Supplementary Information for

Docking Ligands into Flexible and Solvated Macromolecules. 4. Are Popular Scoring Functions Accurate for this Class of Proteins?

Pablo Englebienne, Nicolas Moitessier*

Department of Chemistry, McGill University, 801 Sherbrooke St W, Montreal, QC, Canada H3A 2K6.

Contents

Table S1. Description of the 209 complexes included in Set 1.

PDB	Protein	Ebind (kcal mol^{-1})	MW
1a30	HIV-1 protease	-5.86	375.29
1a85	MMP-8	-6.16	393.45
1adl	Trp RNA-binding attenuation protein	-7.30	303.47
1af6	Maltoporin	-2.48	342.30
1afk	Ribonuclease a	-9.01	504.16
1aid	HIV-1 protease	-6.57	453.07
1ajv	HIV-1 protease	-10.51	574.70
1apw	Penicillopepsin	-10.89	506.64
1b8n	Purine nucleoside phosphorylase	-14.33	282.28
1b8o	Purine nucleoside phosphorylase	-14.48	267.27
1bdl	HIV-1 protease	-7.97	520.68
1biw	MMP-3	-9.51	415.53
1bma	Elastase	-6.25	521.61
1bmn	Thrombin	-11.49	531.66
1bn4	Carbonic anhydrase	-12.67	362.45
1bnn	Carbonic anhydrase	-13.61	374.46
1br5	Ricin	-3.68	253.22
1br6	Ricin	-4.38	311.28
1bxo	Penicillopepsin	-13.61	638.70
1c3x	Pentosyltransferase	-5.01	277.02
1c4u	Thrombin	-14.12	582.51
1c4v	Thrombin	-14.70	531.68
1c5c	chimeric decarboxylase antibody 21d8	-9.48	343.34
1c5q	Trypsin	-8.66	303.15
1cet	Lactate dehydrogenase	-3.93	320.89
1ciz	MMP-3	-10.13	502.62
1d3p	Thrombin	-10.06	545.75
1d3t	Thrombin	-8.90	543.73
1d4p	Thrombin	-8.58	361.47
1d8m	MMP-3	-11.58	419.46
1db1	Vitamin d nuclear receptor	-12.61	416.65
1dmp	HIV-1 protease	-13.00	536.68
1dvz	Transthyretin	-9.68	280.23
1dy4	Cellobiohydrolase I	-5.94	258.34
1e1x	Cyclin dependent kinase 2	-8.02	251.29
1e2k	Thymidine kinase	-6.73	252.27
1e4h	Transthyretin	-11.45	488.62
1e5a	Transthyretin	-10.40	330.82
1e66	Acetylcholinesterase	-13.46	298.82
1eb1	Thrombin	-14.20	439.58
1ecv	PTP 1b	-6.61	333.04
1efy	ADP-ribose polymerase	-11.19	267.29
1ejn	Urokinase-type plasminogen activator	-7.65	342.47
1elc	Elastase	-9.07	507.58
1erb	Retinoic acid receptor rxr-alpha	-9.60	327.51
1evh	Mena evh1 domain	-4.38	594.69
1ezq	Factor xa	-12.32	460.58

