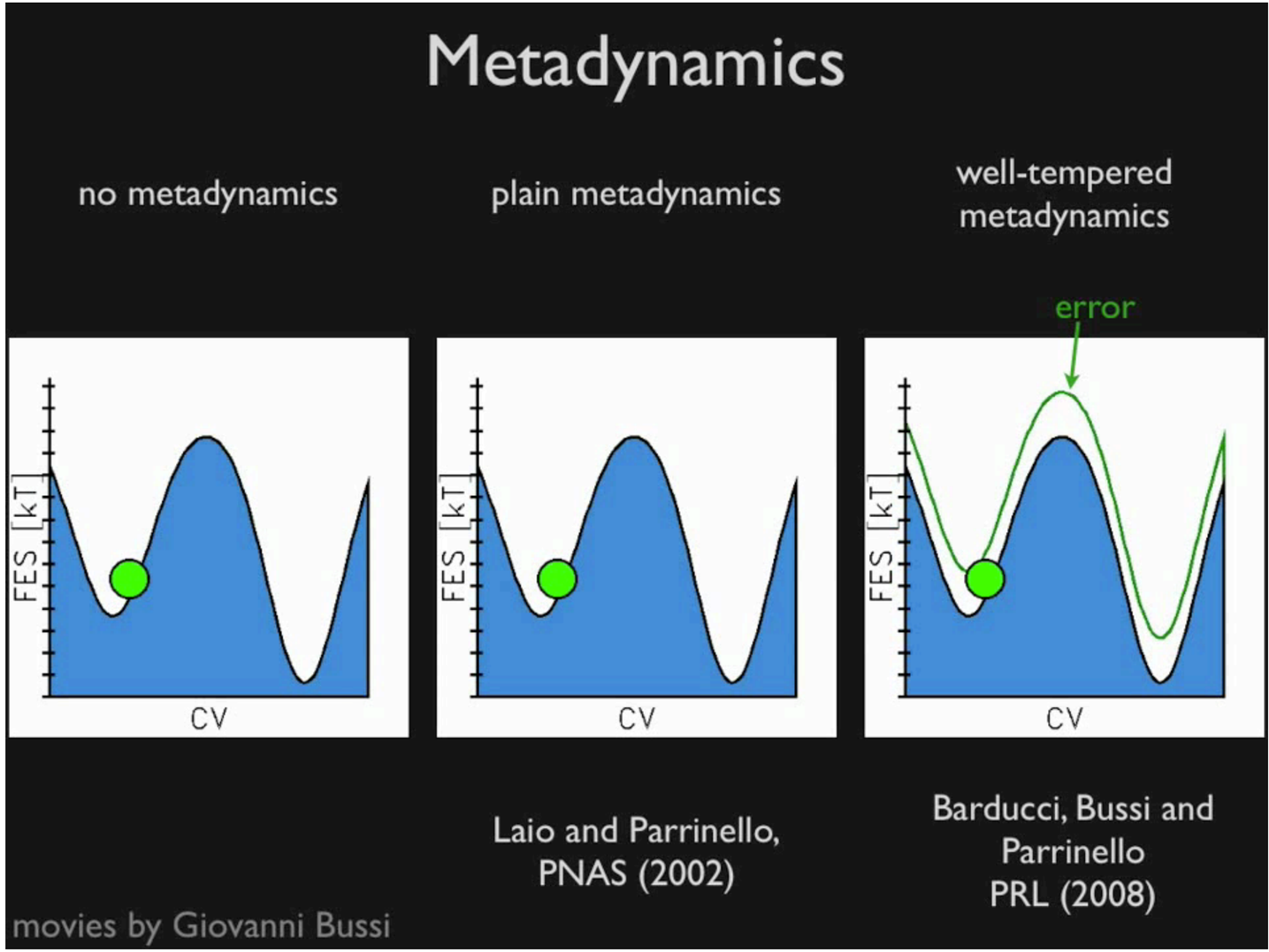


Metadynamics

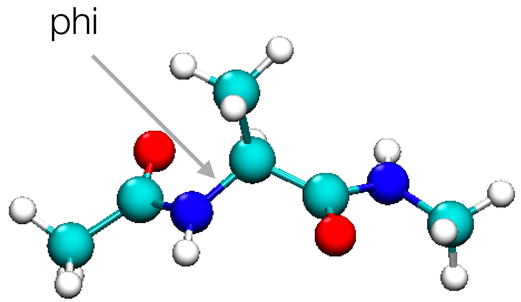
In Well-Tempered Metadynamics the idea is to increase to rescale the probability only of conformations up to some energy defined from a parameter ΔT

