

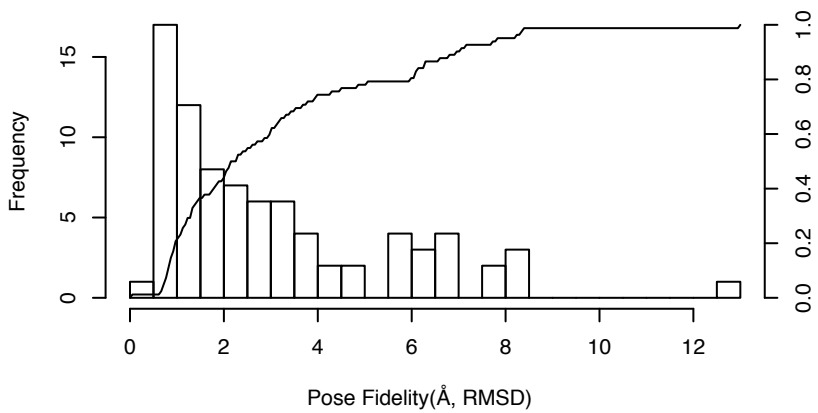
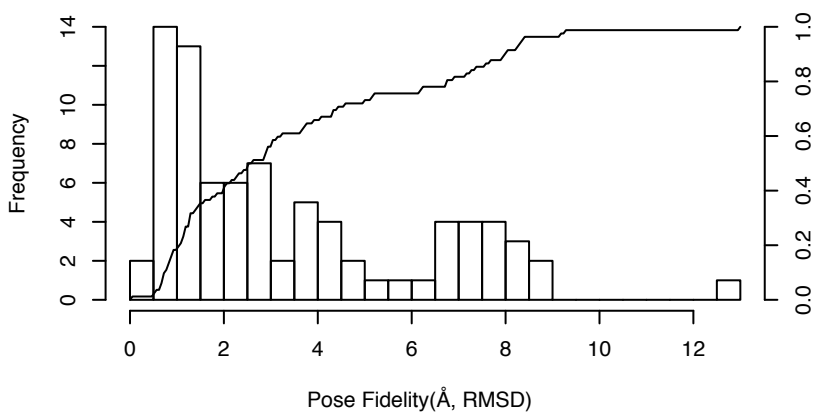
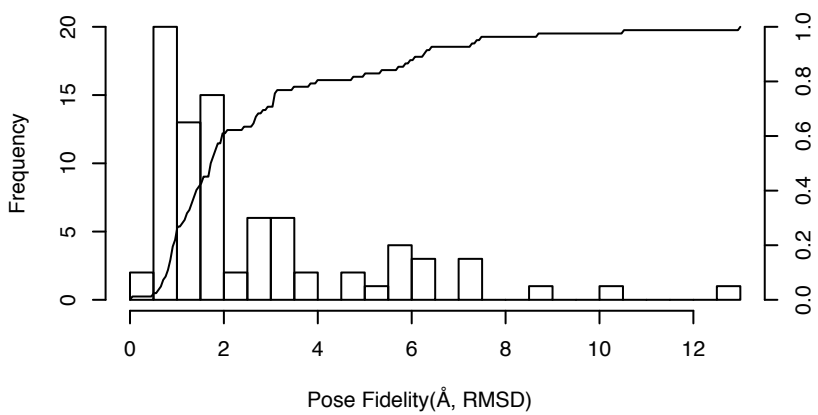
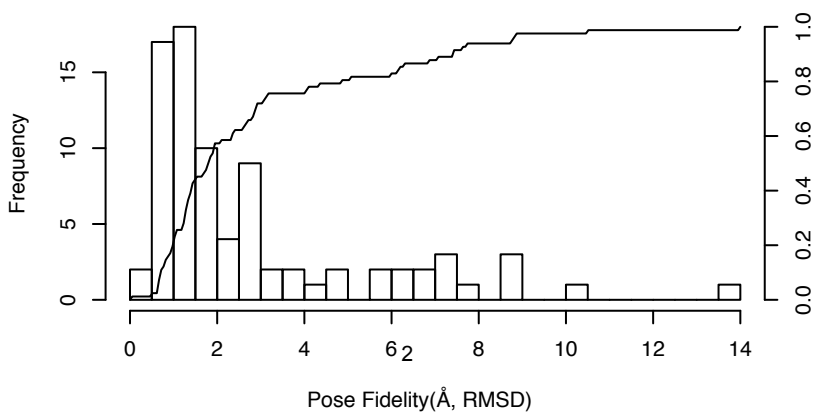
Supporting information for
Automated high-throughput docking: a feasibility study

