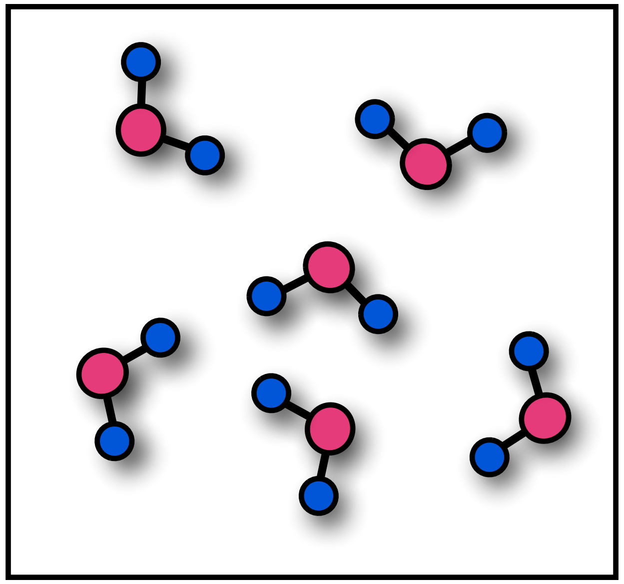
Path-Collective Variables

Winter School for Theoretical Chemistry and Spectroscopy
Han-sur-Lesse, 25-29 November 2019

by
Bernd Ensing

Classical molecular dynamics



Molecular dynamics (MD) simulation

Sampling the distribution and the dynamics of many-particle systems

$$\langle A \rangle = \lim_{t \to \infty} \frac{1}{t} \int_0^t dt' A(\mathbf{r}^N, \mathbf{p}^N)$$

$$\langle A \rangle = \frac{\iint d\mathbf{r}^N d\mathbf{p}^N A(\mathbf{r}^N, \mathbf{p}^N) \exp[-\beta H(\mathbf{r}^N, \mathbf{p}^N)]}{\iint d\mathbf{r}^N d\mathbf{p}^N \exp[-\beta H(\mathbf{r}^N, \mathbf{p}^N)]}$$

time averages
equal the
ensemble averages
(ergodicity hypothesis)

(N,V,T) ensemble

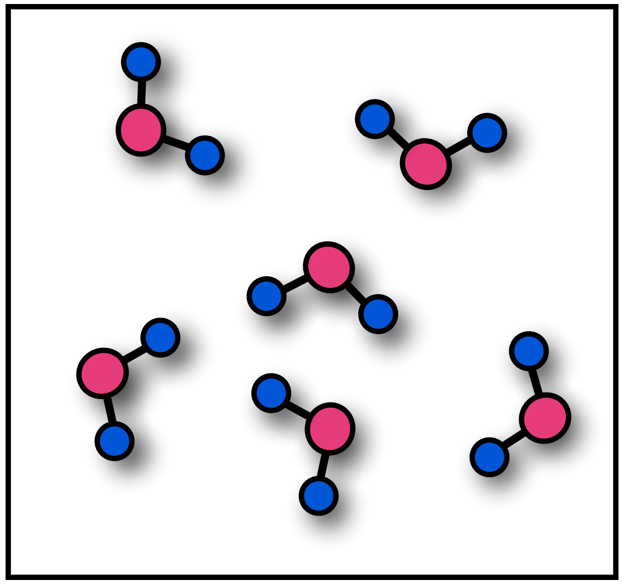
Newtonian equation of motion: $-\frac{dV}{dr_i} = m_i \frac{d^2 r_i}{dt^2}$

Velocity verlet integrator:

$$x(t + \Delta t) = x(t) + v(t)\Delta t + a(t)\frac{(\Delta t)^2}{2}$$

$$v(t + \Delta t) = v(t) + \frac{[a(t) + a(t + \Delta t]}{2}\Delta t$$

Classical molecular dynamics



Molecular dynamics (MD) simulation

Sampling the distribution and the dynamics of many-particle systems

Classical MD means:

1. Newtonian dynamics of the nuclei
2. Interaction potentials are sums of empirical (pair) potentials (forcefield)

$$E^{\mathrm{pot}} = E^{\mathrm{bonded}} + E^{\mathrm{non-bonded}}$$

$$E^{\mathrm{pot}} = \sum_{\mathrm{bonds}} V^{\mathrm{bond}} + \sum_{\mathrm{angles}} V^{\mathrm{angle}} \sum_{\mathrm{dihedrals}} V^{\mathrm{torsion}} + \sum_{\mathrm{i}} \sum_{\mathrm{j \neq i}} V^{\mathrm{LJ}} + \sum_{\mathrm{i}} \sum_{\mathrm{j \neq i}} V^{\mathrm{Coulomb}}$$

$$V^{\mathrm{harm}}(r) = \frac{1}{2} k_b (r - r_0)^2$$

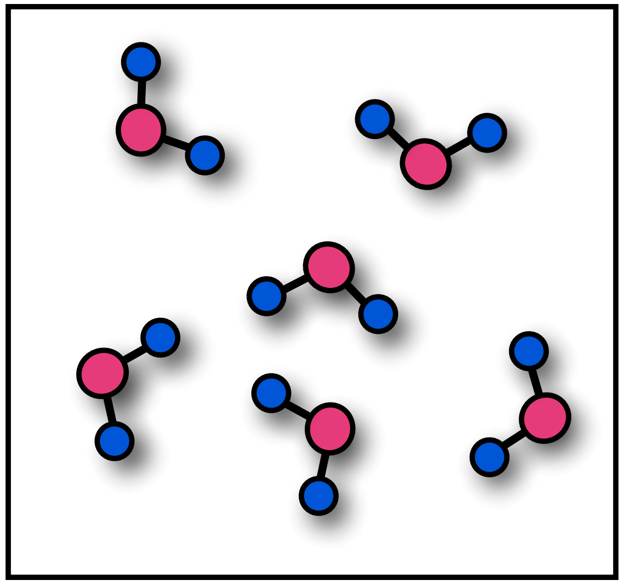
$$V^{\mathrm{harm}}(\theta) = \frac{1}{2} k_\alpha (\theta - \theta_0)^2$$

$$V^{\mathrm{Fourier}}(\omega) = \sum_n \frac{1}{2} V_n [1 + \cos(n\omega - \gamma)]$$

$$V^{\mathrm{LJ}}(r) = -4\epsilon \left(\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6} \right)$$

$$V^{\mathrm{Coulomb}}(r) = \frac{1}{4\pi\epsilon_0} \frac{q_1 q_2}{r}$$

Classical molecular dynamics



Molecular dynamics (MD) simulation

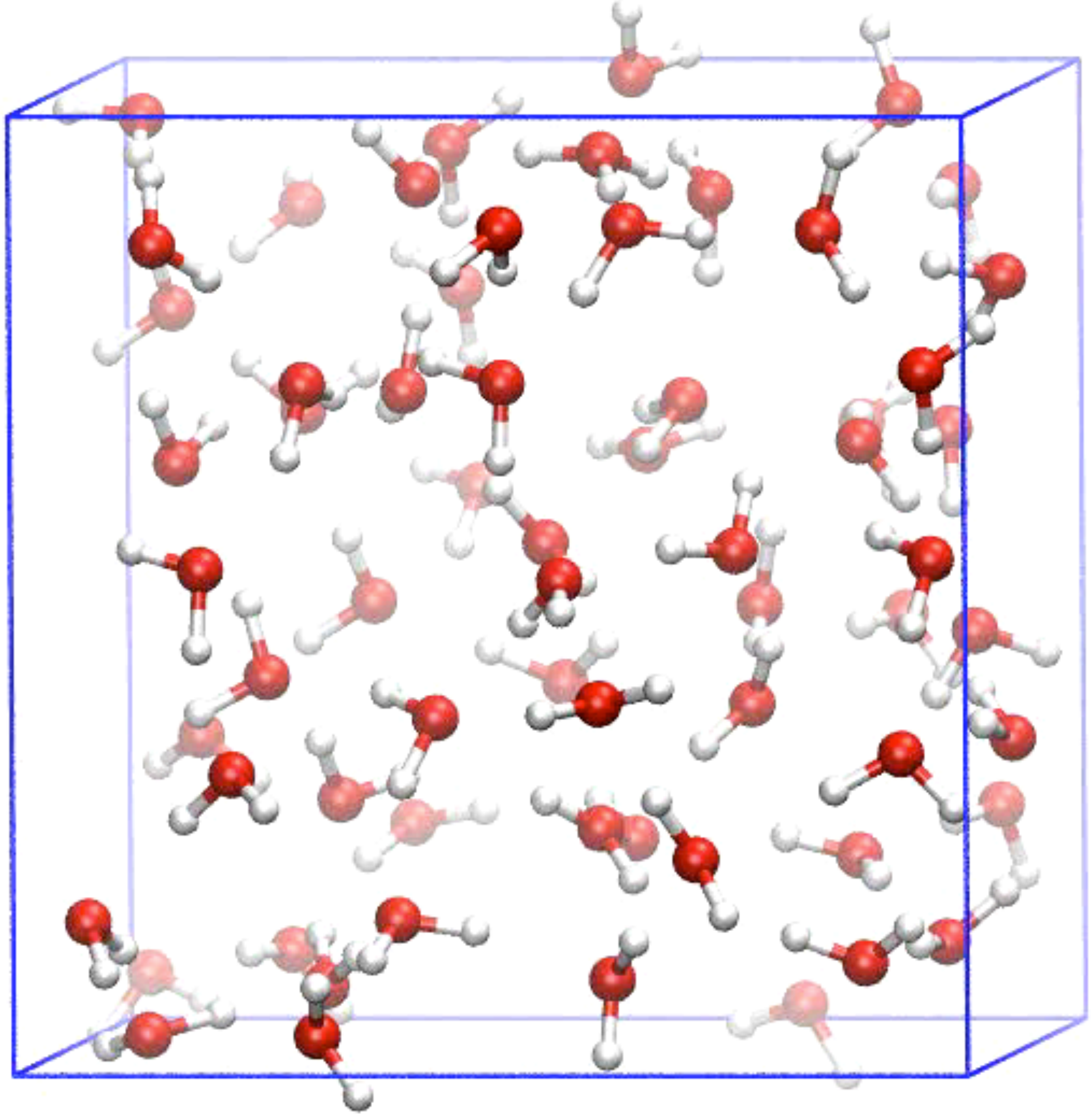
Sampling the distribution and the dynamics of many-particle systems

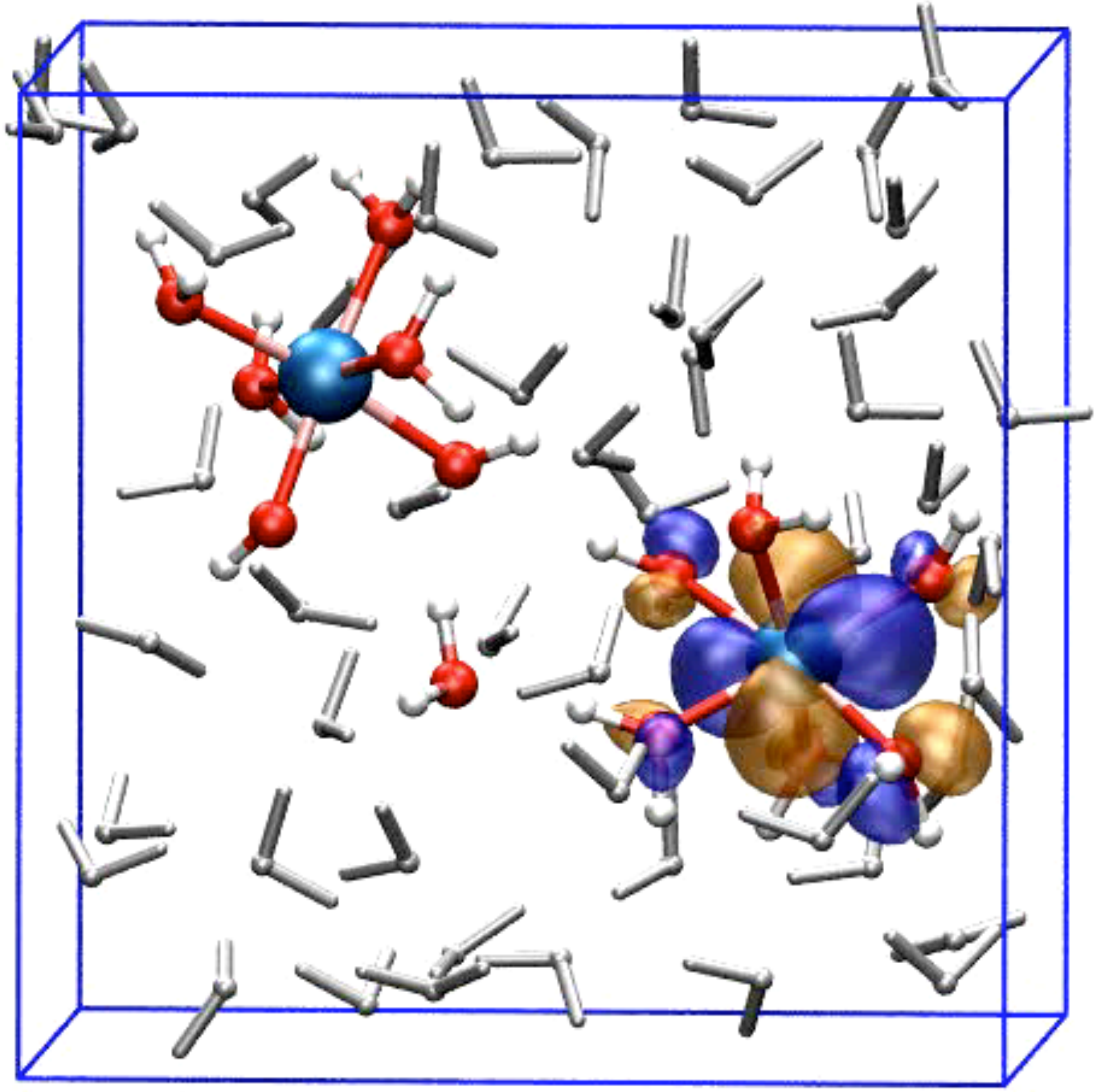
Ab initio MD, First principles MD, DFT-MD, Born-Oppenheimer MD, Car-Parrinello MD

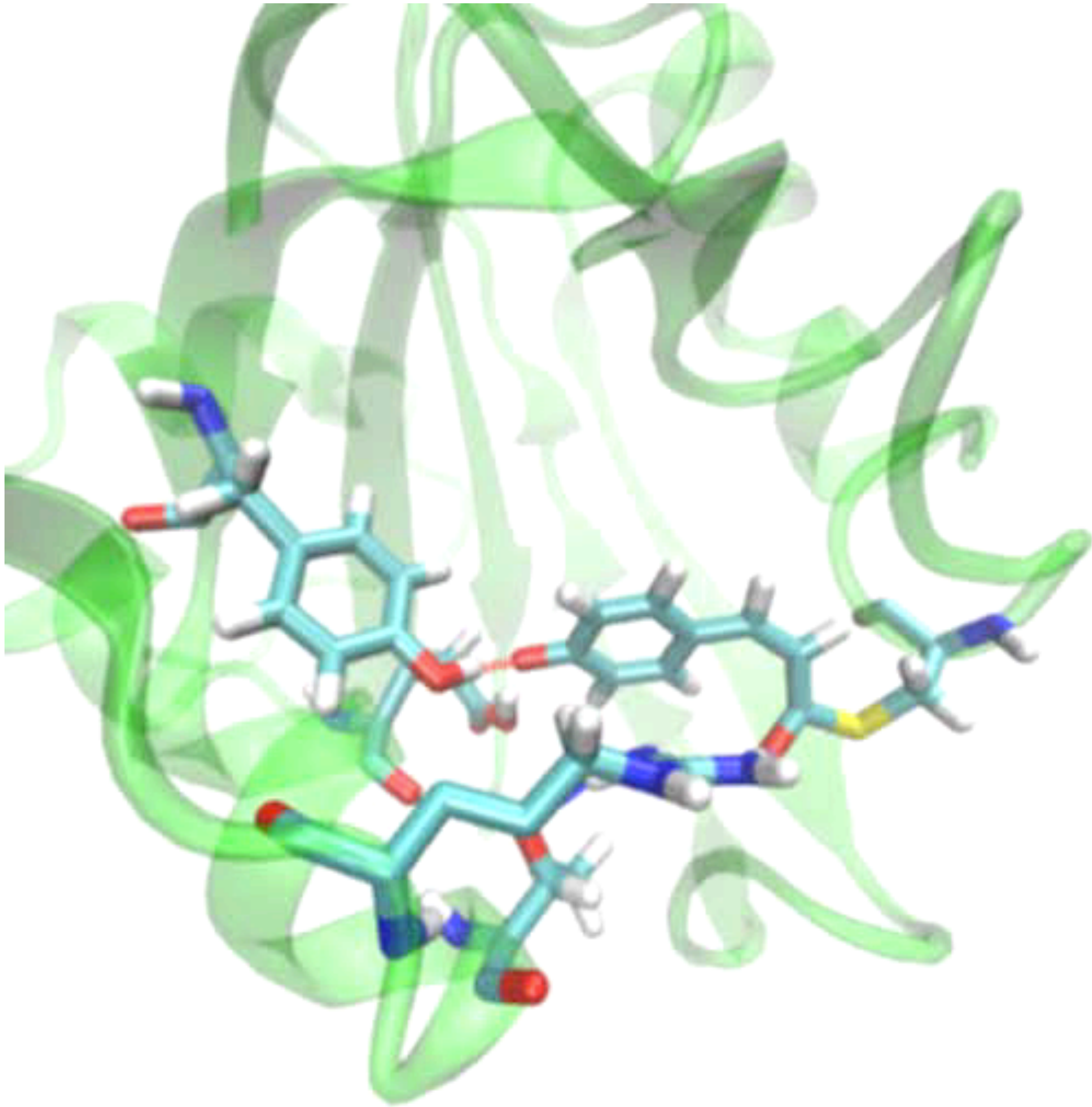
- Interaction potential from electronic structure calculation
- quantum mechanical description of the electronic structure (DFT)
- Born-Oppenheimer approximation (electronic ground-state)
- Nuclear dynamics is still Newtonian!

$$H = -\sum_I \frac{\hbar^2}{2M_I}\nabla_I^2 - \sum_i \frac{\hbar^2}{2m_e}\nabla_i^2 + \frac{1}{4\pi\epsilon_0}\sum_{i<j}\frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|} - \frac{1}{4\pi\epsilon_0}\sum_{I,i}\frac{e^2 Z_I}{|\mathbf{R}_I - \mathbf{r}_i|} + \frac{1}{4\pi\epsilon_0}\sum_{I<J}\frac{e^2}{|\mathbf{R}_I - \mathbf{R}_J|}$$

$$M_I \frac{d^2\mathbf{R}}{dt^2} = -\nabla_I \min_{\psi_0}\{\langle\psi_0|H_e|\psi_0\rangle\}$$

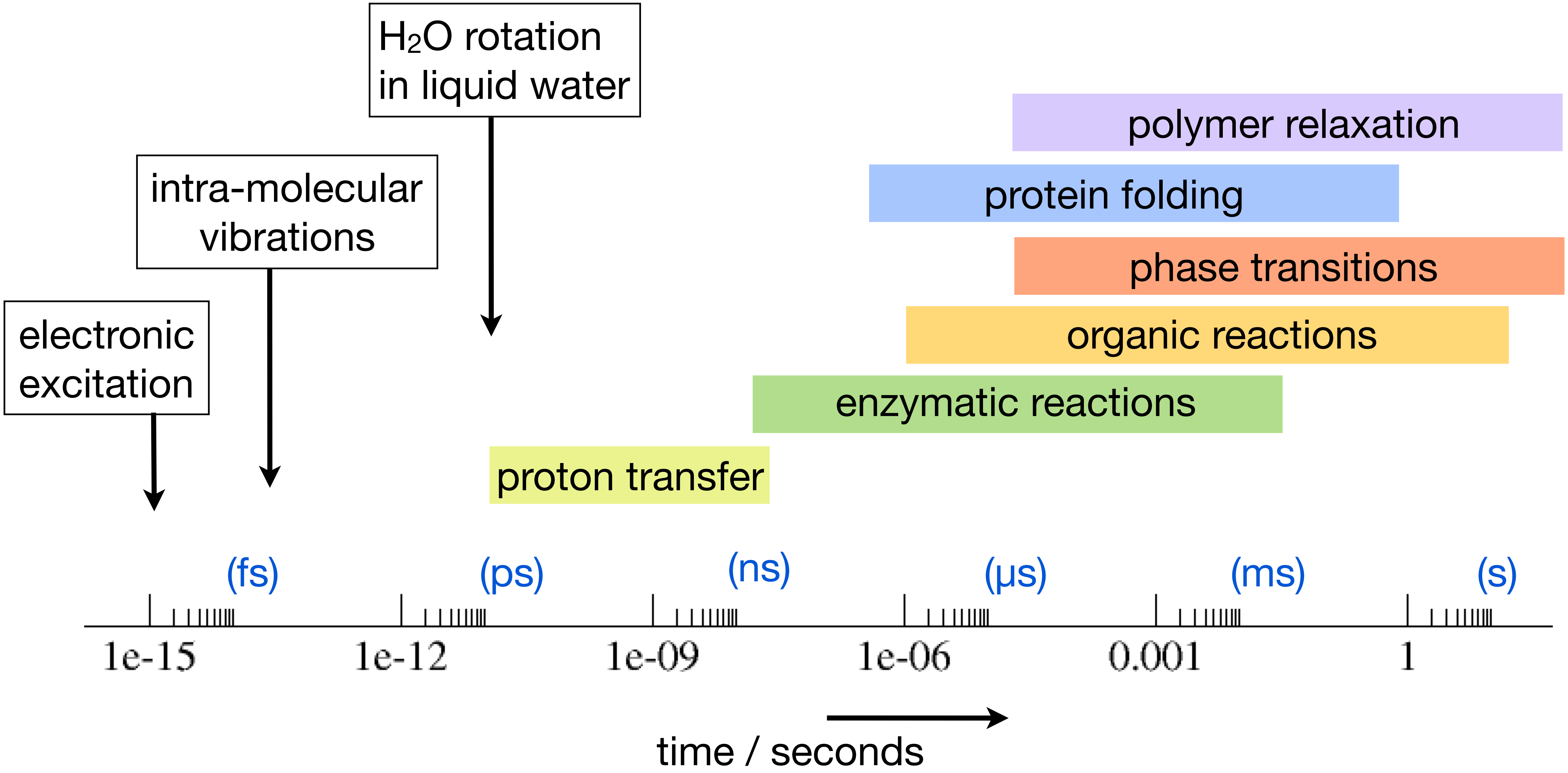






Rare events

In real life (experiment)



H_2O rotation in liquid water

intra-molecular vibrations

electronic excitation

polymer relaxation

protein folding

phase transitions

organic reactions

enzymatic reactions

proton transfer

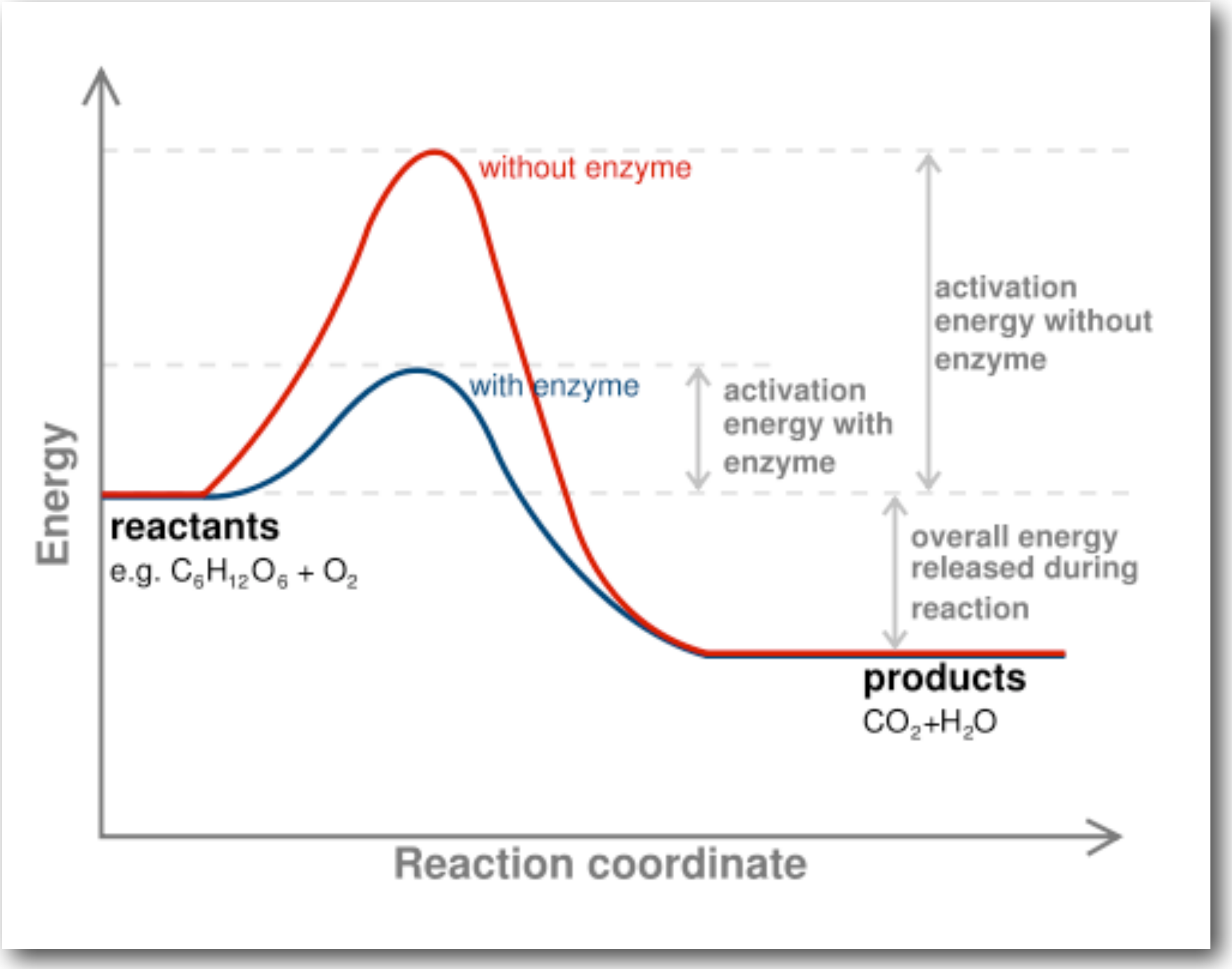
time / seconds

Rare events

In real life (experiment)

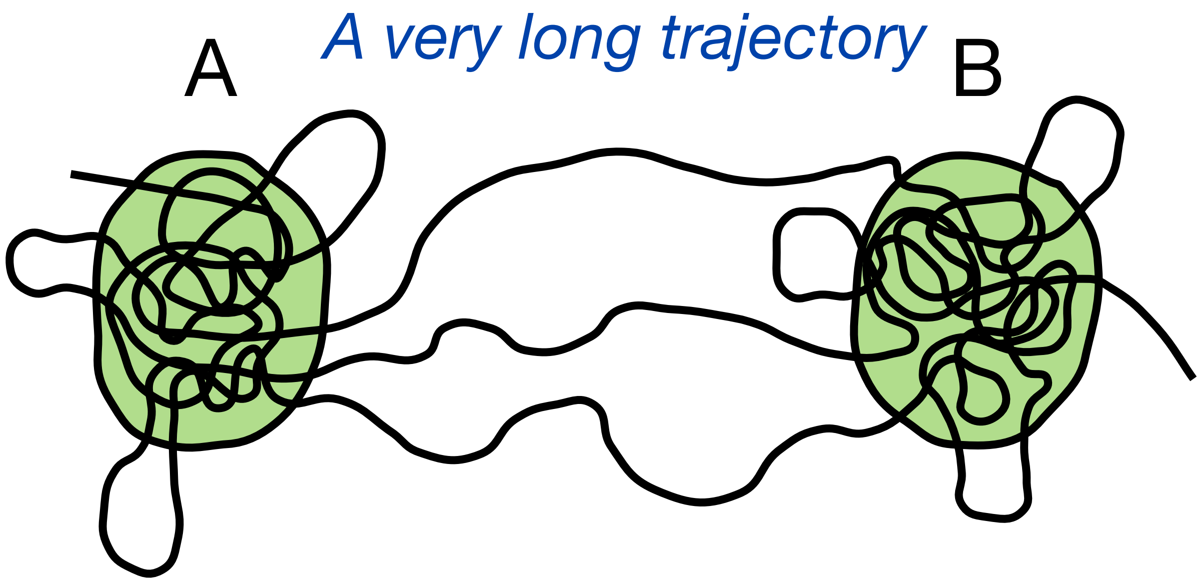
- no direct access to free energy
- reaction rate measurement

$k = k_0\, e^{-\Delta G/kT}$



Rare Events

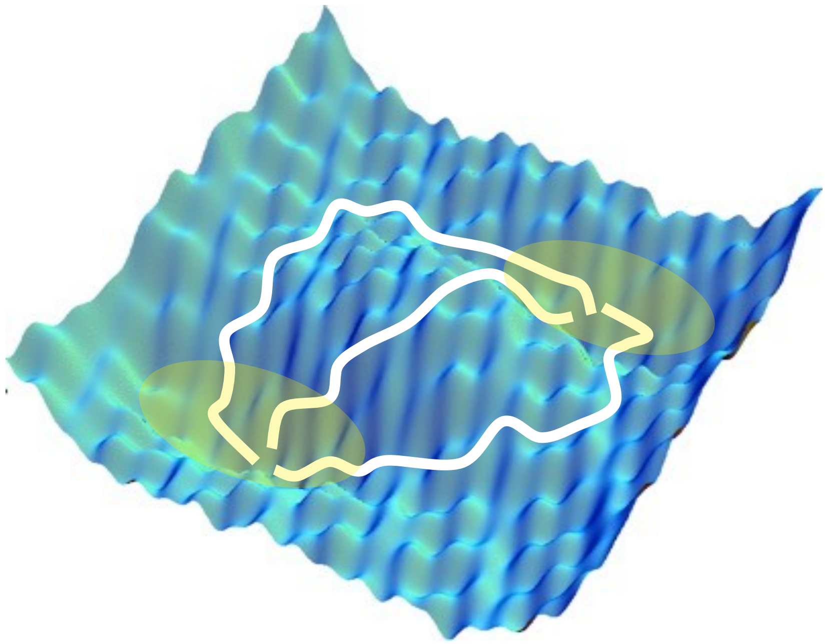
- Molecular transitions are not rare in every day life
- They are only rare events with respect to the femto-second timescale of atomic motions
- Modeling activated transitions by straightforward simulation would be extremely costly (takes forever)
- Advanced methods: Transition state theory (TST), Bennet-Chandler (Reactive Flux) approach, Transition Path Sampling (TPS), Free energy methods, Parallel Tempering, String Method,....



