Supporting Information

Using physics-based pose predictions and Free Energy Perturbation calculations to predict binding poses and relative binding affinities for FXR ligands in the D3R Grand Challenge 2

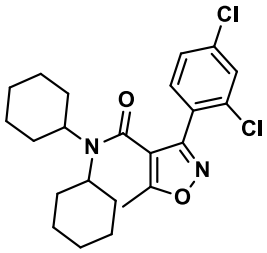
Christina Athanasiou,[1] Sofia Vasilakaki,[1] Dimitris Dellis,[2] Zoe Cournia[1]

1. Biomedical Research Foundation, Academy of Athens, 4 Soranou Ephessiou, 115 27 Athens, Greece

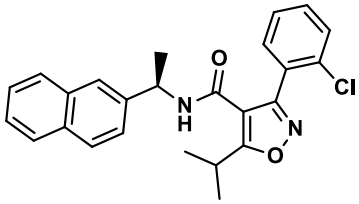
2. Greek Research and Technology Network, S.A., 7 Kifissias Ave, 115 23 Athens, Greece

Table S1. Structures of 36 compounds

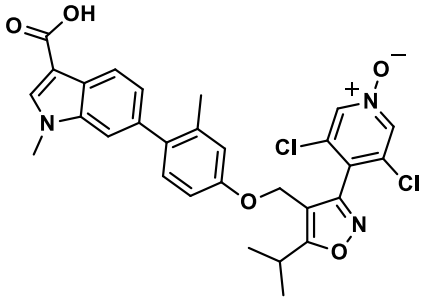
Isoxazoles



FXR_4

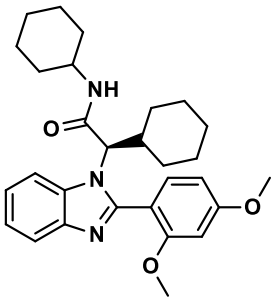


FXR_23

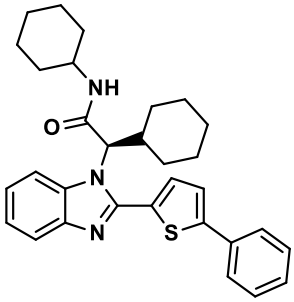


FXR_33

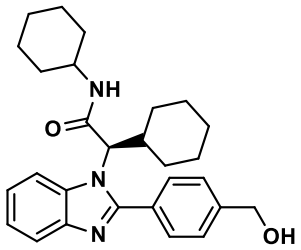
Benzimidazoles



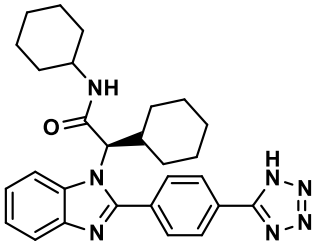
FXR_6



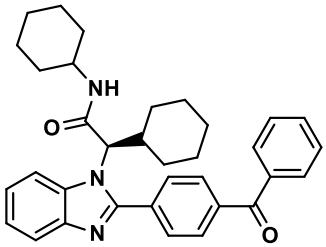
FXR_7



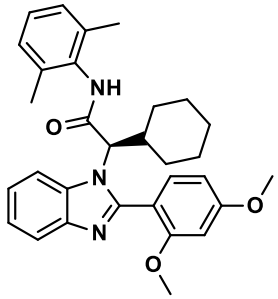
FXR_8



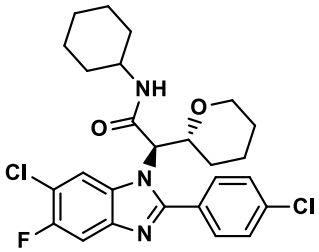
FXR_9



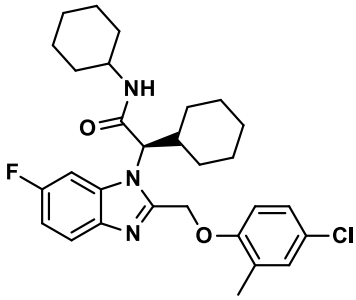
FXR_13



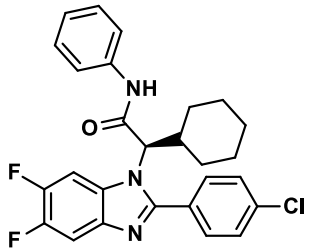
FXR_14



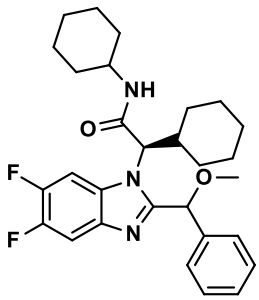
FXR_19



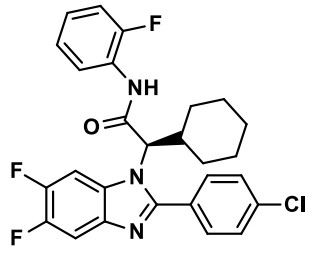
FXR_20



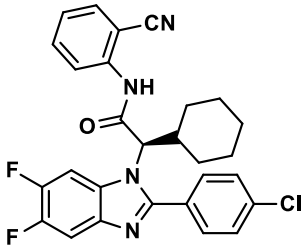
FXR_21



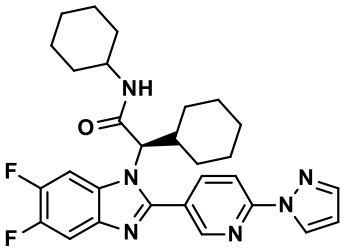
FXR_22



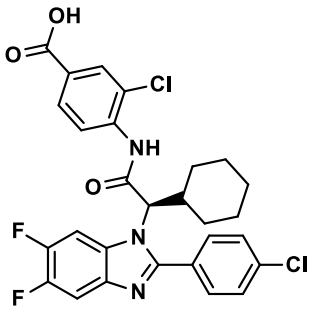
FXR_24



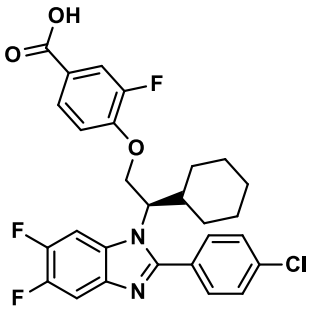
FXR_25



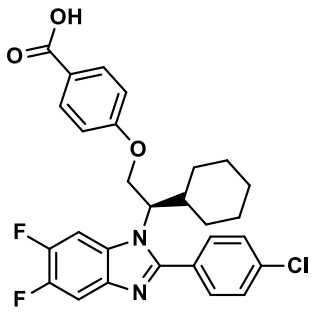
FXR_26



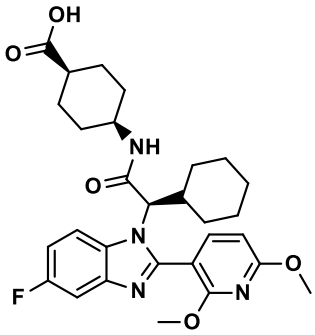
FXR_27



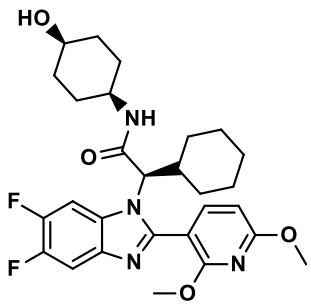
FXR_28



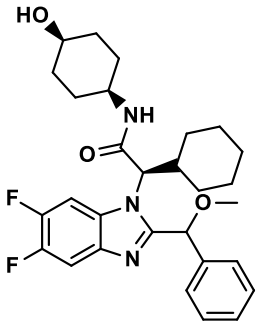
FXR_29



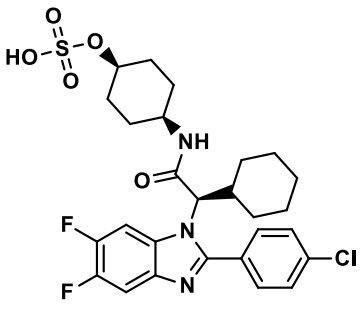
FXR_30



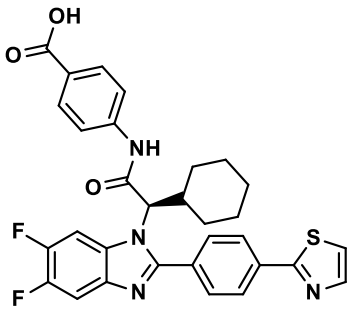
FXR_31



FXR_32

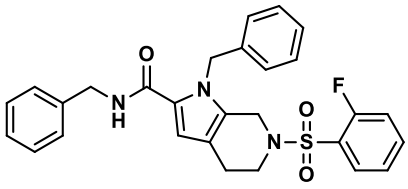


FXR_35

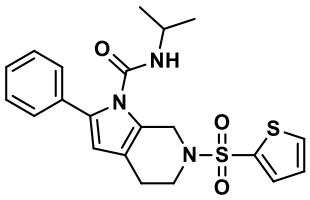


FXR_36

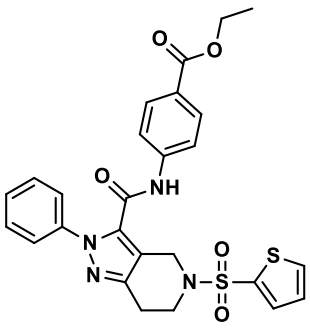
Sulfonamides



FXR_15

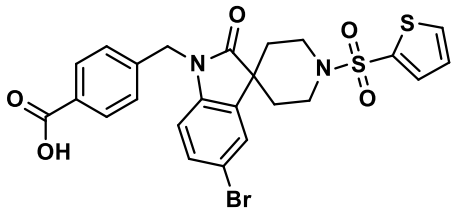


FXR_16

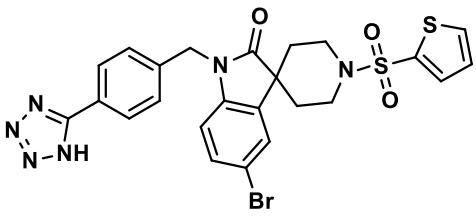


FXR_17

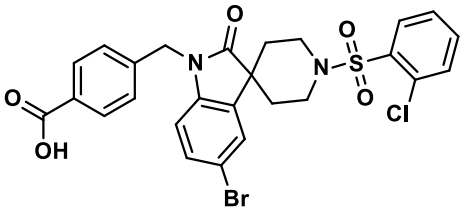
Spiros



FXR_10

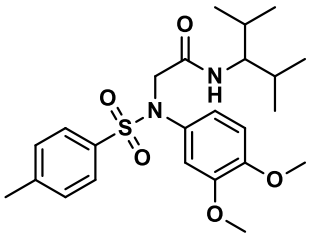


FXR_11

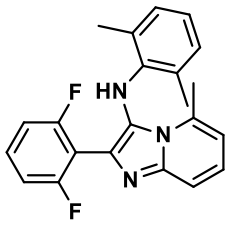


FXR_12

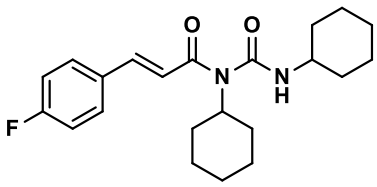
Miscellaneous



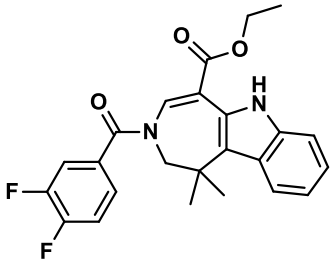
FXR_1



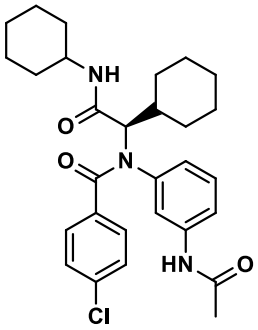
FXR_2



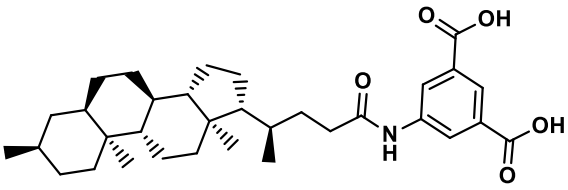
FXR_3



FXR_5



FXR_18



FXR_34

Table S2. Tanimoto similarity of the 36 compounds.

	FXR_1	FXR_2	FXR_3	FXR_4	FXR_5	FXR_6	FXR_7	FXR_8	FXR_9
FXR_1	1.00	0.02	0.05	0.05	0.03	0.06	0.02	0.03	0.03
FXR_2	0.02	1.00	0.02	0.03	0.02	0.06	0.04	0.05	0.05
FXR_3	0.05	0.02	1.00	0.09	0.02	0.05	0.05	0.07	0.06
FXR_4	0.05	0.03	0.09	1.00	0.04	0.06	0.04	0.05	0.04
FXR_5	0.03	0.02	0.02	0.04	1.00	0.03	0.02	0.02	0.02
FXR_6	0.06	0.06	0.05	0.06	0.03	1.00	0.48	0.60	0.55
FXR_7	0.02	0.04	0.05	0.04	0.02	0.48	1.00	0.58	0.55
FXR_8	0.03	0.05	0.07	0.05	0.02	0.60	0.58	1.00	0.81
FXR_9	0.03	0.05	0.06	0.04	0.02	0.55	0.55	0.81	1.00
FXR_10	0.07	0.01	0.02	0.04	0.04	0.02	0.02	0.03	0.02
FXR_11	0.06	0.01	0.02	0.04	0.03	0.03	0.02	0.03	0.07
FXR_12	0.07	0.01	0.02	0.04	0.04	0.03	0.02	0.04	0.03
FXR_13	0.03	0.05	0.06	0.04	0.02	0.57	0.58	0.84	0.78
FXR_14	0.06	0.09	0.03	0.06	0.03	0.76	0.35	0.44	0.41
FXR_15	0.05	0.02	0.02	0.02	0.03	0.03	0.03	0.04	0.03
FXR_16	0.05	0.02	0.02	0.02	0.02	0.03	0.05	0.04	0.04
FXR_17	0.05	0.02	0.02	0.03	0.03	0.02	0.03	0.03	0.03
FXR_18	0.04	0.02	0.10	0.06	0.03	0.10	0.09	0.12	0.11
FXR_19	0.03	0.03	0.06	0.05	0.02	0.27	0.26	0.36	0.33
FXR_20	0.05	0.04	0.06	0.06	0.03	0.26	0.23	0.28	0.26
FXR_21	0.03	0.04	0.04	0.05	0.02	0.24	0.24	0.33	0.30
FXR_22	0.03	0.04	0.06	0.03	0.02	0.22	0.22	0.25	0.24
FXR_23	0.05	0.03	0.03	0.18	0.04	0.05	0.03	0.04	0.04
FXR_24	0.03	0.05	0.04	0.05	0.03	0.24	0.22	0.32	0.29
FXR_25	0.03	0.05	0.04	0.05	0.03	0.23	0.22	0.31	0.29
FXR_26	0.02	0.03	0.05	0.03	0.02	0.26	0.26	0.30	0.28
FXR_27	0.03	0.04	0.04	0.07	0.04	0.23	0.21	0.29	0.27
FXR_28	0.04	0.03	0.04	0.06	0.05	0.18	0.15	0.22	0.20
FXR_29	0.03	0.04	0.04	0.05	0.02	0.18	0.16	0.24	0.22
FXR_30	0.05	0.04	0.04	0.05	0.03	0.36	0.25	0.30	0.28
FXR_31	0.03	0.04	0.04	0.04	0.02	0.32	0.23	0.27	0.25
FXR_32	0.02	0.03	0.04	0.03	0.02	0.18	0.18	0.21	0.20
FXR_33	0.04	0.03	0.01	0.10	0.03	0.04	0.02	0.03	0.02
FXR_34	0.02	0.01	0.02	0.02	0.01	0.01	0.01	0.02	0.01
FXR_35	0.03	0.04	0.05	0.05	0.02	0.27	0.26	0.37	0.34
FXR_36	0.02	0.03	0.04	0.04	0.02	0.22	0.22	0.29	0.28

	FXR_10	FXR_11	FXR_12	FXR_13	FXR_14	FXR_15	FXR_16	FXR_17	FXR_18
FXR_1	0.07	0.06	0.07	0.03	0.06	0.05	0.05	0.05	0.04
FXR_2	0.01	0.01	0.01	0.05	0.09	0.02	0.02	0.02	0.02
FXR_3	0.02	0.02	0.02	0.06	0.03	0.02	0.02	0.02	0.10
FXR_4	0.04	0.04	0.04	0.04	0.06	0.02	0.02	0.03	0.06
FXR_5	0.04	0.03	0.04	0.02	0.03	0.03	0.02	0.03	0.03

FXR_6	0.02	0.03	0.03	0.57	0.76	0.03	0.03	0.02	0.10
FXR_7	0.02	0.02	0.02	0.58	0.35	0.03	0.05	0.03	0.09
FXR_8	0.03	0.03	0.04	0.84	0.44	0.04	0.04	0.03	0.12
FXR_9	0.02	0.07	0.03	0.78	0.41	0.03	0.04	0.03	0.11
FXR_10	1.00	0.85	0.78	0.03	0.02	0.04	0.08	0.08	0.04
FXR_11	0.85	1.00	0.67	0.03	0.03	0.04	0.07	0.07	0.03
FXR_12	0.78	0.67	1.00	0.03	0.03	0.08	0.04	0.05	0.04
FXR_13	0.03	0.03	0.03	1.00	0.42	0.04	0.04	0.04	0.12
FXR_14	0.02	0.03	0.03	0.42	1.00	0.03	0.03	0.02	0.06
FXR_15	0.04	0.04	0.08	0.04	0.03	1.00	0.26	0.09	0.02
FXR_16	0.08	0.07	0.04	0.04	0.03	0.26	1.00	0.16	0.02
FXR_17	0.08	0.07	0.05	0.04	0.02	0.09	0.16	1.00	0.04
FXR_18	0.04	0.03	0.04	0.12	0.06	0.02	0.02	0.04	1.00
FXR_19	0.02	0.02	0.02	0.34	0.19	0.03	0.03	0.02	0.10
FXR_20	0.02	0.02	0.02	0.27	0.18	0.03	0.03	0.02	0.10
FXR_21	0.02	0.03	0.03	0.32	0.25	0.04	0.04	0.04	0.09
FXR_22	0.01	0.01	0.01	0.25	0.15	0.03	0.03	0.02	0.10
FXR_23	0.03	0.03	0.05	0.04	0.04	0.03	0.03	0.03	0.04
FXR_24	0.02	0.03	0.03	0.30	0.28	0.04	0.04	0.03	0.08
FXR_25	0.02	0.03	0.03	0.30	0.28	0.04	0.04	0.03	0.08
FXR_26	0.01	0.02	0.01	0.29	0.19	0.03	0.03	0.02	0.08
FXR_27	0.03	0.03	0.04	0.28	0.26	0.03	0.03	0.04	0.08
FXR_28	0.04	0.03	0.04	0.21	0.17	0.03	0.03	0.03	0.05
FXR_29	0.03	0.03	0.04	0.23	0.17	0.03	0.03	0.03	0.06
FXR_30	0.02	0.02	0.02	0.29	0.31	0.02	0.03	0.02	0.07
FXR_31	0.01	0.02	0.01	0.26	0.27	0.03	0.03	0.02	0.07
FXR_32	0.01	0.01	0.02	0.21	0.15	0.03	0.03	0.02	0.07
FXR_33	0.03	0.03	0.03	0.02	0.05	0.02	0.03	0.03	0.02
FXR_34	0.01	0.01	0.01	0.01	0.02	0.01	0.01	0.01	0.03
FXR_35	0.02	0.03	0.03	0.35	0.22	0.03	0.04	0.03	0.09
FXR_36	0.03	0.03	0.03	0.29	0.23	0.03	0.04	0.05	0.07

	FXR_19	**FXR_20**	**FXR_21**	**FXR_22**	**FXR_23**	**FXR_24**	**FXR_25**	**FXR_26**	**FXR_27**
FXR_1	0.03	0.05	0.03	0.03	0.05	0.03	0.03	0.02	0.03
FXR_2	0.03	0.04	0.04	0.04	0.03	0.05	0.05	0.03	0.04
FXR_3	0.06	0.06	0.04	0.06	0.03	0.04	0.04	0.05	0.04
FXR_4	0.05	0.06	0.05	0.03	0.18	0.05	0.05	0.03	0.07
FXR_5	0.02	0.03	0.02	0.02	0.04	0.03	0.03	0.02	0.04
FXR_6	0.27	0.26	0.24	0.22	0.05	0.24	0.23	0.26	0.23
FXR_7	0.26	0.23	0.24	0.22	0.03	0.22	0.22	0.26	0.21
FXR_8	0.36	0.28	0.33	0.25	0.04	0.32	0.31	0.30	0.29
FXR_9	0.33	0.26	0.30	0.24	0.04	0.29	0.29	0.28	0.27
FXR_10	0.02	0.02	0.02	0.01	0.03	0.02	0.02	0.01	0.03
FXR_11	0.02	0.02	0.03	0.01	0.03	0.03	0.03	0.02	0.03
FXR_12	0.02	0.02	0.03	0.01	0.05	0.03	0.03	0.01	0.04
FXR_13	0.34	0.27	0.32	0.25	0.04	0.30	0.30	0.29	0.28
FXR_14	0.19	0.18	0.25	0.15	0.04	0.28	0.28	0.19	0.26

