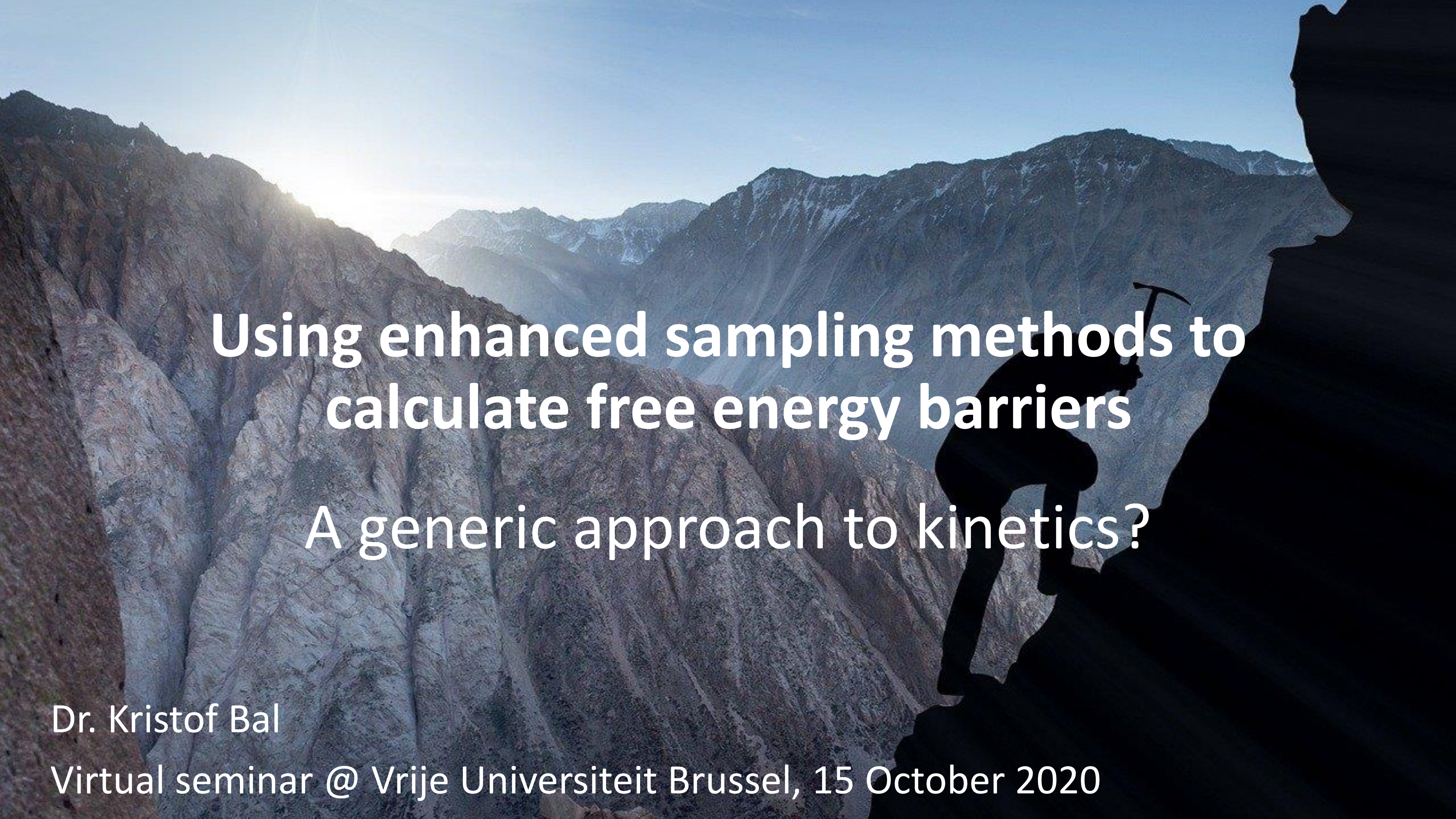


Using enhanced sampling methods to calculate free energy barriers

A generic approach to kinetics?

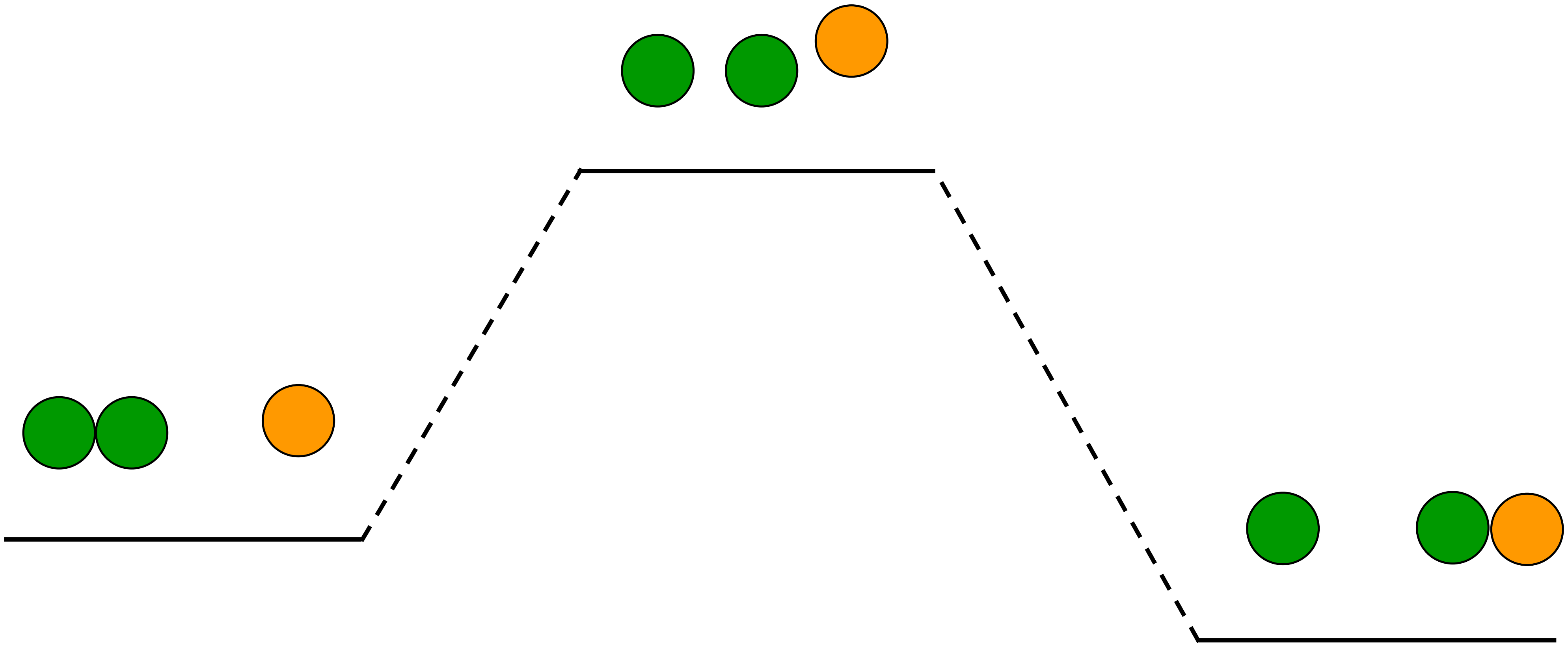
Dr. Kristof Bal

Virtual seminar @ Vrije Universiteit Brussel, 15 October 2020

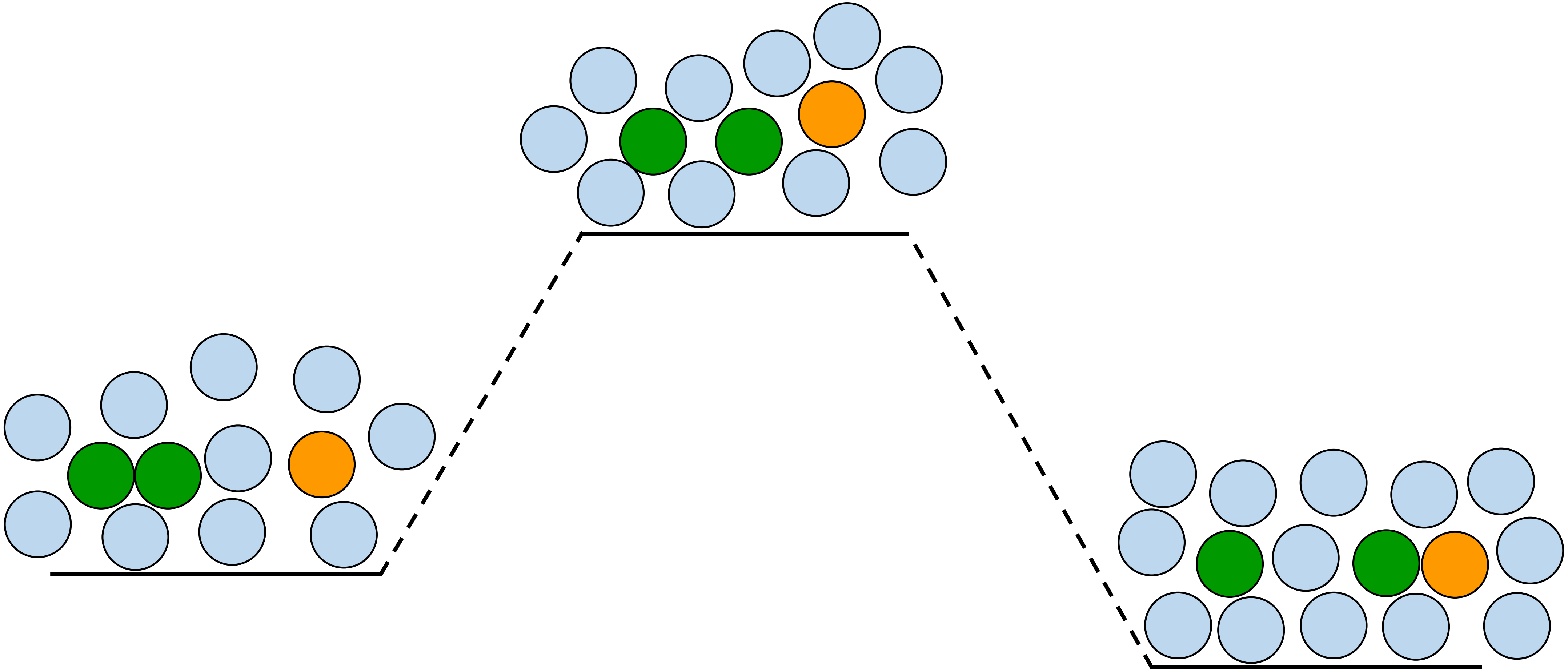
Setting the scene

Free energy, collective variables, enhanced sampling

In many cases, a “static” model of chemistry suffices



In many cases, a “static” model of chemistry suffices



Or not!