By John J. Irwin*, Brian K. Shoichet, Michael M. Mysinger, Niu Huang, Francesco Colizzi, Pascal Wassam
and Yiqun Cao

Department of Pharmaceutical Chemistry, Byers Hall, Box 2550, University of California San Francisco, San
Francisco CA 94158-2330

Figure S1. Histogram and cumulative frequency of cases of good pose fidelity (\AA , RMSD) as a function of pose fidelity, for each parameter set. Astex-85 benchmark.

Figure S2. Plot of pose fidelity (\AA , RMSD) vs % rank compared to about 100 property matched decoys (log scale). For Astex-85, for each parameter set separately.

1.A**1.B****2.A****2.B**

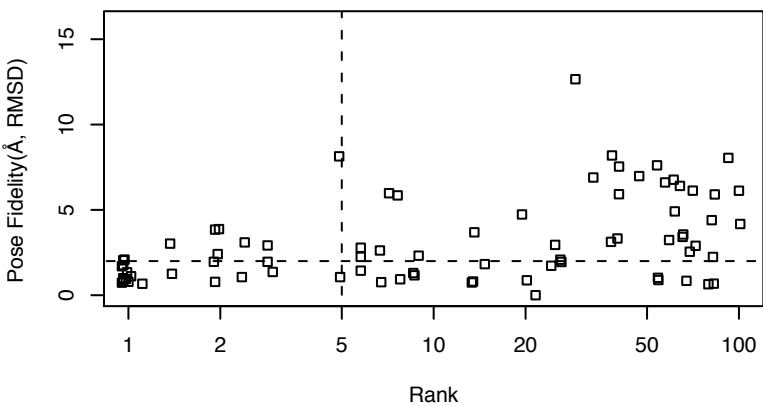
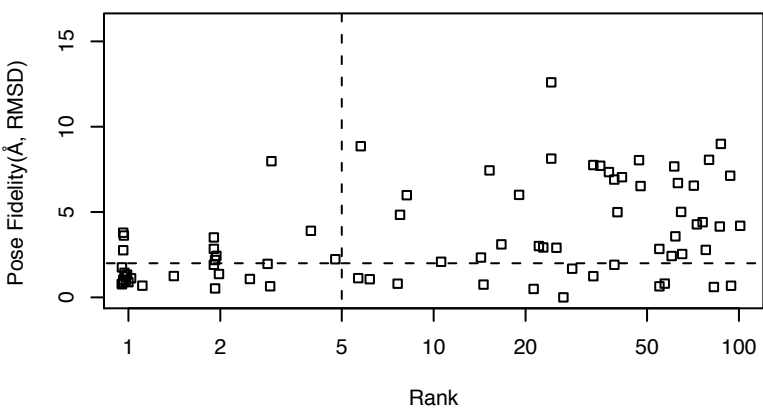
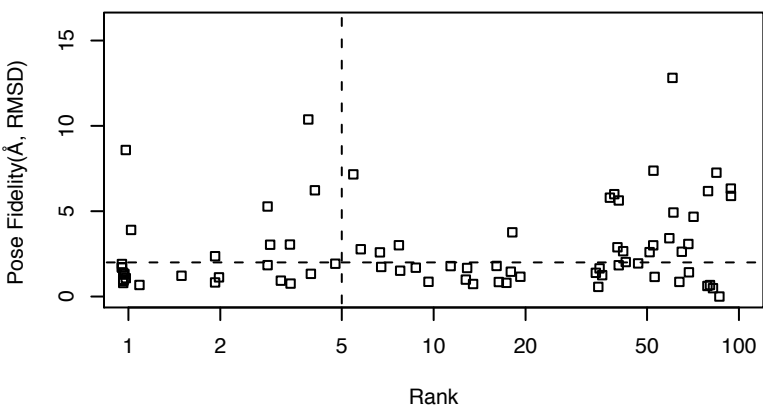
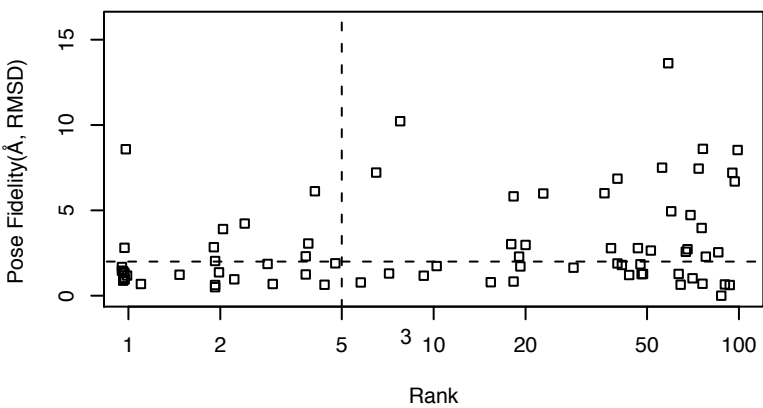
1.A**1.B****2.A****2.B**

