UNIVERSITAT DE BARCELONA

Improving the description of protein-protein association energy

Miguel Alfonso Romero Durana

Universidad de Barcelona
Programa de Doctorado en Biomedicina
EEES HDK05

Improving the description of protein-protein association energy

Memoria presentada por Miguel Alfonso Romero Durana para la obtención del grado de doctor por la Universidad de Barcelona

Miguel Alfonso Romero Durana

Director
Dr. Juan Fernández Recio

Tutor
Dr. Josep Lluís Gelpí Buchaca

Barcelona, 2018

A mis padres

Acknowledgements

Me gustaría comenzar expresando mi más profundo agradecimiento a Juan Fernández Recio. Sin su apoyo, consejo y confianza permanentes esta aventura hubiera sido del todo imposible.

Quiero continuar acordándome de mis compañeros en Life Science, especialmente de Laura y de Manuel. Santi, Valentí, Montse, Leiden, Txema, Víctor fueron mi grupo, sin ser mi grupo, durante los primeros años. En este tiempo he tenido la suerte de conocer a Laia, Romina, Dimitri, Alexis, Pau, Fátima, Emanuele, Ferrán, Gerard, Dani, Ryoji, Ali, Armin, Marina, James, Marta, Mercé, Silvia, Elías. Gracias a todos.

No me puedo olvidar de mis compañeros en el grupo de Protein Interactions and Docking. Solène y Carles fueron los veteranos que me recibieron con los brazos abiertos. La risa permanente de Chiara me ha acompañado durante gran parte del viaje y a Brian, compañero de penas y alegrías le debo, entre otras muchas cosas, todo el Python que sé. Después vinieron Didier, Sergio, Iain, Mark, Jorge, Dhoha, Silvia, Dámaris, Lucia y mi paisano Bruno, con quien compartí algunos proyectos y la suerte de investigar en equipo. De todos ellos guardo un maravilloso recuerdo. A mis compañeros actuales, Luis Ángel y Mireia, quiero agradecerles su apoyo y comprensión a la hora de tomar el testigo en esta última época en que estoy, pero no estoy. Muchísimas gracias a todos. Ha sido un verdadero placer trabajar con vosotros.

Termino recordando a mis amigos, especialmente a Paco, en buena parte responsable de que me decidiera a dar el primer paso del camino que me ha traído hasta aquí.

Y dejo para el final a la familia, sin duda lo más importante. Padres, hermana, abuelos, tíos, primos, Dolo y Marco.

Contents

List of tables

Introduction

Methods

Results

List of figures

Introduction

Methods

Results

Discussion

List of papers

The work of this thesis has produced the following publications:

Jiménez-García, Brian, Jorge Roel-Touris, Miguel Romero-Durana, Miquel Vidal, Daniel Jiménez-González, and Juan Fernández-Recio. (2018). "LightDock: A New Multi-Scale Approach to Protein-Protein Docking." Bioinformatics (Oxford, England) 34 (1): 49–55.

Lensink, Marc F., Sameer Velankar, Andriy Kryshtafovych, Shen-You Huang, Dina Schneidman-Duhovny, Andrej Sali, Joan Segura, et al. (2016). "Prediction of Homoprotein and Heteroprotein Complexes by Protein Docking and Template-Based Modeling: A CASP-CAPRI Experiment." Proteins: Structure, Function, and Bioinformatics 84 (September): 323–48.

Moretti, Rocco, Sarel J. Fleishman, Rudi Agius, Mieczyslaw Torchala, Paul A. Bates, Panagiotis L. Kastritis, João P. G. L. M. Rodrigues, et al. (2013). "Community-Wide Evaluation of Methods for Predicting the Effect of Mutations on Protein–Protein Interactions." Proteins: Structure, Function, and Bioinformatics 81 (11): 1980–87.

Pallara, Chiara, Brian Jiménez-García, Laura Pérez-Cano, Miguel Romero-Durana, Albert Solernou, Solène Grosdidier, Carles Pons, Iain H. Moal, and Juan Fernandez-Recio. (2013). "Expanding the Frontiers of Protein–Protein Modeling: From Docking and Scoring to Binding Affinity Predictions and Other Challenges." Proteins: Structure, Function, and Bioinformatics 81 (12): 2192–2200.

Pallara, Chiara, Brian Jiménez-García, Miguel Romero, Iain H. Moal, and Juan Fernández-Recio. (2017). "PyDock Scoring for the New Modeling Challenges in Docking: Protein–Peptide, Homo-Multimers, and Domain–Domain Interactions." Proteins: Structure, Function, and Bioinformatics 85 (3): 487–96.

Pérez-Cano, Laura, Miguel Romero-Durana, and Juan Fernández-Recio. (2016). "Structural and Energy Determinants in Protein-RNA Docking." Methods (San Diego, Calif.), November.

Romero-Durana, Miguel, Chiara Pallara, Fabian Glaser, and Juan Fernández-Recio. (2017). "Modeling Binding Affinity of Pathological Mutations for Computational Protein Design." In Computational Protein Design, edited by Ilan Samish, 1529:139–59. New York, NY: Springer New York.

Abbreviations

a.u: Arbitrary units

ANM: Anisotropic network model

CG: Coarse grained

FFT: Fast Fourier transform

GB: Generalized Born

MD: Molecular dynamics

NM: Normal modes

NMA: Normal mode analysis

PPI: Protein-protein interactions

TBM: Template-based modeling

1 Introduction

1.1 Introduction to proteins

Proteins are large biomolecules that play a crucial role in virtually all biological processes in our cells, such as enzymatic activity, immune responses, signal transduction, transport mechanisms, cell adhesion, or cell cycle. There are a great variety of proteins, with very different sizes, properties, and functions. Strikingly, all proteins are constructed with the same set of building blocks, namely 20 different amino acids. The combination of these 20 amino acids in multiple ways results in proteins as diverse as enzymes, hormones, and antibodies, which form different tissues like feathers, bones, and muscles (Lehninger, Nelson, and Cox 2008). Amino acids are organic compounds containing an amino and a carboxylic functional group covalently bound to a carbon atom, called alpha carbon or Cα, which in turn can be covalently attached to a side-chain group. The side-chain group defines the amino acid type, providing a particular size, shape, charge, hydrogen-bonding capacity, hydrophobic character and chemical reactivity (Berg et al. 2007). Proteins are linear polymers of amino acids, linked by peptide bonds formed between the carboxyl group of one amino acid and the amino group of another amino acid.

The amino acid sequences of proteins have their origin in the nucleotide sequences of genes. In a process called gene expression, the information contained in a given protein-coding gene is decoded to synthesize a protein. In a first step called transcription, mRNA is synthesized from DNA. Then, ribosomes decode this mRNA to produce a specific amino acid chain in a process named translation. During translation, ribosomes add one amino acid at a time to the end of the synthesized protein. Amino acids polymerize by forming a peptide bond between the carboxyl group of the amino acid in the protein, and the amide group of the amino acid to be added. The repeated amide N, Cα, and carbonyl C atoms of each amino acid residue constitute the backbone of a protein from which the side-chain groups stem out (Lodish et al. 2000). During and after translation, polypeptide chains may be subjected to modifications. Once formed, proteins fold into a distinct 3D structure and may exist for a specific time that ranges from minutes to years, before being degraded and recycled through the process of protein turnover.

1.1.1 Protein structure

Four different protein structure levels are defined.

The primary structure of a protein refers to the linear sequence of its amino acids. By convention, the primary structure of proteins starts at the N-terminal and finishes at the C-terminal, following the order in which ribosomes synthesize them. The primary structure of a protein is specified by the nucleotide sequence of the gene associated with it.

Local, regular, three-dimensional arrangements of the protein backbone constitute the secondary structure of proteins. Hydrogen bonds between the backbone NH, and CO groups, usually stabilize these local arrangements. There are two common types of secondary structure motives: alpha helices and beta sheets. Interestingly, these periodic structures were predicted by Linus Pauling and Robert Corey in 1951, years before they were experimentally confirmed (Berg et al. 2007).

The spatial distribution of secondary structure elements defines the tertiary structure or fold of proteins. Interactions such as hydrogen bonding, van der Waals, salt bridges, hydrophobic packing, and disulfide bonds stabilize tertiary structures.

In some occasions, proteins form assemblies with other molecules. In these cases, the three-dimensional arrangement of the subunits forms its quaternary structure. One example is hemoglobin, which is an assembly of four globular proteins.

In 1961, Anfinsen proposed that the amino acid sequence or primary structure of a protein determines its tertiary structure (Anfinsen et al. 1961). This hypothesis integrates the different structural levels defined above into an emergent process. The existence of intrinsically disordered proteins or the possibility of molecules having more than one conformational state are exceptions to the Anfinsen proposition. Nevertheless, it set the bases from where major research on protein folding developed.

1.2 Protein-protein interactions (PPI)

Protein sequence determines protein structure, and protein structure determines protein function. This is one of the leading mottos of structural biology. However, most proteins do not perform its function in isolation but interacting with biomolecules such as other proteins, lipids, nucleic acids or small molecules (Rual et al. 2005; Stelzl et al. 2005). Therefore, to fully understand protein function, we shall extend the analysis from individual proteins to protein interactions.

1.2.1 PPI experimental characterization

In recent years, the development of high-throughput techniques like yeast two-hybrid (Y2H) screening (Uetz et al. 2000; Ito et al. 2001) and affinity purification-mass spectrometry, AP-MS (Ewing et al. 2007), have dramatically increased the number of reported protein-protein interactions. Nowadays, there is access to global 'snapshots' of the interactome (Sanchez et al. 1999), the set of molecular interactions that take place in a cell or an organism. These studies have provided a global view of the interactions involved in a given process or pathway. They have also shown that PPIs are integrated into highly organized and dynamic interaction networks (Rual et al. 2005; Stelzl et al. 2005). Full understanding of these interaction networks could reveal not only the cellular processes they regulate but the underlying mechanisms of diseases such as cancer, giving clues of possible therapeutic strategies (Jonsson and Bates 2006; Sun and Zhao 2010; Choura and Rebaï 2012).

Useful as they are, these methods can determine whether two proteins are likely to interact, but cannot reveal the molecular mechanisms involved in protein-protein interactions (Skrabanek et al. 2008). To attain a deep understanding of protein function, we need to know the 3D structure and energetics of the formed complexes at atomic resolution (Stein, Mosca, and Aloy 2011), which these high-throughput techniques cannot provide.

1.2.2 Experimental structural determination of complexes

The Protein Data Bank, or PDB (Bernstein et al. 1977), is a database for 3D structural data of biomolecules such as proteins and nucleic acids. Table 1.1 contains the number of structures for proteins and several protein complex types, grouped by experimental technique, deposited in the Protein Data Bank in Europe (PDBe) (Gutmanas et al. 2014; S. Velankar et al. 2012; Sameer Velankar et al. 2011; S. Velankar et al. 2010) as in February 2018. There are almost four times more structures of individual proteins than structures of protein complexes (including protein-protein, protein-RNA, and protein-DNA complexes), even though the total number of estimated interactions can be much higher than the number of individual proteins (Stumpf et al. 2008; Venkatesan et al. 2009), a clear indication of the difficulties encountered when trying to resolve complex structures experimentally. Among the different types of protein interactions, 78% correspond to protein-protein, 14% to protein-DNA and 8% to protein-RNA complexes.

	Protein/protein	Protein/RNA	Protein/DNA	Protein
X-Ray	20749	1629	3984	95568
NMR	861	120	130	9702
CryoEM	866	408	61	594
Other	36	5	1	268
Total	22512	2162	4176	106132

Table 1.1. Number of structures deposited in the PDBe as in February 2018.

Most of the structures in PDB have been solved by X-ray crystallography (Smyth and Martin 2000). While NMR (Wüthrich 1990) is the second preferred technique for single protein structure determination, electron microscopy is very close to NMR in the number of protein complex structures deposited in the PDB.

1.2.2.1 X-ray crystallography

X-ray crystallography has been essential in the development of structural biology (Smyth and Martin 2000). It is a mature technique that provides a very high level of detail of biomolecules, at the scale of atoms and chemical bonds. When a beam of X-rays strikes the regular and symmetric structure of a crystal lattice, it is refracted into a specific pattern. From the diffraction pattern, it is possible to

estimate the density distribution of electrons within the crystal, and from there, the mean positions of the atoms, chemical bond lengths, angles as well as other valuable information. As long as a good crystal is available, X-ray crystallography can be applied to huge complexes (> 100 kDa). The requirement of crystallized samples is the major drawback of this method. Systems for which crystallization is difficult such as membrane proteins, intrinsically disordered proteins, transient or low-affinity complexes are not suitable for this technique. Additionally, there are concerns about to what extent the structural information obtained from crystallized systems can be extrapolated to the in vivo, biological, environment (Ofran and Rost 2003; Ranjit Prasad Bahadur et al. 2004; R. P. Bahadur and Zacharias 2008).

1.2.2.2 NMR spectroscopy

NMR (Wüthrich 1990) spectroscopy is the experimental method with the second largest number of structures deposited in the PDB. This technique exploits a physical phenomenon by which nuclei in a magnetic field absorb and re-emit electromagnetic radiation at specific resonance frequencies. The analysis of NMR spectra can provide detailed information not only about the structure but also about the dynamics of molecules. NMR spectroscopy does not have the limitations imposed by requiring crystallized samples like X-ray crystallography. It can analyze proteins in solution, that is, in closer conditions to the in vivo environment than X-ray crystallography. The primary limiting factor of NMR is its inability to study proteins above 40 kDa (Krishnan and Rupp 2012), although more powerful magnets and technical advances are continuously improving its applicability.

