Supplementary Materials

for the article

Applications of complementarity plot in error detection and structure validation of proteins

Sankar Basu, Dhananjay Bhattacharyya and Rahul Banerjee*

Indian Journal of Biochemistry and Biophysics, Vol 51, June 2014

Table S1—Datasets used in the calculations [Except for the pairs of obsolete and upgraded structures in OUDB, no protein with R-factor > 20% were included in any of the databases. For oligomeric proteins, only the largest polypeptide chain was retained for calculations. In case of multiple occupancies, atoms with the highest occupancy were selected and the first conformer for equal occupancies. For all the databases, homologues were removed at sequence identity of 30% or more. The PDB identifiers for each of the datasets can be found Dataset S1 of the Supplementary Materials]

Database	Resolution range	Chain length (aa)	Number of proteins	Additional criteria	Usage
DB2	≤ 2 Å	75-500	400	No proteins with deeply embedded prosthetic groups, No missing atoms	Training, Parameterization of $CS_{l,}$ *rGb*
UDB	≤ 1 Å	38 – 670	113	-	Computation of $CS_{l,}$ *rGb*
MDB	> 2 ´ , ≤ 2.5 Å	59 – 185	92	-	Same as UDB
LDB	≥ 3 Å	45 – 500	164	-	Same as UDB
OUDB	1.1-3.4 ´	65-900	110 pairs of obsolete and corresponding upgraded structures	Difference in resolution, R-factor between obsolete and upgraded pair: 0.2 Å, 0.02 respectively	Pair-wise Comparison, Detection of errors in Rotamer, Regularization
SDB-1	≤ 2 ´	56-363	20	divided equally among the four major protein classes	Idealization
SDB-2	≤ 2 ´	56-387	30	satisfying all validation filters implemented in Procheck [a]	Detection of low-intensity diffused synthetic errors in main-chain parameters
SDB-3	≤ 1 ´	38 – 670	68	No missing atoms	Idealization, Detection of unbalanced partial charge
SDB-4	≤ 2 ´	57-363	25	satisfying all validation filters implemented in Molprobity [b]	Detection of unbalanced partial charge

[a]Criteria for successful validation in Procheck: greater than -1.0 for all G-factor scores and 'INSIDE' or 'BETTER' recorded for bad contacts

[b]Criteria for successful validation in Molprobity: Ramachandran favored: > 98%, Ramachandran outliers: < 0.05%, Poor Rotamers: < 1%, Bad backbone bonds: 0%, Bad backbone angles: < 0.1%, Clash-score ≤ 20.

Table S2—Sensitivity of CS_l to different values of penalty (*Pen*) [The quantum of penalty *(Pen)* applied to CP1, CP2, CP3 is indicated in the first column of the table. $R_{Sl} = Sl_{zero} / Sl_{non\text{-}zero}$ (see Text)

Pen			R_{Sl}			CS_l
CP1	CP2	CP3	CP1	CP2	CP3	
100	100	100	1.31 (±1.44)	1.75 (±2.22)	2.02 (±1.93)	-0.54 (± 2.33)
75	75	75	0.98 (± 1.08)	1.31 (± 1.66)	1.52 (± 1.45)	0.33 (± 1.75)
50	50	50	0.66 (± 0.72)	0.88 (± 1.11)	1.01 (± 0.96)	1.19 (± 1.17)
30	30	30	0.39 (± 0.43)	0.53 (± 0.66)	0.61 (± 0.58)	1.89 (± 0.71)
25	25	25	0.33 (± 0.36)	0.44 (± 0.55)	0.51 (± 0.48)	2.06 (± 0.59)
20	20	20	0.26 (± 0.29)	0.35 (± 0.44)	0.41 (± 0.39)	2.23 (± 0.48)
15	15	15	0.20 (± 0.22)	0.26 (± 0.33)	0.31 (± 0.29)	2.40 (± 0.36)
10	10	10	0.13 (± 0.14)	0.18 (± 0.22)	0.20 (± 0.19)	2.58 (± 0.25)
5	5	5	0.07 (± 0.07)	0.09 (± 0.11)	0.10 (± 0.10)	2.75 (± 0.14)
30	25	20	0.39 (± 0.43)	0.44 (± 0.55)	0.41 (± 0.39)	2.06 (± 0.60)
25	20	15	0.33 (± 0.36)	0.35 (± 0.44)	0.31 (± 0.29)	2.24 (± 0.48)
20	15	10	0.26 (± 0.29)	0.26 (± 0.33)	0.20 (± 0.19)	2.41 (± 0.37)

Fig. S1—Training and testing of the complementarity and accessibility scores $CS_{i,}$ *rGb* in database DB2 and datasets with different resolution ranges UDB, MDB, LDB (See 'Materials & Methods'). The average (colored filled bars) and standard deviations (error bars) for the two scores (A) CS_i (B) *rGb* have been indicated]

