THE UNIVERSITY of EDINBURGH

Successes and Failures of Docking and Relative Free Energy Calculations on HSP90

Antonia Mey
School of Chemistry
University of Edinburgh
antonia.mey@.ed.ac.uk

UCSD 10/03/2016

Overview

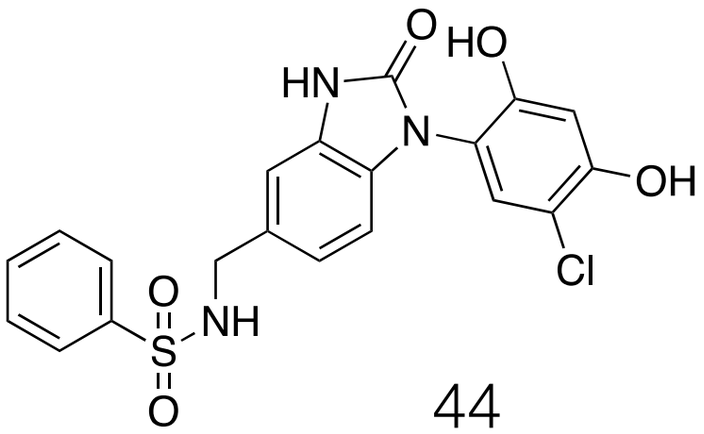
Docking — HSP90 only results

- **Clustering** of available crystal structures
- **rDock and Vina protocols** and results for pose predictions
- **Statistical analysis** of the scoring performance the whole data set.

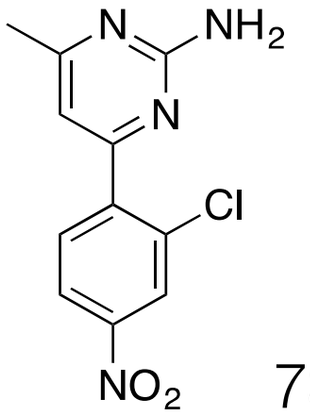
Crystal Data

PDB search for HSP90 returns 401 crystal structures and 195 structures with a resolution of < 2 Å.

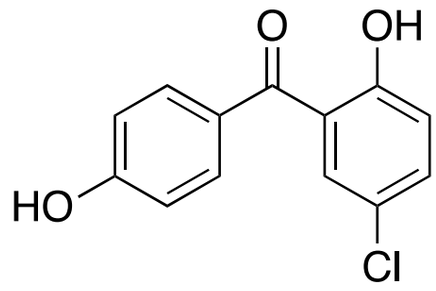
Can we exploit these structures in a way to optimally predict binding poses for the following compounds?



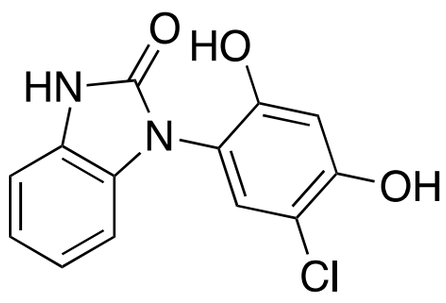
44



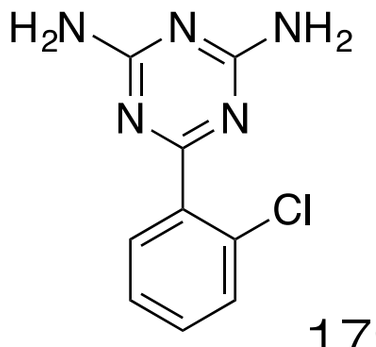
73



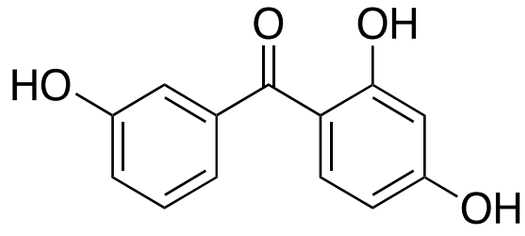
164



40



179



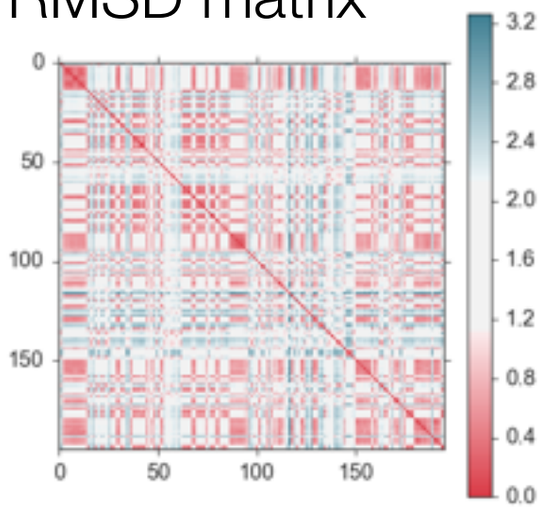
175

Spectral clustering

Aim: Group 195 structures in a small set of manageable clusters with representative structures for docking.

Idea: spectral clustering

RMSD matrix



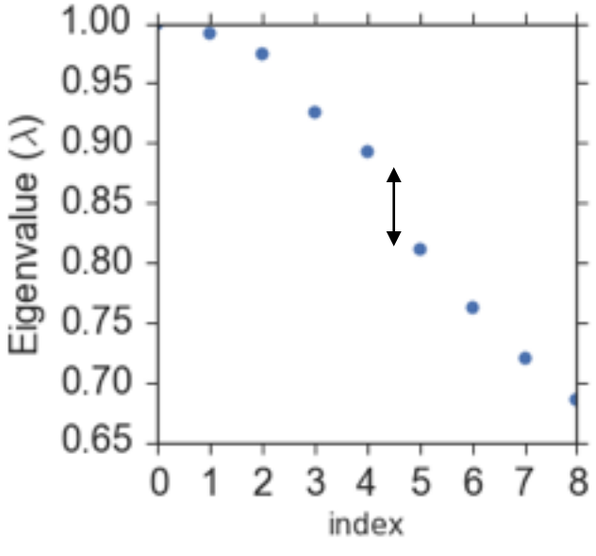
$$L_{ij} = \exp\left(\frac{-d_{ij}^2}{\epsilon}\right)$$

diffusion Kernel

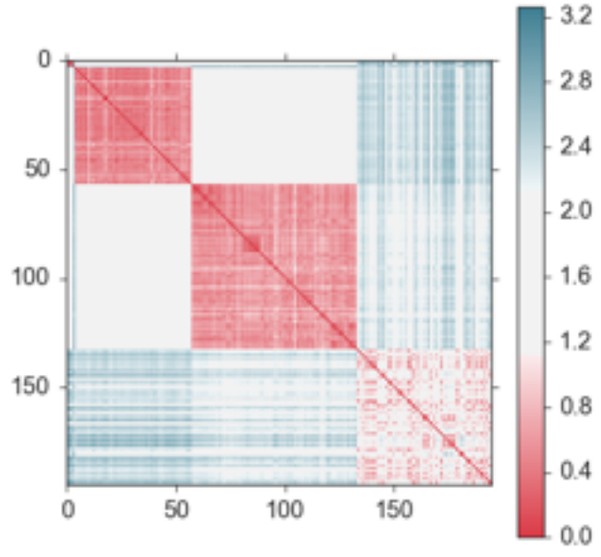
stochastic matrix

$$\mathbf{M} = \mathbf{D}^{-1}\mathbf{L}$$

M_{ij} contains transition probabilities from entry *i* to *j* .

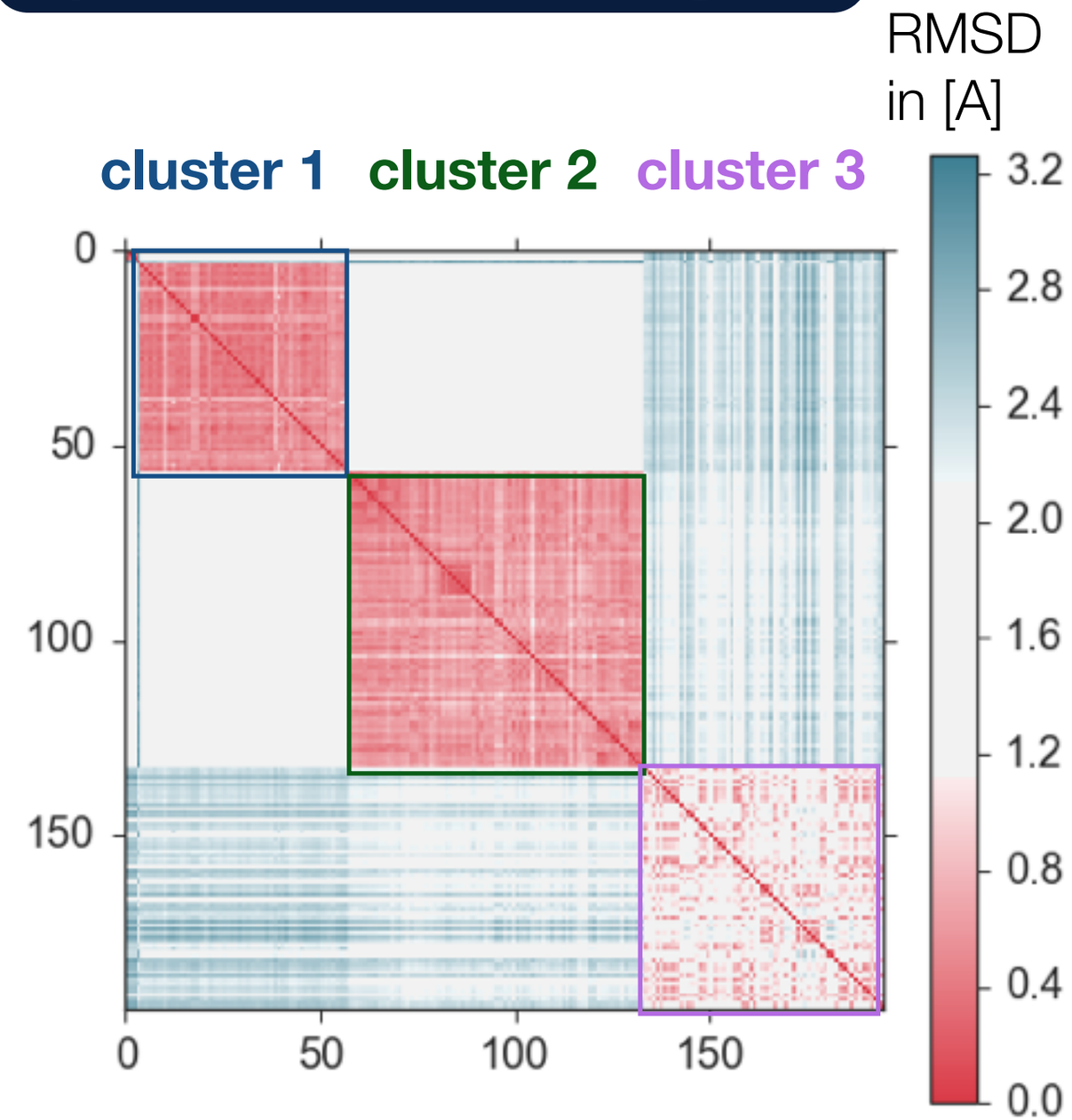


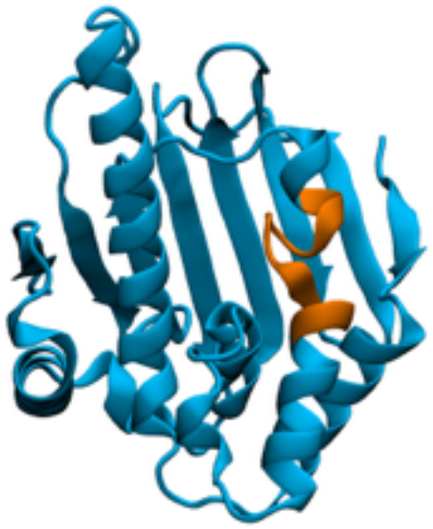
pcca+



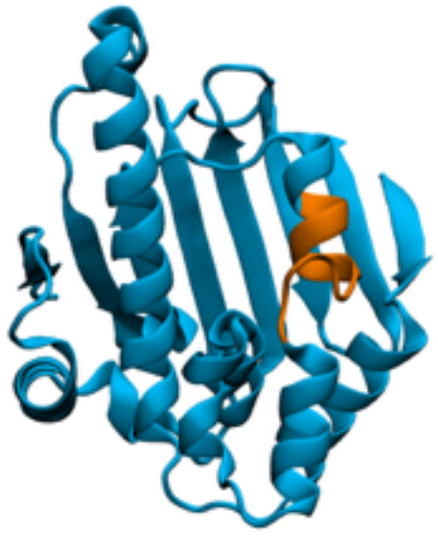
PyEMMA 2, JCTC 1549, 2015

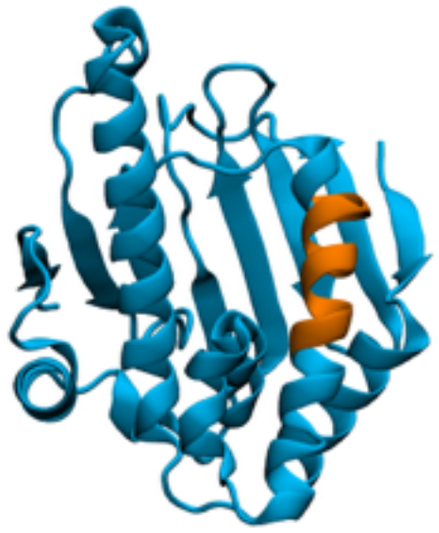
Spectral clustering





cluster 2 — alpha helix 104-111 closed

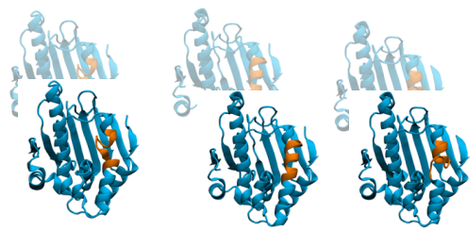




cluster 1 — loop 104-111 closed

cluster 3 — open alpha 104-111

Docking Protocols



Cross docking of all co-crystallised ligands within each cluster using rDock.

Pick crystal structure with the best cross-docking score for each cluster + 2 more crystal structures with **rearranged** **LYS58 ASP54** side chains: **2CCU 2FWZ 4CWF 4L94 4W7T**

Vina ‘default’ docking
5 crystal structure scores
average structure scores
best structure scores
visual structure scores*

rDock: Pharmacophoric restraints
5 crystal structure scores
average structure scores
best structure scores
visual structure scores

* ignores docking ranking, instead visual inspection of first 15 poses out of all 5 crystal structure protocols

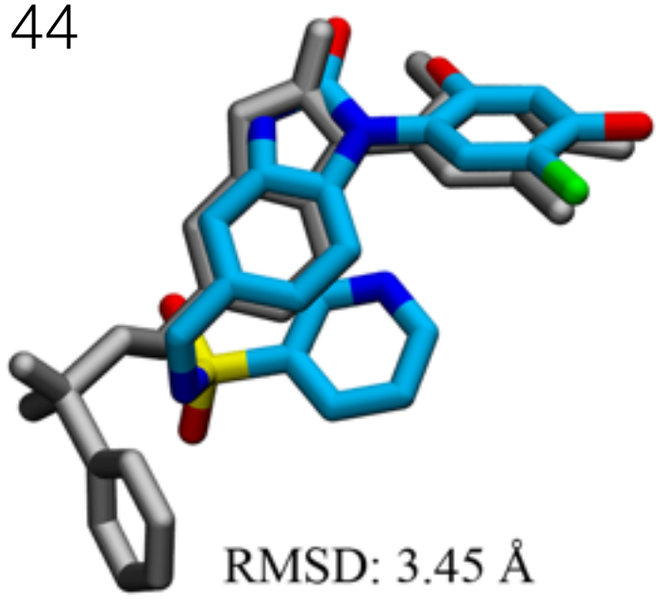
O. Trott, et al, J. Comp. Chem. 31 (2010) 455-461

Ruiz-Carmona et al. PLoS Comput Biol 10(4): e1003571.

