

PLUMED Masterclass 22-9

6 June 2022

Using path collective variables to find reaction mechanisms in complex free energy landscapes

Bernd Ensing

Computational Chemistry Group
AI4Science Laboratory
University of Amsterdam, Amsterdam, The Netherlands

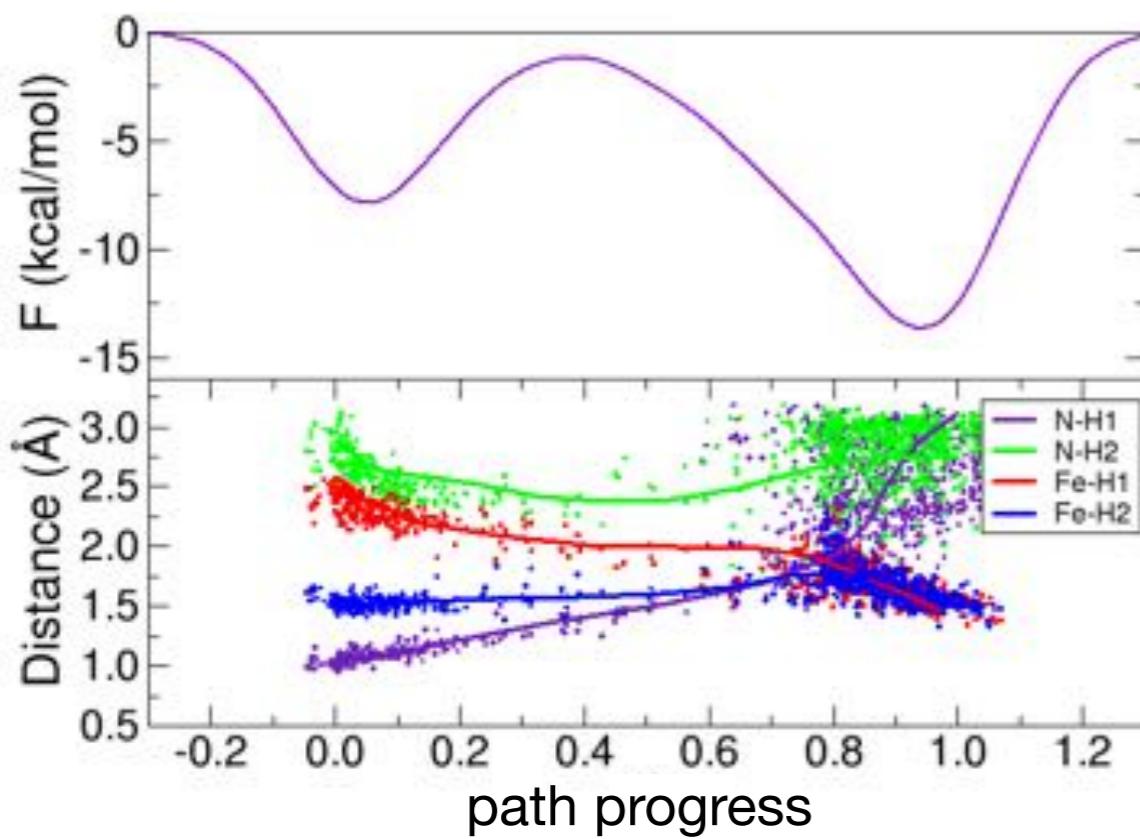
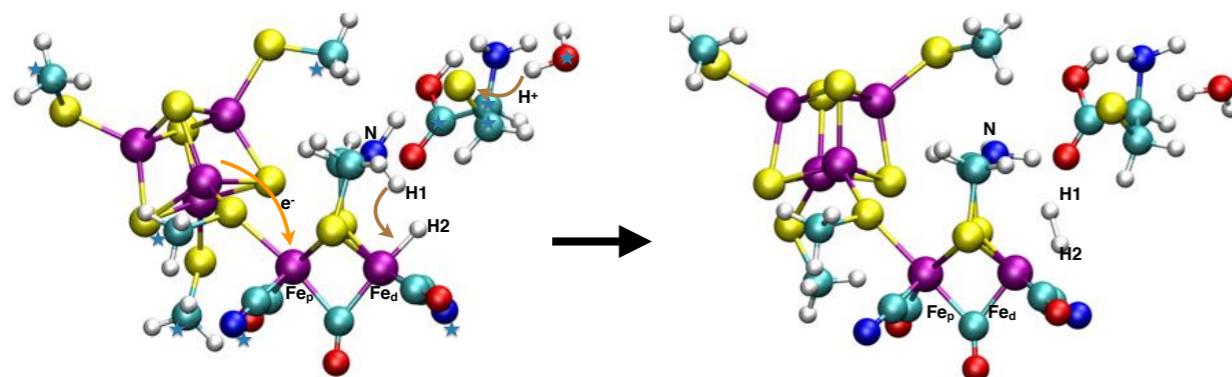


Introduction

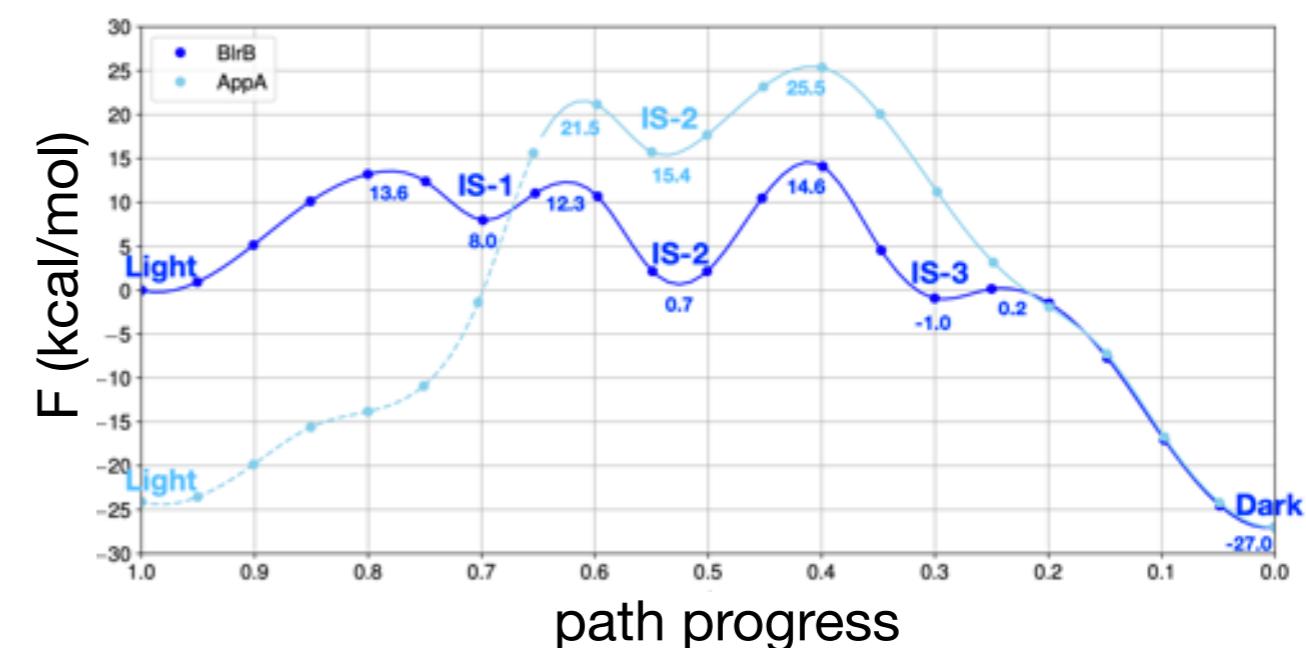
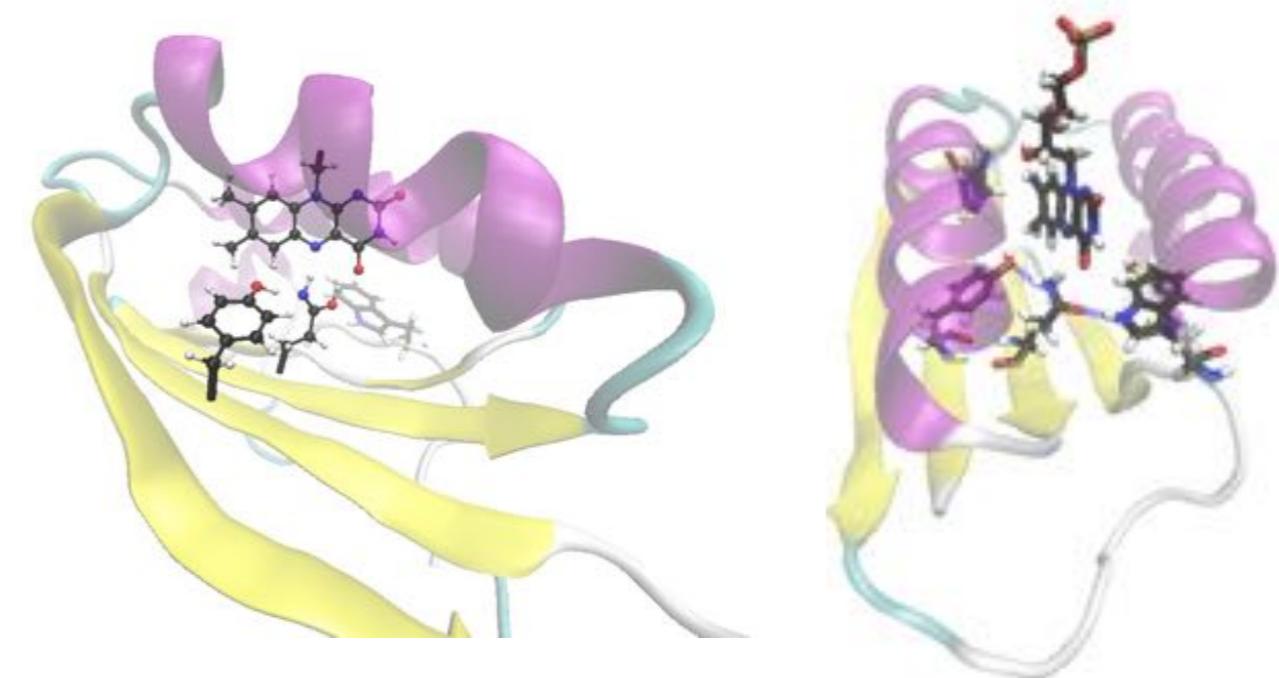
Introduction to
PLUMED Masterclass 22-9
in 5 slides

Why I use Path-CVs

Hydrogen evolution in FeFe H₂-ase enzyme



Dark state recovery in photoactive protein BLUF

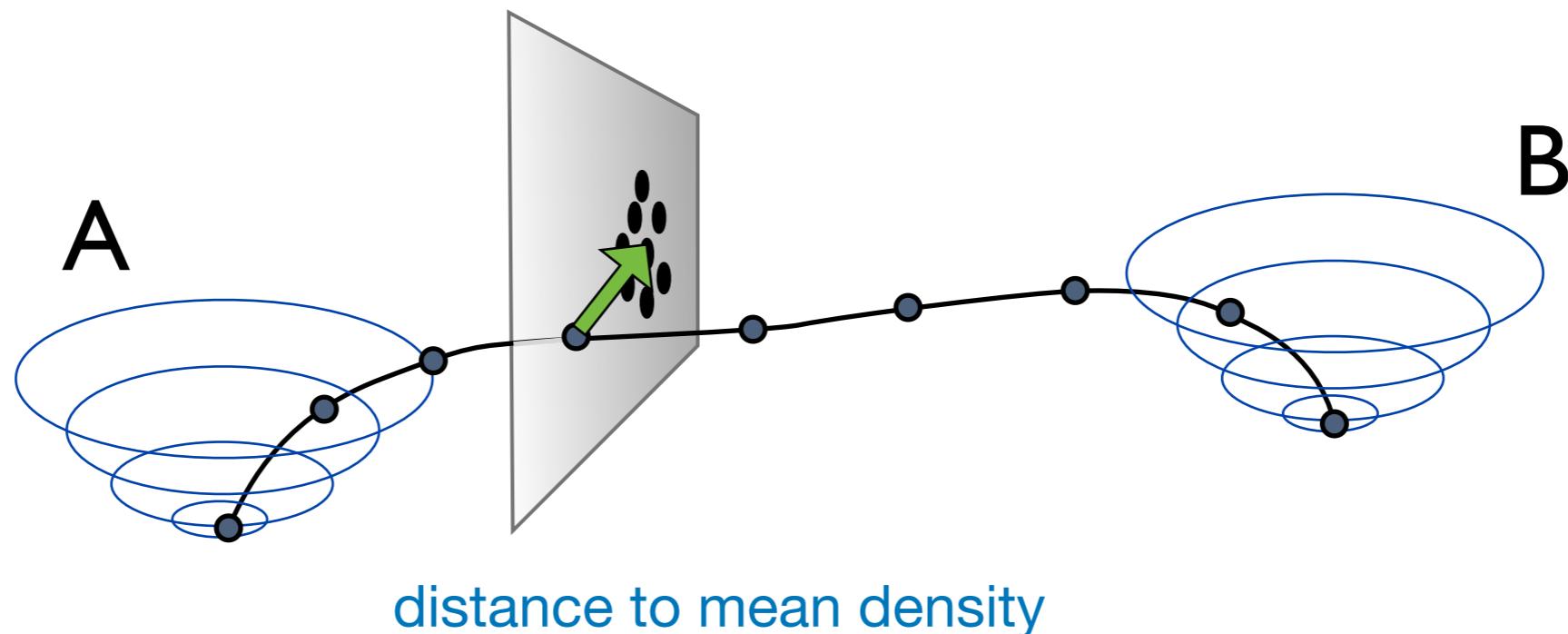


Path-CV in a nutshell

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

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

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



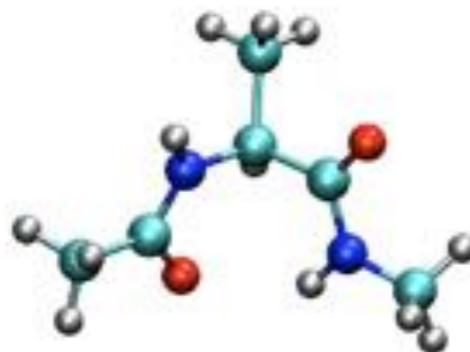
From A to B in free energy space.
Davide Bhandaru, Francesco Luigi Gervasio, and Michele Parrinello.
J. Chem. Phys. **126** (2007), 054103

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

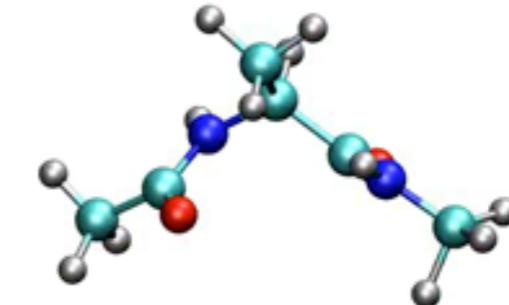
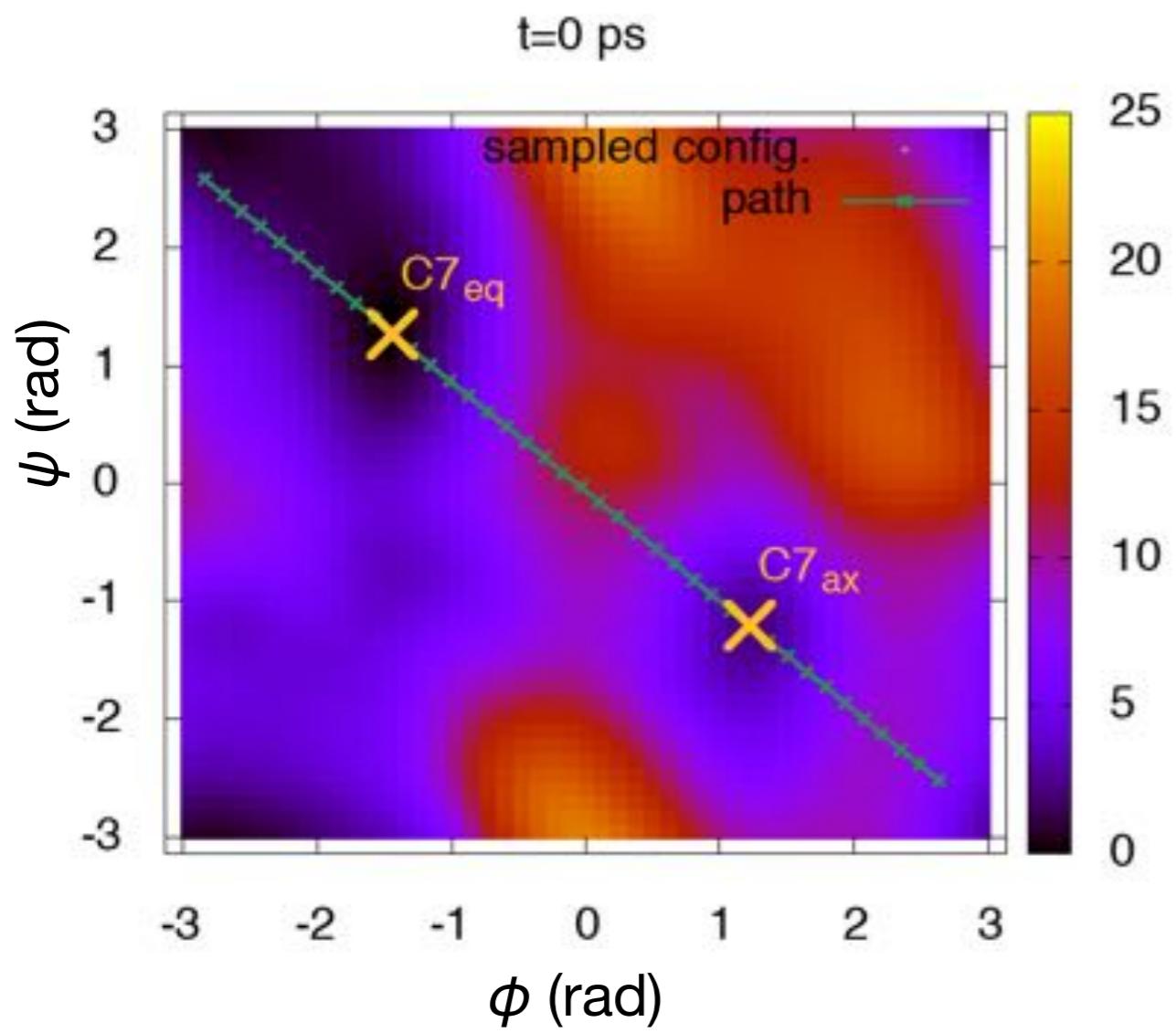
Path finding on high-dimensional free energy landscapes.
Grisell Díaz Leines and Bernd Ensing
Phys. Rev. Lett. **109** (2012), 020601

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

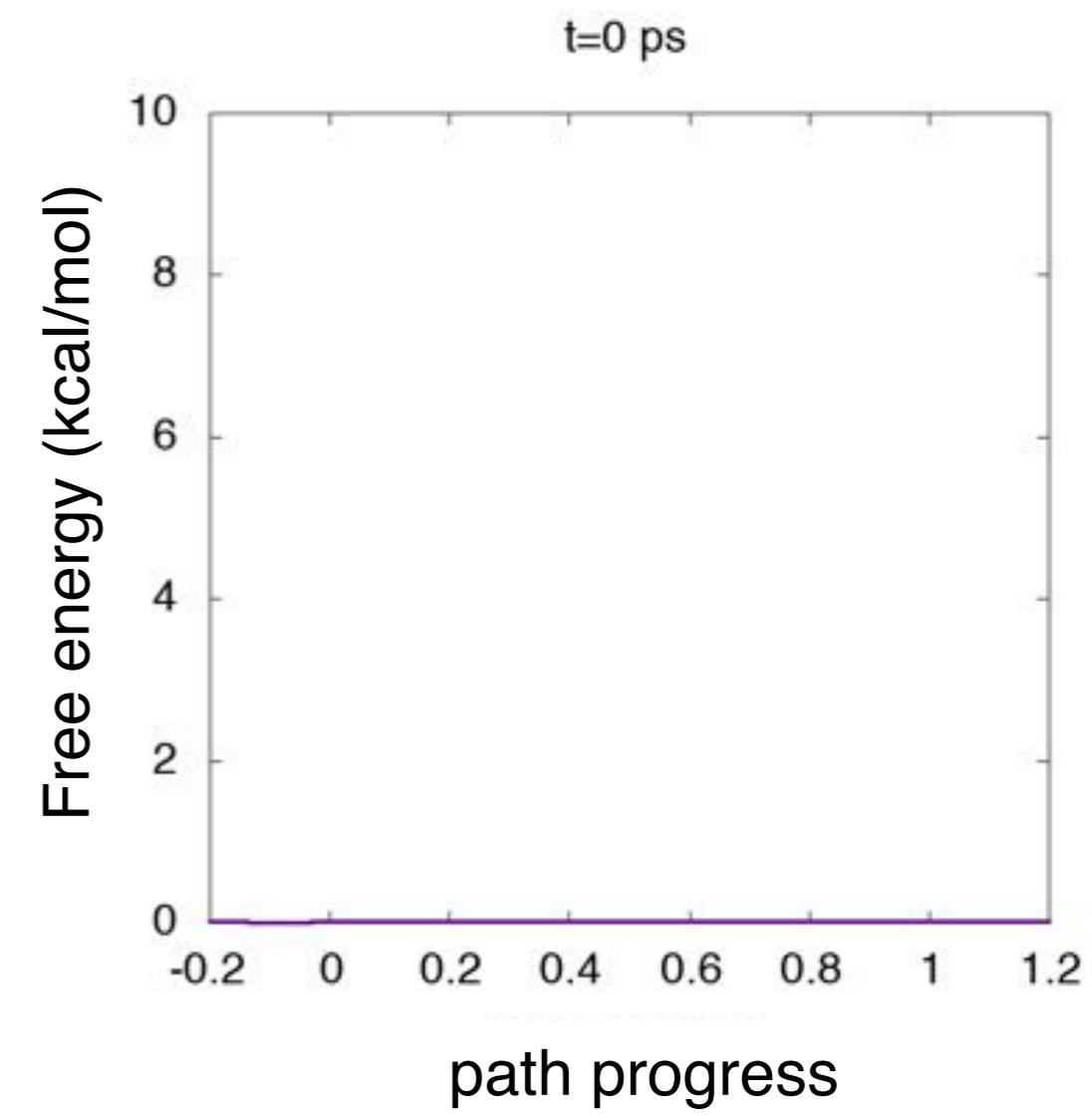
Alanine dipeptide



path evolution



free energy estimation

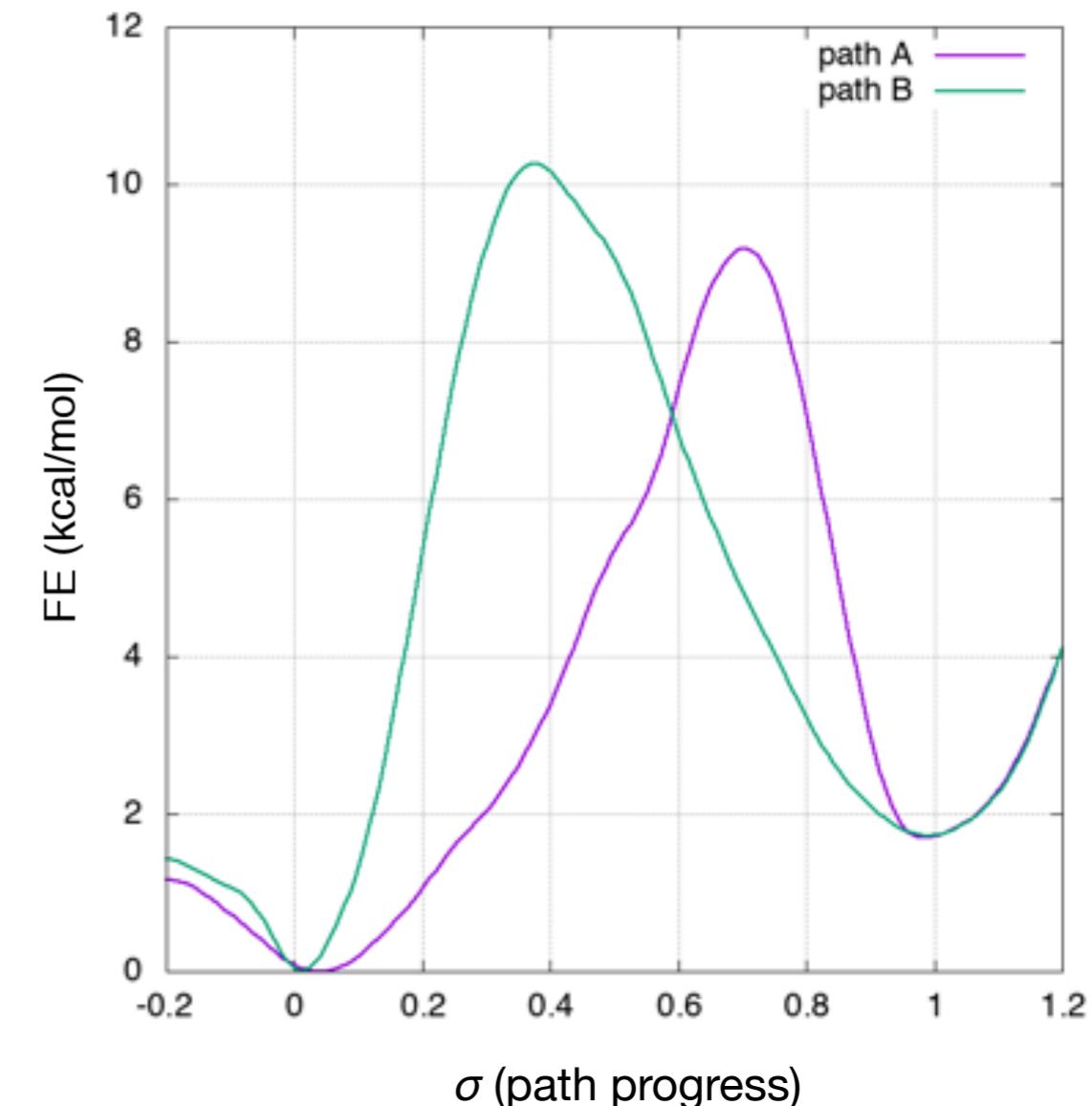
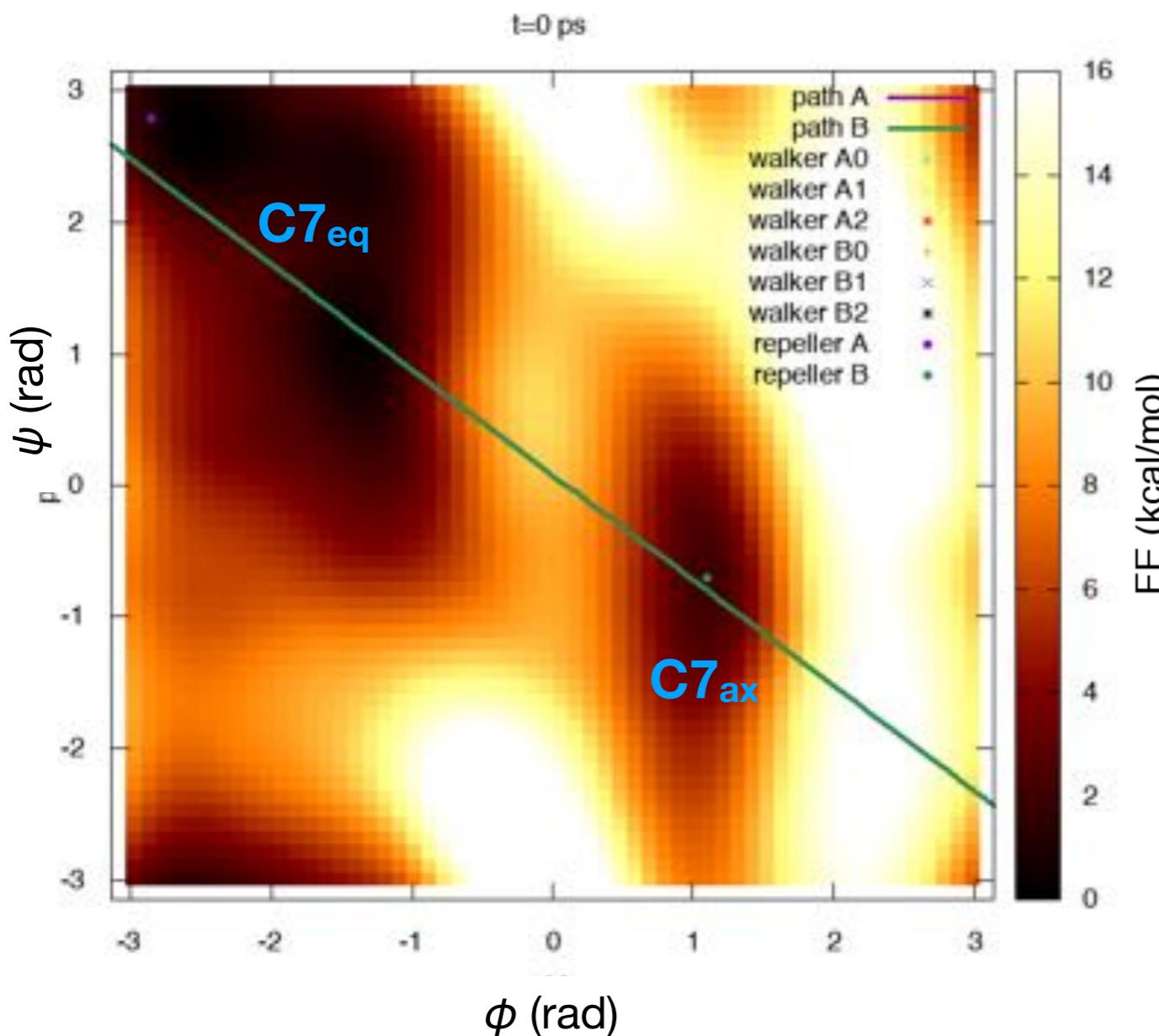


Alanine dipeptide revisited

MuWaMuPaMetaDyn

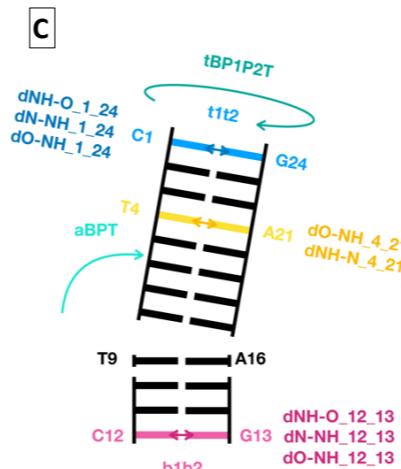
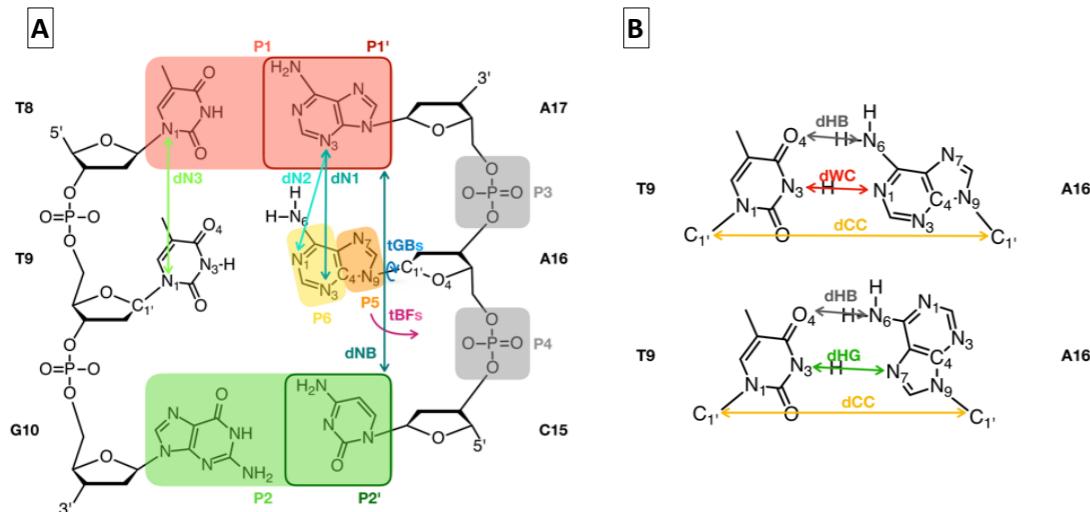
Multiple-Walker Multiple-Path Metadynamics

