Computing free energies with PLUMED 2.0

Davide Branduardi
Formerly at MPI Biophysics, Frankfurt a.M.

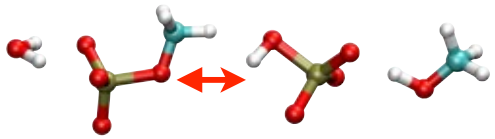
Outline of the talk

- Relevance of free energy computation
- Free-energy from histograms: issues
- Tackling the sampling problem
- Tackling the order-parameter problem

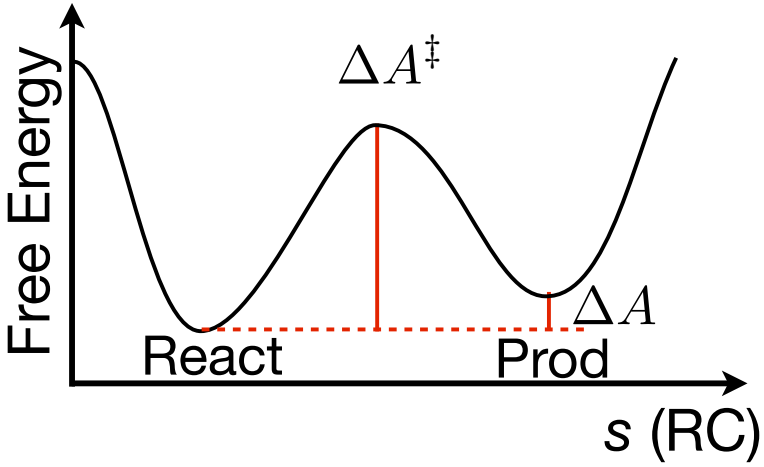
Relevance of free energy calculations

* Free energy is a measure of probability $A(s_0) = -k_B T \ln P(s_0)$

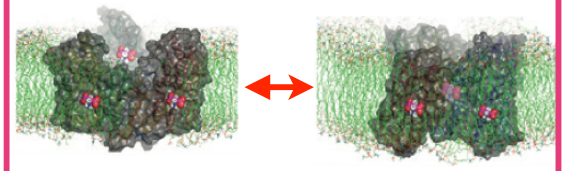
Chemical reactions



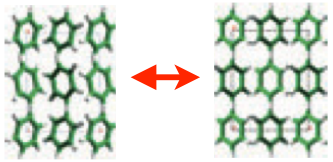
find elusive structures and mechanisms!



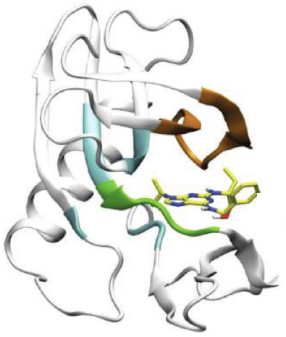
Biophysics



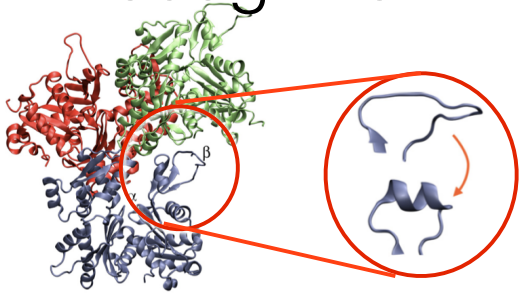
Phase transitions



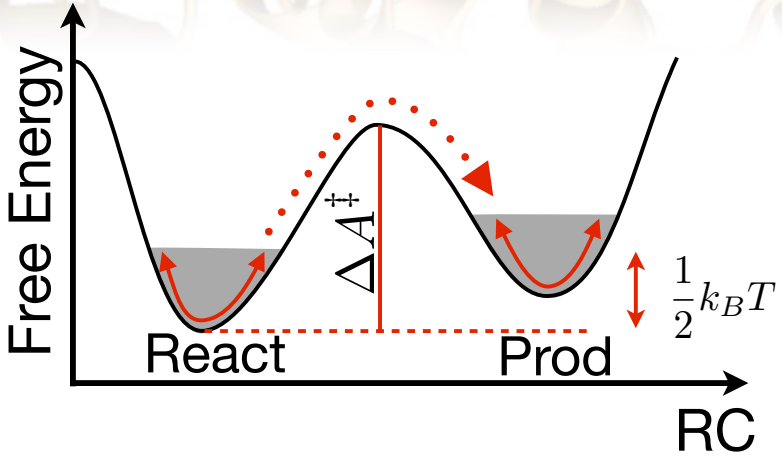
Drug binding



Molecular recognition

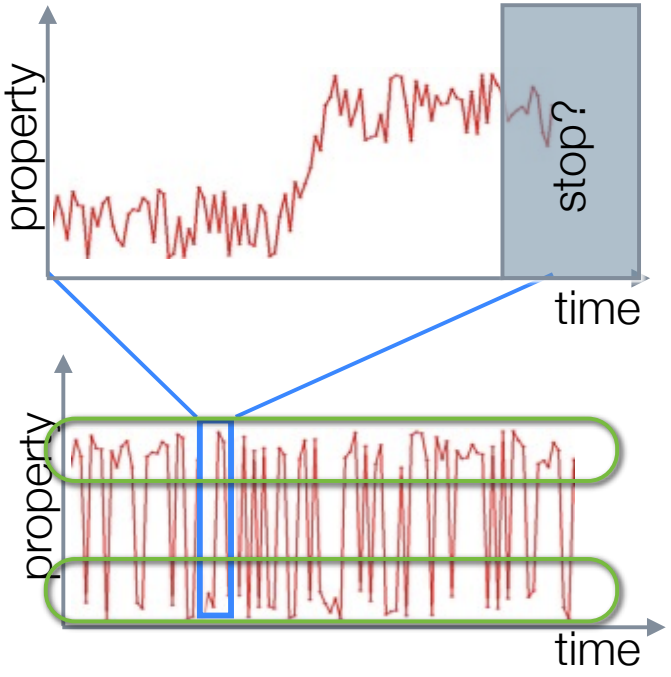


Rare event: troubles with ergodic hypothesis



$$k(T) = \frac{k_B T}{hc^0} e^{-\Delta A^{\ddagger}/RT}$$

Now I sampled all the two states. Is it sufficient?



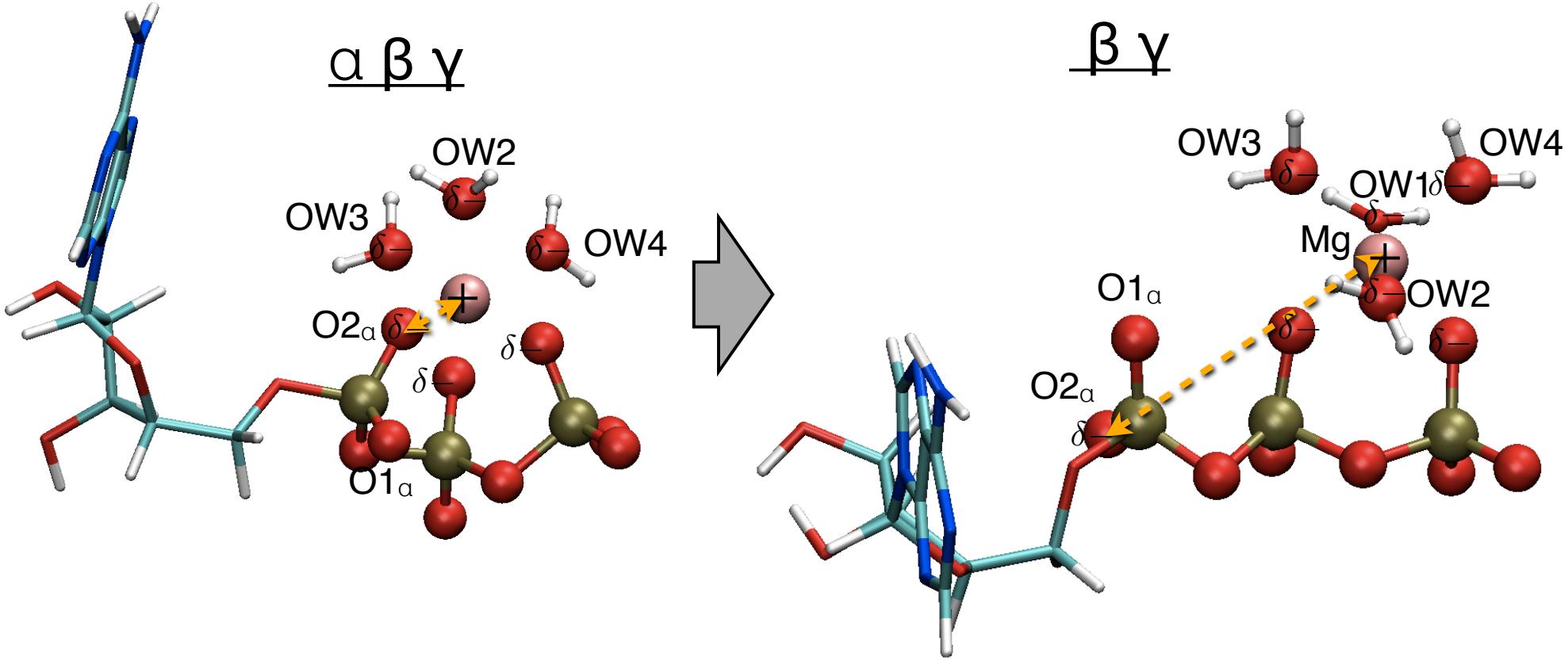
Free energy requires statistics!

