

Virtual Screening Data Fusion Using Both Structure- and Ligand-Based Methods

Fredrik Svensson, Anders Karlén, and Christian Sköld

Supporting information

Table S1. Enrichment factor in top 100 compounds, 1%, 3%, 5%, 10%, 15%, 20%, and 50% of the data sets. Cells marked in red indicate that not enough compounds were scored by that method to calculate an EF for that percentage.

<i>DS1</i>								
CDK2	100 compounds	1%	3%	5%	10%	15%	20%	50%
Glide	57.69	57.11	24.48	16.32	8.57	5.99	4.54	1.92
Phase	14.42	14.28	5.10	3.67	2.45	1.84	1.79	1.65
Rocs	15.45	15.30	6.12	3.88	2.55	2.18	1.94	1.61
Eon	6.18	6.12	4.42	2.65	1.94	1.63	1.68	1.16
Sum rank	17.51	17.34	8.84	7.14	4.79	3.81	3.47	1.94
Rank vote	22.66	22.44	16.32	14.28	8.57	5.78	4.54	1.94
Sum score	21.63	21.42	14.28	11.42	7.75	5.72	4.49	1.94
Pareto	41.20	40.80	22.78	15.50	8.36	5.99	4.64	1.96
Parallel	41.20	41.82	22.10	15.50	8.26	5.85	4.59	1.94
EF Max	100.95	99.95	33.32	19.99	10.00	6.67	5.00	2.00
COX-2	100 compounds	1%	3%	5%	10%	15%	20%	50%
Glide	29.26	29.37	26.04	18.85	9.84	6.58	4.94	1.99
Phase	33.25	33.25	22.38	13.88	7.06	4.75	3.60	1.78
Rocs	33.25	33.25	28.94	19.05	9.84	6.56	4.92	1.99
Eon	32.25	31.96	23.46	18.21	9.80	6.58	4.94	1.98
Sum rank	33.25	33.25	27.12	18.92	9.87	6.60	4.95	1.99
Rank vote	33.25	33.25	30.77	19.76	9.93	6.62	4.97	1.99
Sum score	33.25	33.25	30.77	19.63	9.90	6.62	4.97	1.99
Pareto	32.25	33.25	30.45	19.56	9.90	6.65	4.98	2.00
Parallel	31.92	31.96	28.73	19.76	9.96	6.65	4.98	2.00
EF Max	33.25	33.25	33.25	20.01	10.00	6.67	5.00	2.00
ERα	100 compounds	1%	3%	5%	10%	15%	20%	50%
Glide	34.27	33.94	13.71	10.21	6.09	4.33	3.44	1.76
Phase	26.21	26.62	13.71	10.47	6.42	4.64	3.64	1.72
Rocs	45.03	44.58	17.69	12.06	6.69	4.59	3.51	1.56
Eon	12.77	12.64	6.19	4.51	3.64	2.65	2.25	1.28
Sum rank	40.32	40.59	15.47	10.21	6.22	4.59	3.67	1.68
Rank vote	43.68	43.25	17.46	11.53	6.16	4.42	3.58	1.68
Sum score	45.70	45.25	16.80	10.34	5.96	4.59	3.61	1.67
Pareto	33.60	33.27	16.14	11.27	6.75	4.95	3.87	1.77