Alanine dipeptide in vacuo (AMBER99SB-ILDN ff)

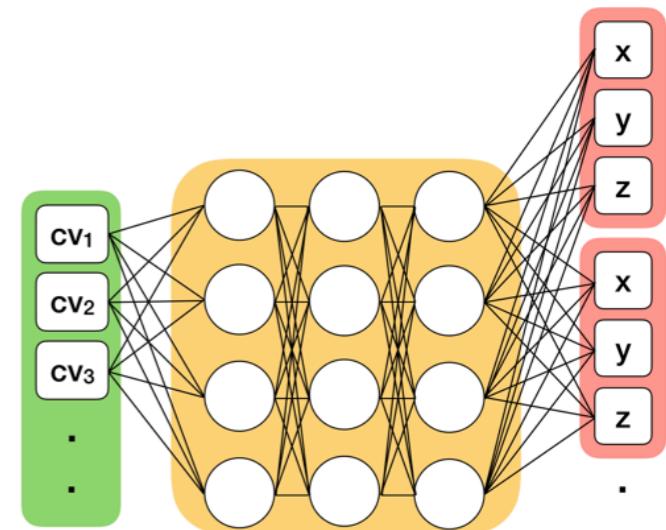


Iterative learning of collective variables

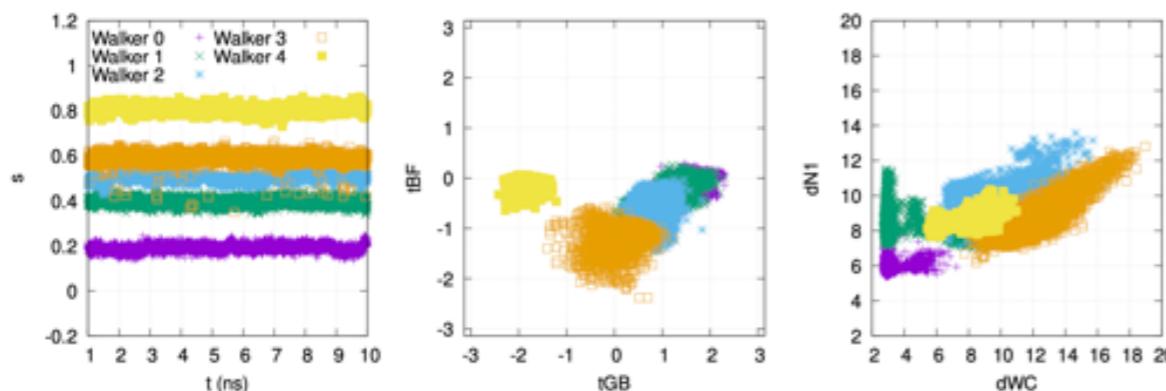
Large pool of collective variables



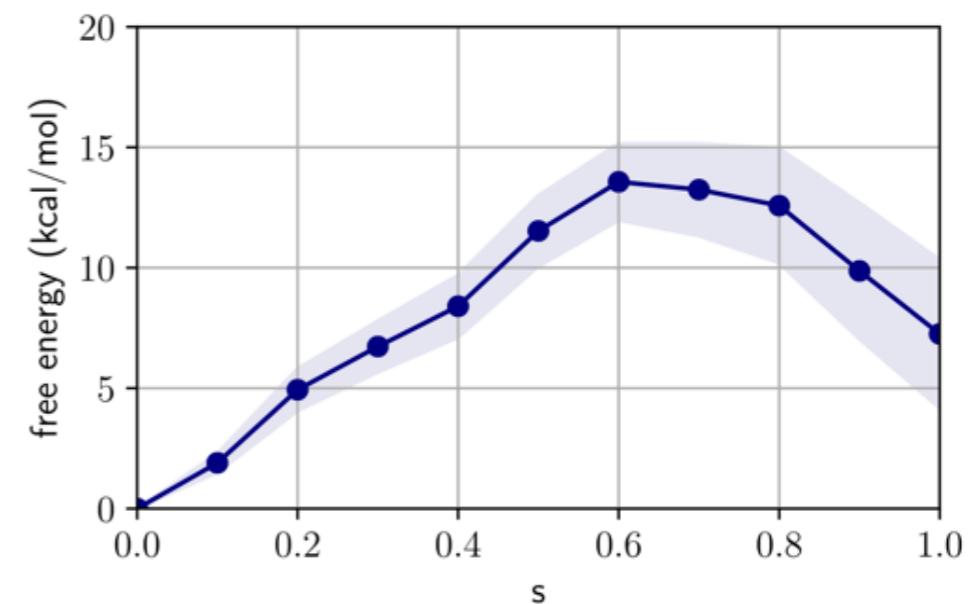
Machine learning



Biased Path-CV run with optimal CV selection



Free energy profile



Finding CVs is a chicken-and-egg problem, solvable with iterative approach.

Acknowledgements

Alberto Pérez de Alba Ortíz

Rakesh Chandran
Puthenkalathil

Ambuj Tiwari



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

Simultaneous sampling of multiple transition channels using adaptive paths of collective variables.
Alberto Pérez de Alba Ortíz and Bernd Ensing
arXiv [cond-mat.stat-mech] 2112.04061 (2021)

Contents

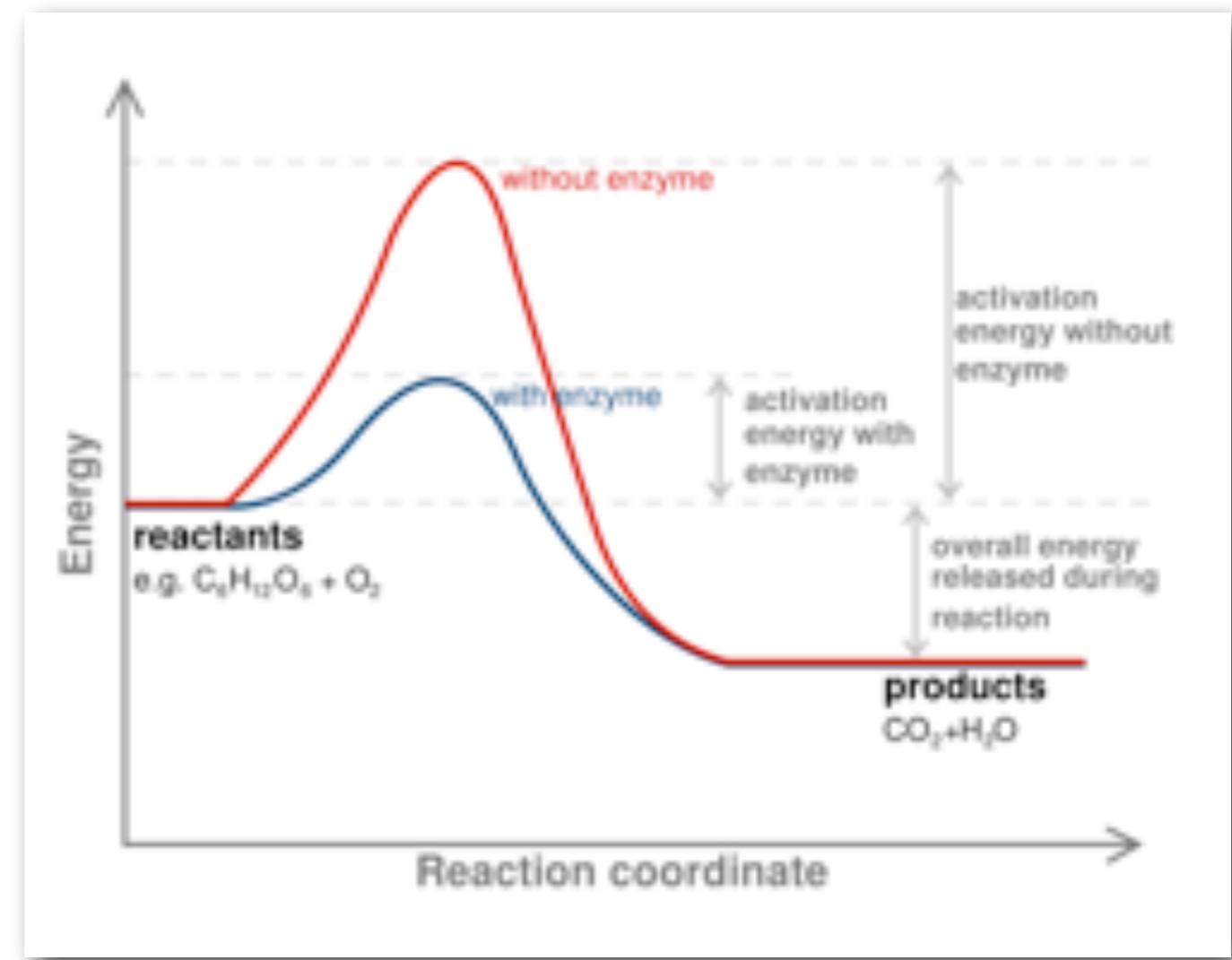
- Rare event simulation
- The metadynamics method
- Path-metadynamics
- Multi paths, multiple walkers
- Combined with machine learning to find CVs
- Exercises

Rare events

In real life (experiment)

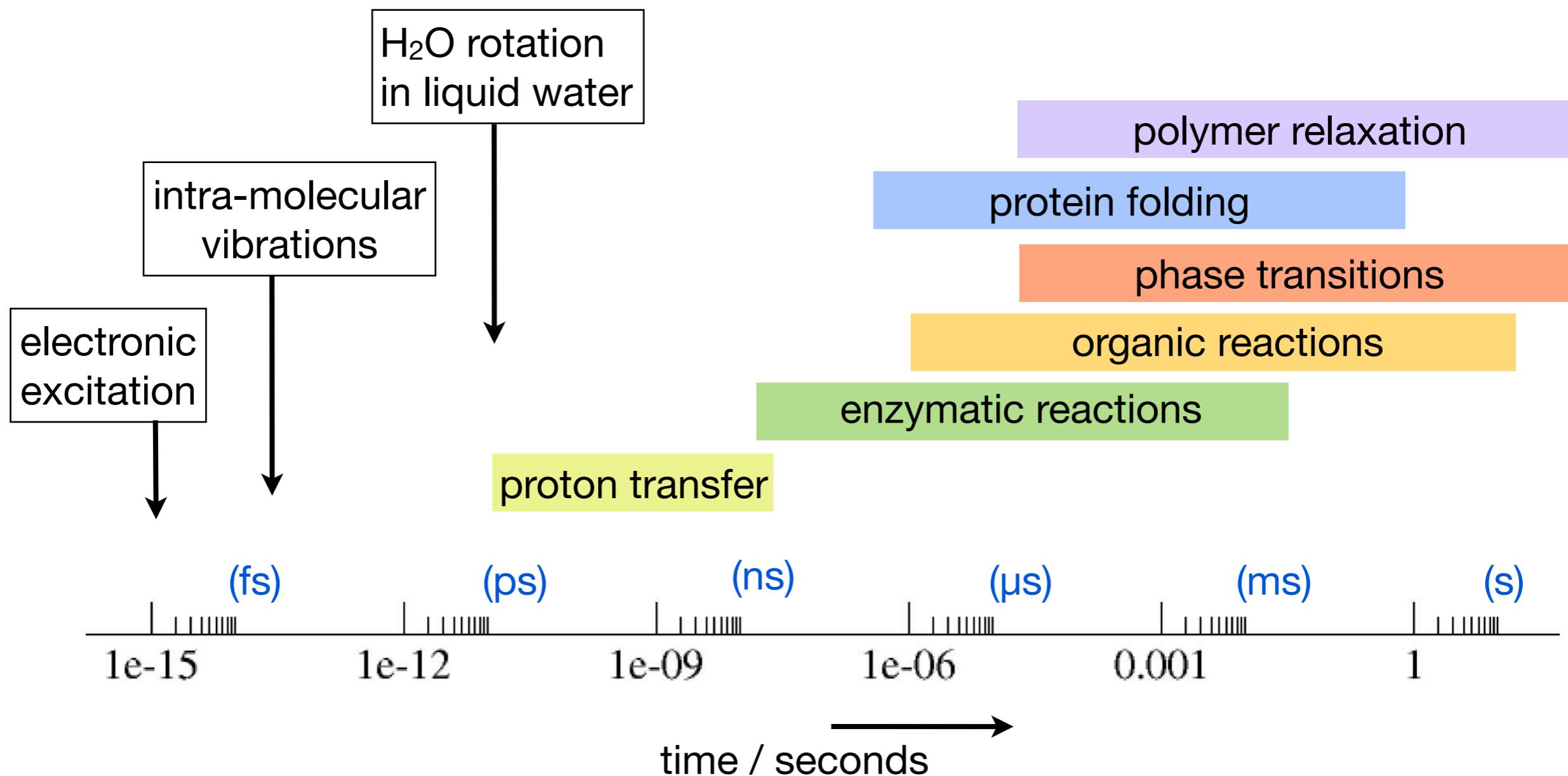
- no direct access to free energy
- reaction rate measurement

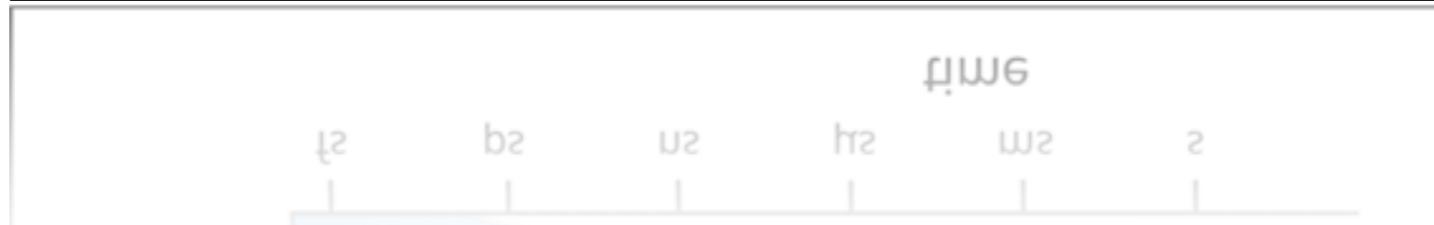
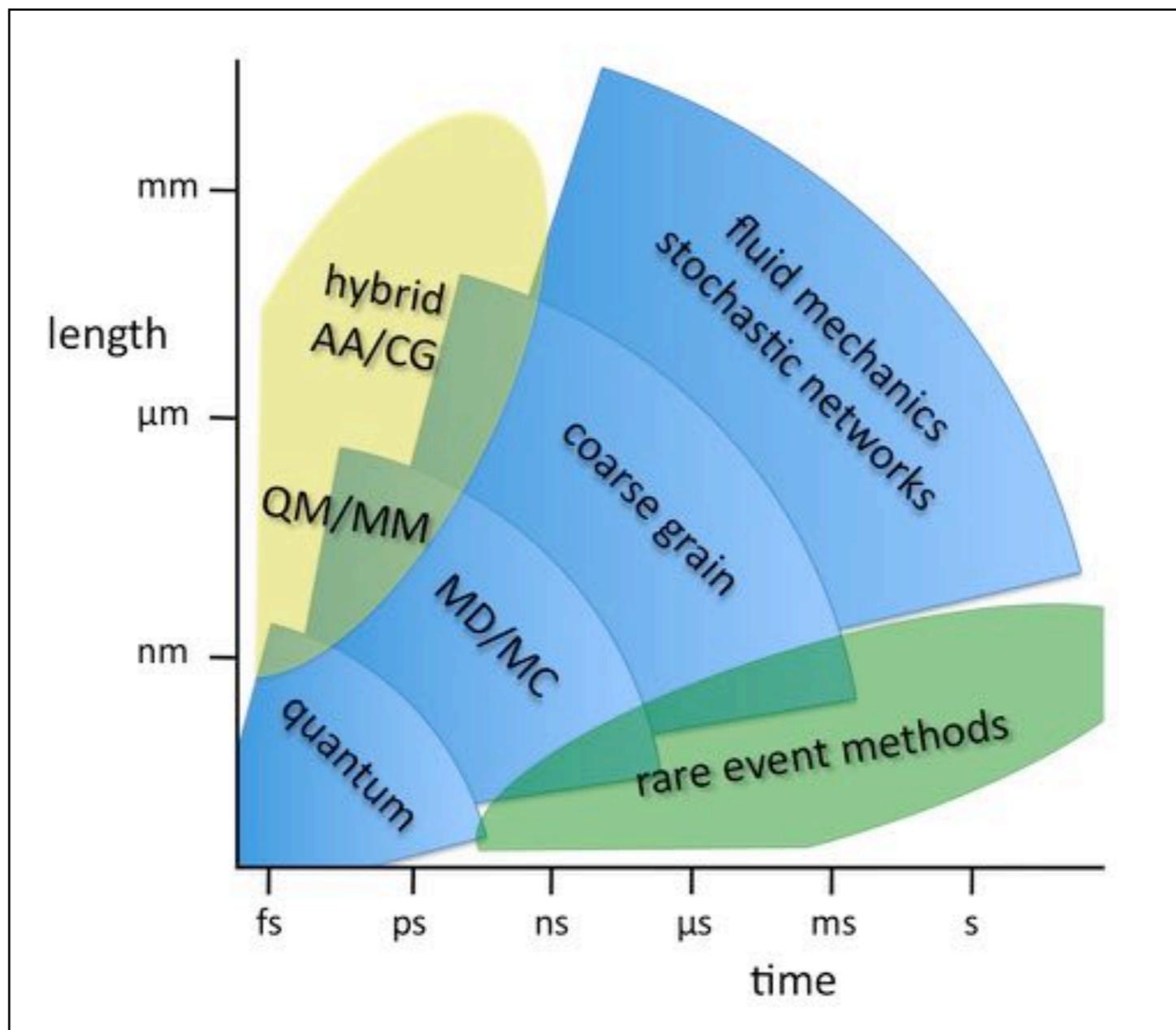
$$k = k_0 e^{-\Delta G/kT}$$



Rare events

In real life (experiment)

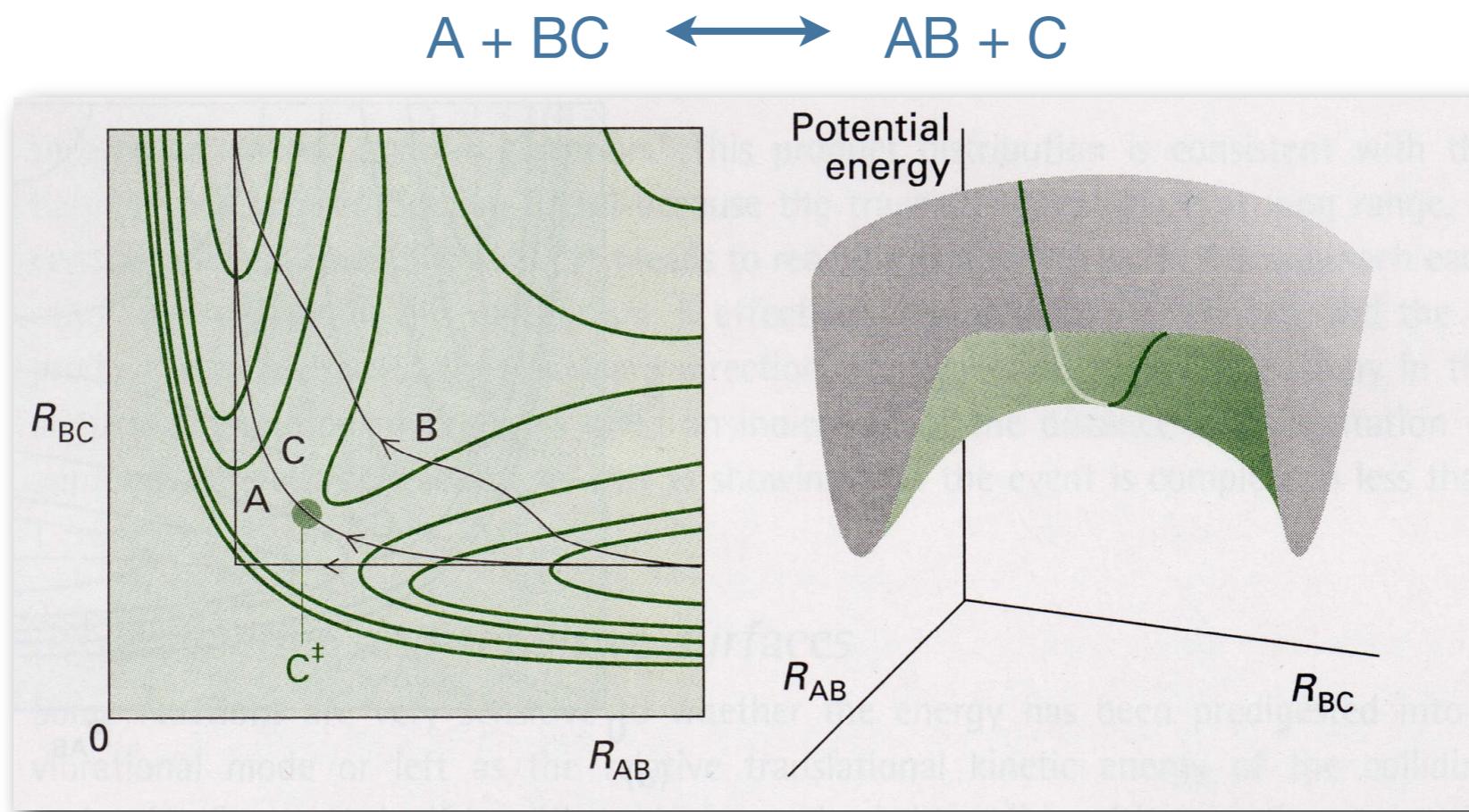




Rare events

Molecular Simulation and Phase Space Sampling

- molecular transitions are rare events
- direct simulation of reaction rate is (usually) impossible
- instead: calculation of relative (free) energies
- requires explicit form of the reaction coordinate



good reaction coordinate is often difficult to define

Some rare event methods

Added bias potential

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

Finding/sampling a reaction path

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

Thermodynamic Integration

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

Raising the temperature

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

Contents

- Rare event simulation
- The metadynamics method
- Path-metadynamics
- Multi paths, multiple walkers
- Combined with machine learning to find CVs
- Exercises

Metadynamics

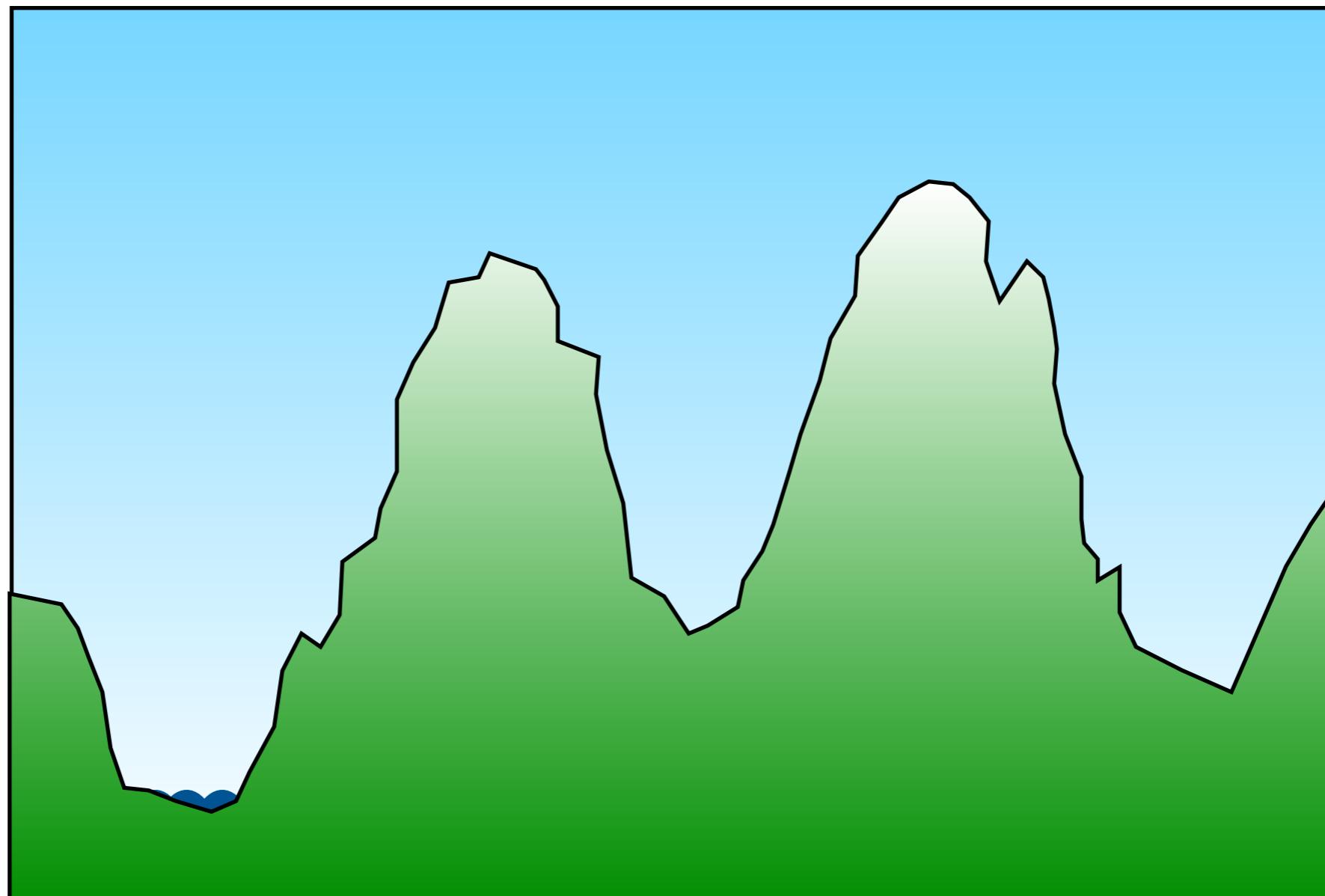
Escaping free-energy minima,
Laio and Parrinello, PNAS (2002)

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

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

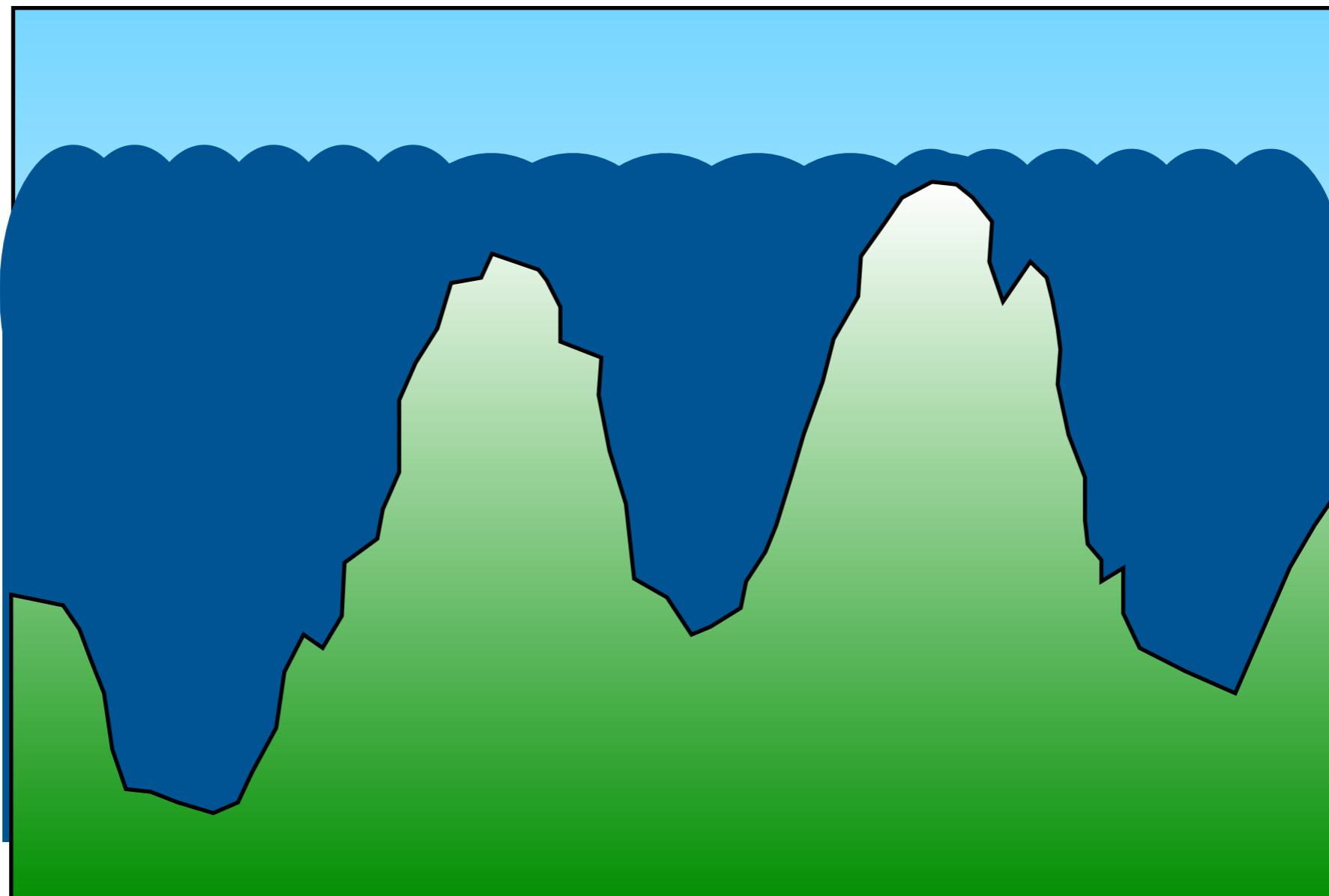
Flooding the landscape with hills

Metadynamics, Laio and Parrinello, PNAS (2002)

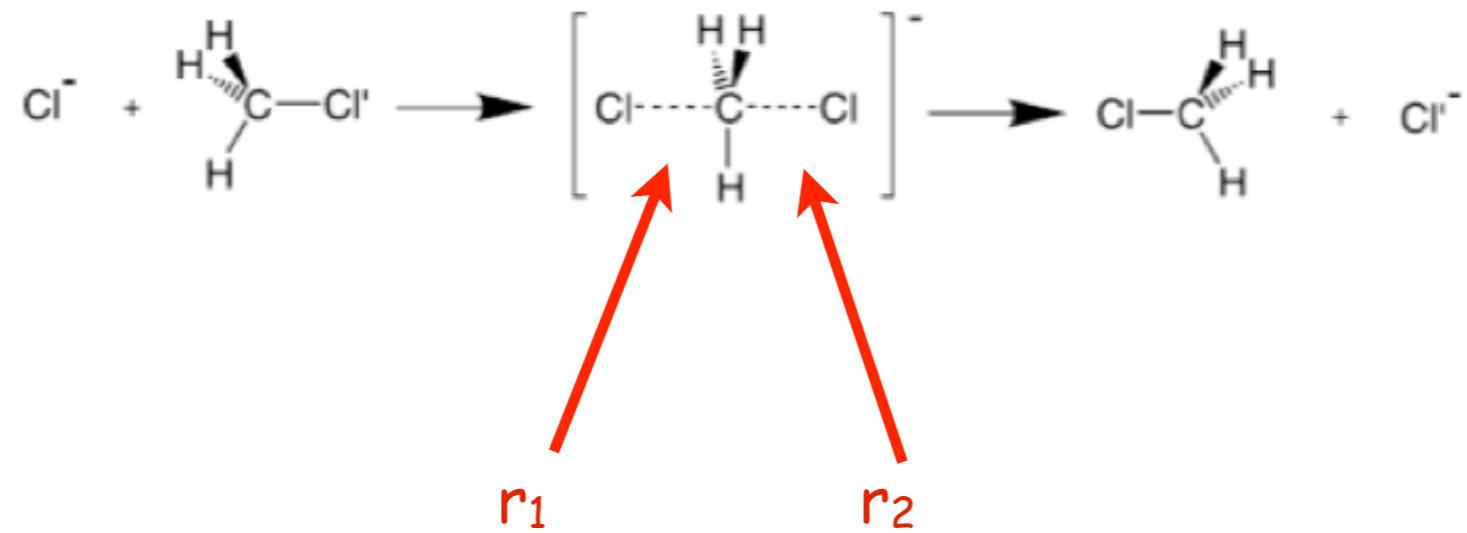


Flooding the landscape with hills

Metadynamics, Laio and Parrinello, PNAS (2002)