$$\Delta A = -\frac{1}{\beta} \ln \frac{N_1}{N_2}$$

A test on ATP-Mg^{2+} in water

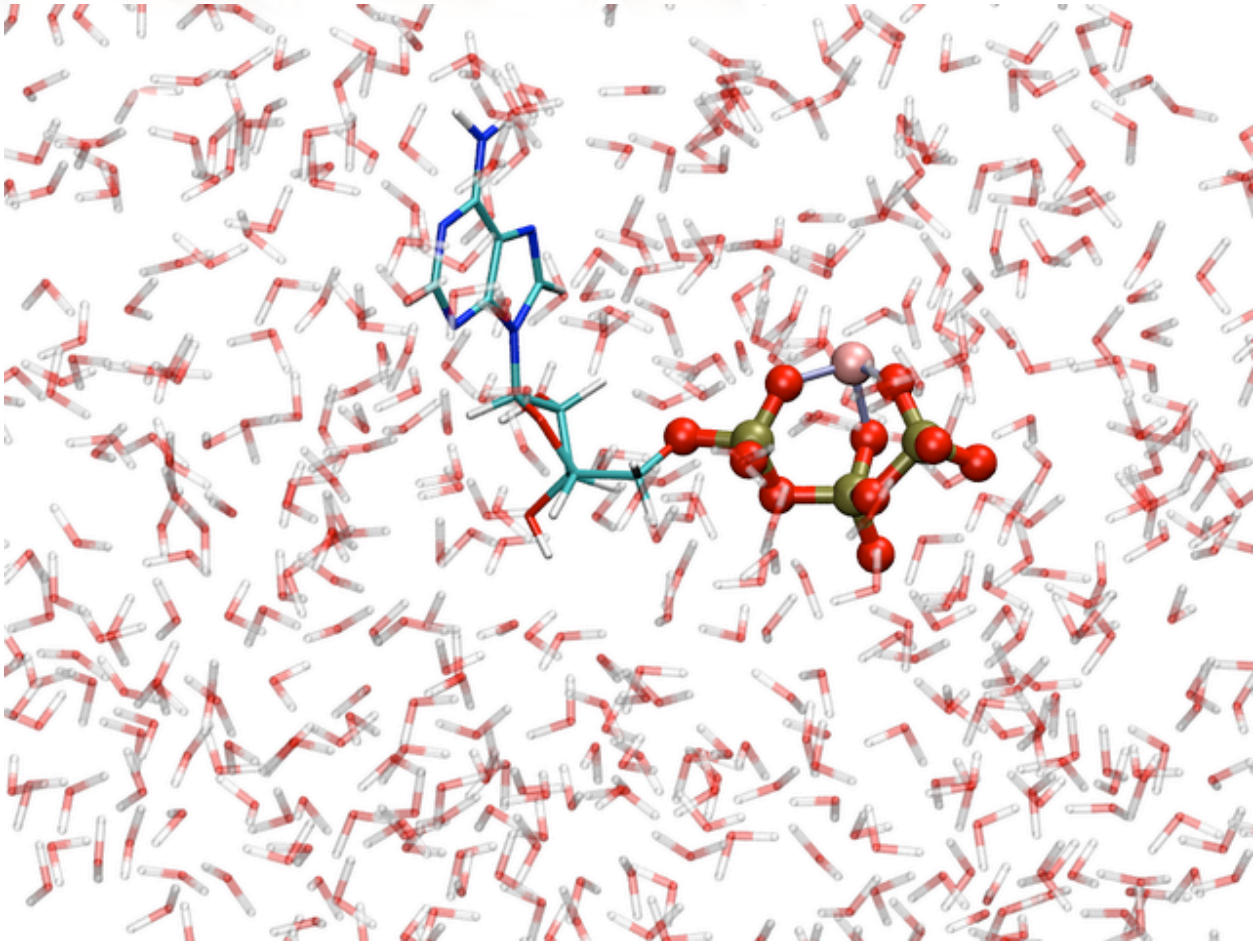
ATP^{4-}-Mg^{2+} is of fundamental importance in bioenergetics and regulation

Free energy of α β γ to β γ



[1] Liao, Sun, Chandler, Oster Eur. Biophys. J. 2004, vol. 33, p. 29

Ergodic hypothesis: am I visiting all the states?



Classical MD, CHARMM27 force field, 915 TIP3P water molecules, GROMACS 4.5.5 code

Possible solutions

Replica exchange methods: multiple simulations at different temperatures or Hamiltonians

Biased dynamics: add a potential function along the degree(s) of freedom you want to accelerate

Biased dynamics + Replica exchange

Thermodynamic-Integration-like approaches

* The ancestor is mean force calculation through harmonic potential

$$F(s) = -k_B T \ln P(s)$$

$$\frac{\partial F(s)}{\partial s}\Big|_{s_0} = -k_B T \frac{1}{\int e^{-\frac{V(x)}{k_B T}} \delta(s'(x) - s)dx} \frac{\partial}{\partial s} \int e^{-\frac{V(x)}{k_B T}} \delta(s'(x) - s)dx \Big|_{s_0}$$

* Dirac delta function ≈ Gaussian function (I like this trick!)

$$\frac{\partial F(s)}{\partial s}\Big|_{s_0} \simeq -k_B T \frac{1}{\int e^{-\frac{V(x)}{k_B T}} e^{-\frac{k}{2k_B T}(s'(x)-s)^2} dx} \frac{\partial}{\partial s} \int e^{-\frac{V(x)}{k_B T}} e^{-\frac{k}{2k_B T}(s'(x)-s)^2} dx \Big|_{s_0}$$

* It is an average of a biased simulation!

$$\frac{\partial F(s)}{\partial s}\Big|_{s_0} \simeq -\frac{1}{\int e^{-\frac{V(x)+k/2(s'(x)-s_0)^2}{k_B T}} dx} \int k(s'(x) - s_0) e^{-\frac{V(x)+k/2(s'(x)-s_0)^2}{k_B T}} dx$$

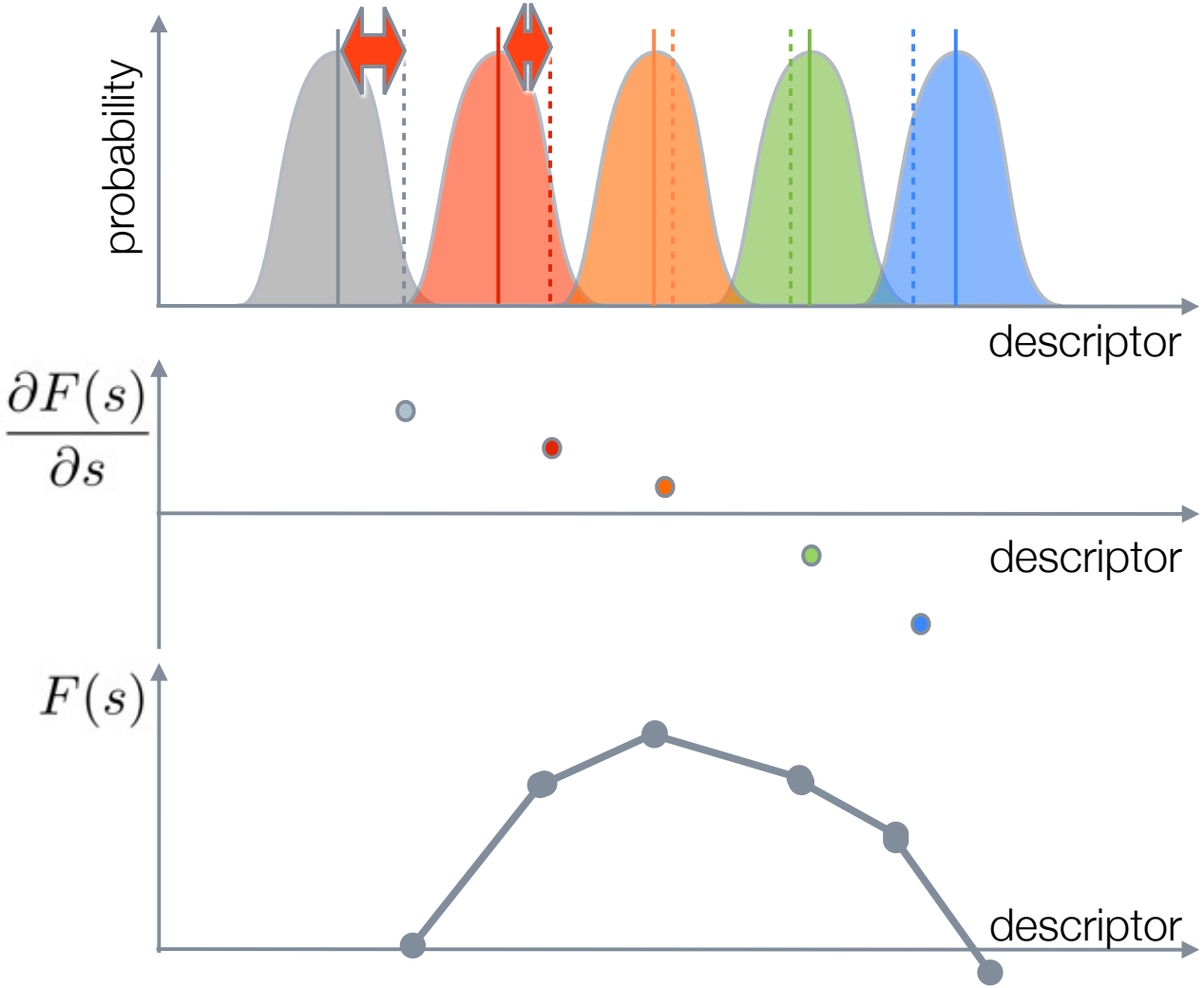
$$= -k\langle s'(x) - s_0 \rangle_{bias,s0}$$ (opposite of the "mean force")

Thermodynamic Integration 101

* Make a constrained simulation (go over barrier!)
* Acquire mean force
* Integrate