1.2.2.3 Cryo-electron microscopy

In the last decade, the development of a new generation of 'direct electron detectors' has led to a revolution in the cryo-electron microscopy (cryo-EM) field (Kühlbrandt 2014; Callaway 2015). Cryo-EM does not require the molecule of interest to be crystallized. Therefore, it can be used to analyze molecules that are difficult to crystallize such as membrane proteins or large multi-protein complexes. The technique consists of firing a beam of electrons at a previously frozen sample. The scattered electrons are then registered on a detector, from which the structure of the sample can be resolved. Cryo-EM is suited to the study of large, stable molecules that are not severely affected by the electron beam. The first high-resolution structures of human ribosomes are a good example (Amunts et al. 2015; Khatter et al. 2015). However, a recent report has shown that cryo-

EM could also resolve the structure of small membrane molecules at near-atomic resolution (Liao et al. 2013).

1.2.3 Structural features of protein-protein interfaces

The experimental techniques described in the previous section have revealed a large number of complex structures from which valuable information can be derived. For example, after a detailed analysis of these structures, several works have tried to decipher whether protein-protein interfaces have unique characteristics that may help to identify them. Unfortunately, it seems that there is no characteristic that, individually, could be used for such purpose.

Protein-protein interfaces are diverse, and this diversity is somehow related to its interaction type. Interfaces of obligate complexes are more hydrophobic than non-obligate complex interfaces. The latter may sometimes show some hydrophilic character (Susan Jones and Thornton 1997). Permanent complexes interfaces are more closely packed, less planar and with fewer inter-subunit hydrogen bonds than those of transient complexes (S. Jones and Thornton 1996).

Some studies conclude protein-protein interfaces are enriched in aromatic (His, Tyr, Trp, Phe) and aliphatic (Leu, Val, Ile, Met) residues. Contrarily, they are depleted in charged residues (Asp, Glu, Lys) other than arginine, which is the residue type with a highest overall contribution to interfaces (R. P. Bahadur and Zacharias 2008; Conte, Chothia, and Janin 1999).

The mean size of protein-protein interfaces is 1600±400 $Å^2$ (Conte, Chothia, and Janin 1999). Data indicates there is a lower limit size for specific recognition. The lack of interfaces with a size below 800 $Å^2$ suggests that complex formation requires a minimum number of contacts and removal of water molecules (R. P. Bahadur and Zacharias 2008).

High complementarity between complex subunits seems to be a common characteristic of protein-protein interfaces of standard size. This feature is extensively exploited by protein-protein docking protocols (see below). Permanent complexes tend to have higher complementarity than transient complexes.

Several studies have focused on the energetic contributions to binding affinity. Some of them report there is a significant hydrophobic contribution to the binding energy (C. J. Tsai et al. 1996, 1997). However, electrostatics seems to play an important role, too (Xu, Lin, and Nussinov 1997; Sheinerman, Norel, and Honig 2000). Other groups have observed van der Waals contacts between non-polar residues, while polar residues establish hydrogen bonds with residues of the

complementary subunit or water molecules (S. Jones and Thornton 1996; Rodier et al. 2005).

1.2.4 Computational methods for structural modeling of PPIs

All the techniques described in section 1.2.2 are non-high-throughput methods that typically require long deployment times. Thus, as Figure 1.1 reflects, structural biology has severe difficulties following the rate at which high-throughput experimental techniques report new protein-protein interaction data. Even though the number of structural models generated by non-high-throughput experimental methods is exponentially increasing, the gap between the number of identified interactions and the number of available 3D complex structures is growing. Given this situation, some computational methods have been developed to incorporate structural information into interactome networks and bridge the gap.

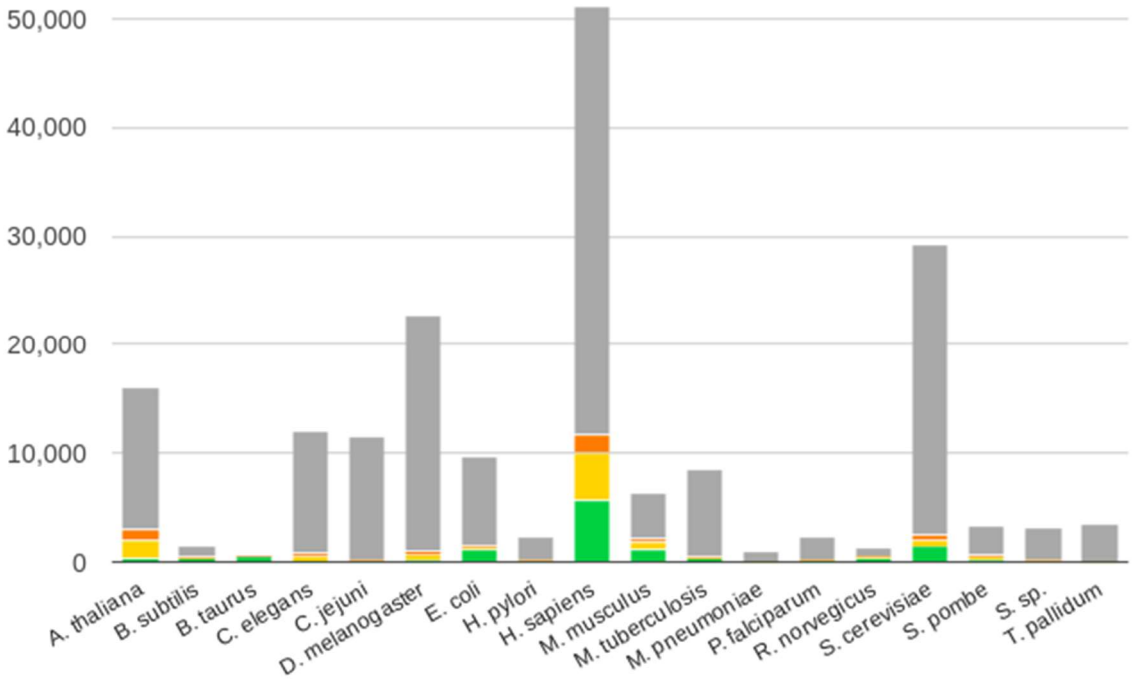

Figure 1.1. Number of interactions in different organisms, with available experimental structure (green), modeled structure from a global template (yellow), modeled structure from a domain-domain template (orange) and without structure (grey). Data from Interactome3D.

1.2.4.1 Molecular dynamics

Classical molecular dynamics (MD) simulation is probably the most widespread method for the study of protein flexibility. Proteins are treated at atomic resolution level using classical force fields, which have been parametrized against

high-quality quantum mechanical calculations or experimental observables obtained from model systems (Orozco 2014).

Since the first molecular dynamics simulation in the late seventies (McCammon, Gelin, and Karplus 1977) the field has experienced sustained progress based on improvements in the parametrization of the force fields, the sampling techniques, and the available computational architectures. Nowadays, it is possible to perform simulations of a standard size protein in the multi-microsecond range (Klepeis et al. 2009) and reach the millisecond regime with special-purpose systems designed for MD (Shaw et al. 2009). Unfortunately, protein dynamics on the microsecond-to-millisecond timescales for large systems is currently out of reach for conventional MD simulations, not to mention the study of protein-protein or protein-membrane interactions. This limitation is even more relevant when some of the most crucial interactions, such as protein folding, protein docking, or concerted domain motions, take place within these timescales and involve multi-molecular complexes of significant size. Simulations within these size and timescale ranges are not computationally feasible due to a size-related-problem coupled with a time-related-problem. As said before, the fast motions of individual atoms translate into collective, wider, and slower movements. To accurately simulate the dynamics of a system, we must compute motions at time-steps of femtoseconds and beyond. For example, to obtain one-millisecond simulation we would have to perform the order of 1012 computations, and this number explodes when we include the size of the systems, which typically can have tens of thousands of atoms. To overcome these limitations, we must substitute the atomistic description of a system for a coarse-grain (CG), lower-resolution description (Orellana 2014).

A large variety of CG approaches have been developed, including normal mode analysis (Case 1994; Tirion 1996; Bahar, Atilgan, and Erman 1997; Atilgan et al. 2001), Gaussian network models (Haliloglu, Bahar, and Erman 1997), FRODA (S. Wells et al. 2005), FIRST (Jacobs et al. 2001) and Go models.

1.2.4.2 Normal mode analysis

Normal mode analysis (NMA) (Case 1994) is one of the most straightforward methods applied to the analysis of near-equilibrium protein motions. NMA assumes the energy of a given molecular structure is quadratic in the vicinity of an energy minimum, and the motions of the molecule can be decomposed into a set of independent harmonic vibrational modes, i.e., normal modes, such that any conformational change can be expressed as a linear combination of them (J. Ma 2004). NMA can be applied with any kind of force field and can also deal with

both atomistic and coarse-grained representations of molecules. However, in practice, CG force fields are preferred due to their significant reduction of computational requirements. A popular CG implementation is the anisotropic network model (ANM) (Atilgan et al. 2001), a class of elastic network model, based on a simplified pairwise harmonic potential, described by a single force constant γ. Molecules are represented as an elastic network of beats connected by springs. Each beat is a residue, represented by its Cα atom, while the springs represent interactions between beats. All springs have the same force constant γ. Only interactions within a given cutoff distance are considered (see Figure 1.2). Despite the simplicity of the model, it has been shown that ANM can reproduce conformational changes observed between bound and unbound protein structures (Petrone and Pande 2006). ANM can also discriminate between low and high-frequency modes. Therefore, ANM is a suited technique to describe collective and large biologically relevant motions associated to the lowest frequency normal modes (Tama and Sanejouand 2001; Zheng and Doniach 2003; Navizet, Lavery, and Jernigan 2004). ANM can further be used for estimating with reasonable accuracy experimental B-factors (Kundu et al. 2002; Kondrashov, Cui, and Phillips 2006), predicting hinge-bending movements (Emekli et al. 2008), and reproducing the flexibility patterns identify by NMR experiments (Yang et al. 2007, 2009) or MD simulations (Rueda, Chacón, and Orozco 2007; Romo and Grossfield 2011; Orellana et al. 2010). Finally, ensembles of discrete conformations generated by ANM have also been used by other computational techniques like protein-protein docking to account for flexibility (Lindahl and Delarue 2005; Dobbins, Lesk, and Sternberg 2008; Rueda, Bottegoni, and Abagyan 2009). Several databases of pre-calculated protein motions like ProMode (Atilgan et al. 2001), MolMovDB (Flores et al. 2006) and iGNM (Yang et al. 2005), and web servers like ElNemo (Suhre and Sanejouand 2004) are available for the scientific community to benefit from this useful technique.

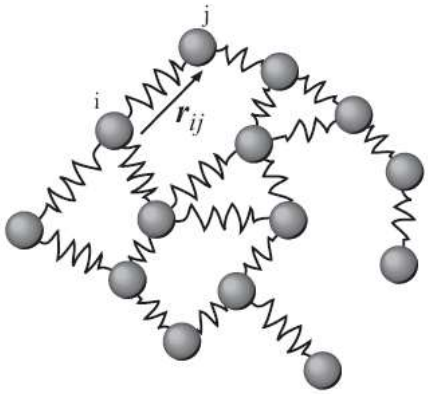

Figure 1.2. Elastic Network Model representation

1.2.5 Protein-protein docking

Neither molecular dynamics nor normal modes are appropriate methods to obtain the structure of a complex from its subunits when no information of the binding mode is available. In these circumstances, many different protein-protein docking methods have been reported in the literature, which can be roughly classified into template-based modeling (TBM) and ab-initio docking (see Figure 1.3).

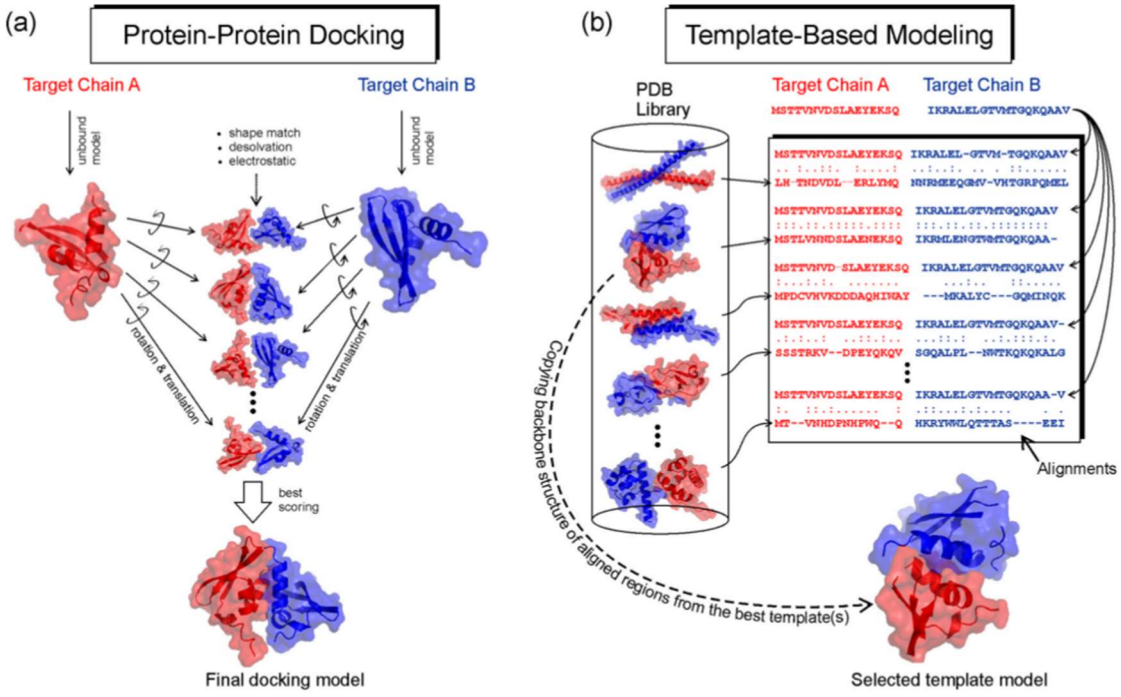

Figure 1.3. Two principal protocols for protein complex structure prediction. Red and blue represent the sequences and structures of two individual chains. a) Rigid-body protein-protein docking, b) template-based modeling. From Szilagyi and Zhang, 2014.

