Specialized Dynamical Properties of Promiscuous Residues Revealed by Simulated Conformational Ensembles

Arianna Fornili, Alessandro Pandini, Hui-Chun Lu, Franca Fraternali

Supporting Information

I. Supplementary Figures:

II Supplementary Tables

Table S10. Average per-residue entropy for $c_{mono\_in\_multi}$ and c_{multi} residues in S^{MD}.

Table S11. Percentage of c_{multi} residues classified as hot spots in 1 to 4 different interfaces.

Table S12. SNP and DisSNP propensity in the S^{Full} and S^{Soc} datasets (values per protein).

Table S13. C^{α} and side chain RMSF Z-scores of $c_{mono\_in\_multi}$ and c_{multi} interface residues of Neddylin.

Table S14. C^{α} and side chain RMSF Z-scores of $c_{mono\_in\_multi}$ and c_{multi} interface residues of SUMO-2.

I. Supplementary Figures

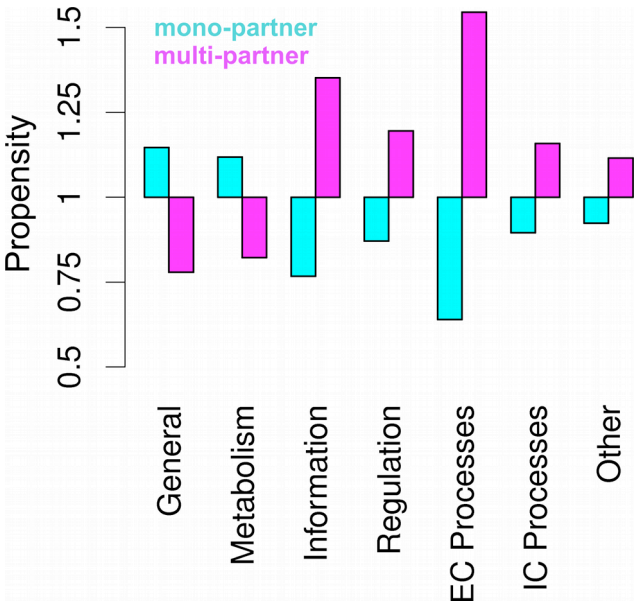

Figure S1. Functional annotation of S^{Full}. Propensity of mono-partner (cyan) and multi-partner (magenta) proteins for 7 general function categories (Vogel et al. (2006) PLoS Comput Biol 2: e48). Propensities are calculated as the abundance of a function category in each protein binding class relative to the average abundance in S^{Full}.

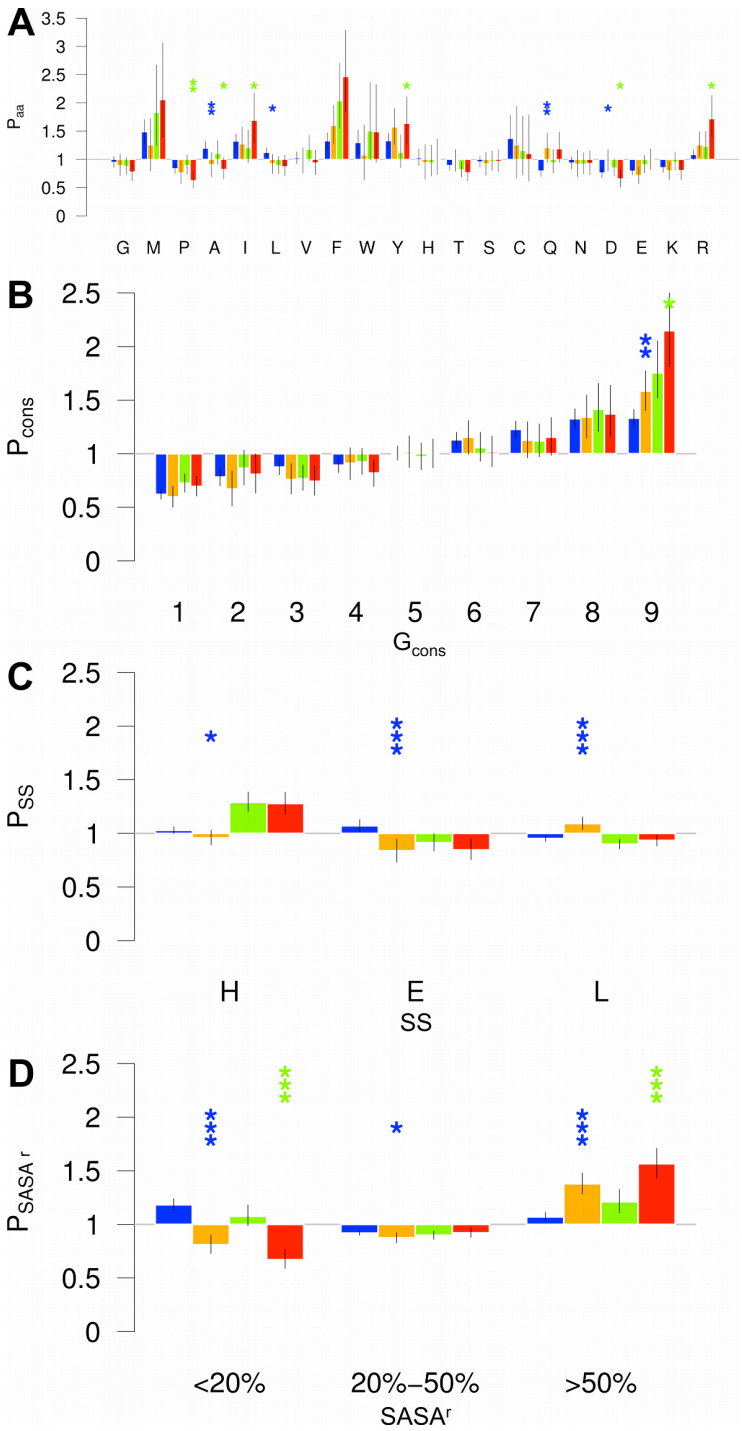

Figure S2. Physicochemical properties of interface residues in the $c_{mono\_in\_multi}$ and c_{multi} binding classes of S^{Soc}. (A-D) Propensities relative to the surface at $c_{mono\_in\_multi}$ (green) and c_{multi} (red) positions of the S^{Soc} dataset. Results from $c_{mono\_in\_multi}$ (blue) and c_{multi} (orange) residues of S^{Full} are reported as reference. Error bars and significance levels (indicated with a star code) were estimated with bootstrap resampling (see Figure 2 legend). The significance levels are reported only for comparisons between residues from the same dataset. (A) Amino acid propensity P_{aa}. (B) Conservation propensity P_{cons}. The conservation is expressed as ConSurf conservation grade (G_{cons}). (C) DSSP secondary structure (SS) propensity P_{SS}. (D) Relative solvent accessibility propensity P_{SASA}^{r}.

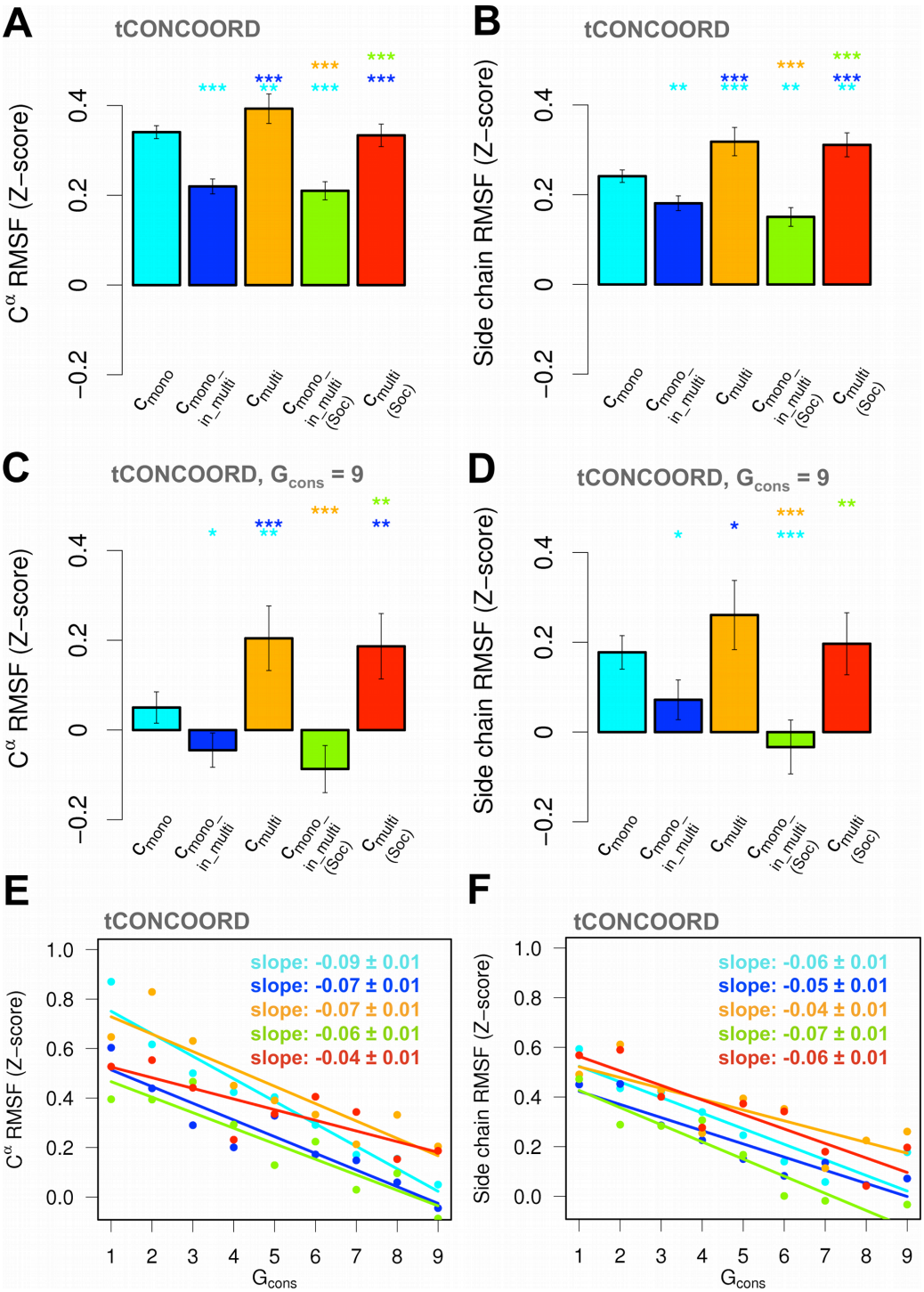

Figure S3. Conformational flexibility of interface residues in S^{Full} and S^{Soc}. (A-D) Average C^{α} (A/C) and side chain (B/D) RMSF Z-scores (tCONCOORD) of all interface residues (A/B) and of the most conserved (G_{cons} = 9) interface residues (C/D) for the three binding classes c_{mono} (cyan), $c_{mono\_in\_multi}$ (blue) and c_{multi} (orange) of S^{Full} and the two binding classes $c_{mono\_in\_multi}$ (green) and c_{multi} (red) of S^{Soc}. The standard error of the mean is represented with an error bar. The significance levels from pair-wise Wilcoxon comparison tests are reported with a star code (see Figure 2 legend). (E/F) Dependence of tCONCOORD average C^{α} (E) and side chain (F) RMSF Z-scores (dots) from the ConSurf conservation grade (G_{cons}). A best-fit linear regression is also reported for each binding class.

