Modeling Affinity in Lead Optimization

Christopher Bayly

CUP XXII 2023 March

CADD Contribution to Lead Opt

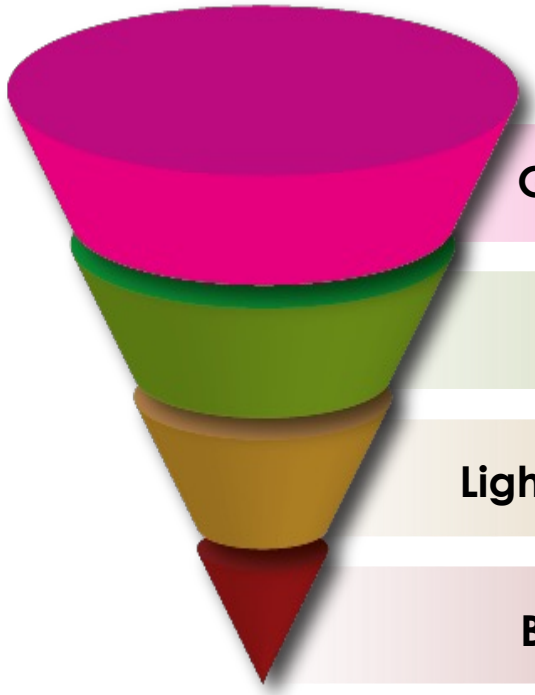
Evaluate

Filter

Generate

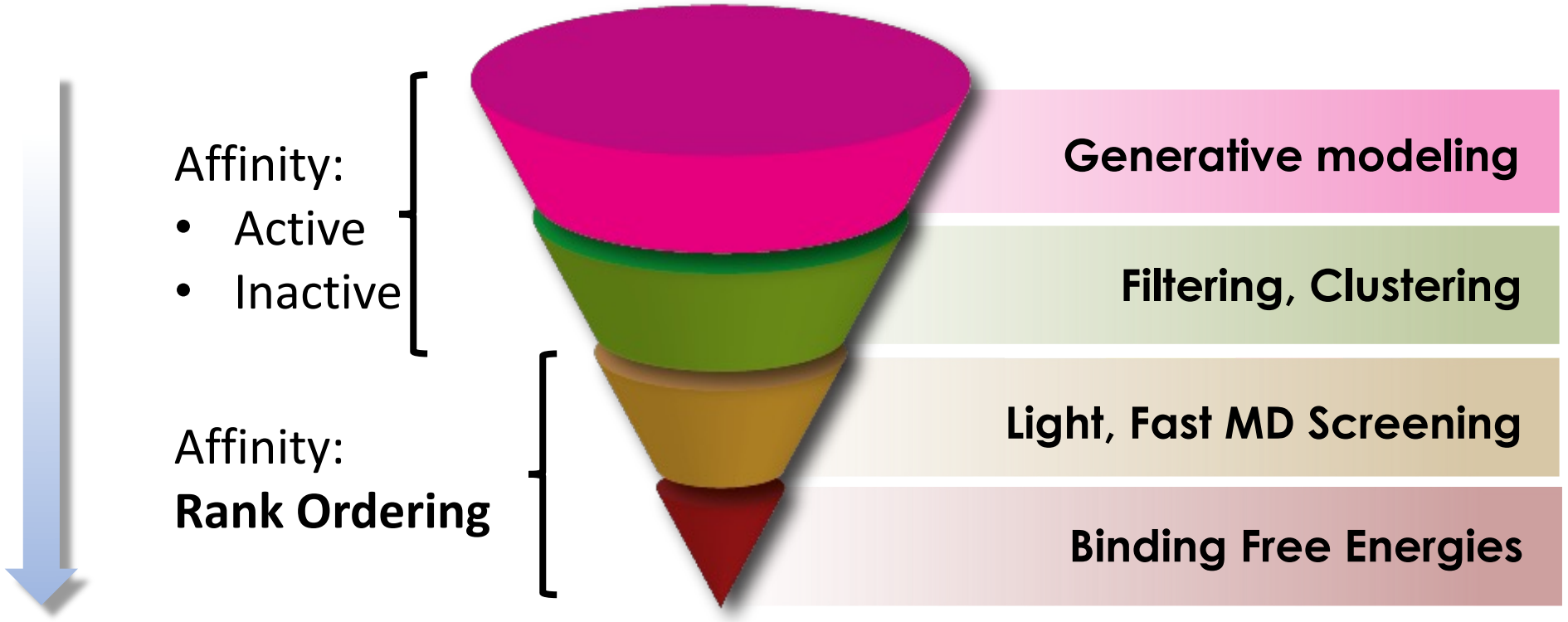
Hypothesize

Prioritize for



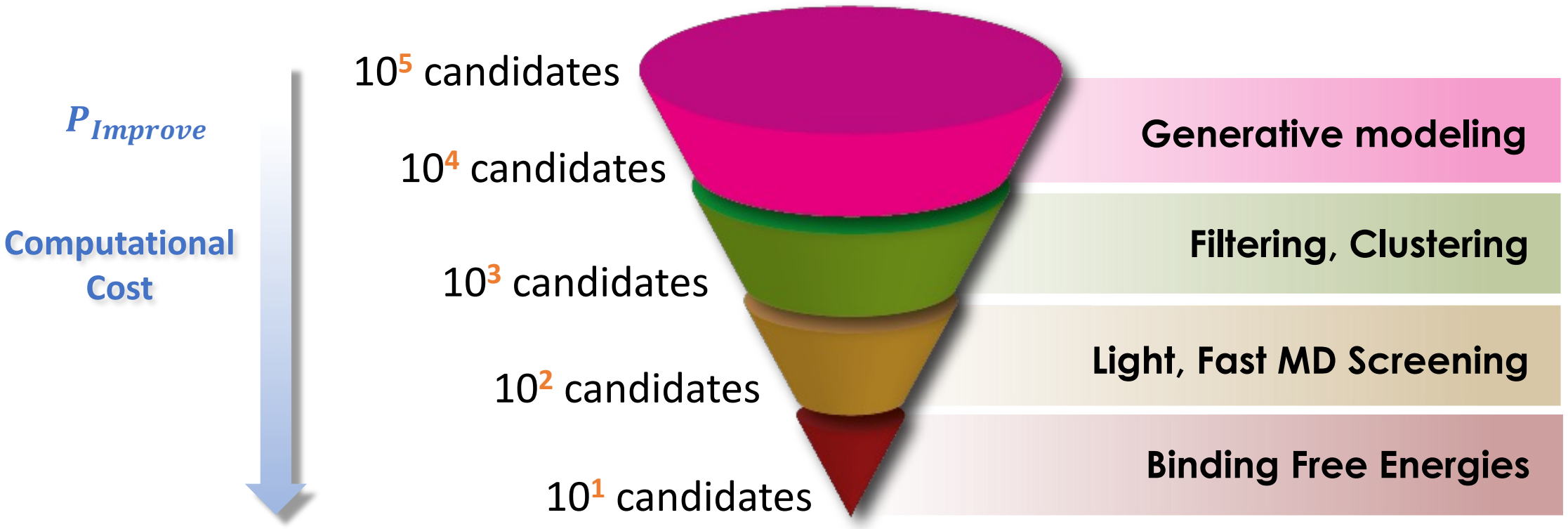
Structure-Based Design Iteration

$P_{Improve}$



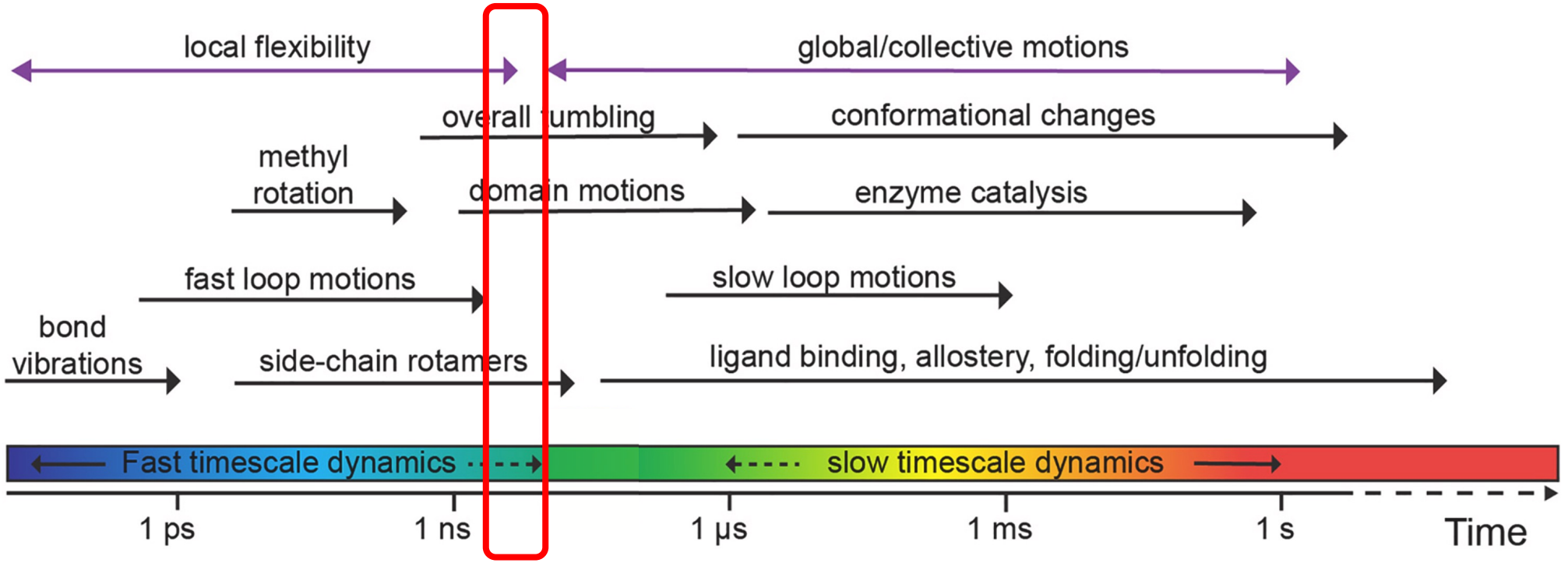
- Improved Affinity scoring required later in design Iteration

Structure-Based Design Iteration



- 1 Structure-Based Design Iteration: 1 – 2 weeks
- High throughput, efficiency needed for MD stages

