

Sampling challenges in studying membrane helix-helix dimerization

Trieste, May 2017

Oxford

Mostly Cloudy

19°

Tuesday Today

20 13

Now	3PM 40%	4PM 60%	5PM 70%	6PM 60%	7PM 30%	8P
19°	19°	19°	18°	17°	17°	17°

Wednesday 16 7

Thursday 17 7

Friday 16 7

Saturday 15 7

Sunday 17 7

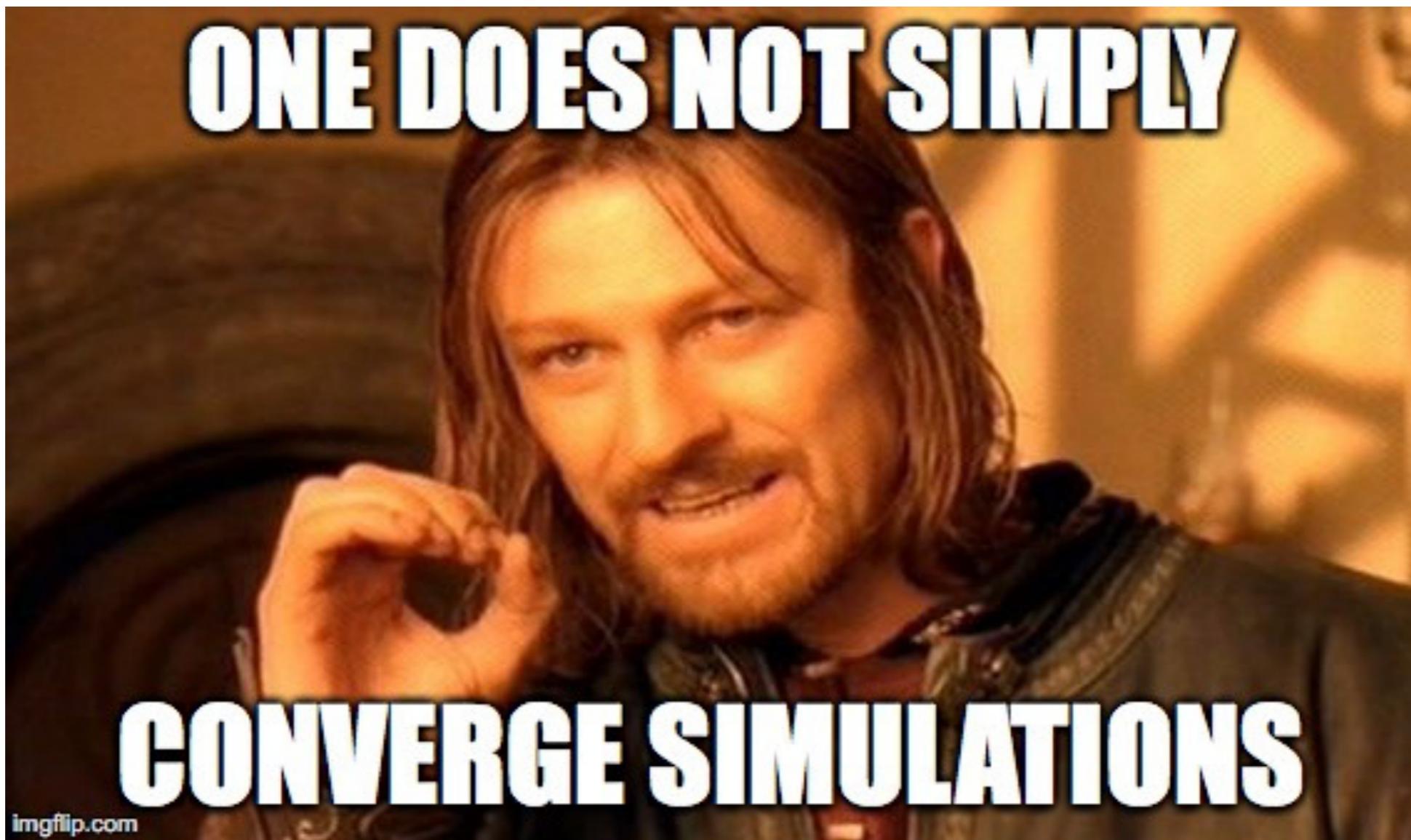
Monday 19 9

Tuesday 16 9

What am I working on?

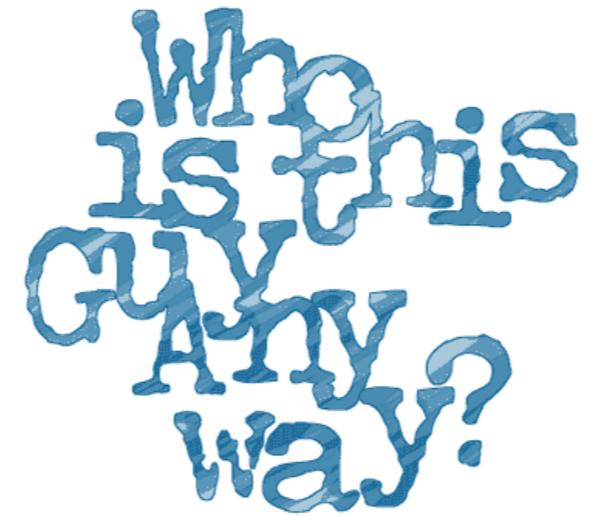
- Protein folding – in membranes
- Simple model systems, to understand if computational models are accurate enough
- Convergence?
- It is also (possibly) my last science talk

Let's keep this fun :)



Key terms

- This is a specialist meeting, so hopefully everybody is familiar with
 - CV – collective variable/RC reaction coordinate
 - PMF – potential of mean force
 - RE – replica exchange
 - US – umbrella sampling



- Currently work in Oxford/NIH
- Will show some cool tricks for what you can do with PLUMED
- Mainly focused on replica exchange umbrella sampling simulations



Robert
Best



Mark
Sansom



Phil
Stansfeld

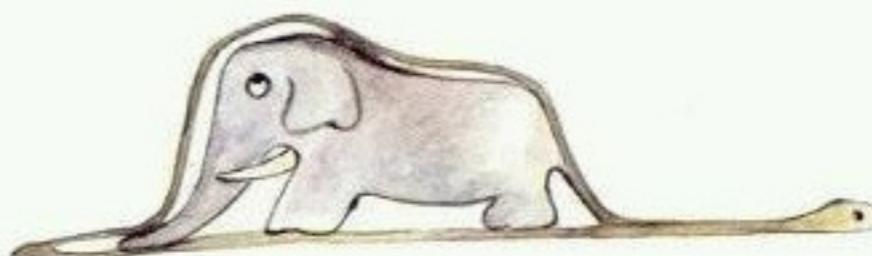
- Used to make my dollar here (very different approach to sampling...)

D E Shaw Research

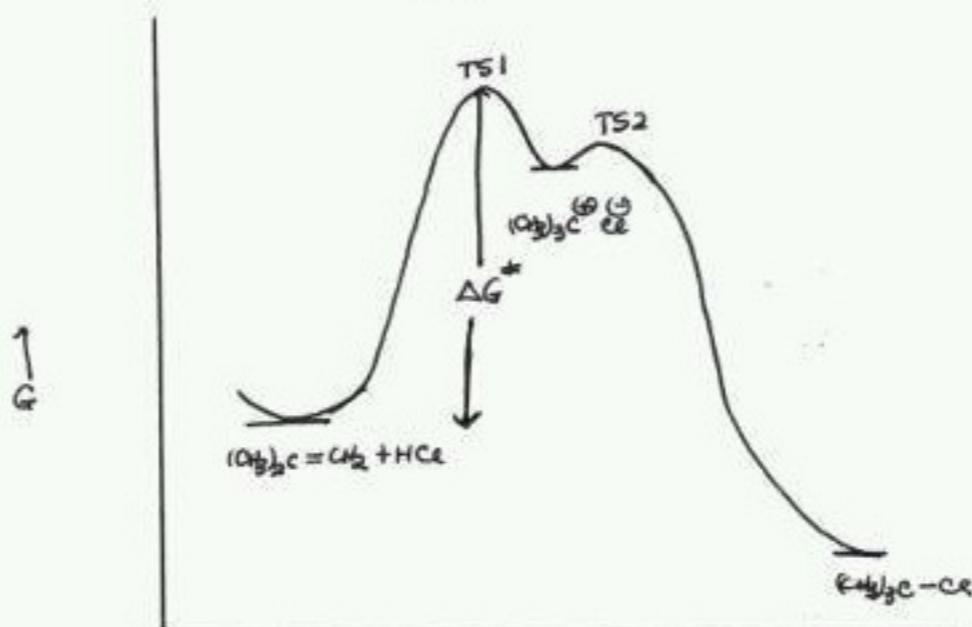
What does this look like to you?



If it looks like a plain hat, you're an adult



If it looks like an elephant being eaten by a
boa constrictor, you're still a child at heart



If it looks like a free-energy
reaction diagram, you're in organic
chemistry and your life is over.

