Electronic Supplementary Material (ESI) for Medicinal Chemistry Communications

Supporting information:

Fragments to link. A multiple docking strategy for second site binders

MedChemComm

*Márton Vass and György M. Keserű**

Discovery Chemistry, Gedeon Richter Plc. Gyömrői út 19-21, H-1103 Budapest, Hungary

gy.keseru@richter.hu

S1. SMILES representation of the docked ligands for the structures used in the study:

PDB ID	1st ligand SMILES	2nd ligand SMILES
1nny	OC(=O)C(=O)N(c1ccccc1C(O)=O)c1cccc2ccccc12	COc1cc2ccccc2cc1C(O)=O
1ph0	OC(=O)C(=O)N(c1ccccc1)c1ccccc1C(O)=O	COC(=O)c1ccccc1O
1pyn	OC(=O)C(Oc1ccccc1)C(O)=O	COC(=O)c1ccccc1O
1q1m	OC(=O)c1cc(on1)c1ccccc1F	COC(=O)c1ccccc1O
2jjk	CCNC(=O)NS(=O)(=O)c1cccc(Cl)c1	CCNC(=O)NS(=O)(=O)c1cccc(Cl)c1
2jt6	CC(=O)NO	Oc1ccc(cc1)c1ccc(cc1)C#N
2ohp	Nc1ccccn1	c1cc2ccccc2[nH]1
2ohq	Nc1ccccn1	COc1cccc(c1)c1ccccc1
2vt5	CCN\C(O)=N\S(=O)(=O)c1ccc(N)cc1	CCN\C(O)=N\S(=O)(=O)c1ccc(N)cc1
2vti	c1n[nH]c2ccccc12	NS(=O)(=O)c1ccccc1
3a7x	NC(=N)c1ccccc1	OCC(O)=O
3daj	CNc1ncnc2[nH]cnc12	CCCCNC(=O)c1cccc(NC(C)=O)c1
3fci	O=C1NC=CC(=O)N1	oC(=O)c1ccccc1
3ftx	CC(C)C[C@H](NC(=O)[C@@H](O)[C@H](N)Cc1ccccc1)C(O)=O	Oc1ccc(CCc2cc(O)cc(O)c2)cc1
3fue	CC(C)C[C@H](NC(=O)[C@@H](O)[C@H](N)Cc1ccccc1)C(O)=O	Clc1ccc2[nH]ccc2c1
3fuf	CC(C)C[C@H](NC(=O)[C@@H](O)[C@H](N)Cc1ccccc1)C(O)=O	Fc1ccc2[nH]ccc2c1
3fuh	CC(C)C[C@H](NC(=O)[C@@H](O)[C@H](N)Cc1ccccc1)C(O)=O	Oc1ccc2[nH]ccc2c1
3img	COc1ccc2[nH]ccc2c1	OC(=O)c1cc2ccccc2o1
3le8	COc1ccc2n(CC(O)=O)ccc2c1	OC(=O)c1cc2ccccc2o1
3lka	CC(=O)NO	COc1ccc(cc1)S(N)(=O)=O
3qc4	O=C1Nc2ccccc2N1	CNC(=O)C1=CC=CN(Cc2ccc(F)c(F)c2)C1=O
3ry2	C[C@H]1CNC(=O)N1	CCCCC(O)=O
4a7i	NC(=O)c1ccc(Cl)s1	CC(C)N1CCC(CC1)NS(C)(=O)=O
4ajh	Cc1nc2ccc(NC(=O)CCNC(N)=O)cc2s1	OC(=O)C(Oc1ccc(Br)cc1)C(O)=O
4ajj	Cc1nc2ccc(NC(=O)CCNC(N)=O)cc2s1	COc1ccc(CC(C(O)=O)C(O)=O)cc1OC
4dm3	Oc1cccc(O)c1	c1c[nH]cn1
2c8w	COc1ccc(cc1)S(=O)(=O)NC[C@@H](O)[C@@H](N)Cc1ccccc1	Clc1ccc(cc1)n1cnnn1
2qfo	Cc1cc(nc(N)n1)C(F)(F)F	O=C1OCC\C1=C/Nc1ccccc1
3hz1	CN(C)c1ncnc2nc[nH]c12	COC(=O)c1c(C)[nH]nc1c1ccco1
2xdu	Nc1ncccn1	Cc1nnc(Cl)c2ccccc12
2yei	COC(=O)Cc1cc(O)cc(O)c1	COC(=O)c1c(C)[nH]nc1c1ccno1
2yej	CSc1nc(C)nc(N)n1	COC(=O)c1c(C)[nH]nc1c1ccno1

S2. Details of the PDB structures used in the study and numeric docking results:

PDB ID	resolution (Å)	N lig.	Target	Method	1st RMSD	1st rank	2nd RMSD	2nd rank
1nny	2.40	1	protein tyrosine phosphatase 1B	linking	0,66	1	0,49	1
1ph0	2.20	1	protein tyrosine phosphatase 1B	linking	1,97	1	0,25	1
1pyn	2.20	1	protein tyrosine phosphatase 1B	linking	1,15	1	0,26	1
1q1m	2.60	1	protein tyrosine phosphatase 1B	linking	0,24	1	0,23	1
2jjk	2.00	1	fructose-1,6-bisphosphatase 1	tethering	0,29	1	0,26	1
2jt6	NMR	1	matrix metalloproteinase 3	linking	1,26	1	0,98	6
2ohp	2.25	1	beta-secretase 1	growing through linker	0,48	1	1,10	2
2ohq	2.10	1	beta-secretase 1	growing through linker	0,25	1	1,22	5
2vt5	2.20	1	fructose-1,6-bisphosphatase 1	tethering	0,83	1	0,23	1
2vti	2.00	1	cyclin dependent kinase 2	growing through linker	0,08	1	0,58	1
3a7x	1.75	1	cationic trypsin	tethering	0,30	1	0,32	12
3daj	2.00	1	aurora A kinase	tethering	0,18	1	1,88	1
3fci	1.27	1	uracil DNA glycosylase	tethering	0,39	1	1,51	1
3ftx	1.96	2	leukotriene A4 hydrolase	ternary complex	0,23	1	0,24	1
3fue	2.38	2	leukotriene A4 hydrolase	ternary complex	1,05	1	0,37	1
3fuf	2.60	2	leukotriene A4 hydrolase	ternary complex	0,41	1	0,63	4
3fuh	1.80	2	leukotriene A4 hydrolase	ternary complex	0,75	1	0,40	3
3img	1.80	2	MTB pantothenate synthetase	linking	0,21	1	0,51	20
3le8	1.70	1	MTB pantothenate synthetase	linking	0,32	1	1,98	1
3lka	1.80	2	matrix metalloproteinase 12	linking	0,32	8	0,43	1
3qc4	1.80	1	3-phosphoinositide dependent kinase 1	tethering	0,20	1	1,05	1
3ry2	0.95	1	streptavidin	deconstruction	0,27	1	0,69	1
4a7i	2.40	1	factor Xa	deconstruction	1,20	1	0,33	1
4ajh	1.93	2	lactate dehydrogenase A	linking	0,32	1	0,50	1
4ajj	1.75	2	lactate dehydrogenase A	linking	0,50	1	0,20	1
4dm3	2.40	2	phenylethanolamine *N*-methyltransferase	ternary complex	0,37	5	0,61	1
Structures with waters included in docking:								
2c8w	1.96	1	thrombin	linking	0,17	1	0,46	1
2qfo	1.68	2	heat shock protein 90-α	linking	0,34	1	0,20	1
3hz1	2.30	2	heat shock protein 90-α	linking	0,51	1	0,41	1
2xdu	1.74	2	heat shock protein 90-α	ternary complex	0,26	1	0,15	1
2yei	2.20	2	heat shock protein 90-α	ternary complex	0,37	1	0,94	1
2yej	2.20	2	heat shock protein 90-α	ternary complex	1,04	1	0,19	1

S3. Numeric cross-docking results (cells contain 1st RMSD, 1st rank, 2nd RMSD and 2nd rank from left to right and top to bottom):

Prot \ Lig	3img	3le8
3img	0.21 1 0.51 20	1.00 1 1.39 14
3le8	0.31 1 1.95 3	0.32 1 1.98 1

Prot \ Lig	4jh	4ajj
4jh	0.32 1 0.50 1	0.37 1 0.41 1
4ajj	0.73 1 0.59 1	0.50 1 0.20 1

Prot \ Lig	2jjk	2vt5
2jjk	0.29 1 0.26 1	0.36 1 0.23 1
2vt5	0.90 1 1.78 1	0.83 1 0.23 1

Prot \ Lig	2ohp	2ohq
2ohp	0.48 1 1.10 2	0.37 1 0.68 1
2ohq	0.49 1 1.48 1	0.25 1 1.22 5

Prot \ Lig	1nny	1ph0	1pyn	1q1m
1nny	0.66 1 0.49 1	1.99 1 1.27 17	1.49 1 - -	1.21 1 - -
1ph0	0.77 1 1.57 1	1.97 1 0.25 1	1.32 1 0.26 2	1.48 1 0.36 1
1pyn	0.73 1 1.30 1	0.35 1 1.98 1	1.15 1 0.26 1	1.49 2 0.35 1
1q1m	- - - -	0.39 9 - -	0.54 1 0.30 1	0.24 1 0.23 1

Prot \ Lig	3ftx	3fue	3fuf	3fuh
3ftx	0.23 1 0.24 1	0.36 1 - -	0.37 1 1.51 13	0.81 1 0.19 5
3fue	0.23 1 1.05 1	1.05 1 0.37 1	0.24 1 1.80 1	0.33 1 0.49 16
3fuf	0.39 1 1.66 1	0.97 3 0.36 1	0.41 1 0.63 4	0.49 1 0.53 7
3fuh	0.26 1 0.62 1	0.55 1 0.96 6	0.40 1 1.45 12	0.75 1 0.40 3

Prot \ Lig	2qfo	3hz1	2xdu	2yei	2yej
2qfo	0.34 1 0.20 1	0.93 12 - -	0.55 1 - -	1.37 3 - -	0.76 3 - -
3hz1	- - - -	0.51 1 0.41 1	0.79 1 0.91 1	1.20 1 - -	1.23 1 - -
2xdu	0.72 1 0.79 7	0.54 1 0.64 1	0.26 1 0.15 1	0.56 1 - -	0.81 1 0.96 9
2yei	0.50 1 1.86 5	0.47 1 0.82 1	0.18 1 0.47 3	0.37 1 0.94 1	0.44 1 0.25 1
2yej	0.75 1 0.54 13	0.44 1 0.60 3	0.50 1 0.38 4	0.38 1 1.93 20	1.04 1 0.19 1