The PLUMED plugin and free energy methods in electronic-structure-based molecular dynamics

Davide Branduardi,
Theoretical Molecular Biophysics Group
Max Planck for Biophysics, Frankfurt am Main, Germany

Outline

- What is free energy and why is important
- Traditional QC approach
- Explicit sampling
- Enhanced sampling
- PLUMED plugin for free energy calculations

Outline

- **What is free energy and why is important**
- Traditional QC approach
- Explicit sampling
- Enhanced sampling
- PLUMED plugin for free energy calculations

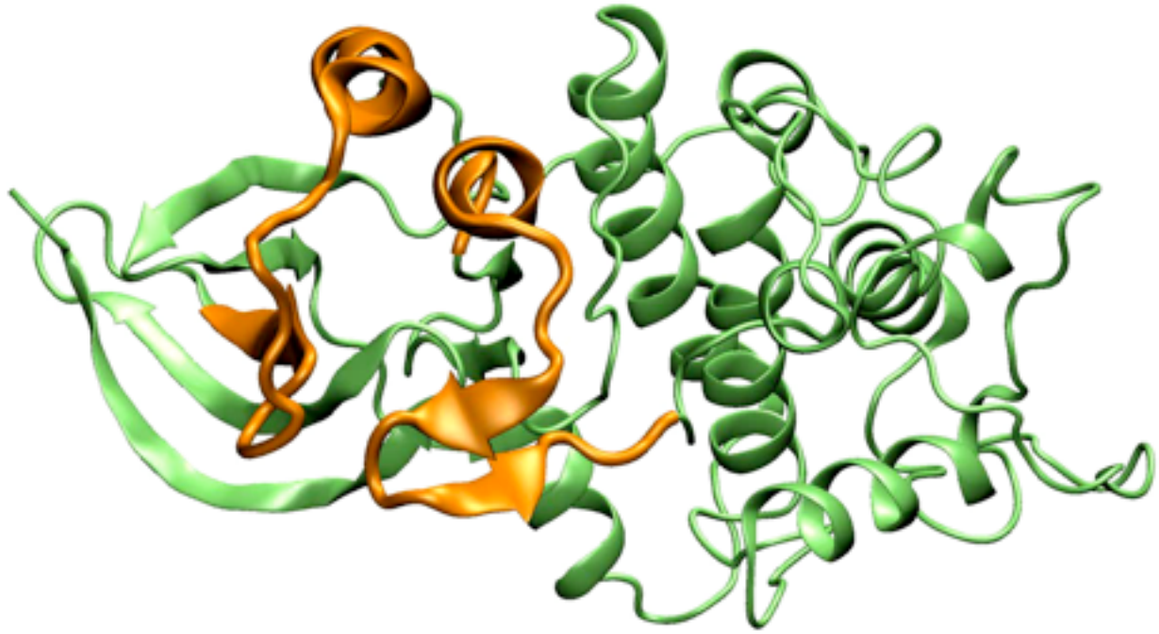
What is free energy and why is important

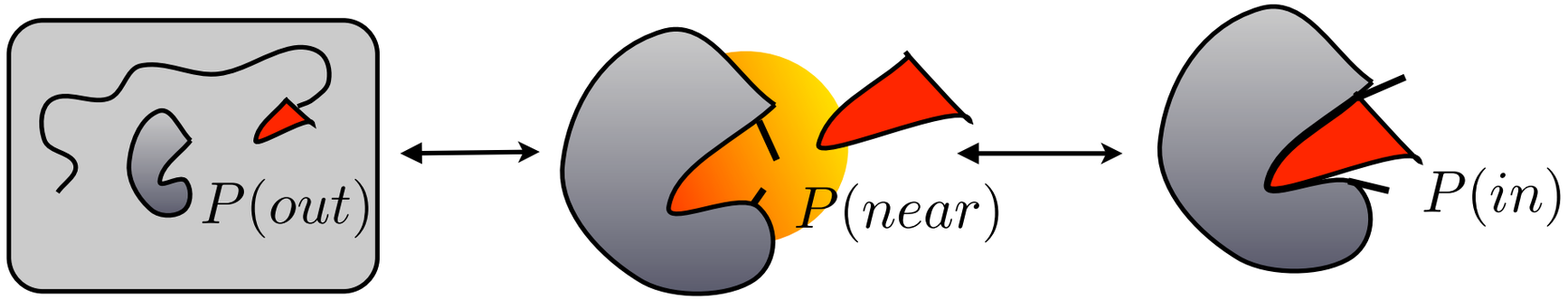
$$F(s) = -k_B T \ln P(s)$$

s= state descriptor (bound/unbound, reactant/products, phaseA/phaseB)

$$P(s) = \frac{\int e^{-\frac{V(x)}{k_b T}} \delta(s'(x) - s) dx}{Q}$$

* conformational transitions and equilibria
* ligand binding
* mechanism of transporters or channels
* chemical reactions and catalysis
* phase transitions
* .....





Outline

- What is free energy and why is important
- **Traditional QC approach**
- Explicit sampling
- Enhanced sampling
- PLUMED plugin for free energy calculations

Traditional QC approach

* Rigid rotor/ harmonic approximation is well known [1]

$$F(s) = H(s) - TS(s)$$

$$S(s) = R + R\ln\boxed{(q_t(s)q_e(s)q_r(s)q_v(s))} + RT\left(\frac{\partial \ln \boxed{q_t(s)q_e(s)q_r(s)q_v(s)}}{\partial T}\right)$$

trasl, elect, rot, vib partition functions

where the partition functions are obtained from the calculated vibrational spectrum (Hessian), moments of inertia, electronic structure

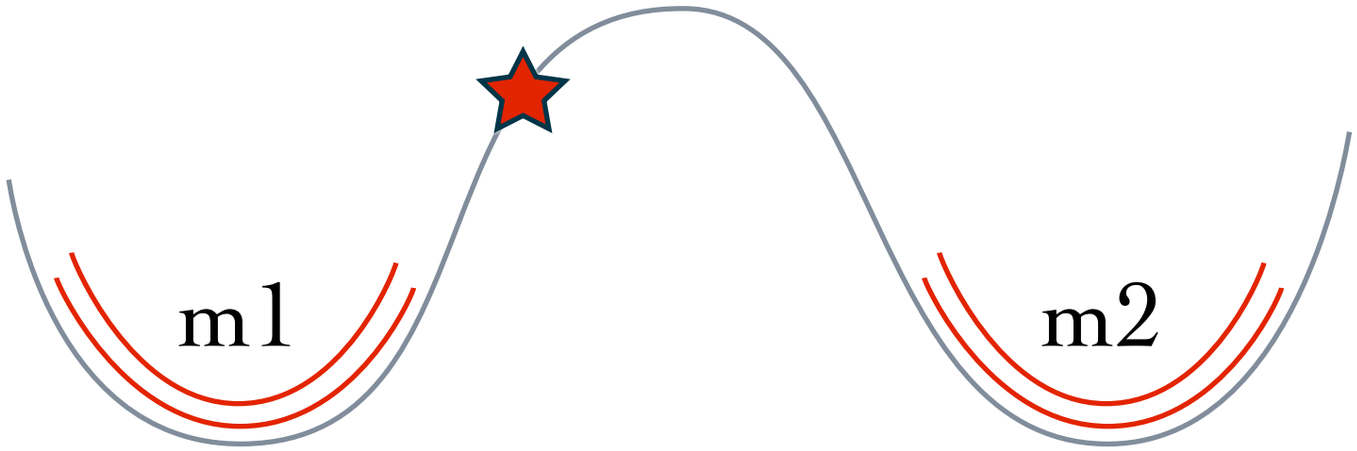
$$H(s) = H_{el}(s) + H_t(s) + H_v(s) + H_r(s)$$

* It requires electronic structure and Hessian in the local minima and transition states: 1 single conformation per state.

[1] McQuarrie, Simon “Physical Chemistry, a molecular approach”

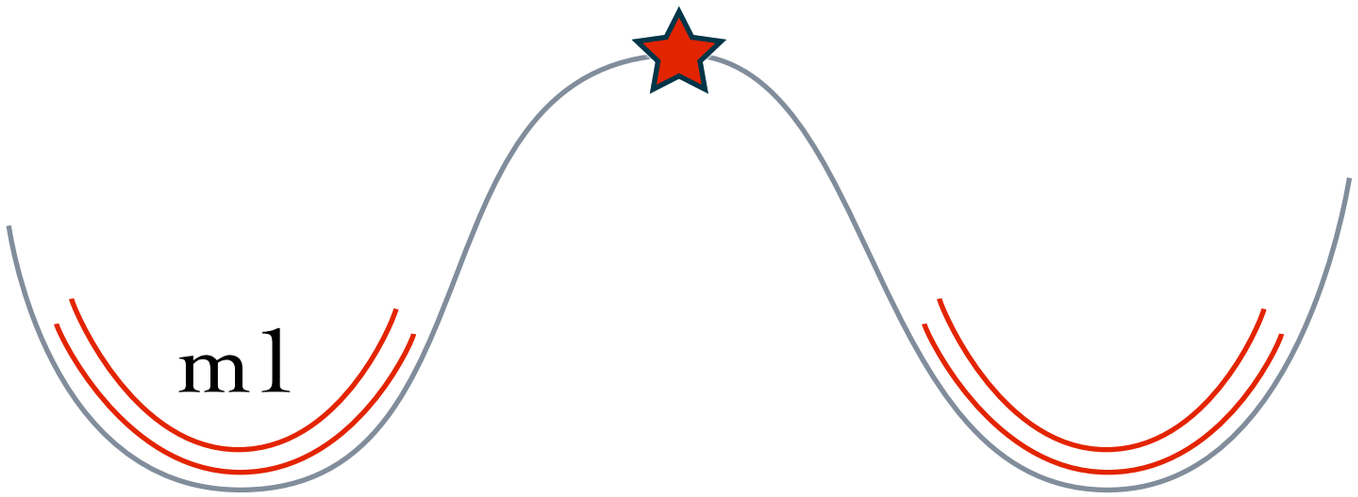
Finding the RC (I)

- Static approach: look for saddle points by using chemical intuition and TS optimization.



Finding the RC

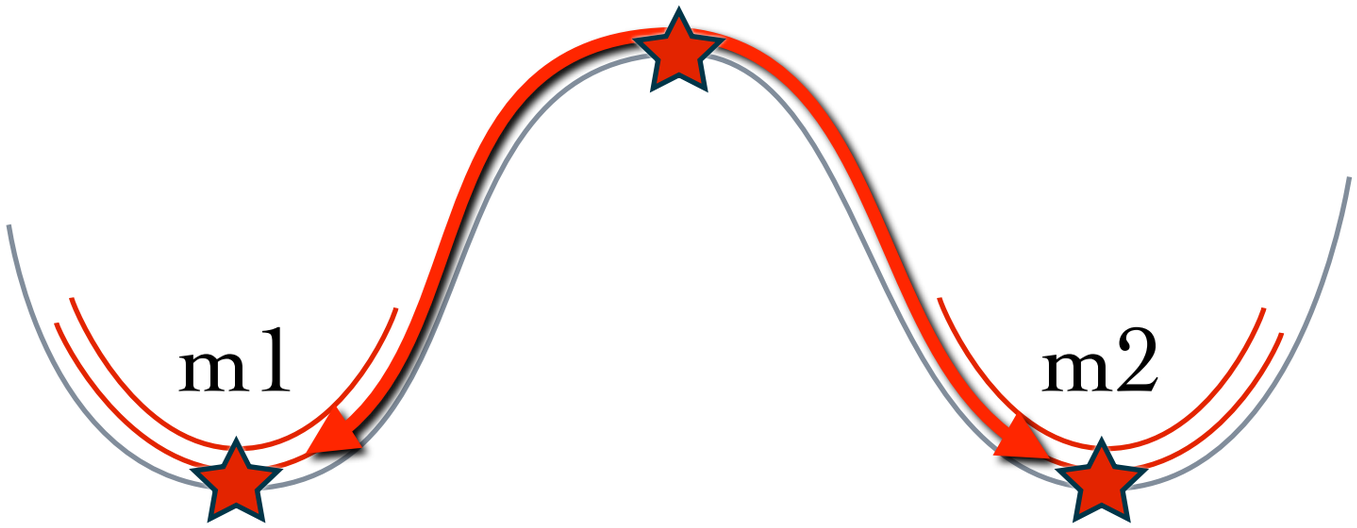
- Static approach: look for saddle points by using chemical intuition and TS optimization.



- Verify then by IRC: move forward and backward to verify to end in minima.

Finding the RC

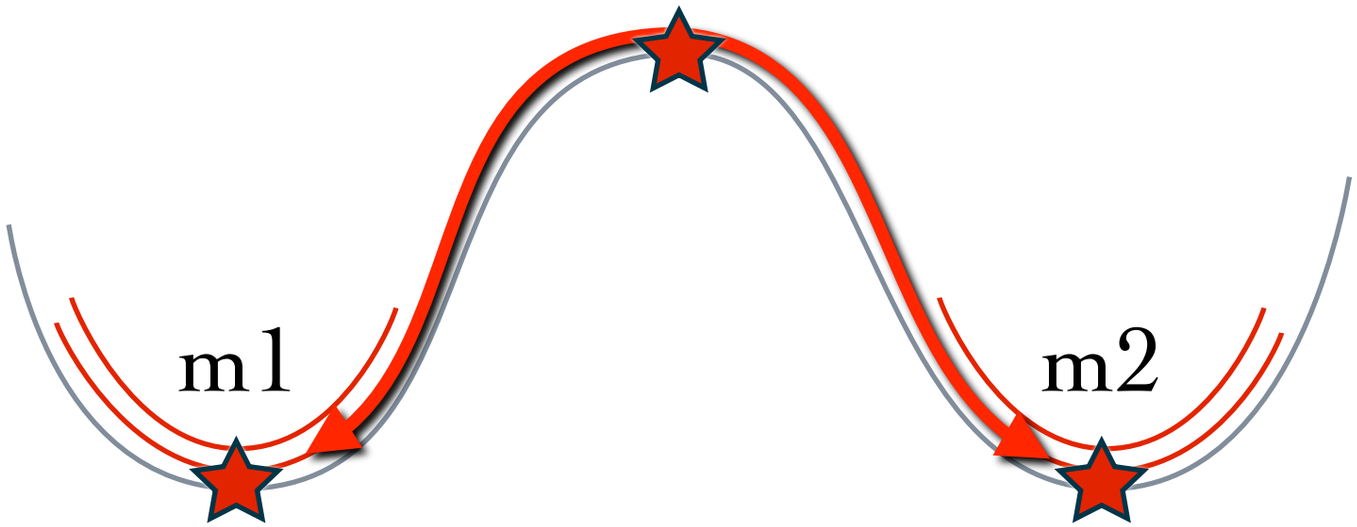
- Static approach: look for saddle points by using chemical intuition and TS optimization.



- Verify then by IRC: move forward and backward to verify to end in minima.

Finding the RC

- Static approach: look for saddle points by using chemical intuition and TS optimization.



- Verify then by IRC: move forward and backward to verify to end in minima.

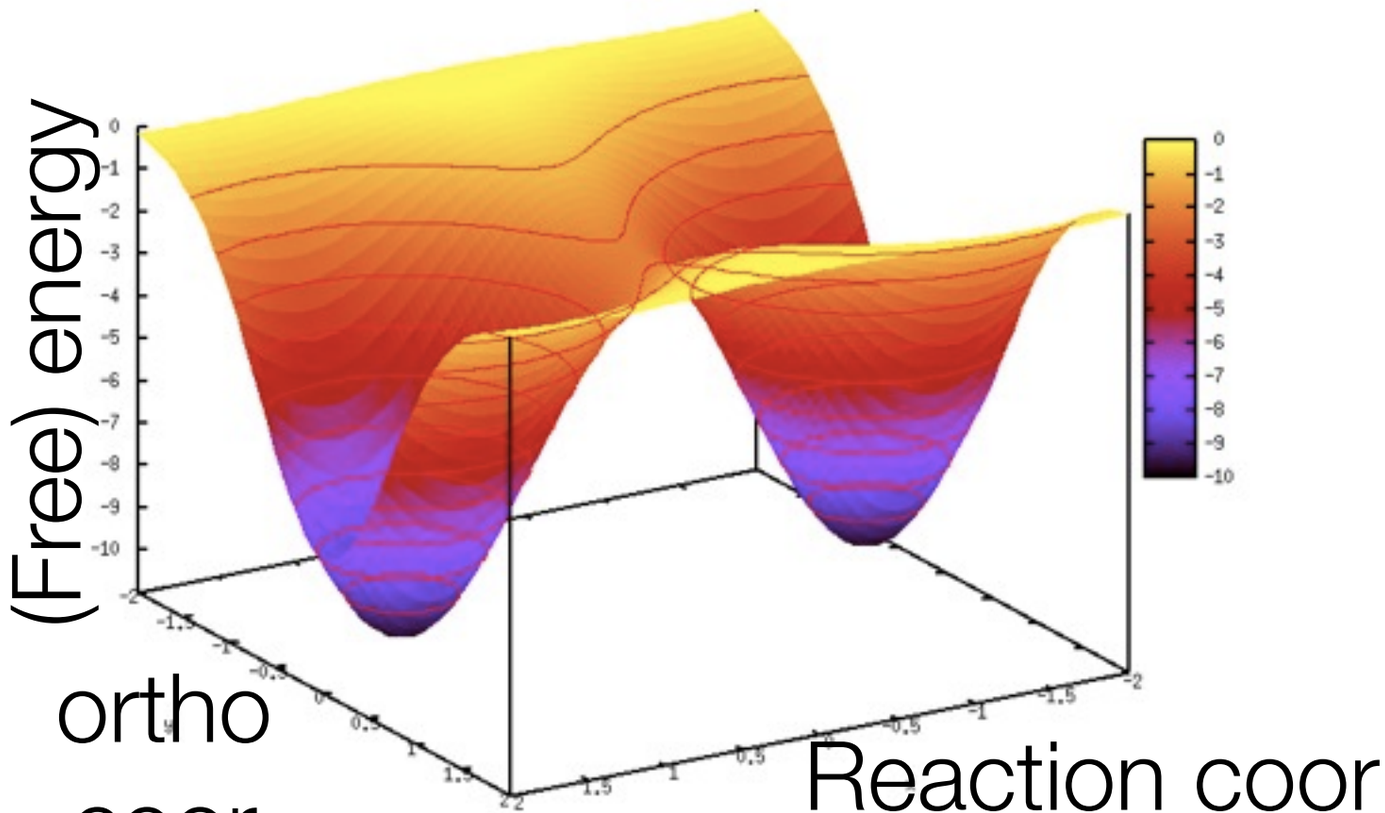
$$\Delta^{\ddagger}G = \Delta^{\ddagger}H - T\Delta^{\ddagger}S$$

- Rates may be derived from Eyring equation[2]: $k(T) = \frac{k_b T}{hc^0} e^{-\Delta^{\ddagger}G/RT}$

[2] Eyring, Chem. Rev. 1935 vol. 17 pp 65

Traditional QC approach

* PRO: one calculation per point (3 single points calculations per rate and free energy differences)
* CONS: optimizing TS is not trivial (redundant coordinates), the landscape must be very simple

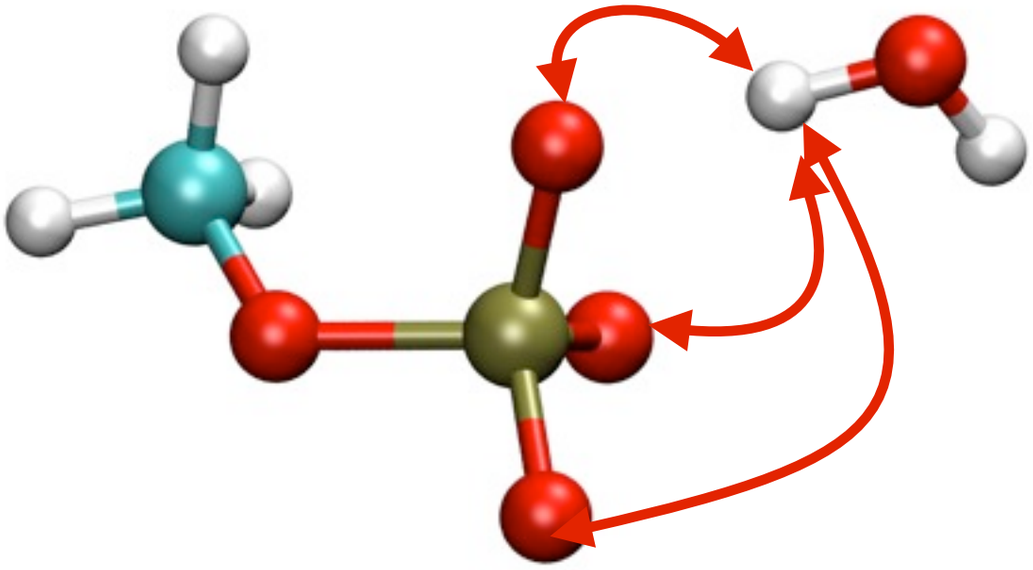


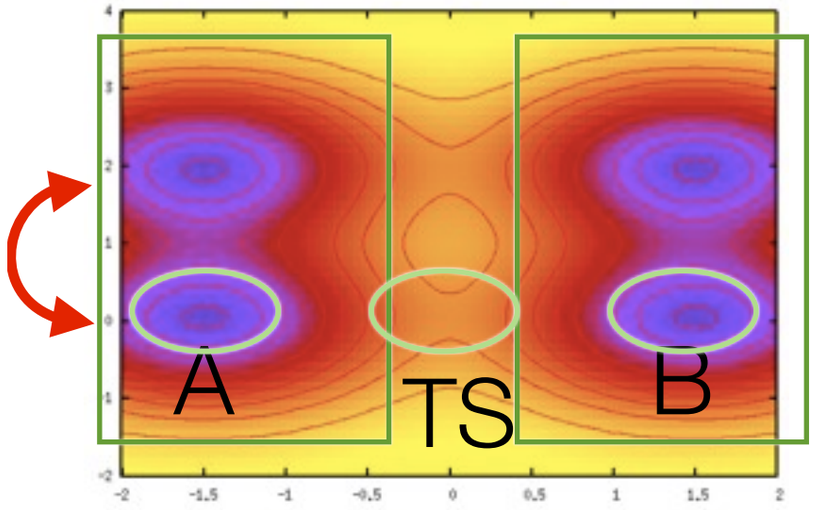
Problems with traditional QC approach

* Many problems are not smooth as before: many parallel valleys (conformers), solvent induced roughness

