HHS Public Access

Author manuscript

Methods Mol Biol. Author manuscript; available in PMC 2021 June 08.

Published in final edited form as:

Methods Mol Biol. 2019 ; 1903: 1–21. doi:10.1007/978-1-4939-8955-3_1.

Methods for Discovering and Targeting Druggable Protein-Protein Interfaces and Their Application to Repurposing

E. Sila Ozdemir[#1], **Farideh Halakou**[#2], **Ruth Nussinov**[3,4], **Attila Gursoy**[2], **Ozlem Keskin**[1]

[1]Department of Chemical and Biological Engineering, Koc University, Istanbul, Turkey.

[2]Department of Computer Engineering, Koc University, Istanbul, Turkey.

[3]Cancer and Inflammation Program, Leidos Biomedical Research, Inc., Frederick National Laboratory for Cancer Research, National Cancer Institute at Frederick, Frederick, MD, USA.

[4]Department of Human Molecular Genetics and Biochemistry, Sackler School of Medicine, Tel Aviv University, Tel Aviv, Israel.

[#] These authors contributed equally to this work.

Abstract

Drug repurposing is a creative and resourceful approach to increase the number of therapies by exploiting available and approved drugs. However, identifying new protein targets for previously approved drugs is challenging. Although new strategies have been developed for drug repurposing, there is broad agreement that there is room for further improvements. In this chapter, we review protein-protein interaction (PPI) interface-targeting strategies for drug repurposing applications. We discuss certain features, such as hot spot residue and hot region prediction and their importance in drug repurposing, and illustrate common methods used in PPI networks to identify drug off-targets. We also collect available online resources for hot spot prediction, binding pocket identification, and interface clustering which are effective resources in polypharmacology. Finally, we provide case studies showing the significance of protein interfaces and hot spots in drug repurposing.

Keywords

Hot spots; Hot regions; Network-based approaches; Interface motifs; Protein interface clustering

1 Introduction

Over the years, studies on drug development are accelerated both in the pharmaceutical industry and in academia; still the increasing demand for new drugs cannot be met. This situation underscores the need for innovative strategies and techniques. However, discovering new drugs and drug targets is challenging. Until recently, drug targets were largely limited to enzymes and receptors. Small molecules that target enzymes mostly mimic and compete with the substrates of the enzymes [1]. The pressing demand for new drugs has led to an increase of investments of pharmaceutical companies in drug development. To

address this challenge, one approach, which is discussed here, involves targeting protein-protein interactions (PPIs) [2]. Proteins typically execute their function by interacting with other proteins. These interactions transmit signaling cues, with the signaling cascading downstream through PPIs. Signaling takes place through transient and permanent PPIs and is key to most (or all) cellular functions, including cell proliferation, motility, and growth [3]. Alterations in PPI interfaces may affect signal transduction, leading to dysfunction and disease [4-6]. There are many computational and a few experimental methods for protein-protein interaction prediction [7]. Given the large number of PPIs, the repurposing possibilities of drugs targeting the interfaces appear high in principle. However, traditional PPI drug discovery has been stymied and challenging. Still, progress has been made with small molecules and with fragment-based approaches [8-10].

Those efforts were guided by the traditional philosophy that coined the “one drug one target” paradigm, reflecting the aim of drug specificity, i.e., low toxicity. However, because drugs are small and typically hydrophobic with aromatic rings [11], they target protein surfaces that have complementing properties. Moreover, the binding regions of unrelated proteins can have similar shapes and surfaces [12, 13]. Therefore, many drugs may have multiple targets, albeit with varied affinity. Polypharmacology aims to capitalize on this and find drugs that bind to multiple protein targets, which, upon further optimization, can be used for repurposing [14, 15]. The functions of those proteins can vary. Overington et al. [16] conducted a comprehensive survey on earlier reports and proposed that clinical drugs act on total 324 drug targets. A recent study also showed that in the human proteome, on average a drug can bind to 329 proteins. This implies that the vast majority of drugs have their own side effects [17].

Protein-protein interfaces are increasingly getting attention in drug discovery [18, 19]. Similar binding sites on protein surfaces can be used to find the potential candidates for a drug. Xie et al. [20] used protein-ligand binding profiles to observe the effects of cholesteryl ester transfer protein (CETP) inhibitors and found its unknown off-targets in genome scale. They used SOIPPA [21] to align the binding site structures and find ligand binding sites similar to the primary target in the network. Two CETP inhibitors are investigated, and their candidate off-targets are mapped to several biological pathways. Based on their results, side effects of CETP inhibitors are involved in immune response and stress control via multiple interconnected pathways. Frigola et al. [22] used the similarity of protein cavities to find the proteins that a ligand can bind in the human proteome. They used BioGPS [23] to investigate all human proteins with available 3D structure, to find potential drug targets based on cavity similarities. Based on their results, similar cavities can be found in distinct unrelated proteins, and, on average, a protein has similar binding sites to seven other proteins [22]. To study the effects of drug combinations on a network scale, they used heat flow analysis. They found that drug combinations could distribute heat in the network at least 25% better than the usage of a single drug in 20 tumor-specific networks.

Developing drugs targeting PPI is challenging [18]. Unlike enzyme binding pockets (Fig. 1a), interfaces usually do not have preexposed cavities, and their large surfaces are flat (Fig. 1b) [2, 24]. This makes the drug design process difficult, since determining where exactly the drugs should bind is crucial [25]. New strategies are being developed to overcome these

challenges [26, 27]. One of these is to identify interface residues playing roles in protein recognition and binding affinity. A small subset of residues in interfaces, which are called hot spots, are the major contributors of the binding energy [28, 29]. Studies showed that hot spots are the main targets of small molecules aiming to disrupt PPIs [6, 30, 31]. Another property related to hot spots is that they are not randomly distributed, but are typically clustered in the interfaces. These densely packed clusters are called hot regions [32]. Hot regions serve as binding platforms for protein partners. This organization of interface residues provides an insight into how small molecules may recognize the interfaces to bind them. Hot spots can be detected by experimental procedures such as alanine scanning mutagenesis [29, 33]; however performing experiments on all known PPIs to detect hot spots is infeasible. Therefore, computational techniques for predicting hot spots are on the rise, and their accuracy increases over the years as well [34]. A number of hot spot prediction algorithms have been developed. Hot spot prediction methods and tools are based on either the structure of the complex or the structure of the unbound proteins. Only a few studies predict hot spots in unbound proteins. Amino acid sequences [35], normalized interface propensity values derived from rigid body docking [36], dynamic fluctuations in high-frequency modes obtained from the Gaussian network model (GNM) [37], and measuring the dynamic exposure of hydrophobic patches [38] can be used to predict the hot spot residues of unbound protein structures. Table 1 summarizes algorithms and tools available for hot spot prediction in unbound proteins.

Most hot spot prediction algorithms focus on interaction/complex-based approaches. One pioneering work proposed a physical model to predict hot spots based on energy measurements of packing interactions, hydrogen bonds, and solvation (Robetta) [39]. Energy measurement-based prediction methods are widely used to develop new tools; for example, energies calculated by FoldX or MutaBind can be used to predict the hot spot residues [40, 41]. Estimating the energetic contribution of interfacial residues to the binding affinity, via identifying non-covalent interactions, is another method used for hot spot prediction [42]. Solvent accessibility and the total contact potential energy of the interface residues can be considered for hot spot prediction [43]. Molecular dynamics (MD) simulations constitute a more detailed and computationally powerful approach for hot spot prediction [44, 45]. Physicochemical properties of interface residues can be considered in hot spot prediction [46]. Some servers investigate the shape specificity, biochemical contact, and plasticity features of the interface residues (KFC and KFC2a) for hot spot prediction [47, 48]. Moreover, some of the atomic features such as mass, polarizability, isoelectric point of residues, and relative ASA can be combined in the prediction [49]. Table 1 outlines some hot spot prediction tools and algorithms, which predict hot spots from the protein complexes. Some of the distinguishing features and websites (if available) are listed in the table. Structure and sequence similarity of the interfaces and conservation of energetically important interface residues, such as hot spots and hot regions, can help repurposing drugs targeting PPI.

Studies showed that despite the vast number of PPIs (approximately 130,000 binary interactions between human proteins [34, 50]), there exist only a limited number of interface architectures [12, 13, 51]. A reasonable strategy to repurpose interface-targeting drugs might be to identify interface motifs sharing similar hot spots [31, 52, 53]. Sequence similarities,

evolutionary conservation, and/or similarities in 3D structures can all be used to cluster similar interfaces, albeit with possibly partially different outcomes [42, 54-58]. Tables 2 and 3 list some representative protein interface databases and binding pocket identification methods which are available online. PIFACE [54] is a database of clustered protein-protein interfaces. It consists of 22,604 unique interface structures derived from 130,209 interfaces which are extracted from protein complexes in PDB [59]. The PIFACE web server can be used to find the interface region in a protein complex and to compare the protein-protein interfaces of two different complexes.

PLIC [60] is a database of protein-ligand interactions in which 84,846 ligand binding sites are grouped into 10,858 clusters. Binding sites are extracted from the protein-ligand complexes in the PDB and compared using the PocketMatch [61] algorithm. The sc-PDB [62] is an up-to-date structure database of ligandable binding sites from the PDB. The binding sites in sc-PDB are extracted from protein complexes having a small ligand and predicted to be ligandable. The database consists of 9283 binding sites corresponding to 3678 unique proteins and 5608 unique ligands. ProtCID [63] is a database of homodimeric and heterodimeric interfaces derived from multiple crystal forms of homologous proteins. It includes chain-chain and domain-domain interactions. The current version of ProtCID, as of December 2017, consists of 125,643 chains and 115,032 domains. 3did [64] is a collection of domain-domain and domain-motif interactions derived from PDB complex structures. The current version of 3did includes 11,200 domain-domain and 702 domain-motif interactions, respectively. Similar interacting domains in 3did are clustered into interaction topologies which can show different modes of binding.

