Supplementary Material

Investigating the Molecular basis of *N*-substituted 1-hydroxy-4-sulfamoyl-2-naphthoates compounds binding to Mcl1

Kalaimathy Singaravelu[¶*], Pavithra K. Balasubramanian[#*] and Parthiban Marimuthu[†§]

[¶]Turku Computer Science and Informatics, Department of Future Technology, University of Turku, FI-20520 Turku, Finland.

[#]Department of Bioscience and Biotechnology, Institute of KU Biotechnology, Konkuk University, Seoul, 05029, South Korea.

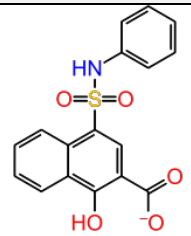
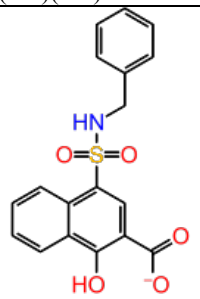
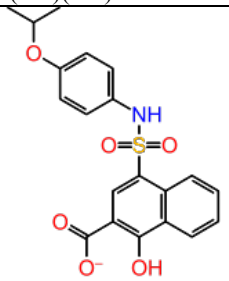
[†]Structural Bioinformatics Laboratory, Biochemistry, Faculty of Science and Engineering, Åbo Akademi University, Tykistökatu 6A, FI-20520 Turku, Finland.

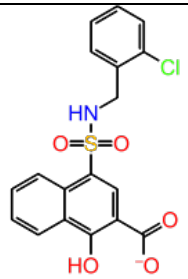
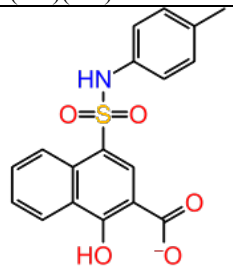
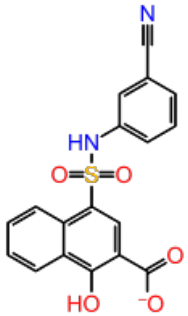
***** – Contributed equally to this work

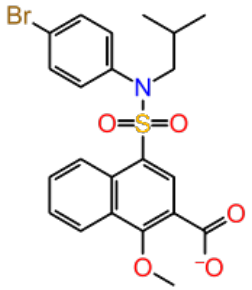
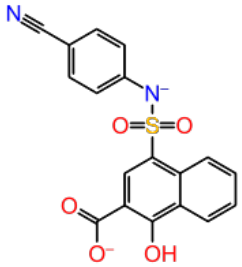
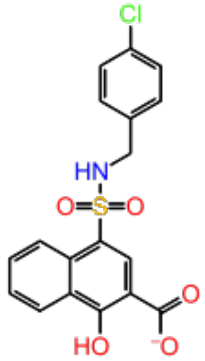
[§]Address correspondence to:

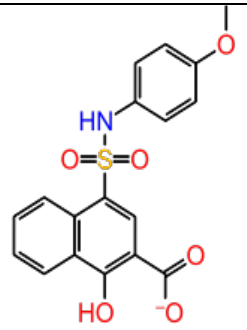
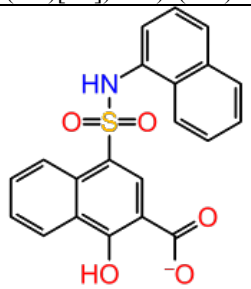
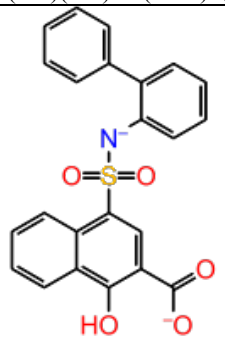
Dr. Parthiban Marimuthu (Ph.D.), Structural Bioinformatics Laboratory (SBL), Biochemistry, Faculty of Science and Engineering, Åbo Akademi University, Tykistökatu 6 A, FI-20520 Turku, Finland. Phone: +358 2 215 4600. E-mail address: parthiban.marimuthu@abo.fi

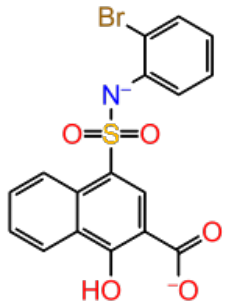
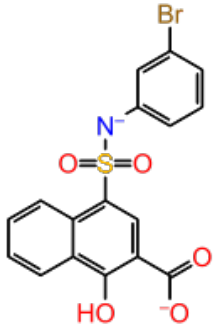
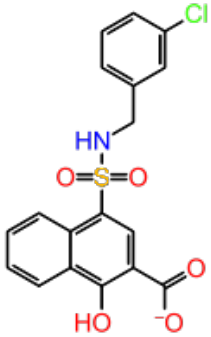
Table S1: The chemical datasets used for 3D-QSAR model constructions along with its activity values (ChEMBL assay id: 3779852).

