ACADEMY OF ATHENS
BRFAA
BIOMEDICAL RESEARCH FOUNDATION
ACADEMY OF ATHENS

grnet

UMASS
AMHERST
Department of Biochemistry
and Molecular Biology

D3R Grand Challenge 2

Free Energy Perturbation calculations to predict relative binding affinities for FXR ligands

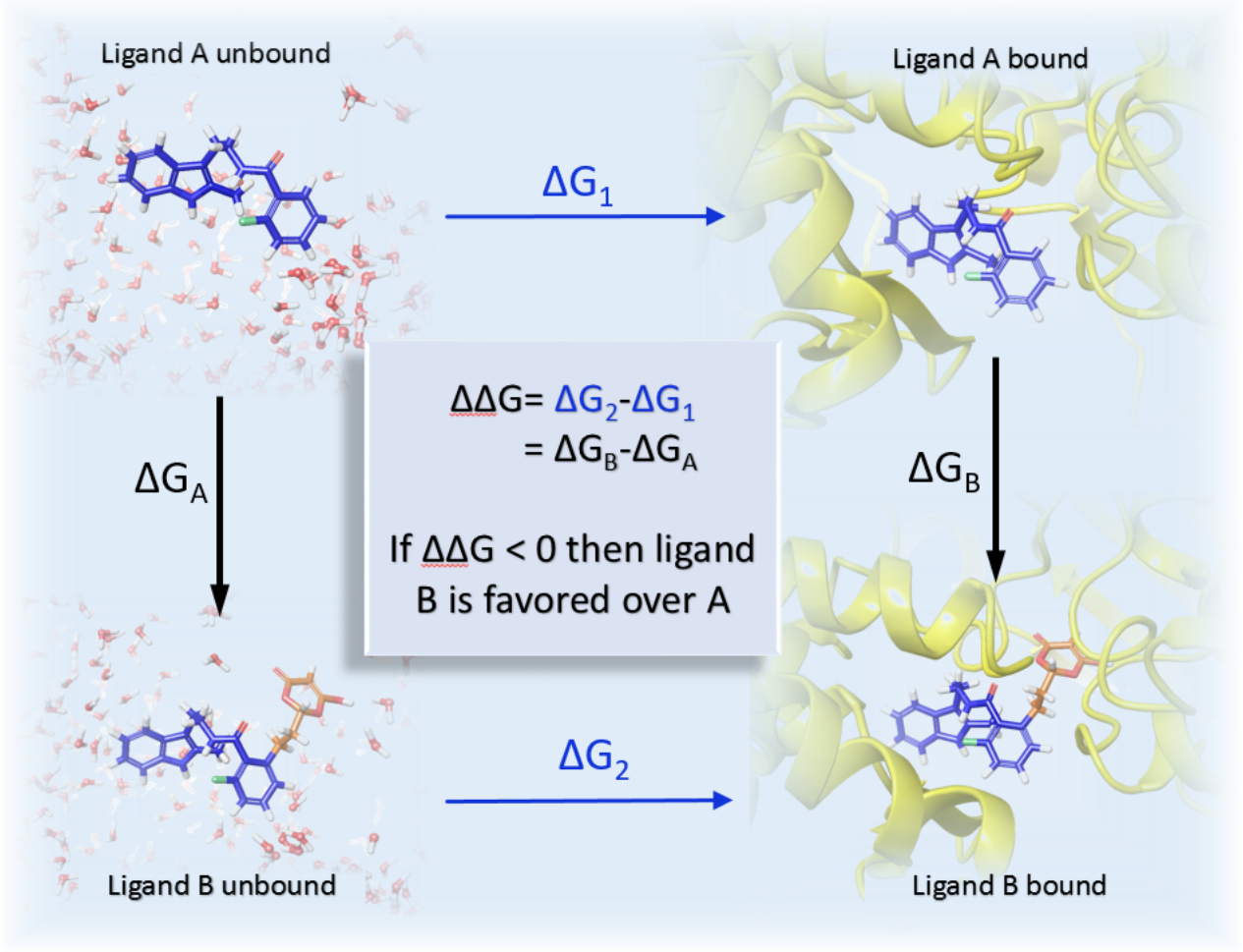
Christina Athanasiou, Sofia Vasilakaki, Dimitris Dellis,

Woody Sherman and Zoe Cournia

27 March 2017

Free Energy Perturbation (FEP) calculations

Zwanzig's formula in complex and in solvent: $\Delta G(A \to B) = G_B - G_A = -kT \ln \left\langle \exp\left(-\frac{V_B - V_A}{kT} \right) \right\rangle_A$ (MBAR used in practice)



ΔG_1 and **ΔG_2** are the free energies of **transfer** of A and B from the unbound to the bound state

ΔG_A and **ΔG_B** are the free energy differences of the **mutation of A into B** in solvent and bound to protein

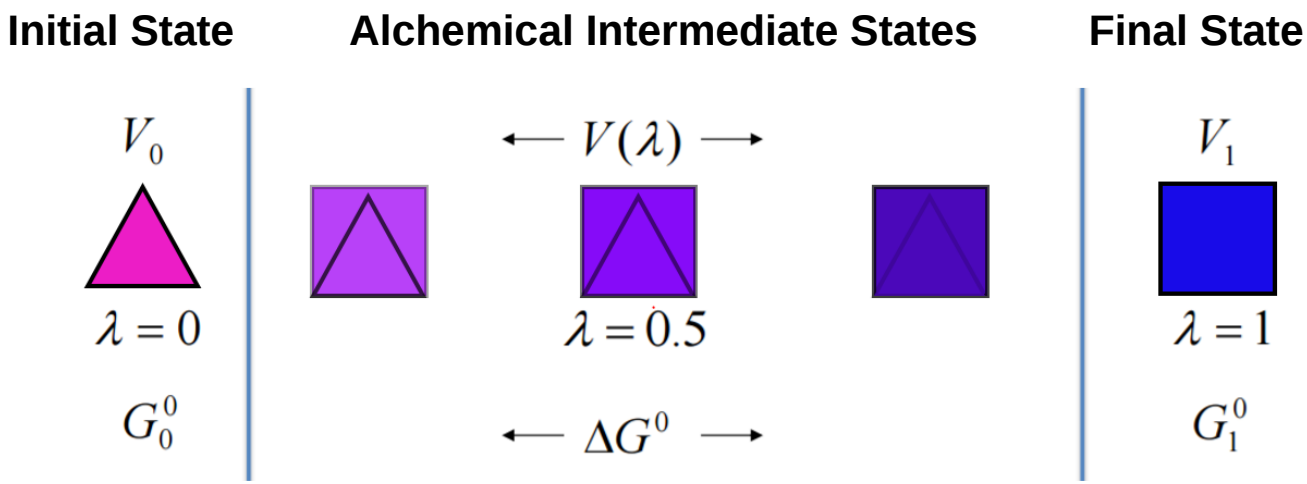
Molecular dynamics - REST

Molecular Dynamics

Time evolution of the system (Desmond/FEP+, Schrödinger)

Allows for fully flexible receptor → Advantage over docking

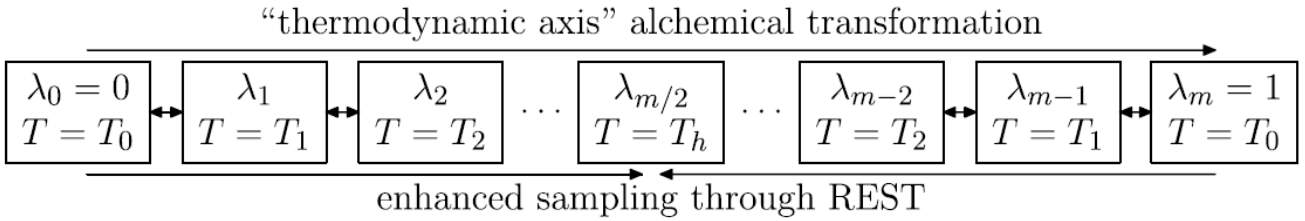
Perturbation is achieved with a λ schedule



REST (Replica Exchange with Solute Tempering)

Usual problem in FEP calculations:
- Inability of convergence
- Need for long computational time

Solution:
- Increase of effective temperature to overcome energetic barriers
- Replica exchange between neighboring λ windows



Wang, Berne, Friesner, On achieving high accuracy and reliability in the calculation of relative protein–ligand binding affinities, PNAS 2012; 109(6): 1937-1942.

Issues for consideration before FEP

- The input structure has to be of sufficiently high quality
- Sufficiently long simulations and methodology to overcome barriers (sampling)
- Cannot change the ligand charge during a mutation
- Sensitive to force field (scoring)
- Examination of buried waters (WaterMap, Schrödinger)
- Ensure perturbations are not too big (normally up to 10 heavy atoms)
- Error of the method ~1 kcal/mol
- Large-scale protein movements cannot be sampled sufficiently within the timeframe of FEP calculations

Setting up FEP calculations

This methodology was followed for both spiros and sulfonamides subsets

1) All ligands must belong in the **same congeneric series**

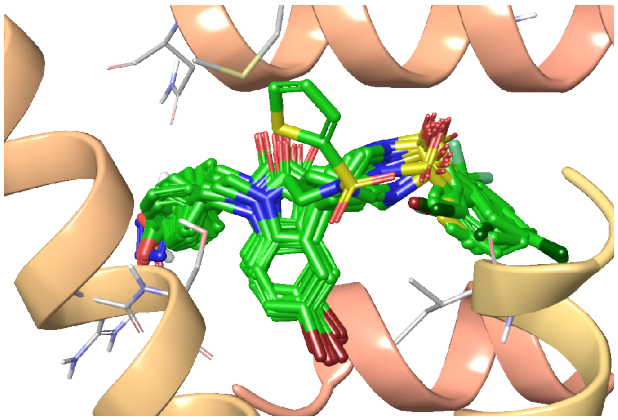
2) Choice of **reference ligand**
✔ It has to be representative of the series

3) Alignment of all ligands into the reference ligand
✔ After this step, a minimization of the complexes is usually needed

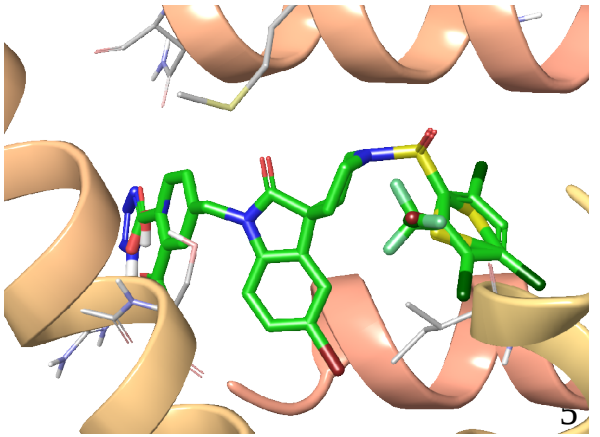
• The input structure has to be of sufficiently high quality: Examples follow with predicted structure and with real crystal structure

Spiros group example

Docking based alignment



Maximum common substructure alignment



Setting up FEP calculations

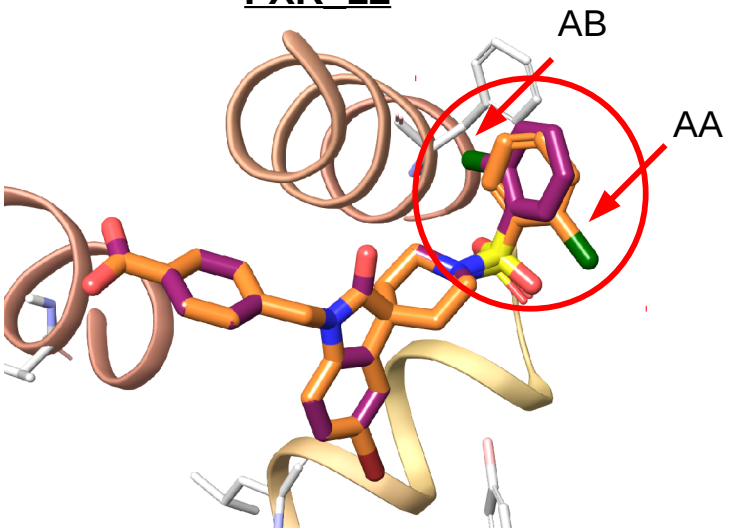
4) If **double occupancy** is plausible, both binding modes should be considered in the calculations

✔ If they rapidly inter-convert during the simulation, the same ΔΔG is expected, and we can just ignore one of them.

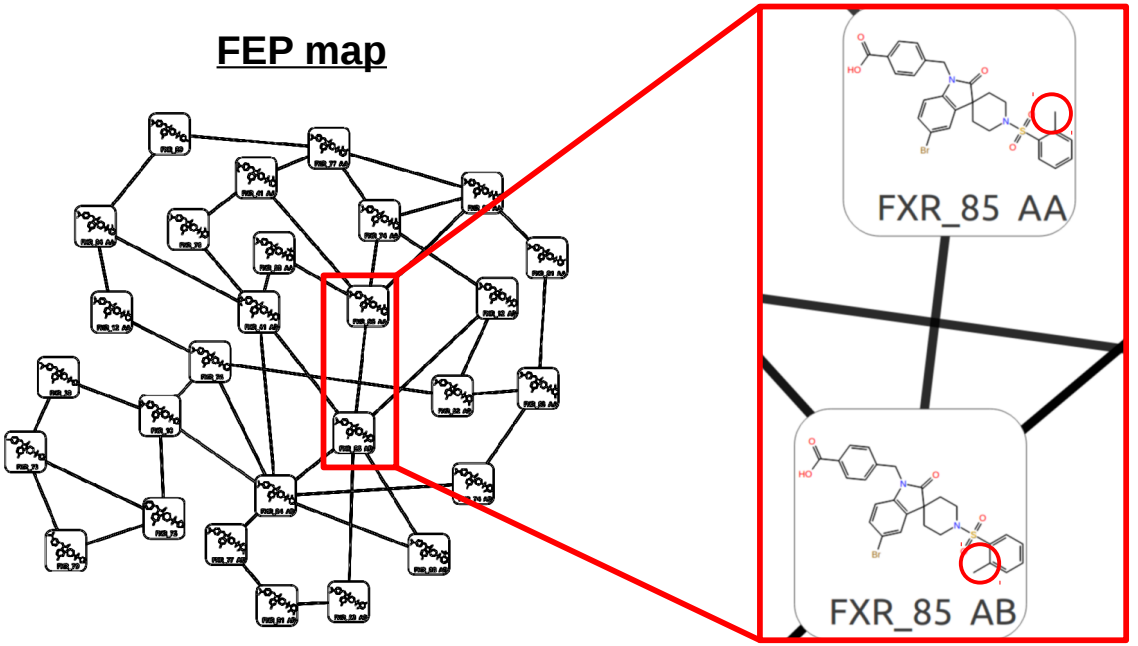
✔ If one pose is significantly less stable, discount it from the results

✔ If both compounds maintain separate binding poses, but result to the same binding free energy, we can correct the binding free energy for multiple poses (Joseph *et al.*, JCTC, 2015).

