Chapter 8
Prediction of Biomolecular Complexes

Anna Vangone, Romina Oliva, Luigi Cavallo and Alexandre M.J.J. Bonvin

Abstract Almost all processes in living organisms occur through specific interactions between biomolecules. Any dysfunction of those interactions can lead to pathological events. Understanding such interactions is therefore a crucial step in the investigation of biological systems and a starting point for drug design. In recent years, experimental studies have been devoted to unravel the principles of biomolecular interactions; however, due to experimental difficulties in solving the three-dimensional (3D) structure of biomolecular complexes, the number of available, high-resolution experimental 3D structures does not fulfill the current needs. Therefore, complementary computational approaches to model such interactions are necessary to assist experimentalists since a full understanding of how biomolecules interact (and consequently how they perform their function) only comes from 3D structures which provide crucial atomic details about binding and recognition processes. In this chapter we review approaches to predict biomolecular complexes, introducing the concept of molecular docking, a technique which uses a combination of geometric, steric and energetics considerations to predict the 3D structure of a biological complex starting from the individual structures of its constituent parts. We provide a mini-guide about docking concepts, its potential and challenges, along with post-docking analysis and a list of related software.

A. Vangone · A.M.J.J. Bonvin (✉)
Computational Structural Biology Group, Bijvoet Center for Biomolecular Research, Faculty of Science—Chemistry, Utrecht University, 3584 Utrecht, The Netherlands
e-mail: a.m.j.j.bonvin@uu.nl

R. Oliva
Department of Sciences and Technologies, University "Parthenope" of Naples, Centro Direzionale Isola C4, 80143 Naples, Italy

L. Cavallo
Kaust Catalysis Center, Physical Sciences and Engineering Division, King Abdullah University of Science and Technology, Thuwal 23955-6900, Saudi Arabia

© Springer Science+Business Media B.V. 2017

D.J. Rigden (ed.), *From Protein Structure to Function with Bioinformatics*,
DOI 10.1007/978-94-024-1069-3_8

Keywords Protein-protein complexes · Protein-peptide complexes · Docking · Searching · Scoring · Data-driven docking · HADDOCK · CAPRI · Flexibility · Binding affinity · PRODIGY · CONSRANK

8.1 Introduction

Biomolecular complexes, such as protein-protein and protein-ligand ones, underlie almost all biological processes in the cell, such as DNA replication, transcription, translation, signaling pathways, immune system response, enzyme inhibition. To implement this wide diversity of (bio)chemical processes, proteins get in touch with other proteins, nucleic acids, sugars, lipids and various other molecules (Jones and Thornton 1996; Alberts 1998). The biological function of a protein is defined by its interactions in the cell. Inappropriate or altered (either inhibited and enhanced) interactions can lead to disease (Stites 1997; Sugiki et al. 2014). For these reasons, research aimed at understanding, disrupting or modulating protein-protein interactions (PPIs) is a crucial step in the investigation of almost all biological processes, ranging from enzyme catalysis and inhibition to signal transduction and gene expression. Accordingly, PPIs are currently receiving considerable attention as targets for rational drug design (González-Ruiz and Gohlke 2006; Metz et al. 2012; Nisius et al. 2012) and as therapeutic agents (Szymkowski 2005; Hwang and Park 2008; Zhou et al. 2013).

In recent years, experimental and theoretical work has been devoted to unravel the principles of protein-protein interactions (Phizicky and Fields 1995; Jones and Thornton 1996). The formation of biological complexes is driven by the free energy of the complex (mostly determined by physicochemical and geometrical interface properties) and the concentration of the protein components. The association of two proteins, in fact, relies on an encounter and possible rearrangement of the interacting surfaces, requiring co-localization in time and space. Generally proteins reside in crowded environments, with many potential binding partners with different surface properties; consequently, during evolution, the interaction surfaces are believed to have evolved to both optimize interaction efficacy and prevent undesired interactions (Ofran and Rost 2003).

In this scenario, it is a must to obtain 3D structural information in order to gain a complete understanding of both the biochemical nature of the process bringing the components together and to facilitate the design of compounds that might influence it. The structural characterization of a protein-protein interface includes in particular the identification of interatomic hydrogen bonds, salt bridges and hydrophobic interactions, the determination of the interaction surface area and possibly the identification of bridging water molecules (Northrup and Erickson 1992; Tsai et al. 1999). The combination of all this information defines the nature of the binding site and of the network of interactions, which makes it possible to pinpoint key residues and contacts for complex formation.

Obtaining 3D structures of biological complexes is therefore of supreme significance for the study of biomolecular interactions and all their possible pharmaceutical and medicinal applications. High-resolution atomic structures are obtained by X-ray crystallography and nuclear magnetic resonance (NMR), while methods like Small-Angle X-ray Scattering (SAXS) (Yang 2014; Chaudhuri 2015) or cryo-Electron Microscopy (cryo-EM) give low-resolution structural data, although the latter, thanks to recent developments in both detector technology and software, is now reaching near atomic resolution (Bai et al. 2015) with, for example, the recent <3 Å high-resolution structure of the ribosome-EF-Tu complex (Fischer et al. 2015). Experimental determination of biomolecules remains, however, difficult, time-consuming and costly (Chruszcz et al. 2010): with X-ray crystallography, dynamics and disorder can impede the crystallization, while (solution) NMR suffers from a size limitation when it comes to studying large macromolecular complexes; and both methods struggle with membrane-resident and membrane-associated complexes. For these reasons, there is relatively little structural information available about biomolecular complexes compared to proteins that exist as single chains or form permanent oligomers (Schreiber and Fersht 1996). As a result, the number of

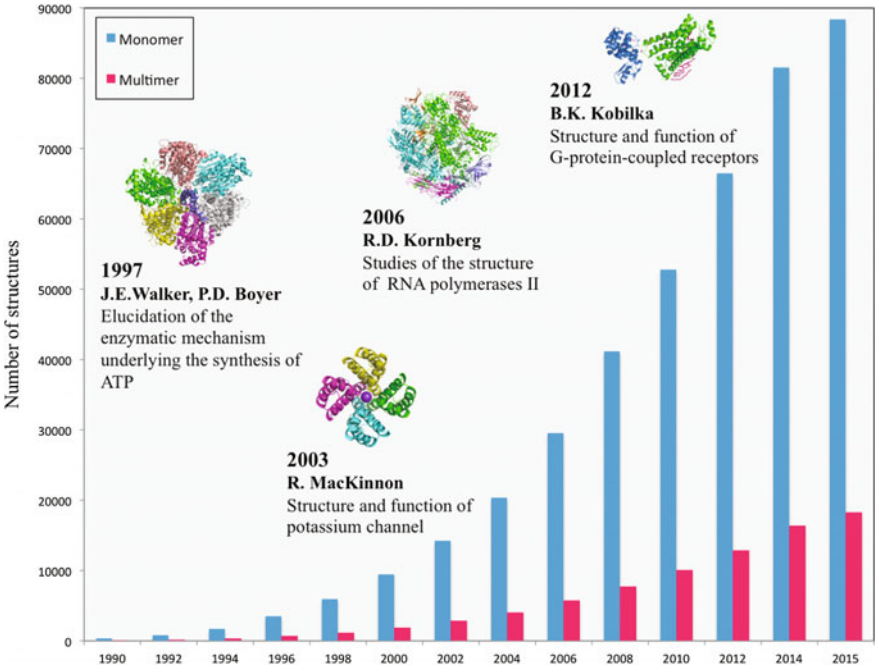


Fig. 8.1 Yearly growth of protein structures number in Protein Data Bank (PDB) from 1990. The PDB was established in 1971, the total number of protein structures grew to 434 in 1990, reaching 106,650 structures on June 2015. The number of single protein structures is reported in *blue*, the number of multiple proteins systems is reported in *magenta*. Some of the Nobel Prized awarded for elucidation of structure (and function) of macromolecular systems are reported in the figure

solved complexes deposited in the Protein Data Bank (PDB) (Bernstein et al. 1977) (http://www.rcsb.org/) is still orders of magnitude smaller than that of individual proteins as shown in Fig. 8.1. Despite this disproportion, the growing number of available experimental structures for protein-protein complexes over the years has allowed statistical studies of the properties and physico-chemical forces that regulate protein-protein interactions (hydrophobicity, hydrogen bonding, electrostatic interactions, van der Waals interactions, and so on). These provide useful information in the development of computational strategies for structure prediction and characterization. Considering the experimental limitations discussed above, computational structural biology is now routinely considered an integral part of research.

Since the pioneering work of Janin and Wodak (Wodak and Janin 1978) who described, more than 30 years ago, the first automated algorithm to predict the 3D interaction between bovine pancreatic trypsin and its natural inhibitor, the docking field (with docking defined as the prediction of protein complexes structures starting from the structures of the free molecules) has advanced considerably (Schlick et al. 2011). The past decades have seen the emergence of a large variety of theoretical algorithms designed to predict the structures of protein-protein and protein-ligand complexes (Smith and Sternberg 2002; Bonvin 2006; Ritchie 2008; Vajda and Kozakov 2009; Moal et al. 2013a).

8.2 Docking

Molecular docking is a computational modeling technique that aims at predicting the 3D structure of a complex (bound form) given the structures of the individual molecules (unbound forms) (Fig. 8.2), hopefully revealing most of the relevant residue-residue contacts involved in the interaction (Smith and Sternberg 2002). It offers a tool for fundamental studies of biomolecular interactions and provides a structural basis for drug design. Docking approaches assume that the native complex is near the global minimum of the energy landscape. In fact, based on the thermodynamic hypothesis, at fixed temperature and pressure the Gibbs free energy of the macromolecule-solvent system reaches its global minimum at the native state of the macromolecule (Ruvinsky and Vakser 2008).

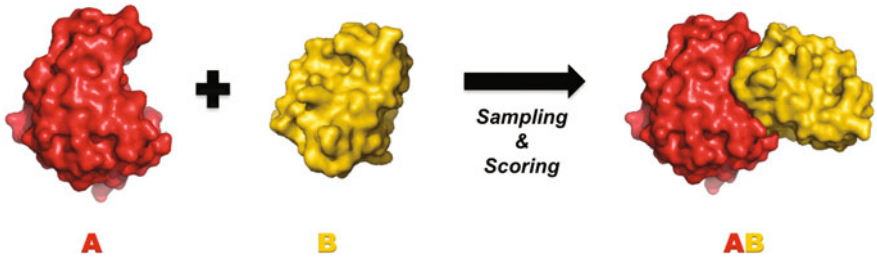


Fig. 8.2 An illustration of protein-protein docking procedure starting from the unbound structures (*A* and *B*), into their final bound form (*AB*). (PDBcode: 1BRS (Buckle et al. 1994), chains *A* and *B*)

Progress in protein-protein docking performance has been monitored over the years with the community wide Critical Assessment of PRedicted Interactions (CAPRI) experiment (Janin et al. 2003; Lensink et al. 2007). Many rounds of blind predictions have highlighted the increasing accuracy of docking methods, in particular for some of them that consistently show good performance (Lensink and Wodak 2013; Lensink et al. 2016) (CAPRI results can be found at the url: http://www.ebi.ac.uk/msd-srv/capri/).

All current docking methods, despite their differences, start from the 3D structures of the unbound components (whether experimentally determined or computationally predicted) and incorporate two crucial steps (Halperin et al. 2002; Vajda and Kozakov 2009):

1. *Searching*, consisting in the generation of thousands of alternative poses to sample the conformational landscape;
2. *Scoring*, consisting in assessing the generated poses using a 'pseudo-energy' function in order to rank them and select the native-like solutions.