Docking Results I

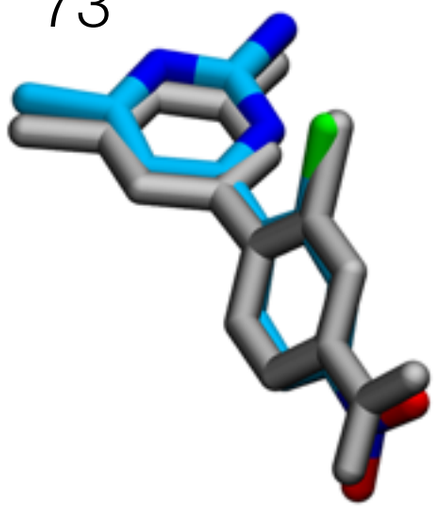
Visual rDock Result — Best Protocol for RMSD

44



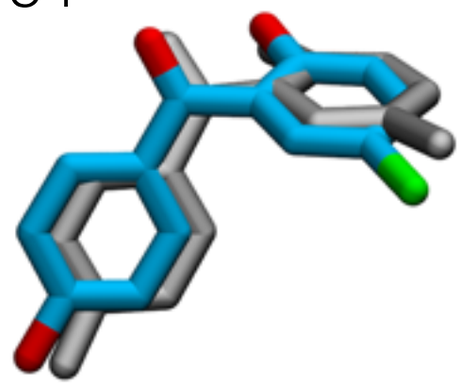
RMSD: 3.45 Å

73



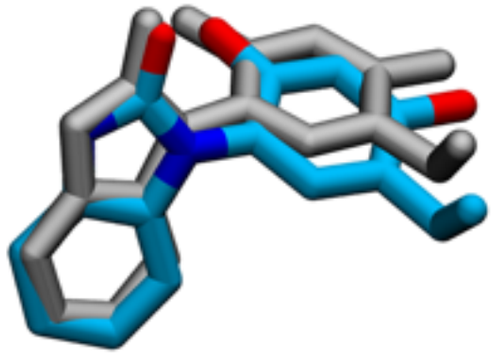
RMSD: 0.68 Å

164



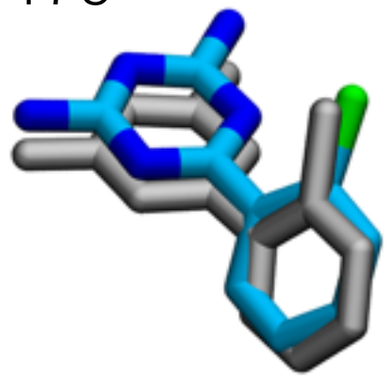
RMSD: 0.63 Å

40



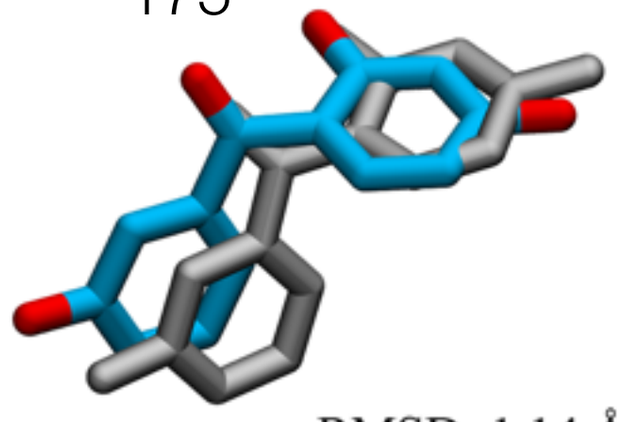
RMSD: 0.67 Å

179



RMSD: 0.63 Å

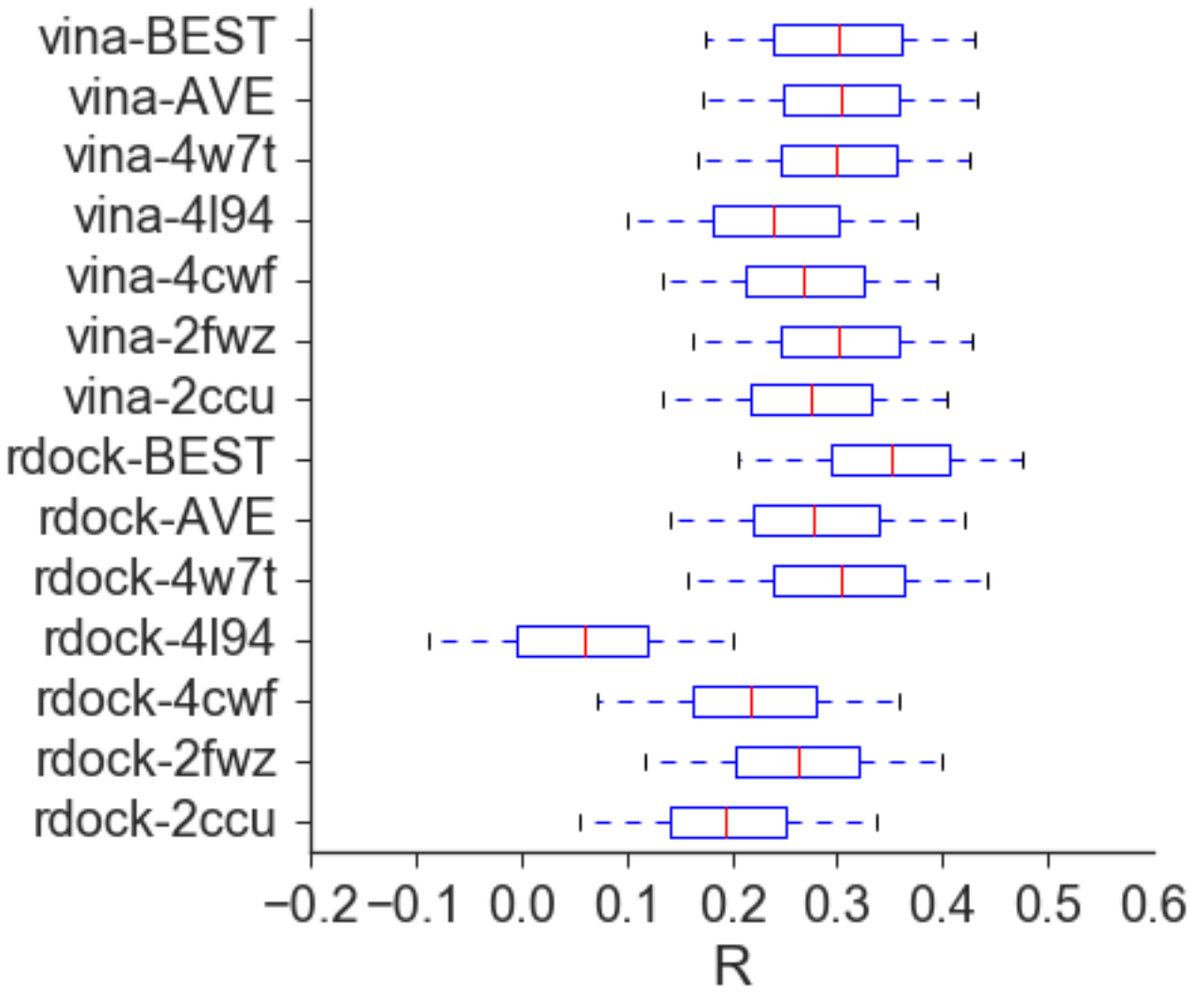
175

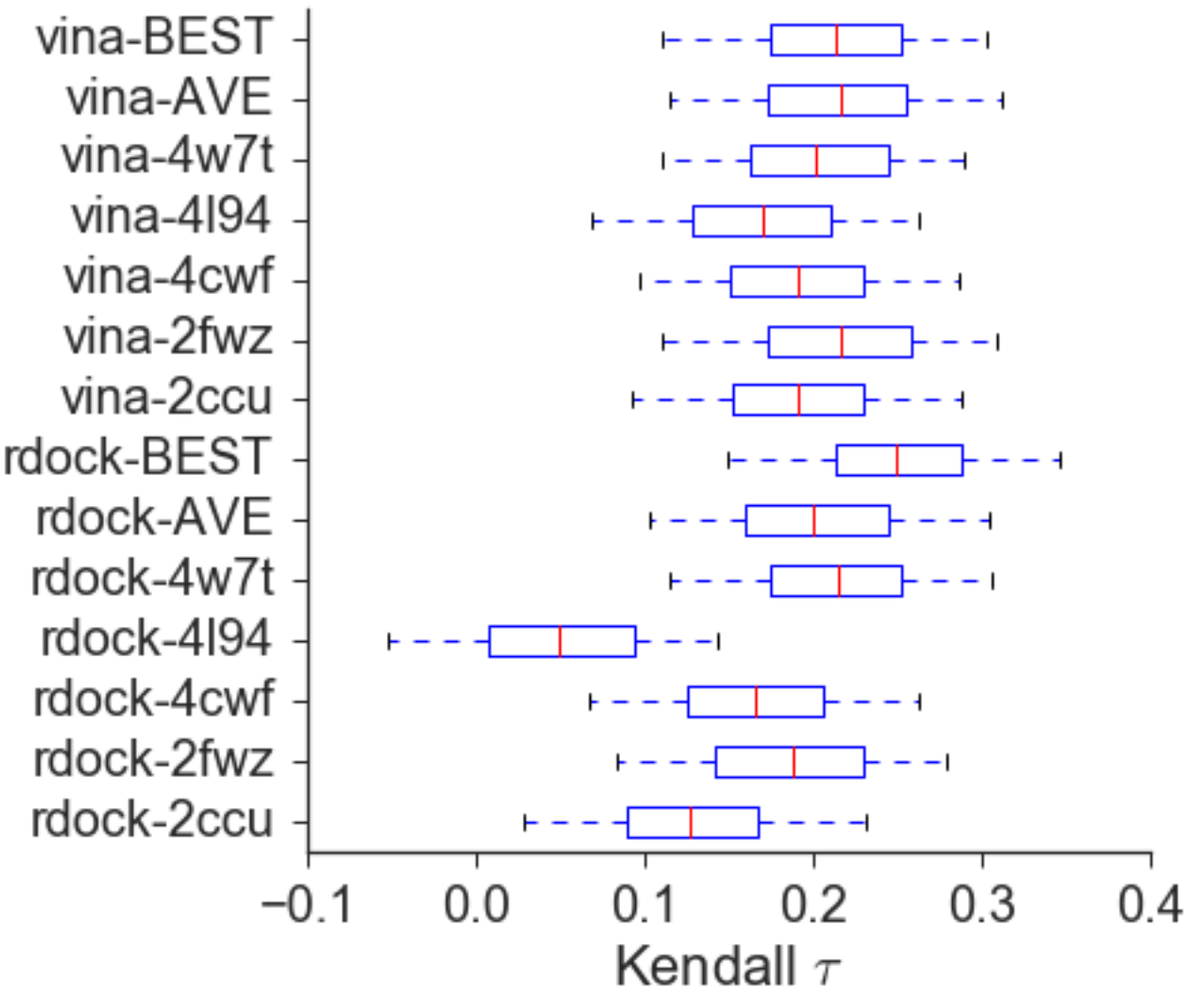


RMSD: 1.14 Å

Docking Results II

1000 bootstrap samples with 95% confidence intervals shown.



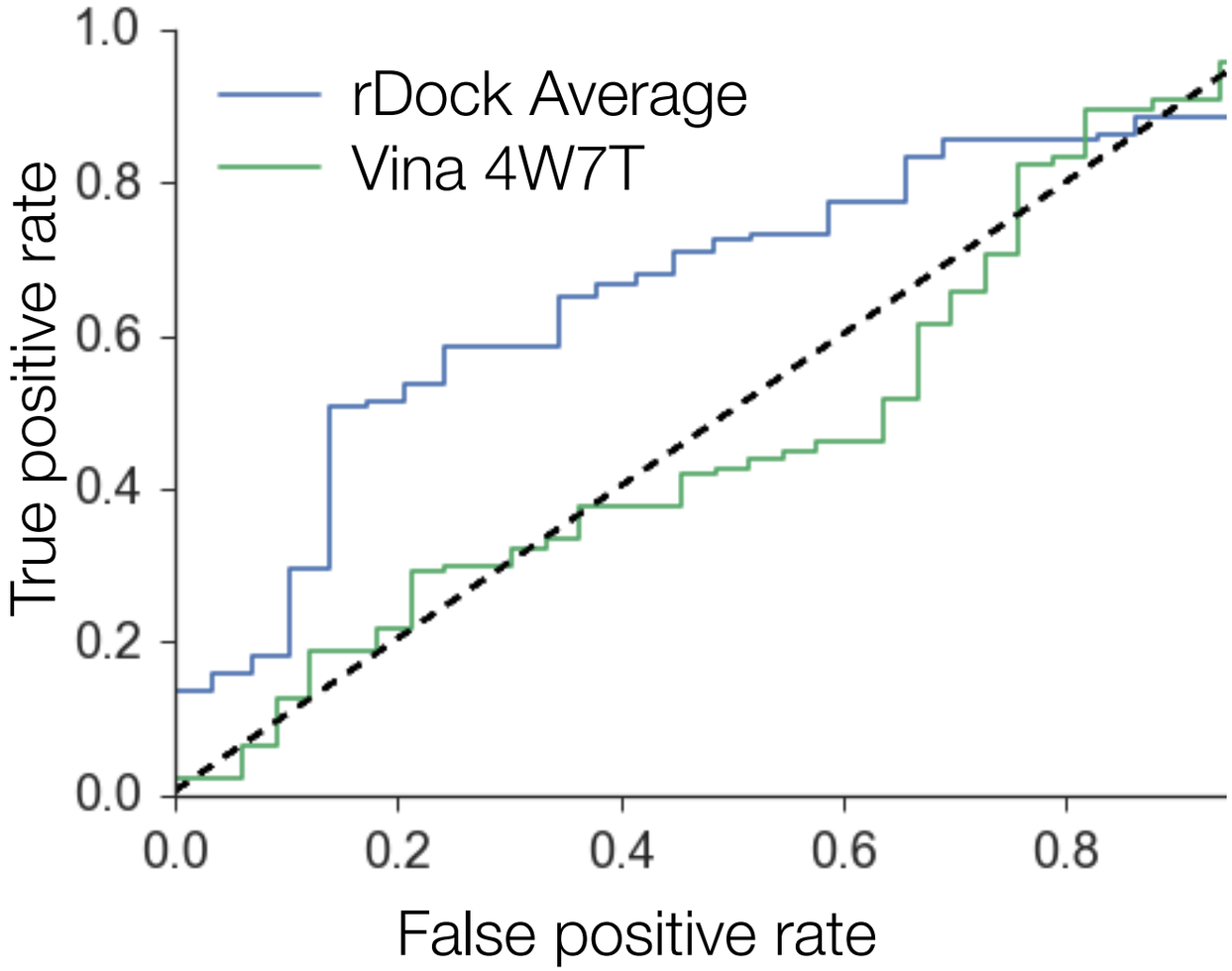


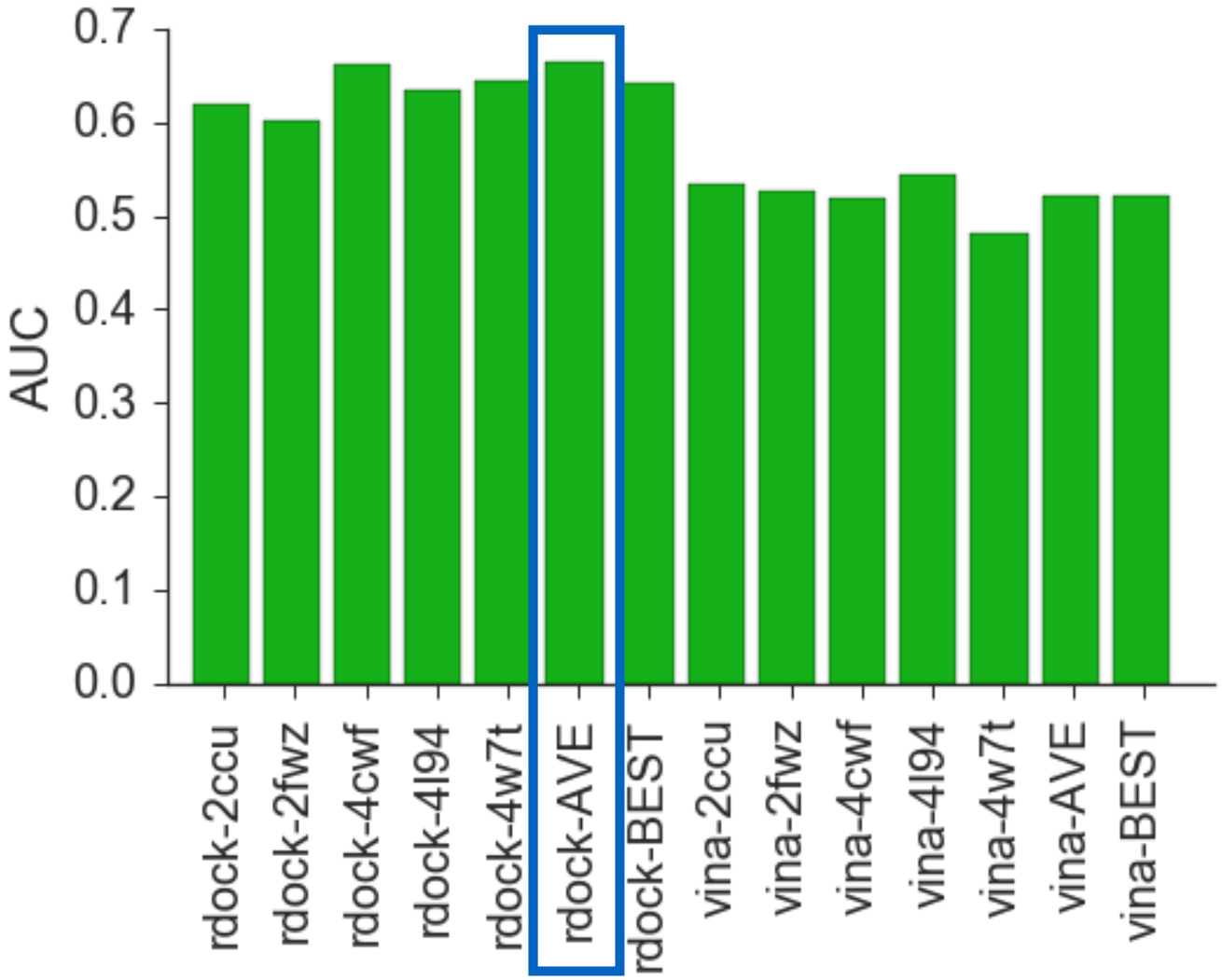
Mean R value and Kendall τ, off by 0.1 from computation of organisers.

Overall docking score preforms the best in comparison to other entries, including most of the dataset.

Docking Results II

active: [c] < 50 μM — 33 are not active





Maximum AUC computed is 0.64, for rDock Average protocol.
Organisers identified the same protocol as the best but computed a value of AUC = 0.73
AUC = 0.73 is also the best overall ranking score.
VINA scores are significantly worse.

Docking Conclusions

6 Pose predictions:

— visual rDock protocol gives the best RMSD results.
— visual rDock performs well in comparison to other submissions.

Scoring of 180 compounds:

— Difficult to **establish best protocol based** on Kendall τ or R value
— **Clear outlier of** protocol docking to crystal structure 4I94 using rDock
— All protocols **perform comparable to or better than** other submissions
— AUC calculations show clear similarities within rDock and Vina protocols, but also clear differences between rDock and Vina.
— Best AUC score performs best amongst other submissions.

