Review

Structure-based protein-protein interaction networks and drug design

Hammad Naveed and Jingdong J. Han*

Chinese Academy of Sciences Key Laboratory of Computational Biology, Chinese Academy of Sciences-Max Planck Partner Institute for Computational Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200031, China
*Correspondence: jdhan@picb.ac.cn

Proteins carry out their functions by interacting with other proteins and small molecules, forming a complex interaction network. In this review, we briefly introduce classical graph theory based protein-protein interaction networks. We also describe the commonly used experimental methods to construct these networks, and the insights that can be gained from these networks. We then discuss the recent transition from graph theory based networks to structure based protein-protein interaction networks and the advantages of the latter over the former, using two networks as examples. We further discuss the usefulness of structure based protein-protein interaction networks for drug discovery, with a special emphasis on drug repositioning.

Keywords: protein-protein interaction; network; structure-based; drug design; drug reposition

INTRODUCTION

Proteins form the basic functional units of a cell. They carry out their functions by interacting with other proteins and small molecules. It is important to characterize the protein-protein interaction interface to gain mechanical insight into these interactions. On a systems level, these interactions form a complex network responsible for responding to both intracellular and extracellular perturbations [1]. A number of experimental techniques have been developed in recent years to comprehensively map such networks. However, due to technical limitations, a significant number of interactions, and in particular the dynamic interactions in such networks, are yet to be discovered.

In this review, we first discuss the experimental and theoretical methods to construct "classical" protein-protein interaction networks. We then summarize recent progress on integrating structural information into these interaction networks. We also discuss how interaction networks are being utilized in rational drug design.

GRAPH THEORY BASED "CLASSICAL" PPI NETWORKS

In the post genomic era, significant effort has been put into identifying and understanding the role of various coding and non-coding regions of the genome [2]. Knockout experiments, targeted mutations, functional assays and other biochemical methods have been used to gain insight into the functions of individual proteins [3]. As most proteins carry out their functions by interacting with other proteins, a number of experimental methodologies, such as yeast two-hybrid (Y2H) [4–8], co-immunoprecipitation [9,10] and co-expression data [11,12], have been used to construct protein-protein interaction networks.

Several databases, including DIP [13], MINT [14], HPRD [15], BioGrid [16], BIND [17] and IntAct [18], have then compiled this data from various sources. A brief description of these databases can be found in Table 1.

Traditionally, PPI networks have been represented as graphs (Fig. 1), where each node represents a protein and

Table 1. A brief description of the databases that integrate protein-protein interaction data from various sources

Name	Description	Website
DIP [13]	Catalogs experimentally determined interactions between proteins. The data are curated both manually by expert curators and also automatically using computational approaches	dip.doe-mbi.ucla.edu/
MINT [14]	Focuses on experimentally verified protein-protein interactions mined from the scientific literature	mint.bio.uniroma2.it/mint/
HPRD [15]	Focuses on visual depictions of the domain architecture, post-translational modifications, interaction networks and disease association for each protein in the human proteome, extracted manually from literature	hprd.org/
BioGRID [16]	Contains collections of protein and genetic interactions from major model organism species. It includes a comprehensive set of interactions reported to date in the primary literature for budding yeast, thale cress and fission yeast	thebiogrid.org/
BIND [17]	Comprises data from peer-reviewed literature and direct submissions	bond.unleashedinformatics.com/
IntAct [18]	All interactions are derived from literature curation or direct user submissions	www.ebi.ac.uk/intact/

interactions between them are shown as edges. Several analyses of these networks have illustrated the built-in robustness of these networks by calculating the degree (number of interactions) of each protein [19–24]. Moreover, the proteins/genes in these networks are not randomly located; instead, proteins associated with a particular function tend to form clusters [25–28], and those associated with a disease have a large number of protein-protein interactions [29,30]. However, the elevated degree observed for disease-associated genes may have some inherent bias because many studies have focused on cancer genes alone and also, in general, disease-associated genes might have higher reported interactions because they attract more research interest [31]. The graphical representation of the network is also useful in tracing potentially perturbed / malfunctioning proteins (nodes). However, this representation ignores the structural information of the protein-protein interaction interface.

STRUCTURE BASED PPI NETWORKS

High-throughput approaches like Y2H, co-immunoprecipitation and co-expression do not provide structural details for protein-protein interactions, and sometimes contain significant false positives [8,32–34]. High-resolution structural protein-protein interaction data can be obtained by X–ray crystallography [35] and NMR spectroscopy [36], while cryo-EM provides low resolution structural data [37]. As of November 2012, more than 86000 structures have been deposited in the Protein Data Bank (PDB) [38]. Although a significant number of structures are now available, a portion of these are monomeric and others may contain non-native packing interactions [39].

Several computational methods have been designed to complement experimental approaches. Docking meth-

odologies are widely used to predict the bound state of two proteins. ZDOCK [40], PIPER [41], ClusPro [42], HADDOCK [43], RosettaDOCK [44] and PatchDock [45] are some of the most commonly used docking methods. They can be broadly divided into two categories: (i) methods that utilize Fast Fourier Transform (FFT) to search for the best interaction conformation during rigid body rotations/translations, (ii) methods that use experimental information, such as interface residues and NMR data. The critical assessment of predicted interactions (CAPRI) is a community wide experiment, held every two years, that aims to judge the performance of existing methods [46]. Homology based methodologies are also widely used, particularly for large-scale studies, as docking methods are time intensive [47–50]. An alternative approach to predict reliable protein-protein interaction is to utilize only the interface information from a homolog protein [51–54]. Interface-based approaches take advantage of the observation that protein interaction sites are more conserved than the remainder of the protein surface [55–57]. A recent study on 231 enzyme families showed that even a sequence identity of 45% between the binding surface of the template protein and the modeled protein could generate the interaction interface successfully [58]. Computational approaches that specialize in identification of the protein-protein interaction sites for a particular type of protein, e.g., membrane proteins, have also been very successful [59–64]. The identification of pockets on the protein surface has also been used successfully by a number of groups to predict protein-protein interaction sites [65–67]. A brief description of computational tools to detect protein-protein interactions can be found in Table 2.