FXR_15	0.03	0.03	0.04	0.03	0.03	0.04	0.04	0.03	0.03
FXR_16	0.03	0.03	0.04	0.03	0.03	0.04	0.04	0.03	0.03
FXR_17	0.02	0.02	0.04	0.02	0.03	0.03	0.03	0.02	0.04
FXR_18	0.10	0.10	0.09	0.10	0.04	0.08	0.08	0.08	0.08
FXR_19	1.00	0.23	0.53	0.36	0.03	0.50	0.48	0.43	0.47
FXR_20	0.23	1.00	0.21	0.27	0.04	0.20	0.19	0.24	0.20
FXR_21	0.53	0.21	1.00	0.38	0.04	0.83	0.80	0.44	0.73
FXR_22	0.36	0.27	0.38	1.00	0.03	0.35	0.34	0.39	0.33
FXR_23	0.03	0.04	0.04	0.03	1.00	0.04	0.04	0.02	0.04
FXR_24	0.50	0.20	0.83	0.35	0.04	1.00	0.92	0.42	0.74
FXR_25	0.48	0.19	0.80	0.34	0.04	0.92	1.00	0.41	0.72
FXR_26	0.43	0.24	0.44	0.39	0.02	0.42	0.41	1.00	0.40
FXR_27	0.47	0.20	0.73	0.33	0.04	0.74	0.72	0.40	1.00
FXR_28	0.37	0.17	0.50	0.25	0.04	0.47	0.46	0.31	0.50
FXR_29	0.41	0.16	0.56	0.27	0.03	0.52	0.50	0.34	0.50
FXR_30	0.25	0.20	0.26	0.20	0.04	0.25	0.24	0.26	0.25
FXR_31	0.39	0.22	0.45	0.35	0.03	0.43	0.42	0.44	0.41
FXR_32	0.31	0.23	0.37	0.87	0.03	0.34	0.33	0.34	0.32
FXR_33	0.03	0.05	0.03	0.02	0.11	0.03	0.02	0.02	0.04
FXR_34	0.01	0.01	0.02	0.01	0.01	0.02	0.02	0.01	0.02
FXR_35	0.59	0.24	0.70	0.42	0.03	0.65	0.63	0.49	0.61
FXR_36	0.45	0.18	0.71	0.33	0.03	0.66	0.64	0.41	0.68

	FXR_28	**FXR_29**	**FXR_30**	**FXR_31**	**FXR_32**	**FXR_33**	**FXR_34**	**FXR_35**	**FXR_36**
FXR_1	0.04	0.03	0.05	0.03	0.02	0.04	0.02	0.03	0.02
FXR_2	0.03	0.04	0.04	0.04	0.03	0.03	0.01	0.04	0.03
FXR_3	0.04	0.04	0.04	0.04	0.04	0.01	0.02	0.05	0.04
FXR_4	0.06	0.05	0.05	0.04	0.03	0.10	0.02	0.05	0.04
FXR_5	0.05	0.02	0.03	0.02	0.02	0.03	0.01	0.02	0.02
FXR_6	0.18	0.18	0.36	0.32	0.18	0.04	0.01	0.27	0.22
FXR_7	0.15	0.16	0.25	0.23	0.18	0.02	0.01	0.26	0.22
FXR_8	0.22	0.24	0.30	0.27	0.21	0.03	0.02	0.37	0.29
FXR_9	0.20	0.22	0.28	0.25	0.20	0.02	0.01	0.34	0.28
FXR_10	0.04	0.03	0.02	0.01	0.01	0.03	0.01	0.02	0.03
FXR_11	0.03	0.03	0.02	0.02	0.01	0.03	0.01	0.03	0.03
FXR_12	0.04	0.04	0.02	0.01	0.02	0.03	0.01	0.03	0.03
FXR_13	0.21	0.23	0.29	0.26	0.21	0.02	0.01	0.35	0.29
FXR_14	0.17	0.17	0.31	0.27	0.15	0.05	0.02	0.22	0.23
FXR_15	0.03	0.03	0.02	0.03	0.03	0.02	0.01	0.03	0.03
FXR_16	0.03	0.03	0.03	0.03	0.03	0.03	0.01	0.04	0.04
FXR_17	0.03	0.03	0.02	0.02	0.02	0.03	0.01	0.03	0.05
FXR_18	0.05	0.06	0.07	0.07	0.07	0.02	0.03	0.09	0.07
FXR_19	0.37	0.41	0.25	0.39	0.31	0.03	0.01	0.59	0.45
FXR_20	0.17	0.16	0.20	0.22	0.23	0.05	0.01	0.24	0.18
FXR_21	0.50	0.56	0.26	0.45	0.37	0.03	0.02	0.70	0.71

FXR_22	0.25	0.27	0.20	0.35	0.87	0.02	0.01	0.42	0.33
FXR_23	0.04	0.03	0.04	0.03	0.03	0.11	0.01	0.03	0.03
FXR_24	0.47	0.52	0.25	0.43	0.34	0.03	0.02	0.65	0.66
FXR_25	0.46	0.50	0.24	0.42	0.33	0.02	0.02	0.63	0.64
FXR_26	0.31	0.34	0.26	0.44	0.34	0.02	0.01	0.49	0.41
FXR_27	0.50	0.50	0.25	0.41	0.32	0.04	0.02	0.61	0.68
FXR_28	1.00	0.79	0.20	0.33	0.25	0.05	0.02	0.48	0.44
FXR_29	0.79	1.00	0.20	0.35	0.27	0.04	0.02	0.54	0.49
FXR_30	0.20	0.20	1.00	0.57	0.24	0.03	0.02	0.34	0.24
FXR_31	0.33	0.35	0.57	1.00	0.41	0.03	0.02	0.57	0.41
FXR_32	0.25	0.27	0.24	0.41	1.00	0.02	0.02	0.47	0.33
FXR_33	0.05	0.04	0.03	0.03	0.02	1.00	0.01	0.03	0.03
FXR_34	0.02	0.02	0.02	0.02	0.02	0.01	1.00	0.02	0.02
FXR_35	0.48	0.54	0.34	0.57	0.47	0.03	0.02	1.00	0.59
FXR_36	0.44	0.49	0.24	0.41	0.33	0.03	0.02	0.59	1.00

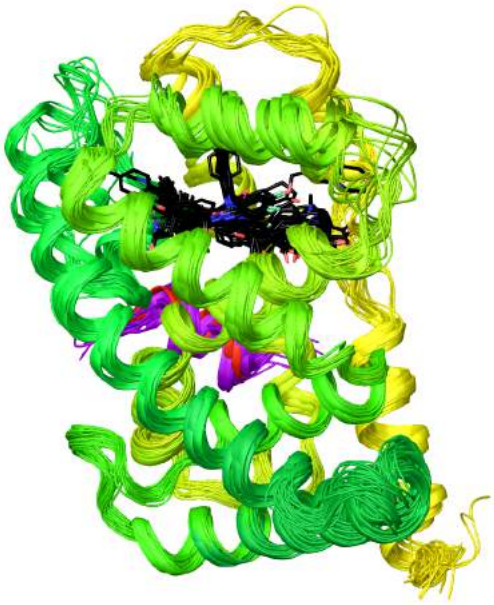


Figure S1. Superposition of all 28 FXR available crystal structures (shown in ribbons) with their co-crystallized ligands (in black) revealed a wide binding site.

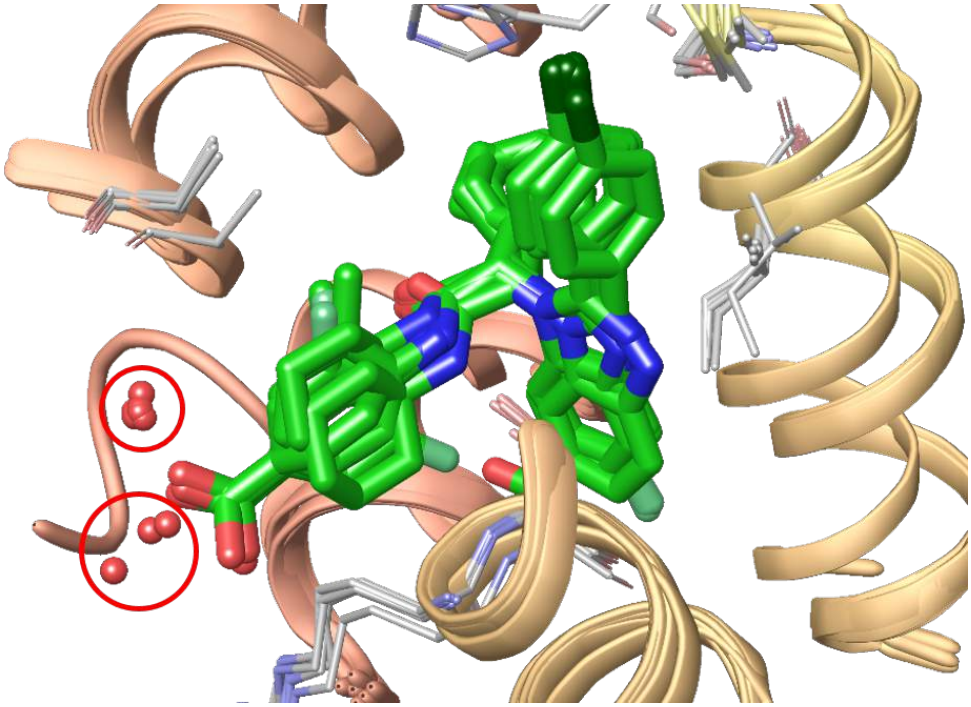


Figure S2. Alignment of benzimidazole crystal structures with PDB ID: 3OLF, 3OMK, 3OMM, 3OOF, 3OOK, 3OKH, 3OKI. Water molecules that appeared in more than three crystal structures were retained. For the benzimidazoles group, two water molecules were consistently crystallized and were retained during the calculations.

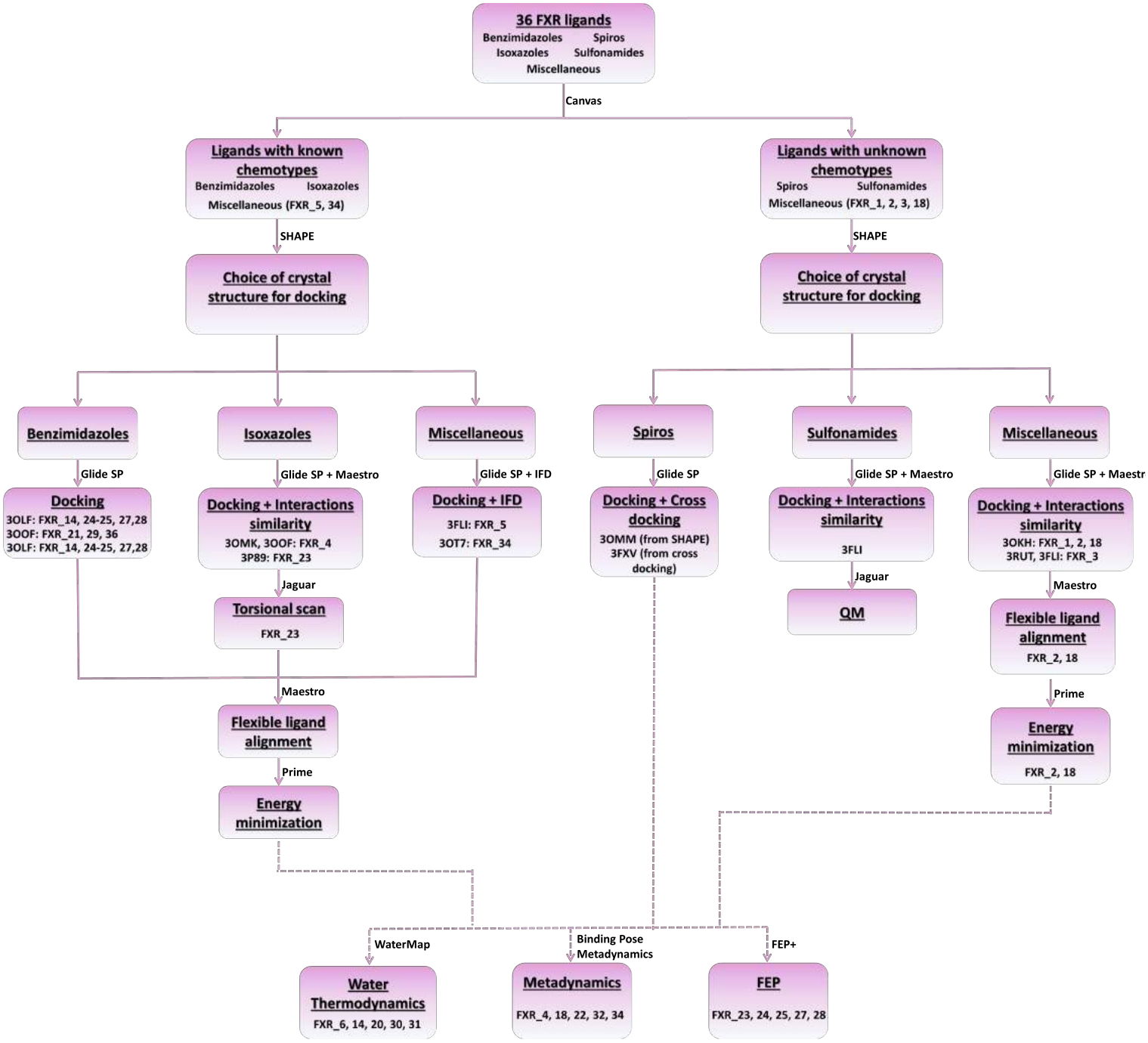
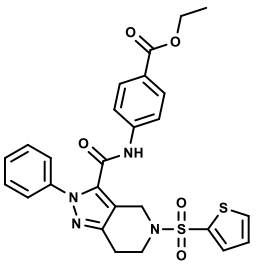


Figure S3. Diagram of the methodology used for the pose predictions. The dashed lines indicate calculations that were performed for only a number of FXR ligands.

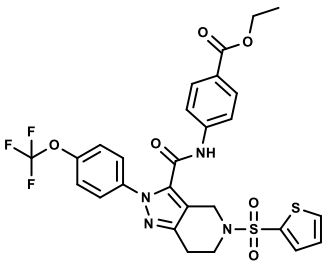
Table S3. Categorization of benzimidazole ligands based on the crystal structure in which they were docked.

3OLF	**3OOF**	**3OKI**
FXR_14	FXR_21	FXR_6-9, FXR_13
FXR_24-25	FXR_29	FXR_19-20, FXR_22
FXR_27-28	FXR_36	FXR_26, FXR_30-32, FXR_35

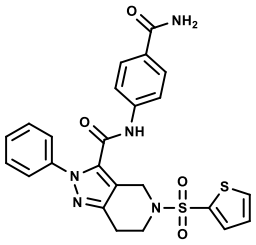
Table S4. Structures of 15 sulfonamides compounds used in relative binding free energy predictions.



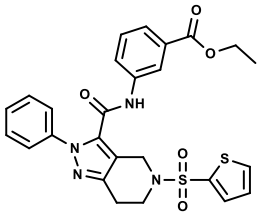
FXR_17



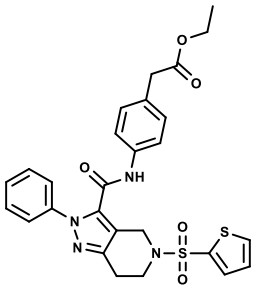
FXR_45



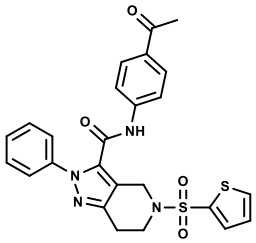
FXR_46



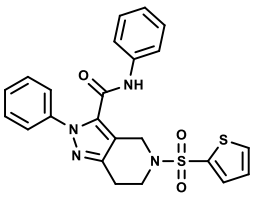
FXR_47



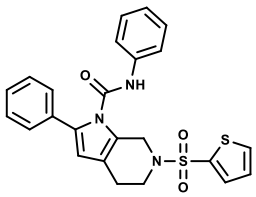
FXR_48



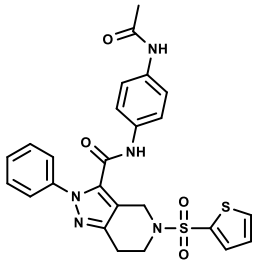
FXR_49



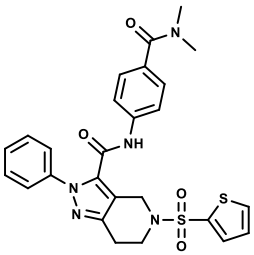
FXR_91



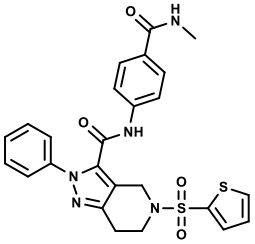
FXR_93



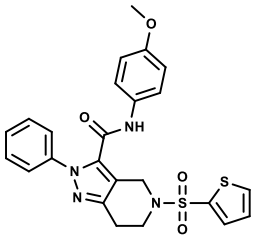
FXR_95



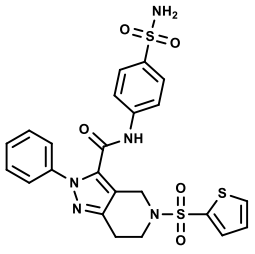
FXR_96



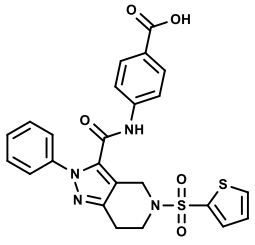
FXR_98



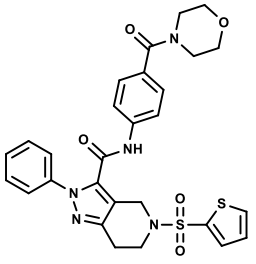
FXR_99



FXR_100

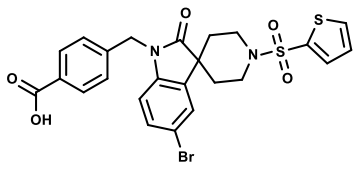


FXR_101

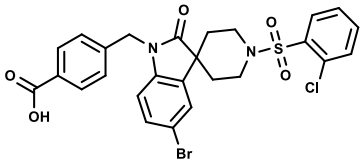


FXR_102

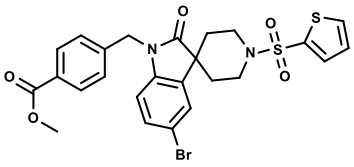
Table S5. Structures of 18 spiros compounds used in relative binding free energy predictions.



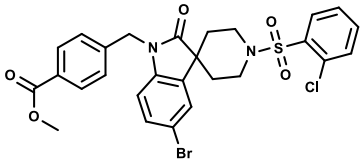
FXR_10



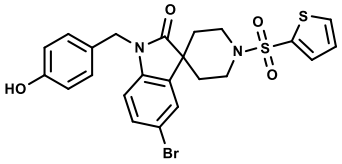
FXR_12



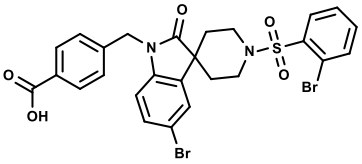
FXR_38



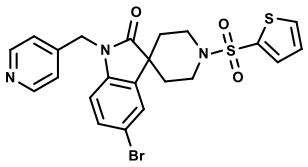
FXR_41



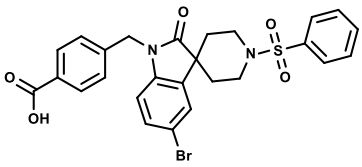
FXR_73



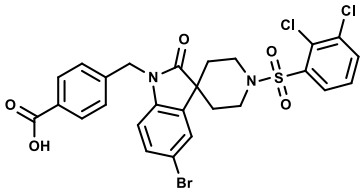
FXR_74



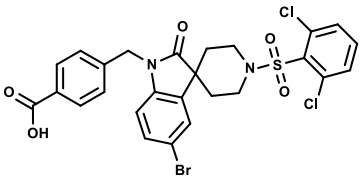
FXR_75



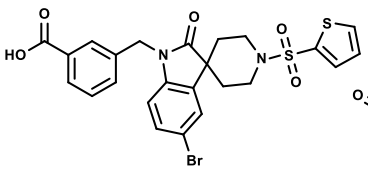
FXR_76



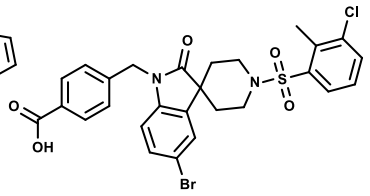
FXR_77



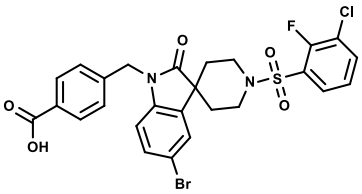
FXR_78



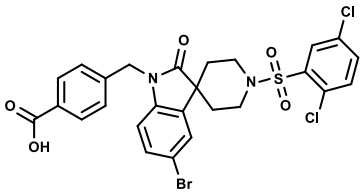
FXR_79



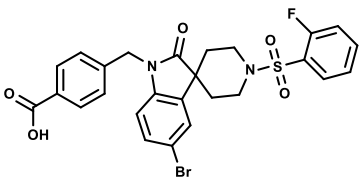
FXR_81



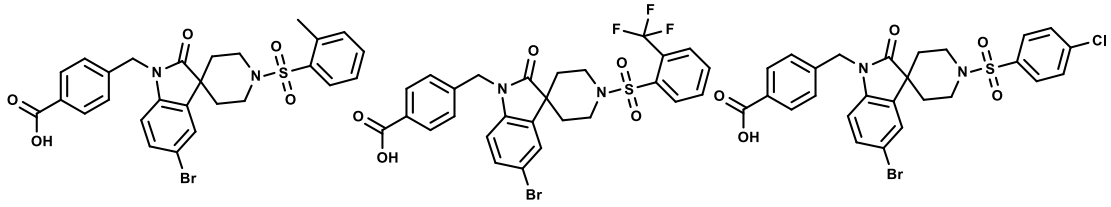
FXR_82



FXR_83



FXR_84



FXR_85 **FXR_88** **FXR_89**

Preparation of ligands for docking

The LigPrep [1] tool of Maestro 10.6 was used for the preparation of the ligand structures. The OPLS3 force field [2] was used and the ionization states of the ligands were determined by generating all possible states at pH 7.0 +/- 1.0 using Epik v3.6 [3-5]. In case of a racemic mixture all combinations of stereoisomers were generated. Other than this, the specified chiralities by the 2D structure were kept. The number of ligands generated by the process was set to seven and the number of low energy ring conformations was set to five. In addition, the ConfGen [6, 7] tool of Maestro 10.6 for conformational search was used to perform a flexible and time-efficient conformational search. The generated number of conformers was set to seven and generated conformers were then subjected to minimization.

