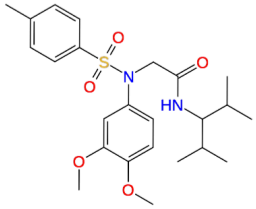
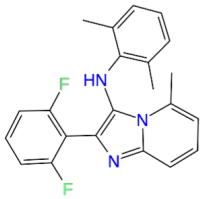
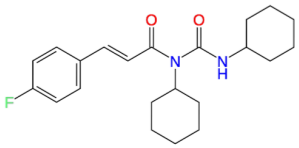
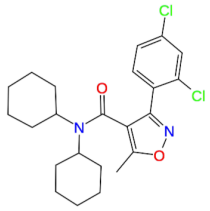
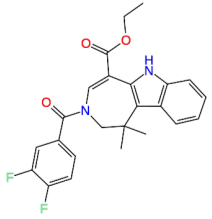
Table S1. Pose prediction dataset, showing each compound's ID, chemical class, the resolution and free R factor of the corresponding co-crystal structures, the Tanimoto similarity coefficient of each ligand with the most similar ligand in any FXR co-crystal structure in the PDB at the start of the challenges, and the PDB IDs of these co-crystal structures. The Tanimoto similarity coefficients were calculated using linear fingerprinting from Canvas, Schrödinger [1, 2].

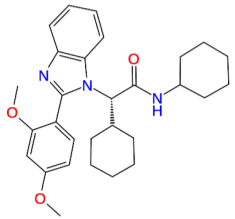
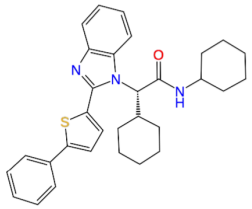
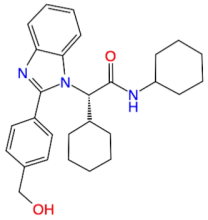
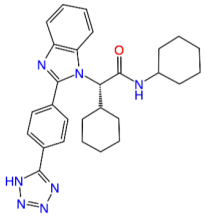
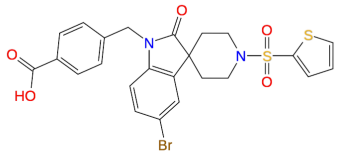
Compound ID	Chemical Class	Resolution (Å)	R_{free}	Maximum Tanimoto Similarity to Ligands in Public FXR Structures	PDB ID of Public Structure with Maximum Tanimoto Similarity
FXR 1	Miscellaneous	1.70	0.226	0.104	4OIV
FXR 2	Miscellaneous	2.00	0.233	0.077	5EIT
FXR 3	Miscellaneous	2.00	0.254	0.069	3OKI
FXR 4	Isoxazole	2.20	0.223	0.131	3DCT
FXR 5	Miscellaneous	2.26	0.258	0.879	3FLI
FXR 6	Benzimidazole	1.98	0.237	0.609	3OKI
FXR 7	Benzimidazole	2.00	0.247	0.587	3OKI
FXR 8	Benzimidazole	2.50	0.263	0.924	3OKI
FXR 9	Benzimidazole	2.20	0.218	0.824	3OKI
FXR 10	Spirocycle	2.10	0.267	0.089	4OIV
FXR 11	Spirocycle	1.84	0.223	0.080	4OIV
FXR 12	Spirocycle	1.90	0.254	0.094	4OIV
FXR 13	Benzimidazole	1.90	0.222	0.858	3OKI
FXR 14	Benzimidazole	2.00	0.235	0.444	3OKI
FXR 15	Tetrahydropyrro-(azo)lopyridines	1.90	0.244	0.136	3DY6
FXR 16	Tetrahydropyrro-(azo)lopyridines	2.10	0.259	0.128	3DY6
FXR 17	Tetrahydropyrro-(azo)lopyridines	2.60	0.295	0.110	3DY6
FXR 18	Miscellaneous	2.20	0.243	0.130	3OKI
FXR 19	Benzimidazole	1.90	0.226	0.520	3OOF
FXR 20	Benzimidazole	2.00	0.284	0.417	3OKH

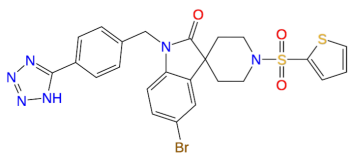
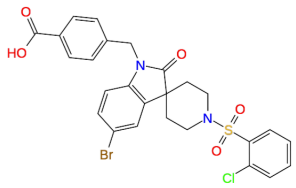
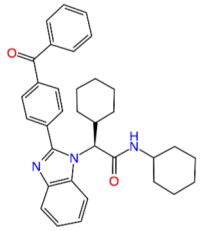
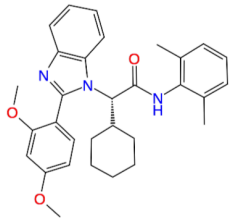
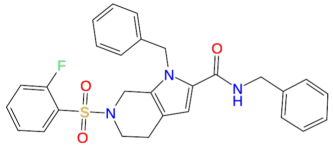
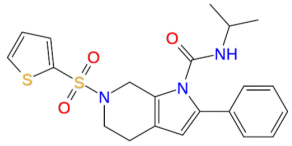
FXR 21	Benzimidazole	1.90	0.244	0.840	3OOF
FXR 22	Benzimidazole	2.00	0.255	0.359	3OOF
FXR 23	Isoxazole	1.95	0.225	0.191	3DCU
FXR 24	Benzimidazole	1.90	0.238	0.943	3OMK
FXR 25	Benzimidazole	2.50	0.26	0.913	3OMK
FXR 26	Benzimidazole	2.20	0.22	0.436	3OOF
FXR 27	Benzimidazole	2.40	0.226	0.939	3OMM
FXR 28	Benzimidazole	1.84	0.247	0.518	3OMM
FXR 29	Benzimidazole	1.80	0.254	0.575	3OOF
FXR 30	Benzimidazole	2.30	0.245	0.300	3OKI
FXR 31	Benzimidazole	2.30	0.264	0.448	3OOF
FXR 32	Benzimidazole	1.90	0.24	0.355	3OOF
FXR 33	Isoxazole	2.30	0.233	0.819	3FXV
FXR 34	Miscellaneous	2.20	0.262	0.830	3W5P
FXR 35	Benzimidazole	1.90	0.218	0.690	3OOF
FXR 36	Benzimidazole	2.30	0.236	0.845	3OOF
Apo		1.80	0.271		

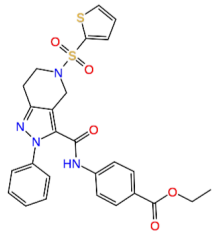
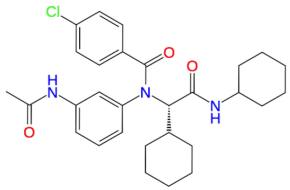
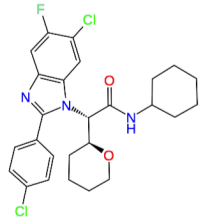
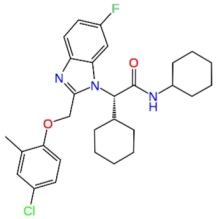
Table S2. FXR potency dataset showing each compound's ID, SMILES string, 2D chemical structure generated in Maestro (Schrödinger Release 2017-2: MS Jaguar, Schrödinger, LLC, New York, NY, 2017), average experimental IC50 as outlined in the "Methods" section, replicate IC50s, information on the stereochemical composition of the compound in the assay, and chemical class. The IC50s were measured with a scintillation proximity assay as described in Methods.

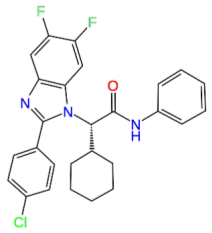
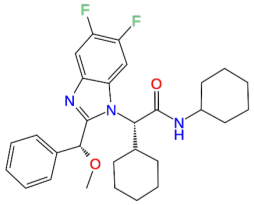
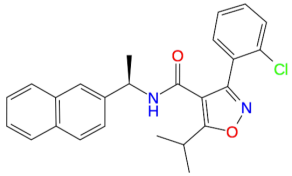
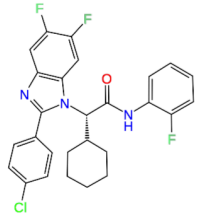
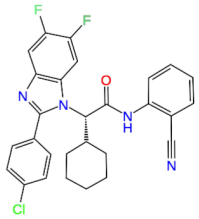
ID	SMILES	Ligand Structure	Average IC50 [μM]	IC50s Standard Deviations [μM]	IC50s [μM]	Stereo	Chemical Class

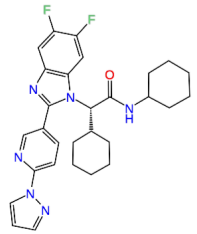
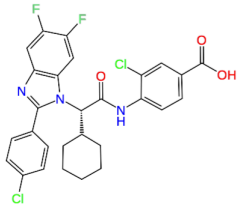
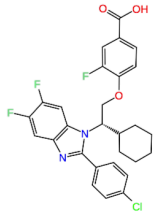
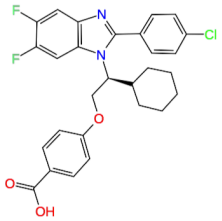
FXR 1	COc1ccc(cc1OC)N(CC(=O)NC(C(C)C)C(C)C)S(=O)(=O)c2ccc(C)cc2		56.1	1.68	26.9, 46, 95.4, >100	Miscellaneous
FXR 2	Cc1cccc(C)c1Nc2c(nc3cccc(C)n23)c4c(F)cccc4F		4.64	2.93	8.31, 0.971	Miscellaneous
FXR 3	Fc1ccc(\C=C\C(=O)N(C2CCCCC2)C(=O)NC3CCCCC3)cc1		263	1.00	263, >100, >100	Miscellaneous
FXR 4	Cc1onc(c2ccc(Cl)cc2Cl)c1C(=O)N(C3CCCCC3)C4CCCCC4		7.19	4.05	9.86, 1.05, 0.66, 17.2	Isoxazole
FXR 5	CCOC(=O)C1=CN(CC(C)(C)c2c1[nH]c3ccccc23)C(=O)c4ccc(F)c(F)c4		1.42	10.01	0.025, 0.017, 2.17, 1.25, 3.65	Miscellaneous

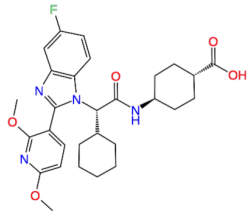
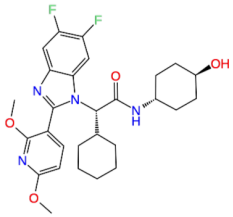
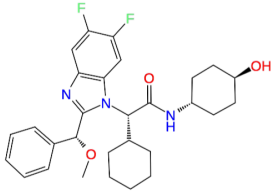
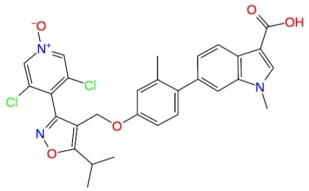
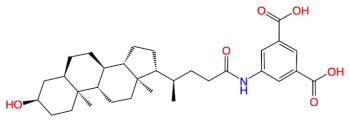
FXR 6	COc1ccc(c(OC)c1)c 2nc3ccccc3n2C(C4C CCCC4)C(=O)NC5C CCCC5		0.048	1.50	0.058, 0.056, 0.055, 0.022	racemate	Benzimidazole
FXR 7	O=C(NC1CCCCC1) C(C2CCCCC2)n3c(n c4ccccc34)c5ccc(s5) c6ccccc6		0.239	1.08	0.258, 0.22	racemate	Benzimidazole
FXR 8	OCc1ccc(cc1)c2nc3 ccccc3n2C(C4CCCC C4)C(=O)NC5CCCC C5		4.36	1.31	5.51, 3.21	racemate	Benzimidazole
FXR 9	O=C(NC1CCCCC1) C(C2CCCCC2)n3c(n c4ccccc34)c5ccc(cc 5)c6nnn[nH]6		7.95	1.08	7.33, 8.57	racemate	Benzimidazole
FXR 10	OC(=O)c1ccc(CN2C (=O)C3(CCN(CC3)S (=O)(=O)c4cccs4)c5 cc(Br)ccc25)cc1		5.64	1.88	2.07, 9.33, 3.13, 8.03		Spirocycle

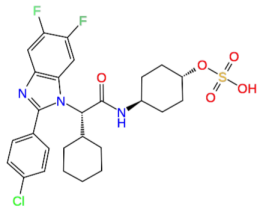
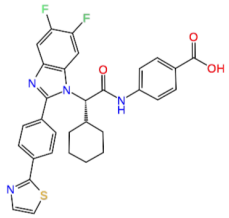
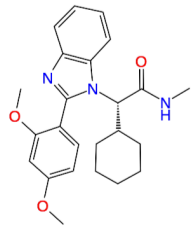
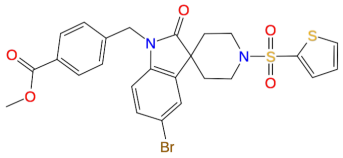
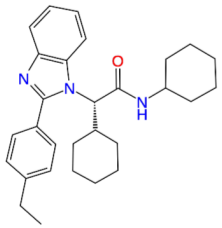
FXR 11	Brc1ccc2N(Cc3ccc(c3)c4nnn[nH]4)C(=O)C5(CCN(CC5)S(=O)(=O)c6cccs6)c2c1		0.12	1.00	0.12		Spirocycle
FXR 12	OC(=O)c1ccc(CN2C(=O)C3(CCN(CC3)S(=O)(=O)c4ccccc4Cl)c5cc(Br)ccc25)cc1		0.058	1.00	0.058		Spirocycle
FXR 13	O=C(NC1CCCCC1)C(C2CCCCC2)n3c(nc4ccccc34)c5ccc(cc5)C(=O)c6ccccc6		0.0244	1.00	0.0244	racemate	Benzimidazole
FXR 14	COc1ccc(c(OC)c1)c2nc3ccccc3n2[C@@H](C4CCCCC4)C(=O)Nc5c(C)cccc5C		0.0096	1.17	0.011, 0.0081		Benzimidazole
FXR 15	Fc1ccccc1S(=O)(=O)N2CCc3cc(C(=O)NCc4ccccc4)n(Cc5ccccc5)c3C2		27	1.01	26.7, 27.4		Tetrahydropyrro-(azo)lopyridines
FXR 16	CC(C)NC(=O)n1c2CN(CCc2cc1c3ccccc3)S(=O)(=O)c4cccs4		29	1.26	35.5, 22.5		Tetrahydropyrro-(azo)lopyridines

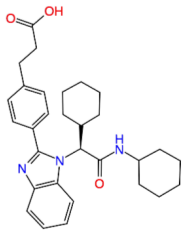
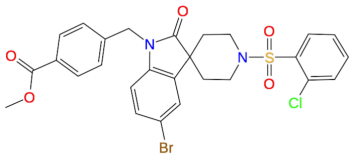
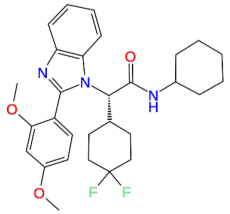
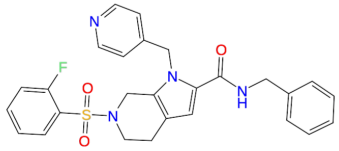
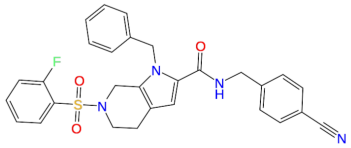
FXR 17	CCOC(=O)c1ccc(NC(=O)c2c3CN(CCc3nn2c4ccccc4)S(=O)(=O)c5cccs5)cc1		0.785	2.28	0.77, 1.85, 0.26, 0.26		Tetrahydropyrro-(azo)lopyridines
FXR 18	CC(=O)Nc1cccc(c1)N(C(C2CCCCC2)C(=O)NC3CCCCC3)C(=O)c4ccc(Cl)cc4		6.28	1.08	6.73, 5.82	racemate	Miscellaneous
FXR 19	[H][C@]1(CCCCO1)[C@@]([H])(C(=O)NC2CCCCC2)n3c4cc(c(cc4nc3c5ccc(cc5)Cl)F)Cl		0.091	1.08	0.098, 0.084		Benzimidazole
FXR 20	Cc1cc(Cl)ccc1OCc2nc3ccc(F)cc3n2C(C4CCCCC4)C(=O)NC5CCCCC5		0.0135	1.12	0.015, 0.012	racemate	Benzimidazole

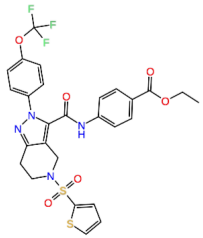
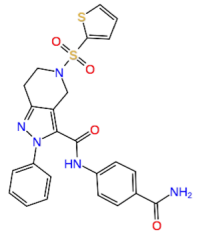
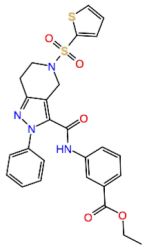
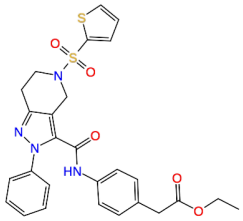
FXR 21	Fc1cc2nc(c3ccc(Cl)cc3)n([C@@H](C4CCCCC4)C(=O)Nc5ccccc5)c2cc1F		0.76	1.10	0.69, 0.83		Benzimidazole
FXR 22	COC(c1ccccc1)c2nc3cc(F)c(F)cc3n2C(C4CCCCC4)C(=O)NC5CCCCC5		0.0089	1.44	0.0058, 0.012	Diastereomer racemic mix	Benzimidazole
FXR 23	CC(C)c1onc(c2ccccc2Cl)c1C(=O)N[C@H](C)c3ccc4ccccc4c3		0.57	1.07	0.53, 0.61		Isoxazole
FXR 24	Fc1ccccc1NC(=O)[C@H](C2CCCCC2)n3c(nc4cc(F)c(F)cc34)c5ccc(Cl)cc5		0.0205	1.19	0.024, 0.017		Benzimidazole
FXR 25	Fc1cc2nc(c3ccc(Cl)cc3)n(C(C4CCCCC4)C(=O)Nc5ccccc5C#N)c2cc1F		0.062	1.12	0.069, 0.055	racemate	Benzimidazole

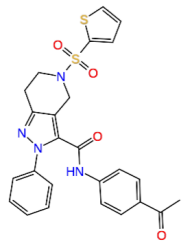
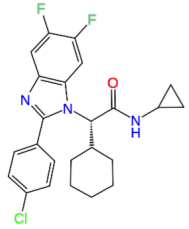
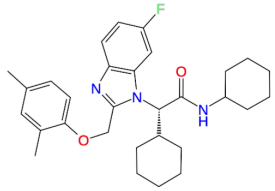
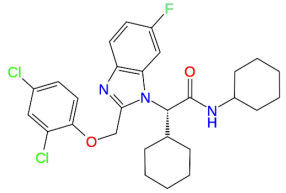
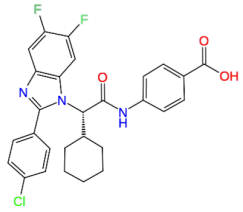
FXR 26	Fc1cc2nc(c3ccc(nc3)n4cccn4)n([C@@H](C5CCCCC5)C(=O)NC6CCCCC6)c2cc1F		1.87	1.12	2.08, 1.65	Benzimidazole
FXR 27	OC(=O)c1ccc(NC(=O)[C@H](C2CCCCC2)n3c(nc4cc(F)c(F)cc34)c5ccc(Cl)cc5)c(Cl)c1		0.0087	1.23	0.0082, 0.0098, 0.0063, 0.0079, 0.0073, 0.0086, 0.0069, 0.01, 0.013	Benzimidazole
FXR 28	[H][C@@](COc1ccc(cc1F)C(=O)O)(C2CCCCC2)n3c4cc(c(cc4nc3c5ccc(cc5)Cl)F)F		0.021	1.15	0.024, 0.018	Benzimidazole
FXR 29	[H][C@@](COc1ccc(cc1)C(=O)O)(C2CCCCC2)n3c4cc(c(cc4nc3c5ccc(cc5)Cl)F)F		0.039	1.05	0.037, 0.041	Benzimidazole