1.2.5.1 Template-based modeling

This technique predicts the unknown 3D structure of protein-protein complexes (targets) using as models the structures of experimentally resolved complexes (templates). TBM was initially developed by the folding community, based on the idea that homologous proteins, with close sequences, also had similar structures (Chothia and Lesk 1986). TMB does not require the structure of the monomer components of the complex. However, its performance strongly depends on the availability of suitable complex structures to use as templates, i.e., the majority of the successful models are for targets that share sequence identity with the templates higher than 40% (Kundrotas et al. 2012; Kundrotas, Vakser, and Janin 2013). Some authors estimate the number of protein interaction types, or "quaternary folds" in nature to be between 10,000 (Aloy and Russell 2004) and 4000 (Garma et al. 2012). Both studies agree that current PDB only covers a small fraction of interaction types and that, given current progress, several decades should pass before a full coverage of the quaternary structure space is available. On the other hand, other authors state that the protein-protein interface space is limited, degenerate and close to being adequately represented in the PDB (M. Gao and Skolnick 2010; Zhang et al. 2010). Several methods (Sinha, Kundrotas, and Vakser 2010; Zhang et al. 2012; Guerler, Govindarajoo, and Zhang 2013; Nurcan Tuncbag et al. 2011; Petrey and Honig 2003; Keskin, Nussinov, and Gursoy 2008) are developed taking advantage of this fact. However, how to exploit interface similarity to model complete protein-protein complexes is still a mostly unsolved issue (Szilagyi and Zhang 2014). TBM methods can be part of genome-wide applications for generating complex models of whole interactomes (Zhang et al. 2012; Singh et al. 2010; Hosur et al. 2012; Mosca, Céol, and Aloy 2013; Lu et al. 2003). Furthermore, template-based modeling and protein-protein docking methods seem to be complementary, and have better results in combination than when they are used individually (Guerler, Govindarajoo, and Zhang 2013; Vreven et al. 2014). In summary, despite its drawbacks, TBM is a promising technique that could achieve excellent results once those drawbacks are overcome.

1.2.5.2 Ab-initio protein-protein docking

Ab-initio protein-protein docking aims to determine the three-dimensional (3D) structure of a protein complex from the 3D structure of its components. Many different methods have been developed since its origins in the late seventies (Wodak 1977; Pincus and Scheraga 1979).

Nowadays, bound protein-protein docking, where the monomers to dock have the same structure as in the native complex, is considered a resolved problem. On the contrary, unbound protein-protein docking where the inputs are structures of the unbound monomers is still an unsolved issue. This fact is important since, in a real situation, the available structures of the protein monomers are not in the bound but the unbound conformation.

Most of the protein-protein docking protocols include the following phases: i) sampling phase, ii) scoring phase and iii) optionally, a final refinement phase.

The goal in the sampling phase is to generate docking poses or conformations as close to the native complex structure as possible. Sampling algorithms should be quick and efficient in covering the conformational space (Halperin et al. 2002). It is essential to bear in mind that the conformational space defined by two interacting proteins has, potentially, 3N dimensions, where N is the total number of atoms of the interacting proteins. Many sampling algorithms have adopted the rigid-body approximation to reduce the number of degrees of freedom of the problem. Within this approximation, interacting proteins are considered rigid bodies during the search. By discarding intrinsic protein flexibility, the conformational space dimension decreases from 3N to six, corresponding to the set of rotations and translations a rigid-body can perform in the Euclidean space.

Once the rigid-body approximation is adopted, it is possible to increase the sampling speed by representing the interacting proteins as grids, and applying Fast Fourier Transform (FFT) algorithms (Katchalski-Katzir et al. 1992) to search the rotational and translational space for those docking poses with the best correlation between grids. This approach, based on the idea that there exists steric complementarity at the protein-protein interfaces, permits an exhaustive search of the full six-dimensional docking space and has become one of the most popular techniques within the docking sampling field. The scheme can incorporate additional grids. For instance, FTDock (Gabb, Jackson, and Sternberg 1997) included an electrostatic grid, PIPER (Kozakov et al. 2006) pairwise interaction potentials, ZDOCK (Chen, Li, and Weng 2003) has incorporated, over the years, electrostatics, desolvation and statistical potential terms (B. G. Pierce, Hourai, and Weng 2011).

The steric complementarity of protein-protein interfaces has also been exploited in other geometric-based methods such as PatchDock (Schneidman-Duhovny et al. 2005a) that, instead of Fourier transforms, incorporates geometric hashing algorithms to compute complementarity between protein surfaces.

Following the sampling phase, in the scoring stage, the objective is to identify

structures similar to the near-native complex within the pool of docking poses previously generated. Most of the docking methods depend on a scoring function based on energetic terms for this task. For instance, pyDock (T. M.-K. Cheng, Blundell, and Fernandez-Recio 2007) scoring function is the combination of electrostatic, desolvation and van der Waals terms. Other methods like SIPPER (Pons et al. 2011) or PIE (Ravikant and Elber 2010) have developed atom or residue knowledge-based statistical potentials. In recent years, some approaches have followed an interesting path by integrating external information into their scoring functions, such as evolutionary data (Andreani, Faure, and Guerois 2013; Ovchinnikov, Kamisetty, and Baker 2014), information coming from SAXS (Jiménez-García et al. 2015) or cryo-EM experiments (de Vries et al. 2016).

Once the scoring function has identified a set of docking candidates, most methods aim to include conformational flexibility by refining them. Usually, this refinement phase tries to incorporate flexibility through normal modes or the optimization of backbone and side-chain atoms. For example, HADDOCK (Dominguez, Boelens, and Bonvin 2003) uses soft potentials and includes water molecules in the refinement phase. FireDock (Andrusier, Nussinov, and Wolfson 2007) restricts the side-chain flexibility to the clashing interface residues and models it by rotamers.

1.2.5.3 Rigid-body approximation: Good and evil

For many years, the rigid-body approximation has made tractable the protein-protein docking problem. In a period where the number of computer resources was much more limited than today, it helped research on the protein-protein interactions field by speeding up both the sampling and the scoring phases to acceptable levels. However, the rigid-body approximation ignores protein flexibility, which is a critical component of protein-protein interactions. Almost all protein-protein interactions are associated with a conformational change. Proteins are entities that interact with the environment to perform their function. In a sense, we could say that proteins modify their environment by modifying themselves. If proteins were rigid bodies, without flexibility, they could not perform their function. Since protein-protein docking is a technique that aims to unravel protein-protein interactions, we could argue it cannot ignore flexibility.

However, the assumption of the rigid-body approximation also has more practical implications. As mentioned above, the bound protein-protein docking problem is considered as solved. However, the unbound protein-protein docking is still an open issue. When differences between the bound and the unbound monomer structures are small, protein-protein methods are likely to get a near-native

solution. The more significant these differences are, the more problems protein-protein methods encounter. Not surprisingly, sampling algorithms and scoring functions have low performances when they are forced to assume rigidity with flexible or very flexible cases. The rigid-body approximation has been fundamental in protein-protein docking research and is still valid for studying instances where proteins change little upon binding. However, if we want to tackle cases with medium to significant structural changes upon binding, our docking algorithms have to incorporate flexibility.

1.2.6 Dynamic aspects of docking

1.2.6.1 Protein flexibility

As mention before, proteins are flexible entities that sample a great number of different conformations. The dynamic character of proteins should not be undervalued. On the contrary, its direct relation to protein function has led to extending the structure-function paradigm to include dynamics. A well-behaved protein must have the specific dynamical properties needed to perform its function. The intrinsic motions a protein can perform are determined by its structure that, ultimately, is defined by the protein sequence. We can follow this same trail to recognize how evolutionary selection acts upon the functional movements, structures, and sequences required for function (Orellana 2014).

Protein dynamics are characterized by the timescale (kinetic component) and the amplitude and directionality of the fluctuations (structural component) (K. Henzler-Wildman and Kern 2007). Table 1.2 shows the amplitude and timescale of protein motions. In one extreme of the spectra, we find the fast, femtosecond bond vibrations. In the other, the slow, long motions that proteins perform while folding, Interestingly, these motion levels can be organized hierarchically. The collective motion of atoms at one level translates into the next, giving rise to slower and broader movements (K. A. Henzler-Wildman et al. 2007). For example, local atomic vibrations are transmitted via hydrogen bonds creating higher amplitude motions (Orellana 2014).

Time scale (s)	Amplitude (Å)	Description
10^{-15} - 10^{-12}	0.001 – 0.1	- Bond stretching, angle bending, dihedral motion - Side chain motions - Loop motions
10^{-12} - 10^{-9}	0.1 – 10	- Helix motions - Subunit and domain motions (hinge bending)
10^{-9} - 10^{-6}	1 – 100	- Small peptide folding - Helix coil transitions
10^{-6} - 10^{-1}	10 – 100	- Folding and unfolding

Table 1.2. Timescales and amplitudes of molecular motions

Protein dynamics can be studied experimentally by combining atomic-resolution or near-atom-resolution techniques like X-ray crystallography, NMR, cryo-electron microscopy and SAXS, with low-resolution spectroscopic and classical biophysical methods like fluorescence, circular dichroism, infrared spectroscopy, absorbance, Raman spectroscopy and electron paramagnetic resonance. The former techniques provide the structural component of protein dynamics, while the latter supply the complementary, kinetic information (K. Henzler-Wildman and Kern 2007). However, the study of protein flexibility by experimental methods is still a challenging, high-demanding task. On the other hand, computational approaches can be perfect companions. Theoretical models can produce valuable information and predictions that may guide new experiments.

Protein flexibility is thus critical to understand protein-protein interactions at the molecular level, and several mechanisms have been proposed to describe the dynamic aspects of protein association.

1.2.6.2 Protein binding mechanism

Several theoretical models have been proposed to explain protein binding mechanism.

The lock-and-key model (Fischer 1894) postulated by Emil Fischer in 1894, states that proteins remain fixed upon complex formation and binding occurs when proteins with geometrical shape complementarity fit precisely into one another. This model could explain binding affinity and specificity. However, contrary to the rigid-body assumption inherent in the lock-and-key formulation, proteins are flexible entities that are subject to conformational changes upon binding. Decades later, the induced-fit model (Koshland 1958) suggested proteins initially interact

from the unbound conformations. Then, recognition between proteins induces a change in their 3D structures until the complex forms. The induced-fit paradigm could explain a certain degree of conformational flexibility (Echols, Milburn, and Gerstein 2003), as well as the promiscuity found in some proteins (Tidow and Nissen 2013). However, some studies (Bosshard 2001) argued that the induced-fit model alone could not justify significant structural changes upon complexion. Furthermore, the induced-fit model could not account for X-ray, cryo-electron microscope images and NMR data showing ensembles of conformations, dynamically fluctuating between them, that contain structures similar to the bound state (Boehr, Nussinov, and Wright 2009; Esteban-Martín, Bryn Fenwick, and Salvatella 2012).

Straub already proposed the idea of selective binding to a specific conformation within the ensemble in 1964 (Straub and Szabolcsi 1964). This hypothesis was supported by Zavodszky experimental results (Závodszky, Abaturov, and Varshavsky 1966). Twenty-five years later, Frauenfelder, Sligar, and Wolynes described the energy landscape of proteins (Frauenfelder, Sligar, and Wolynes 1991). This original work led to the model of 'conformational selection and population shift' (B. Ma et al. 1999; C. J. Tsai et al. 1999; Csermely, Palotai, and Nussinov 2010). According to this model, the native state of a protein is not a single conformation but an ensemble of closely related structures in equilibrium. The most suitable conformers will bind and stabilize with the interacting partner, shifting the equilibrium toward complex formation (Tobi and Bahar 2005).

Recent works show some cases where conformational adjustment often follows conformational selection (Grünberg, Leckner, and Nilges 2004; Wlodarski and Zagrovic 2009). These new data support the development of an extended conformational selection model. The extended conformational selection model describes a general scenario, where both selections, and adjustment-type steps, favored by electrostatics and water-mediated hydrogen bonding, follow each other. Partner proteins can take different paths of consecutive conformational selection and adjustment steps (called binding trajectories (C.-J. Tsai, del Sol, and Nussinov 2008; Antal, Böde, and Csermely 2009), but converge to a common final state. Within the new framework, the lock-and-key, the induced fit, the original conformational selection and the conformational selection plus adjustment models are all special cases of the extended conformational selection model (Csermely, Palotai, and Nussinov 2010).

