REVIEW ARTICLE

INTERNATIONAL JOURNAL OF RESEARCH IN PHARMACEUTICAL SCIENCES

In silico analysis of protein-protein interfaces: Tools and approaches

Christina Nilofer, Arumugam Mohanapriya[*]

School of Biosciences and Technology, Vellore Institute of Technology, Vellore Campus, Tiruvalam Road, Katpadi, Vellore – 632014, Tamil Nadu, India

ABSTRACT

Two or more proteins interact *in vivo* to perform complex molecular functions including catalysis, regulation, assembly, immunity and inhibition through the formation of stable interfaces. This interaction is governed by several factors that are selective, sensitive and specific in nature. Several interface features has been documented since 1975. The study of these interface features of proteins and their dynamicity during interaction with different proteins help understanding the mechanisms underlying diverse molecular functions and its biological processes. Computational tools greatly assist in studying such interface features that determine the interaction between two or more proteins, and in this context, this review enumerates the different interface features reported thus far along with the tools that aid in deciphering protein features (physicochemical characteristics, binding site and interface residue prediction and hotspot residues) along with their approaches that are employed in the prediction these features. Also, the review discusses the advantages and limitations of experimental techniques and computational biological tools deployed for deciphering the protein-protein interactions. Altogether, the review will provide insights into the optimal tools and different strategies involved in protein interaction studies that would facilitate the researchers to understand the protein structural features and molecular principles of protein-protein interaction with known functions.

[*]Corresponding Author

Name: Arumugam Mohanapriya
Phone: +91 - 9843111639
Email: mohanapriyaa@vit.ac.in

INTRODUCTION

Proteins are the fundamental element for any biological process in a living organism. Two or more proteins communicate with one another to carry out specific molecular functions including catalysis, regulation, assembly, immunity and inhibition with high selectivity, specificity and sensitivity through the formation of stable interface (Schreiber and Keating, 2011). Hence, studies about protein-protein interactions (PPI) play a vital role in understanding the molecular mechanisms of several biological processes in a cell.

The developments in high throughput *in-vivo* and *in-vitro* techniques make it feasible to determine the protein-protein interactions experimentally. But, the experimental techniques come with their own limitations including time-consumption; high cost, laborious, high false positive rate and inaccuracy in data generation that affecting the interaction results. These limitations pave way for the utility of *in-silico* methods which produce results that are near to those determined using experimental techniques (Mrowka, 2001).

Protein interactions are been comprehensively analyzed since 1975 for several physicochemical features at the interface, as these features helps in decoding the molecular principal that are underlying different protein functions (Jones and Thornton, 1995). Protein interface features guide proteins to select their partners with high precision through the formation of a stable interface. Protein-protein interface features are studied using the 3D structures available at Protein DataBank (PDB). Protein interfaces are characterized based on the physicochemical features of the interface, which in-turn is based on the varying strength of protein association among different protein complexes.

Proteins associate with multiple partners at varying intensity forming many interfaces resulting in a permanent or transient association. Prediction of binding sites at weak or transient protein interfaces is an intricate process due to the limited availability of transient protein 3D structures in the PDB (Ozbabacan *et al.*, 2011). Interactions between two proteins (dimer) are considered to be the strong as its monomer are rarely functional (Jones and Thornton, 1995). Hence, the structural information of an interaction combined with the deciphered interface features is important in understanding the molecular mechanism of a biological function.

Therefore, in this review we present different protein interface features reported thus far along with the currently available and updated tools that aid in deciphering protein features. Also, we discuss about the advantages and limitations of experimental techniques and computational biological tools that are deployed for deciphering the protein-protein interactions.

Protein-Protein Interface Structural Features

Protein interface (Figure 1) is a part of the surface that comes in contact with the other proteins during binding. Protein interface physicochemical features play an important role in understanding the principle of protein interaction and the disease mechanism. This understanding leads to the discovery of optimal therapeutics for diseases. Hence, we present here a comprehensive review on the interface features that are discussed in the literature since 1975 (Table 1).