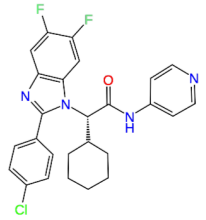
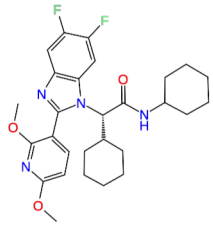
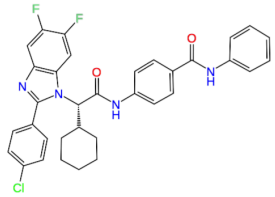
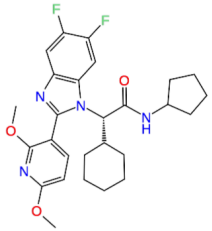
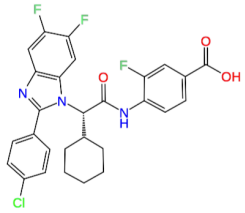
FXR 30	COc1ccc(c(OC)n1)c2nc3cc(F)ccc3n2[C@@H](C4CCCCC4)C(=O)N[C@@H]5CC[C@H](CC5)C(=O)O		0.0725	1.36	0.047, 0.057, 0.1, 0.086		Benzimidazole
FXR 31	COc1nc(c(cc1)c2n(c3c(n2)cc(c(c3)F)F)[C@H](C(=O)N[C@H]4CC[C@@H](CC4)O)C5CCCCC5)OC		0.0084	1.01	0.0084, 0.0083	epimer mix	Benzimidazole
FXR 32	CO[C@H](c1ccccc1)c2nc3cc(F)c(F)cc3n2C(C4CCCCC4)C(=O)N[C@@H]5CC[C@@H](O)CC5		0.0675	1.13	0.059, 0.076		Benzimidazole
FXR 33	CC(C)c1onc(c1COc2ccc(c(C)c2)c3ccc4c(cn(C)c4c3)C(=O)O)c5c(Cl)c[n+]([O-])cc5Cl		0.045	1.17	0.038, 0.052		Isoxazole
FXR 34	C[C@H](CCC(=O)Nc1cc(cc(c1)C(=O)O)C(=O)O)[C@H]2CC[C@H]3[C@@H]4CC[C@@H]5C[C@H](O)CC[C@]5(C)[C@H]4CC[C@]23C		0.0079	1.11	0.0091, 0.0071, 0.0074		Miscellaneous

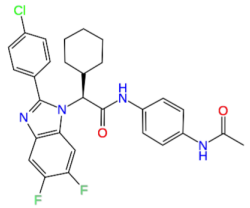
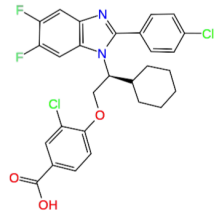
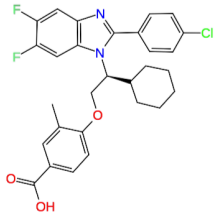
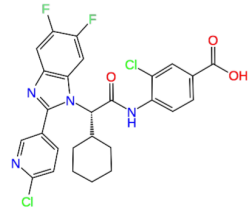
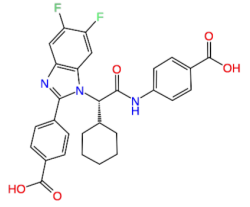
FXR 35	OS(=O)(=O)O[C@@H]1CC[C@H](CC1)NC(=O)C(C2CCCCC2)n3c(nc4cc(F)c(F)cc34)c5ccc(Cl)cc5		0.175	1.15	0.15, 0.2	racemate	Benzimidazole
FXR 36	OC(=O)c1ccc(NC(=O)C(C2CCCCC2)n3c(nc4cc(F)c(F)cc34)c5ccc(cc5)c6nccs6)cc1		0.0487	1.24	0.036, 0.061, 0.049	racemate	Benzimidazole
FXR 37	CNC(=O)C(C1CCCCC1)n2c(nc3ccccc23)c4ccc(OC)cc4OC		25	-	>25, >25	racemate	Benzimidazole
FXR 38	COC(=O)c1ccc(CN2C(=O)C3(CCN(CC3)S(=O)(=O)c4cccs4)c5cc(Br)ccc25)cc1		100	-	>100		Spirocycle
FXR 39	CCc1ccc(cc1)c2nc3ccccc3n2C(C4CCCCC4)C(=O)NC5CCCCC5		0.114	1.32	0.145, 0.083	racemate	Benzimidazole

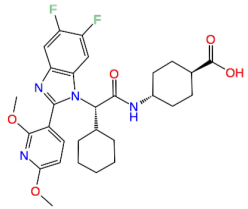
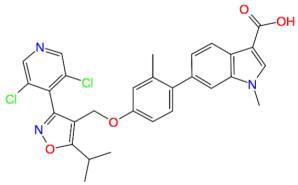
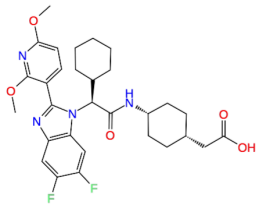
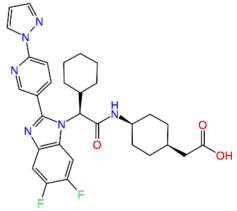
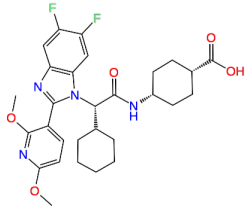
FXR 40	OC(=O)CCc1ccc(cc1)c2nc3ccccc3n2C(C4CCCCC4)C(=O)NC5CCCCC5		9.12	1.58	13, 5.22	racemate	Benzimidazole
FXR 41	COC(=O)c1ccc(CN2C(=O)C3(CCN(CC3)S(=O)(=O)c4ccccc4Cl)c5cc(Br)ccc25)cc1		100	-	>100		Spirocycle
FXR 42	COc1ccc(c(OC)c1)c2nc3ccccc3n2C(C4CCC(F)(F)CC4)C(=O)NC5CCCCC5		0.045	-	0.045	racemate	Benzimidazole
FXR 43	Fc1ccccc1S(=O)(=O)N2CCc3cc(C(=O)NCc4ccccc4)n(Cc5ccnccc5)c3C2		100	-	>100		Tetrahydropyrro-(azo)lopyridines
FXR 44	Fc1ccccc1S(=O)(=O)N2CCc3cc(C(=O)NCc4ccc(cc4)C#N)n(Cc5ccccc5)c3C2		25	-	>25		Tetrahydropyrro-(azo)lopyridines

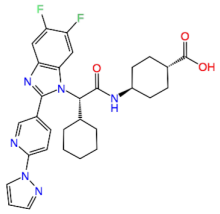
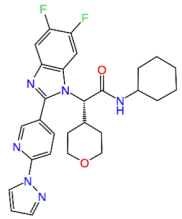
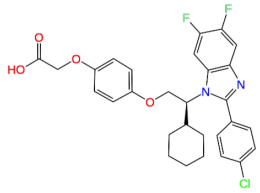
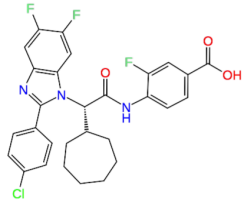
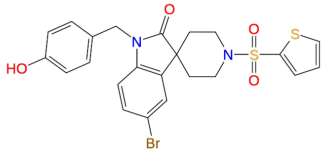
FXR 45	CCOC(=O)c1ccc(NC(=O)c2c3CN(CCc3nn2c4ccc(OC(F)(F)F)cc4)S(=O)(=O)c5cccs5)cc1		28.9	1.02	28.2, 29.5	Tetrahydropyrro-(azo)lopyridines
FXR 46	NC(=O)c1ccc(NC(=O)c2c3CN(CCc3nn2c4ccccc4)S(=O)(=O)c5cccs5)cc1		62.4	1.13	69.8, 54.9	Tetrahydropyrro-(azo)lopyridines
FXR 47	CCOC(=O)c1cccc(NC(=O)c2c3CN(CCc3nn2c4ccccc4)S(=O)(=O)c5cccs5)c1		21	1.14	23.7, 18.2	Tetrahydropyrro-(azo)lopyridines
FXR 48	CCOC(=O)Cc1ccc(NC(=O)c2c3CN(CCc3nn2c4ccccc4)S(=O)(=O)c5cccs5)cc1		100	-	>100	Tetrahydropyrro-(azo)lopyridines

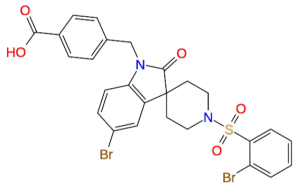
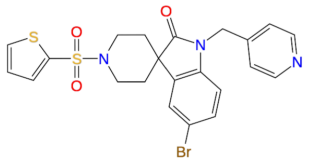
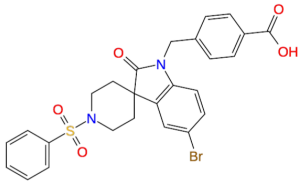
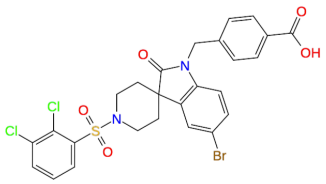
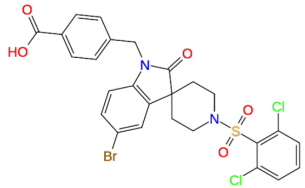
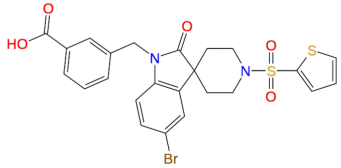
FXR 49	CC(=O)c1ccc(NC(=O)c2c3CN(CCc3nn2c4ccccc4)S(=O)(=O)c5cccs5)cc1		100	-	>100, >100		Tetrahydropyrro-(azo)lopyridines
FXR 50	Fc1cc2nc(c3ccc(Cl)cc3)n(C(C4CCCCC4)C(=O)NC5CC5)c2cc1F		10.3	1.24	8.05, 12.4	racemate	Benzimidazole
FXR 51	Cc1ccc(OCc2nc3ccc(F)cc3n2C(C4CCCCC4)C(=O)NC5CCCCC5)c(C)c1		0.25	1.28	0.31, 0.19	racemate	Benzimidazole
FXR 52	Fc1ccc2nc(COc3ccc(Cl)cc3Cl)n(C(C4CCCCC4)C(=O)NC5CCCCC5)c2c1		0.34	1.03	0.35, 0.33	racemate	Benzimidazole
FXR 53	OC(=O)c1ccc(NC(=O)C(C2CCCCC2)n3c(nc4cc(F)c(F)cc34)c5ccc(Cl)cc5)cc1		0.315	1.02	0.31, 0.32	racemate	Benzimidazole

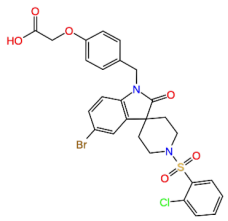
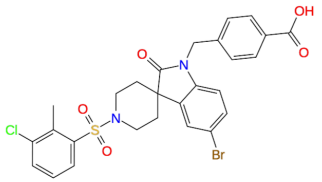
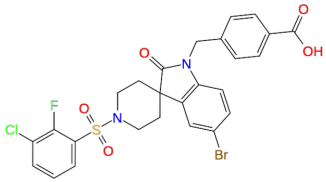
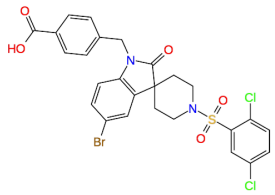
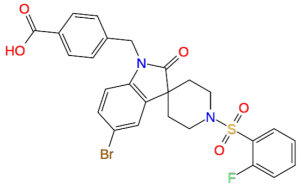
FXR 54	Fc1cc2nc(c3ccc(Cl)cc3)n(C(C4CCCCC4)C(=O)Nc5ccncc5)c2cc1F		1.63	1.04	1.56, 1.69	racemate	Benzimidazole
FXR 55	COc1ccc(c(OC)n1)c2nc3cc(F)c(F)cc3n2C(C4CCCCC4)C(=O)NC5CCCCC5		0.0003	1.14	0.00029, 0.00038	racemate	Benzimidazole
FXR 56	Fc1cc2nc(c3ccc(Cl)cc3)n(C(C4CCCCC4)C(=O)Nc5ccc(cc5)C(=O)Nc6ccccc6)c2cc1F		20.2	1.05	21.2, 19.2	racemate	Benzimidazole
FXR 57	COc1ccc(c(OC)n1)c2nc3cc(F)c(F)cc3n2C(C4CCCCC4)C(=O)NC5CCCC5		0.0091	1.10	0.01, 0.0082	racemate	Benzimidazole
FXR 58	OC(=O)c1ccc(NC(=O)C(C2CCCCC2)n3c(nc4cc(F)c(F)cc34)c5ccc(Cl)cc5)c(F)c1		0.0357	1.09	0.0388, 0.0327	racemate	Benzimidazole

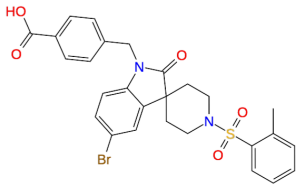
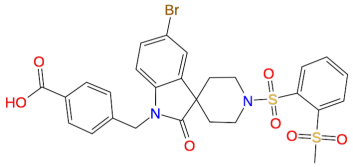
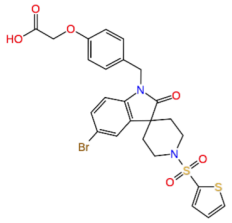
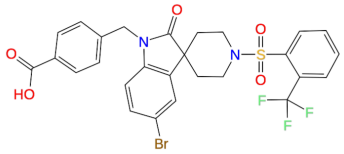
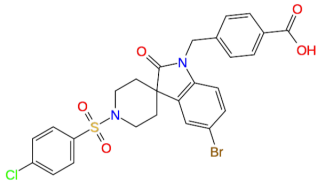
FXR 59	CC(=O)Nc1ccc(NC(=O)C(C2CCCCC2)n3c(nc4cc(F)c(F)cc34)c5ccc(Cl)cc5)cc1		0.59	1.05	0.56, 0.62	racemate	Benzimidazole
FXR 60	OC(=O)c1ccc(OCC(C2CCCCC2)n3c(nc4cc(F)c(F)cc34)c5ccc(Cl)cc5)c(Cl)c1		0.027	1.04	0.028, 0.026	racemate	Benzimidazole
FXR 61	Cc1cc(ccc1OCC(C2CCCCC2)n3c(nc4cc(F)c(F)cc34)c5ccc(Cl)cc5)C(=O)O		0.0225	1.02	0.023, 0.022	racemate	Benzimidazole
FXR 62	OC(=O)c1ccc(NC(=O)C(C2CCCCC2)n3c(nc4cc(F)c(F)cc34)c5ccc(Cl)nc5)c(Cl)c1		0.335	1.02	0.33, 0.34	racemate	Benzimidazole
FXR 63	OC(=O)c1ccc(NC(=O)C(C2CCCCC2)n3c(nc4cc(F)c(F)cc34)c5ccc(cc5)C(=O)O)cc1		56	1.12	62.1, 49.8	racemate	Benzimidazole

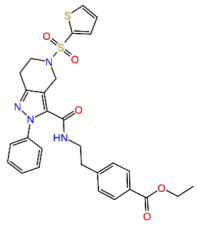
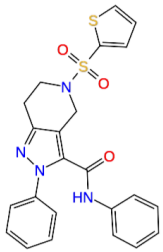
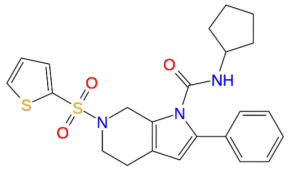
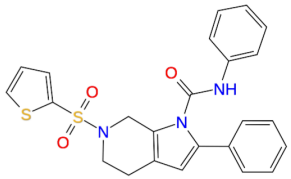
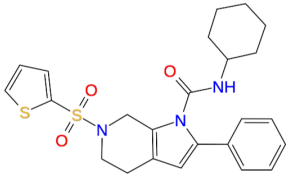
FXR 64	COc1ccc(c(OC)n1)c2nc3cc(F)c(F)cc3n2C(C4CCCCC4)C(=O)N[C@@H]5CC[C@H](CC5)C(=O)O		0.315	1.15	0.36, 0.27	racemate	Benzimidazole
FXR 65	CC(C)c1onc(c1COc2ccc(c(C)c2)c3ccc4c(cn(C)c4c3)C(=O)O)c5c(Cl)cncc5Cl		0.094	1.19	0.078, 0.11		Isoxazole
FXR 66	COc1ccc(c(OC)n1)c2nc3cc(F)c(F)cc3n2C(C4CCCCC4)C(=O)N[C@@H]5CC[C@H](CC(=O)O)CC5		20.8	1.25	16.2, 25.4	racemate	Benzimidazole
FXR 67	OC(=O)C[C@@H]1CC[C@@H](CC1)NC(=O)C(C2CCCCC2)n3c(nc4cc(F)c(F)cc34)c5ccc(nc5)n6cccn6		33.1	1.06	34.8, 30.3, 34.2	racemate	Benzimidazole
FXR 68	COc1ccc(c(OC)n1)c2nc3cc(F)c(F)cc3n2C(C4CCCCC4)C(=O)N[C@@H]5CC[C@@H](CC5)C(=O)O		109	1.18	127, 91.1	racemate	Benzimidazole

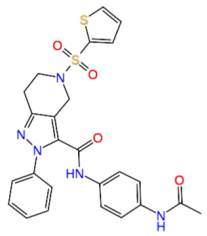
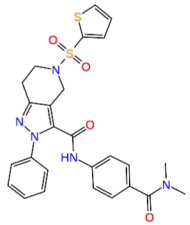
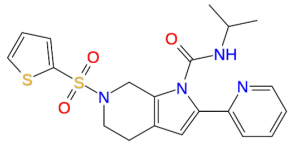
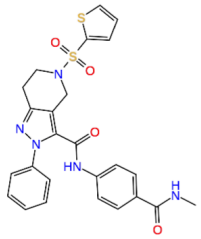
FXR 69	OC(=O)[C@@H]1CC[C@H](CC1)NC(=O)C(C2CCCCC2)n3c(nc4cc(F)c(F)cc34)c5ccc(nc5)n6cccn6		3.85	1.10	3.48, 4.21	racemate	Benzimidazole
FXR 70	Fc1cc2nc(c3ccc(nc3)n4cccn4)n(C(C5CCOCC5)C(=O)NC6CCCCC6)c2cc1F		0.915	1.09	0.99, 0.84	racemate	Benzimidazole
FXR 71	OC(=O)COc1ccc(OCC(C2CCCCC2)n3c(nc4cc(F)c(F)cc34)c5ccc(Cl)cc5)cc1		0.061	1.07	0.066, 0.056, 0.061	racemate	Benzimidazole
FXR 72	OC(=O)c1ccc(NC(=O)C(C2CCCCCC2)n3c(nc4cc(F)c(F)cc34)c5ccc(Cl)cc5)c(F)c1		0.0235	1.07	0.025, 0.022	racemate	Benzimidazole
FXR 73	Oc1ccc(CN2C(=O)C3(CCN(CC3)S(=O)(=O)c4cccs4)c5cc(Br)ccc25)cc1		11.2	1.07	10.5, 12		Spirocycle