Problems with traditional QC approach

* Many problems are not smooth as before: many parallel valleys (conformers), solvent induced roughness





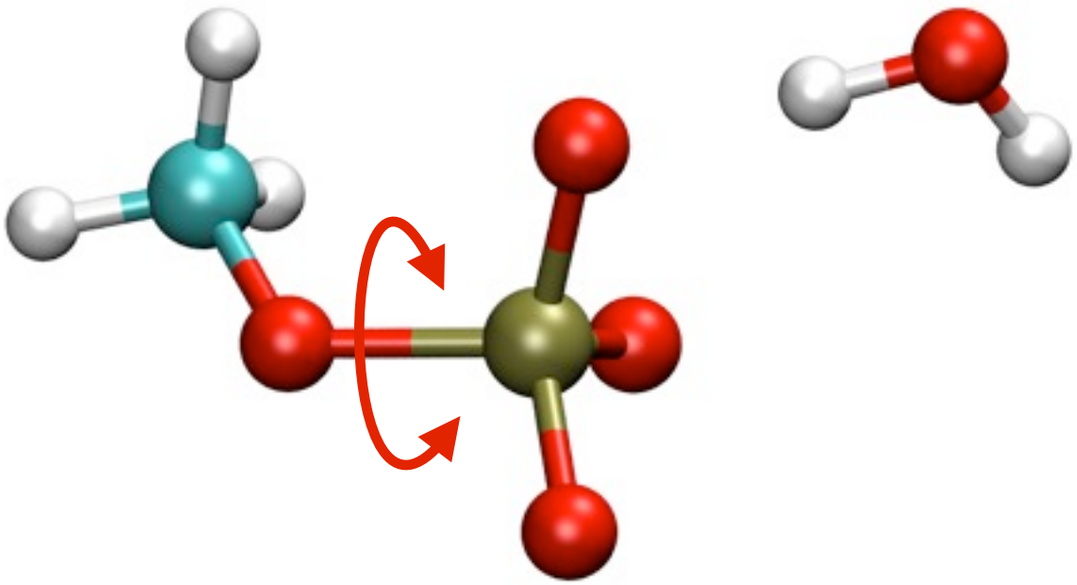
Perfectly parallel basins,
Not too bad
(small correction over
static approaches).
Typical case of symmetry
in a reaction

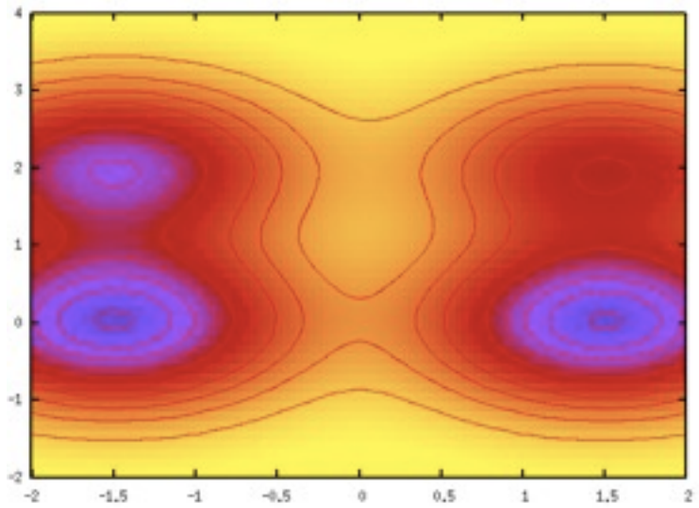
Problems with traditional QC approach

* Many problems are not smooth as before: many parallel valleys (conformers), solvent induced roughness

Problems with traditional QC approach

* Many problems are not smooth as before: many parallel valleys (conformers), solvent induced roughness





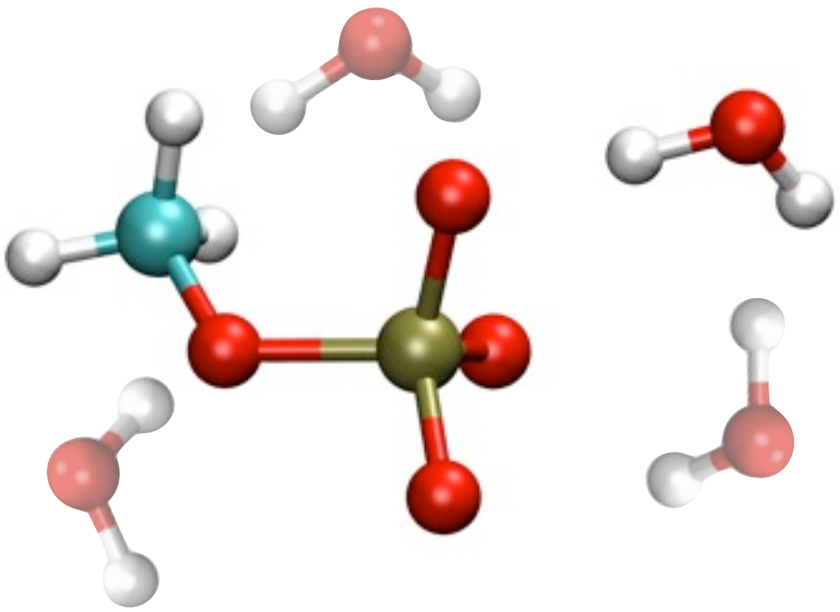
Perfectly parallel basins but energetically asymmetric,
Can be done with static QC but need to sample all the paths

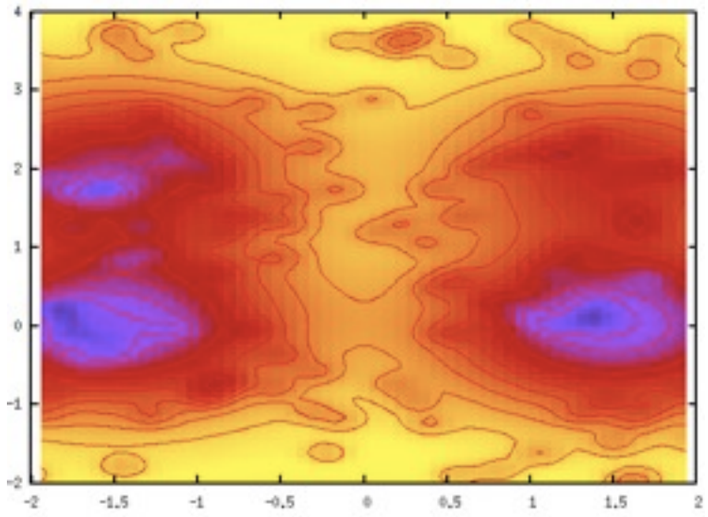
Problems with traditional QC approach

* Many problems are not smooth as before: many parallel valleys (conformers), solvent induced roughness

Problems with traditional QC approach

* Many problems are not smooth as before: many parallel valleys (conformers), solvent induced roughness





Rough paths: vibrations are meaningless, paths from IRC are not representative of reactions:

USE MD BASED METHODS

Problems with traditional QC approach

* Many problems are not smooth as before: many parallel valleys (conformers), solvent induced roughness

Outline

- What is free energy and why is important
- Traditional QM approach
- **Explicit sampling**
- Enhanced sampling
- PLUMED plugin for free energy calculations

The ingredients of the probability picture

* Free energy is connected to the probability of a given state to occur in a single measurement

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

Probability

$$P(s) = \frac{\int e^{-\frac{V(x)}{k_B T}} \delta(s'(x) - s) dx}{\int e^{-\frac{V(x)}{k_B T}} dx}$$

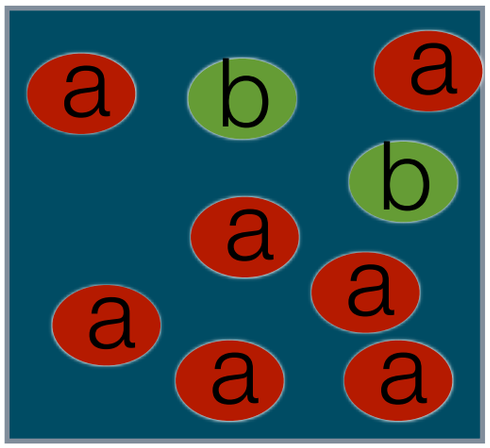
State descriptor

(bond length, folded/ unfolded descr., liquid/ crystal/amorphous)

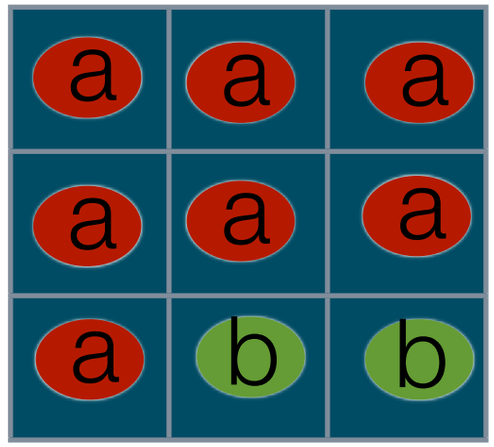
Explicit sampling via molecular dynamics

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

a real system (cuvette)

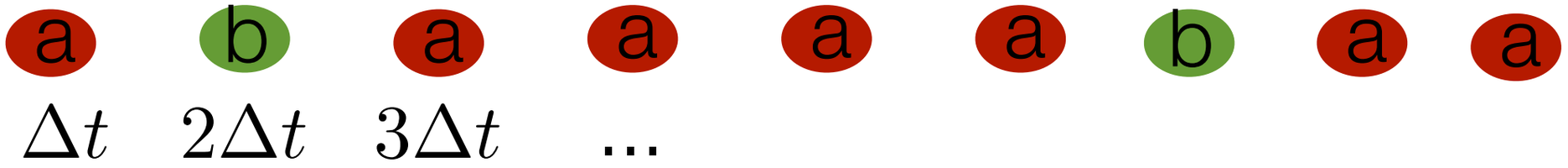


a set of independent objects (ensemble)



MD: evolve in time with Newton's eq. of motion

use ergodic hypothesis

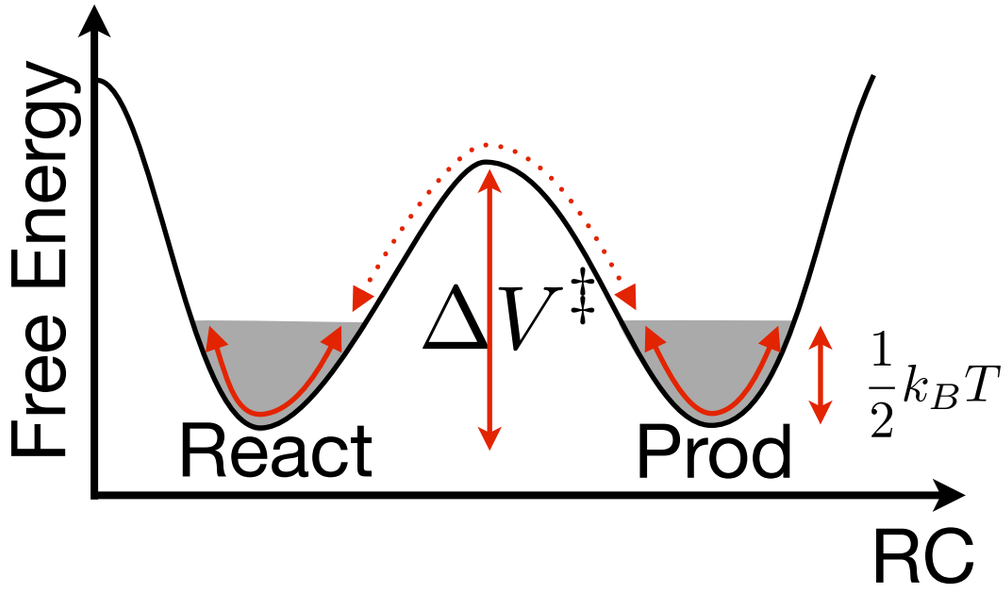


$P(s)$

all the available states can be sampled with t->infinite

Limitations

* Transition of 8 kcal/mol can take up to milliseconds at 300K

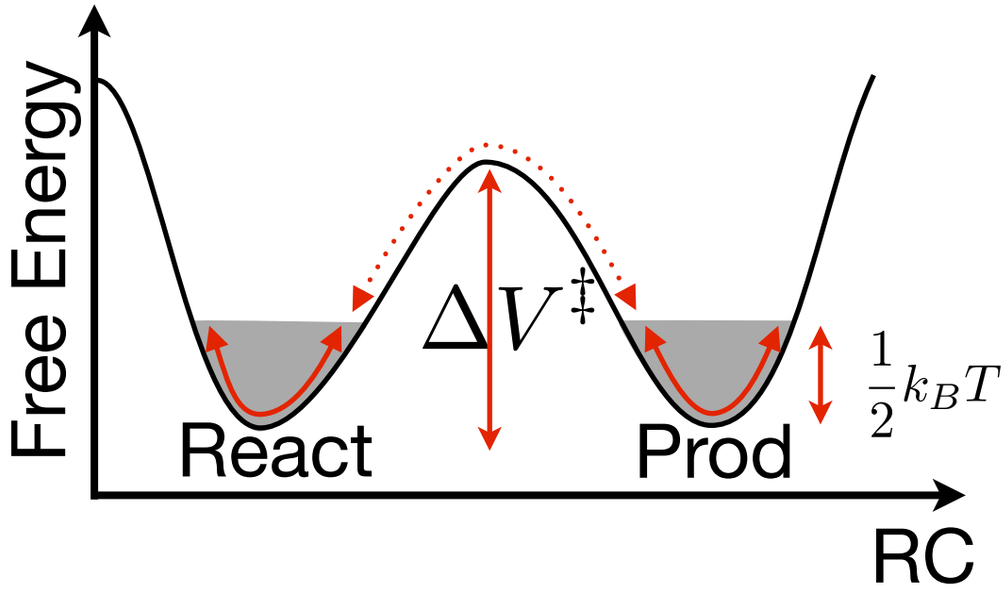


$$\nu = \nu_0 \exp\left(-\beta \Delta V^{\ddagger}\right)$$

$$\nu_0 = 5\ 10^9 s^{-1}$$

Limitations

* Transition of 8 kcal/mol can take up to milliseconds at 300K



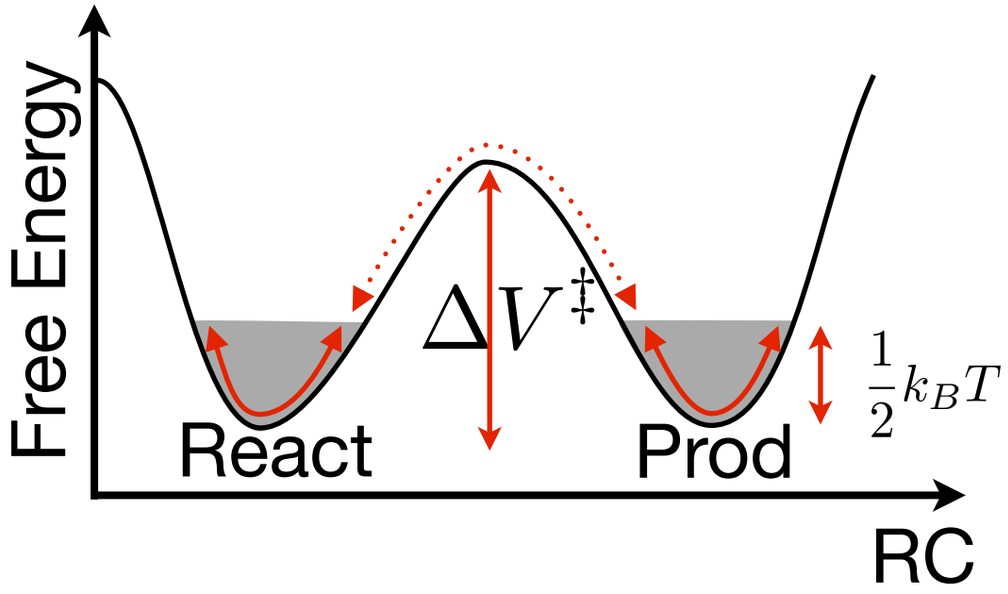
$$\nu = \nu_0 \exp\left(-\beta \Delta V^{\ddagger}\right)$$

$$\nu_0 = 5\ 10^9 s^{-1}$$

* MD (DFT) can reach hundreds of ps.
  Lucky if you see one single event!

Limitations

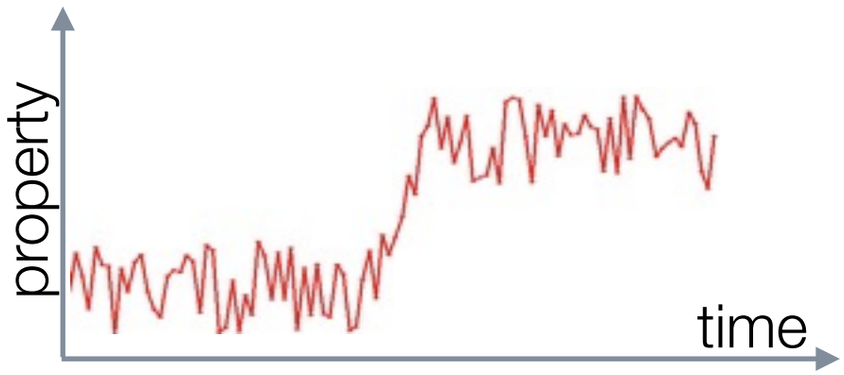
* Transition of 8 kcal/mol can take up to milliseconds at 300K



$$\nu = \nu_0 \exp\left(-\beta \Delta V^{\ddagger}\right)$$

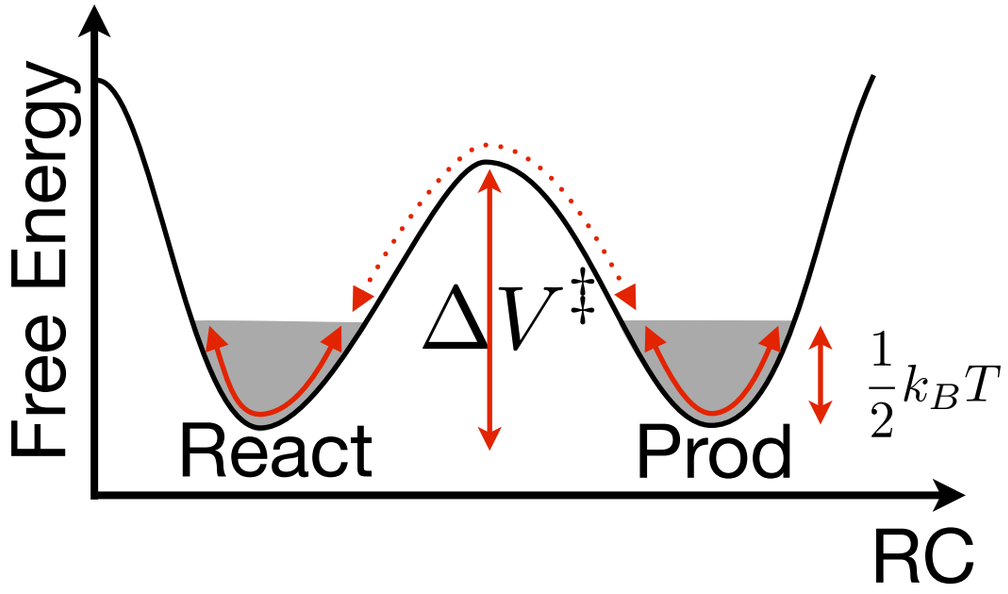
$$\nu_0 = 5\ 10^9 s^{-1}$$

* MD (DFT) can reach hundreds of ps. Lucky if you see one single event!