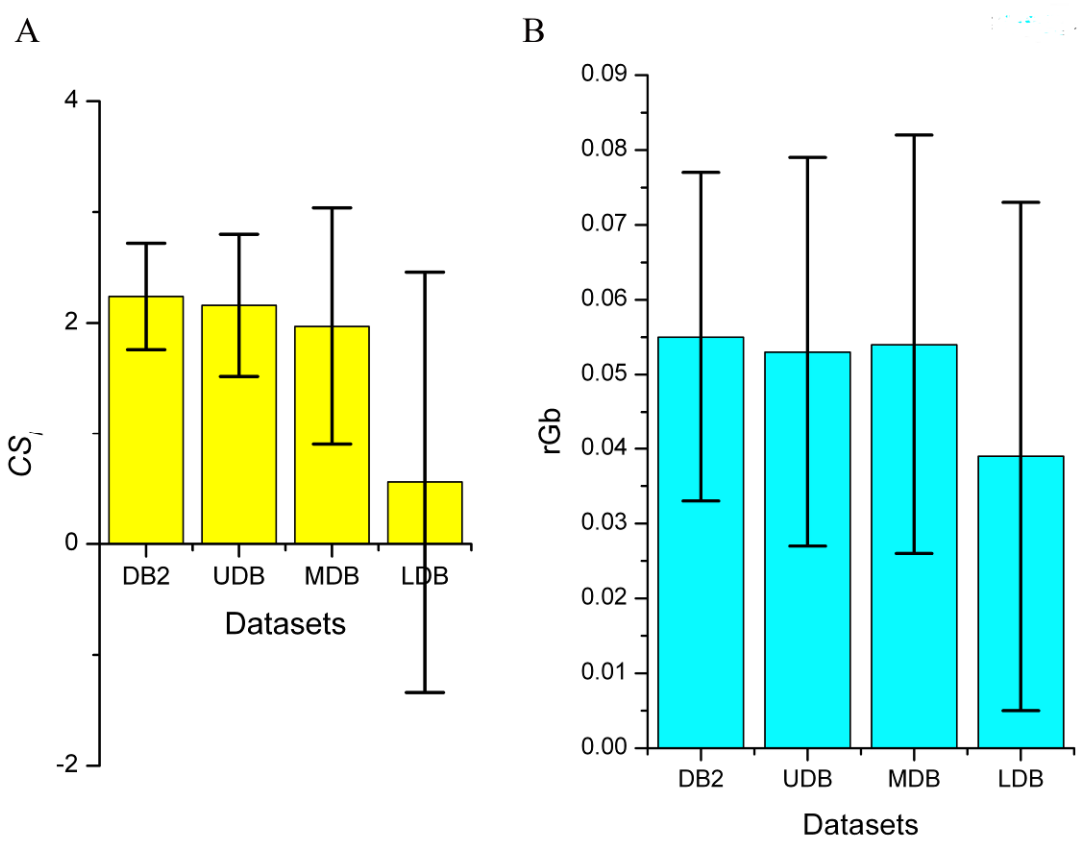


Table S3—Detection of errors in the retracted or suspected structures. Complementarity (CS_b, CS_f) and accessibility scores (*rGb*) along with the clash score (Molprobity), Whatcheck-packing Z-score and Procheck-global score have been given for each structure (see Main-Text for validation criteria). Information regarding these retracted or suspected structures was obtained from http://main.uab.edu/Sites/reporter/articles/71570/ and Read et al., 2011, *Structure* 19, 1395-1412. and ftp://ftp.wwpdb.org/pub/ pdb/data/status/obsolete.dat.

PDB ID	Resolution, R-facor	CS_b	*rGb*	Whatcheck Packing Z-score	P_{sm}	Procheck global score	Clash score
1BEF	2.10, 0.186	-4.08	-0.009	-4.47	-1.351	0.150	105.82
1CMW	2.60, 0.192	-0.40	0.051	-1.64	-0.953	0.290	29.04
1DF9	2.10, 0.199	-2.78	-0.014	-4.72	-1.094	0.090	49.53
2QID	2.10, 0.199	-2.78	-0.014	-4.72	-1.094	0.090	49.53
1G40	2.20, 0.198	-3.98	0.009	-3.90	-1.154	-0.320	57.13
1G44	2.60, 0.234	-4.28	0.007	-4.09	-1.182	-0.430	29.02
1L6L	2.30, 0.198	-1.18	-0.023	-0.92	-0.973	0.340	27.30
1RID	2.10, 0.206	-4.83	0.012	-4.63	-1.295	-0.440	48.45
1Y8E	2.20, 0.195	-3.84	0.009	-4.51	-1.145	-0.440	40.14
2A01	2.40, 0.228	-2.07	0.038	-2.89	-1.395	0.040	91.23
2HR0	2.26, 0.180	0.24	0.047	-2.48	-0.978	-0.160	40.82
1PF4	3.80, 0.240	-2.19	-0.065	-5.41	-1.319	-0.040	47.19
1S7B	3.80, 0.320	1.20	-0.162	-1.56	-1.337	0.160	35.91
1Z2R	4.20, 0.280	-3.84	-0.041	-2.64	-1.307	0.220	42.93
2F2M	3.70, 0.282	2.52	-0.174	-0.76	-1.216	0.290	30.91
2A73	3.30, 0.233	0.04	0.045	-2.28	-0.945	-0.040	20.36
2ADH	2.4, NULL	-6.13	0.063	-3.96	-1.473	-12.340	427.13
2CK9	2.85, 0.187	0.68	0.075	-1.31	-1.089	-0.030	50.20
2MT2	2.30, NULL	-7.76	0.055	-3.77	-1.244	-3.990	149.94
2PZ3	2.42, 0.314	1.10	0.064	-3.35	-0.874	-5.720	42.32
2QNS	3.00, 0.238	-0.28	0.057	-3.02	-0.964	-6.070	100.74
2RA7	1.99, 0.242	1.64	0.067	2.14	-0.813	0.090	7.33
3A00	1.80, 0.222	1.75	0.068	-0.23	-0.896	0.020	16.85
3K78	2.80, 0.274	0.78	0.066	0.45	-0.986	-0.070	30.02
3KJ5	3.00, 0.366	-1.50	0.061	-2.50	-1.195	-0.600	153.5
3O7Y	2.41, 0.180	1.65	0.032	-1.63	-0.844	-0.070	9.02
3O7Z	2.55, 0.183	1.70	0.028	-1.32	-0.819	-0.020	9.4
3O8K	2.70, 0.268	-0.02	0.059	-1.69	-1.074	-0.340	53.68

Table S4—Complementarity and accessibility scores for idealized structures [Average scores ($CS_{l,}$ *rGb*) and standard deviations (in parentheses) obtained for different forms of idealization on the database SDB-1. The same scores have also been tabulated for the native proteins in the original databases DB2 and SDB-1]

Idealization protocol	CS_l	*rGb*
DB2 (≤ 2 Å, 400)	2.24 (0.48)	0.055 (0.022)
SDB-1 (≤ 2 Å, 20)	2.47 (0.41)	0.060 (0.020)
Main-chain bond-lengths[a] , angles[a] and ω[b] idealized	-10.54 (3.48)	0.000 (0.031)
Main-chain bond-lengths[a] , angles[a] and ω[b] idealized and energy-minimized with flexible backbone	-2.58 (2.61)	0.004 (0.030)
Main-chain bond-lengths[a], angles [a] idealized (with native ω)	-10.52 (3.80)	0.009 (0.03)
Main-chain bond-angles[c] idealized (with native ω)	-8.98 (3.87)	0.017 (0.028)
Main-chain bond-angles [c] idealized (with native ω), energy-minimized with rigid backbone	-1.42 (2.59)	0.019 (0.025)
Main-chain bond-lengths [a] idealized	2.45 (0.36)	0.060 (0.020)
Main-chain bond-angles[a] idealized	-10.56 (3.75)	0.010 (0.030)
ω idealized [b]	-7.80 (3.80)	0.022 (0.030)
Main-chain bond-angle: N-C^{α}-C (τ) [a] idealized	-7.80 (3.95)	0.031 (0.027)
Main-chain bond-angle: C^{α}_i-C_i-N_{i+1} [a] idealized	-4.98 (4.73)	0.047 (0.026)
Main-chain bond-angle: C_{i-1}-N_i-C^{α}_i [a] idealized	-3.95 (3.36)	0.037 (0.030)

