

# Theory of Metadynamics (with some examples)

Max Bonomi  
Institut Pasteur - CNRS  
[mbonomi@pasteur.fr](mailto:mbonomi@pasteur.fr)

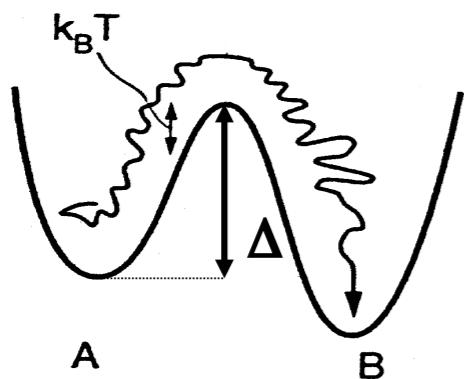


@BonomiMax

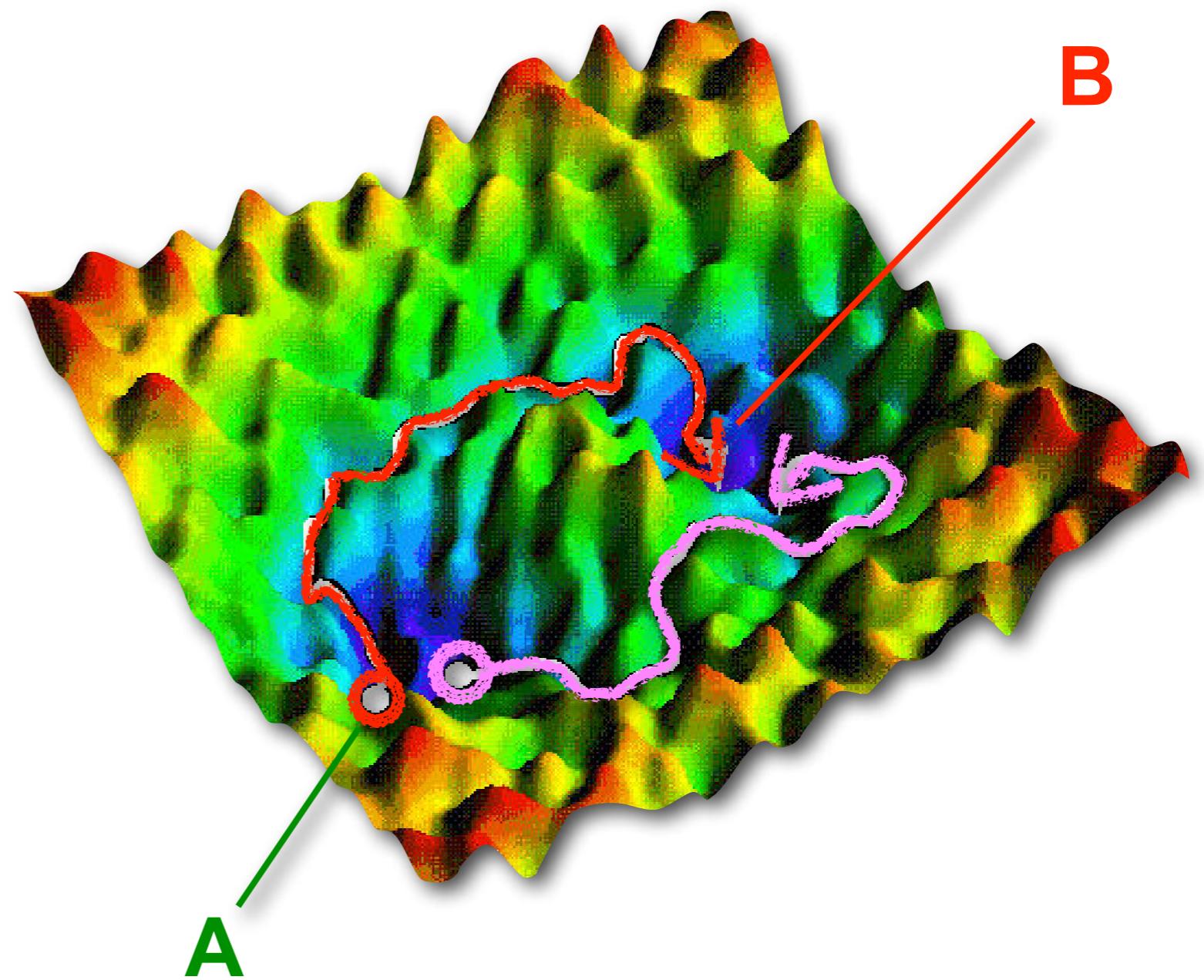
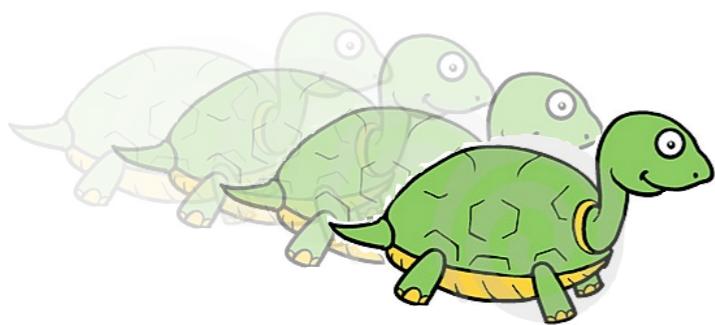
# A time scale problem

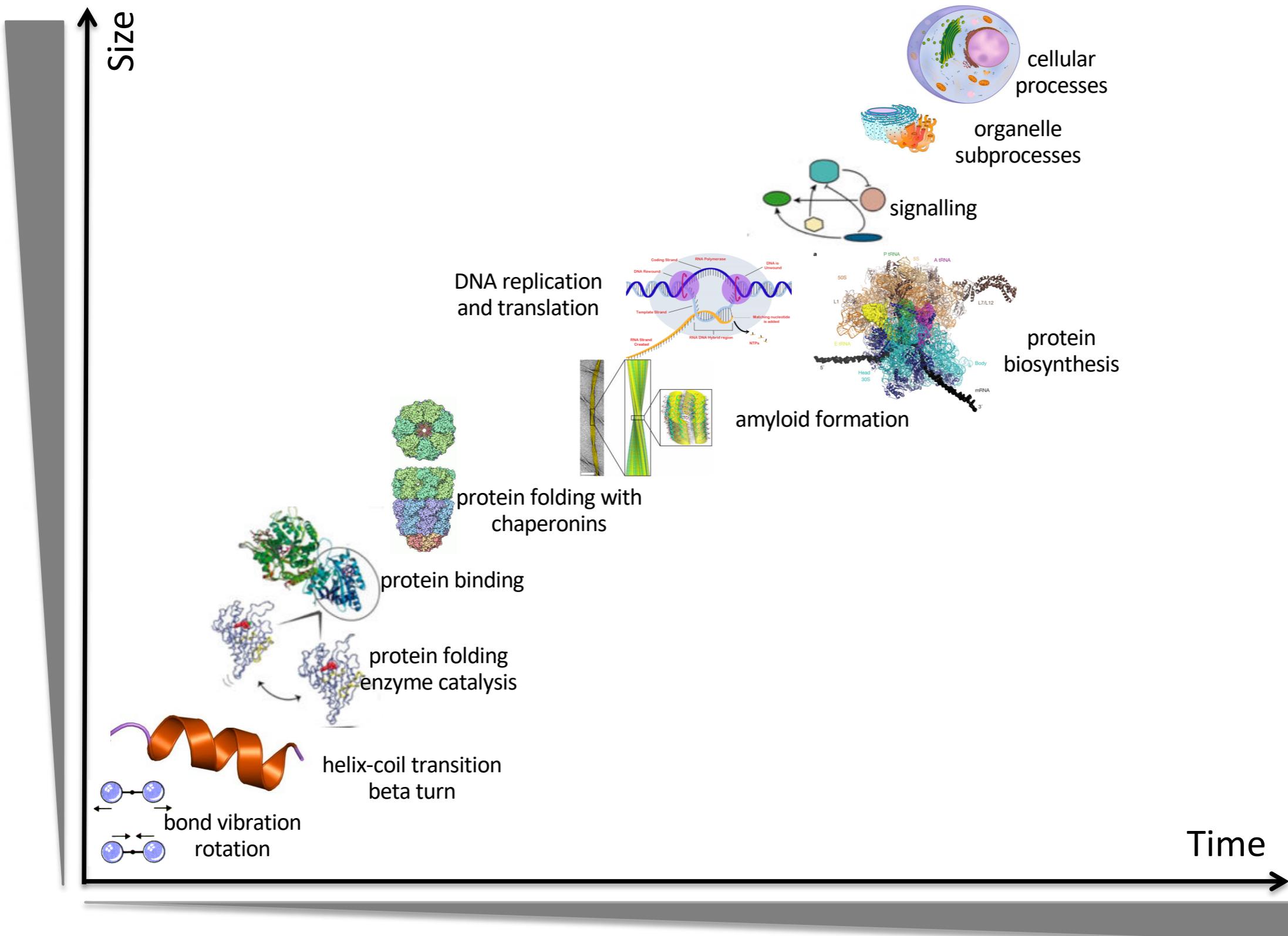
Sampling capability of MD and MC is limited by the time scale accessible in “standard” simulations:

★ Activated events

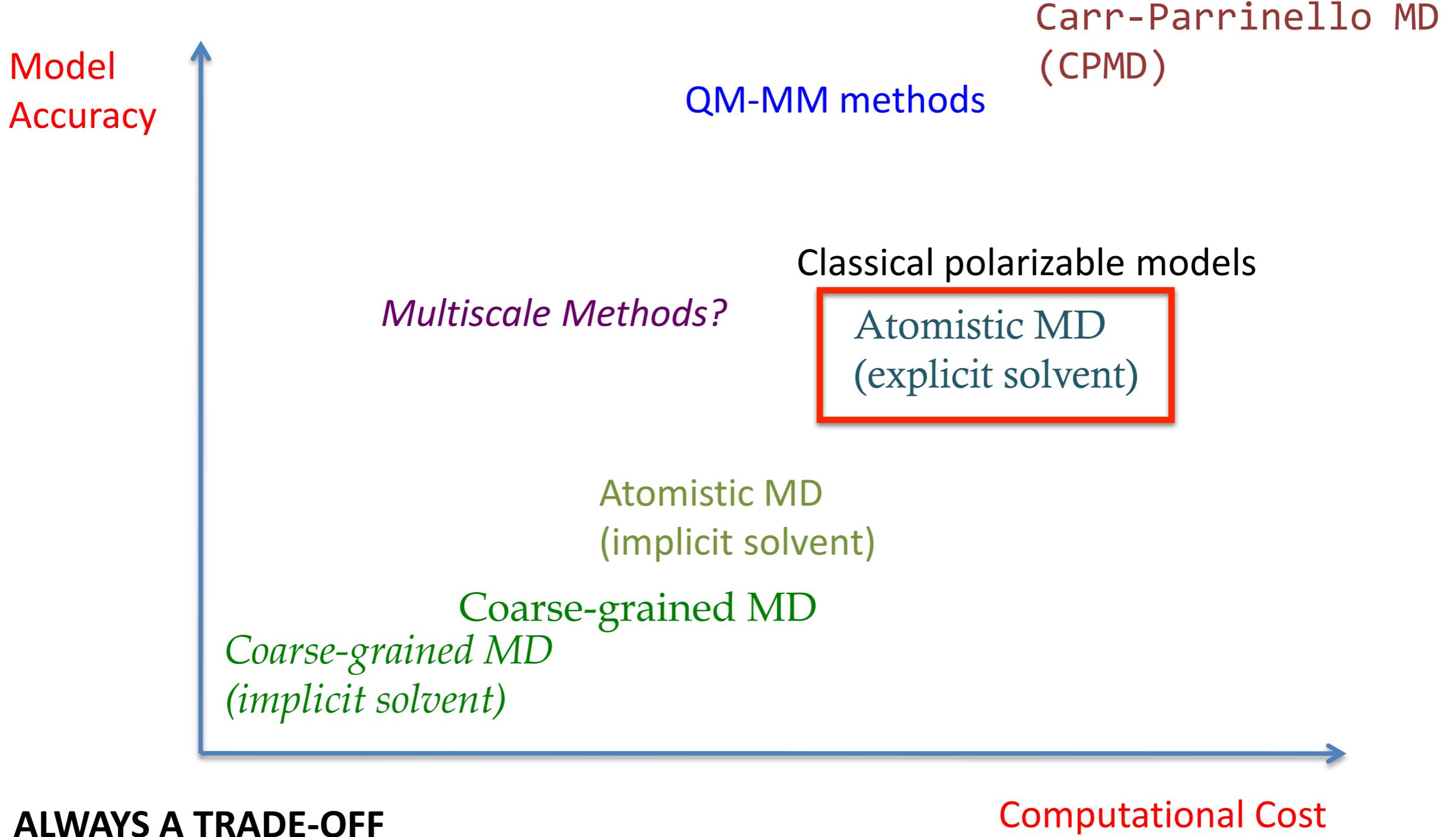


★ Slow diffusion





# Accuracy vs cost



How can we speed up MD simulations?

# Dimensionality reduction

It is often possible to describe a physical/chemical process in terms of a small number of coarse descriptors of the system:

$$\mathbf{S} = \mathbf{S}(\mathbf{R}) = (S_1(\mathbf{R}), \dots, S_d(\mathbf{R}))$$

Key quantity of thermodynamics is the free energy as a function of these variables:

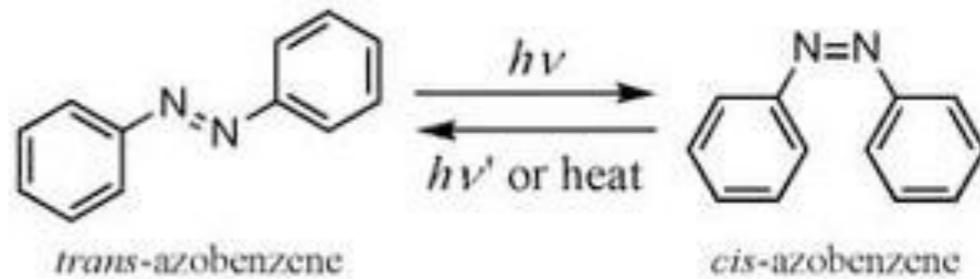
$$F(\mathbf{S}) = -\frac{1}{\beta} \ln P(\mathbf{S}) \quad \text{where} \quad \beta = \frac{1}{k_B T}$$

canonical  
ensemble

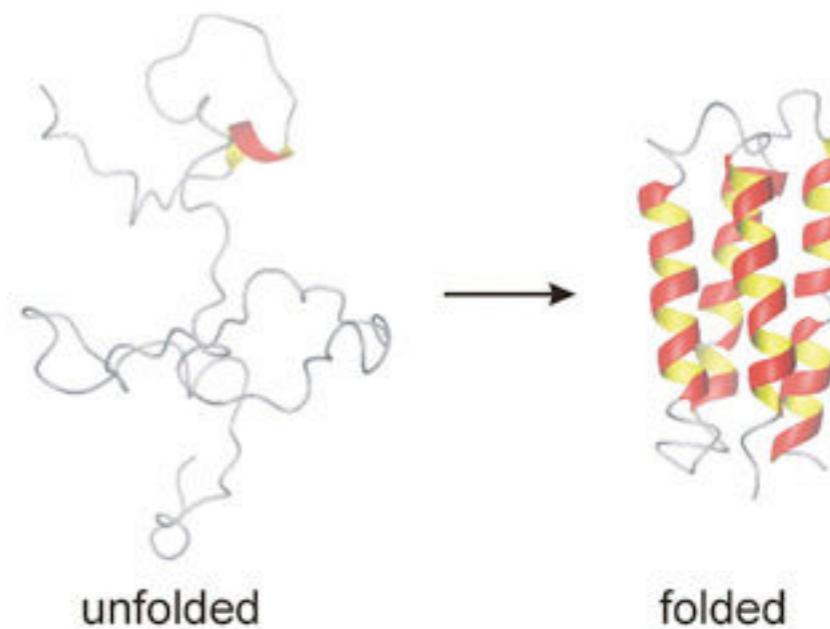
$$P(\mathbf{S}) = \frac{\int d\mathbf{R} \delta(\mathbf{S} - \mathbf{S}(\mathbf{R})) e^{-\beta U(\mathbf{R})}}{\int d\mathbf{R} e^{-\beta U(\mathbf{R})}}$$

# Examples

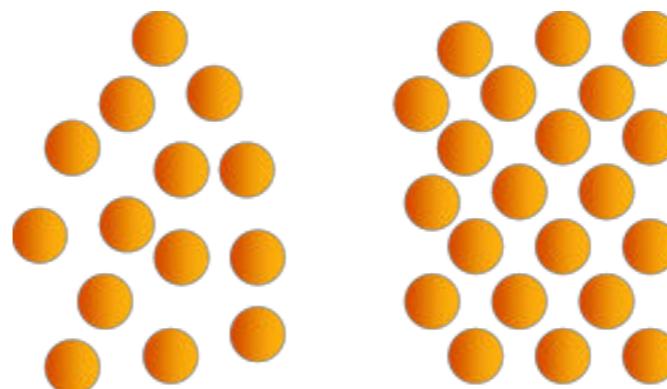
Isomerization:  
dihedral angle



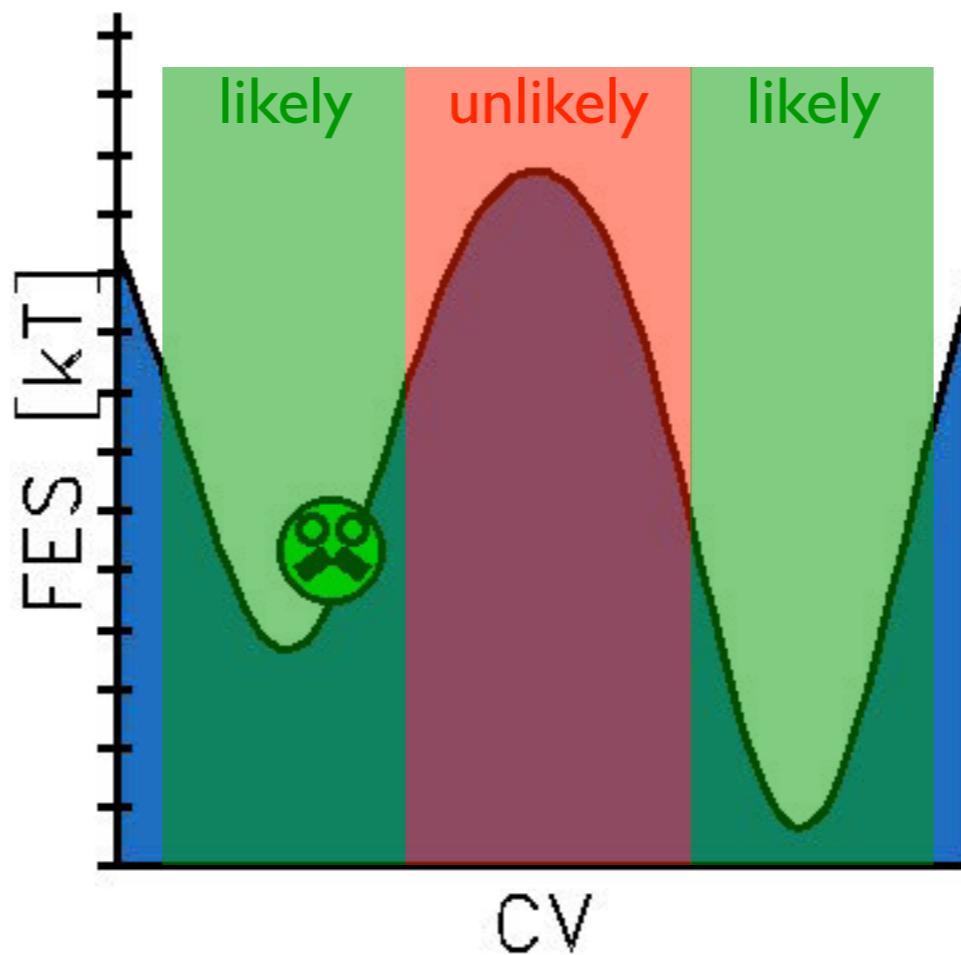
Protein folding:  
gyration radius,  
number of contacts,  
...



Phase transitions:  
lattice vectors,  
bond order parameters,  
...



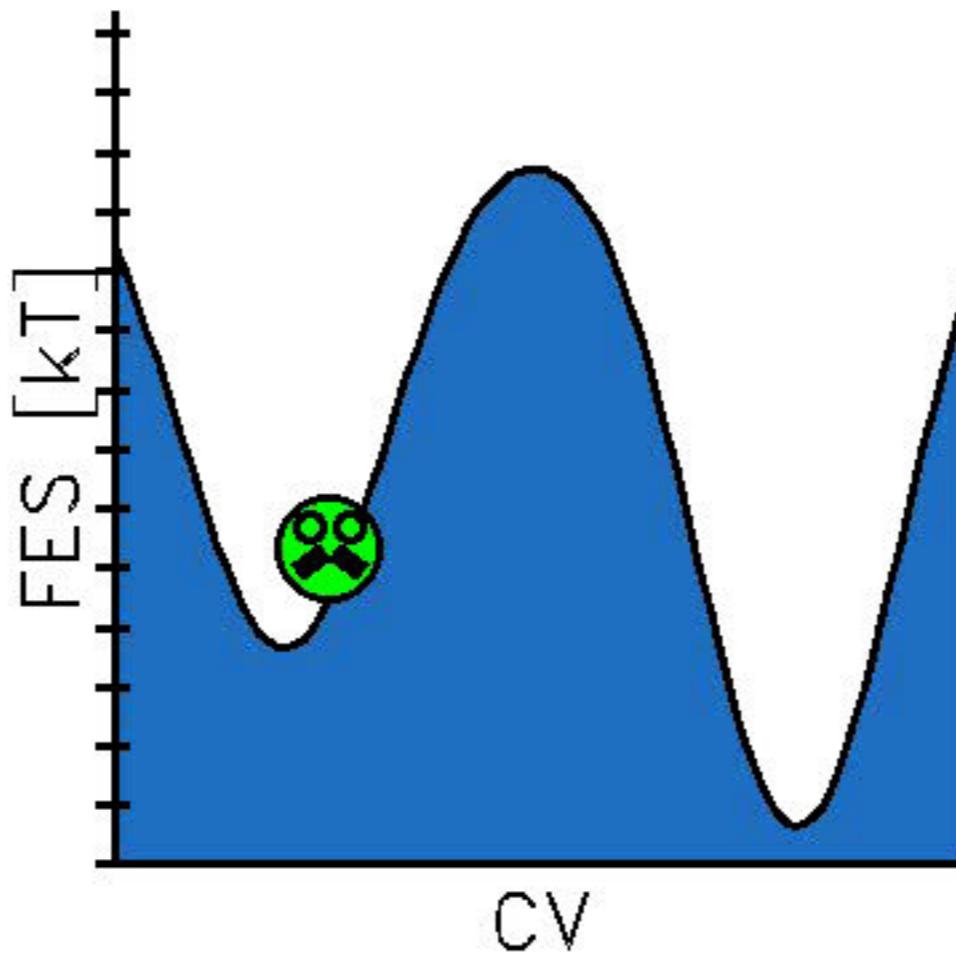
# Rare events simplified



How can we estimate a free energy difference if we never see a transition?

$$F(A) - F(B) = -k_B T \ln \frac{N_A}{N_B}$$

# Biased sampling



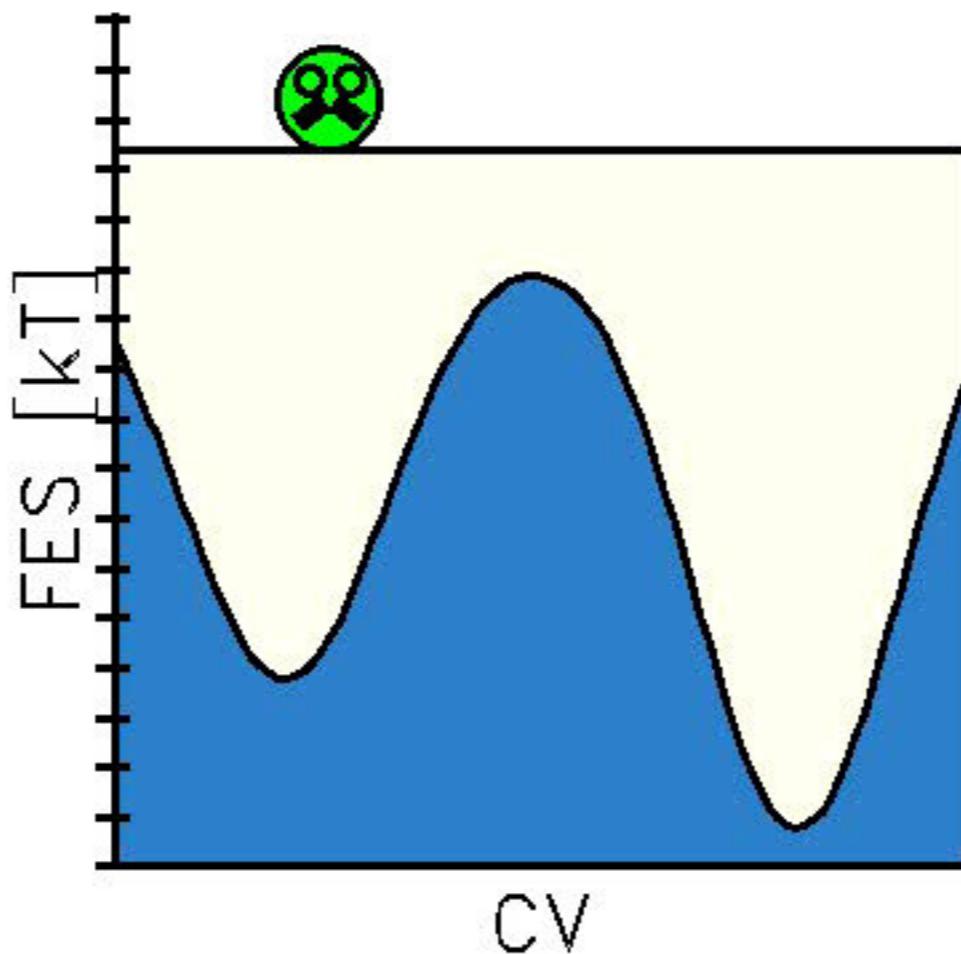
The idea is to add a bias potential that acts on the collective variables:

$$U(\mathbf{R}) \rightarrow U(\mathbf{R}) + V(\mathbf{S}(\mathbf{R}))$$

What is a good choice of bias potential?

# Umbrella sampling

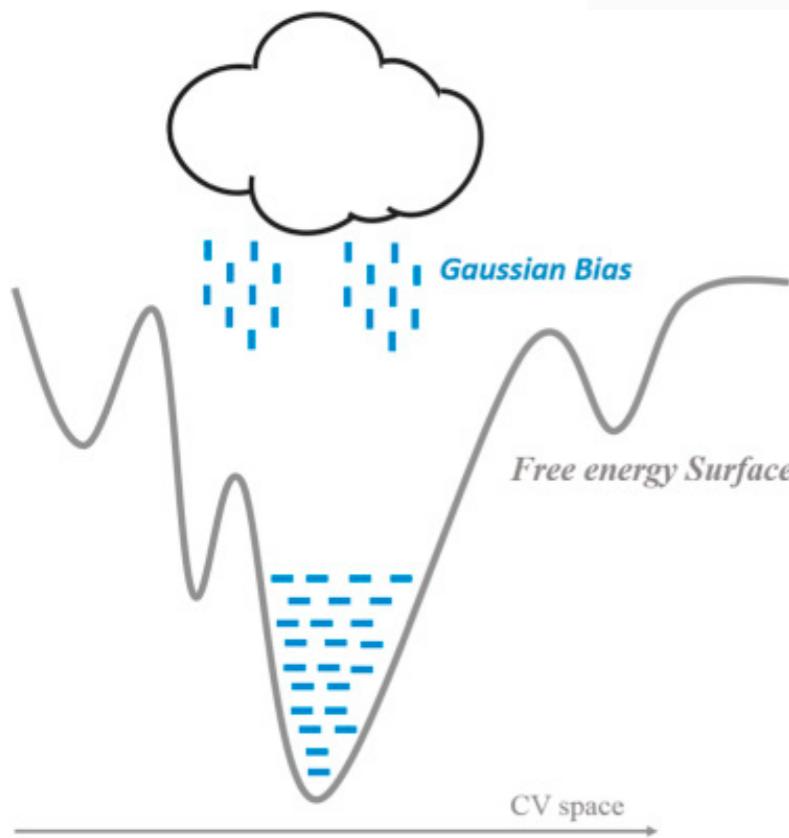
What is a good choice of bias potential?