Compound with double occupancy
FXR_12



FEP map



Setting up FEP calculations

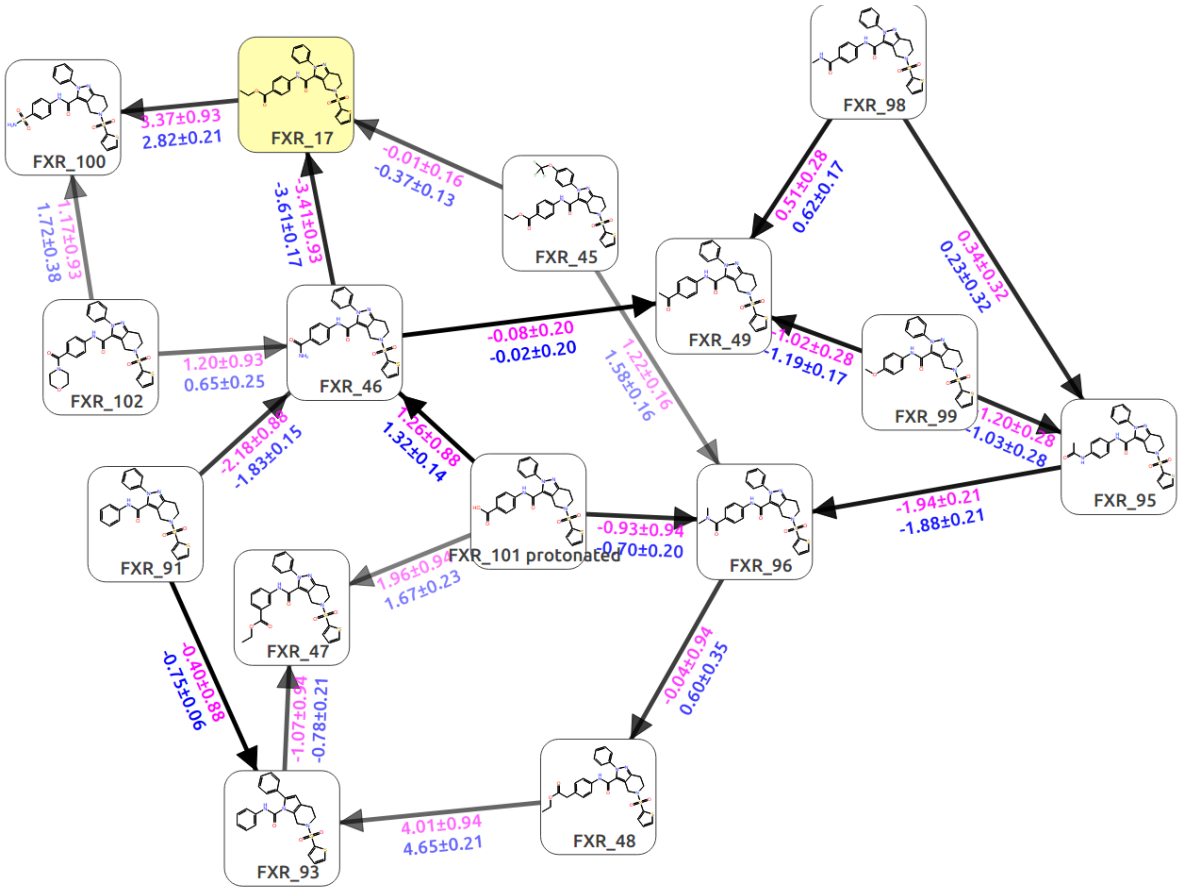
5) Cannot change the ligand charge during a mutation

✔ We chose to calculate both the sulfonamides and spiros subset in their neutral forms in order not to change the charge.

6) FEP Mapper (FEP+)

✔ Initial ligands structures as described above are imported
✔ Ligands are connected through edges based on chemical & binding mode similarity, preservation of ring structure
✔ User can define the cycles
✔ Every molecule must be part of at least one closed thermodynamic cycle
✔ Each edge represents the bound & unbound perturbations

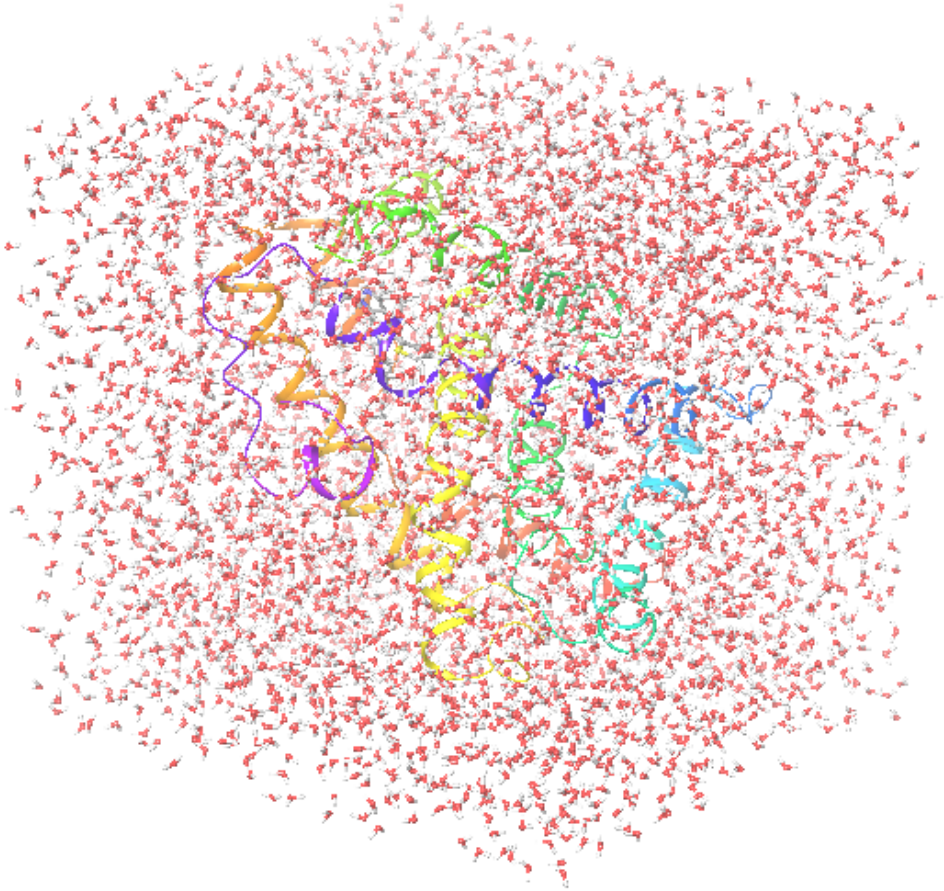
Sulfonamides FEP map



Raw errors are predicted from bootstrapping & analytical error of the BAR free energy
Cycle closure errors are calculated though the hysteresis of each cycle

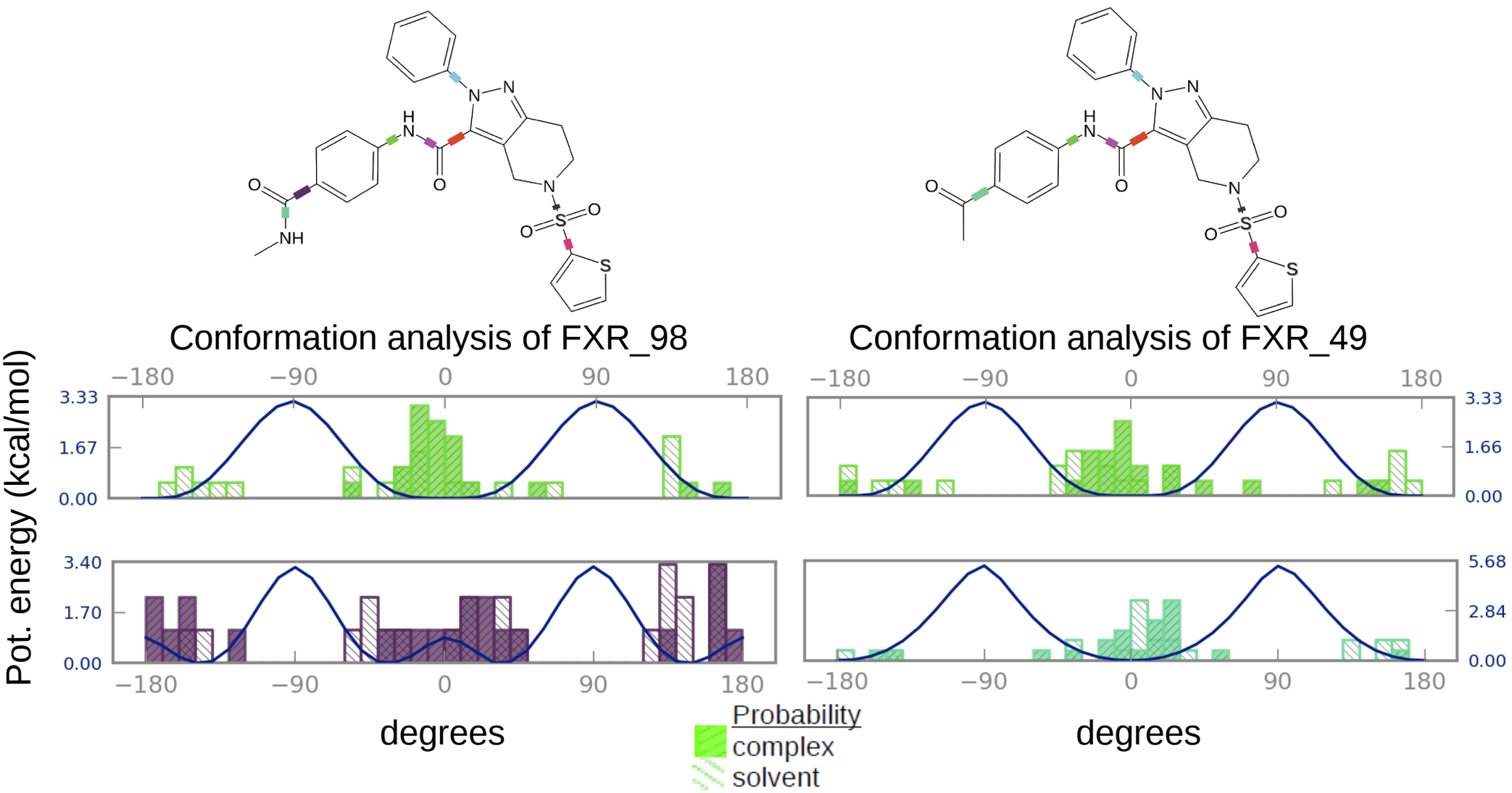
Running FEP calculations

- ✔ Building of final system geometry
- ✔ OPLS3 force field assignment
- ✔ **Equilibration**
  - ✔ Brownian dynamics with restraints on solute heavy atoms (NVT, T = 10 K, 100 ps, force constant = 50 kcal/mol/Å^2)
  - ✔ MD simulation with restraints on solute heavy atoms (NVT, T = 10K, 12 ps)
  - ✔ MD simulation with restraints on solute heavy atoms (NPT, 36 ps)
  - ✔ MD simulation with no restraints (240 ps)
- ✔ **Production Run**
  - ✔ REST MD simulation (NPT, 5 ns)
- ✔ **FEP analysis**



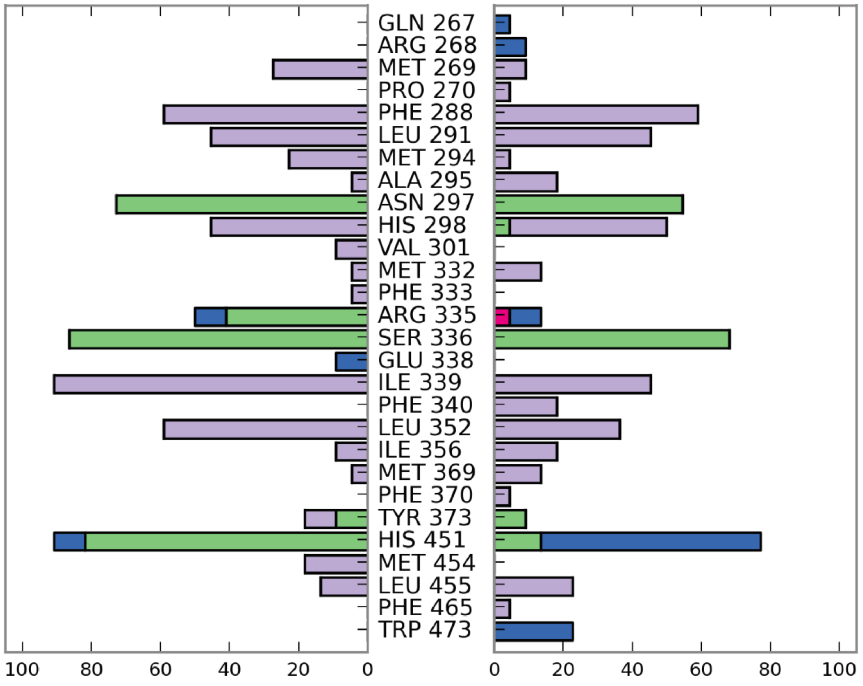
<u>System Size</u>
Spiros: 21,000 atoms
Sulfonamides: 17,000 atoms

Ligand Conformation Analysis



The force field torsional energy profile (blue curves) are shown superimposed with the probability from the FEP calculations in the complex (solid) and solvent (hashes) simulations for each of the two rotatable bonds in the ligands.

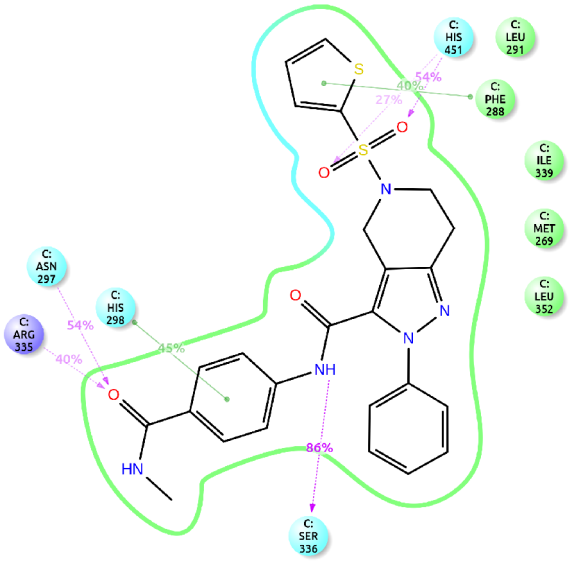
Example of simulation interaction diagram (SID) from FEP calculation



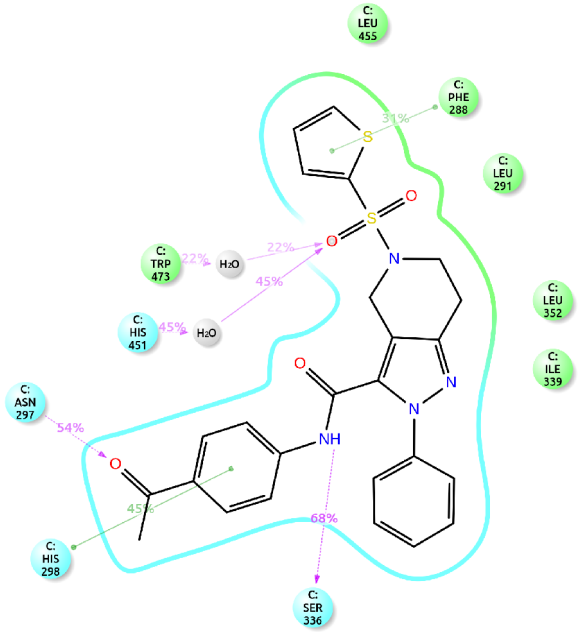
% interactions with FXR_98 % interactions with FXR_49

Interactions:
- Hydrogen bond
- Hydrophobic
- Ionic
- Waters bridges

Comparison of interactions with receptor residues between the two ligands.