Affinity Calculation Timescales

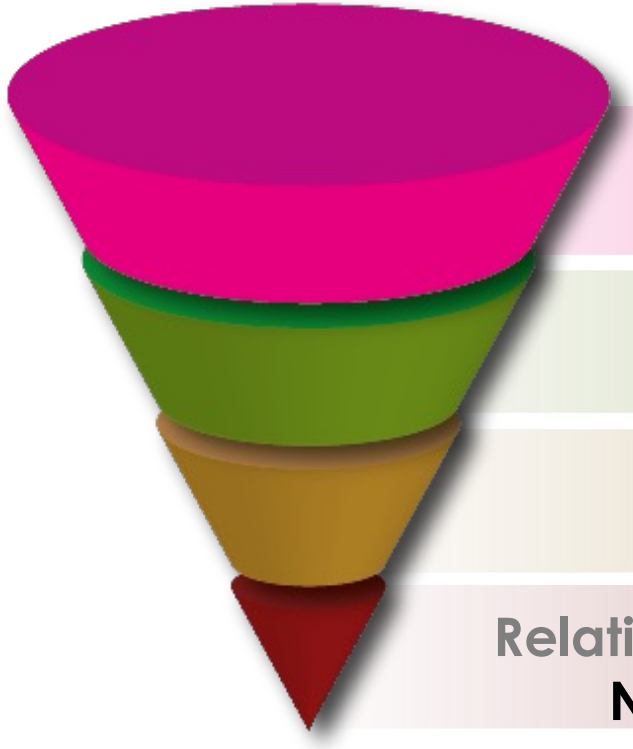


- Sample only short timescale phenomena:
  - Binding thermodynamics, not kinetics
  - Scientific constraint: **assumption of protein similarity** (often (usually?) valid)
  - High throughput, efficient computation

Structure-Based Design Iteration

Maximize
$P_{Improve}$

Minimize
Computational Cost



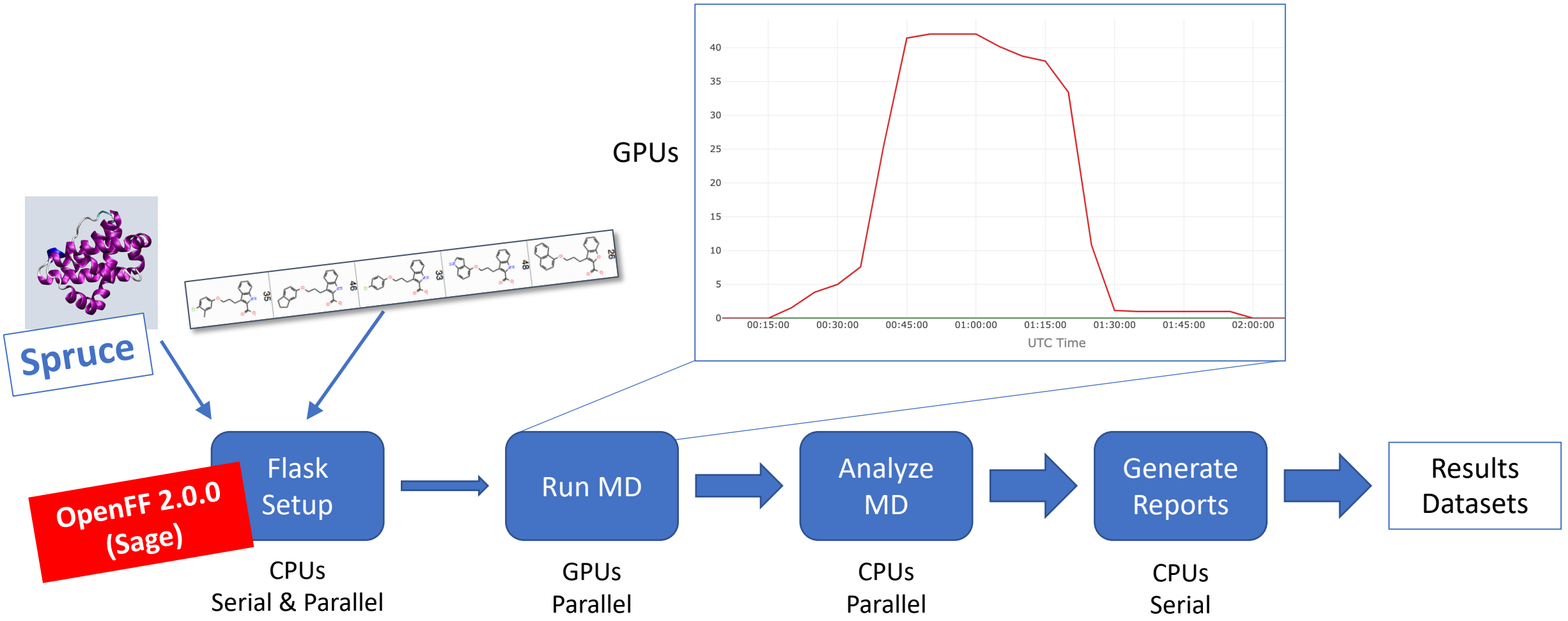
Generative modeling

Filtering, Clustering

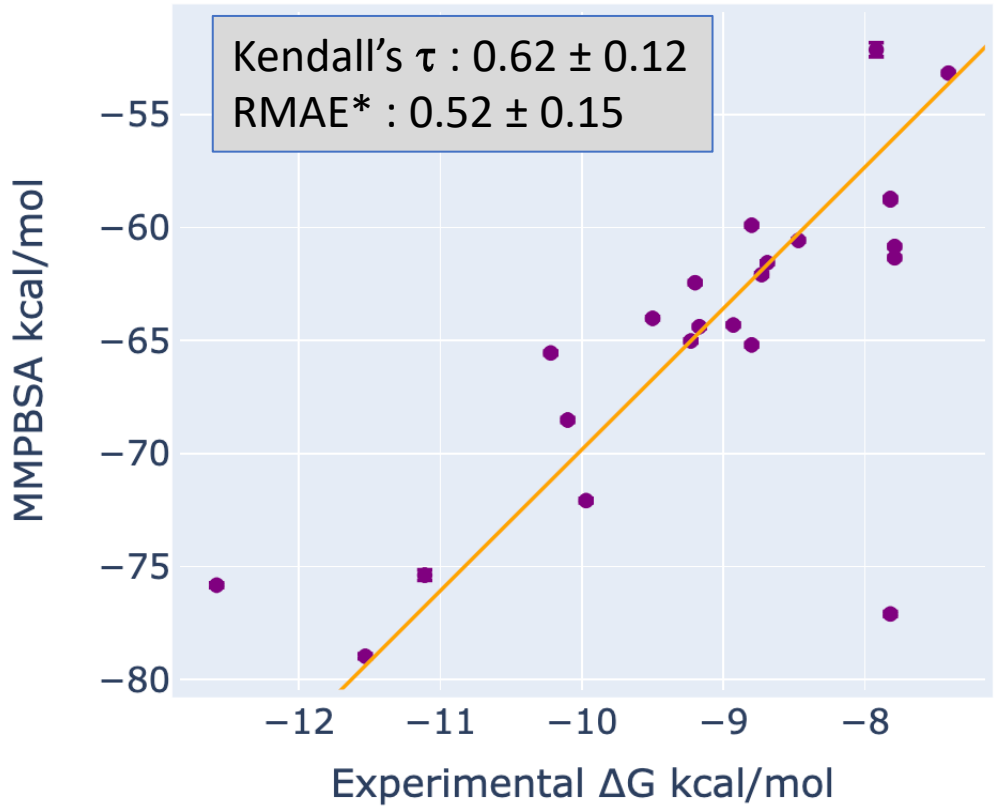
Light, Fast MD Screening:
MMPBSA, BintScore

