Supporting Information

Absolute binding free energies between T4 lysozyme and 141 small molecules: calculations based on multiple rigid receptor structures

Bing Xie, Trung Hai Nguyen, and David D. L. Minh*

Department of Chemistry, Illinois Institute of Technology, Chicago, IL 60616, USA

E-mail: dminh@iit.edu

*To whom correspondence should be addressed

Table S1: Names and SMILES strings[1] of all 141 small molecules

70 active compounds	SMILES
1-methylpyrrole	CN1C=CC=C1
iodobenzene	Ic1ccccc1
p-methylphenyl azide	CC1=CC=C(C=C1)N=[N+]=[N-]
phenyl azide	[N-]=[N+]=NC1=CC=CC=C1
allyl ethyl sulfoxide	CCS(=O)(=O)CC=C
fluorobenzene	Fc1ccccc1
ethylbenzene	CCc1ccccc1
bromoethane	CCBr
benzyl mercaptan	SCc1ccccc1
2-phenylpropanol	CC(CO)C1=CC=CC=C1
anisole	COc1ccccc1
iodoethane	CCI
diisopropyl disulfide	CC(C)SSC(C)C
diethyl disulfide	CCSSCC
cyclohexene	C1CCC=CC1
4-(methylthio)nitrobenzene	1=CC=C(C=C1)[N+](=O)[O-]
2,6-difluorobenzyl bromide	C1=CC(=C(C(=C1)F)CBr)F
thiophenol	Sc1ccccc1
β-chlorophenetole	C1=CC=C(C=C1)OCCCl
cis-3-hexenyl formate	CC/C=C\CCOC=O
6-methyl-1,5-heptadiene	CC(=CCCC=C)C
2-ethoxyphenol	CCOC1=CC=CC=C1O
3-methyl benzyl azide	CC1=CC(=CC=C1)CN=[N+]=[N-]
n-methylaniline	CNC1=CC=CC=C1
1,2-benzenedithiol	C1=CC=C(C(=C1)S)S
1,2-dichlorobenzene	C1=CC=C(C(=C1)Cl)Cl
2-(methylthio)-thiophene	CSc1cccs1
thiophene-2-thiol	Sc1cccs1
2-(thienylthio)-acetone	CC(=O)CSc1cccs1
methyl ethyl cyclopentene	CCC1=C(C)CCC1
allyl ethyl sulfide	CCSCC=C
allyl sulfide	C=CCSCC=C
allyl n-propylether	CCCOCC=C
3-phenyl-1-propanethiol	SCCCc1ccccc1
3-iodotoluene	Cc1cccc(c1)I
3-ethyltoluene	CCc1cccc(C)c1
2-methylbenzofuran	Cc1cc2ccccc2o1
4-iodotoluene	Cc1ccc(cc1)I
4-ethyltoluene	CCc1ccc(C)cc1
3-phenyl-1-propanol	OCCCc1ccccc1
benzofuran	c2cc1ccccc1o2
amylbenzene	CCCCCc1ccccc1
1-phenyl-1-propyne	CC#Cc1ccccc1
1,2-benzisoxazole	c1noc2ccccc12
2-iodotoluene	Cc1ccccc1I
2-iodopropane	CC(C)I

2-ethyltoluene	CCc1ccccc1C
2-ethylfuran	CCc1ccco1
benzene	c1ccccc1
indene	C1C=Cc2ccccc21
indole	c2cc1ccccc1[nH]2
hexafluorobenzene	Fc1c(c(c(c(c1F)F)F)F)F
methylcyclohexane	CC1CCCCC1
m-xylene	Cc1cccc(c1)C
isobutylbenzene	CC(C)Cc1ccccc1
n-hexylbenzene	CCCCCCc1ccccc1
nitrobenzene	[O]N(=O)c1ccccc1
naphthalene	c2ccc1ccccc1c2
n-butylbenzene	CCCCc1ccccc1
phenylacetylene	C#Cc1ccccc1
dipropyl disulfide	CCCSSCCC
n-propylbenzene	CCCc1ccccc1
o-xylene	Cc1ccccc1C
styrene	C=Cc1ccccc1
thianaphthene	c2cc1ccccc1s2
sec-butyl-benzene	CCC(C)c1ccccc1
thioanisole	CSc1ccccc1
toluene	Cc1ccccc1
p-xylene	Cc1ccc(cc1)C
iodopentafluorobenzene	C1(=C(C(=C(C(=C1F)F)I)F)F)F

71 inactive compounds	SMILES
2-iodoethanol	OCCI
2-bromoethanol	OCCBr
benzyl-alcohol	OCc1ccccc1
3-chlorophenol	Oc1cccc(Cl)c1
azulene	c1cc2cccccc2c1
ethanol	CCO
methanol	CO
dimethyl-sulfoxide	CS(=O)C
mesitylene	Cc1cc(C)cc(C)c1
2,5-dimethylfuran	Cc1ccc(C)o1
p-cresol	Cc1ccc(O)cc1
chloroform	ClC(Cl)Cl
pentafluorophenyl-azide	Fc1c(F)c(F)c(N=[N+]=[N-])c(F)c1F
6,7-ifluorobenzotriazole	Fc1ccc2nn[nH]c2c1F
pentafluoroaniline	Nc1c(F)c(F)c(F)c(F)c1F
1,4-diiodobenzene	Ic1ccc(I)cc1
1,2-diiodobenzene	Ic1ccccc1I
trans-cinnamaldehyde	O=CC=Cc1ccccc1
cyclohexane	C1CCCCC1
tert-butylbenzene	CC(C)(C)c1ccccc1
isobutane	CC(C)C
(-)-camphene	CC1(C)C2CCC(C2)C1=C
DL-camphor	CC1(C)C2CCC1(C)C(=O)C2

1-octanol	CCCCCCCCO
1-heptanol	CCCCCCCO
N-heptylbenzene	CCCCCCCc1ccccc1
dibutyl-disulfide	CCCCSSCCCC
1-propanol	CCCO
1,1-diethylurea	CCN(CC)C(=O)N
benzaldehyde	O=Cc1ccccc1
2,6-difluorophenyl azide	FC1=CC=CC(F)=C1N=[N+]=[N-]
benzoic-acid	OC(=O)c1ccccc1
n,n'-dimethylaniline	CN(C)C1=CC=CC=C1
benzaldehyde-oxime	ON=Cc1ccccc1
quinoline	c1ccc2ncccc2c1
furan	c1ccoc1
pyridine	c1ccncc1
2,1-benzisoxazole	c1onc2ccccc12
2,5-diaminophenol	c1(c(N)ccc(c1)N)O
2-methylbenzyl alcohol	c1(CO)c(C)cccc1
4-hydrazinothieno[2,3-D]pyrimidine	c12c(ncnc1scc2)NN
1,6-naphthyridine	c12cnccc1nccc2
8-aminoquinoline	c12cccc(N)c1nccc2
1,2,4-triazolo[1,5-A]pyrimidine	c12ncnn2cccn1
3-methoxymethylindole	c12c(COC)c[nH]c2cccc1
1-methylnaphthalene	c12cccc(C)c1cccc2
2-amino-4-hydroxypteridine	c12c([nH]c(N)nc1nccn2)=O
acenaphthylene	c12c3cccc1cccc2C=C3
2-(methylsulfanyl)-4-quinazolinamine	c12nc(SC)nc(c2cccc1)N
aniline	Nc1ccccc1
D-camphor	O=C1CC2CCC1(C)C2(C)C
3,5-dichloroanisole	COc1cc(Cl)cc(Cl)c1
(+)-camphene	CC1([C@@H]2CC[C@@H](C2)C1=C)C
3,5-difluoro-aniline	Nc1cc(F)cc(F)c1
3-methylyrrole	Cc1cc[nH]c1
2-fluoro-aniline	Nc1ccccc1F
2,4-difluoro-aniline	Nc1ccc(F)cc1F
2-fluorobenzaldehyde	O=Cc1c(cccc1)F
2,4-difluoro-phenol	Oc1ccc(F)cc1F
1-vinylimidazole	C(=C)N1ccnc1
3-fluorobenzonitrile	c1(ccc(cc1)C#N)
5-bromopyrimidine	c1ncncc1Br
4-vinylpyridine	c1(ccncc1)C=C
methyl chlorodifluoroacetate	C(C(OC)=O)(F)(F)Cl
nitrosobenzene	C1=CC=C(C=C1)N=O
2-benzylpyridine	c1(ccccc1)Cc2ccccn2
1-naphthalenemethanol	c12c(cccc1cccc2)CO
2-fluoroaniline	C1=CC=C(C(=C1)N)F
2-phenoxyethanol	C1=CC=C(C=C1)OCCO
1-phenyl-semicarbazide	NC(=O)NNc1ccccc1
phenol	Oc1ccccc1c1

Table S2: **YANK binding free energy estimates** between T4 lysozyme and 24 ligands. The mean value and standard deviation (in parentheses) of three independent simulations are reported.

Ligand names	Ligand structures	ΔG° (kcal/mol)	Ligand names	Ligand structures	ΔG° (kcal/mol)
1-methylpyrrole		-3.65 (0.39)	benzofuran		-5.31 (0.42)
4-ethyltoluene		-6.52 (0.14)	allyl ethyl sulfide		-4.23 (0.34)
benzene		-4.45 (0.25)	hexafluorobenzene		-2.42 (0.27)
indole		-4.48 (0.12)	m-xylene		-5.05 (0.21)
N-hexylbenzene		-7.90 (0.02)	nitrobenzene		-8.58 (0.19)
p-xylene		-5.84 (0.09)	DL-camphor		-7.45 (0.35)
1-propanol		-1.84 (0.19)	1,1-diethylurea		-4.98 (0.33)
1,4-diiodobenzene		-9.70 (0.13)	1,2-diiodobenzene		-11.48 (0.13)
2-bromoethanol		-2.16 (0.28)	2-iodoethanol		-2.28 (0.32)
benzyl alcohol		-3.75 (0.15)	benzaldehyde oxime		-5.53 (0.55)
phenol		-2.88 (0.39)	ethanol		-0.88 (0.14)
methanol		0.38 (0.27)	dimethyl-sulfoxide		-0.82 (0.98)

