LS-align: an atom-level, flexible ligand structural alignment algorithm for high-throughput virtual screening

Jun Hu, Zi Liu, Dong-Jun Yu, and Yang Zhang

Support Information

Supporting Texts

Text S1. Illustration for calculating the $BWJS_i$ score

To illustrate how to calculate the BWJS score in Eq. (4), let's consider *i*th aligned pair of atoms (A_i and B_i) in A and B molecules. Let's further suppose that there are 4 other atoms in A which are connected to A_i through 4 different chemical bonds, i.e., S_{Ai}= {*Single*, *Single, Single*, *Aromatics*}; and there are 3 other atoms in B which are connected to B_i through 3 different chemical bonds, i.e., S_{Bi} = {*Single*, *Double*, *Aromatics*}. Thus, the intersection set $S_{Ii} = S_{Ai} \cap S_{Bi}$ = {*Single*, *Aromatics*}. According to the specific weight for each bond type listed in Table S1, we can calculate the values of $G(S_{Ai})$, $G(S_{Bi})$, and $G(S_{Ii})$ by:

$$\begin{cases} G(S_{Ai}) = w(Single) + w(Single) + w(Single) + w(Aromatic) = 1.0 + 1.0 + 1.0 + 2.0 = 5.0 \\ G(S_{Bi}) = w(Single) + w(Double) + w(Aromatic) = 1.0 + 1.2 + 2.0 = 4.2 \\ G(S_{Ii}) = w(Single) + s(Aromatic) = 1.0 + 2.0 = 3.0 \end{cases}$$

According to Eq. (4), we can then calculate the BWJS value as:

$$BWJS_i = \frac{G(S_{Ii})}{G(S_{Ai}) + G(S_{Bi}) - G(S_{Ii})} = \frac{3.0}{5.0 + 4.2 - 3.0} \approx 0.484$$

Supporting Figures

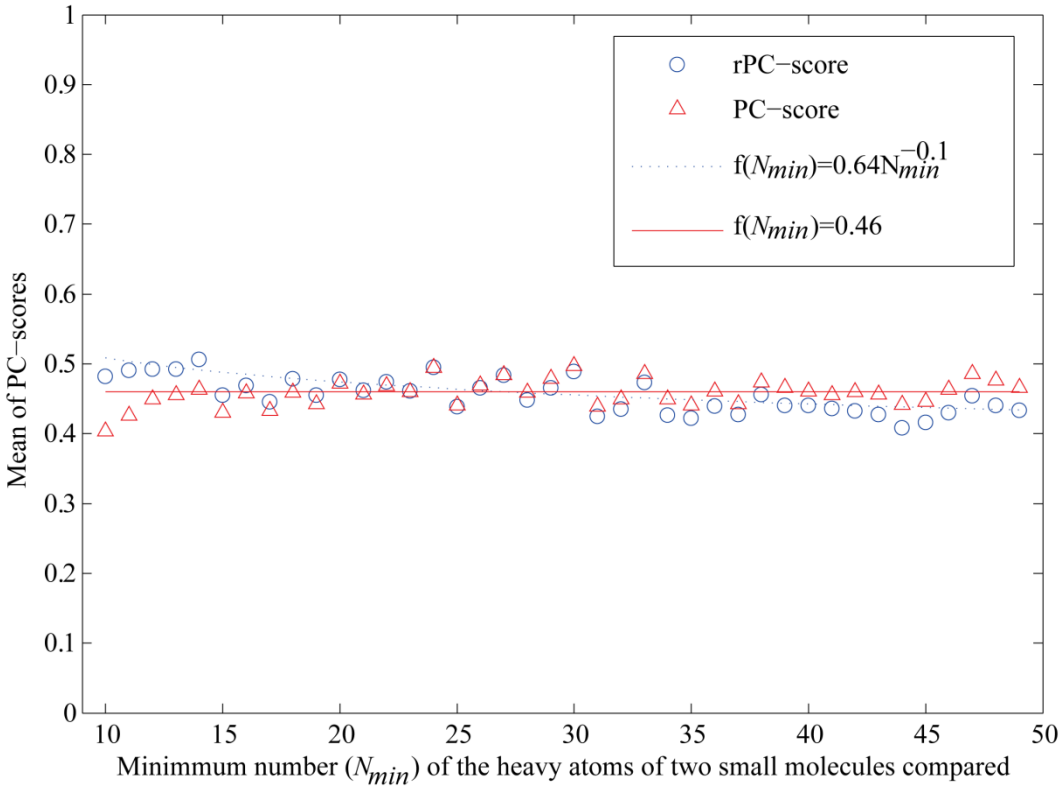

Fig. S1. The average rPC-score (circles) and PC-score (triangles) of 8,000 random ligand pairs as a function of the minimum number (N_{min}) of the heavy atoms of two ligand molecules compared. For the rPC-score, d_0 =1.5 Å; for the PC-score, d_0 is calculated by Eq. (2). The dashed line is a nonlinear least square Marquardt-Levenberg fit of the rPC-score data to a power-law equation $f(N_{min}) = aN_{min}^{b}$. The fit parameters a and b are as indicated. The solid line denotes the horizontal line of PC-score = 0.46.

Fig. S2. Multiple conformers of the ATP (Adenosine-5'-triphosphate) ligand when bound to different protein receptors: (A) ABC transporter HlyB (PDB ID:1xef), (B) RadA C-terminal ATPase (4a6x), (C) Cytoplasmic Asparaginyl-tRNA Synthetase (2xti).