PDB	Protein	Ebind (kcal mol^{-1})	MW
1f0u	Trypsin	-9.75	460.58
1f4e	Thymidylate synthase	-4.03	268.31
1f4g	Thymidylate synthase	-8.82	496.48
1f73	n-Acetyl neuraminate lyase	-3.25	310.28
1f8d	Neuraminidase	-4.63	290.28
1f8e	Neuraminidase	-6.57	290.30
1fcx	Retinoic acid receptor gamma-1	-9.79	387.50
1fcy	Retinoic acid receptor gamma-1	-11.60	385.49
1fcz	Retinoic acid receptor gamma-1	-12.55	361.47
1fd0	Retinoic acid receptor gamma-1	-11.44	400.50
1fh7	Xylanase beta-1	-7.13	266.27
1fh8	Xylanase beta-1	-9.38	250.27
1fh9	Xylanase beta-1	-8.75	294.26
1fhd	Xylanase beta-1	-9.28	302.29
1fjs	Factor xa	-13.56	526.50
1fkh	FK506 binding protein	-11.10	455.64
1fmb	EIAV protease	-13.61	566.77
1g2l	Factor xa	-9.86	526.62
1g30	Thrombin	-9.33	526.62
1g4o	Carbonic anhydrase	-11.23	290.34
1ghv	Thrombin	-5.92	254.27
1ghw	Thrombin	-5.72	253.29
1ghz	Trypsin	-6.53	254.27
1gi1	Trypsin	-6.49	253.29
1gi6	Trypsin	-8.47	252.30
1gi8	Urokinase-type plasminogen activator	-6.87	253.29
1gj4	Thrombin	-5.75	362.84
1gj5	Thrombin	-7.16	329.38
1gj7	Urokinase-type plasminogen activator	-10.74	362.84
1gj9	Urokinase-type plasminogen activator	-10.18	382.46
1gja	Urokinase-type plasminogen activator	-7.38	256.29
1gpn	Acetylcholinesterase	-8.82	257.36
1h0a	Epsin	-7.41	417.07
1h1s	Cyclin dependent kinase 2	-11.19	402.48
1h22	Acetylcholinesterase	-12.38	468.69
1h23	Acetylcholinesterase	-11.36	496.74
1h46	Exoglucanase I	-4.86	257.34
1hpo	HIV-1 protease	-12.55	504.61
1hpv	HIV-1 protease	-12.55	505.64
1htf	HIV-1 protease	-9.30	575.76
1hwr	HIV-1 protease	-11.34	406.53
1hy7	MMP-3	-7.34	402.45
1i73	MMP-8	-7.48	449.47
1igj	Precursor of periplasmic sugar receptor	-13.61	520.67
1ikt	Estradiol 17 beta-dehydrogenase 4	-4.63	352.52
1j8v	Beta-d-glucan glucohydrolase isoenzyme exo1	-4.91	358.37
1jan	MMP-8	-6.43	301.37
1jao	MMP-8	-8.06	323.42

PDB	Protein	Ebind (kcal mol^{-1})	MW
1jaq	MMP-8	-6.10	302.33
1jj9	MMP-8	-7.85	332.38
1jwt	Thrombin	-10.69	493.61
1k22	Thrombin	-11.43	430.53
1k4g	tRNA-guanine transglycosylase	-7.96	288.33
1k9s	Purine nucleoside phosphorylase	-8.88	282.28
1kbc	MMP-8	-10.48	399.53
1kdk	Sex hormone-binding globulin	-12.32	290.45
1kel	28b4 fab	-9.91	345.27
1koj	Glucose-6-phosphate isomerase	-9.12	260.12
1ksn	Factor xa	-12.80	447.52
1kts	Thrombin	-10.93	500.59
1kzk	HIV-1 protease	-14.14	575.73
1l2s	Lactamase beta	-6.25	316.76
1l8b	Eukaryotic translation initiation factor 4e	-9.33	535.20
1lee	Plasmepsin II	-10.54	531.70
1lf2	Plasmepsin II	-10.24	531.70
1lhw	Sex hormone-binding globulin	-11.11	302.42
1lnm	Diga16	-11.84	374.53
1loq	Orotidine 5'-monophosphate decarboxylase	-5.04	323.18
1lor	Orotidine 5'-monophosphate decarboxylase	-15.06	339.18
1m0n	2,2-Dialkylglycine decarboxylase	-3.02	392.24
1m0o	2,2-Dialkylglycine decarboxylase	-3.14	380.23
1m0q	2,2-Dialkylglycine decarboxylase	-4.03	352.18
1m48	Interleukin-2	-6.93	447.56
1mai	Phospholipase C delta-1	-9.09	417.07
1mfa	14-3-3-like protein c	-6.86	486.47
1mfl	Erb-b2 interacting protein	-5.30	540.64
1mmb	MMP-8	-11.30	477.65
1mnc	MMP-8	-12.25	349.43
1moq	Glucosamine 6-phosphate synthase	-4.71	259.15
1mq6	Factor xa	-15.18	568.87
1mrw	Pol polyprotein	-13.21	527.69
1mu6	Thrombin	-11.41	432.41
1mu8	Thrombin	-12.25	446.44
1n3i	Purine nucleoside phosphorylase	-12.10	265.29
1n46	Thyroid hormone receptor beta-1	-14.32	367.41
1n4k	Inositol 1	-13.68	417.07
1nfw	Factor xa	-12.20	436.94
1nfy	Factor xa	-12.10	463.99
1nje	Thymidylate synthase	-5.17	306.19
1njj	Ornithine decarboxylase	-2.86	500.59
1njs	Phosphoribosylglycinamide formyltransferase	-10.65	543.46
1nl9	PTP 1b	-8.11	531.57
1nm6	Thrombin	-13.68	511.07
1no6	PTP 1b	-10.09	333.30
1nq7	Nuclear receptor ror-beta	-9.26	339.50
1nvq	Serine/threonine-protein kinase chk1	-11.23	481.54