Role of Hydrophobicity In Protein-Protein Interaction

In protein-protein interaction the role of hydrophobicity and the structure responsible for the strong binding are studied in the following studies. Chothia and Janin (1975) showed that the interfaces are closely packed and concluded that the interfaces are majorly stabilized by hydrophobicity and that the complementarity to play the selective role in selecting the proteins that are to-be associated. Korn and Burnett (1991) described that the hydropathy distribution is high at interface compared to surface but less than core by reviewing 40 multisubunit proteins and 2 protein-protein complexes. They also devised a tool called hydropathy complementarity to quantify the strength of hydropathy distribution in the protein. Nearly after two decades Jones and Thornton (1995) found that the hydrophobic residues are more at the interface compared to the surface but less than the protein interior by structurally characterizing the protein interface and by comparing the interface residues with the surface and core. Xu *et al.* (1997) discussed about the importance of hydrophobic effect at the interface using a 362 non-redundant protein interfaces and 57 oligomeric interfaces and explored that the effect of hydrophobicity in protein binding is not as strong as it is in protein folding. Therefore these findings show the importance of hydrophobicity, hydrophobic effects, hydrophobic residues and hydrophobic patches at the interface and its impact on stability in protein binding.

Function of Interface

Nearly a decade later the role of the subunit interface in stabilizing the protein-protein binding has been studied. Miller *et al.* (1987) with 23 oligomeric proteins recognized that the surface area within and between the oligomeric proteins are directly proportional to the relative molecular mass and also stated that there is slight variation in the exposure of the surface to the solvent between the core and the interface region. This finding may have a significant role in the stability and activity of the oligomeric protein interfaces.

Caffrey (2004) tested 64 protein interfaces with conservation scores obtained from two multiple sequence alignment types and recognized fewer alignment gaps present at the obligate interfaces when compared to transient interfaces and also reported that the buried interface residues are more conserved than the partially buried residues. Bahadur *et al.* (2004) used interface area, crystal packing, residue propensities and hydrophobic interaction energy to measure specific interfaces (70 protein-protein complexes) and non-specific interfaces (188 monomeric proteins) and observed that the atomic packing is less compact in non-specific interfaces in comparison to the specific interfaces. Vaishnavi *et al.* (2010) reported that the variation in the size of interacting protein partners to be a significant factor in deciding the mode of

interaction using 156 heterodimer protein Bahadur *et al.* (2004) used interface area, crystal packing, residue propensities and hydrophobic interaction energy to measure specific interfaces (70 protein-protein complexes) and non-specific interfaces (188 monomeric proteins) and observed that the atomic packing is less compact in non-specific interfaces in comparison to the specific interfaces. Vaishnavi *et al.* (2010) reported that the variation in the size of interacting protein partners to be a significant factor in deciding the mode of interaction using 156 heterodimer protein interfaces.
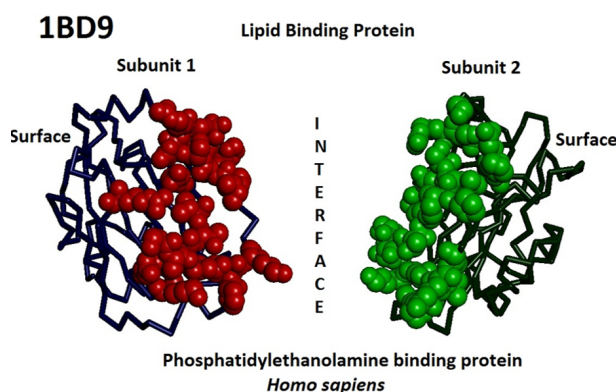

Figure 1: Depiction of surface and interface of a protein subunit-subunit interacting complex shown using PDB ID: 1BD9