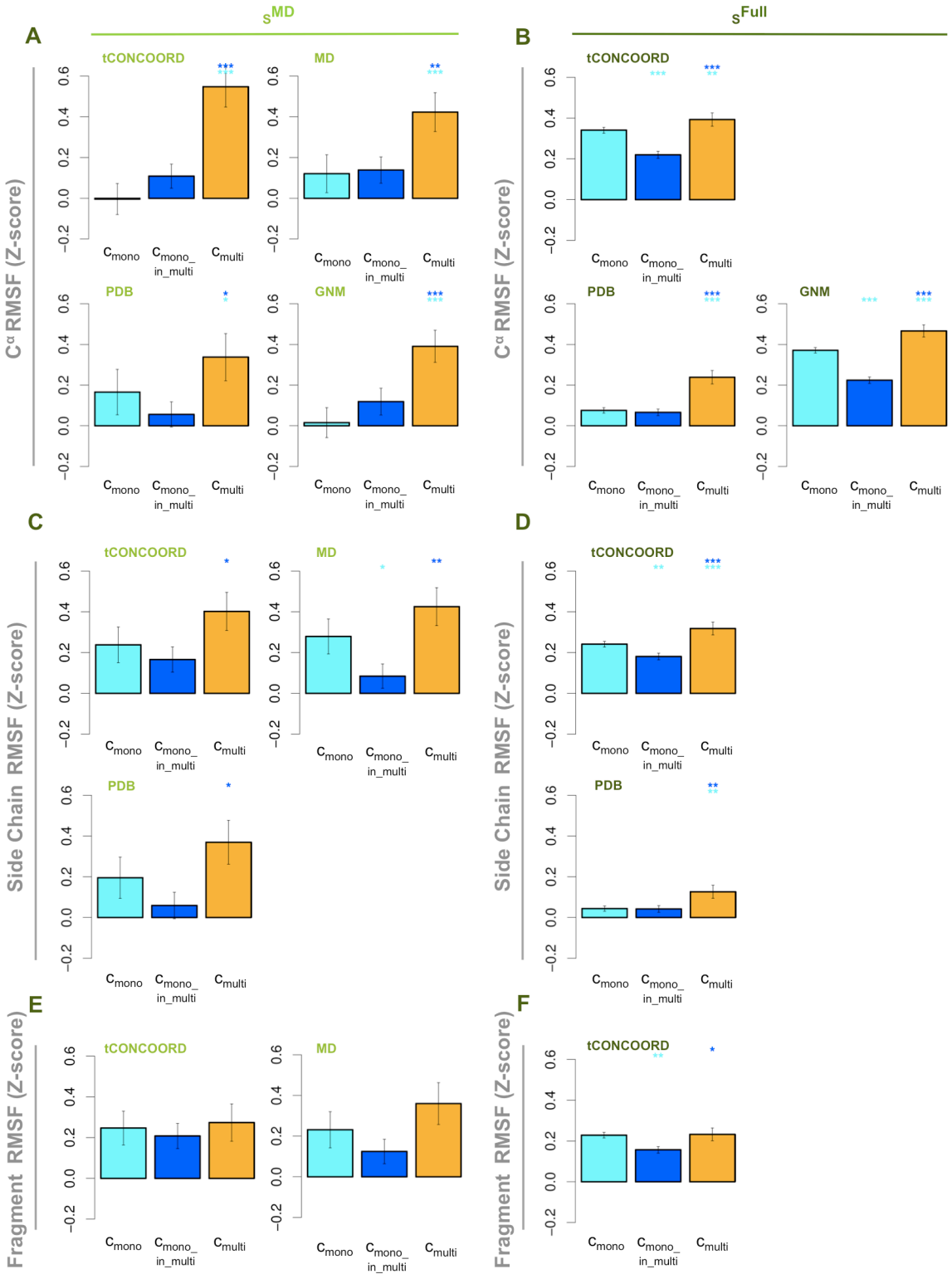

Figure S4. Comparison of conformational flexibility of S^{Full} and S^{MD} interface residues from simulated ensembles (tCONCOORD, MD, GNM) and experimental structures (PDB). Average C^{α} (A/B), side chain (C/D) and fragment (E/F) RMSF Z-scores of interface c_{mono} (cyan), $c_{mono\_in\_multi}$ (blue) and c_{multi} (orange) residues in S^{MD} (left panels) and S^{Full} (right panels). The standard error of the mean is represented with an error bar. The significance levels from pair-wise Wilcoxon comparison tests are reported with a star code (see Figure 2 legend).

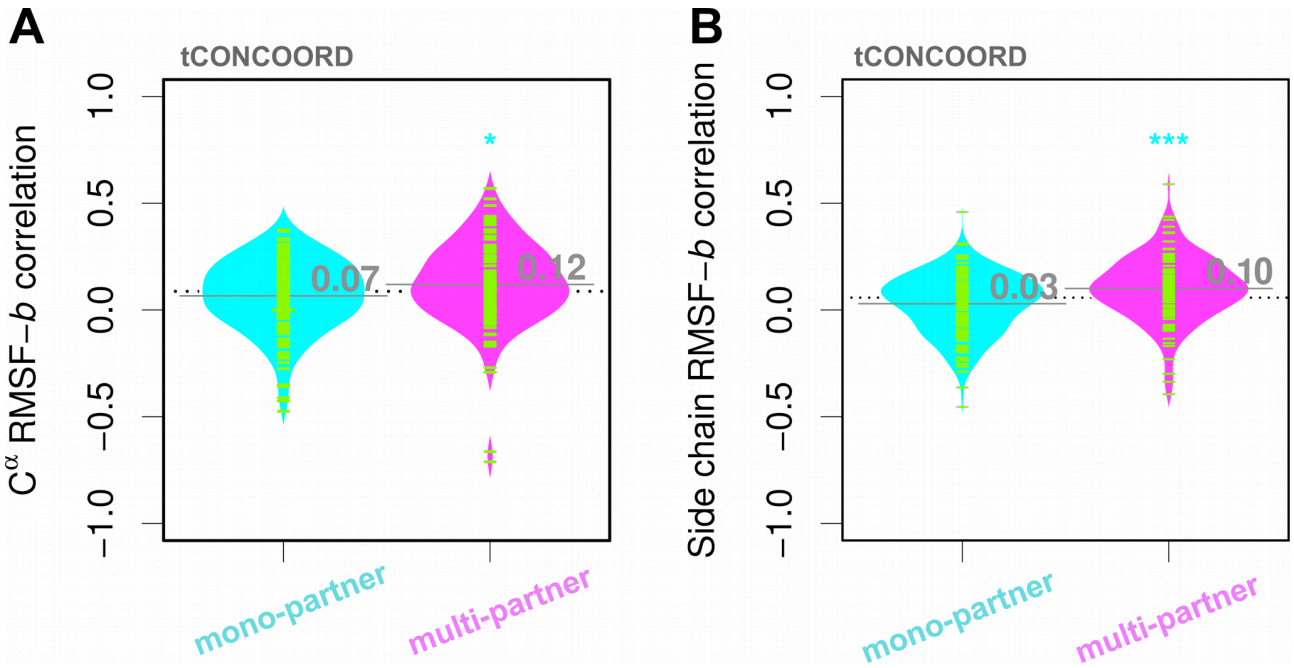

Figure S5. Correlation between RMSF and binding multiplicity profiles in S^{Full}. The distributions of correlation coefficients (Spearman) between RMSF and binding multiplicity (*b*) profiles are reported as beanplots for mono-partner (cyan) and multi-partner (magenta) proteins, using C^{α} (A) and side chain (B) RMSF values from tCONCOORD ensembles. Only exposed residues were considered in the calculation of the correlation coefficient. The averages of the total and of the single distributions are indicated with dotted black and solid grey lines, respectively. Individual observations are represented as green lines. The numeric values of the averages are also reported in grey for each distribution. The significance levels from a Wilcoxon comparison test between the mono- and multi-partner distributions are reported with a star code (see Figure 2 legend).

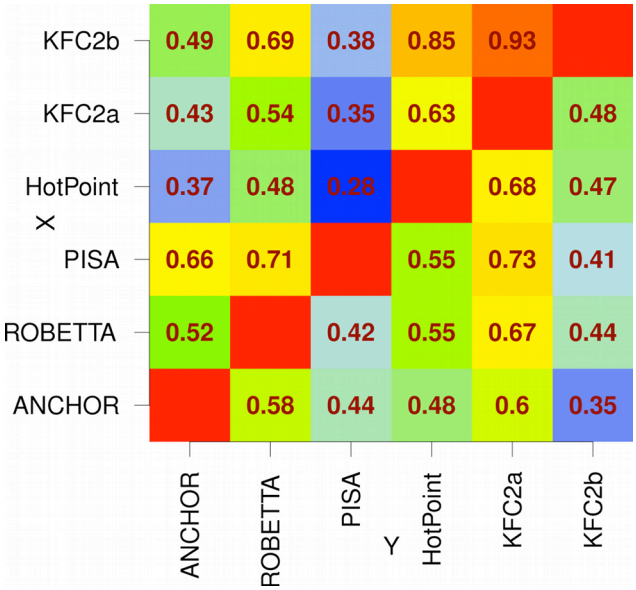

Figure S6. Pairwise comparison of hotspot prediction methods. The overlap between pairs of methods (X, Y) is calculated as the ratio between the number of residues classified as hotspots by both methods and the number of residues classified as hotspots by the X method. The overlap values are reported in red, while the cells are colored according to the overlap from blue (0.0) to red (1.0).