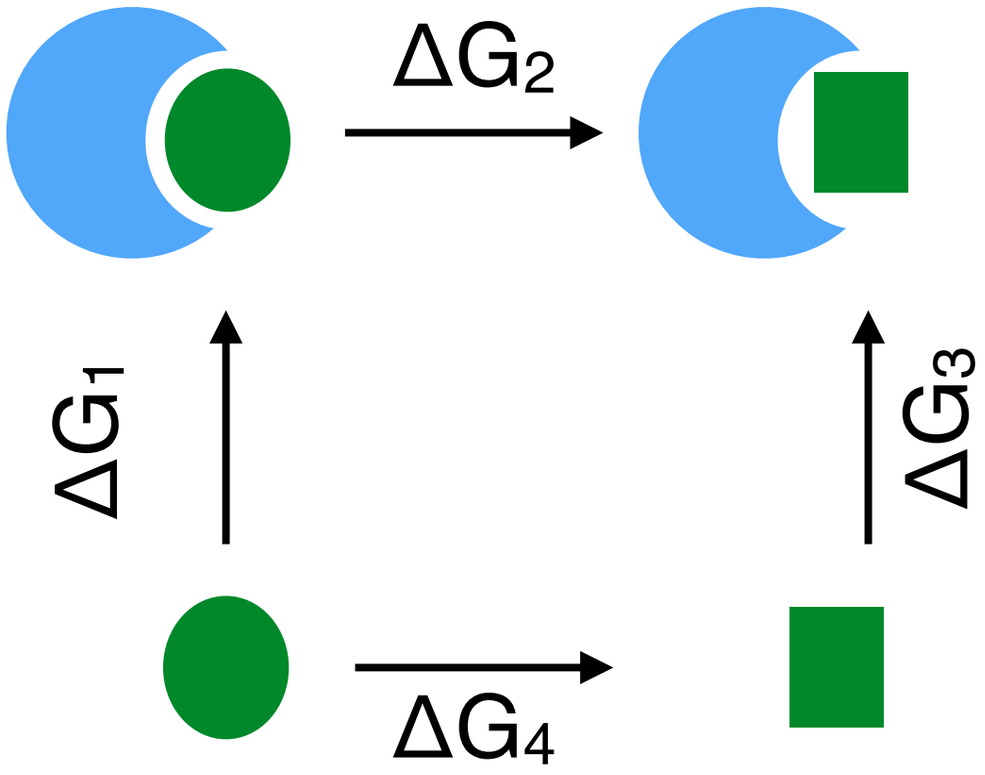
Overview

Free Energy Calculations

- **Free energy protocol** and test data
- **Set 3 results** — a reasonable success
- **Set 2 results** — aka how a water molecule can drastically change your prediction
- **Set 1 results** — what went wrong?

Alchemical free energy

$$\Delta\Delta G_{bind} = \Delta G_4 - \Delta G_2 = \Delta G_1 - \Delta G_3$$

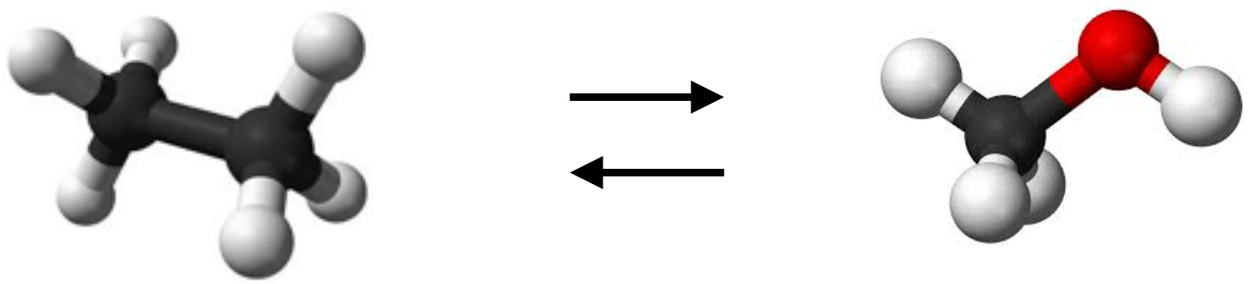


ΔG_1, ΔG_3 are difficult to compute with MD simulations.

ΔG_2, ΔG_4 can be computed via so called alchemical free energy calculations (AFEC)

Alchemical free energy

Relative free energy calculation involves modifying our MD potential with a switching function that allows an artificial perturbation from one molecule to another.



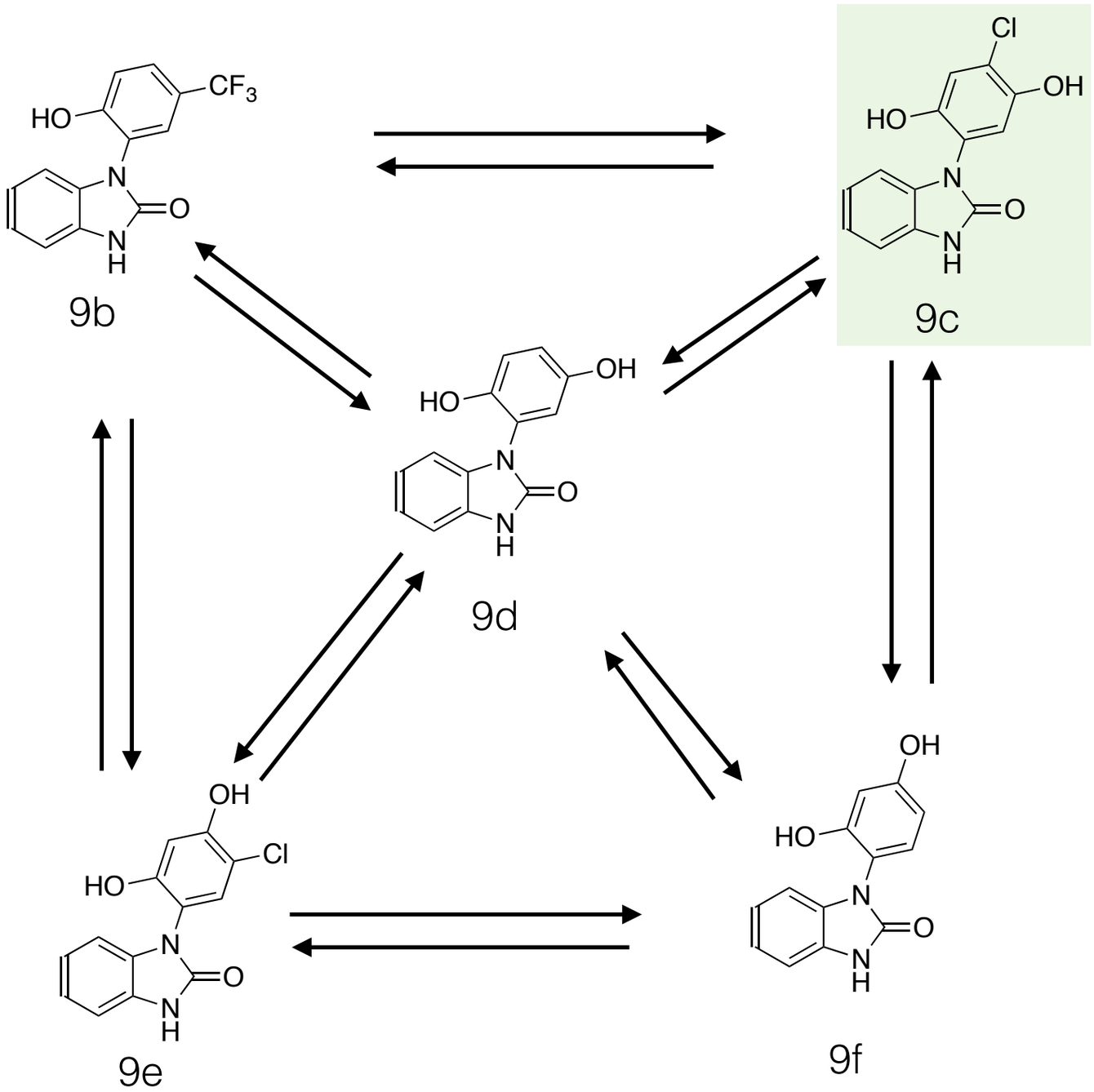
$$U(\mathbf{r}_1 \dots \mathbf{r}_N, \lambda) \equiv f(\lambda) U_A(\mathbf{r}_1 \dots \mathbf{r}_N) + g(\lambda) U_B(\mathbf{r}_1 \dots \mathbf{r}_N)$$

$$f(\lambda) = 1 - \lambda \qquad\qquad g(\lambda) = \lambda$$

Setup is done using FESetup and somdfreenrg from the software package Sire using OpenMM, all small molecules were parametrised using GAFF AM1/BCC.

Test set

Benzimidazolone derivatives as HSP90 inhibitors



Experimental values

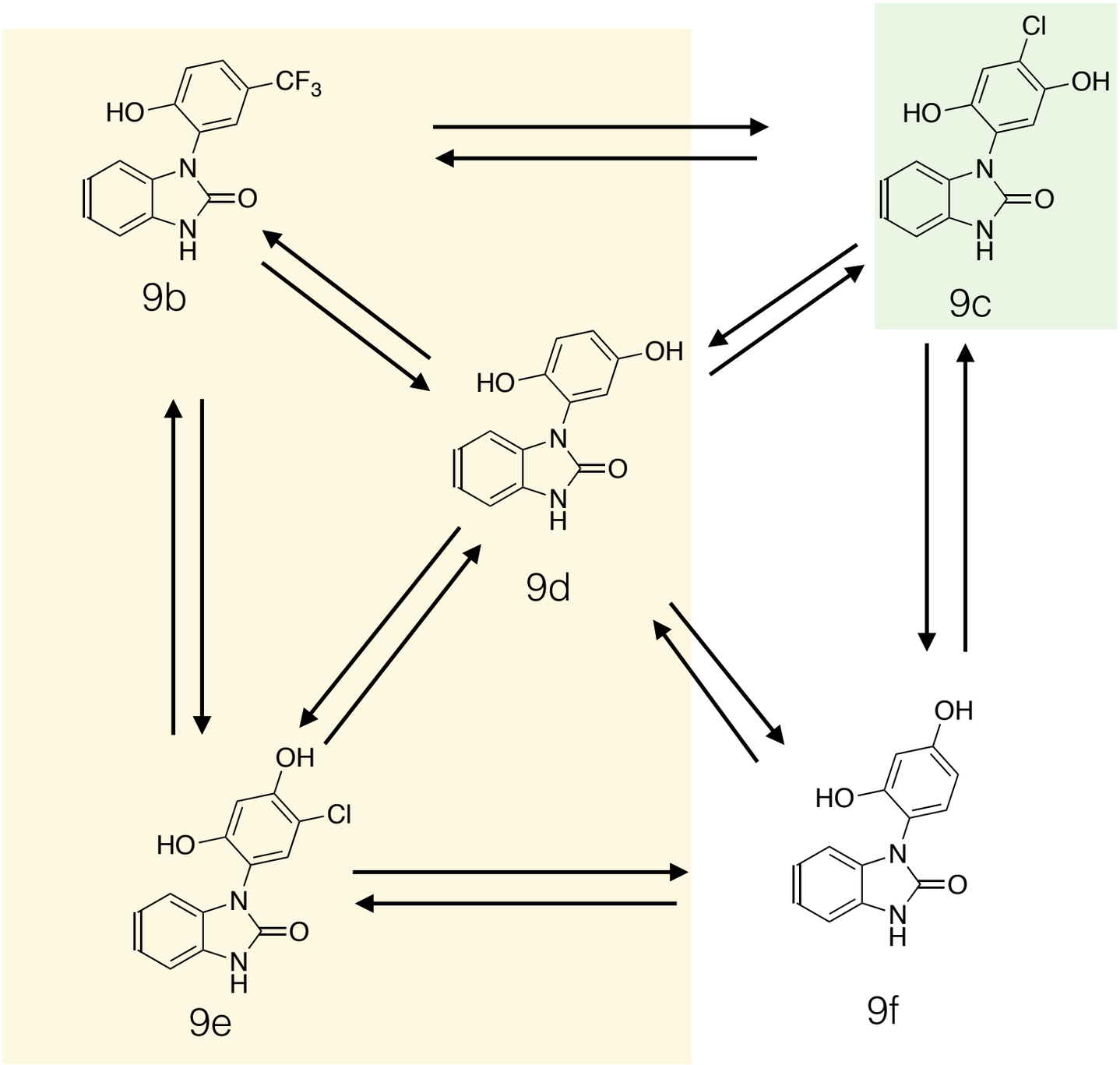
ID	IC50 [μM]
9b	2.32
9c	0.22
9d	34.9
9e	0.054
9f	2.53

9c is co-crystalised in **3OW6** and used as a template for other compounds.

Bruncko *et* al. 2010

Test set

Benzimidazolone derivatives as HSP90 inhibitors



Experimental values

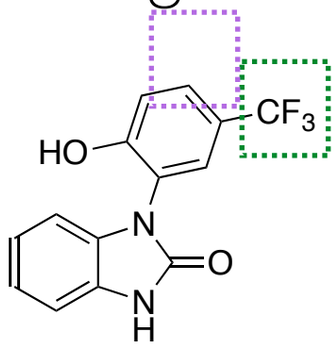
ID	IC50 [μM]
9b	2.32
9c	0.22
9d	34.9
9e	0.054
9f	2.53

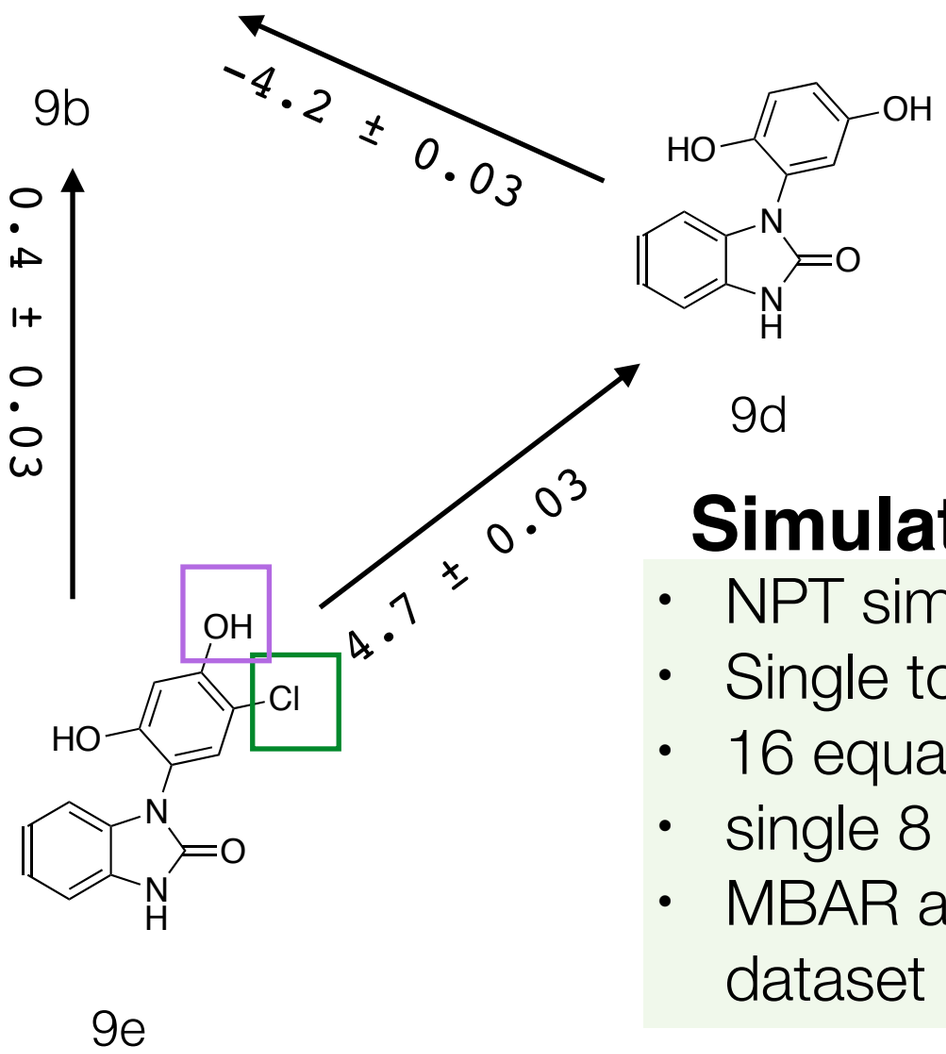
9c is co-crystalised in **3OW6** and used as a template for other compounds.

Bruncko *et* al. 2010

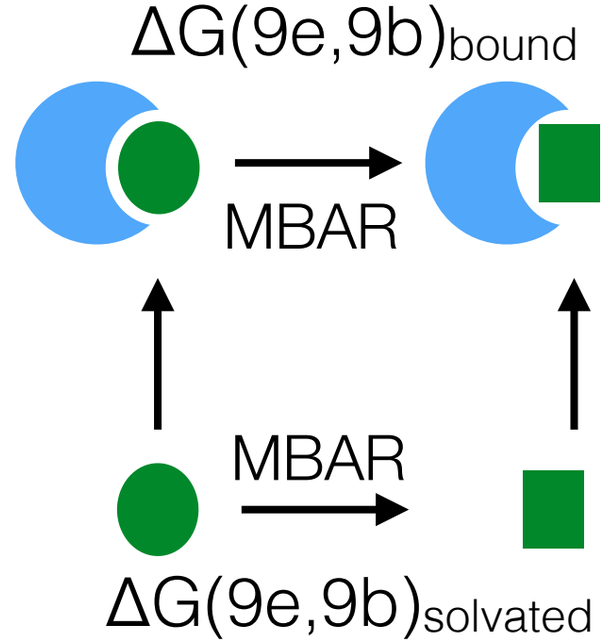
Test set

Getting to the numbers





$\Delta\Delta G(9e,9b) =$
$\Delta G(9e,9b)_{bound} - \Delta G(9e,9b)_{solvated}$

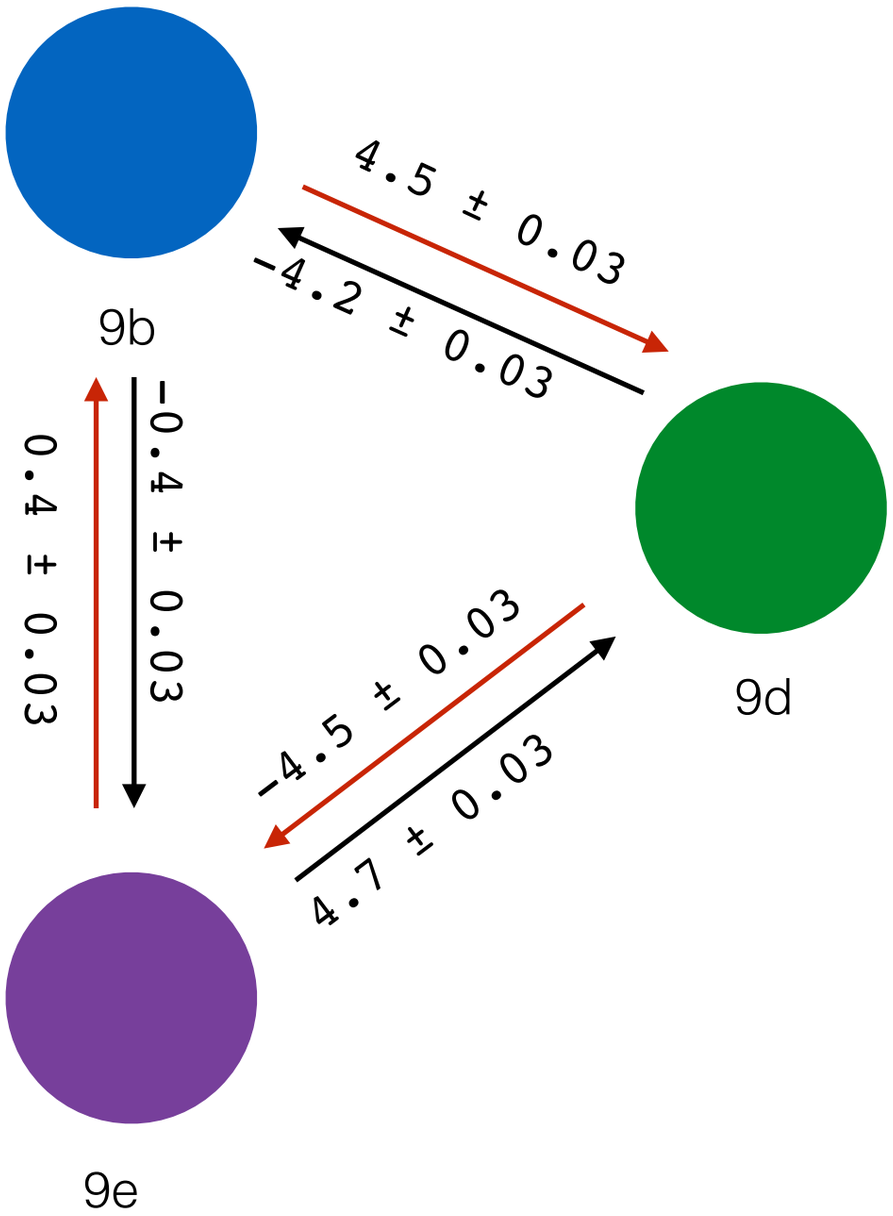


Simulation details:

- NPT simulation, with 4 fs time step at 298K.
- Single topology.
- 16 equally spaced λ windows.
- single 8 ns production run. (Actual data multiple repetitions)
- MBAR analysis after drawing uncorrelated data from simulated dataset using the time series analysis module in pymbar.