Nature of Recognition Sites

Conte *et al.* (1999) analyzed 75 protein-protein interfaces with known structure and observed small interfaces with small conformational changes and large interfaces with large conformational changes and also reported that some interfaces are predominantly polar and some are non-polar with charged residues above average. Jones and Thornton (1997) analyzed protein interfaces using surface patches made of six interface features such as solvation potential, hydrophobicity, accessible surface area (asa), interface residue propensity, planarity and protrusion (SHARP2) to distinguish interface patches from surface patches and to identify the position of the recognition sites. Xu *et al.* (1997) studied about the hydrogen bonds and salt bridges across 319 non-redundant protein-protein interfaces and acknowledged the presence of pattern of charge complementarity and conservations of hydrogen bonds at the interface. Furthermore, the importance of specificity at protein interface and its application in docking and molecular design are also illustrated. Chakrabarti and Janin (2002) described that small interfaces have single patch and large interfaces enclosing multiple patches, where each patch buries interface atoms surrounded by the rim. This was proved using a dataset of 70

protein interfaces. Guharoy and Chakrabarti (2010) utilized 121 homodimers and 392 heterocomplexes and reported that conserved residues at protein interfaces are not random but are distinctly clustered. However, they have also reported that the buried residues at the interfaces are more conserved than the partially buried residues. It has been recently hypothesized that evolutionarily constraints of signal transduction and enzymes interfaces play an important role in characterizing the interacting surface; while energetic constraints are deterministic in immune, inhibitor and structural assembly interfaces (Marchetti *et al.*, 2019). Li *et al.* (2006) found hotspot residues to play a vital role in determining the stability of the interacting proteins. Hotspots possess optimal energy and hence they are most favored residues at the interface. Gromiha *et al.* (2009) designed an energy based approach to understand the mechanism of recognition at protein interface and found that charged and aromatic residues to play an important role in protein binding.

Formation of Homodimer and Heterodimer Interfaces

Jones and Thornton (1995) explored that the obligatory (permanent) protein interfaces are closely packed with fewer hydrogen bonds than non-obligatory (transient) interfaces using the influence of factor in the formation of the protein-protein association in homodimer, heterodimer, enzyme inhibitory complexes and antigen-protein complexes. Zhanhua *et al.* (2005) identified critical heterodimer interface parameters through multi-dimensional scaling in Euclidian space using 65 heterodimers with known 3D structures. Sowmya *et al.* (2015) analyzed 278 heterodimer protein interfaces using five key features such as interface area, interface polar residues abundance, hydrogen bonds, solvation free energy gain and binding energy and found salt bridges to increase with interface area in regulator-inhibitor interfaces. This suggests that regulator-inhibitor interfaces are held together mainly by salt bridges through electrostatics interactions. Nilofer *et al.* (2020) showed that protein interfaces are highly pronounced with van der Waals (vdW) and that H-bonds and electrostatic to play a discerning role towards broad functional specificity where H-bonds increase in obligatory and immune complexes while electrostatic increase in non-obligatory regulator-inhibitors with interface size (number of interface residues). Moreover, they also reported that small interfaces are rich in electrostatics and that they are often found to be linked to regulatory function by analyzing 2557 homodimer and 393 heterodimers interfaces using

Table 1: List of protein interface features documented in the literature since1975 given in chronological order

S. No	Citation	Dataset size	Interface features	Year	Country	Group	Reference
1	Chothia and Janin	3	Hydrophobicity and complementarity	1975	France	J. Janin	Chothia and Janin (1975)
2	Miller et al.	23	Hydrophobic energy, van der Waals forces and hydrogen bonds	1987	UK	J. Janin	Miller *et al.* (1987)
3	Korn and Burneet	42	Distribution of hydropathy	1991	Pennsylvania	Burnett RM	Korn and Burnett (1991)
4	Jones and Thornton	32	Accessible surface area, sphericity, shape, hydrophobicity, amino acid residues, atom distribution, hydrogen bonding, salt bridges, disulphide bond, interface segments, secondary structures and gap volume	1995	UK	JM. Thornton	Jones and Thornton (1995)
5	Xu et al.	319	Pattern of charge complementarity, hydrogen bonds and salt bridges	1997	USA	R. Nussinov	Xu *et al.* (1997)
6	Tsai et al.	419	Hydrophobic effect, charged and polar residues, hydrogen bonds and ion pairs	1997	USA	R. Nussinov	Xu *et al.* (1997)
7	Jones and Thornton	48	Surface patches	1997	UK	JM. Thornton	Jones and Thornton (1997)
8	Lo Conte et al.	75	Polar and non-polar residues	1999	France	J. Janin	Conte *et al.* (1999)
9	Chakrabarti and Janin	70	Small and large interface area	2002	France	J. Janin	Chakrabarti and Janin (2002)