FXR_98



FXR_49

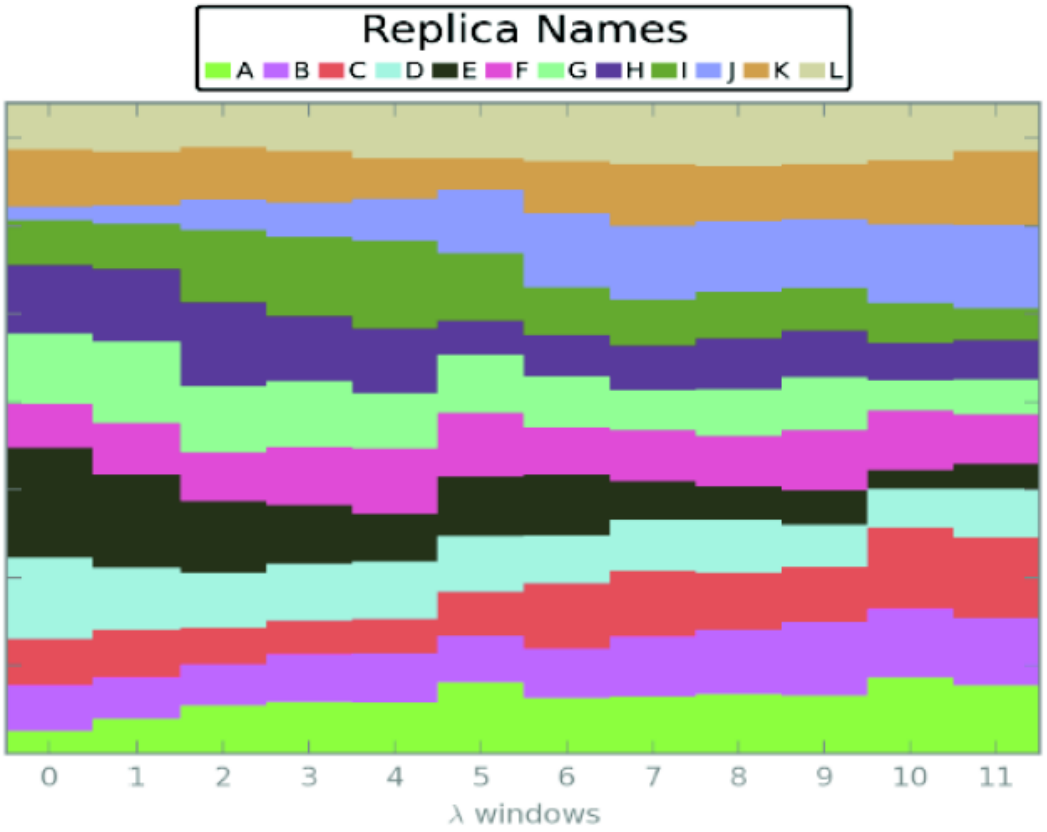
Ligand Interaction Diagram (LID) for each of the ligands with residues contacting the ligands and the percent of the simulation spent making each interaction.

FXR_98 IC_{50} = 13.1 uM

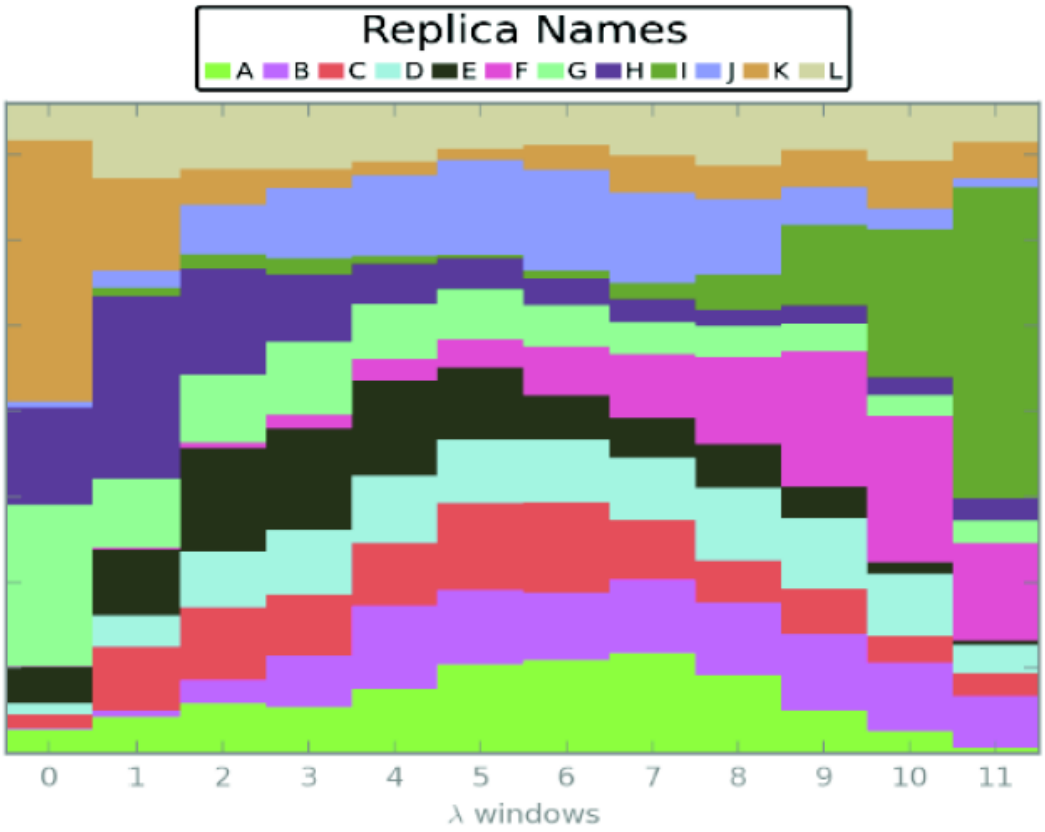
FXR_49 IC_{50} = 100 uM

Exchange density of FEP replicas over λ windows

Solvent Leg



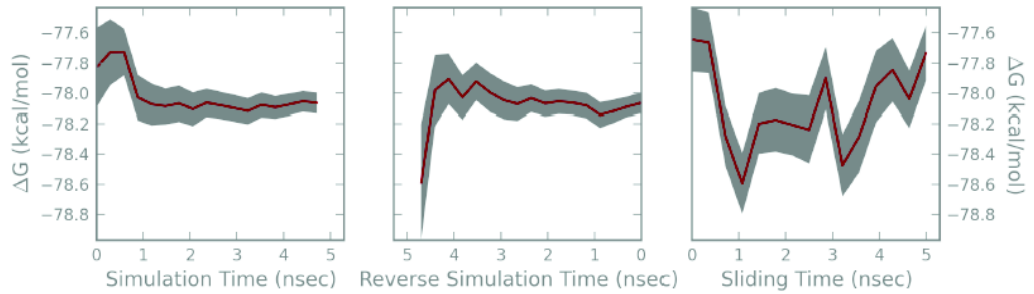
Complex Leg



For both legs of the FXR_98 to FXR_49 simulation, each replica is color coded and the plot shows how it occupies different λ windows during the course of the simulation.

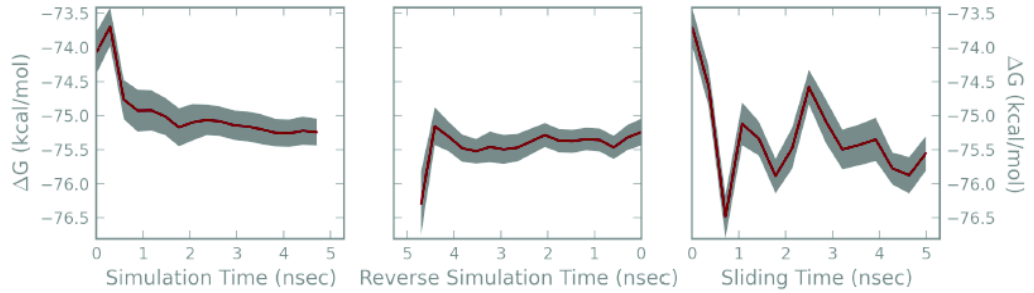
Free energy convergence

Solvent Leg



λ pair:	0,1	1,2	2,3	3,4	4,5	5,6	6,7	7,8	8,9	9,10	10,11	Total
Forward dF:	18.06	12.26	7.99	4.79	2.92	-0.54	-3.78	-14.76	-22.16	-33.08	-49.77	-78.07
Bootstrap STD:	0.026	0.016	0.021	0.014	0.014	0.017	0.017	0.020	0.016	0.019	0.038	0.069
Analytical STD:	0.017	0.015	0.013	0.012	0.016	0.017	0.017	0.015	0.016	0.022	0.036	0.063

Complex Leg



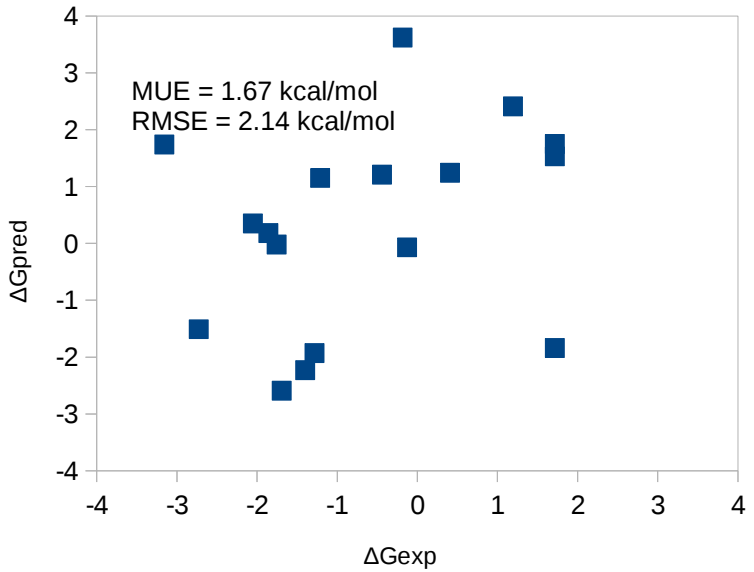
λ pair:	0,1	1,2	2,3	3,4	4,5	5,6	6,7	7,8	8,9	9,10	10,11	Total
Forward dF:	18.51	13.09	8.85	5.53	4.04	-0.44	-4.60	-15.43	-22.99	-33.47	-48.33	-75.24
Bootstrap STD:	0.091	0.063	0.042	0.039	0.034	0.039	0.046	0.044	0.043	0.076	0.096	0.197
Analytical STD:	0.016	0.014	0.013	0.012	0.017	0.018	0.018	0.016	0.015	0.019	0.027	0.057

The total free energy differences between the two ligands (ΔG in kcal/mol) in solvent and complex legs are plotted as a function of time. Three plots for each leg show the accumulated data during different time window schemes; forward; reverse; and sliding window.
The tables report the associated bootstrap and analytical errors estimates from corresponding simulation legs.

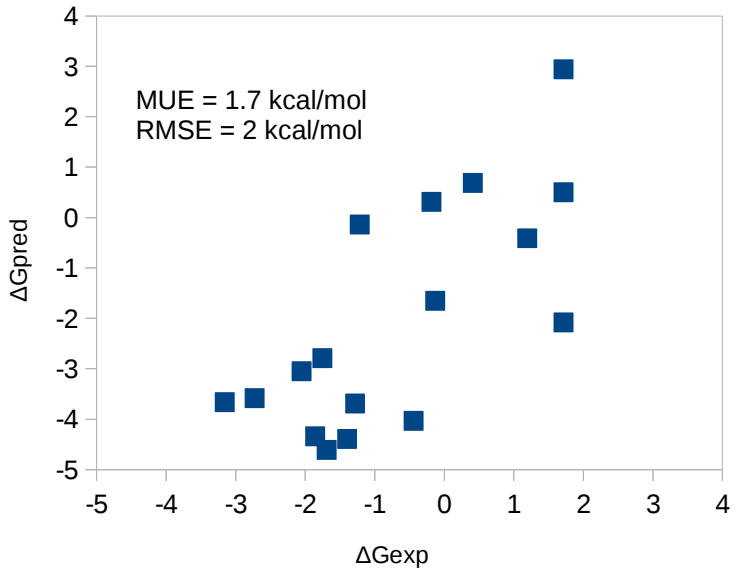
Spiros series results

#LigandID	Predicted 1st stage	Predicted 2nd stage	Experimental	IC50 (uM)
FXR_10	0	0	0	5.64
FXR_12	-1.51	-3.58	-2.73	0.058
FXR_38	1.53	0.5	1.71	100
FXR_41	-1.84	-2.08	1.71	100
FXR_73	1.24	0.69	0.41	11.2
FXR_74	-1.93	-3.69	-1.28	0.655
FXR_75	1.75	2.94	1.71	100
FXR_76	2.41	-0.41	1.19	41.8
FXR_77	0.18	-4.34	-1.86	0.25
FXR_78	1.74	-3.66	-3.16	0.0283
FXR_79	3.62	0.31	-0.18	4.15
FXR_81	1.21	-4.03	-0.44	2.69
FXR_82	0.35	-3.05	-2.05	0.18
FXR_83	-2.59	-4.61	-1.69	0.33
FXR_84	-0.07	-1.65	-0.13	4.54
FXR_85	-0.02	-2.79	-1.75	0.297
FXR_88	-2.23	-4.39	-1.40	0.54
FXR_89	1.15	-0.14	-1.21	0.735