Le Petit Prince



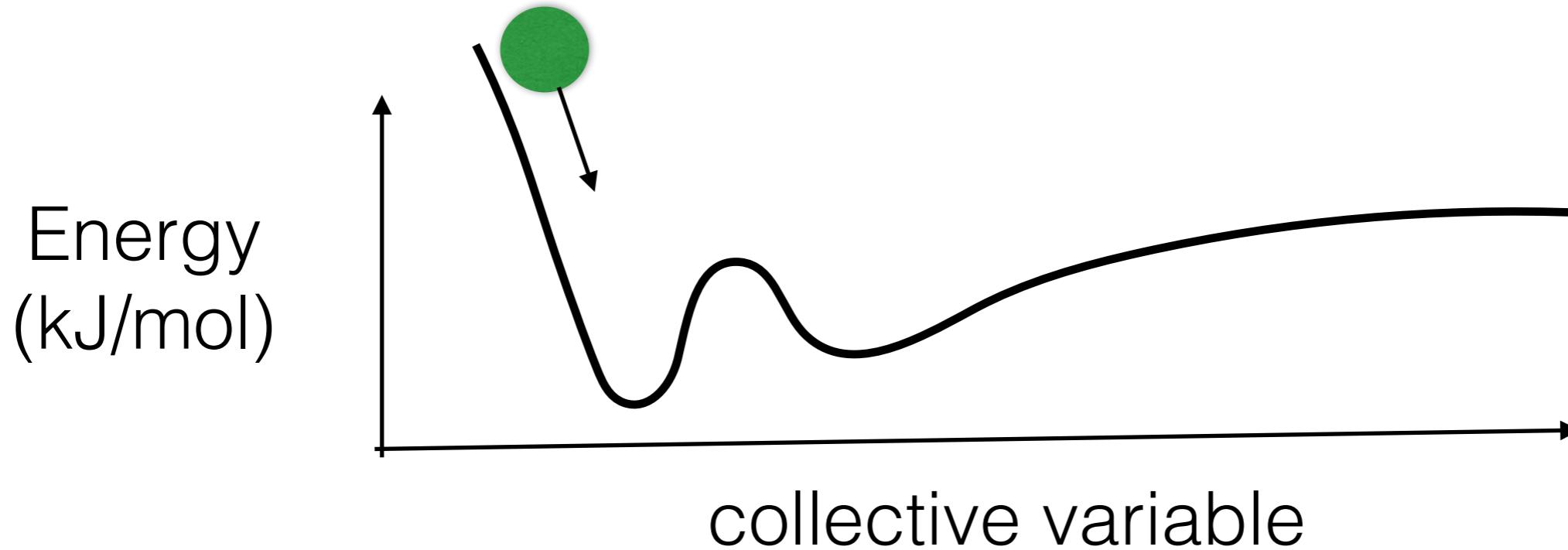
Let's talk convergence

- If simulation is claims to be converged this should be demonstrated... somehow
- “Convergence” is not true/false but should be treated more quantitatively
- Simple principles that apply generally, not just for replica exchange umbrella sampling

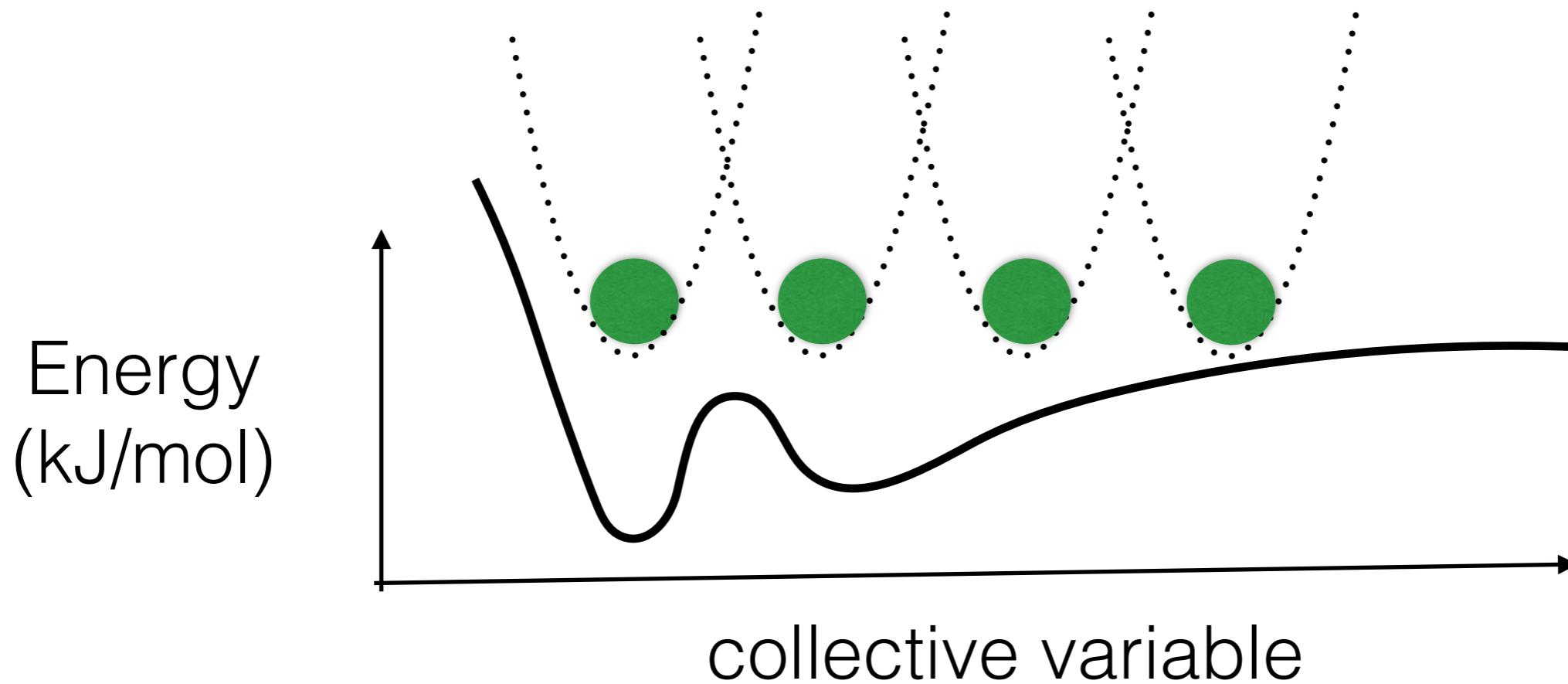
Simple principles

- Independent of initial condition
 - histogram overlap
 - PMF overlap
- Diffusion along the CV
 - After ‘demuxing’
 - When using the negative of the free energy surface should get barrier-free diffusion limited by the inherent viscosity of the system