In this chapter, we explain the importance of hot spot and hot region predictions and outline the HotPoint [65] and HotRegion [66] servers, which predict hot spots and hot regions, respectively. We also detail the method investigating interface similarity to identify drug off-targets in structural PPI networks.

2 The Importance of Hot Spot and Hot Region Prediction in Drug Repurposing

A hot spot is defined as a residue causing an increase of more than 2 kcal/mol in binding free energy upon its mutation to alanine [28]. Further analysis on hot spot residues showed that Tyr, Arg, and Trp amino acids are more favorable to be hot spots compared to other amino acids. These amino acids are more prone to cause higher change in the binding free energy due to their size and conformation [29]. Hot spots are surrounded by a set of energetically less important residues. These residues form structures resembling the O-rings and protect hot spot residues from solvent molecules. The so-called O-ring theory explains that residues contributing more to the binding free energy are largely protected from contact with bulk solvent, with low or no accessible solvent area (ASA) [67, 68]. There is a correlation between the ASA of the residues and their contributions to the binding free energy; the more buried a residue, the more it contributes to the energy. However, this correlation alone is not sufficient to define a residue as a hot spot [67]. Hot regions are also important due to their contribution to the binding free energy and their contribution to

specificity to interfaces [69]. Figure 2 presents hot spots and hot regions in the interface between Cdc42 and GRD2 (PDB ID: 5CJP, chain C, and chain E) showing the 3D organization of these residues. The protein complex has a total of 17 hot spots of which 13 are clustered into 2 hot regions.

Studies imply that interfaces lacking hot spots cannot attain high affinity toward their binding partners, proteins, or specific drugs [70]. Single mutation in only one hot spot may completely abolish interaction [71]. Computational methods confirm the relationship between hot spots and druggability [72]. Drugs targeting hot spots in the protein interfaces increase the possibility of binding to the interface and establishing a stable interaction.

Hot spots are not only energetically important; they are also conserved residues [29]. These conserved residues form complementary binding sites with hot spots from other interfaces. Hot spots of one interface usually pack against hot spots of another and together establish a binding region, which provides important knowledge for drug binding sites [68]. Hot regions usually coevolve with the hot regions of their binding partners, since they consist of hot spots [68]. This also gives a critical insight for drug repurposing. Proteins having similar binding partners are most likely to have similar hot spot and hot region distribution as a result of coevolution. This observation increases the possibility to repurpose a drug targeting an interface that coevolved with other proteins [73].

Moreover, it is possible to experimentally screen FDA-approved drugs. The screening helps to identify the drugs, which bind to the interfaces with similar energetically important residues. Then, the binding affinity and efficiency of the identified drugs to the interfaces can be tested [74-76]. For example, Fang et al. reported a small-molecule antagonist, LF3, for the β-catenin/TCF4 interaction using advanced biochemical screening techniques [75]. In order to identify such a molecule, they effectively docked a library of small molecules onto experimentally identified hot spots of the interaction sites between β-catenin and TCF4. Experimental and computational approaches, which are used for drug repurposing applications, are described in more details in the following sections.

3 Methods

3.1 HotPoint and HotRegion Servers

HotPoint and HotRegion are hot spot and hot region prediction servers, respectively [65, 66]. The prediction algorithm of HotPoint primarily considers solvent accessibility and the total contact potential energy of the interface residues. Interface residues consist of nearby and contacting residues. Contacting residues are defined as two residues from different chains whose distance between any two atoms from two proteins is less than the sum of their van der Waals radii plus 0.5 Å [77]. A non-contacting residue that is closer than 6 Å to an interacting residue in the same chain is defined as nearby residue—the distance between the alpha carbons of two residues [78]. In order to determine the thresholds of HotPoint prediction model, a nonredundant ASEdb data set and a previously compiled data set from Robetta were used as training sets [33, 39]. The data consists of 150 experimentally alanine-mutated residues (58 hot spots and 92 non-hot spots). The conservation and solvent accessibility information are available for all these 150 residues. For training sets, if

mutations change the binding free energy at least 2.0 kcal/mol, these interface residues are considered as experimental hot spots. Residues whose mutations result in a change <0.4 kcal/mol are labeled as experimental non-hot spots. Other residues out of these thresholds are not included in the training. Test set is adopted from Binding Interface Database (BID) [79] which is composed of 112 residues (54 hot spots and 58 non-hot spots).

Prediction criteria such as solvent accessibility, conversation, and contact potentials are integrated to HotPoint algorithm as follows. The ASA of each residue is calculated using Naccess [80] in monomer state and in complex state for both the training and test sets. Then these ASAs are converted into relative accessibility which indicate relative difference ASA between complex and monomer state. Conservation of residues is found by Rate4Site (R4S) algorithm [81]. Contact potentials consider nonbonded interactions which have important role in the stabilization of proteins and complexes [82, 83]. These potentials can be extracted from frequencies of contacts for proteins with known 3D structures. For HotPoint algorithm, knowledge-based solvent-mediated inter-residue potentials are used [84]. To obtain the optimal model of the HotPoint algorithm, several empirical and machine learning methods are trained and tested (*see* Note 1).

The HotRegion server first predicts hot spots using the same algorithm as HotPoint. Following the hot spot prediction, a network of hot spots is constructed. Two hot spot residues are clustered together when the distance between their C_α is smaller than 6.5 Å [32]. This cutoff can be adjusted in “Advanced Search” (*see* Note 2). If the number of hot spots within the cluster is ≥3, the cluster is labeled as a hot region, and the hot spots within the cluster are members of this hot region. Other hot spots, which cannot be clustered within any hot region, can be called singlet hot spots. User can either provide a PDB ID or upload a homology-modeled PDB-formatted file; therefore users are not limited with the structures in PDB (*see* Note 3).

The following case from the literature explains how PIFACE [54], a nonredundant clustered protein-protein interface database, and the HotRegion server can be used to detect interface residues and hot spots on an interface [85]. This example also shows that drug binding sites are compatible with computationally predicted interfaces and hot spots. The human double minute 2 (Hdm2), like its mouse homolog (Mdm2), binds to the tumor suppressor p53 [86]. Therefore, the Hdm2 (and Mdm2) proteins are perfect drug targets to inhibit their binding to p53. It is known that drugs blocking this interaction enhance the tumor suppressor activity of p53 [87]. An experimental study identified three hot spots on p53 of Mdm2-p53 interface

1. For the trainings and testing empirical and machine learning methods, several features with different combinations such as relative ASA in complex and pair potentials, relative difference ASA and conservation, and relative ASA in complex and pair potentials were used. After several trials, an empirical model based on relative accessibility in complex state and total pair potentials gave the best performance. The thresholds to classify a residue as hot spot using this model are the relative ASA in complex state which is ≤20% and total contact potential which is ≥18.0; residues which are out of these thresholds are considered as non-hot spots.
2. 6.5 Å is the default cutoff and described as “Hotregion Neighbor Criteria.” These criteria can be modified in “Advanced Search” part of the HotRegion server [66]. As well as “Hotregion Neighbor Criteria,” users can decide a valid interface extraction threshold which is summed with van der Waals radii of atoms. HotRegion database provides pair potentials of interface residues, ASA and relative ASA values of interface residues of both monomer and complex forms of proteins. In “Advanced Search,” these properties are optionally printed in the result page.
3. User should provide atomic coordinates of the protein complexes in the standard PDB format. If atoms are present in alternative locations, only the first location is considered. For NMR structures, the first model is used. Since HotRegion is specific to protein-protein interfaces, chains corresponding to DNA and RNA structures return no interface solutions.

(Phe19, Trp23, and Leu26), which are also successfully predicted by HotRegion [87]. The Nutlin compound was identified as a strong inhibitor of the Mdm2-p53 complex through high-throughput screening (HTS) and medicinal chemistry methods [88]. To identify interface residues of the Mdm2-p53 complex (PDB ID: 1YCR, chain A and chain B, respectively), the "Interface Search Results" options from PIFACE server can be used. PDB ID and chains involved in interface should be given to server. Then, it can be directly reached to HotRegion server by choosing the interface name (1YCRAB). HotRegion gives information about interface residues, hot spots, and hot regions (Fig. 3a). Mdm2-p53 complex interface is identified by PIFACE, and hot spots on this complex are predicted by HotRegion (Fig. 3b). As well as experimentally identified p53 hot spots, Mdm2 hot spots (Leu57 and Ile61), which are complementary to p53 interface, were predicted. Comparison of this complex with the Mdm2-Nutlin complex (Fig. 3c) reveals that the Nutlin compounds occupy similar regions within the interface as the p53 side chains and these compounds bind to Mdm2 with a greater affinity than p53 [30].

3.2 Drug Target Prediction in PPI Networks

To analyze the protein interfaces on a network scale, Engin et al. [53] proposed a new representation for PPI networks, namely, Protein Interface and Interaction Network (P2IN), in which they marked nodes with interface structures. In this representation, the interactions are shown by edges between the interfaces. This representation has the advantage of showing different interfaces, which a protein pair uses to interact, and different protein pairs having similar interface structures, which may be the targets of a drug. Also, proteins competing to bind to a specific surface region are also detectable. Figure 4 shows a sample network using this representation.