Structural Protein-protein Interaction Networks provide rich mechanistic insight into the regulatory mechanisms of proteins. Not only do they provide information about the important residues involved in the interactions, but
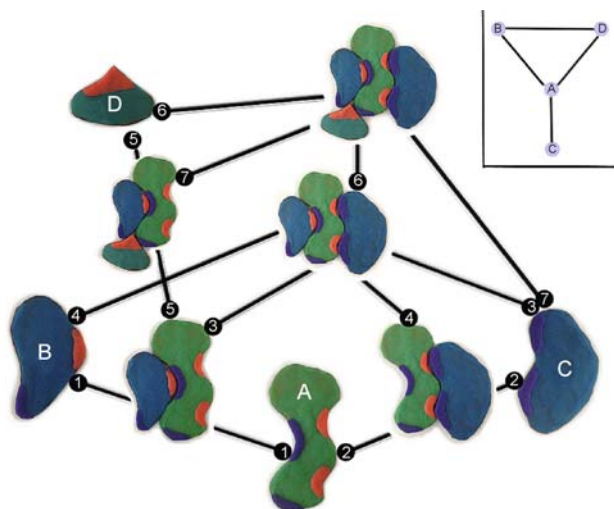
Figure 1. **Classical graph theory based representation of a hypothetical protein-protein interaction network.** The nodes represent the proteins and the interactions between them are shown as edges. Such a representation does not allow for the inclusion of the structural information for the interaction interface. (Main) Structure-based representation of the hypothetical protein-protein interaction network shown in the inset. Mechanistic information, such as where the proteins bind and whether they compete for a binding interaction, is incorporated in this representation. The red patches on the proteins show predominantly positive charge in the active site, while the blue patches represent predominantly negative charge. Some interactions occur only in a particular oligomerization state, e.g., Protein D only interacts with the AB protein complex and not with proteins A or B individually. Such scenarios occur when the binding surface for an interaction is dependent on the interaction of two other proteins. Protein C can bind to Protein A regardless of the binding events of Protein A to Protein B and Protein D. The interactions are numbered for clarity and do not represent a chronological order. The two ends of an interaction are numbered the same to specify which proteins/complexes are involved in the interaction. The complex formed as a result of the interaction is shown in the middle of each interaction (edge).

also indicate whether two proteins might simultaneously interact or compete for a binding partner. If both proteins bind to approximately the same surface on a protein, it is more than likely that they will compete for the binding interaction due to steric hindrances. On the other hand, if the two proteins bind to different parts of a protein surface, it is likely that they can interact with the protein simultaneously. Most proteins interact with only a few other proteins. However, some proteins (named *hub proteins*) have a large number of protein-protein interac-

tions [70,71]. Hub proteins, can include families of enzymes, transcription factors and intrinsically disordered proteins, among others [72,73]. The number of interactions in hub proteins is larger than the number of interaction interfaces. Therefore, hub proteins often reuse their PPI interfaces for multiple interactions. Intrinsically disordered proteins achieve this by sampling the low energy conformation landscape continuously. This enables them to present different interaction interfaces to different binding partners [72]. Studies have shown that hub proteins are more likely to be associated with diseases like cancer than non-hub proteins [30,74].

A human structural interaction network: Wang et al. have used high quality binary interaction data and homology modeling to construct a human structural interaction network (hSIN) that consists of 2816 proteins and 4222 structurally resolved interactions [75]. Utilizing this structurally resolved protein-protein interaction network, they were able to demonstrate that for the corresponding diseases, the in-frame mutations were enriched on the interaction interfaces of the proteins. Moreover, they discuss the basis of pleiotropy of disease genes and locus heterogeneity with experimental case studies on the interactions of WASP protein with CDC42 and VASP proteins [75]. They also predicted 292 candidate genes to have 694 previously unknown disease-to-gene associations by applying the guilt-by-association principle on their structurally resolved interaction network, based on mutations of known disease genes [75].

An extracellular signal-regulated kinase network: PRISM (PRotein Interactions by Structural Matching) is another useful tool for constructing structure based protein-protein interaction networks [51,76,77]. PRISM utilizes structural motifs derived from known non-redundant binary interactions, evolutionary conservation and flexible refinement to predict protein-protein interactions on a proteome wide scale. The structural network of the Extracellular signal-Regulated Kinases (ERK) in the Mitogen-Activated Protein Kinase (MAPK) signaling pathway was constructed using this approach [77]. This network provides rich information about interactions that can occur simultaneously, and those that are mutually exclusive [77]. 64% of the 25 protein-protein interaction interfaces in the network are utilized for two or more interactions. Most notably, ERK protein is involved in seven interactions using seven distinct interfaces [77].

Interacting proteins share at least some subcellular localization. PPI networks have, therefore, also been used to predict the subcellular localization of protein complexes [78,79]. Interestingly, there is some evidence that suggests that information on subcellular localization can be used, in combination with other features, to predict PPIs [80].

Table 2.   A brief description of the tools that can be used to predict protein-protein interactions from protein structures