Most chemical problems are actually about ensemble properties.

→**Molecular dynamics simulations** can sample many configurations

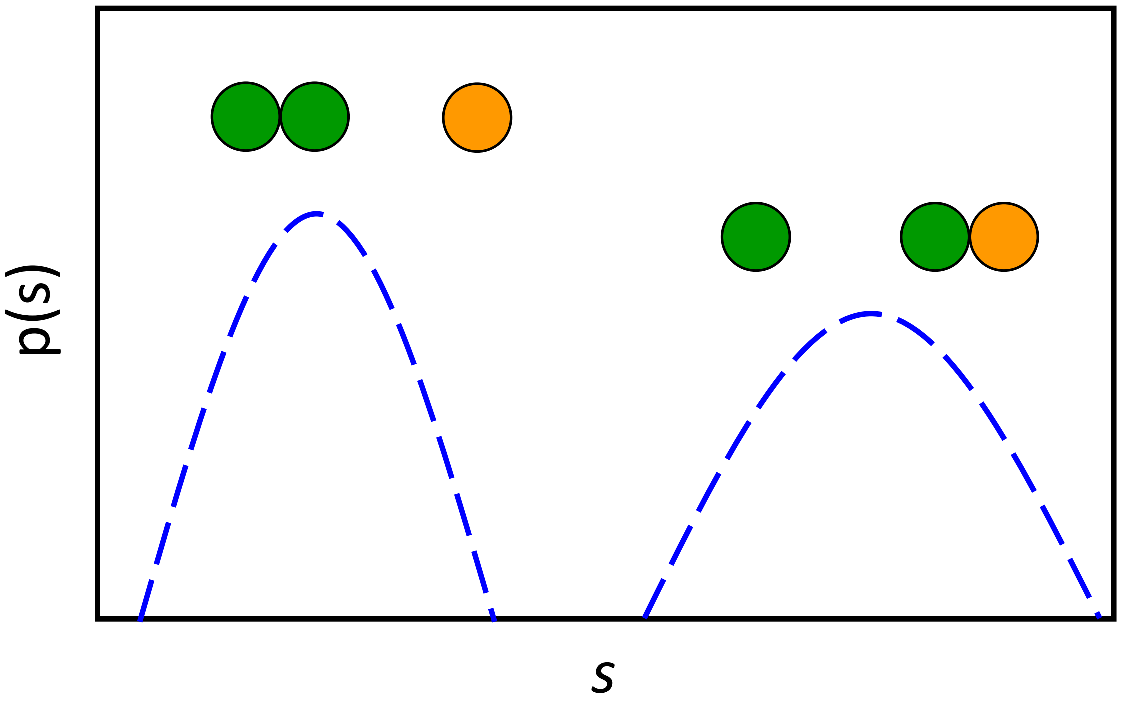
MD simulations produce high-dimensional data – hard to analyse!

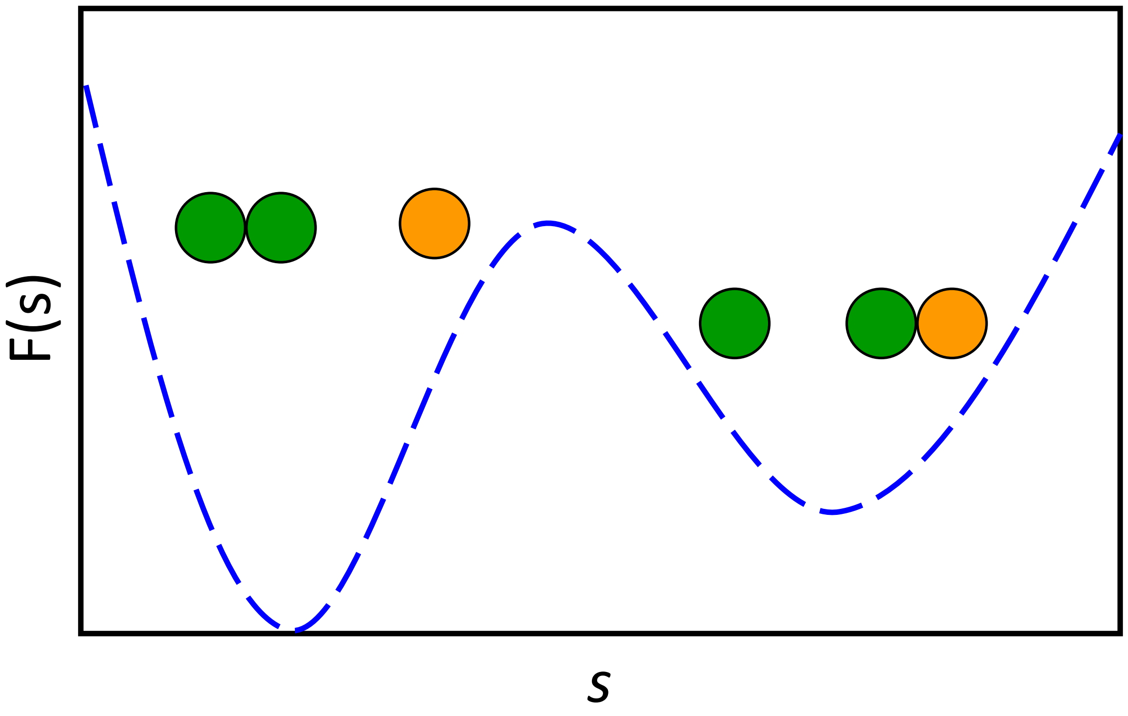
→Reduce dimensionality of the problem using **collective variables**

Set of CVs **s**: order parameters that can distinguish between relevant states of the system (preferably intuitively)

$$p(\mathbf{s}) = \langle \delta[\mathbf{s} - \mathbf{s}(t)] \rangle$$

$$F(\mathbf{s}) = -\frac{1}{\beta} \ln p(\mathbf{s})$$





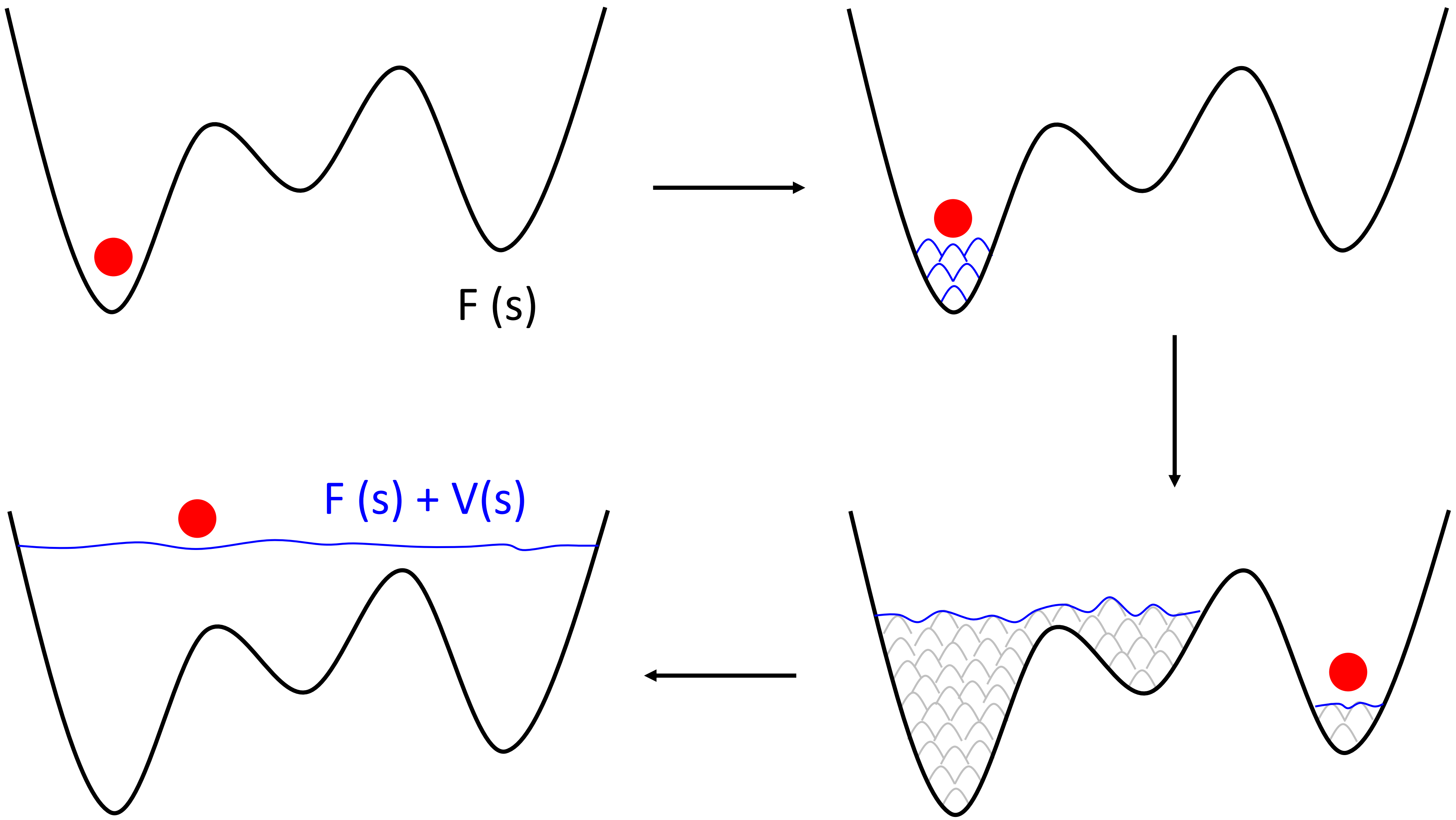
The free energy surface $F(\mathbf{s})$ in principle encodes all thermodynamic information of a transformation along $\mathbf{s}$.

But in practice it’s hard to efficiently sample the histogram $p(\mathbf{s})$!