Engin et al. [53] used this representation to simulate drug effects on the system level and find the side effects of drugs. For this purpose, they defined a new attack model in the networks named "interface attack." The interface attack simulates what a drug can do in PPI networks. Since a drug can bind to all proteins having the similar interface motifs and inhibit their interactions to their physiological partners, an interface attack removes edges between proteins having similar interface structures simultaneously. For example, if a drug is designed to inhibit p4-p7 interaction in Fig. 4, it can also inhibit the interactions p3-p5 and p3-p1.

To create a structural network including protein complexes and their corresponding interfaces, they used PRISM [89]. PRISM is a computational protein docking method which uses the known interface structures extracted from PDB [90] as templates to predict the binding of protein pairs (*see* Note 4). When PRISM predicts that two proteins can bind to each other, the template interface structure used for interaction is known and can be embedded to PPI networks. For some interactions, PRISM may find more than one interface which shows there are different binding modes between them. In these cases, all possible

4. PRISM gets a list of binary interactions, which can be gathered from literature and databases, as the input. The proteins' PDB IDs should be provided in the input list. So if there are more than one PDB structure for a specific protein, all those structures should be investigated. For each binary interaction, PRISM shows the binding interfaces, binding residues list, and binding free energy.

interactions are considered (*see* Note 5). The proteins are discarded if PRISM could not find any interaction between them.

Engin et al. [53] presented two case studies including the creation of the p53 interaction network, represented using P2IN, to find drug side effects. p53 is a tumor suppressor gene and it is a hub protein. p53 is involved in the cell cycle, DNA repair, and apoptosis [91]. p53 protein level is low in normal cells, and its overexpression is construed as a sign of many human cancers [92]. In more than 50% of human tumors, there are p53 mutants, mostly inactivated [93]. p53 interaction network consisted of 81 proteins and 251 interactions in which there were 46 different interface structures based on PRISM results. Among the results, there were two interactions for CDKN2D, with CDK4 and CDK6, which use a similar interface structure. Thus, if there are drugs that target one of these interactions, they may block the other interaction as their side effect. To check this idea, they used five different CDK6 inhibitors which block the G1/S transition of cell, i.e., aminopurvalanol [94], PD-0332991 [95], CHEBI: 792519 [96], CHEBI: 792520 [96], and fisetin [97]. They used AutoDock [98] for docking these drugs to CDK4 and CDK6. Interestingly, they found that these drugs can bind to CDK4 with comparable binding free energies to CDK6. There are studies showing common targets for CDK4 and CDK6 [99, 100]. Superpositioning of the docking results for CDK4 and CDK6 showed that there are several identical hot spots, i.e., Val9, Ala10, Ile12, Arg23, and Phe31 on their interfaces with CDKN2D which intensified their idea (Fig. 5). Therefore, they suggested that the drugs blocking the CDK6-CDKN2D interaction may also interrupt the CDK4-CDKN2D interaction.

The second case study compares interface attack with complete node attack. The Average Inverse Geodesic Length (AIGL) and the Giant Component Size (GCS) [101] are used to measure the robustness of the PPI network after the different attacks. Consecutive interface attacks and complete hub node attacks are performed on p53 P2IN. The complete node attack targets a hub node, which is known to be essential in PPI networks [102], and removes all its interactions simultaneously. Based on AIGL and GCS values, attacking the most frequent interfaces is as destructive as attacking the hub nodes in PPI networks. It should be noted that interface attacks are more realistic in comparison to complete node attacks because even if a drug is designed to target only one specific protein, it may not remove all its interactions at the same time.

4 Conclusions

Research on drug repositioning accelerated recently due to the increase in demand to new drugs, and experimental and computational repositioning strategies are being developed. Targeting similar PPI interfaces with common energetically important residues, which are hot spots and hot regions, is one of these strategies. Simulating drug effects on a network scale using binding site similarities have brought new insights in drug design. These methods can identify candidate off-targets of newly designed drugs and novel applications for existing drugs.

5. The predictions having binding free energies lower than −10 are accepted.

Besides the great benefits of PPI networks in drug repurposing, there are some serious restrictions in this area. One of the major limitations in structural PPI networks is the shortage of 3D structures of proteins and protein complexes [103, 104]. Even for the proteins which have available structures in the PDB, some of them have missing parts, and the structures are incomplete. Homology modeling is a powerful technique in predicting the protein structures. However, the accuracy of the binding sites would be under debate. Furthermore, as proteins dynamically change their conformation based on their environment and form new complexes, there is a need to integrate these information into PPI networks [105, 106]. These challenges will be addressed with the growth of the PDB in the coming years.

Acknowledgments

ESO acknowledges TUBITAK (The Scientific and Technological Research Council of Turkey) for financial support (Scholarship 2211-E). This project has been funded in whole or in part with federal funds from the Frederick National Laboratory for Cancer Research, National Institutes of Health, under contract HHSN261200800001E. This research was supported (in part) by the Intramural Research Program of NIH, Frederick National Lab, Center for Cancer Research. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the US Government.

References

1. Hopkins AL, Groom CR(2002) The druggable genome. Nat Rev Drug Discov 1(9):727–730. 10.1038/nrd892 [PubMed: 12209152]
2. Arkin MR, Wells JA (2004) Small-molecule inhibitors of protein-protein interactions: progressing towards the dream. Nat Rev Drug Discov 3(4):301–317. 10.1038/nrd1343 [PubMed: 15060526]
3. Acuner Ozbabacan SE, Engin HB, Gursoy A, Keskin O (2011) Transient protein-protein interactions. Protein Eng Des Sel 24 (9):635–648. 10.1093/protein/gzr025 [PubMed: 21676899]
4. Petukh M, Kucukkal TG, Alexov E (2015) On human disease-causing amino acid variants: statistical study of sequence and structural patterns. Hum Mutat 36(5):524–534. 10.1002/humu.22770 [PubMed: 25689729]
5. Schuster-Bockler B, Bateman A (2008) Protein interactions in human genetic diseases. Genome Biol 9(1):R9. 10.1186/Gb-2008-9-1-R9 [PubMed: 18199329]
6. Cavga AD, Karahan N, Keskin O, Gursoy A (2015) Taming oncogenic signaling at protein interfaces: challenges and opportunities. Curr Top Med Chem 15(20):2005–2018. 10.2174/1568026615666150519101956 [PubMed: 25986691]
7. Keskin O, Tuncbag N, Gursoy A (2016) Predicting protein-protein interactions from the molecular to the proteome level. Chem Rev 116(8):4884–4909. 10.1021/acs.chemrev.5b00683 [PubMed: 27074302]
8. Scott DE, Ehebauer MT, Pukala T, Marsh M, Blundell TL, Venkitaraman AR, Abell C, Hyvonen M (2013) Using a fragment-based approach to target protein-protein interactions. ChemBioChem 14(3):332–342. 10.1002/cbic.201200521 [PubMed: 23344974]
9. Thomas SE, Mendes V, Kim SY, Malhotra S, Ochoa-Montano B, Blaszczyk M, Blundell TL (2017) Structural biology and the design of new therapeutics: from HIV and cancer to mycobacterial infections: a paper dedicated to John Kendrew. J Mol Biol 429 (17):2677–2693. 10.1016/j.jmb.2017.06.014 [PubMed: 28648615]
10. Jubb HC, Pandurangan AP, Turner MA, Ochoa-Montano B, Blundell TL, Ascher DB (2016) Mutations at protein-protein interfaces: small changes over big surfaces have large impacts on human health. Prog Biophys Mol Biol 128:3. 10.1016/j.pbiomolbio.2016.10.002 [PubMed: 27913149]
11. Aldeghi M, Malhotra S, Selwood DL, Chan AW (2014) Two- and three-dimensional rings in drugs. Chem Biol Drug Des 83 (4):450–461. 10.1111/cbdd.12260 [PubMed: 24472495]