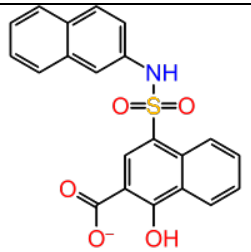
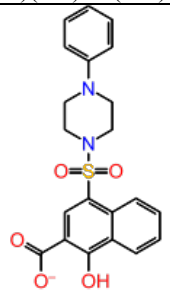
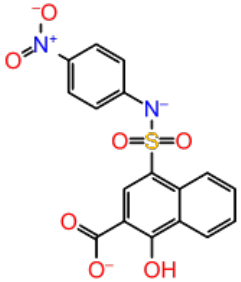
CHEMBL	2D representation of dataset and SMILES	K_i (µM)	Activity	Predicted	Fitness	QSAR Set	Pharm Set
3780724	 c1cccc(c12)c(O)c(C([O-])=O)cc2S(=O)(=O)Nc3ccccc3	106	3.975	4.61	2.13	training	inactive
3780355	 c1cccc(c12)c(O)c(C([O-])=O)cc2S(=O)(=O)NCc3ccccc3	56.3	4.249	4.53	1.79	test	inactive
3780580	 CC(C)Oc1ccc(cc1)NS(=O)(=O)c(cc(C(=O)[O-])c2O)c(c23)cccc3	54	4.268	4.40	2.25	training	inactive

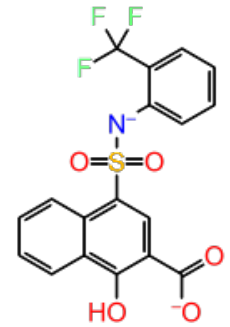
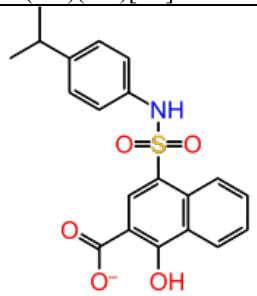
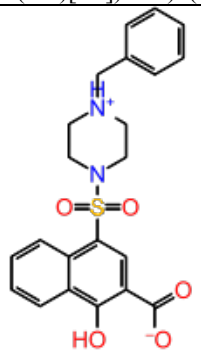
3781306	 c1cccc(c12)c(O)c(C([O-])=O)cc2S(=O)(=O)NCc3ccccc3Cl	50.3	4.298	4.40	1.79	training	inactive
3780094	 Cc1ccc(cc1)NS(=O)(=O)c(cc(C(=O)[O-])c2O)c(c23)cccc3	34.3	4.465	4.82	2.16	test	inactive
3780382	 c1cccc(c12)c(O)c(C([O-])=O)cc2S(=O)(=O)Nc3cccc(c3)C#N	26.9	4.570	4.60	2.12	training	inactive

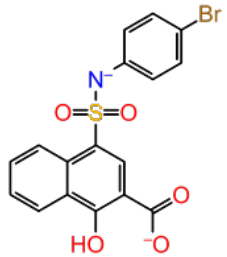
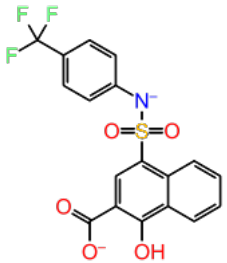
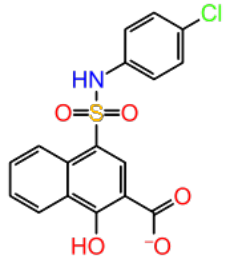
3779915	 c1cccc(c12)c(OC)c(C(=O)[O-])cc2S(=O)(=O)N(CC(C)C)c3ccc(Br)cc3	23.4	4.631	4.70	2.09	training	inactive
3780766	 c1cccc(c12)c(O)c(C([O-])=O)cc2S(=O)(=O)[N-]c3ccc(C#N)cc3	22.2	4.654	5.18	2.37	training	inactive
3780331	 c1cccc(c12)c(O)c(C([O-])=O)cc2S(=O)(=O)NCc3ccc(Cl)cc3	21	4.678	4.52	1.79	test	inactive