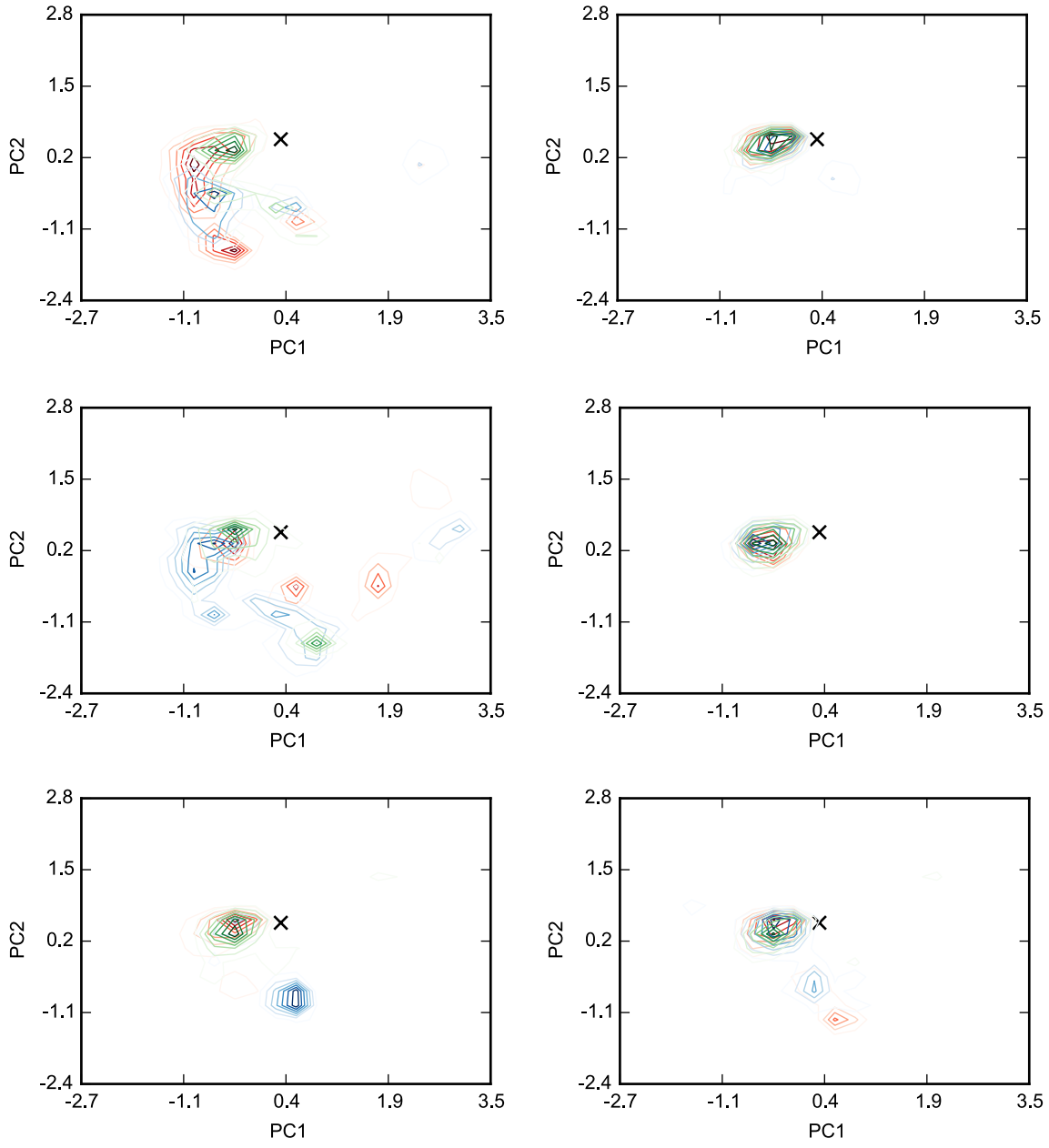

Figure S1: **2D histograms of the first and second principal components** of the unbound (left column) and bound (right column) receptor conformations sampled by YANK for complexes between T4-lysozyme and 24 ligands. Principal components analysis was performed on snapshots from all YANK simulations for heavy atoms within 5 Å of Val 111 in PDB ID 3DMZ. Ligands on this page, from top to bottom, are 1-methylpyrrole, benzene, benzofuran.

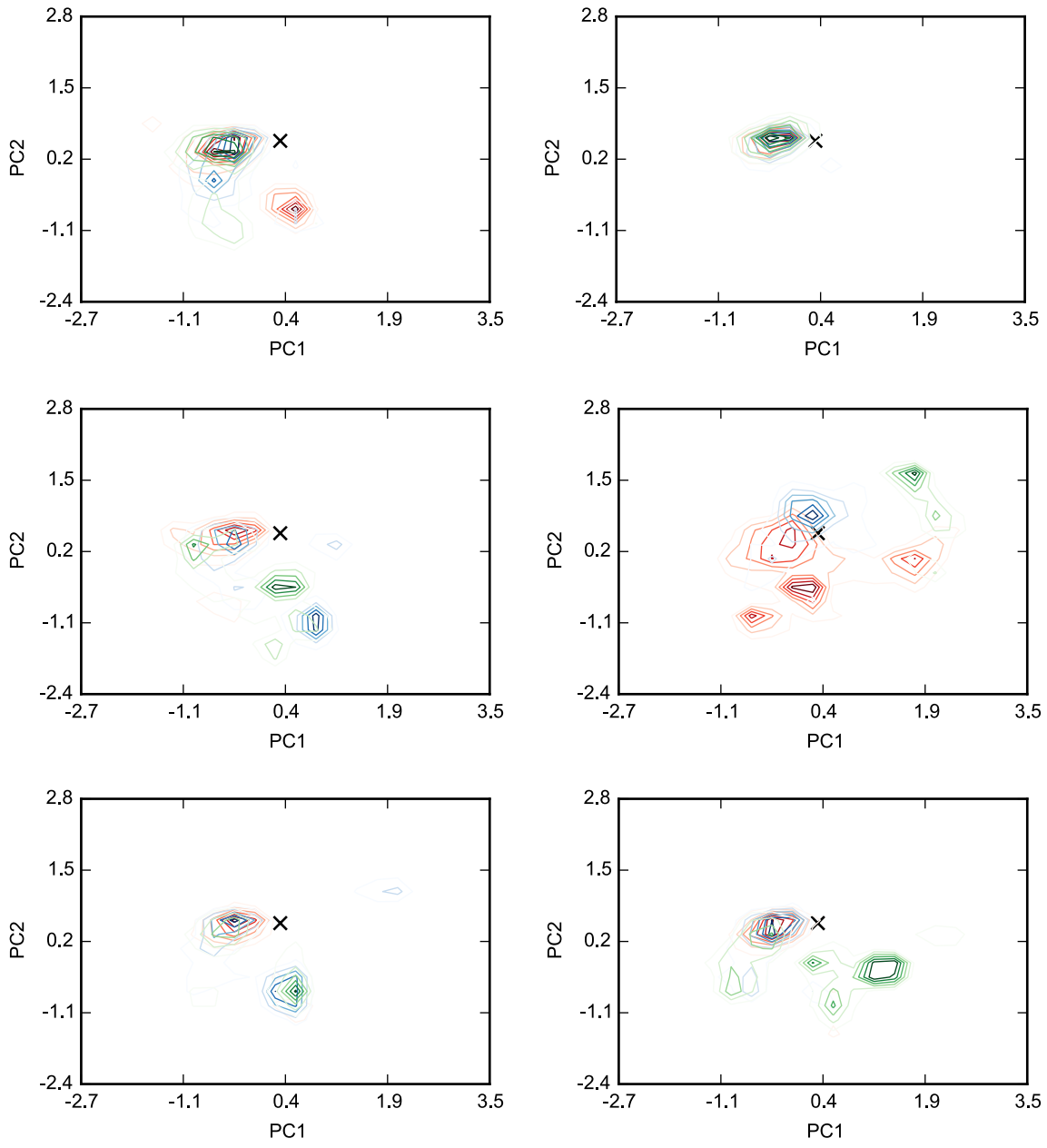

Figure S1: continued. **2D histograms of the first and second principal components** of the unbound (left column) and bound (right column) receptor conformations sampled by YANK for complexes between T4-lysozyme and 24 ligands. Principal components analysis was performed on snapshots from all YANK simulations for heavy atoms within 5 Å of Val 111 in PDB ID 3DMZ. Ligands on this page, from top to bottom, are allyl ethyl sulfide, hexafluorobenzene, indole.

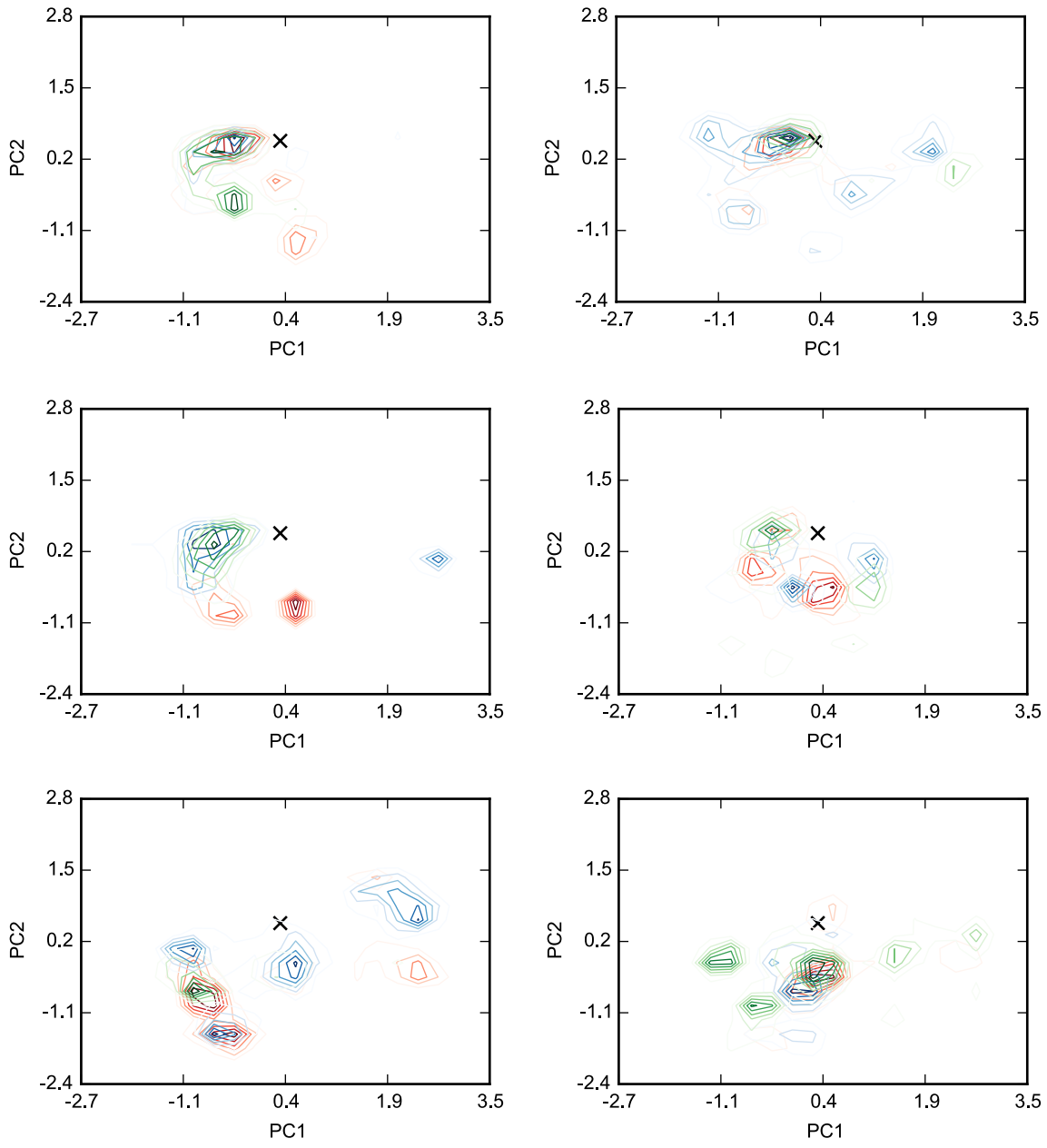

Figure S1: continued. **2D histograms of the first and second principal components** of the unbound (left column) and bound (right column) receptor conformations sampled by YANK for complexes between T4-lysozyme and 24 ligands. Principal components analysis was performed on snapshots from all YANK simulations for heavy atoms within 5 Å of Val 111 in PDB ID 3DMZ. Ligands on this page, from top to bottom, are m-xylene, n-hexylbenzene, nitrobenzene.