Preparation of proteins for docking

The Protein Preparation Wizard tool [8, 9] was used to prepare the protein crystal structures. Bond orders were assigned, missing hydrogen atoms were added, and waters were deleted beyond 5 Å of the binding site. Prime v4.4 [10-12] was run for the placement and optimization of the missing side chains. Epic v3.6 [3-5] was used to generate probable ionization and tautomeric states at physiological pH for all heteroatom groups and the most appropriate were chosen. For the optimization of the hydrogen bond network, the hydrogen bond assignment option of the Protein Preparation Wizard was used. This module optimizes the orientation of the hydroxyl groups by performing 180^{o} flips of the asparagine and glutamine side chains as well as reorients the histidine ring and adjusts its charge state. This is an iterative process, which goes through all groups whose hydrogen bonds need to be optimized multiple times. The “sample water orientations” option was chosen during the hydrogen bond optimization. Minimization of all sampled hydrogens following optimization was also applied and the PROPKA option was used to perform the hydrogen bond optimization with protonation states of residues at a given pH. Finally, a restrained minimization was performed, during which the strain can be relieved without deviating too much from the input geometry. This can be achieved by terminating the minimization when the Root Mean Squared Deviation (RMSD) value of the heavy-atoms displacement during the minimization, reaches a default value of 0.3 Å.

Torsional scan with Jaguar

For the torsional scan, the Relaxed Coordinate Scan tool within Jaguar [13, 14] was used. In a 1D torsional scan, a geometry optimization along fixed increments of a specific dihedral is performed. The default settings for the basis set (6-31G**) and the level of theory options were kept DFT=Becke_3_Parameter/HF+Slater+Becke88+VWN+LYP (B3LYP)) [15]. The maximum interactions for the convergence criteria in the SCF tab were set to 100 and the Z-

matrix coordinates were selected in the Optimization tab. The dihedral type was chosen for the coordinates in the Scan tab and the increment was 5° for the torsional scan.

Introduction to FEP theory

The free energy difference between two states (named 1 and 2) having the same number of particles and potential energies $U_0(x)$ and $U_1(x)$ each can be calculated the following formula:

$$\Delta F = F_1 - F_0 = - kT\ln(Q_1/Q_0) \quad (1)$$

where $b=1/kT$, $\Delta U = U_1(x) - U_0(x)$ is the difference in the potential energies and the average is applied to configurations from state 0. Q_i is the partition function in the Γ-phase space,

$$Qi = \int d\Gamma \exp(-\beta U_i)$$

The partition function Q_1 can be written as

$$Q_1 = \int d\Gamma \exp(-\beta[(U_1 - U_0) + U_0]) = \int d\Gamma \exp(-\beta\Delta U)\exp(-\beta U_0)$$

and thus equation (1) is equivalent to:

$$\Delta F = - kT\ln\langle \exp(-\beta\Delta U)\rangle_0 \quad (2)$$

Hence, the ratio of the partition functions for states 0 and 1 in equation (1) is transformed in equation (2) to the ensemble average of the initial state. Correspondingly, the free energy difference can be expressed in terms of the average over the ensemble of state 1, which is the final state.

$$\Delta F = F_0 - F_1 = - kT\ln(Q_0/Q_1) = - kT\ln\langle \exp(-\beta\Delta U)\rangle_1 \quad (3)$$

Bennett Acceptance Ratio (BAR) method

When calculating the free energy differences by sampling for the reference state 0 and the reference state 1, from equations (2) and (3) respectively, the two estimates are different.

This problem can be partially overcome using the Bennett Acceptance Ratio (BAR) [16] method where both states 0 and 1 appear in equal roles. This method introduces an imaginary intermediate state with potential energy $U_*(x)$ and calculates the differences $U_*(x) - U_0(x)$ and $U_*(x) - U_1(x)$ using data sampled from $U_0(x)$ and $U_1(x)$ respectively. The free energy difference between the two states is given by the following formula:

$$\Delta F = (F_1 - F_*) + (F_* - F_0) = -kT\ln[(Q_1/Q_*)(Q_*/Q_0)] = kT\ln\frac{\langle \exp(-\beta(U_* - U_1))\rangle_1}{\langle \exp(-\beta(U_* - U_0))\rangle_0} \quad (4)$$

Bennett's approach was to find the optimal value for the potential of the imaginary intermediate state which minimizes the expected statistical error in the free energy difference. Eventually, the free energy difference can be estimated by solving the following two equations through iteration:

$$\Delta F^{BAR} = kTln\left[\frac{\sum_{i=1}^{N} f(-x_i)}{\sum_{i=1}^{N} f(x_i)}\right] + C - kTln\left(\frac{N_1}{N_0}\right) \quad (5)$$

$$C = \Delta F^{BAR} + kTln\left(\frac{N_1}{N_0}\right) \quad (6)$$

where N_0 and N_1 are the number of data points sampled from U_0 and U_1,

$f(x) = \frac{1}{1+\exp(x)}$ is the Fermi functions, $x = \frac{\Delta U - C}{k_b T}$ and $C = ln\frac{Q_0 N_1}{Q_1 N_0}$.

The statistical uncertainty of the BAR free energy is estimated by the following formula:

$$\sigma^2(\Delta G^{BAR}) = \frac{1}{N_0\beta}\left[\frac{\langle f^2(x)\rangle_0}{\langle f(x)\rangle_0^2} - 1\right] + \frac{1}{N_1\beta}\left[\frac{\langle f^2(-x)\rangle_1}{\langle f(-x)\rangle_1^2} - 1\right] \quad (7)$$

The error of the BAR free energy can also be calculated via bootstrapping. During this procedure N_0 and N_1 data points are randomly selected from the whole data sampled from U_0 and U_1 respectively, and are repeated (resampled) multiple times. After this step, the BAR free energy is calculated using the resampled data points. Through repetition of this process, the variance of the calculated BAR free energy can be estimated, which gives the bootstrapping estimated error of the BAR free energy.

Cycle closure method

In free energy calculations, it is feasible to calculate the relative binding affinities of a set of ligands with respect to a lead molecule. In theory, given the free energy is a state function, the estimated values of the relative free energy are expected to be independent of the path followed during the mutations.

For example, given the three ligands L1, L2 and L3 of Figure S1, there are two strategies for the calculation of the free energy difference between ligands L1 and L3: (a) by mutating directly L1 to directly L1 to L3 and (b) by first mutating L1 to L2 and then L2 to L3, and adding the two free energy estimates. Theoretically, the calculated free energy differences from the two paths should be the same, i.e. $F_{12}^{BAR} + F_{23}^{BAR} = -F_{31}^{BAR}$ (see Figure S1). In practice, though, due to errors in each calculation, the estimated values from the above two mutation paths are usually somewhat different.

The errors in FEP calculations can be systematic, due to inability of the force field to precisely describe the interactions and molecules motions, and errors coming from the unconverged simulation, either due to random or systematic incomplete sampling of the phase space, or from the BAR estimator.

FEP+ uses the cycle closure method, to satisfy the independence of the calculated free energy difference from the path and estimate the errors [17]. According to this method, the free energy values calculated from the BAR method are corrected in that way, so that the sum of the free energy differences from a closed cycle, equals to zero. The deviation of this sum from zero (called hysteresis of the cycle) can give an estimation of the error.

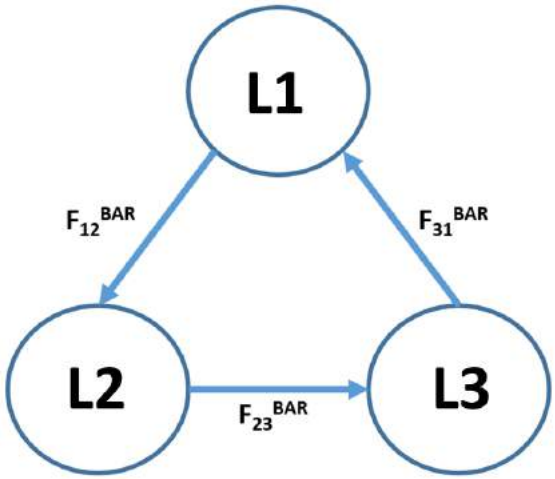


Figure S4. Thermodynamic cycle of three ligands, L1, L2 and L3. Each circle represents a ligands and each arrow (or edge) an FEP calculation. The energy values near the arrows correspond to the calculated with the BAR method binding free energy differences between the two ligands linked by the arrow.

According to the above description, for the simplest case of a cycle consisting of three ligands (see Figure S1) the cycle closure free energy differences are calculated from the following relationships:

$$F_{12} = \frac{2}{3}F_{12}^{BAR} - \frac{1}{3}(F_{23}^{BAR} + F_{31}^{BAR}) \quad (8)$$

$$F_{23} = \frac{2}{3}F_{23}^{BAR} - \frac{1}{3}(F_{12}^{BAR} + F_{31}^{BAR}) \quad (9)$$

$$F_{31} = \frac{2}{3}F_{31}^{BAR} - \frac{1}{3}(F_{12}^{BAR} + F_{23}^{BAR}) \quad (10)$$

where $F_{ij} = F_j - F_i$ are the free energy differences from the cycle closure method

and $F_{ij}^{BAR} = F_j^{BAR} - F_i^{BAR}$ are the free energy differences from the BAR method.

The cycle closure errors are given by the following formula:

$$\sigma = \frac{\Delta}{\sqrt{3}} \quad (11)$$

where Δ is the hysteresis of the cycle, calculated by the formula $\Delta = F_{12}^{BAR} + F_{23}^{BAR} + F_{31}^{BAR}$.

The above model can be easily generalized to more complicated cases, when the cycle is defined by more than three ligands and when an edge is common in two different cycles.

Replica Exchange with Solute Tempering (REST) Enhanced Sampling

As previously described, errors in the FEP can be introduced due to the insufficient sampling and the unconverged simulations. FEP+ uses a combination of a λ schedule, which is important for stratification, and the REST2 enhanced sampling method [18-20]. The latter is a Hamiltonian replica exchange method where only a small region of the system is effectively "heated up". Hence, a small number of replicas is sufficient to achieve the sampling efficiency, in contrast to other replica exchange techniques, where all the system is "heated up", thus requiring a large number of replicas. The region of increased effective temperature (called hot region) includes the ligand and some neighboring residues, because the method assumes that the slow degrees of freedom of the system, which cause the quasi-non ergodicity problem of incomplete sampling, are located within a close proximity of the bound ligand.

The λ schedule is used to increase the overlap in the regions of phase space that the two states (that of potential energy U_0 and the one of potential energy U_1) explore. This is important because one can incur large errors during the estimation of the binding affinities if the potential energies U_0 and U_1 are significantly dissimilar. To overcome this problem, the free energy difference is divided into a series of small steps, which correspond to alchemical intermediate states, and during which the potential energy of the initial state is gradually transformed to the potential energy of the final state. For this, a coupling parameter λ is typically used:

$$U = (1-\lambda)U_0 + \lambda U_1$$

In FEP/REST concomitantly with the alchemical transformation from the initial λ window to the final one, the effective temperature of the aforementioned hot region is gradually increased from a specific value (usually 300 K) in the first lambda window to the highest value (approximately 900 K) in the middle lambda window and subsequently decreases until it reaches again its initial value (300 K) in the final lambda window. The term *effective* temperature is used because in reality only potential energy is scaled by a factor and the term "temperature for replica m, means the effective temperature of the protein with the unscaled potential energy. In fact, all of the replicas run at the same temperature. In this way, the enhanced sampling is achieved by the exchange of configurations of the initial and final λ windows with the intermediate windows at which the potential is scaled.

Metadynamics Simulations

Metadynamics simulations are useful in sampling regions of the phase space not accessible in the conventional molecular dynamics simulations due to high energy barriers. The method introduces a time-dependent bias as a function the slow degrees of freedom of the system, which are called collected variables. These bias urges the system to visit neighboring free energy landscapes. The collective variable used for our calculations is the RMSD deviation from the initial pose.

The metadynamics simulations were run in the NVT ensemble, using a time constant for coupling 0.1 ps for the Berendsen thermostat. The bias potential height was 0.05 kcal/mol with 1 ps interval.

Running FEP+

The running protocol of FEP+ consists of several consecutive steps, which are automatically performed. First, the building of the final system geometry is performed, using an orthorhombic box shape and a 5 Å width buffer. This is succeeded by the assignment of the OPLS3 force field parameters. The final system size for spiros is approximately 21,000 atoms and for sulfonamides 17,000 atoms. Next, in the equilibration phase, a Brownian dynamics simulation with restraints on solute heavy atoms is performed at the NVT ensemble, at T = 10 K for 100 ps. The force constant is set at 50 kcal/mol/Å. Next, an MD simulation is performed, with restraints on solute heavy atoms in the NVT ensemble at T = 10 K for 12 ps. The Langevin thermostat [21] and the Berendsen barostat [22] are used. This is followed by MD simulations, first with restraints on solute heavy atoms in the NPT ensemble, at T = 10 K for 36 ps and then with no restraints for 240 ps in the NPT ensemble. For the production run, the replica exchange with solute tempering (REST) MD [18-20] simulation is performed in the NPT ensemble for 5 ns.

Table S6. Dataset compounds and the PDB IDs structures, which were used to dock them.

Name	PDB	Name	PDB
FXR_1	3OKH	FXR_19	3OKI
FXR_2	3OKH	FXR_20	3OKI
FXR_3	3RUT/ 3FLI	FXR_21	3OOF
FXR_4	3OOF/3OMK	FXR_22	3OKI
FXR_5	3FLI	FXR_23	3P89
FXR_6	3OKI	FXR_24	3OLF
FXR_7	3OKI	FXR_25	3OLF
FXR_8	3OKI	FXR_26	3OKI
FXR_9	3OKI	FXR_27	3OLF
FXR_10	3FXV/ 3OMM	FXR_28	3OLF
FXR_11	3FXV/ 3OMM	FXR_29	3OOF
FXR_12	3FXV/ 3OMM	FXR_30	3OKI
FXR_13	3OKI	FXR_31	3OKI
FXR_14	3OLF	FXR_32	3OKI
FXR_15	3FLI	FXR_33	3FXV
FXR_16	3FLI	FXR_34	1OT7
FXR_17	3FLI	FXR_35	3OKI
FXR_18	3OKH	FXR_36	3OOF

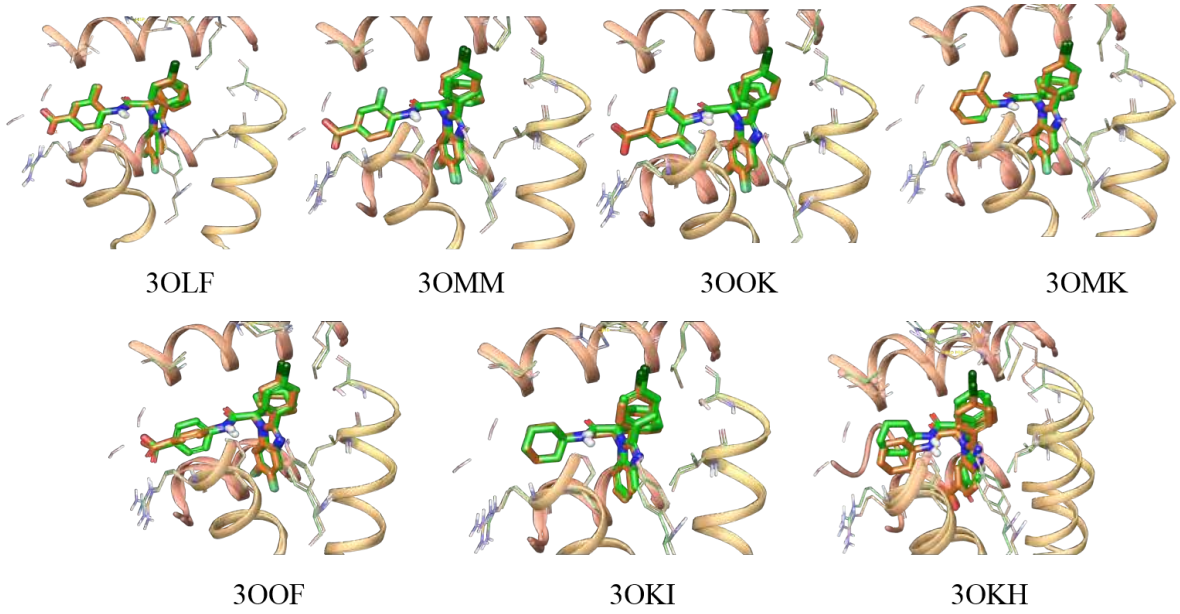


Figure S5. Comparison of docking poses (green) of all benzimidazole native ligands in 3OLF crystal structure with the corresponding crystal structures (orange). Notice the flip phenyl ring bound to the amide linker in the 3OOF structure due to streric clashes.

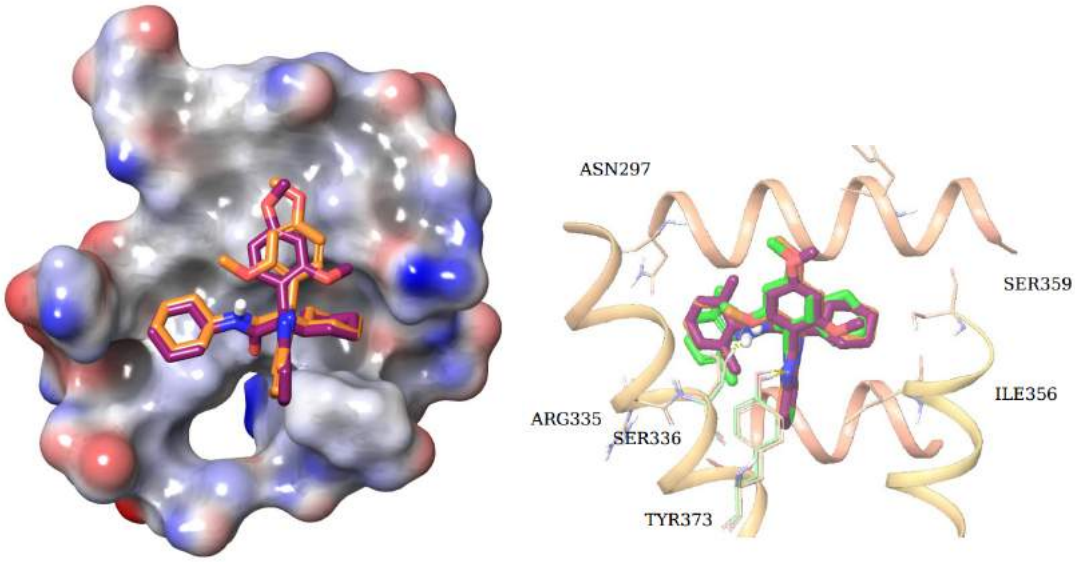


Figure S6. FXR_6 (left) and FXR_14 (right) exhibiting double occupancy.

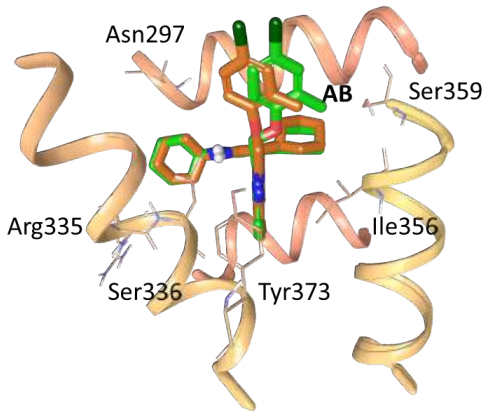


Figure S7. Pose in chain C aligned with the submitted pose (RMSD = 0.823 Å).

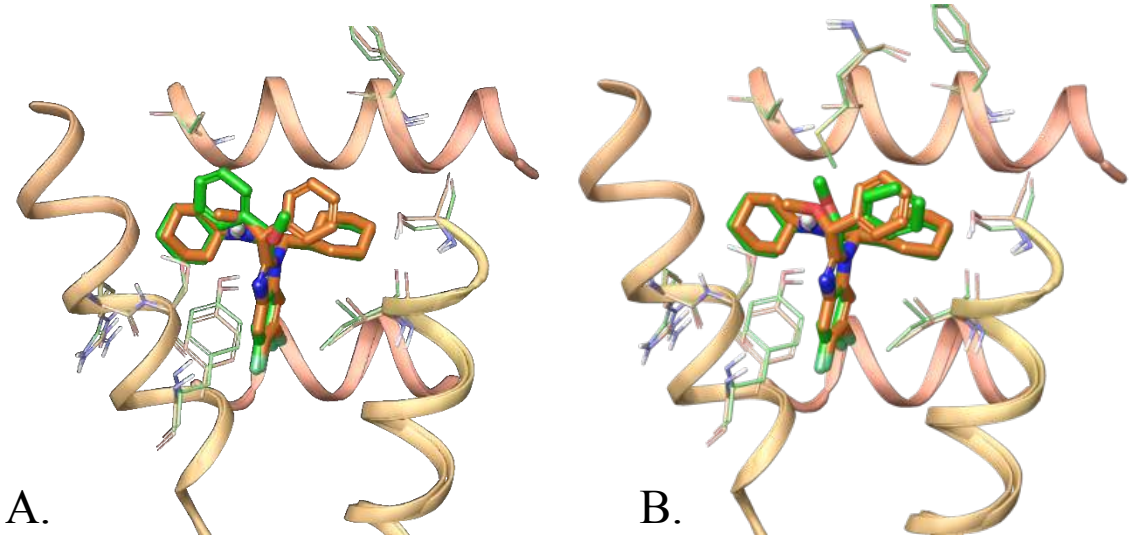


Figure S8. FXR_22 (A.) submitted pose (*green*) according to Metadynamics results aligned to crystal structure (orange) RMSD = 2.20 Å (B.) Alignment of the second pose examined with metadynamics (green) with the crystal one (orange) RMSD = 0.70 Å.