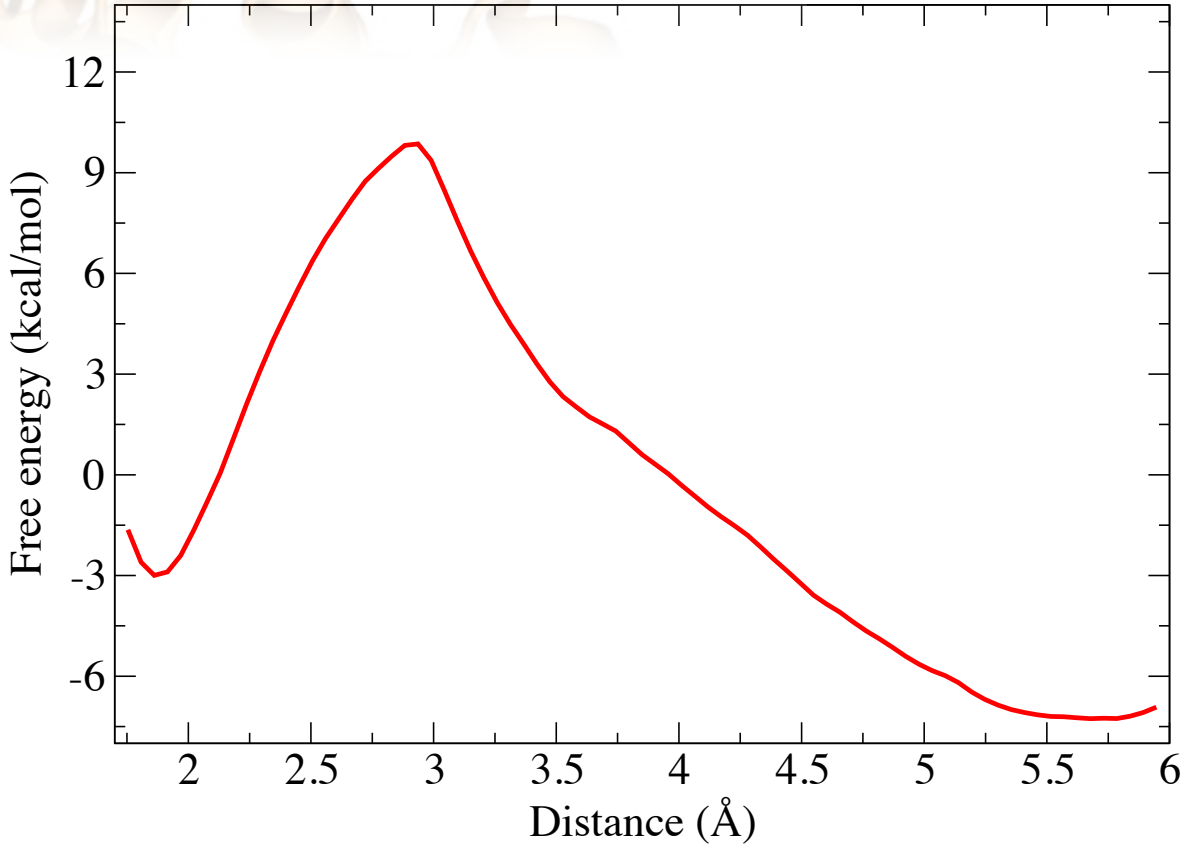
Sloping, slatted, wooden platforms are preferable for sheep dragging.

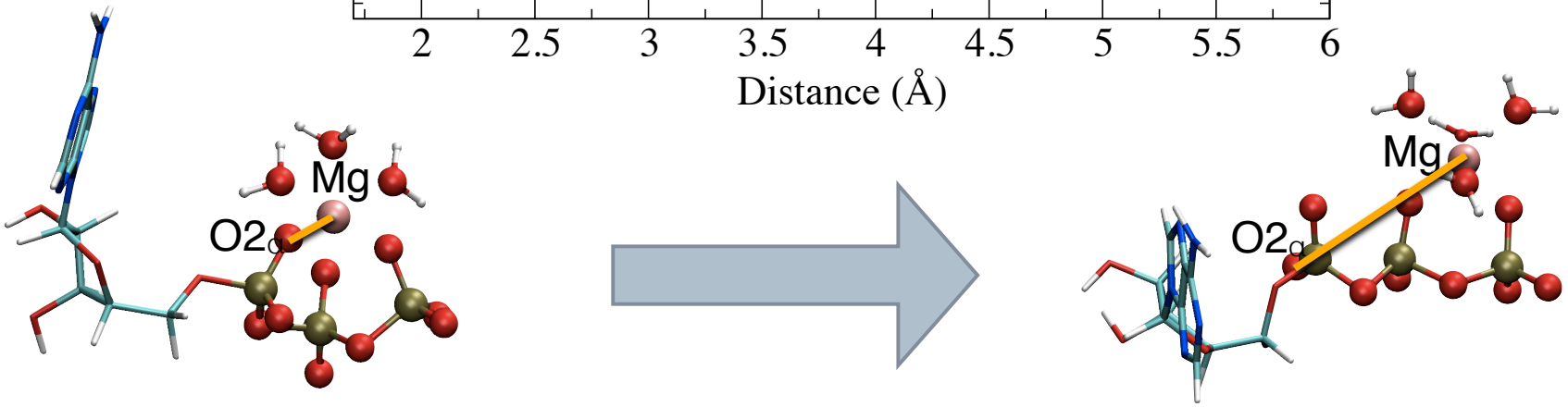
© J. Culvenor



or +WHAM etc...

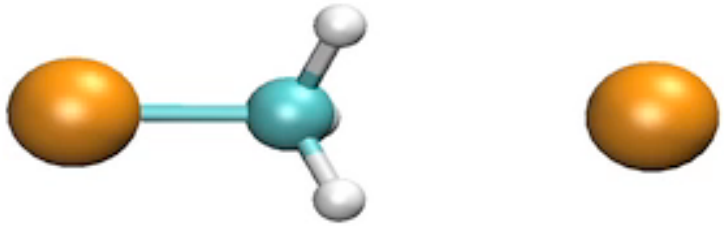
An example with ATP-Mg^{2+}



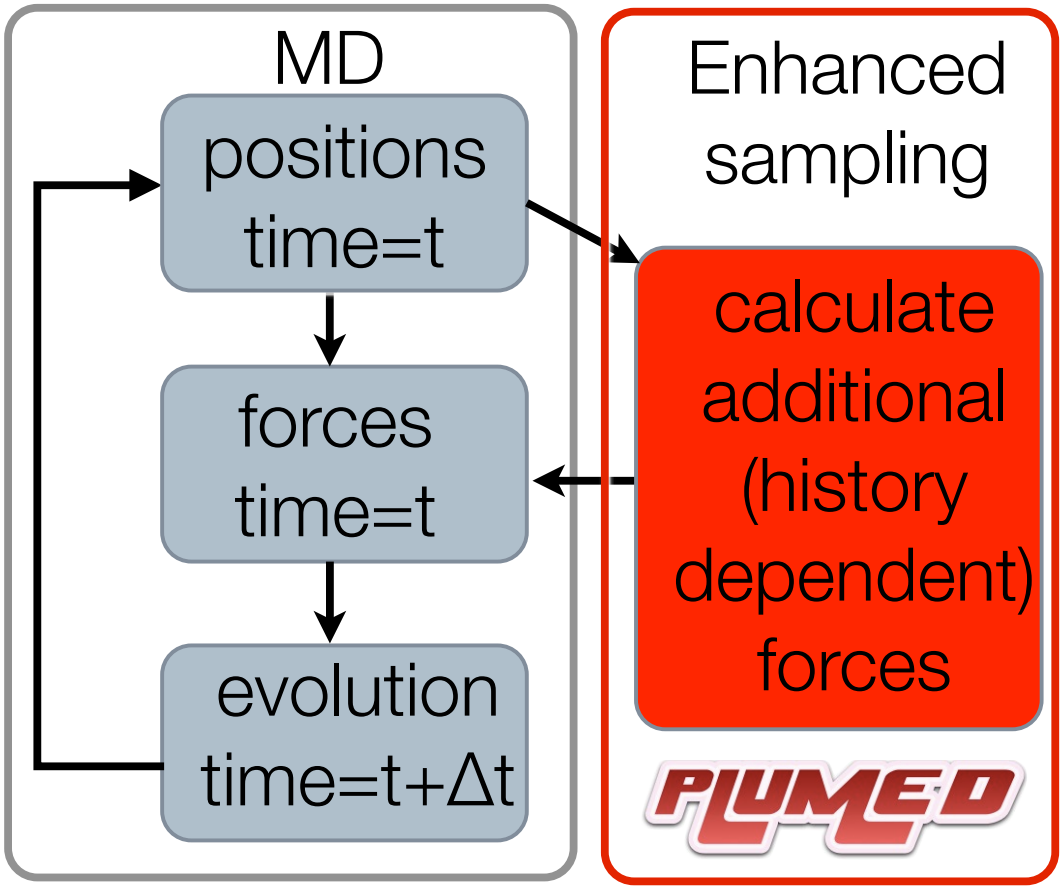


Thermodynamic-Integration-like approaches

* Thermodynamic integration Beveridge, DiCapua, Annu Rev Biophys Biophys Chem 1989, vol 18, pp. 431.
* WHAM Roux, Comp Phys Comm 1995, vol. 91, pp. 275.
* Free energy perturbation Beveridge, DiCapua, Annu Rev Biophys Biophys Chem 1989, vol 18, pp. 431.
* Jarzynski-equation based approaches (steered-MD) Jarzynski Phys Rev Lett 1997, vol. 78, pp. 2690.
* Crooks-equation based approaches (two directions steered-MD) Crooks J Stat Phys 1998, vol. 90, pp. 1481.