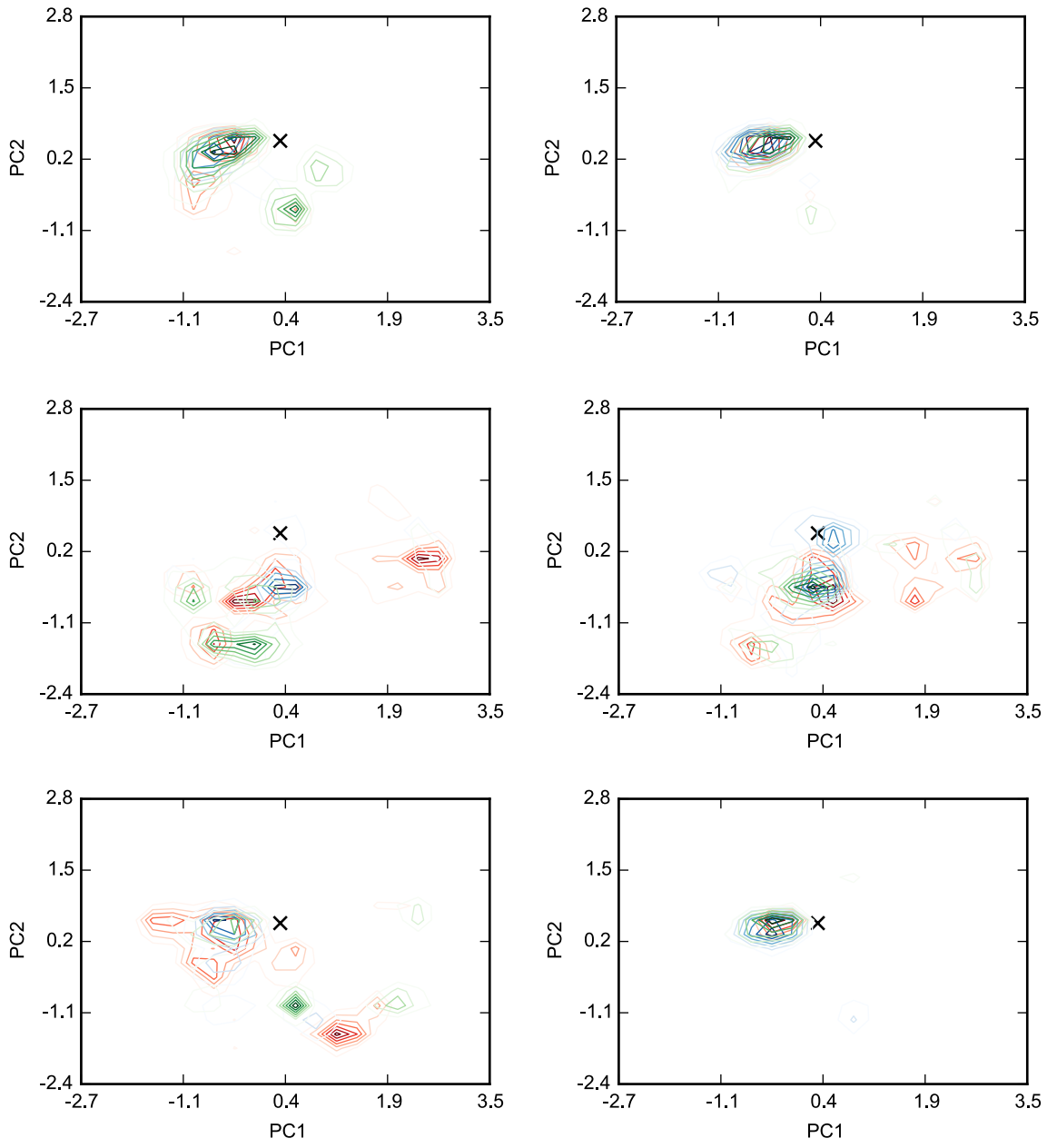

Figure S1: continued. **2D histograms of the first and second principal components** of the unbound (left column) and bound (right column) receptor conformations sampled by YANK for complexes between T4-lysozyme and 24 ligands. Principal components analysis was performed on snapshots from all YANK simulations for heavy atoms within 5 Å of Val 111 in PDB ID 3DMZ. Ligands on this page, from top to bottom, are ethanol, methanol, dimethyl-sulfoxide.

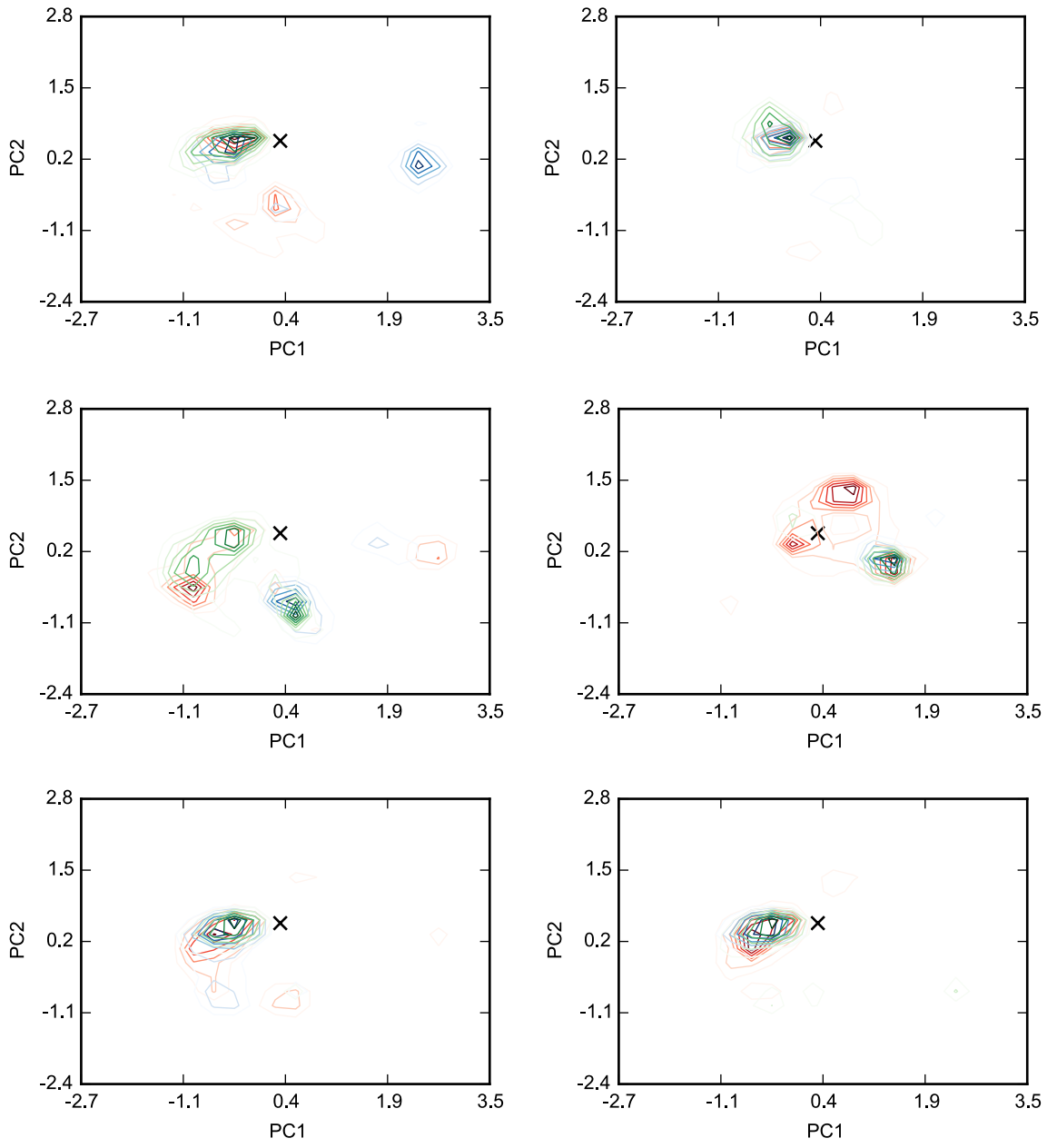

Figure S1: continued. **2D histograms of the first and second principal components** of the unbound (left column) and bound (right column) receptor conformations sampled by YANK for complexes between T4-lysozyme and 24 ligands. Principal components analysis was performed on snapshots from all YANK simulations for heavy atoms within 5 Å of Val 111 in PDB ID 3DMZ. Ligands on this page, from top to bottom, are 4-ethyltoluene, DL-camphor, 1-propanol.

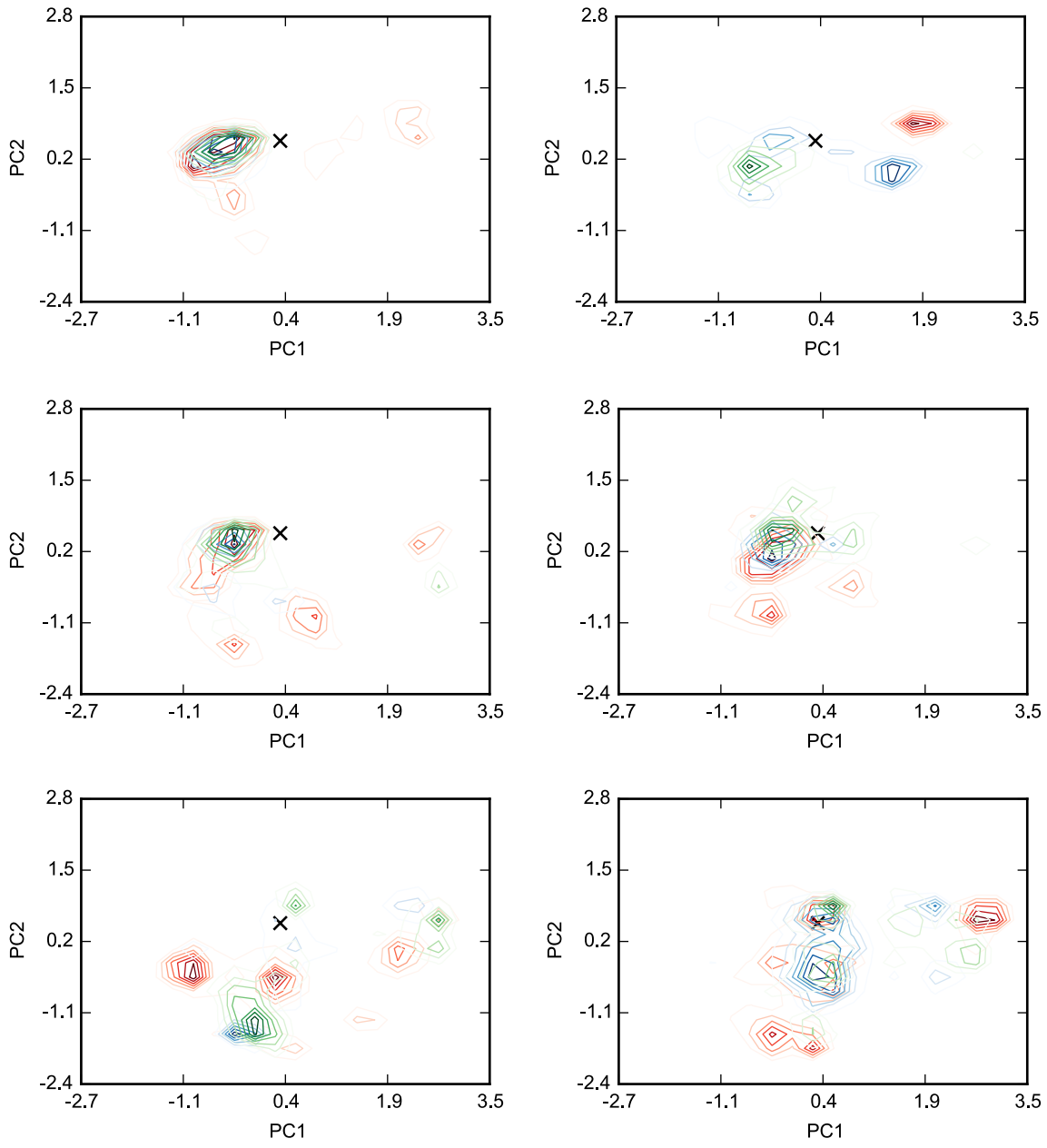

Figure S1: continued. **2D histograms of the first and second principal components** of the unbound (left column) and bound (right column) receptor conformations sampled by YANK for complexes between T4-lysozyme and 24 ligands. Principal components analysis was performed on snapshots from all YANK simulations for heavy atoms within 5 Å of Val 111 in PDB ID 3DMZ. Ligands on this page, from top to bottom, are 1,1-diethylurea, 1,4-diiodobenzene, 1,2-diiodobenzene.

