Supporting Information Section

Rationalizing Tight Ligand Binding Through Cooperative Interaction Networks

*Bernd Kuhn**[†], Julian E. Fuchs[†,§], Michael Reutlinger[†,#], Martin Stahl[†], and Neil R. Taylor**[‡]*

[†] Discovery Chemistry, F. Hoffmann-La Roche AG, CH-4070 Basel, Switzerland

[‡] Desert Scientific Software Pty Ltd, Level 5 Nexus Building, Norwest Business Park, 4 Columbia Court, Baulkham Hills, NSW, 2153, Australia

[§] Present address: Institute of General, Inorganic and Theoretical Chemistry, University of Innsbruck, A-6020 Innsbruck, Austria

[#] Present address: Institute of Pharmaceutical Sciences, Swiss Federal Institute of Technology Zurich, CH-8093 Zurich, Switzerland

to whom correspondence should be addressed: for queries regarding interactions and medicinal chemistry case studies: Dr. Bernd Kuhn, F. Hoffmann-La Roche AG, PRCB 92/2.56B, CH-4070 Basel, Switzerland, Phone +41 616889773, email bernd.kuhn@roche.com; for queries regarding the network concepts and software details: Dr. Neil R. Taylor, Desert Scientific Software Pty Ltd, Level 5 Nexus Building, Norwest Business Park, 4 Columbia Court, Baulkham Hills, NSW, 2153, Australia, Phone +612 8860 6466, email neil.taylor@desertsci.com

Table of Contents

Ligand structures and logarithmic potency values for the Neuraminidase training set S1

Public X-ray structures used in the training set for pose prediction S2

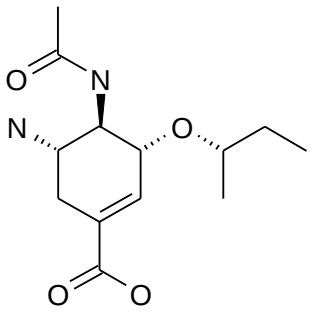
Optimized scoring function parameters for pose prediction ($S_{Scorpion,pose}$) S3

External test of $S_{Scorpion,pose}$ S4

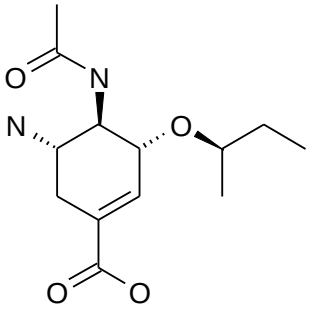
External test of $S_{Scorpion}$ S5

Virtual screening results on targets of the DUD data set with $S_{Scorpion}$ S6

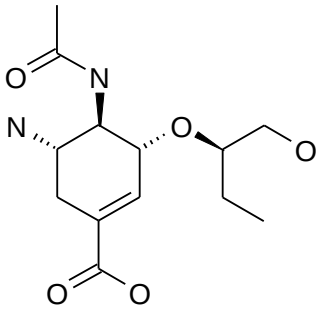
Figure S1: Ligand structures and logarithmic potency values for the Neuraminidase training set



