

ClusPro-DC: Dimer Classification by the Cluspro Server for Protein-Protein Docking

Supplemental Information

Content

Supplemental Results and Discussion

Selecting the threshold T for discrimination between biological and crystallographic dimers

Impact of small interface on prediction success rate

Application to Transient Complexes

Supplemental Figures

Table S1. Training set PDB entries

Table S2. Interface area, number of near-native structures, and inferred probability of being a biological dimer for the proteins in the training set

Table S3. Threshold versus the Matthew correlation coefficient (MCC)

Table S4. Classification of training set dimers by three methods

Table S5. Classification of training set monomers by three methods

Table S6. Interface area, number of near-native structures, and inferred probability of being a biological dimer for the proteins in the DC Bio and DC Xtal sets

Table S7. Classification of DC Bio set by three methods

Table S8. Classification of DC Xtal set by three methods

Table S9. Interface area, number of near-native structures, and inferred probability of being a biological dimer for the proteins in the test set

Table S10. Classification of test set dimers by three methods

Table S11. Classification of test set monomers by three methods

Table S12. Interface area, number of near-native structures, and inferred probability of being a biological dimer for the proteins in the difficult set

Table S13. Classification of difficult set dimers by three methods

Table S14. Classification of test set monomers by three methods

Table S15. Small interface area results

Table S16. Interface area, number of near-native structures, and inferred probability of being a biological dimer for the transitional homodimers by Noreen and Thornton [1]

Table S17. Biological activity and classification the transitional homodimers by Noreen and Thornton [1] using three methods

Table S18. Classification of the transitional homodimers by Noreen and Thornton [1] using three methods: Summary

Table S19. Transitional complexes: Interface area, number of near-native structures, and inferred probability of being a biological dimer for the proteins in the weak homodimer set by Dey et al. [2]

Table S20. Interface area, number of near-native structures, and inferred probability of being a biological dimer for the proteins in the “others” homodimer set by Dey et al. [2]

Table S21. Classification of the weak set of dimers by Dey et al. [2]

Table S22. Classification of the “others” set of dimers by by Dey et al. [2]

Table S23. Classification of the weak set of dimers by Dey et al. [2] using three methods: Summary

Table S24. Classification of the “others” set of dimers by Dey et al. [2] using three methods: Summary

Supplemental Figures

Figure S1. Probability of biological dimer for combined training and test sets

Figure S2. Analysis of CAPRI target T70 (Modulator protein MzrA, PDB ID 4pwu).

Figure S3. Results of the analysis of the interaction between chains A and B in CAPRI target T70 (Modulator protein MzrA, PDB ID 4pwu)

Figure S4. Results of the analysis of the interaction between chains A and C in CAPRI target T70

Supplemental References

Supplementary Results and Discussion

Selecting the threshold T for discrimination between biological and crystallographic dimers

As described in the main text and shown in **Fig. 2c**, based on the receiver operating characteristic (ROC) curve $T = 33$ appears to be a reasonable choice for the threshold number of hits to distinguish between crystallographic and biological dimers. In good agreement with this selection, at $N_{nn} = 33$ the probability P of being a biological dimer is between 0.48 and 0.52, depending on the level of smoothing of the probability curve, and we select $P = 0.5$ as the probability threshold between crystallographic and biological dimers. The predicted probability values are not very sensitive to the changes in the training set. **Fig. S1** shows the fraction of biological dimers as a function of N_{nn} and the smoothed probability curve, now based on unified training and difficult test set, thus including 345 rather than 209 proteins. The curve based on the training set (**Fig. 2b**), and the one based on the combination of training and difficult test sets (**Fig. S1**) are very similar, and the probability P of being biological at the threshold $T = 33$ remains 0.5, the value we use for dimer discrimination throughout the paper. As mentioned, the same threshold value also yields the maximum Matthew correlation coefficient (**Table S3**).

Impact of Small Interface on Prediction Success Rate

It is well known that the interface area is a very important factor for the discrimination between crystallographic and biological dimers as crystallographic interfaces are generally smaller ($<1000 \text{ \AA}^2$) than biologically relevant ones. However, some of the biological interfaces are also small, which makes discrimination difficult. Since docking programs are in general less successful in predicting complexes with small interface areas, it was suggested by one of the reviewers that the successful identification of monomers might be a result of the inability of docking to identify correct small interfaces. In order to investigate this potential artifact, we have analyzed ClusPro, PISA, and EPPIC results for several interface-area ranges. The analysis was performed on the test set (783 proteins). Predictions were separately analyzed for interface areas below 600, 800, and 1000 \AA^2 (**Table S15**). Results show that the identification of small interface area biological dimers is indeed difficult, and for $<600 \text{ \AA}^2$ interface the success rate with ClusPro is only 23.5% (4 out of the 17 cases). However, EPPIC and PISA have exactly the same success rate, thus the low success rate is not caused by the failure of docking. In addition, since the overall percentage of biological dimers with such small interface is low (17 out of 783, thus 2.17%), the overall success rate is over 90%, in spite of the inability to correctly identify most of the dimers. For interface areas below 800 and 1000 \AA^2 finding biological dimers is still difficult, and all three methods still produce relatively weak results, with less than 50% success rates in all but one cases. However, ClusPro actually yields better discrimination for biological dimers than the other two methods (**Table S15**), again confirming that the failures are not due to the difficulty of docking.

Application to Transient Complexes

Transient protein–protein complexes dissociate and associate readily, often depending on the physiological condition or environment, and this behavior frequently plays an important role in biological processes. It is expected that determining the state of such complexes under normal conditions is fairly difficult. We applied the three methods studied in this paper to the sets of transient homomultimers collected by Nooren and Thornton [1] and by Dey et al. [2]. As shown in **Table S16**, the average interface area is 855.1 \AA^2 , and only 3 of the 24 complexes by Nooren and Thornton [1] have interface area exceeding 1000 \AA^2 . Based on Cluspro, the average probability of being a dimer is 0.68, thus slightly shifted toward the dimer state. In **Table S17** we investigate the relationship between the biological active form as provided by Nooren and Thornton, and the assignment of multimeric state by ClusPro, EPPIC, and PISA. The function of the first 3 proteins in the Table (Lambda phage cro, corynebacterium diphtheria tox repressor, and mouse nuclear factor κB (NF κB P50 and P65, each with several structures) is DNA binding, and these proteins are active as dimers. In good agreement with this function, ClusPro predicts predominantly dimeric form for all but one (1nfk)

structures. According to EPPIC, the same structure and one more (5cro) are monomers. PISA predicts two NF κ B structures as monomers, and does not reach conclusion in two more cases. In spite of these differences, all three programs reveal that most of the structures are homodimers. The next six proteins in Table S17 (interleukin-8, stromal cell-derived factor 1 α , monocyte chemoattractant protein 1, stem cell factor cytokine, ciliary neurotrophic factor, and insulin) participate in extracellular, signaling, most likely acting in both monomeric and dimeric forms. Accordingly, ClusPro predict dimerization probabilities between 0.12 and 0.7. The biological active forms are somewhat uncertain for the remaining proteins in Table S17, and ClusPro predictions also vary, but in many cases are close to 0.5, reflecting the uncertainty of dimerization state assignment. As shown in **Table S18**, the three methods predict very similar percentages being in dimeric form.

To study the properties of transient homodimers, Dey et al. [2] built a list of 315 PDB entries containing the atomic coordinates of proteins shown by experiment to be homodimers in solution and in the crystal. Then, they identified cases where the literature indicates that the dimer is in equilibrium with the monomer; the 42 proteins shown in **Table S19** constitute a set of dimers marked “weak” to indicate that they dissociate easily. As for the set by Nooren and Thornton [1], the average interface area is small (817.2 \AA^2), and only 7 of the 42 complexes have interface area exceeding 1000 \AA^2 . The average probability of being a biological dimer is 0.54, indicating almost the same frequency of dimeric and monomeric states. In contrast, the remaining 273 “others” type homodimers, listed in Table S20, have average surface area of 1922.8 \AA^2 , thus more than twice as large. The average probability of being a biological dimer is 0.75. As shown in **Tables S23** and **S24**, in the weak set all three methods predicts around 50 of biological dimers, whereas in the “others” set the fraction is about 85%. In both cases ClusPro predicts slightly more dimers than the other two methods.

Table S1. Training set PDB entries									
A. Biological homodimers^a									
12as	1a3c	1a4i	1a4u	1aa7	1ad3	1ade	1af5	1afw	1ajs
1vlb	1amk	1aor	1aq6	1auo	1b3a	1b5e	1b67	1b8a	1b8j
1bam	1bbh	1bd0	1bif	1biq	1bis	1bjw	1bkp	1bmd	1brw
1bsl	1bsr	1buo	1bxg	1bxk	1cdc	1cg2	1chm	1cmb	1cnz
1coz	1csh	1ctt	1cvu	1czj	1daa	1dor	1dpg	1dqs	1dxg
1e98	1ebh	1f13	1fip	1fro	1gvp	1hhp	1hjr	1hss	1hxp
1icw	1imb	1isa	1ivy	1jhg	1jsg	1kba	1kpf	1lyn	1m6p
1mkb	1mor	1nox	1nse	1nsy	1oac	1opy	1pgt	1pre	1qfh
1qhi	1qr2	1r2f	1reg	1rfb	1rpo	1ses	1slt	1smn	1smt
1sox	1tc1	1tox	1trk	1uby	1utg	1vfr	1vok	1wtl	1xso
2arc	2ccy	2hdh	2ilk	2lig	2mcg	2nac	2ohx	2spc	2sqc
2tct	2tgi	3dap	3sdh	3ssi	4kbp	5csm	5rub	8prk	9wga
B. Monomers (crystallographic dimers)^b									
13pk	1a7v	1ad5	1afk	1ag9	1ah7	1ako	1amu	1atl	1aw7
1ayl	1b1j	1b3j	1bc2	1bea	1bin	1bkz	1bs2	1byo	1c02
1caq	1ck7	1cki	1clu	1cqx	1dsu	1dys	1e0s	1ehy	1epa
1ewf	1feh	1fgk	1fjm	1fkd	1fmt	1g2a	1gar	1gjm	1hf8
2j5x	1ilr	1kpt	1kwa	1mpg	1mss	1naw	1np4	1pbg	1pda
1ppo	1qaz	1qci	1qdm	1qha	1qjp	1qme	1qpa	1qtq	1rb3
1rhs	1rne	1shk	1the	1tht	1toa	1ton	1urp	1vbt	1xgs
256b	256l	2acy	2atj	2bc2	2bls	2erc	2g3p	2ihl	2mbr
2scp	2shp	2tps	2ugi	3mht	3pmg	5tss	830c	8pti	

^a Biological homodimers from Baskaran et al. [3]

^b Monomers (crystallographic dimers) from Baskaran et al. [4]

Table S2. Interface area, number of near-native structures, and inferred probability of being a biological dimer for the proteins in the training set^{a,b}

Biological dimers				Monomers (Crystallographic dimers)			
PDB ID	Interface area (Å ²)	Nnn	Probability (dimer)	PDB ID	Interface area (Å ²)	Nnn	Probability (dimer)
12as	1878.1	156	0.84	13pk	1367.4	81	0.77
1a3c	989.1	173	0.85	1a7v	547.8	13	0.29
1a4i	1354.6	114	0.82	1ad5	1053.1	9	0.24
1a4u	2525.1	195	0.89	1afk	818.5	4	0.18
1aa7	1093.8	46	0.60	1ag9	390.2	0	0.12
1ad3	3940.1	39	0.55	1ah7	949.4	21	0.39
1ade	2699.0	107	0.81	1ako	843.3	0	0.12
1af5	851.3	0	0.12	1amu	679.6	0	0.12
1afw	2388.7	123	0.82	1atl	621.2	0	0.12
1ajs	3436.9	74	0.74	1aw7	1145.9	5	0.19
1vlb	1187.2	145	0.83	1ayl	910.2	16	0.33
1amk	1467.8	205	0.93	1b1j	722.8	178	0.86
1aor	1209.9	19	0.36	1b3j	1630.4	102	0.81
1aq6	2220.5	176	0.86	1bc2	640.6	0	0.12
1auo	663.7	0	0.12	1bea	530.1	0	0.12
1b3a	733.0	200	0.91	1bin	553.7	0	0.12
1b5e	2527.4	117	0.82	1bkz	742.8	144	0.83
1b67	1595.9	283	1.00	1bs2	1113.6	31	0.49
1b8a	4375.3	103	0.81	1byo	483.6	0	0.12
1b8j	3790.5	94	0.79	1c02	866.0	66	0.71
1bam	750.0	128	0.83	1caq	643.3	0	0.12
1bbh	765.0	271	1.00	1ck7	1320.9	0	0.12
1bd0	3042.4	98	0.80	1cki	1690.6	133	0.83
1bif	889.7	6	0.20	1clu	749.1	0	0.12
1biq	3012.6	99	0.80	1cqz	1109.5	0	0.12
1bis	1478.5	205	0.93	1dsu	635.4	1	0.13
1bjw	2945.7	70	0.73	1dys	765.3	25	0.43
1bkp	2206.0	142	0.83	1e0s	859.7	0	0.12
1bmd	1590.4	253	1.00	1ehy	696.4	92	0.79
1brw	1060.5	249	1.00	1epa	769.4	0	0.12
1bsl	1893.9	158	0.84	1ewf	947.4	0	0.12
1bsr	1895.4	277	1.00	1feh	1561.1	10	0.26
1buo	1927.6	165	0.84	1fgk	828.5	2	0.15
1bxg	1039.6	71	0.73	1fjm	641.7	0	0.12
1bxk	1276.6	111	0.82	1fkf	629.5	0	0.12
1cdc	3837.9	70	0.73	1fmt	962.3	30	0.48
1cg2	1298.3	73	0.74	1g2a	573.6	0	0.12
1chm	3230.1	70	0.73	1gar	784.8	0	0.12
1cmb	1796.3	481	1.00	1gjm	895.8	0	0.12
1cnz	2444.9	107	0.81	1hf8	1002.5	127	0.83
1coz	1053.1	192	0.89	2j5x	560.3	0	0.12
1csh	4995.4	61	0.69	1ilr	572.2	0	0.12
1ctt	1933.3	134	0.83	1kpt	432.9	9	0.24
1cvu	2444.7	26	0.44	1kwa	599.4	0	0.12
1czj	741.3	7	0.22	1mpg	987.2	31	0.49

1daa	2301.1	181	0.86	1mss	1045.4	73	0.74
1dor	2196.2	130	0.83	1naw	1192.5	16	0.33
1dpg	2278.7	167	0.84	1np4	731.3	1	0.13
1dqs	1604.6	87	0.78	1pbg	771.0	1	0.13
1dxg	709.8	340	1.00	1pda	881.2	4	0.18
1e98	770.4	230	1.00	1ppo	847.2	0	0.12
1ebh	1781.7	115	0.82	1qaz	797.8	0	0.12
1f13	2563.6	57	0.67	1qci	483.6	0	0.12
1fip	1578.0	209	0.94	1qdm	767.8	0	0.12
1fro	3575.0	96	0.80	1qha	1727.2	7	0.22
1gvp	905.1	192	0.89	1qjp	832.1	3	0.16
1hhp	1596.7	281	1.00	1qme	1271.9	15	0.32
1hjr	962.0	83	0.77	1qpa	822.9	32	0.49
1hss	1095.5	64	0.70	1qtq	715.7	11	0.27
1hxp	3398.7	113	0.82	1rb3	614.9	0	0.12
1icw	989.9	137	0.83	1rhs	595.0	0	0.12
1imb	1648.7	185	0.87	1rne	1077.3	101	0.81
1isa	919.0	180	0.86	1shk	893.9	1	0.13
1ivy	1586.2	141	0.83	1the	1122.0	11	0.27
1jhg	2225.9	349	1.00	1tht	1049.3	134	0.83
1jsg	789.3	146	0.83	1toa	1294.4	27	0.45
1kba	599.5	0	0.12	1ton	712.8	21	0.39
1kpf	1855.0	198	0.90	1urp	906.7	0	0.12
1lyn	942.1	39	0.55	1vbt	642.8	3	0.16
1m6p	1017.7	99	0.80	1xgs	1177.2	74	0.74
1mkb	1598.0	124	0.82	256b	520.5	0	0.12
1mor	2528.4	97	0.80	256l	763.1	9	0.24
1nox	2966.1	122	0.82	2acy	541.4	58	0.68
1nse	2674.6	102	0.81	2atj	576.0	0	0.12
1nsy	2583.3	120	0.82	2bc2	648.1	0	0.12
1oac	7135.2	16	0.33	2bls	706.0	8	0.23
1opy	1044.8	65	0.71	2erc	1364.0	1	0.13
1pgt	1228.4	157	0.84	2g3p	1407.3	31	0.49
1pre	2280.7	33	0.50	2ihl	708.5	47	0.61
1qfh	2275.0	55	0.66	2mbr	750.5	0	0.12
1qhi	1698.8	103	0.81	2scp	1115.1	81	0.77
1qr2	1917.0	130	0.83	2shp	591.3	0	0.12
1r2f	1741.9	127	0.83	2tps	805.7	60	0.69
1reg	676.5	0	0.12	2ugi	738.4	2	0.15
1rfb	2636.9	77	0.75	3mht	667.9	173	0.85
1rpo	1398.9	182	0.86	3pmg	899.1	13	0.29
1ses	2229.6	107	0.81	5tss	1369.6	31	0.49
1slt	547.2	134	0.83	830c	1207.8	69	0.73
1smn	868.4	164	0.84	8pti	688.5	7	0.22
1smt	1969.4	278	1.00				
1sox	1410.1	87	0.78				
1tc1	1505.7	92	0.79				
1tox	3703.5	86	0.78				
1trk	4418.3	54	0.65				
1uby	2162.4	135	0.83				
1utg	1464.2	177	0.86				
1vfr	3355.7	73	0.74				
1vok	1582.6	159	0.84				

1wtl	686.4	96	0.80				
1xso	666.8	102	0.81				
2arc	810.3	83	0.77				
2ccy	807.3	175	0.85				
2hdh	1604.4	260	1.00				
2ilk	4479.7	43	0.58				
2lig	1581.1	47	0.61				
2mcg	1631.8	80	0.76				
2nac	3793.4	180	0.86				
2ohx	1685.1	339	1.00				
2spc	2513.5	57	0.67				
2sqc	805.8	0	0.12				
2tct	2612.3	137	0.83				
2tgi	1260.4	145	0.83				
3dap	2653.4	154	0.84				
3sdh	879.3	104	0.81				
3ssi	838.4	115	0.82				
4kbp	1467.9	125	0.83				
5csm	2000.1	38	0.54				
5rub	2829.1	99	0.80				
8prk	968.0	239	1.00				
9wga	2298.4	151	0.83				
Average	1923.7	129.5	0.77	Average	863.7	25.3	0.32
Std Dev	1097.2	82.2	0.20	Std Dev	294.6	42.4	0.26

^a Biological homodimers from Baskaran et al. [3]

^b Monomers (crystallographic dimers) fro Baskaran et al. [4]

Table S3. MCC value as function of the threshold T^a

Threshold T	MCC
20	0.64
21	0.64
22	0.66
23	0.66
24	0.66
25	0.66
26	0.67
27	0.66
28	0.67
29	0.67
30	0.67
31	0.68
32	0.72
33	0.73
34	0.72
35	0.72
36	0.72
37	0.72
38	0.72
39	0.71
40	0.69
41	0.69
42	0.69
43	0.69
44	0.68
45	0.68
46	0.68
47	0.67
48	0.67
49	0.67
50	0.67
51	0.67
52	0.67
53	0.67
54	0.67
55	0.66
56	0.65
57	0.65
58	0.63
59	0.64
60	0.64

^a MCC: Matthews correlation coefficient

Table S4. Classification of training set dimers by three methods^{a, b}

PDB ID	ClusPro probability of dimer	EPPIC	PISA multimer class	PISA status
12as	0.84	bio	dimer	stable
1a3c	0.85	bio	dimer	stable
1a4i	0.82	bio	dimer	stable
1a4u	0.89	bio	dimer	stable
1aa7	0.6	bio	monomer	stable
1ad3	0.55	bio	dimer	stable
1ade	0.81	bio	dimer	stable
1af5	0.12	bio	dimer	stable
1afw	0.82	bio	dimer	stable
1ajs	0.74	bio	dimer	stable
1vlb	0.83	bio	monomer	stable
1amk	0.93	bio	dimer	stable
1aor	0.36	bio	dimer	stable
1aq6	0.86	bio	dimer	stable
1auo	0.12	bio	dimer	stable
1b3a	0.91	bio	dimer	stable
1b5e	0.82	bio	dimer	stable
1b67	1	xtal	dimer	stable
1b8a	0.81	bio	dimer	stable
1b8j	0.79	bio	dimer	stable
1bam	0.83	xtal	dimer	uncertain
1bbh	1	bio	dimer	stable
1bd0	0.8	bio	dimer	stable
1bif	0.2	xtal	monomer	stable
1biq	0.8	bio	dimer	stable
1bis	0.93	bio	dimer	stable
1bjw	0.73	bio	dimer	stable
1bkp	0.83	bio	dimer	stable
1bmd	1	bio	dimer	stable
1brw	1	bio	dimer	stable
1bsl	0.84	bio	dimer	stable
1bsr	1	bio	tetramer	stable
1buo	0.84	bio	dimer	stable
1bxg	0.73	bio	dimer	uncertain
1bxk	0.82	bio	dimer	stable
1cdc	0.73	bio	dimer	stable
1cg2	0.74	bio	tetramer	stable
1chm	0.73	bio	dimer	stable
1cmb	1	bio	dimer	stable
1cnz	0.81	bio	dimer	stable
1coz	0.89	bio	dimer	stable
1csh	0.69	bio	dimer	stable
1ctt	0.83	bio	dimer	stable
1cvu	0.44	bio	trimer	stable
1czj	0.22	bio	dimer	stable
1daa	0.86	bio	dimer	stable
1dor	0.83	bio	dimer	stable