→ System remains trapped in metastable states during (short) MD time scale

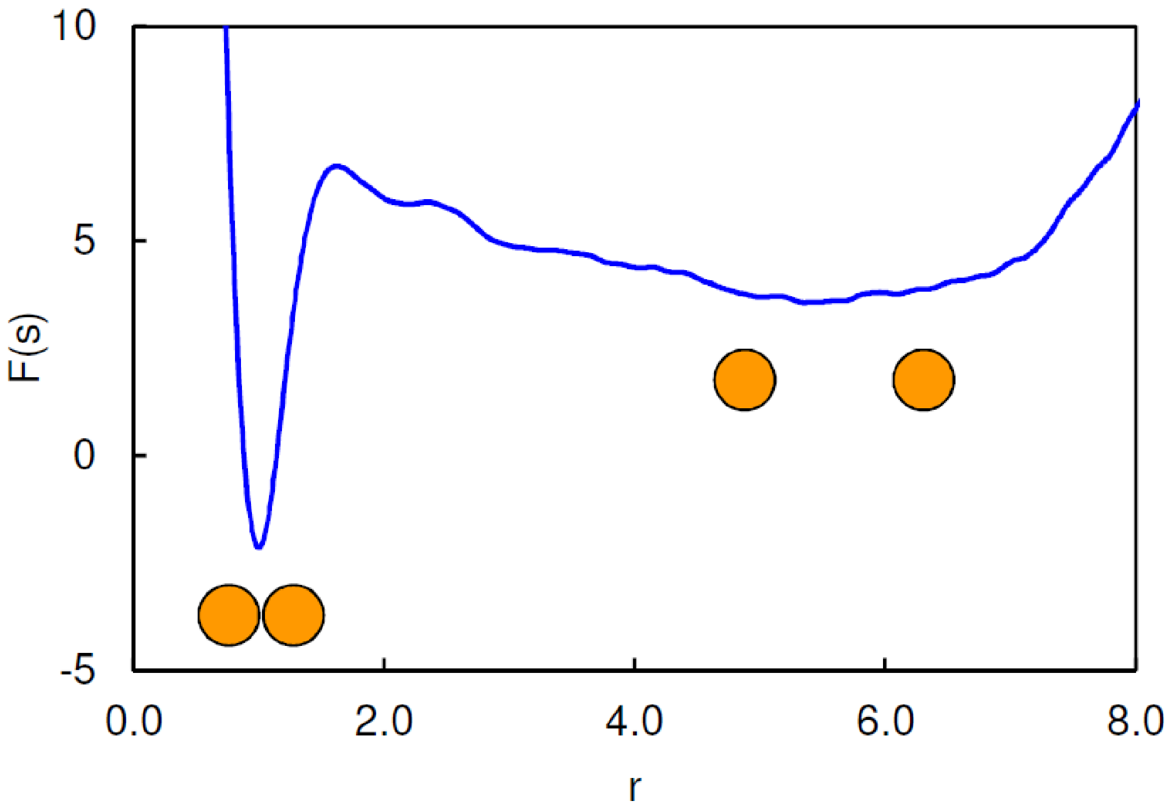
Need for **enhanced sampling methods**, e.g., metadynamics

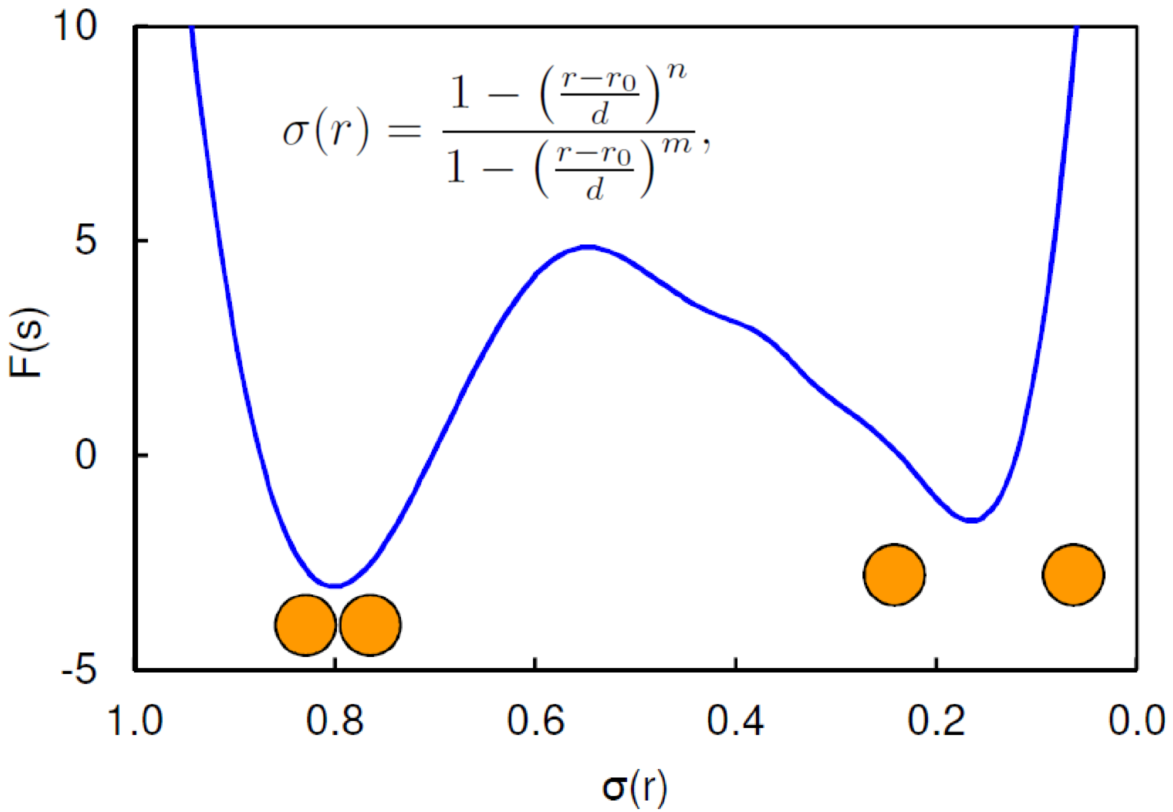
Metadynamics: add bias potential $V(\mathbf{s})$ to accelerate sampling



Need to reweight the histogram: $p(\mathbf{s}) = \langle w \cdot \delta[\mathbf{s} - \mathbf{s}(t)] \rangle_b$

Even for a single degree of freedom, we can choose many parametrizations of *s*.





$$\sigma(r) = \frac{1 - \left(\frac{r - r_0}{d}\right)^n}{1 - \left(\frac{r - r_0}{d}\right)^m},$$

Not an issue: integrate the probability density *p*(**s**) (per unit of **s**)

$$F_A = -\frac{1}{\beta} \ln \int_A \mathrm{d}\mathbf{s}\, e^{-\beta F(\mathbf{s})},$$

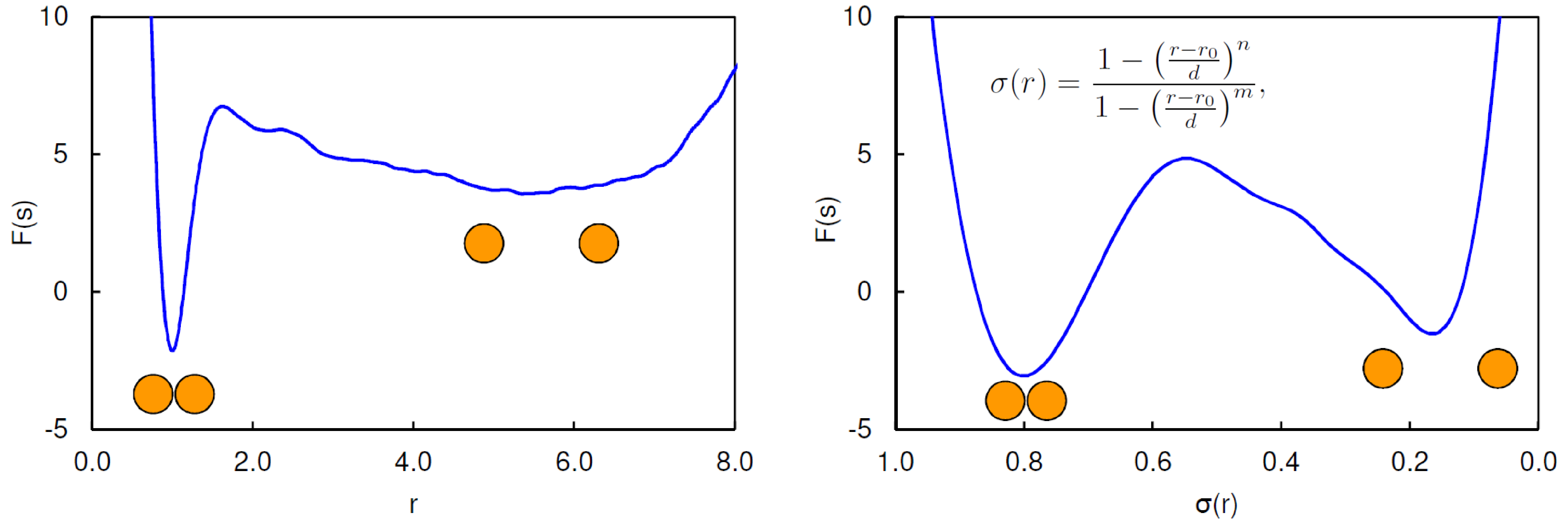
$$F_B = -\frac{1}{\beta} \ln \int_B \mathrm{d}\mathbf{s}\, e^{-\beta F(\mathbf{s})},$$

$$\Delta F_{A \to B} = F_B - F_A.$$

This way, free energy differences between states (and equilibrium constants) can be consistently computed

For this system (dimer in Lennard-Jones fluid), both CVs do give a consistent **binding free energy** Δ*F* = −1.4

So… how about the **free energy barrier** $\Delta^{\ddagger}F$?



$$\sigma(r) = \frac{1-\left(\frac{r-r_0}{d}\right)^n}{1-\left(\frac{r-r_0}{d}\right)^m},$$

Kinetics from the free energy - Pitfall 1

What is the free energy of the transition state?

F_{TS} = $F(\mathbf{s}_{\mathrm{TS}})$ is not the free energy of the transition state, because p_{TS} = $p(\mathbf{s}_{\mathrm{TS}})$ is a probability **density** which is not invariant to **s**.

So $F_{\mathrm{TS}} - F_{\mathrm{A}}$ is a meaningless quantity, and not the barrier!