FXR 74	OC(=O)c1ccc(CN2C(=O)C3(CCN(CC3)S(=O)(=O)c4ccccc4Br)c5cc(Br)ccc25)cc1		0.655	1.01	0.65, 0.66	Spirocycle
FXR 75	Brc1ccc2N(Cc3ccncc3)C(=O)C4(CCN(CC4)S(=O)(=O)c5cccs5)c2c1		100	-	>100	Spirocycle
FXR 76	OC(=O)c1ccc(CN2C(=O)C3(CCN(CC3)S(=O)(=O)c4ccccc4)c5cc(Br)ccc25)cc1		41.8	1.22	49.9, 33.8	Spirocycle
FXR 77	OC(=O)c1ccc(CN2C(=O)C3(CCN(CC3)S(=O)(=O)c4cccc(Cl)c4Cl)c5cc(Br)ccc25)cc1		0.25	1.08	0.23, 0.27	Spirocycle
FXR 78	OC(=O)c1ccc(CN2C(=O)C3(CCN(CC3)S(=O)(=O)c4c(Cl)cccc4Cl)c5cc(Br)ccc25)cc1		0.0283	1.45	0.014, 0.033, 0.035, 0.031	Spirocycle
FXR 79	OC(=O)c1cccc(CN2C(=O)C3(CCN(CC3)S(=O)(=O)c4cccs4)c5cc(Br)ccc25)c1		4.15	1.12	3.69, 4.61	Spirocycle

FXR 80	OC(=O)COc1ccc(CN2C(=O)C3(CCN(CC3)S(=O)(=O)c4ccccc4Cl)c5cc(Br)ccc25)cc1		0.24	1.41	0.32, 0.16	Spirocycle
FXR 81	Cc1c(Cl)cccc1S(=O)(=O)N2CCC3(CC2)C(=O)N(Cc4ccc(cc4)C(=O)O)c5ccc(Br)cc35		2.69	1.31	1.97, 3.4	Spirocycle
FXR 82	OC(=O)c1ccc(CN2C(=O)C3(CCN(CC3)S(=O)(=O)c4cccc(Cl)c4F)c5cc(Br)ccc25)cc1		0.18	1.35	0.15, 0.26, 0.13	Spirocycle
FXR 83	OC(=O)c1ccc(CN2C(=O)C3(CCN(CC3)S(=O)(=O)c4cc(Cl)ccc4Cl)c5cc(Br)ccc25)cc1		0.33	1.46	0.45, 0.21	Spirocycle
FXR 84	OC(=O)c1ccc(CN2C(=O)C3(CCN(CC3)S(=O)(=O)c4ccccc4F)c5cc(Br)ccc25)cc1		4.54	1.07	4.22, 4.86	Spirocycle

FXR 85	Cc1ccccc1S(=O)(=O)N2CCC3(CC2)C(=O)N(Cc4ccc(cc4)C(=O)O)c5ccc(Br)cc35		0.297	-	0.297	Spirocycle
FXR 86	CS(=O)(=O)c1ccccc1S(=O)(=O)N2CCC3(CC2)C(=O)N(Cc4ccc(cc4)C(=O)O)c5ccc(Br)cc35		18.8	1.41	12.6, 25	Spirocycle
FXR 87	OC(=O)COc1ccc(CN2C(=O)C3(CCN(CC3)S(=O)(=O)c4cccs4)c5cc(Br)ccc25)cc1		0.853	1.22	0.71, 0.99, 1.05, 0.66	Spirocycle
FXR 88	OC(=O)c1ccc(CN2C(=O)C3(CCN(CC3)S(=O)(=O)c4ccccc4C(F)(F)F)c5cc(Br)ccc25)cc1		0.54	-	0.54	Spirocycle
FXR 89	OC(=O)c1ccc(CN2C(=O)C3(CCN(CC3)S(=O)(=O)c4ccc(Cl)cc4)c5cc(Br)ccc25)cc1		0.735	1.14	0.83, 0.64	Spirocycle

FXR 90	CCOC(=O)c1ccc(CCNC(=O)c2c3CN(CCc3nn2c4ccccc4)S(=O)(=O)c5cccs5)cc1		100	-	>100	Tetrahydropyrro-(azo)lopyridines
FXR 91	O=C(Nc1ccccc1)c2c3CN(CCc3nn2c4ccccc4)S(=O)(=O)c5cccs5		29.6	1.28	36.8, 22.4	Tetrahydropyrro-(azo)lopyridines
FXR 92	O=C(NC1CCCC1)n2c3CN(CCc3cc2c4ccccc4)S(=O)(=O)c5cccs5		0.505	1.37	0.35, 0.66	Tetrahydropyrro-(azo)lopyridines
FXR 93	O=C(Nc1ccccc1)n2c3CN(CCc3cc2c4ccccc4)S(=O)(=O)c5cccs5		46.7	1.26	36, 57.3	Tetrahydropyrro-(azo)lopyridines
FXR 94	O=C(NC1CCCCC1)n2c3CN(CCc3cc2c4ccccc4)S(=O)(=O)c5cccs5		18.6	1.17	21.6, 15.7	Tetrahydropyrro-(azo)lopyridines

FXR 95	CC(=O)Nc1ccc(NC(=O)c2c3CN(CCc3nn2c4ccccc4)S(=O)(=O)c5cccs5)cc1		32.2	1.03	31.3, 33	Tetrahydropyrro-(azo)lopyridines
FXR 96	CN(C)C(=O)c1ccc(NC(=O)c2c3CN(CCc3nn2c4ccccc4)S(=O)(=O)c5cccs5)cc1		58.9	1.01	59.2, 58.5	Tetrahydropyrro-(azo)lopyridines
FXR 97	CC(C)NC(=O)n1c2CN(CCc2cc1c3ccccn3)S(=O)(=O)c4cccs4		20.1	1.07	21.5, 18.8	Tetrahydropyrro-(azo)lopyridines
FXR 98	CNC(=O)c1ccc(NC(=O)c2c3CN(CCc3nn2c4ccccc4)S(=O)(=O)c5cccs5)cc1		13.1	1.02	12.9, 13.4	Tetrahydropyrro-(azo)lopyridines

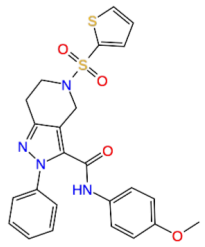
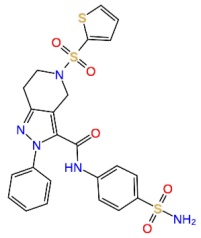
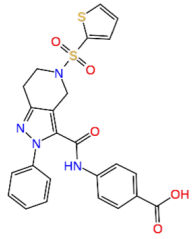
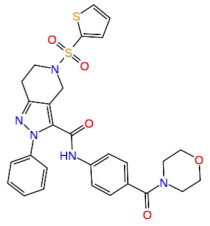
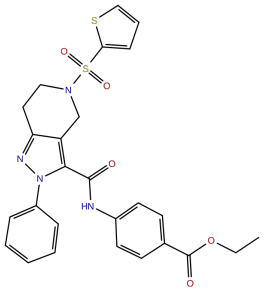
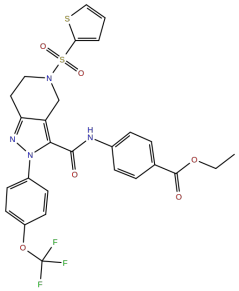
FXR 99	COc1ccc(NC(=O)c2c3CN(CCc3nn2c4ccccc4)S(=O)(=O)c5cccs5)cc1		100	-	>100	Tetrahydropyrro-(azo)lopyridines
FXR 100	NS(=O)(=O)c1ccc(NC(=O)c2c3CN(CCc3nn2c4ccccc4)S(=O)(=O)c5cccs5)cc1		19.1	1.05	20.1, 18.2	Tetrahydropyrro-(azo)lopyridines
FXR 101	OC(=O)c1ccc(NC(=O)c2c3CN(CCc3nn2c4ccccc4)S(=O)(=O)c5cccs5)cc1		27.6	1.32	35.2, 20.1	Tetrahydropyrro-(azo)lopyridines
FXR 102	O=C(Nc1ccc(cc1)C(=O)N2CCOCC2)c3c4CN(CCc4nn3c5ccccc5)S(=O)(=O)c6cccs6		29.2	1.09	26.7, 31.8	Tetrahydropyrro-(azo)lopyridines

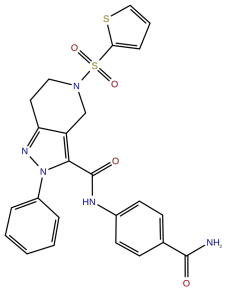
Table S3. Free Energy Set 1. Each panel lists the chemical structure, Compound ID, and experimental binding free energy estimated as RT ln (IC50).



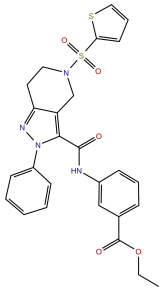
FXR 17
ΔG = -8.38 kcal/mol



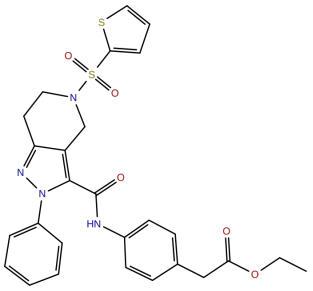
FXR 45
ΔG = -6.23 kcal/mol



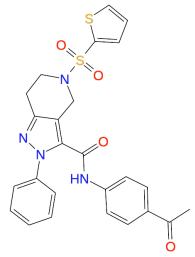
FXR 46
ΔG = -5.77 kcal/mol



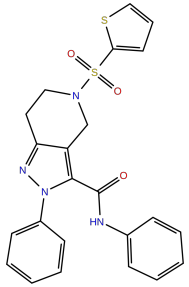
FXR 47
ΔG = -6.42 kcal/mol



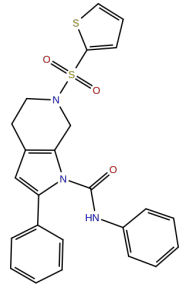
FXR 48
ΔG = -5.49 kcal/mol



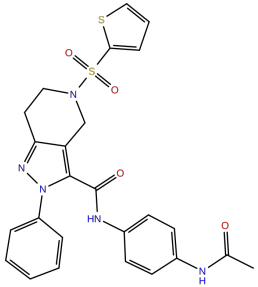
FXR 49
ΔG = -5.49 kcal/mol



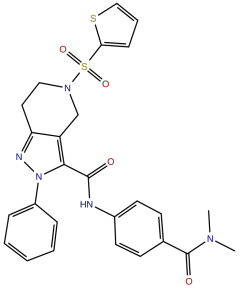
FXR 91
ΔG = -6.22 kcal/mol



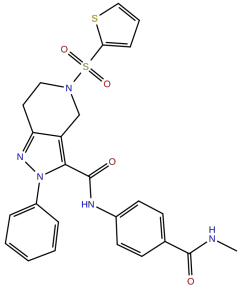
FXR 93
ΔG = -5.94 kcal/mol



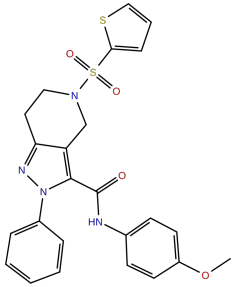
FXR 95
ΔG = -6.17 kcal/mol



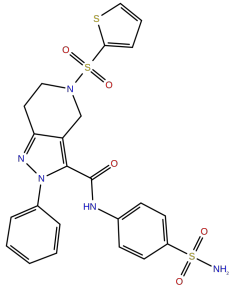
FXR 96
ΔG = -5.81 kcal/mol



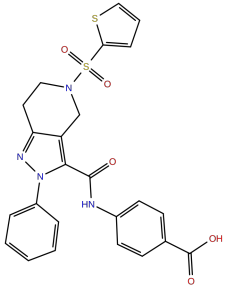
FXR 98
ΔG = -6.70 kcal/mol



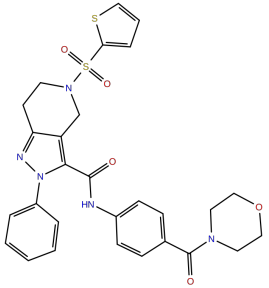
FXR 99
ΔG = -5.49 kcal/mol



FXR 100
ΔG = -6.48 kcal/mol

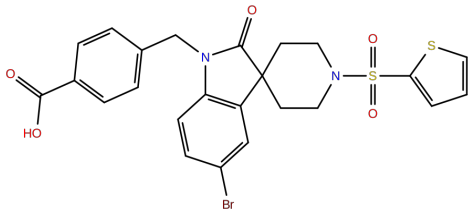


FXR 101
ΔG = -6.26 kcal/mol

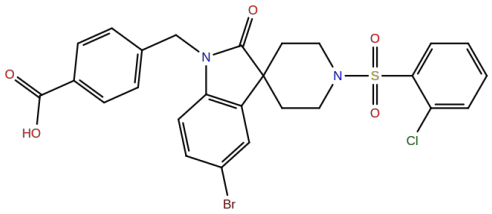


FXR 102
ΔG = -6.22 kcal/mol

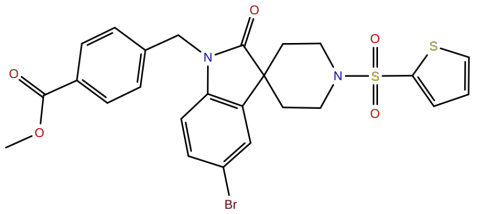
Table S4. Free Energy Set 2. Each panel lists the chemical structure, Compound ID, and experimental binding free energy estimated as RT ln IC50.



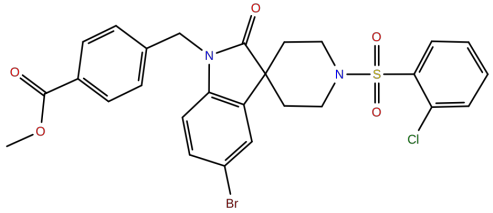
FXR 10
ΔG = -7.20 kcal/mol



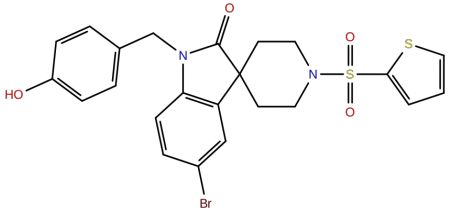
FXR 12
ΔG = -9.93 kcal/mol



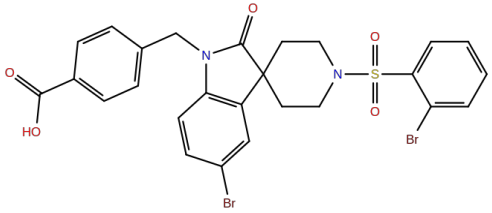
FXR 38
ΔG = -5.49 kcal/mol



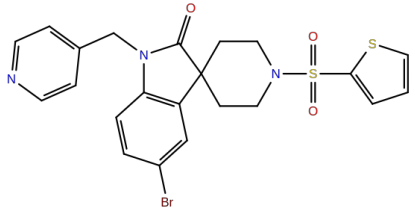
FXR 41
ΔG = -5.49 kcal/mol



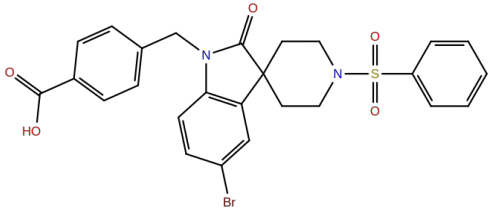
FXR 73
ΔG = -6.79 kcal/mol



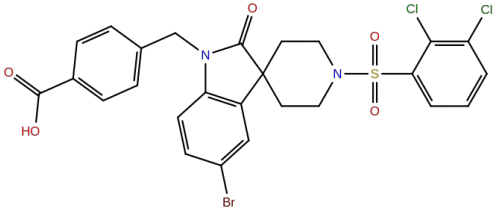
FXR 74
ΔG = -8.49 kcal/mol



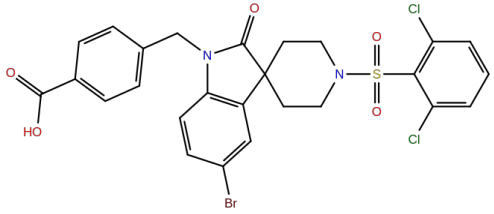
FXR 75
ΔG = -5.49 kcal/mol



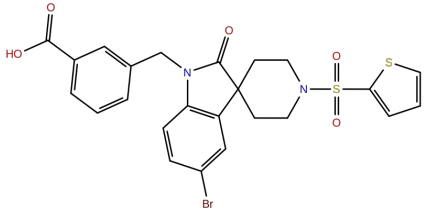
FXR 76
ΔG = -6.01 kcal/mol



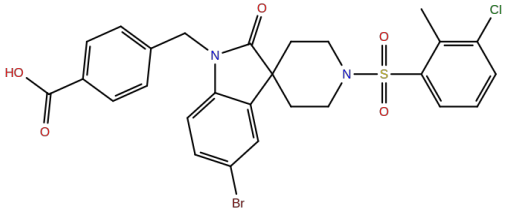
FXR 77
ΔG = -9.06 kcal/mol



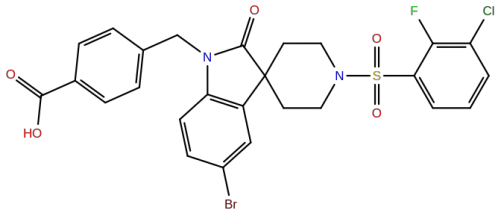
FXR 78
ΔG = -10.36 kcal/mol



FXR 79
ΔG = -7.39 kcal/mol

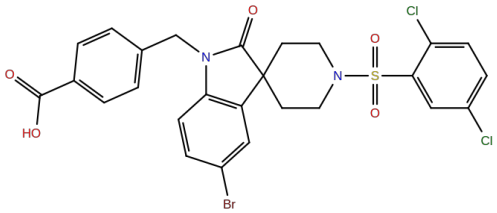


FXR 81
ΔG = -7.65 kcal/mol



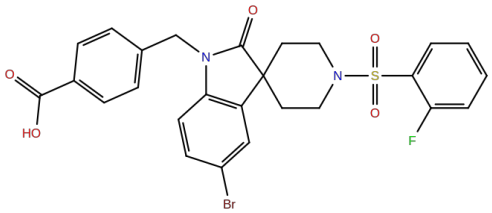
FXR 82

ΔG = -9.26 kcal/mol



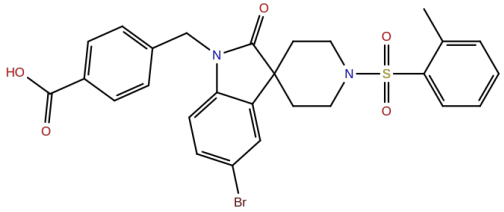
FXR 83

ΔG = -8.90 kcal/mol



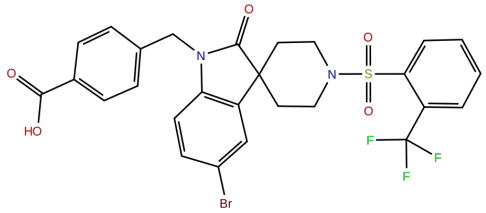
FXR 84

ΔG = -7.33 kcal/mol



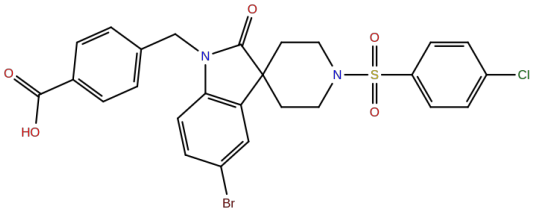
FXR 85

ΔG = -8.96 kcal/mol



FXR 88

ΔG = -8.60 kcal/mol



FXR 89

ΔG = -8.42 kcal/mol

Table S5. D3R refinement statistics on Farnesoid-X-Receptor (FXR) X-ray structures. Work and free R-factors (Rfree) are measures of the fit of the model to a subset of the experimental data that was used in model refinement in the case of Rwork or set aside for experimental validation in the case of Rfree. The overall quality percentile indicates the percentage of PDB entries that are equal or less than the quality indicator of the structure of interest.

