Transition Path Theory

Winterschool on
Theoretical Chemistry and Spectroscopy
9-13 December 2013, Han-sur-Lesse, Belgium

Bernd Ensing
Computation Chemistry
University of Amsterdam

Summary

part 1

A stochastic hopping between states is Markovian if the probability to reach the next state does not depend on the past (no memory)

- reversible Markov chains (detailed balance)
- ergodicity
- spectral decomposition (convergence to equilibrium)

part 2

Markov State Modeling applied as analysis tool of the dynamics of a small peptide

- Analysis of molecular dynamics simulation
- Clustering based on structures
- Relation to free energy landscape based on cluster size
- Transition probability between clusters (states)
- Using the MSM the dynamics can be sampled in a coarse-grain manner, e.g. using kinetic Monte-Carlo

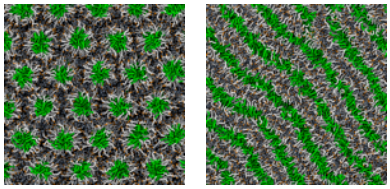
Content

- Lecture 1: Markov processes
- Lecture 2: State clustering
- Lecture 3: Transition Path Theory
- Lecture 4: Sampling Reactive Events

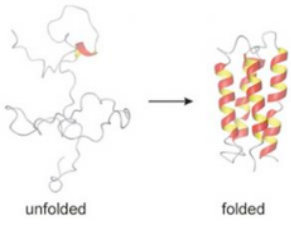
Molecular Transitions are Rare Events

chemical reactions

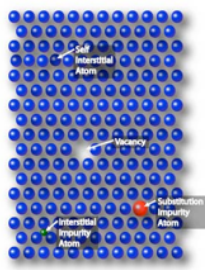
phase transitions



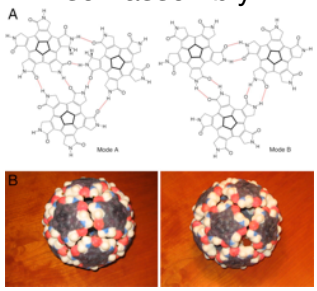
protein folding



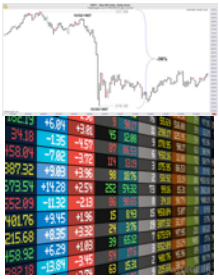
defect diffusion



self-assembly



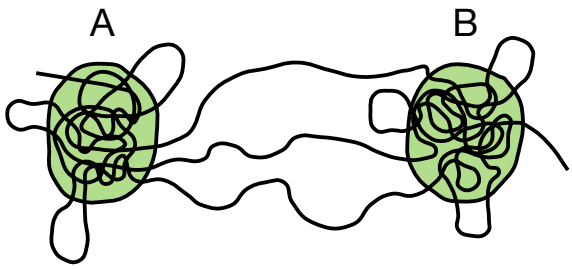
stock exchange crash



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,....

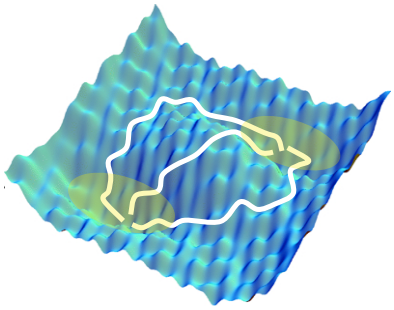
A very long trajectory



- 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

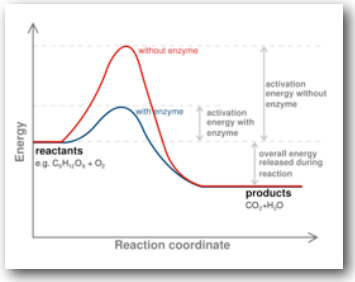
The free energy has local minima separated by barriers

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

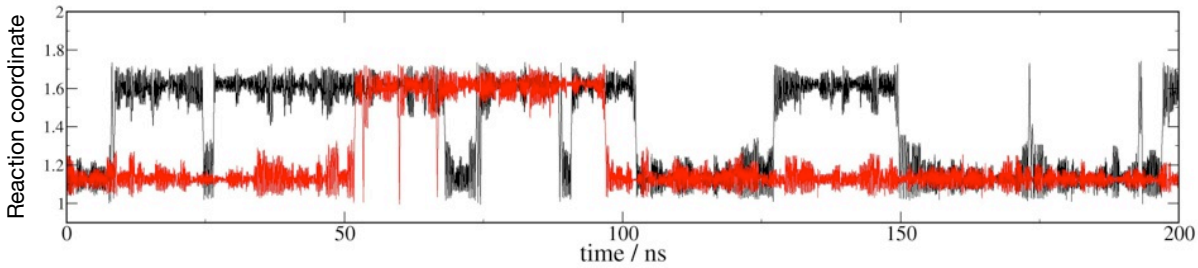
Transition State



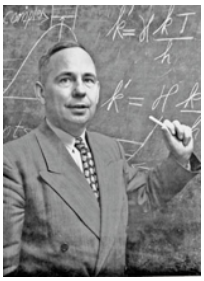
$k = k_0 e^{-\Delta G/kT}$

Arrhenius equation

A long trajectory

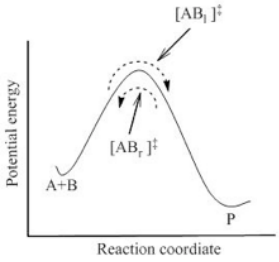


Transition State Theory



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

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

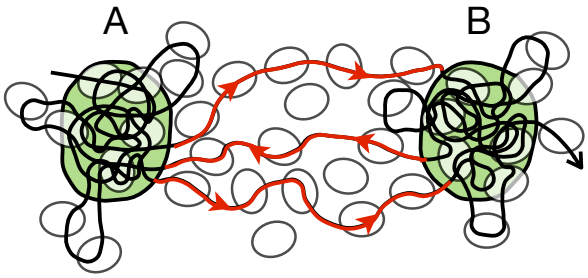
frequency

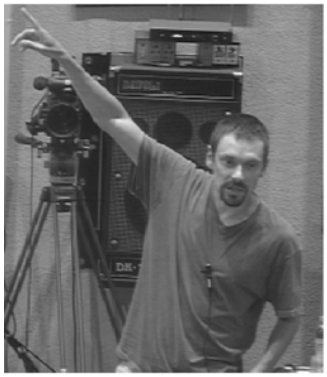
- rate is frequency of unstable mode times the TS probability

- transmission factor is κ=1 in TST (correction for barrier recrossings)

Transition Path Theory

Analysis of reactive events to
- understand reaction mechanisms
- obtain accurate reaction rates





Eric Vanden-Eijnden

Transition paths are Markov chains through a space of intermediate states that depart from A and reach B before A

Some concepts

Lars Onsager: "Splitting probability"

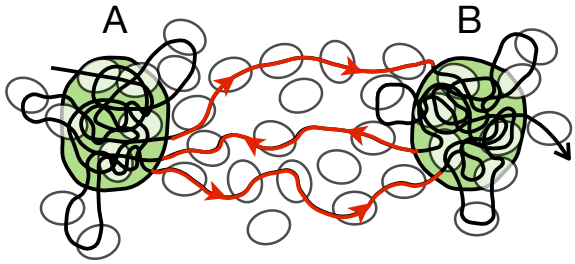
Committor function: q_i is the probability that a trajectory leaving from micro-state i arrives in B before A

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 $0<qi<<1$ and close to B have $0<<qi<1$

$q_i^B = 1 - q_i^A$

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



Friday, 13 December, 13

Some concepts

Lars Onsager: "Splitting probability"

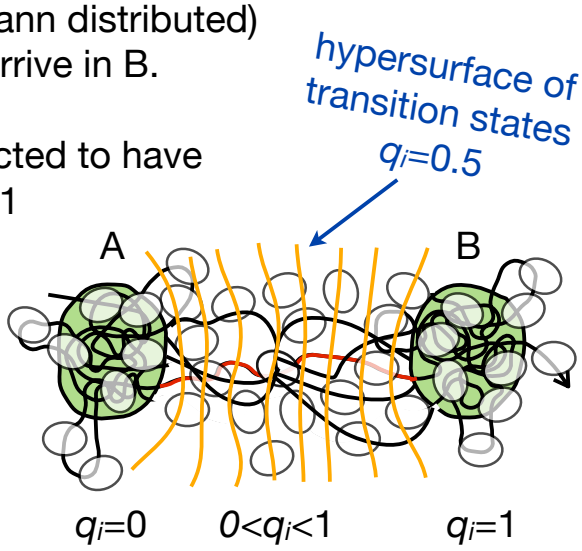
Committor function: q_i is the probability that a trajectory leaving from micro-state i arrives in B before A

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 $0<qi<<1$ and close to B have $0<<qi<1$

$q_i^B = 1 - q_i^A$

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



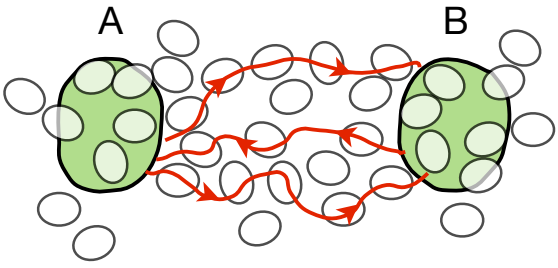
Some concepts

Committor q_i=0.5 connects all micro-states that have a 50/50 chance to go either way. Perfect definition for transition states!