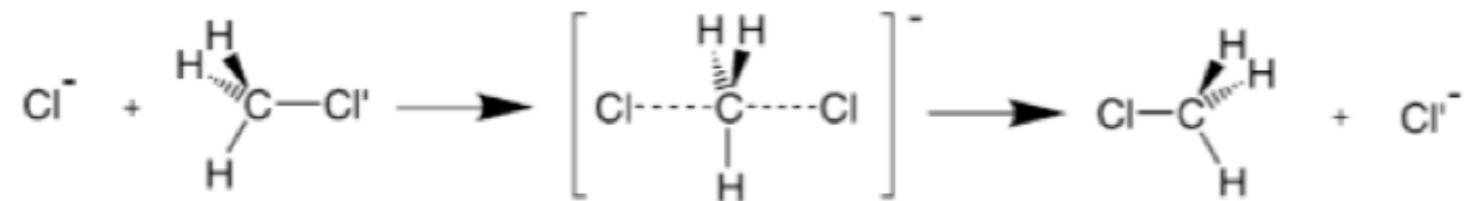


simple example: S_N2 reaction

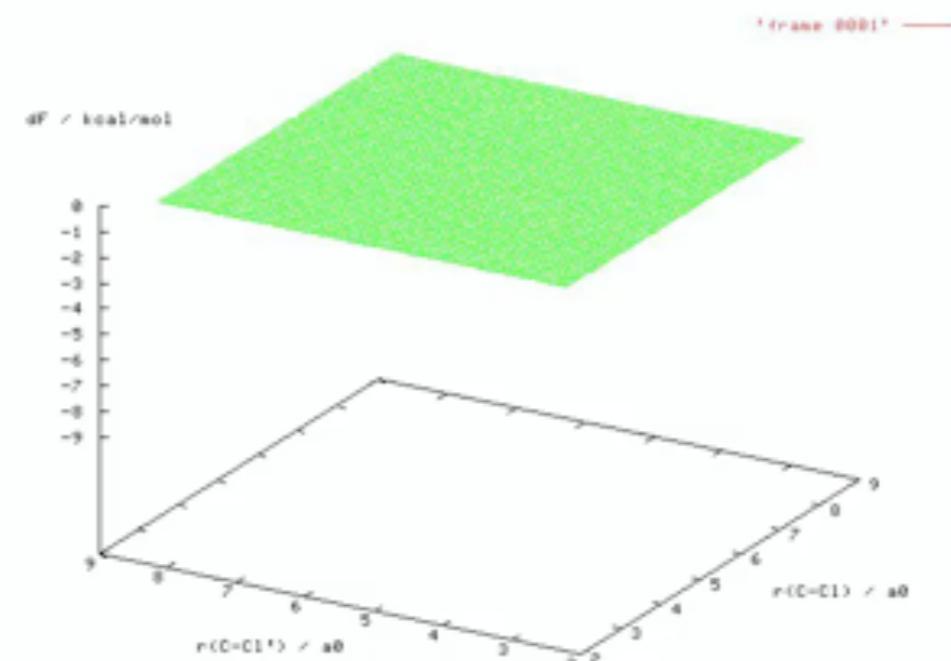
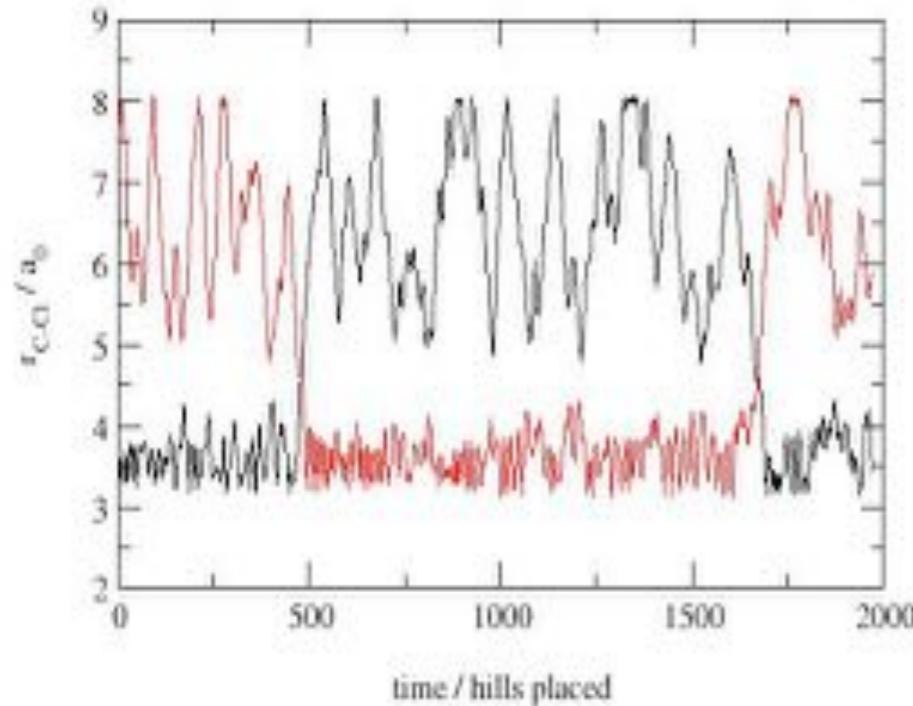


Two collective variables: r_1 and r_2

simple example: S_N2 reaction

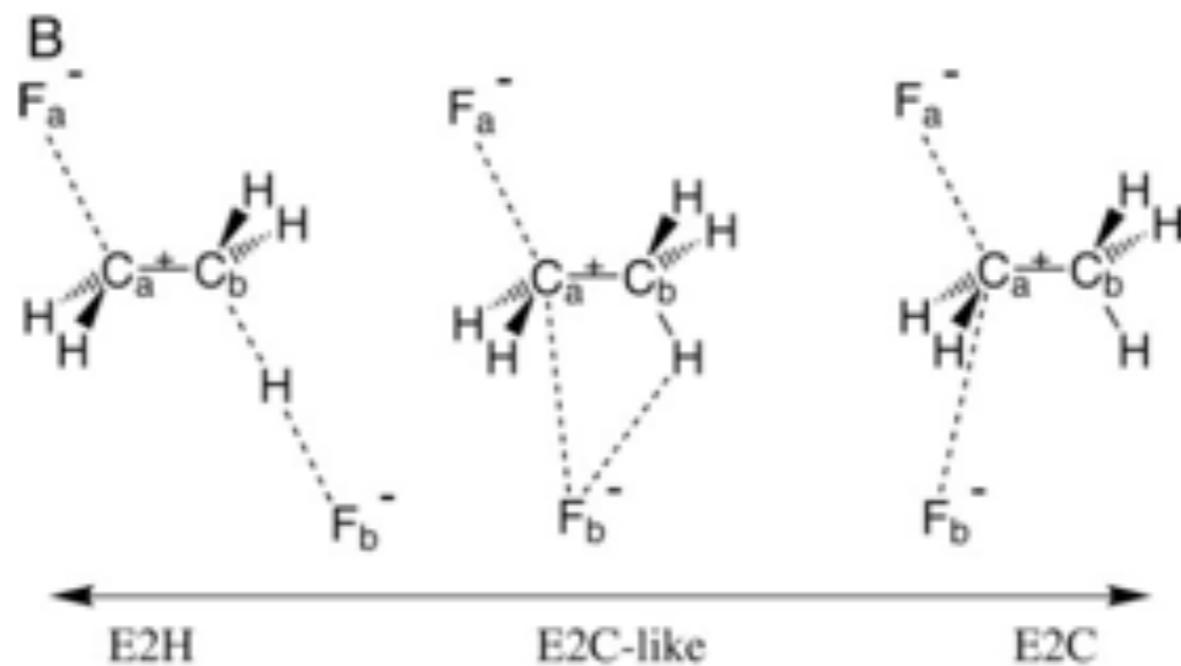
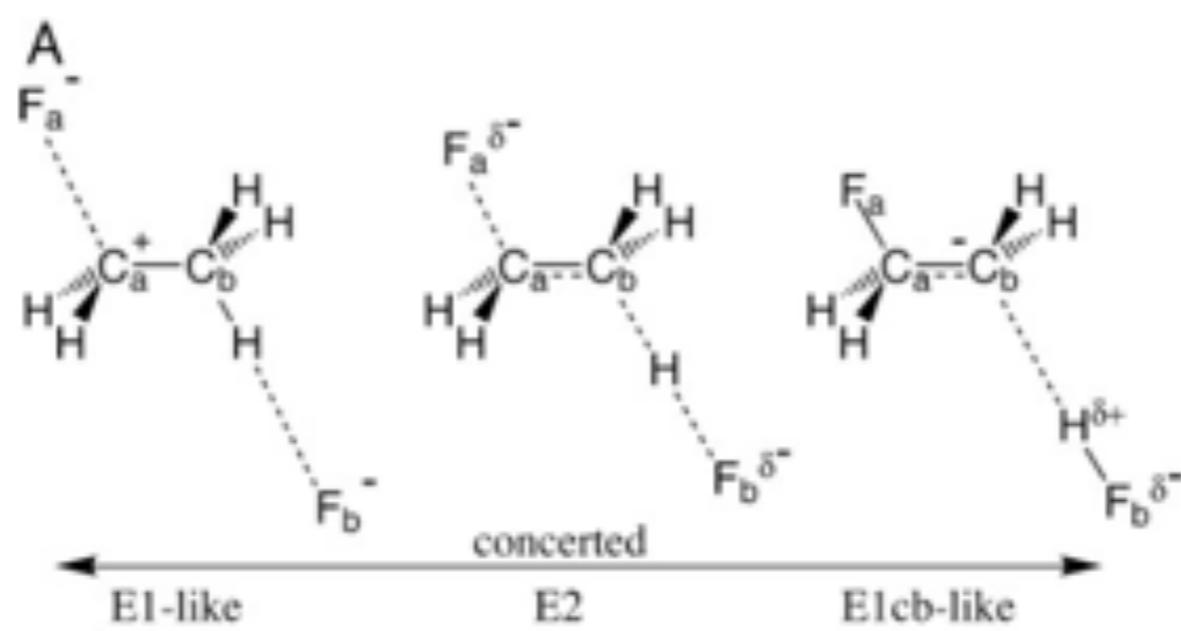


Two collective variables: r_1 and r_2

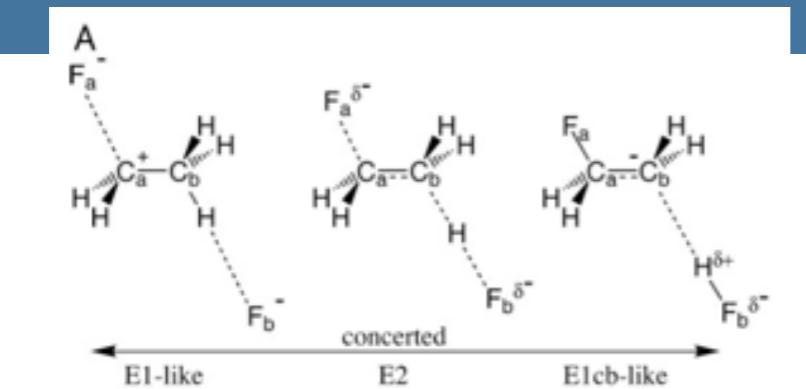


- perfect reaction coordinate not needed
- lowest free energy pathway

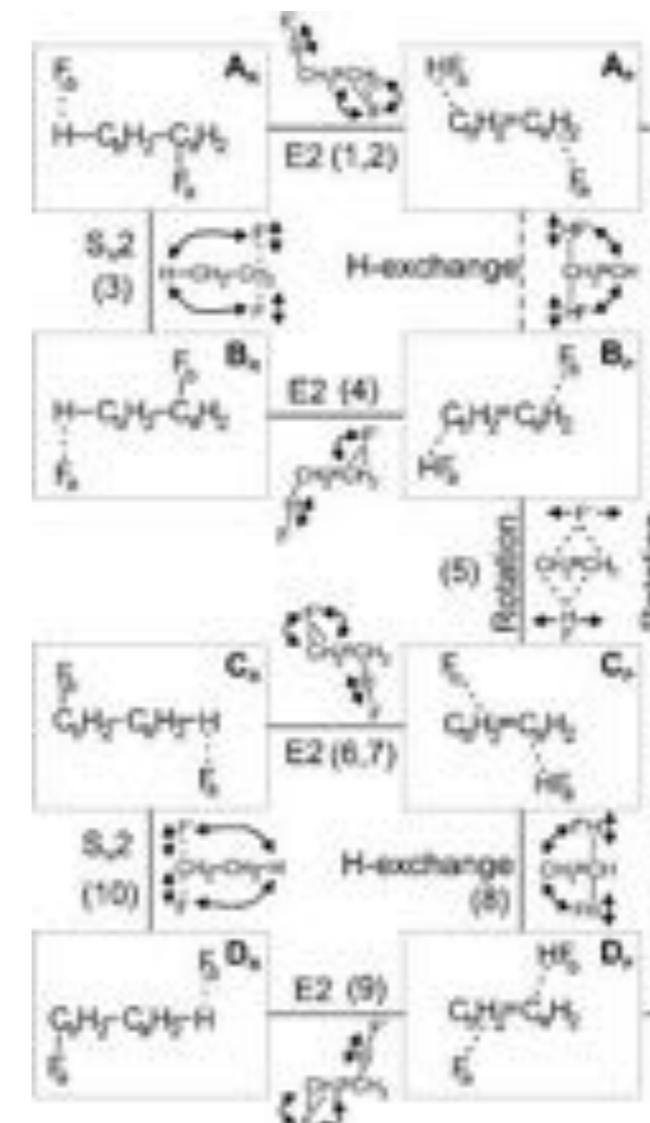
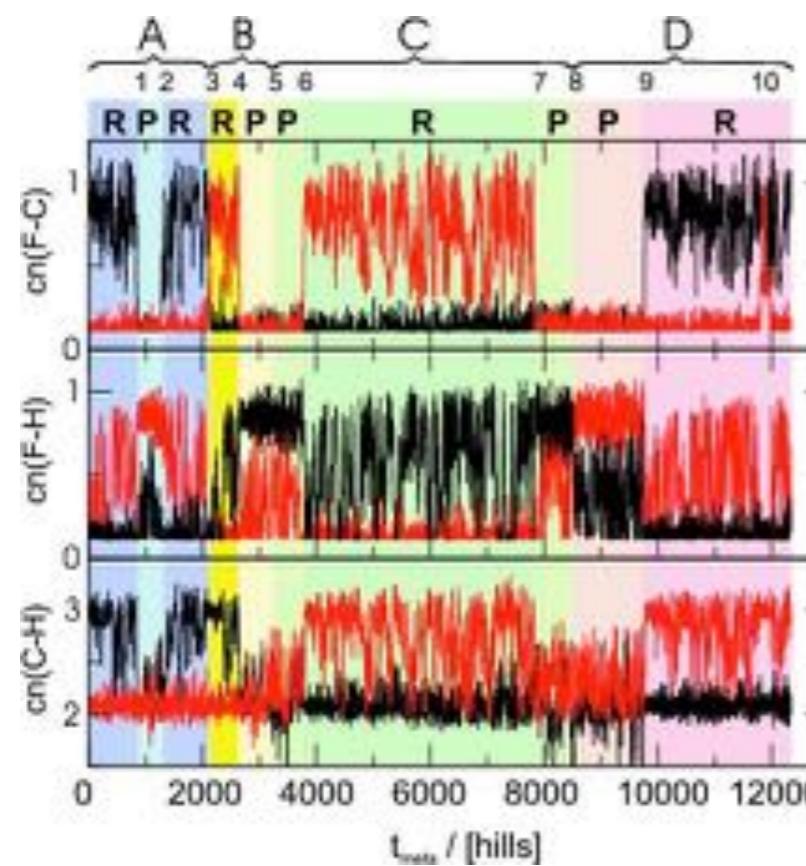
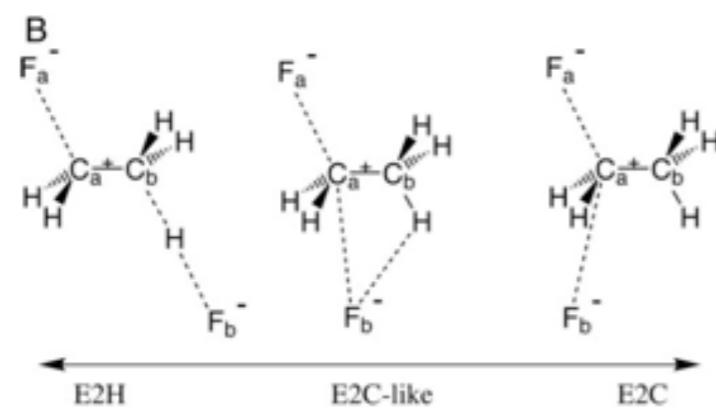
E2 vs S_N2 reaction



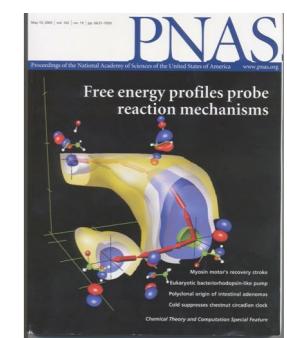
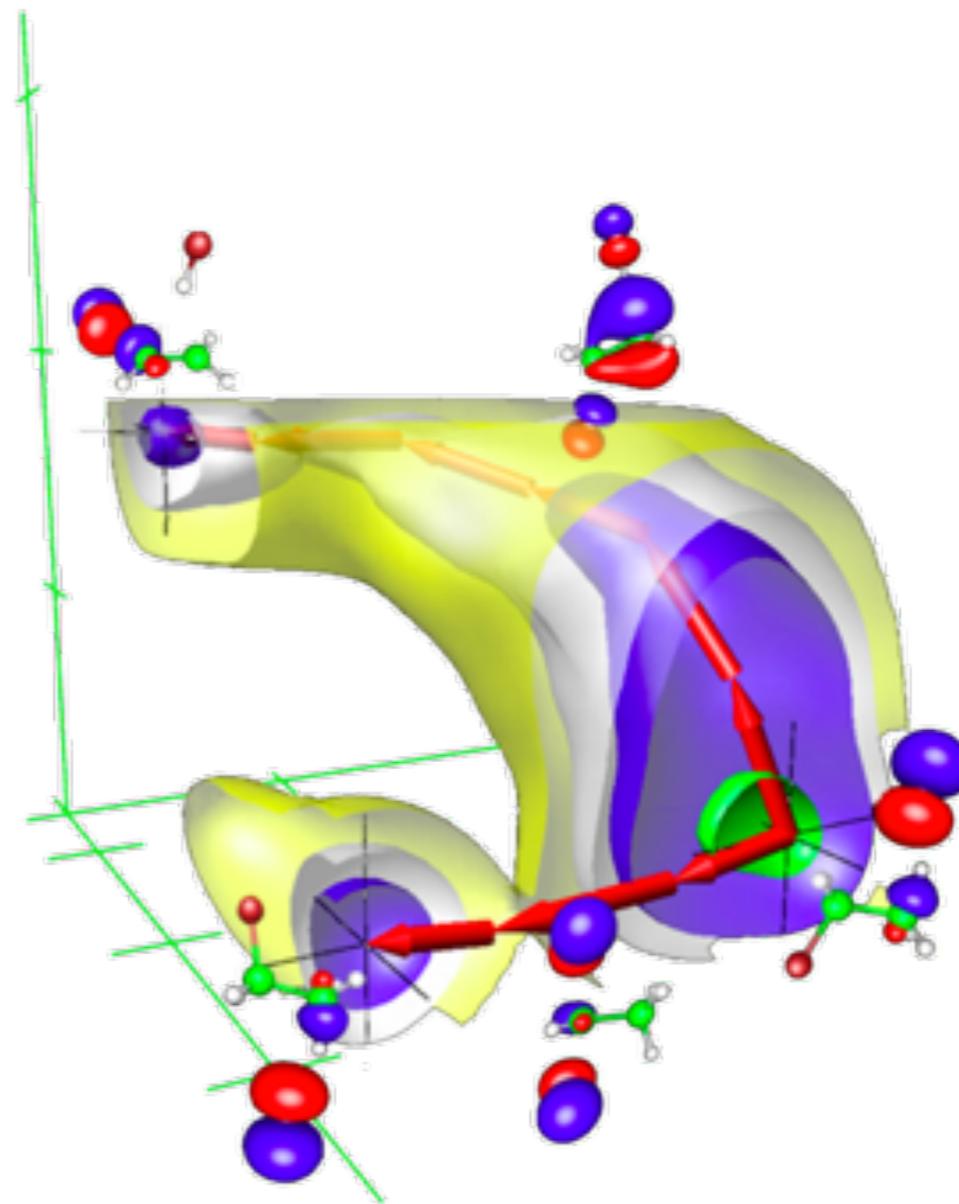
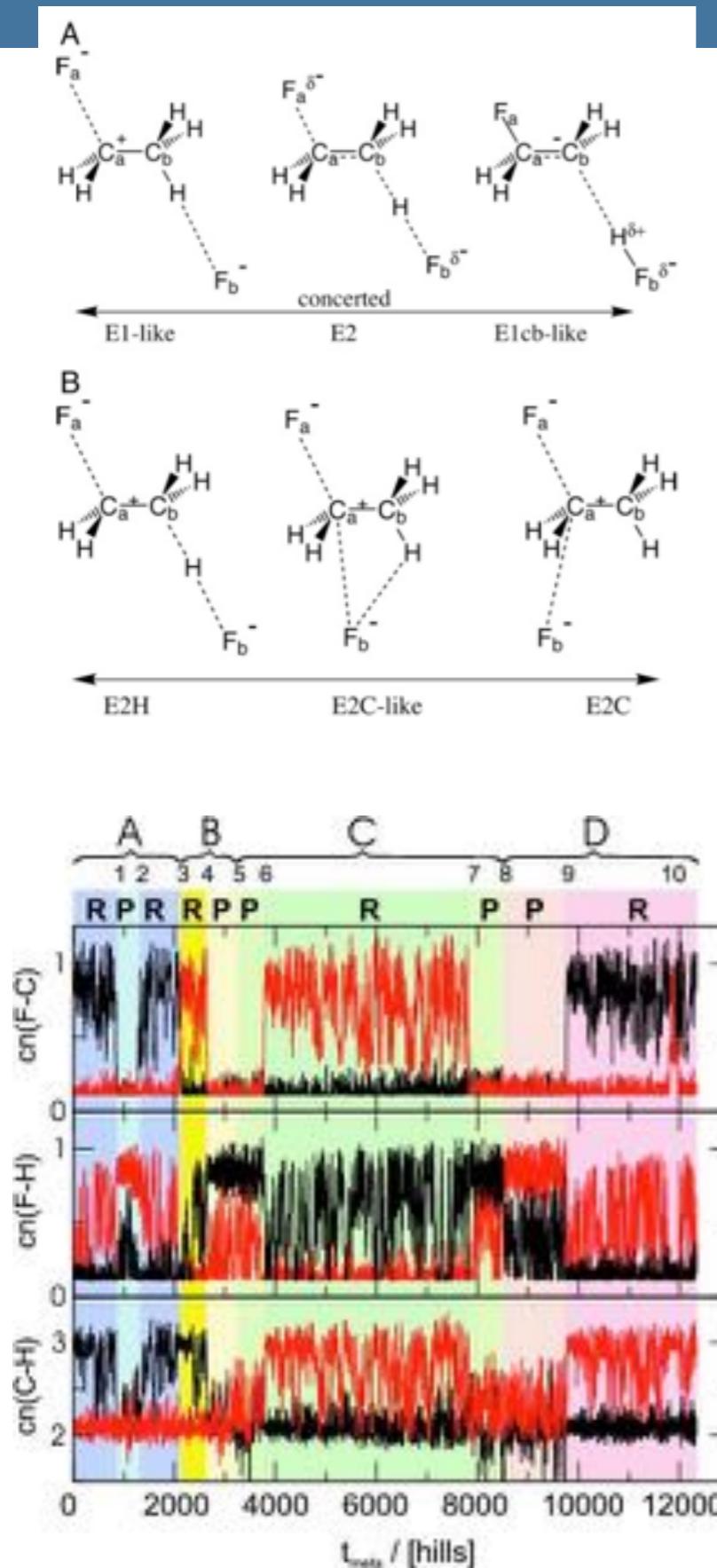
E2 vs S_N2 reaction



8 stable states



E2 vs S_N2 reaction



Accuracy of metadynamics

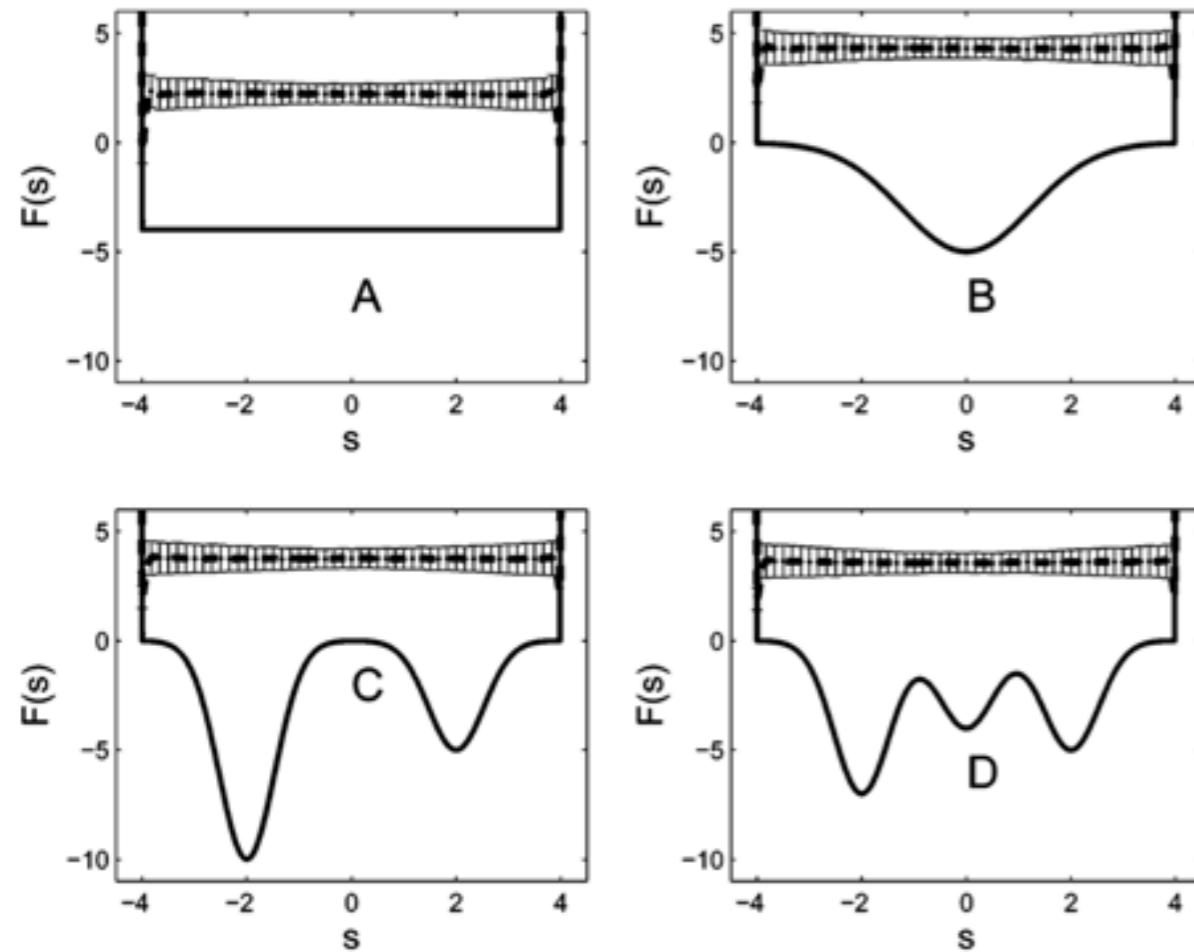
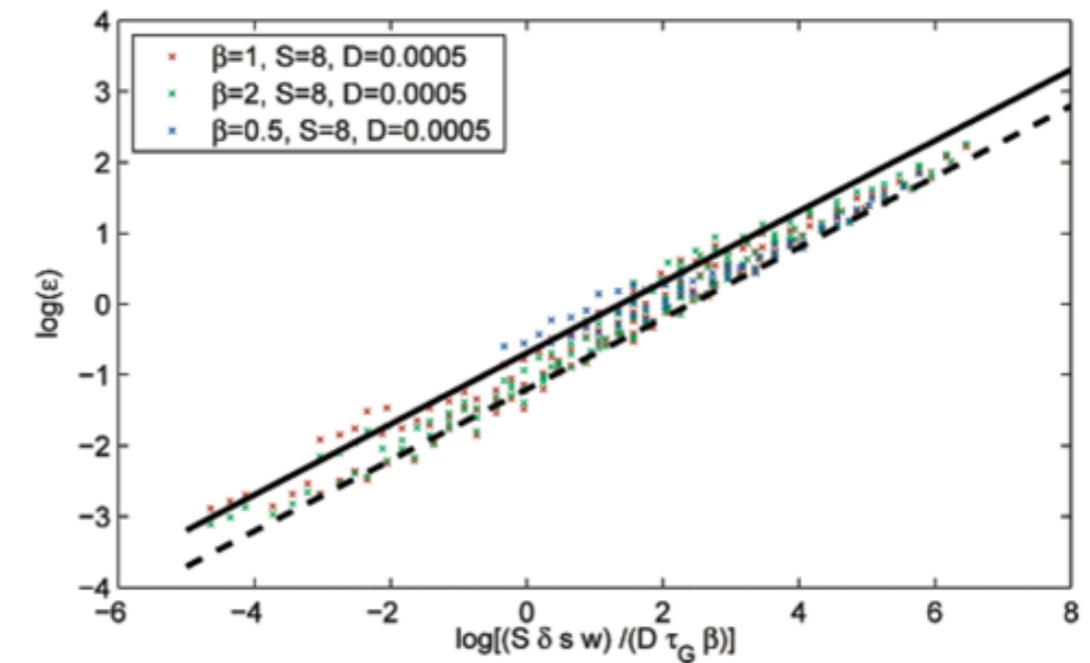


Figure 1. Metadynamics results for four different free energy profiles: (A) $F(s) = -4$; (B) $F(s) = -5 \exp(-(s/1.75)^2)$; (C) $F(s) = -5 \exp(-(s - 2/0.75)^2) - 10 \exp(-(s + 2/0.75)^2)$; (D) $F(s) = -5 \exp(-(s - 2/0.75)^2) - 4 \exp(-(s/0.75)^2) - 7 \exp(-(s + 2/0.75)^2)$. The average $\langle F(s) - F_G(s, t) \rangle$ 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