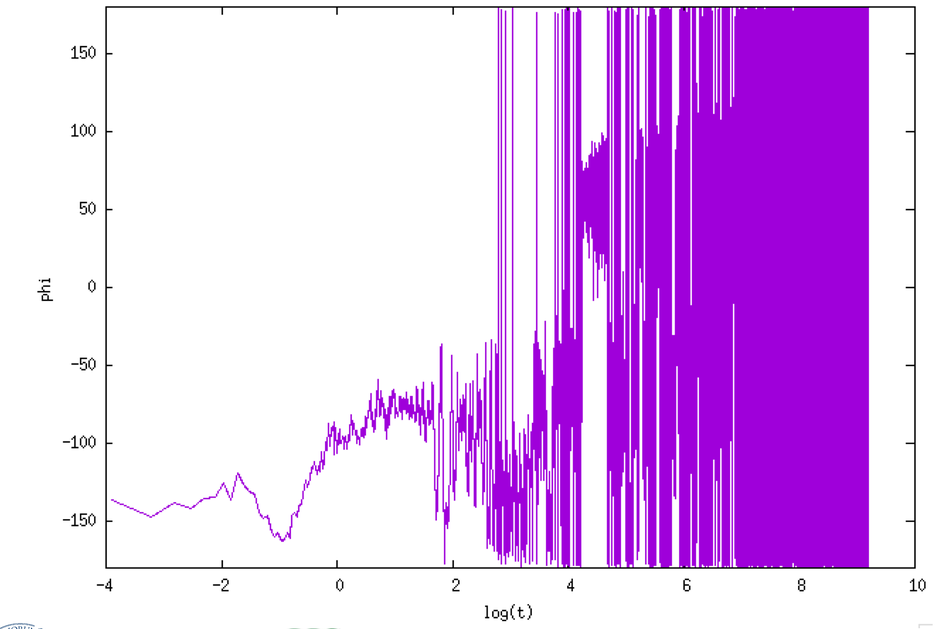
$$\dot{V}(s,t) = \omega e^{-[V(s,t)/\Delta T]} e^{-(s-s(t))^2/2\sigma_s^2}$$

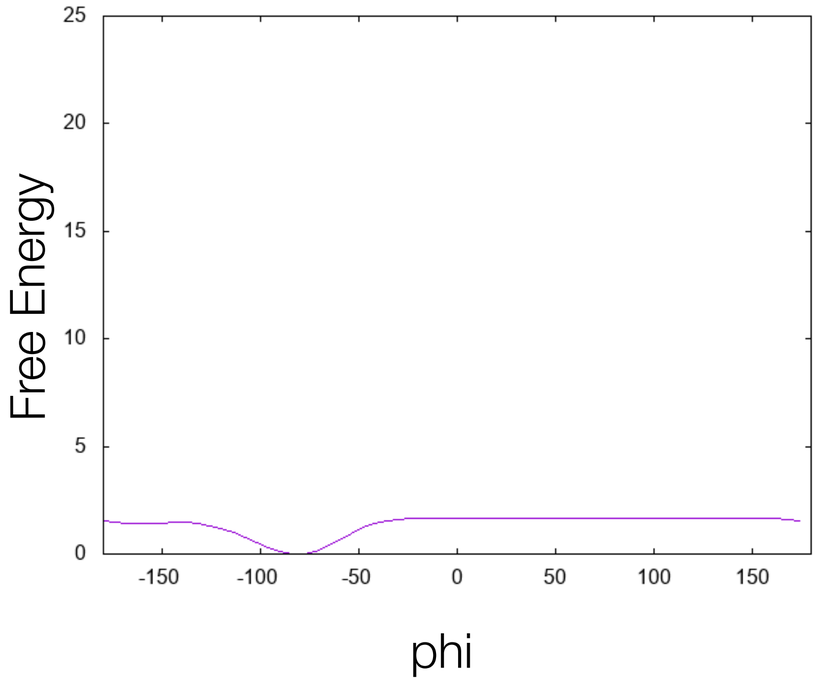
We can add a bias proportional to the time spent in a particular region. But counting every addition as 1/t