- there is not just 1 transition state at finite temperature
- no (free) energy surface saddle points involved
- accounts for dynamics and recrossings

$$q_i = \sum_{j \in B} t_{ij} + \sum_{j \notin A \cup B} t_{ij} q_j = \sum_{j=1}^{N} t_{ij} q_j$$

the committor is easily related to the transition probabilities to reach the next micro-state



Some concepts

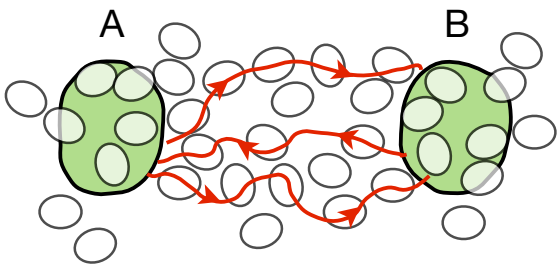
Probability to observe a trajectory in *i* and that it is reactive:

$$\pi_i^R = (1 - q_i)\rho_i q_i$$

peaks in neighborhood of q_i=0.5

Probability that the trajectory visits *i* and *j* consecutively and it is reactive:

$$\pi_{ij}^R = (1 - q_i)\rho_i t_{ij} q_j$$

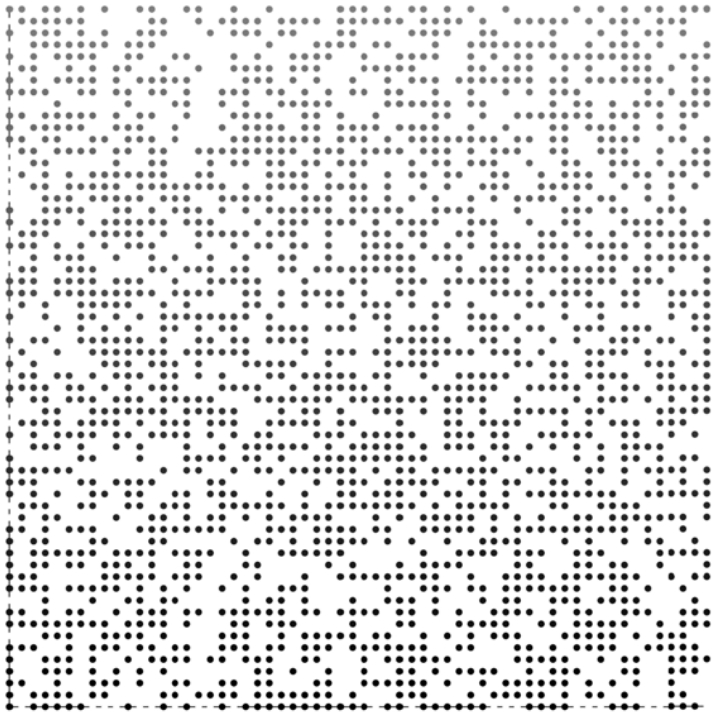


The reactive current through *i* and *j*:

$$F_{ij}^R = \pi_{ij}^R - \pi_{ji}^R = \rho_i t_{ij}(q_j - q_i)$$

Maze example

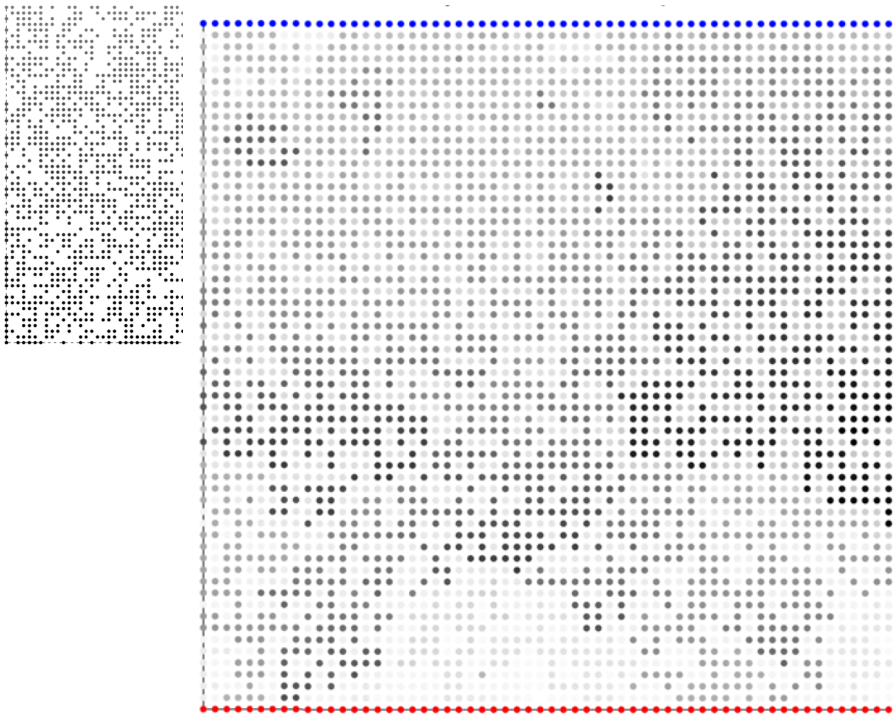
courtesy of Eric Vanden-Eijnden 2013



Maze example

courtesy of Eric Vanden-Eijnden 2013

Energy

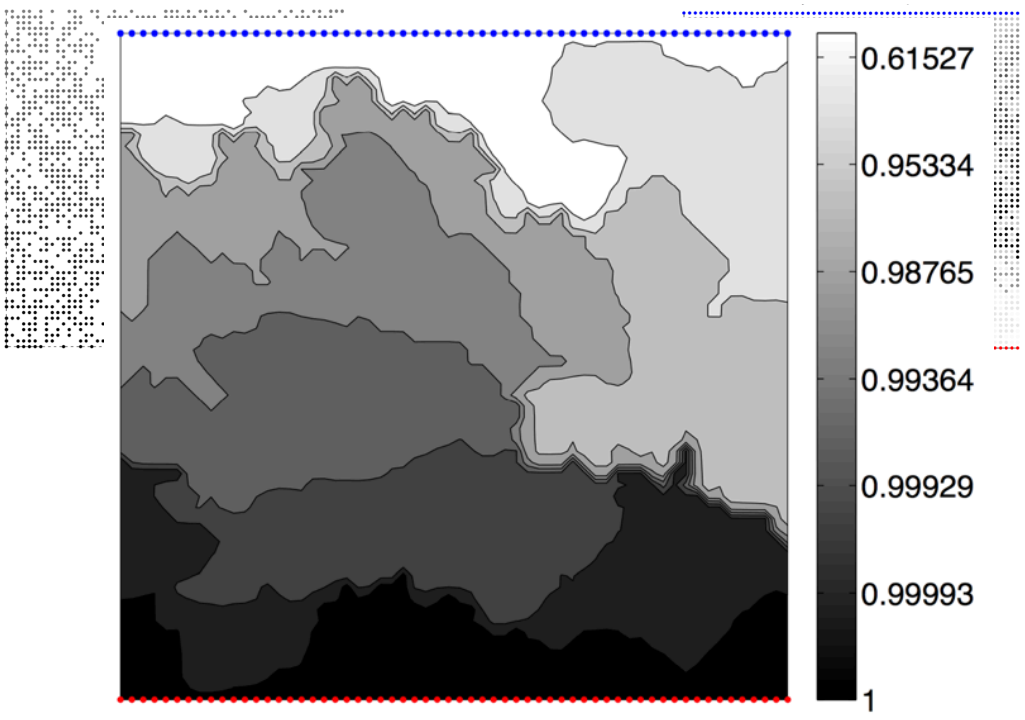


Maze example

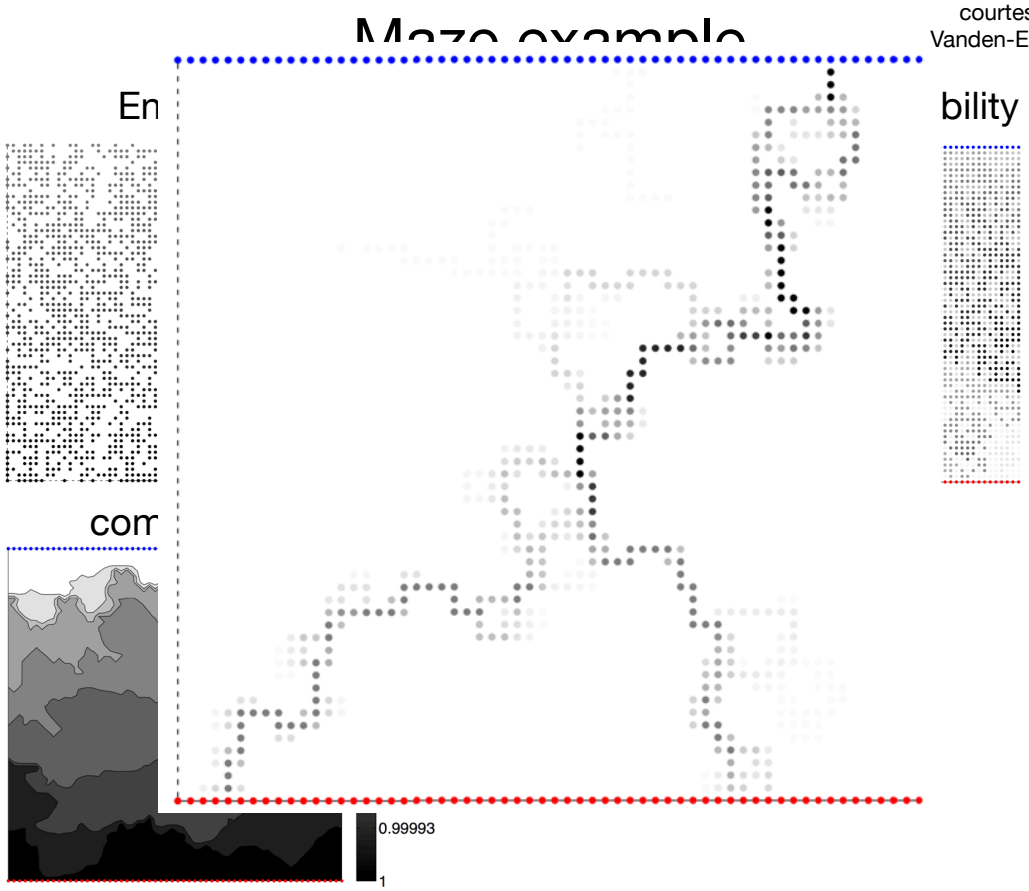
courtesy of Eric Vanden-Eijnden 2013

Energy

Reactive probability



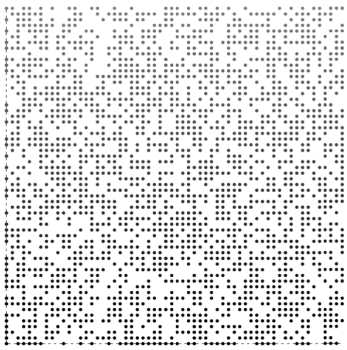
courtesy of Eric Vanden-Eijnden 2013



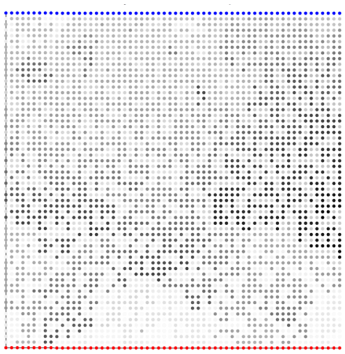
Maze example

courtesy of Eric Vanden-Eijnden 2013

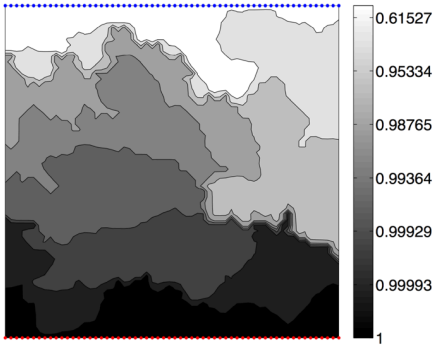
Energy



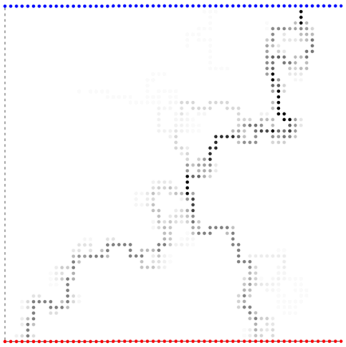
Reactive probability



committor



reactive flux

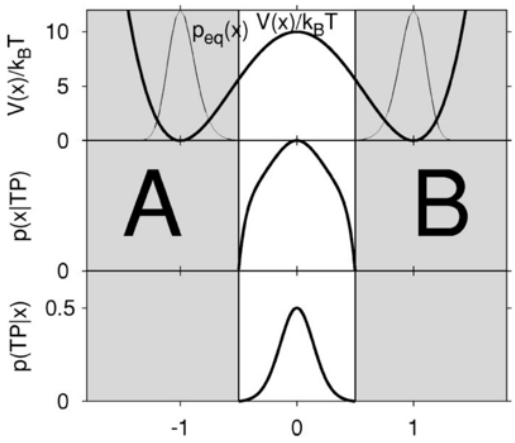


Gerhard Hummer

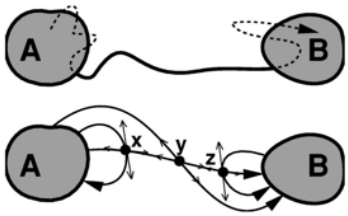
More concepts

Even better definition for a transition state?

A transition state is defined as the state for which the probability that the state is crossed by a reactive trajectory is maximal.



$$p(\mathrm{TP}|x) = 2q_A(x)q_B(x)$$



From transition paths to transition states and rate coefficients
Gerhard Hummer, J. Chem. Phys. 120, 8 (2004)

Summary

part 3

Transition Path Theory is a framework focused on sampling rare events in molecular simulation

- committor function, probability of states to reach product states, defines transition states and reaction coordinate
- reactive current/flux

Content

- Lecture 1: Markov processes
- Lecture 2: State clustering
- Lecture 3: Transition Path Theory
- Lecture 4: Sampling Reactive Events

Sampling Reactive Events

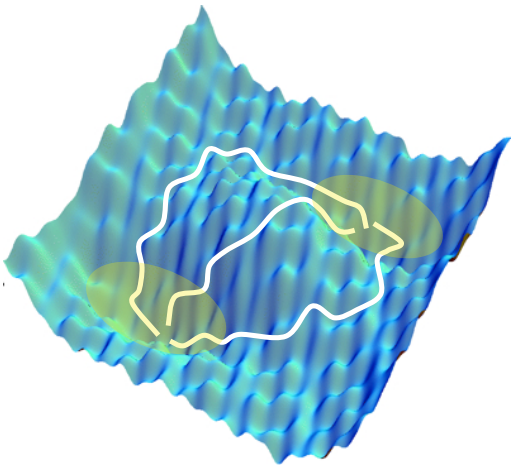
- Generating reactive trajectories
  -Transition Path Sampling, Transition Interface sampling, Forward Flux Sampling, Milestoning,...
- Minimum (free) energy path
  - String method, NEB, dimer method, minimum action
- Free energy methods
  - Umbrella Sampling, Jarzynski (steered) MD, constrained MD, adaptive force bias, metadynamics, path-metadynamics
- Other
  - Parallel tempering, Temperature-accelarated dynamics,...

Transition Path Sampling

Importance sampling of the path ensemble:
all trajectories that lead over barrier and connect stable states.

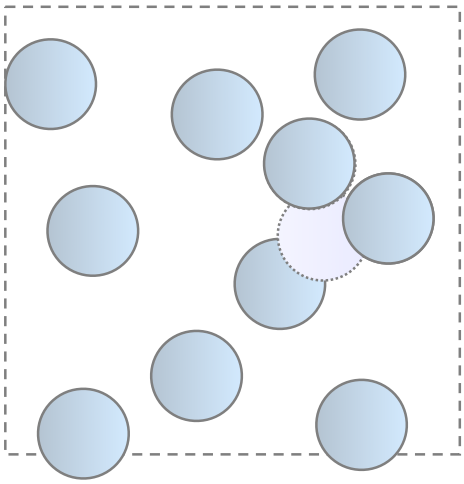
P.G. Bolhuis, D. Chandler, C. Dellago, P.L. Geissler

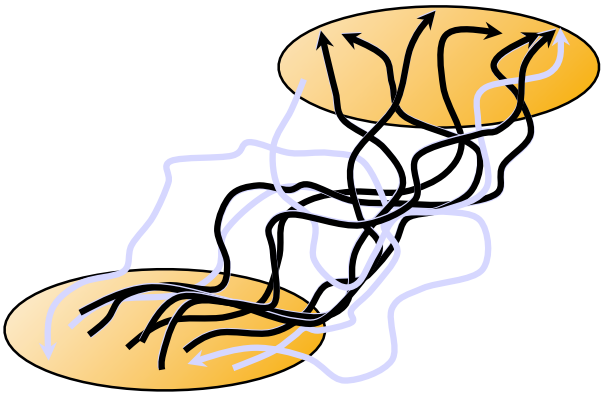
Annu. Rev. Phys. Chem 2002



- No reaction coordinate (only state definition)
- True dynamics
- Works well for 2-state problems
- Needs an initial pathway
- Analysis required
- Transition state ensemble

Transition Path Sampling





Monte Carlo simulation:

trials by random moves

Transition path sampling:

trial paths by shooting move

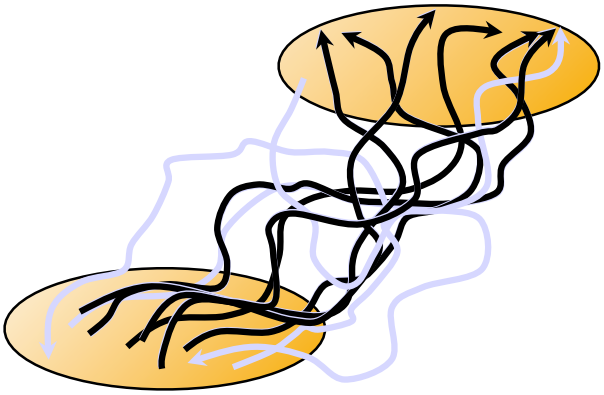
TPS is Monte Carlo sampling of dynamical trajectories

Transition Path Sampling

Algorithm

1. generate an initial reactive path
2. choose a random configuration
3. make a small perturbation
4. integrate backward and forward
5. accept new path if it connects A and B, otherwise reject and go back to previous path
6. go back to 2.