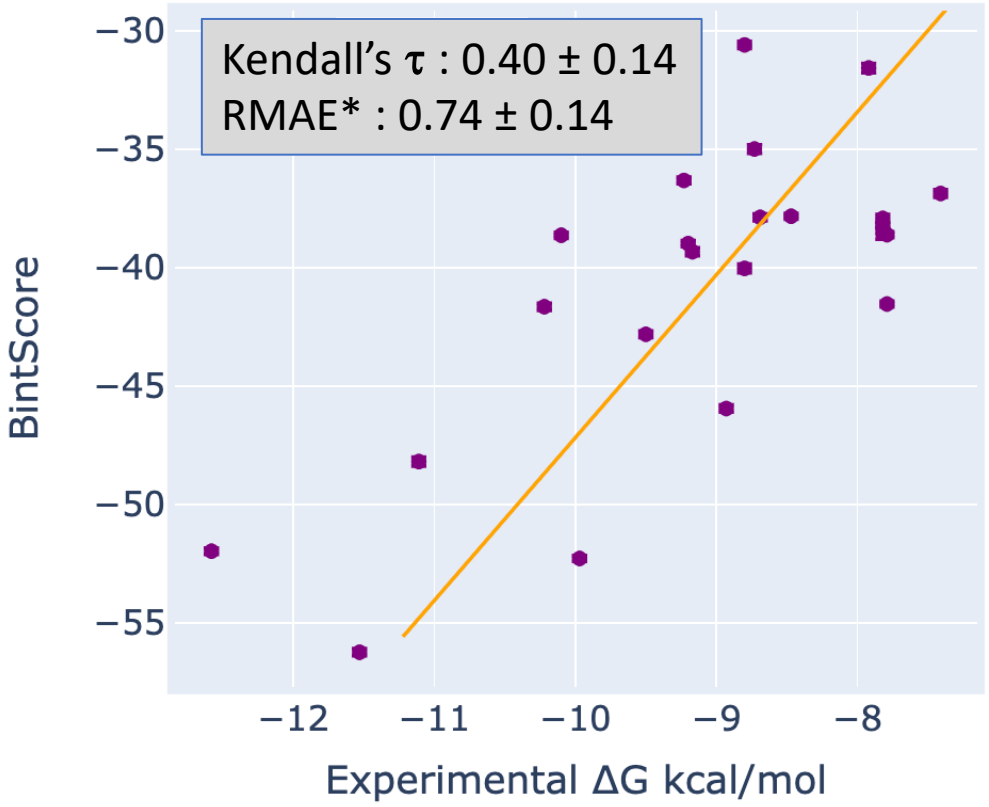
Relative Binding Free Energies:
Non-Equilibrium Switching

Short-Trajectory MD (STMD)



Affinity Models with <MMPBSA> and <BintScore>

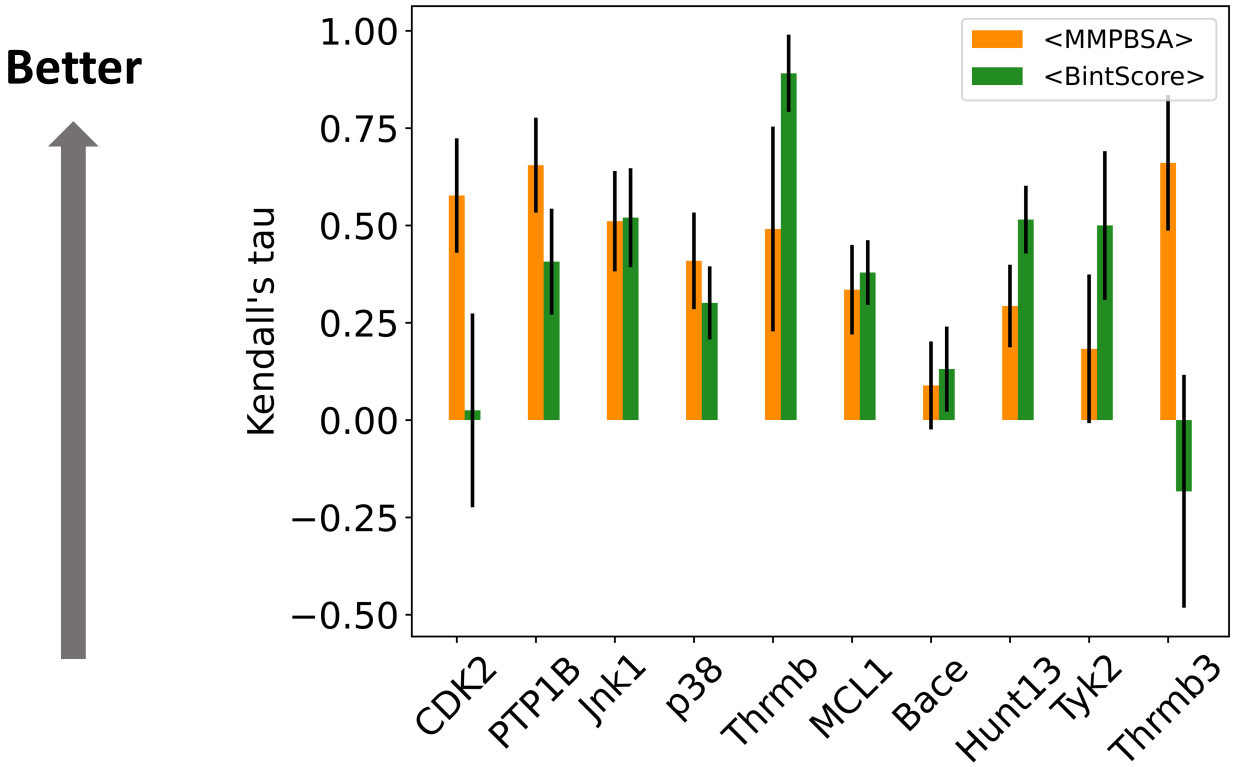




- <MMPBSA> and <BintScore> can produce usable models of affinity

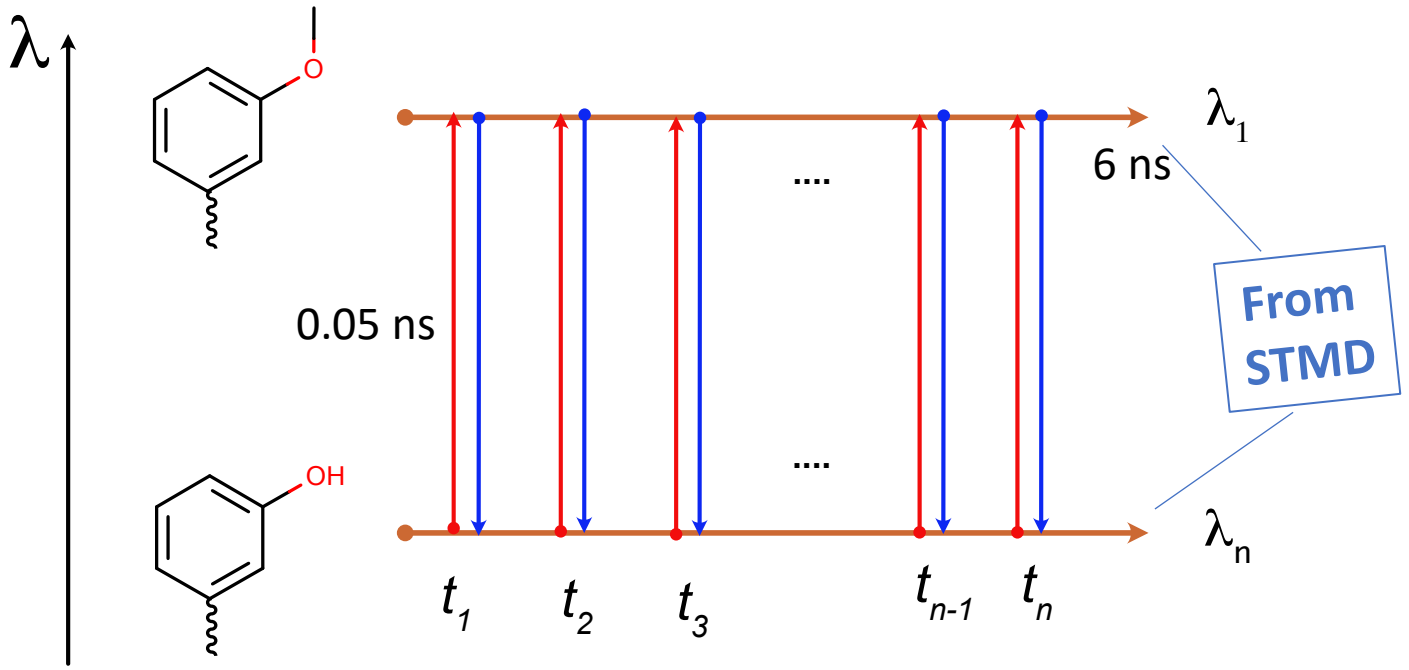
RMAE: Relative MAE = Mean Absolute Error normalized by Mean Absolute Deviation of Experimental Values

Modeling ΔG with <MMPBSA> and <BintScore>



- Robust Linear Models of <MMPBSA> and <BintScore> show usable predictions for some targets.