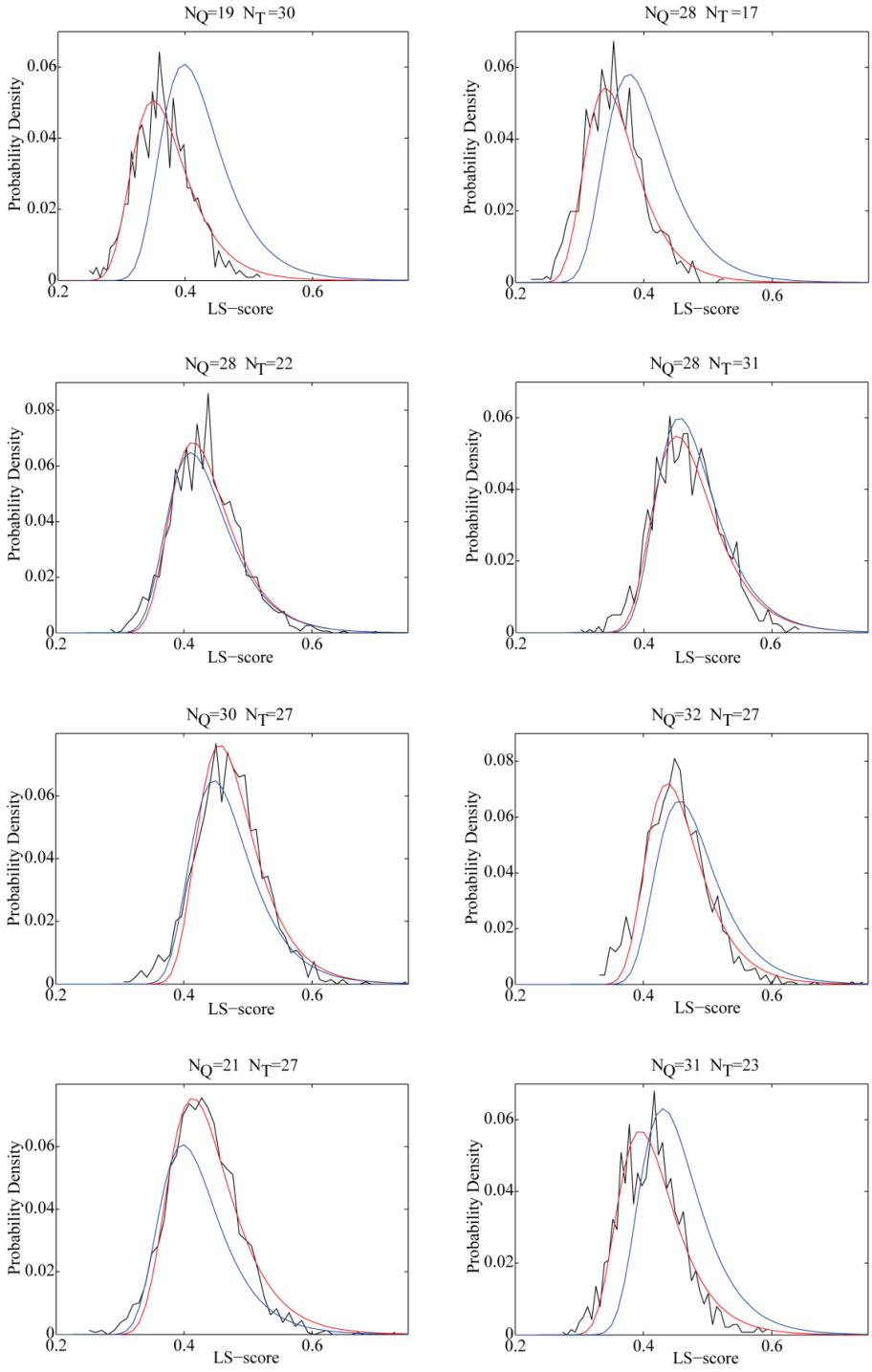

Fig. S3. Distribution of LS-score for random molecule pairs of different sizes. Black lines are the observed probability densities, red lines depict the fitted EVD distributions, and blue lines are for the calculated probability densities using the statistical model described by Equation (9). N_Q and N_T are the numbers of heavy atoms of query and template ligands.

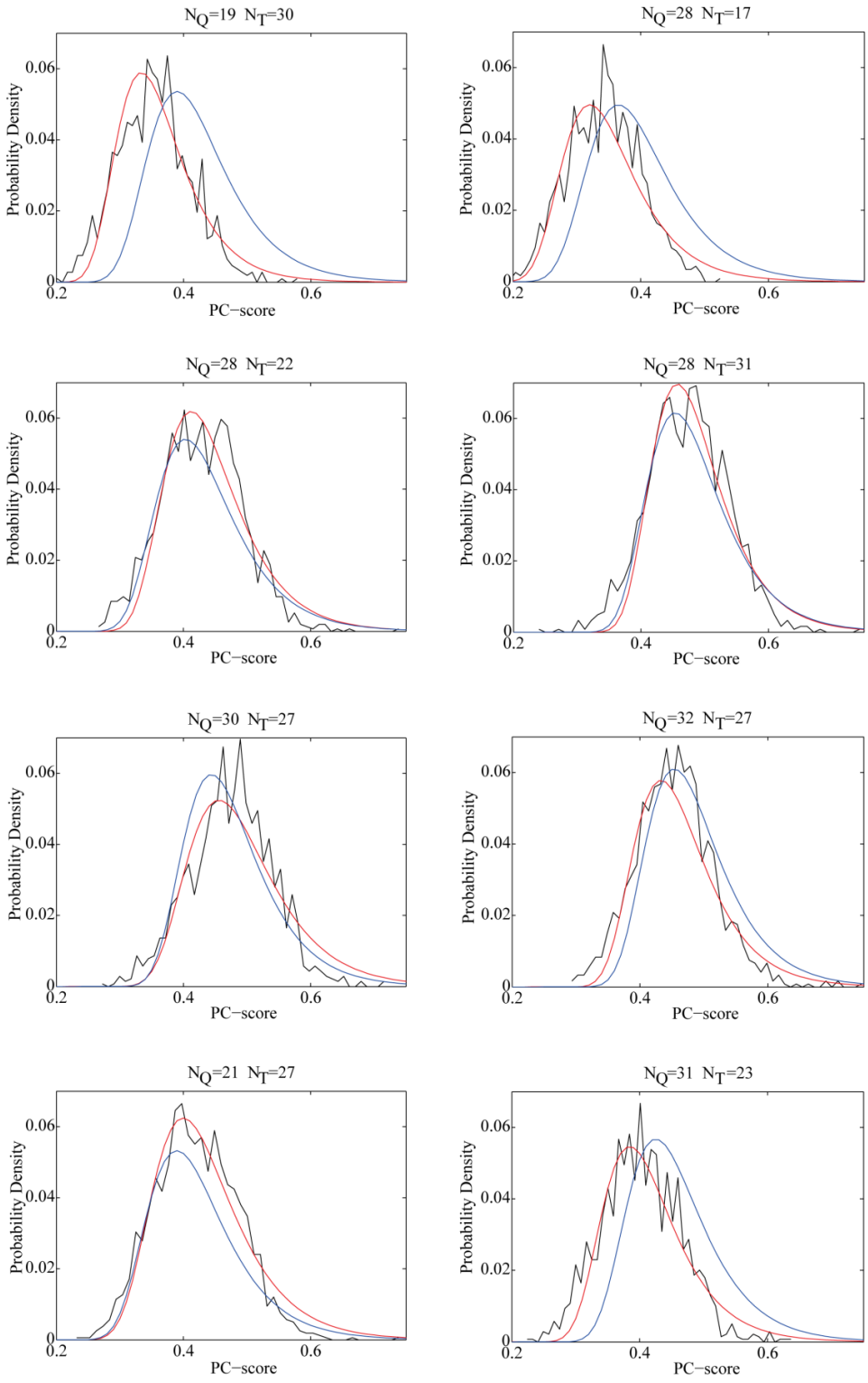

Fig. S4. Distribution of PC-score for random molecule pairs of different sizes. Black lines are the observed probability densities, red lines depict the fitted EVD distributions, and blue lines are for the calculated probability densities using the statistical model described by Equation (9). N_Q and N_T are the numbers of heavy atoms of the query and template ligands.

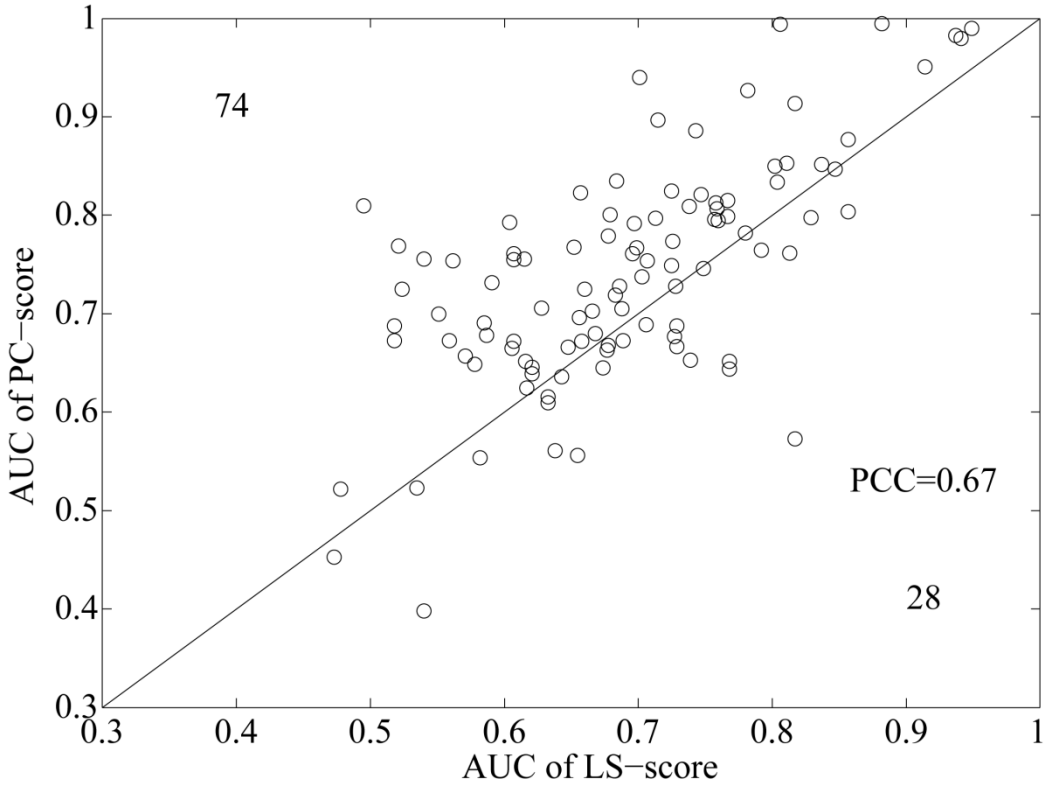

Fig. S5. Head-to-head comparisons of AUC scores between PC-score-based and LS-score-based Rigid-LS-aligns on the DUD-E database. The numbers in each panel means the number of points in the upper and lower triangles, respectively. PCC is the Pearson's correlation coefficient between the AUCs of the two Rigid-LS-aligns.

