

Is structure based drug design ready to predict selectivity?

GSS

12/14/17



STEVEN ALBANESE
CHODERA LAB // MSKCC

What is the selectivity of a compound?

Selectivity on a broad scale

$$S_{\text{inhib}}(3\mu\text{M}) = \frac{\# \text{ kinases bound w/ } K_D < 3\mu\text{M}}{\text{total kinases tested}}$$

Davis et. al. Nat. Biotechnol. 29:11, 2011

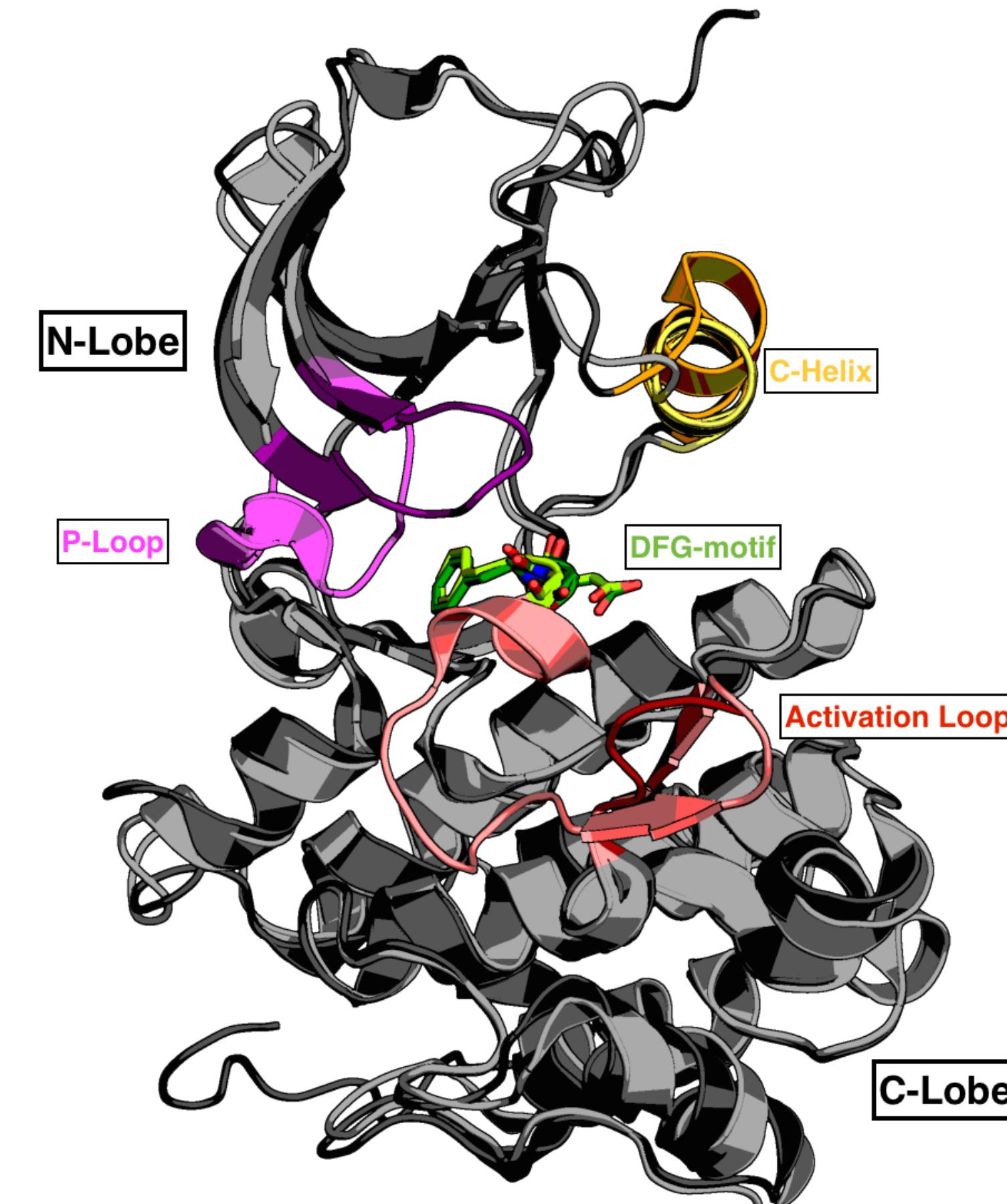
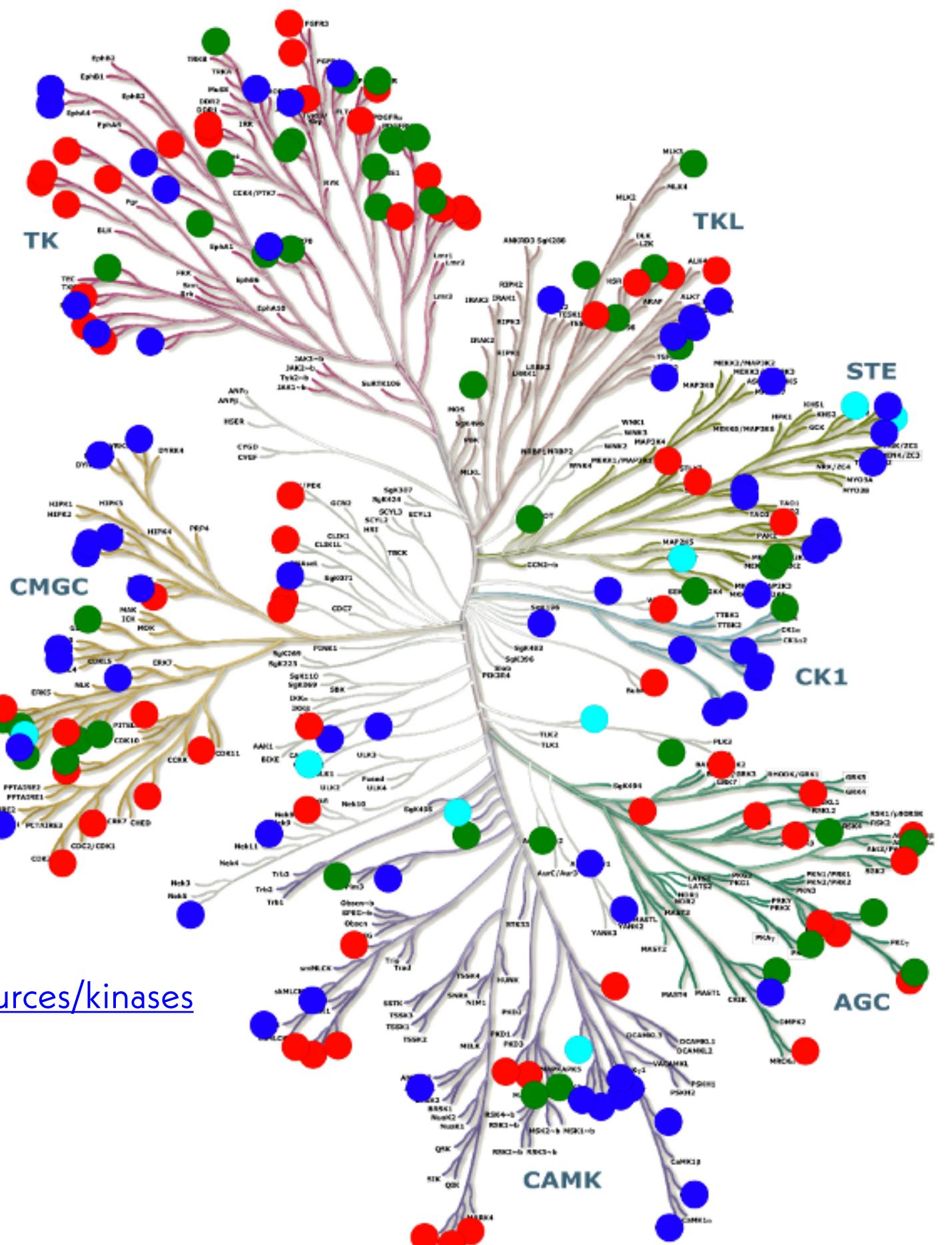
Selectivity on a smaller scale

$$\Delta\Delta G_{\text{selectivity}} = \Delta G_{\text{bind},1} - \Delta G_{\text{bind},2}$$

Kinase inhibitors have potential for a wide range of selectivities

518 human kinases...

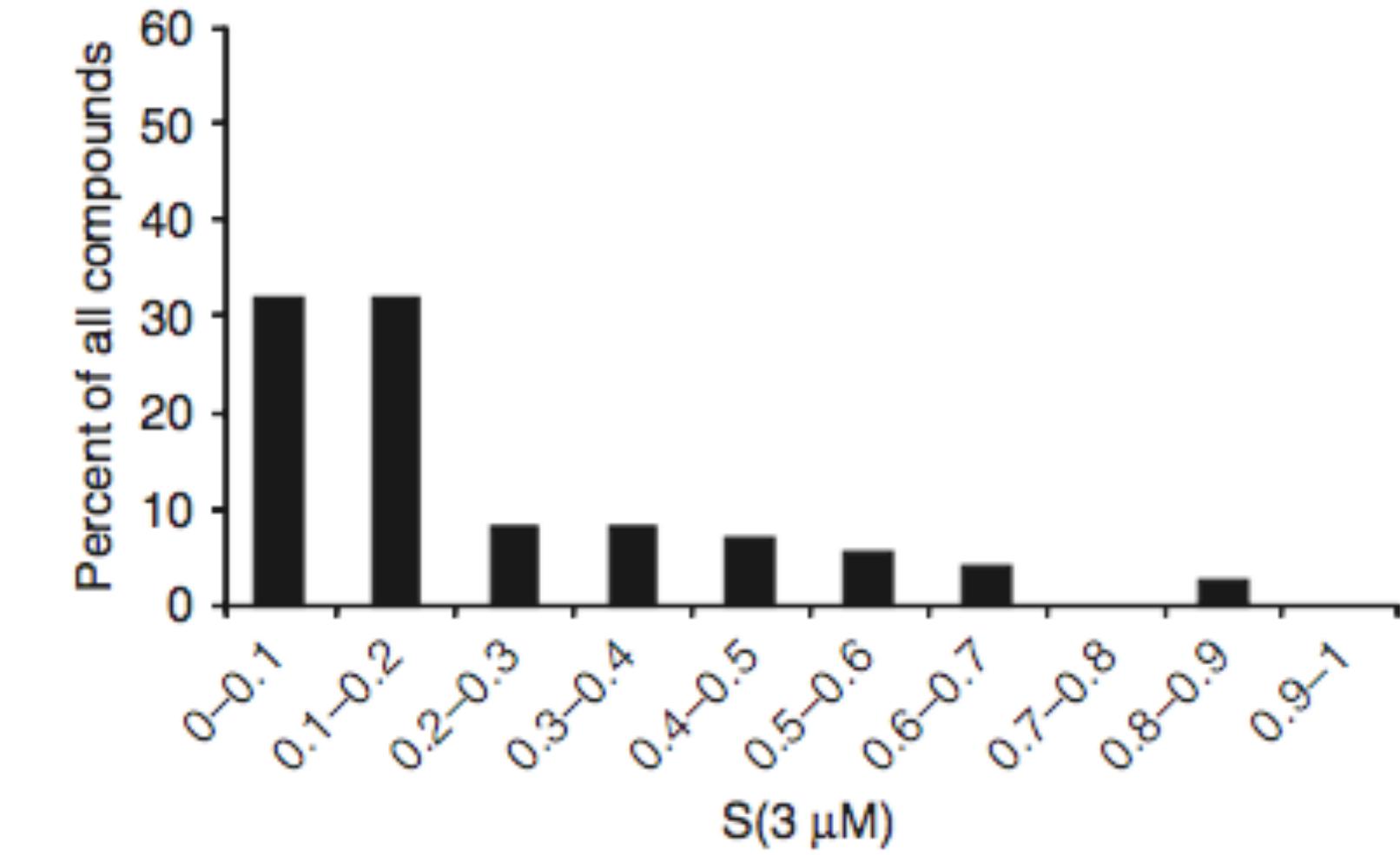
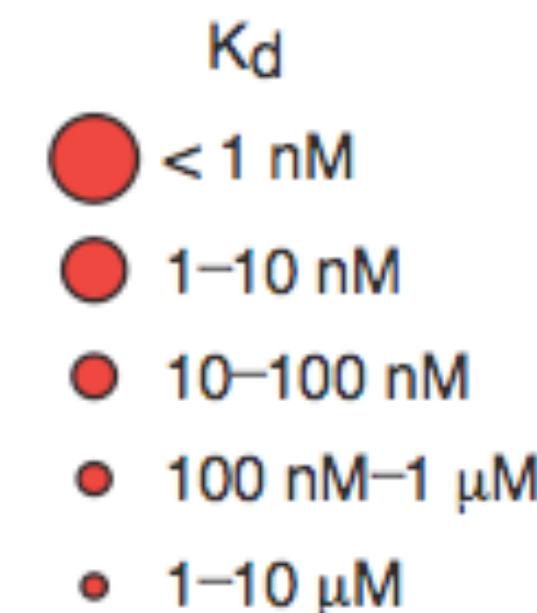
...with a shared fold



PDB: 2OIQ, 2YHH

<http://www.thesgc.org/scientists/resources/kinases>

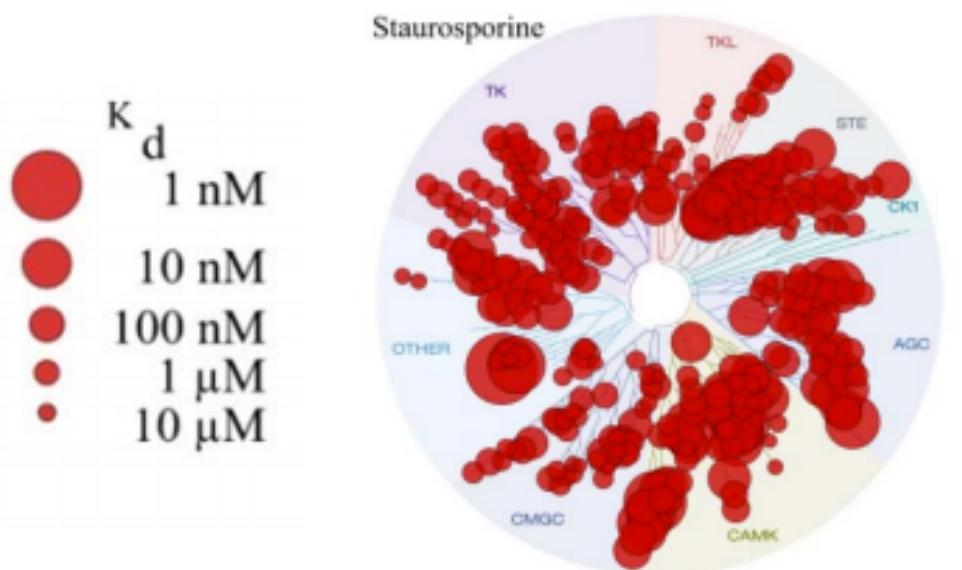
Kinase inhibitors have potential for a wide range of selectivities



$$S_{\text{inhib}}(3 \mu\text{M}) = \frac{\# \text{ kinases bound w/ } K_D < 3 \mu\text{M}}{\text{total kinases tested}}$$

Davis et. al. Nat. Biotechnol. 29:11, 2011

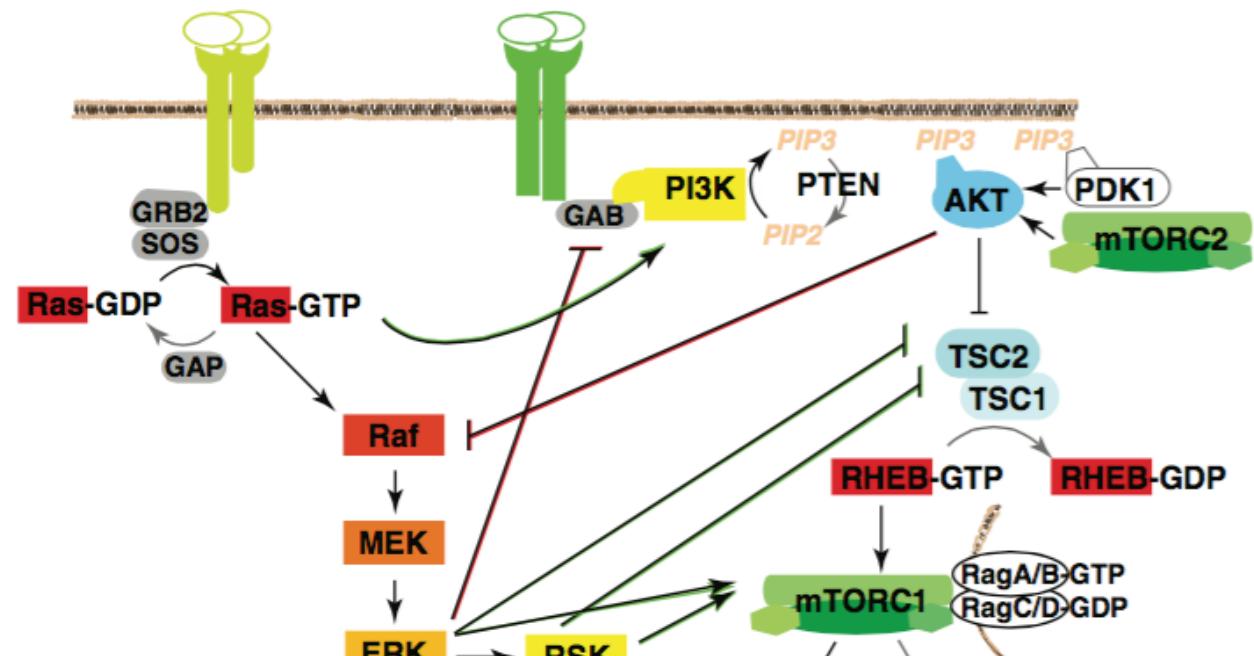
Selectivity is considered for a number of different reasons



Avoid harmful **off-target toxicity** by targeting kinase of interest



Minimize **on-target toxicity** by targeting oncogenic mutant form of kinase

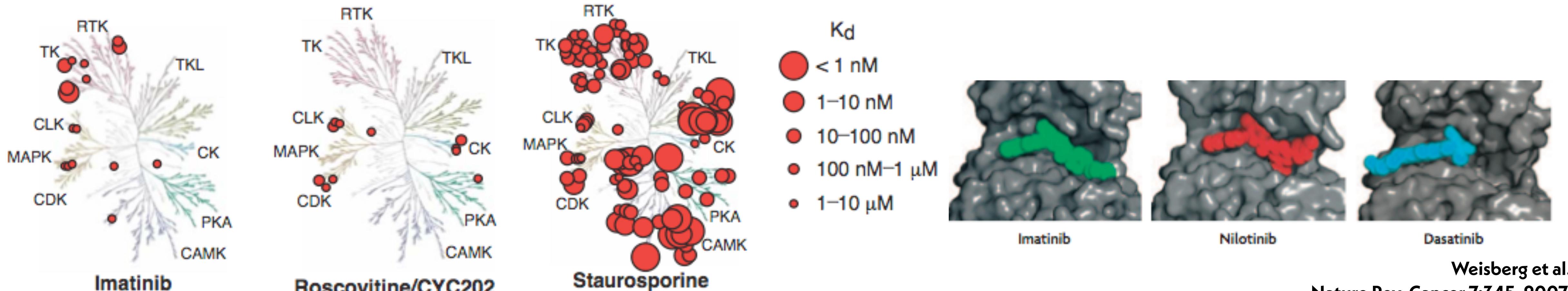


Handle **complex networks of kinase signaling** by targeting multiple kinases

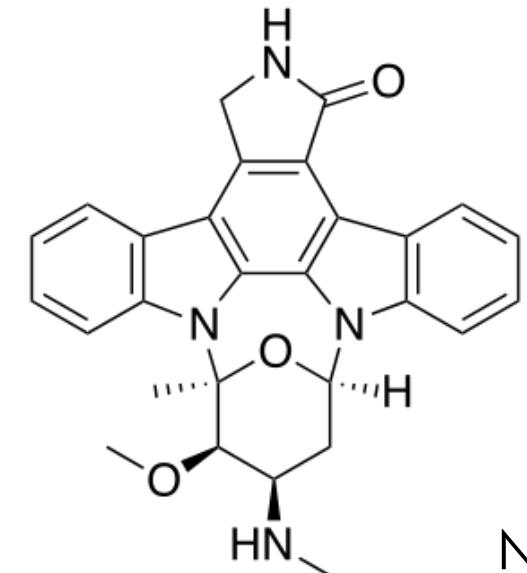
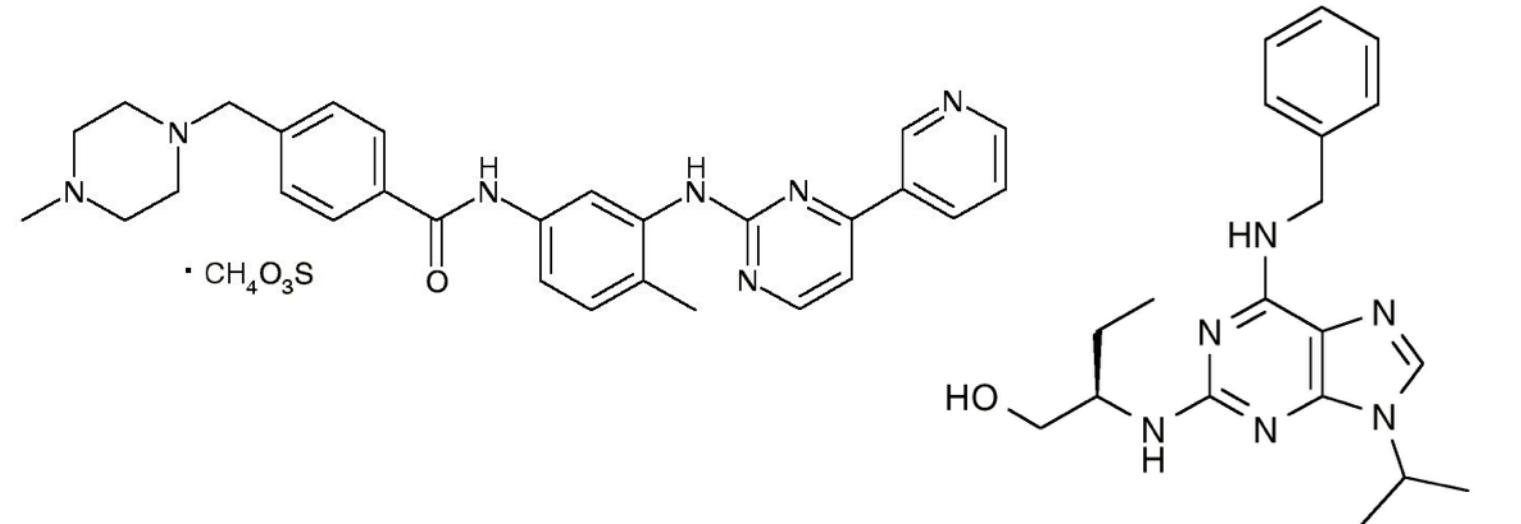
Mendoza et. al. Trends Biochem. Sci. 36:6, 2011

What determines the selectivity of kinase inhibitors?

- * Are particular ligand **scaffolds** privileged with specificity?
- * Are particular **binding modes** better for specificity?
- * Are certain **kinases** inherently more promiscuous?