Name	Description	Website
ZDOCK [40]	Performs full rigid-body searches of docking orientations between two proteins	Zlab.bu.edu/zdock/
ClusPro [41,42]	Includes rigid body docking data from PIPER, selection of docked structures with favorable desolvation, and electrostatics, and clustering of the retained complexes using a pairwise RMSD criterion. It reports the centers of the largest clusters	cluspro.bu.edu/
HADDOCK [43]	Is a docking method that utilizes biochemical and/or biophysical interaction data, such as chemical shift perturbation data resulting from NMR titration experiments, mutagenesis data or bioinformatic predictions	www.nmr.chem.uu.nl/haddock2.1/
RosettaDOCK [44]	Identifies low-energy conformations of a protein-protein interaction by optimizing rigid-body orientation and side-chain conformations.	rosettadock.graylab.jhu.edu/
PatchDock [45]	Uses geometric hashing and pose-clustering matching techniques to match surface patches	bioinfo3d.cs.tau.ac.il/PatchDock/
InterPreTS [48]	Uses alignments of homologs of the interacting proteins, to assess the fitness of any possible interacting pair on the complex by using empirical potentials	www.russelllab.org/interprets/
PrePPI [50]	Predicts interactions based on structural, functional, evolutionary and expression information	bhapp.c2b2.columbia.edu/PrePPI/
PRISM [51]	Uses rigid-body structural comparisons of target proteins to known template protein-protein interfaces, and flexible refinement using a docking energy function	prism.ccbb.ku.edu.tr/prism_protocol/
IBIS [68]	Uses conservation of structural location and sequence patterns of protein-protein binding sites	www.ncbi.nlm.nih.gov/Structure/ibis/ibis.cgi
TMBB-Explorer [64]	Predicts the structure, oligomerization state, protein-protein interaction sites, and thermodynamic properties of the transmembrane domains of beta-barrel membrane proteins using a physical interaction model, a simplified conformational space for efficient enumeration, and an empirical potential function from a detailed combinatorial analysis	tanto.bioengr.uic.edu/TMBB-Explorer/
CASTp [66]	Uses weighted Delaunay triangulation and the alpha complex for shape measurements, and provides identification and measurements of surface accessible pockets, as well as interior inaccessible cavities, for proteins	sts.bioengr.uic.edu/castp/
PASS [69]	Uses geometry to characterize regions of buried volume in proteins, and identified positions likely to represent binding sites based upon the size, shape, and burial extent of these volumes	www.ccl.net/cca/software/UNIX/pass/overview.shtml

It is important to note that apart from PPI networks, other types of networks that depict other cellular activities, for example metabolic networks (KEGG [81], EcoCyc [82], BioCyc [83], and metaTIGER [84]), have also been used extensively in computational and experimental studies.

PPI NETWORKS AND DRUG DESIGN

Structural protein-protein interaction networks are a valuable resource for drug discovery. Proteins function by interacting with other proteins. Therefore, interacting proteins are likely to be involved in the same cellular processes. As a result, perturbing these interactions can result in a number of outcomes, including onset or intensification of a disease such as cancer [85–87]. Perturbing these interactions can often cause *loss of function* or *gain of function* [88]. With the availability of the complete structural information of the interaction interface, it is possible to design peptide inhibitors that mimic the interaction partner and perturb a normal PPI [89,90]. Moreover, the side effects of a drug can be

predicted much more accurately using the structural and topological information embedded in the structure based interaction networks, as compared to just the topological information in the classical graph-based interaction networks [91].

The classical assumption of one drug target for one drug to treat a single disease has been shown to be inaccurate in a number of cases. This assumption might be the reason for the high failure of new drugs in clinical trials as a result of low efficacy and high toxicity [92–94]. So-called "Off-Target" binding generally contributes to side effects and toxicity [95]. However, there have been a few cases where "Off-Target" binding has been beneficial [96]. Each known drug on average binds to 6 known targets, and therefore it is predicted that on average there will be 6 targets, known or unknown, for each newly discovered drug [97]. To understand the side effects and toxicity of rejected drugs, it is important to predict "Off-Target" binding sites. A number of methods have used clustering of proteins into families [98,99], global structure similarity measurement [100,101] and interface similarity measurement [102–105] to predict "Off-Target" binding or to redesign drugs to enhance efficacy. Lounkine et al. have used a similarity ensemble approach to predict off-targets, based on whether a molecule will bind to a target with similar chemical features to those of known targets [106]. They further linked the off-targets to adverse drug reactions (ADR) by using a guilt-by-association pipeline that linked off-targets to the ADRs of drugs for which the off-targets were primary targets [106]. They computationally screened 656 drugs approved for human use against 73 target proteins, and verified their predictions by either searching in protein-ligand databases or performing binding and functional assays [106]. Moreover, they were able to construct a three-way Drug-Target-ADR network that could be an extremely useful starting point for future off-target and ADR predictions [106]. The traditional approach to drug design is to focus on the molecular level, while the phenotypic outcomes in the clinical trial are measured at the organism level. Therefore, in future it will be extremely beneficial to predict off-target binding using structure based PPI networks [85,107], and to predict the drug targets [108], drug response [109,110] and even drug resistance [111] well before entering the clinical trial stage.

Drug repositioning: Due to the high failure rate and huge costs involved with traditional drug design [112,113], a new paradigm has emerged that identifies new targets for existing approved drugs [114,115]. This approach is based on the fact that each drug on average binds more than one target, and that the cost of Phase I clinical trials could be saved by re-using existing approved drugs. The amount of chemical and biologic data has increased exponentially in recent years, and databases have been developed to integrate large amounts of data arising from different sources. PROMISCUOUS is one such database that integrates drug-protein interactions, protein-protein interactions and side effect data for drug repositioning studies [116]. Chem2Bio2RDF is another useful database that integrates chemical and drug data, protein and gene data, chemogenomics data, protein-protein interaction and pathway data and side effects data for designing multiple pathway inhibitors and predicting adverse drug reactions [117]. Drug repositioning can be particularly useful for rare and orphan diseases [118,119]. Successful repositioning of drugs for novel targets has been achieved not only using approved drugs [120] but also using late stage failures [121,122]. With the advent of structure based PPI networks this approach is likely to become much more effective, as lead targets and ADRs can be predicted much more accurately using structure based PPI networks than graph theory based classical networks.

Drug-drug interaction (DDI): Drug-drug interaction (DDI) is often the cause of ADRs, particularly for patient populations regularly taking multiple drugs. DDIs occur when the pharmacologic response of a particular drug is transformed by the action of another drug [123], resulting in potentially harmful clinical effects. A recent algorithm by Huang et al. for the systematic prediction of pharmacodynamic DDIs that considers drug actions and their clinical effects for the first time in the context of complex PPI networks is a significant step forward for predicting ADRs [124]. The authors show that the integration of network topology, cross-tissue gene expression correlations and side effect similarity can predict DDIs with significant success. One of the major avenues for improvement in future studies would be utilization of a structure based PPI network, which would provide a high confidence molecular network reducing the number of false positives. This should also provide a framework to study the mechanisms of Drug-protein interactions. By integrating structural information, it will be possible for the first time to assess the ratio of DDIs occurring as a result of competitive binding, allosteric effects or indirect influence.