This separation into two stages is just one way of describing the docking approach, since sometimes there is no clear separation between these, or they may incorporate multiple different sub-steps. A fundamental point of any docking method is to be computationally efficient both in the search step and in its refinement and scoring scheme in order to be able to evaluate a huge number of candidate solutions and discriminate native-like binding modes from wrong ones in a reasonable computation time.

8.2.1 Step 1: Searching

The search step involves an exhaustive sampling of the conformational space of one protein with respect to the other, resulting in a six-dimensional search (6D) in the case of rigid molecules. Almost all docking programs use a similar approach for the search step: one protein is fixed in space (usually the larger one, named receptor) and the second (named the ligand) is rotated and translated around the first. Various methods have been developed that can efficiently cover the entire conformational space (Vajda and Kozakov 2009) such as:

- *Fast Fourier transforms* (*FFT*)*-based docking*. Despite the huge size of the conformational space to be sampled, the search can be efficiently performed through several FFT calculations, as originally introduced by Katchalski-Katzir et al. (1992). FFT-based methods represent the proteins on a Cartesian grid, with some degree of inter-protein penetration between the ligand and the receptor allowed to account for small conformational changes of mainly side-chains. The shape complementarity is measured using Fourier correlation. Additional terms can be encoded into measure for example electrostatic and hydrophobic matching. Adding such terms in the scoring typically requires multiple FFT

evaluation per pose. Widely used nowadays (Comeau et al. 2007; Pierce et al. 2011; Jiménez-García et al. 2013), such methods efficiently perform an exhaustive rigid-body search.

- *Geometric hashing docking*. First developed in the area of computer vision and implemented in docking by Wolfson and colleagues (Fischer et al. 1992; Mashiach et al. 2010b), this approach allows efficient searching by dividing the biomolecular surface into patches and matching them across the interacting molecules.
- *Spherical harmonics-based docking*. Pioneered by Ritchie and co-workers (Ritchie and Kemp 2000; Macindoe et al. 2010), this uses spherical polar Fourier correlations to accelerate the search, describing the protein shapes as a combination of spherical harmonic functions and calculating the relative orientations via scalar products of rotated and translated coefficient vectors.

Those methods can evaluate very large numbers of interaction poses in a relatively short time amount, making efficient use of computational resources (CPU cores), but other algorithms, although less computationally efficient, can reach high performance as well. HADDOCK (Dominguez et al. 2003) for example uses a gradient-based search method in Cartesian space (rigid-body energy minimization), targeting specific patches on the molecular surface deemed favorable by the energy function used. ATTRACT (Zacharias 2005) pioneered normal-mode analysis into the searching phase and SwarmDock (Moal and Bates 2010) incorporated it into a Particle Swarm Optimization meta-heuristic to perform docking while optimizing conformation, position and orientation simultaneously.

Table 8.1 reports a list of the top-performing docking approaches in CAPRI. For a more complete compilation of the existing docking programs see the latest CAPRI assessment, for recent reviews on the topic see (Moreira et al. 2010; Rodrigues and Bonvin 2014).

8.2.2 Step 2: Scoring

While the goal of sampling is to generate a set of poses, ideally with the highest number of correct conformations (although non-exhaustive sampling might not allow to do that), the goal of scoring is to single out the near-native ones within the pool of models generated. Due to the high complexity of the energetics governing the interaction, scoring is a critical step in docking: for such a reason, a separate challenge to test scoring methods has been added to CAPRI (Lensink et al. 2007).

In an ideal scoring process, one or more descriptors of the docking poses allows to derive a score, which nicely correlates with the model quality (in terms of similarity to the true solution), thus unambiguously distinguishing correct solutions from incorrect ones (Fig. 8.3a–c).

However, current scoring functions are far from reaching perfection, although the CAPRI experiment shows that they are constantly improving (Lensink and

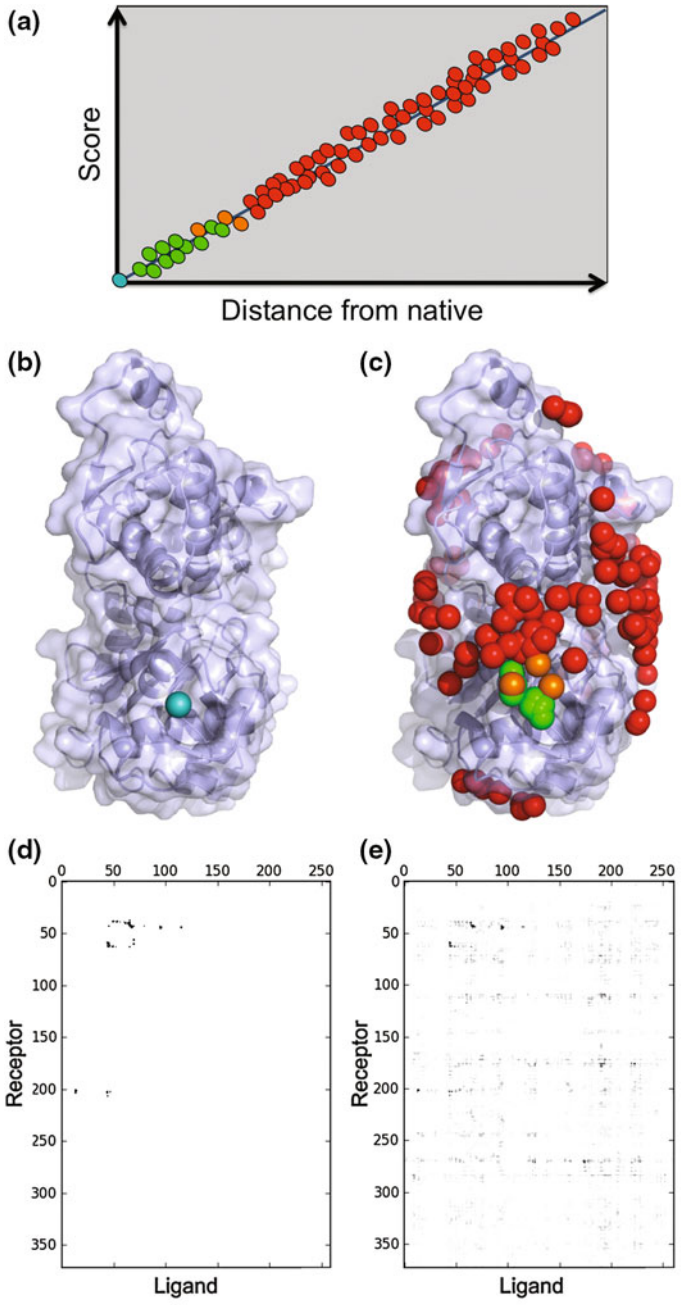
Table 8.1 List of protein-protein docking algorithms

Program name	Searching protocol details	Web-server
ATTRACT (Zacharias 2005)	Energy minimization in translational and rotational space using NMA to allow conformational changes upon binding	None
ClusPro (Comeau et al. 2004b)	Rigid-body search via FFT	http://cluspro.bu.edu
GRAMM-X (Tovchigrechko and Vakser 2006)	Grid-based FFT rigid-body docking	http://vakser.compbio.ku.edu/resources/gramm/grammx/
HADDOCK (de Vries et al. 2010)	Rigid-body energy minimization followed by semi-flexible refinement in torsion angle space	http://haddocking.org
HEX server (Macindoe et al. 2010)	Spherical harmonics, polar FFT	http://hexserver.loria.fr
PatchDock (Schneidman-Duhovny et al. 2005)	Geometric hashing	http://bioinfo3d.cs.tau.ac.il/PatchDock
pyDock (Cheng et al. 2007)	Rigid-body search via FFT	http://life.bsc.es/servlet/pydock/home
RosettaDock (Lyskov and Gray 2008)	Low-resolution, rigid-body MC search	http://antibody.graylab.jhu.edu/docking
SwarmDock (Moal and Bates 2010)	Local docking and particle swarm optimization of position and orientation, NMA	http://bmm.cancerresearchuk.org/~SwarmDock/
ZDOCK (Chen et al. 2003)	FFT-based rigid-body serach	http://zdock.umassmed.edu

FFT fast Fourier transform, *MC* Monte Carlo, *NMA* Normal Mode Analysis

Wodak 2010, 2013). Traditionally, scoring functions for protein-protein docking poses rely on two approaches, both of them widely tested in CAPRI blind tests where they were shown to perform competitively. The first approach uses a linear combination of energy terms, which can be physics-based and/or empirical, such as van der Waals, electrostatics and desolvation energies, buried surface area and terms accounting for shape complementarity (Gray et al. 2003; Cheng et al. 2007; de Vries et al. 2007; Venkatraman et al. 2009; Gong et al. 2010). Weights used in the linear combination are usually optimized to distinguish native-like solutions from non native-like ones.

The second traditional approach is statistics-based or "knowledge-based", as it uses properties derived from experimental structures of protein-protein complexes. Such properties are usually embodied in atom-atom or residue-residue potentials, derived from the statistical occurrences observed in the analyzed database of complexes by means of an inverse Boltzmann equation (the higher the population, the lower the energy) (Moont et al. 1999; Jiang et al. 2002; Lu et al. 2003; Huang and Zou 2008; Kowalsman and Eisenstein 2009; Khashan et al. 2012).


(a)
Score
Distance from native
(b)
(c)
(d)
Receptor
Ligand
0
50
100
150
200
250
300
350
(e)
Receptor
Ligand

◂**Fig. 8.3** **a** Scheme of an ideal scoring process: the score strongly correlates with the distance of the model from the native structure (same color scheme of **b** and **c**). **b**, **d** Actin-DNase I complex [PDB ID: 1ATN (Kabsch et al. 1990)]: surface representation of the receptor (actin, *light blue*) with sphere representation of the center of mass of the ligand (DNase I, teal) interface (**b**) and intermolecular contact map generated by COCOMAPS server (Vangone et al. 2011) (**d**). **c**, **e** An ensemble of 185 predicted docking poses for 1ATN: surface representation of the receptor (*light blue*) with sphere representation of the centre of mass of the model ligand interface (*green*: correct; *red*: incorrect; *orange*: intermediate **c** and 'consensus map' (**e**)

This approach (Viswanath et al. 2013), like the energy-based one, can also take advantage of a training process on extended sets of docking poses, to distinguish correct from incorrect solutions.

The above approaches are, however, not mutually exclusive and in several scoring functions they are indeed combined into a hybrid approach (Pierce and Weng 2007; Andrusier et al. 2007; Vreven et al. 2011). Some of these methods also take advantage of machine learning algorithms in the training process to derive best coefficients to combine the different scoring terms (Champ and Camacho 2007; Fink et al. 2011).

It is important to mention that, as now generally accepted, a native structure is not an isolated event in the global energy landscape and thus native-like models are expected to form "funnels", i.e. clusters of similar low energy solutions. The clustering is often done based on RMSD comparisons between models, but can also efficiently be performed based on the fraction of common contact as introduced by Rodrigues et al. (2012). On these bases, some scoring methods try to characterize funnel-like energy structures on the global energy landscape (Kozakov et al. 2008; London and Schueler-Furman 2008; Moal and Bates 2010; Torchala et al. 2013), also using the concept of transient complex (Qin and Zhou 2013), while others, after scoring, perform a clustering of models in an ensemble of low-energy conformations and select the top ones based on the cluster population (Comeau et al. 2004a). The above approaches implicitly use the concept of consensus, i.e. similarity within an ensemble of docking models. More recently, a "pure" consensus method, CONSRANK, based on the frequency of inter-residue contacts in an ensemble of docking solutions, has been proposed for the ranking of docking solutions. Blind testing in CAPRI Round 30 showed it to perform competitively with classical energy- and knowledge-based approaches.