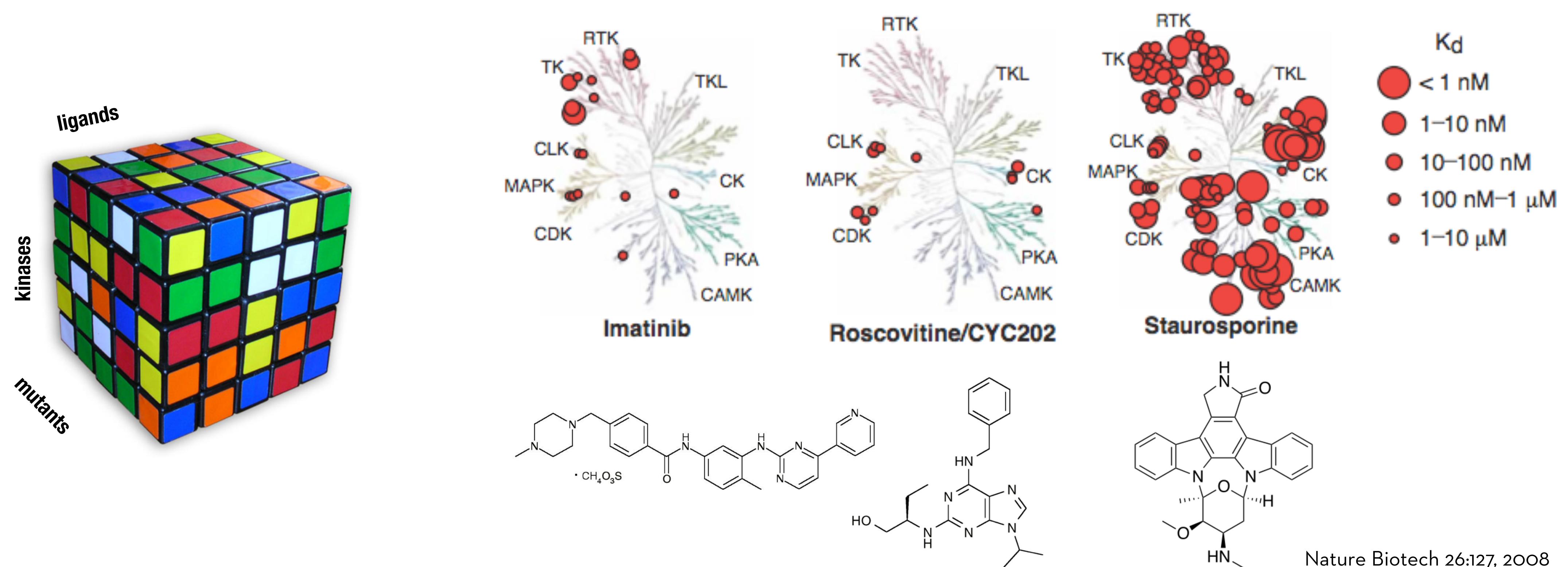
Weisberg et al.
Nature Rev. Cancer 7:345, 2007.



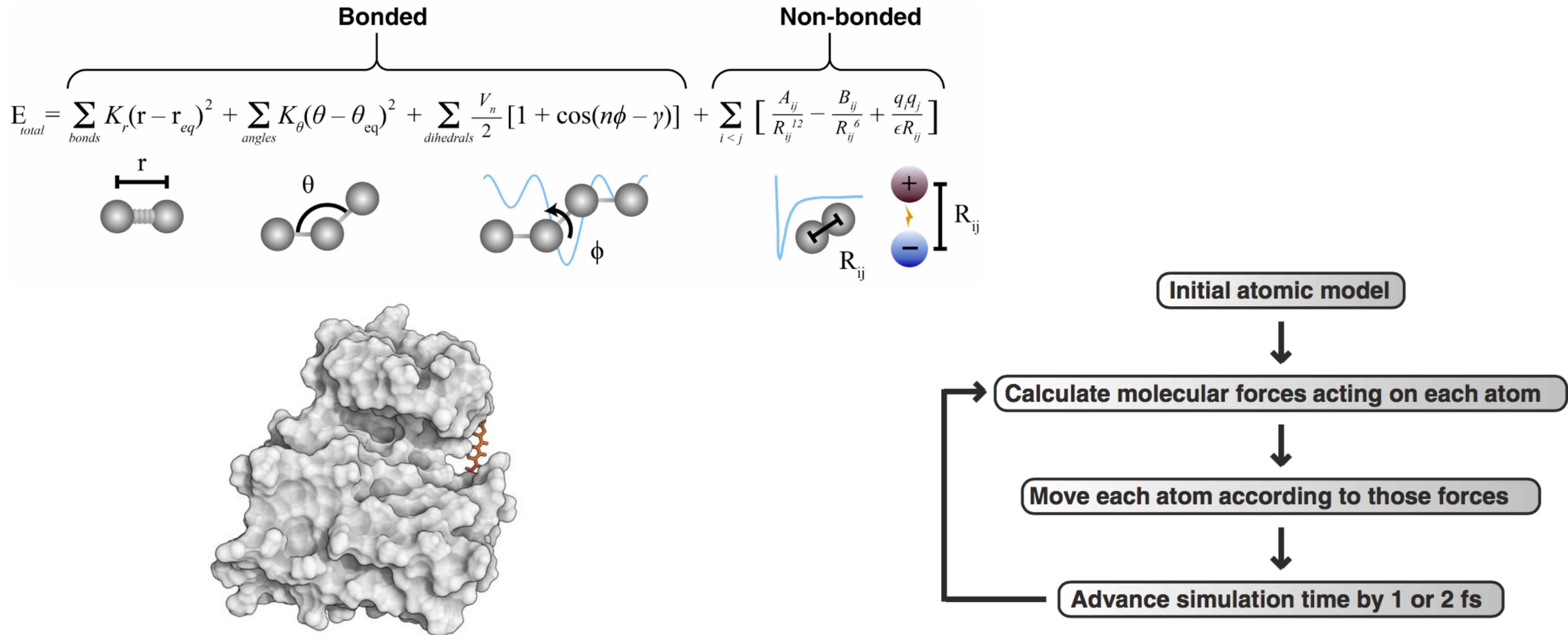
Nature Biotech 26:127, 2008

What determines the selectivity of kinase inhibitors?

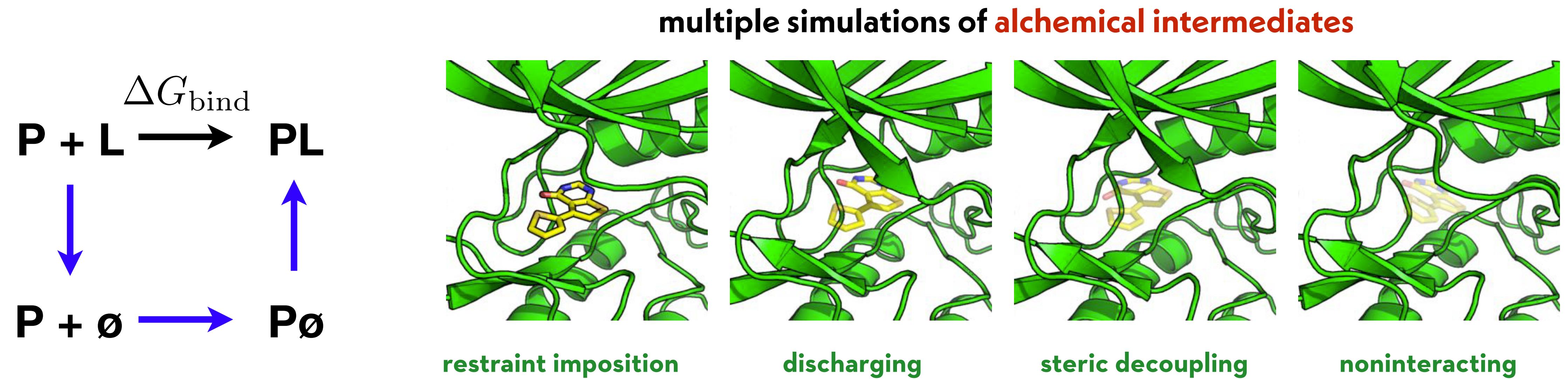
High-throughput fluorescence measurements and free energy calculations can address physical determinants of kinase inhibitor selectivity:



Free energy calculations are a type of atomistic simulation

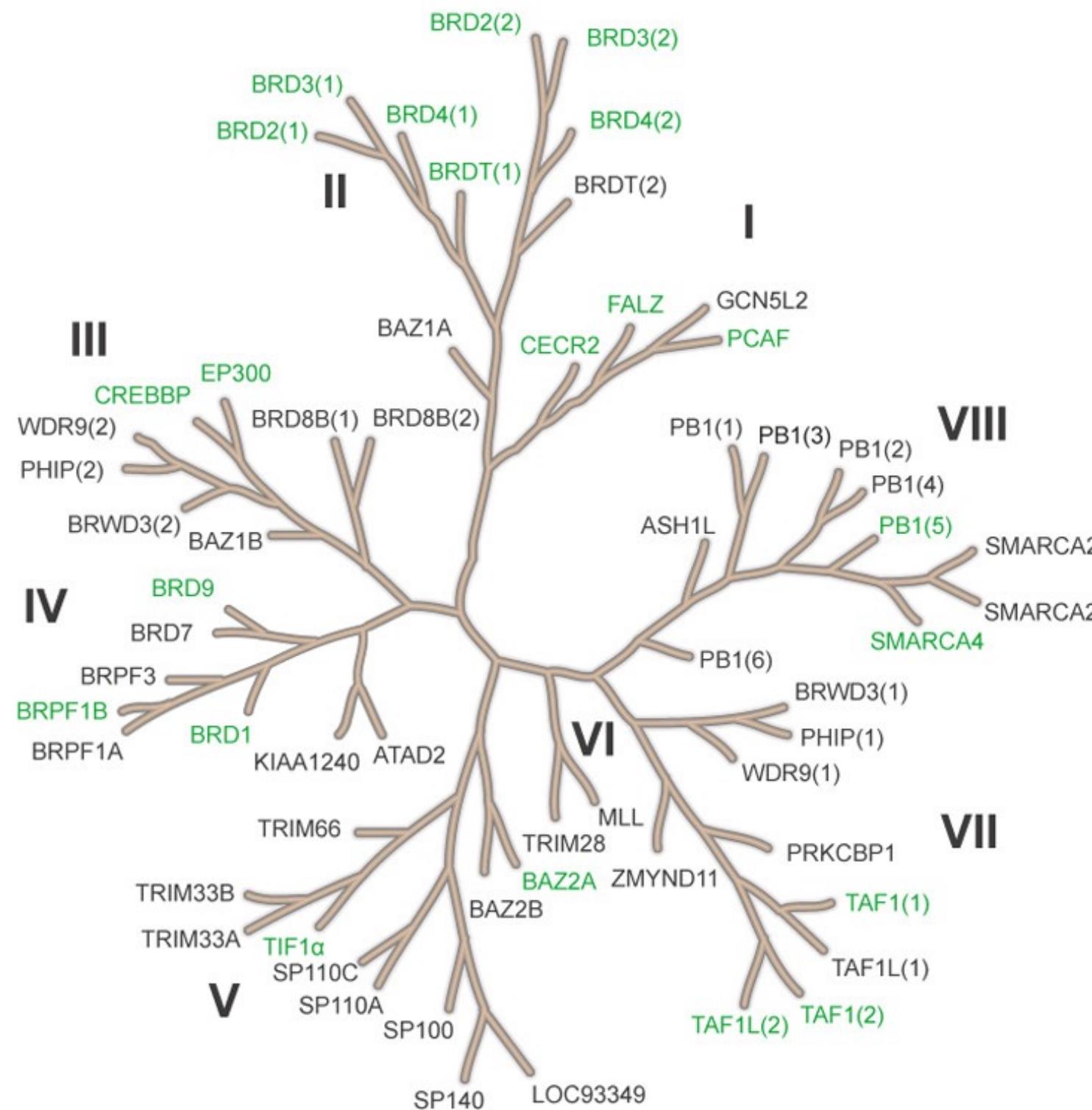
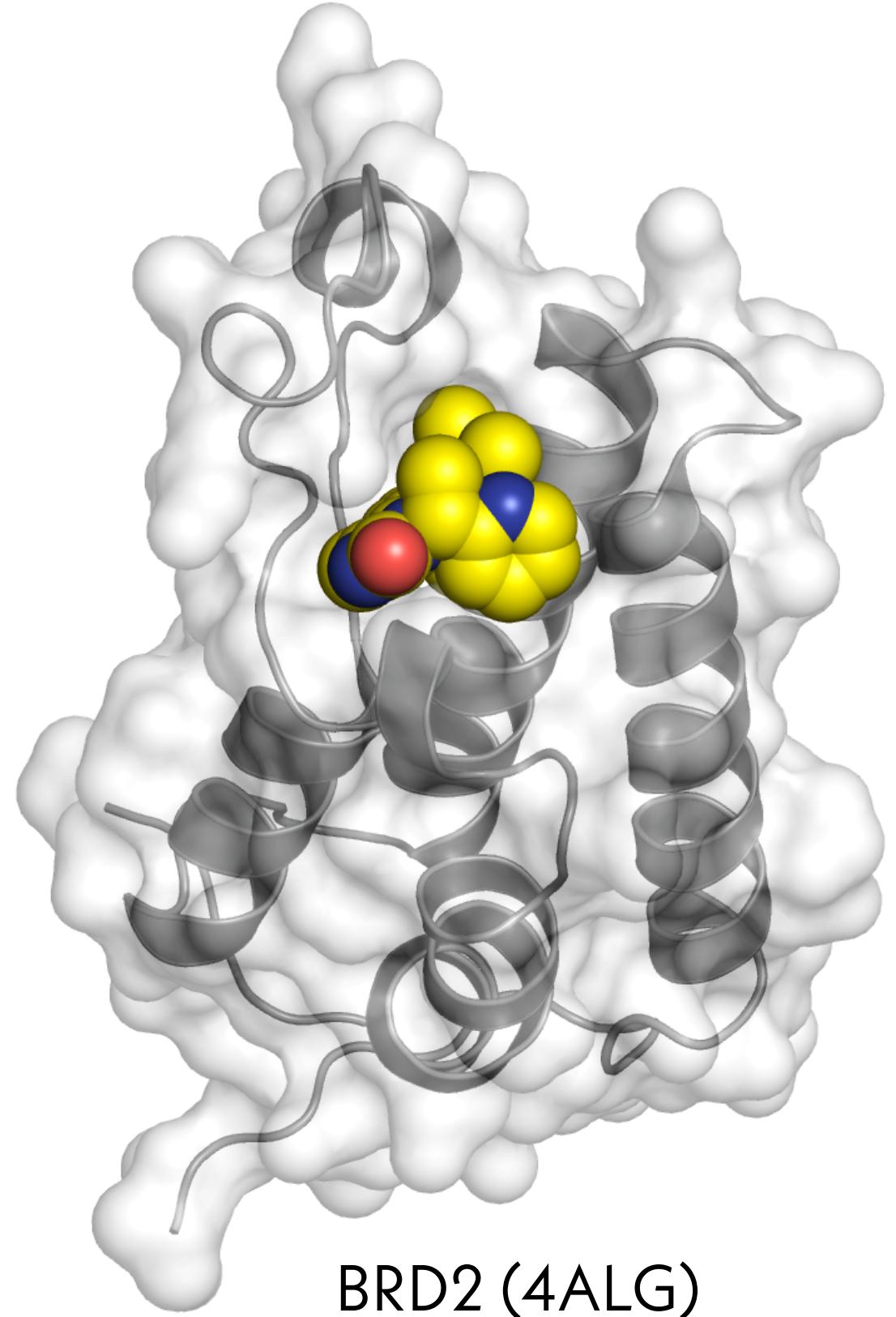


alchemical free energy calculations provide a rigorous way to efficiently compute binding affinities



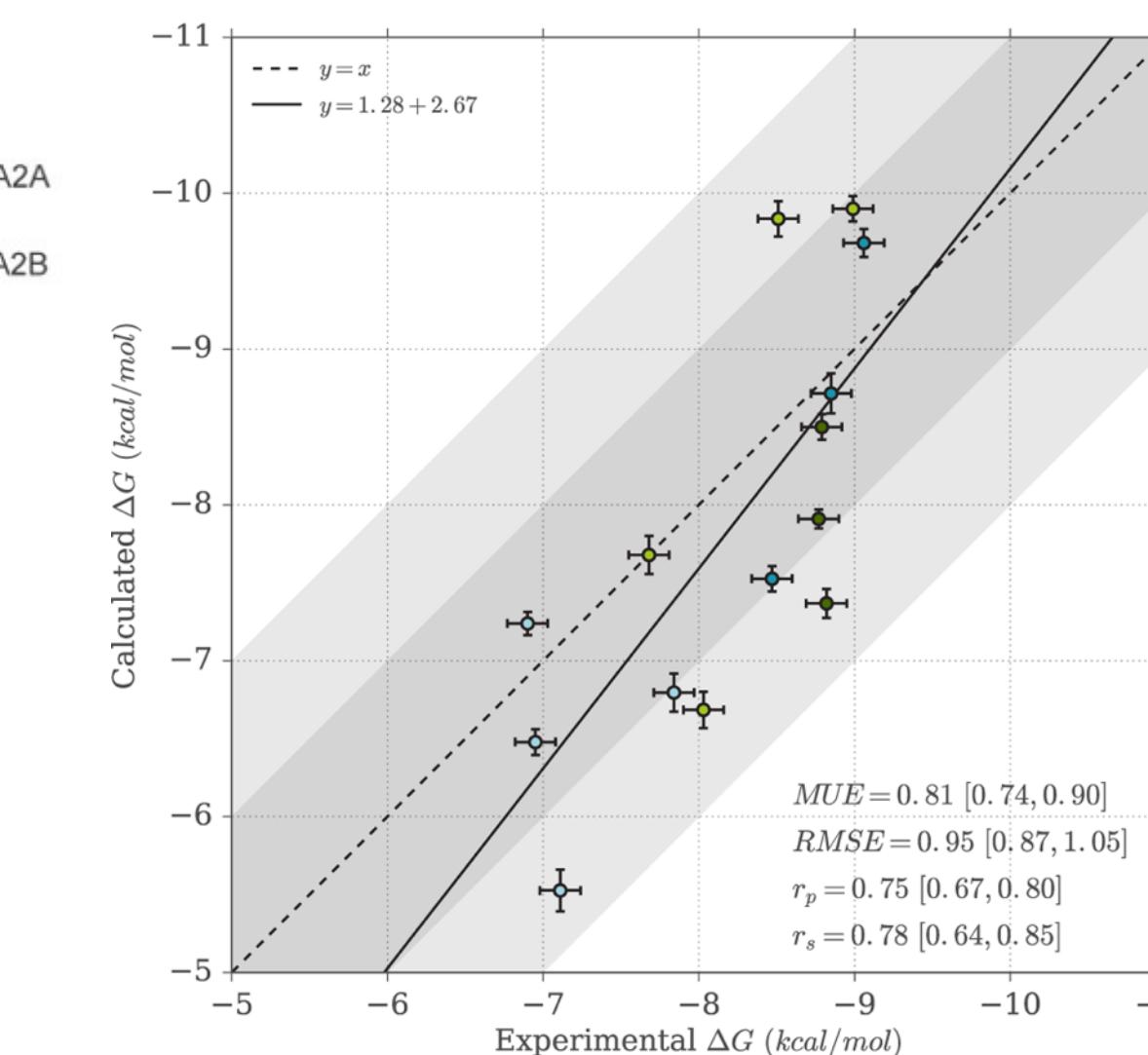
These calculations come in two flavors: absolute (where the entire ligand is alchemically modified) and relative (where shared atoms are NOT alchemically modified)