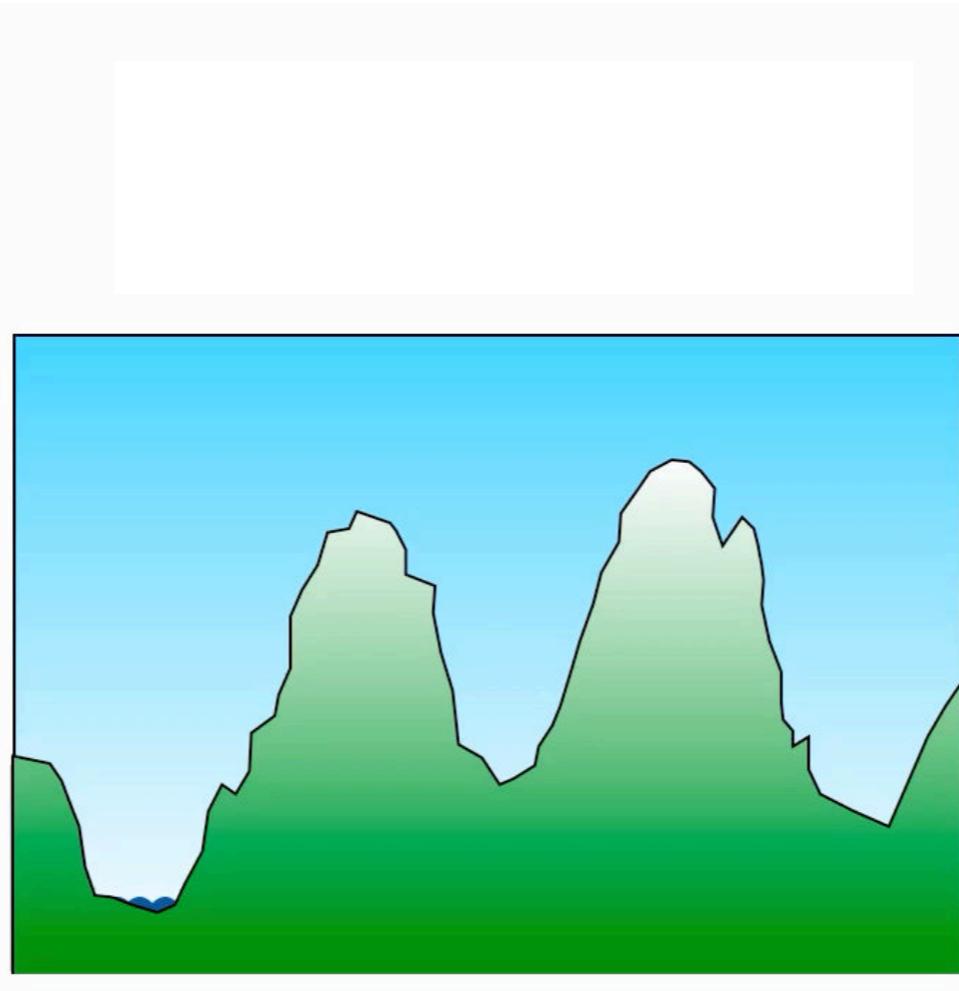
The one that leads to  $F'(\mathbf{S}) = 0 \rightarrow V(\mathbf{S}) = -F(\mathbf{S})$

Let's use an approximation of the free energy as bias potential

# Metadynamics: one of the most portrayed enhanced-sampling technique



The “cloud”  
metadynamics  
by Zheng et al.

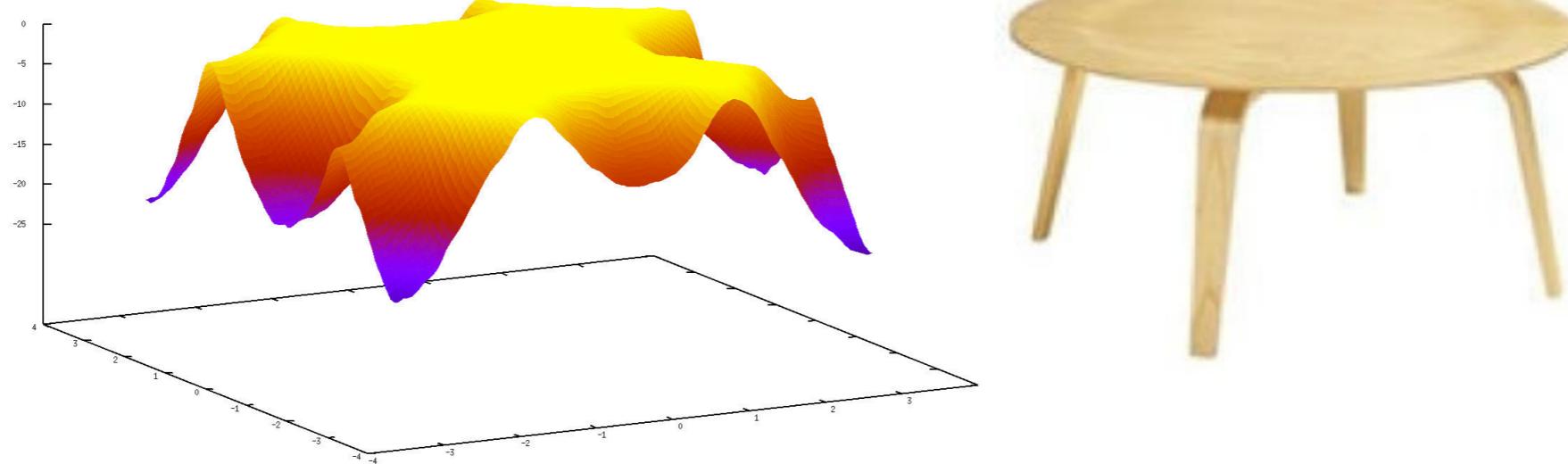


The “dutch”  
metadynamics  
by Bernd Ensing

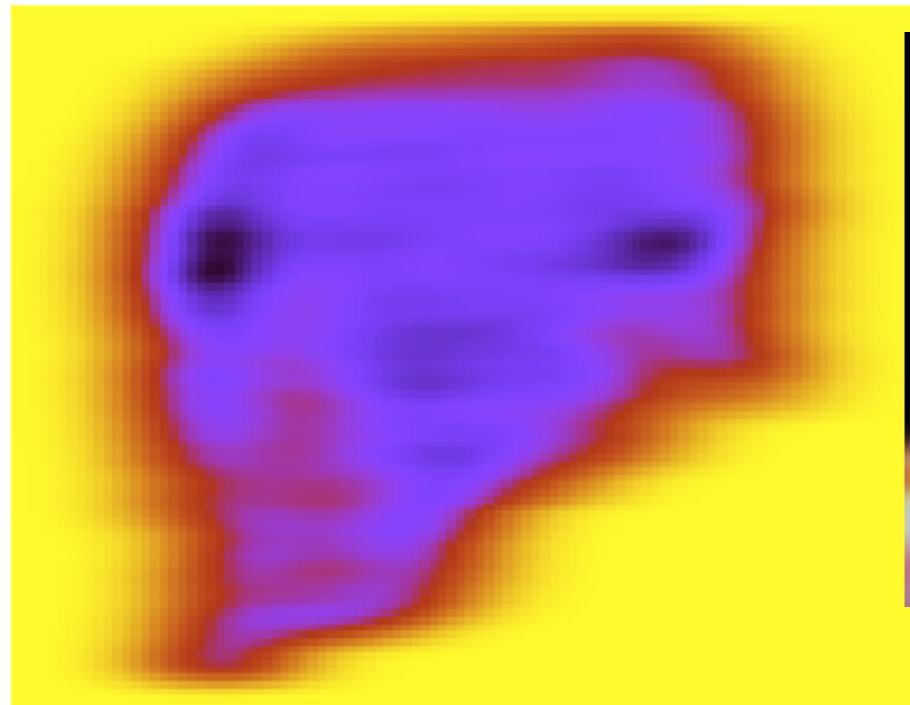


The “bird bath”  
metadynamics  
by Jim Pfaendtner

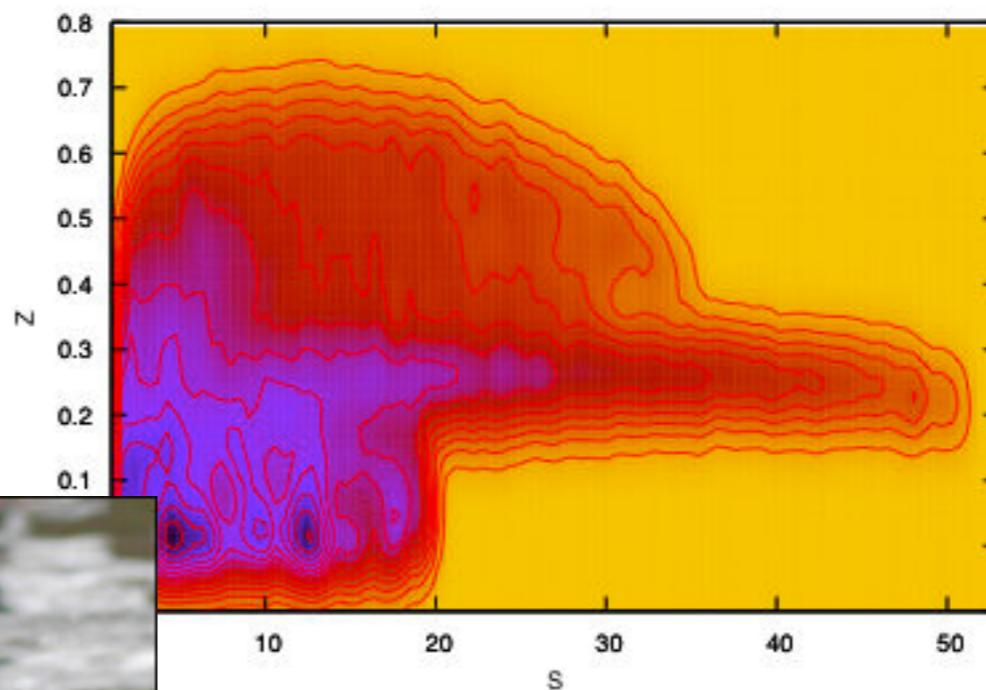
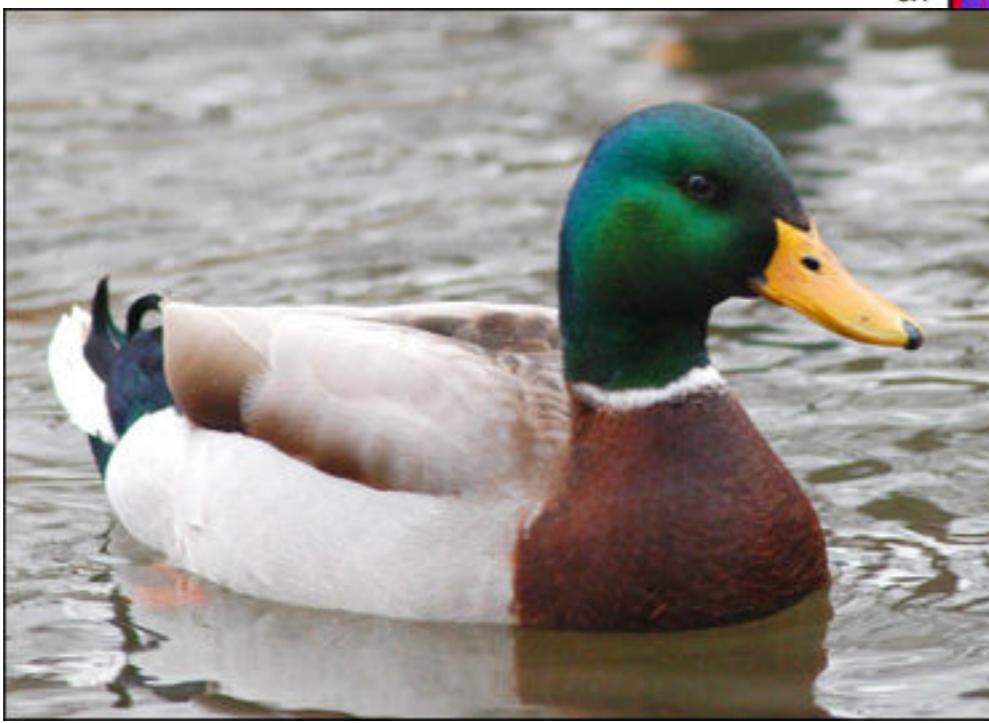
# Metadynamics: a method to create beautiful images for your Nature papers



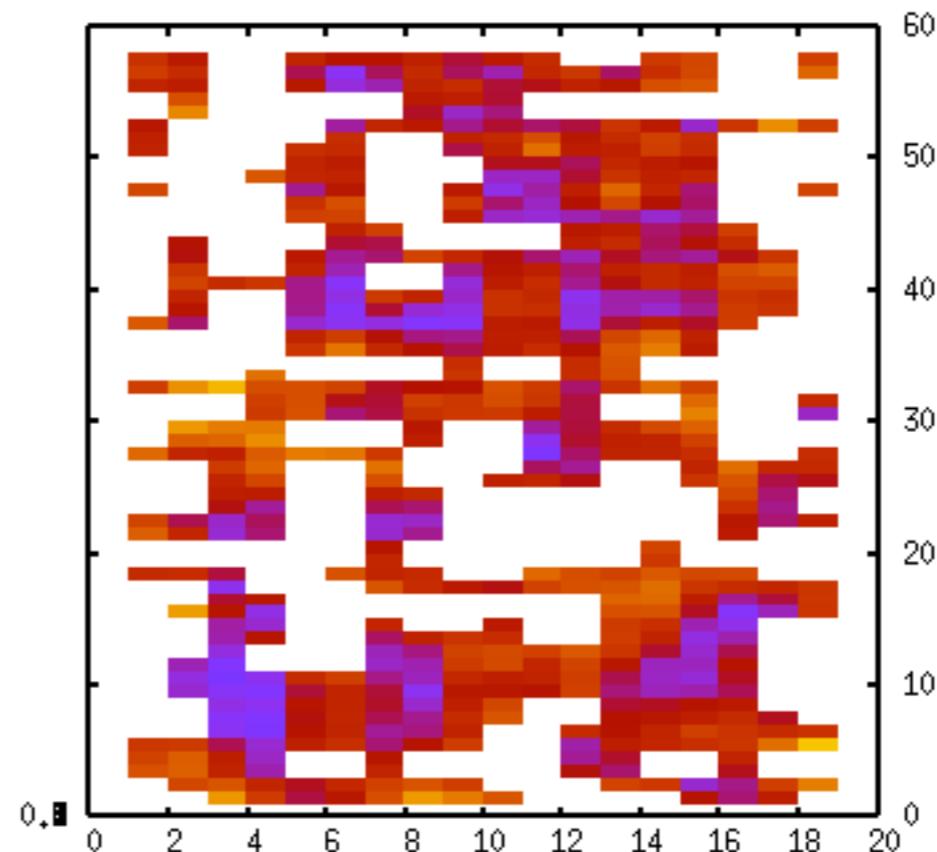
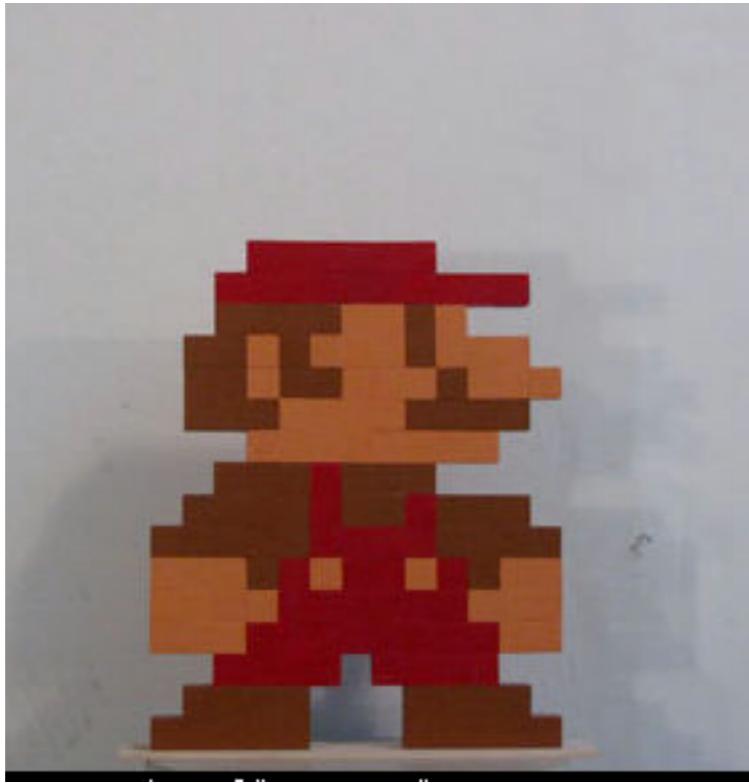
# Metadynamics: a method to create beautiful images for your Nature papers



# Metadynamics: a method to create beautiful images for your Nature papers



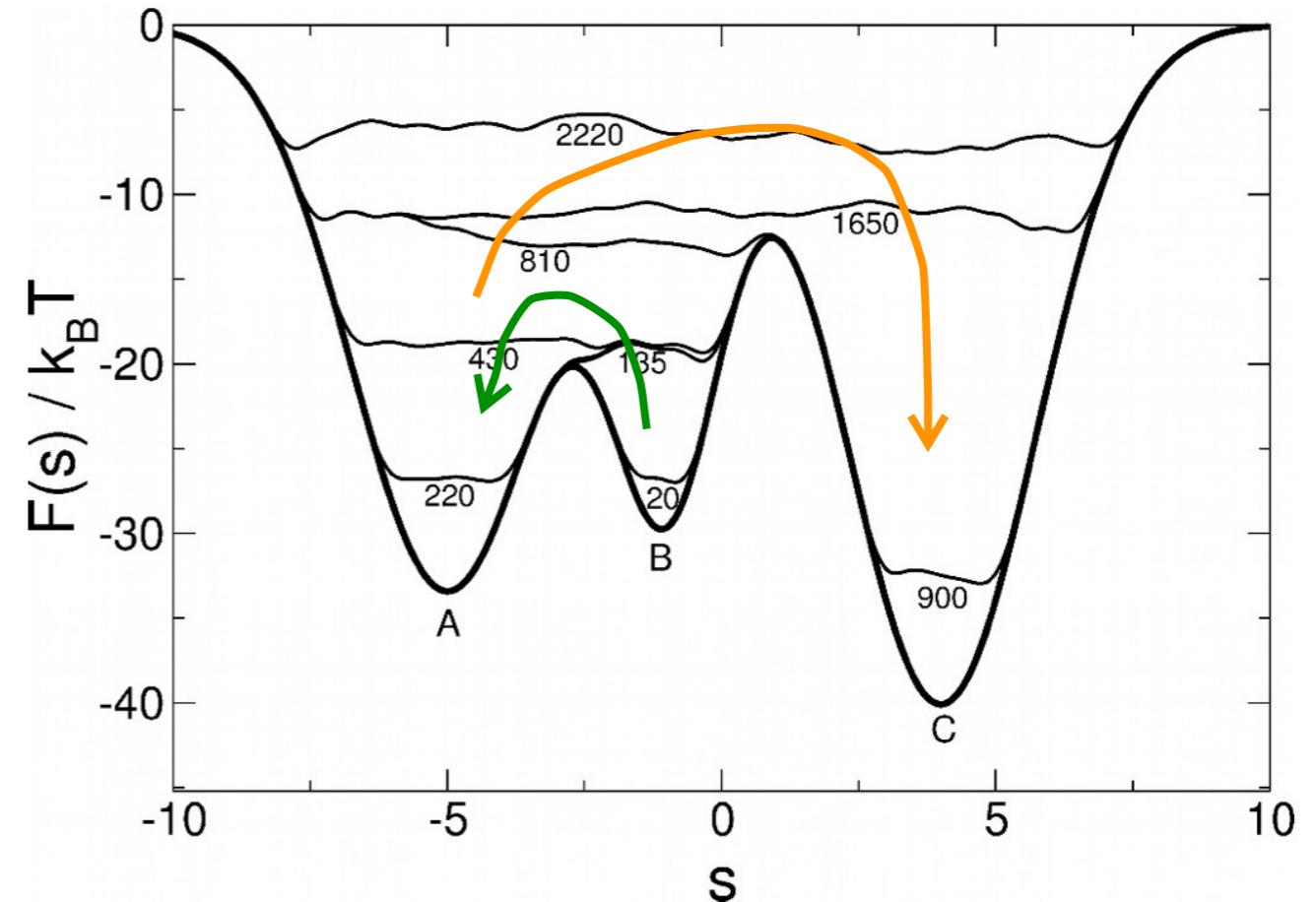
# Metadynamics: a method to create beautiful images for your Nature papers



# Metadynamics

History-dependent bias  
potential acting on selected  
degrees of freedom or  
Collective Variables (CVs)

$$\mathbf{S} = (S_1(\mathbf{R}), \dots, S_d(\mathbf{R}))$$



$$V_G(\mathbf{S}, t) = W \sum_{t'=\tau_G, 2\tau_G, \dots}^{t' < t} \exp \left( - \sum_{i=1}^d \frac{(S_i - S_i(\mathbf{R}(t')))^2}{2\sigma_i^2} \right)$$

Laio & Parrinello PNAS 2002

REVIEW: Barducci, Bonomi, Parrinello WIREs Comput Mol Sci 2011

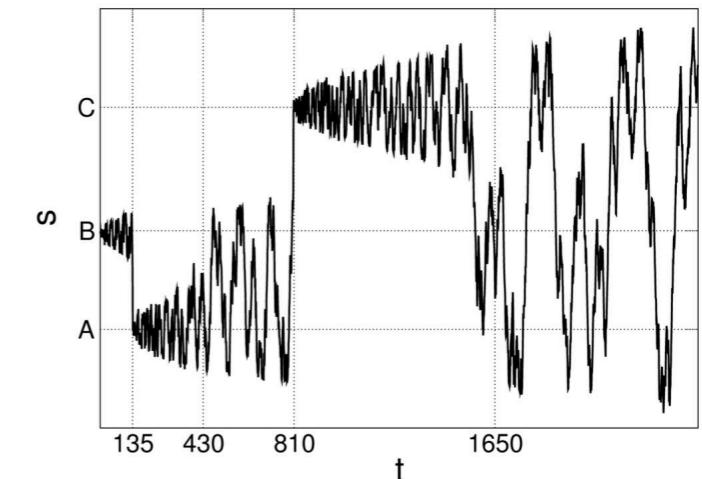
# Pros and Cons



## Advantages

- Enhanced sampling along the CVs
- Reconstruction of the FES:

$$V_G(S, t \rightarrow \infty) = -F(S) + C$$



Bussi, Laio, Parrinello PRL (2006)

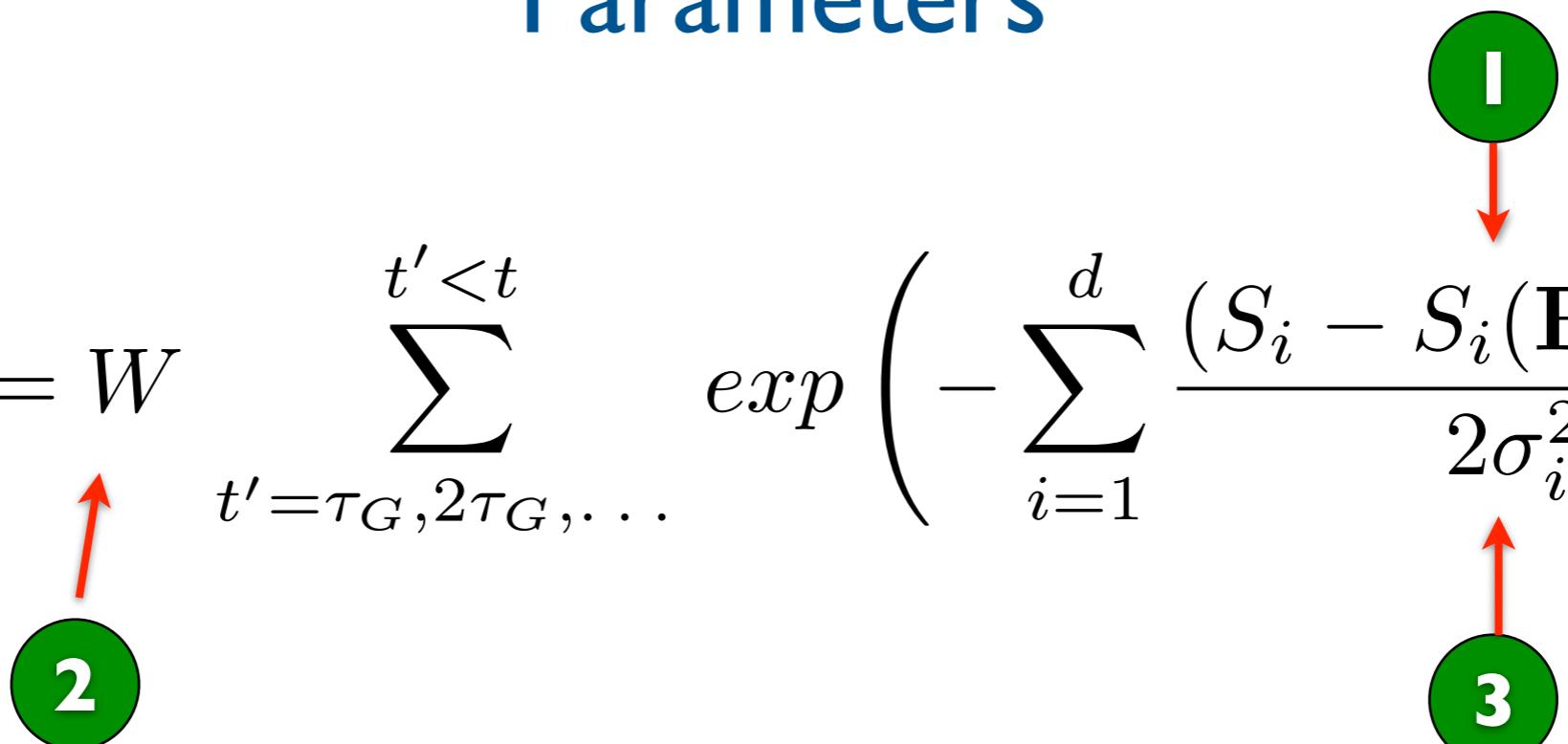
- *A priori* knowledge of the landscape not required



## Disadvantages

- Lack of convergence in a single run
- Overfilling
- The choice of the CVs is not trivial

# Parameters

$$V_G(\mathbf{S}, t) = W \sum_{t'=\tau_G, 2\tau_G, \dots}^{t' < t} \exp \left( - \sum_{i=1}^d \frac{(S_i - S_i(\mathbf{R}(t')))^2}{2\sigma_i^2} \right)$$


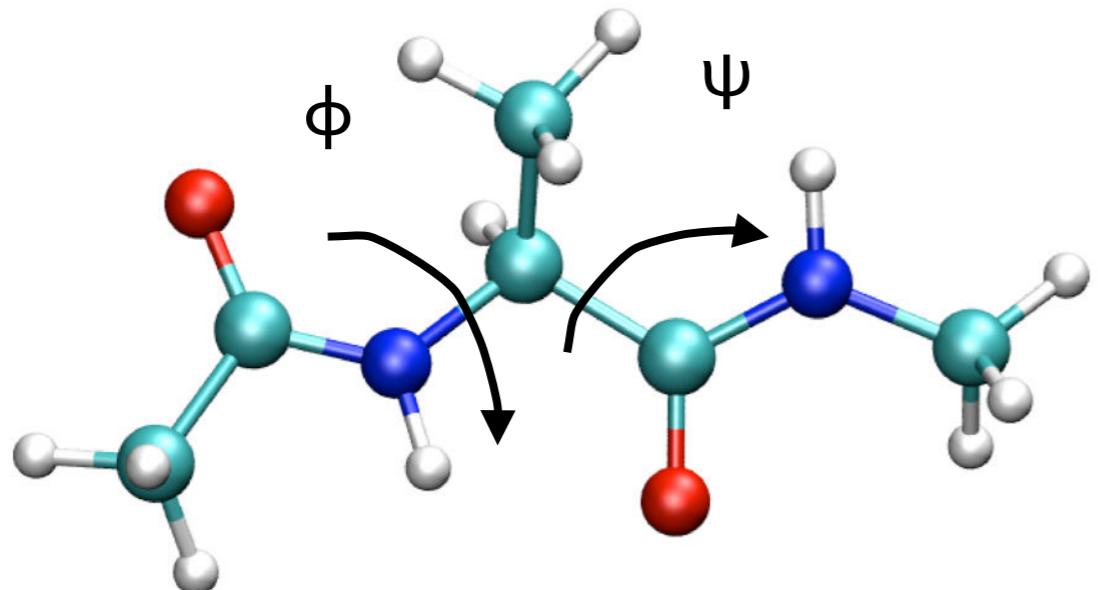
What do we need to choose?