Other approaches to the scoring include methods using evolutionary information (Tress et al. 2005; Andreani et al. 2013; Xue et al. 2014) and methods using experimental information on the complex, when available (de Vries et al. 2007; Gajda et al. 2010; Moreira et al. 2015). For recent reviews on the topic see (Moal et al. 2013a, b).

8.2.3 Data-Driven Docking

Although important progresses in the searching and scoring procedures have been achieved, one of the most useful approaches to improve the quality of the docking simulations is the use of biological information about the interaction regions of the complex when available. As clearly inferable from the latest CAPRI assessment reports (Lensink and Wodak 2013), information (experimentally or computationally derived) on regions and residues involved in the interaction is one of the key points for the improvement of a docking simulation. Many docking programs offer the possibility to integrate data, for example as a scoring bias or as a filter to select solutions at the end, to exclude from the search regions not involved in the interaction or to drive the docking towards the areas known to be involved.

HADDOCK, one of the top performing docking program in the last CAPRI rounds (Lensink and Wodak 2013; Lensink et al. 2016), is the pioneer of data (or information)-driven docking and, in contrast to other docking methods that usually incorporate data at some stage of the protocol, HADDOCK is the only program that uses such data throughout the entire protocol (see Sect. 8.3.1). In HADDOCK the data (experimental and/or predicted) are incorporated into the calculation as an additional restraint energy term, as distance [i.e. mutagenesis, nuclear Overhauser effect, chemical cross-links, electron paramagnetic resonance distances, or even co-evolution-derived distances (Hopf et al. 2014)], orientation [e.g. NMR residual dipolar coupling (van Dijk et al. 2005), pseudo-contact shifts (Schmitz and Bonvin 2011)] or relaxation anisotropy (van Dijk et al. 2006) restraints (Schmitz et al. 2012) or even recently shape information [e.g. cryo-EM data (van Zundert et al. 2015)]. HADDOCK implements the concept of highly ambiguous distance restraints to incorporate information which define patches of interacting residues but no specific pairwise interactions between them (like in the case of NMR chemical shift perturbations).

Most traditionally successful methods in CAPRI also offer the possibility to integrate data into the protocol: FFT-based approaches [ClusPro (Comeau et al. 2004b), GRAMM-X (Tovchigrechko and Vakser 2006), pyDock (Cheng et al. 2007), ZDOCK (Chen et al. 2003) and HEX (Macindoe et al. 2010)] use data to bias the score toward models that satisfy it, or as a filter at the end. Thus, SwarmDock (Moal and Bates 2010) uses the data to pre-orientate the molecules such as the identified or predicted interfaces face each other while PatchDock (Schneidman-Duhovny et al. 2005) allows the definition of interacting or non-interacting regions, and also the setting of distance constraints. The RosettaDock (Lyskov and Gray 2008) program includes data as distance-filters to bias the Monte Carlo search whereas the most recent version of ATTRACT now also supports ambiguous distance restraints and allows docking using Cryo-EM density maps (de Vries and Zacharias 2012). Finally, ZDOCK (Chen et al. 2003) includes specific knowledge-based scoring functions in the protocol.

The quality of models coming out of data-driven docking approaches will depend on the quality of the data used. The most common experiments that give

information about interface residues involved in the binding are mutagenesis, NMR chemical shift perturbation and cross-saturation and hydrogen/deuterium exchange, while techniques such as nuclear Overhauser effect in NMR and cross-link experiments in mass spectromety provide distance information. This experimental information can be complemented or even replaced by bioinformatics predictions. These are mostly based on the study of sequence/structure conservation of key residues, co-evolution principles allowing to derive residue pairs in predicted proximity, propensity of residues to be surface-exposed, or the combination of such information as consensus and partner-specific methods (Neuvirth et al. 2004; de Vries et al. 2006; Porollo and Meller 2006; Negi et al. 2007; Qin and Zhou 2007; Ashkenazy et al. 2010; Ahmad and Mizuguchi 2011; de Vries and Bonvin 2011; Zhang et al. 2011; Zellner et al. 2012; Xue et al. 2014). However, the predictions have to be analyzed critically and combined with experimental information when available.

8.3 The Challenges of Docking: Flexibility and Binding Affinity

8.3.1 Changes upon Binding: The Flexible Docking Challenge

Although docking programs have improved their performance over the years according to CAPRI, predicting the structure of biomolecular complexes remains a difficult problem with, at the moment, two major challenges: the identification of correct solutions within a pool of models (scoring) and the treatment of proteins with substantial conformational change upon binding (flexibility).

Proteins are not rigid, and during the association process they usually undergo conformational changes that include both backbone and side-chains movements (Betts and Sternberg 1999). As a result, the conformation of the proteins within the complex/bound form might be different from the one they have in the free form. Therefore, incorporating flexibility in docking algorithms is necessary to predict the native associations and reach high accuracy of the solutions. In the cases where structural changes occurring upon binding are minimal, the difference between bound and free forms can be neglected so the rigid body docking is sufficient. A major problem here is that, in general, one can not know a priori if conformational changes will take place or not, nor their extent. Properly dealing with flexibility in docking is therefore one of the main challenges in the field (Smith et al. 2005a; Bonvin 2006; Lensink et al. 2007).

A major problem of incorporating flexibility in docking, compared to performing rigid-body docking only, is the considerable increase in the number of degrees of freedom and, consequently, in the search space. This also often goes together with a higher rate of false-positive solutions, since all might be refined to some local

energy minimum, which thus complicates the identification of correct solutions (Andrusier et al. 2008).

Flexibility can be introduced at several levels:

- *Implicitly*. Implicit flexibility can be incorporated by soft-docking, by smoothing the protein surface or allowing some degrees of interpenetration or overlap of atoms (Palma et al. 2000; Heifetz and Eisenstein 2003) [although one of the drawbacks of such an approach is that severe steric clashes can be introduced (Smith et al. 2005b)], or with cross-docking by performing rigid-body docking of ensembles of conformations, taken for example from NMR structures or MD simulations or any other conformational sampling method (de Groot et al. 1997). Depending on the implementation this can lead to a significant increase in computing time. It has, on the other hand, the advantage that rather large conformational changes can be pre-sampled in that way.
- *Explicitly*. In the past few years, flexibility has been explicitly introduced into the docking process by allowing side-chains and/or backbone to move. The docking programs allowing side-chain flexibility (Fernández-Recio et al. 2003; Zacharias 2005; de Vries et al. 2007; Lyskov and Gray 2008; de Vries et al. 2010) use different approaches, like Monte Carlo (MC) optimization of the ligand (ICM-DISCO) (Fernández-Recio et al. 2003), sampling the known populated rotamers of the side-chains followed by energy minimization steps (ATTRACT) (Zacharias 2005), using MD simulated annealing for refinement of both receptor and ligand side-chains (HADDOCK) (de Vries et al. 2010), or repacking and optimization of side-chains in a MC search (RosettaDock) (Lyskov and Gray 2008).

In contrast with side-chains flexibility, which is easier to model, backbone flexibility is currently one of the main challenges in docking.

In addition to conformational changes upon binding, some programs have been developed to tackle the challenge of large domain motions, such as the flexible multi-domain docking approach proposed by Karaca and Bonvin (Karaca and Bonvin 2011) that can describe large domain motion-type conformational changes. The proper treatment of flexibility in protein-protein docking and also for peptide docking (see Sect. 8.4) remains an active area of research. In small-molecule docking (like protein-ligand docking), in which flexibility plays a major role, the problem is more tractable, but no less challenging (Brooijmans and Kuntz 2003; Erickson et al. 2004).

8.3.2 *The 'Perfect' Scoring Function and the Binding Affinity Problem*

Scoring approaches typically attempt to fish the most likely model of a complex from a set of poses but are not designed to predict how strongly the proteins bind,

i.e. their free energy of binding $\Delta G_{binding}$, or whether they bind at all [as showed by cross-docking simulations (Sacquin-Mora et al. 2008; Wass et al. 2011a, b, Martin and Lavery 2012)]. That is because scoring (ranking) and binding affinity prediction (ΔG) are two different things. The $\Delta G_{binding,}$ or Gibbs free energy of the complex can be determined by measuring the dissociation constant as:

$$\Delta G = RT \ln K_d$$

where R is the gas constant, T is the temperature and K_d is the dissociation constant. It reflects the natural inclination of molecules entities to associate and is a key thermodynamic quantity for understanding recognition and association phenomena, and possible dysfunctions thereof.

Accurately predicting binding free energies with a general scoring function, while a very ambitious goal, would revolutionize the efficiency of docking methods. Different methods aimed at predicting binding affinity in protein complexes have been proposed throughout the years, taking into account different structural and energetic features of the complex and varying greatly in terms of accuracy and computational cost. Based on the initial observation of Chothia and Janin (1975) in the 1970s and described by Horton and Lewis (1992) in 1992, the buried surface area (BSA), i.e. the surface that is buried upon complex formation, has been the first descriptor to be related to the binding affinity. Since then, many methods have been proposed. Exact methods such as free energy perturbation and thermodynamics integration can be very accurate, but due to their computational costs their application is extremely limited (mostly to low throughput studies and mainly for small drug binding or mutations). Methods based on empirical functions (empirical, force field-based potentials, statistical potentials, scoring functions used in docking) are much faster (Jiang et al. 2002; Ma et al. 2002; Zhang et al. 2005; Audie and Scarlata 2007; Su et al. 2009; Bai et al. 2011; Qin et al. 2011; Moal and Bates 2012; Tian et al. 2012; Luo et al. 2014; Kastritis et al. 2014). However, even if some have been very successful on small training sets (Horton and Lewis 1992; Audie and Scarlata 2007), most published models still fail to systematically predict the binding affinity (Kastritis and Bonvin 2010, 2013a, b) for large datasets or discriminate between binders from non-binders (Sacquin-Mora et al. 2008; Fleishman et al. 2011). Usually, factors such as conformational changes occurring upon binding, allosteric regulation, solvent and co-factor effects, which may contribute to the binding strength, are neglected, which entails their main weaknesses. Using a large binding affinity benchmark consisting of 144 complexes (Kastritis et al. 2011) [updated version of the benchmark now available in (Vreven et al., 2015)], Kastritis et al. (2014) demonstrated that non-interacting surface properties like percentages of charged and polar residues do also contribute to binding affinity. These rather surprising finding were corroborated in a recent study by Cazal et al. in which this contribution from the non-interaction surface was reproduced (Marillet et al. 2015). New advances were made by Vangone and Bonvin (Vangone and Bonvin 2015; Xue et al. 2016) who recently showed that the network of contacts made at the

interface in a protein-protein complex is a better structural descriptor of the binding affinity then the BSA.