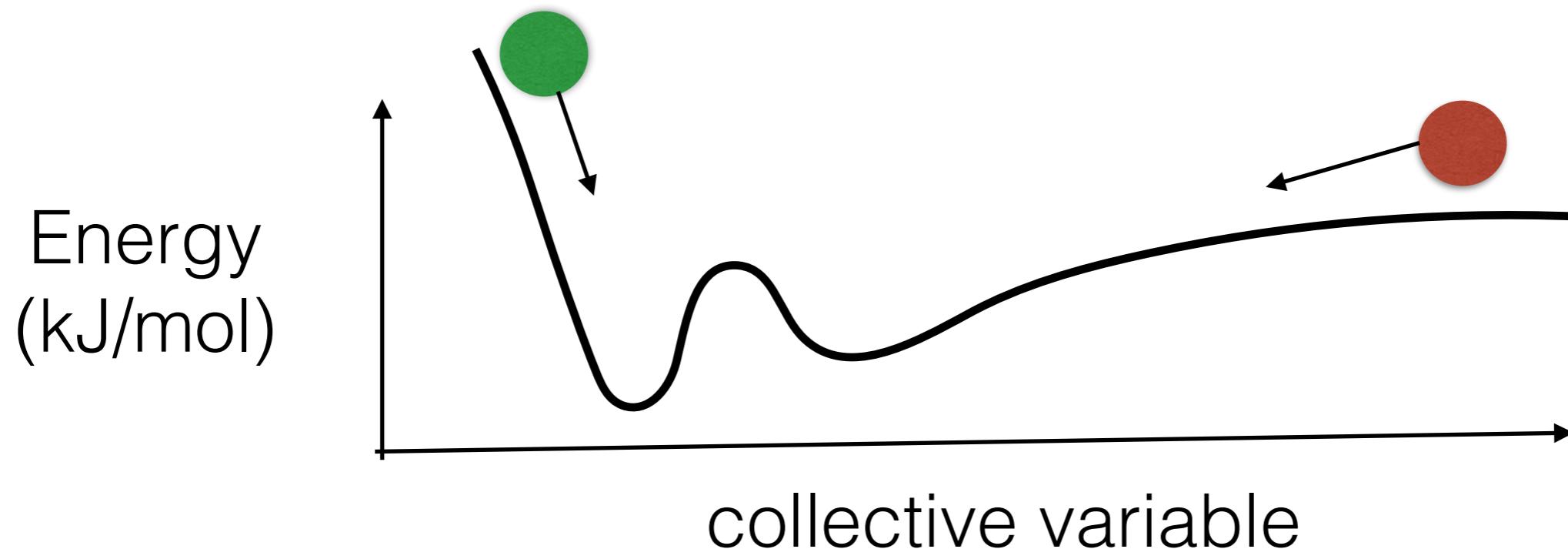
Suppose a CV and an energy surface



In replica exchange umbrella sampling

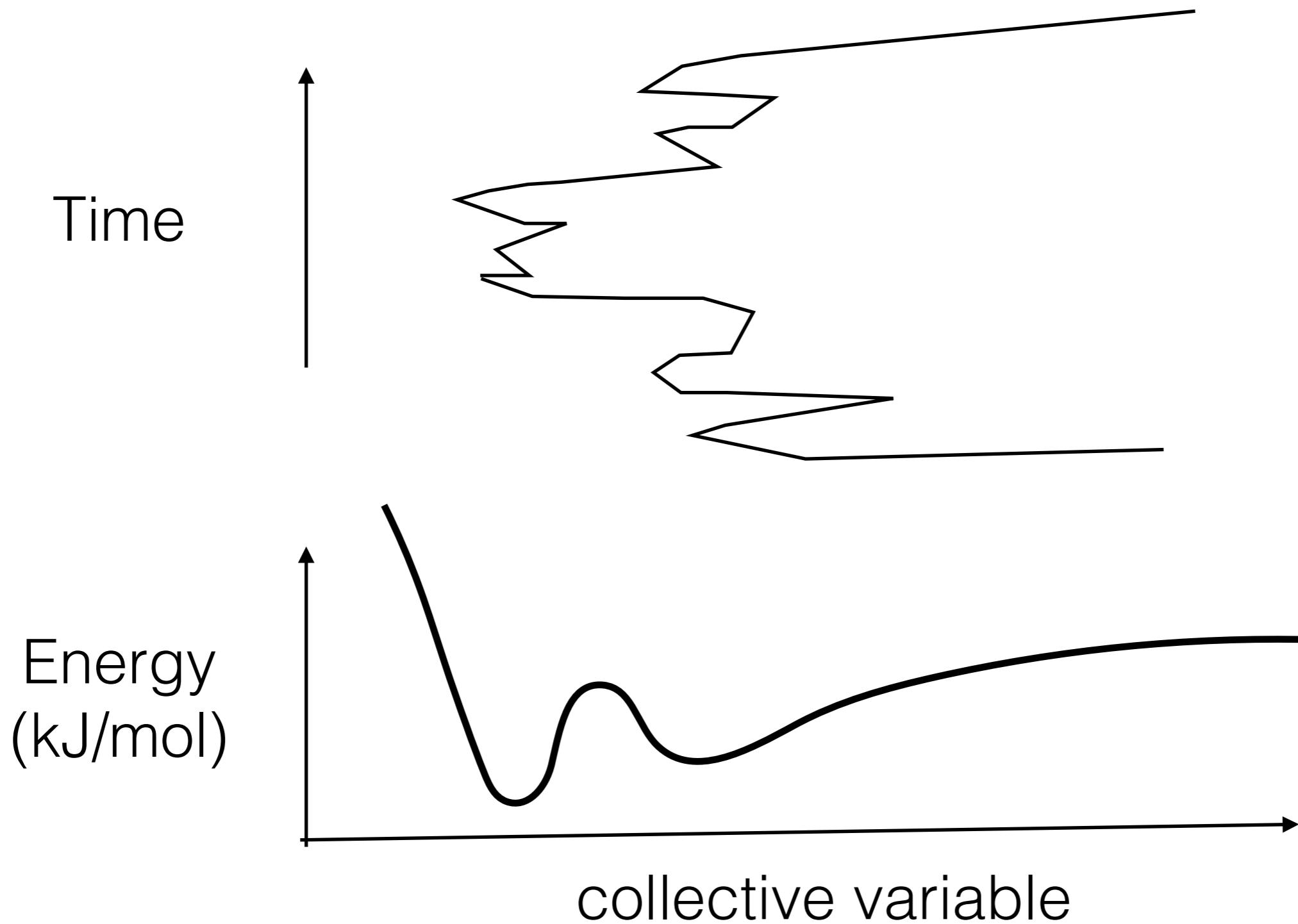


Independent of starting configuration

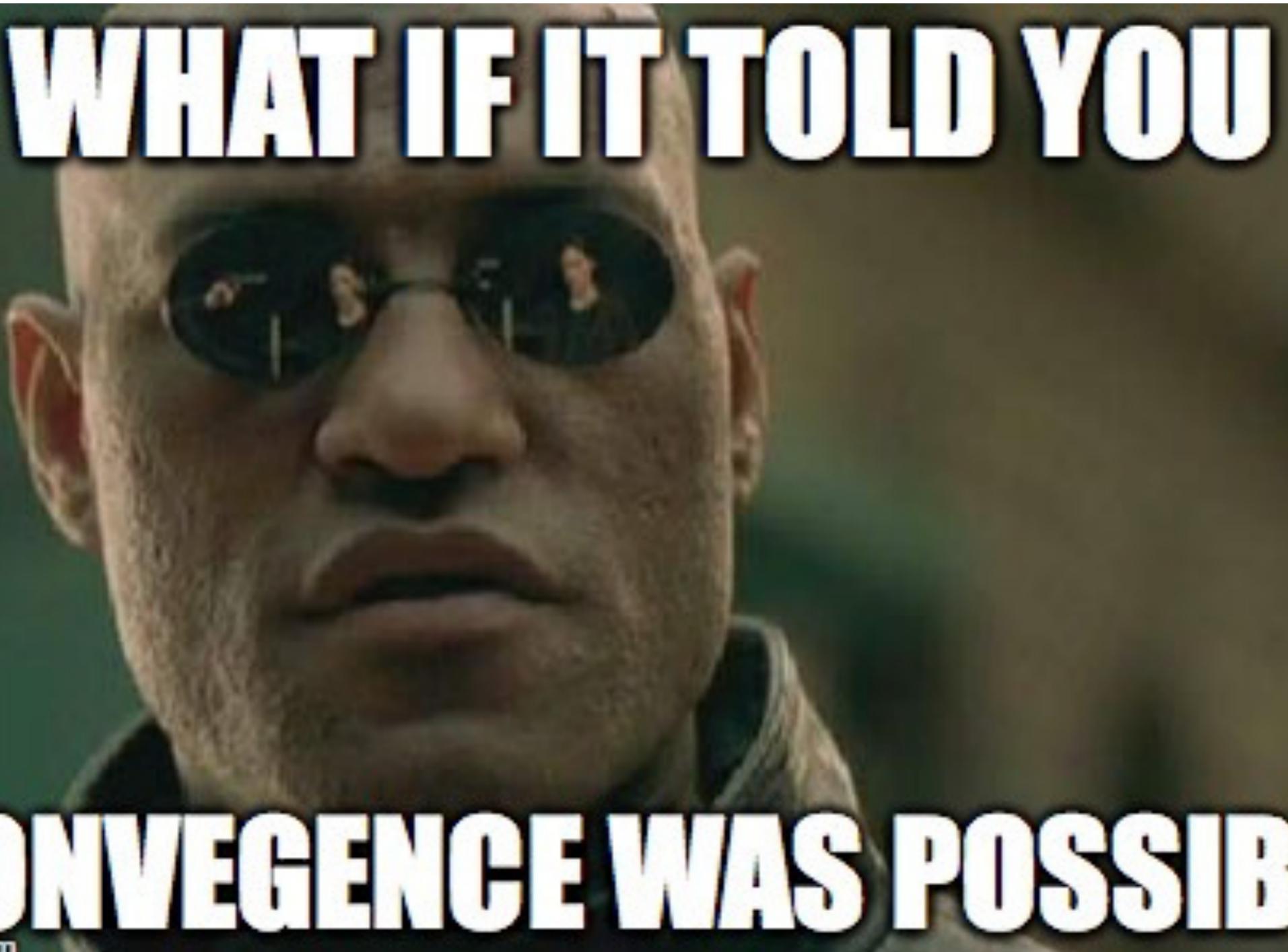


Results obtained starting from ● should be the same as starting from ●

Multiple transition/diffusivity over the CV



WHAT IF IT TOLD YOU



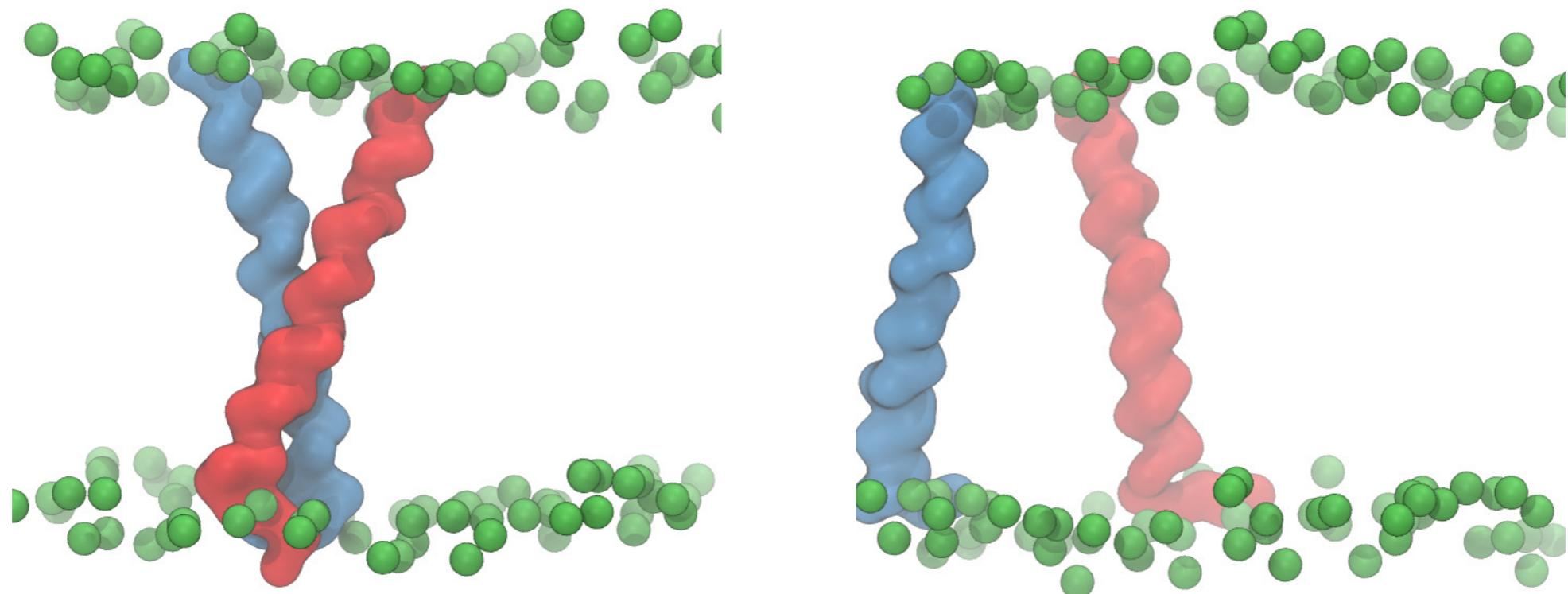
CONVEGENCE WAS POSSIBLE

imgflip.com

... in helix-helix dimerization simulations

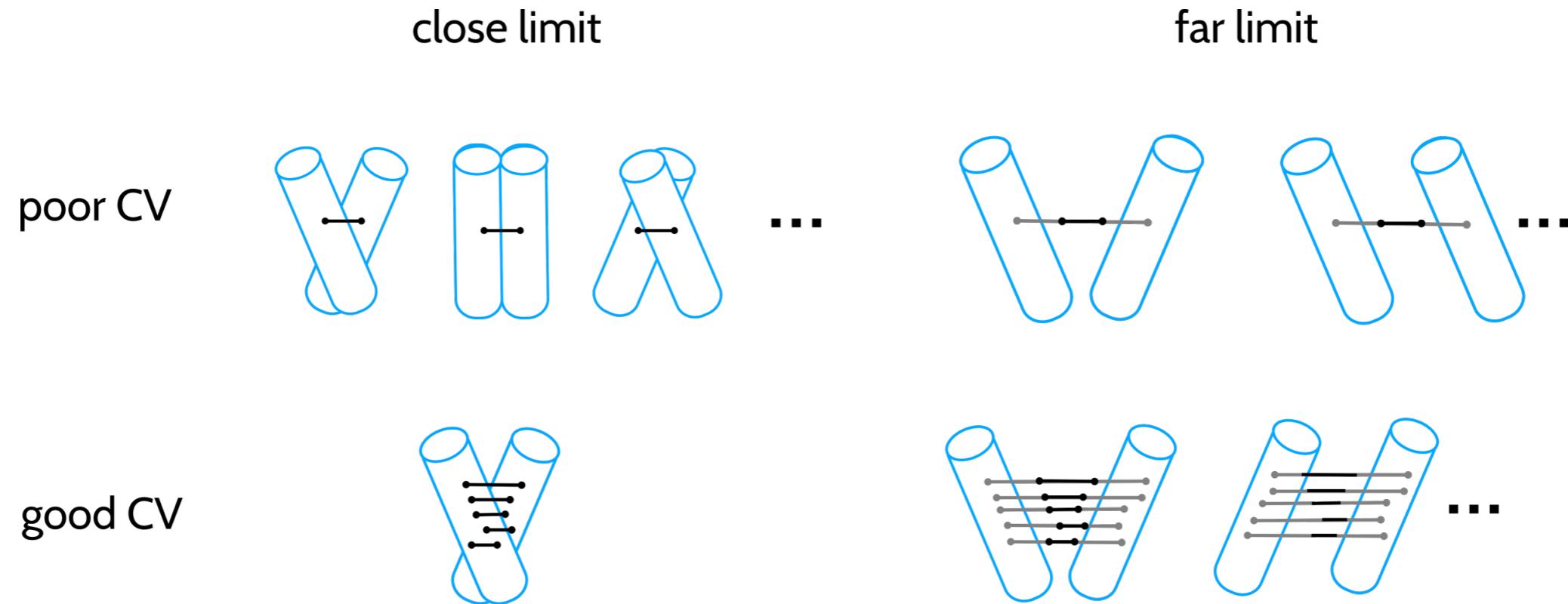
Membrane protein folding

- Similar to soluble protein folding
- Membrane is more viscous and slower to diffuse in than water (challenging, use coarse-grained MARTINI force-field)
- Glycophorin A is a model system