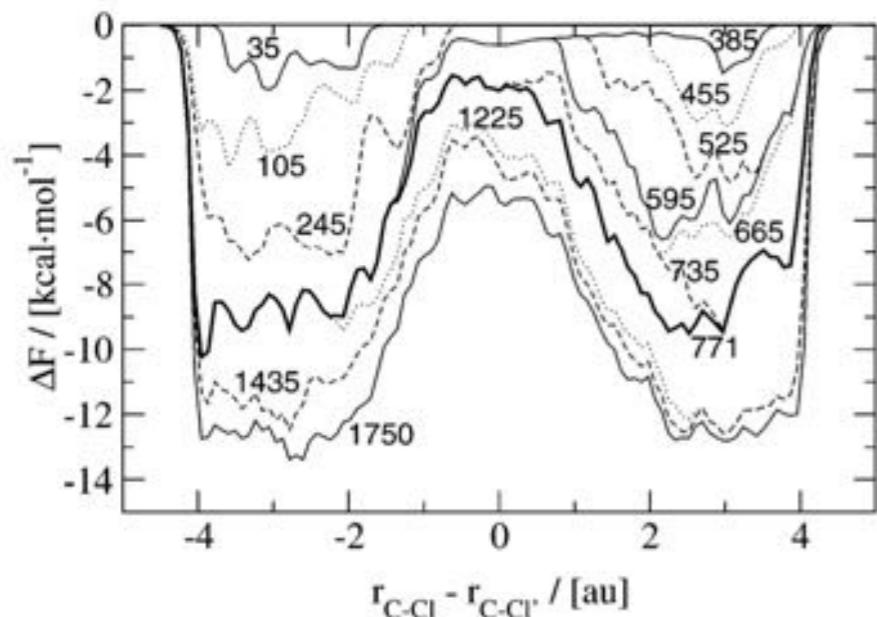
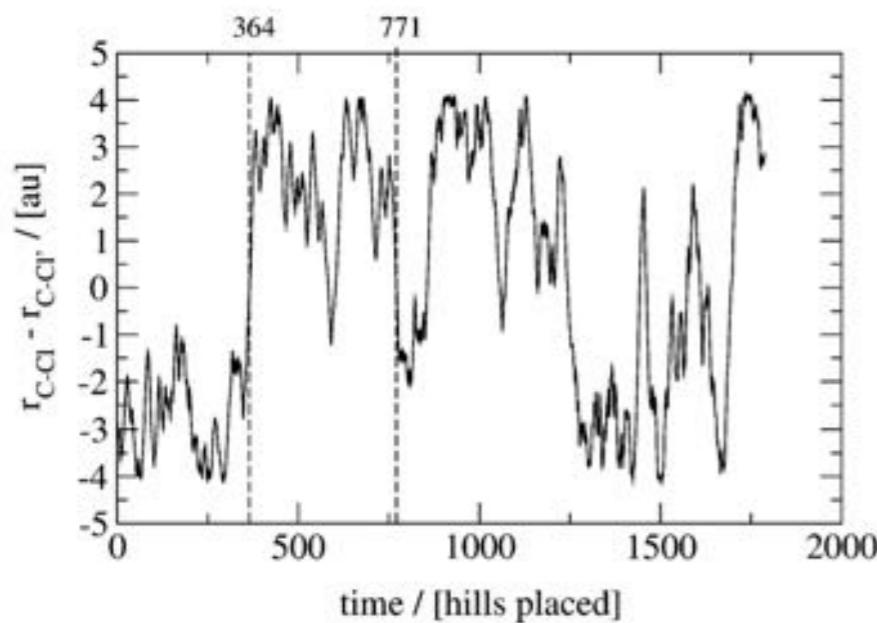
D : diffusion coefficient

S : size of landscape

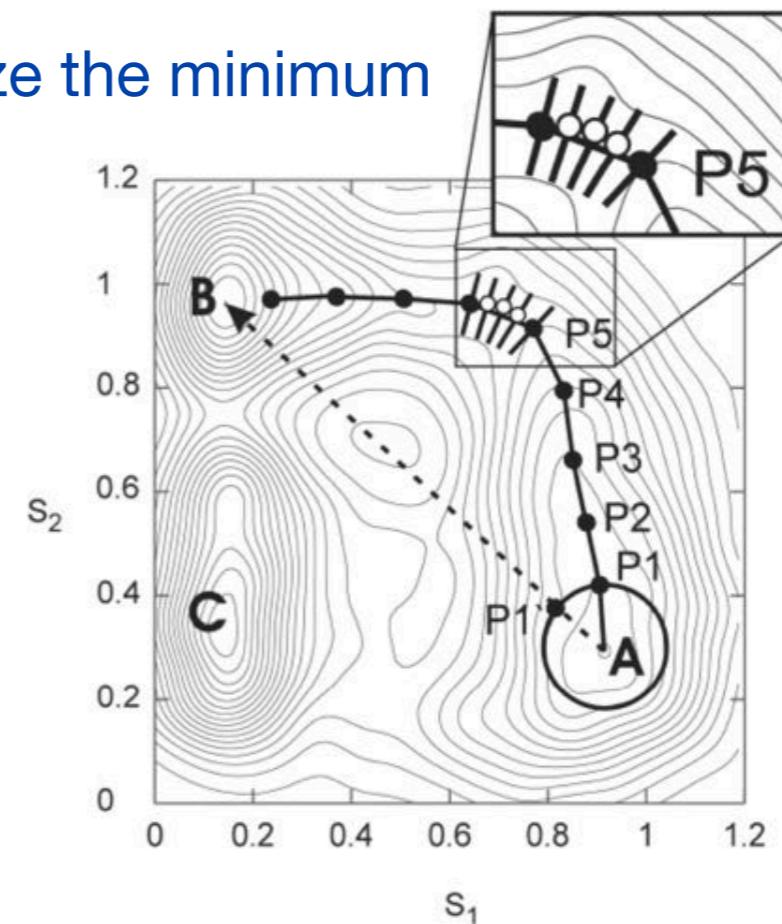
Dashed line is error: $\epsilon = C(d) \sqrt{\frac{HWS}{\tau D \beta}}$

A recipe

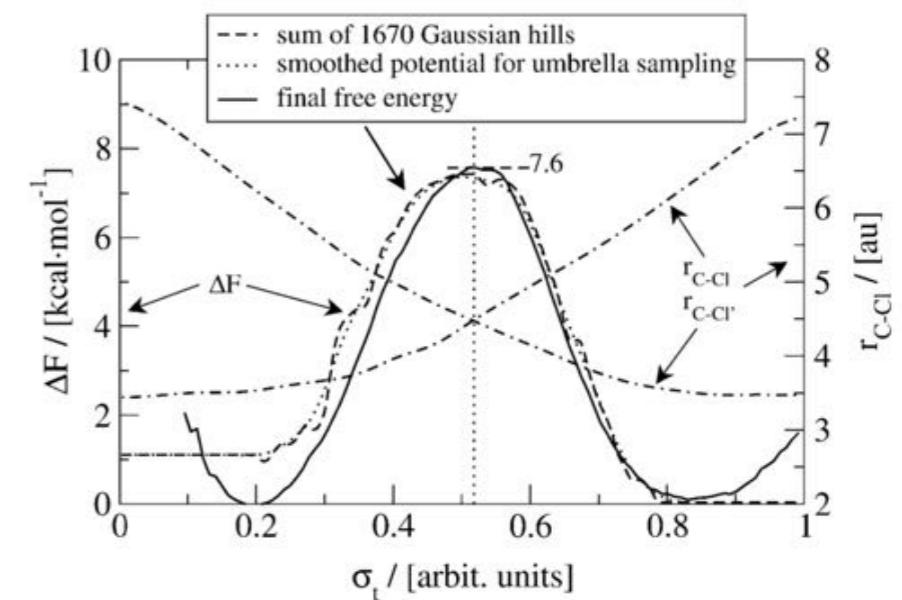
(1) Decrease hills with every recrossing for convergence



(2) Localize the minimum FE path

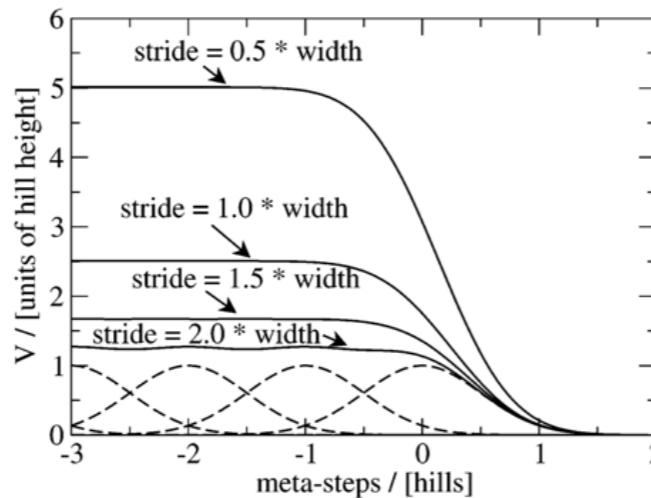


(3) Switch to 1D to converge further

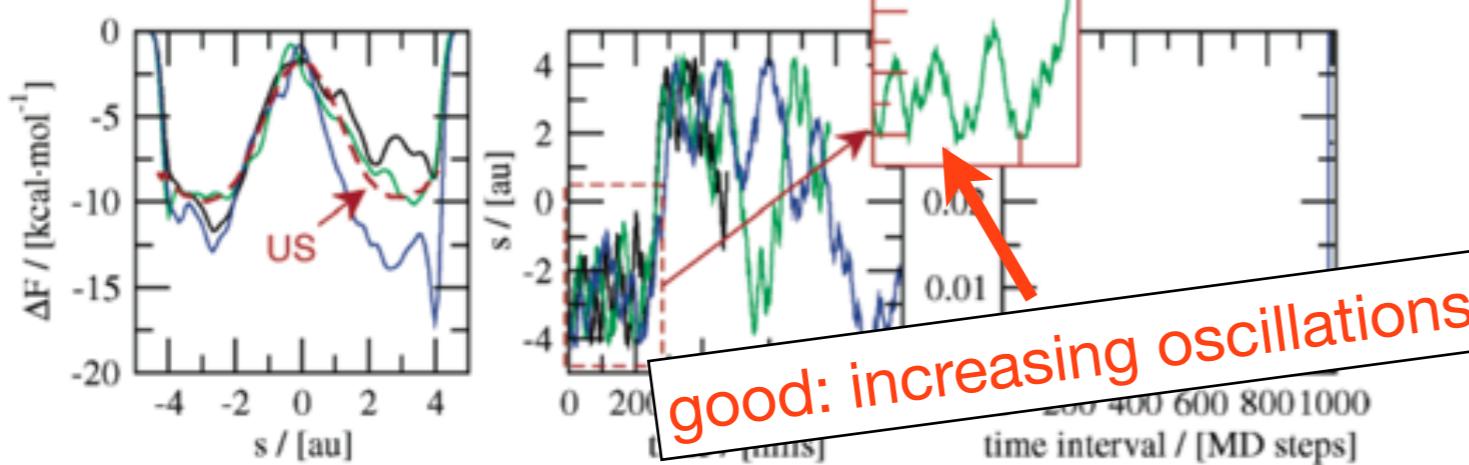
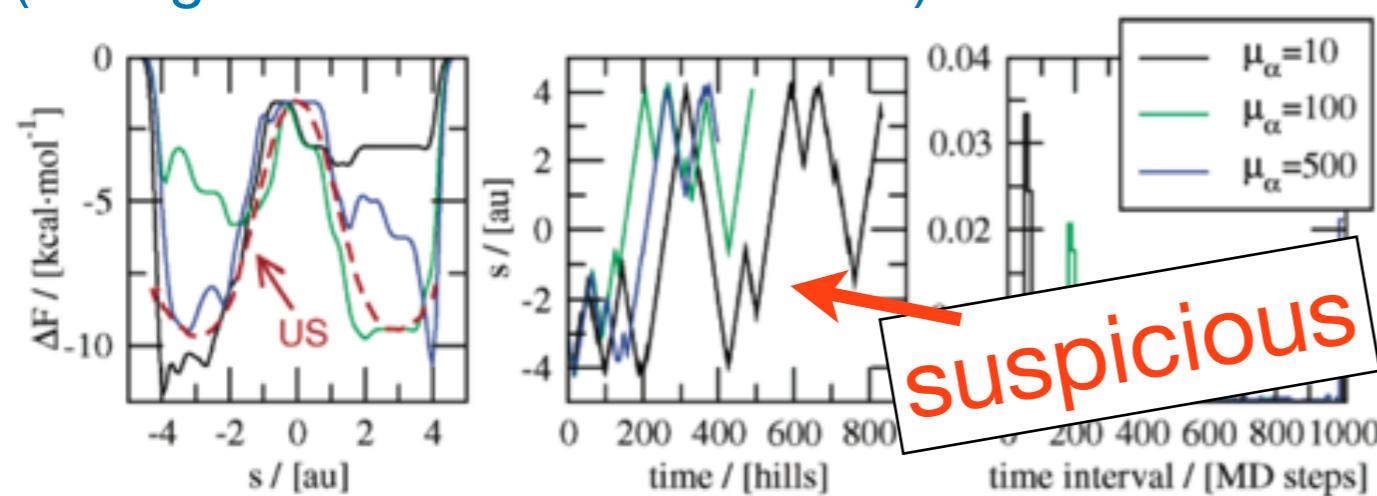


Possible pitfalls

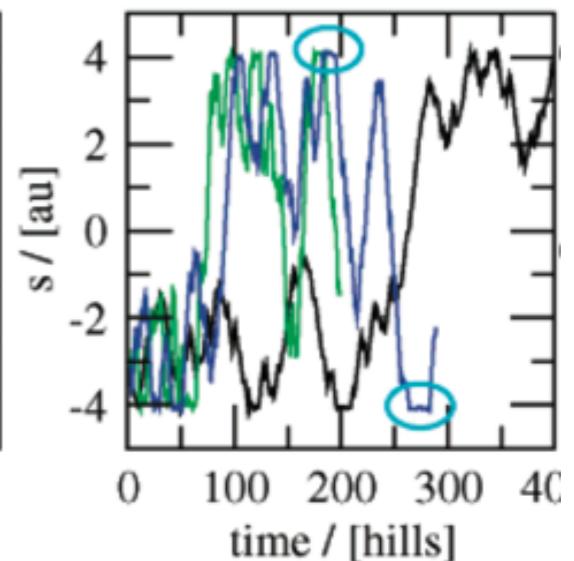
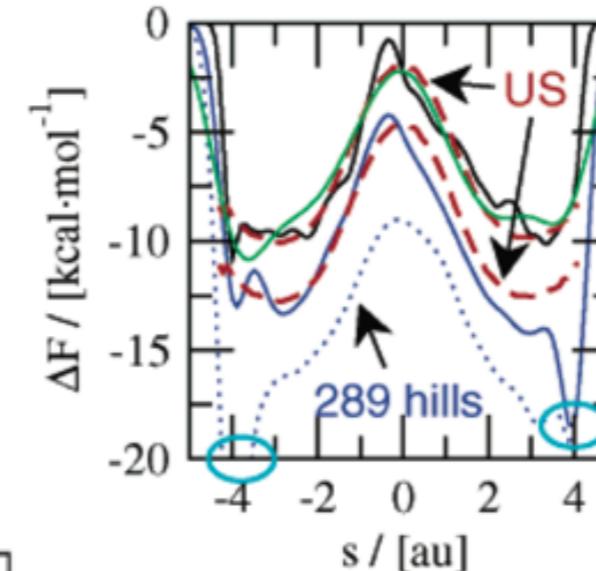
(1) “hill surfing” (too large hills/small time interval)



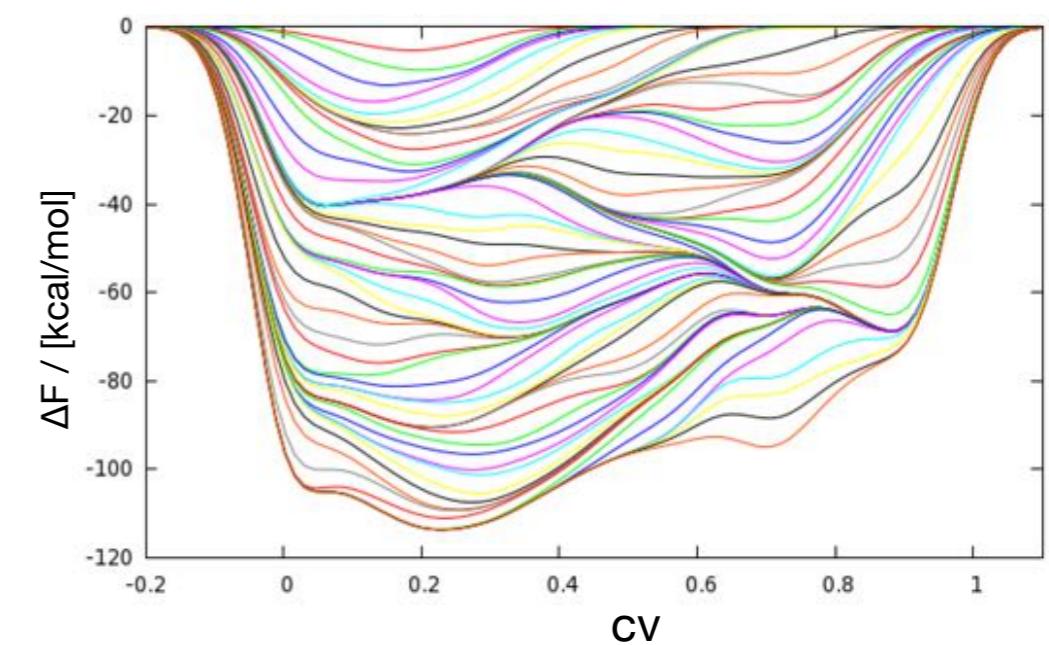
(2) strange dynamics (wrong masses/force-constants)



(3) “sticking (too hard walls)”

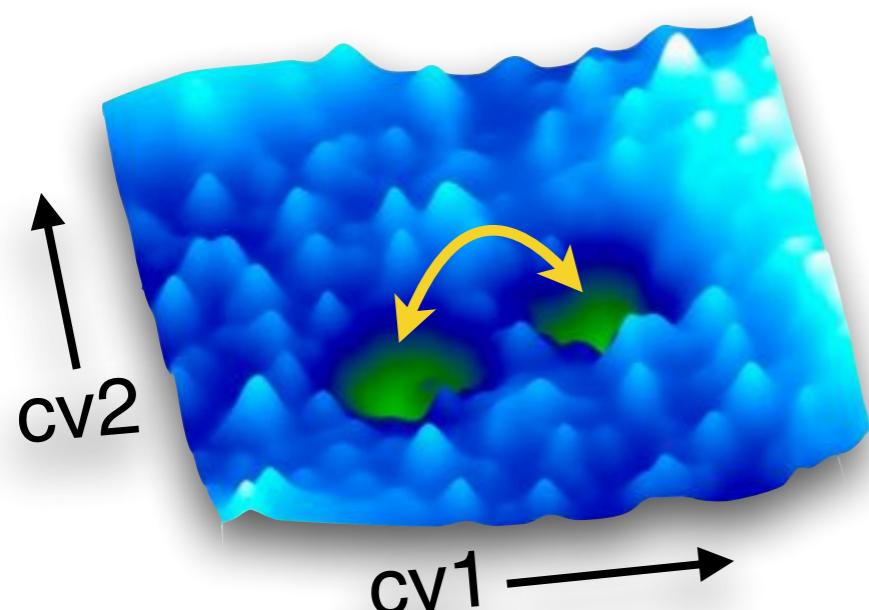
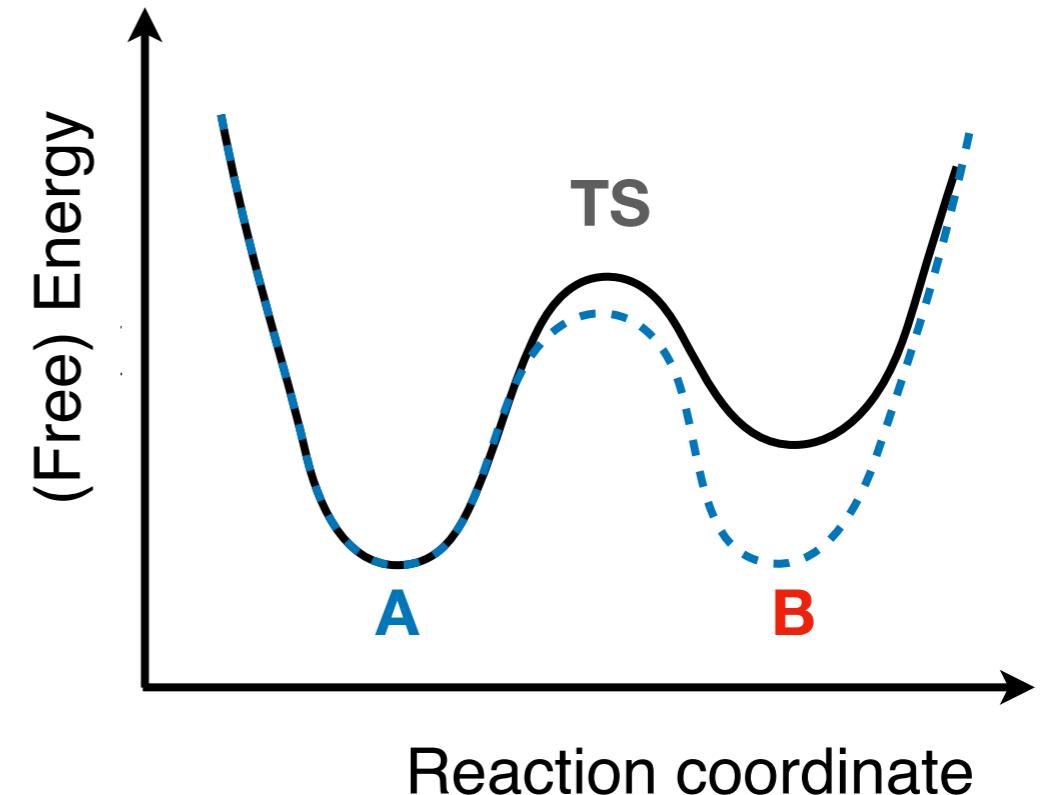


(4) hysteresis (wrong collective variables)

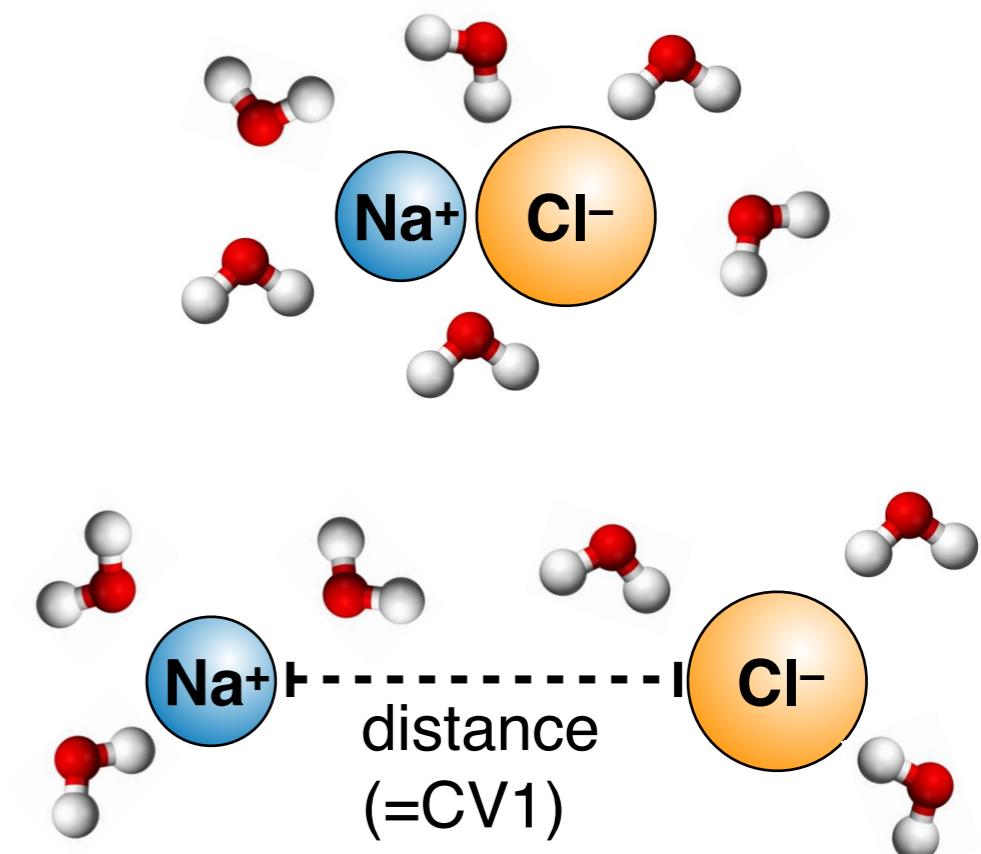


Do we have all relevant CVs?

- Observing activated transitions in a simulation is **very improbable**
- **Enhanced sampling** methods allow for barrier crossing
- Requires a **reaction coordinate** or a small set of **collective variables** (CVs)



Bonded NaCl molecule
↓
Dissociated Na^+ + Cl^- ions



Example of disaster

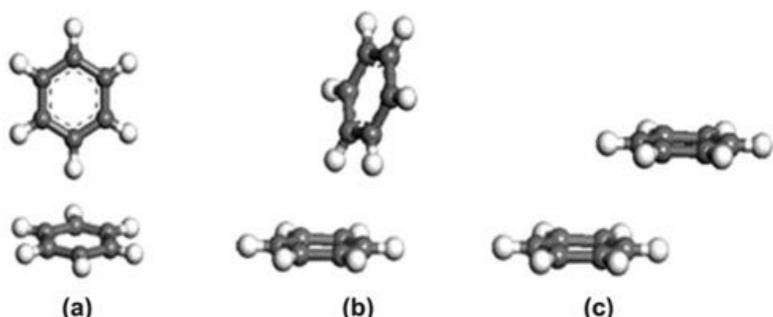
THE JOURNAL OF CHEMICAL PHYSICS 139, 201102 (2013)

Communication: Benzene dimer—The free energy landscape

Anil Kumar Tummanapelli and Sukumaran Vasudevan^{a)}

Department of Inorganic and Physical Chemistry, Indian Institute of Science, Bangalore 560012, India

(Received 24 September 2013; accepted 14 November 2013; published online 26 November 2013)

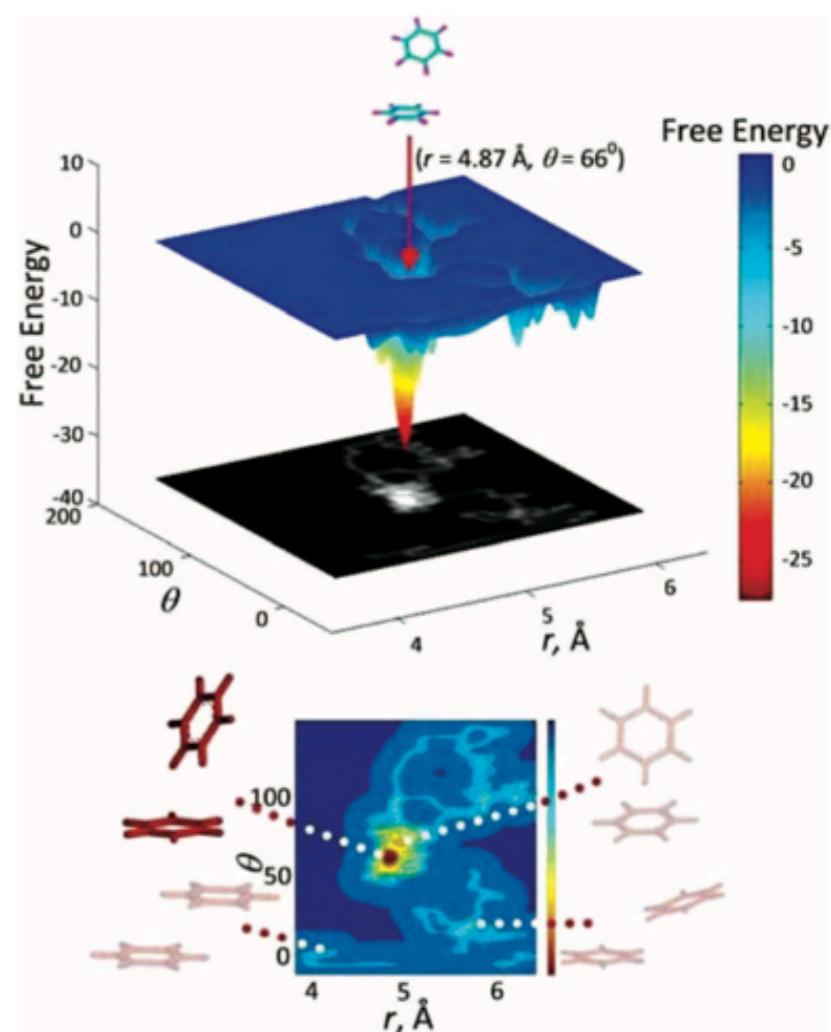


Establishing the relative orientation of the two benzene molecules in the dimer has remained an enigmatic challenge. Consensus has narrowed the choice of structures to either a T-shape, that may be tilted, or a parallel displaced arrangement, but the relatively small energy differences makes identifying the global minimum difficult. Here we report an *ab initio* Car-Parrinello Molecular Dynamics based metadynamics computation of the free-energy landscape of the benzene dimer. Our calculations show that although competing structures may be isoenergetic, free energy always favors a tilted T-shape geometry at all temperatures where the bound benzene dimer exist. © 2013 AIP Publishing LLC. [http://dx.doi.org/10.1063/1.4834855]

Known energy difference between T-shaped and parallel benzene dimer is about **0.2 kcal/mol**

CPMD metadynamics free energy difference using 2 collective variables (distance and angle): **27 kcal/mol** at T=2 K ???

Used Gaussian height: 3.1 kcal/mol (Aha!)