8.4 Protein-Peptide Docking

In eukaryotes more than 40% of the interactions are estimated to be mediated by peptides, for example in signal transduction, protein degradation, transcription regulation and immune response (Petsalaki and Russell 2008). Due to their involvement in many biological pathways, peptide interactions are implicated in many diseases and in cancer (Petsalaki and Russell 2008; Naider and Anglister 2009), making them of high interest in the development of new therapeutics and for drug design (Vaara 2009). Indeed, alongside small-molecule inhibitors, peptides are large enough to competitively inhibit protein-protein interactions and can mimic protein binding domains. However, the experimental structure determination of protein-peptide recognition remains a challenging task also in this case, mainly due of two factors: peptides are highly flexible and they usually show transient interactions with the substrate. From a structural point of view, peptides are short chains ranging from 5 to 30 amino acids, often lacking a well-defined fold in their free form. They might not necessarily be independent molecules, but can appear as disordered regions of proteins (for example at termini), and they can show multiplicity in their interaction, as for example in the case of the tumor suppressor p53 (Russell and Gibson 2008).

Complementary computational prediction methods like docking are therefore urgently required to model those systems, as also reflected by the recent addition of protein-peptides cases in CAPRI. Peptides' peculiar characteristics represent, however, a unique challenge for computational predictions. Conventional protein-protein docking struggles with the high flexibility of peptides while ligand-docking protocols have only been successfully applied to short peptide, due to the significant higher number of peptide rotatable bonds than in drug-like small molecules (Hetényi and van der Spoel 2002; Sousa et al. 2006; Rubinstein and Niv 2009; London et al. 2013). Over the last years a number of new algorithms or ad hoc adaptations have been developed with the aim of modelling protein-peptide complexes (Petsalaki et al. 2009; Antes 2010; Raveh et al. 2010; Ben-Shimon and Eisenstein 2010; Raveh et al. 2011; Donsky and Wolfson 2011; Dagliyan et al. 2011; Trellet et al. 2013; Verschueren et al. 2013; Lavi et al. 2013; Ben-Shimon and Niv 2015; Kurcinski et al. 2015). Similarly to protein-protein docking, there are two main steps: (i) identification of the binding site on the protein surface (which might include the use of experimental or bioinformatics data when available; see also Chap. 10) and (ii) docking and refinement of the peptide into the binding site.

Several high-resolution approaches have been successfully applied to unbound protein-peptide datasets. FlexPepDock (Raveh et al. 2010, 2011), the first generic algorithm released to model protein-peptide complexes, uses fragment-based sampling for the generations of different peptide backbone conformations, and then

allows full flexibility of the peptide and to the protein side chains within a defined docking site. HADDOCK (Trellet et al. 2013) overcomes the problem of the indetermination of the peptide free structure by using as input an ensemble of three different conformation of the peptide: alpha-helix, polyproline-II and extended. Taken together, these conformations cover about 80% of the observed protein-peptide structures in the PDB (Diella et al. 2008). This is followed by flexible refinement step in which more flexibility is given to the peptide. This protocol mimics the conformational selection mechanism/induced fit recognition mechanism (Weikl and Deuster 2009; Hammes et al. 2009; Csermely et al. 2010; Changeux and Edelstein 2011). Lately Blaszczyk and co-workers implemented CABS-dock, an ab initio protein-peptide modelling approach (Kurcinski et al. 2015) that performs the search for the binding site and docking (giving flexibility) simultaneously using a coarse grained representation of the system. Additional ab initio algorithms or tools aimed to predict candidate sites of interaction on the protein surface (Trabuco et al. 2012) have been implemented lately to overcome the lack of information on the peptide binding site (Ben-Shimon and Niv 2015). which is, in addition to the high flexibility of peptides, the main challenge to overcome in protein-peptide docking.

Despite the recent progresses, this is a field that still is its infancy with further development and extensive evaluation required, for example in CAPRI challenges. For further information please check (London et al. 2013; Trellet et al. 2015).

8.5 Post-docking: Interface Prediction from Docking Results and Use of Docking-Derived Contacts for Clustering and Ranking

It is now over ten years since Fernandez-Recio et al. (2004) proposed to predict residues at the protein-protein interface from results of docking simulations (Fig. 8.4a). They analyzed the rigid-body docking energy landscape in several training sets, in search of protein recognition areas, showing that the energy profile for the ensemble of found docked poses can be used to determine accurately interaction sites on protein surfaces. In particular, they defined a normalized interface propensity (NIP) parameter, which represents the tendency of a given residue to be located at the interface, based on the buried surface area in docking poses from rigid docking simulations. Based on the NIP definition, more recently Fernadez-Recio and Grosdidier derived a method for hot-spot residues prediction, achieving up to 80% positive predictive value (Grosdidier and Fernández-Recio 2008).

In 2010, based on their experience as assessors in the CAPRI experiment, Lensink and Wodak confirmed the potential of docking techniques for the prediction of protein interfaces (Lensink et al. 2014). By analyzing docking models submitted in CAPRI by 76 participants for 46 interfaces in 20 targets, they found

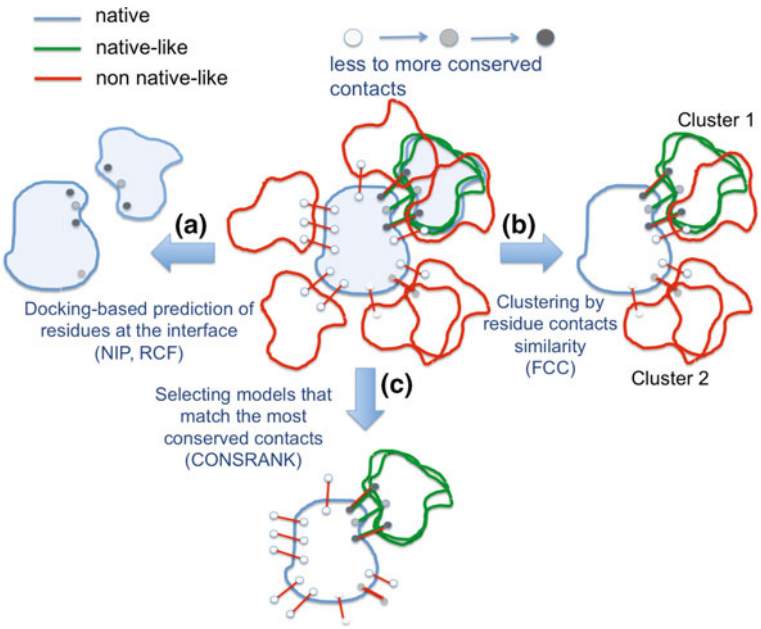


Fig. 8.4 Scheme of the use of docking results for: **a** predicting residues at the interface, and for: **b** models clustering and **c** ranking. Figure adapted from Oliva et al. (2013)

that the best performing groups were able to predict residues at the interface with precision and sensitivity levels around 60% for the majority of the analyzed cases, thus reaching a performance competitive with the most successful non-docking based methods in the field. The main finding of this analysis was thus that, apparently, models ranked highly by docking procedures are more enriched in correct interfaces than in correct complexes. In fact, prediction of correct interfaces is also contributed by incorrect (according to the CAPRI assessment) models, which were found to feature one quarter of correct interfaces (with precision and sensitivity above 50%), contributing to 70% of the overall correct interface predictions.

de Vries and Bsonvin also showed that, after improving the performance of docking predictions with HADDOCK by a consensus monomer-based interface prediction, the interface prediction itself could be further improved by post-prediction based on top-scored docking results (de Vries and Bonvin 2011). Following these findings, Weng and colleagues recently developed RCF (residue contact frequency), another method to predict interface residues from models generated by docking algorithms (Hwang et al. 2014) (Fig. 8.4a). They used RCF to predict the binding interfaces of proteins that bind to multiple partners, finding that it correctly predicts interface residues unique for the respective binding partners. They also showed that the combination of RCF with monomer-based interface prediction methods, through a support vector machine, improved performance compared to both separated approaches. RCF was also used by the Weng's group to

analyze their docking results in the CAPRI rounds 20–26, where selection of final models for submission was in fact guided by RCF (Vreven et al. 2013).

Besides the identification of residues likely involved in the interface from results of docking simulations, specific inter-residue contacts observed in docking poses have been recently used to guide their clustering, analysis and ranking. As the native structure of a complex is not expected to be an isolated position in the energy landscape, docking experiments often incorporate one clustering step in their protocols, which is classically based on time-consuming (live memory, RAM) and size-dependent RMSD measures (Janin 2010). In this context, Bonvin and colleagues proposed the use of the fraction of common contacts (FCC) within models as a similarity description to base their clustering on (Rodrigues et al. 2012) (Fig. 8.4b). They showed that FCC is an efficient measure of the structure similarity for protein complexes, greatly reducing the computation time while generating clusters of similar quality with the state-of-the art RMSD-based methods. Further, it is particularly suited for flexible docking approaches, multicomponent assemblies and heterogeneous systems like protein-DNA complexes.

Oliva and colleagues proposed instead to analyse an ensemble of protein-protein docking models, by deriving a consensus based on the conservation within them of the inter-residue contacts (Vangone et al. 2012). Such a consensus can also be visualized as a "consensus contact map", i.e. an intermolecular contact map where the conservation of contacts is reported on a gray scale (see an example in Fig. 8.3e, compared to the intermolecular contact map of the corresponding crystal structure, Fig. 8.3d). Analysis of prediction sets of docking models for seven CAPRI targets showed that a significant fraction of native contacts was included within the contacts with highest conservation rate, even in the cases where only a small percentage of solutions were correct. This suggests that incorrect models can contribute to the correct prediction not only of residues, but also of specific inter-residue contacts at the complexes interface. A natural extension of this approach was the development of CONSRANK (CONSensus-RANKing) (Oliva et al. 2013; Vangone et al. 2013), a consensus method for the scoring of docking models, which ranks models based on their ability to match the most conserved contacts in the ensemble they belong to (Fig. 8.4c).

8.5.1 Web Tools for the Post-docking Processing

As discussed previously (Sect. 8.2.2), a scoring/filtering step is normally included in a docking procedure. However, to date no program can provide a single docking solution with a high enough confidence to be correct. Docking programs instead generally provide the user with an ensemble of models, corresponding to a subset (usually refined) of the solutions they generated in the conformational sampling step, which possibly contain native-like models. These models have thus to be analyzed to attempt to single out the correct ones. Some tools have been specifically devoted to the post-docking processing, i.e. the analysis, scoring and ranking of

Table 8.2 List of available web servers for the post-docking processing

Server name	Algorithm	Analyses	URL
CCharPPI (Moal et al. 2015)	Energy/knowledge-based	109 parameters including FireDock, PyDock, RosettaDock, SIPPER & ZRANK scores	http://life.bsc.es/pid/ccharppi/
CONSRANK (Chermak et al. 2014)	Consensus-based	Contacts analysis and visualization; re-scoring	https://www.molnac.unisa.it/BioTools/consrank/
DOCKRANK (Xue et al. 2014)	Evolution-based	Prediction of the interface; re-scoring	http://einstein.cs.iastate.edu/DockRank/
FastContact (Champ and Camacho 2007)	Energy/knowledge-based	Energy minimization; prediction of residue contact free energies; re-scoring	http://structure.pitt.edu/servers/fastcontact/
FiberDock (Mashiach et al. 2010a, 2008)	Energy/knowledge-based	Flexible refinement; re-scoring	http://bioinfo3d.cs.tau.ac.il/FiberDock/ http://bioinfo3d.cs.tau.ac.il/FireDock/
FILTREST3D (Gajda et al. 2010)	User-defined restraints from experimental data	Re-scoring	http://filtrest3d.genesilico.pl/filtrest3d/
FunHunt (London and Schueler-Furman 2008)	Energy-based	Characterization of local energy landscape	http://funhunt.furmanlab.cs.huji.ac.il/
PROCOS (Fink et al. 2011)	Energy/knowledge-based	Re-scoring	http://compdiag.uni-regensburg.de/procos/

models representing the output of docking programs. Several of these post-processing tools are publicly available as web servers and are listed in Table 8.2, together with the corresponding URLs. The scoring approaches they mainly rely on, reflecting the ones described above (Sect. 8.2.2), are also reported in Table 8.2.