Test set

If there is no experimental data available how can we know that an alchemical calculation gives reasonable results?



Kirchhoff's law:

$$\sum_{i=1}^{n} \phi_i = 0$$

For example: cycle

ΔΔG(9d,9e) + ΔΔG(9e,9b)+ΔΔG(9b,9d)=0

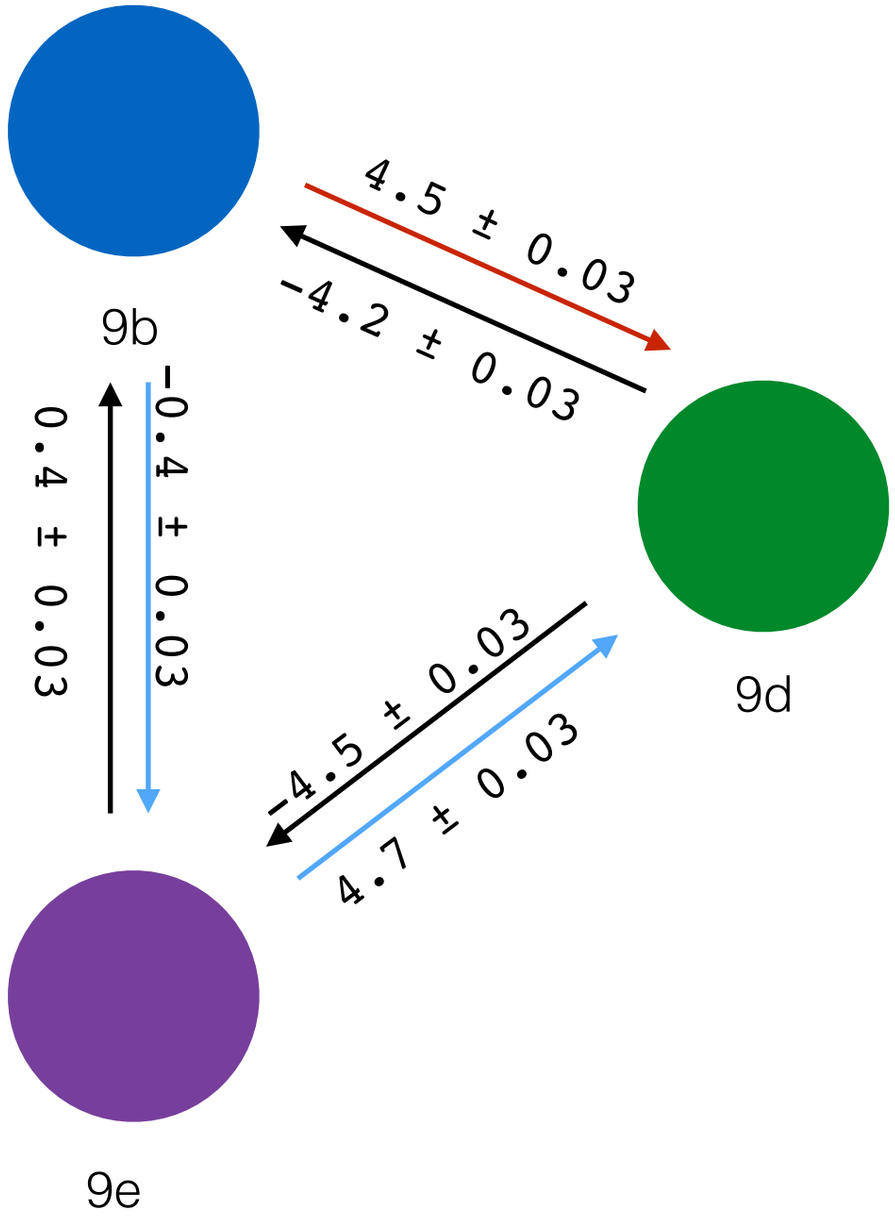
Also:

ΔΔG(9d,9e) = - ΔΔG(9e,9d)

Bruncko *et* al. 2010

Test set

In a directed graph there are multiple paths to compute the same relative free energy:



$\Delta\Delta G(9b,9d)_{p1}$ = →

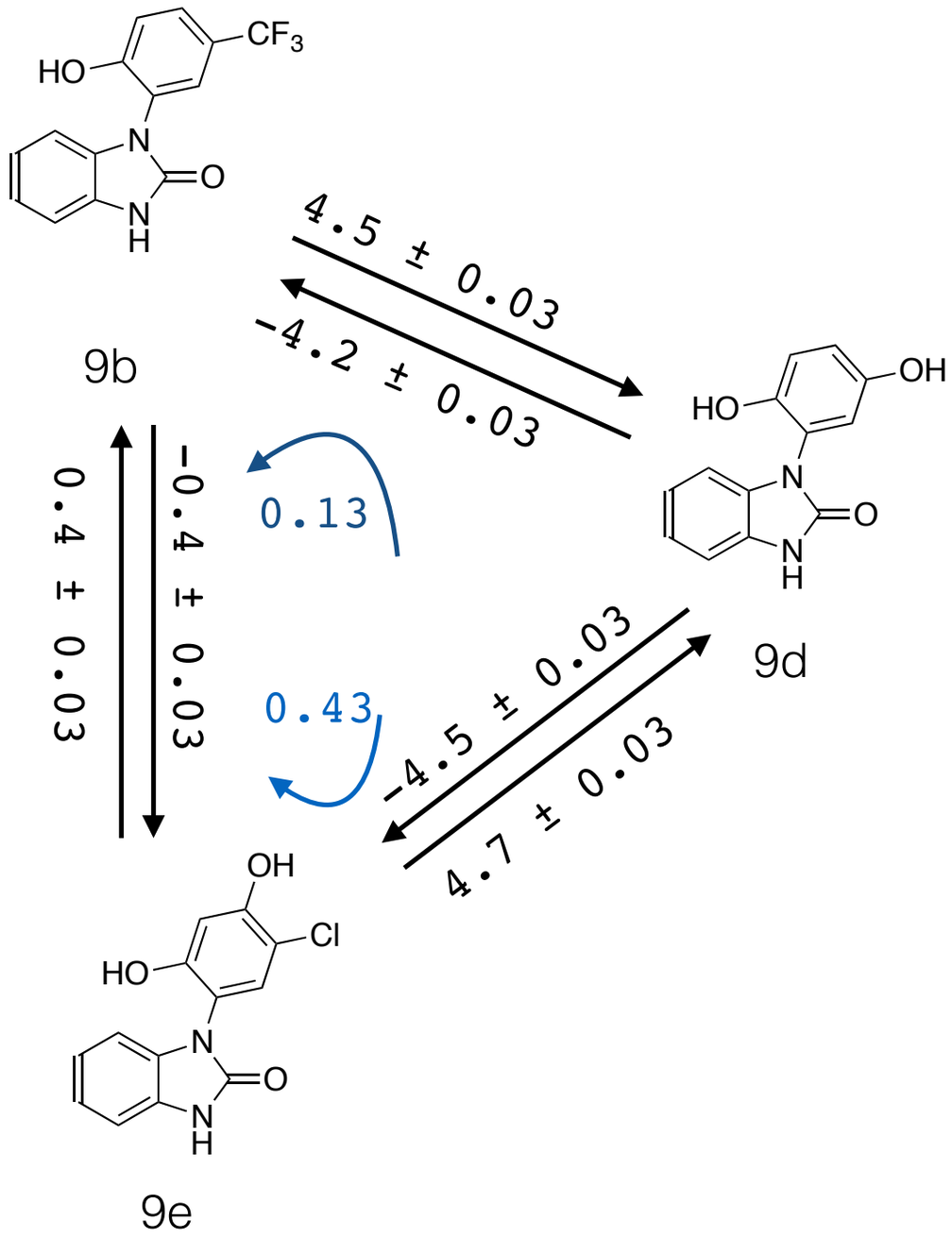
\+ $\Delta\Delta G(9b,9d)_{p2}$ = → + →

$\langle\Delta\Delta G(9d,9f)\rangle \pm \sigma(\Delta\Delta G(9d,9f))$

For D3R results, we computed the average over all possible paths and the resulting standard deviation.