Avoid recoding with PLUMED



Each single engine had only few enhanced sampling techniques (according developers' interests)

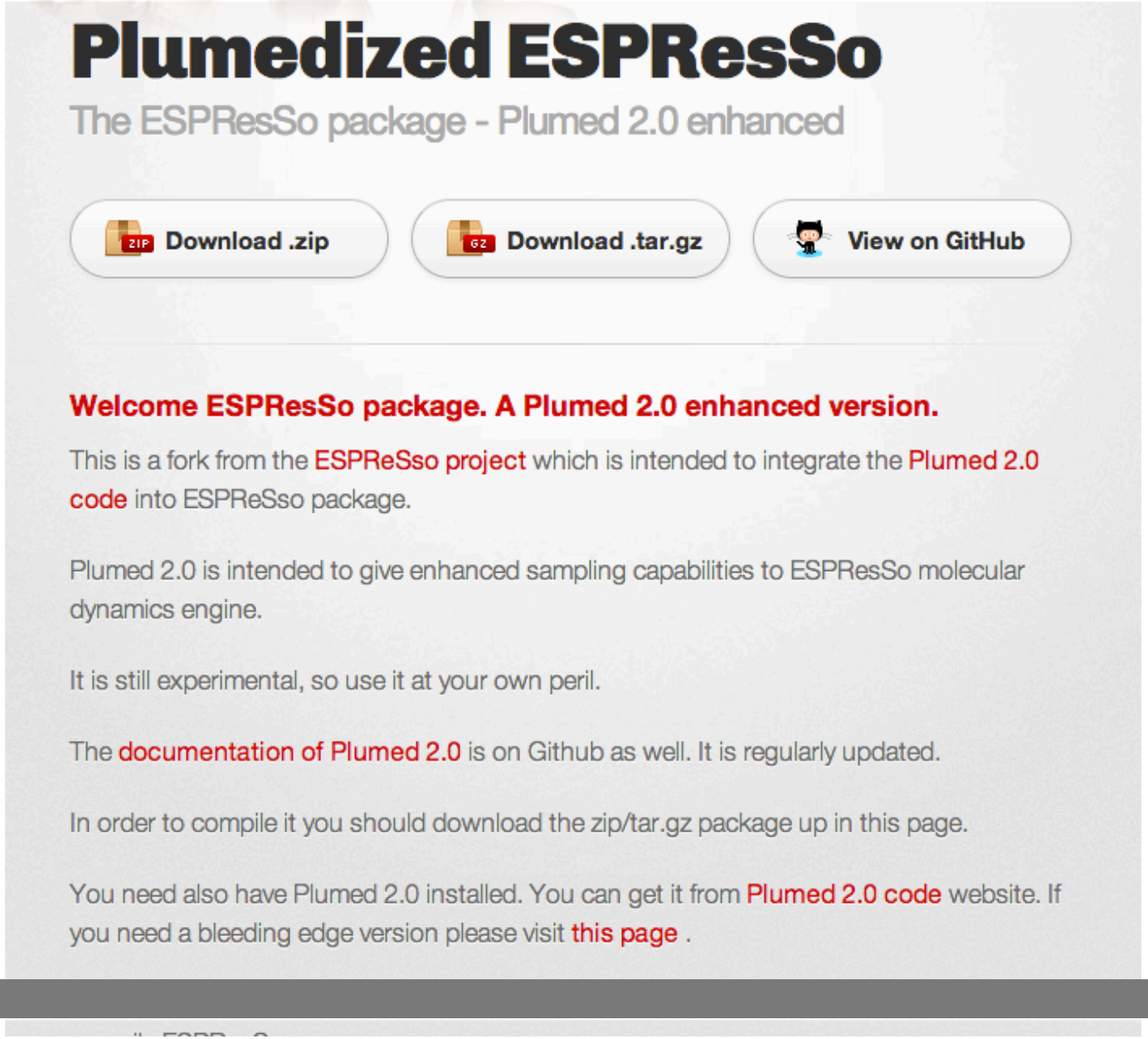
in NAMD, GROMACS, AMBER, LAMMPS, QUANTUM-ESPRESSO

https://github.com/plumed/plumed2

Bonomi, Branduardi, Bussi, Camilloni, Provasi, Raiteri, Donadio, Marinelli, Pietrucci, Broglia, Parrinello, M. Comp. Phys. Commun. 2009, 180, 1961–1972.

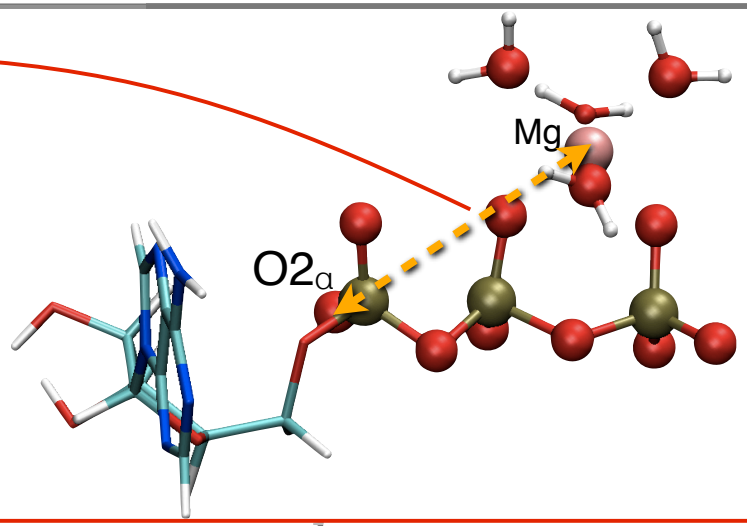
Tribello, Bonomi, Branduardi, Camilloni, Bussi Comp. Phys. Commun. (2013) in press.

Now in ESPResSo: http://davidebr.github.io/espresso/



How to set this up in PLUMED?

```
# MG - O  distance
DISTANCE ...
 ATOMS=44,34
 LABEL=d
... DISTANCE
# a moving restraint, in steps
MOVINGRESTRAINT ...
ARG=d
STEP0=0 AT0=0.170000 KAPPA0=200000
STEP1=50000 AT1=0.170000
STEP2=50300 AT2=0.175375
.
.
.
STEP160=4024000 AT160=0.600000
STEP161=4074000 AT161=0.600000
... MOVINGRESTRAINT
#
# calculate the coordination of counterions phosphate group
#
COORDINATION ...
        LABEL=coor
        GROUPA=41,42,43,37,38,33,34
        GROUPB=45,46
        NN=8 MM=12 D_0=0.0 R_0=0.5
... COORDINATION
#
# prevent the ions to stick on the phosphate
#
UPPER_WALLS ARG=coor AT=0.2 KAPPA=10000  EXP=4 LABEL=uwall
#
# do some printout
#
PRINT ARG=* STRIDE=100  FILE=COLVAR
ENDPLUMED
```



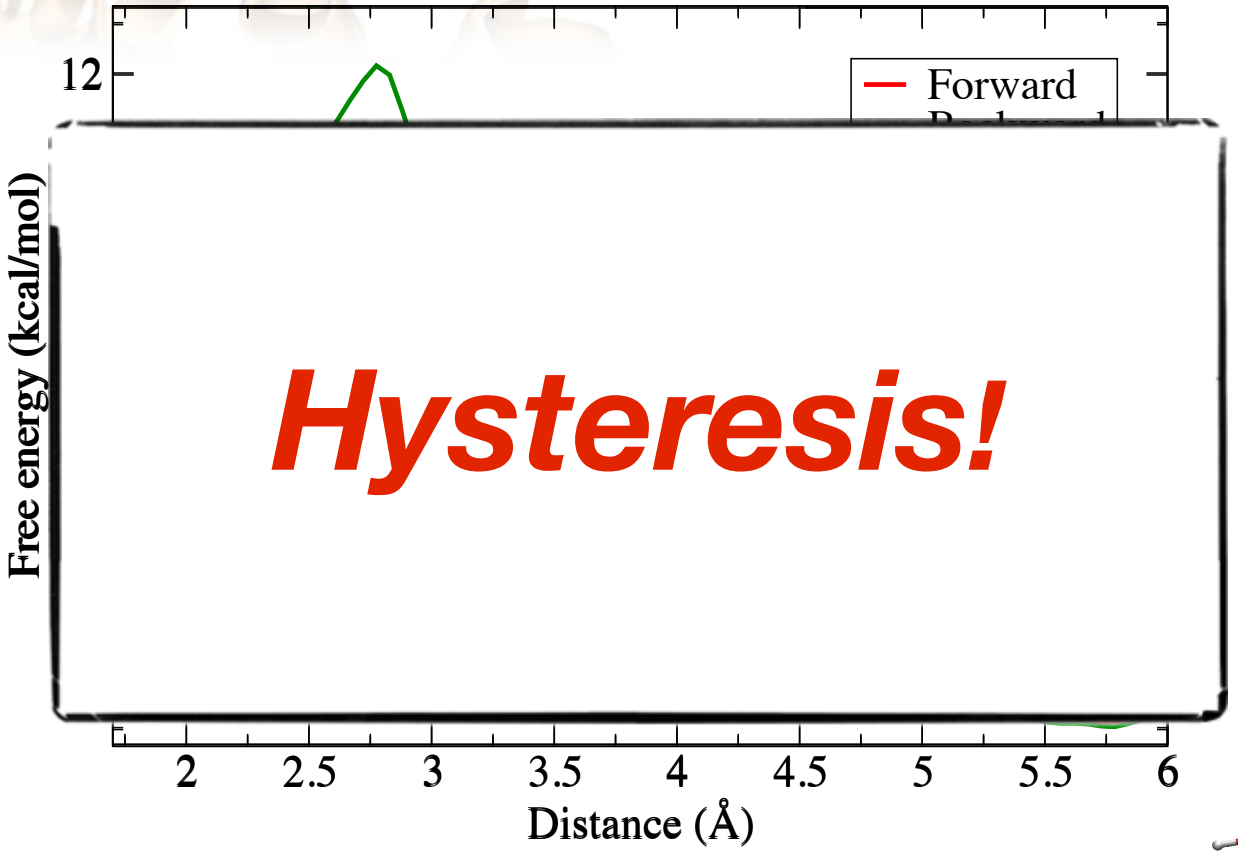
$$m(\mathbf{X},t) = \frac{k}{2}(d(\mathbf{X}) - x_0(t))^2$$

$$c(\mathbf{X}) = \sum_i^{SOD} \sum_j^{PHO} \frac{1 - \left(\frac{|\mathbf{x}_i - \mathbf{x}_j| - d_0}{r_0}\right)^{NN}}{1 - \left(\frac{|\mathbf{x}_i - \mathbf{x}_j| - d_0}{r_0}\right)^{MM}}$$

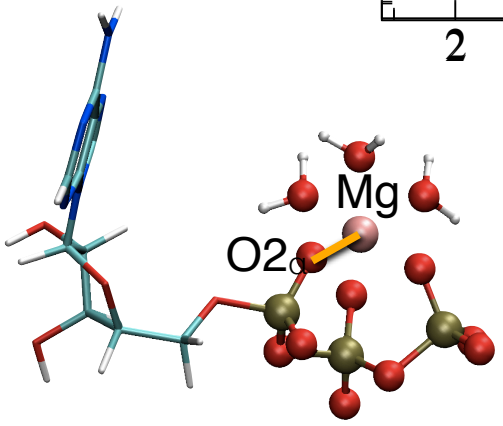
$$c(\mathbf{X}) > 0.2 \ : \ U(\mathbf{X}) = k\,(c(\mathbf{X}) - 0.2)^4$$

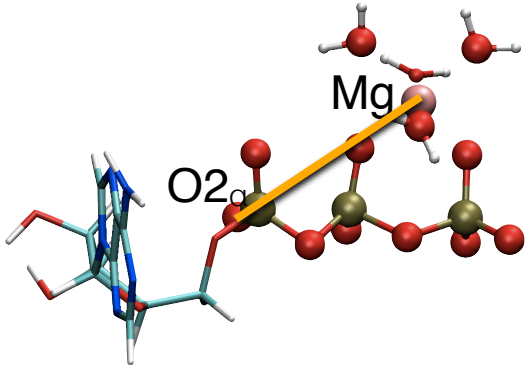
$$c(\mathbf{X}) < 0.2 \ : \ U(\mathbf{X}) = 0.$$

An example with ATP-Mg^{2+}



Hysteresis!