12. Keskin O, Nussinov R (2005) Favorable scaffolds: proteins with different sequence, structure and function may associate in similar ways. Protein Eng Des Sel 18(1):11–24. 10.1093/protein/gzh095 [PubMed: 15790576]
13. Tuncbag N, Kar G, Gursoy A, Keskin O, Nussinov R (2009) Towards inferring time dimensionality in protein-protein interaction networks by integrating structures: the p53 example. Mol Biosyst 5(12):1770–1778. 10.1039/b905661k [PubMed: 19585003]
14. Medina-Franco JL, Giulianotti MA, Welmaker GS, Houghten RA (2013) Shifting from the single to the multitarget paradigm in drug discovery. Drug Discov Today 18 (9-10):495–501. 10.1016/j.drudis.2013.01.008 [PubMed: 23340113]
15. Paolini GV, Shapland RH, van Hoorn WP, Mason JS, Hopkins AL (2006) Global mapping of pharmacological space. Nat Biotechnol 24(7):805–815. 10.1038/nbt1228 [PubMed: 16841068]
16. Overington JP, Al-Lazikani B, Hopkins AL (2006) How many drug targets are there? Nat Rev Drug Discov 5(12):993–996. 10.1038/nrd2199 [PubMed: 17139284]
17. Zhou H, Gao M, Skolnick J (2015) Comprehensive prediction of drug-protein interactions and side effects for the human proteome. Sci Rep 5:11090. 10.1038/srep11090 [PubMed: 26057345]
18. Scott DE, Bayly AR, Abell C, Skidmore J (2016) Small molecules, big targets: drug discovery faces the protein-protein interaction challenge. Nat Rev Drug Discov 15 (8):533–550. 10.1038/nrd.2016.29 [PubMed: 27050677]
19. Gurung AB, Bhattacharjee A, Ali MA, Al-Hemaid F, Lee J (2017) Binding of small molecules at interface of protein-protein complex - a newer approach to rational drug design. Saudi J Biol Sci 24(2):379–388. 10.1016/j.sjbs.2016.01.008 [PubMed: 28149177]
20. Xie L, Li J, Xie L, Bourne PE (2009) Drug discovery using chemical systems biology: identification of the protein-ligand binding network to explain the side effects of CETP inhibitors. PLoS Comput Biol 5(5):e1000387. 10.1371/journal.pcbi.1000387 [PubMed: 19436720]
21. Xie L, Bourne PE (2008) Detecting evolutionary relationships across existing fold space, using sequence order-independent profile-profile alignments. Proc Natl Acad Sci U S A 105(14):5441–5446. 10.1073/pnas.0704422105 [PubMed: 18385384]
22. Duran-Frigola M, Siragusa L, Ruppin E, Barril X, Cruciani G, Aloy P (2017) Detecting similar binding pockets to enable systems polypharmacology. PLoS Comput Biol 13 (6):e1005522. 10.1371/journal.pcbi.1005522 [PubMed: 28662117]
23. Siragusa L, Cross S, Baroni M, Goracci L, Cruciani G (2015) BioGPS: navigating biological space to predict polypharmacology, off-targeting, and selectivity. Proteins 83 (3):517–532. 10.1002/prot.24753 [PubMed: 25556939]
24. Fry DC (2006) Protein-protein interactions as targets for small molecule drug discovery. Biopolymers 84(6):535–552. 10.1002/bip.20608 [PubMed: 17009316]
25. Arkin MR, Randal M, DeLano WL, Hyde J, Luong TN, Oslob JD, Raphael DR, Taylor L, Wang J, McDowell RS, Wells JA, Braisted AC (2003) Binding of small molecules to an adaptive protein-protein interface. Proc Natl Acad Sci U S A 100(4):1603–1608. 10.1073/pnas.252756299 [PubMed: 12582206]
26. Li J, Zheng S, Chen B, Butte AJ, Swamidass SJ, Lu Z (2016) A survey of current trends in computational drug repositioning. Brief Bioinform 17(1):2–12. 10.1093/bib/bbv020 [PubMed: 25832646]
27. Choi SH, Choi KY (2017) Screening-based approaches to identify small molecules that inhibit protein-protein interactions. Expert Opin Drug Discovery 12(3):293–303. 10.1080/17460441.2017.1280456
28. Clackson T, Wells JA (1995) A hot-spot of binding-energy in a hormone-receptor interface. Science 267(5196):383–386. 10.1126/science.7529940 [PubMed: 7529940]
29. Bogan AA, Thorn KS (1998) Anatomy of hot spots in protein interfaces. J Mol Biol 280 (1):1–9. 10.1006/jmbi.1998.1843 [PubMed: 9653027]
30. Wells JA, McClendon CL (2007) Reaching for high-hanging fruit in drug discovery at protein-protein interfaces. Nature 450 (7172):1001–1009. 10.1038/nature06526 [PubMed: 18075579]
31. Thangudu RR, Bryant SH, Panchenko AR, Madej T (2012) Modulating protein-protein interactions with small molecules: the importance of binding hotspots. J Mol Biol 415 (1):443–453. 10.1016/j.jmb.2011.12.026 [PubMed: 22198293]

32. Keskin O, Ma B, Nussinov R (2005) Hot regions in protein--protein interactions: the organization and contribution of structurally conserved hot spot residues. J Mol Biol 345 (4):1281–1294. 10.1016/j.jmb.2004.10.077 [PubMed: 15644221]
33. Thorn KS, Bogan AA (2001) ASEdb: a database of alanine mutations and their effects on the free energy of binding in protein interactions. Bioinformatics 17(3):284–285. 10.1093/bioinformatics/17.3.284 [PubMed: 11294795]
34. Rosell M, Fernandez-Recio J (2018) Hot-spot analysis for drug discovery targeting protein-protein interactions. Expert Opin Drug Discovery 13:327–338. 10.1080/17460441.2018.1430763
35. Ofran Y, Rost B (2007) Protein-protein interaction hotspots carved into sequences. PLoS Comput Biol 3(7):e119. 10.1371/journal.pcbi.0030119 [PubMed: 17630824]
36. Grosdidier S, Fernandez-Recio J (2008) Identification of hot-spot residues in protein-protein interactions by computational docking. BMC Bioinformatics 9:447. 10.1186/1471-2105-9-447 [PubMed: 18939967]
37. Ozbek P, Soner S, Haliloglu T (2013) Hot spots in a network of functional sites. PLoS One 8(9):e74320. 10.1371/journal.pone.0074320 [PubMed: 24023934]
38. Agrawal NJ, Helk B, Trout BL (2014) A computational tool to predict the evolutionarily conserved protein-protein interaction hot-spot residues from the structure of the unbound protein. FEBS Lett 588 (2):326–333. 10.1016/j.febslet.2013.11.004 [PubMed: 24239538]
39. Kortemme T, Baker D (2002) A simple physical model for binding energy hot spots in protein-protein complexes. Proc Natl Acad Sci U S A 99(22):14116–14121. 10.1073/pnas.202485799 [PubMed: 12381794]
40. Guerois R, Nielsen JE, Serrano L (2002) Predicting changes in the stability of proteins and protein complexes: a study of more than 1000 mutations. J Mol Biol 320(2):369–387. 10.1016/S0022-2836(02)00442-4 [PubMed: 12079393]
41. Li MH, Simonetti FL, Goncearenco A, Panchenko AR (2016) MutaBind estimates and interprets the effects of sequence variants on protein-protein interactions. Nucleic Acids Res 44(W1):W494–W501. 10.1093/nar/gkw374 [PubMed: 27150810]
42. Gao Y, Wang R, Lai L (2004) Structure-based method for analyzing protein-protein interfaces. J Mol Model 10(1):44–54. 10.1007/s00894-003-0168-3 [PubMed: 14634848]
43. Tuncbag N, Gursoy A, Keskin O (2009) Identification of computational hot spots in protein interfaces: combining solvent accessibility and inter-residue potentials improves the accuracy. Bioinformatics 25 (12):1513–1520. 10.1093/bioinformatics/btp240 [PubMed: 19357097]
44. Gonzalez-Ruiz D, Gohlke H (2006) Targeting protein-protein interactions with small molecules: challenges and perspectives for computational binding epitope detection and ligand finding. Curr Med Chem 13 (22):2607–2625. 10.2174/092986706778201530 [PubMed: 17017914]
45. Rajamani D, Thiel S, Vajda S, Camacho CJ (2004) Anchor residues in protein-protein interactions. Proc Natl Acad Sci U S A 101 (31):11287–11292. 10.1073/pnas.0401942101 [PubMed: 15269345]
46. Burgoyne NJ, Jackson RM (2006) Predicting protein interaction sites: binding hot-spots in protein-protein and protein-ligand interfaces. Bioinformatics 22(11):1335–1342. 10.1093/bioinformatics/bt/079 [PubMed: 16522669]
47. Darnell SJ, LeGault L, Mitchell JC (2008) KFC Server: interactive forecasting of protein interaction hot spots. Nucleic Acids Res 36:W265–W269. 10.1093/nar/gkn346 [PubMed: 18539611]
48. Zhu XL, Mitchell JC (2011) KFC2: a knowledge-based hot spot prediction method based on interface solvation, atomic density, and plasticity features. Prot Struct Funct Bioinf 79(9):2671–2683. 10.1002/prot.23094
49. Wang LC, Hou YQ, Quan HH, Xu WW, Bao YL, Li YX, Fu Y, Zou SX (2013) A compound-based computational approach for the accurate determination of hot spots. Protein Sci 22(8):1060–1070. 10.1002/pro.2296 [PubMed: 23776011]
50. Bonetta L (2010) Protein-protein interactions: interactome under construction. Nature 468(7325):851–854. 10.1038/468851a [PubMed: 21150998]