Parallel	41.00	40.59	17.02	10.47	6.49	4.90	3.91	1.75
EF Max	67.21	67.21	33.38	20.02	10.00	6.67	5.00	2.00
fXa	100 compounds	1%	3%	5%	10%	15%	20%	50%
Glide	49.78	49.44	21.69	14.57	8.07	5.68	4.32	1.91
Phase	10.48	10.27	4.72	4.13	2.13	-	-	-
Rocs	9.17	8.99	3.65	2.58	2.52	2.84	3.36	1.87
Eon	12.44	12.84	10.95	8.77	6.52	5.85	4.81	1.99
Sum rank	15.72	15.41	8.59	10.19	8.20	5.76	4.48	2.00
Rank vote	15.06	14.77	9.66	13.15	8.32	6.41	4.94	2.00
Sum score	14.41	14.13	7.95	8.64	8.65	6.28	4.94	2.00
Pareto	33.40	32.75	18.04	13.92	8.71	6.36	4.94	2.00
Parallel	25.54	25.69	18.04	13.67	8.13	5.93	4.71	2.00
EF Max	65.50	65.50	33.29	19.98	10.00	6.67	5.00	2.00
MMP3	100 compounds	1%	3%	5%	10%	15%	20%	50%
Glide	65.63	64.98	29.56	18.40	9.47	6.43	4.83	2.00
Phase	23.95	24.59	12.00	7.89	4.47	3.27	2.68	1.77
Rocs	47.89	47.42	19.61	12.09	6.40	4.74	3.90	1.98
Eon	57.65	57.96	22.25	13.67	7.28	5.15	4.12	1.86
Sum rank	48.78	48.30	20.20	13.50	8.07	5.96	4.74	1.98
Rank vote	60.31	60.59	24.88	17.70	9.12	6.20	4.78	2.00
Sum score	62.09	61.47	24.88	15.78	9.12	6.26	4.83	2.00
Pareto	59.42	58.84	25.76	17.00	9.04	6.08	4.69	1.98
Parallel	56.76	57.08	28.10	17.53	9.12	6.20	4.65	1.96
EF Max	88.69	88.69	33.37	19.98	9.47	6.67	5.00	2.00
NA	100 compounds	1%	3%	5%	10%	15%	20%	50%
Glide	64.30	64.30	24.70	17.21	9.40	6.40	4.80	2.00
Phase	52.24	52.24	22.03	14.81	8.40	6.40	-	-
Rocs	84.39	84.39	30.04	18.01	10.00	6.67	5.00	2.00
Eon	82.39	82.39	29.37	18.01	9.20	6.67	5.00	2.00
Sum rank	90.42	90.42	33.38	20.01	10.00	6.67	5.00	2.00
Rank vote	94.44	94.44	33.38	20.01	10.00	6.67	5.00	2.00
Sum score	98.46	98.46	33.38	20.01	10.00	6.67	5.00	2.00
Pareto	88.41	88.41	33.38	20.01	10.00	6.67	5.00	2.00
Parallel	90.42	90.42	31.38	20.01	10.00	6.67	5.00	2.00
EF Max	100.47	100.47	33.38	20.01	10.00	6.67	5.00	2.00
<i>DS2</i>								
ACE	100 compounds	1%	3%	5%	10%	15%	20%	50%
Glide	174.89	30.62	16.32	10.20	6.12	5.03	4.18	1.92
Phase	310.92	42.87	17.68	12.65	6.94	4.90	3.78	-
Rocs	408.09	59.20	23.13	13.88	7.14	5.03	3.88	1.76
Eon	291.49	63.28	23.13	14.69	9.59	6.53	4.90	2.00
Sum rank	388.65	61.24	23.13	14.69	7.96	5.44	4.18	2.00
Rank vote	388.65	61.24	25.17	15.92	8.16	5.44	4.18	1.88
Sum score	408.09	65.32	23.81	15.10	7.76	5.44	4.18	1.88
Pareto	427.52	59.20	23.81	14.29	7.76	5.44	4.18	2.00
Parallel	427.52	59.20	23.13	15.10	7.55	5.17	4.18	2.00
EF Max	952.2	100.02	33.33	20.00	10.00	6.67	5.00	2.00
AChE	100 compounds	1%	3%	5%	10%	15%	20%	50%
Glide	0	3.81	5.40	6.29	5.14	4.00	3.38	1.70
Phase	9.07	2.86	2.86	1.90	1.05	0.76	0.62	0.65