- 1 Collective variables  $S = (S_1(\mathbf{R}), \dots, S_d(\mathbf{R}))$
- 2 Energy rate  $\omega = \frac{W}{\tau_G}$
- 3 Gaussian width

# MetaD with PLUMED

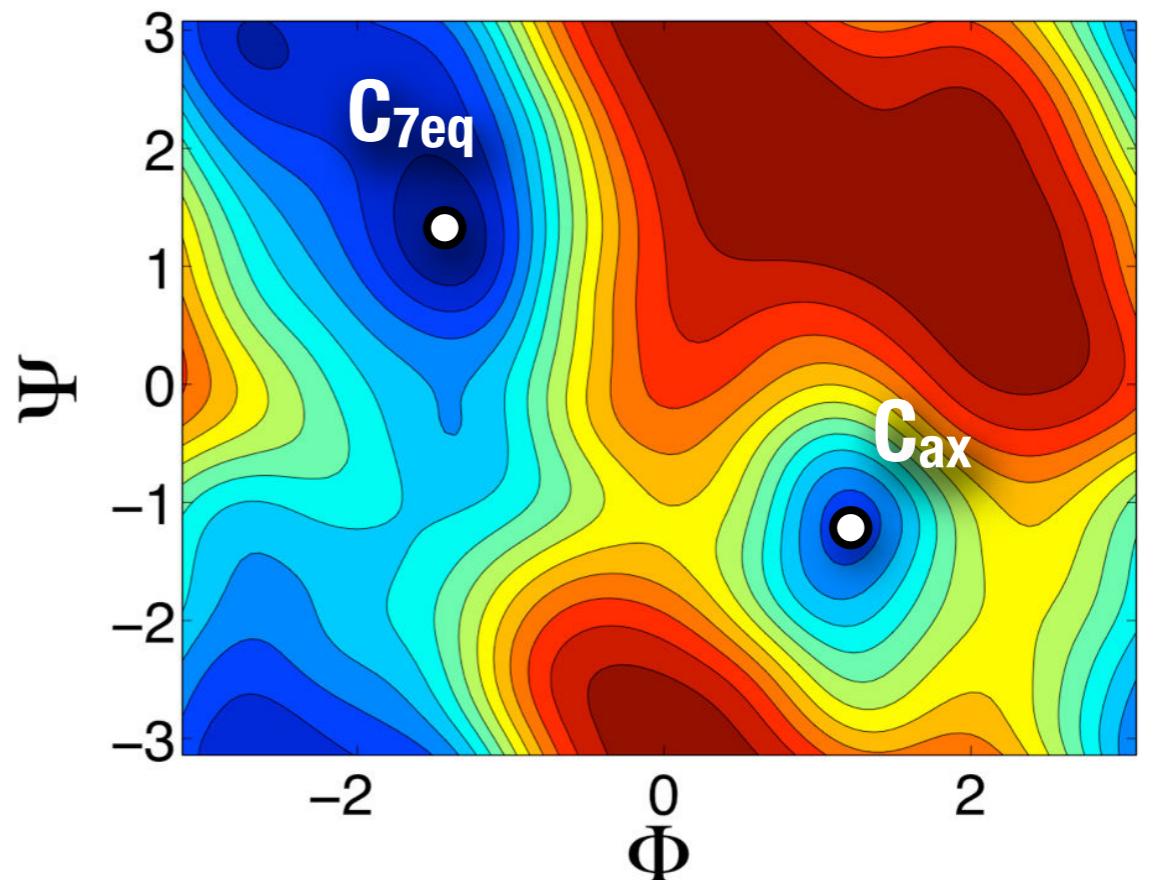
Alanine dipeptide in vacuo:

- two metastable states in the Ramachandran plot



$$(\Phi, \Psi)_{C_{7eq}} = (-1.45, 1.30)$$

$$(\Phi, \Psi)_{C_{ax}} = (1.22, -1.22)$$



The isomerization is a rare event: barrier of 6-8 kcal/mol at 300K  
(CHARMM27 force field)

# MetaD with PLUMED

```
# declare collective variables
phi: TORSION ATOMS=5,7,9,15
psi: TORSION ATOMS=7,9,15,17

# activate metadynamics
metad: METAD ARG=phi,psi HEIGHT=1.2 SIGMA=0.35,0.35 PACE=500

# print out collective variables and metadynamics bias
PRINT ARG=phi,psi,metad.bias STRIDE=100 FILE=COLVAR
```

## Units of measure:

- Height in kJoule/mol
- Stride in time steps
- Sigma in CV units

# MetaD with PLUMED

The main output is the HILLS file

```
#! FIELDS time phi psi sigma_phi sigma_psi height biasf  
#! SET multivariate false  
#! SET min_phi -pi  
#! SET max_phi pi  
#! SET min_psi -pi  
#! SET max_psi pi
```

HEADER

1.0000	1.1087	-0.8443	0.3500	0.3500	1.2000	-1.0000
2.0000	0.8819	-0.4241	0.3500	0.3500	1.2000	-1.0000
3.0000	1.2015	-0.2560	0.3500	0.3500	1.2000	-1.0000
4.0000	1.0359	-0.1549	0.3500	0.3500	1.2000	-1.0000
5.0000	0.9848	-0.6522	0.3500	0.3500	1.2000	-1.0000
6.0000	0.9970	-1.2823	0.3500	0.3500	1.2000	-1.0000
7.0000	0.8488	-1.0816	0.3500	0.3500	1.2000	-1.0000
8.0000	0.6605	-1.2918	0.3500	0.3500	1.2000	-1.0000
9.0000	1.1637	-1.4699	0.3500	0.3500	1.2000	-1.0000
10.0000	1.0504	-0.6997	0.3500	0.3500	1.2000	-1.0000
11.0000	1.0553	-0.8905	0.3500	0.3500	1.2000	-1.0000
12.0000	0.9541	-0.4194	0.3500	0.3500	1.2000	-1.0000

Time (ps)	CV (rad)	Sigma (rad)	Height (kJ/mol)	Bias Factor
--------------	-------------	----------------	--------------------	----------------

# Well-Tempered Metadynamics

Idea. New form for the bias potential:

$$V(\mathbf{S}, t) = k_B \Delta T \ln \left( 1 + \frac{\omega N(\mathbf{S}, t)}{k_B \Delta T} \right)$$

histogram  
collected during  
the simulation

“CV temperature”

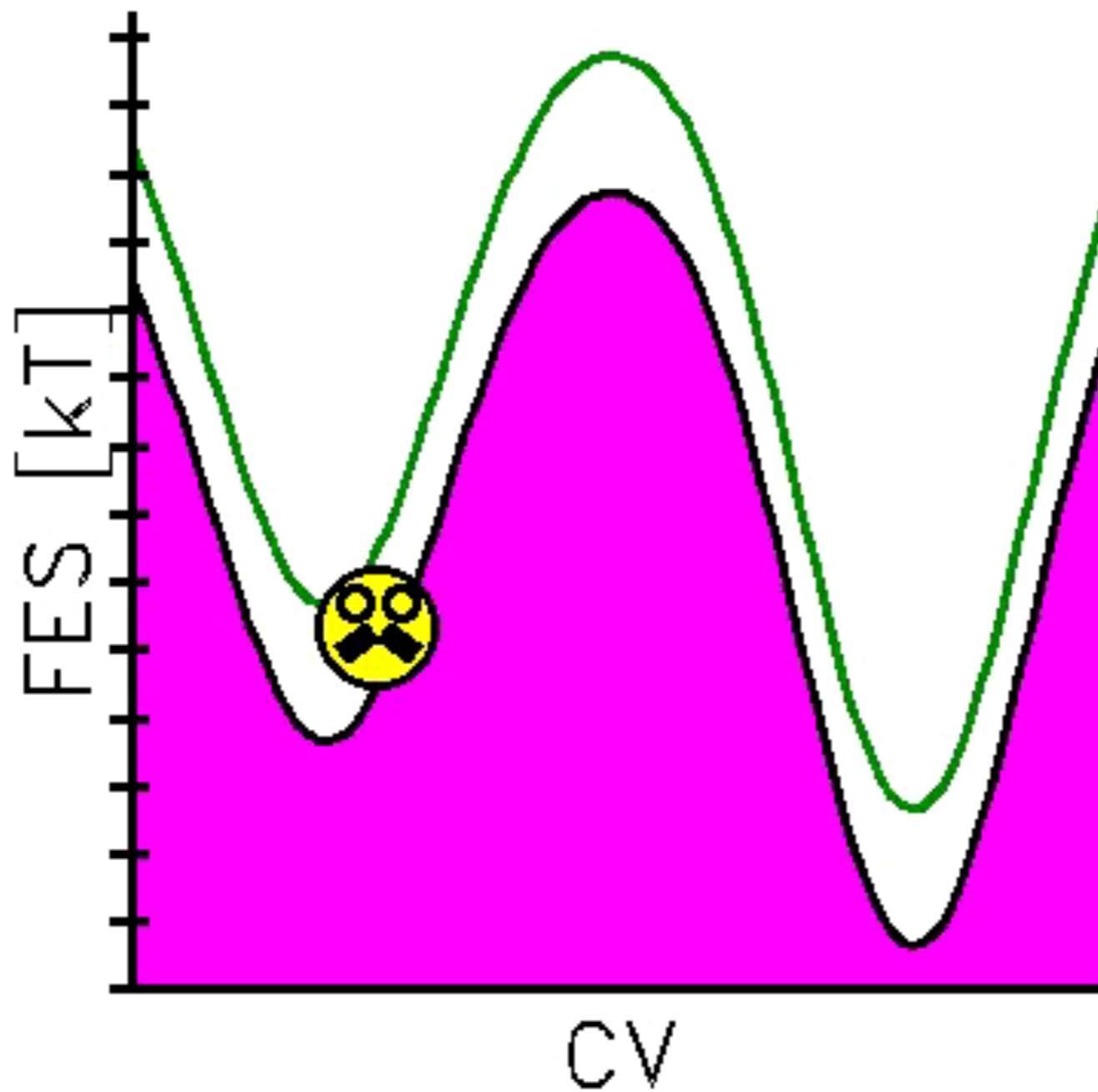
$$\dot{V}(\mathbf{S}, t) = \frac{\omega \delta_{\mathbf{S}, \mathbf{S}(t)}}{1 + \frac{\omega N(\mathbf{S}, t)}{k_B \Delta T}} = \omega e^{-\frac{V(\mathbf{S}, t)}{k_B \Delta T}} \delta_{\mathbf{S}, \mathbf{S}(t)}$$

Implementation.

Rescale the Gaussian height:

$$W = \omega \tau_G e^{-\frac{V_G(\mathbf{S}, t)}{k_B \Delta T}}$$

# Well-Tempered Metadynamics



movie by G. Bussi

- Single run convergence:

$$V_G(\mathbf{S}, t \rightarrow \infty) = -\frac{\Delta T}{T + \Delta T} F(\mathbf{S}) + C$$

- Overfilling issue solved.  $\Delta T$  to limit the exploration to the physically relevant regions.
- Error goes to zero in a single run:  $\epsilon(t) \propto 1/\sqrt{t}$
- The dynamics of all the microscopic variables becomes closer to equilibrium as the simulation proceeds:

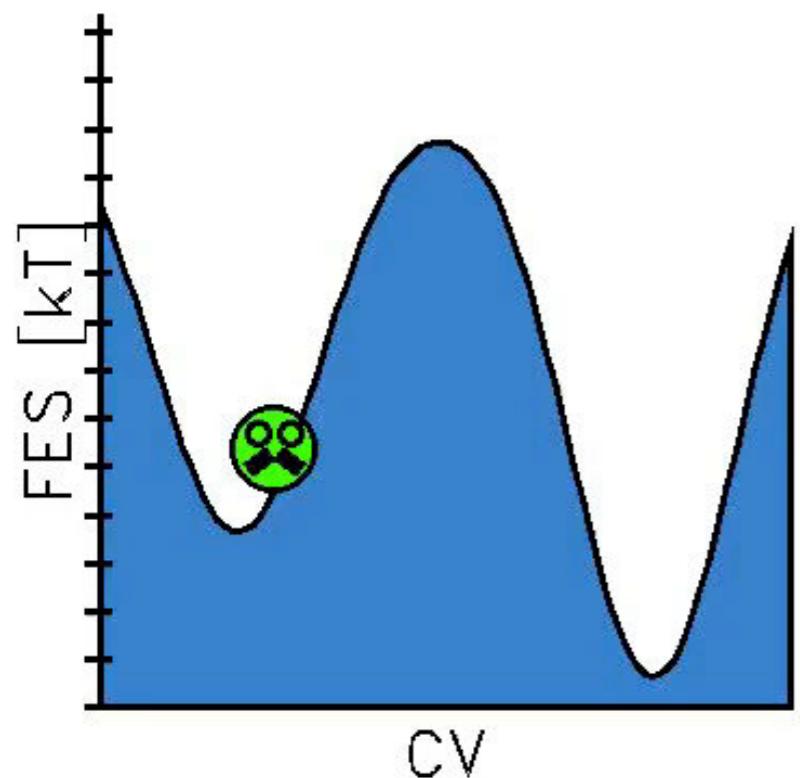
$$\dot{V}_G(\mathbf{S}, t \rightarrow \infty) \propto 1/t \rightarrow 0$$

- Distribution of variables other than the CVs can be easily reconstructed

Bonomi, Barducci, Parrinello *J. Comp. Chem.* (2009)  
 Branduardi, Bussi, Parrinello *JCTC* (2012)  
 Tiwary, Parrinello *JPCB* (2015)

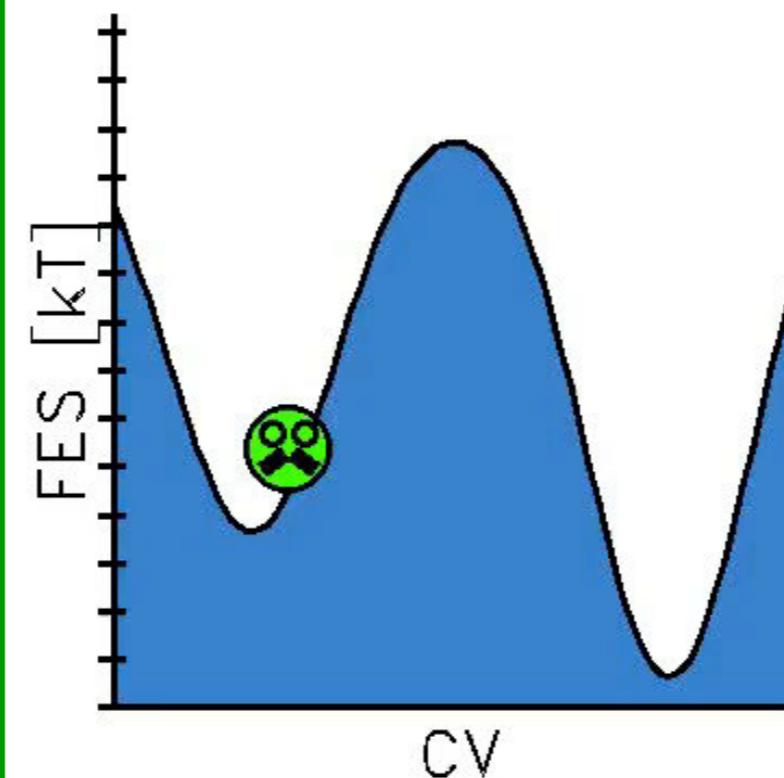
# An optimal recipe

TOO SLOW



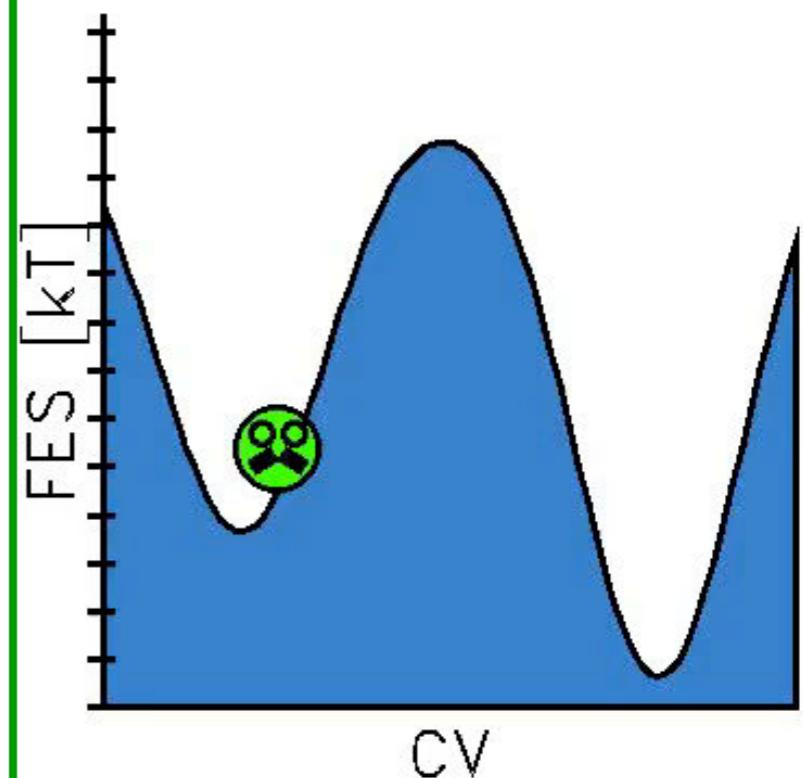
$$w(t) \propto \frac{1}{\log t}$$

OPTIMAL



$$w(t) \propto \frac{1}{t}$$

TOO FAST



$$w(t) \propto \frac{1}{t^2}$$

# Parameters

$$W = \omega \tau_G e^{-\frac{V_G(S,t)}{k_B \Delta T}}$$


1

2

1 Initial Gaussian height + deposition stride

2 How fast the amount of bias added decreases to zero

$$\left\{ \begin{array}{l} T + \Delta T \\ \frac{T + \Delta T}{T} \end{array} \right.$$

CV  
TEMPERATURE

BIAS  
FACTOR

# WT-MetaD with PLUMED

```
# declare collective variables
phi: TORSION ATOMS=5,7,9,15
psi: TORSION ATOMS=7,9,15,17

# activate well-tempered metadynamics
metad: METAD ARG=phi,psi HEIGHT=1.2 SIGMA=0.35,0.35 PACE=500 BIASFACTOR=8

# print out collective variables and metadynamics bias
PRINT ARG=phi,psi,metad.bias STRIDE=100 FILE=COLVAR
```

# WT-MetaD with PLUMED

The main output is the HILLS file

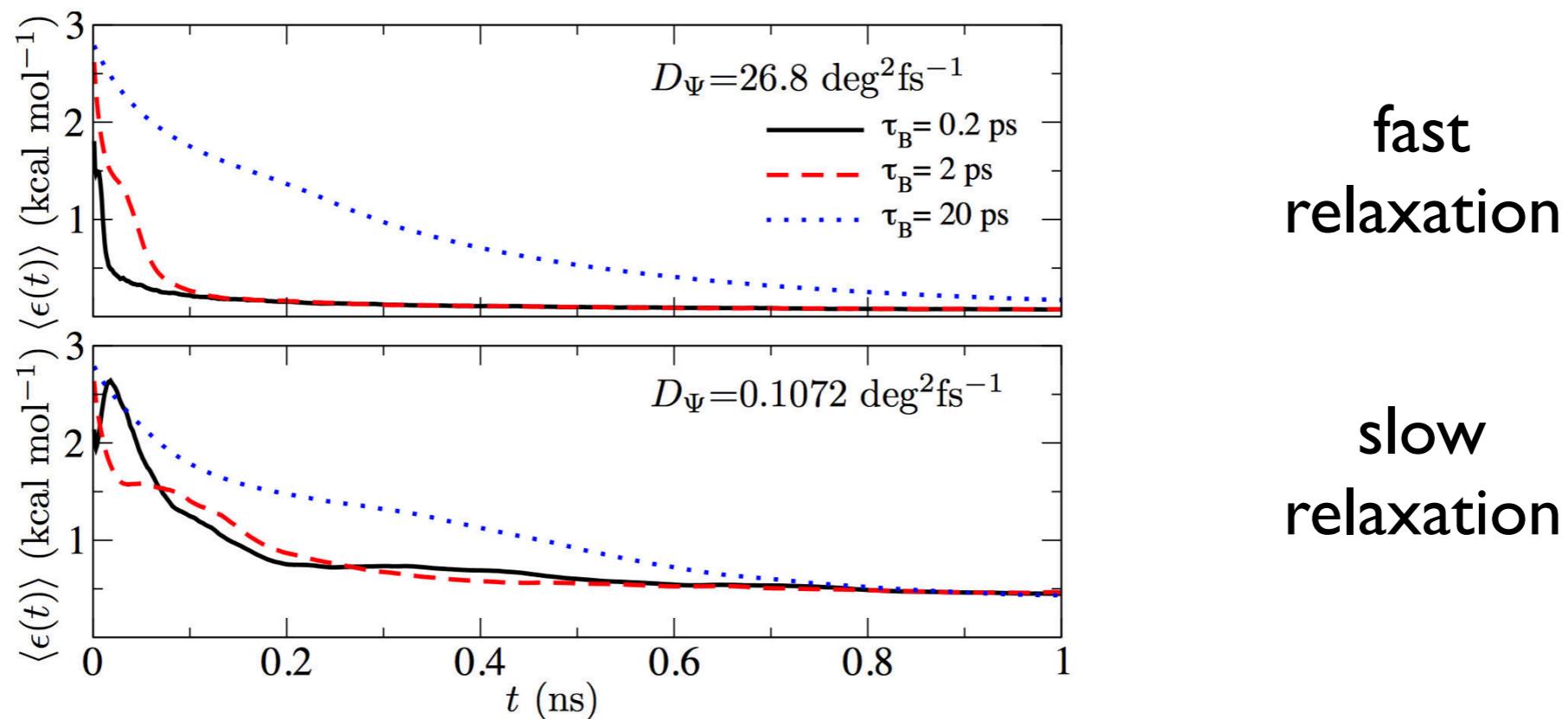
```
#! FIELDS time phi psi sigma_phi sigma_psi height biasf
#! SET multivariate false
#! SET min_phi -pi
#! SET max_phi pi
#! SET min_psi -pi
#! SET max_psi pi
```

1.0000	1.1087	-0.8443	0.3500	0.3500	1.3714	8.0000
2.0000	0.8819	-0.4241	0.3500	0.3500	1.3348	8.0000
3.0000	1.1897	-0.2275	0.3500	0.3500	1.3007	8.0000
4.0000	0.9427	-0.0746	0.3500	0.3500	1.2513	8.0000
5.0000	0.9916	-0.7114	0.3500	0.3500	1.1933	8.0000
6.0000	1.1558	-1.1060	0.3500	0.3500	1.2525	8.0000
7.0000	1.1366	-0.7487	0.3500	0.3500	1.0996	8.0000
8.0000	1.0933	-1.0890	0.3500	0.3500	1.1288	8.0000
9.0000	1.0794	-0.7793	0.3500	0.3500	1.0000	8.0000
10.0000	1.1538	-0.4934	0.3500	0.3500	1.0038	8.0000
11.0000	0.7173	-0.7677	0.3500	0.3500	1.0751	8.0000
12.0000	1.0673	-0.3857	0.3500	0.3500	0.9596	8.0000

Time (ps)	CV (rad)	Sigma (rad)	Height (kJ/mol)	Bias Factor
--------------	-------------	----------------	--------------------	----------------

$$V_G(\mathbf{S}, t \rightarrow \infty) = -\frac{\Delta T}{T + \Delta T} F(\mathbf{S}) + C$$

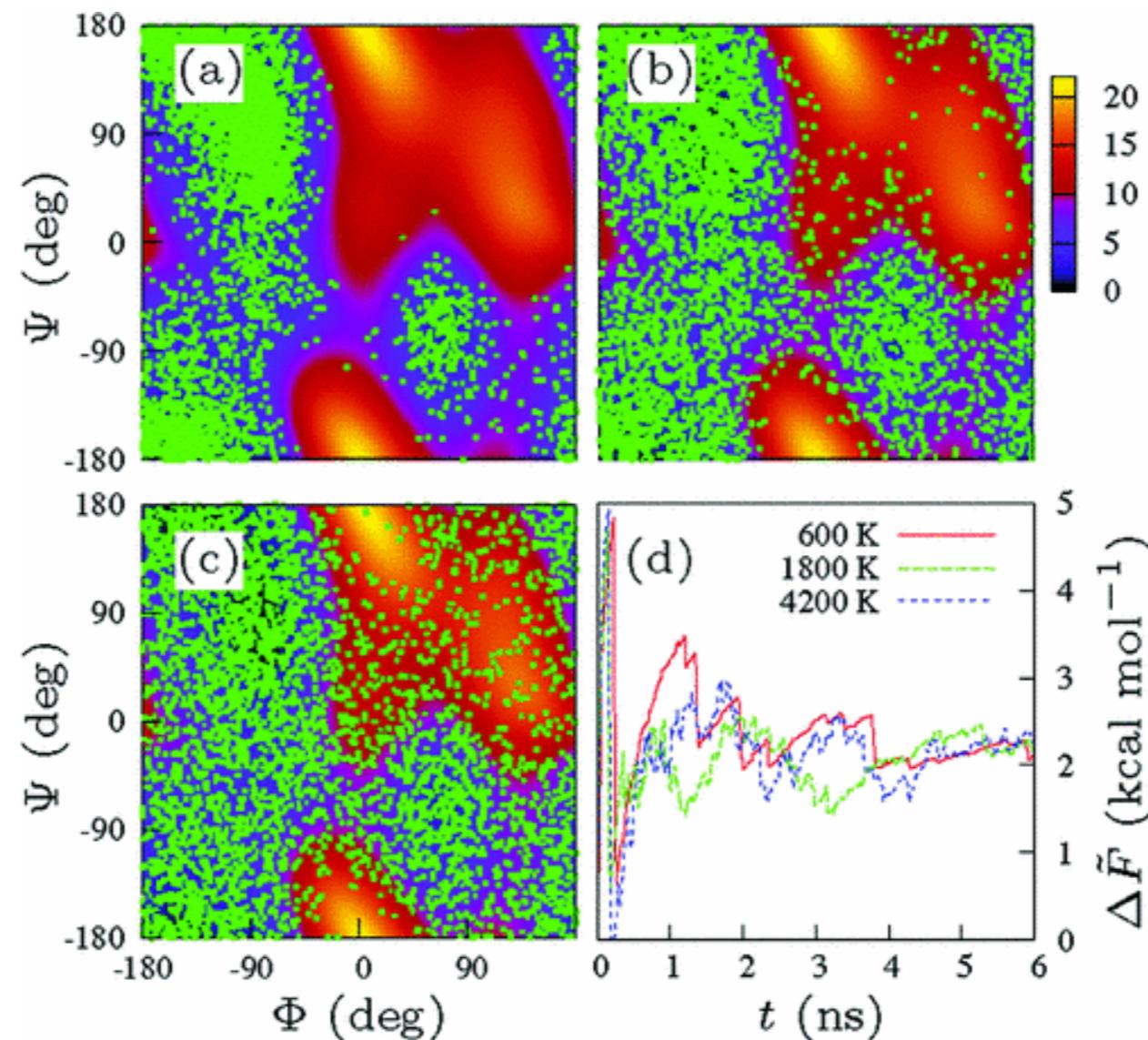
# Guidelines for choosing the initial deposition rate



