Protein Model **Refinement** via **Iterative** Molecular Dynamics simulations

Lim Heo and Michael Feig

Department of Biochemistry and Molecular Biology
Michigan State University

MD simulation and refinement

Numerically solving
Newtonian equation!

$$F = ma = -\nabla V(r)$$

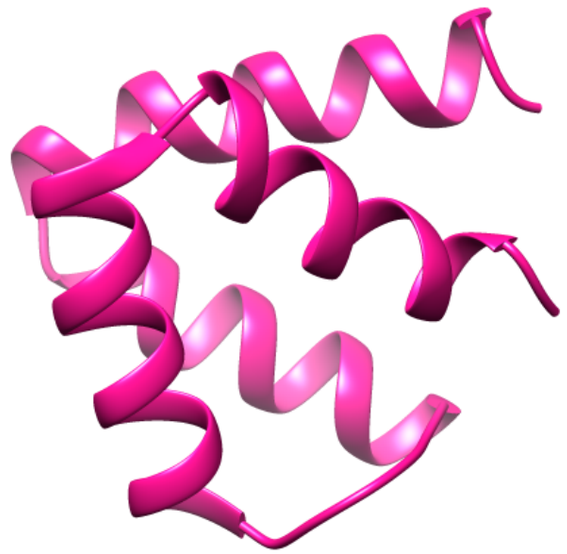
Physical atomistic force field (CHARMM 36m)

MD simulation and refinement

Numerically solving Newtonian equation!

$$F = ma = -\nabla V(r)$$

Physical atomistic force field (CHARMM 36m)

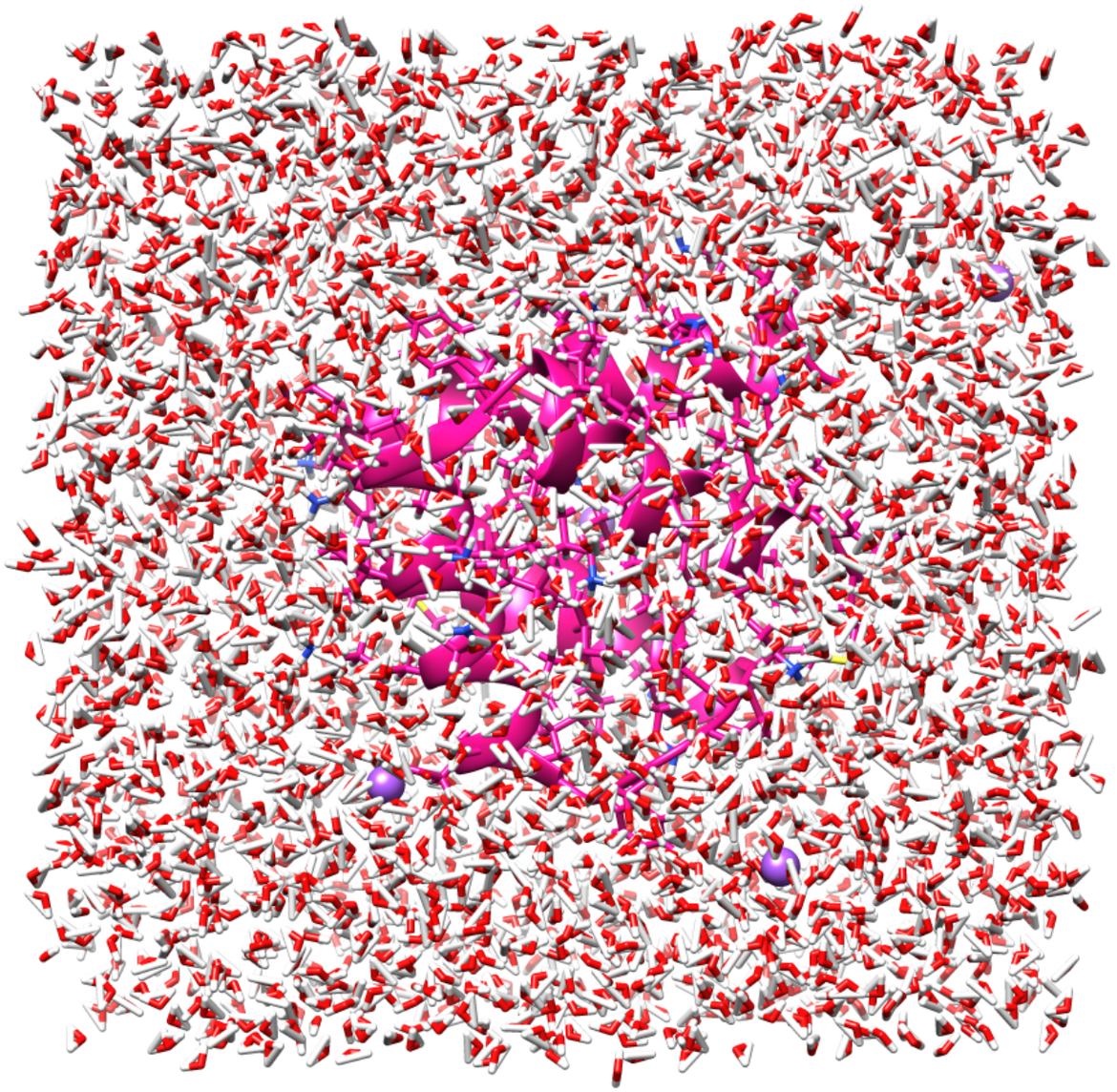


MD simulation and refinement

Numerically solving Newtonian equation!

$$F = ma = -\nabla V(r)$$

Physical atomistic force field (CHARMM 36m)



MD simulation and refinement

Numerically solving Newtonian equation!

$$F = ma = -\nabla V(r)$$

Physical atomistic force field (CHARMM 36m)

MD simulation and refinement

Numerically solving Newtonian equation!

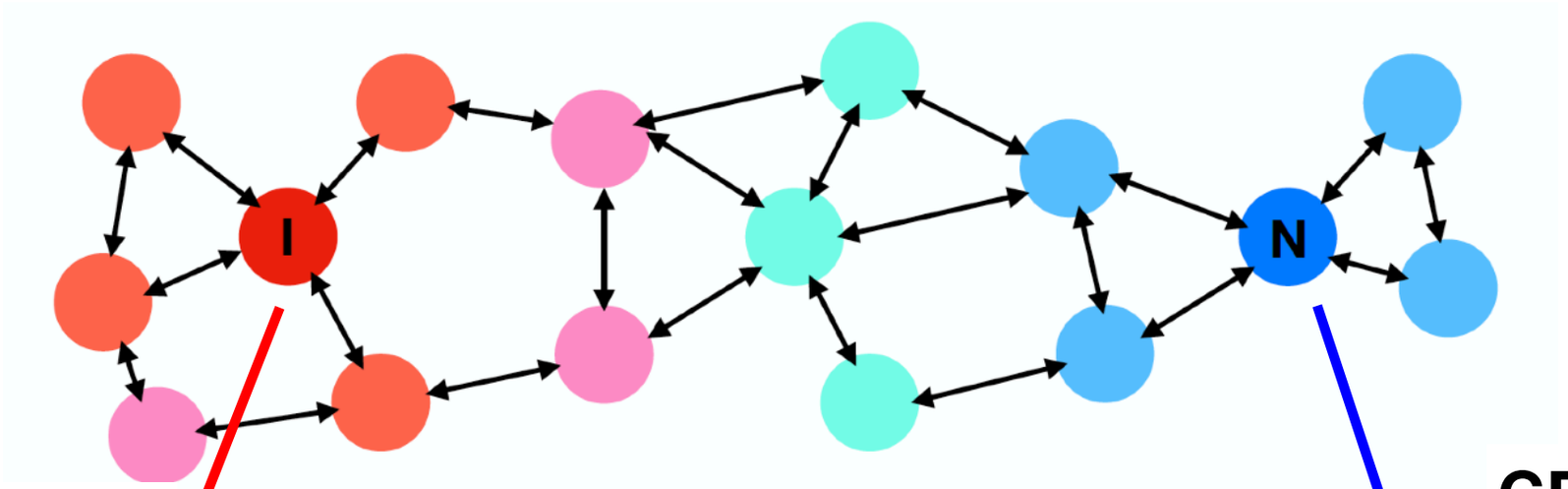
$$F = ma = -\nabla V(r)$$

Physical atomistic force field (CHARMM 36m)

Pre-study

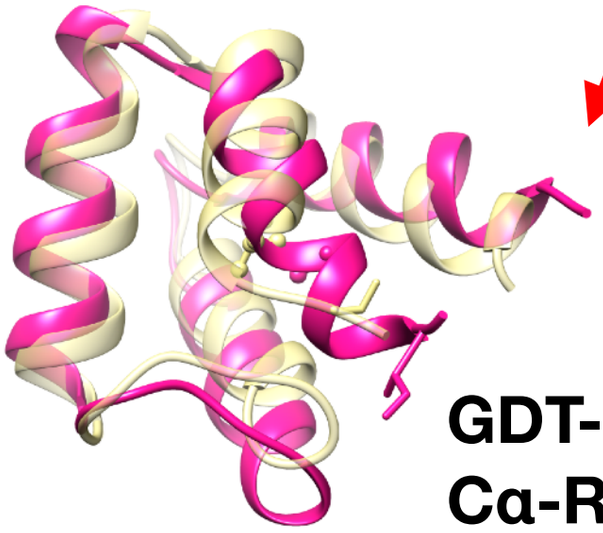
Markov state modeling for refinement

10s μs MD simulations/target with c36m
to identify "**refinement pathway**"

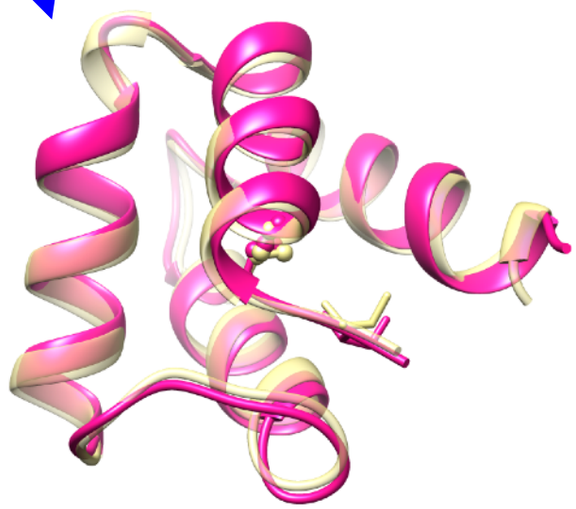


TR816

GDT-HA = 87
Cα-RMSD = 0.8 Å



GDT-HA = 51
Cα-RMSD = 2.5 Å



Lim Heo and Michael Feig, *PNAS* (**2018**)
https://www.pnas.org/content/early/2018/12/07/1811364115.

Pre-study

Native states have the lowest free energies

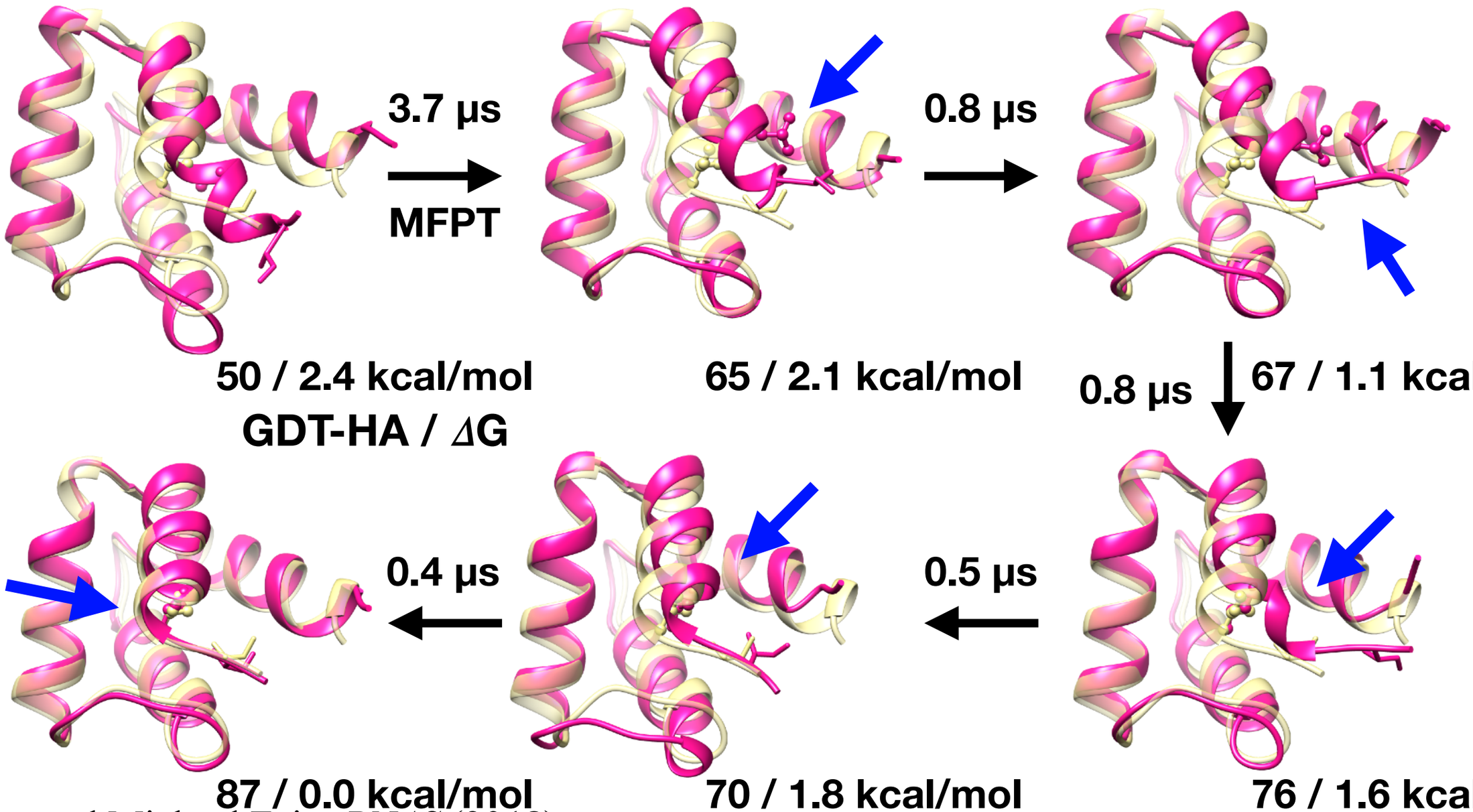
Target	Initial model RMSD [Å]	Native state RMSD [Å]	$\Delta\Delta G$ [kcal/mol]	MFPT [μs]	# of transitions
TR816	2.53	**0.80**	**-2.46**	1.7	5
TR837	2.95	**0.88**	**-2.60**	43.4	7
TR854	2.27	1.04	**-1.59**	1.5	3
TR782	1.93	**0.94**	**-0.65**	39.6	5
TR872	5.59	1.97	**-0.83**	2.9	2
TR921	3.51	**0.90**	**-0.85**	637.9	15
TR769	1.74	1.14	**-1.10**	0.8	2
TR894	2.23	**0.85**	**-1.71**	6.0	5

Lim Heo and Michael Feig, *PNAS* (**2018**)
https://www.pnas.org/content/early/2018/12/07/1811364115.

Pre-study

Partial unfolding and refolding is required

It has to be **partially unfolded and refolded** to get to the native.



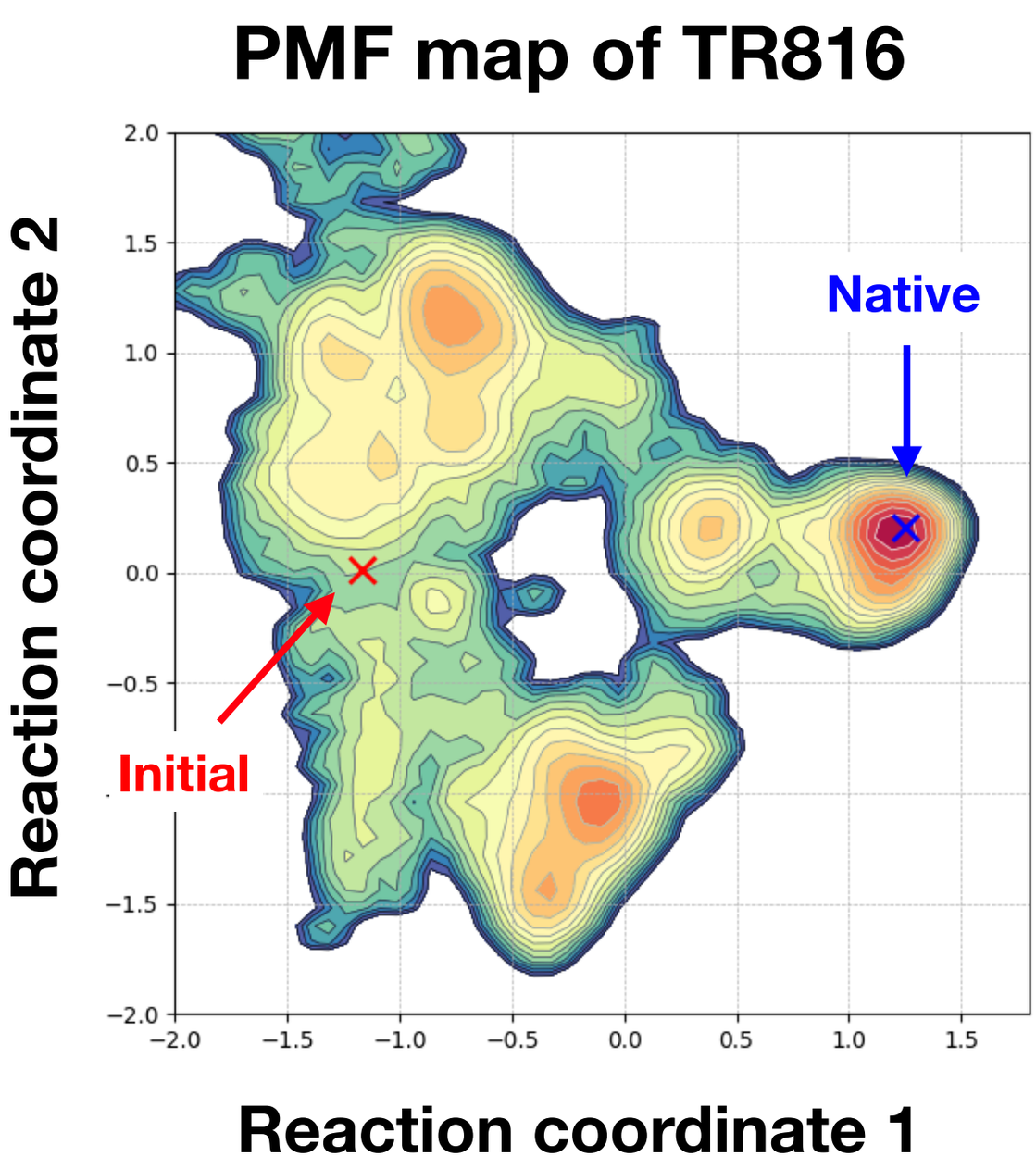
Lim Heo and Michael Feig, *PNAS* (**2018**)
https://www.pnas.org/content/early/2018/12/07/1811364115.