51. Gao M, Skolnick J (2010) Structural space of protein-protein interfaces is degenerate, close to complete, and highly connected. Proc Natl Acad Sci U S A 107(52):22517–22522. 10.1073/pnas.1012820107 [PubMed: 21149688]
52. Haupt VJ, Daminelli S, Schroeder M (2013) Drug promiscuity in PDB: protein binding site similarity is key. PLoS One 8(6):e65894. 10.1371/journal.pone.0065894 [PubMed: 23805191]
53. Engin HB, Keskin O, Nussinov R, Gursoy A (2012) A strategy based on protein-protein interface motifs may help in identifying drug off-targets. J Chem Inf Model 52 (8):2273–2286. 10.1021/ci300072q [PubMed: 22817115]
54. Cukuroglu E, Gursoy A, Nussinov R, Keskin O (2014) Non-redundant unique interface structures as templates for modeling protein interactions. PLoS One 9(1):e86738. 10.1371/journal.pone.0086738 [PubMed: 24475173]
55. Aloy P, Russell RB (2004) Ten thousand interactions for the molecular biologist. Nat Biotechnol 22(10):1317–1321. 10.1038/nbt1018 [PubMed: 15470473]
56. Tyagi M, Thangudu RR, Zhang D, Bryant SH, Madej T, Panchenko AR (2012) Homology inference of protein-protein interactions via conserved binding sites. PLoS One 7(1): e28896. 10.1371/journal.pone.0028896 [PubMed: 22303436]
57. De S, Krishnadev O, Srinivasan N, Rekha N (2005) Interaction preferences across protein-protein interfaces of obligatory and non-obligatory components are different. BMC Struct Biol 5:15. 10.1186/1472-6807-5-15 [PubMed: 16105176]
58. Keseru GM, Erlanson DA, Ferenczy GG, Hann MM, Murray CW, Pickett SD (2016) Design principles for fragment libraries: maximizing the value of learnings from pharma fragment-based drug discovery (FBDD) programs for use in academia. J Med Chem 59 (18):8189–8206. 10.1021/acs.jmedchem.6b00197 [PubMed: 27124799]
59. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000) The protein data bank. Nucleic Acids Res 28(1):235–242 [PubMed: 10592235]
60. Anand P, Nagarajan D, Mukherjee S, Chandra N (2014) PLIC: protein-ligand interaction clusters. Database (Oxford) 2014:bau029. 10.1093/database/bau029 [PubMed: 24763918]
61. Yeturu K, Chandra N (2008) PocketMatch: a new algorithm to compare binding sites in protein structures. BMC Bioinformatics 9:543. 10.1186/1471-2105-9-543 [PubMed: 19091072]
62. Desaphy J, Bret G, Rognan D, Kellenberger E (2015) sc-PDB: a 3D-database of ligandable binding sites--10 years on. Nucleic Acids Res 43(Database issue):D399–D404. 10.1093/nar/gku928 [PubMed: 25300483]
63. Xu Q, Dunbrack RL Jr (2011) The protein common interface database (ProtCID)--a comprehensive database of interactions of homologous proteins in multiple crystal forms. Nucleic Acids Res 39(Database issue): D761–D770. 10.1093/nar/gkq1059 [PubMed: 21036862]
64. Mosca R, Ceol A, Stein A, Olivella R, Aloy P (2014) 3did: a catalog of domain-based interactions of known three-dimensional structure. Nucleic Acids Res 42(D1):D374–D379. 10.1093/nar/gkt887 [PubMed: 24081580]
65. Tuncbag N, Keskin O, Gursoy A (2010) HotPoint: hot spot prediction server for protein interfaces. Nucleic Acids Res 38: W402–W406. 10.1093/nar/gkq323 [PubMed: 20444871]
66. Cukuroglu E, Gursoy A, Keskin O (2012) HotRegion: a database of predicted hot spot clusters. Nucleic Acids Res 40(D1):D829–D833. 10.1093/nar/gkr929 [PubMed: 22080558]
67. Guharoy M, Chakrabarti P (2005) Conservation and relative importance of residues across protein-protein interfaces. Proc Natl Acad Sci U S A 102(43):15447–15452. 10.1073/pnas.0505425102 [PubMed: 16221766]
68. Moreira IS, Fernandes PA, Ramos MJ (2007) Hot spots--a review of the protein-protein interface determinant amino-acid residues. Proteins 68(4):803–812. 10.1002/prot.21396 [PubMed: 17546660]
69. Cukuroglu E, Gursoy A, Keskin O (2010) Analysis of hot region organization in hub proteins. Ann Biomed Eng 38(4):2068–2078. 10.1007/s10439-010-0048-9 [PubMed: 20437205]
70. Hajduk PJ, Huth JR, Fesik SW (2005) Druggability indices for protein targets derived from NMR-based screening data. J Med Chem 48(7):2518–2525. 10.1021/jm049131r [PubMed: 15801841]

71. Ozdemir ES, Jang H, Gursoy A, Keskin O, Li Z, Sacks DB, Nussinov R (2018) Unraveling the molecular mechanism of interactions of the Rho GTPases Cdc42 and Rac1 with the scaffolding protein IQGAP2. J Biol Chem 293:3685. 10.1074/jbc.RA117.001596 [PubMed: 29358323]
72. Hall DR, Kozakov D, Whitty A, Vajda S (2015) Lessons from hot spot analysis for fragment-based drug discovery. Trends Pharmacol Sci 36(11):724–736. 10.1016/j.tips.2015.08.003 [PubMed: 26538314]
73. Li X, Keskin O, Ma BY, Nussinov R, Liang J (2004) Protein-protein interactions: hot spots and structurally conserved residues often locate in complemented pockets that pre-organized in the unbound states: implications for docking. J Mol Biol 344 (3):781–795. 10.1016/j.jmb.2004.09.051 [PubMed: 15533445]
74. Tsao DHH, Sutherland AG, Jennings LD, Li YH, Rush TS, Alvarez JC, Ding WD, Dushin EG, Dushin RG, Haney SA, Kenny CH, Malakian AK, Nilakantan R, Mosyak L (2006) Discovery of novel inhibitors of the ZipA/FtsZ complex by NMR fragment screening coupled with structure-based design. Bioorg Med Chem 14 (23):7953–7961. 10.1016/j.bmc.2006.07.050 [PubMed: 16919463]
75. Fang L, Zhu Q, Neuenschwander M, Specker E, Wulf-Goldenberg A, Weis WI, von Kries JP, Birchmeier W (2016) A small-molecule antagonist of the beta-catenin/TCF4 interaction blocks the self-renewal of cancer stem cells and suppresses tumorigenesis. Cancer Res 76(4):891–901. 10.1158/0008-5472.CAN-15-1519 [PubMed: 26645562]
76. Shin WH, Christoffer CW, Kihara D (2017) In silico structure-based approaches to discover protein-protein interaction-targeting drugs. Methods 131:22–32. 10.1016/j.ymeth.2017.08.006 [PubMed: 28802714]
77. Keskin O, Tsai CJ, Wolfson H, Nussinov R (2004) A new, structurally nonredundant, diverse data set of protein-protein interfaces and its implications. Protein Sci 13 (4):1043–1055. 10.1110/ps.03484604 [PubMed: 15044734]
78. Tuncbag N, Gursoy A, Guney E, Nussinov R, Keskin O (2008) Architectures and functional coverage of protein-protein interfaces. J Mol Biol 381(3):785–802. 10.1016/j.jmb.2008.04.071 [PubMed: 18620705]
79. Fischer TB, Arunachalam KV, Bailey D, Mangual V, Bakhru S, Russo R, Huang D, Paczkowski M, Lalchandani V, Ramachandra C, Ellison B, Galer S, Shapley J, Fuentes E, Tsai J (2003) The binding interface database (BID): a compilation of amino acid hot spots in protein interfaces. Bioinformatics 19(11):1453–1454. 10.1093/bioinformatics/btg163 [PubMed: 12874065]
80. Hubbard SJ, Thornton J (1993) Naccess version 2.1.1. Computer program. Department of Biochemistry and Molecular Biology, University College London
81. Pupko T, Bell RE, Mayrose I, Glaser F, Ben-Tal N (2002) Rate4Site: an algorithmic tool for the identification of functional regions in proteins by surface mapping of evolutionary determinants within their homologues. Bioinformatics 18(Suppl 1):S71–S77 [PubMed: 12169533]
82. Jernigan RL, Bahar I (1996) Structure-derived potentials and protein simulations. Curr Opin Struct Biol 6(2):195–209. 10.1016/S0959-440x(96)80075-3 [PubMed: 8728652]
83. Godzik A, Skolnick J (1992) Sequence structure matching in globular-proteins - application to supersecondary and tertiary structure determination. Proc Natl Acad Sci U S A 89(24):12098–12102. 10.1073/pnas.89.24.12098 [PubMed: 1465445]
84. Keskin O, Bahar I, Badretdinov AY, Ptitsyn OB, Jernigan RL (1998) Empirical solvent-mediated potentials hold for both intra-molecular and inter-molecular inter-residue interactions. Protein Sci 7(12):2578–2586. 10.1002/pro.5560071211 [PubMed: 9865952]
85. Acuner Ozbabacan SE, Gursoy A, Keskin O, Nussinov R (2010) Conformational ensembles, signal transduction and residue hot spots: application to drug discovery. Curr Opin Drug Discov Devel 13(5):527–537
86. Levine AJ, Hu W, Feng Z (2006) The P53 pathway: what questions remain to be explored? Cell Death Differ 13 (6):1027–1036. 10.1038/sj.cdd.4401910 [PubMed: 16557269]
87. Picksley SM, Vojtesek B, Sparks A, Lane DP (1994) Immunochemical analysis of the interaction of p53 with MDM2;--fine mapping of the MDM2 binding site on p53 using synthetic peptides. Oncogene 9(9):2523–2529 [PubMed: 8058315]

88. Vassilev LT, Vu BT, Graves B, Carvajal D, Podlaski F, Filipovic Z, Kong N, Kammlott U, Lukacs C, Klein C, Fotouhi N, Liu EA (2004) In vivo activation of the p53 pathway by small-molecule antagonists of MDM2. Science 303(5659):844–848. 10.1126/science.1092472 [PubMed: 14704432]

89. Tuncbag N, Gursoy A, Nussinov R, Keskin O (2011) Predicting protein-protein interactions on a proteome scale by matching evolutionary and structural similarities at interfaces using PRISM. Nat Protoc 6(9):1341–1354. 10.1038/nprot.2011.367 [PubMed: 21886100]

90. Sussman JL, Lin D, Jiang J, Manning NO, Prilusky J, Ritter O, Abola EE (1998) Protein Data Bank (PDB): database of three-dimensional structural information of biological macromolecules. Acta Crystallogr D Biol Crystallogr 54(Pt 6 Pt 1):1078–1084 [PubMed: 10089483]

91. Aylon Y, Oren M (2011) New plays in the p53 theater. Curr Opin Genet Dev 21(1):86–92. 10.1016/j.gde.2010.10.002 [PubMed: 21317061]

92. Levine AJ, Oren M (2009) The first 30 years of p53: growing ever more complex. Nat Rev Cancer 9(10):749–758. 10.1038/nrc2723 [PubMed: 19776744]

93. Rivlin N, Brosh R, Oren M, Rotter V (2011) Mutations in the p53 tumor suppressor gene: important milestones at the various steps of tumorigenesis. Genes Cancer 2(4):466–474. 10.1177/1947601911408889 [PubMed: 21779514]