- A and B are (meta-)stable states, i.e. attractive basins
- transitions between A and B are rare
- transitions can happen fast
- system looses memory in A and B

Rare Events

Free energy landscape



P.G. Bolhuis, D. Chandler,
C. Dellago, P.L. Geissler

Annu. Rev. Phys. Chem 2002

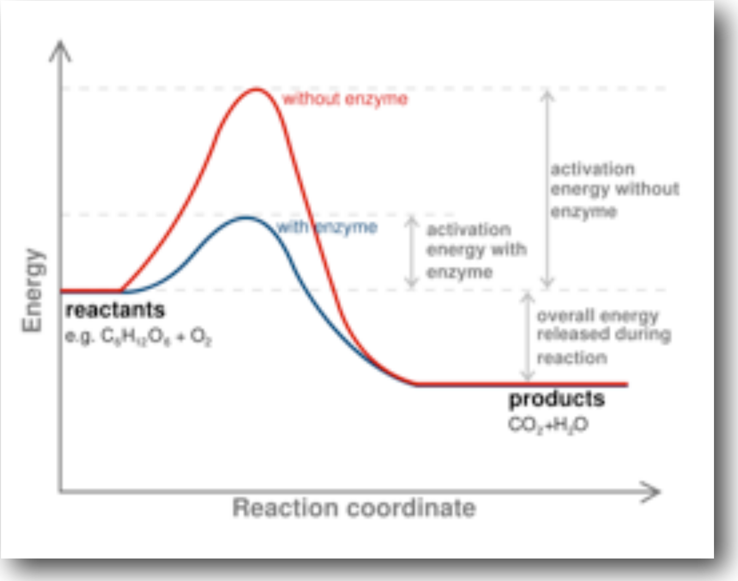
The free energy has local minima separated by barriers

Many possible transition paths via meta-stable states

Projection on reaction coordinate shows FE profile with transition state barrier

Reaction rate depends exponentially on the barrier height

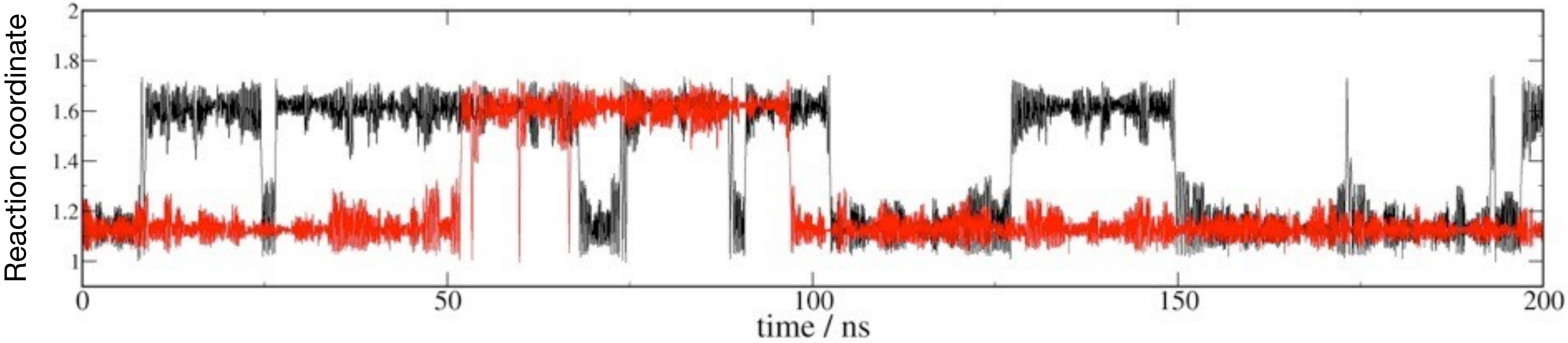
Transition State



$$k = k_0\, e^{-\Delta G/kT}$$

Arrhenius equation

A long trajectory



Some rare event methods

Added bias potential

(adaptive) Umbrella Sampling
Conformational flooding
Local Elevation
Taboo search
Wang-Landau sampling
Metadynamics

Thermodynamic Integration

Constrained MD (Blue Moon)
Steered MD (Jarzynski's method)
Adaptive Bias Force
Free energy perturbation

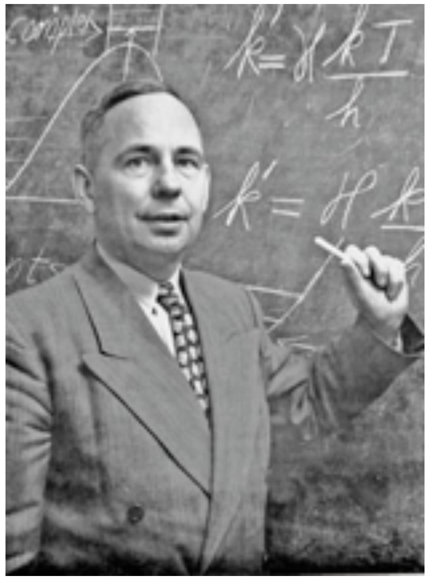
Finding/sampling a reaction path

String method
Minimum action paths
Mile-stoning
Transition Path Sampling
Transition Interface Sampling
Forward Flux Sampling

Raising the temperature

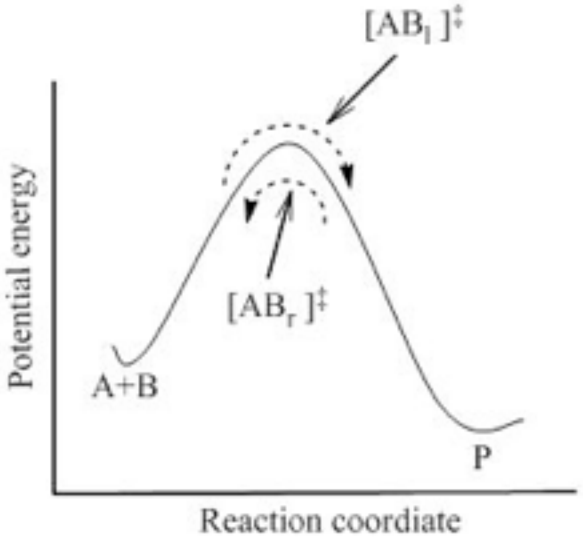
Replica Exchange (Parallel Tempering)
Temperature-accelarated dynamics
Adiabatic molecular dynamics
Multicanonical methods

Transition State Theory



Henry Eyring

- TST ≠ TPT
- Derived by Henry Eyring and by Meredith Gwynne Evans and Michael Polanyi in 1935
- Explains Arrhenius equation (1889) $k = Ae^{-E_a/RT}$
- Activated complex is in quasi equilibrium with reactants
- All reactants that reach the activated complex state, continue to products



$$A + B \leftrightarrow [AB]^\ddagger \rightarrow C \qquad K^\ddagger = \frac{[AB]^\ddagger}{[A][B]}$$

probability to reach the transition state

$$k = k^\ddagger K^\ddagger = \kappa \frac{k_B T}{h} e^{-\Delta G^\ddagger/RT}$$

frequency

- rate is frequency of unstable mode times the TS probability

- transmission factor is $\kappa=1$ in TST (correction for barrier recrossings)

Some concepts

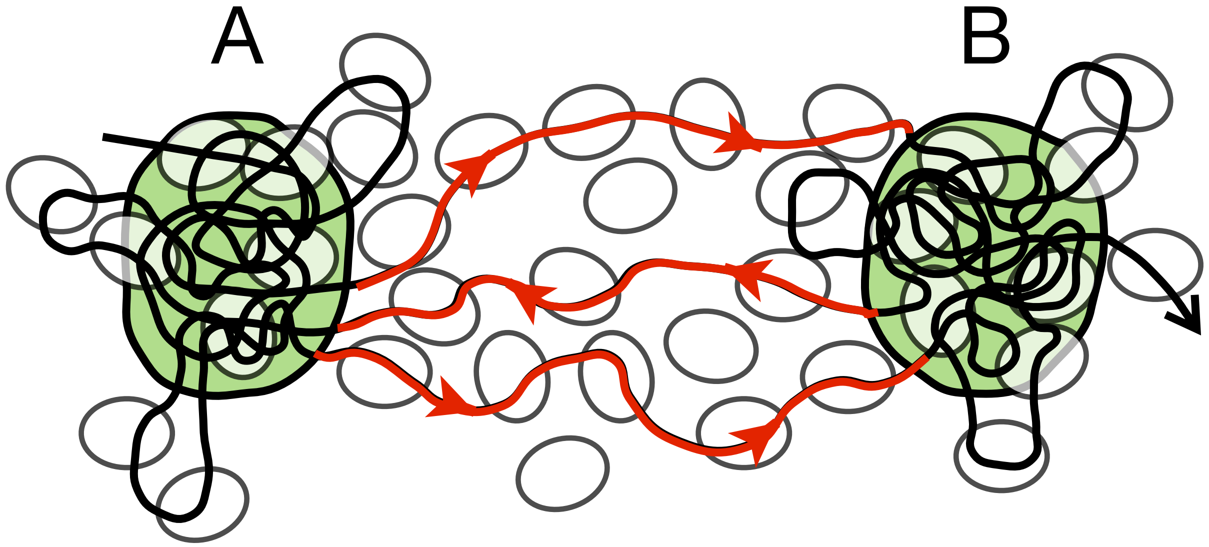
Lars Onsager: "Splitting probability"

Committor function: p_B is the probability that a trajectory leaving from micro-state i arrives in B before A

The committor of a micro-state (or configuration) is computed by initializing a large number of trajectories from this state with random (Boltzmann distributed) velocities and recording how many arrive in B.

Micro-states close to A can be expected to have $0<p_B<<1$ and close to B have $0<<p_B<1$

$p_B = 1 - p_A$



In a sense, the committor function is the perfect reaction coordinate (although expensive and without mechanistic insight).

Some concepts

Lars Onsager:
“Splitting probability”

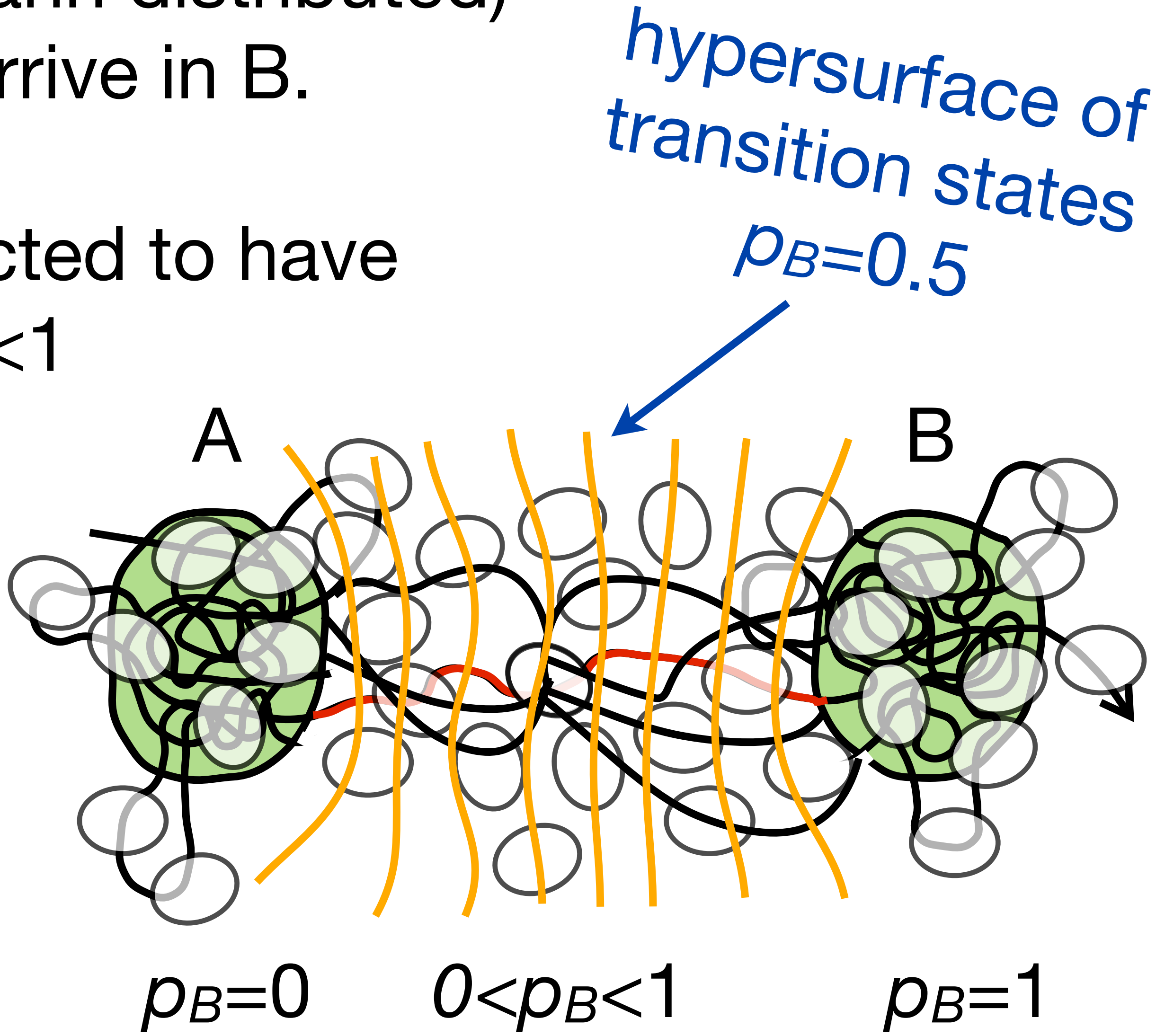
Committor function: p_B is the probability that a trajectory leaving from micro-state i arrives in B before A

The committor of a micro-state (or configuration) is computed by initializing a large number of trajectories from this state with random (Boltzmann distributed) velocities and recording how many arrive in B.

Micro-states close to A can be expected to have $0<p_B<<1$ and close to B have $0<<p_B<1$

$p_B = 1 - p_A$

In a sense, the committor function is the perfect reaction coordinate (although expensive and without mechanistic insight).



Rare event simulation

Macroscopic phenomenological theory

Chemical reaction: $\mathrm{A} \longleftrightarrow \mathrm{B}$

$$\frac{dc_A(t)}{dt} = -k_{A\to B}c_A(t) + k_{B\to A}c_A(t)$$

$$\frac{dc_B(t)}{dt} = +k_{A\to B}c_A(t) - k_{B\to A}c_A(t)$$

Total number of molecules:

$$\frac{d[c_A(t) + c_B(t)]}{dt} = 0$$

Equilibrium:

$$\frac{dc_A(t)}{dt} = \frac{dc_B(t)}{dt} = 0 \qquad \frac{\langle c_A \rangle}{\langle c_B \rangle} = \frac{k_{B\to A}}{k_{A\to B}}$$

Rare event simulation

Macroscopic phenomenological theory

Make a small perturbation:

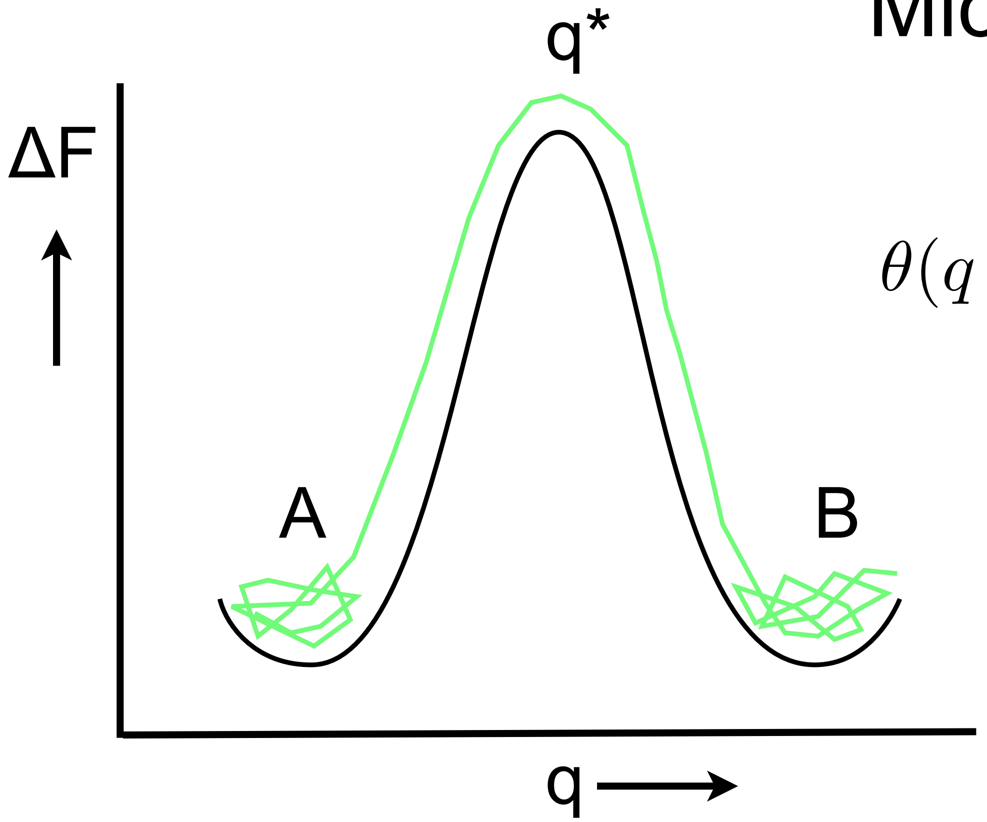
$$c_A(t) = \langle c_A \rangle + \Delta c_A(t) \qquad\qquad c_B(t) = \langle c_B \rangle + \Delta c_A(t)$$

$$\frac{d\Delta c_A(t)}{dt} = -k_{A\to B}\Delta c_A(t) - k_{B\to A}\Delta c_A(t)$$

$$\begin{aligned} \Delta c_A(t) &= \Delta c_A(0)\exp[-(k_{A\to B} + k_{B\to A})t] \\ &= \Delta c_A(0)\exp[-t/\tau] \end{aligned}$$

$$\begin{aligned} \tau &= \left(k_{A\to B} + k_{B\to A}\right)^{-1} \\ &= k_{A\to B}^{-1}\left(1 + \langle c_A \rangle / \langle c_B \rangle\right)^{-1} = \frac{\langle c_B \rangle}{k_{A\to B}} \end{aligned}$$

Rare event simulation



Microscopic linear response theory

$$\theta(q-q^*) = \begin{cases} 0 & \text{if} \quad q-q^* < 0 \quad \text{(Reactant A)} \\ 1 & \text{if} \quad q-q^* > 0 \quad \text{(Product B)} \end{cases}$$

Perturbation

add bias to increase concentration c_A

$$H = H_0 - \epsilon g_A(q-q^*)$$

$$g_A(q-q^*) = 1 - \theta(q-q^*) = \theta(q^*-q)$$

$$\Delta c_A = \langle c_A \rangle_\epsilon - \langle c_A \rangle_0$$

$\langle g_A \rangle$ → probability to be in state A

$$\Delta c_A = \langle g_A \rangle_\epsilon - \langle g_A \rangle_0$$

Reactive Flux Method

$$R \longleftrightarrow P$$

$$\frac{dc_R}{dt} = -k_R c_R(t) + k_P c_P(t)$$

$$\frac{dc_P}{dt} = k_R c_R(t) - k_P c_P(t)$$

equilibrium

$$k_R \langle c_R \rangle = k_P \langle c_P \rangle$$

$$\Theta(x - x^{\ddagger}) = \begin{cases} 0 & \text{at the reactant side of the transition state } x^{\ddagger} \\ 1 & \text{at the product side of the transition state } x^{\ddagger} \end{cases}$$

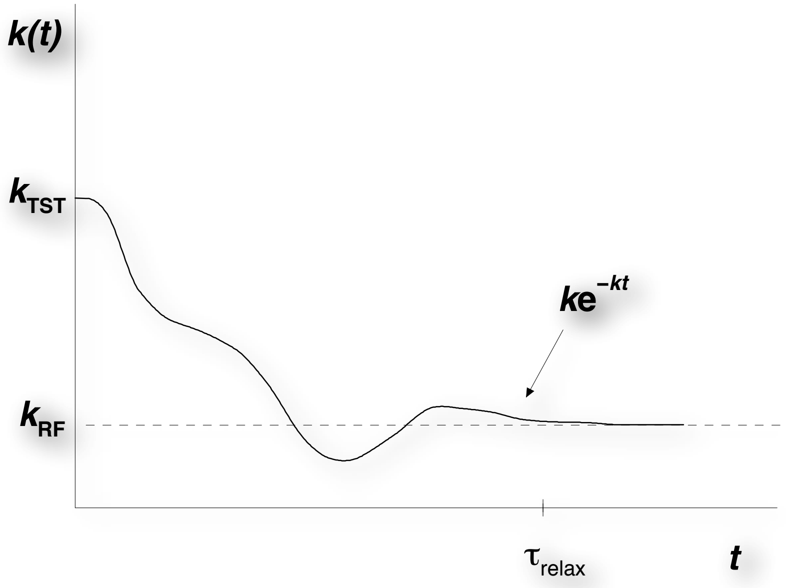
$\left\langle \Theta(x - x^{\ddagger}) \right\rangle$ equilibrium

$\Delta\Theta(x(0) - x^{\ddagger})$ deviation at t=0

$$\Delta\Theta(x - x^{\ddagger}) = \Theta(x - x^{\ddagger}) - \left\langle \Theta(x - x^{\ddagger}) \right\rangle$$

$$\Delta\Theta(x(t) - x^{\ddagger}) = \Delta\Theta(x(0) - x^{\ddagger}) \exp[-(k_R + k_P)t]$$

$$k_R(t) = \frac{\left\langle \delta(x(0) - x^{\ddagger})\dot{x}(0)\Theta(x(t) - x^{\ddagger}) \right\rangle}{\left\langle \Theta(-x(0) + x^{\ddagger}) \right\rangle}$$



Reactive Flux Method

$$k_R(t) = \frac{\left\langle \delta(x(0)-x^\ddagger)\dot{x}(0)\Theta(x(t)-x^\ddagger)\right\rangle}{\left\langle \Theta(-x(0)+x^\ddagger)\right\rangle} \; \frac{\left\langle \delta(x(0)-x^\ddagger)\dot{x}(0)\Theta(\dot{x}(0))\right\rangle \left\langle \delta(x(0)-x^\ddagger)\right\rangle}{\left\langle \delta(x(0)-x^\ddagger)\dot{x}(0)\Theta(\dot{x}(0))\right\rangle \left\langle \delta(x(0)-x^\ddagger)\right\rangle}$$

$$k_R(t) = \frac{\left\langle \delta(x(0)-x^\ddagger)\dot{x}(0)\Theta(x(t)-x^\ddagger)\right\rangle}{\left\langle \delta(x(0)-x^\ddagger)\dot{x}(0)\Theta(\dot{x}(0))\right\rangle} \times$$

$$\times \frac{\left\langle \delta(x(0)-x^\ddagger)\dot{x}(0)\Theta(\dot{x}(0))\right\rangle}{\left\langle \delta(x(0)-x^\ddagger)\right\rangle} \times \frac{\left\langle \delta(x(0)-x^\ddagger)\right\rangle}{\left\langle \Theta(-x(0)+x^\ddagger)\right\rangle}$$

$$k_R = \kappa \frac{1}{2} \langle |\dot{x}| \rangle_\ddagger \frac{\exp\left[-\beta \Delta G(x^\ddagger)\right]}{\int_{-\infty}^{x^\ddagger} \mathrm{d}x \exp\left[-\beta \Delta G(x)\right]}$$

(Chandler notation)

Transition State Theory: $\kappa = 1$

Free energy methods

It is very (very) difficult to compute (or measure) absolute thermodynamic properties, such as free energy and entropy, that depend on the size of the phase space.

But we can compute relative free energies, in particular free energy differences between thermodynamic states.

Reaction equilibrium constants $A \longleftrightarrow B$

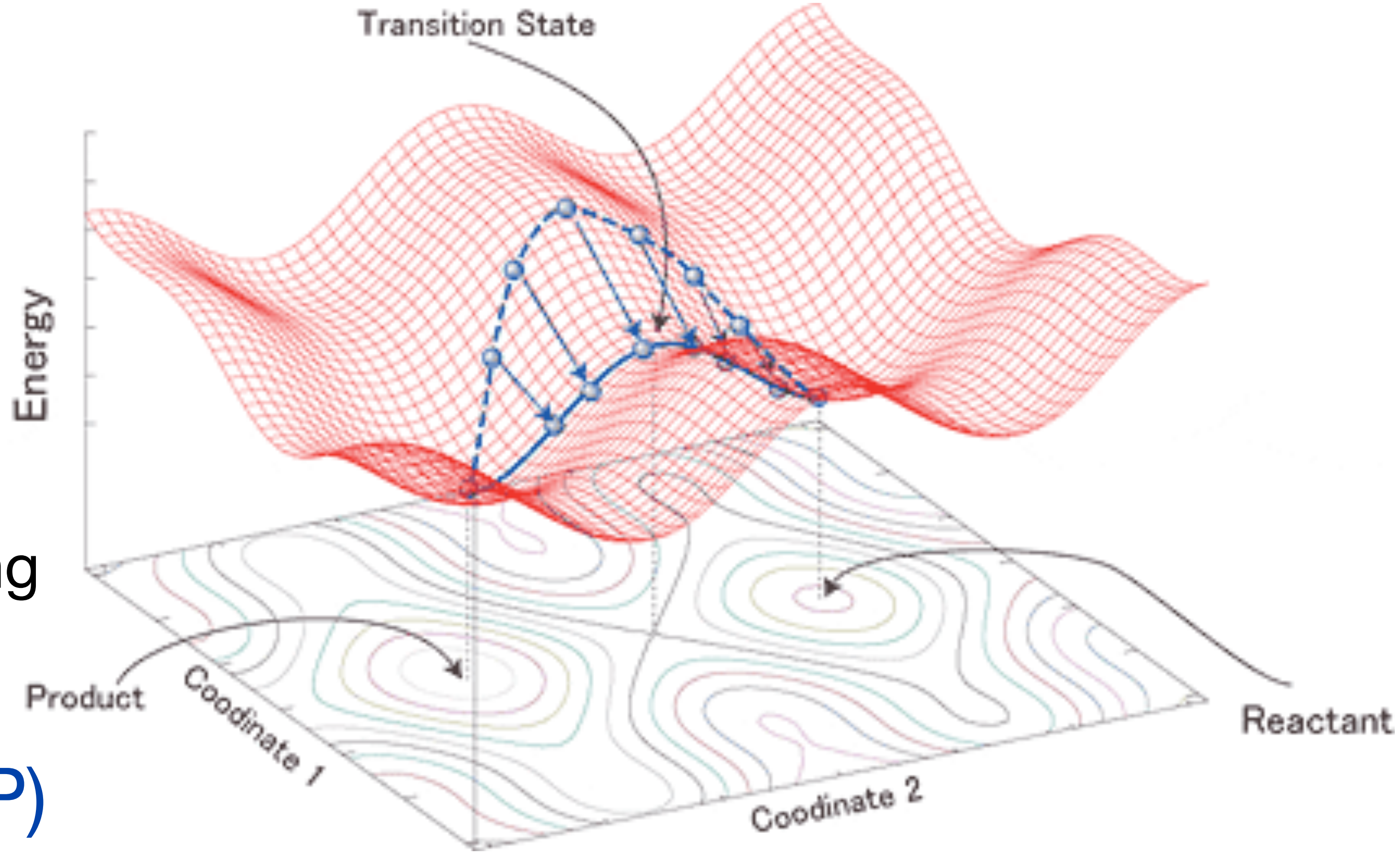
$$K = \frac{[B]}{[A]} = \frac{p_B}{p_A} = \exp\left[-\beta(G_B - G_A)\right]$$

Examples:
- Chemical reactions, catalysis, isomerization, etc...
- Protein folding, ligand binding affinity, protein-protein association
- Phase diagrams, coexistence lines, critical points, transitions

Path finding

Minimum Energy Path (MEP)

- string method, Nudge Elastic Band, ...
- start from guess path
- string of equidistant nodes from A to B
- compute gradients at nodes
- move nodes steepest descent
- maintain equidistant nodes
- converged: forces perpendicular to string are zero



Minimum Free Energy Path (MFEP)

- finite temperature string method
- perform constrained MD simulation at each node
- compute average force of constraint
- steepest descent perpendicular to string
- converged: perpendicular forces are zero
- free energy profile estimated nu integration of constraint force along the string from A to B

Exploring the free energy landscape

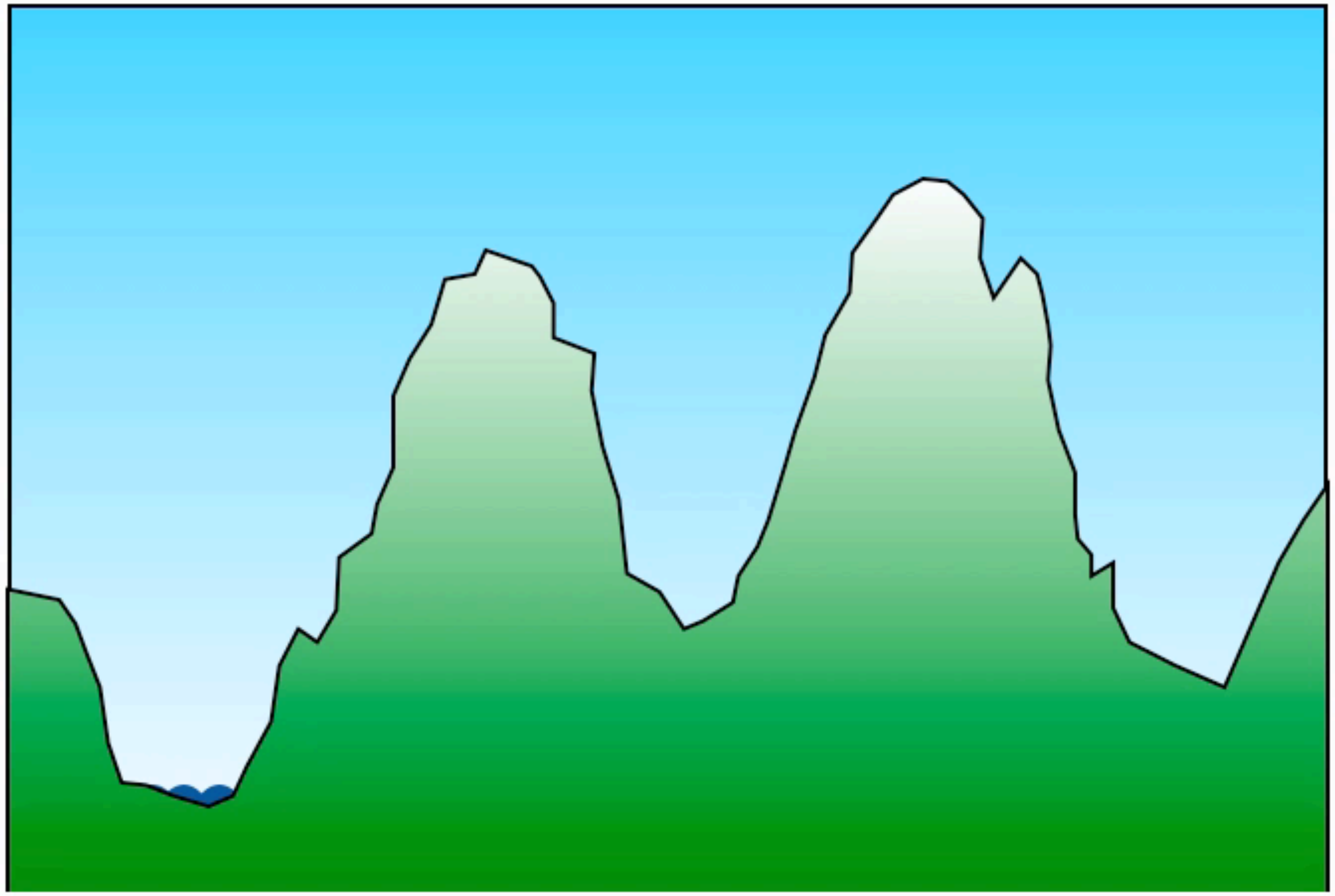
Escaping free-energy minima,
Laio and Parrinello, PNAS (2002)

$$V(t,s) = \sum_{t'<t} H_{t'} \prod_{\alpha} \exp\left[\frac{-(s_\alpha - s_\alpha^{t'})^2}{2\delta_\alpha^2 W^2}\right]$$