Ligand ID	PDB ID	Resolution (Å)	Initial Model		Final Model		RMSD Initial *versus* Final (Å)		Overall Quality Percentile	
			Rwork	Rfree	Rwork	Rfree	Protein	Ligand	Initial	Final
FXR_1	5Q0I	1.7	0.185	0.236	0.185	0.226	0.079	0.089	35.10%	73.30%
FXR_2	5Q12	2	0.178	0.245	0.184	0.233	0.106	0.158	68.00%	66.10%
FXR_3	5Q1G	2	0.208	0.264	0.205	0.254	0.263	0.126	7.70%	10.60%
FXR_4	5Q11	2.2	0.184	0.265	0.178	0.223	0.245	0.118	25.80%	68.90%
FXR_5	5Q0Z	2.26	0.243	0.296	0.236	0.258	0.291	0.18	1.20%	19.40%
FXR_6	5Q19	1.98	0.235	0.279	0.215	0.237	0.231	0.094	3.80%	22.50%
FXR_7	5Q0J	2	0.226	0.284	0.195	0.247	0.149	0.105	3.90%	33.60%
FXR_8	5Q0L	2.5	0.229	0.301	0.209	0.263	0.266	0.131	1.20%	14.40%
FXR_9	5Q1H	2.2	0.219	0.262	0.18	0.218	0.183	0.093	26.60%	80.40%
FXR_10	5Q17	2.1	0.193	0.274	0.218	0.267	0.139	0.233	12.60%	19.00%
FXR_11	5Q1E	1.85	0.19	0.232	0.188	0.223	0.16	0.189	32.60%	74.60%
FXR_12	5Q0W	1.9	0.206	0.26	0.208	0.254	0.17	0.109	5.50%	23.90%

FXR_13	5Q0O	1.9	0.182	0.225	0.191	0.222	0.185	0.071	36.20%	52.10%
FXR_14	5Q1A	2	0.228	0.273	0.206	0.235	0.188	0.097	8.00%	43.80%
FXR_15	5Q0R	1.91	0.214	0.273	0.208	0.244	0.232	0.143	1.60%	22.30%
FXR_16	5Q0T	2.14	0.202	0.289	0.191	0.259	0.271	0.129	1.60%	14.20%
FXR_17	5Q0Q	2.6	0.253	0.318	0.246	0.295	0.292	0.204	1.50%	8.80%
FXR_18	5Q10	2.2	0.191	0.26	0.189	0.243	0.196	0.156	14.70%	50.30%
FXR_19	5Q13	1.9	0.167	0.232	0.171	0.226	0.22	0.109	48.90%	82.80%
FXR_20	5Q16	2	0.254	0.304	0.263	0.284	0.196	0.118	0.40%	9.50%
FXR_21	5Q1D	1.89	0.208	0.249	0.206	0.244	0.193	0.074	27.50%	61.80%
FXR_22	5Q15	1.9	0.241	0.271	0.231	0.255	0.178	0.083	2.20%	32.80%
FXR_23	5Q1I	1.95	0.202	0.239	0.199	0.225	0.125	0.169	8.80%	29.20%
FXR_24	5Q0V	1.87	0.215	0.257	0.216	0.238	0.164	0.082	16.60%	52.90%
FXR_25	5Q0S	2.5	0.189	0.274	0.195	0.26	0.247	0.084	34.40%	62.40%
FXR_26	5Q0Y	2.2	0.185	0.238	0.183	0.22	0.22	0.093	46.40%	68.10%
FXR_27	5Q0N	2.4	0.205	0.252	0.188	0.226	0.261	0.092	38.30%	76.90%
FXR_28	5Q14	1.85	0.209	0.255	0.212	0.247	0.243	0.09	33.20%	45.40%
FXR_29	5Q0P	1.8	0.226	0.269	0.223	0.254	0.162	0.074	37.00%	67.50%

FXR_30	5Q1F	2.3	0.231	0.285	0.221	0.245	0.235	0.109	4.40%	35.00%
FXR_31	5Q1C	2.3	0.235	0.284	0.228	0.264	0.222	0.1	11.60%	41.20%
FXR_32	5Q18	1.9	0.227	0.28	0.213	0.24	0.188	0.122	4.20%	47.70%
FXR_33	5Q0X	2.26	0.188	0.242	0.176	0.233	0.265	0.1	24.90%	80.60%
FXR_34	5Q0M	2.2	0.2	0.276	0.183	0.262	0.187	0.119	17.30%	45.20%
FXR_35	5Q0U	1.9	0.208	0.252	0.196	0.218	0.167	0.085	12.30%	63.80%
FXR_36	5Q1B	2.3	0.216	0.265	0.214	0.236	0.259	0.103	7.30%	34.10%
Apo	5Q0K	1.8	0.211	0.277	0.21	0.271	0.217	N/A*	15.50%	27.80%

Table S6. Summary of results for pose prediction challenge. For each submission, the table lists the Receipt ID (a unique ID given to each prediction upon submission); the number of ligands for which pose predictions were provided; the median RMSD (Å), relative to the experimental structures, for the poses marked as the best guess across all 35 ligands (Pose 1); the median RMSD (Å) of Pose 1 predictions only for ligands with a similarity coefficient to publicly available ligands of less than 0.3 (FXRs 1, 2, 3, 4, 10, 11, 12, 15, 16, 17, 18, 23); the mean RMSD of the Pose 1 ligands relative to the experimental structures; and a list of software used, according to the protocol file, as well as other methodological notes. Also listed is whether the protocol file mentioned any use of ligand similarity between the challenge ligands and publicly available co-crystallized FXR ligands, and whether visual inspection played a role in the method.

Receipt ID	Number Ligands Docked	Median RMSD	Median RMSD (ligands with Tanimoto Similarity to Publicly Available Ligands <0.3)	Mean RMSD	Software Used	Used similarity	Visual Inspection
touhi	18	0.68	N/A	0.81	Glide, CCDC Gold, Desmond (MD)	Yes	No
oky3v	35	0.99	5.66	2.34	WaterMap, SHAPE Screening, Structural Interaction Fingerprint, DFT/B3LYP/6-31G*, GLIDE-SP-XP, Induced-fit-docking, Emodel/GlideScore-SP, Binding Pose Metadynamics	Yes	No
mgxbc	35	1.00	5.72	2.76	OMEGA, SHAFTS, NoDocking, Amber11	Yes	No
7ltmc	35	1.02	3.74	2.19	GLIDE-CCDC-GOLD, Amber14, MMGBSA	Not mentioned	No
hciq4	35	1.13	3.75	2.19	Glide, Gold, Induced-fit-docking	Not mentioned	No
cfn8u	35	1.16	3.85	2.37	Glide-XP	No	No
ixnzu	35	1.17	3.27	1.95	Molsoft ICM	Yes	No
5rqrx	35	1.19	6.38	2.92	ROCS, Omega, Glide-SP-XP	Yes	No
5cf33	35	1.20	4.25	2.76	Modeller, Gromacs (minimization), surflex-sim, Clusterizer and DockAccessor, Vina	Yes	No
psiuj	35	1.20	6.71	3.16	Forge, Smina (minimization only), Glide-SP	Yes	Yes
wax1j	35	1.20	7.15	3.08	MOE (most similar molecule), Confgen, Forge (alignment), Smina (minimization), Manual ranking	Yes	Yes

txyzj	35	1.27	2.73	2.12	Autodock Vina, GROMACS, in-house LIE (Linear Interaction Energy Model)	Not mentioned	No
jz0em	35	1.39	3.06	2.27	Wilma (flexible ligand docking algorithm), scoring function SIE (force field based scoring function)	Not mentioned	No
wzu5m	35	1.41	7.11	3.13	Glide, CCDC Gold, Xscore (re-score)	Yes	No
qfrw4	35	1.59	6.55	3.44	Simulated Annealing, CDOCKER, Vina (re-score)	Not mentioned	No
x7jp3	35	1.63	7.81	3.79	Smina	Not mentioned	No
to4tt	35	1.75	5.49	3.53	Smina	Not mentioned	No
gfifa	35	1.75	7.12	3.92	ROCS TanimotoComboShapeSimilarity, DockBench: autodock-ga, autodock-lga, autodock-ls, glide-sp, gold-asp, gold-chemscore, gold-goldscore, gold-plp, moe-affinitydg, moe-gbviwsa, moe-londondg, plants-chemplp, plants-plp, plants-plp95, rdock-solv, rdock-std	Yes	No
qbvvg	35	1.75	7.15	3.26	OMEGA (ligand conformational sampling), SHAFTS (structural similarity score), AutoDock Vina (pose sampling), ITScore_v1_TF (knowledge-based scoring function)	Yes	No
4i2mb	35	1.75	7.73	3.79	OMEGA (ligand conformational sampling), SHAFTS (structural similarity score), AutoDock Vina (pose sampling), ITScore_v2_TF (knowledge-based scoring function)	Yes	No
piwli	34	1.76	5.16	3.31	Glide	Not mentioned	No
qfu33	35	1.79	6.28	3.32	Induced-fit-docking, Binding pose metadynamics (rescoring)	Not mentioned	No
mmj0e	35	1.80	7.15	3.77	OMEGA (ligand conformational sampling), SHAFTS (structural similarity score), AutoDock Vina (pose sampling), ITScore_TF (knowledge-based scoring function)	Yes	No
gdhv4	35	1.80	7.76	4.12	OMEGA (ligand conformational sampling), SHAFTS (structural similarity score),	Yes	No

					AutoDock Vina (pose sampling), ITScore_v1 (knowledge-based scoring function)		
elm2r	35	1.91	6.64	3.65	AutoDock Vina, Convex-PL scoring function	Yes	No
utpcl	35	1.94	5.78	3.26	Surflex, Ichem-GRIM (interaction pattern similarity), HYDE	Yes	No
piwlh	35	2.07	6.19	3.48	Induced-fit-docking	Not mentioned	No
yyjmb	35	2.08	7.41	3.59	Surflex, Ichem-GRIM (interaction pattern similarity), HYDE	Yes	No
6tnqb	35	2.25	4.33	3.43	Smina	Yes	Yes
h67ea	35	2.25	4.33	3.43	Smina	Yes	Yes
tbuot	35	2.69	5.61	3.38	Surflex, Ichem-GRIM (interaction pattern similarity), HYDE	Yes	No
04hag	34	2.69	5.71	3.68	Glide-XP (No scoring, one pose outputted for each ligand)	Not mentioned	No
am6an	35	3.28	6.43	4.17	Schrodinger QM-polarized ligand docking, Glide XP, AutoDock Vina	Not mentioned	No
3rfl5	35	3.74	7.68	4.3	OMEGA (ligand conformational sampling), SHAFTS (structural similarity score), AutoDock Vina (pose sampling), ITScore_v2_TF (knowledge-based scoring function)	Yes	No
02sil	35	4.33	6.81	4.33	AutoDock 4, SeeSAR (minimization), HYDE scoring function	Yes (for FXR 5 and 33 only)	Yes
byf51	35	4.57	5.15	4.25	AutoDock Vina, Gromacs	Not mentioned	No
0lxp5	35	4.70	5.32	4.53	AutoDock Vina	Not mentioned	No
l835g	35	4.73	5.15	4.39	AutoDock Vina, Gromacs, Reconnaissance metadynamic method (enhanced sampling)	Not mentioned	No
m00nf	34	5.72	5.50	6.37	MOE	Not mentioned	No
7rmfb	35	5.78	6.82	5.13	Smina, idock-RF-v3, Yasara (simulated annealing)	Not mentioned	Yes
3rj71	35	5.90	5.21	6.34	OMEGA (ligand conformational sampling), AutoDock Vina	Not mentioned	No

41tia	33	5.90	6.77	5.68	CCDC-Gold, In-house machine learning scoring method	Not mentioned	No
0xk7w	35	5.98	6.09	5.07	Smina, gnina-cnn (rescoring, trained using CSAR and virtual screening data from DUDE)	Not mentioned	No
mx1v2	35	6.09	6.74	5.28	Smina, gnina-cnn (rescoring, trained using CSAR and FXR ligand poses from PDB)	Not mentioned	No
00ulb	33	6.21	5.37	6.19	SILCS (Site Identification by Ligand Competitive Saturation)	Not mentioned	No
qxrku	35	6.39	6.03	6.15	HADDOCK2.2	Not mentioned	No
qt334	35	6.60	6.68	6.34	Smina, gnina-cnn (rescoring, trained using CSAR)	Not mentioned	No
e0aar	35	6.94	6.94	6.03	CCDC-Gold, In-house machine learning scoring method	Not mentioned	No
clzrj	35	6.96	7.81	6.74	Smina, RF-Score-VS v2 and v3	Yes	No
knz3v	35	7.02	7.02	7.67	Cross-Docking Benchmark, Gold-goldscore	Used Apo Only	No
ypz46	35	7.28	7.80	6.24	Smina, RF-Score-VS v2 and v3	Yes	No
8mocv	35	7.53	19.5	11.87	Rhodium380E6-4 (proprietary)	Not mentioned	No

Table S7. Kendall's tau statistics and software used for FXR affinity ranking predictions during Stage 1. For each participant, the table lists Receipt ID, the number of ligands for which predictions were made, Kendall's tau and Kendall's tau uncertainty, method type (structure- vs ligand-based), and software used. Asterisks indicate the use of visual inspection.

Receipt ID	Number of Ligands	Kendall's Tau	Kendall's Tau Uncertainty	Method Type	Software Used
g7q2q	102	0.44	0.05	Structure-based	AutoDock Vina
87x7c	102	0.44	0.05	Structure-based	AutoDock Vina
4ynsp	102	0.44	0.06	Structure-based	Ichem-GRIM, HYDE
6pcik	102	0.43	0.05	Structure-based	idock-RF-v3*

gzd7a	102	0.43	0.05	Structure-based	Smina
6xixa	36	0.43	0.09	Structure-based	Glide, Gold, Amber-MMGBSA
0kg1e	102	0.42	0.05	Structure-based	Smina
ttgw7	102	0.41	0.05	Structure-based	Glide
aaveo	102	0.40	0.05	Structure-based	Smina*
2o6iv	102	0.38	0.06	Structure-based	Ichem-GRIM, HYDE
f48cf	102	0.37	0.05	Structure-based	Smina
nvoyn	102	0.36	0.06	Structure-based	RF-Score-VS, Smina
3qyiy	102	0.36	0.05	Structure-based	Amber-MMPBSA
5f688	102	0.35	0.05	Structure-based	gnina_cnn
1t2p5	102	0.35	0.06	Structure-based	Ichem-GRIM, HYDE
f30wc	102	0.35	0.06	Structure-based	RF-Score-VS, Smina
h2w3q	102	0.34	0.06	Structure-based	SeeSAR
ec84y	102	0.34	0.06	Structure-based	xScore
mseke	102	0.33	0.05	Structure-based	Smina
6pf5g	102	0.31	0.05	Structure-based	Glide-XP
too1u	26	0.30	0.14	Structure-based	Rhodium HTS
fzr1f	102	0.30	0.05	Structure-based	SIE (Solvated Interaction Energy)
c20xb	102	0.30	0.06	Structure-based	ITScore_TF
l0xpi	102	0.30	0.06	Structure-based	Molsoft ICM
ukdfw	102	0.29	0.06	Structure-based	Smina
4ivv5	102	0.28	0.06	Structure-based	ITScore_v2_TF
yprv4	102	0.28	0.06	Structure-based	AutoDock Vina, Clusterizer-DockAccessor
bwdnh	102	0.28	0.06	Structure-based	gnina_cnn
33a8g	102	0.28	0.06	Structure-based	ITScore_v2
1un0y	102	0.28	0.05	Structure-based	gnina_cnn
pgbuh	102	0.28	0.06	Structure-based	CDOCKER (pose prediction) + Autodock Vina (scoring)
6fzf6	102	0.27	0.06	Structure-based	HADDOCK2.2 web server

q76s3	102	0.27	0.06	Structure-based	ITScore_v1_TF
k2um8	102	0.26	0.06	Structure-based	qpld, Glide-XP, Glide-SP, AutoDock Vina, AutoDock 4
dlvjf	102	0.26	0.06	Structure-based	qpld, Glide-XP, Glide-SP, AutoDock Vina, AutoDock 4
mzwwt	102	0.25	0.06	Structure-based	ITScore_v2
pcy26	102	0.24	0.06	Structure-based	Molsoft ICM
kz0dz	102	0.20	0.07	Ligand-based	3D QSAutogrid/R
fvfe7	102	0.19	0.06	Structure-based	ITScore_v1
lpcmd	102	0.18	0.07	Ligand-based	Glide-XP
o3nlu	102	0.18	0.07	Structure-based	Glide-XP
a2w8d	102	0.15	0.07	Structure-based	Smina
l8rmr	102	0.14	0.06	Structure-based	LIE (Linear Interaction Energy Model)
mf3fc	102	0.13	0.06	Structure-based	GoldScore, ChemScore, ChemPLP, ASP
h1on0	102	0.13	0.07	Structure-based	ITScore_v1
r885q	102	0.10	0.07	Structure-based	ITScore_v2_TF
5nim5	102	0.08	0.05	Structure-based	Smina
dnil6	102	0.07	0.06	Structure-based	In-house machine learning score
vskkl	102	0.07	0.07	Structure-based	AutoDock Vina
sn46i	102	0.07	0.06	Structure-based	GoldScore, ChemScore, ChemPLP, ASP, RMSD similarity to reference fragment
vu0d0	102	0.02	0.06	Structure-based	ITScore_TF
cwip8	102	-0.01	0.07	Structure-based	SILCS (Site Identification by Ligand Competitive Saturation)
c0l1t	102	-0.03	0.07	Structure-based	In-house machine learning score
bhtq4	102	-0.04	0.07	Structure-based	ITScore_v1_TF
tdbsk	102	-0.05	0.07	Structure-based	SILCS (Site Identification by Ligand Competitive Saturation)
cs2lm	102	-0.09	0.06	Structure-based	Amber-MMGBSA
keum2	102	-0.11	0.06	Structure-based	GBVI/WSA scoring
2umsq	102	-0.30	0.06	Structure-based	SILCS (Site Identification by Ligand Competitive Saturation)
jr0oc	102	-0.39	0.06	Structure-based	Gold

Table S8. Kendall's tau statistics and software used for FXR affinity ranking predictions during Stage 2. For each participant, the table lists Receipt ID, number of ligands for which predictions were made, Kendall's tau and Kendall's tau uncertainty, method type (structure- vs ligand-based), and software used. Asterisks indicate use of visual inspection.