Recent RBFE Method: Non-Equilibrium Switching



Bert L. De Groot

Vytautas Gapsys

Gapsys et al., *Chem. Sci.* (2020) **11**, 1140-1152

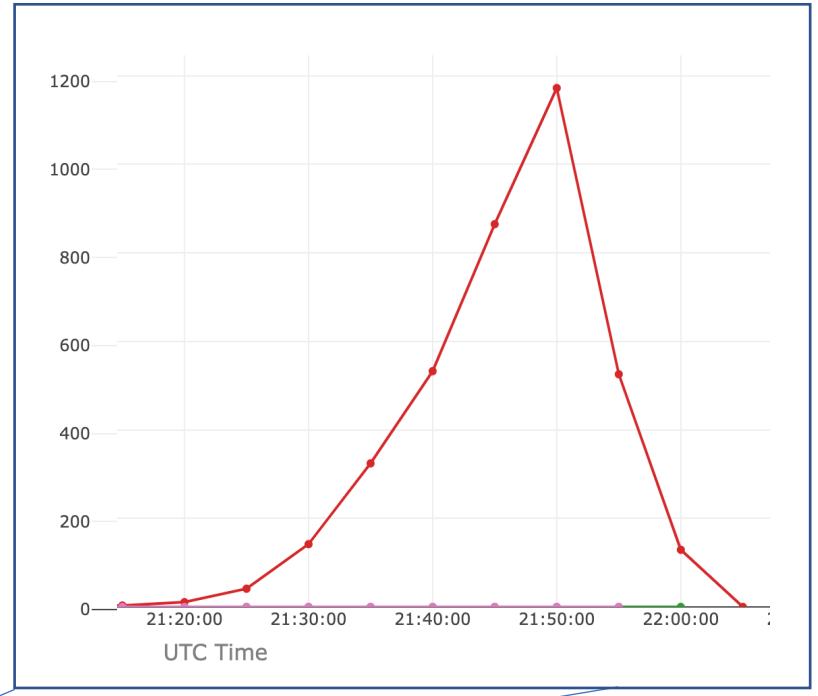
- Endpoint MD equilibration (~6ns) once per ligand (from STMD)
  - n= 80 independent starting points t_n selected for switching
  - λ is quickly (0.05ns) and continuously switched from A to B
  - Expense: 4ns per edge
- 320 independent short simulations ***per edge***

