Supporting Information

Combination of pose and rank consensus in docking-based virtual screening: the best of both worlds

Valeria Scardino[1,2], Mariela Bollini[3], Claudio N. Cavasotto[2,4,5,*]

[1] Meton AI, Inc., Wilmington, DE, 19801
[2] Austral Institute for Applied Artificial Intelligence, Universidad Austral, Pilar, Buenos Aires, Argentina
[3] Centro de Investigaciones en BioNanociencias (CIBION), Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Ciudad de Buenos Aires, Argentina
[4] Computational Drug Design and Biomedical Informatics Laboratory, Translational Medicine Research Institute (IIMT), Universidad Austral-CONICET, Pilar, Buenos Aires, Argentina
[5] Facultad de Ciencias Biomédicas, and Facultad de Ingeniería, Universidad Austral, Pilar, Buenos Aires, Argentina

* Corresponding Author
CCavasotto@austral.edu.ar; cnc@cavasotto-lab.net

Table S1: The 34 target proteins used in the molecular docking calculations

Receptor	Receptor code	PDB entry	Resolution (Å)	Co-factors[a]	Water molecules[b]
Thymidine kinase	KITH	2B8T	2.0	-	-
Phospholipase A2	PA2GA	1KVO	2.0	Ca^{2+}	-
Coagulation factor VII	FA7	1W7X	1.8	-	1
Hexokinase type IV	HXK4	3F9M	1.5	-	-
Cyclin-dependent kinase 2	CDK2	1FVV	2.8	-	-
Cyclooxygenase-1	COX1	2OYU	2.7	-	-
Fatty acid-binding protein 4	FABP4	2NNQ	1.8	-	4
Heat shock protein 90 alpha	HSP90a	1UYG	2.0	-	4
Estrogen receptor alpha	ESR1	3ERT	1.9	-	-
Neuraminidase	NRAM	1B9V	2.3	Ca^{2+}	-
β_2 adrenergic receptor (agonist bound)	ADRB2	4LDO	3.2	-	-
HMG-CoA reductase	HMDH	3CCW	2.1	-	-
Dopamine D_3 receptor (antagonist bound)	DRD3	3PBL	2.8	-	-
Histone deacetylase 2	HDAC2	5IXO	1.7	Zn^{2+},Ca^{2+}	-
Leukocyte Function Associated Antigen-1	LFA1	2ICA	1.6	-	-
Leukotriene A4 hydrolase	LKHA4	3CHP	2.1	Zn^{2+}	2
Urokinase-type plasminogen activator	UROK	1SQT	1.9	-	2
Tyrosine-protein kinase ABL	ABL1	2HZI	1.7	-	-
Protein-tyrosine phosphatase 1B	PTN1	2AZR	2.0	-	-
Inhibitor of apoptosis protein 3	XIAP	3HL5	1.8	-	-
Androgen Receptor	ANDR	2AM9	1.6	-	-
Renin	RENI	3G6Z	2.0	-	-
Glutamate receptor ionotropic, AMPA 2	GRIA2	3KGC	1.5	-	-
Aldose reductase	ALDR	2HV5	1.6	Napd	-
Dihydrofolate reductase	DYR	3NXO	1.4	Nadph	-
Dihydroorotate dehydrogenase	PYRD	1D3G	1.6	Fmn	-
11-beta-hydroxysteroid dehydrogenase 1	DHI1	3FRJ	2.3	Napd	-
Angiotensin-converting enzyme	ACE	3BKL	2.2	Zn^{2+}	-
Progesterone receptor	PRGR	3KBA	2.0	-	-
Human immunodeficiency virus type 1 reverse transcriptase	HIVRT	3LAN	2.5	-	-
Purine nucleoside phosphorylase	PNPH	3BGS	2.1	-	-
Protein kinase C beta	KPCB	2I0E	2.6	-	-
Insulin-like growth factor I receptor	IGF1R	2OJ9	2.0	-	-
Phosphodiesterase 5A	PDE5A	1UDT	2.3	Zn^{2+}, Mg^{2+}	2

[a]Within 8 Å of the crystallographic ligand
[b]Within 4 Å of the crystallographic ligand