Adaptive methods: Torrie & Valleau[3]

Free energy:

$$F(s_0) = -\frac{1}{\beta}\ln P(s_0)$$
$$= -\frac{1}{\beta}\ln \frac{\int e^{-\beta V(x)}\delta(s(x)-s_0)dx}{Q}$$

Free energy for a biased system:

$$F_b(s_0) = -\frac{1}{\beta}\ln P_b(s_0)$$
$$= -\frac{1}{\beta}\ln \frac{\int e^{-\beta(V(x)+V_b(s(x)))}\delta(s(x)-s_0)dx}{Q_b}$$
$$= V_b(s_0) - \frac{1}{\beta}\ln \frac{\int e^{-\beta V(x)}\delta(s(x)-s_0)dx}{Q_b}$$

use biasing potential to overcome barriers and retrieve free energies

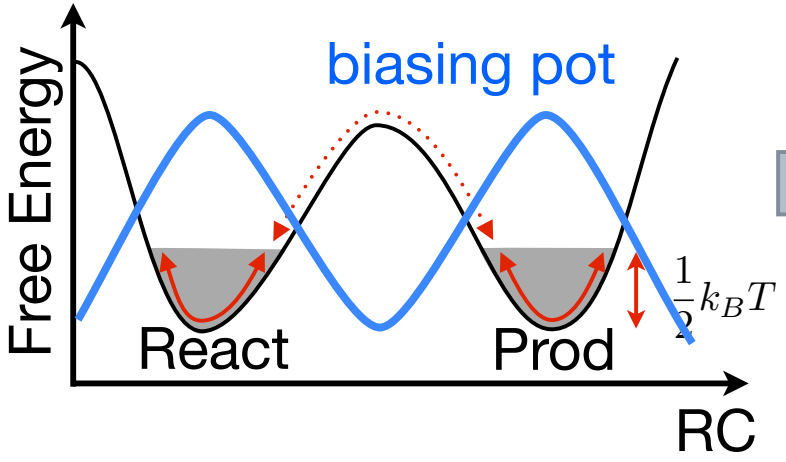
$$F_b(s_0) = V_b(s_0) - \frac{1}{\beta}\ln P(s)\frac{Q}{Q_b}$$

$$F(s_0) = F_b(s_0) - V_b(s_0) + C$$

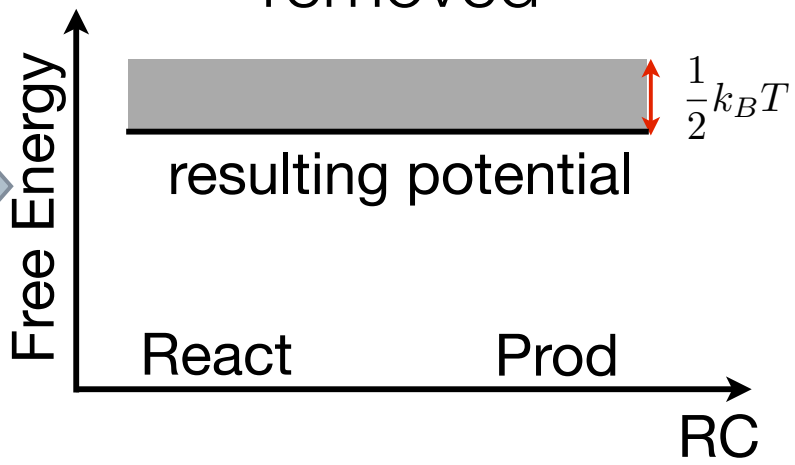
[3] Torrie, Valleau J. Comp. Phys, 1977 vol. 23, pp 187

Lessons from Torrie & Valleau

Biasing potentials:



Metastabilities are removed



* allow to measure free energies: useful
* allow to overcome barriers: cheap
* need only an additional energy term to standard force field: easy

[3] Torrie, Valleau J. Comp. Phys, 1977 vol. 23, pp 187

Adaptive sampling methods

* Adaptive Biasing Force Darve & Pohorille, J. Chem. Phys. 2001, vol. 115, pp. 9169.
* Adaptive Umbrella Sampling Mezei, M. J Comput Phys 1987, vol. 68, pp. 237.
* Self Healing Umbrella Sampling Marsili et. al J. Chem. Phys. B vol. 110, pp. 14011.
* Metadynamics Laio & Parrinello, PNAS 2002, vol. 20, pp. 12562.
* Conformational Flooding Grubmüller, Phys Rev E 1995, vol. 52, pp. 2893.
* ...

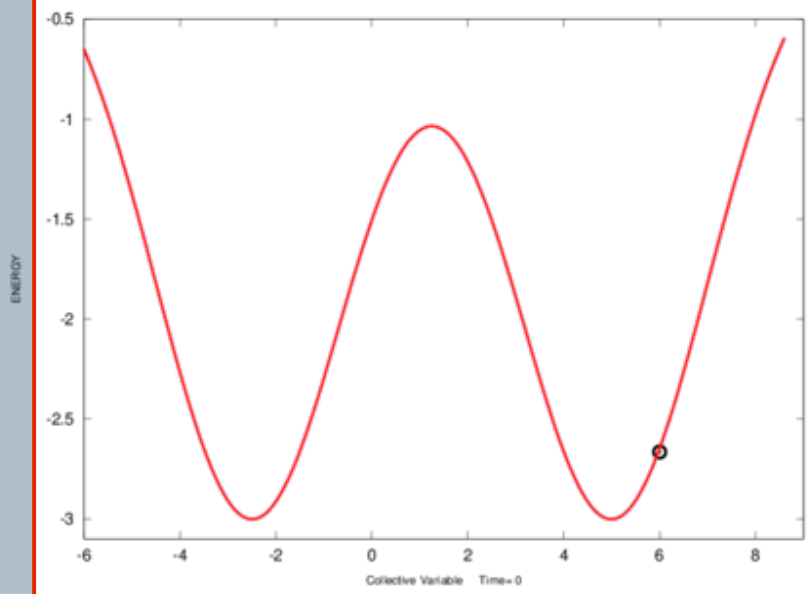
Enhanced sampling: metadynamics

- Apply a repulsive gaussian potential up to compensate the underlying free energy:

$$V_g(t, s(\mathbf{x})) = \sum_{i=1}^{N_{hills}} W \; e^{-\frac{(s(\mathbf{x}) - s_i)}{2\sigma^2}}$$

$$\widetilde{H}(t, \mathbf{x}) = H(\mathbf{x}) + V_g(t, s(\mathbf{x}))$$

- Produces a non markovian dynamics. The negative of the potential deposited is proven to converge to the correct free energy surface [6]
- Transition between metastable states are enhanced by artificial removal of metastable states



[4] Laio & Parrinello, PNAS 2002, vol. 20, p. 12562.
[5] Bussi & Laio, Parrinello Phys. Rev. Lett. 2006, vol 96, p. 90601
[6] Branduardi, Bussi, Parrinello, J. Chem. Theory Comput. 2012, vol 8, p. 2247

How to set this up in PLUMED?

```
# MG - oxygen
DISTANCE ...
 ATOMS=44,34
 LABEL=d
... DISTANCE
#
# Metadynamics: add a hill of 0.6
# every 500 ps, large 0.015nm, and use a
# grid (to make it faster)
#
METAD ...
   ARG=d
   PACE=500 SIGMA=0.015 HEIGHT=0.6
   GRID_MIN=0.16 GRID_MAX=0.7 GRID_BIN=500
   GRID_WFILE=mygrid.dat GRID_WSTRIDE=5000
... METAD
#
# prevent the system to explore unwanted regions
#
UPPER_WALLS ARG=d AT=0.65 KAPPA=100 EPS=0.3 EXP=4 LABEL=dwall
# coordination PHO + SOD
COORDINATION ...
 LABEL=coor
 GROUPA=41,42,43,37,38,33,34
 GROUPB=45,46
 NN=8 MM=12 D_0=0.0 R_0=0.5
... COORDINATION
# wall on the coordination
UPPER_WALLS ARG=coor AT=0.2 KAPPA=10000  EXP=4 LABEL=uwall
# some printouts
PRINT ARG=* STRIDE=100  FILE=COLVAR
ENDPLUMED
```

$$V_g(t, s(\mathbf{x})) = \sum_{i=1}^{N_{hills}} W \; e^{-\frac{(s(\mathbf{x})-s_i)}{2\sigma^2}}$$