FEIGLAB group CASP13 refinement protocol

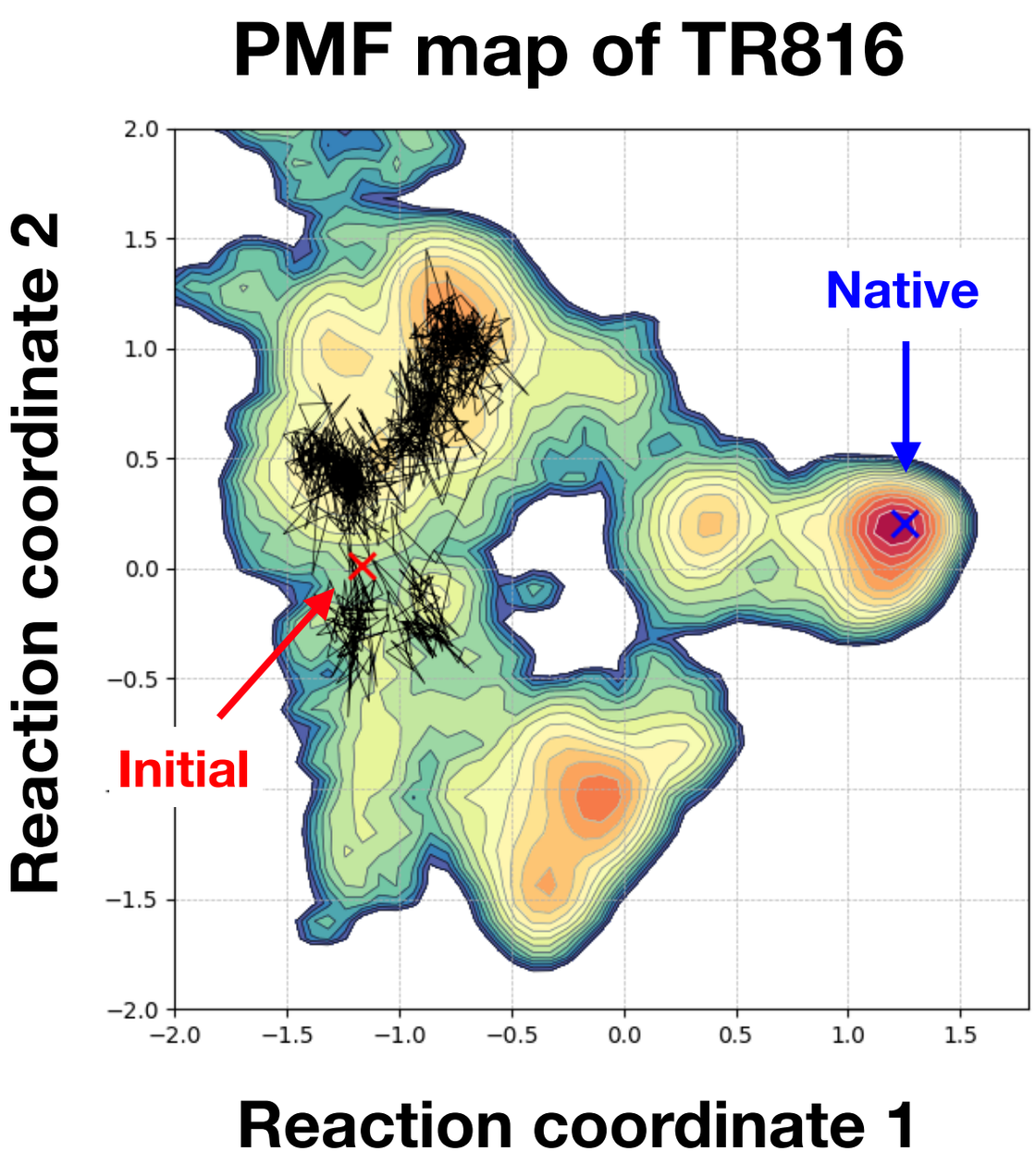
- Main protocol (20 / 31 targets)
  - **Iterative** runs of **MD** simulations
    (2 μs/target; 94 GPU hrs on GTX1080Ti for R0949 (129 aa))
  - **Flat-bottom** harmonic restraints
  - **c36m** with modifications to accelerate sampling
  - Scoring with **Rosetta,** structure **averaging**, and **locPREFMD**

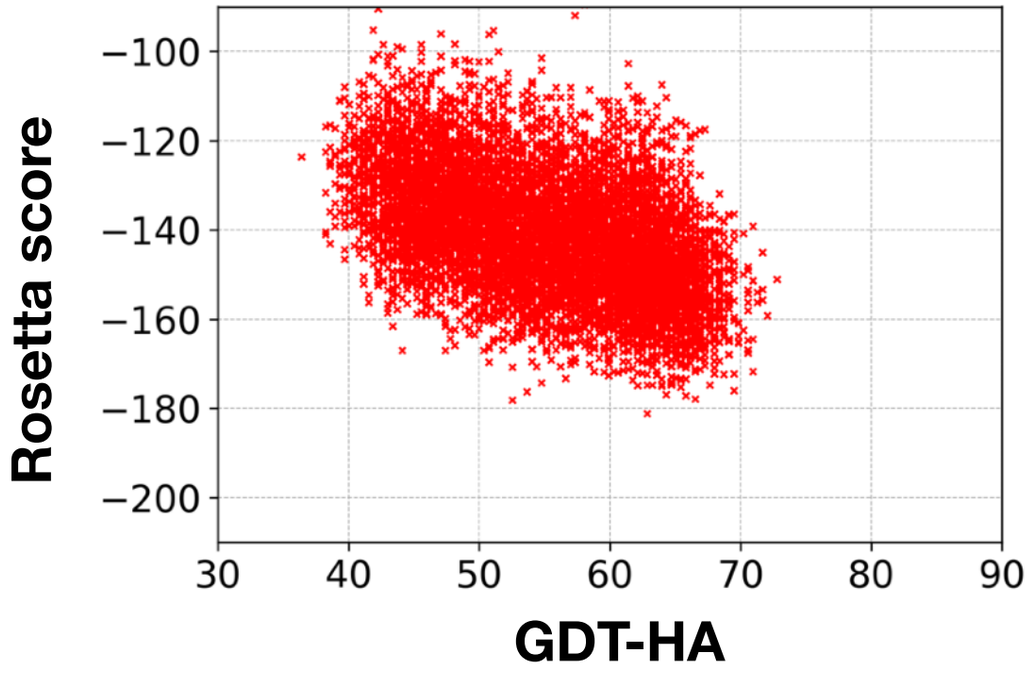
- Conservative protocol (11 / 31 targets)
  - Analogous to **our previous protocol** with reduced simulation time
    (250 ns; 23 GPU hrs on GTX1080Ti for R0949 (129 aa))