Receipt ID	Number of Ligands	Kendall's Tau	Kendall's Tau Uncertainty	Method Type	Software Used
f2wjs	102	0.46	0.05	Structure-based	Smina
5zcmb	102	0.44	0.05	Structure-based	Rhodium HTS
7c2tp	102	0.44	0.05	Structure-based	Smina
fo0p4	102	0.42	0.05	Structure-based	Smina
dh2du	102	0.41	0.05	Structure-based	SeeSAR, HYDE, MMGBSA
wfe3c	102	0.39	0.05	Structure-based	Vina
naex2	102	0.38	0.06	Ligand-based	In-house QSAR script
0f7u7	102	0.38	0.06	Structure-based	PRODIGY webserver
hj31e	102	0.37	0.05	Ligand-based	In-house QSAR script
aewpt	102	0.36	0.05	Structure-based	Smina
erdfw	102	0.36	0.05	Structure-based	MMPBSA
6nrpf	102	0.35	0.05	Structure-based	Vinardo Score
15ub2	102	0.35	0.05	Structure-based	Smina
vovuk	102	0.33	0.05	Structure-based	MMGBSA
35yg0	102	0.33	0.05	Structure-based	Rhodium HTS
riyj0	102	0.32	0.05	Structure-based	Smina, CNN Model Scoring
mbbz2	102	0.31	0.06	Structure-based	Smina, Vinardo scoring function
6mjkt	102	0.31	0.06	Structure-based	Smina
hvjjb	102	0.31	0.06	Structure-based	Smina, CNN Model Scoring
66nbk	102	0.3	0.07	Structure-based	SIE (solvated interaction energy) with FISH solvation model
6fsmu	102	0.29	0.06	Structure-based	Smina, CNN Model Scoring
hekzj	102	0.28	0.06	Structure-based	Smina, CNN Model Scoring

ljdjm	102	0.28	0.06	Structure-based	Glide-XP
gnu3p	102	0.27	0.06	Ligand-based	ChemmineR
tbxzq	102	0.26	0.06	Structure-based	Glide
0pgek	102	0.25	0.07	Structure-based	Ligands ranked manually
li83b	102	0.25	0.05	Structure-based	Molsoft ICM
p8rak	102	0.24	0.05	Structure-based	MMGBSA_modified
x8jt0	102	0.23	0.06	Structure-based	eTOX ALLIES
qokw3	102	0.23	0.06	Structure-based	Xscore
v4uoy	102	0.23	0.06	Structure-based	Rhodium HTS
20shi	102	0.23	0.06	Structure-based	ITScore_v2_TF
pr2fp	102	0.22	0.06	Structure-based	ITScore_v2
0qoih	102	0.22	0.06	Structure-based	Xscore
sge2j	102	0.21	0.06	Structure-based	CDOCKER (pose prediction) + Autodock Vina (scoring)
4rbjk	102	0.2	0.06	Structure-based	RI-Score
n55eq	102	0.19	0.06	Structure-based	RI-Score
vxvhq	102	0.18	0.07	Structure-based	MMGBSA
xv7qs	102	0.18	0.06	Structure-based	SIE (solvated interaction energy)
5hri3	102	0.17	0.07	Structure-based	ITScore_v1
rjyhz	102	0.17	0.06	Structure-based	GoldScore, ChemScore, ChemPLP, ASP, lowestRMSDtoRefFragment
cvxs6	102	0.16	0.07	Structure-based	Smina, Random Forest Model
jdr8n	102	0.15	0.06	Structure-based	MMGBSA
fww4f	102	0.15	0.06	Structure-based	AutoDock Vina, Clusterizer & DockAccessor
w8p2f	102	0.15	0.06	Structure-based	RI-Score
pny4b	102	0.14	0.07	Structure-based	ITScore_TF
vbzci	102	0.14	0.05	Structure-based	LIE (Linear Interaction Energy Model)
jtsy2	102	0.12	0.07	Structure-based	RI-Score
rtqum	102	0.12	0.06	Structure-based	RI-Score

a8r3i	102	0.12	0.06	Structure-based	ITScore_v1_TF
jajar	102	0.11	0.07	Structure-based	ITScore_v2_TF
bw4pj	102	0.11	0.07	Structure-based	RI-Score
bxn4v	102	0.11	0.07	Structure-based	ITScore_v2
86ozh	102	0.10	0.07	Structure-based	RI-Score
35sno	102	0.10	0.06	Structure-based	Rhodium HTS
0aggj	102	0.10	0.06	Structure-based	RI-Score
hgqci	102	0.09	0.07	Structure-based	Smina
sb1dg	102	0.09	0.07	Structure-based	ITScore_v1
g4bd3	102	0.08	0.07	Structure-based	Smina, in-house scoring function
p2wkt	102	0.07	0.07	Structure-based	Chemscore
hxm7v	102	0.07	0.06	Structure-based	Glide-SP, GIST scoring 2
moii1	102	0.06	0.06	Structure-based	RI-Score
mtbxd	102	0.05	0.07	Structure-based	RI-Score
y5gbl	102	0.05	0.06	Structure-based	Glide-SP, GIST scoring 3
15tvu	102	0.04	0.07	Structure-based	ITScore_TF
bckkg	102	0.03	0.05	Structure-based	Glide-SP
uyp2k	102	0.01	0.07	Structure-based	ITScore_v1_TF
0zno2	102	0	0.06	Structure-based	Smina, CNN Model Scoring
327r0	102	-0.01	0.07	Structure-based	BEDAM (Binding Energy Distribution Analysis Method)
awgis	102	-0.01	0.06	Structure-based	Glide-SP, GIST (solvent displacement score) scoring 1
21jpg	102	-0.02	0.07	Ligand-based	3D QSAutogrid/R
4m842	102	-0.02	0.07	Ligand-based	3D QSAutogrid/R
tgmx1	102	-0.03	0.06	Structure-based	Smina, CNN Model Scoring
22juj	102	-0.03	0.07	Structure-based	BEDAM (Binding Energy Distribution Analysis Method)
bxfza	102	-0.03	0.06	Structure-based	Smina, CNN Model Scoring
xe4uo	102	-0.04	0.08	Structure-based	Smina
qcbm1	102	-0.07	0.06	Structure-based	Xscore

xkgbb	102	-0.07	0.07	Structure-based	Smina, CNN Model Scoring
gia3u	102	-0.18	0.06	Structure-based	SILCS (Site Identification by Ligand Competitive Saturation)
7wx7v	102	-0.20	0.06	Structure-based	SILCS (Site Identification by Ligand Competitive Saturation)
ska85	102	-0.30	0.06	Structure-based	MMGBSA
rwmcz	102	-0.33	0.06	Structure-based	SILCS (Site Identification by Ligand Competitive Saturation)

Table S9. $RMSE_c$ statistics and software used for free energy predictions of FE Set 1 during Stage 1. Statistics are calculated for predictions from scoring methods with free energy estimates in the affinity ranking component of the challenge, and free energy methods in the free energy component of the challenge. For each participant, the table lists Receipt ID, number of ligands used, $RMSE_c$ and $RMSE_c$ uncertainty, method type [scoring (structure- vs ligand-based) vs alchemical methods], and software used. Asterisks indicate the use of visual inspection.

Receipt ID	Number of Ligands	$RMSE_c$	$RMSE_c$ Uncertainty	Method Type	Software Used
pgbuh	15	0.66	0.20	Scoring (structure-based with FE estimates)	CDOCKER (pose prediction) + Autodock Vina (scoring)
5bvwx	15	0.68	0.20	Scoring (structure-based with FE estimates)	In-house machine learning score
c0l1t	15	0.68	0.20	Scoring (structure-based with FE estimates)	In-house machine learning score
2umsq	15	0.69	0.21	Scoring (structure-based with FE estimates)	SILCS (Site Identification by Ligand Competitive Saturation)
1bhkb	15	0.70	0.23	Scoring (structure-based with FE estimates)	Vina
2yqgz	15	0.70	0.23	Scoring (structure-based with FE estimates)	Vina
87x7c	15	0.70	0.23	Scoring (structure-based with FE estimates)	AutoDock Vina
g7q2q	15	0.70	0.24	Scoring (structure-based with FE estimates)	AutoDock Vina
dnil6	15	0.75	0.18	Scoring (structure-based with FE estimates)	In-house machine learning score
xr834	15	0.75	0.19	Scoring (structure-based with FE estimates)	In-house machine learning score
fzr1f	15	0.77	0.14	Scoring (structure-based with FE estimates)	SIE (Solvated Interaction Energy)
gzd7a	15	0.79	0.17	Scoring (structure-based with FE estimates)	Smina
pcy26	15	0.80	0.20	Scoring (structure-based with FE estimates)	Molsoft ICM

cwip8	15	0.85	0.22	Scoring (structure-based with FE estimates)	SILCS (Site Identification by Ligand Competitive Saturation)
nrdge	15	0.90	0.19	Scoring (structure-based with FE estimates)	Quasi exact method
l8rmr	15	0.92	0.15	Scoring (structure-based with FE estimates)	LIE (Linear Interaction Energy Model)
mseke	15	0.98	0.24	Scoring (structure-based with FE estimates)	Smina
axxmu	15	1.06	0.19	Scoring (structure-based with FE estimates)	MMPBSA
3qyiy	15	1.06	0.19	Scoring (structure-based with FE estimates)	Amber-MMPBSA
6pf5g	15	1.07	0.26	Scoring (structure-based with FE estimates)	Glide-XP
yprv4	15	1.09	0.19	Scoring (structure-based with FE estimates)	AutoDock Vina, Clusterizer-DockAccessor
ec84y	15	1.15	0.35	Scoring (structure-based with FE estimates)	xScore
dlvjf	15	1.28	0.18	Scoring (structure-based with FE estimates)	qpld, Glide-XP, Glide-SP, AutoDock Vina, AutoDock 4
k2um8	15	1.28	0.18	Scoring (structure-based with FE estimates)	qpld, Glide-XP, Glide-SP, AutoDock Vina, AutoDock 4
keum2	15	1.74	0.27	Scoring (structure-based with FE estimates)	GBVI/WSA scoring
oo7gj	15	1.74	0.27	Scoring (structure-based with FE estimates)	MOE
pbjwu	15	1.85	0.30	Scoring (structure-based with FE estimates)	Ligands ranked manually
4ynsp	15	1.88	0.37	Scoring (structure-based with FE estimates)	Ichem-GRIM, HYDE
0psyy	14	1.91	0.37	Alchemical Free Energy	FESetup, NetworkX
aaveo	15	1.97	0.39	Scoring (structure-based with FE estimates)	Smina*
1t2p5	15	2.17	0.35	Scoring (structure-based with FE estimates)	Ichem-GRIM, HYDE
bolbu	15	2.25	0.20	Alchemical Free Energy	MCPro
lpcmd	15	2.25	0.20	Scoring (ligand-based with FE estimates)	Glide-XP
o3nlu	15	2.25	0.20	Scoring (structure-based with FE estimates)	Glide-XP
tdbsk	15	2.34	0.41	Scoring (structure-based with FE estimates)	SILCS (Site Identification by Ligand Competitive Saturation)
2o6iv	15	2.44	0.43	Scoring (structure-based with FE estimates)	Ichem-GRIM, HYDE
o8bk4	15	2.55	0.29	Alchemical Free Energy	FEP-Desmond*

7hw77	15	2.95	0.87	Alchemical Free Energy	FESetup, NetworkX
a3c8k	15	3.38	1.07	Alchemical Free Energy	FESetup, NetworkX
vskkl	15	3.95	0.73	Scoring (structure-based with FE estimates)	AutoDock Vina
cs2lm	15	4.61	0.65	Scoring (structure-based with FE estimates)	Amber-MMGBSA
8nbnv	15	5.50	1.23	Scoring (structure-based with FE estimates)	MMGBSA

Table S10. $RMSE_c$ statistics and software used for free energy predictions of FE Set 2 during Stage 1. Statistics are calculated for predictions from scoring methods with free energy estimates in the affinity ranking component of the challenge, and free energy methods in the free energy component of the challenge. For each participant, the table lists Receipt ID, number of ligands for which predictions were made, $RMSE_c$ and $RMSE_c$ uncertainty, method type [scoring (structure- vs ligand-based) vs alchemical methods], and software used. Asterisks indicate use of visual inspection.

Receipt ID	Number of Ligands	$RMSE_c$	$RMSE_c$ Uncertainty	Method Type	Software Used
f6een	14	0.98	0.18	Alchemical Free Energy	FESetup, NetworkX
gzd7a	18	1.07	0.17	Scoring (structure-based with FE estimates)	Smina
87x7c	18	1.18	0.16	Scoring (structure-based with FE estimates)	AutoDock Vina
g7q2q	18	1.18	0.16	Scoring (structure-based with FE estimates)	AutoDock Vina
5nsef	18	1.18	0.16	Scoring (structure-based with FE estimates)	Vina
eg8rg	18	1.18	0.16	Scoring (structure-based with FE estimates)	Vina
2efa1	18	1.25	0.23	Scoring (structure-based with FE estimates)	Quasi exact method
l8rmr	18	1.35	0.24	Scoring (structure-based with FE estimates)	LIE (Linear Interaction Energy Model)
sndmm	18	1.36	0.18	Alchemical Free Energy	MCPro
lpcmd	18	1.36	0.19	Scoring (structure-based with FE estimates)	Glide-XP
o3nlu	18	1.36	0.19	Scoring (structure-based with FE estimates)	Glide-XP

yprv4	18	1.36	0.27	Scoring (structure-based with FE estimates)	AutoDock Vina, Clusterizer-DockAccessor
fzr1f	18	1.39	0.19	Scoring (structure-based with FE estimates)	SIE (Solvated Interaction Energy)
cut86	18	1.40	0.19	Scoring (structure-based with FE estimates)	In-house machine learning score
c0l1t	18	1.40	0.19	Scoring (structure-based with FE estimates)	In-house machine learning score
mseke	18	1.49	0.28	Scoring (structure-based with FE estimates)	Smina
2umsq	18	1.50	0.19	Scoring (structure-based with FE estimates)	SILCS (Site Identification by Ligand Competitive Saturation)
6pf5g	18	1.61	0.28	Scoring (structure-based with FE estimates)	Glide-XP
dnil6	18	1.61	0.20	Scoring (structure-based with FE estimates)	In-house machine learning score
b5mvu	18	1.61	0.20	Scoring (structure-based with FE estimates)	In-house machine learning score
cwip8	18	1.69	0.24	Scoring (structure-based with FE estimates)	SILCS (Site Identification by Ligand Competitive Saturation)
pcy26	18	1.69	0.25	Scoring (structure-based with FE estimates)	Molsoft ICM
aaveo	18	1.76	0.33	Scoring (structure-based with FE estimates)	Smina*
ec84y	18	1.79	0.32	Scoring (structure-based with FE estimates)	xScore
1t2p5	18	1.87	0.25	Scoring (structure-based with FE estimates)	Ichem-GRIM, HYDE
ukb3f	18	1.88	0.35	Alchemical Free Energy	FEP-Desmond*
tdbsk	18	1.92	0.40	Scoring (structure-based with FE estimates)	SILCS (Site Identification by Ligand Competitive Saturation)
dlvjf	18	1.92	0.22	Scoring (structure-based with FE estimates)	qpld, Glide-XP, Glide-SP, AutoDock Vina, AutoDock 4
k2um8	18	1.92	0.22	Scoring (structure-based with FE estimates)	qpld, Glide-XP, Glide-SP, AutoDock Vina, AutoDock 4
4ynsp	18	1.96	0.38	Scoring (structure-based with FE estimates)	Ichem-GRIM, HYDE
fxtpq	18	2.02	0.42	Scoring (structure-based with FE estimates)	Ligands ranked manually
2o6iv	18	2.04	0.38	Scoring (structure-based with FE estimates)	Ichem-GRIM, HYDE
j507m	18	2.09	0.37	Scoring (structure-based with FE estimates)	MMPBSA
3qyiy	18	2.09	0.37	Scoring (structure-based with FE estimates)	Amber-MMPBSA
cs2lm	18	2.21	0.35	Scoring (structure-based with FE estimates)	Amber-MMGBSA

pgbuh	18	2.30	0.55	Scoring (structure-based with FE estimates)	CDOCKER (pose prediction) + Autodock Vina (scoring)
keum2	18	3.55	0.76	Scoring (structure-based with FE estimates)	GBVI/WSA scoring
ga5bt	18	3.55	0.76	Scoring (structure-based with FE estimates)	MOE
vskkl	18	3.90	0.67	Scoring (structure-based with FE estimates)	AutoDock Vina
qvnq5	18	4.07	0.72	Alchemical Free Energy	FESetup, NetworkX
fyt14	18	4.23	0.77	Scoring (structure-based with FE estimates)	MMGBSA

Table S11. $RMSE_c$ statistics and software used for free energy predictions of FE Set 1 during Stage 2. Statistics are calculated for predictions from scoring methods with free energy estimates in the affinity ranking component of the challenge, and free energy methods in the free energy component of the challenge. For each participant, the table lists Receipt ID, number of ligands for which predictions were made, $RMSE_c$ and $RMSE_c$ uncertainty, method type [scoring (structure- vs ligand-based) vs alchemical methods], and software used. Asterisks indicate the use of visual inspection.