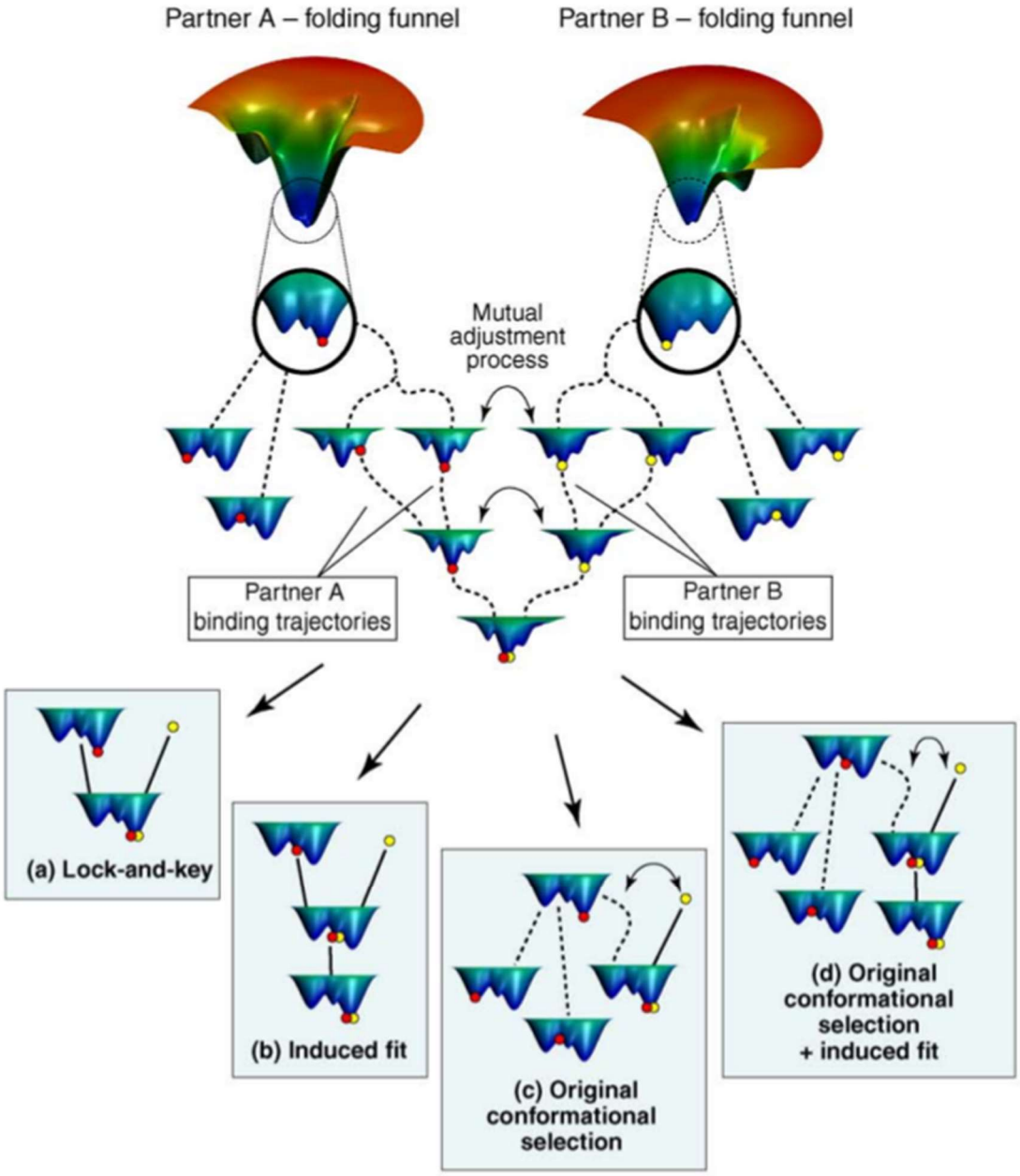

Figure 1.4. Schematic representation of the extended conformational-selection model. a) classical lock-and-key model, b) classical induced-fit model, c) classical conformational-selection model, d) conformational-selection-plus-induce-fit model. From Csermely et al., 2010.

1.2.7 Including flexibility in docking

Incorporating flexibility in docking represents one of the current challenges in the protein structure modeling field. This flexibility affects both backbone and side-chain conformational changes. The high number of degrees of freedom of the systems not only increases the required computational resources but also results in a higher rate of false positive conformations (Andrusier et al. 2008). Although there are methods that handle backbone and side-chain flexibility simultaneously, most groups treat them separately.

Most of the docking algorithms deal with side-chain flexibility in the refinement stage once the docked solutions have been generated. Most of them use rotamer libraries derived from statistical analysis of side-chain conformations in resolved protein structures (Lovell et al. 2000; Shapovalov and Dunbrack 2011). In this case, side-chain prediction can be treated as a combinatorial optimization problem, where the goal is to find the residue rotamer combination that minimizes the energy of the system. It has been proved this is an NP-hard (N. A. Pierce and Winfree 2002) and inapproximable (Chazelle, Kingsford, and Singh 2004) problem. Several algorithms like SCWRL (Canutescu, Shelenkov, and Dunbrack 2003) or FireDock (Andrusier, Nussinov, and Wolfson 2007) have been developed to tackle this issue. To simplify the complexity of the problem, a family of methods prun rotamers based on the dead-end elimination (DEE) method (Desmet et al. 1992). In the same direction, the Residue-Rotamer-Reduction (R3) method, in addition, use a residue reduction procedure (Xie and Sahinidis 2006) to decrease the resources and computational time required to solve the side-chain conformation problem.

A great variety of methods have been developed to handle backbone flexibility. Some docking methods use a soft interface approach by which some degree of interpenetration between the docked molecules is allowed. These methods can only model side-chain and small backbone rearrangements. One of the drawbacks of this strategy is that steric clashes are frequently introduced into the docked results, and a further refinement stage is necessary. For example, many methods rely on a short energy minimization of the docked solutions in a final, refinement stage, using algorithms based on all-atom force fields like CHARMM (Brooks et al. 2009) or AMBER (Case et al. 2005). This final step that we could relate to the induced-fit binding mechanism with small conformational changes reduces the number of steric clashes and improves the energetic description of the solutions, which may affect the scoring performance positively. Other families of methods have been developed to tackle cases were substantial conformational changes are present. For example, programs like FlexDock (Schneidman-Duhovny et al. 2005b), have focused on identifying hinge regions. Once the hinge regions have been detected, the proteins are dissected into their rigid subdomains, and a two-body docking is performed with the various combinations of subdomains (Bonvin 2006). Finally, a group of methods, that we could associate to the conformational selection binding mechanism, follow an ensemble docking approach. These techniques perform a rigid body docking of ensembles of conformations previously generated. The ensembles may be taken from experimental sources like crystal and NMR structures, or computational sampling methods based on molecular

dynamics (MD), normal modes, essential dynamics, etc. The conformers of the ensembles may be docked one by one (cross docking), or all together using algorithms based on the mean-field approach where the conformers are assigned weights that change according to the Boltzmann criterion (Bastard et al. 2003).

1.2.8 Protein-protein docking assessment

1.2.8.1 Protein-protein docking benchmarks

Benchmarks are useful tools that serve as curated repositories where researchers can efficiently gather data to perform their studies. They also provide a common ground on which to compare the performance of the different algorithms.

During the years, a protein-protein docking benchmark initially developed in Z. Weng lab has been steadily growing (Chen et al. 2003; Mintseris et al. 2005; Hwang et al. 2008, 2010; Vreven et al. 2015). Currently, the benchmark is a repository of 230 non-redundant, high-quality 3D-structures of protein-protein complexes and the unbound structures of their components. The benchmark incorporated, in its last version 5.0 (Vreven et al. 2015), experimental binding affinity values for 179 cases, addressing the growing interest of the docking community in developing algorithms to predict not only the structure of protein-protein complexes but in the thermodynamics and energetic aspects of PPIs.

DOCKGROUND (Douguet et al. 2006; Y. Gao et al. 2007) is another useful, automatically updated benchmark for protein-protein docking development. It contains bound, and unbound structures, and lately has included a set of docking decoys (Shiyong Liu, Gao, and Vakser 2008) that can be used to improve docking algorithms performance.

SKEMPI (Moal and Fernández-Recio 2012a) is a manually curated database containing over 3000 experimentally measured changes in binding free energy upon mutation in heterodimeric complexes with at least one structure available in the PDB (Bernstein et al. 1977). SKEMPI also contains, when possible, data on changes in entropy, enthalpy and rate constants.

1.2.8.2 CAPRI

Following CASP example (Moult et al. 1995), the docking community established in 2001 the Critical Assessment of PRedicted Interactions, or CAPRI (Janin et al. 2003), a community-wide, worldwide experiment for protein-protein docking.

CAPRI rounds start whenever experimentalists offer suitable targets. In each round, the organizers ask the participants to model a set of target protein complexes, whose structures have been experimentally resolved but are not yet

publicly available. CAPRI is a double-blind experiment, i.e., the participants do not know the complex structure to solve, and the organizers do not know the correspondence between the submissions and the participant groups. Each round comprises two separate competitions: predictors and scorers. In the predictor's contest, participants are asked to submit a total of 100 complex models per target. Depending on the target, organizers may provide structures, models or sequences of the complex components for the participants to build their models. Models should be ranked, i.e., best models should be assigned the lowest ranks. The ten best-ranked models are selected for evaluation by the organizers. In the scorer's competition, participants are asked to evaluate a pool of docking complex models made up from the 100 models each predictor's group submitted in the predictor's contest. Each scorers group must provide a ranked list of their ten best models. For both competitions, group predictions are evaluated according to the number of native contacts, ligand RMSD (LigRMSD), interface RMSD (IntRMSD) and rank of the near-native solutions (Méndez et al. 2003; Lensink, Méndez, and Wodak 2007).

A total of 6 evaluation CAPRI meetings have been celebrated, corresponding to more than 100 targets. CAPRI context has not only proved to be a valuable tool to assess protein-protein docking algorithms. It has fostered the creation of a docking community and has guided it, suggesting the way forward and the new challenges to address. For example, besides protein-protein complex targets, CAPRI has proposed targets involving binding affinity, sugar binding, and interface water molecule prediction (Lensink and Wodak 2013a; Moretti et al. 2013; Lensink et al. 2014). In the same spirit, the first joint CASP-CAPRI initiative took place in CAPRI round 30 (Lensink et al. 2016). The experiment encouraged participants to improve their docking protocols in the challenging area of docking homology models.

1.3 Energetics of protein-protein interactions

1.3.1 Binding affinity

The binding of two proteins can be interpreted as a reversible process. When this process has reached equilibrium, it is governed by the law of mass action (Kastritis and Bonvin 2013). For example, we could represent a simple reaction between two proteins A and B as follows:

$$[A] + [B] \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} [AB]$$

Where [A] and [B] are the concentration of the unbound proteins (reactants), and [AB] is the concentration of the bound complex (product). k_{on} and k_{off} denote, respectively, the association and dissociation rate constants, and are typically measured in $M^{-1}s^{-1}$.

In this context, the binding affinity is the strength of the binding interaction between two or more molecules. Binding affinity can be expressed by the equilibrium dissociation constant (K_d). At equilibrium, we can write K_d as a function of the concentrations of reactants and product, and also in terms of the dissociation and association rate constants:

$$K_d = \frac{[A][B]}{[AB]} = \frac{k_{off}}{k_{on}}$$

Additionally, the binding affinity can be associated with the Gibbs free energy (ΔG) by the following equation:

$$\Delta G = -RT \ln K_d = \Delta H - T\Delta S$$

where R is the gas constant, T is the absolute temperature, and ΔH and ΔS represent the changes in enthalpy and entropy, respectively.

1.3.1.1 Binding affinity determination by experimental methods

Different experimental methods have been developed for determining binding kinetics (Vuignier et al. 2010). These methods can be grouped into direct and indirect techniques. Direct methods like gel filtration, ultracentrifugation or equilibrium dialysis, measure the actual concentration of bound and unbound proteins. Indirect methods, such as absorbance, resonance or fluorescence spectroscopy, infer the concentrations from a signal that is measured, assuming that the measured signal is directly proportional to the concentrations of the proteins. Three of the most frequently used methods to measure binding affinity belong to the indirect methods group: isothermal titration calorimetry (ITC) (Ladbury and Chowdhry 1996), surface plasmon resonance (SPR) (Willander and Al-Hilli 2009) and fluorescence-based methods (Masi et al. 2010).

1.3.1.2 Binding affinity prediction by computational methods

Calculation of protein-protein binding affinities is also a fundamental topic in structural modeling. This time, the focus is not determining the atom coordinates of protein complexes but understanding the energy determinants of protein binding. This question has vast practical implications in fields such as protein engineering (Kortemme and Baker 2004; Sharabi et al. 2011), de novo interface design (Fleishman, Whitehead, Ekiert, et al. 2011), computational mutagenesis (Ben-Shimon and Eisenstein 2010) and peptide therapeutics (Rao and Kumar

2008; Kumar et al. 2011).