1dpg	0.84	bio	dimer	stable
1dqs	0.78	bio	dimer	stable
1dxg	1	bio	dimer	stable
1e98	1	bio	dimer	stable
1ebh	0.82	bio	dimer	stable
1f13	0.67	bio	dimer	stable
1fip	0.94	bio	tetramer	stable
1fro	0.8	bio	dimer	stable
1gvp	0.89	bio	dimer	stable
1hhp	1	bio	dimer	stable
1hjr	0.77	bio	dimer	stable
1hss	0.7	bio	dimer	stable
1hxp	0.82	bio	dimer	stable
1icw	0.83	bio	dimer	stable
1imb	0.87	xtal	dimer	stable
1isa	0.86	bio	dimer	stable
1ivy	0.83	bio	dimer	stable
1jhg	1	bio	dimer	stable
1jsg	0.83	bio	tetramer	stable
1kba	0.12	bio	dimer	stable
1kpf	0.9	bio	dimer	stable
1lyn	0.55	bio	dimer	stable
1m6p	0.8	bio	dimer	stable
1mkb	0.82	bio	dimer	stable
1mor	0.8	bio	dimer	stable
1nox	0.82	bio	dimer	stable
1nse	0.81	bio	dimer	stable
1nsy	0.82	bio	dimer	stable
1oac	0.33	bio	dimer	stable
1opy	0.71	xtal	dimer	stable
1pgt	0.84	bio	dimer	stable
1pre	0.5	bio	dimer	stable
1qfh	0.66	bio	dimer	stable
1qhi	0.81	bio	dimer	stable
1qr2	0.83	bio	dimer	stable
1r2f	0.83	bio	dimer	stable
1reg	0.12	bio	dimer	stable
1rfb	0.75	bio	dimer	stable
1rpo	0.86	bio	dimer	stable
1ses	0.81	bio	dimer	stable
1slt	0.83	xtal	monomer	stable
1smn	0.84	xtal	dimer	uncertain
1smt	1	bio	dimer	stable
1sox	0.78	bio	dimer	stable
1tc1	0.79	bio	dimer	stable
1tox	0.78	bio	dimer	stable
1trk	0.65	bio	dimer	stable
1uby	0.83	bio	dimer	stable
1utg	0.86	bio	dimer	stable
1vfr	0.74	bio	dimer	stable
1vok	0.84	bio	tetramer	stable
1wtl	0.8	bio	dimer	stable
1xso	0.81	bio	dimer	stable

2arc	0.77	bio	monomer	stable
2ccy	0.85	bio	dimer	stable
2hdh	1	bio	dimer	stable
2ilk	0.58	bio	dimer	stable
2lig	0.61	bio	dimer	stable
2mcg	0.76	bio	dimer	stable
2nac	0.86	bio	dimer	stable
2ohx	1	xtal	dimer	stable
2spc	0.67	bio	dimer	stable
2sqc	0.12	xtal	monomer	stable
2tct	0.83	bio	dimer	stable
2tgi	0.83	bio	dimer	stable
3dap	0.84	bio	dimer	stable
3sdh	0.81	bio	dimer	stable
3ssi	0.82	bio	dimer	stable
4kbp	0.83	bio	tetramer	stable
5csm	0.54	bio	dimer	stable
5rub	0.8	bio	dimer	stable
8prk	1	bio	dimer	stable
9wga	0.83	bio	dimer	stable

^a yellow: incorrect classification

^b biological homodimers from Baskaran et al. [3]

Table S5. Classification of training set monomers by three methods^{a,b}

PDB ID	ClusPro Probability of dimer	EPPIC	PISA multimer class	PISA status
13pk	0.77	xtal	dimer	stable
1a7v	0.29	xtal	monomer	stable
1ad5	0.24	xtal	dimer	uncertain
1afk	0.18	xtal	dimer	uncertain
1ag9	0.12	xtal	dimer	stable
1ah7	0.39	bio	monomer	stable
1ako	0.12	xtal	dimer	stable
1amu	0.12	xtal	monomer	stable
1atl	0.12	xtal	monomer	stable
1aw7	0.19	xtal	monomer	stable
1ayl	0.33	xtal	monomer	stable
1b1j	0.86	xtal	monomer	stable
1b3j	0.81	xtal	dimer	uncertain
1bc2	0.12	bio	monomer	stable
1bea	0.12	xtal	tetramer	uncertain
1bin	0.12	xtal	dimer	uncertain
1bkz	0.83	bio	monomer	stable
1bs2	0.49	xtal	dimer	uncertain
1byo	0.12	xtal	monomer	stable
1c02	0.71	bio	dimer	stable
1caq	0.12	bio	dimer	stable
1ck7	0.12	xtal	dimer	stable
1cki	0.83	xtal	dimer	stable
1clu	0.12	xtal	monomer	stable
1cqx	0.12	bio	dimer	stable
1dsu	0.13	xtal	monomer	stable
1dys	0.43	xtal	monomer	stable
1e0s	0.12	xtal	dimer	stable
1ehy	0.79	bio	tetramer	stable
1epa	0.12	xtal	dimer	stable
1ewf	0.12	xtal	monomer	stable
1feh	0.26	xtal	dimer	uncertain
1fgk	0.15	xtal	monomer	stable
1fjm	0.12	bio	monomer	stable
1fkd	0.12	xtal	tetramer	stable
1fmt	0.48	xtal	monomer	stable
1g2a	0.12	xtal	monomer	stable
1gar	0.12	xtal	dimer	stable
1gjm	0.12	xtal	monomer	stable
1hf8	0.83	xtal	dimer	stable
2j5x	0.12	xtal	monomer	stable
1ilr	0.12	xtal	monomer	stable
1kpt	0.24	xtal	monomer	stable
1kwa	0.12	bio	tetramer	stable
1mpg	0.49	xtal	dimer	stable
1mss	0.74	bio	dimer	stable

1naw	0.33	bio	tetramer	stable
1np4	0.13	xtal	dimer	uncertain
1pbg	0.13	xtal	dimer	uncertain
1pda	0.18	xtal	monomer	stable
1ppo	0.12	xtal	dimer	stable
1qaz	0.12	xtal	monomer	stable
1qci	0.12	xtal	monomer	stable
1qdm	0.12	xtal	trimer	uncertain
1qha	0.22	bio	monomer	stable
1qjp	0.16	bio	monomer	stable
1qme	0.32	xtal	monomer	stable
1qpa	0.49	xtal	dimer	stable
1qtq	0.27	xtal	dimer	stable
1rb3	0.12	xtal	monomer	stable
1rhs	0.12	xtal	monomer	stable
1rne	0.81	xtal	dimer	stable
1shk	0.13	xtal	dimer	stable
1the	0.27	xtal	dimer	stable
1tht	0.83	xtal	dimer	stable
1toa	0.45	bio	dimer	stable
1ton	0.39	xtal	dimer	stable
1urp	0.12	xtal	monomer	stable
1vbt	0.16	bio	monomer	stable
1xgs	0.74	xtal	monomer	stable
256b	0.12	xtal	monomer	stable
256l	0.24	xtal	monomer	stable
2acy	0.68	bio	monomer	stable
2atj	0.12	xtal	monomer	stable
2bc2	0.12	bio	dimer	stable
2bls	0.23	xtal	dimer	uncertain
2erc	0.13	xtal	dimer	stable
2g3p	0.49	xtal	dimer	stable
2ihl	0.61	xtal	monomer	stable
2mbr	0.12	bio	monomer	stable
2scp	0.77	bio	dimer	uncertain
2shp	0.12	xtal	monomer	stable
2tps	0.69	xtal	monomer	stable
2ugi	0.15	xtal	dimer	stable
3mht	0.85	xtal	trimer	stable
3pmg	0.29	xtal	monomer	stable
5tss	0.49	xtal	dimer	uncertain
830c	0.73	bio	tetramer	stable
8pti	0.22	bio	dimer	stable

^a yellow: incorrect classification

^b monomers (crystallographic dimers) from Baskaran et al. [4]

Table S5. MCC value as function of the threshold T^a

Threshold T	MCC
20	0.64
21	0.64
22	0.66
23	0.66
24	0.66
25	0.66
26	0.67
27	0.66
28	0.67
29	0.67
30	0.67
31	0.68
32	0.72
33	0.73
34	0.72
35	0.72
36	0.72
37	0.72
38	0.72
39	0.71
40	0.69
41	0.69
42	0.69
43	0.69
44	0.68
45	0.68
46	0.68
47	0.67
48	0.67
49	0.67
50	0.67
51	0.67
52	0.67
53	0.67
54	0.67
55	0.66
56	0.65
57	0.65
58	0.63
59	0.64
60	0.64

^a MCC: Matthews correlation coefficient

Table S6. Interface area, number of near-native structures, and inferred probability of being a biological dimer for the proteins in the DC Bio and DC Xtal sets^a

Biological dimers				Crystallographic dimers			
PDB ID	Interface area (Å ²)	Nnn	Probability (dimer)	PDB ID	Interface area (Å ²)	Nnn	Probability (dimer)
2fwv	2088.9	225	1.00	1ejd	1215.9	16	0.33
1v2x	1684.6	178	0.86	1gpi	1043.3	51	0.64
1pkh	1624.2	93	0.79	1j96	1217.0	37	0.54
1r5y	1551.5	150	0.83	1lqt	1099.9	0	0.12
1kq3	1487.5	129	0.83	1lxx	1028.7	1	0.13
1jq5	1325.9	191	0.88	1n45	1089.3	11	0.27
2h7i	1468.8	99	0.80	1pp3	1106.8	23	0.41
2nzi	1299.0	177	0.86	1s83	925.1	1	0.13
1uj6	1257.5	137	0.83	1so7	1189.4	15	0.32
1ju3	1266.4	49	0.62	1ueb	1150.0	26	0.44
1sml	1253.3	37	0.54	1w9q	1000.7	57	0.67
3fah	1236.5	153	0.83	1wly	1144.8	57	0.67
1lzl	1207.4	219	0.99	1xgk	1014.5	5	0.19
3iue	1106.1	101	0.81	1ynq	1100.6	5	0.19
1vk5	1110.5	69	0.73	2cki	1203.5	79	0.76
2exb	1895.8	41	0.57	2eyi	1112.3	0	0.12
3a2q	1745.5	156	0.84	2f37	1163.9	0	0.12
1s2z	1619.4	170	0.85	2gas	1561.6	89	0.79
2z1n	1578.5	91	0.79	2hlq	1071.7	21	0.39
2vef	1544.2	204	0.92	2ipi	1330.5	0	0.12
1bs1	1455.5	237	1.00	2j0p	1016.7	189	0.88
3lw6	1013.0	200	0.91	2j46	1152.0	6	0.20
2vr4	1278.8	55	0.66	2ow9	1179.4	65	0.71
1x7v	1190.5	220	1.00	2q7d	1782.7	46	0.60
1f2d	1890.4	63	0.70	2w20	1090.3	0	0.12
1lw4	1772.9	87	0.78	2wbf	1204.3	2	0.15
1n8p	1932.3	1	0.13	2wbq	1521.9	86	0.78
1qop	1536.3	85	0.78	2wsa	1002.4	0	0.12
2aq6	1188.1	59	0.68	2xov	1144.4	0	0.12
2bhs	1484.0	96	0.80	2yz1	1372.0	199	0.91
2cft	946.6	92	0.79	2z6o	1156.5	18	0.35
2e7j	1671.2	113	0.82	3aap	1382.8	139	0.83
2ecq	1626.5	68	0.72	3b37	1150.4	2	0.15
2rkb	1104.5	134	0.83	3c8y	1525.0	13	0.29
1eej	856.9	177	0.86	3cj1	1122.9	37	0.54
1o17	932.1	143	0.83	3cu9	1085.4	2	0.15
3d36	1660.8	97	0.80	3fwk	1044.4	0	0.12
3r0n	884.7	82	0.77	3gkj	992.4	32	0.49
3da8	867.3	11	0.27	3go5	1049.8	0	0.12
2c4w	872.9	15	0.32	3gvo	1208.0	65	0.71
3jrz	932.7	232	1.00	3h30	1310.1	4	0.18
2wxd	954.3	82	0.77	3hzi	1000.0	38	0.54
2d0d	1003.5	118	0.82	3kk8	1287.3	143	0.83
3cm3	1014.1	115	0.82	3m66	1006.6	2	0.15
3o1n	1007.0	307	1.00	3mg1	1095.8	0	0.12

3h0n	993.9	129	0.83	3mhz	1530.8	0	0.12
2dvn	1045.0	249	1.00	1lf2	2139.9	109	0.81
3ovp	1067.4	129	0.83	2qb5	1797.0	49	0.62
2i7d	1076.0	158	0.84	1zfq	1653.2	13	0.29
2car	1100.5	186	0.87	2yvw	1516.4	65	0.71
3gus	1215.6	167	0.84	1d3h	1464.7	0	0.12
3jyo	1224.2	279	1.00	2e1v	1657.8	161	0.84
1uz3	1196.3	105	0.81	3n5c	1455.7	147	0.83
2wtm	1241.9	125	0.83	2wbm	1441.3	62	0.70
2a5l	1217.2	43	0.58	3c1d	1429.8	49	0.62
2y39	1285.5	125	0.83	1ndb	1383.0	0	0.12
2ab0	1290.6	97	0.80	2x26	1369.3	158	0.84
3epw	1302.4	31	0.49	3kh7	1329.5	114	0.82
2vvt	1328.8	16	0.33	1toa	1294.4	27	0.45
1p5f	1339.3	156	0.84	1ffr	1223.4	76	0.75
3itf	1392.3	80	0.76	1vqq	1173.4	2	0.15
3kd2	1365.8	149	0.83	1n4g	1161.5	101	0.81
2y27	1355.8	154	0.84	3ita	1139.6	13	0.29
				1woq	1130.0	120	0.82
				1cqx	1109.5	0	0.12
				2eqa	1106.3	17	0.34
				3mhj	1098.5	8	0.23
				2fgz	1092.2	1	0.13
				1fpo	1080.0	0	0.12
				3lvd	996.1	46	0.60
				3irb	1075.3	64	0.70
				1utj	986.3	0	0.12
				3f0o	1040.1	9	0.24
				2h44	1041.1	312	1.00
				1t8g	1026.1	0	0.12
				1g6a	1021.4	9	0.24
				2zyr	1010.6	32	0.49
				3els	949.9	15	0.32
Average	1309.0	126.0	0.78	Average	1212.5	43.1	0.42
Std Dev	293.8	68.0	0.18	Std Dev	225.4	58.8	0.28

^a DC Bio and DC Xtal sets from Duarte et al. [5]

Table S7. Classification of DC Bio set by three methods^{a, b}

PDB ID	ClusPro probability of dimer	EPPIC	PISA multimer class	PISA status
2fwv	1.00	bio	tetramer	stable
1v2x	0.86	bio	dimer	stable
1pkh	0.79	bio	hexamer	stable
1r5y	0.83	bio	dimer	stable
1kq3	0.83	bio	octomer	stable
1jq5	0.88	bio	octomer	stable
2h7i	0.80	bio	tetramer	stable
2nzl	0.86	bio	tetramer	stable
1uj6	0.83	bio	dimer	stable
1ju3	0.62	xtal	dimer	stable
1sml	0.54	bio	tetramer	stable
3fah	0.83	bio	dimer	stable
1lzl	0.99	bio	dimer	stable
3iue	0.81	bio	monomer	stable
1vk5	0.73	bio	dimer	stable
2exb	0.57	bio	dimer	stable
3a2q	0.84	bio	dimer	stable
1s2z	0.85	bio	tetramer	stable
2z1n	0.79	bio	tetramer	stable
2vef	0.92	bio	dimer	stable
1bs1	1.00	bio	dimer	stable
3lw6	0.91	bio	monomer	stable
2vr4	0.66	bio	dimer	uncertain
1x7v	1.00	bio	hexamer	stable
1f2d	0.70	bio	dimer	stable
1lw4	0.78	bio	tetramer	stable
1n8p	0.13	bio	tetramer	stable
1qop	0.78	bio	tetramer	stable
2aq6	0.68	bio	dimer	stable
2bhs	0.80	bio	dimer	stable
2cft	0.79	bio	dimer	stable
2e7j	0.82	bio	dimer	stable
2ecq	0.72	xtal	dimer	stable
2rkb	0.83	bio	dimer	stable
1eej	0.86	bio	dimer	stable
1o17	0.83	xtal	dimer	stable
3d36	0.80	bio	hexamer	stable
3r0n	0.77	bio	dimer	stable
3da8	0.27	bio	dimer	stable
2c4w	0.32	bio	dodecamer	stable
3jrz	1.00	bio	dimer	stable
2wxd	0.77	bio	dimer	stable
2d0d	0.82	bio	dimer	uncertain
3cm3	0.82	bio	dimer	stable
3o1n	1.00	bio	dimer	stable
3h0n	0.83	bio	dimer	stable
2dvn	1.00	bio	tetramer	stable

3ovp	0.83	bio	dimer	stable
2i7d	0.84	bio	dimer	stable
2car	0.87	bio	dimer	stable
3gus	0.84	bio	dimer	stable
3jyo	1.00	bio	dimer	stable
1uz3	0.81	bio	dimer	stable
2wtm	0.83	xtal	dimer	stable
2a5l	0.58	bio	tetramer	stable
2y39	0.83	bio	dimer	stable
2ab0	0.80	bio	dimer	stable
3epw	0.49	bio	dimer	stable
2vvt	0.33	bio	dimer	stable
1p5f	0.84	bio	dimer	stable
3itf	0.76	bio	dimer	stable
3kd2	0.83	bio	dimer	stable
2y27	0.84	bio	dimer	stable

^a yellow: incorrect classification

^a DC Bio and DC Xtal sets from Duarte et al. [5]

Table S8. Classification of DC Xtal set by three methods^{a,b}

PDB ID	ClusPro probability of dimer	EPPIC	PISA multimer class	PISA status
1ejd	0.33	bio	tetramer	stable
1gpi	0.64	xtal	monomer	stable
1j96	0.54	xtal	monomer	stable
1lqt	0.12	xtal	monomer	stable
1lxx	0.13	xtal	monomer	stable
1n45	0.27	xtal	monomer	stable
1pp3	0.41	xtal	monomer	stable
1s83	0.13	xtal	monomer	stable
1so7	0.32	xtal	monomer	stable
1ueb	0.44	xtal	dimer	stable
1w9q	0.67	bio	trimer	stable
1wly	0.67	bio	dimer	stable
1xgk	0.19	bio	dimer	stable
1ynq	0.19	xtal	dimer	stable
2cki	0.76	xtal	dimer	stable
2eyi	0.12	xtal	monomer	stable
2f37	0.12	xtal	monomer	stable
2gas	0.79	xtal	dimer	stable
2hlq	0.39	xtal	monomer	stable
2ipi	0.12	bio	monomer	stable
2j0p	0.88	bio	monomer	stable
2j46	0.20	xtal	monomer	stable
2ow9	0.71	bio	tetramer	stable
2q7d	0.60	xtal	dimer	stable
2w20	0.12	xtal	monomer	stable
2wbf	0.15	xtal	monomer	stable
2wbq	0.78	bio	dimer	stable
2wsa	0.12	xtal	monomer	stable
2xov	0.12	bio	monomer	stable
2yz1	0.91	bio	dimer	stable
2z6o	0.35	xtal	monomer	stable
3aap	0.83	bio	monomer	stable
3b37	0.15	xtal	monomer	stable
3c8y	0.29	xtal	monomer	stable
3cj1	0.54	xtal	monomer	stable
3cu9	0.15	xtal	dimer	stable
3fwk	0.12	xtal	monomer	stable
3gkj	0.49	bio	monomer	stable
3go5	0.12	xtal	monomer	stable
3gvo	0.71	bio	monomer	stable
3h30	0.18	xtal	dimer	stable
3hzi	0.54	bio	dimer	uncertain
3kk8	0.83	xtal	monomer	stable
3m66	0.15	xtal	monomer	stable
3mg1	0.12	xtal	monomer	stable
3mhz	0.12	xtal	monomer	stable
1lf2	0.81	xtal	tetramer	stable
2qb5	0.62	xtal	monomer	stable