Table S1. Processing summary using DOCK Blaster against four benchmarking sets.

	Astex-85	Gold-114	DUD-38	PDB-9050
Number of protein targets	85	114	38	9,050
Launched OK, PDB code only	83	75	37	9,050
Launched OK, when ligand specified	1	19	1	N/A
Ligand docked and scored	83	94	36	7,750
Total elapsed time (h, 400 CPUs max)	3	4	20	250
Total CPU time (h)	130	140	130	13,000

Table S2. Number of good poses (<2 Å RMSD) achieved by DOCK Blaster against four benchmarking sets.

Class	Description	Astex-85	Gold-114	DUD-38	PDB-9050
	Ligand docked and scored	82	94	36	7,750
1	Good pose achieved	51	58	23	3,020

Table S3. DUD Enrichment, comparison between an expert (the DUD paper) and fully automatic docking with DOCK Blaster.

Protein Name	PDB code	Expert (DUD paper)			Automatic (DOCK Blaster)		
		EF1	EF20	EFmax	EF1	EF20	EFmax
AR	2ao6	33.5	3.8	60.2	4	1.83	35.96
ER-agonist	1l2i	19.2	4.5	29.6	7.49	2.01	35.97
ER-antagonist	3ert	12.7	1.3	101.6	15.34	1.92	48.56
GR	1m2z	8.9	1.4	31.7	6.43	1.15	36
MR	2aa2	46.2	3.7	330.0	36	3.33	36
PPAR γ	1fm9	1.0	1.0	1.0	1.23	0.56	2.72
PR	1sr7	0.0	2.0	2.9	3.6	0.74	5.14
RXR α	1mvc	24.8	2.2	148.5	30.86	3.5	36
CDK2	1ckp	13.9	1.4	19.8	N/A	N/A	N/A
EGFr	1m17	2.1	2.4	3.8	1.58	1.54	1.71
FGFr1	1agw	0.0	0.2	1.0	0	1.23	1.31