Continued on next page

Table 1 continued

S. No	Citation	Dataset size	Interface features	Year	Country	Group	Reference
10	Caffrey et al.	64	Obligate and transient interface	2004	USA	ES. Huang	Caffrey (2004)
11	Bahadur et al.	192	Hydrophobic interactions	2004	France	J. Janin	Bahadur *et al.* (2004)
12	Zhanhua et al.	65	Interface area, Hydrogen bonds, hydrophobicity and residue propensity	2005	Singapore	P. Kangueane	Zhanhua *et al.* (2005)
13	Li et al.	296	Hotspot residues	2006	Singapore	P. Kangueane	Li *et al.* (2006)
14	Pal et al.	204	Interface area and peptide segment	2007	France	J. Janin	Pal *et al.* (2007)
15	Gromiha et al.	153	Charged and aromatic residues, hydrophobic residues, hydrogen bond and polar residues	2009	Japan	K. Fukui	Gromiha *et al.* (2009)
16	Guharoy and Chakrabarti	326	Conserved residues, hydrophobic, aromatic and charged residues	2010	India	P. Chakrabarti	Guharoy and Chakrabarti (2010)
17	Sowmya et al.	278	Interface area, Interface polar residues abundance, hydrogen bonds, solvation free energy gain, binding energy and charged residues	2015	Australia	S. Ranganathan	Sowmya *et al.* (2015)
18	Nilofer et al.	2950	Large interface, small interface, vdW dominant, vdW sub-dominant, Interface area, interface size, van der Waals, electrostatics and hydrogen bonds	2019	India	P. Kangueane	Nilofer *et al.* (2020)

Table 2: Tools specific to protein interface and hot-spot analysis are listed

S. No	Tool name	Description	University, Country	Type	Reference
Tools for Physicochemical characteristics prediction					
1	SHARP2	Predicts potential protein-protein recognition sites using protein structures	University of Sussex, UK	Web Server (not updated)	Jones and Thornton (1997)
2	ProFace	Dissects protein interface and derive various physicochemical parameters	Bose Institute, India	Web Server (not updated)	Saha et al. (2006)
3	PIC	Protein Interactions Calculator	Indian Institute of Science, India	Web Server	Tina et al. (2007)
4	ProtorP	Calculates physicochemical parameters at protein recognition sites	University of Sussex, UK	Web Server (not updated)	Reynolds et al. (2009)
5	PPCheck	Identifies non-covalent interactions at the interface	National Centre for Biological Sciences, India	Web Server	Sukhwal and Sowdhamini (2015)
6	ProFunc	Identifies protein function using protein 3D structure	EMBL-EBI, UK	Web Server	Laskowski et al. (2005)
7	PDBsum	Provides structural information for protein structures deposited in PDB	EMBL-EBI, UK	Web Server	Laskowski et al. (2018)
Tools for the prediction of binding site and interface residues at the interface					
8	HBPLUS	Hydrogen bond calculator	University College London, UK	Program	McDonald and Thornton (1994)
9	ConSurf Server	Server for the Identification of Functional Regions in Proteins	Tel Aviv University, Israel	Web Sever	Landau et al. (2005)
10	PIER	Protein IntErface Recognition	Scripps Research Institute, USA	Web Server	Kufareva et al. (2007)
11	SPPIDER	Solvent accessibility based Protein-Protein Interface iDEntification and Recognition	Children's Hospital Research Foundation, USA	Web Server	Porollo and Meller (2006)
12	iFrag	Protein -protein binding site predicting server	Universitat Pompeu Fabra, Barcelona	Web Server	Garcia-Garcia et al. (2017)
Tools for the prediction of Hot Spot Residues at the interface					
13	PCRPi-W	Predicting Critical Residues at Protein interface - Webserver	University of Leeds, UK	Web Server	Mora et al. (2010)
14	SpotOn	Determines hotspots at protein interface	Utrecht University, Netherlands	Web Server	Moreira et al. (2017)

vdW, hydrogen bonds and electrostatics.