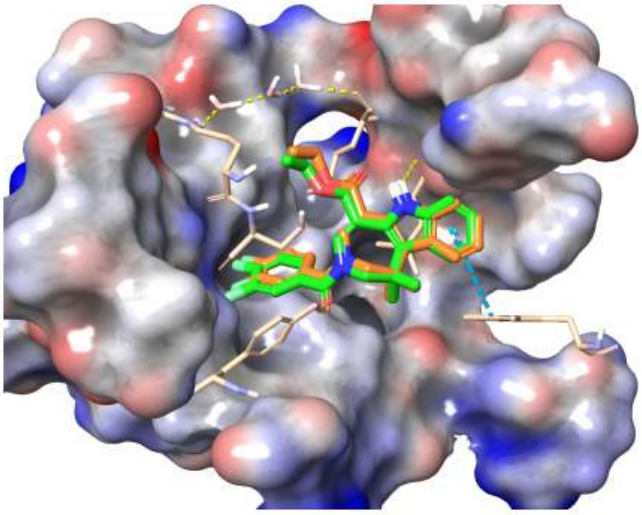


Figure S9. Comparison of FXR_5 predicted pose with crystal structure (RMSD = 0.68 Å).

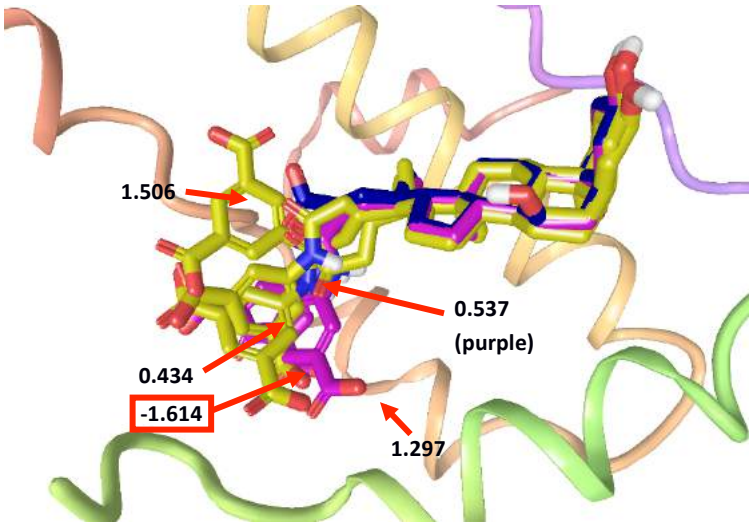


Figure S10. FXR_34 poses that were submitted to metadynamics calculation. Three IFD poses (yellow), two Glide SP poses (purple). The native 1OT7 ligand is depicted in dark blue.

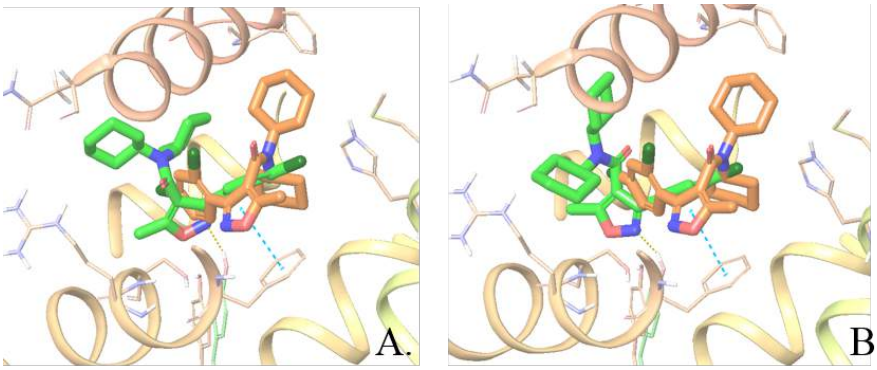


Figure S11. FXR_4 crystal structure (orange) superimposed with the pose prediction (green) in 3OOF (A) (RMSD = 6.77 Å) and in 3OMK (B) (RMSD = 6.96 Å).

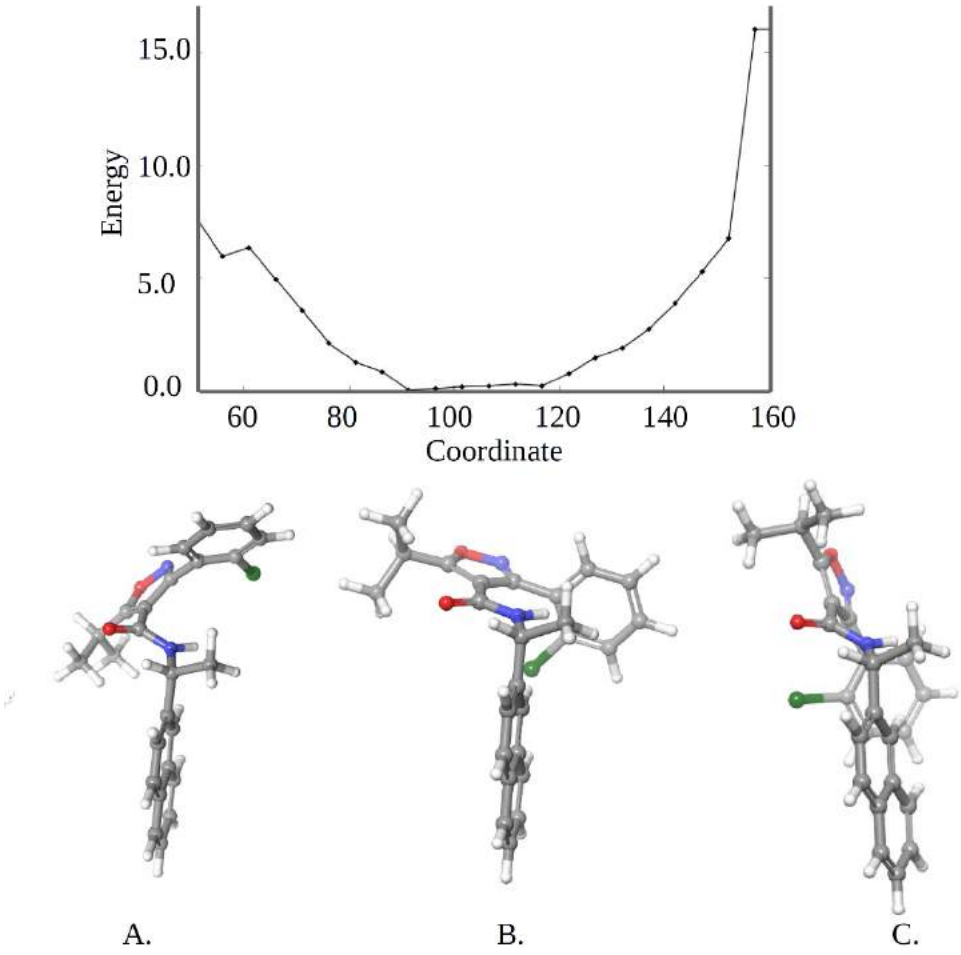


Figure S12. Torsional scan of the bond between the isoxazole ring and the amide in FXR_23 compound.

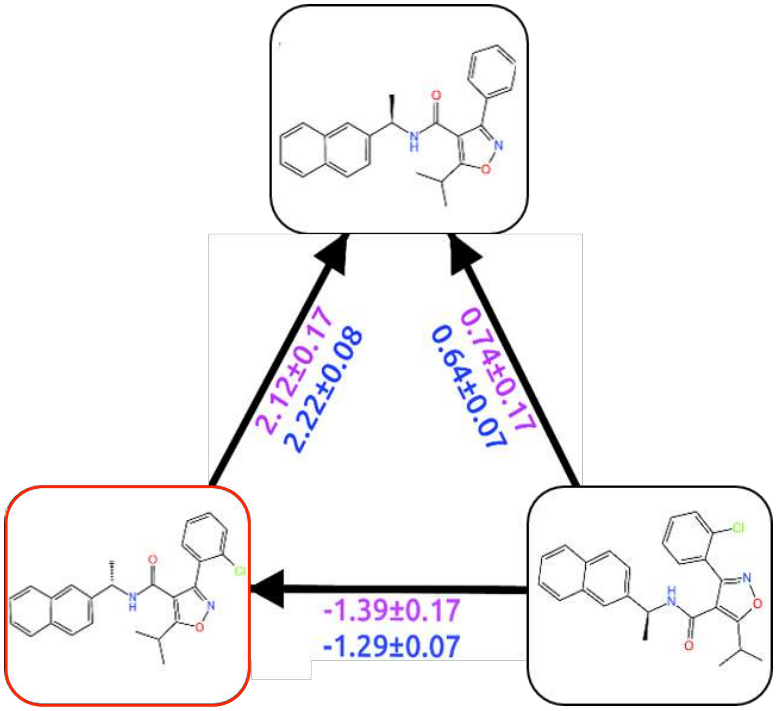


Figure S13. FEP map for FXR_23 compound.

Table S7. Shape similarity results.

	FXR_2	FXR_4	FXR_10	FXR_11	FXR_12	FXR_18
3OMM	0.526	0.612	0.538	0.458	0.526	0.558
3OKH	0.570	0.590	0.486	0.460	0.430	0.599
3P89	0.554	0.544	0.366	0.366	0.388	0.428
3RVF	0.581	0.560	0.390	0.390	0.380	0.450
1OSH	0.677	0.645	0.490	0.524	0.450	0.517
4OIV	0.557	0.458	0.550	0.466	0.507	0.540
4QE8	0.557	0.458	0.550	0.466	0.507	0.446
3L1B	0.41	0.420	0.400	0.400	0.400	0.510
1OT7	0.55	0.520	0.560	0.640	0.560	0.520

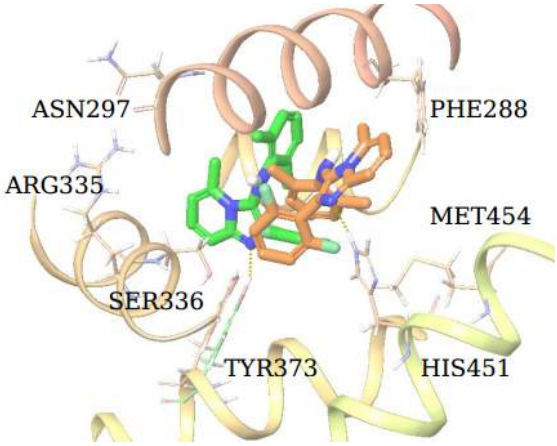


Figure S14. FXR_2 predicted (green) and crystal (orange) binding pose (RMSD = 7.47 Å).

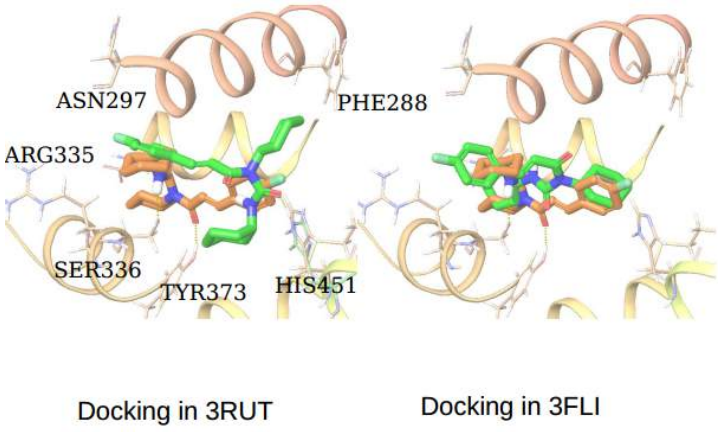


Figure S15. FXR_3 submitted poses (green) superimposd to crystal structure (orange). Left: RMSD = 8.37 Å, Right: RMSD = 6.91 Å.

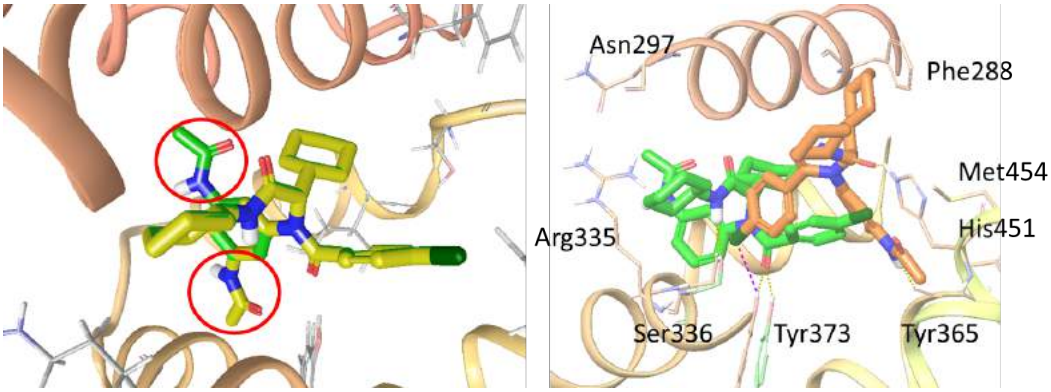


Figure S16. Left: FXR_18 poses submitted to binding pose metadynamics calculations. Right: Overlap of the predicted with the actual crystal pose (RMSD = 8.42 Å).

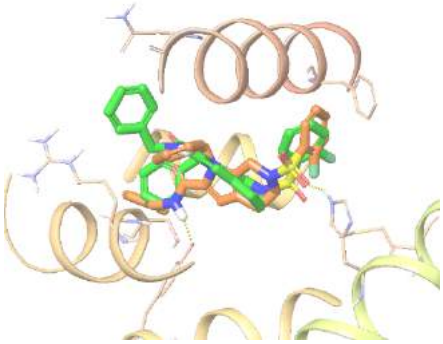


Figure S17: FXR_15 crystal structure (orange) superimposed with the pose prediction (green) in 1FLI (RMSD = 5.66 Å).

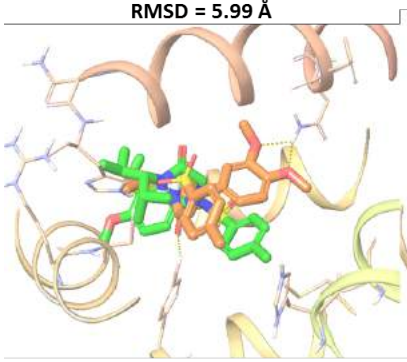
Table S8. Interaction fingerprints tanimoto similarity for spiros ligands.

FXR ligand	Tanimoto similarity
3OMM native ligand	1.000
FXR_79	0.654
FXR_84	0.630
FXR_85	0.630
FXR_81	0.621
FXR_74	0.607
FXR_12	0.607
FXR_76	0.607
FXR_78	0.607
FXR_77	0.607
FXR_83	0.607
FXR_89	0.600
FXR_10	0.593
FXR_88	0.586
FXR_82	0.552

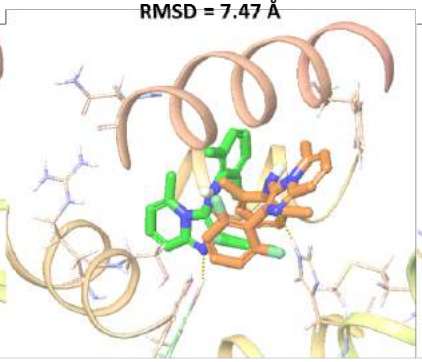
Table S9. The Interaction Similarity score for each compound of the sulfonamides subset to the native ligand of 3OOF crystal structure.

Ligand	**Interaction Similarity score**
3OOF native ligand	1.000
FXR_91	0.654
FXR_93	0.630
FXR_95	0.630
FXR_100	0.621
FXR_101	0.607
FXR_17	0.607
FXR_47	0.607
FXR_102	0.607
FXR_46	0.607
FXR_99	0.607
FXR_96	0.600
FXR_98	0.593
FXR_49	0.586
FXR_45	0.552
FXR_48	0.464