orion

OpenEye SCIENTIFIC

CADENCE MOLECULAR SCIENCES

NES Floe in Orion



Equilibrium MD runs from STMD

Map of Ligand transformations

Select starting points for Switches

CPUs
Serial

Run NES Switches

GPUs
Parallel

Analyze Switches

CPUs
Serial

Generate Report

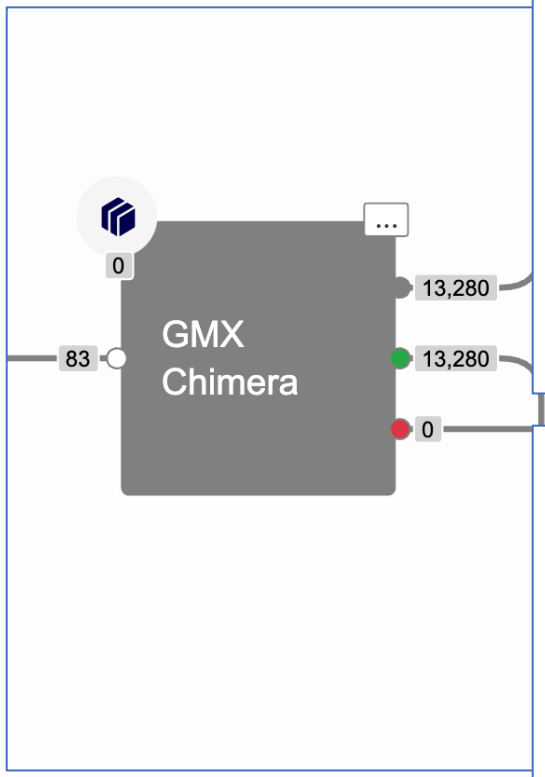
CPUs
Serial

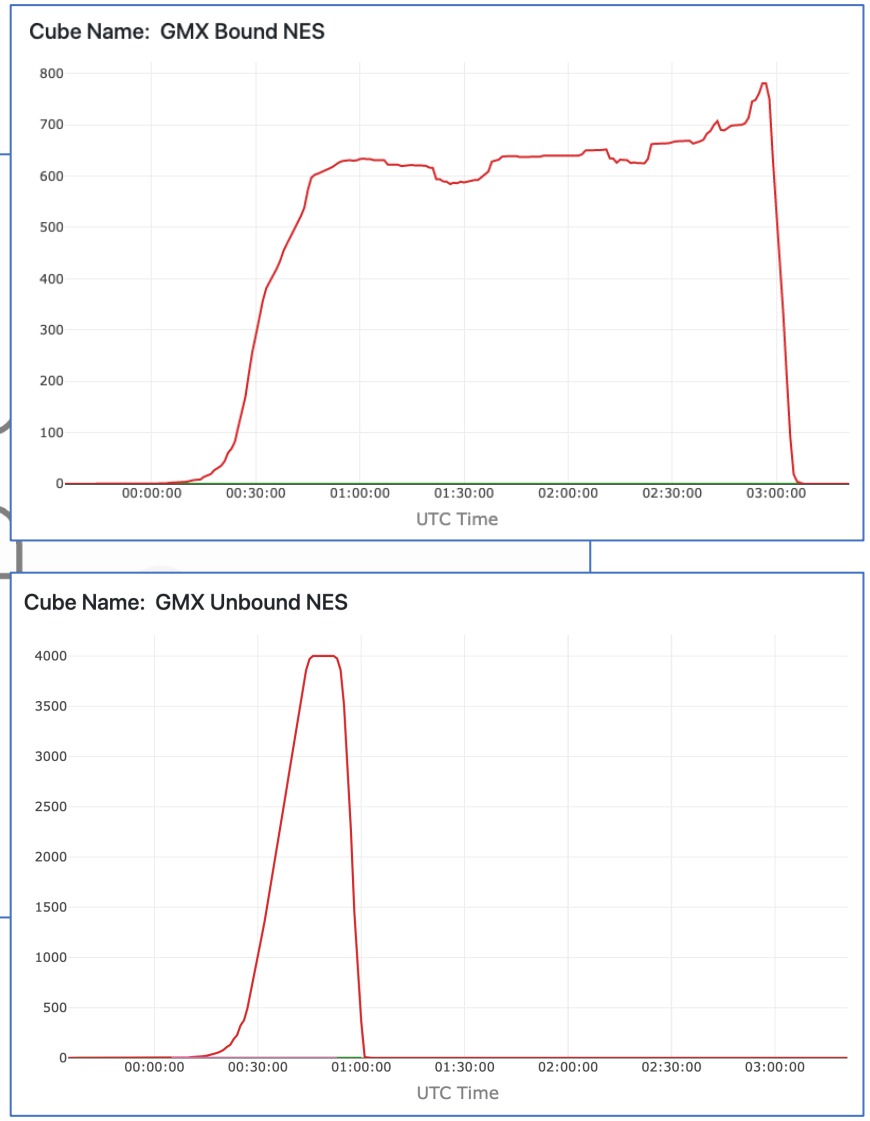
Results Datasets

GROMACS

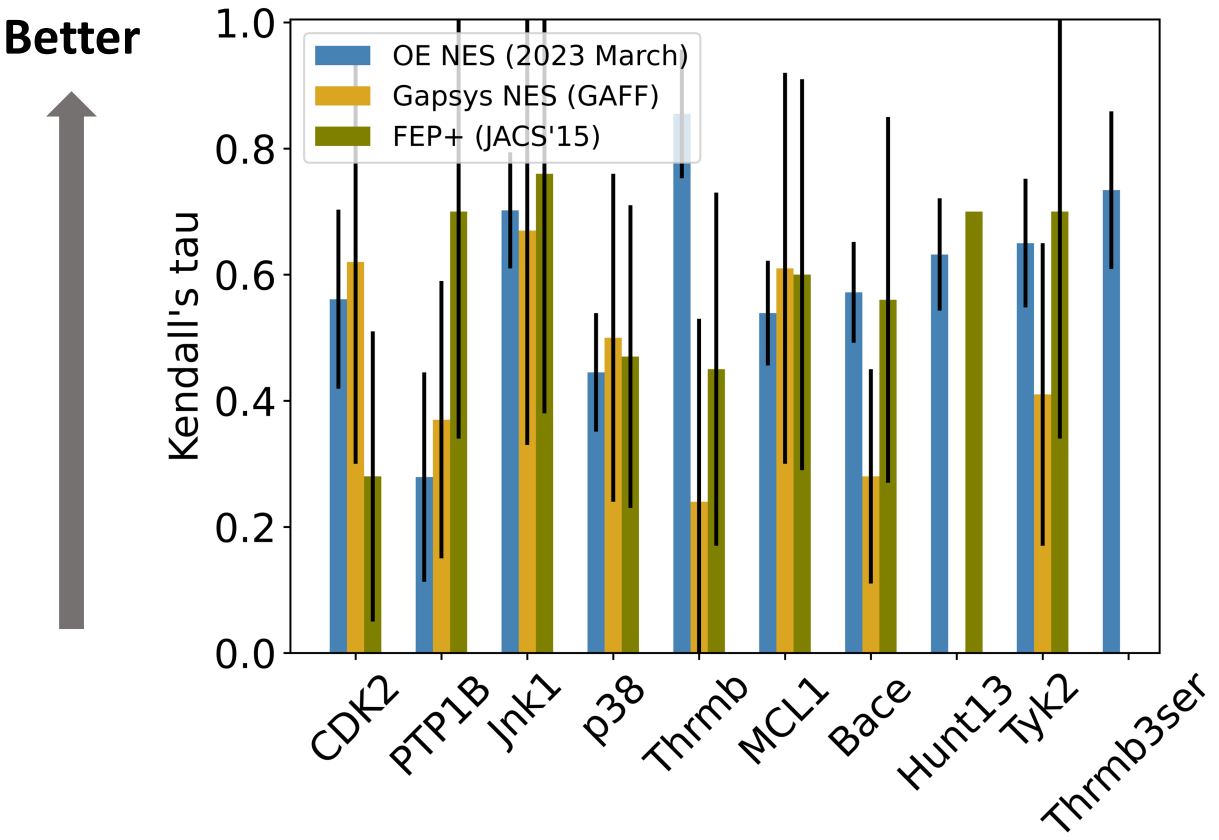
Inside a large NES run

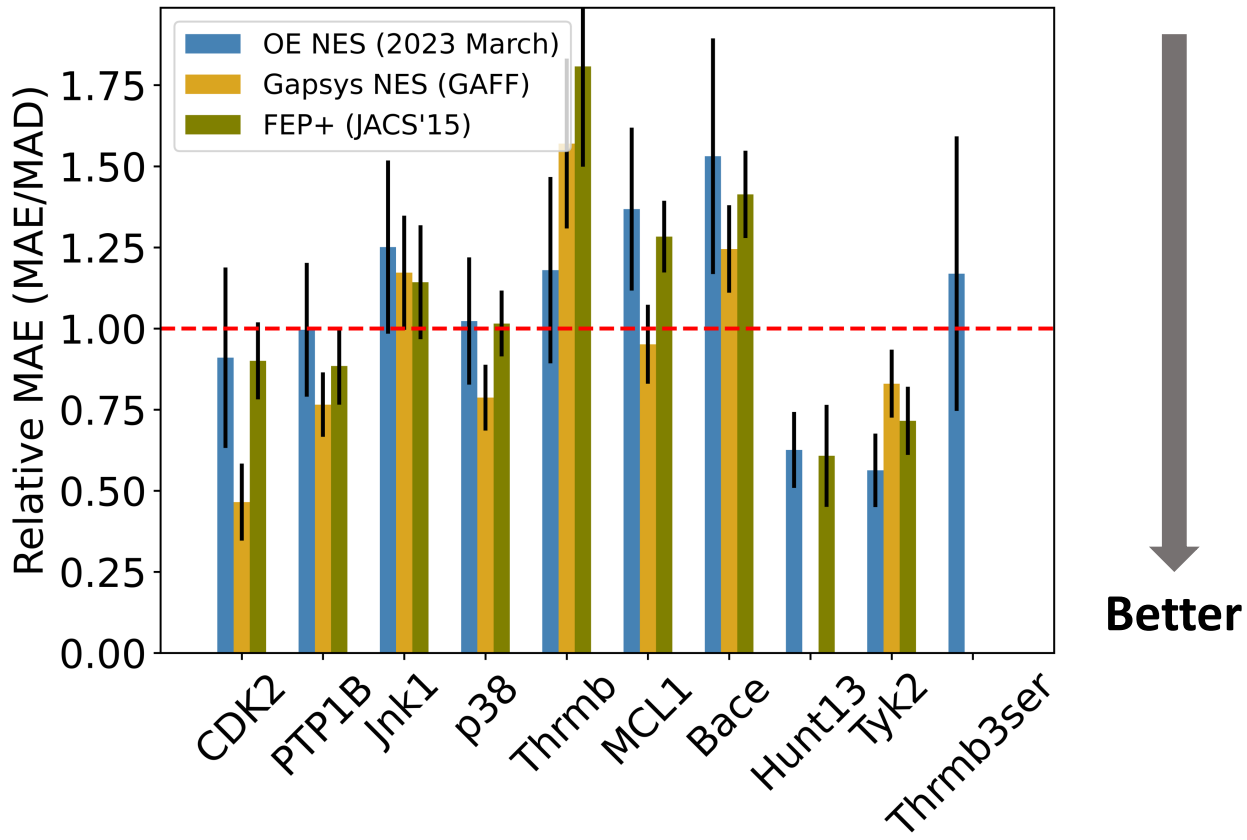
- 58 Ligands
- 83 Edges
- Parallel tasks: 13280
- **GPUs** in parallel
  - GMX **Bound**: 600+
- **CPUs** in parallel
  - GMX **Unbound**: 4000
- Runtime **< 4h**
- Cost $814
  - **< $10/edge**





Direct Predictions of ΔG : 9 Datasets





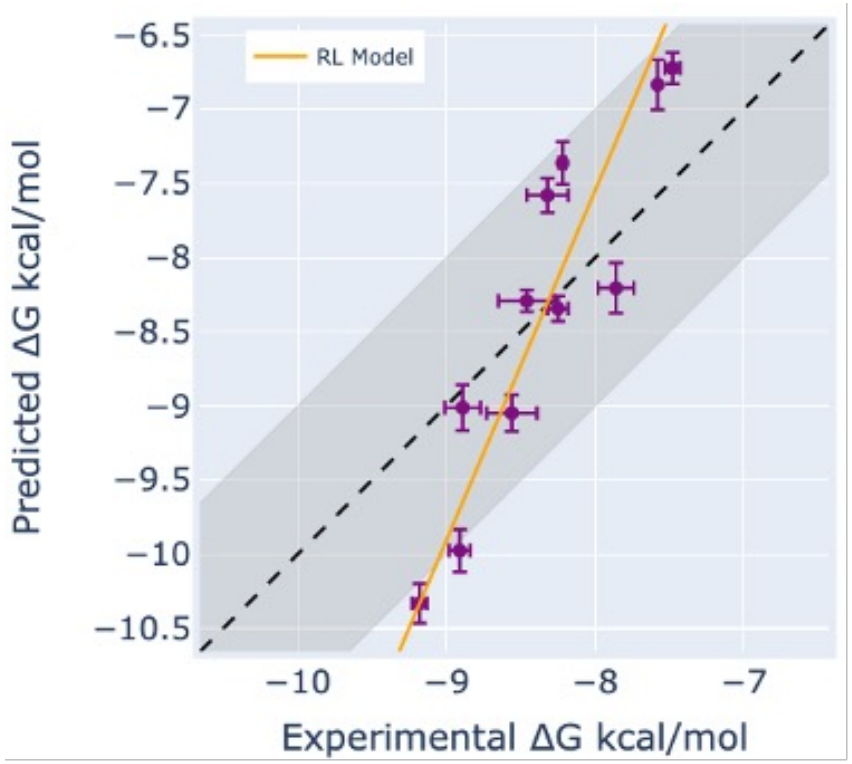
- OE NES has comparable accuracy to literature RBFE benchmarks
- RBFE performance often worse than experimental MAD