1zlq	0.29	xtal	dimer	stable
2yvw	0.71	bio	dimer	stable
1d3h	0.12	xtal	monomer	stable
2e1v	0.84	xtal	dimer	stable
3n5c	0.83	bio	dimer	stable
2wbm	0.70	bio	tetramer	stable
3c1d	0.62	xtal	dimer	stable
1ndb	0.12	xtal	monomer	stable
2x26	0.84	xtal	monomer	stable
3kh7	0.82	bio	dimer	stable
1toa	0.45	bio	dimer	stable
1ffr	0.75	bio	monomer	stable
1vqq	0.15	xtal	dimer	stable
1n4g	0.81	bio	dimer	stable
3ita	0.29	xtal	dimer	stable
1woq	0.82	bio	dimer	stable
1cqx	0.12	bio	dimer	stable
2eqa	0.34	bio	monomer	stable
3mhj	0.23	xtal	monomer	stable
2fgz	0.13	xtal	monomer	stable
1fpo	0.12	bio	trimer	stable
3lvd	0.60	xtal	monomer	stable
3irb	0.70	xtal	tetramer	stable
1utj	0.12	xtal	monomer	stable
3f0o	0.24	bio	dimer	stable
2h44	1.00	xtal	dimer	stable
1t8g	0.12	xtal	octomer	stable
1g6a	0.24	bio	dimer	stable
2zyr	0.49	xtal	monomer	stable
3els	0.32	xtal	hexamer	stable

^a yellow: incorrect classification

^a DC Bio and DC Xtal sets from Duarte et al. [5]

Table S9. Interface area, number of near-native structures, and inferred probability of being a biological dimer for the proteins in the test set^a

Biological dimers				Crystallographic dimers			
PDB ID	Interface area (Å ²)	Nnn	Probability (dimer)	PDB ID	Interface area (Å ²)	Nnn	Probability (dimer)
3w66	529.6	0	0.12	3w6r	859.9	54	0.65
3w8g	760.8	183	0.87	3w94	750.3	0	0.12
3we1	529.3	1	0.13	3w9g	604.3	8	0.23
3wfi	1028.6	316	1.00	3wbi	1467.0	0	0.12
3wjk	1722.6	192	0.89	3wbj	1195.5	59	0.68
3wkw	2891.9	81	0.77	3wdf	546.3	42	0.58
3wmd	1081.4	47	0.61	3wea	445.6	0	0.12
3wo8	2446.6	104	0.81	3who	567.0	1	0.13
3wpx	992.8	181	0.86	3wj9	588.6	0	0.12
3wsc	1566.2	1	0.13	3wo9	710.6	0	0.12
3wtz	690.6	74	0.74	3wob	460.8	1	0.13
3wu0	687.0	88	0.78	3wp3	644.0	0	0.12
3wva	1517.9	140	0.83	3wp9	716.5	23	0.41
3wvb	1521.0	155	0.84	3wpj	566.9	0	0.12
3wvo	1540.6	78	0.76	3wpy	1100.1	5	0.19
3wvz	2607.7	82	0.77	3wpz	1118.5	25	0.43
3wx4	1220.4	159	0.84	3wrt	490.3	0	0.12
3wxi	1416.9	148	0.83	3wuq	798.5	0	0.12
3wy7	2592.5	114	0.82	3wuy	872.8	0	0.12
3wyb	3216.7	96	0.80	3wv4	514.6	0	0.12
3wyd	542.7	68	0.72	3wvf	451.6	0	0.12
3wzg	1066.5	164	0.84	3wvt	589.1	0	0.12
4bhy	1988.9	37	0.54	3wx5	403.3	0	0.12
4c7v	3998.0	78	0.76	3wzt	718.4	13	0.29
4c88	1107.9	94	0.79	3x0w	595.8	0	0.12
4cbv	1777.7	54	0.65	3zfm	333.2	0	0.12
4chk	743.5	49	0.62	3zfy	404.9	0	0.12
4cij	1458.2	37	0.54	3zk0	589.7	7	0.22
4ckk	1761.7	55	0.66	3zlh	1825.2	97	0.80
4ckn	1371.7	82	0.77	3zm0	334.0	0	0.12
4cmz	3689.9	220	1.00	3zm1	314.0	0	0.12
4cop	489.0	131	0.83	3zm2	324.7	0	0.12
4cti	2596.6	62	0.70	3zm3	381.8	0	0.12
4cu5	706.1	481	1.00	3zny	564.5	0	0.12
4cw9	635.1	0	0.12	3zq3	748.1	10	0.26
4cwc	1505.3	52	0.64	4bg0	892.2	0	0.12
4cwe	1378.3	30	0.48	4bob	354.0	0	0.12
4cxj	2001.5	137	0.83	4bod	453.0	0	0.12
4d28	1404.9	34	0.51	4bp8	1075.0	51	0.64
4d5s	758.2	142	0.83	4bpf	580.6	2	0.15
4d6k	2270.1	227	1.00	4bpy	426.5	0	0.12
4fb8	2682.2	108	0.81	4bxm	584.7	0	0.12
4ihe	1746.5	118	0.82	4c41	620.9	0	0.12
4j0e	1685.2	213	0.96	4c5f	599.5	0	0.12
4j0f	1644.4	216	0.97	4c6r	710.5	11	0.27
4j7a	1116.5	110	0.82	4cbe	896.0	0	0.12

4j7r	1911.8	232	1.00	4cbh	1171.6	0	0.12
4jgm	698.2	0	0.12	4cbi	1177.1	0	0.12
4jp2	1483.3	29	0.47	4cbl	1332.7	6	0.20
4ju5	1810.7	214	0.97	4cbm	1180.1	0	0.12
4jup	389.5	20	0.38	4cdo	458.8	20	0.38
4k2e	1349.0	163	0.84	4cfi	628.2	0	0.12
4kde	1615.9	281	1.00	4ch7	579.2	0	0.12
4ke2	1787.5	4	0.18	4chj	930.5	144	0.83
4km3	1157.9	125	0.83	4cil	783.5	2	0.15
4kn8	1201.8	255	1.00	4co7	515.3	0	0.12
4kp2	1756.8	41	0.57	4cso	1146.4	5	0.19
4kz1	695.3	104	0.81	4cx8	746.1	0	0.12
4l2w	1072.2	47	0.61	4cxi	2001.1	145	0.83
4l4q	2598.3	74	0.74	4cye	684.5	148	0.83
4l69	1287.4	132	0.83	4d2o	1026.0	0	0.12
4l9e	1306.0	391	1.00	4d4q	533.7	0	0.12
4le5	2448.8	46	0.60	4d5r	790.3	109	0.81
4lft	517.9	0	0.12	4d70	445.7	0	0.12
4ljk	791.5	194	0.89	4d7r	523.9	0	0.12
4ljl	710.3	213	0.96	4h9w	438.6	2	0.15
4lpu	2061.2	98	0.80	4hic	542.7	0	0.12
4lpw	1916.8	74	0.74	4icu	545.8	0	0.12
4lpx	2109.4	50	0.63	4ino	904.8	0	0.12
4ltb	5267.9	18	0.35	4iqc	581.3	1	0.13
4luy	3263.7	115	0.82	4irf	674.0	22	0.40
4lwo	1645.7	102	0.81	4it7	466.5	0	0.12
4m0d	4320.9	55	0.66	4iyc	773.1	0	0.12
4m0q	1449.4	83	0.77	4ize	764.7	13	0.29
4m23	1547.5	86	0.78	4izl	772.1	0	0.12
4m36	3405.2	47	0.61	4j0o	798.6	0	0.12
4m6a	2105.6	117	0.82	4j1s	602.8	3	0.16
4m8z	965.3	103	0.81	4j2r	620.9	0	0.12
4mac	488.9	106	0.81	4jkq	787.4	0	0.12
4mdu	943.7	5	0.19	4jmp	454.7	0	0.12
4mea	1505.2	351	1.00	4k1s	473.8	0	0.12
4mei	682.9	14	0.31	4k40	692.2	0	0.12
4mfu	963.5	0	0.12	4k51	874.1	0	0.12
4mfv	896.3	0	0.12	4k5q	626.7	0	0.12
4mid	1266.1	192	0.89	4k82	729.0	0	0.12
4mir	1560.6	187	0.87	4k83	565.0	39	0.55
4mis	1493.7	206	0.93	4k9s	727.3	0	0.12
4ml3	811.0	52	0.64	4kex	449.1	0	0.12
4mld	758.2	57	0.67	4kxo	1758.6	186	0.87
4mmn	831.4	311	1.00	4l1d	447.5	25	0.43
4mpl	1249.3	144	0.83	4l6u	635.4	0	0.12
4mt8	2024.5	47	0.61	4l8i	827.8	0	0.12
4mtx	2122.5	44	0.59	4l8l	871.1	141	0.83
4muo	1129.0	109	0.81	4l9h	1109.1	85	0.78
4mxm	1122.6	130	0.83	4lcm	814.0	0	0.12
4myc	6853.3	10	0.26	4lee	564.5	0	0.12
4myz	1566.4	402	1.00	4lhf	1034.8	148	0.83
4mzm	1257.7	246	1.00	4lk4	595.0	9	0.24
4mzp	1277.0	219	0.99	4lpt	788.0	40	0.56

4mzt	1239.1	182	0.86	4lrk	1351.9	126	0.83
4n06	1254.8	56	0.66	4lso	524.0	0	0.12
4n1k	871.5	22	0.40	4lth	406.7	0	0.12
4n1y	1624.5	185	0.87	4lw5	681.9	1	0.13
4n92	584.4	0	0.12	4lzi	660.2	64	0.70
4n9h	1857.9	155	0.84	4m64	828.6	0	0.12
4na2	2607.9	78	0.76	4m7g	499.9	0	0.12
4ncd	855.9	151	0.83	4m80	602.7	0	0.12
4ndj	2363.8	70	0.73	4m8g	577.1	20	0.38
4ndk	1414.9	151	0.83	4m90	630.5	0	0.12
4ndw	2010.2	199	0.91	4m9p	903.2	17	0.34
4nj0	882.7	229	1.00	4mc0	641.1	0	0.12
4nj1	924.9	217	0.98	4mck	1005.3	26	0.44
4nk7	785.6	45	0.60	4mee	423.1	0	0.12
4nkk	1467.8	331	1.00	4mel	1516.8	187	0.87
4nn1	3044.8	243	1.00	4mem	992.7	24	0.42
4nom	813.5	0	0.12	4mfi	957.9	20	0.38
4npu	1565.1	382	1.00	4mfj	742.9	1	0.13
4nqf	1203.6	32	0.49	4mgf	1031.0	27	0.45
4nqt	1286.7	67	0.72	4mgs	415.5	0	0.12
4nsr	1126.8	138	0.83	4mhp	712.7	0	0.12
4nsw	5938.3	14	0.31	4mhr	1037.8	260	1.00
4ny0	735.5	0	0.12	4mq3	525.1	0	0.12
4o0q	1498.6	237	1.00	4mt5	743.5	9	0.24
4o1h	731.2	319	1.00	4mt6	713.4	0	0.12
4o2h	919.3	20	0.38	4mt7	458.5	0	0.12
4o60	1437.3	1	0.13	4mzd	399.1	0	0.12
4o8b	1438.0	274	1.00	4n08	683.6	0	0.12
4oaf	1130.5	28	0.46	4n1l	691.4	9	0.24
4oft	1298.0	255	1.00	4n3x	940.2	122	0.82
4ogd	2516.7	71	0.73	4n40	654.7	0	0.12
4ogf	1330.1	149	0.83	4n4a	1357.1	5	0.19
4oh9	1447.5	132	0.83	4n5a	856.9	0	0.12
4oi3	1069.6	416	1.00	4n5c	944.0	0	0.12
4okc	2395.3	129	0.83	4n5q	877.0	2	0.15
4oky	811.4	26	0.44	4n6l	796.5	6	0.20
4om7	671.5	32	0.49	4n6t	555.7	0	0.12
4omt	1234.1	12	0.28	4n6u	591.8	0	0.12
4ooi	1425.5	219	0.99	4n75	486.6	0	0.12
4or8	3048.9	108	0.81	4n7f	626.1	142	0.83
4otm	2117.8	137	0.83	4n8h	1844.3	36	0.53
4oun	1301.0	168	0.85	4n8k	1885.0	42	0.58
4p08	1192.2	39	0.55	4n8l	1813.1	30	0.48
4p12	1514.5	37	0.54	4nbo	626.7	36	0.53
4p22	2447.4	55	0.66	4nc9	920.1	25	0.43
4p61	1613.3	276	1.00	4nci	594.3	53	0.65
4p6k	544.4	8	0.23	4ndl	376.7	0	0.12
4p6l	535.0	1	0.13	4nfb	772.8	0	0.12
4p96	1702.6	97	0.80	4nfc	641.8	33	0.50
4p9f	1175.1	105	0.81	4nfx	1640.6	226	1.00
4pd0	3058.9	40	0.56	4ng0	407.9	382	1.00
4pdp	738.4	0	0.12	4nj6	657.2	57	0.67
4pgg	4733.5	121	0.82	4nj7	462.7	0	0.12

4ph6	990.4	145	0.83	4nj8	364.8	0	0.12
4pmu	955.2	117	0.82	4njc	726.7	368	1.00
4pn0	2112.5	172	0.85	4nk6	807.5	17	0.34
4pon	1508.6	226	1.00	4nk8	803.6	84	0.77
4prs	1471.2	120	0.82	4nl7	1254.8	0	0.12
4pu7	1034.3	282	1.00	4nlb	611.1	0	0.12
4pu8	1047.0	229	1.00	4nng	613.8	0	0.12
4puh	705.7	327	1.00	4nnh	602.0	11	0.27
4pus	653.0	0	0.12	4nnk	673.3	0	0.12
4pvc	1038.6	97	0.80	4noj	473.9	0	0.12
4pz9	792.6	15	0.32	4npf	1404.4	182	0.86
4pzc	1882.4	201	0.91	4nq0	624.3	0	0.12
4q0c	4048.2	69	0.73	4nsp	695.8	0	0.12
4q20	1582.8	106	0.81	4nu2	651.3	0	0.12
4q25	1306.1	54	0.65	4nv8	837.6	73	0.74
4q2o	1481.0	282	1.00	4nx9	1116.7	58	0.68
4q2q	1345.1	420	1.00	4nzz	497.2	0	0.12
4q75	3807.5	100	0.80	4o2x	524.9	15	0.32
4q76	3645.9	100	0.80	4o3x	376.2	0	0.12
4q97	977.5	26	0.44	4o4c	814.6	23	0.41
4q9b	727.8	140	0.83	4o5p	352.7	0	0.12
4qfb	656.8	0	0.12	4o5s	382.7	0	0.12
4qhr	2359.5	137	0.83	4o6g	579.5	1	0.13
4qjn	1902.4	166	0.84	4o6k	880.3	6	0.20
4ql6	1293.0	0	0.12	4o6x	616.0	151	0.83
4qr0	1428.3	229	1.00	4o7q	357.0	0	0.12
4qr1	1396.3	235	1.00	4o8s	340.7	1	0.13
4qr2	1398.5	230	1.00	4o8v	613.5	0	0.12
4qva	1921.1	15	0.32	4o9f	621.2	0	0.12
4qvg	3542.2	190	0.88	4o9l	660.7	5	0.19
4qvj	1923.1	15	0.32	4oa6	410.0	0	0.12
4qxz	689.0	117	0.82	4ob4	577.0	64	0.70
4qyx	532.0	7	0.22	4od4	650.6	0	0.12
4r0b	1180.1	97	0.80	4od6	423.7	0	0.12
4r1i	1813.7	12	0.28	4odd	638.8	78	0.76
4r1n	1837.7	254	1.00	4of0	898.4	0	0.12
4r27	890.3	0	0.12	4of5	565.9	0	0.12
4r3q	2250.6	94	0.79	4ofv	574.2	0	0.12
4r8e	2890.5	89	0.79	4ohd	1073.1	0	0.12
4ra8	670.7	2	0.15	4ohe	788.4	0	0.12
4rct	1569.6	96	0.80	4ohh	873.1	0	0.12
4rd9	1799.3	77	0.75	4ohi	825.7	0	0.12
4rdj	1762.0	90	0.79	4ohl	652.1	0	0.12
4rg7	2210.3	199	0.91	4ohr	1356.2	306	1.00
4rie	1498.0	76	0.75	4oj0	1524.5	139	0.83
4rmb	2540.6	211	0.95	4ok0	290.2	0	0.12
4rni	1077.7	29	0.47	4om9	794.9	0	0.12
4rnj	772.9	14	0.31	4oq1	1078.4	0	0.12
4rns	2014.2	191	0.88	4oqw	1470.9	225	1.00
4rsj	2314.0	33	0.50	4orw	520.5	0	0.12
4rsw	1615.5	34	0.51	4os0	1024.8	452	1.00
4rvs	1189.0	49	0.62	4os3	1046.9	333	1.00
4ry2	5221.6	20	0.38	4os8	1010.5	369	1.00

4ryj	476.7	0	0.12	4otd	731.8	0	0.12
4ryp	3238.5	111	0.82	4ou0	280.0	7	0.22
4s03	851.7	49	0.62	4ow6	482.5	0	0.12
4s0j	849.4	50	0.63	4oyl	381.5	0	0.12
4to7	1040.6	340	1.00	4oyy	539.1	0	0.12
4tq7	774.9	108	0.81	4ozk	393.0	0	0.12
4tqj	949.5	305	1.00	4p09	778.0	0	0.12
4tr7	1139.1	28	0.46	4p0f	1184.7	19	0.36
4trh	2343.9	16	0.33	4p0j	758.0	50	0.63
4trt	1084.6	26	0.44	4p0k	730.5	94	0.79
4ttw	624.3	0	0.12	4p0l	748.4	6	0.20
4ttx	826.7	283	1.00	4p0m	844.4	5	0.19
4u2i	913.8	420	1.00	4p0o	1113.1	23	0.41
4u2u	3170.2	65	0.71	4p0y	485.3	0	0.12
4u3q	651.0	0	0.12	4p0z	476.1	0	0.12
4u4i	925.8	11	0.27	4p17	755.8	1	0.13
4u5h	590.7	11	0.27	4p39	783.3	22	0.40
4u64	1516.7	129	0.83	4p47	730.3	0	0.12
4u9c	430.5	0	0.12	4p48	312.2	0	0.12
4ua2	1791.7	77	0.75	4p5u	575.6	0	0.12
4uds	2686.6	183	0.87	4p7o	1213.8	3	0.16
4up8	4485.2	82	0.77	4pa1	1179.0	57	0.67
4uu3	1420.2	178	0.86	4pbd	917.9	43	0.58
4uvj	1053.6	13	0.29	4pcf	657.7	0	0.12
4uzn	550.4	0	0.12	4pd3	1330.8	8	0.23
4v1p	637.6	187	0.87	4pej	1102.3	46	0.60
4v2f	2260.3	121	0.82	4pek	665.9	0	0.12
4w4r	996.8	34	0.51	4pgr	605.0	0	0.12
4w69	859.5	57	0.67	4pgs	644.6	0	0.12
4w6a	721.1	126	0.83	4pgu	550.2	2	0.15
4w6c	921.9	79	0.76	4pgv	572.7	1	0.13
4w6g	903.5	138	0.83	4ph1	682.9	70	0.73
4w6h	714.0	69	0.73	4pjg	697.2	6	0.20
4w6i	675.5	110	0.82	4pk1	480.4	0	0.12
4w6k	770.7	78	0.76	4pk9	2832.5	86	0.78
4w6l	959.2	18	0.35	4pka	2944.0	109	0.81
4w6n	817.4	67	0.72	4pkx	510.9	0	0.12
4w6o	754.8	119	0.82	4pmh	582.2	0	0.12
4w6p	915.5	57	0.67	4pn7	702.5	0	0.12
4w6r	835.5	93	0.79	4po4	622.8	0	0.12
4w7x	819.9	83	0.77	4poy	441.2	0	0.12
4w8i	3009.0	87	0.78	4ppj	977.6	0	0.12
4wd9	3030.4	47	0.61	4ppk	921.5	0	0.12
4whj	1018.7	56	0.66	4ppl	1001.2	0	0.12
4wic	2466.7	17	0.34	4pqo	488.6	0	0.12
4wij	4371.7	26	0.44	4pqz	377.1	0	0.12
4wik	4164.8	28	0.46	4psd	472.6	0	0.12
4wja	1393.0	369	1.00	4psf	595.2	8	0.23
4wnz	1286.8	53	0.65	4pug	826.8	0	0.12
4wo7	1830.4	241	1.00	4pui	476.4	0	0.12
4wpe	3979.6	51	0.64	4pwq	802.1	1	0.13
4wse	3250.3	45	0.60	4pzp	538.0	0	0.12
4wt5	1363.3	109	0.81	4q2s	467.6	0	0.12