Different computational strategies have been developed to address the problem including Monte Carlo conformational searches (Abagyan and Totrov 1994), free-energy perturbations (Kollman 1993), Poisson-Boltzmann (Honig and Nicholls 1995), generalized-Born solvation (Qiu et al. 1997), and atomic continuum electrostatic calculations (Schaefer and Karplus 1996). All these techniques require a lot of computational resources and cannot be applied in the context of protein-protein docking, where a significant number of structural models have to be evaluated (Kastritis and Bonvin 2010). Less sophisticated methods have been suggested (Horton and Lewis 1992; Baker and Murphy 1998; X. H. Ma et al. 2002). One of the proposed alternatives consists on predicting binding affinity with docking scoring functions. After all, many docking scoring functions are based on energetic terms and, ideally, binding affinity prediction and docking scoring should converge at some point in the future (Kastritis and Bonvin 2013). The development of databases with binding affinity data (Kastritis et al. 2011; Vreven et al. 2015) has allowed to develop and test these coarse-grain alternatives. At present, docking scoring functions cannot predict experimentally measured binding affinity with sufficient precision (Kastritis and Bonvin 2010). These scoring functions have been designed with the main goal of identifying the models structurally similar to the native complex but fail to reproduce the underlying energetics during complex formation. Several reasons may explain the difficulties encountered when predicting binding affinity (Kastritis and Bonvin 2013). First, the quality of the experimental data has been questioned. For example, atom coordinates from X-ray crystallography might be ambiguous sometimes. It has been also reasoned that predicting experimental affinity values obtained in solution, from crystal structures can introduce some noise. On the other hand, most of the current models do not consider the conformational changes that are taking place during binding. They also ignore variables like pH, temperature, and concentration of reactants, or implicitly assume the lock-and-key binding model, overlooking more complicated binding mechanisms like allosteric regulation, induced-fit or conformational selection. Finally, it should be noted that the performance of these methods depends on the quality and size of the experimental data used for learning and testing. We remark above the crucial role of a few databases with experimental binding affinity data in the development of novel predictors. That being true, we should also note that the data compiled in these databases is still limited and does represent a tiny portion of protein-protein binding heterogeneity. In this context, several algorithms showing high correlations with the data available at the time, have seen their performance

decrease when tested against new experimental measurements (Kastritis and Bonvin 2010).

1.3.2 Protein binding hot-spots residues

Protein-protein interface studies have revealed the existence of a set of residues, called hot-spots, which despite comprising only a small fraction of the interfaces, confer most of the binding energy of protein-protein interaction, thus being crucial for the stability of complexes. Typically, hot-spot residues are defined as those residues contributing in more than two kcal/mol to the binding energy of the complex (Thorn and Bogan 2001; Moreira, Fernandes, and Ramos 2007b). Several studies have focused on the characterization of hot-spot residues. Some works have estimated that 9.5% of interface residues are hot-spots. Hot-spot composition is enriched in tryptophan (21%), arginine (13.3%), and tyrosine (12.3%) residues (Lichtarge, Bourne, and Cohen 1996; Bogan and Thorn 1998). They frequently appear surrounded by a set of less energetically important residues, adopting an O-ring disposition, whose purpose seemed to be to occlude hot-spots from water (Bogan and Thorn 1998). Hot-spots are not randomly distributed across the protein interfaces, but clustered in what have been called hot regions (Keskin, Ma, and Nussinov 2005). They are structurally conserved (Schreiber and Fersht 1995; Lockless and Ranganathan 1999; Keskin, Ma, and Nussinov 2005; Caffrey Daniel R. et al. 2009), and have a cooperative character that may explain the extraordinary capacity of proteins to modulate its binding affinity and specificity to different partners (Cukuroglu et al. 2014). Hot-spots research may reveal essential mechanisms of protein-protein interactions and is also very important in drug design. For years it was thought that protein-protein interactions were not druggable since they are usually large, planar and lack the pockets targeted by classic protein inhibitors. However, the discovery of the role of hot-spots gave raised to a set of studies that demonstrated that binding affinity and specificity could be attained by epitopes consisting in a small number of residues (J. A. Wells 1991; J. A. Wells and de Vos 1993; J. A. Wells 1996; DeLano 2002). Experimentally, hot-spots can be identified by alanine scanning mutagenesis (J. A. Wells 1991), alanine shaving (Jin and Wells 1994), and residue grafting (Jin and Wells 1994). These techniques are expensive, time-consuming, labor-intensive, and cannot be applied on a large scale. Therefore, different computational methods for hot-spot identification have been developed.

Reported hot-spot prediction algorithms are based on evolutionary information (Buyong Ma et al. 2003), the energetic contribution of residues to the binding energy (Guerois, Nielsen, and Serrano 2002), the structural features of hot-spots

(Landon et al. 2007; Nurcan Tuncbag, Gursoy, and Keskin 2009; N. Tuncbag, Keskin, and Gursoy 2010), protein-protein docking (Solene Grosdidier and Fernandez-Recio 2008), machine learning techniques (Ofran and Rost 2007; Assi et al. 2010; Melo et al. 2016) or molecular dynamics (MD) (Moreira, Fernandes, and Ramos 2007a). It should be noted that most of them require the structure of the complex to compute their predictions. Noticeably, pyDockNIP (Solene Grosdidier and Fernandez-Recio 2008), is among the few ones that does not require the complex structure and admits the unbound structures or models of the complex subunits as valid inputs. Interestingly, this method makes its predictions based on the normalized interface propensity (NIP) values of the residues, derived from rigid-body docking simulations.

1.4 Other protein interactions: Protein-RNA docking

Proteins interact not only with other proteins but with different biomolecules. Protein-RNA interactions constitute a good example. Protein-RNA interactions play a crucial role in critical cellular processes such as regulation, RNA translation, RNA transport and localization or RNA post-transcriptional modification, e.g., splicing and polyadenylation. This functional diversity extends to the variety of RNA types involved, including transfer-RNA (tRNA), ribosomal-RNA (rRNA), messenger-RNA (mRNA) and micro-RNA (miRNA) (Susan Jones 2016). RNA can be found in flexible single strands (ssRNA), but also in complex tertiary folds in the form of double-stranded RNA (dsRNA). Although protein-RNA complexes are essential for cell function, their structures are underrepresented in the PDB (Bernstein et al. 1977), as shown in Table 1.1. Protein-RNA complexes are difficult to crystallize due to their conformational flexibility. Standard NMR techniques can be applied to solve RNA-protein complexes below 50 kDa, but the size of large macromolecular complexes such as the spliceosome or the ribosome is a limiting factor for this technique. On the other hand, the latest hardware and software developments in cryo-EM may help to convert this technique in the standard for RNA-protein structural characterization. However, given that experimental structural determination of protein-RNA complexes at high resolution is challenging, computational approaches may help to fill the gap and foster the understanding of the processes where they are involved.

Target 34 was the first protein-RNA case in a CAPRI context (Janin 2010b; Lensink and Wodak 2010b). It was the first time the docking community evaluated the potential of their protocols, designed for protein-protein docking,

to deal with protein-RNA complexes. Initially, several docking methods initially developed for docking proteins were adapted to predict protein-RNA models such as GRAMM (Katchalski-Katzir et al. 1992), Haddock (van Zundert et al. 2016), Hex (Ritchie and Kemp 2000), PatchDock (Schneidman-Duhovny et al. 2005a), and FTDock (Gabb, Jackson, and Sternberg 1997). Later, specific protein-RNA scoring potentials were developed. They are based on reverse Boltzmann statistics as DARS-RNP and QUASI-RNP (Tuszynska and Bujnicki 2011), residue propensities like OPRA (Pérez-Cano and Fernández-Recio 2010) and statistical mechanics as ITScore (Huang and Zou 2014). Several benchmarks (Barik et al. 2012; Pérez-Cano, Jiménez-García, and Fernández-Recio 2012) and docking algorithms (Pérez-Cano et al. 2009; Guilhot-Gaudeffroy et al. 2014; Huang and Zou 2014; Tuszynska et al. 2015) are now available. In general, the performance of these methods is acceptable in complexes between proteins and double-stranded RNA, but quickly decreases when they have to deal with the flexibility of single-stranded RNA (ssRNA). New methods try to address this issue by modeling the ssRNA structures from small sequence fragments (Beauchene, Vries, and Zacharias 2016). Even though structural biology has seen substantial steps forward in the field of protein-RNA interactions in the last decade, much more work is needed to unravel protein-RNA interactions.

2 Objectives

Protein-protein interactions (PPI) are essential in virtually all biological processes, and its study at the molecular level is a field of high scientific relevance and therapeutic interest. Several computational approaches have been developed to model protein interactions by docking. Despite all the advances in recent years, computational protein docking faces significant challenges ahead, such as dealing with highly flexible interacting proteins or weak interactions. Thus, new strategies for better energetic description and exploration of the protein docking landscapes are necessary in order to overcome current limitations.

This Ph.D. thesis has focused on developing new computational tools to deal efficiently with major challenges in protein docking, with the general goal of achieving a better understanding of binding energetics and mechanism.

The specific objectives of this thesis are:

1. Analyse the state-of-the-art of protein-protein docking methods and current limitations.
2. Develop new methodologies for fast exploration of docking energy landscapes, by improving energy minimization and flexible refinement.
3. Improve description of docking landscapes by conformational ensembles and clustering analysis.

4. Compute docking energy at the residue level for the description of mutational effects
5. Study the energetics of protein-RNA interactions by docking.

3 Methods

3.1 Standard protein-protein docking with pyDock

In this thesis, we have used, evaluated and helped to optimize the rigid-body protein-protein docking protocol pyDock, which can be structured in two different phases: sampling and scoring.

3.1.1 Sampling: FTDock

In the sampling phase, we used the FFT-based docking program FTDock 2.0 (Gabb, Jackson, and Sternberg 1997), with electrostatics and a grid resolution of 0.7 Å, to generate 10,000 docking poses. FTDock is based on the algorithm developed by Katchalski-Katzir et al. (Katchalski-Katzir et al. 1992). FTDock discretizes the interacting molecules onto grids. Keeping the bigger molecule (receptor) fixed and moving the smaller one (ligand), FTDock performs a global scan of the translational and rotational space. The scoring method is based on the shape complementarity and favorable electrostatic interactions between the two grids. The scoring function is designed to permit some degree of surface overlapping to account for side-chain flexibility and allow to perform the docking of unbound structures. FTDock uses the fast Fourier transform (FFT) and Fourier correlation theory to reduce the overall computation required.

3.1.2 Scoring: pyDock scoring function

In the scoring phase, the docking poses generated in the sampling stage were scored with pyDock scoring function (T. M.-K. Cheng, Blundell, and Fernandez-Recio 2007), composed of three energetic terms: electrostatics, desolvation, and van der Waals (see Eq. 3.1).

$$E_{pyDock} = E_{ele} + E_{desolv} + w_{vdw}E_{vdw} \quad \text{Eq. 3.1}$$

The pairwise electrostatics energy (Eq. 3.2) is based on the Coulombic electrostatics energy with a distance-dependent dielectric constant ($\mathcal{E} = 4.0d$), which was explicitly calculated for all intermolecular *ij* atom pairs separated a distance d_{ij}, with q atomic charges taken from the AMBER 94 force field (Cornell et al. 1995). Pairwise electrostatic interaction energy values were truncated to a maximum and minimum value of +1.0 and -1.0 kcal/mol, respectively (to avoid excessive dependence on incorrect geometries).

$$E_{ele} = 332.0 \sum_{i}^{rec} \sum_{j}^{lig} \frac{q_i q_j}{4d_{ij}^2} \quad \text{Eq. 3.2}$$

Desolvation energy (Eq. 3.3) represents the effective water-to-interface atomic desolvation energy of protein molecules, and it is based on the changes in the atomic accessible surface areas (ASAs) upon binding, with atomic desolvation parameters (ADPs) initially optimized for rigid-body protein-protein docking (Fernández-Recio, Totrov, and Abagyan 2003).

$$E_{desolv} = \sum_{i}^{rec} ADP_i \cdot \Delta ASA_i + \sum_{j}^{lig} ADP_j \cdot \Delta ASA_j \quad \text{Eq. 3.3}$$

If we defined the buried surface area (BSA) of a given atom as

$$BSA_i = \Delta ASA_i = ASA_i^{unbound} - ASA_i^{docked} \quad \text{Eq. 3.4}$$

we can express Eq. 3.3 as

$$E_{desolv} = \sum_{i}^{rec} ADP_i \cdot BSA_i + \sum_{j}^{lig} ADP_j \cdot BSA_j \quad \text{Eq. 3.5}$$

The Lennard-Jones van der Waals energy (Eq. 3.6) was explicitly calculated for all intermolecular *ij* atom pairs separated a distance d_{ij}, with atomic parameters for equilibrium radii (r) and well depth (e) taken from AMBER 94 force field (Cornell et al. 1995). Atomic van der Waals interaction energy values were truncated to a maximum of 1.0 kcal/mol (again to avoid excessive noise from the docking of rigid-body surfaces). pyDock energy function (T. M.-K. Cheng, Blundell, and Fernandez-Recio 2007) is defined in Eq. 3.1, where E_{VDW} is scaled by a weight w_{VDW} set to 0.1.

$$E_{vdw} = \sum_{i}^{rec} \sum_{j}^{lig} \sqrt{e_i e_j} \left(\left(\frac{r_i + r_j}{d_{ij}} \right)^{12} - 2 \left(\frac{r_i + r_j}{d_{ij}} \right)^{6} \right) \qquad \text{Eq. 3.6}$$

3.1.3 Protein-protein docking benchmark

We used the protein-protein docking benchmark version 4 (Hwang et al. 2010) to test our protein docking methods. The benchmark contains structures of proteins with a high-resolution structure in both the unbound and bound states deposited in the PDB (Bernstein et al. 1977). All cases must have a sequence length longer than 30 amino acids. X-ray structures must have a resolution better than 3.25 Å. Redundancy of the protein complexes at the family level has been avoided using the Structural Classification of Proteins (SCOP) database (Murzin et al. 1995). Biological assembly information from the PDB has been used to distinguish between crystal contacts and biological complexes.