Collective variable choice

- Helix-helix center of mass distance? simple but degenerate when helices are close
- Native contacts? could still be dimer and not have any native contacts

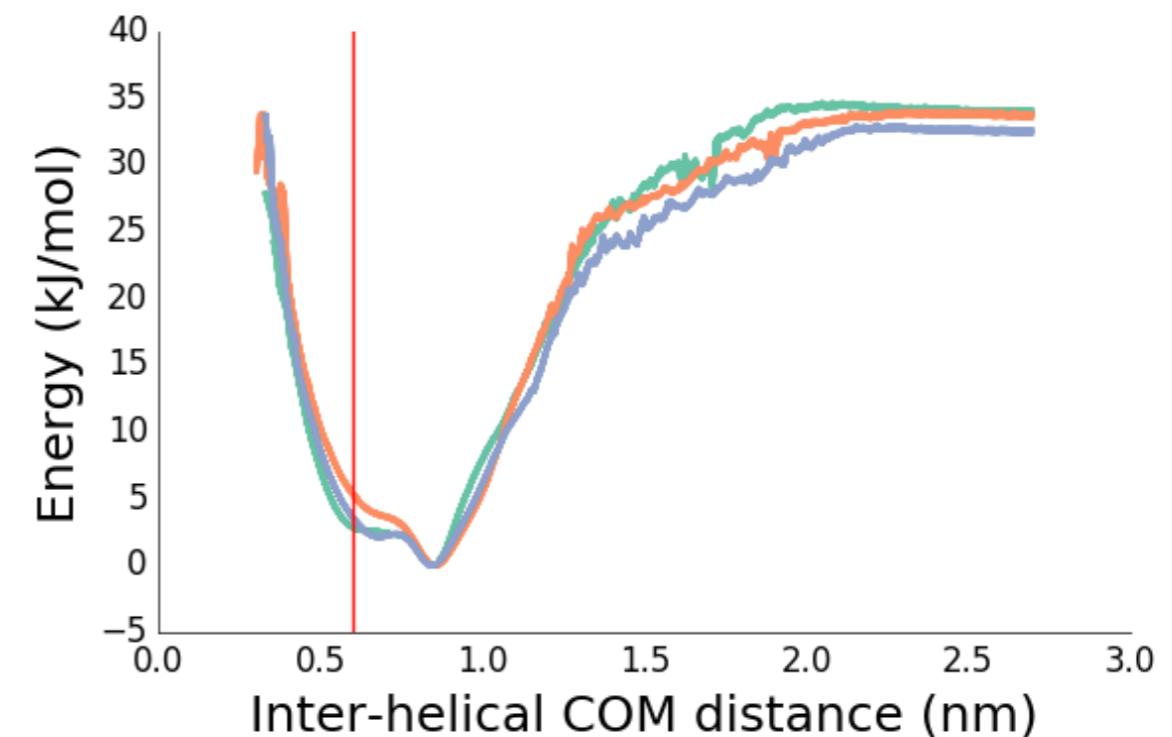
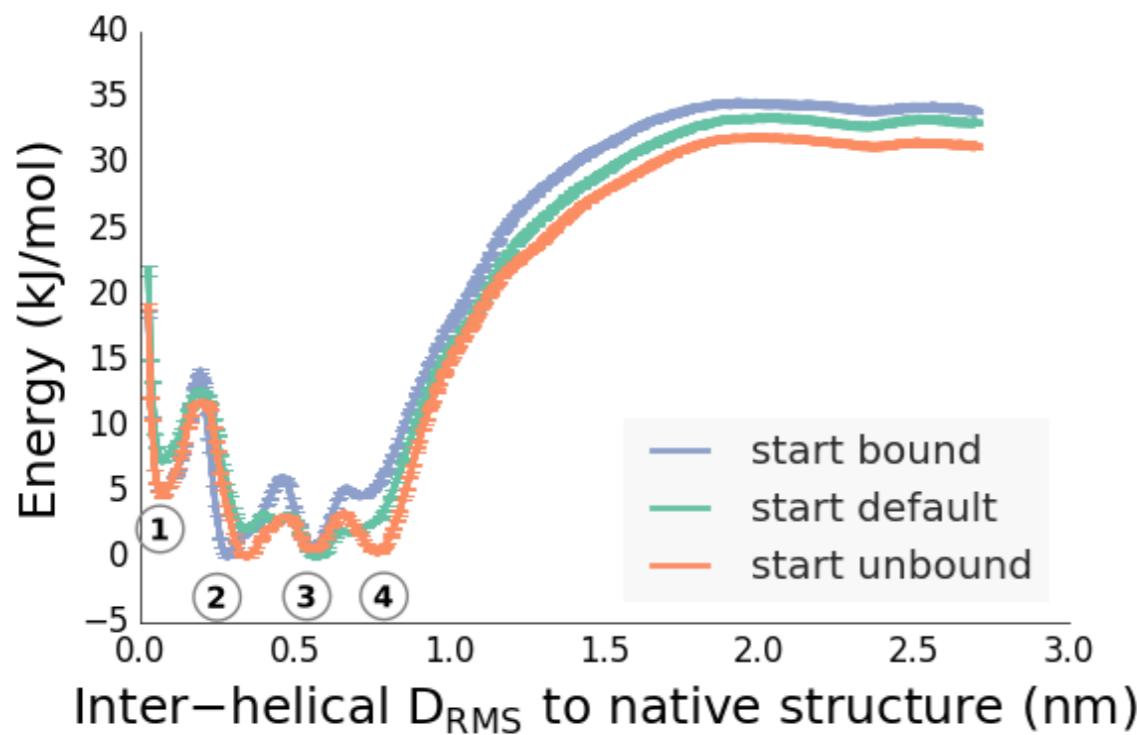


Collective variable choice

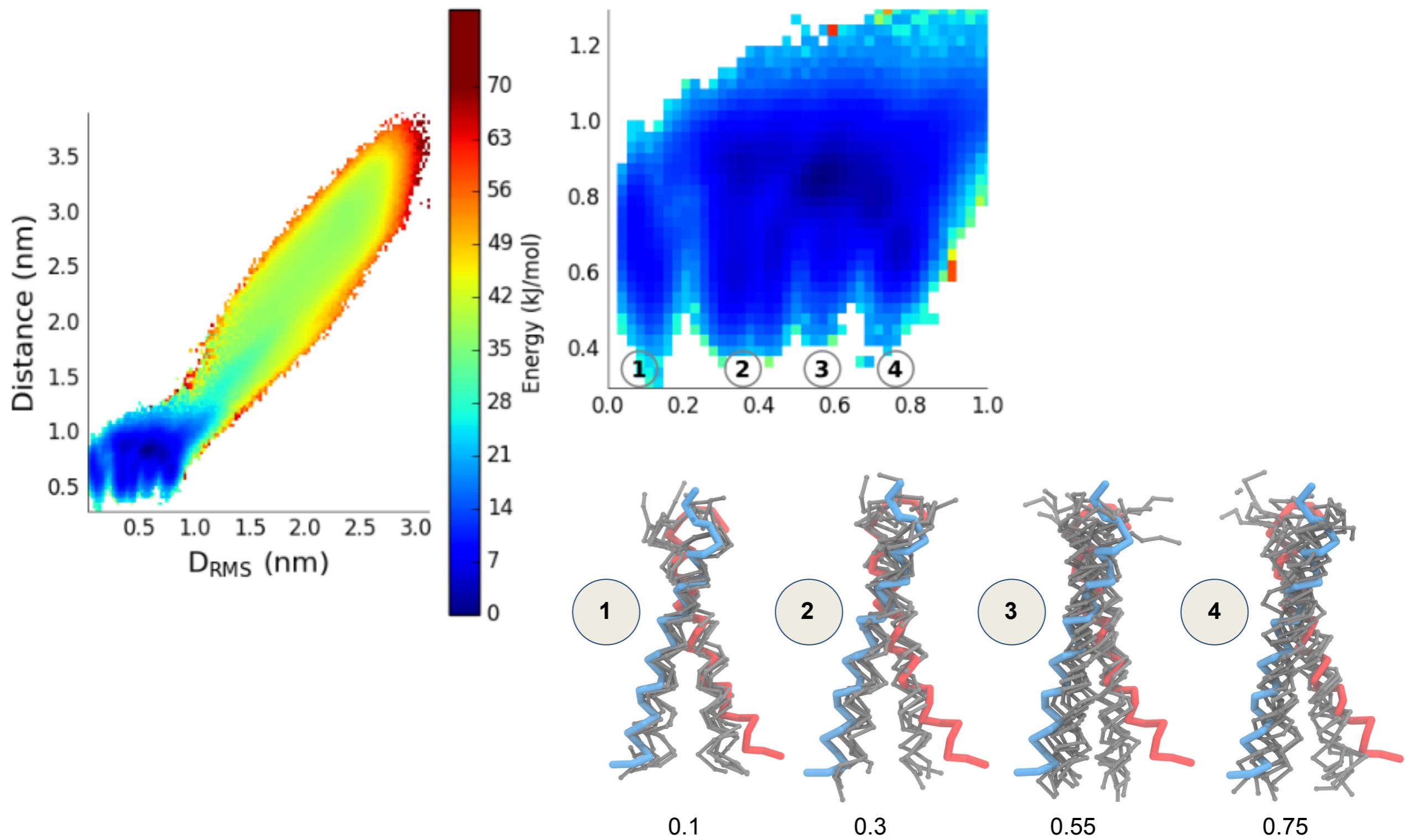
```
drms: INTERDRMSD REFERENCE=gpa_together_noH_nos_long.pdb LOWER_CUTOFF=0.1 UPPER_CUTOFF=0.6
```