Yet not all DDIs are bad side effects of drugs; some DDIs provide a useful *Drug Combination* strategy. Identification of drugs that could be prescribed simultaneously may improve efficacy and reduce side effects. This strategy is particularly effective in accounting for pathway redundancy. Several drug combinations have already been reported, particularly for various types of cancer [125–127] and Human Immunodeficiency Virus (HIV) [128–130]. In the near future, the use of structure

based PPI networks will provide more detailed information to bring bioinformatics studies in the field of Drug Combination strategies to the next level.

CONCLUDING REMARKS

Protein-protein interaction interfaces are a rich resource for gaining insight into the mechanism of how a protein carries out its functions. A complete structurally resolved interaction network of all the proteins will be an invaluable resource for not only understanding the complex and essential functions associated with these proteins, but will also help in designing novel therapeutic strategies for the diseases associated with these proteins. Due to experimental limitations, computational methods are extremely useful for completing such interaction networks, and for using these networks to predict the side effects of new drugs and to reposition existing drugs.

ACKNOWLEDGEMENTS

CONFLICT OF INTEREST

The authors Hammad Naveed and Jingdong J. Han declare that they have no conflict of interests.

REFERENCES

1. Hartwell, L. H., Hopfield, J. J., Leibler, S., and Murray, A. W. (1999) From molecular to modular cell biology. Nature, 402, C47–C52.

2. Dunham, I., Kundaje, A., Aldred, S. F., Collins, P. J., Davis, C. A., Doyle, F., Epstein, C. B., Frietze, S., Harrow, J., Kaul, R., et al. and the ENCODE Project Consortium. (2012) An integrated encyclopedia of DNA elements in the human genome. Nature, 489, 57–74.

3. Whisstock, J. C. and Lesk, A. M. (2003) Prediction of protein function from protein sequence and structure. Q. Rev. Biophys., 36, 307–340.

4. Fields, S. Uetz, P. Giot, L. Cagney, G. Mansfield, T. A. Judson, R. S. Knight, J. R. Lockshon, D. Narayan, V. Srinivasan, M. et al. (2000) A comprehensive analysis of protein-protein interactions in *Saccharomyces cerevisiae*. Nature, 403, 623–627.

5. Ito, T. Chiba, T. Ozawa, R. Yoshida, M. Hattori, M. and Sakaki, Y. (2001) A comprehensive two-hybrid analysis to explore the yeast protein interactome. Proc. Natl. Acad. Sci. U.S.A., 98, 4569–4574.

6. Li, S. Armstrong, C. M. Bertin, N. Ge, H. Milstein, S. Boxem, M. Vidalain, P. O. Han, J. D. Chesneau, A. Hao, T. et al. (2004) A map of the interactome network of the metazoan *C. elegans*. Science, 303, 540–543.

7. Rual, J.-F. Venkatesan, K. Hao, T. Hirozane-Kishikawa, T. Dricot, A. Li, N. Berriz, G. F. Gibbons, F. D. Dreze, M. Ayivi-Guedehoussou, N. et al. (2005) Towards a proteome-scale map of the human protein-protein interaction network. Nature, 437, 1173–1178.

8. Rajagopala, S. V. and Uetz, P. (2011) Analysis of protein-protein interactions using high-throughput yeast two-hybrid screens. Methods Mol. Biol., 781, 1–29.

9. Gavin, A. C. Aloy, P. Grandi, P. Krause, R. Boesche, M. Marzioch, M. Rau, C. Jensen, L. J. Bastuck, S. Dümpelfeld, B. et al. (2006) Proteome survey reveals modularity of the yeast cell machinery. Nature, 440, 631–636.

10. Krogan, N. J. Cagney, G. Yu, H. Zhong, G. Guo, X. Ignatchenko, A. Li, J. Pu, S. Datta, N. Tikuisis, A. P. et al. (2006) Global landscape of protein complexes in the yeast *Saccharomyces cerevisiae*. Nature, 440, 637–643.

11. Soong, T. T. Wrzeszczynski, K. O. and Rost, B. (2008) Physical protein-protein interactions predicted from microarrays. Bioinformatics, 24, 2608–2614.

12. Liu, C. T., Yuan, S. and Li, K. C. (2009) Patterns of co-expression for protein complexes by size in *Saccharomyces cerevisiae*. Nucleic Acids Res., 37, 526–532.

13. Salwinski, L. Miller, C. S. Smith, A. J. Pettit, F. K. Bowie, J. U. and Eisenberg, D. (2004) The database of interacting Proteins: 2004 update. Nucleic Acids Res., 32, D449–D451.

14. Licata, L. Briganti, L. Peluso, D. Perfetto, L. Iannuccelli, M. Galeota, E. Sacco, F. Palma, A. Nardozza, A. P. Santonico, E. et al. (2012) MINT the molecular interaction database: 2012 update. Nucleic Acids Res., 40, D857–D861.

15. Keshava Prasad, T. S. Goel, R. Kandasamy, K. Keerthikumar, S. Kumar, S. Mathivanan, S. Telikicherla, D. Raju, R. Shafreen, B. Venugopal, A. et al. (2009) Human Protein Reference Database–2009 update. Nucleic Acids Res., 37, D767–D772.

16. Stark, C. Breitkreutz, B. J. Chatr-Aryamontri, A. Boucher, L. Oughtred, R. Livstone, M. S. Nixon, J. Van Auken, K. Wang, X. Shi, X. et al. (2011) The BioGRID Interaction Database: 2011 update. Nucleic Acids Res., 39, D698–D704.

17. Willis, R. C. and Hogue, C. W. (2006) Searching viewing and visualizing data in the Biomolecular Interaction Network Database (BIND) Curr. Protoc. Bioinformatics, 8.9.1–8.9.30.

18. Kerrien, S. Aranda, B. Breuza, L. Bridge, A. Broackes-Carter, F. Chen, C. Duesbury, M. Dumousseau, M. Feuermann, M. Hinz, U. et al. (2012) The IntAct molecular interaction database in 2012. Nucleic Acids Res., 40, D841–D846.

19. Jeong, H. Mason, S. P. Barabási, A.-L. and Oltvai, Z. N. (2001) Lethality and centrality in protein networks. Nature, 411, 41–42.