It turns out that:

$$\Delta^{\ddagger} F_{A\to B} = F_{\mathrm{TS}} + \frac{1}{\beta} \ln \frac{\langle |\nabla s| \rangle_{s_{\mathrm{TS}}}^{-1}}{h} \sqrt{\frac{2\pi m}{\beta}} - F_A$$

$$\Delta^{\ddagger} F_{A\to B} = F_{\mathrm{TS}} + \frac{1}{\beta}\ln \boxed{\frac{\langle|\nabla s|\rangle_{s_{\mathrm{TS}}}^{-1}}{h}\sqrt{\frac{2\pi m}{\beta}}} - F_A$$

We see the partition function in the segment *s*

Essentially, we have to take out the degree of freedom along *s*

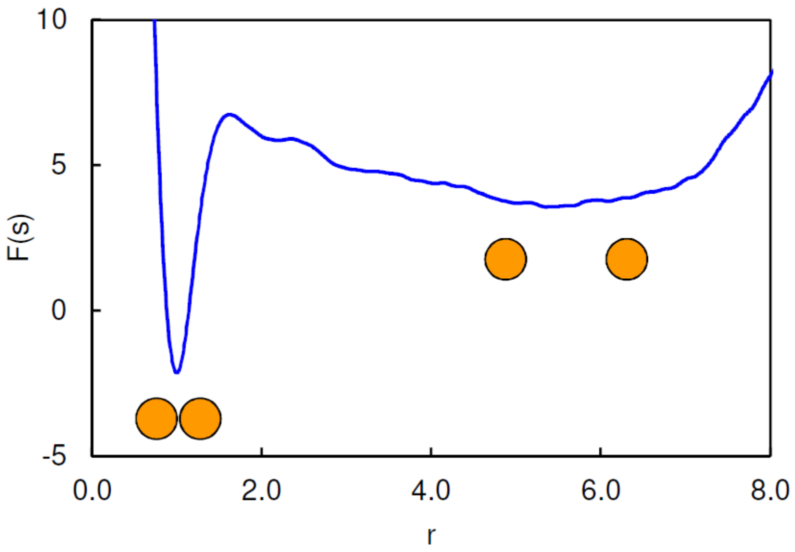
because there is no density along this direction

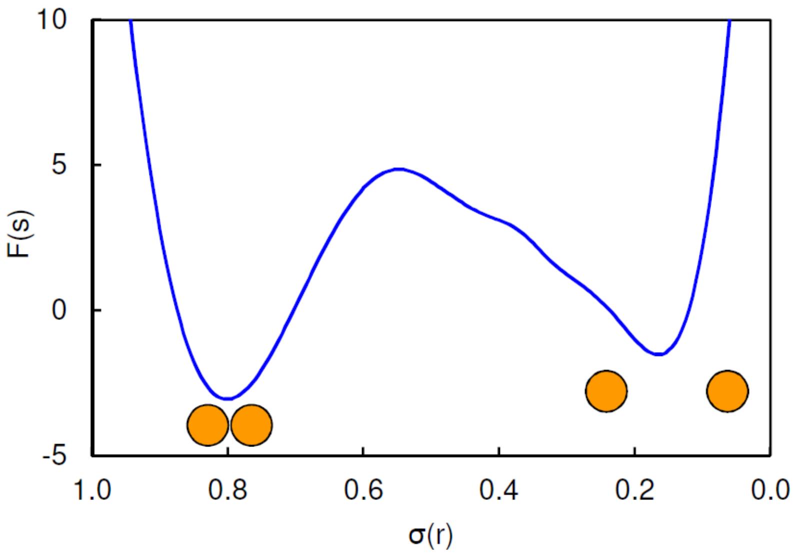
We can use transition state theory (TST) to get rates

$$k_{A\to B} = (h\beta)^{-1} e^{-\beta \Delta^{\ddagger} F_{A\to B}}$$

Consistency!

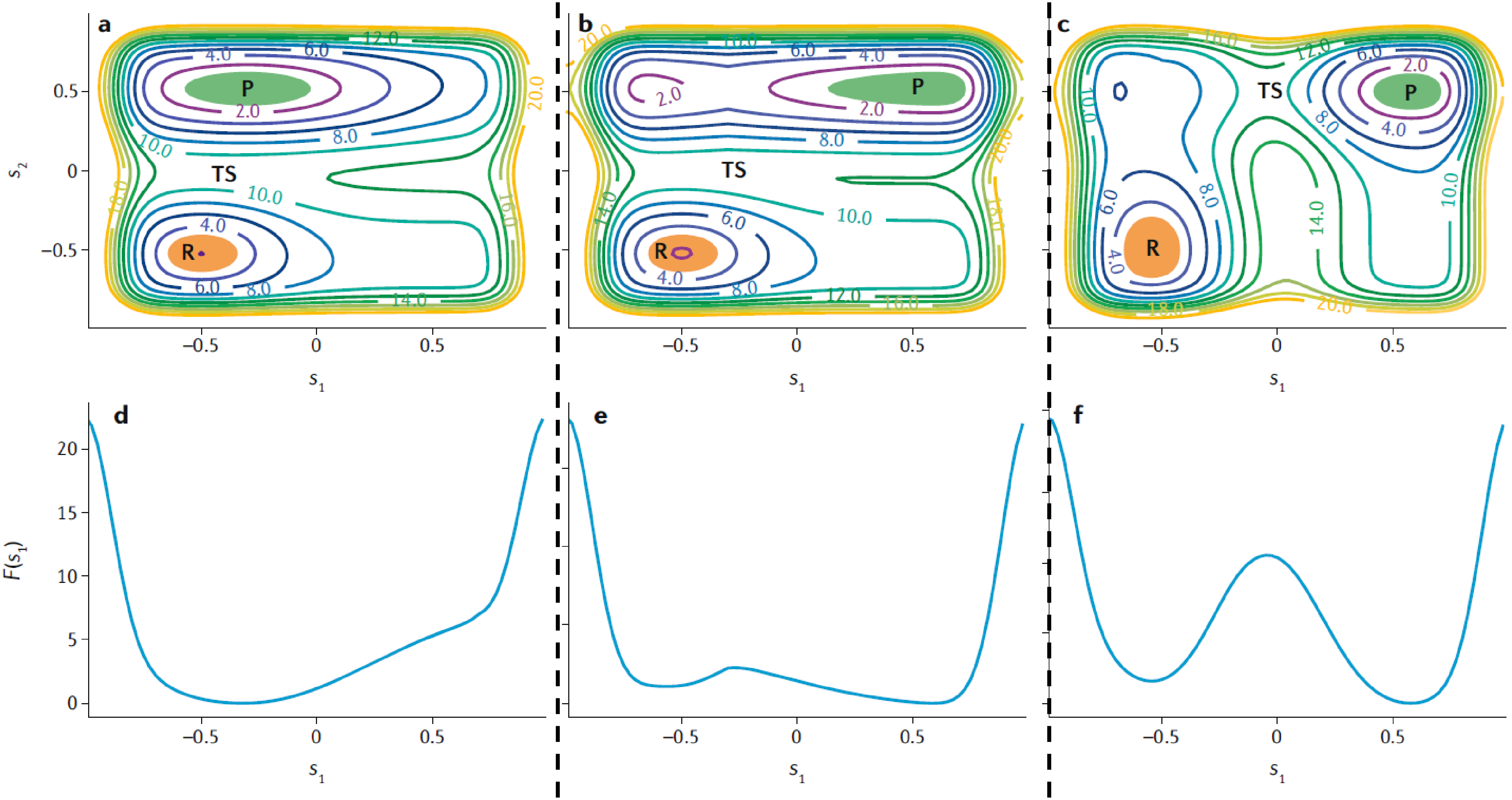
energy/CV	**r**	**σ(r)**
ΔF	−1.39	−1.38
$F_{TS} - F_A$	5.04	3.15
$\Delta^{\ddagger}F$	3.44	3.41





Kinetics from the free energy - Pitfall 2

Is our CV actually the reaction coordinate?



useless!

only thermodynamics

inefficient biasing

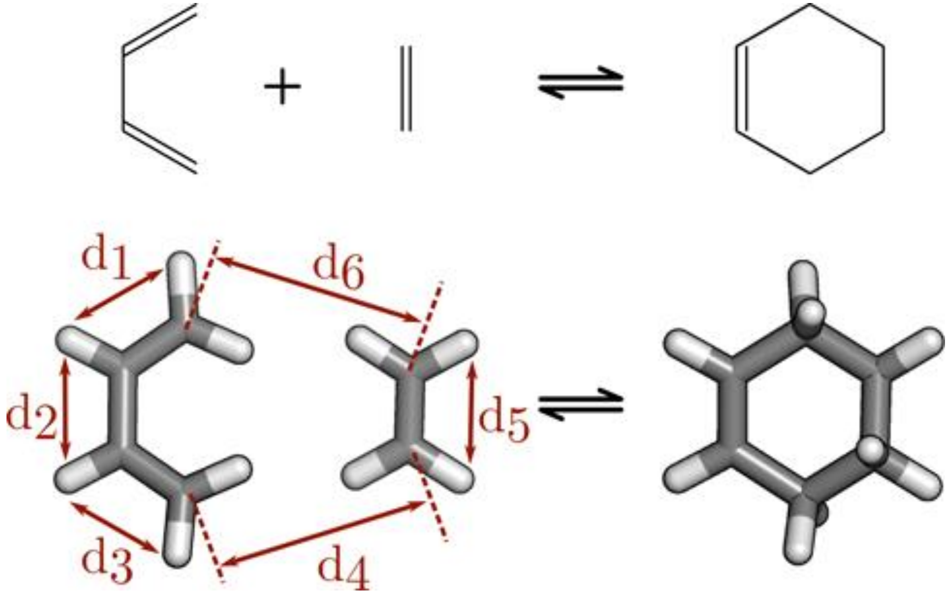
good barrier

best for biasing

Bussi & Laio, Nat. Rev. Phys. **2**, 200 (2020)

Many new methods promise to produce good CVs systematically:

1. Start with large set of candidate CVs $\mathbf{s} = (s_1, s_2, s_3, \ldots)$
2. Assume true RC = $f(\mathbf{s})$ (linear expansion, neural network, …)
3. Construct *a priori*, or refine iteratively during simulation