Rocs	435.55	69.51	23.49	14.29	7.71	5.14	3.86	1.56
Eon	444.62	68.55	25.40	15.62	8.09	5.71	4.38	1.81
Sum rank	81.67	16.19	6.98	5.90	6.48	5.02	3.86	1.66
Rank vote	326.66	65.70	23.81	15.24	7.81	5.21	3.95	1.66
Sum score	190.55	61.89	23.49	15.24	7.81	5.21	3.95	1.66
Pareto	154.26	54.27	23.49	14.67	7.71	5.27	4.00	1.73
Parallel	326.66	64.75	23.81	14.67	7.90	5.33	4.05	1.79
EF Max	907.39	99.97	33.34	20.00	10.00	6.67	5.00	2.00
COX-2	100 compounds	1%	3%	5%	10%	15%	20%	50%
Glide	118.03	52.60	23.85	15.57	8.56	5.90	4.48	1.89
Phase	236.05	52.02	20.21	12.76	7.01	4.98	3.88	1.74
Rocs	260.76	72.14	27.68	16.95	8.68	5.88	4.47	1.88
Eon	208.60	35.93	16.76	11.15	6.75	5.19	4.11	1.85
Sum rank	236.05	69.55	25.76	15.86	8.65	5.90	4.50	1.89
Rank vote	249.78	74.73	27.68	16.95	8.74	5.86	4.43	1.86
Sum score	255.27	75.88	27.49	16.78	8.65	5.80	4.38	1.86
Pareto	203.12	74.15	27.68	17.36	8.85	6.00	4.55	1.93
Parallel	216.84	70.13	27.87	17.01	8.85	6.02	4.57	1.94
EF Max	274.48	100.02	33.33	20.00	10.00	6.67	5.00	2.00
EGFr	100 compounds	1%	3%	5%	10%	15%	20%	50%
Glide	4.31	4.96	5.33	5.86	4.66	3.93	3.23	1.61
Phase	195.97	70.51	26.43	16.76	8.99	6.17	4.66	1.97
Rocs	189.51	76.36	27.56	17.43	9.08	6.11	4.61	1.90
Eon	99.06	25.68	11.86	8.47	5.18	3.93	3.24	1.71
Sum rank	144.28	55.64	21.93	14.10	8.29	5.95	4.58	1.97
Rank vote	180.89	74.56	27.26	16.80	8.69	6.05	4.64	1.96
Sum score	167.97	72.31	26.51	16.35	8.58	6.01	4.61	1.96
Pareto	86.14	57.22	27.86	17.66	9.53	6.56	4.92	2.00
Parallel	124.90	75.46	29.06	18.29	9.48	6.41	4.90	1.99
EF Max	215.35	100.02	33.34	20.00	10.00	6.67	5.00	2.00
fXa	100 compounds	1%	3%	5%	10%	15%	20%	50%
Glide	234.93	54.23	23.24	15.49	8.24	5.73	4.44	1.92
Phase	6.71	1.41	1.17	0.85	0.92	0.85	1.02	1.21
Rocs	26.85	4.93	1.64	1.69	1.34	1.17	1.27	1.07
Eon	13.42	1.41	0.70	0.42	0.21	0.19	0.14	0.35
Sum rank	6.71	2.11	1.64	0.99	0.85	0.94	1.13	1.54
Rank vote	53.70	39.44	13.38	8.17	4.30	3.15	2.43	1.65
Sum score	33.56	7.75	4.93	3.66	2.96	2.54	2.25	1.62
Pareto	33.56	21.13	13.85	9.44	5.70	4.51	3.73	1.80
Parallel	87.26	39.44	17.61	11.97	6.97	5.12	4.01	1.80
EF Max	671.22	100.01	33.34	20.00	10.00	6.67	5.00	2.00
HIVRT	100 compounds	1%	3%	5%	10%	15%	20%	50%
Glide	23.80	10.00	9.17	7.00	5.00	3.50	3.13	1.70
Phase	190.42	20.00	8.33	5.00	3.50	2.67	2.50	1.45
Rocs	166.62	17.50	6.67	5.00	2.75	2.50	2.50	1.70
Eon	142.82	17.50	7.50	5.00	3.25	2.67	2.13	1.80
Sum rank	142.82	20.00	11.67	8.00	5.25	3.83	3.13	1.70
Rank vote	166.62	22.50	8.33	8.50	5.00	3.67	2.88	1.35
Sum score	166.62	20.00	8.33	8.50	5.00	3.67	2.88	1.35
Pareto	166.62	22.50	7.50	6.50	5.25	4.17	3.25	1.70