Table S2: Docking chemical libraries

Receptor code	Actives	Decoys	Source
KITH	132	2866	DUD-E[1]
PA2GA	127	5215	DUD-E[1]
FA7	185	6300	DUD-E[1]
HXK4	127	4802	DUD-E[1]
CDK2	72	2074	DUD[2]
COX1	210	6955	DUD-E[1]
FABP4	57	2855	DUD-E[1]
HSP90a	125	4942	DUD-E[1]
ESR1	133	6555	NRLiSt[3]
NRAM	222	6227	DUD-E[1]
ADRB2	206	8034	GLL/GDD[4]
HMDH	299	8884	DUD-E[1]
DRD3	317	12363	GLL/GDD[4]
HDAC2	238	10366	DUD-E[1]
LFA1	233	8690	DUD-E[1]
LKHA4	244	9477	DUD-E[1]
UROK	306	9933	DUD-E[1]
ABL1	295	10885	DUD-E[1]
PTN1	225	7433	DUD-E[1]
XIAP	129	5213	DUD-E[1]
ANDR	523	14503	DUD-E[1]
RENI	387	6984	DUD-E[1]
GRIA2	297	12061	DUD-E[1]
ALDR	220	9136	DUD-E[1]
DYR	566	17383	DUD-E[1]
PYRD	134	6648	DUD-E[1]
DHI1	519	19623	DUD-E[1]
ACE	802	17150	DUD-E[1]
PRGR	445	15813	DUD-E[1]
HIVRT	639	19773	DUD-E[1]
PNPH	233	7016	DUD-E[1]
KPCB	245	8847	DUD-E[1]
IGF1R	226	9407	DUD-E[1]
PDE5A	705	28532	DUD-E[1]

Table S3: EF values with Active/Selected (A/S) molecule rate for a pose consensus alone strategy of at least two matching poses (2 MPs), three (3 MPs) and four (4 MPs). The best EF for each target is highlighted. This table is an extension of Table 3.

Receptor	A/S	2 MPs	A/S	3 MPs	A/S	4 MPs
KITH	105/1654	1.4	58/512	2.6	17/83	4.7
PA2GA	83/2146	1.6	33/411	3.4	9/81	4.7
FA7	119/2568	1.6	56/571	3.4	9/100	3.2
HXK4	97/3009	1.3	30/828	1.4	6/135	1.7
CDK2	63/1597	1.2	44/674	1.9	25/225	3.3
COX1	192/5999	1.1	116/2945	1.3	51/1144	1.5
FABP4	49/2160	1.2	25/862	1.5	8/270	1.5
HSP90a	98/3822	1.0	49/1568	1.3	29/534	2.2
ESR1	95/4187	1.2	56/1575	1.9	37/470	4.1
NRAM	170/2760	1.8	90/559	4.7	14/72	5.6
ADRB2	160/5512	1.2	49/1581	1.2	3/274	0.4
HMDH	105/2395	2.3	40/404	5.2	8/59	7.1
DRD3	215/7563	1.1	57/2168	1.1	10/421	0.9
HDAC2	100/4625	1.2	27/723	2.1	3/93	1.8
LFA1	63/4220	0.9	28/1259	1.4	13/290	2.8
LKHA4	123/4489	1.5	29/863	1.9	4/128	1.8
UROK	113/4518	1.5	54/885	3.8	22/137	9.9
ABL1	135/5936	1.4	44/1606	1.6	13/280	2.8
PTN1	80/3609	1.3	27/791	1.9	3/129	1.3
XIAP	33/1277	1.4	8/103	4.1	1/7	7.5
ANDR	230/10506	1.2	114/3331	1.9	41/487	4.6
RENI	27/988	1.8	8/58	9.3	3/7	28.9
GRIA2	111/6066	1.4	68/1493	3.5	31/307	7.7
ALDR	145/7049	1.2	88/2482	2.0	41/585	4.0
DYR	150/8481	1.3	40/1794	1.7	10/303	2.5
PYRD	87/4001	1.3	44/1014	2.6	11/193	3.4
DHI1	247/13090	1.1	113/3954	1.7	41/908	2.7
ACE	95/4355	1.3	27/328	5.0	1/11	5.6
PRGR	278/12903	1.2	172/5092	1.8	82/1258	3.5
HIVRT	250/10477	1.4	78/2404	1.8	26/389	3.8
PNPH	96/5285	1.2	62/1823	2.3	35/431	5.6
KPCB	90/4404	1.3	41/1034	2.6	19/188	6.6
IGF1R	65/3059	1.4	15/420	2.3	1/39	1.6
PDE5A	310/11479	1.9	122/1787	4.8	39/193	14.2
Average	129/5182[a]	1.4	56/1409[a]	2.7	20/301[a]	4.8

