Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.
This journal is © the Owner Societies 2023

Supplementary Information

Target-ligand binding affinity from single point enthalpy calculation and elemental composition

Viktor Szél, Balázs Zoltán Zsidó, Norbert Jeszenői, and Csaba Hetényi*
*Pharmacoinformatics Unit, Department of Pharmacology and Pharmacotherapy, Medical School, University of Pécs, Szigeti út 12, 7624 Pécs, Hungary. *Corresponding author.*

Table of contents

Figures

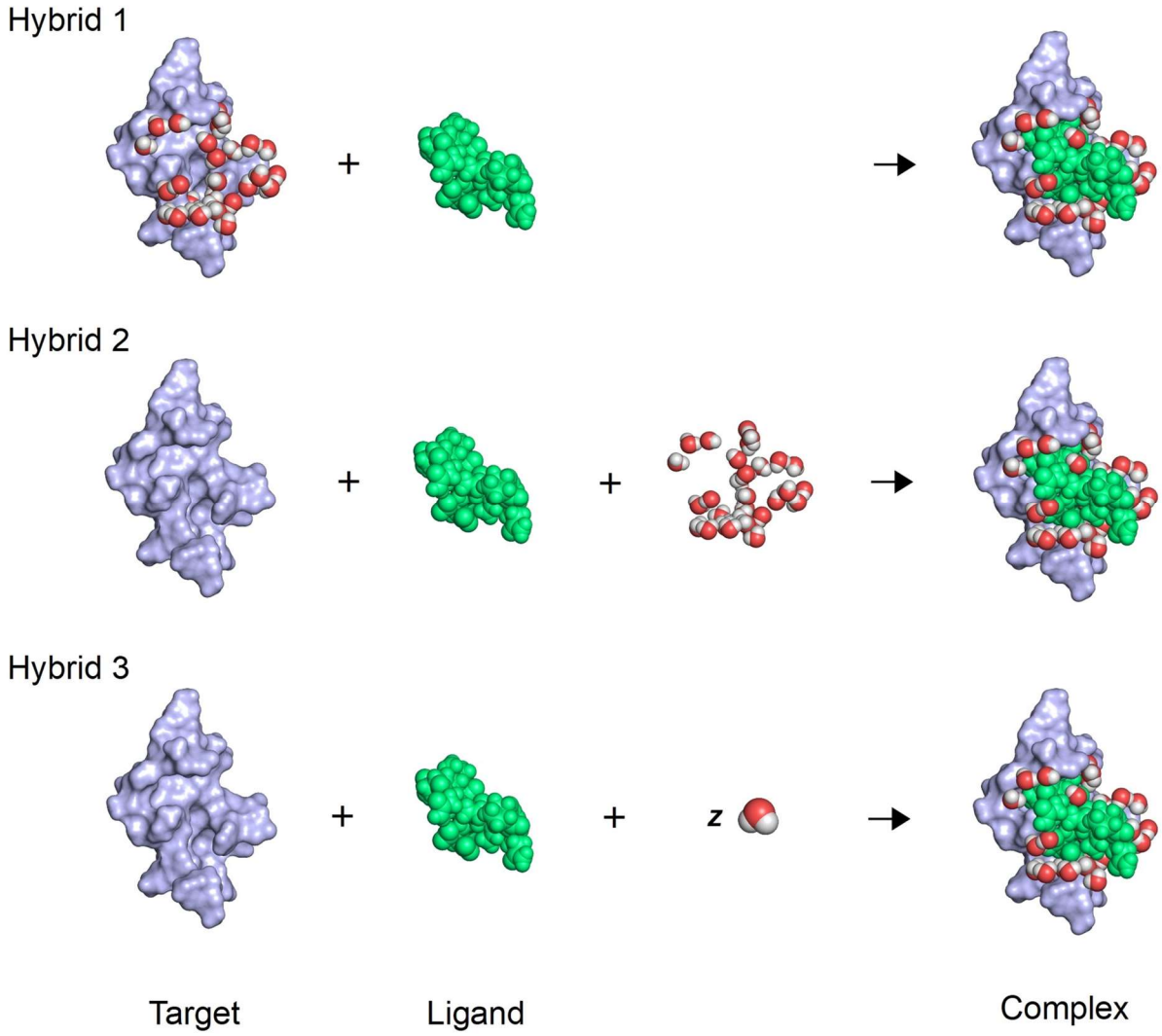

Fig. S1 Assignment of water molecules in the Hybrid Models

The number of waters assigned to the target, ligand or to the bulk are denoted by x, y, z respectively in **Eq. 2**. In Hybrid Model 1, interfacial waters are assigned to the target molecule ($y = z = 0$, $x \neq 0$) in their ligand-bound orientation on the reactant (left) side. Upon complex formation, the ligand binds to the hydrated target structure. This model does not account for the re-orientation of water molecules during ligand binding. In Hybrid Model 2 ($x = y = 0$, $z \neq 0$:waters in cluster), water molecules are treated together as a cluster on the reactant (left) side, and the complex is formally assembled from three components, the target, the ligand, and the water cluster. This model does not account for the enthalpy change corresponding to interfacial water-water interactions. In Hybrid Model 3 ($x = y = 0$, $z \neq 0$:individual waters), no interaction between interfacial waters and the target was assumed on the reactant (left) side.

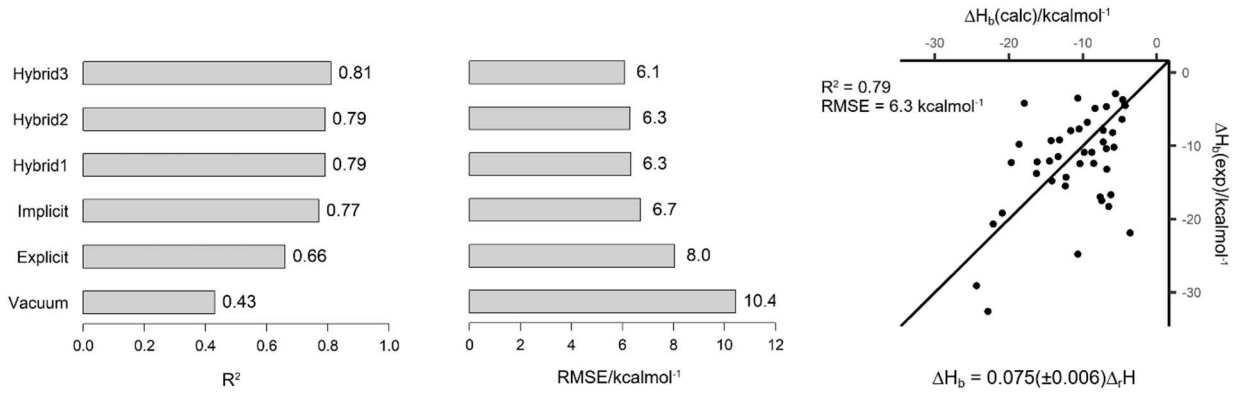

Fig. S2 Comparison of the performance of hydration models used for single point calculation of binding enthalpy. R^2 (left) and RMSE (middle) values of linear regressions (β=0) between calculated and experimental binding enthalpies are shown for the full (N=43) data set. On the right correlation plots between calculated and experimental binding enthalpies obtained are also shown for the best model (Hybrid 1). Regression equations with values α(±Δα) are displayed below.

Tables

Table S1 Target-ligand systems

System (PDB)	Resolution [Å]	Target	Ligand	MW (ligand)	#Waters[f]	ΔH_b(exp) [kcal/mol]	ΔG_b(exp) [kcal/mol]
3ptb	1.7	beta trypsin	benzamidine	121.2	7	-4.5[c,1]	-6.4[1]
3ptb_pme[a]	1.7	beta trypsin	p-methyl-benzamidine	135.2	6	-4.4[c,1]	-6.6[1]
3ptb_pam[a]	1.7	beta trypsin	p-amino-benzamidine	136.2	7	-6.4[c,1]	-7.0[1]
3ptb_pmo[a]	1.7	beta trypsin	p-methoxy-benzamidine	151.2	7	-3.7[c,1]	-6.0[1]
3ptb_pad[a,g]	1.7	beta trypsin	p-amidino-benzamidine	164.2	10	-2.9[c,1]	-5.7[1]
1dpp[g]	3.2	dipeptide bonding protein	GL	188.2	7	-10.41[2]	-8.08[2]
1k1l	2.5	bovine trypsin	NAP-piperazine	467.6	15	-7.9[3]	-9.8[3]
1k1i	2.2	bovine trypsin	NAP-D-pipecholinic acid	508.6	15	-10.9[3]	-8.7[3]
1k1m	2.2	bovine trypsin	NAP-4-acetyl-piperazine	508.6	16	-8.2[3]	-10.2[3]
1k1j	2.2	bovine trypsin	NAP-isopipecholinic acid methyl ester	523.6	16	-9.5[3]	-10.4[3]
1tp5[h]	1.54	PSD-95 PDZ3 domain	KKETWV	790.9	33	-3.5[4]	-7.6[4]
1hcs	NMR	human pp60c-src SH2 domain	Ac-pY-EEIE-NH_2	797.7	34	-6.8[5]	-10.1[5]
1e4x[h]	1.9	anti-Tgf alpha antibody fab fragment	VVSHFND	815.9	27	-24.8[b,6]	-10.3[6]
1jyr	1.55	Grb2 SH2 domain	Ac-S-pY-VNVQ-NH_2	827.8	26	-7.94[7]	-9.12[7]
1py1	2.6	ADP-ribosylation factor binding GGA1 protein	ADDI-pS-LLK	951.1	33	-10.9[c,8]	-6.1[c,8]
1bbz[h]	1.65	Abl tyrosine kinase SH3 domain	Ac-APSYSPPPPP-NH_2	1050.2	20	-21.9[b,9]	-7.7[9]
1abo[h]	2	Abl tyrosine kinase SH3 domain	Ac-APTMPPPLPP-NH_2	1058.3	31	-16.7[b,9]	-6.1[9]
2e7l	2.5	cytotoxic T cell receptor	QLSPFPFDL	1062.2	30	-13.8[10]	-9.0[10]
2oi9[h]	2.35	cytotoxic T cell receptor	QLSPFPFDL	1062.2	33	-4.2[10]	-7.6[10]
3e3q	2.95	cytotoxic T cell receptor	QLSPFPFDL	1062.2	32	-12.3[10]	-8.7[10]
2x6m	1.62	camelid antibody	GYQDYEPEA	1068.1	21	-12.44[11]	-9.16[11]
4j9g[h]	1.8	Abl tyrosine kinase SH3 domain	Ac-APTYPPPPPP-NH_2	1074.2	27	-18.3[b,9]	-7.6[9]
4j9i[h]	2.2	Abl tyrosine kinase SH3 domain	Ac-APTYSPPLPP-NH_2	1080.2	31	-17.5[b,9]	-6.9[9]
1rlq	NMR	C-src tyrosine kinase SH3 domain	RALPPLPRY	1084.3	27	-10.2[12]	-6.4[12]
4j9f[h]	1.09	Abl tyrosine kinase SH3 domain	Ac-APTYPPPLPP-NH_2	1090.3	32	-17.0[b,9]	-6.9[9]
1sem[g]	2.2	SEM-5	Ace-PPPVPPRRR-NH_2	1115.4	27	-4.9[13]	-4.7[13]
2ke1	NMR	autoimmun regulator	ARTKQTARKS	1150.3	27	-9.2[d,14]	-7.0[d,14]
1mw4	NMR	Grb7 SH2 domain	PQPE-pY-VNQPD	1262.2	33	-4.66[15]	-7.70[15]
2gsi[g]	2.81	monoclonal antibody 260.33.12	TKHPKKGVEKY	1317.6	33	-7.7[b,16]	-8.9[16]
2qbx[g]	2.3	Ephrine B receptor 4	SNEWIQPRLPQH	1504.7	43	-9.8[17]	-7.1[17]
1lcj	1.8	P56-lck tyrosine kinase SH2 domain	Ac-EPQ-pY-EEIPIYL-NH_2	1509.5	36	-9.3[18]	-9.6[18]
2bba	1.65	Ephrine B receptor 4	NYLFSPNGPIARAW	1606.8	27	-15.5[19]	-9.8[19]