Parallel	214.22	22.50	7.50	5.00	4.25	3.67	3.00	1.70
EF Max	952.11	100.01	33.34	20.00	10.00	6.67	5.00	2.00
InhA	100 compounds	1%	3%	5%	10%	15%	20%	50%
Glide	190.51	31.75	11.76	8.00	4.12	2.75	2.29	1.34
Phase	145.69	42.33	15.68	9.65	4.94	3.45	2.71	1.48
Rocs	358.61	49.39	16.86	10.35	5.29	3.92	2.94	1.58
Eon	156.89	28.22	10.59	6.82	3.41	2.43	1.88	1.27
Sum rank	313.78	36.45	13.33	8.94	5.29	3.53	2.76	1.32
Rank vote	313.78	42.33	14.90	9.18	5.06	3.53	2.65	1.27
Sum score	313.78	35.28	13.72	8.47	5.06	3.53	2.65	1.27
Pareto	302.58	43.51	16.47	10.35	5.29	3.69	2.82	1.32
Parallel	336.20	42.33	17.25	10.82	5.41	3.76	2.82	1.27
EF Max	952.56	99.95	33.33	20.00	10.00	6.67	5.00	2.00
p38	100 compounds	1%	3%	5%	10%	15%	20%	50%
Glide	0	0.39	1.30	1.41	1.56	1.82	1.76	1.43
Phase	52.19	12.11	4.43	2.66	1.45	1.25	1.09	0.98
Rocs	48.46	7.81	4.04	2.66	1.60	1.38	1.41	1.23
Eon	14.91	6.64	3.52	2.58	2.07	2.01	2.17	1.41
Sum rank	22.37	8.99	5.08	3.59	2.77	2.08	1.93	1.36
Rank vote	52.19	11.33	5.21	4.30	2.89	2.42	2.13	1.34
Sum score	26.09	10.55	5.47	4.30	2.89	2.42	2.13	1.34
Pareto	41.00	11.33	6.25	4.61	2.66	2.03	1.97	1.26
Parallel	48.46	11.33	6.64	5.00	2.73	2.01	1.62	1.22
EF Max	372.76	100.03	33.33	20.00	10.00	6.67	5.00	2.00
PDE5	100 compounds	1%	3%	5%	10%	15%	20%	50%
Glide	0	1.96	0.65	0.39	0.39	0.65	0.59	1.10
Phase	112.03	17.65	8.50	5.88	3.53	2.48	1.96	1.22
Rocs	130.70	17.65	5.88	3.53	2.55	1.83	1.96	1.41
Eon	56.01	9.81	5.23	3.14	1.96	1.96	1.67	1.06
Sum rank	37.34	11.77	5.88	3.92	2.35	2.09	1.96	1.29
Rank vote	149.37	19.61	7.84	5.10	2.94	2.22	1.76	1.29
Sum score	130.70	17.65	7.19	4.71	2.94	2.22	1.76	1.29
Pareto	74.68	21.57	7.84	4.71	3.14	2.22	1.76	1.25
Parallel	130.70	19.61	7.19	5.10	3.53	2.48	2.06	1.37
EF Max	952.22	100.02	33.33	20.00	10.00	6.67	5.00	2.00
SRC	100 compounds	1%	3%	5%	10%	15%	20%	50%
Glide	36.90	8.39	3.87	2.71	2.06	1.68	1.48	0.90
Phase	0	0.65	0.86	0.77	0.71	0.60	0.71	0.49
Rocs	0	0.00	0.22	0.52	0.52	0.56	0.45	0.63
Eon	0	0.00	0.65	1.03	0.65	0.56	0.42	0.31
Sum rank	0	0.00	0.00	0.26	0.19	0.30	0.35	0.48
Rank vote	0	5.16	1.94	1.16	0.77	0.73	0.68	0.49
Sum score	0	0.65	0.65	0.65	0.52	0.65	0.68	0.50
Pareto	12.30	3.87	1.94	1.16	1.16	0.82	0.81	0.77
Parallel	12.30	5.16	2.80	1.81	1.42	1.38	1.35	0.90
EF Max	615.01	100.03	33.33	20.00	10.00	6.67	5.00	2.00

Table S2. Average relative EF for DS1

Relative EF	100 compounds	1%	3%	5%	10%	15%	20%	50%
Glide	0.68	0.68	0.70	0.80	0.87	0.89	0.90	0.97
Phase	0.41	0.42	0.40	0.46	0.52	0.44	0.59	0.87
Rocs	0.56	0.55	0.53	0.56	0.64	0.69	0.75	0.92
Eon	0.48	0.48	0.48	0.55	0.65	0.71	0.76	0.86
Sumrank	0.58	0.58	0.57	0.67	0.79	0.83	0.88	0.97
Rankvote	0.62	0.62	0.66	0.80	0.88	0.90	0.93	0.97
Sumscore	0.63	0.63	0.64	0.72	0.86	0.90	0.93	0.97
Pareto	0.66	0.66	0.73	0.81	0.89	0.92	0.94	0.98
Parallel	0.65	0.65	0.73	0.81	0.87	0.90	0.93	0.97
EF Max	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

Table S3. Average relative EF for DS2

Relative EF	100 compounds	1%	3%	5%	10%	15%	20%	50%
Glide	0.13	0.20	0.30	0.36	0.46	0.52	0.58	0.78
Phase	0.27	0.26	0.32	0.34	0.39	0.42	0.46	0.56
Rocs	0.36	0.37	0.41	0.43	0.47	0.50	0.55	0.74
Eon	0.24	0.26	0.32	0.34	0.41	0.47	0.50	0.68
Sumrank	0.26	0.28	0.35	0.38	0.48	0.53	0.57	0.76
Rankvote	0.34	0.42	0.47	0.51	0.54	0.57	0.59	0.74
Sumscore	0.31	0.37	0.42	0.47	0.52	0.56	0.59	0.74
Pareto	0.25	0.37	0.47	0.50	0.57	0.61	0.64	0.79
Parallel	0.32	0.41	0.49	0.52	0.58	0.62	0.65	0.80
EF Max	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

Table S4. Average relative EF for all the data sets.