Ideal values for pre-selected geometrical parameters were obtained from
[a]Engh and Huber, 2001[6]
[b]Whatif (Vriend, 1990)[21]
[c]Conformation Dependent Library (CDL) (Berkholz et al., 2009)[23]

Fig. S2—Effect of CDL-idealization probed by CP [Distribution for (A) the native polypeptide chain (1PGS) and (B) its corresponding idealized structure generated utilizing CDL (Conformation Dependent Library) ideal values]

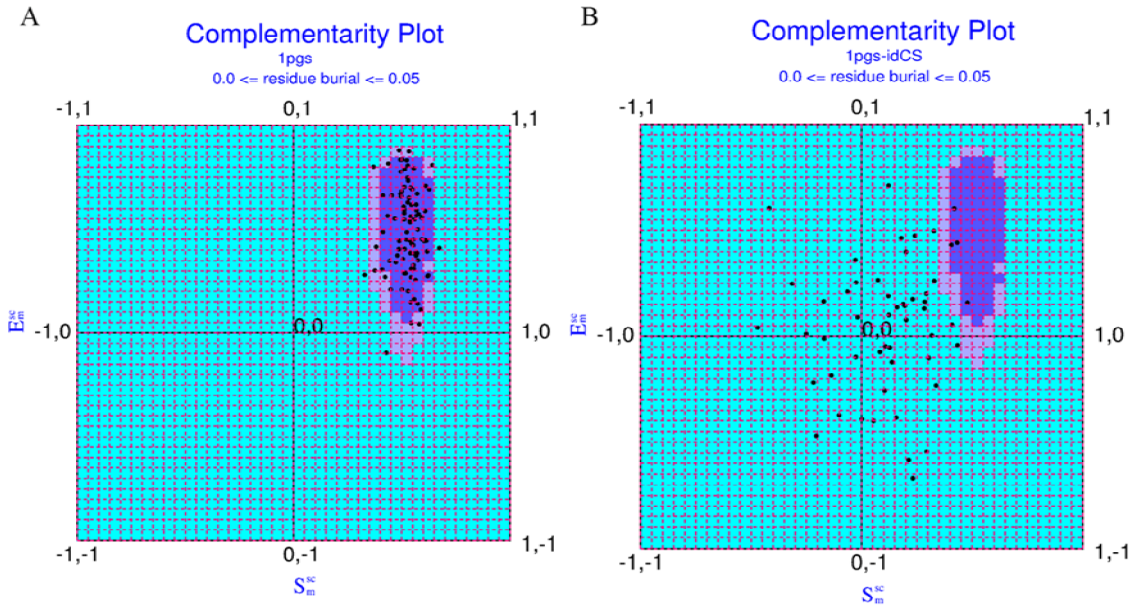


Table S5—Structural distortions due to idealization as reflected in the RMSDs

PDB ID	RMSD (Å) [a]	
	Idealized vs. native	Idealized and energy minimized vs. native [b]
1AKO	13.98	-[c]
1BGF	7.30	6.77
1CEM	16.61	17.35
1CHD	22.42	22.18
1CKA	3.05	3.10
1ERZ	24.44	22.78
1HBQ	22.60	-
1IFC	11.30	11.48
1LMB	4.56	4.56
1MKB	-	-
1MLA	23.33	22.05
1PDO	4.56	5.29
1PGS	-	-
1SFP	-	-
1SRV	13.82	-
1STN	18.02	18.10
1UBI	4.31	4.14
2CPL	12.52	12.58
2END	7.64	7.41
2LIS	9.09	9.11

[a]RMSDs calculated between C^{α} atoms of idealized (all main-chain bond lengths, bond angles and ω) and the native coordinates (calculated at a one-to-one atomic correspondence) subsequent to superposition by Dali server.
[b]The same calculation was repeated for energy minimized coordinates subsequent to idealization.
[c] '-' stands for non-superposable structures.

Table S6—Complementarity and accessibility scores for idealized structures of ultra-high resolution

Parameters used for Idealization [a]	CS_l	*rGb*
Unimodal ideal values [b]	-9.82	-0.009
	(3.75)	(0.032)
CDL ideal values [c]	-6.64	0.024
	(4.12)	(0.003)

[a] Structures idealized by different methods from a database of 68 ultra-high resolution structures (SDB-3).
[b] Engh and Huber, 2001[6]
[c] a Conformation Dependent Library (CDL) (Berkholz *et al*., 2009)[23]
Average scores (CS_l, *rGb*) standard deviations (in parentheses) for the idealized structures.

Fig S3—***CS$_l$* scores for native and corresponding redesigned structures [The native *CS$_l$* scores for 93 structures (from SDB-3 and SDB-4) plotted in 'red' along with those subsequent to the 'hydrophobic to hydrophilic' transitions (and *vice-versa*) plotted in 'blue'. The structures in the database have been numbered sequentially]