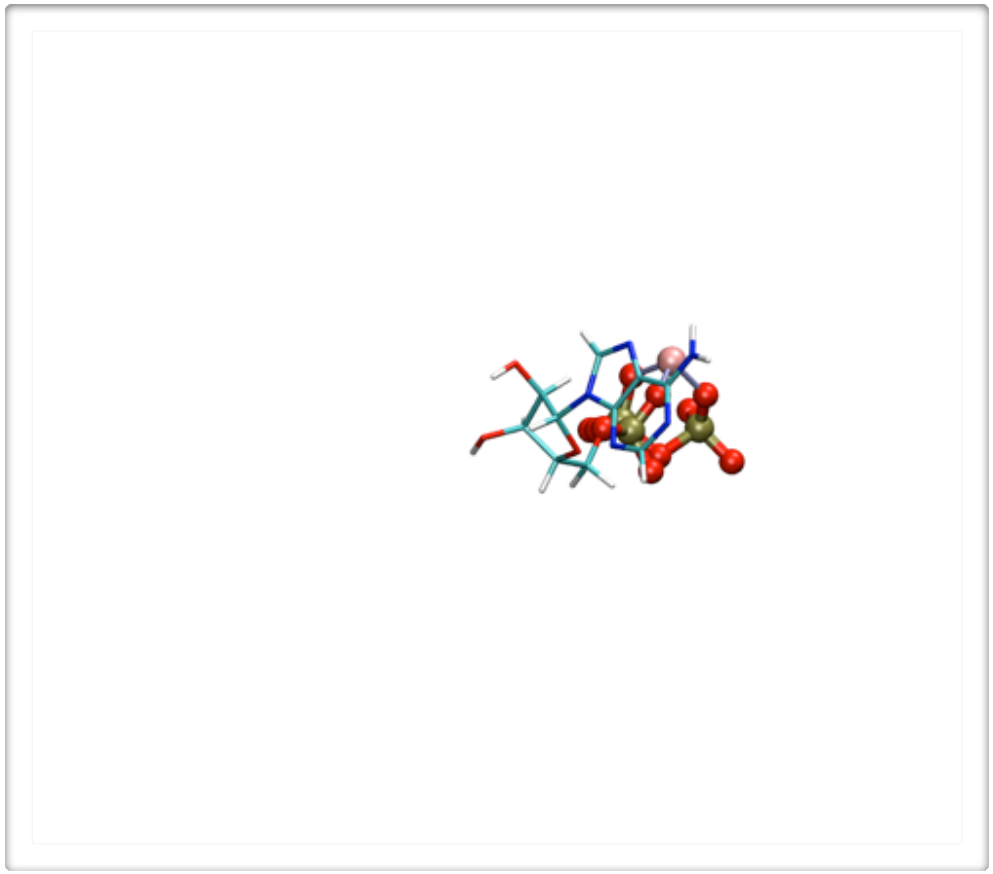
$$\widetilde{H}(t, \mathbf{x}) = H(\mathbf{x}) + V_g(t, s(\mathbf{x}))$$

$$c(\mathbf{X}) > 0.65 \;:\; U(\mathbf{X}) = 100 \left(\frac{d(\mathbf{X}) - 0.65}{0.3} \right)^4$$

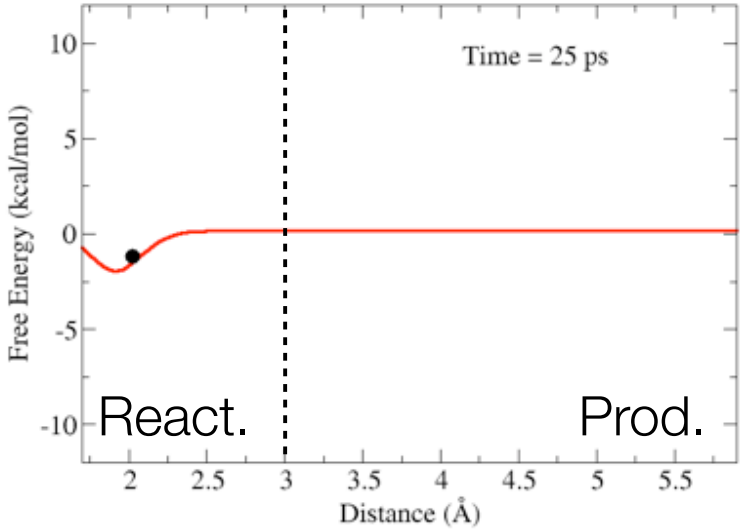
$$c(\mathbf{X}) < 0.65 \;:\; U(\mathbf{X}) = 0.$$

A test on ATP-Mg^{2+} in water: meta with distance

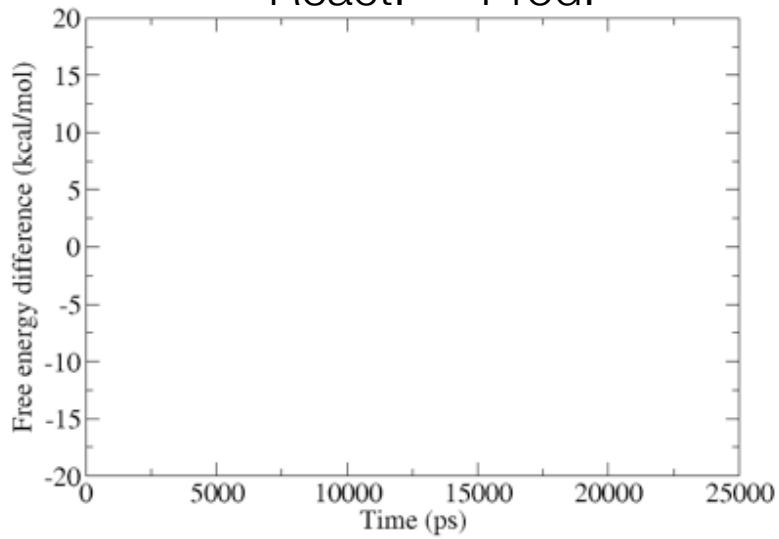
Choose the CV(s): maximum 3!



Free energy estimate



$A_{React.}$-$A_{Prod.}$



Adaptive sampling vs TI-like methods

Adaptive sampling:

* The problem is multidimensional but probably the accessible phase space is limited: they explore only what you need
* The free energy landscape has competitive, parallel reactive paths (solvent degrees of freedom, rotations) that can be overcome at some point

TI-like:

* In 1-d (max 2-d, via WHAM or, better with rbf fitting schemes) very effective
* Sometimes trivially parallel (many umbrellas at same time)

Misconceptions in free energy calculations

Technique (TI, WHAM, SMD, MetaD, ABF?)

Choice of order parameter

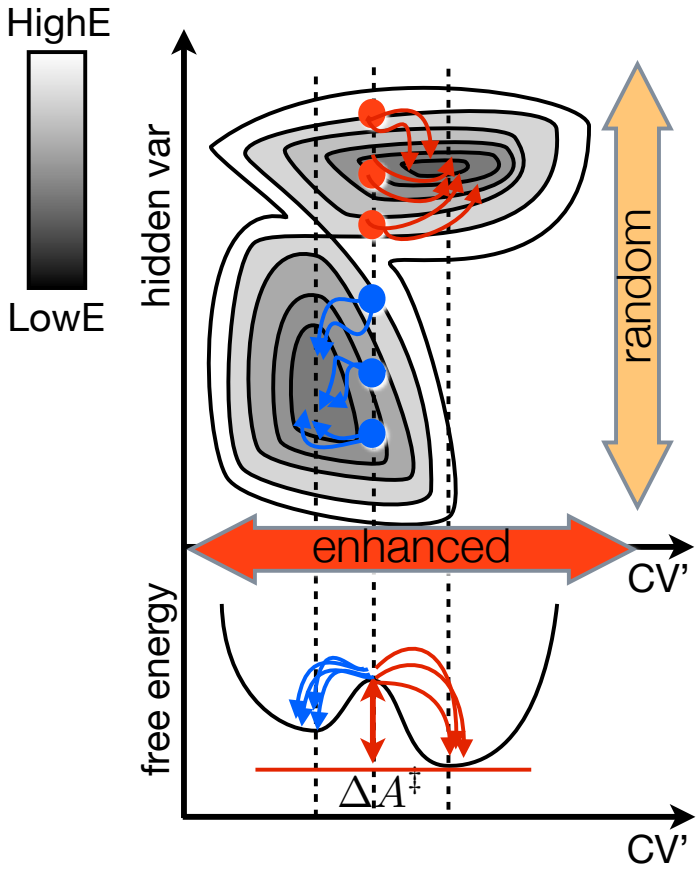
Misconceptions in free energy calculations

Technique (TI, WHAM, SMD, MetaD, ABF?)