Relative EF	100 compounds	1%	3%	5%	10%	15%	20%	50%
Glide	0.34	0.44	0.50	0.58	0.66	0.71	0.74	0.87
Phase	0.33	0.34	0.36	0.40	0.46	0.43	0.52	0.71
Rocs	0.43	0.46	0.47	0.50	0.55	0.60	0.65	0.83
Eon	0.33	0.37	0.40	0.45	0.53	0.59	0.63	0.77
Sum rank	0.38	0.43	0.46	0.52	0.64	0.68	0.72	0.86
Rank vote	0.45	0.52	0.56	0.66	0.71	0.74	0.76	0.85
Sum score	0.43	0.50	0.53	0.59	0.69	0.73	0.76	0.85
Pareto	0.40	0.51	0.60	0.66	0.73	0.76	0.79	0.88
Parallel	0.44	0.53	0.61	0.67	0.73	0.76	0.79	0.89
EF Max	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

Examples of data fusion calculations

The top ten glide hits from the ACE data set are used to illustrate the different data fusion methods.

Table S5. Parallel selection calculated from the raw scores

Compound title	Active	Glide score	Phase score	ROCS score	EON score	Parallel Selection
ZINC03798786	No	-11.5095907	1.78537551	0.994	1.163	1
ZINC04218120	No	-11.4582187	1.64951414	0.83	0.663	5
ZINC01492143	No	-11.4508012	1.23655119	0.755	0.945	8
ZINC03988734	No	-11.4422131	1.06001352	0.755	0.945	9
ZINC01554316	No	-11.3383462	1.30499251	0.743	0.851	10
ZINC03874386	No	-11.1849068	1.49706598	0.982	1.061	4
ZINC02726694	No	-11.1587683	1.49124425	1.029	1.048	6
ZINC03814175	Yes	-11.1361274	1.81640255	1.246	1.098	3
ZINC04097772	No	-11.1182701	1.3335404	0.845	1.059	7
ZINC03814183	Yes	-11.0207759	2.02068832	1.363	1.229	2

Table S6. Pareto rank and sum rank calculated from the rank lists.

Compound title	Glide rank	Phase rank	ROCS rank	EON rank	Pareto rank	Sum rank
ZINC03798786	1	3	4	2	1	1
ZINC04218120	2	4	7	10	2	6
ZINC01492143	3	9	8	7	2	7
ZINC03988734	4	10	8	7	3	9
ZINC01554316	5	8	10	9	2	10
ZINC03874386	6	5	5	4	2	4
ZINC02726694	7	6	3	6	1	5
ZINC03814175	8	2	2	3	1	3
ZINC04097772	9	7	6	5	4	7
ZINC03814183	10	1	1	1	1	2

Table S7. Rank vote calculated from the rank list.

Compound title	Glide votes	Phase votes	ROCS votes	EON votes	Total votes	Rank vote
ZINC03798786	1	1	1	1	4	1
ZINC04218120	1	1	0	0	2	4
ZINC01492143	1	0	0	0	1	7
ZINC03988734	1	0	0	0	1	8
ZINC01554316	0	0	0	0	0	10
ZINC03874386	0	0	0	1	1	5
ZINC02726694	0	0	1	0	1	6
ZINC03814175	0	1	1	1	3	3
ZINC04097772	0	0	0	0	0	9
ZINC03814183	0	1	1	1	3	2

To exemplify each method is given four votes to calculate the rank voting ranks.