The initial deposition rate affects the transient period

Typical value for biological applications  $\sim 0.5\text{-}1.0 \text{ k}_\text{B}\text{T}$  per ps

# Guidelines for choosing the biasfactor



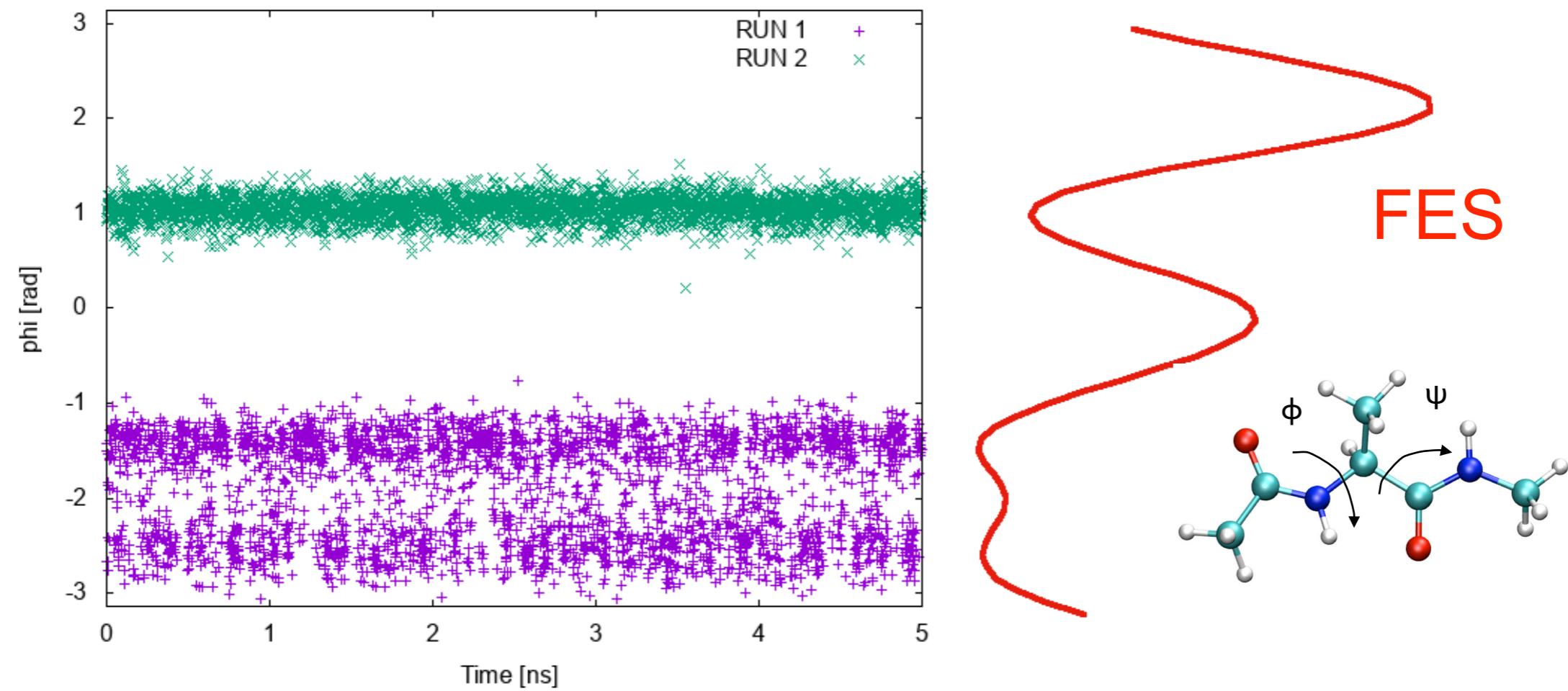
Determine the extent of free-energy exploration

Typical value for biological applications  $\sim 10$

# Guidelines for choosing sigma

The width of the Gaussian should be commensurate to the typical size of the “features” of the free energy landscape

Typically, the Gaussian width is estimated from the fluctuations of the collective variables in an unbiased run:



# Adaptive Gaussians

Branduardi, Bussi, Parrinello *JCTC* (2015)

Automatic way to determine Gaussian sigma that accounts for

- fluctuations varying across the free-energy landscape
- CVs being correlated

It uses multivariate, non-diagonal  
Gaussians:

$$\begin{aligned}\dot{V}(s, t) = \omega e^{-V(s(t), t)/\Delta T} \\ \times \exp\left(-\frac{1}{2} \sum_{ij} [s_i - s_i(t)] \sigma_{ij}^{-2} [s_j - s_j(t)]\right)\end{aligned}$$

with covariance-matrix elements determined by:

$$\bar{s}_i(t) = \frac{1}{\tau_D} \int_0^t dt' s_i(t') e^{-(t-t')/\tau_D}$$

and

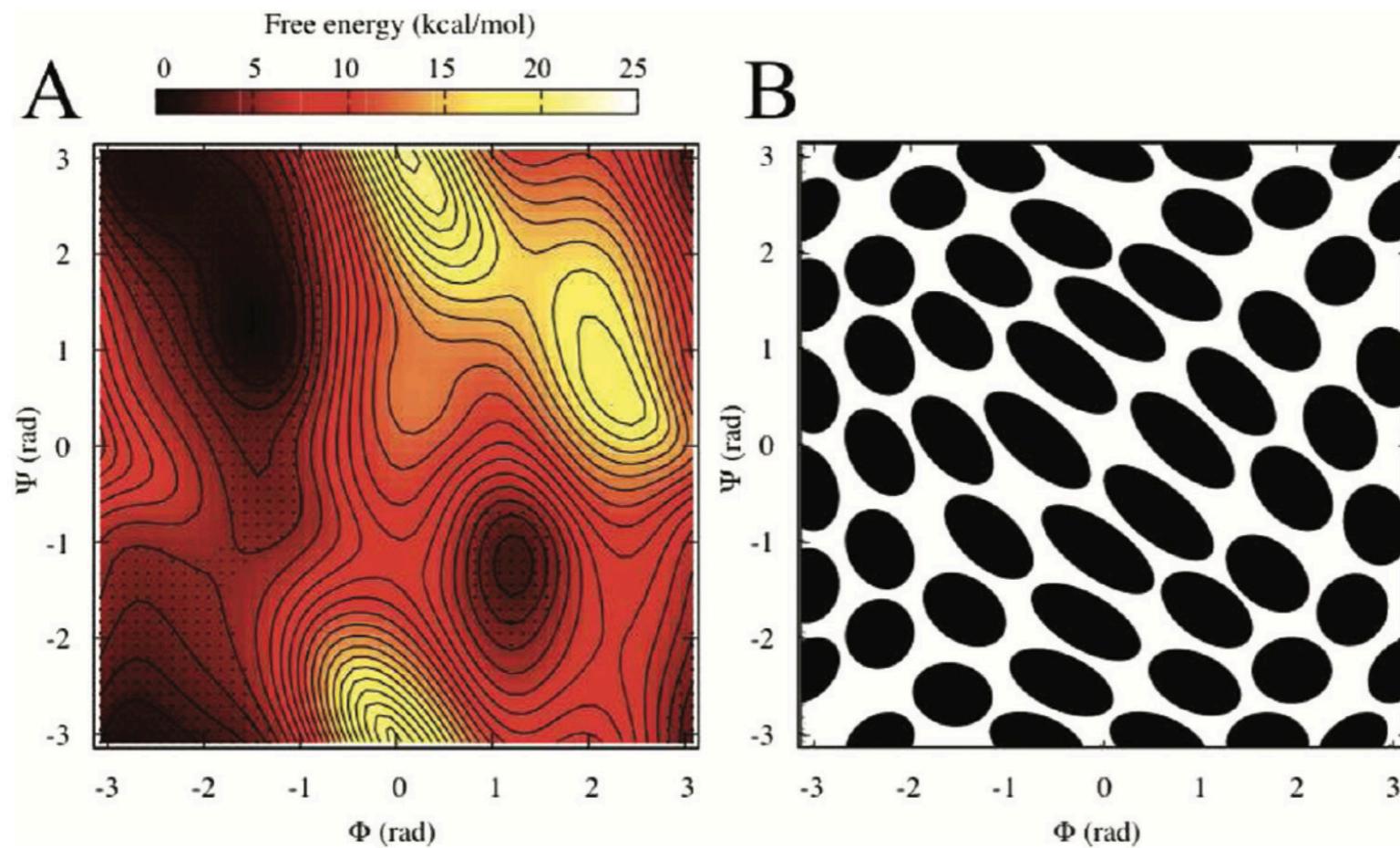
$$\begin{aligned}\sigma_{ij}^2(t) = \frac{1}{\tau_D} \int_0^t dt' [s_i(t') - \bar{s}_i(t')][s_j(t') - \bar{s}_j(t')] \\ \times e^{-(t-t')/\tau_D}\end{aligned}$$

dynamically-adapted

or

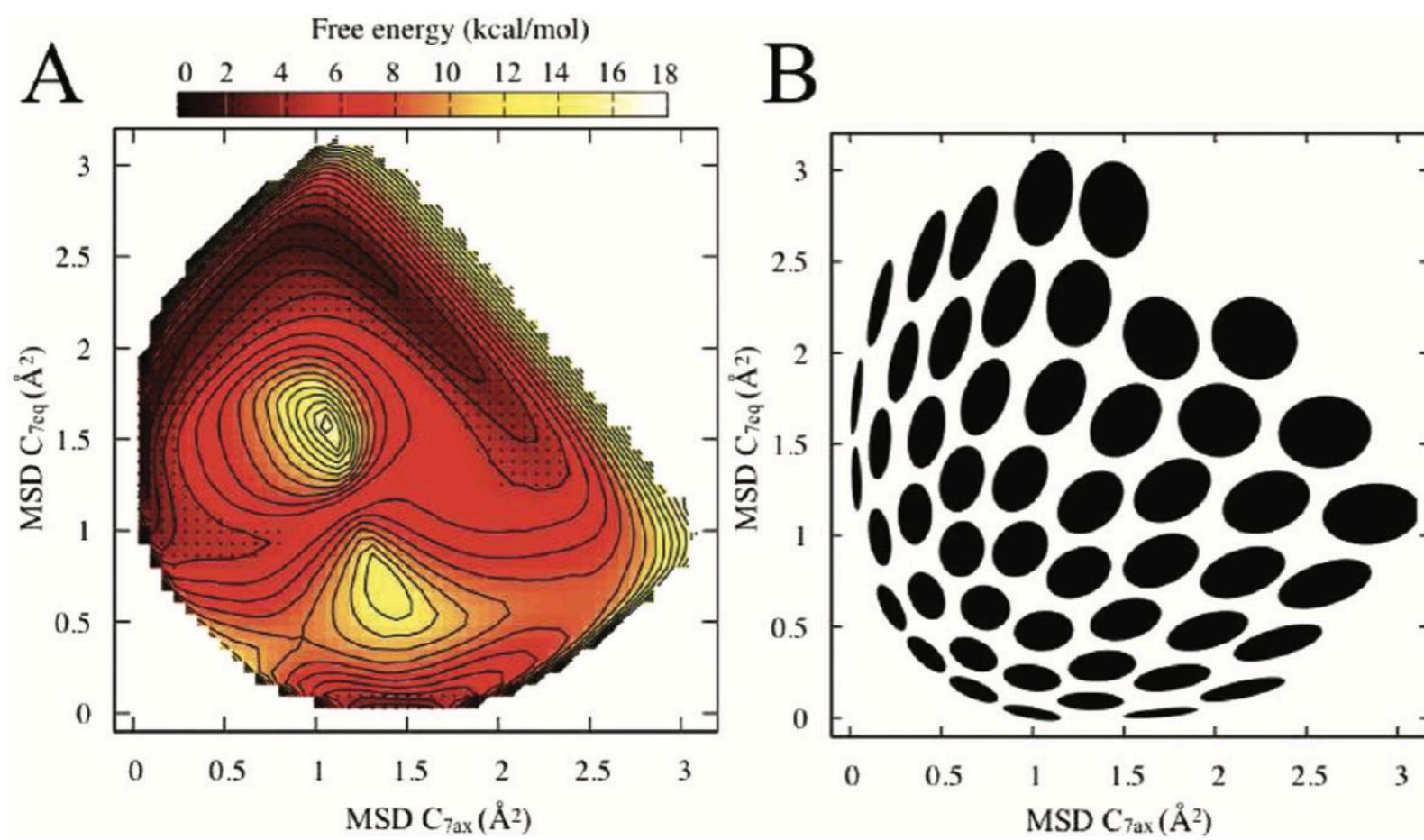
$$\sigma_{ij}^2(q) = \sigma_G^2 \sum_{\alpha} \frac{\partial s_i}{\partial q_{\alpha}} \frac{\partial s_j}{\partial q_{\alpha}}$$

geometrically-adapted



CVs

backbone  
dihedrals



MSD from  
local minima

# Guidelines for choosing the CVs

A good set of CVs for metadynamics (and other biasing techniques) should:

- Discriminate between relevant metastable states
- Include all the slow modes of a process
- Be as small as possible

Metadynamics is inefficient with a large number of CVs.



Possible strategies:

- design smart CVs
- devise automatic protocols to find good CVs
- improve metadynamics to deal with a large number of CVs
- couple metadynamics with other methods, such as REM

# How to calculate free-energies

## Free-energy reconstruction with SUM\_HILLS

$$V(\mathbf{S}, t \rightarrow \infty) = -\frac{\Delta T}{T + \Delta T} F(\mathbf{S}) + C$$

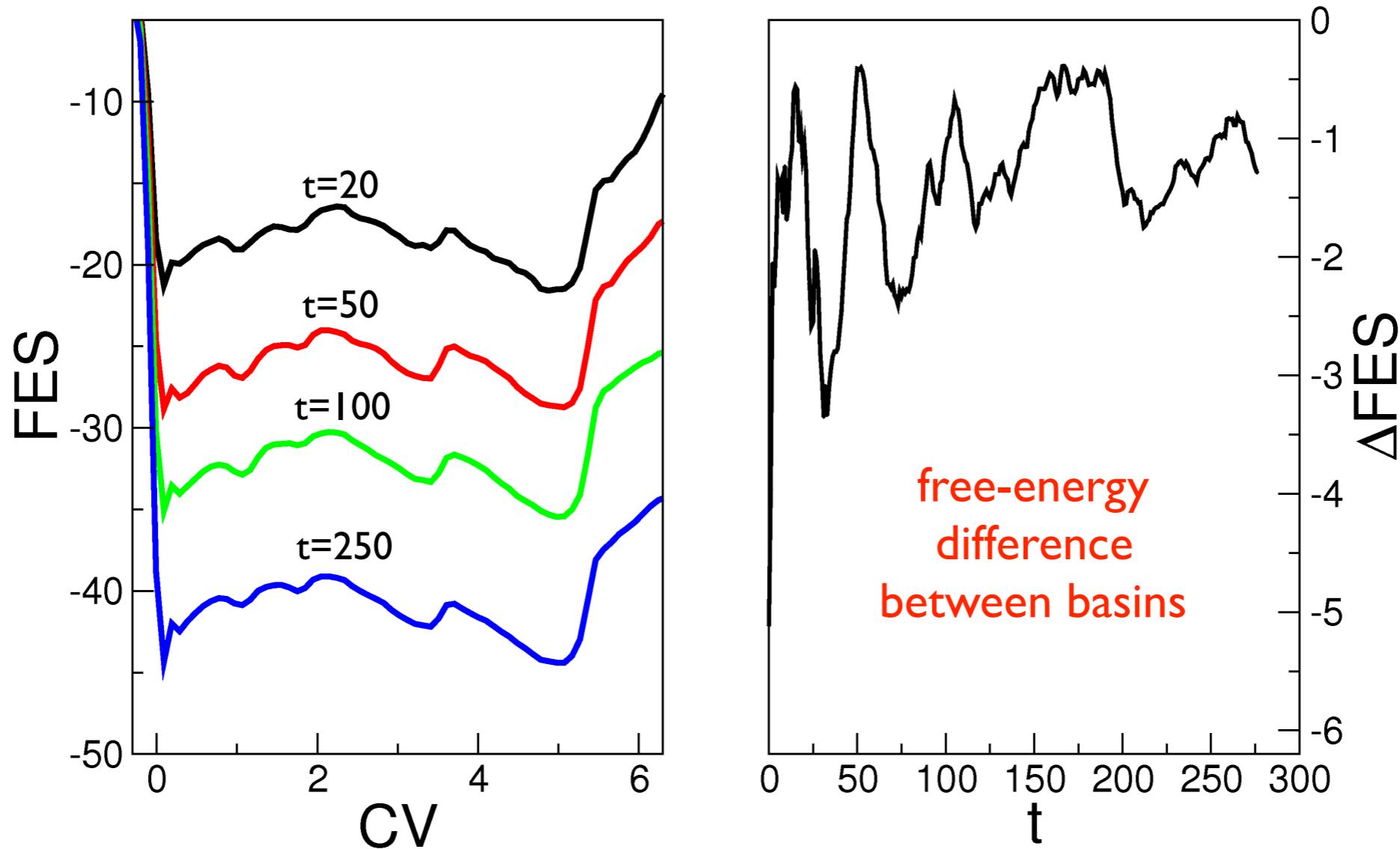
Usage: sum\_hills [options]

The following options are available

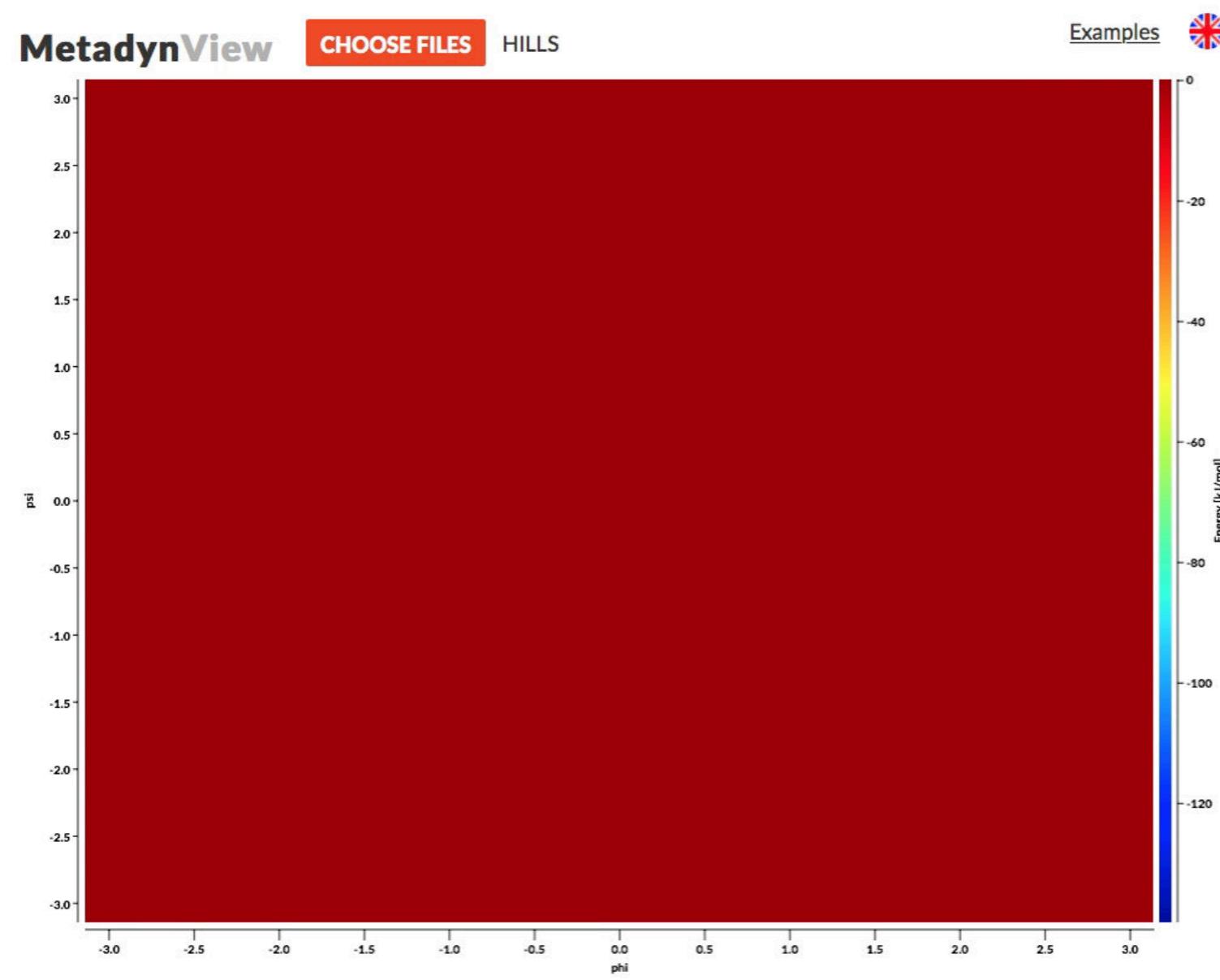
```
--help/-h - ( default=off ) print this help
--help-debug - ( default=off ) print special options that can be used to
               create regtests
--negbias - ( default=off ) print the negative bias instead of the free
               energy (only needed with welltempered runs and flexible hills)
--nohistory - ( default=off ) to be used with --stride: it splits the
               bias/histogram in pieces without previous history
--mintozero - ( default=off ) it translate all the minimum value in
               bias/histogram to zero (usefull to compare results)
--hills - specify the name of the hills file
--histo - specify the name of the file for histogram a colvar/hills
          file is good
--stride - specify the stride for integrating hills file (default
          0=never)
```

# How to monitor convergence

Qualitative indication of convergence from the analysis of the free-energy estimate as a function of time

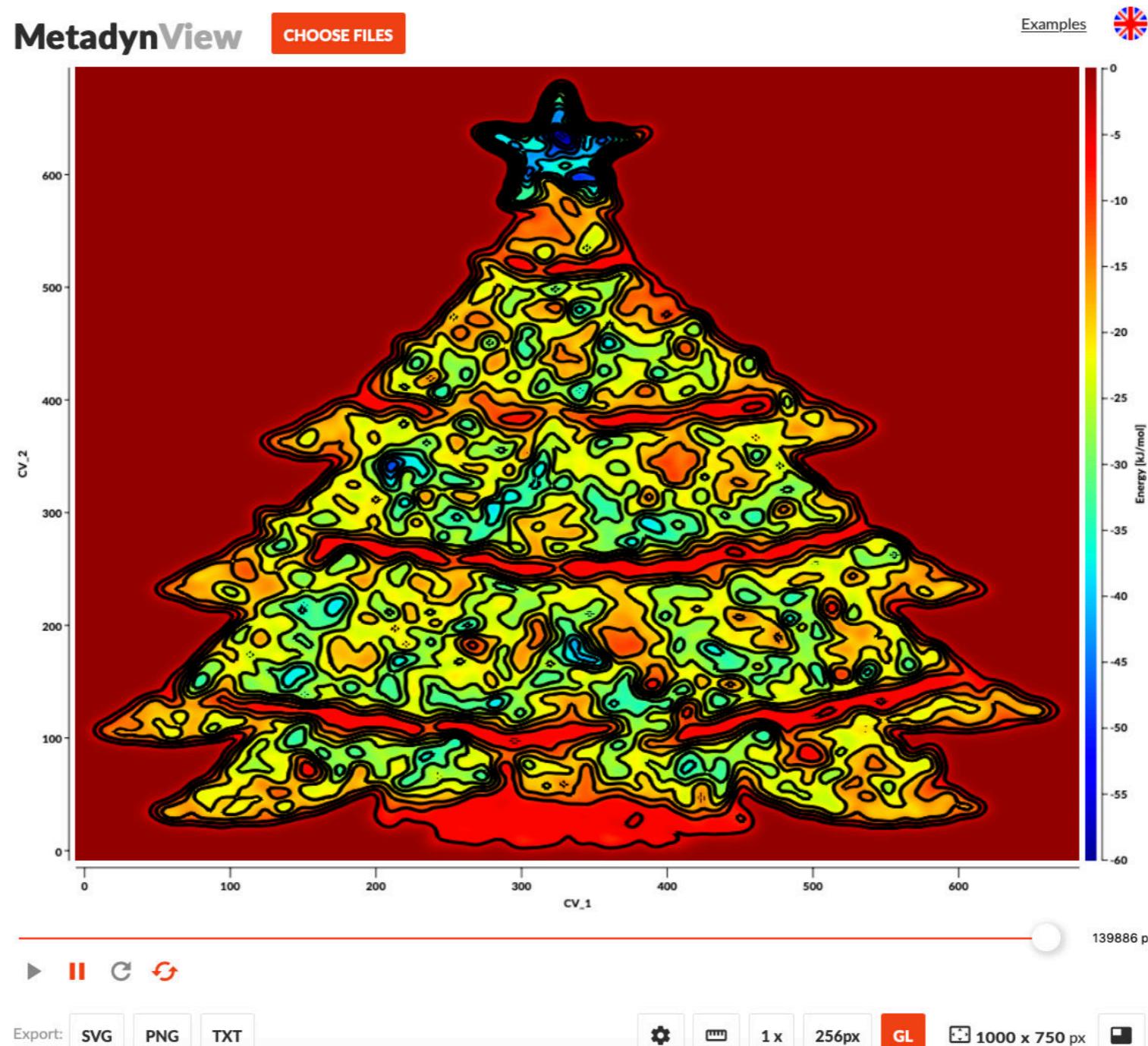


# How to monitor convergence or with a cool web server, like Metadyn View



# How to monitor convergence

## or with a cool web server, like Metadyn View



# How to quantitatively estimate convergence

We need a quantitative measure of convergence!