Well-tempered metadynamics

Rigorous convergence to the free energy surface

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

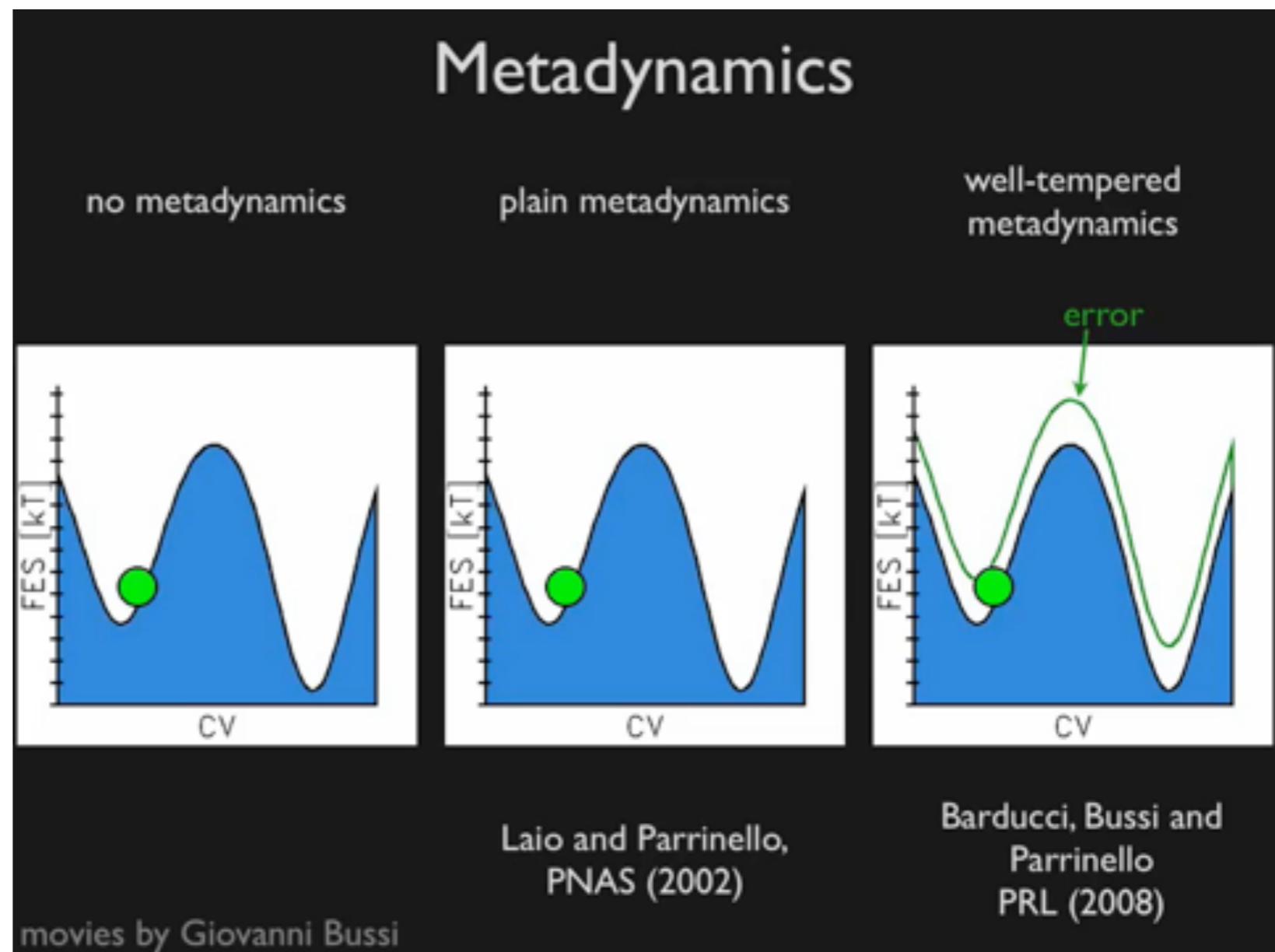
$$V(s, t) = \Delta T \ln \left(1 + \frac{\omega N(s, t)}{\Delta T} \right)$$

- $\Delta T = 0 \rightarrow$ bias is zero
- $\Delta T = \infty \rightarrow$ normal metdyn.
(no convergence in these limits)

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

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

Youtube movie by Giovanni Bussi



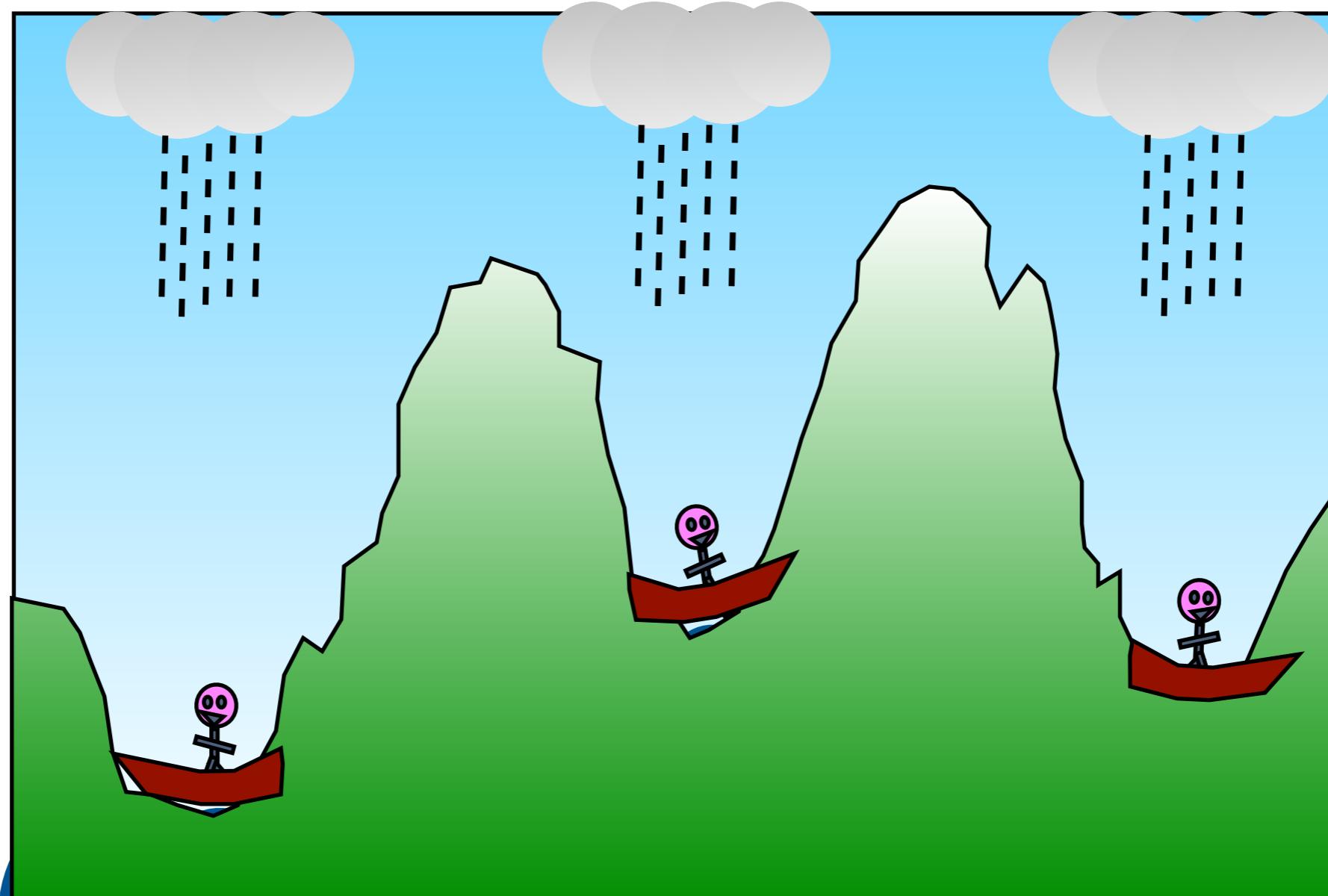
- A. Barducci, G. Bussi, M. Parrinello, Well-tempered metadynamics: A smoothly converging and tunable free-energy method, Phys. Rev. Lett. 100 (2008) 020603.

- 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

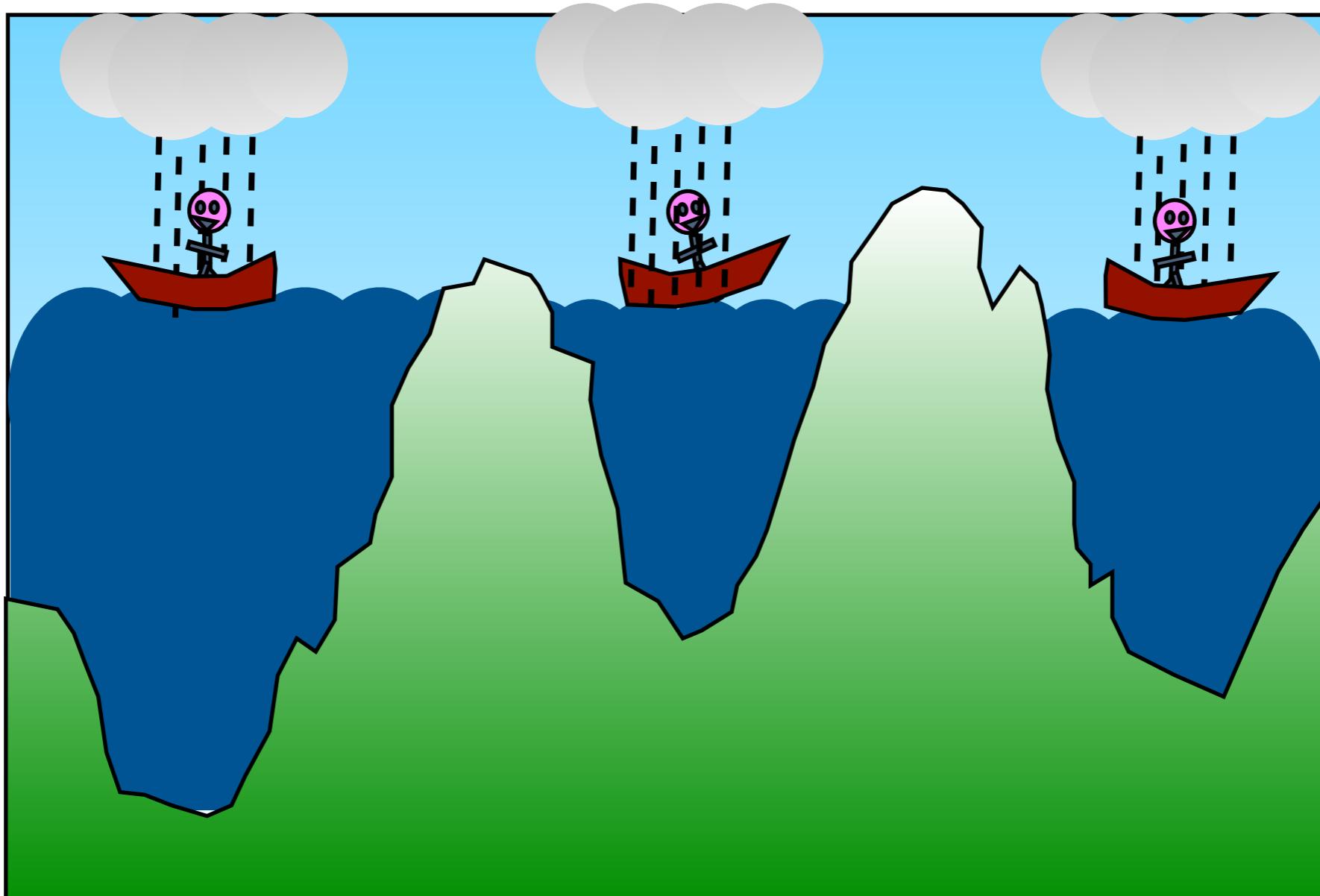
- several metadynamics simulations running
- sharing and adding to a common V_{bias} potential (runs “feel” each other)
- runs “repel” each other and explore different regions
- trivial speedup and scaling



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



Contents

- Rare event simulation
- The metadynamics method
- Path-metadynamics
- Multi paths, multiple walkers
- Combined with machine learning to find CVs
- Exercises

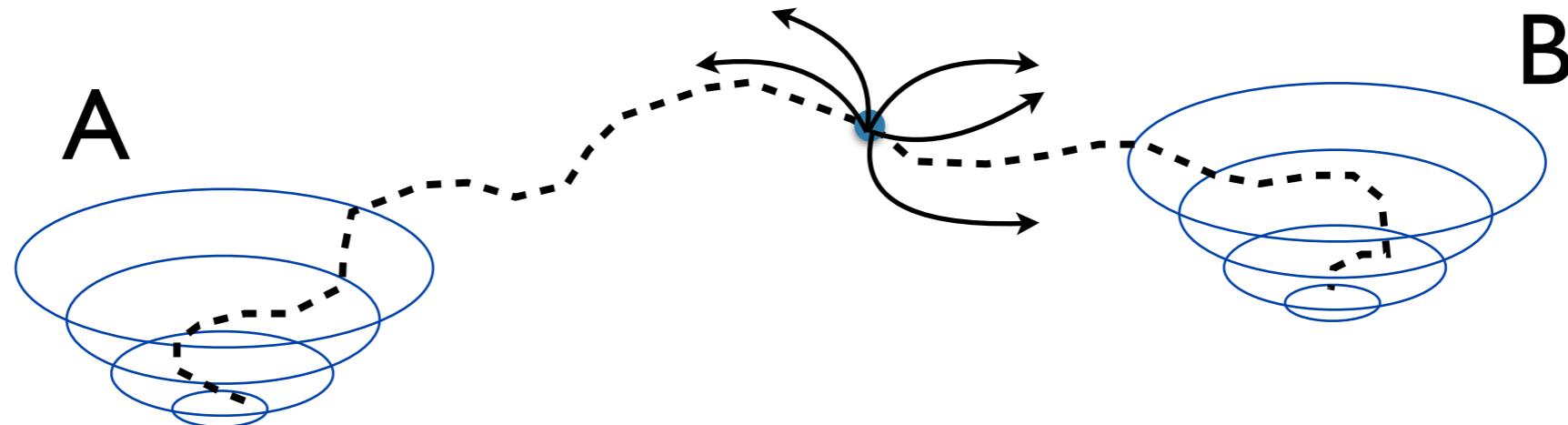
path-metadynamics

Dealing with many collective variables

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

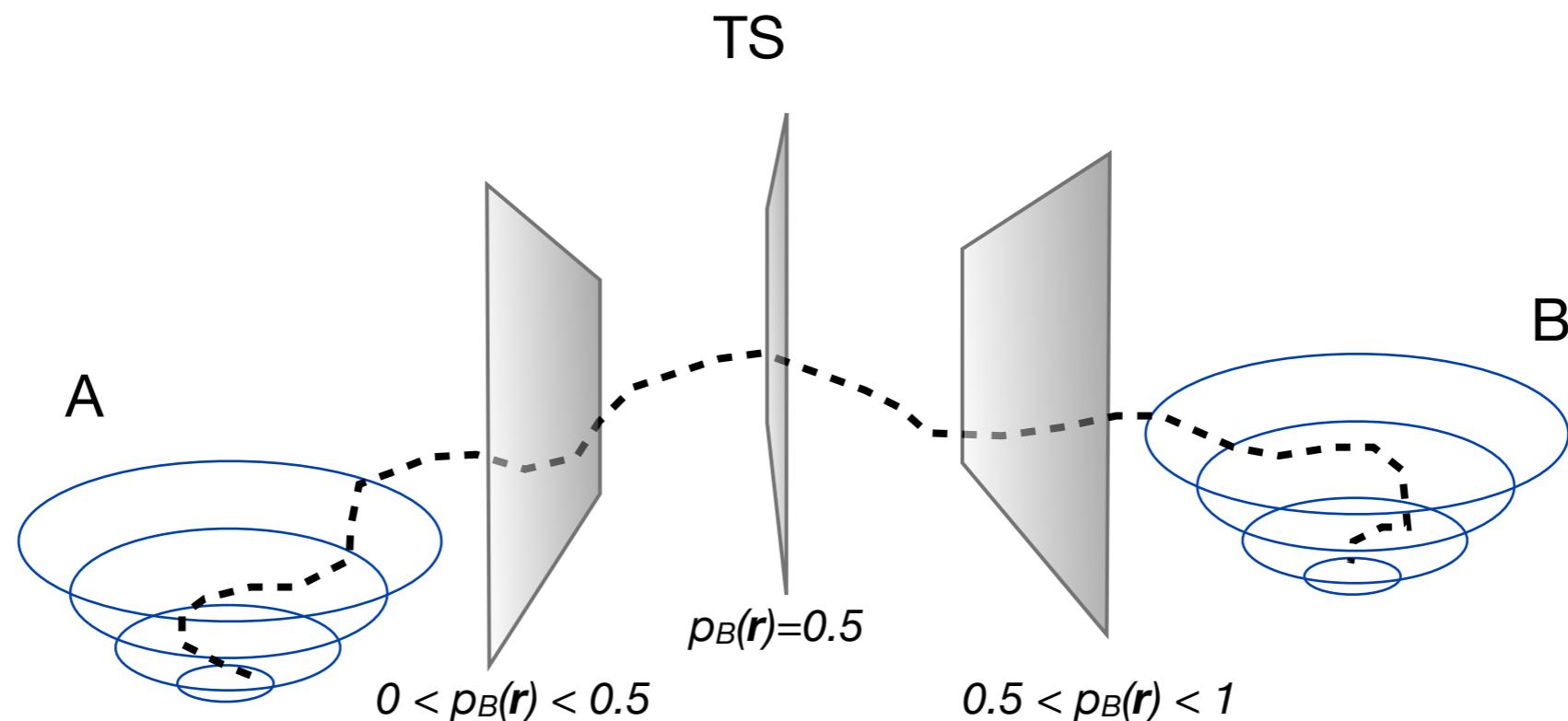
Committer probability



A transition state is:
The configuration with equal probability
to go (commit) to states A or B

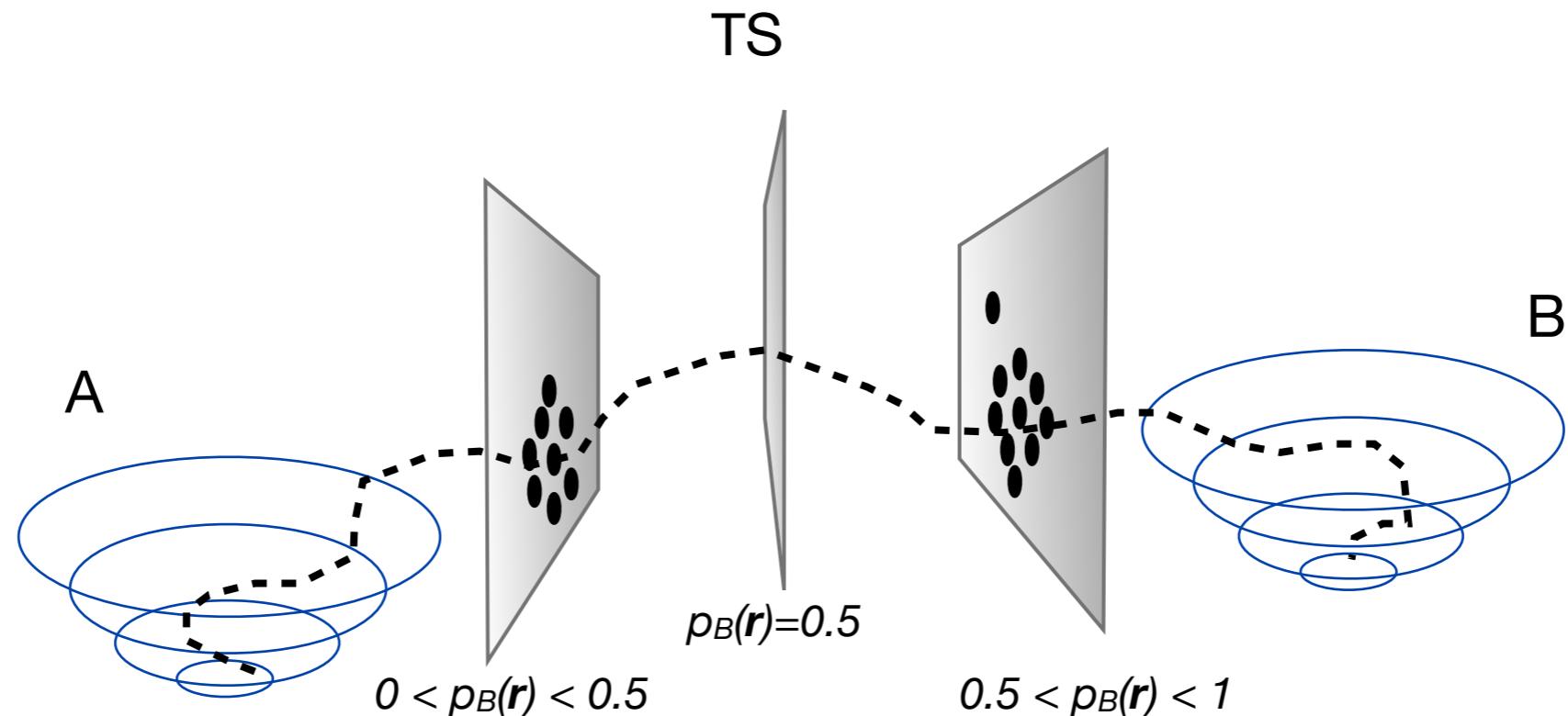
path-metadynamics

iso-committor surfaces



path-metadynamics

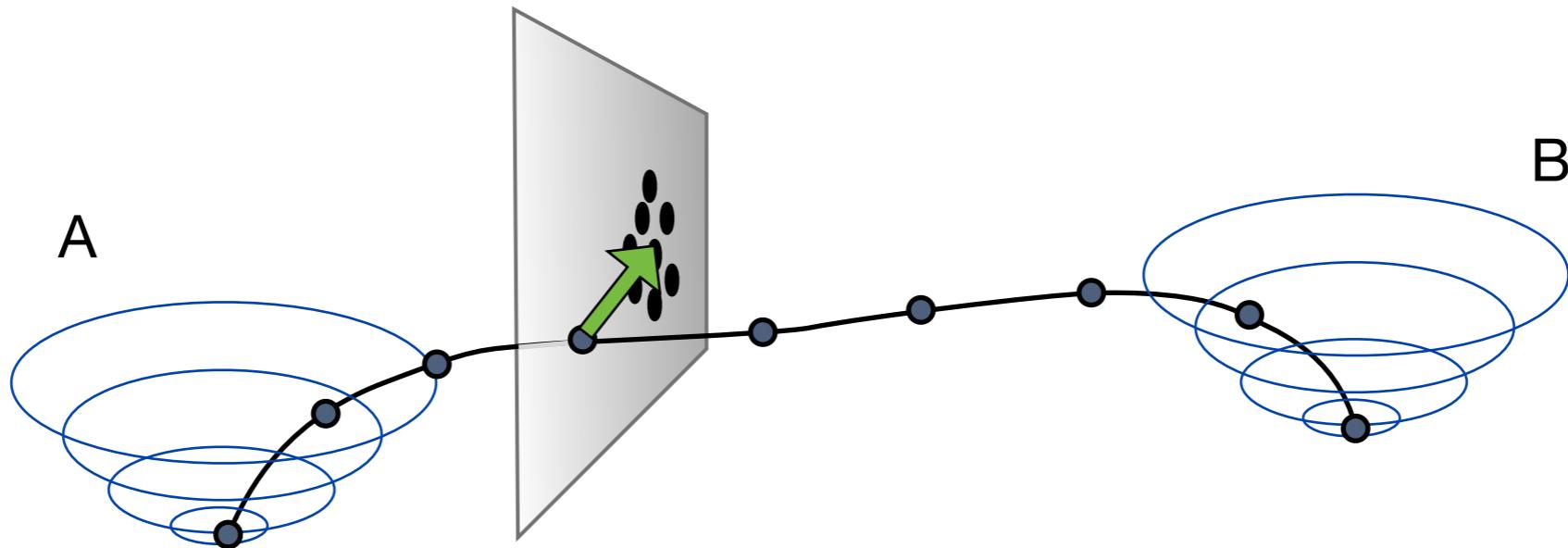
iso-committor surfaces



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

practical implementation

leverage measurement of d
between closest nodes

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

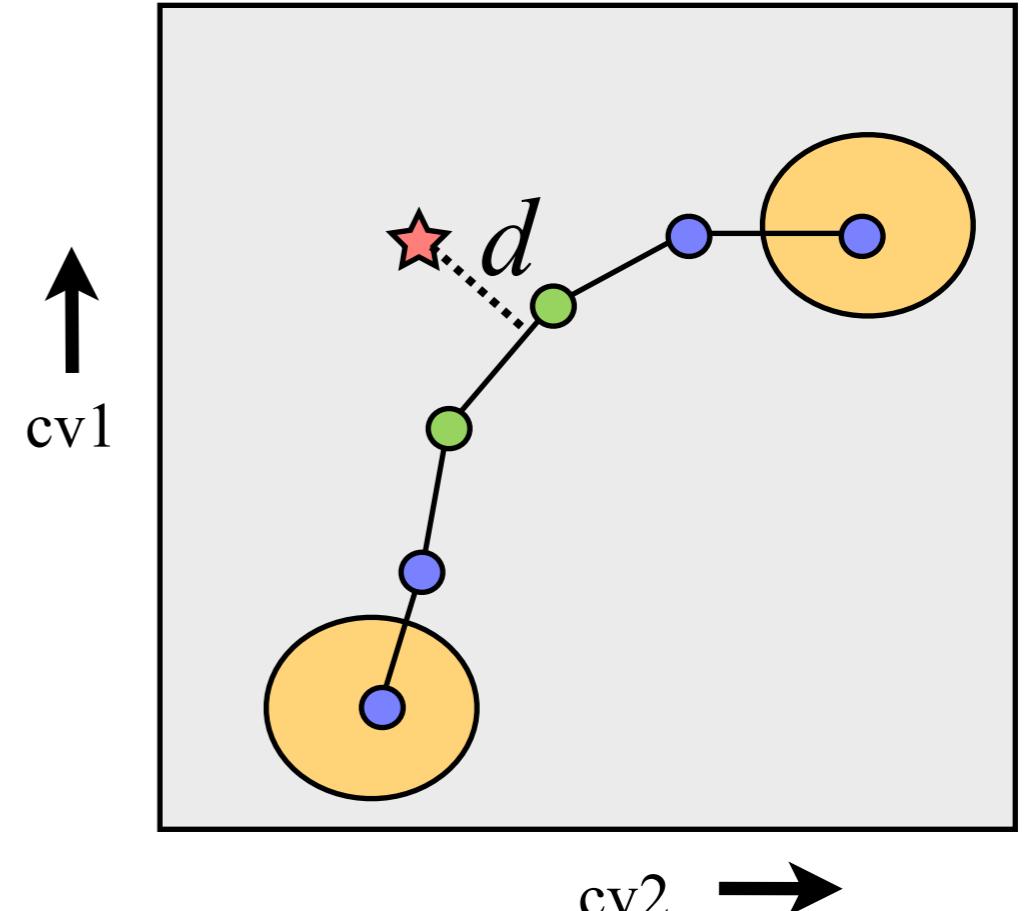
$$w_k = \max \left(0, \left(1 - \frac{|\mathbf{s}_j^{t_i} - \mathbf{s}^{t_i}(\sigma(\mathbf{z}_k))|}{|\mathbf{s}_j^{t_i} - \mathbf{s}_{j+1}^{t_i}|} \right) \right)$$

path update

move nodes to d (every step)

set $d=0$

redistribute nodes along path



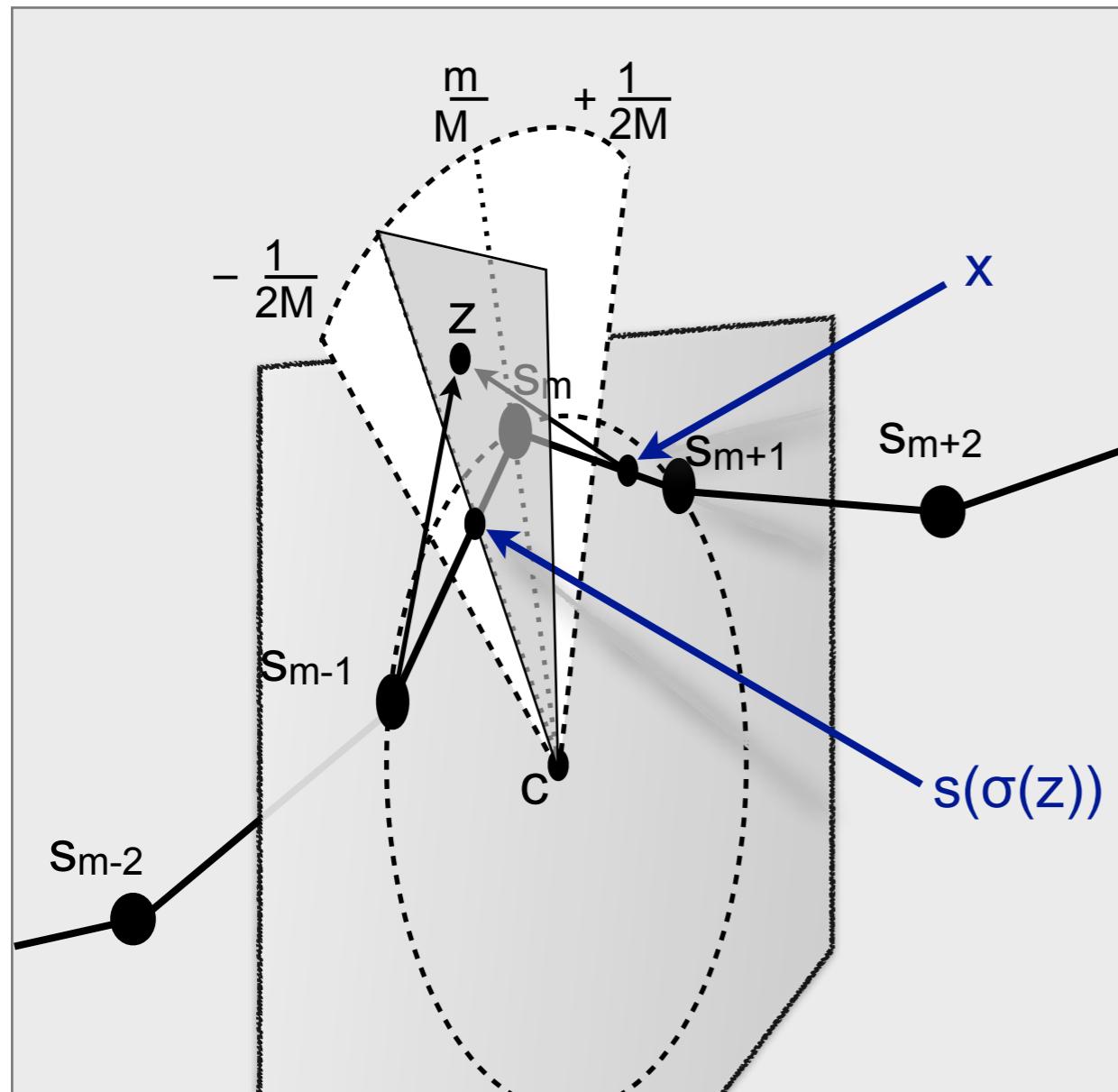
half life time of data

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

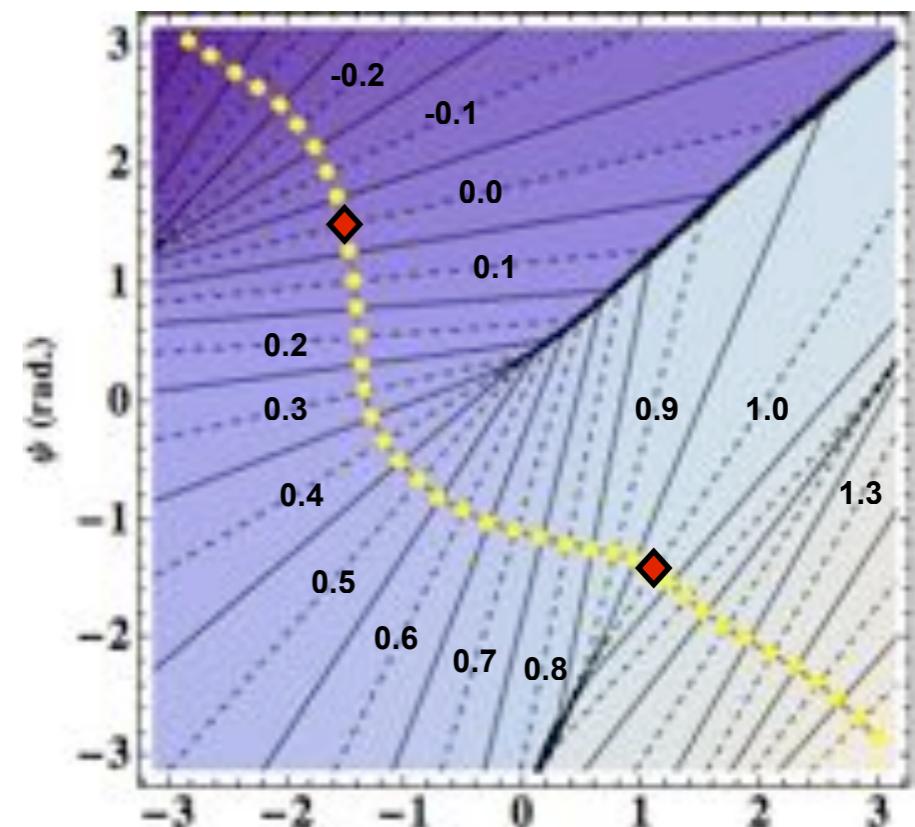
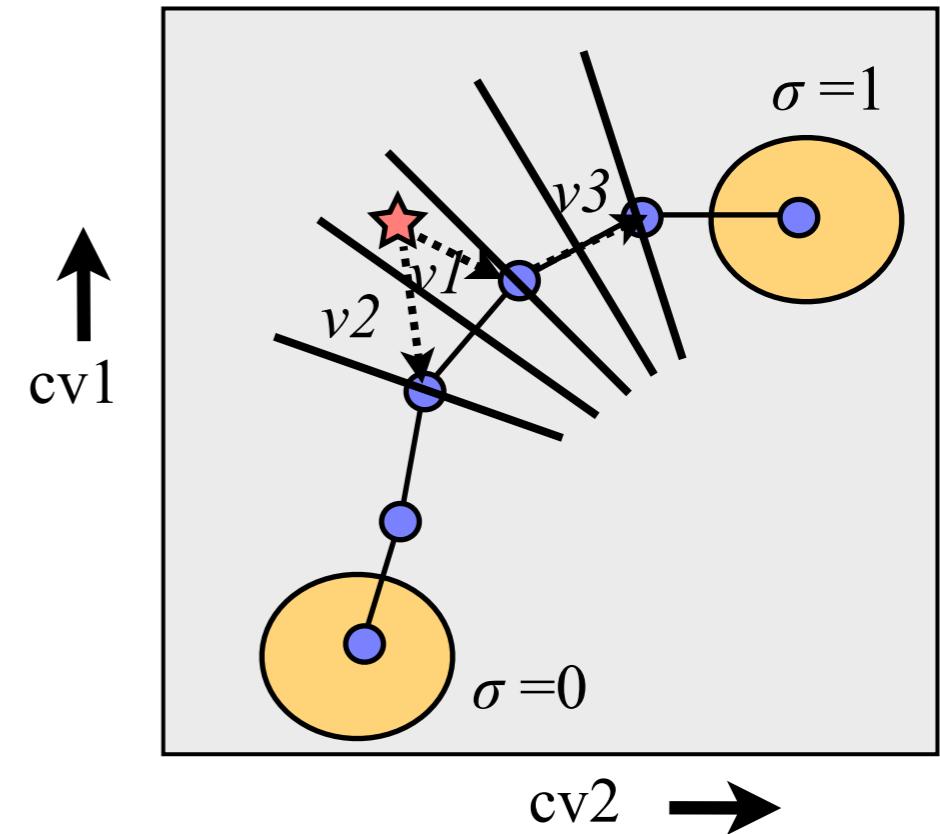
optional tube potential