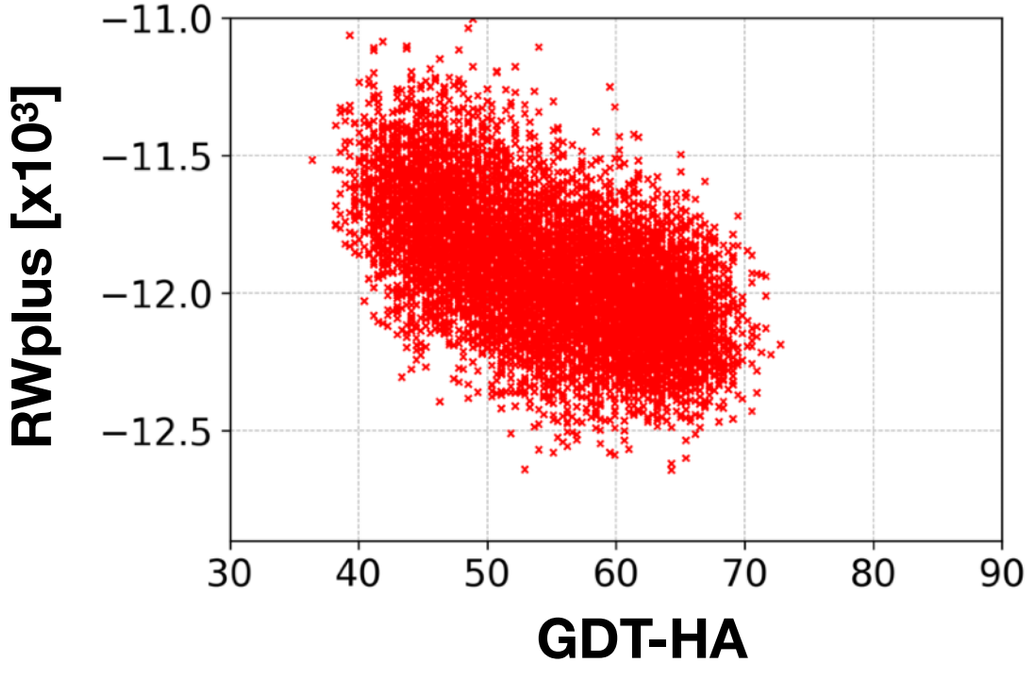
Iterative MD sampling and scoring



Iterative MD sampling and scoring

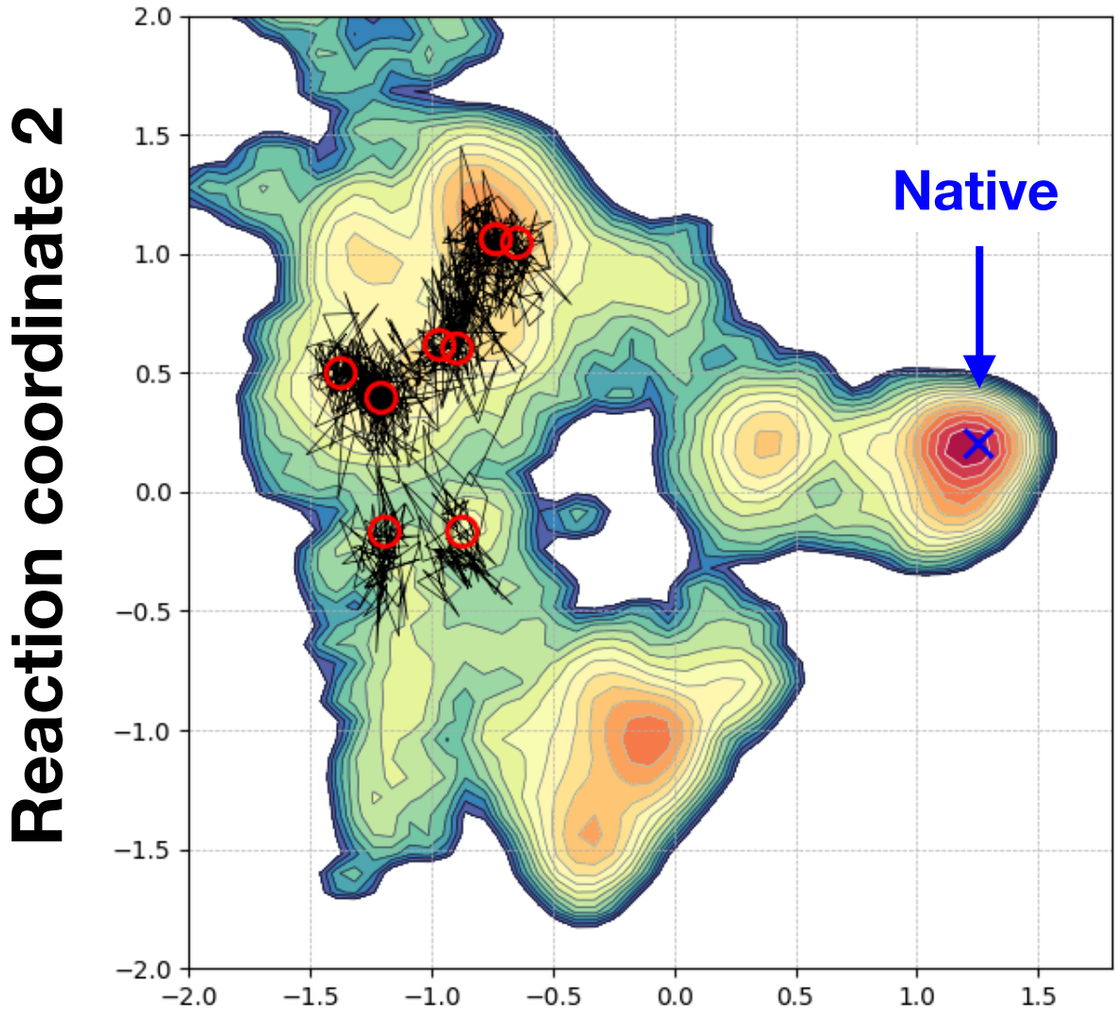


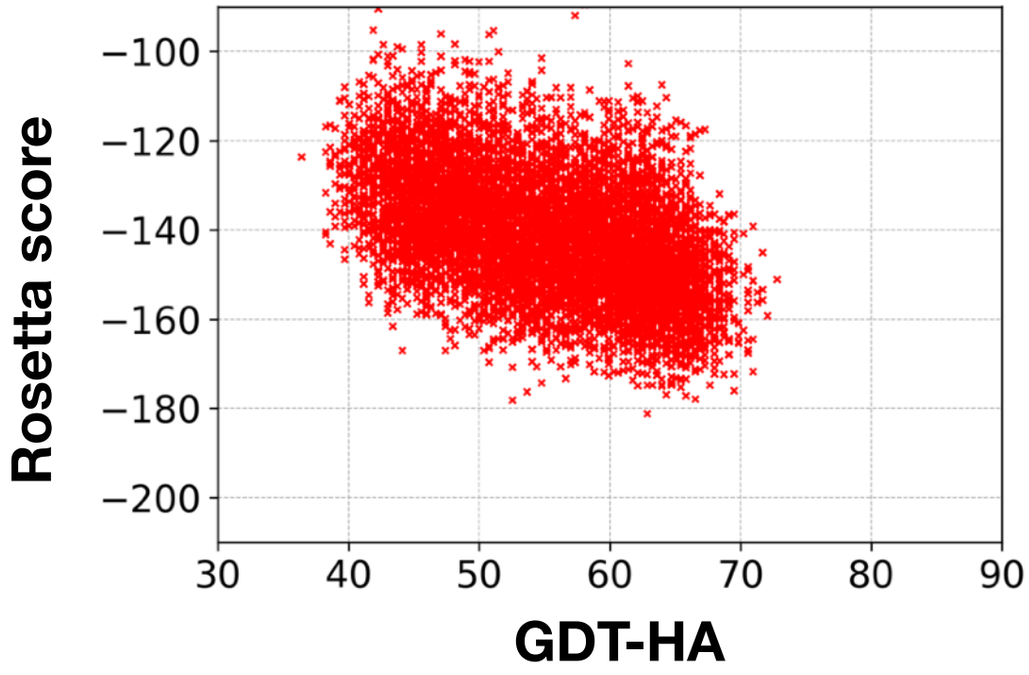


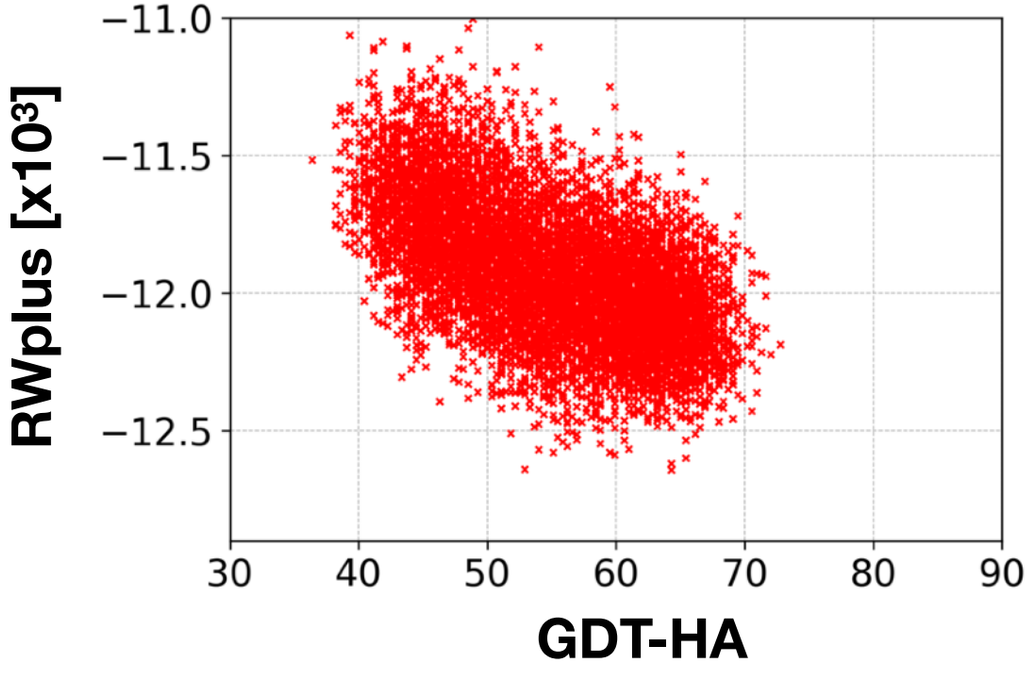


Iterative MD sampling and scoring

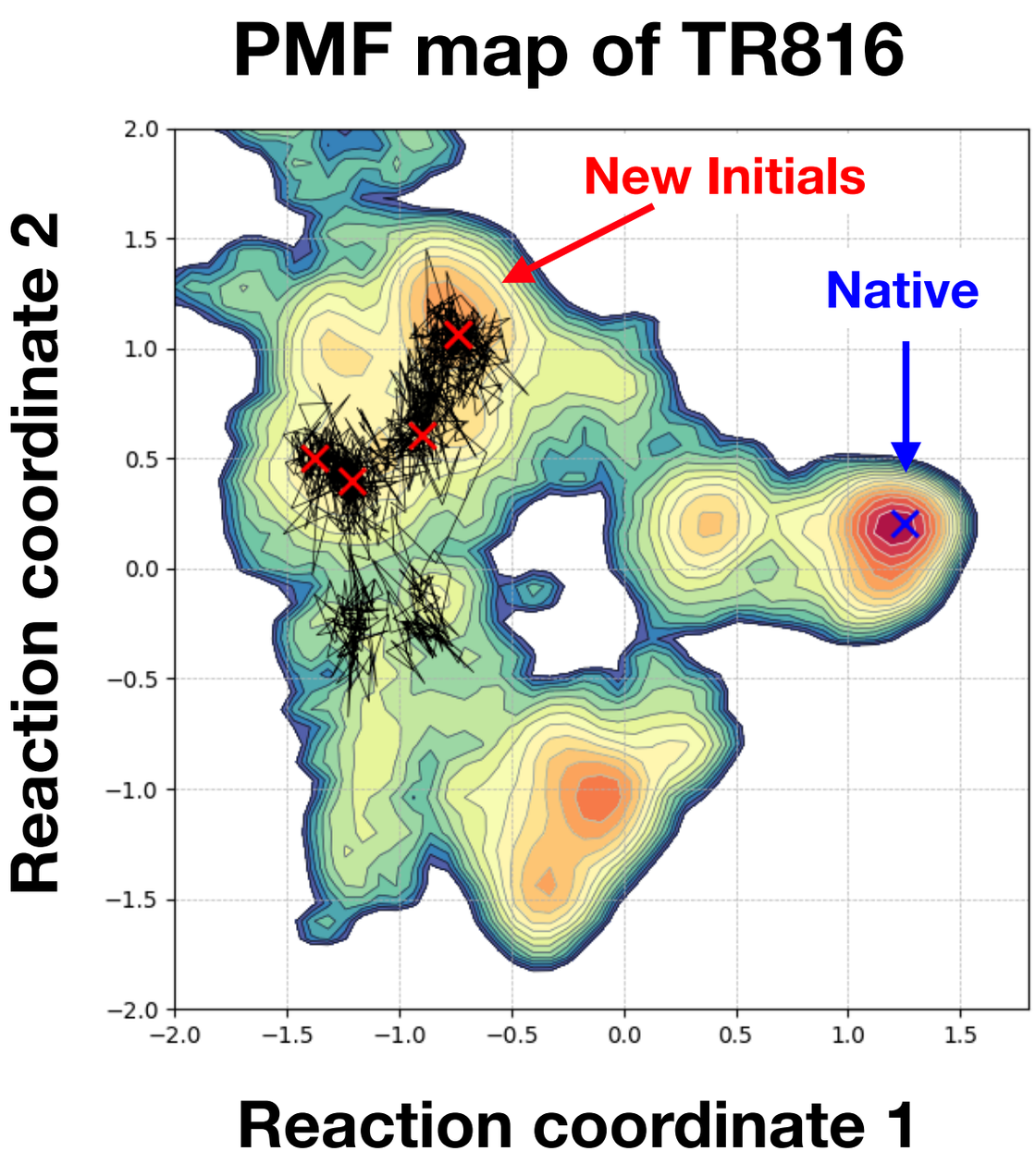
PMF map of TR816

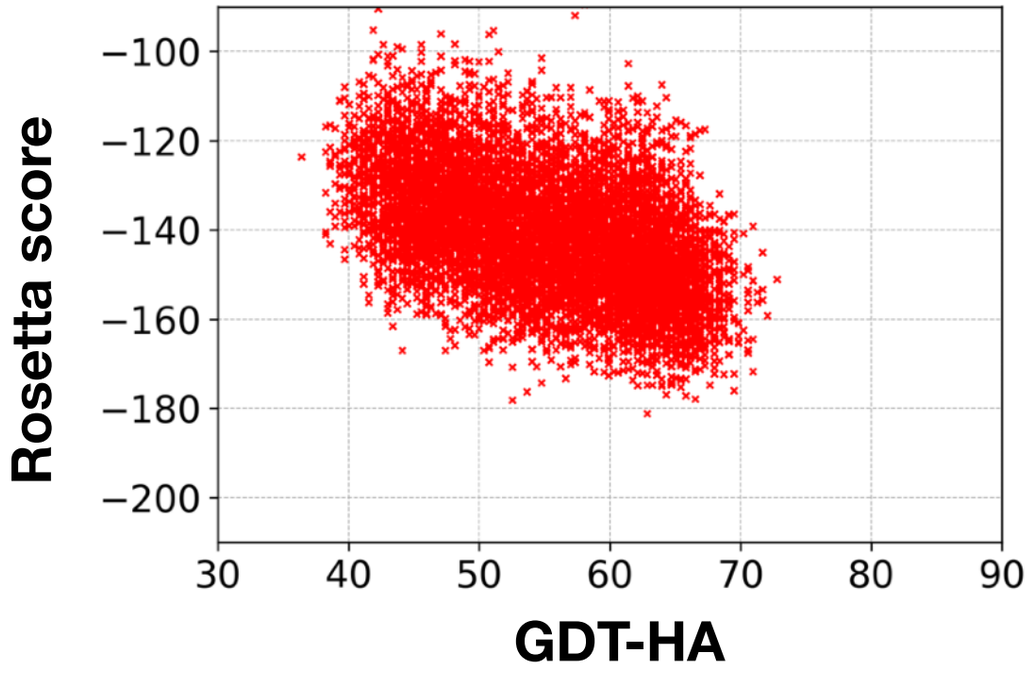


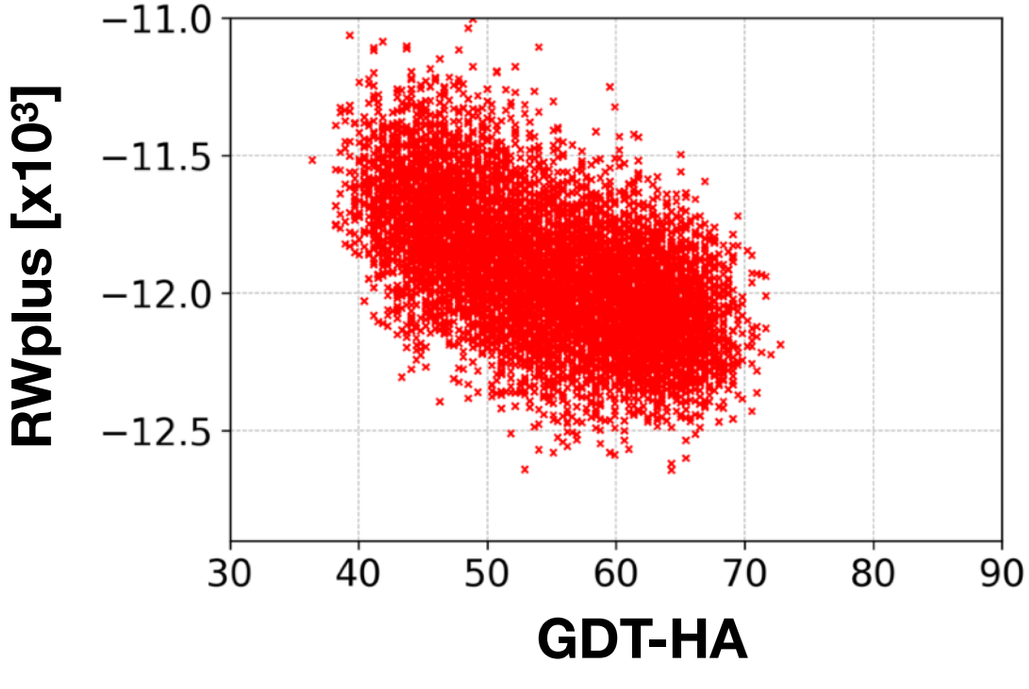




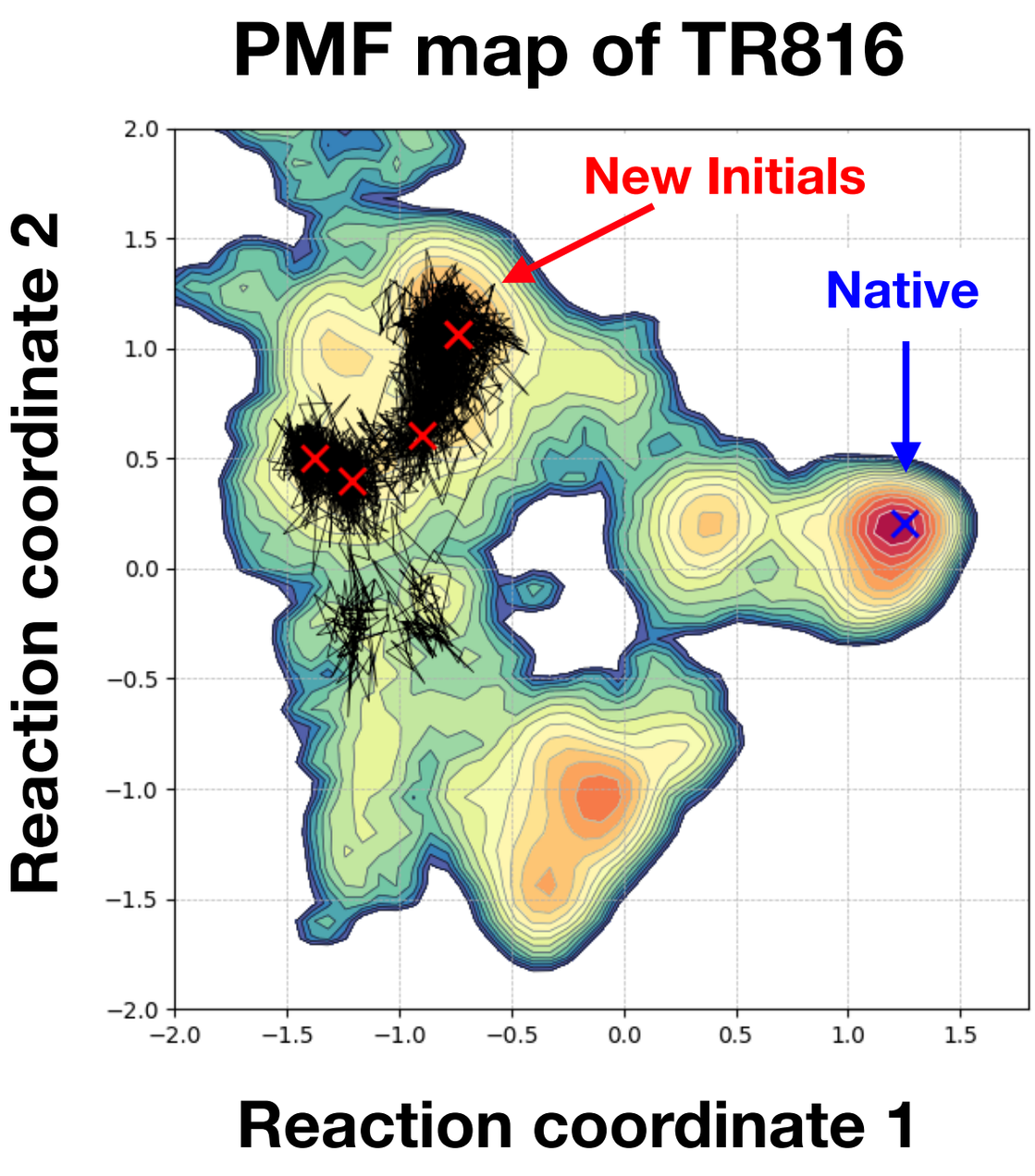
Iterative MD sampling and scoring

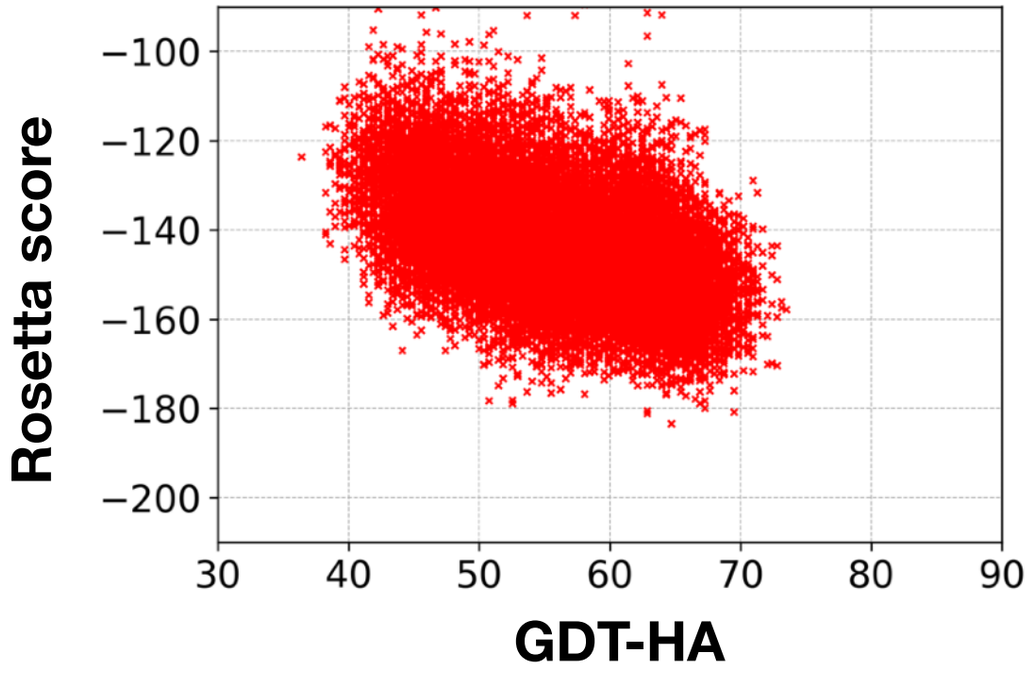


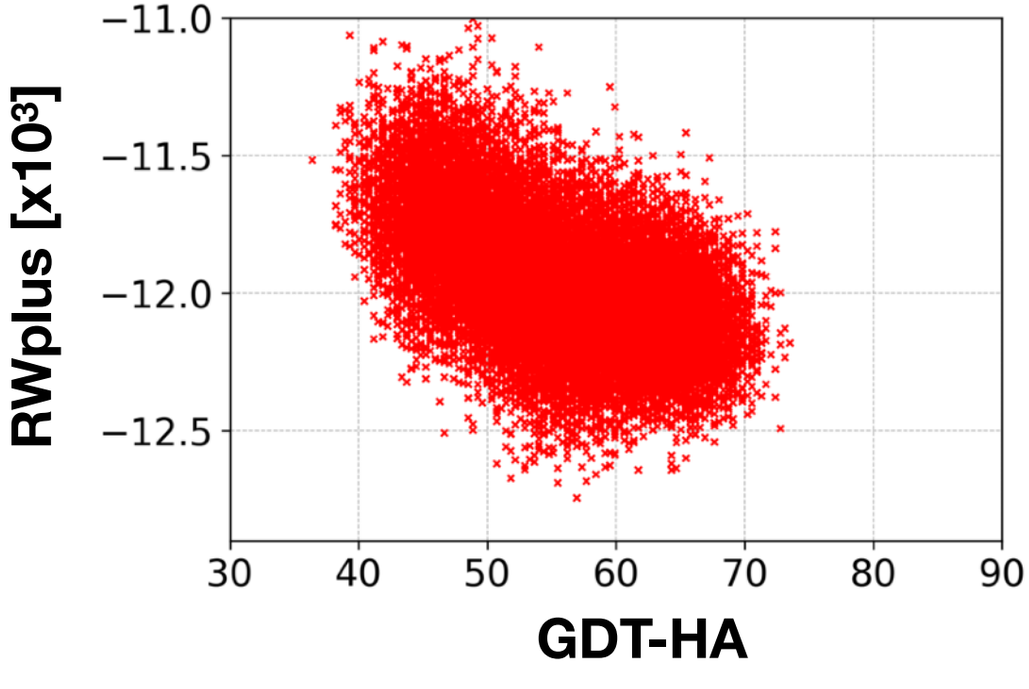