94. Lu H, Schulze-Gahmen U (2006) Toward understanding the structural basis of cyclin-dependent kinase 6 specific inhibition. J Med Chem 49(13):3826–3831. 10.1021/jm0600388 [PubMed: 16789739]

95. Baughn LB, Di Liberto M, Wu K, Toogood PL, Louie T, Gottschalk R, Niesvizky R, Cho H, Ely S, Moore MA, Chen-Kiang S (2006) A novel orally active small molecule potently induces G1 arrest in primary myeloma cells and prevents tumor growth by specific inhibition of cyclin-dependent kinase 4/6. Cancer Res 66(15):7661–7667. 10.1158/0008-5472.CAN-06-1098 [PubMed: 16885367]

96. Cho YS, Borland M, Brain C, Chen CHT, Cheng H, Chopra R, Chung K, Groarke J, He G, Hou Y, Kim S, Kovats S, Lu YP, O'Reilly M, Shen JQ, Smith T, Trakshel G, Vogtle M, Xu M, Xu M, Sung MJ (2010) 4-(Pyrazol-4-yl)-pyrimidines as selective inhibitors of cyclin-dependent kinase 4/6. J Med Chem 53(22):7938–7957. 10.1021/jm100571n [PubMed: 21038853]

97. Lu H, Chang DJ, Baratte B, Meijer L, Schulze-Gahmen U (2005) Crystal structure of a human cyclin-dependent kinase 6 complex with a flavonol inhibitor, fisetin. J Med Chem 48(3):737–743. 10.1021/jm049353p [PubMed: 15689157]

98. Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, Olson AJ (2009) AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. J Comput Chem 30 (16):2785–2791. 10.1002/jcc.21256 [PubMed: 19399780]

99. Fry DW, Harvey PJ, Keller PR, Elliott WL, Meade M, Trachet E, Albassam M, Zheng X, Leopold WR, Pryer NK, Toogood PL (2004) Specific inhibition of cyclin-dependent kinase 4/6 by PD 0332991 and associated antitumor activity in human tumor xenografts. Mol Cancer Ther 3(11):1427–1438 [PubMed: 15542782]

100. Gunther S, Kuhn M, Dunkel M, Campillos M, Senger C, Petsalaki E, Ahmed J, Urdiales EG, Gewiess A, Jensen LJ, Schneider R, Skoblo R, Russell RB, Bourne PE, Bork P, Preissner R (2008) SuperTarget and Matador: resources for exploring drug-target relationships. Nucleic Acids Res 36(Database issue):D919–D922. 10.1093/nar/gkm862 [PubMed: 17942422]

101. Holme P, Kim BJ, Yoon CN, Han SK(2002) Attack vulnerability of complex networks. Phys Rev E Stat Nonlin Soft Matter Phys 65 (5 Pt 2):056109. 10.1103/PhysRevE.65.056109 [PubMed: 12059649]

102. He X, Zhang J (2006) Why do hubs tend to be essential in protein networks? PLoS Genet 2(6):e88. 10.1371/journal.pgen.0020088 [PubMed: 16751849]

103. Vakser IA (2014) Protein-protein docking: from interaction to interactome. Biophys J 107(8):1785–1793. 10.1016/j.bpj.2014.08.033 [PubMed: 25418159]

104. Szilagyi A, Zhang Y (2014) Template-based structure modeling of protein-protein interactions. Curr Opin Struct Biol 24:10–23. 10.1016/j.sbi.2013.11.005 [PubMed: 24721449]

105. Halakou F, Kilic ES, Cukuroglu E, Keskin O, Gursoy A (2017) Enriching traditional protein-protein interaction networks with alternative conformations of proteins. Sci Rep 7(1):7180. 10.1038/s41598-017-07351-0 [PubMed: 28775330]
106. Ozgur B, Ozdemir ES, Gursoy A, Keskin O (2017) Relation between protein intrinsic normal mode weights and pre-existing conformer populations. J Phys Chem B 121 (15):3686–3700. 10.1021/acs.jpcb.6b10401 [PubMed: 28181436]
107. Kussie PH, Gorina S, Marechal V, Elenbaas B, Moreau J, Levine AJ, Pavletich NP (1996) Structure of the MDM2 oncoprotein bound to the p53 tumor suppressor transactivation domain. Science 274(5289):948–953 [PubMed: 8875929]
108. Kim DE, Chivian D, Baker D (2004) Protein structure prediction and analysis using the Robetta server. Nucleic Acids Res 32(Web Server):W526–W531. 10.1093/nar/gkh468 [PubMed: 15215442]
109. Xia JF, Zhao XM, Song J, Huang DS (2010) APIS: accurate prediction of hot spots in protein interfaces by combining protrusion index with solvent accessibility. BMC Bioinformatics 11:174. 10.1186/1471-2105-11-174 [PubMed: 20377884]
110. Deng L, Zhang QC, Chen ZG, Meng Y, Guan JH, Zhou SG (2014) PredHS: a web server for predicting protein-protein interaction hot spots by using structural neighborhood properties. Nucleic Acids Res 42(W1): W290–W295. 10.1093/nar/gku437 [PubMed: 24852252]
111. Shulman-Peleg A, Shatsky M, Nussinov R, Wolfson HJ (2007) Spatial chemical conservation of hot spot interactions in protein-protein complexes. BMC Biol 5:43. 10.1186/1741-7007-5-43 [PubMed: 17925020]
112. Meireles LMC, Domling AS, Camacho CJ (2010) ANCHOR: a web server and database for analysis of protein-protein interaction binding pockets for drug discovery. Nucleic Acids Res 38:W407–W411. 10.1093/nar/gkq502 [PubMed: 20525787]
113. Assi SA, Tanaka T, Rabbitts TH, Fernandez-Fuentes N (2010) PCRPi: presaging critical residues in protein interfaces, a new computational tool to chart hot spots in protein interfaces. Nucleic Acids Res 38(6):e86. 10.1093/nar/gkp1158 [PubMed: 20008102]
114. Jimenez J, Doerr S, Martinez-Rosell G, Rose AS, De Fabritiis G (2017) DeepSite: protein-binding site predictor using 3D-convolutional neural networks. Bioinformatics 33(19):3036–3042. 10.1093/bioinformatics/btx350 [PubMed: 28575181]
115. Greener JG, Sternberg MJE (2015) AlloPred: prediction of allosteric pockets on proteins using normal mode perturbation analysis. BMC Bioinformatics 16:335. 10.1186/s12859-015-0771-1 [PubMed: 26493317]
116. Borrel A, Regad L, Xhaard H, Petitjean M, Camproux AC (2015) PockDrug: a model for predicting pocket druggability that overcomes pocket estimation uncertainties. J Chem Inf Model 55(4):882–895. 10.1021/ci5006004 [PubMed: 25835082]
117. Huang BD, Schroeder M (2006) LIGSITE (csc): predicting ligand binding sites using the Connolly surface and degree of conservation. BMC Struct Biol 6:19. 10.1186/1472-6807-6-19 [PubMed: 16995956]
118. Zhang ZM, Li Y, Lin BY, Schroeder M, Huang BD (2011) Identification of cavities on protein surface using multiple computational approaches for drug binding site prediction. Bioinformatics 27(15):2083–2088. 10.1093/bioinformatics/btr331 [PubMed: 21636590]
119. Yu J, Zhou Y, Tanaka I, Yao M (2010) Roll: a new algorithm for the detection of protein pockets and cavities with a rolling probe sphere. Bioinformatics 26(1):46–52. 10.1093/bioinformatics/btp599 [PubMed: 19846440]

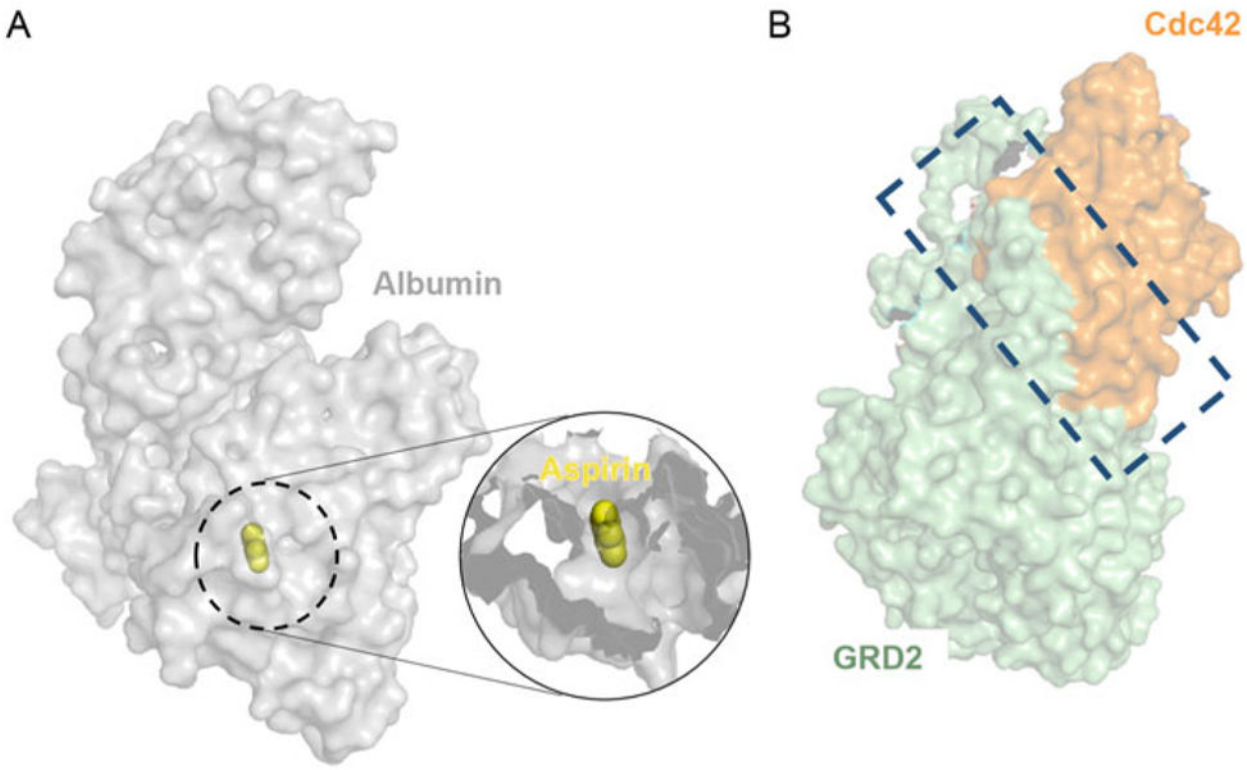

Fig. 1.