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

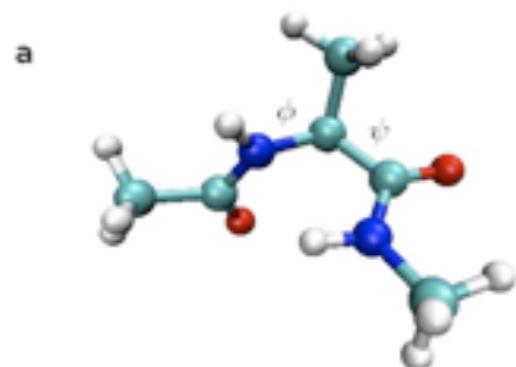
Projecting on a string of nodes



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



alanine dipeptide



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

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

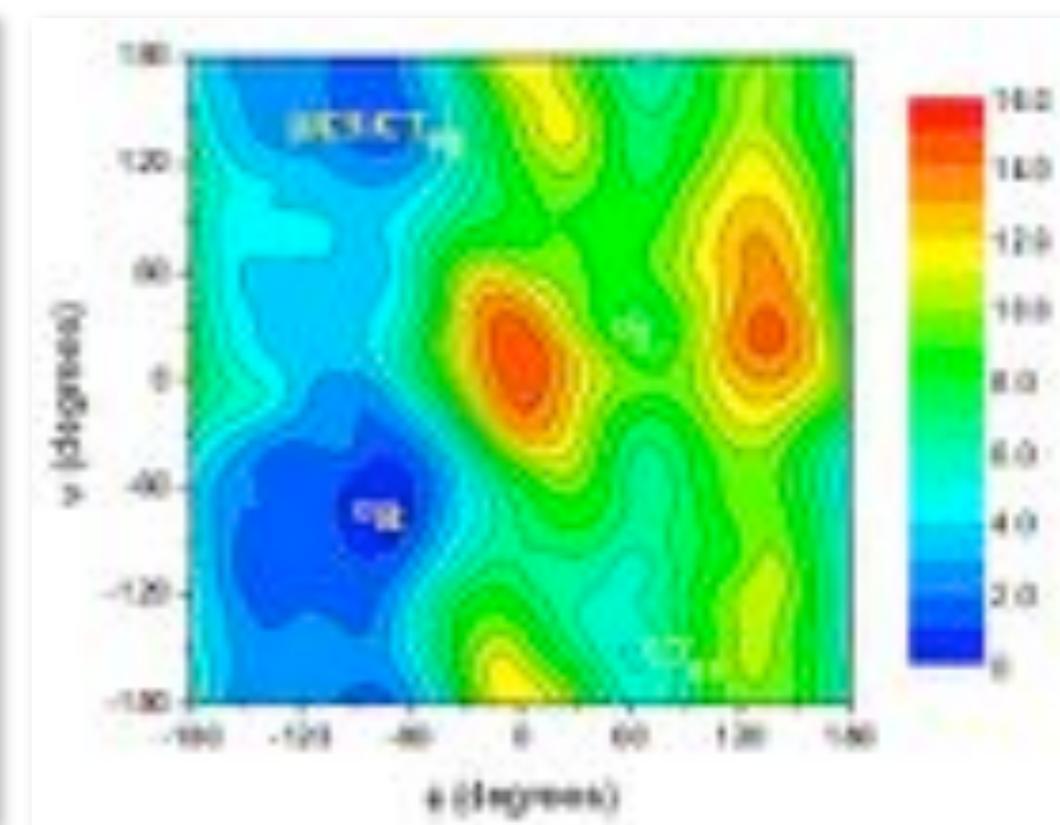
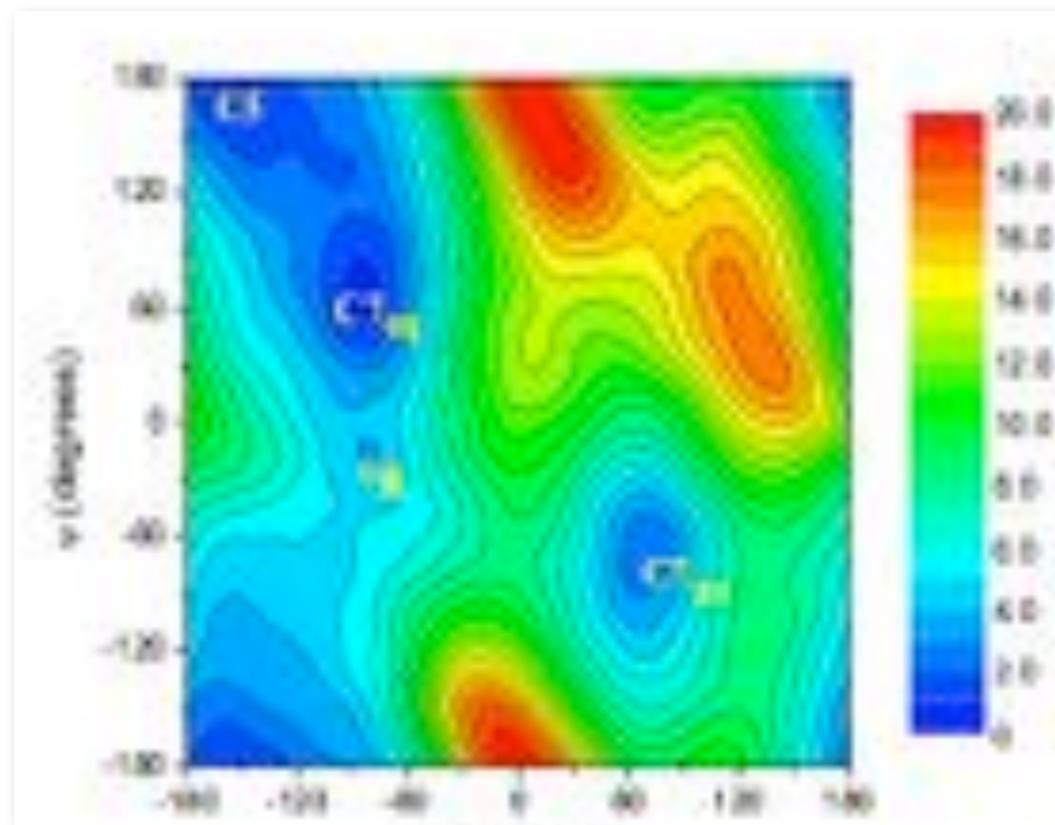


Alanine dipeptide + 216 water
CHARMM27 forcefield (modified internal params water)
Cubic periodic box L=18.8 Angstrom
NVT ensemble, T=298K

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

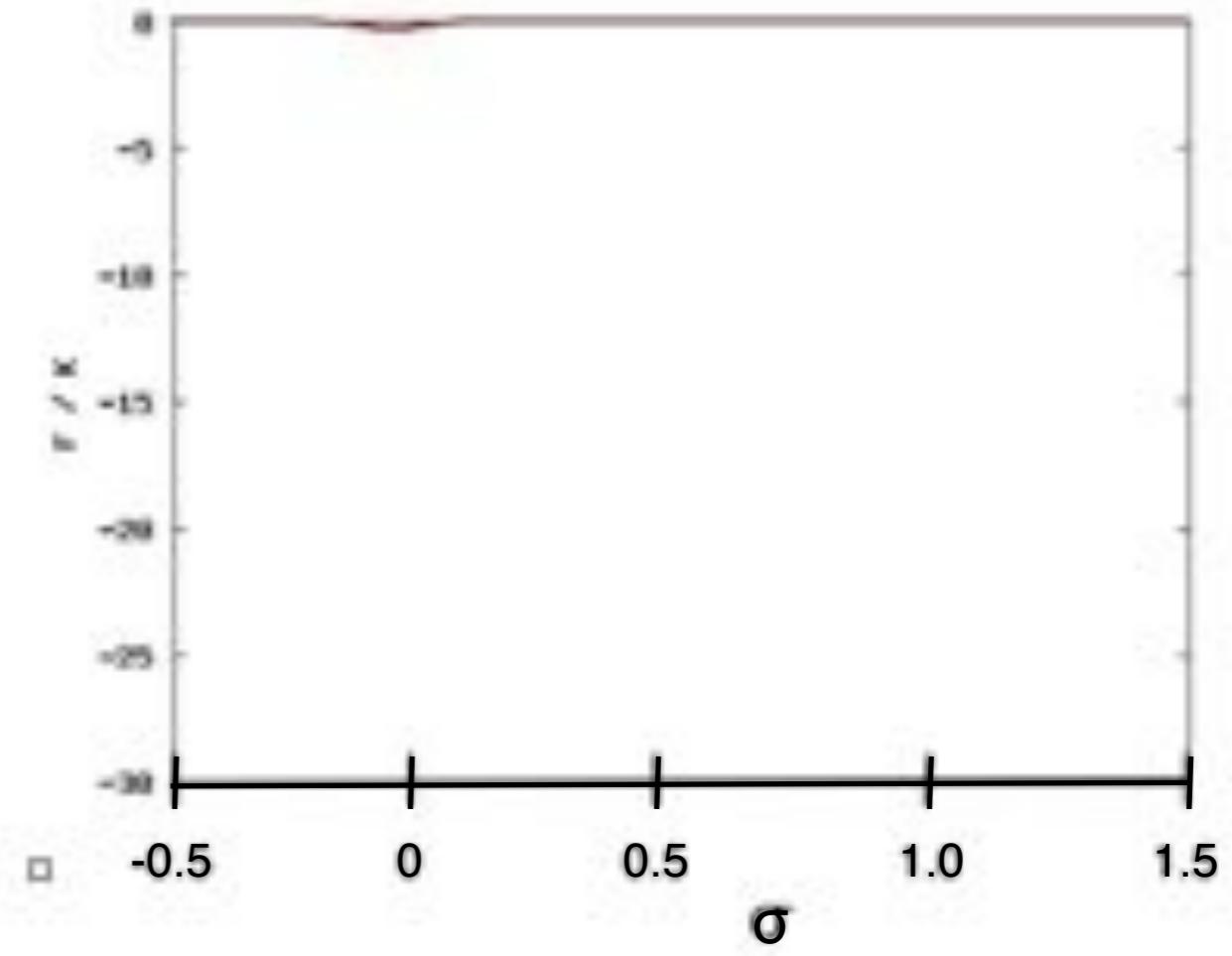
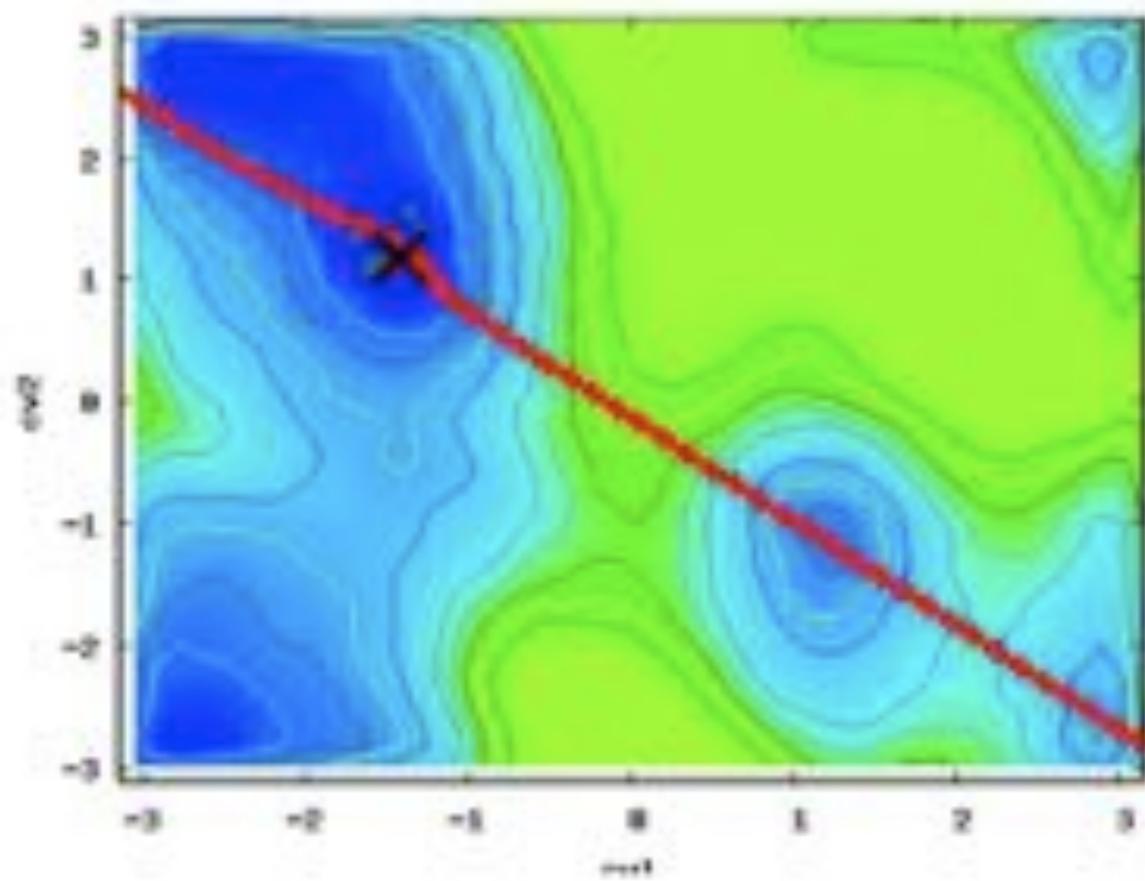
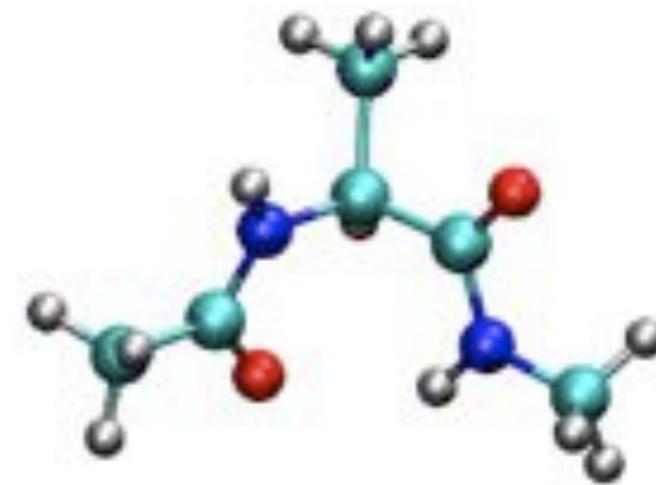
alanine dipeptide

Parameters:

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

every recrossing:

$H \times 50\%$
 $W \times 50\%$
 $\tau \times 100$



Path Finding on High-Dimensional Free Energy Landscapes.

Grisell Díaz Leines and Bernd Eising

Phys. Rev. Lett. **109** (2012), 020601

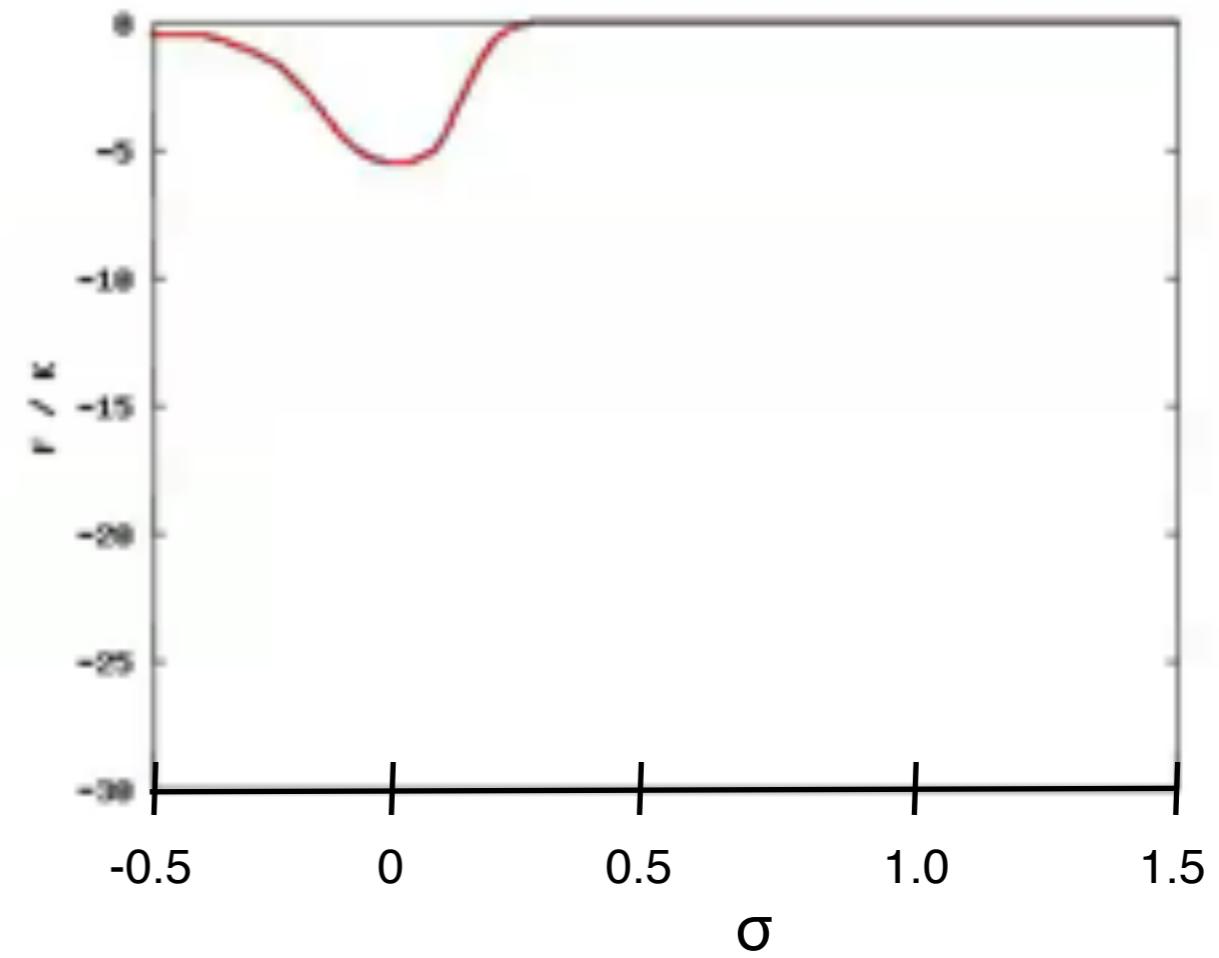
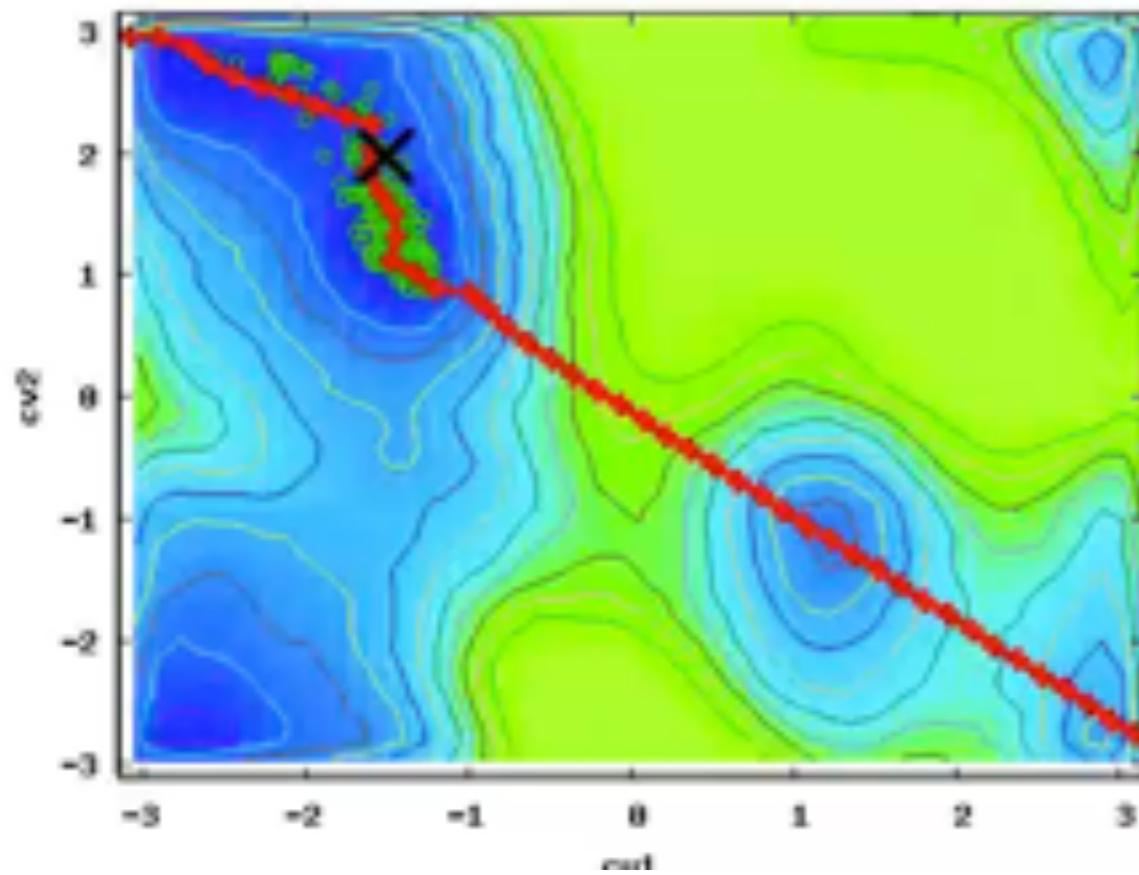
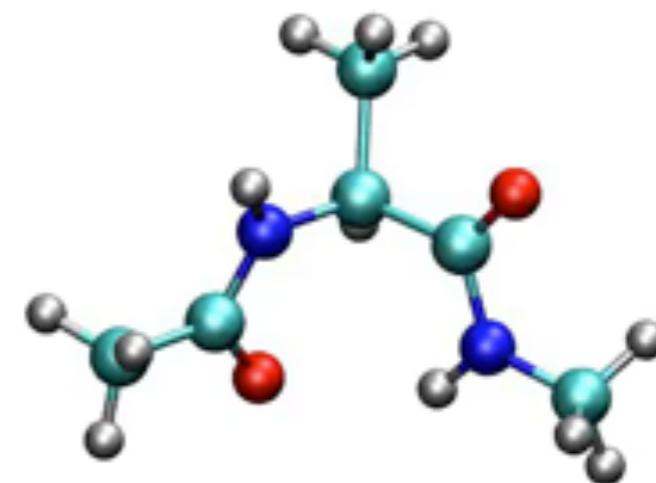
alanine dipeptide

Parameters:

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

every recrossing:

$H \times 50\%$
 $W \times 50\%$
 $\tau \times 100$



Path Finding on High-Dimensional Free Energy Landscapes.

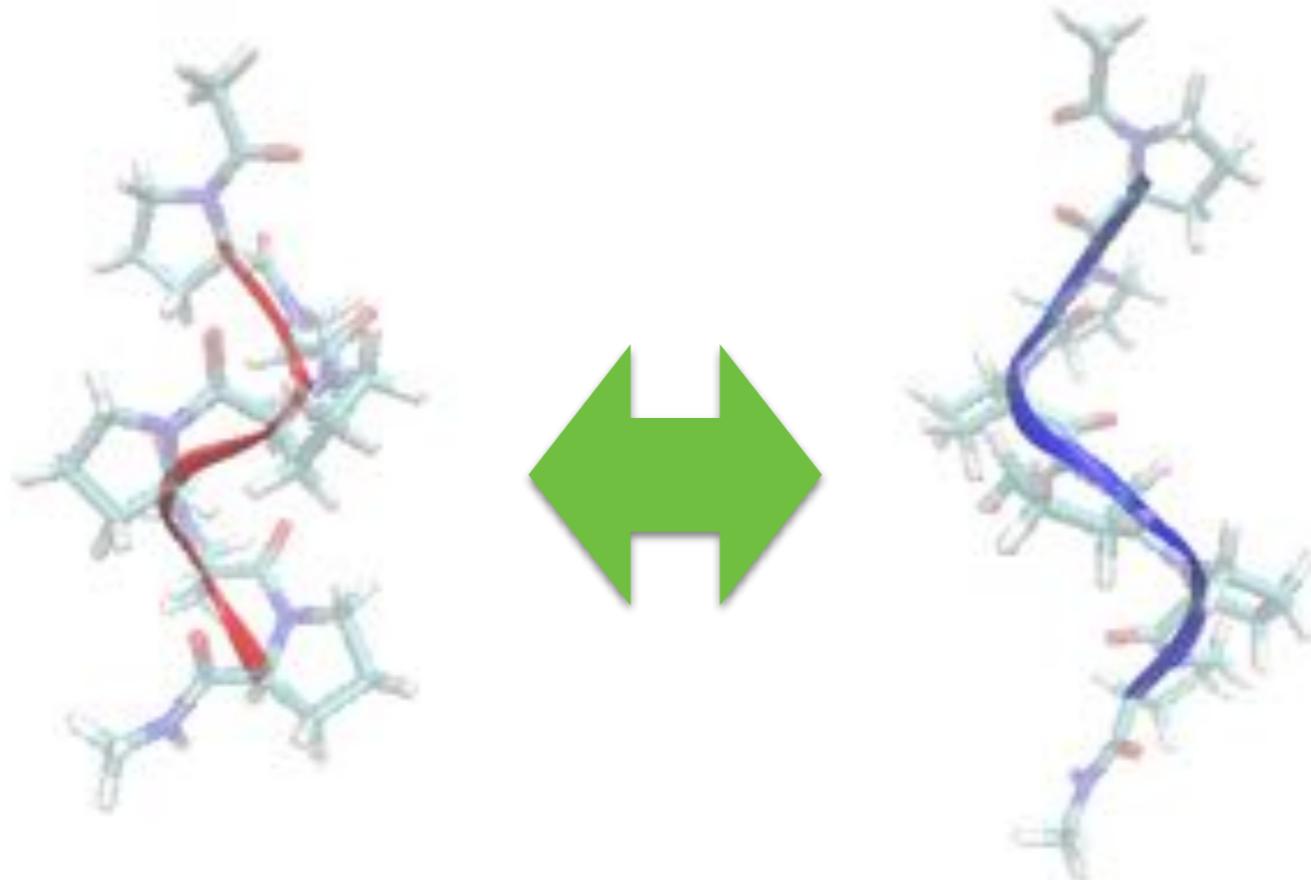
Grisell Díaz Leines and Bernd Eising

Phys. Rev. Lett. **109** (2012), 020601

Contents

- Rare event simulation
- The metadynamics method
- Path-metadynamics
- Multi paths, multiple walkers
- Combined with machine learning to find CVs
- Exercises

Polyproline oligomers



PMD Settings:

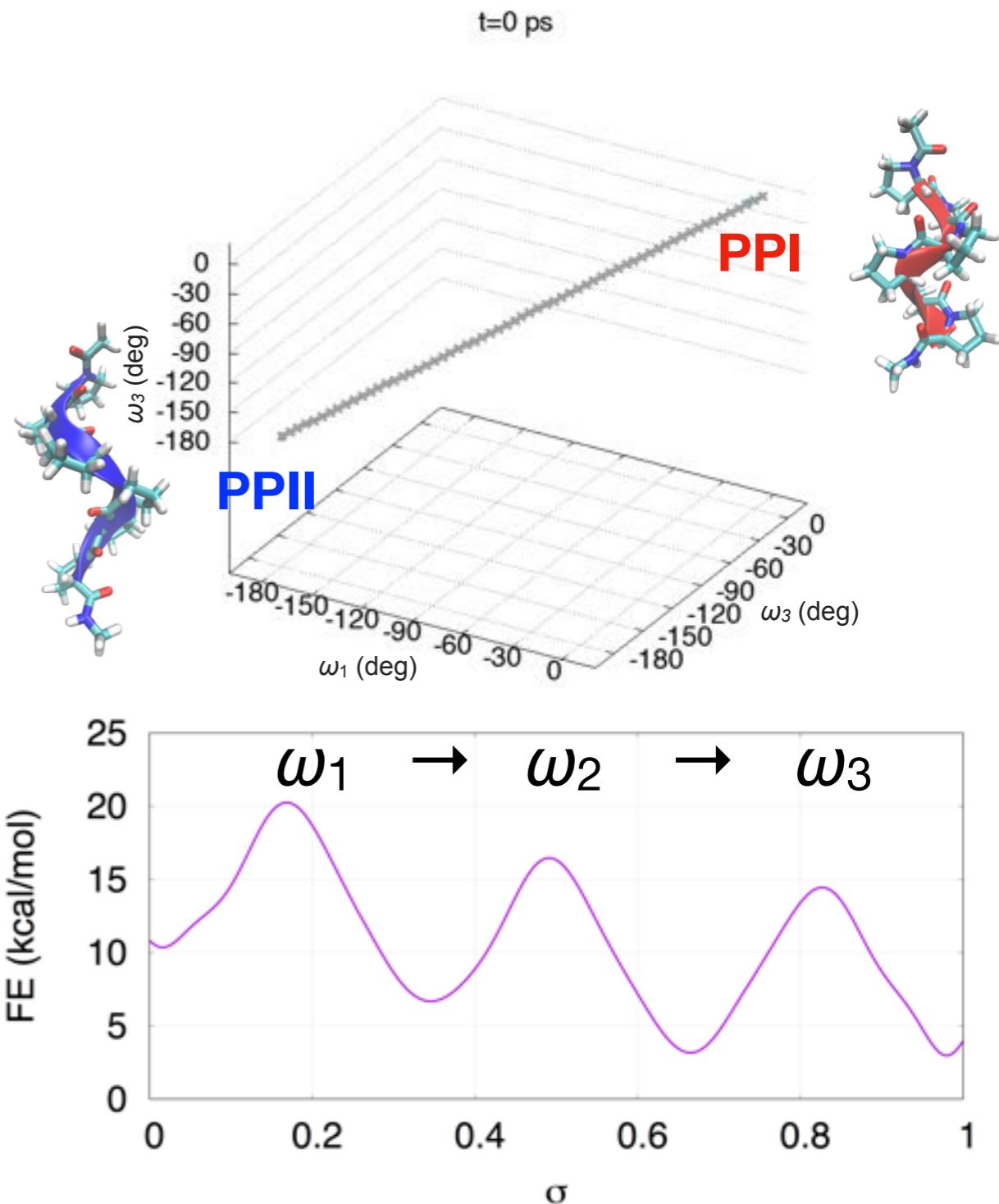
- 20 path nodes per CV + ends
- path update: 0.5 ps
- half live: 2.0 ps
- tube + wall potentials

Metadynamics settings:

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

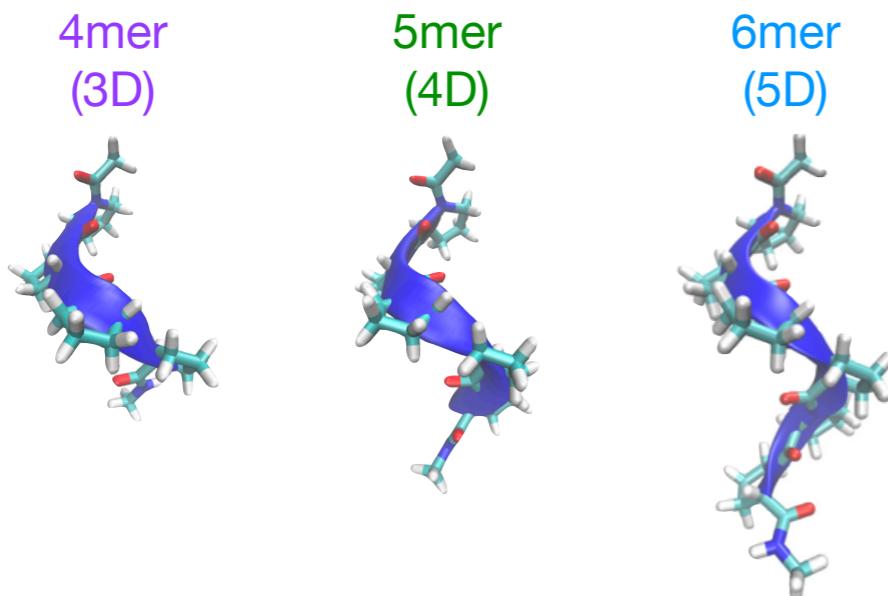
Multiple-walker PMD

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

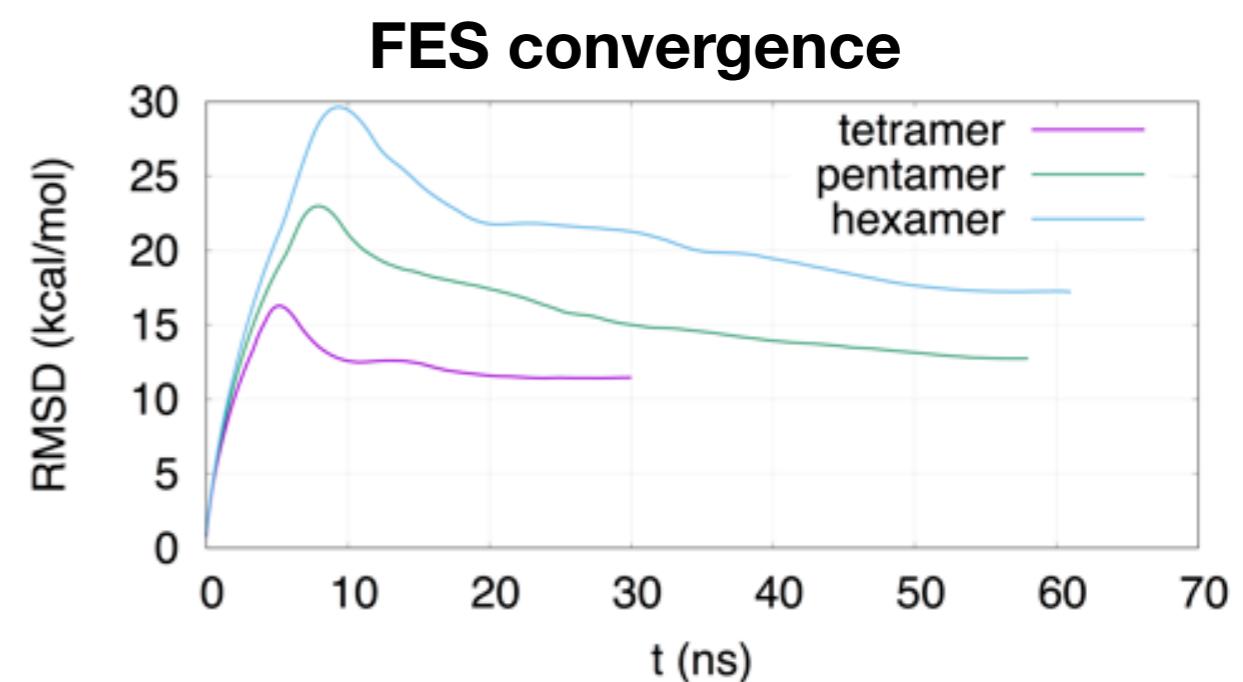
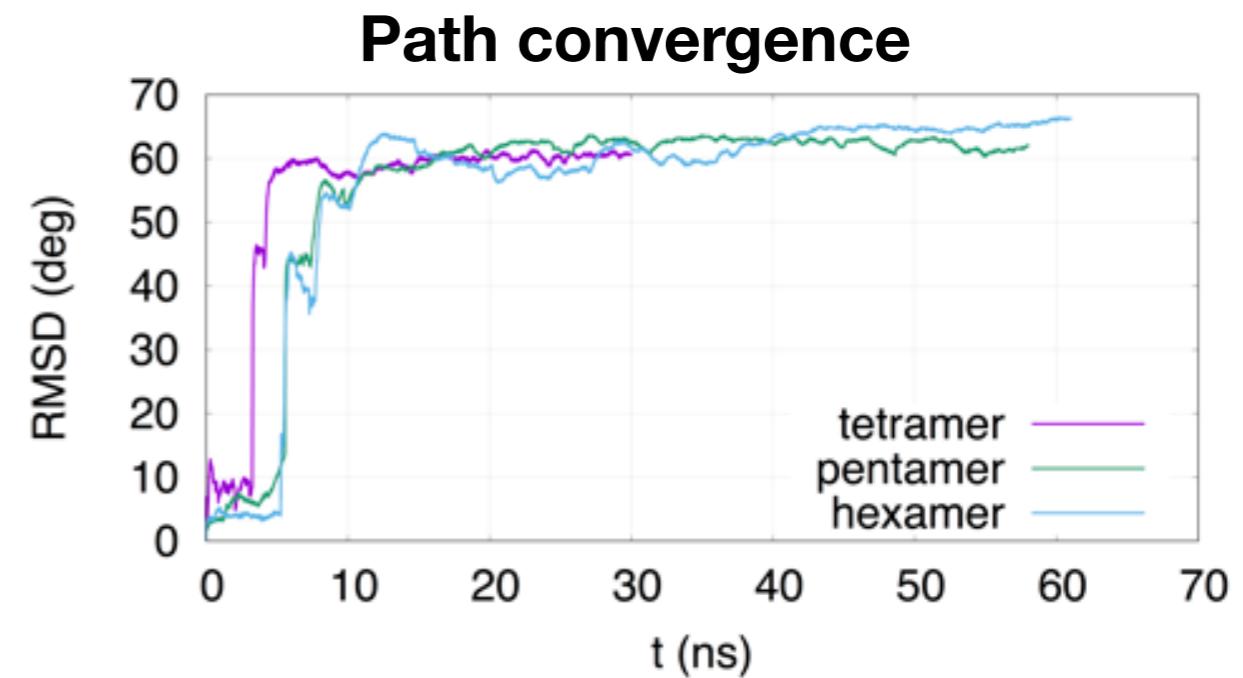


PMD performance

- 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

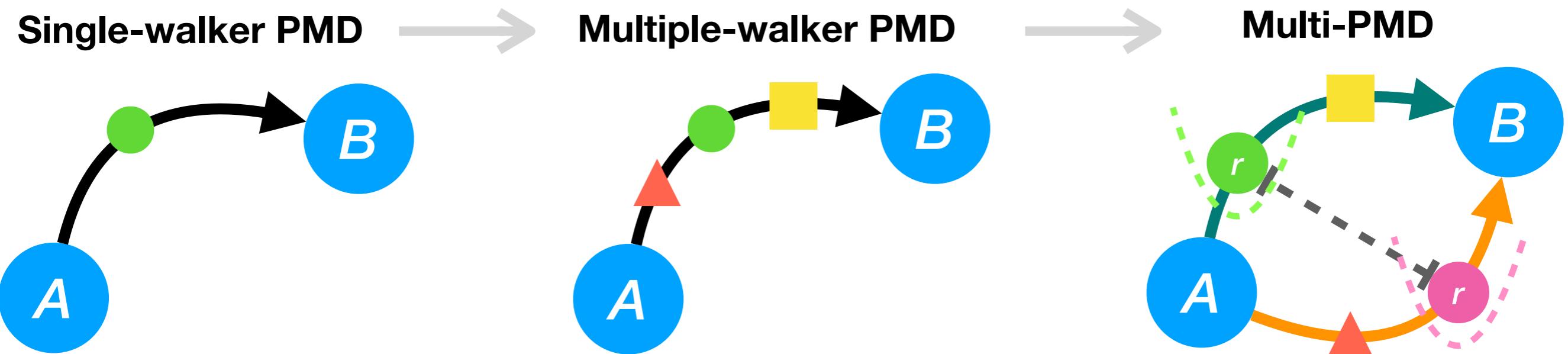


- But the number of possible paths grows **factorially** with CV-dimensionality!

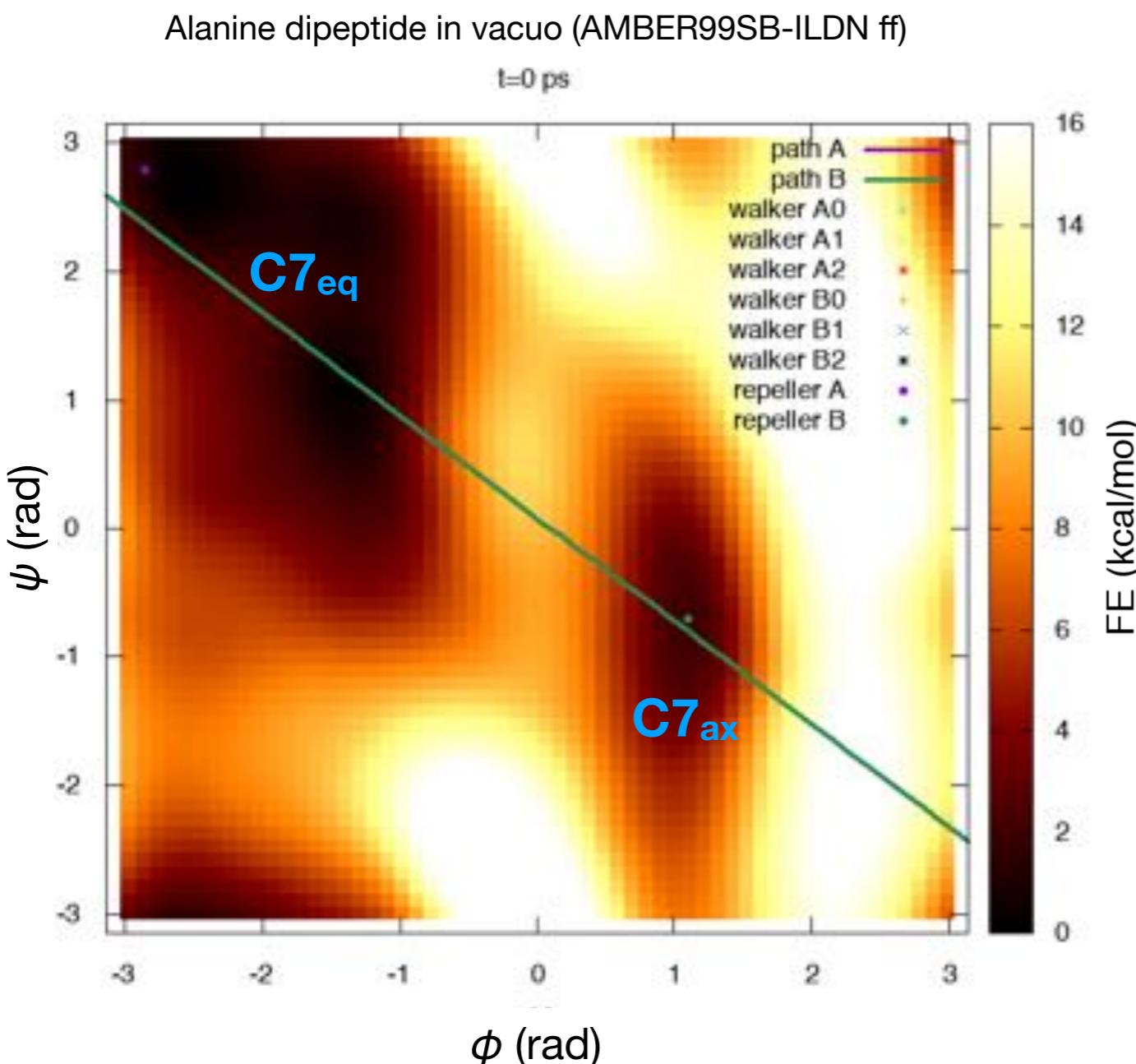


MuWaMuPaMetaDyn

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



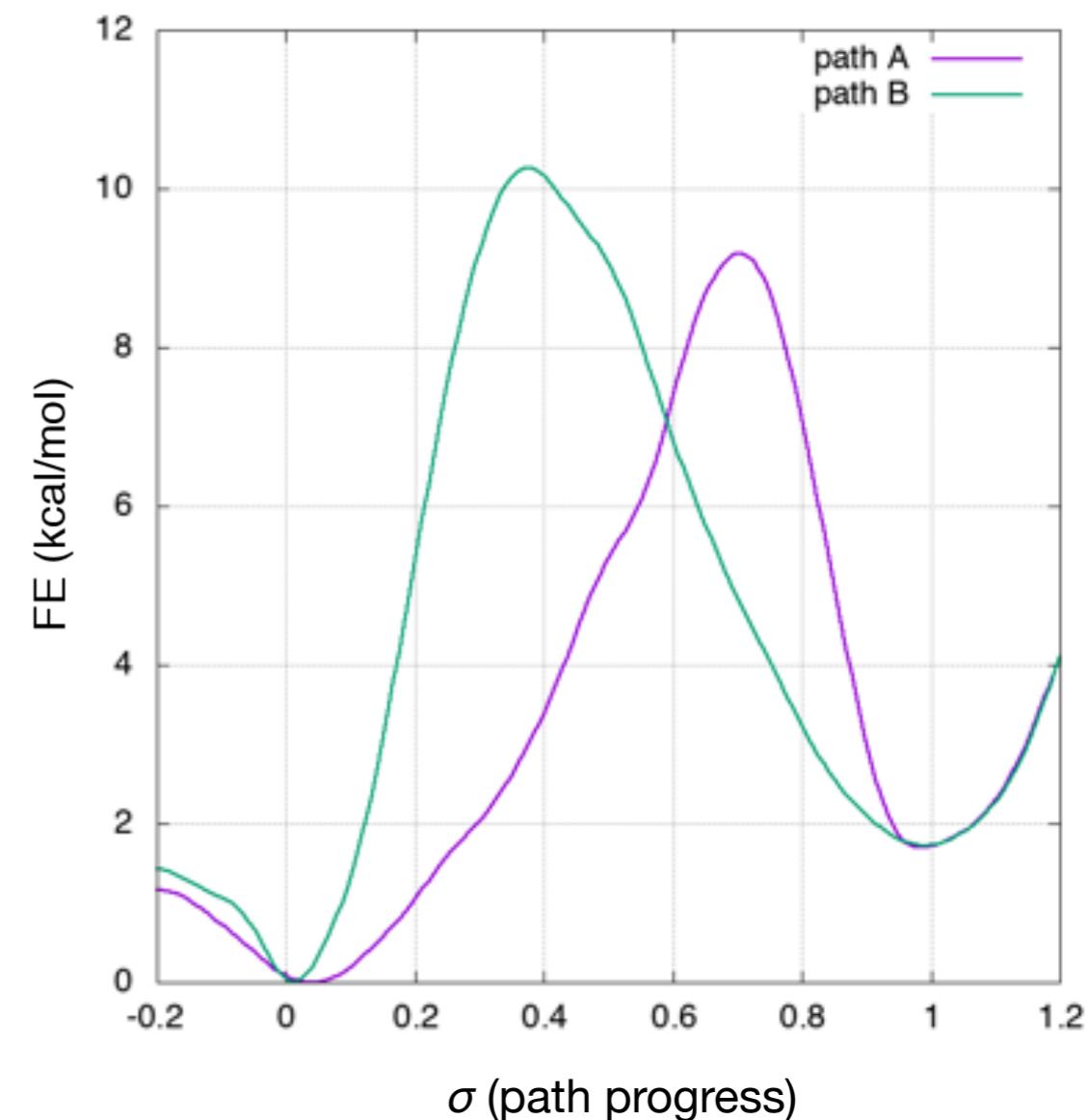
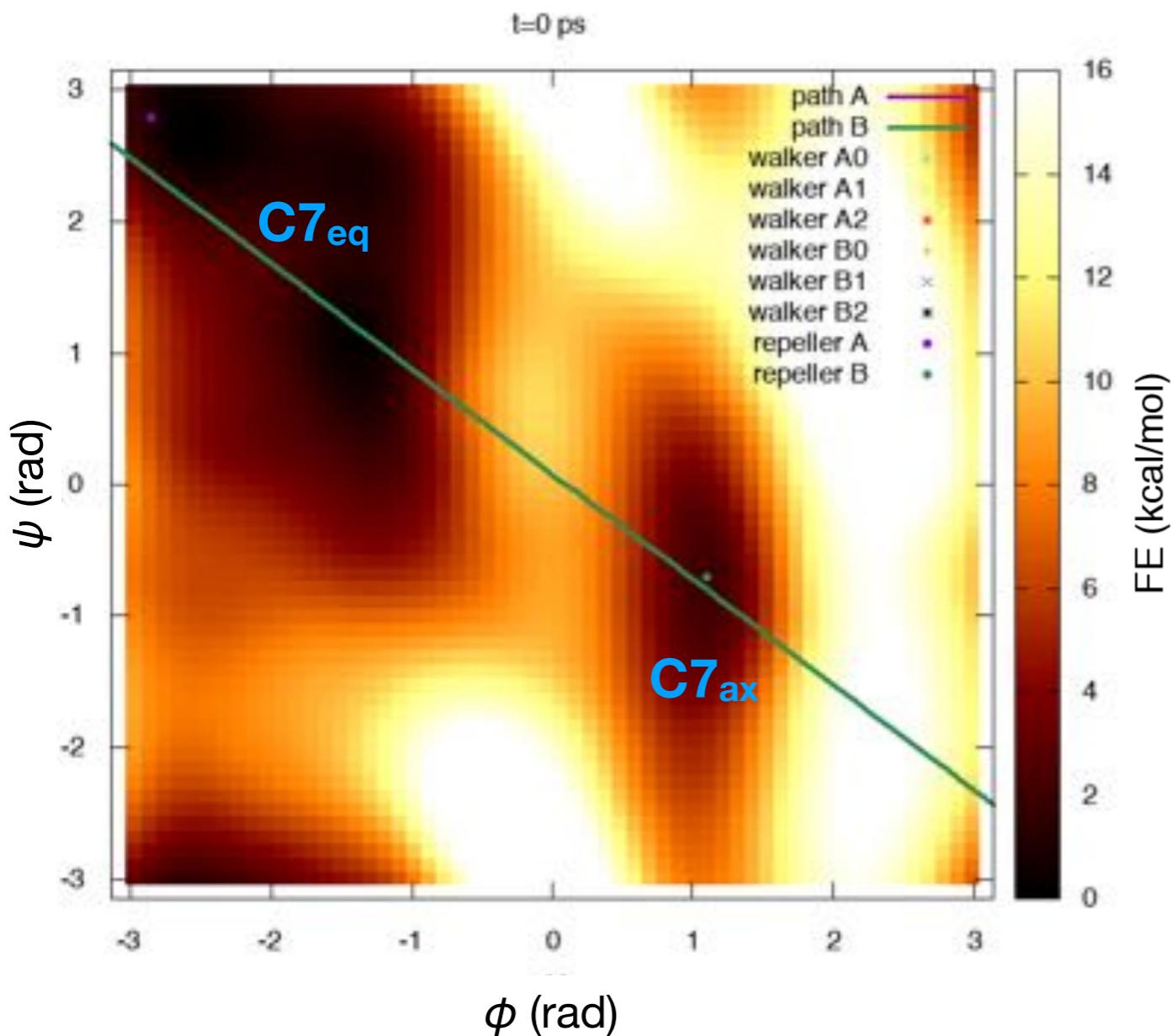
Alanine dipeptide revisited



- Two $C7_{eq} \rightarrow C7_{ax}$ paths
- 3 walkers & 1 repeller per path
- Repellers steered to $\sigma = 0.5$
- Min. distance² between repellers restrained to 0.2 rad

Alanine dipeptide revisited

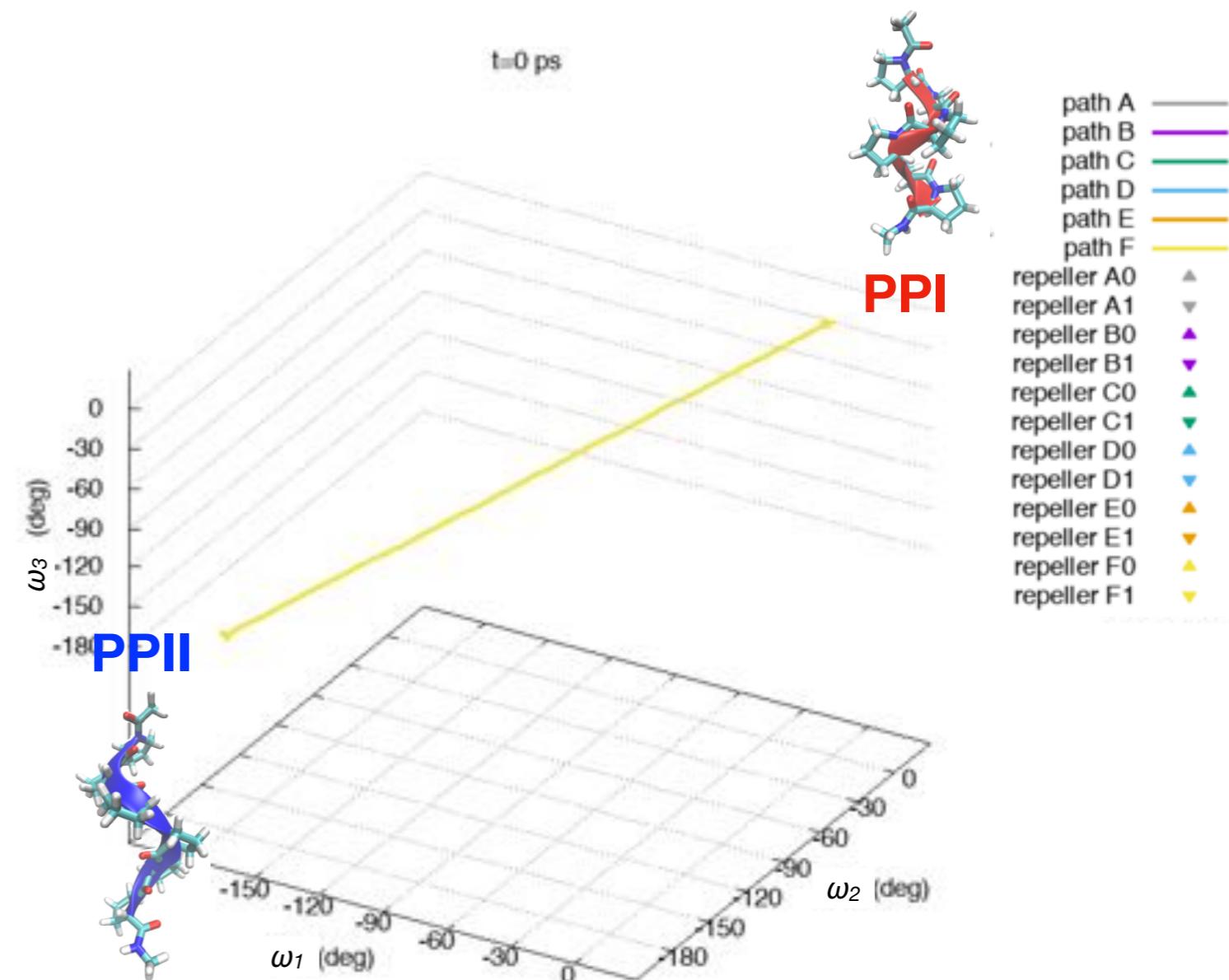
Alanine dipeptide in vacuo (AMBER99SB-ILDN ff)



MuWaMuPaMetaDyn

Multiple-Walker Multiple-Path Metadynamics

- Six PPI \rightarrow 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





PLUMED-NEST

The public repository of the PLUMED consortium

[Home](#)[News](#)[PLUMED](#)[Consortium](#)[Contribute](#)[Cite](#)[Browse](#)

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

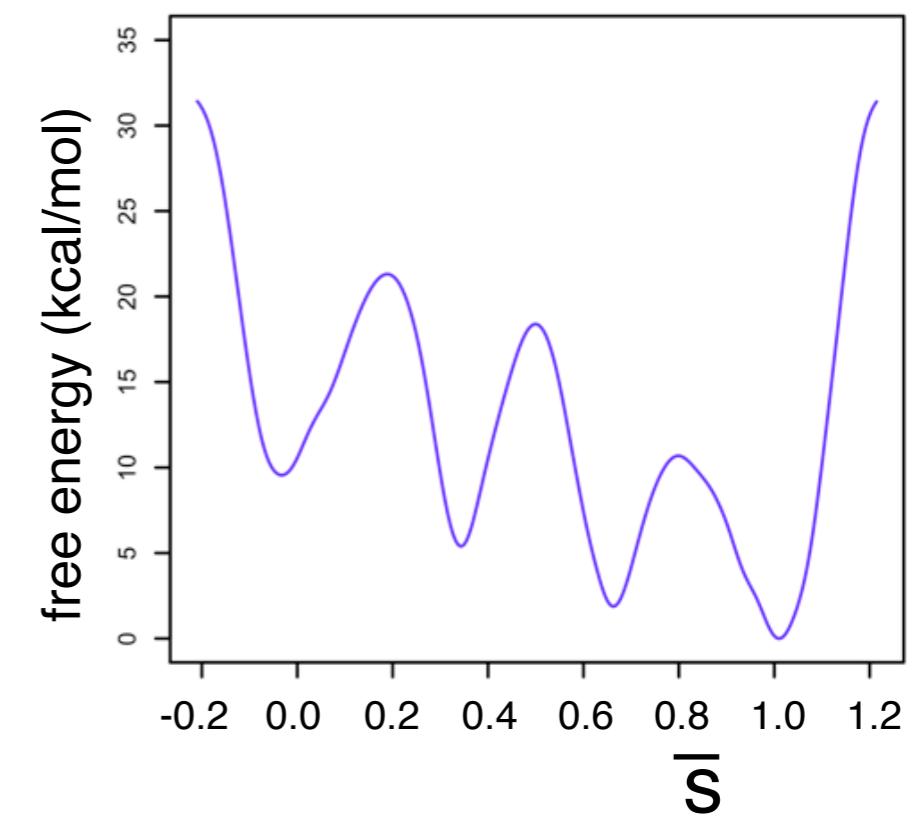
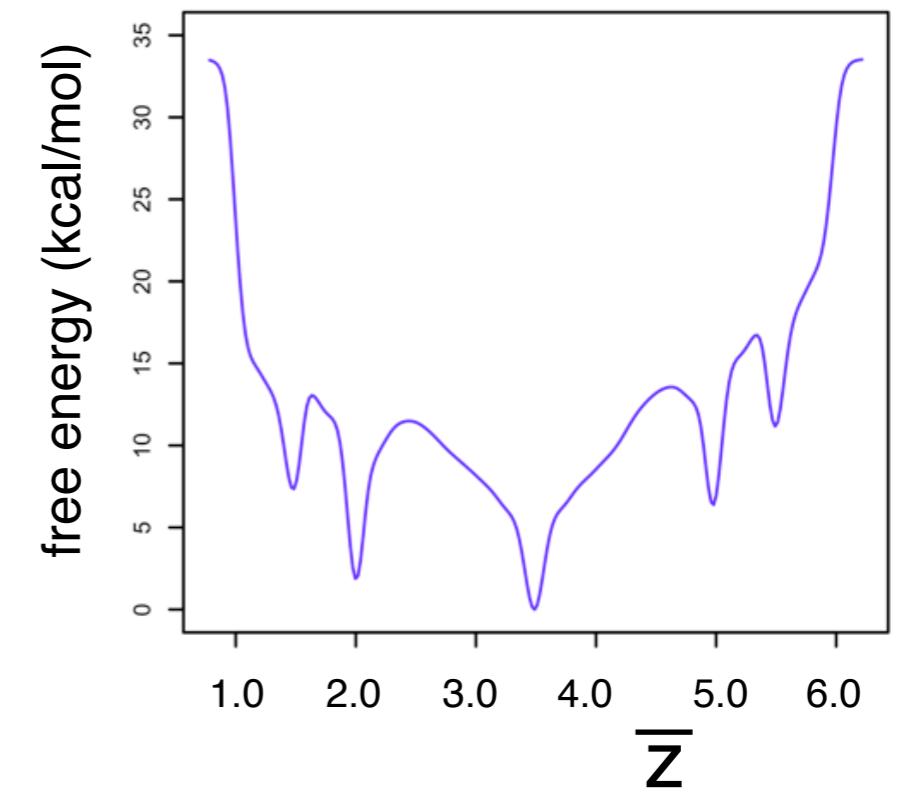
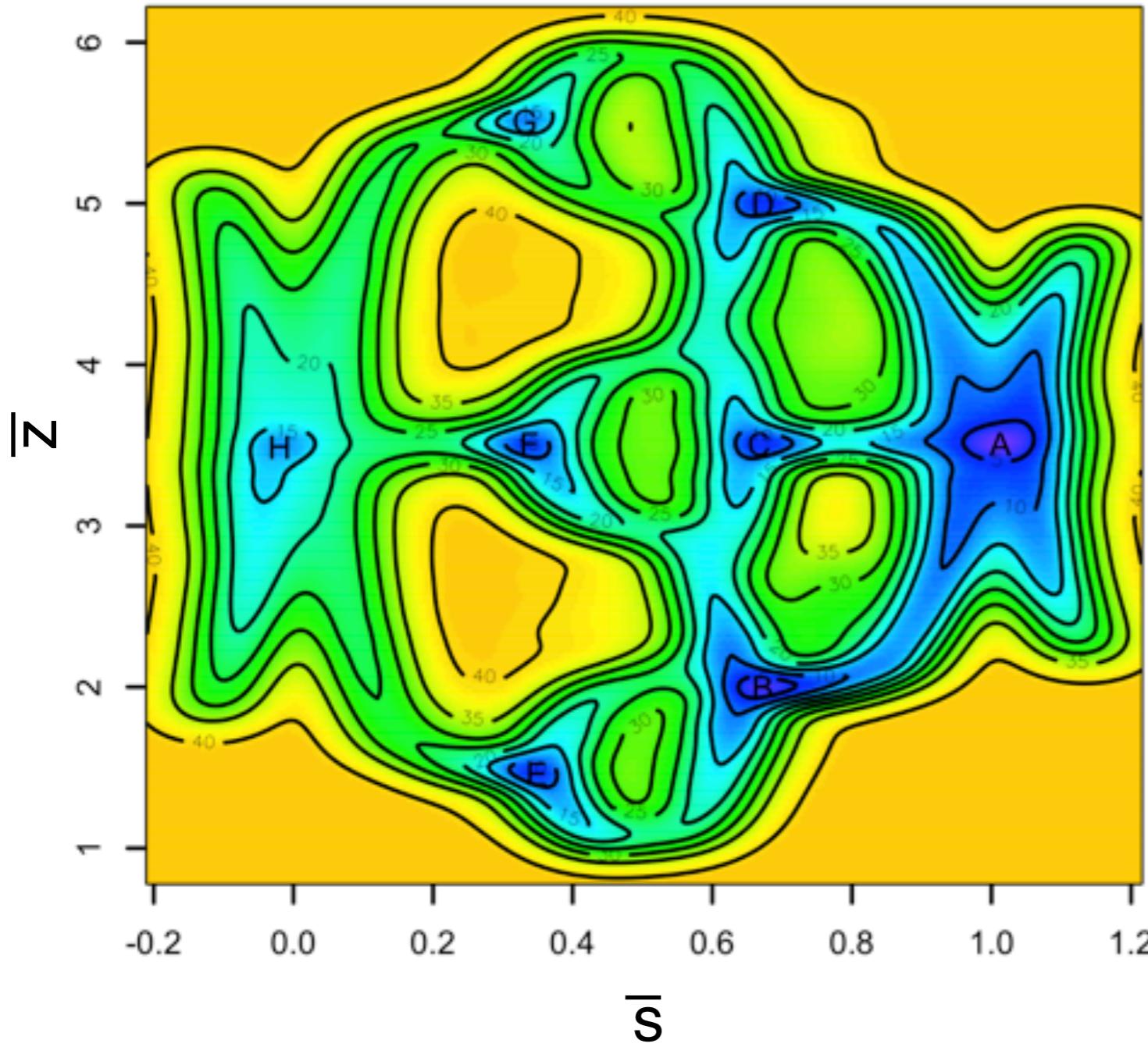
Last tested: 19 Jul 2019, 15:40:32

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.