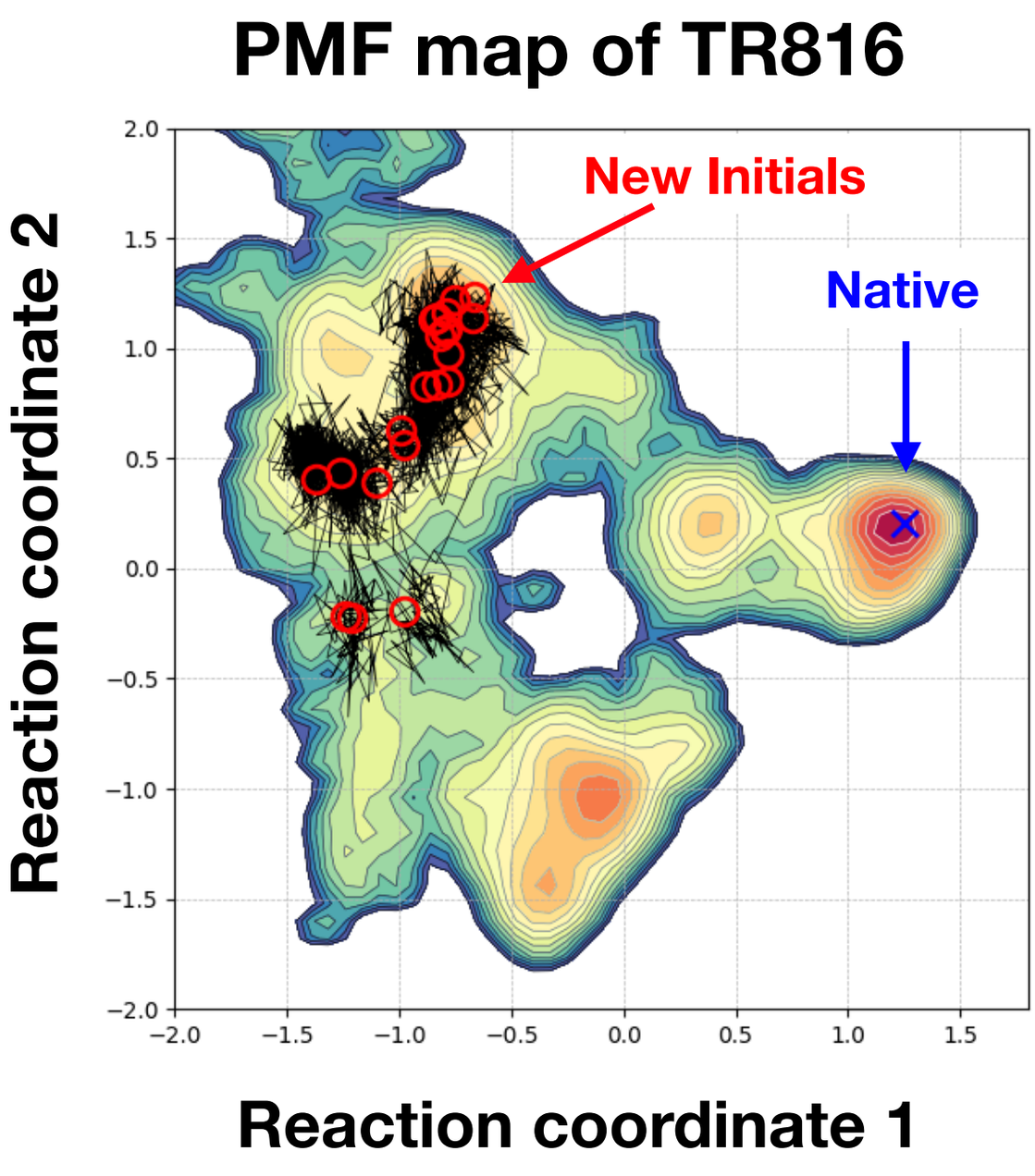
Iterative MD sampling and scoring

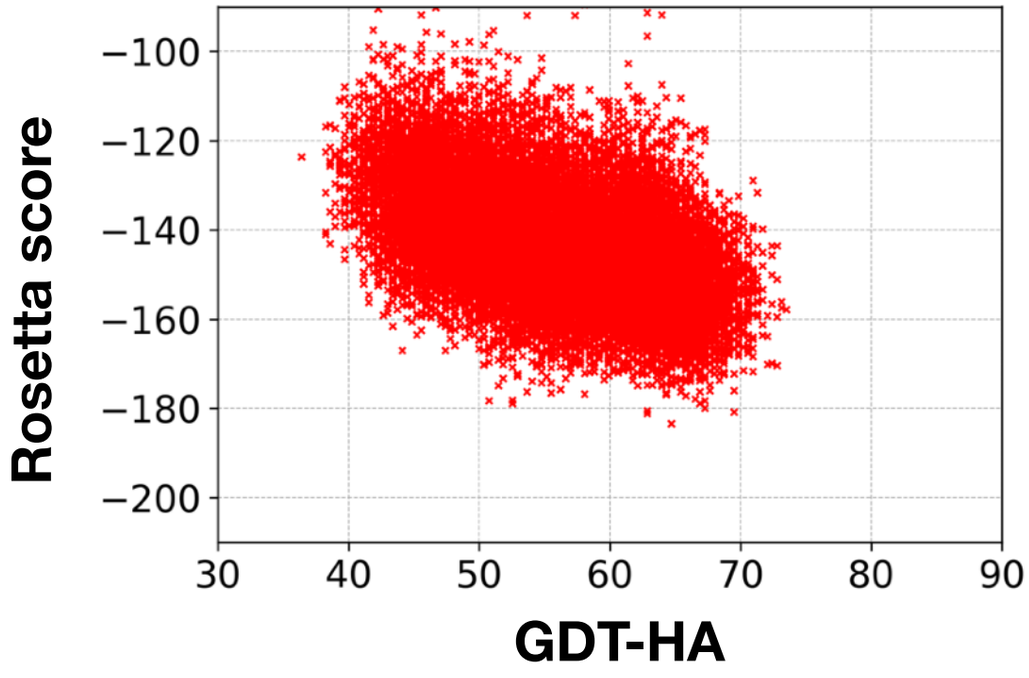


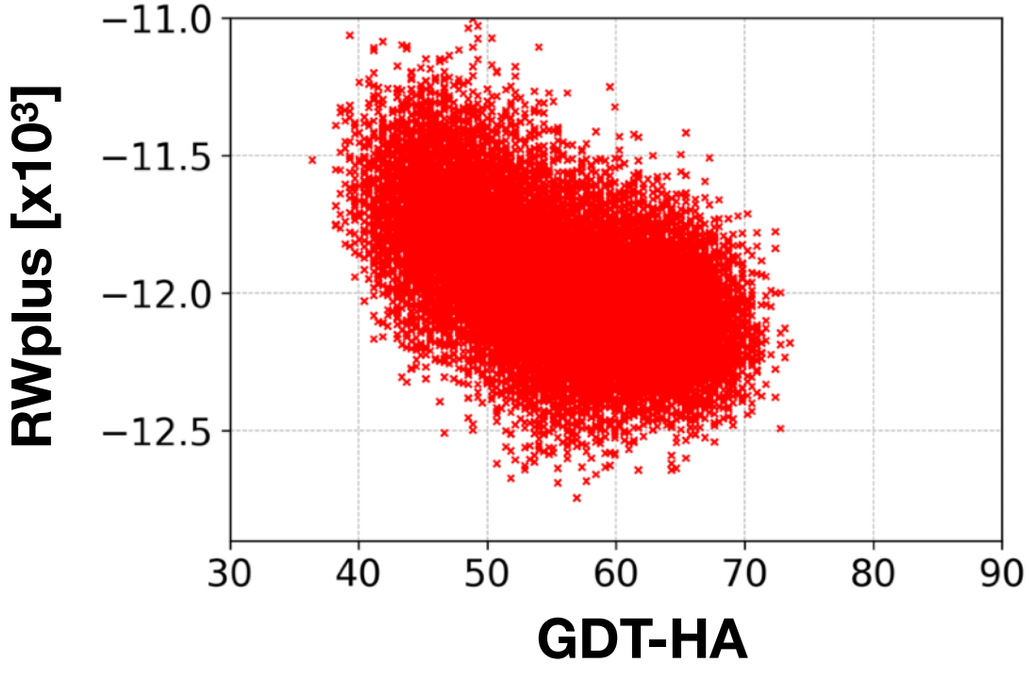




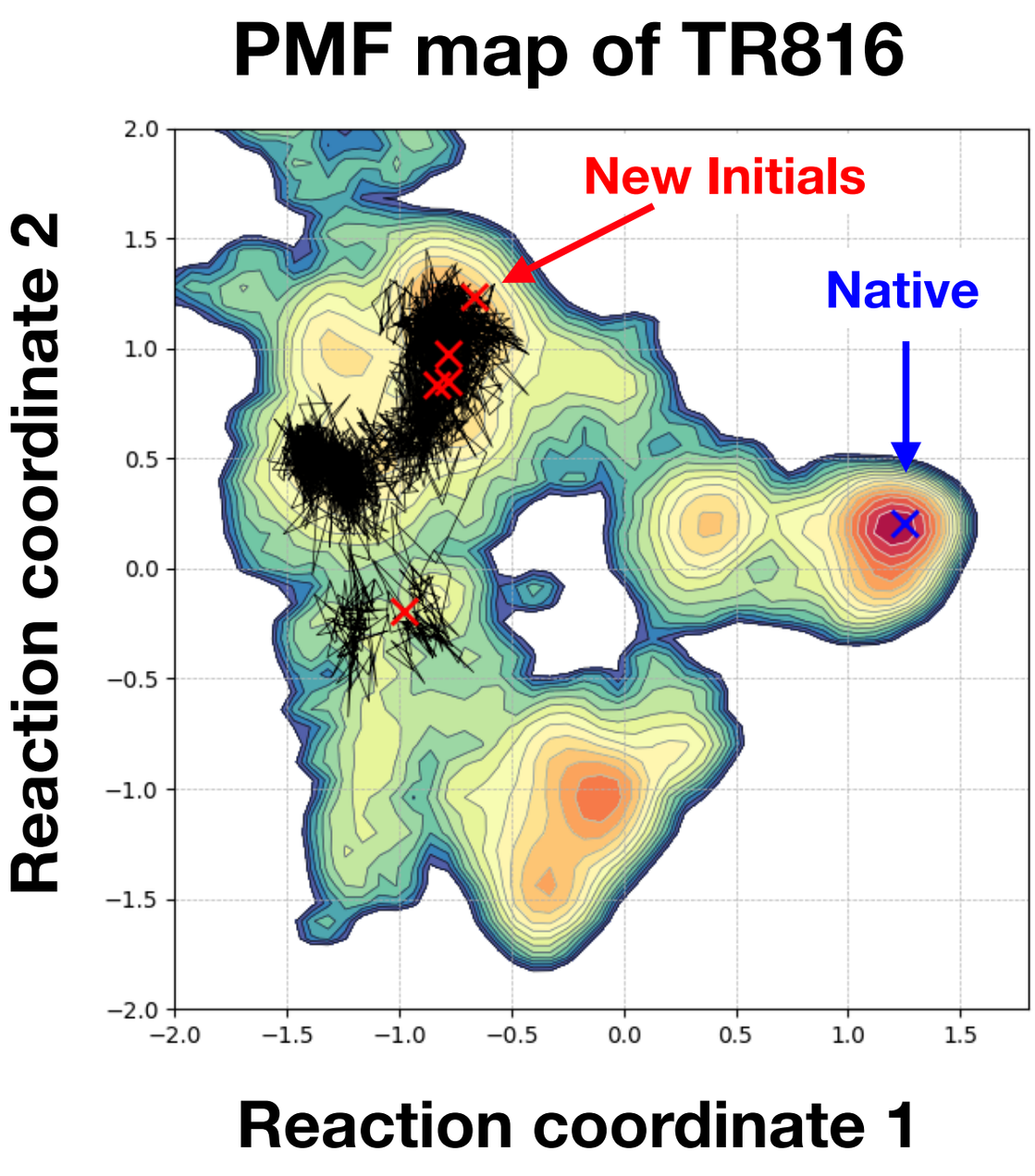
Iterative MD sampling and scoring

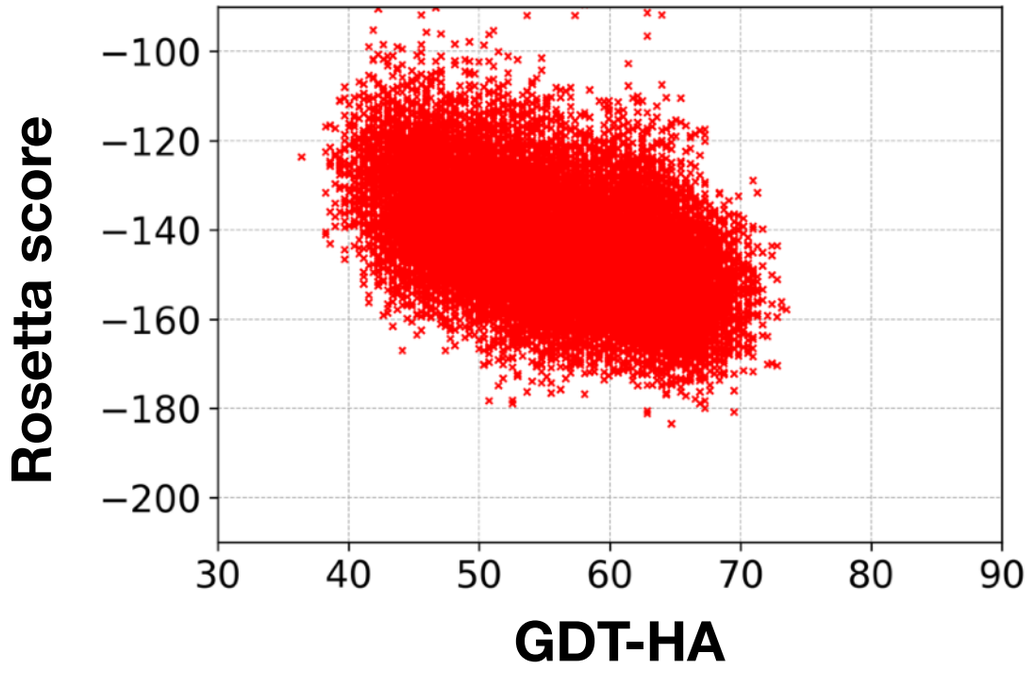


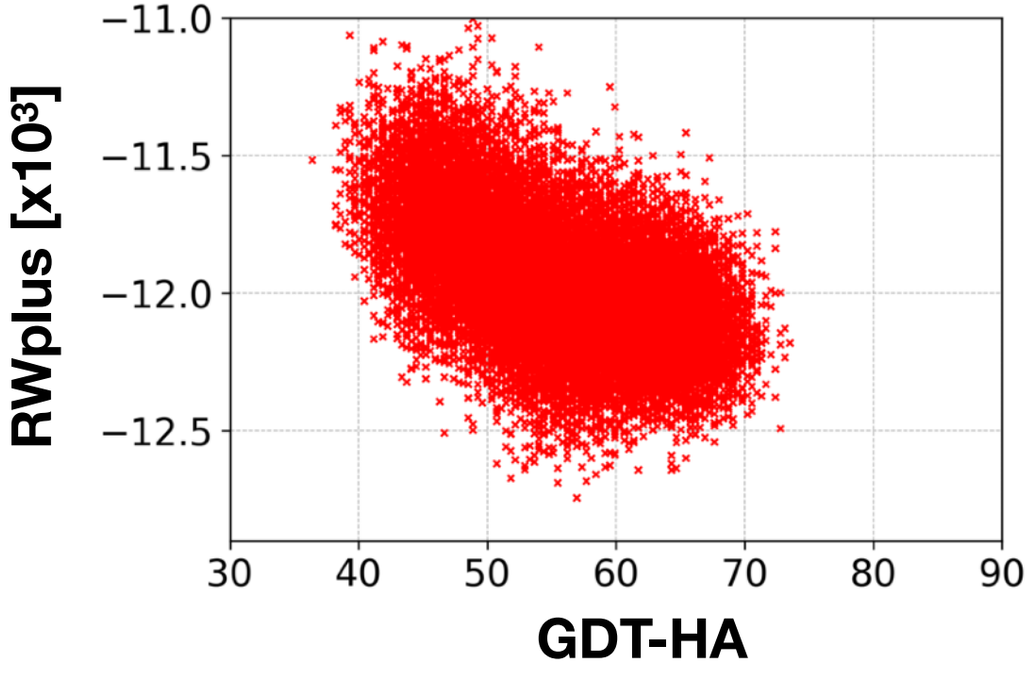




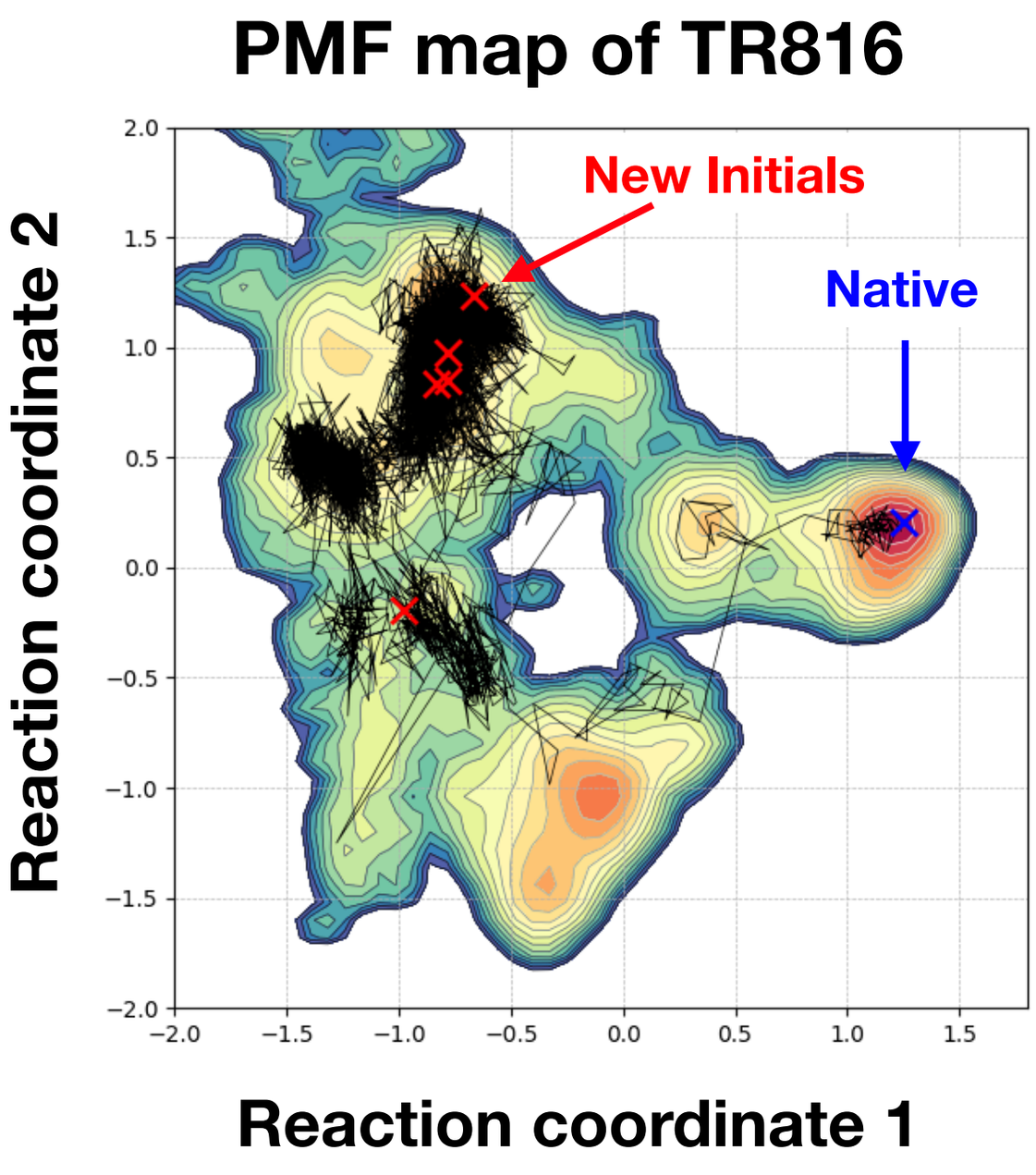
Iterative MD sampling and scoring

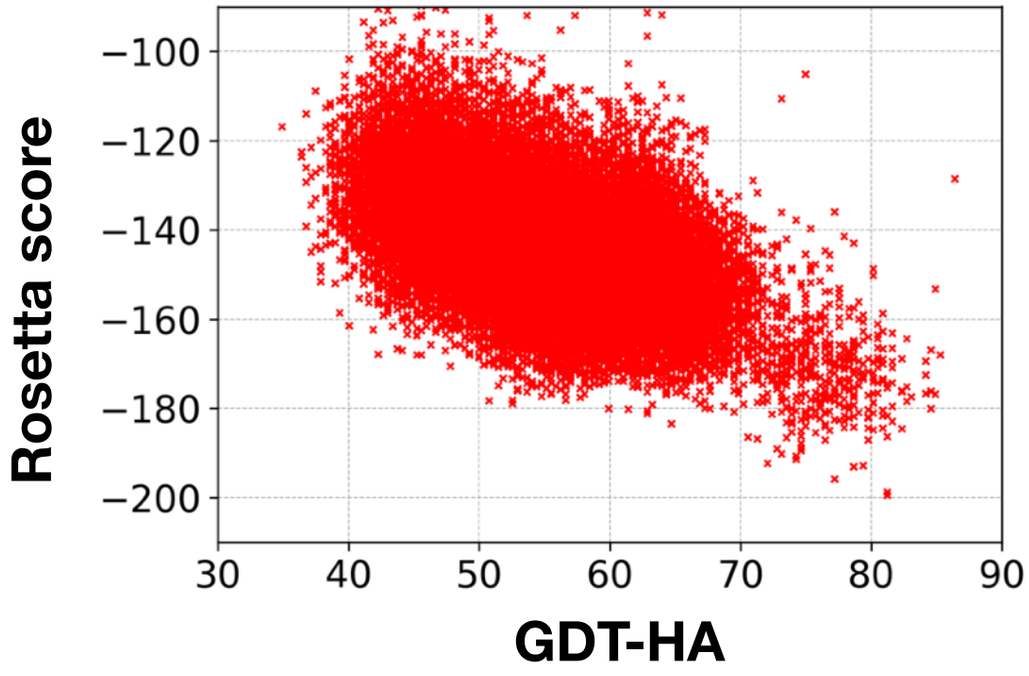


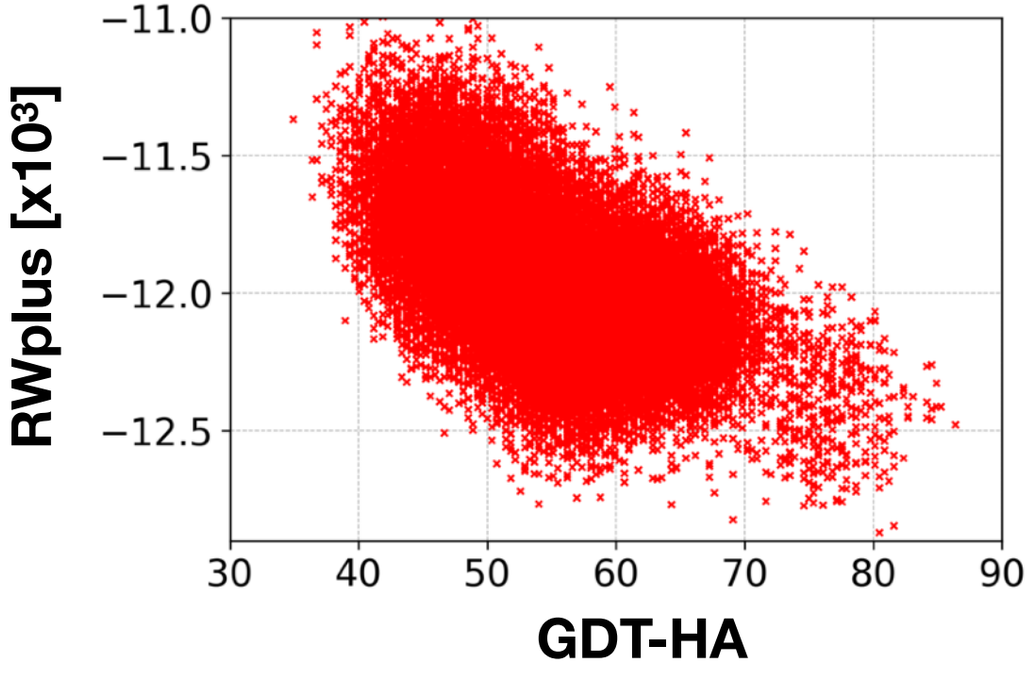




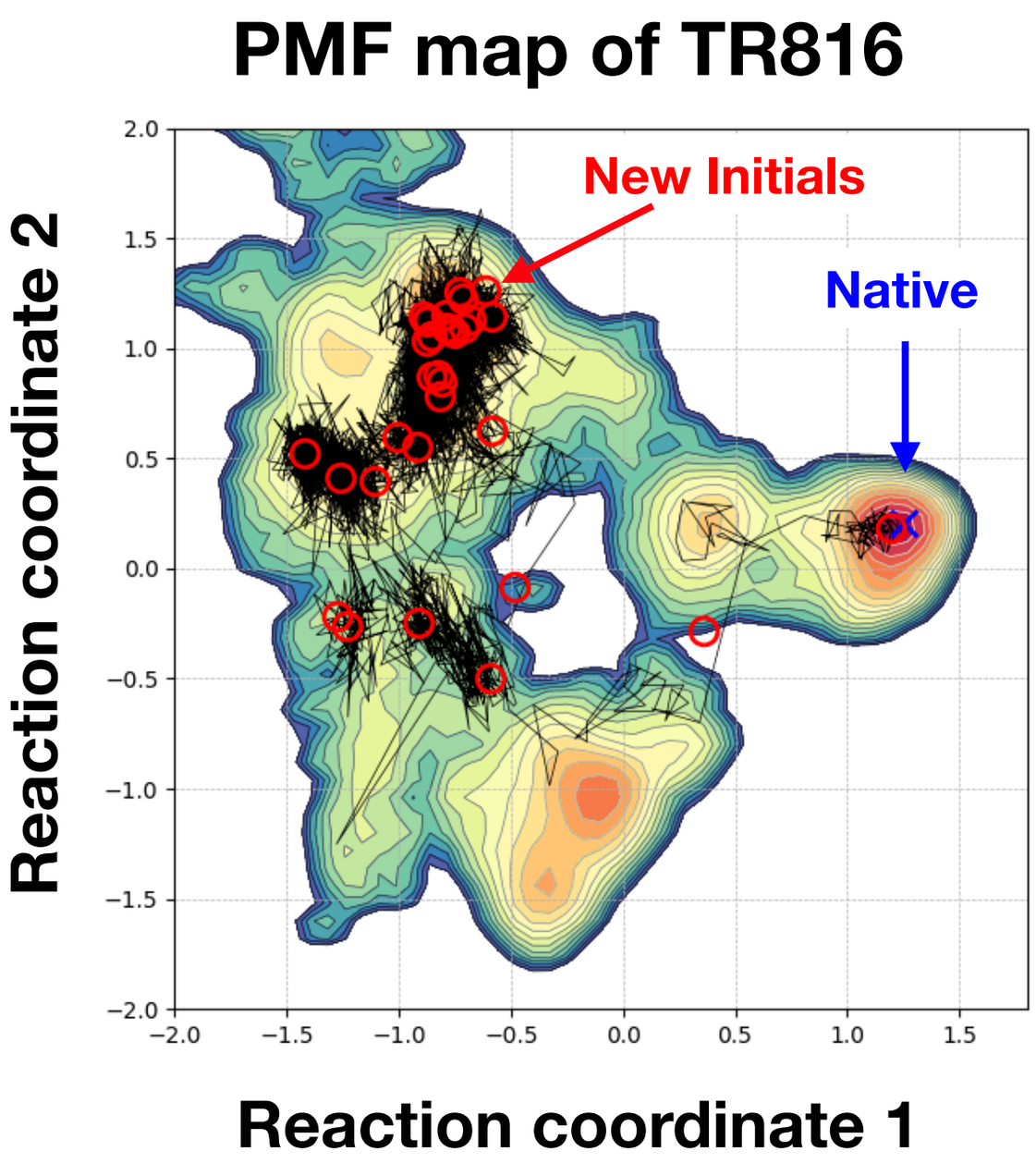