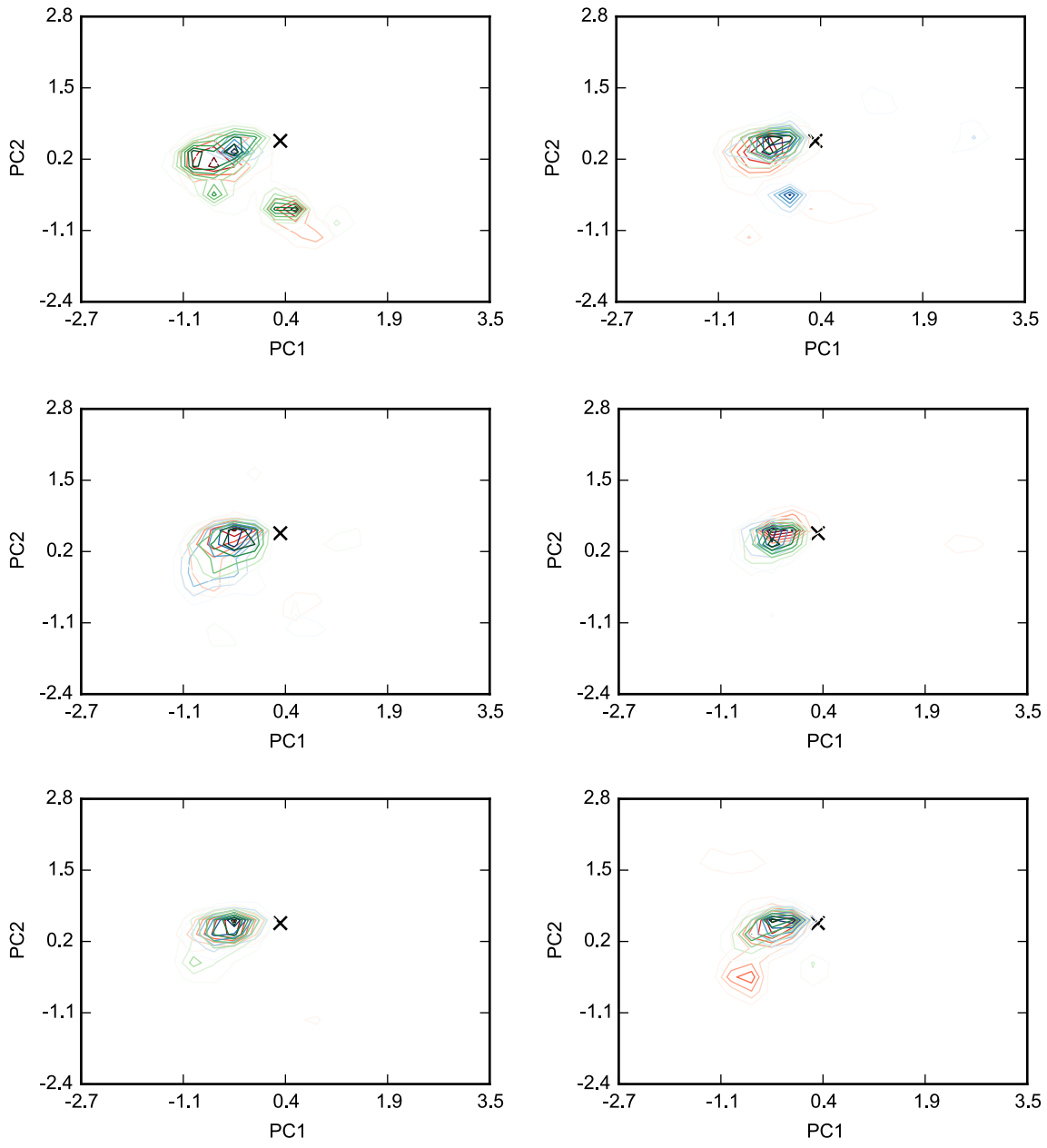

Figure S1: continued. **2D histograms of the first and second principal components** of the unbound (left column) and bound (right column) receptor conformations sampled by YANK for complexes between T4-lysozyme and 24 ligands. Principal components analysis was performed on snapshots from all YANK simulations for heavy atoms within 5 Å of Val 111 in PDB ID 3DMZ. Ligands on this page, from top to bottom, are 2-bromoethanol, 2-iodoethanol, benzyl-alcohol.

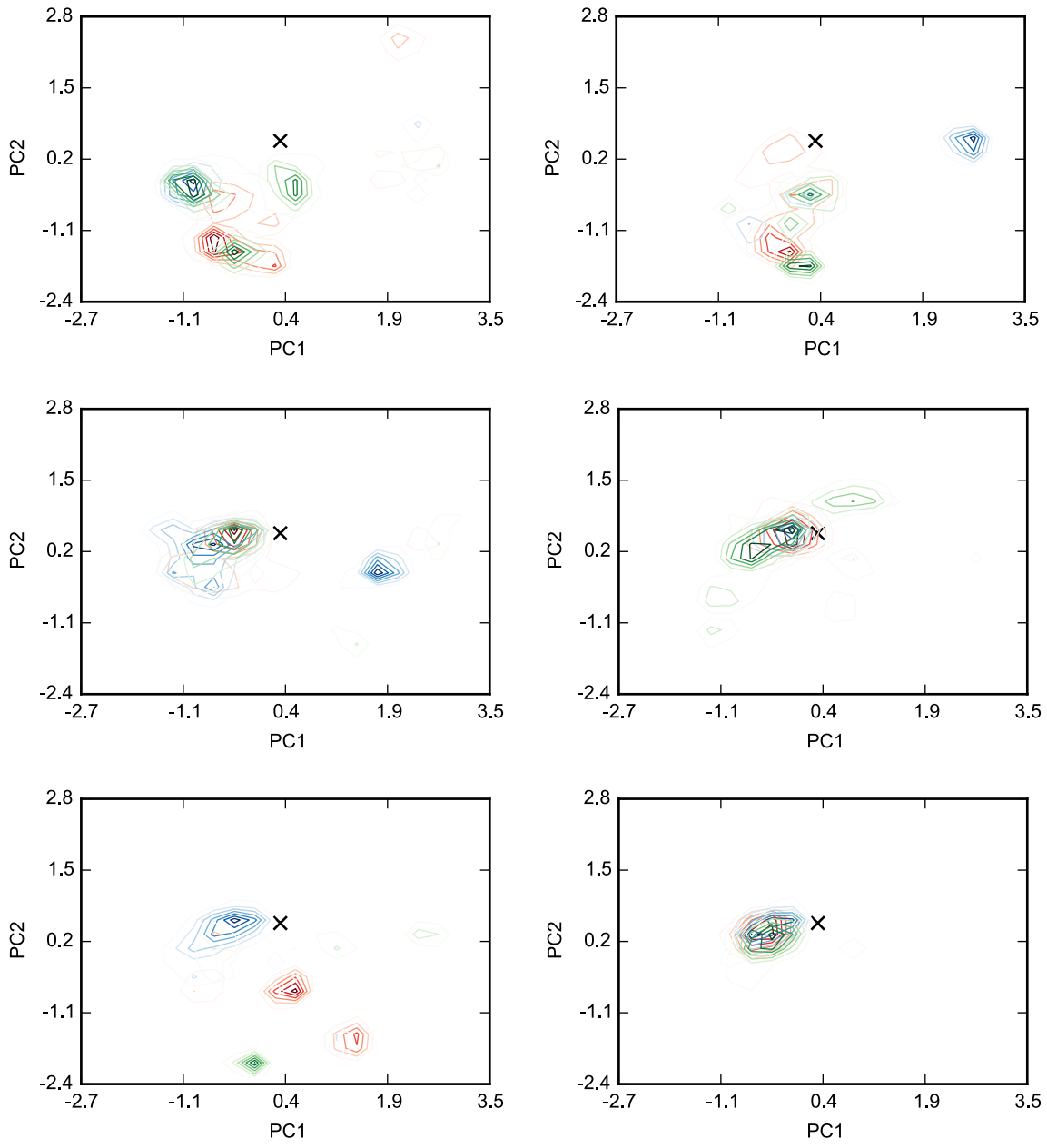

Figure S1: continued. **2D histograms of the first and second principal components** of the unbound (left column) and bound (right column) receptor conformations sampled by YANK for complexes between T4-lysozyme and 24 ligands. Principal components analysis was performed on snapshots from all YANK simulations for heavy atoms within 5 Å of Val 111 in PDB ID 3DMZ. Ligands on this page, from top to bottom, are benzaldehyde-oxime, p-xylene, phenol.

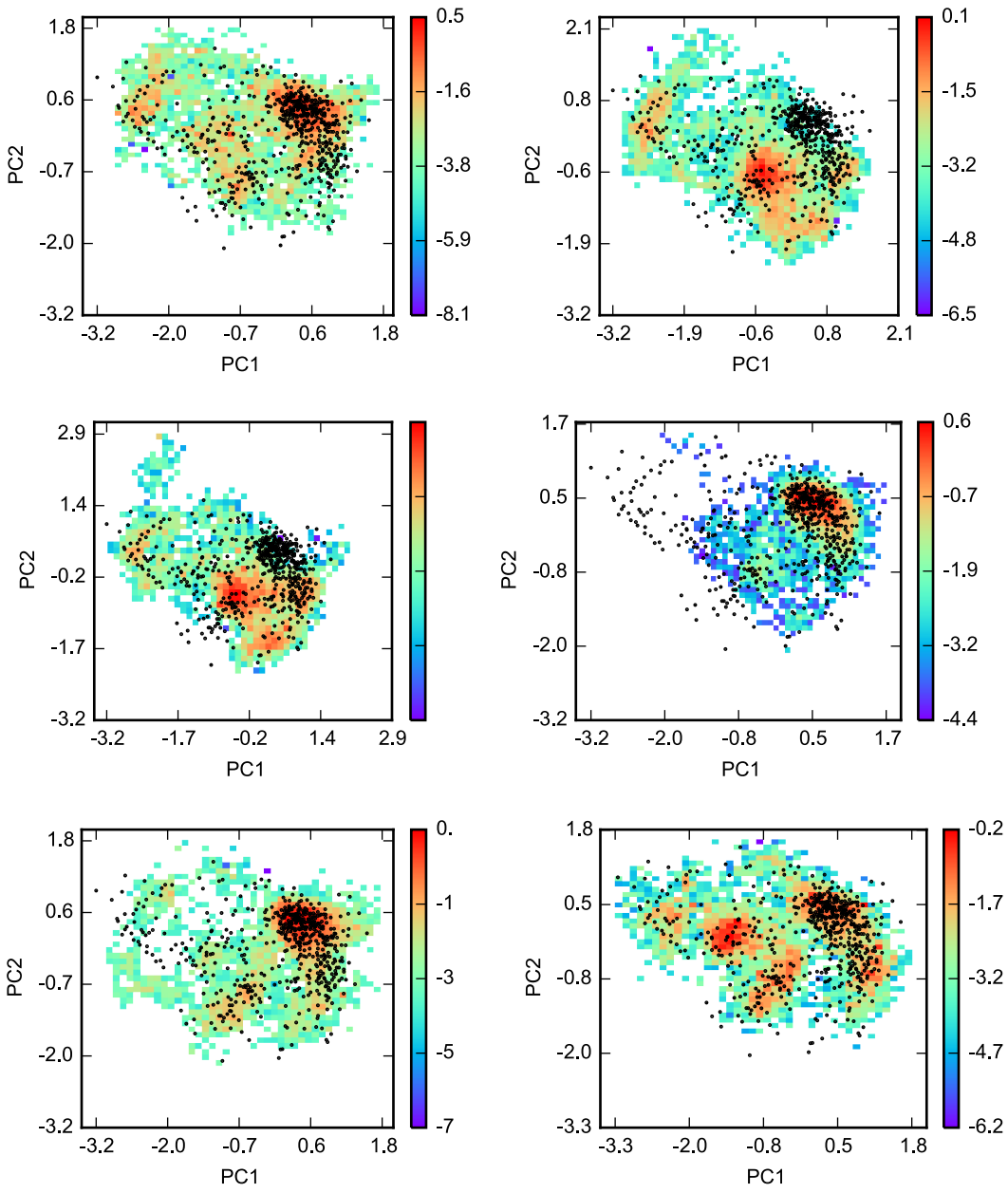

Figure S2: **Configuration space comparison** between YANK and AlGDock. Principal components analysis was performed on snapshots from all YANK simulations for heavy atoms within 5 Å of Val 111 in PDB ID 3DMZ. 2D histograms (plotted on a log scale) were created by projecting YANK snapshots on the first two PCs and weighting points by Eq. 5 of the main text. The black dots are projections of the 576 snapshots used in AlGDock calculations onto the same eigenvectors. Progressing from left to right and from the top to the bottom, the complexes have a decreasing absolute difference in binding free energy estimates between YANK and AlGDock. The ligand names are 1,1-diethylurea, methanol, benzaldehyde-oxime, ethanol, dimethyl-sulfoxide, and DL-camphor.