8.6 Concluding Remarks

In view of the growing interest in protein-protein interactions for pharmaceutical and medical applications, and the persistent disproportion between experimental structures available for single proteins and multiple protein systems, the relevance of molecular docking as the method of choice for modelling the structure of protein-protein complexes is set to increase.

In the last 15 years, the CAPRI blind assessment has shown that docking techniques can be successfully applied to a variety of cases, with biological information on the interface, when available, further improving results, by driving the search of allowed configurations and helping in filtering out incorrect solutions. At the same time, the development of web servers characterized by a user-friendly interface, for performing both docking predictions and post-docking analyses, is in fact making the use of this technique accessible also to a non-specialized audience.

That notwithstanding, to further extend its confident applicability to critical cases, protein-protein docking needs to face a number of challenges in the near future. First of all, the flexibility of the two interacting proteins has to be more confidently coped with, possibly by exploring novel approaches to the sampling of the conformational space. In this regard, it is remarkable that, in the latest CAPRI rounds, scorer groups have been shown to achieve overall a better prediction performance than predictor groups. In other words, the same groups typically recognized more correct solutions from ensembles of models obtained by a variety of techniques, rather than from their own generated models ensemble. This suggests that the bottleneck in a docking procedure still resides in an efficient sampling of the conformational space and that application of different docking strategies to the target system could help overcoming the issue—a kind of consensus docking strategy using various approaches. Other challenges that need to be addressed include a reliable identification of native-like models, with possibly an estimation of the binding affinity of the complex. In addition, when one of the interacting partners is a peptide, docking protocols have to deal with further challenges, such as the high flexibility and the undefined folding of peptides.

Finally, the prediction of the 3D structure of a biomolecular complex, which is fundamental for understanding biological processes, can also help in advancement of related fields. Indeed, it is becoming increasingly clear that results of docking simulations can also be used as an intermediate step for other applications, such as the interface prediction itself, which can be very valuable for experimentalists to guide their work (e.g. to target mutagenesis to interesting regions on the surface of a protein). Further, three-dimensional structural information can also be useful to identify pair on interacting proteins/peptide motifs with the final goal to predict the full network of protein-protein interactions governing the cells (Zhang et al. 2012; Chen et al. 2015).

Acknowledgements AV was supported by Marie Skłodowska-Curie Individual Fellowship H2020 MSCA-IF-2015 [BAP-659025]. RO was supported by Regione Campania [LR5-AF2008].

References

Ahmad S, Mizuguchi K (2011) Partner-aware prediction of interacting residues in protein-protein complexes from sequence data. PLoS ONE 6:e29104. doi:10.1371/journal.pone.0029104

Alberts B (1998) The cell as a collection of protein machines: preparing the next generation of molecular biologists. Cell 92:291–294. doi:10.1016/s0092-8674(00)80922-8

Andreani J, Faure G, Guerois R (2013) InterEvScore: a novel coarse-grained interface scoring function using a multi-body statistical potential coupled to evolution. Bioinformatics 29:1742–1749. doi:10.1093/bioinformatics/btt260

Andrusier N, Mashiach E, Nussinov R, Wolfson HJ (2008) Principles of flexible protein-protein docking. Proteins 73:271–289. doi:10.1002/prot.22170

Andrusier N, Nussinov R, Wolfson HJ (2007) FireDock: fast interaction refinement in molecular docking. Proteins 69:139–159. doi:10.1002/prot.21495

Antes I (2010) DynaDock: a new molecular dynamics-based algorithm for protein-peptide docking including receptor flexibility. Proteins 78:1084–1104. doi:10.1002/prot.22629

Ashkenazy H, Erez E, Martz E, Pupko T, Ben-Tal N (2010) ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. Nucleic Acids Res 38:W529–W533. doi:10.1093/nar/gkq399

Audie J, Scarlata S (2007) A novel empirical free energy function that explains and predicts protein-protein binding affinities. Biophys Chem 129:198–211. doi:10.1016/j.bpc.2007.05.021

Bai H, Yang K, Yu D, Zhang C, Chen F, Lai L (2011) Predicting kinetic constants of protein-protein interactions based on structural properties. Proteins 79:720–734. doi:10.1002/prot.22904

Bai X-C, McMullan G, Scheres SHW (2015) How cryo-EM is revolutionizing structural biology. Trends Biochem Sci 40:49–57. doi:10.1016/j.tibs.2014.10.005

Ben-Shimon A, Eisenstein M (2010) Computational mapping of anchoring spots on protein surfaces. J Mol Biol 402:259–277. doi:10.1016/j.jmb.2010.07.021

Ben-Shimon A, Niv MY (2015) AnchorDock: blind and flexible anchor-driven peptide docking. Structure 23:929–940. doi:10.1016/j.str.2015.03.010

Bernstein FC, Koetzle TF, Williams GJ, Meyer EF, Brice MD, Rodgers JR, Kennard O, Shimanouchi T, Tasumi M (1977) The Protein Data Bank: a computer-based archival file for macromolecular structures. J Mol Biol 112:535–542. doi:10.1016/s0022-2836(77)80200-3

Betts MJ, Sternberg MJ (1999) An analysis of conformational changes on protein-protein association: implications for predictive docking. Protein Eng 12:271–283. doi:10.1093/protein/12.4.271

Bonvin AMJJ (2006) Flexible protein-protein docking. Curr Opin Struct Biol 16:194–200. doi:10.1016/j.sbi.2006.02.002

Brooijmans N, Kuntz ID (2003) Molecular recognition and docking algorithms. Annu Rev Biophys Biomol Struct 32:335–373. doi:10.1146/annurev.biophys.32.110601.142532

Buckle AM, Schreiber G, Fersht AR (1994) Protein-protein recognition: crystal structural analysis of a barnase-barstar complex at 2.0-A resolution. Biochemistry 33:8878–8889. doi:10.1021/bi00196a004

Champ PC, Camacho CJ (2007) FastContact: a free energy scoring tool for protein-protein complex structures. Nucleic Acids Res 35:W556–W560. doi:10.1093/nar/gkm326

Changeux J-P, Edelstein S (2011) Conformational selection or induced fit? 50 years of debate resolved. F1000 Biol Rep 3:19. doi:10.3410/B3-19

Chaudhuri BN (2015) Emerging applications of small angle solution scattering in structural biology. Protein Sci 24:267–276. doi:10.1002/pro.2624

Chen R, Li L, Weng Z (2003) ZDOCK: an initial-stage protein-docking algorithm. Proteins 52:80–87. doi:10.1002/prot.10389

Chen TS, Petrey D, Garzon JI, Honig B (2015) Predicting peptide-mediated interactions on a genome-wide scale. PLoS Comput Biol 11:e1004248. doi:10.1371/journal.pcbi.1004248

Cheng TM-K, Blundell TL, Fernández-Recio J (2007) pyDock: electrostatics and desolvation for effective scoring of rigid-body protein-protein docking. Proteins 68:503–515. doi:10.1002/prot.21419

Chermak E, Petta A, Serra L, Vangone A, Scarano V, Cavallo L, Oliva R (2014) CONSRANK: a server for the analysis, comparison and ranking of docking models based on inter-residue contacts. Bioinformatics 31:1481–1483. doi:10.1093/bioinformatics/btu837

Chothia C, Janin J (1975) Principles of protein–protein recognition. Nature 256:705–708. doi:10.1038/256705a0

Chruszcz M, Domagalski M, Osinski T, Wlodawer A, Minor W (2010) Unmet challenges of structural genomics. Curr Opin Struct Biol 20:587–597. doi:10.1016/j.sbi.2010.08.001

Comeau SR, Gatchell DW, Vajda S, Camacho CJ (2004a) ClusPro: an automated docking and discrimination method for the prediction of protein complexes. Bioinformatics 20:45–50. doi:10.1093/bioinformatics/btg371

Comeau SR, Gatchell DW, Vajda S, Camacho CJ (2004b) ClusPro: a fully automated algorithm for protein-protein docking. Nucleic Acids Res 32:W96–W99. doi:10.1093/nar/gkh354

Comeau SR, Kozakov D, Brenke R, Shen Y, Beglov D, Vajda S (2007) ClusPro: performance in CAPRI rounds 6–11 and the new server. Proteins 69:781–785. doi:10.1002/prot.21795

Csermely P, Palotai R, Nussinov R (2010) Induced fit, conformational selection and independent dynamic segments: an extended view of binding events. Trends Biochem Sci 35:539–546. doi:10.1016/j.tibs.2010.04.009

Dagliyan O, Proctor EA, D'Auria KM, Ding F, Dokholyan NV (2011) Structural and dynamic determinants of protein-peptide recognition. Structure 19:1837–1845. doi:10.1016/j.str.2011.09.014

de Groot BL, van Aalten DM, Scheek RM, Amadei A, Vriend G, Berendsen HJ (1997) Prediction of protein conformational freedom from distance constraints. Proteins 29:240–251. doi:10.1002/(SICI)1097-0134(199710)29:2<240:AID-PROT11>3.0.CO;2-O

de Vries SJ, Bonvin AMJJ (2011) CPORT: a consensus interface predictor and its performance in prediction-driven docking with HADDOCK. PLoS ONE 6:e17695. doi:10.1371/journal.pone.0017695

de Vries SJ, van Dijk ADJ, Bonvin AMJJ (2006) WHISCY: what information does surface conservation yield? Application to data-driven docking. Proteins 63:479–489. doi:10.1002/prot.20842

de Vries SJ, van Dijk ADJ, Krzeminski M, van Dijk M, Thureau A, Hsu V, Wassenaar T, Bonvin AMJJ (2007) HADDOCK versus HADDOCK: new features and performance of HADDOCK2.0 on the CAPRI targets. Proteins 69:726–733. doi:10.1002/prot.21723

de Vries SJ, van Dijk M, Bonvin AMJJ (2010) The HADDOCK web server for data-driven biomolecular docking. Nat Protoc 5:883–897. doi:10.1038/nprot.2010.32

de Vries SJ, Zacharias M (2012) ATTRACT-EM: a new method for the computational assembly of large molecular machines using cryo-EM maps. PLoS ONE 7:e49733. doi:10.1371/journal.pone.0049733

Diella F, Haslam N, Chica C, Budd A, Michael S, Brown NP, Trave G, Gibson TJ (2008) Understanding eukaryotic linear motifs and their role in cell signaling and regulation. Front Biosci 13:6580–6603. doi:10.2741/3175

Dominguez C, Boelens R, Bonvin AMJJ (2003) HADDOCK: a protein-protein docking approach based on biochemical or biophysical information. J Am Chem Soc 125:1731–1737. doi:10.1021/ja026939x

Donsky E, Wolfson HJ (2011) PepCrawler: a fast RRT-based algorithm for high-resolution refinement and binding affinity estimation of peptide inhibitors. Bioinformatics 27:2836–2842. doi:10.1093/bioinformatics/btr498