# Reweighting WT-MetaD simulations

At convergence, we have  $F(\mathbf{S})$

From this, the ensemble average of any function of the CVs can be calculated:

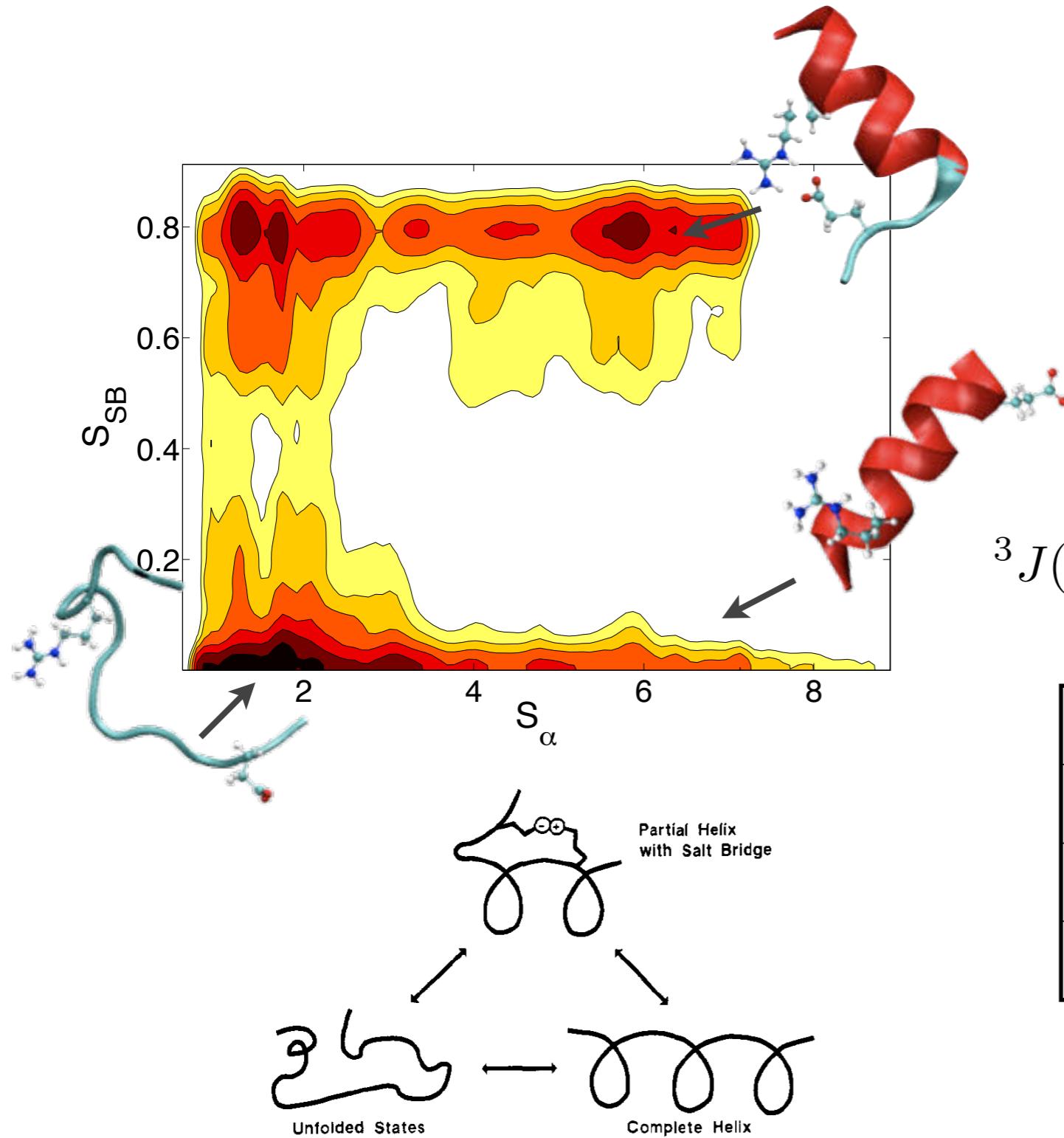
$$P_B(\mathbf{S}) \propto e^{-\beta F(\mathbf{S})} \quad \langle g(\mathbf{S}) \rangle = \frac{\int d\mathbf{S} g(\mathbf{S}) e^{-\beta F(\mathbf{S})}}{\int d\mathbf{S} e^{-\beta F(\mathbf{S})}}$$

What about the other “fast” degrees of freedom?  
Their probability distribution is distorted in a non-trivial way

- With a static bias:  $P_B(\mathbf{R}) \propto P(\mathbf{R})e^{\beta V(S(\mathbf{R}))}$
- MetaD bias is time-dependent!



# Linking simulations with experiments



Calculate average scalar  
couplings  ${}^3J$  from WT-MetaD  
simulations

$${}^3J(\Phi) = A\cos^2(\Phi + \Delta) + B\cos(\Phi + \Delta) + C$$

	$\chi^2$
PTMetaD Open	0.88
PTMetaD Closed	0.92
$\langle MD \rangle$	3.29

Caballero et al. Biophys J 2005

Barducci, Bonomi, Parrinello Biophys J 2010

# Reweighting schemes

## Three possible approaches for WT-MetaD\*

### Reconstructing the Equilibrium Boltzmann Distribution from Well-Tempered Metadynamics

JCC 2009

M. BONOMI, A. BARDUCCI, M. PARRINELLO

*Computational Science, Department of Chemistry and Applied Biosciences, ETH Zurich,  
c/o USI Campus, via Buffi 13, CH-6900 Lugano, Switzerland*

2

### A Time-Independent Free Energy Estimator for Metadynamics

Pratyush Tiwary\* and Michele Parrinello

*Department of Chemistry and Applied Biosciences, ETH, 8092 Zurich, Switzerland  
Facoltà di Informatica, Istituto di Scienze Computazionali, Università della Svizzera Italiana, Via G. Buffi 13, 6900 Lugano, Switzerland*

JPCB 2014

3

### Metadynamics with Adaptive Gaussians

Davide Branduardi,<sup>†</sup> Giovanni Bussi,<sup>\*,‡</sup> and Michele Parrinello<sup>§,||</sup>

<sup>†</sup>Theoretical Molecular Biophysics Group, Max Planck Institute for Biophysics, Max-von-Laue strasse 5, 60438, Frankfurt am Main, Germany

<sup>‡</sup>SISSA - Scuola Internazionale Superiore di Studi Avanzati, via Bonomea 265, 34136, Trieste, Italy

<sup>§</sup>Department of Chemistry and Applied Biosciences, ETH Zurich, Via G. Buffi 13, 6900, Lugano, Switzerland

<sup>||</sup>Facoltà di Informatica, Istituto di Scienze Computazionali, Università della Svizzera Italiana, Via G. Buffi 13, 6900, Lugano, Switzerland

JCTC 2012

## Plumed I.3

### Method I

Assuming a slow-growth regime, we evolve the “biased” probability distribution during the simulation:

$$P(\mathbf{R}, t + \Delta t) = e^{-\beta(\dot{V}_G(\mathbf{S}(\mathbf{R}), t) - \langle \dot{V}_G(\mathbf{S}(\mathbf{R}), t) \rangle) \Delta t} P(\mathbf{R}, t)$$

↑  
the new Gaussian      ↑  
average  
calculated in the  
biased ensemble

Finally we estimate the Boltzmann probability distribution like in umbrella sampling reweighting:

$$P_B(\mathbf{R}) \propto e^{+\beta V_G(\mathbf{S}(\mathbf{R}), t)} P(\mathbf{R}, t)$$

## Method 2

The average of any function of the microscopic coordinates is obtained from the bias potential and a reweighing factor:

$$\langle O(\mathbf{R}) \rangle_0 = \langle O(\mathbf{R}) e^{\beta(V(s(\mathbf{R}, t) - c(t)))} \rangle$$

where the reweighing factor is obtained from the bias potential:

$$\begin{aligned} e^{\beta c(t)} &\approx \frac{\tau(t + \Delta t) - \tau(t)}{\Delta t} = \frac{k_B \Delta T}{\gamma \omega \Delta t (2\pi)^{d/2} \det \sigma} \\ &= \int d\mathbf{s} [e^{\gamma V(s, t + \Delta t) / k_B \Delta T} - e^{\gamma V(s, t) / k_B \Delta T}] \end{aligned}$$

## Method 3

This method is based on the following relation between free energy, bias potential, and histogram accumulated in the biased ensemble:

$$\begin{aligned}\tilde{F}_N(s, t) = & -T \ln N(s, t) - V(s, t) \\ & + T \ln \int ds' N(s', t)\end{aligned}$$

Strictly-true when bias is time-independent and coincides with the umbrella-sampling reweighting recipe

In practice, the bias at the end of the simulation (at convergence) is used to reweigh all conformations generated by the metadynamics simulation using the relation above

## Summary II

- Theory of metadynamics and its well-tempered variant
- Setup a metadynamics simulation with PLUMED:
  - the PLUMED input file
  - choosing the basic parameters
  - choosing the collective variables
- Calculate free energies
- Overview of methods to reweigh metadynamics simulations
- Assessing convergence

# Guidelines for choosing the CVs

A good set of CVs for metadynamics (and other biasing techniques) should:

- Discriminate between relevant metastable states
- Include all the slow modes of a process
- Be as small as possible

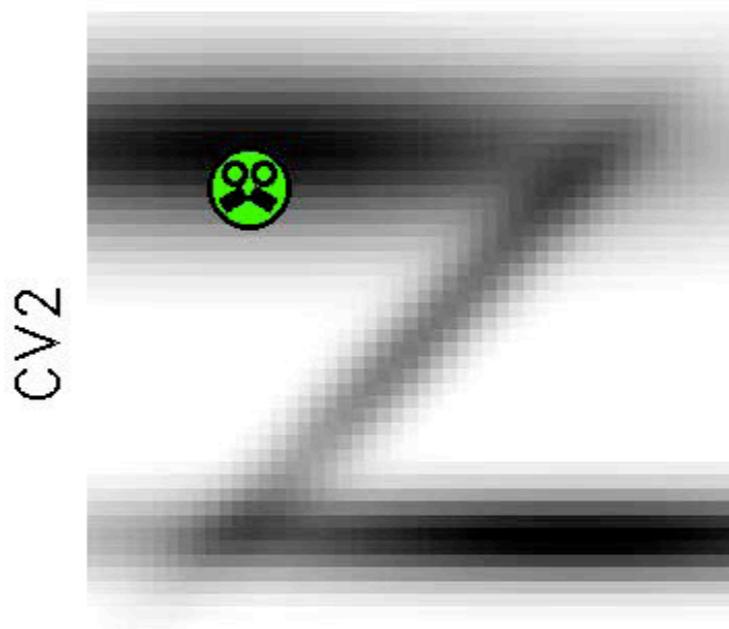
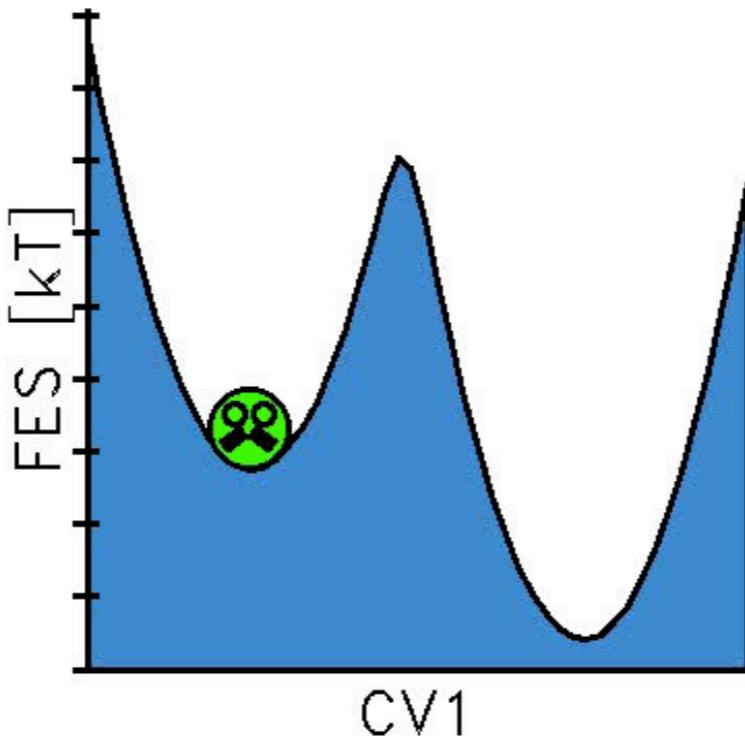
Metadynamics is inefficient with a large number of CVs.



Possible strategies:

- design smart CVs
- devise automatic protocols to find good CVs
- improve metadynamics to deal with a large number of CVs
- couple metadynamics with other methods, such as REM

# Hidden degrees of freedom



# Path Collective Variables

Branduardi, Gervasio, Parrinello JCP 2007

Map a configuration in terms of:

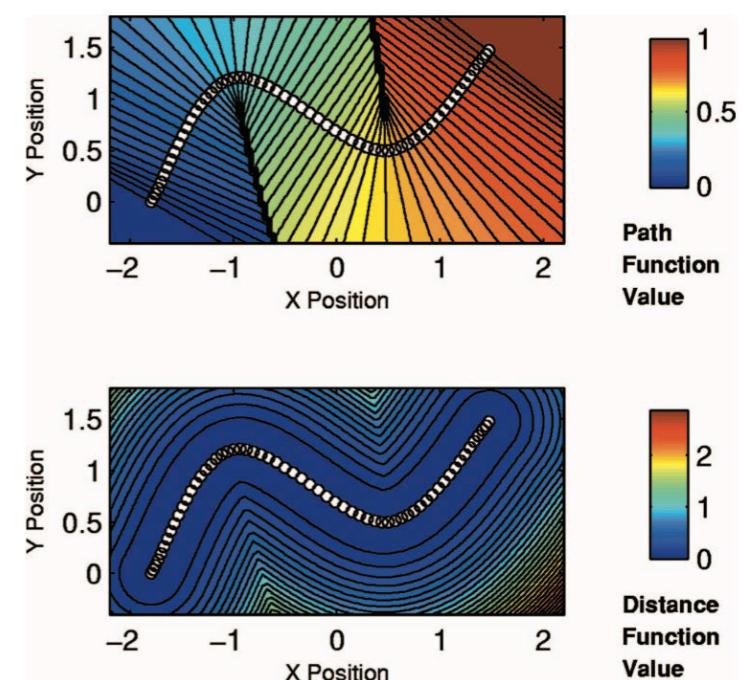
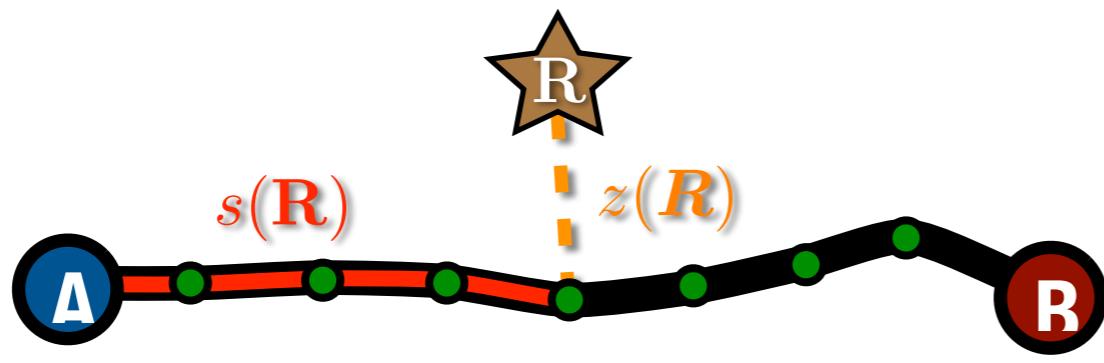
→ PROGRESS

$$s(\mathbf{R}) = \frac{1}{P-1} \frac{\sum_{l=1}^P (l-1) e^{-\lambda \|S(\mathbf{R}) - S(l)\|^2}}{\sum_{l=1}^P e^{-\lambda \|S(\mathbf{R}) - S(l)\|^2}}$$

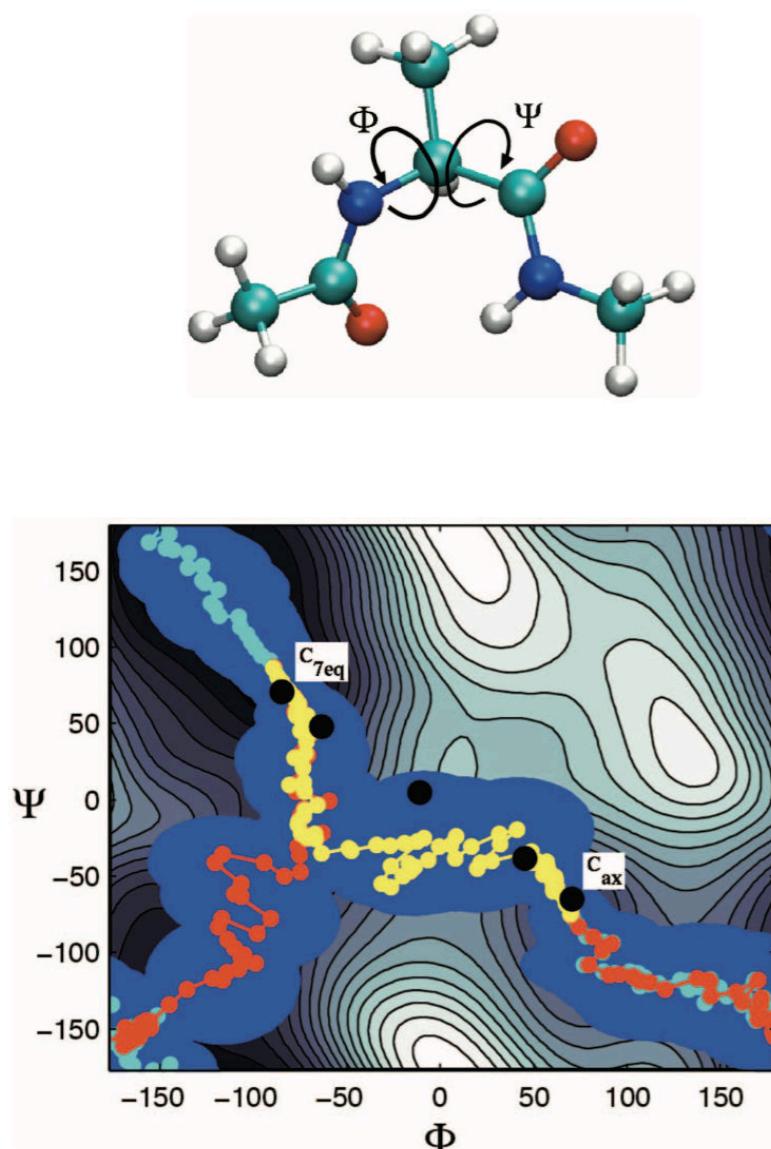
→ DISTANCE

$$z(\mathbf{R}) = -\frac{1}{\lambda} \ln \left( \sum_{l=1}^P e^{-\lambda \|S(\mathbf{R}) - S(l)\|^2} \right)$$

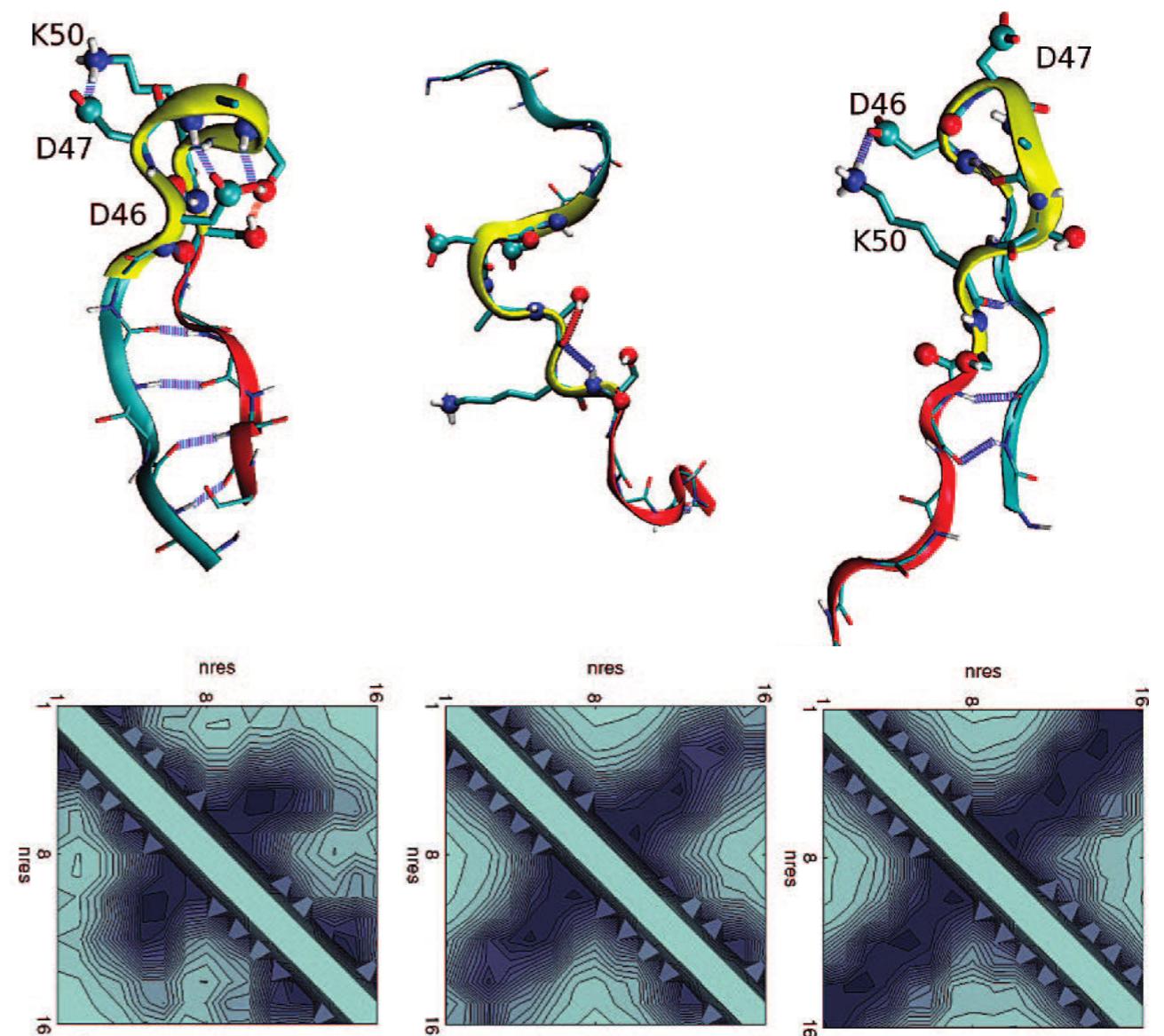
with respect to a reference path going from A to B



# RMSD path variables



# Contact Map path variables



Branduardi, Gervasio, Parrinello *JCP* (2007)

Bonomi, Branduardi, Gervasio, Parrinello *JACS* (2008)

# Replica exchange

Consider  $N$  non-interacting copies of the system:

$$R_1, \dots, R_N$$

in the canonical ensemble with different potential energy functions:

$$U_1, \dots, U_N$$

For the moment, let's assume that all copies have the same temperature

Swendsen & Wang PRL 1986  
Hansmann Chem Phys Lett 1997  
Sugita & Okamoto Chem Phys Lett 1999



Usually  $U_1$  is the “real” potential energy function, the others:

- are modified versions in which barrier crossing is accelerated
- might be completely unphysical

In REM we want to exchange configurations in order for all them to benefit from the good potentials

Swendsen and Wang PRL 1986  
 Hansmann Chem. Phys. Lett. 1997  
 Sugita and Okamoto Chem. Phys. Lett. 1999

# How should we exchange?

The exchange is regulated by a transition probability:

$$w(j \rightarrow k)$$

For the exchange process to converge to an equilibrium distribution, it is sufficient to impose the *detailed balance* to the transition probability

$$P(R_1, \dots, R_j, R_k, \dots, R_N) w(j \rightarrow k) = P(R_1, \dots, R_k, R_j, \dots, R_N) w(k \rightarrow j)$$



The probability of finding the extended system in a given state is:

$$P(R_1, \dots, R_N) \propto \prod_i^N e^{-\beta U_i(R_i)}$$

Before and after an exchange between configuration j and k:

$$P(R_1, \dots, R_j, R_k, \dots, R_N) \propto e^{-\beta[U_j(R_j) + U_k(R_k)]} \prod_{i \neq j, k}^N e^{-\beta U_i(R_i)} \quad \text{BEFORE}$$



$$P(R_1, \dots, R_k, R_j, \dots, R_N) \propto e^{-\beta[U_j(R_k) + U_k(R_j)]} \prod_{i \neq j, k}^N e^{-\beta U_i(R_i)} \quad \text{AFTER}$$

This leads to:

$$\frac{w(j \rightarrow k)}{w(k \rightarrow j)} = e^{-\Delta}$$

with

$$\Delta = \beta [U_j(R_k) + U_k(R_j) - U_j(R_j) - U_k(R_k)]$$

The usual Metropolis criterion satisfies this requirement

$$w(j \rightarrow k) = \begin{cases} 1, & \Delta \leq 0 \\ e^{-\Delta}, & \Delta > 0 \end{cases}$$

$$\Delta = \beta [U_j(R_k) + U_k(R_j) - U_j(R_j) - U_k(R_k)]$$

The acceptance depends on differences of energies:

- overlap between energy distributions

To keep in mind:

- a)** multiple copies of the system must be simulated
- b)** most of them might have an unphysical energy function

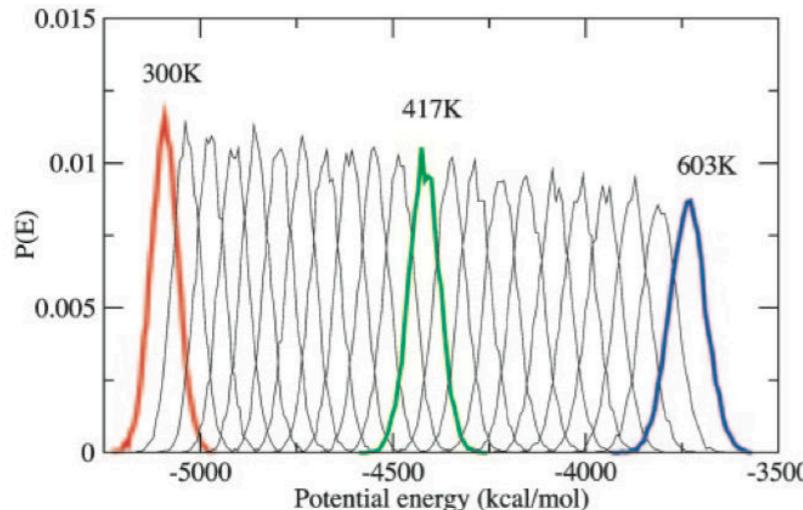
# Example

## Replica exchange with solute tempering: A method for sampling biological systems in explicit water

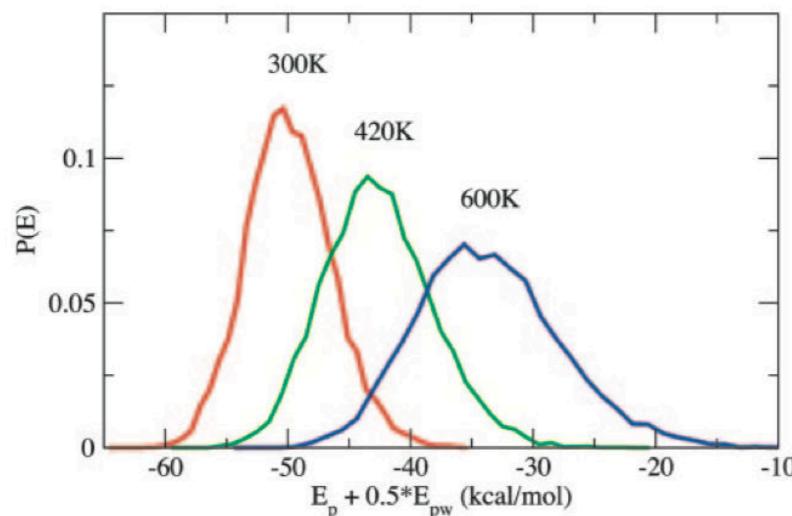
Pu Liu\*, Byungchan Kim\*, Richard A. Friesner, and B. J. Berne<sup>†</sup> PNAS 2005

$$E_m(X) = E_p(X) + \left[ \frac{\beta_0}{\beta_m} \right] E_{ww}(X) + \left[ \frac{\beta_0 + \beta_m}{2\beta_m} \right] E_{pw}(X),$$