4wum	2951.4	86	0.78	4q2y	1153.2	8	0.23
4wy8	1280.4	75	0.75	4q3h	460.7	172	0.85
4x0r	1026.9	168	0.85	4q4g	422.2	0	0.12
4x5n	1515.3	43	0.58	4q4n	425.5	0	0.12
4x5u	3721.3	34	0.51	4q5y	1028.2	8	0.23
4x8f	1205.5	354	1.00	4q65	58.7	0	0.12
4xa1	3974.0	11	0.27	4q6l	600.1	0	0.12
4xa3	3674.4	13	0.29	4q6v	559.9	0	0.12
4xa4	2472.8	24	0.42	4q6z	625.4	0	0.12
4xa6	4370.0	9	0.24	4q7k	355.4	0	0.12
4xbx	1614.6	202	0.92	4q7m	1390.7	166	0.84
4xlx	1062.3	77	0.75	4q7t	528.1	0	0.12
4xo1	1576.0	209	0.94	4q9h	661.2	0	0.12
4xo2	1581.3	220	1.00	4q9p	775.3	58	0.68
4xoh	1083.9	6	0.20	4qaj	584.1	46	0.60
4xws	1050.8	49	0.62	4qaz	634.2	1	0.13
4xyg	3638.3	188	0.88	4qb0	467.1	0	0.12
4xz7	2973.9	89	0.79	4qbe	634.5	25	0.43
4y0m	1294.5	92	0.79	4qcj	472.0	14	0.31
4y9a	1662.8	218	0.99	4qdr	682.2	0	0.12
4ylt	1916.7	186	0.87	4qdt	562.9	0	0.12
4ymr	2237.6	80	0.76	4qdu	659.9	0	0.12
4ynh	1528.6	297	1.00	4qgy	573.2	0	0.12
4yub	2778.3	74	0.74	4qlo	1517.2	96	0.80
4zyz	1008.9	39	0.55	4qlr	618.0	2	0.15
4z27	1546.0	99	0.80	4qmd	653.3	0	0.12
4z6j	1571.1	157	0.84	4qmi	856.6	0	0.12
4zbh	510.6	171	0.85	4qmj	902.9	0	0.12
4zbw	472.6	0	0.12	4qs8	630.1	0	0.12
4zm8	3174.8	43	0.58	4qt4	599.0	0	0.12
4zqa	1794.3	58	0.68	4qtz	630.7	0	0.12
5a0o	974.7	0	0.12	4quc	200.4	0	0.12
5a2n	1096.1	0	0.12	4quk	419.9	0	0.12
5a48	778.2	1	0.13	4qx0	1356.7	12	0.28
5bto	876.8	22	0.40	4qx1	1328.2	6	0.20
				4qx2	1335.0	11	0.27
				4qx3	1326.8	9	0.24
				4qxl	385.4	0	0.12
				4r05	588.3	0	0.12
				4r0e	452.6	30	0.48
				4r0x	671.7	0	0.12
				4r0y	793.2	0	0.12
				4r19	706.5	0	0.12
				4r1a	468.7	0	0.12
				4r1b	536.5	0	0.12
				4r1c	549.8	0	0.12
				4r58	574.3	0	0.12
				4r61	710.9	31	0.49
				4r6f	782.0	0	0.12
				4r7v	757.5	4	0.18
				4r8r	680.7	0	0.12
				4rcw	1336.8	22	0.40
				4rdp	1381.7	111	0.82

4reh	643.8	0	0.12
4rej	614.6	0	0.12
4rf6	896.8	0	0.12
4rid	1471.5	4	0.18
4rl1	661.6	0	0.12
4rl9	506.7	0	0.12
4rlb	816.1	0	0.12
4rlm	560.4	0	0.12
4rln	568.5	0	0.12
4ro1	226.8	0	0.12
4roa	729.3	27	0.45
4rs6	869.0	122	0.82
4rsq	767.9	28	0.46
4rsx	933.4	0	0.12
4rtd	1339.0	1	0.13
4rth	1038.2	32	0.49
4rwi	1184.5	55	0.66
4rwq	645.6	0	0.12
4rwr	1023.5	86	0.78
4rwu	734.0	46	0.60
4rwz	895.4	91	0.79
4ry4	939.0	0	0.12
4ry6	963.0	0	0.12
4ry7	958.1	0	0.12
4rz2	769.4	0	0.12
4s11	565.3	0	0.12
4s3l	495.4	2	0.15
4s3p	2064.0	46	0.60
4tks	1025.6	43	0.58
4tlp	587.8	0	0.12
4tlw	863.4	0	0.12
4tn9	489.6	0	0.12
4tnm	642.4	0	0.12
4tql	785.7	0	0.12
4tr5	1340.3	222	1.00
4trk	536.2	0	0.12
4tse	1057.9	26	0.44
4ttl	125.0	0	0.12
4ttp	530.3	0	0.12
4tum	515.8	11	0.27
4tut	161.3	96	0.80
4tuu	517.7	0	0.12
4tve	480.4	2	0.15
4tx2	695.4	0	0.12
4txp	464.4	0	0.12
4tza	906.5	9	0.24
4tzg	1452.9	96	0.80
4tzi	605.0	0	0.12
4tzj	640.7	0	0.12
4u0l	753.8	0	0.12
4u1d	1336.3	3	0.16
4u2m	664.7	17	0.34
4u2n	698.3	224	1.00

4u3a	1320.8	202	0.92
4u3h	795.9	0	0.12
4u3t	871.0	2	0.15
4u3v	631.7	0	0.12
4u48	4122.1	21	0.39
4u4h	1423.8	1	0.13
4u4j	4204.9	24	0.42
4u59	4234.2	24	0.42
4u5a	629.0	0	0.12
4u9q	724.9	28	0.46
4ud5	886.9	0	0.12
4uex	709.7	0	0.12
4ulw	859.8	132	0.83
4umg	601.3	9	0.24
4ums	460.4	6	0.20
4uos	526.4	22	0.40
4uts	752.4	0	0.12
4uu4	881.4	17	0.34
4uv2	1979.0	80	0.76
4uv6	799.7	72	0.74
4uvk	744.7	0	0.12
4uw2	1029.0	0	0.12
4uww	441.0	250	1.00
4uy3	418.9	0	0.12
4uz2	739.0	5	0.19
4v0t	978.3	152	0.83
4v14	423.1	0	0.12
4v16	789.0	0	0.12
4v2b	539.8	0	0.12
4v2d	696.1	0	0.12
4v2e	1539.3	0	0.12
4v2u	357.2	0	0.12
4v38	372.1	0	0.12
4v39	359.2	0	0.12
4w4m	595.1	0	0.12
4w5x	437.2	18	0.35
4w60	399.3	37	0.54
4w6v	673.6	0	0.12
4w7g	525.1	0	0.12
4w8v	2072.4	302	1.00
4w8w	1381.4	116	0.82
4w8z	1533.3	143	0.83
4wau	818.2	16	0.33
4wbz	673.2	22	0.40
4wdc	428.5	0	0.12
4we2	291.9	0	0.12
4wfi	551.9	0	0.12
4wj1	1240.1	33	0.50
4wj2	595.3	0	0.12
4wjs	491.1	0	0.12
4wmg	599.5	0	0.12
4wo8	844.0	0	0.12
4woy	1029.5	44	0.59

4wp6	558.9	0	0.12
4wt3	425.1	0	0.12
4wtx	501.6	3	0.16
4wu0	868.2	11	0.27
4wz0	701.2	48	0.62
4x0j	912.8	0	0.12
4x2m	562.7	0	0.12
4x2w	1095.6	44	0.59
4x2x	580.8	4	0.18
4x2y	805.9	38	0.54
4x3x	712.4	0	0.12
4x43	643.0	0	0.12
4x8h	974.2	6	0.20
4x8m	978.8	0	0.12
4x9u	375.5	0	0.12
4xah	703.5	311	1.00
4xeh	5944.1	35	0.52
4xfu	599.9	0	0.12
4xfv	758.8	0	0.12
4xh3	767.8	0	0.12
4xhm	554.3	3	0.16
4xkk	495.8	1	0.13
4xm4	1840.7	79	0.76
4xp8	487.4	17	0.34
4xpk	578.7	0	0.12
4xq7	607.0	0	0.12
4xy3	578.5	0	0.12
4xzg	666.3	0	0.12
4y21	1200.0	28	0.46
4y4q	858.6	9	0.24
4y5j	1107.5	16	0.33
4yeq	874.7	17	0.34
4yf2	735.9	0	0.12
4yfe	383.1	0	0.12
4yfs	692.9	0	0.12
4yhe	555.5	0	0.12
4ykc	961.7	22	0.40
4ykd	532.1	0	0.12
4yno	880.1	0	0.12
4ynv	644.0	203	0.92
4ynw	651.0	246	1.00
4ypj	838.6	0	0.12
4yuo	353.0	0	0.12
4yvm	1688.5	20	0.38
4yvo	482.7	0	0.12
4z29	938.1	29	0.47
4z99	583.3	0	0.12
4z9p	781.4	67	0.72
4zf3	739.3	15	0.32
4zf4	587.1	22	0.40
4zf5	749.1	38	0.54
4zg1	539.5	0	0.12
4zj2	454.2	0	0.12

4zj3	404.1	0	0.12
4zmd	382.5	0	0.12
4zot	532.0	0	0.12
4zov	1138.3	158	0.84
4zoy	734.9	0	0.12
4zuz	1005.3	5	0.19
4zwj	1081.6	0	0.12
5a0g	720.4	0	0.12
5a0z	1330.2	37	0.54
5aem	898.4	4	0.18
5af5	291.9	0	0.12
5aio	872.5	104	0.81
5aj8	671.6	0	0.12
5avg	710.4	13	0.29
5avh	4608.7	89	0.79
5brl	535.8	5	0.19
5bth	695.4	12	0.28
5bvq	607.2	3	0.16
5c00	915.8	149	0.83
5c86	431.7	0	0.12
Average	1635.0	111.7	0.68
Std Dev	1029.7	98.7	0.27
Average	793.6	26.8	0.28
Std Dev	528.2	62.5	0.26

^a Test set selection was based on the following criteria: (1) PDB release date between January 2014 and August 2015; (2) no ligands in structure; (3) only a single type of protein in the structure, i.e., no heterodimers; and (4) the PDB file describes author determined biological units to assess the biological assembly as suggested by the authors.

Table S10. Classification of test set dimers by three methods^{a,b}

PDB ID	ClusPro probability of dimer	EPPIC	PISA multimer class	PISA status
3w66	0.12	xtal	dimer	uncertain
3w8g	0.87	xtal	dimer	stable
3we1	0.13	xtal	monomer	stable
3wfi	1.00	bio	dimer	stable
3wjk	0.89	bio	dimer	stable
3wkw	0.77	bio	dimer	stable
3wmd	0.61	bio	dimer	stable
3wo8	0.81	bio	dimer	stable
3wpx	0.86	bio	dimer	uncertain
3wsc	0.13	bio	dimer	stable
3wtz	0.74	xtal	dimer	stable
3wu0	0.78	xtal	dimer	stable
3wva	0.83	bio	dimer	stable
3wvb	0.84	bio	dimer	stable
3wvo	0.76	bio	dimer	stable
3wvz	0.77	bio	tetramer	stable
3wx4	0.84	xtal	dimer	stable
3wxi	0.83	bio	dimer	stable
3wy7	0.82	bio	dimer	stable
3wyb	0.80	bio	dimer	stable
3wyd	0.72	xtal	dimer	stable
3wzg	0.84	bio	dimer	stable
4bhy	0.54	bio	tetramer	stable
4c7v	0.76	bio	dimer	stable
4c88	0.79	bio	16-mer	stable
4cbv	0.65	bio	dimer	stable
4chk	0.62	bio	monomer	stable
4cij	0.54	bio	dimer	stable
4ckk	0.66	bio	dimer	stable
4ckn	0.77	xtal	dimer	stable
4cmz	1.00	bio	dimer	stable
4cop	0.83	xtal	dimer	uncertain
4cti	0.70	bio	dimer	stable
4cu5	1.00	bio	dimer	stable
4cw9	0.12	xtal	dimer	stable
4cwc	0.64	bio	tetramer	stable
4cwe	0.48	bio	dimer	stable
4cxj	0.83	bio	dimer	stable
4d28	0.51	xtal	dimer	stable
4d5s	0.83	bio	dimer	stable
4d6k	1.00	bio	dimer	stable
4fb8	0.81	bio	dimer	stable
4ihe	0.82	bio	tetramer	stable
4j0e	0.96	bio	dimer	stable
4j0f	0.97	bio	dimer	stable
4j7a	0.82	bio	tetramer	stable
4j7r	1.00	bio	dimer	stable

4jgm	0.12	bio	dimer	uncertain
4jp2	0.47	bio	tetramer	stable
4ju5	0.97	bio	dimer	stable
4jup	0.38	xtal	dimer	uncertain
4k2e	0.84	bio	dimer	stable
4kde	1.00	bio	dimer	stable
4ke2	0.18	xtal	hexamer	stable
4km3	0.83	xtal	dimer	stable
4kn8	1.00	bio	dimer	stable
4kp2	0.57	bio	dimer	stable
4kz1	0.81	bio	dimer	stable
4l2w	0.61	xtal	tetramer	stable
4l4q	0.74	bio	tetramer	stable
4l69	0.83	bio	dimer	stable
4l9e	1.00	bio	dimer	stable
4le5	0.60	bio	dimer	stable
4lft	0.12	bio	monomer	stable
4ljk	0.89	bio	dimer	uncertain
4ljl	0.96	bio	dimer	stable
4lpu	0.80	bio	tetramer	stable
4lpw	0.74	bio	hexamer	stable
4lpx	0.63	bio	hexamer	stable
4ltb	0.35	bio	dimer	stable
4luy	0.82	bio	dimer	stable
4lwo	0.81	bio	dimer	stable
4m0d	0.66	bio	tetramer	stable
4m0q	0.77	xtal	dimer	stable
4m23	0.78	bio	dimer	stable
4m36	0.61	bio	dimer	stable
4m6a	0.82	bio	dimer	stable
4m8z	0.81	bio	dimer	stable
4mac	0.81	bio	monomer	stable
4mdu	0.19	xtal	dimer	stable
4mea	1.00	bio	dimer	stable
4mei	0.31	bio	dimer	stable
4mfu	0.12	xtal	monomer	stable
4mfv	0.12	xtal	monomer	stable
4mid	0.89	bio	dimer	stable
4mir	0.87	bio	dimer	stable
4mis	0.93	bio	dimer	stable
4ml3	0.64	bio	dimer	uncertain
4mld	0.67	bio	dodecamer	stable
4mmn	1.00	bio	dimer	stable
4mpl	0.83	bio	dimer	stable
4mt8	0.61	bio	tetramer	stable
4mtx	0.59	bio	octomer	stable
4muo	0.81	bio	dimer	stable
4mxm	0.83	bio	dimer	stable
4myc	0.26	bio	dimer	stable
4myz	1.00	bio	dimer	stable
4mzm	1.00	bio	dimer	stable
4mzp	0.99	bio	dimer	stable
4mzt	0.86	bio	dimer	stable

4n06	0.66	bio	dimer	stable
4n1k	0.40	bio	tetramer	stable
4n1y	0.87	bio	dimer	stable
4n92	0.12	xtal	monomer	stable
4n9h	0.84	bio	dimer	stable
4na2	0.76	bio	dimer	stable
4ncd	0.83	xtal	monomer	stable
4ndj	0.73	bio	tetramer	stable
4ndk	0.83	bio	dimer	stable
4ndw	0.91	bio	dimer	stable
4nj0	1.00	xtal	dimer	stable
4nj1	0.98	xtal	dimer	stable
4nk7	0.60	bio	tetramer	uncertain
4nkk	1.00	bio	dimer	stable
4nn1	1.00	bio	dimer	stable
4nom	0.12	xtal	dimer	uncertain
4npu	1.00	bio	dimer	stable
4nqf	0.49	xtal	dimer	stable
4nqt	0.72	bio	dimer	stable
4nsr	0.83	xtal	dimer	stable
4nsw	0.31	bio	dimer	stable
4ny0	0.12	xtal	dimer	stable
4o0q	1.00	bio	dimer	stable
4o1h	1.00	bio	dimer	stable
4o2h	0.38	xtal	dimer	stable
4o60	0.13	bio	dimer	stable
4o8b	1.00	bio	dimer	stable
4oaf	0.46	bio	monomer	stable
4oft	1.00	bio	dimer	stable
4ogd	0.73	bio	dimer	stable
4ogf	0.83	bio	dimer	stable
4oh9	0.83	bio	dimer	stable
4oi3	1.00	bio	dimer	stable
4okc	0.83	bio	dimer	stable
4oky	0.44	xtal	dimer	stable
4om7	0.49	bio	dimer	stable
4omt	0.28	bio	monomer	stable
4ooi	0.99	bio	dimer	stable
4or8	0.81	bio	dimer	stable
4otm	0.83	bio	dimer	stable
4oun	0.85	bio	dimer	stable
4p08	0.55	xtal	dimer	stable
4p12	0.54	bio	dimer	stable
4p22	0.66	bio	dimer	stable
4p61	1.00	bio	dimer	stable
4p6k	0.23	bio	tetramer	stable
4p6l	0.13	xtal	dimer	stable
4p96	0.80	bio	dimer	stable
4p9f	0.81	bio	hexamer	stable
4pd0	0.56	bio	dimer	stable
4pdp	0.12	xtal	monomer	stable
4pgg	0.82	bio	dimer	stable
4ph6	0.83	bio	dimer	stable

4pmu	0.82	bio	dimer	stable
4pn0	0.85	bio	dimer	stable
4pon	1.00	bio	dimer	stable
4prs	0.82	bio	dimer	stable
4pu7	1.00	xtal	dimer	stable
4pu8	1.00	xtal	dimer	stable
4puh	1.00	bio	dimer	uncertain
4pus	0.12	xtal	dimer	uncertain
4pvc	0.80	xtal	monomer	stable
4pz9	0.32	bio	dimer	uncertain
4pzc	0.91	bio	hexamer	stable
4q0c	0.73	bio	dimer	stable
4q20	0.81	bio	tetramer	stable
4q25	0.65	bio	dimer	stable
4q2o	1.00	bio	dimer	stable
4q2q	1.00	bio	dimer	stable
4q75	0.80	bio	dimer	stable
4q76	0.80	bio	dimer	stable
4q97	0.44	bio	dimer	stable
4q9b	0.83	xtal	dimer	stable
4qfb	0.12	xtal	monomer	stable
4qhr	0.83	bio	dimer	stable
4qjn	0.84	bio	dimer	stable
4ql6	0.12	bio	dimer	uncertain
4qr0	1.00	bio	dimer	stable
4qr1	1.00	bio	dimer	stable
4qr2	1.00	bio	dimer	stable
4qva	0.32	bio	dimer	stable
4qvg	0.88	bio	dimer	stable
4qvj	0.32	bio	dimer	stable
4qxz	0.82	bio	dimer	uncertain
4qyx	0.22	xtal	monomer	stable
4r0b	0.80	bio	dimer	stable
4r1i	0.28	bio	dimer	stable
4r1n	1.00	bio	hexamer	stable
4r27	0.12	bio	monomer	stable
4r3q	0.79	bio	hexamer	stable
4r8e	0.79	bio	dimer	stable
4ra8	0.15	xtal	dimer	stable
4rct	0.80	bio	dimer	stable
4rd9	0.75	bio	dimer	stable
4rdj	0.79	xtal	dimer	stable
4rg7	0.91	bio	dimer	stable
4rie	0.75	bio	dimer	stable
4rmb	0.95	bio	dimer	stable
4rni	0.47	bio	dimer	stable
4rnj	0.31	xtal	monomer	stable
4rns	0.88	bio	dimer	stable
4rsj	0.50	bio	dimer	stable
4rsw	0.51	bio	dimer	stable
4rvs	0.62	bio	dimer	stable
4ry2	0.38	bio	dimer	stable
4ryj	0.12	xtal	dimer	stable

4ryp	0.82	bio	dimer	stable
4s03	0.62	xtal	dimer	uncertain
4s0j	0.63	xtal	dimer	uncertain
4to7	1.00	xtal	monomer	stable
4tq7	0.81	xtal	dimer	uncertain
4tqj	1.00	xtal	monomer	stable
4tr7	0.46	bio	dimer	stable
4trh	0.33	bio	dimer	stable
4trt	0.44	bio	dimer	stable
4ttw	0.12	xtal	dimer	stable
4ttx	1.00	bio	dimer	stable
4u2i	1.00	bio	dimer	uncertain
4u2u	0.71	bio	dimer	stable
4u3q	0.12	xtal	dimer	uncertain
4u4i	0.27	bio	monomer	stable
4u5h	0.27	xtal	dimer	stable
4u64	0.83	xtal	dimer	stable
4u9c	0.12	xtal	monomer	stable
4ua2	0.75	bio	octomer	stable
4uds	0.87	bio	dimer	stable
4up8	0.77	bio	dimer	stable
4uu3	0.86	bio	dimer	stable
4uvj	0.29	xtal	dimer	stable
4uzn	0.12	xtal	monomer	stable
4v1p	0.87	xtal	dimer	stable
4v2f	0.82	bio	dimer	stable
4w4r	0.51	bio	dimer	stable
4w69	0.67	xtal	monomer	stable
4w6a	0.83	xtal	monomer	stable
4w6c	0.76	bio	dimer	stable
4w6g	0.83	bio	monomer	stable
4w6h	0.73	xtal	monomer	stable
4w6i	0.82	xtal	monomer	stable
4w6k	0.76	xtal	dimer	uncertain
4w6l	0.35	xtal	tetramer	stable
4w6n	0.72	xtal	monomer	stable
4w6o	0.82	xtal	tetramer	stable
4w6p	0.67	xtal	monomer	stable
4w6r	0.79	xtal	monomer	stable
4w7x	0.77	xtal	monomer	stable
4w8i	0.78	bio	dimer	stable
4wd9	0.61	bio	dimer	stable
4whj	0.66	bio	dimer	stable
4wic	0.34	bio	dimer	stable
4wij	0.44	bio	dimer	stable
4wik	0.46	bio	dimer	stable
4wja	1.00	bio	dimer	stable
4wnz	0.65	bio	monomer	stable
4wo7	1.00	bio	dimer	stable
4wpe	0.64	bio	dimer	stable
4wse	0.60	bio	dimer	stable
4wt5	0.81	bio	dimer	stable
4wum	0.78	bio	dimer	stable