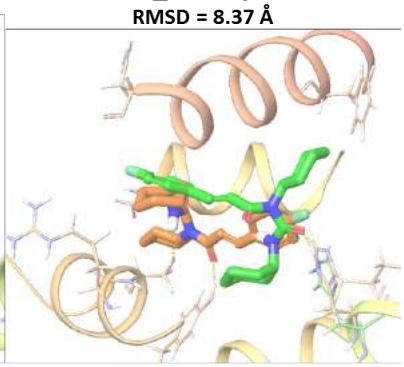
FXR_1
RMSD = 5.99 Å



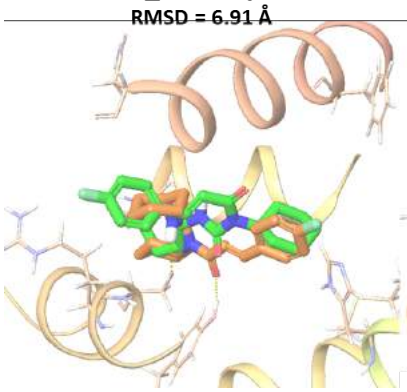
FXR_2
RMSD = 7.47 Å



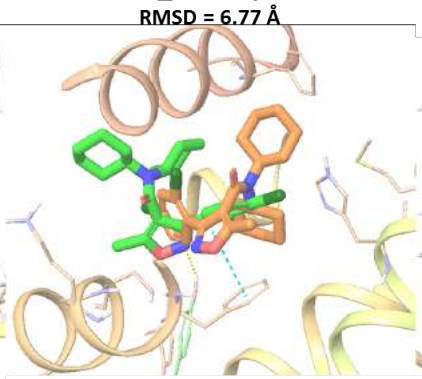
FXR_3 - 1st pose
RMSD = 8.37 Å



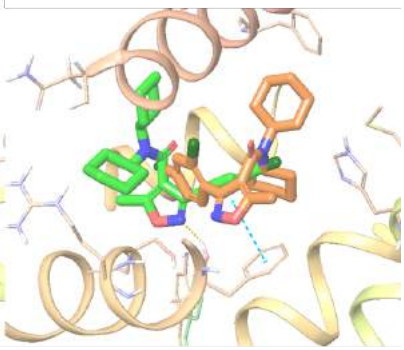
FXR_3 – 2nd pose
RMSD = 6.91 Å



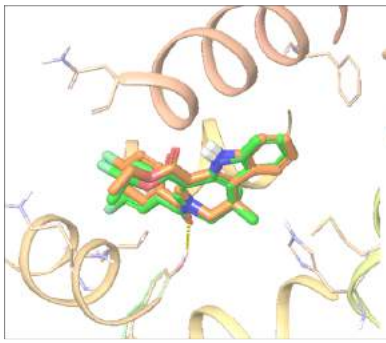
FXR_4 – 1st pose
RMSD = 6.77 Å



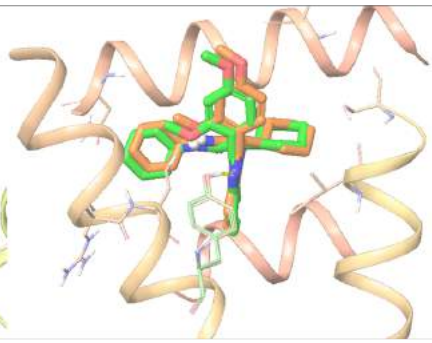
FXR_4 – 2nd pose
RMSD = 6.96 Å



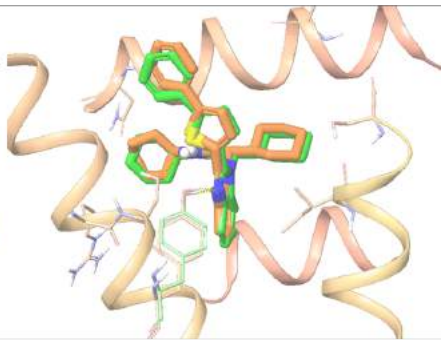
FXR_5
RMSD = 0.379 Å



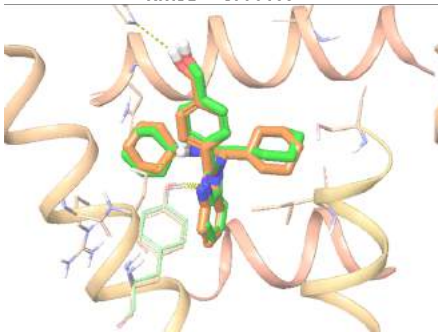
FXR_6
RMSD = 0.528 Å



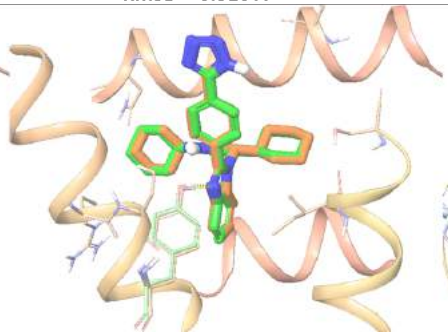
FXR_7
RMSD = 0.517 Å



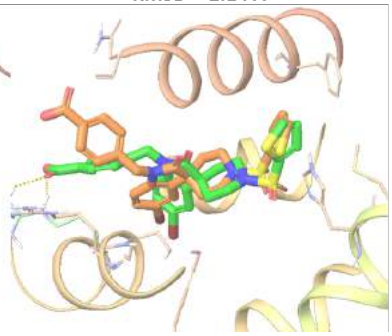
FXR_8
RMSD = 0.444 Å



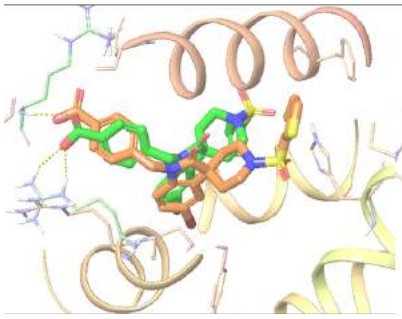
FXR_9
RMSD = 0.316 Å



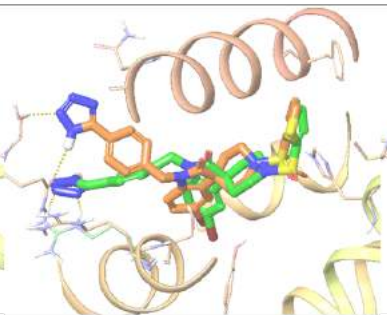
FXR_10 – 1st pose
RMSD = 2.14 Å



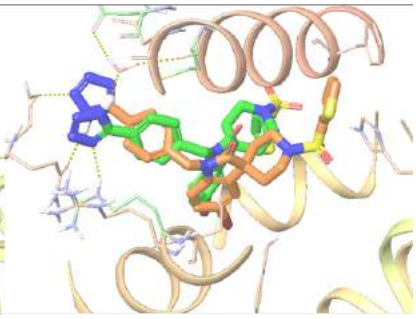
FXR_10 – 2nd pose
RMSD = 4.43 Å



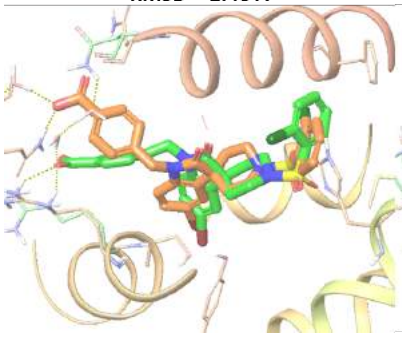
FXR_11 – 1st pose
RMSD = 2.49 Å



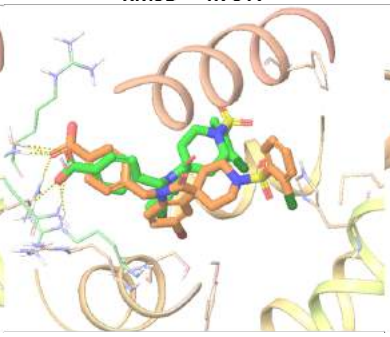
FXR_11 – 2nd pose
RMSD = 4.38 Å



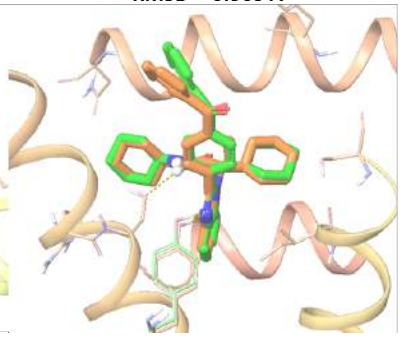
FXR_12 – 1st pose
RMSD = 2.48 Å



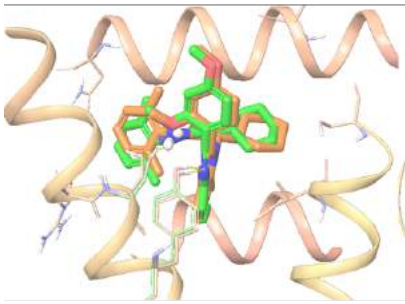
FXR_12 – 2nd pose
RMSD = 4.78 Å



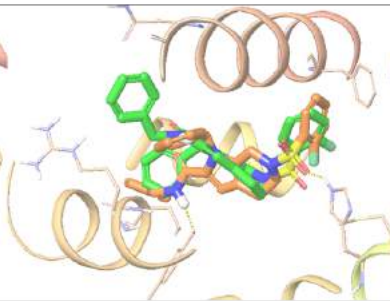
FXR_13
RMSD = 0.565 Å



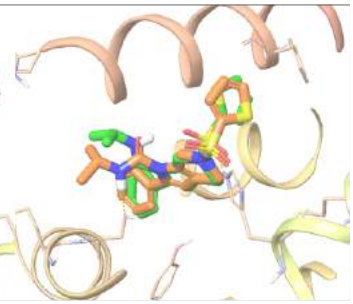
FXR_14
RMSD = 0.691 Å



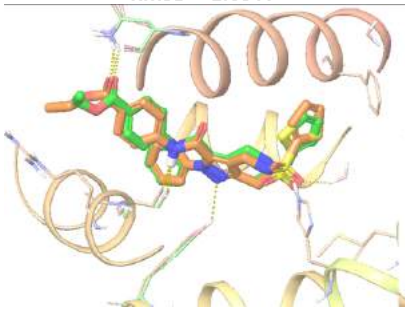
FXR_15
RMSD = 5.66 Å



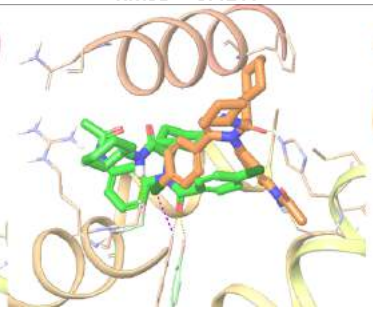
FXR_16
RMSD = 1.57 Å



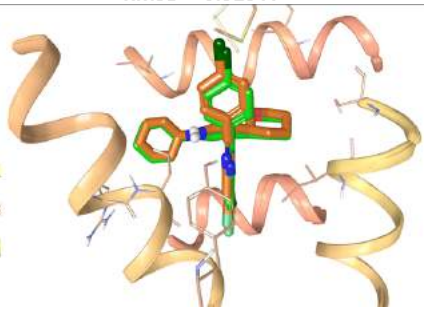
FXR_17
RMSD = 1.63 Å



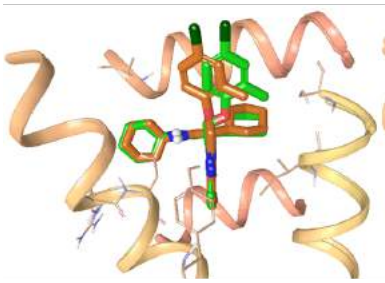
FXR_18
RMSD = 8.42 Å



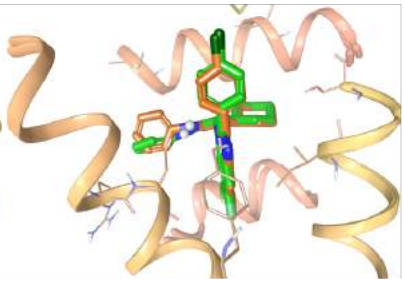
FXR_19
RMSD = 0.525 Å



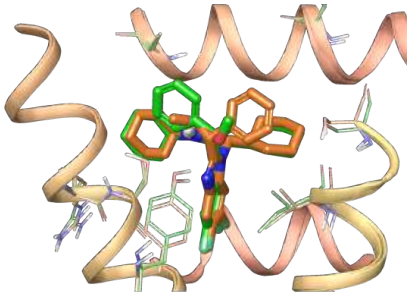
FXR_20
RMSD = 0.823 Å



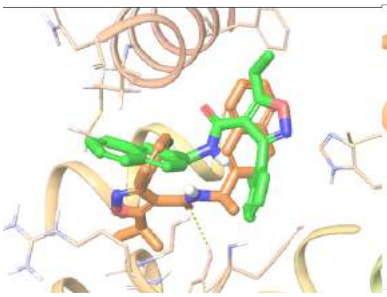
FXR_21
RMSD = 0.798 Å



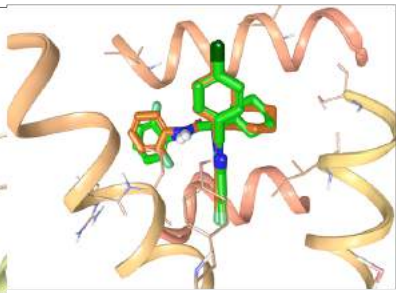
FXR_22
RMSD = 2.2 Å



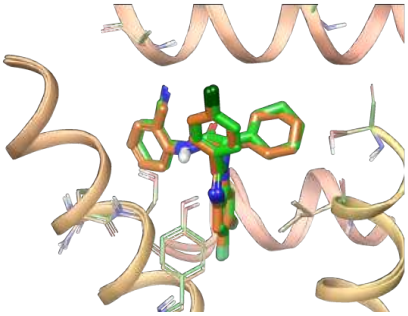
FXR_23
RMSD = 7.76 Å



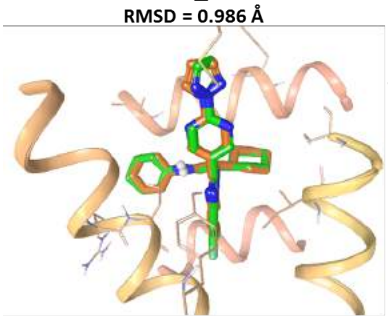
FXR_24
RMSD = 1.19 Å



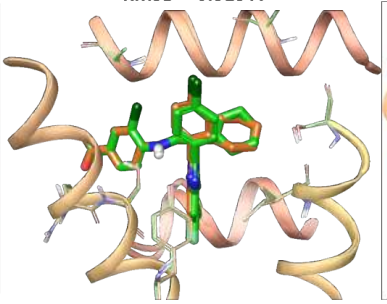
FXR_25
RMSD = 0.328 Å



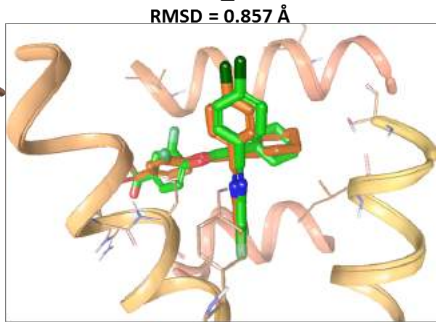
FXR_26
RMSD = 0.986 Å



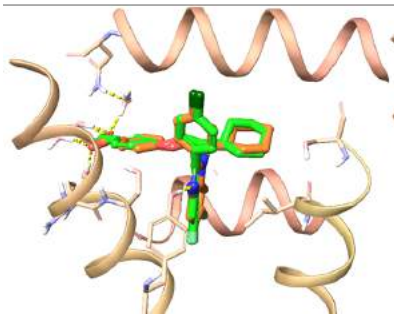
FXR_27
RMSD = 0.313 Å



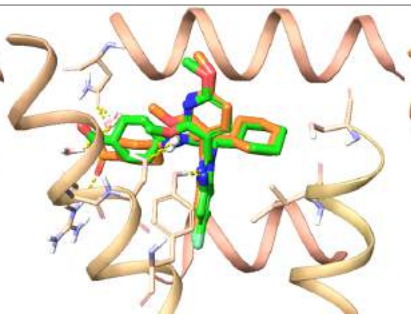
FXR_28
RMSD = 0.857 Å



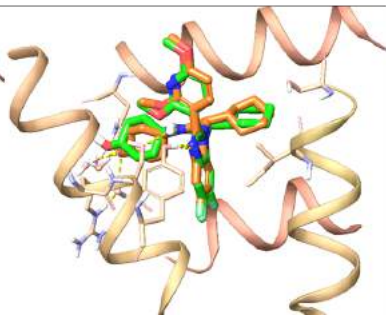
FXR_29
RMSD = 0.393 Å



FXR_30
RMSD = 0.893 Å



FXR_31
RMSD = 0.542 Å



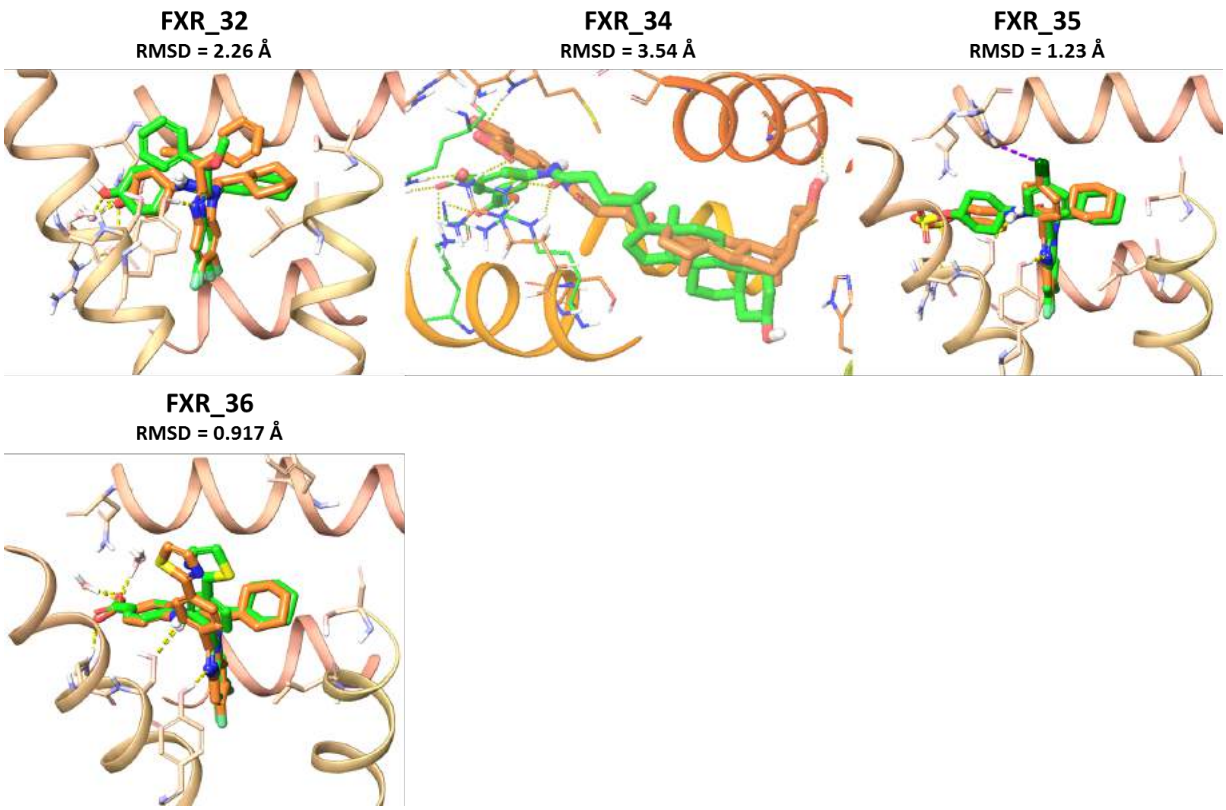


Figure S18. Superposition of all 36 ligands predicted poses (except FXR_33 for which the crystal structure was not solved) with the respective crystal structures.

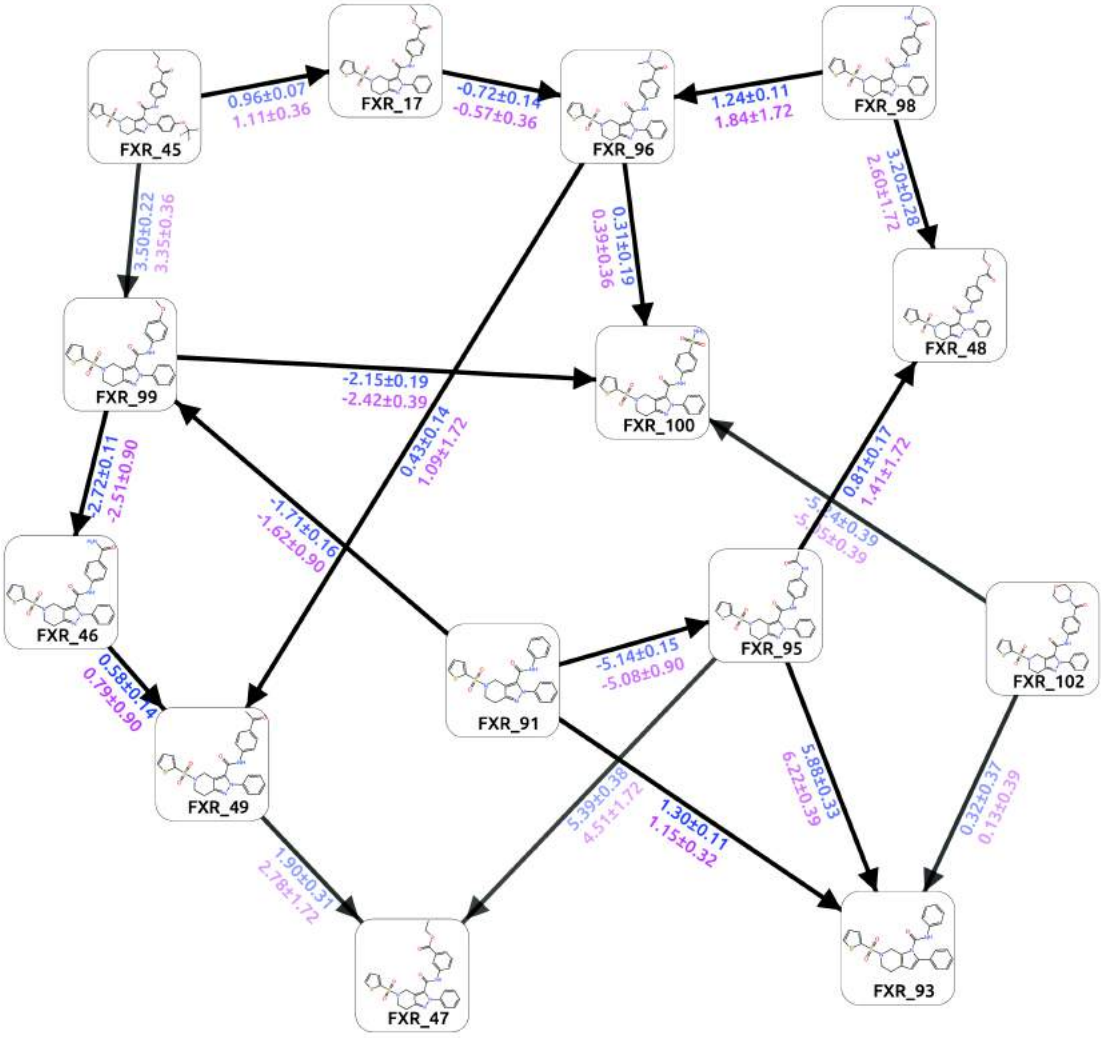


Figure S19. FEP Map as generated by FEP+ for the connections between the sulfonamide subset.

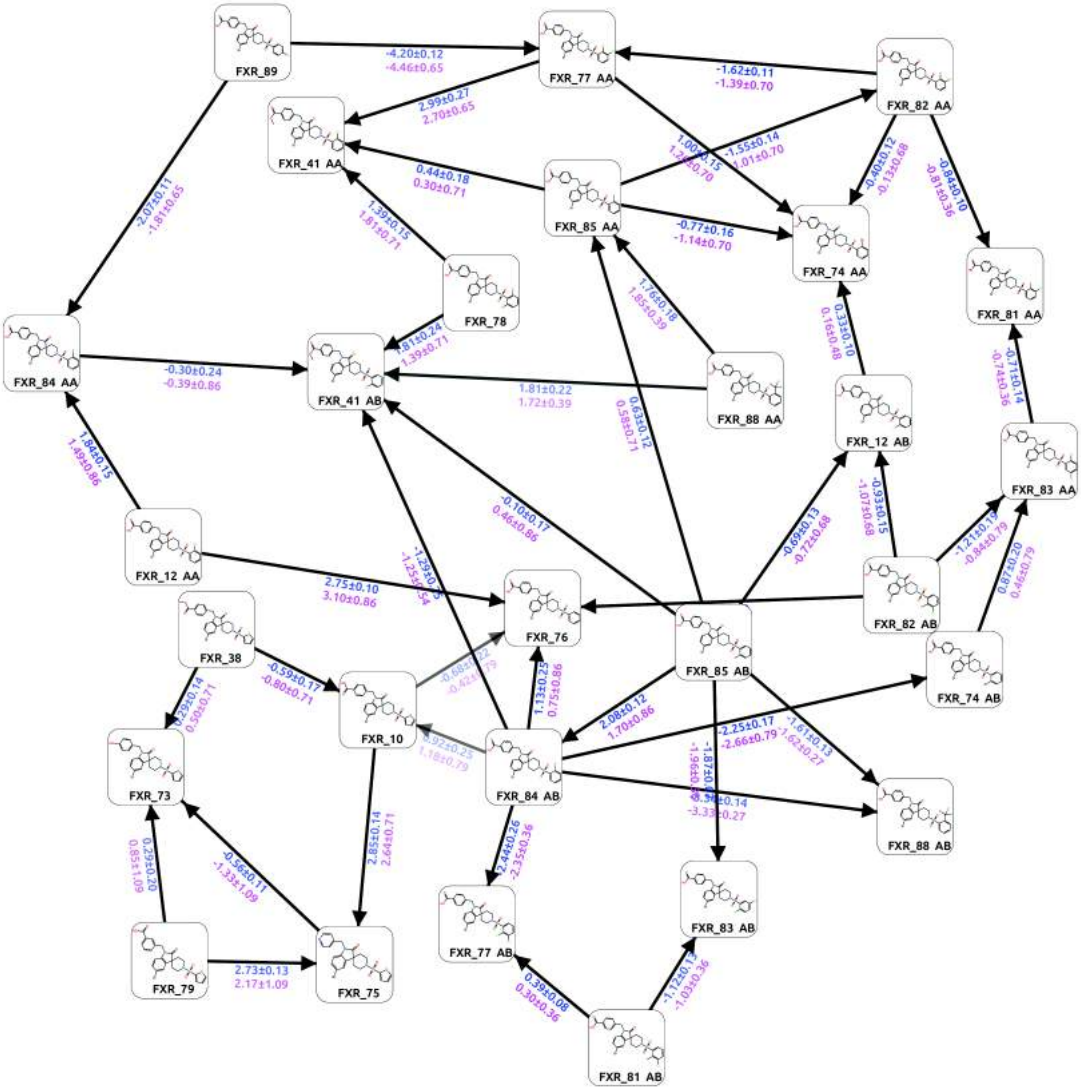


Figure S20. FEP Map as generated by FEP+ for the connections between the spiros subset.