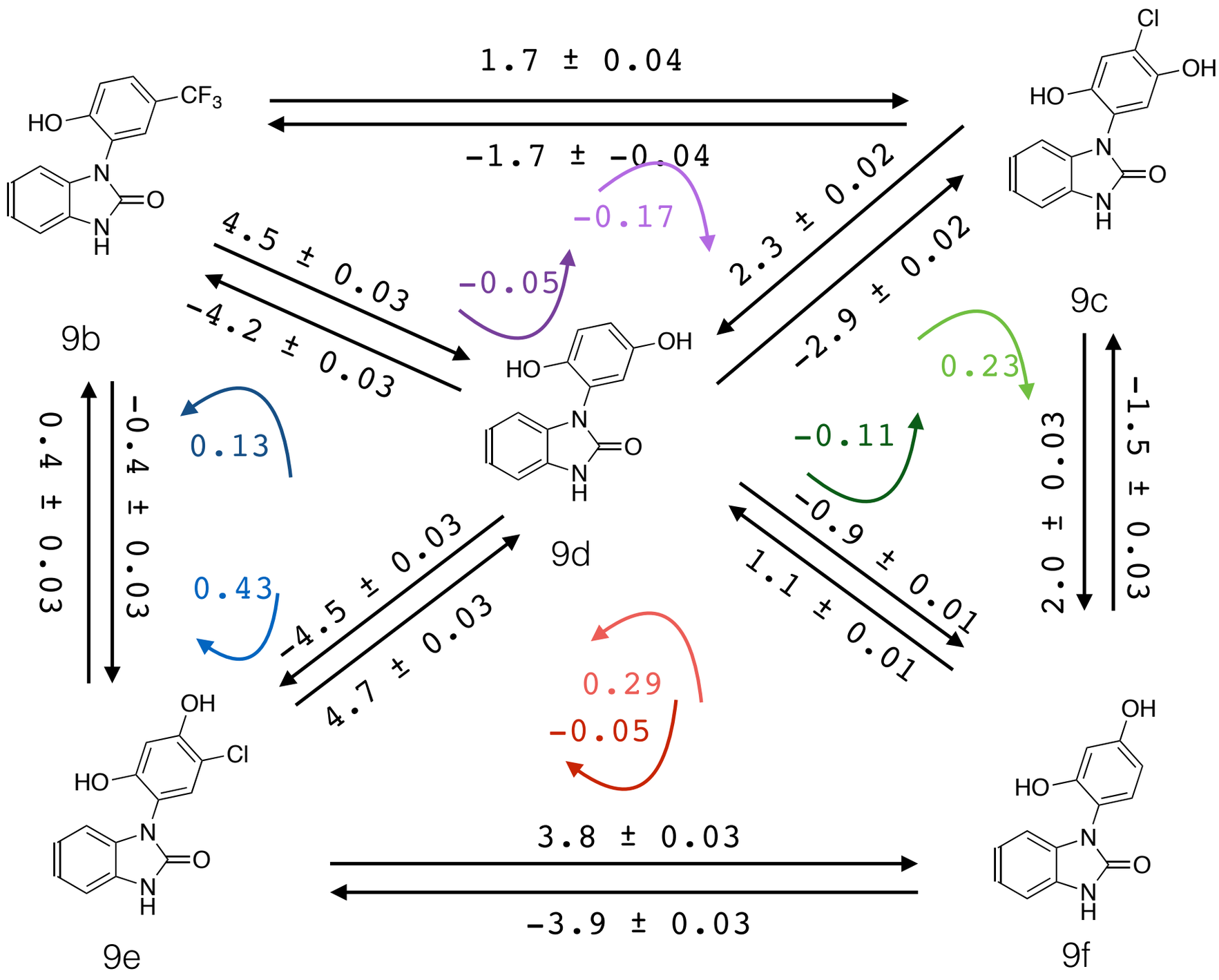
Bruncko *et* al. 2010

Test set



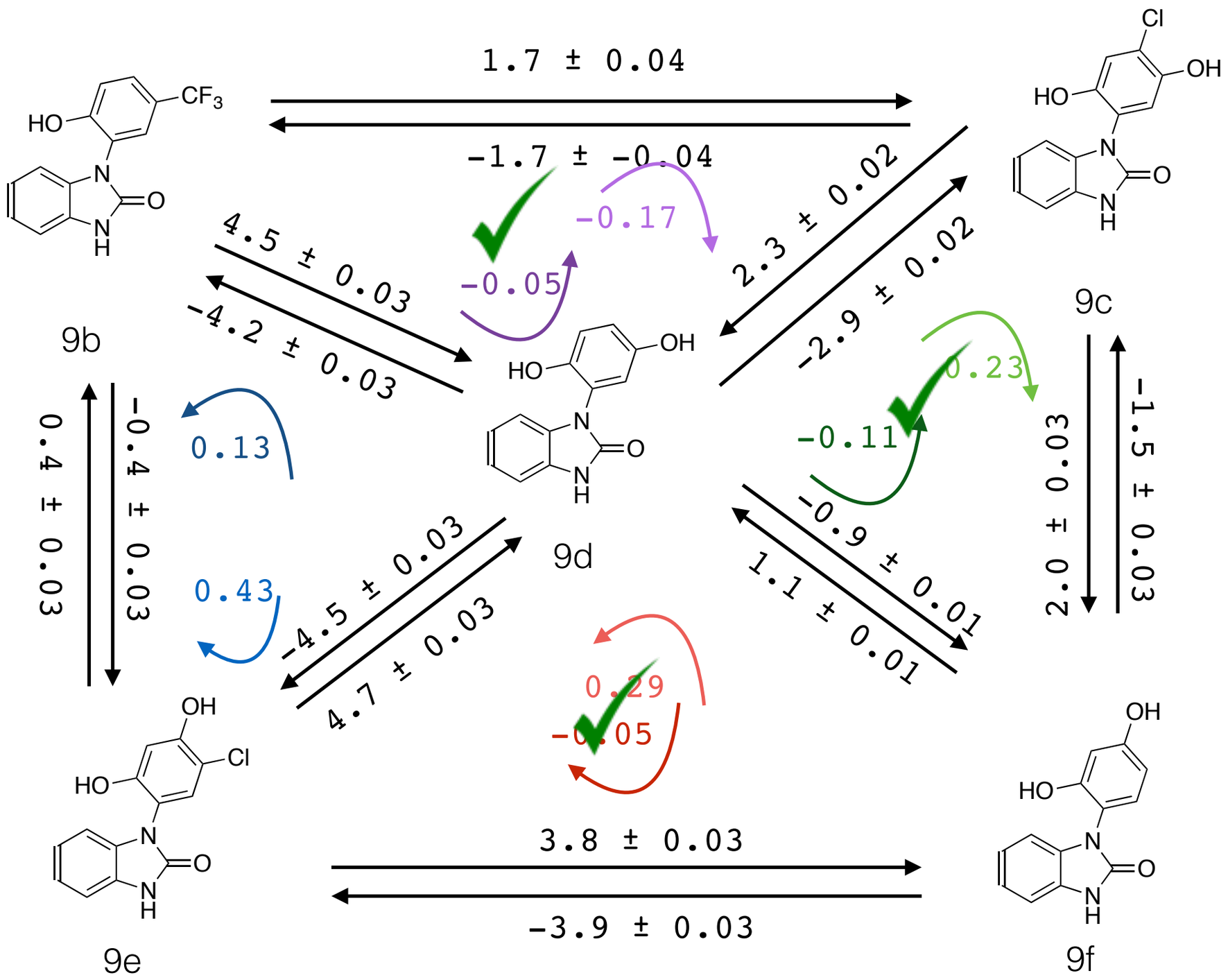
Bruncko *et* al. 2010

Test set



Bruncko *et* al. 2010

Test set

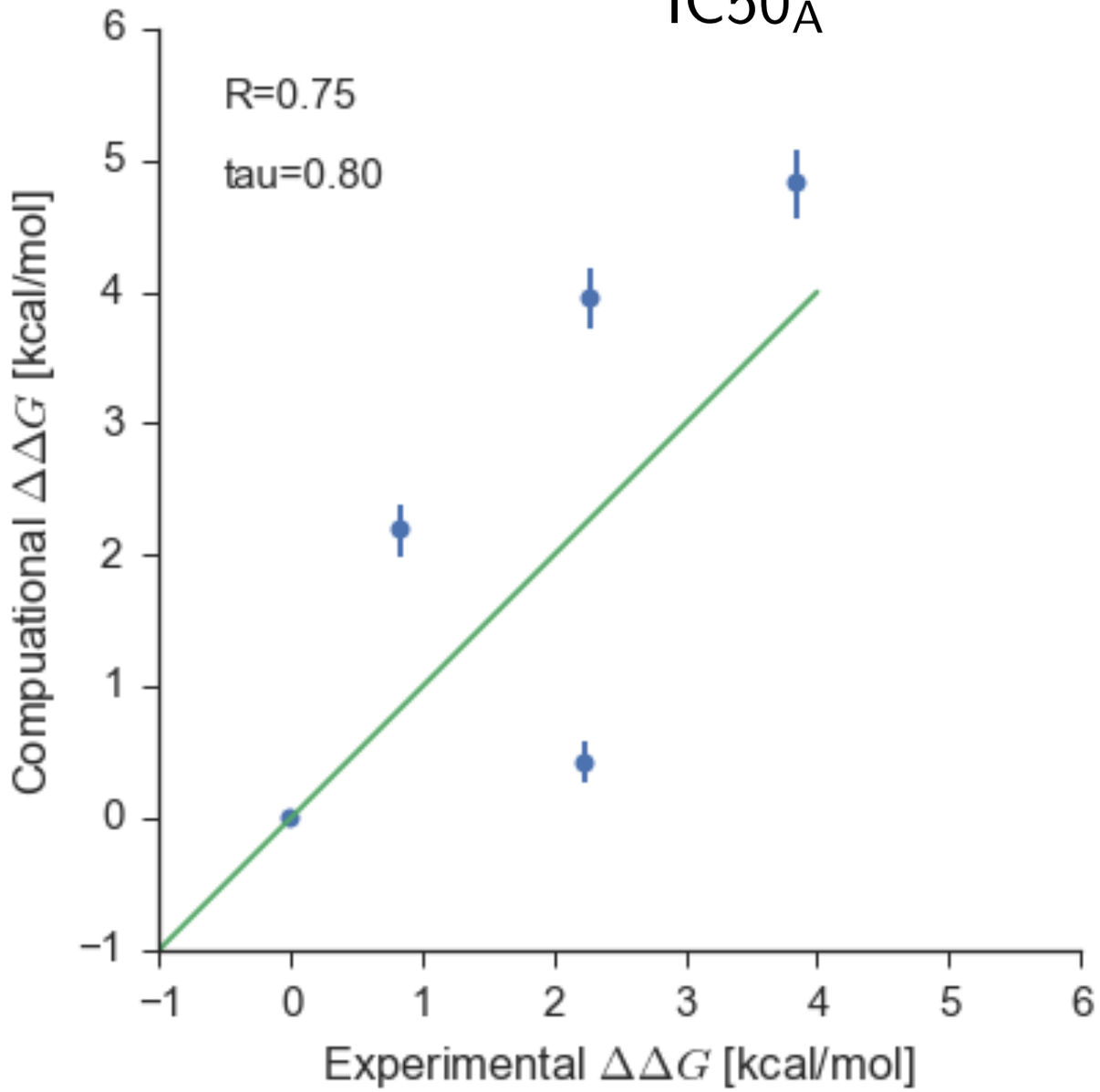


Bruncko *et* al. 2010

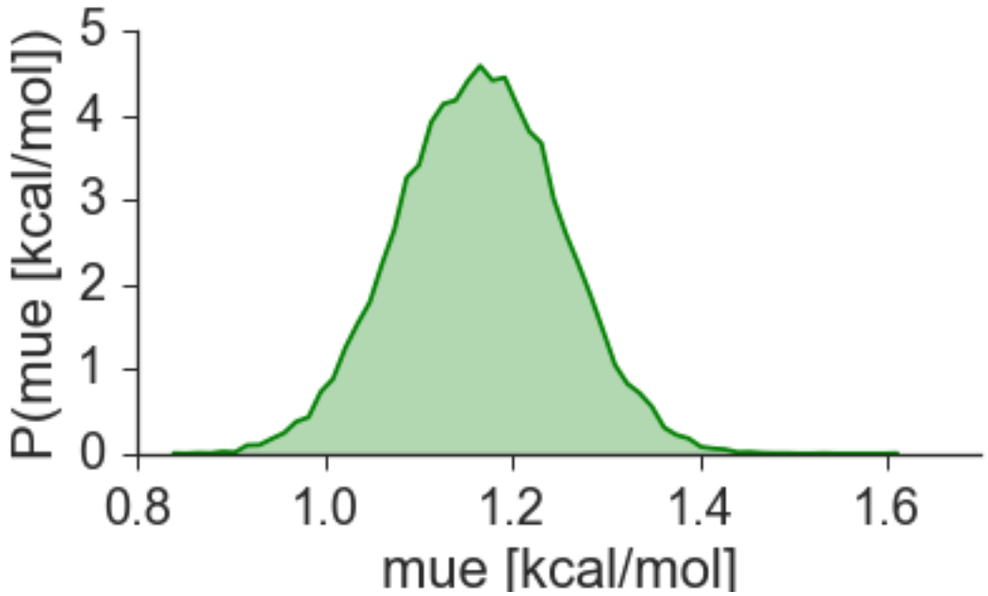
Test set

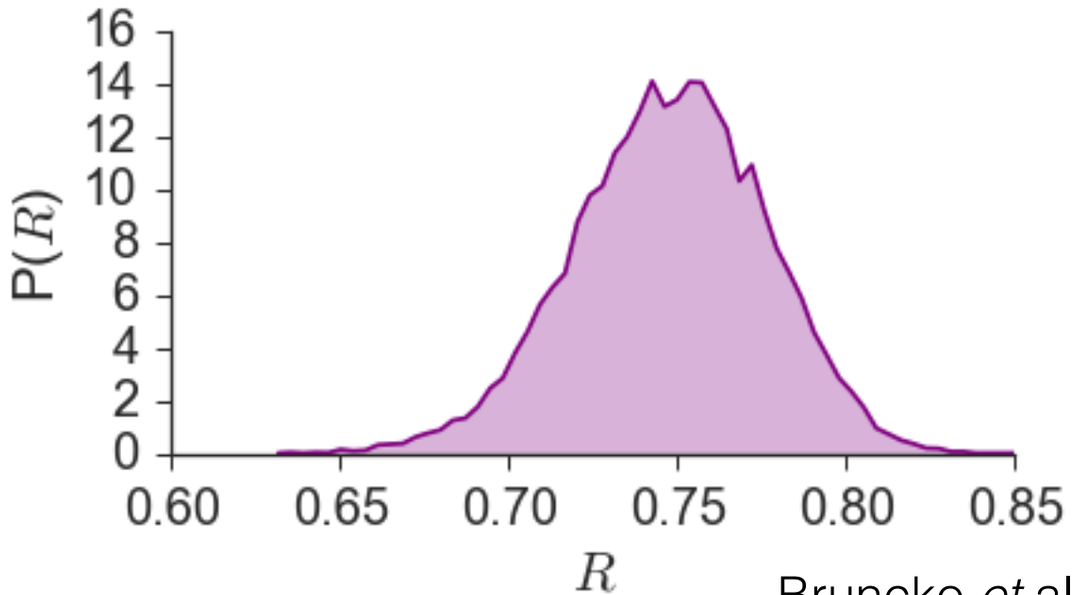
Comparing to experiment:

$$\Delta\Delta G_{A,B} = k_B T \ln(\frac{\text{IC50}_\text{B}}{\text{IC50}_\text{A}})$$



Bootstrapping computational predictions, looking at variation of different observables.



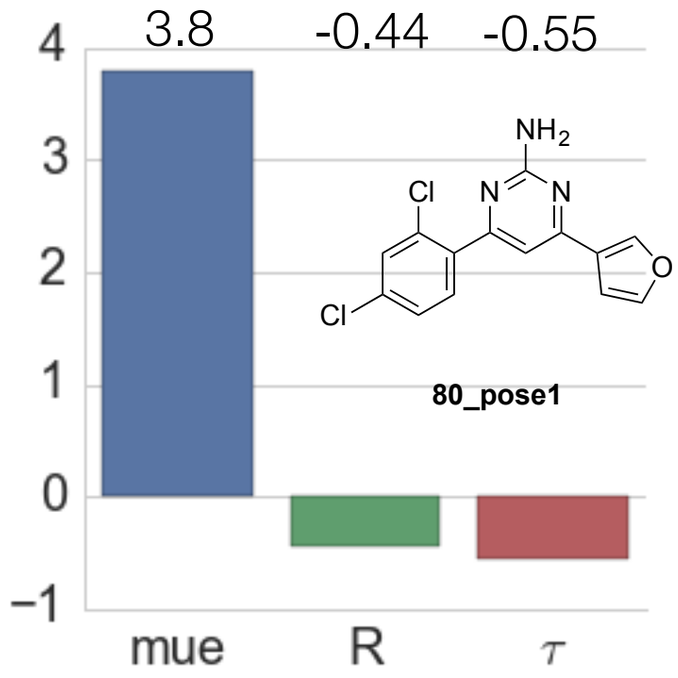


Bruncko *et* al. 2010

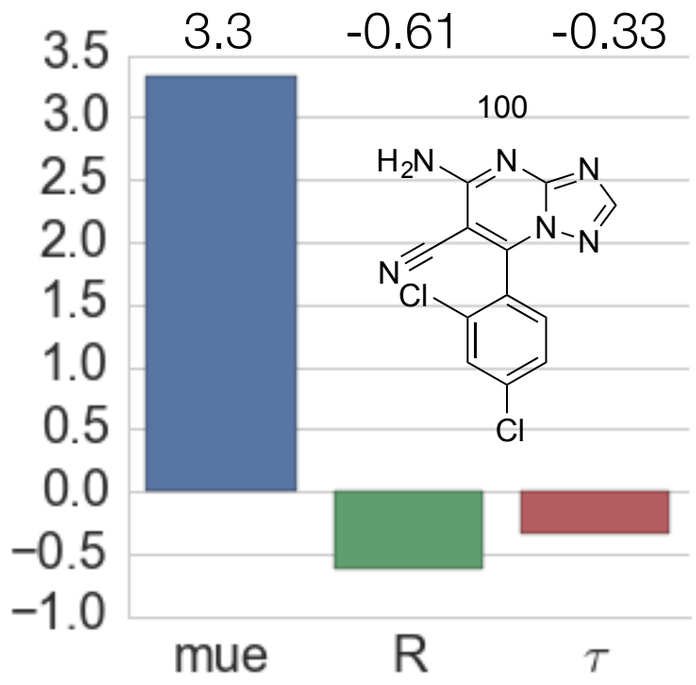
Set 1-3— how we did

Summary of bootstrapped results comparing to experimental data:

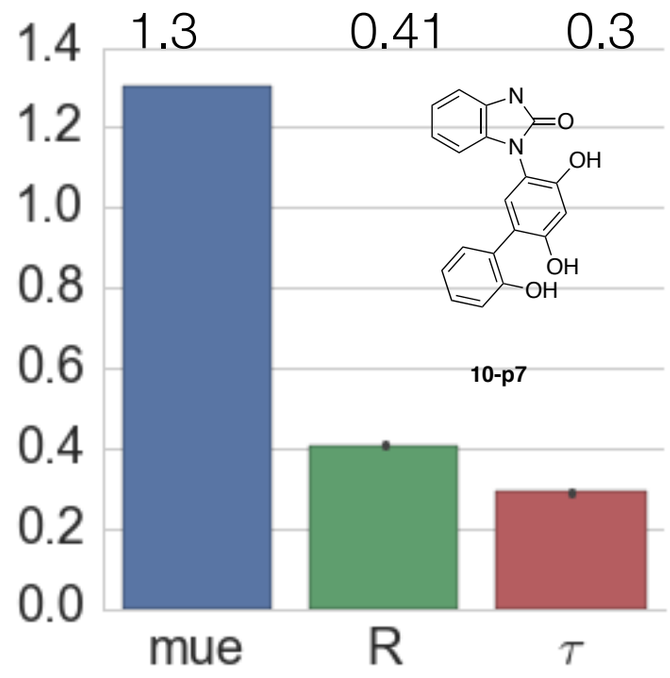
Set 1



Set 2



Set 3

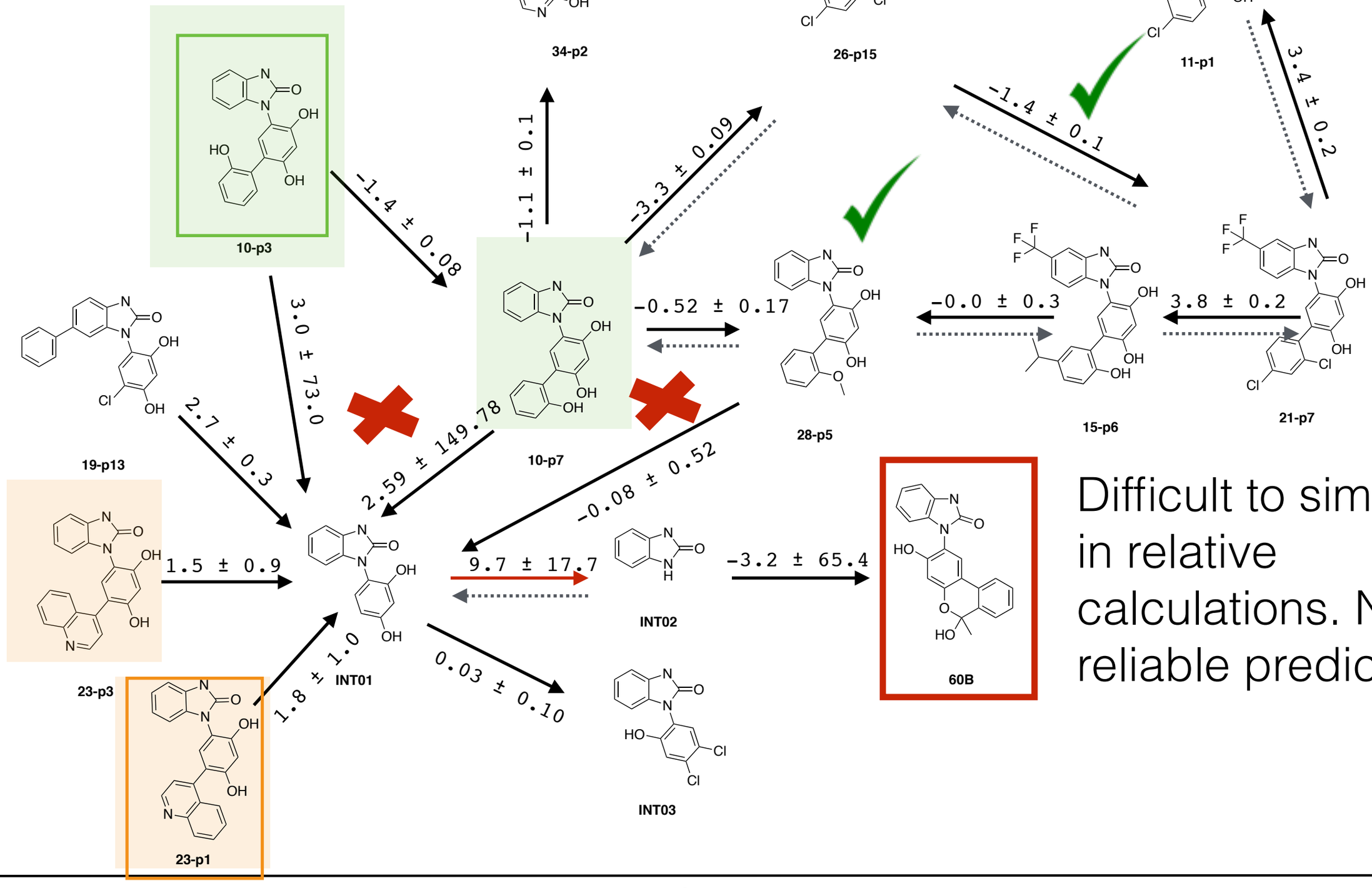


Organisers analysis:

Pearson R	-0.80 / 41st (44)	-0.40 / 11th (18)	0.42 / 3rd (20)
RMS error val/ Rank (# submissions)	2.67 / 12th (44)	2.00 / 10th (18)	1.43 / 1st (20)

Set 3 — how we did

Averages over two runs.