4wy8	0.75	bio	dimer	stable
4x0r	0.85	bio	dimer	stable
4x5n	0.58	bio	dimer	stable
4x5u	0.51	bio	dimer	stable
4x8f	1.00	bio	dimer	stable
4xa1	0.27	bio	dimer	stable
4xa3	0.29	bio	dimer	stable
4xa4	0.42	bio	dimer	stable
4xa6	0.24	bio	octomer	stable
4xbx	0.92	bio	dimer	stable
4xlx	0.75	bio	dimer	stable
4xo1	0.94	bio	dimer	stable
4xo2	1.00	bio	dimer	stable
4xoh	0.20	bio	dimer	stable
4xws	0.62	bio	dimer	stable
4xyg	0.88	bio	dimer	stable
4xz7	0.79	bio	dimer	stable
4y0m	0.79	bio	dimer	stable
4y9a	0.99	bio	dimer	stable
4ylt	0.87	bio	dimer	stable
4ymr	0.76	bio	dimer	stable
4ynh	1.00	bio	dimer	stable
4yub	0.74	bio	dimer	stable
4zyz	0.55	bio	dimer	stable
4z27	0.80	bio	dimer	stable
4z6j	0.84	bio	octomer	stable
4zbh	0.85	bio	dimer	uncertain
4zbw	0.12	xtal	monomer	stable
4zm8	0.58	bio	dimer	stable
4zqa	0.68	xtal	dimer	stable
5a0o	0.12	bio	dimer	stable
5a2n	0.12	xtal	dimer	stable
5a48	0.13	bio	dimer	uncertain
5bto	0.40	bio	monomer	stable

^a yellow: incorrect classification

^b Criteria for test selection are given in Table S9

Table S11. Classification of test set monomers by three methods^{a,b}

PDB ID	ClusPro probability of dimer	EPPIC	PISA multimer class	PISA status
3w6r	0.65	xtal	monomer	stable
3w94	0.12	xtal	monomer	stable
3w9g	0.23	xtal	dimer	stable
3wbi	0.12	bio	monomer	stable
3wbj	0.68	xtal	monomer	stable
3wdf	0.58	xtal	monomer	stable
3wea	0.12	xtal	dimer	stable
3who	0.13	bio	monomer	stable
3wj9	0.12	xtal	monomer	stable
3wo9	0.12	bio	monomer	stable
3wob	0.13	xtal	monomer	stable
3wp3	0.12	xtal	monomer	stable
3wp9	0.41	xtal	dimer	stable
3wpj	0.12	bio	monomer	stable
3wpy	0.19	xtal	monomer	stable
3wpz	0.43	xtal	monomer	stable
3wrt	0.12	xtal	monomer	stable
3wuq	0.12	xtal	monomer	stable
3wuy	0.12	bio	dimer	stable
3wv4	0.12	xtal	monomer	stable
3wvf	0.12	xtal	monomer	stable
3wvt	0.12	xtal	monomer	stable
3wx5	0.12	xtal	monomer	stable
3wzt	0.29	xtal	monomer	stable
3x0w	0.12	bio	monomer	stable
3zfm	0.12	xtal	monomer	stable
3zfy	0.12	xtal	monomer	stable
3zk0	0.22	xtal	monomer	stable
3zlh	0.80	bio	octomer	stable
3zm0	0.12	xtal	monomer	stable
3zm1	0.12	xtal	monomer	stable
3zm2	0.12	xtal	monomer	stable
3zm3	0.12	xtal	monomer	stable
3zny	0.12	xtal	monomer	stable
3zq3	0.26	xtal	monomer	stable
4bg0	0.12	xtal	monomer	stable
4bob	0.12	xtal	monomer	stable
4bod	0.12	xtal	monomer	stable
4bp8	0.64	bio	dimer	stable
4bpf	0.15	xtal	monomer	stable
4bpy	0.12	xtal	monomer	stable
4bxm	0.12	xtal	monomer	stable
4c41	0.12	bio	monomer	stable
4c5f	0.12	xtal	monomer	stable
4c6r	0.27	xtal	monomer	stable
4cbe	0.12	xtal	monomer	stable
4cbh	0.12	xtal	monomer	stable

4cbi	0.12	xtal	monomer	stable
4cbl	0.20	bio	dimer	stable
4cbm	0.12	bio	dimer	uncertain
4cdo	0.38	xtal	monomer	stable
4cfi	0.12	bio	monomer	stable
4ch7	0.12	xtal	monomer	stable
4chj	0.83	xtal	tetramer	uncertain
4cil	0.15	xtal	monomer	stable
4co7	0.12	xtal	monomer	stable
4cso	0.19	xtal	monomer	stable
4cx8	0.12	bio	dimer	stable
4cxi	0.83	bio	dimer	stable
4cye	0.83	bio	dimer	stable
4d2o	0.12	xtal	monomer	stable
4d4q	0.12	xtal	monomer	stable
4d5r	0.81	bio	dimer	stable
4d70	0.12	xtal	monomer	stable
4d7r	0.12	xtal	monomer	stable
4h9w	0.15	xtal	monomer	stable
4hic	0.12	xtal	monomer	stable
4icu	0.12	xtal	monomer	stable
4ino	0.12	xtal	monomer	stable
4iqc	0.13	xtal	monomer	stable
4irf	0.40	xtal	monomer	stable
4it7	0.12	xtal	monomer	stable
4iyc	0.12	xtal	monomer	stable
4ize	0.29	xtal	monomer	stable
4izl	0.12	xtal	monomer	stable
4j0o	0.12	xtal	monomer	stable
4j1s	0.16	xtal	monomer	stable
4j2r	0.12	xtal	monomer	stable
4jkq	0.12	bio	monomer	stable
4jmp	0.12	xtal	monomer	stable
4k1s	0.12	xtal	monomer	stable
4k40	0.12	xtal	monomer	stable
4k51	0.12	xtal	trimer	stable
4k5q	0.12	xtal	monomer	stable
4k82	0.12	xtal	monomer	stable
4k83	0.55	xtal	monomer	stable
4k9s	0.12	xtal	monomer	stable
4kex	0.12	xtal	monomer	stable
4kxo	0.87	bio	dimer	stable
4l1d	0.43	xtal	monomer	stable
4l6u	0.12	xtal	monomer	stable
4l8i	0.12	xtal	dimer	stable
4l8l	0.83	bio	dimer	uncertain
4l9h	0.78	bio	monomer	stable
4lcm	0.12	xtal	dimer	stable
4lee	0.12	xtal	monomer	stable
4lhf	0.83	bio	monomer	stable
4lk4	0.24	bio	monomer	stable
4lpt	0.56	bio	hexamer	stable
4lrk	0.83	bio	dimer	stable

4lso	0.12	bio	monomer	stable
4ltt	0.12	xtal	monomer	stable
4lw5	0.13	xtal	monomer	stable
4lzi	0.70	xtal	monomer	stable
4m64	0.12	xtal	monomer	stable
4m7g	0.12	bio	monomer	stable
4m80	0.12	xtal	monomer	stable
4m8g	0.38	xtal	monomer	stable
4m90	0.12	xtal	monomer	stable
4m9p	0.34	xtal	monomer	stable
4mc0	0.12	xtal	monomer	stable
4mck	0.44	xtal	monomer	stable
4mee	0.12	xtal	monomer	stable
4mel	0.87	bio	dimer	stable
4mem	0.42	bio	dimer	stable
4mfi	0.38	bio	monomer	stable
4mfj	0.13	xtal	monomer	stable
4mgf	0.45	xtal	monomer	stable
4mgs	0.12	xtal	monomer	stable
4mhp	0.12	xtal	monomer	stable
4mhr	1.00	bio	dimer	stable
4mq3	0.12	bio	monomer	stable
4mt5	0.24	xtal	monomer	stable
4mt6	0.12	xtal	monomer	stable
4mt7	0.12	xtal	monomer	stable
4mzd	0.12	xtal	monomer	stable
4n08	0.12	xtal	monomer	stable
4n1l	0.24	xtal	monomer	stable
4n3x	0.82	xtal	tetramer	stable
4n40	0.12	xtal	trimer	stable
4n4a	0.19	bio	dimer	stable
4n5a	0.12	xtal	monomer	stable
4n5c	0.12	xtal	monomer	stable
4n5q	0.15	xtal	monomer	stable
4n6l	0.20	xtal	monomer	stable
4n6t	0.12	xtal	monomer	stable
4n6u	0.12	xtal	monomer	stable
4n75	0.12	xtal	monomer	stable
4n7f	0.83	xtal	dimer	stable
4n8h	0.53	bio	monomer	stable
4n8k	0.58	bio	monomer	stable
4n8l	0.48	bio	monomer	stable
4nbo	0.53	xtal	monomer	stable
4nc9	0.43	bio	dimer	stable
4nci	0.65	xtal	monomer	stable
4ndl	0.12	xtal	monomer	stable
4nfb	0.12	xtal	monomer	stable
4nfc	0.50	xtal	monomer	stable
4nfx	1.00	bio	dimer	stable
4ng0	1.00	xtal	monomer	stable
4nj6	0.67	bio	dimer	uncertain
4nj7	0.12	xtal	monomer	stable
4nj8	0.12	xtal	monomer	stable

4njc	1.00	xtal	monomer	stable
4nk6	0.34	bio	dimer	stable
4nk8	0.77	bio	dimer	stable
4nl7	0.12	xtal	dimer	stable
4nlb	0.12	xtal	monomer	stable
4nng	0.12	xtal	monomer	stable
4nnh	0.27	xtal	monomer	stable
4nnk	0.12	xtal	monomer	stable
4noj	0.12	xtal	monomer	stable
4npf	0.86	bio	dimer	uncertain
4nq0	0.12	xtal	monomer	stable
4nsp	0.12	xtal	monomer	stable
4nu2	0.12	xtal	monomer	stable
4nv8	0.74	xtal	dimer	stable
4nx9	0.68	bio	monomer	stable
4nzz	0.12	xtal	monomer	stable
4o2x	0.32	xtal	monomer	stable
4o3x	0.12	xtal	monomer	stable
4o4c	0.41	xtal	monomer	stable
4o5p	0.12	xtal	monomer	stable
4o5s	0.12	xtal	monomer	stable
4o6g	0.13	xtal	monomer	stable
4o6k	0.20	xtal	monomer	stable
4o6x	0.83	xtal	dimer	uncertain
4o7q	0.12	xtal	monomer	stable
4o8s	0.13	xtal	monomer	stable
4o8v	0.12	xtal	monomer	stable
4o9f	0.12	xtal	monomer	stable
4o9l	0.19	xtal	monomer	stable
4oa6	0.12	xtal	monomer	stable
4ob4	0.70	bio	dimer	stable
4od4	0.12	xtal	dimer	stable
4od6	0.12	xtal	monomer	stable
4odd	0.76	xtal	monomer	stable
4of0	0.12	xtal	monomer	stable
4of5	0.12	bio	monomer	stable
4ofv	0.12	bio	monomer	stable
4ohd	0.12	xtal	monomer	stable
4ohe	0.12	xtal	monomer	stable
4ohh	0.12	xtal	monomer	stable
4ohi	0.12	xtal	monomer	stable
4ohl	0.12	xtal	monomer	stable
4ohr	1.00	bio	dimer	stable
4oj0	0.83	xtal	dimer	stable
4ok0	0.12	xtal	monomer	stable
4om9	0.12	xtal	monomer	stable
4oq1	0.12	xtal	monomer	stable
4oqw	1.00	bio	dimer	stable
4orw	0.12	xtal	monomer	stable
4os0	1.00	xtal	dimer	stable
4os3	1.00	xtal	dimer	stable
4os8	1.00	xtal	dimer	stable
4otd	0.12	xtal	monomer	stable

4ou0	0.22	xtal	monomer	stable
4ow6	0.12	xtal	monomer	stable
4oyl	0.12	xtal	trimer	uncertain
4oyy	0.12	xtal	trimer	stable
4ozk	0.12	xtal	monomer	stable
4p09	0.12	xtal	monomer	stable
4p0f	0.36	xtal	dimer	stable
4p0j	0.63	xtal	monomer	stable
4p0k	0.79	xtal	monomer	stable
4p0l	0.20	xtal	monomer	stable
4p0m	0.19	xtal	hexamer	stable
4p0o	0.41	xtal	monomer	stable
4p0y	0.12	xtal	monomer	stable
4p0z	0.12	xtal	monomer	stable
4p17	0.13	xtal	monomer	stable
4p39	0.40	bio	tetramer	stable
4p47	0.12	xtal	monomer	stable
4p48	0.12	xtal	monomer	stable
4p5u	0.12	xtal	monomer	stable
4p7o	0.16	xtal	dimer	stable
4pa1	0.67	xtal	dimer	stable
4pbd	0.58	xtal	monomer	stable
4pcf	0.12	bio	trimer	uncertain
4pd3	0.23	bio	monomer	stable
4pej	0.60	bio	dimer	stable
4pek	0.12	bio	monomer	stable
4pgr	0.12	xtal	monomer	stable
4pgs	0.12	xtal	dimer	stable
4pgu	0.15	xtal	monomer	stable
4pgv	0.13	xtal	monomer	stable
4ph1	0.73	xtal	trimer	stable
4pjq	0.20	bio	dimer	stable
4pk1	0.12	xtal	monomer	stable
4pk9	0.78	bio	dimer	stable
4pka	0.81	bio	dimer	stable
4pkx	0.12	xtal	monomer	stable
4pmh	0.12	xtal	monomer	stable
4pn7	0.12	bio	monomer	stable
4po4	0.12	xtal	monomer	stable
4poy	0.12	xtal	monomer	stable
4ppj	0.12	bio	monomer	stable
4ppk	0.12	bio	monomer	stable
4ppl	0.12	xtal	monomer	stable
4pqo	0.12	xtal	monomer	stable
4pqz	0.12	xtal	dimer	uncertain
4psd	0.12	xtal	monomer	stable
4psf	0.23	bio	monomer	stable
4pug	0.12	bio	dimer	stable
4pui	0.12	xtal	monomer	stable
4pwq	0.13	xtal	dimer	uncertain
4pzp	0.12	xtal	monomer	stable
4q2s	0.12	xtal	monomer	stable
4q2y	0.23	xtal	monomer	stable

4q3h	0.85	xtal	monomer	stable
4q4g	0.12	xtal	monomer	stable
4q4n	0.12	xtal	monomer	stable
4q5y	0.23	xtal	dimer	stable
4q65	0.12	xtal	monomer	stable
4q6l	0.12	xtal	monomer	stable
4q6v	0.12	xtal	monomer	stable
4q6z	0.12	xtal	dimer	uncertain
4q7k	0.12	xtal	monomer	stable
4q7m	0.84	bio	dimer	stable
4q7t	0.12	xtal	monomer	stable
4q9h	0.12	xtal	monomer	stable
4q9p	0.68	xtal	tetramer	uncertain
4qaj	0.60	xtal	monomer	stable
4qaz	0.13	xtal	monomer	stable
4qb0	0.12	xtal	monomer	stable
4qbe	0.43	xtal	monomer	stable
4qcj	0.31	bio	monomer	stable
4qdr	0.12	xtal	monomer	stable
4qdt	0.12	xtal	monomer	stable
4qdu	0.12	xtal	monomer	stable
4qgy	0.12	bio	monomer	stable
4qlo	0.80	bio	dimer	stable
4qlr	0.15	xtal	monomer	stable
4qmd	0.12	xtal	monomer	stable
4qmi	0.12	xtal	monomer	stable
4qmj	0.12	xtal	monomer	stable
4qs8	0.12	xtal	monomer	stable
4qt4	0.12	xtal	monomer	stable
4qtz	0.12	xtal	monomer	stable
4quc	0.12	xtal	monomer	stable
4quk	0.12	xtal	monomer	stable
4qx0	0.28	xtal	monomer	stable
4qx1	0.20	xtal	monomer	stable
4qx2	0.27	xtal	monomer	stable
4qx3	0.24	xtal	monomer	stable
4qxl	0.12	xtal	monomer	stable
4r05	0.12	xtal	monomer	stable
4r0e	0.48	xtal	monomer	stable
4r0x	0.12	xtal	dimer	stable
4r0y	0.12	xtal	monomer	stable
4r19	0.12	xtal	monomer	stable
4r1a	0.12	xtal	monomer	stable
4r1b	0.12	xtal	monomer	stable
4r1c	0.12	xtal	monomer	stable
4r58	0.12	xtal	monomer	stable
4r61	0.49	xtal	monomer	stable
4r6f	0.12	xtal	monomer	stable
4r7v	0.18	bio	dimer	stable
4r8r	0.12	xtal	monomer	stable
4rcw	0.40	bio	dimer	uncertain
4rdp	0.82	bio	monomer	stable
4reh	0.12	xtal	monomer	stable

4rej	0.12	xtal	monomer	stable
4rf6	0.12	xtal	monomer	stable
4rid	0.18	bio	tetramer	stable
4rl1	0.12	xtal	monomer	stable
4rl9	0.12	xtal	tetramer	stable
4rlb	0.12	bio	tetramer	stable
4rlm	0.12	xtal	monomer	stable
4rln	0.12	bio	monomer	stable
4ro1	0.12	xtal	monomer	stable
4roa	0.45	xtal	monomer	stable
4rs6	0.82	bio	hexamer	stable
4rsq	0.46	xtal	monomer	stable
4rsx	0.12	xtal	monomer	stable
4rtd	0.13	xtal	monomer	stable
4rth	0.49	xtal	monomer	stable
4rwi	0.66	xtal	dimer	stable
4rwq	0.12	xtal	monomer	stable
4rwr	0.78	xtal	monomer	stable
4rwu	0.60	xtal	monomer	stable
4rwz	0.79	xtal	monomer	stable
4ry4	0.12	xtal	monomer	stable
4ry6	0.12	xtal	monomer	stable
4ry7	0.12	xtal	monomer	stable
4rz2	0.12	xtal	monomer	stable
4s11	0.12	xtal	monomer	stable
4s3l	0.15	xtal	monomer	stable
4s3p	0.60	xtal	tetramer	stable
4tks	0.58	bio	monomer	stable
4tlp	0.12	xtal	monomer	stable
4tlw	0.12	xtal	monomer	stable
4tn9	0.12	xtal	monomer	stable
4tnm	0.12	xtal	monomer	stable
4tql	0.12	xtal	monomer	stable
4tr5	1.00	bio	monomer	stable
4trk	0.12	xtal	monomer	stable
4tse	0.44	xtal	dimer	stable
4ttl	0.12	xtal	nopred	n/a
4ttp	0.12	xtal	monomer	stable
4tum	0.27	xtal	monomer	stable
4tut	0.80	xtal	monomer	stable
4tuu	0.12	xtal	monomer	stable
4tve	0.15	xtal	monomer	stable
4tx2	0.12	xtal	monomer	stable
4txp	0.12	xtal	monomer	stable
4tza	0.24	xtal	monomer	stable
4tzg	0.80	xtal	monomer	stable
4tzi	0.12	xtal	monomer	stable
4tzj	0.12	xtal	monomer	stable
4u0l	0.12	xtal	monomer	stable
4u1d	0.16	xtal	monomer	stable
4u2m	0.34	xtal	tetramer	stable
4u2n	1.00	xtal	dimer	stable
4u3a	0.92	bio	dimer	stable