Protein-protein benchmark 4 contains a total of 176 cases. According to their biochemical function, these cases can be classified as enzyme-inhibitor (52), antibody-antigen (25) and other functions (99). The authors have also classified the cases according to the expected difficulty for docking algorithms as easy/rigid body (121), medium (30) and difficult (25). This classification is based on the structural differences between the bound and the unbound conformations, as measured by the interface Cα-RMSD (Cα-IntRMSD) and the fraction of non-native residue contacts ($f_{non\text{-}nat}$) (Méndez et al. 2003). Rigid-body cases are those cases with Cα-IntRMSD ≤ 1.5 Å and $f_{non\text{-}nat} \leq 0.4$, medium cases are cases with 1.5 Å < Cα-IntRMSD ≤ 2.2 Å, or Cα-IntRMSD ≤ 1.5 Å and $f_{non\text{-}nat} > 0.4$ Å, and difficult cases those with Cα-IntRMSD > 2.2 Å. To compute these quantities, we first should superimpose the unbound structures of receptor and ligand into the structure of the complex. Cα-IntRMSD is defined as the RMSD of the interface Cα atoms between the superimposed unbound and bound structures. The interface is defined as all residues with at least one atom located within 10 Å of the other protein (Chen and Weng 2002). $f_{non\text{-}nat}$ is defined as the number of non-native

residue-residue contacts between the unbound structures divided by the total number of contacts in the complex. A pair of residues of receptor and ligand were considered in contact if any of their atoms were within 5 Å (Méndez et al. 2003).

3.1.4 Assessment of protein-protein docking results

Assessment of the docking solutions was done by calculating their Cα ligand root-mean-square deviation (Cα-LigRMSD) with respect to the reference complex structure. We computed Cα-LigRMSD as the RMSD of the Cα atoms of the model and target ligands after superimposing the receptor subunits by least-square minimization on its equivalent Cα atoms. In oligomeric proteins, in which additional molecules of the ligand protein could bind to symmetric sites of the receptor, we calculated the Cα-LigRMSD with respect to all the different symmetric positions and chose the smallest value. The predictive success rates were computed as the percentage of cases of the benchmark set for which a near-native docking solution was found within the top N docking models as sorted by the scoring function. We considered a solution was near-native if its Cα-LigRMSD with respect to the reference complex was not greater than 10 Å.

3.2 CAPRI evaluation protocol

The Critical Assessment of PRedicted Interactions (CAPRI) (Janin et al. 2003) is a community-wide experiment designed to test the reliability and accuracy of protein docking protocols. Each CAPRI round comprises two separate competitions: predictors and scorers. In predictors, the participants receive information from the individual components of a protein complex or target that they have to model. In most cases, target complexes are formed by two proteins. Typically, the largest is called receptor and the smallest ligand. In scorers, the participants are not asked to build models of the protein complex, but to evaluate a set of models made up from the structures submitted in predictors and identify the near-native ones. For each group, the ten best-ranked models are selected for assessment by the organizers both in predictors and scorers. Table 3.1 summarizes the criteria used to evaluate the quality of the models submitted by the participants. The same rules apply for both competitions. These rules try to assess both the correct geometry of the models compare to experimental structure, and their biological relevance. The quantities computed for each predicted model are the number of clashes between the docked proteins, the fraction of native contacts f_{nat}, the RMSD of the backbone atoms of the ligand (BB-LigRMSD) and the RMSD of the backbone atoms of the interface residues (BB-IntRMSD). Models with clashes above a certain threshold are rejected. The threshold is set to the

average number of clashes in all the models submitted by the different groups for the given target, plus two times the standard deviation. A clash is defined as a pair of interacting atoms less than 3 Å apart (Lensink and Wodak 2010a). The geometric quality of the models is derived from the BB-LigRMSD, computed as the RMSD of the backbone atoms, i.e., C, Cα, O, and N, of the model and target ligands after superimposing the receptor subunits by least-square minimization on its equivalent residues. As the BB-LigRMSD might give poor estimates when the ligands are large, the BB-IntRMSD is also computed. BB-IntRMSD is defined as the RMSD between the equivalent backbone atoms of the superimposed interfaces of model and target, where the interface is given by all residues with atoms less than 5 Å apart in the target structure (Janin 2010a). The biological relevance of the models is evaluated computing f_{nat}, i.e., the fraction of correctly identified residue-residue contacts. Two interacting residues are considered in contact if any of its atoms are within 5 Å apart. The f_{nat}, BB-LigRMSD and BB-IntRMSD values of a submitted model determine its classification as a high, medium, acceptable or incorrect model according to the criteria shown in Table 3.1.

Quality	Criteria
High	$f_{nat} \geq 0.5$ AND (BB-LigRMSD $\leq$ 1 Å OR BB-IntRMSD $\leq$ 1 Å)
Medium	$(0.3 \leq f_{nat} < 0.5)$ AND (BB-LigRMSD $\leq$ 5 Å OR BB-IntRMSD $\leq$ 2 Å)
	$f_{nat} \geq 0.5$ AND BB-LigRMSD $>$ 1 Å AND BB-IntRMSD $>$ 1 Å
Acceptable	$(0.1 \leq f_{nat} < 0.3)$ AND (BB-LigRMSD $\leq$ 10 Å OR BB-IntRMSD $\leq$ 4 Å)
	$f_{nat} \geq 0.3$ AND BB-LigRMSD $>$ 5 Å AND BB-IntRMSD $>$ 2 Å
Incorrect	$f_{nat} < 0.1$ OR (BB-LigRMSD $>$ 10 Å AND BB-IntRMSD $>$ 4 Å)

Table 3.1. Quality criteria used to evaluate the submitted models (Lensink, Méndez, and Wodak 2007).

3.2.1 Generation of rigid-body docking poses for the predictors experiments

In all targets, except for T100-101, we used FTDock (Gabb, Jackson, and Sternberg 1997) with electrostatics and 0.7 Å grid resolution and ZDOCK 2.1 (Chen and Weng 2003) to generate 10,000 and 2000 rigid-body docking poses, respectively, in the same conditions as previously described (Solène Grosdidier et

al. 2007). For targets T53, T54, T57, T58 in the fifth CAPRI edition, and T59, T96-97, T103-105 in the sixth CAPRI edition we generated an additional pool of flexible docking poses using SwarmDock standard protocol (Moal and Bates 2010; Torchala et al. 2013; Li, Moal, and Bates 2010). As SwarmDock objective function we used DFIRE score (Song Liu et al. 2004), but without the final rescoring phase. In T46 we generated an additional pool of 10,000 solutions using FTDock without electrostatics and 1.2 Å grid resolution. In T46 and T47, we used RotBUS (Solernou and Fernandez-Recio 2010) to generate 59,112 and 41,021 additional docking poses, respectively. In T50, given the large size of H1N1 influenza virus hemagglutinin protein, we increased the number of translations selected from each rotation from 3 (default) to 10, generating a total of 92,432 FTDock docking poses. For target T106 we used LightDock (Jiménez-García et al. 2018a) to generate around 3000 additional flexible docking poses with DFIRE (Song Liu et al. 2004) as the objective function. Cofactors, water molecules, and solvent ions were not included in our docking calculations.

3.2.2 Scoring of the docking poses for both the predictors and the scorers experiment

We scored the docking models generated by the methods mentioned above, by applying pyDock default protocol described in section 3.1. For some targets, we found experimental information on possible interface residues, which were included in the final scoring as distance restraints with pyDockRST (Chelliah, Blundell, and Fernández-Recio 2006) in targets T60-64, or used as a final distance-based filtering step in targets T104-105. We used the same protocol in the scorer experiment to score all the docking models that were proposed, except for target T46, where we did not include van der Waals, and target T59, where final RMSD-based filtering was applied only as scorers. Cofactors, water molecules, and solvent ions were not considered for scoring.

3.2.3 Removal of the redundant docking poses

After scoring, we applied BSAS clustering algorithm (Theodoridis and Koutroumbas 1999) with a distance cutoff of 4.0 Å to eliminate redundant predictions as previously described (Pons, Solernou, et al. 2010). In target T47, since the resulting solutions looked correct, according to a highly homologous complex structure (PDB code 2WPT), we reduced the cutoff to 0.5 Å.

3.2.4 Minimization of final models

The final ten selected poses were minimized to improve the quality of the docking models and reduce the number of interatomic clashes. In most targets of the fifth CAPRI edition we used TINKER (Ponder and Richards 1987) as previously described (Pons, Solernou, et al. 2010; Pons, Grosdidier, et al. 2010). In targets T53 and T54, we used CHARMM (50 steps conjugate gradient, 500 steps adopted-basis Newton-Raphson and 50 steps steepest descent, with the CHARMM19 force field (Brooks et al. 2009)). In T58 and all targets of the sixth CAPRI edition, we applied AMBER10 with AMBER parm99 force field (Case et al. 2005; Wang, Cieplak, and Kollman 2000). The minimization protocol consisted of a 500-cycle steepest descent minimization with harmonic restraints applied at a force constant of 25 kcal/(mol·Å-2) to all the backbone atoms, followed by another 500-cycle conjugate gradient minimization without restrains.

3.2.5 Modeling of subunits with no available structure

In several targets, the structures of the subunits were not available, and we had to model them. In most of the targets, we used MODELLER 9v6 with default parameters (Sali and Blundell 1993) based on the template/s suggested by the organizers or other homolog proteins found by BLAST (Altschul et al. 1990) search. The final selected model was that with the lowest DOPE score (Shen and Sali 2006). For targets T53 and T54 we applied POPULUS (Offman, Fitzjohn, and Bates 2006) with default template selection and model building settings. We used HHpred server (Söding, Biegert, and Lupas 2005) to model the artificial alpha-repeat eGFP A in target T96, and the missing carboxy-terminal peptide (313-329 residues) of the Ubiquitin carboxyl-terminal hydrolase L5 (UCH-L5) in targets T98 and T99. MUSTER server (Wu and Zhang 2008) was used to model the UBE2Z protein in target T103.

3.2.6 Servers experiment

For the servers experiment, we ran our pyDockWeb server (Jiménez-García, Pons, and Fernández-Recio 2013) in 15 of the evaluated targets (T59-67, T96-97, T103-105, T107). Sampling and scoring were done automatically with FTDock and pyDock, respectively. In those cases with additional experimental data available, we added distance restraints with the pyDockRST module included in the server. The best-scored server predictions were clustered with BSAS algorithm and minimized with AMBER10 as previously described.

3.2.7 Modeling of protein-peptide complexes

For the prediction of the complexes between importin-α and nuclear signaling peptides (T60–64), we applied two different strategies. On the one side, the initial peptide structures were modeled by 500-cycle minimization with generalized Born (GB) model using AMBER12 package (Case et al. 2005) and AMBER parm99 force field (Cheatham, Cieplak, and Kollman 1999) followed by 20 ns unrestrained molecular dynamics (MD), from which five representative snapshots were selected. Then, these peptide structures were used for docking simulations with our standard protocol for predictors, after which the results of the independent docking runs were merged, scored by pyDock and clustered with BSAS algorithm. On the other side, we applied a template-based approach. We first superimposed 27 peptide-bound importin-α structures and identified the residue correspondence in the peptides at both binding sites. We threaded the target sequences through the peptide sequence and identified alignments which gave good agreements with the residue propensities in the homologs. We then used this as a basis for template modeling. For each target/alignment pair, peptide fragments were joined together with averaging of the atomic coordinates of overlapping fragments, keeping the side-chain conformations where possible. The amalgamated partial models were then superimposed into position in the binding sites of all 27 homologs; missing side-chains were rebuilt with SCWRL4 (Krivov, Shapovalov, and Dunbrack 2009) and the structures minimized with CHARMM (Brooks et al. 2009). The large set of models was then scored with pyDock. For each target, we submitted the five best models generated by each of these two strategies. The server submissions were automatically built by the pyDockWeb docking server, using as input the conformations of the peptides generated by homologous templates (PDB 3UL1 and 3UKZ), followed by side-chain rebuilding with SCWRL in the context of the PDB 1EJL complex, and a subsequent 500-cycle minimization with GB model with Amber using AMBER12 package (Case et al. 2005) and AMBER parm99 force field (Cheatham, Cieplak, and Kollman 1999). For the prediction of the rest of protein-peptide complexes (T65–67), we applied an ad hoc template-based homology protocol. For T65 and T66, we rigidly docked by FTDock and then scored the DIPF binding motif of the SBB peptide, which is structurally conserved in other SSB interactions (3C94, 3Q8D, 3SXU, 3UF7, and 3UFM). For our 10 top hits, we then built the missing WMDFDD fragment by iteratively building toward the N terminus by sampling putative conformations from neighbour-dependent u and w distributions (Ting et al. 2010) and a backbone-dependent rotamer library (Shapovalov and Dunbrack 2011), selecting configurations using DFIRE (Song Liu et al. 2004). A similar protocol was undertaken for T67 after

docking of the PSY domain of Commissureless (2EZ5) and building both the N- and C- terminal flanking residue outwards from this motif. We scored the final models with pyDock. For the server submissions with pyDockWeb, we used FTDock with the peptide models obtained by template-based homology modelling (for T65–66: eight templates with the following PDB code and chain ID: 3C94_B, 3C94_C, 3Q8D_E, 3Q8D_F, 3UF7_B, 3UF7_C, 3SXU_C, and 3UFM_B; for T67: two templates with the following PDB code and chain ID: 2KQ0_B, 2KPZ_B) or five representative models from 100 ns MD trajectories. The results from all docking runs were merged, and we submitted the top 10 models as scored by pyDock.