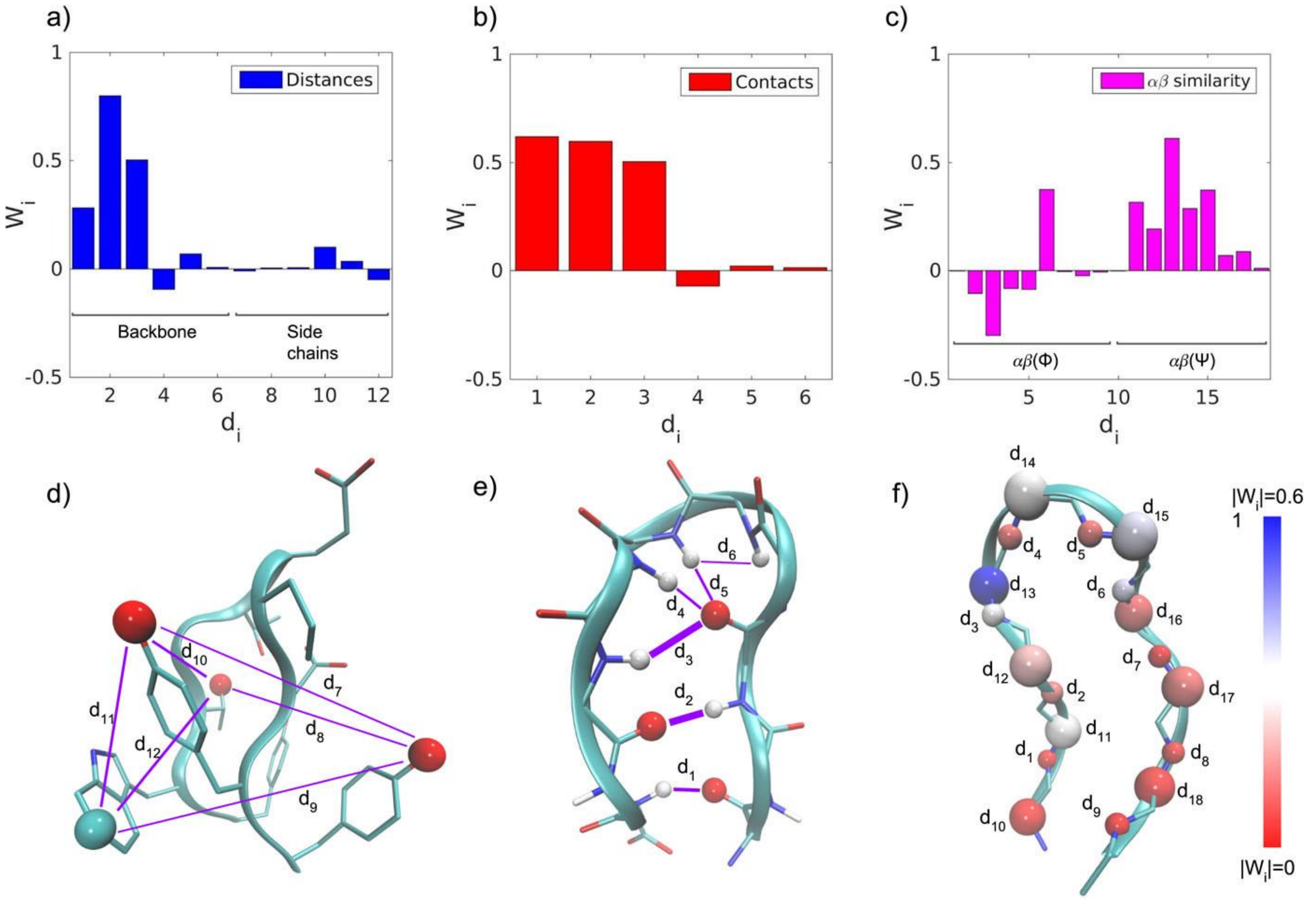


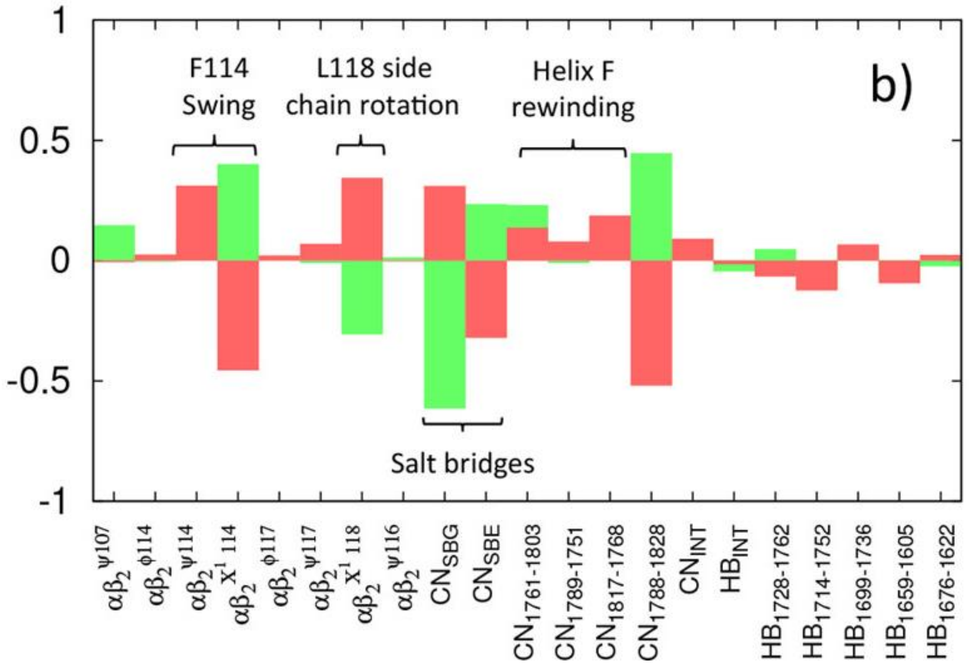
d_1	d_2	d_3	d_4	d_5	d_6
0.05	−0.21	0.08	−0.69	0.00	−0.68

Mendels *et al.,* J. Phys. Chem. Lett. **9**, 2776 (2018)

Need for good CVs: bug or feature?

Good CV = good sampling, but also physical insight!

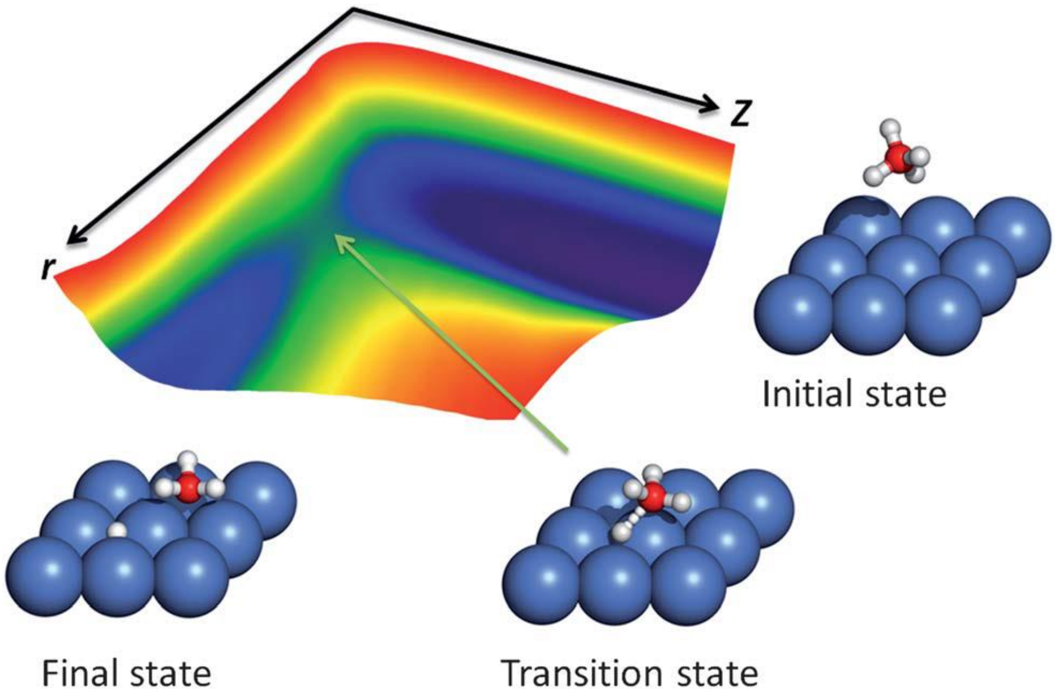


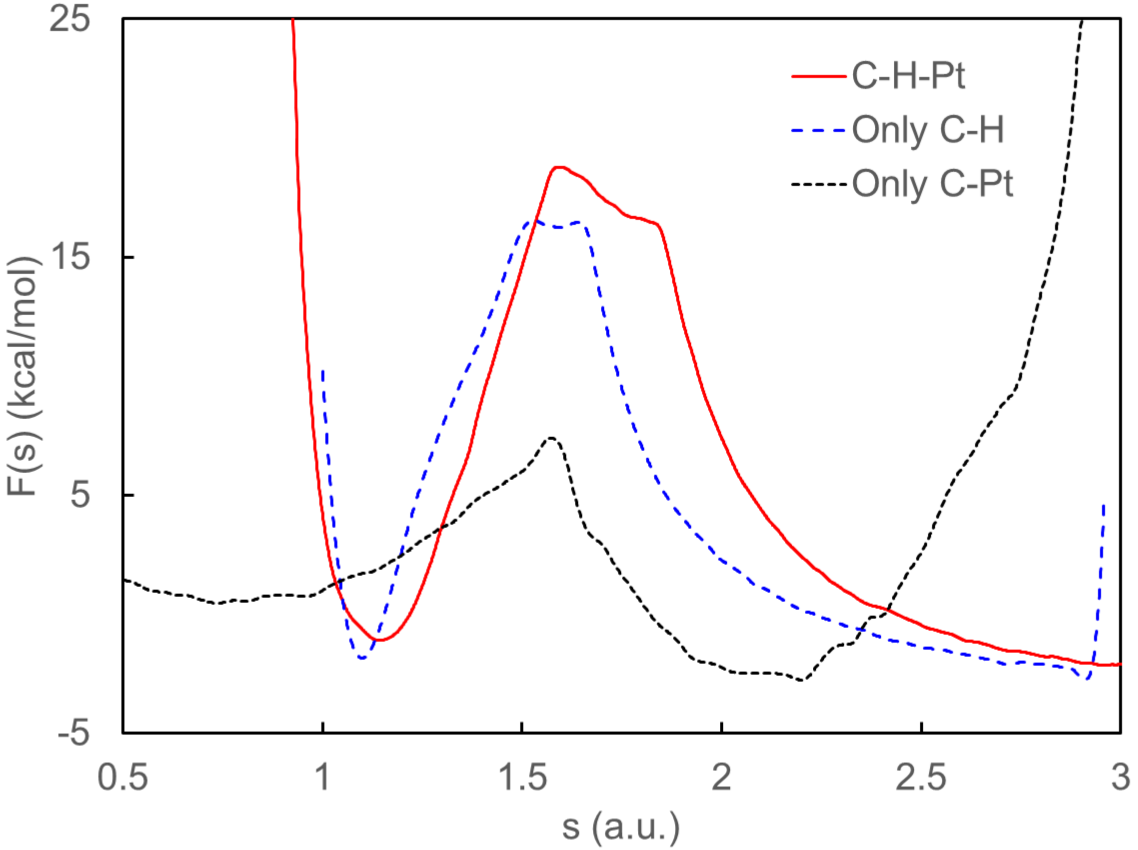


Mendels *et al.,* J. Chem. Phys. **149**, 194113 (2018)