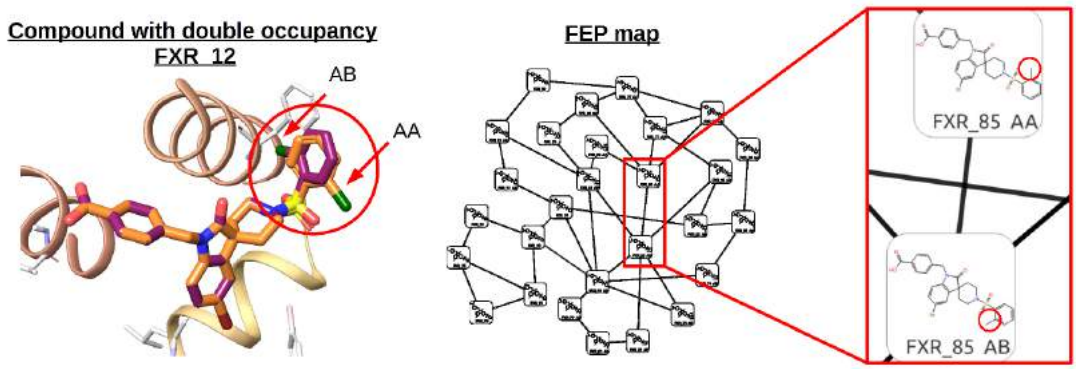


Figure S21. Left: FXR_12 displaying double occupancy depending chlorine orientation. Right: Incorporation of both poses (termed AA and AB) in the FEP map.

Calculation of MUE and RMSE errors

According to error propagation theory, the uncertainties of the independent variables affect the error of the dependent value as in the following formula:

$$\delta f(x,y,z,\ldots) = \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 \delta x + \left(\frac{\partial f}{\partial y}\right)^2 \delta y + \left(\frac{\partial f}{\partial z}\right)^2 \delta z + \ldots}$$

where δf, δx, δy, δz,… are the errors of f,x,y,z,… values.

The Mean Unsigned Error (MUE) value is given by the formula:

$$\text{MUE} = \frac{\sum \left|\Delta G_{exp_i} - \Delta G_{pred_i}\right|}{N}$$, where N is the number of sets of predicted and experimental values.

If we assume that the error in MUE comes only from the predicted values and that the experimental error is 0, then the error of MUE is calculated by this formula:

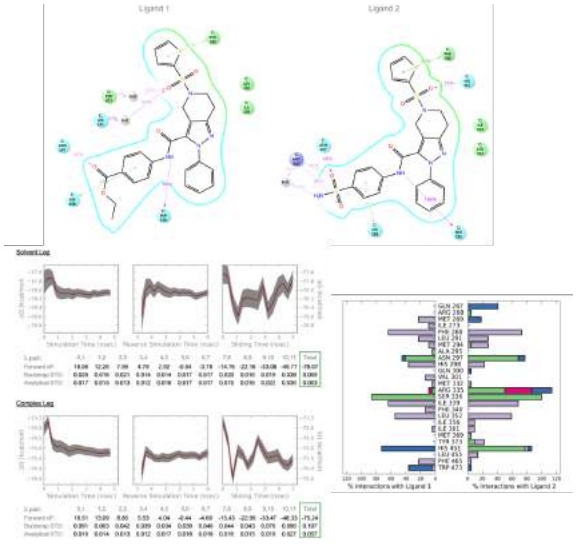
$$\delta\text{MUE} = \frac{1}{N}\sqrt{\sum_{i=1}^{N}\left(\Delta G_{pred_i}\delta\Delta G_{pred_i}\right)^2}$$

Accordingly, RMSE's function is:

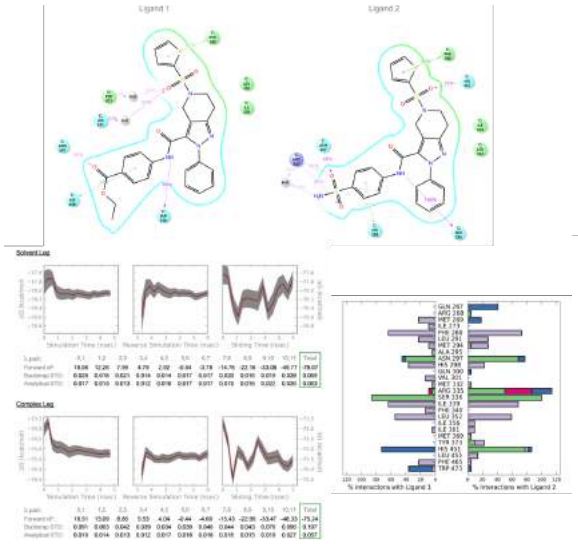
$$\text{RMSE} = \sqrt{\frac{\sum\left(\Delta G_{exp_i} - \Delta G_{pred_i}\right)^2}{N}}$$

and its error is: $\delta\text{RMSE} = \sqrt{\frac{1}{N}}$

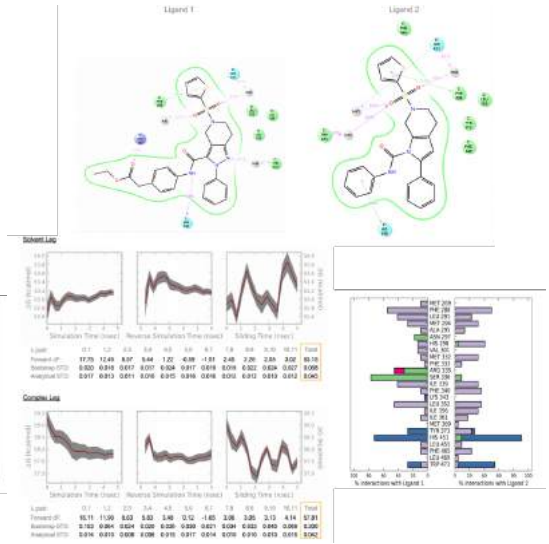
All convergence plots similar to Figure 17 in the main text for all mutation pairs, follow.



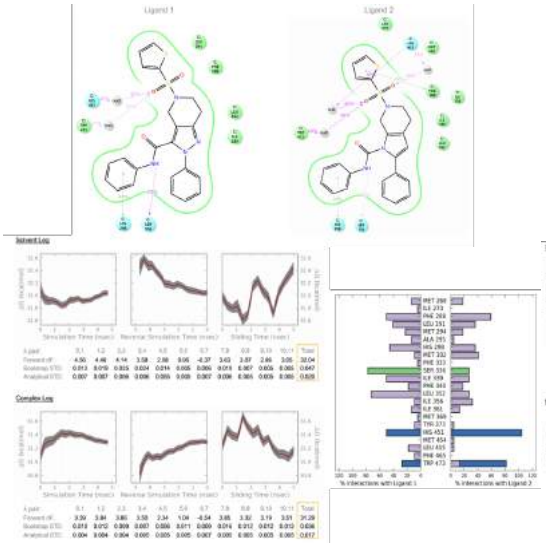
FXR_17 to FXR_100



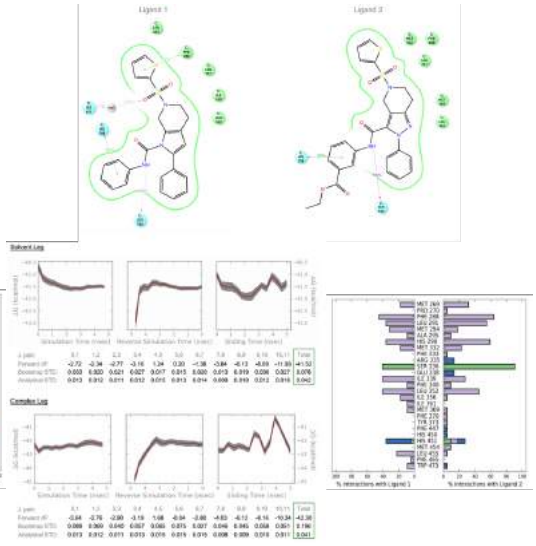
FXR_48 to FXR_93



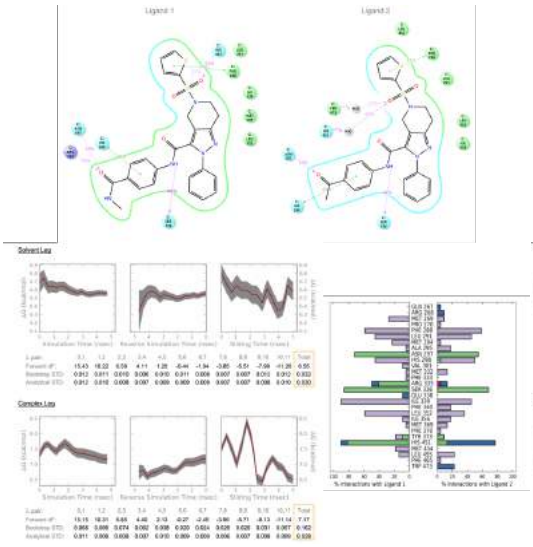
FXR_91 to FXR_93



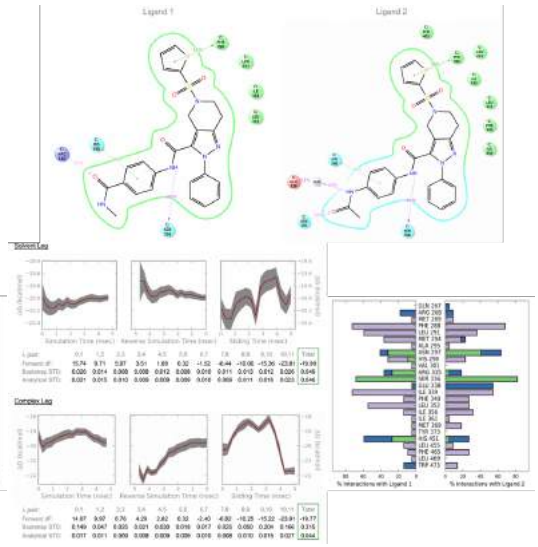
FXR_93 to FXR_47



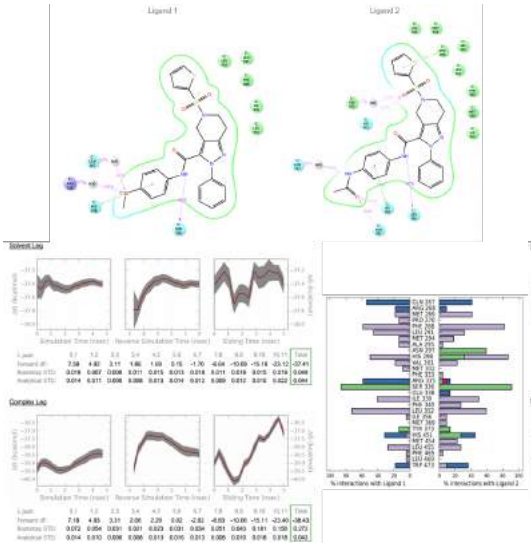
FXR_98 to FXR_49



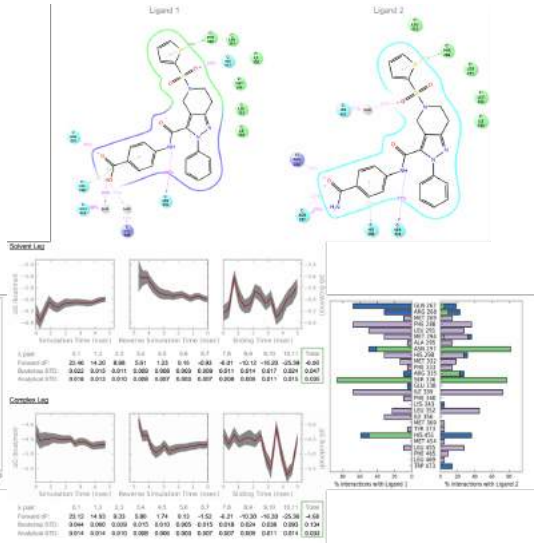
FXR_98 to FXR_95



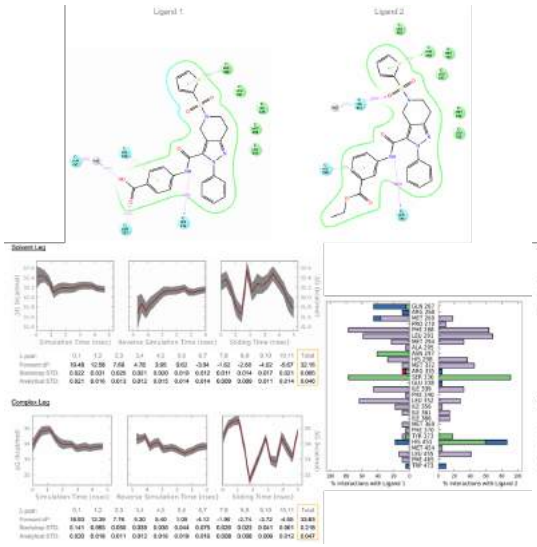
FXR_99 to FXR_95



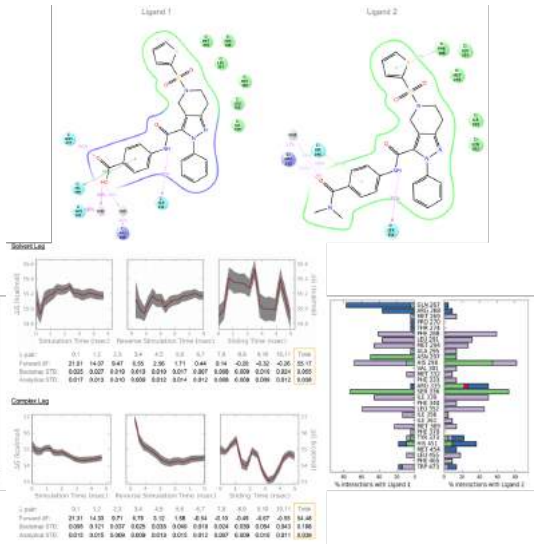
FXR_101 to FXR_46



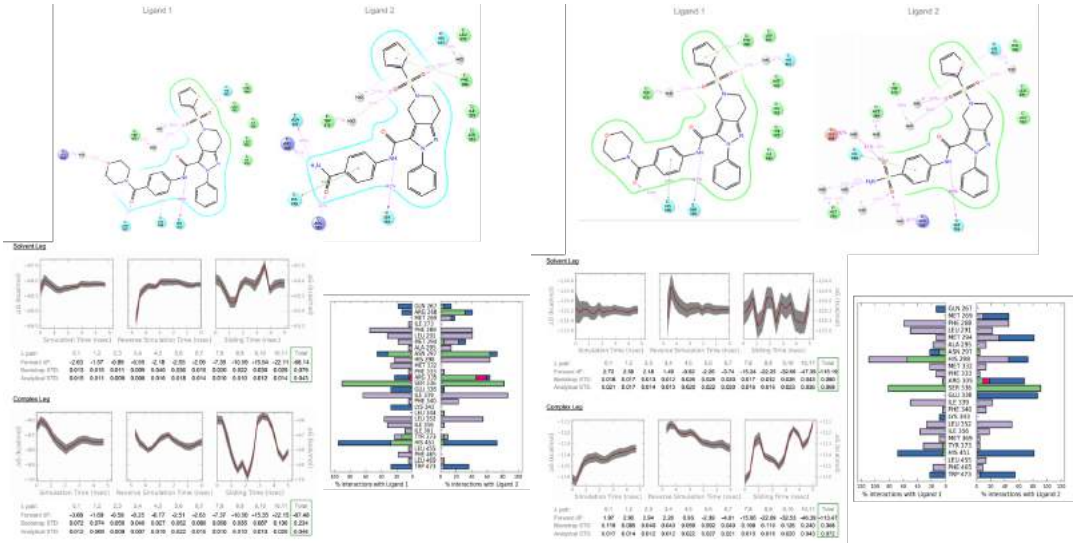
FXR_101 to FXR_47



FXR_101 to FXR_96

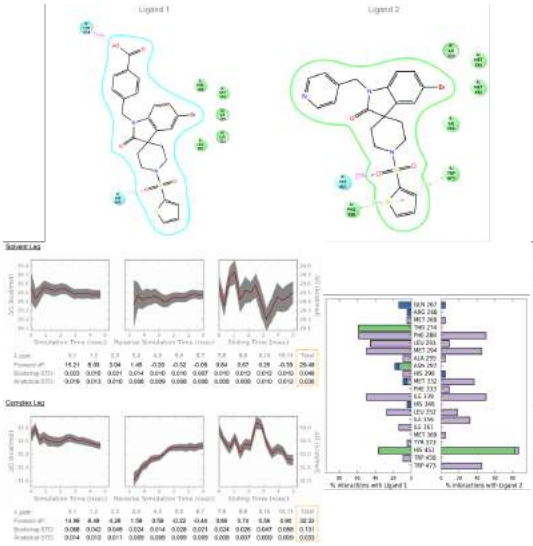


FXR_102 to FXR_46

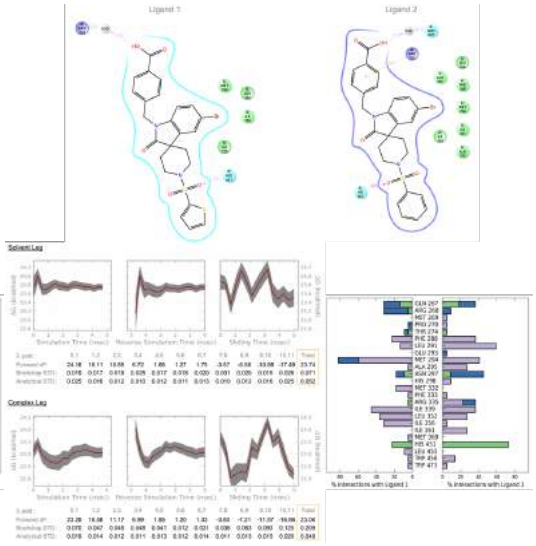


FXR_46 to FXR_17 FXR_102 to FXR_100

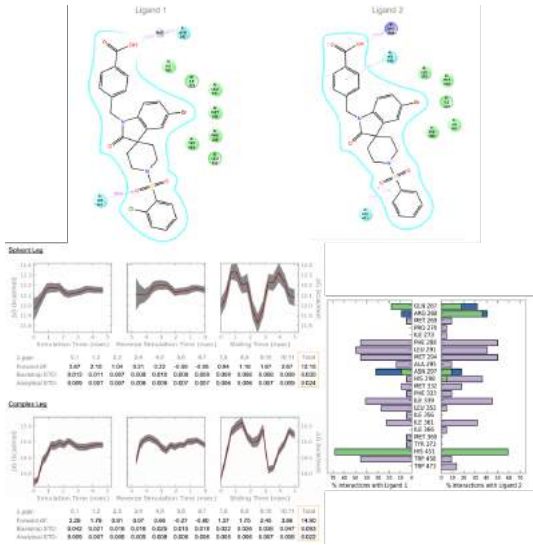
Figure S22. The Free energy convergence and the Protein-Ligand interactions for End-point λ-replicas plots for the sulfonamides analogues.