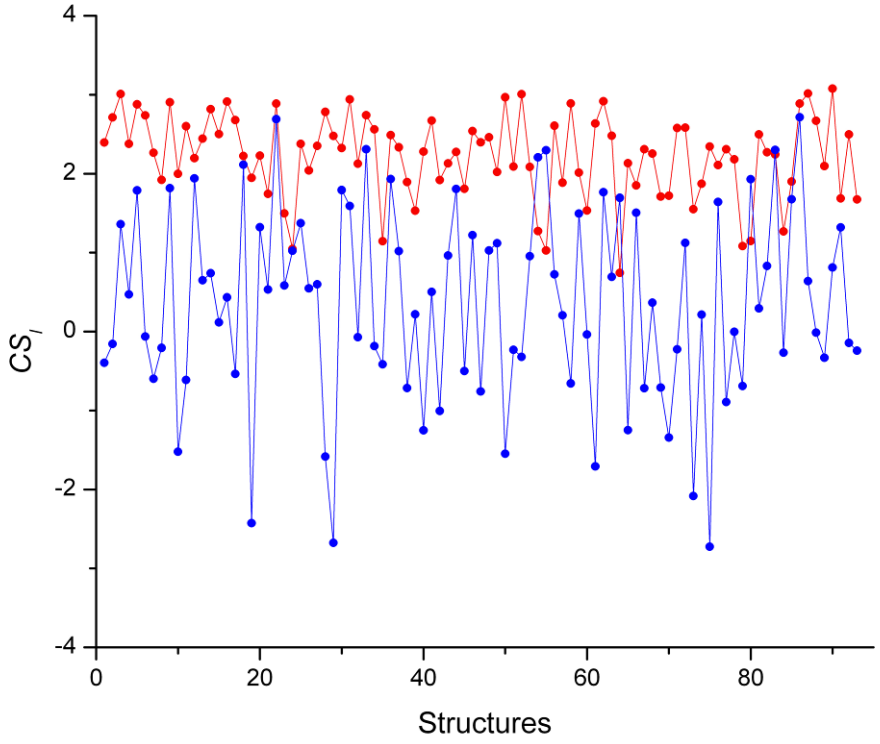


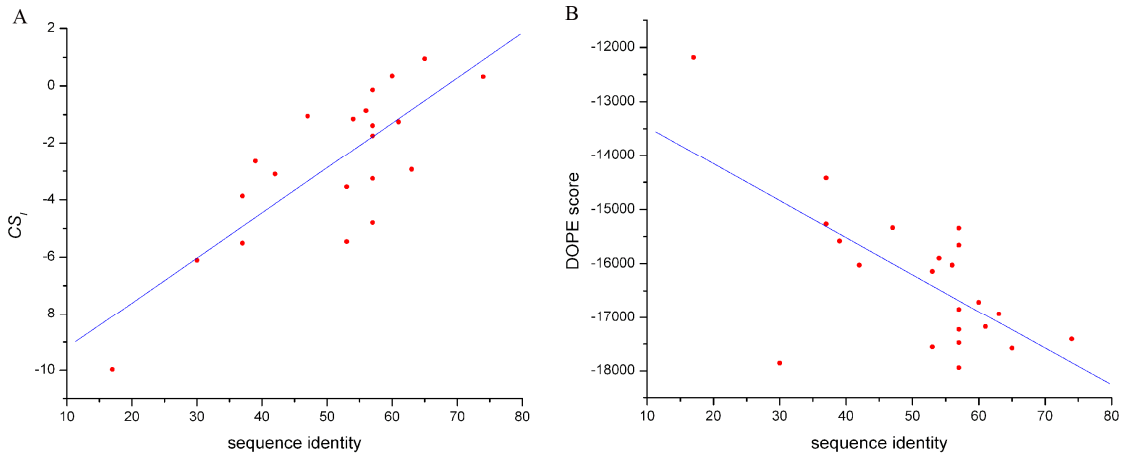
Table S7—Performance of CP in quality assessment of homology models with varying sequence identity w.r.t. the template. For each fold, the template along with 6 modeled sequences which includes the native sequence (PDB IDs tabulated) and their corresponding complementarity scores (CSl, CSf, rGb). The sequence identity and similarity of the modeled sequence w.r.t. the template has also been given.

Fold	Template	Modeled Sequence	Identity (%)	Similarity (%)	CSl	rGb	QMEAN global
Cyclophilin-like	2HAQ-native (A:22-187)				2.54	0.059	0.83
	2HAQ	2HAQ	100	100	2.55	0.060	0.84
	2HAQ	1DYW	54.9	70.9	1.23	0.050	0.64
	2HAQ	2POE	35.8	54.7	0.28	0.027	0.59
	2HAQ	2NUL	32.8	49.7	0.28	0.050	0.55
	2HAQ	2OSE	24.2	45.5	-0.04	0.005	0.37
	2HAQ	3KOP	16.5	39.4	-1.79	-0.009	0.41
4-helical Cytokines	1WU3-native (I:1-161)				1.62	0.053	0.71
	1WU3	1WU3	100	100	2.35	0.051	0.72
	1WU3	1AU1	48.5	69.8	2.91	0.056	0.64
	1WU3	3OQ3	26.7	55.2	1.63	0.042	0.56
	1WU3	1B5L	29.8	54.1	0.91	0.061	0.64
	1WU3	3SE3	27.5	51.7	0.57	0.026	0.54
	1WU3	3PIV	15.2	48.9	-1.14	0.032	0.47
ALDH-like	1BPW-native (A:1-503)				1.32	0.045	0.69
	1BPW	1BPW	100	100	1.48	0.046	0.71
	1BPW	2XDR	48.7	70.0	1.21	0.037	0.65
	1BPW	1ZUM	40.7	64.4	1.12	0.032	0.65
	1BPW	3RH9	35.4	57.8	0.61	0.043	0.64
	1BPW	1EUH	29.1	53.5	-0.96	0.035	0.61
	1BPW	3MY7	19.7	42.8	-1.46	0.005	0.46
Tim-barrel	1AMY-native (A:1-346)				1.75	0.043	0.61
	1AMY	1AMY	100	100	1.57	0.041	0.64
	1AMY	1HT6	77.2	88.5	1.80	0.035	0.64
	1AMY	1GCY	27.7	47.3	0.45	0.026	0.58
	1AMY	1W9X	24.1	44.4	-0.64	-0.002	0.32
	1AMY	1HVX	23.0	43.4	-1.68	-0.018	0.27
	1AMY	1BVZ	20.6	40.0	-1.89	0.001	0.41
EF-hand-like	1PSR-native, (A:1-100, Ca+2)				1.53	0.063	0.88
	1PSR	1PSR	100	100	0.69	0.065	0.81
	1PSR	1E8A	25.2	51.5	1.19	0.046	0.54
	1PSR	3C1V	22.9	47.7	0.19	0.028	0.58
	1PSR	4ICB	24.3	44.7	0.89	0.077	0.98
	1PSR	1YTZ	21.0	36.4	1.04	0.001	0.39
	1PSR	1TCF	18.1	34.5	-1.08	0.012	0.34
Nudix	2FVV-native (A:8-142)				2.43	0.062	0.79
	2FVV	2FVV	100	100	2.71	0.058	0.79
	2FVV	2DUK	83.7	91.9	1.81	0.073	0.78
	2FVV	3MCF	76.5	83.3	2.65	0.074	0.74
	2FVV	1KT9	23.8	49.0	-0.17	-0.004	0.41
	2FVV	2B06	20.4	44.4	-1.39	0.010	0.45
	2FVV	1VC8	24.3	42.6	1.07	-0.012	0.56
Pheromone-sensitive-neuron	2ERB-native (A:3-125)				2.93	0.080	0.81
	2ERB	2ERB	100	100	2.13	0.072	0.77
	2ERB	3OGN	90.3	94.4	2.75	0.062	0.79
	2ERB	3K1E	81.6	92.8	2.73	0.076	0.79
	2ERB	3R1O	24.1	48.9	-0.08	0.048	0.63
	2ERB	2FJY	22.0	44.0	-0.52	-0.012	0.32
	2ERB	3B7A	16.3	41.5	-0.56	0.045	0.48