[a]Average(A)/Average(S)

Table S4: Possible combinations of the number of required MPs and ranking thresholds considered. The average EF value obtained within the 34 benchmark targets is presented along with the minimum and maximum EFs, and the average number of actual ligands (actives) retrieved. The selected combinations (Options A, B and C) are highlighted.

# MPs	Ranking threshold (%)	Average EF	Min EF	Max EF	Average number of actives
2	5[i]	17.9	1.8	48.2	34
2	10	11.1	2.6	28.9	49
2	20	6.1	1.7	12.7	67
3	5	29.1	0.0	67.5	10
3	10[ii]	23.6	0.0	60.0	16
3	20	15.4	1.6	49.1	25
4	5	29.1	0.0	75.8	2
4	10	29.4	0.0	75.8	4
4	15	26.9	0.0	75.8	5
4	20[iii]	25.7	0.0	69.6	6
4	25	22.2	0.0	67.5	7

i. Option A; ii. Option B; iii. Option C

Table S5: Different Pose/Rank filtering strategies that were evaluated in order to assess the best way of combining both consensus techniques. The average and minimum EF for each combination is displayed, along with the average number of actives retrieved. The selected combination (Option D) is highlighted.

Ranking thresholds for 2, 3 and 4 MPs (%)	Average EF	Min EF	Average number of actives
(5, 10, 25)	19.5	3.3	28
(5, 10, 20)[iv]	19.9	3.4	27
(5, 10, 15)	19.9	3.5	25
(5, 12, 25)	19.1	3.1	29
(5, 12, 20)	19.4	3.2	28
(5, 12, 15)	19.5	3.3	26
(5, 14, 25)	18.6	3.0	30
(5, 14, 20)	18.9	3.1	29
(5, 14, 15)	18.9	3.1	27
(6, 10, 25)	17.9	3.0	30
(6, 10, 20)	18.2	3.1	28
(6, 10, 15)	18.2	3.1	27
(6, 12, 25)	17.6	2.9	30
(6, 12, 20)	17.9	2.9	29
(6, 12, 15)	17.8	2.9	28
(6, 14, 25)	17.2	2.8	31
(6, 14, 20)	17.4	2.8	30
(6, 14, 15)	17.4	2.8	29
(7, 10, 25)	16.5	2.8	31
(7, 10, 20)	16.7	2.8	30
(7, 10, 15)	16.7	2.8	28
(7, 12, 25)	16.3	2.7	32
(7, 12, 20)	16.5	2.7	31
(7, 12, 15)	16.4	2.7	29
(7, 14, 25)	16.0	2.6	33
(7, 14, 20)	16.2	2.6	31
(7, 14, 15)	16.1	2.6	30

iv. Option D

Table S6: Comparison of the HR at 1% (HR1) of ECR and the HR of the PRC. The fold increase (PRC HR/ECR HR1) is displayed in the last column.

Receptor	ECR HR1	PRC HR	Fold Increase
KITH	0.55	0.87	1.58
PA2GA	0.60	0.75	1.24
FA7	0.98	0.98	0.99
HXK4	0.14	0.39	2.76
CDK2	0.62	0.65	1.04
COX1	0.09	0.17	1.71
FABP4	0.79	0.80	1.01
HSP90a	0.12	0.38	3.14
ESR1	0.67	0.62	0.93
NRAM	0.16	0.47	3.07
ADRB2	0.61	0.58	0.96
HMDH	0.33	0.47	1.43
DRD3	0.07	0.13	1.56
HDAC2	0.32	0.49	2.04
LFA1	0.17	0.19	1.06
LKHA4	0.27	0.26	0.97
UROK	0.72	0.92	1.28
ABL1	0.42	0.44	1.04
PTN1	0.52	0.42	0.81
XIAP	0.38	0.5	1.30
ANDR	0.17	0.25	1.50
RENI	0.26	0.82	3.20
GRIA2	0.26	0.38	1.47
ALDR	0.58	0.63	1.08
DYR	0.34	0.34	0.99
PYRD	0.45	0.59	1.32
DHI1	0.14	0.15	1.05
ACE	0.23	0.37	1.59
PRGR	0.17	0.32	1.88
HIVRT	0.27	0.29	1.09
PNPH	0.54	0.51	0.94
KPCB	0.69	0.82	1.19
IGF1R	0.29	0.61	2.11
PDE5A	0.25	0.42	1.71
Average	0.39	0.50	1.50[a]