Experimental Methods and Its Limitations

Determination of protein protein interactions (PPI) using experimental methods depends on the physical and chemical characteristics of protein interface (Szilágyi *et al.*, 2005). Advancements in experimental methods are promising, yet with limited precision and coverage. Hence a consensus results from different experimental approaches are to be considered for PPI prediction. Experimental methods include *in-vivo* and *in-vitro* techniques. *In-vivo* techniques include Yeast two hybrid and affinity purification succeeded by mass spectrometry and *in-vitro* techniques include co-immuno precipitation, pull down assay and label transfer assay are previously discussed elsewhere (Fields and kyu Song, 1989; Piehler, 2005). Therefore, computational biological approaches play a pivotal role in choosing the important data for high throughput experimental analysis, thereby minimizing the cost, time consumed and high false positive rate.

Protein-Protein Interface Tools and Approaches

Protein interface is the significant part of the protein-protein interaction. Therefore analyzing protein interfaces are important in understanding the molecular mechanism underlying biological function. Protein-protein interface is made up of physicochemical features, interface and hotspot residues (Table 2). Hence, we present here a comparative discussion about the available tools and their different approaches under each category.

Tools for The Prediction of Physicochemical Features at the Interface

ProFace (Saha *et al.*, 2006) is a web based suite of programs used for calculating physicochemical parameters across interface formed between two or more subunits. It also helps in identifying the interface residues at the protein recognition sites. PIC (Protein Interactions Calculator) (Tina *et al.*, 2007) estimates various interactions including disulphide bonds, hydrophobic interactions, ionic interactions, hydrogen bonds, aromatic-aromatic, aromatic-sulphur and cation - π interactions at inter-protein and intra-protein complexes using standard criteria. Moreover, the server also analyzes the accessible surface area and distance of a residue from the surface of a protein. Each interaction can be visualized using Jmol or Rasmol visualization tool. ProtorP (Reynolds *et al.*, 2009) is a web server helps in predicting the physicochemical features (size and shape, intermolecular bonding, residue and atom composition and secondary structure contributions) at the protein recognition

sites that contribute to the overall binding energy of the interaction. PPCheck (Sukhwal and Sowdhamini, 2015) is a webserver aid in computing total stabilizing energy, hydrogen bonds, hydrophobic interactions, salt bridges, favourable electrostatic interactions, unfavourable electrostatic interactions and short contacts by assigning pseudo-energies to quantify the strength of protein-protein interaction. Moreover, the server also calculates hotspot residues, computational alanine scanning, residue conservation and the prediction of right docking pose. ProFunc (Laskowski *et al.*, 2005) is a web based predicting server used to predict the function of a protein with known structure. PDBsum (Laskowski *et al.*, 2018) is a web server providing structural details on a given protein. The output is generally is in image format and the structural details include protein secondary structure, protein-ligand and protein-DNA interactions, PROCHECK analyses of structural quality, and many others. The output image files can be viewed using RasMol, PyMOL and JavaScript Viewer.

SHARP2, ProtorP and ProFace are eminent web servers in predicting the physicochemical features at protein interface. However, these web servers are not updated and non-functional. PIC is the only webserver calculating the physicochemical properties for intra and inter-protein interaction when compared to other web servers. PPCheck gives a complete analysis for protein interaction in one single window including residue conservation, computing the strength of protein interactions, identifying the hotspot residues, performing computational alanine scanning and predicting the right docking pose. Therefore PPCheck is an independent platform for the analysis of protein protein interaction. However MAPPIS also recognizes spatially conserved physicochemical features and identify hot spot residues using multiple sequence alignment. PIC is restricted to the analysis of physicochemical properties, whereas PPCheck is extended to predict the other aspects of protein protein interaction. PDBsum serves as a finest web server in providing the structural information of a protein in an image format.