3.3 Minimization

3.3.1 Mapping of the rigid body transformation manifold

A rigid body transformation can be described by a rotation matrix R, member of the Special Orthogonal group SO(3), and a three-dimensional translation vector t, defined in $\mathbb{R}^3$. A matrix R is a rotation matrix if and only if it is a square matrix such that $R^T = R^{-1}\ and\ \det R = 1$. The rigid body transformations can be considered the direct product of SO(3) and $\mathbb{R}^3$, SO(3) x $\mathbb{R}^3$.

As described below, we can use the exponential parametrization to locally map the nonlinear manifold onto the tangent space to the manifold, which is an Euclidean space. The exponential map relates the neighborhood of a manifold point to the tangent plane at that point.

For example, the tangent space of SO(3) at I, the identity of the group of rotations, denoted by so(3) can be identified with the space of 3x3 skew-symmetric matrices,

$$[\omega] = \begin{bmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{bmatrix} \qquad \text{Eq. 3.7}$$

with $\omega = (\omega_1, \omega_2, \omega_3)^T \in \mathbb{R}^3$.

We can now map the tangent space so(3) of skew-symmetric matrices to SO(3) with the exponential map at identity $I \in SO(3)$ defined as

$$\exp_I(\omega) = e^{[\omega]} \qquad \text{Eq. 3.8}$$

where the right-hand side of the equation can be rewritten as the Rodrigues formula

$$e^{[\omega]} = I + \frac{sin(\|\omega\|)}{\|\omega\|}[\omega] + \frac{(1 - cos(\|\omega\|))}{\|\omega\|^2}[\omega]^2 \quad \text{Eq. 3.9}$$

where $\|\omega\|$ is the Euclidean norm of ω.

We can easily derive the exponential map of $SO(3) \times \mathbb{R}^3$from that of $SO(3)$. The tangent space of the product group $SO(3) \times \mathbb{R}^3$ at the identity (I, $\mathbf{0}$) can be identified with the set of points $(\omega, v) \in \mathbb{R}^6$, such that

$$\exp_{(I,0)}(\omega, v) = (e^{[\omega]}, v) \quad \text{Eq. 3.10}$$

Therefore,

$$\exp_{(I,0)}: \mathbb{R}^6 \to SO(3) \times \mathbb{R}^3 \quad \text{Eq. 3.11}$$

defines a local mapping for $SO(3) \times \mathbb{R}^3$ in the neighborhood of (I, $\mathbf{0}$).

In practice, this means that we can perform the minimization at the tangent space of $SO(3) \times \mathbb{R}^3$, which is a Euclidean space $\mathbb{R}^6$, and map the minimized solution in $\mathbb{R}^6$ back to a rotation-translation pair (R, t) in $SO(3) \times \mathbb{R}^3$ (see Figure 3.1).

Thus, we reduce the total number of parameters to optimize to six, three for the rotation vector and three for the translation vector, dramatically decreasing the resources and time required to perform the minimization.

We performed the minimization with the Powell optimization algorithm (Powell 1964) implemented in the Python package Scipy (E. Jones, Oliphant, and Peterson 2015). Powell minimization is a conjugate direction method that does not require the optimization function to be differentiable.

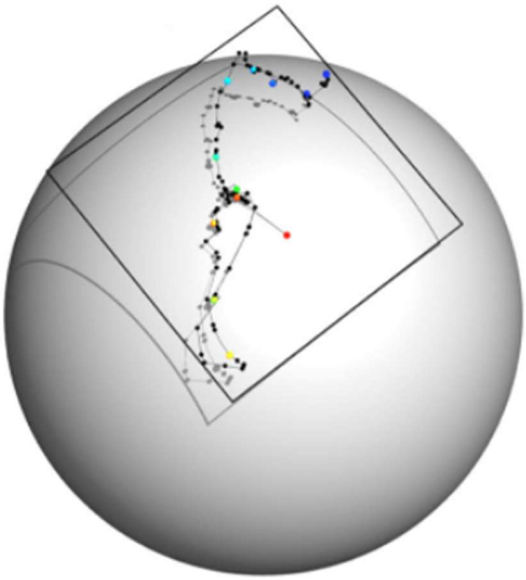

Figure 3.1. Plot of a minimization run. The sphere represents the rigid body transformation space $SO(3) \times \mathbb{R}^3$ and the plane the tangent space at the identity. The minimization takes place in the tangent space. Each optimization step is represented as a dot in the plane. Color dots are plotted every ten steps. Colors represent the energy values at the minimization steps. Red/blue dots are associated with high/low energies. The shadow dots in the sphere represent the minimization steps mapped onto $SO(3) \times \mathbb{R}^3$ with the exponential map (From Mirzaei et al. 2012).

3.4 Analysis of binding affinity changes upon mutation

3.4.1 *In-silico* alanine scanning mutagenesis with AMBER

The method requires the structure of the complex that is being studied and combines molecular dynamics simulations and binding energy calculations with the MM-GBSA method (Miller et al. 2012). Before running the molecular dynamics simulations, the system was solvated, minimized and equilibrated as follows. First, original PDB coordinates were stripped of hydrogen atoms, monovalent ions, and all water molecules. Missing side-chain atoms and hydrogen atoms were added from AMBER residue libraries using the LEAP AMBER tool. The resulting complex structure was then immersed in a periodic truncated octahedron box containing a 12 Å buffer of TIP3P water molecules, and Na+ and Cl- counterions were added to the solvent bulk to maintain the neutrality of the system and reach 50 mM NaCl ionic strength. Each solvated system underwent a short solvent minimization and five-step equilibration protocol. First, a 500-cycle steepest descent and a 500-cycle conjugate gradient minimization were performed, applying harmonic restraints with force constant of 50 kcal/(mol·Å^2) to all protein atoms in order to minimize the solvent molecules. Then, the five-step equilibration

was performed by applying periodic boundary conditions and computing long-range electrostatics by the particle-mesh Ewald method. At each stage, the integration time step was set to 2 fs, the system pressure to 1 atm, and the nonbonding cutoff distance to 12 Å. The five steps were:

1) A 40 ps MD simulation was run applying harmonic restraints to all the protein atoms with force constant of 25 kcal/mol·Å^2, raising the temperature to 300 K by Langevin dynamics approach with a collision frequency of 1 fs.
2) A 20 ps step was performed, setting the temperature to 300 K and reducing system restraints to 10 kcal/mol·Å^2.
3) Another 20 ps simulation was run with 10 kcal/mol·Å^2 restraints only to the protein backbone atoms.
4) A further 20 ps simulation was performed, decreasing protein backbone restraints to 5 kcal/ mol·Å^2.
5) A final 100 ps unrestrained MD simulation was run without any restraint.

Finally, a 5 ns MD simulation was performed in an isothermal-isobaric ensemble, setting pressure to 1 atm and temperature to 300 K.

The AMBER MMPBSA.py script was used to perform Computational Alanine scanning calculation on 200 snapshots extracted from the last 2ns of the 5-ns-long MD trajectory. All the MEK1 and BRAF residues were mutated to alanine, and then the binding free energy change ($\Delta\Delta G$) was estimated as the difference between the binding ΔG (MM-GBSA method) of the wild-type and that of the mutated complex. Give its high computational cost, the contribution of conformational entropy was not included, but that should not significantly affect the comparison of mutant and the wild-type free energies.

3.4.2 Benchmarking predictive performance

We checked the accuracy of the binding affinity changes upon mutation predictions with the data stored in SKEMPI v1.1, the Structural database of Kinetics and Energetics of Mutant Protein Interactions (Moal and Fernández-Recio 2012b). SKEMPI is among the largest databases with experimentally measured changes in binding free energy upon mutation. SKEMPI contains 3047 $\Delta\Delta G$ measurements of 2792 unique mutations or sets of mutations found for 158 structures of 85 protein-protein complexes. All its data has been manually curated and come from multiple studies reported in the literature. Many different types of mutations have been collected. Given that a large number of mutations come from alanine scanning experiments, a third of the data correspond to mutations to alanine.

3.5 Structural and energy determinants in protein-RNA docking

3.5.1 Protein-RNA docking protocol

The general docking protocol described in section 3.1 was followed. However, some modifications were necessary to adapt the protocol to RNA molecules. In the sampling phase, FTDock 2.0 (Gabb, Jackson, and Sternberg 1997) was run without electrostatics and a grid size resolution of 1.2 Å. A total number of 10,000 docking poses were generated for each benchmark case.

Rigid-body orientations were scored with pyDock scoring function, defined by equations Eq. 3.1-Eq. 3.6 and described in section 3.1. Atomic desolvation parameters (ADPs) were adapted for protein-RNA complexes. We used the same ADPs as for protein-protein complexes, assigning the ADP of carboxylic O to RNA phosphate O atoms, and null ADP to RNA P atoms, to minimize the errors when computing their contribution (Pons, Solernou, et al. 2010). Parameters like atomic charges (q), equilibrium radii (r) and well depth (e) were taken from AMBER 94 force field (Cornell et al. 1995). Electrostatics and van der Waals terms were truncated as explained in section 3.1 to avoid excessive noise coming from docking rigid-body surfaces.

3.5.2 Statistical potentials for scoring

We also scored the protein-RNA rigid-body docking orientations with the pairwise protein-RNA statistical potentials previously developed within our group (Pérez-Cano et al. 2009) (see Figure 3.2). For every residue-ribonucleotide pq pair at the interface of each docking pose i (considering as pairs those that have at least one atom within 4 Å distance from each other), the corresponding propensity value according to its type was assigned. The propensity-based values of all pairs were summed up to compute the final score of the given docking pose i, as shown in Eq. 3.12. Finally, all docking solutions were ranked according to these propensity-based scores.

$$\Delta G_i^{stat} = \sum_{pq} \Delta G_{pq}^{stat} \qquad \text{Eq. 3.12}$$

For the sake of comparison, success rates of pairwise protein-protein statistical potentials in the scoring of bound protein-protein rigid-body docking poses were extracted from the previous work of Pons et al., 2011 (Pons et al. 2011).

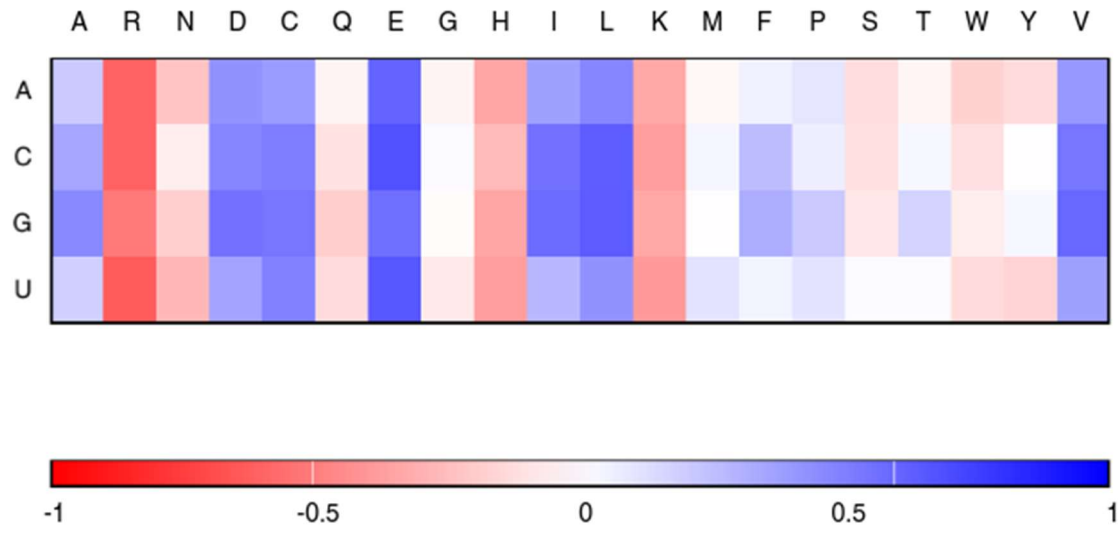

Figure 3.2. Pairwise protein-RNA statistical potentials (favorable pairs are in red; disfavored in blue). From Pérez-Cano et al. 2009.