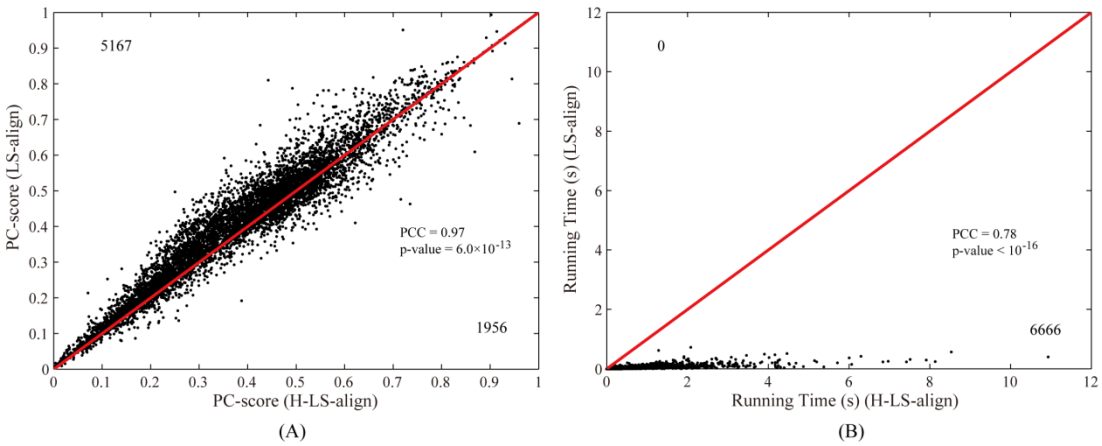

Fig. S6. Head-to-head comparisons of PC-score (A) and running time (B) between LS-align and H-LS-align on the 8,000 ligand pairs. The numbers in each panel represent the number of points in the upper and lower triangles, respectively. 'PCC' is the Pearson's correlation coefficient between the two methods. 'p-value' is calculated in the Wilcoxon signed rank *t*-test.

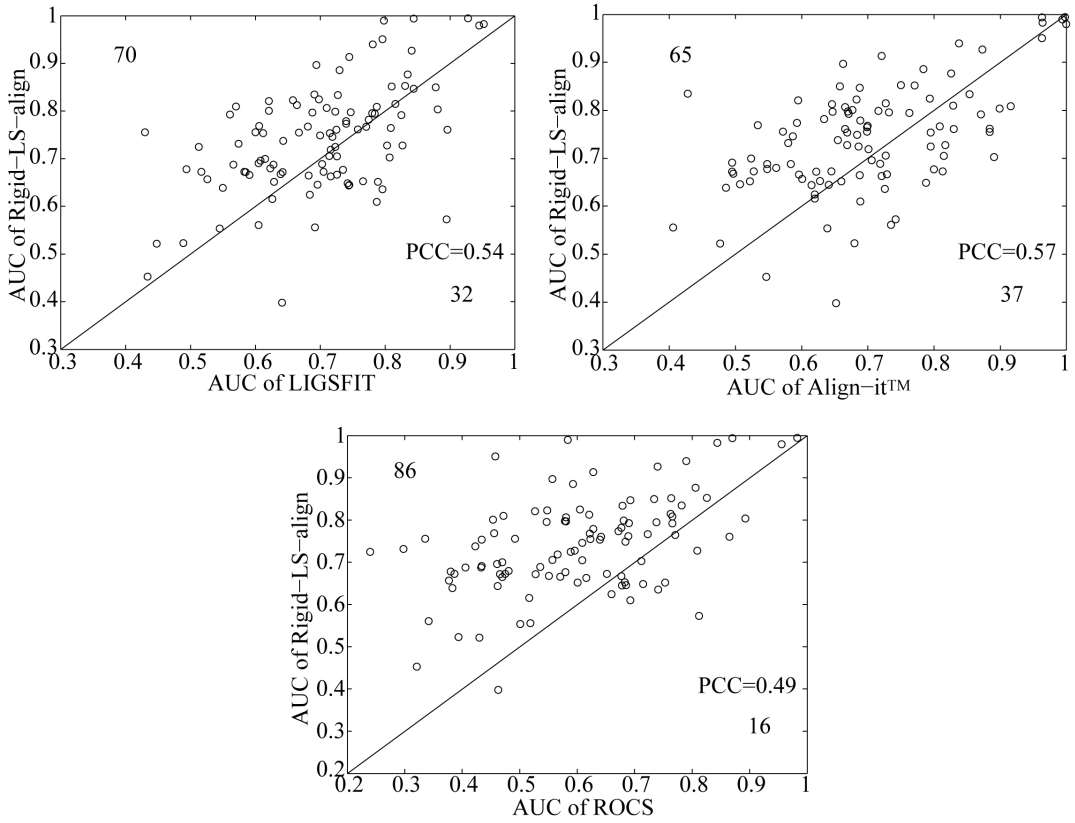

Fig. S7. Head-to-head comparisons of AUC values between Rigid-LS-align and the three control methods from LIGSIFT, Align-It™, and ROCS, using the single rigid conformers of DUD-E database molecules. The numbers in each panel means the number of points in the upper and lower triangles, respectively. PCC is the Pearson's correlation coefficient between the AUCs of the two compared methods.