Tools for Predicting Protein Interface Residues

HBPLUS (McDonald and Thornton, 1994) is a web based program to calculate the hydrogen bonding and its position at the interface. ConSurf Server (Landau *et al.*, 2005) is an automated web tool to recognize the binding residues at the interface using evolutionary conservation score. These scores are generated using Bayesian method. PIER (Protein IntErface Recognition) (Kufareva *et al.*,

2007) uses local statistical properties on the surface to identify the binding residues at atomic level for monomers. SPPIDER (Solvent accessibility based Protein-Protein Interface identification and Recognition) (Porollo and Meller, 2006) is a protein interface residue prediction method using RSA relative solvent accessibility of an amino acid residue as a fingerprint. SPPIDER is developed using machine learning approaches including Support Vector Machines and Neural networks and other informative features. iFrag (Garcia-Garcia et al., 2017) is a sequence based protein interface residue predicting server using minimal sequence fragments. iFrag accepts FASTA sequence as input.

iFrag is the only server that accept sequence as input whereas other methods are depended on protein's structural information. SPPIDER's RSA fingerprint method is unique and novel method and yield better results than other methods. ConSurf server is based on evolutionary conservation and structural information of a protein complex for prediction.

Tools Predicting the Hotspot Residues at the Interface

Hotspot residues are those residues at the interface playing an important role in binding affinity thereby increasing the total binding energy. Prediction of hotspot residues holds a great impact in the field of drug discovery and protein design. Here are some currently active hotspot residue predictor with different approaches and visualization.

PCRPi-W - Presaging Critical Residues in Protein interfaces-Web Server (Mora et al., 2010) identifies protein interfacial hotspot residues based on the integration of structural, energetic, and evolutionary-based measures by using Bayesian Networks (BNs).

SpotOn (Moreira et al., 2017) identifies and classify the interface residues as hotspots and null spots with high accuracy and sensitivity. The Spot On algorithm was developed using machine learning approach.

PCRPi-W accepts only PDB structural data as input, while 3D structural data are not available for all proteins. However, SpotOn accepts both structure and sequence data as an input. Accuracy, sensitivity and specificity play a vital role in protein interfacial hotspot prediction.

SpotOn server provide hotspot residues in a table format and can also be visualized using Jmol. In addition, it provides sequence viewer to tabulate the probability of hotspot prediction. Furthermore, PCRPi-W uses evolutionary approach measured using Bayesian Networks to identify the hotspot residues.

Limitations in the Usage of Computational Biological Tools

Computational biological tools are available as commercial (licensed) and free softwares (educational purpose) on the web. Commercial tools are all-in-one platform type, expensive however properly maintained, updated and provide customer care support. The free trial version of commercial tools is available with minimal features for 30 days upon payment agreement. Contrarily, most of the free softwares are neither properly maintained nor updated and the issues faced with the working of the software are not properly addressed. In majority of cases the published contact person's e-mail address does not work due to change of work place (laboratory) or the completion of project tenure. On the other hand, free software request for educational e-mail address for usage, and the accessibility is restricted for independent researchers or industrial officials. Few computational tools links (for example SHARP2, ProtorP and ProFace) given in the published paper are not working. Furthermore there are few tools that are created using CUI (Character User Interface), in such case the software are operated using command line. Researchers with no coding knowledge find it difficult with the software operation. Unlike wet-lab experiments there is no one-platform one-solution situation with computational biology tools as there are multiple tools with multiple approaches for one biological problem. Hence scientists have to run their query in multiple tools to trust their results accuracy. Computational biology tools require high-end computers/servers for operation nevertheless the results are not considered to be accurate until verified by wet-lab experiments. Computational tools also have reasonable high false positive and false negative rates. These limitations are to be considered and rectified while developing computational biological tools.

CONCLUSION

Protein-protein interactions play a fundamental role in all molecular processes to perform a biological function. While, protein interface features play a pivotal role in selecting a selective, specific and sensitive partner for interaction. Protein interfaces are described using several features such as hydrophobic effect, hydropathy, hydration, H-bonds, charge and shape complementarity, peptide segments, ion-pairs, single and multiple patches, hydrophobicity, hydrophilicity, interface residue propensity, charged, aromatic and Arginine residues, binding

energy, hotspot residues, van der Waals, electrostatics and salt bridges. These features help us in deciphering the driving force for protein associations. Advancements in protein-protein docking methods, algorithms, tools and databases provide deep insights about the protein interaction but poses difficulty in mimicking these features in *invivo* condition with high accuracy. Hence, understanding the fundamentals of protein-protein interactions with interface features will provide us with deep knowledge and clarity about the stable interface formation. An examination and combination of both atomic and residue features at the interface compared to surface in different functional protein complexes is required to differentiate protein interfaces. Therefore, studying the distinguishable features at different protein interfaces with known structure and consequently analyzing the resulting recognition pattern of the interface are critical in identifying the protein partners for association.