Absolute free energy calculations can predict the potency of a compound

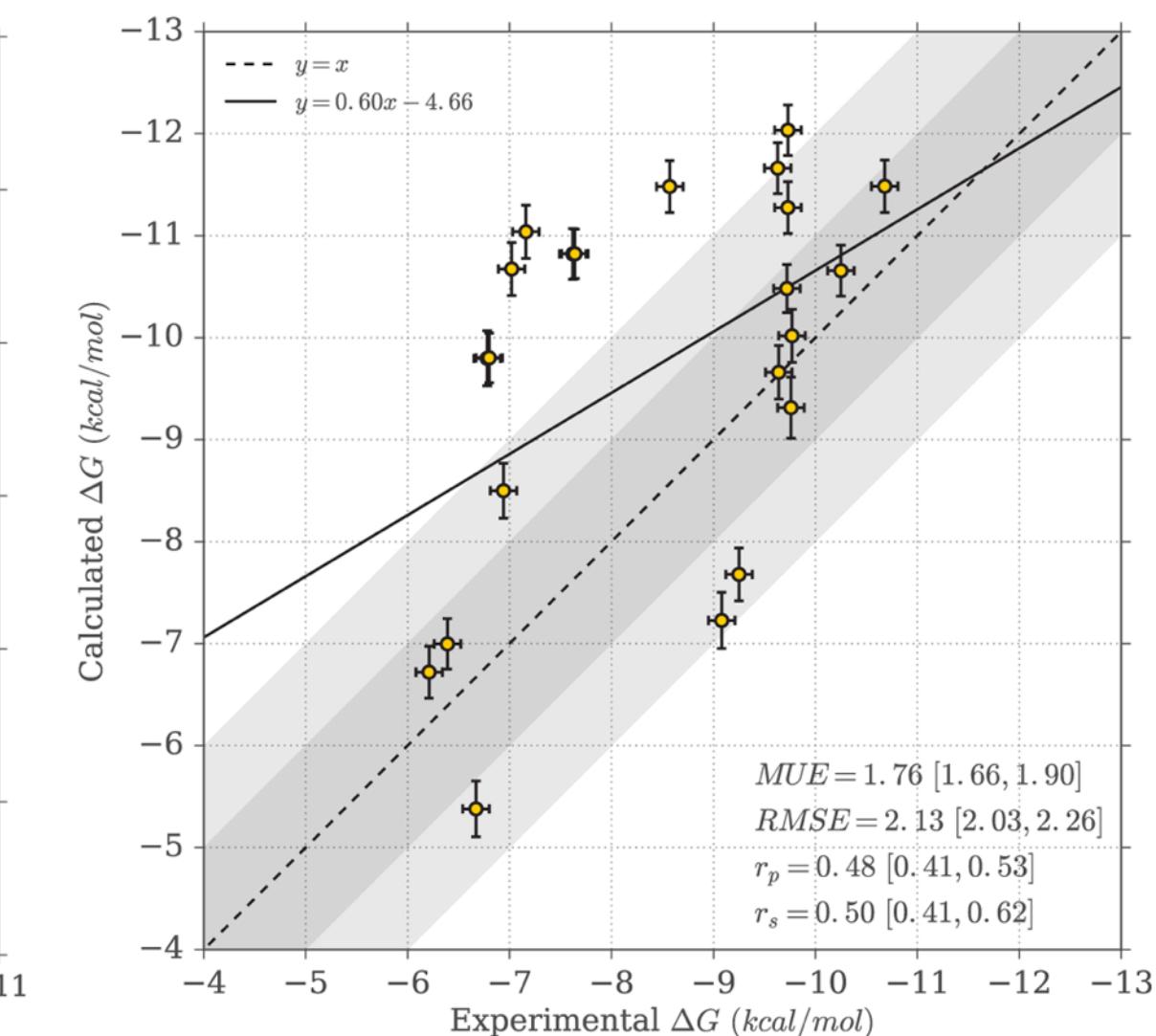


Absolute free energy calculations

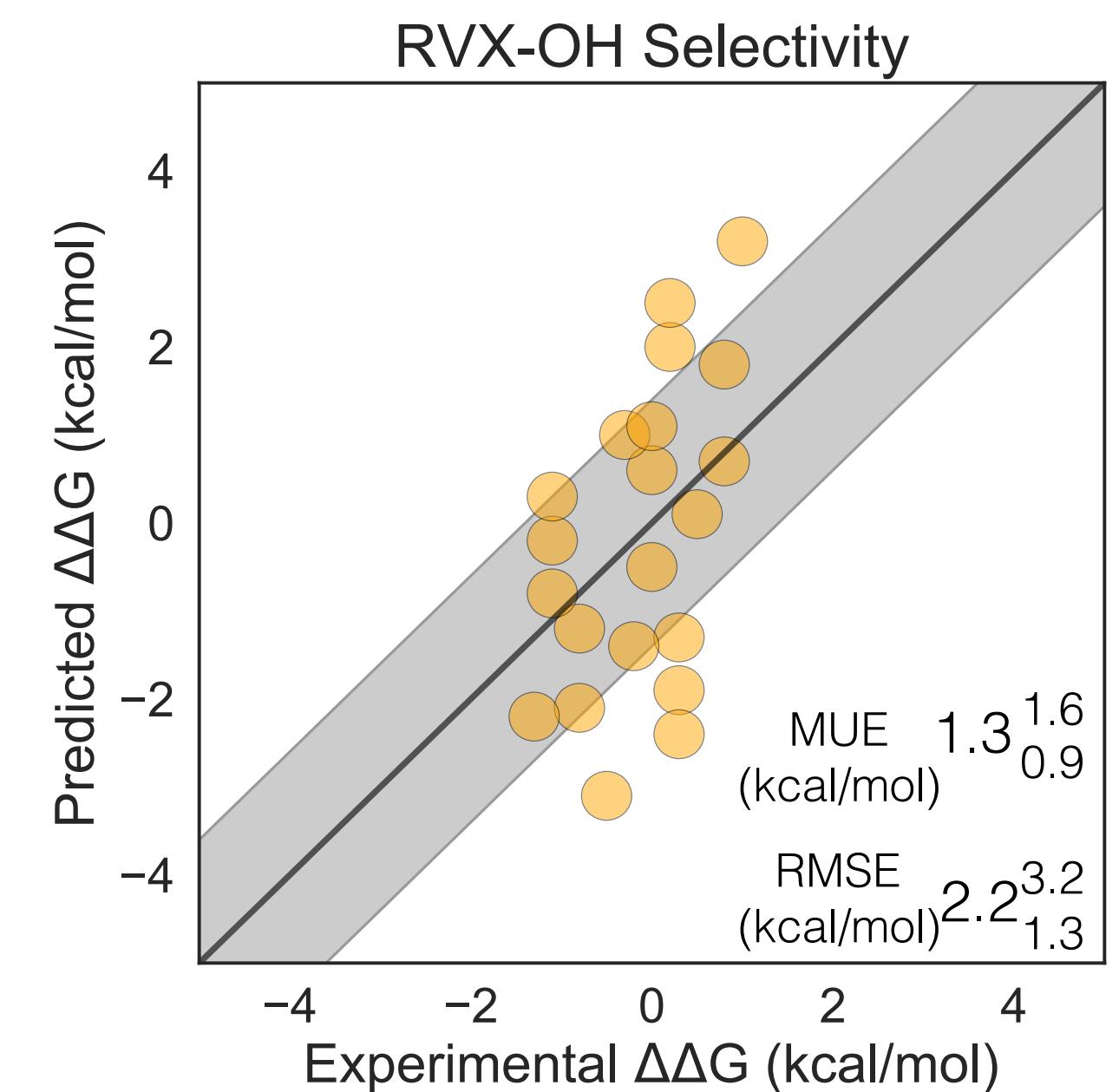
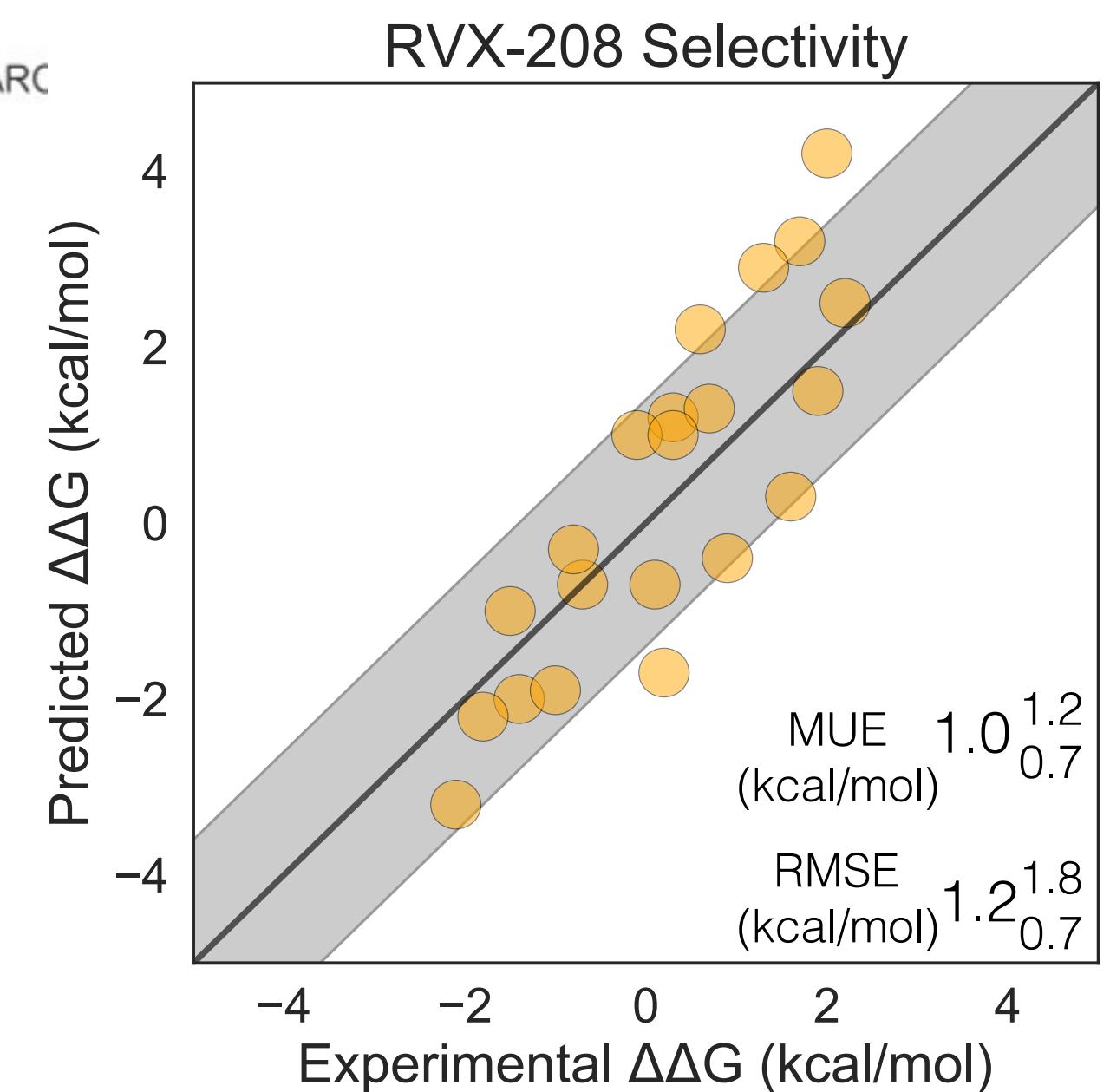
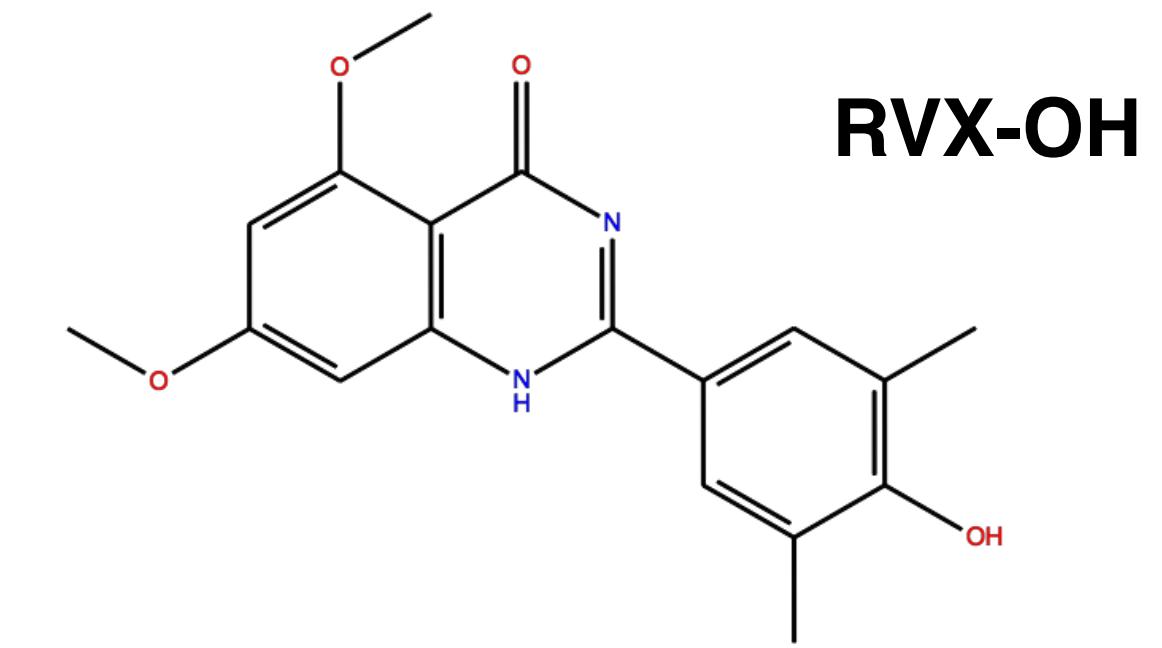
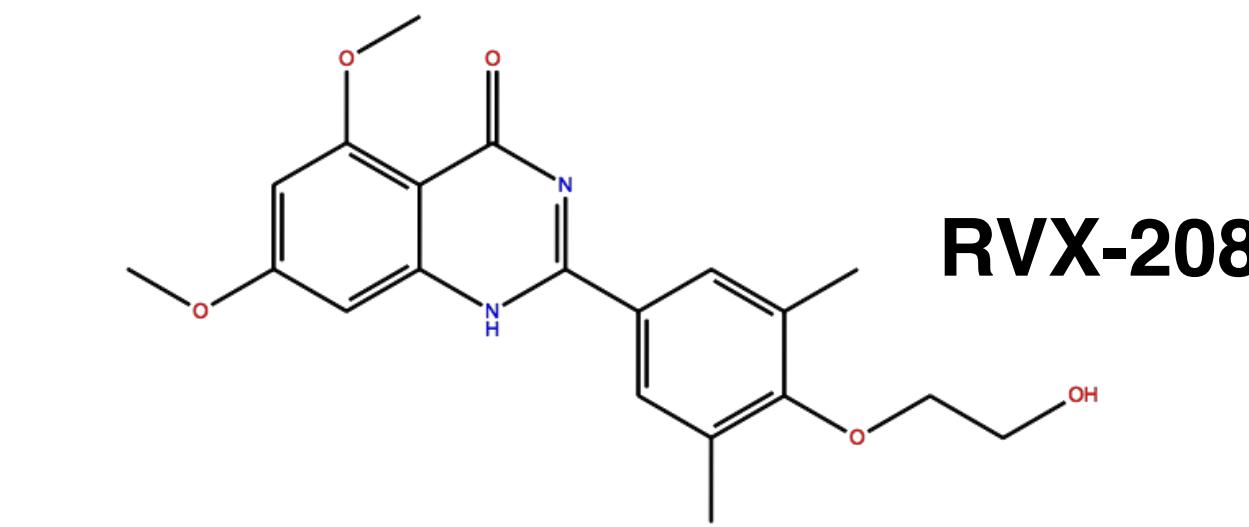
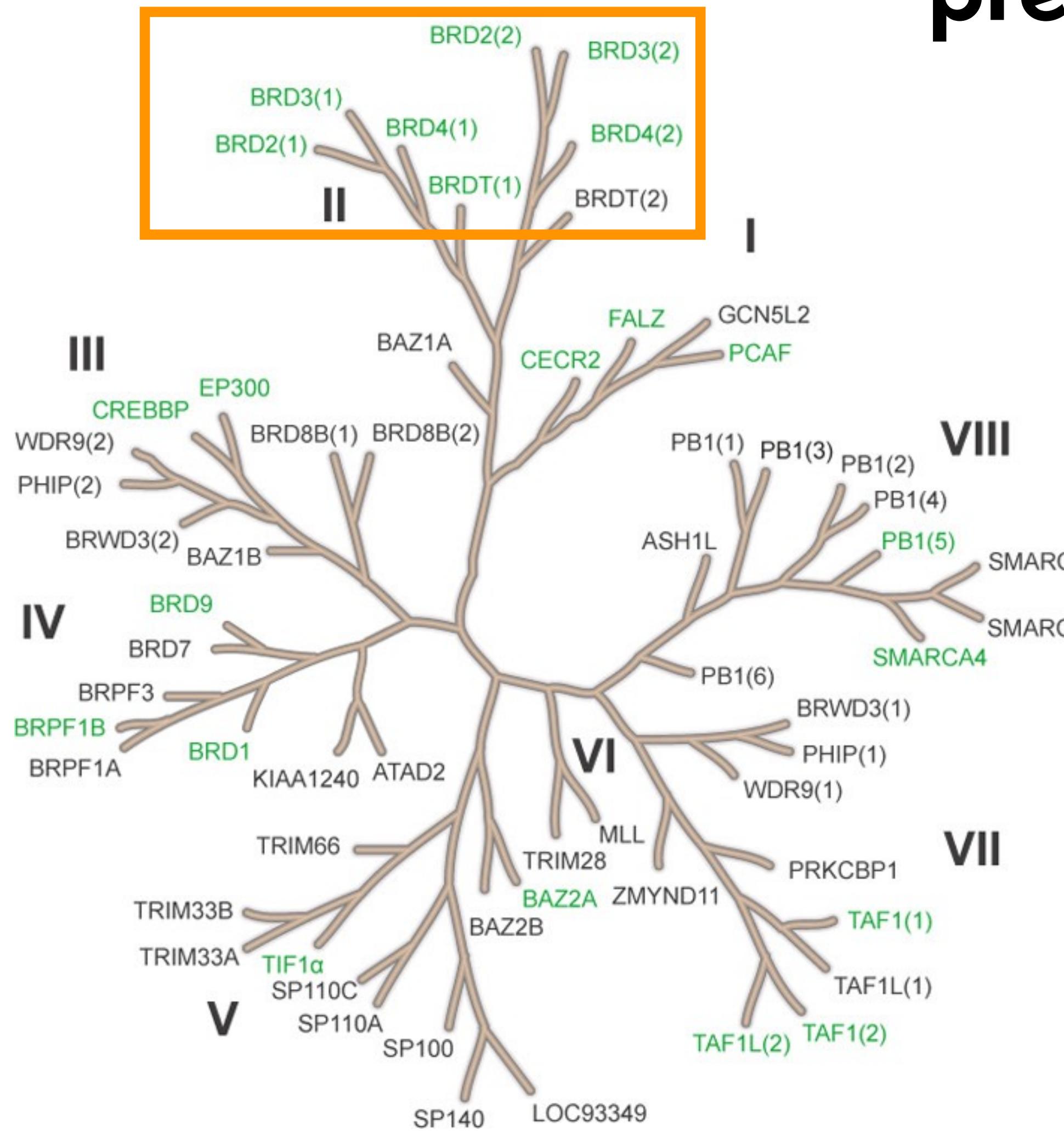
RVX-208



Bromosporine



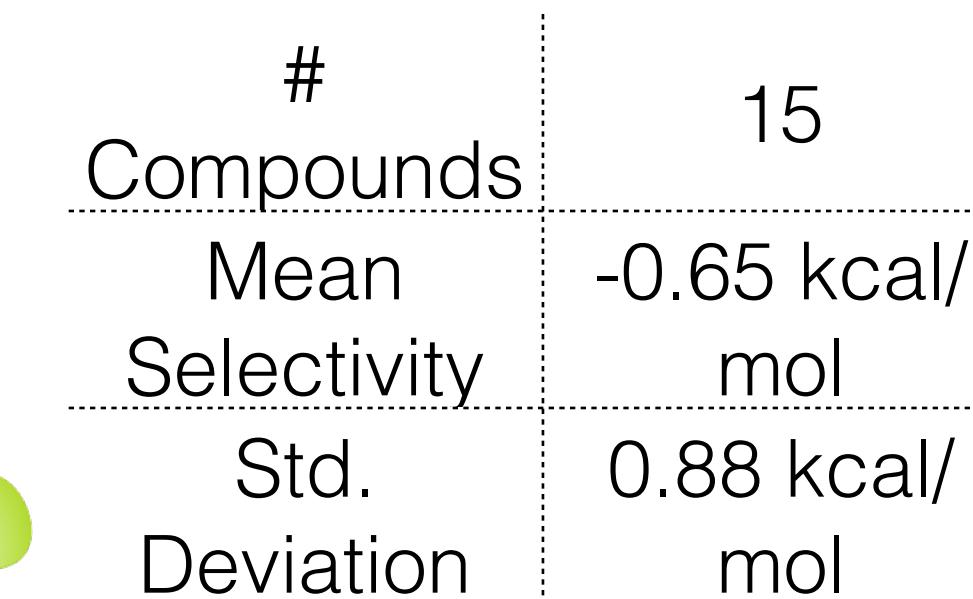
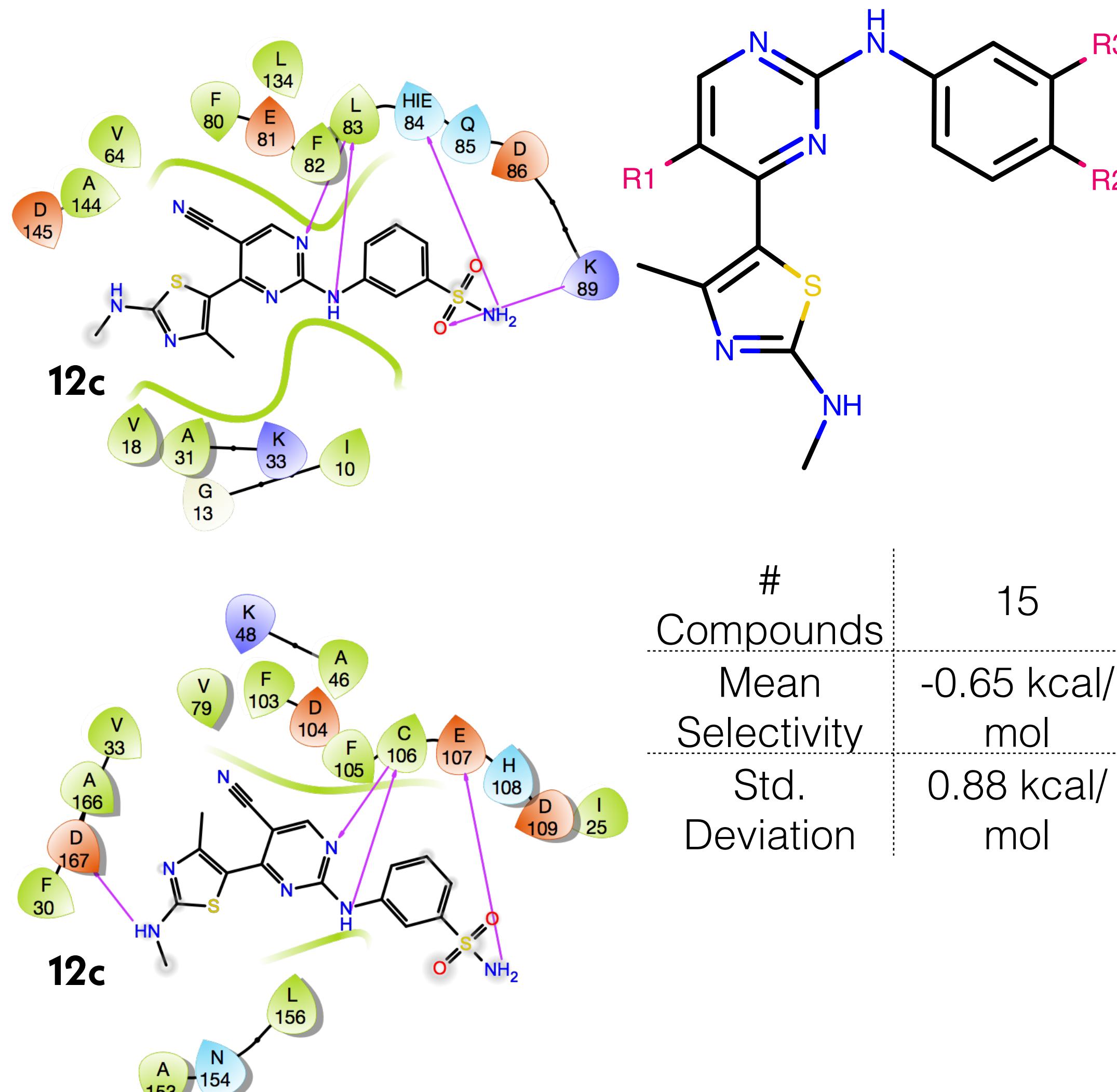
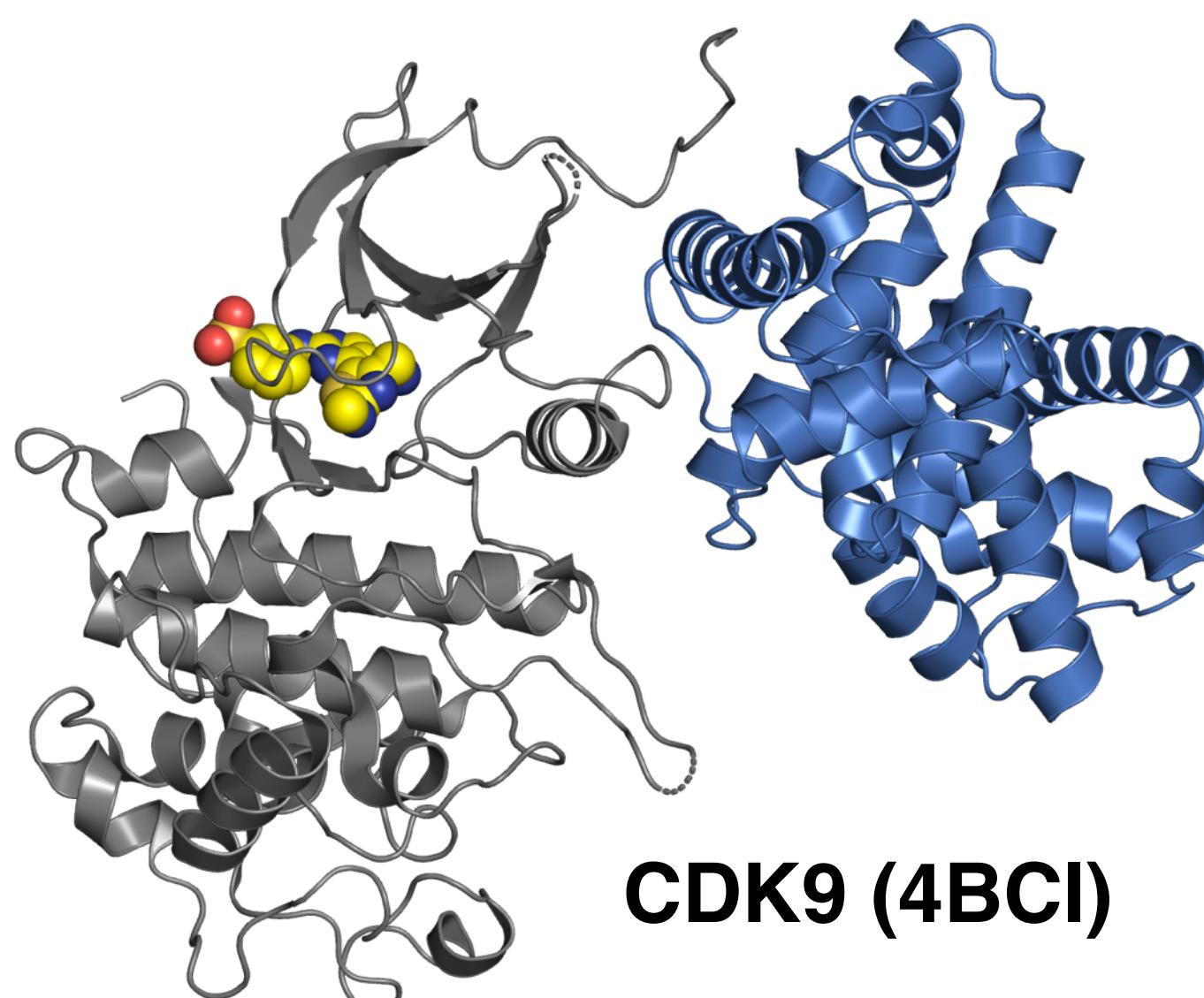
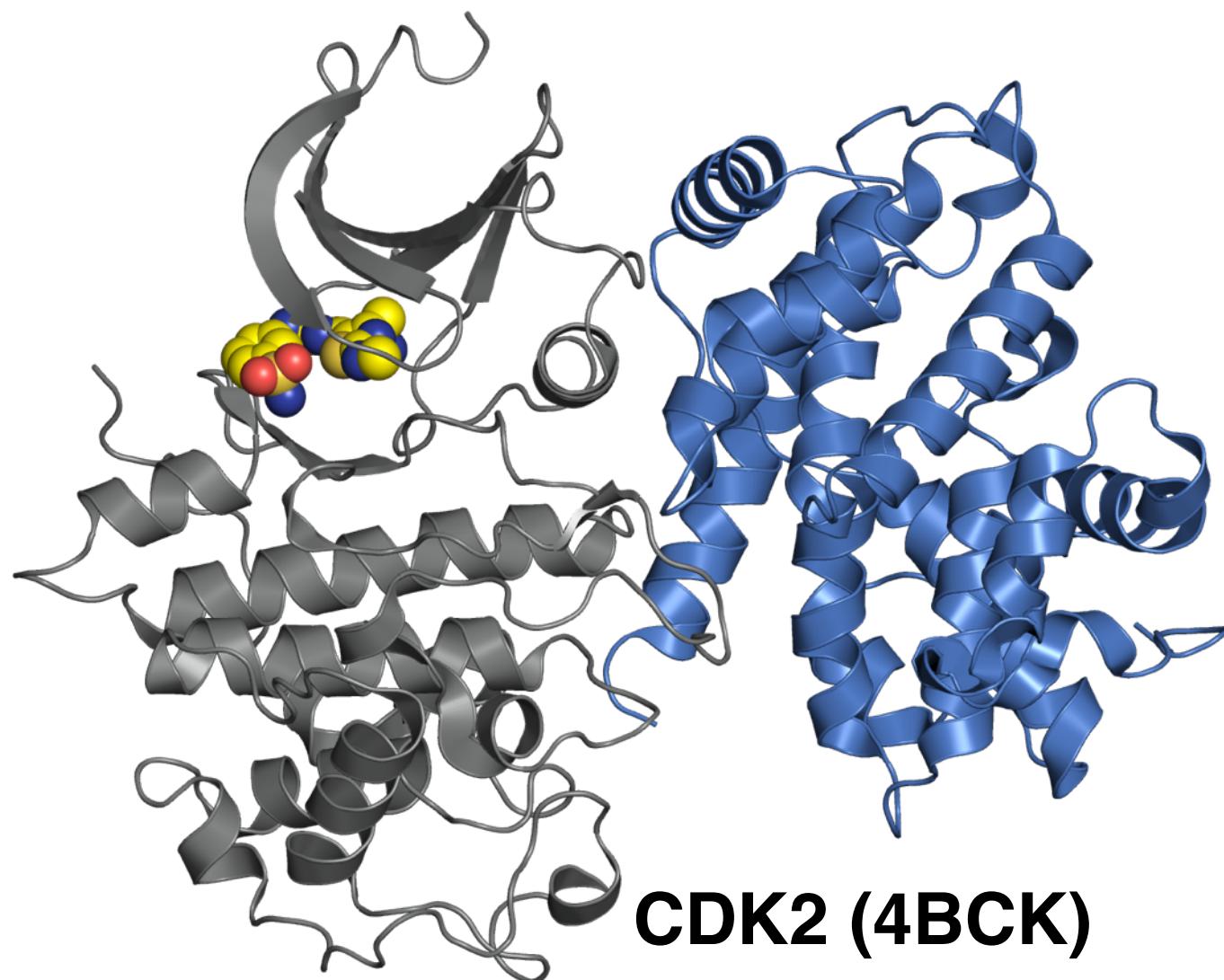
Predicting selectivity is difficult, even with good potency predictions