PT



REST



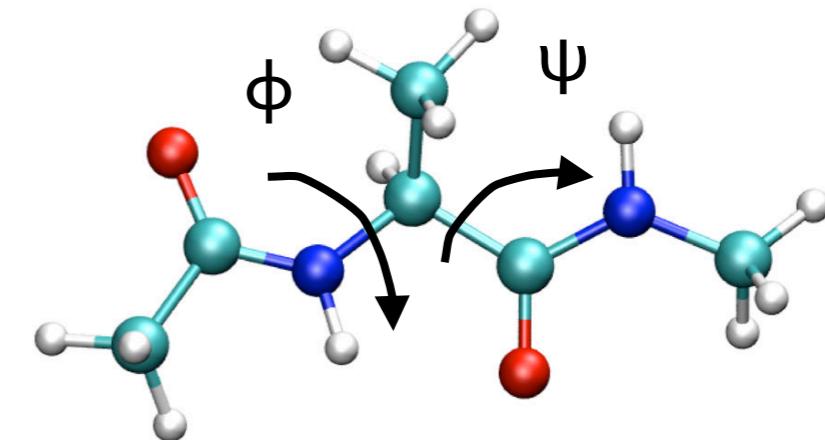
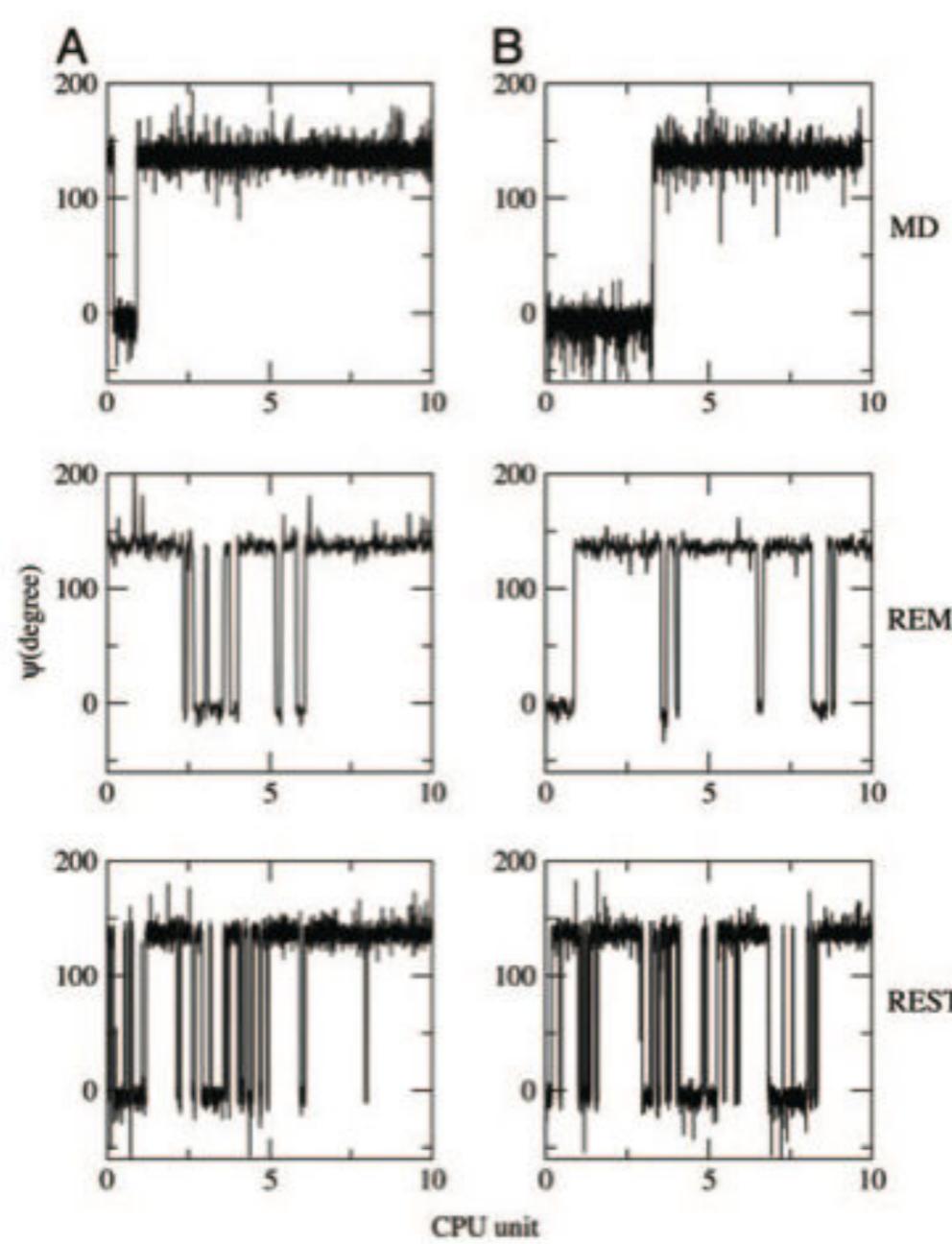
The acceptance does not depend on water-water interactions

$$\Delta_{nm} = (\beta_n - \beta_m)[(E_p(X_m) + \frac{1}{2}E_{pw}(X_m)) - (E_p(X_n) + \frac{1}{2}E_{pw}(X_n))].$$

# Example

## Replica exchange with solute tempering: A method for sampling biological systems in explicit water

Pu Liu\*, Byungchan Kim\*, Richard A. Friesner, and B. J. Berne<sup>†</sup> PNAS 2005



Alanine dipeptide in water:

- REM: 22 replicas 300-600K
- REST: 3 replicas, 300K, 420K and 600K

# Example

## Replica exchange with solute tempering: A method for sampling biological systems in explicit water

Pu Liu\*, Byungchan Kim\*, Richard A. Friesner, and B. J. Berne<sup>†</sup> PNAS 2005

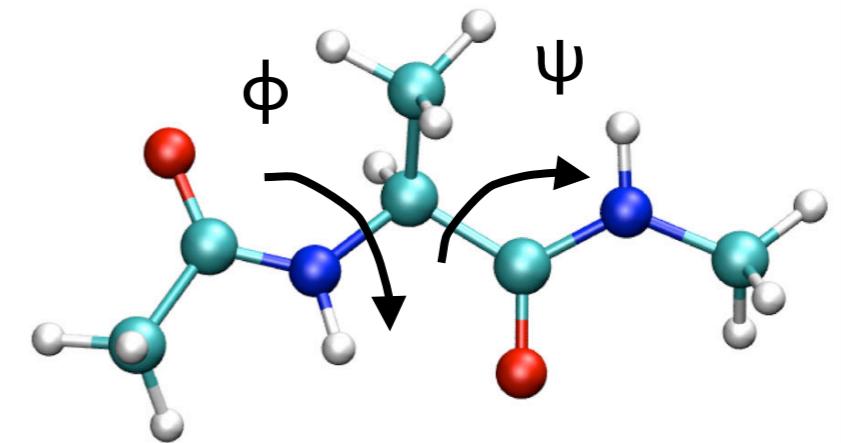
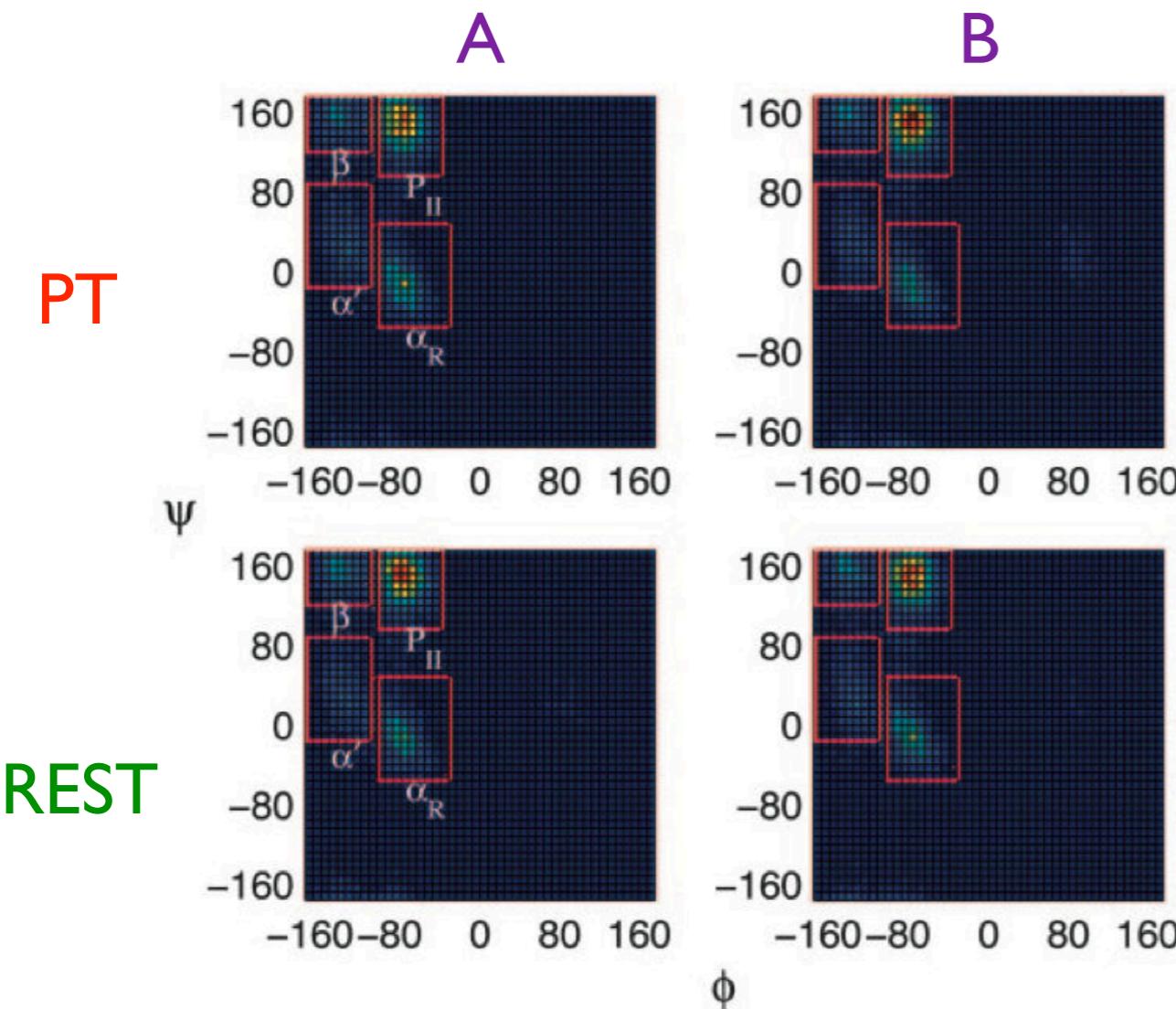


Table 1. Comparison of the percent population of four regions

Population region	REM		REST	
	A	B	A	B
$P_{II}$	0.430	0.463	0.467	0.466
$\alpha_R$	0.233	0.214	0.212	0.214
$\beta$	0.182	0.181	0.183	0.183
$\alpha'$	0.155	0.142	0.138	0.137

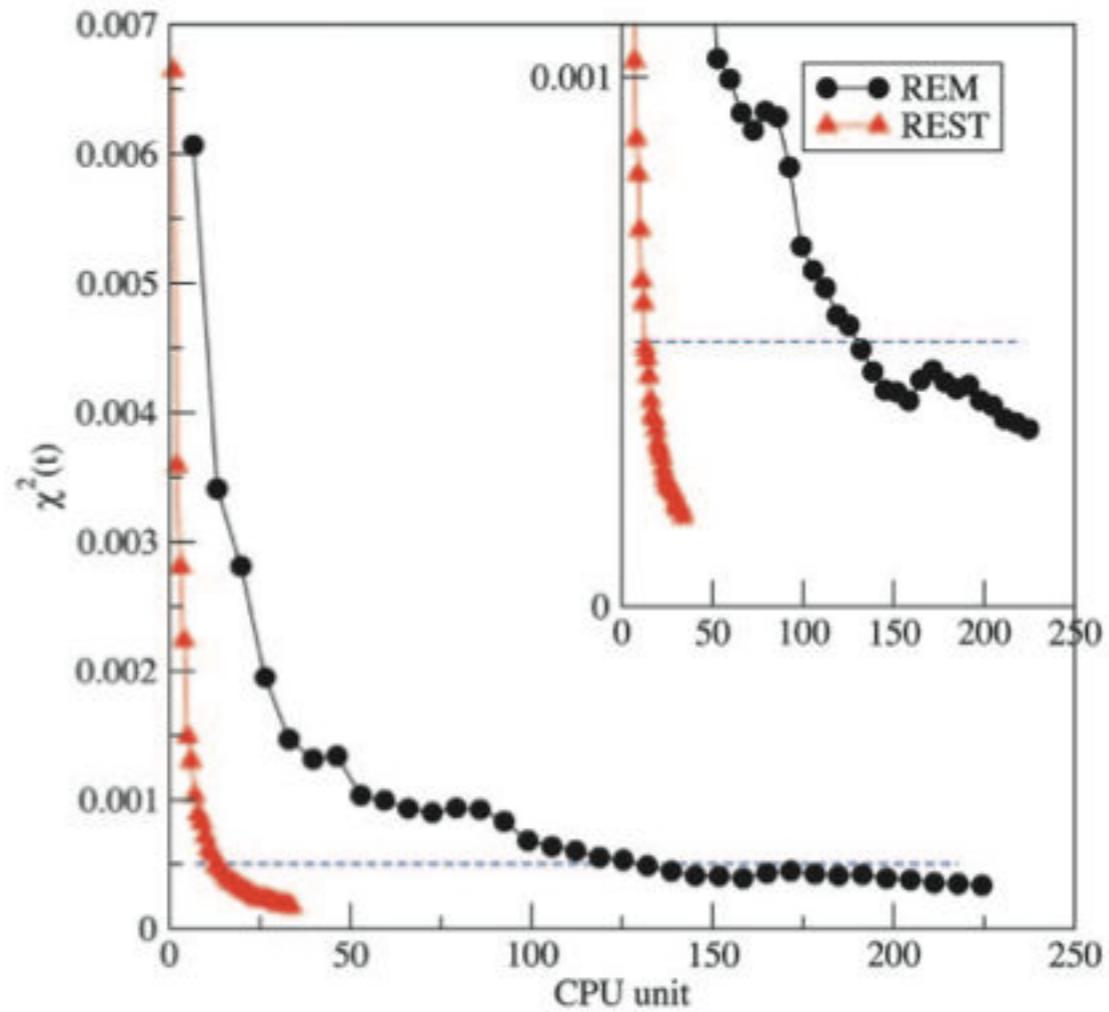
The results are collected over 10 ns for standard REM and REST. The errors listed are the standard deviations of the percent population.

The free energy was properly reconstructed at a fraction of the computational cost

# Example

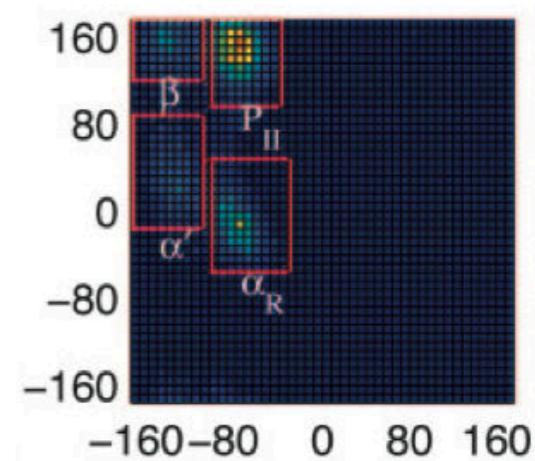
## Replica exchange with solute tempering: A method for sampling biological systems in explicit water

Pu Liu\*, Byungchan Kim\*, Richard A. Friesner, and B. J. Berne<sup>†</sup> PNAS 2005



Ergodic measure\*:

$$\chi^2(t) = \frac{1}{mn} \sum_{i=1, j=1}^{m, n} (\bar{R}_{i,j}^A - \bar{R}_{i,j}^B)^2.$$



\*Thirumalai et al PRA 1989

# A popular REM: Parallel Tempering

Temperature is just a scaling factor of the potential energy

Consider  $N$  non-interacting copies of the system

$$R_1, \dots, R_N$$

in the canonical ensemble at different temperatures

$$T_1, \dots, T_N$$

Same considerations for the transition probability lead to

$$w(j \rightarrow k) = \begin{cases} 1, & \Delta \leq 0 \\ e^{-\Delta}, & \Delta > 0 \end{cases} \quad \Delta = (\beta_j - \beta_i) [U(R_i) - U(R_j)]$$

# Early days applications

- Monte Carlo on Spin Glasses: exchange of partial configuration between replicas at different temperatures

Swendsen & Wang PRL 1986

- Swap of entire conformations

Geyer 1991

- First Monte Carlo simulation of a small peptide (Tyr-Gly-Gly-Phe-Met)

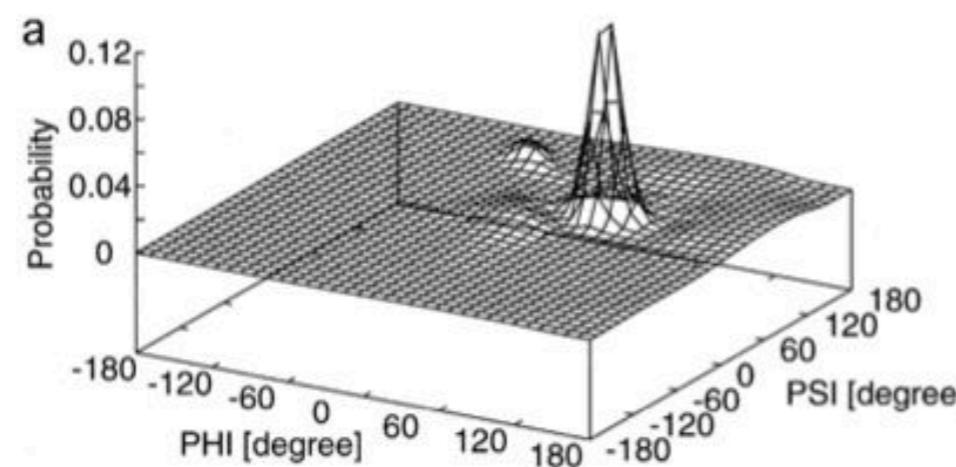
Hansmann Chem Phys Lett 1997

- Two years later, first application in combination with MD...

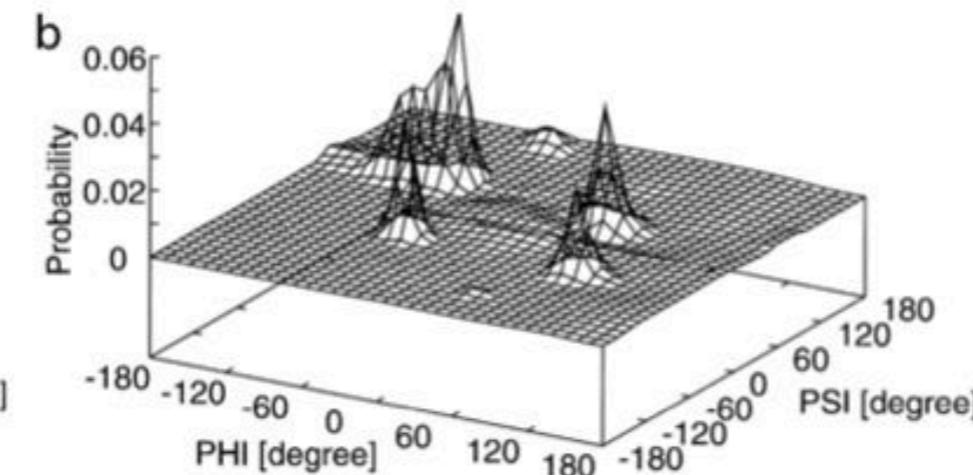
# PT and Molecular Dynamics

Gly-2

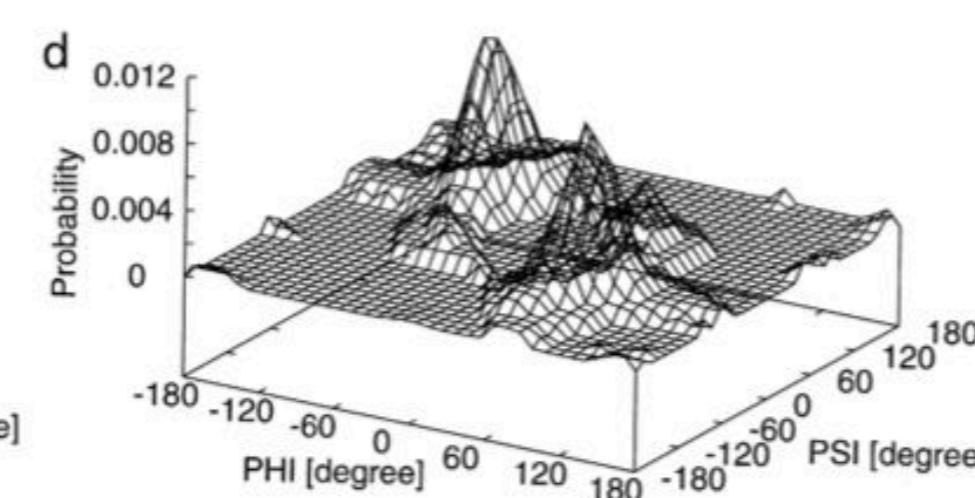
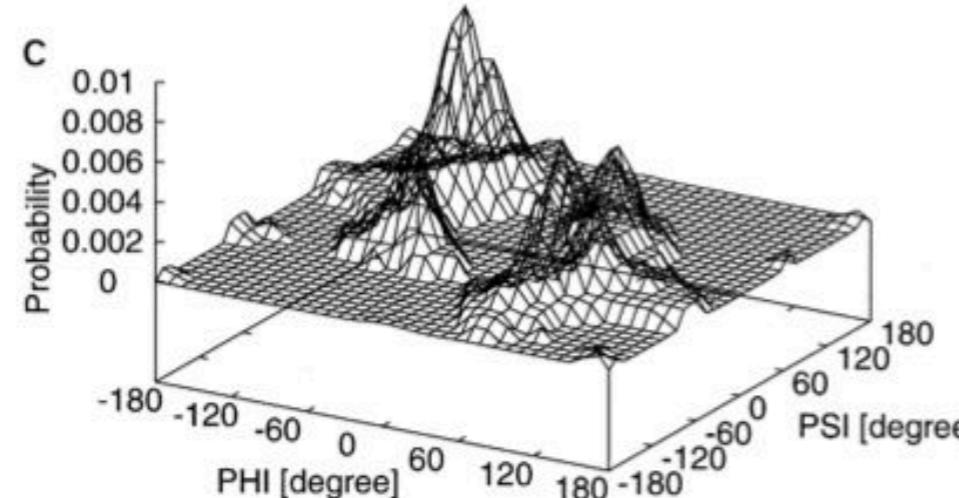
MD



PT



T=200K



T=600K

# PT and Molecular Dynamics

When sampling with MD, usually there is a rescaling of the velocities after an exchange is accepted

Exchange       $i \rightarrow j$

$$v'_i = \sqrt{\frac{T_j}{T_i}} v_i$$

$$v'_j = \sqrt{\frac{T_i}{T_j}} v_j$$



after



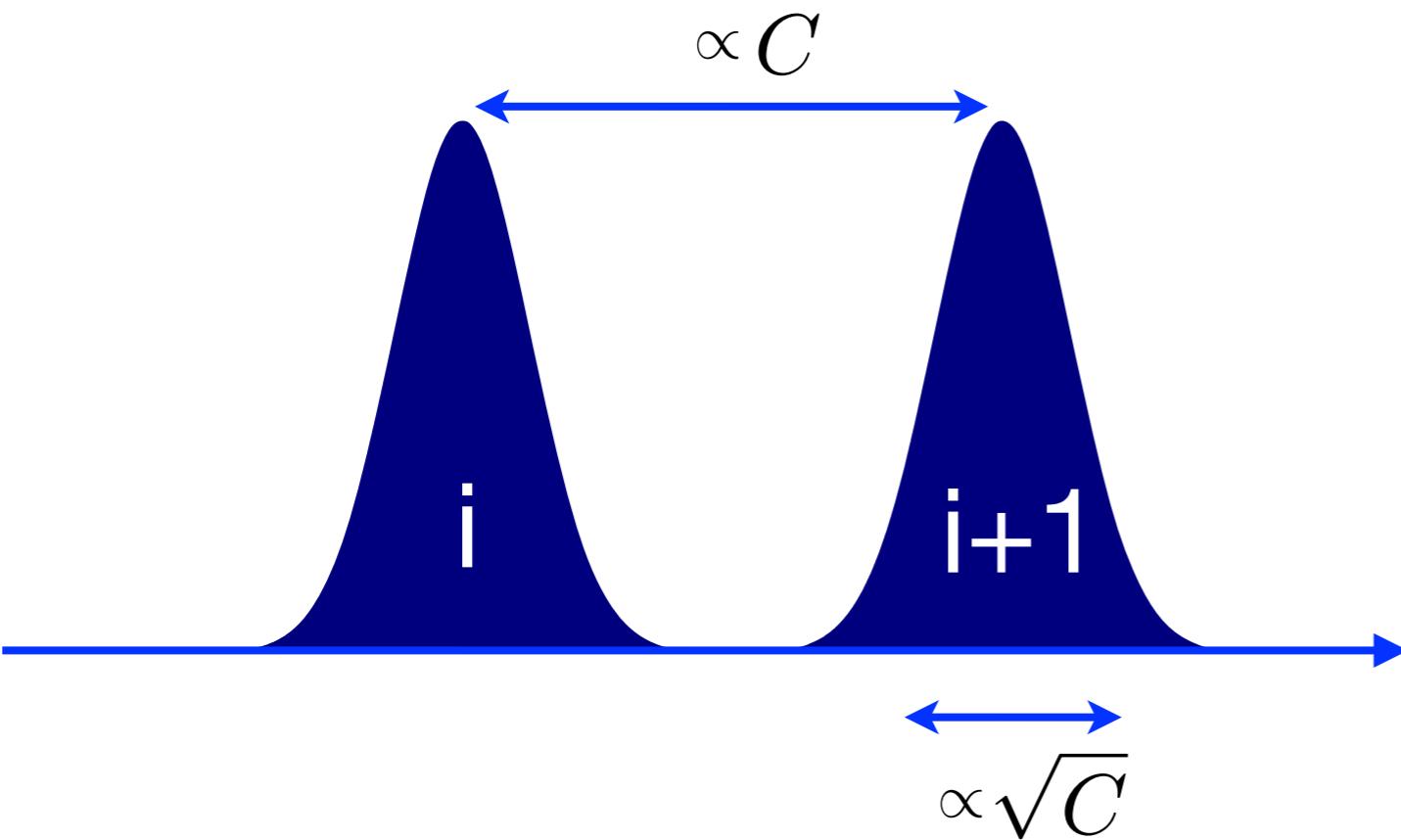
before

# The exchange acceptance

$$\Delta = (\beta_j - \beta_i) [U(R_i) - U(R_j)]$$

To exchange efficiently, an overlap between energy distributions at different temperatures is needed!

$$U_i = \langle U \rangle_{\beta_i} \pm \sqrt{\langle \Delta U^2 \rangle_{\beta_i}} = \langle U \rangle_{\beta_i} \pm \sqrt{-\frac{\partial \langle U \rangle_{\beta}}{\partial \beta}}$$



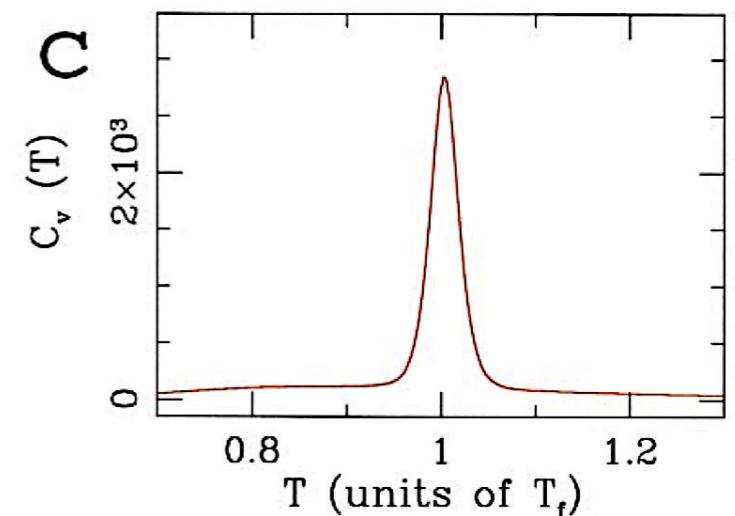
Number of replicas  
increases with  $\sqrt{C}$