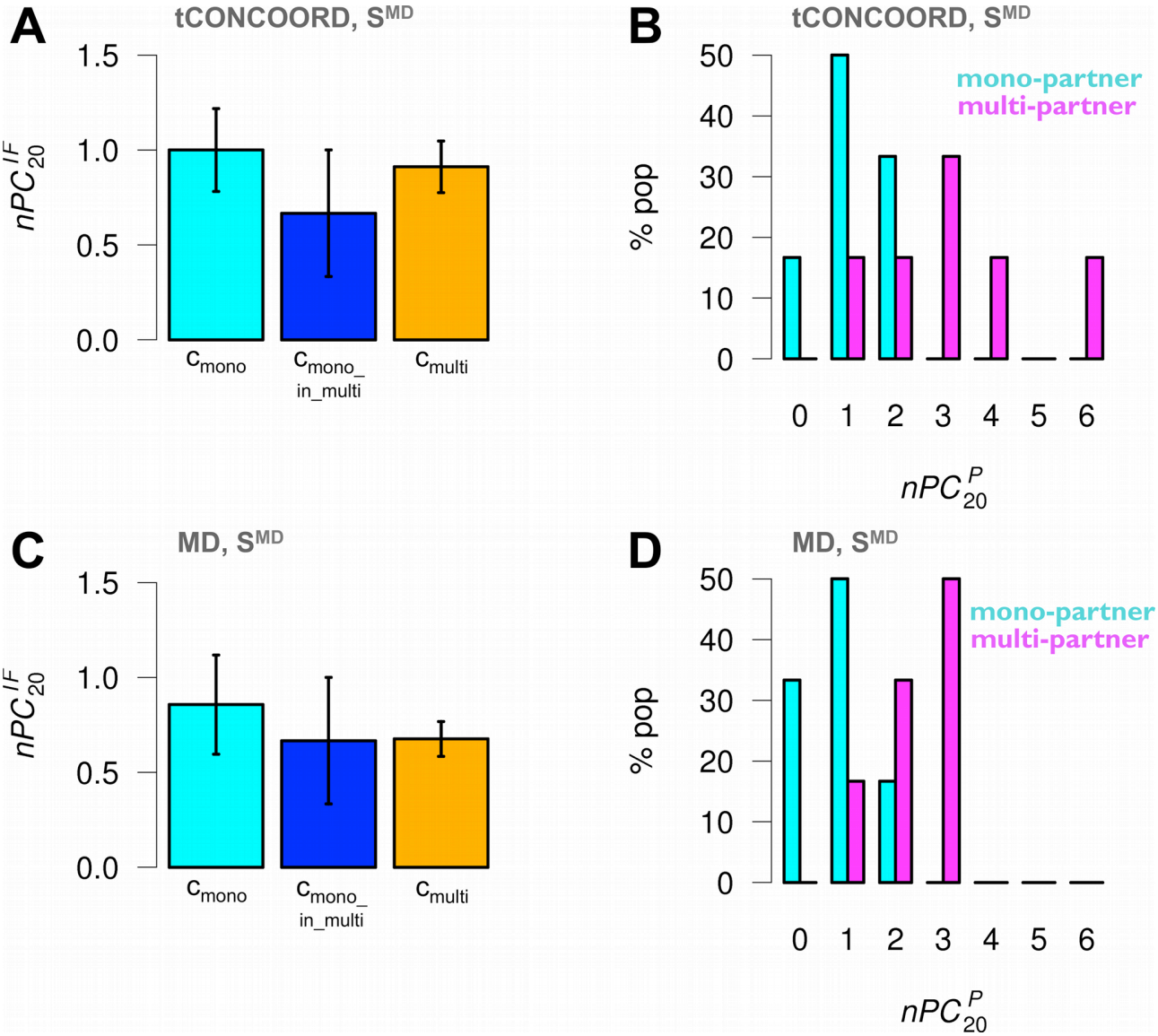

Figure S7. **Analysis of correlated motions in S^{MD}.** The correlation between collective motions and interfaces is analysed by functional mode analysis (FMA) on tCONCOORD and MD ensembles of the S^{MD} dataset. (A/C) Number of tCONCOORD (A) and MD (C) PCs accounting for at least 20% of the R_g^{IF} variance of a given interface (nPC_{20}^{IF}). Average values calculated over interfaces of the three binding classes c_{mono1} (cyan), $c_{mono\_in\_multi}$ (blue) and c_{multi} (orange) are reported. The standard error of the mean is represented with an error bar. The significance levels from pair-wise Wilcoxon comparison tests are reported with a star code (see Figure 2 legend). (B/D) Distributions of tCONCOORD (B) and MD (D) nPC_{20} per protein (nPC_{20}^{P}), calculated as the number of unique PCs accounting for at least 20% of the R_g^{IF} variance of any interface in the protein. Values for mono-partner (cyan) and multi-partner (magenta) proteins are reported.

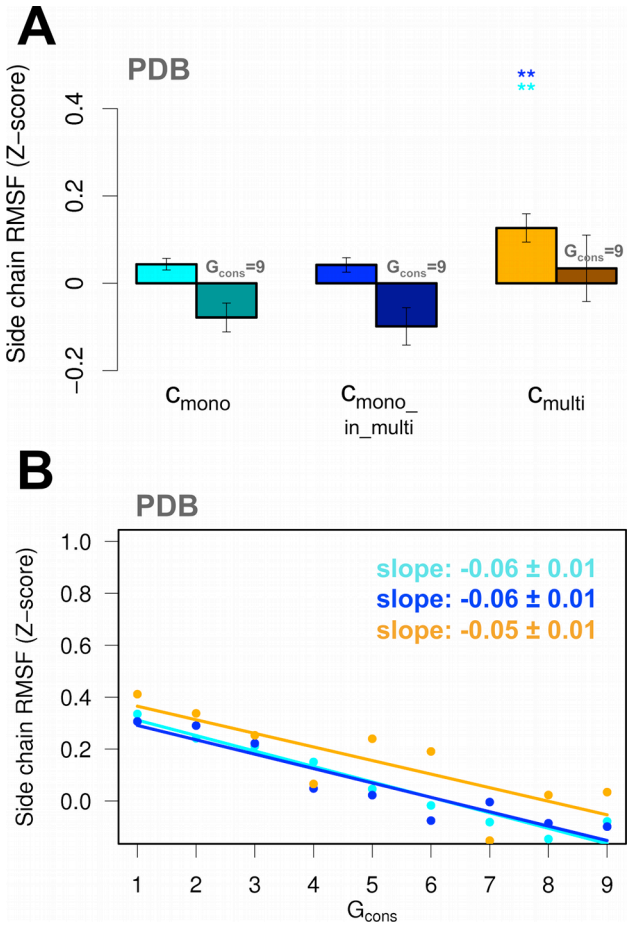

Figure S8. Analysis of the conformational variability within PDB ensembles in S^{Full} (side chains). (A) Average side chain RMSF Z-scores (PDB ensembles) calculated over c_{mono} (cyan), $c_{mono\_in\_multi}$ (blue) and c_{multi} (orange) residues in the S^{Full} dataset. The standard error of the mean is represented with an error bar. The significance levels from pair-wise Wilcoxon comparison tests are reported with a star code (see Figure 2 legend). Averages and significance levels calculated considering only the residues with the highest ConSurf conservation grade (G_{cons} = 9) are also reported in dark colours. (B) Dependence of PDB average side chain RMSF Z-scores (dots) from evolutionary conservation for c_{mono} (cyan), $c_{mono\_in\_multi}$ (blue), c_{multi} (orange) residues. Residues are partitioned into 9 groups according to their ConSurf conservation grade (G_{cons}). A best-fit linear regression is also reported for each binding class.

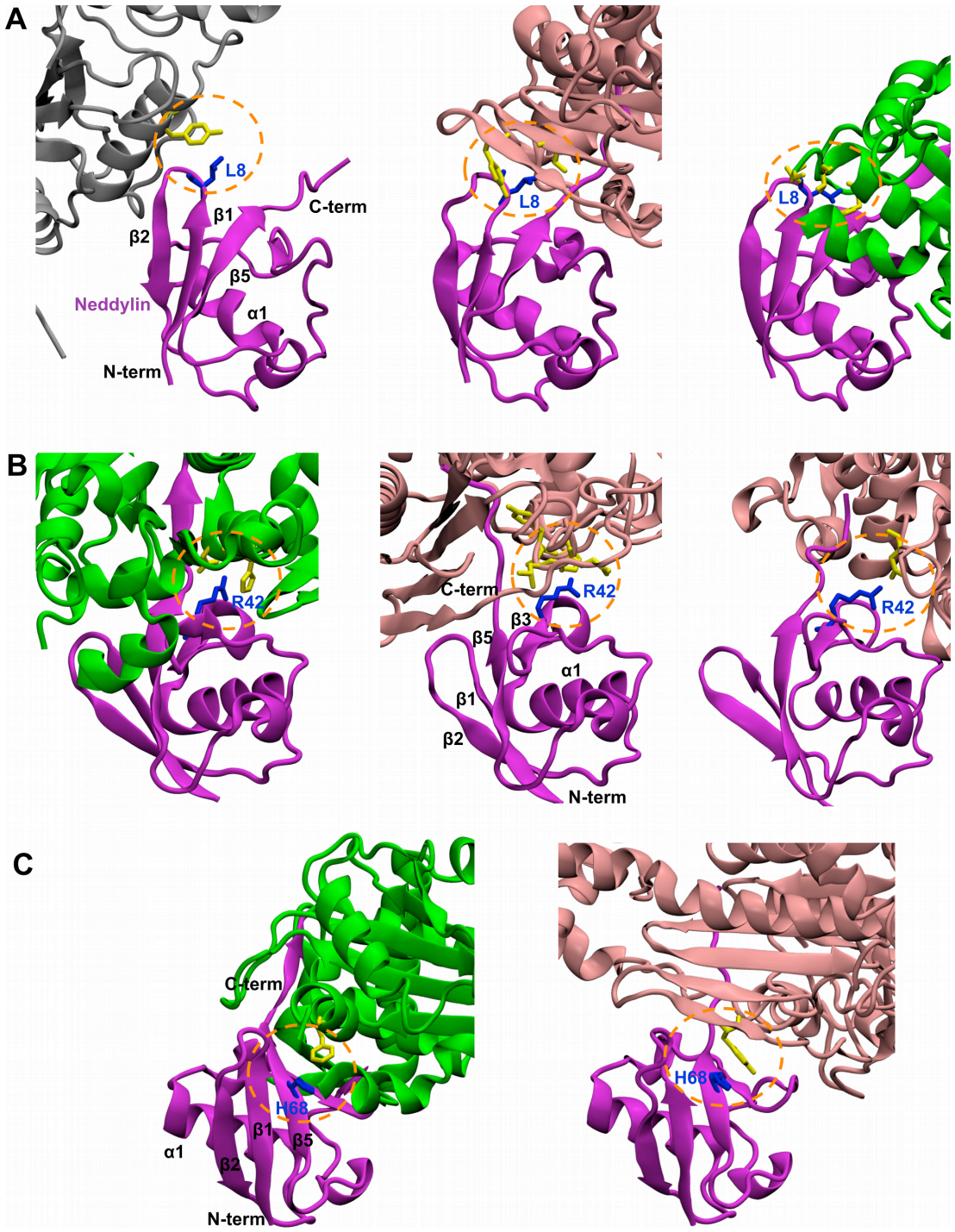

Figure S9. Examples of multi-partner residues in Neddylin. Cartoon representation of the complexes between Neddylin (magenta) and selected partners (NEDD8-conjugating enzyme Ubc12 in grey from PDB ID 2nvu, chain J; NEDD8-activating enzyme E1 regulatory subunit in pink from PDB ID 1r4n, chain B, and from PDB ID 2nvu, chain A; sentrin-specific protease 8 in green from PDB ID 1xt9, chain A). The three panels highlight the conformations adopted in the different complexes by L8 (A), R42 (B) and H68 (C), represented as blue sticks. In each panel, the partner residues within 4.0 Å from any atom of the highlighted Neddylin residue are represented as yellow sticks.