Are the errors for each target correlated with each other?

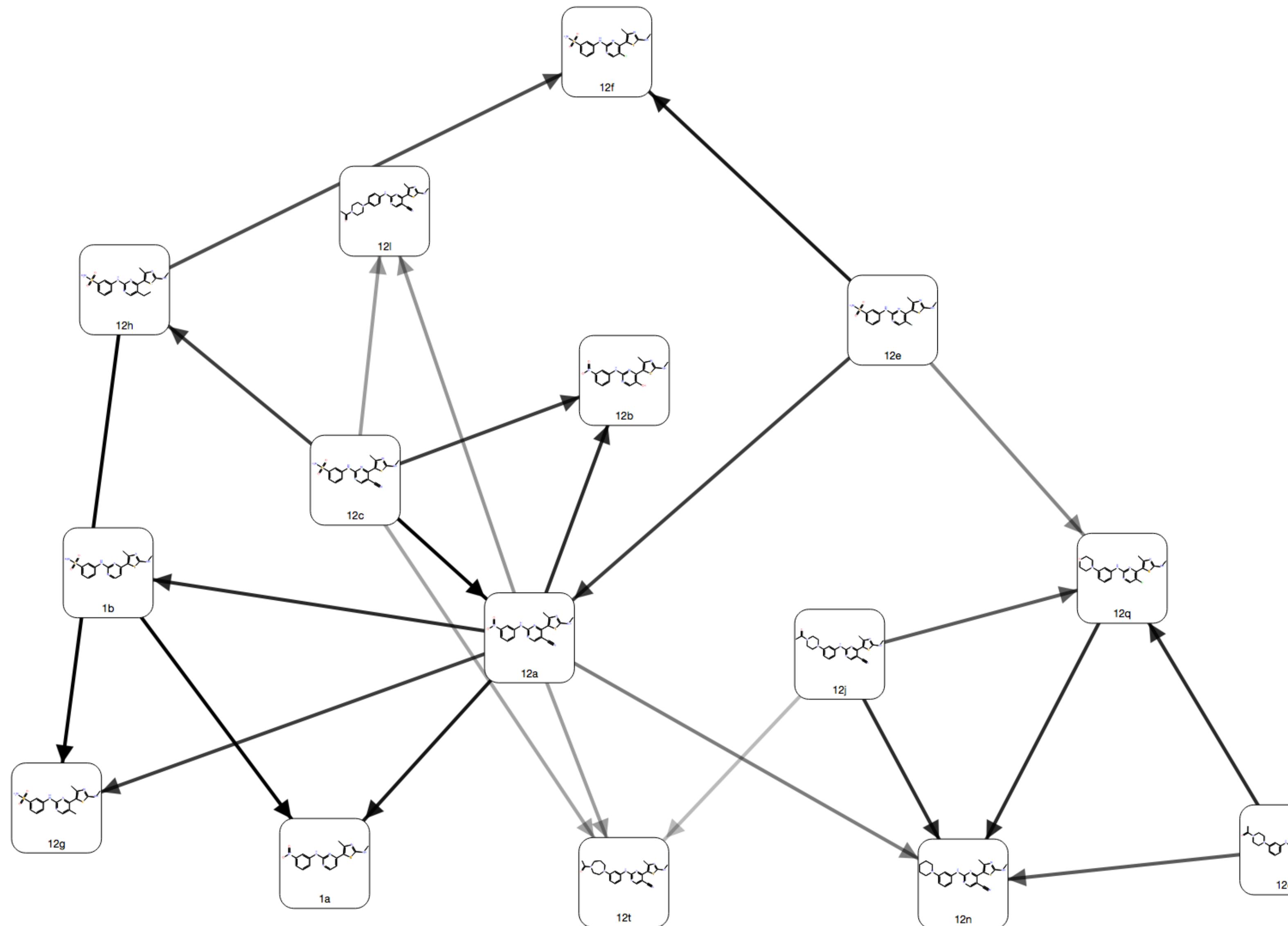
- If the errors are correlated with each other, there will be some cancellation of error and a smaller dynamic range of selectivities required to predict selective molecules with confidence
- If the errors are uncorrelated with each other, a much larger dynamic range of selectivity is required to predict selective molecules with confidence
- Answering this question gives us an idea of the utility of the methodology once working with compounds with unknown experimental measurements

A CDK2/CDK9 dataset as a test case for relative free energy calculations between closely related targets



Ligand	R1	R2	R3	ΔG CDK2 (kcal/mol)	ΔG CDK9 (kcal/mol)	ΔΔG (kcal/mol)
12a	CN	H		-12.27	-11.21	-1.64
12b	OH	H		-7.23	-8.22	-1.57
12c	CN	H		-11.45	-11.21	-1.57
12e	F	H		-11.62	-11.45	-1.57
12f	Cl	H		-10.91	-10.85	-2.36
12g	Methyl	H		-10.18	-11.32	-1.97
12h	Ethyl	H		-8.28	-9.56	-2.37
12j	CN	H		-10.04	-11.12	-1.56
12l	CN		H	-10.34	-10.44	-1.34
12n	CN	H		-10.06	-10.97	-2.47
12o	F	H		-10.06	-11.12	-0.75
12q	F	H		-10.91	-11.62	-2.31
12t	CN	H		-9.38	-11.12	-1.91
1a	H	H		-11.62	-11.86	-2.77
1b	H	H		-11.45	-11.86	-1.77

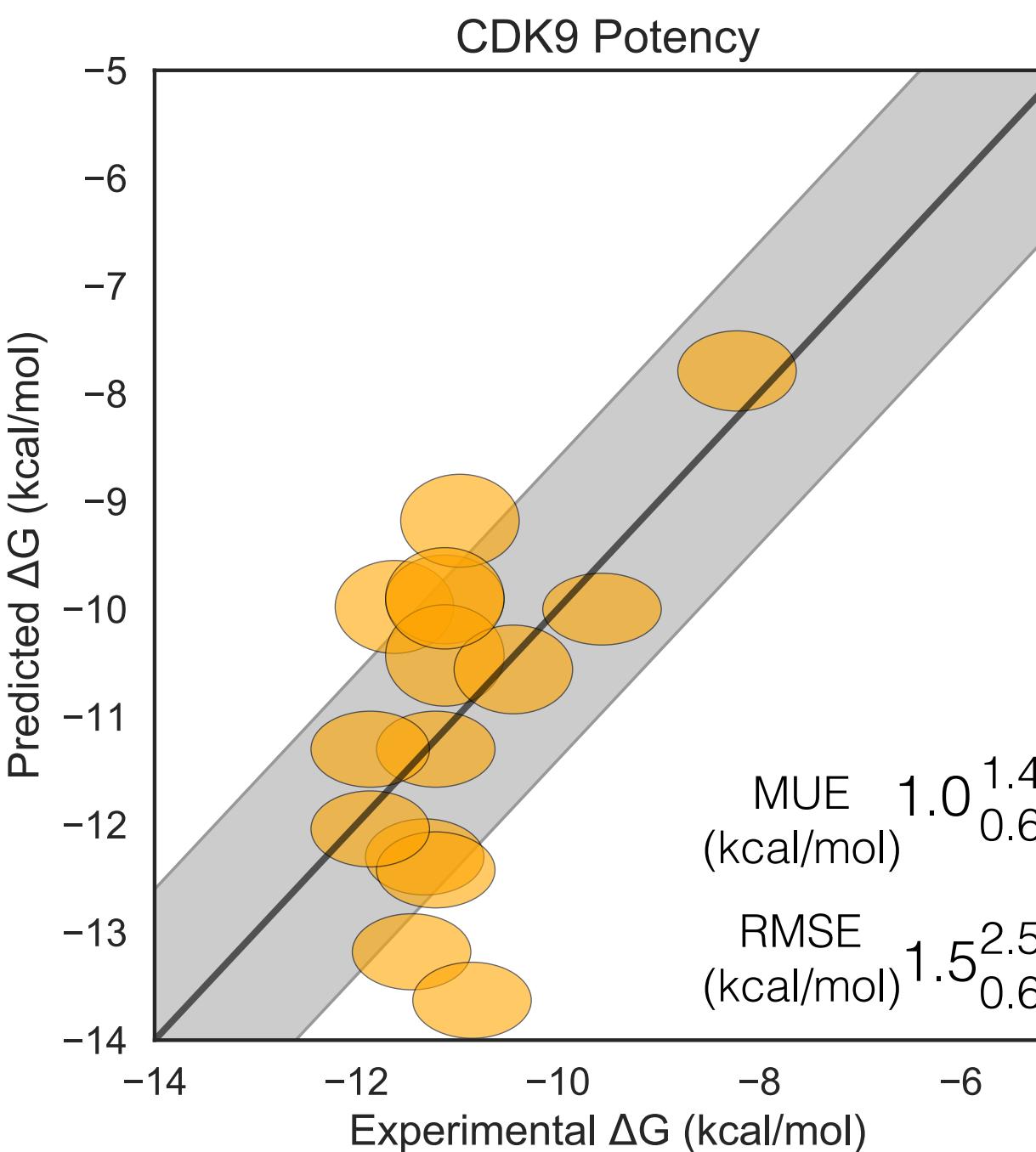
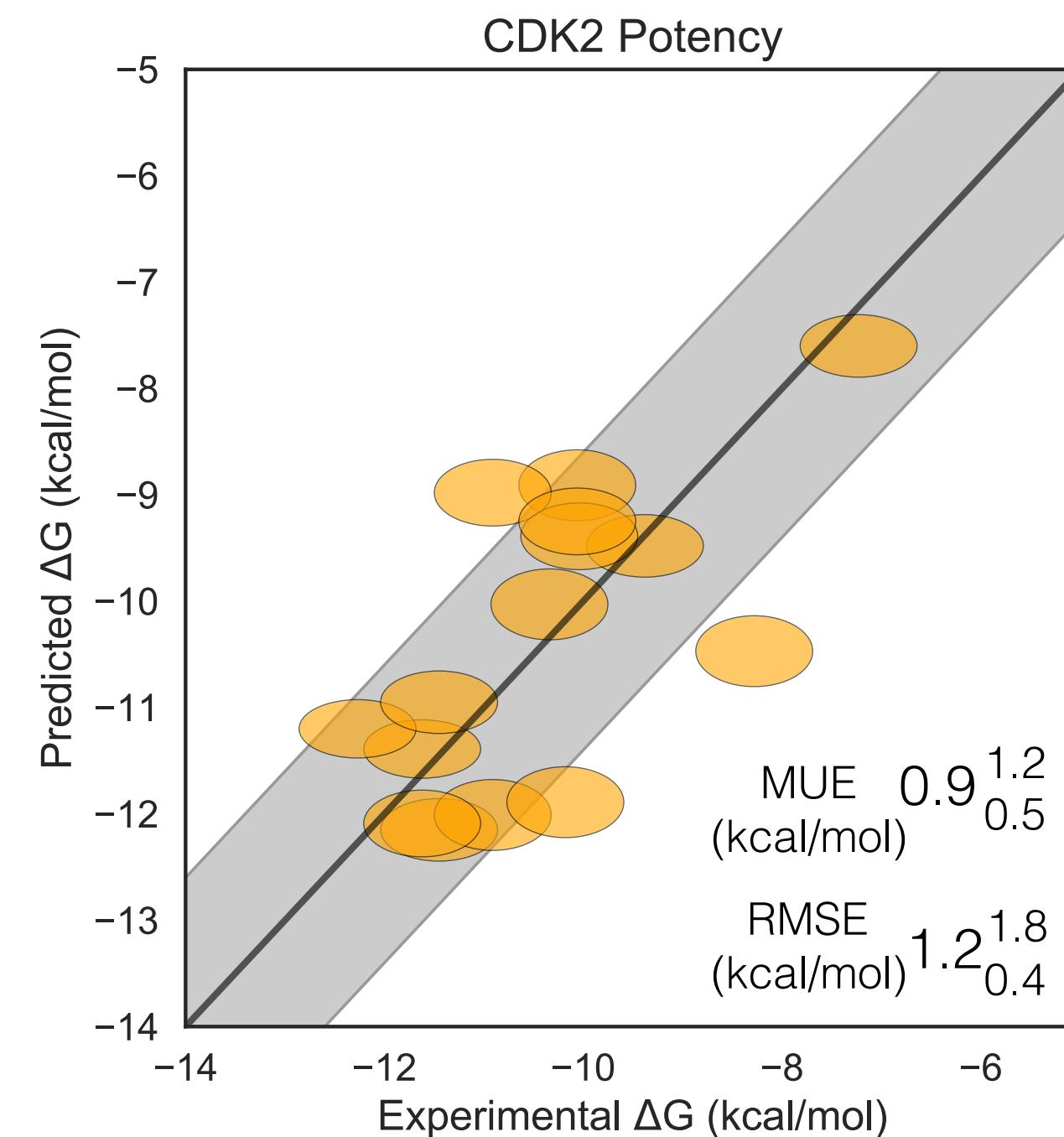
Relative free energy calculations calculate the $\Delta\Delta G$ between two different ligands



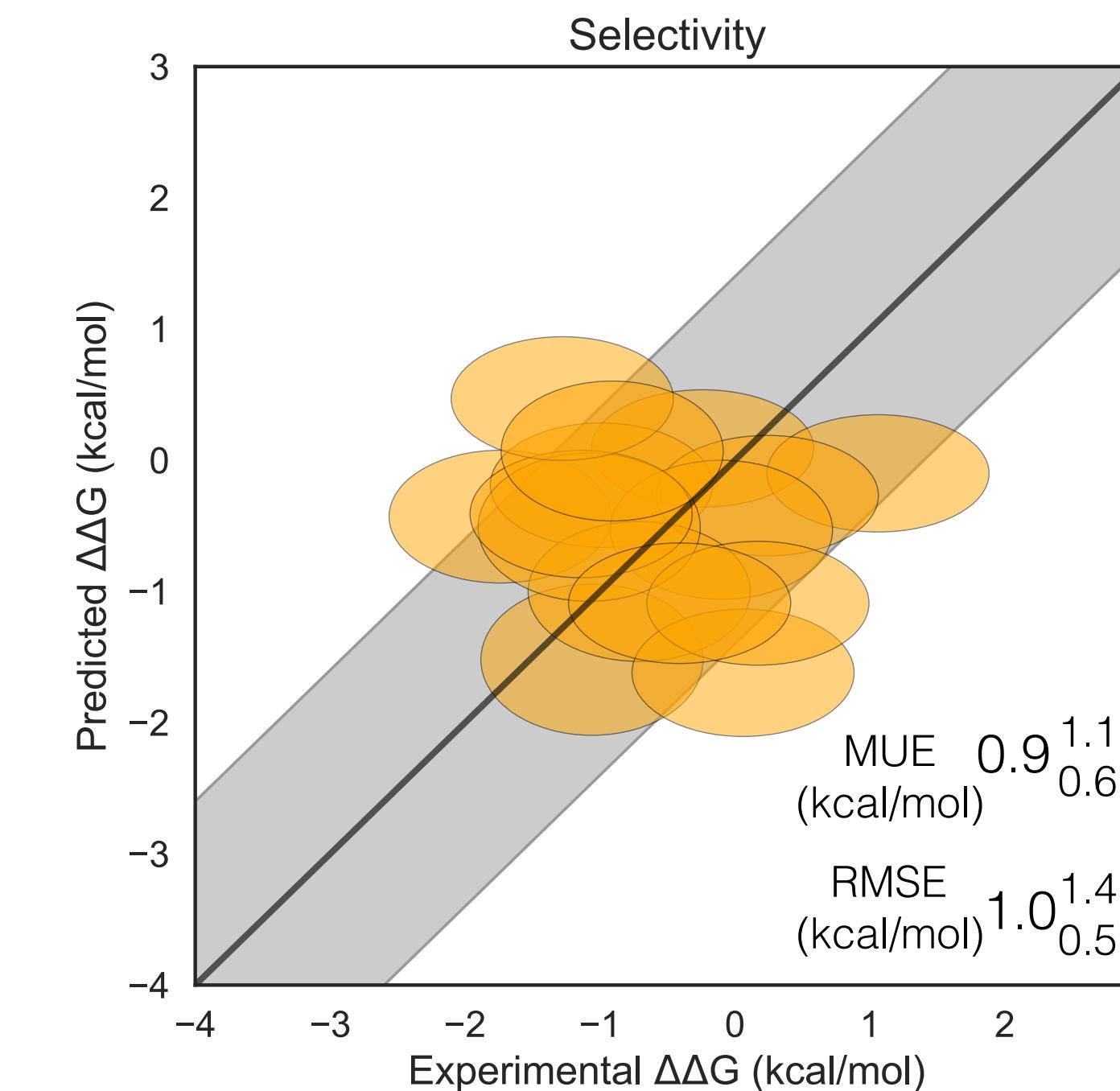
- Each edge represents a relative free energy calculation
- the lighter the edge, the less similar the connected compounds
- Knowing the experimental value for one (or more) compounds allows you to calculate ΔG for the rest

Relative free energy calculations achieve low error when predicting selectivity, but high uncertainty limits utility of the method

