Supporting Information

Table S2. The 110 descriptors used to characterize protein–protein interface

No.	Descriptor	Type	Software
1	The number of residues	Constitutional descriptor	ProtorP [1]
2	The number of polar residues	Constitutional descriptor	ProtorP [1]
3	The number of nonpolar residues	Constitutional descriptor	ProtorP [1]
4	The number of neutral residues	Constitutional descriptor	ProtorP [1]
5	The number of positively charged residues	Constitutional descriptor	ProtorP [1]
6	The number of negatively charged residues	Constitutional descriptor	ProtorP [1]
7	The number of aromatic residues	Constitutional descriptor	ProtinDB [2]
8	The number of hotspot residues	Constitutional descriptor	KFC2 [3]
9	The number of critical residues	Constitutional descriptor	PCRPi-W [4]
10	The percentage of buried residues	Constitutional descriptor	ProScale [5]
11	The percentage of accessible residues	Constitutional descriptor	ProScale [5]
12	The number of hydroxyl groups	Constitutional descriptor	PSAIA [6]
13	The number of amides	Constitutional descriptor	PSAIA [6]
14	The number of carboxyl groups	Constitutional descriptor	PSAIA [6]
15	The number of ammonium salts	Constitutional descriptor	PSAIA [6]
16	The number of sulfhydryl groups	Constitutional descriptor	PSAIA [6]
17	The number of benzene rings	Constitutional descriptor	PSAIA [6]
18	The number of atoms	Constitutional descriptor	ProtorP [1]
19	The number of heavy atoms	Constitutional descriptor	ProFace [7]
20	The number of hydrogen atoms	Constitutional descriptor	ProFace [7]
21	The number of carbon atoms	Constitutional descriptor	ProFace [7]
22	The number of oxygen atoms	Constitutional descriptor	ProFace [7]
23	The number of nitrogen atoms	Constitutional descriptor	ProFace [7]
24	The number of sulfur atoms	Constitutional descriptor	ProFace [7]
25	The relative quantity of α-helices	Constitutional descriptor	DSSP [8]
26	The relative quantity of β-strands	Constitutional descriptor	DSSP [8]
27	The relative quantity of coils	Constitutional descriptor	DSSP [8]
28	Number of pocket	Constitutional descriptor	Q-SiteFinder [9]
29	Sum of atomic contact vectors	Contacting descriptor	NOXclass [10]
30	Atomic contact energy	Contacting descriptor	FastContact [11]
31	Residue contact propensity	Contacting descriptor	FastContact [11]
32	Pairwise interaction	Contacting descriptor	HORI [12]
33	Triplet interaction	Contacting descriptor	HORI [12]
34	Quadruple interaction	Contacting descriptor	HORI [12]
35	The number of residue-pairs in contact	Contacting descriptor	CMAPpro [13]
36	The number of atom-pairs in contact	Contacting descriptor	SPACE [14]
37	Residue match index	Contacting descriptor	HotPatch [15]
38	Total empirical contact potential	Contacting descriptor	PISA [16]
39	Residue interaction index	Contacting descriptor	PISA [16]
40	Core steric contact	Contacting descriptor	Probe [17]
41	Interface packing rate	Contacting descriptor	Probe [17]
42	Complementarity Index	Contacting descriptor	Probe [17]
43	Hydrophobic packing index	Contacting descriptor	PSAIA [6]
44	Total molecular weight	Geometrical descriptor	ProScale [5]
45	Average buried area	Geometrical descriptor	MSMS [18]
46	Average exposed area	Geometrical descriptor	MSMS [18]
47	Relative size of interface to surface	Geometrical descriptor	MSMS [18]
48	Accessible surface area of interface	Geometrical descriptor	MSMS [18]
49	Connectivity index	Geometrical descriptor	EMBnet [19]
50	Connectivity correlation	Geometrical descriptor	EMBnet [19]
51	Planarity	Geometrical descriptor	SHARP2 [20]
52	Eccentricity	Geometrical descriptor	ProtorP [1]