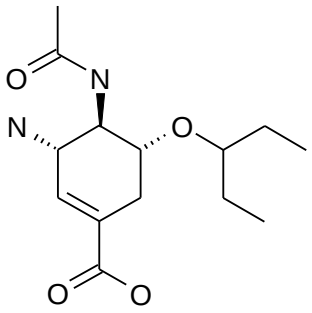
A: 8.0
B: 8.8



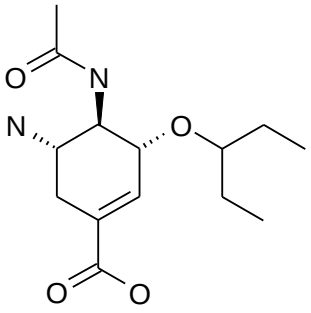
A: 8.0
B: 8.2



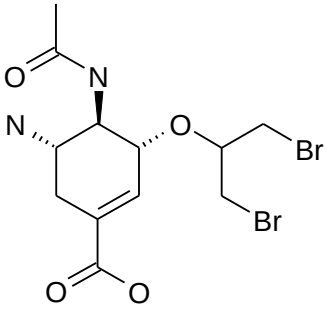
A: 7.6



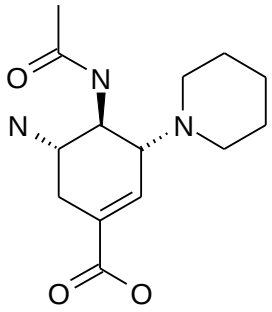
A: 7.5



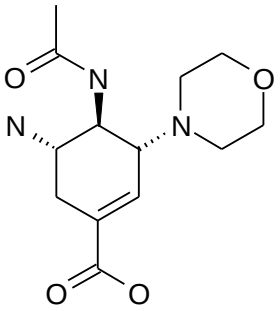
A: 8.9
B: 8.5



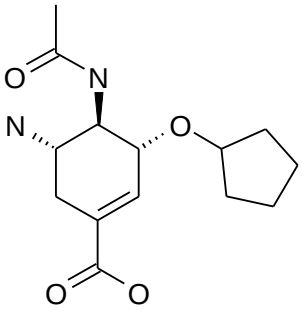
A: 9.2



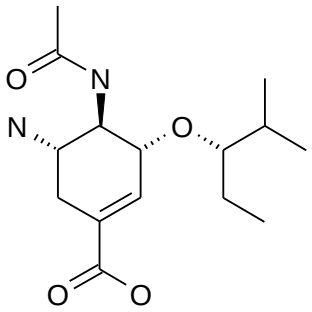
A: 7.5



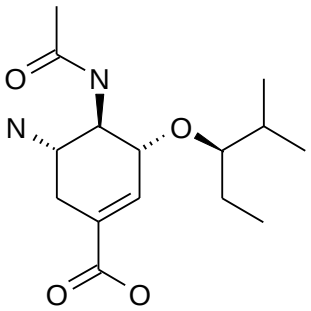
A: 6.7



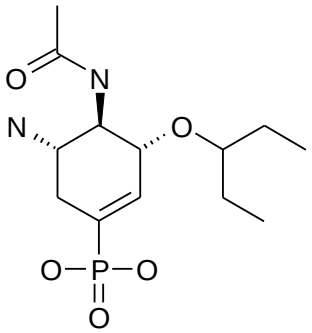
A: 7.7



A: 8.7
B: 8.1

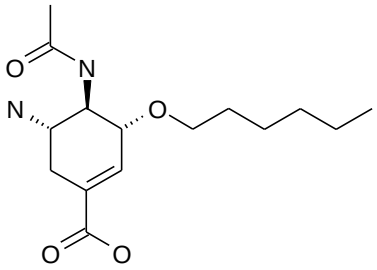


A: 9.5

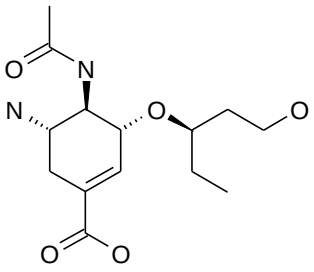


A: 8.7
B: 7.1

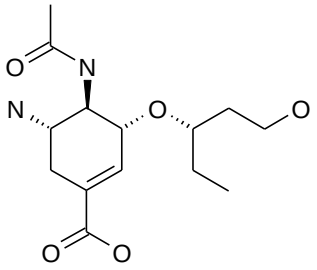
Figure S1 (continued)