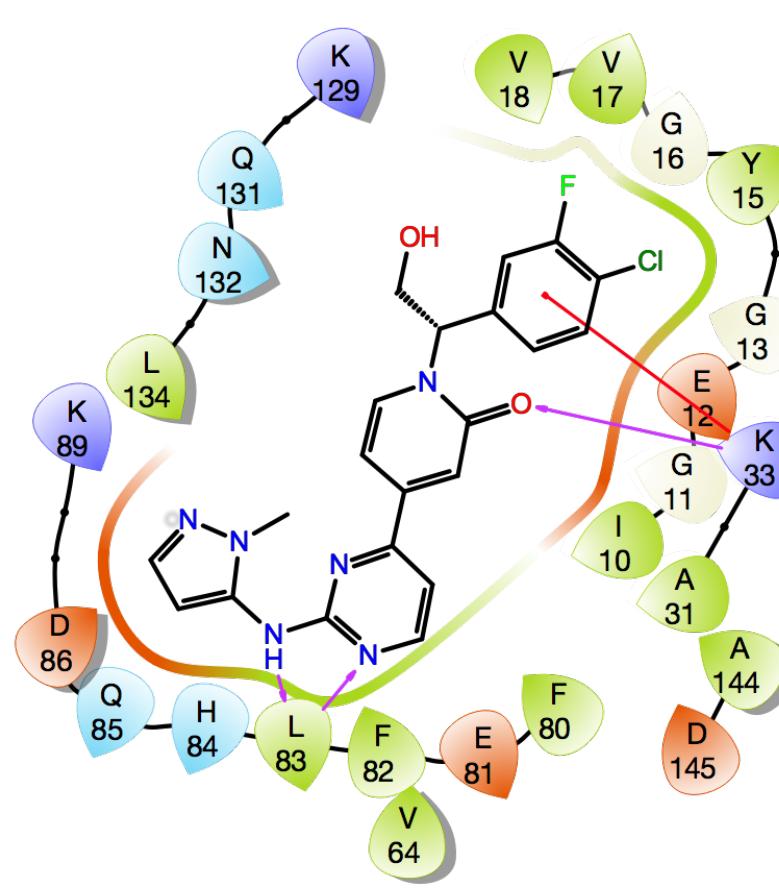
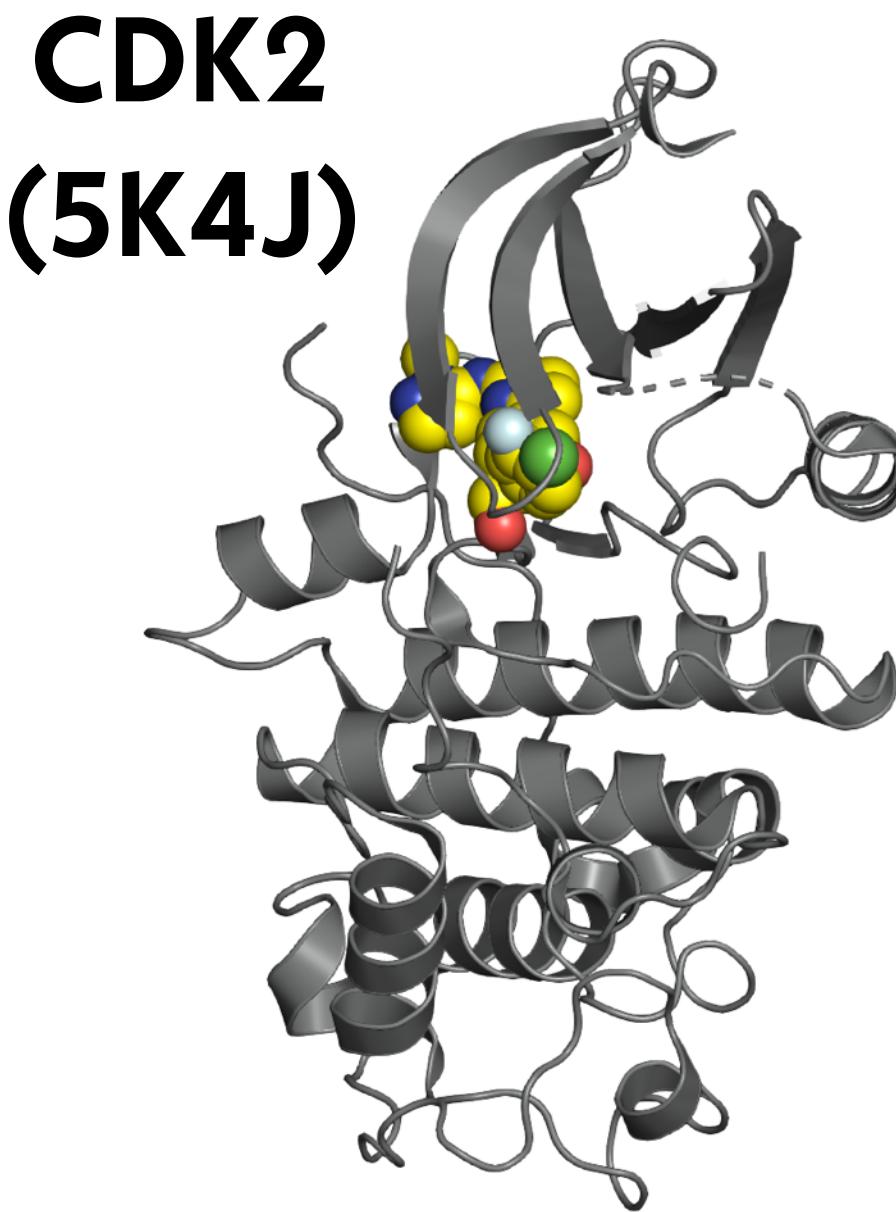
ΔG



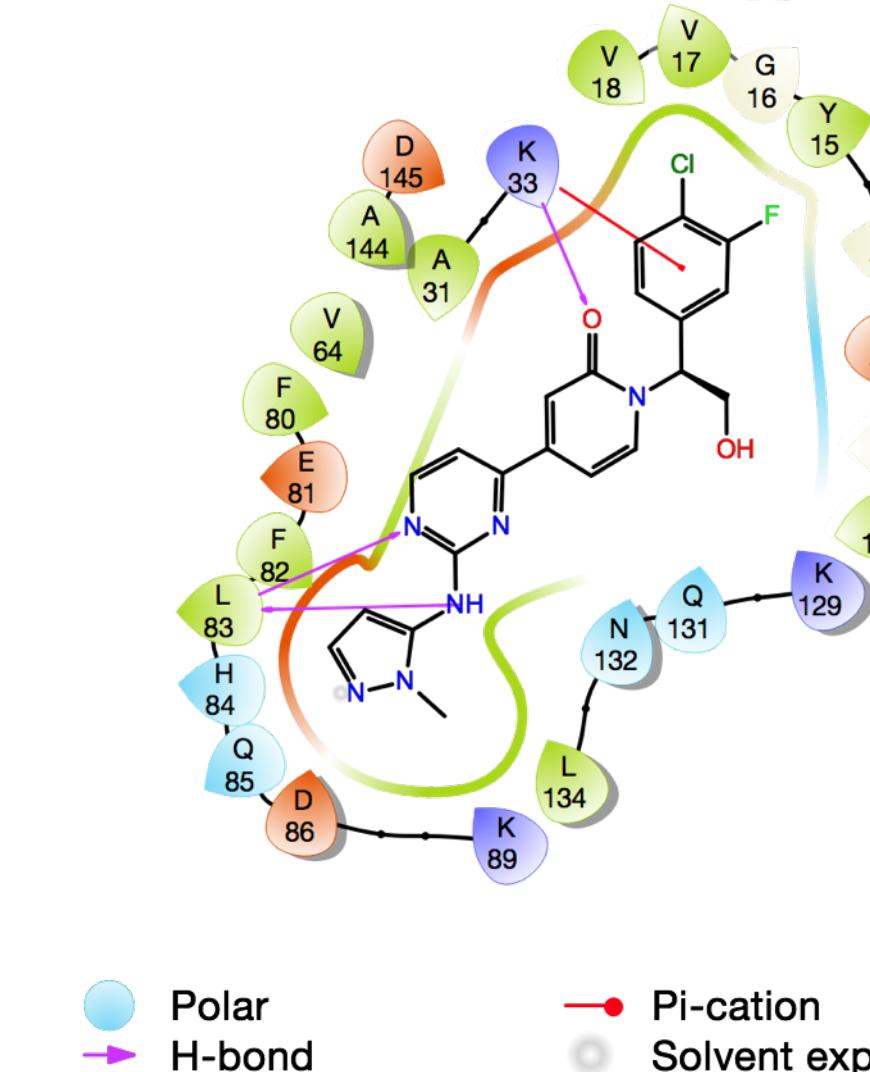
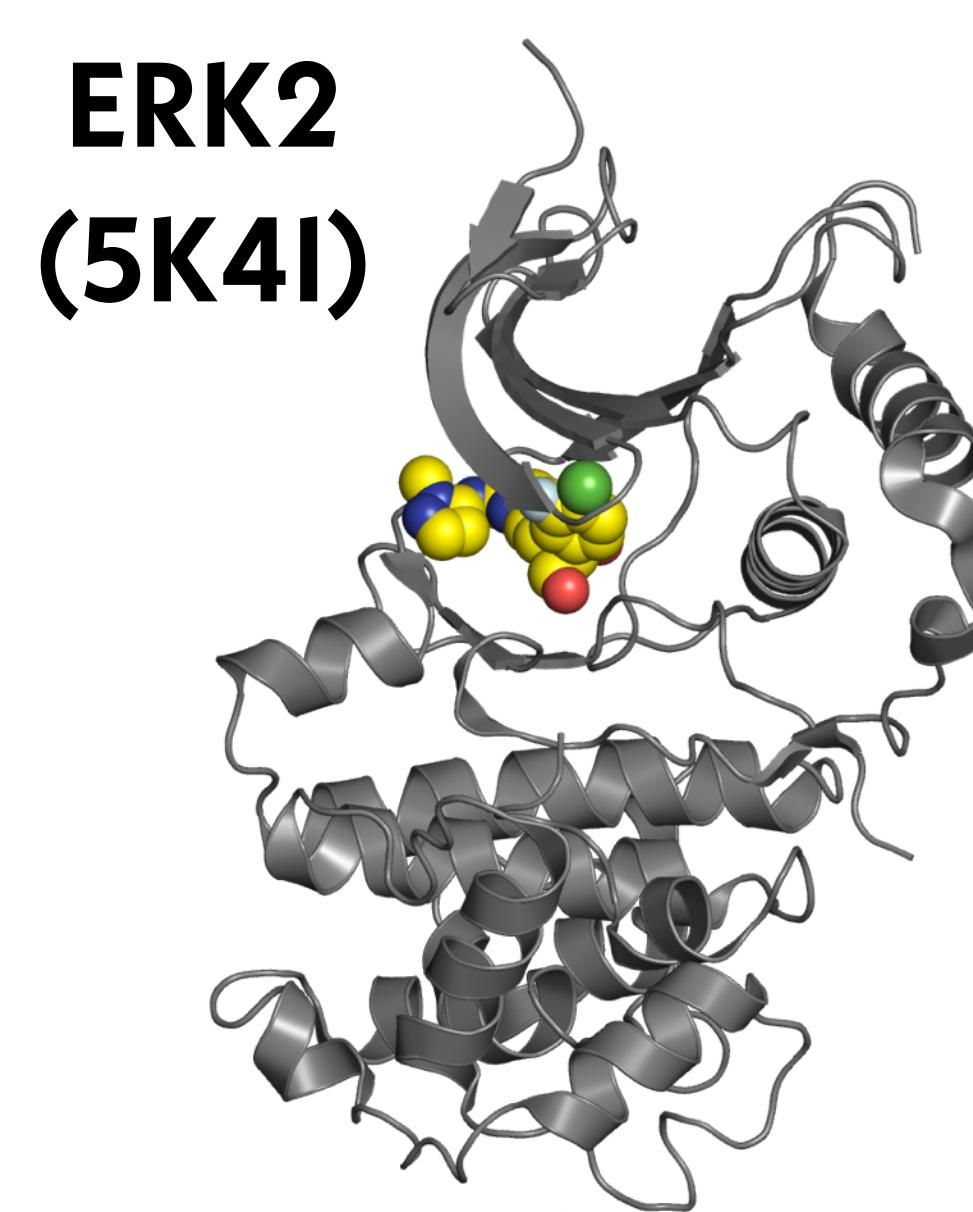
$\Delta\Delta G$ (CDK9 - CDK2)



A CDK2/ERK2 dataset as a test case for relative free energy calculations when moderate selectivity has been achieved



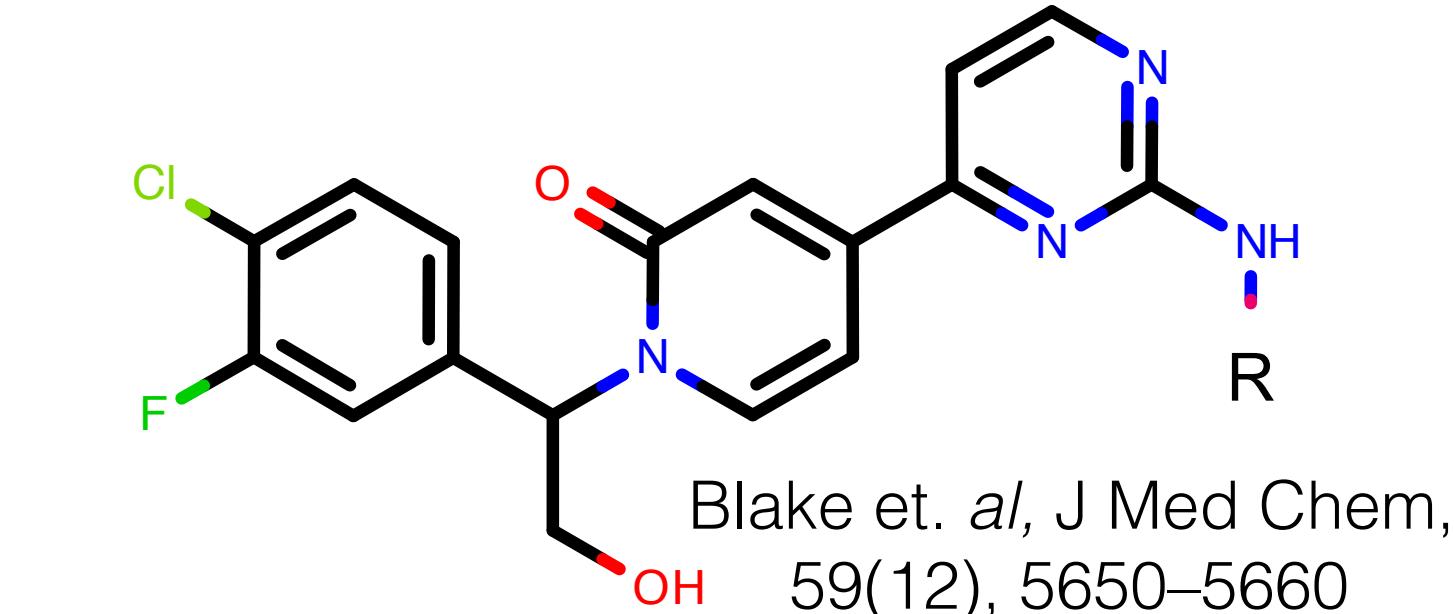
● Charged (negative)
● Charged (positive)
● Polar H-bond
● Glycine
● Hydrophobic



● Polar H-bond
● Pi-cation Solvent exposure
● Glycine
● Hydrophobic

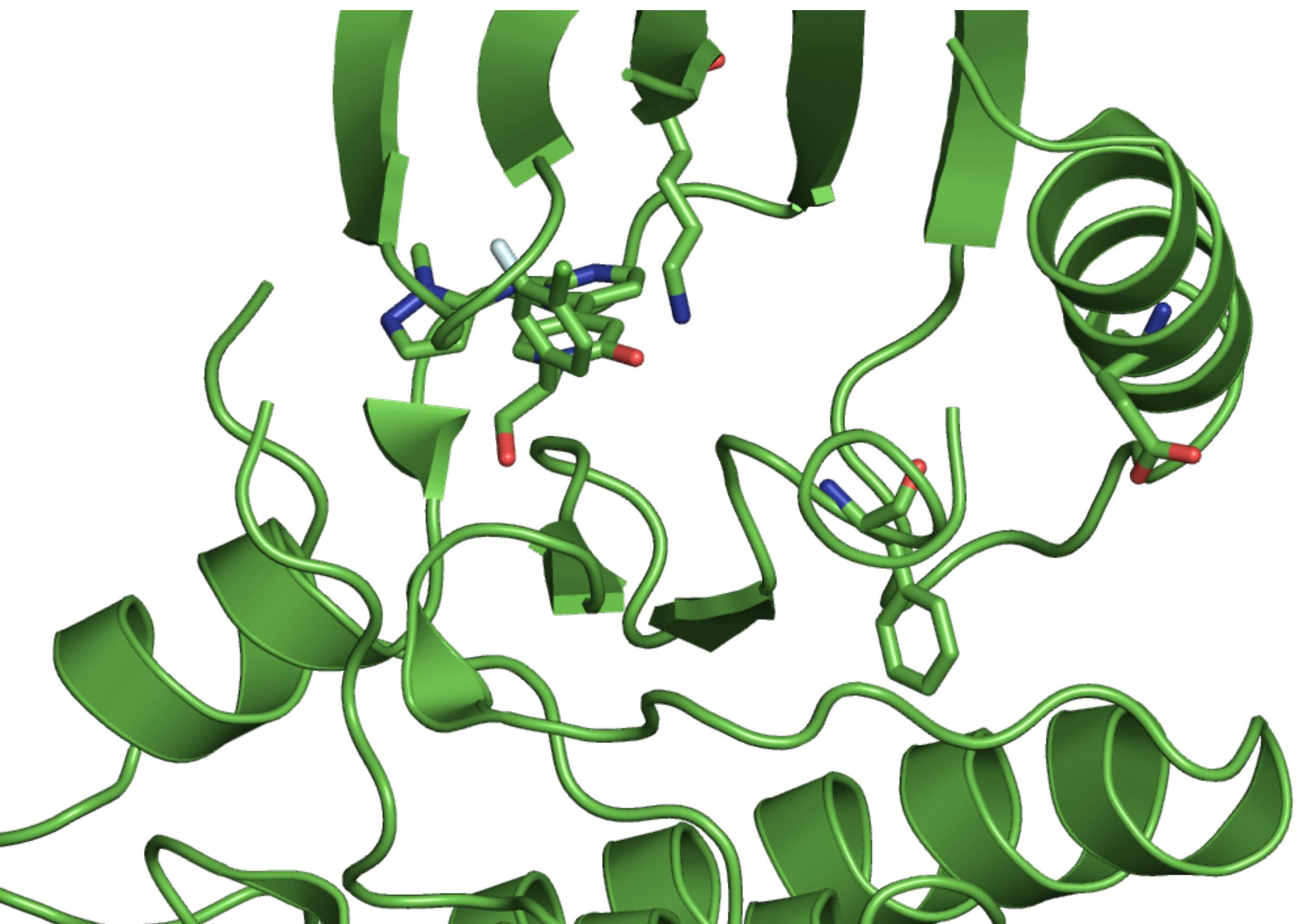
# Compounds	19
Mean Selectivity	-1.74 kcal/mol
Std. Deviation	0.56 kcal/mol

Ligand		ΔG CDK2 (kcal/mol)	ΔG ERK2 (kcal/mol)	$\Delta\Delta G$ (kcal/mol)
6		-10.57	-12.21	-1.64
9		-10.74	-12.31	-1.57
10		-9.87	-11.44	-1.57
11		-8.71	-10.28	-1.57
12		-8.86	-11.22	-2.36
13		-8.44	-10.41	-1.97
14		-8.69	-11.06	-2.37
15		-10.36	-11.92	-1.56
16		-9.57	-10.91	-1.34
17		-9.86	-12.33	-2.47

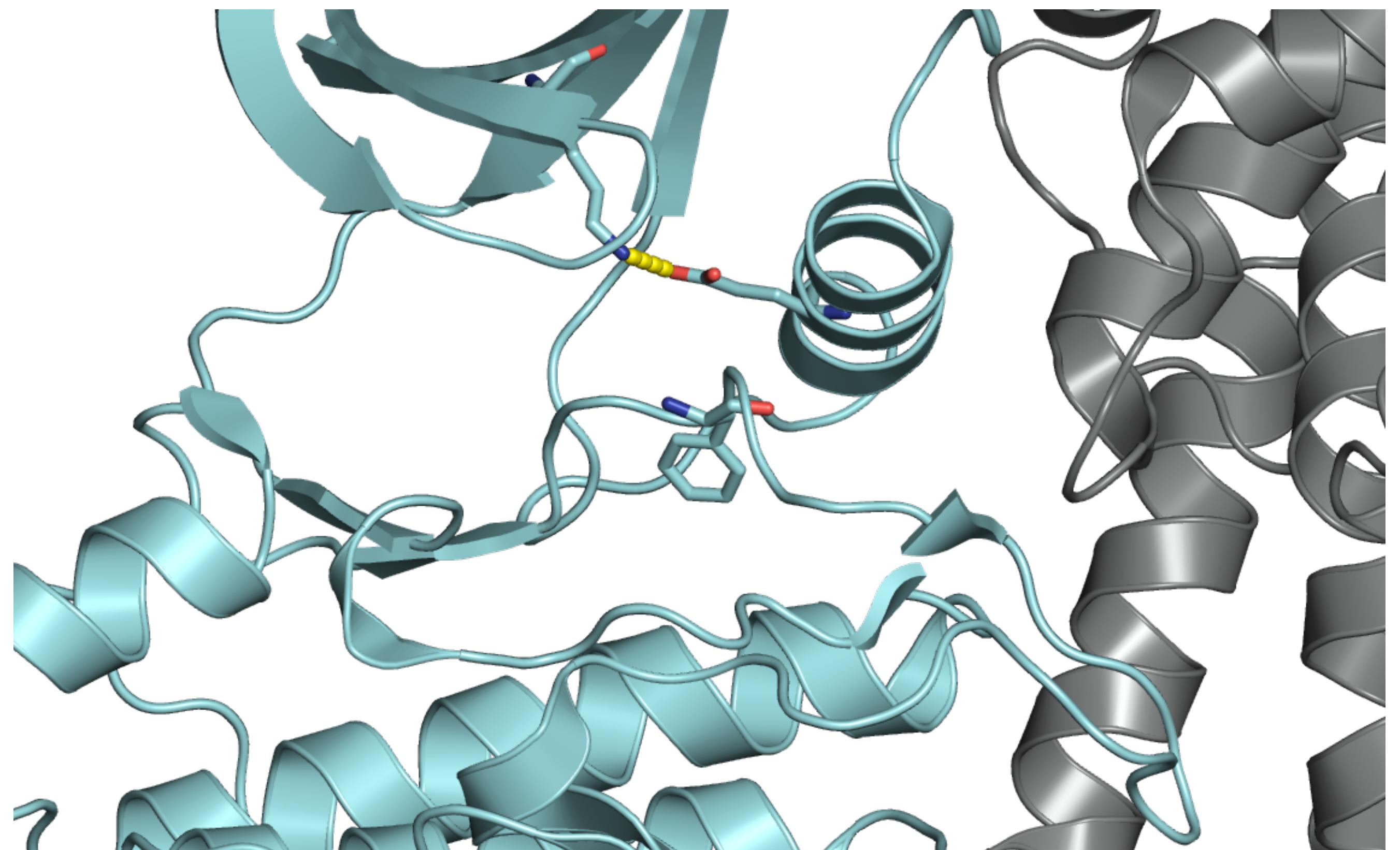


Ligand		ΔG CDK2 (kcal/mol)	ΔG ERK2 (kcal/mol)	$\Delta\Delta G$ (kcal/mol)
18		-8.28	-9.03	-0.75
21		-7.82	-10.13	-2.31
22		-9.69	-11.6	-1.91
23		-7.86	-10.63	-2.77
24		-8.98	-10.75	-1.77
25		-7.87	-9.61	-1.74
26		-8.92	-10.59	-1.67
27		-8.01	-8.6	-0.59
28		-8.76	-9.93	-1.17

The CDK2 structure adopts a DFG-in, inactive conformation in the absence of its cyclin