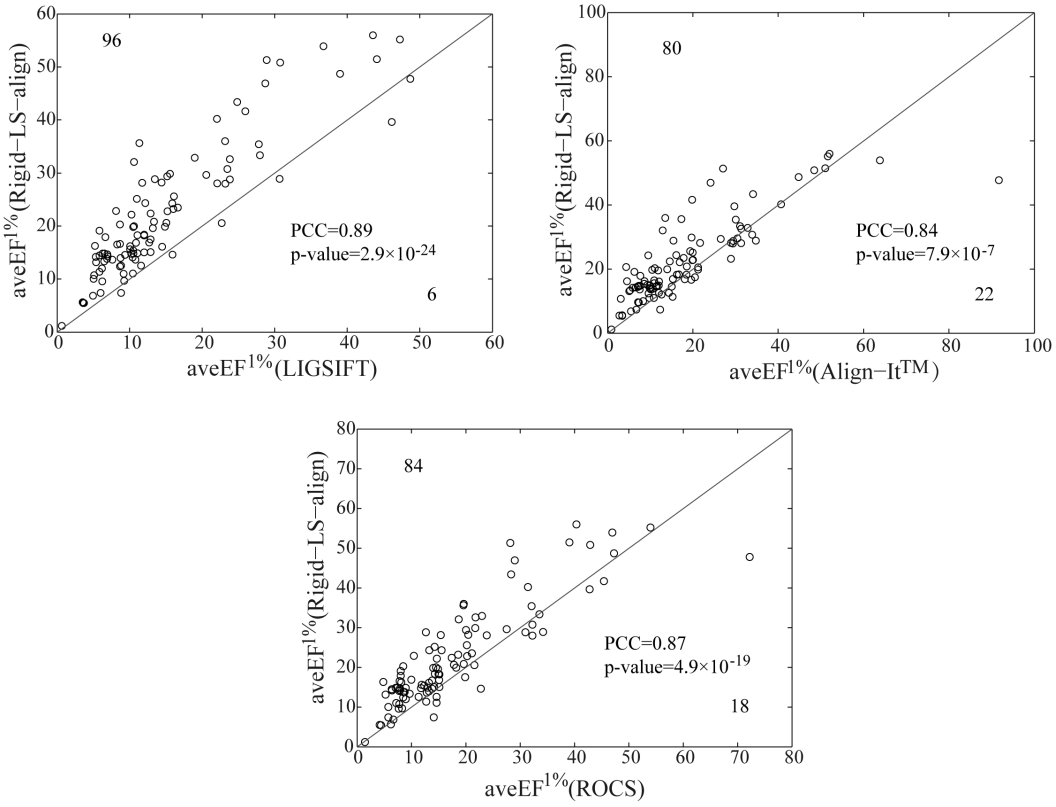

Fig. S8. Head-to-head comparisons of the average $EF^{1\%}$ values between Rigid-LS-align and LIGSIFT, Align-It™, and ROCS. Here, each of the nearly 224 active ligands was used as the seed to rank all other ligands associated to the protein target in the DUD-E database. The average $EF^{1\%}$ ($aveEF^{1\%}$) is calculated over all the active ligands for each of the protein targets. The numbers in each panel means the number of points in the upper and lower triangles, respectively. PCC is the Pearson's correlation coefficient between the $EF^{1\%}$ of the two compared methods. 'p-value' is calculated in the student *t*-test.

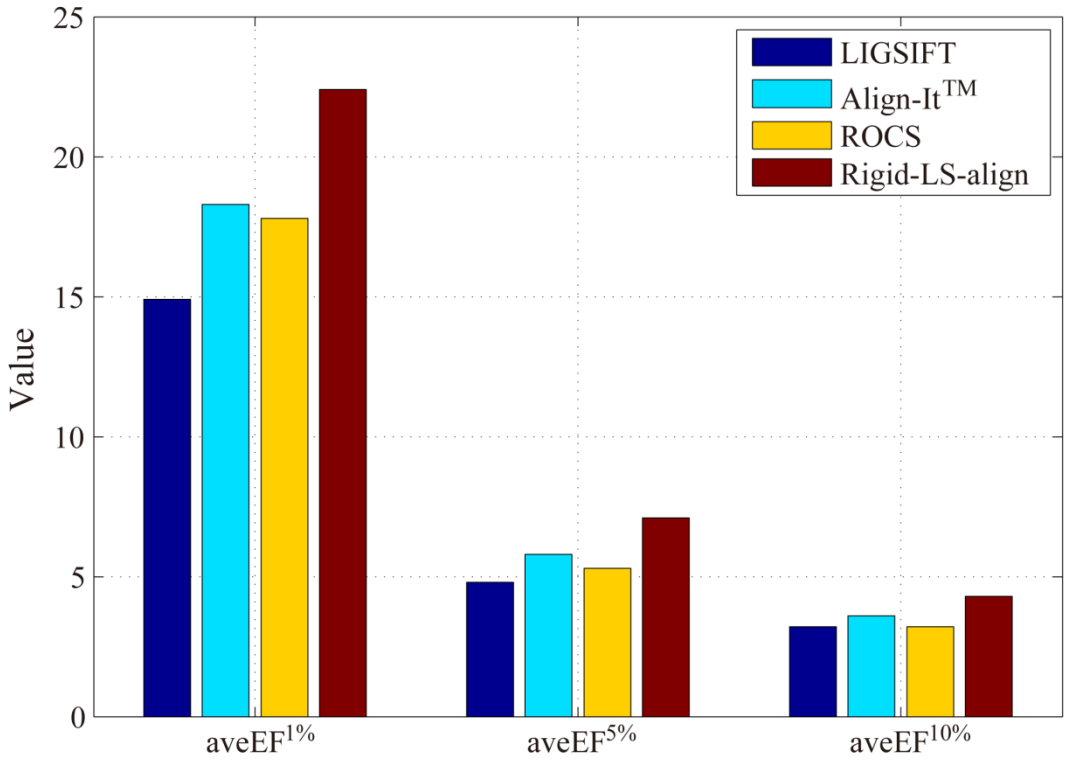

Fig. S9. The average aveEF$^{1\%}$, aveEF$^{5\%}$, and aveEF$^{10\%}$ of LIGSIFT, Align-It^{TM}, ROCS, and Rigid-LS-align over all 102 proteins, using the single rigid conformers of DUD-E database molecules under the new experimental scheme in Fig. S8.

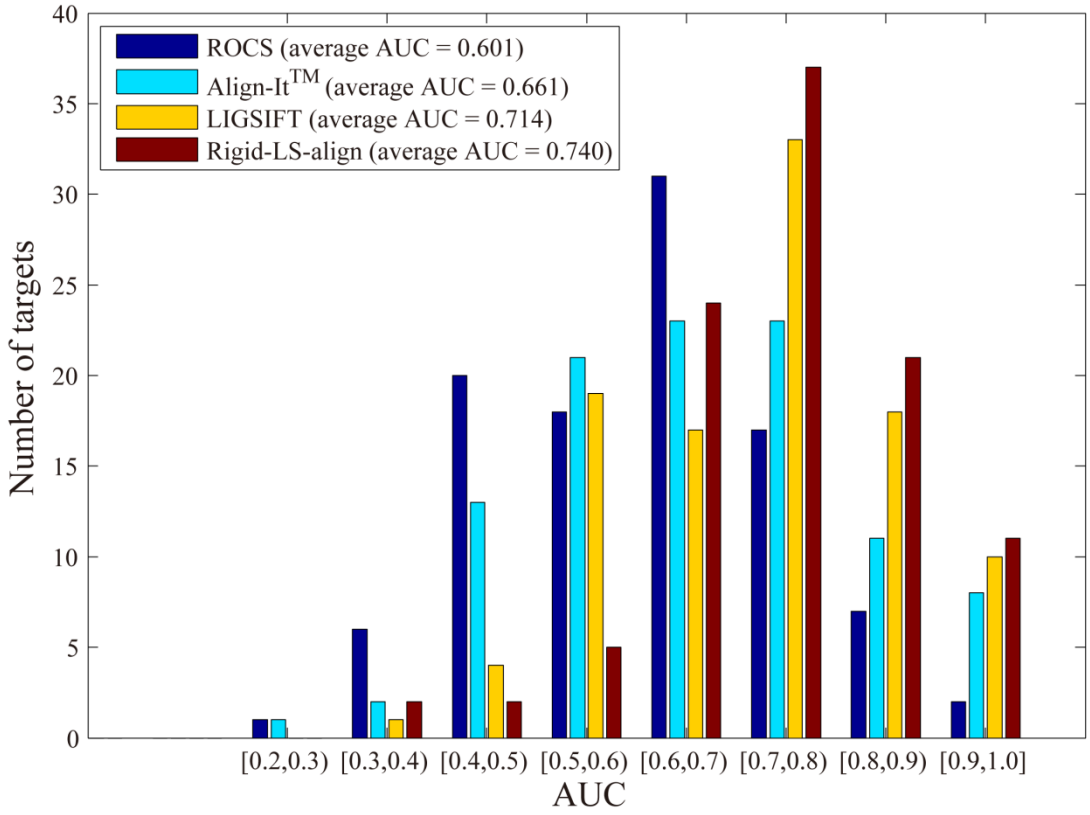

Fig. S10. AUC distributions of ROCS, Align-It™, LIGSIFT, and Rigid-LS-align, respectively, using 50 OMEGA-generated conformers of database molecules on 102 DUD-E targets