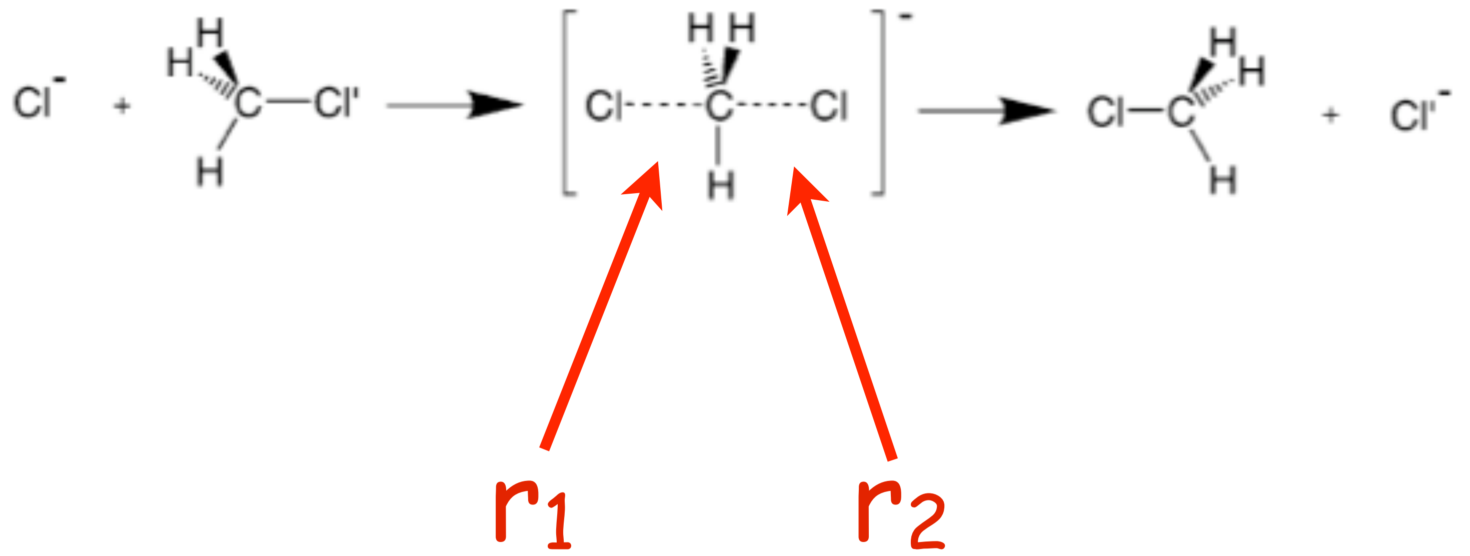
- enhance sampling
- probe free energy landscape
- multiple collective variables

A Dutch perspective on escaping free energy minima

Metadynamics, Laio and Parrinello, PNAS (2002)

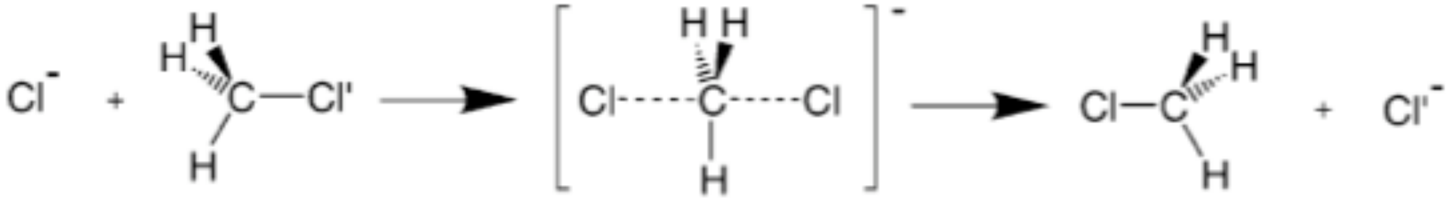


simple example: S_N2 reaction

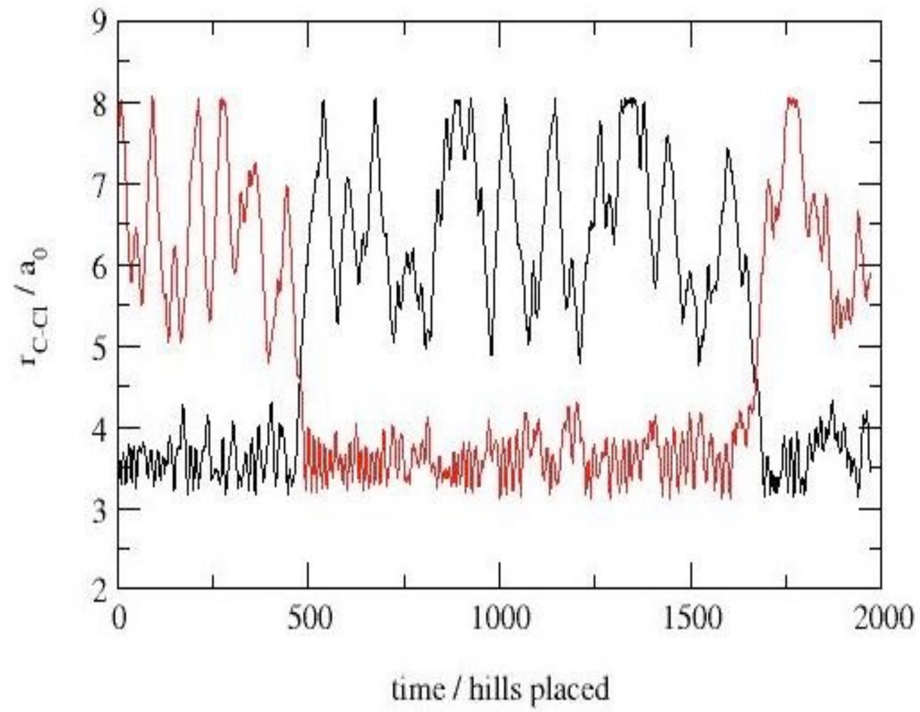


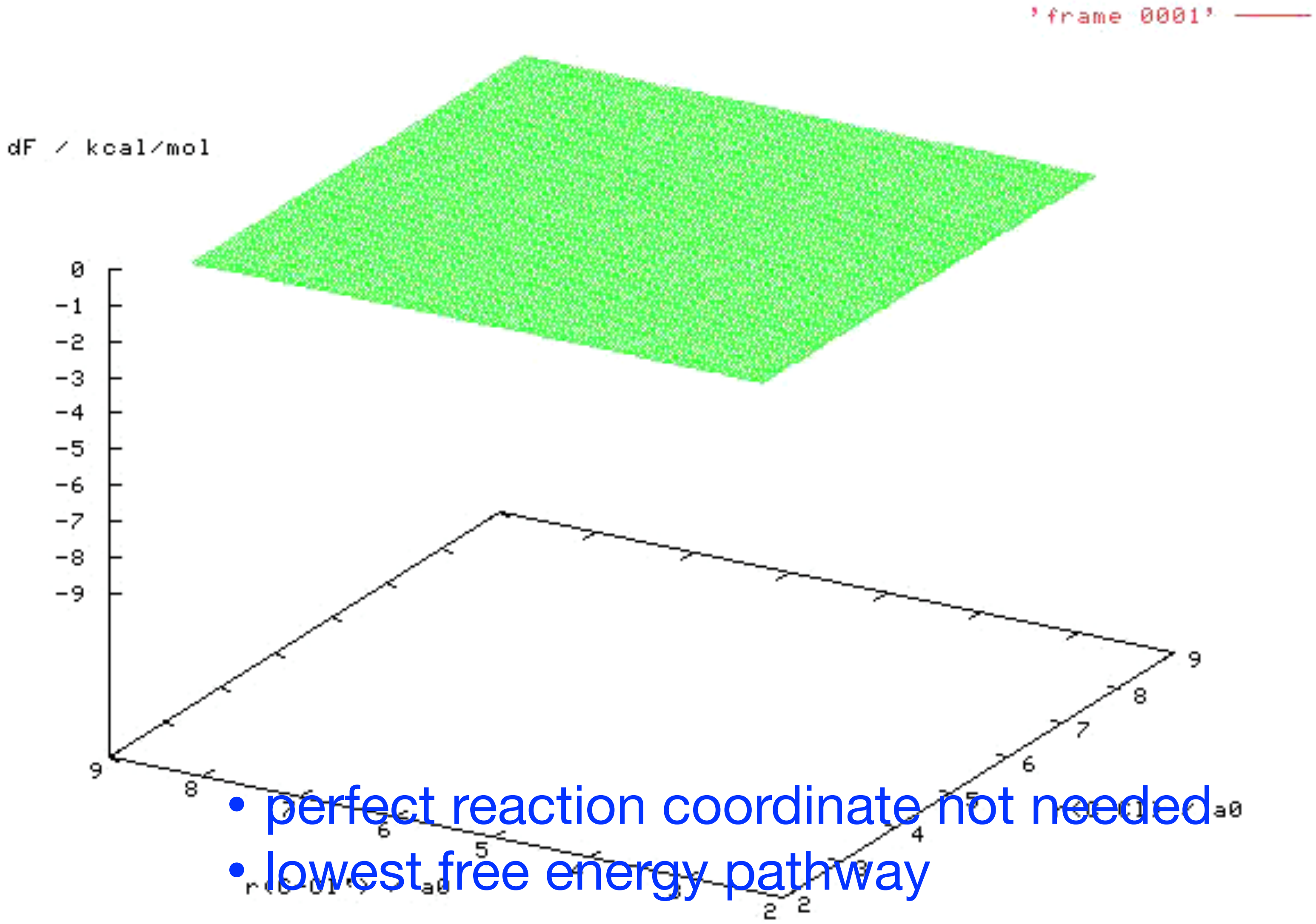
Two collective variables: r_1 and r_2

simple example: S_N2 reaction



Two collective variables: r_1 and r_2





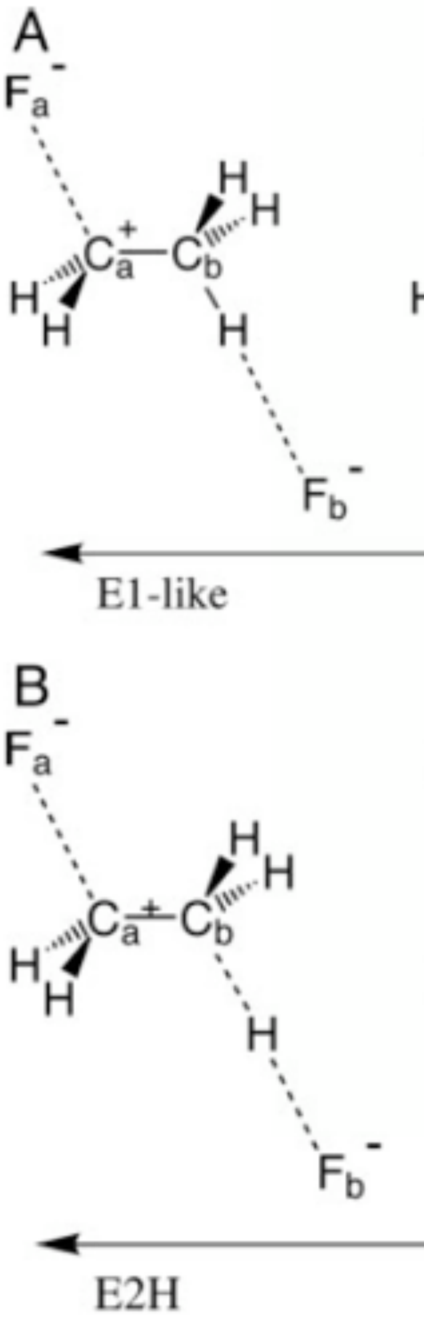
- perfect reaction coordinate not needed
- lowest free energy pathway

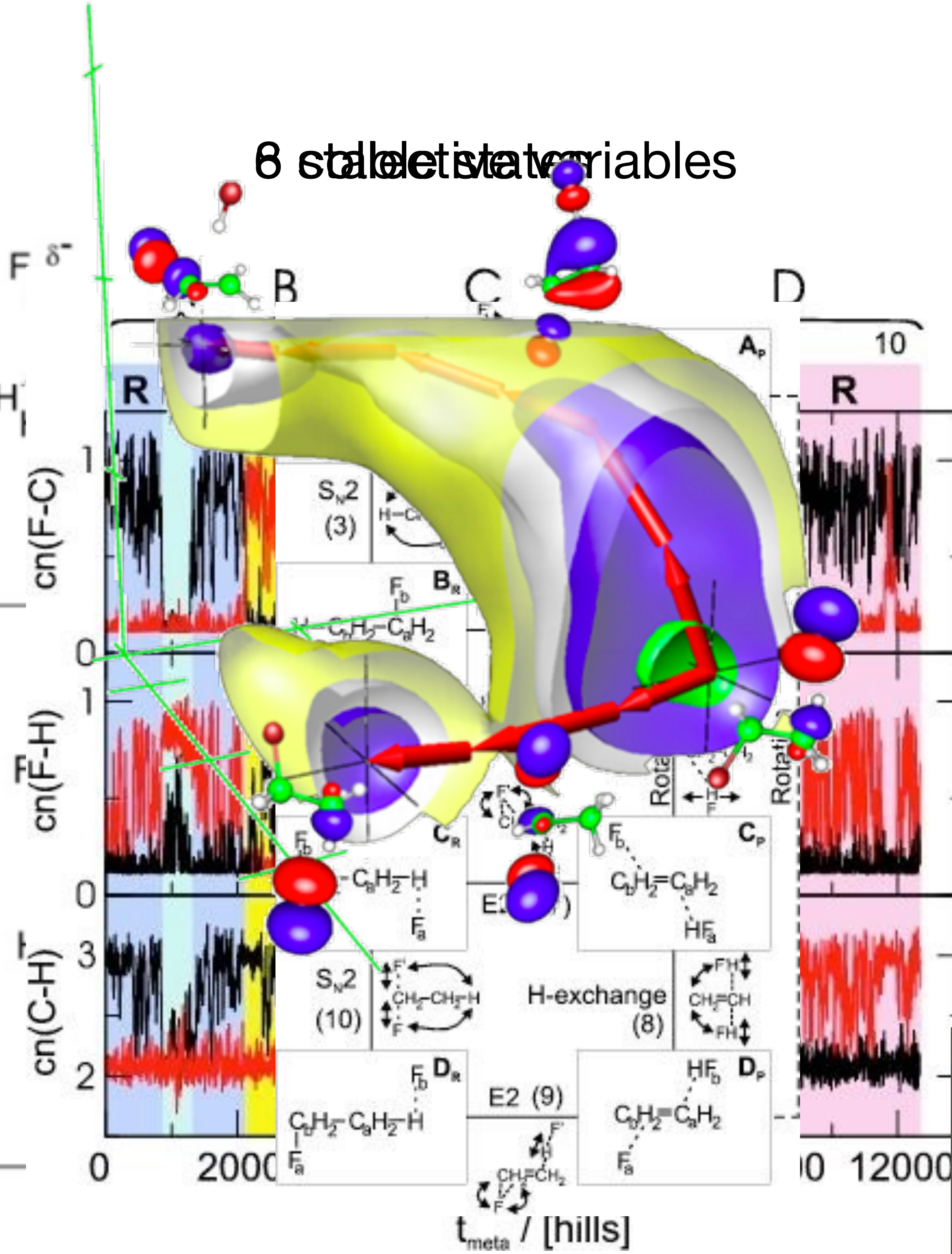
Bernd Ensing, Alessandro Laio, Michele Parrinello and Michael L. Klein
J. Phys. Chem. B **109** (2005), 6676-6687

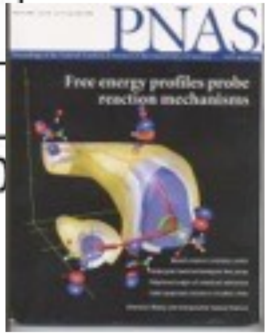
E2 vs S_N2 reaction

3 collective variables

6 stable states







Bernd Ensing and Michael L. Klein
Proc. Natl. Proc. Acad., USA **102** (2005), 6755-6759

Accuracy of metadynamics

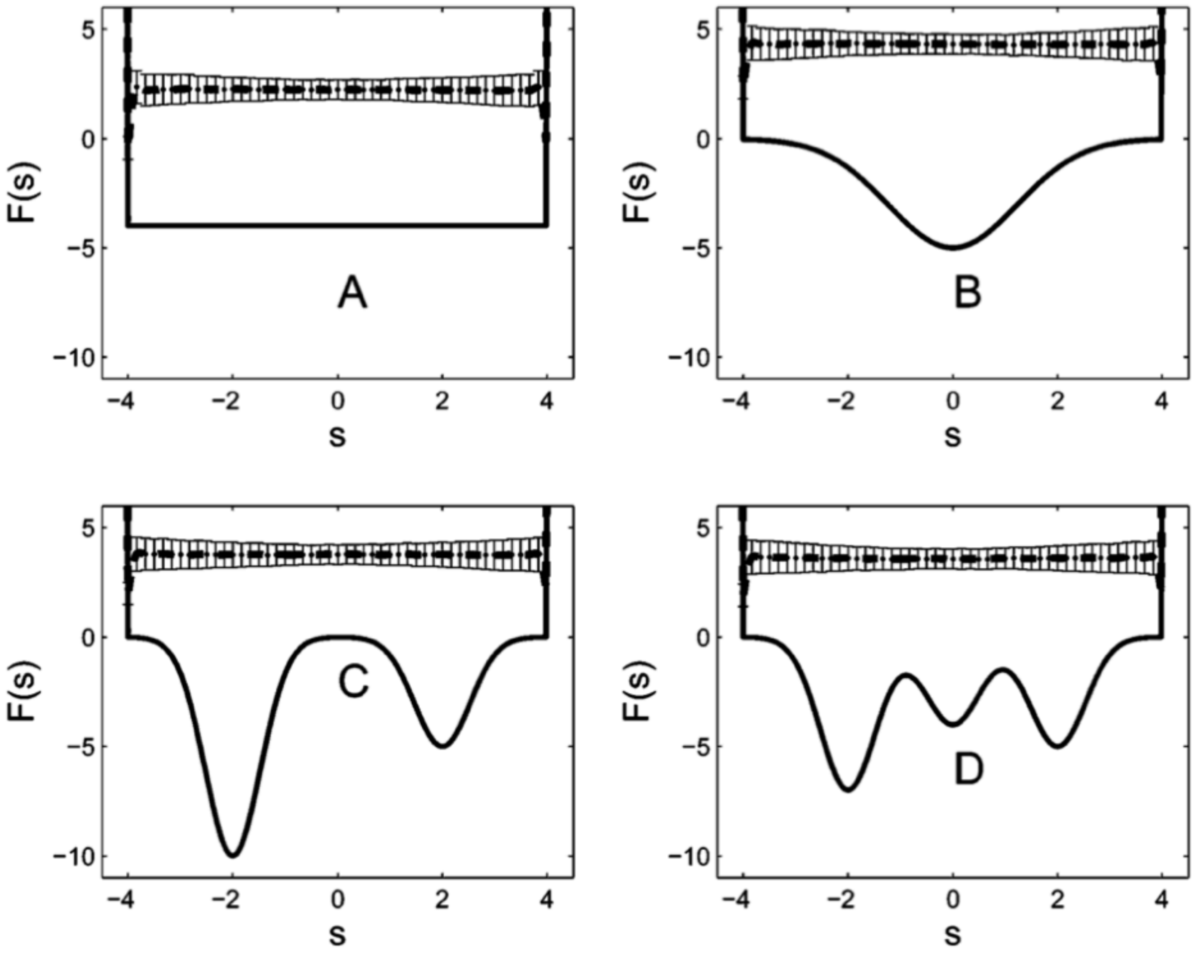
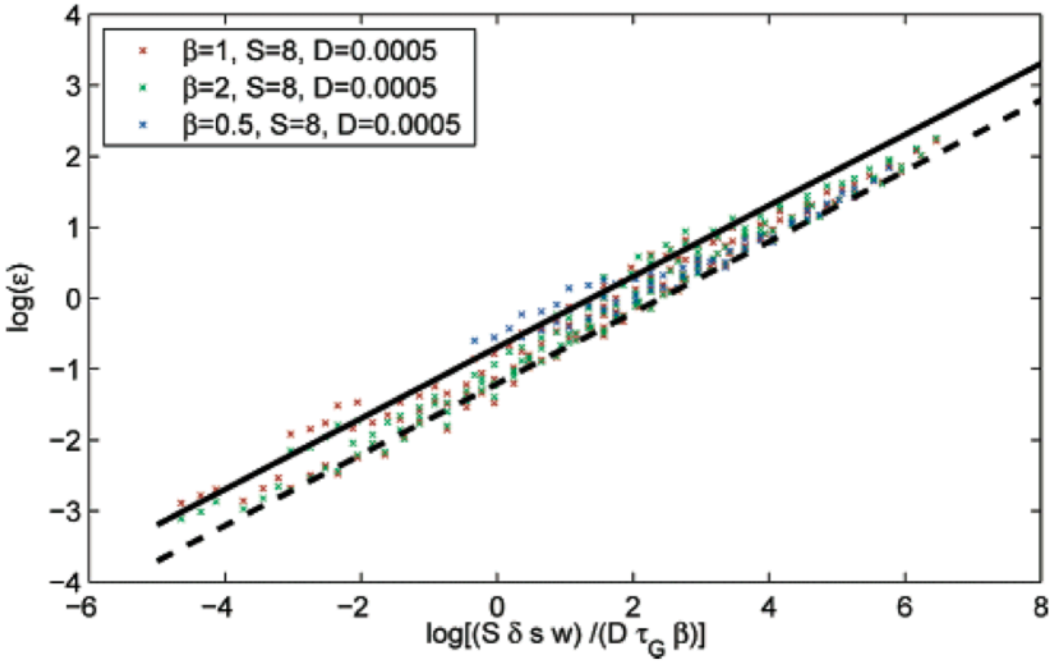


Figure 1. Metadynamics results for four different free energy profiles: (A) $F(s) = -4$; (B) $F(s) = -5\exp(-(s/1.75)^2)$; (C) $F(s) = -5\exp(-(s - 2/0.75)^2) - 10\exp(-(s + 2/0.75)^2)$; (D) $F(s) = -5\exp(-(s - 2/0.75)^2) - 4\exp(-(s/0.75)^2) - 7\exp(-(s + 2/0.75)^2)$. The average $\langle F(s) - F_G(s, t)\rangle$ computed over 1000 independent trajectories is represented as a dashed line, with the error bar given by

$$\epsilon(s,t) = \sqrt{\langle (F_G(s,t) - F(s) - \langle F_G(s,t) - F(s)\rangle)^2\rangle}$$



Error as a function of the metadynamics parameters *H*, τ and W and of β, *D* and *S* for dimension = 2 collective variables.

H: Gaussian height
τ : Gaussian deposit interval
W: Gaussian width
β : 1/temperature
D : diffusion coefficient
S : size of landscape

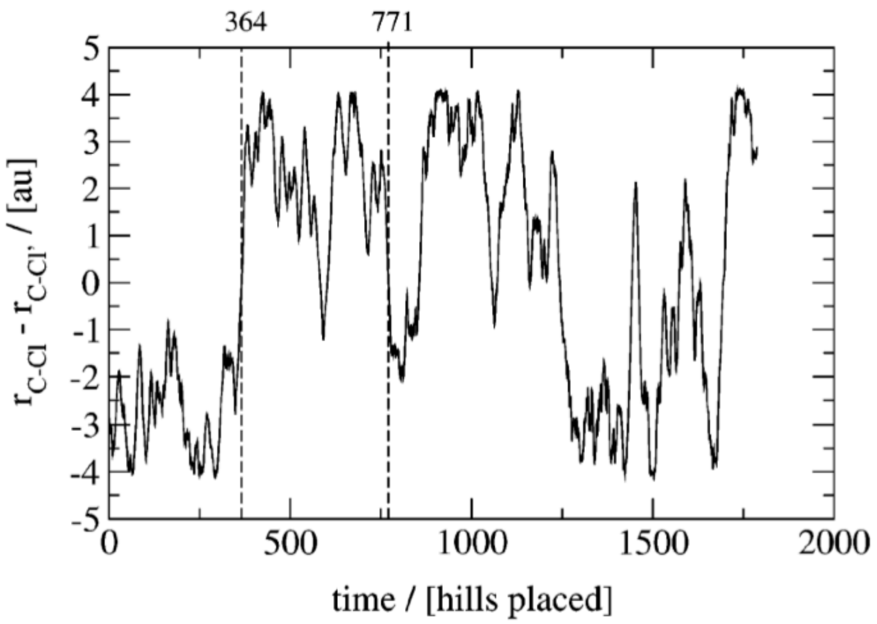
Dashed line is error: $\epsilon = C(d)\sqrt{\dfrac{HWS}{\tau D\beta}}$

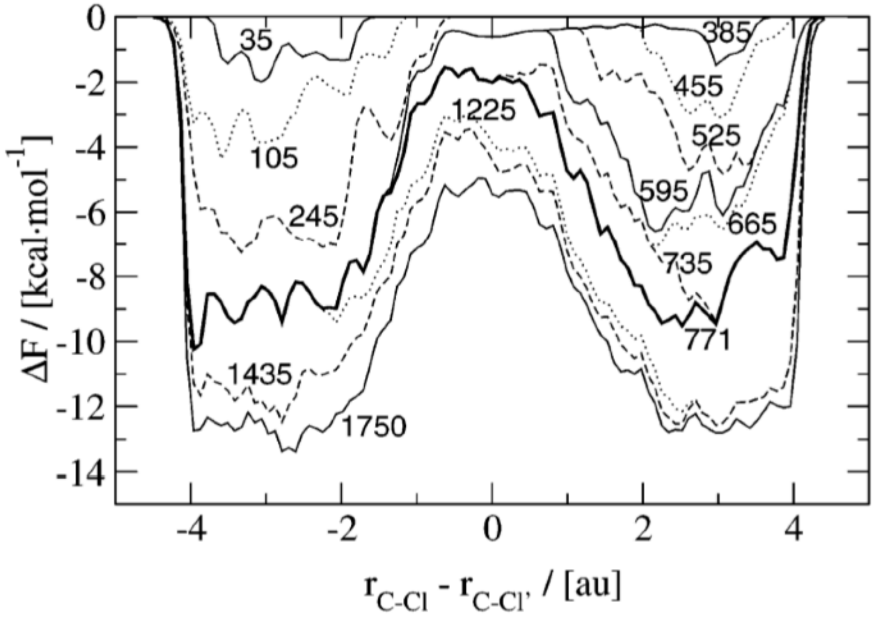
Laio, A.; Rodriguez-Fortea, A.; Gervasio, F. L.; Ceccarelli, M.; Parrinello, M. Assessing the accuracy of metadynamics. J. Phys. Chem. B 2005, 109, 6714-6721.

More collective variables?

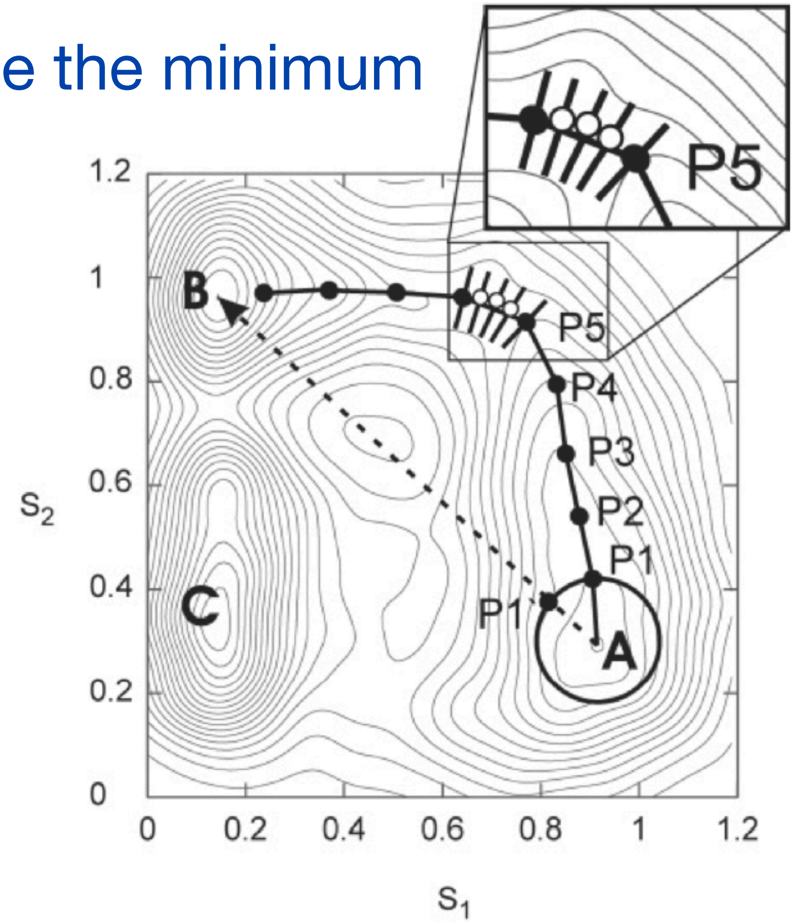
A recipe

(1) Decrease hills with every recrossing for convergence

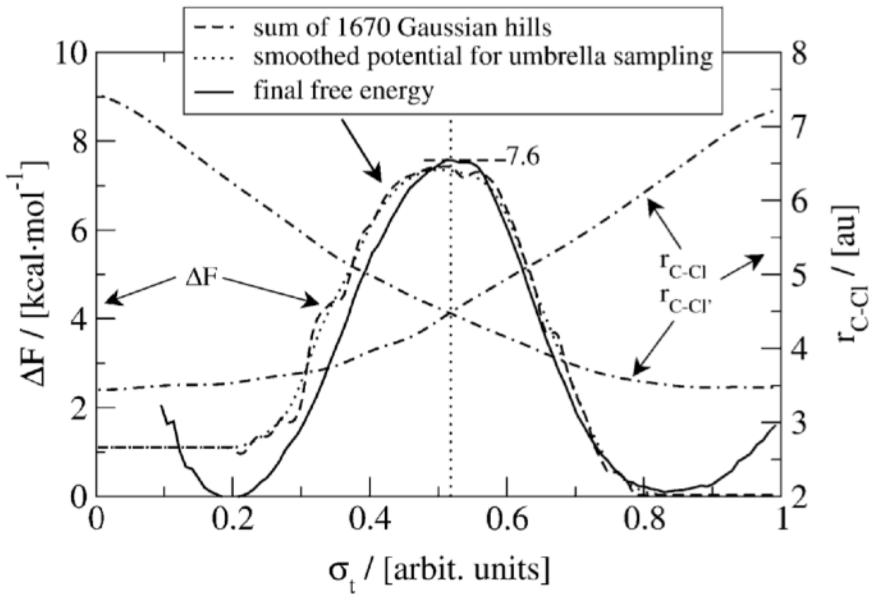




(2) Localize the minimum FE path



(3) Switch to 1D to converge further



Ensing, B.; Laio, A.; Parrinello, M.; Klein, M. L. A recipe for the computation of the free energy barrier and the lowest free energy path of concerted reactions. J. Phys. Chem. B 2005, 109, 6676.

Well-tempered metadynamics

Rigorous convergence to the free energy surface

- extra parameter: ΔT
- (ω is deposit rate)
- Gaussian size depends on previously added Gaussians

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

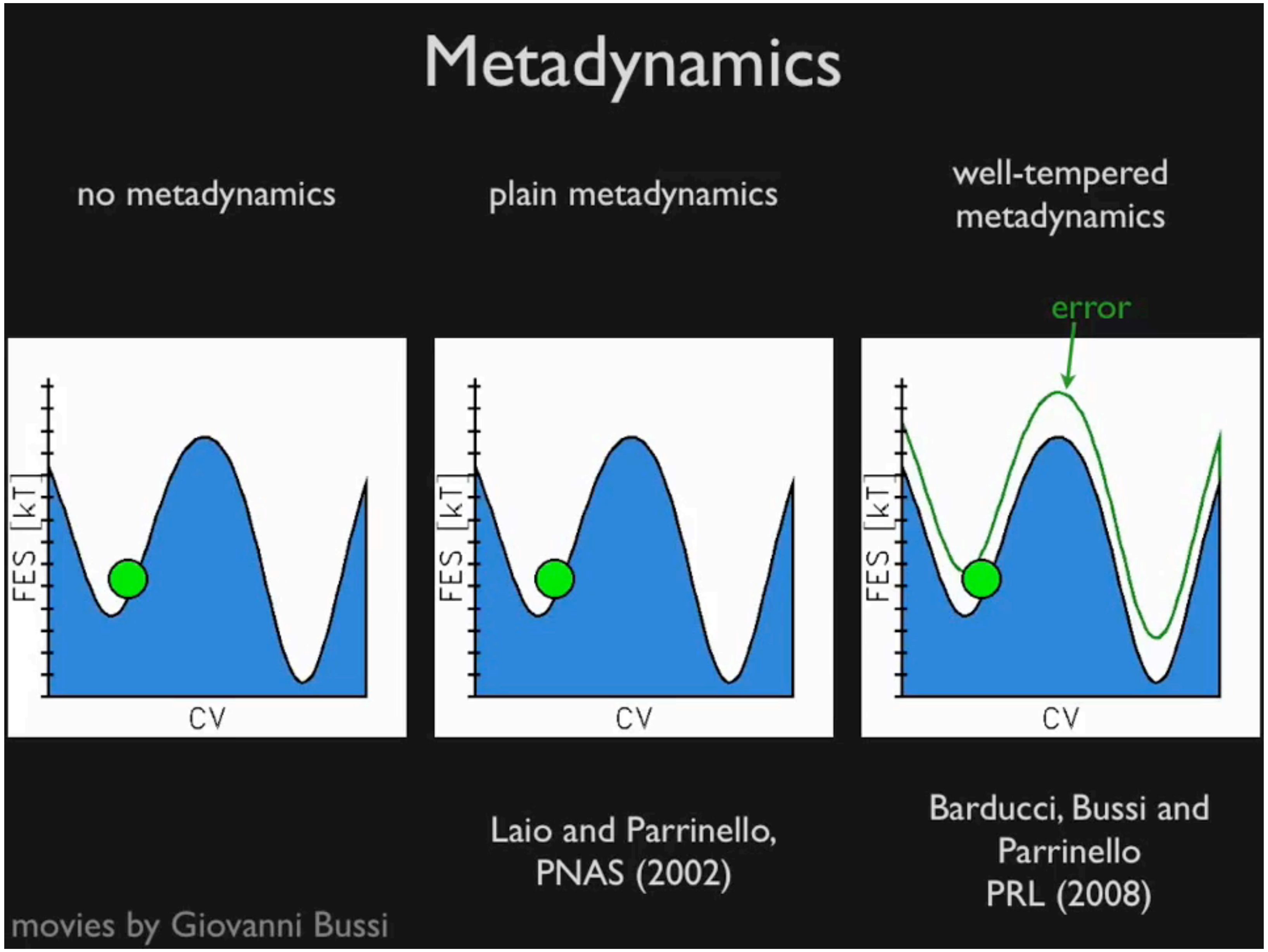
- ΔT= 0 --> bias is zero
- ΔT= ∞ --> normal metadyn.