20. Barabási, A.-L. and Bonabeau, E. (2003) Scale-free networks. Sci. Am., 288, 60–69.

21. Han, J.-D. Bertin, N. Hao, T. Goldberg, D. S. Berriz, G. F. Zhang, L. V. Dupuy, D. Walhout, A. J. Cusick, M. E. Roth, F. P. et al. (2004) Evidence for dynamically organized modularity in the yeast protein-protein interaction network. Nature, 430, 88–93.

22. Yook, S. H. Oltvai, Z. N. and Barabási, A. L. (2004) Functional and topological characterization of protein interaction networks. Proteo-

mics, 4, 928–942.

23. Lim, J. Hao, T. Shaw, C. Patel, A. J. Szabá, G. Rual, J. F. Fisk, C. J. Li, N. Smolyar, A. Hill, D. E. et al. (2006) A protein-protein interaction network for human inherited ataxias and disorders of Purkinje cell degeneration. Cell, 125, 801–814.

24. Huang, H. Jedynak, B. M. and Bader, J. S. (2007) Where have all the interactions gone? Estimating the coverage of two-hybrid protein interaction maps. PLoS Comput. Biol., 3, e214.

25. Shen-Orr, S. S. Milo, R. Mangan, S. and Alon, U. (2002) Network motifs in the transcriptional regulation network of *Escherichia coli*. Nat. Genet., 31, 64–68.

26. Milo, R. Shen-Orr, S. Itzkovitz, S. Kashtan, N. Chklovskii, D. and Alon, U. (2002) Network motifs: simple building blocks of complex networks. Science, 298, 824–827.

27. Said, M. R. Begley, T. J. Oppenheim, A. V. Lauffenburger, D. A. and Samson, L. D. (2004) Global network analysis of phenotypic effects: protein networks and toxicity modulation in *Saccharomyces cerevisiae*. Proc. Natl. Acad. Sci. U.S.A., 101, 18006–18011.

28. Shachar, R. Unger, L. Kupiec, M. Ruppin, R. and Sharan, R. (2008) A systems-level approach to mapping the telomere-length maintenance gene circuitry Mol. Syst. Biol., 4, 172.

29. Wachi, S. Yoneda, K. and Wu, R. (2005) Interactome-transcriptome analysis reveals the high centrality of genes differentially expressed in lung cancer tissues. Bioinformatics, 21, 4205–4208.

30. Jonsson, P. F. and Bates, P. A. (2006) Global topological features of cancer proteins in the human interactome. Bioinformatics, 22, 2291–2297.

31. Goh, K. I. Cusick, M. E. Valle, D. Childs, B. Vidal, M. and Barabási, A. L. (2007) The human disease network. Proc. Natl. Acad. Sci. U.S.A., 104, 8685–8690.

32. Nguyen, T. N. and Goodrich, J. A. (2006) Protein-protein interaction assays: eliminating false positive interactions. Nat. Methods, 3, 135–139.

33. Fullwood, M. J. and Ruan, Y. (2009) ChIP-based methods for the identification of long-range chromatin interactions. J. Cell. Biochem., 107, 30–39.

34. Cusick, M. E. Yu, H. Smolyar, A. Venkatesan, K. Carvunis, A. R. Simonis, N. Rual, J. F. Borick, H. Braun, P. Dreze, M. et al. (2009) Literature-curated protein interaction datasets. Nat. Methods 6, 39–46.

35. Ennifar, E. (2012) X-ray crystallography as a tool for mechanism-of-action studies and drug discovery. Curr. Pharm. Biotechnol., PMID: 22429136

36. Guerry, P. and Herrmann, T. (2011) Advances in automated NMR protein structure determination. Q. Rev. Biophys., 44, 257–309.

37. Glaeser, R. M. and Hall, R. J. (2011) Reaching the information limit in cryo-EM of biological macromolecules: experimental aspects. Biophys. J., 100, 2331–2337.

38. Berman, H. M. Westbrook, J. Feng, Z. Gilliland, G. Bhat, T. N. Weissig, H. Shindyalov, I. N. and Bourne, P. E. (2000) The protein data bank. Nucleic Acids Res., 28, 235–242.

39. Dunitz, J. D. and Gavezzotti, A. (2005) Molecular recognition in organic crystals: directed intermolecular bonds or nonlocalized bonding? Angew. Chem. Int. Ed. Engl., 44, 1766–1787.

40. Chen, R., Li, L. and Weng, Z. (2003) ZDOCK: an initial-stage protein-docking algorithm. Proteins, 52, 80–87.

41. Kozakov, D. Brenke, R. Comeau, S. R. and Vajda, S. (2006) PIPER: an FFT-based protein docking program with pairwise potentials. Proteins, 65, 392–406.

42. Kozakov, D. Hall, D. R. Beglov, D. Brenke, R. Comeau, S. R. Shen, Y. Li, K. Zheng, J. Vakili, P. Paschalidis, I. C. et al. (2010) Achieving reliability and high accuracy in automated protein docking: Cluspro PIPER SDU and stability analysis in CAPRI rounds 13–19, Proteins. Proteins, 78, 3124–3130.

43. de Vries, S. J. van Dijk, M. and Bonvin, A. M. (2010) The HADDOCK web server for data-driven biomolecular docking. Nat. Protoc., 5, 883–897.

44. Lyskov, S. and Gray, J. J. (2008) The RosettaDock server for local protein-protein docking. Nucleic Acids Res., 36, W233-8.

45. Schneidman-Duhovny, D. Inbar, Y. Nussinov, R. and Wolfson, H. J. (2005) PatchDock and SymmDock: servers for rigid and symmetric docking. Nucleic Acids Res., 33, W363-7.

46. Fernández-Recio, J. and Sternberg, M. J. E. (2010) The 4th meeting on the Critical Assessment of Predicted Interaction (CAPRI) held at the Mare Nostrum Barcelona Proteins, 78, 3065–3066.

47. Aloy, P. and Russell, R. B. (2002) Interrogating protein interaction networks through structural biology. Proc. Natl. Acad. Sci. U.S.A., 99, 5896–5901.