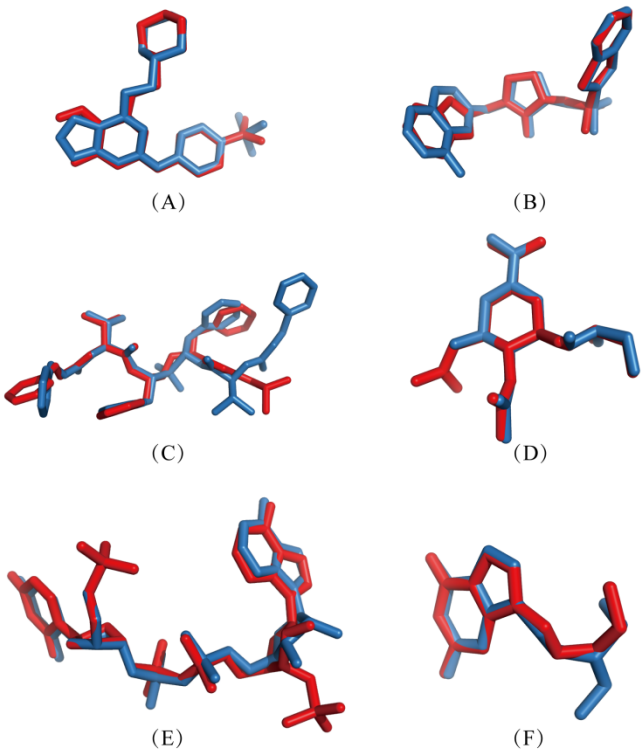

Fig. S11. Examples of scaffold hopping identified by LS-align for different ligand molecules. (A) CDK2; (B) FXa ligands; (C) HIV-1p; (D) NA; (E) RNAse; (F) TK. All superimpositions are generated by Rigid-LS-align. The six examples are selected from the validation set of Sperandio and coworkers (Sperandio, O. et al, J Chem Info Mod, 2007; 47: 1097), which is composed of 52 molecules active against six different proteins: cyclin-dependent kinase 2 (CDK2), coagulation factor Xa (FXa), HIV-1 protease (HIV-1p), neuraminidase (NA), ribonuclease A (RNase), and thymidine kinase (TK).

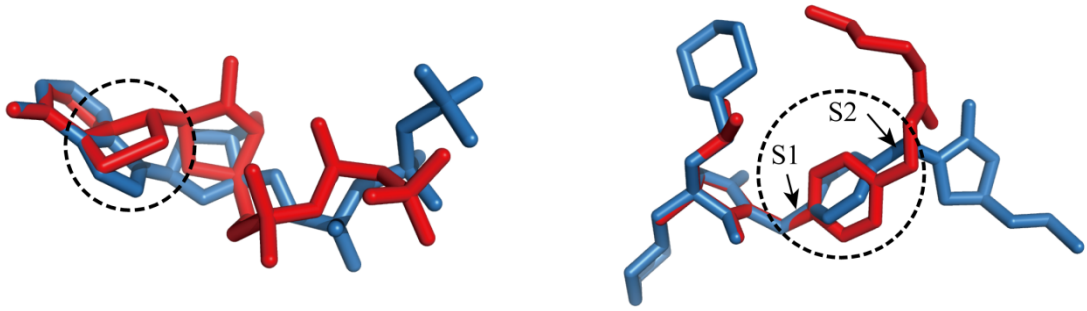

Fig. S12. Two examples highlighting the issues of the current Flexi-LS-align program. (A) Alignments of two ATP ligand structures from PDBID 4TL7 (blue) and 1GOL (red), in which Flexi-LS-align could achieve a better alignment in the highlighted region when the ring-level flexibility was considered. (B) The alignment between the seed ligand (red) of the DUD-E protein target BACE1 and the active ligand (CHEML583875, blue), in which the optimal match was not generated in the highlighted region due to the big-size of the angle changes in flexibility search. S1 and S2 are two rotatable single bonds.

Supporting Tables

Table S1. Specific weights for Single, Double, Triple, Amide, Aromatic and Dummy bond types to calculate the *BWJS* value.

Bond type	Specific weight
Single	1.0
Double	1.2
Triple	1.4
Amide	1.8
Aromatic	2.0
Dummy	1.6

Table S2. Parameters obtained from the fitting of type-I extreme value distribution of Eqs. (9) and (10) with the observed probability density of the alignment scores of random ligand pairs collected from PDB library in Figs. S3-S4.

Scoring function	*a*	*b*	*c*	*d*	*e*	*f*
LS-score	-0.449436	0.136696	0.130899	0.057542	-0.002017	-0.001779
PC-score	-0.563358	0.152792	0.147930	0.110170	-0.007444	-0.008468

Table S3. List of DUD-E protein targets used for analyzing virtual screening performance. Num_{act} and Num_{dec} are the numbers of active and decoy ligands respectively.

DUD-E target	(Num_{act}, Num_{dec})	DUD-E target	(Num_{act}, Num_{dec})	DUD-E target	(Num_{act}, Num_{dec})
AA2AR	(482, 31498)	FABP4	(47, 2749)	MMP13	(572, 37126)
ABL1	(182, 10746)	FAK1	(100, 5350)	MP2K1	(121, 8147)
ACE	(282, 16860)	FGFR1	(139, 333)	NOS1	(100, 8050)
ACES	(453, 26233)	FKB1A	(111, 5800)	NRAM	(98, 6199)
ADA	(93, 5449)	FNTA	(592, 51430)	PA2GA	(99, 5146)
ADA17	(532, 35809)	FPPS	(85, 8822)	PARP1	(508, 30035)
ADRB1	(247, 15842)	GCR	(258, 14987)	PDE5A	(398, 27521)
ADRB2	(231, 14993)	GLCM	(54, 3799)	PGH1	(195, 10797)
AKT1	(293, 16426)	GRIA2	(158, 11832)	PGH2	(435, 23135)
AKT2	(117, 6893)	GRIK1	(101, 6547)	PLK1	(107, 6797)
ALDR	(159, 8995)	HDAC2	(185, 10299)	PNPH	(103, 6950)
AMPC	(48, 2832)	HDAC8	(170, 10448)	PPARA	(373, 19356)
ANDR	(269, 14343)	HIVINT	(100, 6644)	PPARD	(240, 12223)
AOFB	(122, 6900)	HIVPR	(536, 35688)	PPARG	(484, 25256)
BACE1	(283, 18080)	HIVRT	(338, 18879)	PRGR	(293, 15642)
BRAF	(152, 9942)	HMDH	(170, 8743)	PTN1	(130, 7243)
CAH2	(492, 31133)	HS90A	(88, 4848)	PUR2	(50, 2694)
CASP3	(199, 10692)	HXK4	(92, 4696)	PYGM	(77, 3940)
CDK2	(474, 27830)	IGF1R	(148, 9291)	PYRD	(111, 6446)
COMT	(41, 3848)	INHA	(43, 2300)	RENI	(104, 6955)
CP2C9	(120, 7446)	ITAL	(138, 8487)	ROCK1	(100, 6297)
CP3A4	(170, 11796)	JAK2	(107, 6495)	RXRA	(131, 6935)
CSF1R	(166, 12144)	KIF11	(116, 6848)	SAHH	(63, 3450)
CXCR4	(40, 3406)	KIT	(166, 10447)	SRC	(524, 34454)
DEF	(102, 5696)	KITH	(57, 2850)	TGFR1	(133, 8498)
DHI1	(330, 19340)	KPCB	(135, 8692)	THB	(103, 7441)
DPP4	(533, 40916)	LCK	(420, 27374)	THRB	(461, 26948)
DRD3	(480, 34022)	LKHA4	(171, 9448)	TRY1	(449, 25914)
DYR	(231, 17170)	MAPK2	(101, 6147)	TRYB1	(148, 7643)
EGFR	(542, 35020)	MCR	(94, 5146)	TYSY	(109, 6738)
ESR1	(383, 20663)	MET	(166, 11240)	UROK	(162, 9841)
ESR2	(367, 20182)	MK01	(79, 4548)	VGFR2	(409, 24927)
FA10	(537, 20023)	MK10	(104, 6599)	WEE1	(102, 6148)
FA7	(114, 6245)	MK14	(578, 35810)	XIAP	(100, 5145)

Table S4. Comparison of virtual screening results by LS-score, Mass-score, BWJS-score, and PC-score on DUD-E database.