53	Length	Geometrical descriptor	ProtorP [1]
54	Breadth	Geometrical descriptor	ProtorP [1]
55	Gap volume	Geometrical descriptor	ProtorP [1]
56	Gap volume index	Geometrical descriptor	ProtorP [1]
57	Interface accessibility	Geometrical descriptor	ProtorP [1]
58	Voronoi volume	Geometrical descriptor	DiMoVo [21]
59	van der Waals volume	Geometrical descriptor	VEGA [22]
60	Hotspot size	Geometrical descriptor	KFC2 [3]
61	Total pocket size	Geometrical descriptor	Q-SiteFinder [9]
62	The degree of curvature	Geometrical descriptor	CASTp [23]
63	Average diameter	Geometrical descriptor	Plasmod-PPI [24]
64	Interface roughness	Geometrical descriptor	ProtFract [25]
65	Interface topology	Geometrical descriptor	CASTp [23]
66	Bulkness	Geometrical descriptor	ProScale [5]
67	Average volume of buried residue	Geometrical descriptor	DEPTH [26]
68	Electrostatic potential	Physicochemical descriptor	DelPhi [27]
69	Interfacial polarity	Physicochemical descriptor	ProtParam [28]
70	Molar refractivity	Physicochemical descriptor	VEGA [22]
71	Average hydrophobicity	Physicochemical descriptor	2D-GraLab [29]
72	Average hydrophilicity	Physicochemical descriptor	2D-GraLab [29]
73	Net charge	Physicochemical descriptor	ProtParam [28]
74	Positive net charge	Physicochemical descriptor	ProtParam [28]
75	Negative net charge	Physicochemical descriptor	ProtParam [28]
76	Dipolar moment	Physicochemical descriptor	ProtParam [28]
77	Average solubility	Physicochemical descriptor	ProtParam [28]
78	Amphipathicity	Physicochemical descriptor	ProtParam [28]
79	Total polarity	Physicochemical descriptor	ProScale [5]
80	Degree of atomic buriedness	Physicochemical descriptor	DEPTH [26]
81	Average flexibility	Physicochemical descriptor	ProScale [5]
82	Interfacial pI	Physicochemical descriptor	ProtParam [28]
83	Instability index	Physicochemical descriptor	ProtParam [28]
84	Aliphatic index	Physicochemical descriptor	ProtParam [28]
85	Relative mutability	Physicochemical descriptor	ProScale [5]
86	Membrane tendency	Physicochemical descriptor	ProScale [5]
87	Refractivity	Physicochemical descriptor	ProScale [5]
88	Hydrophobic moment	Physicochemical descriptor	ProtParam [28]
89	Extinction coefficient	Physicochemical descriptor	ProtParam [28]
90	Average residue depth	Physicochemical descriptor	DEPTH [26]
91	Average z-scale, vector 1	Physicochemical descriptor	PROFEAT [31]
92	Average z-scale, vector 2	Physicochemical descriptor	PROFEAT [31]
93	Average z-scale, vector 3	Physicochemical descriptor	PROFEAT [31]
94	The number of hydrogen bonds	Nonbonded descriptor	HBplus [32]
95	The number of water-mediated hydrogen bonds	Nonbonded descriptor	HBplus [32]
96	Hydrogen bond potential	Nonbonded descriptor	2D-GraLab [29]
97	The number of salt bridges	Nonbonded descriptor	2D-GraLab [29]
98	Salt bridge energy	Nonbonded descriptor	2D-GraLab [29]
99	The number of π-π stacking	Nonbonded descriptor	2D-GraLab [29]
100	Desolvation energy	Nonbonded descriptor	2D-GraLab [29]
101	Conformational entropic penalty	Nonbonded descriptor	2D-GraLab [29]
102	Hydrophobic packing	Nonbonded descriptor	2D-GraLab [29]
103	The electrostatic potential of ion pairs	Nonbonded descriptor	2D-GraLab [29]
104	The number of steric interactions	Nonbonded descriptor	Probe [17]
105	The number of aromatic-sulphur interactions	Nonbonded descriptor	PIC [32]
106	The number of cation-π interactions	Nonbonded descriptor	PIC [32]
107	Long-range electrostatic potential	Nonbonded descriptor	DelPhi [27]
108	Polar solvation contribution	Nonbonded descriptor	DelPhi [27]
109	Nonpolar solvation contribution	Nonbonded descriptor	2D-GraLab [29]