System (PDB)	Resolution [Å]	Target	Ligand	MW (ligand)	#Waters[f]	ΔH_b(exp) [kcal/mol]	ΔG_b(exp) [kcal/mol]
3ask	2.90	Ubiquitin-like containing plant homeodomain and RING finger domains 1	ARTKQTAR-Kme3-STGGKA	1607.9	53	-12.1[20]	-8.6[20]
3ql9	0.93	Trascriptional regulator ATRX	ARTKQTAR-Kme3-STGGKA	1607.9	45	-12.22[21]	-8.60[21]
2v8c[h]	1.98	mouse profilin 2A	GPPPPPGPPPPPGPPPPPGL	1816.1	70	-12.4[c,22]	-6.4[c,22]
2v8f[h]	1.1	mouse profilin 2A	IPPPPPLPGVASIPPPPPLPG	2007.4	60	-13.2[c,22]	-6.6[c,22]
1jgn	NMR	poli(A)-binding protein PABC domain	VVKSNLNPNAKEFVPGVKYGNI	2389.8	48	-14.8[23]	-8.81[23]
1jh4	NMR	poli(A)-binding protein PABC domain	VLMSKLSVNAPEFYPSGYSSSY	2426.7	71	-11.5[23]	-7.9[23]
1axc[h]	2.6	PCNA antibody	GRKRRQTSMTDFYHSKRRLIFS	2777.2	68	-29.1[c24]	-9.8[c24]
2rod	NMR	mieloid leukemia differentiation protein	AELPPEFAAQLRKIGDKVYCTWSAPD	2905.3	75	-20.7[25]	-10.10[25]
2jm6[h]	NMR	mieloid leukemia differentiation protein	PADLKDECAQLRRIGDKVNLRQKLLN	3009.5	65	-32.6[25]	-9.41[25]
2roc[g]	NMR	mieloid leukemia differentiation protein	EEEWAREIGAQLRRIADDLNAQYERR	3186.4	65	-14.3[25]	-12.50[25]
2nl9[g]	1.55	mieloid leukemia differentiation protein	DMRPEIWIAQELRRIGDEFNAYYARR	3269.7	55	-19.2[25]	-11.84[25]

[a]Systems were built with computational modeling from parent system 3ptb modifying benzamidine ligand[26]

[b]Enthalpies corrected for protonation effects with buffer.

[c]Measurement at 303.15 K. Default is 298.15 K.

[d]Measurement at 296.15 K. Default is 298.15 K.

[e]Enthalpies converted from 298.25 K using heat capacities.

[f]Number of calculated interfacial water molecules.

[g]System not included in the tailored set (ΔG_b model **Fig. 4a**)

[h]System not included in the tailored set. (ΔH_b model **Fig. 3a**)

Table S2 Charges of fragmented complexes and their target-ligand constituents separately

System	Target	Ligand	Complex
1abo	-3	0	-3
1axc	-8	6	-2
1bbz	-2	0	-2
1dpp	-1	0	-1
1e4x	-2	-1	-3
1hcs	3	-5	-2
1jgn	-1	2	1
1jh4	-2	0	-2
1jyr	1	-2	-1
1k1i	-1	0	-1
1k1j	1	1	2
1k1l	0	2	2
1k1m	1	1	2
1lcj	3	-5	-2
1mw4	1	-4	-3
1py1	4	-3	1
1rlq	-3	2	-1
1sem	-3	3	0
1tp5	0	1	1
2bba	2	1	3
2e7l	3	-1	2
2gsi	0	3	3
2jm6	-1	2	1
2ke1	-5	4	-1
2nl9	3	0	3
2oi9	2	-1	1
2qbx	-1	0	-1
2roc	1	-2	-1
2rod	1	-1	0
2v8c	-1	0	-1
2v8f	1	0	1
2x6m	2	-3	-1
3ask	-10	5	-5
3e3q	2	-1	1
3ptb	0	1	1
3ptb_pad	-1	1	0
3ptb_pam	0	1	1
3ptb_pme	-1	1	0
3ptb_pmo	0	1	1
3ql9	-8	5	-3
4j9f	-3	0	-3
4j9g	-3	0	-3
4j9i	-3	0	-3

Table S3 Correlation matrix (R^2 values) for shortlisted descriptors[a], both their parent and daughter (NA normalized) versions. ΔG_b and $\Delta_r H$ values are also included for comparison.

	NHA[b]/NA	NHB[c]/NB	MW[d]/NA	NHA[b]	NHB	MW[c]	ΔG_b	$\Delta_r H$
NHA[b]/NA	1.00	0.93	0.91	0.03	0.03	0.03	0.14	0.01
NHB[c]/NB	0.93	1.00	0.82	0.02	0.02	0.02	0.10	0.02
MW[d]/NA	0.91	0.82	1.00	0.01	0.00	0.00	0.21	0.00
NHA[b]	0.03	0.02	0.01	1.00	1.00	1.00	0.21	0.51
NHB	0.03	0.02	0.00	1.00	1.00	1.00	0.20	0.49
MW[c]	0.03	0.02	0.00	1.00	1.00	1.00	0.21	0.51
ΔG_b	0.14	0.10	0.21	0.21	0.20	0.21	1.00	0.23
$\Delta_r H$	0.01	0.02	0.00	0.51	0.49	0.51	0.23	1.00

[a]Descriptors are detailed in Yap et al (2011)[27].
[b]NHA stands for number of heavy atoms.
[c]NHB stands for number bonds between heavy atoms.
[d]MW stands for molecular weight.

Table S4 Regression parameters and residuals for the different binding enthalpy models assuming β=0 (N=31, tailored set)

System	Vacuum	Explicit	Implicit	Hybrid 1	Hybrid 2	Hybrid 3
	ε					
1dpp	-7.17	-7.26	-0.82	-4.51	-4.89	-5.69
1hcs	3.06	6.15	-6.43	1.32	1.63	2.68
1jgn	-9.90	-5.18	-10.56	-2.51	-2.77	-3.18
1jh4	-6.77	-0.46	-3.82	0.05	0.69	1.97
1jyr	-2.06	-0.77	0.26	2.13	1.14	1.03
1k1i	-8.59	-7.19	-1.57	-3.30	-2.44	-2.98
1k1j	-7.44	-6.48	1.00	-3.24	-2.63	-2.97
1k1l	-5.66	-4.39	2.26	-1.63	-0.61	-1.09
1k1m	-6.31	-5.30	0.38	-3.03	-2.58	-2.48
1lcj	3.05	3.71	2.63	3.07	1.65	1.78
1mw4	0.69	4.32	-7.22	1.23	1.40	1.65
1py1	-1.63	-0.50	-2.63	-2.40	-2.52	-2.92
1rlq	-6.37	-4.29	-4.44	-5.18	-2.65	-2.45
1sem	0.02	1.48	2.72	2.34	2.78	2.61
2bba	-10.17	-7.97	-5.13	-4.81	-4.24	-4.97
2e7l	-7.33	-4.25	-1.34	0.30	2.45	0.90
2gsi	-1.94	0.43	3.29	1.40	1.13	1.35
2ke1	3.93	3.32	1.32	2.19	2.14	0.75
2nl9	-23.03	-9.05	-1.05	-1.09	-2.83	-3.18
2qbx	-5.22	-0.85	5.58	6.32	5.62	5.72
2roc	-11.43	-5.40	-6.23	-3.68	-3.77	-2.06
2rod	-14.98	-7.59	-9.02	-1.53	-2.63	-1.66
2x6m	-4.77	-4.69	-4.25	-3.45	-5.89	-6.24
3ask	10.20	10.20	-1.16	0.45	0.21	1.75
3e3q	-5.20	-2.07	5.10	4.73	5.06	4.23
3ptb	-2.74	-2.33	2.13	-0.81	0.21	-0.07
3ptb_pad	-0.36	0.20	4.88	1.92	3.34	2.88
3ptb_pam	-4.63	-4.21	0.03	-2.33	-0.92	-1.75
3ptb_pme	-2.10	-2.18	2.91	-0.66	-0.09	-0.56
3ptb_pmo	-1.76	-1.57	3.59	0.28	1.40	0.69
3ql9	6.71	8.26	0.39	1.80	1.71	2.34
R^2	0.51	0.78	0.85	0.93	0.90	0.93
R^2(cv)	0.43	0.73	0.82	0.92	0.89	0.92
RMSE	7.59	5.09	4.24	2.84	2.84	2.91
α	1.38E-02	9.50E-03	1.31E-01	6.50E-02	4.81E-02	4.10E-02
Δα	2.47E-03	9.24E-04	1.02E-02	3.23E-03	2.38E-03	2.09E-03
t_α	5.58E+00	1.03E+01	1.29E+01	2.02E+01	2.02E+01	1.97E+01
p_α	4.59E-06	2.36E-11	9.59E-14	5.41E-19	5.21E-19	1.08E-18

Table S5 The construction of the tailored set of the ΔH_b model

A systematic exclusion of data points was performed on the full set of points to avoid a bias of the least squares fit algorithm by possible outliers. The outliers were excluded one-by-one, and in each step a simple linear regression ($\beta \neq 0$) was calculated according to **Eq. 2** and the data point with the largest residual was considered as an outlier and excluded in the next round.