Metadynamics

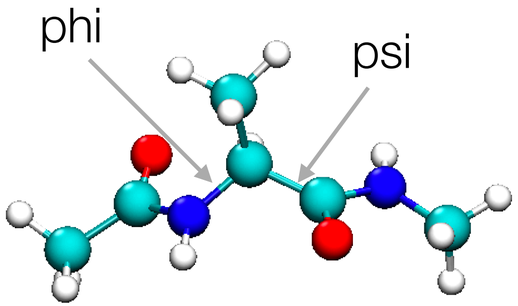




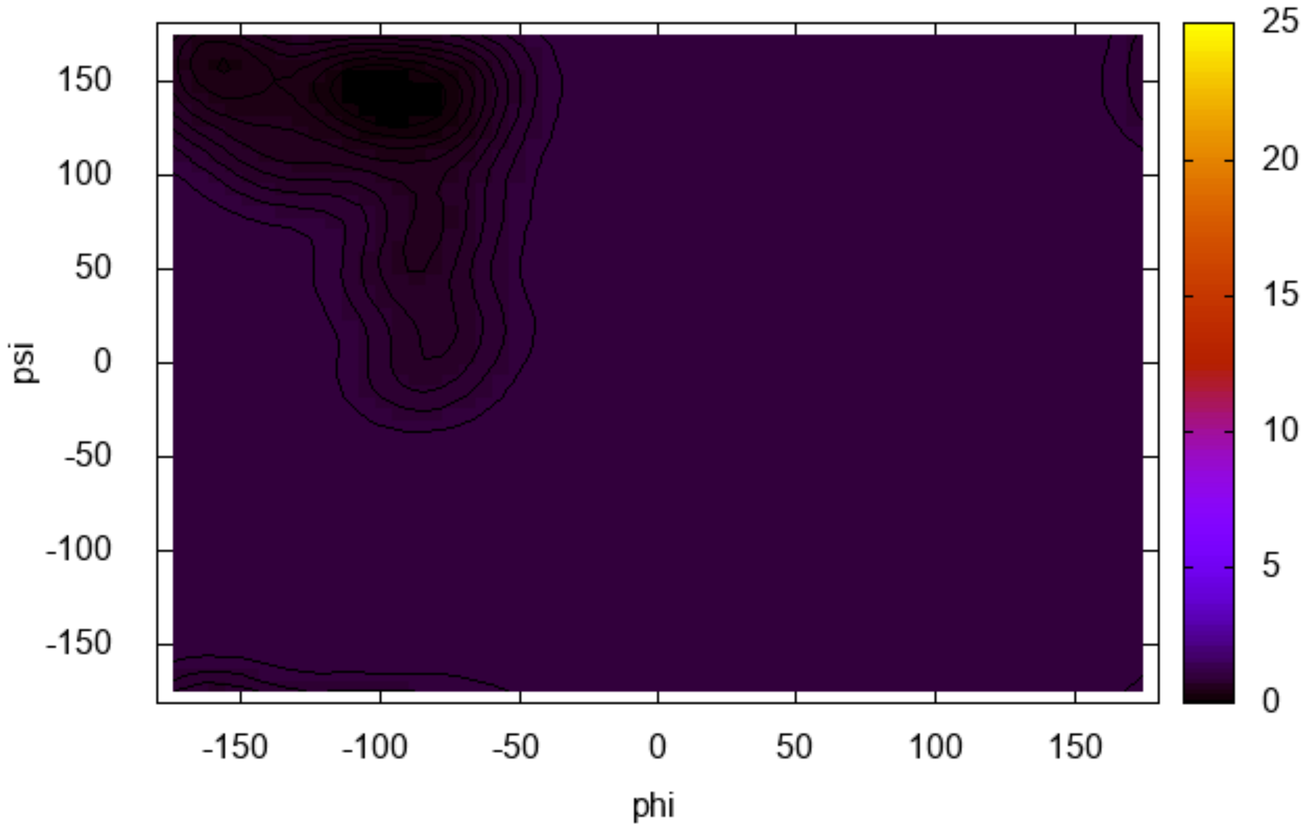


Now this is done with a single 10ns simulation (so 8x faster than PT or US)

Metadynamics

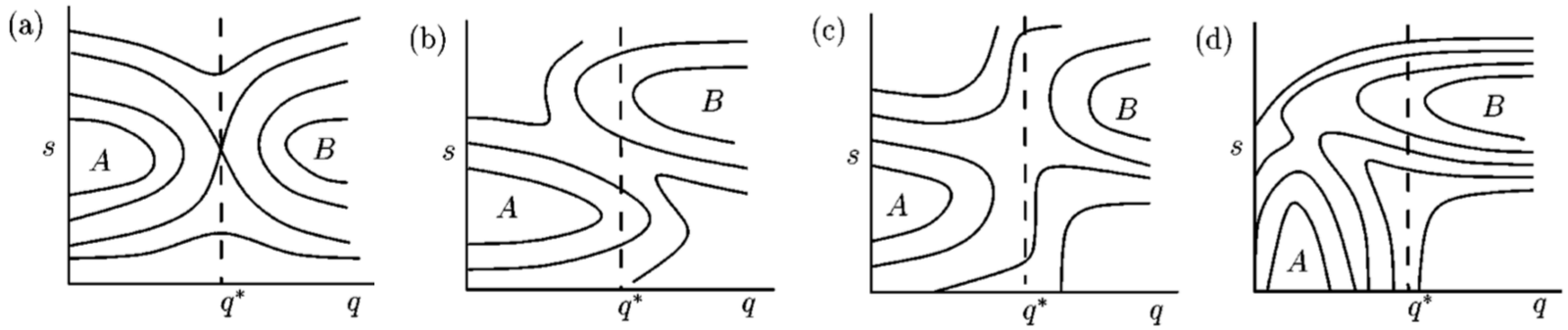


We can now easily run it in more than 1D (2 or 3, not much more)



Choosing CV can be tricky

Projections are tricky:



With US and MTD the big issue is not anymore how to speed up the sampling but how to choose a good reaction coordinate. We cannot choose many because the method is exponentially slower with the number of CVs.