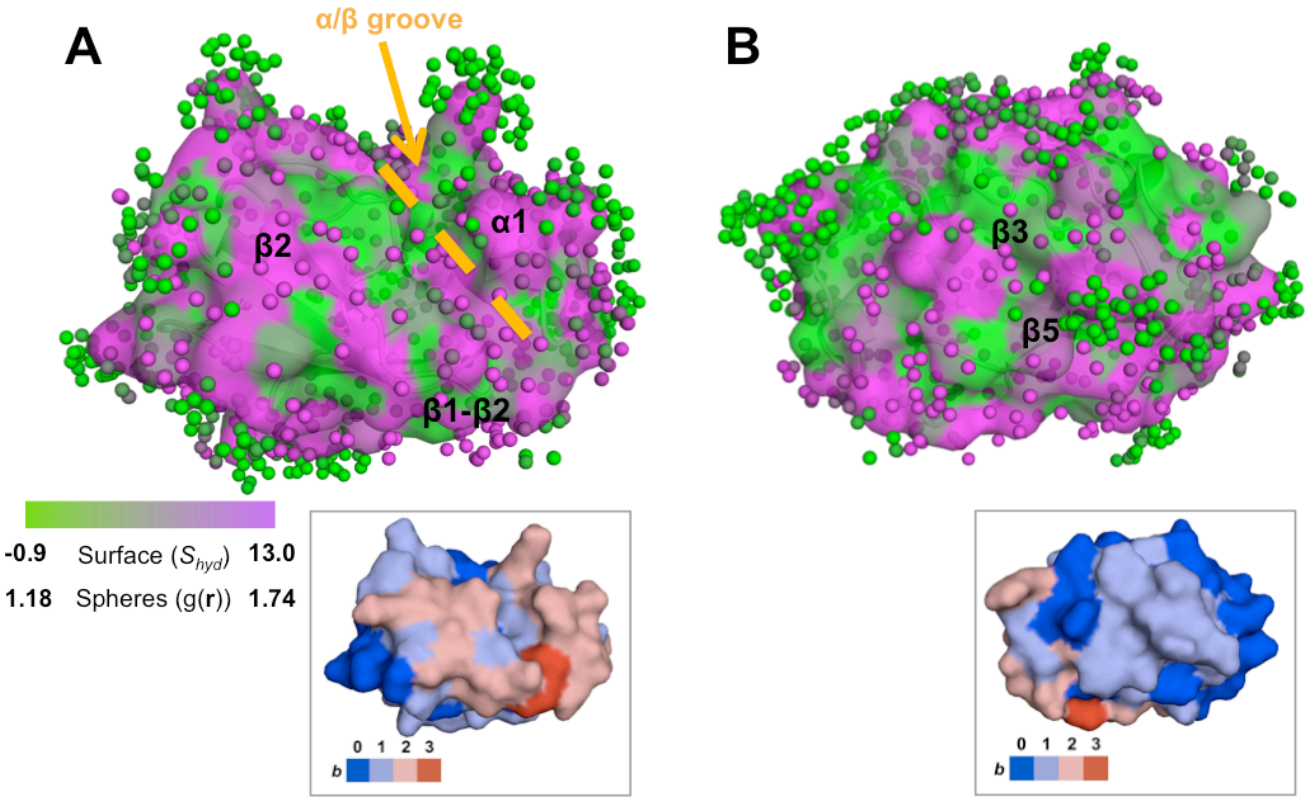

Figure S10. Water distribution around the SUMO-2 protein. The local maxima (hydration sites) of the water density map used for the calculation of the hydration score S_{hyd} are represented as spheres, coloured from green to magenta according to the increasing g(**r**) value. The protein surface is coloured from green to magenta according to the increasing hydration score. Hydrophobic atoms (green surface) have low hydration scores and are poor in hydration sites. The colour scale used for S_{hyd} (surface) and g(**r**) (spheres) is reported on the left. Two different orientations are shown, highlighting the multi-partner (A) and mono-partner (B) side of the molecule. The binding multiplicity is colour-mapped onto each side in the two insets.

II. Supplementary Tables

Table S1. List of the 251 proteins in the S^{Full} dataset.

FR[a]	SR[b]	Protein Name	np_s[c]	nres[d]	resol[e]	SCOP class
1ewdA	1aldA	Fructose-1,6-bisphosphate aldolase	1	363	2	c
1d2fA	1d2fA	Modulator in mal gene expression, MalY	1	361	2.5	c
1ck2A	1nmuD	Eukaryotic ribosomal protein L30 (L30e)	1	103	2.31	d
1orjA	1orjD	Flagellar export chaperone FliS	1	125	2.25	a
3c62A	256bB	Cytochrome b562	1	106	1.4	a
1n9sA	1n9rG	Small nuclear ribonucleoprotein F, Smf	1	68	2.8	b
1etkA	1f36B	FIS protein	1	89	2.65	a
1psrA	1psrB	Calcyclin (S100)	1	100	1.05	a
1dhrA	1hdrA	Dihydropteridin reductase (pteridine reductase)	1	236	2.5	c
1otkA	1otkB	Phenylacetic acid degradation protein PaaC	1	242	2	a
1gg1A	1n8fD	3-deoxy-D-arabino-heptulosonate-7-phosphate synthase (DAHP synthase, AroG)	1	343	1.75	c
2av6A	1l5wB	Maltodextrin phosphorylase (MALP)	1	796	1.8	c
1i1rB	1i1rB	Interleukin-6	1	167	2.4	a
2bsxA	1q1gF	Putative uridine phosphorylase	1	243	2.02	c
2a78B	2c8eE	MONO-ADP-RIBOSYLTRANSFERASE C3	1	200	1.6	d
1bsvA	1e6uA	GDP-4-keto-6-deoxy-d-mannose epimerase/reductase (GDP-fucose synthetase)	1	315	1.45	c
1a4pA	1a4pA	Calcyclin (S100)	1	92	2.25	a
1x8dA	1x8dD	L-rhamnose mutarotase YiiL	1	104	1.8	d
1i7fA	1hw7A	Heat shock protein 33, Hsp33	1	229	2.2	d
1fjjA	1fjjA	Hypothetical protein YbhB	1	153	1.66	b
1surA	1surA	Phosphoadenylyl sulphate (PAPS) reductase	1	215	2	c
2incC	2incC	Toluene, o-xylene monooxygenase oxygenase subunit TouB	1	83	1.85	d
2i3sA	1yfqA	Cell cycle arrest protein BUB3	1	342	1.1	b
1g57A	1iezA	3,4-dihydroxy-2-butanone 4-phosphate synthase, DHBP synthase, RibB	1	217	NMR	d
1eu5A	1euwA	Deoxyuridine 5'-triphosphate nucleotidohydrolase (dUTPase)	1	135	1.05	b
2dgkA	2dgkF	Glutamate decarboxylase beta	1	438	1.9	c
1cgqA	1gd0C	Microphage migration inhibition factor (MIF)	1	118	1.5	d
1cviA	1nd6D	Prostatic acid phosphatase	1	342	2.4	c
1ixnA	1m5wH	Pyridoxine 5'-phosphate synthase	1	242	1.96	c
1aroL	1aroL	Bacteriophage T7 lysozyme (Zn amidase)	1	149	2.8	d
2g46A	2g46B	Viral histone H3 Lysine 27 Methyltransferase	1	119	NMR	b
1dfgA	1qsgG	Enoyl-ACP reductase	1	259	1.75	c
1xviA	1xviB	Putative mannosyl-3-phosphoglycerate phosphatase MPGP (YedP)	1	234	2.26	c
1aiuA	2ifqB	Thioredoxin	1	105	1.2	c

1av8A	1mxrB	Ribonucleotide reductase R2	1	339	1.42	a
1eo6A	1eo6A	Golgi-associated ATPase enhancer of 16 kD, Gate-16	1	116	1.8	d
1qvvA	1rw7A	Hypothetical protein Ydr533Cp	1	235	1.8	c
1b4uA	1bouC	LigA subunit of an aromatic-ring-opening dioxygenase LigAB	1	132	2.2	a
1a99A	1a99D	Putrescine receptor (PotF)	1	341	2.2	c
2hj9A	1jx6A	Quorum-sensing signal (autoinducer-2) binding protein LuxP	1	338	1.5	c
1wmiB	1wmiD	Hypothetical protein PHS014	1	61	2.3	a
1izyA	1n57A	HSP31 (HchA; YedU)	1	272	1.6	c
2e98A	1uehB	Undecaprenyl diphosphate synthase	1	225	1.73	c
2ae0X	2ae0X	Membrane-bound lytic murein transglycosylase A, MLTA	1	335	2	b
1hnoA	1k39B	Dienoyl-CoA isomerase (delta3-delta2-enoyl-CoA isomerase)	1	271	3.29	c
3c4oB	2g2uB	beta-lactamase-inhibitor protein, BLIP	1	165	1.6	d
1cmaA	1cmcB	Met repressor, MetJ (MetR)	1	104	1.8	a
2bdxA	1jk7A	Protein phosphatase-1 (PP-1)	1	294	1.9	d
1af6A	1af6C	Maltoporin (also LamB protein)	1	421	2.4	f
2qw7A	2qw7J	Carbon dioxide concentrating mechanism protein CcmL	1	96	2.4	b
2d5rA	2d5rA	CCR4-NOT transcription complex subunit 7, CAF1	1	252	2.5	c
1c0nA	1jf9A	NifS-like protein/selenocysteine lyase	1	405	2	c
2cluA	2cihA	FERRITIN HEAVY CHAIN	1	172	1.5	a
1b9lA	1b9lH	7,8-dihydroneopterin triphosphate epimerase	1	119	2.9	d
1hslA	1hslB	Histidine-binding protein	1	239	1.89	c
1bh5A	1qipD	Glyoxalase I (lactoylglutathione lyase)	1	183	1.72	d
1ex6A	1ex7A	Guanylate kinase	1	186	1.9	c
1bykA	1bykB	Trehalose repressor, C-terminal domain	1	255	2.5	c
1r5tA	1r5tD	mono-domain cytidine deaminase	1	134	2	c
1gynA	1dosB	Fructose-bisphosphate aldolase (FBP aldolase)	1	358	1.67	c
1dhsA	1rozA	Deoxyhypusine synthase, DHS	1	336	2.21	c
1cb7A	1ccwC	Glutamate mutase, small subunit	1	137	1.6	c
1ga7A	1g0sB	ADP-ribose pyrophosphatase	1	202	1.9	d
1jdiA	1k0wF	L-ribulose-5-phosphate 4-epimerase	1	223	2.1	c
2avyD	2qanD	Ribosomal protein S4	1	205	3.21	d
2avyM	2qanM	Ribosomal protein S13	1	113	3.21	a
2avyL	2qanL	Ribosomal protein S12	1	123	3.21	b
2avyO	2qanO	30S ribosomal protein S15	1	88	3.21	a
1fwqA	1fwqA	RabGEF Mss4	1	115	NMR	b
2c4nA	2c4nA	NagD	1	250	1.8	c
1ejcA	1uaeA	UDP-N-acetylglucosamine enolpyruvyl transferase (EPT, MurA, MurZ)	1	418	1.8	d
2cf2E	1q7bD	beta-keto acyl carrier protein reductase	1	243	2.05	c
1euiC	1ugiH	Uracil-DNA glycosylase inhibitor protein	1	83	1.55	d