Brotzakis & Parrinello, J. Chem. Theory Comput. **15,** 1393 (2019)

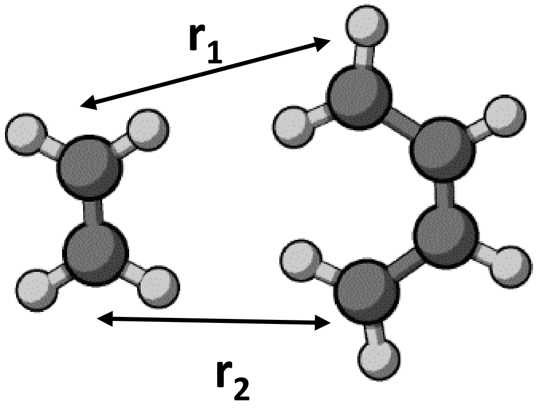
Constructing a CV for CH_4 chemisorption @ Pt(111)

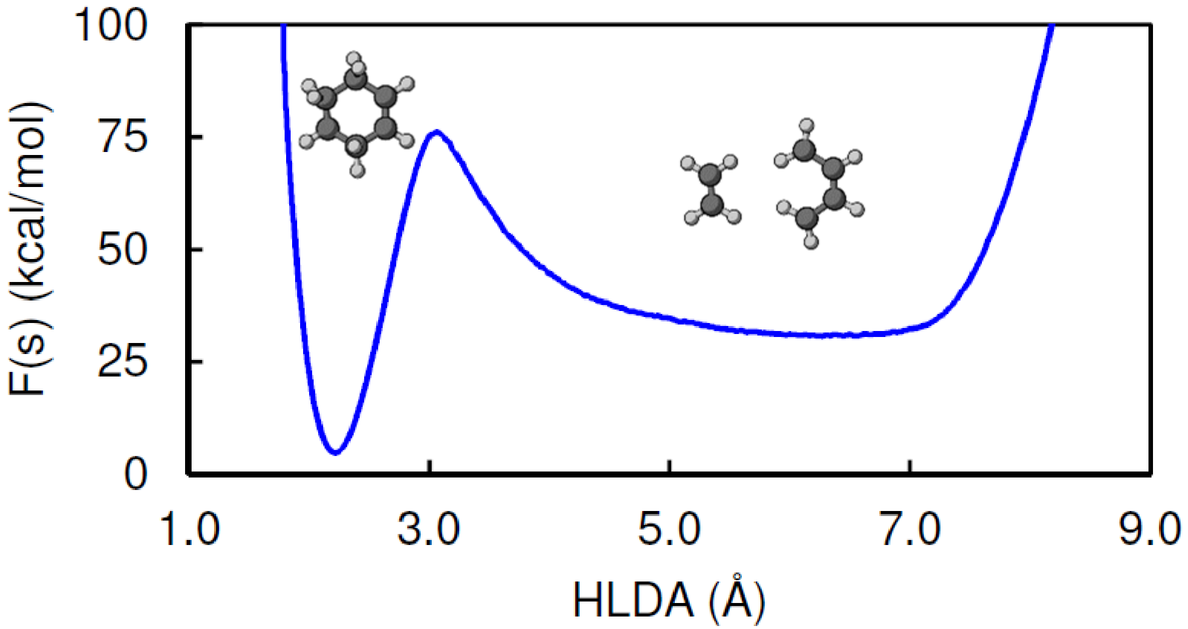




TST give the classical upper bound of the true rate

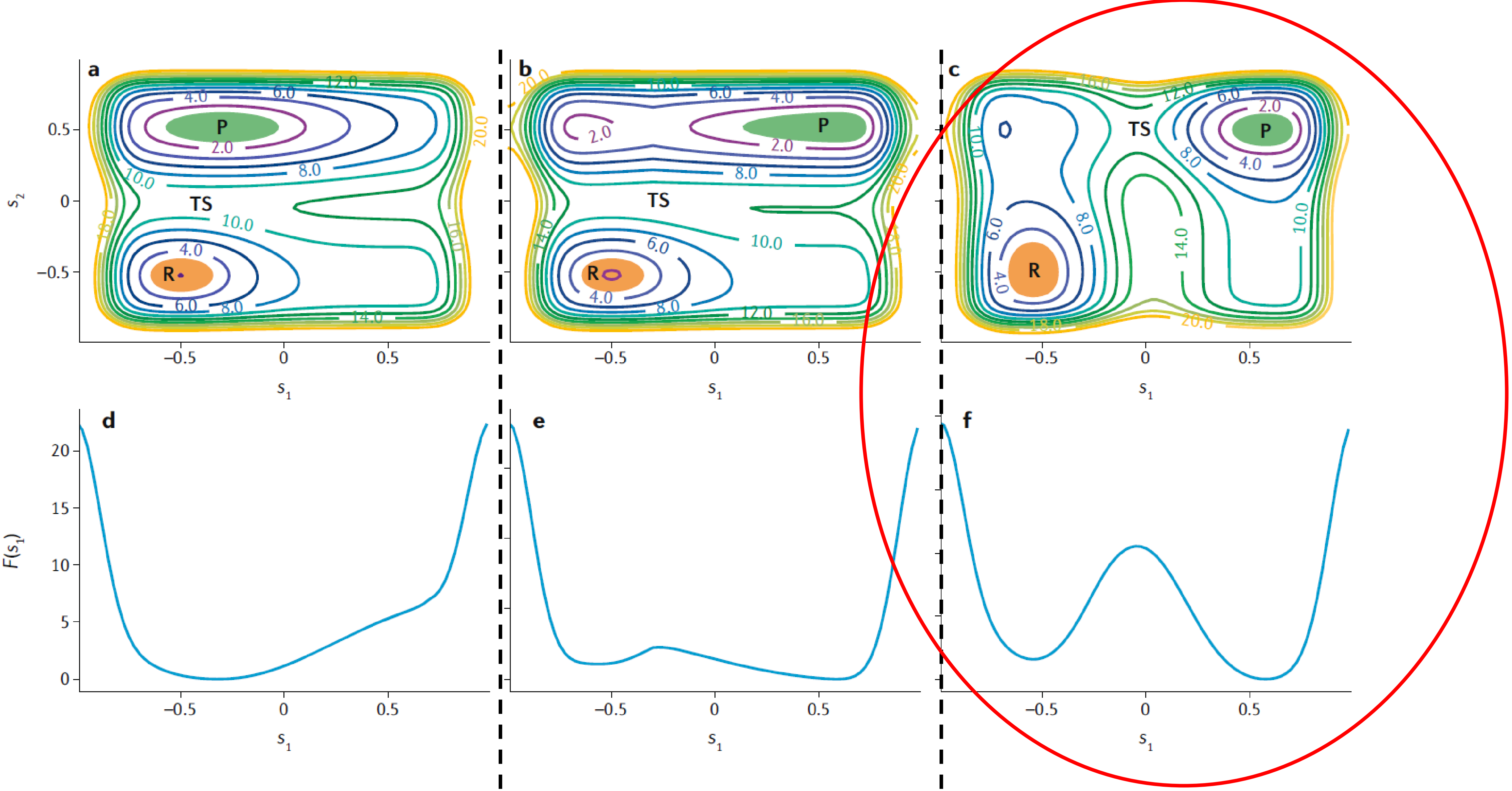
So: the best CV maximizes the barrier





Project FES on several trial CVs

- Two body vs. manybody
- 1D vs. 2D



useless!

only thermodynamics

inefficient biasing

good barrier

best for biasing

Bussi & Laio, Nat. Rev. Phys. **2**, 200 (2020)

	HLDA	r_1	SPRINT1	SPRINT2	$\overline{\sigma}$	(r_1, r_2)
ΔF	-23.25	-23.34	-23.25	-23.25	-23.25	-23.25
$\Delta^{\ddagger} F$	47.01	29.01	32.66	46.51	47.16	47.23

Two body r_1 is an all-round bad CV

Several manybody CVs are decent 1D reaction coordinate

If we “give up” and go 2D, we get the best (highest) barrier

Ultimately, things are quite robust w/r/t the CV

How good are rates?

- We (indirect): $(7.8 \pm 1.7) \cdot 10^{-5}\ \mathrm{s}^{-1}$
- Direct approach (Fu *et al.*): $(5.8 \pm 0.7) \cdot 10^{-5}\ \mathrm{s}^{-1}$

Direct approach = explicit sampling of reactions (with bias)

But we need way shorter MD runs!