(**a**) Small drug binding pocket. Aspirin (yellow) binds to small binding cavity of albumin (gray). (**b**) Large and flat PPI interfaces. Flat binding interface of Cdc42 (orange) and GRD2 (green)

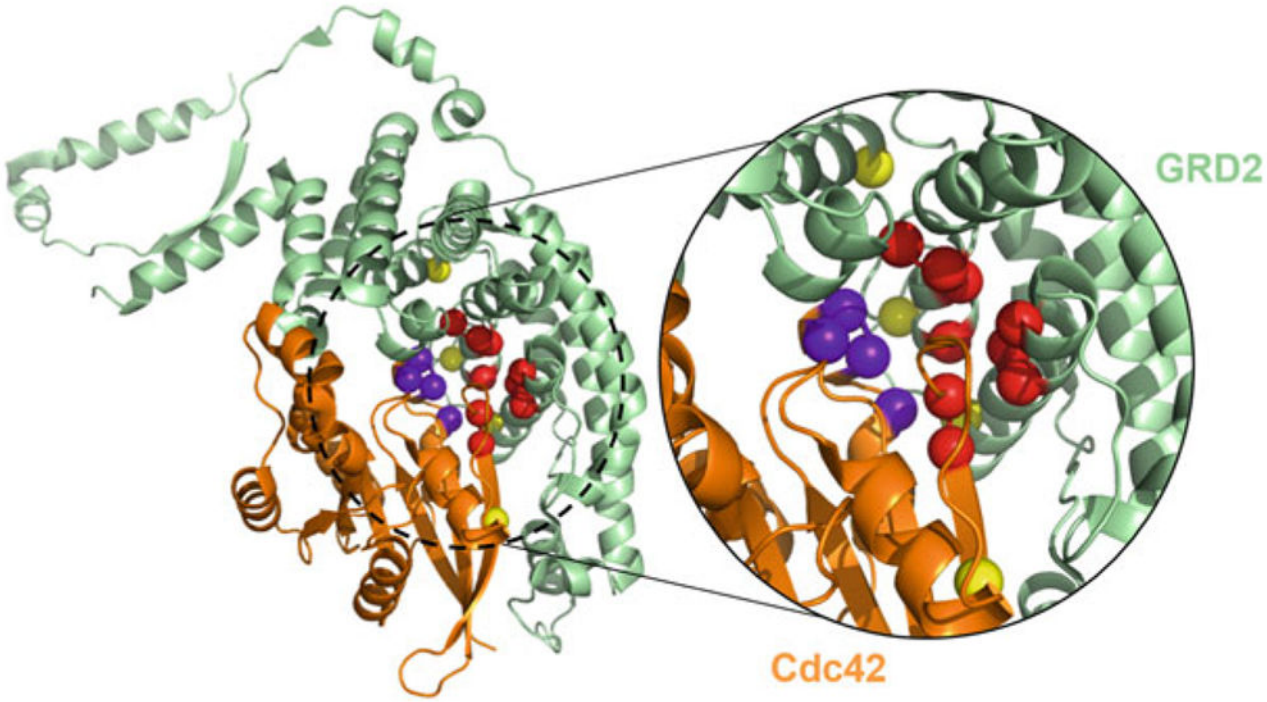

Fig. 2.

Hot spots and hot regions located in the interface of a small GTPase. Purple and red balls represent first and second hot regions, respectively. Yellow ball is a hot spot residue that is not included in any hot region

A

Interface Name	Residue Number	Residue Type	Chain	Relative Complex ASA	Relative Monomer ASA	Pair Potential	Hotspot Status	Hotregion Status	Complex ASA	Monomer ASA
1YCRAB	25	GLU	A	77.89	89.32	7.6	NH		134.17	153.86
1YCRAB	26	THR	A	34.64	52.78	7.5	NH		48.24	73.5
1YCRAB	50	MET	A	0	10.42	36.36	H	1	0	20.23
1YCRAB	51	LYS	A	42.63	67.16	7.29	NH		85.61	134.87
1YCRAB	54	LEU	A	7.98	46.68	30.22	H	1	14.25	83.39
1YCRAB	57	LEU	A	0.12	4.07	68.22	H	1	0.22	7.27
1YCRAB	58	GLY	A	0.37	21.99	10.55	NH		0.3	17.61
1YCRAB	61	ILE	A	0	12.91	48.64	H	-	0	22.61
1YCRAB	62	MET	A	26.43	59.66	19.3	NH		51.31	115.83
1YCRAB	67	TYR	A	18.72	24.08	15.09	NH		39.83	51.23
1YCRAB	72	GLN	A	18.31	58.93	8.19	NH		32.69	105.19
1YCRAB	73	HIS	A	14.87	23.55	17.4	NH		27.19	43.06
1YCRAB	75	VAL	A	0	1.08	32.89	H	-	0	1.64
1YCRAB	93	VAL	A	0.15	42.54	41.71	H	0	0.22	64.43
1YCRAB	94	LYS	A	53.64	68.89	10.77	NH		107.71	138.34
1YCRAB	96	HIS	A	36.87	77.13	11.5	NH		67.42	141.06
1YCRAB	100	TYR	A	25.87	50.82	12.51	NH		55.04	108.13
1YCRAB	17	GLU	B	59.84	99.8	13.78	NH		103.08	171.9
1YCRAB	18	THR	B	49.06	59.56	7.72	NH		68.32	82.95
1YCRAB	19	PHE	B	1.53	71.33	37.16	H	0	3.05	142.28
1YCRAB	20	SER	B	24.72	46.67	8.3	NH		28.8	54.37
1YCRAB	22	LEU	B	9.92	42.8	22.75	H	0	17.72	76.46
1YCRAB	23	TRP	B	4.6	59.65	38.57	H	0	11.48	148.74
1YCRAB	25	LEU	B	67.8	82.27	4.85	NH		121.12	146.95
1YCRAB	26	LEU	B	4.93	47.48	31.38	H	0	8.81	84.81
1YCRAB	27	PRO	B	39.98	90.59	8.47	NH		54.43	123.32
1YCRAB	28	GLU	B	55.17	92.9	2.83	NH		95.03	160.02
1YCRAB	29	ASN	B	49.61	136.89	8.16	NH		71.41	197.04

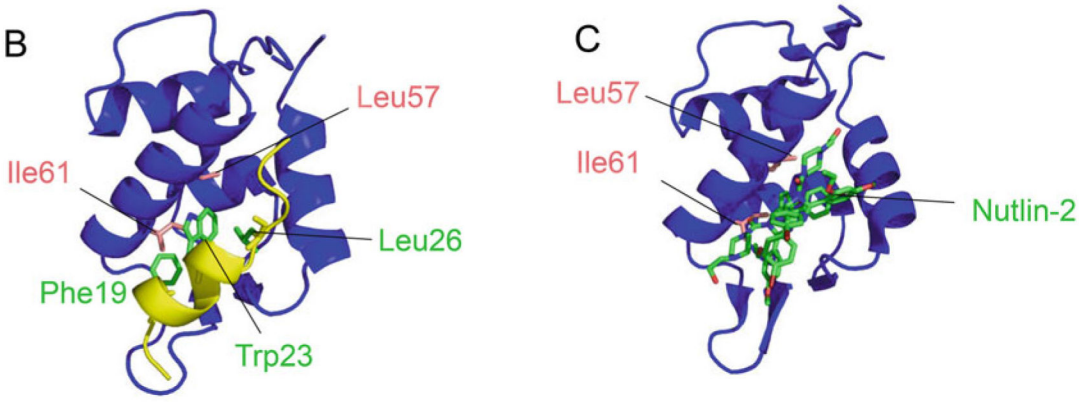

Fig. 3.