2e2dC	2e2dC	TIMP-2	1	179	2	b
2fjcK	2fjcL	Dodecameric ferritin homolog	1	151	2.5	a
1dzuP	1e4cP	L-fuculose-1-phosphate aldolase	1	206	1.66	c
1dqwA	1dqwD	Orotidine 5'-monophosphate decarboxylase (OMP decarboxylase)	1	267	2.1	c
1f5mA	1f5mB	Hypothetical protein ykl069wp	1	177	1.9	d
1l6wA	1l6wJ	Decameric fructose-6-phosphate aldolase/transaldolase	1	220	1.93	c
1cddA	1jkxB	Glycinamide ribonucleotide transformylase, GART	1	209	1.6	c
1f05A	2e1dB	Transaldolase	1	321	2	c
1tuvA	1tuvA	Hypothetical protein YgiN	1	103	1.7	d
1e8oA	1e8oC	Signal recognition particle 9KDa protein, SRP9	1	71	3.2	d
1poiB	1poiD	Glutaconate:CoA transferase beta	1	260	2.5	c
1gggA	1wdnA	Glutamine-binding protein	1	224	1.94	c
2es4A	1tahD	Lipase	1	318	3	c
1qsmA	1qsmD	Histone acetyltransferase HPA2	1	152	2.4	d
1oroA	1oroA	Orotate PRTase	1	212	2.4	c
1fteA	1fthA	Holo-(acyl carrier protein) synthase ACPS	1	117	1.9	d
1d6kA	1dfuP	Ribosomal protein L25	1	94	1.8	b
1v74B	1v74B	Colicin D immunity protein	1	87	2	a
2c0aA	2c0aC	KDPG aldolase	1	213	1.55	c
1a03A	1k96A	Calcyclin (S100)	1	89	1.44	a
1qsdA	1qsdB	Tubulin chaperone cofactor A	1	102	2.2	a
1s5uA	1s5uE	Hypothetical protein YbgC	1	136	1.7	d
2hiqA	2hiqB	Hypothetical protein YdhR	1	96	2	d
1nmoA	1nmpE	Hypothetical protein YbgI	1	247	2.2	c
1ajaA	1y6vB	Alkaline phosphatase	1	449	1.6	c
1erbA	1jydA	Retinol binding protein	1	174	1.7	b
1pugA	1pugA	Hypothetical upf0133 protein YbaB	1	94	2.2	d
11asA	12asB	Asparagine synthetase	1	328	2.2	d
1bs1A	1byiA	Dethiobiotin synthetase	1	224	0.97	c
2aw4Z	2awbZ	Ribosomal protein L31p	1	70	3.46	d
1n2zA	1n2zB	Vitamin B12 binding protein BtuF	1	245	2	c
1dj0A	1dj0B	Pseudouridine synthase I TruA	1	264	1.5	d
1av5A	1kpfA	Protein kinase C inhibitor-1, PKCI-1	1	111	1.5	d
1cl1A	2fq6B	Cystathionine beta-lyase, CBL	1	392	1.78	c
1p91A	1p91A	rRNA methyltransferase RlmA	1	261	2.8	c
1bd9A	1behB	Phosphatidylethanolamine binding protein, PEBP	1	183	1.75	b
2cmjA	1t0lD	NADP-dependent isocitrate dehydrogenase	1	414	2.41	c
2ht0A	1owfA	Integration host factor alpha subunit (IHFA)	1	96	1.95	a
1ai2A	1pb3A	Isocitrate dehydrogenase, ICDH	1	416	1.7	c
1jn5A	1jkgA	NTF2-related export protein 1 (p15)	1	139	1.9	d
1cksA	1cksA	CksHs2	1	74	2.1	d
1itqA	1ituB	Renal dipeptidase	1	369	2	c

3cptB	3cptB	Mitogen-activated protein-binding protein-interacting protein	1	116	1.9	d
1k4kA	1k4mC	Nicotinamide mononucleotide (NMN) adenylyltransferase	1	213	1.9	c
1xouB	1xouB	EspA chaperone CesA	1	84	2.8	a
1yacA	1yacB	YcaC	1	204	1.8	c
2bnfA	2bndA	Uridylate kinase PyrH	1	237	2.6	c
1lbqA	1lbqB	Ferrochelatase	1	354	2.4	c
1cm5A	1h16A	Pyruvate formate-lyase, PFL	1	759	1.53	c
2dznA	2dznE	26S proteasome non-ATPase regulatory subunit 10, gankyrin	1	227	2.2	d
1lxaA	2qiaA	UDP N-acetylglucosamine acyltransferase	1	262	1.74	b
1y0gA	1y0gD	Lipid binding protein YceI	1	169	2.2	b
2f93A	1h2sA	Sensory rhodopsin II	1	225	1.93	f
1b6tA	1qjcB	Phosphopantetheine adenylyltransferase	1	157	1.64	c
2g1aA	2b82B	Class B acid phosphatase, AphA	1	211	1.25	c
2b95A	1z09A	Dynein light chain 2A, cytoplasmic	1	96	NMR	d
1wpbA	1wpbM	Hypothetical protein YfbU	1	167	2	a
2i2tX	2qaoZ	Ribosomal protein L28 (L28p)	1	77	3.21	d
2a0wA	3bgsA	Purine nucleoside phosphorylase, PNP	1	286	2.1	c
1keyA	1j1jD	Translin	1	217	2.2	a
1gs5A	1gs5A	N-acetyl-l-glutamate kinase	1	258	1.5	c
1eumA	1eumF	Non-hem ferritin	1	161	2.05	a
1ynfA	1ynhD	Succinylarginine dihydrolase	1	439	1.95	d
1fuoA	1furA	Fumarase	1	456	1.95	a
1b09A	1b09E	C-reactive protein (CRP)	1	206	2.5	b
2b2kA	1hztA	Isopentenyl diphosphate isomerase	1	153	1.45	d
2esfA	1i6pA	beta-carbonic anhydrase	1	214	2	c
1bxiA	2vlqA	ImmE9 protein (Im9)	1	84	1.6	a
1uhnA	2zfdA	Calcineurin B-like protein 2	1	183	1.2	a
2ca6A	2ca6B	Rna1p (RanGAP1), N-terminal domain	1	344	2.2	c
1dj7B	1dj7B	Ferredoxin thioredoxin reductase (FTR), alpha (variable) chain	1	73	1.6	b
2pcdA	3pcgF	Protocatechuate-3,4-dioxygenase, alpha chain	1	200	1.96	b
1vetA	1vetA	MEK binding partner 1, MP1	1	122	1.9	d
1go3F	1go3N	RNA polymerase II subunit RBP4 (RpoF)	1	106	1.75	a
1n4pB	1n4qL	Protein farnesyltransferase, beta-subunit	1	346	2.4	a
1ek5A	1ek6B	Uridine diphosphogalactose-4-epimerase (UDP-galactose 4-epimerase)	1	345	1.5	c
2bkqA	2bkrA	Sentrin-specific protease 8, SENP8	1	212	1.9	d
1bwpA	1es9A	Platelet-activating factor acetylhydrolase	1	212	1.3	c
2bkuA	1wa5A	Ran	10	170	2	c
1a4rA	2ngrA	CDC42	11	191	1.9	c
2av1B	1k5nB	beta2-microglobulin	12	99	1.09	b
1il6A	1il6A	Interleukin-6	2	166	NMR	a
1eg9B	1o7nB	Naphthalene 1,2-dioxygenase beta subunit	2	193	1.4	d

1i2mB	1a12C	Regulator of chromosome condensation RCC1	2	401	1.7	b
1ar0A	1gy6B	Nuclear transport factor-2 (NTF2)	2	123	1.6	d
3eipA	3eipB	Colicin E3 immunity protein	2	84	1.8	d
1xrsA	1xrsA	D-lysine 5,6-aminomutase alpha subunit, KamD	2	516	2.8	c
1azzC	1ifgA	Ecotin, trypsin inhibitor	2	140	2	b
1dj1A	2eutA	Cytochrome c peroxidase, CCP	2	291	1.12	a
2incB	2incB	Toluene, o-xylene monooxygenase oxygenase subunit TouE	2	322	1.85	a
1nbwB	1nbwD	Glycerol dehydratase reactivase, beta subunit	2	113	2.4	c
1ni4A	2ozlC	E1-beta subunit of pyruvate dehydrogenase (PP module)	2	361	1.9	c
1b4uB	1bouD	LigB subunit of an aromatic-ring-opening dioxygenase LigAB	2	298	2.2	c
1wmiA	1wmiC	Hypothetical protein PHS013	2	88	2.3	d
3b6fD	1kx5H	Histone H2B	2	121	1.94	a
1c48A	1dm0K	Verotoxin-1/shiga-toxin, B-pentamer	2	69	2.5	b
1fl7B	1xwdE	Follicle stimulating hormone, follitropin, beta chain	2	105	2.92	g
1jwhC	1rqfJ	Casein kinase II beta subunit	2	173	2.89	g
3cirD	1kf6P	Fumarate reductase subunit FrdD	2	119	2.7	f
2hyeD	1ldjB	RIGG-box protein 1 (RBX1) of SCF ubiquitin ligase complex	2	88	3	g
2avyG	2qanG	Ribosomal protein S7	2	152	3.21	a
2avyI	2qanI	Ribosomal protein S9	2	127	3.21	d
2avyQ	2qanQ	Ribosomal protein S17	2	81	3.21	b
2avyS	2qanS	Ribosomal protein S19	2	80	3.21	d
1jr5A	1jr5B	Anti-sigma factor AsiA	2	90	NMR	a
1qo0A	1peaA	Amide receptor/negative regulator of the amidase operon (AmiC)	2	368	2.1	c
3biiD	1fm0D	Molybdopterin synthase subunit MoaD	2	81	1.45	d
3biiE	1nvjB	Molybdopterin synthase subunit MoaE	2	126	2.15	d
3b9uA	1rg8B	Acidic FGF (FGF1)	2	141	1.1	b
1djrD	1djrH	Heat-labile toxin	2	103	1.3	b
1jyoA	1jyoB	Virulence effector SptP secretion chaperone SicP	2	130	1.9	d
1f3cA	1cmiA	Dynein light chain 1 (DLC1)	2	85	2.5	d
1c54A	1lniB	RNase Sa	2	96	1	d
1a0uA	2dn2C	Hemoglobin, alpha-chain	2	141	1.25	a
1aonO	1svtU	Chaperonin-10 (GroES)	2	97	2.81	b
1is7K	1jg5E	GTP cyclohydrolase I feedback regulatory protein, GFRP	2	83	2.6	d
1b6bA	1kuxA	Serotonin N-acetyltranferase	2	166	1.8	d
1cb7B	1ccwD	Glutamate mutase, large subunit	2	483	1.6	c
2d1pB	2d1pB	tRNA 2-thiouridine synthesizing protein C, TusC	2	119	2.15	c
2avuA	1g8eA	Flagellar transcriptional activator FlhD	2	98	1.8	a
1fq1A	1fpzB	Kinase associated phosphatase (kap)	2	176	2	c