GpA dimerization PMF

- Via replica exchange umbrella sampling – along native inter-helical DRMSD
- 16 windows, linearly spaced 0.0–2.5 nm

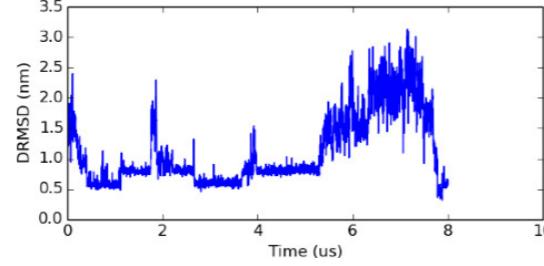
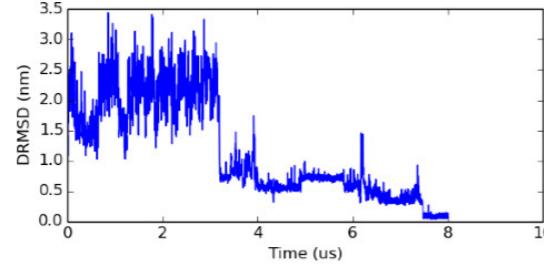
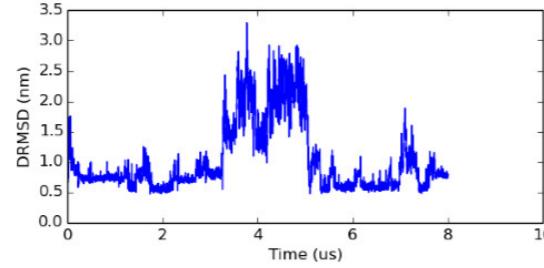
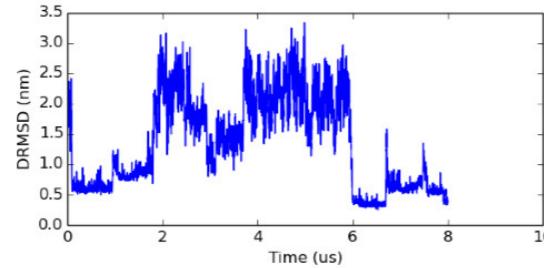
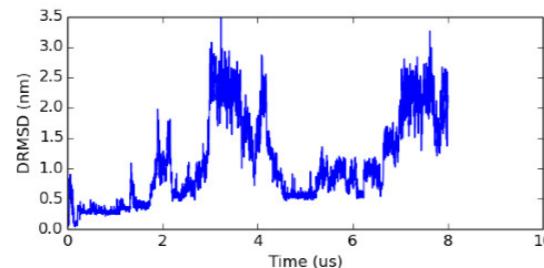


GpA dimerization – a more detailed look

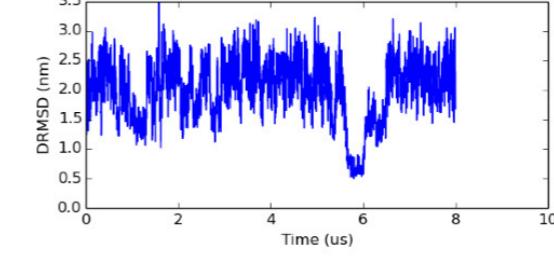
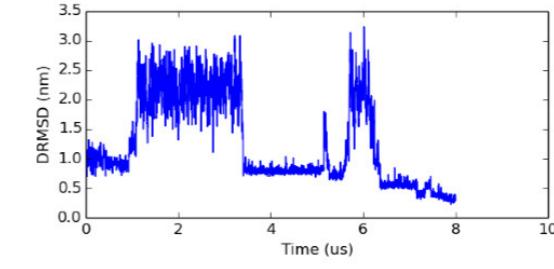
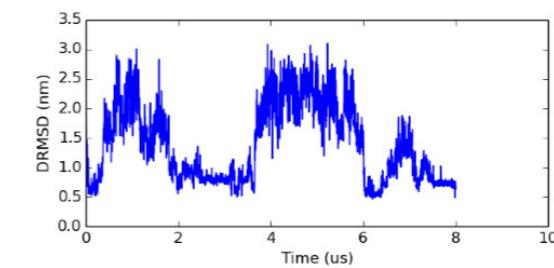
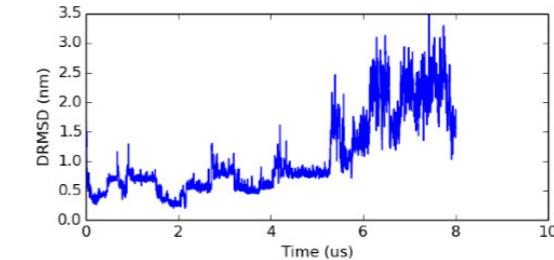
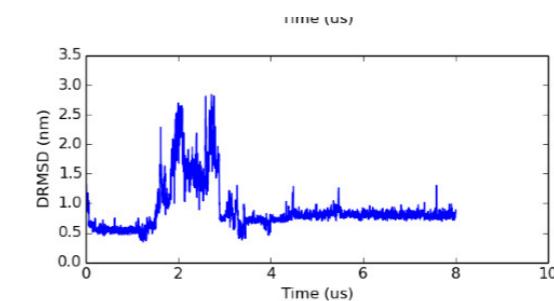


Diffusive over the CV?

Bound



Unbound



Convergence and Sampling in Determining Free Energy Landscapes for Membrane Protein Association

Jan Domański^{†‡}, George Hedger[†], Robert B. Best[‡], Phillip J. Stansfeld[†], and Mark S. P. Sansom^{*†} 

[†] Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, U.K.

[‡] Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892-0520, United States

J. Phys. Chem. B, 2017, 121 (15), pp 3364–3375

DOI: 10.1021/acs.jpcb.6b08445

Publication Date (Web): November 3, 2016

Copyright © 2016 American Chemical Society

 bestlab / SamplingPaper

 Unwatch ▾

4

 Star

0

 Fork

0

 Code

 Issues 0

 Pull requests 0

 Projects 0

 Wiki

 Settings

 Insights ▾

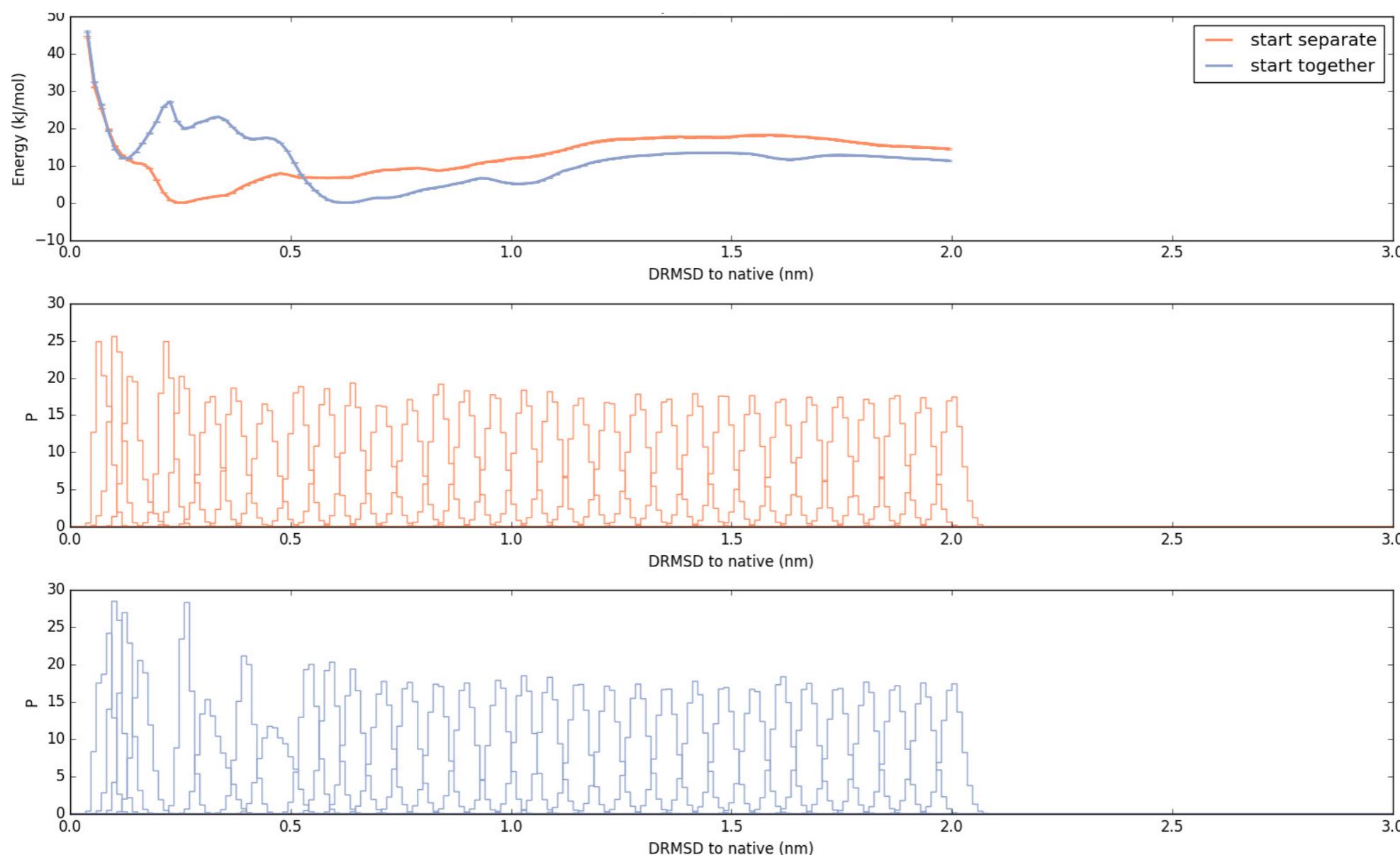
Setup files and scripts to reproduce the sampling paper