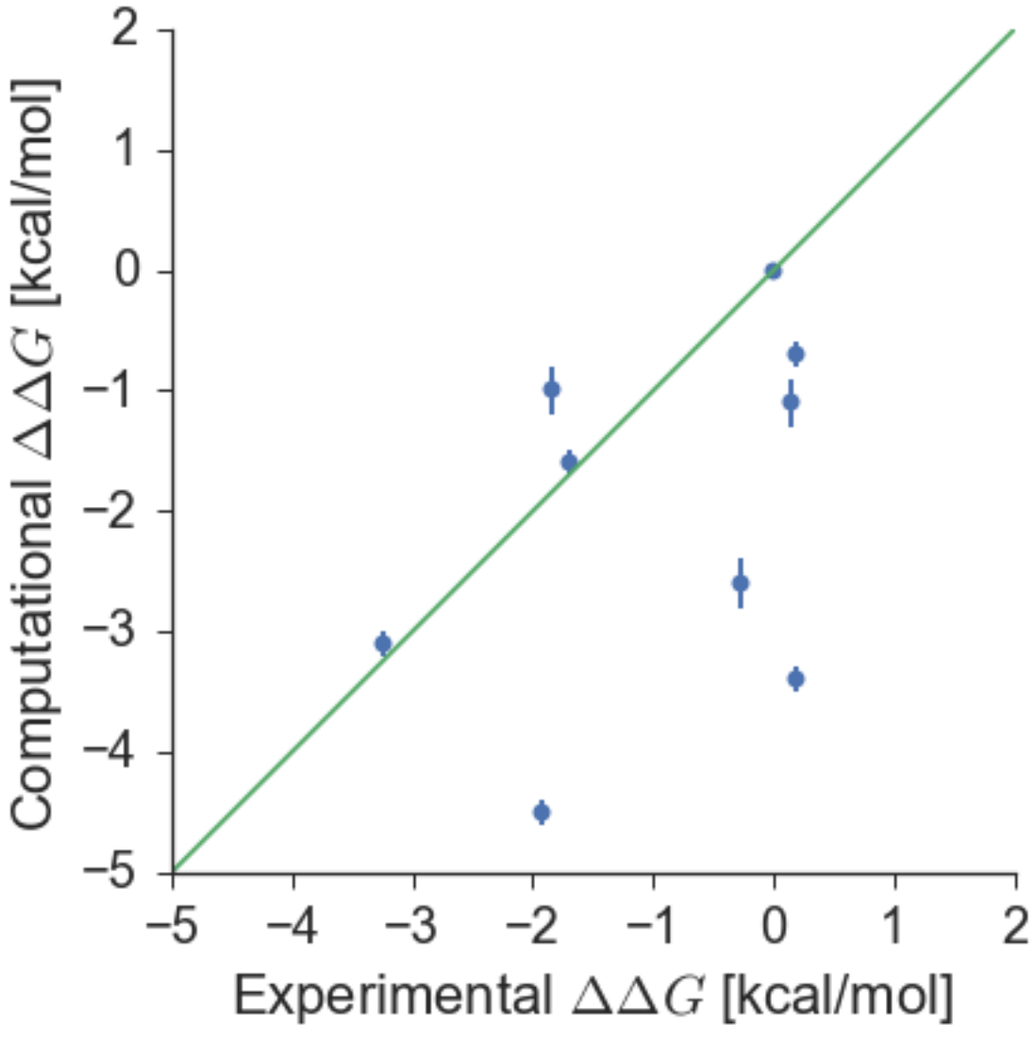


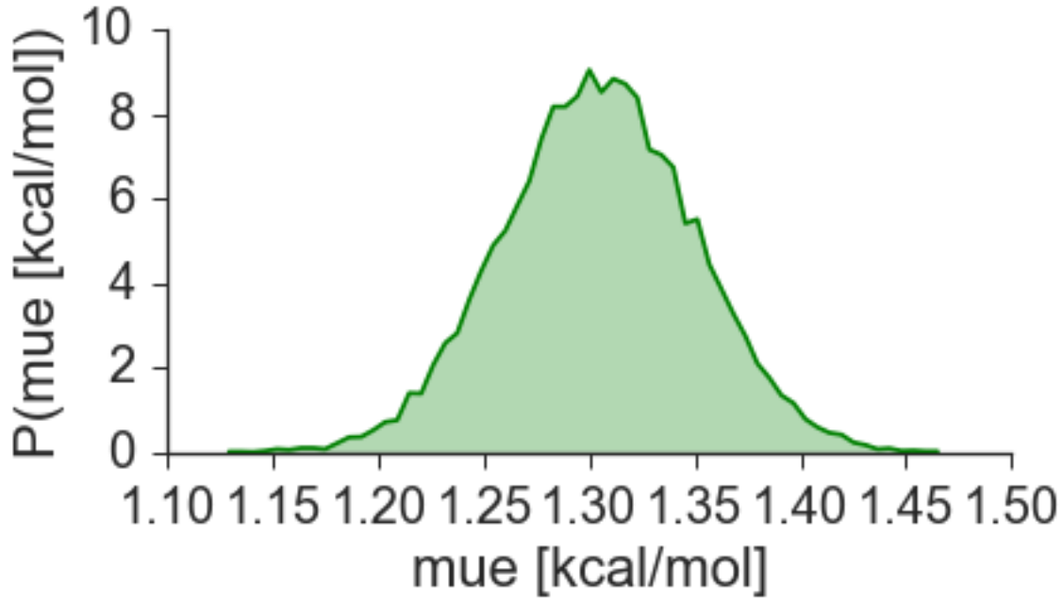
Difficult to simulate in relative calculations. No reliable prediction.

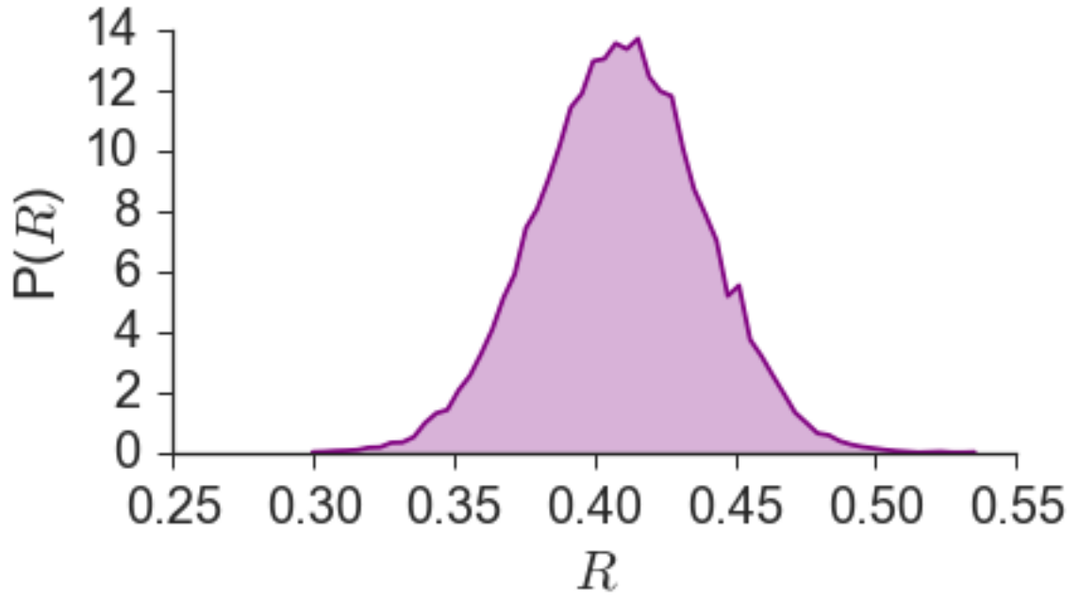
Set 3 — how we did

Comparing to experiment:

$$\Delta\Delta G_{A,B} = k_B T \ln(\frac{\text{IC50}_\text{B}}{\text{IC50}_\text{A}})$$

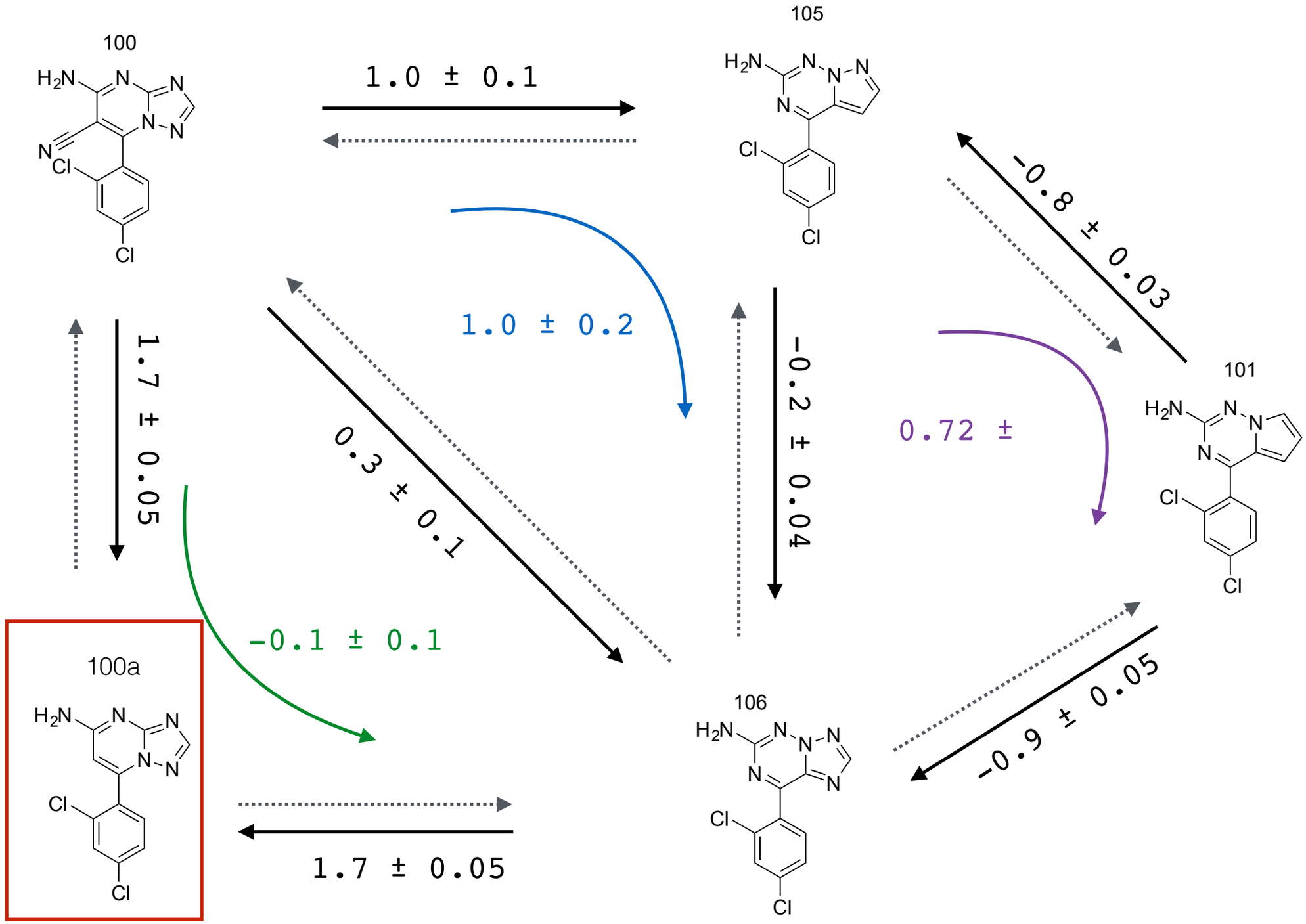




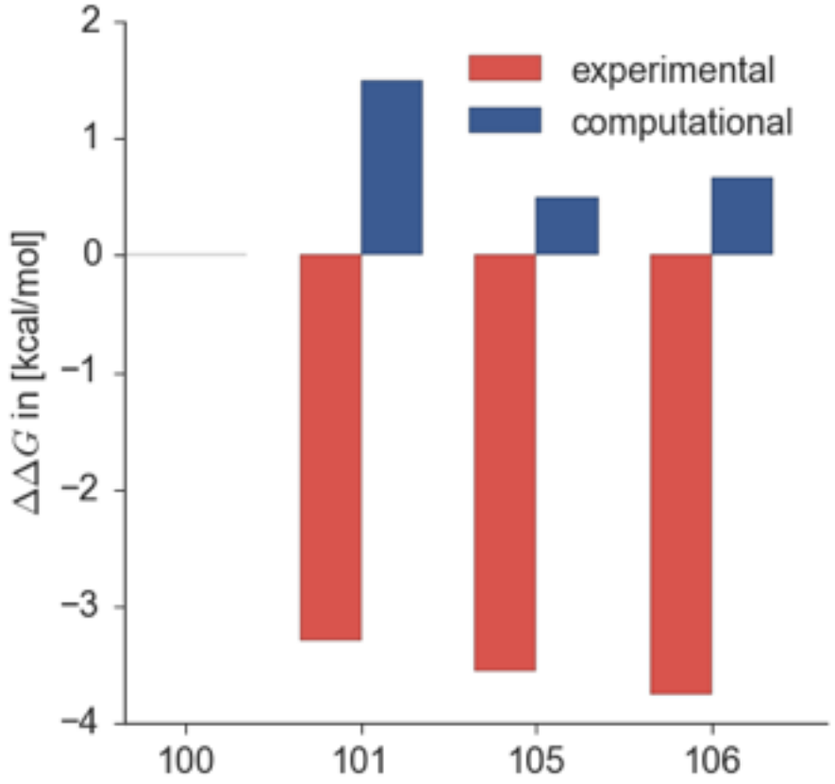


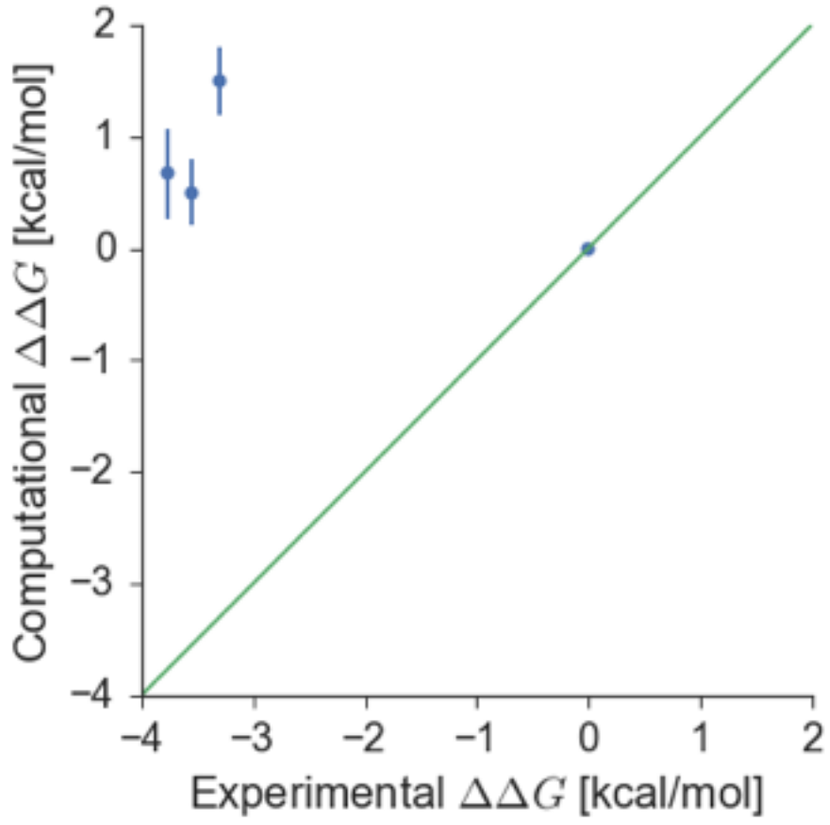
Set 2 — how we did

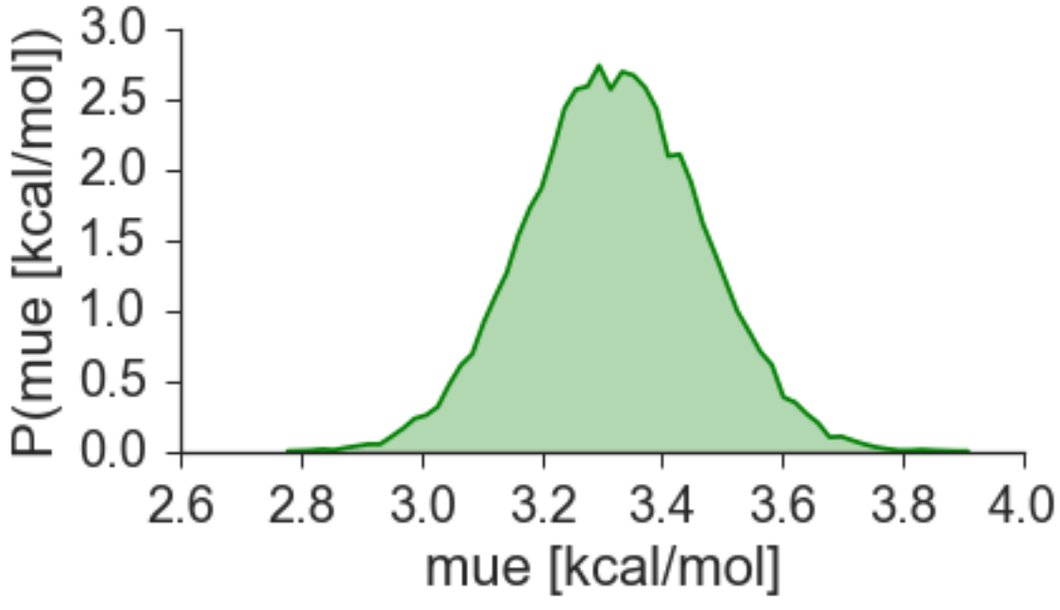
amino pyrimidine derivatives, averages over two runs.