Wang et al., *JACS* 2015, **137**, 2695

Gapsys et al., *Chem. Sci.*, 2020, **11**, 1140-1152

Thrombin

Ligands	**11**
Edges	16
NES Cost	$163.51
Run time	2 h 1 min

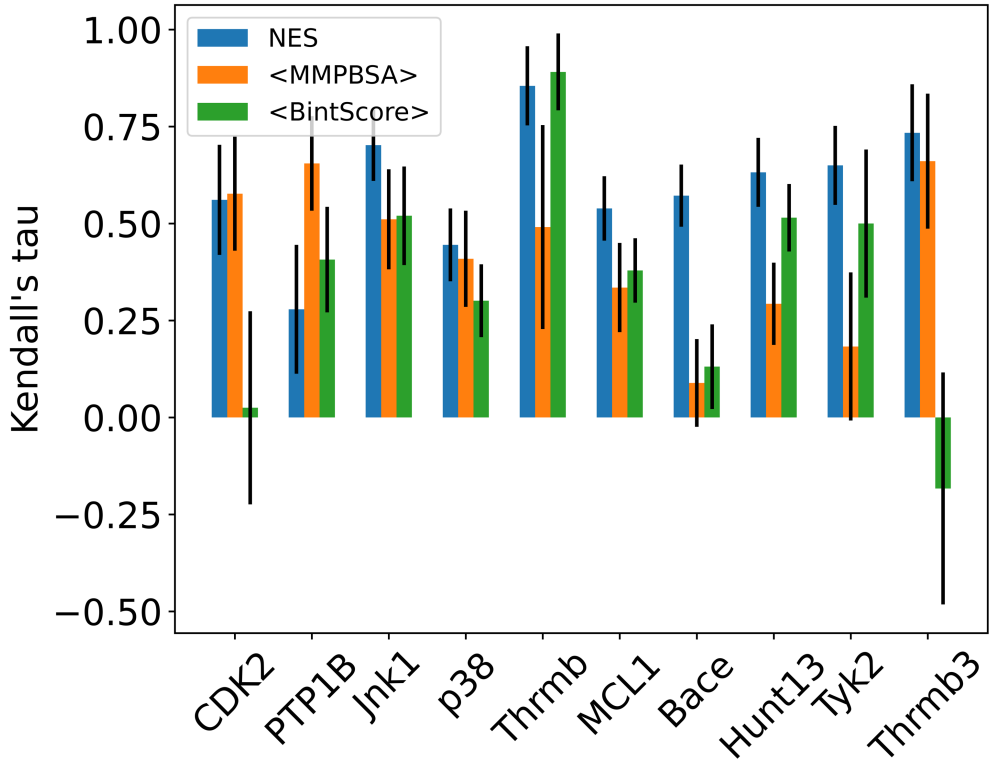


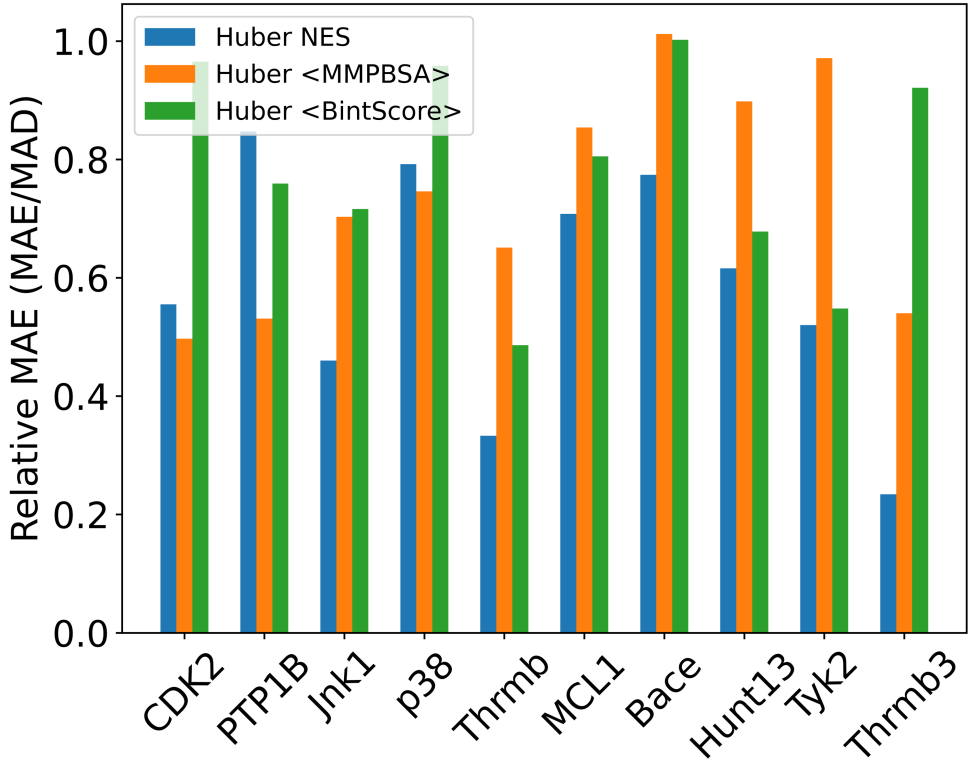
Pearson's r^2	0.824 ± 0.114
Kendall's τ	0.818 ± 0.127
MAE[a]	0.594 ± 0.113
RMAE[b]	1.412 ± 0.381

[a]Mean Absolute Error in kcal/mol.
[b]MAE divided by the Mean Absolute Deviation of Experimental ΔG.

Modeling ΔG with <MMPBSA>, <BintScore>, and ΔG_NES

Better



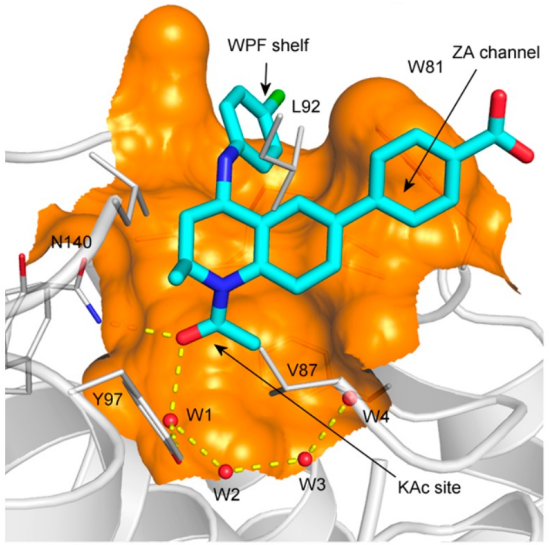


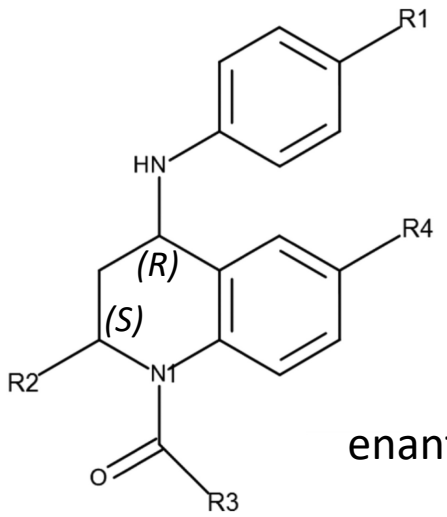
Better

- Huber RL Models of ΔG_NES are better than <MMPBSA> ,<BintScore> in **aggregate** performance.
  - Substantial variation between targets.