1st stage



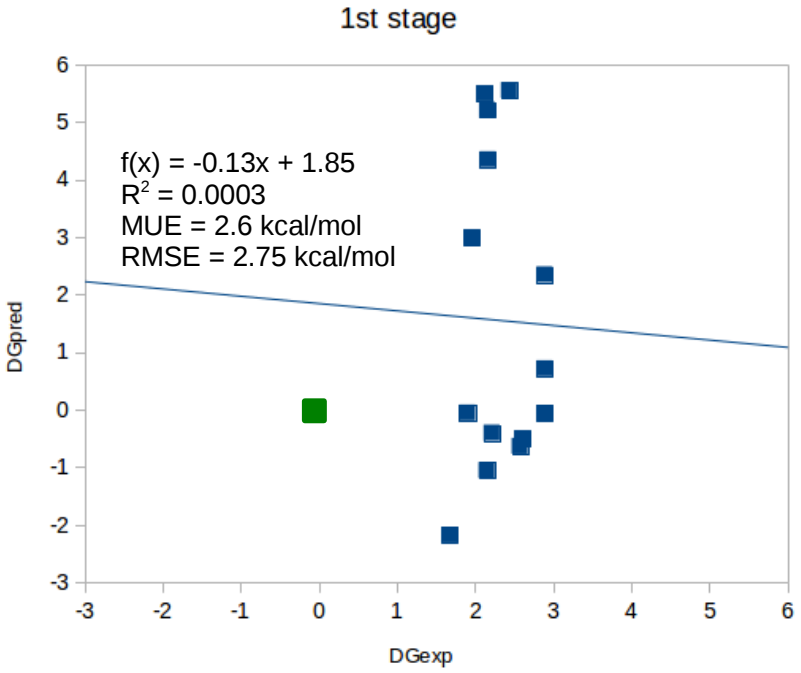
2nd stage

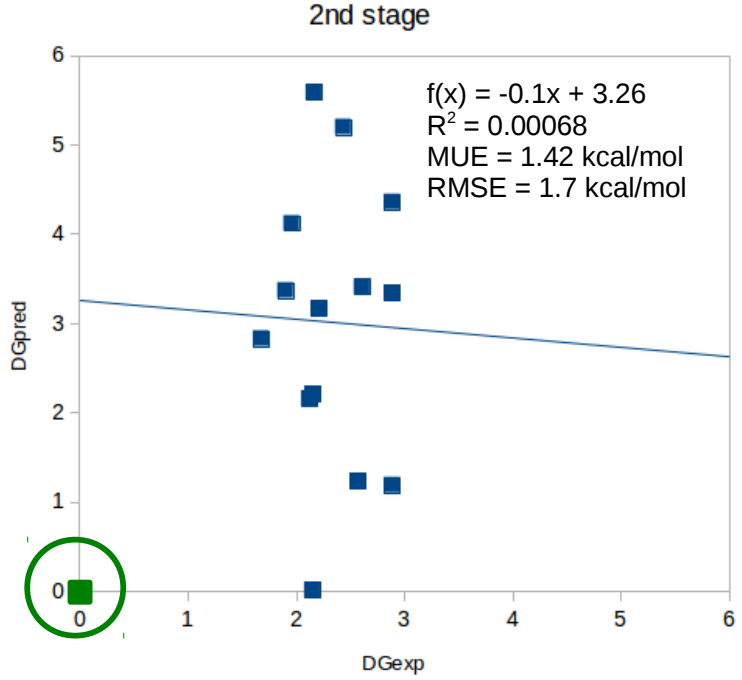


- Slope improves with crystal structure
- **85% true positives**
- Only 2/13 false positives

Sulfonamides series results

#LigandID	Predicted 1st stage	Predicted 2nd stage	Experimental	IC50 (uM)
FXR_17	0	0	0.00	0.785
FXR_45	-1.05	0.01	2.15	28.9
FXR_46	-0.5	3.41	2.61	62.4
FXR_47	2.99	4.12	1.96	21
FXR_48	0.72	1.19	2.89	100
FXR_49	-0.06	3.34	2.89	100
FXR_91	4.35	5.59	2.16	29.6
FXR_93	5.55	5.2	2.44	46.7
FXR_95	-0.4	3.17	2.21	32.2
FXR_96	-0.63	1.23	2.57	58.9
FXR_98	-2.17	2.83	1.68	13.1
FXR_99	2.35	4.36	2.89	100
FXR_100	-0.05	3.37	1.90	19.1
FXR_101	5.5	2.16	2.12	27.6
FXR_102	5.21	2.21	2.16	29.2





- Narrow range of experimental binding free energies
- **100% success rate** in predicting less active binders than the reference compound (true negatives)

FEP conclusions

- FEP+ is predictive given a good initial structure
- A specific protocol has to be followed:
  - ✔ ligand alignment to a reference ligand structure
  - ✔ investigation of buried waters
  - ✔ carefully selecting the mutations
  - ✔ investigation of double occupancy
- Cannot change the charge during a FEP mutation
- Correlations should not be expected for a narrow range (1-3 kcal/mol) of experimental binding free energies because the error of FEP is ~1 kcal/mol
- High success rate in classifying true positives and true negatives
- Reasonable throughput for lead optimization
  - ✔ 18 spiros, 21,000 atoms, 28 edges, 22 GPUs (Tesla K40m), 29 h
  - ✔ 15 sulfonamides, 17,000 atoms, 19 edges, 22 GPUs, 20 h

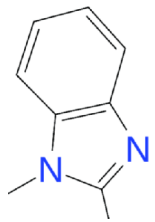
D3R Grand Challenge 2

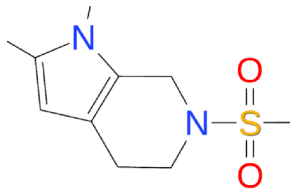
Physics-based pose predictions guided by native ligands

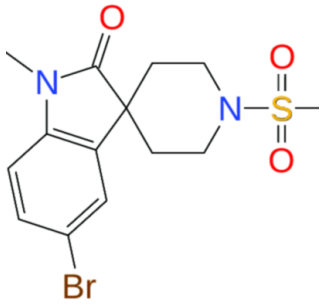
Christina Athanasiou, Sofia Vasilakaki, Dimitris Dellis,
Woody Sherman and Zoe Cournia

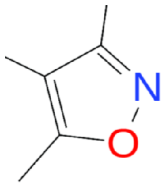
36 structures for pose prediction

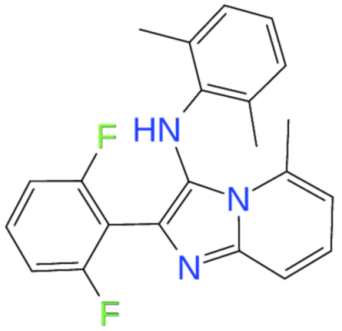
Overall rank: 5 out of 46 complete entries











	Benzimidazoles	Sulfonamides	Spiros	Isoxazoles	Miscellaneous
	6-9, 13-14, 19-22, 24-32, 35-36	15-17	10-12	4, 23, 33	1-3, 5, 18, 34
Mean RMSD:	**0.84 Å**	**2.95 Å**	**3.45 Å**	**4.94 Å**	**5.57 Å**
Mean Rank:	**8**	**6**	**9**	**43**	**24**

- ✔ Known chemotype in crystal structures
- ✔ Docking, alignment, minimization worked really well
- ✔ Cross docking predicted unknown binding mode
- ✔ Cross docking predicted unknown binding mode
- ✔ Diversity in binding modes did not allow for accurate prediction
- ✔ Cross docking did not work

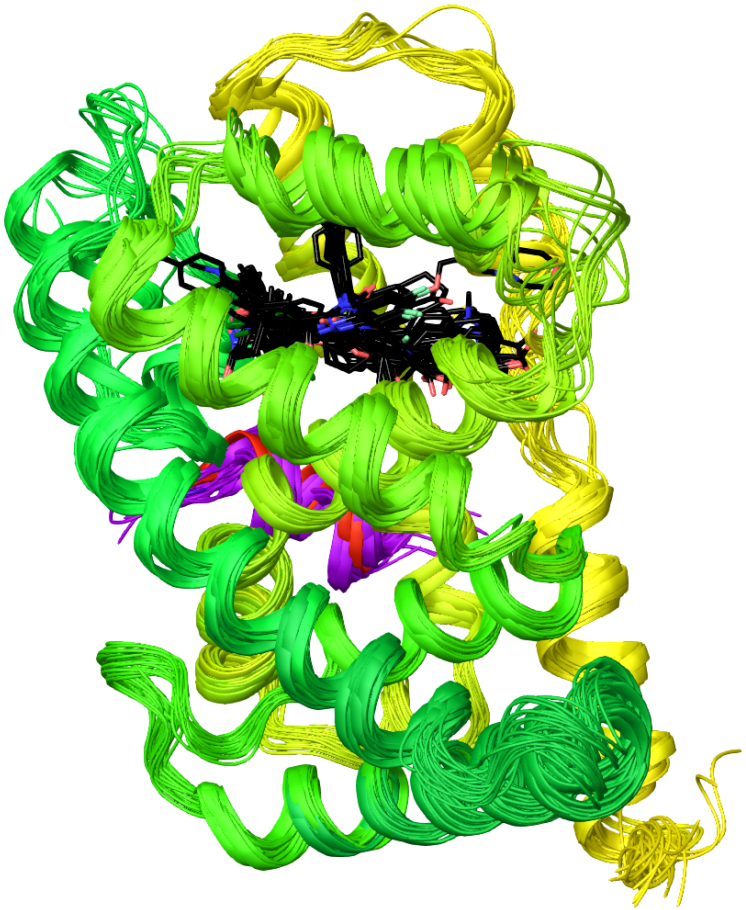
28 relevant PDB crystal structures available

The crystal structure were clustered based on the co-crystallized inhibitor structure:

- Isoxazole derivatives
- Steroid derivative
- Benzimidazole derivatives
- Indole derivatives
- Others
- APO crystal structure was provided by the D3R group

- **Wide binding pocket**

- **Not all ligands can be docked to the same crystal structure**

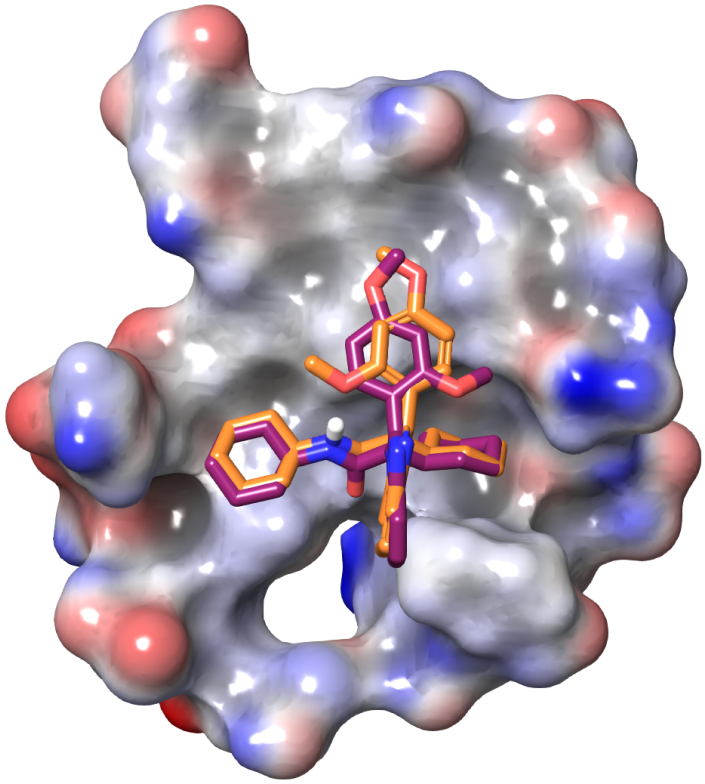


3RUT 3P89 3RUU 3P88 3RVF 3HC6 3HC5 3GD2 3DCT 3FXV 3DCU 4QE6 3BEJ 1OSV 1OT7 4QE8 4OIV 1OSH 3OLF 3OMK 3OMM 3OOF 3OOK 3OKH 3OKI 3L1B 3FLI

Methodology

1) Pose prediction for compounds with known chemotype in crystal structures

- Choice of crystal structure according to known chemotype
- Water molecules that were persistent in crystal structures were kept
- Ligand docking (Glide)
- Alignment to the native ligand (Maestro)
- Minimization of the complex (Maestro)



In case of double occupancy possibility:

- Water thermodynamics in binding pocket (WaterMap)
- Binding pose metadynamics (Desmond)
- FEP calculations (FEP+)

Benzimidazoles: FXR_6-9, FXR_13-14, FXR_20-22, FXR_24-32, FXR_35-36
Miscellaneous: FXR_5, FXR_34

Methodology

2) Pose prediction for compounds with unknown chemotype in crystal structures

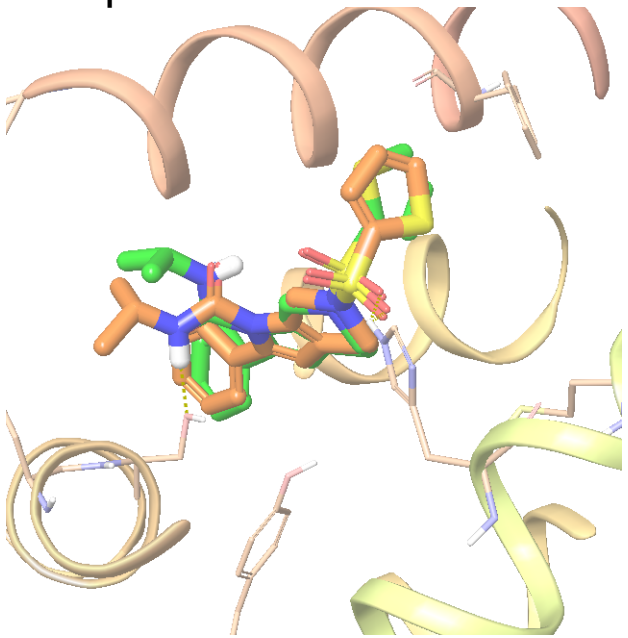
- Choice of crystal structure based on

  a) Shape similarity with native ligands (SHAPE)
  b) Cross docking in all 28 crystal structures (xglide.py)
  c) Interaction fingerprints (Maestro)

- Water molecules that were persistent in crystal structures were kept
- Docking (Glide)
- Alignment with the native ligand when a common core was present (Maestro)
- Minimization of the structure in case of alignment (Maestro)

In case of double occupancy possibility:

- Metadynamics calculations were used



Isoxazoles: FXR_4, FXR_23, FXR_33
Sulfonamides: FXR_15-17
Spiros: FXR_10-12
Miscellaneous: FXR_1-3, FXR_18

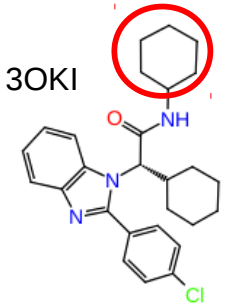
Pose prediction for compounds with known chemotypes

Benzimidazoles were categorized based on choice of crystal structure

- **3OKI**

  1) Saturated ring

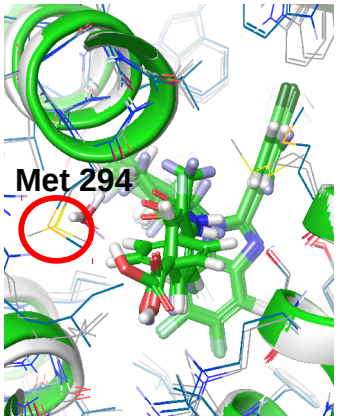
  FXR_6, FXR_7, FXR_8, FXR_9, FXR_13, FXR_19, FXR_20, FXR_22, FXR_26, FXR_30, FXR_31, FXR_32, FXR_35