Generated e-pharmacophores

DSI

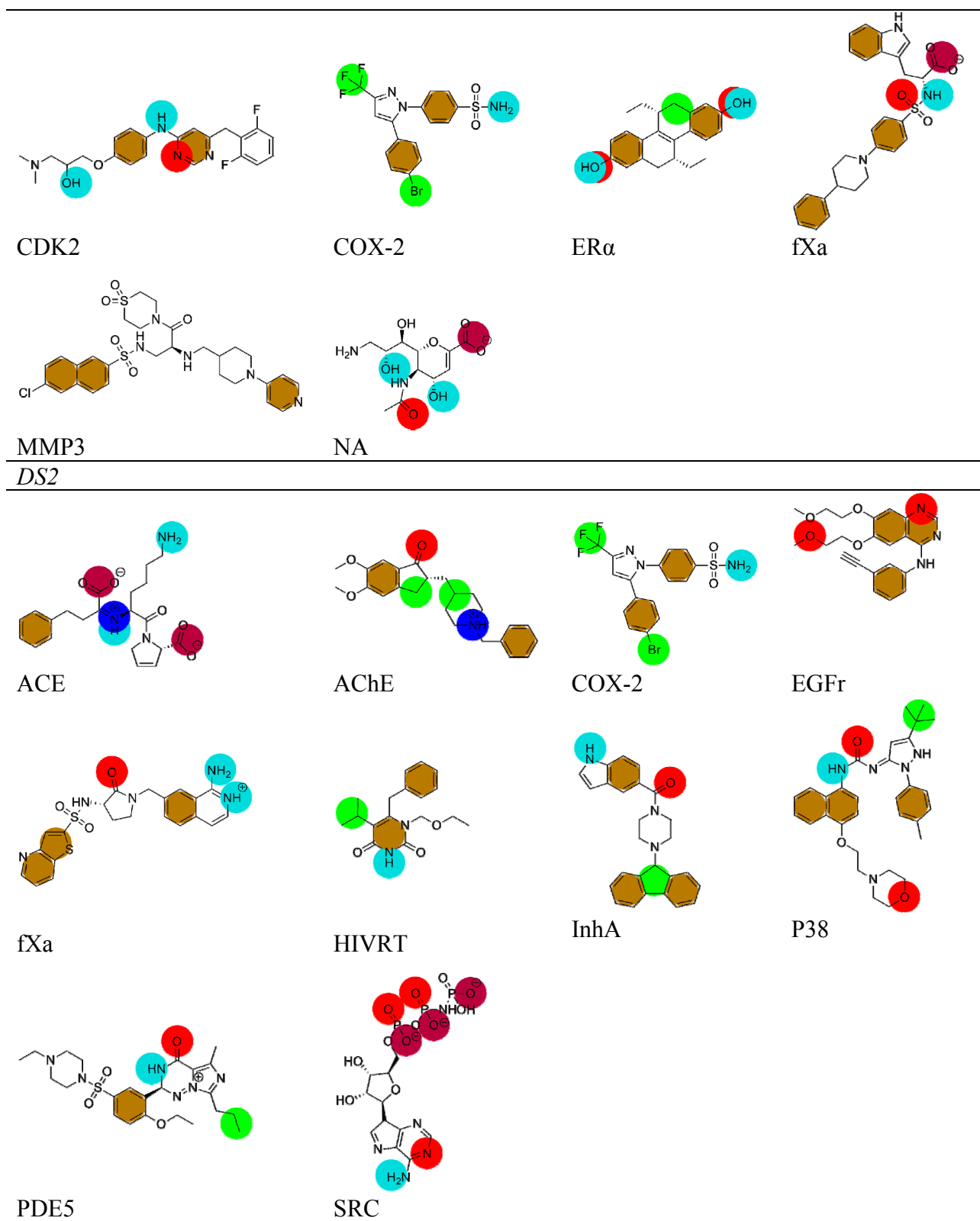


Figure S1. The energetically favorable protein interaction sites as calculated by Glide XP of the respective ligands. These sites were used as pharmacophore sites. Hydrogen bond acceptors are