Erickson JA, Jalaie M, Robertson DH, Lewis RA, Vieth M (2004) Lessons in molecular recognition: the effects of ligand and protein flexibility on molecular docking accuracy. J Med Chem 47:45–55. doi:10.1021/jm030209y

Fernández-Recio J, Totrov M, Abagyan R (2003) ICM-DISCO docking by global energy optimization with fully flexible side-chains. Proteins 52:113–117. doi:10.1002/prot.10383

Fernández-Recio J, Totrov M, Abagyan R (2004) Identification of protein-protein interaction sites from docking energy landscapes. J Mol Biol 335:843–865. doi:10.1016/j.jmb.2003.10.069

Fink F, Hochrein J, Wolowski V, Merkl R, Gronwald W (2011) PROCOS: computational analysis of protein-protein complexes. J Comput Chem 32:2575–2586. doi:10.1002/jcc.21837

Fischer D, Bachar O, Nussinov R, Wolfson H (1992) An efficient automated computer vision based technique for detection of three dimensional structural motifs in proteins. J Biomol Struct Dyn 9:769–789. doi:10.1080/07391102.1992.10507955

Fischer N, Neumann P, Konevega AL, Bock LV, Ficner R, Rodnina MV, Stark H (2015) Structure of the *E. coli* ribosome-EF-Tu complex at <3 Å resolution by Cs-corrected cryo-EM. Nature 520:567–570. doi:10.1038/nature14275

Fleishman SJ, Whitehead TA, Strauch E-M, Corn JE, Qin S, Zhou H-X, Mitchell JC, Demerdash ONA, Takeda-Shitaka M, Terashi G, Moal IH, Li X, Bates PA, Zacharias M, Park H, Ko J-S, Lee H, Seok C, Bourquard T, Bernauer J, Poupon A, Azé J, Soner S, Ovalı ŞK, Ozbek P, Tal NB, Haliloglu T, Hwang H, Vreven T, Pierce BG, Weng Z, Pérez-Cano L, Pons C, Fernández-Recio J, Jiang F, Yang F, Gong X, Cao L, Xu X, Liu B, Wang P, Li C, Wang C, Robert CH, Guharoy M, Liu S, Huang Y, Li L, Guo D, Chen Y, Xiao Y, London N, Itzhaki Z, Schueler-Furman O, Inbar Y, Potapov V, Cohen M, Schreiber G, Tsuchiya Y, Kanamori E, Standley DM, Nakamura H, Kinoshita K, Driggers CM, Hall RG, Morgan JL, Hsu VL, Zhan J, Yang Y, Zhou Y, Kastritis PL, Bonvin AMJJ, Zhang W, Camacho CJ, Kilambi KP, Sircar A, Gray JJ, Ohue M, Uchikoga N, Matsuzaki Y, Ishida T, Akiyama Y, Khashan R, Bush S, Fouches D, Tropsha A, Esquivel-Rodríguez J, Kihara D, Stranges PB, Jacak R, Kuhlman B, Huang S-Y, Zou X, Wodak SJ, Janin J, Baker D (2011) Community-wide assessment of protein-interface modeling suggests improvements to design methodology. J Mol Biol 414:289–302. doi:10.1016/j.jmb.2011.09.031

Gajda MJ, Tuszynska I, Kaczor M, Bakulina AY, Bujnicki JM (2010) FILTREST3D: discrimination of structural models using restraints from experimental data. Bioinformatics 26:2986–2987. doi:10.1093/bioinformatics/btq582

Gong X, Wang P, Yang F, Chang S, Liu B, He H, Cao L, Xu X, Li C, Chen W, Wang C (2010) Protein-protein docking with binding site patch prediction and network-based terms enhanced combinatorial scoring. Proteins 78:3150–3155. doi:10.1002/prot.22831

González-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 13:2607–2625. doi:10.2174/092986706778201530

Gray JJ, Moughon S, Wang C, Schueler-Furman O, Kuhlman B, Rohl CA, Baker D (2003) Protein-protein docking with simultaneous optimization of rigid-body displacement and side-chain conformations. J Mol Biol 331:281–299. doi:10.1016/s0022-2836(03)00670-3

Grosdidier S, Fernández-Recio J (2008) Identification of hot-spot residues in protein-protein interactions by computational docking. BMC Bioinformatics 9:447. doi:10.1186/1471-2105-9-447

Halperin I, Ma B, Wolfson H, Nussinov R (2002) Principles of docking: an overview of search algorithms and a guide to scoring functions. Proteins 47:409–443. doi:10.1002/prot.10115

Hammes GG, Chang Y-C, Oas TG (2009) Conformational selection or induced fit: a flux description of reaction mechanism. Proc Natl Acad Sci 106:13737–13741. doi:10.1073/pnas.0907195106

Heifetz A, Eisenstein M (2003) Effect of local shape modifications of molecular surfaces on rigid-body protein-protein docking. Protein Eng 16:179–185. doi:10.1093/proeng/gzg021

Hetényi C, van der Spoel D (2002) Efficient docking of peptides to proteins without prior knowledge of the binding site. Protein Sci 11:1729–1737. doi:10.1110/ps.0202302

Hopf TA, Schärfe CPI, Rodrigues JPGLM, Green AG, Kohlbacher O, Sander C, Bonvin AMJJ, Marks DS (2014) Sequence co-evolution gives 3D contacts and structures of protein complexes. Elife. doi:10.7554/eLife.03430

Horton N, Lewis M (1992) Calculation of the free energy of association for protein complexes. Protein Sci 1:169–181. doi:10.1002/pro.5560010117

Huang S-Y, Zou X (2008) An iterative knowledge-based scoring function for protein-protein recognition. Proteins 72:557–579. doi:10.1002/prot.21949

Hwang H, Vreven T, Weng Z (2014) Binding interface prediction by combining protein-protein docking results. Proteins 82:57–66. doi:10.1002/prot.24354

Hwang I, Park S (2008) Computational design of protein therapeutics. Drug Discov Today Technol 5:e43–e48. doi:10.1016/j.ddtec.2008.11.004

Janin J (2010) Protein-protein docking tested in blind predictions: the CAPRI experiment. Mol BioSyst 6:2351–2362. doi:10.1039/c005060c

Janin J, Henrick K, Moult J, Eyck LT, Sternberg MJE, Vajda S, Vakser I, Wodak SJ, Critical Assessment of PRedicted Interactions (2003) CAPRI: a Critical Assessment of PRedicted Interactions. Proteins 52:2–9. doi:10.1002/prot.10381

Jiang L, Gao Y, Mao F, Liu Z, Lai L (2002) Potential of mean force for protein-protein interaction studies. Proteins 46:190–196. doi:10.1002/prot.10031

Jiménez-García B, Pons C, Fernández-Recio J (2013) pyDockWEB: a web server for rigid-body protein-protein docking using electrostatics and desolvation scoring. Bioinformatics 29:1698–1699. doi:10.1093/bioinformatics/btt262

Jones S, Thornton JM (1996) Principles of protein-protein interactions. Proc Natl Acad Sci USA 93:13–20. doi:10.1073/pnas.93.1.13

Kabsch W, Mannherz HG, Suck D, Pai EF, Holmes KC (1990) Atomic structure of the actin: DNase I complex. Nature 347:37–44. doi:10.1038/347037a0

Karaca E, Bonvin AMJJ (2011) A multidomain flexible docking approach to deal with large conformational changes in the modeling of biomolecular complexes. Structure 19:555–565. doi:10.1016/j.str.2011.01.014

Kastritis PL, Bonvin AMJJ (2013a) Molecular origins of binding affinity: seeking the Archimedean point. Curr Opin Struct Biol 23:868–877. doi:10.1016/j.sbi.2013.07.001

Kastritis PL, Bonvin AMJJ (2013b) On the binding affinity of macromolecular interactions: daring to ask why proteins interact. J R Soc Interface 10:20120835. doi:10.1098/rsif.2012.0835

Kastritis PL, Bonvin AMJJ (2010) Are scoring functions in protein-protein docking ready to predict interactomes? Clues from a novel binding affinity benchmark. J Proteome Res 9:2216–2225. doi:10.1021/pr9009854

Kastritis PL, Moal IH, Hwang H, Weng Z, Bates PA, Bonvin AMJJ, Janin J (2011) A structure-based benchmark for protein-protein binding affinity. Protein Sci 20:482–491. doi:10.1002/pro.580

Kastritis PL, Rodrigues JPGLM, Folkers GE, Boelens R, Bonvin AMJJ (2014) Proteins feel more than they see: fine-tuning of binding affinity by properties of the non-interacting surface. J Mol Biol 426:2632–2652. doi:10.1016/j.jmb.2014.04.017

Katchalski-Katzir E, Shariv I, Eisenstein M, Friesem AA, Aflalo C, Vakser IA (1992) Molecular surface recognition: determination of geometric fit between proteins and their ligands by correlation techniques. Proc Natl Acad Sci USA 89:2195–2199. doi:10.1073/pnas.89.6.2195

Khashan R, Zheng W, Tropsha A (2012) Scoring protein interaction decoys using exposed residues (SPIDER): a novel multibody interaction scoring function based on frequent geometric patterns of interfacial residues. Proteins 80:2207–2217. doi:10.1002/prot.24110

Kowalsman N, Eisenstein M (2009) Combining interface core and whole interface descriptors in postscan processing of protein-protein docking models. Proteins 77:297–318. doi:10.1002/prot.22436

Kozakov D, Schueler-Furman O, Vajda S (2008) Discrimination of near-native structures in protein-protein docking by testing the stability of local minima. Proteins 72:993–1004. doi:10.1002/prot.21997

Kurcinski M, Jamroz M, Blaszczyk M, Kolinski A, Kmiecik S (2015) CABS-dock web server for the flexible docking of peptides to proteins without prior knowledge of the binding site. Nucleic Acids Res 43:W419–W424. doi:10.1093/nar/gkv456

Lavi A, Ngan CH, Movshovitz-Attias D, Bohnuud T, Yueh C, Beglov D, Schueler-Furman O, Kozakov D (2013) Detection of peptide-binding sites on protein surfaces: the first step toward the modeling and targeting of peptide-mediated interactions. Proteins 81:2096–2105. doi:10.1002/prot.24422

Lensink MF, Méndez R, Wodak SJ (2007) Docking and scoring protein complexes: CAPRI 3rd Edition. Proteins 69:704–718. doi:10.1002/prot.21804

Lensink MF, Moal IH, Bates PA, Kastritis PL, Melquiond ASJ, Karaca E, Schmitz C, van Dijk M, Bonvin AMJJ, Eisenstein M, Jiménez-García B, Grosdidier S, Solernou A, Pérez-Cano L, Pallara C, Fernández-Recio J, Xu J, Muthu P, Praneeth Kilambi K, Gray JJ, Grudinin S, Derevyanko G, Mitchell JC, Wieting J, Kanamori E, Tsuchiya Y, Murakami Y, Sarmiento J, Standley DM, Shirota M, Kinoshita K, Nakamura H, Chavent M, Ritchie DW, Park H, Ko J, Lee H, Seok C, Shen Y, Kozakov D, Vajda S, Kundrotas PJ, Vakser IA, Pierce BG, Hwang H, Vreven T, Weng Z, Buch I, Farkash E, Wolfson HJ, Zacharias M, Qin S, Zhou H-X, Huang S-Y, Zou X, Wojdyla JA, Kleanthous C, Wodak SJ (2014) Blind prediction of interfacial water positions in CAPRI. Proteins 82:620–632. doi:10.1002/prot.24439