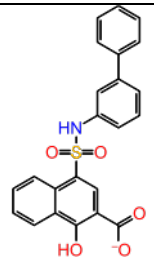
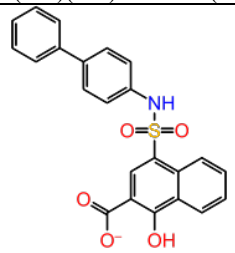
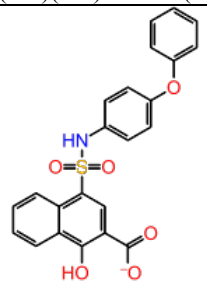
3780070	 COc1ccc(cc1)NS(=O)(=O)c(cc(C(=O)[O-])c2O)c(c23)cccc3	11.9	4.924	4.54	2.20	training	inactive
3780357	 c1cccc(c12)c(O)c(C([O-])=O)cc2S(=O)(=O)Nc(ccc3)c(c34)cccc4	11.2	4.951	5.08	2.07	training	inactive
3780451	 c1cccc(c12)c(O)c(C([O-])=O)cc2S(=O)(=O)[N-]c3ccccc3-c4ccccc4	8.5	5.071	5.48	2.21	test	

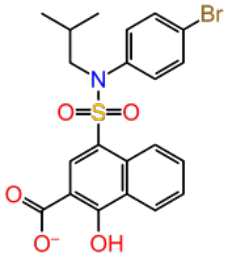
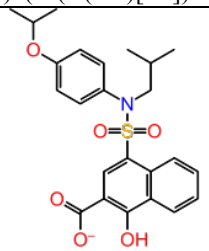
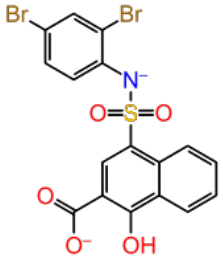
3781478	 c1cccc(c12)c(O)c(C([O-])=O)cc2S(=O)(=O)[N-]c3ccccc3Br	6.84	5.165	5.23	2.31	test	
3780899	 c1cccc(c12)c(O)c(C([O-])=O)cc2S(=O)(=O)[N-]c3cccc(Br)c3	5.64	5.249	5.26	2.32	training	
3781868	 c1cccc(c12)c(O)c(C([O-])=O)cc2S(=O)(=O)NCc3cccc(Cl)c3	5.03	5.298	4.94	1.81	training	

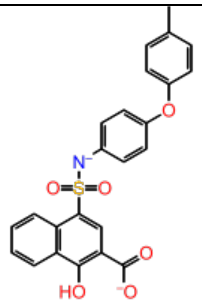
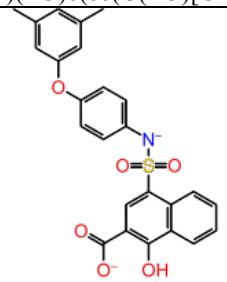
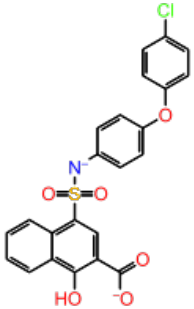
3781502	 c1cccc(c12)c(O)c(C([O-])=O)cc2S(=O)(=O)Nc(cc3)cc(c34)cccc4	4.58	5.339	5.36	2.20	training	
3780385	 c1cccc(c12)c(O)c(C([O-])=O)cc2S(=O)(=O)N3CCN(CC3)c4ccccc4	4.54	5.343	5.28	1.35	training	
3781264	 c1cccc(c12)c(O)c(C([O-])=O)cc2S(=O)(=O)[N-]c3ccc([N+](=O)[O-])cc3	4.49	5.348	4.91	2.38	test	

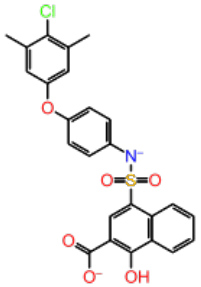
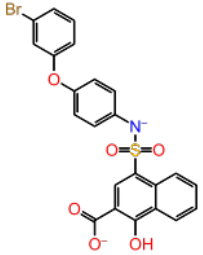
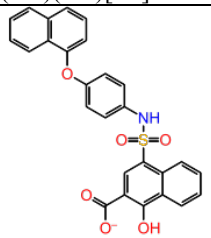
3781823	 c1cccc(c12)c(O)c(C([O-])=O)cc2S(=O)(=O)[N-]c3ccccc3C(F)(F)F	4.05	5.393	5.42	2.25	training	
3781692	 CC(C)c1ccc(cc1)NS(=O)(=O)c(cc(C(=O)[O-])c2O)c(c23)cccc3	3.86	5.413	5.15	2.20	training	
3781593	 c1cccc(c12)c(O)c(C([O-])=O)cc2S(=O)(=O)N3CC[NH+](CC3)Cc4ccccc4	3.41	5.467	5.54	1.34	training	