(no convergence in these limits)

V_{bias} does not become flat!
CV distribution is sampled at T+ΔT and needs to be rescaled:

$$F(s,t) = -\frac{T+\Delta T}{\Delta T} V(s,t)$$

Youtube movie by Giovanni Bussi



Metadynamics

no metadynamics

plain metadynamics

well-tempered metadynamics

Laio and Parrinello, PNAS (2002)

Barducci, Bussi and Parrinello PRL (2008)

movies by Giovanni Bussi

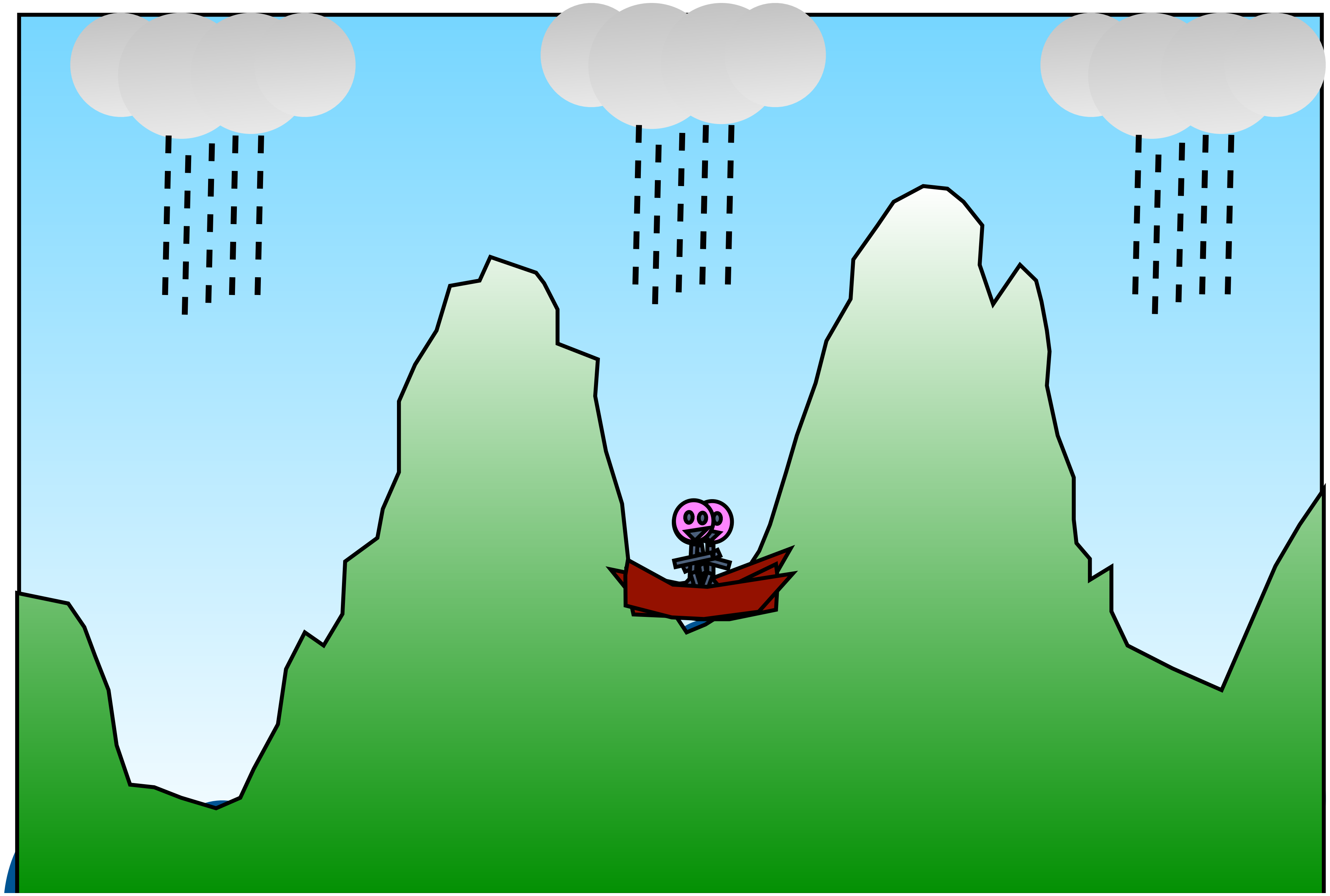
- A. Barducci, G. Bussi, M. Parrinello, Well-tempered metadynamics: A smoothly converging and tunable free-energy method, Phys. Rev. Lett. 100 (2008) 020603.

- J. F. Dama, M. Parrinello, and G. A. Voth Well-Tempered Metadynamics Converges Asymptotically, Phys. Rev. Lett. 112, (2014) 240602

Multiple walkers

A parallel version of metadynamics

- several metadynamics simulations running
- sharing and adding to a common V_{bias} potential (runs "feel" each other)
- runs "repel" each other and explore different regions
- trivial speedup and scaling



P. Raiteri, A. Laio, F. L. Gervasio, C. Micheletti, and M. Parrinello, Efficient Reconstruction of Complex Free Energy Landscapes by Multiple Walkers Metadynamics, J. Phys. Chem. B 2006, 110, 3533-3539

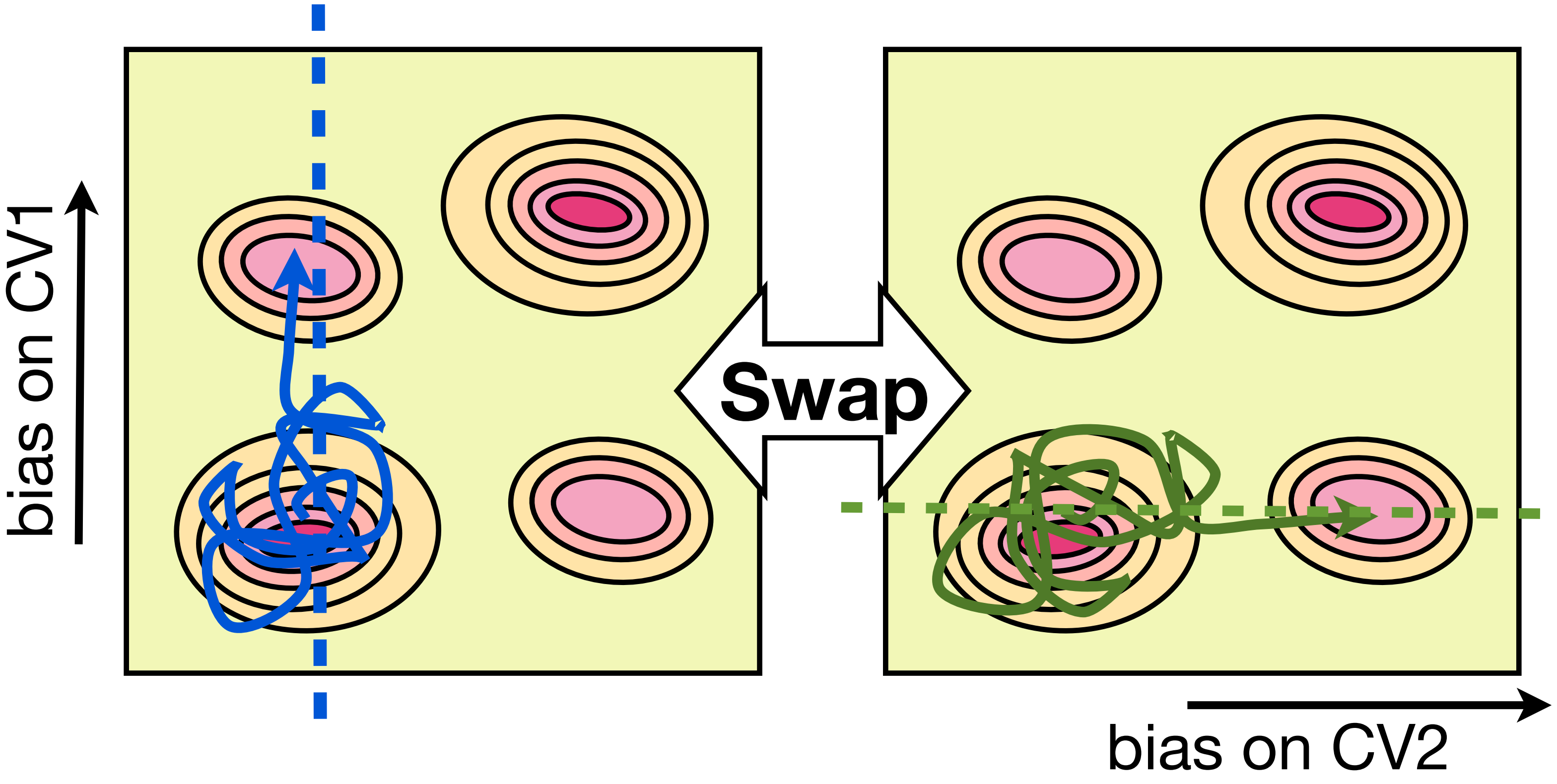
Dealing with many collective variables (approach 1)

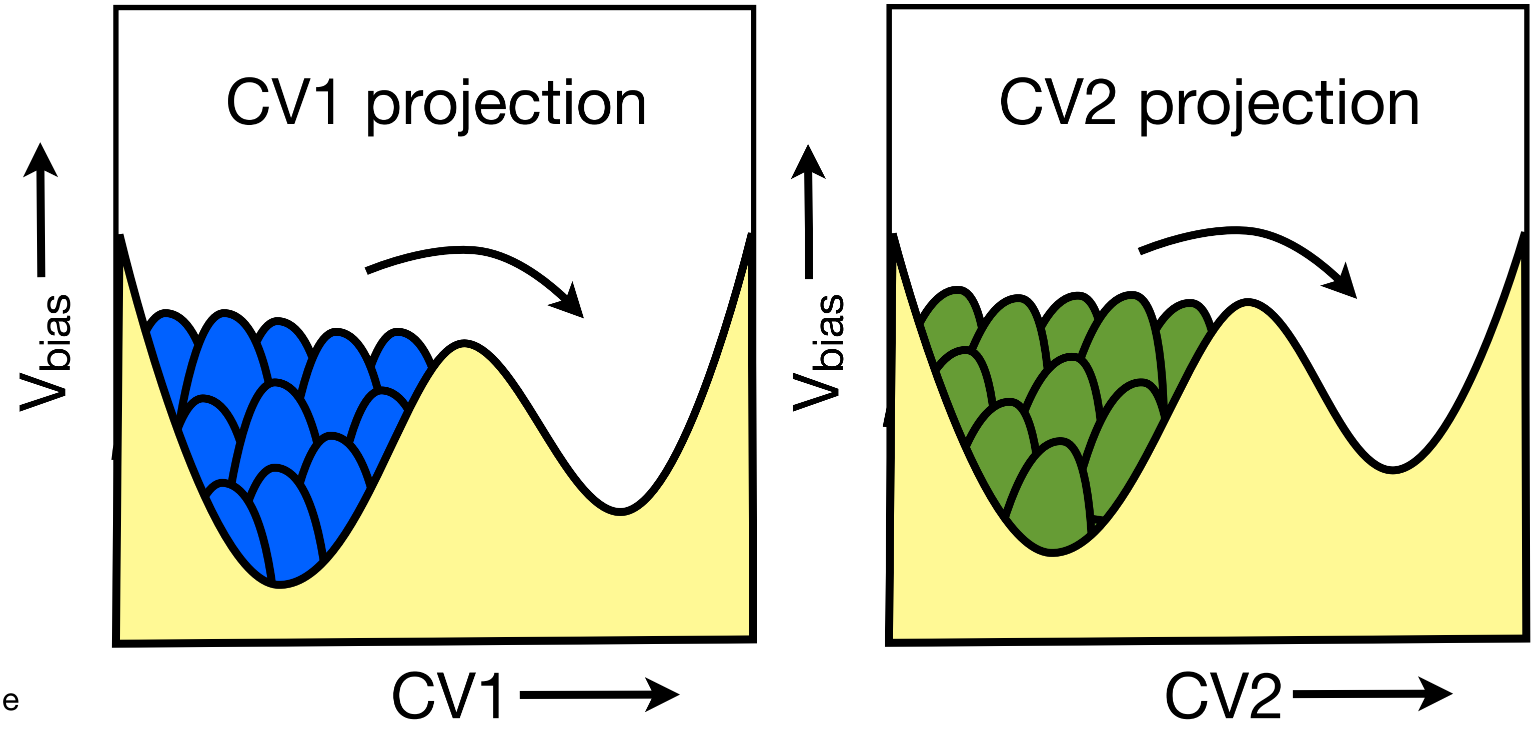
Bias-exchange metadynamics

- run many metadynamics simulations in parallel
- each builds a bias on 1 CV
- periodic swaps of configurations between runs
- reconstruct free energy landscape afterward

- non-trivial FE reconstruction
- non-trivial convergence (V_{bias} growth cannot stop)

- many CVs / complex systems
- also dynamics/diffusion information can be obtained.





S Piana, A Laio, A bias-exchange approach to protein folding
J. Phys. Chem. B 111, 4553 (2007)

F Marinelli, F Pietrucci, A Laio, S Piana, A Kinetic Model of Trp-Cage Folding from Multiple Biased Molecular Dynamics Simulations
PLOS Comput. Bio., 5, e1000452 (2009)

metadynamics

What about a path collective variable?

Escaping free energy minima

Laio A. and Parrinello M.
Proc. Natl. Acad. Sci. USA **99**, 12562–12566 (2002)

$$V(t,s) = \sum_{t'<t} H_{t'} \prod_{\alpha} \exp\left[\frac{-(s_\alpha - s_\alpha^{t'})^2}{2\delta_\alpha^2 W^2}\right]$$

String method in collective variables: Minimum free energy paths and isocommittor surfaces

Maragliano L., Fischer A., Vanden-Eijnden E., and Ciccotti G.
J. Chem. Phys. **125**, 024106 (2006)

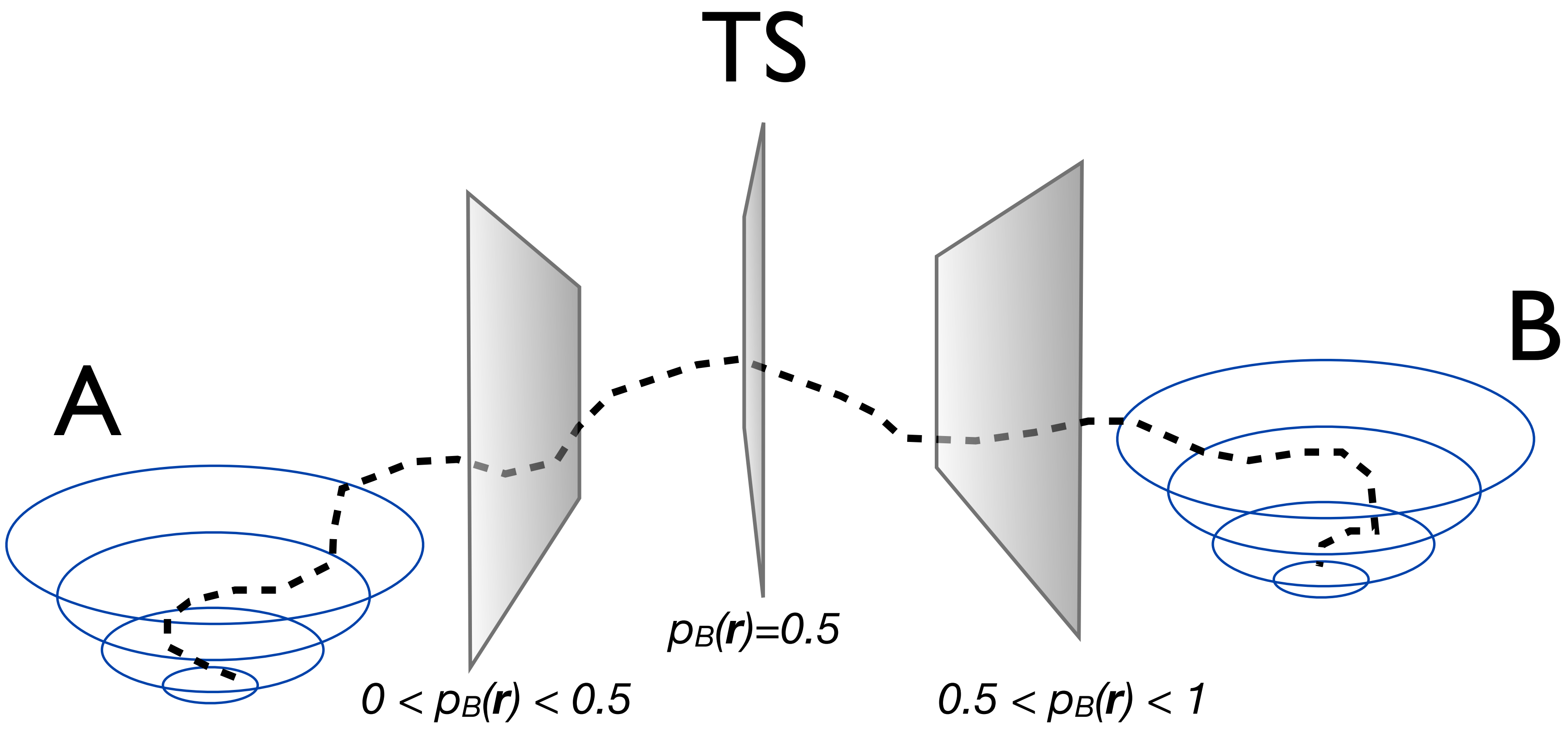
From A to B in free energy space
Branduardi D., Gervasio F.L., and Parrinello M.
J. Chem. Phys. **126**, 054103 (2007)

$$s(\mathbf{r}) = \frac{1}{P-1}\frac{\sum_{i=1}^{P}(i-1)e^{-\lambda(\mathbf{r}-\mathbf{r}(i))^2}}{\sum_{i=1}^{P} e^{-\lambda(\mathbf{r}-\mathbf{r}(i))^2}}$$

$$z(\mathbf{r}) = -\frac{1}{\lambda}\ln\left(\sum_{i=1} e^{-\lambda(\mathbf{r}-\mathbf{r}(i))^2}\right)$$

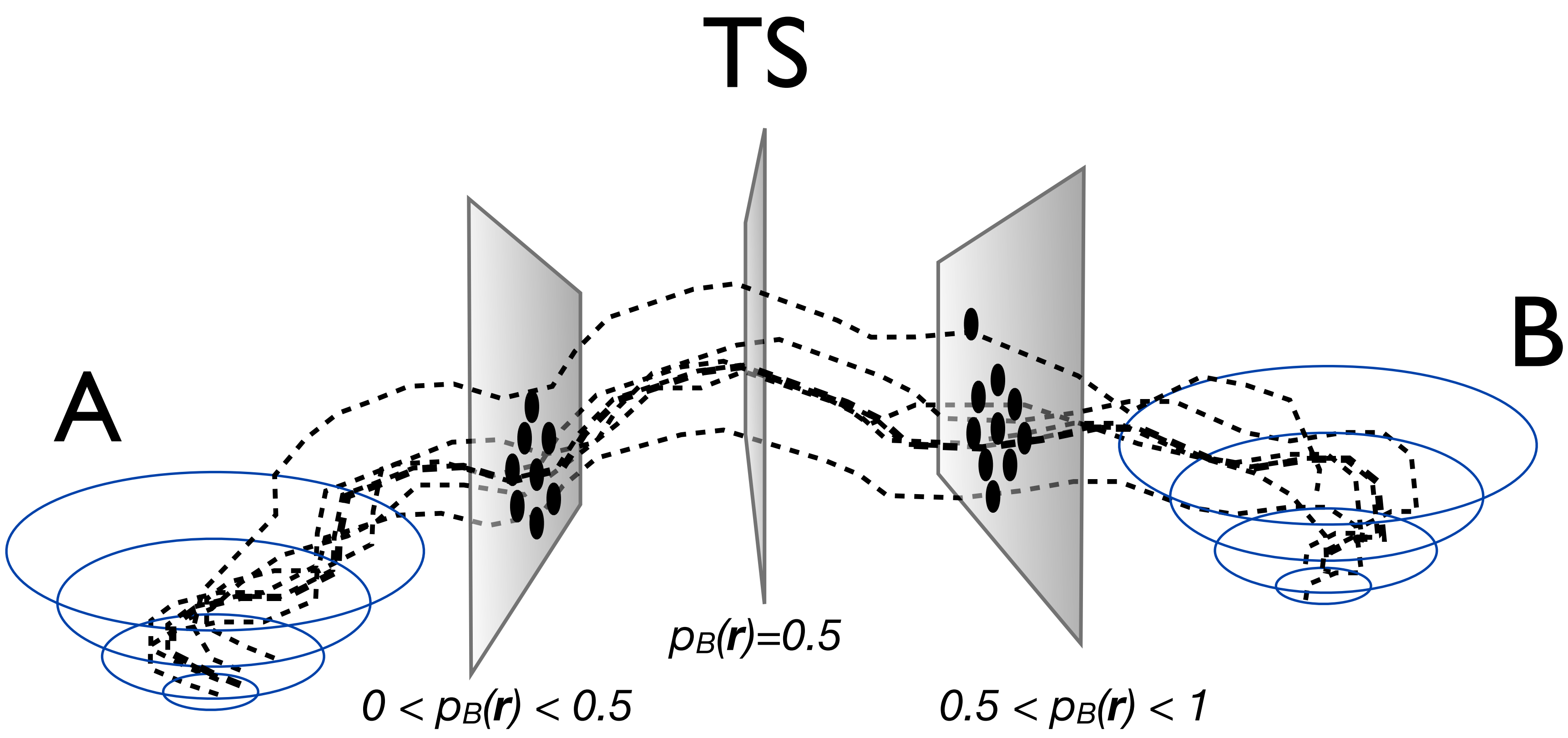
transition path sampling and path-metadynamics

iso-committor surfaces



transition path sampling and path-metadynamics

iso-committor surfaces



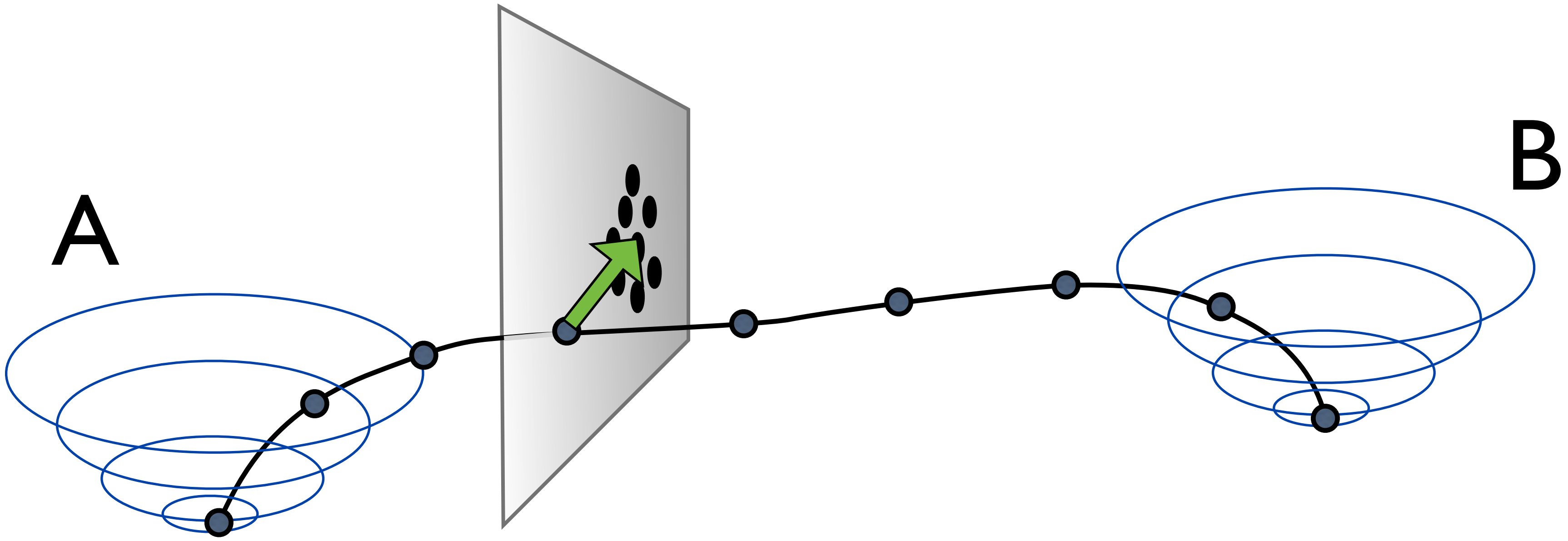
the average transition pathway
(in CV space)

Path-metadynamics (PMD)

- start from guess path
- bias dynamics along path
- move nodes to the mean density
- maintain equidistant nodes

$$s = i_2 + \mathrm{sign}(i_2 - i_1)\frac{\sqrt{(\mathbf{v}_1 \cdot \mathbf{v}_2)^2 - |\mathbf{v}_3|^2(|\mathbf{v}_1|^2 - |\mathbf{v}_2|^2)}}{2|\mathbf{v}_3|^2} - \frac{\mathbf{v}_1 \cdot \mathbf{v}_3 - |\mathbf{v}_3|^2}{2|\mathbf{v}_3|^2}$$

$$z = \sqrt{\left[|\mathbf{v}_1|^2 - |\mathbf{v}_2|\left(\frac{\sqrt{(\mathbf{v}_1 \cdot \mathbf{v}_2)^2 - |\mathbf{v}_3|^2(|\mathbf{v}_1|^2 - |\mathbf{v}_2|^2)}}{2|\mathbf{v}_3|^2} - \frac{\mathbf{v}_1 \cdot \mathbf{v}_3 - |\mathbf{v}_3|^2}{2|\mathbf{v}_3|^2}\right)\right]^2}$$



distance to mean density

From A to B in free energy space.
Davide Branduardi, Francesco Luigi Gervasio, and Michele Parrinello.
J. Chem. Phys. **126** (2007), 054103

String method in collective variables: Minimum free energy paths and isocommittor surfaces
Luca Maragliano, Alexander Fischer, Eric Vanden-Eijnden, and Giovanni Ciccotti
J. Chem. Phys. **125** (2006), 024106

Path finding on high-dimensional free energy landscapes.
Grisell Díaz Leines and Bernd Ensing
Phys. Rev. Lett. **109** (2012), 020601

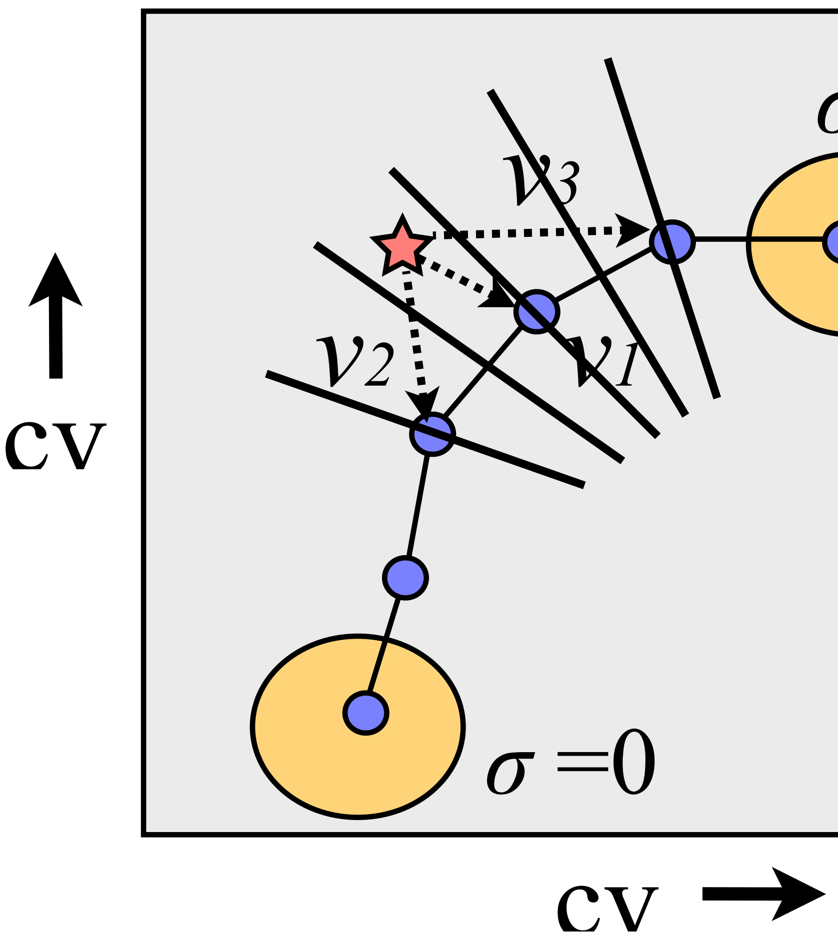
Advances in enhanced sampling along adaptive paths of collective variables.
Alberto Pérez de Alba Ortíz, Ambuj Tiwari, Rakesh C. Puthenkalathil, and Bernd Ensing
J. Chem. Phys. **149** (2018), 072320

MFEP as the collective variable

Metadynamics

- add extra collective variable: σ
- σ is a function of all other coll. variables
- biasing potential is only working on σ
- σ function adapts on the fly