In case of flexible path lengths, adjust acceptance criterion to maintain detailed balance.



Transition path sampling:

trial paths by shooting move

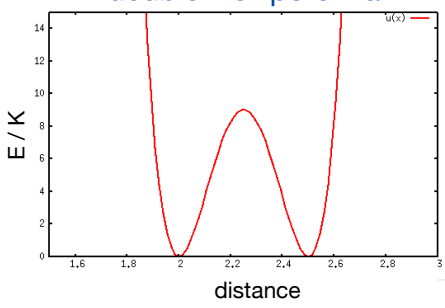
TPS is Monte Carlo sampling of dynamical trajectories

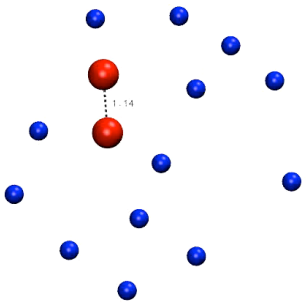
Example: exercise from Molsim2014 tutorial in Amsterdam

Transition Path Sampling

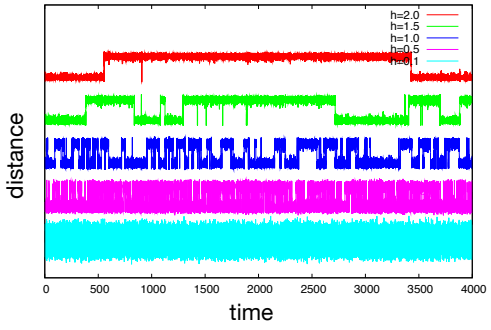
- 15 WCA particles in 2D box
- particles 1 and 2 are connected by double-well potential
- initial path from heigh-temperature MD simulation

double well potential



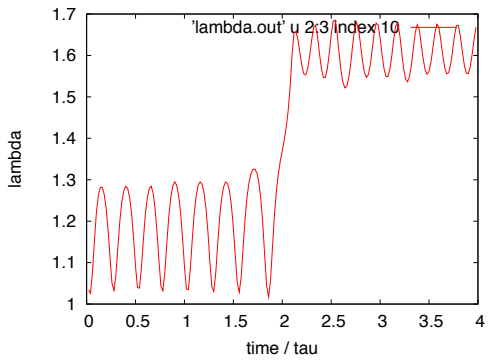


dynamics for different barrier heights

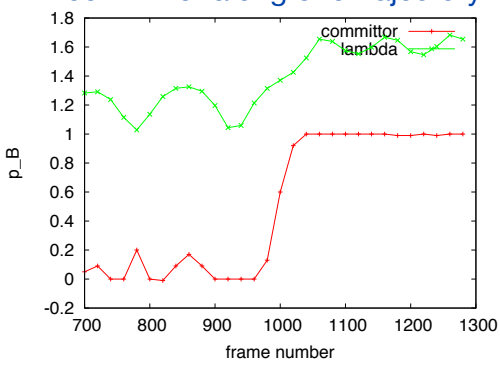


Transition Path Sampling

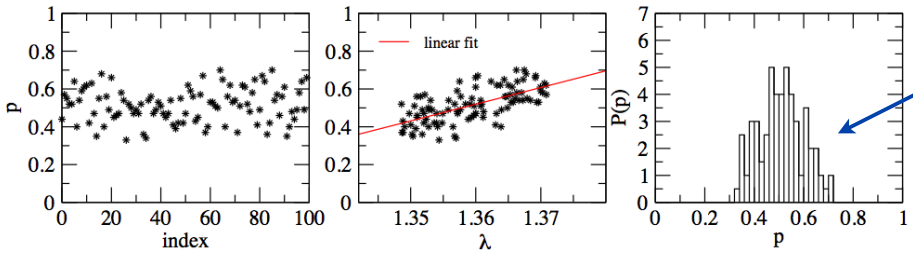
distance vs time



committor along one trajectory

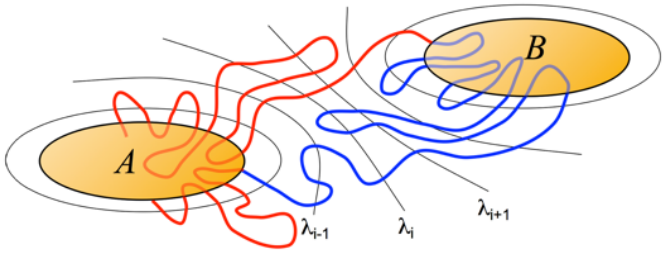


Reaction coordinate test: committor distribution



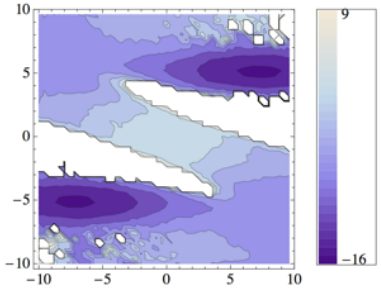
Peaked committor probability means that the reaction coordinate is good

Transition Interface Sampling

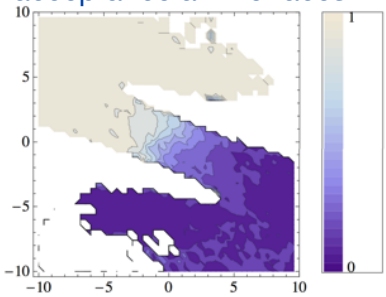


TPS between intermediate interfaces to sample “long” diffusive processes (protein folding) – no strong double-well behavior.

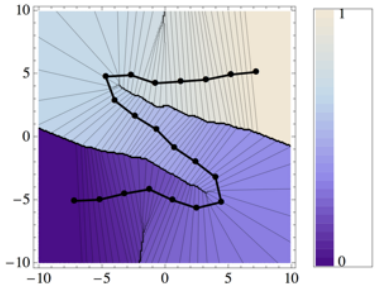
Difficult test: the Z-potential



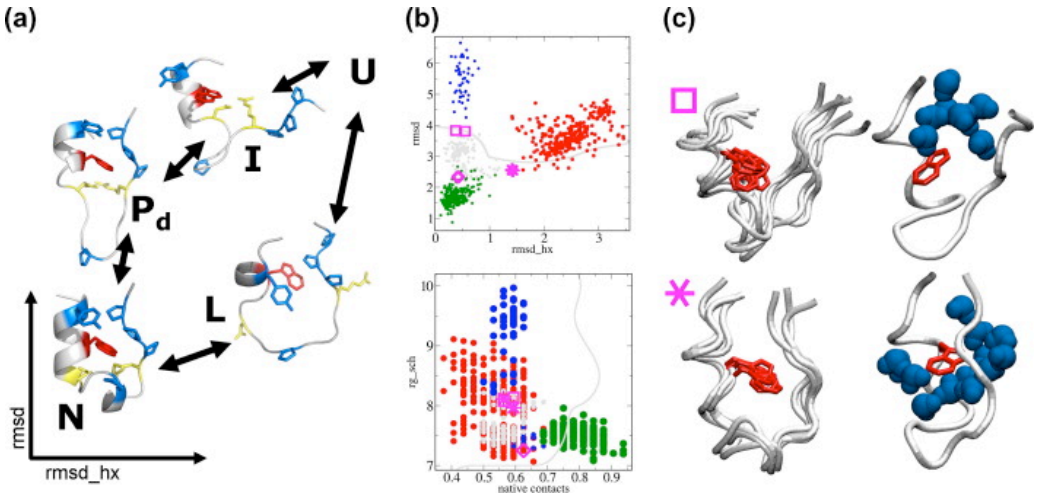
committor from acceptance at interfaces



most likely pathway



Protein folding with TPS/TIS



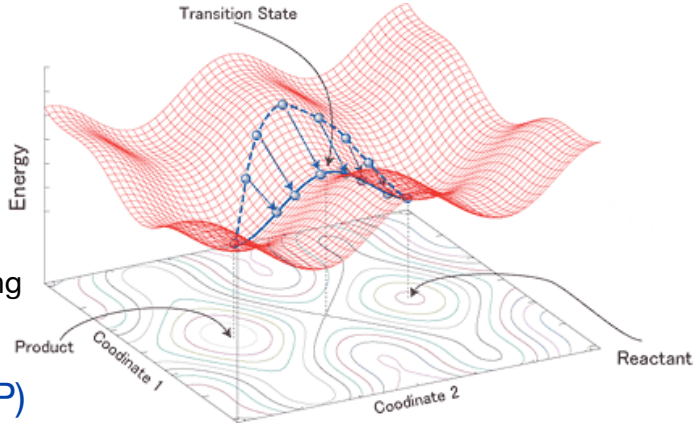
(a) Schematic depiction of the folding mechanism of Trp-cage, elucidated with TPS simulations. (b) Stable states and transition states obtained by committor calculation, plotted as rgsch versus ρ and rmsd versus rmsdhx. Transition states for the NL and NPd transitions are plotted in magenta as stars and squares respectively. The scatter points are taken from the corresponding TPS trajectories. Native states (N) are plotted in green, loop structures (L) in red, close to native structures with Pro-12 detached (Pd) in gray and the I-state in blue. The region on the bottom of the gray line on the rmsd/rmsdhx (to the right on the ρ/rgsch) plot is the part of the configuration space not reachable by our REMD simulation starting from an unfolded configuration. (c) Left: superimposed TS structures for four different NL (star) and six for PdI paths (square). Right: one of the TS structures for both routes in a side view plotted as a ribbon. Central tryptophan residue is plotted in licorice representation (red), water molecules within 5 Å of Trp-6 in space-filling representation in blue. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Transition path sampling of protein conformational changes
J. Juraszek, J. Vreede and P.G. Bolhuis
Chem. Phys. **396** (2012), 30-44