Limitations

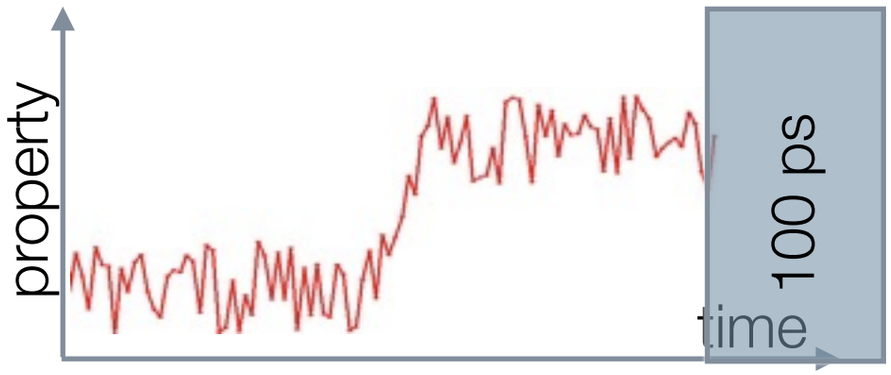
* Transition of 8 kcal/mol can take up to milliseconds at 300K



$$\nu = \nu_0 \exp\left(-\beta \Delta V^{\ddagger}\right)$$

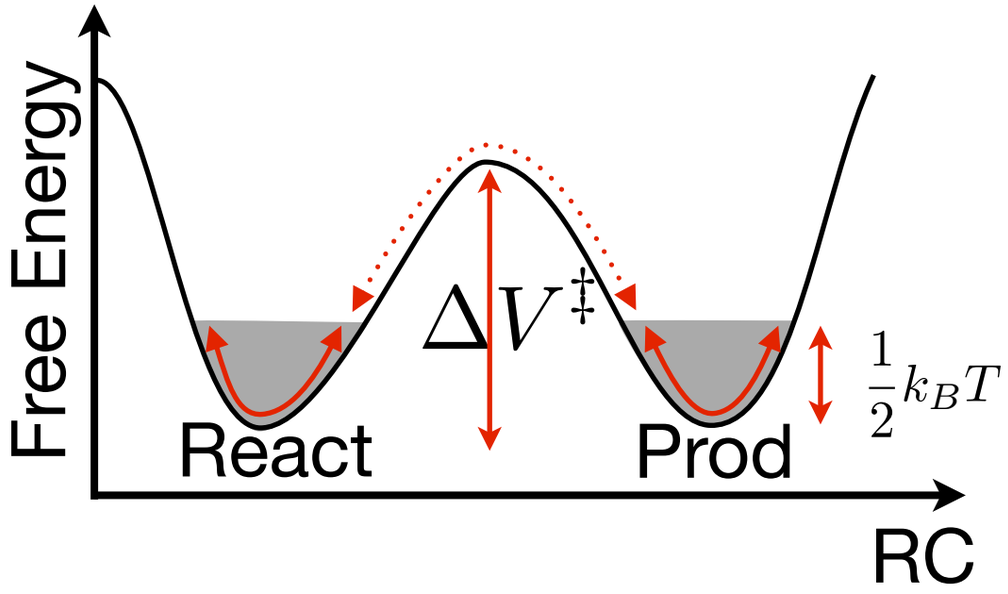
$$\nu_0 = 5\ 10^9 s^{-1}$$

* MD (DFT) can reach hundreds of ps. Lucky if you see one single event!



Limitations

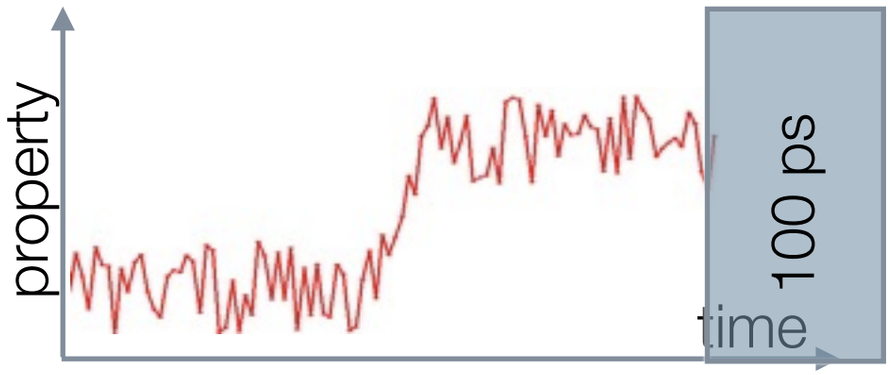
* Transition of 8 kcal/mol can take up to milliseconds at 300K



$$\nu = \nu_0 \exp\left(-\beta \Delta V^{\ddagger}\right)$$

$$\nu_0 = 5\ 10^9 s^{-1}$$

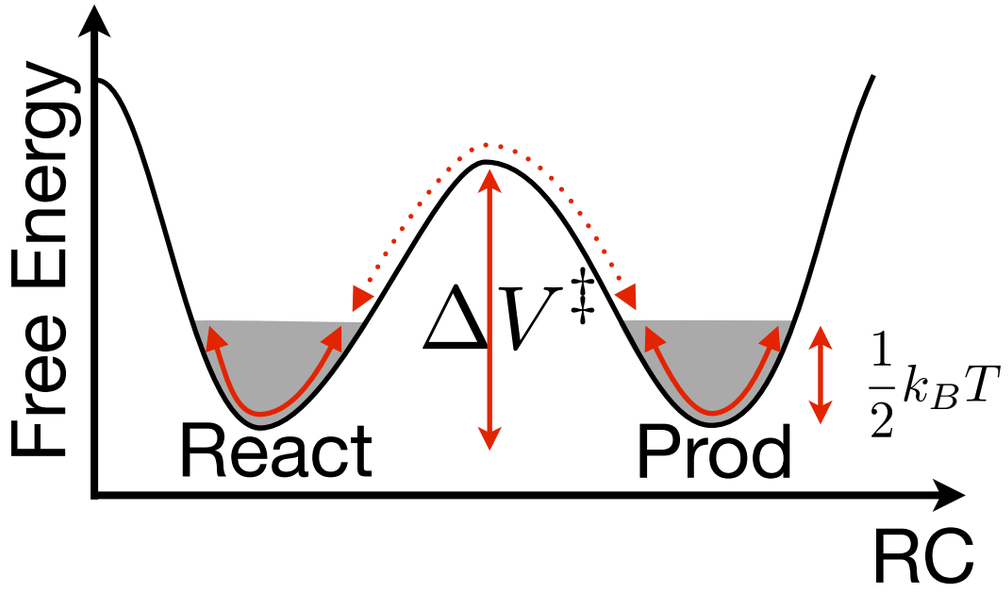
* MD (DFT) can reach hundreds of ps. Lucky if you see one single event!



* If you want thermodynamics you must average over events

Limitations

* Transition of 8 kcal/mol can take up to milliseconds at 300K

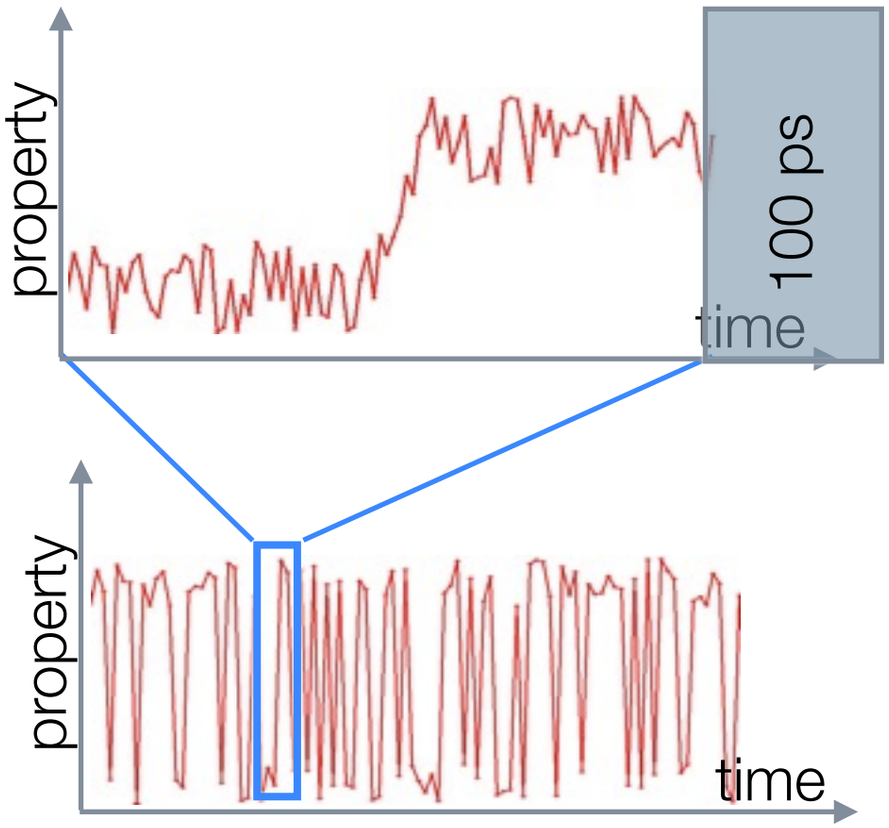


$$\nu = \nu_0 \exp\left(-\beta \Delta V^{\ddagger}\right)$$

$$\nu_0 = 5\ 10^9 s^{-1}$$

* MD (DFT) can reach hundreds of ps. Lucky if you see one single event!

* If you want thermodynamics you must average over events



Limitations

* 300K

* M
  Lu

* If y
  m

This is not your solution

Outline

- What is free energy and why is important
- Traditional QM approach
- Explicit sampling
- **Enhanced sampling**
- PLUMED plugin for free energy calculations

The ancestor: 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

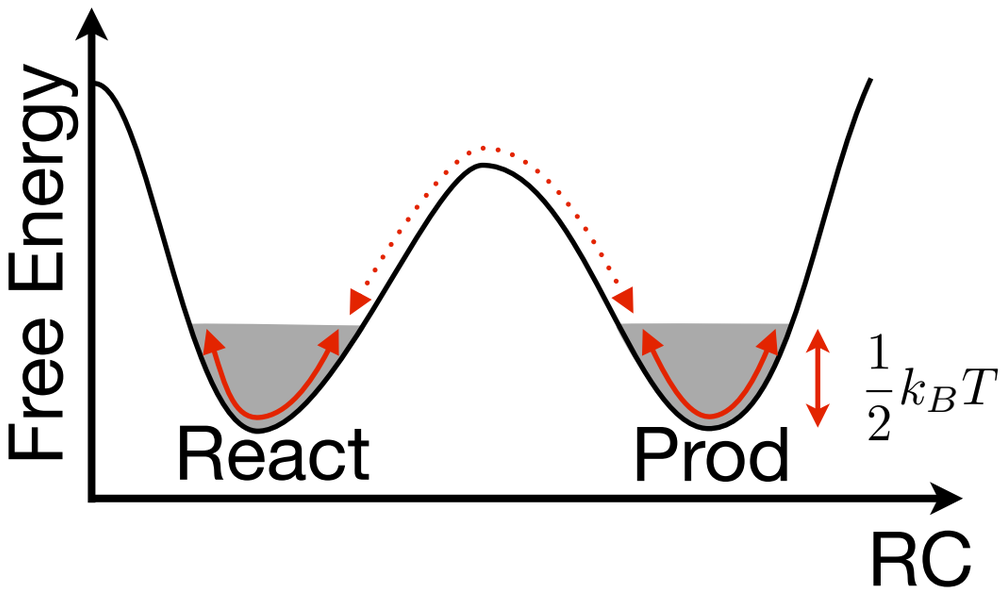
$$F_b(s_0) = V_b(s_0) - \frac{1}{\beta}\ln P(s)\frac{Q}{Q_b}$$

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

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

Lessons from Torrie & Valleau

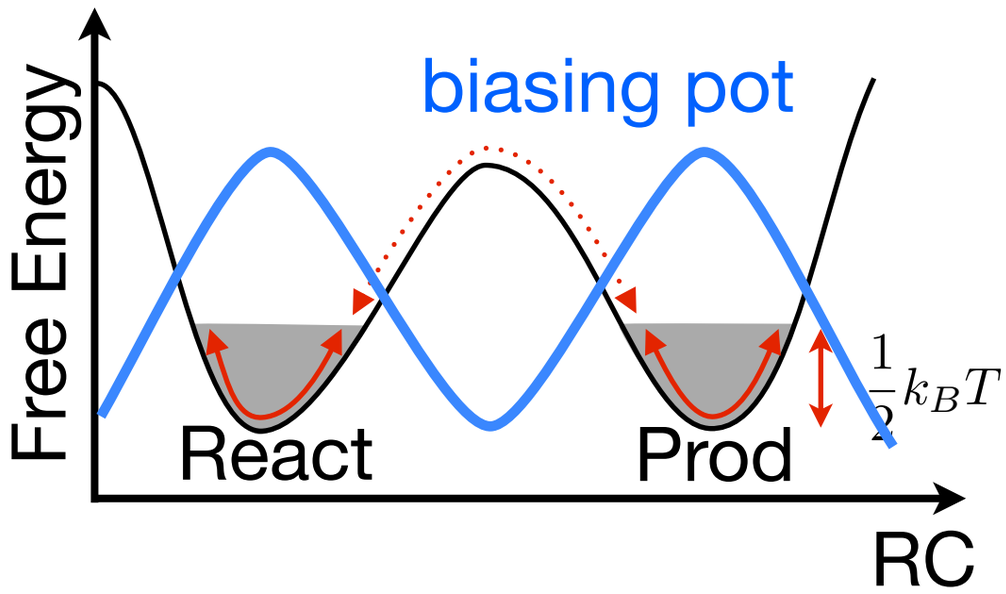
Biasing potentials:



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

Lessons from Torrie & Valleau

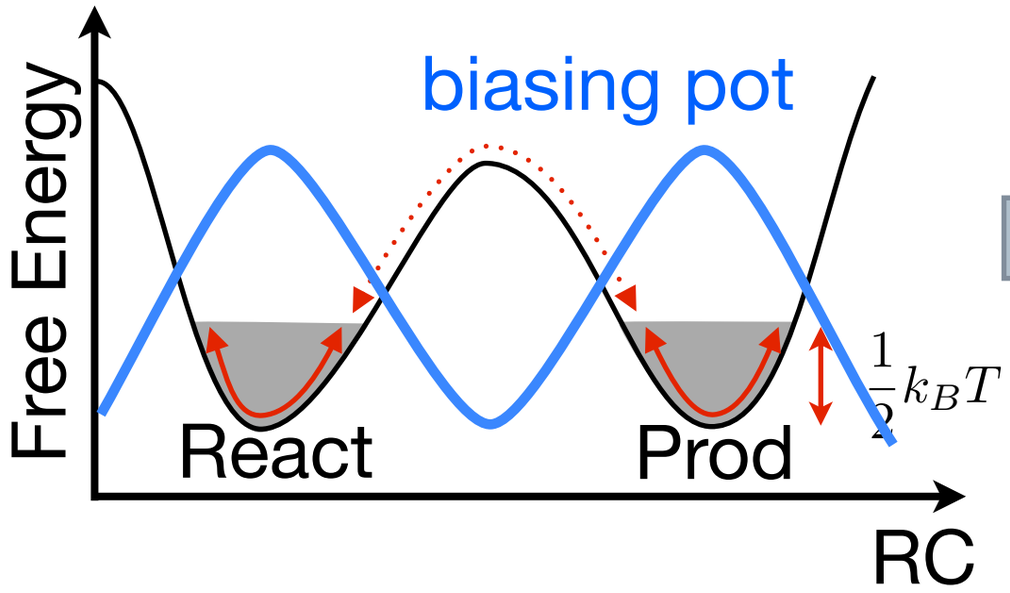
Biasing potentials:



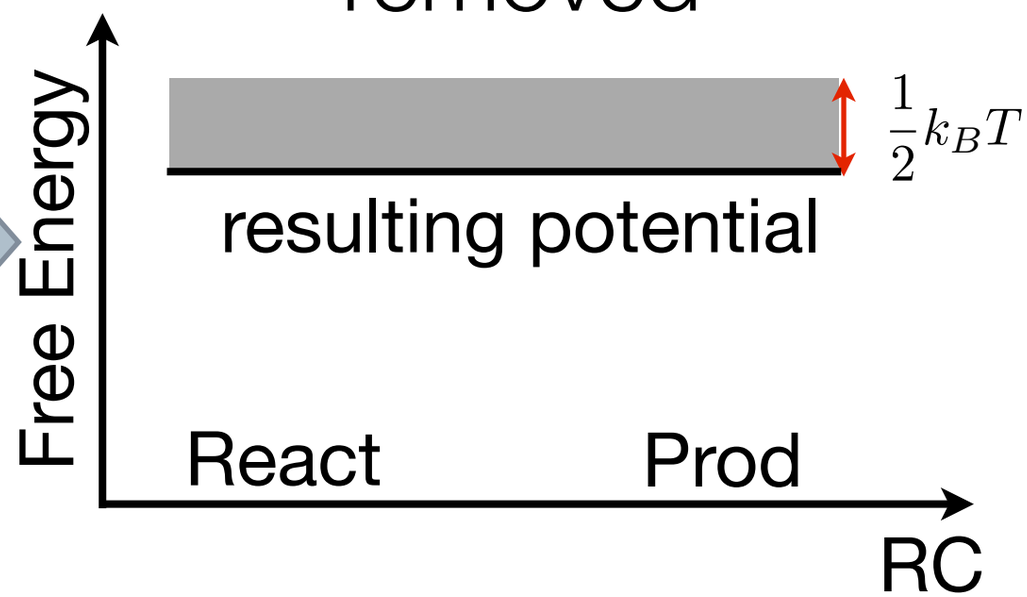
[3] Torrie, Valleau J. Comp. Phys, 1977 vol. 23, pp 187

Lessons from Torrie & Valleau

Biasing potentials:



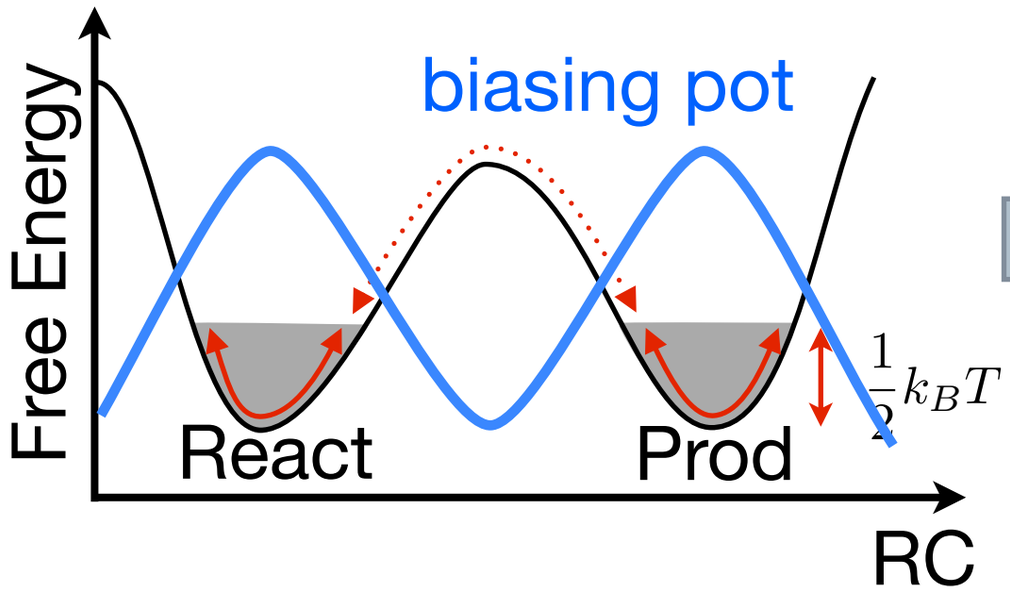
Metastabilities are removed



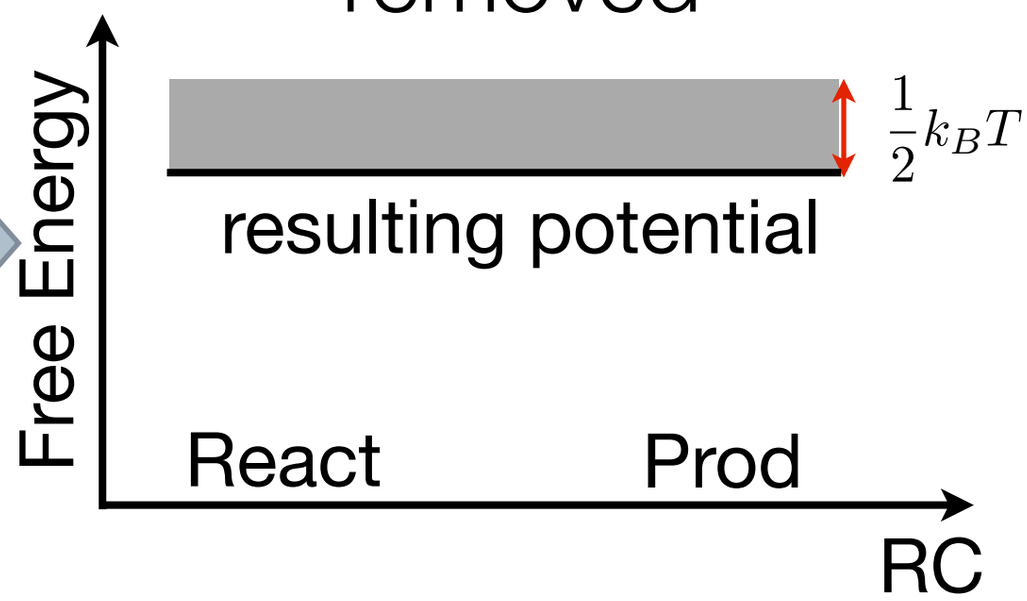
[3] Torrie, Valleau J. Comp. Phys, 1977 vol. 23, pp 187