PDB	Protein	Ebind (kcal mol^{-1})	MW
1nvr	Serine/threonine-protein kinase chk1	-11.04	467.55
1nw4	Uridine phosphorylase	-12.35	267.27
1nwl	PTP 1b	-3.25	468.58
1o0m	Ribonuclease a	-7.01	323.18
1o0n	Ribonuclease a	-5.57	323.18
1o0o	Ribonuclease a	-6.94	425.19
1o2h	Trypsin	-8.37	336.42
1o2p	Trypsin	-6.60	359.84
1o36	Trypsin	-8.11	488.44
1o3d	Trypsin	-9.71	329.38
1o3f	Trypsin	-10.84	328.40
1ocq	Endoglucanase 5a	-7.07	310.33
1oe8	Glutathione s-transferase	-7.51	306.32
1okl	Carbonic anhydrase	-8.21	250.32
1ony	PTP 1b	-9.22	588.64
1osv	Bile acid receptor	-4.74	419.63
1ow4	Pheromone binding protein	-7.73	299.35
1p57	Serine protease hepsin	-5.99	253.29
1pkx	Bifunctional purine biosynthesis protein purh	-9.42	363.20
1pme	Erk2 map kinase	-12.80	377.44
1pr5	Purine nucleoside phosphorylase	-5.34	266.26
1pvn	Inosine-5'-monophosphate dehydrogenase	-13.37	338.19
1pzi	Heat-labile enterotoxin b subunit	-5.74	331.28
1pzp	Beta-lactamase tem	-4.51	305.34
1q1g	Uridine phosphorylase putative	-11.67	297.36
1q66	Queuine trna-ribosyltransferase	-7.86	313.40
1qb9	Trypsin	-10.13	490.63
1qbv	Thrombin	-7.33	382.49
1qk4	Hypoxanthine phosphoribosyltransferase	-5.73	347.20
1rbo	Ribulose bisphosphate carboxylase/oxygenase (rubisco)	-17.32	353.09
1sl3	Thrombin	-16.13	552.33
1sqa	Urokinase-type plasminogen activator	-12.54	413.49
1sqn	Progesterone receptor	-12.80	300.44
1sr7	Progesterone receptor	-13.75	523.46
1sri	Streptavidin	-8.28	269.28
1stc	Cyclin dependent kinase 2	-11.03	467.55
1ta6	Thrombin	-12.43	480.03
1tom	Thrombin	-11.30	388.56
1ttm	Carbonic anhydrase	-10.00	309.34
1tyr	Transthyretin	-9.53	303.47
1ur8	Chitinase b	-5.92	422.39
1urg	Maltose-binding protein	-7.92	342.30
1usn	MMP-3	-10.54	427.38
1v2k	Trypsin	-8.43	499.04
1v2o	Trypsin	-6.44	433.51
1v48	Purine nucleoside phosphorylase	-10.61	335.21
1vpo	Anti-testosterone (light chain)	-12.96	288.43

PDB	Protein	Ebind (kcal mol^{-1})	MW
1vyf	14 kda fatty acid binding protein	-10.96	281.46
1w1d	3-phosphoinositide dependent protein kinase-1	-8.88	496.05
1xkk	Epidermal growth factor receptor	-11.60	582.08
1yej	Ig antibody d2.3 (light chain)	-10.16	400.33
2bpv	HIV-1 protease	-10.45	622.84
2gss	Glutathione s-transferase	-6.73	302.14
2qwe	Neuraminidase	-10.19	332.32
2qwf	Neuraminidase	-7.71	341.37
2rkm	Oligo-peptide binding protein	-5.31	276.38
2std	Scytalone dehydratase	-13.41	334.68
3gst	Glutathione S-transferase	-9.15	500.55
3std	Scytalone dehydratase	-15.12	364.45
4sga	Proteinase a	-9.94	471.54
4std	Scytalone dehydratase	-14.06	338.18
5std	Scytalone dehydratase	-14.28	375.42
6cpa	Carboxypeptidase	-15.69	476.43
6std	Scytalone dehydratase	-11.76	413.17
7std	Scytalone dehydratase	-14.59	334.68
8cpa	Carboxypeptidase	-12.45	462.40

Table S2. Listing of 87 complexes included in Set 2.