Peptidase_S7	2GGV-native (B:10-178)				2.37	0.044	0.80
	2GGV	2GGV	100	100	2.11	0.034	0.81
	2GGV	2WV9	75.4	88.6	2.44	0.019	0.81
	2GGV	2WHX	56.4	73.9	1.61	0.047	0.73
	2GGV	3L6P	48.7	65.8	0.67	0.053	0.71
	2GGV	3J0C	20.8	42.7	-2.04	0.004	0.39
	2GGV	1XVM	18.8	33.3	-3.18	-0.025	0.31
Thioredoxin-like	1EEJ-native (B:61-216)				1.84	0.063	0.88
	1EEJ	1EEJ	100	100	2.08	0.078	0.89
	1EEJ	1T3B	46.8	67.1	0.87	0.067	0.81
	1EEJ	3GV1	29.6	55.6	-0.25	0.054	0.67
	1EEJ	3GYK	24.6	43.9	-1.65	0.028	0.54
	1EEJ	1BED	19.0	42.6	-2.29	0.014	0.33
	1EEJ	1UVZ	22.0	39.6	-2.51	0.008	0.44
Macro-domain-like	2ACF-native (C:184-351)				2.22	0.056	0.79
	2ACF	2ACF	100	100	2.74	0.069	0.79
	2ACF	3EWQ	34.0	56.5	0.68	0.054	0.58
	2ACF	3EJG	32.1	56.0	0.14	0.045	0.39
	2ACF	1HJZ	28.1	51.2	-1.04	0.020	0.44
	2ACF	1SPV	30.1	47.6	-0.83	0.046	0.49
	2ACF	2X47	22.6	45.8	-2.10	0.027	0.41
	2ACF	1ZR5	17.6	42.1	-0.48	-0.001	0.32
Aldo-keto-reductase (family)	3V36-native (A:1-315)				2.25	0.064	0.76
	3V36	3V36	100	100	2.08	0.060	0.74
	3V36	1DLA	86.3	94.3	2.04	0.061	0.74
	3V36	1ZUA	70.4	85.5	1.77	0.062	0.75
	3V36	3O3R	68.7	84.5	1.55	0.066	0.71
	3V36	3UWE	46.7	70.3	1.22	0.051	0.73
	3V36	1A80	33.3	54.3	1.12	0.030	0.59
	3V36	1YNP	19.5	45.5	-1.72	0.020	0.47
Citrate-synthase-like	2IBP-native (A: 35-386)				2.60	0.053	0.78
	2IBP	2IBP	100	100	2.02	0.051	0.75
	2IBP	1AJ8	44.1	64.9	1.57	0.065	0.71
	2IBP	2C6X	37.0	59.6	0.79	0.042	0.64
	2IBP	1IOM	40.1	59.1	0.29	0.038	0.65
	2IBP	4E6Y	31.8	49.6	0.45	0.036	0.59
	2IBP	1AMZ	23.9	45.7	-0.05	-0.013	0.43
Phosphoglycerate mutase-like	1RII-native (A:2-244)				2.32	0.034	0.73
	1RII	1RII	100	100	2.39	0.033	0.72
	1RII	4EO9	86.5	94.3	2.35	0.037	0.75
	1RII	3PGM	57.6	71.8	1.89	0.057	0.67
	1RII	1XQ9	46.9	65.8	1.44	0.045	0.64
	1RII	1EBB	28.0	46.8	0.30	-0.013	0.43
	1RII	1H2E	27.0	46.4	0.14	-0.013	0.41
SAM-dependent methyl trasferase like	2AVD-native (B:44-262)				2.24	0.042	0.79
	2AVD	2AVD	100	100	2.05	0.045	0.77
	2AVD	3TR6	47.1	64.0	1.34	0.033	0.69
	2AVD	3CBG	39.1	59.6	1.45	0.040	0.71
	2AVD	2HNK	33.8	56.2	1.28	0.029	0.66
	2AVD	3C3P	22.6	49.3	1.11	0.042	0.71
	2AVD	3BWY	21.2	45.1	0.41	-0.008	0.51
Tripsin-like-serine-protease	2GGV-native (10-178)				2.15	0.044	0.80
	2GGV	2GGV	100	100	2.11	0.034	0.81
	2GGV	2WV9	75.4	88.6	2.44	0.019	0.81
	2GGV	2WHX	55.8	73.3	0.78	0.052	0.73
	2GGV	2FOM	50.0	67.8	1.89	0.027	0.66
	2GGV	1WCZ	22.4	45.6	-2.96	-0.021	0.27
	2GGV	1AGI	16.1	37.2	-1.48	0.022	0.43
Glutathione-S-transeferase	2IMI-native (B:1-220)				2.80	0.065	0.76
	2IMI	2IMI	100	100	2.44	0.063	0.78
	2IMI	1PN9	41.3	62.7	1.52	0.031	0.62