Lessons from Torrie & Valleau

Biasing potentials:



Metastabilities are removed



* allow to measure free energies: useful
* allow to overcome barriers: cheap
* need only an additional term to standard DFT forces: easy

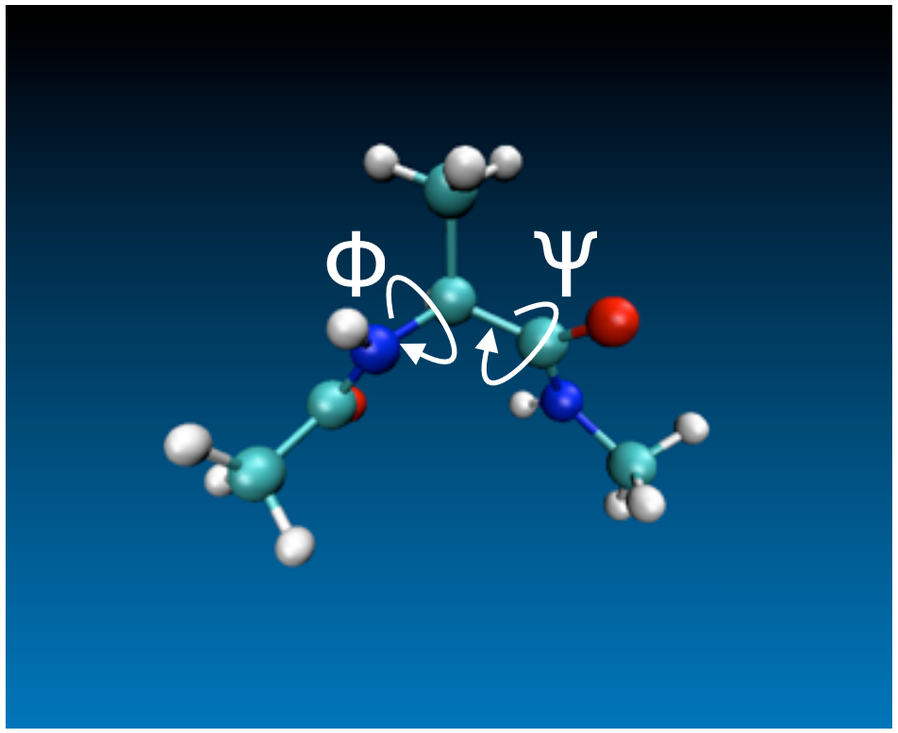
[3] Torrie, Valleau J. Comp. Phys, 1977 vol. 23, pp 187

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.
* MetadynamicsLaio & Parrinello, PNAS 2002, vol. 20, pp. 12562.
* Conformational Flooding Grubmüller, Phys Rev E 1995, vol. 52, pp. 2893.
* ...

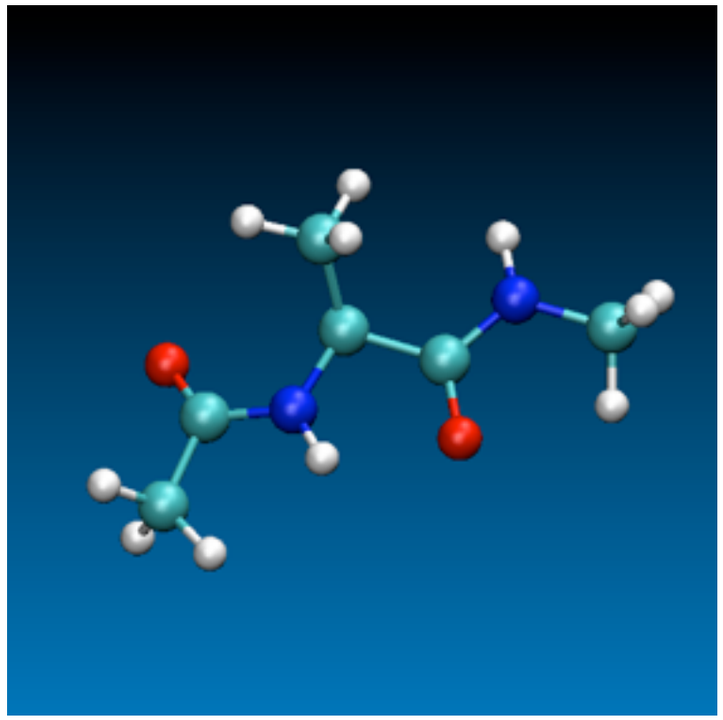
Adaptive sampling methods (e.g. metadynamics)

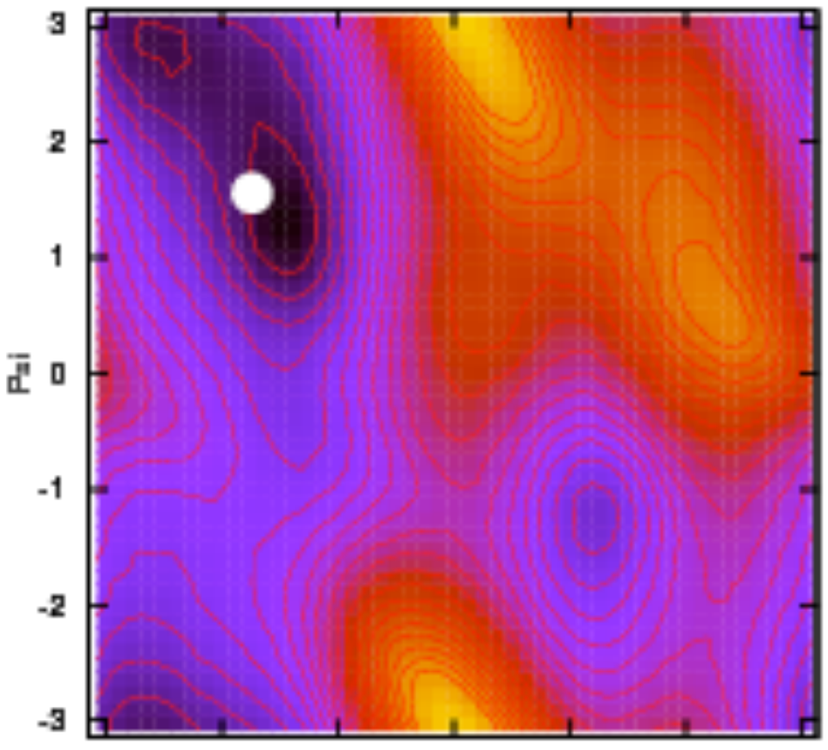
without ES



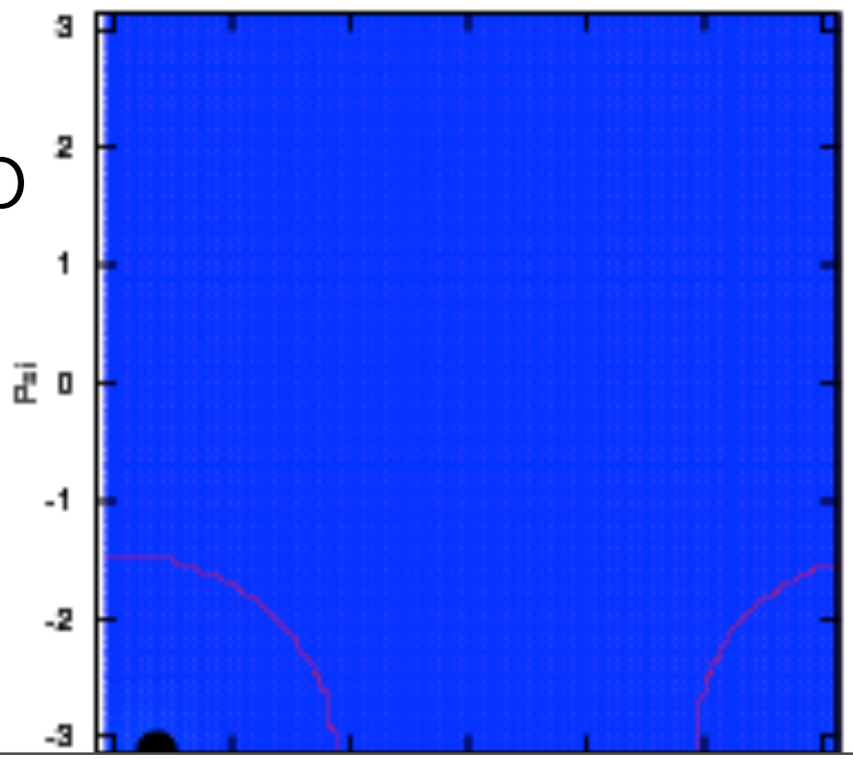
from milli secs...

with ES





to nano secs!



Thermodynamic-Integration-like approaches

* The ancestor is mean force calculation through harmonic potential

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

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

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

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

* It is an average of a biased simulation!

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

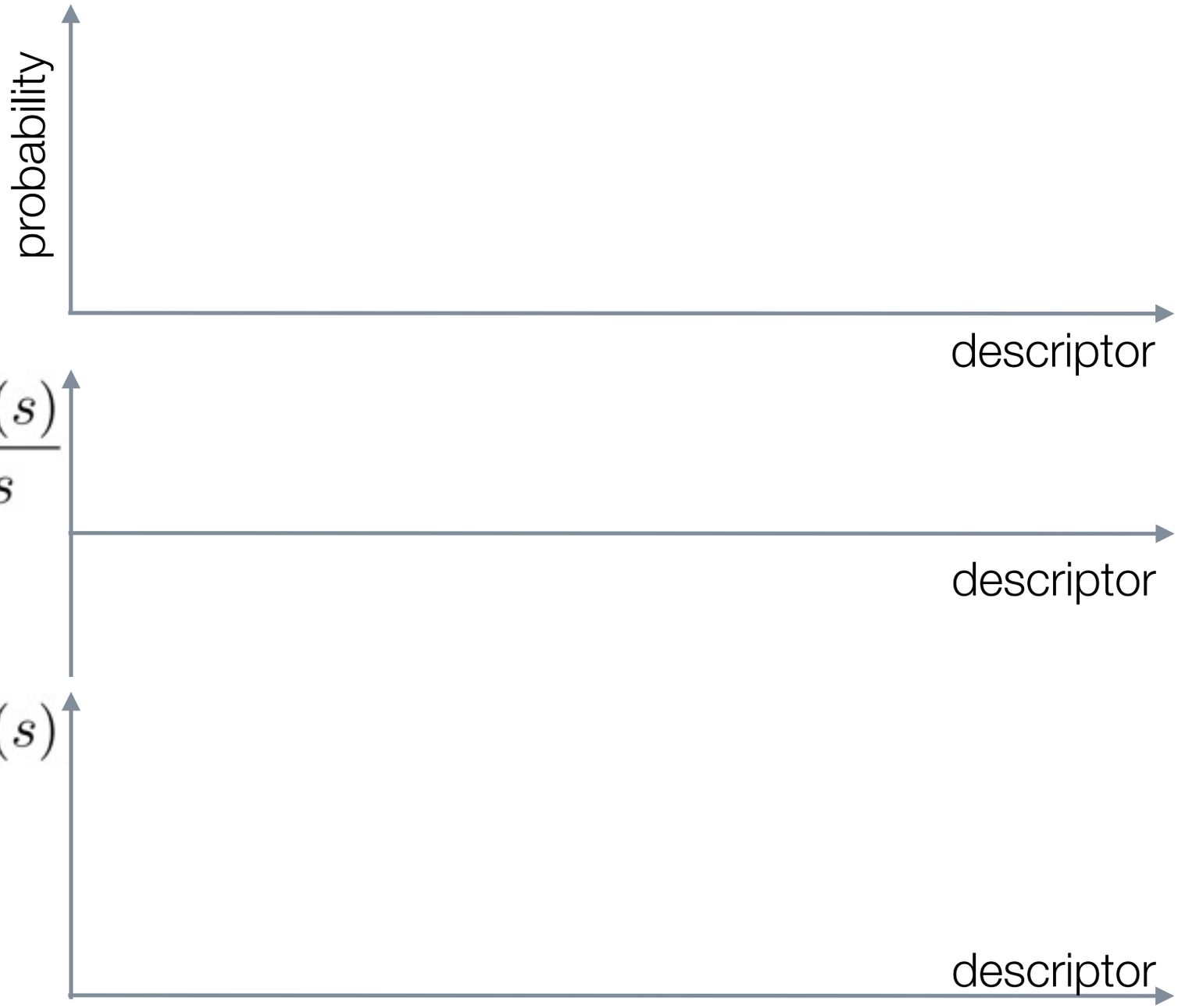
$$= -k\langle s'(x) - s_0 \rangle_{bias,s0}$$ (opposite of the “mean force”)

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



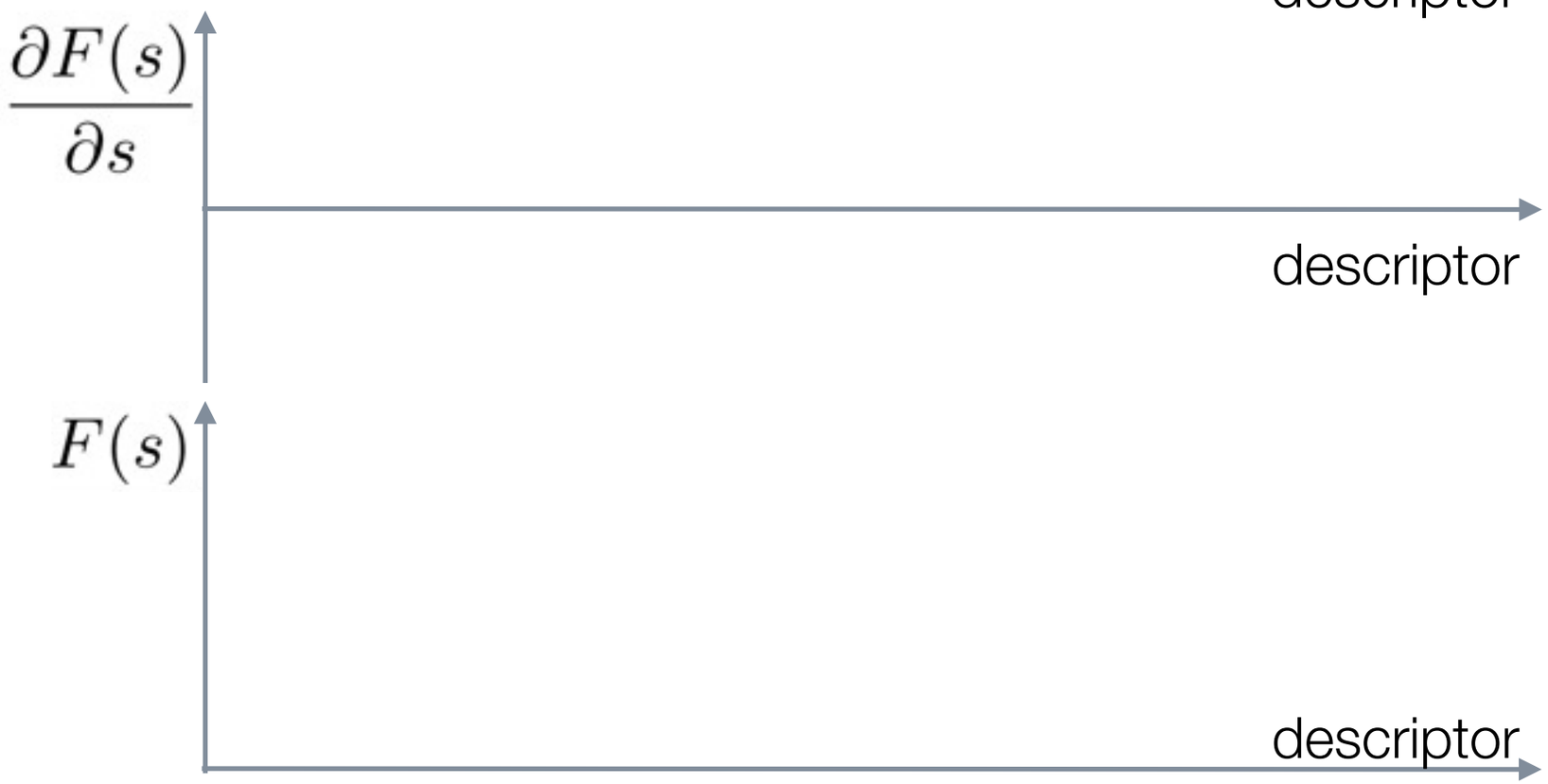
+WHAM etc...

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



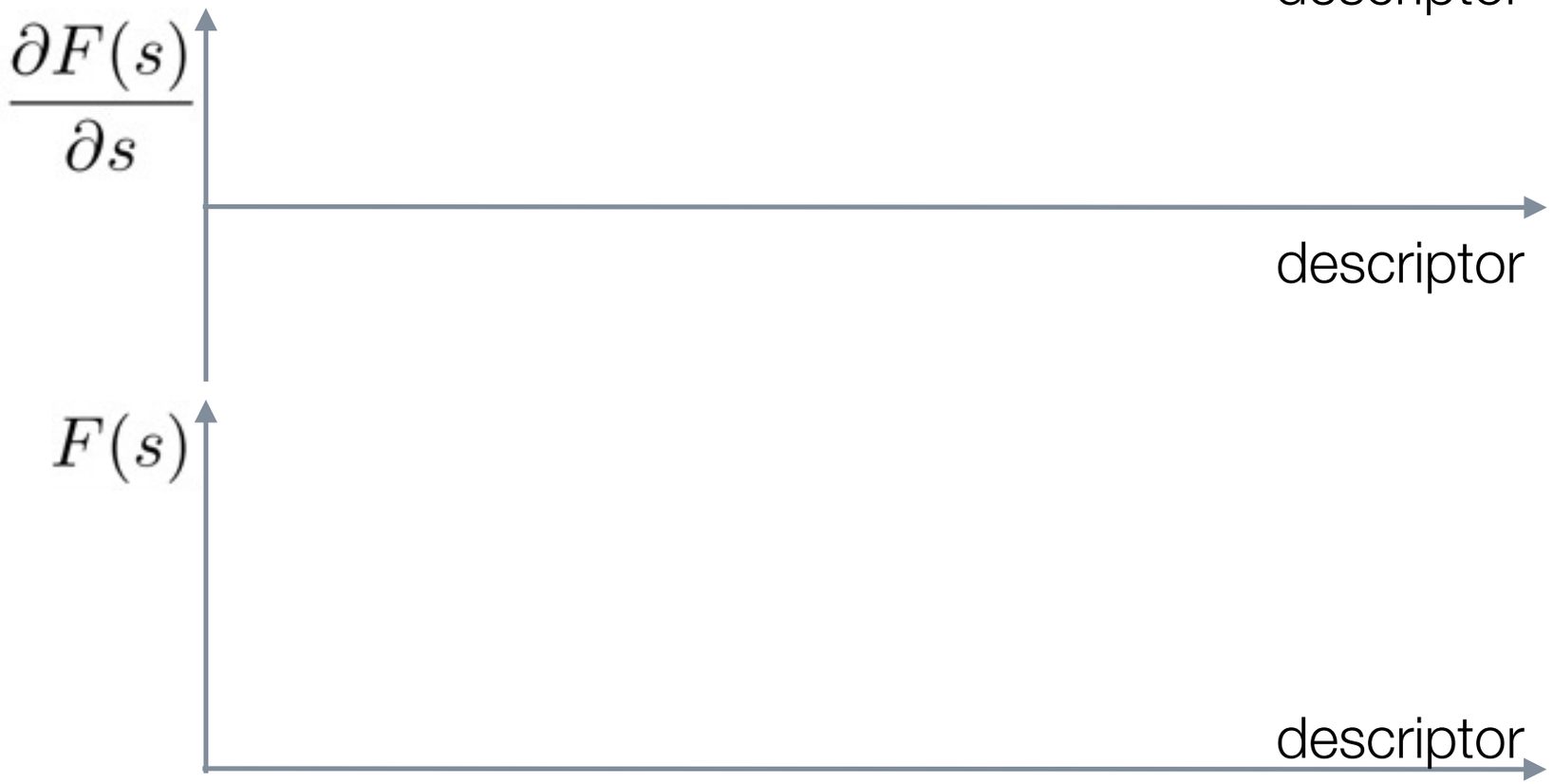
+WHAM etc...