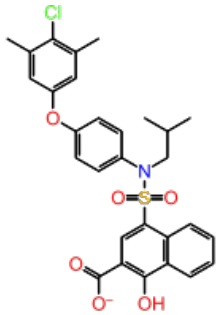
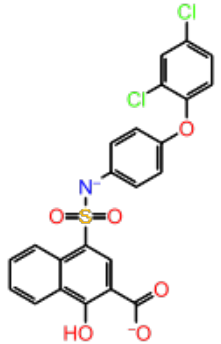
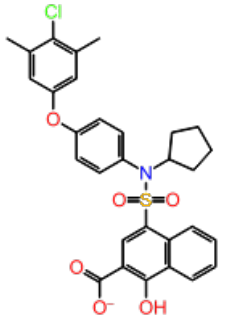
3781059	 c1cccc(c12)c(O)c(C([O-])=O)cc2S(=O)(=O)[N-]c3ccc(Br)cc3	2.76	5.559	5.55	2.35	test	
3781065	 c1cccc(c12)c(O)c(C([O-])=O)cc2S(=O)(=O)[N-]c3ccc(C(F)(F)F)cc3	2.5	5.602	5.70	2.37	training	
3780290	 c1cccc(c12)c(O)c(C([O-])=O)cc2S(=O)(=O)Nc3ccc(Cl)cc3	1.91	5.719	5.03	2.16	training	

3780990	 c1cccc(c12)c(O)c(C([O-])=O)cc2S(=O)(=O)Nc3cccc(c3)-c4ccccc4	1.88	5.726	5.91	1.91	test	
3780617	 c1cccc(c12)c(O)c(C([O-])=O)cc2S(=O)(=O)Nc3ccc(cc3)-c4ccccc4	1.54	5.812	6.19	2.14	training	
3781946	 c1cccc(c12)c(O)c(C([O-])=O)cc2S(=O)(=O)Nc3ccc(cc3)Oc4ccccc4	1.15	5.939	6.03	2.54	training	

3781019	 c1cc(Br)ccc1N(CC(C)C)S(=O)(=O)c(cc(C(=O)[O-])c2O)c(c23)cccc3	0.566	6.247	6.27	2.10	training	active
3780576	 c1cccc(c12)c(cc(C(=O)[O-])c2O)S(=O)(=O)N(CC(C)C)c3ccc(cc3)OC(C)C	0.487	6.312	5.93	2.17	training	active
3780113	 c1cccc(c12)c(O)c(C([O-])=O)cc2S(=O)(=O)[N-]c3ccc(Br)cc3Br	0.42	6.377	5.74	2.34	test	active

3781396	 Cc1ccc(cc1)Oc2ccc(cc2)[N-]S(=O)(=O)c(cc(C(=O)[O-])c3O)c(c34)cccc4	0.335	6.475	6.44	3.00	training	active
3781894	 Cc1cc(C)cc(c1)Oc2ccc(cc2)[N-]S(=O)(=O)c(cc(C(=O)[O-])c3O)c(c34)cccc4	0.284	6.547	6.83	2.96	training	active
3780719	 c1cccc(c12)c(O)c(C([O-])=O)cc2S(=O)(=O)[N-]c3ccc(cc3)Oc4ccc(Cl)cc4	0.173	6.762	6.50	3.00	test	active

3780473	 Clc1c(C)cc(cc1C)Oc2ccc(cc2)[N-]S(=O)(=O)c(cc(C(=O)[O-])c3O)c(c34)cccc4	0.117	6.932	6.89	2.97	training	active
3780311	 c1cccc(c12)c(O)c(C([O-])=O)cc2S(=O)(=O)[N-]c3ccc(cc3)Oc4cccc(Br)c4	0.114	6.943	6.67	2.97	training	active
3780225	 c1cccc(c12)c(O)c(C([O-])=O)cc2S(=O)(=O)Nc3ccc(cc3)Oc(ccc4)c(c45)cccc5	0.082	7.086	6.73	2.45	training	active