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

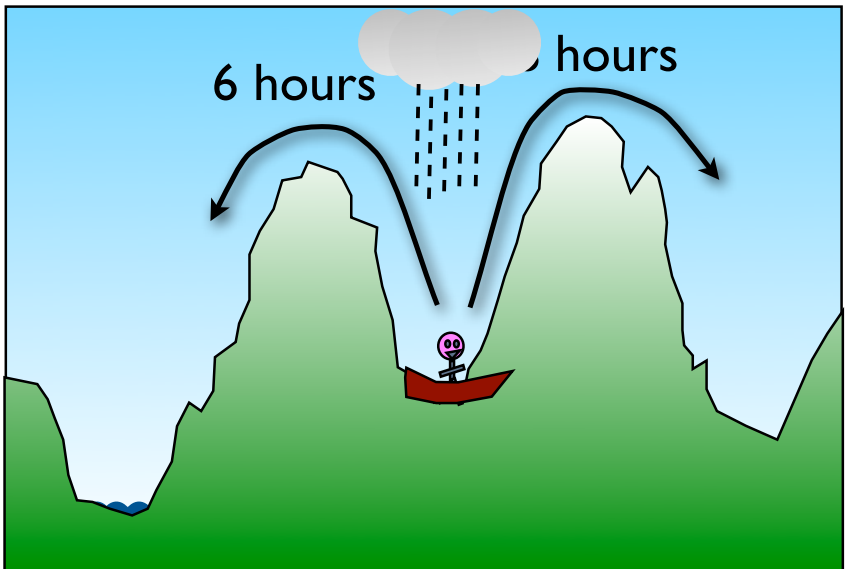
Escaping free-energy minima,
Laio and Parrinello, PNAS (2002)

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

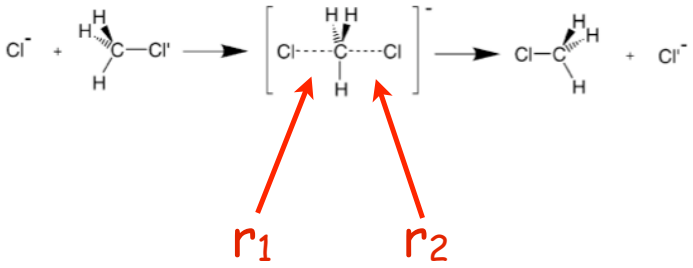
- enhance sampling
- probe free energy landscape
- multiple collective variables

Metadynamics

Metadynamics, Laio and Parrinello, PNAS (2002)

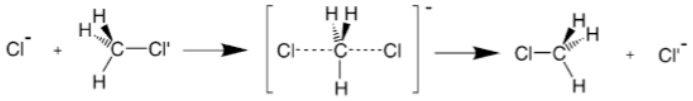


simple example: S_N2 reaction

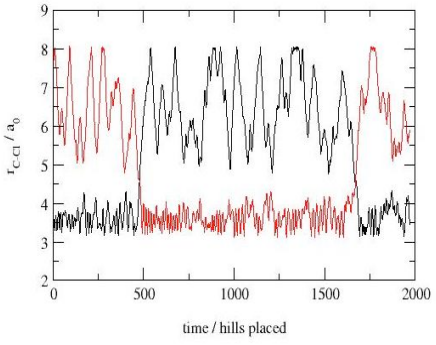


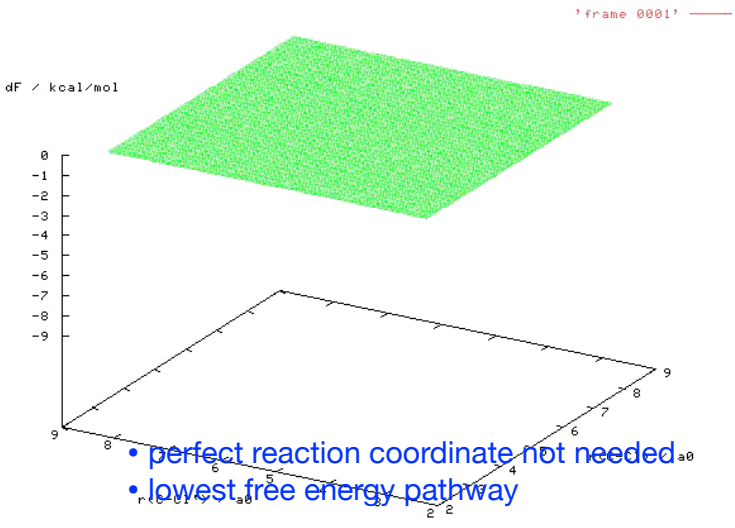
Two collective variables: r_1 and r_2

simple example: S_N2 reaction



Two collective variables: r_1 and r_2

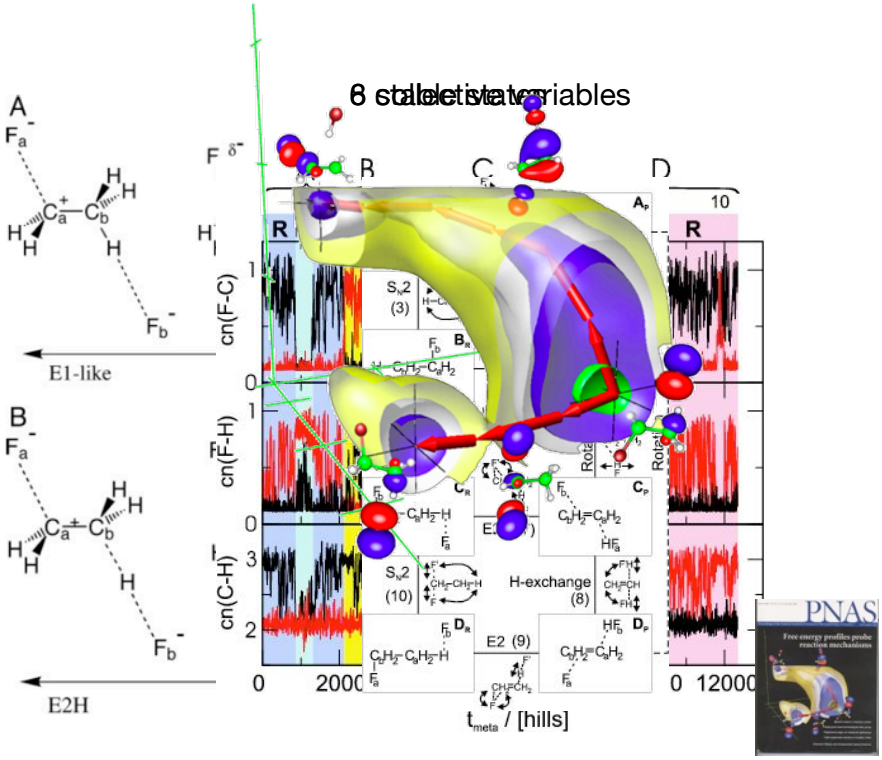




Bernd Ensing, Alessandro Laio, Michele Parrinello and Michael L. Klein
J. Phys. Chem. B **109** (2005), 6676-6687

- perfect reaction coordinate not needed
- lowest free energy pathway

E2 vs S_N2 reaction

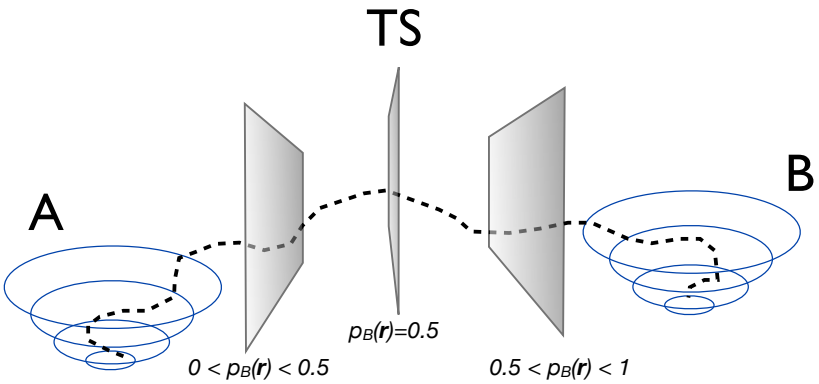


Bernd Ensing and Michael L. Klein
Proc. Natl. Proc. Acad., USA **102** (2005), 6755-6759

More collective variables?