 Edit

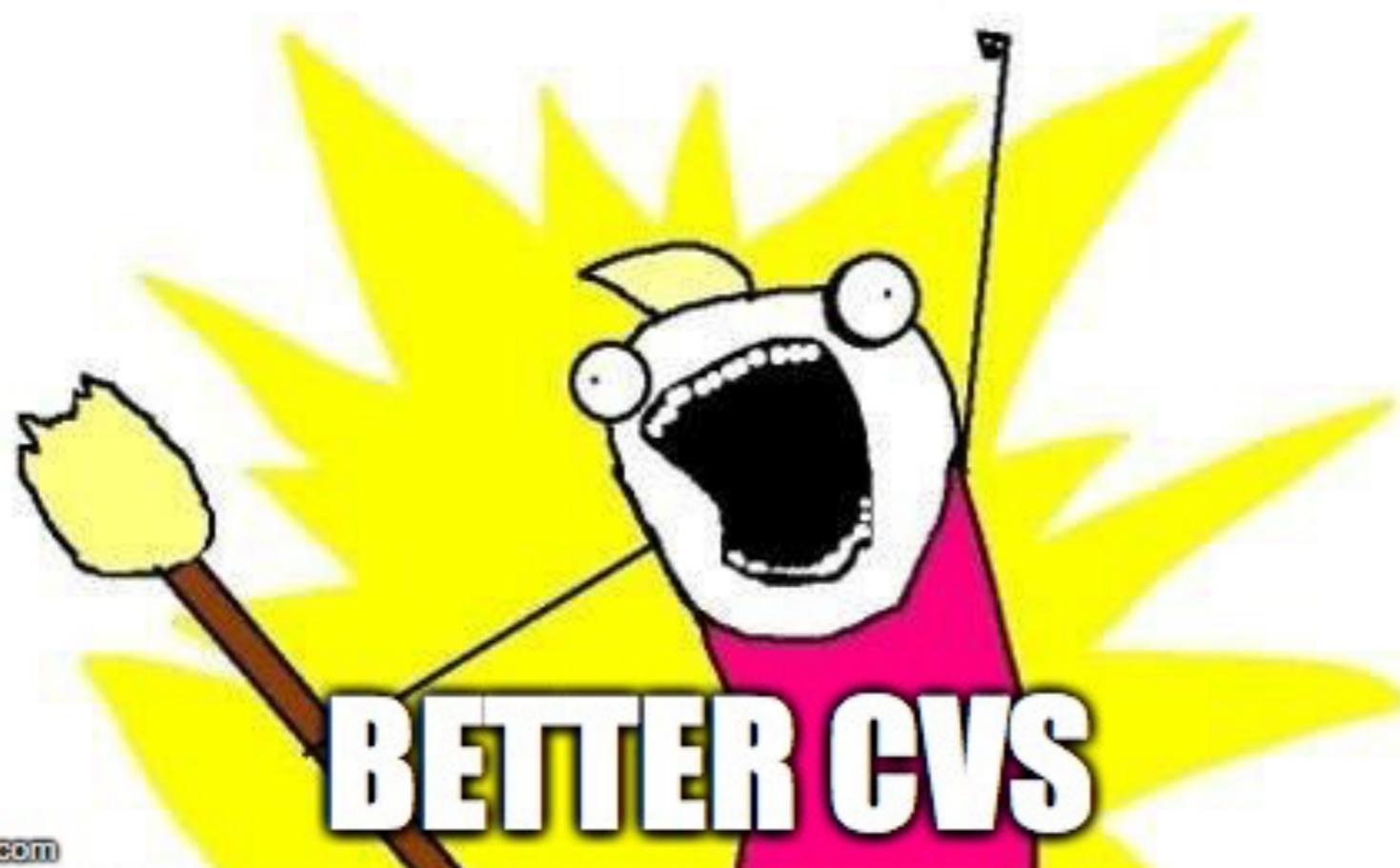
Add topics

What if it goes wrong?

- Different force-field, shows a highly non-converged 1D PMF

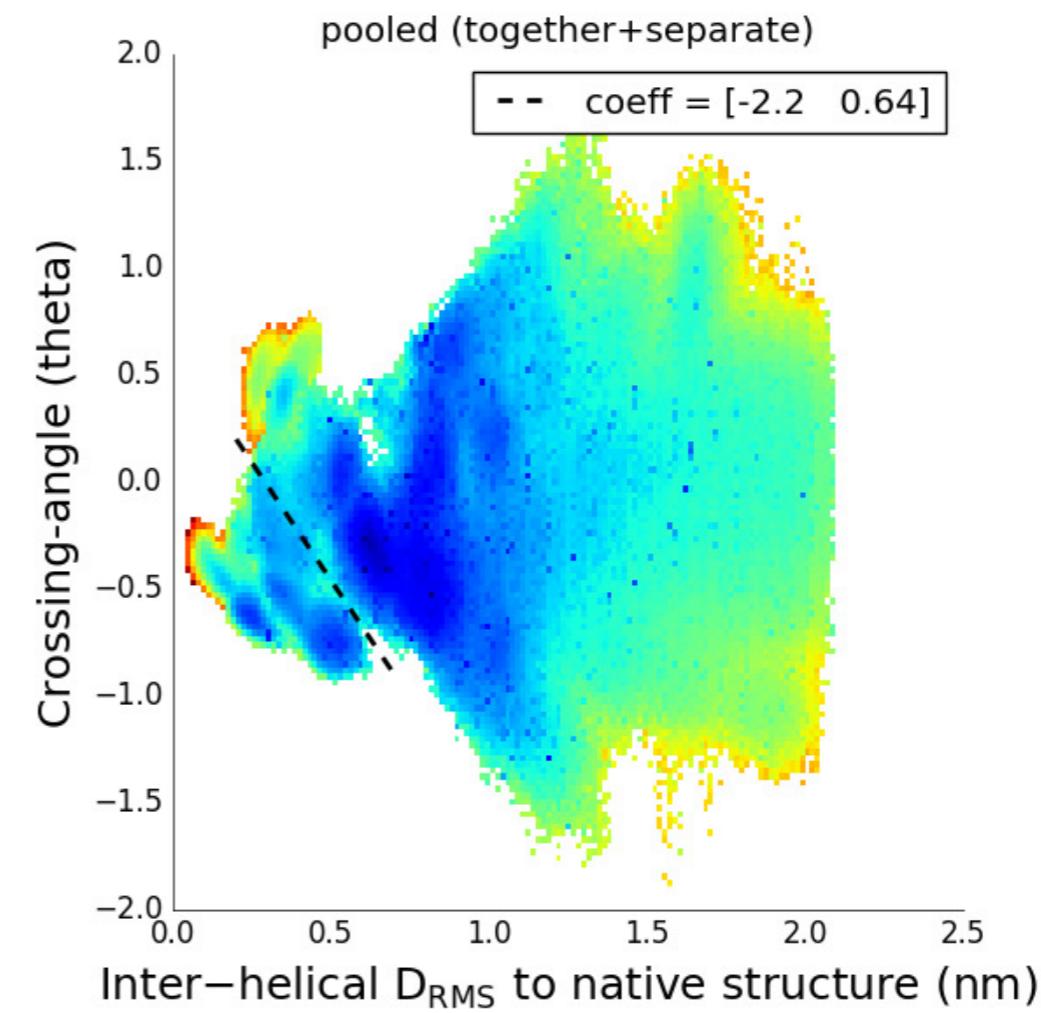
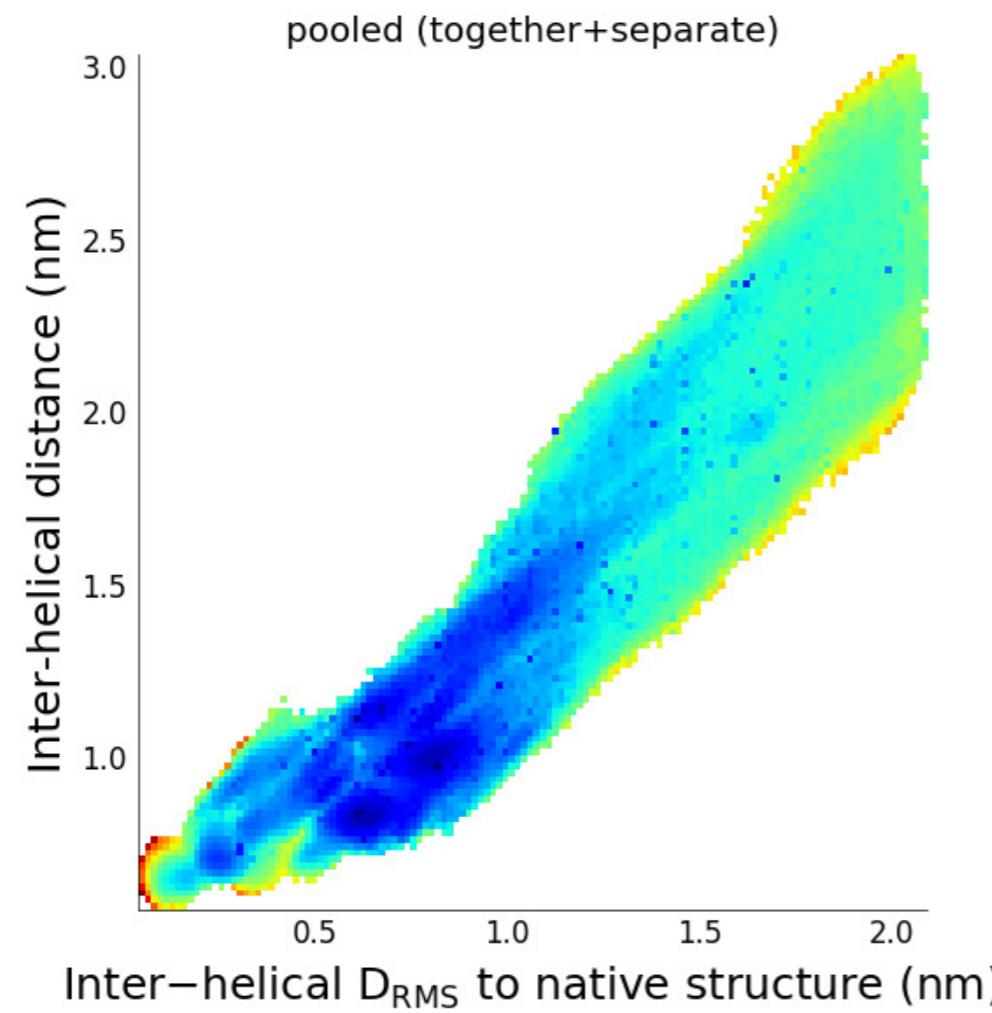


WHAT DO WE WANT?