- bad scaling with system size
- more replicas around critical points/peaks in specific heat

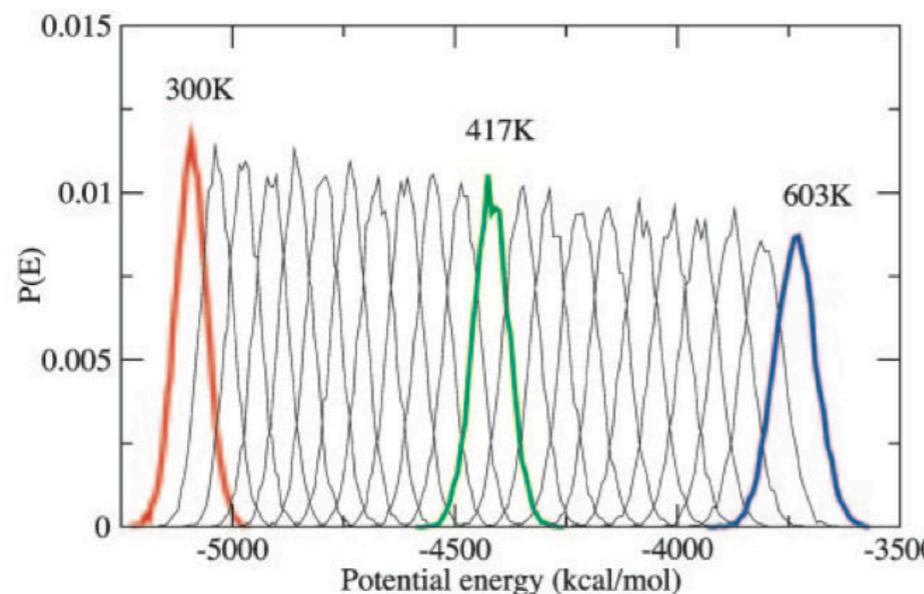
## Two problems of REM

- In solvated systems, regular distribution of replicas, their number scales (poorly) with the square root of the degrees of freedom
- In systems with a peak in the specific heat, replicas distribution must be optimized around the peak

Rathore *et al.* J Chem Phys 2005  
Katzgraber *et al.* J Stat Mech-Theory 2006  
Gront, Kolinsky J.Phys Condens Matter 2007

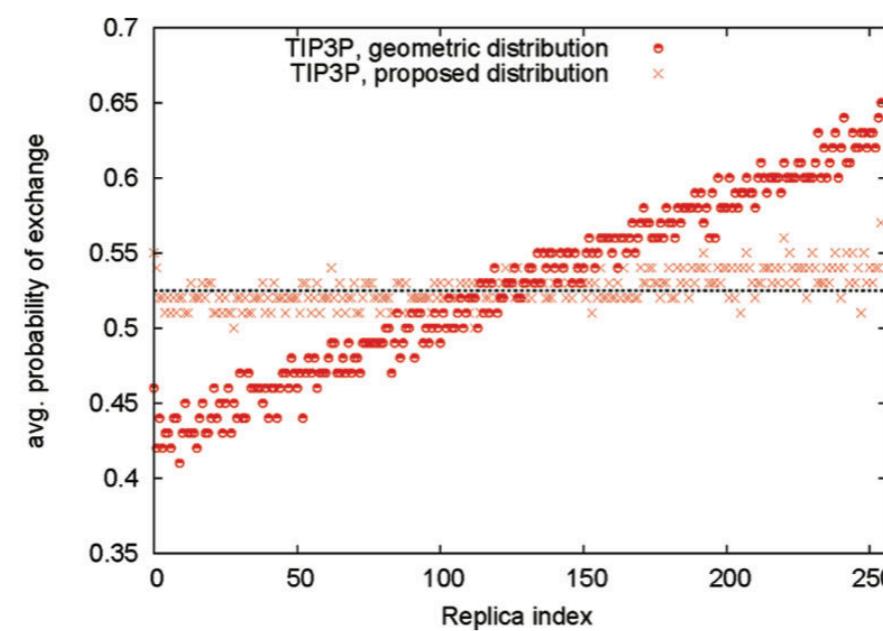
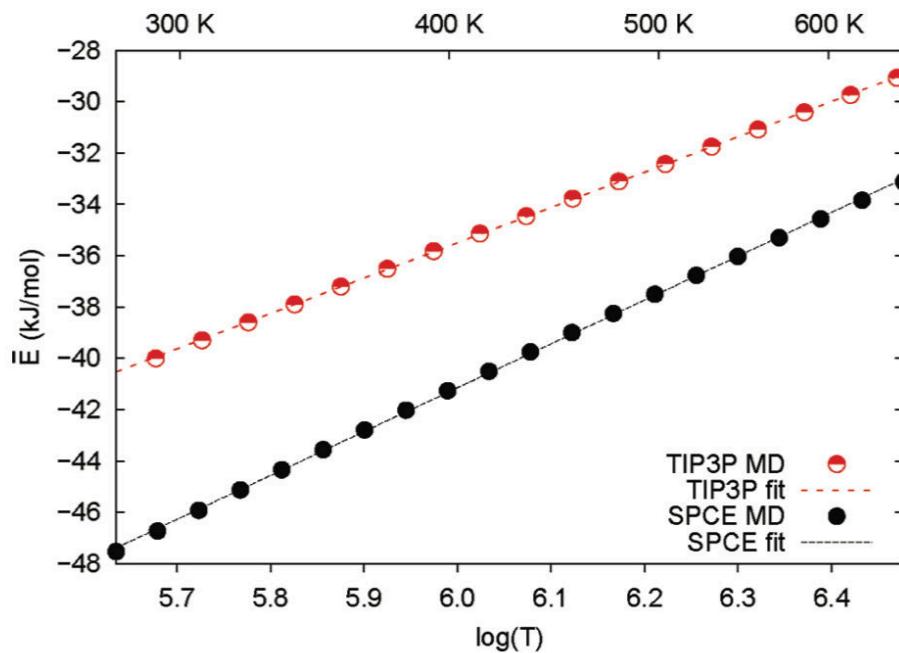


# Example of well-behaved system is proteins in explicit solvent



- specific heat constant\*
- $P(E)$  close to Gaussian
- number of replicas  $\sqrt{N}$

\*Specific heat with popular all-atom explicit water force fields is not that constant

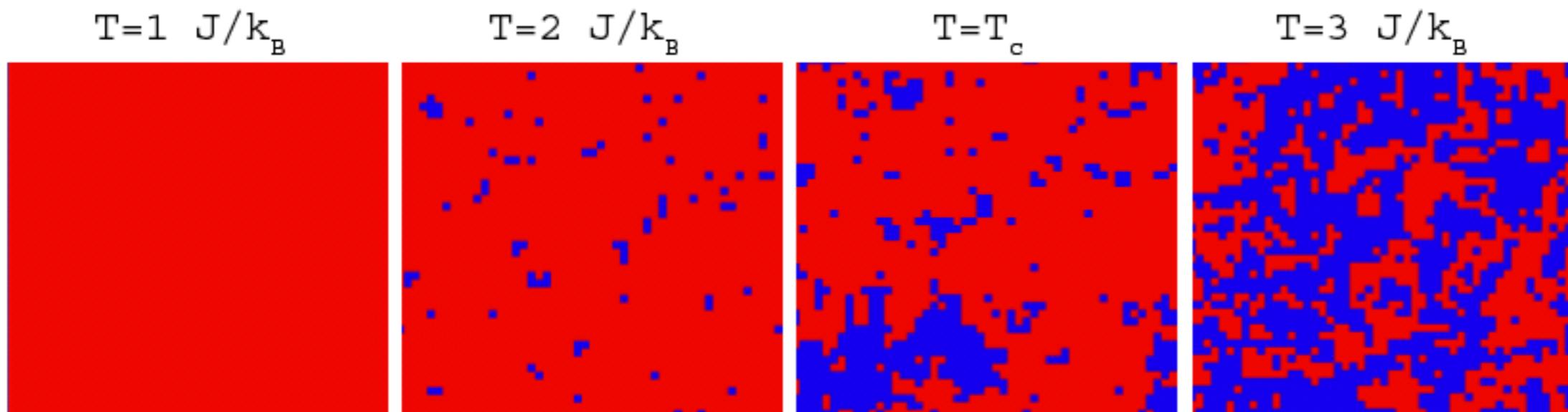
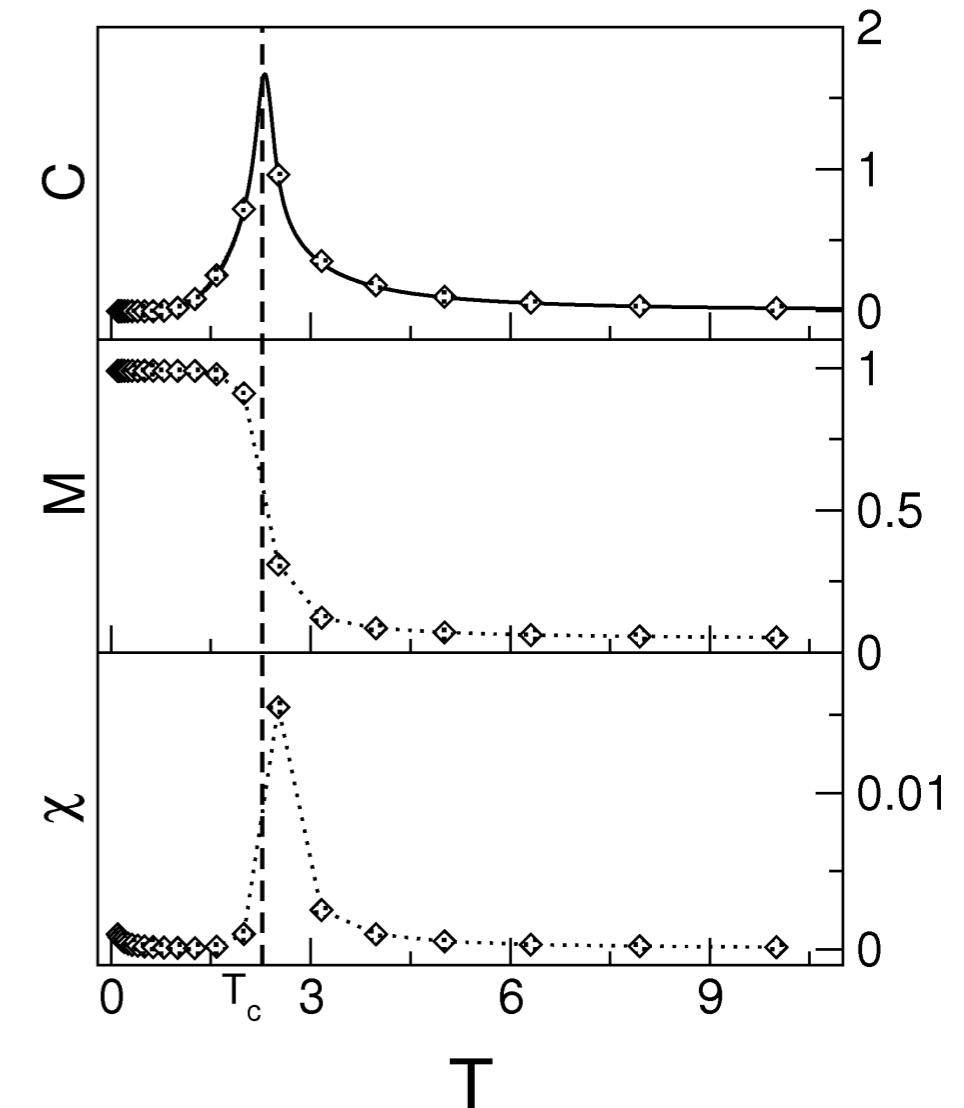


$$\frac{1}{T_i} = \frac{1}{T_{i-1}} - \sqrt{\frac{c/a}{T_{i-1}}}$$

Potentially bad systems are proteins in implicit solvents, coarse grained potentials, structure-based potentials...

2D ferromagnetic Ising model

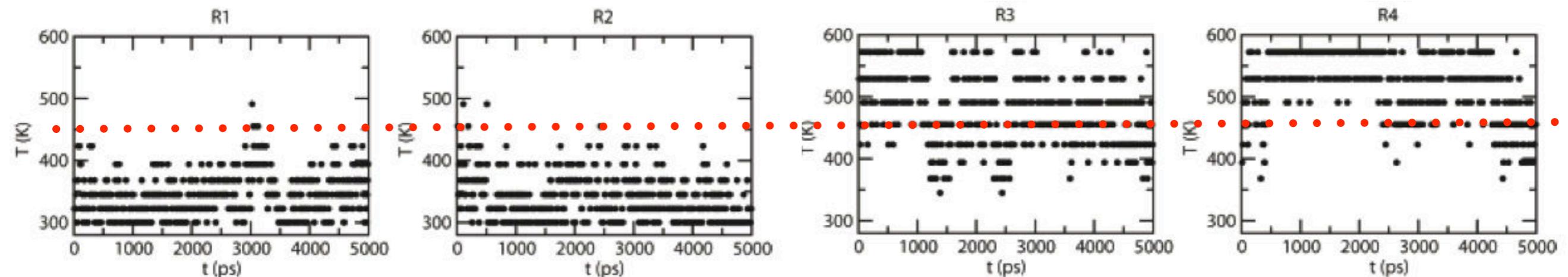
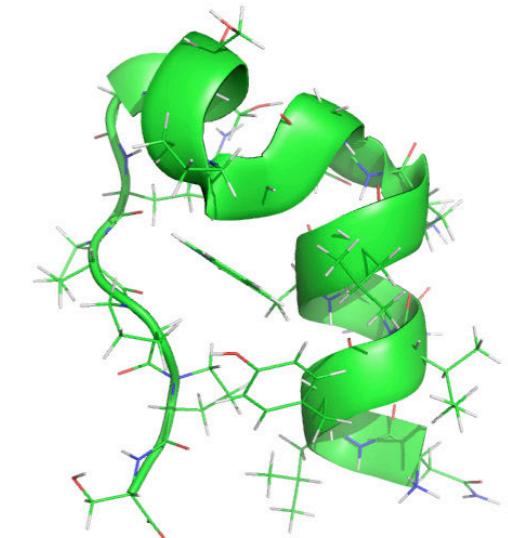
$$\mathcal{H} = -J \sum_{\langle i,j \rangle} S_i S_j \quad J = 1 \quad S = \pm 1$$



Easy to verify that something is wrong by monitoring the diffusion of replicas in temperature

### Replica Exchange Solute Tempering on TrpCage miniprotein

$$E_m(X) = E_p(X) + \left[ \frac{\beta_0}{\beta_m} \right] E_{ww}(X) + \left[ \frac{\beta_0 + \beta_m}{2\beta_m} \right] E_{pw}(X),$$



Huang *et al.* J Phys Chem B 2007

Replica diffusion in temperature is blocked close to the peak of the specific heat!

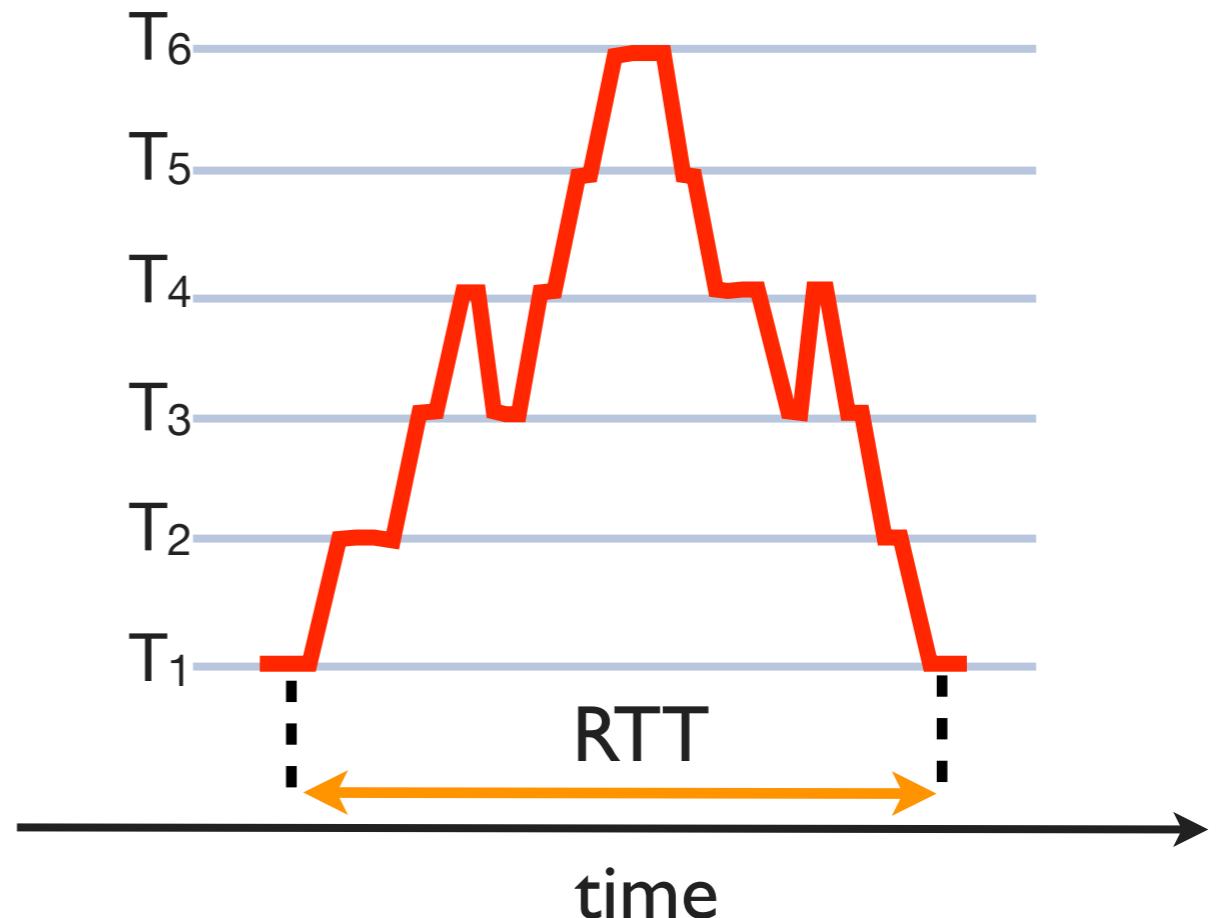
# Parallel Tempering and...

- N replicas at different  $T_i$
- Exchange configurations with a Metropolis criteria:

$$p(j \rightarrow k) = \min\{1, e^{\Delta_{j,k}}\}$$

with:

$$\Delta_{j,k} = (\beta_j - \beta_k)(U(\mathbf{R}_j) - U(\mathbf{R}_k))$$



- Prevent cold replicas to get trapped in local energy minima

# ... Metadynamics together

Bussi, Gervasio, Laio, Parrinello JACS 2006

- N replicas at different T
- Same CVs, different bias potential
- Modified exchange probability:

$$\begin{aligned}\Delta_{j,k} = & (\beta_j - \beta_k)(U(\mathbf{R}_j) - U(\mathbf{R}_k)) \\ & + \beta_j[V_G^{(j)}(\mathbf{S}(\mathbf{R}_j), t) - V_G^{(j)}(\mathbf{S}(\mathbf{R}_k), t)] \\ & + \beta_k[V_G^{(k)}(\mathbf{S}(\mathbf{R}_k), t) - V_G^{(k)}(\mathbf{S}(\mathbf{R}_j), t)]\end{aligned}$$

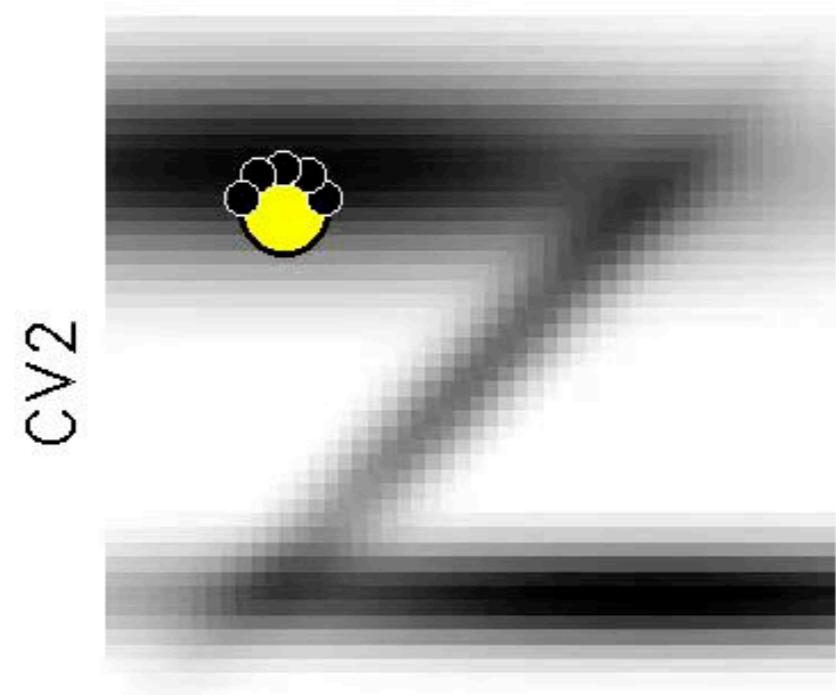
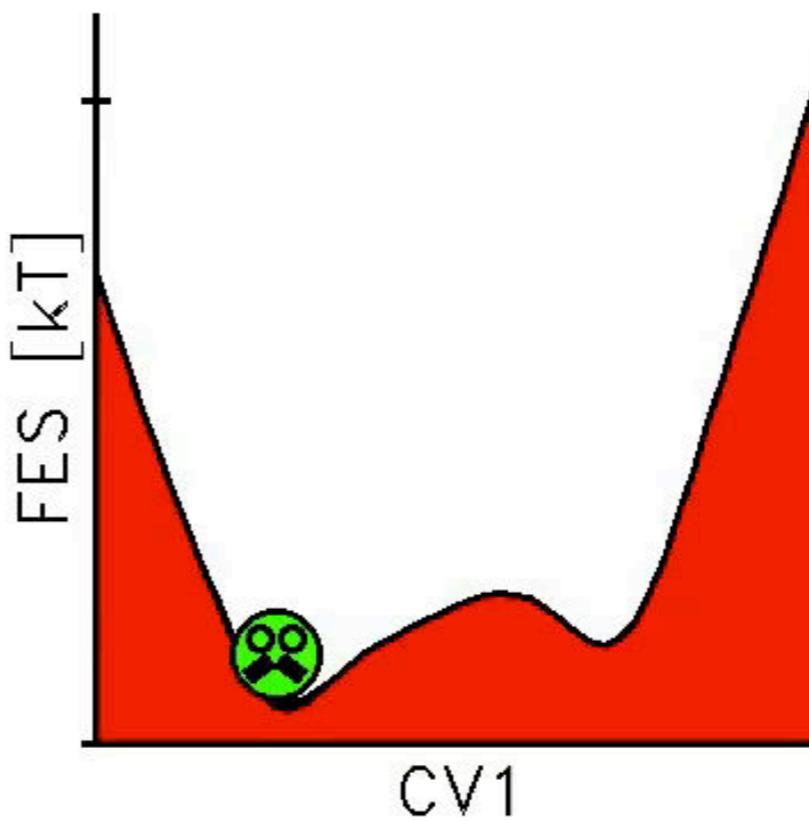
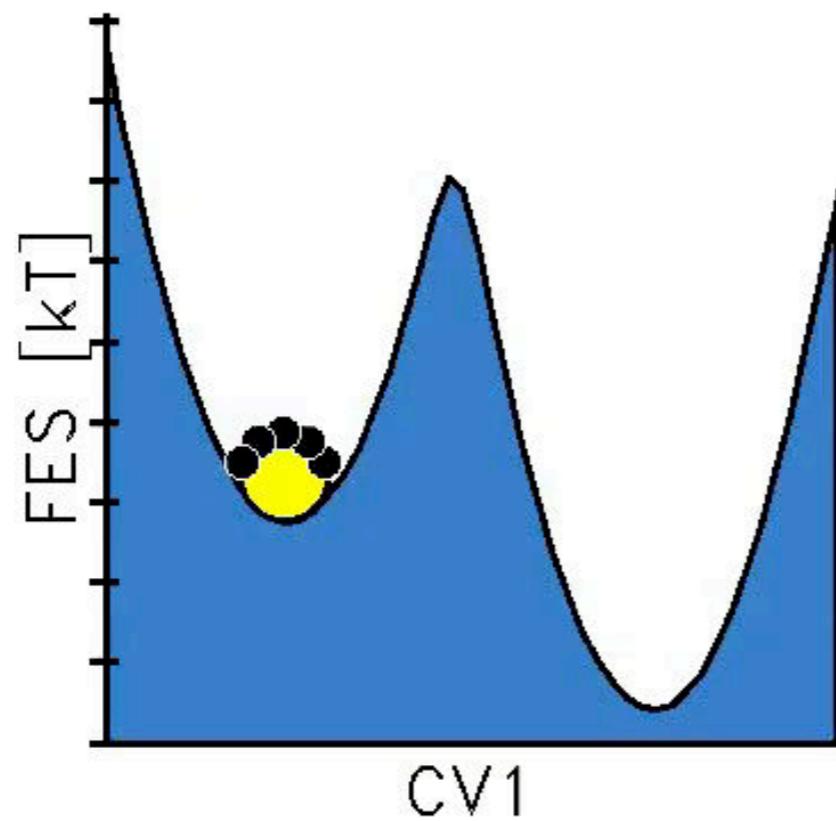
**MetaD**: cross high barriers in a few selected degrees of freedom



**PT**: cross moderate barriers along all degrees of freedom



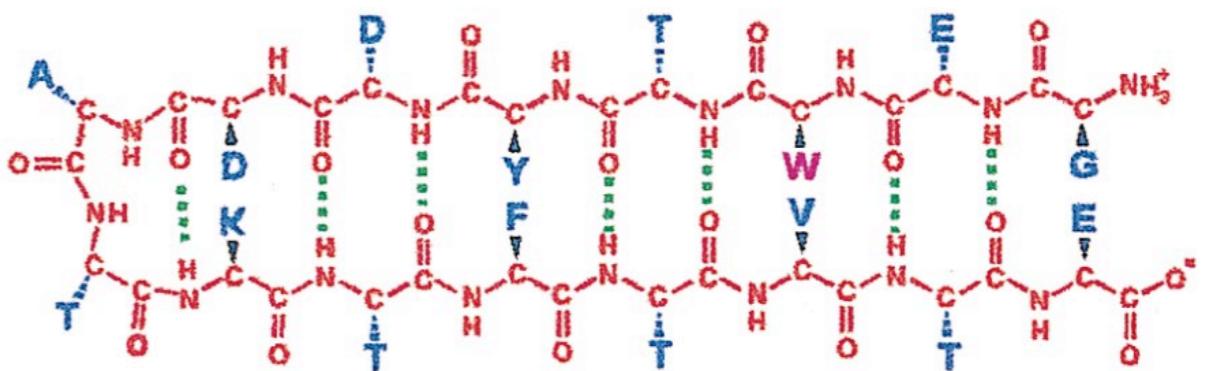
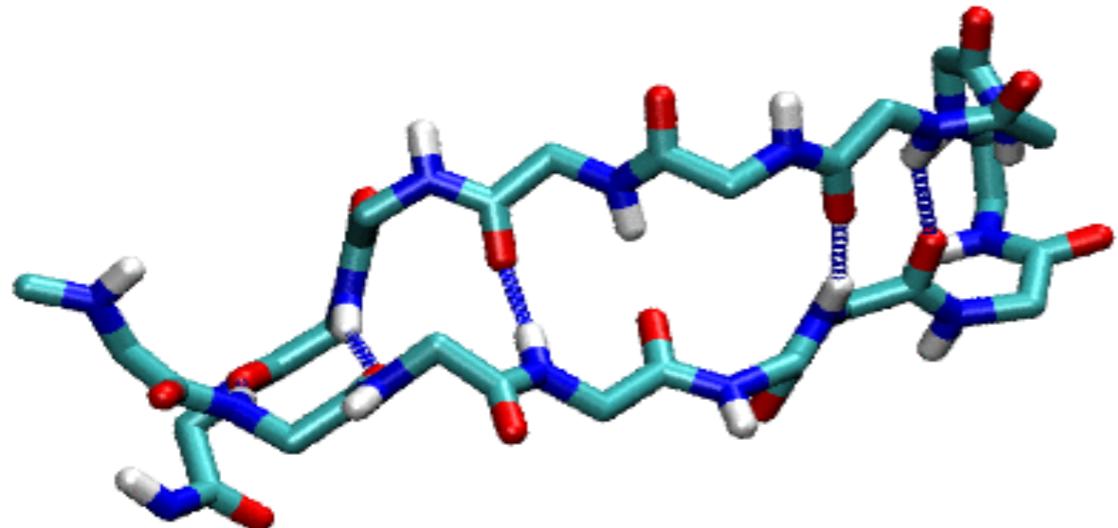
The combination improves both methods!