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



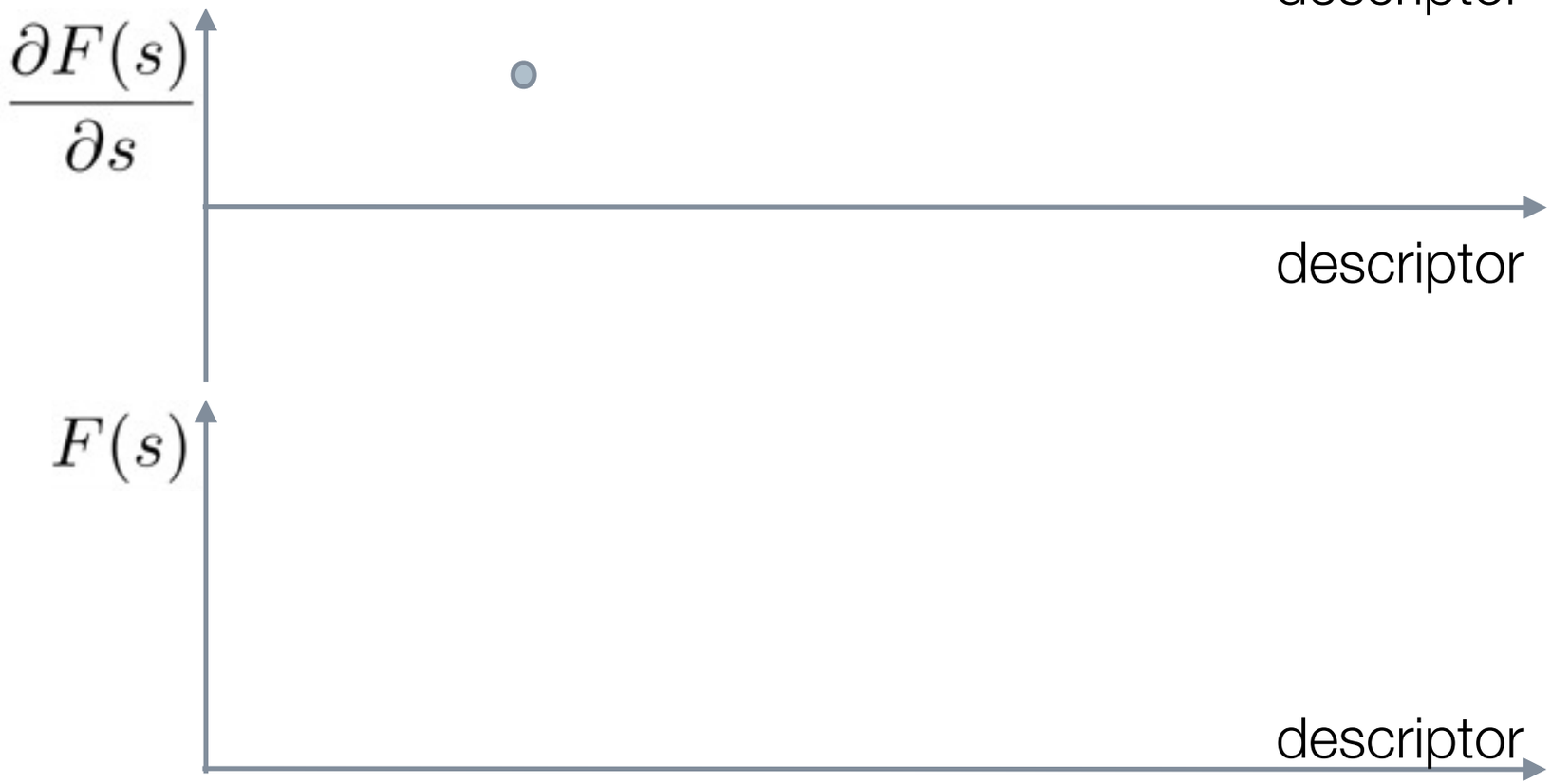
+WHAM etc...

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



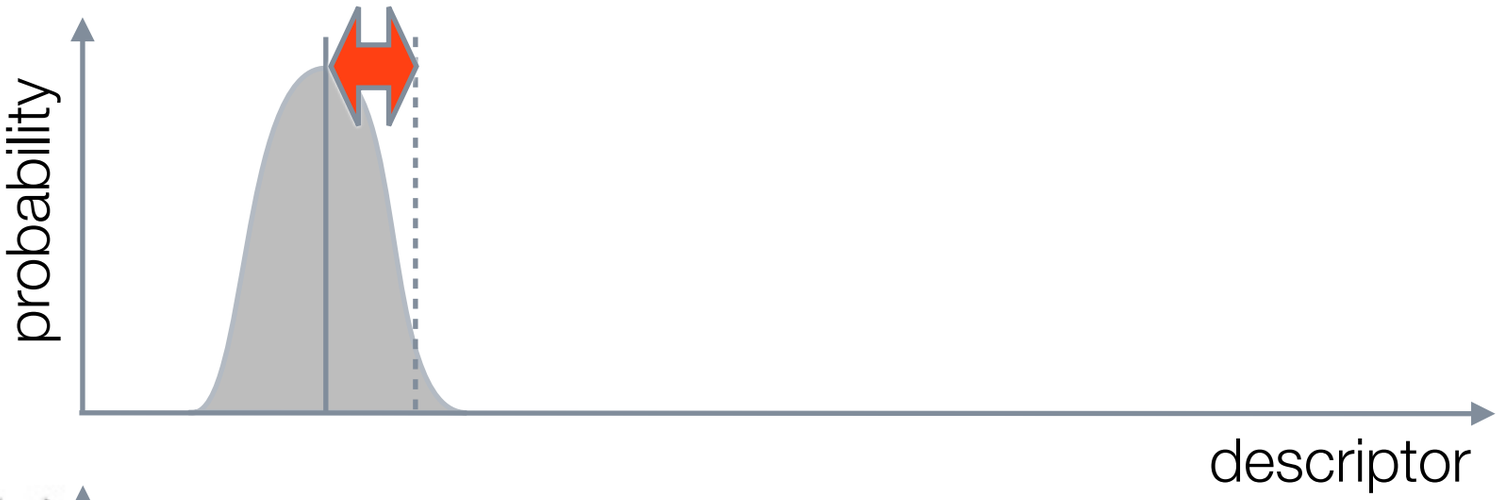
+WHAM etc...

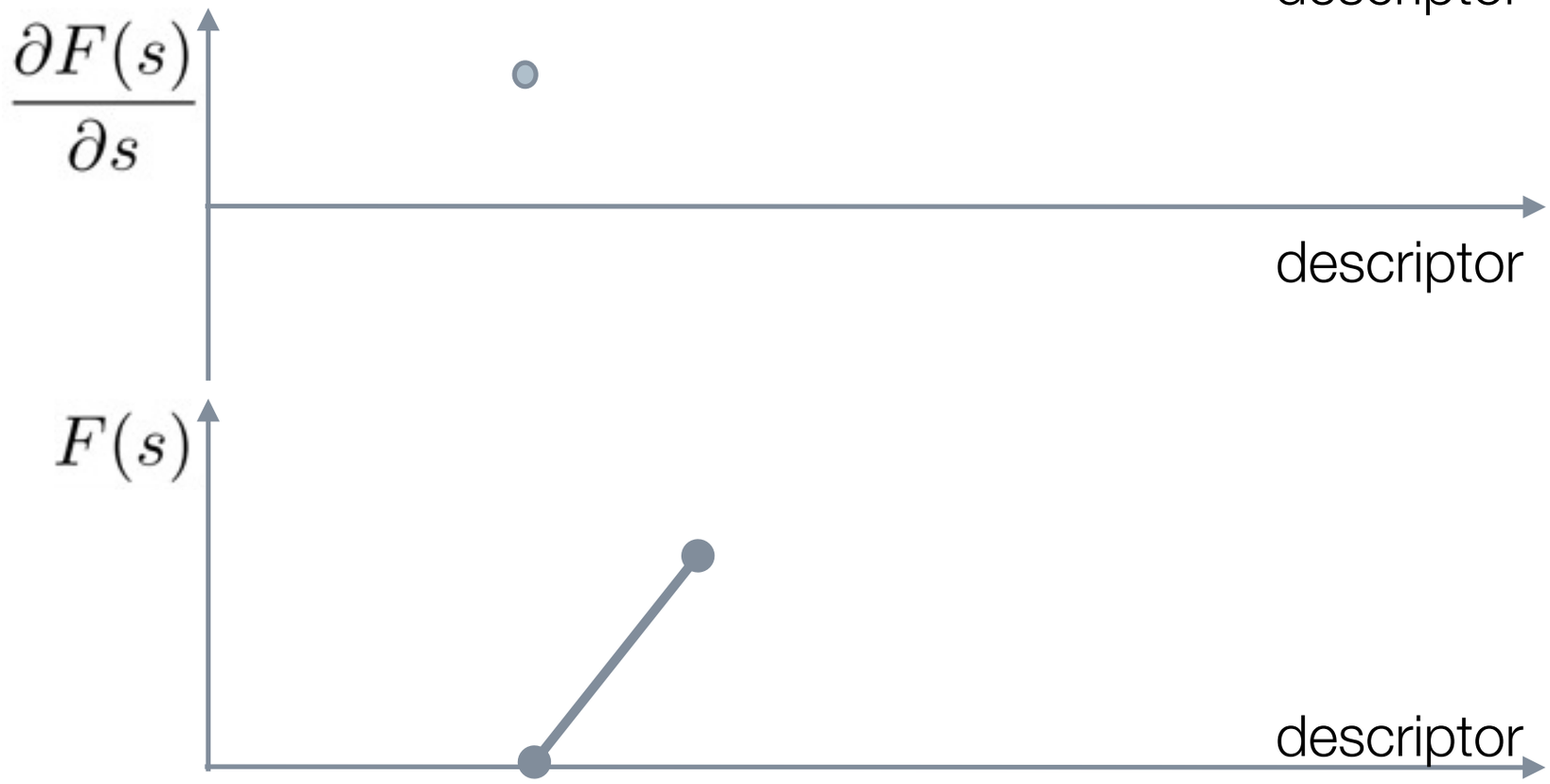
Thermodynamic Integration 101

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

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

© J. Culvenor





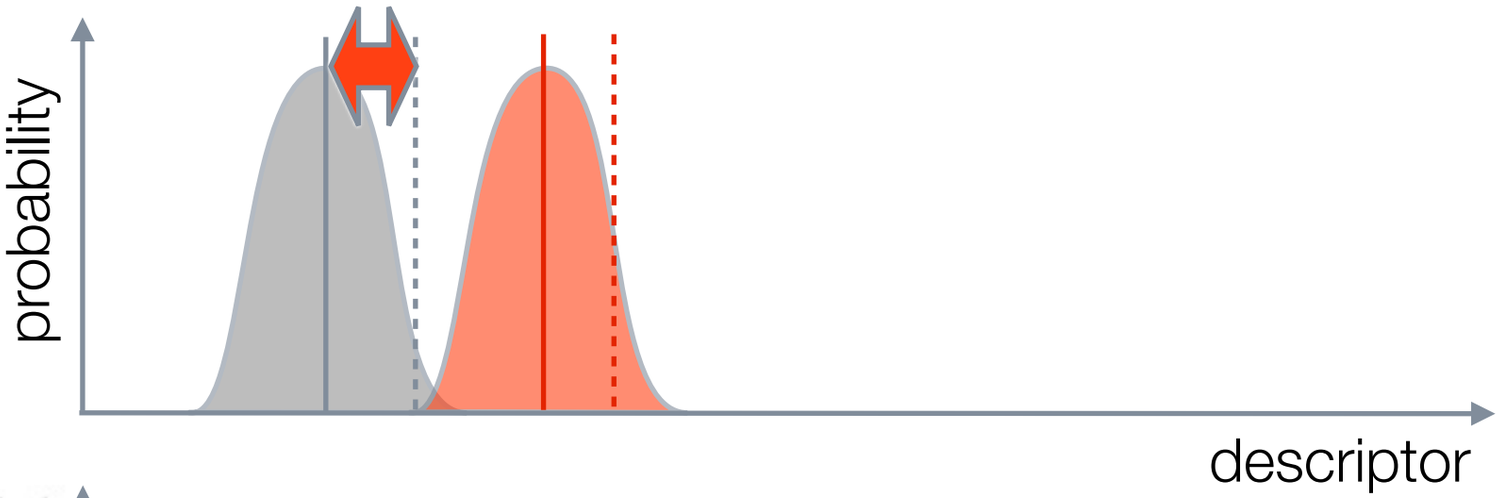
+WHAM etc...

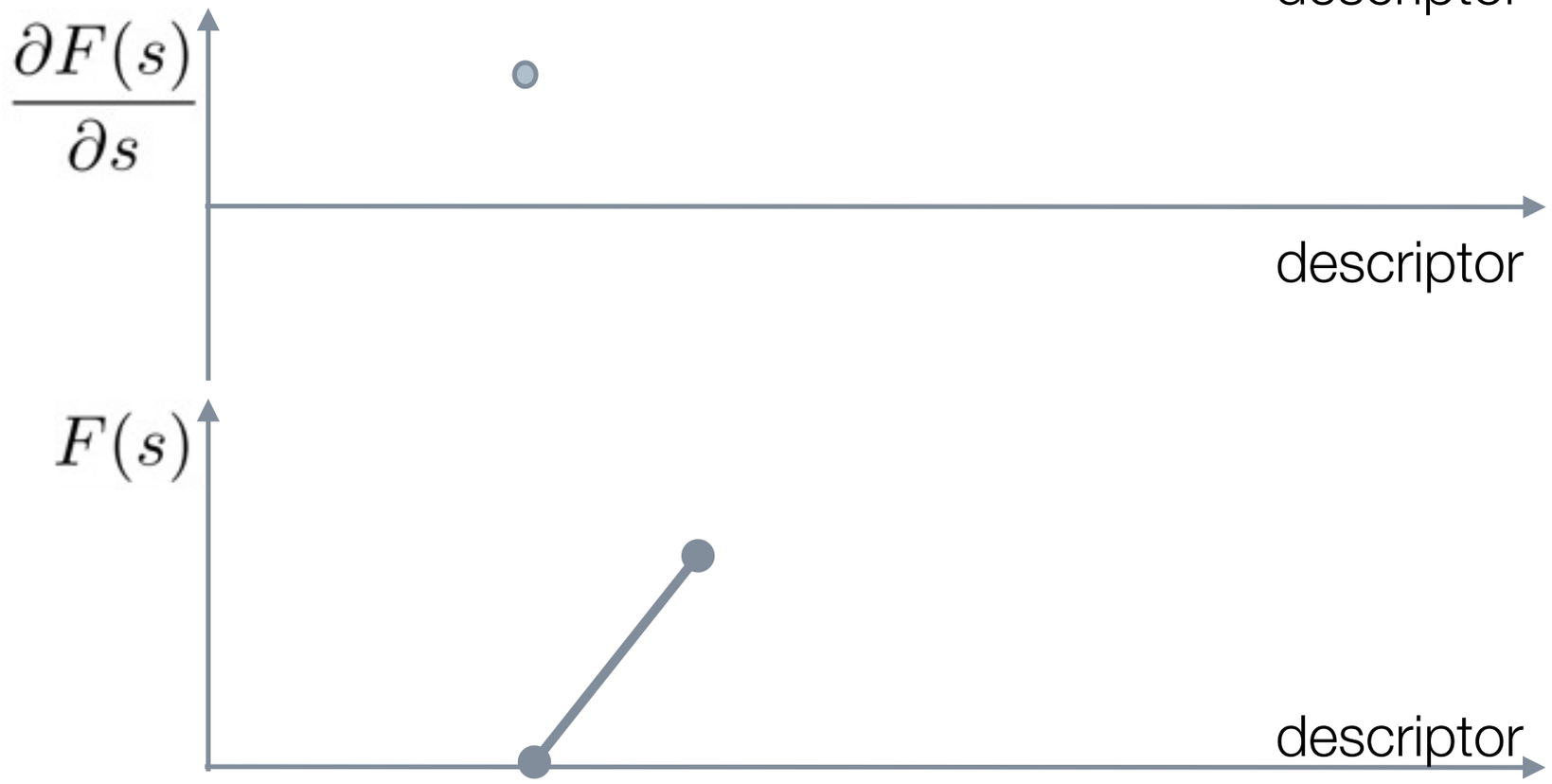
Thermodynamic Integration 101

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

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

© J. Culvenor





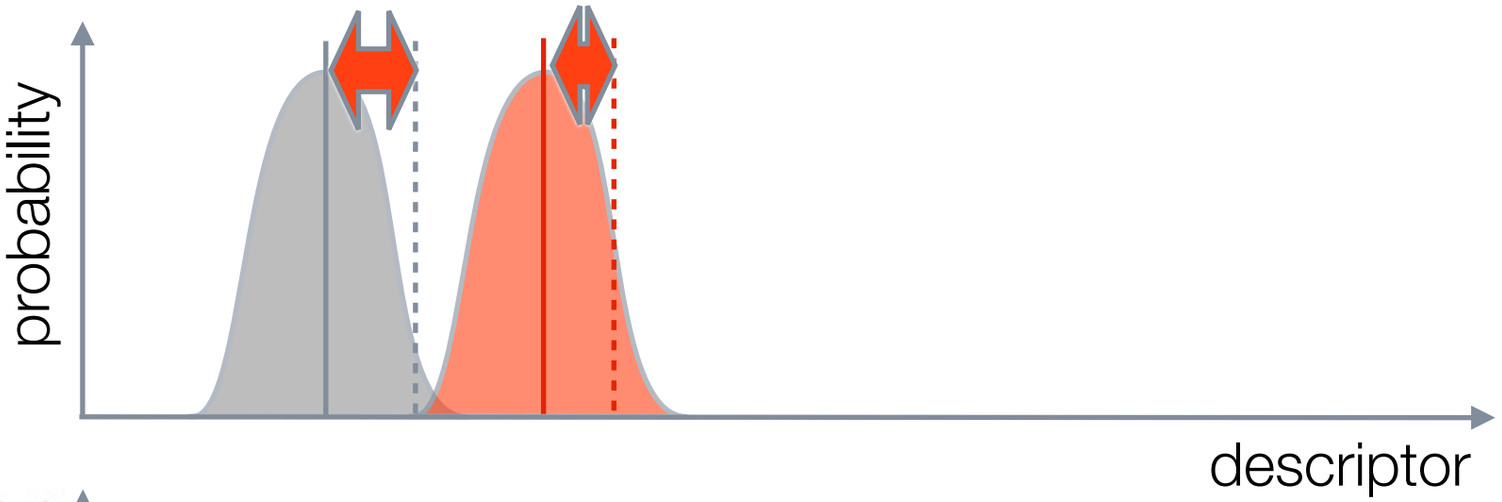
+WHAM etc...