Interface, hot spot, and hot region residues of Mdm2-p53 complex. (**a**) The residues listed in HotRegion are interface residues. Hot spots and hot regions can be identified from “Hotspot Status” and “Hotregion Status” columns. (**b**) The structure (PDB identifier: 1YCR) of a complex between Mdm2- (blue) and a p53-derived peptide (yellow) [107]. Pink and green sticks represent hot spots, which also correspond to Nutlin binding site, of Mdm2 and the p53-derived peptide, respectively. (**c**) The structure (PDB identifier: 1RV1) of a complex between Mdm2 (blue) and a Nutlin-2 (green) [88]. Pink sticks represent hot spots of Mdm2. The hot spots of the p53-derived peptide (Phe19, Trp23, and Leu26) were determined experimentally [87], whereas the hot spots for Mdm2 (Leu57 and Ile61) were predicted by HotRegion

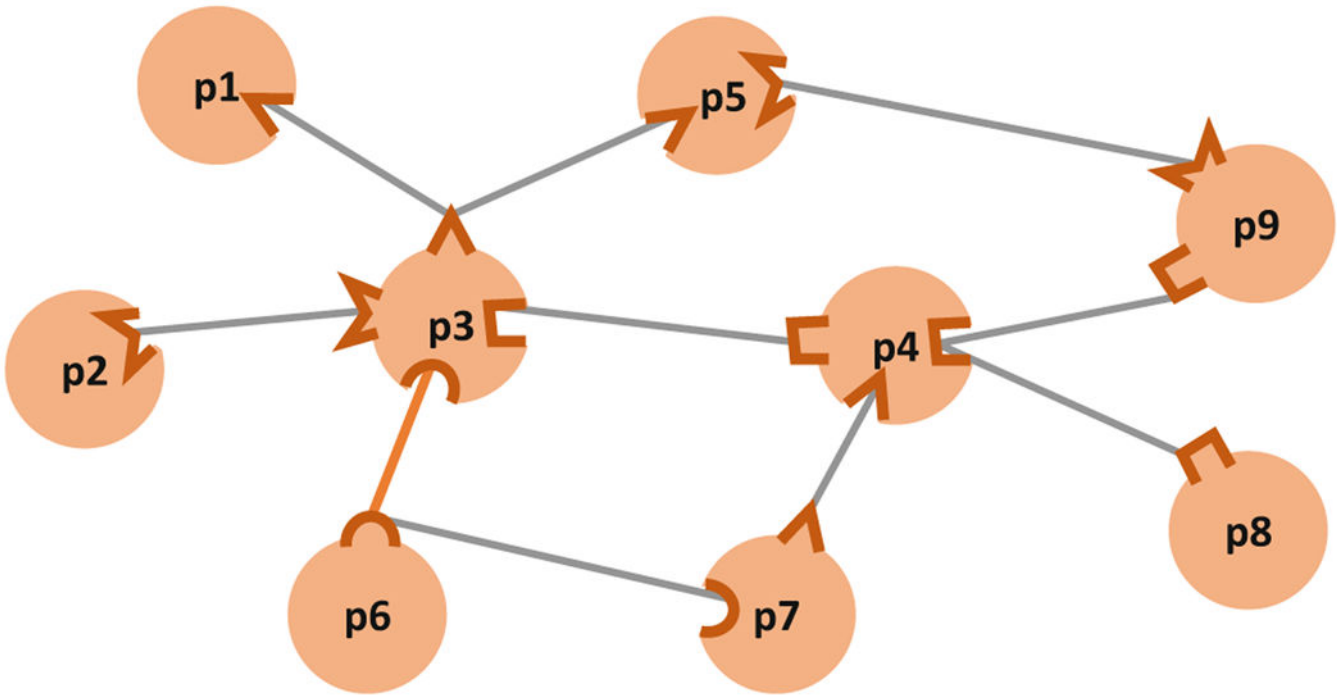

Fig. 4.

A sample protein-protein interaction network using P2IN representation. Protein interfaces are shown in dark orange color

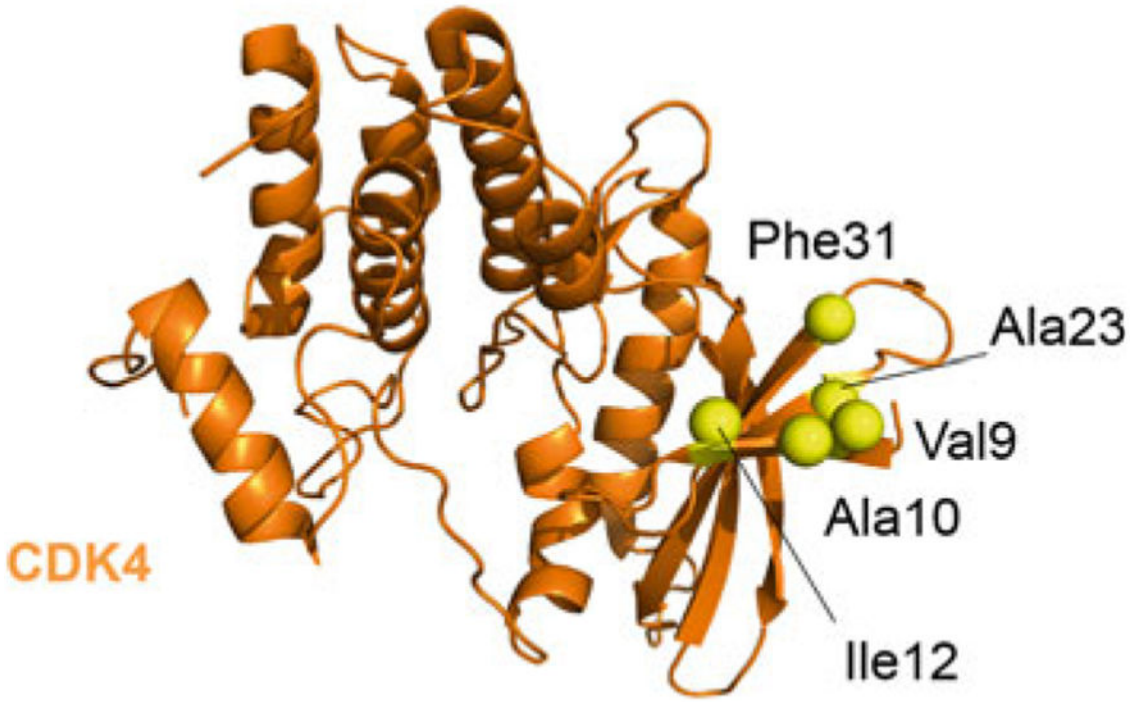

Fig. 5.

Hot spots in CDK4 interface targeted with CDK6 inhibitors. Hot spots Val9, Ala10, Ile12, Ala23, and Phe31 are targeted to inhibit CDK4 interactions (PDB ID: 2W96, chain B). Yellow balls represent hot spots

Table 1

Hot spot prediction tools/algorithms

Name	Features	Website
Prediction from unbound proteins		
ISIS [35]	Uses amino acid sequences to predict the hot spot residues	https://www.rostlab.org/services/isis
pyDockNIP [36]	Uses normalized interface propensity values derived from rigid body docking	
GNM-based predictions [37]	Measures dynamic fluctuations in high-frequency modes	
SIM [38]	Measures the dynamic exposure of hydrophobic patches on the protein surfaces	
Prediction from the protein complex		
Robetta [108]	Measures energies of packing interactions, hydrogen bonds, and solvation	http://www.robetta.org/alascansubmit.jsp
KFC/KFC 2 [48]	Considers shape specificity, biochemical contact, and plasticity features of the interface residues	http://mitchell-lab.biochem.wisc.edu/KFC_Server
APIS [109]	Combines protrusion index with solvent accessibility	http://home.ustc.edu.cn/~jfxia/hotspot.html
HotPoint [65]	Considers the solvent accessibility and the total contact potential of the interface residues	http://prism.ccbb.ku.edu.tr/hotpoint/
PredHS [110]	Uses machine learning algorithm to optimize structural and energetic features	http://www.predhs.org
FOLDEF [40]	Uses FoldX energies to predict the hot spot residues	http://fold-x.embl-heidelberg.de
MutaBind [41]	Calculates binding energy changes, which can be used to predict hot spots, based on molecular mechanics force fields	http://www.ncbi.nlm.nih.gov/research/mutabind
MAPPIS [111]	Compares physicochemical interactions of PPIs with multiple alignment	http://bioinfo3d.cs.tau.ac.il/mappis/
ANCHOR [112]	Calculates the change in solvent-accessible surface area upon binding for each side chain	http://structure.pitt.edu/anchor/
PCRPi [113]	Integrates diverse metrics into a unique probabilistic measure by using Bayesian networks	http://www.bioinsilico.org/PCRPi
HotRegion [66]	Predicts hot spots using the same algorithm as HotPoint and predicts hot regions	http://prism.ccbb.ku.edu.tr/hotregion/

Table 2

Representative online protein interface databases

Name	Web server	Input type						
		Protein name	PDB ID	Pfam ID	Sequence	UniProt ID	GO ID	HETATM code
PIFACE [54]	http://prism.ccbb.ku.edu.tr/piface	–	✓	✓	–	–	–	–
PLIC [60]	http://proline.biochem.iisc.ernet.in/PLIC/index.php	✓	✓	✓	–	–	✓	✓
sc-PDB [62]	http://bioinfo-pharma.u-strasbg.fr/scPDB	✓	✓	–	–	✓	–	–
ProtCID [63]	http://dunbrack2.fccc.edu/ProtCiD/default.aspx	–	✓	✓	✓	✓	–	–
3did [64]	https://3did.irbbarcelona.org	–	✓	✓	–	–	✓	–

Table 3

Representative online protein binding pocket prediction methods

Name	Web server	Features
DeepSite [114]	http://www.playmolecule.org/deepsite	Uses neural network to predict ligand binding pockets on proteins
AlloPred [115]	http://www.sbg.bio.ic.ac.uk/allopred/home	Investigates normal mode perturbation analysis and pocket features to predict allosteric pockets on proteins
PockDrug [116]	http://pockdrug.rpbs.univparis-diderot.fr/cgi-bin/index.py	Uses a combination of pocket estimation methods and pocket properties to predict pocket druggability
LIGSITE^{csc} [117]	http://projects.biotec.tudresden.de/cgi-bin/index.php	Identifies pockets on protein surface using Connolly surface and degree of conservation
MetaPocket [118]	http://projects.biotec.tudresden.de/metapocket	Combines the predicted binding sites from eight different methods to identify ligand binding sites on protein surface
POCASA [119]	http://altair.sci.hokudai.ac.jp/g6/service/pocasa	Predicts protein binding sites by rolling a sphere to detect pockets and cavities on protein surface