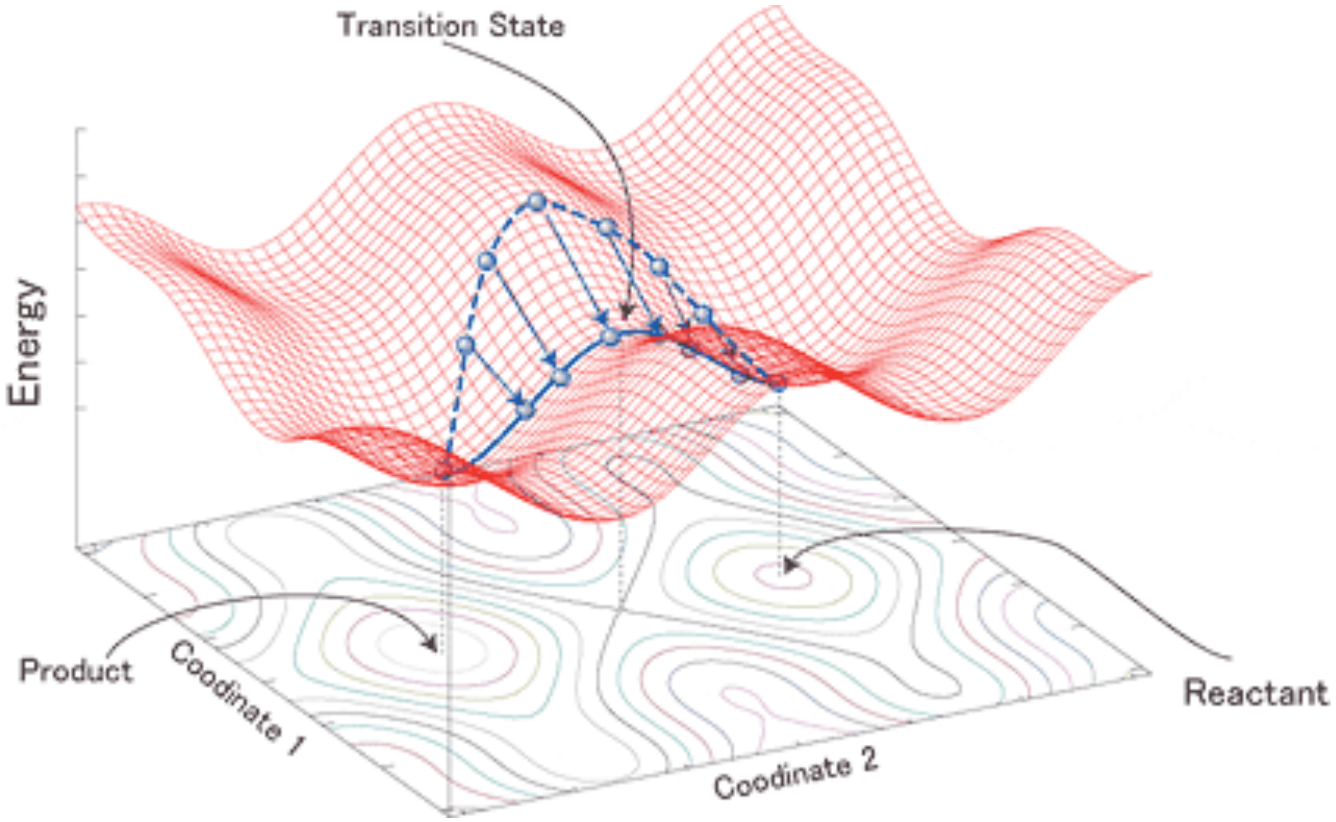
MFEP as the collective variable

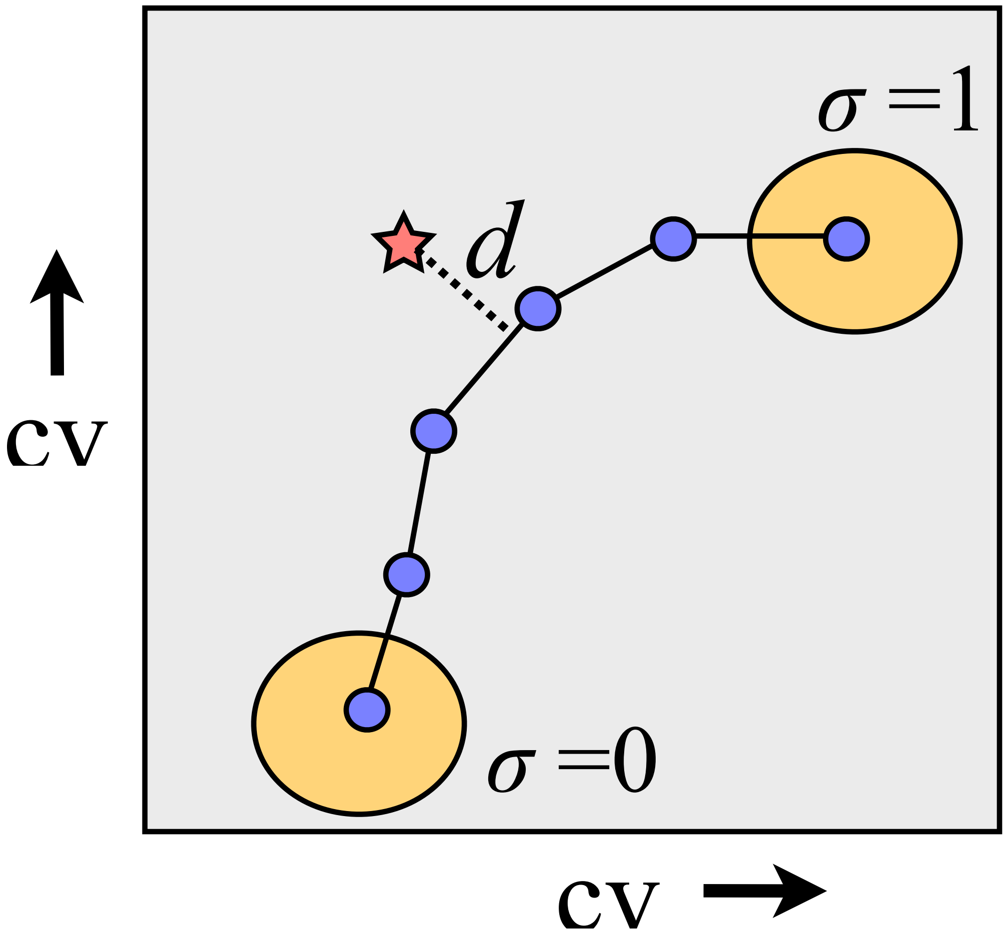


MFEP CV

$$\sigma = n^{-1}\left(i_{min} \pm \frac{\sqrt{(\bar{v}_1\bar{v}_3)^2 - \bar{v}_3\bar{v}_3(\bar{v}_1\bar{v}_1 - \bar{v}_2\bar{v}_2)} - \bar{v}_1\bar{v}_3}{2(\bar{v}_3\bar{v}_3 - 1)}\right)$$

MFEP as the collective variable





histogram distance to the path for each node
maximum of *P(d)* should be at *d*=0

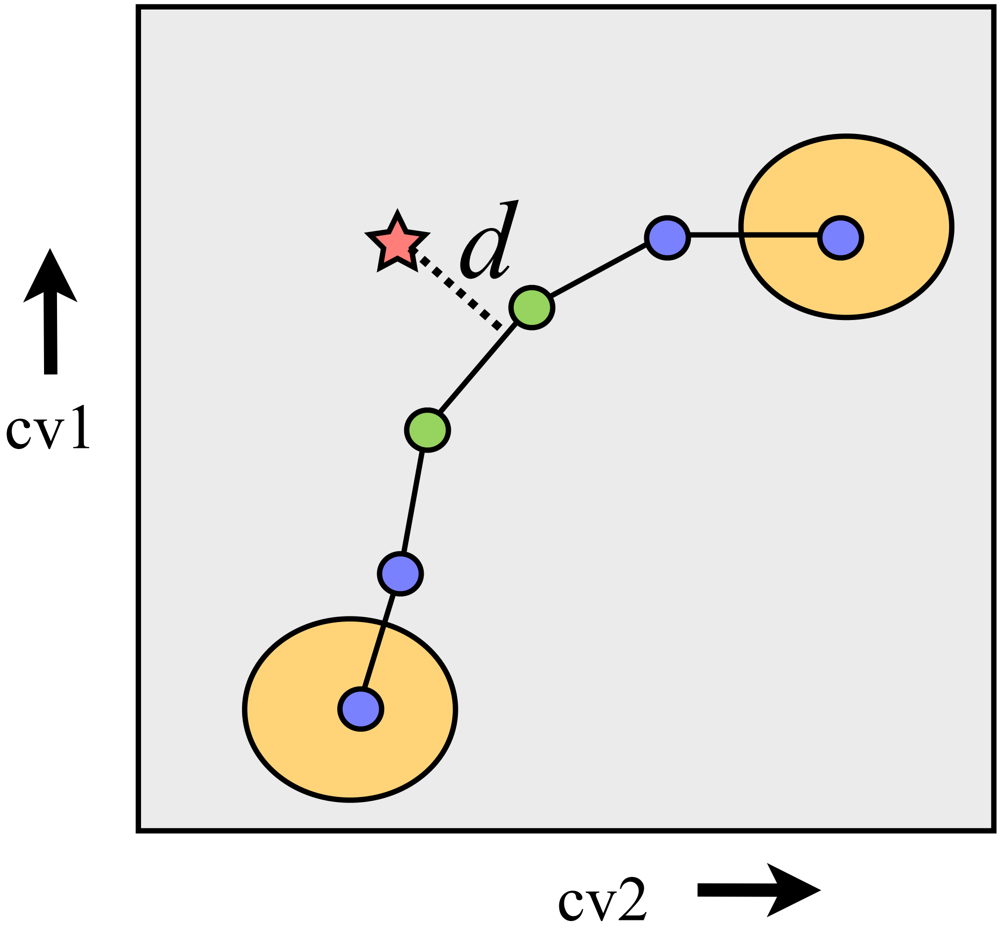
or accumulate average *d for each node*

practical implementation

leverage measurement of *d*
between closest nodes

$$\mathbf{s}_j^{t_{i+1}} = \mathbf{s}_j^{t_i} + \sum_k w_k \cdot |\mathbf{s}^{t_i}(\sigma(\mathbf{z}_k)) - \mathbf{z}_k| / \sum_k w_k$$

$$w_k = \max\left(0, \left(1 - \frac{|\mathbf{s}_j^{t_i} - \mathbf{s}^{t_i}(\sigma(\mathbf{z}_k))|}{|\mathbf{s}_j^{t_i} - \mathbf{s}_{j+1}^{t_i}|}\right)\right)$$



path update

move nodes to *d* (every step)
set *d*=0
redistribute nodes along path

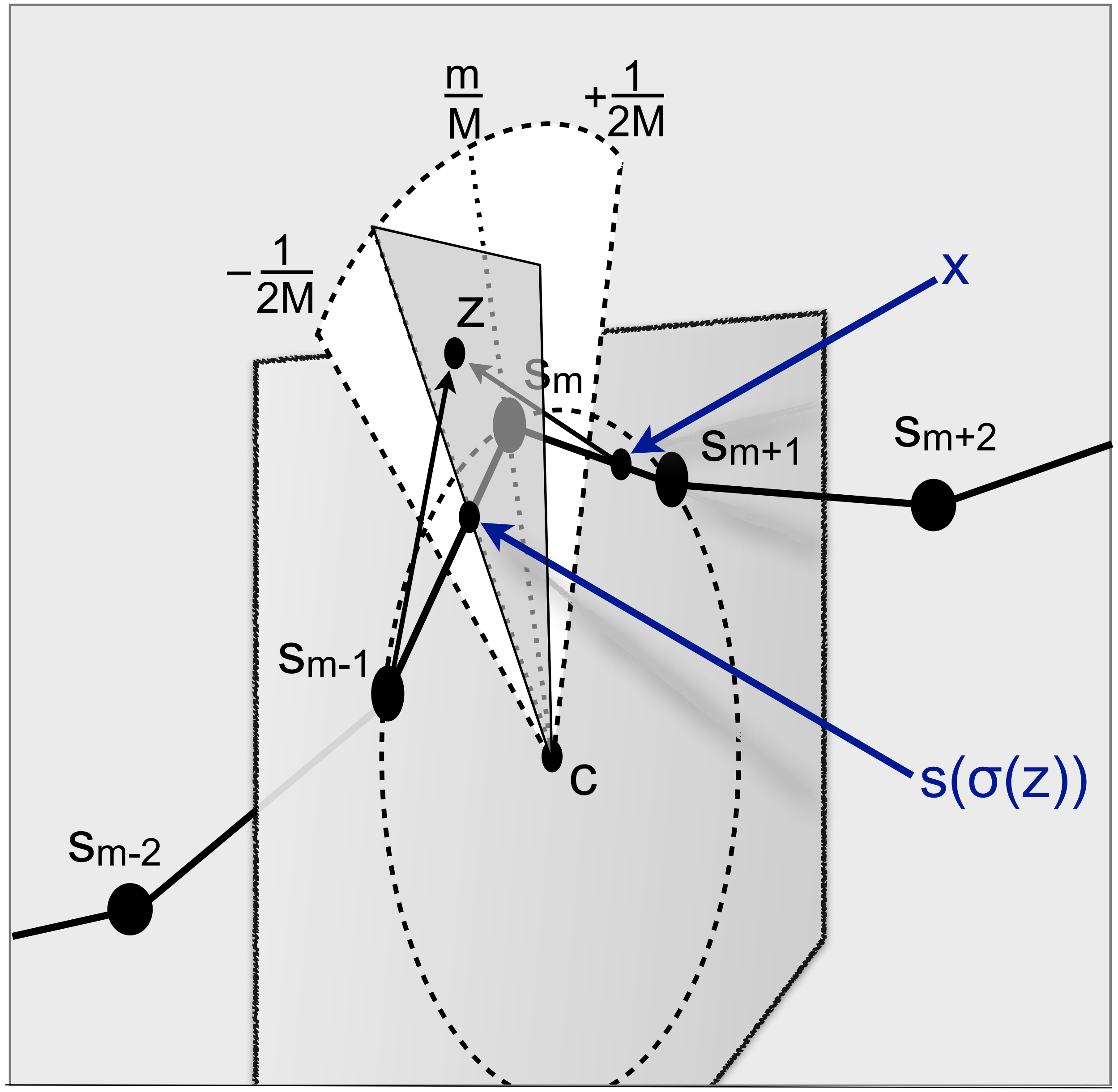
half life time of data

$$f = \exp[\tau^{-1} * \ln \frac{1}{2}]$$

optional tube potential

$$V^{\text{tube}} = \frac{k}{2} d^2 \, .$$

functional form of the path-CV



parameterized curve

$$\sigma = n^{-1}\left(i_{min} \pm \frac{\sqrt{(\bar{v}_1\bar{v}_3)^2 - \bar{v}_3\bar{v}_3(\bar{v}_1\bar{v}_1 - \bar{v}_2\bar{v}_2)} - \bar{v}_1\bar{v}_3}{2(\bar{v}_3\bar{v}_3 - 1)}\right)$$

$$|\mathbf{z} - \mathbf{s}_{m-1}| = |f(\mathbf{s}_{m+1} - \mathbf{s}_m) + \mathbf{s}_m - \mathbf{z}| \qquad (4)$$

Introducing $\mathbf{v}_1 = \mathbf{s}_m - \mathbf{z}$, $\mathbf{v}_2 = \mathbf{z} - \mathbf{s}_{m-1}$ and $\mathbf{v}_3 = \mathbf{s}_{m+1} - \mathbf{s}_m$, this equation is simplified to

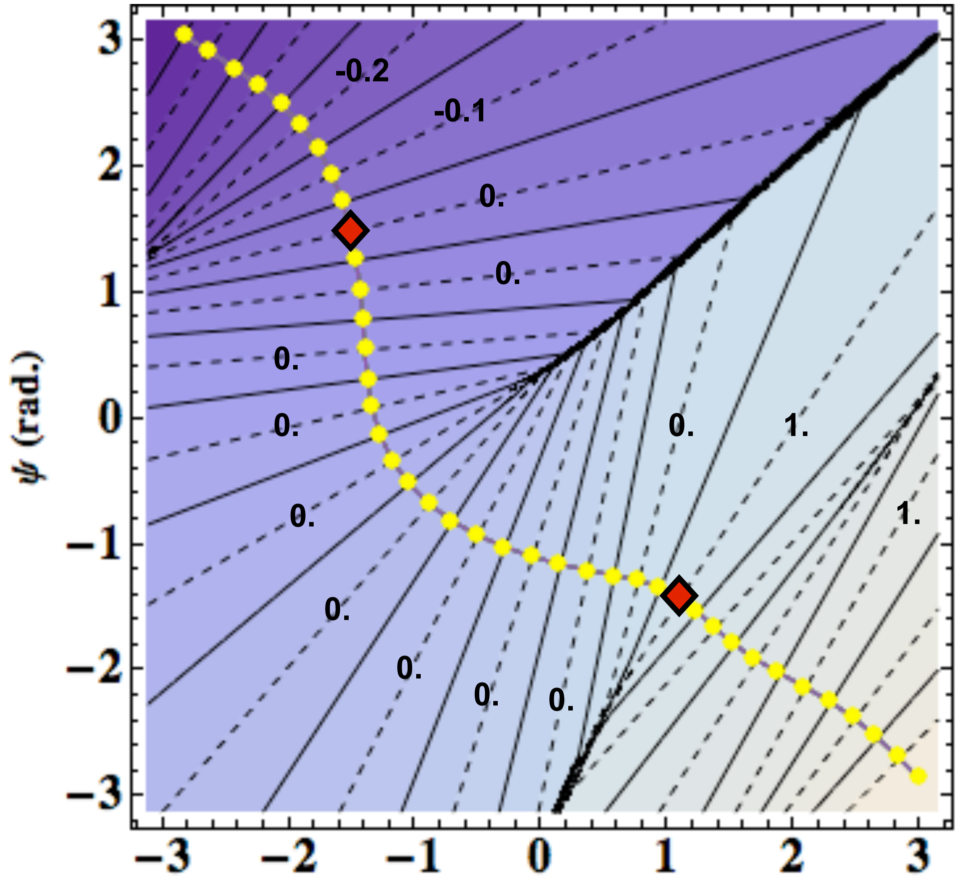
$$|\mathbf{v}_2| = |f\mathbf{v}_3 + \mathbf{v}_1|. \qquad (5)$$

Rewriting to solve for f results in a quadratic equation,

$$|\mathbf{v}_3|^2 f^2 + 2(\mathbf{v}_1 \cdot \mathbf{v}_3) f + (|\mathbf{v}_1|^2 - |\mathbf{v}_2|^2) = 0 \qquad (6)$$

with (relevant) root:

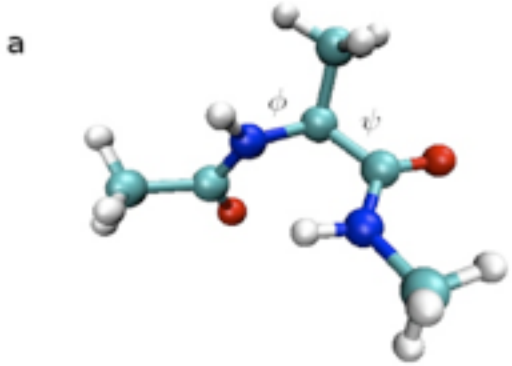
$$f = \frac{\sqrt{(\mathbf{v}_1 \cdot \mathbf{v}_3)^2 - |\mathbf{v}_3|^2(|\mathbf{v}_1|^2 - |\mathbf{v}_2|^2)} - (\mathbf{v}_1 \cdot \mathbf{v}_3)}{|\mathbf{v}_3|^2}. \qquad (7)$$



- Zhiwei Liu
- Preston Moore

University of the Sciences, Philadelphia

alanine dipeptide



Classical Molecular Dynamics
(CM3D code by Preston Moore, USP)

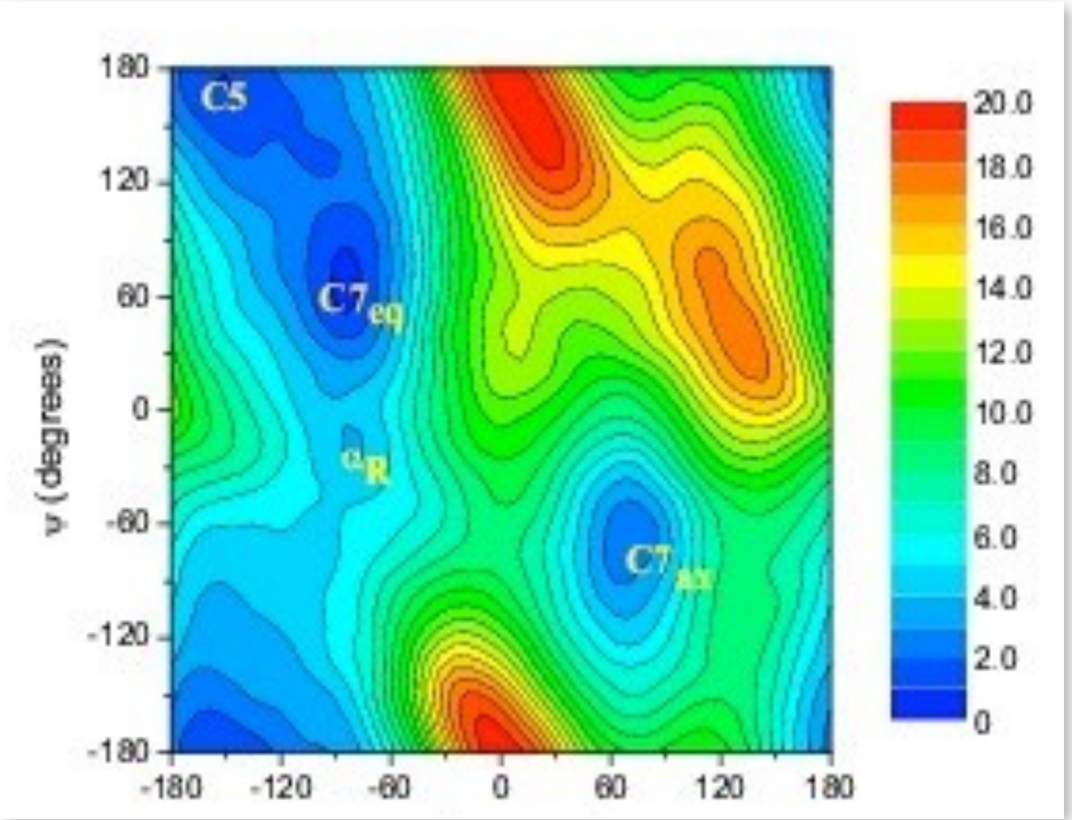
Alanine dipeptide + 216 water
CHARMM27 forcefield (modified internal params water)
Cubic periodic box L=18.8 Angstrom
NVT ensemble, T=298K

Hill size: W=0.2 rad, H=0.02 kcal/mol
Hill stride: 100 MD steps (=50 fs)

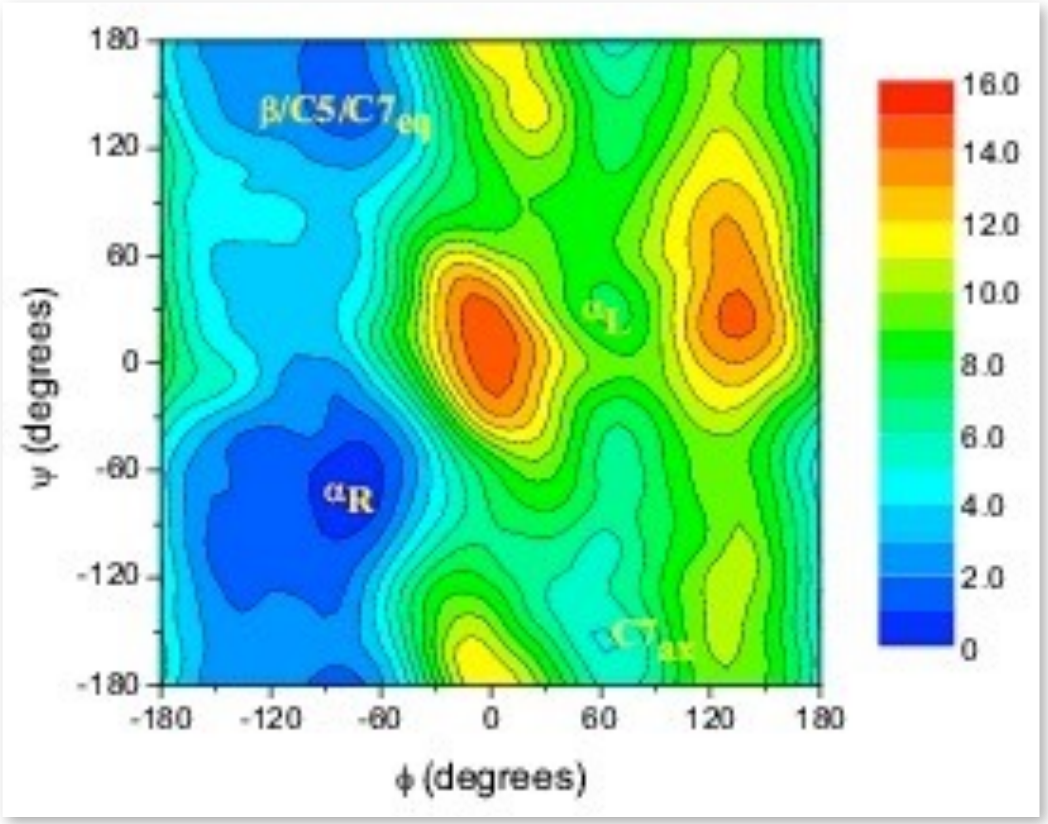
Total simulation time: ca. 5 ns

Error in energy: 0.3-0.5 kcal/mol
Error in angles: 3-10 degrees

Gas-phase



Aqueous solution



Metadynamics as a tool for exploring the free energy landscape of chemical reactions.
Bernd Ensing, Marco De Vivo, Zhiwei Liu, Preston Moore, and Michael L. Klein
Acc. Chem. Res. **39** (2006), 73-81

alanine dipeptide

Parameters:

$T = 300$ K

$H_{gaussian} = 10$ K

$W_{gaussian} = 0.05$

$\Delta t_{gaussian} = 100$ MD steps

$n = 20 + 20 + 20$

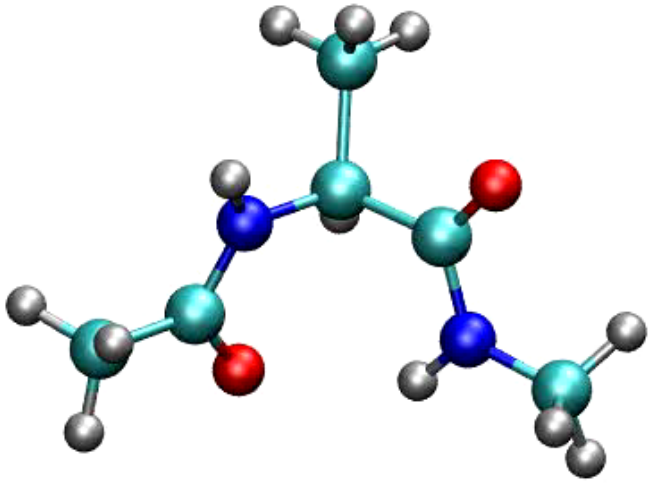
$\tau = 1000$ MD steps

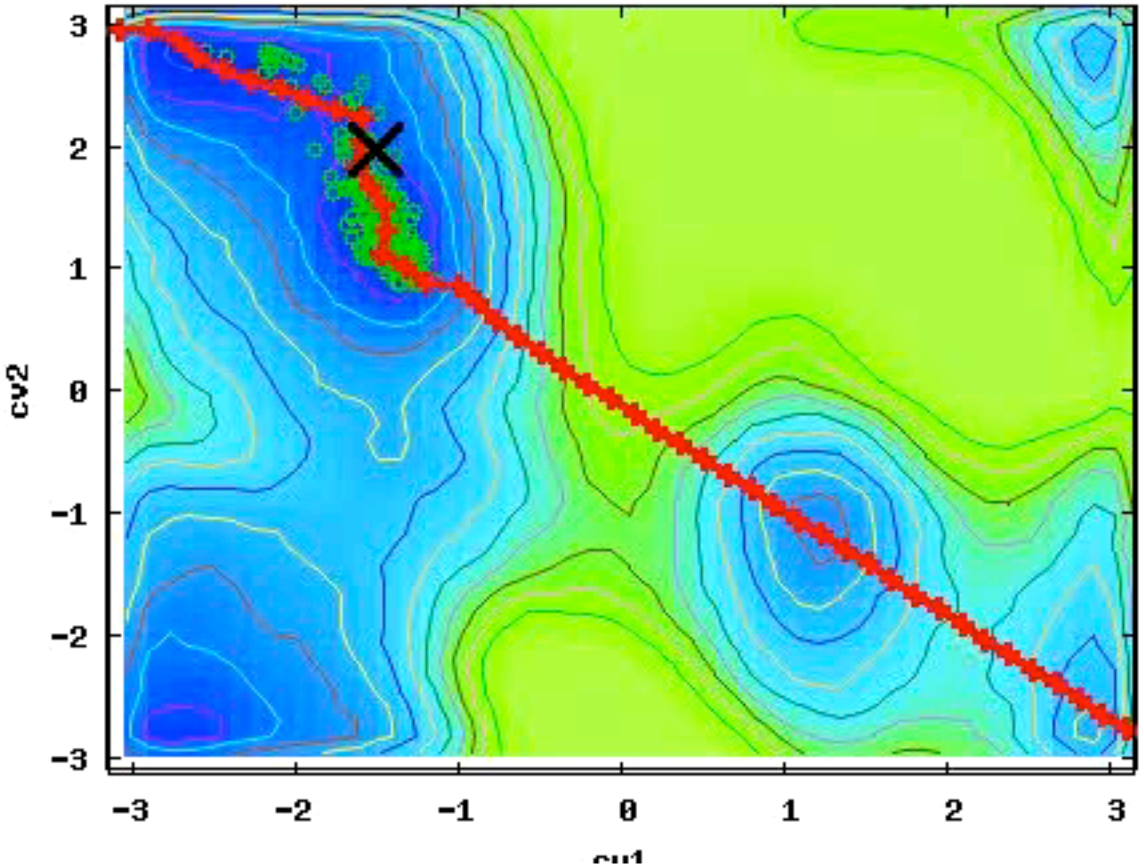
every recrossing:

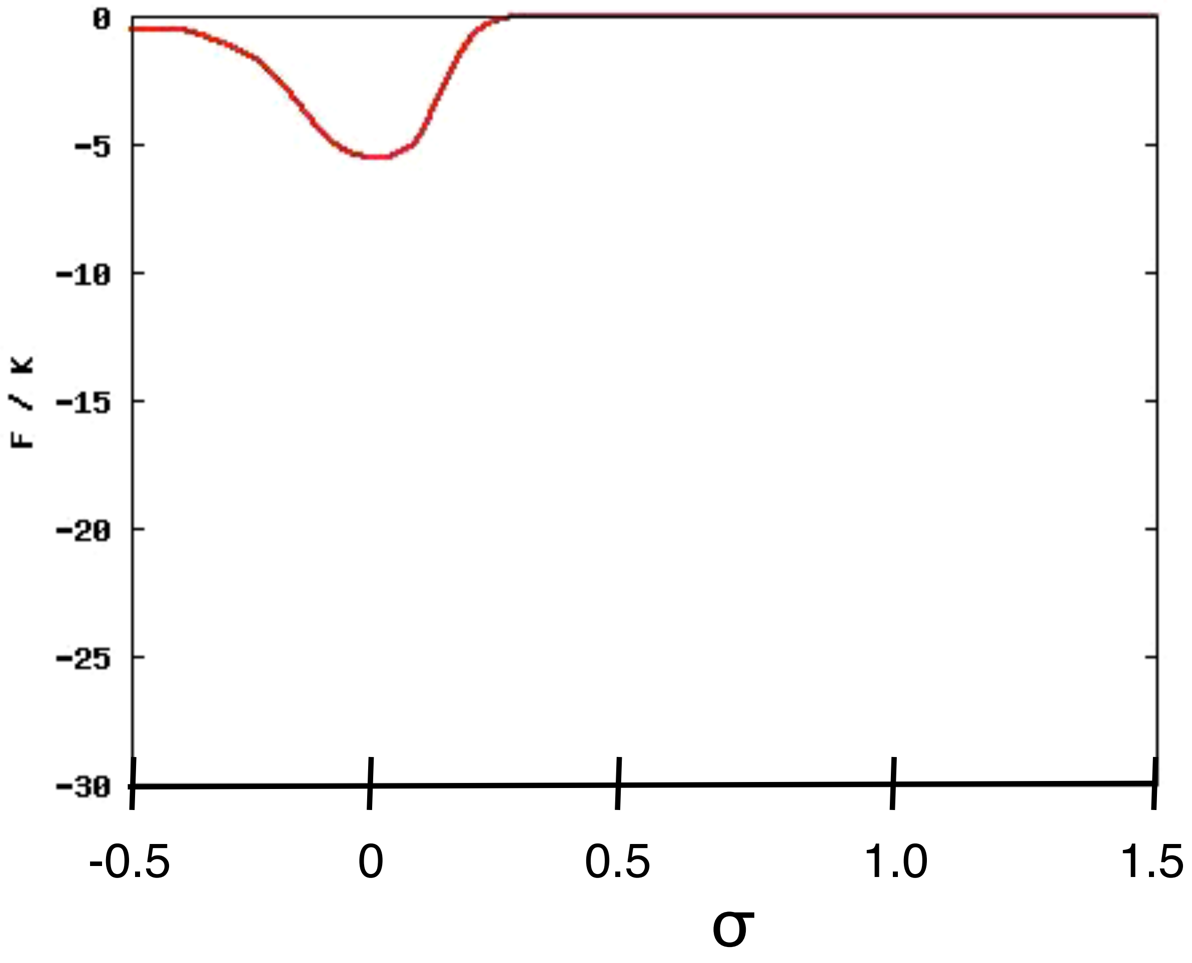
H x 50%

W x 50%

τ x 100



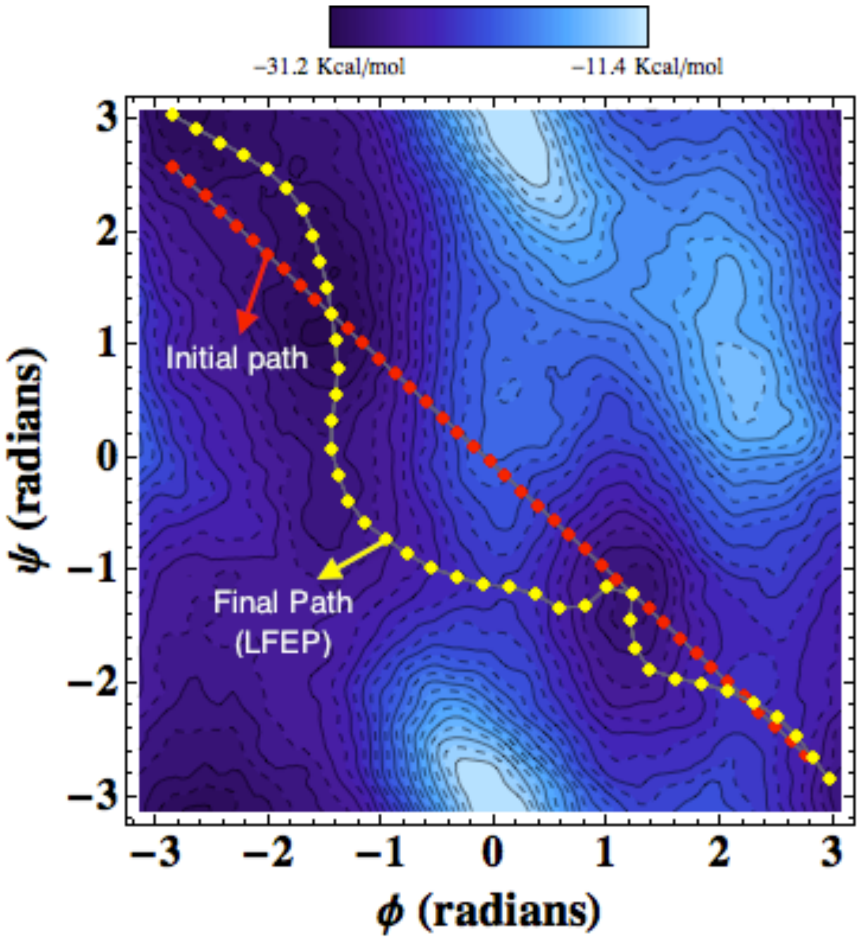




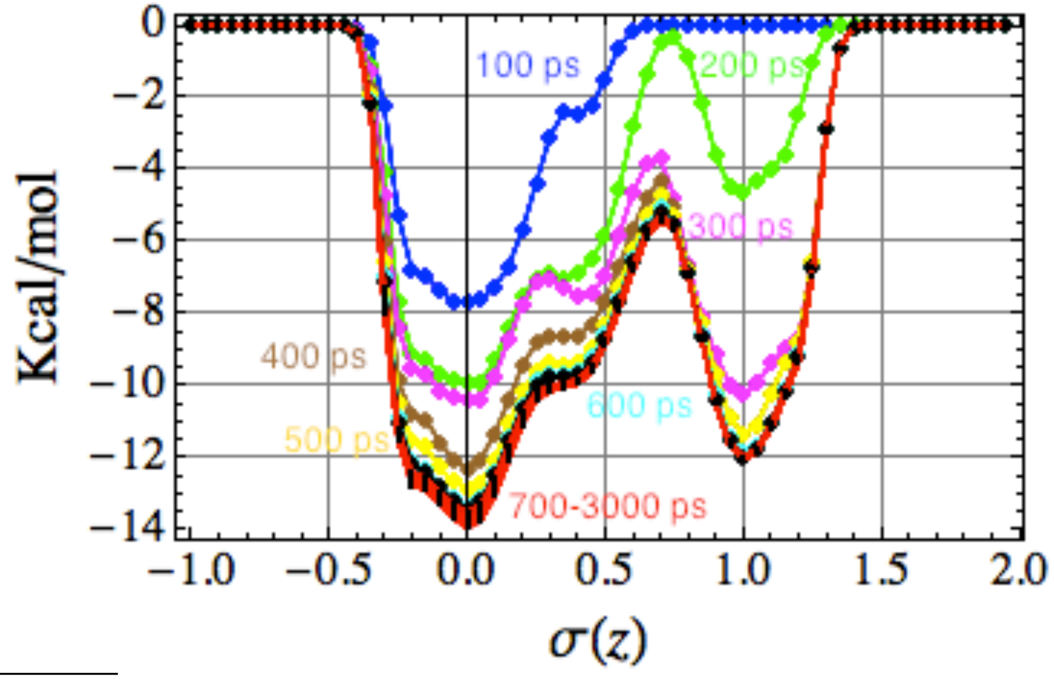
Path Finding on High-Dimensional Free Energy Landscapes.
Grisell Díaz Leines and Bernd Ensing
Phys. Rev. Lett. **109** (2012), 020601

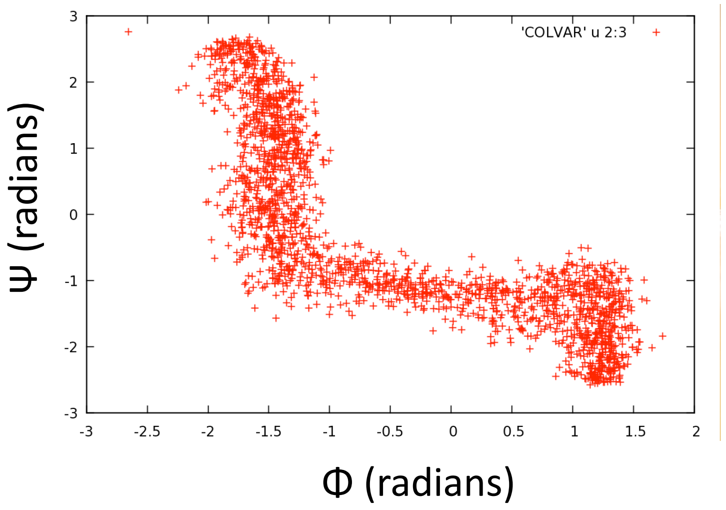
path-metadynamics convergence

Monitoring convergence of the path

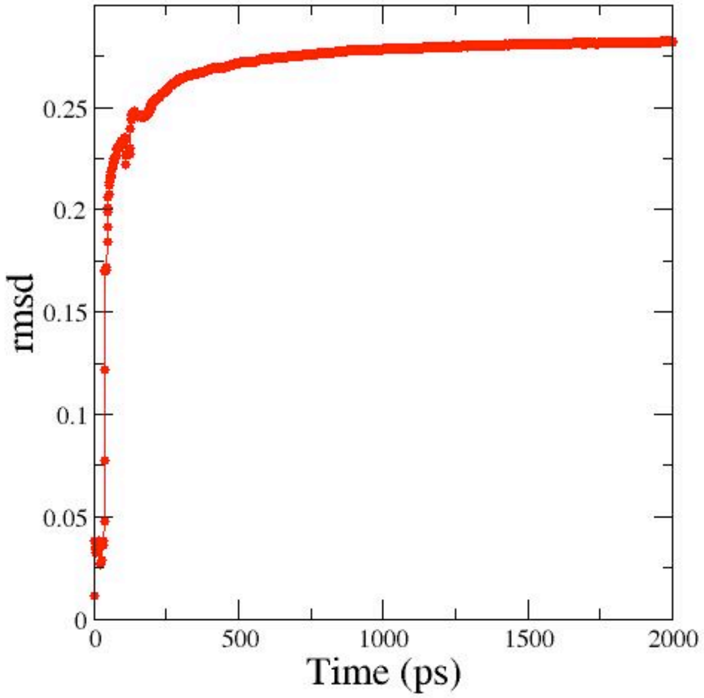


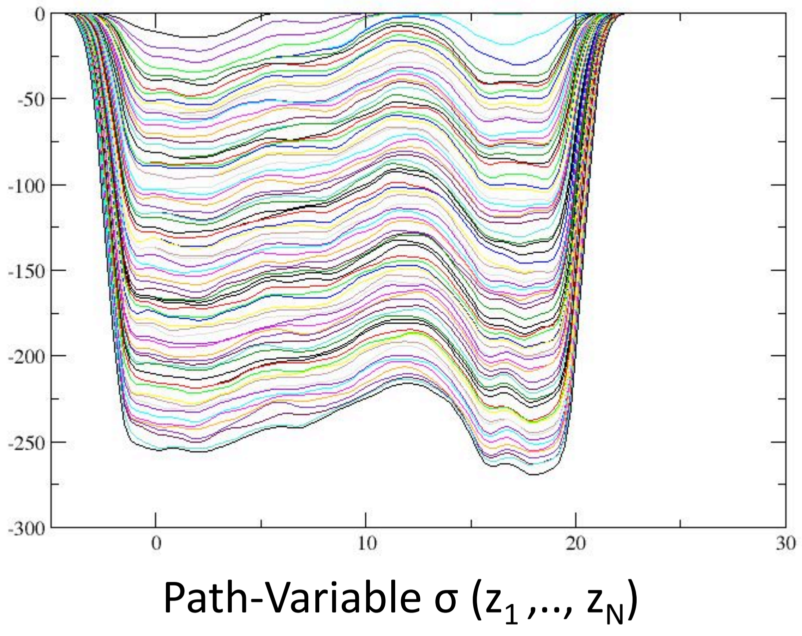
Free Energy along the path

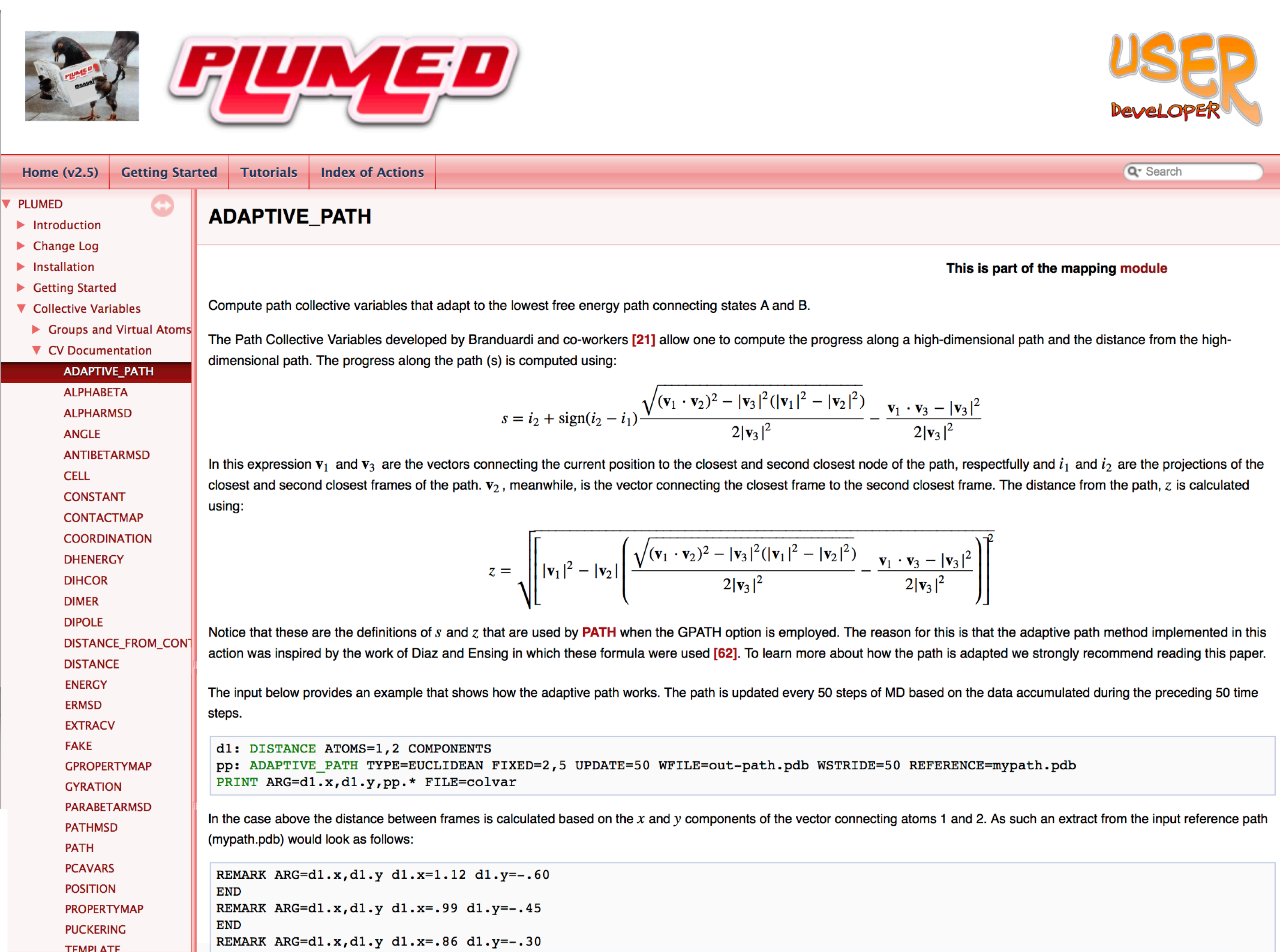




$$d(t)=\frac{1}{R}\sqrt{\sum_{p=1}^{R}\left\|s^{p}(t)-s^{p}(0)\right\|}$$







ADAPTIVE_PATH

This is part of the mapping module

Compute path collective variables that adapt to the lowest free energy path connecting states A and B.

The Path Collective Variables developed by Branduardi and co-workers [21] allow one to compute the progress along a high-dimensional path and the distance from the high-dimensional path. The progress along the path (s) is computed using:

$$s = i_2 + \mathrm{sign}(i_2 - i_1)\frac{\sqrt{(\mathbf{v}_1\cdot\mathbf{v}_2)^2 - |\mathbf{v}_3|^2(|\mathbf{v}_1|^2 - |\mathbf{v}_2|^2)}}{2|\mathbf{v}_3|^2} - \frac{\mathbf{v}_1\cdot\mathbf{v}_3 - |\mathbf{v}_3|^2}{2|\mathbf{v}_3|^2}$$

In this expression $\mathbf{v}_1$ and $\mathbf{v}_3$ are the vectors connecting the current position to the closest and second closest node of the path, respectfully and i_1 and i_2 are the projections of the closest and second closest frames of the path. $\mathbf{v}_2$, meanwhile, is the vector connecting the closest frame to the second closest frame. The distance from the path, z is calculated using:

$$z = \sqrt{\left[|\mathbf{v}_1|^2 - |\mathbf{v}_2|\left(\frac{\sqrt{(\mathbf{v}_1\cdot\mathbf{v}_2)^2 - |\mathbf{v}_3|^2(|\mathbf{v}_1|^2 - |\mathbf{v}_2|^2)}}{2|\mathbf{v}_3|^2} - \frac{\mathbf{v}_1\cdot\mathbf{v}_3 - |\mathbf{v}_3|^2}{2|\mathbf{v}_3|^2}\right)\right]^2}$$

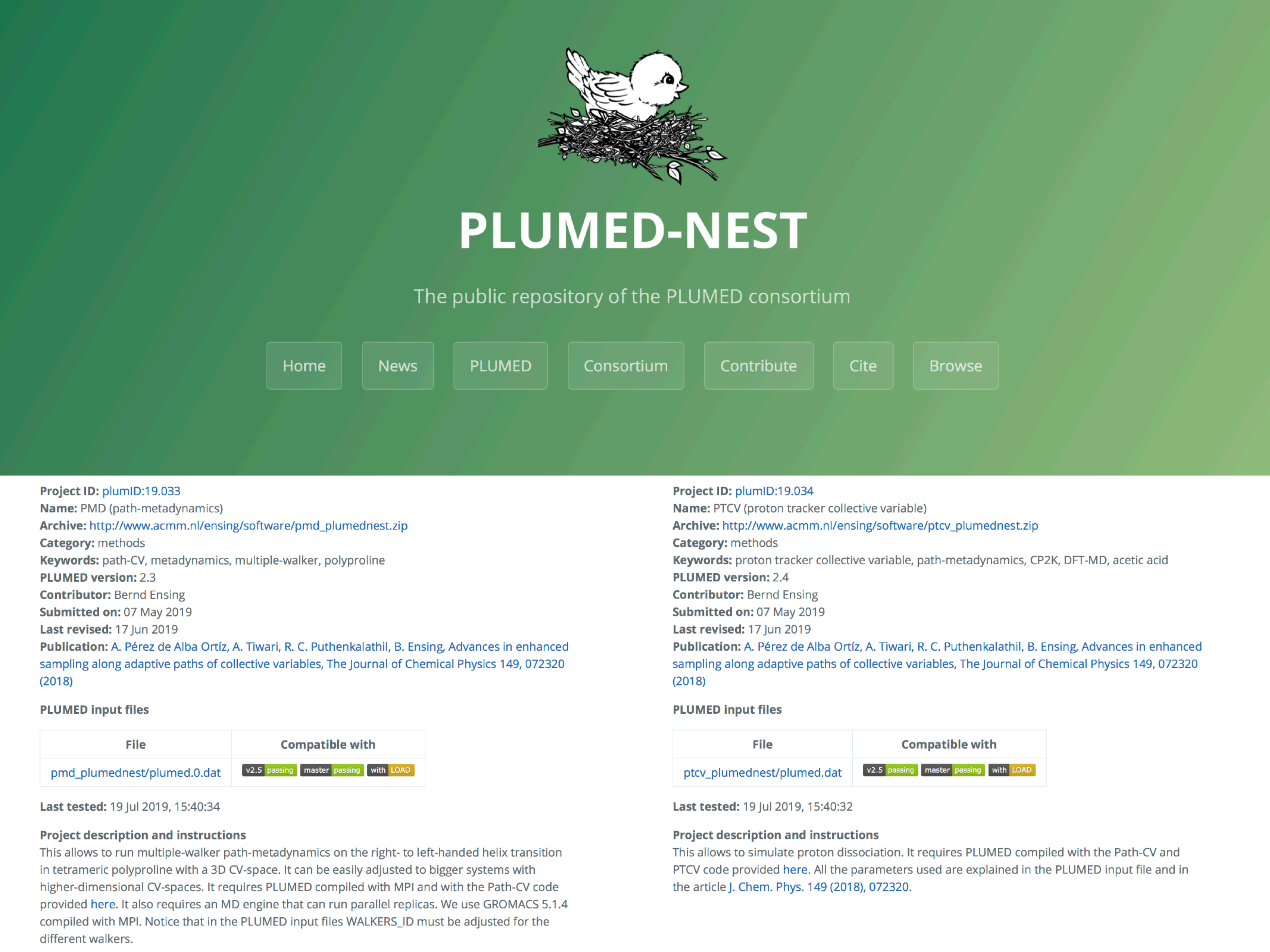
Notice that these are the definitions of s and z that are used by **PATH** when the GPATH option is employed. The reason for this is that the adaptive path method implemented in this action was inspired by the work of Diaz and Ensing in which these formula were used [62]. To learn more about how the path is adapted we strongly recommend reading this paper.

The input below provides an example that shows how the adaptive path works. The path is updated every 50 steps of MD based on the data accumulated during the preceding 50 time steps.

```
d1: DISTANCE ATOMS=1,2 COMPONENTS
pp: ADAPTIVE_PATH TYPE=EUCLIDEAN FIXED=2,5 UPDATE=50 WFILE=out-path.pdb WSTRIDE=50 REFERENCE=mypath.pdb
PRINT ARG=d1.x,d1.y,pp.* FILE=colvar
```

In the case above the distance between frames is calculated based on the x and y components of the vector connecting atoms 1 and 2. As such an extract from the input reference path (mypath.pdb) would look as follows:

```
REMARK ARG=d1.x,d1.y d1.x=1.12 d1.y=-.60
END
REMARK ARG=d1.x,d1.y d1.x=.99 d1.y=-.45
END
REMARK ARG=d1.x,d1.y d1.x=.86 d1.y=-.30
```



PLUMED-NEST

The public repository of the PLUMED consortium

Project ID: plumID:19.033
Name: PMD (path-metadynamics)
Archive: http://www.acmm.nl/ensing/software/pmd_plumednest.zip
Category: methods
Keywords: path-CV, metadynamics, multiple-walker, polyproline
PLUMED version: 2.3
Contributor: Bernd Ensing
Submitted on: 07 May 2019
Last revised: 17 Jun 2019
Publication: A. Pérez de Alba Ortíz, A. Tiwari, R. C. Puthenkalathil, B. Ensing, Advances in enhanced sampling along adaptive paths of collective variables, The Journal of Chemical Physics 149, 072320 (2018)

PLUMED input files

File	Compatible with
pmd_plumednest/plumed.0.dat	v2.5 passing master passing with LOAD

Last tested: 19 Jul 2019, 15:40:34

Project description and instructions
This allows to run multiple-walker path-metadynamics on the right- to left-handed helix transition in tetrameric polyproline with a 3D CV-space. It can be easily adjusted to bigger systems with higher-dimensional CV-spaces. It requires PLUMED compiled with MPI and with the Path-CV code provided here. It also requires an MD engine that can run parallel replicas. We use GROMACS 5.1.4 compiled with MPI. Notice that in the PLUMED input files WALKERS_ID must be adjusted for the different walkers.

Project ID: plumID:19.034
Name: PTCV (proton tracker collective variable)
Archive: http://www.acmm.nl/ensing/software/ptcv_plumednest.zip
Category: methods
Keywords: proton tracker collective variable, path-metadynamics, CP2K, DFT-MD, acetic acid
PLUMED version: 2.4
Contributor: Bernd Ensing
Submitted on: 07 May 2019
Last revised: 17 Jun 2019
Publication: A. Pérez de Alba Ortíz, A. Tiwari, R. C. Puthenkalathil, B. Ensing, Advances in enhanced sampling along adaptive paths of collective variables, The Journal of Chemical Physics 149, 072320 (2018)

PLUMED input files

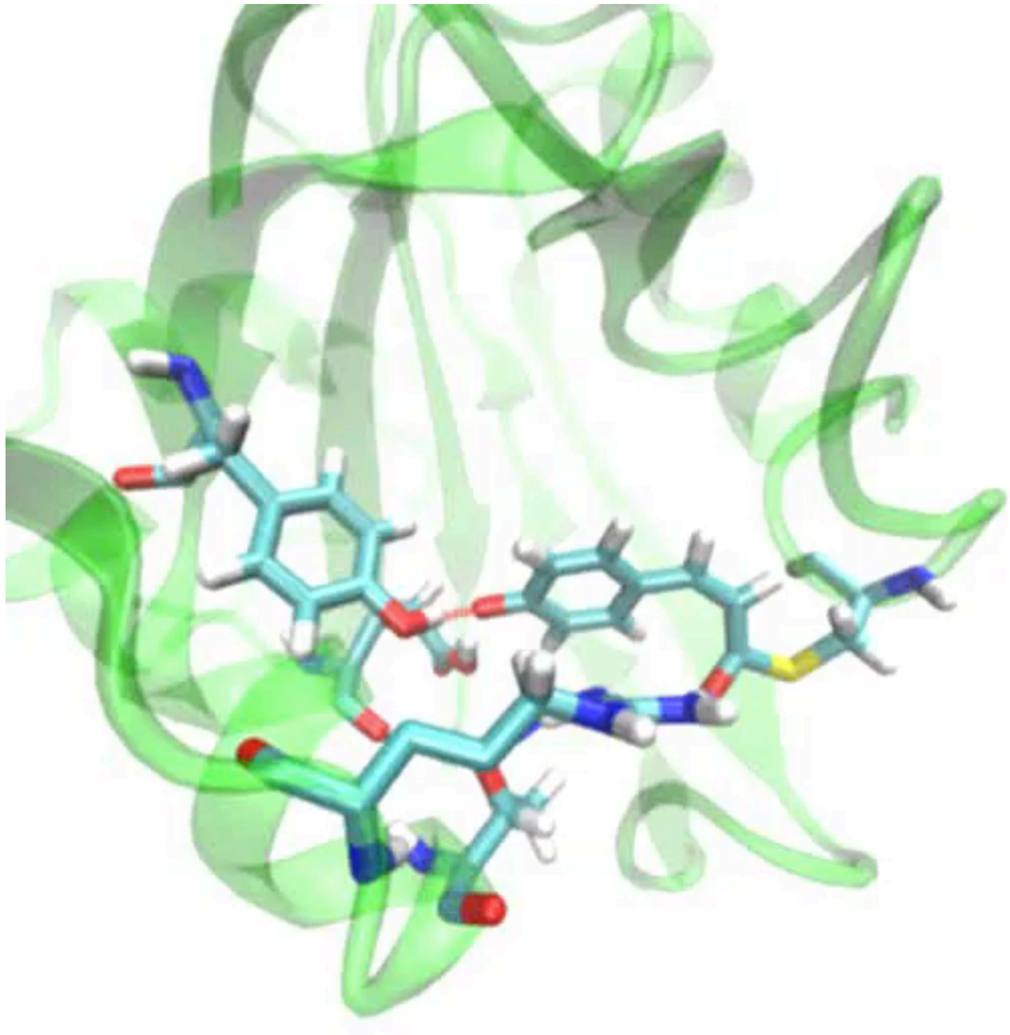
File	Compatible with
ptcv_plumednest/plumed.dat	v2.5 passing master passing with LOAD

Last tested: 19 Jul 2019, 15:40:32

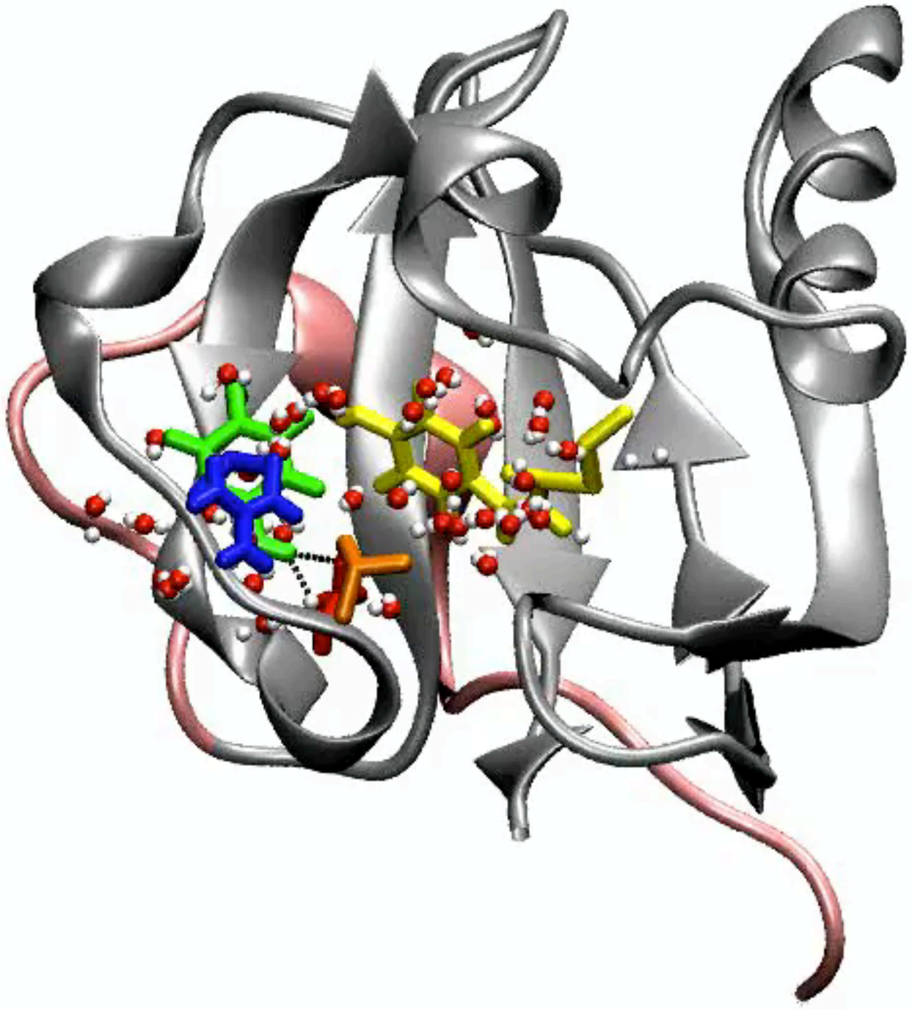
Project description and instructions
This allows to simulate proton dissociation. It requires PLUMED compiled with the Path-CV and PTCV code provided here. All the parameters used are explained in the PLUMED input file and in the article J. Chem. Phys. 149 (2018), 072320.