Iterative MD sampling and scoring

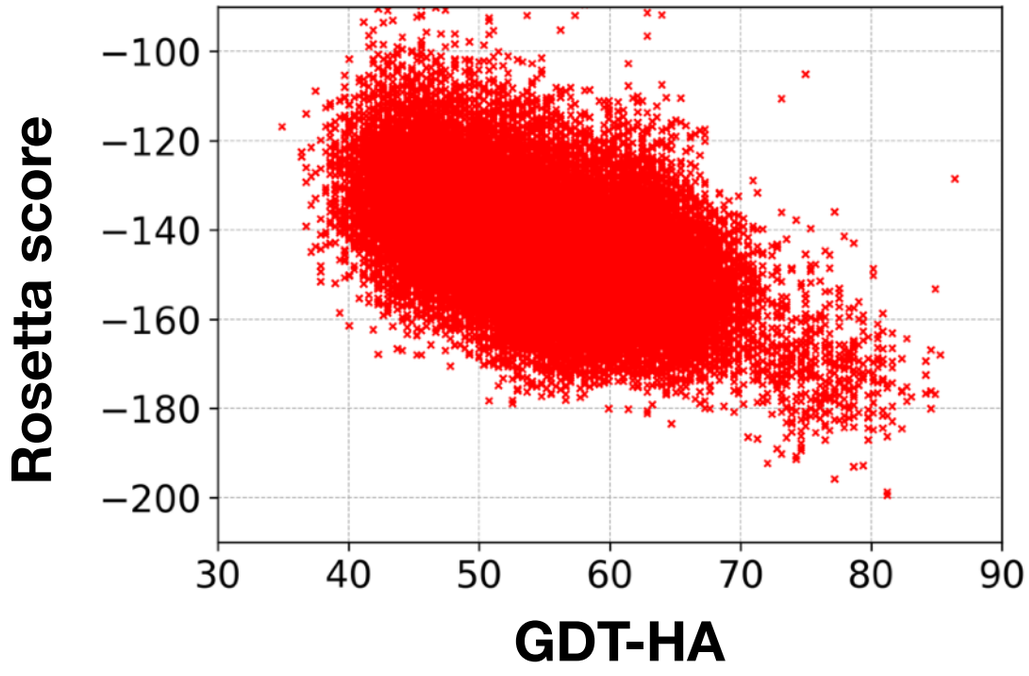


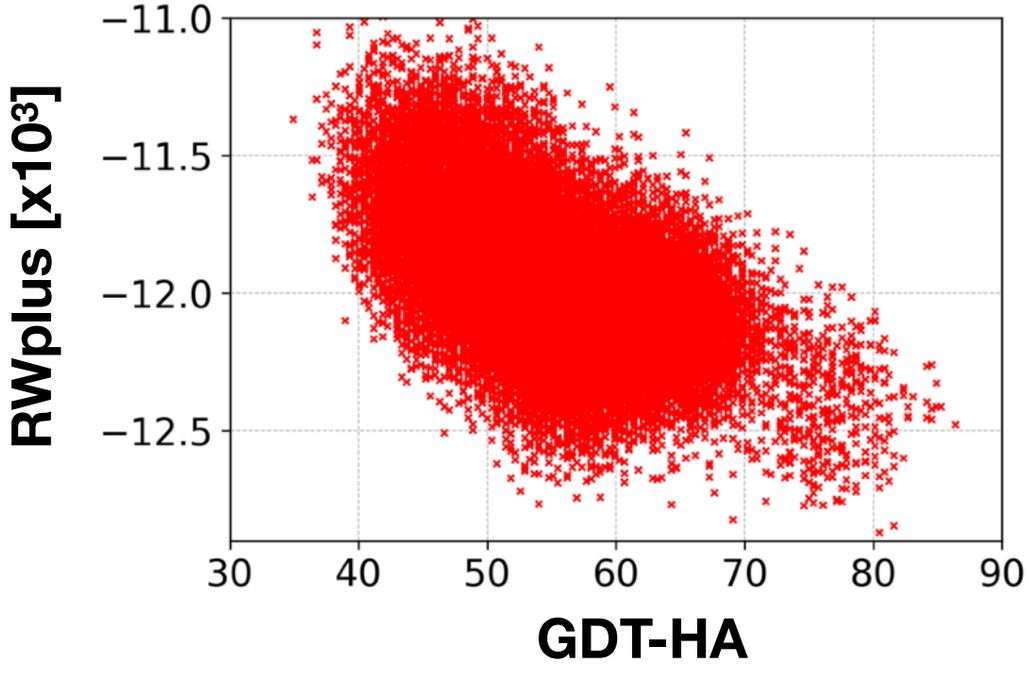




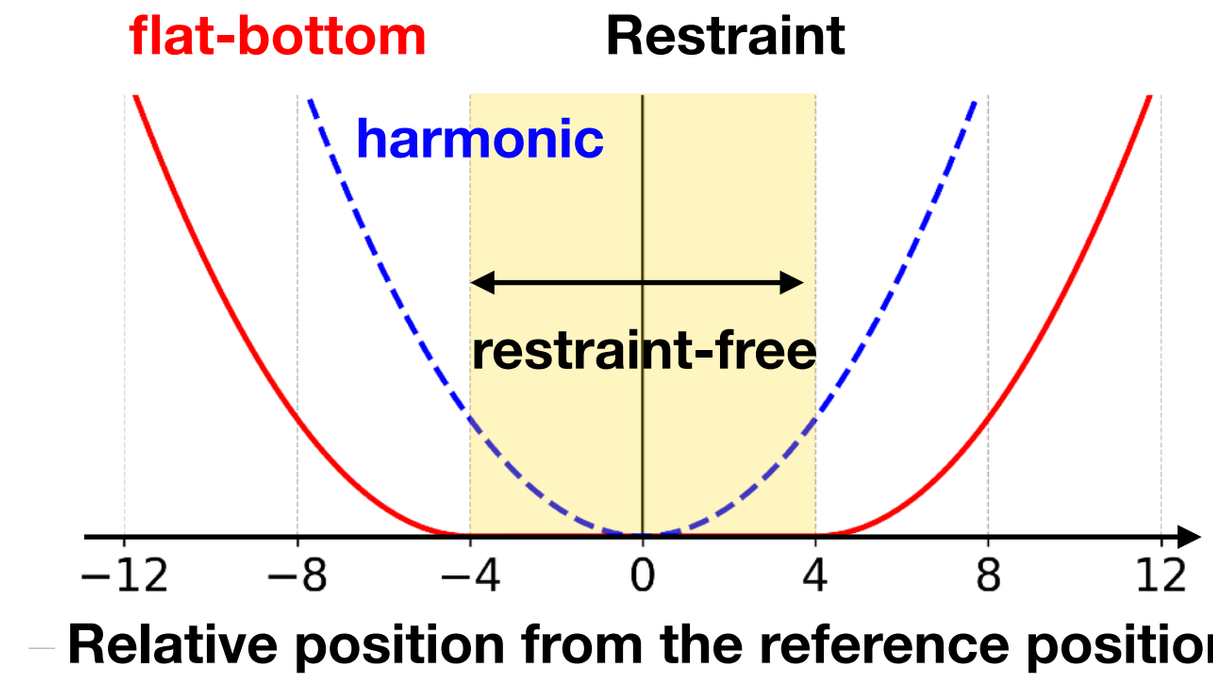
Iterative MD sampling and scoring

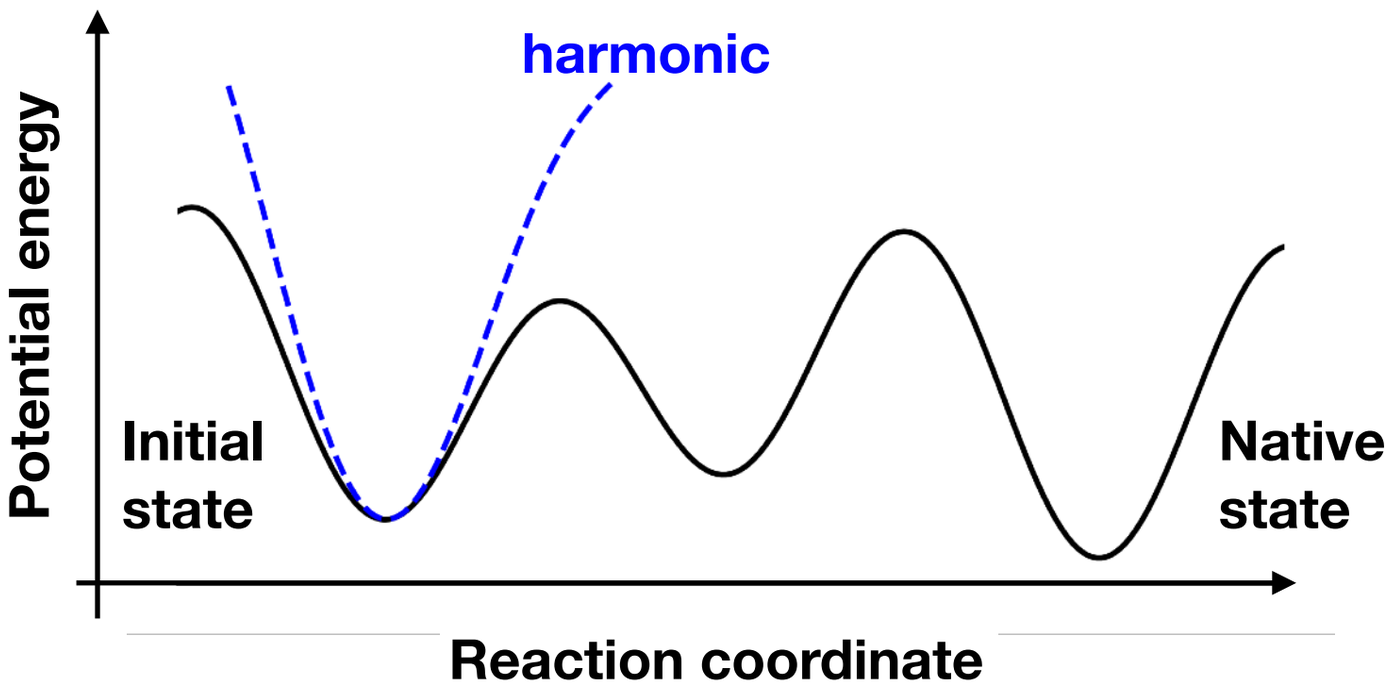




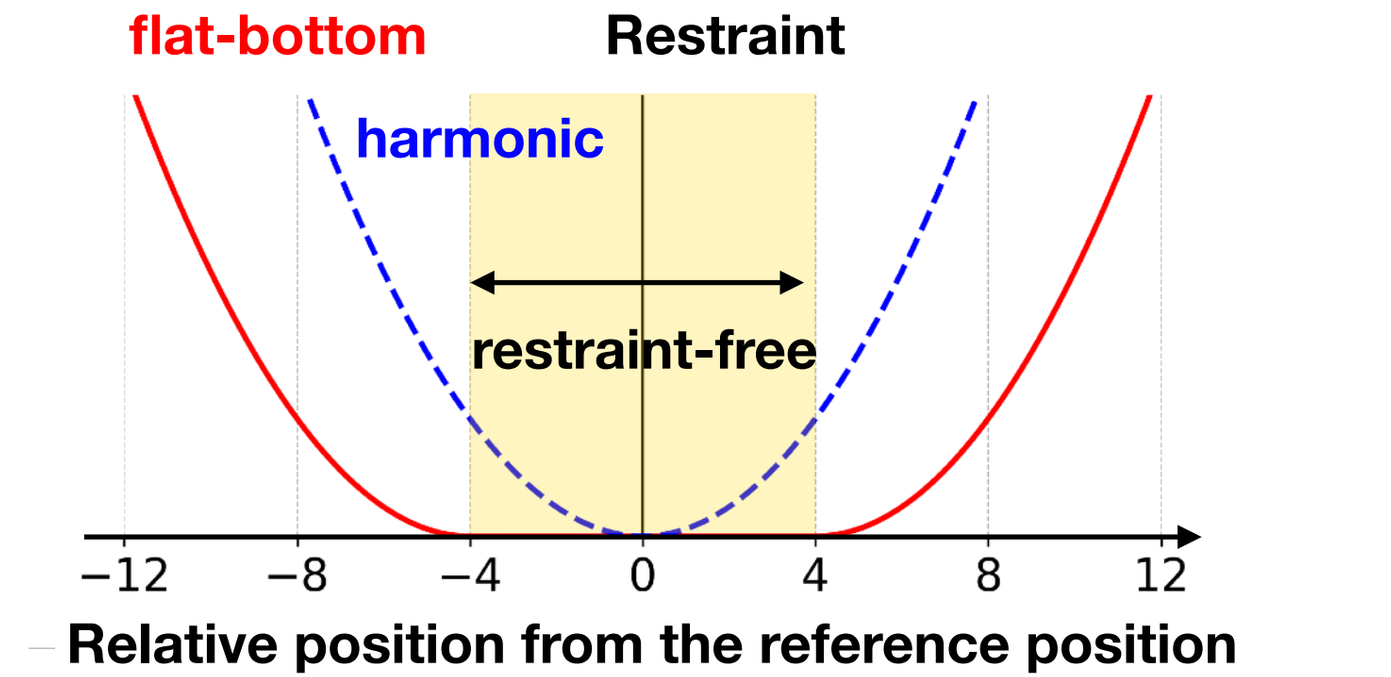


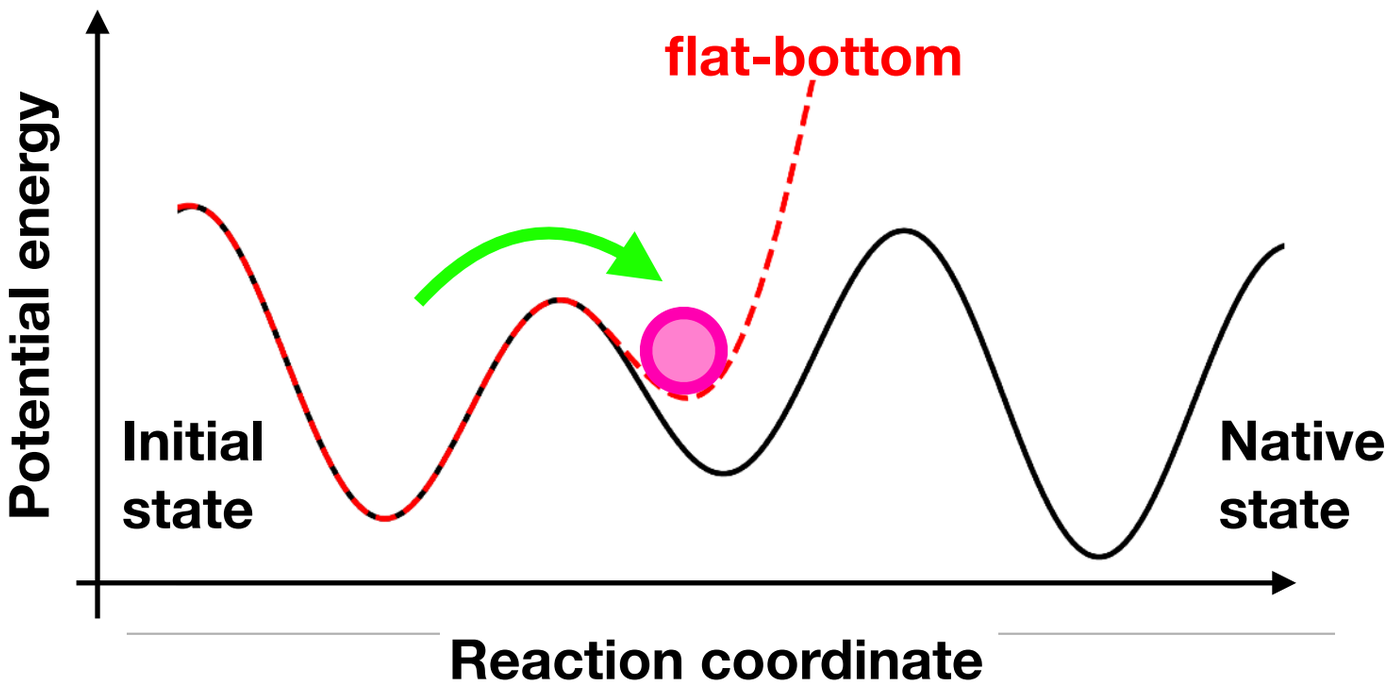
Flat-bottom harmonic restraints



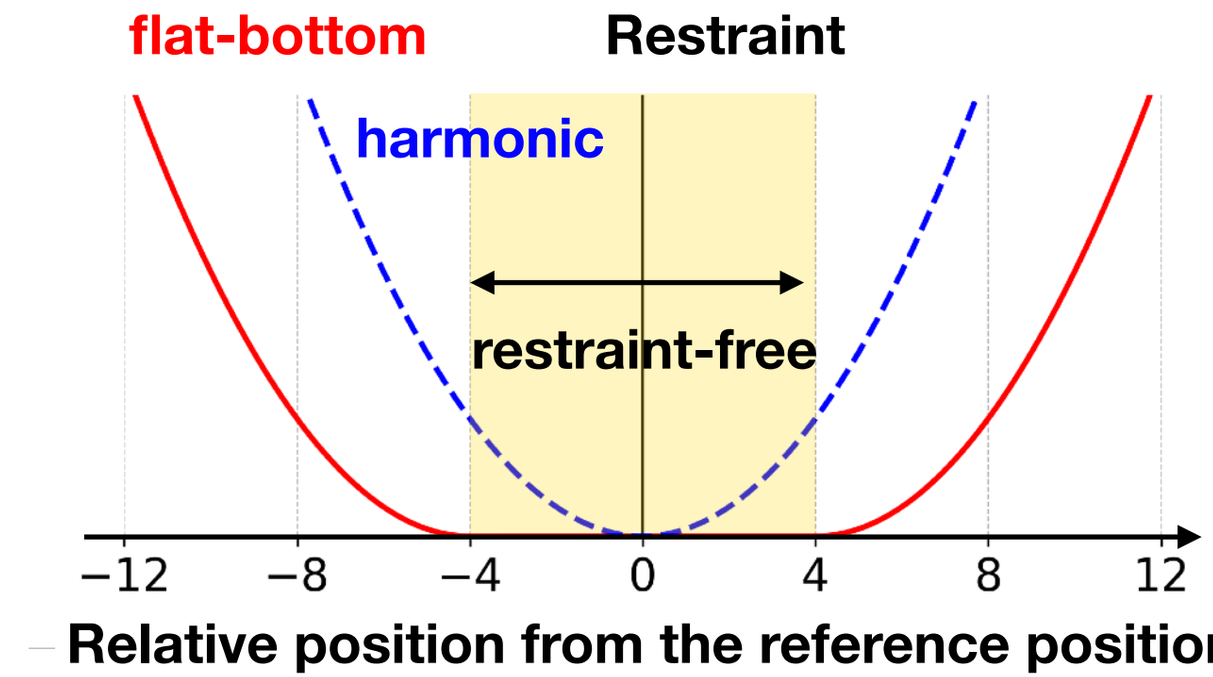


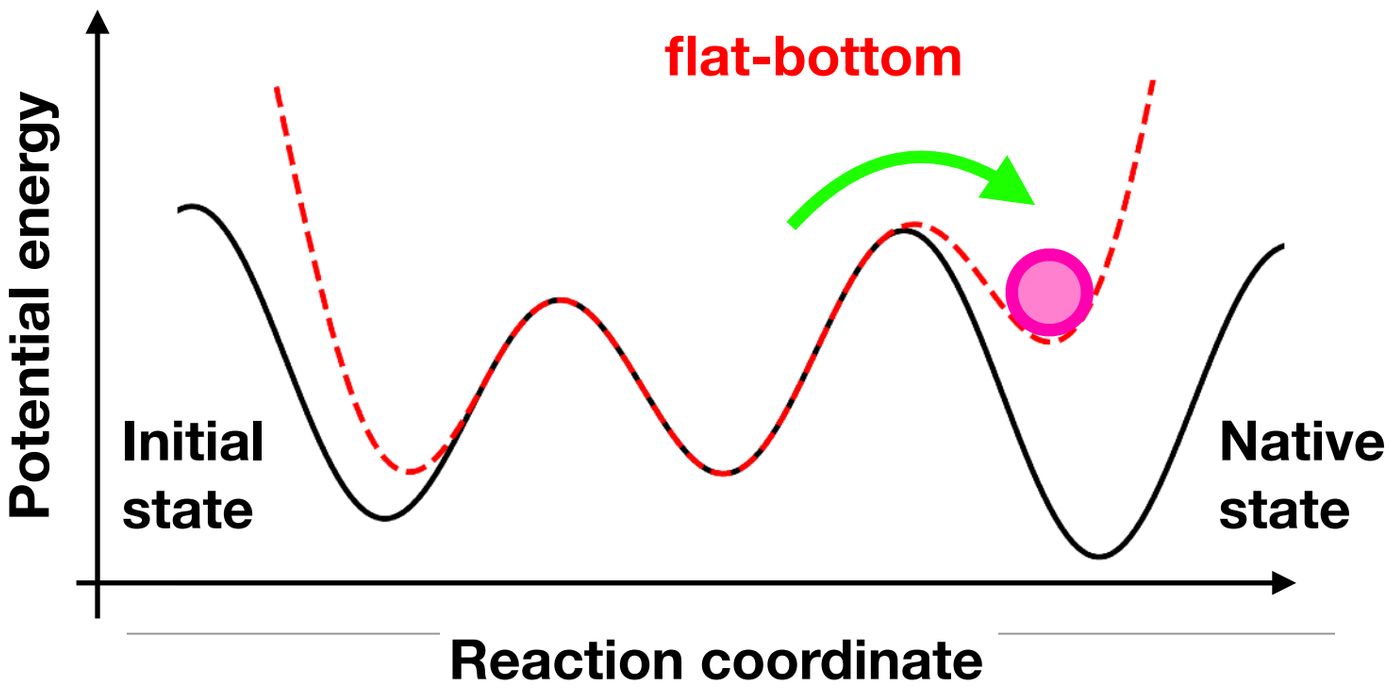
Flat-bottom harmonic restraints



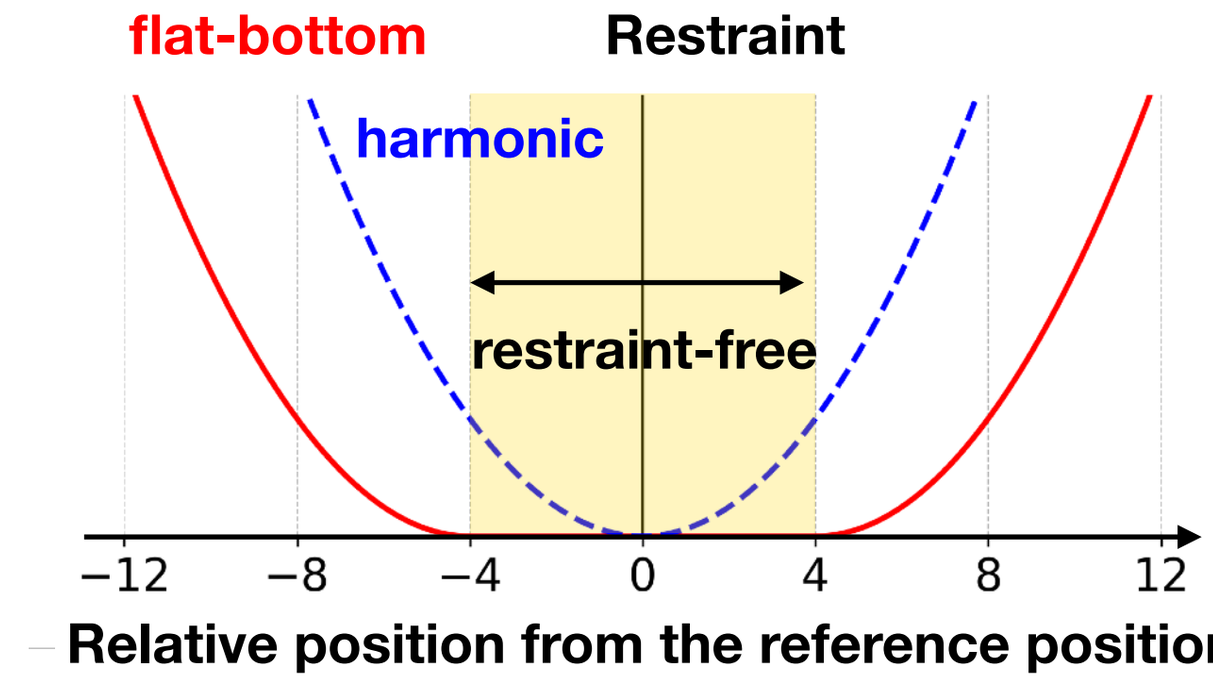


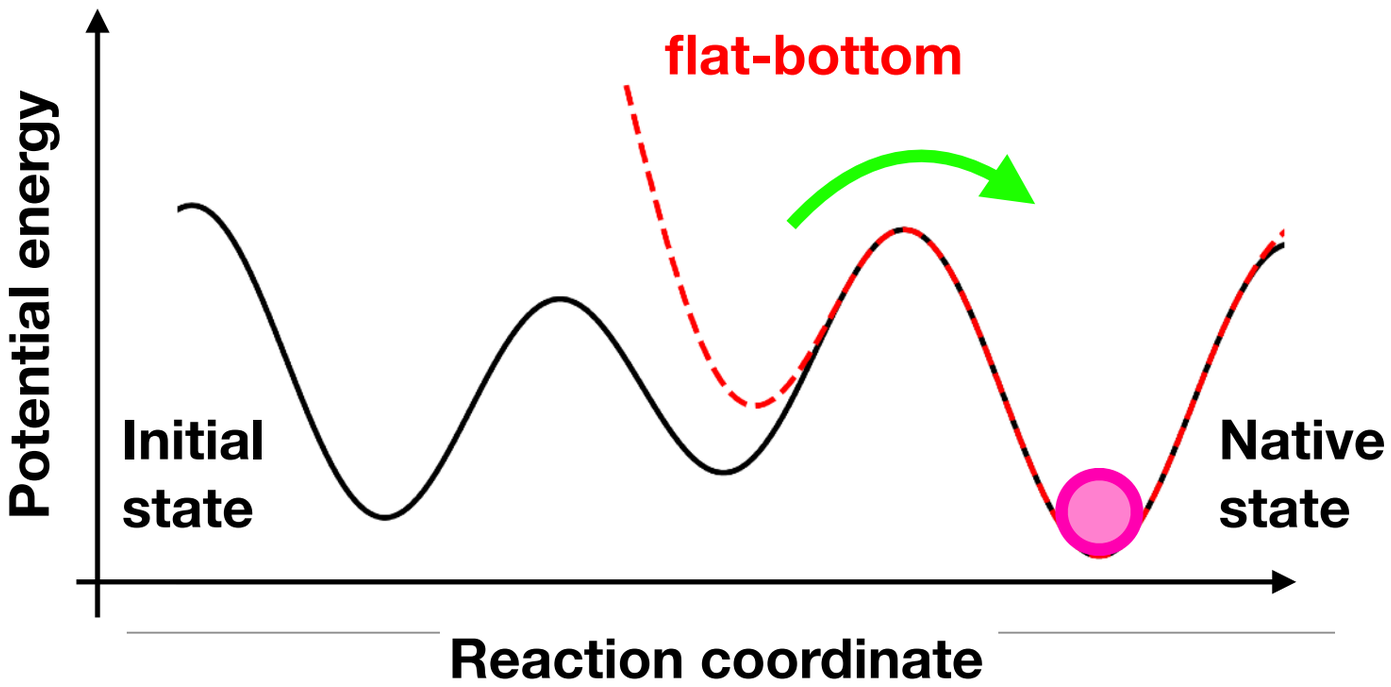
Flat-bottom harmonic restraints