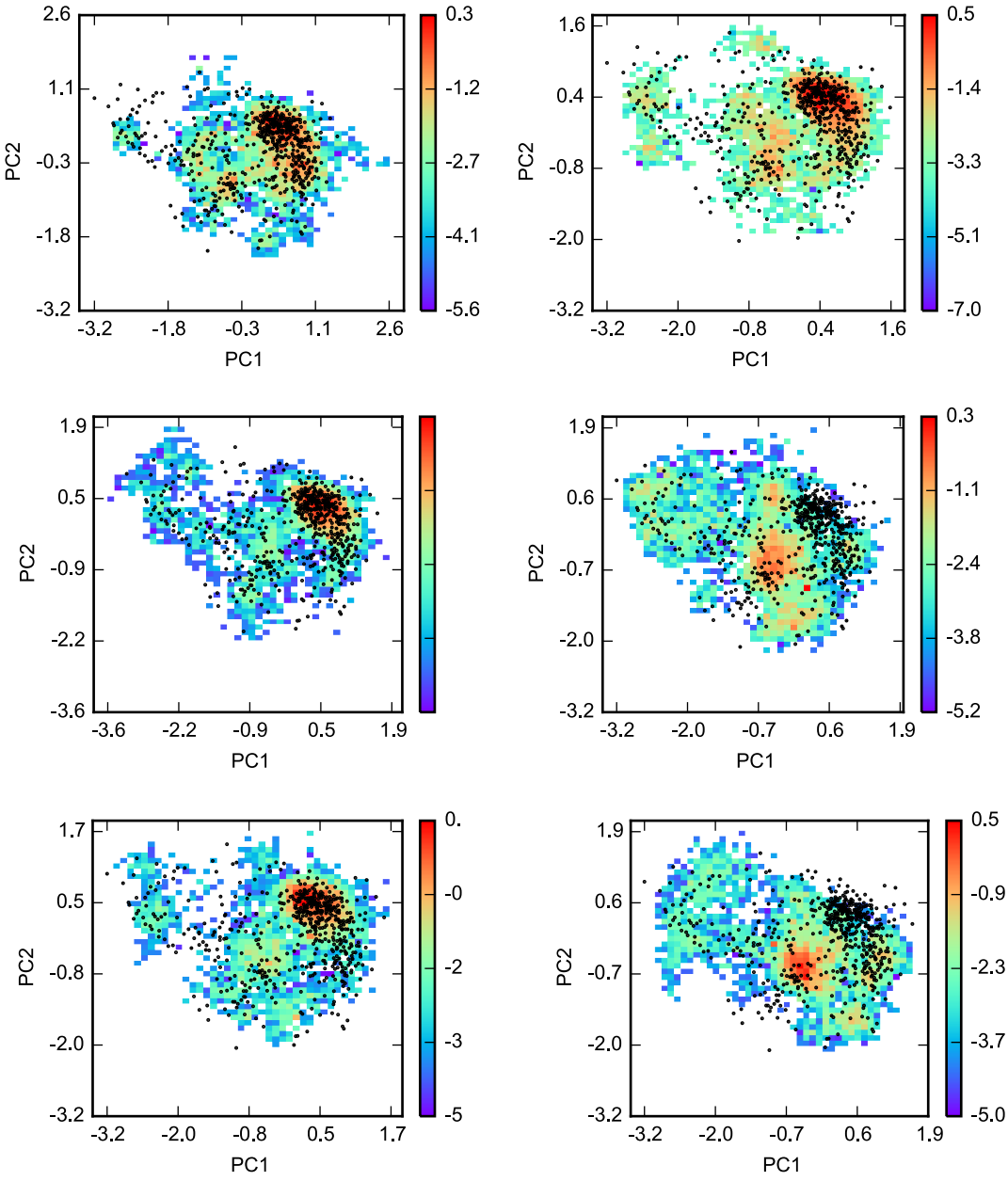

Figure S2: continued. **Configuration space comparison** between YANK and AlGDock. Principal components analysis was performed on snapshots from all YANK simulations for heavy atoms within 5 Å of Val 111 in PDB ID 3DMZ. 2D histograms (plotted on a log scale) were created by projecting YANK snapshots on the first two PCs and weighting points by Eq. 5 of the main text. The black dots are projections of the 576 snapshots used in AlGDock calculations onto the same eigenvectors. Progressing from left to right and from the top to the bottom, the complexes have a decreasing absolute difference in binding free energy estimates between YANK and AlGDock. The ligand names are n-hexylbenzene, 1-propanol, benzene, 1,2-diiodobenzene, 4-iodotoluene, and nitrobenzene.

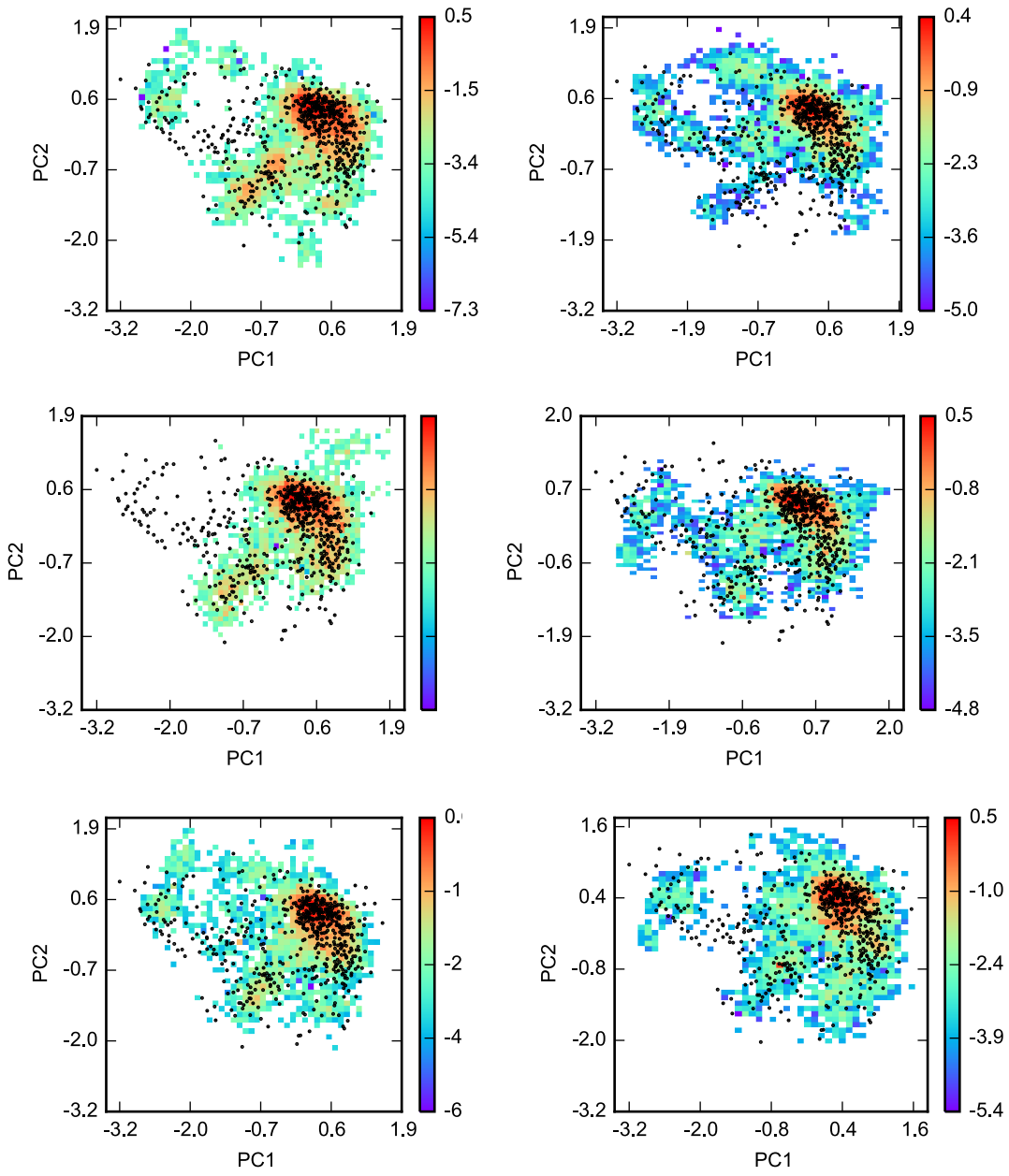

Figure S2: continued. **Configuration space comparison** between YANK and AlGDock. Principal components analysis was performed on snapshots from all YANK simulations for heavy atoms within 5 Å of Val 111 in PDB ID 3DMZ. 2D histograms (plotted on a log scale) were created by projecting YANK snapshots on the first two PCs and weighting points by Eq. 5 of the main text. The black dots are projections of the 576 snapshots used in AlGDock calculations onto the same eigenvectors. Progressing from left to right and from the top to the bottom, the complexes have a decreasing absolute difference in binding free energy estimates between YANK and AlGDock. The ligand names are 2-bromoethanol, p-xylene, benzyl-alcohol, m-xylene, 2-iodoethanol, and 1,4-diiodobenzene.