(family)	2IMI	1R5A	35.7	62.1	1.38	0.033	0.61
	2IMI	1E6B	28.0	49.4	0.32	0.017	0.61
	2IMI	2CZ2	23.1	47.4	-0.41	0.024	0.53
	2IMI	3AY9	13.5	31.7	-3.47	-0.027	0.16
Chelatase-class II	2H1V-native (A:2-310)				2.61	0.093	0.85
(super family)	2H1V	2H1V	100	100	2.45	0.090	0.85
	2H1V	2C8J	71.8	85.8	2.54	0.082	0.8
	2H1V	2QD3	26.7	53.9	1.07	0.053	0.64
	2H1V	2PO7	26.4	53.3	0.99	0.053	0.67
	2H1V	1L8X	23.9	51.6	1.29	0.049	0.55
	2H1V	2AFR	17.3	41.0	-3.05	0.001	0.33
Alpha-beta knot	1V2X-native (A:1-191)				1.72	0.059	0.89
	1V2X	1V2X	100	100	2.21	0.055	0.88
	1V2X	1ZJR	49.0	67.3	1.87	0.055	0.79
	1V2X	3N4J	24.5	43.5	0.80	0.023	0.39
	1V2X	1IPA	24.3	39.1	-0.62	-0.009	0.43
	1V2X	3NK6	21.2	39.4	-4.11	-0.018	0.28
	1V2X	3IC6	18.5	39.7	-1.37	0.021	0.39
Periplasmic binding protein	1XS5-native (A:5-244)				2.30	0.065	0.86
like II	1XS5	1XS5	100	100	2.46	0.063	0.79
	1XS5	3TQW	36.6	63.0	1.58	0.044	0.74
	1XS5	3K2D	35.9	61.2	1.51	0.051	0.67
	1XS5	4EF1	31.8	64.1	1.66	0.072	0.73
	1XS5	2X26	22.2	39.2	-3.49	0.001	0.28
	1XS5	4ESW	17.3	37.4	-3.63	-0.049	0.16
SH3-like barrel	1CKA-native (A:134-190)				3.01	0.09	0.95
	1CKA	1CKA	100	100	2.98	0.08	0.99
	1CKA	1PWT	31.1	62.3	2.79	0.07	0.85
	1CKA	1EFN	37.9	62.1	0.71	0.03	0.66
	1CKA	1Y57	41.4	60.3	2.69	0.04	0.78
	1CKA	1CSK	31	54.9	-0.51	0.04	0.36
	1CKA	1QWE	34.4	53.3	-0.81	0.03	0.31

Fig. S4—***CS_I* and DOPE scores derived from Modeller for homology models built on the template 2HAQ as a function of sequence identity [The Pearson's correlation between the sequence identity and (A) *CS_I* ; (B) DOPE-score are 0.79 and -0.66 respectively]**



Dataset S1—The databases of protein structures used in the study. The PDB identifiers and the chain identifiers (underscored) in case of oligomeric proteins constituting each database are listed below. For obsolete structures in the database OUDB, '-' is given for PDB files with resolution and/or R-factor mentioned as 'NULL'. The training database, DB2 has already been used in a previous calculation (Basu et al, 2012[1]) with satisfactory results.

UDB:

1EB6_A, 1EJG_A, 1F94, 1GA6, 1I1W_A, 1IEE, 1K5C, 1KTH, 1L9L_A, 1M40_A, 1MJ5, 1MUW, 1MWQ, 1NQJ, 1OAI_A, 1OD3_A, 1P9G, 1R6J_A, 1TG0_A, 1TT8, 1UCS, 1UFY, 1VBW_A, 1W0N_A, 1X6X, 1ZLB, 1ZUU, 1ZZK, 2A6Z_A, 2B97_B, 2BW4_A, 2CHH_A, 2CWS_A, 2DDX_A, 2F01, 2FMA, 2FVY, 2FWH_A, 2GGC, 2GKG_A, 2GUD, 2H3L_A, 2I4A, 2IXT_B, 2O7A_A, 2OV0, 2P5K_A, 2P74, 2PND_A, 2QCP, 2QSK_A, 2R31_A, 2RBK, 2RH2, 2V1M_A, 2V8T_A, 2VHK_A, 2X46_A, 2XFR_A, 2XJP_A, 2XOM_A, 2XU3_A, 2Y78_A, 3A02_A, 3A38, 3AGN_A, 3AJ4_A, 3AKS_A, 3BWH_A, 3CCD, 3D43, 3EA6_A, 3F1L_A, 3FSA, 3FYM_A, 3G21_A, 3GOE_A, 3H31_A, 3IP0_A, 3JU4_A, 3KLR_A, 3KS3_A, 3L8W, 3MFJ, 3NE0_A, 3NED_A, 3O4P_A, 3ODV_B, 3PSM, 3PUC_A, 3Q46_A, 3QL9, 3QPA, 3RCG, 3RQ9_B, 3RWN_A, 3S6E_A, 3SOJ, 3U7Q_D, 3U7T_A, 3UI4_A, 3V1A_A, 3VII_A, 3VLA_A, 3VN3_A, 3ZR8_X, 4ACJ_A, 4AXO_B, 4DRQ, 4EGU, 4F1V_A, 4GNR_A, 7A3H_A.

MDB:

1DCF_A, 1G43_A, 1IGU_A, 1J6V_A, 1JEO_A, 1LXI_A, 1NE8_A, 1NFJ_A, 1ORE_A, 1OSD_B, 1R6Y_A, 1T3X_A, 1T4A_A, 1T4Q_A, 1TJN_A, 1TWU_A, 1U2Q_A, 1U61_A, 1WQG_A, 1WU3_I, 1XJC_A, 1YIG_A, 1YOY_A, 1YU9_A, 1Z9F_A, 1Z9W_A, 1ZUI_A, 2APL_A, 2B45_X, 2B91_A, 2CZT_A, 2EH1_A, 2EH9_A, 2FB0_A, 2FBI_A, 2FE1_A, 2FIW_A, 2FLS_A, 2FU2_A, 2FZT_A, 2GUX_A, 2HYZ_A, 2I9V_A, 2IL5_A, 2OBB_A, 2OGG_A, 2P71_A, 2PIM_A, 2Q0V_A, 2QG8_A, 2QYB_A, 2QYJ_A, 2R18_A, 2R39_A, 2RDP_A, 2REA_A, 2ZBN_A, 3AW6_A, 3BJV_A, 3BW6_A, 3BZ6_A, 3CDN_A, 3CZC_A, 3D4T_A, 3DJN_B, 3DXR_A, 3EFA_A, 3FF4_A, 3FFY_A, 3FNI_A, 3G1G_A, 3HMH_A, 3K3D_A, 3KAB_A, 3KCW_A, 3L93_A, 3LDT_A, 3LMF_A, 3MSE_B, 3NQJ_A, 3O8Z_A, 3P7J_B, 3R2E_A, 3RA6_A, 3RH1_A, 3RKV_A, 3T9Y_A, 3TNJ_A, 3TS9_A, 3U2R_A, 3UH8_A, 4DN8_A.