Score function	$EF^{1\%}$	$EF^{5\%}$	$EF^{10\%}$	$HR^{1\%}$	$HR^{5\%}$	$HR^{10\%}$	AUC
LS-score	12.0	4.7	3.3	19.8	23.7	33.2	0.69
Mass-score	6.3	3.1	2.3	10.4	15.8	23.3	0.60
BWJS-score	11.6	4.6	3.1	19.9	23.4	31.3	0.68
PC-score	20.1	6.9	4.3	33.0	34.8	43.5	0.74

Table S5. P-values in student's t-test for the difference between Rigid-LS-align and the three control methods in EF, HR, and AUC values using single conformers of database molecules on DUD-E database. 'Shape sTC', 'chem sTC', and 'combo sTC' are different modules used in LIGSIFT.

Method	in $EF^{1\%}$	in $EF^{5\%}$	in $EF^{10\%}$	in $HR^{1\%}$	in $HR^{5\%}$	in $HR^{10\%}$	in AUC
LIGSIFT (shape sTC)	4.9×10^{-15}	1.1×10^{-16}	5.7×10^{-17}	5.4×10^{-15}	6.7×10^{-17}	3.8×10^{-17}	3.0×10^{-14}
LIGSIFT (chem sTC)	1.8×10^{-8}	1.6×10^{-8}	8.0×10^{-8}	1.3×10^{-8}	9.5×10^{-9}	6.3×10^{-8}	3.7×10^{-4}
LIGSIFT (combo sTC)	1.4×10^{-10}	3.8×10^{-11}	1.4×10^{-10}	1.3×10^{-9}	2.6×10^{-11}	1.0×10^{-10}	9.8×10^{-7}
Align-It™	1.1×10^{-7}	5.2×10^{-6}	5.3×10^{-5}	1.9×10^{-6}	2.6×10^{-6}	6.5×10^{-4}	3.1×10^{-4}
ROCS	3.2×10^{-8}	1.3×10^{-11}	6.0×10^{-14}	5.0×10^{-8}	7.4×10^{-12}	3.9×10^{-14}	3.2×10^{-17}

Table S6. EF of Flexi-LS-align and the control methods (LIGSIFT, Align-It^{TM}, and ROCS) in top 1%, and 5%, using Flexi-LS-align-generated multiple conformers of DUD-E database molecules.

DUD-E	Flexi-LS-align		LIGSIFT (chem sTC)		Align-It^{TM}		ROCS	
targets	$EF^{1\%}$	$EF^{5\%}$	$EF^{1\%}$	$EF^{5\%}$	$EF^{1\%}$	$EF^{5\%}$	$EF^{1\%}$	$EF^{5\%}$
AA2AR	28.91	8.34	11.44	4.65	0.00	0.41	13.31	4.07
ABL1	12.67	4.40	6.61	2.86	9.92	5.72	7.16	3.41
ACE	22.04	7.52	10.31	5.18	9.95	6.10	12.44	4.75
ACES	31.23	7.55	3.54	1.90	7.53	2.83	3.1	1.28
ADA	39.01	9.47	16.25	6.02	14.09	6.88	15.17	5.59
ADA17	18.82	8.38	10.54	7.07	3.76	3.20	7.72	3.87
ADRB1	15.06	5.43	10.99	3.40	7.33	2.75	7.33	2.92
ADRB2	16.48	6.76	3.90	2.34	5.64	2.25	6.94	2.34
AKT1	5.81	2.32	2.05	1.16	8.20	3.49	2.73	0.89
AKT2	10.27	3.08	6.85	2.23	11.13	3.25	10.27	3.25
ALDR	26.57	7.18	19.61	5.17	15.82	5.04	20.88	5.42
AMPC	27.86	6.67	6.43	7.08	30.00	8.33	8.57	6.67
ANDR	25.67	6.62	17.11	5.21	11.16	4.09	16.37	5.58
AOFB	1.64	2.30	4.11	1.80	1.64	0.98	1.64	0.98
BACE1	2.13	1.41	1.77	1.91	2.84	2.12	1.42	0.28
BRAF	15.94	6.32	13.95	6.19	21.25	6.06	21.91	5.53
CAH2	7.73	6.87	8.34	3.78	4.68	4.88	4.07	2.07
CASP3	18.24	8.05	1.52	0.91	4.05	2.21	1.01	1.01
CDK2	2.74	3.71	4.64	2.95	3.17	2.32	4.64	3.17
COMT	34.95	16.13	29.95	9.78	42.43	12.22	44.93	9.29
CP2C9	10.09	4.17	2.52	2.00	1.68	0.33	1.68	1.67
CP3A4	11.24	3.41	1.18	0.71	2.96	1.29	0.59	0.71
CSF1R	25.32	7.72	21.70	5.79	23.51	5.91	24.12	5.91
CXCR4	27.87	7.01	7.60	3.51	2.53	3.51	12.67	5.01
DEF	37.90	9.83	29.92	9.24	30.91	9.83	22.94	7.47
DHI1	15.81	5.03	4.26	3.15	1.82	1.33	8.82	4.12
DPP4	9.58	4.05	8.64	4.24	4.13	3.12	6.95	3.15
DRD3	3.54	1.08	2.08	1.04	0.63	1.13	1.25	0.83
DYR	14.72	6.41	15.15	5.37	14.29	5.11	29.44	7.62
EGFR	23.10	8.08	12.20	4.61	21.99	9.26	12.75	4.98
ESR1	43.70	11.18	23.29	7.89	24.07	9.66	28.78	8.83
ESR2	19.94	8.45	22.40	7.20	13.93	6.71	26.49	8.34
FA10	8.78	4.32	13.82	5.40	6.91	2.57	11.02	4.8
FA7	28.33	11.09	7.08	4.05	2.66	2.46	7.08	3.34
FABP4	35.25	6.85	8.81	4.71	26.44	6.85	13.22	4.71
FAK1	51.47	15.03	18.17	8.42	20.19	7.21	20.19	6.61
FGFR1	0.85	1.33	0.00	0.15	0.00	0.44	0.85	0.3