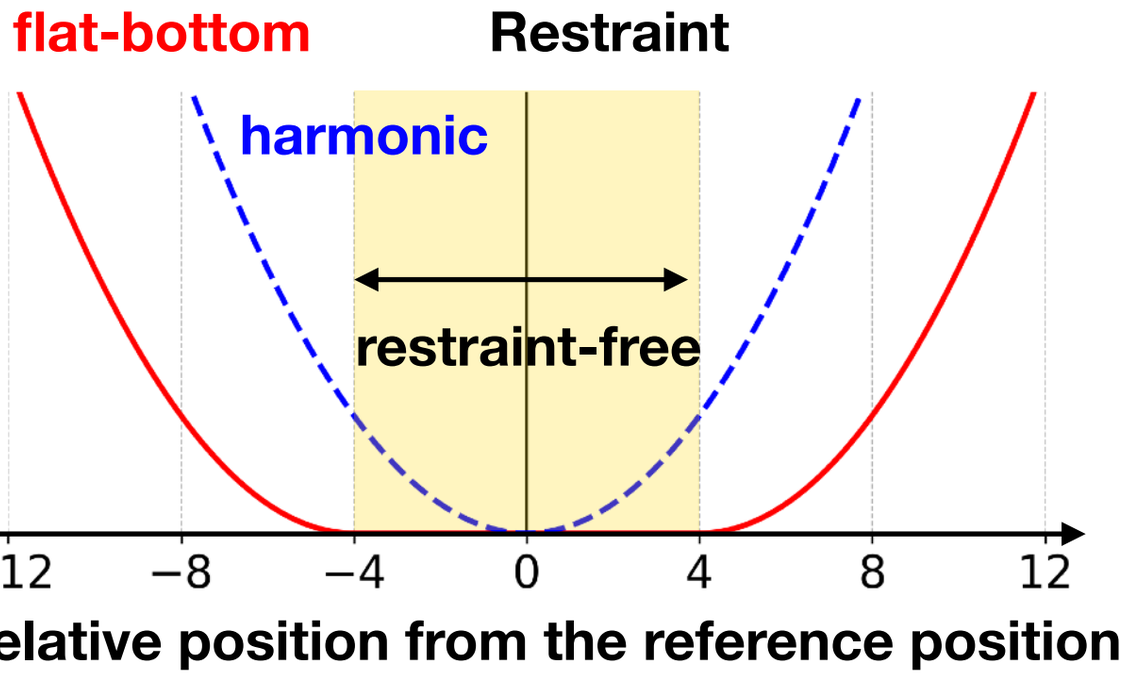


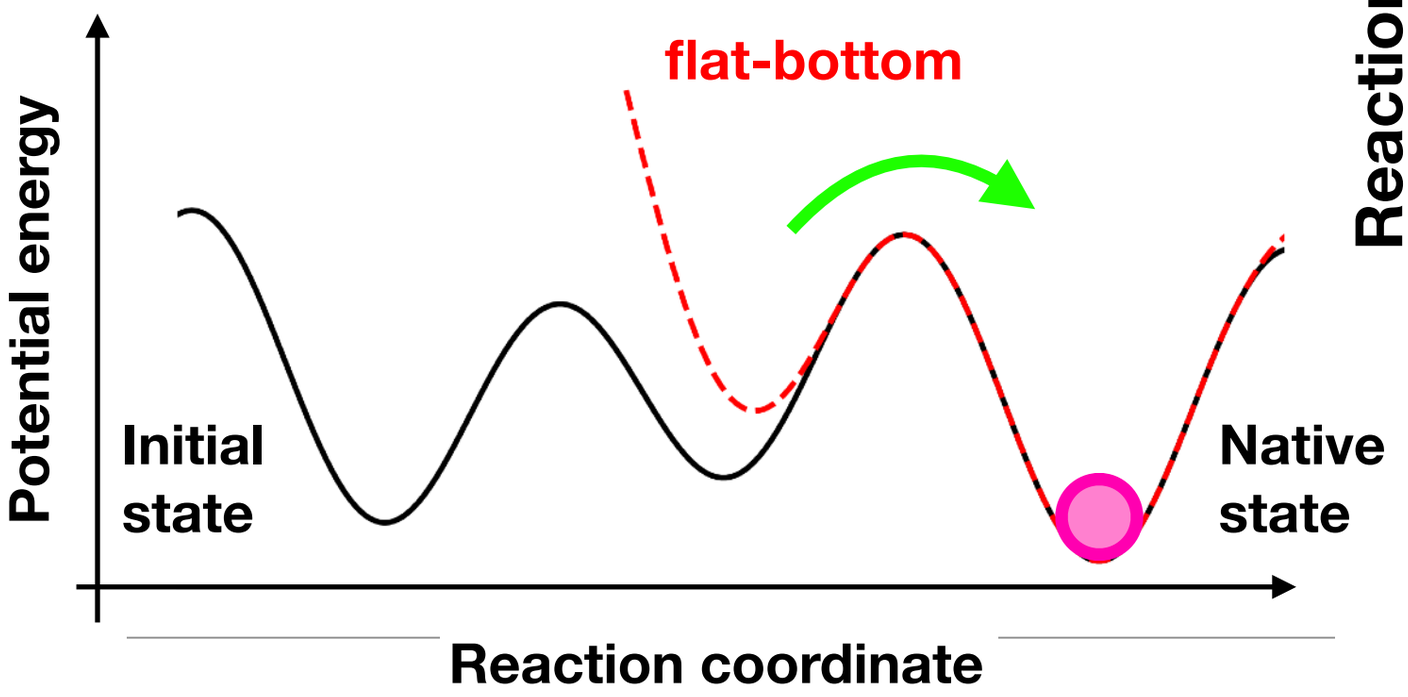
Flat-bottom harmonic restraints

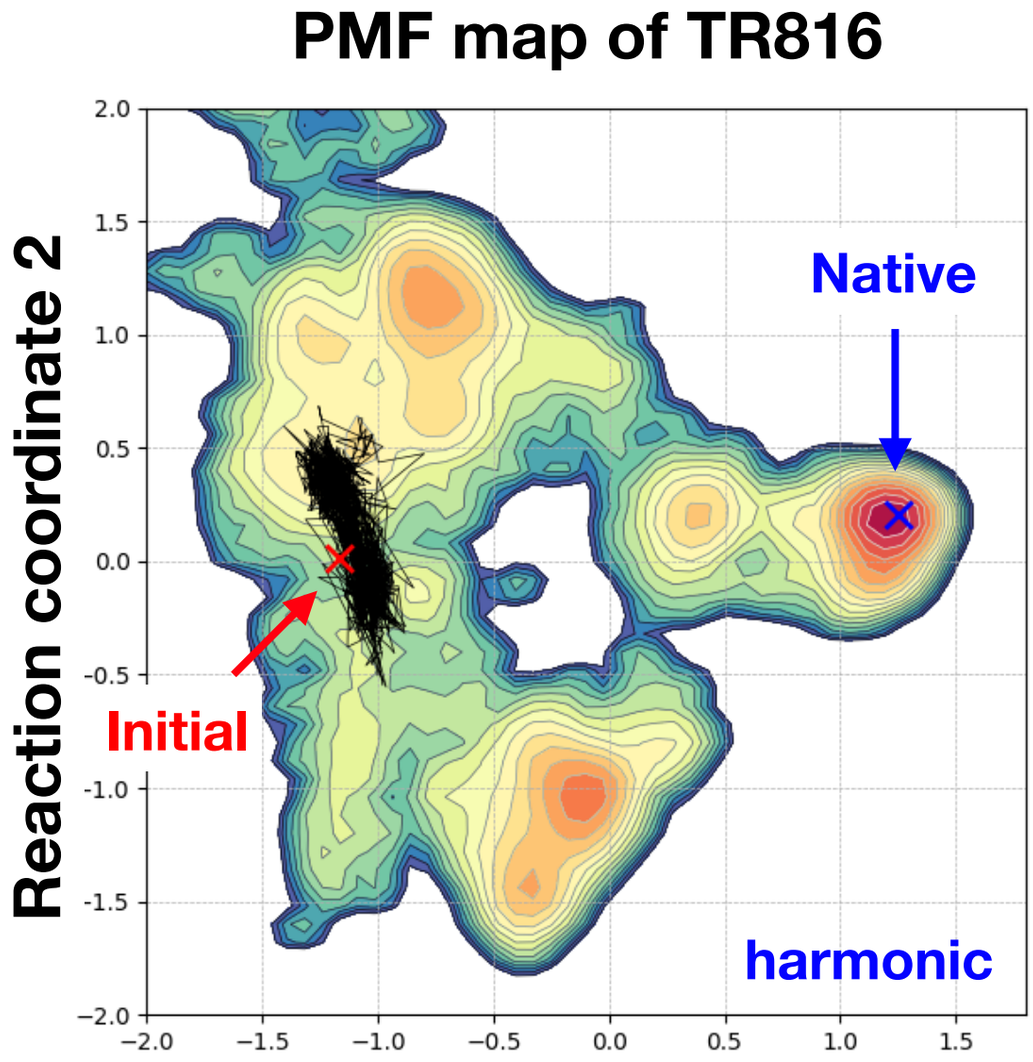




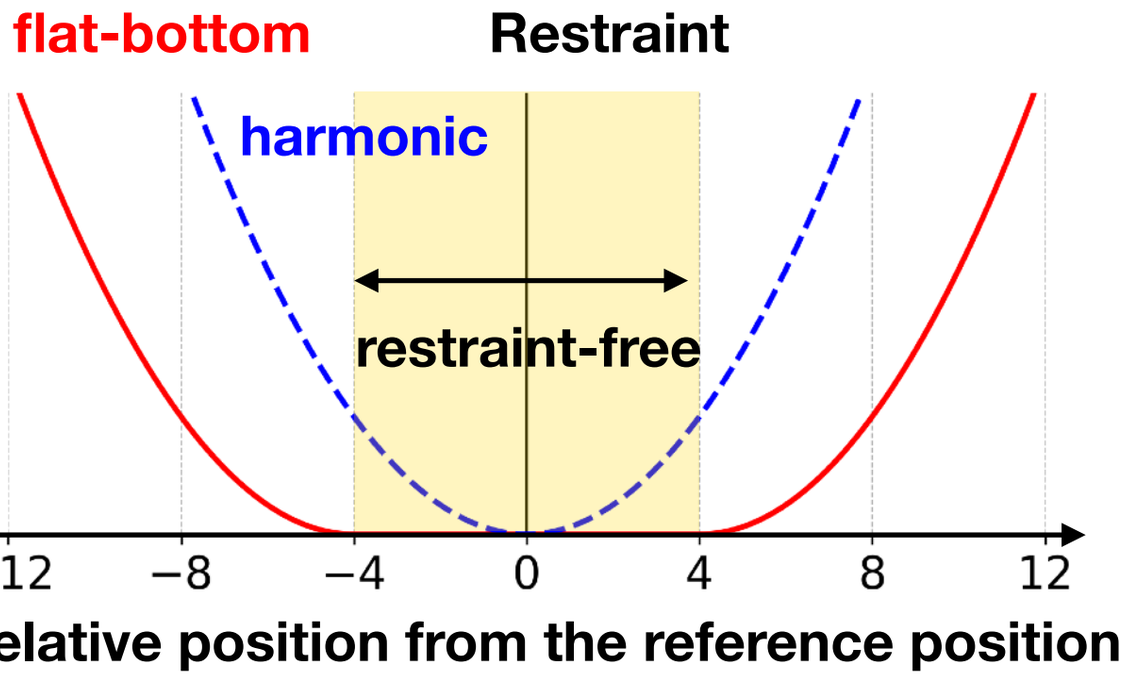
Flat-bottom harmonic restraints

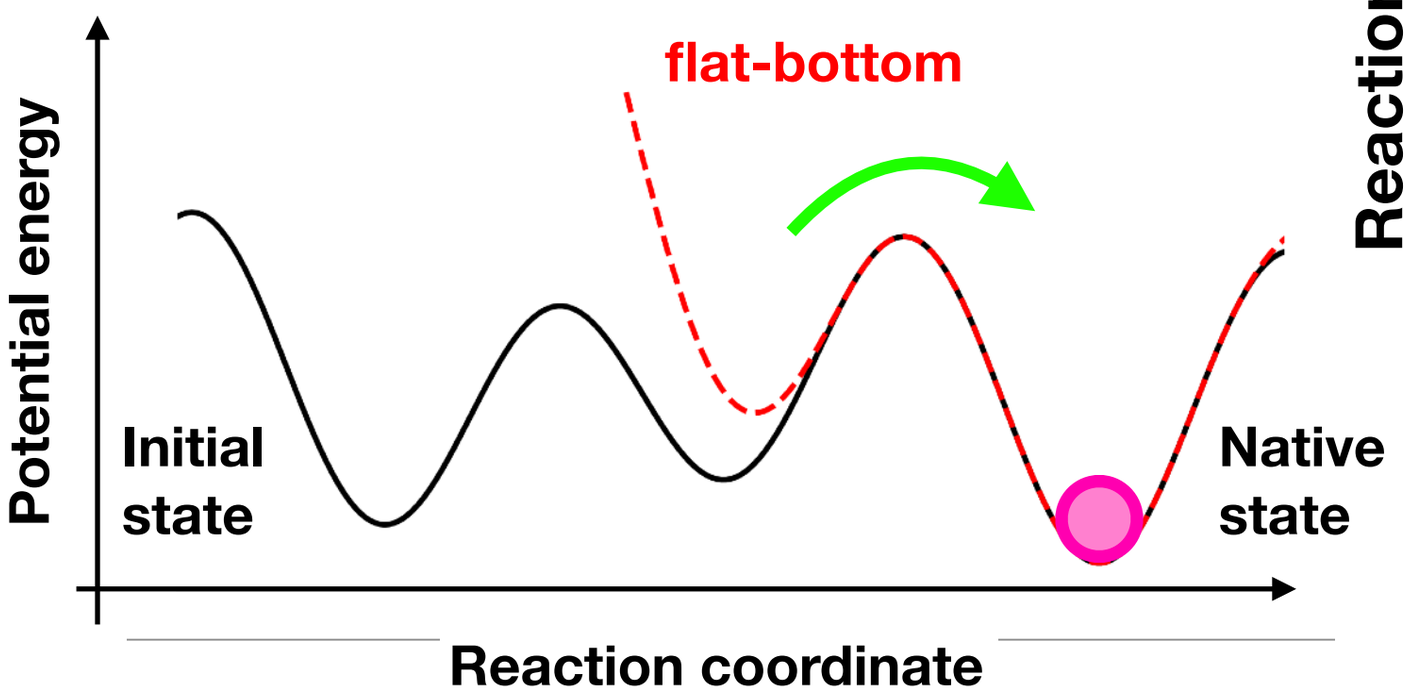


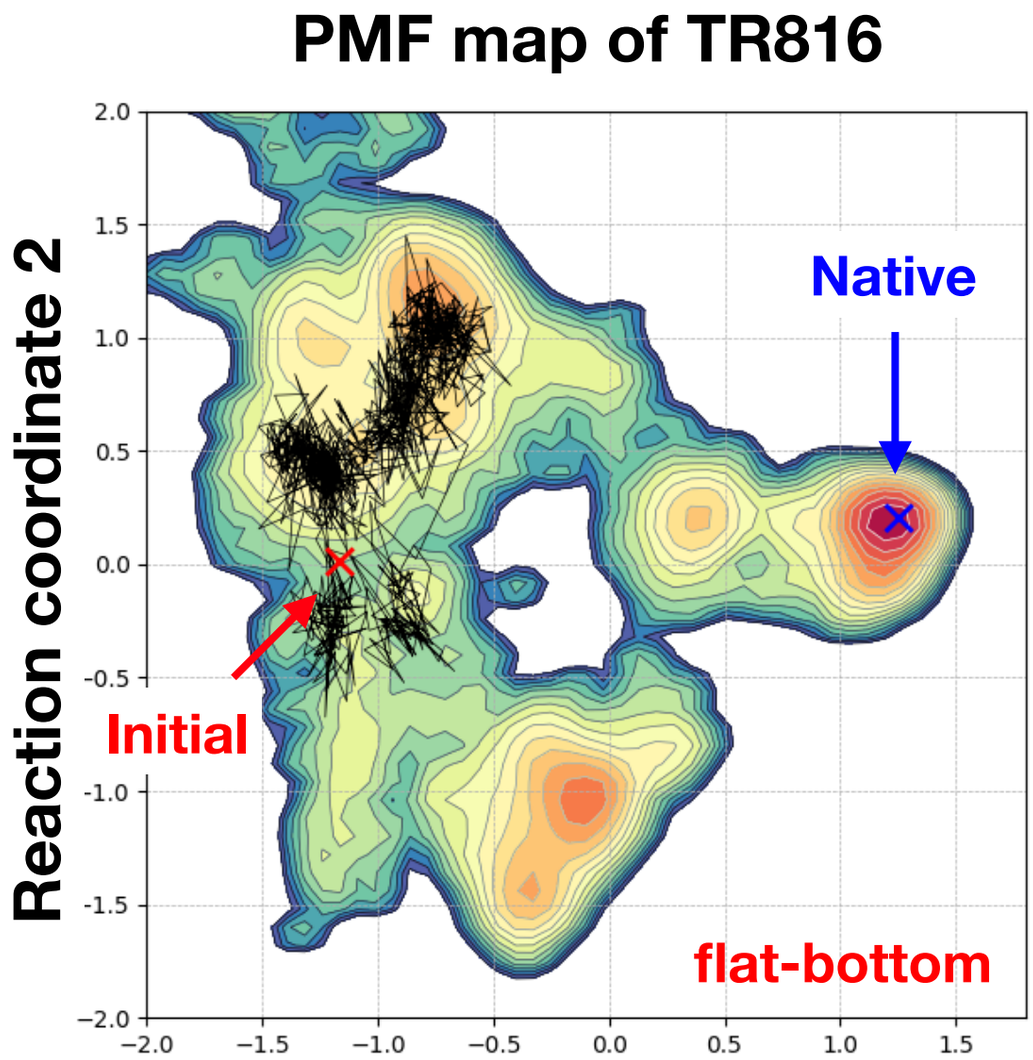




Flat-bottom harmonic restraints

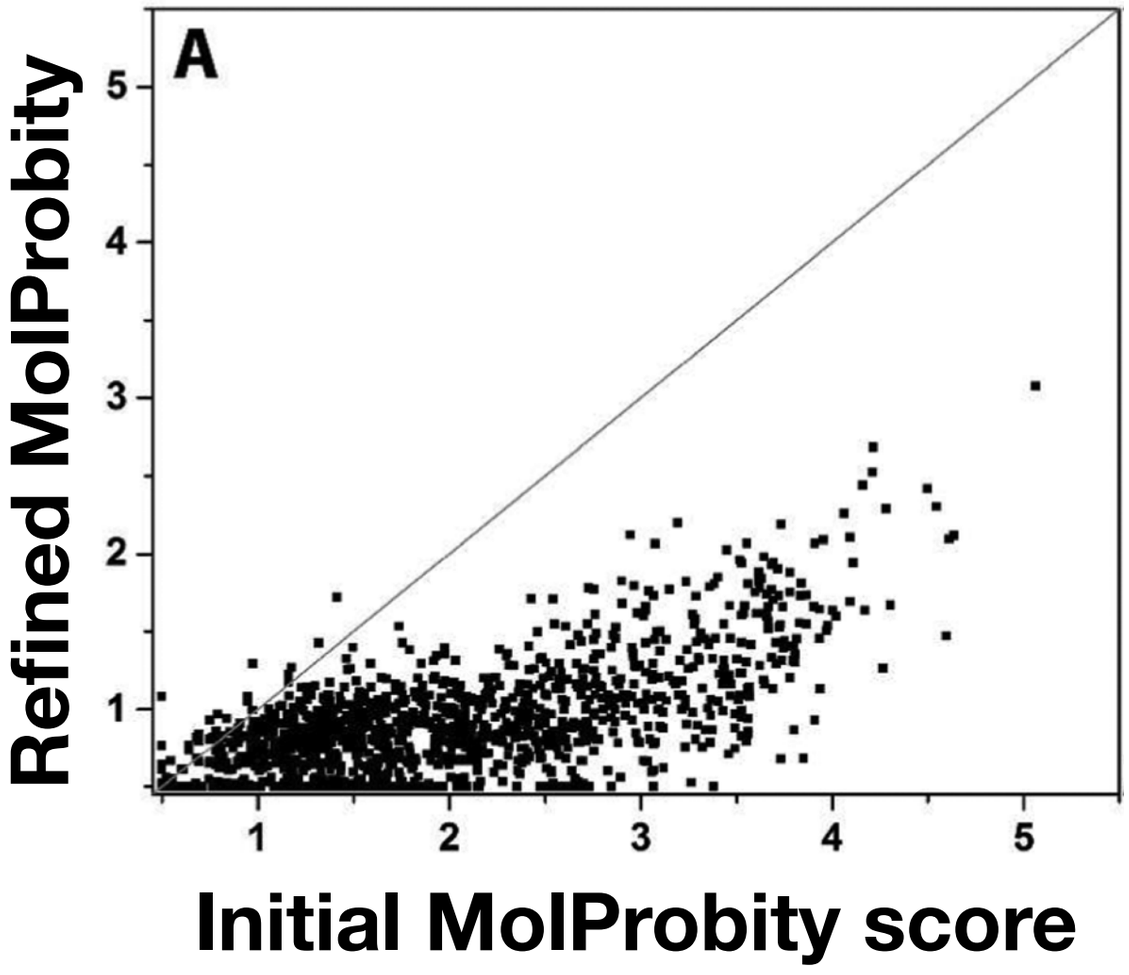






locPREFMD

- Quickly resolves stereochemical problems
- http://feiglab.org/locprefmd

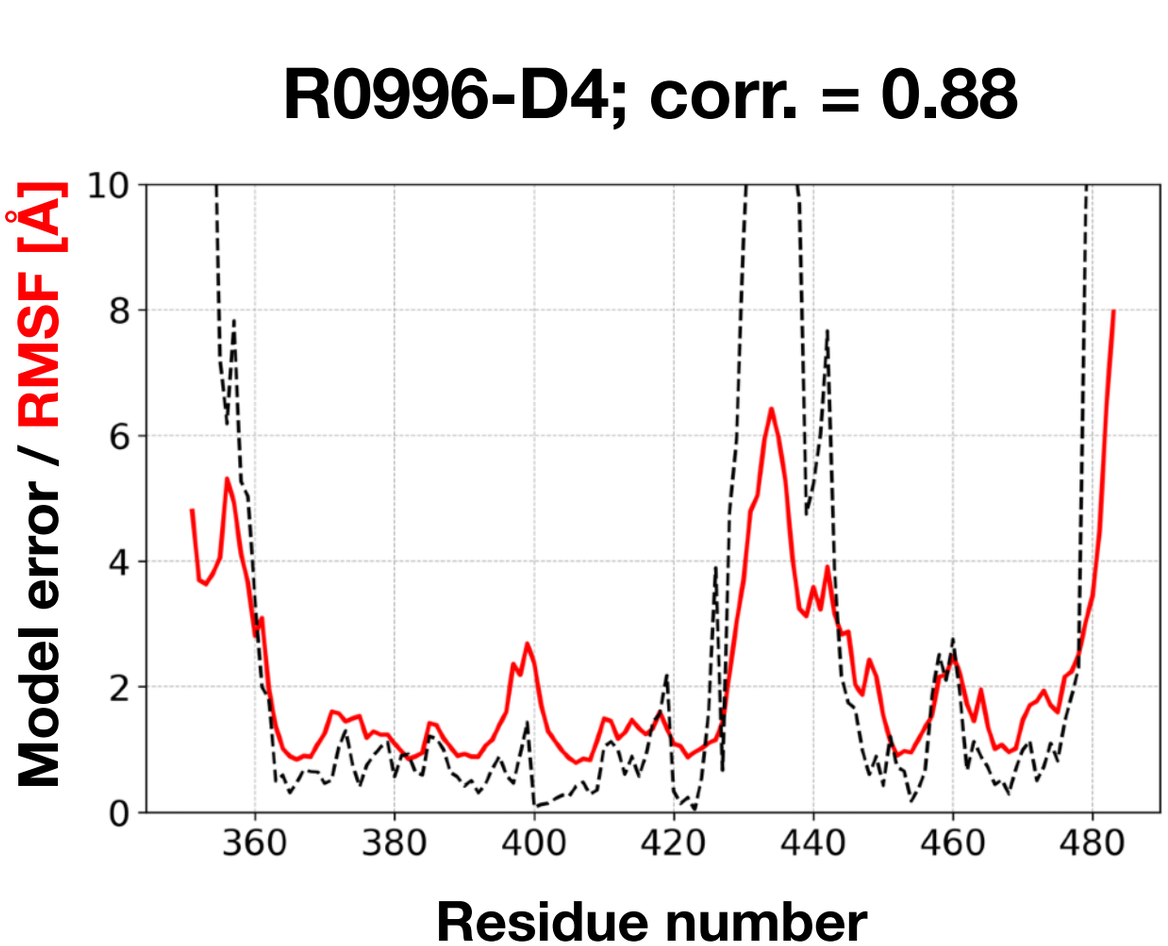


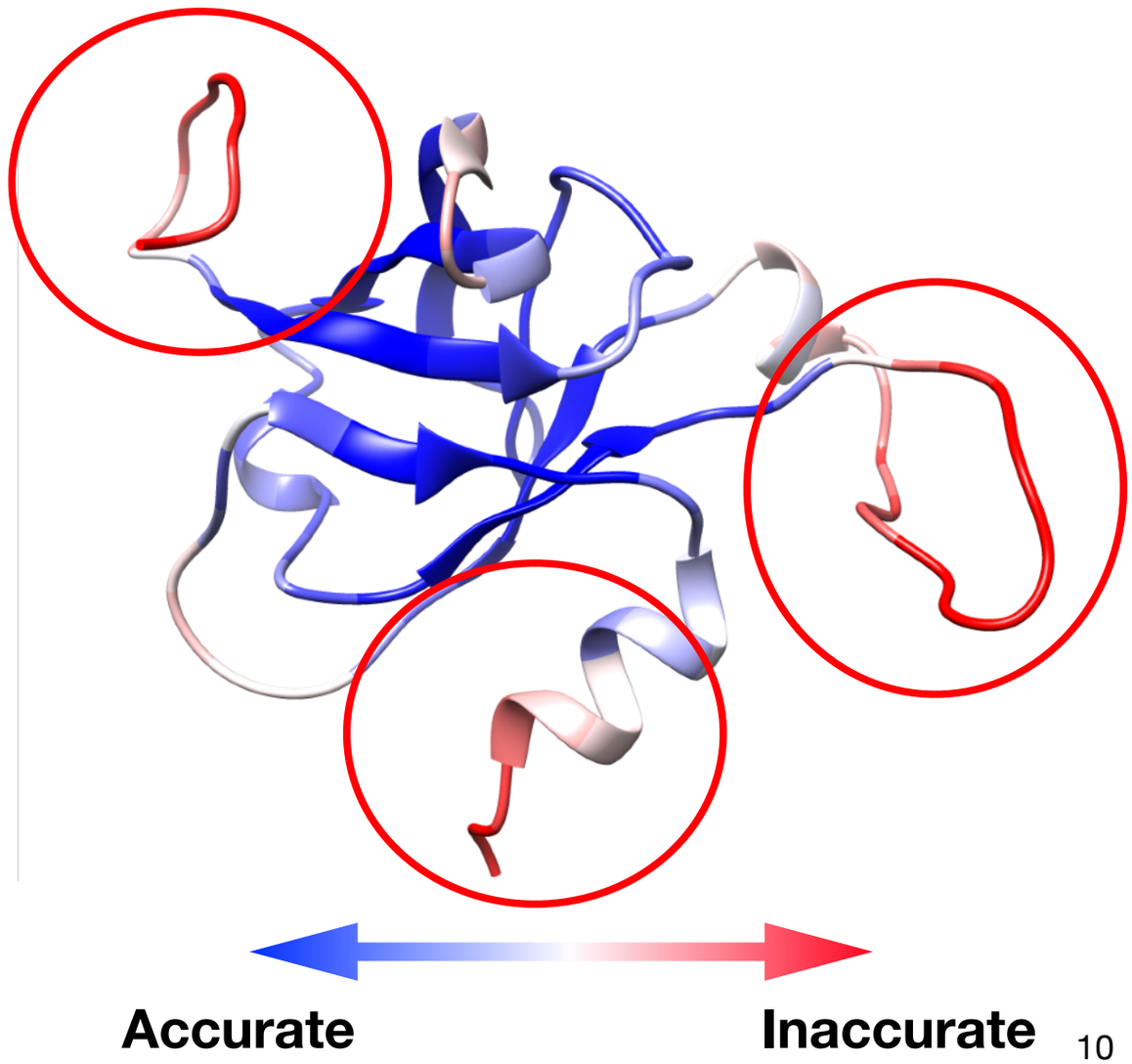