N_o[a]	System[b]	R^2	RMSE	β	Δβ	t(β)	p(β)	α_1	$\Delta\alpha_1$	t(α_1)	p(α_1)
	Hybrid1										
0	-	0.24	5.81	-5.55E+00	2.03E+00	-2.73E+00	9.18E-03	4.47E-02	1.24E-02	3.61E+00	8.21E-04
1	1bbz	0.32	5.43	-4.24E+00	1.96E+00	-2.17E+00	3.62E-02	5.12E-02	1.18E-02	4.33E+00	9.70E-05
2	1e4x	0.36	5.08	-3.86E+00	1.84E+00	-2.10E+00	4.24E-02	5.16E-02	1.11E-02	4.66E+00	3.59E-05
3	2jm6	0.28	4.64	-5.06E+00	1.73E+00	-2.93E+00	5.75E-03	4.11E-02	1.07E-02	3.83E+00	4.70E-04
4	1axc	0.16	4.29	-6.48E+00	1.69E+00	-3.84E+00	4.62E-04	2.90E-02	1.09E-02	2.66E+00	1.15E-02
5	4j9g	0.21	4.07	-5.78E+00	1.63E+00	-3.55E+00	1.09E-03	3.22E-02	1.04E-02	3.09E+00	3.85E-03
6	2oi9	0.29	3.80	-5.18E+00	1.54E+00	-3.36E+00	1.89E-03	3.83E-02	1.01E-02	3.81E+00	5.41E-04
7	4j9i	0.35	3.57	-4.61E+00	1.47E+00	-3.14E+00	3.51E-03	4.07E-02	9.51E-03	4.28E+00	1.45E-04
8	1abo	0.43	3.29	-3.85E+00	1.39E+00	-2.78E+00	8.86E-03	4.43E-02	8.88E-03	4.99E+00	1.91E-05
9	4j9f	0.51	2.98	-3.24E+00	1.27E+00	-2.55E+00	1.58E-02	4.68E-02	8.08E-03	5.79E+00	2.01E-06
10	1tp5	0.55	2.80	-3.44E+00	1.20E+00	-2.86E+00	7.43E-03	4.67E-02	7.62E-03	6.13E+00	8.37E-07
11	2v8f	0.60	2.66	-2.96E+00	1.17E+00	-2.53E+00	1.68E-02	4.89E-02	7.32E-03	6.68E+00	2.10E-07
12	2v8c	0.62	2.61	-2.70E+00	1.16E+00	-2.33E+00	2.70E-02	4.98E-02	7.20E-03	6.91E+00	1.35E-07
	Hybrid3										
0	-	0.27	5.72	-4.77E+00	2.10E+00	-2.28E+00	2.80E-02	3.11E-02	8.04E-03	3.86E+00	3.91E-04
1	2jm6	0.21	5.28	-5.78E+00	1.97E+00	-2.93E+00	5.56E-03	2.52E-02	7.72E-03	3.27E+00	2.24E-03
2	1e4x	0.25	4.89	-5.33E+00	1.83E+00	-2.90E+00	6.04E-03	2.57E-02	7.16E-03	3.60E+00	8.97E-04
3	1bbz	0.35	4.42	-4.12E+00	1.70E+00	-2.42E+00	2.03E-02	2.94E-02	6.57E-03	4.48E+00	6.73E-05
4	2oi9	0.42	4.13	-3.72E+00	1.60E+00	-2.33E+00	2.54E-02	3.23E-02	6.25E-03	5.17E+00	8.41E-06
5	4j9g	0.48	3.88	-3.12E+00	1.52E+00	-2.05E+00	4.81E-02	3.38E-02	5.91E-03	5.72E+00	1.67E-06
6	4j9i	0.52	3.67	-2.64E+00	1.46E+00	-1.81E+00	7.93E-02	3.49E-02	5.62E-03	6.20E+00	4.17E-07
7	4j9f	0.57	3.47	-2.19E+00	1.39E+00	-1.57E+00	1.26E-01	3.59E-02	5.33E-03	6.73E+00	9.96E-08
8	1abo	0.62	3.23	-1.70E+00	1.31E+00	-1.29E+00	2.06E-01	3.69E-02	4.98E-03	7.41E+00	1.63E-08
9	1axc	0.49	3.04	-3.00E+00	1.36E+00	-2.21E+00	3.46E-02	3.04E-02	5.47E-03	5.56E+00	3.86E-06
10	1tp5	0.51	2.93	-3.29E+00	1.32E+00	-2.49E+00	1.82E-02	2.99E-02	5.29E-03	5.65E+00	3.36E-06
11	2v8f	0.54	2.84	-2.93E+00	1.30E+00	-2.25E+00	3.20E-02	3.08E-02	5.17E-03	5.96E+00	1.56E-06
12	2v8c	0.58	2.74	-2.46E+00	1.28E+00	-1.91E+00	6.57E-02	3.21E-02	5.05E-03	6.37E+00	5.83E-07

[a]Number of points left out from full set.

[b]The system that was left out in the actual step.

Table S6 Measurement methods of protein concentration for ITC experiments

PDB code	Measurement method	References
3ptb; 3ptb_pme; 3ptb_pam; 3ptb_pmo; 3ptb_pad; 1dpp; 1hcs; 1tp5; 1e4x; 1bbz; 1abo; 1jyr; 4j9f; 4j9g; 4j9i; 2x6m; 2ke1; 1mw4; 3ask; 2qbx; 2bba; 3ql9; 2rod; 2jm6; 2nl9; 2roc	UV spectrophotometry with calculated or previously measured extinction coefficient.	Talhout et al. (2001)[1]; Payne et al. (2000)[2]; Charifson et al. (1997)[5]; Kaufmann et al. (2011)[4]; Hahn et al. (2001)[6]; Palencia et al. (2004)[9]; McNemar et al. (1997)[7]; De Genst et al. (2010)[11]; Org et al. (2008)[14]; Ivancic et al. (2005)[15]; Arita et al. (2012)[20]; Chrencik et al. (2006)[19]; Iwase et al. (2011)[21]; Day et al. (2008)[25]
1rlq	Fitted from ligand concentration and known complex stoichiometry.	Wang et al. (2001)[12]
1k1i; 1k1j; 1k1l; 1k1m; 1py1; 2e7l; 3e3q; 2oi9; 1sem; 2gsi; 1lcj; 2v8c; 2v8f; 1jgn; 1jh4; 1axc	Not specified in the article.	Dullweber et al. (2001)[3]; He et al. (2003)[8]; Jones et al. (2008)[10]; Ferreon et al. (2003)[13]; Armstrong et al. (2013)[16]; Chung et al. (1998)[18]; Kursula et al. (2008)[22]; Kozlov et al. (2004)[23]; Bruning et al. (2004)[24]

Table S7 Regression parameters and residuals for the different binding enthalpy models assuming β=0 (N=43, full set)

System	Vacuum	Explicit	Implicit	Hybrid 1	Hybrid 2	Hybrid 3
			ε			
1abo	-15.06	-11.44	-11.08	-10.52	-9.17	-8.20
1axc	-8.82	-3.60	-3.46	-4.74	-2.75	-1.60
1bbz	-20.63	-18.66	-17.01	-18.31	-17.58	-16.75
1dpp	-6.36	-6.61	1.00	-3.60	-4.03	-4.93
1e4x	-23.00	-19.26	-16.00	-14.14	-14.60	-14.65
1hcs	5.51	8.85	-6.36	2.57	2.94	4.21
1jgn	-8.68	-3.18	-9.76	-0.63	-0.89	-1.30
1jh4	-5.59	1.83	-2.37	1.82	2.60	4.14
1jyr	-0.59	0.72	1.81	3.67	2.56	2.48
1k1i	-8.01	-6.41	0.20	-2.13	-1.12	-1.70
1k1j	-6.93	-5.85	2.99	-2.28	-1.56	-1.92
1k1l	-5.10	-3.66	4.18	-0.67	0.53	0.01
1k1m	-5.83	-4.70	2.01	-2.24	-1.70	-1.56
1lcj	6.12	6.42	4.89	4.96	3.36	3.57
1mw4	2.02	6.19	-7.71	2.14	2.34	2.67
1py1	0.67	1.66	-1.06	-1.10	-1.22	-1.63
1rlq	-5.42	-3.07	-3.34	-4.41	-1.47	-1.20
1sem	1.24	2.81	4.16	3.45	3.98	3.82
1tp5	2.92	5.88	3.77	7.16	5.88	5.80
2bba	-8.85	-6.41	-3.16	-3.17	-2.48	-3.27
2e7l	-5.73	-2.27	1.02	2.46	4.98	3.27
2gsi	-0.51	2.12	5.37	2.79	2.51	2.82
2jm6	-23.25	-16.71	-11.31	-9.79	-12.95	-12.16
2ke1	7.20	5.93	3.31	3.94	3.91	2.35
2nl9	-23.99	-6.94	2.39	1.68	-0.28	-0.59
2oi9	4.01	7.09	11.33	13.69	14.35	12.86
2qbx	-4.08	1.01	8.49	8.79	8.03	8.22
2roc	-10.72	-3.55	-4.71	-2.05	-2.13	-0.09
2rod	-13.56	-4.87	-6.81	1.41	0.18	1.41
2v8c	-10.23	-7.61	-1.96	-3.88	-6.10	-5.34
2v8f	-11.42	-8.81	-3.40	-6.46	-5.96	-5.14
2x6m	-2.87	-3.08	-2.69	-2.07	-4.87	-5.24
3ask	15.73	14.84	0.91	2.38	2.12	3.98
3e3q	-3.44	0.05	8.40	7.34	7.77	6.89
3ptb	-2.30	-1.88	3.38	-0.25	0.95	0.65
3ptb_pad	0.27	0.84	6.35	2.65	4.31	3.81
3ptb_pam	-4.19	-3.76	1.25	-1.71	-0.07	-1.00
3ptb_pme	-1.53	-1.72	4.29	-0.09	0.58	0.06
3ptb_pmo	-1.27	-1.12	4.97	0.89	2.19	1.40
3ql9	11.41	12.51	2.78	3.95	3.88	4.69
4j9f	-14.91	-11.75	-9.41	-9.37	-8.77	-8.33
4j9g	-16.38	-13.73	-10.76	-11.82	-10.63	-10.68
4j9i	-15.42	-12.41	-9.42	-10.07	-9.44	-9.07
R^2	0.43	0.66	0.77	0.79	0.79	0.81

System	Vacuum	Explicit	Implicit	Hybrid 1	Hybrid 2	Hybrid 3
R^2(cv)	0.31	0.58	0.71	0.75	0.75	0.78
RMSE	10.44	8.03	6.70	6.32	6.30	6.07
α	1.72E-02	1.15E-02	1.56E-01	7.50E-02	5.56E-02	4.76E-02
Δα	3.06E-03	1.26E-03	1.34E-02	5.95E-03	4.39E-03	3.59E-03
t_α	5.62E+00	9.08E+00	1.17E+01	1.26E+01	1.27E+01	1.33E+01
p_α	1.42E-06	1.83E-11	8.47E-15	7.44E-16	6.36E-16	1.30E-16

Table S8 Regression parameters and residuals for the different binding enthalpy models assuming β≠0 (N=43, full set)