FKB1A	23.47	8.12	2.71	1.26	1.81	1.81	0	0.54
FNTA	16.22	7.33	3.38	1.93	0.51	0.74	6.76	3.38
FPPS	76.53	19.31	73.00	18.13	91.84	20.02	65.93	18.37
GCR	31.10	9.07	5.83	2.40	4.66	2.17	5.44	1.63
GLCM	0.00	0.74	1.88	1.49	7.51	4.46	11.27	3.72
GRIA2	22.32	6.97	26.78	7.60	24.87	8.99	30.61	7.85
GRIK1	3.99	2.78	18.95	6.54	8.98	4.56	4.99	6.15
HDAC2	12.53	4.11	6.54	3.68	8.17	2.16	3.81	1.84
HDAC8	2.36	1.77	16.50	5.54	12.37	5.07	11.2	3.77
HIVINT	0.00	3.00	0.00	1.20	4.03	2.00	0	1
HIVPR	31.58	10.12	6.35	4.37	1.49	2.76	3.17	3.1
HIVRT	9.18	3.85	5.33	2.07	6.22	3.02	5.92	2.19
HMDH	47.13	11.43	15.91	4.59	20.03	9.66	14.73	4.12
HS90A	17.17	9.35	4.58	2.28	10.30	6.38	9.16	3.65
HXK4	25.47	7.62	32.11	12.63	17.72	6.75	36.54	8.27
IGF1R	16.28	6.63	13.57	4.87	14.93	4.20	14.93	3.39
INHA	28.43	5.59	0.00	0.93	2.37	0.93	0	0.93
ITAL	15.26	3.34	13.08	3.19	2.91	3.05	11.63	2.61
JAK2	25.24	7.85	15.89	4.67	19.63	5.80	20.57	6.17
KIF11	40.02	9.49	15.66	6.38	7.83	4.14	20.01	6.73
KIT	3.62	2.05	2.41	1.57	1.21	0.60	3.02	1.21
KITH	51.00	16.18	17.59	6.33	24.62	12.31	14.07	4.57
KPCB	52.01	12.31	42.35	11.12	23.03	9.79	49.04	12.01
LCK	22.93	7.43	1.91	1.29	7.64	3.57	2.87	1.91
LKHA4	22.85	7.38	10.55	5.63	7.62	3.98	15.23	5.27
MAPK2	20.95	5.55	21.95	9.91	13.97	5.95	23.95	7.34
MCR	17.15	6.60	7.50	3.40	8.58	2.55	10.72	3.83
MET	50.63	12.90	28.93	7.84	53.64	13.86	25.92	7.47
MK01	31.83	10.40	21.65	6.59	31.83	6.85	21.65	6.59
MK10	2.89	0.77	3.85	2.50	1.92	1.54	3.85	1.54
MK14	6.07	3.74	2.77	1.49	5.72	2.53	5.03	2.73
MMP13	37.33	13.50	18.58	6.93	15.42	6.58	15.07	4.9
MP2K1	16.67	3.47	11.67	3.81	1.67	1.65	7.5	3.31
NOS1	3.02	1.20	5.03	2.80	2.01	2.60	2.01	1.2
NRAM	1.04	1.43	22.80	7.78	2.07	1.64	17.62	7.57
PA2GA	2.04	2.43	3.06	2.22	2.04	2.22	6.11	2.63
PARP1	16.95	6.69	17.54	5.63	17.94	5.24	17.35	5.47
PDE5A	18.10	5.03	8.05	2.72	19.86	6.03	3.02	1.21
PGH1	4.65	2.57	2.59	1.75	1.55	0.51	5.69	2.77
PGH2	32.51	10.76	17.98	6.16	13.37	5.20	26.98	9.11
PLK1	12.16	4.11	0.00	1.50	3.74	3.18	2.81	1.12
PNPH	68.48	19.84	41.09	10.89	55.76	16.92	48.91	11.87
PPARA	19.87	8.42	11.28	5.20	11.28	5.15	11.55	5.2

PPARD	12.14	4.83	5.44	3.17	4.61	1.92	3.77	2.33
PPARG	27.32	8.22	9.52	3.76	12.00	5.41	6.41	2.89
PRGR	6.50	2.94	1.37	1.91	0.34	0.96	5.47	3.35
PTN1	10.10	4.62	9.32	2.77	8.55	3.54	2.33	0.92
PUR2	54.88	18.83	36.59	13.22	38.62	17.63	20.33	12.02
PYGM	0.00	2.09	0.00	0.00	0.00	0.00	0	0
PYRD	48.17	10.48	43.62	10.30	46.35	9.57	47.26	11.02
RENI	24.24	7.71	7.76	3.28	0.97	2.31	13.58	4.05
ROCK1	10.15	5.01	1.02	1.40	1.02	0.40	1.02	1
RXRA	13.87	8.25	6.16	2.90	0.00	0.15	13.87	8.86
SAHH	55.76	18.80	47.80	15.29	55.76	19.76	55.76	19.44
SRC	9.56	4.01	4.21	2.71	8.80	2.79	2.3	1.22
TGFR1	29.43	12.05	10.56	3.16	14.34	5.72	15.85	4.82
THB	41.99	11.66	46.88	13.99	27.34	7.77	41.02	12.63
THRB	0.43	2.52	1.08	0.74	0.87	1.17	1.3	0.91
TRY1	15.63	7.22	2.46	1.96	1.79	1.34	2.46	1.51
TRYB1	9.57	3.38	2.73	1.22	0.68	1.22	0	0.68
TYSY	42.49	11.20	14.78	7.35	31.41	11.02	9.24	4.41
UROK	16.05	6.05	4.32	1.73	9.88	4.45	7.41	2.1
VGFR2	18.12	6.65	2.94	2.15	5.63	2.30	2.94	1.52
WEE1	61.27	18.85	61.27	17.28	61.27	18.07	61.27	14.34
XIAP	47.41	15.41	21.18	5.20	7.06	2.80	27.23	7.81
Average	22.01	7.23	13.03	4.76	13.19	4.91	13.7	4.6
Standard deviation	16.68	4.46	13.47	3.66	15.59	4.31	14.09	3.72

Table S7. EF of Rigid-LS-align and compared methods (LIGSIFT, Align-It^{TM}, and ROCS) in top 1%, and 5% using 50 OMEGA-generated conformers of DUD-E database molecules

DUD-E target	Rigid-LS-align (PC)		LIGSIFT (chem sTC)		Align-It^{TM}		ROCS	
	$EF^{1\%}$	$EF^{5\%}$	$EF^{1\%}$	$EF^{5\%}$	$EF^{1\%}$	$EF^{5\%}$	$EF^{1\%}$	$EF^{5\%}$
AA2AR	29.33	7.68	19.14	6.60	0.42	0.58	11.73	4.68
ABL1	17.08	3.96	20.38	5.28	11.02	5.17	16.25	5.81
ACE	20.62	7.38	14.93	6.24	12.09	5.89	22.94	7.47
ACES	31.00	7.73	3.10	1.99	8.64	3.14	29.44	7.79
ADA	39.01	9.68	44.42	13.55	24.92	11.40	7.97	3.34
ADA17	25.78	9.93	15.81	8.27	11.29	3.87	6.61	1.78
ADRB1	13.84	5.51	15.88	4.78	9.36	3.08	3.02	1.21
ADRB2	15.18	5.89	6.94	2.51	6.94	2.25	2.39	1.86
AKT1	6.49	3.76	4.44	1.30	6.49	3.69	11.79	3.83
AKT2	10.27	2.74	11.98	2.91	11.13	3.77	24.71	7.35
ALDR	29.74	8.31	26.57	7.43	18.98	6.05	2.49	1.22
AMPC	27.86	7.08	27.86	7.50	36.43	8.33	39.06	12.63
ANDR	24.56	6.25	18.97	5.58	14.14	4.76	7.71	3.52
AOFB	1.64	1.64	2.47	1.64	1.64	0.98	2.88	1.28
BACE1	3.55	2.47	3.55	2.54	4.96	2.05	2.73	0.96
BRAF	15.27	5.01	26.56	7.38	24.57	5.67	10.27	3.25
CAH2	6.51	4.72	10.78	5.24	2.24	2.72	20.88	5.54
CASP3	16.72	8.55	1.52	2.21	4.05	1.41	8.57	6.67
CDK2	6.33	3.80	13.08	4.94	10.76	4.56	16.37	5.88
COMT	34.95	12.22	27.46	9.78	42.43	9.78	1.64	1.15
CP2C9	5.88	2.67	6.73	2.50	3.36	1.67	21.91	5.53
CP3A4	7.69	3.53	1.77	1.41	4.14	1.53	4.48	2.2
CSF1R	25.32	6.63	28.94	8.68	25.32	5.55	4.85	3.25
CXCR4	32.94	8.51	32.94	8.51	25.34	6.01	44.93	9.29
DEF	47.87	10.23	45.87	12.78	35.90	14.16	9.12	3.88
DHI1	19.46	5.58	8.21	3.76	1.52	1.82	7.33	3.19
DPP4	9.96	3.60	20.29	7.32	5.82	3.98	1.25	0.75
DRD3	2.50	1.08	1.67	0.88	1.04	1.42	13.31	5.06
DYR	9.52	5.45	16.88	6.06	16.02	5.54	28.78	8.78
EGFR	23.10	7.42	17.93	6.86	23.66	9.34	26.22	8.56
ESR1	42.39	11.13	35.85	10.45	25.12	10.71	10.83	4.77
ESR2	19.12	8.72	28.13	9.21	18.57	8.12	21.19	7.21
FA10	11.58	4.77	13.63	6.48	7.66	3.31	8.28	3.45
FA7	25.68	10.21	17.71	4.58	14.17	3.87	63.58	18.37
FABP4	35.25	6.85	35.25	8.13	35.25	6.85	11.27	3.34
FAK1	50.46	15.03	52.48	16.63	23.21	5.61	17.67	4.71