Next for NES…

BRD4: N-terminal Bromodomain





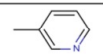
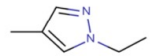
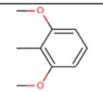
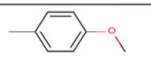
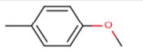
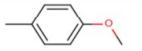
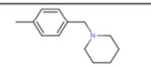
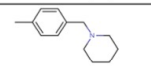
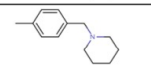
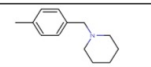
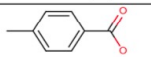
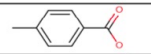
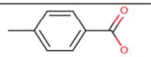
Wan et al., *J. Chem. Theory Comput.* 2017, 13, 784−795

neutral

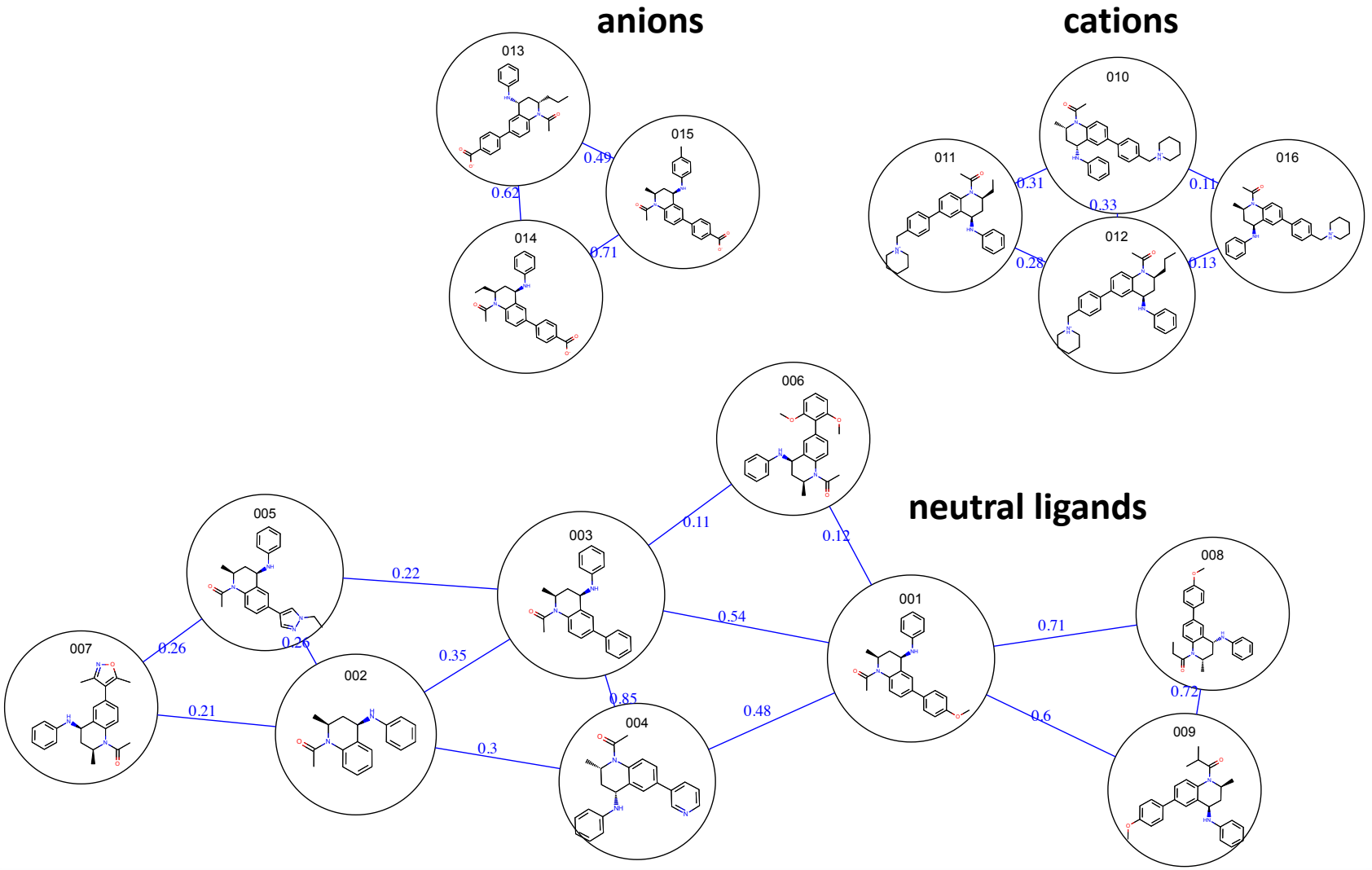
enantiomers

cations

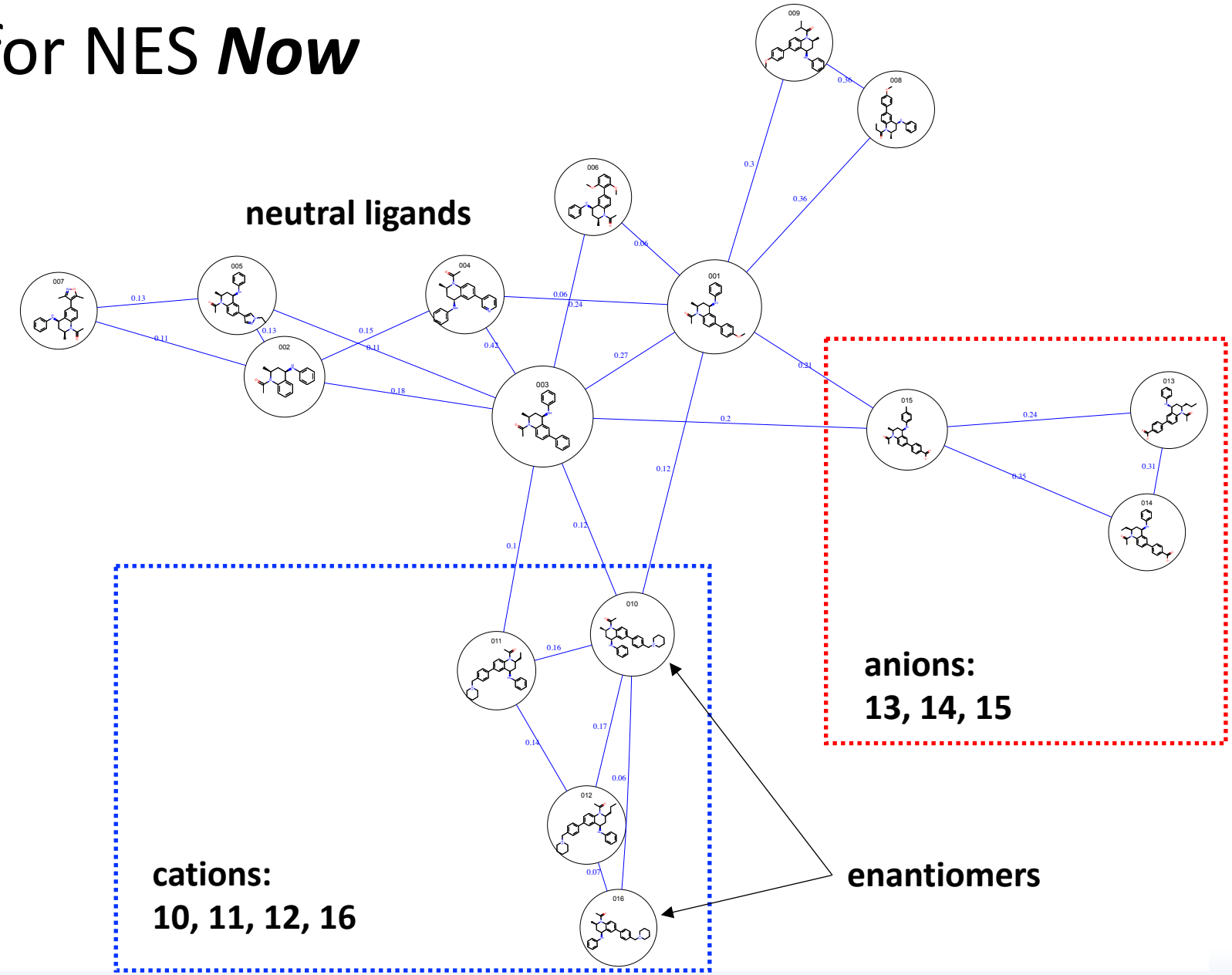
anions

Compound[a]	R1	R2	R3	R4	pIC_{50}[b]	pIC_{50} n	ΔG[b]
2	H	Me	Me	H	5.6 ± 0.06	6	-7.7 ± 0.08
3	H	Me	Me		6.8 ± 0.04	2	-9.3 ± 0.05
4	H	Me	Me		6.8 ± 0.09	3	-9.3 ± 0.12
5	H	Me	Me		7.9 ± 0.03	2	-10.8 ± 0.04
6	H	Me	Me		5.6 ± 0.01	2	-7.7 ± 0.01
7	H	Me	Me		5.8 ± 0.07	4	-8.0 ± 0.10
1	H	Me	Me		7.0 ± 0.05	3	-9.6 ± 0.07
8	H	Me	Et		6.5 ± 0.01	2	-8.9 ± 0.01
9	H	Me	i-Pr		< 4.3[c]	4	> -5.9
10	H	Me	Me		7.6 ± 0.1	16	-10.4 ± 0.14
16[d]	H	Me	Me		5.4 ± 0.28	5	-7.4 ± 0.38
11	H	Et	Me		6.8 ± 0.3	4	-9.3 ± 0.41
12	H	Pr	Me		5.5 ± 0.01	4	-7.5 ± 0.01
13	H	Pr	Me		5.4 ± 0.01	2	-7.4 ± 0.01
14	H	Et	Me		6.7 ± 0.2	4	-9.2 ± 0.27
15	Cl	Me	Me		7.8 ± 0.1	44	-10.7 ± 0.14