Lensink MF, Velankar S, Kryshtafovych A, Huang SY, Schneidman-Duhovny D, Sali A, Segura J, Fernandez-Fuentes N, Viswanath S, Elber R, Grudinin S, Popov P, Neveu E, Lee H, Baek M, Park S, Heo L, Rie Lee G, Seok C, Qin S, Zhou HX, Ritchie DW, Maigret B, Devignes MD, Ghoorah A, Torchala M, Chaleil RA, Bates PA, Ben-Zeev E, Eisenstein M, Negi SS, Weng Z, Vreven T, Pierce BG, Borrman TM, Yu J, Ochsenbein F, Guerois R, Vangone A, Rodrigues JP, van Zundert G, Nellen M, Xue L, Karaca E, Melquiond AS, Visscher K, Kastritis PL, Bonvin AM, Xu X, Qiu L, Yan C, Li J, Ma Z, Cheng J, Zou X, Shen Y, Peterson LX, Kim HR, Roy A, Han X, Esquivel-Rodriguez J, Kihara D, Yu X, Bruce NJ, Fuller JC, Wade RC, Anishchenko I, Kundrotas PJ, Vakser IA, Imai K, Yamada K, Oda T, Nakamura T, Tomii K, Pallara C, Romero-Durana M, Jimenez-Garcia B, Moal IH, Fernandez-Recio J, Joung JY, Kim JY, Joo K, Lee J, Kozakov D, Vajda S, Mottarella S, Hall DR, Beglov D, Mamonov A, Xia B, Bohnuud T, Del Carpio CA, Ichiishi E, Marze N, Kuroda D, Roy Burman SS, Gray JJ, Chermak E, Cavallo L, Oliva R (2016) Prediction of homo- and hetero-protein complexes by protein docking and template-based modeling: a CASP-CAPRI experiment. Proteins. doi:10.1002/prot.25007

Lensink MF, Wodak SJ (2013) Docking, scoring, and affinity prediction in CAPRI. Proteins 81:2082–2095. doi:10.1002/prot.24428

Lensink MF, Wodak SJ (2010) Docking and scoring protein interactions: CAPRI 2009. Proteins 78:3073–3084. doi:10.1002/prot.22818

London N, Raveh B, Schueler-Furman O (2013) Peptide docking and structure-based characterization of peptide binding: from knowledge to know-how. Curr Opin Struct Biol 23:894–902. doi:10.1016/j.sbi.2013.07.006

London N, Schueler-Furman O (2008) FunHunt: model selection based on energy landscape characteristics. Biochem Soc Trans 36:1418–1421. doi:10.1042/BST0361418

Lu H, Lu L, Skolnick J (2003) Development of unified statistical potentials describing protein-protein interactions. Biophys J 84:1895–1901. doi:10.1016/S0006-3495(03)74997-2

Luo J, Guo Y, Zhong Y, Ma D, Li W, Li M (2014) A functional feature analysis on diverse protein-protein interactions: application for the prediction of binding affinity. J Comput Aided Mol Des 28:619–629. doi:10.1007/s10822-014-9746-y

Lyskov S, Gray JJ (2008) The RosettaDock server for local protein-protein docking. Nucleic Acids Res 36:W233–W238. doi:10.1093/nar/gkn216

Ma XH, Wang CX, Li CH, Chen WZ (2002) A fast empirical approach to binding free energy calculations based on protein interface information. Protein Eng 15:677–681. doi:10.1093/protein/15.8.677

Macindoe G, Mavridis L, Venkatraman V, Devignes M-D, Ritchie DW (2010) HexServer: an FFT-based protein docking server powered by graphics processors. Nucleic Acids Res 38:W445–W449. doi:10.1093/nar/gkq311

Marillet S, Boudinot P, Cazals F (2015) High resolution crystal structures leverage protein binding affinity predictions

Martin J, Lavery R (2012) Arbitrary protein-protein docking targets biologically relevant interfaces. BMC Biophys 5:7. doi:10.1186/2046-1682-5-7

Mashiach E, Nussinov R, Wolfson HJ (2010a) FiberDock: a web server for flexible induced-fit backbone refinement in molecular docking. Nucleic Acids Res 38:W457–W461. doi:10.1093/nar/gkq373

Mashiach E, Schneidman-Duhovny D, Andrusier N, Nussinov R, Wolfson HJ (2008) FireDock: a web server for fast interaction refinement in molecular docking. Nucleic Acids Res 36:W229–W232. doi:10.1093/nar/gkn186

Mashiach E, Schneidman-Duhovny D, Peri A, Shavit Y, Nussinov R, Wolfson HJ (2010b) An integrated suite of fast docking algorithms. Proteins 78:3197–3204. doi:10.1002/prot.22790

Metz A, Ciglia E, Gohlke H (2012) Modulating protein-protein interactions: from structural determinants of binding to druggability prediction to application. Curr Pharm Des 18:4630–4647. doi:10.2174/138161212802651553

Moal IH, Bates PA (2010) SwarmDock and the use of normal modes in protein-protein docking. Int J Mol Sci 11:3623–3648. doi:10.3390/ijms11103623

Moal IH, Bates PA (2012) Kinetic rate constant prediction supports the conformational selection mechanism of protein binding. PLoS Comput Biol 8:e1002351. doi:10.1371/journal.pcbi.1002351

Moal IH, Jiménez-García B, Fernández-Recio J (2015) CCharPPI web server: computational characterization of protein-protein interactions from structure. Bioinformatics 31:123–125. doi:10.1093/bioinformatics/btu594

Moal IH, Moretti R, Baker D, Fernández-Recio J (2013a) Scoring functions for protein–protein interactions. Curr Opin Struct Biol 23:862–867. doi:10.1016/j.sbi.2013.06.017

Moal IH, Torchala M, Bates PA, Fernández-Recio J (2013b) The scoring of poses in protein-protein docking: current capabilities and future directions. BMC Bioinformatics 14:286. doi:10.1186/1471-2105-14-286

Moont G, Gabb HA, Sternberg MJ (1999) Use of pair potentials across protein interfaces in screening predicted docked complexes. Proteins 35:364–373. doi:10.1002/(SICI)1097-0134(19990515)35:3<364:AID-PROT11>3.0.CO;2-4

Moreira IS, Fernandes PA, Ramos MJ (2010) Protein-protein docking dealing with the unknown. J Comput Chem 31:317–342. doi:10.1002/jcc.21276

Moreira IS, Martins JM, Coimbra JTS, Ramos MJ, Fernandes PA (2015) A new scoring function for protein-protein docking that identifies native structures with unprecedented accuracy. Phys Chem Chem Phys 17:2378–2387. doi:10.1039/c4cp04688a

Naider F, Anglister J (2009) Peptides in the treatment of AIDS. Curr Opin Struct Biol 19:473–482. doi:10.1016/j.sbi.2009.07.003

Negi SS, Schein CH, Oezguen N, Power TD, Braun W (2007) InterProSurf: a web server for predicting interacting sites on protein surfaces. Bioinformatics 23:3397–3399. doi:10.1093/bioinformatics/btm474

Neuvirth H, Raz R, Schreiber G (2004) ProMate: a structure based prediction program to identify the location of protein-protein binding sites. J Mol Biol 338:181–199. doi:10.1016/j.jmb.2004.02.040

Nisius B, Sha F, Gohlke H (2012) Structure-based computational analysis of protein binding sites for function and druggability prediction. J Biotechnol 159:123–134. doi:10.1016/j.jbiotec.2011.12.005

Northrup SH, Erickson HP (1992) Kinetics of protein-protein association explained by Brownian dynamics computer simulation. Proc Natl Acad Sci USA 89:3338–3342. doi:10.1073/pnas.89.8.3338

Ofran Y, Rost B (2003) Analysing six types of protein-protein interfaces. J Mol Biol 325:377–387. doi:10.1016/s0022-2836(02)01223-8

Oliva R, Vangone A, Cavallo L (2013) Ranking multiple docking solutions based on the conservation of inter-residue contacts. Proteins 81:1571–1584. doi:10.1002/prot.24314

Palma PN, Krippahl L, Wampler JE, Moura JJ (2000) BiGGER: a new (soft) docking algorithm for predicting protein interactions. Proteins 39:372–384. doi:10.1002/(SICI)1097-0134(20000601)39:4<372:AID-PROT100>3.0.CO;2-Q

Petsalaki E, Russell RB (2008) Peptide-mediated interactions in biological systems: new discoveries and applications. Curr Opin Biotechnol 19:344–350. doi:10.1016/j.copbio.2008.06.004

Petsalaki E, Stark A, García-Urdiales E, Russell RB (2009) Accurate prediction of peptide binding sites on protein surfaces. PLoS Comput Biol 5:e1000335. doi:10.1371/journal.pcbi.1000335

Phizicky EM, Fields S (1995) Protein-protein interactions: methods for detection and analysis. Microbiol Rev 59:94–123

Pierce B, Weng Z (2007) ZRANK: reranking protein docking predictions with an optimized energy function. Proteins 67:1078–1086. doi:10.1002/prot.21373

Pierce BG, Hourai Y, Weng Z (2011) Accelerating protein docking in ZDOCK using an advanced 3D convolution library. PLoS ONE 6:e24657. doi:10.1371/journal.pone.0024657

Porollo A, Meller J (2006) Prediction-based fingerprints of protein-protein interactions. Proteins 66:630–645. doi:10.1002/prot.21248

Qin S, Pang X, Zhou H-X (2011) Automated prediction of protein association rate constants. Structure 19:1744–1751. doi:10.1016/j.str.2011.10.015

Qin S, Zhou H-X (2013) Using the concept of transient complex for affinity predictions in CAPRI rounds 20-27 and beyond. Proteins 81:2229–2236. doi:10.1002/prot.24366

Qin S, Zhou H-X (2007) meta-PPISP: a meta web server for protein-protein interaction site prediction. Bioinformatics 23:3386–3387. doi:10.1093/bioinformatics/btm434

Raveh B, London N, Schueler-Furman O (2010) Sub-angstrom modeling of complexes between flexible peptides and globular proteins. Proteins 78:2029–2040. doi:10.1002/prot.22716

Raveh B, London N, Zimmerman L, Schueler-Furman O (2011) Rosetta FlexPepDock ab-initio: simultaneous folding, docking and refinement of peptides onto their receptors. PLoS ONE 6:e18934. doi:10.1371/journal.pone.0018934

Ritchie DW (2008) Recent progress and future directions in protein-protein docking. Curr Protein Pept Sci 9:1–15. doi:10.2174/138920308783565741

Ritchie DW, Kemp GJ (2000) Protein docking using spherical polar Fourier correlations. Proteins 39:178–194. doi:10.1002/(SICI)1097-0134(20000501)39:2<178:AID-PROT8>3.0.CO;2-6

Rodrigues JPGLM, Bonvin AMJJ (2014) Integrative computational modeling of protein interactions. FEBS J 281:1988–2003. doi:10.1111/febs.12771

Rodrigues JPGLM, Trellet M, Schmitz C, Kastritis P, Karaca E, Melquiond ASJ, Bonvin AMJJ (2012) Clustering biomolecular complexes by residue contacts similarity. Proteins 80:1810–1817. doi:10.1002/prot.24078