	Vacuum	Implicit	Explicit	Hybrid 1	Hybrid 2	Hybrid 3
	ε					
1abo	-4.90	-5.97	-7.34	-7.47	-6.82	-6.38
1axc	-16.17	-11.72	-9.30	-9.03	-7.76	-6.40
1bbz	-10.12	-11.84	-12.92	-14.21	-13.97	-13.76
1dpp	1.54	-0.16	1.97	-0.80	-1.23	-2.06
1e4x	-12.99	-13.98	-13.78	-12.90	-13.29	-13.41
1hcs	5.65	7.34	-0.15	4.33	4.43	5.15
1jgn	-2.73	-1.98	-5.75	-0.81	-1.04	-1.22
1jh4	0.56	1.88	-0.31	1.99	2.38	3.47
1jyr	4.21	3.90	3.57	4.53	3.75	3.63
1k1i	0.98	-0.43	1.31	-0.13	0.35	-0.13
1k1j	2.36	0.70	3.44	0.35	0.63	0.22
1k1l	3.97	2.49	4.83	1.96	2.53	2.03
1k1m	3.64	1.95	3.55	0.90	1.05	0.91
1lcj	3.33	4.86	4.53	4.75	3.70	3.87
1mw4	7.45	7.90	0.17	4.94	4.90	4.89
1py1	1.50	2.22	0.66	0.49	0.29	-0.08
1rlq	1.79	1.14	-0.20	-1.20	0.41	0.44
1sem	7.17	6.63	6.25	5.63	5.80	5.56
1tp5	8.59	8.58	6.72	8.40	7.51	7.34
2bba	-3.40	-3.51	-2.64	-2.60	-2.28	-2.75
2e7l	-1.61	-1.01	0.35	1.44	2.93	2.10
2gsi	4.44	4.52	5.54	4.10	3.81	3.93
2jm6	-20.33	-18.38	-15.07	-13.45	-15.34	-14.50
2ke1	3.49	4.77	3.75	4.18	4.08	3.11
2nl9	-7.79	-6.17	-1.51	-1.20	-2.39	-2.29
2oi9	8.00	8.51	10.33	12.02	12.39	11.70
2qbx	2.25	2.75	6.17	6.83	6.35	6.73
2roc	-2.38	-1.77	-2.87	-1.45	-1.59	-0.26
2rod	-8.57	-6.50	-7.03	-1.97	-2.69	-1.51
2v8c	-0.57	-1.83	-0.53	-1.77	-3.27	-3.02
2v8f	-1.39	-2.76	-1.66	-3.63	-3.50	-3.17
2x6m	-0.16	-0.37	-0.93	-0.71	-2.53	-2.97
3ask	1.29	5.75	1.11	2.08	1.86	3.16
3e3q	-0.06	0.76	4.93	4.96	5.21	4.99
3ptb	7.33	5.36	6.04	3.58	4.11	3.63
3ptb_pad	8.99	7.33	8.35	5.96	6.79	6.25
3ptb_pam	5.43	3.46	4.01	1.94	2.75	1.90
3ptb_pme	7.47	5.47	6.56	3.72	3.93	3.28
3ptb_pmo	8.15	6.14	7.25	4.59	5.19	4.40
3ql9	0.91	4.91	2.03	2.97	2.88	3.58
4j9f	-5.17	-6.28	-6.61	-6.91	-6.69	-6.57
4j9g	-6.48	-7.80	-7.94	-8.89	-8.33	-8.55
4j9i	-5.67	-6.83	-6.86	-7.53	-7.30	-7.23
R^2	0.00	0.09	0.17	0.24	0.22	0.27

	Vacuum	Implicit	Explicit	Hybrid 1	Hybrid 2	Hybrid 3
R^2(cv)	0.00	0.03	0.09	0.16	0.14	0.19
RMSE	6.66	6.36	6.08	5.81	5.90	5.72
α	1.04E-03	3.78E-03	8.16E-02	4.47E-02	3.38E-02	3.11E-02
$\Delta\alpha$	2.88E-03	1.87E-03	2.82E-02	1.24E-02	9.95E-03	8.04E-03
t_α	3.62E-01	2.02E+00	2.89E+00	3.61E+00	3.40E+00	3.86E+00
p_α	7.19E-01	4.96E-02	6.05E-03	8.21E-04	1.52E-03	3.91E-04
β	-1.17E+01	-8.99E+00	-6.42E+00	-5.55E+00	-5.30E+00	-4.77E+00
$\Delta\beta$	1.51E+00	1.83E+00	2.18E+00	2.03E+00	2.20E+00	2.10E+00
t_β	-7.73E+00	-4.92E+00	-2.95E+00	-2.73E+00	-2.41E+00	-2.28E+00
p_β	1.59E-09	1.45E-05	5.25E-03	9.18E-03	2.07E-02	2.80E-02

Table S9 Regression parameters and residuals for the different binding enthalpy models assuming β≠0 (N=31, tailored set)

System	Vacuum	Implicit	Explicit	Hybrid 1	Hybrid 2	Hybrid 3
			ε			
1dpp	-0.64	-2.48	-0.18	-3.19	-3.86	-4.25
1hcs	3.50	5.11	-1.39	2.12	2.14	3.09
1jgn	-4.90	-4.24	-7.37	-2.70	-2.90	-3.23
1jh4	-1.61	-0.37	-2.27	0.04	0.53	1.51
1jyr	2.04	1.63	1.56	2.46	1.54	1.55
1k1i	-1.21	-2.74	-0.81	-2.38	-1.94	-2.24
1k1j	0.17	-1.62	1.21	-2.01	-1.84	-1.93
1k1l	1.79	0.18	2.63	-0.40	0.11	-0.11
1k1m	1.46	-0.37	1.50	-1.54	-1.57	-1.26
1lcj	1.20	2.64	2.15	2.87	1.71	1.84
1mw4	5.28	5.64	-0.78	2.55	2.33	2.74
1py1	-0.65	-0.02	-1.36	-1.69	-2.00	-2.19
1rlq	-0.39	-1.15	-1.97	-3.66	-1.98	-1.67
1sem	5.00	4.35	4.30	3.34	3.43	3.44
2bba	-5.57	-5.79	-4.86	-4.62	-4.23	-4.79
2e7l	-3.77	-3.27	-2.07	-0.31	1.57	0.18
2gsi	2.27	2.25	3.26	1.96	1.57	1.85
2ke1	1.36	2.54	1.52	2.22	2.13	1.05
2nl9	-10.00	-8.43	-4.49	-2.64	-3.74	-4.18
2qbx	0.07	0.49	3.46	5.24	4.89	4.82
2roc	-4.56	-4.03	-4.87	-3.47	-3.63	-2.25
2rod	-10.73	-8.73	-9.38	-3.33	-3.84	-3.32
2x6m	-2.32	-2.64	-2.94	-2.86	-5.04	-5.12
3ask	-0.80	3.60	-1.16	0.21	0.03	1.21
3e3q	-2.22	-1.49	2.01	3.44	3.98	3.11
3ptb	5.15	3.04	4.18	1.02	1.39	1.43
3ptb_pad	6.81	5.01	6.38	3.49	4.24	4.09
3ptb_pam	3.25	1.15	2.18	-0.59	0.11	-0.30
3ptb_pme	5.29	3.16	4.64	1.16	1.16	1.07
3ptb_pmo	5.96	3.82	5.33	2.05	2.50	2.20
3ql9	-1.19	2.74	-0.41	1.21	1.24	1.65
R^2	0.01	0.23	0.25	0.62	0.58	0.58
R^2(cv)	0.00	0.17	0.18	0.55	0.53	0.53
RMSE	4.22	3.73	3.66	2.61	2.73	2.74
α	1.11E-03	3.86E-03	6.87E-02	4.98E-02	3.95E-02	3.21E-02
Δα	2.11E-03	1.32E-03	2.19E-02	7.20E-03	6.18E-03	5.05E-03
t_α	5.26E-01	2.91E+00	3.14E+00	6.91E+00	6.39E+00	6.37E+00
p_α	6.03E-01	6.80E-03	3.84E-03	1.35E-07	5.48E-07	5.83E-07
β	-9.51E+00	-6.66E+00	-5.22E+00	-2.70E+00	-2.01E+00	-2.46E+00
Δβ	1.18E+00	1.33E+00	1.66E+00	1.16E+00	1.34E+00	1.28E+00
t_β	-8.05E+00	-5.00E+00	-3.15E+00	-2.33E+00	-1.50E+00	-1.91E+00
p_β	7.09E-09	2.53E-05	3.80E-03	2.70E-02	1.45E-01	6.57E-02

Table S10 Comparison of the statistical parameters obtained by 1SCF vs. QM-minimization (model Hybrid 3, β≠0). The comparison of the results based on $\Delta_r H$ values from 1SCF and QM minimization shows no improvement for the QM-minimized models over the 1SCF ones neither for the full set (**Row 0**), nor for the final, tailored set (**Row 12**).

	1SCF			Minimization		
N_o[a]	System[b]	R^2	RMSE	System	R^2	RMSE
0	**-**	**0.27**	**5.72**	**-**	**0.25**	**5.80**
1	2jm6	0.21	5.28	2jm6	0.24	5.17
2	1e4x	0.25	4.89	1e4x	0.28	4.78
3	1bbz	0.35	4.42	1bbz	0.35	4.41
4	2oi9	0.42	4.13	1axc	0.24	4.09
5	4j9g	0.48	3.88	4j9g	0.29	3.86
6	4j9i	0.52	3.67	4j9i	0.37	3.57
7	4j9f	0.57	3.47	4j9f	0.39	3.42
8	1abo	0.62	3.23	1abo	0.43	3.25
9	1axc	0.49	3.04	2oi9	0.46	3.14
10	1tp5	0.51	2.93	1tp5	0.48	3.00
11	2v8f	0.54	2.84	1mw4	0.51	2.88
12	**2v8c**	**0.58**	**2.74**	**1sem**	**0.54**	**2.76**

[a]Number of points left out from full set.
[b]The system that was left out in the actual step.

Table S11 Comparison of linear regression results for different semi-empirical parametric methods for model 1SCF Hybrid 3

	PM7		PM6-D3H4X		PM6-DH2X	
	Full set	**Tailored set**	**Full set**	**Tailored set**	**Full set**	**Tailored set**
$\beta \neq 0$						
R^2	0.27	0.58	0.21	0.53	0.22	0.47
RMSE	5.72	2.74	5.94	2.92	5.90	3.09
α	3.11E-02	3.21E-02	3.77E-02	4.39E-02	4.24E-02	4.70E-02
$\Delta\alpha$	8.04E-03	5.05E-03	1.15E-02	7.76E-03	1.25E-02	9.29E-03
β	-4.77E+00	-2.46E+00	-5.77E+00	-2.58E+00	-5.48E+00	-2.61E+00
$\Delta\beta$	2.10E+00	1.28E+00	2.15E+00	1.41E+00	2.16E+00	1.56E+00
$\beta = 0$						
R^2	0.81	0.93	0.78	0.92	0.79	0.91
RMSE	6.07	2.91	6.45	3.09	6.35	3.24
α	4.76E-02	4.10E-02	6.57E-02	5.70E-02	7.11E-02	6.14E-02
$\Delta\alpha$	3.59E-03	2.09E-03	5.35E-03	3.09E-03	5.67E-03	3.51E-03

Table S12 Ligand-based descriptors[a] generated with PaDEL program package. For all descriptors a normalized version was calculated by division with the number of atoms (NA).