LDB:

1A0D_B, 1AIP_E, 1AUM_A, 1AY9_A, 1AYB_A, 1AZY_B, 1B26_B, 1B6U_A, 1B8H_B, 1C1A_A, 1C41_G, 1C8B_B, 1DOV_A, 1DWN_B, 1E0J_B, 1E57_B, 1E69_B, 1EFR_C, 1FLL_B, 1FOU_L, 1FP9_A, 1FQX_B, 1GMO_C, 1HES_A, 1I0E_B, 1IJ5_A, 1JH5_G, 1JJK_G, 1JJO_C, 1JL9_B, 1K0U_G, 1K1D_G, 1KPL_B, 1L8I_J, 1LJ7_G, 1LM7_B, 1LQM_G, 1M5Y_D, 1MO2_B, 1MPQ_B, 1N9S_L, 1NOV_C, 1NY6_G, 1NY7_2, 1O96_B, 1OJL_D, 1P75_B, 1PFC_A, 1PGR_F, 1PP6_B, 1PVJ_B, 1PW4_A, 1Q9C_I, 1QB3_B, 1QBE_B, 1QMO_G, 1QOL_G, 1QQK_A, 1QZ2_C, 1R8I_A, 1RH5_A, 1RLV_B, 1RW3_A, 1RXT_A, 1RXU_G, 1S94_B, 1S9C_G, 1SC5_A, 1SLQ_A, 1SRU_D, 1SXI_L, 1T62_B, 1T8B_B, 1TIJ_A, 1TLY_A, 1UFU_A, 1V1S_B, 1V6O_G, 1VGZ_A, 1VYH_G, 1VZ7_B, 1WCD_J, 1X24_B, 1X3G_B, 1X9J_G, 1XB4_C, 1XFB_L, 1XGO_A, 1XIQ_B, 1XN1_G, 1XXH_C, 1Y6E_B, 1YAB_B, 1YCF_B, 1YKH_B, 1YPO_G, 1Z7N_B, 1ZCF_G, 1ZP2_A, 1ZTM_C, 1ZYE_L, 2A81_B, 2A8Z_A, 2ABM_G, 2AHD_B, 2ALA_A, 2ANC_C, 2AV5_B, 2AW6_A, 2AXT_b, 2AYU_A, 2B26_A, 2B3T_B, 2BHV_A, 2BT8_A, 2BVG_B, 2BWE_T, 2FT3_F, 2FXT_A, 2G0B_B, 2G4C_C, 2GFP_B, 2GHU_B, 2GQQ_C, 2HAE_B, 2HSM_A, 2HZ6_A, 2I5K_B, 2IJZ_L, 2J8A_A, 2JGT_A, 2NOT_B, 2NSB_A, 2NTY_B, 2O6U_B, 2O96_B, 2OGK_B, 2OGU_A, 2OKH_B, 2OM5_A, 2OWM_B, 2P5T_D, 2P6G_B, 2PJY_A, 2POP_A,

2PRO_B, 2PV3_B, 2Q06_A, 2Q19_X, 2QYG_B, 2R1A_C, 2R5I_L, 2UU7_L, 2UZZ_B, 2V4D_F, 2VF9_B, 2VFJ_D, 2VH2_B, 2VHI_G, 2VMK_B, 2VR9_A, 2VV5_B, 2VZA_G, 2VZN_A.

SDB-1:

1AKO_A, 1ARB_A, 1BGF_A, 1CEM_A, 1CHD_A, 1CKA_A, 1CUS_A, 1DSB_A, 1DTO_A, 1ERZ_A, 1EUR_A, 1IFC_A, 1JR2_A, 1LAU_E, 1LMB_4, 1LKK_A, 1MKB_A, 1MLA_A, 1PDO_A, 1RVA_A, 1TTA_A, 1UBI_A, 1WHI_A, 1XYZ_A, 2BBK_H, 2CPL_A, 2LIS_A, 2OVO_A, 3PTE_A, 6XIA_A.

SDB-2:

1EJG_A, 1F94, 1I1W_A, 1IQZ, 1K2A, 1L9L_A, 1M40_A, 1MJ5, 1MWQ, 1NQJ, 1OAI_A, 1OD3_A, 1P9G, 1R6J_A, 1TG0_A, 1TT8, 1UCS, 1UG6_A, 1VBW_A, 2A6Z_A, 2B97_B, 2CWS_A, 2FMA, 2FVY, 2FWH_A, 2GUD, 2H3L_A, 2I4A, 2P74, 2PND_A, 2QCP, 2QSK_A, 2R31_A, 2RBK, 2RH2, 2VHK_A, 2XU3_A, 2Y78_A, 3A02_A, 3A38, 3AGN_A, 3AJ4_A, 3AKS_A, 3BWH_A, 3CCD, 3F1L_A, 3G21_A, 3GOE_A, 3H31_A, 3KLR_A, 3L8W, 3MFJ, 3NE0_A, 3ODV_B, 3PSM, 3QPA, 3RCG, 3RQ9_B, 3RWN_A, 3SOJ, 3U7T_A, 3V1A_A, 3VII_A, 3VLA_A, 3ZR8_X, 4AXO_B, 4F1V_A, 7A3H_A

SDB-3:

1C1K_A, 1CEM_A, 1CKA_A, 1DQZ, 1INL, 1J3W, 1LAU_E, 1MKB_A, 1MLA_A, 1NTV, 1NWA_A, 1RKU_B, 1STN_A, 1TUA_A, 2CPL_A, 2ERF, 2G8O, 2GXG, 2HAQ, 2I3F, 2LIS_A, 2NR7, 2Q0L_B, 3B4U, 3BU1.