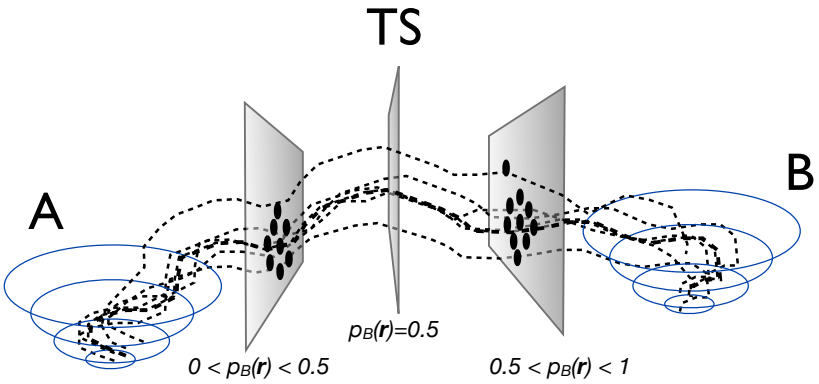
transition path sampling and path-metadynamics

iso-committor surfaces



transition path sampling and path-metadynamics

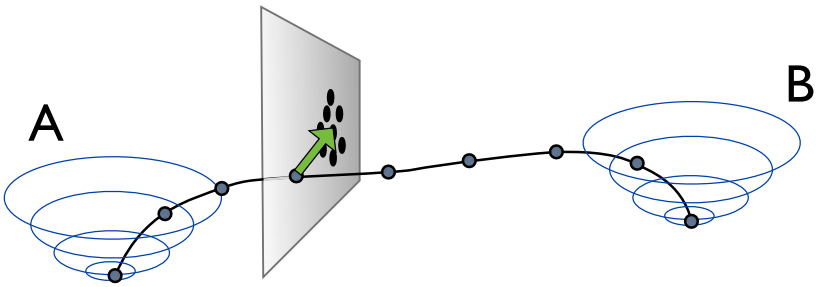
iso-committor surfaces



the average transition pathway
(in CV space)

path-metadynamics

distance to mean density



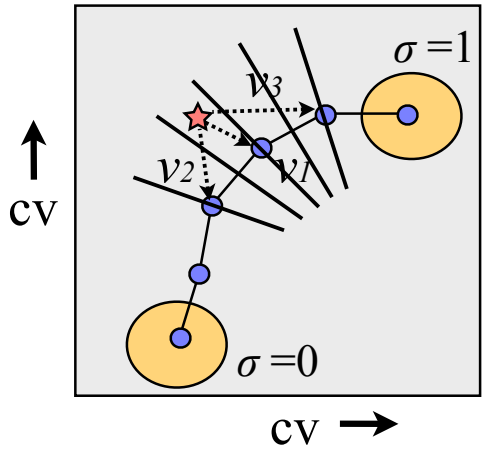
- start from guess path
- bias dynamics along path
- move nodes to the mean density
- maintain equidistant nodes

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

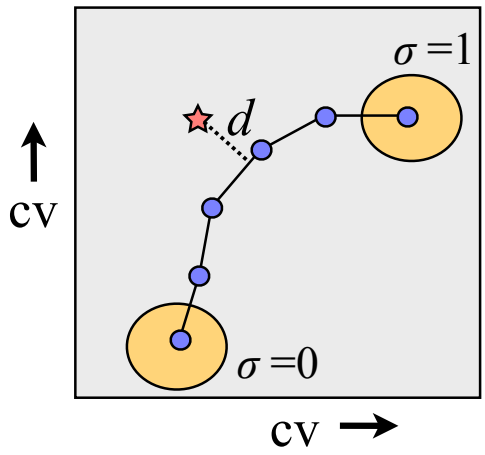
MFEP as the collective variable



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

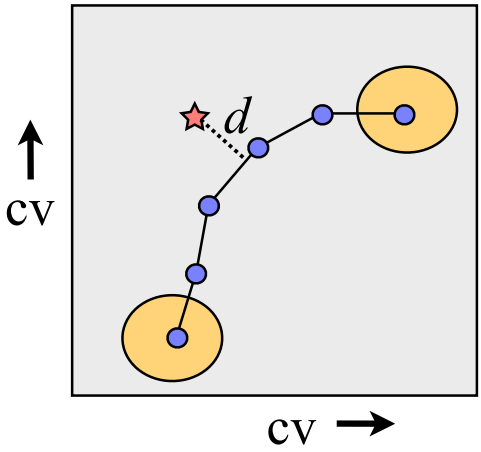
or accumulate average *d for each node*

MFEP as the collective variable

$$< d_i >= \frac{\sum_0^\tau w_i * d_i}{\sum_0^\tau w_i}$$

$$w_i(t) = (\sigma(n-1) - i_{min}) f^{\tau - t}$$

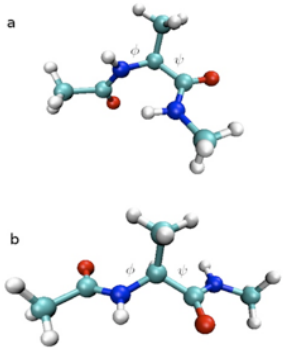
$$f = \exp[\tau^{-1} * \ln \frac{1}{2}]$$



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

alanine dipeptide

• Zhiwei Liu
• Preston Moore
University of the Sciences, Philadelphia



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

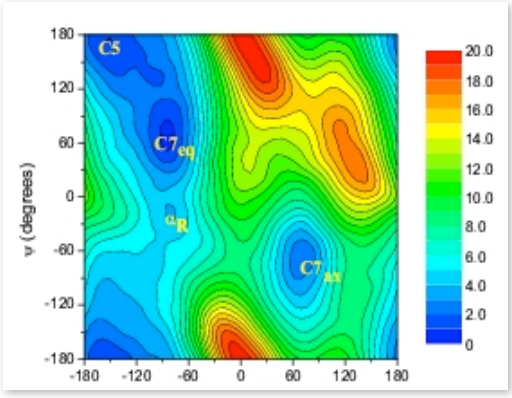
Alanine dipeptide + 216 water
CHARMM27 forcefield (modified internal params water)
Cubic periodic box L=18.8 Angstrom
NVT ensemble, T=298K

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

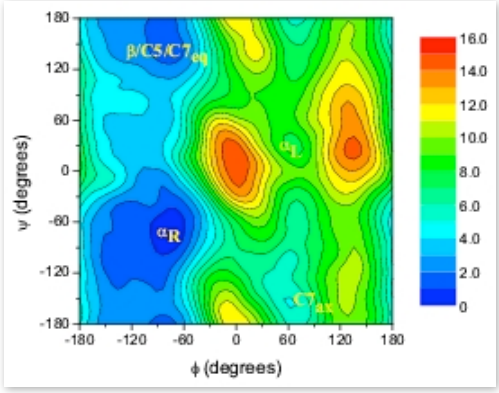
Total simulation time: ca. 5 ns

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

Gas-phase



Aqueous solution

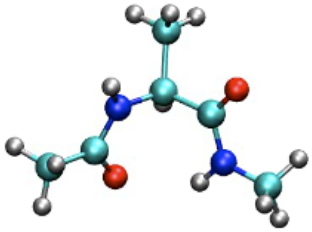


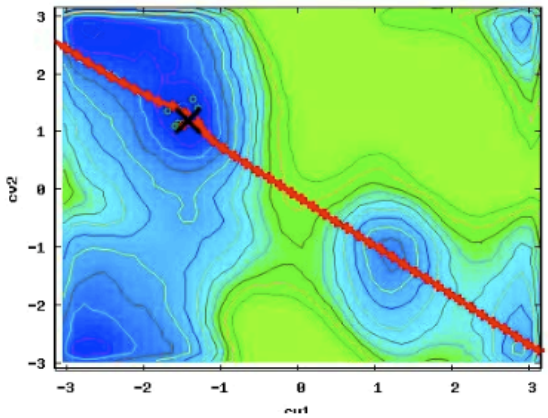
Metadynamics as a tool for exploring the free energy landscape of chemical reactions.
Bernd Ensing, Marco De Vivo, Zhiwei Liu, Preston Moore, and Michael L. Klein
Acc. Chem. Res. **39** (2006), 73-81

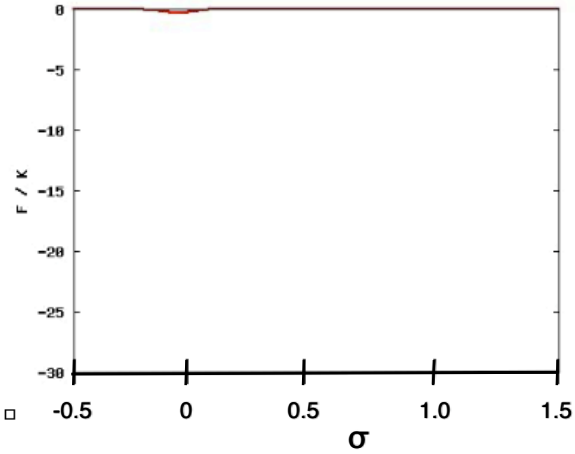
alanine dipeptide

Parameters:
T = 300 K
$H_{gaussian}$ = 10 K
$W_{gaussian}$ = 0.05
$\Delta t_{gaussian}$ = 100 MD steps
n = 20 + 20 + 20
τ = 1000 MD steps

every recrossing:
H x 50%
W x 50%
τ x 100





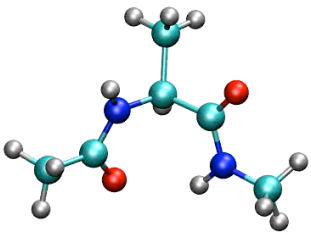


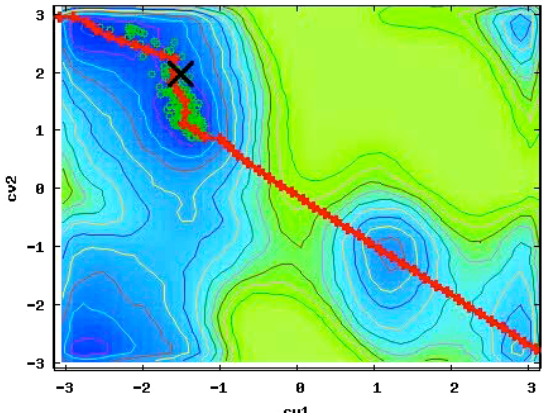
Path Finding on High-Dimensional Free Energy Landscapes.
Grisell Díaz Leines and Bernd Ensing
Phys. Rev. Lett. **109** (2012), 020601