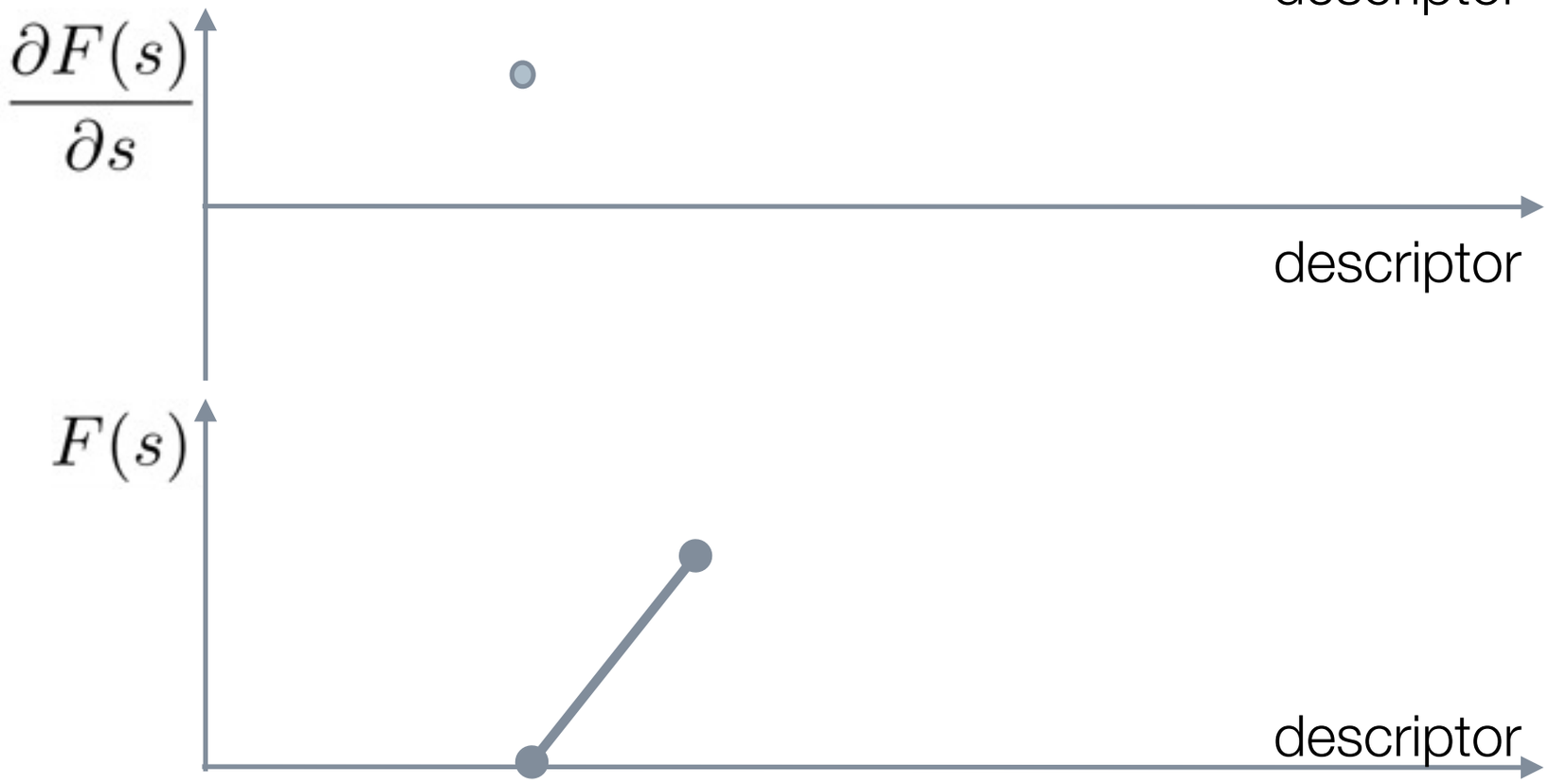
Thermodynamic Integration 101

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

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

© J. Culvenor





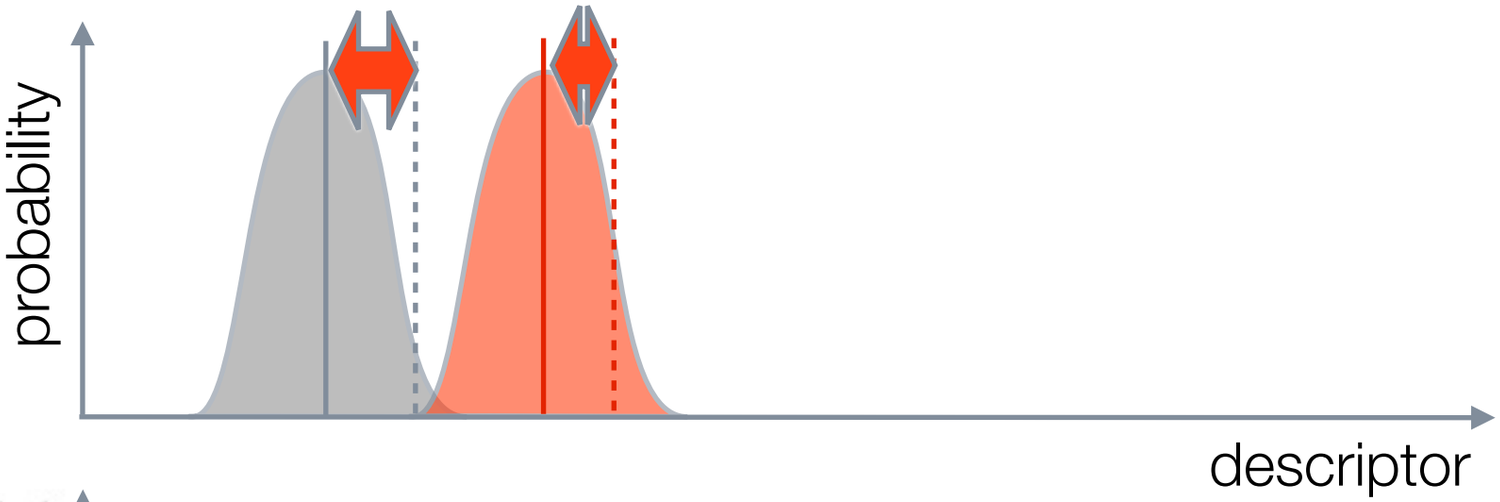
+WHAM etc...

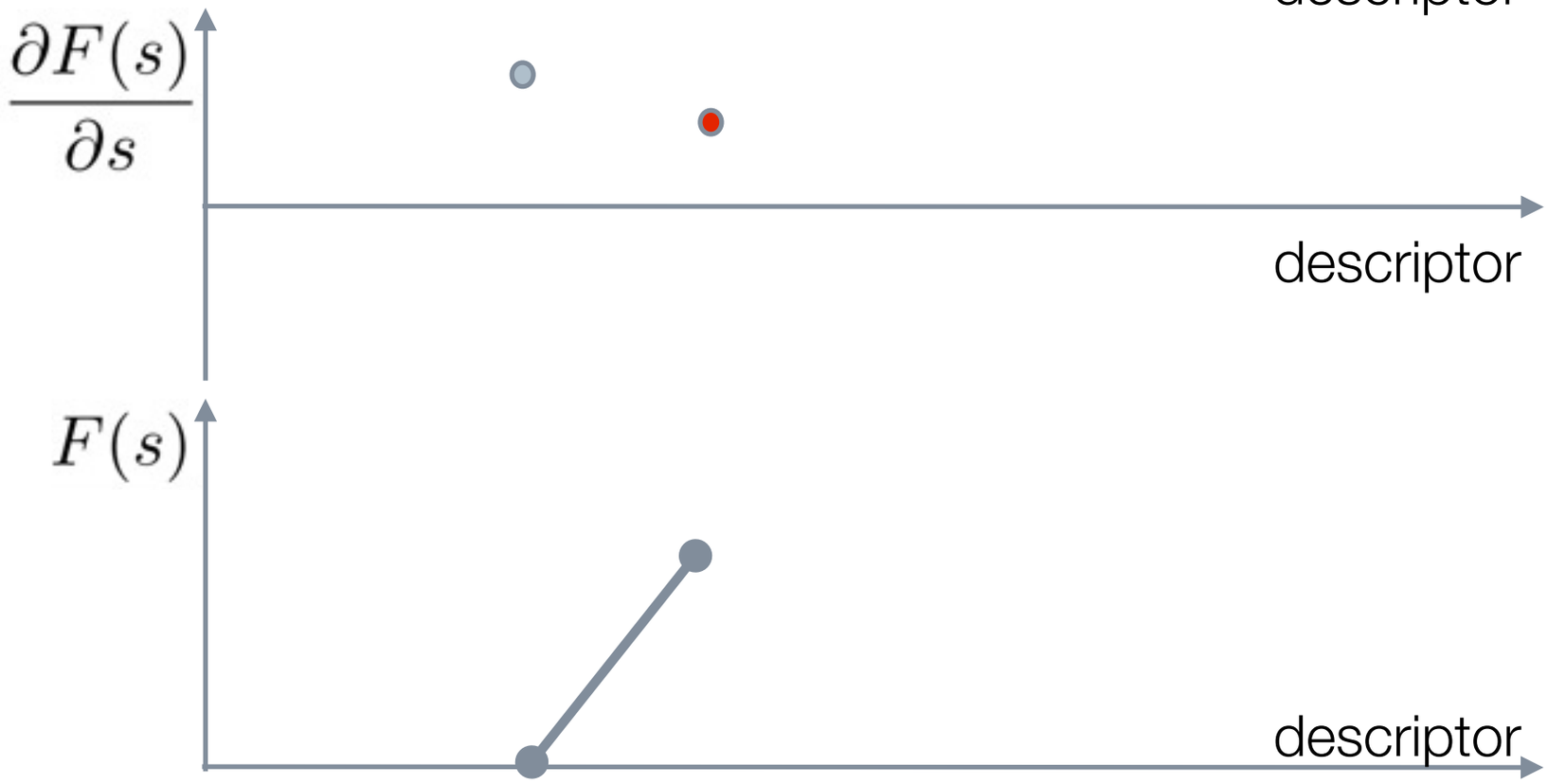
Thermodynamic Integration 101

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

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

© J. Culvenor





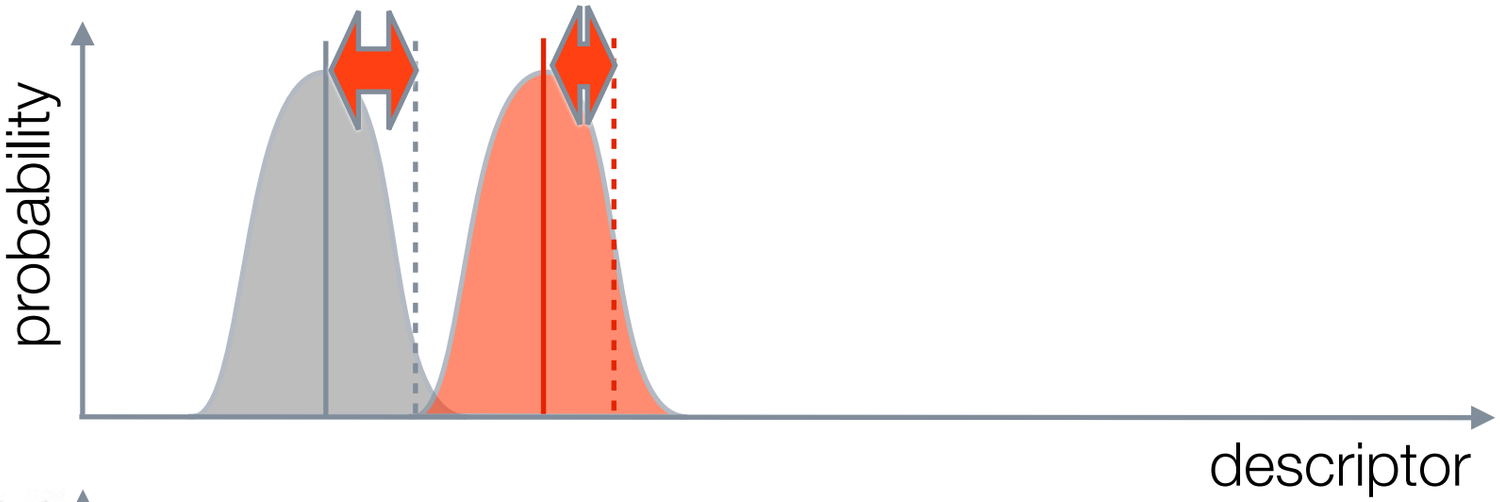
+WHAM etc...

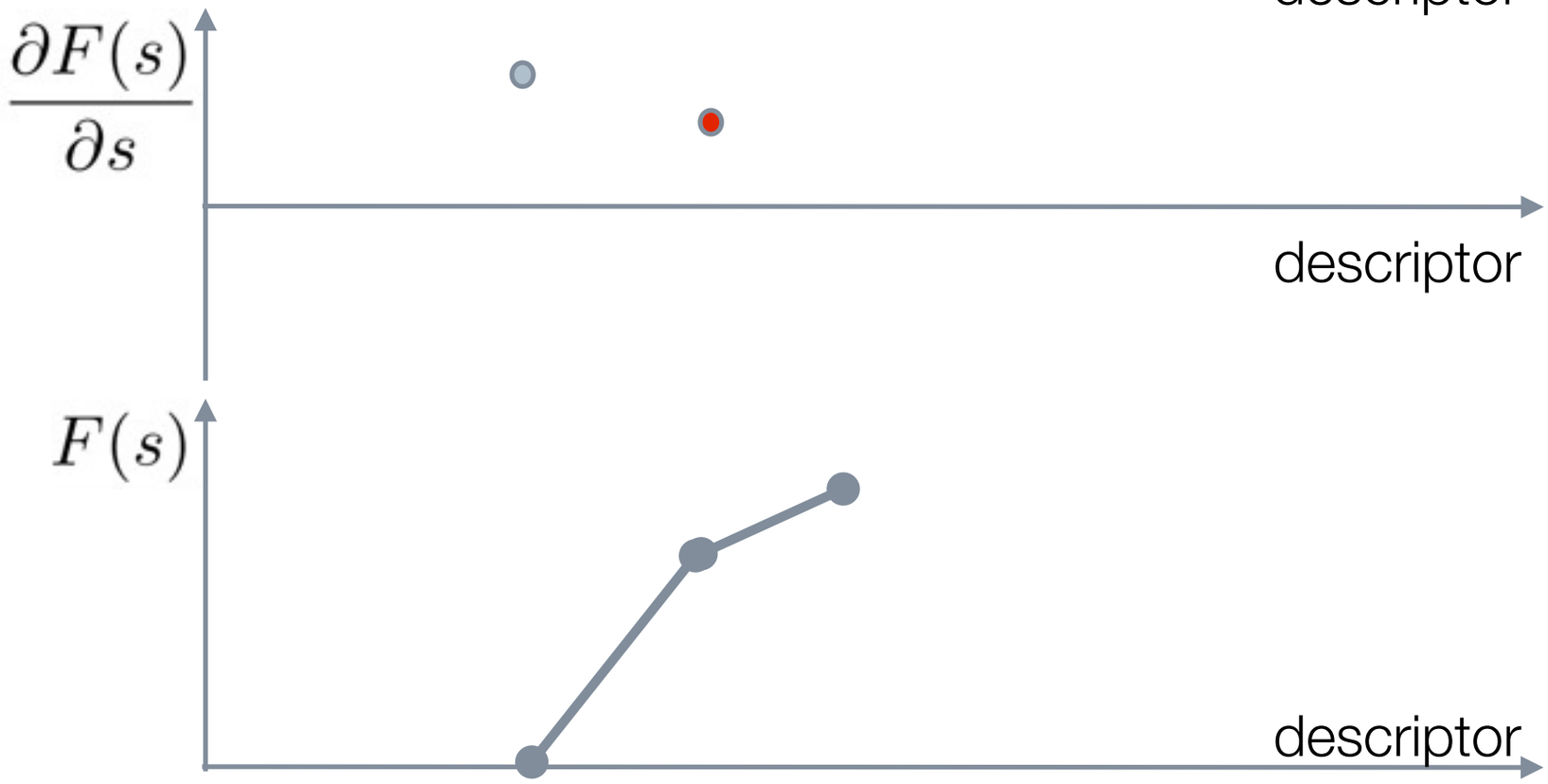
Thermodynamic Integration 101

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

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

© J. Culvenor





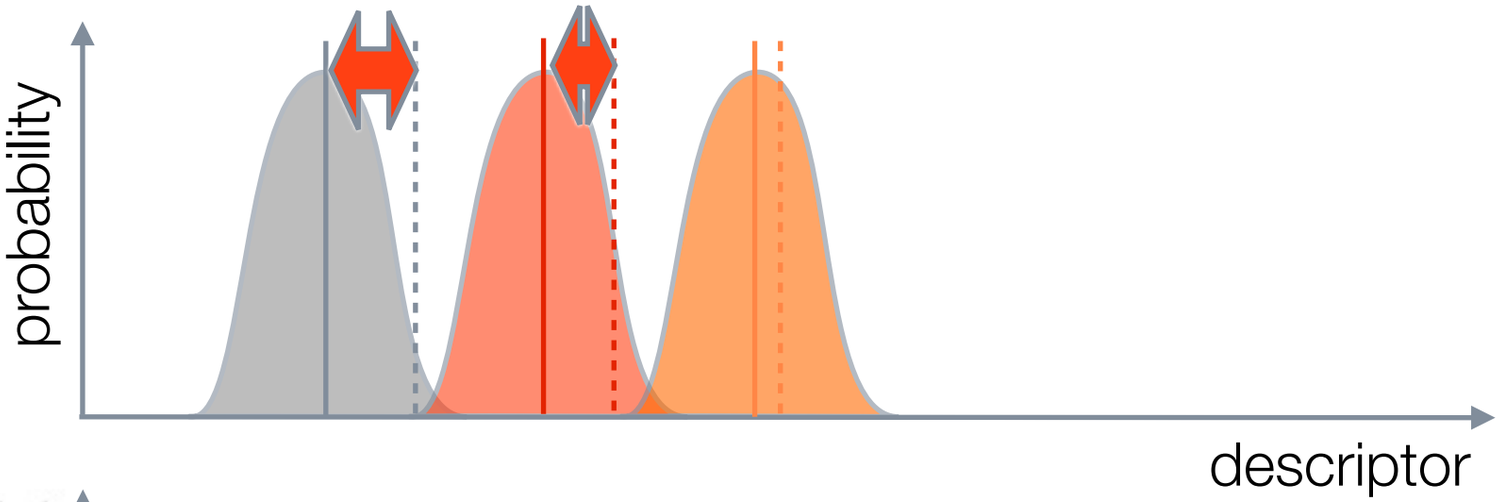
+WHAM etc...

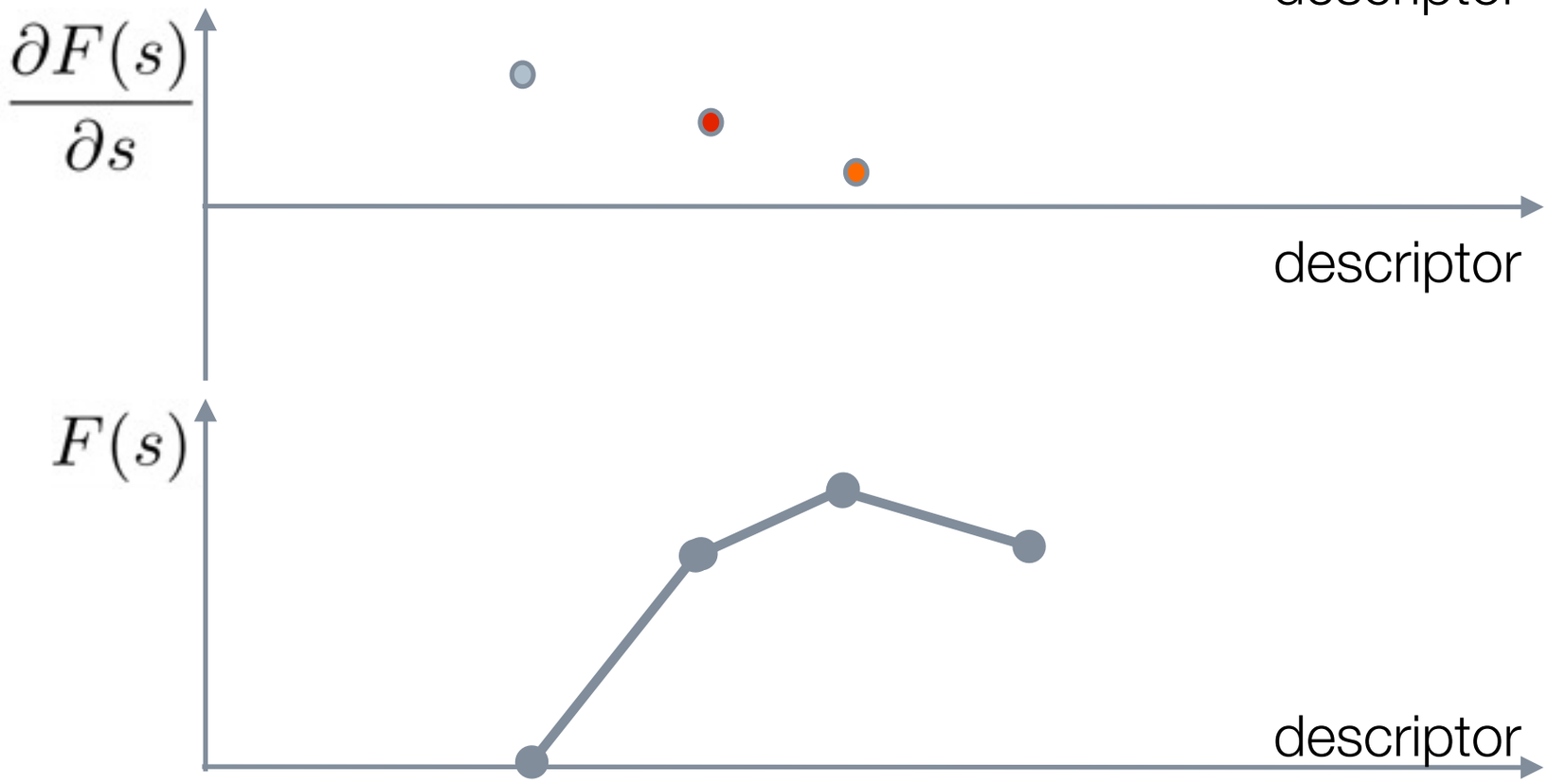
Thermodynamic Integration 101

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

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

© J. Culvenor





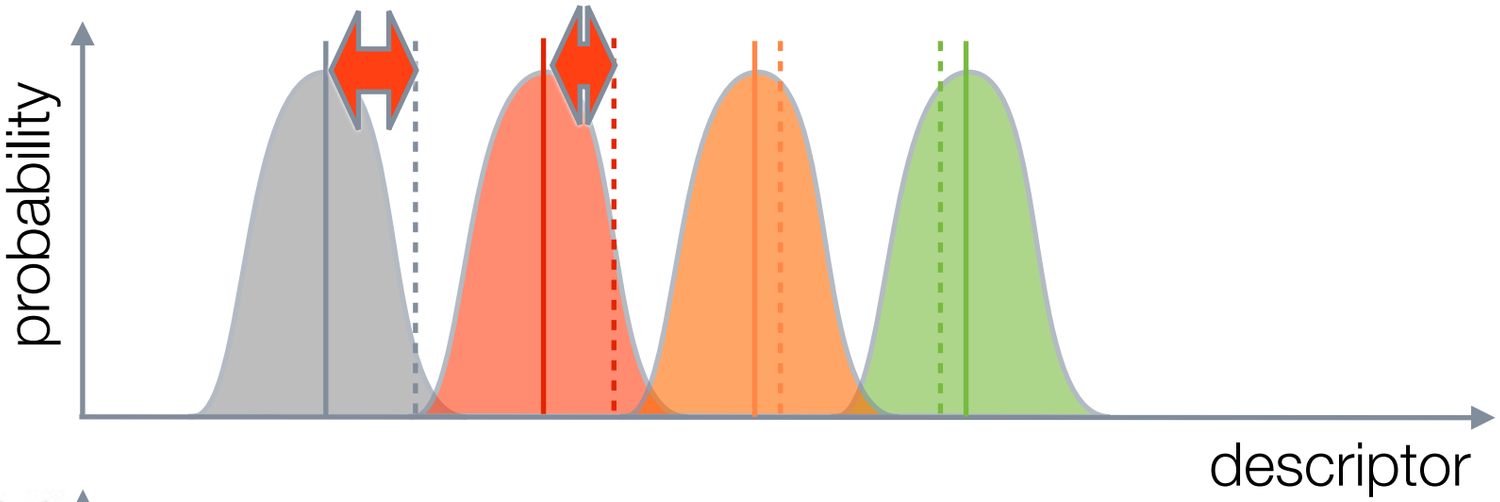
+WHAM etc...