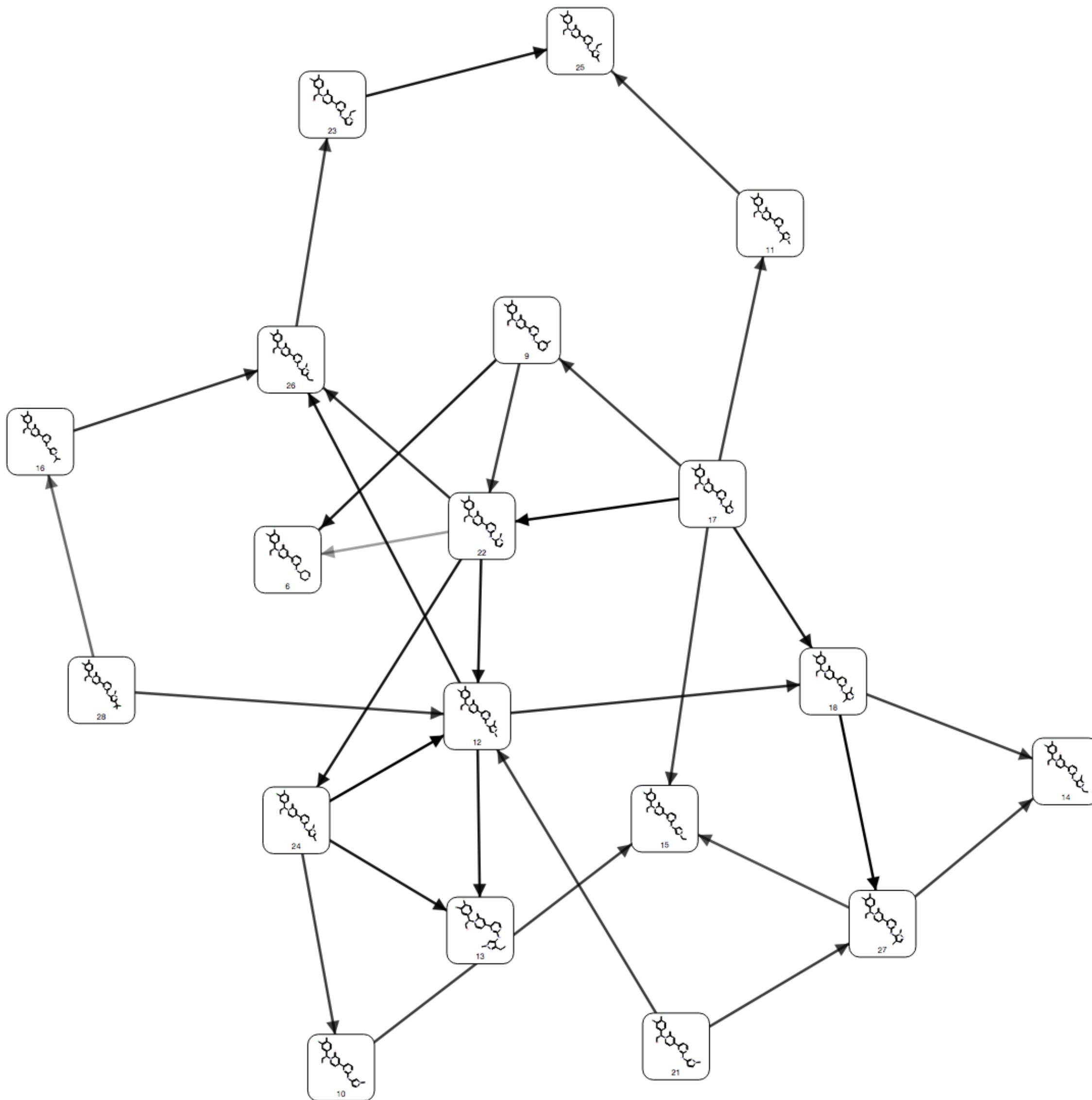
PDB: 5K4J



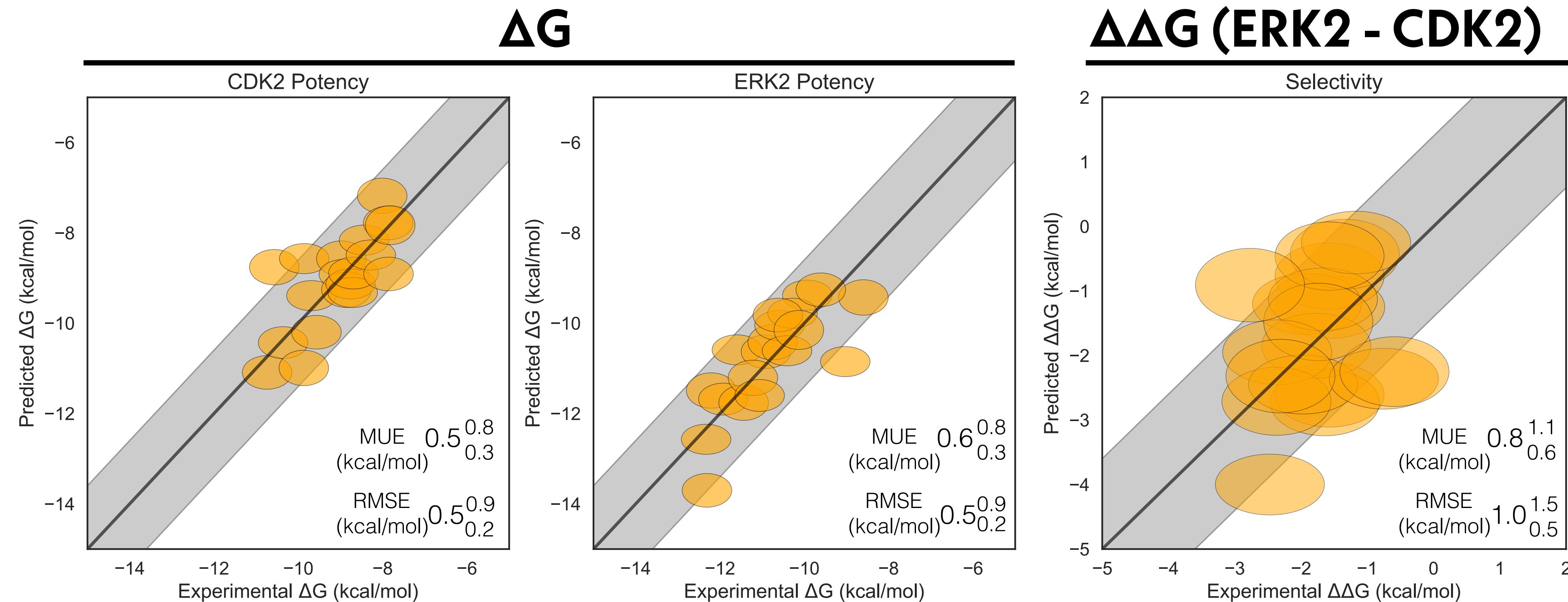
PDB: 4BCK

Relative free energy calculations can take advantage of cancellation of errors

- If a certain conformation contributes roughly equally to the binding affinity of all ligands in the series, neglecting it will not change the relative ordering of the compounds

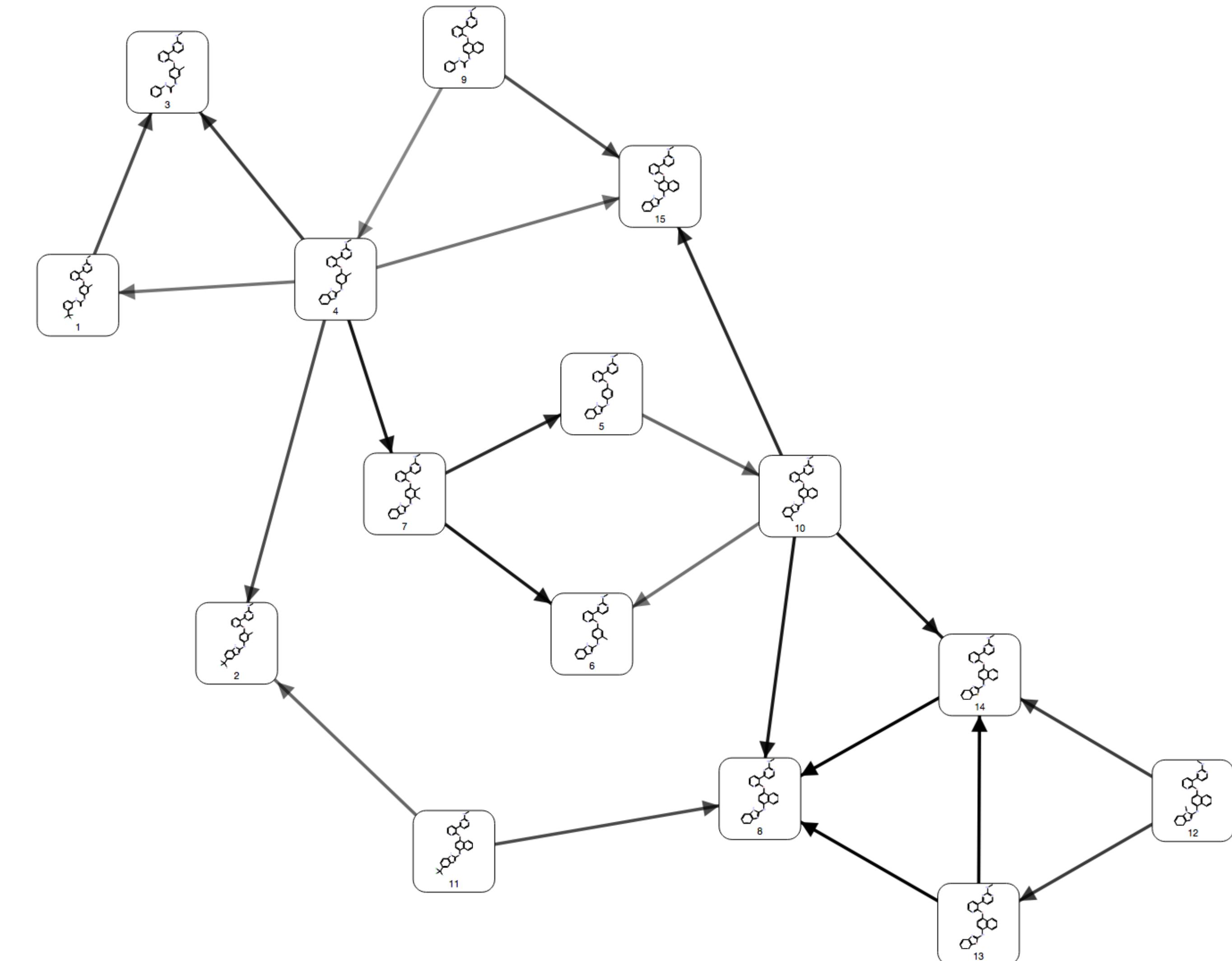


Relative free energy calculations achieve low error when predicting selectivity, but high uncertainty limits utility of the method



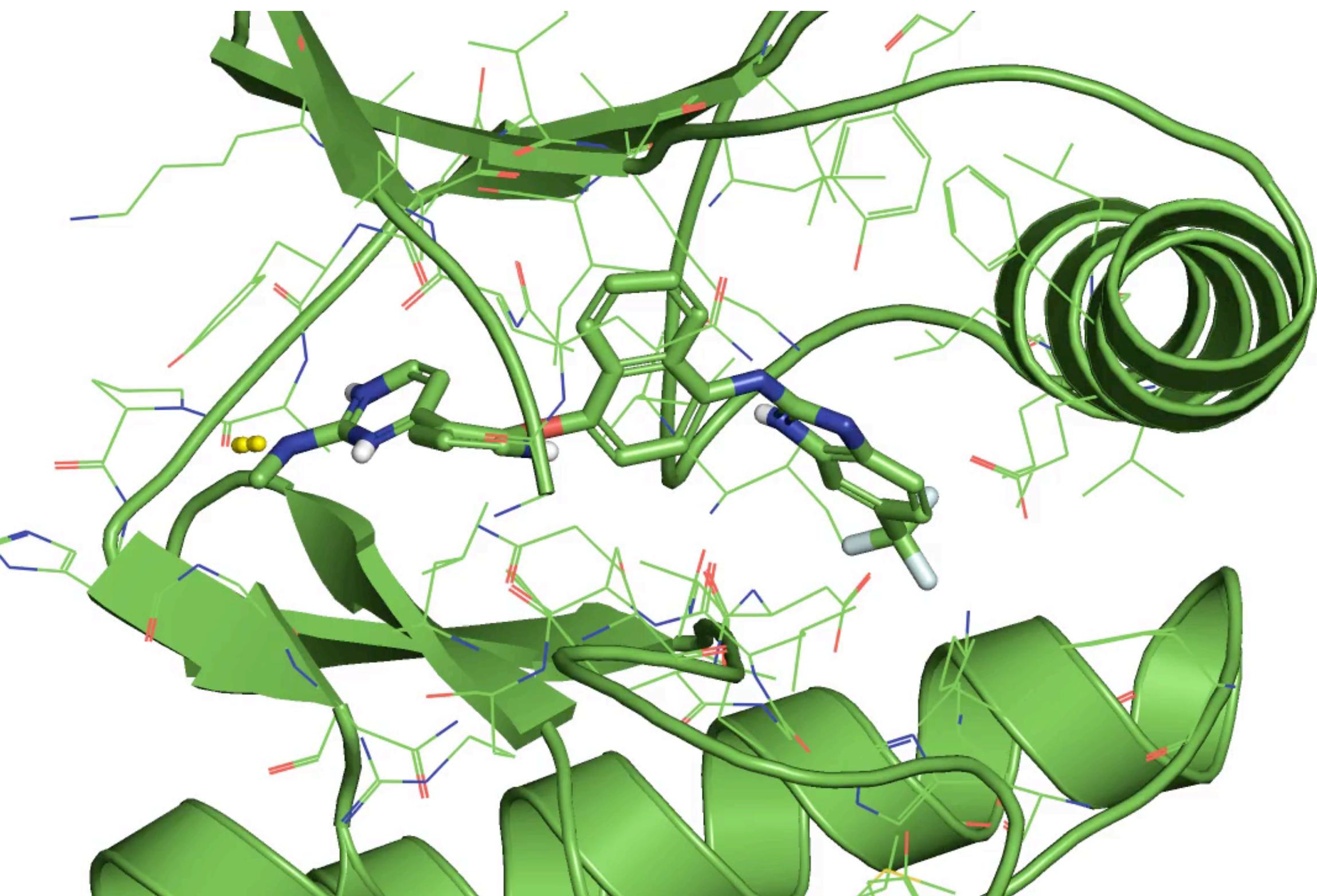
Tie2/KDR: a lead series with a higher dynamic range

- 15 closely related compounds with simple modifications
- Crystal structure for KDR/compound 11
- Need to generate a binding pose for Tie2 with compound 11



Generating a Tie2 structure

- A Tie2 structure bound to a related ligand suggests that Tie2 adopts a very similar conformation to KDR when bound to ligands
- Protocol:
 - align Tie2 structure to KDR using the binding site residues
 - transfer compound 11 from KDR to Tie2
 - Run 20 ns of MD to gauge stability of the transferred binding pose



A simple approach to estimating correlation shows mostly uncorrelated errors

$\{\Delta G_{\text{exp},i} - \Delta G_{\text{pred},i}\}$

vs

$\{\Delta G_{\text{exp},i}^2 - \Delta G_{\text{pred},i}^2\}$

Target	R [95% CI]
BRDs — RVX-OH	-0.2 [-0.6, 0.3]
BRDS — RVX-208	-0.2 [-0.5, 0.2]
BRDS — Bromosporine	0.0 [-0.2, 0.1]
BRDs — Mod Bromosporine	0 [-0.2, 0.1]
CDK2/CDK9	0.7 [0.4, 0.8]
ERK2/CDK2	0.1 [-0.4, 0.4]

Bayesian inference allows us to better estimate the errors and correlation

Priors

$$\Delta\Delta G_i^{\text{true}} \sim \mathcal{N}(\mu_{\text{sel}}, \sigma_{\text{sel}}^2) \quad i = 1, \dots, M$$

$$\mu_{\text{sel}} \sim U(-6, +6)$$

$$\sigma_{\text{sel}} \propto \sigma^{-1}$$

Data Likelihood

$$\Delta\Delta G_{\text{exp}} \sim \mathcal{N}(\Delta\Delta G_i, \sigma_{\text{exp}}^2)$$

$$\Delta\Delta G_{\text{calc},i} \sim \mathcal{N}(\Delta\Delta G_i, \sigma_{\text{calc},i}^2)$$

Unknown parameters

$$\Delta\Delta G_{\text{calc}}^{\text{true}} \sim \mathcal{N}(\Delta\Delta G_i^{\text{true}}, \sigma_{\text{RMSE}}^2)$$

```
# Build the Bayesian model for alchemical free energy calculations
fep_model = pm.Model()

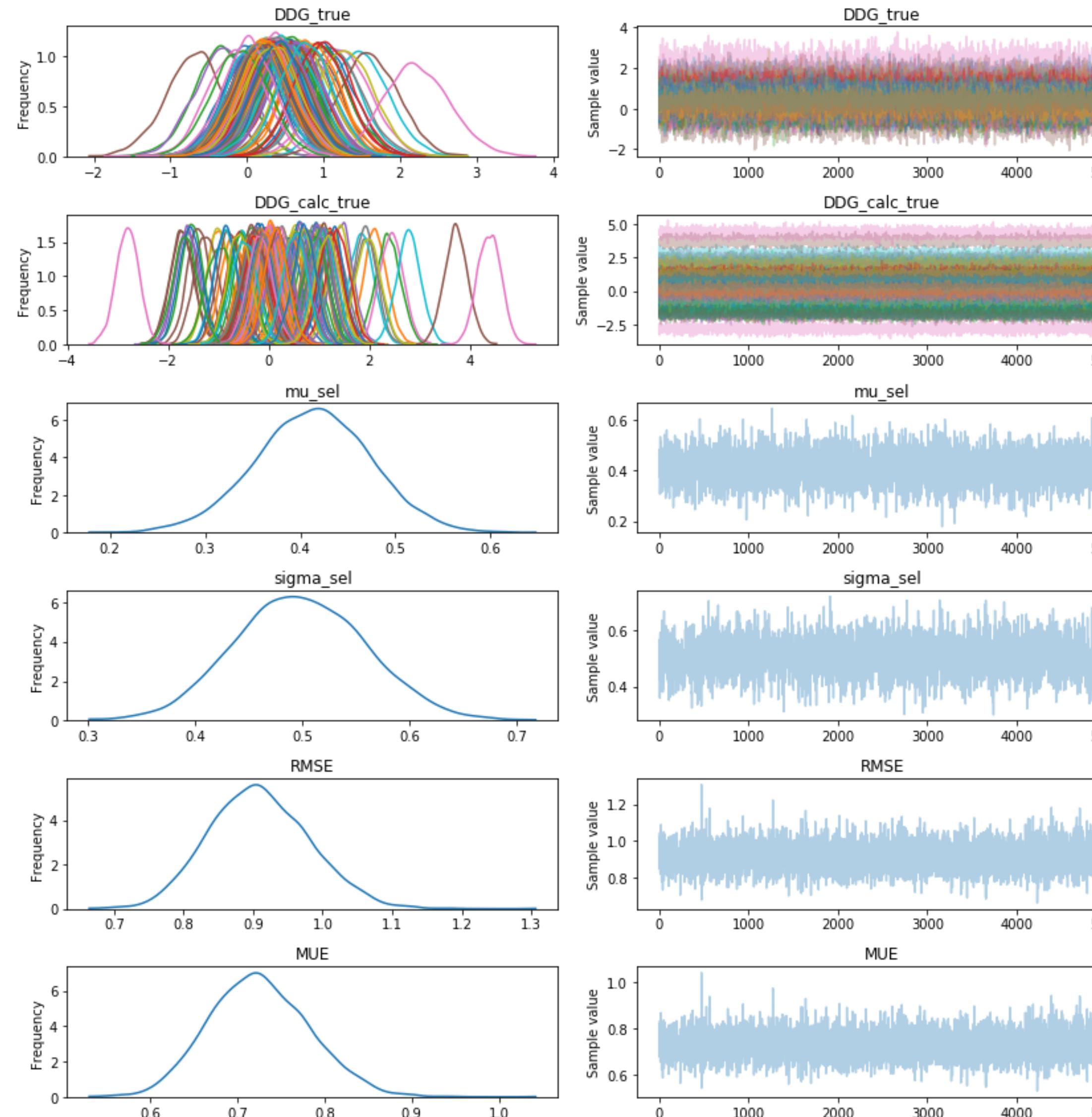
with fep_model:
    # Priors on nuisance parameters (that will be marginalized out)
    mu_sel = pm.Uniform('mu_sel', -6, +6) # kcal/mol
    sigma_sel = pm.HalfFlat('sigma_sel', testval=1) # kcal/mol
    DDG_true = pm.Normal('DDG_true', mu=mu_mut, sd=sigma_mut, shape=nmutants)

    # Priors on unknown values of interest
    RMSE_true = pm.HalfFlat('RMSE', testval=1) # kcal/mol, uninformative prior for nonnegative values of the RMSE
    MUE_true = pm.Deterministic('MUE', RMSE_true*np.sqrt(2.0/np.pi)) # store MUE estimate alongside RMSE
    # Unknown true computed values
    DDG_calc_true = pm.Normal('DDG_calc_true', mu=(DDG_true), sd=RMSE_true, shape=nmutants)

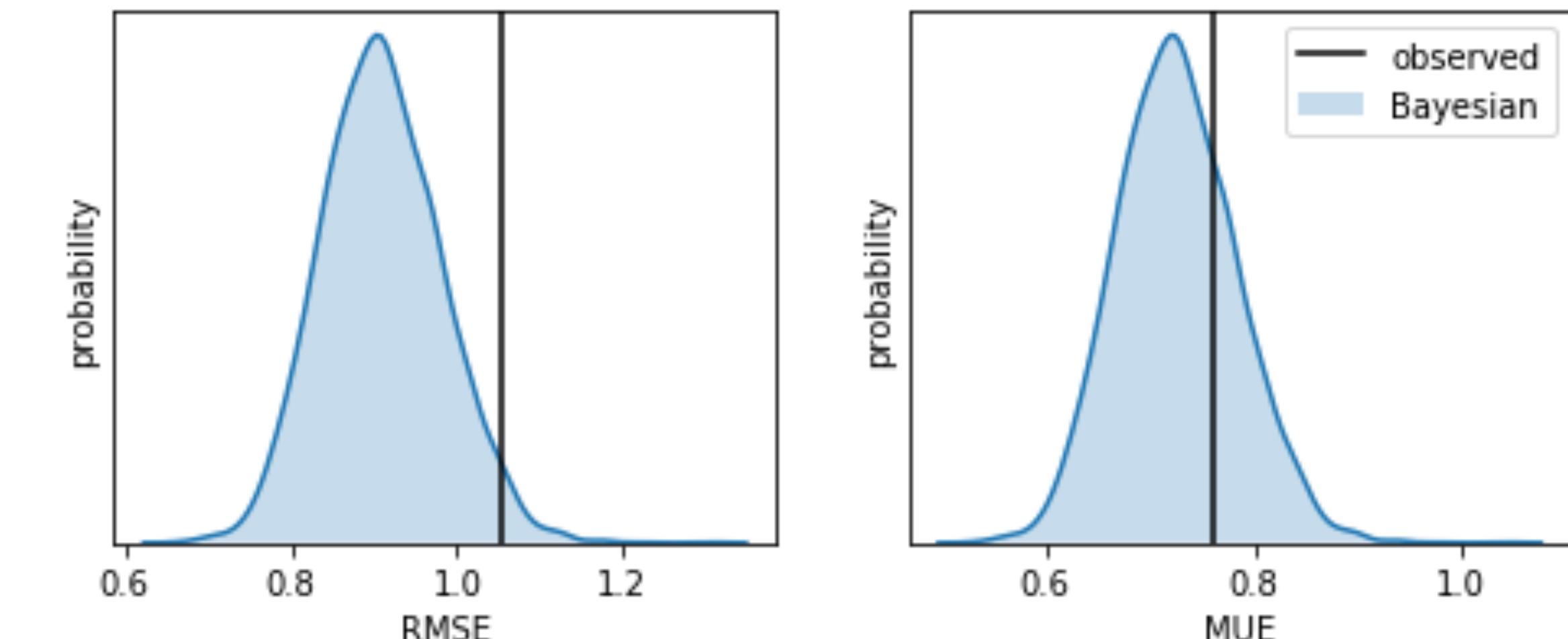
    # Data likelihood for observed data
    DDG_exp = pm.Normal('DDG_exp', mu=DDG_true, sd=sigma_exp, shape=nmutants, observed=DDG_exp_obs)
    DDG_calc = list()
    for replicate in range(nreplicates):
        DDG_calc.append(pm.Normal('DDG_calc_%d' % replicate, mu=DDG_calc_true,
                                  sd=dDDG_calc_obs[:,replicate], shape=nmutants, observed=DDG_calc_obs[:,replicate]))

with fep_model:
    trace = pm.sample(draws=5000)
```