PDB code	Proteins
1bn4	Carbonic anhydrase
1bnn	Carbonic anhydrase
1okl	Carbonic anhydrase
1ttm	Carbonic anhydrase
1ezq	Factor xa
1fjs	Factor xa
1g2l	Factor xa
1ksn	Factor xa
1mq6	Factor xa
1nfw	Factor xa
1nfy	Factor xa
1a30	HIV-1 protease
1aid	HIV-1 protease
1ajv	HIV-1 protease
1bdl	HIV-1 protease
1dmp	HIV-1 protease
1hpo	HIV-1 protease
1hpv	HIV-1 protease
1htf	HIV-1 protease
1hwr	HIV-1 protease
1kzk	HIV-1 protease
2bpv	HIV-1 protease
1biw	MMP-3
1ciz	MMP-3
1d8m	MMP-3
1usn	MMP-3
1a85	MMP-8
1i73	MMP-8
1jan	MMP-8
1jao	MMP-8
1jaq	MMP-8
1jj9	MMP-8
1kbc	MMP-8
1mmb	MMP-8
1mnc	MMP-8
1ecv	PTP 1b
1nl9	PTP 1b
1no6	PTP 1b
1nwl	PTP 1b
1ony	PTP 1b
2std	Scytalone dehydratase
3std	Scytalone dehydratase
4std	Scytalone dehydratase
5std	Scytalone dehydratase
6std	Scytalone dehydratase
7std	Scytalone dehydratase
1bmn	Thrombin
1c4u	Thrombin

1c4v	Thrombin
1d3p	Thrombin
1d3t	Thrombin
1d4p	Thrombin
1eb1	Thrombin
1g30	Thrombin
1ghv	Thrombin
1ghw	Thrombin
1gj4	Thrombin
1gj5	Thrombin
1jwt	Thrombin
1k22	Thrombin
1kts	Thrombin
1mu6	Thrombin
1mu8	Thrombin
1nm6	Thrombin
1qbv	Thrombin
1sl3	Thrombin
1ta6	Thrombin
1tom	Thrombin
1c5q	Trypsin
1f0u	Trypsin
1ghz	Trypsin
1gi1	Trypsin
1gi6	Trypsin
1o2h	Trypsin
1o2p	Trypsin
1o36	Trypsin
1o3d	Trypsin
1o3f	Trypsin
1qb9	Trypsin
1v2k	Trypsin
1v2o	Trypsin
1ejn	Urokinase-type plasminogen activator
1gi8	Urokinase-type plasminogen activator
1gj7	Urokinase-type plasminogen activator
1gj9	Urokinase-type plasminogen activator
1gja	Urokinase-type plasminogen activator
1sqa	Urokinase-type plasminogen activator

Table S3. Accuracy of the scoring functions on 5 subsets of set 1. Each cell presents the ranked correlation coefficient τ for each scoring function within a subset of proteins.

	Thrombin (22)	HIV-1P (11)	Trypsin (13)	MMP-3/8 (14)	trypsin/thrombin /FXa (42)
Min E_{bind}	-16.1	-14.1	-10.8	-12.2	-16.1
Max E_{bind}	-5.7	-5.8	-6.4	-6.1	-5.7
ChemScore (Sybyl)	0.056	0.624	0.564	0.341	0.266
DockScore (Sybyl)	0.368	0.624	0.385	0.297	0.459
$\text{DrugScore}^{\text{CSD}}$	0.42	0.55	0.462	0.077	0.443
$\text{DrugScore}^{\text{PDB}}$	0.558	0.477	0.179	0.077	0.556
eHiTS SF	0.455	0.367	0.077	0.209	0.388
FlexXScore (Sybyl)	-0.022	0.661	0.385	0.209	0.075
GlideScore	0.307	0.514	0.308	0.165	0.333
GoldScore (Sybyl)	0.55	0.44	0.179	0.077	0.503
Hammerhead (Cerius2)	0.304	0.477	0.487	0.253	0.313
LigScore1 (Cerius2)	0.429	0.44	0.179	0.088	0.369
LigScore2 (Cerius2)	0.451	0.514	0.245	0.165	0.419
PLP1 (Cerius2)	0.481	0.587	0.308	0.099	0.487
PLP2 (Cerius2)	0.394	0.587	0.308	0.143	0.447
PMF (Sybyl)	0.29	-0.183	-0.026	-0.209	0.303
PMF (Cerius2)	0.247	0.55	0.359	-0.209	0.354
RankScore	0.68	0.55	0.359	0.099	0.605
Surflex SF	0.177	0.477	0.308	0.231	0.298
XScore	0.312	0.734	0.436	0.199	0.414
MW	0.489	0.257	0.128	0.341	0.459

Table S4. Accuracy (τ) of the scoring functions on three subsets of Set 1.