Receipt ID	Number of Ligands	$RMSE_c$	$RMSE_c$ Uncertainty	Method Type	Software Used
p8rak	15	0.57	0.14	Scoring (structure-based with FE estimates)	MMGBSA modified
li83b	15	0.57	0.14	Scoring (structure-based with FE estimates)	Molsoft ICM
4rbjk	15	0.66	0.21	Scoring (structure-based with FE estimates)	RI-Score
bw4pj	15	0.67	0.21	Scoring (structure-based with FE estimates)	RI-Score
moii1	15	0.68	0.19	Scoring (structure-based with FE estimates)	RI-Score
rwmcz	15	0.68	0.22	Scoring (structure-based with FE estimates)	SILCS (Site Identification by Ligand Competitive Saturation)
wfe3c	15	0.69	0.13	Scoring (structure-based with FE estimates)	Vina
jtsy2	15	0.69	0.19	Scoring (structure-based with FE estimates)	RI-Score
rtqum	15	0.71	0.25	Scoring (structure-based with FE estimates)	RI-Score
hj31e	15	0.72	0.21	Scoring (ligand-based w/ FE estimates)	In-house QSAR script
w8p2f	15	0.72	0.23	Scoring (structure-based with FE estimates)	RI-Score

0aggj	15	0.72	0.21	Scoring (structure-based with FE estimates)	RI-Score
n55eq	15	0.74	0.22	Scoring (structure-based with FE estimates)	RI-Score
gia3u	15	0.74	0.23	Scoring (structure-based with FE estimates)	SILCS (Site Identification by Ligand Competitive Saturation)
86ozh	15	0.75	0.23	Scoring (structure-based with FE estimates)	RI-Score
mtbxd	15	0.77	0.23	Scoring (structure-based with FE estimates)	RI-Score
vbzci	15	0.81	0.23	Scoring (structure-based with FE estimates)	LIE (Linear Interaction Energy Model)
xv7qs	15	0.86	0.15	Scoring (structure-based with FE estimates)	SIE (solvated interaction energy)
3idpo	15	0.95	0.16	Scoring (structure-based with FE estimates)	Quasi-exact method
bqhi0	15	0.96	0.13	Scoring (structure-based with FE estimates)	MMGBSA
jdr8n	15	0.96	0.13	Scoring (structure-based with FE estimates)	MMGBSA
sge2j	15	0.96	0.28	Scoring (structure-based with FE estimates)	CDOCKER (pose prediction) + Autodock Vina (scoring)
6nrpf	15	1.03	0.18	Scoring (structure-based with FE estimates)	Vinardo Score
fww4f	15	1.09	0.34	Scoring (structure-based with FE estimates)	AutoDock Vina, Clusterizer & DockAccessor
ljdjm	15	1.15	0.24	Scoring (structure-based with FE estimates)	Glide-XP
6mjkt	15	1.24	0.16	Scoring (structure-based with FE estimates)	Smina
erdfw	15	1.29	0.26	Scoring (structure-based with FE estimates)	MMPBSA
qpv1a	15	1.29	0.26	Scoring (structure-based with FE estimates)	MMPBSA
66nbk	15	1.47	0.28	Scoring (structure-based with FE estimates)	SIE (solvated interaction energy) with FISH solvation model
ck8kc	15	1.48	0.27	Alchemical Free Energy	Schrodinger-FEP
2612s	15	1.51	0.29	Alchemical Free Energy	Amber
pyxiv	15	1.52	0.25	Alchemical Free Energy	Schrodinger-FEP
g4bd3	15	1.59	0.24	Scoring (structure-based with FE estimates)	Smina, in-house scoring function
olv52	14	1.81	0.46	Alchemical Free Energy	FESetup, NetworkX
lh0zq	15	1.86	0.36	Alchemical Free Energy	CHARMM
327r0	15	1.96	0.26	Scoring (structure-based with FE estimates)	BEDAM (Binding Energy Distribution Analysis Method)
zkx4x	15	2.00	0.34	Alchemical Free Energy	CHARMM

7wx7v	15	2.09	0.38	Scoring (structure-based with FE estimates)	SILCS (Site Identification by Ligand Competitive Saturation)
22juj	15	2.21	0.29	Scoring (structure-based with FE estimates)	BEDAM (Binding Energy Distribution Analysis Method)
eexcx	15	2.24	0.59	Alchemical Free Energy	CHARMM
nszkx	13	2.36	0.58	Alchemical Free Energy	Gromacs
4botu	14	2.38	0.45	Alchemical Free Energy	FESetup, NetworkX
inspj	15	2.39	0.42	Alchemical Free Energy	FESetup, NetworkX
mbbz2	15	3.04	0.82	Scoring (structure-based with FE estimates)	Smina, Vinardo scoring function
07tpe	15	3.76	1.41	Alchemical Free Energy	FESetup, NetworkX
b2xt3	15	4.99	0.73	Scoring (structure-based with FE estimates)	MMGBSA
g18vf	15	5.70	0.94	Scoring (structure-based with FE estimates)	MMGBSA
vxvhq	15	6.24	1.30	Scoring (structure-based with FE estimates)	MMGBSA
e2nsc	15	6.51	0.96	Scoring (structure-based with FE estimates)	MMGBSA
eta0e	15	6.93	1.84	Scoring (structure-based with FE estimates)	QMMM
vovuk	15	10.72	1.80	Scoring (structure-based with FE estimates)	MMGBSA
gzx7q	15	11.13	4.60	Scoring (structure-based with FE estimates)	QMMM
ska85	15	14.32	1.92	Scoring (structure-based with FE estimates)	MMGBSA
veadz	15	18.22	2.58	Scoring (structure-based with FE estimates)	SeeSAR/HYDE
t7htp	15	99.69	26.27	Scoring (structure-based with FE estimates)	Total Energy

Table S12. $RMSE_c$ statistics and software used for free energy predictions of FE Set 2 during Stage 2. Statistics are calculated for predictions from scoring methods with free energy estimates in the affinity ranking component of the challenge, and free energy methods in the free energy component of the challenge. For each participant, the table lists Receipt ID, number of ligands for which predictions were made, $RMSE_c$ and $RMSE_c$ uncertainty, method type [scoring (structure- vs ligand-based) vs alchemical methods], and software used. Asterisks indicate the use of visual inspection.

Receipt ID	Number of Ligands	$RMSE_c$	$RMSE_c$ Uncertainty	Method Type	Software Used

xk67c	18	0.94	0.15	Absolute free energy method	Jarzynski umbrella sampling
ljdjm	18	1.10	0.16	Scoring (structure-based with FE estimates)	Glide-XP
vbzci	18	1.18	0.15	Scoring (structure-based with FE estimates)	LIE (Linear Interaction Energy Model)
li83b	18	1.22	0.17	Scoring (structure-based with FE estimates)	Molsoft ICM
p8rak	18	1.22	0.18	Scoring (structure-based with FE estimates)	MMGBSA_modified
fww4f	18	1.23	0.20	Scoring (structure-based with FE estimates)	AutoDock Vina, Clusterizer & DockAccessor
g4bd3	18	1.24	0.19	Scoring (structure-based with FE estimates)	Smina, in-house scoring function
6mjkt	18	1.26	0.17	Scoring (structure-based with FE estimates)	Smina
c1nbt	14	1.29	0.21	Alchemical Free Energy	FESetup, NetworkX
rtqum	18	1.29	0.17	Scoring (structure-based with FE estimates)	RI-Score
86ozh	18	1.30	0.17	Scoring (structure-based with FE estimates)	RI-Score
81n55	18	1.31	0.21	Alchemical Free Energy	Schrodinger-FEP
sge2j	18	1.34	0.22	Scoring (structure-based with FE estimates)	CDOCKER (pose prediction) + Autodock Vina (scoring)
7wx7v	18	1.34	0.19	Scoring (structure-based with FE estimates)	SILCS (Site Identification by Ligand Competitive Saturation)
mtbxd	18	1.34	0.19	Scoring (structure-based with FE estimates)	RI-Score
0jz8u	14	1.36	0.24	Alchemical Free Energy	FESetup, NetworkX
66nbk	18	1.36	0.28	Scoring (structure-based with FE estimates)	SIE (solvated interaction energy) with FISH solvation model
0aggj	18	1.37	0.19	Scoring (structure-based with FE estimates)	RI-Score
n55eq	18	1.37	0.17	Scoring (structure-based with FE estimates)	RI-Score
w8p2f	18	1.39	0.18	Scoring (structure-based with FE estimates)	RI-Score
rvm67	18	1.39	0.17	Scoring (structure-based with FE estimates)	Quasi-exact method
4rbjk	18	1.40	0.18	Scoring (structure-based with FE estimates)	RI-Score
xv7qs	18	1.44	0.22	Scoring (structure-based with FE estimates)	SIE (solvated interaction energy)
jzrt5	18	1.45	0.22	Alchemical Free Energy	FESetup, NetworkX
bw4pj	18	1.46	0.20	Scoring (structure-based with FE estimates)	RI-Score
adxx8	18	1.46	0.20	Alchemical Free Energy	Amber
k32u2	18	1.47	0.17	Scoring (structure-based with FE estimates)	Quasi-exact method

moii1	18	1.48	0.19	Scoring (structure-based with FE estimates)	RI-Score
rwmcz	18	1.48	0.19	Scoring (structure-based with FE estimates)	SILCS (Site Identification by Ligand Competitive Saturation)
jdr8n	18	1.49	0.21	Scoring (structure-based with FE estimates)	MMGBSA
4pnzi	18	1.49	0.22	Scoring (structure-based with FE estimates)	MMGBSA
x2j7p	18	1.49	0.19	Alchemical Free Energy	Schrodinger-FEP
jtsy2	18	1.49	0.20	Scoring (structure-based with FE estimates)	RI-Score
gia3u	18	1.53	0.19	Scoring (structure-based with FE estimates)	SILCS (Site Identification by Ligand Competitive Saturation)
hj31e	18	1.55	0.19	Scoring (ligand-based w/ FE estimates)	In-house QSAR script
erdfw	18	1.56	0.21	Scoring (structure-based with FE estimates)	MMPBSA
14qhp	18	1.56	0.21	Scoring (structure-based with FE estimates)	MMPBSA
6nrpf	18	1.59	0.23	Scoring (structure-based with FE estimates)	Vinardo Score
izhdb	18	1.61	0.28	Alchemical Free Energy	CHARMM
wfe3c	18	1.65	0.30	Scoring (structure-based with FE estimates)	Vina
mbbz2	18	1.67	0.22	Scoring (structure-based with FE estimates)	Smina, Vinardo scoring function
22juj	18	2.32	0.39	Scoring (structure-based with FE estimates)	BEDAM (Binding Energy Distribution Analysis Method)
327r0	18	2.40	0.37	Scoring (structure-based with FE estimates)	BEDAM (Binding Energy Distribution Analysis Method)
vxvhq	18	2.50	0.38	Scoring (structure-based with FE estimates)	MMGBSA
d8fcq	18	2.56	0.47	Scoring (structure-based with FE estimates)	QMMM
bdcdh	18	3.27	0.75	Scoring (structure-based with FE estimates)	MMGBSA
szjdn	18	3.58	0.90	Scoring (structure-based with FE estimates)	MMGBSA
qt771	18	4.11	0.74	Alchemical Free Energy	FESetup, NetworkX
2ytv8	18	4.37	1.34	Alchemical Free Energy	Gromacs
2etlu	18	7.72	1.03	Scoring (structure-based with FE estimates)	MMGBSA
vovuk	18	13.45	1.71	Scoring (structure-based with FE estimates)	MMGBSA
ska85	18	14.04	2.30	Scoring (structure-based with FE estimates)	MMGBSA
aau3q	18	24.35	4.60	Scoring (structure-based with FE estimates)	QMMM

x7ib3	18	27.72	4.61	Scoring (structure-based with FE estimates)	SeeSAR/HYDE
67a3e	18	73.54	8.21	Scoring (structure-based with FE estimates)	Total Energy

Table S13. Kendall's tau statistics and software used for free energy predictions of FE Set 1 during Stage 1. Statistics are calculated for predictions from all scoring and free energy methods. For each participant, the table lists Receipt ID, number of ligands for which predictions were made, Kendall's tau and Kendall's tau uncertainty, method type [scoring (structure- vs ligand-based) vs alchemical methods], and software used. Asterisks indicate the use of visual inspection.

Receipt ID	Number of Ligands	Kendall's Tau	Kendall's Tau Uncertainty	Method Type	Software Used
5nim5	15	0.44	0.16	Scoring (structure-based)	Smina
nrdge	15	0.39	0.16	Scoring (structure-based)	Quasi exact method
f48cf	15	0.39	0.17	Scoring (structure-based)	Smina
kz0dz	15	0.35	0.16	Scoring (ligand-based)	3D QSAutogrid/R
ukdfw	15	0.35	0.2	Scoring (structure-based)	Smina
r885q	15	0.31	0.16	Scoring (structure-based)	ITScore_v2_TF
3qyiy	15	0.29	0.16	Scoring (structure-based)	Amber-MMPBSA
axxmu	15	0.29	0.16	Scoring (structure-based)	MMPBSA
c20xb	15	0.29	0.19	Scoring (structure-based)	ITScore_TF
h2w3q	15	0.29	0.22	Scoring (structure-based)	SeeSAR
nvoyn	15	0.29	0.18	Scoring (structure-based)	RF-Score-VS, Smina
l8rmr	15	0.27	0.16	Scoring (structure-based)	LIE (Linear Interaction Energy Model)
fvfe7	15	0.23	0.2	Scoring (structure-based)	ITScore_v1
mf3fc	15	0.23	0.17	Scoring (structure-based)	GoldScore, ChemScore, ChemPLP, ASP
sn46i	15	0.23	0.17	Scoring (structure-based)	GoldScore, ChemScore, ChemPLP, ASP, RMSD similarity to reference fragment

tdbsk	15	0.19	0.17	Scoring (structure-based)	SILCS (Site Identification by Ligand Competitive Saturation)
bolbu	15	0.17	0.18	Alchemical free energy	MCPro
lpcmd	15	0.17	0.18	Scoring (ligand-based)	Glide-XP
mzwwt	15	0.17	0.17	Scoring (structure-based)	ITScore_v2
o3nlu	15	0.17	0.19	Scoring (structure-based)	Glide-XP
pgbuh	15	0.17	0.22	Scoring (structure-based)	CDOCKER (pose prediction) + Autodock Vina (scoring)
5f688	15	0.15	0.19	Scoring (structure-based)	gnina_cnn
6pcik	15	0.15	0.24	Scoring (structure-based)	idock-RF-v3*
bhtq4	15	0.15	0.19	Scoring (structure-based)	ITScore_v1_TF
cwip8	15	0.15	0.21	Scoring (structure-based)	SILCS (Site Identification by Ligand Competitive Saturation)
fzr1f	15	0.15	0.16	Scoring (structure-based)	SIE (Solvated Interaction Energy)
vu0d0	15	0.15	0.22	Scoring (structure-based)	ITScore_TF
33a8g	15	0.12	0.15	Scoring (structure-based)	ITScore_v2
5bvwx	15	0.12	0.2	Scoring (structure-based)	In-house machine learning score
c0l1t	15	0.12	0.2	Scoring (structure-based)	In-house machine learning score
1bhkb	15	0.11	0.2	Scoring (structure-based)	Vina
2yqgz	15	0.11	0.2	Scoring (structure-based)	Vina
cs2lm	15	0.08	0.18	Scoring (structure-based)	Amber-MMGBSA
ec84y	15	0.08	0.21	Scoring (structure-based)	xScore
pbjwu	15	0.06	0.2	Scoring (structure-based)	Ligands ranked manually
2umsq	15	0.06	0.18	Scoring (structure-based)	SILCS (Site Identification by Ligand Competitive Saturation)
87x7c	15	0.06	0.19	Scoring (structure-based)	AutoDock Vina
g7q2q	15	0.06	0.19	Scoring (structure-based)	AutoDock Vina
o8bk4	15	0.06	0.17	Alchemical free energy	FEP-Desmond*

0kg1e	15	0.04	0.17	Scoring (structure-based)	Smina
7hw77	15	0.04	0.18	Alchemical free energy	FESetup, NetworkX
8nbnv	15	0.04	0.16	Scoring (structure-based)	MMGBSA
gzd7a	15	0.04	0.19	Scoring (structure-based)	Smina
jr0oc	15	0.04	0.19	Scoring (structure-based)	Gold
ttgw7	15	0.04	0.18	Scoring (structure-based)	Glide
0psyy	14	0.03	0.22	Alchemical free energy	FESetup, NetworkX
dnil6	15	0.02	0.18	Scoring (structure-based)	In-house machine learning score
h1on0	15	0.02	0.2	Scoring (structure-based)	ITScore_v1
xr834	15	0.02	0.18	Scoring (structure-based)	In-house machine learning score
6fzf6	15	0	0.2	Scoring (structure-based)	HADDOCK2.2 web server
bwdnh	15	0	0.19	Scoring (structure-based)	gnina_cnn
4ivv5	15	-0.02	0.19	Scoring (structure-based)	ITScore_v2_TF
mseke	15	-0.02	0.2	Scoring (structure-based)	Smina
a3c8k	15	-0.03	0.2	Alchemical free energy	FESetup, NetworkX
f30wc	15	-0.04	0.18	Scoring (structure-based)	RF-Score-VS, Smina
l0xpi	15	-0.04	0.19	Scoring (structure-based)	Molsoft ICM
q76s3	15	-0.06	0.19	Scoring (structure-based)	ITScore_v1_TF
vskkl	15	-0.06	0.22	Scoring (structure-based)	AutoDock Vina
pcy26	15	-0.08	0.19	Scoring (structure-based)	Molsoft ICM
2o6iv	15	-0.10	0.22	Scoring (structure-based)	Ichem-GRIM, HYDE
4ynsp	15	-0.10	0.22	Scoring (structure-based)	Ichem-GRIM, HYDE
dlvjf	15	-0.10	0.17	Scoring (structure-based)	qpld, Glide-XP, Glide-SP, AutoDock Vina, AutoDock 4
k2um8	15	-0.10	0.18	Scoring (structure-based)	qpld, Glide-XP, Glide-SP, AutoDock Vina, AutoDock 4
1t2p5	15	-0.15	0.17	Scoring (structure-based)	Ichem-GRIM, HYDE
6pf5g	15	-0.15	0.2	Scoring (structure-based)	Glide-XP

a2w8d	15	-0.25	0.15	Scoring (structure-based)	Smina
yprv4	15	-0.25	0.18	Scoring (structure-based)	AutoDock Vina, Clusterizer-DockAccessor
1un0y	15	-0.27	0.17	Scoring (structure-based)	gnina_cnn
keum2	15	-0.37	0.17	Scoring (structure-based)	GBVI/WSA scoring
oo7gj	15	-0.37	0.16	Scoring (structure-based)	MOE
aaveo	15	-0.46	0.18	Scoring (structure-based)	Smina*

Table S14. Kendall's tau statistics and software used for free energy predictions of FES Set 2 during Stage 1. Statistics are calculated for predictions from all scoring and free energy methods. For each participant, the table lists Receipt ID , number of ligands for which predictions were made, Kendall's tau and Kendall's tau uncertainty, method type [scoring (structure- vs ligand-based) vs alchemical methods], and software used. Asterisks indicate the use of visual inspection.