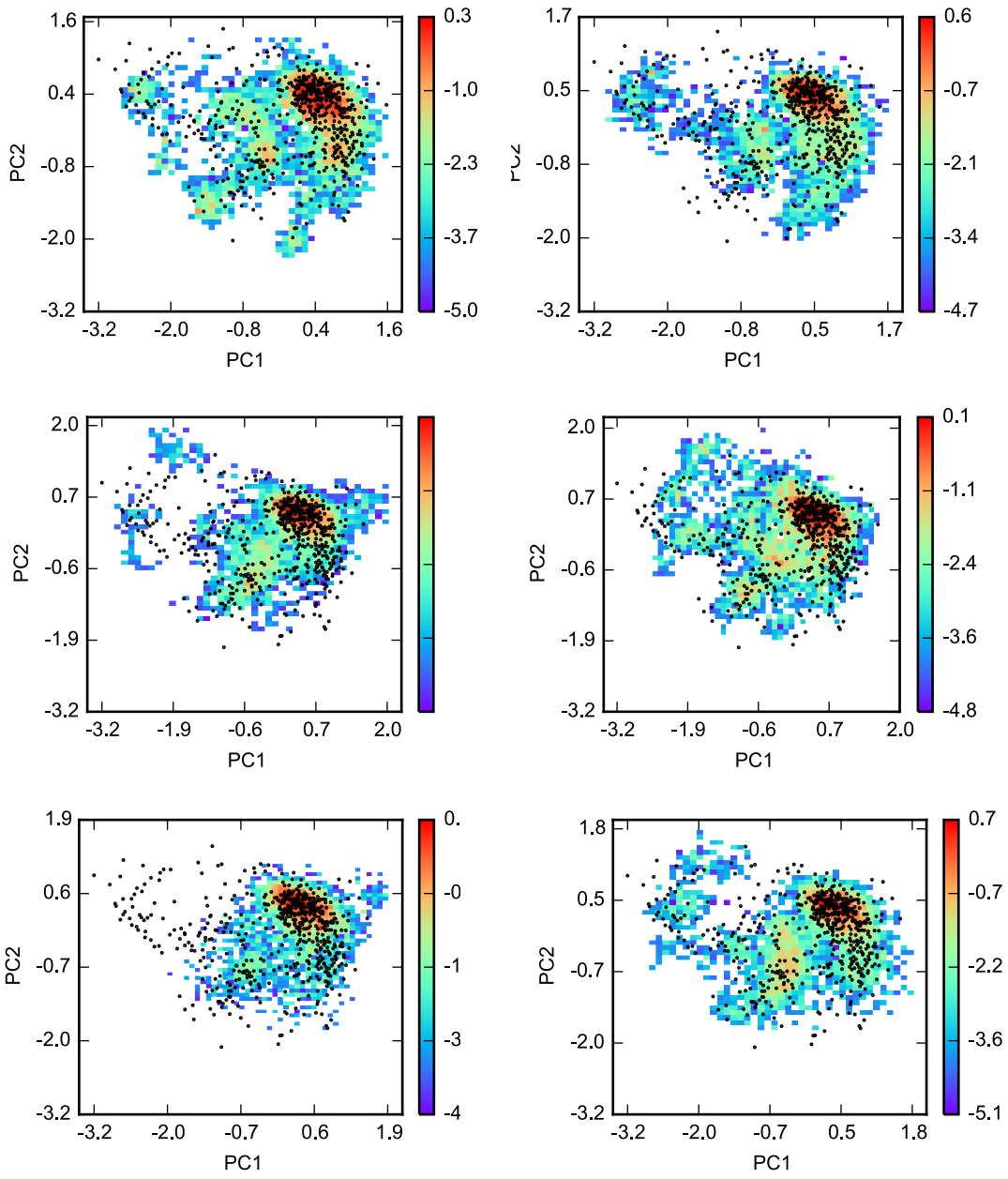

Figure S2: continued. **Configuration space comparison** between YANK and AlGDock. Principal components analysis was performed on snapshots from all YANK simulations for heavy atoms within 5 Å of Val 111 in PDB ID 3DMZ. 2D histograms (plotted on a log scale) were created by projecting YANK snapshots on the first two PCs and weighting points by Eq. 5 of the main text. The black dots are projections of the 576 snapshots used in AlGDock calculations onto the same eigenvectors. Progressing from left to right and from the top to the bottom, the complexes have a decreasing absolute difference in binding free energy estimates between YANK and AlGDock. The ligand names are phenol, 1-methylpyrrole, benzofuran, hexafluorobenzene, allyl ethyl sulfide, and indole.

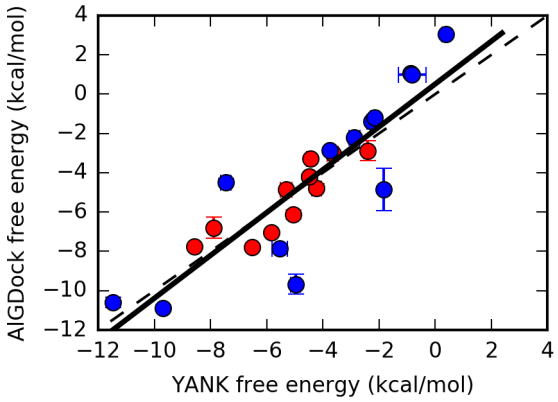

Figure S3: **Binding free energies for 24 ligands** estimated using YANK (x-axis) and AlGDock (y-axis). Only snapshots from YANK simulations with 4 active ligands were included. The snapshots were weighted with scheme (c).

Table S3: **Comparing AlGDock and YANK** binding free energy estimates. The AlGDock free energies were calculated using different weighting schemes. RMSE is in units of kcal/mol. In contrast to Table 1 of the main text, only BPMFs from simulations with active ligands were used.

	Full (24)			Active (11)		Inactive (13)	
	Linear Regression	RMSE	R	RMSE	R	RMSE	R
a	y=1.06x+0.83	1.82(0.21)	0.87(0.01)	1.25(0.19)	0.77(0.04)	2.19(0.37)	0.88(0.04)
b	y=1.07x+0.47	1.83(0.24)	0.86(0.01)	1.35(0.23)	0.72(0.05)	2.15(0.39)	0.88(0.03)
c	y=1.09x+0.52	1.75(0.22)	0.87(0.01)	0.89(0.11)	0.87(0.04)	2.23(0.36)	0.87(0.04)
d	y=1.10x+0.13	1.87(0.31)	0.87(0.02)	1.07(0.20)	0.86(0.03)	2.34(0.48)	0.87(0.04)
e	y=1.19x+0.95	1.52(0.18)	0.92(0.02)	0.95(0.36)	0.93(0.01)	1.87(0.21)	0.92(0.03)
f	y=1.19x+0.67	1.50(0.21)	0.93(0.02)	1.04(0.31)	0.94(0.02)	1.80(0.26)	0.92(0.03)

Table S4: **Solvation and binding free energies for all 141 ligands.** The column labeled B indicates whether the ligand had thermal shift activity (Y) or not (N). The columns for solvation (ΔG_{solv}) and binding (ΔG°) are experimental (E) or based on OBC2 (O) or PBSA (P) implicit solvent. Binding free energy estimates are based on AlGDock BPMF estimates for 384 receptor snapshots weighted by scheme (c). Ligands are listed in order of increasing PBSA binding free energy estimate. All free energies are in units of kcal/mol.

Ligand	B	ΔG_{solv}			ΔG°		
		E	O	P	E	O	P
(+)-camphene	N		2.33	5.52		-10.13(1.10)	-6.8(1.90)
n-propylbenzene	Y	-2.13	0.26	2.66	-6.55	-7.78(0.22)	-6.62(1.14)
methylcyclohexane	Y		2.42	5.30		-7.66(0.28)	-6.59(0.73)
cyclohexene	Y		1.46	3.52		-5.76(0.22)	-5.96(0.47)
diisopropyl disulfide	Y		-0.87	3.32		-6.13(0.41)	-5.87(1.47)
(-)-camphene	N		2.33	5.53		-8.34(0.92)	-5.86(1.99)
1,1-diethylurea	N		0.81	2.11		-9.67(0.53)	-5.61(0.59)
methyl ethyl cyclopentene	Y		2.49	5.15		-8.41(0.26)	-5.48(0.35)
cyclohexane	N		2.12	4.61		-6.64(0.27)	-5.32(0.26)
6-methyl-1,5-heptadiene	Y		2.38	5.63		-7.72(0.24)	-5.19(0.66)
2-(methylthio)-thiophene	Y		-0.85	0.84		-6.06(0.14)	-5.16(0.65)
1-phenyl-semicarbazide	N		0.16	0.36		-10.29(0.44)	-5.14(0.84)
dipropyl disulfide	Y		-0.10	3.88		-6.73(0.52)	-4.93(0.82)
4-ethyltoluene	Y	-1.65	0.56	2.70	-5.42	-7.80(0.18)	-4.68(0.81)
1,2-benzisoxazole	Y		-1.38	-1.39		-6.94(0.10)	-4.34(0.78)
isobutylbenzene	Y	-2.83	0.52	3.44	-6.51	-7.55(0.18)	-4.25(1.06)
amylbenzene	Y		0.88	4.24		-7.51(0.40)	-4.25(0.94)
diethyl disulfide	Y		-1.04	1.97		-5.75(0.36)	-4.08(0.32)
p-xylene	Y	-1.40	0.09	1.72	-4.67	-7.05(0.11)	-3.97(0.78)
n-butylbenzene	Y	-2.91	0.57	3.45	-6.70	-7.52(0.32)	-3.9(0.86)
2-ethyltoluene	Y	-1.65	0.33	2.41	-4.56	-7.57(0.25)	-3.79(0.84)
toluene	Y	-0.82	-0.51	0.90	-5.52	-4.80(0.12)	-3.75(0.75)
1-heptanol	N		-2.35	0.50		-7.39(0.62)	-3.56(1.05)
1,2-dichlorobenzene	Y		-0.27	0.61	-6.37	-7.76(0.18)	-3.55(0.50)
sec-butyl-benzene	Y		0.57	3.57		-8.04(0.48)	-3.5(0.82)
2-ethylfuran	Y		-0.56	1.28		-4.97(0.16)	-3.4(0.58)
1-vinylimidazole	N		-0.12	-0.05		-5.86(0.09)	-3.32(0.53)
indene	Y	-1.08	-1.77	-0.56	-5.13	-5.97(0.17)	-3.29(0.53)
2,1-benzisoxazole	N		-3.57	-3.41		-5.88(0.12)	-3.26(1.03)
2-methylbenzofuran	Y		-2.86	-1.86		-7.19(0.28)	-3.19(1.30)
thioanisole	Y		-1.24	0.67		-6.51(0.26)	-3.18(0.53)
allyl n-propylether	Y		0.21	2.35		-5.33(0.28)	-3.12(0.61)
fluorobenzene	Y		-0.87	0.69		-3.97(0.12)	-3.09(0.40)
thianaphthene	Y	-1.35	-2.39	-1.32	-5.71	-6.03(0.19)	-3.03(0.68)
benzofuran	Y	-1.24	-3.87	-3.25	-5.46	-4.85(0.12)	-3.03(0.98)
1-octanol	N		-2.05	1.22		-5.42(0.18)	-2.94(1.43)