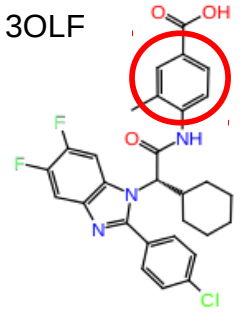


- **3OLF**

  1) Benzene ring, 2) Ortho substituted

  FXR_14, FXR_24, FXR_25, FXR_27, FXR_28

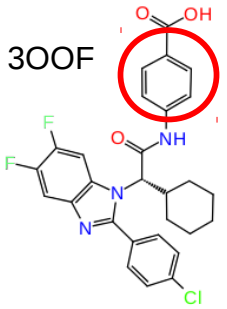


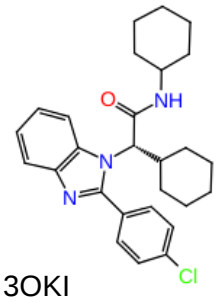


- **3OOF**

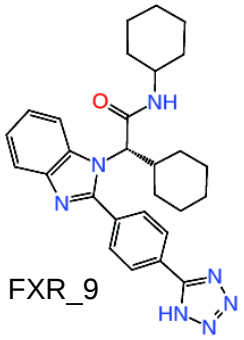
  1) Benzene ring, 2) Non ortho substituted

  FXR_21, FXR_29, FXR_36





3OKI

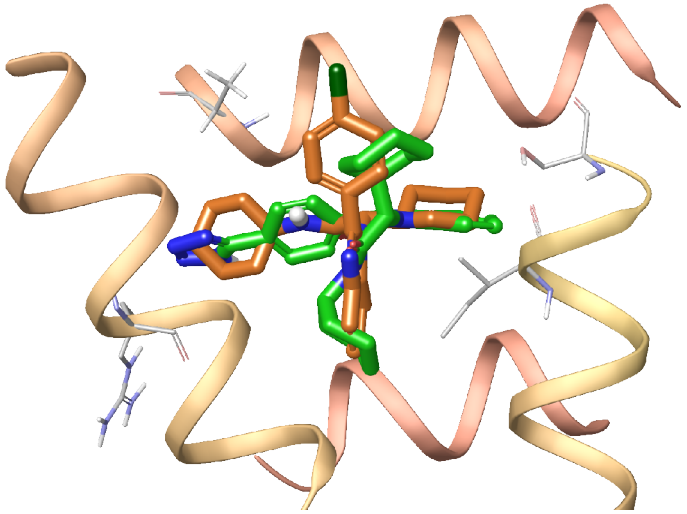


FXR_9

FXR_9

1) Docking in 3OKI

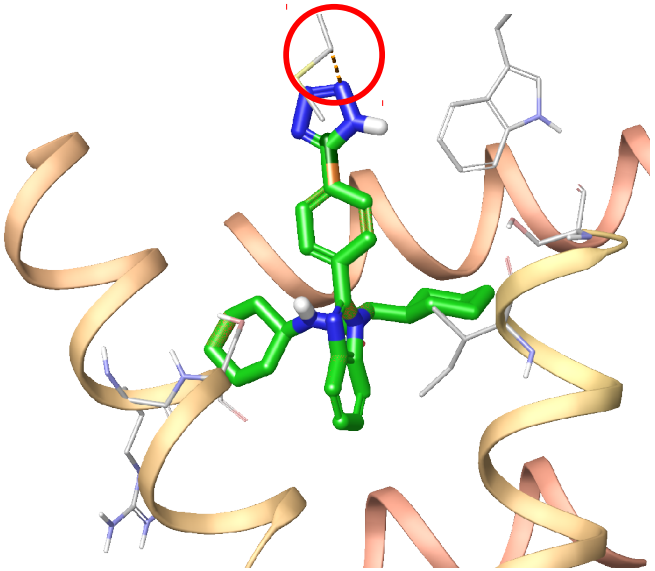
Not consistent binding mode



Orange: native ligand
Green: docked ligand

2) Alignment to 3OKI ligand

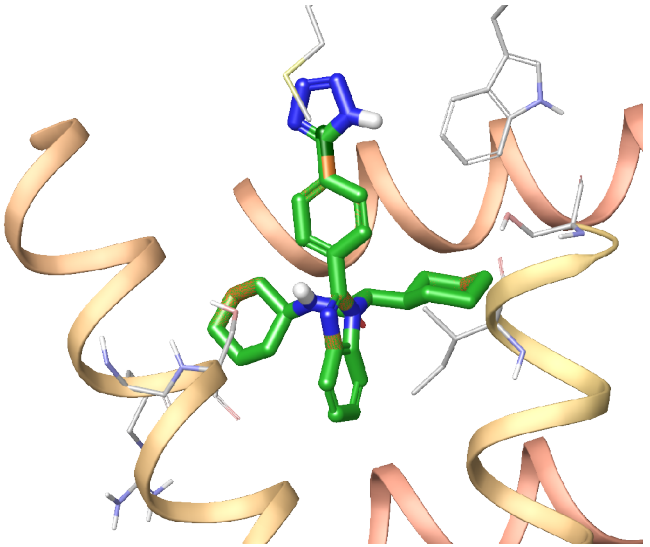
Bad contact



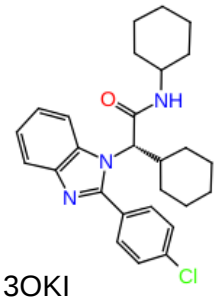
Orange: native ligand
Green: aligned ligand to native

3) Minimization of the complex: aligned ligand - protein

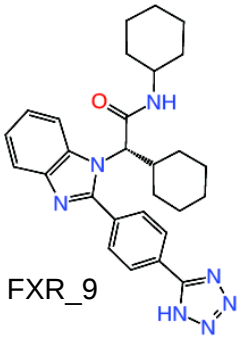
Alleviation of clashes



Orange: native ligand
Green: aligned ligand, complex minimization

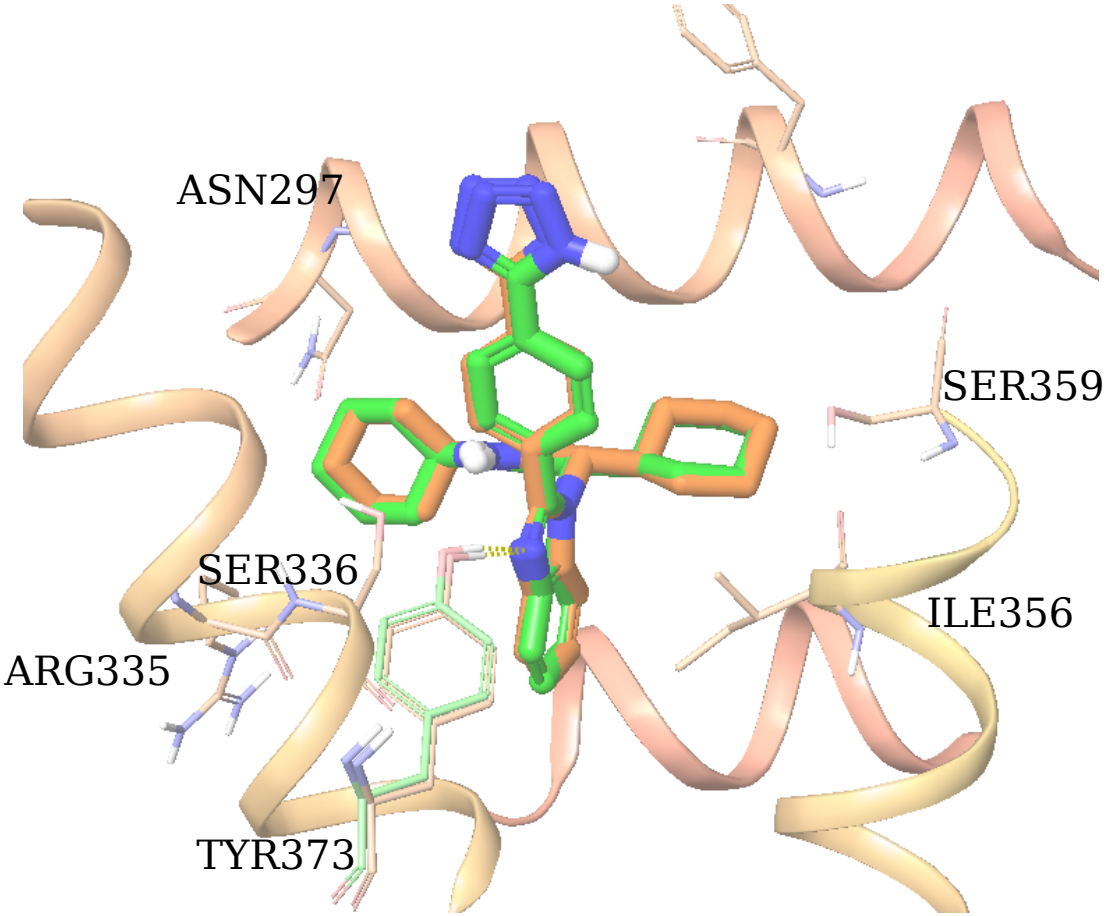


3OKI



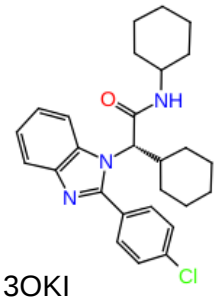
FXR_9

FXR_9 (1ytut)

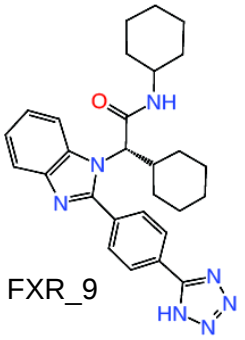


Orange: crystal structure
Green: predicted pose

RMSD = 0.316 Å



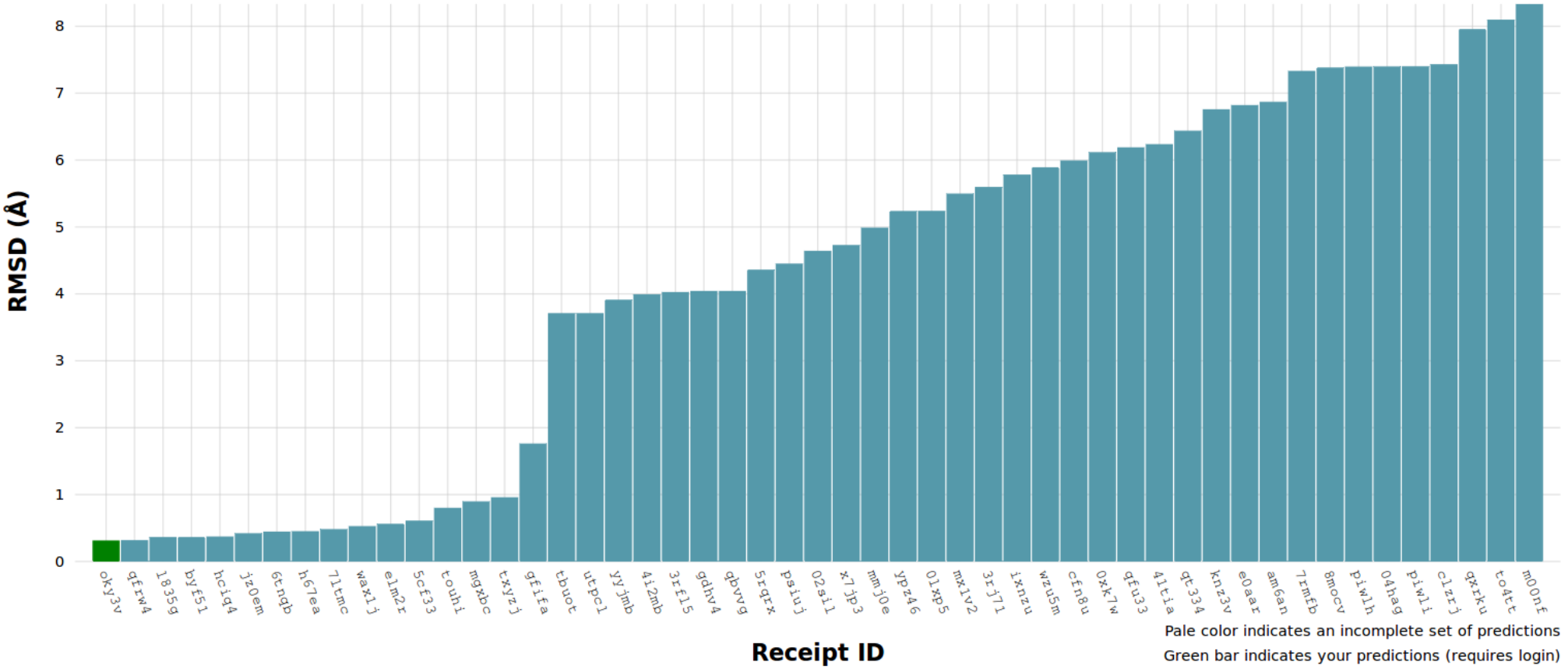
3OKI



FXR_9

FXR_9 (1ytut)