	Initial	Refined
MolP	2.00	0.91
Clash	23.27	0.11
Rotamer outlier	3.6%	1.1%
Rama outlier	2.0%	1.1%

Michael Feig, *J. Chem. Inf. Model.* (**2016**) 56(7): 1304-1312.

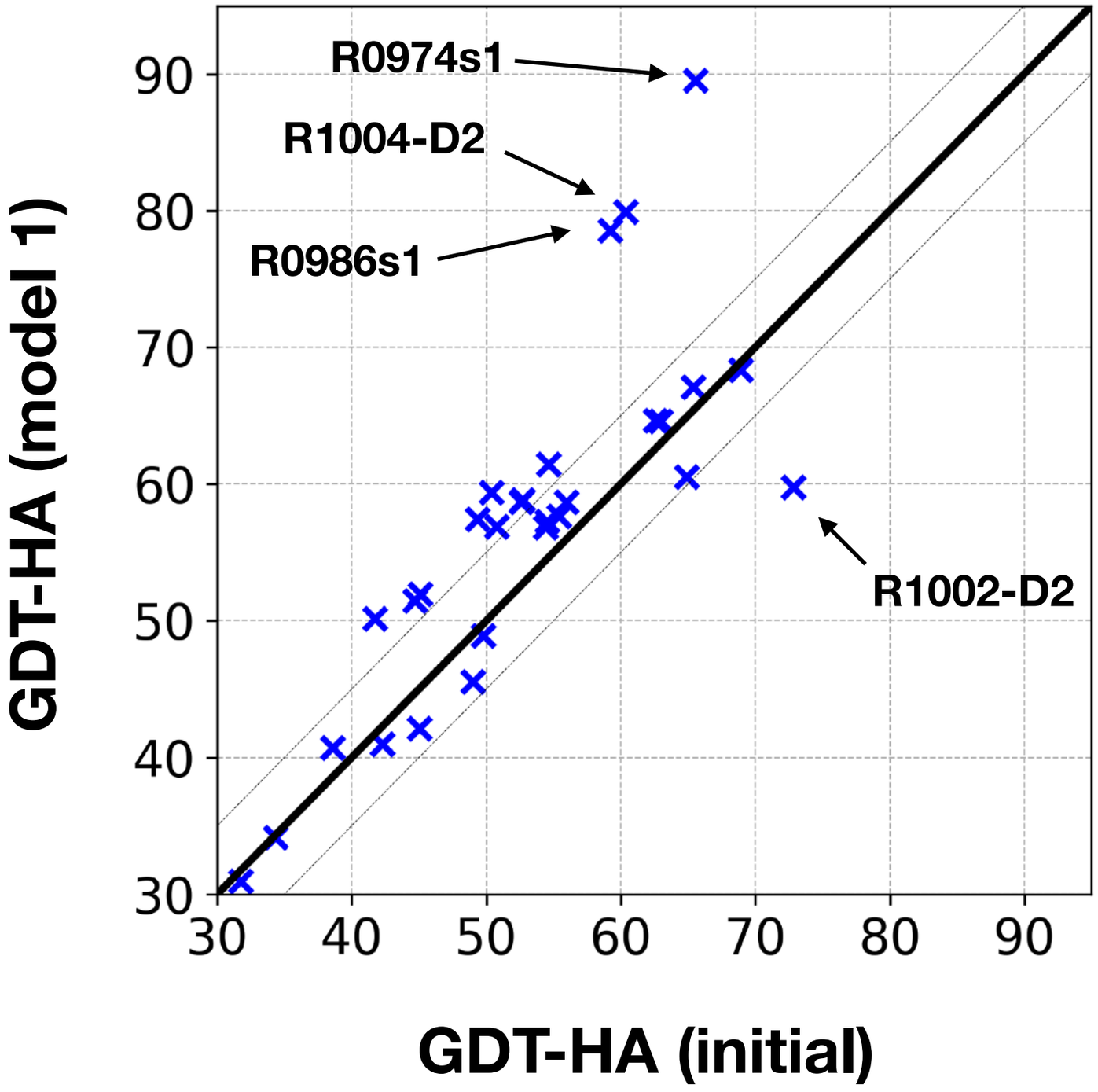
Error estimations

- Running 2 x 5 ns MD simulations without restraint for a model
- Residue-wise **RMSF** is highly correlated with residue-wise error!





CASP13 results with the current protocol

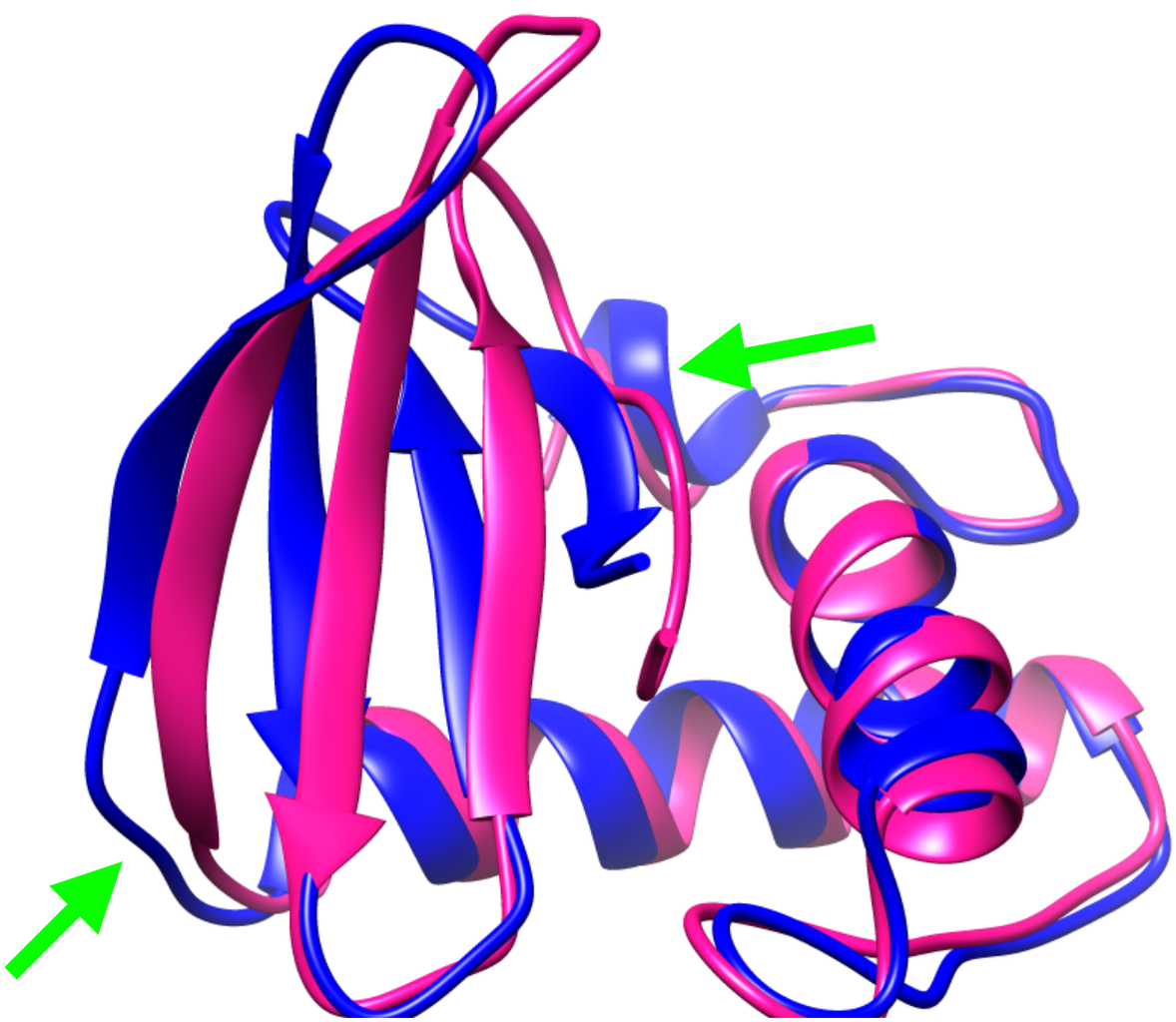


20/29 (69%) improved

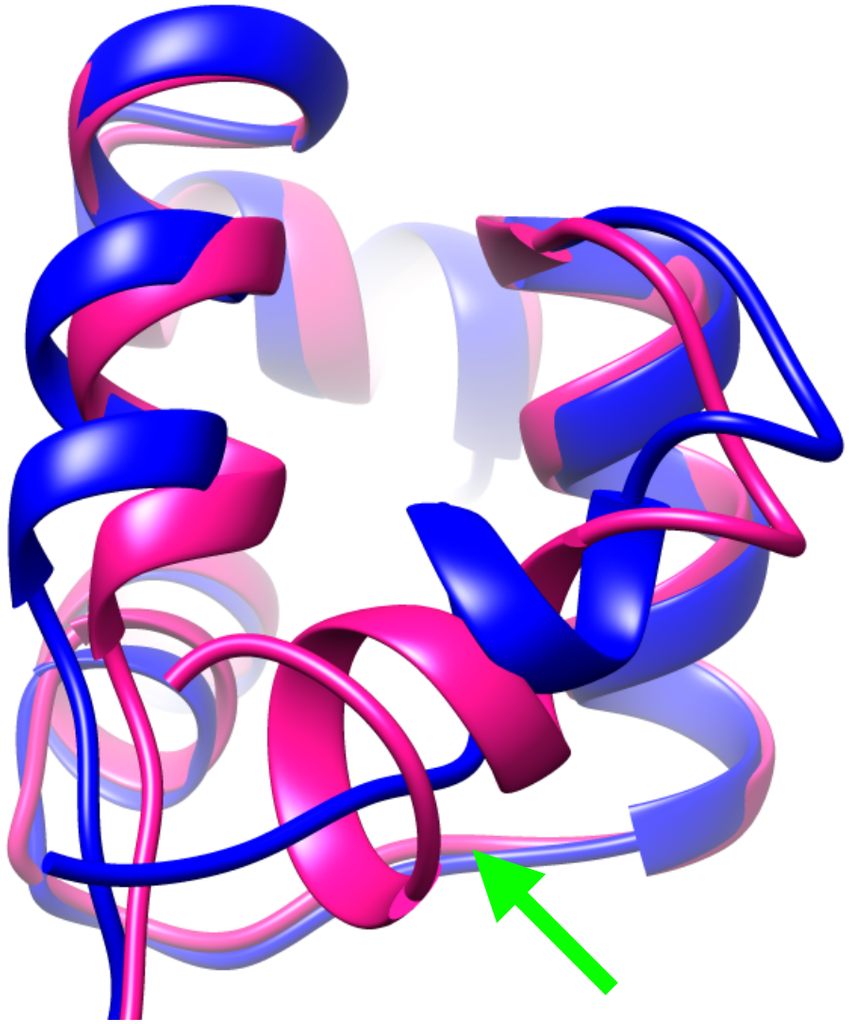
ΔGDT-HA = +3.99

What are improved?