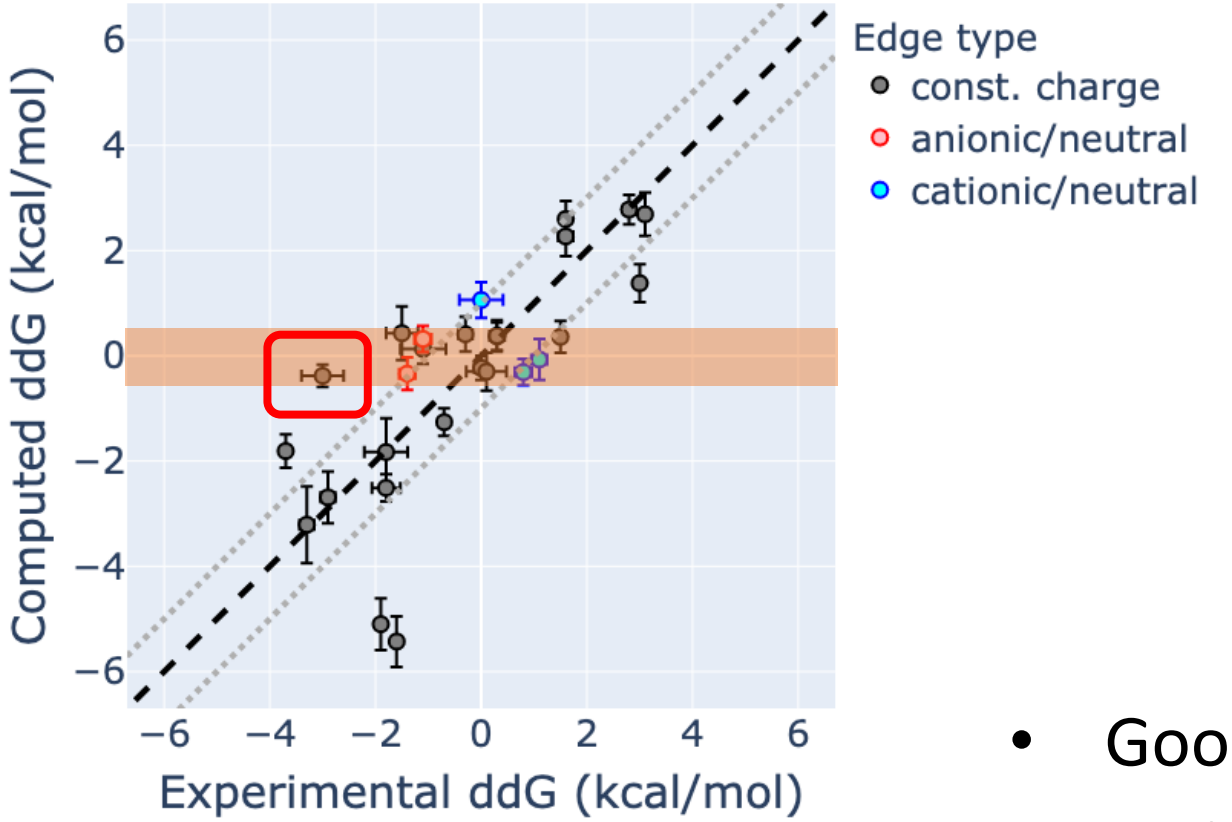
BRD4 Map for NES ***Before 2022***: No Charge-changes Allowed



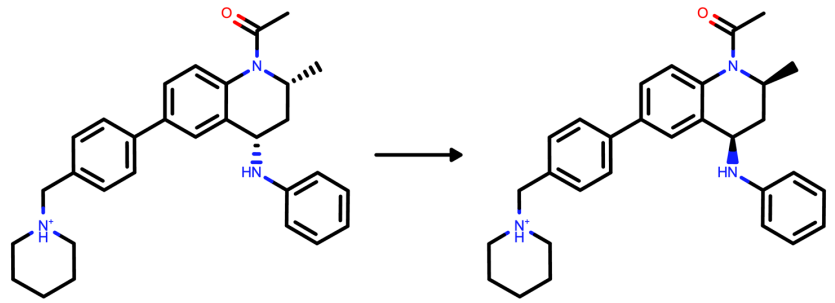
BRD4 Map for NES ***Now***



NES on BRD4: Good news, Bad news

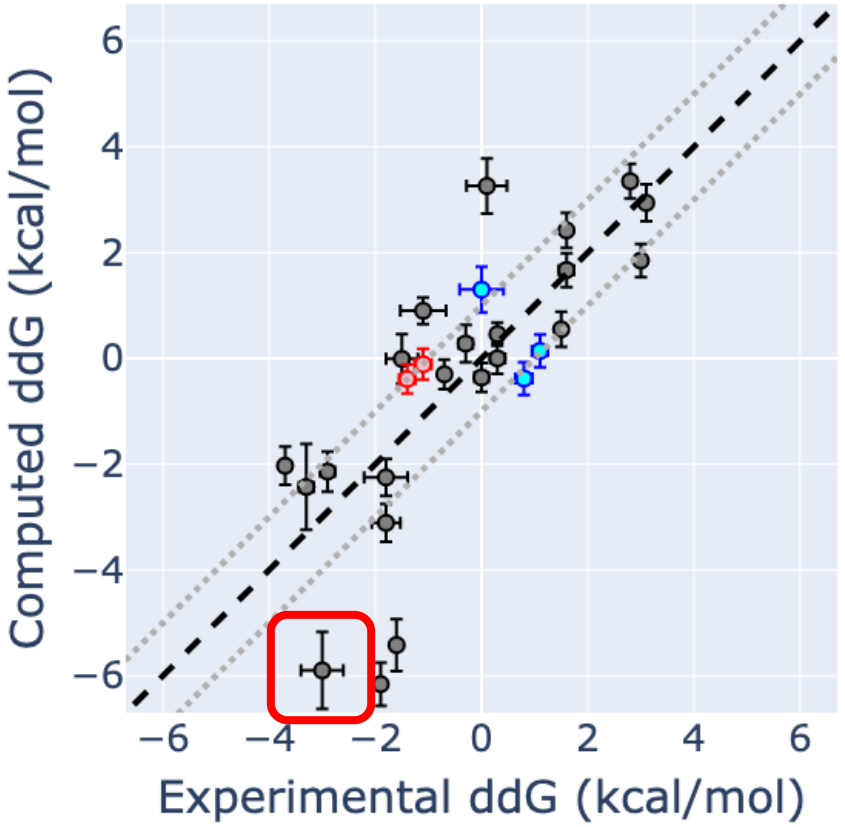


016 to 010

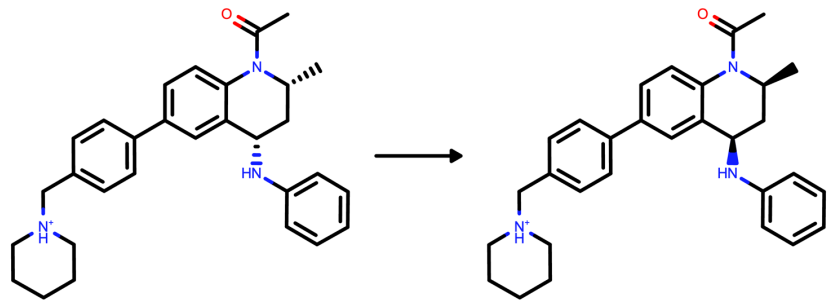


- Good: Charge-change edges handled
- Bad: Stereochemistry bug

… Stereochemistry bug fixed.



016 to 010



- Stereochemistry now correctly handled

NES: Further down the road

- Edge Map Improvements
- Implementing NES in OpenMM

New Group members:
Bing Xie
Christopher Neale

- NES treatment for trapped waters
  - MobleyLab Collaboration: Swapnil Wagle's poster tomorrow night
- Separated topologies
  - MobleyLab: Hannah Baumann's poster tomorrow night

Acknowledgements

- NES: Gaetano Calabro, Christopher Neale
- STMD, IFP: Hyesu Jang, Bing Xie
- Jack Delany (Gromacs compilation)
- Application Sciences (Gunther Stahl)
  - **Mireille Krier**, Suhani Nagpal, Adam Green, Lukas Eberlain

OpenEye CMS Orion and Toolkit Developers

- David Mobley & lab (UC Irvine)
- John Chodera & lab (MSKCC)
- Bert de Groot & Vytautas Gapsys (Max Planck Inst.)