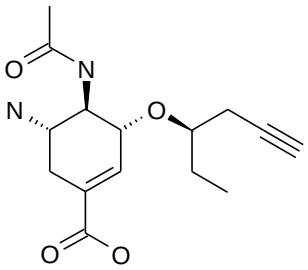
A: 6.8



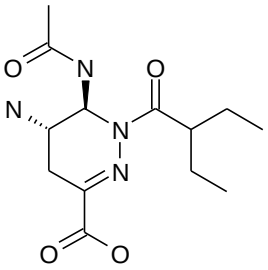
A: 7.5



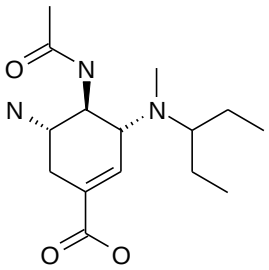
A: 7.4



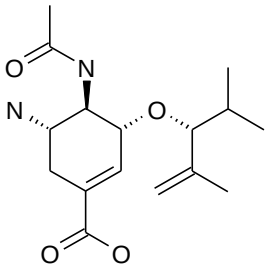
A: 8.2
B: 9.1



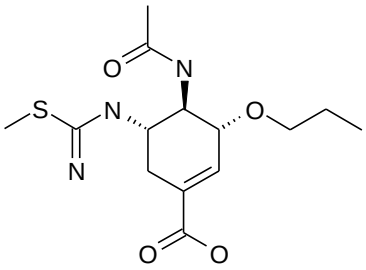
A: 8.3



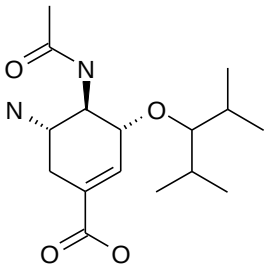
A: 8.2



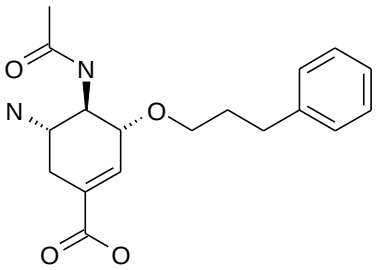
A: 9.4



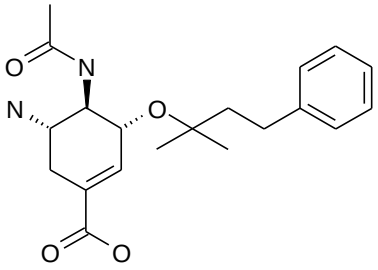
A: 7.5



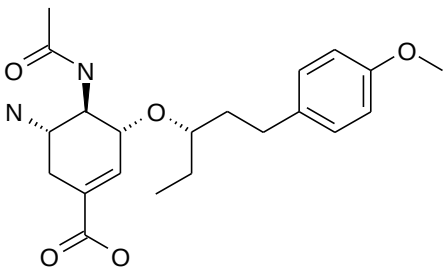
A: 10.0
B: 8.6



A: 7.0



A: 9.2



A: 9.3

Table S2: Public X-ray structures used in the training set for pose prediction.

PDB code	Protein
1b9v	neuraminidase
1f8b	neuraminidase
1h1s	cdk2
1lpz	factor Xa
1o5r	adenosine deaminase
1of1	thymidine kinase
1oi9	cdk2
1tow	fatty acid binding protein 4
1uu9	pdk1
1x70	dpp4
1y2c	pde4
1ydr	pkaca
1yds	pkaca
1z5m	pdk1
1z95	androgen receptor
2axa	androgen receptor
2c68	cdk2
2c93	thrombin
2ctc	carboxypeptidase a
2gnj	pkaca
2hha	dpp4
2hny	hiv reverse transcriptase
2ihq	androgen receptor
2iiv	dpp4
2j34	factor Xa
2jdo	akt2
2oag	dpp4
2oph	dpp4
2pog	estrogen receptor α

Table S3: Optimized scoring function parameters for pose prediction ($S_{Scorpion,pose}$), see also Equation (3) of the manuscript.

Interaction type (i)	Pairwise interaction coefficient (p_i)	Network coefficient (n_i)	Network threshold ($n_{thres,i}$)
hydrogen bond	0.90	–	–
vdW	0.097	0.20	4
π-π	0.34	0.20	4
ionic	0.40	–	–
cation-π	0.41	–	–
unf_hydrogen bond	–1.38	–	–
unf_dipole	–0.81	–	–
clash_apolar	–1.25	–	–
clash_polar	–1.25	–	–
desolv_donor	–1.14	–	–
desolv_acceptor	–0.97	–	–

Table S4: External test of $S_{Scorpion,pose}$ in reproducing the experimentally observed binding mode within a root mean square distance (RMSD) of 2.0Å of the top-ranked pose. Binding poses and results from other scoring functions were taken from the data set of Wang et al.,[1] which contains 100 computer-generated binding poses for each of 100 protein structures. For each of the ligand poses we calculated ligand strain energies as detailed in the Methods section. Poses with a strain energy > 2 kcal/mol were eliminated.

Scoring function	Success rate (%)
PLP (Cerius2)	76
FScore (Sybyl)	74
LigScore (Cerius2)	74
DrugScore	72
$S_{Scorpion,pose}$	**70**
Ludi (Cerius2)	67
XScore	66
AutoDock	62
PMF (Cerius2)	52
GScore (Sybyl)	42
ChemScore (Sybyl)	35
DScore (Sybyl)	26

Table S5: External test of $S_{Scorpion}$ in ranking binding affinities. The four subsets are taken from Englebienne and Moitessier (see Table S3 in this publication).[2] Kendall's τ is used as a rank correlation measure. There is zero overlap in complexes between our training and the external test sets.