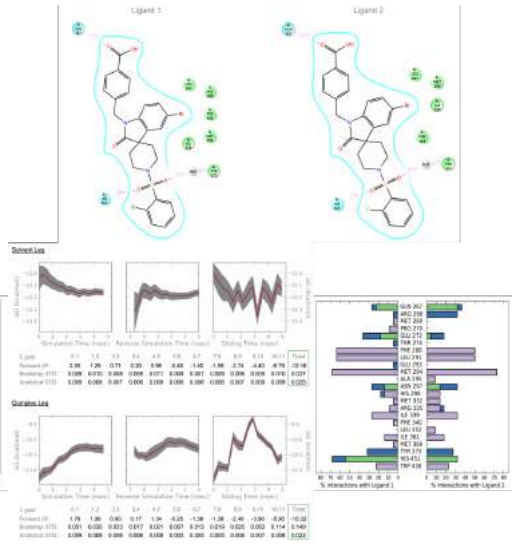
FXR_10 to FXR_75



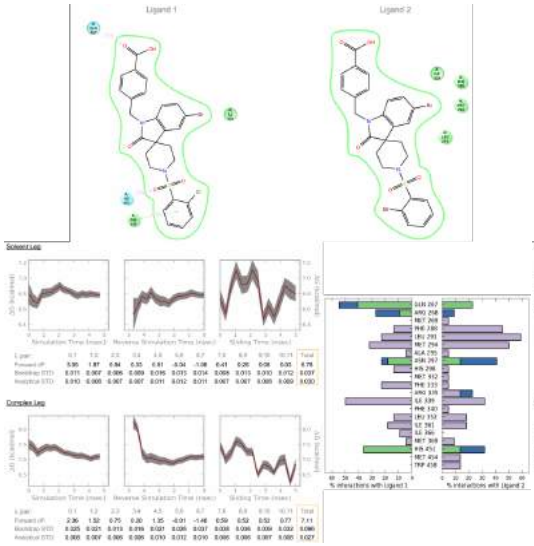
FXR_10 to FXR_76



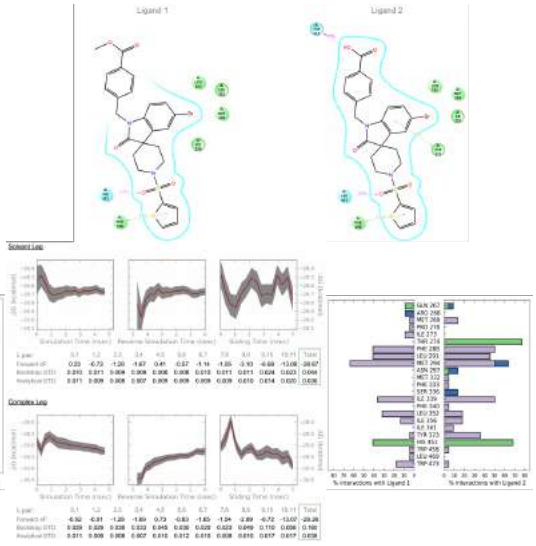
FXR_12AA to FXR_76



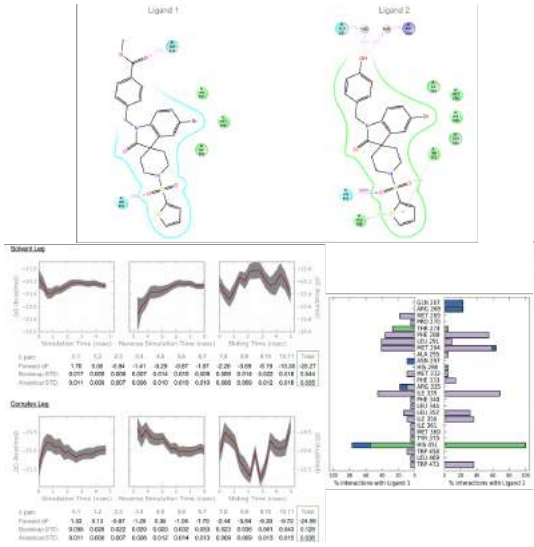
FXR_12AA to FXR_84AA



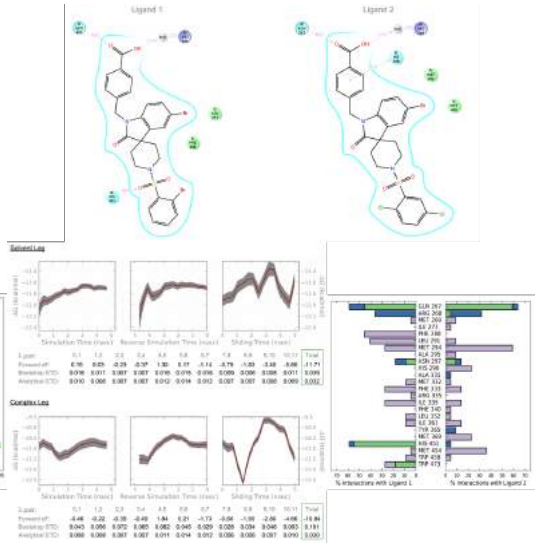
FXR_12AB to FXR_74AA



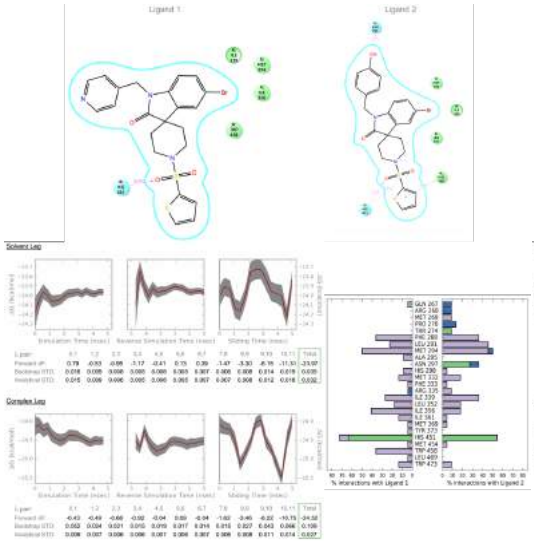
FXR_38 to FXR_10



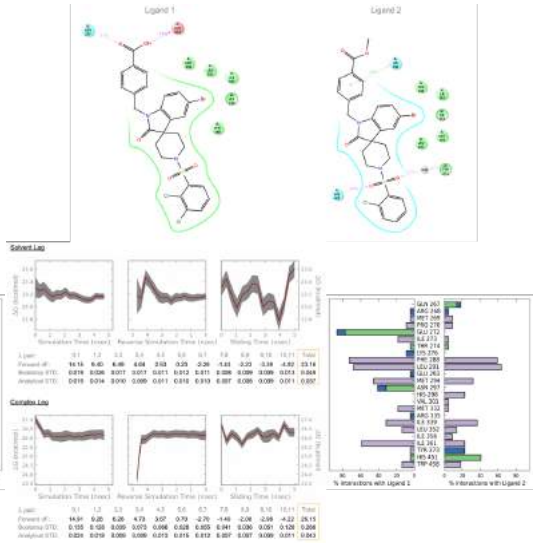
FXR_38 to FXR_73



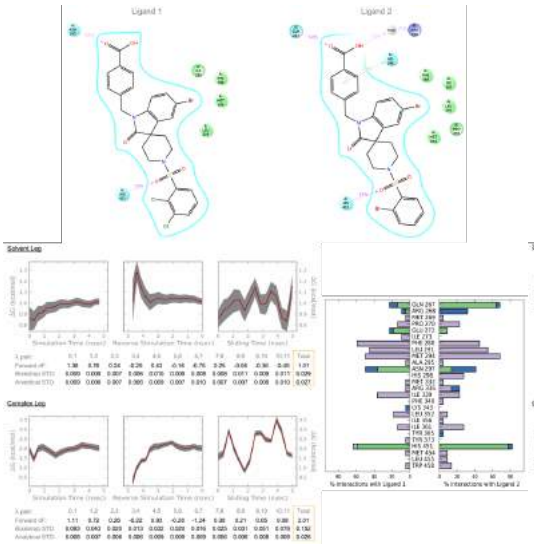
FXR_74AB to FXR_83AA



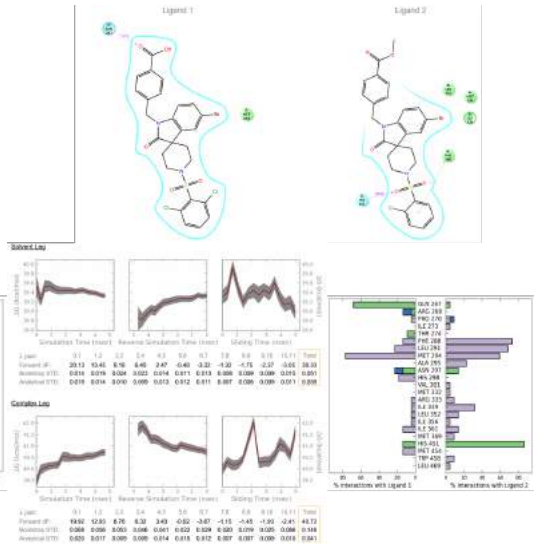
FXR_75 to FXR_73



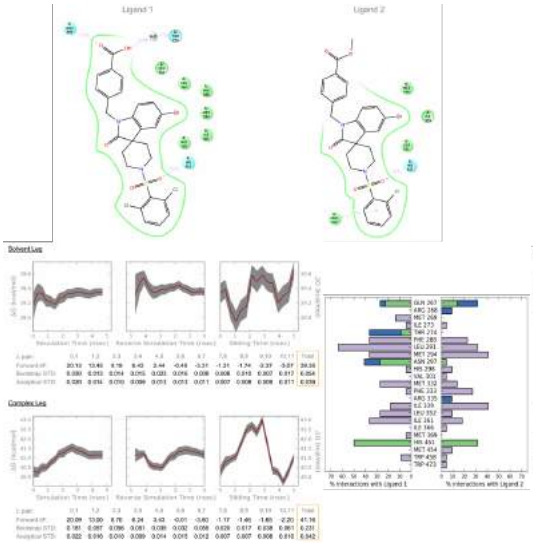
FXR_77AA to FXR_41AA



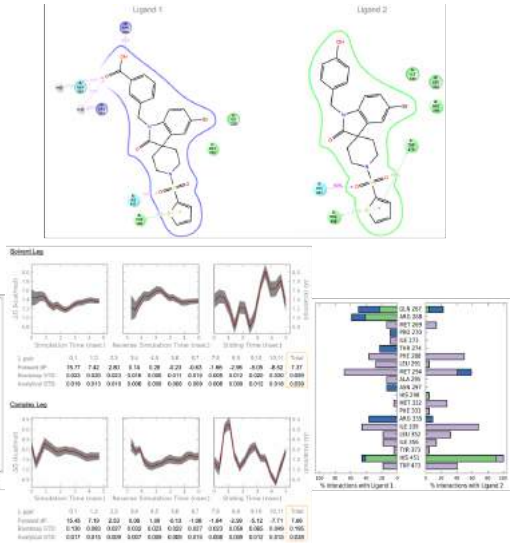
FXR_77AA to FXR_74AA



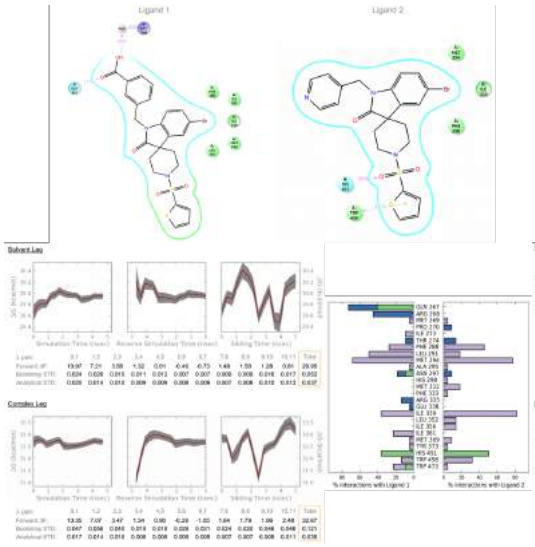
FXR_78 to FXR_41AA



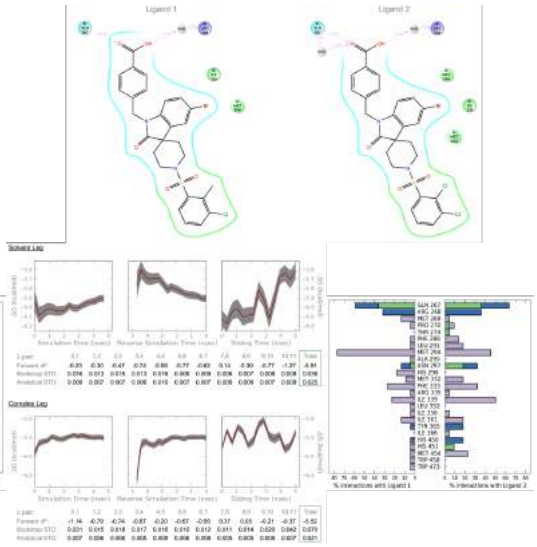
FXR_78 to FXR_41AB



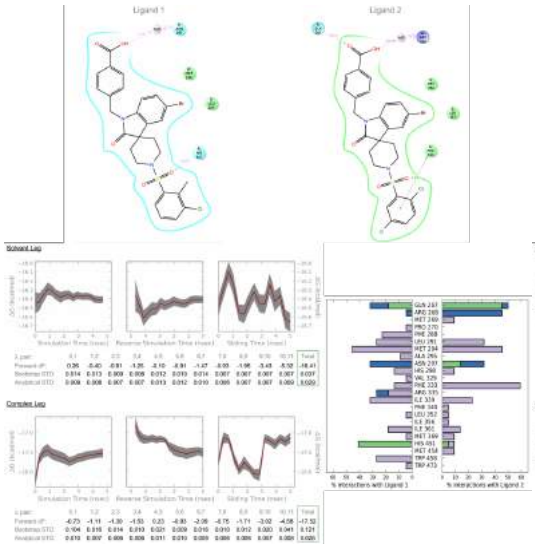
FXR_79 to FXR_73

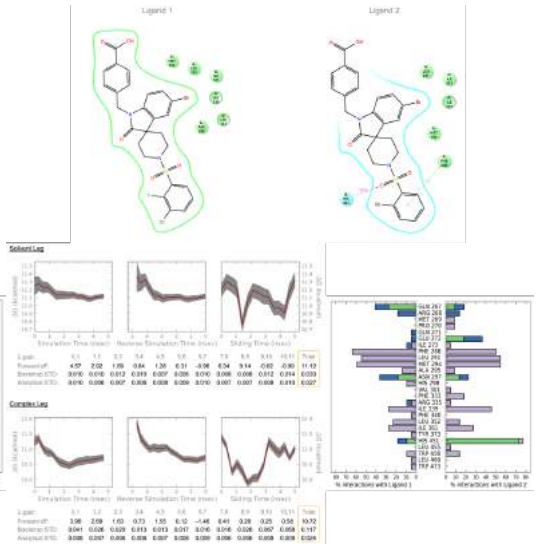


FXR_79 to FXR_75

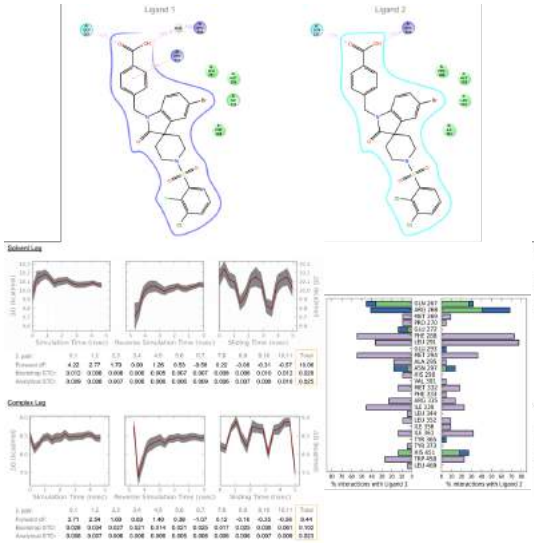


FXR_81AB to FXR_77AB

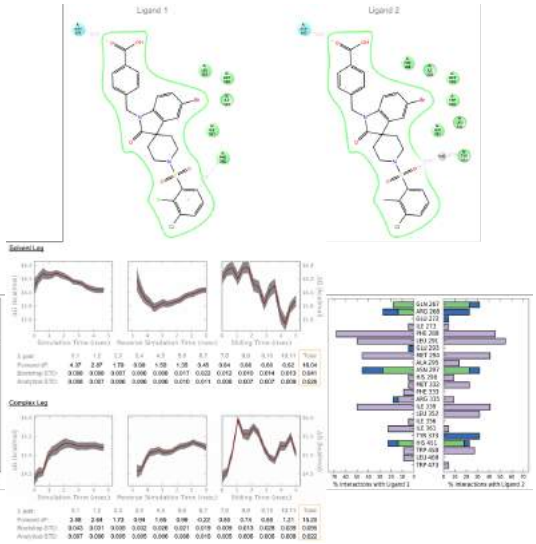




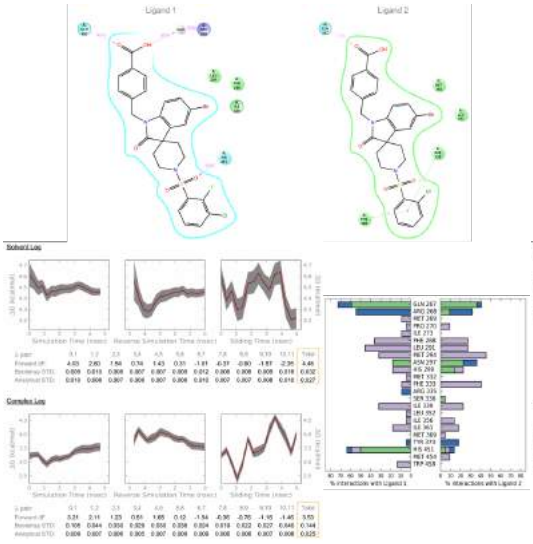
FXR_81AB to FXR_83AB



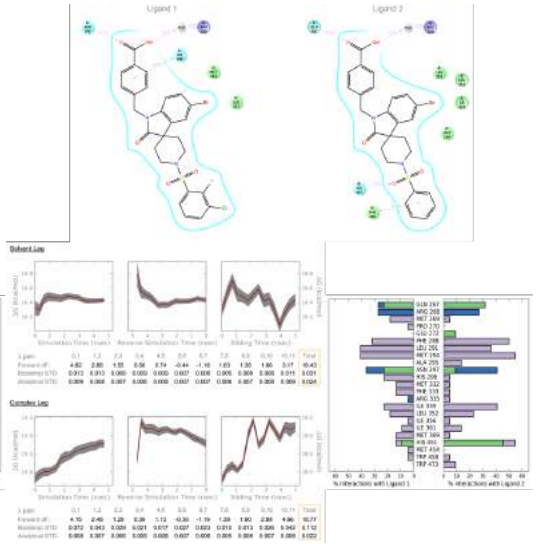
FXR_82AA to FXR_74AA



FXR_82AA to FXR_77AA

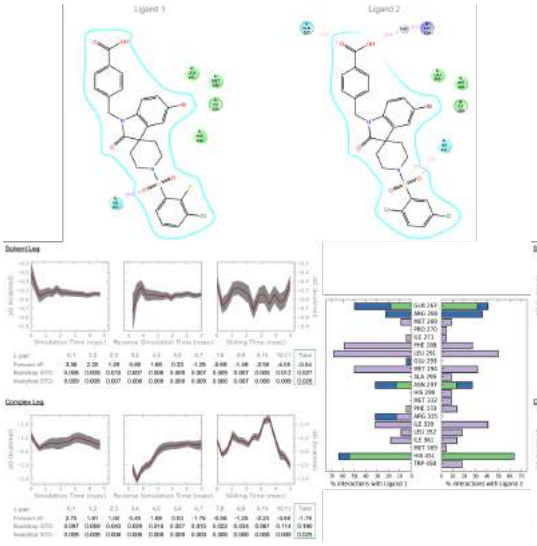


FXR_82AA to FXR_81AA

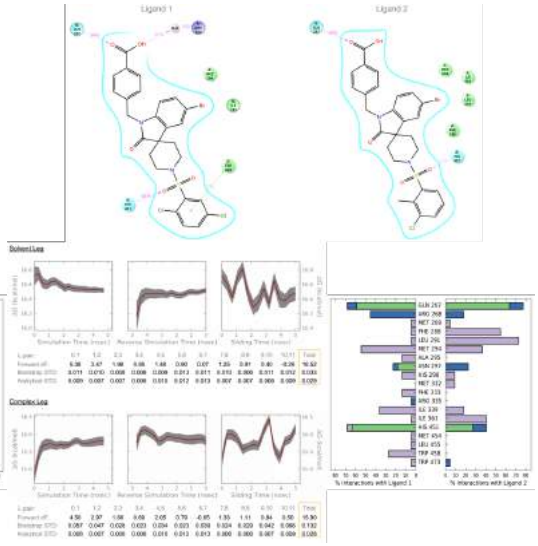


FXR_82AB to FXR_12AB

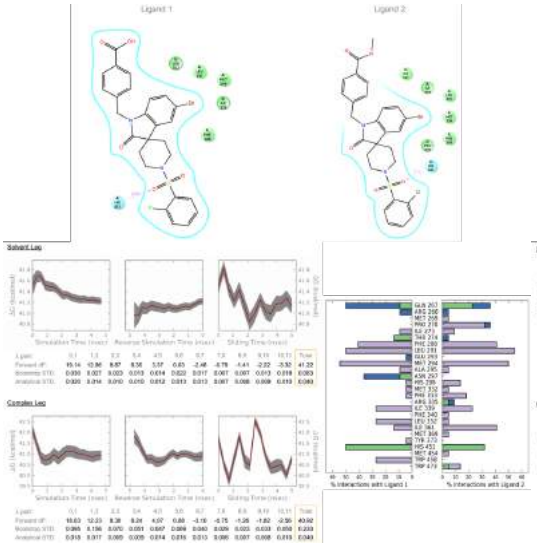
FXR_82AB to FXR_76



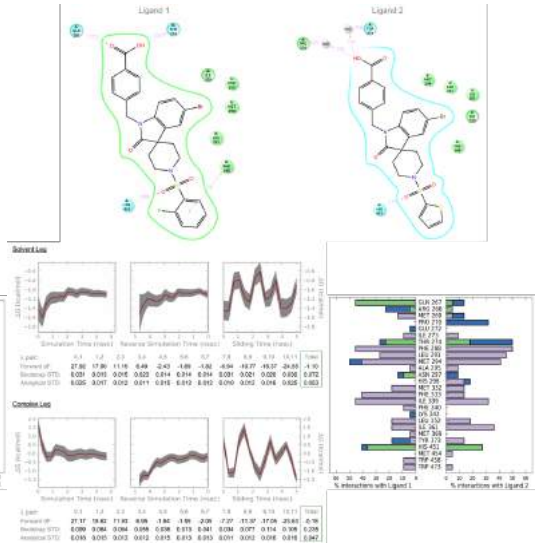
FXR_82AB to FXR_83AA



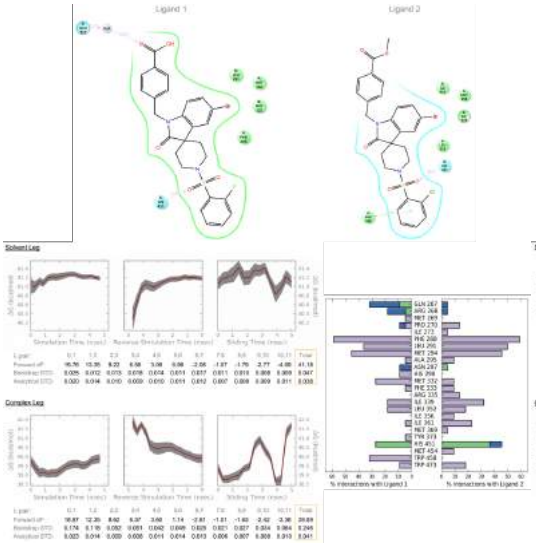
FXR_83AA to FXR_81AA



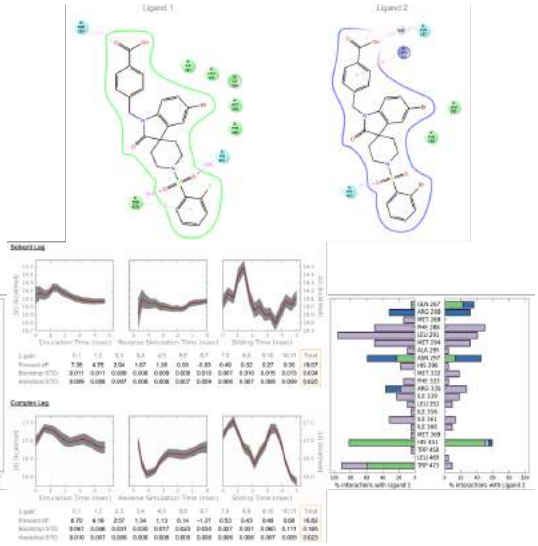
FXR_84AA to FXR_41AB



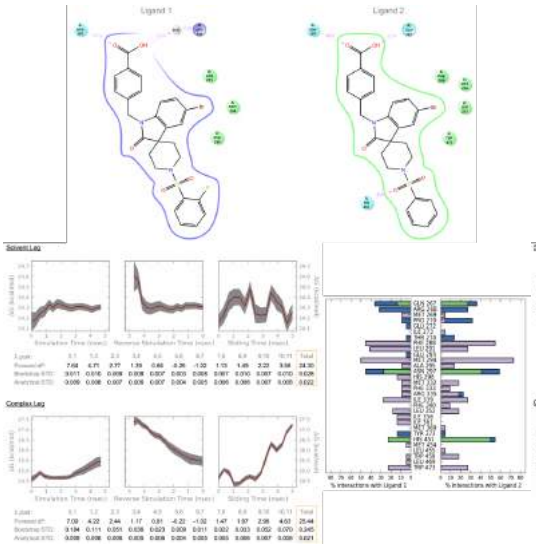
FXR_84AB to FXR_10



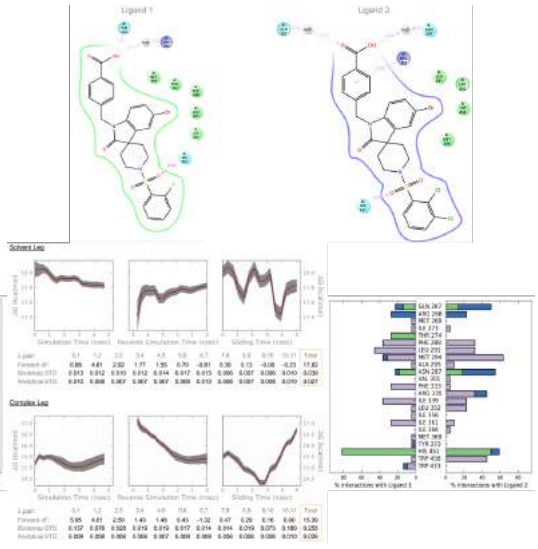
FXR_84AB to FXR_41AB



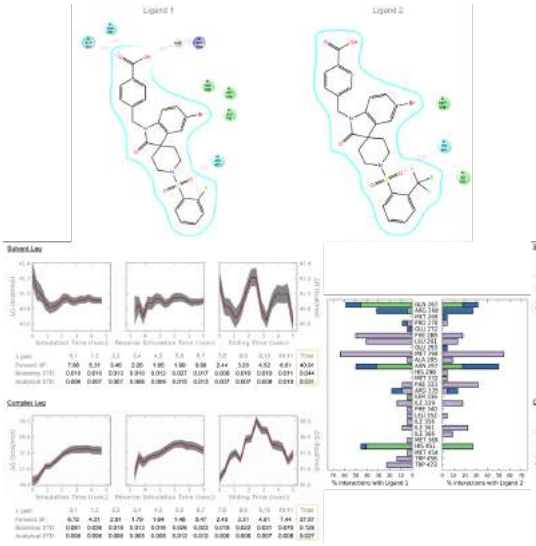
FXR_84AB to FXR_74AB



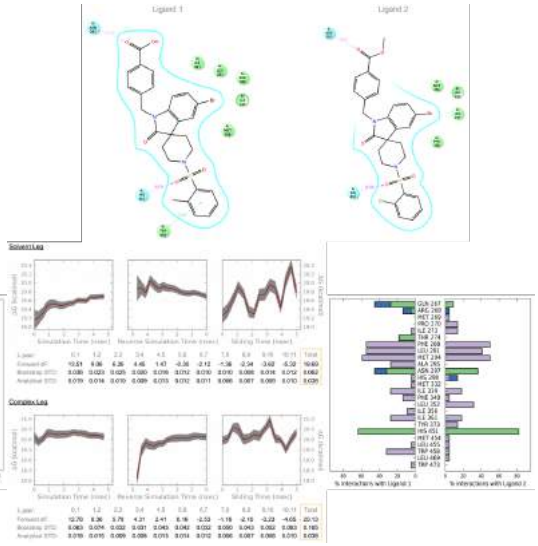
FXR_84AB to FXR_76



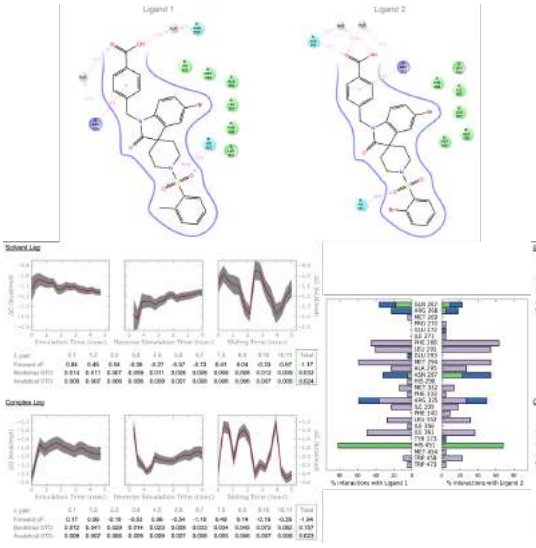
FXR_84AB to FXR_77AB



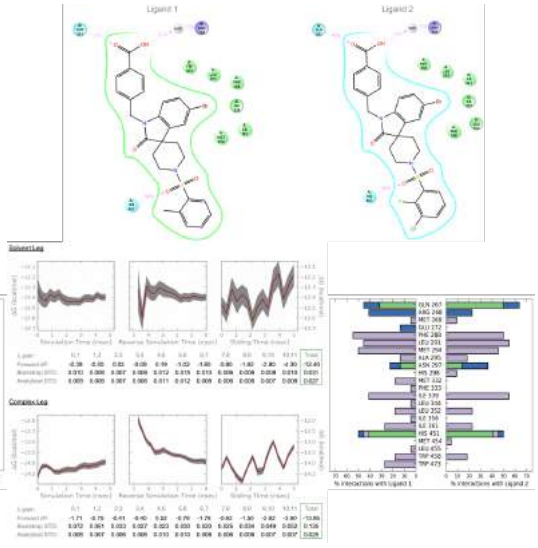
FXR_84AB to FXR_88AB



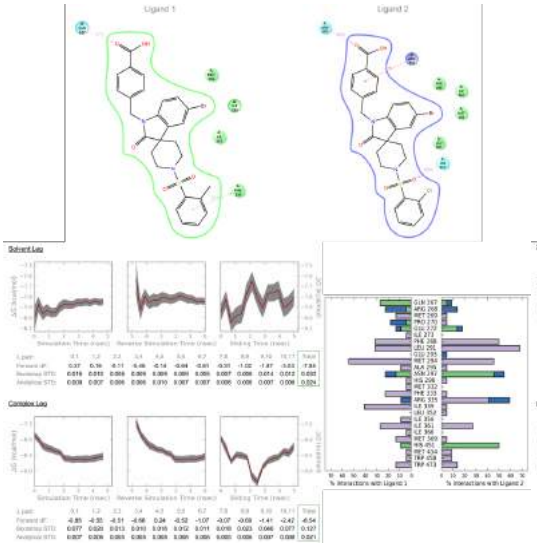
FXR_85AA to FXR_41AA



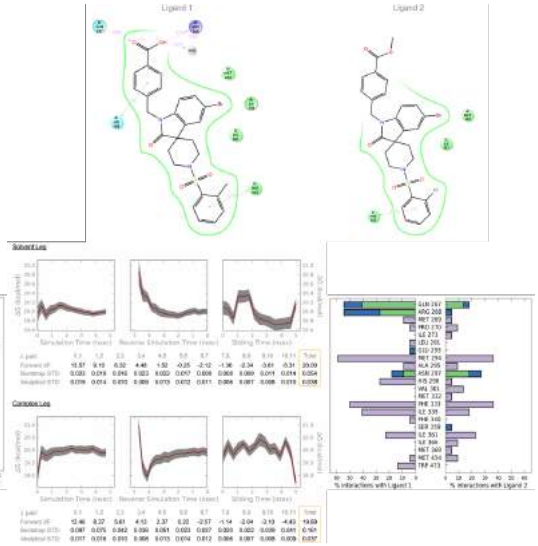
FXR_85AA to FXR_74AA



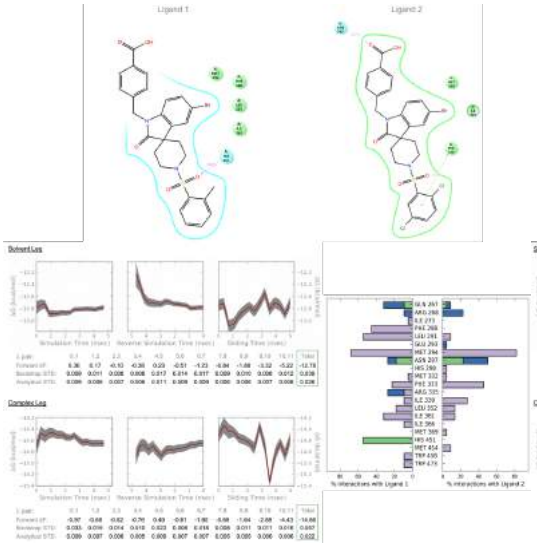
FXR_85AA to FXR_82AA



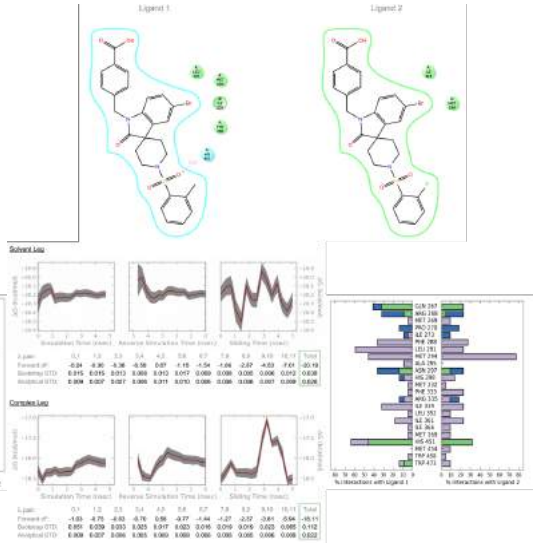
FXR_85AB to FXR_12AB



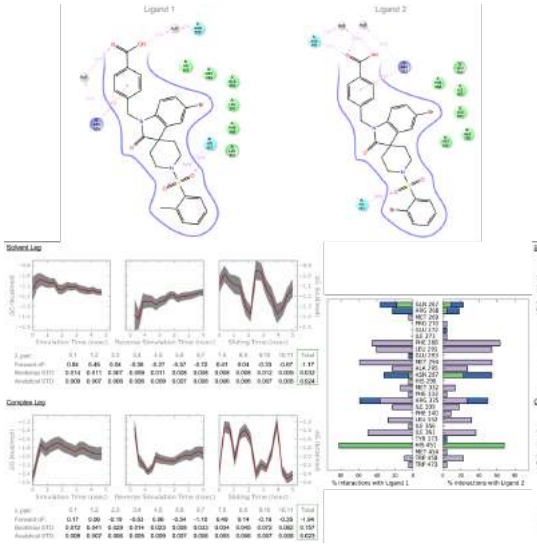
FXR_85AB to FXR_41AB



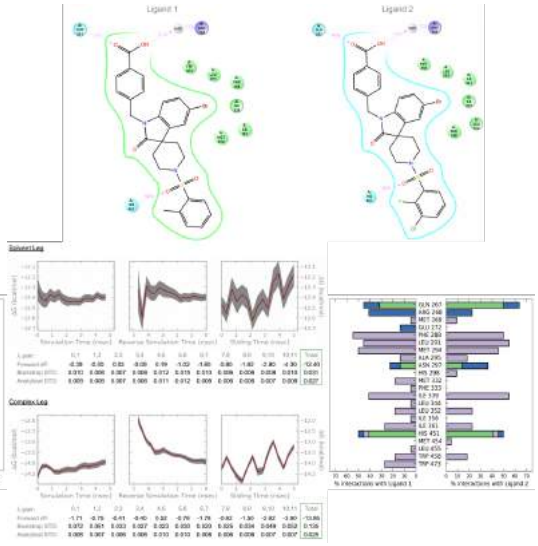
FXR_85AB to FXR_83AB



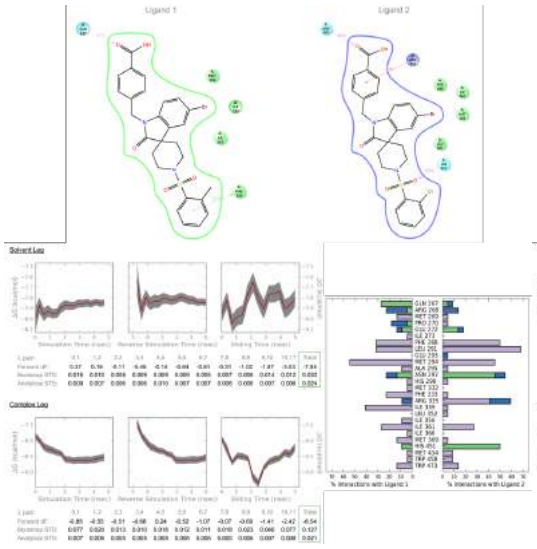
FXR_ 85AB to FXR_84AB



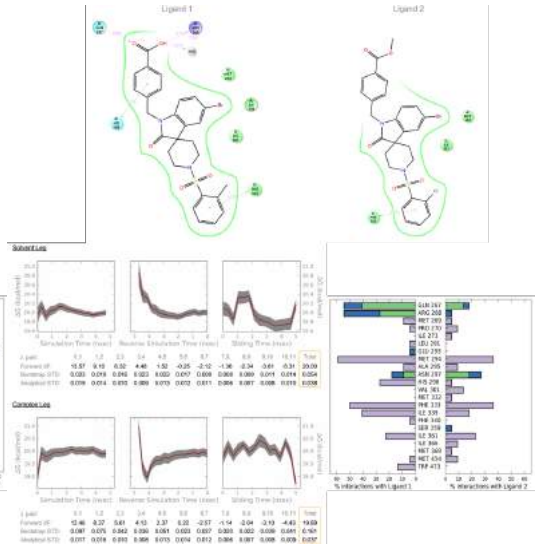
FXR_85AB to FXR_85AA



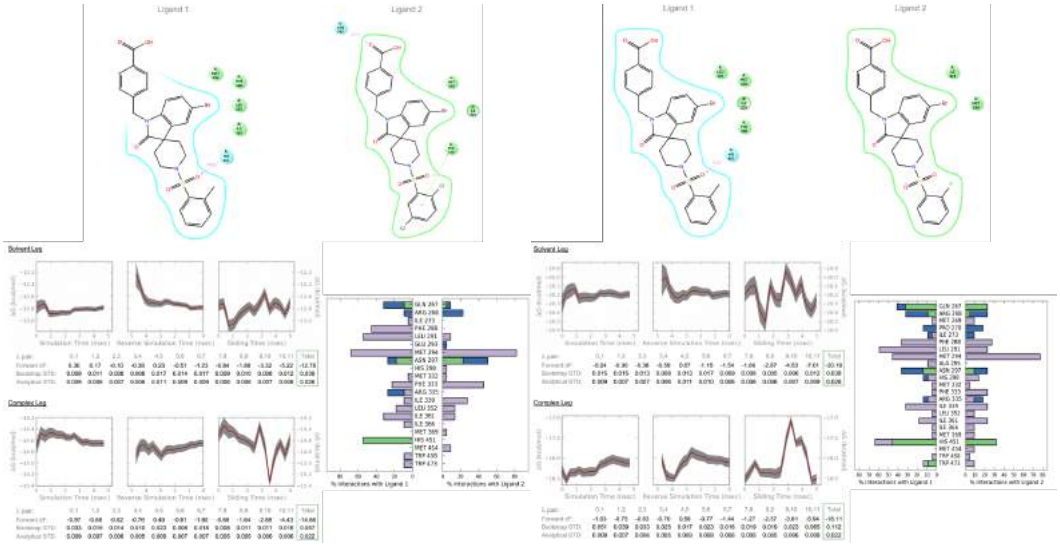
FXR_85AB to FXR_88AB



FXR_88AA to FXR_41AB



FXR_88AA to FXR_85AA



FXR_89 to FXR_77AA FXR_89 to FXR_84AA

Figure S23. The Free energy convergence and the Protein-Ligand interactions for End-point λ-replicas plots for the spiros analogues.

References

1. *Schrödinger Release 2016-2: LigPrep, Schrödinger, LLC, New York, NY, 2016.*
2. Harder, E., et al., *OPLS3: A Force Field Providing Broad Coverage of Drug-like Small Molecules and Proteins.* J Chem Theory Comput, 2016. **12**(1): p. 281-96.
3. *Schrödinger Release 2016-2: Epik, Schrödinger, LLC, New York, NY, 2016.*