HSP90	1uy6	8.6	2.0	10.8	4.49	1.53	8.99
P38MAP	1kv2	2.1	2.4	4.1	0	0.63	1.19
SRC	2src	1.2	1.5	3.1	N/A	N/A	N/A
TK	1kim	54.0	5.0	63.0	0	2.28	2.3
FXa	1f0r	14.6	3.8	34.9	19.19	4.04	24.11
Thrombin	3biu	13.7	2.9	18.3	23.5	3.92	39.17
Trypsin	1bju	22.5	2.6	22.5	11.31	4.55	40.73
ACE	1o86	40.4	3.7	141.4	0	1.63	1.74
ADA	1ndw	12.9	2.4	21.5	0	0.44	1.56
COMT	1h1d	0.0	3.3	11.8	0	1.37	3.92
PDE5	1xp0	11.8	2.3	29.1	5.96	1.3	7.6
DHFR	3dfr	21.7	3.5	28.7	53.31	4.3	70.25
GART	1c2t	42.4	3.3	70.7	30.41	4.76	35.48
AChE	1eve	1.9	2.0	3.1	0.95	2.48	2.52
ALR2	1ah3	38.1	2.3	76.2	13.48	2.18	35.96
AmpC	1xgj	17.1	4.7	23.6	0	1.91	2.22
COX-1	1q4g	4.0	1.6	9.9	15.98	3.13	18.65
COX-2	1cx2	20.1	3.3	29.1	27.31	3.36	43.12
GPB	1a8i	22.8	4.1	28.6	0	0.87	2.25
HIVPR	1hpx	3.7	2.2	9.3	0	0.76	1.34
HIVRT	1rt1	5.0	3.0	49.5	5.14	1.79	38.56
HMGR	1hw8	33.9	2.1	198.0	11.02	2.14	23.89
InhA	1p44	0.0	0.3	1.0	8.31	2.29	36.81
NA	1a4g	20.2	3.3	60.6	13.96	3.37	19.94
PARP	1efy	6.0	3.6	6.3	6.18	3.78	18.53
PNP	1b8o	31.7	4.4	158.4	4.16	2.4	37.4
SAHH	1a7a	78.0	5.0	120.0	15.98	3.13	18.65

Table S4. Compare single ligand metric vs multi ligand metric (DUD, adjusted logAUC). DUD Enrichment results compared with the single ligand metric as a function of parameter set. Showing the agreement between single ligand and multi ligand metrics.