Scoring function	HIV protease	thrombin	trypsin	thrombin/ trypsin/ factorXa	**Average**
No. of complexes	11	22	13	42	
RankScore	0.55	0.68	0.36	0.61	**0.55**
XScore	0.73	0.31	0.44	0.41	**0.47**
$DrugScore^{CSD}$	0.55	0.42	0.46	0.44	**0.47**
PLP1 (Cerius2)	0.59	0.48	0.31	0.49	**0.47**
$S_{Scorpion}$	**0.59**	**0.49**	**0.23**	**0.51**	**0.46**
DockScore (Sybyl)	0.62	0.37	0.39	0.46	**0.46**
$DrugScore^{PDB}$	0.48	0.56	0.18	0.56	**0.44**
PLP2 (Cerius2)	0.59	0.39	0.31	0.45	**0.43**
GoldScore (Sybyl)	0.44	0.55	0.18	0.50	**0.42**
LigScore2 (Cerius2)	0.51	0.45	0.25	0.42	**0.41**
Hammerhead (Cerius2)	0.48	0.30	0.49	0.31	**0.40**
PMF (Cerius2)	0.55	0.25	0.36	0.35	**0.38**
ChemScore (Sybyl)	0.62	0.06	0.56	0.27	**0.38**
GlideScore	0.51	0.31	0.31	0.33	**0.37**
LigScore1 (Cerius2)	0.44	0.43	0.18	0.37	**0.35**
MW	0.26	0.49	0.13	0.46	0.33
eHiTS SF	0.37	0.46	0.08	0.39	**0.32**
Surflex SF	0.48	0.18	0.31	0.30	**0.32**
FlexXScore (Sybyl)	0.66	–0.02	0.39	0.08	**0.27**
PMF (Sybyl)	–0.18	0.29	–0.03	0.30	**0.10**

Table S6: Comparison of ROC virtual screening enrichments[3] at several early false positive rates for different targets of the DUD data set.[4] SP stands for Glide/SP docking and scoring. Results for $S_{Scorpion}$ were obtained from rescoring the top-ranked Glide/SP docking poses. ACHE: Acetylcholine esterase, FGFr1: Fibroblast growth factor receptor kinase, FXa: Factor Xa, HIVPR: HIV protease, NA: Neuraminidase, P38: P38 mitogen activated protein kinase, PDGFrb: Platelet derived growth factor receptor kinase, PR: Progesterone receptor.

	0.5%		1.0%		2.0%	
	SP	$S_{Scorpion}$	SP	$S_{Scorpion}$	SP	$S_{Scorpion}$
ACHE	0.0	**0.0**	0.0	**0.0**	0.0	**0.5**
FGFr1	1.6	**0.0**	1.6	**0.8**	2.5	**1.2**
FXa	42.1	**20.1**	30.5	**11.6**	19.2	**7.5**
HIVPR	12.0	**15.3**	7.8	**8.0**	6.4	**4.1**
NA	68.8	**3.8**	40.3	**6.0**	23.1	**8.0**
P38	0.4	**0.9**	0.4	**1.1**	0.6	**1.8**
PDGFrb	1.2	**3.9**	0.6	**2.0**	0.3	**1.3**
PR	0.0	**15.2**	0.0	**11.4**	1.8	**5.7**

References

1. Wang, R.; Lu, Y. P.; Wang, S. M., Comparative evaluation of 11 scoring functions for molecular docking. *J. Med. Chem.* **2003,** *46*, 2287-2303.
2. Englebienne, P.; Moitessier, N., Docking Ligands into Flexible and Solvated Macromolecules. 4. Are Popular Scoring Functions Accurate for this Class of Proteins? *J. Chem. Inf. Model.* **2009,** *49*, 1568-1580.
3. Nicholls, A., What do we know and when do we know it? *J. Comput.-Aided Mol. Des.* **2008,** *22*, 239-255.
4. Huang, N.; Shoichet, B. K.; Irwin, J. J., Benchmarking Sets for Molecular Docking. *J. Med. Chem.* **2006,** *49*, 6789-6801.