allyl ethyl sulfide	Y		-0.48	2.09		-4.78(0.21)	-2.92(0.39)
phenyl azide	Y		-1.09	-1.05		-7.66(0.15)	-2.89(0.69)
benzene	Y	0.00	-1.10	0.10	-5.19	-3.29(0.13)	-2.85(0.63)
p-methylphenyl azide	Y		-0.41	-0.02		-8.20(0.15)	-2.82(0.81)
ethylbenzene	Y	-1.40	-0.06	1.83	-5.76	-6.11(0.18)	-2.77(0.42)
nitrosobenzene	N		-1.51	-0.99		-6.47(0.09)	-2.73(0.47)
1-phenyl-1-propyne	Y		-0.78	-0.10		-8.74(0.37)	-2.69(0.47)
m-xylene	Y	-1.38	0.11	1.76	-4.75	-6.12(0.22)	-2.56(0.36)
isobutane	N		2.30	4.43		-3.86(0.19)	-2.47(0.20)
3-ethyltoluene	Y	-2.01	0.56	2.73	-5.12	-7.48(0.22)	-2.44(1.09)
tert-butylbenzene	N		0.33	3.00		-9.19(0.78)	-2.38(0.73)
3-methyl benzyl azide	Y		-0.86	-0.34		-9.11(0.19)	-2.23(0.95)
cis-3-hexenyl formate	Y		-3.11	0.06		-5.70(0.33)	-2.22(0.85)
allyl sulfide	Y		-0.44	2.33		-5.13(0.22)	-2.19(0.43)
benzaldehyde-oxime	N		0.29	0.87		-7.86(0.15)	-2.18(0.36)
nitrobenzene	Y		-2.91	-2.87		-7.74(0.13)	-2.15(0.62)
2-fluoro-aniline	N		-5.54	-3.86		-2.36(0.13)	-2.08(0.94)
styrene	Y		-1.01	0.30		-5.21(0.13)	-2.06(0.41)
pentafluorophenyl-azide	N		-0.37	1.06		-6.67(0.56)	-2.05(0.88)
indole	Y	-0.02	-5.26	-4.97	-4.89	-4.20(0.16)	-1.94(0.92)
2,5-dimethylfuran	N		-0.59	1.14		-4.97(0.12)	-1.79(0.38)
2-fluoroaniline	N		-5.56	-3.91		-2.26(0.13)	-1.78(0.85)
3-phenyl-1-propanethiol	Y		-2.69	0.04		-6.57(0.25)	-1.73(0.57)
anisole	Y		-2.53	-1.20		-4.89(0.13)	-1.69(0.41)
1,2,4-triazolo[1,5-A]pyrimidine	N		-5.45	-6.26		-6.66(0.10)	-1.66(0.66)
2,4-difluoro-aniline	N		-5.01	-3.02		-2.97(0.16)	-1.6(0.78)
1-methylpyrrole	Y		-2.37	-1.45	-4.44	-3.01(0.11)	-1.55(0.42)
4-(methylthio)nitrobenzene	Y		-2.65	-1.88		-10.14(0.35)	-1.54(0.51)
phenylacetylene	Y		-1.18	-0.81		-7.61(0.13)	-1.45(1.07)
dibutyl-disulfide	N		0.57	5.55		-5.39(0.39)	-1.42(0.54)
hexafluorobenzene	Y		0.48	3.32		-2.89(0.46)	-1.4(0.40)
naphthalene	Y		-2.25	-1.26		-6.81(0.31)	-1.13(0.43)
2,4-difluoro-phenol	N		-3.85	-2.13		-3.35(0.15)	-1.01(0.48)
benzyl mercaptan	Y		-2.98	-1.39		-4.75(0.14)	-0.95(0.48)
3,5-dichloroanisole	N		-0.82	0.34		-8.51(0.69)	-0.92(0.86)
azulene	N		-2.74	-1.65		-6.06(0.33)	-0.91(0.65)
o-xylene	Y	-1.35	-0.10	1.46	-4.60	-6.15(0.21)	-0.89(0.54)
β-chlorophenetole	Y		-3.22	-1.44		-6.59(0.19)	-0.75(0.70)
2-ethoxyphenol	Y		-4.40	-2.83		-5.90(0.24)	-0.65(0.70)
2,6-difluorophenyl azide	N		-1.52	-0.46		-6.39(0.21)	-0.62(0.52)
2-(thienylthio)-acetone	Y		-5.75	-3.34		-5.68(0.34)	-0.58(0.57)

p-cresol	N	-5.22	-4.30		-4.11(0.11)	-0.56(0.88)
4-iodotoluene	Y	-0.71	-3.69		-9.41(0.12)	-0.51(1.00)
mesitylene	N	0.76	2.73		-6.35(0.66)	-0.5(1.28)
3-methylyrrole	N	-4.04	-3.33		-2.09(0.10)	-0.43(0.35)
aniline	N	-6.01	-4.80		-1.35(0.12)	-0.22(0.52)
quinoline	N	-4.92	-3.94		-6.16(0.23)	-0.15(0.50)
3,5-difluoro-aniline	N	-4.95	-3.10		-2.81(0.21)	-0.15(0.61)
furan	N	-1.31	-0.44		-2.08(0.11)	-0.13(0.32)
phenol	N	-5.89	-5.13		-2.21(0.11)	-0.02(0.54)
2,6-difluorobenzyl bromide	Y	-1.49	-1.45		-7.52(0.20)	0.02(0.33)
methyl chlorodifluoroacetate	N	-4.81	-1.49		-0.61(0.21)	0.09(0.22)
pyridine	N	-4.65	-3.79		-1.63(0.10)	0.16(0.51)
pentafluoroaniline	N	-3.97	-1.38		-2.83(0.40)	0.24(0.50)
2-methylbenzyl alcohol	N	-4.31	-3.35		-5.04(0.18)	0.26(0.46)
3-phenyl-1-propanol	Y	-4.95	-3.05		-5.38(0.20)	0.27(0.57)
2-fluorobenzaldehyde	N	-5.44	-3.13		-3.50(0.17)	0.32(0.60)
3-fluorobenzonitrile	N	-3.95	-3.14		-6.38(0.12)	0.34(0.26)
DL-camphor	N	-2.43	1.64		-4.49(0.29)	0.45(1.35)
2-phenylpropanol	Y	-4.40	-2.46		-5.92(0.61)	0.73(0.88)
6,7-ifluorobenzotriazole	N	-10.61	-10.33		-3.81(0.25)	0.77(0.99)
benzyl-alcohol	N	-5.12	-4.35		-2.85(0.13)	0.77(0.57)
n-methylaniline	Y	-4.32	-3.60	-4.70	-3.49(0.12)	0.79(0.42)
3-chlorophenol	N	-5.13	-4.57		-4.48(0.17)	0.8(0.37)
n-hexylbenzene	Y	1.17	5.03		-6.79(0.58)	0.86(0.60)
iodopentafluorobenzene	Y	0.64	-1.20	-4.45	-7.02(0.44)	0.88(0.96)
4-vinylpyridine	N	-4.49	-3.51		-3.72(0.14)	0.89(0.56)
n,n'-dimethylaniline	N	-2.57	-2.34		-4.30(0.17)	1.05(0.71)
D-camphor	N	-2.40	1.67		-4.27(0.41)	1.13(1.76)
1-methylnaphthalene	N	-1.67	-0.40		-7.11(0.59)	1.19(0.78)
2-iodotoluene	Y	-0.79	-3.54		-8.37(0.20)	1.24(0.48)
iodobenzene	Y	-1.30	-4.62	-5.95	-7.56(0.17)	1.45(0.56)
3-iodotoluene	Y	-0.64	-3.55		-8.61(0.22)	1.49(0.65)
2-phenoxyethanol	N	-5.72	-4.28		-4.60(0.24)	1.5(0.63)
bromoethane	Y	-0.16	-1.74		-2.96(0.10)	1.59(0.22)
benzaldehyde	N	-6.51	-4.62		-2.74(0.11)	1.6(0.36)
1,6-naphthyridine	N	-8.05	-7.56		-3.83(0.15)	1.85(0.67)
1-propanol	N	-3.61	-2.46		-4.86(1.09)	2.01(0.19)
3-methoxymethylindole	N	-5.16	-4.68		-7.64(0.52)	2.57(0.64)
5-bromopyrimidine	N	-5.77	-7.71		-2.75(0.09)	2.84(0.45)
2-iodopropane	Y	-2.27	-4.85		-3.67(0.20)	3.25(0.26)
ethanol	N	-4.09	-3.32		1.05(0.10)	3.3(0.22)
N-heptylbenzene	N	1.46	5.80		-12.50(1.64)	3.5(0.95)
8-aminoquinoline	N	-8.80	-7.50		-3.50(0.33)	3.54(0.52)

2-benzylpyridine	N	-5.05	-2.70	-8.84(1.37)	3.62(0.93)
allyl ethyl sulfoxide	Y	-12.63	-8.19	0.94(0.26)	3.74(0.61)
acenaphthylene	N	-3.71	-2.99	-6.34(0.74)	3.83(0.83)
trans-cinnamaldehyde	N	-6.92	-4.86	-4.13(0.45)	4.13(0.24)
1-naphthalenemethanol	N	-5.44	-4.74	-5.84(0.70)	4.15(0.78)
iodoethane	Y	-2.70	-5.98	-2.50(0.10)	4.71(0.20)
dimethyl-sulfoxide	N	-8.62	-7.47	0.98(0.17)	4.85(0.21)
methanol	N	-4.78	-4.17	3.05(0.08)	4.88(0.15)
1,2-diiodobenzene	N	-1.27	-8.59	-9.59(0.50)	6.26(1.19)
1,4-diiodobenzene	N	-1.23	-9.01	-10.46(0.38)	5.48(1.05)
2,5-diaminophenol	N	-14.48	-13.75	0.88(0.19)	5.09(0.91)
2-bromoethanol	N	-5.49	-7.37	-0.87(0.10)	5.15(0.23)
2-(methylsulfanyl)-4-quinazolinamine	N	-8.28	-8.55	-5.13(0.62)	5.35(0.78)
4-hydrazinothieno[2,3-D]pyrimidine	N	-13.00	-12.70	-0.98(0.29)	6.97(0.69)
2-iodoethanol	N	-7.70	-11.25	-0.44(0.11)	8.14(0.22)
benzoic-acid	N	-77.35	-75.10	2.87(0.65)	14.45(1.47)
thiophenol	Y	-71.46	-68.21	2.32(0.76)	16.24(0.72)
thiophene-2-thiol	Y	-67.00	-64.79	1.70(0.44)	16.31(0.79)
2-amino-4-hydroxypteridine	N	-24.63	-27.44	-0.09(0.82)	16.45(0.87)
1,2-benzenedithiol	Y	-232.92	-226.54	18.14(0.25)	38.64(2.53)
chloroform	N	inf	inf	inf(nan)	inf(nan)

Table S5: **AmberTools radii versus Bondi radii.** A comparison of binding free energies for molecules containing chlorine, bromine, or iodine from experiment (ΔG°_{exp}), based on radii from AmberTools (ΔG°_{AT}), or based on Bondi radii (ΔG°_{Bondi}). (AmberTools incorrectly specifies the Bondi radii of most halogens at 1.5 Å.) The columns specify whether estimates are based on OBC2 (O) or PBSA (P) implicit solvent. The column labeled B indicates whether the ligand had thermal shift activity (Y) or not (N).

Ligand name	B	ΔG°_{exp}	ΔG°_{AT}		ΔG°_{Bondi}	
		E	O	P	O	P
1,2-dichlorobenzene	Y	-6.37	-7.78	-2.49	-7.76	-3.55
β-chlorophenetole	Y		-6.62	-0.81	-6.59	-0.75
3,5-dichloroanisole	N		-8.57	-0.94	-8.51	-0.92
methyl chlorodifluoroacetate	N		-0.64	0.75	-0.61	0.09
3-chlorophenol	N		-4.49	1.57	-4.48	0.8
chloroform	N		inf	inf	inf	inf
bromoethane	Y		-3.25	2.61	-2.96	1.59
2,6-difluorobenzyl bromide	Y		-7.96	-0.10	-7.52	0.02
5-bromopyrimidine	N		-2.88	3.89	-2.75	2.84
2-bromoethanol	N		-1.21	5.72	-0.87	5.15
iodoethane	Y		-3.13	5.49	-2.50	4.71
2-iodopropane	Y		-4.45	3.84	-3.67	3.25
2-iodotoluene	Y		-8.84	1.29	-8.37	1.24
3-iodotoluene	Y		-8.95	1.56	-8.61	1.49
4-iodotoluene	Y		-9.60	1.42	-9.41	-0.51
iodobenzene	Y	-5.95	-7.72	2.91	-7.56	1.45
1,4-diiodobenzene	N		-10.90	5.27	-10.46	5.48
1,2-diiodobenzene	N		-10.61	4.82	-9.59	6.26
2-iodoethanol	N		-1.40	7.66	-0.44	8.14