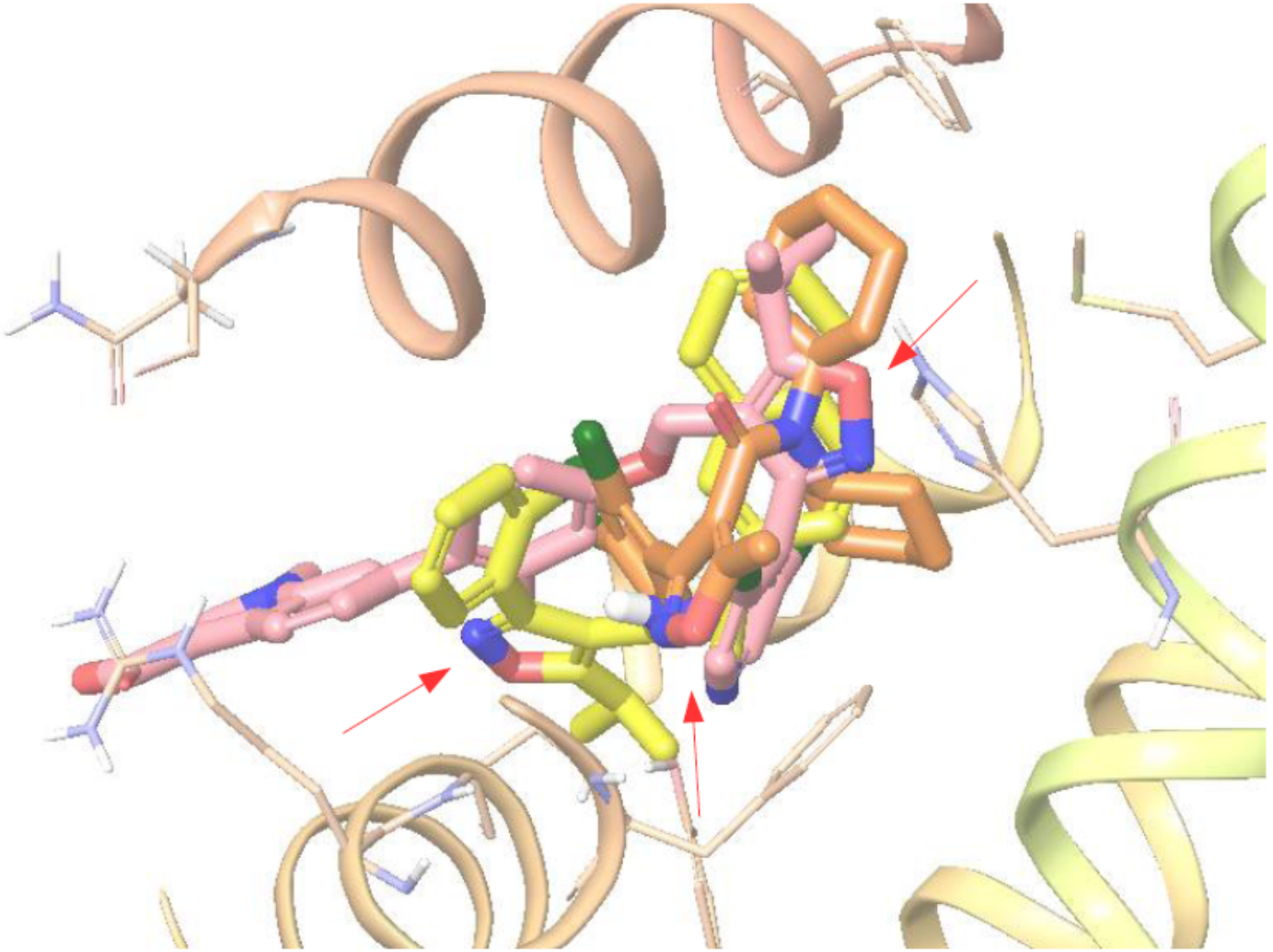
RMSD = 0.316 Å



Pose prediction for compounds with unknown chemotypes

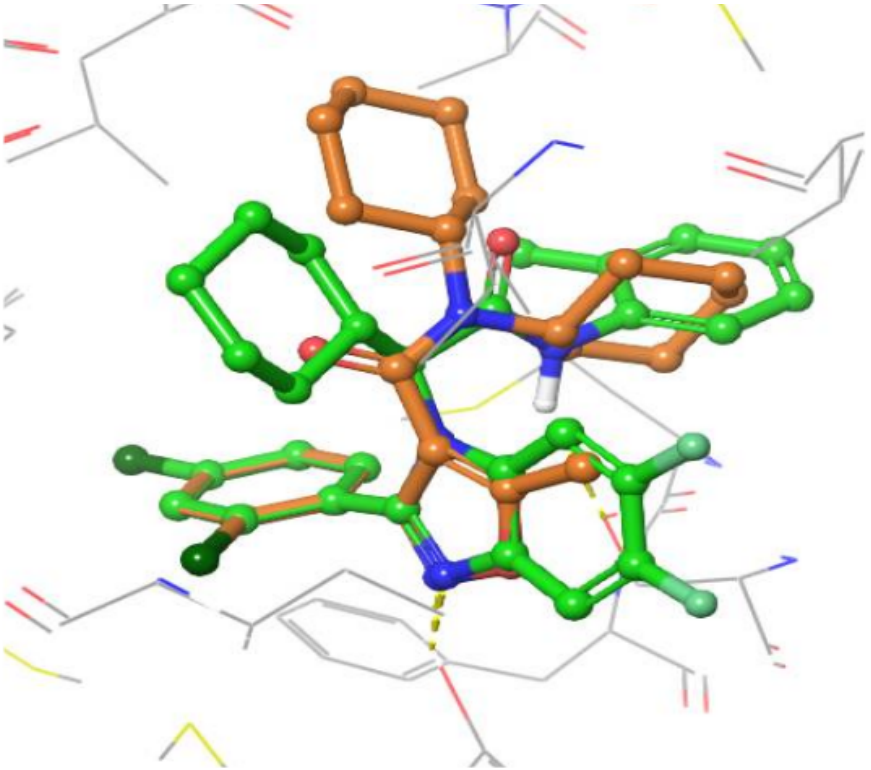
Isoxazoles

Alignment of isoxazoles

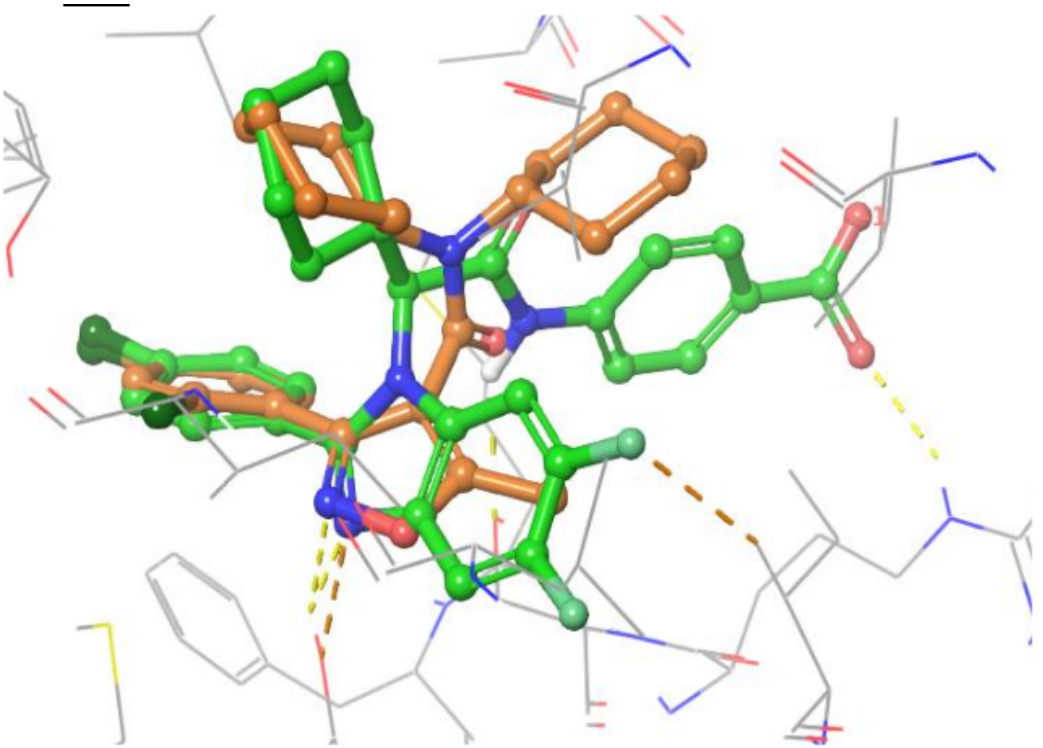


The isoxazole ring is located at the active site quite differently in the released crystal structures (*orange*: FXR_4, *yellow*: FXR_23, *pink*: FXR_33).

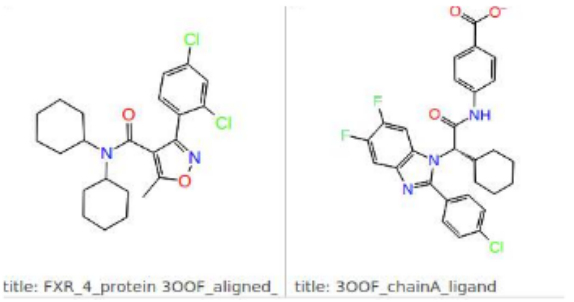
FXR_4



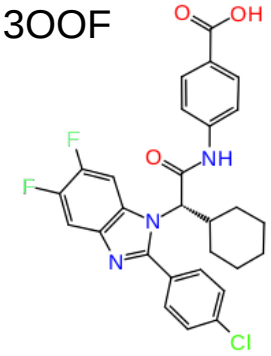
FXR_4 (orange) aligned with 3OMK native ligand (green) and Prime minimized. Fingerprints Similarity 0.46



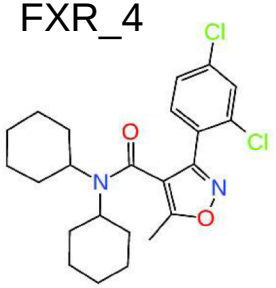
Binding pose metadynamics of FXR_4 (orange) aligned with 3OOF native ligand (green) and Prime minimized. Fingerprints Similarity 0.58



3OOF

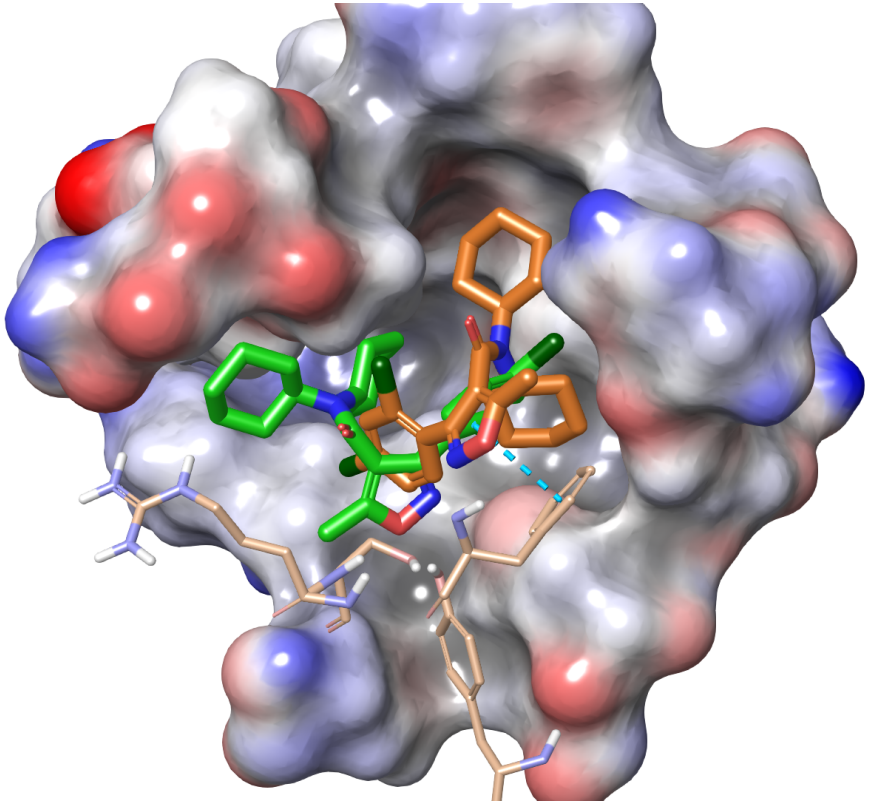


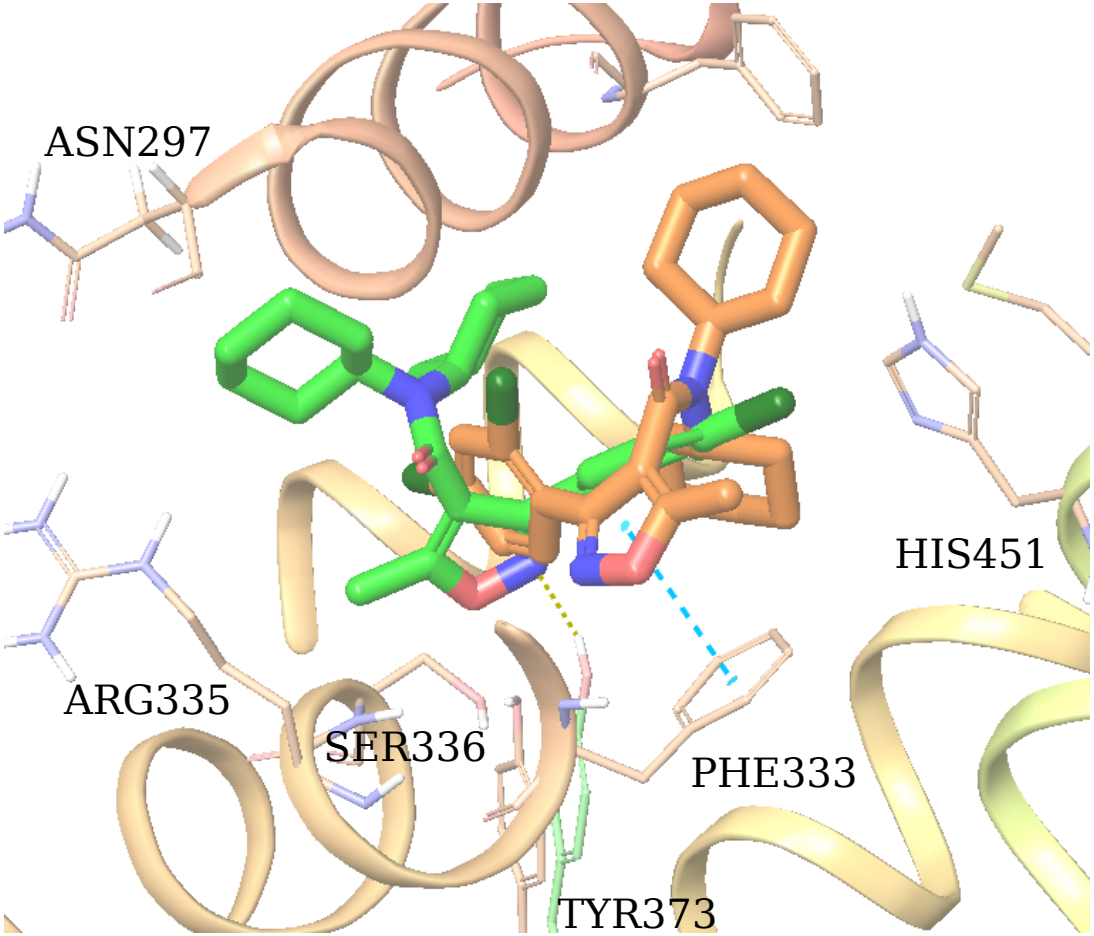
FXR_4



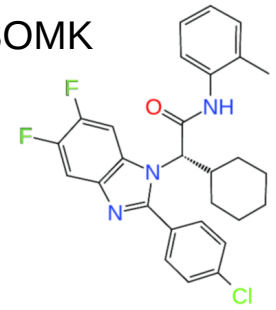
FXR_4 (1pdbc)