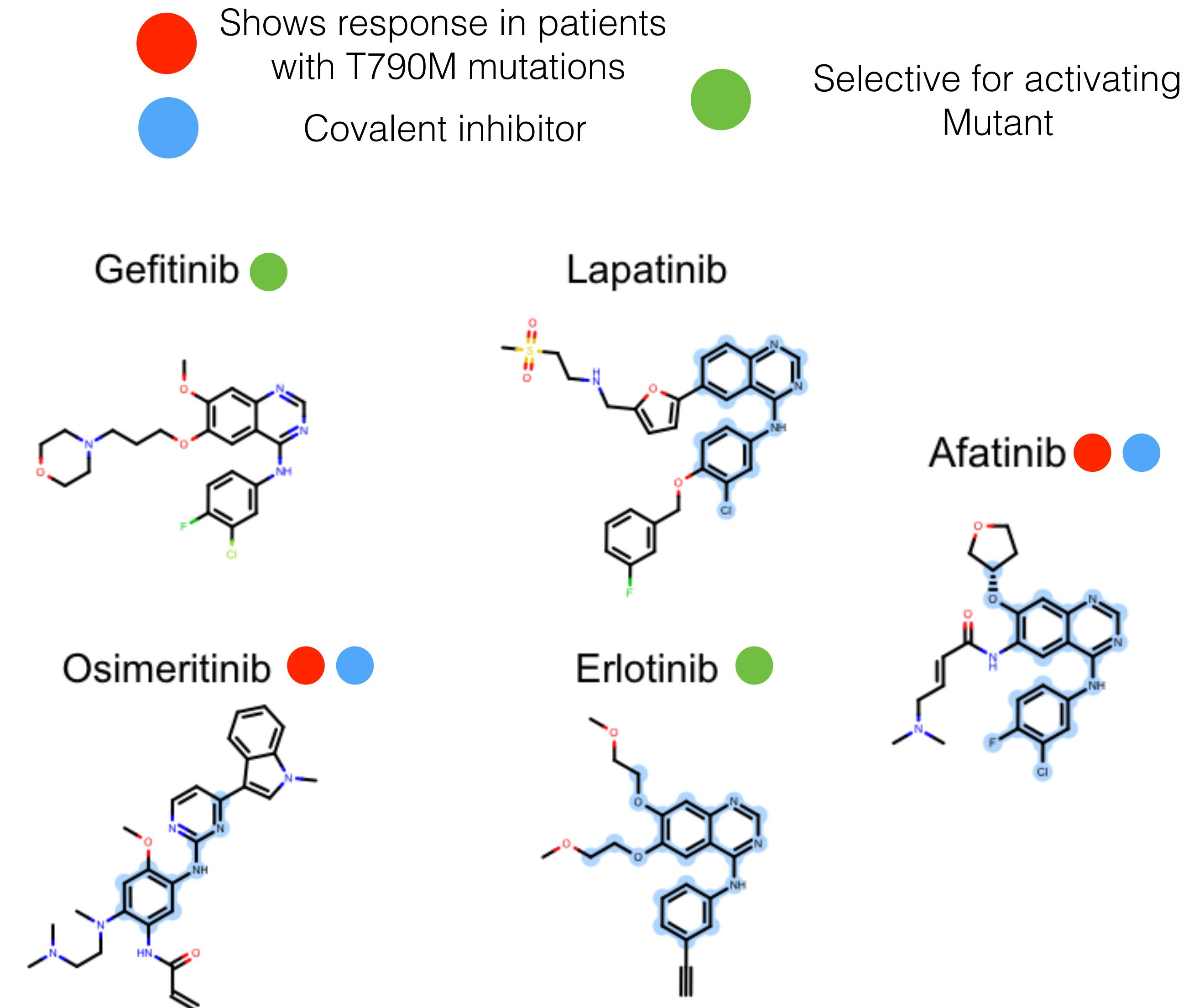
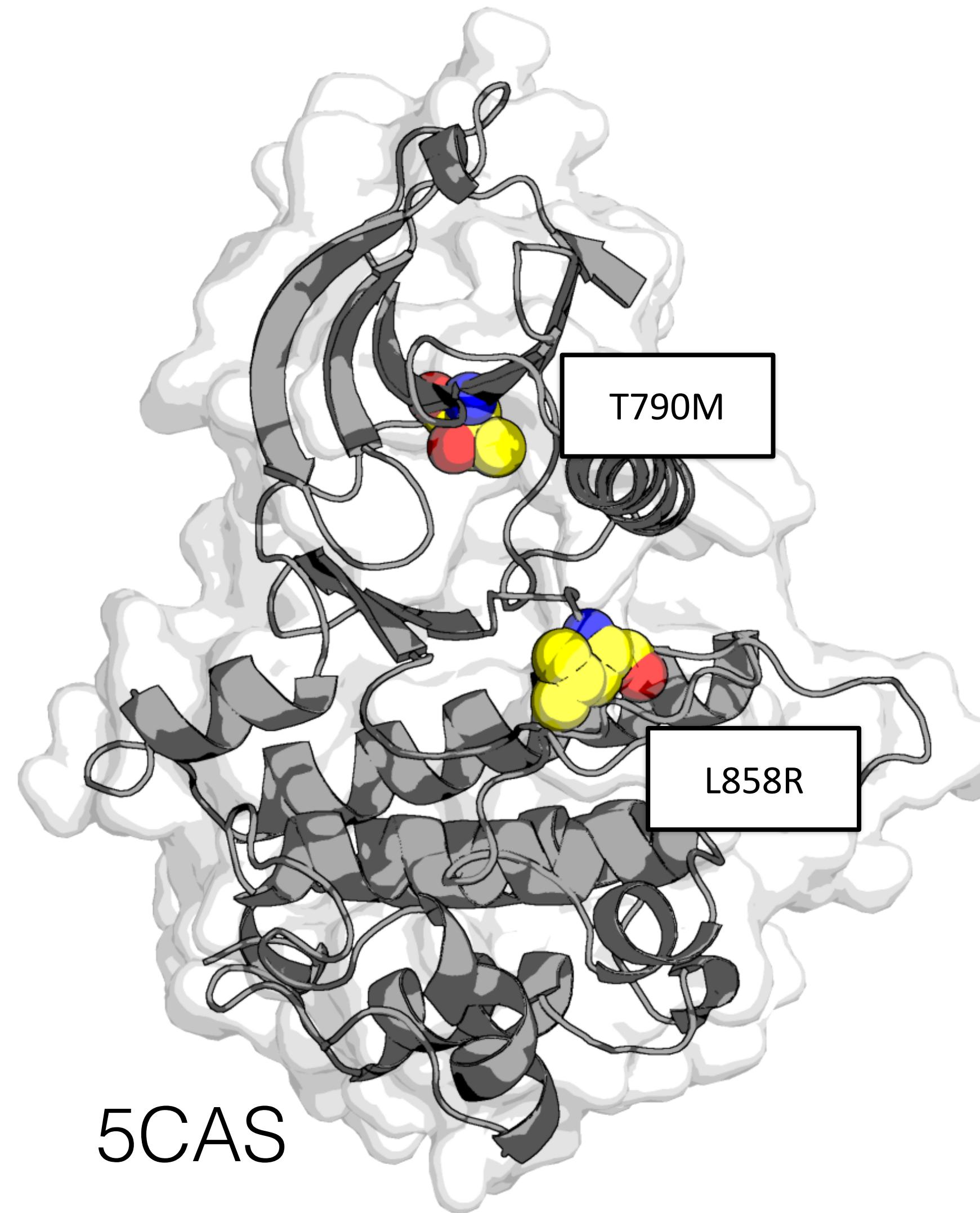
Bayesian inference allows us to better estimate the errors and correlation



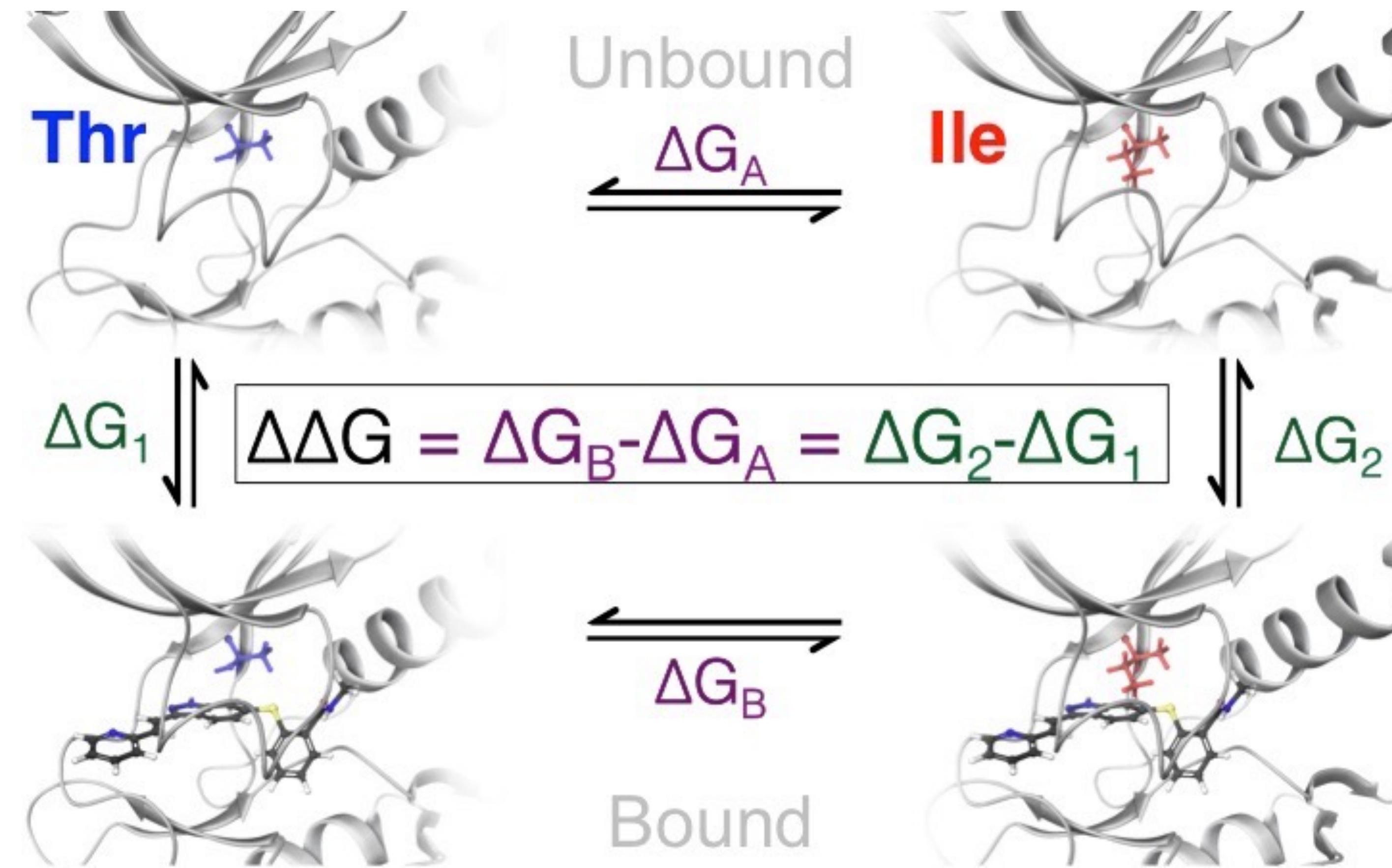
True RMSE: 0.912 [0.778, 1.051] kcal/mol (95% CI)
True MUE: 0.727 [0.620, 0.838] kcal/mol (95% CI)



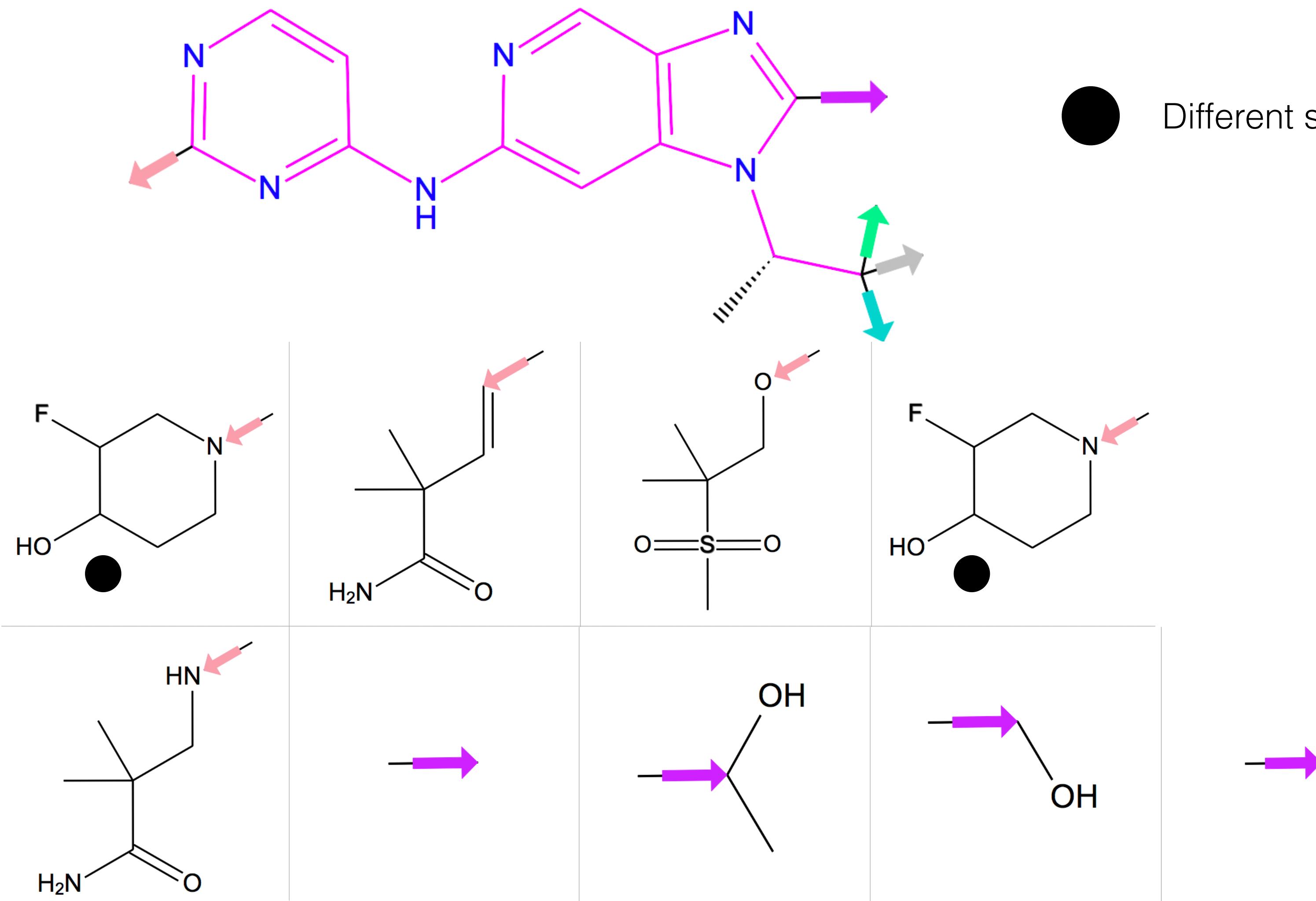
Predicting mutant selectivity and using protein mutation FEP to enable cycle closure between targets



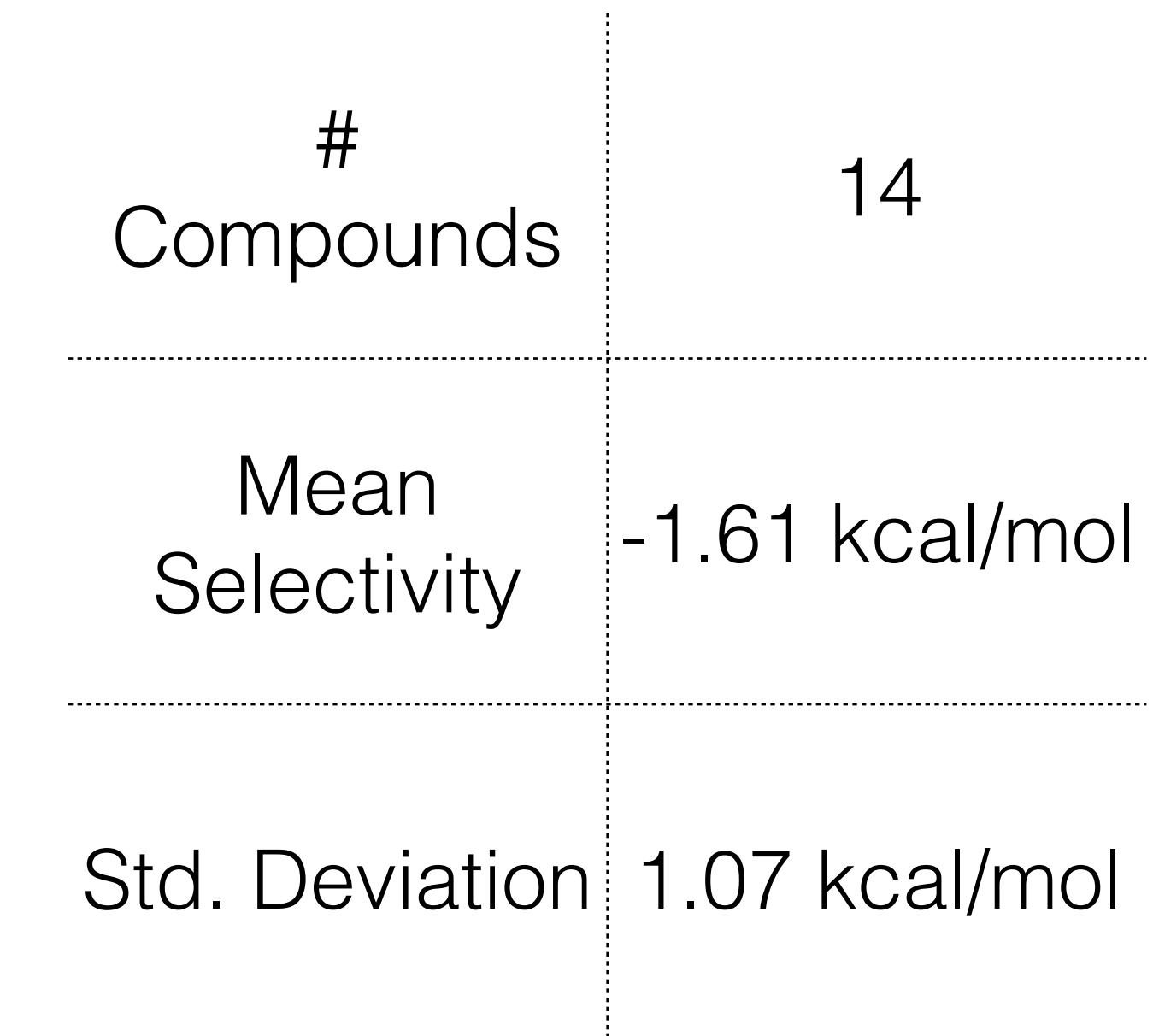
Protein mutation FEP uses a thermodynamic cycle to estimate $\Delta\Delta G_{\text{mutation}}$