Photoactive Yellow Protein

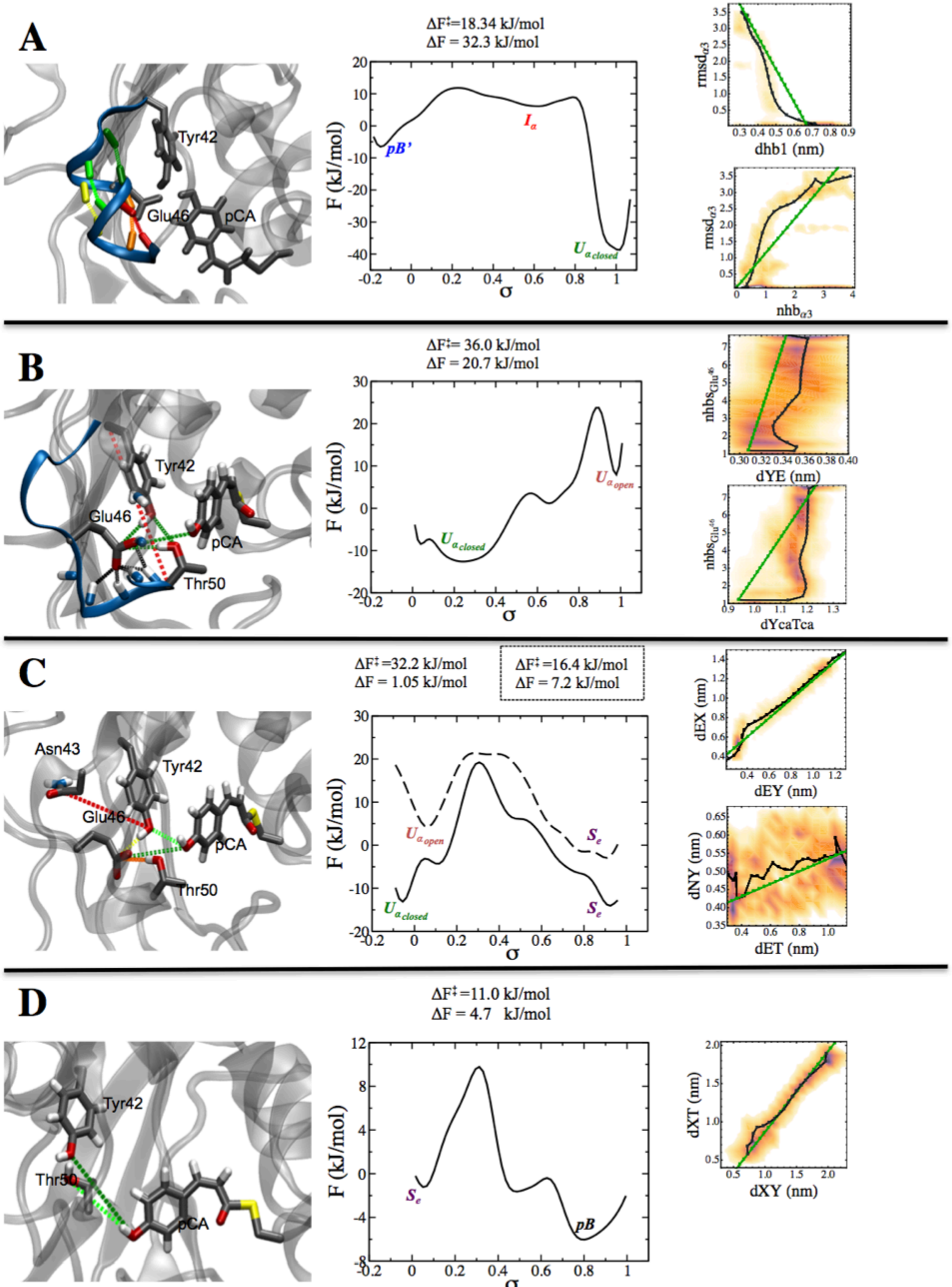
Proton transfer step



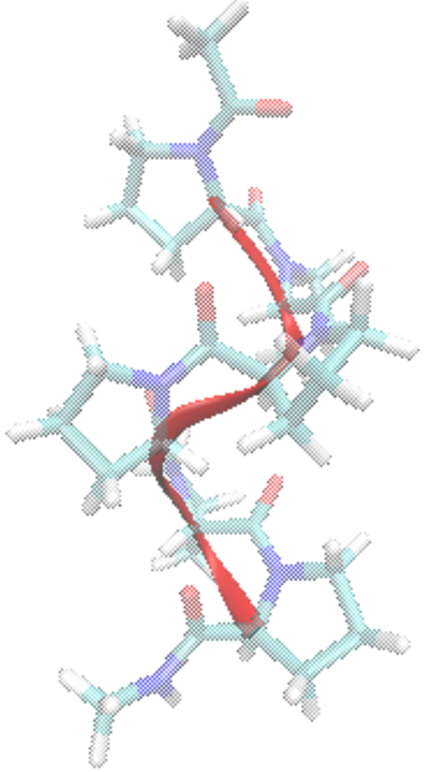
Conformational transformation



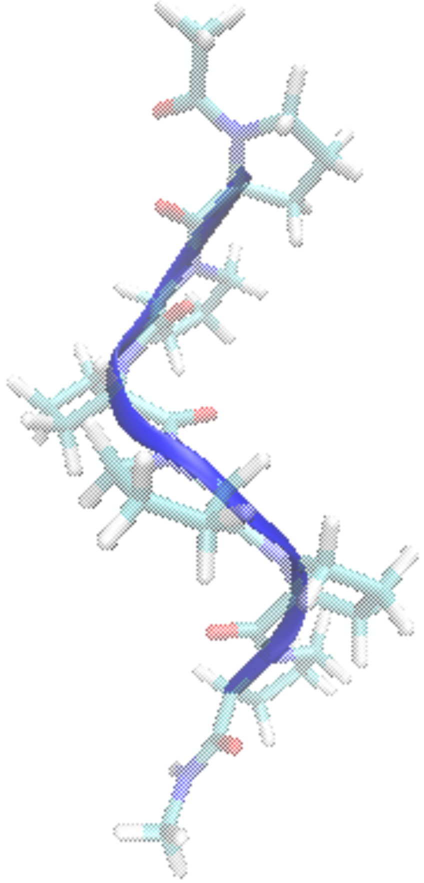
Predicting the signaling state of Photoactive Yellow Protein.
J. Vreede, W. Crielaard, K.J. Hellingwerf, and P.G. Bolhuis
Biophys J. **88** (2005), 3525 DOI: 10.1529/biophysj.104.055103


A
ΔF‡=18.34 kJ/mol
ΔF = 32.3 kJ/mol
Tyr42
Glu46
pCA
pB'
I_α
$U_{\alpha closed}$
F (kJ/mol)
σ
$rmsd_{\alpha 3}$
dhb1 (nm)
$nhb_{\alpha 3}$
B
ΔF‡= 36.0 kJ/mol
ΔF = 20.7 kJ/mol
Tyr42
Glu46
pCA
Thr50
$U_{\alpha closed}$
$U_{\alpha open}$
F (kJ/mol)
σ
$nhbs_{Glu^{46}}$
dYE (nm)
dYcaTca
C
ΔF‡=32.2 kJ/mol
ΔF = 1.05 kJ/mol
ΔF‡=16.4 kJ/mol
ΔF = 7.2 kJ/mol
Asn43
Tyr42
Glu46
pCA
Thr50
$U_{\alpha open}$
$U_{\alpha closed}$
S_e
F (kJ/mol)
σ
dEX (nm)
dEY (nm)
dNY (nm)
dET (nm)
D
ΔF‡=11.0 kJ/mol
ΔF = 4.7 kJ/mol
Tyr42
Thr50
pCA
S_e
pB
F (kJ/mol)
σ
dXT (nm)
dXY (nm)

polyproline oligomers



PPI
right-handed helix
$\phi=-75^{o}$, $\psi=160^{o}$, $\omega=0^{o}$



PPII
left-handed helix
$\phi=-75^{o}$, $\psi=146$, $\omega=180^{o}$

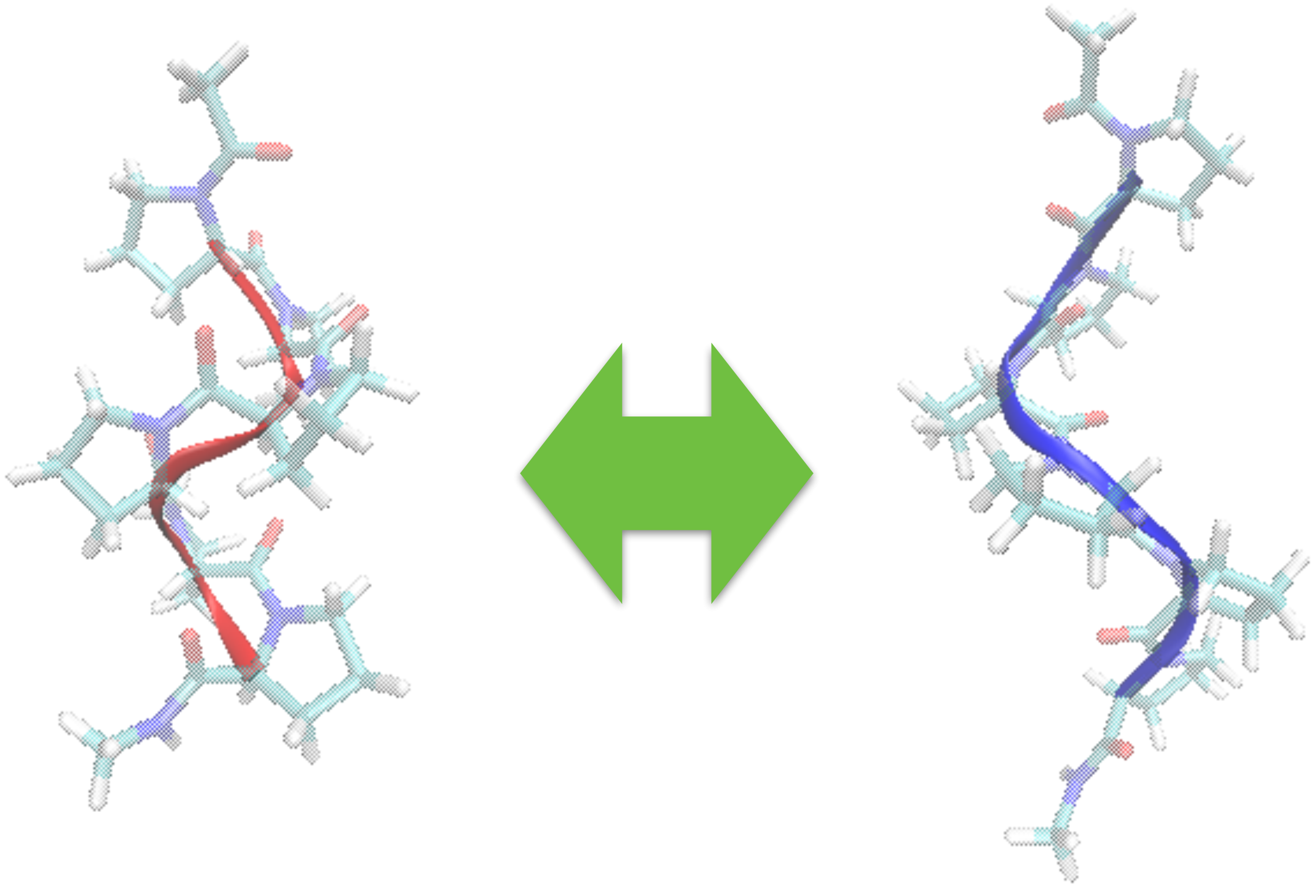
Settings:

- Gromacs + Plumed
- Amber ff99SB, no solvent
- NVT, T=300 K

Previous work:

- ABMD, SMD, HT-REMD,…
Moradi, Sagui, Babin, Roland, et al.
- Exp. Shi, Holliday, Clemmer, et al.

polyproline oligomers



PPI

right-handed helix

$\phi=-75^{o}$, $\psi=160^{o}$, $\omega=0^{o}$

PPII

left-handed helix

$\phi=-75^{o}$, $\psi=146$, $\omega=180^{o}$

PMD Settings:

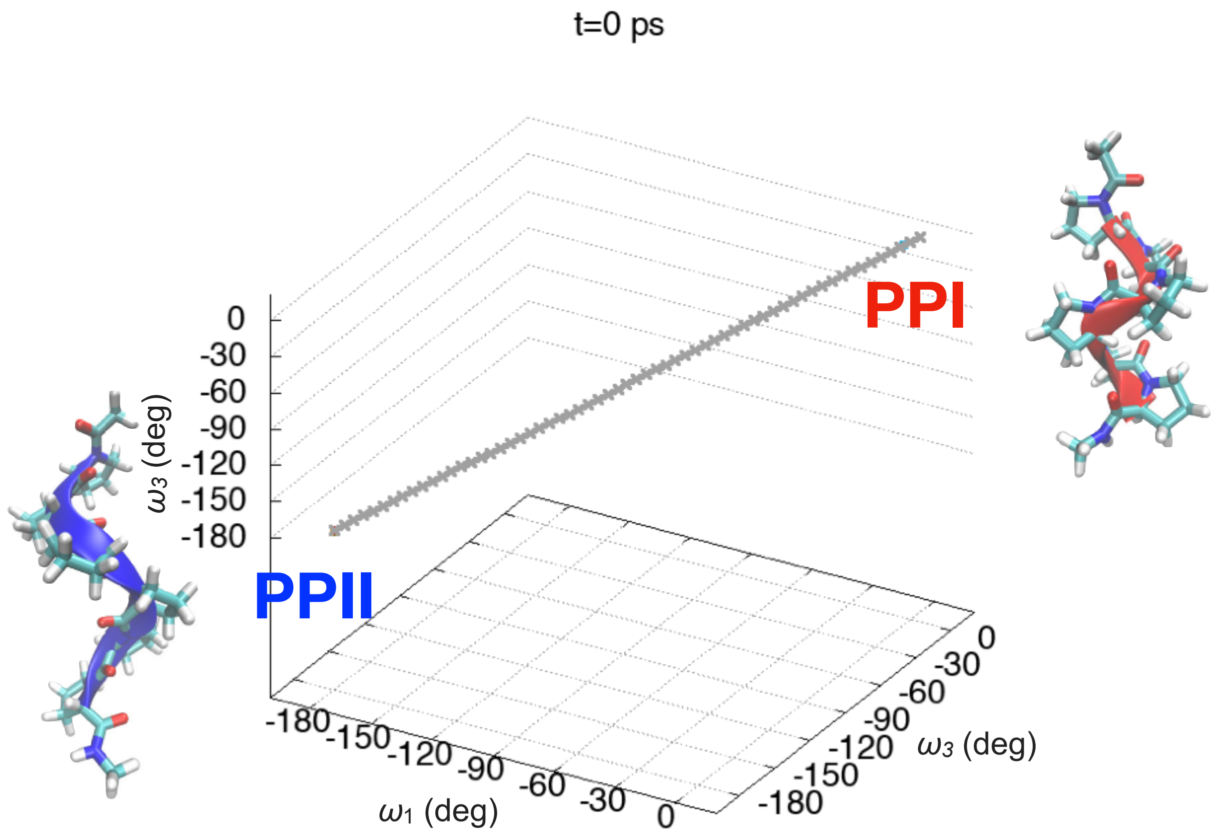
- 20 path nodes per CV + ends
- path update: 0.5 ps
- half live: 2.0 ps
- tube + wall potentials

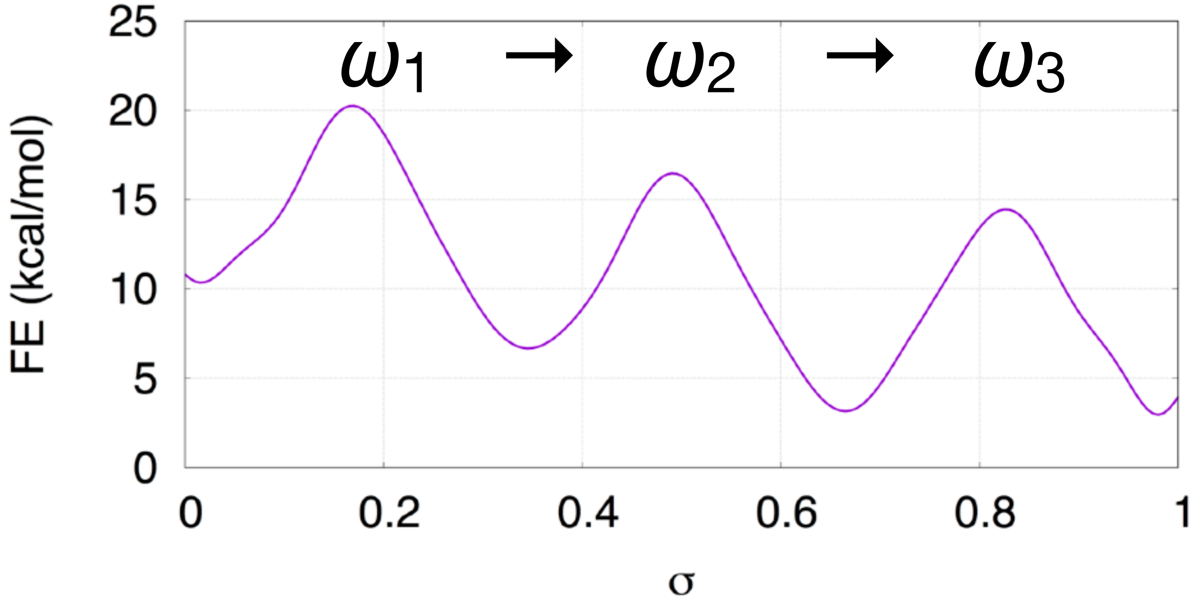
Metadynamics settings:

- H=0.05 kcal/mol
- W=0.05
- dt=1.0 ps
- potential on a grid

Multiple-walker PMD

- Polyproline chains Ace-(Pro)$_n$-Nme
- Two distinct helical structures:
  - **PPI:** right-handed (***ω*=0º**)
  - **PPII:** left-handed (***ω*=±180º**)
- 8 walkers update the path and the free energy in **parallel**
- **Zipper-like** mechanisms starting from Ace (with *n* = 4, 5, 6...) also seen in experiments

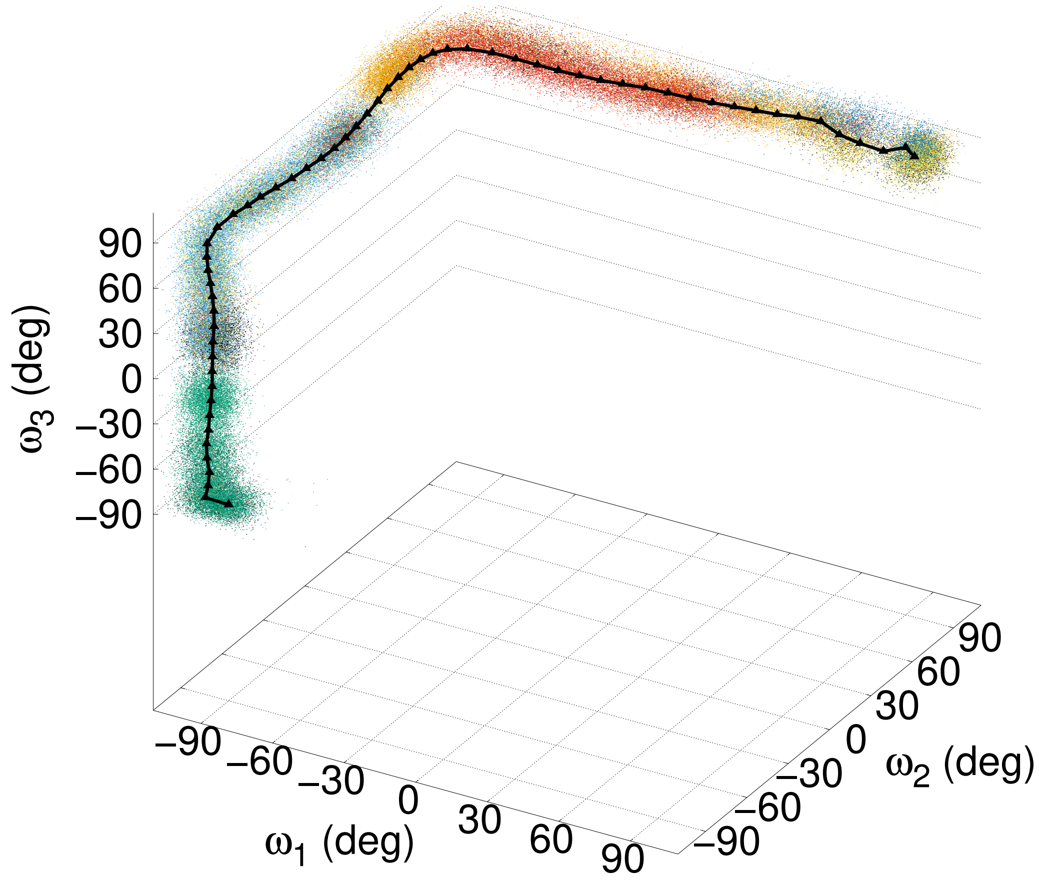




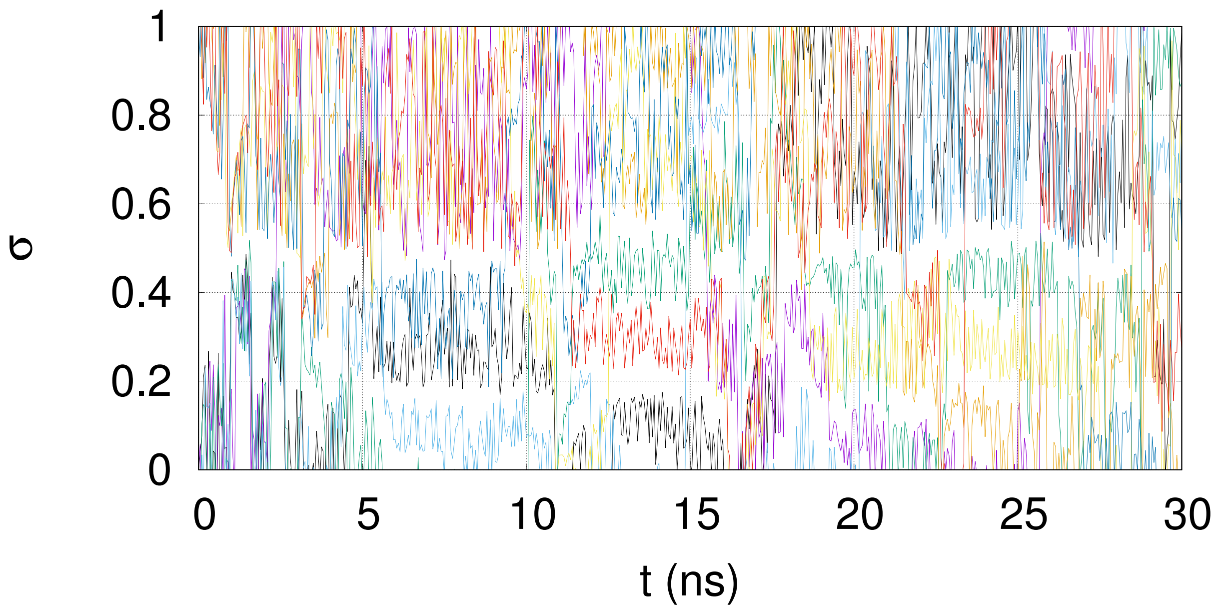
Moradi et al., 2010; Shi et al., 2014;
Pérez de Alba Ortíz, Tiwari, Puthenkalathil and Ensing, 2018

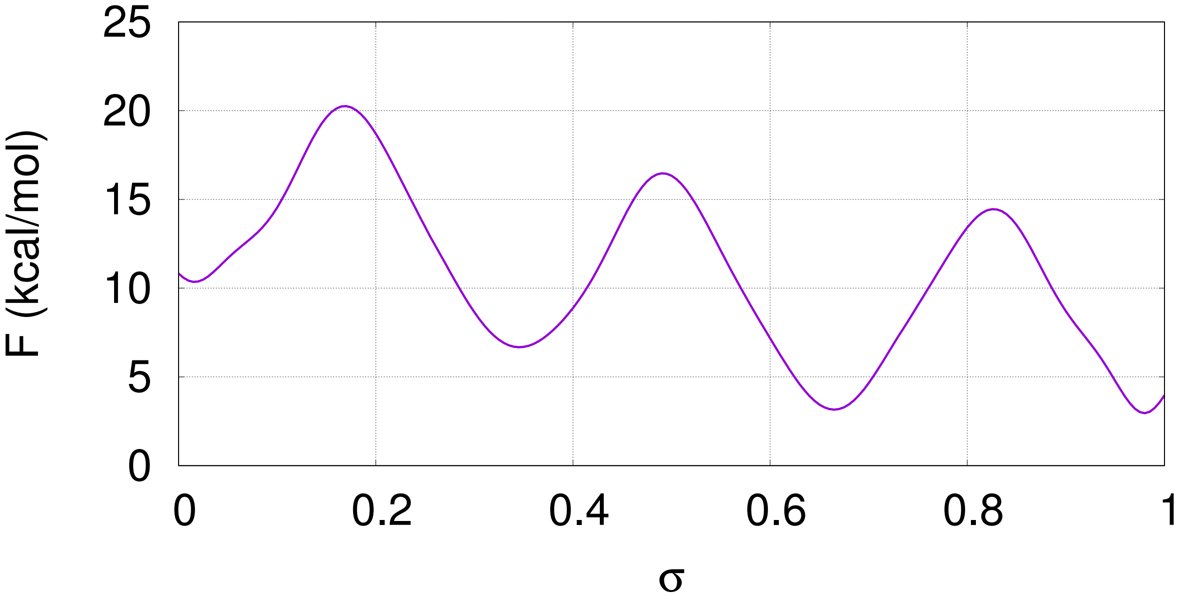
Multiple Walker PMD

- tetramer oligomer
- 8 walkers
- zipper mechanism observed
- ω1 → ω2 → ω3
- barriers: ca. 10 kcal/mol
- dF: 7 kcal/mol



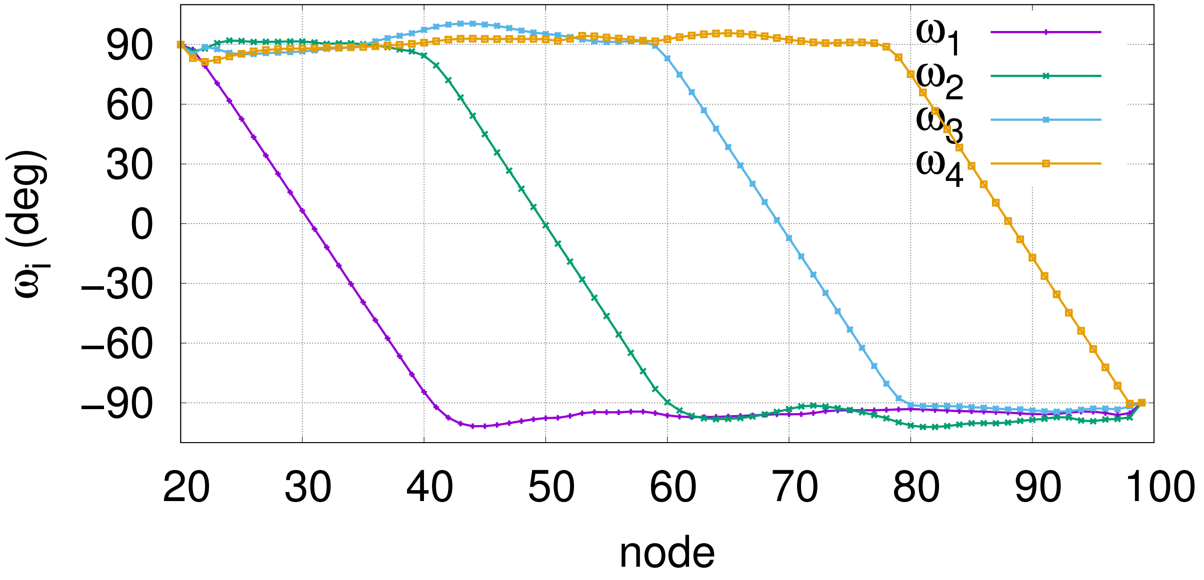
n-oligomer, n-1 dihedrals
(n-1)! pathways

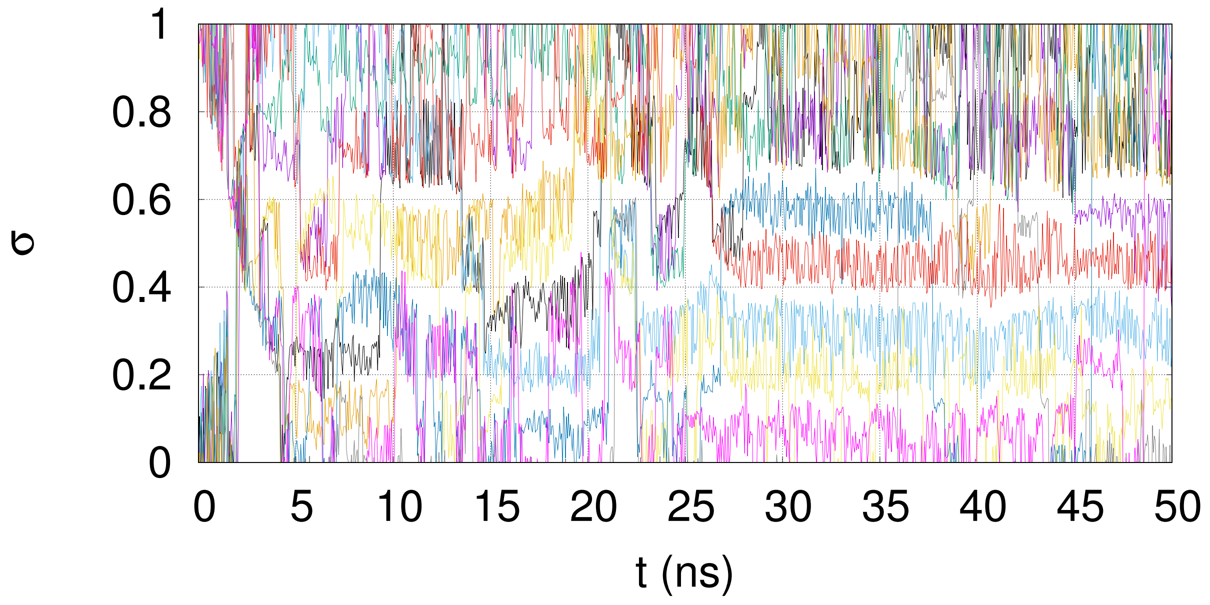


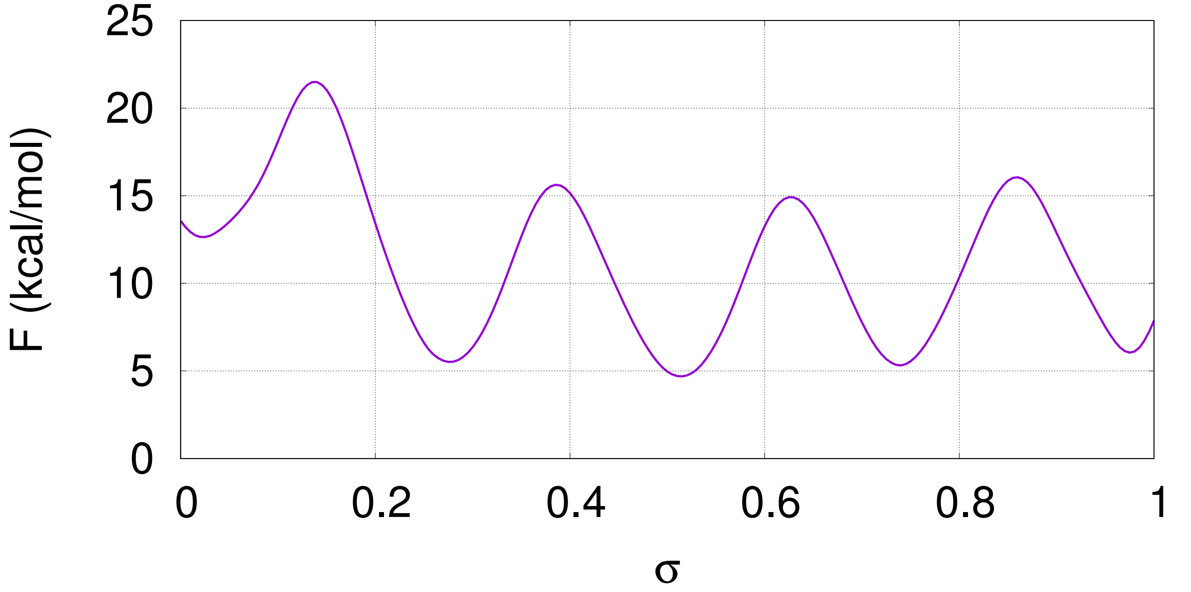


Pentamer

- pentamer oligomer
- 10 walkers
- zipper mechanism observed
- ω1 → ω2 → ω3 → ω4
- different mechanism for smaller or less frequent Gaussians
- barriers: ca. 10 kcal/mol
- dF: 7 kcal/mol