References

[1] C. Reynolds, D. Damerell, S. Jones, ProtorP: a protein–protein interaction analysis server. *Bioinformatics* 2009, 25, 413–414.

[2] R. A Jordan, Y. EL-Manzalawy, D. Dobbs, V. Honavar, Predicting protein–protein interface residues using local surface structural similarity. *BMC Bioinformatics* 2012, 13, 41.

[3] X. Zhu, J. C. Mitchell, KFC2: a knowledge-based hot spot prediction method based on interface solvation, atomic density, and plasticity features. *Proteins* 2011, 79, 2671–2683.

[4] J. S. Mora, S. A. Assi, N. Fernandez-Fuentes, Presaging critical residues in protein interfaces-web server (PCRPi-W): a web server to chart hot spots in protein interfaces. *PLoS ONE* 2010, 5, e12352.

[5] E. Gasteiger, C. Hoogland, A. Gattiker, S. Duvaud, M. R. Wilkins, R. D. Appel, A. Bairoch, *Protein Identification and Analysis Tools on the ExPASy Server*; (In) John M. Walker (ed): The Proteomics Protocols Handbook, Humana Press, 2005, pp.571–607.

[6] J. Mihel, M. Šikić, S. Tomić, B. Jeren, K. Vlahoviček, PSAIA — protein structure and interaction analyzer. *BMC Struct. Biol.* 2008, 8, 21.

[7] R.P. Saha, R.P. Bahadur, A. Pal, S. Mandal, P. Chakrabarti, ProFace: a server for the analysis of the physicochemical features of protein–protein interfaces. *BMC Struct. Biol.* 2006, 6, 11.

[8] W. Kabsch, C. Sander, Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* 1983, 22, 2577–2637.

[9] A. T. Laurie, R. M. Jackson, Q-SiteFinder: an energy-based method for the prediction of protein-ligand binding sites. Bioinformatics 2005, 21, 1908–1916.

[10] H. Zhu, F. S. Domingues, I. Sommer, T. Lengauer, NOXclass: prediction of protein–protein interaction types. *BMC Bioinformatics* 2006, 7, 27.

[11] P. C. Champ, C. J. Camacho, FastContact: a free energy scoring tool for protein–protein complex structures. *Nucleic Acids Res.* 2007, 35, W556–W560.

[12] P. Sundaramurthy, K. Shameer, R. Sreenivasan, S. Gakkhar, R. Sowdhamini, HORI: a web server to compute higher order residue interactions in protein structures. *BMC Bioinformatics* 2010, 11, S24.

[13] P. Di Lena, K. Nagata, P. Baldi, Deep architectures for protein contact map prediction. *Bioinformatics* 2012, 28, 2449–2457.

[14] V. Sobolev, E. Eyal, S. Gerzon, V. Potapov, M. Babor, J. Prilusky, M. Edelman, SPACE: a suite of tools for protein structure prediction and analysis based on complementarity and environment. *Nucleic Acids Res.* 2005, 33, W39–W43.

[15] F. K. Pettit, E. Bare, A. Tsai, J. U. Bowie, HotPatch: a statistical approach to finding biologically relevant features on protein surfaces. *J. Mol. Biol.* 2007, 369, 863–879.