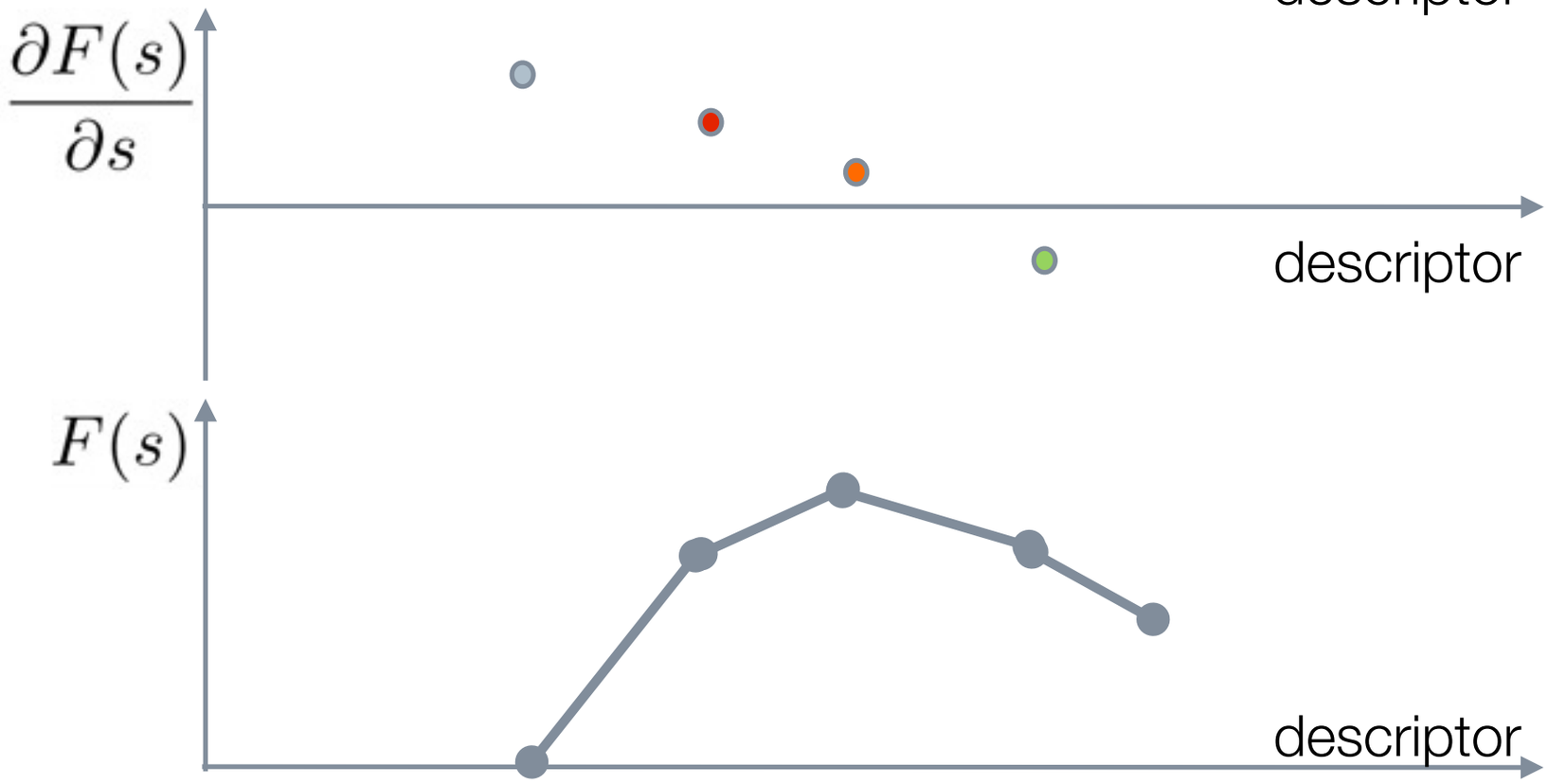
Thermodynamic Integration 101

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

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

© J. Culvenor





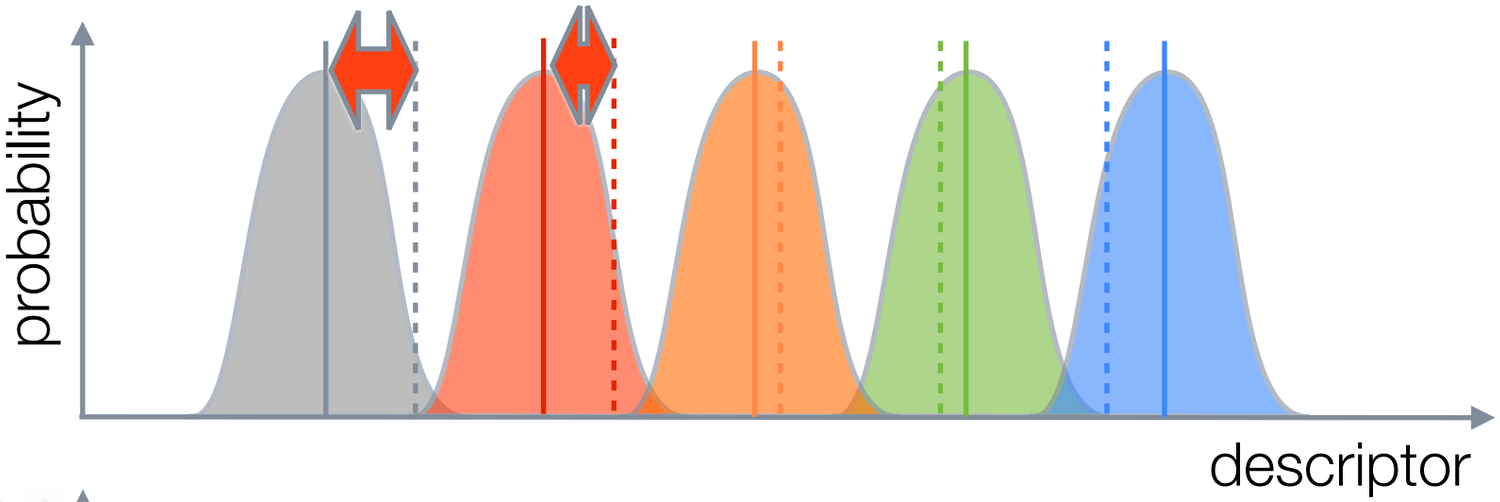
+WHAM etc...

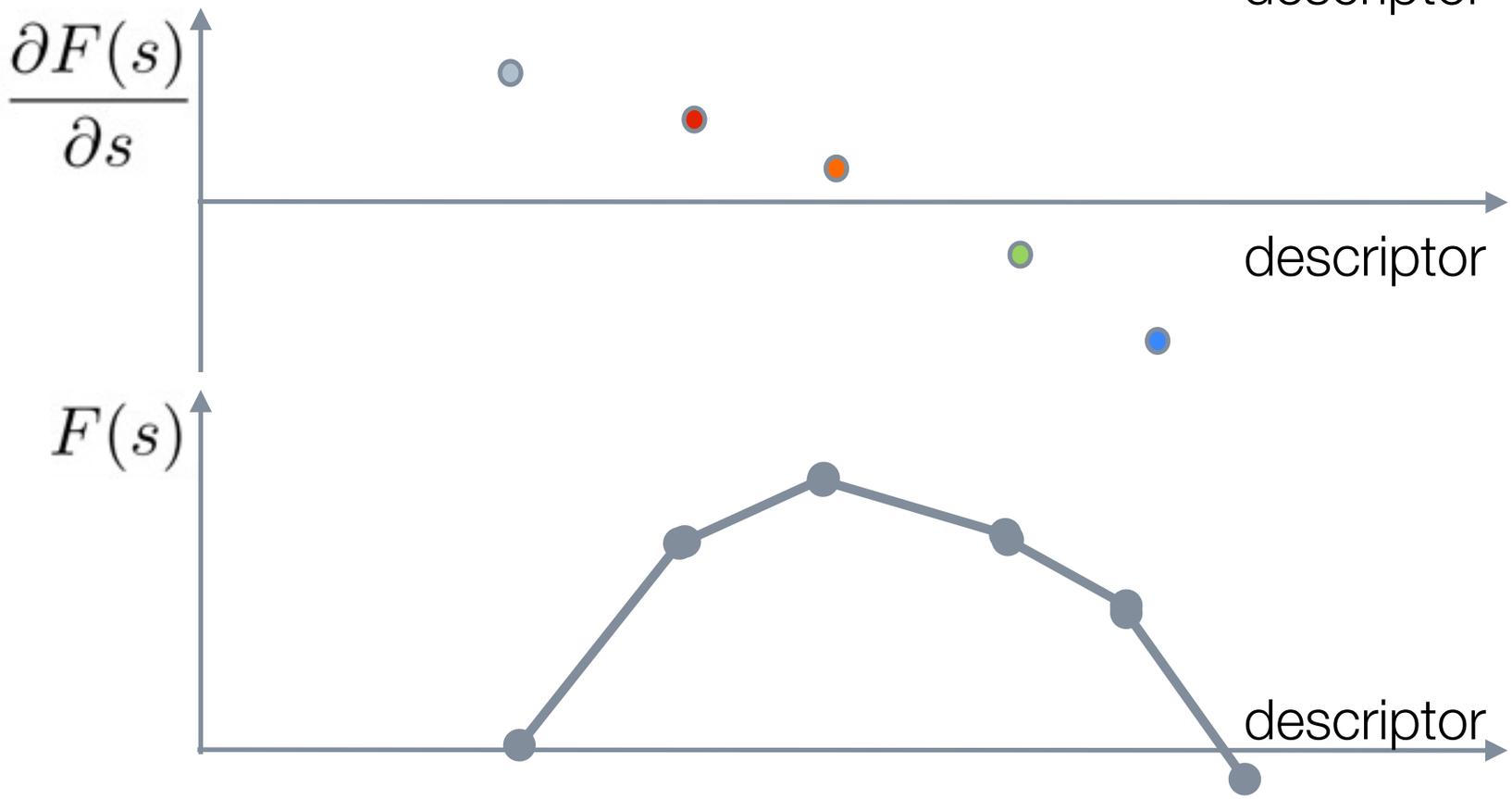
Thermodynamic Integration 101

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

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

© J. Culvenor

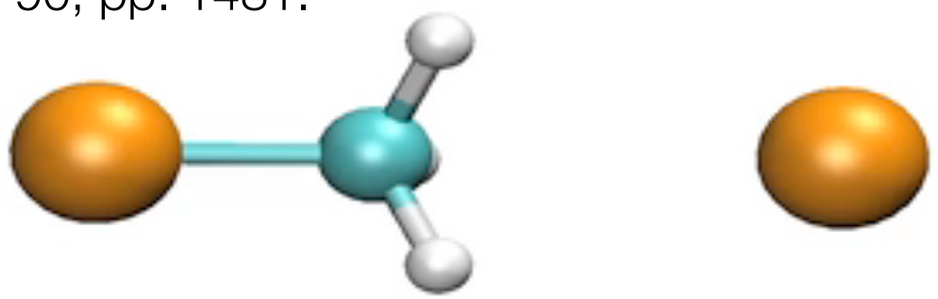




+WHAM etc...

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.



SN2 reaction: if you like this movie come to the tutorial to see how to make it!

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

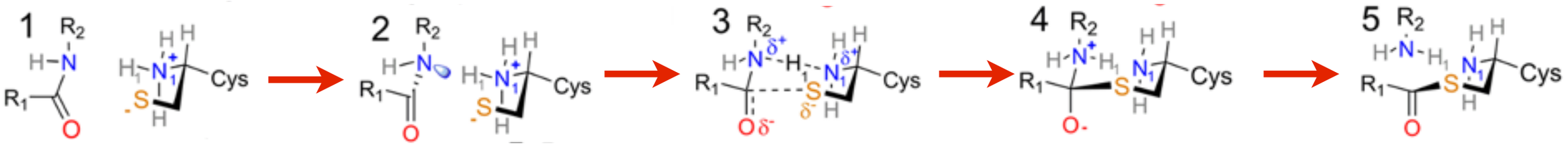
TI-like:

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

The problem of the descriptors

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

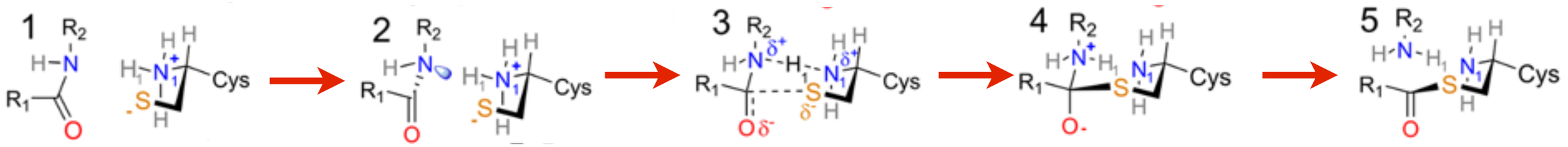
* A multidimensional landscape: turn into monodimensional