4u3h	0.12		xtal	monomer	stable
4u3t	0.15		xtal	monomer	stable
4u3v	0.12		xtal	monomer	stable
4u48	0.39		bio	monomer	stable
4u4h	0.13		bio	dimer	stable
4u4j	0.42		bio	monomer	stable
4u59	0.42		bio	monomer	stable
4u5a	0.12		xtal	monomer	stable
4u9q	0.46		xtal	dimer	uncertain
4ud5	0.12		xtal	dimer	stable
4uex	0.12		bio	monomer	stable
4ulw	0.83		xtal	monomer	stable
4umg	0.24		bio	dimer	uncertain
4ums	0.20		xtal	monomer	stable
4uos	0.40		xtal	monomer	stable
4uts	0.12		xtal	monomer	stable
4uu4	0.34		bio	monomer	stable
4uv2	0.76		bio	16-mer	stable
4uv6	0.74		xtal	monomer	stable
4uvk	0.12		xtal	monomer	stable
4uw2	0.12		xtal	monomer	stable
4uww	1.00		xtal	monomer	stable
4uy3	0.12		xtal	monomer	stable
4uz2	0.19		xtal	octomer	stable
4v0t	0.83		bio	dimer	stable
4v14	0.12		xtal	monomer	stable
4v16	0.12		xtal	monomer	stable
4v2b	0.12		xtal	monomer	stable
4v2d	0.12		xtal	monomer	stable
4v2e	0.12		bio	monomer	stable
4v2u	0.12		xtal	monomer	stable
4v38	0.12		xtal	monomer	stable
4v39	0.12		xtal	monomer	stable
4w4m	0.12		bio	monomer	stable
4w5x	0.35		xtal	dimer	stable
4w60	0.54		xtal	dimer	stable
4w6v	0.12		xtal	monomer	stable
4w7g	0.12		xtal	monomer	stable
4w8v	1.00		bio	tetramer	stable
4w8w	0.82		bio	monomer	stable
4w8z	0.83		bio	dimer	stable
4wau	0.33		xtal	dimer	uncertain
4wbz	0.40		xtal	monomer	stable
4wdc	0.12		xtal	monomer	stable
4we2	0.12		xtal	monomer	stable
4wfi	0.12		xtal	monomer	stable
4wj1	0.50		bio	dimer	stable
4wj2	0.12		bio	monomer	stable
4wjs	0.12		xtal	monomer	stable
4wmg	0.12		bio	monomer	stable
4wo8	0.12		xtal	monomer	stable
4woy	0.59		bio	tetramer	stable
4wp6	0.12		xtal	monomer	stable

4wt3	0.12	xtal	monomer	stable
4wtx	0.16	xtal	monomer	stable
4wu0	0.27	xtal	monomer	stable
4wz0	0.62	xtal	dimer	stable
4x0j	0.12	xtal	monomer	stable
4x2m	0.12	xtal	monomer	stable
4x2w	0.59	bio	tetramer	stable
4x2x	0.18	bio	monomer	stable
4x2y	0.54	xtal	monomer	stable
4x3x	0.12	xtal	dimer	stable
4x43	0.12	bio	monomer	uncertain
4x8h	0.20	bio	monomer	stable
4x8m	0.12	xtal	monomer	stable
4x9u	0.12	xtal	monomer	stable
4xah	1.00	xtal	monomer	stable
4xeh	0.52	bio	dimer	stable
4xfu	0.12	xtal	dimer	stable
4xfv	0.12	xtal	dimer	uncertain
4xh3	0.12	bio	dimer	uncertain
4xhm	0.16	xtal	monomer	stable
4xkk	0.13	xtal	monomer	stable
4xm4	0.76	bio	dimer	stable
4xp8	0.34	xtal	monomer	stable
4xpk	0.12	xtal	monomer	stable
4xq7	0.12	xtal	monomer	stable
4xy3	0.12	xtal	monomer	stable
4xzg	0.12	xtal	monomer	stable
4y21	0.46	bio	monomer	stable
4y4q	0.24	xtal	dimer	stable
4y5j	0.33	bio	dimer	stable
4yeq	0.34	xtal	monomer	stable
4yf2	0.12	xtal	monomer	stable
4yfe	0.12	xtal	monomer	stable
4yfs	0.12	xtal	monomer	stable
4yhe	0.12	xtal	monomer	stable
4ykc	0.40	xtal	tetramer	stable
4ykd	0.12	xtal	monomer	stable
4yno	0.12	xtal	monomer	stable
4ynv	0.92	xtal	monomer	stable
4ynw	1.00	bio	dimer	stable
4ypj	0.12	xtal	monomer	stable
4yuo	0.12	xtal	monomer	stable
4yvm	0.38	xtal	dimer	stable
4yvo	0.12	xtal	monomer	stable
4z29	0.47	xtal	monomer	stable
4z99	0.12	xtal	monomer	stable
4z9p	0.72	xtal	monomer	stable
4zf3	0.32	xtal	monomer	stable
4zf4	0.40	xtal	monomer	stable
4zf5	0.54	xtal	monomer	stable
4zg1	0.12	xtal	hexamer	uncertain
4zj2	0.12	xtal	monomer	stable
4zj3	0.12	xtal	monomer	stable

4zmd	0.12	xtal	monomer	stable
4zot	0.12	xtal	monomer	stable
4zov	0.84	xtal	dimer	uncertain
4zoy	0.12	xtal	monomer	stable
4zuz	0.19	xtal	monomer	stable
4zwj	0.12	xtal	monomer	stable
5a0g	0.12	xtal	monomer	stable
5a0z	0.54	bio	tetramer	stable
5aem	0.18	xtal	monomer	stable
5af5	0.12	xtal	monomer	stable
5aio	0.81	xtal	monomer	stable
5aj8	0.12	xtal	monomer	stable
5avg	0.29	xtal	monomer	stable
5avh	0.79	bio	tetramer	stable
5brl	0.19	xtal	monomer	stable
5bth	0.28	bio	monomer	stable
5bvq	0.16	xtal	monomer	stable
5c00	0.83	bio	monomer	stable
5c86	0.12	xtal	monomer	stable

^a yellow: incorrect classification

^b Criteria for test selection are given in Table S9

Table S12. Interface area, number of near-native structures, and inferred probability of being a biological dimer for the proteins in the “difficult” subset of the test set^a

Biological dimers				Crystallographic dimers			
PDB ID	Interface area (Å ²)	N _{nn}	Probability (dimer)	PDB ID	Interface area (Å ²)	N _{nn}	Probability (dimer)
3w66	529.6	0	0.05	3w9g	604.3	8	0.17
3w8g	760.8	183	0.91	3wea	445.6	0	0.05
3wtz	690.6	74	0.75	3wp9	716.5	23	0.36
3wx4	1220.4	159	0.89	3wpj	566.9	0	0.05
3wyd	542.7	68	0.73	4c41	620.9	0	0.05
4chk	743.5	49	0.61	4cbm	1180.1	0	0.05
4ckn	1371.7	82	0.79	4cfi	628.2	0	0.05
4cop	489.0	131	0.87	4chj	930.5	144	0.88
4cw9	635.1	0	0.05	4d5r	790.3	109	0.85
4cwc	1505.3	52	0.63	4k51	874.1	0	0.05
4d28	1404.9	34	0.48	4l8i	827.8	0	0.05
4jgm	698.2	0	0.05	4l9h	1109.1	85	0.80
4jup	389.5	20	0.33	4lcm	814.0	0	0.05
4ke2	1787.5	4	0.11	4lhf	1034.8	148	0.88
4km3	1157.9	125	0.87	4lk4	595.0	9	0.18
4kp2	1756.8	41	0.54	4lso	524.0	0	0.05
4l2w	1072.2	47	0.59	4m7g	499.9	0	0.05
4lft	517.9	0	0.05	4mfi	957.9	20	0.33
4m0q	1449.4	83	0.79	4n3x	940.2	122	0.87
4m23	1547.5	86	0.80	4n40	654.7	0	0.05
4mdu	943.7	5	0.12	4n7f	626.1	142	0.88
4nj0	882.7	229	1.00	4n8h	1844.3	36	0.50
4nqf	1203.6	32	0.46	4nl7	1254.8	0	0.05
4nsr	1126.8	138	0.88	4npf	1404.4	182	0.91
4ny0	735.5	0	0.05	4nv8	837.6	73	0.75
4o2h	919.3	20	0.33	4nx9	1116.7	58	0.67
4oaf	1130.5	28	0.42	4o6x	616.0	151	0.88
4oky	811.4	26	0.40	4od4	650.6	0	0.05
4p08	1192.2	39	0.53	4of5	565.9	0	0.05
4p12	1514.5	37	0.51	4oj0	1524.5	139	0.88
4p6l	535.0	1	0.06	4os0	1024.8	452	1.00
4pu7	1034.3	282	1.00	4oyl	381.5	0	0.05
4pus	653.0	0	0.05	4oyy	539.1	0	0.05
4q9b	727.8	140	0.88	4p0f	1184.7	19	0.32
4r27	890.3	0	0.05	4p0m	844.4	5	0.12
4ra8	670.7	2	0.08	4p39	783.3	22	0.35
4rd9	1799.3	77	0.77	4p7o	1213.8	3	0.09
4rdj	1762.0	90	0.81	4pa1	1179.0	57	0.66
4ryj	476.7	0	0.05	4pcf	657.7	0	0.05
4s03	851.7	49	0.61	4pek	665.9	0	0.05
4u3q	651.0	0	0.05	4pgs	644.6	0	0.05
4u4i	925.8	11	0.21	4pn7	702.5	0	0.05
4u5h	590.7	11	0.21	4ppj	977.6	0	0.05

4u64	1516.7	129	0.87	4pqz	377.1	0	0.05
4uvj	1053.6	13	0.24	4psf	595.2	8	0.17
4v1p	637.6	187	0.91	4pug	826.8	0	0.05
4w6g	903.5	138	0.88	4pwq	802.1	1	0.06
4w6k	770.7	78	0.77	4q5y	1028.2	8	0.17
4wnz	1286.8	53	0.64	4q6z	625.4	0	0.05
4xoh	1083.9	6	0.14	4q9p	775.3	58	0.67
4zqa	1794.3	58	0.67	4qcj	472.0	14	0.25
5a2n	1096.1	0	0.05	4qgy	573.2	0	0.05
5a48	778.2	1	0.06	4r7v	757.5	4	0.11

4rdp	1381.7	111	0.85
4rid	1471.5	4	0.11
4rl9	506.7	0	0.05
4rln	568.5	0	0.05
4rs6	869.0	122	0.87
4rwi	1184.5	55	0.65
4tks	1025.6	43	0.56
4tr5	1340.3	222	1.00
4tse	1057.9	26	0.40
4u2n	698.3	224	1.00
4u48	4122.1	21	0.34
4u9q	724.9	28	0.42
4ud5	886.9	0	0.05
4uex	709.7	0	0.05
4uu4	881.4	17	0.29
4uz2	739.0	5	0.12
4w5x	437.2	18	0.30
4w60	399.3	37	0.51
4w8w	1381.4	116	0.86
4wau	818.2	16	0.28
4wj2	595.3	0	0.05
4wmg	599.5	0	0.05
4woy	1029.5	44	0.57
4wz0	701.2	48	0.60
4x2w	1095.6	44	0.57
4x43	643.0	0	0.05
4x8h	974.2	6	0.14
4ynw	651.0	246	1.00
5bth	695.4	12	0.22
5c00	915.8	149	0.88

Average	994.3	67.2	0.52	Average	934.1	47.1	0.40
Std Dev	391.7	81.5	0.31	Std Dev	618.9	75.5	0.30

^a Subset of the test set in Table S9, containing only proteins' results from EPPIC and PISA are conflicting, thus one considers the dimer biological the other crystallographic, or the classification by PISA is uncertain.

Table S13. Classification of difficult set dimers by three methods^{a,b}

PDB ID	ClusPro probability of dimer	EPPIC	PISA multimer class	PISA status
3w66	0.12	xtal	dimer	uncertain
3w8g	0.87	xtal	dimer	stable
3wtz	0.74	xtal	dimer	stable
3wx4	0.84	xtal	dimer	stable
3wyd	0.72	xtal	dimer	stable
4chk	0.62	bio	monomer	stable
4ckn	0.77	xtal	dimer	stable
4cop	0.83	xtal	dimer	uncertain
4cw9	0.12	xtal	dimer	stable
4cwc	0.64	bio	tetramer	stable
4d28	0.51	xtal	dimer	stable
4jgm	0.12	bio	dimer	uncertain
4jup	0.38	xtal	dimer	uncertain
4ke2	0.18	xtal	hexamer	stable
4km3	0.83	xtal	dimer	stable
4kp2	0.57	bio	dimer	stable
4l2w	0.61	xtal	tetramer	stable
4lft	0.12	bio	monomer	stable
4m0q	0.77	xtal	dimer	stable
4m23	0.78	bio	dimer	stable
4mdu	0.19	xtal	dimer	stable
4nj0	1.00	xtal	dimer	stable
4nqf	0.49	xtal	dimer	stable
4nsr	0.83	xtal	dimer	stable
4ny0	0.12	xtal	dimer	stable
4o2h	0.38	xtal	dimer	stable
4oaf	0.46	bio	monomer	stable
4oky	0.44	xtal	dimer	stable
4p08	0.55	xtal	dimer	stable
4p12	0.54	bio	dimer	stable
4p6l	0.13	xtal	dimer	stable
4pu7	1.00	xtal	dimer	stable
4pus	0.12	xtal	dimer	uncertain
4q9b	0.83	xtal	dimer	stable
4r27	0.12	bio	monomer	stable
4ra8	0.15	xtal	dimer	stable
4rd9	0.75	bio	dimer	stable
4rdj	0.79	xtal	dimer	stable
4ryj	0.12	xtal	dimer	stable
4s03	0.62	xtal	dimer	uncertain
4to7	1.00	xtal	monomer	stable
4tqj	1.00	xtal	monomer	stable

4u3q	0.12	xtal	dimer	uncertain
4u4i	0.27	bio	monomer	stable
4u5h	0.27	xtal	dimer	stable
4u64	0.83	xtal	dimer	stable
4uvj	0.29	xtal	dimer	stable
4v1p	0.87	xtal	dimer	stable
4w6g	0.83	bio	monomer	stable
4w6k	0.76	xtal	dimer	uncertain
4wnz	0.65	bio	monomer	stable
4xoh	0.20	bio	dimer	stable
4zbx	0.12	xtal	monomer	stable
4zqa	0.68	xtal	dimer	stable
5a2n	0.12	xtal	dimer	stable
5a48	0.13	bio	dimer	uncertain

^a Yellow: incorrect classification

^b Criteria for the selection of the “difficult” subset are given in Table S12

Table S14. Classification of test set monomers by three methods^{a,b}

PDB ID	ClusPro probability of dimer	EPPIC	PISA multimer class	PISA status
3w9g	0.23	xtal	dimer	stable
3wea	0.12	xtal	dimer	stable
3wp9	0.41	xtal	dimer	stable
3wpj	0.12	bio	monomer	stable
4c41	0.12	bio	monomer	stable
4cbm	0.12	bio	dimer	uncertain
4cfi	0.12	bio	monomer	stable
4chj	0.83	xtal	tetramer	uncertain
4d5r	0.81	bio	dimer	stable
4k51	0.12	xtal	trimer	stable
4l8i	0.12	xtal	dimer	stable
4l9h	0.78	bio	monomer	stable
4lcm	0.12	xtal	dimer	stable
4lhf	0.83	bio	monomer	stable
4lk4	0.24	bio	monomer	stable
4lso	0.12	bio	monomer	stable
4m7g	0.12	bio	monomer	stable
4mfi	0.38	bio	monomer	stable
4n3x	0.82	xtal	tetramer	stable
4n40	0.12	xtal	trimer	stable
4n7f	0.83	xtal	dimer	stable
4n8h	0.53	bio	monomer	stable
4nl7	0.12	xtal	dimer	stable
4npf	0.86	bio	dimer	uncertain
4nv8	0.74	xtal	dimer	stable
4nx9	0.68	bio	monomer	stable
4o6x	0.83	xtal	dimer	uncertain
4od4	0.12	xtal	dimer	stable
4of5	0.12	bio	monomer	stable
4oj0	0.83	xtal	dimer	stable
4os0	1.00	xtal	dimer	stable
4oyl	0.12	xtal	trimer	uncertain
4oyy	0.12	xtal	trimer	stable
4p0f	0.36	xtal	dimer	stable
4p0m	0.19	xtal	hexamer	stable
4p39	0.40	bio	tetramer	stable
4p7o	0.16	xtal	dimer	stable
4pa1	0.67	xtal	dimer	stable
4pcf	0.12	bio	trimer	uncertain
4pek	0.12	bio	monomer	stable
4pgs	0.12	xtal	dimer	stable
4pn7	0.12	bio	monomer	stable
4ppj	0.12	bio	monomer	stable

4pqz	0.12	xtal	dimer	uncertain
4psf	0.23	bio	monomer	stable
4pug	0.12	bio	dimer	stable
4pwq	0.13	xtal	dimer	uncertain
4q5y	0.23	xtal	dimer	stable
4q6z	0.12	xtal	dimer	uncertain
4q7m	0.84	bio	dimer	stable
4q9p	0.68	xtal	tetramer	uncertain
4qcj	0.31	bio	monomer	stable
4qgy	0.12	bio	monomer	stable
4r7v	0.18	bio	dimer	stable
4rdp	0.82	bio	monomer	stable
4rid	0.18	bio	tetramer	stable
4rl9	0.12	xtal	tetramer	stable
4rln	0.12	bio	monomer	stable
4rs6	0.82	bio	hexamer	stable
4rwi	0.66	xtal	dimer	stable
4tks	0.58	bio	monomer	stable
4tr5	1.00	bio	monomer	stable
4tse	0.44	xtal	dimer	stable
4u2n	1.00	xtal	dimer	stable
4u48	0.39	bio	monomer	stable
4u9q	0.46	xtal	dimer	uncertain
4ud5	0.12	xtal	dimer	stable
4uex	0.12	bio	monomer	stable
4uu4	0.34	bio	monomer	stable
4uz2	0.19	xtal	octomer	stable
4w5x	0.35	xtal	dimer	stable
4w60	0.54	xtal	dimer	stable
4w8w	0.82	bio	monomer	stable
4wau	0.33	xtal	dimer	uncertain
4wj2	0.12	bio	monomer	stable
4wmg	0.12	bio	monomer	stable
4woy	0.59	bio	tetramer	stable
4wz0	0.62	xtal	dimer	stable
4x2w	0.59	bio	tetramer	stable
4x43	0.12	bio	monomer	uncertain
4x8h	0.20	bio	monomer	stable
4xm4	0.76	bio	dimer	stable
4ynw	1.00	bio	dimer	stable
5avh	0.79	bio	tetramer	stable
5bth	0.28	bio	monomer	stable
5c00	0.83	bio	monomer	stable

^a Yellow: incorrect classification

^b Criteria for the selection of the “difficult” subset are given in Table S12

Table S15. Impact of small interface on success rates^a

Test set (783)							
Dimers							
Interface area (Å ²)	Cases	ClusPro	ClusPro %	EPPIC	EPPIC %	PISA	PISA %
<600	17	4	23.5%	4	23.5%	4	23.5%
<800	53	26	49.1%	20	37.7%	20	37.7%
<1000	88	48	54.5%	37	42.0%	34	38.6%
Monomers							
Interface area (Å ²)	Cases	ClusPro	ClusPro %	EPPIC	EPPIC %	PISA	PISA %
<600	183	173	94.5%	165	90.2%	172	94.0%
<800	333	299	89.8%	294	88.3%	291	87.4%
<1000	399	351	88.0%	344	86.2%	335	84.0%
Total							
Interface area (Å ²)	Cases	ClusPro	ClusPro %	EPPIC	EPPIC %	PISA	PISA %
<600	200	177	88.5%	169	84.5%	176	88.0%
<800	386	325	84.2%	314	81.3%	311	80.6%
<1000	487	399	81.9%	381	78.2%	369	75.8%

^b Criteria for test selection are given in Table S9

Table S16. Transitional complexes: Interface area, number of near-native structures, and inferred probability of being a biological dimer for the proteins in the Nooren set^a

PDB ID	Interface area (Å ²)	N _{nn}	Probability (dimer)
1a15	739.9	0	0.12
1a78	512.7	99	0.80
1bbh	765.0	323	1.00
1beb	552.5	112	0.82
1bfs	685.1	267	1.00
1bft	688.1	275	1.00
1bi2	894.5	123	0.82
1c0w	862.3	96	0.80
1cmi	883.6	43	0.58
1cnt	981.2	62	0.70
1cpq	599.6	0	0.12
1dom	845.8	99	0.80
1edh	900.3	30	0.48
1ikn	1141.9	316	1.00
1nfk	683.4	32	0.49
1ram	896.2	80	0.76
1scf	828.4	13	0.29
1trz	793.8	19	0.36
1vqx	1068.0	83	0.77
2dtr	876.1	164	0.84
2lyn	946.1	33	0.50
3il8	918.6	33	0.50
5cro	638.1	126	0.83
6cro	1821.4	244	1.00
Average	855.1	111.3	0.68
Std Dev	257.9	101.3	0.27

^aTransitional complexes from Nooren and Thornton [1]

Table S17. Biological activity and classification of Nooren set dimers by three methods^a