Descriptor type	Count	Descriptor(s)
Acidic group count	1	nAcid
ALOGP	3	ALogP, ALogp2, AMR
APol	1	apol
Aromatic atoms count	1	naAromAtom
Aromatic bonds count	1	nAromBond
Atom count	14	nAtom, nHeavyAtom, nH, nB, nC, nN, nO, nS, nP, nF, nCl, nBr, nI, nX
Autocorrelation	346	ATS0m, ATS1m, ATS2m, ATS3m, ATS4m, ATS5m, ATS6m, ATS7m, ATS8m, ATS0v, ATS1v, ATS2v, ATS3v, ATS4v, ATS5v, ATS6v, ATS7v, ATS8v, ATS0e, ATS1e, ATS2e, ATS3e, ATS4e, ATS5e, ATS6e, ATS7e, ATS8e, ATS0p, ATS1p, ATS2p, ATS3p, ATS4p, ATS5p, ATS6p, ATS7p, ATS8p, ATS0i, ATS1i, ATS2i, ATS3i, ATS4i, ATS5i, ATS6i, ATS7i, ATS8i, ATS0s, ATS1s, ATS2s, ATS3s, ATS4s, ATS5s, ATS6s, ATS7s, ATS8s, AATS0m, AATS1m, AATS2m, AATS3m, AATS4m, AATS5m, AATS6m, AATS7m, AATS8m, AATS0v, AATS1v, AATS2v, AATS3v, AATS4v, AATS5v, AATS6v, AATS7v, AATS8v, AATS0e, AATS1e, AATS2e, AATS3e, AATS4e, AATS5e, AATS6e, AATS7e, AATS8e, AATS0p, AATS1p, AATS2p, AATS3p, AATS4p, AATS5p, AATS6p, AATS7p, AATS8p, AATS0i, AATS1i, AATS2i, AATS3i, AATS4i, AATS5i, AATS6i, AATS7i, AATS8i, AATS0s, AATS1s, AATS2s, AATS3s, AATS4s, AATS5s, AATS6s, AATS7s, AATS8s, ATSC0c, ATSC1c, ATSC2c, ATSC3c, ATSC4c, ATSC5c, ATSC6c, ATSC7c, ATSC8c, ATSC0m, ATSC1m, ATSC2m, ATSC3m, ATSC4m, ATSC5m, ATSC6m, ATSC7m, ATSC8m, ATSC0v, ATSC1v, ATSC2v, ATSC3v, ATSC4v, ATSC5v, ATSC6v, ATSC7v, ATSC8v, ATSC0e, ATSC1e, ATSC2e, ATSC3e, ATSC4e, ATSC5e, ATSC6e, ATSC7e, ATSC8e, ATSC0p, ATSC1p, ATSC2p, ATSC3p, ATSC4p, ATSC5p, ATSC6p, ATSC7p, ATSC8p, ATSC0i, ATSC1i, ATSC2i, ATSC3i, ATSC4i, ATSC5i, ATSC6i, ATSC7i, ATSC8i, ATSC0s, ATSC1s, ATSC2s, ATSC3s, ATSC4s, ATSC5s, ATSC6s, ATSC7s, ATSC8s, AATSC0c, AATSC1c, AATSC2c, AATSC3c, AATSC4c, AATSC5c, AATSC6c, AATSC7c, AATSC8c, AATSC0m, AATSC1m, AATSC2m, AATSC3m, AATSC4m, AATSC5m, AATSC6m, AATSC7m, AATSC8m, AATSC0v, AATSC1v, AATSC2v, AATSC3v, AATSC4v, AATSC5v, AATSC6v, AATSC7v, AATSC8v, AATSC0e, AATSC1e, AATSC2e, AATSC3e, AATSC4e, AATSC5e, AATSC6e, AATSC7e, AATSC8e, AATSC0p, AATSC1p, AATSC2p, AATSC3p, AATSC4p, AATSC5p, AATSC6p, AATSC7p, AATSC8p, AATSC0i, AATSC1i, AATSC2i, AATSC3i, AATSC4i, AATSC5i, AATSC6i, AATSC7i, AATSC8i, AATSC0s, AATSC1s, AATSC2s, AATSC3s, AATSC4s, AATSC5s, AATSC6s, AATSC7s, AATSC8s, MATS1c, MATS2c, MATS3c, MATS4c, MATS5c, MATS6c, MATS7c, MATS8c, MATS1m, MATS2m, MATS3m, MATS4m, MATS5m, MATS6m, MATS7m, MATS8m, MATS1v, MATS2v, MATS3v, MATS4v, MATS5v, MATS6v, MATS7v, MATS8v, MATS1e, MATS2e, MATS3e, MATS4e, MATS5e, MATS6e, MATS7e, MATS8e, MATS1p, MATS2p, MATS3p, MATS4p, MATS5p, MATS6p, MATS7p, MATS8p, MATS1i, MATS2i, MATS3i, MATS4i, MATS5i, MATS6i, MATS7i, MATS8i, MATS1s, MATS2s, MATS3s, MATS4s, MATS5s, MATS6s, MATS7s, MATS8s, GATS1c, GATS2c, GATS3c, GATS4c, GATS5c, GATS6c, GATS7c, GATS8c, GATS1m, GATS2m, GATS3m, GATS4m, GATS5m, GATS6m, GATS7m, GATS8m, GATS1v, GATS2v, GATS3v,

Descriptor type	Count	Descriptor(s)
		GATS4v, GATS5v, GATS6v, GATS7v, GATS8v, GATS1e, GATS2e, GATS3e, GATS4e, GATS5e, GATS6e, GATS7e, GATS8e, GATS1p, GATS2p, GATS3p, GATS4p, GATS5p, GATS6p, GATS7p, GATS8p, GATS1i, GATS2i, GATS3i, GATS4i, GATS5i, GATS6i, GATS7i, GATS8i, GATS1s, GATS2s, GATS3s, GATS4s, GATS5s, GATS6s, GATS7s, GATS8s
Barysz matrix	91	SpAbs_DzZ, SpMax_DzZ, SpDiam_DzZ, SpAD_DzZ, SpMAD_DzZ, EE_DzZ, SM1_DzZ, VE1_DzZ, VE2_DzZ, VE3_DzZ, VR1_DzZ, VR2_DzZ, VR3_DzZ, SpAbs_Dzm, SpMax_Dzm, SpDiam_Dzm, SpAD_Dzm, SpMAD_Dzm, EE_Dzm, SM1_Dzm, VE1_Dzm, VE2_Dzm, VE3_Dzm, VR1_Dzm, VR2_Dzm, VR3_Dzm, SpAbs_Dzv, SpMax_Dzv, SpDiam_Dzv, SpAD_Dzv, SpMAD_Dzv, EE_Dzv, SM1_Dzv, VE1_Dzv, VE2_Dzv, VE3_Dzv, VR1_Dzv, VR2_Dzv, VR3_Dzv, SpAbs_Dze, SpMax_Dze, SpDiam_Dze, SpAD_Dze, SpMAD_Dze, EE_Dze, SM1_Dze, VE1_Dze, VE2_Dze, VE3_Dze, VR1_Dze, VR2_Dze, VR3_Dze, SpAbs_Dzp, SpMax_Dzp, SpDiam_Dzp, SpAD_Dzp, SpMAD_Dzp, EE_Dzp, SM1_Dzp, VE1_Dzp, VE2_Dzp, VE3_Dzp, VR1_Dzp, VR2_Dzp, VR3_Dzp, SpAbs_Dzi, SpMax_Dzi, SpDiam_Dzi, SpAD_Dzi, SpMAD_Dzi, EE_Dzi, SM1_Dzi, VE1_Dzi, VE2_Dzi, VE3_Dzi, VR1_Dzi, VR2_Dzi, VR3_Dzi, SpAbs_Dzs, SpMax_Dzs, SpDiam_Dzs, SpAD_Dzs, SpMAD_Dzs, EE_Dzs, SM1_Dzs, VE1_Dzs, VE2_Dzs, VE3_Dzs, VR1_Dzs, VR2_Dzs, VR3_Dzs
Basic group count	1	nBase
BCUT	6	BCUTw-1l, BCUTw-1h, BCUTc-1l, BCUTc-1h, BCUTp-1l, BCUTp-1h
Bond count	10	nBonds, nBonds2, nBondsS, nBondsS2, nBondsS3, nBondsD, nBondsD2, nBondsT, nBondsQ, nBondsM
BPol	1	bpol
Burden modified eigenvalues	96	SpMax1_Bhm, SpMax2_Bhm, SpMax3_Bhm, SpMax4_Bhm, SpMax5_Bhm, SpMax6_Bhm, SpMax7_Bhm, SpMax8_Bhm, SpMin1_Bhm, SpMin2_Bhm, SpMin3_Bhm, SpMin4_Bhm, SpMin5_Bhm, SpMin6_Bhm, SpMin7_Bhm, SpMin8_Bhm, SpMax1_Bhv, SpMax2_Bhv, SpMax3_Bhv, SpMax4_Bhv, SpMax5_Bhv, SpMax6_Bhv, SpMax7_Bhv, SpMax8_Bhv, SpMin1_Bhv, SpMin2_Bhv, SpMin3_Bhv, SpMin4_Bhv, SpMin5_Bhv, SpMin6_Bhv, SpMin7_Bhv, SpMin8_Bhv, SpMax1_Bhe, SpMax2_Bhe, SpMax3_Bhe, SpMax4_Bhe, SpMax5_Bhe, SpMax6_Bhe, SpMax7_Bhe, SpMax8_Bhe, SpMin1_Bhe, SpMin2_Bhe, SpMin3_Bhe, SpMin4_Bhe, SpMin5_Bhe, SpMin6_Bhe, SpMin7_Bhe, SpMin8_Bhe, SpMax1_Bhp, SpMax2_Bhp, SpMax3_Bhp, SpMax4_Bhp, SpMax5_Bhp, SpMax6_Bhp, SpMax7_Bhp, SpMax8_Bhp, SpMin1_Bhp, SpMin2_Bhp, SpMin3_Bhp, SpMin4_Bhp, SpMin5_Bhp, SpMin6_Bhp, SpMin7_Bhp, SpMin8_Bhp, SpMax1_Bhi, SpMax2_Bhi, SpMax3_Bhi, SpMax4_Bhi, SpMax5_Bhi, SpMax6_Bhi, SpMax7_Bhi, SpMax8_Bhi, SpMin1_Bhi, SpMin2_Bhi, SpMin3_Bhi, SpMin4_Bhi, SpMin5_Bhi, SpMin6_Bhi, SpMin7_Bhi, SpMin8_Bhi, SpMax1_Bhs, SpMax2_Bhs, SpMax3_Bhs, SpMax4_Bhs, SpMax5_Bhs, SpMax6_Bhs, SpMax7_Bhs, SpMax8_Bhs, SpMin1_Bhs, SpMin2_Bhs, SpMin3_Bhs, SpMin4_Bhs, SpMin5_Bhs, SpMin6_Bhs, SpMin7_Bhs, SpMin8_Bhs
Carbon types	9	C1SP1, C2SP1, C1SP2, C2SP2, C3SP2, C1SP3, C2SP3, C3SP3, C4SP3
Chi chain	10	SCH-3, SCH-4, SCH-5, SCH-6, SCH-7, VCH-3, VCH-4, VCH-5, VCH-6, VCH-7