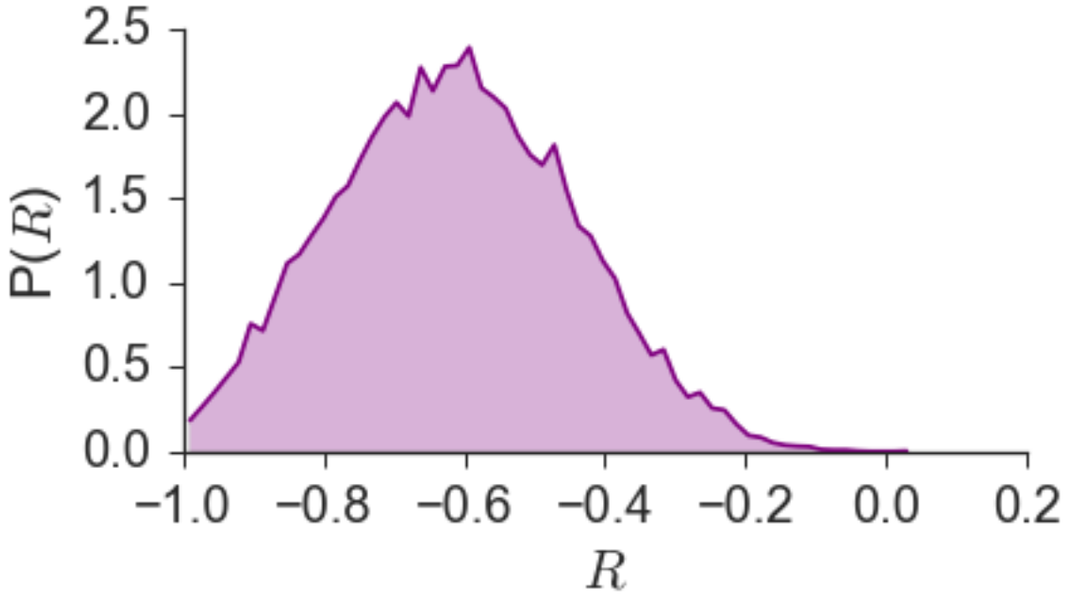


Set 2 — how we did



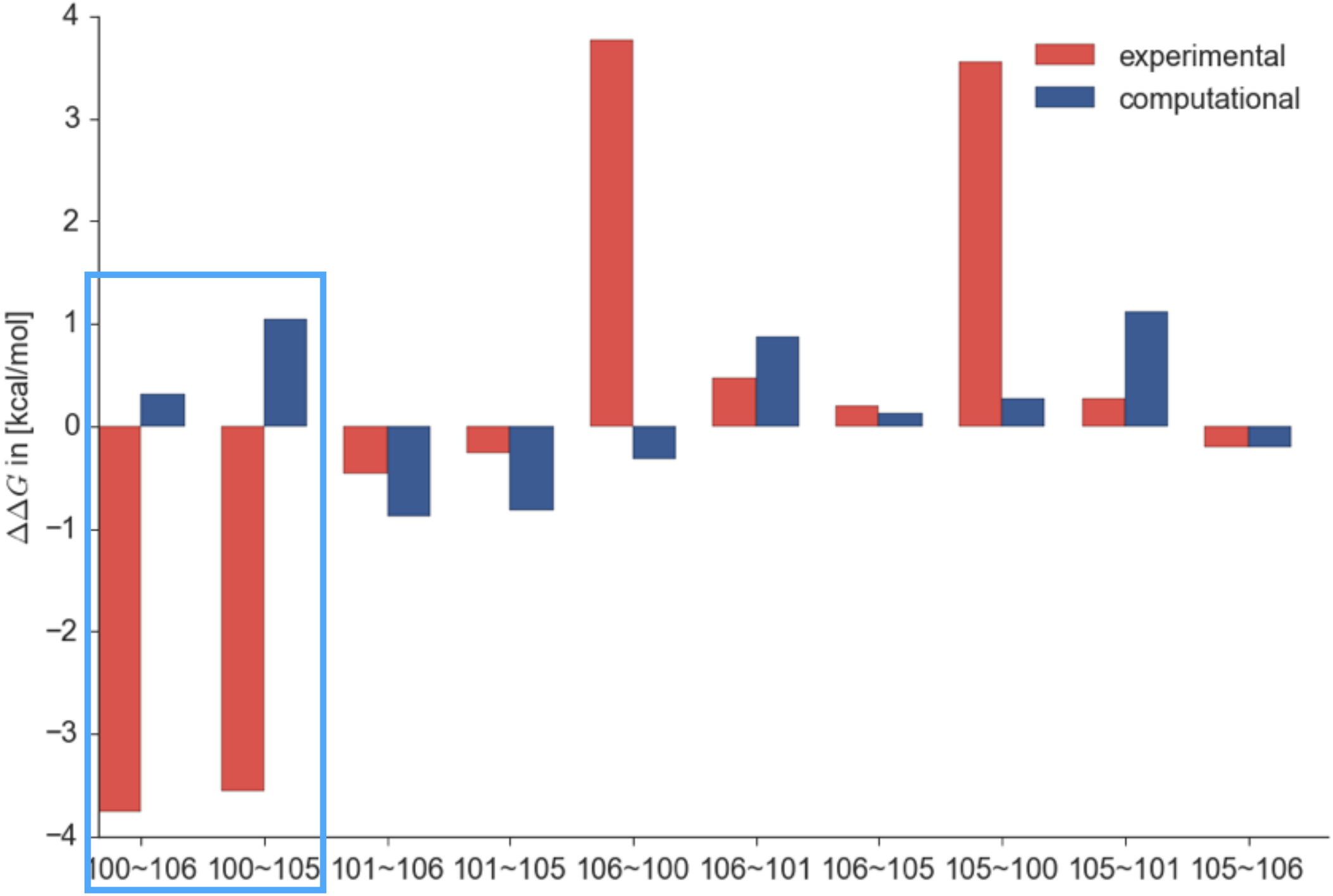






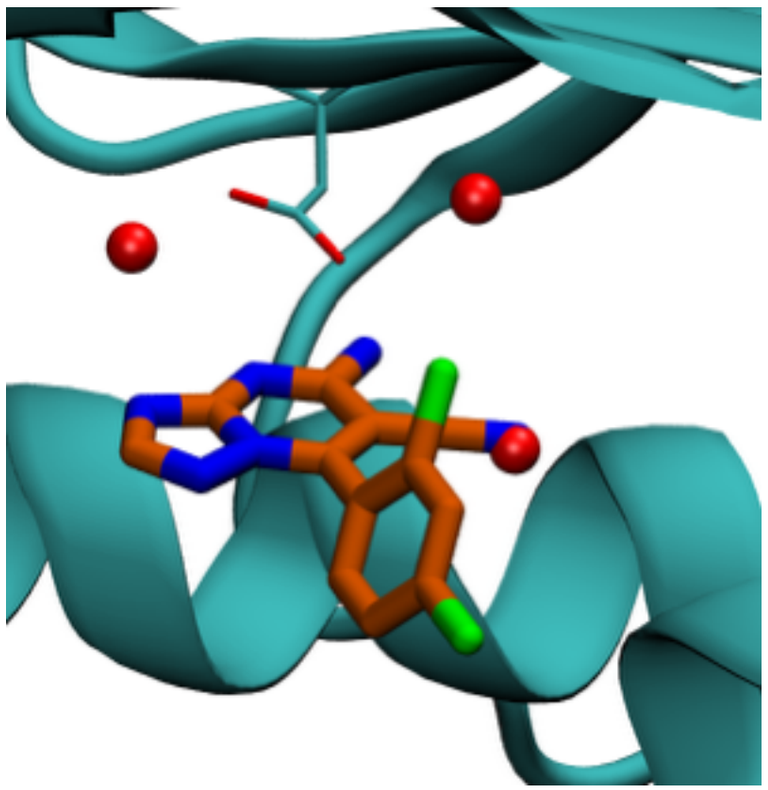
Set 2 — how we did

Averaged, computed and experimental ΔΔG, for simulated pairs.

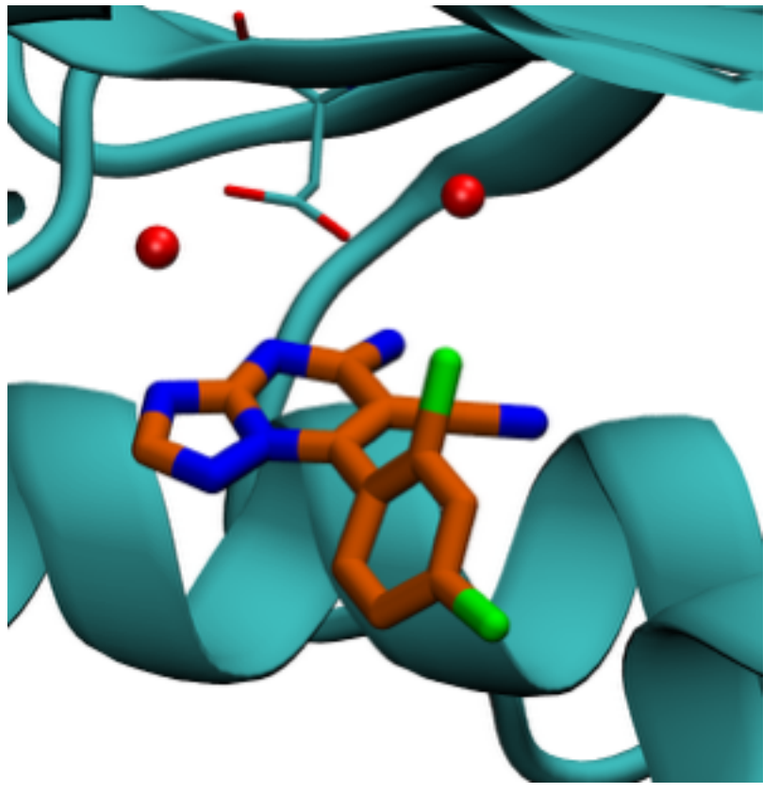


Set 2 — how we did

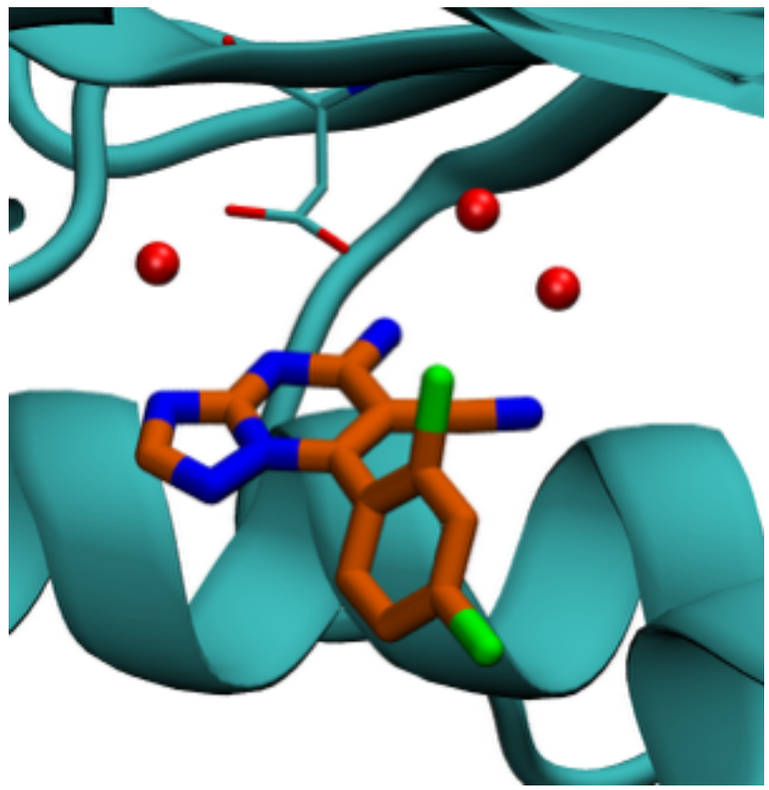
What if we had included crystal water in hsp90_100 starting compounds?



clashing crystal water

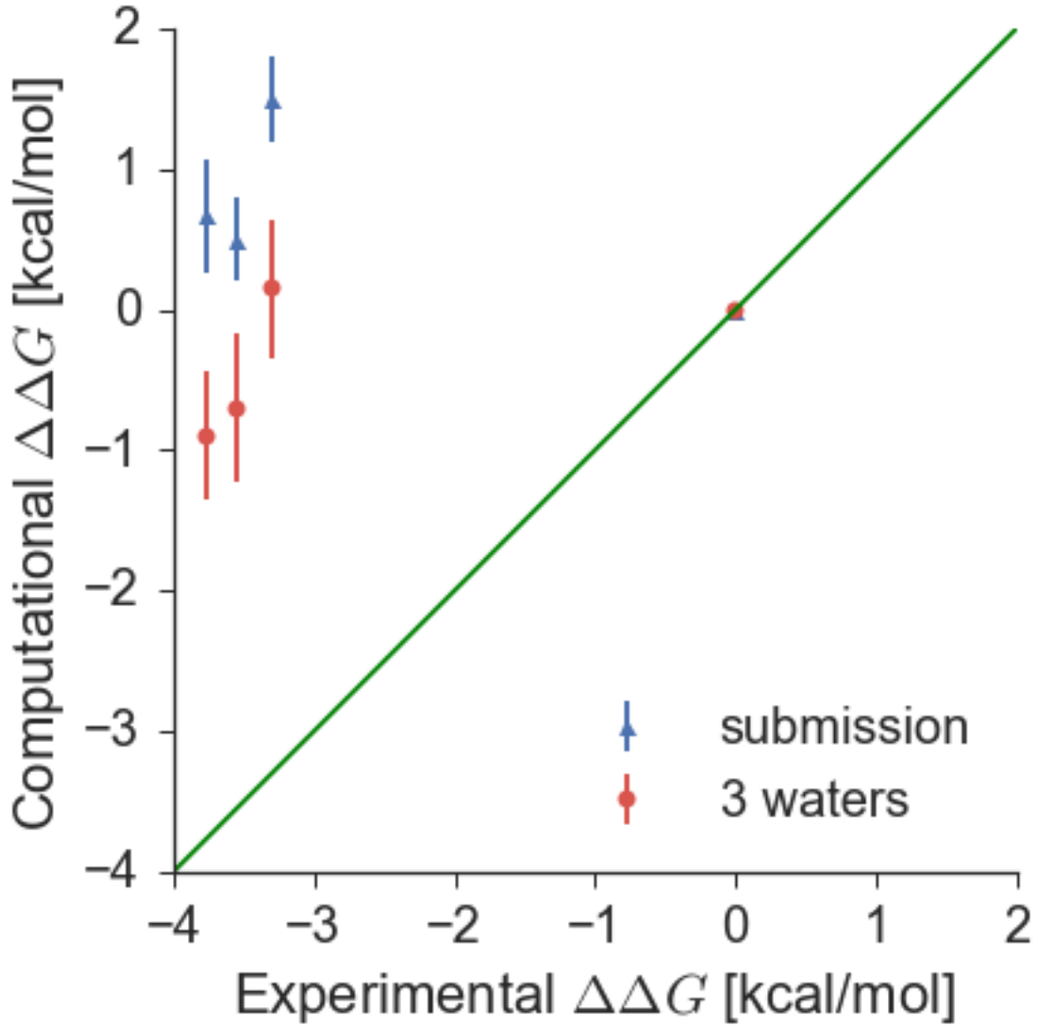


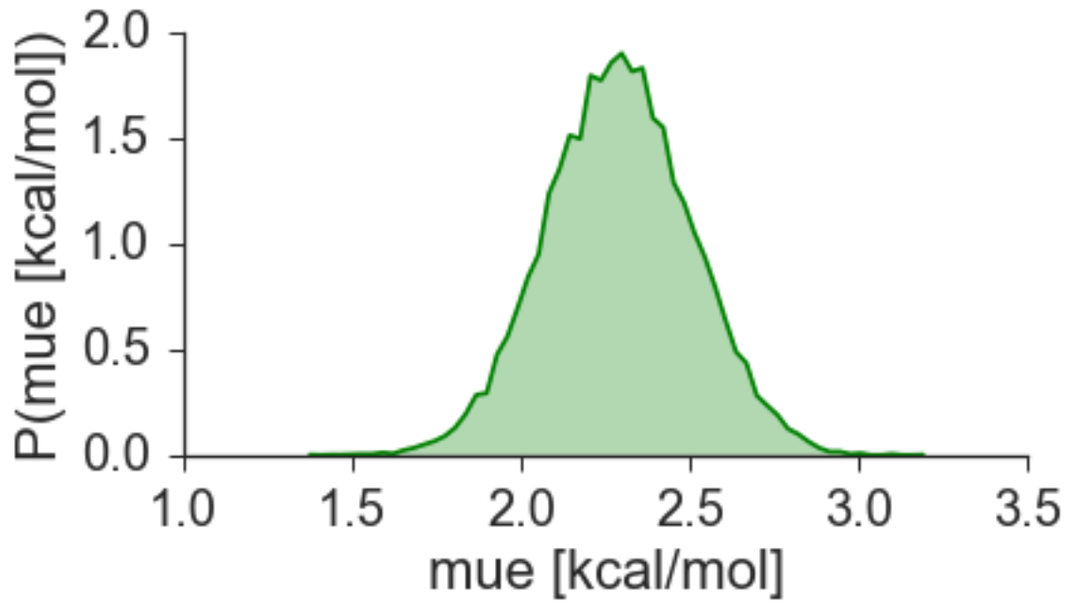
original hsp90_100 starting structure for submission