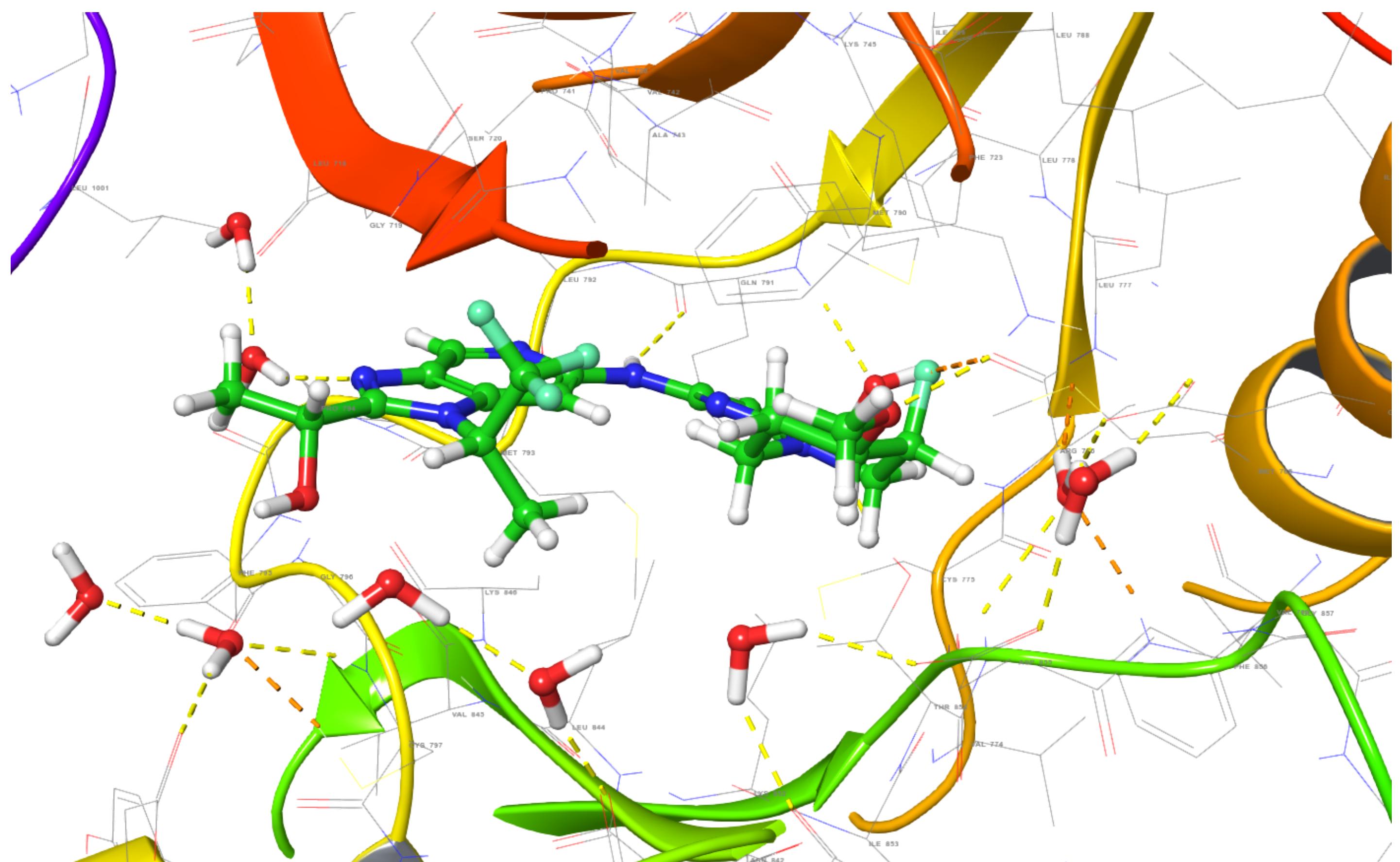
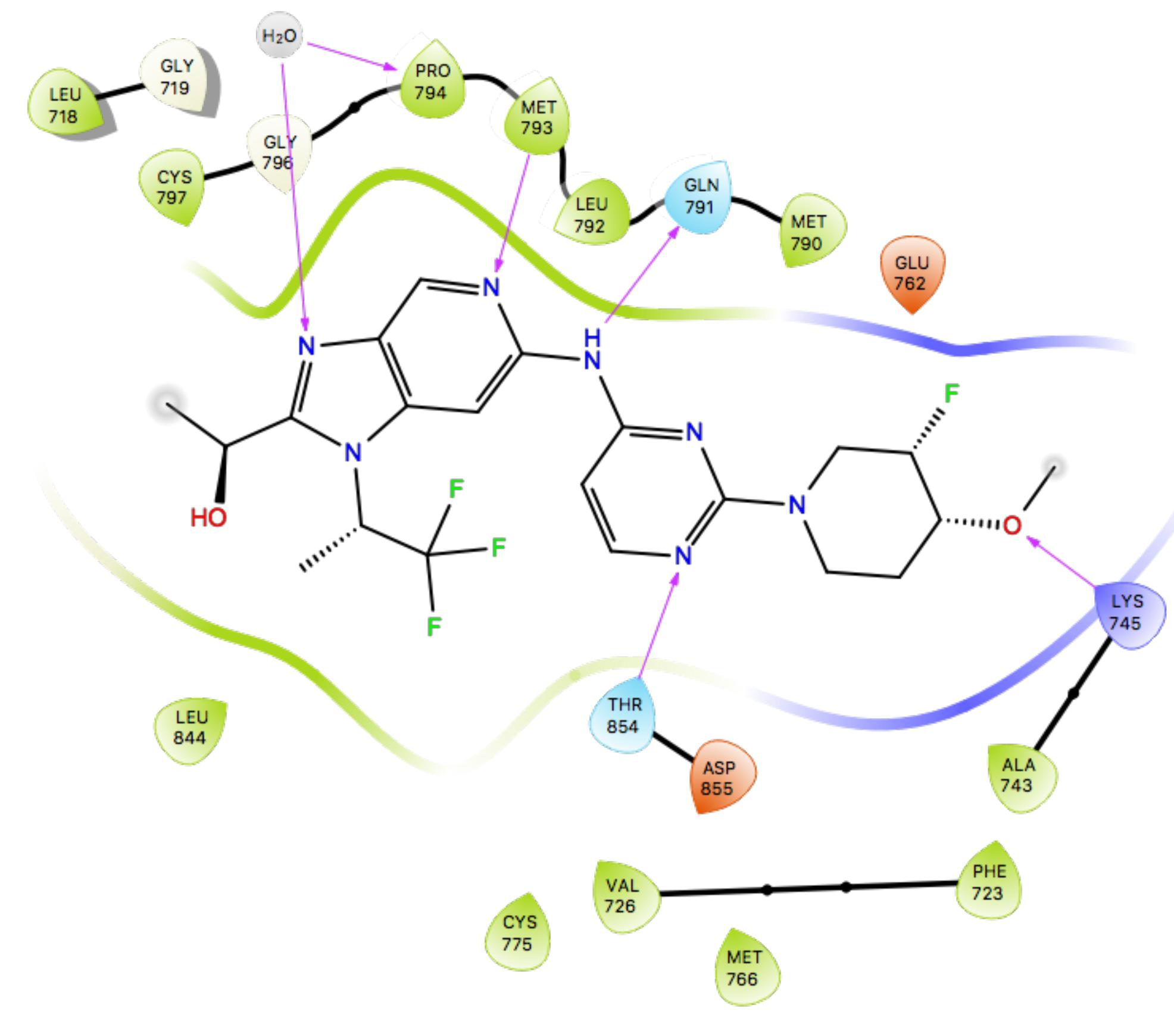
A lead series of compounds for wild type and double mutant EGFR



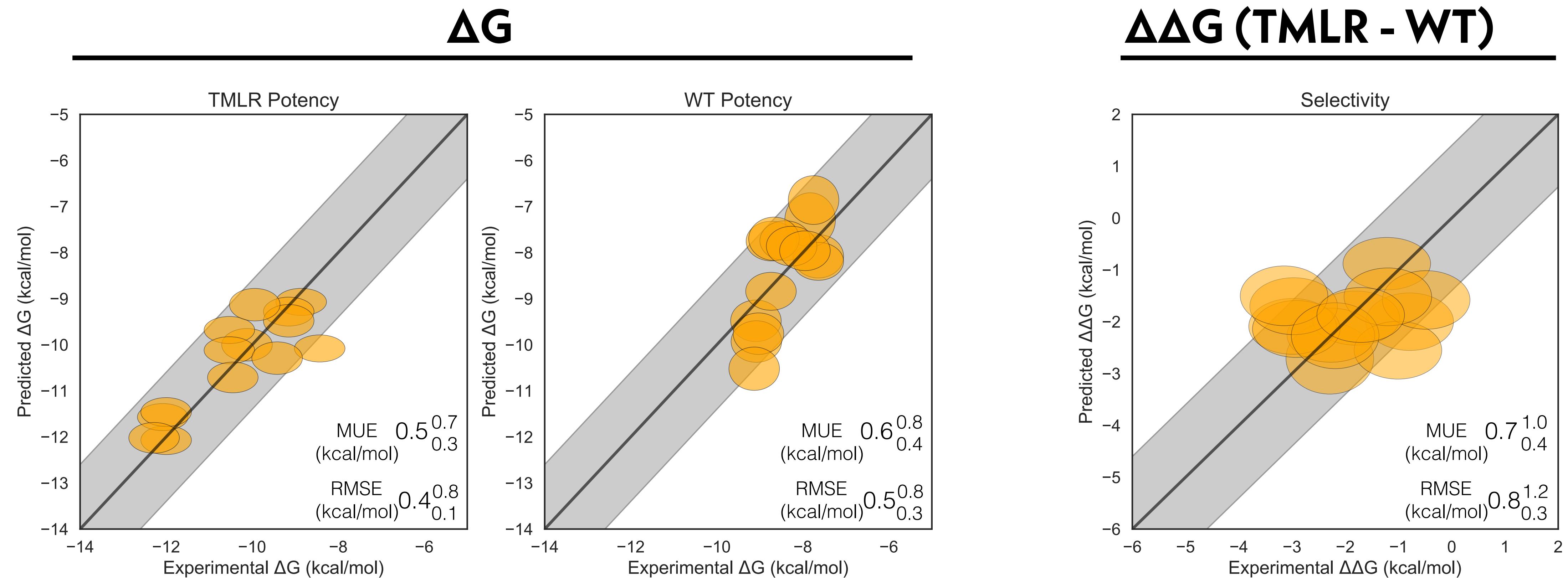
● Different stereochemistry



A lead series of compounds for wild type and double mutant EGFR

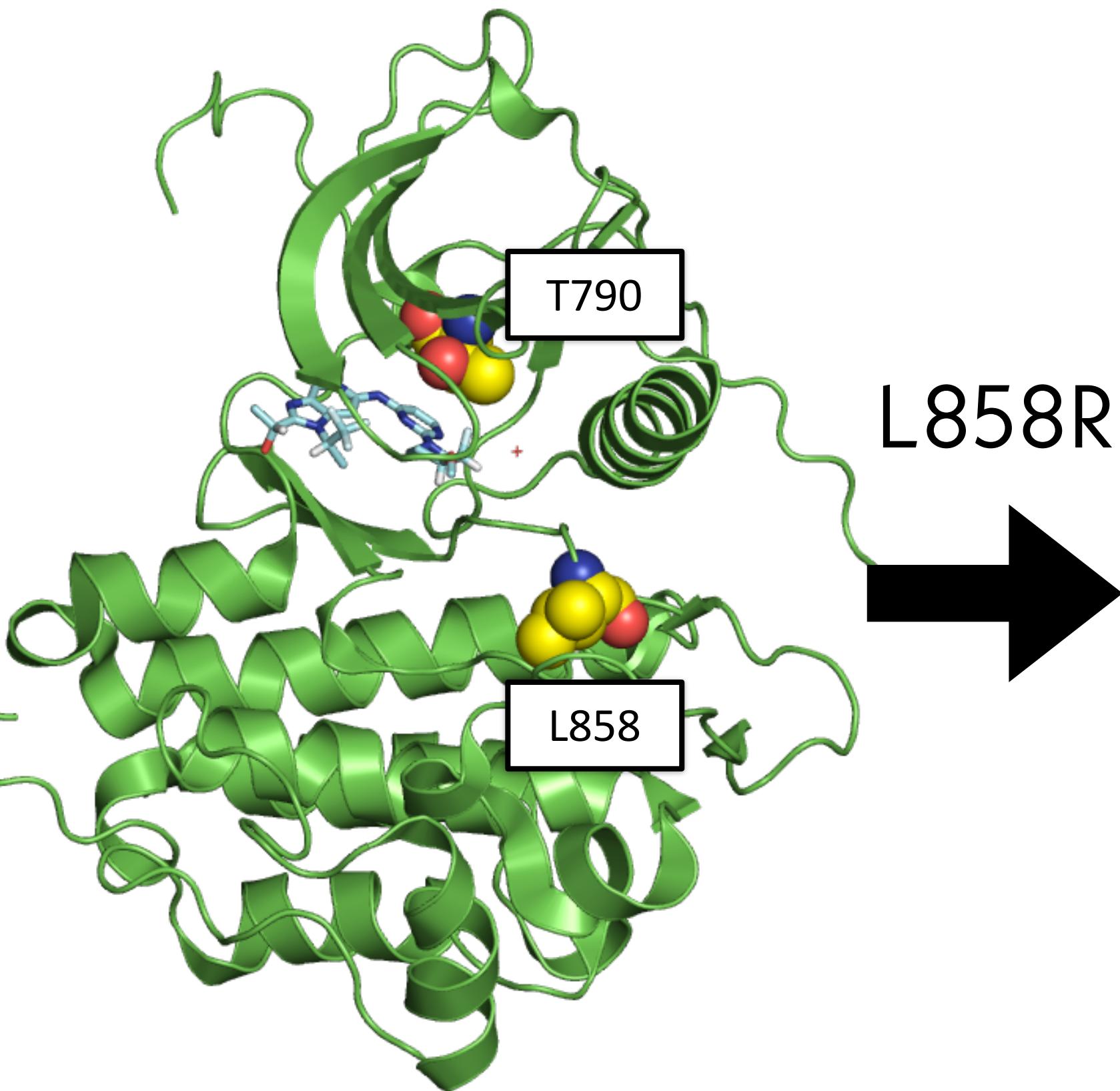


Predicting mutant selectivity for wtEGFR/TMLR



Protein mutation FEP enables cycle closure between targets and could reduce the uncertainty between targets

wtEGFR (5CAV)



Chemical structures of the compounds:

Structure 9: 2-(2-methylimidazo[1,2-b]pyridin-3(2H)-yl)acetonitrile

Structure 10: 2-(2-methylimidazo[1,2-b]pyridin-3(2H)-yl)acetamide

Structure 11: 2-(2-methylimidazo[1,2-b]pyridin-3(2H)-yl)propanamide

Structure 12: 2-(2-methylimidazo[1,2-b]pyridin-3(2H)-yl)methylcyclobutanemethanone

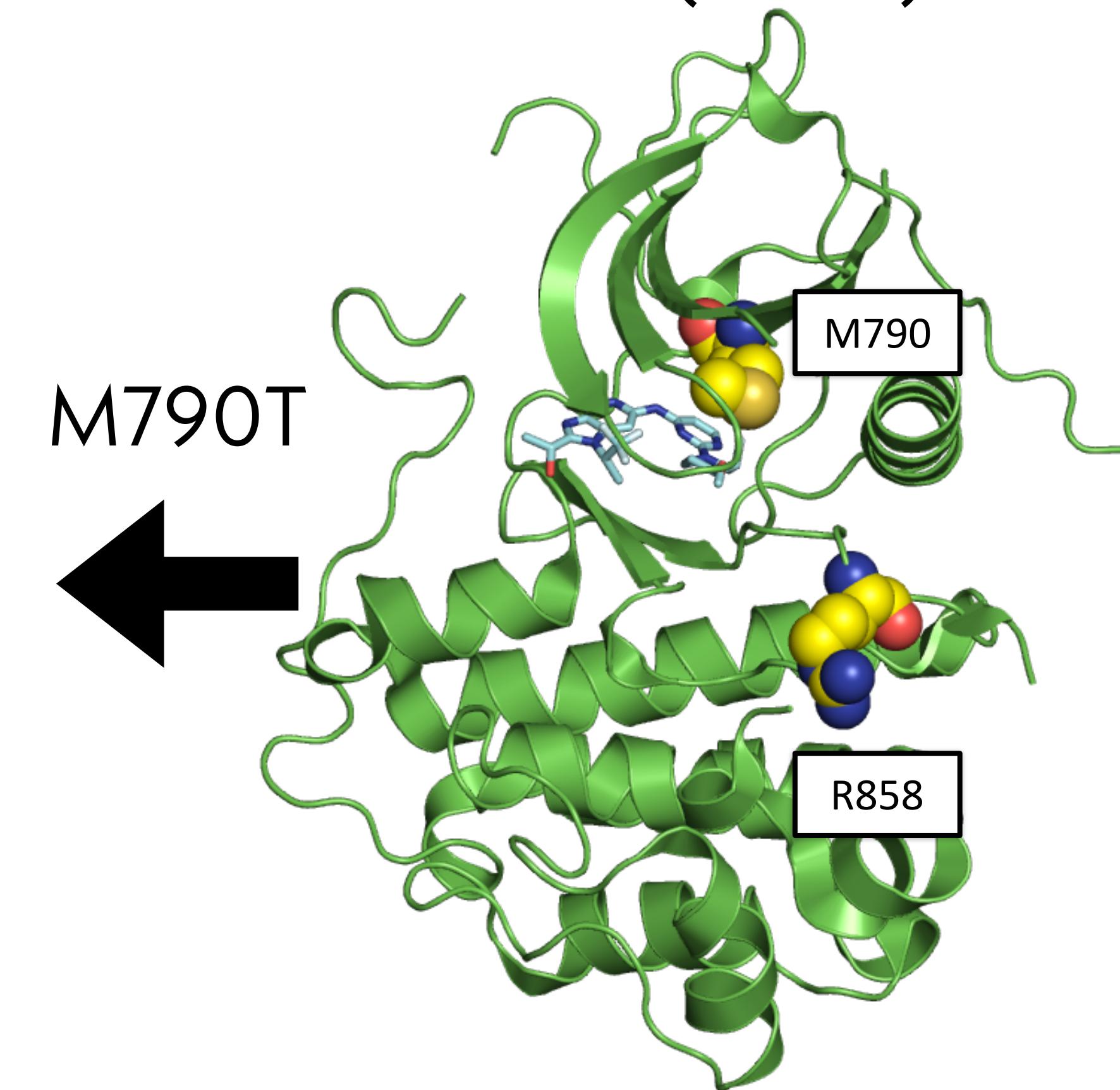
Structure 13: 2-(2-methylimidazo[1,2-b]pyridin-3(2H)-yl)methylcyclohexanemethanone

Structure A: 2-(2-methylimidazo[1,2-b]pyridin-3(2H)-yl)sulfone

Structure B: 2-(2-methylimidazo[1,2-b]pyridin-3(2H)-yl)methylcyclohexane

Cmpd	X	R ¹	R ²	TMLR Ki _{app} (nM) ^a	TMdel Ki _{app} (nM) ^a	L858R Ki _{app} (nM) ^a	del ₇₄₆₋₇₅₀ Ki _{app} (nM) ^a	wtEGFR Ki _{app} (nM) ^a
9	CH	A		1.1	0.35	4.1	1.2	7.8
10	CH	B		5.3	27	363	323	167
11	N	B		29	-	-	-	>1000
12	CH	A		1.1	2.0	6.4	2.0	4.6
13	CH	A		1.2	1.6	3.8	1.4	3.4

TMLR (5CAS)



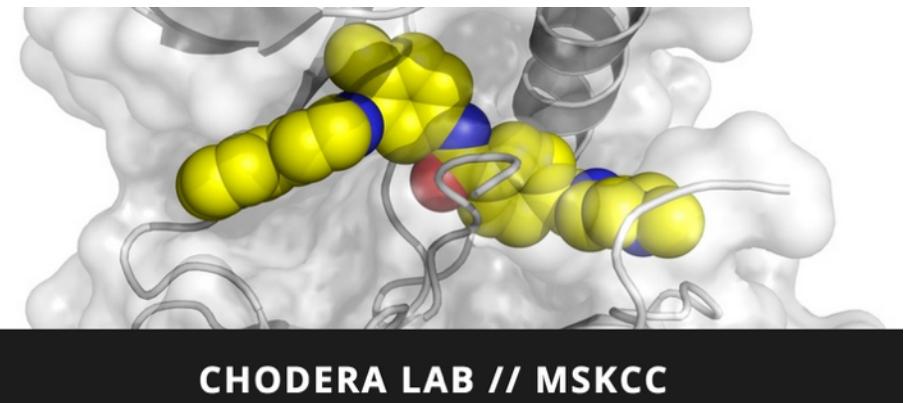
Future Directions

- **Outlier analysis for KDR and finish calculations for Tie2**
- **Analyze second set of EGFR wild type and double mutant small molecule FEP+ calculations**
- **Complete protein mutation FEP calculations and develop cycle closure analysis**
- **Further develop Bayesian model to assess how correlated the errors in each target are with each other**
- **Quantify the dynamic range of selectivity needed to confidently identify a selective compound (i.e., the utility of this method for predicting this property)**

Takeaway Message

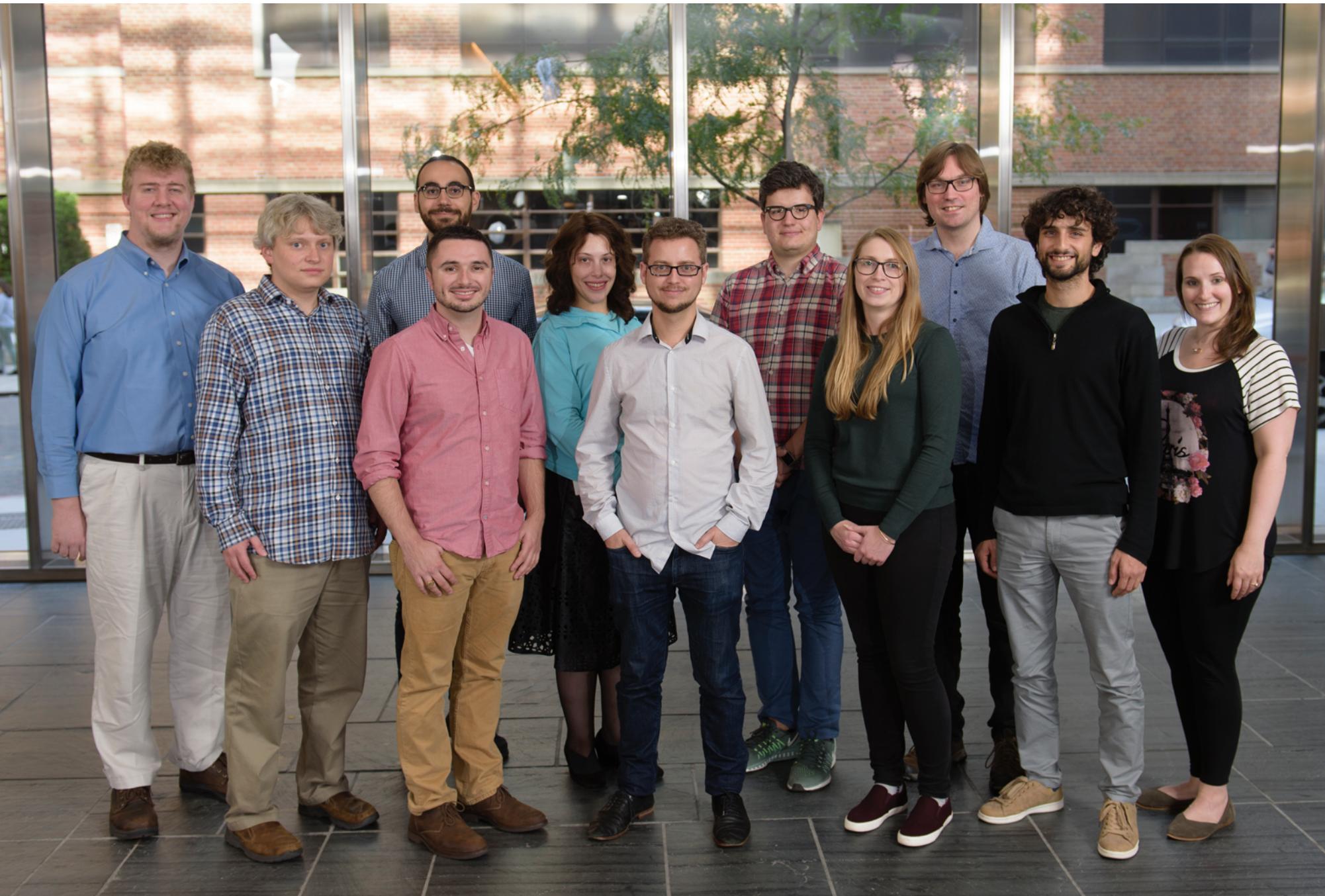
- **Selectivity is difficult to predict due to the larger uncertainties when considering multiple targets (both experimental and predicted), even when the errors for the individual targets are small**
- **At first pass, the errors in calculation largely seem uncorrelated (i.e, random) even for closely related targets**
- **Novel cycle closure methods may be able to lower the uncertainty of our predictions**

Acknowledgements



[John Chodera](#)
[Levi Naden](#)
[Patrick Grinaway](#)
[Chaya Stern](#)
[Bas Rustenburg](#)
[Mehtap Isik](#)
[Rafal Wiewiora](#)
[Josh Fass](#)
[Andrea Rizzi](#)
[Lucelenie Rodriguez](#)

Former Members
[Sonya Hanson](#)
[Gregory Ross](#)
[Julie Behr](#)
[Danny Parton](#)



Schrödinger

Robert Abel
[Lingle Wang](#)
[Kevin Hauser](#)
[Simon Peng](#)

Gunner Lab
[Marilyn gunner](#)
[Salah Salah](#)
[Elizabeth Rosenzweig](#)

Thesis Committee

Sarat Chandarlapaty
Daniel Heller
Robert Abel

Kinase experimentalists
Levinson Lab
Seeliger Lab

Clinical Mentor
[Helena Yu](#)