RMSD = 6.77 Å

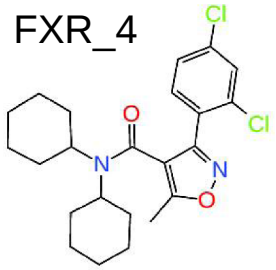




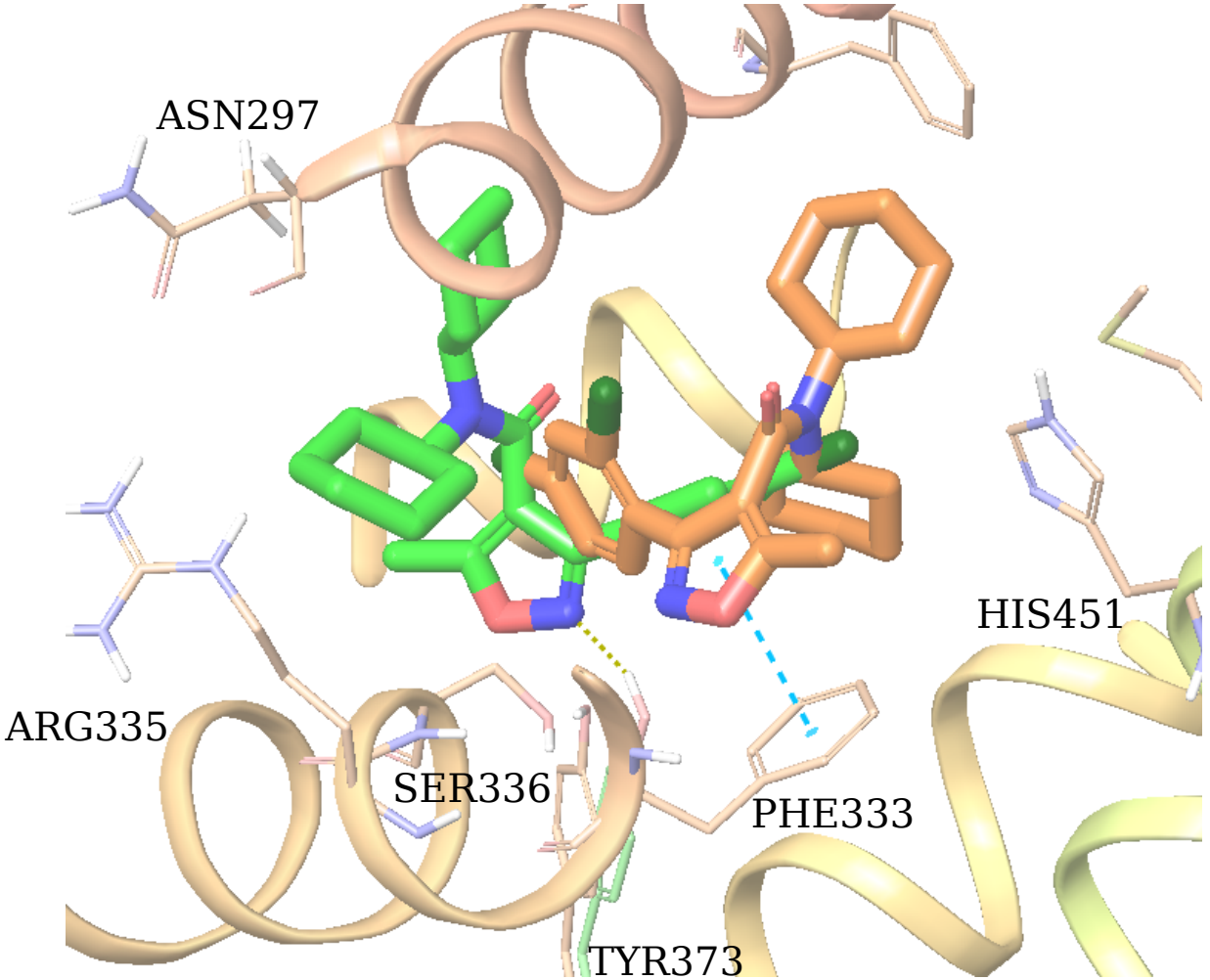
3OMK



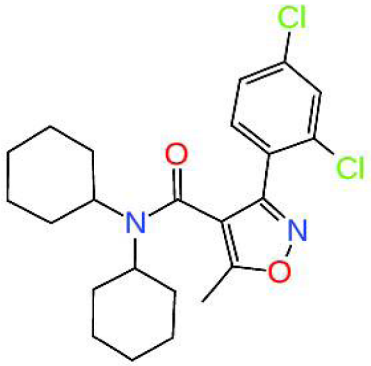
FXR_4



FXR_4 (1pdbc)

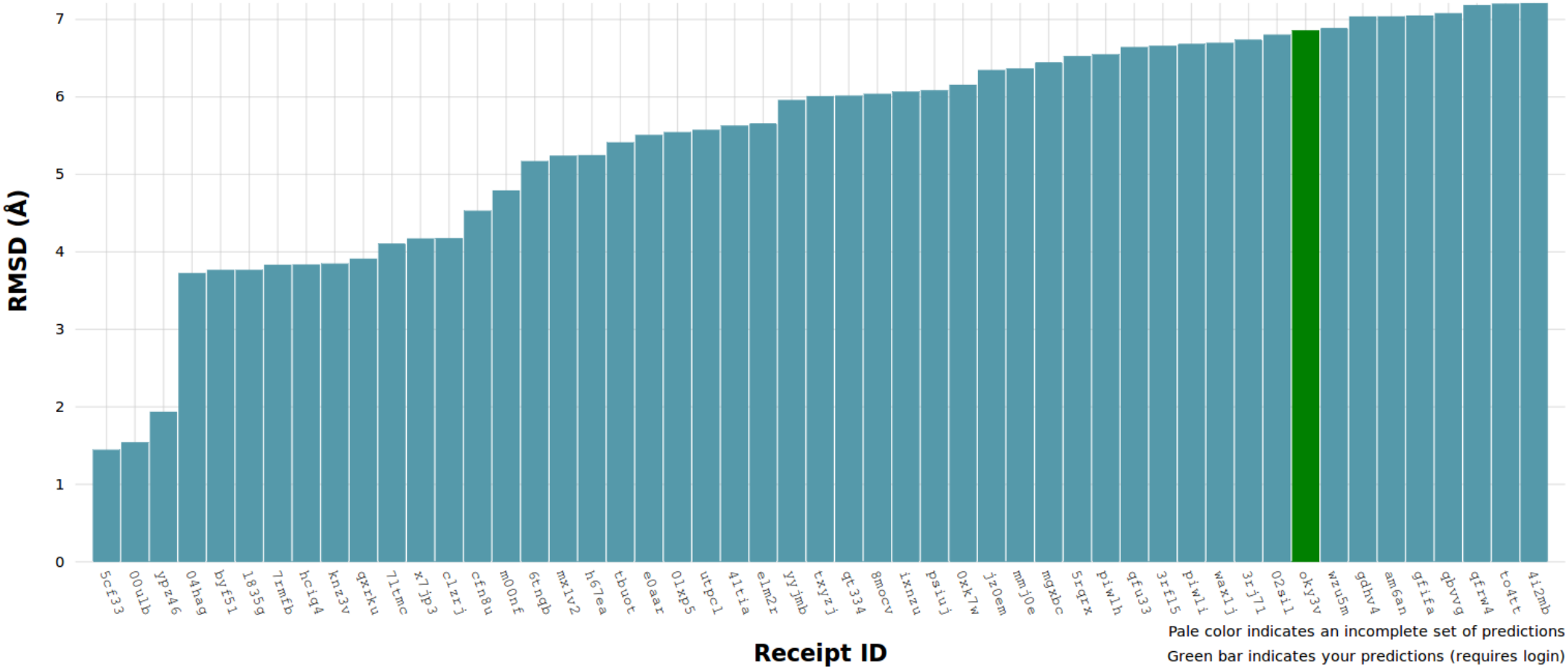


RMSD = 6.96 Å



FXR_4 (1pdbc)

RMSD = 6.96 Å

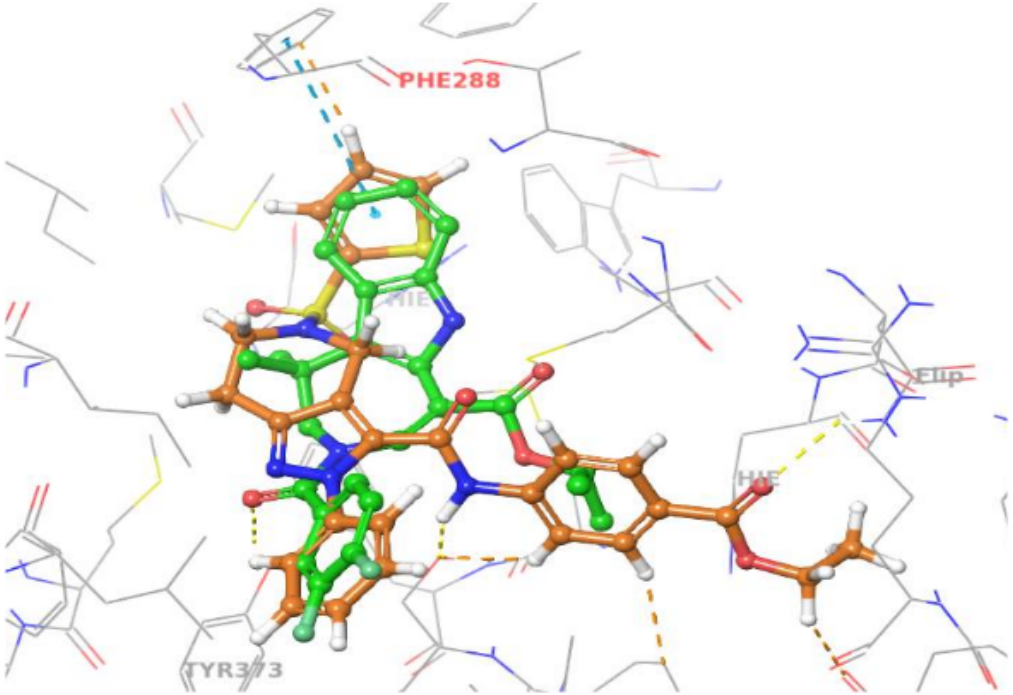


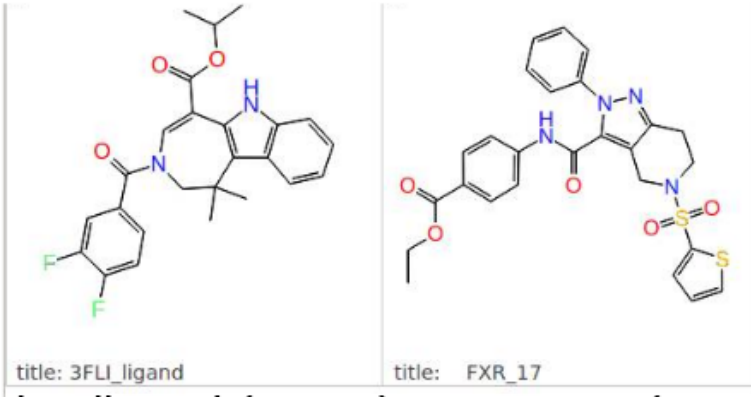
Sulfonamides

FXR_17

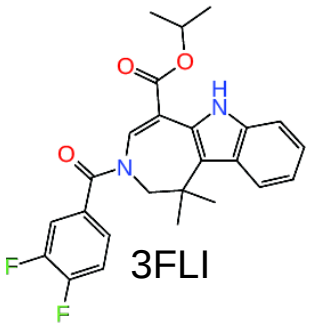
XGlide was used and FXR_17 docking pose in 3FLI had the best overlap and the best interaction fingerprints with the native ligand. The docking pose was optimized in Jaguar and re-docked (SP and XP).

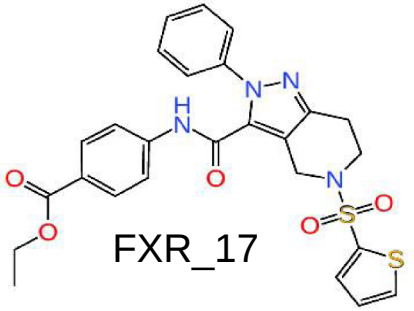
Interactions similarity: 0.614



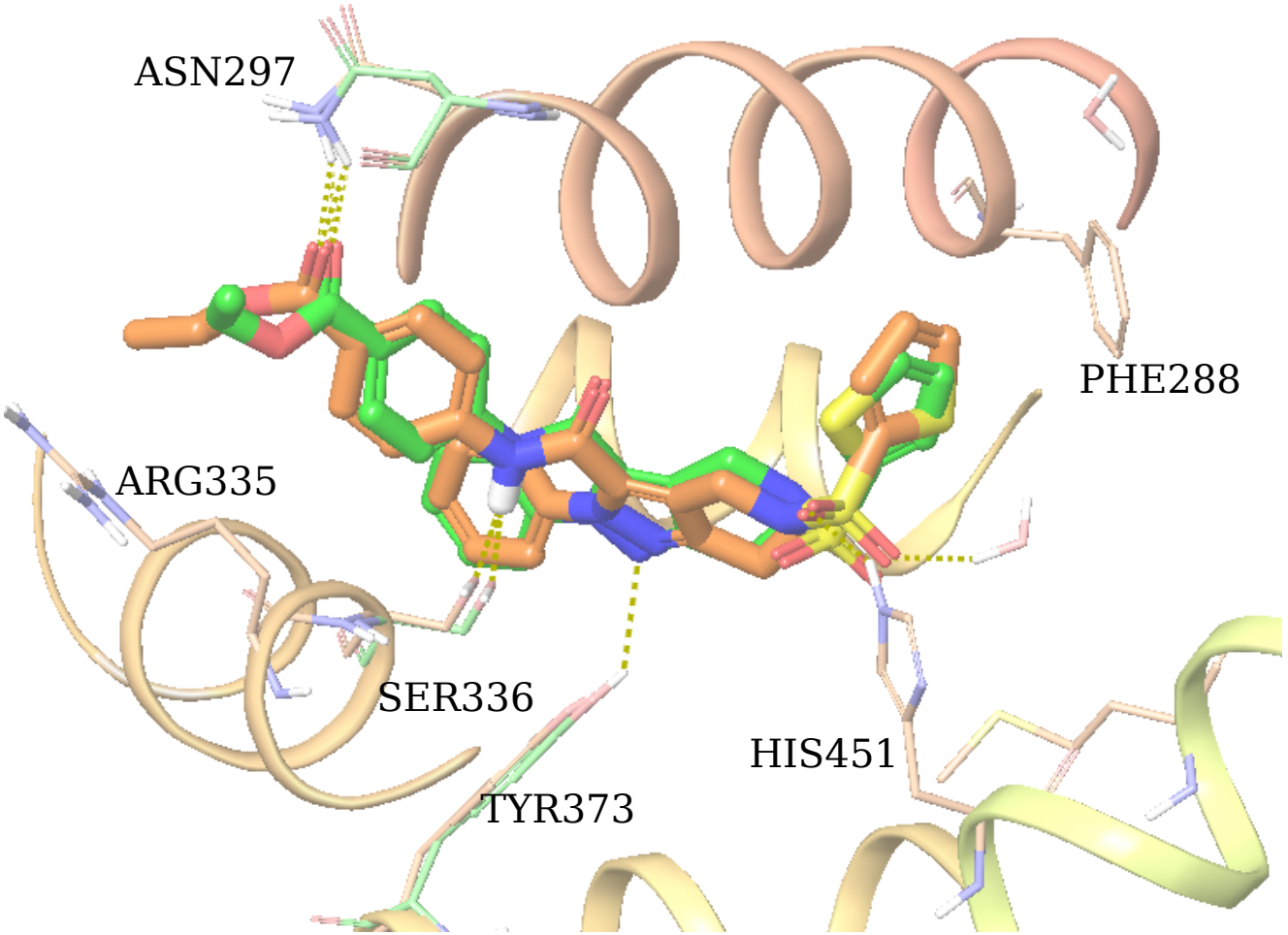


native ligand (green) FXR_17 (orange)