Diagnosing problems

- Let's look closer at 2D PMF projections on different CVs
- Histogram across all COLVAR, weight related the US window energy (from WHAM) + restraint.bias (from PLUMED)



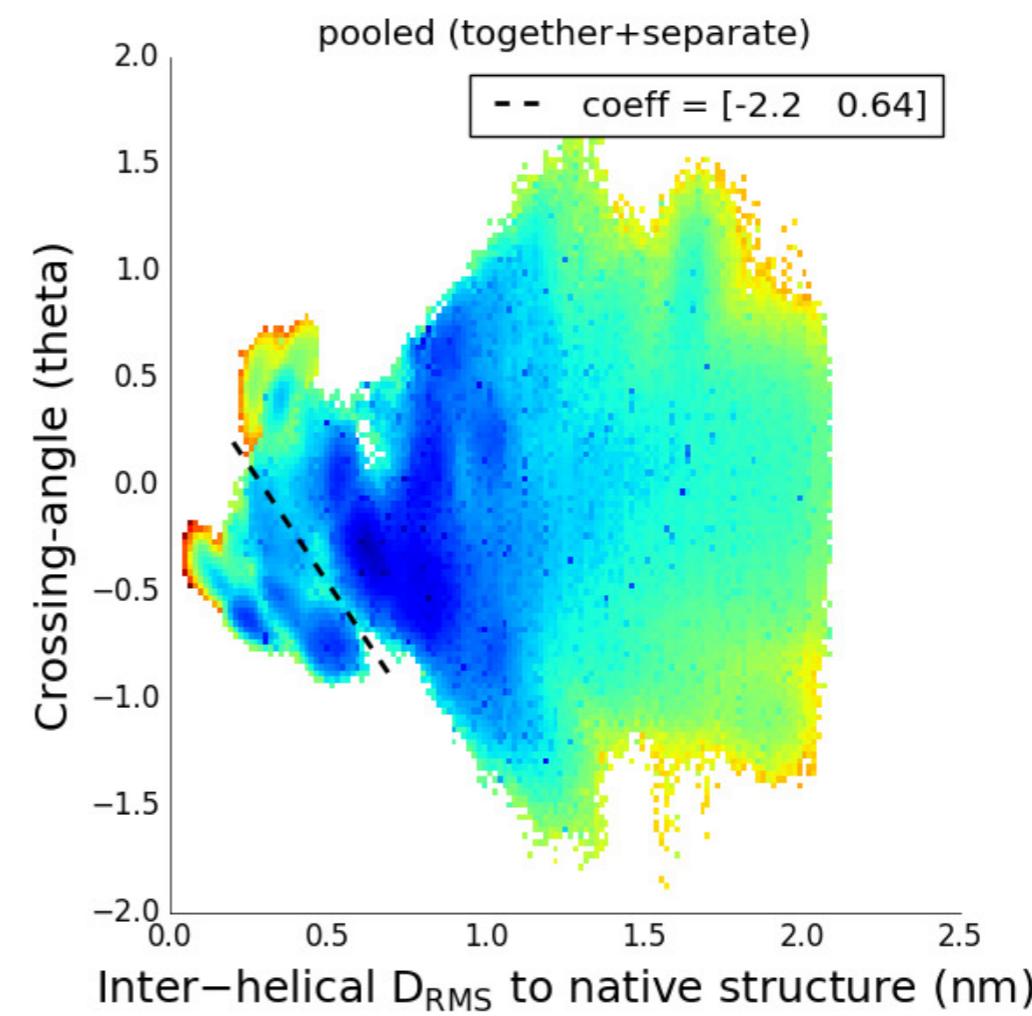
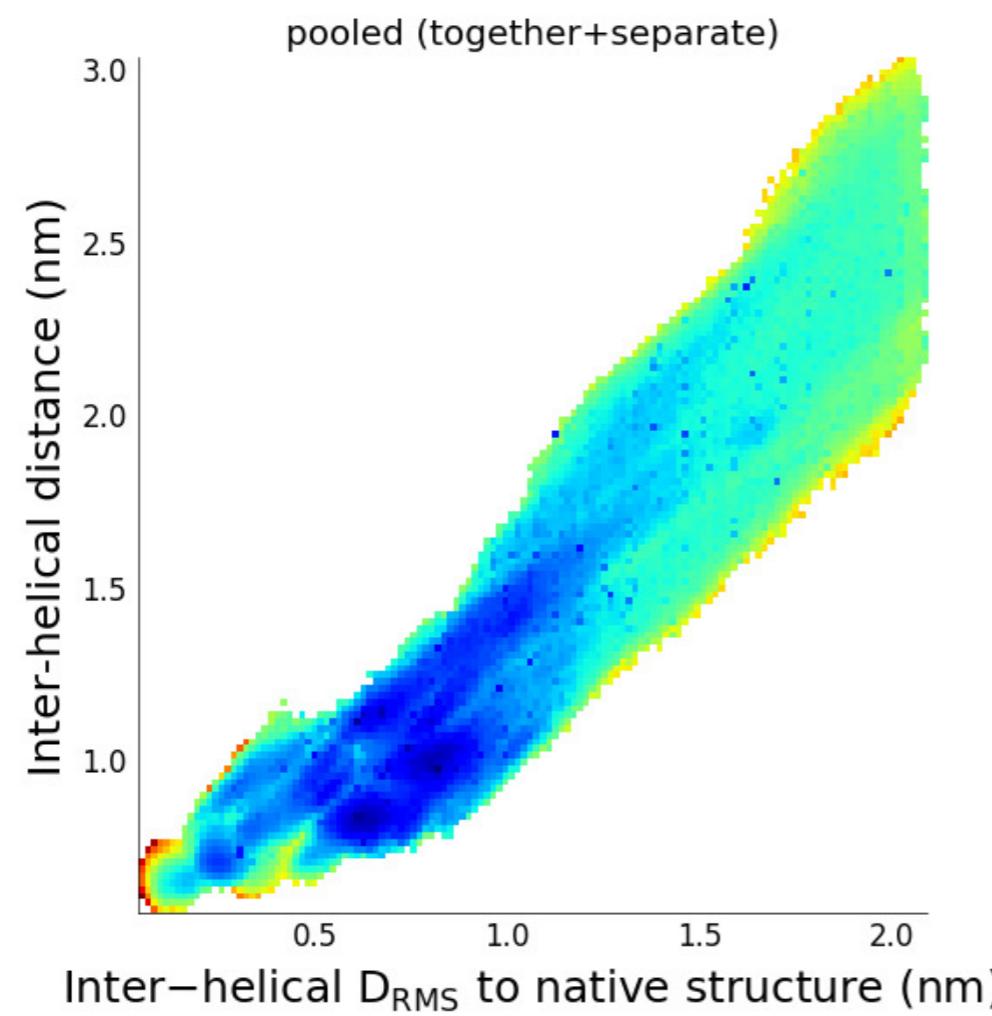
Diagnosing problems

- Clearly DRMS is not a perfect CV!
 - Better than distance
 - Degenerate with respect to theta, helix-helix crossing angle
- Let's build a new CV called "X" where $X = \text{DRMS} + \lambda * \text{THETA}$
- 'lambda' is some mixing content from the slope of the curve dividing bound and unbound basins