Receipt ID	Number of Ligands	Kendall's Tau	Kendall's Tau Uncertainty	Method Type	Software Used
fvfe7	18	0.51	0.10	Scoring (structure-based)	ITScore_v1
gzd7a	18	0.50	0.12	Scoring (structure-based)	Smina
33a8g	18	0.49	0.14	Scoring (structure-based)	ITScore_v2
6pcik	18	0.49	0.16	Scoring (structure-based)	idock-RF-v3*
q76s3	18	0.49	0.13	Scoring (structure-based)	ITScore_v1_TF
4ivv5	18	0.45	0.15	Scoring (structure-based)	ITScore_v2_TF
4ynsp	18	0.44	0.18	Scoring (structure-based)	Ichem-GRIM, HYDE
ttgw7	18	0.44	0.18	Scoring (structure-based)	Glide
f30wc	18	0.41	0.14	Scoring (structure-based)	RF-Score-VS, Smina
mzwwt	18	0.41	0.17	Scoring (structure-based)	ITScore_v2
mf3fc	18	0.40	0.14	Scoring (structure-based)	GoldScore, ChemScore, ChemPLP, ASP
sn46i	18	0.40	0.14	Scoring (structure-based)	GoldScore, ChemScore, ChemPLP, ASP, RMSD similarity to reference fragment
c20xb	18	0.38	0.18	Scoring (structure-based)	ITScore_TF
sndmm	18	0.37	0.16	Alchemical free energy	MCPro

lpcmd	18	0.37	0.16	Scoring (ligand-based)	Glide-XP
nvoyn	18	0.37	0.14	Scoring (structure-based)	RF-Score-VS, Smina
o3nlu	18	0.37	0.16	Scoring (structure-based)	Glide-XP
3qyiy	18	0.36	0.16	Scoring (structure-based)	Amber-MMPBSA
6fzf6	18	0.36	0.17	Scoring (structure-based)	HADDOCK2.2 web server
yprv4	18	0.36	0.17	Scoring (structure-based)	AutoDock Vina, Clusterizer-DockAccessor
j507m	18	0.36	0.16	Scoring (structure-based)	MMPBSA
5nsef	18	0.35	0.16	Scoring (structure-based)	Vina
eg8rg	18	0.35	0.16	Scoring (structure-based)	Vina
87x7c	18	0.34	0.16	Scoring (structure-based)	AutoDock Vina
g7q2q	18	0.34	0.16	Scoring (structure-based)	AutoDock Vina
l0xpi	18	0.34	0.17	Scoring (structure-based)	Molsoft ICM
tdbsk	18	0.34	0.17	Scoring (structure-based)	SILCS (Site Identification by Ligand Competitive Saturation)
f6een	14	0.34	0.16	Alchemical free energy	FESetup, NetworkX
2efa1	18	0.34	0.15	Scoring (structure-based)	Quasi exact method
c0l1t	18	0.32	0.18	Scoring (structure-based)	In-house machine learning score
fzr1f	18	0.32	0.16	Scoring (structure-based)	SIE (Solvated Interaction Energy)
cut86	18	0.31	0.18	Scoring (structure-based)	In-house machine learning score
bwdnh	18	0.30	0.19	Scoring (structure-based)	gnina_cnn
2o6iv	18	0.29	0.14	Scoring (structure-based)	Ichem-GRIM, HYDE
aaveo	18	0.29	0.16	Scoring (structure-based)	Smina*
l8rmr	18	0.29	0.18	Scoring (structure-based)	LIE (Linear Interaction Energy Model)
cs2lm	18	0.28	0.16	Scoring (structure-based)	Amber-MMGBSA
mseke	18	0.28	0.18	Scoring (structure-based)	Smina
fxtpq	18	0.24	0.20	Scoring (structure-based)	Ligands ranked manually
ukb3f	18	0.24	0.18	Alchemical free energy	FEP-Desmond*
5f688	18	0.21	0.18	Scoring (structure-based)	gnina_cnn
6pf5g	18	0.20	0.18	Scoring (structure-based)	Glide-XP

ec84y	18	0.20	0.20	Scoring (structure-based)	xScore
1un0y	18	0.18	0.18	Scoring (structure-based)	gnina_cnn
f48cf	18	0.18	0.20	Scoring (structure-based)	Smina
1t2p5	18	0.17	0.18	Scoring (structure-based)	Ichem-GRIM, HYDE
a2w8d	18	0.15	0.15	Scoring (structure-based)	Smina
h1on0	18	0.15	0.18	Scoring (structure-based)	ITScore_v1
vu0d0	18	0.15	0.19	Scoring (structure-based)	ITScore_TF
fyt14	18	0.11	0.18	Scoring (structure-based)	MMGBSA
0kg1e	18	0.07	0.20	Scoring (structure-based)	Smina
keum2	18	0.07	0.18	Scoring (structure-based)	GBVI/WSA scoring
ga5bt	18	0.07	0.18	Scoring (structure-based)	MOE
r885q	18	0.05	0.20	Scoring (structure-based)	ITScore_v2_TF
bhtq4	18	-0.03	0.18	Scoring (structure-based)	ITScore_v1_TF
h2w3q	18	-0.03	0.16	Scoring (structure-based)	SeeSAR
pcy26	18	-0.08	0.17	Scoring (structure-based)	Molsoft ICM
dnil6	18	-0.11	0.17	Scoring (structure-based)	In-house machine learning score
b5mvu	18	-0.11	0.16	Scoring (structure-based)	In-house machine learning score
qvnq5	18	-0.14	0.20	Alchemical free energy	FESetup, NetworkX
pgbuh	18	-0.16	0.15	Scoring (structure-based)	CDOCKER (pose prediction) + Autodock Vina (scoring)
k2um8	18	-0.22	0.17	Scoring (structure-based)	qpld, Glide-XP, Glide-SP, AutoDock Vina, AutoDock 4
kz0dz	18	-0.24	0.19	Scoring (ligand-based)	3D QSAutogrid/R
5nim5	18	-0.25	0.16	Scoring (structure-based)	Smina
ukdfw	18	-0.26	0.16	Scoring (structure-based)	Smina
vskkl	18	-0.30	0.14	Scoring (structure-based)	AutoDock Vina
cwip8	18	-0.33	0.19	Scoring (structure-based)	SILCS (Site Identification by Ligand Competitive Saturation)
2umsq	18	-0.34	0.17	Scoring (structure-based)	SILCS (Site Identification by Ligand Competitive Saturation)
dlvjf	18	-0.36	0.14	Scoring (structure-based)	qpld, Glide-XP, Glide-SP, AutoDock Vina, AutoDock 4
jr0oc	18	-0.51	0.13	Scoring (structure-based)	Gold

too1u	5	-0.60	0.37	Scoring (structure-based)	Rhodium HTS

Table S15. Kendall's tau statistics and software used for free energy predictions of FE Set 1 during Stage 2. Statistics are calculated for predictions from all scoring and free energy methods. For each participant, the table lists Receipt ID, number of ligands for which predictions were made, Kendall's tau and Kendall's tau uncertainty, method type [scoring (structure- vs ligand-based) vs alchemical methods], and software used. Asterisks indicate the use of visual inspection.

Receipt ID	Number of Ligands	Kendall's Tau	Kendall's Tau Uncertainty	Method Type	Software Used
4rbjk	15	0.41	0.17	Scoring (structure-based)	RI-Score
bw4pj	15	0.35	0.16	Scoring (structure-based)	RI-Score
jtsy2	15	0.33	0.19	Scoring (structure-based)	RI-Score
n55eq	15	0.33	0.18	Scoring (structure-based)	RI-Score
0aggj	15	0.31	0.17	Scoring (structure-based)	RI-Score
0zno2	15	0.31	0.19	Scoring (structure-based)	Smina, CNN Model Scoring
pr2fp	15	0.31	0.18	Scoring (structure-based)	ITScore_v2
tgmx1	15	0.31	0.19	Scoring (structure-based)	Smina, CNN Model Scoring
eta0e	15	0.31	0.18	Scoring (structure-based)	QMMM
35yg0	15	0.29	0.20	Scoring (structure-based)	Rhodium HTS
15ub2	15	0.27	0.19	Scoring (structure-based)	Smina
20shi	15	0.25	0.17	Scoring (structure-based)	ITScore_v2_TF
hj31e	15	0.25	0.15	Scoring (ligand-based)	In-house QSAR script
li83b	15	0.25	0.21	Scoring (structure-based)	Molsoft ICM
p8rak	15	0.25	0.21	Scoring (structure-based)	MMGBSA_modified
rtqum	15	0.25	0.19	Scoring (structure-based)	RI-Score
nszkx	13	0.25	0.19	Alchemical free energy	Gromacs
bxfza	15	0.23	0.21	Scoring (structure-based)	Smina, CNN Model Scoring
hekzj	15	0.23	0.19	Scoring (structure-based)	Smina, CNN Model Scoring
pny4b	15	0.23	0.19	Scoring (structure-based)	ITScore_TF

0pgek	15	0.21	0.18	Scoring (structure-based)	Ligands ranked manually
21jpg	15	0.21	0.19	Scoring (ligand-based)	3D QSAutogrid/R
fo0p4	15	0.21	0.21	Scoring (structure-based)	Smina
vxvhq	15	0.21	0.21	Scoring (structure-based)	MMGBSA
xkgbb	15	0.21	0.20	Scoring (structure-based)	Smina, CNN Model Scoring
5hri3	15	0.19	0.17	Scoring (structure-based)	ITScore_v1
6fsmu	15	0.19	0.2	Scoring (structure-based)	Smina, CNN Model Scoring
6nrpf	15	0.19	0.23	Scoring (structure-based)	Vinardo Score
7wx7v	15	0.19	0.18	Scoring (structure-based)	SILCS (Site Identification by Ligand Competitive Saturation)
86ozh	15	0.19	0.19	Scoring (structure-based)	RI-Score
pyxiv	15	0.17	0.21	Alchemical free energy	Schrodinger-FEP
wfe3c	15	0.17	0.22	Scoring (structure-based)	Vina
4botu	14	0.16	0.16	Alchemical free energy	FESetup, NetworkX
bqhi0	15	0.15	0.19	Scoring (structure-based)	MMGBSA
4m842	15	0.15	0.19	Scoring (ligand-based)	3D QSAutogrid/R
6mjkt	15	0.15	0.18	Scoring (structure-based)	Smina
gia3u	15	0.15	0.18	Scoring (structure-based)	SILCS (Site Identification by Ligand Competitive Saturation)
gnu3p	15	0.15	0.18	Scoring (ligand-based)	ChemmineR
jdr8n	15	0.15	0.19	Scoring (structure-based)	MMGBSA
moii1	15	0.15	0.18	Scoring (structure-based)	RI-Score
w8p2f	15	0.15	0.19	Scoring (structure-based)	RI-Score
y5gbl	15	0.15	0.19	Scoring (structure-based)	Glide-SP, GIST scoring 3
7c2tp	15	0.14	0.19	Scoring (structure-based)	Smina
aewpt	15	0.14	0.19	Scoring (structure-based)	Smina
awgis	15	0.14	0.21	Scoring (structure-based)	Glide-SP, GIST (solvent displacement score) scoring 1
qcbm1	15	0.14	0.18	Scoring (structure-based)	Xscore
qokw3	15	0.14	0.2	Scoring (structure-based)	Xscore
rwmcz	15	0.14	0.16	Scoring (structure-based)	SILCS (Site Identification by Ligand Competitive Saturation)

3idpo	15	0.14	0.19	Scoring (structure-based)	Quasi-exact method
b2xt3	15	0.14	0.19	Scoring (structure-based)	MMGBSA
inspj	15	0.12	0.15	Alchemical free energy	FESetup, NetworkX
gzx7q	15	0.12	0.19	Scoring (structure-based)	QMMM
uyp2k	15	0.12	0.2	Scoring (structure-based)	ITScore_v1_TF
xe4uo	15	0.12	0.18	Scoring (structure-based)	Smina
olv52	14	0.11	0.18	Alchemical free energy	FESetup, NetworkX
rjyhz	15	0.10	0.22	Scoring (structure-based)	GoldScore, ChemScore, ChemPLP, ASP, lowestRMSDtoRefFragment
bckkg	15	0.08	0.22	Scoring (structure-based)	Glide-SP
bxn4v	15	0.08	0.18	Scoring (structure-based)	ITScore_v2
riyj0	15	0.08	0.21	Scoring (structure-based)	Smina, CNN Model Scoring
sb1dg	15	0.08	0.18	Scoring (structure-based)	ITScore_v1
g18vf	15	0.08	0.17	Scoring (structure-based)	MMGBSA
e2nsc	15	0.08	0.19	Scoring (structure-based)	MMGBSA
qpv1a	15	0.06	0.20	Scoring (structure-based)	MMPBSA
eexcx	15	0.06	0.19	Alchemical free energy	CHARMM
erdfw	15	0.06	0.20	Scoring (structure-based)	MMPBSA
hxm7v	15	0.06	0.21	Scoring (structure-based)	Glide-SP, GIST scoring 2
sge2j	15	0.06	0.21	Scoring (structure-based)	CDOCKER (pose prediction) + Autodock Vina (scoring)
vbzci	15	0.06	0.19	Scoring (structure-based)	LIE (Linear Interaction Energy Model)
07tpe	15	0.04	0.18	Alchemical free energy	FESetup, NetworkX
a8r3i	15	0.04	0.19	Scoring (structure-based)	ITScore_v1_TF
f2wjs	15	0.04	0.22	Scoring (structure-based)	Smina
p2wkt	15	0.04	0.20	Scoring (structure-based)	Chemscore
x8jt0	15	0.04	0.20	Scoring (structure-based)	eTOX ALLIES
xv7qs	15	0.04	0.19	Scoring (structure-based)	SIE (solvated interaction energy)
0qoih	15	0.02	0.17	Scoring (structure-based)	Xscore

66nbk	15	0.02	0.19	Scoring (structure-based)	SIE (solvated interaction energy) with FISH solvation model
cvxs6	15	0.02	0.21	Scoring (structure-based)	Smina, Random Forest Model
2612s	15	0.02	0.20	Alchemical free energy	Amber
22juj	15	0	0.24	Scoring (structure-based)	BEDAM (Binding Energy Distribution Analysis Method)
ck8kc	15	0	0.21	Alchemical free energy	Schrodinger-FEP
327r0	15	-0.02	0.23	Scoring (structure-based)	BEDAM (Binding Energy Distribution Analysis Method)
5zcmb	15	-0.02	0.19	Scoring (structure-based)	Rhodium HTS
hvjjb	15	-0.02	0.20	Scoring (structure-based)	Smina, CNN Model Scoring
naex2	15	-0.02	0.18	Scoring (ligand-based)	In-house QSAR script
ska85	15	-0.02	0.21	Scoring (structure-based)	MMGBSA
jajar	15	-0.04	0.17	Scoring (structure-based)	ITScore_v2_TF
zkx4x	15	-0.04	0.19	Alchemical free energy	CHARMM
0f7u7	15	-0.06	0.18	Scoring (structure-based)	PRODIGY webserver
dh2du	15	-0.06	0.22	Scoring (structure-based)	SeeSAR, HYDE, MMGBSA
g4bd3	15	-0.06	0.20	Scoring (structure-based)	Smina, in-house scoring function
mtbxd	15	-0.06	0.21	Scoring (structure-based)	RI-Score
veadz	15	-0.06	0.22	Scoring (structure-based)	SeeSAR/HYDE
15tvu	15	-0.08	0.18	Scoring (structure-based)	ITScore_TF
mbbz2	15	-0.08	0.16	Scoring (structure-based)	Smina, Vinardo scoring function
v4uoy	15	-0.10	0.23	Scoring (structure-based)	Rhodium HTS
35sno	15	-0.12	0.23	Scoring (structure-based)	Rhodium HTS
t7htp	15	-0.12	0.19	Scoring (structure-based)	Total Energy
lh0zq	15	-0.14	0.2	Alchemical free energy	CHARMM
vovuk	15	-0.14	0.24	Scoring (structure-based)	MMGBSA
ljdjm	15	-0.15	0.19	Scoring (structure-based)	Glide-XP
tbxzq	15	-0.17	0.21	Scoring (structure-based)	Glide
hgqci	15	-0.23	0.18	Scoring (structure-based)	Smina
fww4f	15	-0.27	0.21	Scoring (structure-based)	AutoDock Vina, Clusterizer & DockAccessor

Table S16. Kendall's tau statistics and software used for free energy predictions of FE Set 2 during Stage 2. Statistics are calculated for predictions from all scoring and free energy methods. For each participant, the table lists Receipt ID, number of ligands for which predictions were made, Kendall's tau and Kendall's tau uncertainty, method type [scoring (structure- vs ligand-based) vs alchemical methods], and software used. Asterisks indicate the use of visual inspection.

Receipt ID	Number of Ligands	Kendall's Tau	Kendall's Tau Uncertainty	Method Type	Software Used
xk67c	18	0.62	0.12	Absolute Free Energy	Jarzynski umbrella sampling
hj31e	18	0.55	0.12	Scoring (ligand-based)	In-house QSAR script
81n55	18	0.55	0.12	Alchemical free energy	Schrodinger-FEP
ljdjm	18	0.53	0.12	Scoring (structure-based)	Glide-XP
67a3e	18	0.53	0.15	Scoring (structure-based)	Total Energy
qokw3	18	0.51	0.11	Scoring (structure-based)	Xscore
tbxzq	18	0.51	0.12	Scoring (structure-based)	Glide
x2j7p	18	0.50	0.11	Alchemical free energy	Schrodinger-FEP
naex2	18	0.50	0.16	Scoring (ligand-based)	In-house QSAR script
sb1dg	18	0.50	0.13	Scoring (structure-based)	ITScore_v1
adxx8	18	0.48	0.13	Alchemical free energy	Amber
6mjkt	18	0.45	0.13	Scoring (structure-based)	Smina
vovuk	18	0.45	0.15	Scoring (structure-based)	MMGBSA
rtqum	18	0.44	0.14	Scoring (structure-based)	RI-Score
sge2j	18	0.44	0.14	Scoring (structure-based)	CDOCKER (pose prediction) + Autodock Vina (scoring)
7wx7v	18	0.42	0.17	Scoring (structure-based)	SILCS (Site Identification by Ligand Competitive Saturation)
cvxs6	18	0.42	0.14	Scoring (structure-based)	Smina, Random Forest Model
gnu3p	18	0.42	0.12	Scoring (ligand-based)	ChemmineR
fww4f	18	0.41	0.18	Scoring (structure-based)	AutoDock Vina, Clusterizer & DockAccessor
g4bd3	18	0.41	0.13	Scoring (structure-based)	Smina, in-house scoring function

y5gbl	18	0.41	0.12	Scoring (structure-based)	Glide-SP, GIST scoring 3
li83b	18	0.40	0.14	Scoring (structure-based)	Molsoft ICM
mtbxd	18	0.40	0.16	Scoring (structure-based)	RI-Score
p8rak	18	0.40	0.14	Scoring (structure-based)	MMGBSA_modified
hxm7v	18	0.38	0.15	Scoring (structure-based)	Glide-SP, GIST scoring 2
jzrt5	18	0.37	0.14	Alchemical free energy	FESetup, NetworkX
d8fcq	18	0.37	0.16	Scoring (structure-based)	QMMM
0jz8u	14	0.36	0.17	Alchemical free energy	FESetup, NetworkX
0aggj	18	0.36	0.17	Scoring (structure-based)	RI-Score
66nbk	18	0.36	0.19	Scoring (structure-based)	SIE (solvated interaction energy) with FISH solvation model
86ozh	18	0.36	0.13	Scoring (structure-based)	RI-Score
f2wjs	18	0.36	0.15	Scoring (structure-based)	Smina
awgis	18	0.34	0.16	Scoring (structure-based)	Glide-SP, GIST (solvent displacement score) scoring 1
dh2du	18	0.34	0.17	Scoring (structure-based)	SeeSAR, HYDE, MMGBSA
x7ib3	18	0.34	0.18	Scoring (structure-based)	SeeSAR/HYDE
0f7u7	18	0.33	0.14	Scoring (structure-based)	PRODIGY webserver
bxn4v	18	0.33	0.17	Scoring (structure-based)	ITScore_v2
vbzci	18	0.33	0.16	Scoring (structure-based)	LIE (Linear Interaction Energy Model)
bckkg	18	0.32	0.19	Scoring (structure-based)	Glide-SP
pr2fp	18	0.32	0.14	Scoring (structure-based)	ITScore_v2
vxvhq	18	0.30	0.17	Scoring (structure-based)	MMGBSA
c1nbt	14	0.30	0.20	Alchemical free energy	FESetup, NetworkX
5zcmb	18	0.29	0.16	Scoring (structure-based)	Rhodium HTS
w8p2f	18	0.29	0.16	Scoring (structure-based)	RI-Score
izhdb	18	0.28	0.17	Alchemical free energy	CHARMM
mbbz2	18	0.28	0.14	Scoring (structure-based)	Smina, Vinardo scoring function
n55eq	18	0.28	0.17	Scoring (structure-based)	RI-Score