Choice of order parameter

The importance of the order parameter

Different order parameters give different landscapes



Transition state is mixed pre post

Poor convergence of free energy

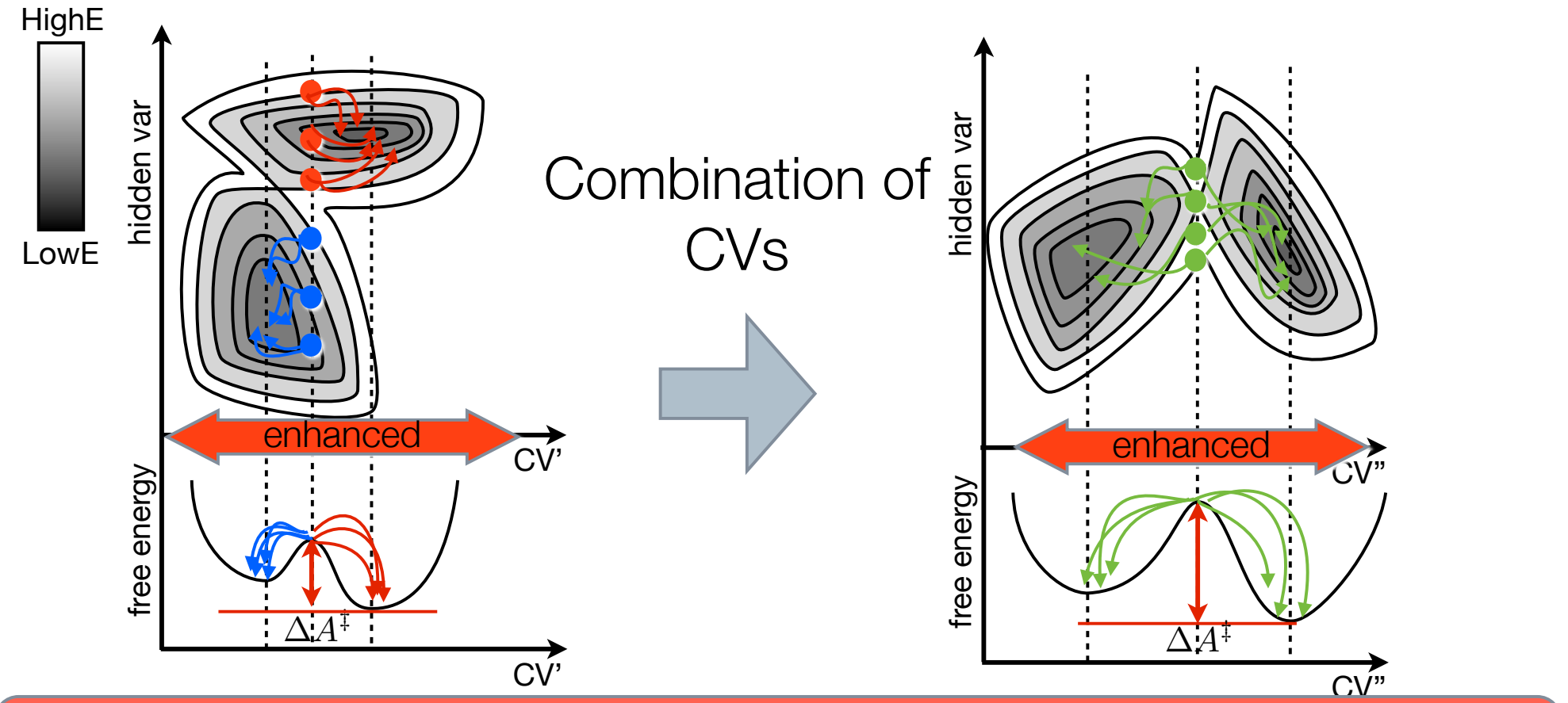
Unreliable transition rates

$$k(T) = \frac{k_B T}{hc^0} e^{-\Delta A^{\ddagger}/RT}$$

Hard to compare with Expt!

How to choose the order parameter

Different order parameters give different landscapes



PRO: the best perspective has good convergence and rates!
CONS: it is a trial and error procedure!

Path free energy calculations

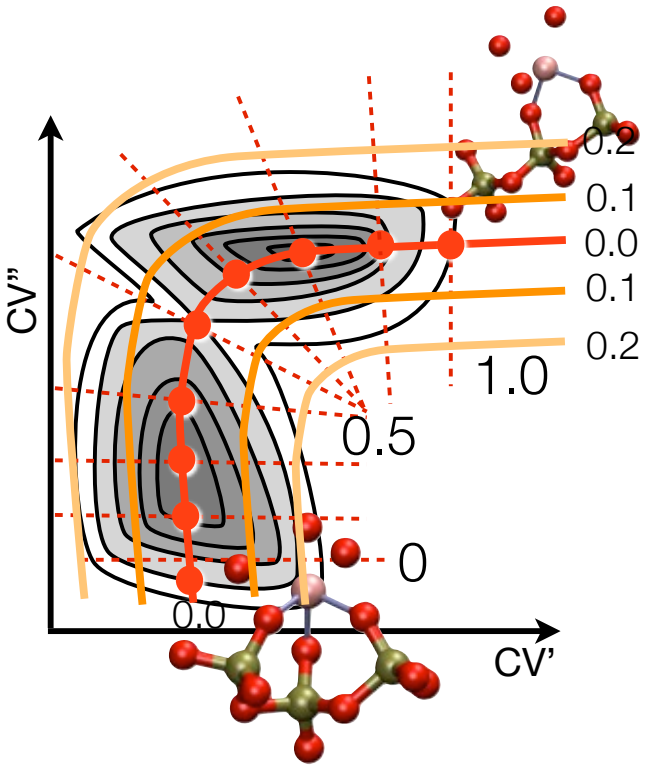
* Monodimensional nonlinear mapping of the transition in terms of many coordinates

Distance from the path

selected path snapshots in a given representation

actual MD snapshot

$$z(\mathbf{x}) = -\frac{1}{\lambda} \ln \sum_{i}^{P} e^{-\lambda |\mathbf{x} - \mathbf{x}_i|}$$

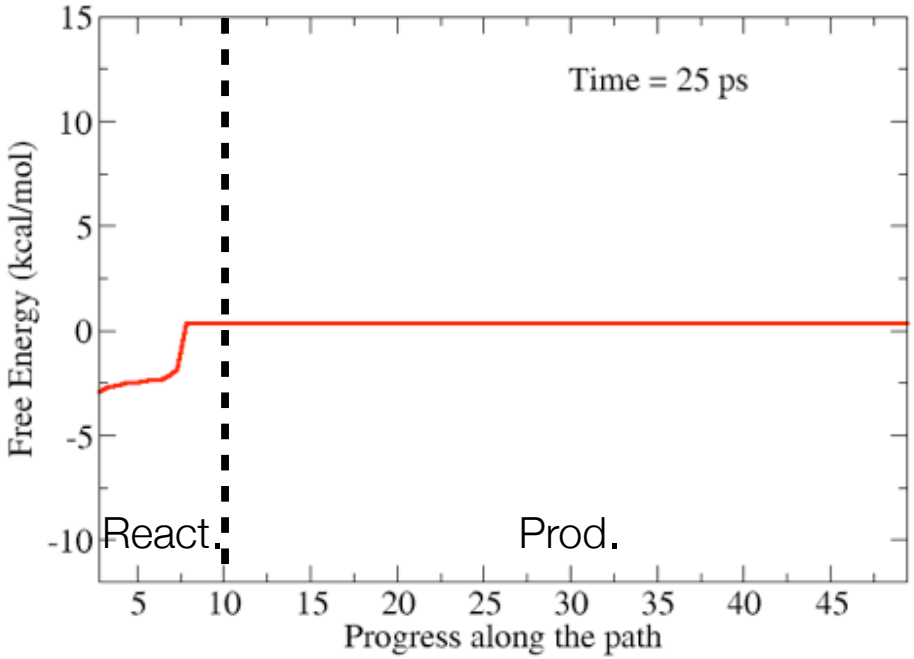


* Adaptive (it uses information from a previous transition)
* Preserves low dimensional description for efficiency
* Allows extensions (distance from path)
* Beyond metadynamics: Steered molecular dynamics and others

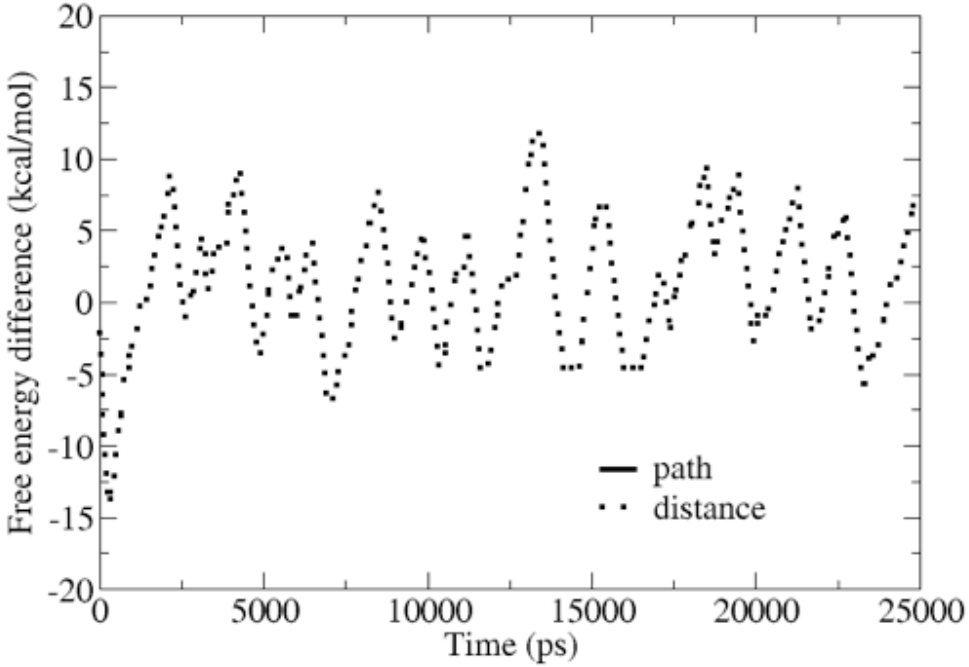
[8] Branduardi, Gervasio, Parrinello,J. Chem. Phys. 2007, vol.126, p. 054103.

Path collective variables: ATP-Mg^{2+}

Free energy estimate



$A_{React.}$-$A_{Prod.}$



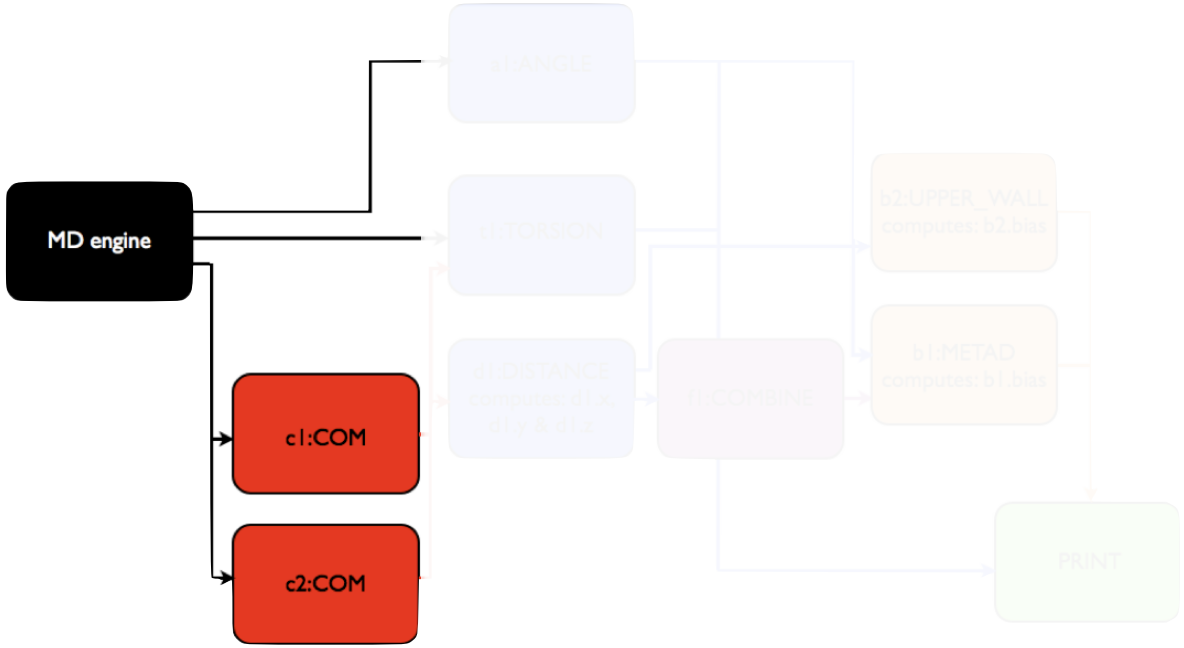
the shape is conserved, the free energy difference has smaller fluctuations!