4. Greenwood, J.R., et al., *Towards the comprehensive, rapid, and accurate prediction of the favorable tautomeric states of drug-like molecules in aqueous solution.* J Comput Aided Mol Des, 2010. **24**(6-7): p. 591-604.
5. Shelley, J.C., et al., *Epik: a software program for pK(a) prediction and protonation state generation for drug-like molecules.* J Comput Aided Mol Des, 2007. **21**(12): p. 681-91.
6. *Schrödinger Release 2016-2: ConfGen, Schrödinger, LLC, New York, NY, 2016.*
7. Watts, K.S., et al., *ConfGen: a conformational search method for efficient generation of bioactive conformers.* J Chem Inf Model, 2010. **50**(4): p. 534-46.
8. *Schrödinger Release 2016-2: Schrödinger Suite 2016-2 Protein Preparation Wizard; Epik, Schrödinger, LLC, New York, NY, 2016; Impact, Schrödinger, LLC, New York, NY, 2016; Prime, Schrödinger, LLC, New York, NY, 2016.*
9. Sastry, G.M., et al., *Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments.* J Comput Aided Mol Des, 2013. **27**(3): p. 221-34.
10. *Schrödinger Release 2016-2: Prime, Schrödinger, LLC, New York, NY, 2016.*
11. Jacobson, M.P., et al., *On the Role of the Crystal Environment in Determining Protein Side-chain Conformations.* Journal of Molecular Biology, 2002. **320**(3): p. 597-608.
12. Jacobson, M.P., et al., *A hierarchical approach to all-atom protein loop prediction.* Proteins, 2004. **55**(2): p. 351-67.
13. *Schrödinger Release 2016-2: Jaguar, Schrödinger, LLC, New York, NY, 2016.*

14. Bochevarov, A.D., et al., *Jaguar: A high-performance quantum chemistry software program with strengths in life and materials sciences.* International Journal of Quantum Chemistry, 2013. **113**(18): p. 2110-2142.
15. Kim, K. and K.D. Jordan, *Comparison of Density Functional and MP2 Calculations on the Water Monomer and Dimer.* The Journal of Physical Chemistry, 1994. **98**(40): p. 10089-10094.
16. Bennett, C.H., *Efficient estimation of free energy differences from Monte Carlo data.* Journal of Computational Physics, 1976. **22**(2): p. 245-268.
17. Wang, L., et al., *Modeling Local Structural Rearrangements Using FEP/REST: Application to Relative Binding Affinity Predictions of CDK2 Inhibitors.* Journal of Chemical Theory and Computation, 2013. **9**(2): p. 1282-1293.
18. Liu, P., et al., *Replica exchange with solute tempering: a method for sampling biological systems in explicit water.* Proc Natl Acad Sci U S A, 2005. **102**(39): p. 13749-54.
19. Wang, L., B.J. Berne, and R.A. Friesner, *On achieving high accuracy and reliability in the calculation of relative protein-ligand binding affinities.* Proc Natl Acad Sci U S A, 2012. **109**(6): p. 1937-42.
20. Wang, L., R.A. Friesner, and B.J. Berne, *Replica Exchange with Solute Scaling: A More Efficient Version of Replica Exchange with Solute Tempering (REST2).* The Journal of Physical Chemistry B, 2011. **115**(30): p. 9431-9438.
21. Pastor, R.W., B.R. Brooks, and A. Szabo, *An analysis of the accuracy of Langevin and molecular dynamics algorithms.* Molecular Physics, 1988. **65**(6): p. 1409-1419.
22. Berendsen, H.J.C., et al., *Molecular dynamics with coupling to an external bath.* The Journal of Chemical Physics, 1984. **81**(8): p. 3684-3690.