Protein	PDB ID	Biological activity		ClusPro probability of dimer	EPPIC	PISA multimer class	PISA status
		Mono-mer	Dimer				
Lambda phage cro repressor	5cro	-	+	0.83	xtal	24-mer	stable
	6cro			1	bio	nopred	n/a
Corynebacterium diphtheria tox repressor	1bi2	-	+	0.82	bio	dimer	stable
	2dtr			0.84	bio	dimer	stable
	1c0w			0.8	bio	nopred	n/a
	1bfs	-	+	1	bio	monomer	stable
Mouse nuclear factor κB (NFκB) P50 and P65	1bft			1	bio	monomer	stable
	1ikn			1	bio	trimer	stable
	1nfk			0.49	xtal	tetramer	stable
	1ram			0.76	bio	tetramer	stable
	1vkx			0.77	bio	tetramer	stable
Human interleukin-8 (IL-8), type I	3il8	+	+	0.5	bio	dimer	stable
Human stromal cell-derived factor 1α (SDF1α), type I	1a15	+	?	0.12	xtal	dimer	stable
Human monocyte chemoattractant protein 1 (MCP-1), type II	1dom	+	+	0.8	xtal	nopred	n/a
Human stem cell factor cytokine (SCF) (secreted soluble form)	1scf	?	+	0.29	bio	tetramer	stable
Human ciliary neurotrophic factor (CNTF)	1cnt	+	?	0.7	xtal	dimer	stable
Human insulin	1trz	+	?	0.36	bio	dodecamer	stable
Chromatium vinosum cytochrome c' type Ie (CVCP)	1bbh	+	?	1	bio	dimer	stable
Rhodobacter capsulatus cytochrome c' type IIe (RCCP)	1cpq	+	?	0.12	xtal	monomer	stable
Cow β-lactoglobulin	1beb	+	+	0.82	bio	monomer	stable
California red abalone sperm lysin	2lyn	+	+	0.5	xtal	tetramer	stable
Mouse epithelial cadherin, first two N-terminal extracellular domain (ECAD12)	1edh	?	+	0.48	bio	monomer	stable
Human LC8 light chain subunit of dynein complex	1cmi	?	+	0.58	bio	tetramer	stable
Toad ovary galectin-1	1a78	?	?	0.8	xtal	dimer	uncertain

^a Transitional complexes from Nooren and Thornton [1]

Table S18. Classification of the Nooren set dimers by three methods: Summary

method	pisa	eppic	cluspro
dimers (24)	17	16	18
dimers %	70.8	66.7%	75.0%

^a Transitional complexes from Nooren and Thornton [1]

Table S19. Transitional complexes: Interface area, number of near-native structures, and inferred probability of being a biological dimer for the proteins in the weak homodimer set selected by Dey et al. [2]

PDB ID	Interface area (Å ²)	N _{nn}	Probability (dimer)
1a15	739.9	0	0.12
1a45	552.5	1	0.13
1a78	512.7	99	0.80
1b8e	648.1	1	0.13
1bbh	765.0	323	1.00
1bfs	685.1	267	1.00
1cbi	962.9	4	0.18
1cnt	981.2	62	0.70
1cxq	754.5	7	0.22
1dok	969.5	141	0.83
1dto	983.3	0	0.12
1dv8	405.3	0	0.12
1e87	866.0	163	0.84
1eaj	675.2	0	0.12
1edh	900.3	30	0.48
1eeq	679.5	143	0.83
1erv	511.7	317	1.00
1exz	846.3	4	0.18
1gml	786.9	33	0.50
1h65	1292.2	172	0.85
1hgx	1046.1	43	0.58
1ihk	1204.9	140	0.83
1ii5	1014.6	22	0.40
1itv	644.8	75	0.75
1jr9	854.4	142	0.83
1lwj	1318.8	119	0.82
1neu	475.2	0	0.12
1nrv	397.6	0	0.12
1o7z	613.5	0	0.12
1oal	754.5	98	0.80
1p9y	1223.9	30	0.48
1py9	903.7	0	0.12
1sj1	367.4	0	0.12
1tip	922.3	132	0.83
1uq5	911.8	14	0.31
1zop	1068.3	175	0.85
2dtr	876.1	164	0.84
2lyn	946.1	33	0.50
3bri	823.7	150	0.83
3il8	918.6	33	0.50
3sdh	879.3	104	0.81
5cro	638.1	126	0.83
Average	817.2	80.2	0.54
Std Dev	234.1	88.0	0.32

Table S20. Interface area, number of near-native structures, and inferred probability of being a biological dimer for the proteins in the “others” homodimer set by Dey et al. [2]

PDB ID	Interface area (Å ²)	Nnn	Probability (dimer)
12as	1878.1	156	0.84
1a3c	989.1	173	0.85
1a4i	1354.6	114	0.82
1aa7	1093.8	46	0.60
1ad3	3940.1	39	0.55
1ade	2699.0	107	0.81
1af5	851.3	0	0.12
1afw	2388.7	123	0.82
1ajs	3436.9	74	0.74
1aor	1209.9	19	0.36
1aoz	1816.2	57	0.67
1aq6	2220.5	176	0.86
1auo	663.7	0	0.12
1b16	2518.4	176	0.86
1b3a	733.0	200	0.91
1b5e	2527.4	117	0.82
1b67	1595.9	283	1.00
1b8a	4375.3	103	0.81
1b8j	3790.5	94	0.79
1b8z	1889.6	242	1.00
1b9h	2858.4	125	0.83
1bam	750.0	128	0.83
1bd0	3042.4	98	0.80
1bdm	1588.2	259	1.00
1bif	889.7	6	0.20
1biq	3012.6	99	0.80
1bis	1478.5	205	0.93
1bk5	1055.9	6	0.20
1bkp	2206.0	142	0.83
1brw	1060.5	249	1.00
1bsl	1893.9	158	0.84
1bsr	1895.4	277	1.00
1buo	1927.6	165	0.84
1bxg	1039.6	71	0.73
1byf	844.0	104	0.81
1byi	1398.3	223	1.00
1byk	1514.5	76	0.75
1c4k	6254.2	36	0.53
1cd8	1025.3	178	0.86
1chm	3230.1	70	0.73
1cmb	1796.3	481	1.00
1cnz	2444.9	107	0.81
1coz	1053.1	192	0.89
1csh	4995.4	61	0.69
1ctt	1933.3	134	0.83

1czj	741.3	7	0.22
1d0c	2753.6	92	0.79
1d2f	1757.3	103	0.81
1d9c	3633.0	82	0.77
1daa	2301.1	181	0.86
1db3	1140.6	118	0.82
1dbq	1532.4	56	0.66
1dcf	1057.3	70	0.73
1dor	2196.2	130	0.83
1dos	2333.0	145	0.83
1dpg	2278.7	167	0.84
1dqn	1724.8	88	0.78
1dqs	1604.6	87	0.78
1dxg	709.8	340	1.00
1e1z	1235.7	17	0.34
1e4m	946.6	85	0.78
1e58	725.7	219	0.99
1e6b	1144.8	52	0.64
1e6u	1499.9	87	0.78
1e7n	811.9	247	1.00
1ebh	1781.7	115	0.82
1ecy	1562.5	63	0.70
1elu	2940.7	85	0.78
1ew3	1024.9	218	0.99
1ex0	2414.9	52	0.64
1eyv	1067.3	183	0.87
1f3h	524.5	346	1.00
1f6b	743.7	0	0.12
1f89	1407.8	134	0.83
1fc4	4891.9	65	0.71
1fg7	3227.5	73	0.74
1fi4	1373.5	140	0.83
1fip	1578.0	209	0.94
1fjj	873.0	134	0.83
1fp1	4526.7	86	0.78
1fp2	4159.7	80	0.76
1fp5	982.5	96	0.80
1fro	3575.0	96	0.80
1g5c	2129.7	231	1.00
1g60	1926.3	141	0.83
1g8l	2120.0	80	0.76
1gd9	2637.7	154	0.84
1glj	1588.3	258	1.00
1gnw	1156.5	102	0.81
1gqi	1935.6	50	0.63
1gt3	2338.9	171	0.85
1gta	1504.9	96	0.80
1gte	11367.4	11	0.27
1gu7	894.7	188	0.88
1gvj	1194.9	120	0.82
1gvp	905.1	192	0.89
1gy6	1241.9	175	0.85
1gyo	610.9	0	0.12

1h16	1890.0	157	0.84
1h1y	945.2	119	0.82
1h4v	3708.2	69	0.73
1h7e	1361.9	173	0.85
1h9r	1386.3	66	0.71
1hf2	1041.6	208	0.94
1hjr	962.0	83	0.77
1hpx	1731.3	273	1.00
1hq8	1016.7	151	0.83
1hss	1095.5	64	0.70
1hw1	1593.0	138	0.83
1hxp	3398.7	113	0.82
1hzj	976.3	188	0.88
1i07	1960.6	201	0.91
1i24	2423.5	101	0.81
1i2k	2225.1	175	0.85
1i4u	1285.5	66	0.71
1i52	1822.5	126	0.83
1i69	1069.1	58	0.68
1i86	2974.1	75	0.75
1imb	1648.7	185	0.87
1is3	748.3	66	0.71
1iup	987.0	153	0.83
1ivy	1586.2	141	0.83
1ix9	907.8	115	0.82
1iyz	948.2	68	0.72
1j3b	973.1	0	0.12
1j5p	2146.1	172	0.85
1j98	1869.9	125	0.83
1jd0	1121.6	0	0.12
1jdw	1473.4	101	0.81
1jhd	1547.2	44	0.59
1jhg	2225.9	349	1.00
1jm6	1096.3	7	0.22
1jsc	2200.4	113	0.82
1jsg	789.3	146	0.83
1jtv	2603.8	141	0.83
1jye	1528.6	88	0.78
1jys	1249.1	162	0.84
1k04	2356.6	107	0.81
1k3y	1405.9	104	0.81
1k51	2340.5	133	0.83
1k66	1248.4	191	0.88
1k94	1820.6	272	1.00
1kc7	2030.4	105	0.81
1kcm	1510.7	148	0.83
1kic	942.7	0	0.12
1kjq	2304.9	99	0.80
1kkj	3700.6	115	0.82
1kll	2050.8	220	1.00
1kly	1642.4	226	1.00
1knq	1029.8	165	0.84
1kpf	1855.0	198	0.90

1kso	1680.6	291	1.00
1kta	2339.6	159	0.84
1ktb	989.9	2	0.15
1kzh	2096.1	103	0.81
1lc5	1943.0	231	1.00
1llf	681.6	0	0.12
1lom	3217.1	40	0.56
1lxe	2444.8	61	0.69
1m6p	1017.7	99	0.80
1m7y	2701.7	118	0.82
1mh9	1132.0	137	0.83
1mkb	1598.0	124	0.82
1mo9	4712.5	95	0.80
1mor	2528.4	97	0.80
1mqi	787.9	24	0.42
1mr8	1289.3	387	1.00
1mvc	1084.3	64	0.70
1mze	1614.9	416	1.00
1n1b	1177.6	212	0.96
1n26	1112.5	3	0.16
1n2s	1374.1	155	0.84
1n55	1397.6	189	0.88
1n57	945.8	162	0.84
1n7f	633.1	44	0.59
1nf9	1320.5	230	1.00
1nkd	1377.7	154	0.84
1nn5	769.8	232	1.00
1noj	2540.4	79	0.76
1nox	2966.1	122	0.82
1ns5	1107.9	214	0.97
1nsj	1420.5	326	1.00
1nsy	2583.3	120	0.82
1o8b	1077.0	170	0.85
1oaa	1932.7	137	0.83
1oac	7135.2	16	0.33
1oc2	1300.9	132	0.83
1oh0	1038.7	89	0.79
1oi6	1646.7	139	0.83
1oke	1929.2	12	0.28
1on2	1257.2	120	0.82
1oqz	990.0	9	0.24
1ov3	2449.9	143	0.83
1oyb	735.5	44	0.59
1oys	1387.3	127	0.83
1p5f	1339.3	156	0.84
1pfq	2176.4	89	0.79
1pre	2280.7	33	0.50
1psr	1199.3	291	1.00
1puc	2148.8	199	0.91
1pv9	1663.1	58	0.68
1pxf	2202.5	243	1.00
1q3o	604.3	0	0.12
1q4g	2503.8	30	0.48

1q6o	1444.6	71	0.73
1qfh	2275.0	55	0.66
1qhi	1698.8	103	0.81
1qi9	5330.3	36	0.53
1qj4	844.8	102	0.81
1qjb	973.9	101	0.81
1qkk	923.9	51	0.64
1qks	1288.6	28	0.46
1qo8	1233.0	22	0.40
1qr2	1917.0	130	0.83
1qvz	549.1	0	0.12
1qyc	1146.8	300	1.00
1r2f	1741.9	127	0.83
1r31	4779.1	88	0.78
1r3s	1233.5	78	0.76
1r7a	958.3	103	0.81
1rdo	663.4	294	1.00
1reg	676.5	0	0.12
1rrm	1153.9	95	0.80
1ryt	1638.5	152	0.83
1s0a	5189.6	85	0.78
1s3e	2042.9	23	0.41
1sb8	1176.5	82	0.77
1se8	1097.1	19	0.36
1ses	2229.6	107	0.81
1sjy	1526.2	140	0.83
1smn	868.4	164	0.84
1smt	1969.4	278	1.00
1sox	1410.1	87	0.78
1t0i	1317.6	261	1.00
1t5h	1329.7	48	0.62
1tc1	1505.7	92	0.79
1tox	3703.5	86	0.78
1trb	2735.2	170	0.85
1trk	4418.3	54	0.65
1tw3	3376.7	177	0.86
1ual	3517.1	110	0.82
1uby	2162.4	135	0.83
1ui8	6595.9	8	0.23
1ukg	838.3	25	0.43
1utg	1464.2	177	0.86
1uyr	5132.3	10	0.26
1v25	2850.8	129	0.83
1v2z	841.3	142	0.83
1vfr	3355.7	73	0.74
1vhz	2820.5	148	0.83
1vls	1254.6	44	0.59
1vok	1582.6	159	0.84
1xso	666.8	102	0.81
2arc	810.3	83	0.77
2gsa	4576.7	85	0.78
2hdh	1604.4	260	1.00
2ilk	4479.7	43	0.58

2mjp	950.0	252	1.00
2nac	3793.4	180	0.86
2ohx	1685.1	339	1.00
2spc	2513.5	57	0.67
2tct	2612.3	137	0.83
2tgi	1260.4	145	0.83
3dap	2653.4	154	0.84
3lad	3364.1	120	0.82
3ssi	838.4	115	0.82
3vub	1003.5	296	1.00
4kbp	1467.9	125	0.83
5csm	2000.1	38	0.54
5rub	2829.1	99	0.80
8prk	968.0	239	1.00
9wga	2298.4	151	0.83
Average	1922.8	125.9	0.75
Std Dev	1260.2	83.2	0.22

Table S21. Classification of the weak set of dimers by Dey et al. [2]

PDB ID	Kd (μM)	ClusPro probability of dimer	EPPIC	PISA multimer class	PISA status
1a15	150	0.12	xtal	dimer	stable
1a45	—	0.13	xtal	monomer	stable
1a78	7	0.80	xtal	dimer	uncertain
1b8e	5–200	0.13	bio	monomer	stable
1bbh	2000	1.00	bio	dimer	stable
1bfs	0.7	1.00	bio	monomer	stable
1cbi	> 100	0.18	xtal	dimer	stable
1cnt	\approx 40	0.70	xtal	dimer	stable
1cxq	10–50	0.22	bio	dimer	stable
1dok	1–10	0.83	xtal	tetramer	stable
1dto	0.03	0.12	xtal	monomer	stable
1dv8	—	0.12	xtal	dimer	stable
1e87	—	0.84	xtal	monomer	stable
1eaj	16	0.12	bio	dimer	stable
1edh	170–720	0.48	bio	monomer	stable
1eeq	0.012	0.83	xtal	dimer	stable
1erv	6–160	1.00	bio	dimer	uncertain
1exz	< 0.1	0.18	bio	tetramer	stable
1gml	—	0.50	bio	dimer	uncertain
1h65	—	0.85	xtal	dimer	stable
1hgx	—	0.58	xtal	dimer	stable
1ihk	0.7	0.83	xtal	dimer	stable
1ii5	0.8	0.40	bio	dimer	uncertain
1itv	—	0.75	bio	monomer	stable
1jr9	—	0.83	bio	dimer	stable
1lwj	—	0.82	bio	dimer	stable
1neu	900	0.12	xtal	monomer	stable
1nrv	2	0.12	xtal	monomer	stable
1o7z	9	0.12	xtal	dimer	stable
1oal	0.3	0.80	bio	dimer	stable
1p9y	18	0.48	bio	monomer	stable
1py9	—	0.12	bio	dimer	stable
1sj1	—	0.12	xtal	monomer	stable
1tip	—	0.83	xtal	monomer	stable
1uq5	—	0.31	xtal	dimer	stable
1zop	—	0.85	bio	dimer	stable
2dtr	0.033	0.84	bio	dimer	stable
2lyn	\approx 1	0.50	xtal	tetramer	stable
3bri	12	0.83	bio	dimer	stable
3ii8	0.3–18	0.50	bio	dimer	stable
3sdh	0.013	0.81	bio	dimer	stable
5cro	0.3	0.83	xtal	24-mer	stable

Table S22. Classification of the “others” set of dimers by Dey et al. [2]

PDB ID	ClusPro probability of dimer	EPPIC	PISA multimer class	PISA status
12as	0.84	bio	dimer	stable
1a3c	0.85	bio	dimer	stable
1a4i	0.82	bio	dimer	stable
1aa7	0.60	bio	monomer	stable
1ad3	0.55	bio	dimer	stable
1ade	0.81	bio	dimer	stable
1af5	0.12	bio	dimer	stable
1afw	0.82	bio	dimer	stable
1ajs	0.74	bio	dimer	stable
1aor	0.36	bio	dimer	stable
1aoz	0.67	xtal	tetramer	stable
1aq6	0.86	bio	dimer	stable
1auo	0.12	bio	dimer	stable
1b16	0.86	bio	dimer	stable
1b3a	0.91	bio	dimer	stable
1b5e	0.82	bio	dimer	stable
1b67	1.00	xtal	dimer	stable
1b8a	0.81	bio	dimer	stable
1b8j	0.79	bio	dimer	stable
1b8z	1.00	bio	dimer	stable
1b9h	0.83	bio	dimer	stable
1bam	0.83	xtal	dimer	uncertain
1bd0	0.80	bio	dimer	stable
1bdm	1.00	bio	dimer	stable
1bif	0.20	xtal	monomer	stable
1biq	0.80	bio	dimer	stable
1bis	0.93	bio	dimer	stable
1bk5	0.20	bio	monomer	stable
1bkp	0.83	bio	dimer	stable
1brw	1.00	bio	dimer	stable
1bsl	0.84	bio	dimer	stable
1bsr	1.00	bio	tetramer	stable
1buo	0.84	bio	dimer	stable
1bxg	0.73	bio	dimer	uncertain
1byf	0.81	bio	dimer	stable
1byi	1.00	bio	dimer	stable
1byk	0.75	bio	dimer	stable
1c4k	0.53	bio	dimer	stable
1cd8	0.86	bio	dimer	stable
1chm	0.73	bio	dimer	stable
1cmb	1.00	bio	dimer	stable
1cnz	0.81	bio	dimer	stable
1coz	0.89	bio	dimer	stable

1csh	0.69	bio	dimer	stable
1ctt	0.83	bio	dimer	stable
1czj	0.22	bio	dimer	stable
1d0c	0.79	bio	dimer	stable
1d2f	0.81	bio	dimer	stable
1d9c	0.77	bio	dimer	stable
1daa	0.86	bio	dimer	stable
1db3	0.82	bio	hexamer	stable
1dbq	0.66	bio	dimer	stable
1dcf	0.73	bio	dimer	stable
1dor	0.83	bio	dimer	stable
1dos	0.83	bio	dimer	stable
1dpg	0.84	bio	dimer	stable
1dqn	0.78	bio	dimer	stable
1dqs	0.78	bio	dimer	stable
1dxg	1.00	bio	dimer	stable
1e1z	0.34	bio	octomer	stable
1e4m	0.78	bio	dimer	stable
1e58	0.99	xtal	dimer	stable
1e6b	0.64	xtal	dimer	stable
1e6u	0.78	bio	dimer	stable
1e7n	1.00	bio	dimer	uncertain
1ebh	0.82	bio	dimer	stable
1ecy	0.70	bio	monomer	stable
1elu	0.78	bio	dimer	stable
1ew3	0.99	xtal	dimer	stable
1ex0	0.64	bio	dimer	stable
1eyv	0.87	xtal	dimer	stable
1f3h	1.00	xtal	dimer	stable
1f6b	0.12	xtal	monomer	stable
1f89	0.83	bio	dimer	stable
1fc4	0.71	bio	dimer	stable
1fg7	0.74	bio	dimer	stable
1fi4	0.83	bio	dimer	stable
1fip	0.94	bio	tetramer	stable
1fjj	0.83	bio	monomer	stable
1fp1	0.78	bio	dimer	stable
1fp2	0.76	bio	dimer	stable
1fp5	0.80	bio	dimer	stable
1fro	0.80	bio	dimer	stable
1g5c	1.00	bio	dimer	stable
1g60	0.83	bio	dimer	stable
1g8l	0.76	xtal	dimer	stable
1gd9	0.84	bio	dimer	stable
1glj	1.00	bio	dimer	stable
1gnw	0.81	bio	dimer	stable
1gqi	0.63	bio	monomer	stable
1gt3	0.85	bio	dimer	stable
1gta	0.80	bio	dimer	stable