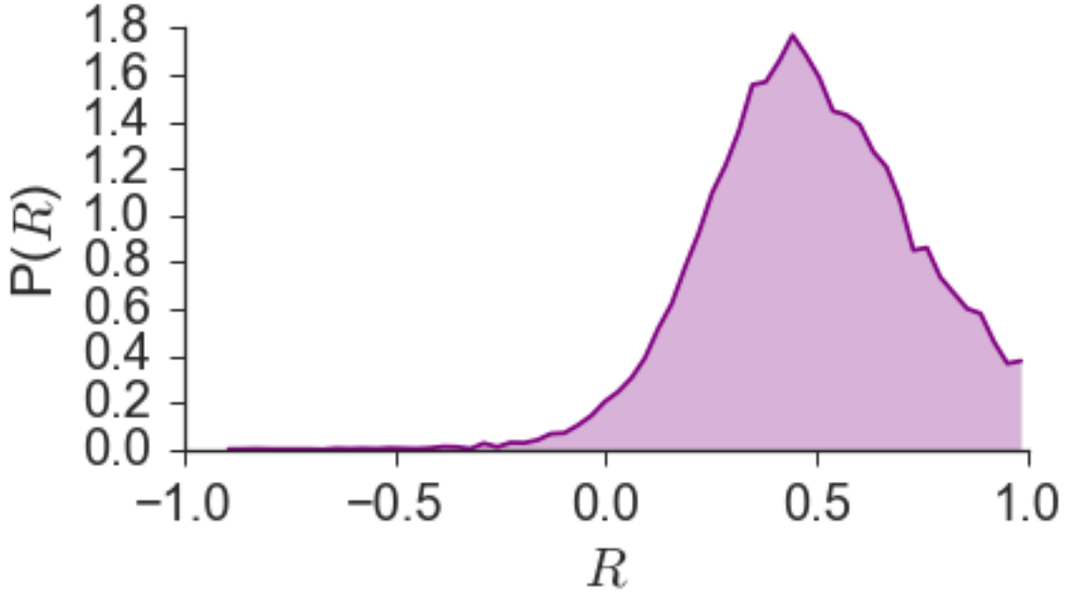


moved and preserved crystal water

Set 2 — how we did





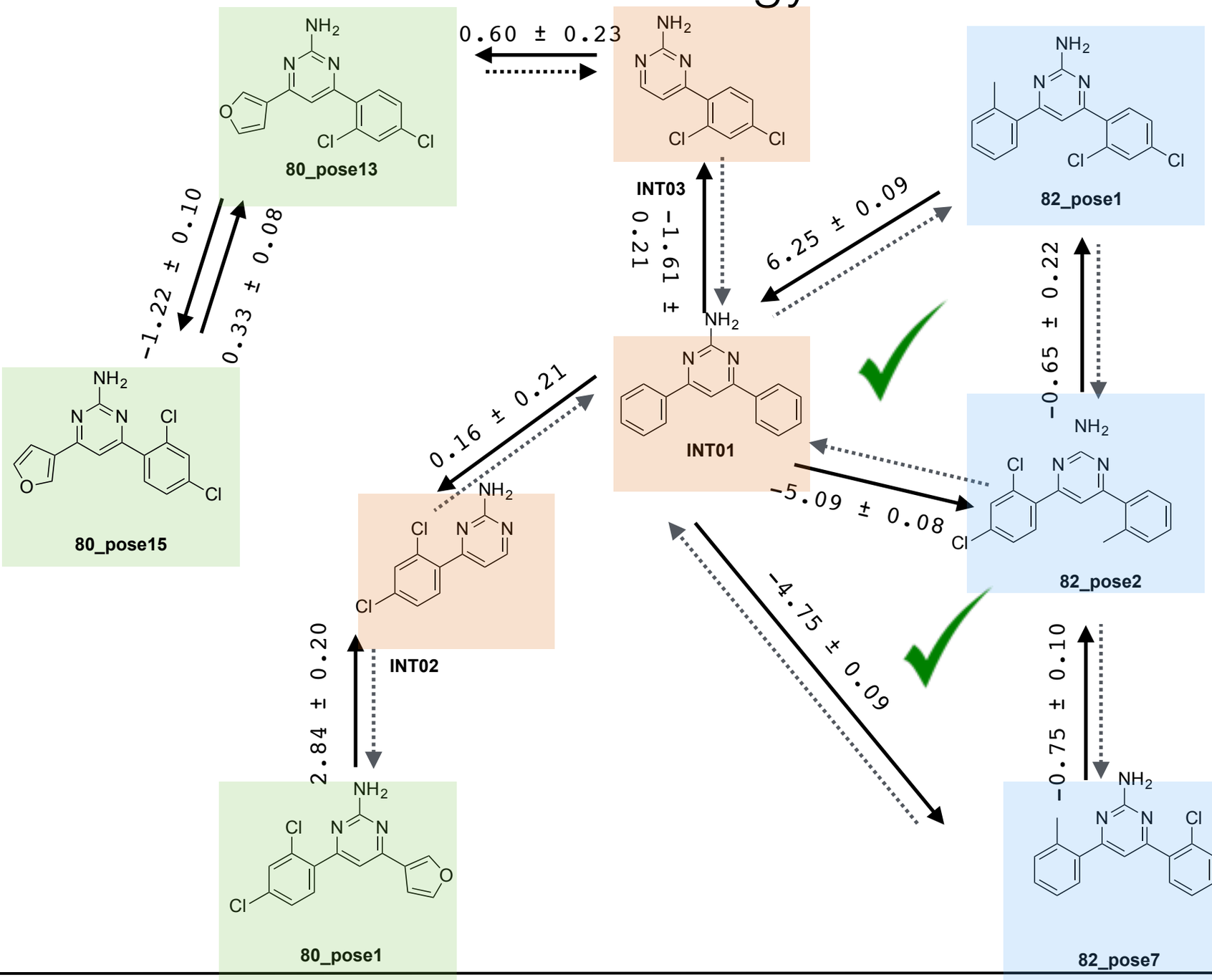


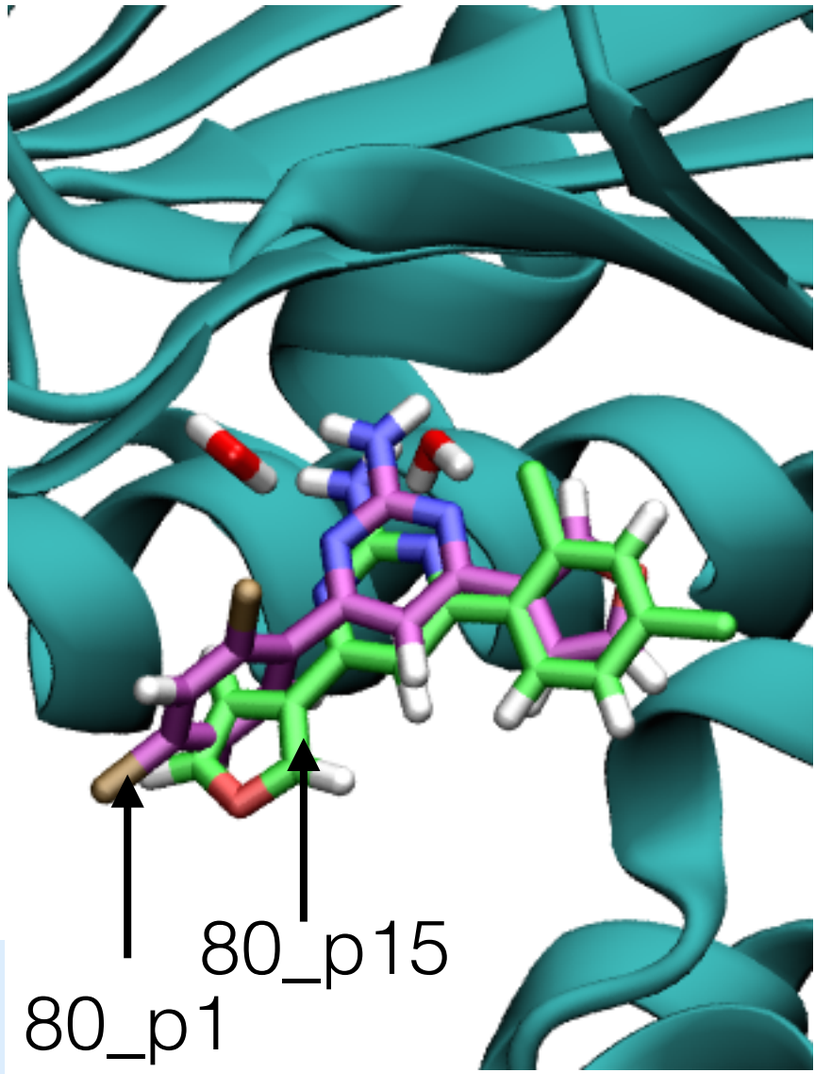
Including 3rd crystal water in calculation improved the predictions, giving an R value of 0.56 and τ = 0.66.

Set 1

From the docking, it was not clear which are the correct poses for dataset 1.

Idea: Use the relative free energy calculations to determine the best pose.

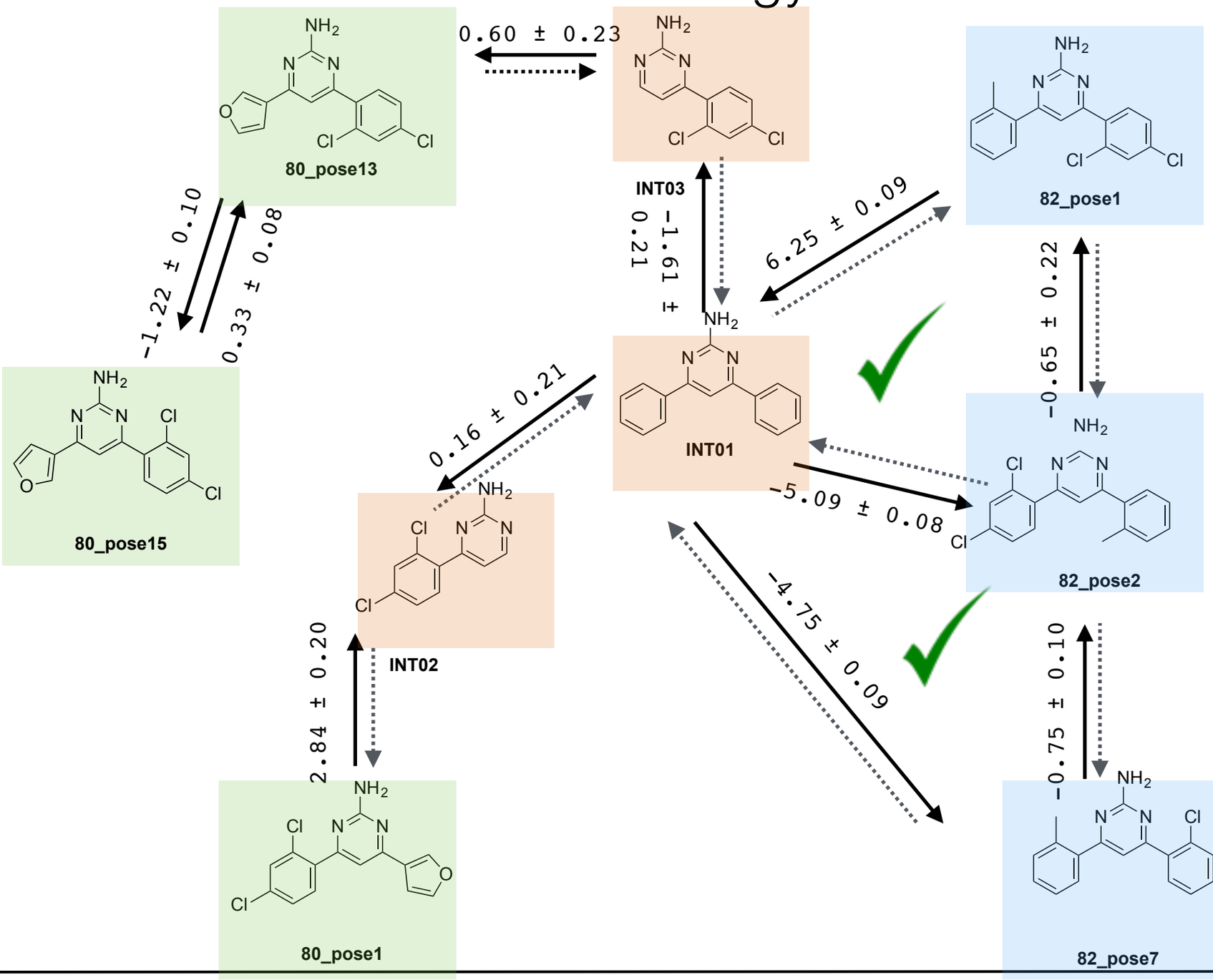




Set 1

From the docking, it was not clear which are the co .

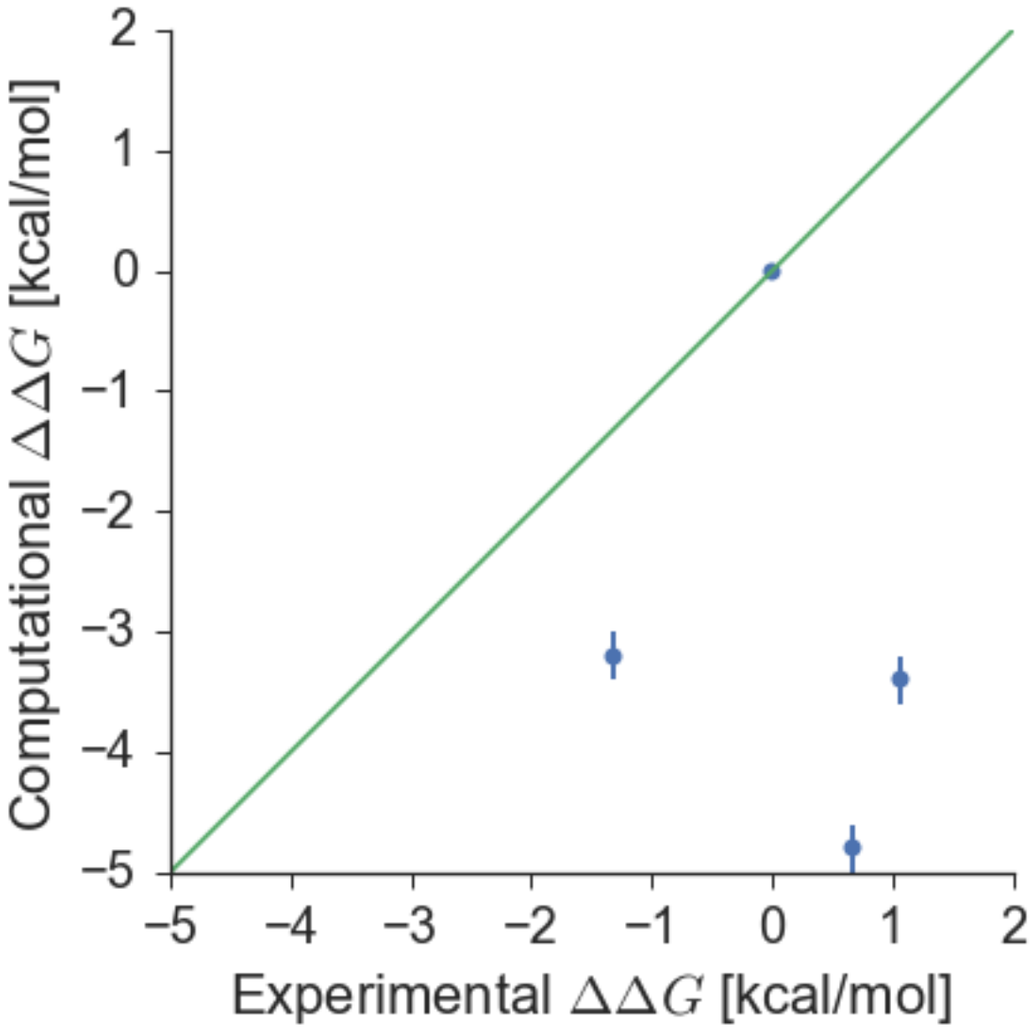
Idea: Use the relative free energy calculations to de

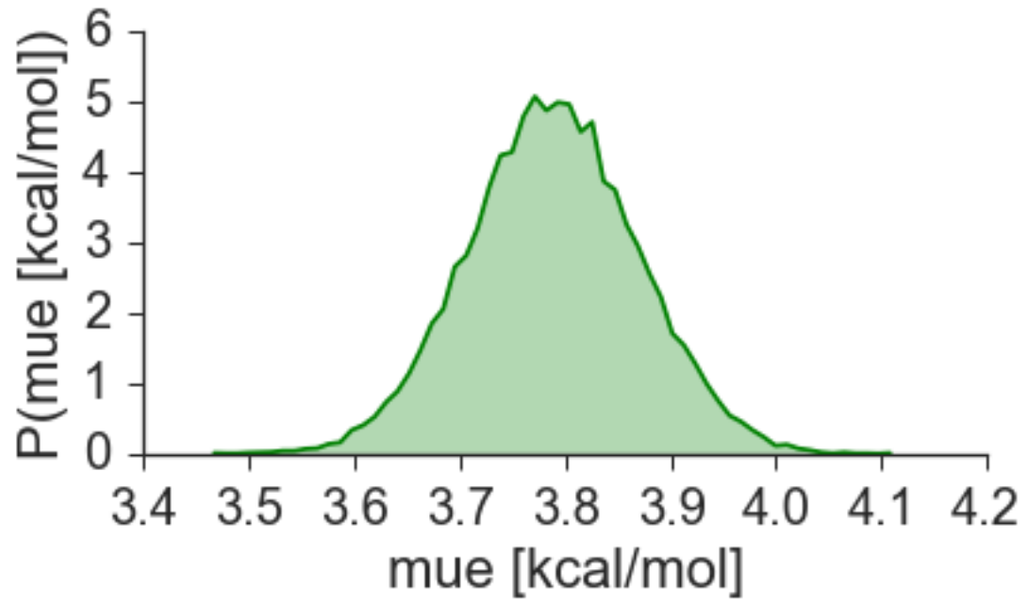


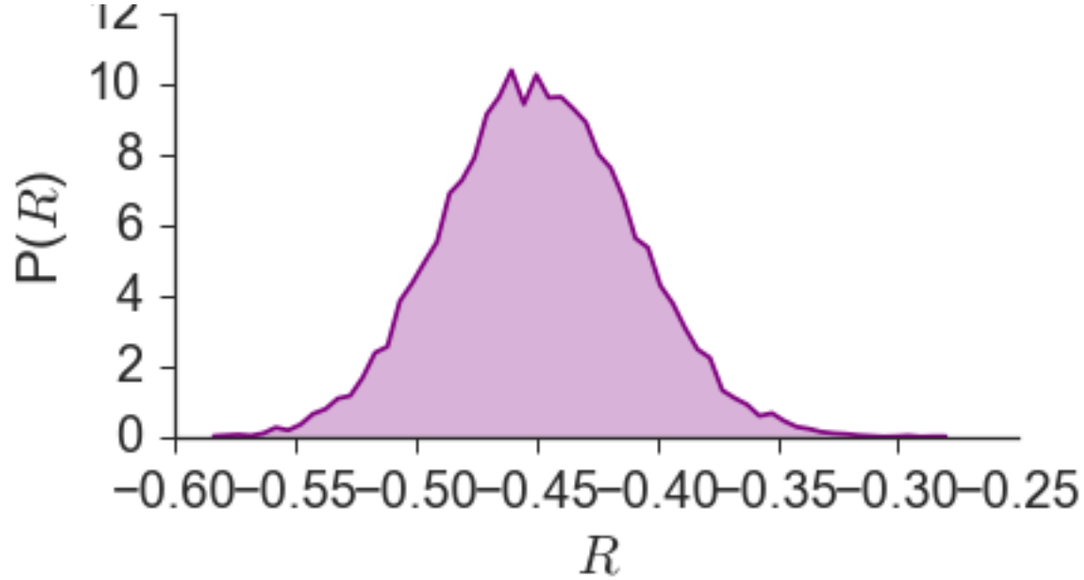
average		
structure	ΔΔG	error
80-p1	0.00	0.00
80-p13	1.66	0.00
80-p15	0.44	0.00
81-p1	-3.19	0.22
81-p3	-2.30	0.22
81-p11	-2.69	0.22
82-p1	-3.37	0.22
82-p2	-2.72	0.22
82-p7	-1.97	0.22
83-p1	-5.94	0.22
83-p17	0.01	0.22
84-p1	-4.84	0.22
84-p6	-4.82	1.26
84-p8	-1.57	1.26
INT01	2.67	0.00
INT02	2.84	0.00
INT03	1.06	0.00

Set 1

Comparing to experiment:







With the actual data given, does a not selected pose perform better?
No, there is no clear pose which should be the right one resulting in better agreement with experimental free energy differences.

Conclusions

Test data from crystal structures perform much better in predicting binding free energies than the D3R challenge datasets.

Conclusions

Test data from crystal structures perform much better in predicting binding free energies than the D3R challenge datasets.

Set 3 binding free energies from relative free energy calculations performed comparatively well to other submissions, but are far from accurate.

Conclusions

Test data from crystal structures perform much better in predicting binding free energies than the D3R challenge datasets.

Set 3 binding free energies from relative free energy calculations performed comparatively well to other submissions, but are far from accurate.

Overall, ranking/free energy prediction in a relative free energy calculation can be **skewed from one badly estimated compound**, i.e here the missing water in **Set 2**.

Conclusions

Test data from crystal structures perform much better in predicting binding free energies than the D3R challenge datasets.

Set 3 binding free energies from relative free energy calculations performed comparatively well to other submissions, but are far from accurate.

Overall, ranking/free energy prediction in a relative free energy calculation can be **skewed from one badly estimated compound**, i.e here the missing water in **Set 2**.

Set 1: With unknown correct binding poses, the protocol failed to estimate the correct pose from the free energy calculation.

Conclusions

Test data from crystal structures perform much better in predicting binding free energies than the D3R challenge datasets.

Set 3 binding free energies from relative free energy calculations performed comparatively well to other submissions, but are far from accurate.

Overall, ranking/free energy prediction in a relative free energy calculation can be **skewed from one badly estimated compound**, i.e here the missing water in **Set 2**.

Set 1: With unknown correct binding poses, the protocol failed to estimate the correct pose from the free energy calculation.

High variability in correct predictions despite employing identical simulation protocols.

Acknowledgments

Jordi Juárez-Jiménez

Docking

Alexis Hennessy

Julien Michel

The rest of the Michel group.

Funding:

European Research Council
Established by the European Commission
Supporting top researchers from anywhere in the world

Questions?

www.xkcd.com