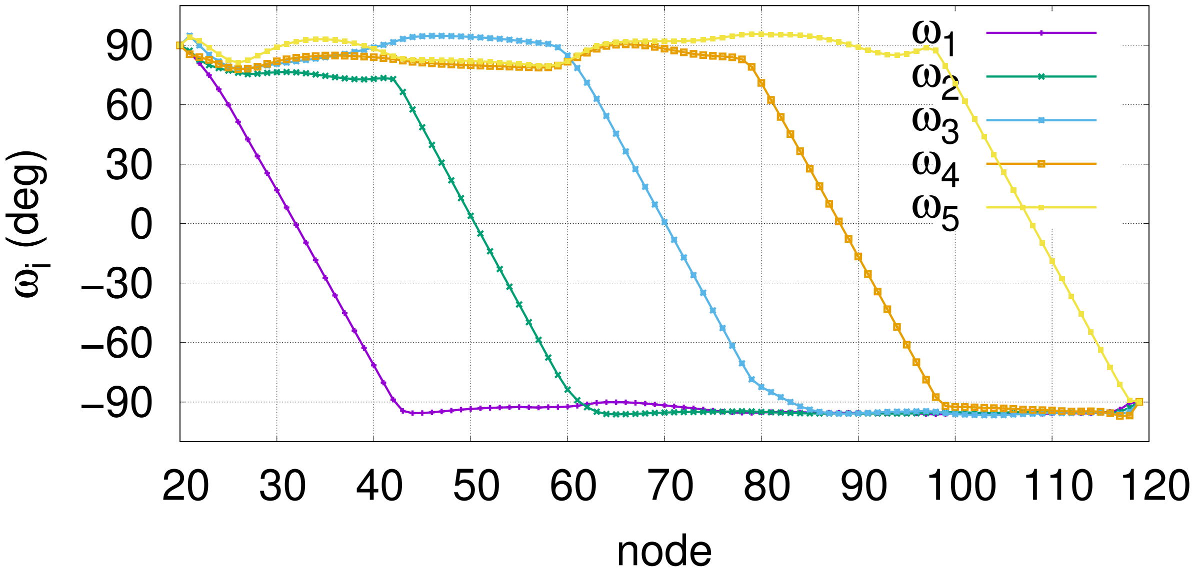


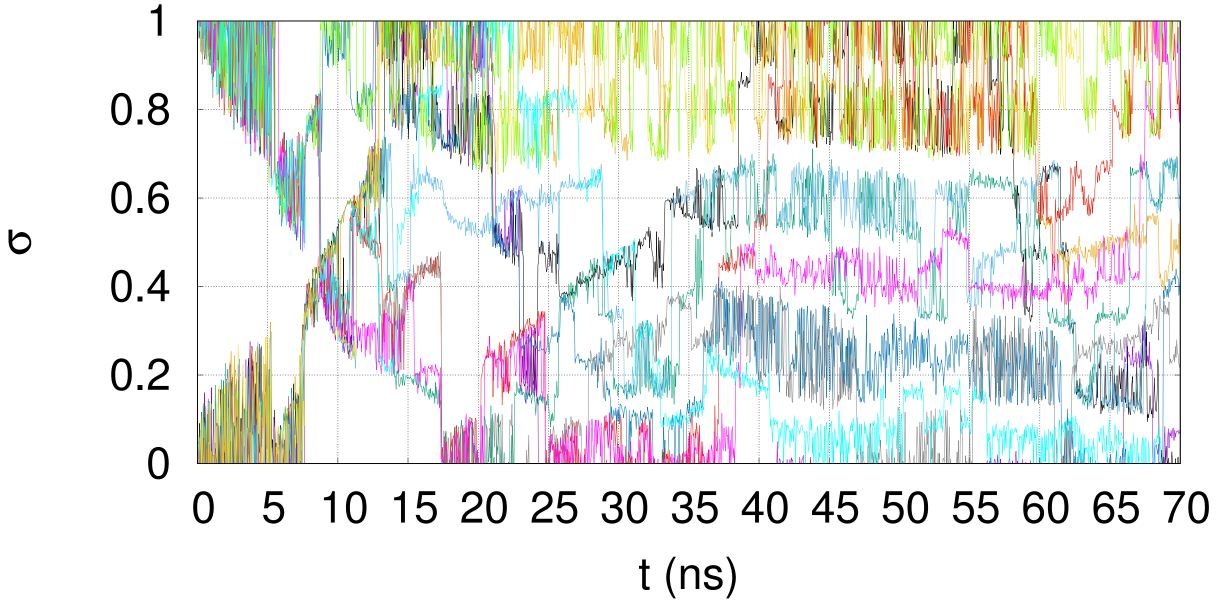


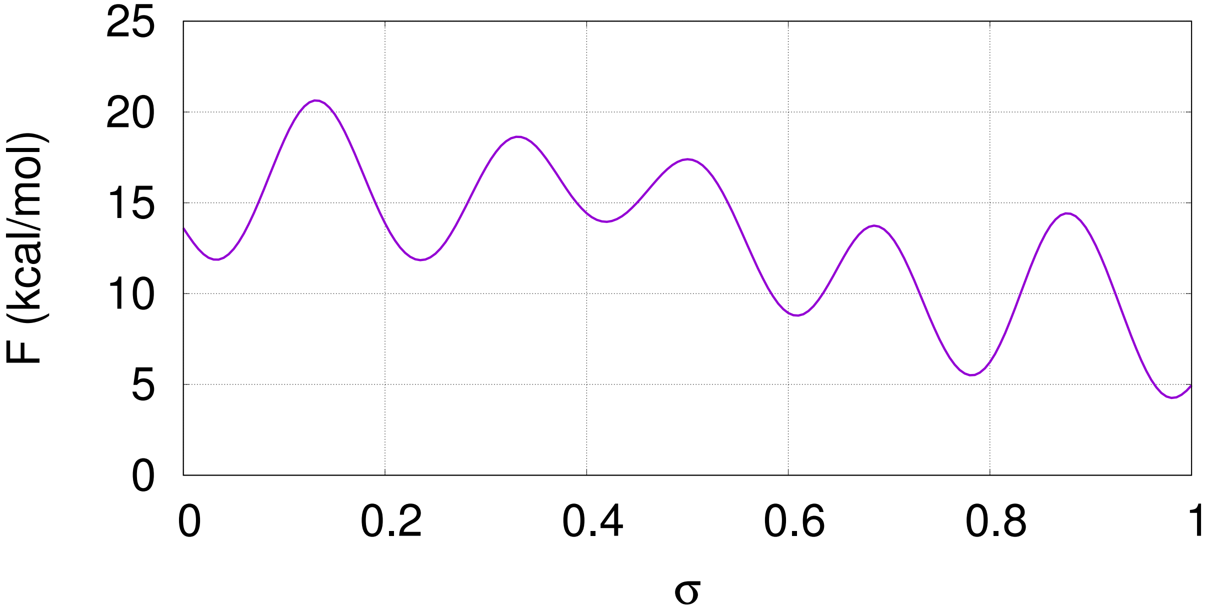
Hexamer

- pentamer oligomer
- 10 walkers
- different zipper mechanisms observed, depending on initial conditions

path-degeneracy limits finding the global minimum path

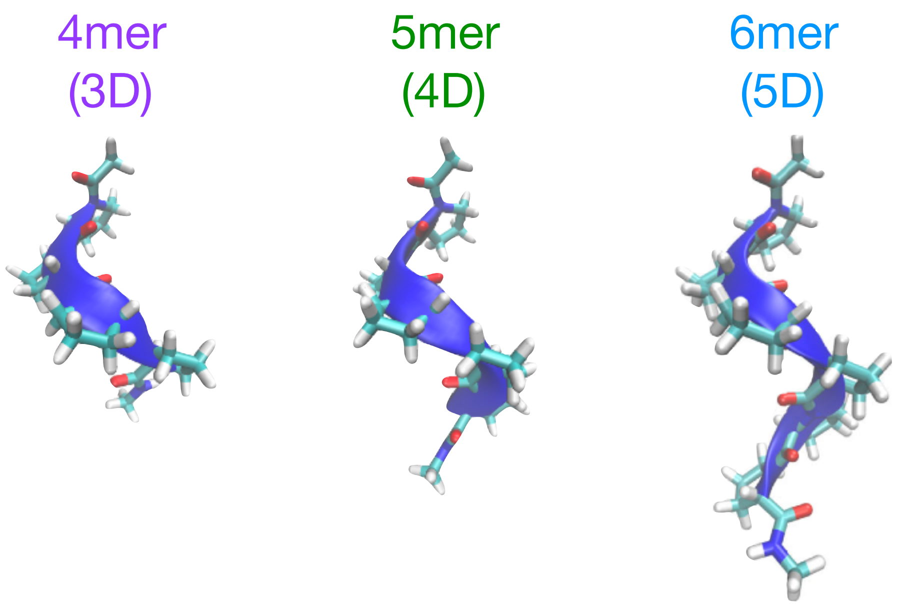




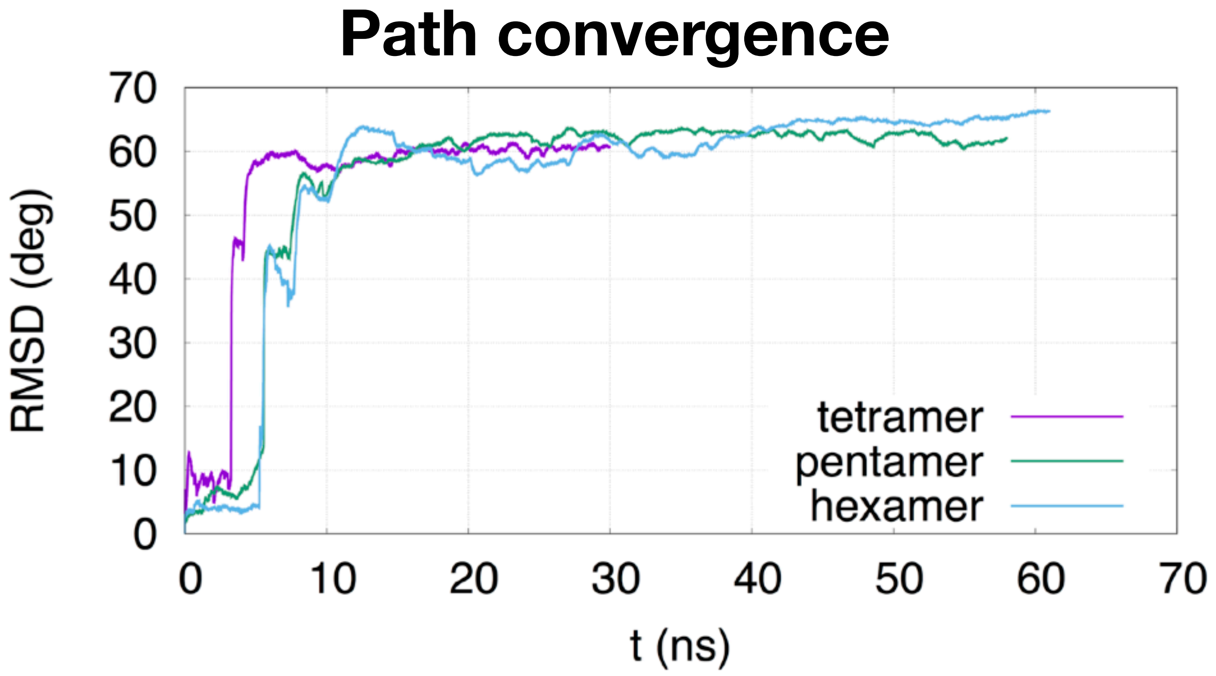


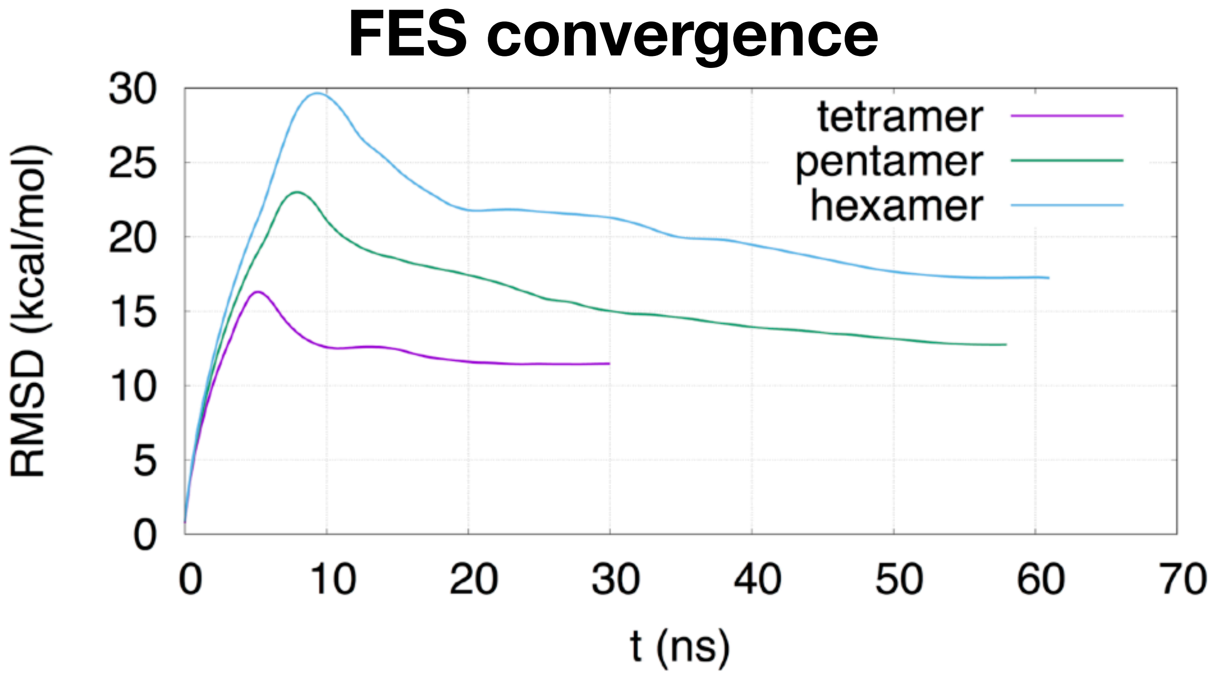
PMD performance

- PMD is able to converge an average transition path and the free energy along it, with a **sublinear** rise in cost w.r.t. CV dimensionality



- But the number of possible paths grows factorially with CV-dimensionality!

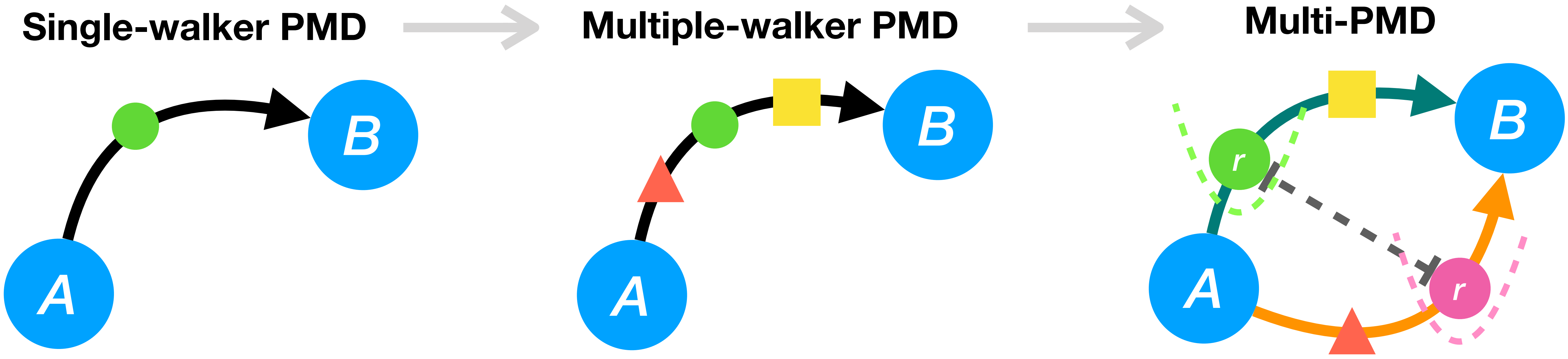




Pérez de Alba Ortíz, Tiwari, Puthenkalathil and Ensing, 2018

MuWaMuPaMetaDyn

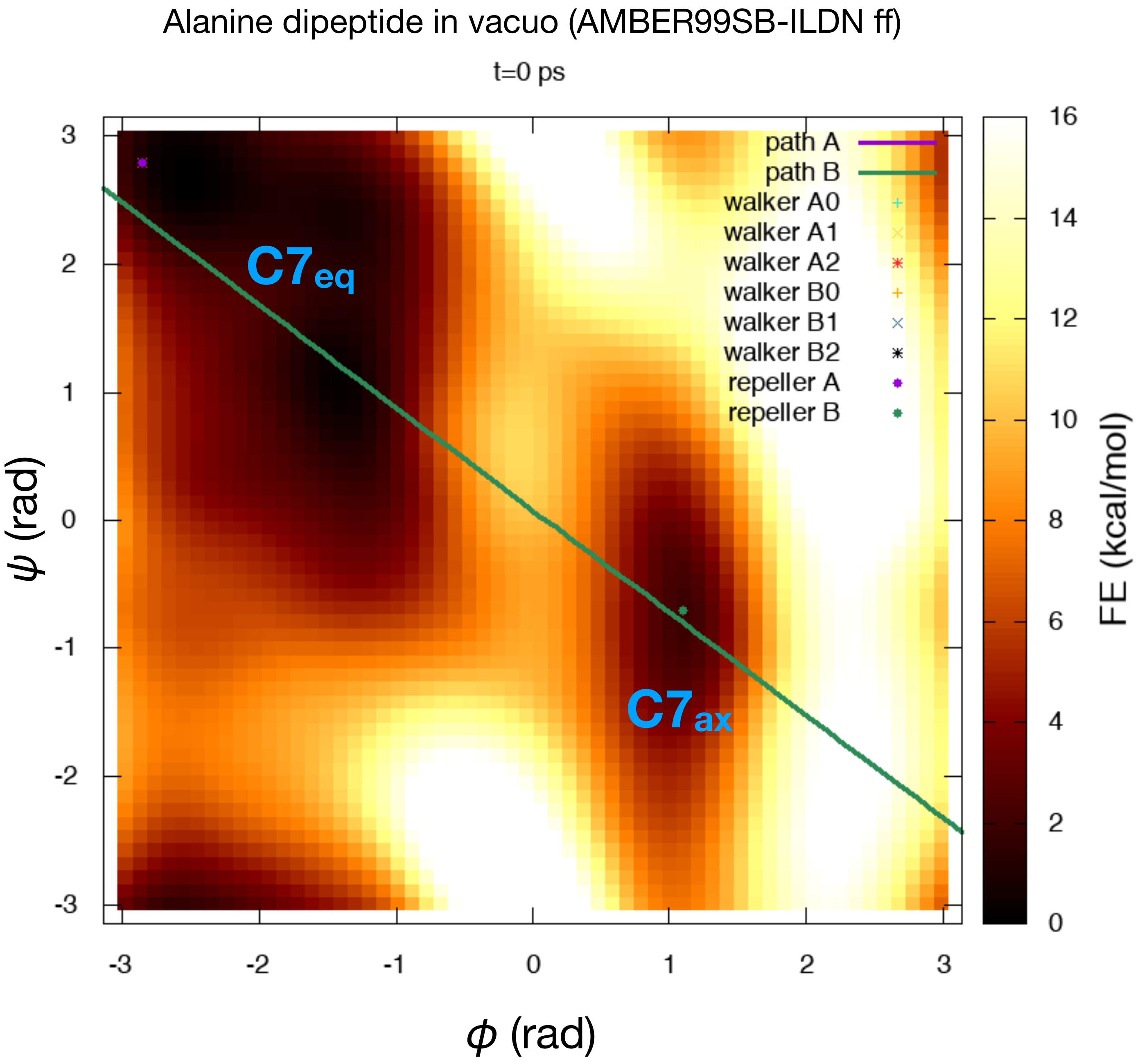
- **Multiple walkers** can simultaneously explore **multiple paths** representing different **mechanisms**
- Some walkers can be used as **repellers**

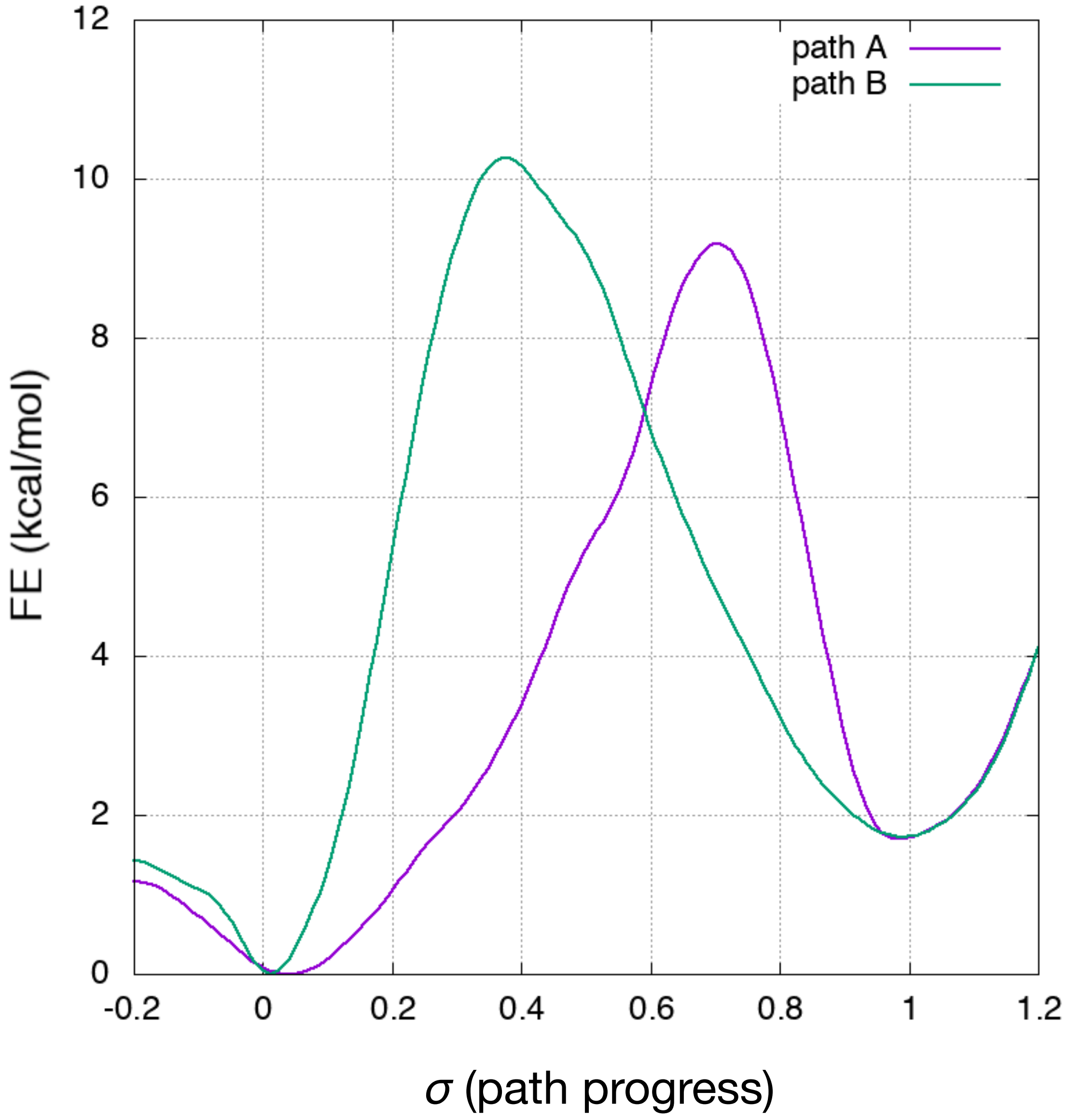


Single-walker PMD → **Multiple-walker PMD** → **Multi-PMD**

Pérez de Alba Ortíz and Ensing, 2019, *In press;*
Pérez de Alba Ortíz and Ensing, *In prep.*

Alanine dipeptide revisited

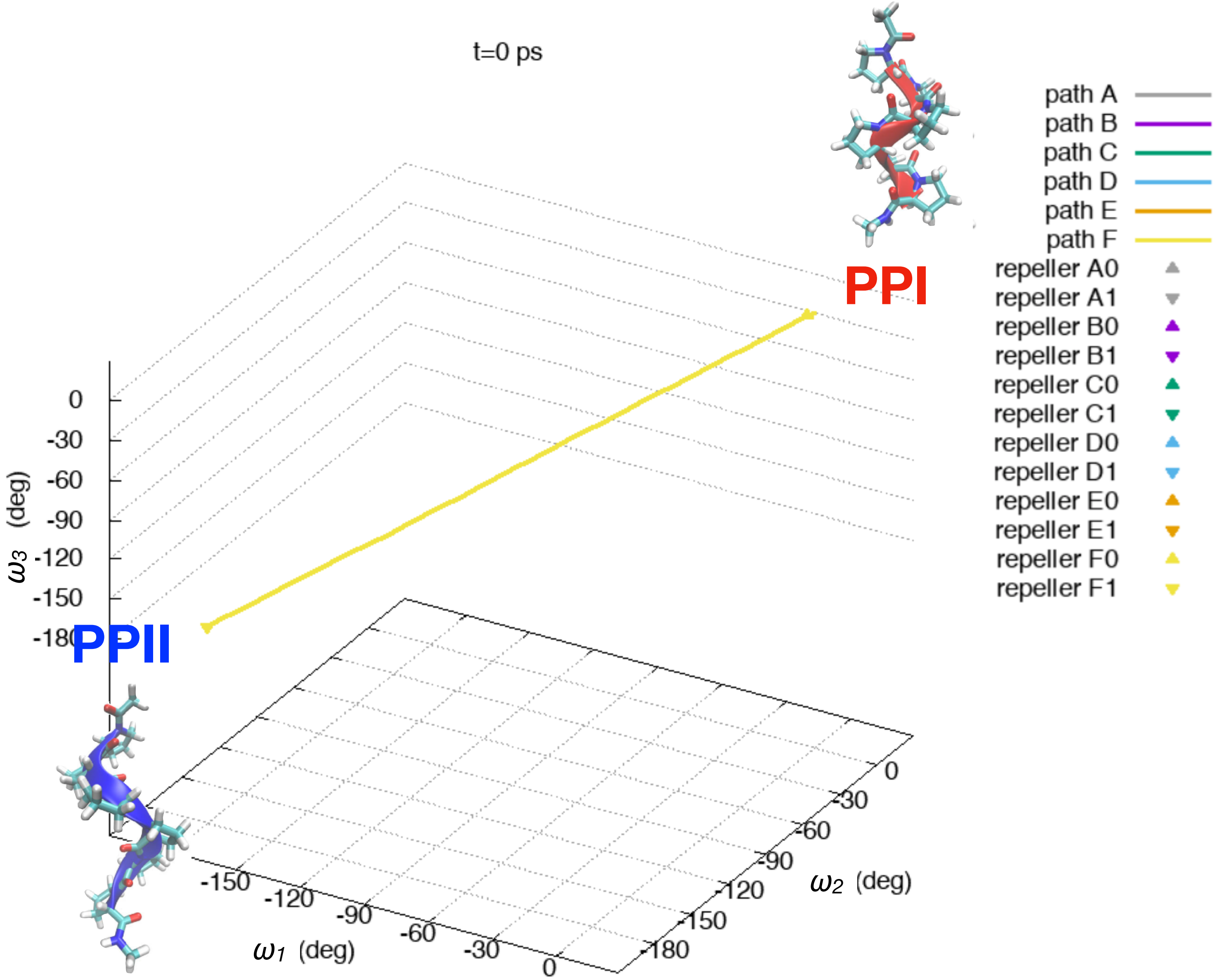




Pérez de Alba Ortíz and Ensing, *In prep.*

Polyproline revisited

- Six PPI→PPII paths
- 2 repellers per path
- Repellers steered to $\sigma = 0.33$ and $\sigma = 0.66$
- Min. distance^2 between repeller-pair average CV-positions restrained to 1.4 rad



Pérez de Alba Ortíz and Ensing, *In prep.*

Summary

- Molecular Dynamics (MD) simulations allow for sampling molecular configurations and visualising trajectories
- Ensemble averages are obtained as time averages. Dynamical properties can be computed using time-correlation functions.
- Chemical reactions and other activated molecular transitions are rare events.
- The Bennett-Chandler approach, or reactive flux method, allows for accurate estimation of reaction rates, by combining a free energy calculation (giving the transition state theory estimate of the rate) with a correction using the transmission coefficient.
- Free energy methods require a collective variable (or reaction coordinate).
- Metadynamics builds on the fly a repulsive bias potential to accelerate sampling and probe the free energy landscape as a function of one or more (max 3) collective variables.
- Path-metadynamics finds the reaction coordinate as a function of several (many) collective variables and computes the (1D) free energy profile along the mean transition path.

Alberto Pérez de Alba Ortíz

Rakesh Chandran
Puthenkalathil

Ambuj Tiwari

bibliography

Laio, A.; Parrinello, M. Escaping free-energy minima. Proc. Natl. Acad. Sci. U.S.A. 2002, 99, 12562-12566.

Iannuzzi, M.; Laio, A.; Parrinello, M. Efficient exploration of reactive potential energy surfaces using car-parrinello molecular dynamics. Phys. Rev. Lett. 2003, 90, No. 238302.

Martonak, R.; Laio, A.; Parrinello, M. Predicting crystal structures: the Parrinello-Rahman method revisited. Phys. Rev. Lett. 2003, 90, No. 075503.

Bernd Ensing, Alessandro Laio, Francesco L. Gervasio, Michele Parrinello, and Michael L. Klein A minimum free energy reaction path for the E2 reaction between fluoro ethane and a fluoride ion. J. Am. Chem. Soc. 126 (2004), 9492-9493

Micheletti, C.; Laio, A.; Parrinello, M. Reconstructing the density of states by history-dependent metadynamics. Phys. Rev. Lett. 2004, 92, No. 170601.

Ensing, B., Klein, M. L. Perspective of the reactions between F- and CH3CH2F: the free energy landscape of the E2 and SN2 reaction channels. Proc. Natl. Acad. Sci. U.S.A. 2005, 102, 6755-6759.

Donadio, D.; Raiteri, P.; Parrinello, M. Topological defects and bulk melting of hexagonal ice. J. Phys. Chem. B 2005, 109, 5421-5424.

Ensing, B.; Laio, A.; Parrinello, M.; Klein, M. L. A recipe for the computation of the free energy barrier and the lowest free energy path of concerted reactions. J. Phys. Chem. B 2005, 109, 6676.

Laio, A.; Rodriguez-Fortea, A.; Gervasio, F. L.; Ceccarelli, M.; Parrinello, M. Assessing the accuracy of metadynamics. J. Phys. Chem. B 2005, 109, 6714-6721.

D. Donadio, P. Raiteri, M. Parrinello, Topological defects and bulk melt- ing of hexagonal ice, J. Phys. Chem. B 109 (12) (2005) 5421–5424.

Bernd Ensing, Marco De Vivo, Zhiwei Liu, Preston Moore, and Michael L. Klein. Metadynamics as a tool for exploring the free energy landscape of chemical reactions. Acc. Chem. Res. 39 (2006), 73-81

P. Raiteri, A. Laio, F. Gervasio, C. Micheletti, M. Parrinello, Efficient reconstruction of complex free energy landscapes by multiple walkers metadynamics, J. Phys. Chem. B 110 (2006) 3533–3539.

G. Bussi, F. L. Gervasio, A. Laio, M. Parrinello, Free-energy landscape for beta hairpin folding from combined parallel tempering and metady- namics, J. Am. Chem. Soc. 128 (41) (2006) 13435–41.

S. Piana, A. Laio, A bias-exchange approach to protein folding, J. Phys. Chem. B 111 (17) (2007) 4553–9.

D. Branduardi, F. L. Gervasio, M. Parrinello, From A to B in free energy space, J. Chem. Phys. 126 (5) (2007) 054103.

A. Barducci, G. Bussi, M. Parrinello, Well-tempered metadynamics: A smoothly converging and tunable free-energy method, Phys. Rev. Lett. 100 (2008) 020603.

A. Laio, F. L. Gervasio, Metadynamics: a method to simulate rare events and reconstruct the free energy in biophysics, chemistry and material science, Rep. Prog. Phys. 71 (2008) 126601.

M. Bonomi, D. Branduardi, G. Bussi, C. Camilloni, D. Provasi, P. Raiteri, D. Donadio, F. Marinelli, F. Pietrucci, R.A. Broglia and M. Parrinello. PLUMED: a portable plugin for free-energy calculations with molecular dynamics, Comp. Phys. Comm. 2009 vol. 180 (10) pp. 1961-1972.

C. Michel, A. Laio, A. Milet, Tracing the entropy along a reactive path- way: The energy as a generalized reaction coordinate, J. Chem. Theory Comput. 5 (9) (2009) 2193–2196.

M. Bonomi, A. Barducci, M. Parrinello, Reconstructing the equilibrium Boltzmann distribution from well-tempered metadynamics, J. Comput. Chem. 30 (11) (2009) 1615–1621.

Y. Crespo, F. Marinelli, F. Pietrucci, A. Laio, Metadynamics con- vergence law in a multidimensional system, Phys. Rev. E 81 (2010) 055701

A. Barducci, M. Bonomi, M. Parrinello, Metadynamics, WIREs Comput. Mol. Sci. 1 (2011) 826.

G. A. Tribello, J. Cuny, H. Eshet, M. Parrinello, Exploring the free energy surfaces of clusters using reconnaissance metadynamics, J. Chem. Phys 135 (11) (2011) 114109.

Grisell Díaz Leines and Bernd Ensing. Path finding on high-dimensional free energy landscapes. Phys. Rev. Lett. 109 (2012), 020601