riyj0	18	0.28	0.19	Scoring (structure-based)	Smina, CNN Model Scoring
szjdn	18	0.28	0.18	Scoring (structure-based)	MMGBSA
5hri3	18	0.26	0.20	Scoring (structure-based)	ITScore_v1
rvm67	18	0.26	0.18	Scoring (structure-based)	Quasi-exact method
22juj	18	0.25	0.15	Scoring (structure-based)	BEDAM (Binding Energy Distribution Analysis Method)
7c2tp	18	0.25	0.19	Scoring (structure-based)	Smina
aewpt	18	0.25	0.18	Scoring (structure-based)	Smina
xe4uo	18	0.25	0.20	Scoring (structure-based)	Smina
14qhp	18	0.24	0.17	Scoring (structure-based)	MMPBSA
bdcdh	18	0.24	0.22	Scoring (structure-based)	MMGBSA
4rbjk	18	0.24	0.18	Scoring (structure-based)	RI-Score
erdfw	18	0.24	0.17	Scoring (structure-based)	MMPBSA
p2wkt	18	0.24	0.18	Scoring (structure-based)	Chemscore
wfe3c	18	0.24	0.19	Scoring (structure-based)	Vina
xkgbb	18	0.24	0.18	Scoring (structure-based)	Smina, CNN Model Scoring
0qoih	18	0.22	0.18	Scoring (structure-based)	Xscore
fo0p4	18	0.22	0.21	Scoring (structure-based)	Smina
0pgek	18	0.21	0.18	Scoring (structure-based)	Ligands ranked manually
bxfza	18	0.21	0.19	Scoring (structure-based)	Smina, CNN Model Scoring
4pnzi	18	0.18	0.18	Scoring (structure-based)	MMGBSA
jdr8n	18	0.18	0.18	Scoring (structure-based)	MMGBSA
k32u2	18	0.18	0.19	Scoring (structure-based)	Quasi-exact method
hgqci	18	0.17	0.19	Scoring (structure-based)	Smina
6nrpf	18	0.16	0.20	Scoring (structure-based)	Vinardo Score
uyp2k	18	0.16	0.15	Scoring (structure-based)	ITScore_v1_TF
xv7qs	18	0.16	0.17	Scoring (structure-based)	SIE (solvated interaction energy)
6fsmu	18	0.15	0.16	Scoring (structure-based)	Smina, CNN Model Scoring
v4uoy	18	0.15	0.19	Scoring (structure-based)	Rhodium HTS

hekzj	18	0.13	0.18	Scoring (structure-based)	Smina, CNN Model Scoring
hvjjb	18	0.13	0.18	Scoring (structure-based)	Smina, CNN Model Scoring
20shi	18	0.12	0.20	Scoring (structure-based)	ITScore_v2_TF
a8r3i	18	0.12	0.17	Scoring (structure-based)	ITScore_v1_TF
jtsy2	18	0.12	0.20	Scoring (structure-based)	RI-Score
2etlu	18	0.11	0.15	Scoring (structure-based)	MMGBSA
15ub2	18	0.11	0.18	Scoring (structure-based)	Smina
35yg0	18	0.11	0.20	Scoring (structure-based)	Rhodium HTS
rwmcz	18	0.11	0.17	Scoring (structure-based)	SILCS (Site Identification by Ligand Competitive Saturation)
bw4pj	18	0.09	0.15	Scoring (structure-based)	RI-Score
moii1	18	0.09	0.13	Scoring (structure-based)	RI-Score
35sno	18	0.05	0.20	Scoring (structure-based)	Rhodium HTS
jajar	18	0.04	0.18	Scoring (structure-based)	ITScore_v2_TF
327r0	18	0.01	0.16	Scoring (structure-based)	BEDAM (Binding Energy Distribution Analysis Method)
qcbm1	18	0.01	0.20	Scoring (structure-based)	Xscore
tgmx1	18	0.01	0.19	Scoring (structure-based)	Smina, CNN Model Scoring
ska85	18	0	0.19	Scoring (structure-based)	MMGBSA
0zno2	18	-0.03	0.19	Scoring (structure-based)	Smina, CNN Model Scoring
15tvu	18	-0.05	0.20	Scoring (structure-based)	ITScore_TF
pny4b	18	-0.12	0.19	Scoring (structure-based)	ITScore_TF
qt771	18	-0.14	0.20	Alchemical free energy	FESetup, NetworkX
2ytv8	18	-0.15	0.22	Alchemical free energy	Gromacs
rjyhz	18	-0.15	0.17	Scoring (structure-based)	GoldScore, ChemScore, ChemPLP, ASP, lowestRMSDtoRefFragment
aau3q	18	-0.17	0.19	Scoring (structure-based)	QMMM
21jpg	18	-0.21	0.17	Scoring (ligand-based)	3D QSAutogrid/R
gia3u	18	-0.21	0.19	Scoring (structure-based)	SILCS (Site Identification by Ligand Competitive Saturation)
x8jt0	18	-0.21	0.14	Scoring (structure-based)	eTOX ALLIES

4m842	18	-0.22	0.17	Scoring (ligand-based)	3D QSAutogrid/R

Table S17. List of non-anonymous participants and their corresponding receipt IDs.

First Name	Last Name	Organization	Receipt ID
Alexandre	Bonvin	Utrecht University	gnu3p, 0f7u7 , 6fzf6 , qxrku ,
Alfonso T.	Garcia-Sosa	University of Tartu, Estonia	k2um8 , dlvjf , am6an
Andrey	Voronkov	Digital BioPharm ltd.	g4bd3, mbbz2 , cvxs6
Anne	Brown	Virginia Tech	vskkl
Ashutosh	Kumar	RIKEN, Japan	tbxzq , ttgw7 , 5rqrx
Bentley	Wingert	University of Pittsburgh	rvm67, k32u2 , 2efa1 , 3idpo , nrdge , fo0p4 , hgqci , f2wjs , 15ub2 , 7c2tp , aewpt , 0pgek , xe4uo , a2w8d , f48cf, ukdfw , 5nim5 , 0kg1e , 6tnqb , h67ea
Bogdan	Iorga	CNRS-ICSN	2ytv8 , rjyhz , mf3fc , sn46i , nszkx
Christina	Athanasiou	Biomedical Research Foundation, Academy of Athens	x2j7p , ukb3f , pyxiv , o8bk4 , oky3v
David	Koes	University of Pittsburgh	6nrpf , wfe3c , mseke , 6fsmu , hvjjb , bxfza , hekzj , 0zno2 , xkgbb , tgmx1 , riyj0 , bwdnh , 1un0y , 5f688 , to4tt , qt334 , 0xk7w , mx1v2
Didier	ROGNAN	University of Strasbourg	2o6iv , 1t2p5 , 4ynsp , tbuot , utpcl , yyjmb
doha	naga	Paris Diderot University	87x7c , eg8rg , 1bhkb
Enrico	Purisima	National Research Council Canada	xv7qs , 66nbk , fzr1f , jz0em
Flavio	Ballante	Washington University School of Medicine in St. Louis	fww4f , yprv4 , 21jpg , 4m842 , kz0dz , 5cf33

Ho Leung	Ng		6pcik , 7rmfb
Hongwu	Wang		hciq4
Jamal	Shamsara	Mashhad University of Medicula Sciences	p2wkt
Jonathan	Bohmann	Southwest Research Institute	5zcmb , 35sno , 35yg0 , v4uoy , 8mocv , too1u
Julien	Michel	University of Edinburgh	jzrt5 , qt771 , fxtpq , qvnq5 , inspj , 07tpe , a3c8k , pbjwu , 7hw77 , 0psyy , 0jz8u , c1nbt , 4botu , olv52 , f6een
Martin	Olsson	Lund University	adxx8 , 2612s
Matthew	Baumgartner	Eli Lilly and Company	6mjkt , hj31e , vovuk , aaveo , gzd7a , naex2 , wax1j , x7jp3 , piwlh , qfu33 , psiuj
Maxim	Totrov	Molsoft	li83b , p8rak , pcy26 , l0xpi , ixnzu
Oleksandr	Yakovenko	ifowonco	vbzci , xk67c , l8rmr , txyzj
Olivier	Bequignon	Université Paris Diderot	g7q2q , 5nsef , 2yqgz
Pär	Söderhjelm	Lund University	0lxp5 , byf51 , l835g
Pinaki	Saha	Shiv Nadar University	keum2 , ga5bt , oo7gj , m00nf
Rajat Kumar	Pal	CUNY	327r0 , 22juj
Sergei	Grudinin	Inria / CNRS	elm2r
Sirish	Lakkaraju	University of Maryland, Baltimore	7wx7v , rwmcz , gia3u , tdbsk , 2umsq , cwip8 , 00ulb
Stefano	Moro	UNIPD	vxvhq , cs2lm , jr0oc , gfifa , knz3v
Steven	Ramsey	Lehman College	awgis , bckkg , hxm7v , y5gbl
Wei	Yang	Florida State University	lh0zq , eexcx
Xiaoqin	Zou	University of Missouri - Columbia	erdfw , jdr8n , 14qhp , 4pnzi , 3qyiy , j507m , qpv1a , bqhi0 , axxmu , jajar , 15tvu , bxn4v , 5hri3 , a8r3i , 20shi , sb1dg , uyp2k , pny4b , pr2fp , c20xb , fvfe7 , q76s3 , 33a8g , 4ivv5 , h1on0 ,

			bhtq4 , mzwwt , r885q , vu0d0 , mmj0e , gdhv4 , 4i2mb , 3rfl5 , mgxbc , qbvvg
Xinqiang	Ding	University of Michigan	sge2j , izhdb , pgbuh , zkx4x , qfrw4 ,
Yuan	Hu	Merck Research Laboratories	bdcdh , szjdn , fyt14 , b2xt3 , g18vf , 8nbnv , 7ltmc , 6xixa
Zhaoping	Xiong	ShanghaiTech University	lpcmd , o3nlu , sndmm , bolbu , 3rj71

References

1. Duan J, Dixon SL, Lowrie JF, Sherman W (2010) Analysis and comparison of 2D fingerprints: Insights into database screening performance using eight fingerprint methods. J Mol Graph Model 29:157–170. doi: 10.1016/j.jmgm.2010.05.008

2. Sastry M, Lowrie JF, Dixon SL, Sherman W (2010) Large-Scale Systematic Analysis of 2D Fingerprint Methods and Parameters to Improve Virtual Screening Enrichments. J Chem Inf Model 50:771–784. doi: 10.1021/ci100062n

Supplementary text

Submission and validation of prediction files

To enable automated processing and basic validation of their submissions (such as verification of the presence of chemically valid ligands in pose predictions), participants were required to generate files with specific formats, much as done in GC2015, but with a few changes and additions [1]. In GC2, participants were not required to follow a naming convention for their submissions as was done with GC2015 to identify the challenge components. Instead, the submission webpage provided radio buttons to specify the challenge component to be submitted: pose prediction, affinity rankings, relative binding free energy calculation. Unlike GC2015, participants were also given the option of submitting a supplementary information folder ("suppinfo") to include any additional information pertaining to their methodology or calculations. Additionally, a MOL file validation was also added to the initial validation stage that required MOL files to be V2000 compliant. Submitters were immediately notified if a submission did not pass basic validation tests, and were able to consult with project staff and to delete any problematic files and substitute corrected ones, until the submission deadline.

Structural re-refinement

The program sf_convert was used to extract required diffraction data and convert it to the appropriate refinement program file formats. The program eLBOW (PHENIX 1.11.1-2575), with the AM1 optimizer [2] and Mogul [3] reference values from the CSD database, was used to generate restraints for the ligands used in co-crystal structure refinement based on contributor-provided SMILES strings, and the results were inspected manually to ensure options "-opt" (use AM1 optimizer) and "-mogul" (use Mogul reference) interoperate appropriately.

A python script named "D3R_refine" was used to automate the entire refinement process, utilizing three different refinement programs including PHENIX [4], REFMAC [5], and BUSTER [6], with facilitating CCP4 package utilities [7]. Results from different programs were assessed and compared to obtain the model of the best quality. Individual anisotropic B-factor refinement was used when the ratio of reflections/atom>30. Individual isotropic B-factor refinement was used when the ratio of reflections/atom<13. Between 30 and 13 reflections/atom, both isotropic and anisotropic B-factor refinement approaches were tested, using automated weighting, and the final B-factor model was selected on the basis of the lowest free R-factor (Rfree) [8].

Non-Crystallographic Symmetry (NCS) refinement was not used for structures with resolution better than 1.7 Å, and only local NCS restraints were used during refinement of lower resolution structures (>2.0 Å). For resolution limits in the range of 1.7-2.0 Å, refinements with and without NCS were evaluated, and the best model was selected on the basis of refinement statistics. Because NCS can affect the value of Rfree, we used multiple statistics to judge model quality, as reflected in the final wwPDB validation

reports. Translation-Libration-Screw (TLS) corrections [9] were used to improve outcomes in the final stages of refinement. Typically, 15 cycles of TLS refinement were computed. TLS rigid groups were selected for each chain when using REFMAC for refinement. Auto TLS options were invoked for refinement by PHENIX and BUSTER. Weight matrices were, in most cases, automatically calculated by each of the three refinement programs. Rarely, manual adjustments of the weight matrix were required to obtain better protein geometry. Although hydrogen atoms make little contribution to protein crystal diffraction data, they were included during refinement.

Following initial refinement of the protein structure, the best-refined protein model from the three parallel refinements (i.e., lowest R-free values) was selected, and polypeptide chain geometry validations were performed as described below. Following any necessary manual corrections to that atomic model, bound ligand structures were built and refined to convergence. Ligand geometry was validated as described below. In some cases, ligand geometry restraints files were modified to ensure appropriate ligand stereochemistry. In the final round of refinement, water molecules were automatically added to the atomic model, using default PHENIX criteria. Water molecules found in the vicinity of bound ligands were manually checked.

For each refined structure, a wwPDB Validation Report [10] was generated and reviewed, with particular attention to the value of Rfree, non-bonded atom-atom clashes assessed by a scaled Clashscore [11], % Ramachandran Outliers [12], % Sidechain Rotamer Outliers [11], and % Real Space R-factor Z-Score (RSRZ) Outliers [13]. Extreme outliers were examined and errors corrected as required. The Real Space R factor (RSR) [14], Real Space Correlation Coefficient (RSCC) [15], and Real Space Z-Score of Difference map (RSZD) [16] were calculated using the DCC program for main chain and sidechains at the individual residue level. Difference electron density maps were examined and any problematic residues rebuilt using COOT [17].

Mogul [3] was used to validate bound ligand geometric features, including bond lengths, bond angles, planarity, and chirality using small molecule X-ray crystal structures in the Cambridge Structural Database [18]. Geometric restraints were modified and the ligand structures re-refined until satisfactory outcomes were obtained, based on visual inspection. Electron density map features corresponding to bound ligands were evaluated using RSR, RSCC and RSZD (RSZD+ and RSZD-), Occupancy Weighted Average B factor (OWAB) quality metrics, and errors corrected using COOT. For higher resolution structures, any alternate conformations were detected by visual inspection of electron density maps and are then included in structure refinements. Finally, intermolecular interactions between protein and bound ligands were visually inspected for plausibility, with particular focus on pseudo-symmetric ligands that could be modeled in opposite orientations without penalty in terms of electron density map quality measures.

Quality improvements were observed for all structures following re-refinement. wwPDB validation reports are available for every PDB structure (https://www.wwpdb.org/validation/2016/XrayValidationReportHelp), and the overall quality at a glance presents five quality indicators, each representing the percentile of the structure relative to the entire PDB archive (e.g., 70% in Rfree means the Rfree of

the structure is better than 70% of the entire archive, with higher value indicating better quality). Compared with the Roche structures, the post-refinement Rfree percentiles improved, which indicates that the post-refinement models fit the diffraction data better. In terms of protein geometry, the post-refinement percentiles for Clashscore and sidechain Rotamer outliers also improved, indicating a better atomic arrangement and sidechain orientation, including those at the ligand binding site. The overall quality of any structure in the PDB can be expressed using a one-dimensional (1-D) metric [19]. This metric was calculated for both the Roche structures and the post-refinement structures, and is provided in the last two columns Table S5. On average, the 1-D metrics improved by 27%. For example, Roche 1-D metric for 1btoj=35th percentile (i.e., better than ~35% of PDB archive) *versus* post-refinement 1-D metric for 1Q0I=73rd percentile (i.e., better than ~73% of PDB archive).

In terms of ligands, better fitting between ligand atomic model and electron density was also observed, as judged by improvements real-space R factor (Section 6.4 of the wwPDB validation report) for the ligands in most FXR structures. As for ligand geometry, during re-refinement, the new ligand chemical restrains were generated and optimized “with the AM1 optimizer [2] and Mogul [3] reference values from the CSD database” as indicated above. After re-refinement, the ligand model was assessed with Mogul and outliers reviewed, then either verified or corrected. Use of Mogul reference data improved the chemical consistency of the ligand model, as verified by improved ligand geometry (Section 5.6 in the validation report online), with fewer bond length and bond angle outliers for ligands in most of the re-refined structures *versus* the original structures.

References

1. Gathiaka S, Liu S, Chiu M, et al (2016) D3R grand challenge 2015: Evaluation of protein–ligand pose and affinity predictions. J Comput Aided Mol Des 30:651–668 . doi: 10.1007/s10822-016-9946-8

2. Moriarty NW, Grosse-Kunstleve RW, Adams PD (2009) electronic Ligand Builder and Optimization Workbench (eLBOW): a tool for ligand coordinate and restraint generation. Acta Crystallogr D Biol Crystallogr 65:1074–1080 . doi: 10.1107/S0907444909029436

3. Bruno IJ, Cole JC, Kessler M, et al (2004) Retrieval of Crystallographically-Derived Molecular Geometry Information. J Chem Inf Comput Sci 44:2133–2144 . doi: 10.1021/ci049780b

4. Adams PD, Afonine PV, Bunkóczi G, et al (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr 66:213–221 . doi: 10.1107/S0907444909052925

5. Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of Macromolecular Structures by the Maximum-Likelihood Method. Acta Crystallogr D Biol Crystallogr 53:240–255 . doi: 10.1107/S0907444996012255

6. Bricogne G, Blanc E, Brandl M, et al (2016) BUSTER version 2.10.3. Cambridge, United Kingdom, Global Phasing Ltd.:

7. Winn MD, Ballard CC, Cowtan KD, et al (2011) Overview of the CCP4 suite and current developments. Acta Crystallogr D Biol Crystallogr 67:235–242 . doi: 10.1107/S0907444910045749

8. Brünger AT (1992) Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. Nature 355:472–475

9. Painter J, Merritt EA (2005) A molecular viewer for the analysis of TLS rigid-body motion in macromolecules. Acta Crystallogr D Biol Crystallogr 61:465–471 . doi: 10.1107/S0907444905001897

10. Young JY, Westbrook JD, Feng Z, et al (2017) OneDep: Unified wwPDB System for Deposition, Biocuration, and Validation of Macromolecular Structures in the PDB Archive. Structure 25:536–545 . doi: 10.1016/j.str.2017.01.004

11. Chen VB, Arendall WB, Headd JJ, et al (2010) MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr D Biol Crystallogr 66:12–21 . doi: 10.1107/S0907444909042073

12. Ramachandran GN, Ramakrishnan C, Sasisekharan V (1963) Stereochemistry of polypeptide chain configurations. J Mol Biol 7:95–99

13. Kleywegt GJ, Harris MR, Zou J, et al (2004) The Uppsala Electron-Density Server. Acta Crystallogr D Biol Crystallogr 60:2240–2249 . doi: 10.1107/S0907444904013253

14. Jones TA, Zou J-Y, Cowan SW, Kjeldgaard M (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. Acta Crystallogr A 47:110–119 . doi: 10.1107/S0108767390010224

15. Brändén C-I, Jones TA (1990) Between objectivity and subjectivity. Nature 343:687–689 . doi: 10.1038/343687a0

16. Tickle IJ (2012) Statistical quality indicators for electron-density maps. Acta Crystallogr D Biol Crystallogr 68:454–467 . doi: 10.1107/S0907444911035918

17. Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. Acta Crystallogr D Biol Crystallogr 66:486–501 . doi: 10.1107/S0907444910007493

18. Groom CR, Bruno IJ, Lightfoot MP, Ward SC (2016) The Cambridge Structural Database. Acta Crystallogr Sect B Struct Sci Cryst Eng Mater 72:171–179 . doi: 10.1107/S2052520616003954

19. Shao C, Yang H, Westbrook JD, et al (2017) Multivariate Analyses of Quality Metrics for Crystal Structures in the PDB Archive. Structure 25:458–468 . doi: 10.1016/j.str.2017.01.013