Protein Name	PDB code	AdjustedLogAUC				Rank (%)				Pose fidelity (A, RMSD)			
		1A	1B	2A	2B	1A	1B	2A	2B	1A	1B	2A	2B
AR	2AO6	4.3	6.2	3.9	5.8	1	1	1	1	0.55	0.54	0.55	0.52
ER-agonist	1L2I	12.3	13.5	12.9	13.6	2	1	1	1	0.87	0.88	0.85	0.87
ER-antagonist	3ERT	15.3	13.3	16.1	16.0	1	1	1	1	1.51	1.47	1.49	1.49
GR	1M2Z	3.2	4.1	3.1	4.1	1	1	1	1	0.67	0.69	0.68	0.69
MR	2AA2	37.8	44.2	35.1	41.8	1	1	1	1	0.56	0.63	0.59	0.61
PPARg	1FM9	-0.1	-0.1	0.2	-0.2	29	32	48	47	13.77	13.76	15.73	15.75
PR	1SR7	-2.8	-0.3	-2.0	0.8	4	8	4	4	1.06	1.12	1.05	1.28
RXRa	1MVC	37.7	36.7	36.4	34.7	7	6	2	6	1.01	2.09	0.86	0.97
CDK2	1CKP	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
EGFr	1M17	2.9	1.3	3.6	1.9	14	5	64	67	1.70	4.14	8.41	7.52
FGFr1	1AGW	8.7	5.8	6.9	6.1	25	56	28	15	8.21	9.32	7.84	7.74
HSP90	1UY6	-1.9	-1.5	-1.0	-0.6	17	22	62	38	2.65	2.95	0.94	7.62
TK	1KIM	3.6	4.1	2.3	3.7	62	61	33	26	3.79	3.77	1.17	1.17
FXa	1F0R	23.9	21.3	23.9	21.3	2	2	2	3	1.23	7.87	1.11	8.23
Thrombin	3BIU	21.3	35.3	16.9	31.0	11	1	10	1	4.29	4.36	4.21	4.16
Trypsin	1BJU	18.6	21.8	23.2	25.0	5	1	1	1	2.49	2.51	1.90	2.34
ACE	1O86	3.1	2.1	0.2	0.9	3	5	4	8	2.88	2.81	2.81	2.11
ADA	1NDW	-1.0	-1.0	-0.8	-0.8	20	20	28	28	2.86	2.86	3.44	3.44
COMT	1H1D	5.3	4.2	7.2	5.0	84	81	90	82	10.01	10.25	6.14	6.05
PDE5	1XP0	5.7	5.0	6.9	7.0	11	16	13	16	7.68	7.56	7.62	7.59
DHFR	3DFR	45.4	29.4	42.6	27.1	2	4	5	6	5.22	5.15	5.22	5.22
GART	1C2T	49.1	30.4	41.0	27.3	2	2	1	1	9.11	9.32	9.38	9.39
AChE	1EVE	7.6	7.1	7.4	5.6	10	15	9	15	1.48	1.44	1.94	2.04
ALR2	1AH3	17.2	16.7	15.4	17.0	4	3	6	4	3.24	1.90	3.34	1.93
AmpC	1XGJ	4.9	1.5	2.6	-0.8	54	16	57	20	5.37	6.26	5.85	5.76
COX-1	1Q4G	12.8	10.6	14.2	12.7	3	3	3	3	1.96	1.96	1.84	1.86
COX-2	1CX2	27.8	19.8	26.7	18.7	2	2	2	2	0.86	0.94	0.89	0.89
GPB	1A8I	1.6	-2.2	0.4	-5.0	30	54	33	53	0.65	0.69	0.67	0.70
HIVPR	1HPX	-1.8	-2.2	-3.8	-2.9	81	37	74	55	8.44	8.56	14.58	14.75
HIVRT	1RT1	8.0	4.2	6.4	3.7	6	11	8	15	1.05	1.55	1.51	1.44
HMGR	1HW8	13.5	10.2	10.1	8.3	1	1	1	1	2.56	2.55	2.60	2.53
InhA	1P44	10.8	9.0	12.4	9.9	1	1	2	1	1.30	1.28	1.28	1.30
NA	1A4G	29.4	29.1	29.7	27.4	43	33	4	3	1.35	1.39	1.55	1.54
PARP	1EFY	21.6	10.4	20.9	9.8	4	24	3	14	0.76	3.63	2.54	6.58
PNP	1B8O	12.1	11.9	12.7	13.5	22	39	29	40	2.28	2.28	2.36	2.31
SAHH	1A7A	21.5	23.0	20.9	20.3	13	18	17	26	0.65	0.91	0.82	0.96

Table S5. Summary of docking successes, condensing data from tables at <http://data.docking.org/>.

Outcome	Astex-85	Gold-114	DUD-38	PDB-9050
Targets attempted	85	114	38	9,050
Ligand re-docked and scored	82	94	36	7,750
Re-dock within 0-2A, insensitive to parameters	29	34	13	1,528
Re-dock within 0-2A, sensitive to parameters	22	24	10	1,527
Re-dock within 2-3A	11	17	4	1,344
Re-dock within 0-2A, <5% rank, insensitive to parameters	18	17	8	626
Re-dock within 0-2A, <5% rank, sensitive to parameters	11	10	7	772
Re-dock within 0-2A, 5-15% rank	9	11	4	576