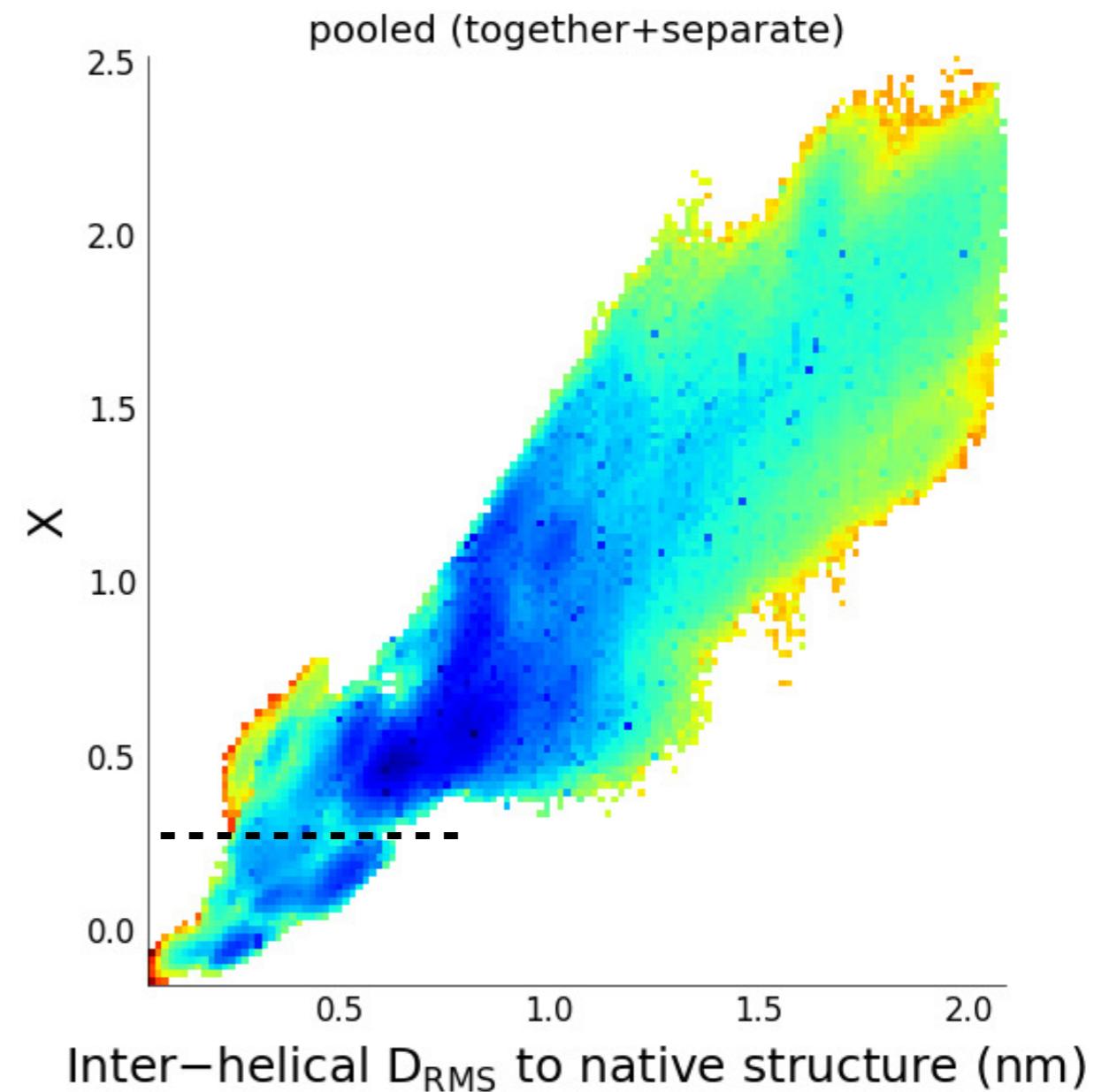
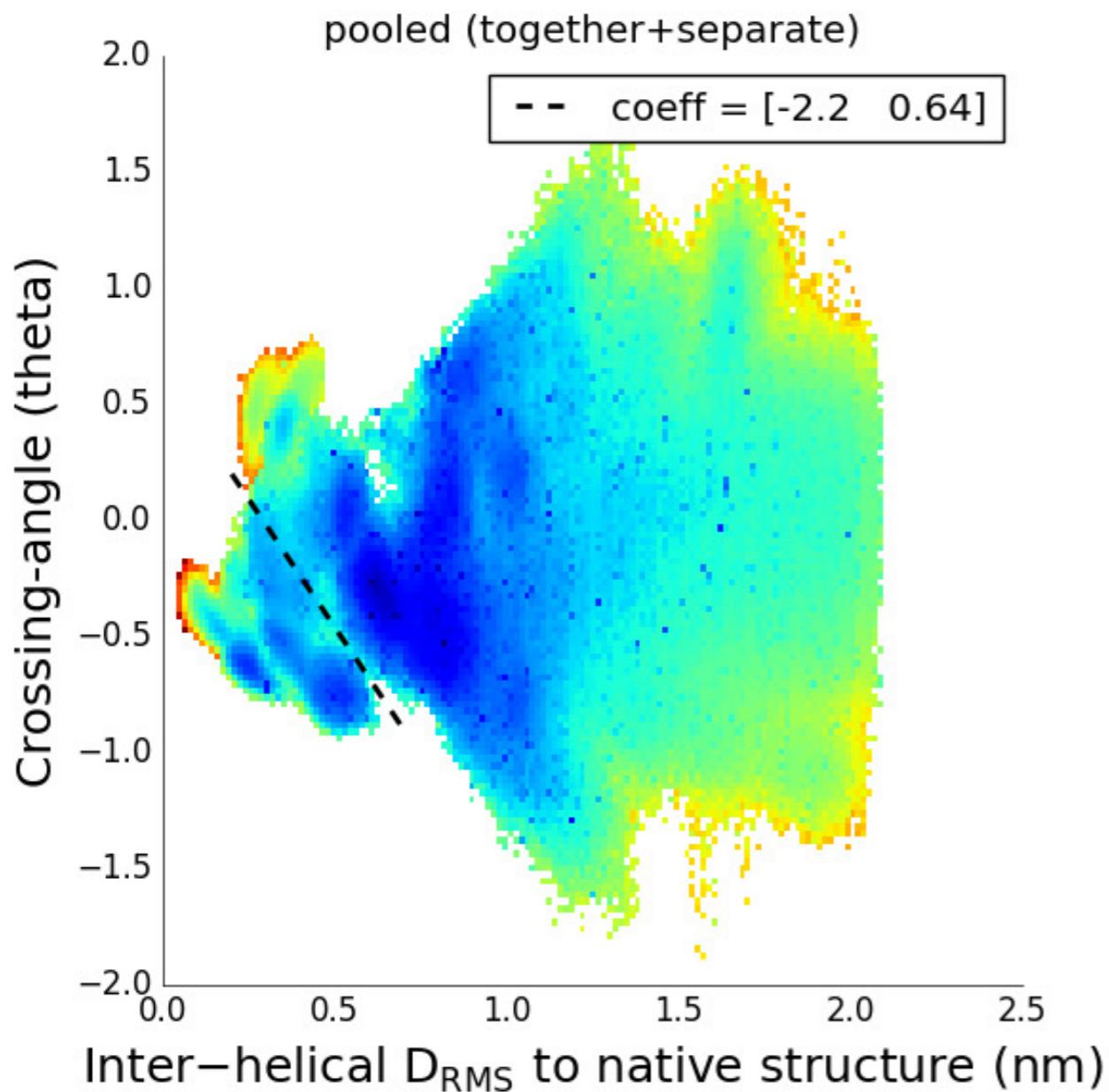
```
combine: COMBINE ARG=drms,t COEFFICIENTS=1,0.5 POWERS=1,1 PERIODIC=N0
```

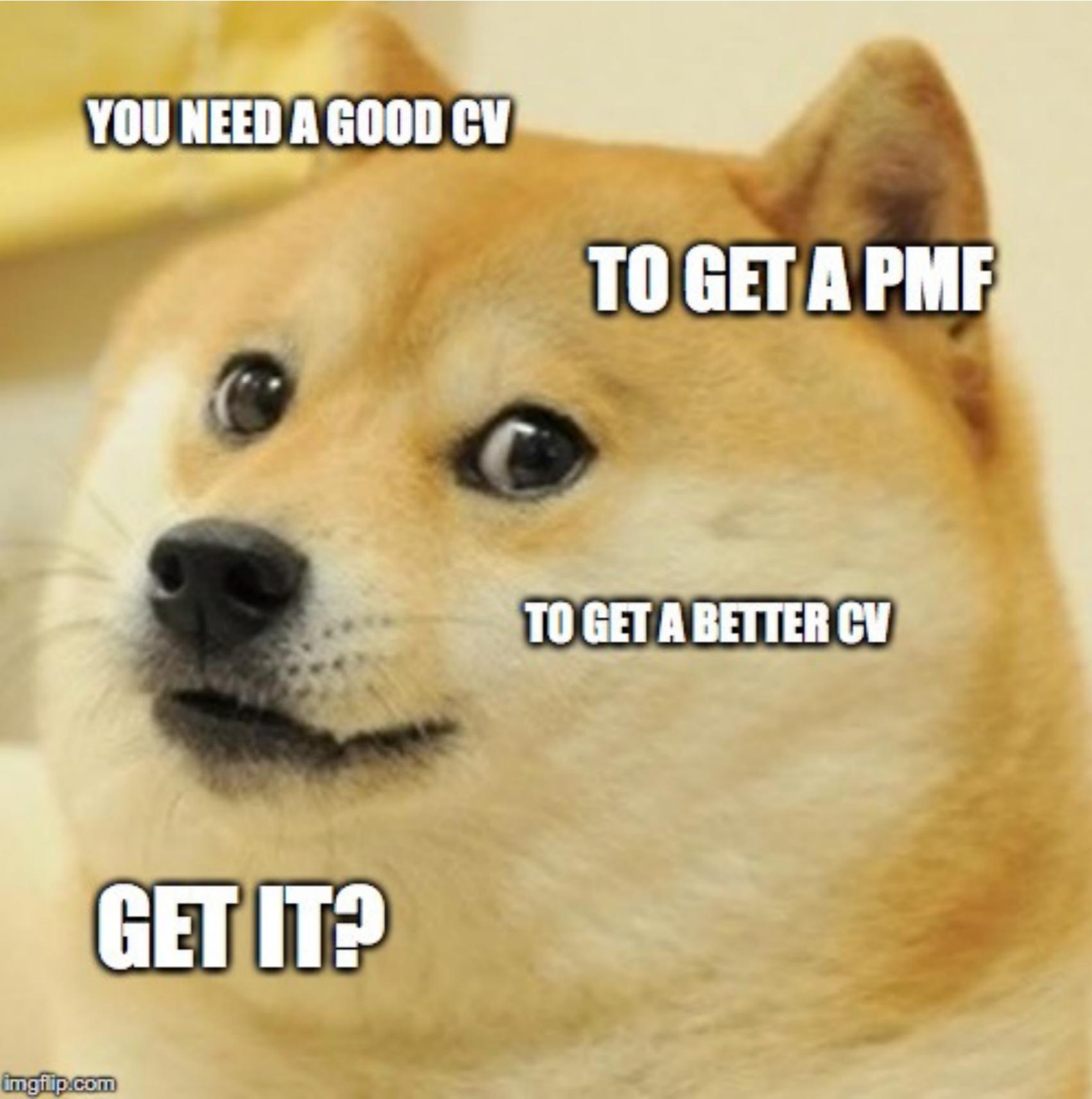
Diagnosing problems

- Let's look closer at 2D PMF projections on different CVs
- Histogram across all COLVAR, weight is the US window energy (from WHAM) + restraint.bias from PLUMED



Choosing the mixing constant ‘lambda’





YOU NEED A GOOD CV

TO GET A PMF

TO GET A BETTER CV

GET IT?

Challenges

- Even for this simple system, finding a good 1D CV is a tough (think multi-span protein like a transporter!)
- Projections/mixing CV can introduce new degenerate regions
 - How can we move forward? Metadynamics, 2D umbrella sampling?

ANY OTHER GOOD NEWS

I HATE YOU

Take aways

- Principles for assessing convergence (independent starting point, diffusivity)
- More meaningful benchmarking standards (we need them!)
 - Ala-dipeptide is not the answer can do brute-force now
- Even simple systems can have complex energy landscapes, hard to find a non-degenerate 1D CV

Acknowledgements



Robert
Best



Mark
Sansom



Phil
Stansfeld

- Best Lab Members:
 - Wenwei Zheng
 - Pengfei Tian
 - Mathias Bellaiche and Travis Hoppe
- Giovanni Bussi
- Rich Pastor and Alex Sodt
- David Dotson and Oliver Beckstein
- Wonpil Im
- Jettain Mittal
- Thomas Piggot
- Biowulf2 supercomputing center



National Institute of
Diabetes and Digestive
and Kidney Diseases

wellcome trust