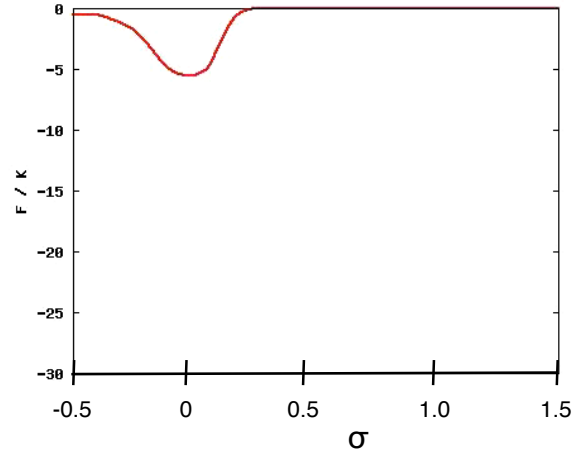
alanine dipeptide

Parameters:
T = 300 K
$H_{gaussian}$ = 10 K
$W_{gaussian}$ = 0.05
$\Delta t_{gaussian}$ = 100 MD steps
n = 20 + 20 + 20
τ = 1000 MD steps

every recrossing:
H x 50%
W x 50%
τ x 100





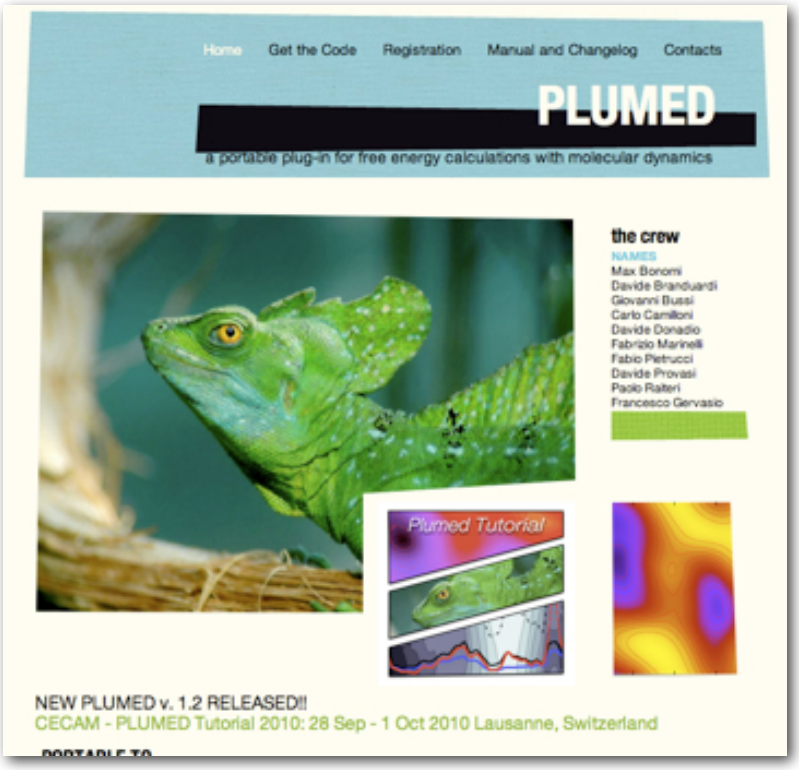


Path Finding on High-Dimensional Free Energy Landscapes.
Grisell Díaz Leines and Bernd Ensing
Phys. Rev. Lett. **109** (2012), 020601

Friday, 13 December, 13

Grisell Diaz Leines

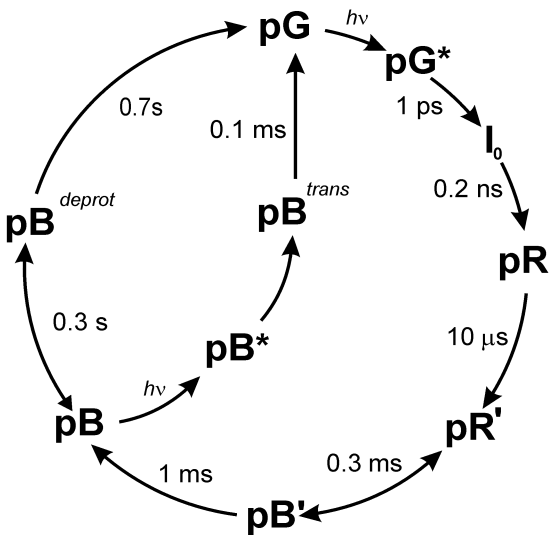
path-metadynamics

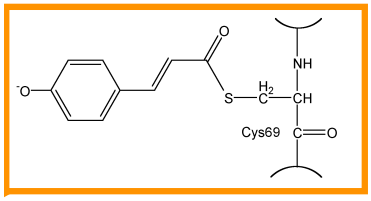


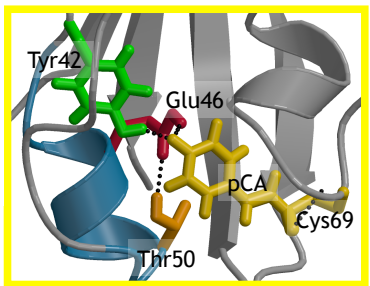
Davide Branduardi
MPI-Biophys, Frankfurt

Photoactive Yellow Protein

blue-light sensor from H. halophila

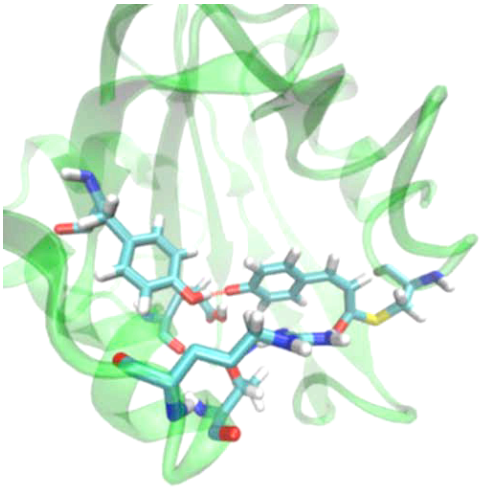




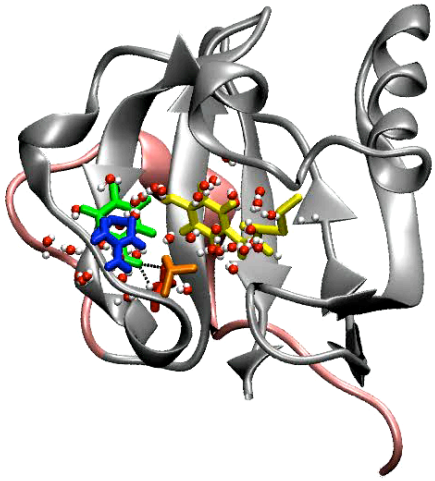


Photoactive Yellow Protein

Proton transfer step



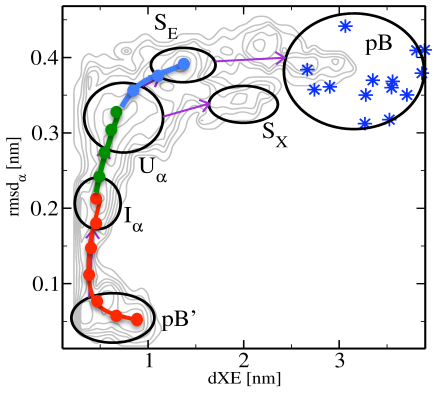
Conformational transformation

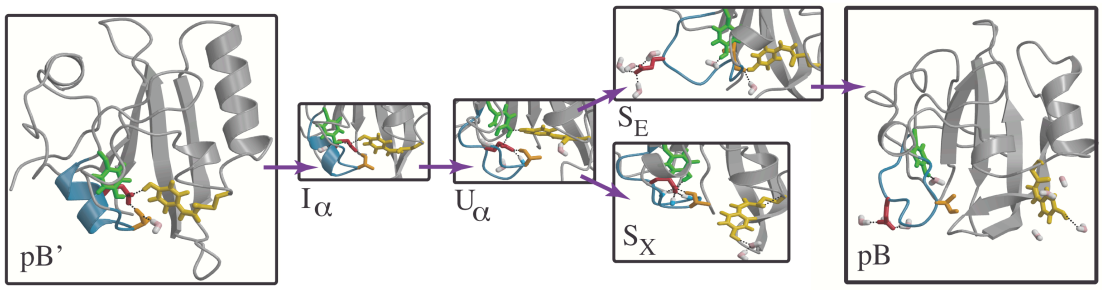


Predicting the signaling state of Photoactive Yellow Protein.
J. Vreede, W. Crielaard, K.J. Hellingwerf, and P.G. Bolhuis
Biophys J. **88** (2005), 3525 DOI: 10.1529/biophysj.104.055103