PLUMED provides many CVs: are they enough?

ANGLE	Calculate an angle.
CELL	Calculate the components of the simulation cell
CONSTANT	Return a constant quantity.
CONTACTMAP	Calculate the distances between a number of pairs of atoms and transform each distance by a switching function.
COORDINATION	Calculate coordination numbers.
CS2BACKBONE	This collective variable calculates the backbone chemical shifts for a protein (CA, CB, C', H, HA, N) using the CamShift method Kohlhoff:2009us. The backcalculated chemical shifts are then compared with a set of provided chemical shifts to generate a score Robustelli:2010dn Granata:2013dk.
DHENERGY	Calculate Debye-Huckel interaction energy among GROUPA and GROUPB.
DIPOLE	Calcualte the dipole moment for a group of atoms.
DISTANCE	Calculate the distance between a pair of atoms.
DRMSD	Calculate the distance RMSD with respect to a reference structure.
ENERGY	Calculate the total energy of the simulation box.
FAKE	This is a fake colvar container used by cltools or various other actions and just support input and period definition
GYRATION	Calculate the radius of gyration, or other properties related to it.
PATHMSD	This Colvar calculates path collective variables.
PROPERTYMAP	Calculate generic property maps.
RMSD	Calculate the RMSD with respect to a reference structure.
TEMPLATE	This file provides a template for if you want to introduce a new CV.
TORSION	Calculate a torsional angle.
VOLUME	Calculate the volume of the simulation box.
ALPHARMSD	Probe the alpha helical content of a protein structure.
ANTIBETARMSD	Probe the antiparallel beta sheet content of your protein structure.
PARABETARMSD	Probe the parallel beta sheet content of your protein structure.

Plumed provides many CVs : combine with functions

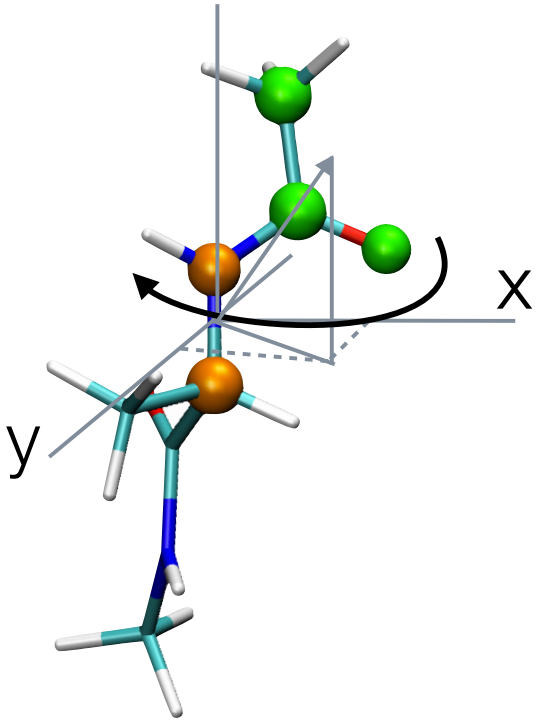
```
c1: COM ATOMS=1-10
c2: COM ATOMS=30-40
d1: DISTANCE ATOMS=c1,c2 COMPONENTS
a1: ANGLE ATOMS=14,15,16
f1: COMBINE ARG=d1.x,d1.y,d1.z POWERS=2,2,2
t1: TORSION ATOMS=20,c1,c2,23
b1: METAD ARG=f1,a1 PACE=20 HEIGHT=0.5 SIGMA=0.05,0.1
b2: UPPER_WALL ARG=d1.z AT=1.0 KAPPA=0.1
PRINT ARG=a1,t1,b1.bias,b2.bias FILE=colvar STRIDE=100
```



Plumed provides many CVs : matheval

```
# time and functions of time
t: TIME
cos: MATHEVAL ARG=t VAR=t FUNC=cos(t/2.-1.02849029) PERIODIC=NO
sin: MATHEVAL ARG=t VAR=t FUNC=sin(t/2.-1.02849029) PERIODIC=NO
# the center of mass of  two groups:the axis
c3: COM ATOMS=7,9
# the rotating group: center of mass
c4: COM ATOMS=1,5,6
# this produces the components x,y,z
d: DISTANCE COMPONENTS ATOMS=c4,c3
## running cosine and sine values
mycos:  MATHEVAL ARG=d.x,d.y  VAR=x,y   FUNC=x/sqrt(x*x+y*y) PERIODIC=NO
mysin:  MATHEVAL ARG=d.x,d.y  VAR=x,y   FUNC=y/sqrt(x*x+y*y) PERIODIC=NO
## this is the potential function
MATHEVAL ...
  ARG=mycos,mysin,cos,sin
  VAR=mc,ms,c,s
  FUNC=100*((mc-c)^2+(ms-s)^2)
  PERIODIC=NO
  LABEL=vv1
... MATHEVAL
## add a bias on the value calculated by vv1
cc: BIASVALUE ARG=vv1
PRINT ARG=cos,sin,mycos,mysin,cc.bias.vv1 STRIDE=1 FILE=colvar FMT=%8.4f
ENDPLUMED
```

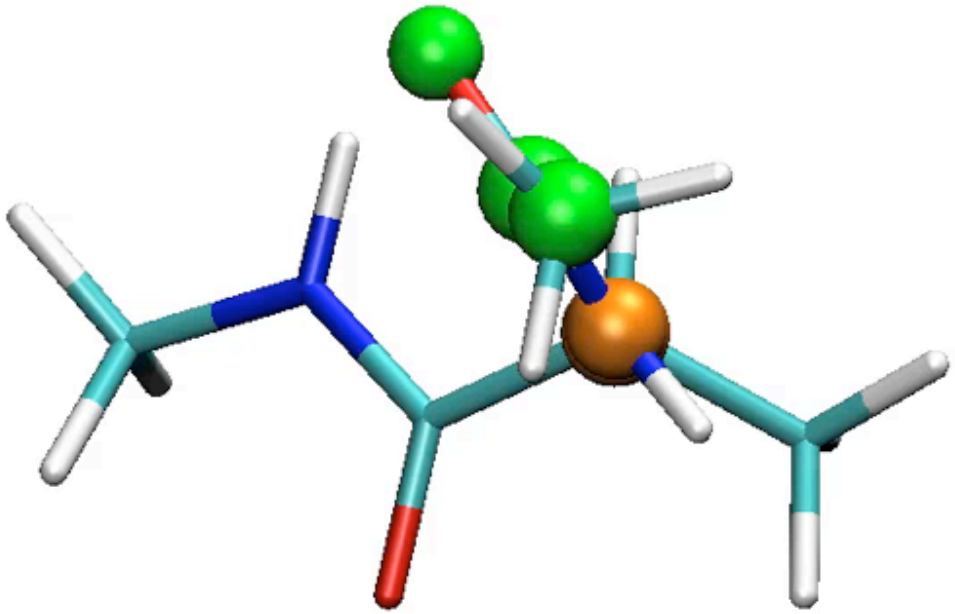
$$c(t) = cos(0.5t - 1.028)$$
$$s(t) = sin(0.5t - 1.028)$$



$$U(\mathbf{X}, t) = 100 * [(s(t) - sin(\mathbf{X}))^2 + (c(t) - cos(\mathbf{X}))^2]$$

http://www.gnu.org/software/libmatheval/

Plumed provides many CVs : matheval



Plumed is also a set of tools

```
davidebranduardi: ~/Programs/plumed2 >plumed help
Usage: plumed [options] [command] [command options]
  plumed [command] -h       : to print help for a specific command
Options:
  [help|-h|--help]          : to print this help
  [--has-mpi]               : fails if plumed is running without MPI
  [--has-matheval]          : fails if plumed is compiled without matheval
  [--has-almost]            : fails if plumed is compiled without almost
  [--has-dlopen]            : fails if plumed is compiled without dlopen
  [--load LIB]              : loads a shared object (typically a plugin library)
  [--standalone-executable] : tells plumed not to look for commands implemented as scripts
Commands:
  plumed driver : analyze trajectories with plumed
  plumed driver-float : analyze trajectories with plumed (single precision version)
  plumed gentemplate : print out a template input for a particular action
  plumed info : provide informations about plumed
  plumed manual : print out a description of the keywords for an action in html
  plumed simplemd : run lj code
  plumed sum_hills : sum the hills with  plumed
  plumed mklib : compile a .cpp file into a shared library
  plumed newcv : create a new collective variable from a template
  plumed patch : patch an MD engine
davidebranduardi: ~/Programs/plumed2 >
```

Summary

* Enhanced sampling method makes you cross barriers
* Free energies can be calculated from biased dynamics
* Free energies depends on the technique and from the order parameter
* Order parameters are as important as the accelerated sampling method
* PLUMED offers a flexible tool to access a variety of methods and collective variables
* Beyond what I shown here: mixed replica exchange/biased dynamics (PT-WTE, PT-MetaD, Bias Exchange)
* Even more: String methods

Acknowledgements

Giovanni Bussi (SISSA, Trieste)

Max Bonomi (UCSF, San Francisco)

Gareth Tribello (Queen's University, Belfast)

Carlo Camilloni (Cambridge University)

Michele Parrinello (ETH, Zurich)