CV1

movie by G. Bussi

# GB1 $\beta$ -hairpin



- 16 residues (GEWTYDDATKTFTVTE)
- Prototype of “fast”-folder ( $6\mu\text{s}^*$ )
- Well studied (experiments\* and simulations\*\*)

Bussi, Gervasio, Laio & Parrinello JACS 2006

\*Munoz, Eaton, et al, Nature 1997, PNAS 1998

\*\*e.g. Zhou, Berne, Germain PNAS 2001; Bolhuis PNAS 2003

# Simulation details

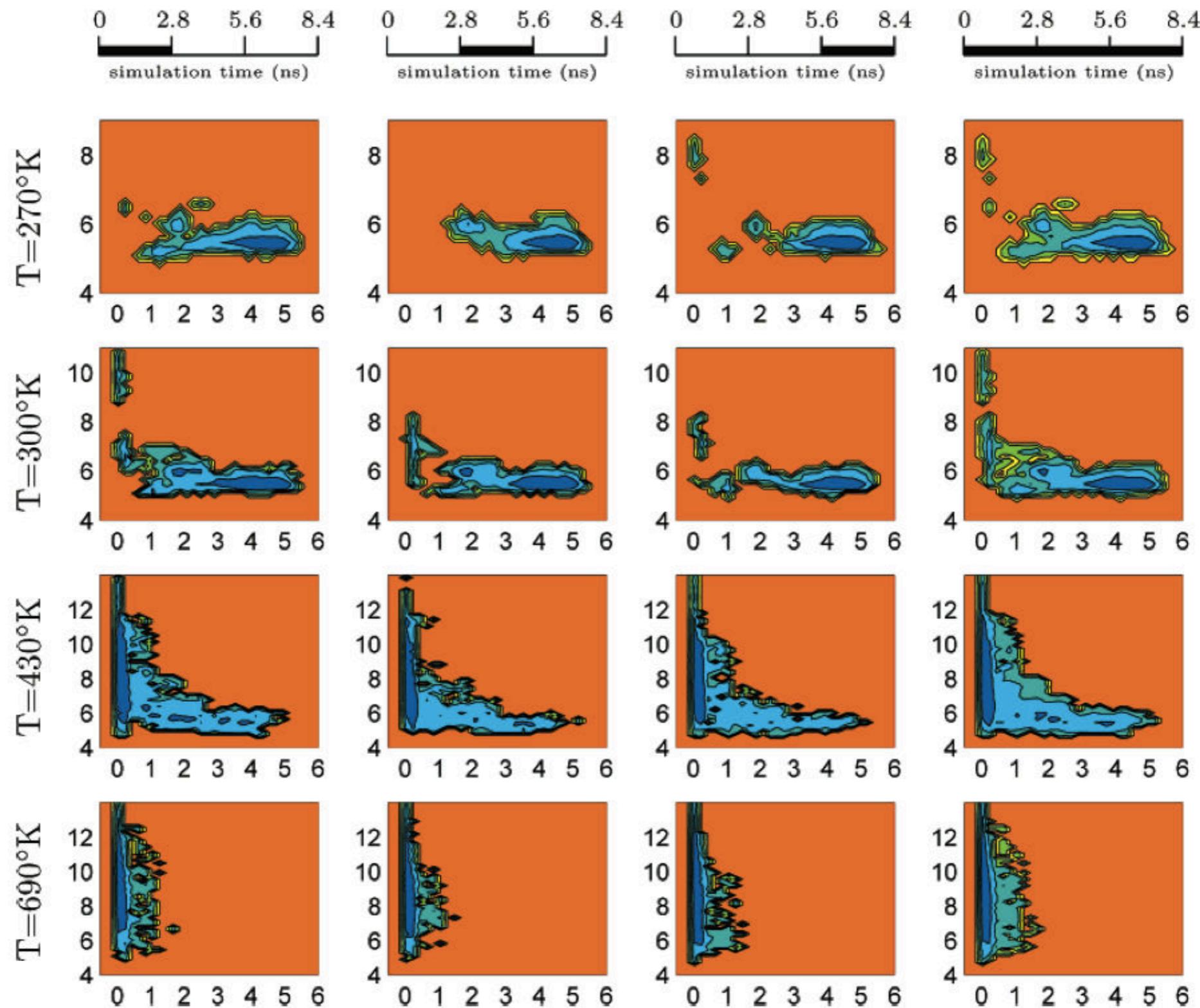
- ORAC MD code\* (+MetaD+PT)
- OPLSAA force field
- Explicit water (TIP3P, 1559 molecules)
- 64 replicas (270 to 695 K)
- Two collective variables:
  - hydrophobic core gyration radius

$$\text{Gyr} = \sqrt{\sum_i \left( \mathbf{r}_i - \frac{1}{N_b} \sum_j \mathbf{r}_j \right)^2}$$

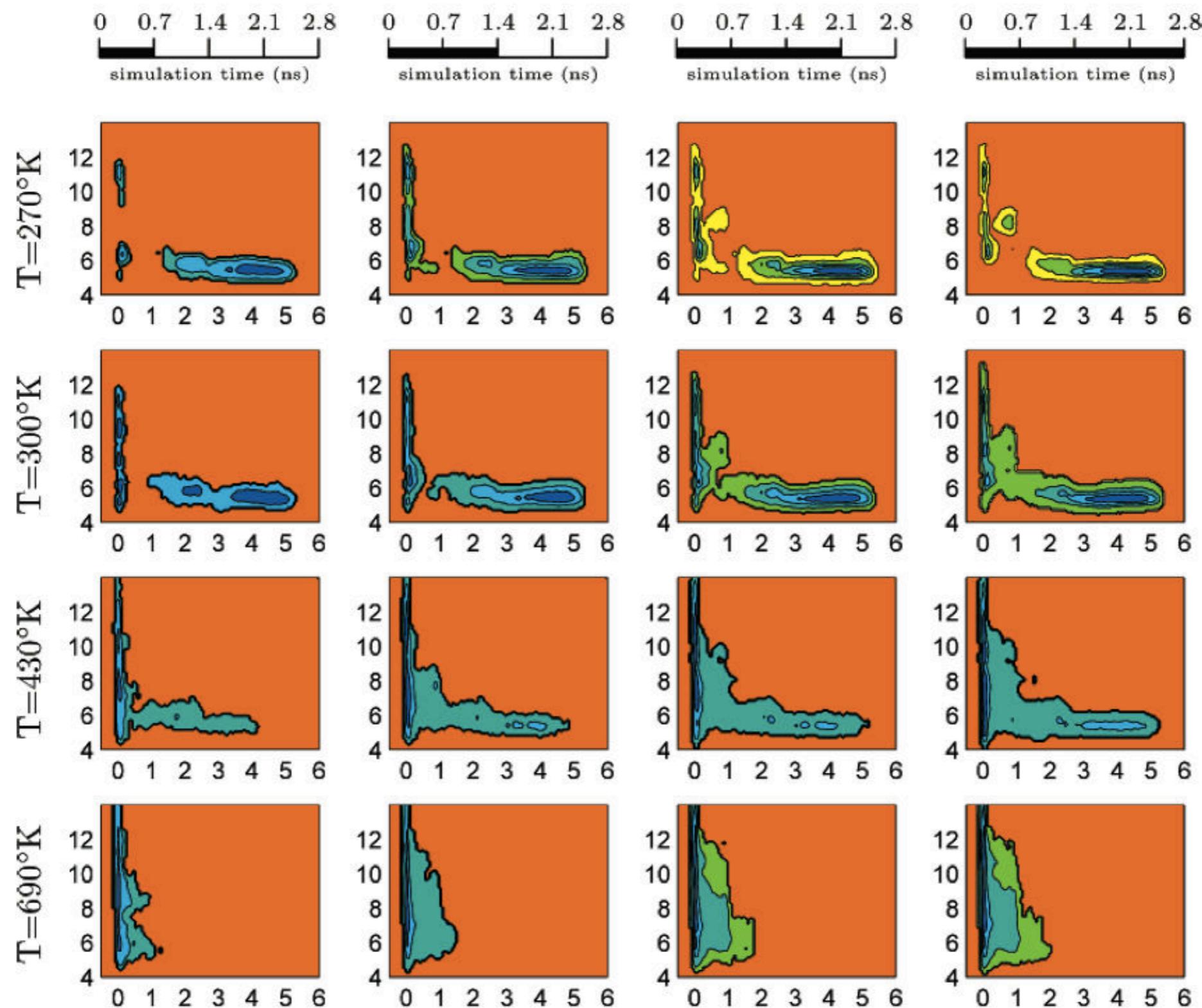
- backbone-backbone hydrogen bonds

$$N_H = \sum_{i \in O} \sum_{j \in H} \frac{1 - \left( \frac{\mathbf{r}_i - \mathbf{r}_j}{d_0} \right)^6}{1 - \left( \frac{\mathbf{r}_i - \mathbf{r}_j}{d_0} \right)^{12}}$$

# Reference PT

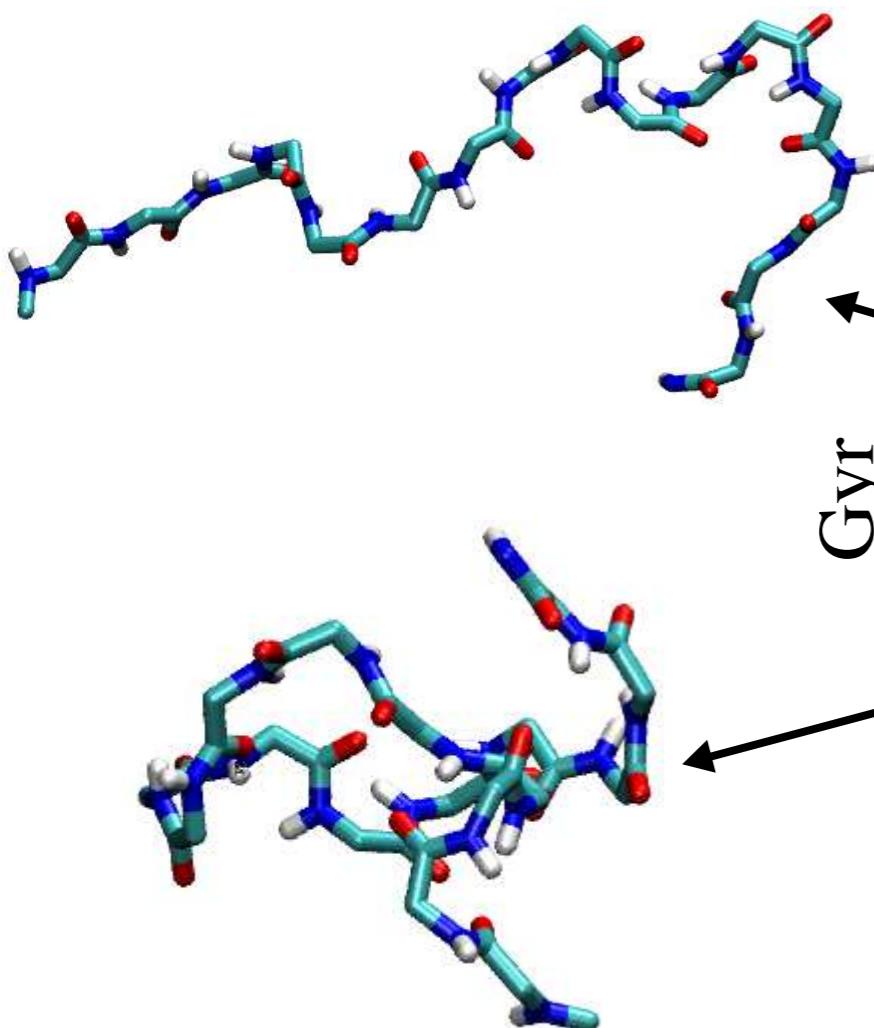


# PTMetaD

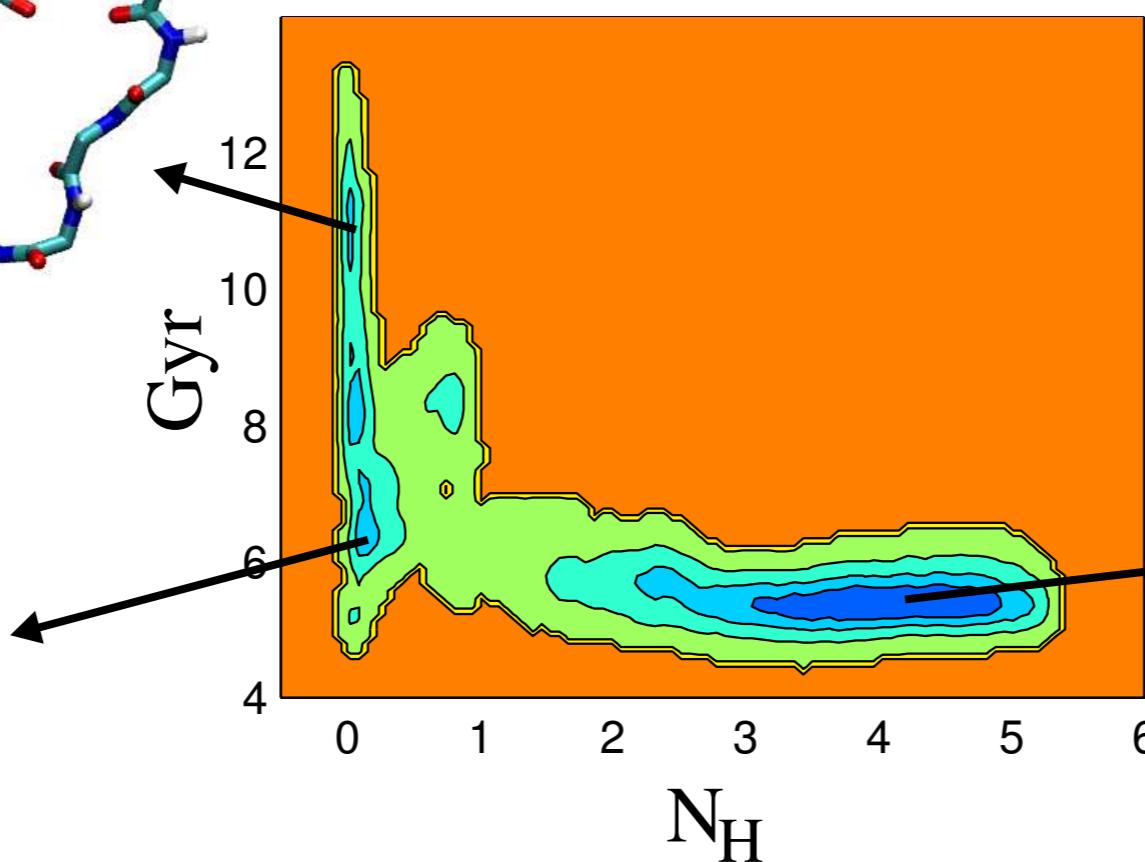


# The free energy

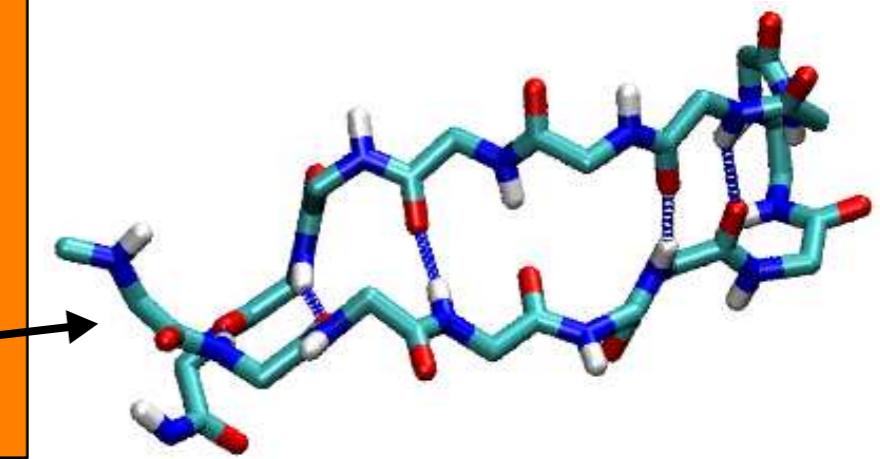
unfolded



molten  
globule



spacing is  $2 k_B T$



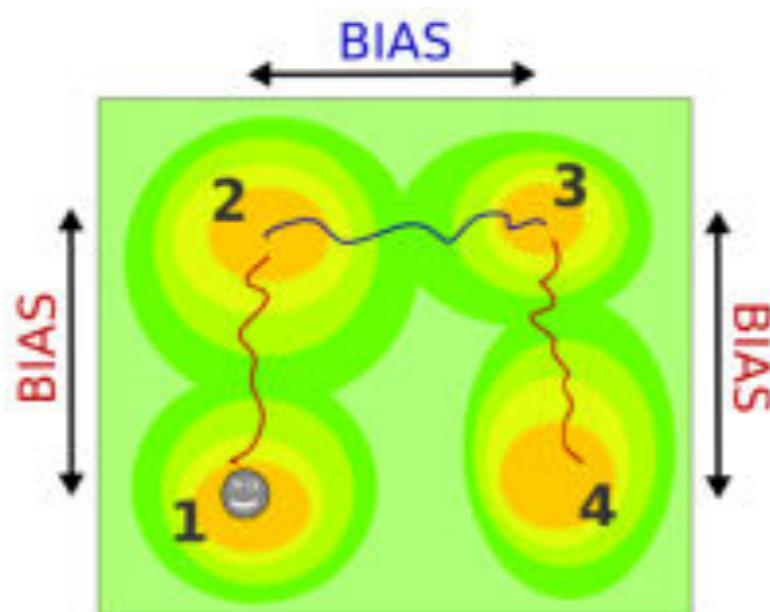
folded

# Bias Exchange Metadynamics

Piana & Lai J Phys Chem B 2007

- N replicas at the same temperature T
- Different CVs and bias potentials
- Modified exchange probability:

$$\Delta_{j,k} = \cancel{(\beta_j - \beta_k)(U(\mathbf{R}_j) - U(\mathbf{R}_k))}$$
$$+ \beta_j [V_G^{(j)}(\mathbf{S}(\mathbf{R}_j), t) - V_G^{(j)}(\mathbf{S}(\mathbf{R}_k), t)]$$
$$+ \beta_k [V_G^{(k)}(\mathbf{S}(\mathbf{R}_k), t) - V_G^{(k)}(\mathbf{S}(\mathbf{R}_j), t)]$$



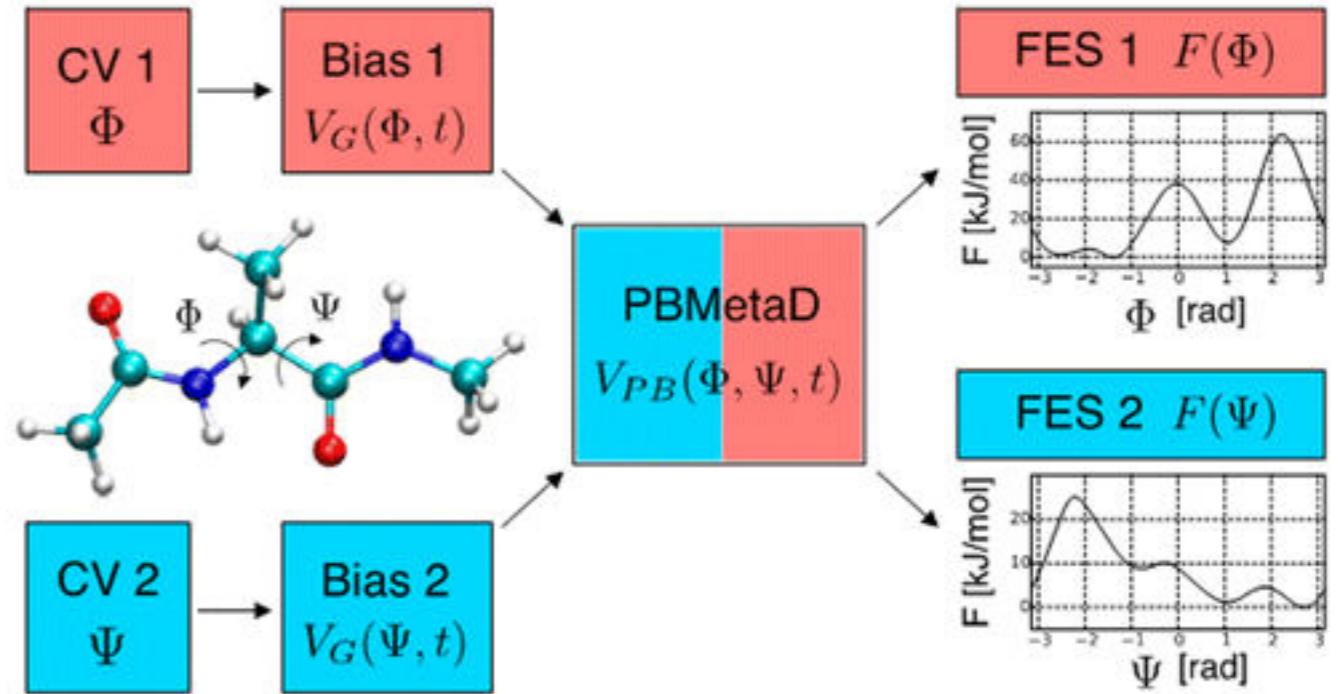
# Parallel Bias Metadynamics

Pfaendtner & Bonomi JCTC 2015

In PBMetaD, we apply multiple low-dimensional bias potentials:

$$V(S_1, t), \dots, V(S_N, t)$$

one at a time:



$$P_t(\mathbf{R}, \boldsymbol{\eta}) \propto \exp \left[ -\beta \left( U(\mathbf{R}) + \sum_i \eta_i V(S_i, t) \right) \right]$$

where  $\boldsymbol{\eta} = (\eta_1, \dots, \eta_N)$  switches on and off (and allows updating) one bias potential at a time

Each bias potential converges to the corresponding free energy:

$$V(S_i, t) \rightarrow -\frac{\Delta T}{T + \Delta T} F(S_i)$$

# Parallel Bias Metadynamics

Pfaendtner & Bonomi JCTC 2015

Since we are not interested in the  $\eta$ -distribution, we can marginalize this variable:

$$P_t(\mathbf{R}) = \int d\eta P_t(\mathbf{R}, \eta) \propto \exp [-\beta (U(\mathbf{R}) + V_{PB}(\mathbf{S}, t))]$$

where:

$$V_{PB}(\mathbf{S}, t) = -\frac{1}{\beta} \log \sum_{i=1}^N \exp - [\beta V(S_i, t)]$$

In order for each bias potential to converge to the corresponding free energy, we need a new rescaling rule:

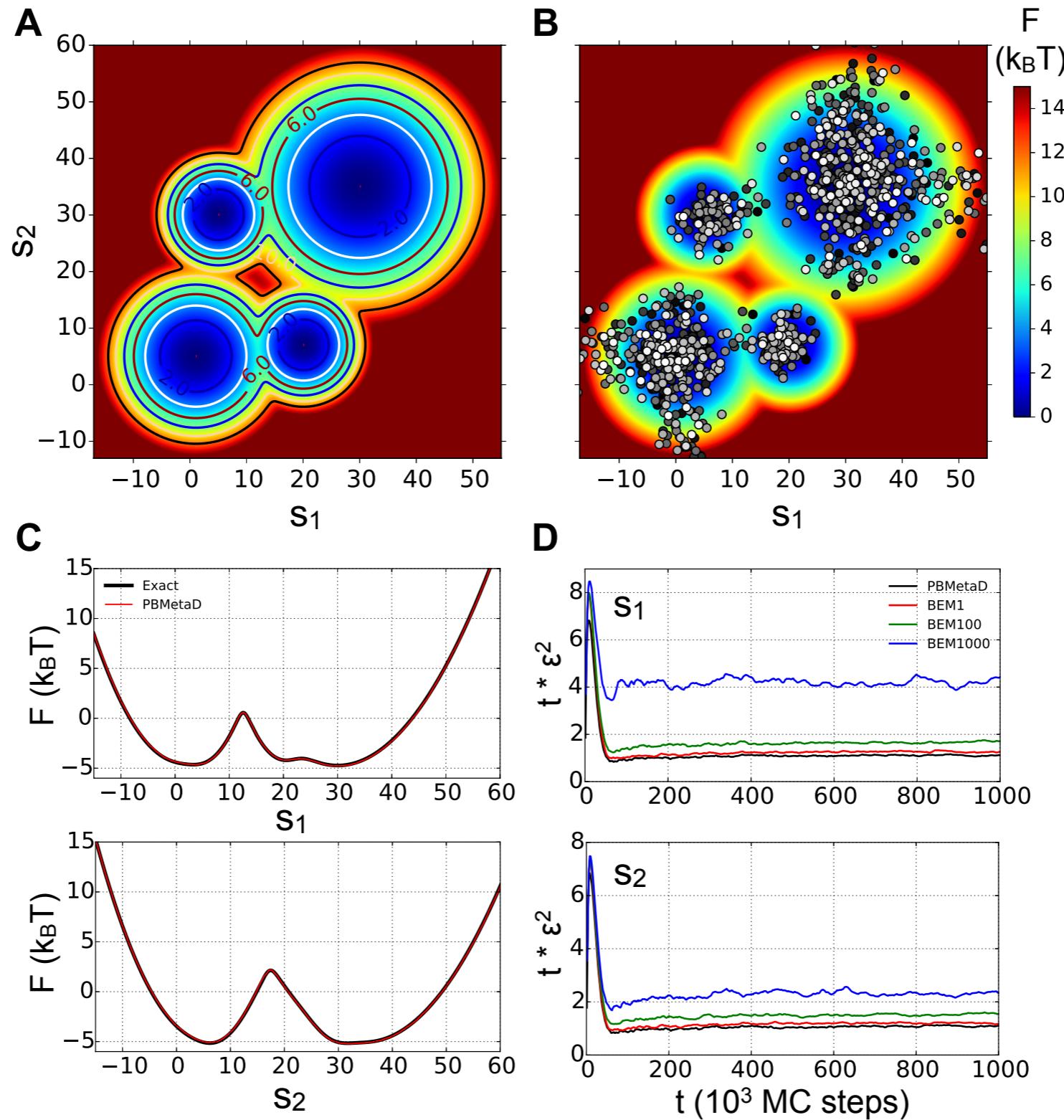
$$\omega_i = \omega_{0,i} e^{-\frac{V(S_i, t)}{k_B \Delta T_i}} P(\eta_i = 1 | \mathbf{R})$$

where:

$$P(\eta_i = 1 | \mathbf{R}) = \frac{\exp [-\beta V(S_i, t)]}{\sum_{j=1}^N \exp [-\beta V(S_j, t)]}$$

# Benchmark on a model system

Pfaendtner & Bonomi JCTC 2015



## Summary III

- Designing smart CVs
- A popular REM: Parallel Tempering
- Combining PT and MetaD
- Using a large number of CVs I: Bias Exchange Metadynamics
- Using a large number of CVs II: Parallel Bias Metadynamics