OUDB:

Obsolete	Resolution	R-factor	Upgraded	Resolution	R-factor
2GCR	-	-	1A45	2.3	0.186
1ABP	2.4	-	1ABE	1.7	0.137
1CAC	-	-	1CA2	2.0	0.173
1CD4	-	-	1CDH	2.3	0.193
1B5C	-	-	1CYO	1.5	0.16
1FNR	-	-	1FNB	1.7	0.179
3GAP	-	-	1G6N	2.1	0.207
1LH4	-	-	1GDJ	1.7	0.165
1DB0	3.4	0.242	1H1V	2.99	0.219

1LUS	2.11	0.242	1N0N	1.82	0.224
1KYB	2.8	0.217	1O0R	2.3	0.203
1LW8	2.5	0.221	1O4V	2.0	0.219
1DHX	2.9	0.225	1P2Z	2.2	0.178
14PS	2.6	0.246	1QJB	2.0	0.21
1JJY	1.63	0.161	1SU8	1.1	0.146
1GDO	1.8	-	1XFF	1.8	0.161
156B	2.5	-	256B	1.4	0.164
1ACE	-	-	2ACE	2.5	0.199
1ACT	2.8	-	2ACT	1.7	0.171
1J5V	2.3	0.206	2AFB	2.05	0.177
1AFN	2.6	0.176	2AFN	2.0	0.158
1ALP	2.8	-	2ALP	1.7	0.131
1APR	-	-	2APR	1.8	0.143
1AYH	2.0	0.164	2AYH	1.6	0.143
1AZA	-	-	2AZA	1.8	0.157
1BAA	2.8	0.2	2BAA	1.8	0.18
1C2C	-	-	2C2C	2.0	0.172
1OKU	2.9	0.209	2C5V	2.9	0.175
1CAB	-	-	2CAB	2.0	0.193
1CI2	-	-	2CI2	2.0	0.198
1CPK	-	-	2CPK	2.7	0.18
1CYP	-	-	2CYP	1.7	0.202
1DTR	2.8	0.21	2DTR	1.9	0.171
1EIP	2.7	0.209	2EIP	2.2	0.183
1F19	-	-	2F19	2.8	0.182
1FB4	-	-	2FB4	1.9	0.189
1FBJ	-	-	2FBJ	1.95	0.194
1GCH	1.9	-	2GCH	1.9	0.181
1HFT	2.4	0.224	2HFT	1.69	0.204

1HMG	-	-	2HMG	3.0	0.222
1HMM	-	-	2HMQ	1.66	0.189
1HMZ	-	-	2HMZ	1.66	0.176
1HVM	2.2	0.169	2HVM	1.8	0.149
1IG2	-	-	2IG2	3.0	0.207
1ISD	2.5	-	2ISD	2.5	0.227
1LHB	-	-	2LHB	2.0	0.142
1LZM	-	-	2LZM	1.7	0.193
1MBR	2.7	0.222	2MBR	1.8	0.203
1MBW	-	-	2MBW	1.5	0.179
1MHB	-	-	2MHB	2.0	0.23
1OMF	2.4	0.208	2OMF	2.4	0.167
1ADT	2.6	0.197	2WB0	1.95	0.187
151C	-	-	351C	1.6	0.195
1APP	-	-	3APP	1.8	0.126
1BJL	-	-	3BJL	2.3	0.165
1COX	-	-	3COX	1.8	0.156
1DFR	-	-	3DFR	1.7	0.152
2IJP	2.14	0.237	3EFZ	2.08	0.227
3BCL	-	-	3EOJ	1.3	0.137
3DBB	2.15	0.212	3F70	2.1	0.203
2GRS	-	-	3GRS	1.54	0.186
2HFL	-	-	3HFL	2.65	0.196
2HFL	-	-	1YQV	1.7	0.195
2HSC	-	-	3HSC	1.93	0.21
1ICB	-	-	3ICB	2.3	0.178
1M80	2.35	0.223	3M10	1.73	0.192
3AAH	2.4	0.201	4AAH	2.4	0.152
2APE	-	-	4APE	2.1	0.158
1GCR	-	-	4GCR	1.47	0.181

1HHB	-	-	4HHB	1.74	0.135
2MDH	-	-	4MDH	2.5	0.167
7ATC	-	-	5AT1	2.6	0.16
1CPV	1.85	-	5CPV	1.6	0.187
2FD1	-	-	5FD1	1.9	0.215
1LYM	-	-	5LYM	1.8	0.189
1PTP	-	-	5PTP	1.34	0.152
1FDC	2.1	0.228	6FDR	1.4	0.175
2LDH	-	-	6LDH	2.0	0.202
1FAB	-	-	7FAB	2.0	0.169
3CAT	-	-	8CAT	2.5	0.191
2ADK	-	-	3ADK	2.1	0.193
1B2N	1.9	0.183	1G72	1.9	0.161
1B8B	2.75	0.244	1H7A	2.75	0.215
1BFL	2.27	0.19	1CNQ	2.27	0.168
1BOV	2.2	0.177	2XSC	2.05	0.145
1DGT	2.9	0.254	1V9P	2.9	0.232
1DN1	2.6	0.239	3C98	2.6	0.208
1DOO	2.81	0.301	1E94	2.8	0.254
1DYV	1.7	0.232	1UN2	2.4	0.211
1EIM	2.0	0.187	1KW9	1.95	0.162
1END	1.6	0.196	2END	1.45	0.161
1GGA	3.2	0.176	2X0N	3.2	0.153
1GLY	2.2	0.23	3GLY	2.2	0.141
1GSR	2.3	0.241	2GSR	2.11	0.165
1GST	2.2	0.171	6GST	2.2	0.149
1GT2	2.91	0.227	2X0R	2.91	0.189
1HYY	1.8	0.228	2DQU	1.7	0.207
1J5R	1.4	0.212	1O2D	1.3	0.139
1J6N	1.8	0.211	1O58	1.8	0.159

1J8W	2.0	0.231	1NF5	2.0	0.19
1J8X	2.0	0.246	1NKH	2.0	0.196
1J92	2.0	0.241	1NQI	2.0	0.211
1J94	2.5	0.241	1NHE	2.5	0.197
1JN8	2.0	0.234	1NMM	2.0	0.189
1JNA	2.3	0.258	1NWG	2.32	0.208
1JVF	2.5	0.254	2PD3	2.5	0.233
1JW7	2.3	0.262	2PD4	2.3	0.224
1KSA	2.2	0.19	3ENI	2.2	0.166
1L7W	2.1	0.227	1OQM	2.1	0.194
1MFH	3.0	0.295	2AKQ	3.0	0.214

[1]Basu, S., Bhattacharyya, D., Banerjee, R. (2012). *Biophys. J.* **102**, 2605-2614.