1kqfC	1kqfC	Formate dehydrogenase N, cytochrome (gamma) subunit	2	216	1.6	f
1dvcA	1cyvA	Cystatin A (stefin A)	2	98	NMR	d
2f0yB	2h6fB	Protein farnesyltransferase, beta-subunit	2	410	1.5	a
1egmB	1eexE	Diol dehydratase, beta subunit	2	178	1.7	c
2b9aA	1f86B	Transthyretin (synonym: prealbumin)	2	115	1.1	b
2awb3	2qao3	Ribosomal protein L35p	2	64	3.21	d
1go4A	2vfxE	MITOTIC SPINDLE ASSEMBLY CHECKPOINT PROTEIN MAD2A	2	205	1.95	d
1mvfD	1ub4C	MazE	2	75	1.7	b
1e9jA	1xwdD	Glycoprotein hormones alpha chain (Gonadotropin A, Follitropin alpha)	3	87	2.92	g
2aczD	1nekD	Succinate dehydrogenase subunit SdhD	3	113	2.6	f
2aczC	1nekC	Succinate dehydrogenase subunit SdhC	3	129	2.6	f
2incA	2incA	Toluene, o-xylene monooxygenase oxygenase subunit TouA	3	491	1.85	a
1l1oA	2pi2G	Replication protein A 14 KDa (RPA14) subunit	3	118	2	b
3b6fA	1kx5E	Histone H3	3	135	1.94	a
1f3hA	2qfaA	Anti-apoptotic protein survivin	3	137	1.4	g
1bksB	1qopB	Tryptophan synthase, beta-subunit	3	390	1.4	c
1mkfA	1mkfB	Viral chemokine binding protein m3	3	371	2.1	b
3cirC	1kf6O	Fumarate reductase subunit FrdC	3	130	2.7	f
2avyH	1s03H	Ribosomal protein S8	3	127	2.7	d
2avyK	2qanK	Ribosomal protein S11	3	117	3.21	c
2avyJ	2qanJ	Ribosomal protein S10	3	98	3.21	d
1b2uF	1ay7B	Barstar (barnase inhibitor)	3	89	1.7	c
1l0oA	1thnC	Anti-sigma factor spoIIab	3	136	2.5	d
1b3uA	1b3uB	Constant regulatory domain of protein phosphatase 2a, pr65alpha	3	588	2.3	a
3bdm0	1z7qN	Proteasome beta subunit (catalytic)	3	233	3.22	d
2vglS	2vglS	AP1 clathrin adaptor core	3	142	2.59	i
1a4oA	2o02B	zeta isoform	3	230	1.5	a
2b63J	1twfJ	RNA polymerase subunit RPB10	3	65	2.3	a
2bbkH	2bbkJ	Methylamine dehydrogenase, H-chain	3	355	1.75	b
1fntF	1rypT	Proteasome alpha subunit (non-catalytic)	4	233	1.9	d
3bdmA	1rypB	Proteasome alpha subunit (non-catalytic)	4	250	1.9	d
3bdmG	1rypO	Proteasome alpha subunit (non-catalytic)	4	243	1.9	d
3bdmJ	1z7qY	Proteasome beta subunit (catalytic)	4	198	3.22	d
3bdmI	1z7qX	Proteasome beta subunit (catalytic)	4	204	3.22	d
3bdmL	1z7qM	Proteasome beta subunit (catalytic)	4	222	3.22	d
1jb0D	1jb0D	Photosystem I subunit PsaD	4	138	2.5	d
2b90A	2d48A	Interleukin-4	4	129	1.65	a
2b63K	1twfK	RPB11	4	114	2.3	d
2ckhB	1wm3A	SUMO-2	4	72	1.2	d
2bkrB	1nddD	Nedd8	4	73	1.6	d
3b6fC	1tzyE	Histone H2A	4	104	1.9	a

1id3B	2hueC	Histone H4	4	82	1.7	a
2dyrG	2occT	Mitochondrial cytochrome c oxidase subunit VIa	4	84	2.3	f
1f3gA	1f3zA	Glucose-specific factor III (glsIII)	4	150	1.98	b
2a06J	2a06W	Subunit X (non-heme 7 kDa protein) of cytochrome bc1 complex (Ubiquinol-cytochrome c reductase)	4	62	2.1	f
2h24A	2ilkA	Interleukin-10 (cytokine synthesis inhibitory factor, CSIF)	4	155	1.6	a
2dyrH	1v54U	Cytochrome c oxidase subunit h	5	79	1.8	a
2hyiA	1oo0A	Mago nashi protein	5	144	1.85	d
2dyrI	1v54V	Mitochondrial cytochrome c oxidase subunit VIc	5	72	1.8	f
2a06F	2a06S	14 kDa protein of cytochrome bc1 complex (Ubiquinol-cytochrome c reductase)	5	99	2.1	f
1b6cA	2ppnA	FK-506 binding protein (FKBP12), an immunophilin	5	107	0.92	d
3ezaB	1opdA	Histidine-containing phosphocarrier protein (HPr)	6	85	1.5	d
1azfA	2vb1A	Lysozyme	6	129	0.65	d
2dyrF	1v54S	Cytochrome c oxidase Subunit F	6	98	1.8	g
2a06G	2a06T	Ubiquinone-binding protein QP-C of cytochrome bc1 complex (Ubiquinol-cytochrome c reductase)	6	76	2.1	f
1b9xA	1gotB	beta1-subunit of the signal-transducing G protein heterotrimer	6	339	2	b
3cx5H	3cx5S	Ubiquinone-binding protein QP-C of cytochrome bc1 complex (Ubiquinol-cytochrome c reductase)	7	93	1.9	f

[a]PDB ID + chain ID of the Family Representative (FR).
[b]PDB ID + chain ID of the Structural Representative (SR).
[c]Number of structural partners.
[d]Number of residues of the SR.
[e]X-ray resolution in Å (or 'NMR' for NMR structures).

Table S2. List of the 12 proteins in the S^{MD} dataset.

FR[a]	SR[b]	Protein	np_S[c]	np[d]	SCOP class
1orjA	1orjD	Flagellar export chaperone FliS	1	1	α
1v74B	1v74B	Colicin D immunity protein	1	1	α
2d5rA	2d5rA	CCR4-NOT transcription complex subunit 7, CAF1	1	1	α/β
3c4oB	2g2uB	Beta-lactamase-inhibitor protein, BLIP	1	1	α+β
1euiC	1ugiH	Uracil-DNA glycosylase inhibitor protein	1	1	α+β
3cptB	3cptB	Mitogen-activated protein-binding protein-interacting protein	1	1	α+β
2b90A	2d48A	Interleukin-4	4	4	α
1fntF	1rypT	Proteasome alpha subunit (non-catalytic)	4	31	α+β
2ckhB	1wm3A	SUMO-2	4	58	α+β
2hyiA	1oo0A	Mago nashi protein	5	16	α+β
2dyrH	1v54U	Cytochrome c oxidase subunit h	5	7	α
2bkuA	1wa5A	Ran	10	22	α/β

[a]PDB ID + chain ID of the Family Representative (FR).
[b]PDB ID + chain ID of the Structural Representative (SR).
[c]Number of structural partners.
[d]Number of partners (Intact + PiSite).

Table S3. List of the 69 sociable proteins in the S^{Soc} dataset.

FR[a]	Protein Name	np_s[b]	nres[c]	resol[d]	SCOP[e] class
3bhvD	CYCLIN-A2	3	262	2.1	a, a
2pvoA	FERREDOXIN-THIOREDOXIN REDUCTASE, CATALYTICCHAIN	3	110	3.4	g
2racA	PROTEIN (AMICYANIN)	3	105	1.3	b
2h9gS	TUMOR NECROSIS FACTOR RECEPTOR SUPERFAMILYMEMBER 10B PRECURSOR	3	64	2.32	g, g, g
1z5yD	THIOL:DISULFIDE INTERCHANGE PROTEIN DSBD	3	118	1.94	b
1wo2A	ALPHA-AMYLASE, PANCREATIC	3	495	2.01	b, c
1kmcD	X-LINKED INHIBITOR OF APOPTOSIS PROTEIN	3	17	2.9	g
2izvB	TRANSCRIPTION ELONGATION FACTOR B POLYPEPTIDE 2	3	105	2.55	d
3bvdB	CYTOCHROME C OXIDASE SUBUNIT 2	3	166	3.37	b, f
2ntiG	DNA POLYMERASE SLIDING CLAMP B	3	249	2.5	NA
1lqbC	VON HIPPEL-LINDAU DISEASE TUMOR SUPRESSOR	3	150	2	b
2pcxA	CELLULAR TUMOR ANTIGEN P53	4	205	1.54	b
2yccA	CYTOCHROME C	4	108	1.9	a
2jixE	ERYTHROPOIETIN RECEPTOR	4	216	3.2	b, b
2iadA	MHC CLASS II I-AD	4	187	2.4	b, d
1u81A	ADP-RIBOSYLATION FACTOR 1	4	164	NMR	c
5sicE	SUBTILISIN BPN'	4	275	2.2	c
3bpnA	INTERLEUKIN-4	4	126	3.02	a
2z6kD	REPLICATION PROTEIN A 14 KDA SUBUNIT	4	117	3	b
2vpfH	VASCULAR ENDOTHELIAL GROWTH FACTOR	4	94	1.93	g
2o93O	ACTOR OF ACTIVATED T-CELLS, CYTOPLASMIC 2	4	287	3.05	b, b
2nz1Y	SMALL INDUCIBLE CYTOKINE A2	4	64	2.5	d
2nljC	VOLTAGE-GATED POTASSIUM CHANNEL	4	103	2.52	f
1uexC	VON WILLEBRAND FACTOR	4	202	2.85	c
1u0nD	PLATELET GLYCOPROTEIN IB	4	265	2.95	c
1dzfA	DNA-DIRECTED RNA POLYMERASES I, II, AND III 27KD POLYPEPTIDE	4	211	1.9	c, d
5cyhA	CYCLOPHILIN A	4	164	2.1	b
2p6bC	CALMODULIN-DEPENDENT CALCINEURIN A SUBUNIT ALPHAISOFORM	4	357	2.3	d
2f3gB	GLUCOSE-SPECIFIC PHOSPHOCARRIER	4	150	2.13	b
2c1tB	IMPORTIN ALPHA SUBUNIT	4	423	2.6	a
2angA	ANGIOGENIN	4	123	2	d
2iy0C	RAN GTPASE-ACTIVATING PROTEIN 1	4	156	2.77	a
2io5B	HISTONE H3.1	4	76	2.7	a
2ilkA	INTERLEUKIN-10	4	155	1.6	a
2z8wB	APICAL MEMBRANE ANTIGEN 1	4	335	2.45	NA

2nu4I	OVOMUCOID	5	51	1.75	g
1jckD	STAPHYLOCOCCAL ENTEROTOXIN C3	5	239	3.5	b, d
1btgC	BETA NERVE GROWTH FACTOR	5	109	2.5	g
4nn9A	NEURAMINIDASE N9	5	388	2.3	b
3bmpA	PROTEIN (BONE MORPHOGENETIC PROTEIN 2 (BMP-2))	5	106	2.7	g
2uyzA	SUMO-CONJUGATING ENZYME UBC9	5	156	1.4	d
2trxB	THIOREDOXIN	5	108	1.68	c
2f43A	ATP SYNTHASE ALPHA CHAIN, MITOCHONDRIAL	5	480	3	a, b, c
4fapA	FK506-BINDING PROTEIN	5	107	2.8	d
2ghwA	SPIKE GLYCOPROTEIN	5	191	2.3	d
2pxrC	GAG-POL POLYPROTEIN (PR160GAG-POL)	5	145	1.5	a
1ytfA	PROTEIN (TATA BINDING PROTEIN (TBP))	5	180	2.5	d, d
1qfwA	GONADOTROPIN ALPHA SUBUNIT	5	87	3.5	g
1urqB	SYNTAXIN 1A	6	63	2	h
4mbpA	MALTODEXTRIN BINDING PROTEIN	6	370	1.7	c
8cpaA	CARBOXYPEPTIDASE A	6	307	2	c
3ezeB	PROTEIN (PHOSPHOTRANSFERASE SYSTEM, HPR)	6	85	NMR	d
2trcB	TRANSDUCIN	6	340	2.4	b
2atpC	T-CELL SURFACE GLYCOPROTEIN CD8 ALPHA CHAIN	6	119	2.4	b
2hueC	HISTONE H4	6	82	1.7	a
2e76A	CYTOCHROME B6	6	215	3.41	f
2uyzB	SMALL UBIQUITIN-RELATED MODIFIER 1	7	78	1.4	d
2aq3G	T-CELL RECEPTOR BETA CHAIN V	7	109	2.3	b
3hlaB	BETA 2-MICROGLOBULIN	8	99	2.6	b
9ptiA	BOVINE PANCREATIC TRYPSIN INHIBITOR	8	58	1.22	g
2iffY	HEN EGG WHITE LYSOZYME	8	129	2.65	d
4clnA	CALMODULIN	8	148	2.2	a
2hn7A	HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, A-11ALPHA CHAIN	9	274	1.6	b, d
2ql1A	IGHM PROTEIN	10	209	2.53	b, b
3cjcA	ACTIN, ALPHA SKELETAL MUSCLE	11	372	3.9	c, c
2vrwA	RAS-RELATED C3 BOTULINUM TOXIN SUBSTRATE 1	13	177	1.85	c
1mz4A	CYTOCHROME C550	13	131	1.8	a
1za3L	FAB-YSD1 LIGHT CHAIN	16	213	3.35	b, b
2gbrC	UBIQUITIN	27	76	2	d

[a]PDB ID + chain ID of the Family Representative (FR).
[b]Number of structural partners
[c]Number of residues of the SR.
[d]X-ray resolution in Å (or 'NMR' for NMR structures).
[e]SCOP class. A multiple entry is reported for multi-domain FRs. In two cases, no SCOP annotation was found ('NA').