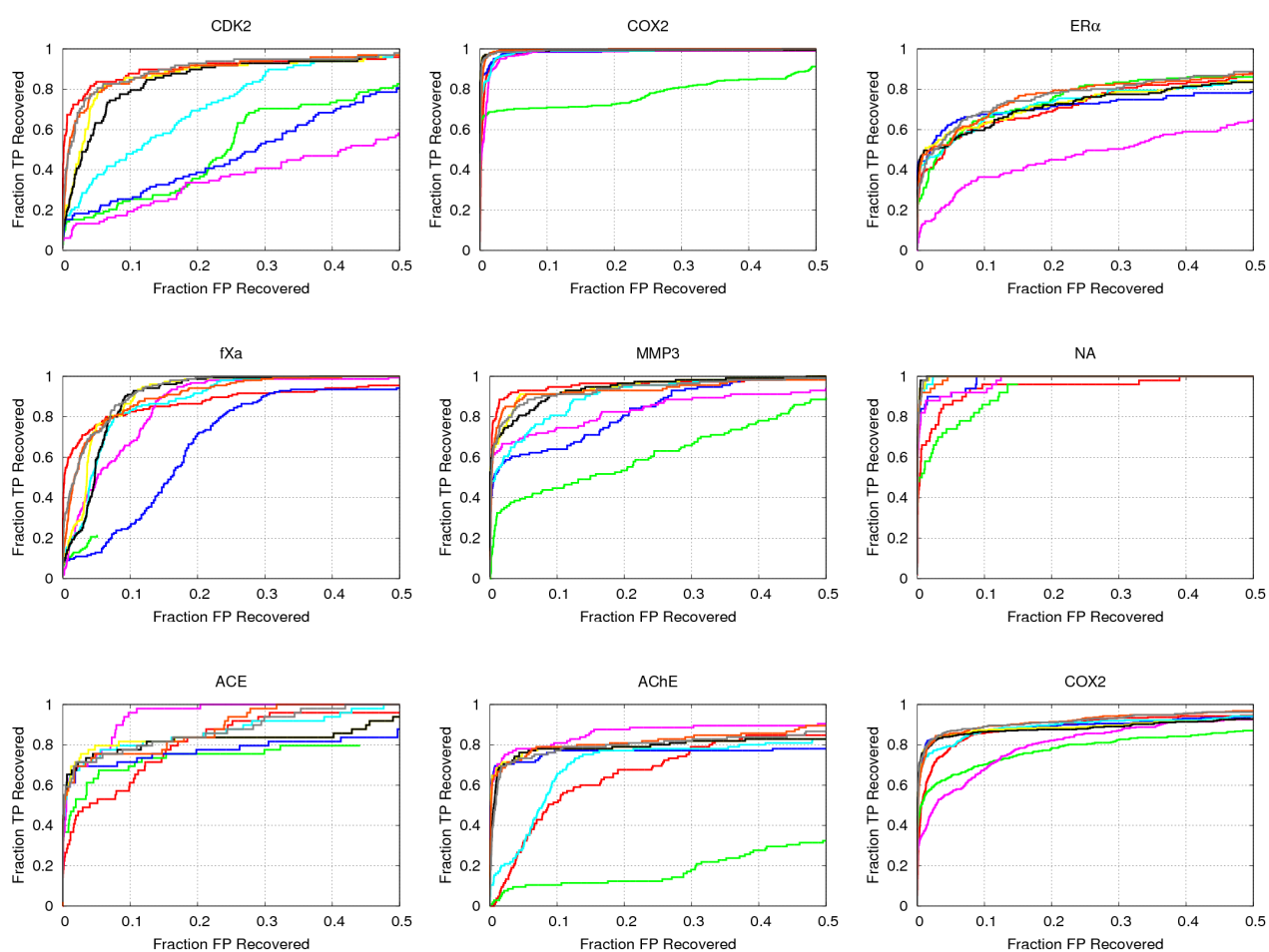
pictured in red, donors in blue, hydrophobic features in green, aromatic features in brown and positive and negative charges in dark blue and dark red, respectively.

Omega settings

Flags used when running OMEGA:

`-rms 0.6 -ewindow 5 -maxconfs 1000`

ROC plots



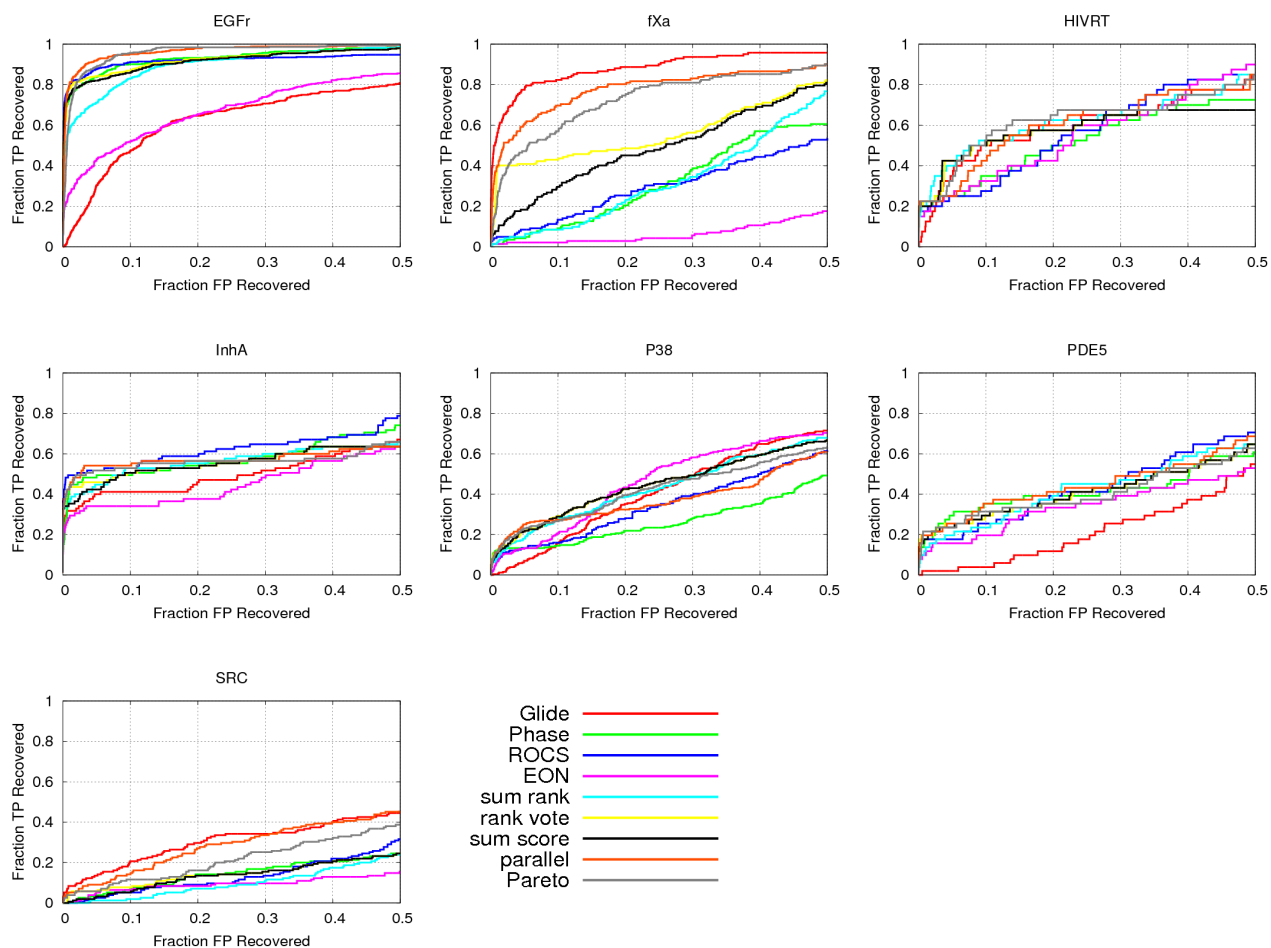


Figure S2. ROC curves showing all the results for all methods in the same plot for all the datasets.

Number of top enrichment factors for each method

Table S8. The number of times each method gave the top one, top two, and top three highest EF in top 100 compounds, 1%, 3%, 5%, 10%, 15%, 20%, and 50% of the data sets.

	100 comp.			1 %			3 %		
	Top 1	Top 2	Top 3	Top 1	Top 2	Top 3	Top 1	Top 2	Top 3
Glide	5	5	5	5	5	5	5	6	6
Phase	3	4	4	2	2	3	1	1	2
ROCS	3	7	10	4	5	5	1	3	4
EON	1	1	1	0	2	2	1	1	1
Sum rank	1	1	3	1	1	3	2	2	2
Rank vote	3	4	12	2	8	14	3	7	9
Sum score	3	6	9	5	6	6	2	3	5
Pareto rank	1	4	5	3	6	8	1	8	13
Parallel	2	7	11	1	7	10	4	9	11

	5 %			10 %			15 %		
	Top 1	Top 2	Top 3	Top 1	Top 2	Top 3	Top 1	Top 2	Top 3
Glide	5	5	5	4	4	5	4	4	5
Phase	1	1	1	1	1	1	1	1	3
ROCS	1	2	4	1	3	4	2	2	2
EON	1	1	1	2	2	2	3	3	3
Sum rank	1	1	2	2	3	5	1	3	5
Rank vote	4	8	11	3	6	10	3	4	8
Sum score	2	4	6	2	4	7	2	4	8
Pareto rank	2	7	10	6	7	12	6	9	14
Parallel	5	11	13	5	10	11	4	10	13

	20 %			50 %		
	Top 1	Top 2	Top 3	Top 1	Top 2	Top 3
Glide	3	5	6	5	7	10
Phase	0	1	2	0	1	1
ROCS	2	3	3	3	4	5
EON	4	4	4	4	5	5
Sum rank	1	4	6	3	4	8
Rank vote	2	4	7	3	3	5
Sum score	3	5	6	3	3	5
Pareto rank	6	10	13	7	10	12
Parallel	5	12	12	6	12	13