Rubinstein M, Niv MY (2009) Peptidic modulators of protein-protein interactions: progress and challenges in computational design. Biopolymers 91:505–513. doi:10.1002/bip.21164

Russell RB, Gibson TJ (2008) A careful disorderliness in the proteome: sites for interaction and targets for future therapies. FEBS Lett 582:1271–1275. doi:10.1016/j.febslet.2008.02.027

Ruvinsky AM, Vakser IA (2008) Interaction cutoff effect on ruggedness of protein-protein energy landscape. Proteins 70:1498–1505. doi:10.1002/prot.21644

Sacquin-Mora S, Carbone A, Lavery R (2008) Identification of protein interaction partners and protein-protein interaction sites. J Mol Biol 382:1276–1289. doi:10.1016/j.jmb.2008.08.002

Schlick T, Collepardo-Guevara R, Halvorsen LA, Jung S, Xiao X (2011) Biomolecularmodeling and simulation: a field coming of age. Q Rev Biophys 44:191–228. doi:10.1017/S0033583510000284

Schmitz C, Bonvin AMJJ (2011) Protein-protein HADDocking using exclusively pseudocontact shifts. J Biomol NMR 50:263–266. doi:10.1007/s10858-011-9514-4

Schmitz C, Melquiond ASJ, de Vries SJ, Karaca E, van Dijk M, Kastritis PL, Bonvin AMJJ (2012) Protein–protein docking with HADDOCK. Towards Mech Syst Biol 520–535. doi:10.1002/9783527644506.ch32

Schneidman-Duhovny D, Inbar Y, Nussinov R, Wolfson HJ (2005) PatchDock and SymmDock: servers for rigid and symmetric docking. Nucleic Acids Res 33:W363–W367. doi:10.1093/nar/gki481

Schreiber G, Fersht AR (1996) Rapid, electrostatically assisted association of proteins. Nat Struct Biol 3:427–431. doi:10.1038/nsb0596-427

Smith GR, Fitzjohn PW, Page CS, Bates PA (2005a) Incorporation of flexibility into rigid-body docking: applications in rounds 3–5 of CAPRI. Proteins 60:263–268. doi:10.1002/prot.20568
Smith GR, Sternberg MJE (2002) Prediction of protein-protein interactions by docking methods. Curr Opin Struct Biol 12:28–35. doi:10.1016/s0959-440x(02)00285-3
Smith GR, Sternberg MJE, Bates PA (2005b) The relationship between the flexibility of proteins and their conformational states on forming protein-protein complexes with an application to protein-protein docking. J Mol Biol 347:1077–1101. doi:10.1016/j.jmb.2005.01.058
Sousa SF, Fernandes PA, Ramos MJ (2006) Protein-ligand docking: current status and future challenges. Proteins 65:15–26. doi:10.1002/prot.21082
Stites WE (1997) Proteinminus signProtein interactions: interface structure, binding thermodynamics, and mutational analysis. Chem Rev 97:1233–1250. doi:10.1021/cr960387h
Su Y, Zhou A, Xia X, Li W, Sun Z (2009) Quantitative prediction of protein-protein binding affinity with a potential of mean force considering volume correction. Protein Sci 18:2550–2558. doi:10.1002/pro.257
Sugiki T, Fujiwara T, Kojima C (2014) Latest approaches for efficient protein production in drug discovery. Expert Opin Drug Discov 9:1189–1204. doi:10.1517/17460441.2014.941801
Szymkowski DE (2005) Creating the next generation of protein therapeutics through rational drug design. Curr Opin Drug Discov Devel 8:590–600
Tian F, Lv Y, Yang L (2012) Structure-based prediction of protein-protein binding affinity with consideration of allosteric effect. Amino Acids 43:531–543. doi:10.1007/s00726-011-1101-1
Torchala M, Moal IH, Chaleil RAG, Agius R, Bates PA (2013) A Markov-chain model description of binding funnels to enhance the ranking of docked solutions. Proteins 81:2143–2149. doi:10.1002/prot.24369
Tovchigrechko A, Vakser IA (2006) GRAMM-X public web server for protein-protein docking. Nucleic Acids Res 34:W310–W314. doi:10.1093/nar/gkl206
Trabuco LG, Lise S, Petsalaki E, Russell RB (2012) PepSite: prediction of peptide-binding sites from protein surfaces. Nucleic Acids Res 40:W423–W427. doi:10.1093/nar/gks398
Trellet M, Melquiond ASJ, Bonvin AMJJ (2013) A unified conformational selection and induced fit approach to protein-peptide docking. PLoS ONE 8:e58769. doi:10.1371/journal.pone.0058769
Trellet M, Melquiond ASJ, Bonvin AMJJ (2015) Information-driven modeling of protein-peptide complexes. Methods Mol Biol 1268:221–239. doi:10.1007/978-1-4939-2285-7_10
Tress M, de Juan D, Graña O, Gómez MJ, Gómez-Puertas P, González JM, López G, Valencia A (2005) Scoring docking models with evolutionary information. Proteins 60:275–280. doi:10.1002/prot.20570
Tsai CJ, Kumar S, Ma B, Nussinov R (1999) Folding funnels, binding funnels, and protein function. Protein Sci 8:1181–1190. doi:10.1110/ps.8.6.1181
Vaara M (2009) New approaches in peptide antibiotics. Curr Opin Pharmacol 9:571–576. doi:10.1016/j.coph.2009.08.002
Vajda S, Kozakov D (2009) Convergence and combination of methods in protein-protein docking. Curr Opin Struct Biol 19:164–170. doi:10.1016/j.sbi.2009.02.008
van Dijk ADJ, Fushman D, Bonvin AMJJ (2005) Various strategies of using residual dipolar couplings in NMR-driven protein docking: application to Lys48-linked di-ubiquitin and validation against 15N-relaxation data. Proteins 60:367–381. doi:10.1002/prot.20476
van Dijk ADJ, Kaptein R, Boelens R, Bonvin AMJJ (2006) Combining NMR relaxation with chemical shift perturbation data to drive protein-protein docking. J Biomol NMR 34:237–244. doi:10.1007/s10858-006-0024-8
van Zundert GCP, Melquiond ASJ, Bonvin AMJJ (2015) Integrative modeling of biomolecular complexes: HADDOCKing with cryo-electron microscopy data. Structure 23:949–960. doi:10.1016/j.str.2015.03.014
Vangone A, Bonvin AM (2015) Contacts-based prediction of binding affinity in protein-protein complexes. Elife 4:e07454. doi:10.7554/eLife.07454
Vangone A, Cavallo L, Oliva R (2013) Using a consensus approach based on the conservation of inter-residue contacts to rank CAPRI models. Proteins 81:2210–2220. doi:10.1002/prot.24423

Vangone A, Oliva R, Cavallo L (2012) CONS-COCOMAPS: a novel tool to measure and visualize the conservation of inter-residue contacts in multiple docking solutions. BMC Bioinformatics 13(Suppl 4):S19. doi:10.1186/1471-2105-13-S4-S19

Vangone A, Spinelli R, Scarano V, Cavallo L, Oliva R (2011) COCOMAPS: a web application to analyze and visualize contacts at the interface of biomolecular complexes. Bioinformatics 27:2915–2916. doi:10.1093/bioinformatics/btr484

Venkatraman V, Yang YD, Sael L, Kihara D (2009) Protein-protein docking using region-based 3D Zernike descriptors. BMC Bioinformatics 10:407. doi:10.1186/1471-2105-10-407

Verschueren E, Vanhee P, Rousseau F, Schymkowitz J, Serrano L (2013) Protein-peptide complex prediction through fragment interaction patterns. Structure 21:789–797. doi:10.1016/j.str.2013.02.023

Viswanath S, Ravikant DVS, Elber R (2013) Improving ranking of models for protein complexes with side chain modeling and atomic potentials. Proteins 81:592–606. doi:10.1002/prot.24214

Vreven T, Hwang H, Weng Z (2011) Integrating atom-based and residue-based scoring functions for protein-protein docking. Protein Sci 20:1576–1586. doi:10.1002/pro.687

Vreven T, Pierce BG, Hwang H, Weng Z (2013) Performance of ZDOCK in CAPRI rounds 20-26. Proteins 81:2175–2182. doi:10.1002/prot.24432

Vreven T, Moal IH, Vangone A, Pierce BG, Kastritis PL, Torchala M, Chaleil R, Jimenez-Garcia B, Bates PA, Fernandez-Recio J, Bonvin AM, Weng Z (2015) Updates to the integrated protein-protein interaction benchmarks: docking benchmark version 5 and affinity benchmark version 2. J Mol Biol 427:3031–3041. doi:10.1016/j.jmb.2015.07.016

Wass MN, David A, Sternberg MJE (2011a) Challenges for the prediction of macromolecular interactions. Curr Opin Struct Biol 21:382–390. doi:10.1016/j.sbi.2011.03.013

Wass MN, Fuentes G, Pons C, Pazos F, Valencia A (2011b) valencia2011. Mol Syst Biol 7:1–8. doi:10.1038/msb.2011.3

Weikl TR, von Deuster C (2009) Selected-fit versus induced-fit protein binding: kinetic differences and mutational analysis. Proteins 75:104–110. doi:10.1002/prot.22223

Wodak SJ, Janin J (1978) Computer analysis of protein-protein interaction. J Mol Biol 124:323–342. doi:10.1016/0022-2836(78)90302-9

Xue LC, Rodrigues JP, Kastritis PL, Bonvin AM, Vangone A (2016) PRODIGY: a web server for predicting the binding affinity of protein-protein complexes. Bioinformatics. doi:10.1093/bioinformatics/btw514

Xue LC, Jordan RA, El-Manzalawy Y, Dobbs D, Honavar V (2014) DockRank: ranking docked conformations using partner-specific sequence homology-based protein interface prediction. Proteins 82:250–267. doi:10.1002/prot.24370

Yang S (2014) Methods for SAXS-based structure determination of biomolecular complexes. Adv Mater Weinheim 26:7902–7910. doi:10.1002/adma.201304475

Zacharias M (2005) ATTRACT: protein-protein docking in CAPRI using a reduced protein model. Proteins 60:252–256. doi:10.1002/prot.20566

Zellner H, Staudigel M, Trenner T, Bittkowski M, Wolowski V, Icking C, Merkl R (2012) PresCont: predicting protein-protein interfaces utilizing four residue properties. Proteins 80:154–168. doi:10.1002/prot.23172

Zhang C, Liu S, Zhu Q, Zhou Y (2005) A knowledge-based energy function for protein-ligand, protein-protein, and protein-DNA complexes. J Med Chem 48:2325–2335. doi:10.1021/jm049314d

Zhang QC, Deng L, Fisher M, Guan J, Honig B, Petrey D (2011) PredUs: a web server for predicting protein interfaces using structural neighbors. Nucleic Acids Res 39:W283–W287. doi:10.1093/nar/gkr311

Zhang QC, Petrey D, Deng L, Qiang L, Shi Y, Thu CA, Bisikirska B, Lefebvre C, Accili D, Hunter T, Maniatis T, Califano A, Honig B (2012) Structure-based prediction of protein-protein interactions on a genome-wide scale. Nature 490:556–560. doi:10.1038/nature11503

Zhou P, Wang C, Ren Y, Yang C, Tian F (2013) Computational peptidology: a new and promising approach to therapeutic peptide design. Curr Med Chem 20:1985–1996. doi:10.2174/09298673113201500005