48. Aloy, P. and Russell, R. B. (2003) InterPreTS: protein interaction prediction through tertiary structure. Bioinformatics, 19, 161–162.

49. Kundrotas, P. J. Lensink, M. F. and Alexov, E. (2008) Homology-based modeling of 3D structures of protein-protein complexes using alignments of modified sequence profiles. Int. J. Biol. Macromol., 43, 198–208.

50. Zhang, Q. C. Petrey, D. Deng, L. Qiang, L. Shi, Y. Thu, C. A. Bisikirska, B. Lefebvre, C. Accili, D. Hunter, T. et al. (2012) Structure-based prediction of protein-protein interactions on a genome-wide scale. Nature, 490, 556–560.

51. Ogmen, U. Keskin, O. Aytuna, A. S. Nussinov, R. and Gursoy, A. (2005) PRISM: protein interactions by structural matching. Nucleic Acids Res., 33, W331-6.

52. Gunther, S. May, P. Hoppe, A. Frommel, C. and Preissner, R. (2007) Docking without docking: ISEARCH-prediction of interactions using known interfaces. Proteins, 69, 839–844.

53. Sinha, R. Kundrotas, P. J. and Vakser, I. A. (2010) Docking by structural similarity at protein-protein interfaces. Proteins, 78, 3235–3241.

54. Tyagi, M. Thangudu, R. R. Zhang, D. Bryant, S. H. Madej, T. and Panchenko, A. R. (2012) Homology inference of protein-protein interactions via conserved binding sites. PLoS ONE, 7, e28896.

55. Fraser, H. B. Wall, D. P. and Hirsh, A. E. (2003) A simple dependence between protein evolution rate and the number of protein-protein interactions. BMC Evol. Biol., 3, 11.

56. Fraser, H. B. (2005) Modularity and evolutionary constraint on proteins. Nat. Genet., 37, 351–352.

57. Choi, Y. S. Yang, J. S. Choi, Y. Ryu, S. H. and Kim, S. (2009) Evolutionary conservation in multiple faces of protein interaction. Proteins, 77, 14–25.

58. Zhao, J. Dundas, J. Kachalo, S. Ouyang, Z. and Liang, J. (2011) Accuracy of functional surfaces on comparatively modeled protein structures. J. Struct. Funct. Genomics, 12, 97–107.

59. Naveed, H. Jackups, R. Jr and Liang, J. (2009) Predicting weakly stable regions oligomerization state and protein-protein interfaces in transmembrane domains of outer membrane proteins. Proc. Natl. Acad. Sci. U. S. A., 106, 12735–12740.

60. Bordner, A. J. (2009) Predicting protein-protein binding sites in membrane proteins. BMC Bioinformatics, 10, 312.

61. Gessmann, D. Mager, F. Naveed, H. Arnold, T. Weirich, S. Linke, D. Liang, J. and Nussberger, S. (2011) Improving the resistance of a eukaryotic β-barrel protein to thermal and chemical perturbations. J. Mol. Biol., 413, 150–161.

62. Geula, S. Naveed, H. Liang, J. and Shoshan-Barmatz, V. (2012) Structure-based analysis of VDAC1 protein: defining oligomer contact sites. J. Biol. Chem., 287, 2179–2190.

63. Naveed, H. Jimenez-Morales, D. Tian, J. Pasupuleti, V. Kenney, L. J. and Liang, J. (2012) Engineered oligomerization state of OmpF protein through computational design decouples oligomer dissociation from unfolding. J. Mol. Biol., 419, 89–101.

64. Naveed, H. and Liang, J. (2012) TMBB-Explorer: A Webserver to Predict the Structure, Oligomerization State PPI Interface and Thermodynamic Properties of the Transmembrane Domains of Outer Membrane Proteins. Biophys. J., 102, 469a.

65. Levitt, D. G. and Banaszak, L. J. (1992) POCKET: a computer graphics method for identifying and displaying protein cavities and their surrounding amino acids. J. Mol. Graph., 10, 174–177.

66. Dundas, J. Ouyang, Z. Tseng, J. Binkowski, A. Turpaz, Y. and Liang, J. (2006) CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues. Nucleic Acids Res., 34, W116-8.

67. Eyrisch, S. and Helms, V. (2007) Transient pockets on protein surfaces involved in protein-protein interaction. J. Med. Chem., 50, 3457–3464.

68. Shoemaker, B. A. Zhang, D. Tyagi, M. Thangudu, R. R. Fong, J. H. Marchler-Bauer, A. Bryant, S. H. Madej, T. and Panchenko, A. R. (2012) IBIS (Inferred Biomolecular Interaction Server) reports predicts and integrates multiple types of conserved interactions for proteins. Nucleic Acids Res., 40, D834–D840.

69. Brady, G. P. Jr and Stouten, P. F. W. (2000) Fast prediction and visualization of protein binding pockets with PASS. J. Comput. Aided Mol. Des., 14, 383–401.

70. Jeong, H. Tombor, B. Albert, R. Oltvai, Z. N. and Barabási, A.-L. (2000) The large-scale organization of metabolic networks. Nature, 407, 651–654.

71. Mirzarezaee, M. Araabi, B. N. and Sadeghi, M. (2010) Features analysis for identification of date and party hubs in protein interaction network of Saccharomyces cerevisiae. BMC Syst. Biol., 4, 172.

72. Dunker, A. K. Cortese, M. S. Romero, P. Iakoucheva, L. M. and Uversky, V. N. (2005) Flexible nets. The roles of intrinsic disorder in protein interaction networks. FEBS J., 272, 5129–5148.

73. Sarmady, M. Dampier, W. and Tozeren, A. (2011) HIV protein sequence hotspots for crosstalk with host hub proteins. PLoS ONE, 6, e23293.

74. Matthews, L. R. Vaglio, P. Reboul, J. Ge, H. Davis, B. P. Garrels, J. Vincent, S. and Vidal, M. (2001) Identification of potential interaction networks using sequence-based searches for conserved protein-protein interactions or "interologs". Genome Res., 11, 2120–2126.