- We: 5 ns (250 ps for quick ‘n’ dirty)
- Fu *et al.*: up to 2 µs

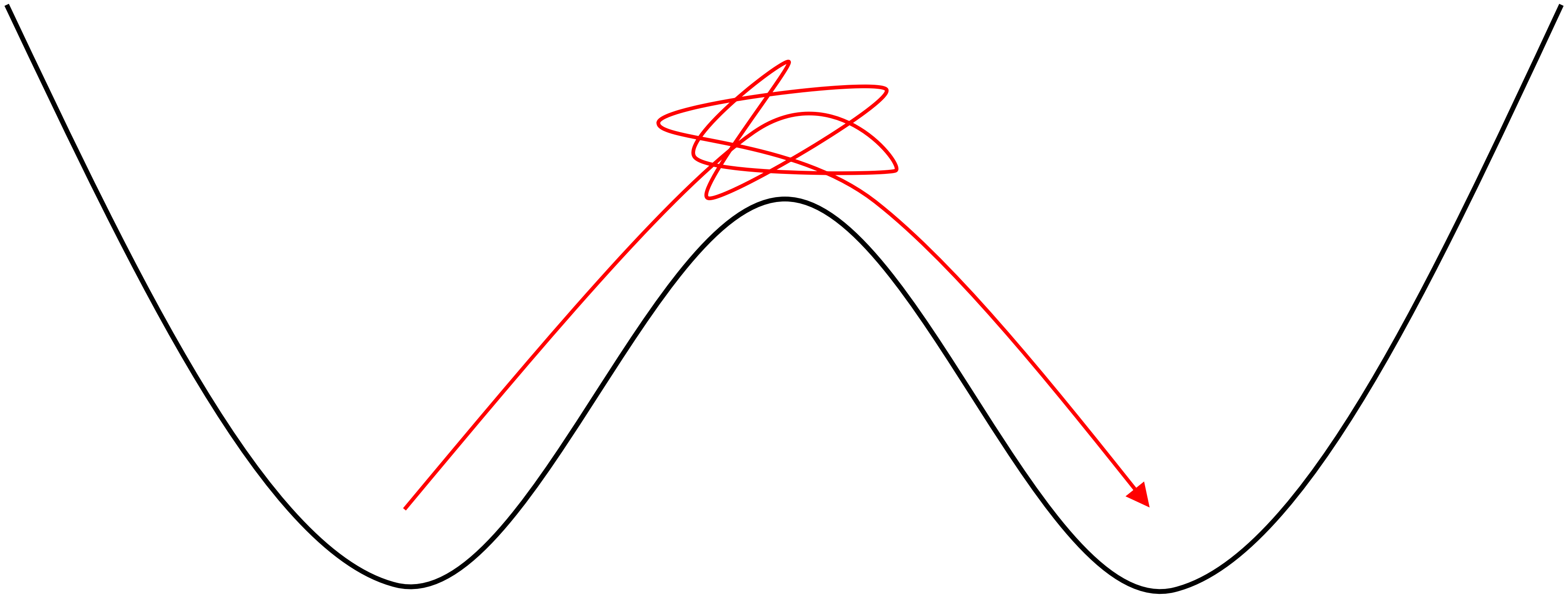
In principle we could go *ab initio*!

Fu *et al.,* J. Chem. Theory Comput. **13**, 968 (2017)

Kinetics from the free energy - Pitfall 3

Do we move in a straight line?

TST rate = frequency of crossing TS in a certain direction



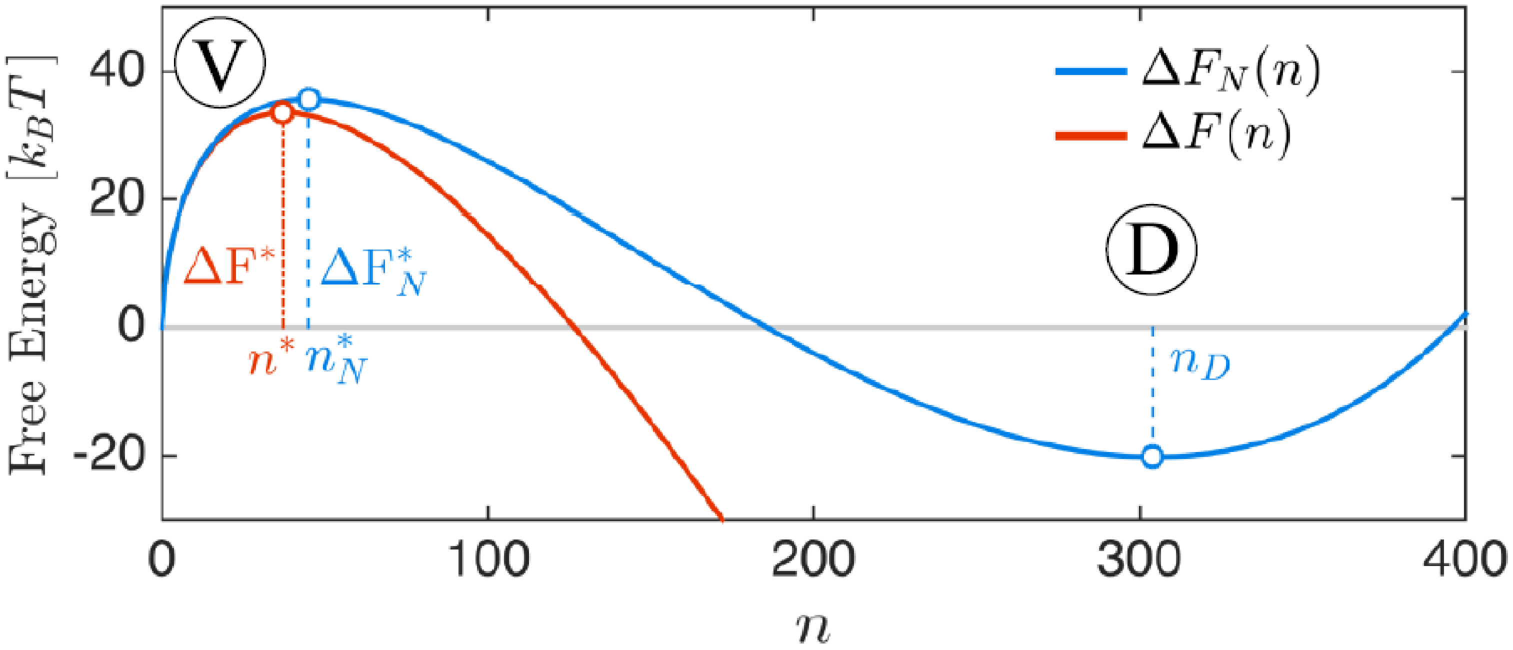
Not every TS crossing corresponds to an actual transition.

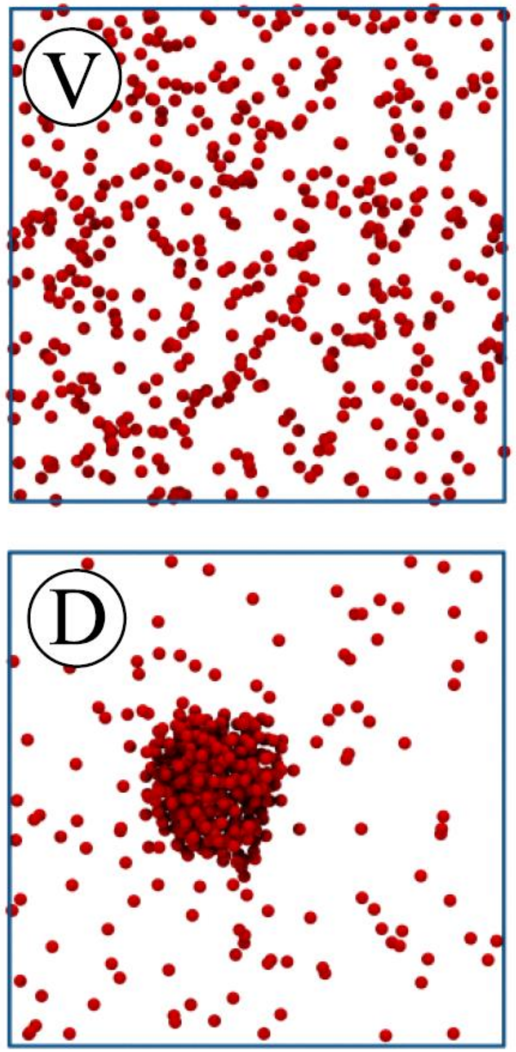
Correct the Eyring equation with a transmission coefficient κ ≤ 1

$$k_{A\to B} = \frac{\kappa}{h\beta} e^{-\beta \Delta^{\ddagger} F_{A\to B}}$$

Droplet (D) nucleation from supersaturated Ar vapor (V)

No way that TST holds, right?





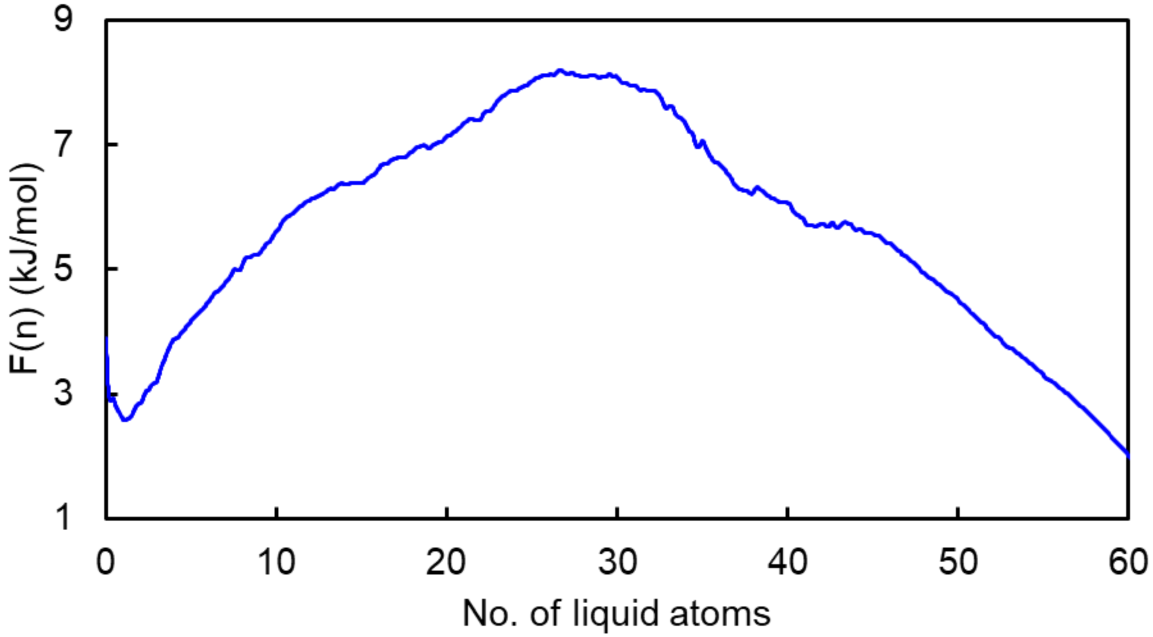
Salvalaglio *et al.*, J. Chem. Phys. **145**, 211925 (2016)

$\Delta^{\ddagger}F$ = 6.3 kJ/mol for nucleation @ 80.07 K (supersaturation of 9.87)

Average time to nucleation

- Direct approach: 690 ns
- TST: 8 ns

So **κ ≈ 0.012**



Short MD runs starting at TS show about 70 crossing attempts

before a liquid state is finally reached: **κ = 0.014 ± 0.004**

It works… and orders of magnitude faster than direct approach!