Magenta: Initial
Blue: Refined



R0986s1
GDT-HA: 59.2 → 78.5 (+19.3)
RMSD: 1.8 Å → 1.0 Å (-0.8 Å)

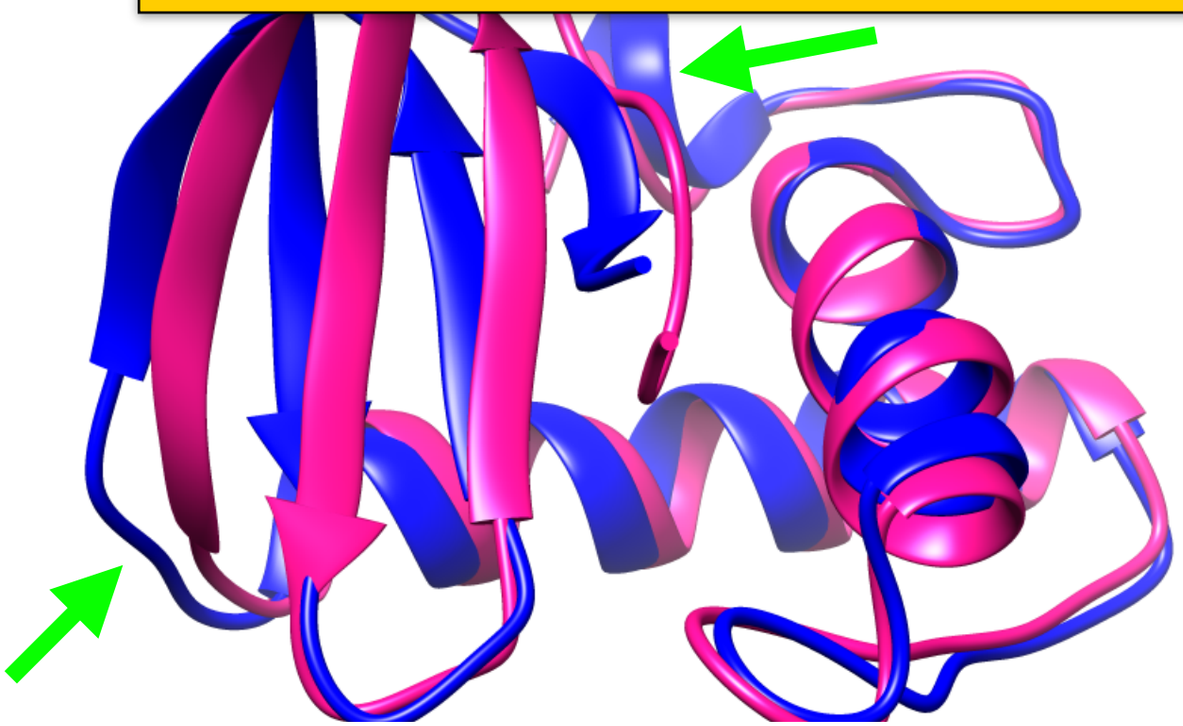


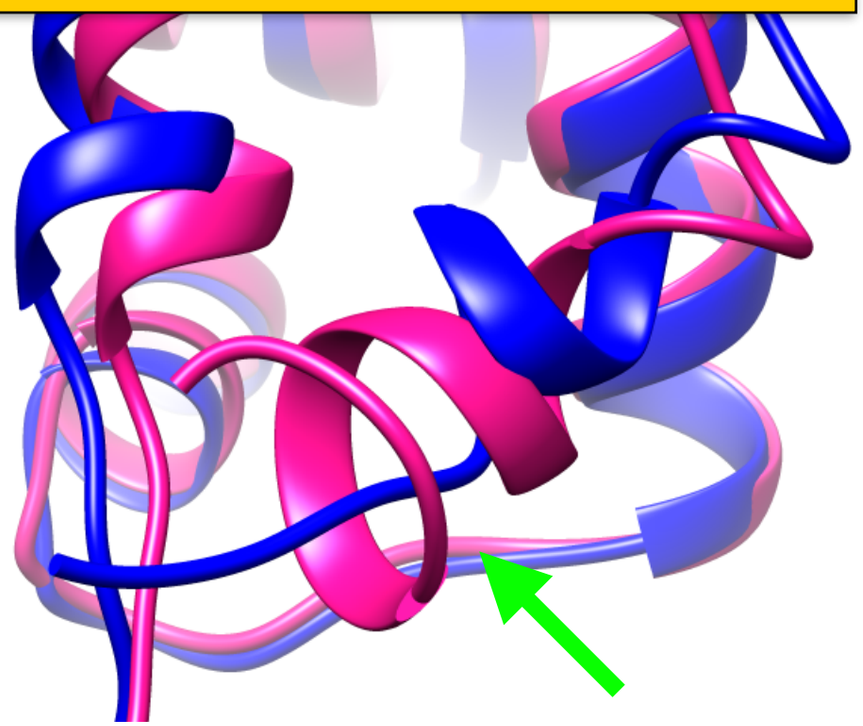
R0974s1
GDT-HA: 65.6 → 89.5 (+23.9)
RMSD: 2.2 Å → 0.8 Å (-1.4 Å)

What are improved?

Magenta: Initial
Blue: Refined

Secondary structure rearrangement
Partial folding/unfolding of structure elements



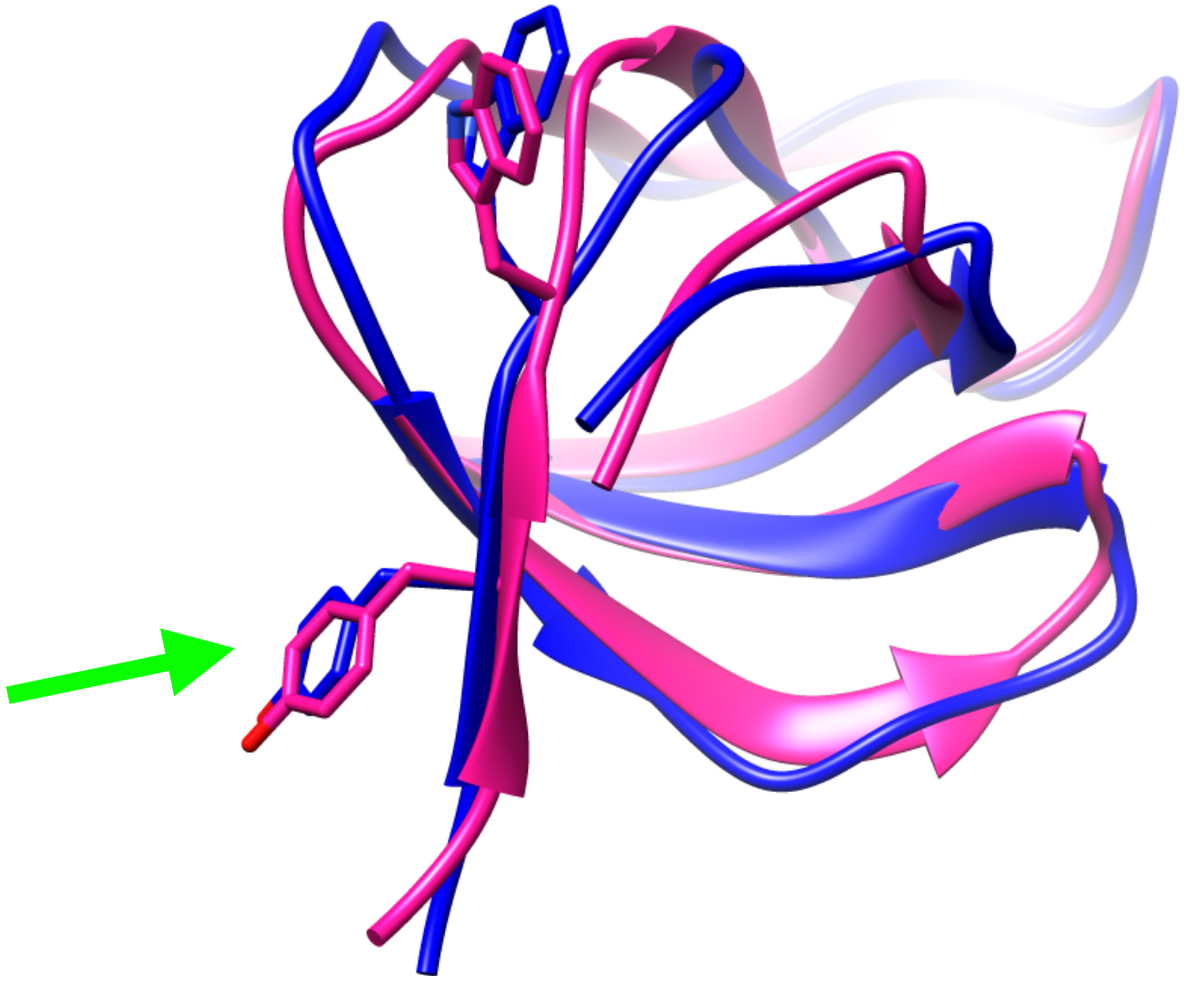


R0986s1
GDT-HA: 59.2 → 78.5 (+19.3)
RMSD: 1.8 Å → 1.0 Å (-0.8 Å)

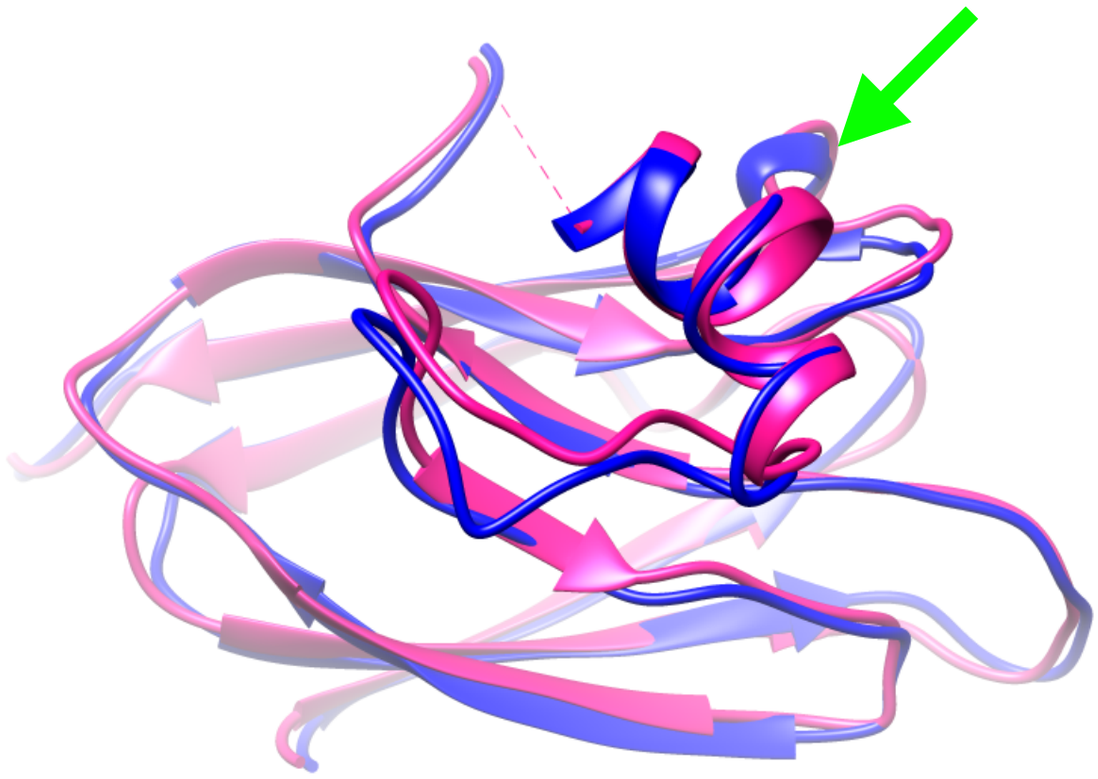
R0974s1
GDT-HA: 65.6 → 89.5 (+23.9)
RMSD: 2.2 Å → 0.8 Å (-1.4 Å)

What went wrong?

Magenta: Initial
Blue: Refined



R1002-D2
GDT-HA: 72.9 → 59.8 (-13.1)
RMSD: 1.9 Å → 2.3 Å (+0.4 Å)

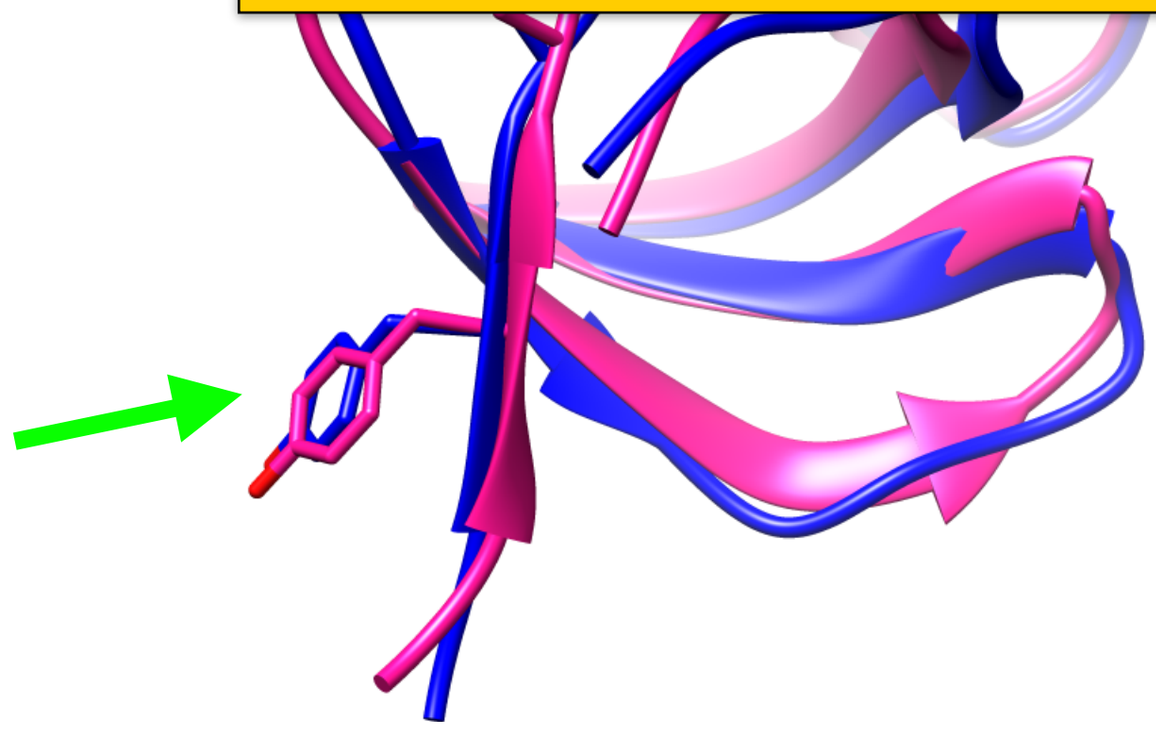


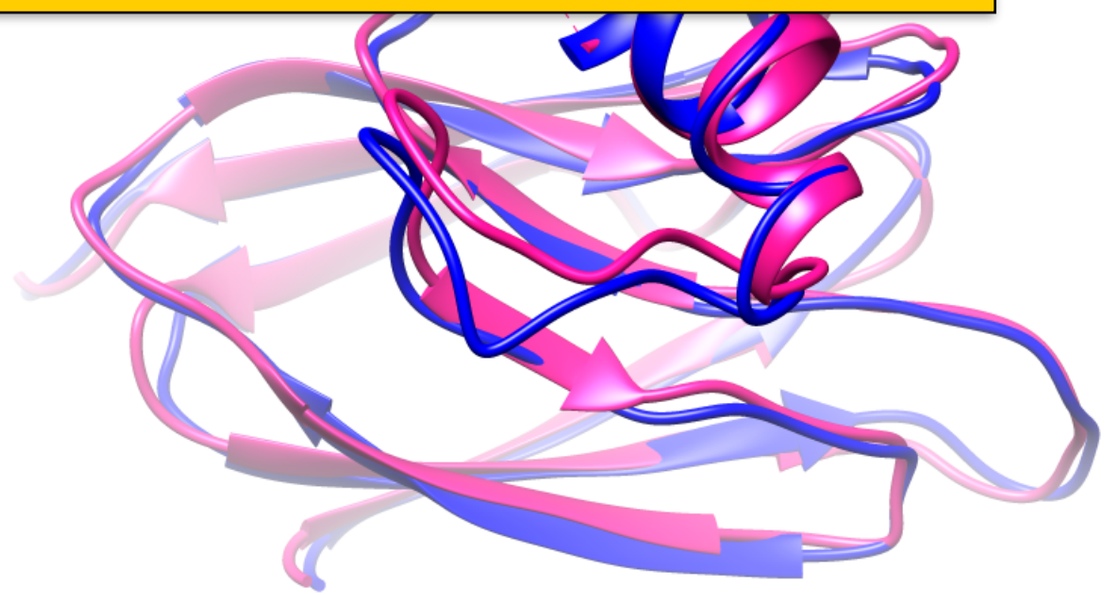
R0949
GDT-HA: 49.0 → 45.5 (-3.5)
RMSD: 7.0 Å → 7.1 Å (+0.1 Å)

What went wrong?

Magenta: Initial
Blue: Refined

Incorrect sequence alignment
Too far from the native structure





R1002-D2
GDT-HA: 72.9 → 59.8 (-13.1)
RMSD: 1.9 Å → 2.3 Å (+0.4 Å)

R0949
GDT-HA: 49.0 → 45.5 (-3.5)
RMSD: 7.0 Å → 7.1 Å (+0.1 Å)

Summary

- Purely **physics**-based method

- Successes
  - **Smaller** targets via **iterative** MD sampling
  - **Partial folding, unfolding, and rearrangement** allowed by **flat-bottom** harmonic restraints
- Failures
  - **Incorrect topology** (fixing sequence alignment errors) because of **insufficient** sampling

Acknowledgements

- Feig Computational Biophysics Lab MICHIGAN STATE UNIVERSITY

- CASP participants, assessors, target providers, and organizers

- NIH fundings (R01 GM084953 and R35 GM126948)
- XSEDE computing resources (TG-MCB090003)