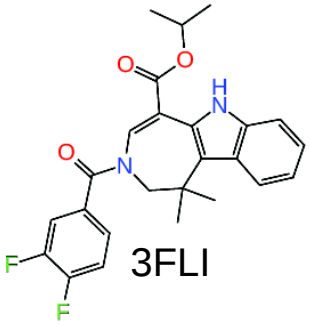


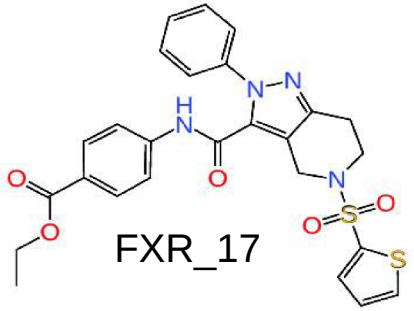
FXR_17



Orange: crystal structure
Green: predicted pose

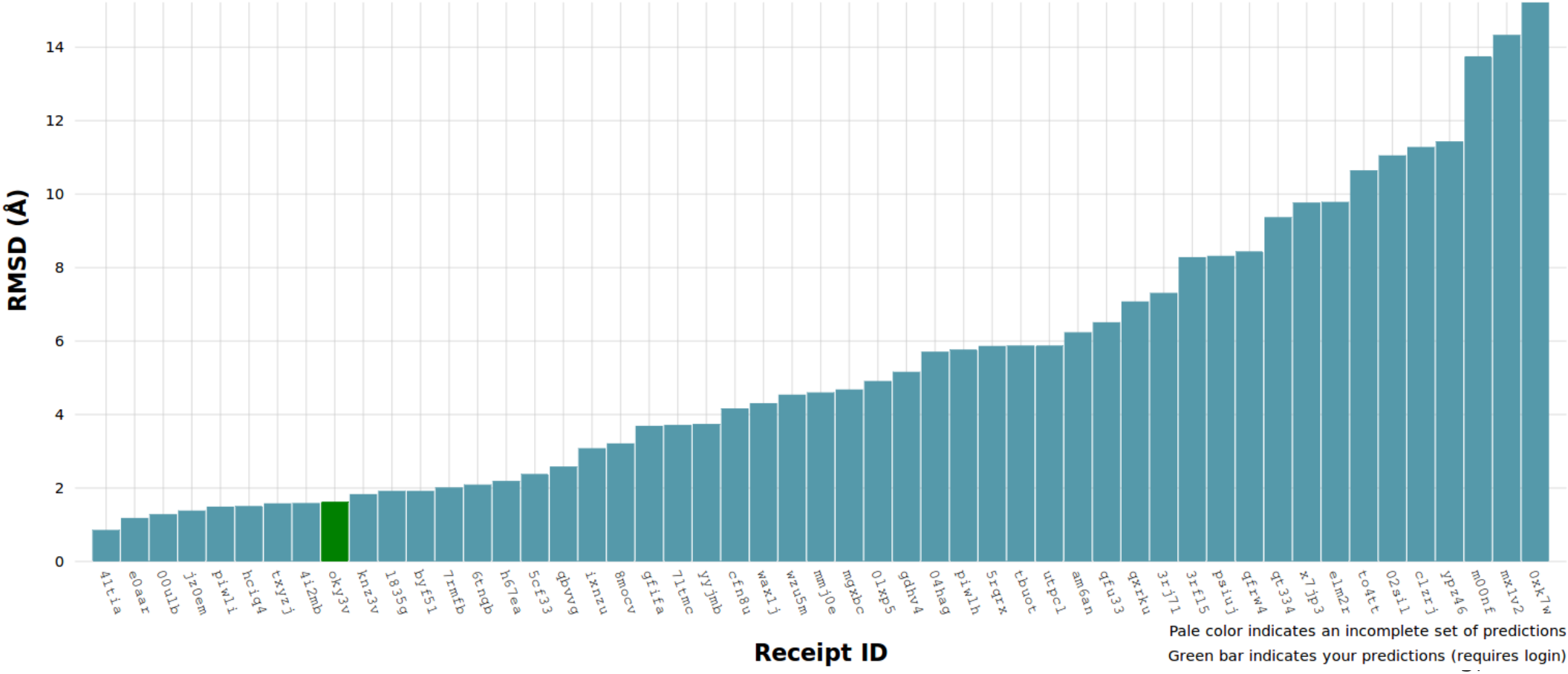
RMSD = 1.63 Å





FXR_17

RMSD = 1.63 Å



Spiros

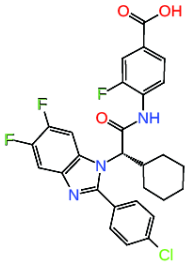
Spiros

- Spiros compounds were initially docked in **3OMM**, which was indicated by SHAPE analysis.

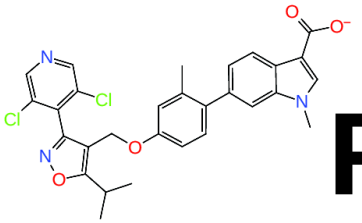
- Subsequently, they were docked in all 28 crystal structures (cross docking).
  The pose of docking in **3FXV** crystal structure was the best, with a Glide Score of ~ -11 kcal/mol.

Shape Results

	10	11	12
3OMM	0.538	0.458	0.526
3OKH	0.486	0.46	0.43
3P89	0.366	0.366	0.388
3RVF	0.39	0.39	0.38
1OSH	0.490	0.524	0.450
4OIV	0.550	0.466	0.507
4QE8	0.550	0.466	0.507
3L1B	0.40	0.40	0.40
1OT7	0.56	0.64	0.56

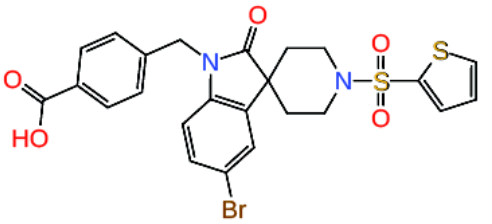


3OMM



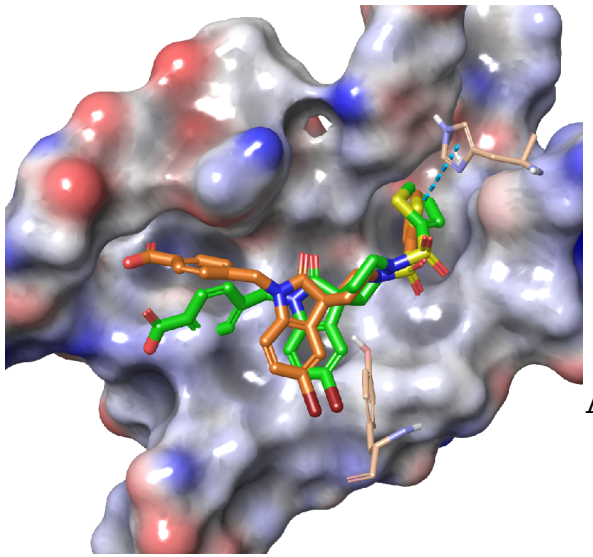
3FXV

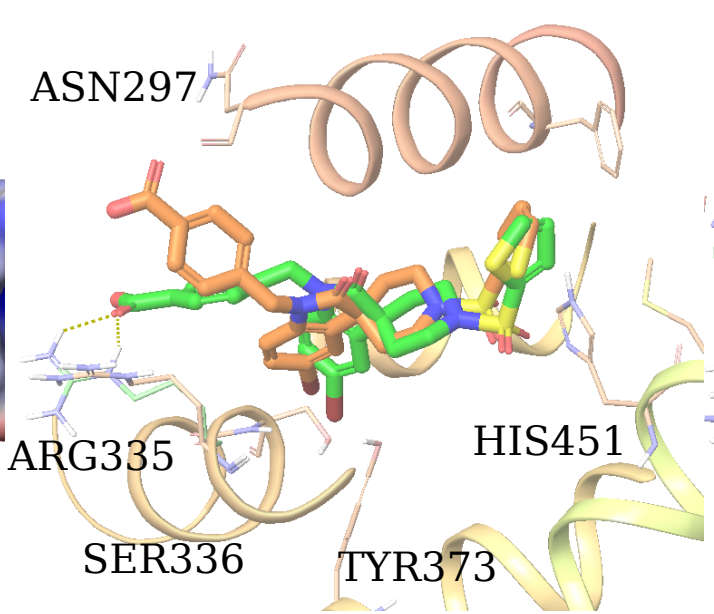
FXR_10 (1sjpr)

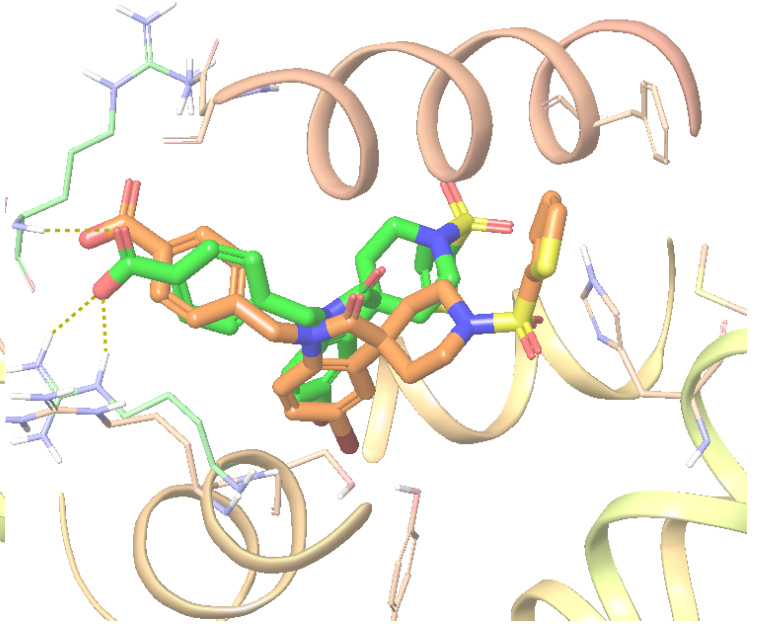


FXR_10

Orange: crystal structure
Green: predicted pose







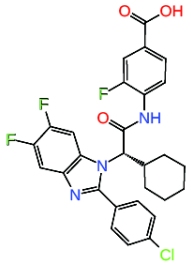
Docking in 3FXV – 1st pose

RMSD = 2.14 Å

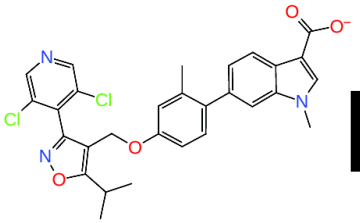
Docking in 3OMM – 2nd pose

RMSD = 4.43 Å

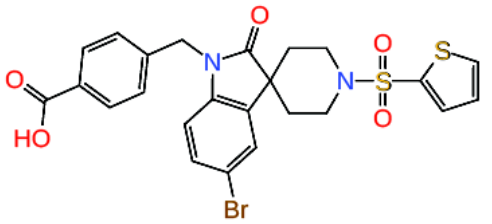
FXR_10 (1sjpr)



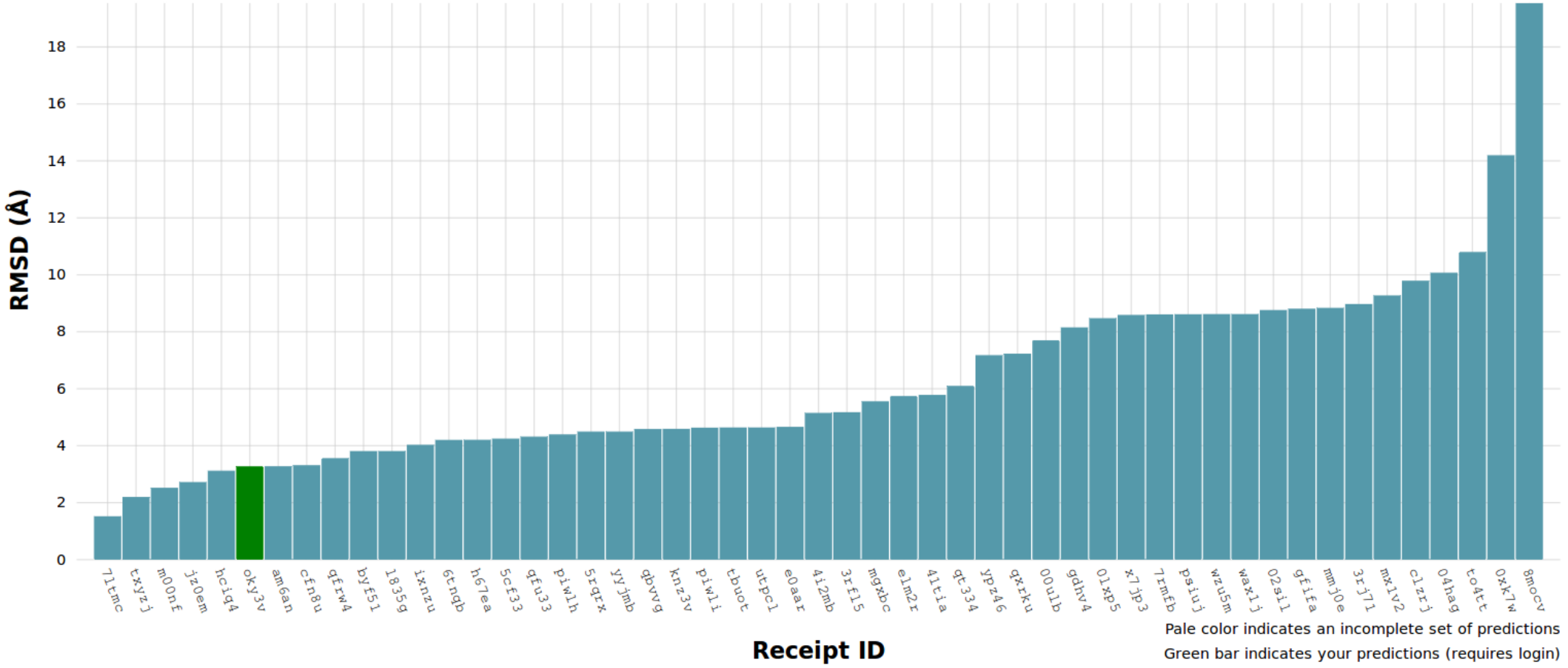
3OMM



3FXV



FXR_10



Conclusions

- Pose predictions were accurate, when a crystal structure with common chemotype native ligand was available.
  In this case, docking, alignment to native ligand and minimization performed well.

- Methodology needs improvements, in case a crystal structure with common chemotype native ligand is not available.
  In this case, cross docking and interaction fingerprints performed well for some compounds.

- Difficulty in predicting isoxazoles poses due to diversity in binding modes

Acknowledgements

Dr. Stephan Ehrlich
Dr. Thomas Steinbrecher

SCHRÖDINGER.

D3R organizers

Cournia lab members,
Biomedical Research Foundation,
Academy of Athens

http://www.drugdesign.gr/