Wrapping up

Conclusions

Conclusions

Studying kinetics using free energy approaches is almost as generic as studying thermodynamics (and can be done simultaneously)

Some challenges, but:

1. We learn more about our process of interest
2. Mostly shared with the field at large (so active development)

Applications only limited by imagination, including **yours**

Key reference: Bal *et al.,* J. Chem. Phys. **153**, 114118 (2020)

Acknowledgments

Thanks to Satoru Fukuhara for starting this

I took simulation ideas and set-ups from Christopher Fu, Pablo Piaggi, GiovanniMaria Piccini, Dan Mendels, Matteo Salvalaglio, …

Erik Neyts for the support

Universiteit Antwerpen

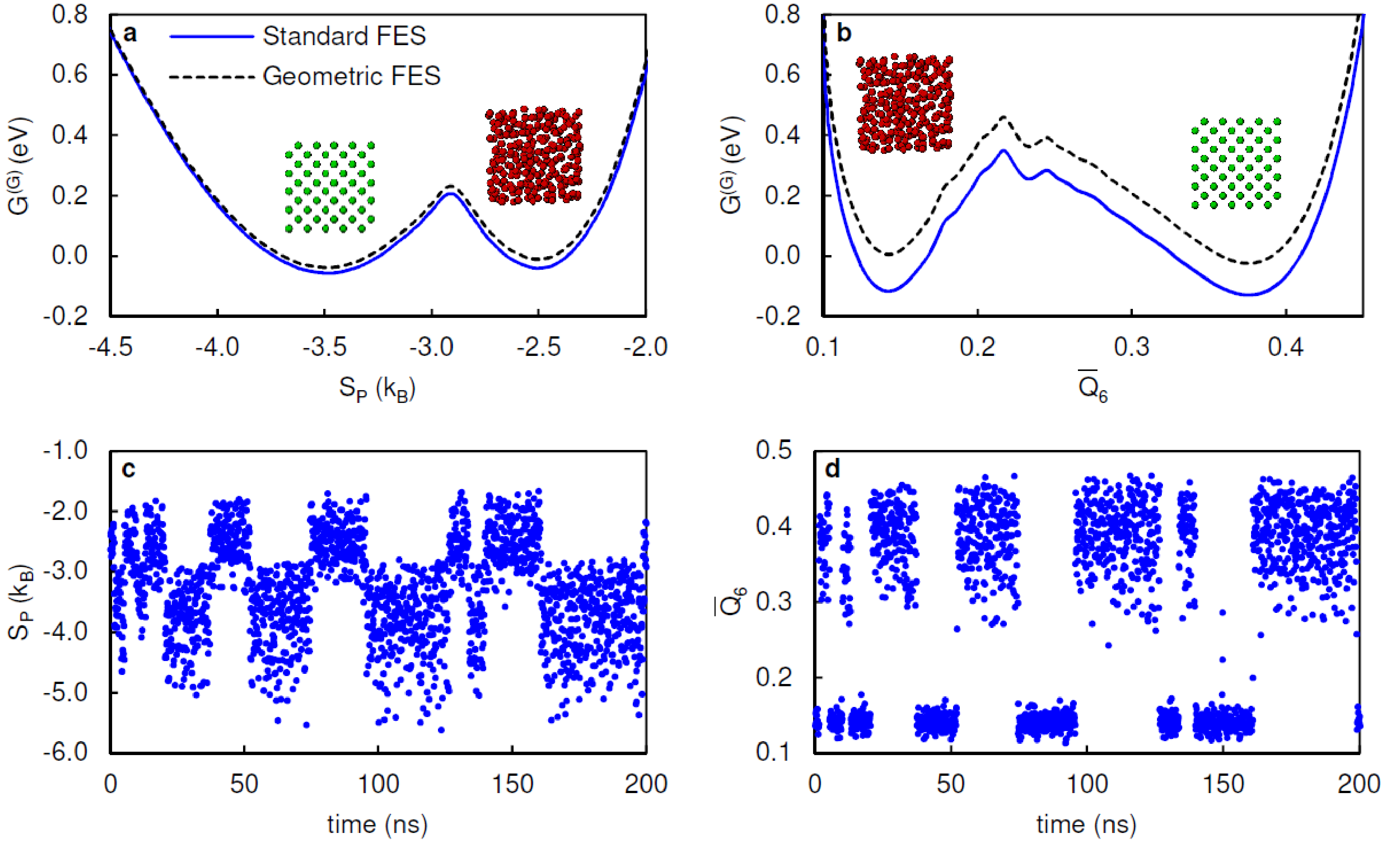
fwo

VSC Vlaams Supercomputer Centrum

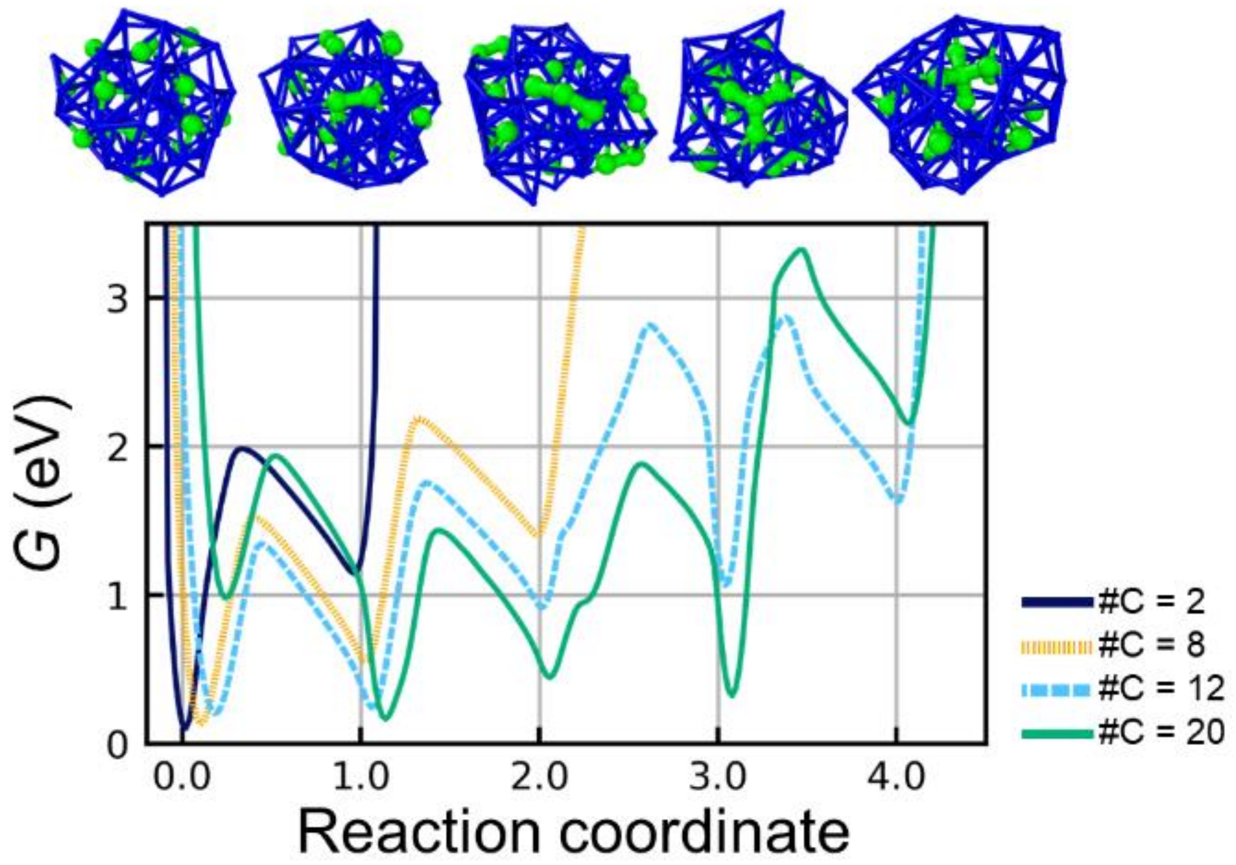
Encore

Some extra examples

Another manybody CV example: liquid Na crystallization

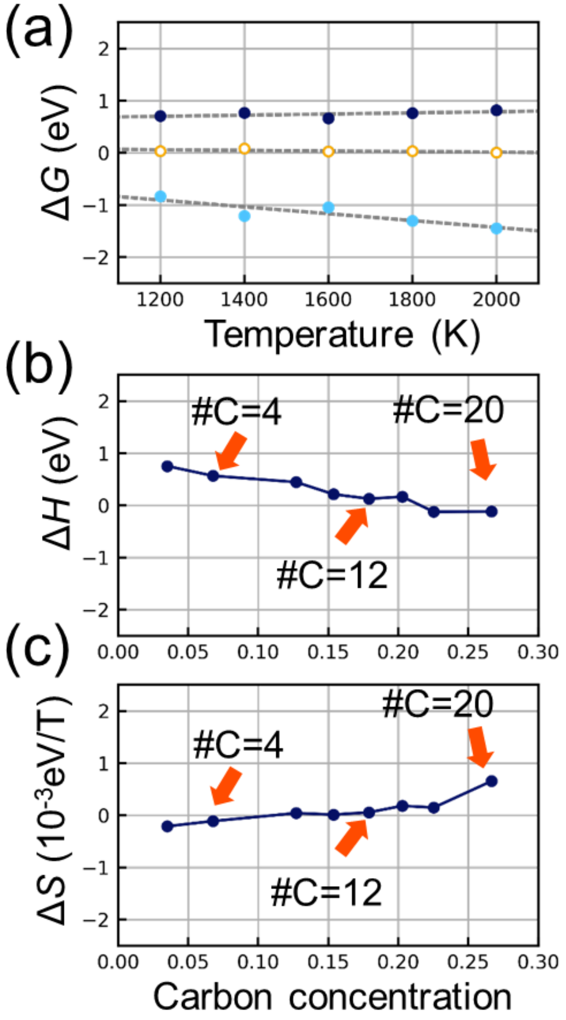


Initial stages of carbon nanotube nucleation (S. Fukuhara)



Initial stages of carbon nanotube nucleation (S. Fukuhara)

reaction **energy**



reaction **barrier**