Funding Support

The authors declare that they have no funding support for this study.

Conflict of Interest

The authors declare that they have no conflict of interest for this study.

REFERENCES

Bahadur, R. P., Chakrabarti, P., Rodier, F., Janin, J. 2004. A Dissection of Specific and Non-specific Protein–Protein Interfaces. *Journal of Molecular Biology*, 336(4):943–955.

Caffrey, D. R. 2004. Are protein-protein interfaces more conserved in sequence than the rest of the protein surface? *Protein Science*, 13(1):190–202.

Chakrabarti, P., Janin, J. 2002. Dissecting protein-protein recognition sites. *Proteins: Structure, Function, and Genetics*, 47(3):334–343.

Chothia, C., Janin, J. 1975. Principles of protein–protein recognition. *Nature*, 256(5520):705–708.

Conte, L. L., Chothia, C., Janin, J. 1999. The atomic structure of protein-protein recognition sites 1 1Edited by A. R. Fersht. *Journal of Molecular Biology*, 285(5):2177–2198.

Fields, S., kyu Song, O. 1989. A novel genetic system to detect protein–protein interactions. *Nature*, 340(6230):245–246.

Garcia-Garcia, J., Valls-Comamala, V., Guney, E., Andreu, D., Muñoz, F. J., Fernandez-Fuentes, N., Oliva, B. 2017. iFrag: A Protein–Protein Interface Prediction Server Based on Sequence Fragments. *Journal of Molecular Biology*, 429(3):382–389.

Gromiha, M. M., Yokota, K., Fukui, K. 2009. Energy based approach for understanding the recognition mechanism in protein–protein complexes. *Molecular BioSystems*, 5(12):1779.

Guharoy, M., Chakrabarti, P. 2010. Conserved residue clusters at protein-protein interfaces and their use in binding site identification. *BMC Bioinformatics*, 11(1):286.

Jones, S., Thornton, J. M. 1995. Protein-protein interactions: A review of protein dimer structures. *Progress in Biophysics and Molecular Biology*, 63(1):31–65.

Jones, S., Thornton, J. M. 1997. Analysis of protein-protein interaction sites using surface patches 1 1Edited by G.Von Heijne. *Journal of Molecular Biology*, 272(1):121–132.

Korn, A. P., Burnett, R. M. 1991. Distribution and complementarity of hydropathy in mutisunit proteins. *Proteins: Structure, Function, and Genetics*, 9(1):37–55.

Kufareva, I., Budagyan, L., Raush, E., Totrov, M., Abagyan, R. 2007. PIER: Protein interface recognition for structural proteomics. *Proteins: Structure, Function, and Bioinformatics*, 67(2):400–417.

Landau, M., Mayrose, I., Rosenberg, Y., Glaser, F., Martz, E., Pupko, T., Ben-Tal, N. 2005. ConSurf 2005: the projection of evolutionary conservation scores of residues on protein structures. *Nucleic Acids Research*, 33(Web Server):W299–W302.

Laskowski, R. A., Jabłońska, J., Pravda, L., Vařeková, R. S., Thornton, J. M. 2018. PDBsum: Structural summaries of PDB entries. *Protein Science*, 27(1):129–134.

Laskowski, R. A., Watson, J. D., Thornton, J. M. 2005. ProFunc: a server for predicting protein function from 3D structure. *Nucleic Acids Research*, 33(Web Server):W89–W93.

Li, L., Zhao, B., Cui, Z., Gan, J., Sakharkar, M. K., Kangueane, P. 2006. Identification of hot spot residues at protein-protein interface. *Bioinformation*, 1(4):121–126.

Marchetti, F., Capelli, R., Rizzato, F., Laio, A., Colombo, G. 2019. The Subtle Trade-Off between Evolutionary and Energetic Constraints in Protein–Protein Interactions. *The Journal of Physical Chemistry Letters*, 10(7):1489–1497.