FGFR1	1.70	1.18	0.00	0.44	0.00	0.44	36.54	7.84
FKB1A	23.47	7.58	4.51	2.35	9.93	4.69	0	0.93
FNTA	16.90	7.16	4.90	2.84	2.03	2.30	12.35	2.76
FPPS	56.52	16.95	88.30	18.60	77.71	20.02	22.44	6.92
GCR	29.16	9.23	20.99	4.96	1.17	1.94	14.07	4.57
GLCM	0.00	0.00	3.76	1.86	9.39	2.23	47.55	12.16
GRIA2	22.32	7.22	40.81	9.50	19.77	7.47	21.65	6.85
GRIK1	2.99	3.37	22.94	7.53	6.98	4.56	2.89	1.54
HDAC2	9.81	4.33	9.26	4.22	6.54	2.60	5.2	2.77
HDAC8	2.36	1.41	17.68	7.54	18.86	7.19	3.02	1.4
HIVINT	0.00	2.80	0.00	1.80	3.02	2.40	19.69	7.78
HIVPR	31.18	10.15	8.03	5.41	6.91	4.22	6.21	3.08
HIVRT	8.00	3.02	9.18	2.96	7.40	3.32	26.52	9.06
HMDH	49.48	12.14	41.24	12.96	37.70	12.72	2.81	1.31
HS90A	17.17	10.26	11.45	7.07	9.16	3.88	46.95	11.48
HXK4	27.68	10.23	40.97	16.11	27.68	11.76	4.79	3.42
IGF1R	16.28	6.36	22.39	6.91	14.93	3.93	2.33	0.92
INHA	26.06	5.59	9.48	5.59	11.85	3.73	22.36	12.42
ITAL	14.53	3.48	17.44	3.77	7.99	2.61	0	0
JAK2	29.92	8.41	27.11	8.41	19.63	6.36	47.26	11.02
KIF11	36.54	9.49	34.80	11.04	7.83	3.11	13.58	4.05
KIT	3.02	1.69	3.02	1.45	1.21	0.24	16.95	8.86
KITH	51.00	17.59	35.17	9.14	51.00	15.48	55.76	19.44
KPCB	46.81	12.60	54.24	13.05	40.87	10.23	1.3	0.87
LCK	24.13	7.53	6.21	2.72	9.79	3.53	2.68	1.6
LKHA4	25.78	8.09	28.13	13.24	12.89	5.98	12.93	4.78
MAPK2	24.94	6.94	36.92	12.29	18.96	5.55	7.41	2.1
MCR	12.86	6.17	9.65	4.47	4.29	2.98	61.27	14.34
MET	53.64	13.38	44.00	11.09	57.86	14.83	27.23	7.61
MK01	31.83	11.16	30.56	7.35	33.10	7.10	12.9	4.02
MK10	1.92	0.77	7.70	2.12	5.77	1.15	7.34	3.83
MK14	6.24	3.81	6.59	3.15	3.99	2.15	6.92	2.92
MMP13	37.16	12.98	29.80	9.93	20.86	7.66	6.94	2.34
MP2K1	11.67	3.47	12.50	4.80	4.17	2.81	1.77	0.35
NOS1	4.02	1.00	8.05	4.00	0.00	1.60	1.01	1.01
NRAM	1.04	1.23	35.24	13.51	4.15	2.25	1.68	1.5
PA2GA	6.11	4.25	4.08	3.44	8.15	7.48	0.59	0.71
PARP1	16.36	6.58	24.25	7.52	21.29	5.16	23.51	5.79
PDE5A	17.85	4.88	21.12	5.08	25.39	6.99	12.67	5.51
PGH1	6.21	2.46	3.62	2.36	1.03	0.62	17.63	5.14
PGH2	34.82	11.18	26.05	9.34	19.37	6.72	0.85	0.74
PLK1	7.48	4.11	4.68	2.62	1.87	1.12	0.9	1.08
PNPH	68.48	19.45	56.74	15.95	57.72	15.56	30.61	7.85

PPARA	19.33	9.66	15.57	6.87	14.50	6.54	2.99	5.55
PPARD	11.73	5.75	7.12	3.58	3.77	1.67	3.81	1.84
PPARG	31.45	8.43	22.35	8.14	21.31	6.65	11.2	3.89
PRGR	5.47	2.73	2.39	1.50	0.34	1.02	3.17	3.02
PTN1	13.21	4.62	6.99	3.70	3.88	2.47	5.33	2.43
PUR2	54.88	18.83	50.81	14.02	46.75	13.62	9.16	3.65
PYGM	5.22	5.48	0.00	1.04	0.00	0.00	14.93	3.39
PYRD	48.17	11.02	47.26	12.28	48.17	10.30	18.27	6.38
RENI	28.12	6.94	22.30	5.40	8.73	3.66	14.65	5.27
ROCK1	5.08	3.61	1.02	1.00	1.02	0.40	25.94	7.34
RXRA	16.18	8.10	18.49	7.03	0.00	0.31	15.78	5.14
SAHH	55.76	18.16	55.76	19.76	55.76	19.76	7.5	3.31
SRC	8.22	3.74	10.14	4.24	10.90	3.13	7.13	3.24
TGFR1	25.66	12.20	16.60	4.97	23.39	7.53	18.53	5.79
THB	40.04	13.79	64.45	16.32	27.34	9.13	3.02	1.06
THRB	0.22	1.74	1.30	1.13	0.65	1.30	11.81	5.31
TRY1	12.95	6.99	4.02	2.32	0.67	1.16	3.77	2.25
TRYB1	8.20	3.65	4.79	2.03	4.10	1.08	6.41	2.98
TYSY	42.49	10.84	22.17	9.37	29.56	9.73	1.02	1
UROK	20.38	6.42	9.26	3.58	13.58	5.19	15.85	4.82
VGFR2	18.61	6.80	6.12	3.03	4.16	2.35	0	0.68
WEE1	61.27	17.87	61.27	19.64	61.27	18.07	2.94	1.57
XIAP	45.39	14.61	21.18	11.61	12.10	6.61	0	1
Average	21.9	7.3	20.3	6.7	16.0	5.4	13.9	4.7
Standard deviation	16.3	4.4	17.2	4.6	16.1	4.4	13.9	3.7

Table S8. EF values by TC-FP, Rigid-LS-align, LIGSIFT, Align-It^{TM} and ROCS, as well as the composite methods of TC-FP+LIGSIFT, TC-FP+Align-It^{TM}, TC-FP+ROCS, and TC-FP+Rigid-LS-align, using 50 OMEGA-generated conformers of database molecules on 102 DUD-E targets. In the composite methods, the EF values correspond to the best combination with *w* being the weight of TC-FP, i.e., $score = w \times S_{TC-FP} + (1-w) \times S_{3D-method}$. Bold fonts highlight the highest value in each category.

Method (best weight *w* on TC-FP)	**EF$^{1\%}$**	**EF$^{5\%}$**	**EF$^{10\%}$**
Individual Methods			
TC-FP	20.5	7.0	**4.6**
LIGSIFT (chem sTC)	20.3	6.7	4.1
Align-ItTM	16.0	5.4	3.4
ROCS	13.9	4.7	2.9
Rigid-LS-align	**21.9**	**7.3**	4.5
Combined methods			
TC-FP+LIGSIFT (chem sTC) (0.3)	25.0	8.3	5.1
TC-FP+Align-ItTM (0.6)	23.1	7.8	4.8
TC-FP+ROCS (0.8)	21.9	7.6	4.8
TC-FP+Rigid-LS-align (0.5)	**26.0**	**8.6**	**5.2**