[16] E. Krissinel, K. Henrick, Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* 2007, 372, 774–797.

[17] J. M. Word, S. C. Lovell, T. H. LaBean, H. C. Taylor, M. E. Zalis, B. K. Presley, J. S. Richardson, D. C. Richardson, Visualizing and quantifying molecular goodness-of-fit: small-probe contact dots with explicit hydrogen atoms. *J. Mol. Biol.* 1999, 285, 1711–1733.

[18] M. F. Sanner, A. J. Olson, J. C. Spehner, Reduced surface: an efficient way to compute molecular surfaces. *Biopolymers* 1996, 38, 305–320.

[19] L. Falquet, L. Bordoli, V. Ioannidis, M. Pagni, C. V. Jongeneel, Swiss EMBnet node web server. *Nucleic Acids Res.* 2003, 31, 3782–3783.

[20] Y. Murakami, S. Jones, SHARP2: protein–protein interaction predictions using patch analysis. *Bioinformatics* 2006, 22, 1794–1795.

[21] J. Bernauer, R. P. Bahadur, F. Rodier, J. Janin, A. Poupon, DiMoVo: a Voronoi tessellation-based method for discriminating crystallographic and biological protein–protein interactions. *Bioinformatics* 2008, 24, 652–658.

[22] A. Pedretti, L. Villa, G. Vistoli, VEGA — an open platform to develop chemo-bio-informatics applications, using plug-in architecture and script programming. *J. Comput. Aid. Mol. Des.* 2004, 18, 167–173.

[23] T. A. Binkowski, S. Naghibzadeh, J. Liang, CASTp: computed atlas of surface topography

of proteins. *Nucleic Acids Res.* 2003, 31, 3352–3355.
[24] Y. Rodriguez-Soca, C. R. Munteanu, J. Doradob, J. Rabuñal, A. Pazos, H. González-Díaz, Plasmod-PPI: a web-server predicting complexbiopolymer targets in plasmodium with entropymeasures of protein–proteininteractions. *Polymer* 2010, 51, 264–273.
[25] A. Banerji, I. Ghosh, Revisiting the myths of protein interior: studying proteins with mass-fractal hydrophobicity-fractal and polarizability-fractal dimensions. *PLoS ONE* 2009, 4, e7361.
[26] K. P. Tan, R. Varadarajan, M. S. Madhusudhan, DEPTH: a web server to compute depth and predict small-molecule binding cavities in proteins. *Nucleic Acids Res.* 2011, 39: W242–W248.
[27] N. Smith, S. Witham, S. Sarkar, J. Zhang, L. Li, C. Li, E. Alexov, DelPhi web server v2: incorporating atomic-style geometrical figures into the computational protocol. *Bioinformatics* 2012, 28, 1655–1657.
[28] E. Gasteiger, C. Hoogland, A. Gattiker, S. Duvaud, M. R. Wilkins, R. D. Appel, A. Bairoch, *Protein Identification and Analysis Tools on the ExPASy Server*; (In) J. M. Walker (ed): The Proteomics Protocols Handbook, Humana Press, 2005, p571–607.
[29] P. Zhou, F. Tian, Z. Shang, 2D depiction of nonbonding interactions for protein complexes. *J. Comput. Chem.* 2009, 30, 940–951.
[30] Z. R. Li, H. H. Lin, L. Y. Han, L. Jiang, X. Chen, Y. Z. Chen, FEAT: a web server for computing structural and physicochemical features of proteins and peptides from amino acid sequence. *Nucleic Acids Res.* 2006, 34, W32–W37.
[31] R. W. W. Hooft, C. Sander, G. Vriend, Positioning hydrogen atoms by optimizing hydrogen-bond networks in protein structures. *Proteins* 1996, 26, 363–376.
[32] K. G. Tina, PIC: protein interactions calculator. *Nucleic Acids Res.* 2007, 35, W473–W476.