Table S4. Distribution of residues and interfaces over the three binding classes c_{mono}, $c_{mono\_in\_multi}$ and c_{multi} for the S^{Full}, S^{MD} and S^{Soc} datasets.

	Residues			Interfaces	
	S^{Full}	S^{MD}	S^{Soc}	S^{Full}	S^{MD}
c_{mono}	6821	165	-	302	7
$c_{mono\_in\_multi}$	4540	297	2780	71	3
c_{multi}	1261	146	1910	322	34
Tot[a]	12622	608	4690	695	44

[a] Total number of interface residues and interfaces for all the proteins in each dataset.

Table S5. Criteria used for the different hot spot prediction methods.

Method	Hot spot prediction criteria
ANCHOR	Binding energy[a] <= 2.0 kcal/mol
Robetta	$\Delta\Delta G_{bind}$[b] >= 1 kcal/mol
PISA	ΔG_{int}[c] <= -1 kcal/mol
HotPoint	Relative accessibility in the complex <= 20 % and contact potential[d] >= 18
KFC2a	Confidence of KFC2a prediction > 0
KFC2b	Confidence of KFC2b prediction > 0

[a] Residue contribution to the binding energy as evaluated from the FastContact potential (Meireles et al. (2010) Nucleic Acids Res 38: W407).
[b] Change in the binding free energy when the residue is replaced by alanine.
[c] Residue contribution to the binding energy as given by $\Delta G_{solv}+\Delta G_{cont}$ (Krissinel et al. (2007) J Mol Biol 372: 774).
[d] Tuncbag et al. (2010) Nucleic Acids Res 38: W402.

Table S6. Composition of the 12 systems used for MD simulations of the S^{MD} dataset

SR[a]	n_{res}[b]	n_{wat}[c]	n_{ions}[d]	n_{atoms}[e]	b_{size}[f]
1orjD	125	16897	6 (Na^+)	52016	81.3
1v74B	87	10952	2 (Na^+)	33756	70.3
2d5rA	252	16211	14 (Na^+)	51301	81.0
2g2uB	165	13009	1 (Na^+)	40631	74.9
1ugiH	83	8183	12 (Na^+)	25358	64.0
3cptB	116	15166	1 (Na^+)	46614	77.8
2d48A	129	12492	6 (Cl^-)	38841	73.8
1rypT	233	19084	1 (Cl^-)	59555	85.1
1wm3A	72	7431	1 (Na^+)	23043	61.9
1oo0A	144	13026	4 (Na^+)	40627	74.9
1v54U	79	10305	5 (Cl^-)	31805	68.9
1wa5A	170	12914	10 (Cl^-)	40603	75.0

[a]PDB ID + chain ID of the Structural Representative (SR).
[b]Number of protein residues
[c]Number of water molecules
[d]Number of counterions (the ion type is indicated in parentheses).
[e]Total number of atoms in the system
[f]Side length (Å) of the cubic box after equilibration.

Table S7. Analysis of pairwise RMSD distributions of tCONCOORD and MD ensembles.

	tCONCOORD			MD		
SR[a]	**pairRMSDave[b]**	**pairRMSDmin[c]**	**pairRMSDmax[d]**	**pairRMSDave[b]**	**pairRMSDmin[c]**	**pairRMSDmax[d]**
1orjD	1.79	0.44	5.50	1.12	0.50	2.19
1v74B	2.09	0.49	4.42	1.65	0.59	3.28
2d5rA	1.82	0.71	4.16	1.85	0.67	3.70
2g2uB	3.01	1.00	8.31	2.19	0.64	3.99
1ugiH	1.91	0.65	4.64	2.03	0.57	4.79
3cptB	2.28	0.69	5.35	2.46	0.66	5.06
2d48A	1.72	0.63	4.17	1.59	0.54	3.20
1rypT	3.40	0.96	8.71	1.80	0.63	3.33
1wm3A	1.38	0.48	3.02	1.94	0.55	3.70
1oo0A	1.78	0.57	4.07	1.72	0.58	3.64
1v54U	2.60	0.69	6.26	3.11	0.54	5.96
1wa5A	2.23	0.86	4.30	1.52	0.62	2.68
average	**2.17**	**0.68**	**5.24**	**1.92**	**0.59**	**3.79**

[a]PDB ID + chain ID of the Structural Representative (SR).
[b]Average pairwise RMSD (Å). Pairwise RMSD values are calculated over all the possible pairs of conformations in the ensemble.
[c]Minimum pairwise RMSD (Å).
[d]Maximum pairwise RMSD (Å).

Table S8. Comparison between X/Y pairs of RMSF profiles, with X(Y) = tCON(COORD), MD, GNM or PDB.

	tCON/MD[a]	**tCON/PDB**	**tCON/GNM**	**MD/PDB**	**MD/GNM**	**GNM/PDB**
C^{α} RMSF (S^{Full})	-	0.50 (± 0.20)	0.75 (± 0.11)	-	-	0.47 (± 0.22)
C^{α} RMSF (S^{MD})	0.75 (± 0.08)	0.57 (± 0.14)	0.70 (± 0.18)	0.55 (± 0.16)	0.69 (± 0.16)	0.52 (± 0.17)
SC RMSF (S^{Full})	-	0.60 (± 0.10)	-	-	-	-
SC RMSF (S^{MD})	0.71 (± 0.06)	0.60 (± 0.12)	-	0.63 (± 0.11)	-	-

[a]Average Spearman correlation coefficient between C^{α} or side chain (SC) RMSF profiles calculated over the entire dataset (S^{Full}) or the MD dataset (S^{MD}). The standard deviation is reported in parentheses.

Table S9. Overlap between the tCONCOORD, MD and PDB essential spaces.

SR[a]	nPDB[b]	CovMatOver[c] MD/tCON	CumOver15[d] MD/tCON	MaxInpr[e] MD/tCON	MaxInpr[e] PDB/tCON	MaxInpr[e] PDB/MD
1orjD	5	0.349	0.537	0.759 (3, 2)	0.424 (2, 4)	0.267 (2, 6)
1v74B	3	0.309	0.463	0.614 (2, 1)	0.529 (1, 3)	0.389 (1, 8)
2d5rA	1	0.360	0.402	0.545 (2, 2)	-	-
2g2uB	6	0.363	0.499	0.852 (3, 1)	0.478 (3, 2)	0.404 (3, 10)
1ugiH	30	0.421	0.542	0.603 (3, 1)	0.499 (1, 2)	0.387 (1, 7)
3cptB	6	0.376	0.489	0.631 (3, 2)	0.558 (2, 3)	0.445 (2, 4)
2d48A	17	0.356	0.495	0.336 (1, 5)	0.34 (2, 4)	0.392 (2, 7)
1rypT	22	0.274	0.458	0.589 (3, 1)	0.315 (1, 1)	0.297 (1, 4)
1wm3A	12	0.305	0.434	0.423 (2, 1)	0.42 (2, 13)	0.343 (1, 1)
1oo0A	11	0.332	0.418	0.578 (3, 2)	0.319 (2, 8)	0.402 (2, 4)
1v54U	28	0.280	0.472	0.424 (2, 8)	0.335 (2, 1)	0.372 (2, 14)
1wa5A	29	0.347	0.481	0.434 (3, 5)	0.414 (1, 6)	0.409 (1, 5)
	average	**0.339**	**0.474**	**0.566**	**0.421**	**0.373**

[a]PDB ID + chain ID of the Structural Representative (SR).
[b]Number of PDB structures in the ensemble.
[c]Normalized overlap of the covariance matrices.
[d]Cumulative overlap between the first 15 PCs from MD and tCONCOORD ensembles.
[e]Maximum inner product between the first 3 PCs of ensemble X and the first 15 PCs of ensemble Y. The X/Y pair is indicated in the column label. The indices of the PCs from the X and Y ensembles with the largest inner product value are indicated in parentheses. Only PDB ensembles with nPDB > 1 were considered.

Table S10. Average per-residue configurational entropy for $c_{mono\_in\_multi}$ and c_{multi} residues in S^{MD}.

	TS[b]			TS (main chain)[c]		
SR[a]	**$c_{mono\_in\_multi}$**	**c_{multi}**	**Δ(TS)**	**$c_{mono\_in\_multi}$**	**c_{multi}**	**Δ(TS)**
2d48A	12.66	14.18	1.52	5.61	6.26	0.65
1rypT	13.83	14.82	0.99	6.47	7.02	0.55
1wm3A	14.31	15.21	0.90	6.38	6.92	0.54
1oo0A	13.56	14.79	1.23	6.09	6.53	0.43
1v54U	14.78	15.46	0.68	7.25	7.15	-0.10
1wa5A	12.64	13.58	0.94	5.56	6.13	0.56
		average	**1.04**		average	**0.44**

[a]PDB ID + chain ID of the Structural Representative (SR).
[b]Per-residue entropy term TS (T = 300K) in kcal/mol averaged over $c_{mono\_in\_multi}$ and c_{multi} residues of each multi-partner protein in S^{MD}. To make entropy values of different residue types comparable, the entropy of each residue is normalised by the total number of its atoms and multiplied by 10 (average number of atoms in a residue).
[c]Per-residue entropy term TS (T = 300K) in kcal/mol calculated considering main chain atoms only and averaged over $c_{mono\_in\_multi}$ and c_{multi} residues.

Table S11. Percentage of c_{multi} residues classified as hot spots in 1 to 4 different interfaces.