Descriptor type	Count	Descriptor(s)
Chi cluster	8	SC-3, SC-4, SC-5, SC-6, VC-3, VC-4, VC-5, VC-6
Chi path cluster	6	SPC-4, SPC-5, SPC-6, VPC-4, VPC-5, VPC-6
Chi path	32	SP-0, SP-1, SP-2, SP-3, SP-4, SP-5, SP-6, SP-7, ASP-0, ASP-1, ASP-2, ASP-3, ASP-4, ASP-5, ASP-6, ASP-7, VP-0, VP-1, VP-2, VP-3, VP-4, VP-5, VP-6, VP-7, AVP-0, AVP-1, AVP-2, AVP-3, AVP-4, AVP-5, AVP-6, AVP-7
Constitutional	12	Sv, Sse, Spe, Sare, Sp, Si, Mv, Mse, Mpe, Mare, Mp, Mi
Crippen logP and MR	2	CrippenLogP, CrippenMR
Detour matrix	11	SpMax_Dt, SpDiam_Dt, SpAD_Dt, SpMAD_Dt, EE_Dt, VE1_Dt, VE2_Dt, VE3_Dt, VR1_Dt, VR2_Dt, VR3_Dt
Eccentric connectivity index	1	ECCEN
Atom type electrotopological state	489	nHBd, nwHBd, nHBa, nwHBa, nHBint2, nHBint3, nHBint4, nHBint5, nHBint6, nHBint7, nHBint8, nHBint9, nHBint10, nHsOH, nHdNH, nHsSH, nHsNH2, nHssNH, nHaaNH, nHsNH3p, nHssNH2p, nHsssNHp, nHtCH, nHdCH2, nHdsCH, nHaaCH, nHCHnX, nHCsats, nHCsatu, nHAvin, nHother, nHmisc, nsLi, nssBe, nssssBem, nsBH2, nssBH, nsssB, nssssBm, nsCH3, ndCH2, nssCH2, ntCH, ndsCH, naaCH, nsssCH, nddC, ntsC, ndssC, naasC, naaaC, nssssC, nsNH3p, nsNH2, nssNH2p, ndNH, nssNH, naaNH, ntN, nsssNHp, ndsN, naaN, nsssN, nddsN, naasN, nssssNp, nsOH, ndO, nssO, naaO, naOm, nsOm, nsF, nsSiH3, nssSiH2, nsssSiH, nssssSi, nsPH2, nssPH, nsssP, ndsssP, nddsP, nssssssP, nsSH, ndS, nssS, naaS, ndssS, nddssS, nsssssssS, nSm, nsCl, nsGeH3, nssGeH2, nsssGeH, nssssGe, nsAsH2, nssAsH, nsssAs, ndsssAs, nddsAs, nssssssAs, nsSeH, ndSe, nssSe, naaSe, ndssSe, nsssssssSe, nddssSe, nsBr, nsSnH3, nssSnH2, nsssSnH, nssssSn, nsI, nsPbH3, nssPbH2, nsssPbH, nssssPb, SHBd, SwHBd, SHBa, SwHBa, SHBint2, SHBint3, SHBint4, SHBint5, SHBint6, SHBint7, SHBint8, SHBint9, SHBint10, SHsOH, SHdNH, SHsSH, SHsNH2, SHssNH, SHaaNH, SHsNH3p, SHssNH2p, SHsssNHp, SHtCH, SHdCH2, SHdsCH, SHaaCH, SHCHnX, SHCsats, SHCsatu, SHAvin, SHother, SHmisc, SsLi, SssBe, SssssBem, SsBH2, SssBH, SsssB, SssssBm, SsCH3, SdCH2, SssCH2, StCH, SdsCH, SaaCH, SssssCH, SddC, StsC, SdssC, SaasC, SaaaC, SssssC, SsNH3p, SsNH2, SssNH2p, SdNH, SssNH, SaaNH, StN, SsssNHp, SdsN, SaaN, SsssN, SddsN, SaasN, SssssNp, SsOH, SdO, SssO, SaaO, SaOm, SsOm, SsF, SsSiH3, SssSiH2, SssssSiH, SssssSi, SsPH2, SssPH, SsssP, SdsssP, SddsP, SssssssP, SsSH, SdS, SssS, SaaS, SdssS, SddssS, SssssssS, SSm, SsCl, SsGeH3, SssGeH2, SsssGeH, SssssGe, SsAsH2, SssAsH, SsssAs, SdsssAs, SddsAs, SssssssAs, SsSeH, SdSe, SssSe, SaaSe, SdssSe, SssssssSe, SddssSe, SsBr, SsSnH3, SssSnH2, SsssSnH, SssssSn, SsI, SsPbH3, SssPbH2, SsssPbH, SssssPb, minHBd, minwHBd, minHBa, minwHBa, minHBint2, minHBint3, minHBint4, minHBint5, minHBint6, minHBint7, minHBint8, minHBint9, minHBint10, minHsOH, minHdNH, minHsSH, minHsNH2, minHssNH, minHaaNH, minHsNH3p, minHssNH2p, minHsssNHp, minHtCH, minHdCH2, minHdsCH, minHaaCH, minHCHnX, minHCsats, minHCsatu, minHAvin, minHother, minHmisc, minsLi, minssBe, minssssBem, minsBH2, minssBH, minsssB, minssssBm, minsCH3, mindCH2, minssCH2, mintCH, mindsCH, minaaCH, minsssCH, minddC, mintsC, mindssC, minaasC, minaaaC, minssssC, minsNH3p, minsNH2, minssNH2p, mindNH, minssNH, minaaNH, mintN, minsssNHp, mindsN, minaaN, minsssN, minddsN, minaasN, minssssNp, minsOH, mindO, minssO, minaaO, minaOm, minsOm, minsF, minsSiH3, minssSiH2, minsssSiH, minssssSi, minsPH2, minssPH, minsssP, mindsssP, minddsP, minsssssP, minsSH, mindS,

Descriptor type	Count	Descriptor(s)
		minssS, minaaS, mindssS, minddssS, minsssssS, minSm, minsCl, minsGeH3, minssGeH2, minsssGeH, minssssGe, minsAsH2, minssAsH, minsssAs, mindsssAs, minddsAs, minsssssAs, minsSeH, mindSe, minssSe, minaaSe, mindssSe, minsssssSe, minddssSe, minsBr, minsSnH3, minssSnH2, minsssSnH, minssssSn, minsI, minsPbH3, minssPbH2, minsssPbH, minssssPb, maxHBd, maxwHBd, maxHBa, maxwHBa, maxHBint2, maxHBint3, maxHBint4, maxHBint5, maxHBint6, maxHBint7, maxHBint8, maxHBint9, maxHBint10, maxHsOH, maxHdNH, maxHsSH, maxHsNH2, maxHssNH, maxHaaNH, maxHsNH3p, maxHssNH2p, maxHsssNHp, maxHtCH, maxHdCH2, maxHdsCH, maxHaaCH, maxHCHnX, maxHCsats, maxHCsatu, maxHAvin, maxHother, maxHmisc, maxsLi, maxssBe, maxssssBem, maxsBH2, maxssBH, maxsssB, maxssssBm, maxsCH3, maxdCH2, maxssCH2, maxtCH, maxdsCH, maxaaCH, maxsssCH, maxddC, maxtsC, maxdssC, maxaasC, maxaaaC, maxssssC, maxsNH3p, maxsNH2, maxssNH2p, maxdNH, maxssNH, maxaaNH, maxtN, maxsssNHp, maxdsN, maxaaN, maxsssN, maxddsN, maxaasN, maxssssNp, maxsOH, maxdO, maxssO, maxaaO, maxaaOm, maxsOm, maxsF, maxsSiH3, maxssSiH2, maxsssSiH, maxssssSi, maxsPH2, maxssPH, maxsssP, maxdsssP, maxddsP, maxssssssP, maxsSH, maxdS, maxssS, maxaaS, maxdssS, maxddssS, maxsssssssS, maxSm, maxsCl, maxsGeH3, maxssGeH2, maxsssGeH, maxssssGe, maxsAsH2, maxssAsH, maxsssAs, maxdsssAs, maxddsAs, maxsssssAs, maxsSeH, maxdSe, maxssSe, maxaaSe, maxdssSe, maxsssssssSe, maxddssSe, maxsBr, maxsSnH3, maxssSnH2, maxsssSnH, maxssssSn, maxsI, maxsPbH3, maxssPbH2, maxsssPbH, maxssssPb, sumI, meanI, hmax, gmax, hmin, gmin, LipoaffinityIndex, MAXDN, MAXDP, DELS, MAXDN2, MAXDP2, DELS2
Extended topochemical atom	43	ETA_Alpha, ETA_AlphaP, ETA_dAlpha_A, ETA_dAlpha_B, ETA_Epsilon_1, ETA_Epsilon_2, ETA_Epsilon_3, ETA_Epsilon_4, ETA_Epsilon_5, ETA_dEpsilon_A, ETA_dEpsilon_B, ETA_dEpsilon_C, ETA_dEpsilon_D, ETA_Psi_1, ETA_dPsi_A, ETA_dPsi_B, ETA_Shape_P, ETA_Shape_Y, ETA_Shape_X, ETA_Beta, ETA_BetaP, ETA_Beta_s, ETA_BetaP_s, ETA_Beta_ns, ETA_BetaP_ns, ETA_dBeta, ETA_dBetaP, ETA_Beta_ns_d, ETA_BetaP_ns_d, ETA_Eta, ETA_EtaP, ETA_Eta_R, ETA_Eta_F, ETA_EtaP_F, ETA_Eta_L, ETA_EtaP_L, ETA_Eta_R_L, ETA_Eta_F_L, ETA_EtaP_F_L, ETA_Eta_B, ETA_EtaP_B, ETA_Eta_B_RC, ETA_EtaP_B_RC
FMFDescriptor	1	FMF
Fragment complexity	1	fragC
Hbond acceptor count	4	nHBAcc, nHBAcc2, nHBAcc3, nHBAcc_Lipinski
Hbond donor count	2	nHBDon, nHBDon_Lipinski
Hybridization ratio	1	HybRatio
Information content	42	IC0, IC1, IC2, IC3, IC4, IC5, TIC0, TIC1, TIC2, TIC3, TIC4, TIC5, SIC0, SIC1, SIC2, SIC3, SIC4, SIC5, CIC0, CIC1, CIC2, CIC3, CIC4, CIC5, BIC0, BIC1, BIC2, BIC3, BIC4, BIC5, MIC0, MIC1, MIC2, MIC3, MIC4, MIC5, ZMIC0, ZMIC1, ZMIC2, ZMIC3, ZMIC4, ZMIC5
Kappa shape indices	3	Kier1, Kier2, Kier3
Largest chain	1	nAtomLC
Largest Pi system	1	nAtomP
Longest aliphatic chain	1	nAtomLAC