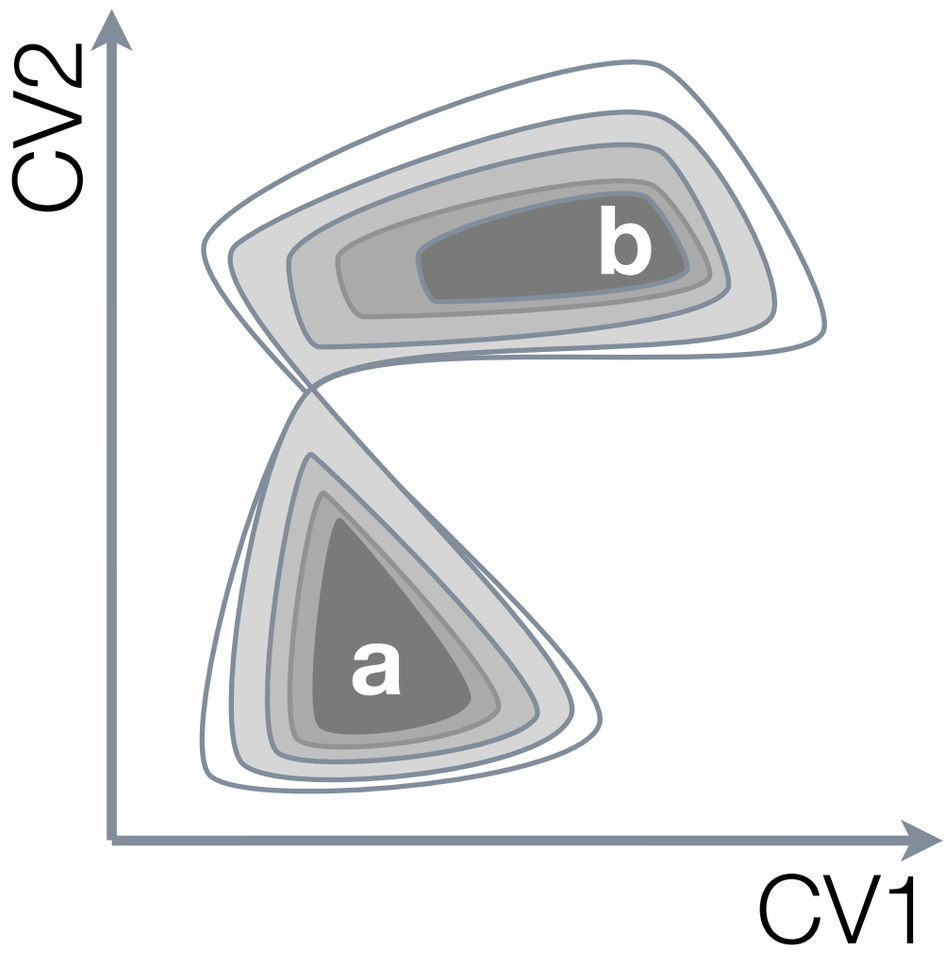


The problem of the descriptors

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

* A multidimensional landscape: turn into monodimensional

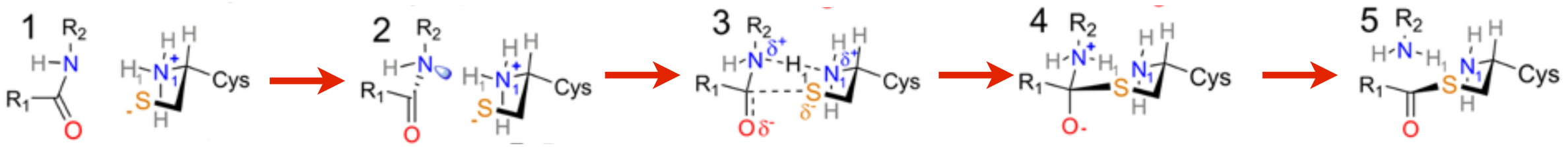


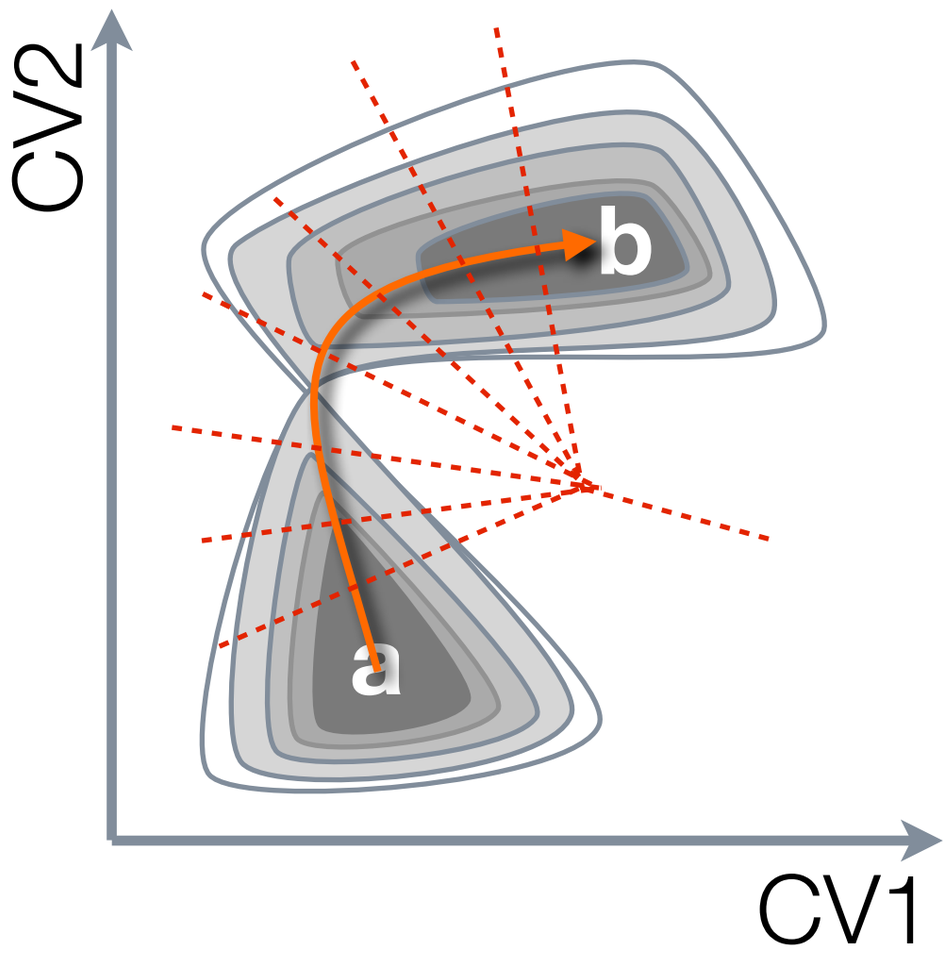


The problem of the descriptors

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

* A multidimensional landscape: turn into monodimensional

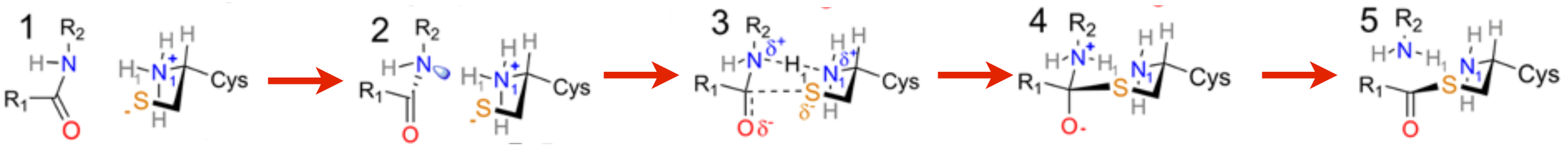


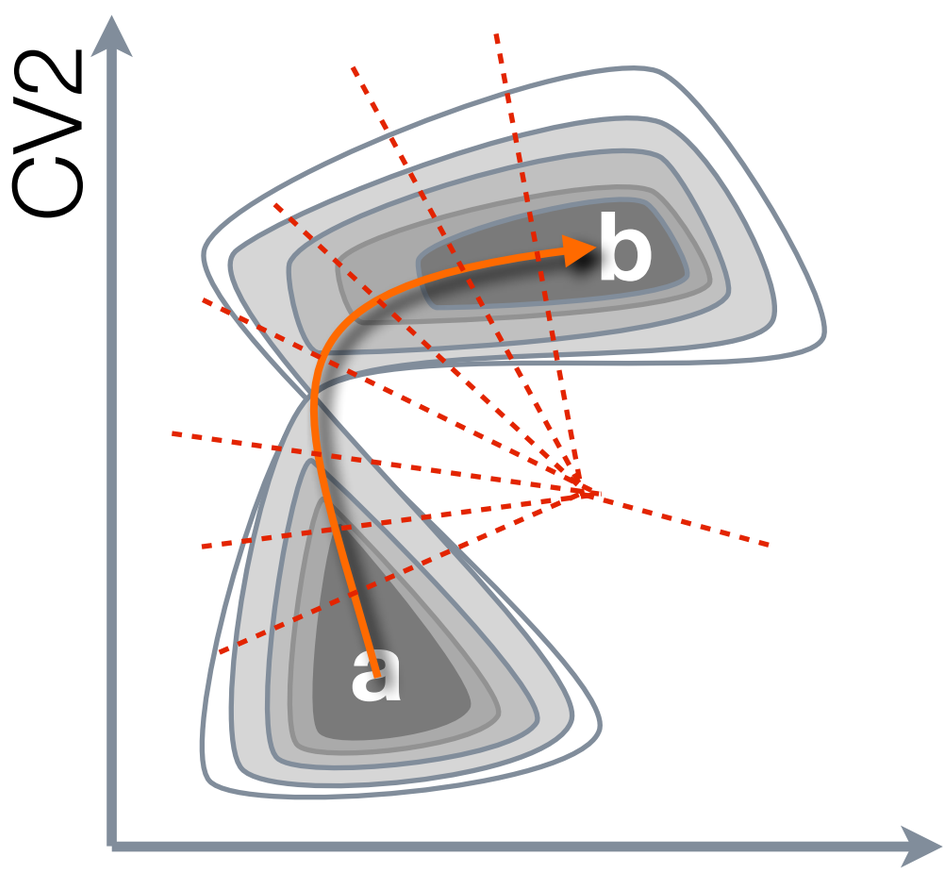


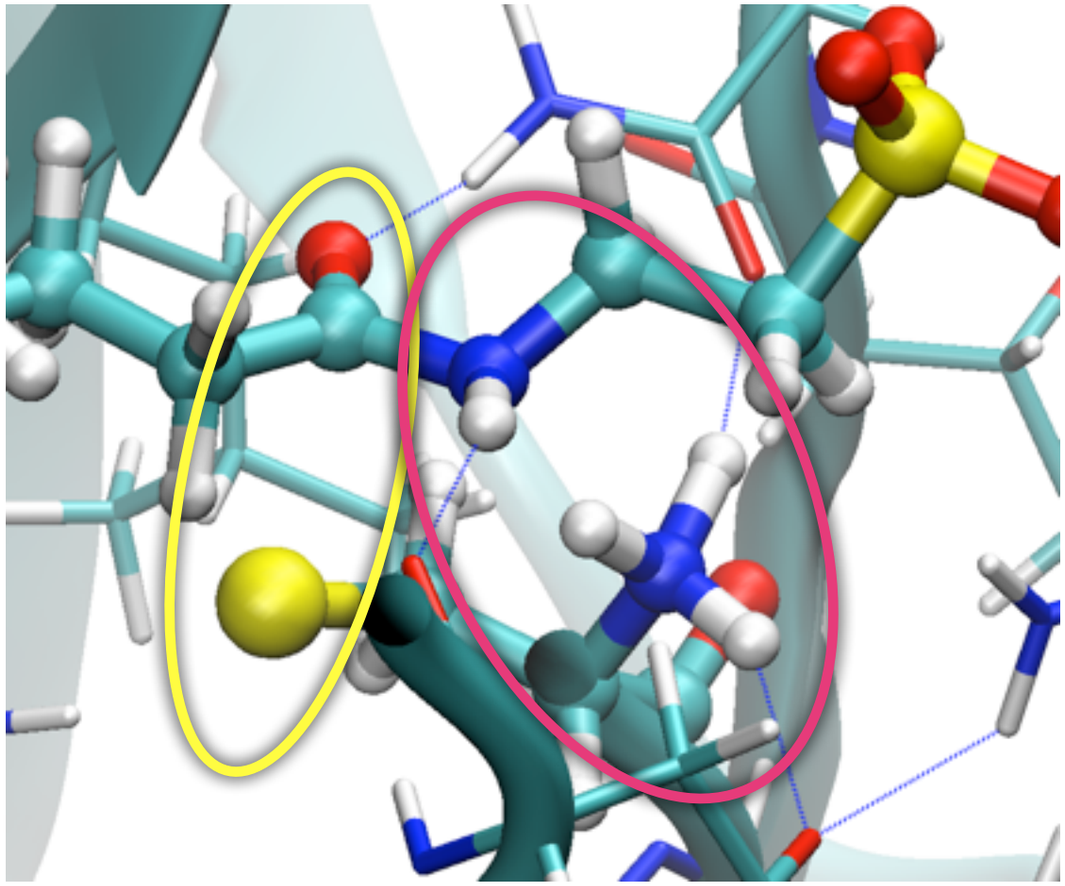
The problem of the descriptors

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

* A multidimensional landscape: turn into monodimensional



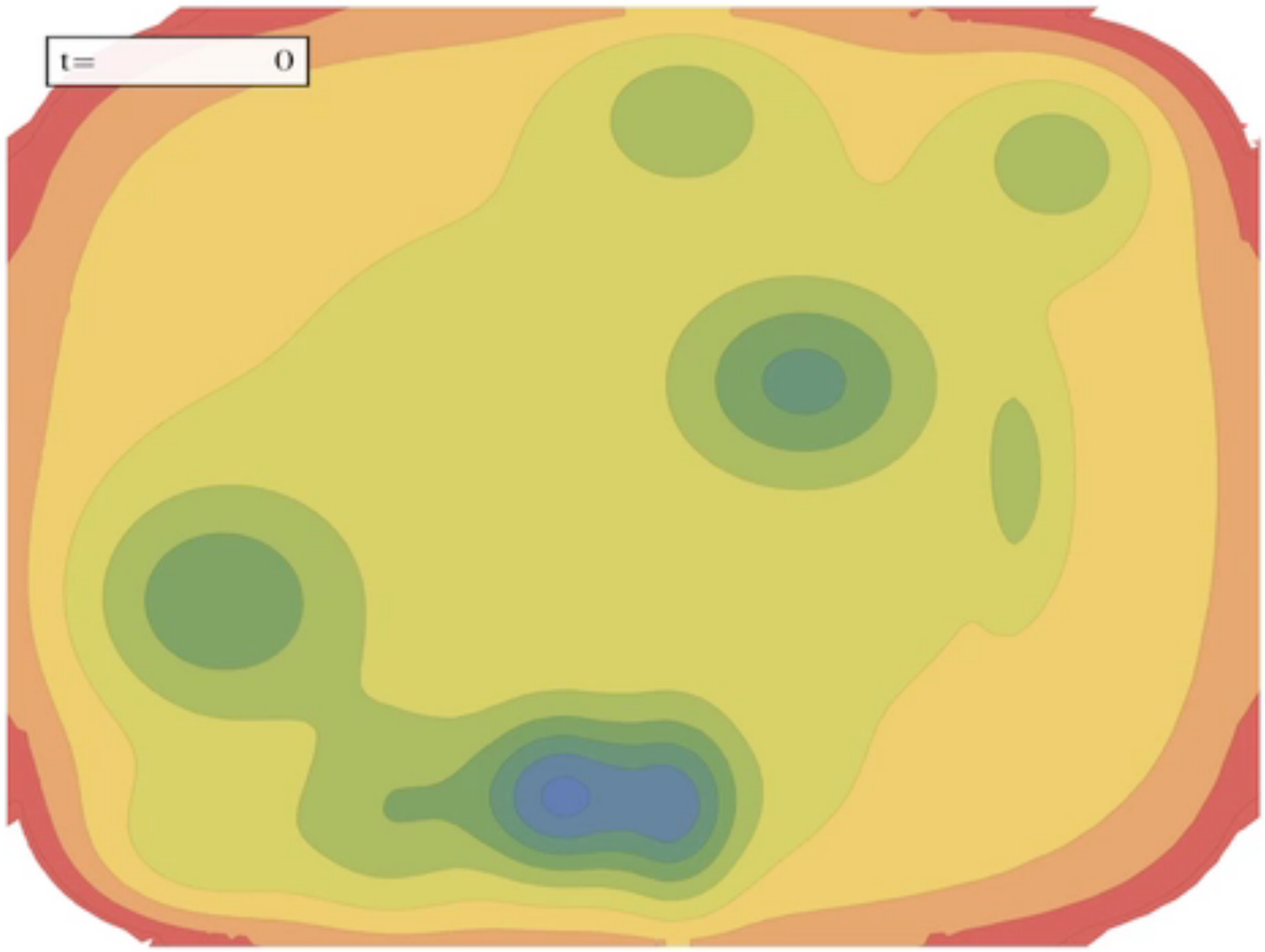




Lodola, Branduardi, De Vivo, Capoferri, Mor, Piomelli, Cavalli, PlosOne, 2012, vol 7, pp. e32397

The problem of the descriptors

* Use machine learning approaches Tribello et. al. PNAS 2010, vol. 107, pp. 17509.

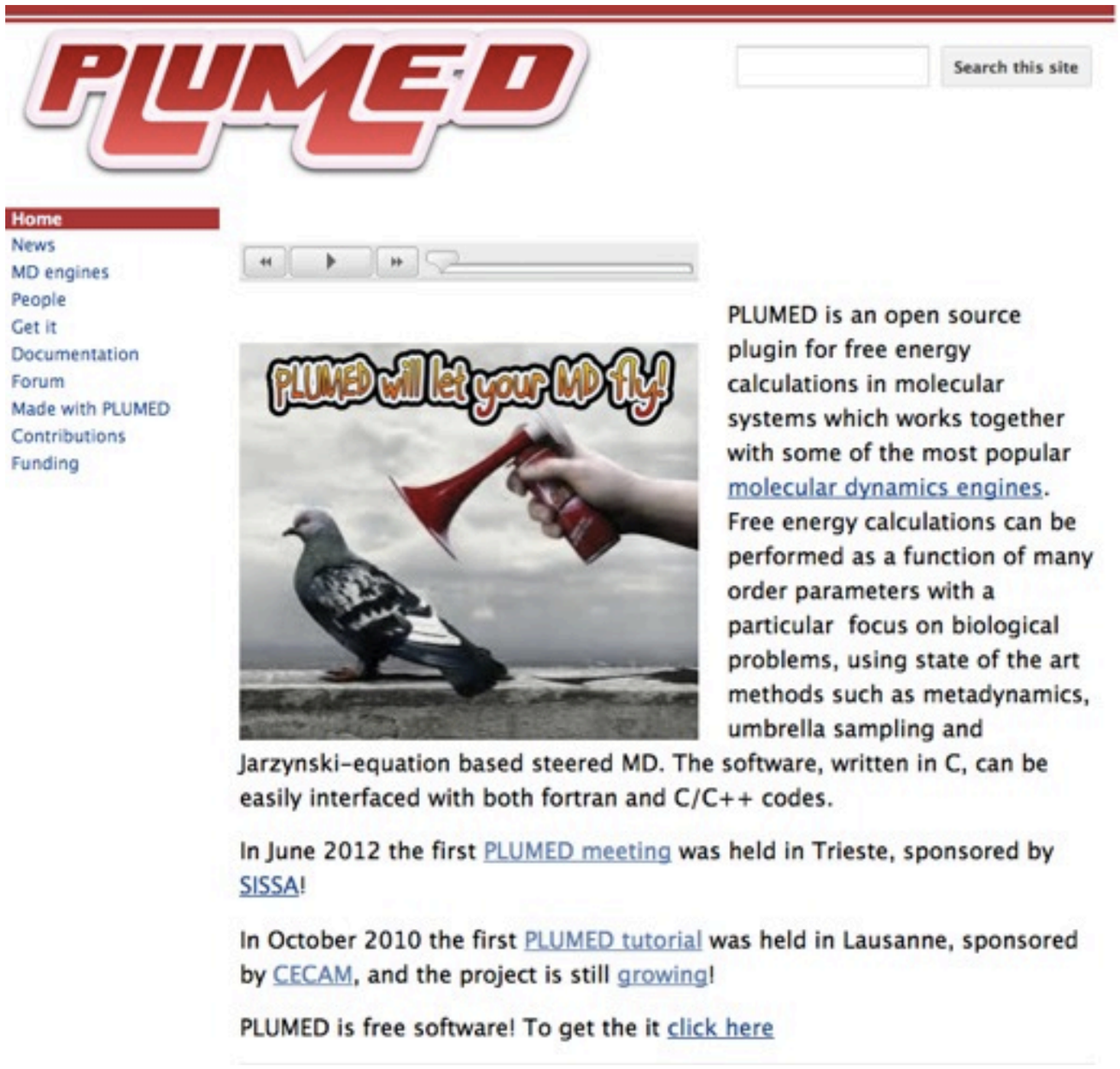


Outline

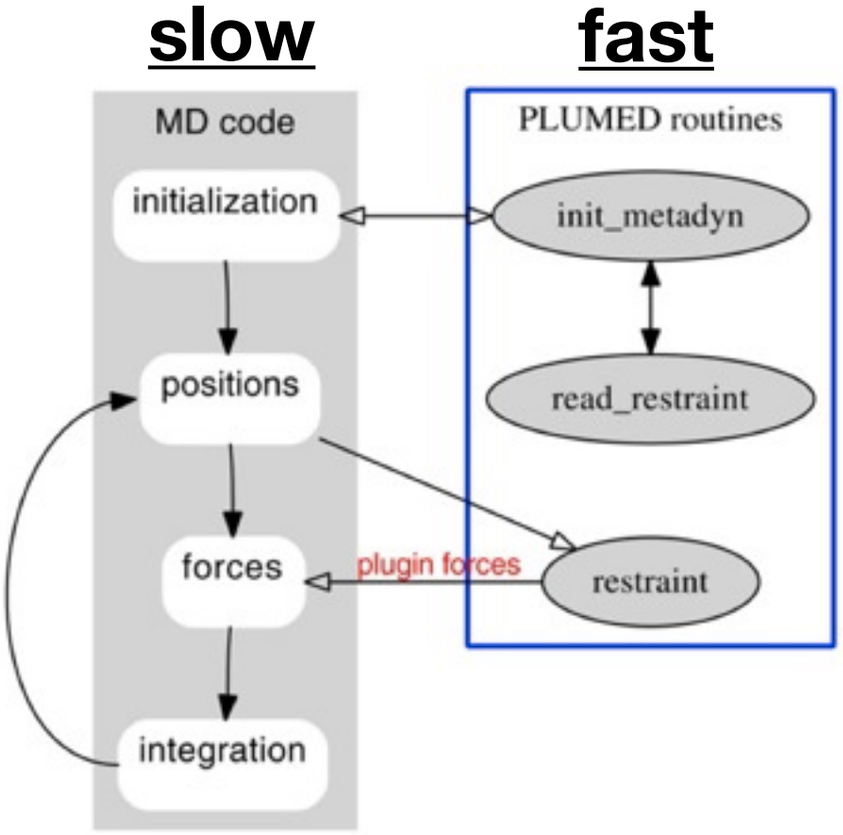
- What is free energy and why is important
- Traditional QM approach
- Explicit sampling
- Enhanced sampling
- **PLUMED plugin for free energy calculations**

The PLUMED plugin: www.plumed-code.org

Bonomi, Branduardi et al., Comp. Phys. Comm. 2009 vol. 180 pp.1961



small connection with MD code



Already in many classical MD codes (LAMMPS, DLPOLY, SANDER, GROMACS, NAMD) **now in FHI-aims**!

PLUMED: what it contains

Methods

- Constrained relaxation
- Umbrella sampling
- Metadynamics (well-tempered)
- d-AFED
- Steered MD
- Harmonic walls
- Ratcheting
- Reweighting schemes
- String method

Descriptors (with der.)

- Distances, angles, dihedrals
- Coordination numbers
- Contact maps
- Path collective variables
- Energy
- Function of variables
- SPRINT variables
- .....

+ MetaGUI analysis interface and PLUMED preparation GUI in VMD+analysis tools!

In FHI-aims

* compile separately: Makefile.meta

* in control.in

```
MD_run   500.0  NVT_parrinello 300 0.01
MD_time_step  0.002
MD_clean_rotations  .true.
MD_restart    .false.
output_level MD_light
MD_maxsteps -1
MD_MB_init  300
MD_RNG_seed 12345
plumed .true.
```

this enables PLUMED

* plumed.dat

```
DISTANCE LIST 1 <g1> SIGMA 0.11
g1->
3 4 5 6
g1<-
DISTANCE LIST 2 <g1> SIGMA 0.11
UWALL CV 1 LIMIT 7. KAPPA 0.5
UWALL CV 2 LIMIT 7. KAPPA 0.5
HILLS HEIGHT 0.00047 W_STRIDE 50
PRINT W_STRIDE 2
ENDMETA
```

set a distance as descriptor

set anotherdistance as descriptor

set repulsive boundaries

set metadynamics

Acknowledgements

* Luca Ghiringhelli (FHI, Berlin)
* The PLUMED developers (ETHZ, USILU, Cambridge, UCSF, SISSA, Mount Sinai, Curtin, EPFL, MPI-BP)
* Giovanni Bussi (SISSA, Triest)
* Michele Parrinello (ETHZ/USILU, Lugano)
* José Faraldo-Gómez (MPI-BP, Frankfurt/Main)

Thank you!