[a]Average of the fold increase values

Table S7: EF values and Active/Selected (A/S) molecule rate for a pose consensus alone strategy using free docking programs using at least two (2 MPs), three (3 MPs), and four matching poses (4 MPs).

Receptor	A/S	2 MPs	A/S	3 MPs	A/S	4 MPs
KITH	96/1481	1.5	42/380	2.5	10/54	4.2
PA2GA	85/2443	1.5	42/450	3.9	8/65	5.2
FA7	115/2549	1.6	49/488	3.5	10/90	3.9
HXK4	105/3186	1.3	44/920	1.9	7/142	1.9
CDK2	61/1513	1.2	39/549	2.1	22/167	3.9
COX1	187/5988	1.1	119/2894	1.4	51/1045	1.7
FABP4	41/1915	1.1	22/714	1.6	7/235	1.5
HSP90a	102/3725	1.1	53/1457	1.5	29/468	2.5
ESR1	103/4446	1.2	55/1598	1.8	28/481	3.1
NRAM	155/2871	1.6	54/513	3.1	9/73	3.6
ADRB2	163/6497	1.0	55/2419	0.9	6/532	0.5
HMDH	106/2976	1.9	30/432	3.6	6/67	4.7
DRD3	196/7289	1.1	47/1854	1.0	10/358	1.1
HDAC2	102/4559	1.3	22/781	1.6	3/91	1.9
LFA1	66/4322	1.0	27/1136	1.5	13/266	3.1
LKHA4	107/4476	1.3	23/780	1.7	3/113	1.5
UROK	116/5144	1.4	62/1084	3.5	39/207	11.6
ABL1	143/6913	1.2	46/2107	1.3	13/450	1.7
PTN1	82/3837	1.2	28/906	1.8	4/183	1.2
XIAP	50/1423	1.8	13/157	4.3	2/3	35.0
ANDR	250/12261	1.1	143/5445	1.4	60/1068	3.0
RENI	23/1174	1.3	4/80	3.4	1/4	16.9
GRIA2	117/6535	1.4	66/1734	2.9	33/337	7.4
ALDR	136/7239	1.1	84/2733	1.8	29/688	2.4
DYR	149/9509	1.2	44/2058	1.6	15/345	3.3
PYRD	96/4362	1.3	48/1218	2.3	12/243	2.9
DHI1	270/13867	1.2	118/4683	1.5	41/1161	2.1
ACE	80/4432	1.1	11/401	1.7	1/29	2.1
PRGR	279/13928	1.1	181/6572	1.5	111/2307	2.6
HIVRT	261/12491	1.2	96/3357	1.6	34/635	3.1
PNPH	93/5295	1.2	62/2026	2.1	33/497	4.5
KPCB	96/5007	1.3	49/1329	2.4	25/318	5.1
IGF1R	79/3169	1.6	20/476	2.7	3/50	3.8
PDE5A	299/13145	1.6	114/2484	3.2	34/310	7.7
Average	130/5587[a]	1.3	56/1653[a]	2.2	21/385[a]	4.7

[a]Average(A)/Average(S)

References

1.   M. M. Mysinger, M. Carchia, J. J. Irwin and B. K. Shoichet, *J. Med. Chem.*, 2012, **55**, 6582-6594.
2.   N. Huang, B. K. Shoichet and J. J. Irwin, *J. Med. Chem.*, 2006, **49**, 6789-6801.
3.   N. Lagarde, N. Ben Nasr, A. Jeremie, H. Guillemain, V. Laville, T. Labib, J. F. Zagury and M. Montes, *J. Med. Chem.*, 2014, **57**, 3117-3125.
4.   E. A. Gatica and C. N. Cavasotto, *J. Chem. Inf. Model.*, 2012, **52**, 1-6.