Descriptor type	Count	Descriptor(s)
Mannhold LogP	1	MLogP
McGowan volume	1	McGowan_Volume
Molecular distance edge	19	MDEC-11, MDEC-12, MDEC-13, MDEC-14, MDEC-22, MDEC-23, MDEC-24, MDEC-33, MDEC-34, MDEC-44, MDEO-11, MDEO-12, MDEO-22, MDEN-11, MDEN-12, MDEN-13, MDEN-22, MDEN-23, MDEN-33
Molecular linear free energy relation	6	MLFER_A, MLFER_BH, MLFER_BO, MLFER_S, MLFER_E, MLFER_L
Path counts	22	MPC2, MPC3, MPC4, MPC5, MPC6, MPC7, MPC8, MPC9, MPC10, TPC, piPC1, piPC2, piPC3, piPC4, piPC5, piPC6, piPC7, piPC8, piPC9, piPC10, TpiPC, R_TpiPCTPC
Petitjean number	1	PetitjeanNumber
Ring count	68	nRing, n3Ring, n4Ring, n5Ring, n6Ring, n7Ring, n8Ring, n9Ring, n10Ring, n11Ring, n12Ring, nG12Ring, nFRing, nF4Ring, nF5Ring, nF6Ring, nF7Ring, nF8Ring, nF9Ring, nF10Ring, nF11Ring, nF12Ring, nFG12Ring, nHeteroRing, n3HeteroRing, n4HeteroRing, n5HeteroRing, n6HeteroRing, n7HeteroRing, n8HeteroRing, n9HeteroRing, n10HeteroRing, n11HeteroRing, n12HeteroRing, nG12HeteroRing, nFHeteroRing, nF4HeteroRing, nF5HeteroRing, nF6HeteroRing, nF7HeteroRing, nF8HeteroRing, nF9HeteroRing, nF10HeteroRing, nF11HeteroRing, nF12HeteroRing, nFG12HeteroRing, nTHeteroRing, nT4HeteroRing, nT5HeteroRing, nT6HeteroRing, nT7HeteroRing, nT8HeteroRing, nT9HeteroRing, nT10HeteroRing, nT11HeteroRing, nT12HeteroRing, nTG12HeteroRing
Rotatable bonds count	4	nRotB, RotBFrac, nRotBt, RotBtFrac
Rule of five	1	LipinskiFailures
Topological	3	topoRadius, topoDiameter, topoShape
Topological charge	21	GGI1, GGI2, GGI3, GGI4, GGI5, GGI6, GGI7, GGI8, GGI9, GGI10, JGI1, JGI2, JGI3, JGI4, JGI5, JGI6, JGI7, JGI8, JGI9, JGI10, JGT
Topological distance matrix	11	SpMax_D, SpDiam_D, SpAD_D, SpMAD_D, EE_D, VE1_D, VE2_D, VE3_D, VR1_D, VR2_D, VR3_D
Topological polar surface area	1	TopoPSA
Van der Waals volume	1	VABC
Vertex adjacency information (magnitude)	1	vAdjMat
Walk counts	20	MWC2, MWC3, MWC4, MWC5, MWC6, MWC7, MWC8, MWC9, MWC10, TWC, SRW2, SRW3, SRW4, SRW5, SRW6, SRW7, SRW8, SRW9, SRW10, TSRW
Weight	2	MW, AMW
Weighted path	5	WTPT-1, WTPT-2, WTPT-3, WTPT-4, WTPT-5
Wiener numbers	2	WPATH, WPOL
XLogP	1	XLogP
Zagreb index	1	Zagreb

[a]Descriptors are detailed in Yap et al (2011)[27].

Table S13 Correlation of filtered descriptors with binding free energy (ΔG_b) and calculated reaction enthalpy ($\Delta_r H$) expressed in R^2 values

Descriptor	Correlation with ΔG_b (R^2)	Correlation with $\Delta_r H$ (R^2)
ATSC3e	0.28	2.51E-02
AATS0m	0.25	1.85E-02
maxHdsCH	0.24	2.27E-02
maxHother	0.24	2.27E-02
ATSC4m	0.23	8.47E-03
AATSC0m	0.22	2.11E-02
nAtomP	0.22	1.49E-03
MW/NA	**0.21**	**5.16E-06**
minHdsCH	0.19	5.56E-03
minHother	0.19	5.56E-03
AATS1m	0.19	2.52E-06
AATS3m	0.17	3.72E-03
GATS5c	0.16	4.80E-03
AATS2m	0.16	3.23E-05
VE1_Dzm	0.15	3.17E-02
VE1_DzZ	0.15	3.17E-02
AATSC4e	0.15	3.94E-02
NHA/NA	**0.14**	**1.12E-02**
Mv	0.14	5.60E-04
AATSC2e	0.13	2.85E-02
GATS1m	0.13	3.67E-02
AATS4p	0.12	9.23E-05
AATS0v	0.12	6.04E-03
AATS4v	0.11	2.83E-02
AATS1v	0.11	1.17E-02
NHB/NB	**0.10**	**2.15E-02**
BIC5	0.10	8.53E-03
AATS6i	0.10	3.78E-02
SIC5	0.10	3.97E-04
nHeteroRing	0.09	2.01E-02
ATSC7i	0.09	1.31E-02

Table S14 Multiple linear regression results for ΔG_b models with calculated $\Delta_r H$ (α_1 in Eq. 3) and selected ligand-based descriptors (α_2) NHA/NA or NHB/NB. A systematic procedure (see **Table S5**) was applied to identify outliers and form tailored sets of the remaining points. Parameters of the final tailored sets featured in the main text are highlighted in boldface.

N_o[a]	System[b]	R^2	RMSE	α_1	$\Delta\alpha_1$	$t(\alpha_1)$	α_2	$\Delta\alpha_2$	$t(\alpha_2)$	β	$\Delta\beta$	$t(\beta)$
β≠0							NHA[c]/NA					
0	-	0.41	1.30	1.21E-02	2.81E-03	4.30E+00	-2.58E+01	7.30E+00	-3.54E+00	6.63E+00	3.77E+00	1.76E+00
1	2roc	0.46	1.16	1.18E-02	2.53E-03	4.66E+00	-2.49E+01	6.56E+00	-3.79E+00	6.20E+00	3.39E+00	1.83E+00
2	3ptb_pad	0.49	1.10	1.10E-02	2.42E-03	4.57E+00	-2.79E+01	6.35E+00	-4.39E+00	7.54E+00	3.26E+00	2.31E+00
3	2qbx	0.55	1.04	1.22E-02	2.34E-03	5.22E+00	-2.88E+01	6.02E+00	-4.79E+00	8.14E+00	3.09E+00	2.63E+00
4	1dpp	0.61	0.98	1.30E-02	2.24E-03	5.81E+00	-3.35E+01	6.04E+00	-5.54E+00	1.07E+01	3.13E+00	3.41E+00
5	**1sem**	**0.60**	**0.94**	**1.26E-02**	**2.16E-03**	**5.80E+00**	**-3.09E+01**	**5.93E+00**	**-5.22E+00**	**9.25E+00**	**3.08E+00**	**3.01E+00**
6	2oi9	0.65	0.90	1.35E-02	2.11E-03	6.40E+00	-3.15E+01	5.66E+00	-5.57E+00	9.63E+00	2.94E+00	3.28E+00
7	1py1	0.67	0.86	1.33E-02	2.02E-03	6.60E+00	-3.10E+01	5.42E+00	-5.72E+00	9.28E+00	2.82E+00	3.29E+00
8	2nl9	0.65	0.82	1.21E-02	2.04E-03	5.94E+00	-3.03E+01	5.21E+00	-5.81E+00	8.80E+00	2.71E+00	3.24E+00
9	2gsi	0.68	0.79	1.22E-02	1.95E-03	6.28E+00	-3.22E+01	5.09E+00	-6.33E+00	9.83E+00	2.65E+00	3.70E+00
β=0												
0	-	0.97	1.35	1.10E-02	2.82E-03	3.92E+00	-1.31E+01	9.10E-01	-1.44E+01	-	-	-
1	2roc	0.98	1.21	1.08E-02	2.54E-03	4.25E+00	-1.30E+01	8.21E-01	-1.58E+01	-	-	-
2	3ptb_pad	0.98	1.18	9.99E-03	2.50E-03	3.99E+00	-1.33E+01	8.18E-01	-1.63E+01	-	-	-
3	2qbx	0.98	1.13	1.10E-02	2.47E-03	4.46E+00	-1.31E+01	7.94E-01	-1.65E+01	-	-	-
4	1dpp	0.98	1.13	1.12E-02	2.47E-03	4.53E+00	-1.30E+01	8.01E-01	-1.63E+01	-	-	-
5	1sem	0.98	1.06	1.09E-02	2.31E-03	4.71E+00	-1.32E+01	7.53E-01	-1.76E+01	-	-	-
6	2oi9	0.99	1.03	1.17E-02	2.30E-03	5.07E+00	-1.31E+01	7.38E-01	-1.78E+01	-	-	-
7	1py1	0.99	0.99	1.16E-02	2.21E-03	5.23E+00	-1.32E+01	7.12E-01	-1.86E+01	-	-	-
8	2nl9	0.99	0.95	1.03E-02	2.22E-03	4.63E+00	-1.35E+01	6.95E-01	-1.94E+01	-	-	-
9	2gsi	0.99	0.94	1.02E-02	2.21E-03	4.62E+00	-1.35E+01	6.95E-01	-1.94E+01	-	-	-
β≠0							NHB[d]/NB					
0	-	0.38	1.33	1.24E-02	2.90E-03	4.28E+00	-2.08E+01	6.56E+00	-3.17E+00	4.28E+00	3.47E+00	1.24E+00
1	2roc	0.43	1.19	1.21E-02	2.59E-03	4.66E+00	-2.05E+01	5.87E+00	-3.49E+00	4.18E+00	3.10E+00	1.35E+00
2	2qbx	0.49	1.13	1.33E-02	2.54E-03	5.23E+00	-2.15E+01	5.63E+00	-3.82E+00	4.81E+00	2.98E+00	1.61E+00
3	1dpp	0.56	1.06	1.45E-02	2.43E-03	5.98E+00	-2.79E+01	5.85E+00	-4.77E+00	8.34E+00	3.12E+00	2.67E+00
4	3ptb_pad	0.59	1.01	1.39E-02	2.33E-03	5.97E+00	-2.94E+01	5.61E+00	-5.25E+00	8.99E+00	2.98E+00	3.02E+00
5	**1sem**	**0.59**	**0.96**	**1.33E-02**	**2.23E-03**	**5.99E+00**	**-2.73E+01**	**5.41E+00**	**-5.05E+00**	**7.75E+00**	**2.89E+00**	**2.69E+00**
6	2oi9	0.64	0.91	1.43E-02	2.16E-03	6.64E+00	-2.81E+01	5.14E+00	-5.47E+00	8.25E+00	2.75E+00	3.01E+00
7	2v8c	0.66	0.87	1.41E-02	2.08E-03	6.78E+00	-2.85E+01	4.94E+00	-5.76E+00	8.36E+00	2.64E+00	3.17E+00
8	2nl9	0.64	0.84	1.29E-02	2.10E-03	6.12E+00	-2.78E+01	4.77E+00	-5.83E+00	7.88E+00	2.55E+00	3.09E+00
9	2gsi	0.68	0.80	1.31E-02	2.02E-03	6.47E+00	-2.95E+01	4.66E+00	-6.33E+00	8.82E+00	2.49E+00	3.53E+00
β=0												
0	-	0.97	1.35	1.14E-02	2.81E-03	4.07E+00	-1.28E+01	8.94E-01	-1.43E+01	-	-	-
1	2roc	0.98	1.22	1.12E-02	2.53E-03	4.41E+00	-1.27E+01	8.04E-01	-1.57E+01	-	-	-
2	2qbx	0.98	1.17	1.22E-02	2.49E-03	4.88E+00	-1.25E+01	7.81E-01	-1.60E+01	-	-	-
3	1dpp	0.98	1.16	1.24E-02	2.47E-03	5.01E+00	-1.24E+01	7.77E-01	-1.59E+01	-	-	-
4	3ptb_pad	0.98	1.13	1.17E-02	2.44E-03	4.79E+00	-1.27E+01	7.76E-01	-1.63E+01	-	-	-
5	1sem	0.98	1.05	1.13E-02	2.27E-03	4.99E+00	-1.29E+01	7.28E-01	-1.77E+01	-	-	-
6	2oi9	0.99	1.02	1.21E-02	2.25E-03	5.39E+00	-1.28E+01	7.10E-01	-1.80E+01	-	-	-
7	2v8c	0.99	0.99	1.19E-02	2.20E-03	5.39E+00	-1.29E+01	7.00E-01	-1.85E+01	-	-	-
8	2nl9	0.99	0.95	1.06E-02	2.21E-03	4.79E+00	-1.32E+01	6.83E-01	-1.93E+01	-	-	-
9	2gsi	0.99	0.95	1.05E-02	2.20E-03	4.79E+00	-1.31E+01	6.82E-01	-1.93E+01	-	-	-