transition path sampling

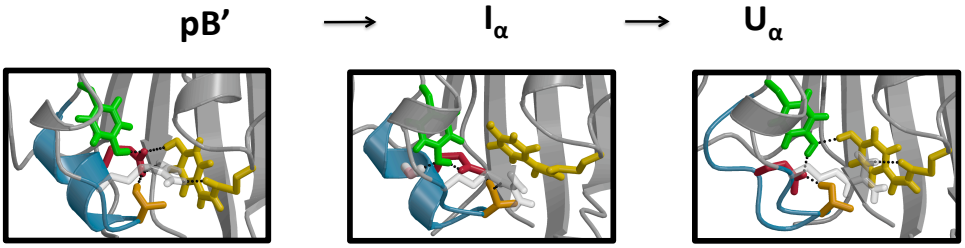
REMD
TPS





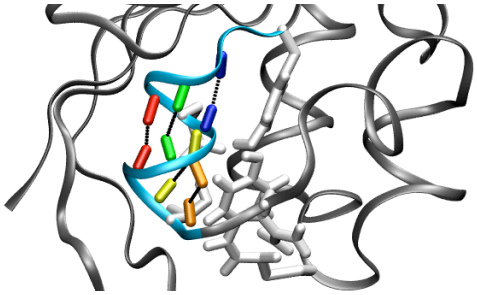
path-metadynamics study of PYP

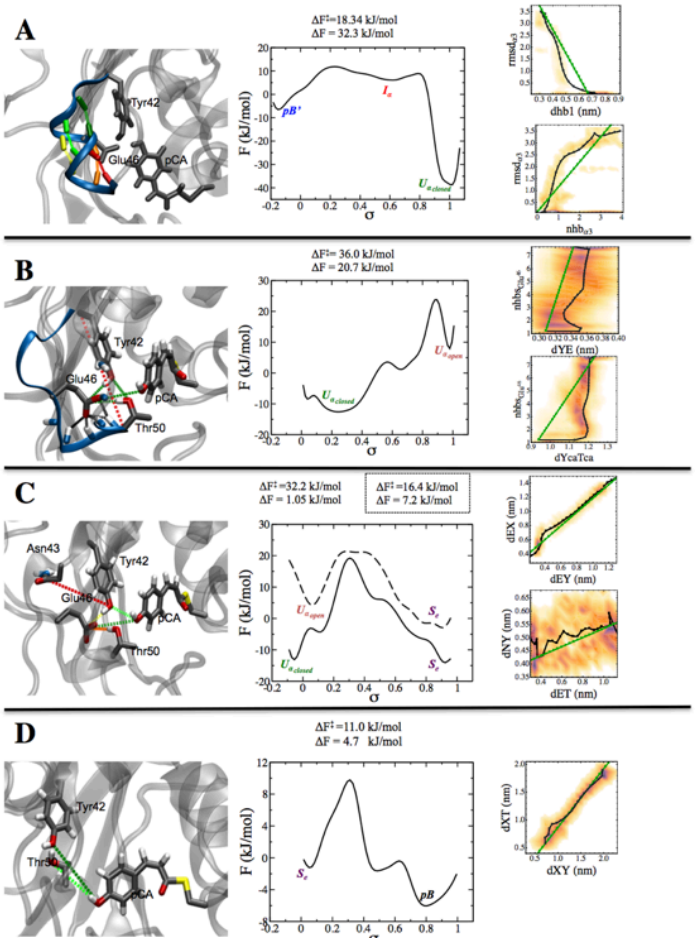
unfolding of helix 43-51



Collective variables:

- Asn43-O---Gly47-H = $dhb_{\alpha 3}{}^{1}$
- Ala44-O----Asp48-H=$dhb_{\alpha 3}{}^{2}$
- Ala45-O----Ile49-H= $dhb_{\alpha 3}{}^{3}$
- Glu46-O----Thr50-H=$dhb_{\alpha 3}{}^{4}$
- Gly47-O----Gly51-H=$dhb_{\alpha 3}{}^{5}$





Content

- Lecture 1: Markov processes
- Lecture 2: State clustering
- Lecture 3: Transition Path Theory
- Lecture 4: Sampling Reactive Events

Summary

part 1

A stochastic hopping between states is Markovian if the probability to reach the next state does not depend on the past (no memory)

- reversible Markov chains (detailed balance)
- ergodicity
- spectral decomposition (convergence to equilibrium)

part 2

Markov State Modeling applied as analysis tool of the dynamics of a small peptide

- Analysis of molecular dynamics simulation
- Clustering based on structures
- Relation to free energy landscape based on cluster size
- Transition probability between clusters (states)
- Using the MSM the dynamics can be sampled in a coarse-grain manner, e.g. using kinetic Monte-Carlo

Summary

part 3

Transition Path Theory is a framework focused on sampling rare events in molecular simulation

- committor function, probability of states to reach product states, defines transition states and reaction coordinate
- reactive current/flux

part 4

Advanced sampling methods for rare events focus on sampling reactive trajectories, the free energy landscape, or an average transition pathway

- Transition path sampling (TPS/TIS/FFS)
- String method
- Metadynamics, path-metadynamics

Further Reading

- **The life and work of A.A. Markov,** Basharin, G.P., Langville, A.N., Naumov, V.A, Linear Algebra and its Applications 386 (2004), 3-26.
- **The activated complex and the absolute rate of chemical reactions.** Eyring H. Chem. Rev. 17 (1935), 65–77
- **The transition state method.** Wigner E. Trans. Faraday Soc. 34 (1938), 29–41
- **Understanding Molecular Simulation: From Algorithms to Applications.** Frenkel, D.; Smit, B., Academic, San Diego, CA, 2 edition, 2002**.**
- **Transition-Path Theory and Path-Finding Algorithms for the Study of Rare Events** Weinan E and Eric Vanden-Eijnden Annu. Rev. Phys. Chem. 61 (2010), 391–420
- **String method for the study of rare events.** E W, Ren W, Vanden-Eijnden E. Phys.Rev.B 66 (2002), 052301
- **Finite temperature string method for the study of rare events.** E W, Ren W, Vanden-Eijnden E. J. Phys. Chem. B 109 (2005), 6688–93
- **Reactive flux and folding pathways in network models of coarse-grained protein dynamics.** Alexander Berezhkovskii, Gerhard Hummer, and Attila Szabo J. Chem. Phys. 130 (2009), 205102
- **A growing string method for determining transition states: Comparison to the nudged elastic band and string methods.** Peters, B.; Heyden, A.; Bell A. T.; Chakraborty A. J. Chem. Phys. 2004, 120, 7877-7886.
- **Transition path sampling and the calculation of rate constants.** Dellago, C.; Bolhuis, P. G.; Csajka, F. S.; Chandler, D.. J. Chem. Phys. 108 (1998), 1964-1977.
- **Non-linear reaction coordinate analysis in the reweighted path ensemble.** Wolfgang Lechner, Jutta Rogal, Jarek Juraszek, Bernd Ensing, and Peter Bolhuis J. Chem. Phys. 133 (2010), 174110
- **Non-linear reaction coordinate analysis in the reweighted path ensemble.** Wolfgang Lechner, Jutta Rogal, Jarek Juraszek, Bernd Ensing, and Peter Bolhuis J. Chem. Phys. 133 (2010), 174110
- **From transition paths to transition states and rate coefficients** Gerhard Hummer, J. Chem. Phys. 120 (2004), 8
- **The reweighted path ensemble.** Jutta Rogal, Wolfgang Lechner, Jarek Juraszek, Bernd Ensing, and Peter Bolhuis J. Chem. Phys. 133 (2010), 174109
- **Nudged elastic band method for finding minimum energy paths of transitions.** In Classical and Quantum Dynamics in Condensed Phase Simulations, Jonsson H, Mills G, Jacobsen K W.1998. ed. B J Berne, G Ciccotti, D F Coker, pp. 385–404. Singapore: World Sci.
- **A dimer method for finding saddle points on high dimensional potential surfaces using only first derivatives.** Henkelman G, Jonsson H. J. Chem. Phys. 111 (1999), 7010–22
- **Conjugate peak refinement: an algorithm for finding reaction paths and accurate transition states in systems with many degrees of freedom.** Fischer S, Karplus M. Chem. Phys. Lett. 194 (1992), 252–61
- **A milestoning study of the kinetics of anallosteric transition: atomically detailed simulations of deoxy Scapharca hemoglobin.** Elber R. Biophys. J. 92 (2007), L85–87
- **Forward flux sampling-type schemes for simulating rare events: efficiency analysis.** Allen R J, Frenkel D, ten Wolde P R. J. Chem. Phys. 124 (2006.), 194111
- **Finding transition pathways using the string method with swarms of trajectories.** Pan AC, Sezer D, Roux B. J. Phys. Chem. B 112 (2008), 3432–40
- **Escaping free-energy minima.** Laio, A.; Parrinello, M. Proc. Natl. Acad. Sci. USA, 99 (2002), 12562-12566.
- **A recipe for the computation of the free energy barrier and the lowest free energy path of concerted reactions.** Ensing, B.; Laio, A.; Parrinello, M.; Klein, M. L. J. Phys. Chem. B 109 (2005), 6676-6687
- **Path Finding on High-Dimensional Free Energy Landscapes.** Grisell Díaz Leines and Bernd Ensing Phys. Rev. Lett. 109 (2012), 020601