Number of interfaces	ANCHOR	ROBETTA	PISA	HotPoint	KFC2a	KFC2b
1	79.4	81.6	88.4	81.6	85.6	83.8
2	16.5	15.3	9.1	14.1	11.5	12.8
3	3.2	2.4	1.4	4.1	2.1	3.4
4	0.9	0.7	1.1	0.3	0.9	0.0

Table S12. SNP and DisSNP propensity in the S^{Full} and S^{Soc} datasets (values per protein).

c_{mono}	P_{SNP}[a]	P_{DisSNP}[a]	$c_{mono\ in\ multi}$	P_{SNP}[b]	P_{DisSNP}[b]	c_{multi}	P_{SNP}[b]	P_{DisSNP}[b]
2i3sA	0.00	-	**1b3uA**	1.55	-	**1b3uA**	0.00	-
1gggA	0.00	-	**1f3hA**	2.04	-	**1f3hA**	0.52	-
1ai2A	0.31	-	**2aczC**	1.00	-	**2aczC**	0.00	-
1bwpA	2.42	-	**3b9uA**	1.01	-	**3b9uA**	0.00	-
1aiuA	0.00	-	**3b6fC**	0.58	-	**3b6fC**	0.00	-
1ajaA	0.95	-	**2bkrB**	0.00	-	**2bkrB**	1.39	-
1cddA	0.73	-	**1fntF**	0.00	-	**1fntF**	0.00	-
2av6A	1.34	1.20	**3bdmG**	1.69	-	**3bdmG**	0.00	-
1a03A	1.26	-	**3bdmL**	0.37	-	**3bdmL**	0.00	-
1psrA	1.68	-	**2bkuA**	2.29	2.29	**2bkuA**	1.18	1.18
2ca6A	3.72	-	**2b63K**	1.12	-	**2b63K**	0.00	-
1av5A	1.53	-	**2ckhB**	0.00	-	**2ckhB**	0.00	-
2dgkA	1.05	-	**1b9xA**	0.00	-	**1b9xA**	0.00	-
1n4pB	2.54	-	**1a4rA**	0.91	-	**1a4rA**	0.43	-
1bsvA	0.92	0.60	**1a0uA**	0.92	1.19	**1a0uA**	1.48	0.75
1ek5A	0.70	0.00	**1ni4A**	0.85	1.41	**1ni4A**	0.00	0.00
1d2fA	0.00	-	**1i2mB**	2.77	-	**1i2mB**	-	-
1ex6A	0.00	-	**2f0yB**	0.00	-	**2f0yB**	2.70	-
			1dvcA	1.11	-	**1dvcA**	0.00	-
			2avyH	3.81	-	**2avyH**	0.00	-

Propensities are calculated with respect to the surface of each protein.
[a] Only the mono-partner proteins of S^{Full} with a human homologue and SNP annotation were considered (see Methods).
[b] Only the multi-partner proteins of S^{Full} with a human homologue and SNP annotation were considered (see Methods).

Table S12 - continued

$c_{mono\ in\ multi(Soc)}$	P_{SNP}[c]	P_{DisSNP}[c]	$c_{multi(Soc)}$	P_{SNP}[c]	P_{DisSNP}[c]
3cjcA	0.87	1.30	**3cjcA**	0.00	0.31
2atpC	5.45	-	**2atpC**	0.00	-
2aq3G	0.99	-	**2aq3G**	1.25	-
2iadA	0.73	-	**2iadA**	1.49	-
2jixE	0.48	-	**2jixE**	0.00	-
4fapA	0.00	-	**4fapA**	2.63	-
2c1tB	0.61	-	**2c1tB**	0.00	-
2nljC	-	0.94	**2nljC**	-	1.23
1uexC	0.00	-	**1uexC**	0.00	-
2pcxA	1.31	0.99	**2pcxA**	1.14	3.42
2trxB	4.82	-	**2trxB**	2.52	-
2p6bC	2.59	-	**2p6bC**	0.00	-
1dzfA	1.22	-	**1dzfA**	0.00	-
2uyzA	1.26	-	**2uyzA**	1.60	-
2f43A	1.60	-	**2f43A**	0.00	-
2vpfH	1.01	-	**2vpfH**	1.60	-
2trcB	0.00	-	**2trcB**	0.00	-
2vrwA	0.00	-	**2vrwA**	0.00	-
2nz1Y	0.73	-	**2nz1Y**	1.21	-
2hn7A	1.02	-	**2hn7A**	1.19	-
4clnA	1.42	-	**4clnA**	0.79	-
3bmpA	0.55	-	**3bmpA**	1.56	-
2uyzB	0.00	-	**2uyzB**	0.00	-
8cpaA	1.71	-	**8cpaA**	2.00	-
9ptiA	1.70	-	**9ptiA**	0.59	-

[c] Only the multi-partner proteins of S^{Full} with a human homologue and SNP annotation were considered (see Methods).

Table S13. C^{α} and side chain (SC) RMSF Z-scores of $c_{mono\_in\_multi}$ and c_{multi} interface residues of Neddylin.

resid[a]	b[b]	DSSP[c]	tCON ensemble[d]		PDB ensemble[e]	
			C^{α} RMSF	SC RMSF	C^{α} RMSF	SC RMSF
$c_{mono\_in\_multi}$						
E28	1	H	-0.65	0.61	-0.14	-0.27
R29	1	H	-0.64	0.25	0.39	2.03
E31	1	H	-0.53	1.12	-0.15	-1.00
E32	1	H	-0.17	1.02	0.36	0.86
E34	1	H	-0.13	0.64	-0.22	0.42
G35	1	~	0.11	-	-0.22	-
I36	1	~	-0.02	-0.83	-0.16	0.04
Q39	1	G	1.51	0.67	0.34	1.08
Y45	1	E	0.02	0.61	-0.80	0.37
S46	1	T	1.64	-0.80	0.50	-
M50	1	~	-0.58	-0.33	-0.26	0.22
D52	1	T	-0.31	0.83	-0.08	0.59
E53	1	T	0.29	0.78	-0.60	0.94
Y59	1	T	-0.09	0.57	-0.29	-0.84
		average	0.03 (± 0.72)	0.40 (± 0.65)	-0.09 (± 0.38)	0.37 (± 0.84)
c_{multi}						
K6	2	E	-0.70	0.37	-0.10	-0.01
T7	2	~	0.56	-0.31	2.37	0.09
L8	**3**	**T**	**2.56**	**0.43**	**4.64**	**0.47**
T9	4	T	3.73	0.15	2.65	-1.07
G10	2	S	3.47	-	2.00	-
K11	2	~	1.98	2.88	0.53	-
Q40	2	G	0.66	-0.08	-0.68	2.12
R42	**3**	**E**	**-0.75**	**0.67**	**-0.64**	**1.49**
I44	**2**	**E**	**-0.82**	**-0.21**	**-0.67**	**-0.12**
G47	**2**	**T**	**1.78**	**-**	**0.61**	**-**
K48	**3**	**E**	**1.06**	**2.09**	**0.69**	**1.85**
Q49	**2**	**E**	**0.40**	**1.69**	**-0.47**	**0.94**
N51	**2**	**~**	**-0.33**	**-0.48**	**-0.34**	**-0.24**
K54	2	S	0.02	0.51	-0.55	-0.14
H68	**2**	**E**	**-1.02**	**0.17**	**-0.50**	**-0.77**
V70	**2**	**E**	**-0.22**	**0.18**	**0.77**	**0.49**
L71	**2**	**E**	**1.07**	**0.91**	**4.03**	**0.14**
		average	0.79 (± 1.48)	0.60 (± 0.94)	0.84 (± 1.70)	0.37 (± 0.94)

[a] Residues identified as hot spots by at least two prediction methods are reported in bold.
[b] Binding multiplicity.
[c] DSSP secondary structure (H = α helix, G = 3_{10} helix, E = β strand, T = turn, S = bend, ~ = irregular).
[d] RMSF Z-scores from tCONCOORD ensemble. Standard deviations are reported in parentheses.
[e] RMSF Z-scores from PDB ensemble. Standard deviations are reported in parentheses.

Table S14. C^{α} and side chain (SC) RMSF Z-scores of $c_{mono\_in\_multi}$ and c_{multi} interface residues of SUMO-2.

resid[a]	b[b]	DSSP[c]	tCON ensemble[d]		PDB ensemble[e]	
			C^{α} RMSF	SC RMSF	C^{α} RMSF	SC RMSF
$c_{mono\_in\_multi}$						
N3	1	E	-0.28	-0.59	-	-
K5	1	E	-0.78	0.36	-1.20	-0.15
A7	1	E	-0.76	-1.17	-1.23	-1.64
V14	**1**	**E**	**-0.68**	**-0.96**	**-0.91**	**-1.43**
K19	1	~	-0.26	0.74	-1.03	1.33
T22	1	S	-0.16	-1.13	-0.94	-0.61
L27	1	H	-0.98	-0.92	-0.44	-0.87
K29	1	H	-0.59	0.58	-0.05	0.73
Y31	**1**	**H**	**-0.54**	**0.71**	**-0.09**	**1.18**
C32	1	H	-0.33	-0.80	0.68	-
L37	1	~	0.44	-0.43	0.45	-0.38
S38	1	~	0.62	-1.08	0.94	-0.04
M39	1	T	0.36	-0.31	1.26	0.74
R43	**1**	**E**	**-0.32**	**1.01**	**-0.65**	**1.92**
R45	**1**	**E**	**-0.77**	**0.87**	**-0.73**	**0.96**
F46	1	E	-0.55	-0.07	-0.28	-0.90
D47	**1**	**T**	**0.78**	**1.76**	**0.17**	**0.13**
G48	1	T	0.98	-	2.04	-
Q49	1	E	0.43	0.04	0.51	0.91
P50	1	E	0.22	-1.04	0.60	-1.56
N52	1	~	-0.14	0.43	0.53	-0.21
D55	1	~	-0.29	-0.18	-0.55	-0.46
L60	1	T	-0.31	-0.80	2.21	0.09
D66	1	~	-0.38	0.33	-1.56	-0.93
D69	**1**	**E**	**-0.94**	**-0.16**	**-0.92**	**-0.07**
		average	-0.21 (± 0.55)	-0.12 (± 0.81)	-0.05 (± 1.01)	-0.06 (± 0.98)
c_{multi}						
Q9	2	T	2.83	2.35	-0.56	1.18
D10	**3**	**T**	**4.24**	**0.17**	**1.23**	**0.29**
G11	3	S	3.49	-	2.76	-
S12	2	~	1.72	-0.17	-0.60	-0.62
V13	**2**	**E**	**0.15**	**-0.13**	**-0.66**	**-0.52**
Q15	**2**	**E**	**-0.66**	**0.03**	**-0.67**	**0.92**
F16	**2**	**E**	**-0.78**	**0.16**	**-0.31**	**-0.55**
K17	**2**	**E**	**-0.49**	**1.06**	**-0.35**	**0.87**
I18	**2**	**E**	**-0.59**	**-0.33**	**-0.83**	**0.30**
K26	2	H	-0.67	0.84	-0.37	2.07
A30	2	H	-0.48	-1.36	-0.19	-1.62
E33	2	H	0.15	0.87	0.47	0.58
R34	**2**	**H**	**0.42**	**1.95**	**-0.07**	**2.63**

Q35	**2**	**H**	**0.46**	**0.01**	**0.41**	**1.32**
G36	2	T	0.70	-	0.08	-
R40	**2**	**T**	**1.73**	**1.75**	**0.84**	**0.93**
		average	0.76 (± 1.59)	0.51 (± 1.01)	0.07 (± 0.92)	0.56 (± 1.13)

[a] Residues identified as hot spots by at least two prediction methods are reported in bold.
[b] Binding multiplicity.
[c] DSSP secondary structure (H = α helix, E = β strand, T = turn, S = bend, ~ = irregular).
[d] RMSF Z-scores from tCONCOORD ensemble. Standard deviations are reported in parentheses.
[e] RMSF Z-scores from PDB ensemble. Standard deviations are reported in parentheses.