75. Wang, X. Wei, X. Thijssen, B. Das, J. Lipkin, S. M. and Yu, H. (2012) Three-dimensional reconstruction of protein networks provides insight into human genetic disease. Nat. Biotechnol., 30, 159–164.

76. Tuncbag, N. Gursoy, A. Nussinov, R. and Keskin, O. (2011) Predicting protein-protein interactions on a proteome scale by matching evolutionary and structural similarities at interfaces using PRISM. Nat. Protoc., 6, 1341–1354.

77. Kuzu, G. Keskin, O. Gursoy, A. and Nussinov, R. (2012) Constructing structural networks of signaling pathways on the proteome scale Curr. Opin. Struct. Biol., 22, 367–377.

78. Shin, C. J. Wong, S. Davis, M. J. and Ragan, M. A. (2009) Protein-protein interaction as a predictor of subcellular location. BMC Syst. Biol., 3, 28.

79. Jiang, J. Q. and Wu, M. (2012) Predicting multiplex subcellular localization of proteins using protein-protein interaction network: a comparative study. BMC Bioinformatics, 13, S20.

80. Liu, J. Zhao, H. Tan, J. Luo, D. Yu, W. Harner, E. J. and Shih, W. J. (2008) Is subcellular localization informative for modeling protein-protein interaction signal? Research Letters in Signal Processing, DOI: 10.1155/2008/365152

81. Kanehisa, M. Goto, S. Sato, Y. Furumichi, M. and Tanabe, M. (2012) KEGG for integration and interpretation of large-scale molecular data sets. Nucleic Acids Res., 40, D109–D114.

82. Keseler, I. M. Mackie, A. Peralta-Gil, M. Santos-Zavaleta, A. Gama-Castro, S. Bonavides-Martínez, C. Fulcher, C. Huerta, A. M. Kothari, A. Krummenacker, M. et al. (2013) EcoCyc: fusing model organism databases with systems biology. Nucleic Acids Res., 41, D605–D612.

83. Caspi, R. Altman, T. Dale, J. M. Dreher, K. Fulcher, C. A. Gilham, F. Kaipa, P. Karthikeyan, A. S. Kothari, A. Krummenacker, M. et al. (2010) The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. Nucleic Acids Res., 38, D473–D479.

84. Whitaker, J. W. Letunic, I. McConkey, G. A. and Westhead, D. R. (2009) MetaTIGER: a metabolic evolution resource. Nucleic Acids Res., 37, D531–D538.

85. Ideker, T. and Sharan, R. (2008) Protein networks in disease. Genome Res., 18, 644–652.

86. Wong, J. M. Ionescu, D. and Ingles, C. J. (2003) Interaction between BRCA2 and replication protein A is compromised by a cancer-predisposing mutation in BRCA2. Oncogene, 22, 28–33.

87. Bruncko, M. Oost, T. K. Belli, B. A. Ding, H. Joseph, M. K. Kunzer, A. Martineau, D. McClellan, W. J. Mitten, M. Ng, S. C. et al. (2007) Studies leading to potent dual inhibitors of Bcl-2 and Bcl-xL. J. Med. Chem., 50, 641–662.

88. Schuster-Böckler, B. and Bateman, A. (2008) Protein interactions in human genetic diseases. Genome Biol., 9, R9.

89. London, N. Raveh, B. Movshovitz-Attias, D. and Schueler-Furman, O. (2010) Can self-inhibitory peptides be derived from the interface of globular protein-protein interactions? Proteins, 78, 3140–3149.

90. Eldar-Finkelman, H. and Eisenstein, M. (2009) Peptide inhibitors targeting protein kinases. Curr. Pharm. Des., 15, 2463–2470.

91. Xie, L. Xie, L. and Bourne, P. E. (2011) Structure-based systems biology for analyzing off-target binding. Curr. Opin. Struct. Biol., 21, 189–199.

92. Nwaka, S. and Hudson, A. (2006) Innovative lead discovery strategies for tropical diseases. Nat. Rev. Drug Discov., 5, 941–955.

93. Arrowsmith, J. (2011) Trial watch: phase III and submission failures: 2007–2010. Nat. Rev. Drug Discov., 10, 87.

94. Arrowsmith, J. (2011) Trial watch: Phase II failures: 2008–2010. Nat. Rev. Drug Discov., 10, 328–329.

95. Major, E. O. (2010) Progressive multifocal leukoencephalopathy in patients on immunomodulatory therapies. Annu. Rev. Med., 61, 35–47.

96. Booth, B. and Zemmel, R. (2003) Quest for the best. Nat. Rev. Drug Discov., 2, 838–841.

97. Mestres, J. Gregori-Puigjané, E. Valverde, S. and Solé, R. V. (2009)

The topology of drug-target interaction networks: implicit dependence on drug properties and target families. Mol. Biosyst., 5, 1051–1057.

98. Müller, G. (2003) Medicinal chemistry of target family-directed masterkeys. Drug Discov. Today, 8, 681–691.

99. Hopkins, A. L. and Groom, C. R. (2002) The druggable genome. Nat. Rev. Drug Discov., 1, 727–730.

100. Bisson, W. H. Cheltsov, A. V. Bruey-Sedano, N. Lin, B. Chen, J. Goldberger, N. May, L. T. Christopoulos, A. Dalton, J. T. Sexton, P. M. et al. (2007) Discovery of antiandrogen activity of nonsteroidal scaffolds of marketed drugs. Proc. Natl. Acad. Sci. U.S.A., 104, 11927–11932.

101. Liu, C. I. Liu, G. Y. Song, Y. Yin, F. Hensler, M. E. Jeng, W. Y. Nizet, V. Wang, A. H. and Oldfield, E. (2008) A cholesterol biosynthesis inhibitor blocks *Staphylococcus aureus* virulence. Science, 319, 1391–1394.

102. Specker, E. Böttcher, J. Lilie, H. Heine, A. Schoop, A. Müller, G. Griebenow, N. and Klebe, G. (2005) An old target revisited: two new privileged skeletons and an unexpected binding mode for HIV-protease inhibitors. Angew. Chem. Int. Ed. Engl., 44, 3140–3144.