3780803	 c1cccc(c12)c(cc(C(=O)[O-])c2O)S(=O)(=O)N(CC(C)C)c3ccc(cc3)Oc4cc(C)c(Cl)c(C)c4	0.08	7.097	7.31	2.45	training	active
3781669	 c1cccc(c12)c(O)c(C([O-])=O)cc2S(=O)(=O)[N-]c3ccc(cc3)Oc4ccc(Cl)cc4Cl	0.079	7.102	6.62	2.99	test	active
3780397	 Clc1c(C)cc(cc1C)Oc2ccc(cc2)N(C3CCCC3)S(=O)(=O)c(cc(C(=O)[O-])c4O)c(c45)cccc5	0.033	7.481	7.43	2.42	training	active

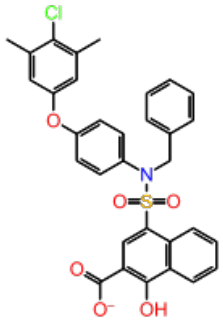
3781822	 Clc1c(C)cc(cc1C)Oc2ccc(cc2)N(Cc3ccccc3)S(=O)(=O)c(cc(C(=O)[O-])c4O)c(c45)cccc5	0.031	7.509	7.56	2.37	training	active

Table S2: Count and percentage of Decoys

%Decoys	1%	2%	5%	10%	20%
#Actives	28	28	28	28	36
%Actives	68.3	68.3	68.3	68.3	87.8

Table S3: Count and percentage of actives

%Results	1%	2%	5%	10%	20%
#Actives	5	11	24	28	28
%Actives	12.2	26.8	58.5	68.3	68.3

Table S4: Enrichment factors with respect to N% sample size

%Sample	1%	2%	5%	10%	20%
EF	13	13	12	6.8	3.4
EF*	70	35	14	6.8	4.4
EF'	82	70	35	16	7.5
DEF	8.8	9.2	12	6.8	3.4
DEF*	69	35	14	6.8	4.4
DEF'	82	70	35	16	7.5
Eff	0.971	0.943	0.864	0.745	0.629

Supplementary Figure

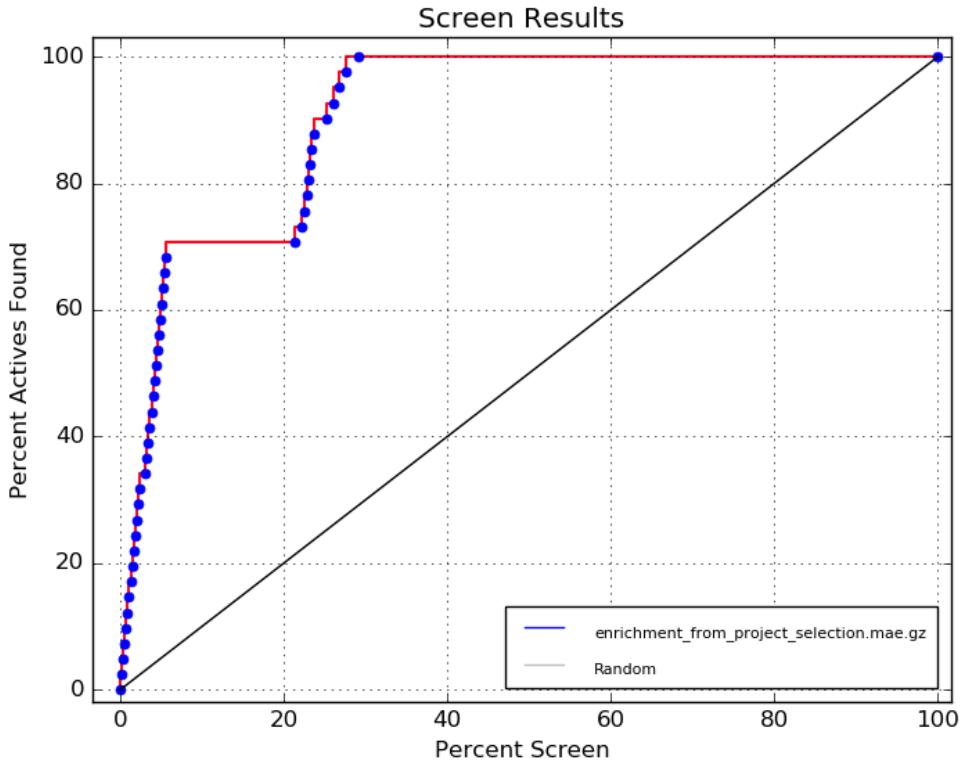


Figure S1: ROC curve obtained for the model NNRRR against randomly curve**.**