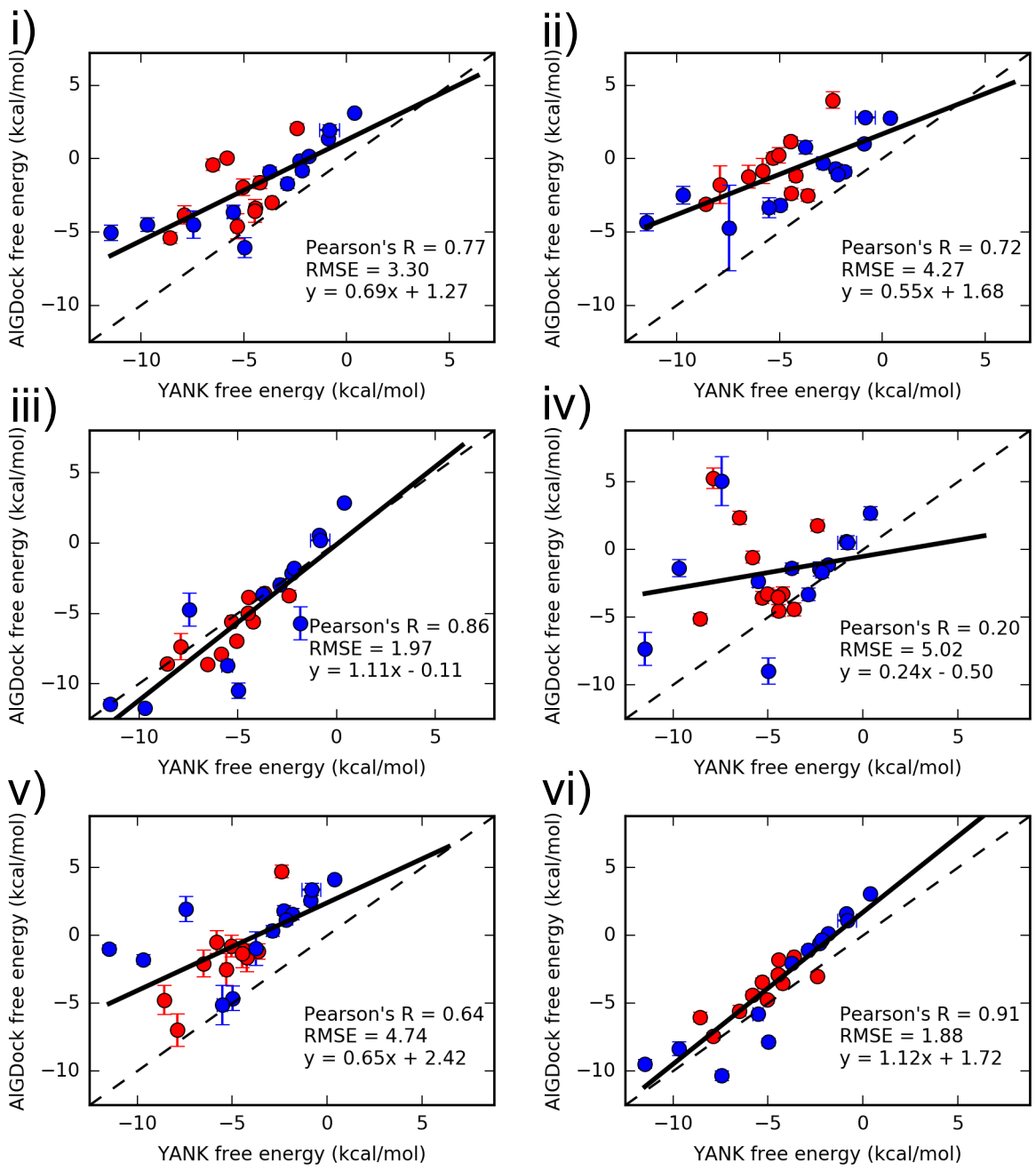

Figure S4: **Binding free energies for 24 ligands** estimated using YANK (x-axis) and AlGDock (y-axis). The AlGDock free energies were calculated using sets of receptor snapshots obtained from separate YANK simulations for ligands (i) to (vi) in complex with T4 Lysozyme.

Table S6: **Comparison of measured and calculated binding free energies.** Previously reported values include isothermal titration calorimetry measurements from Morton et al.[2], Mobley et al.[3], and Liu et al.[4] (ΔG_{exp}) and estimates from Mobley et al.[3] (M), Gallicchio et al.[5] (G), Purisima and Hogues[6] (P), and Ucisik et al.[7] (U). AlGDock BPMF estimates were weighted by scheme (c). All free energies are in units of kcal/mol.

Ligand name	ΔG_{exp}	M	G	P	U	OBC2	PBSA
benzofuran	-5.46	-3.53		-6.99	-15.51	-4.85	-3.03
benzene	-5.19	-4.56	-4.01	-5.32	-10.45	-3.29	-2.85
ethylbenzene	-5.76	-6.36				-6.11	-2.77
indene	-5.13	-1.75		-5.14	-15.86	-5.97	-3.29
indole	-4.89	-0.42	-3.75	-3.47	-14.04	-4.20	-1.94
iodobenzene	-5.95					-7.56	1.45
iodopentafluorobenzene	-4.45					-7.02	0.88
isobutylbenzene	-6.51	-5.01	-5.21	-7.35	-19.90	-7.55	-4.25
N-butylbenzene	-6.70	-4.87		-7.50	-19.84	-7.52	-3.90
N-propylbenzene	-6.55	-5.88				-7.78	-6.62
o-xylene	-4.60	-1.27		-3.38	-13.68	-6.15	-0.89
p-xylene	-4.67	-3.54		-3.63	-13.98	-7.05	-3.97
toluene	-5.52	-4.58	-3.80			-4.80	-3.75
1-methylpyrrole	-4.44	-4.32				-3.01	-1.55
1,2-dichlorobenzene	-6.37	-5.66				-7.76	-3.55
N-methylaniline	-4.7	-5.37				-3.49	0.79
m-xylene	-4.75					-6.12	-2.56
2-ethyltoluene	-4.56					-7.57	-3.79
3-ethyltoluene	-5.12					-7.48	-2.44
4-ethyltoluene	-5.42					-7.80	-4.68
thianaphthene	-5.71					-6.03	-3.03

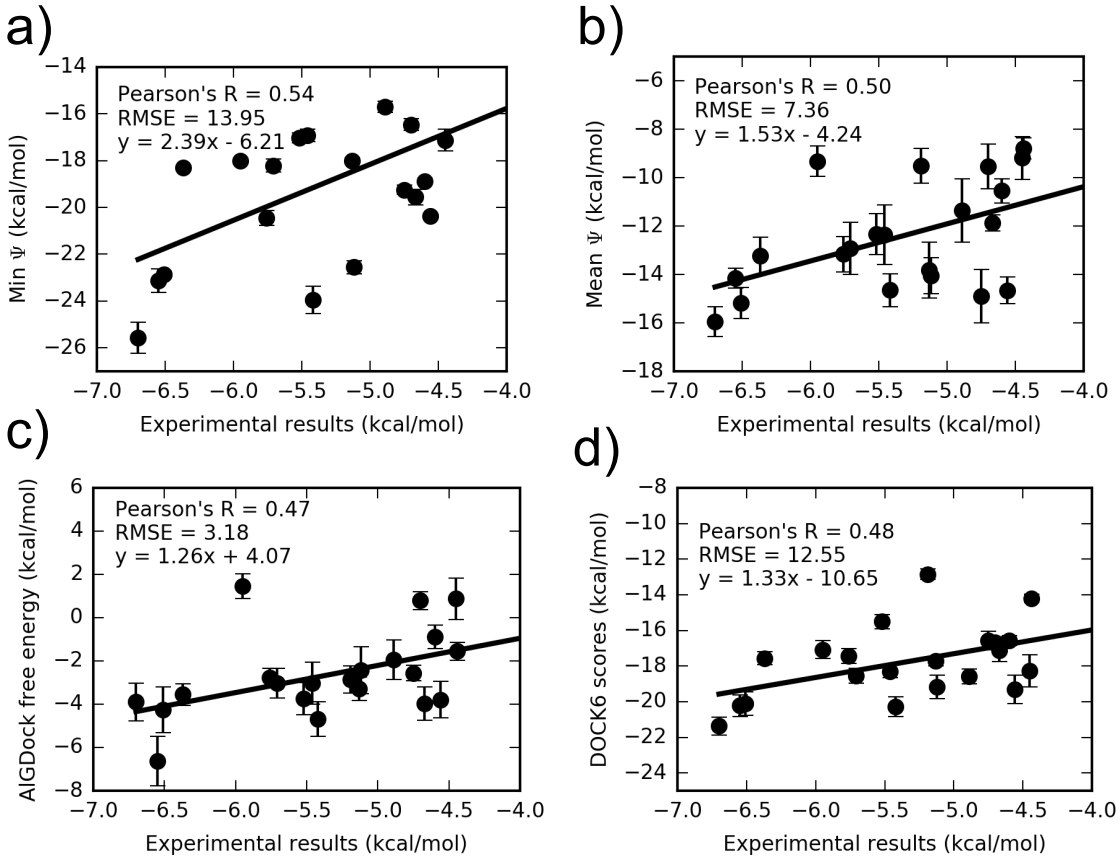

Figure S5: **Comparing different free energy estimates with experiment.** AlGDock free energy estimates in PBSA implicit solvent were based on BPMFs calculated with either the (a) minimum interaction energy, (b) mean interaction energy, or (c) MBAR estimator, using weighting scheme (c). A comparison of the average UCSF DOCK 6 grid score with experiment is shown in (d). Note that the y axes have different limits. In contrast to Fig. 10 of the main text, this figure includes the outlier iodobenzene.

Table S7: **Comparing different free energy estimates with experiment.** AlGDock calculations were based on 384 snapshots from simulations with active ligands, weighted according to scheme (c). As DOCK 6 does not have an implicit solvent, OBC2 and PBSA describe the receptor potential energies used for snapshot weights. In contrast with Table 3 of the main text, this analysis includes the outlier, iodobenzene.

Method	Implicit solvent	Spearman ρ	Kendall τ	Pearson R	RMSE
DOCK 6	OBC2	0.33(0.00)	0.22(0.01)	0.43(0.04)	11.98(0.36)
DOCK 6	PBSA	0.41(0.01)	0.30(0.01)	0.48(0.04)	12.56(0.37)
min Ψ	OBC2	0.41(0.00)	0.30(0.01)	0.54(0.03)	15.42(0.59)
min Ψ	PBSA	0.43(0.00)	0.33(0.01)	0.54(0.03)	13.95(0.60)
mean Ψ	OBC2	0.41(0.01)	0.32(0.01)	0.51(0.04)	9.62(0.47)
mean Ψ	PBSA	0.47(0.01)	0.36(0.01)	0.50(0.04)	7.36(0.41)
MBAR-based	OBC2	0.39(0.01)	0.27(0.02)	0.46(0.05)	1.54(0.18)
MBAR-based	PBSA	0.48(0.00)	0.38(0.01)	0.47(0.06)	3.18(0.44)

References

(1) Weininger, D. *J. Chem. Inf. Comput. Sci.* **1988**, *28*, 31–36.

(2) Morton, A.; Baase, W. A.; Matthews, B. W. *Biochemistry* **1995**, *34*, 8564–8575.

(3) Mobley, D. L.; Graves, A. P.; Chodera, J. D.; McReynolds, A. C.; Shoichet, B. K.; Dill, K. A. *J. Mol. Biol.* **2007**, *371*, 1118–1134.

(4) Liu, L.; Baase, W. A.; Matthews, B. W. *J. Mol. Biol.* **2009**, *385*, 595–605.

(5) Gallicchio, E.; Lapelosa, M.; Levy, R. M. *J. Chem. Theory Comput.* **2010**, *6*, 2961–2977.

(6) Purisima, E. O.; Hogues, H. *J. Phys. Chem. B* **2012**, *116*, 6872–9.

(7) Ucisik, M. N.; Zheng, Z.; Faver, J. C.; Merz, K. M. *J. Chem. Theory Comput.* **2014**, *10*, 1314–1325.