1gte	0.27	bio	dimer	stable
1gu7	0.88	bio	dimer	stable
1gvj	0.82	xtal	dimer	stable
1gvp	0.89	bio	dimer	stable
1gy6	0.85	bio	dimer	stable
1gyo	0.12	xtal	dimer	stable
1h16	0.84	bio	dimer	stable
1h1y	0.82	bio	dimer	stable
1h4v	0.73	bio	dimer	stable
1h7e	0.85	bio	dimer	stable
1h9r	0.71	bio	nopred	n/a
1hf2	0.94	bio	tetramer	stable
1hjr	0.77	bio	dimer	stable
1hpx	1.00	bio	dimer	stable
1hq8	0.83	xtal	dimer	uncertain
1hss	0.70	bio	dimer	stable
1hw1	0.83	bio	dimer	stable
1hxp	0.82	bio	dimer	stable
1hzj	0.88	bio	dimer	stable
1i07	0.91	bio	dimer	stable
1i24	0.81	bio	tetramer	stable
1i2k	0.85	bio	dimer	stable
1i4u	0.71	xtal	dimer	stable
1i52	0.83	bio	dimer	stable
1i69	0.68	bio	dimer	stable
1i86	0.75	bio	dimer	stable
1imb	0.87	xtal	dimer	stable
1is3	0.71	bio	dimer	uncertain
1iup	0.83	bio	monomer	stable
1ivy	0.83	bio	dimer	stable
1ix9	0.82	xtal	dimer	stable
1iyz	0.72	bio	dimer	stable
1j3b	0.12	bio	dimer	stable
1j5p	0.85	bio	dimer	stable
1j98	0.83	bio	dimer	stable
1jd0	0.12	xtal	monomer	stable
1jdw	0.81	bio	dimer	stable
1jhd	0.59	bio	dimer	stable
1jhg	1.00	bio	dimer	stable
1jm6	0.22	bio	dimer	stable
1jsc	0.82	bio	dimer	stable
1jsg	0.83	bio	tetramer	stable
1jtv	0.83	bio	dimer	stable
1jye	0.78	bio	dimer	stable
1jys	0.84	bio	dimer	stable
1k04	0.81	bio	dimer	stable
1k3y	0.81	bio	dimer	stable
1k51	0.83	bio	dimer	stable
1k66	0.88	xtal	dimer	stable

1k94	1.00	bio	dimer	stable
1kc7	0.81	bio	dimer	stable
1kcm	0.83	bio	dimer	stable
1kic	0.12	bio	dimer	stable
1kjq	0.80	bio	dimer	stable
1kkj	0.82	bio	dimer	stable
1kl	1.00	xtal	dimer	stable
1kly	1.00	bio	dimer	stable
1knq	0.84	bio	dimer	uncertain
1kpf	0.90	bio	dimer	stable
1kso	1.00	bio	dimer	stable
1kta	0.84	bio	dimer	stable
1ktb	0.15	bio	dimer	stable
1kzh	0.81	bio	dimer	stable
1lc5	1.00	bio	dimer	stable
1llf	0.12	xtal	dimer	stable
1lom	0.56	bio	dimer	stable
1lxe	0.69	bio	dimer	stable
1m6p	0.80	bio	dimer	stable
1m7y	0.82	bio	dimer	stable
1mh9	0.83	bio	dimer	stable
1mkb	0.82	bio	dimer	stable
1mo9	0.80	bio	dimer	stable
1mor	0.80	bio	dimer	stable
1mqi	0.42	xtal	dimer	uncertain
1mr8	1.00	bio	tetramer	stable
1mvc	0.70	bio	dimer	stable
1mze	1.00	bio	dimer	stable
1n1b	0.96	xtal	dimer	stable
1n26	0.16	xtal	dimer	stable
1n2s	0.84	bio	dimer	stable
1n55	0.88	bio	dimer	stable
1n57	0.84	bio	dimer	stable
1n7f	0.59	xtal	tetramer	stable
1nf9	1.00	bio	dimer	stable
1nkd	0.84	bio	dimer	stable
1nn5	1.00	bio	dimer	stable
1noj	0.76	bio	dimer	stable
1nox	0.82	bio	dimer	stable
1ns5	0.97	bio	dimer	stable
1nsj	1.00	bio	dimer	stable
1nsy	0.82	bio	dimer	stable
1o8b	0.85	xtal	dimer	stable
1oaa	0.83	bio	dimer	stable
1oac	0.33	bio	dimer	stable
1oc2	0.83	bio	dimer	stable
1oh0	0.79	bio	dimer	stable
1oi6	0.83	bio	dimer	stable
1oke	0.28	xtal	dimer	stable

1on2	0.82	bio	dimer	stable
1oqz	0.24	xtal	monomer	stable
1ov3	0.83	bio	octomer	stable
1oyb	0.59	bio	monomer	stable
1oys	0.83	bio	dimer	uncertain
1p5f	0.84	bio	dimer	stable
1pfq	0.79	bio	dimer	stable
1pre	0.50	bio	dimer	stable
1psr	1.00	bio	dimer	stable
1puc	0.91	bio	dimer	stable
1pv9	0.68	bio	dimer	stable
1pxf	1.00	bio	dimer	stable
1q3o	0.12	xtal	dimer	stable
1q4g	0.48	bio	dimer	stable
1q6o	0.73	bio	tetramer	stable
1qfh	0.66	bio	dimer	stable
1qhi	0.81	bio	dimer	stable
1qi9	0.53	bio	dimer	stable
1qj4	0.81	xtal	dimer	stable
1qjb	0.81	bio	tetramer	stable
1qkk	0.64	bio	tetramer	stable
1qks	0.46	bio	dimer	stable
1qo8	0.40	bio	monomer	stable
1qr2	0.83	bio	dimer	stable
1qvz	0.12	xtal	monomer	stable
1qyc	1.00	xtal	dimer	stable
1r2f	0.83	bio	dimer	stable
1r31	0.78	bio	hexamer	stable
1r3s	0.76	bio	dimer	stable
1r7a	0.81	bio	dimer	uncertain
1rdo	1.00	xtal	dimer	stable
1reg	0.12	bio	dimer	stable
1rrm	0.80	bio	dimer	stable
1ryt	0.83	bio	tetramer	stable
1s0a	0.78	bio	dimer	stable
1s3e	0.41	bio	dimer	stable
1sb8	0.77	bio	dimer	stable
1se8	0.36	bio	dimer	stable
1ses	0.81	bio	dimer	stable
1sjy	0.83	bio	dimer	stable
1smn	0.84	xtal	dimer	uncertain
1smt	1.00	bio	dimer	stable
1sox	0.78	bio	dimer	stable
1t0i	1.00	bio	tetramer	stable
1t5h	0.62	bio	monomer	stable
1tc1	0.79	bio	dimer	stable
1tox	0.78	bio	dimer	stable
1trb	0.85	bio	dimer	stable
1trk	0.65	bio	dimer	stable

1tw3	0.86	bio	dimer	stable
1ual	0.82	bio	dimer	stable
1uby	0.83	bio	dimer	stable
1ui8	0.23	bio	dimer	stable
1ukg	0.43	xtal	dimer	uncertain
1utg	0.86	bio	dimer	stable
1uyr	0.26	bio	dimer	stable
1v25	0.83	bio	dimer	stable
1v2z	0.83	bio	dimer	stable
1vfr	0.74	bio	dimer	stable
1vhz	0.83	bio	dimer	stable
1vls	0.59	bio	dimer	stable
1vok	0.84	bio	tetramer	stable
1xso	0.81	bio	dimer	stable
2arc	0.77	bio	monomer	stable
2gsa	0.78	bio	dimer	stable
2hdh	1.00	bio	dimer	stable
2ilk	0.58	bio	dimer	stable
2mjp	1.00	bio	dimer	stable
2nac	0.86	bio	dimer	stable
2ohx	1.00	xtal	dimer	stable
2spc	0.67	bio	dimer	stable
2tct	0.83	bio	dimer	stable
2tgi	0.83	bio	dimer	stable
3dap	0.84	bio	dimer	stable
3lad	0.82	bio	dimer	stable
3ssi	0.82	bio	dimer	stable
3vub	1.00	bio	dimer	stable
4kbp	0.83	bio	tetramer	stable
5csm	0.54	bio	dimer	stable
5rub	0.80	bio	dimer	stable
8prk	1.00	bio	dimer	stable
9wga	0.83	bio	dimer	stable

Table S23. Classification of the weak set of dimers by Dey et al. [2] using three methods: Summary

method	pisa	eppic	cluspro
dimers (42)	22	21	24
dimers %	52.4%	50.0%	57.1%

Table S24. Classification of the “others” set of dimers by Dey et al. [2] using three methods: Summary

method	pisa	eppic	cluspro
dimers (272)	226	236	240
dimers %	83.1%	86.8%	88.2%

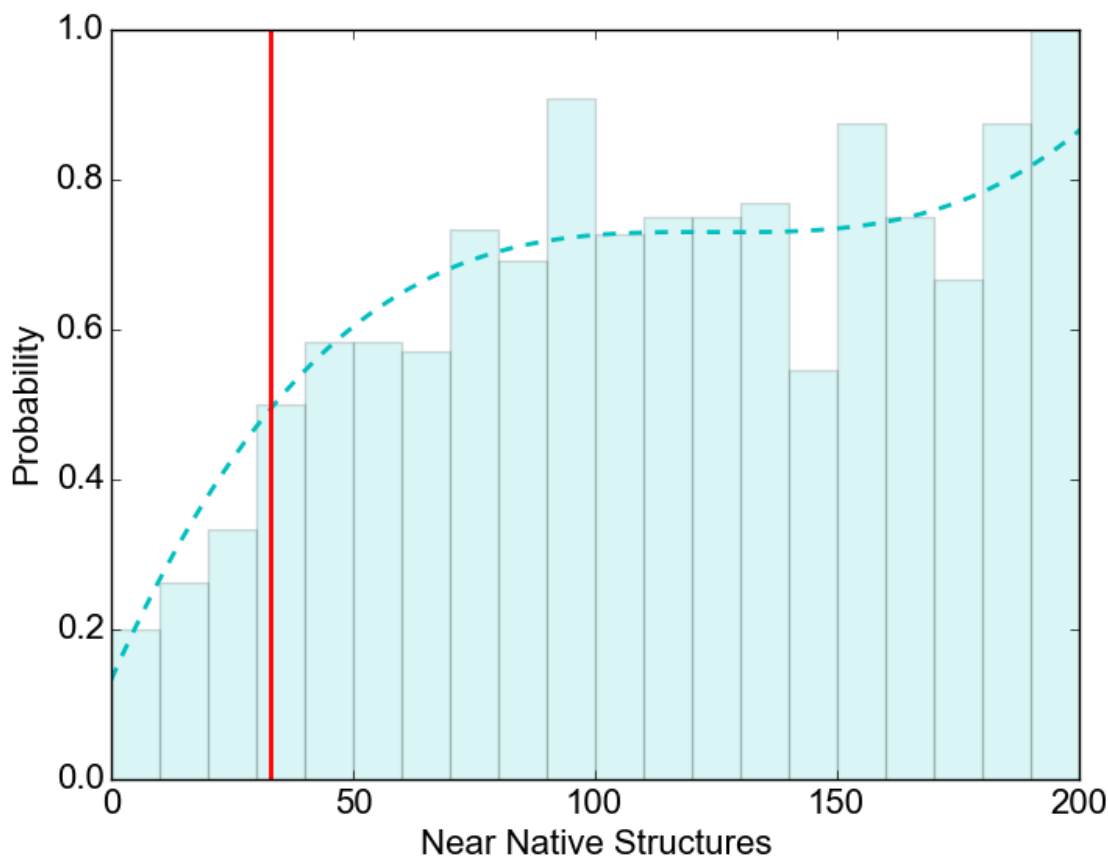


Figure S1.

Figure S1. Probability of biological dimer for combined training and test sets

Figure S2. (Page 12) Analysis of CAPRI target T70 (Modulator protein MzrA, PDB ID 4pwu). **(a)** Input for the analysis of the interface between chains A and B. **(b)** Status page showing the input and processed PDB files.

Figure S3. (Page 13) Results of the analysis of the interaction between chains A and B in CAPRI target T70 (Modulator protein MzrA, PDB ID 4pwu). **(a)** The result page shows the number N_{nn} of near-native docked structures among the lowest energy 1000 structures, indicated by a vertical red line, and the implied probability of the interface being in a biological dimer. Since for this dimer $N_{nn} = 214$, it is off the scale of the figure. Note the label for downloading the PyMOL session at the bottom of the page. **(b)** The PyMOL session downloaded shows the subunit 1 considered the receptor (solid grey), subunit 2 considered the ligand (solid green), and the top 100 lowest energy structures (green transparent cartoons). A large number but not all overlap with the native ligand positions.

[sign out](#)

a

Dock

Job Name:

Accepted PDB Input: The 20 standard amino acids.

Protein

PDB ID:

[Upload PDB](#)

Chain 1:

Chain 2:

[sign out](#)

b

171138: T70:4pwu:A_B

Status finished ([View Models](#))

Submitted 2016-06-26 13:40:07

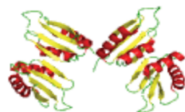
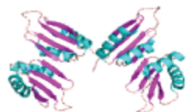
Errors (none reported)

Rotations 70000

user input

rec

lig



chains:
A

chains:
B

processed files

rec

lig

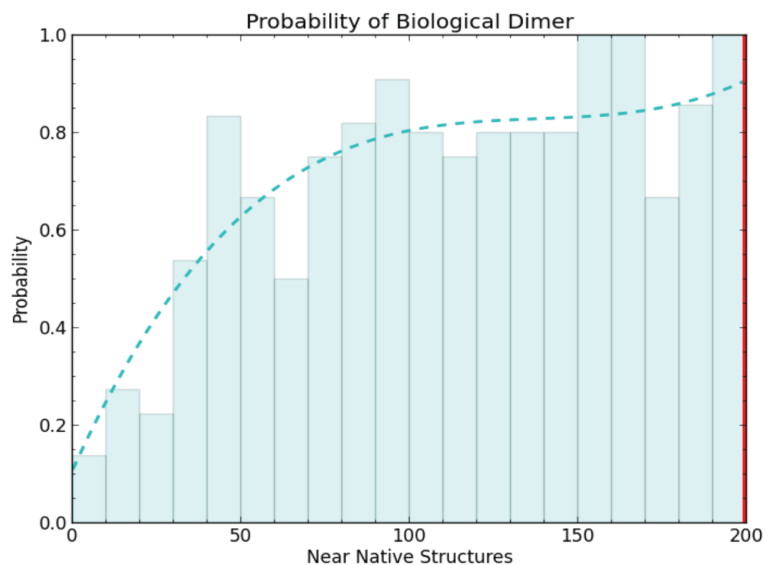


Figure S2

a

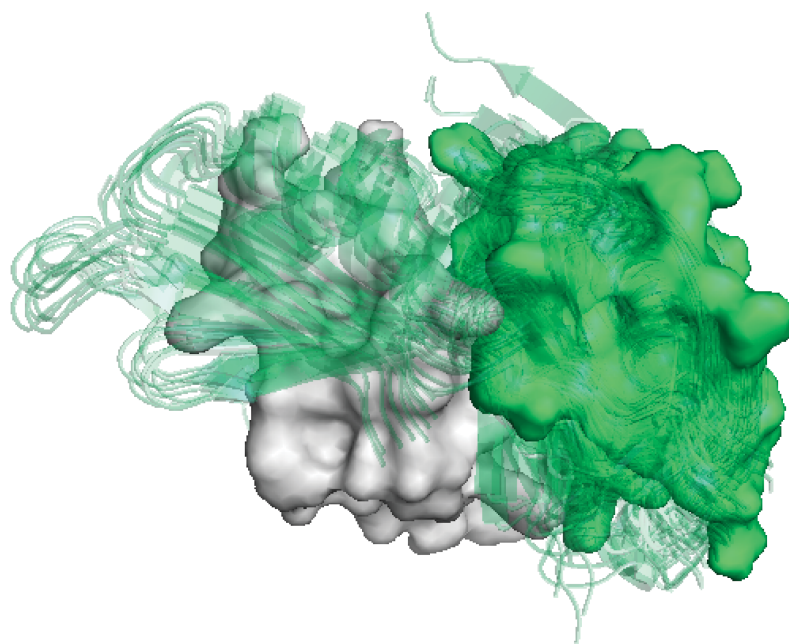
Result: Likely Biological

Near-native count: 214
Probability of being a biological dimer: 97%



[Download PyMOL Session](#)

b



all	A	S	H	L	C
rec	A	S	H	L	C
lig	A	S	H	L	C
lig.0	A	S	H	L	C
lig.1	A	S	H	L	C
lig.2	A	S	H	L	C
lig.3	A	S	H	L	C
lig.4	A	S	H	L	C
lig.5	A	S	H	L	C
lig.6	A	S	H	L	C
lig.7	A	S	H	L	C
lig.8	A	S	H	L	C
lig.9	A	S	H	L	C
lig.10	A	S	H	L	C
lig.11	A	S	H	L	C
lig.12	A	S	H	L	C
lig.13	A	S	H	L	C
lig.14	A	S	H	L	C
lig.15	A	S	H	L	C
lig.16	A	S	H	L	C
lig.17	A	S	H	L	C
lig.18	A	S	H	L	C
lig.19	A	S	H	L	C
lig.20	A	S	H	L	C
lig.21	A	S	H	L	C

Mouse Mode 3-Button Viewing
Buttons L M R Wheel
& Keys Rota Move MovZ Slab
Shft +Box -Box Clip MovS
Ctrl +/- PkAt Pk1 MvSZ
CtSh Sele Orig Clip MovZ
SnglClk +/- Cent Menu
DblClk Menu - PkAt
Selecting Residues
State 1/ 1

PyMOL>_

Figure S3

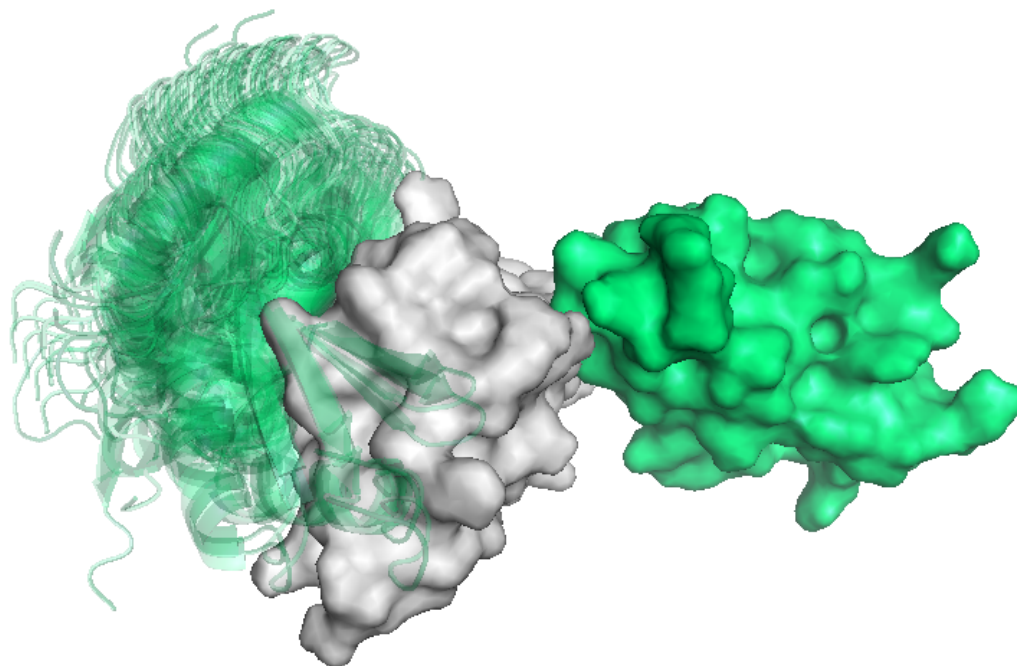


Figure S4. Results of the analysis of the interaction between chains A and C in CAPRI target T70 (Modulator protein MzrA, PDB ID 4pwu). The docked poses show strong interactions on the other side of subunit A. PyMOL generates a symmetry mate at that position, and the ClusPro result indicates a biological interface in addition to the one in the A:B dimer. This results in the stable A2:B2 tetramer (see text).

Supplementary References

- [1] I.M. Nooren, J.M. Thornton, Structural characterisation and functional significance of transient protein-protein interactions, *J. Mol. Biol.* 325 (2003) 991-1018.
- [2] S. Dey, A. Pal, P. Chakrabarti, J. Janin, The subunit interfaces of weakly associated homodimeric proteins, *J. Mol. Biol.* 398 (2010) 146-60.
- [3] R.P. Bahadur, P. Chakrabarti, F. Rodier, J. Janin, Dissecting subunit interfaces in homodimeric proteins, *Proteins* 53 (2003) 708-19.
- [4] R.P. Bahadur, P. Chakrabarti, F. Rodier, J. Janin, A dissection of specific and non-specific protein - Protein interfaces, *J. Mol. Biol.* 336 (2004) 943-55.
- [5] J.M. Duarte, A. Srebniak, M.A. Scharer, G. Capitani, Protein interface classification by evolutionary analysis, *BMC Bioinformatics* 13 (2012) 334.