103. Weber, A. Casini, A. Heine, A. Kuhn, D. Supuran, C. T. Scozzafava, A. and Klebe, G. (2004) Unexpected nanomolar inhibition of carbonic anhydrase by COX-2-selective celecoxib: new pharmacological opportunities due to related binding site recognition. J. Med. Chem., 47, 550–557.

104. Stauch, B. Hofmann, H. Perkovic, M. Weisel, M. Kopietz, F. Cichutek, K. Münk, C. and Schneider, G. (2009) Model structure of APOBEC3C reveals a binding pocket modulating ribonucleic acid interaction required for encapsidation. Proc. Natl. Acad. Sci. U.S.A., 106, 12079–12084.

105. Amaro, R. E. Schnaufer, A. Interthal, H. Hol, W. Stuart, K. D. and McCammon, J. A. (2008) Discovery of drug-like inhibitors of an essential RNA-editing ligase in *Trypanosoma brucei*. Proc. Natl. Acad. Sci. U.S.A., 105, 17278–17283.

106. Lounkine, E. Keiser, M. J. Whitebread, S. Mikhailov, D. Hamon, J. Jenkins, J. L. Lavan, P. Weber, E. Doak, A. K. Côté, S. et al. (2012) Large-scale prediction and testing of drug activity on side-effect targets. Nature 486, 361–367.

107. Butcher, E. C. Berg, E. L. and Kunkel, E. J. (2004) Systems biology in drug discovery. Nat. Biotechnol., 22, 1253–1259.

108. Yildirim, M. A. Goh, K. I. Cusick, M. E. Barabási, A. L. and Vidal, M. (2007) Drug-target network. Nat. Biotechnol., 25, 1119–1126.

109. Taylor, I. W. Linding, R. Warde-Farley, D. Liu, Y. Pesquita, C. Faria, D. Bull, S. Pawson, T. Morris, Q. and Wrana, J. L. (2009) Dynamic modularity in protein interaction networks predicts breast cancer outcome. Nat. Biotechnol., 27, 199–204.

110. Chang, R. L. Xie, L. Xie, L. Bourne, P. E. and Palsson, B. Ø. (2010) Drug off-target effects predicted using structural analysis in the context of a metabolic network model. PLOS Comput. Biol., 6, e1000938.

111. Raman, K. and Chandra, N. (2008) *Mycobacterium tuberculosis* interactome analysis unravels potential pathways to drug resistance. BMC Microbiol., 8, 234.

112. DiMasi, J. A. Hansen, R. W. and Grabowski, H. G. (2003) The price of innovation: new estimates of drug development costs. J. Health Econ., 22, 151–185.

113. Hughes, J. P. Rees, S. Kalindjian, S. B. and Philpott, K. L. (2011) Principles of early drug discovery. Br. J. Pharmacol., 162, 1239–1249.

114. Ashburn, T. T. and Thor, K. B. (2004) Drug repositioning: identifying and developing new uses for existing drugs. Nat. Rev. Drug Discov., 3, 673–683.

115. Dudley, J. T. Deshpande, T. and Butte, A. J. (2011) Exploiting drug-disease relationships for computational drug repositioning. Brief. Bioinform., 12, 303–311.

116. von Eichborn, J. Murgueitio, M. S. Dunkel, M. Koerner, S. Bourne, P. E. and Preissner, R. (2011) PROMISCUOUS: a database for network-based drug-repositioning. Nucleic Acids Res., 39, D1060–D1066.

117. Chen, B. Dong, X. Jiao, D. Wang, H. Zhu, Q. Ding, Y. and Wild, D. J. (2010) Chem2Bio2RDF: a semantic framework for linking and data mining chemogenomic and systems chemical biology data. BMC Bioinformatics, 11, 255.

118. Ekins, S. Williams, A. J. Krasowski, M. D. and Freundlich, J. S. (2011) In silico repositioning of approved drugs for rare and neglected diseases. Drug Discov. Today, 16, 298–310.

119. Sardana, D. Zhu, C. Zhang, M. Gudivada, R. C. Yang, L. and Jegga, A. G. (2011) Drug repositioning for orphan diseases. Brief. Bioinform., 12, 346–356.

120. Campillos, M. Kuhn, M. Gavin, A. C. Jensen, L. J. and Bork, P. (2008) Drug target identification using side-effect similarity. Science 321, 263–266.

121. Duran-Frigola, M. and Aloy, P. (2012) Recycling side-effects into clinical markers for drug repositioning. Genome Med., 4 3.

122. Rajkumar, S. V. (2004) Thalidomide: tragic past and promising future. Mayo Clin. Proc., 79, 899–903.

123. Tatro, D. S. (1992) Drug Interaction Facts 1992. St Louis: Facts and Comparisons.

124. Huang, J. Niu, C. Green, C. D. Yang, L. Mei, H. and Han, J.-D. (2013) Systematic prediction of pharmacodynamic drug-drug interactions through protein-protein-interaction network. PLoS Comput. Biol., 9, e1002998.

125. Greco, F. and Vicent, M. J. (2009) Combination therapy: opportunities and challenges for polymer-drug conjugates as anticancer nanomedicines. Adv. Drug Deliv. Rev., 61, 1203–1213.

126. Breen, E. C. and Walsh, J. J. (2010) Tubulin-targeting agents in hybrid drugs. Curr. Med. Chem., 17, 609–639.

127. Altundag, O. Dursun, P. and Ayhan, A. (2010) Emerging drugs in endometrial cancers. Expert Opin. Emerg. Drugs, 15, 557–568.

128. De Clercq, E. (2012) Where rilpivirine meets with tenofovir, the start of a new anti-HIV drug combination era. Biochem. Pharmacol., 84, 241–248.

129. Schuelter, E. Luebke, N. Jensen, B. Zazzi, M. Sönnerborg, A. Lengauer, T. Incardona, F. Camacho, R. Schmit, J. Clotet, B. et al. (2012) Etravirine in protease inhibitor-free antiretroviral combination therapies. J. Int. AIDS Soc., 15, 18260.

130. Bogojeska, J. and Lengauer, T. (2012) Hierarchical Bayes Model for predicting effectiveness of HIV combination therapies. Stat. Appl. Genet. Mol. Biol., 11, 11.