[a]Number of points left out from full set.
[b]The system that was left out in the actual step.
[c]NHA stands for number of heavy atoms.
[d]NHB stands for number of bonds between heavy atoms.

Table S15 Values of descriptors used in regression models for binding free energy estimation

System	NHA	NHA/NA	NHB	NHB/NB
1abo	74	0.4837	79	0.5000
1axc	195	0.4912	198	0.4950
1bbz	75	0.5137	81	0.5329
1dpp	13	0.4483	12	0.4286
1e4x	58	0.5321	59	0.5364
1hcs	55	0.5670	55	0.5670
1jgn	169	0.4884	172	0.4928
1jh4	171	0.5104	176	0.5176
1jyr	57	0.5327	57	0.5327
1k1i	36	0.5625	39	0.5821
1k1j	37	0.5441	40	0.5634
1k1l	33	0.5323	36	0.5538
1k1m	36	0.5455	39	0.5652
1lcj	106	0.5274	109	0.5343
1mw4	88	0.5500	91	0.5583
1py1	65	0.5000	64	0.4961
1rlq	77	0.4753	80	0.4848
1sem	79	0.4731	83	0.4854
1tp5	56	0.4828	57	0.4872
2bba	115	0.5134	120	0.5240
2e7l	76	0.5101	79	0.5197
2gsi	93	0.4769	95	0.4822
2jm6	210	0.4817	210	0.4817
2ke1	80	0.4678	79	0.4647
2nl9	231	0.5099	236	0.5153
2oi9	76	0.5101	79	0.5197
2qbx	107	0.5144	111	0.5236
2roc	225	0.5137	227	0.5159
2rod	205	0.5062	211	0.5134
2v8c	130	0.5000	144	0.5255
2v8f	143	0.4799	154	0.4984
2x6m	76	0.5630	78	0.5693
3ask	112	0.4667	111	0.4644
3e3q	76	0.5101	79	0.5197
3ptb	9	0.5000	9	0.5000
3ptb_pad	12	0.5455	12	0.5455
3ptb_pam	10	0.5000	10	0.5000
3ptb_pme	10	0.4762	10	0.4762
3ptb_pmo	11	0.5000	11	0.5000
3ql9	112	0.4667	111	0.4644
4j9f	78	0.4968	84	0.5153
4j9g	77	0.5066	84	0.5283
4j9i	77	0.5000	82	0.5157

References

1. Talhout R, Engberts JBFN. Thermodynamic analysis of binding of p-substituted benzamidines to trypsin. European Journal of Biochemistry. 2001;268(6).

2. Payne JW, Grail BM, Gupta S, Ladbury JE, Marshall NJ, O'Brien R, et al. Structural basis for recognition of dipeptides by peptide transporters. Archives of Biochemistry and Biophysics. 2000;384(1).

3. Dullweber F, Stubbs MT, Musil D, Stürzebecher J, Klebe G. Factorising ligand affinity: A Combined thermodynamic and crystallographic study of trypsin and thrombin inhibition. Journal of Molecular Biology. 2001;313(3).

4. Kaufmann K, Shen N, Mizoue L, Meiler J. A physical model for PDZ-domain/peptide interactions. Journal of Molecular Modeling. 2011;17(2).

5. Charifson PS, Shewchuk LM, Rocque W, Hummel CW, Jordan SR, Mohr C, et al. Peptide ligands of pp60(c-src) SH2 domains: A thermodynamic and structural study. Biochemistry. 1997;36(21):6283–93.

6. Hahn M, Winkler D, Welfle K, Misselwitz R, Welfle H, Wessner H, et al. Cross-reactive binding of cyclic peptides to an anti-TGFα antibody Fab fragment: An X-ray structural and thermodynamic analysis. Journal of Molecular Biology. 2001;314(2).

7. McNemar C, Snow ME, Windsor WT, Prongay A, Mui P, Zhang R, et al. Thermodynamic and structural analysis of phosphotyrosine polypeptide binding to Grb2-SH2. Biochemistry. 1997;36(33).

8. He X, Zhu G, Koelsch G, Rodgers KK, Zhang XC, Tang J. Biochemical and Structural Characterization of the Interaction of Memapsin 2 (β-Secretase) Cytosolic Domain with the VHS Domain of GGA Proteins. Biochemistry. 2003;42(42).

9. Palencia A, Cobos ES, Mateo PL, Martínez JC, Luque I. Thermodynamic Dissection of the Binding Energetics of Proline-rich Peptides to the Abl-SH3 Domain: Implications for Rational Ligand Design. Journal of Molecular Biology. 2004;336(2).

10. Jones LL, Colf LA, Bankovich AJ, Stone JD, Gao YG, Choi MC, et al. Different thermodynamic binding mechanisms and peptide fine specificities associated with a panel of structurally similar high-affinity T cell receptors. Biochemistry. 2008;47(47).

11. De Genst EJ, Guilliams T, Wellens J, Day EM, Waudby CA, Meehan S, et al. Structure and properties of a complex of α-synuclein and a single-domain camelid antibody. Journal of Molecular Biology. 2010;402(2).

12. Wang C, Pawley NH, Nicholson LK. The role of backbone motions in ligand binding to the c-Src SH3 domain. Journal of Molecular Biology. 2001;313(4).

13. Ferreon JC, Hilser VJ. Ligand-induced changes in dynamics in the RT loop of the C-terminal SH3 domain of Sem-5 indicate cooperative conformational coupling. Protein Science. 2003;12(5).

14. Org T, Chignola F, Hetényi C, Gaetani M, Rebane A, Liiv I, et al. The autoimmune regulator PHD finger binds to non-methylated histone H3K4 to activate gene expression. EMBO

Reports. 2008;9(4).

15. Ivancic M, Spuches AM, Guth EC, Daugherty MA, Wilcox DE, Lyons BA. Backbone nuclear relaxation characteristics and calorimetric investigation of the human Grb7-SH2/erbB2 peptide complex. Protein science. 2005;14(6):1556–69.

16. Armstrong A, Hildreth JEK, Amzel LM. Structural and thermodynamic insights into the recognition of native proteins by anti-peptide antibodies. Journal of Molecular Biology. 2013;425(11).

17. Chrencik JE, Brooun A, Recht MI, Nicola G, Davis LK, Abagyan R, et al. Three-dimensional structure of the EphB2 receptor in complex with an antagonistic peptide reveals a novel mode of inhibition. Journal of Biological Chemistry. 2007;282(50).

18. Chung E, Henriques D, Renzoni D, Zvelebil M, Bradshaw JM, Waksman G, et al. Mass spectrometric and thermodynamic studies reveal the role of water molecules in complexes formed between SH2 domains and tyrosyl phosphopeptides. Structure. 1998;6(9).

19. Chrencik JE, Brooun A, Kraus ML, Recht MI, Kolatkar AR, Gye WH, et al. Structural and biophysical characterization of the EphB4·EphrinB2 protein-protein interaction and receptor specificity. Journal of Biological Chemistry. 2006;281(38).

20. Arita K, Isogai S, Oda T, Unoki M, Sugita K, Sekiyama N, et al. Recognition of modification status on a histone H3 tail by linked histone reader modules of the epigenetic regulator UHRF1. Proceedings of the National Academy of Sciences of the United States of America. 2012;109(32).

21. Iwase S, Xiang B, Ghosh S, Ren T, Lewis PW, Cochrane JC, et al. ATRX ADD domain links an atypical histone methylation recognition mechanism to human mental-retardation syndrome. Nature Structural and Molecular Biology. 2011;18(7).

22. Kursula P, Kursula I, Massimi M, Song YH, Downer J, Stanley WA, et al. High-resolution Structural Analysis of Mammalian Profilin 2a Complex Formation with Two Physiological Ligands: The Formin Homology 1 Domain of mDia1 and the Proline-rich Domain of VASP. Journal of Molecular Biology. 2008;375(1).

23. Kozlov G, De Crescenzo G, Lim NS, Siddiqui N, Fantus D, Kahvejian A, et al. Structural basis of ligand recognition by PABC, a highly specific peptide-binding domain found in poly(A)-binding protein and a HECT ubiquitin ligase. EMBO Journal. 2004;23(2).

24. Bruning JB, Shamoo Y. Structural and thermodynamic analysis of human PCNA with peptides derived from DNA polymerase-δ p66 subunit and flap endonuclease-1. Structure. 2004;12(12).

25. Day CL, Smits C, Fan FC, Lee EF, Fairlie WD, Hinds MG. Structure of the BH3 Domains from the p53-Inducible BH3-Only Proteins Noxa and Puma in Complex with Mcl-1. Journal of Molecular Biology. 2008;380(5).

26. Horváth I, Jeszenői N, Bálint M, Paragi G, Hetényi C. A fragmenting protocol with explicit hydration for calculation of binding enthalpies of target-ligand complexes at a quantum mechanical level. International Journal of Molecular Sciences. 2019;20(18).

27. Yap CW. PaDEL-descriptor: An open source software to calculate molecular descriptors and fingerprints. Journal of computational chemistry. 2011;32(7):1466–74.