	Hydrophobic proteins (85 complexes)	Hydrophilic proteins (93 complexes)	Metallo enzymes (23 complexes, all hydrophobic)
ChemScore (Sybyl)	0.301	0.277	0.043
DockScore (Sybyl)	0.230	0.282	0.083
DrugScore^{CSD}	0.218	0.357	-0.036
DrugScore^{PDB}	0.176	0.346	-0.107
eHiTS SF	0.227	0.365	0.02
FlexXScore	0.223	0.211	0.083
GlideScore	0.288	0.240	0.04
GoldScore (Sybyl)	0.278	0.315	0.02
Hammerhead (Cerius2)	0.292	0.221	0.043
LigScore1 (Cerius2)	0.109	0.109	0.083
LigScore2 (Cerius2)	0.190	0.317	0.051
PMF (Sybyl)	0.245	0.380	-0.051
PMF (Cerius2)	0.269	0.352	-0.043
PLP1 (Cerius2)	0.109	-0.064	0.075
PLP2 (Cerius2)	0.255	0.004	0.051
RankScore	0.223	0.402	-0.012
Surflex SF	0.282	0.238	0.067
XScore	0.307	0.361	0.1
MW	0.153	0.289	0.249

Table S5. Accuracy of the scoring functions (τ) on set 2 when waters are kept for docking and scoring, kept for docking and removed for scoring, removed for both docking and scoring, or made displaceable.

	native protein conformation	Non-native protein conformation[a]			
	Waters kept	Wet/wet	Wet/dry	Dry/dry	Displaceable
ChemScore (Sybyl)	0.338	0.307	0.307	0.328	0.330
DockScore (Sybyl)	0.280	0.273	0.274	0.267	0.281
DrugScore^{CSD}	0.351	0.295	0.296	0.320	0.326
DrugScore^{PDB}	0.333	0.336	0.340	0.337	0.342
eHiTS SF	0.366	0.359	0.380	0.380	0.386
FlexXScore (Sybyl)	0.164	0.081	0.048	0.069	0.100
GlideScore	0.307	0.071	0.076	0.126	0.133
GoldScore (Sybyl)	0.259	0.274	0.274	0.272	0.286
Hammerhead SF (Cerius2)	0.176	0.116	-0.009	0.001	0.116
LigScore1 (Cerius2)	0.114	-0.017	-0.026	0.007	0.015
LigScore2 (Cerius2)	0.266	0.076	0.064	0.101	0.135
PLP1 (Cerius2)	0.289	0.261	0.253	0.276	0.280
PLP2 (Cerius2)	0.305	0.260	0.241	0.261	0.270
PMF (Sybyl)	0.230	0.209	0.209	0.203	0.212
PMF (Cerius2)	0.315	0.297	0.288	0.292	0.296
RankScore	0.295	0.242	0.243	0.286	0.286
Surflex SF	0.218	0.050	0.055	0.056	0.073
XScore	0.413	0.403	0.405	0.412	0.412
MW			0.276		

[a] random sampling of cross-docked structures, 1 protein per ligand; maximum and minimum correlation values displayed for 10,000 trials.

Table S6. Accuracy of the scoring functions (τ) on the complete set 2 when protein conformational ensembles were considered.

	Self-docking	Protein conformational ensemble[a]			
Water molecules	Wet/wet	Wet/wet	Wet/dry	Dry/dry	displaceable
ChemScore (Sybyl)	0.338	0.317	0.317	0.290	0.305
DockScore (Sybyl)	0.280	0.260	0.260	0.235	0.261
DrugScore^{CSD}	0.351	0.296	0.297	0.311	0.315
DrugScore^{PDB}	0.333	0.329	0.333	0.320	0.336
eHiTS SF	0.366	0.451	0.441	0.441	0.451
FlexXScore (Sybyl)	0.164	0.102	0.052	0.058	0.098
GlideScore	0.307	0.228	0.219	0.267	0.276
GoldScore (Sybyl)	0.259	0.287	0.287	0.251	0.282
Hammerhead SF (Cerius2)	0.176	0.135	0.007	0.002	0.135
LigScore1 (Cerius2)	0.114	0.072	0.046	0.048	0.057
LigScore2 (Cerius2)	0.266	0.170	0.113	0.129	0.177
PLP1 (Cerius2)	0.289	0.294	0.280	0.281	0.290
PLP2 (Cerius2)	0.305	0.298	0.279	0.280	0.298
PMF (Sybyl)	0.230	0.215	0.215	0.215	0.219
PMF (Cerius2)	0.315	0.281	0.280	0.272	0.277
RankScore	0.295	0.302	0.302	0.305	0.314
Surflex SF	0.218	0.105	0.075	0.075	0.112
XScore	0.413	0.391	0.391	0.390	0.390
MW			0.277		

[b] best scored complex for each ligand.