McDonald, I. K., Thornton, J. M. 1994. Satisfying Hydrogen Bonding Potential in Proteins. *Journal of Molecular Biology*, 238(5):777–793.

Miller, S., Lesk, A. M., Janin, J., Chothia, C. 1987. The accessible surface area and stability of oligomeric proteins. *Nature*, 328(6133):834–836.

Mora, J. S., Assi, S. A., Fernandez-Fuentes, N. 2010. Presaging Critical Residues in Protein interfaces-Web Server (PCRPi-W): A Web Server to Chart Hot Spots in Protein Interfaces. *PLoS ONE*, 5(8):e12352.

Moreira, I. S., Koukos, P. I., Melo, R., Almeida, J. G., Preto, A. J., Schaarschmidt, J., Trellet, M., Gümüş, Z. H., Costa, J., Bonvin, A. M. J. J. 2017. SpotOn: High Accuracy Identification of Protein-Protein Interface Hot-Spots. *Scientific Reports*, 7(1):8007.

Mrowka, R. 2001. Is There a Bias in Proteome Research? *Genome Research*, 11(12):1971–1973.

Nilofer, C., Sukhwal, A., Mohanapriya, A., Sakharkar, M. K., Kangueane, P. 2020. Small protein–protein interfaces rich in electrostatic are often linked to regulatory function. *Journal of Biomolecular Structure and Dynamics*, 38(11):3260–3279.

Ozbabacan, S. E. A., Engin, H. B., Gursoy, A., Keskin, O. 2011. Transient protein-protein interactions. *Protein Engineering Design and Selection*, 24(9):635–648.

Pal, A., Chakrabarti, P., Bahadur, R., Rodier, F., Janin, J. 2007. Peptide segments in protein-protein interfaces. *Journal of Biosciences*, 32(1):101–111.

Piehler, J. 2005. New methodologies for measuring protein interactions in vivo and in vitro. *Current Opinion in Structural Biology*, 15(1):4–14.

Porollo, A., Meller, J. 2006. Prediction-based fingerprints of protein-protein interactions. *Proteins: Structure, Function, and Bioinformatics*, 66(3):630–645.

Reynolds, C., Damerell, D., Jones, S. 2009. ProtorP: a protein-protein interaction analysis server. *Bioinformatics*, 25(3):413–414.

Saha, R. P., Bahadur, R. P., Pal, A., Mandal, S., Chakrabarti, P. 2006. ProFace: a server for the analysis of the physicochemical features of protein-protein interfaces. *BMC Structural Biology*, 6:11.

Schreiber, G., Keating, A. E. 2011. Protein binding specificity versus promiscuity. *Current Opinion in Structural Biology*, 21(1):50–61.

Sowmya, G., Breen, E. J., Ranganathan, S. 2015. Linking structural features of protein complexes and biological function. *Protein Science*, 24(9):1486–1494.

Sukhwal, A., Sowdhamini, R. 2015. PPCheck: A Webserver for the Quantitative Analysis of Protein-Protein Interfaces and Prediction of Residue Hotspots. *Bioinformatics and Biology Insights*, 9:BBI.S25928.

Szilágyi, A., Grimm, V., Arakaki, A. K., Skolnick, J.

2005. Prediction of physical protein–protein interactions. *Physical Biology*, 2(2):S1–S16.

Tina, K. G., Bhadra, R., Srinivasan, N. 2007. PIC: Protein Interactions Calculator. *Nucleic Acids Research*, 35(Web Server):W473–W476.

Vaishnavi, A., Sowmya, G., Kalaivanii, J., Ilakya, S., Kangueane, U., Kangueane, P. 2010. Interaction modes at protein hetero-dimer interfaces. *Bioinformation*, 4(7):310–319.

Xu, D., Tsai, C. J., Nussinov, R. 1997. Hydrogen bonds and salt bridges across protein-protein interfaces. *Protein Engineering Design and Selection*, 10(9):999–1012.

Zhanhua, C., Gan, J. G.-K., lei, L., Sakharkar, M. K., Kangueane, P. 2005. Protein subunit interfaces: heterodimers versus homodimers. *Bioinformation*, 1(2):28–39.