Making a PathMap

projections



Contents

- Rare event simulation
- The metadynamics method
- Path-metadynamics
- Multi paths, multiple walkers
- Combined with machine learning to find CVs
- Exercises

Recipe for using a Path-CV

- Characterize the stable states
- Choose CVs to use as path-CV arguments
- Choose initial guess path
 - linear interpolation
 - construct by hand, use plumed driver on previous trajectory
 - choose number of nodes, trailing nodes
- Choose biasing method (umbrella sampling, metadynamics, OPES, ...)
- Multiple walkers?
- Monitor path evolution, convergence
- Monitor free energy profile evolution, convergence
- Need for tube potential?

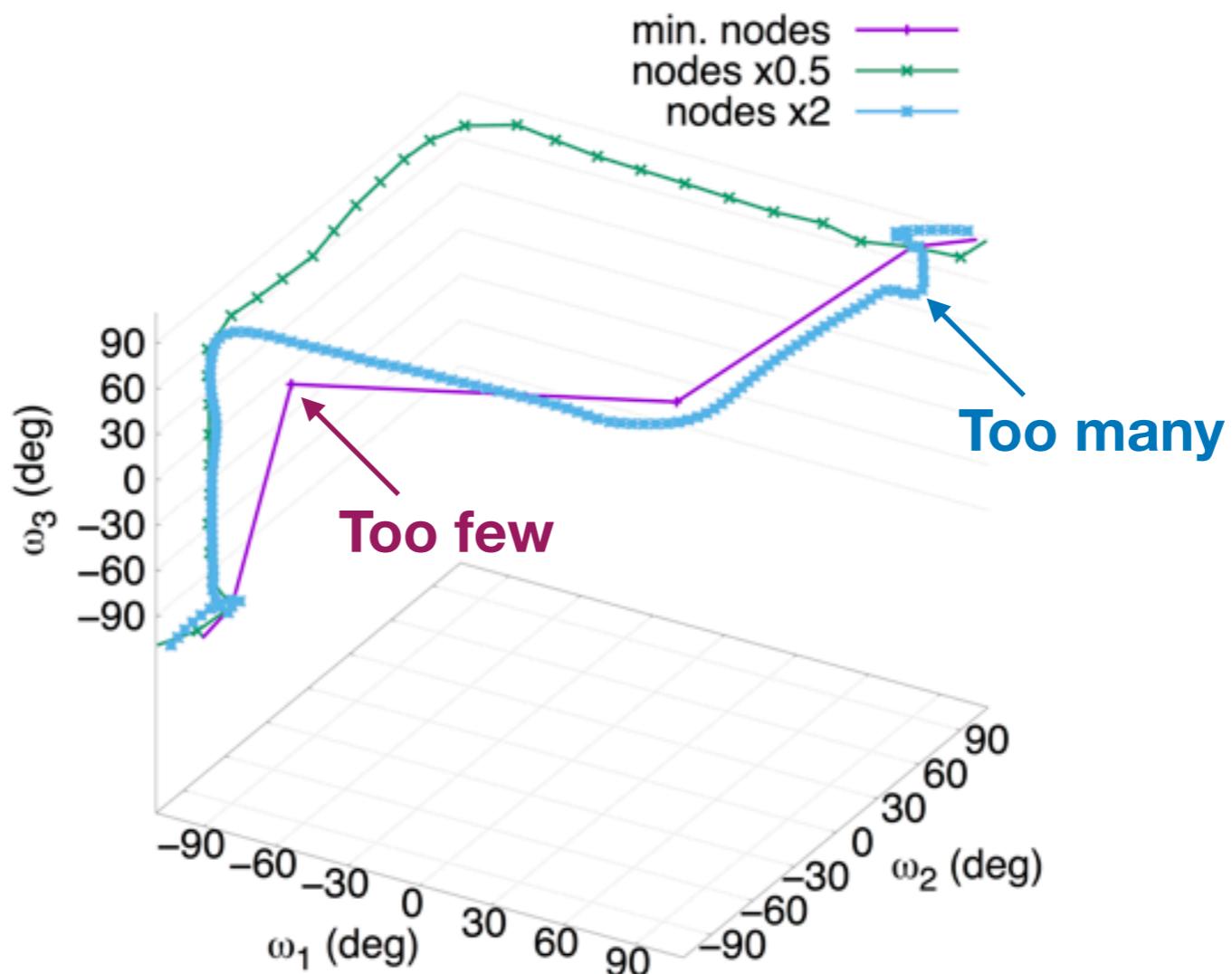
Which biasing method?

- Metadynamics (path-metadynamics)
- Well-tempered metadynamics
- Umbrella sampling
- Constrained (restrained) MD
- Steered MD (Jarzynski)
- Adaptive Biasing Force, OPES
- Transition Path Sampling, Transition Interface Sampling
- ...

How to obtain an initial guess path

- Linear interpolation
- Construct guess path by hand
 - Reparametrized for equidistant nodes
- Use the plumed driver
 - needs initial reactive trajectory
 - feed trajectory to driver until path convergence

How many path nodes?



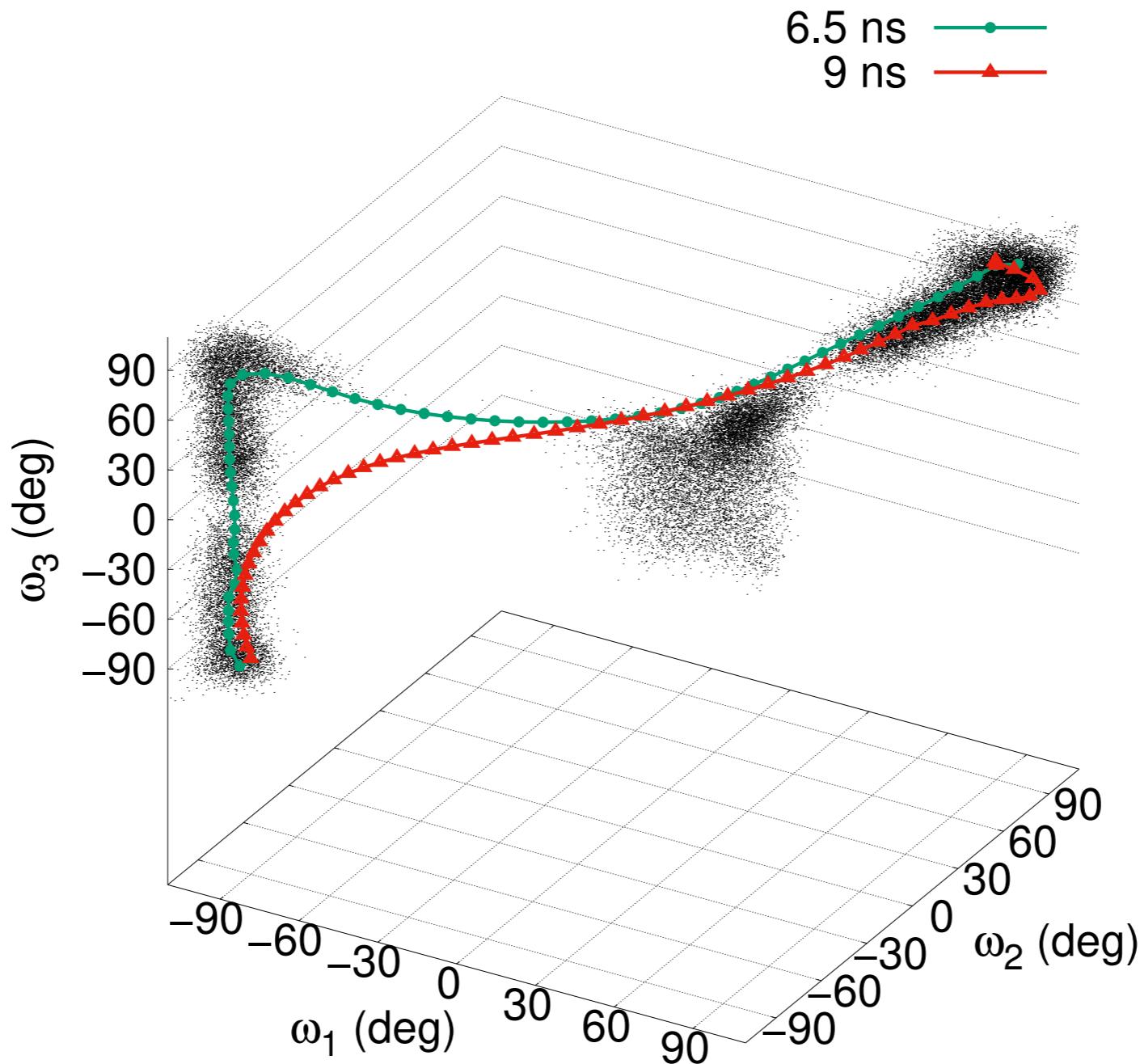
Which CVs as arguments in the path-CV?

- Not too few, not too many
- Rescale the CVs, using COMBINE directive in PLUMED
- Most path and FES convergence problems due to incomplete CV description of the transition
- Second cause of problems: path too flexible

What can possibly go wrong?

- Missing CVs, incomplete description of transition
- Path too stiff
- Path is too flexible
 - fewer nodes; add tube potential; less frequent updates; improve guess path
- Metadynamics potential update much slower/faster than path update
- ...

the corner cutting problem



- the path update algorithm tends to a straight line
- more complex free energy landscapes make it worse

The exercises

- Five masterclass exercises (<https://www.plumed.org/doc-master/user-doc/html/masterclass-22-9.html>)
 - 0_md
 - 1_metadynamics
 - 2_fixedpathmetadynamics
 - 3_adaptivepathmetadynamics
 - 4_multiplewalkers
- Download from github
 - git clone <https://github.com/Ensing-Laboratory/masterclass-22-09>
- Comes with modified Plumed-2.3.0
 - but most works also with more recent plumed versions! (except multiple walkers)
- Advanced exercises: try (your) molecular system with MD software!
 - e.g. plumed nest: plumID 21.049 and plumID 21.033
- Ask questions on the Slack channel

PLUMED INPUT

```
PATHCV LABEL=pcv ARG=d3,d4,t2,dNB INFILE=path.input FIXED=11,60
SCALE=0.5,0.5,1.0,0.5 OUTFILE=path.out STRIDE=500 PACE=500 HALFLIFE=5000

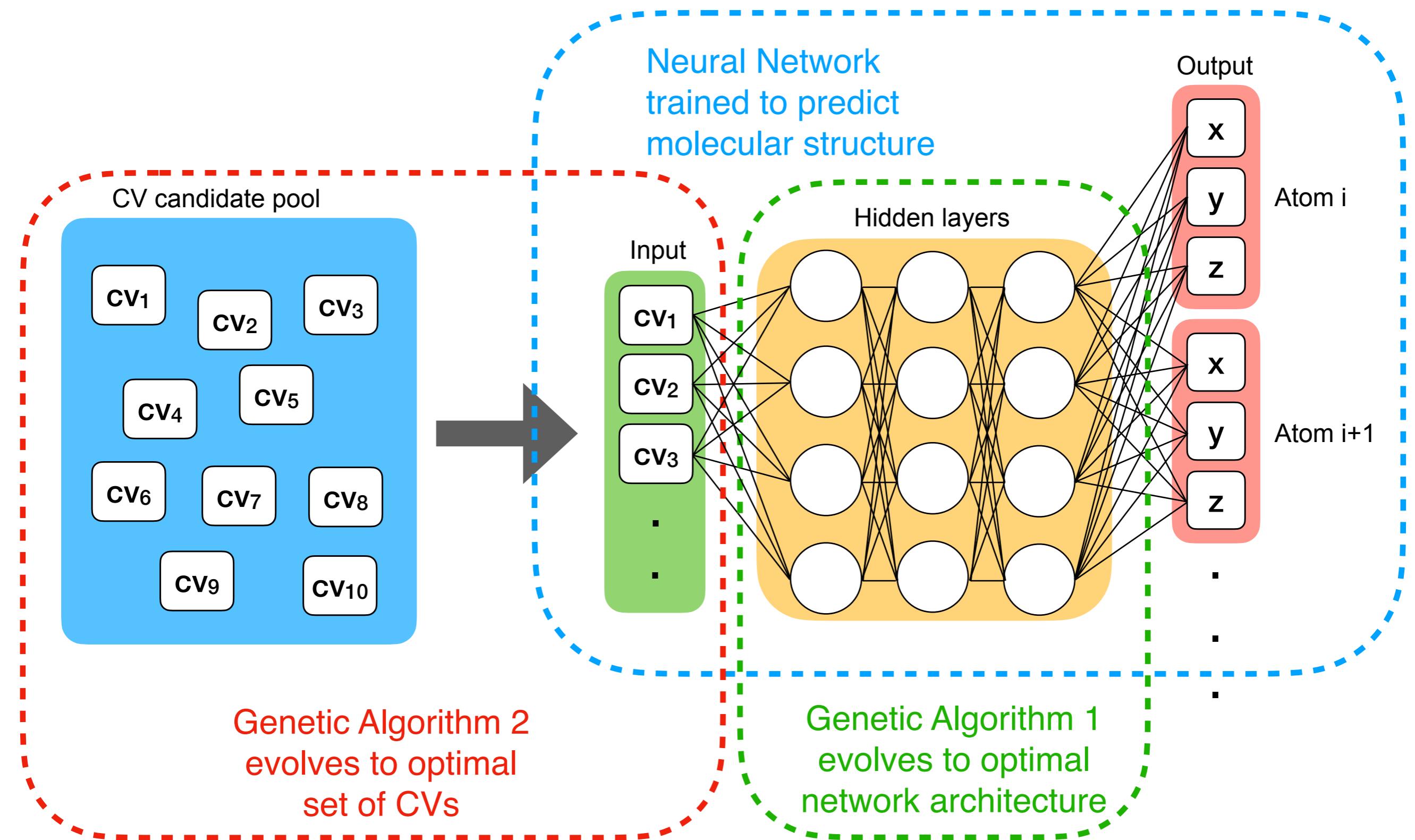
METAD ARG=pcv.s SIGMA=0.2 HEIGHT=0.035 PACE=500 LABEL=metadyn

RESTRAINT ARG=pcv.z LABEL=tube KAPPA=70.0 AT=0.0

LOWER_WALLS ARG=pcv.s AT=-0.2 KAPPA=500.0 EXP=2 EPS=1 OFFSET=0
LABEL=lwall

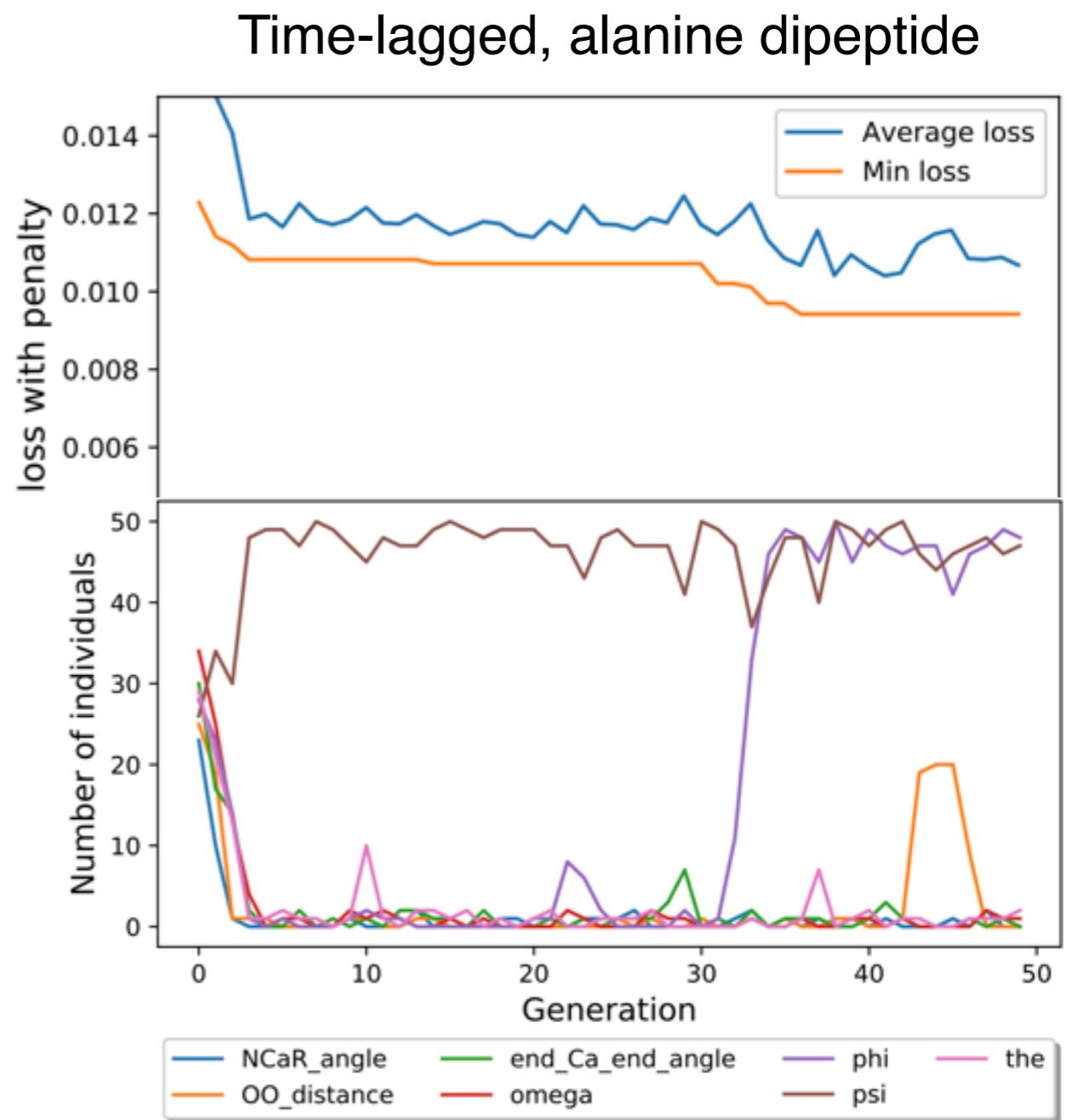
UPPER_WALLS ARG=pcv.s AT=1.2 KAPPA=500.0 EXP=2 EPS=1 OFFSET=0
LABEL=uwall
```

Machine learning framework



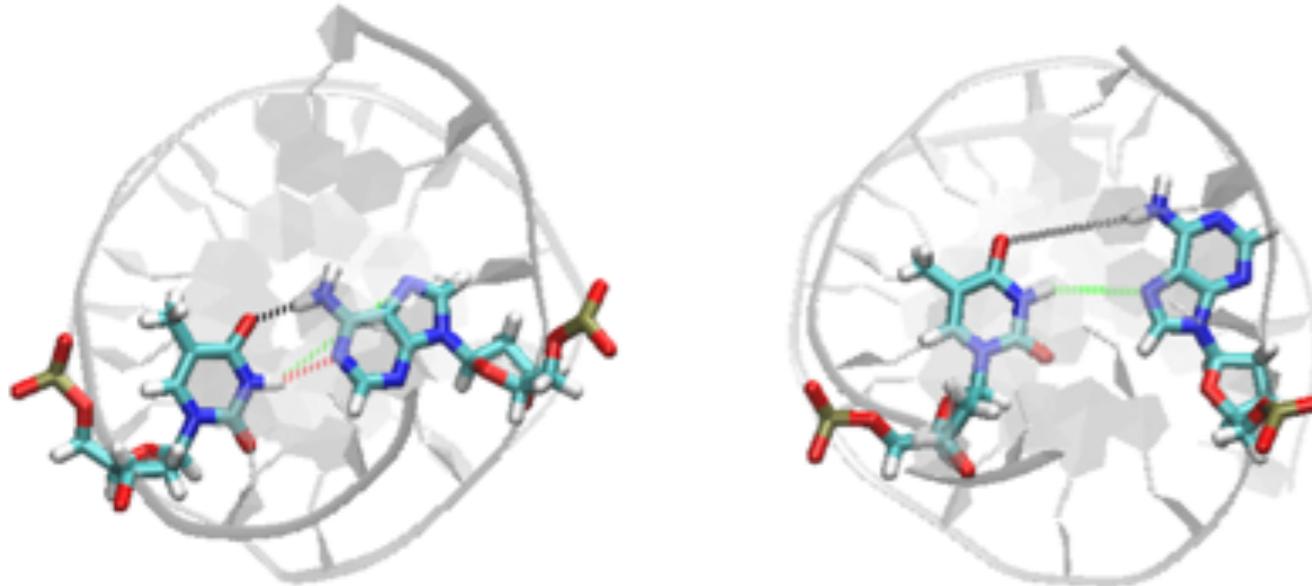
Machine learning framework

- Two networks types: “current time” and “time lagged”
- Atom coordinates: Cartesian works better than Z-matrix
- Configurations are aligned by translation and rotation
- All features (CVs) are normalized
- Input data from Transition Path Sampling trajectories or alternating with enhanced sampling (path-CV)
- Loss function: MAE of positions + penalty for #CVs
- Optimal CVs ranked by increase in MAE-loss when each CV is replaced by noise



Watson-Crick-Franklin to Hoogsteen transition in DNA

Watson-Crick-Franklin \longleftrightarrow Hoogsteen

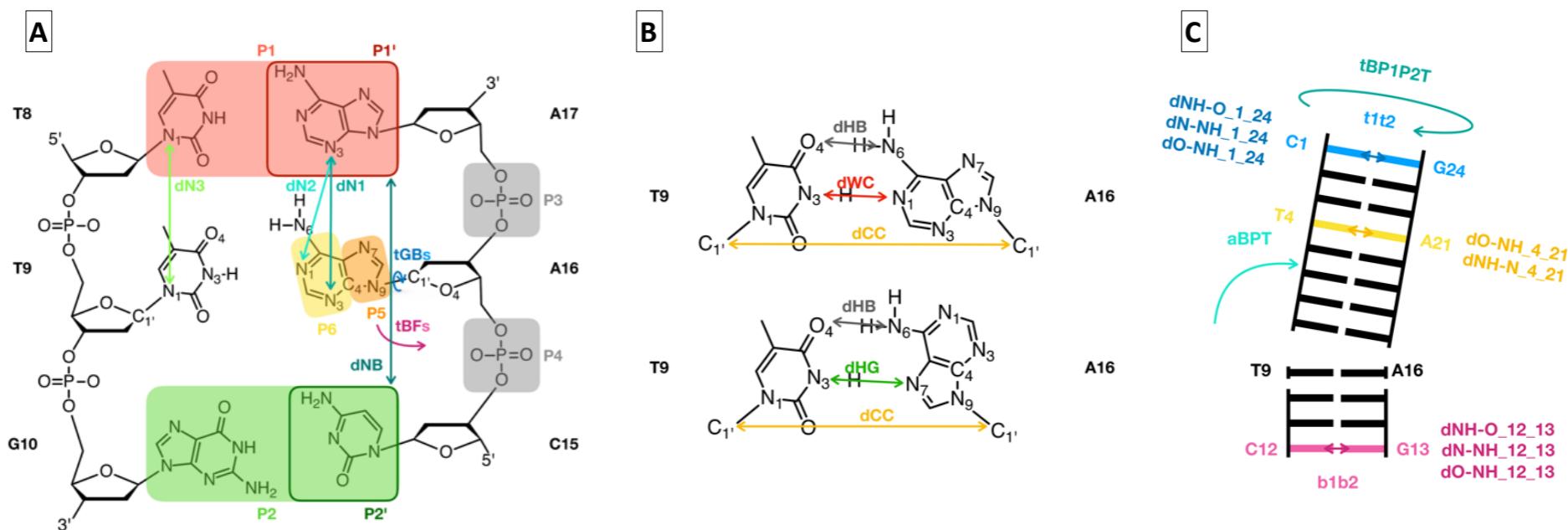


TPS trajectories by Jocelyne Vreede



J. Vreede, A. Pérez de Alba Ortíz, P.G. Bolhuis, & D.W.H. Swenson
Nucleic Acids Research **47** (2019) 11069-11076.

large pool of descriptors

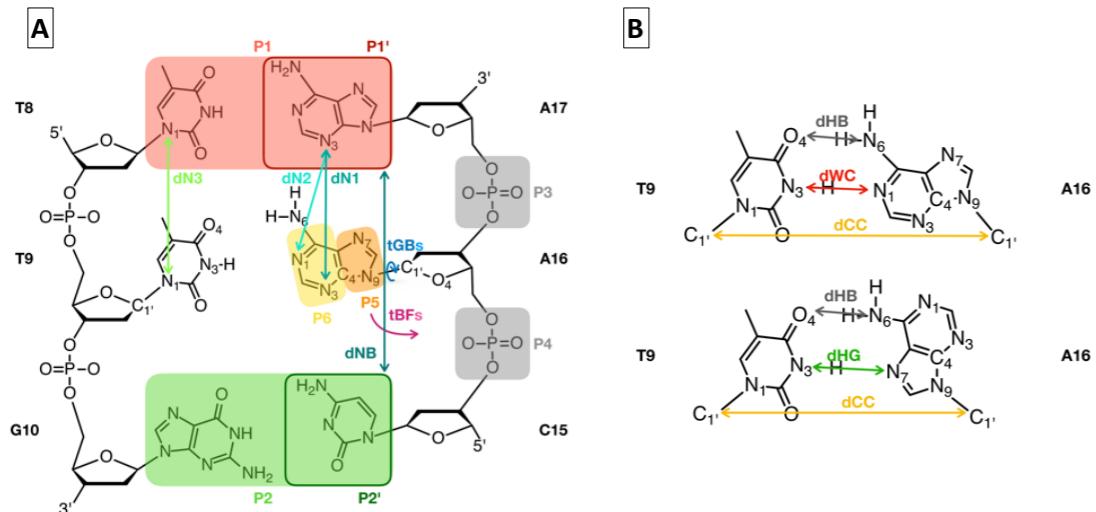


Pool of candidate CVs to describe Watson-Crick-Franklin to Hoogsteen transition in B-DNA

CV	type	description
dWC	distance	distance of the H-bond between A16 (N1) and T9 (N3), characteristic for Watson-Crick
dHG	distance	distance of the H-bond between A16 (N7) and T9 (N3), characteristic for Hoogsteen
dHB	distance	distance of the H-bond between A16 (N6) and T9 (O4), present both in WC and in HG
dCC	distance	distance between A16 (C1) and T9 (C1)
dNB	distance	distance between the centers of mass P1 and P2
dN1	distance	distance between A16 (N3) and A17 (N3)
dN2	distance	distance between A16 (N1) and A17 (N3)
dN3	distance	distance between T8 (N1) and T9 (N1)
dNH-O_12_13	distance	distance between C12 (N4) and G13 (O6)
dN-NH_12_13	distance	distance between C12 (N3) and G13 (N1)
dO-NH_12_13	distance	distance between C12 (O2) and G13 (N2)
dNH-O_1_24	distance	distance between C1 (N4) and G24 (O6)
dN-NH_1_24	distance	distance between C1 (N3) and G24 (N1)
dO-NH_1_24	distance	distance between C1 (O2) and G24 (N2)
dO-NH_4_21	distance	distance between T4 (O4) and A21 (N6)
dNH-N_4_21	distance	distance between T4 (N3) and A21 (N1)
aBPT	angle	angle of the entire DNA segment with the A16-T9 base pair as the vertex, defined by the centers of mass b1b2, p_all, and t1t2
tGB	torsion	torsion around the glycosidic bond defined by the pseudodihedral angle formed by the axis A16 (C1-N9) and the vectors P2-P1 and P5-P6
tBF	torsion	base flipping torsion defined by the pseudodihedral angle (P1+P2)-P3-P4-P5
tGBs	torsion	simpler version of the glycosidic bond torsion defined by the dihedral angle A16 (O4'-C1-N9-C4)
tBPs	torsion	simpler version of the base flipping torsion defined by the dihedral angle by the axis A16(C4'-C3') and the vectors T9(C3')-A16(C4') and A16(C3'-N1)
tBP1P2T	torsion	torsion of the entire segment of DNA defined by the pseudodihedral angle b1b2-p1-p2-t1t2
atan_dWCdHG	function	evaluation of the function atan2(dWC, dHG)

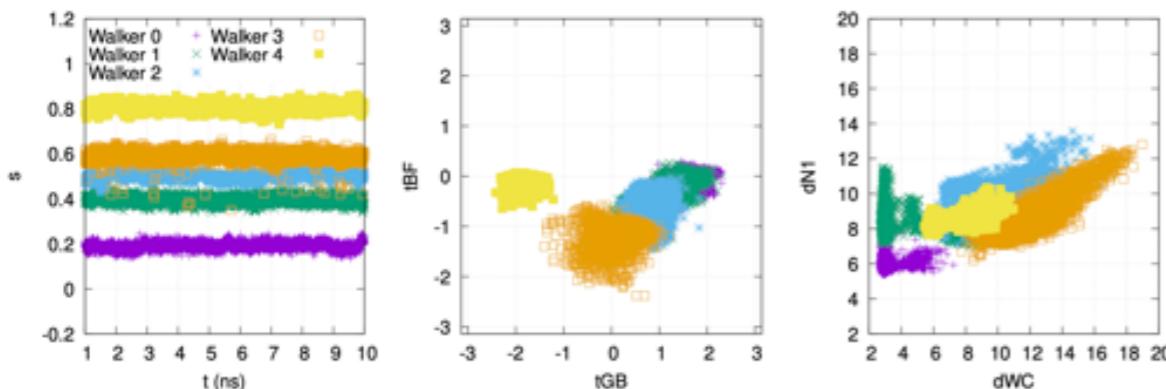
Iterative learning of collective variables

Large pool of collective variables



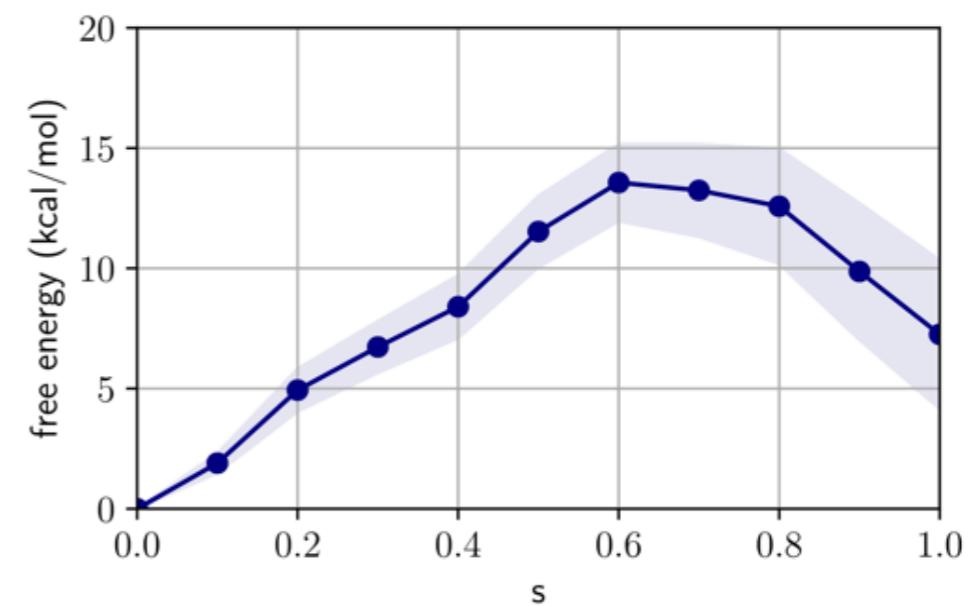
Machine learning

Biased Path-CV run with optimal CV selection



CV	MAE loss (10^{-2})	Importance
tBF	0.27	3.3
dHB	0.18	2.0
tBFs	0.12	1.0
tGB	0.10	0.60
dN1	0.10	0.59
dNB	0.08	0.23

Free energy profile



Acknowledgements

Alberto Pérez de Alba Ortíz

Rakesh Chandran
Puthenkalathil

Ambuj Tiwari



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

Simultaneous sampling of multiple transition channels using adaptive paths of collective variables.
Alberto Pérez de Alba Ortíz and Bernd Ensing
arXiv [cond-mat.stat-mech] 2112.04061 (2021)

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-Fornea, 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 melting 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 metadynamics, 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 convergence 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