3.5.3 Explicit consideration of shape complementarity for scoring

As previously reported, considering shape complementarity when scoring protein-RNA docking poses can improve the predictive rates (Pérez-Cano et al. 2009). Here we scored the rigid-body docking orientations by using the FTDock 2.0 SCscore, which is a correlation score describing the shape complementarity of the components represented as a discrete 3D grid (Katchalski-Katzir et al. 1992). We then aimed to combine the SCscore with the electrostatics and van der Waals energy terms. Since the variation of the distribution of values for these three terms ranged within the same order, we decided not to use any weighting factor to avoid potential over-fitting problems. Therefore, the three terms were combined in a single scoring function just by adding the scores, after changing the sign of the SCscore values to be consistent with the other terms (see Eq. 3.13).

$$E_{comb} = E_{ele} + E_{vdw} - SCscore \qquad \text{Eq. 3.13}$$

3.5.4 Benchmarking and evaluation of scoring results

The performances of the different scoring functions analyzed here for bound protein-RNA docking were evaluated on the bound structures of the 106 cases of the protein-RNA docking benchmark v1.0 (Pérez-Cano, Jiménez-García, and Fernández-Recio 2012). The benchmark set is composed of experimental protein-RNA complex structures, for which the unbound proteins had available structures (or could be homology-based modeled). It includes 5 unbound-unbound, 4

unbound-pseudo-unbound, 62 unbound-bound, 5 unbound-model, 8 model-unbound, 19 model-bound and 3 model-model cases. The benchmark covers all major functional categories and contains cases with different degrees of difficulty for docking, as far as protein and RNA flexibility is concerned. The benchmark set and more detailed information are available online at https://life.bsc.es/pid/protein-rna-benchmark/.

On the other hand, the unbound protein-RNA docking was tested on a subset of cases from the same benchmark with unbound or modeled structures available for both protein and RNA components. This ‘‘unbound” data set comprises a mixture of 5 unbound-unbound, 4 unbound-pseudo-unbound, 5 unbound-model, 8 model-unbound and 3 model-model cases. The total of 25 ‘‘unbound” cases were grouped in 6 easy, 13 medium and 6 difficult cases, defined according to the unbound-to-bound conformational deformation of their interfaces as those cases with protein and RNA average IntRMSD below 2.5 Å, between 2.5 and 5.0 Å, and above 5.0 Å, respectively (Pérez-Cano, Jiménez-García, and Fernández-Recio 2012). For comparison, protein-protein docking calculations were evaluated on the protein-protein docking benchmark 3.0 (Hwang et al. 2008).

The assessment of the docking solutions was done by calculating the ligand root-mean-square deviation (LigRMSD) with respect to the reference complex structure, defined as the RMSD between the P atoms of the docking and reference RNA molecules, after superimposing the Cα atoms of the docking and reference protein molecules. In oligomeric proteins, in which additional molecules of RNA could bind in symmetric sites, we calculated the LigRMSD with respect to all the different symmetric positions and chose the smallest value. The docking solutions were classified as excellent-accuracy (LigRMSD $\leq$ 1 Å), high-accuracy (LigRMSD $\leq$ 2 Å), medium-accuracy (LigRMSD $\leq$ 5 Å), acceptable (LigRMSD $\leq$ 10 Å), or incorrect (LigRMSD $>$ 10 Å). Then, the predictive success rates for a given scoring function were computed as the percentage of cases, among those within the corresponding benchmark set, in which an acceptable (or any other better accuracy criteria) docking solution was found within the top N docking models as sorted by such scoring function. For comparison, we computed the success rates expected by random, based on aleatory permutation of the rank of the generated docking solutions for each case.

4 Results

4.1 Rigid-body docking: assessment of the state-of-the-art and current limitations in CAPRI

We have participated in CAPRI editions 5th and 6th, where we submitted predictions for all the proposed targets. The assessment of our submissions by the CAPRI organization is summarized in Table 4.1 and Table 4.2. In the fifth CAPRI edition, we submitted acceptable models (or better) for 6 out of 9 targets as predictors and 4 out 7 as scorers. In the sixth CAPRI edition, excluding the CASP-CAPRI cases, we submitted acceptable models (or better) for 7 out of 18 targets as predictors, and 4 out of 11 targets as scorers. Additionally, in the 25 cases of the CASP-CAPRI experiment we submitted acceptable models or better for 11 out of 25 targets as predictors, and 14 out of 25 targets as scorers. Hereinafter, we describe in detail our submissions for each of the targets

		Predictors			Scorers		
Target	Type	Submission rank[a]	Quality[b]	Successful groups[c]	Submission rank[a]	Quality[b]	Successful groups[c]
T46	HH	-	-	2 (40)	-	-	8 (16)
T47	HU	1	***	25 (29)	2[d]	***	13 (14)
T48	UU	3	*	14 (32)	No scorers	No scorers	No scorers
T49	UU	4	*	14 (33)	6	*	7 (13)
T50	UH	1	**	18 (40)	4	**	12 (17)
T51	DHD	-	-	3 (46)	-	-	5 (13)
T53	UH	3[e]	**	20 (42)	1	**	11 (13)
T54	UH	-	-	4 (41)	-	-	0 (13)
T58	UU	5	**	11 (23)	No scorers	No scorers	No scorers

Table 4.1. Results of our pyDock protocol for all protein–protein targets of fifth CAPRI edition.

U, unbound; H, homology-based model; D, domain.

[a] Rank of the best model within our submission to CAPRI.

[b] Quality of our best model according to CAPRI criteria.

[c] Number of successful groups for each target; in brackets, the total number of participants.

[d] Model Rank 1 had medium accuracy (**).

[e] Model Rank 1 had acceptable accuracy (*).

Target[a]	Type[b]	Predictors[c]	Servers[c]	Scorers[c]
T59	Prot/Prot (U/H)	0	0	0
T60-64 (major binding site)	Prot/Pep (U/H)	**M02** (T60)** **M03** (T61)**, M04* (T61) **M03** (T62)**, M04* (T62) **M03** (T63)**, M04* (T63) **M03** (T64)**, M04* (T64)	**M08* (T63)**	N/A
T60-64 (minor binding site)	Prot/Pep (U/H)	**M01* (T63)**	0	N/A
T60-64 (minor binding site, six-residue)	Prot/Pep (U/H)	**M01** (T60)** **M01** (T61)** **M01** (T62)** **M01** (T63)** **M01** (T64)**	**M03* (T61)** **M02** (T62)**, M08** (T62) **M06* (T62)**, M10* (T62) **M01* (T63)**, M07* (T63)	N/A
T65	Prot/Pep (U/H)	0	0	N/A
T66	Prot/Pep (U/H)	**M01* (EF)**	**M04* (EF)**, M08* (EF)	N/A
T67	Prot/Pep (U/H)	**M10* (all)**	**M06* (PPxY)**, M07*	N/A
T95	Prot-DNA/Prot (U/U)	0	X	N/A
T96	Prot/Prot (H/H)	0	0	0
T97	Prot/Prot (H/H)	0	**M10***	M05*, M09*, M10*
T98	Prot/Prot (U/U)	0	X	0
T99	Prot/Prot (U/U)	0	X	0
T100	Prot/Prot (U/H)	0	X	0
T101	Prot/Prot (U/H)	0	X	0
T103	Prot/Prot (H/H)	0	0	**M03* (Ct)**, M05* (Ct)
T104	Prot/Prot (H/H)	M03*, **M06****, M07*, M10**	0	**M01**-02***, M03**, M04***, M05**, M06***, M07-08**, M09***, M10**
T105	Prot/Prot (H/H)	**M02****, M10**	M02*, **M06****	**M01****, M02*, M03-05**, M07-10**
T107	Prot/Prot (U/U)	0	0	0

Table 4.2. Results of our pyDock protocol for all protein–protein targets of sixth CAPRI edition.
[a] Underscored: target of special difficulty, with only three or fewer groups that submitted correct models.
[b] B: bound; U: unbound; H: homology-based model.
[c] Correct models submitted to CAPRI by our group. Each model is numbered according to its rank within our submission. The quality of each model is indicated, following CAPRI criteria: * acceptable; ** medium quality; * high quality. In bold, our best-quality model for each target. "0": no correct model submitted. "X": no submissions. "N/A": experiment not available (i.e., the target was not proposed for the scorers experiment).**

4.1.1 Fifth CAPRI

4.1.1.1 Successful predictions

4.1.1.1.1 Target T47 (model/pseudo-unbound)

Target T47 consisted in the structural prediction of the complex between the DNase domain of colicin E2 and the immunity protein Im2. The colicin E2 was modeled based on the structure of colicin E9 (85% sequence identity) in complex with Im9 immunity protein (PDB 1EMV (Kühlmann et al. 2000)). The coordinates of the immunity protein Im2 were extracted from its structure in complex with colicin E9 (PDB 2WPT (Meenan et al. 2010)). Even though the binding mode for target T47 was easy to derive by template-based docking from the homologous colicin E9/Im2 complex structure (PDB 2WPT), we performed a template-free docking to assess our standard docking protocol. We applied distance restraints selecting those docking poses in which two key residues, Im2 Y54 and colicin E2 F85 (equivalent to colicin E9 F86 in 2WPT), were within 6 Å, the default restraint distance of pyDockRST module (T. M. K. Cheng, Blundell, and Fernandez-Recio 2008). We submitted five correct models (one high accuracy, one medium accuracy, and three acceptable). In the scoring experiment, we evaluated 1051 models with pyDock scoring function and applied the same distance filter that we used in predictors. All the models we submitted were successful (four high accuracy and six medium accuracy).

4.1.1.1.2 T48 (unbound/unbound)

Target T48 was the structural prediction of the complex between the diiron-hydroxylase toluene 4-monooxygenase and the Rieske-type ferredoxin T4moC protein (PDB 1VM9 (Moe et al. 2006)). As the organizers suggested, we built the heterohexameric biological unit of the diiron-hydroxylase by applying crystal symmetry operations to its trimeric structure in complex with the T4moD effector protein (PDB 3DHH (Bailey et al. 2008)). After running the standard pyDock protocol we selected those docking poses that had any of the diiron-hydroxylase

Fe^{2+} and ferredoxin S_2Fe_2 atoms within 23 Å distance to allow for the electron transfer between these groups (Bailey et al. 2008) (expected distance of 16 Å in 3DHH plus an arbitrary margin of 7 Å to allow the inclusion of some low-energy docking solutions). We submitted three models of acceptable quality.

4.1.1.1.3 T49 (unbound/unbound)

Target T49 was the same complex as T48 but with a different hexameric conformation for diiron-hydroxylase toluene 4-monooxygenase. We ran the same protocol as for target T48 and submitted four acceptable quality models. For the scoring experiment, we also applied the same protocol based on pyDock scoring and electron transfer distance filtering and submitted one acceptable model.

4.1.1.1.4 T50 (unbound/model)

Target T50 consisted in the structural prediction of the complex between the 1918 H1N1 influenza virus hemagglutinin and the HB36.3 de novo designed protein. The coordinates of the hemagglutinin were taken from its structure in complex with an antibody (PDB 3GBN (Ekiert et al. 2009)) and the biological hexamer was rebuilt by applying symmetry operations. We modeled the HB36.3 based on the crystal structure of the homologous (83% sequence identity) protein APC36109 from Bacillus stearothermophilus (PDB 1U84), using the target-template protein alignment offered by the organizers. Our submission as predictors contained nine successful models (five acceptable and four medium-quality). For the scoring experiment, we evaluated the 1451 models with the same protocol as in predictors. We found five acceptable and one medium-quality model.

4.1.1.1.5 T53 (unbound/model)

Target T53 was a complex between two artificial alpha helicoidal repeat proteins (alpha-Rep), alpha-rep4 (PDB 3LTJ (Urvoas et al. 2010)) and alpha-rep2, both designed on the basis of thermostable HEAT-like repeats. The ligand alpha-rep2 was built using as template alpha-rep4 (PDB 3LTJ), with 77% sequence identity. We generated the docking poses with ZDock, FTDock, and SwarmDock and evaluated them with pyDock. We submitted four successful predictions (three acceptable and one medium-quality model). For the scoring experiment, we applied the same protocol as in predictors. We found three acceptable and one medium-quality model.

4.1.1.1.6 Target T58 (unbound/unbound)

This target was a complex between the unbound G-Type Lysozyme (PDB 3MGW (Kyomuhendo et al. 2010)) and the unbound Escherichia coli Plig lysozyme

inhibitor (PDB 4DY3 (Leysen et al. 2012)). There was available small-angle X-ray scattering (SAXS) data for this complex, which we used for scoring with our module pyDockSAXS, previously developed to combine pyDock scoring and fitting to SAXS data (Pons, D'Abramo, et al. 2010). Also, there was some available information indicating a central role of the G-type lysozyme E73, D86, and D97 residues and the E. coli Plig lysozyme inhibitor R119 and Y47 residues (Helland et al. 2009). Based on these residues, we imposed ambiguous distance restraints with our module pyDockRST (Chelliah, Blundell, and Fernández-Recio 2006). We submitted one medium-accuracy and two acceptable models.

4.1.1.2 Unsuccessful predictions

In three of the protein-protein cases (T46, T51, and T54) we were not able to submit any correct model, either as predictors or as scorers. These cases seemed to be highly challenging for the majority of participants since in all of them there were no more than three successful groups as predictors or as scorers or both (see Table 4.1). In target T46 (model/model), the interacting subunits Mtq2 and Trm112 were modeled based on the homolog templates with low sequence identity (Mtq2 was based on a template with PDB code 1T43, 28% sequence identity; Trm112 was based on a template with PDB code 2J6A, 36% sequence identity). The inaccuracies in the modeling added too much error, and the docking was not successful. Target T51 (bound/model/unbound) was a complicated case of a multidomain protein, with interactions between GH5-CBM6/CBM13/Fn3 domains. This could be divided into two different docking cases both involving CBM13 domain, which needed to be modeled based on a template with PDB code 1KNL (38% sequence identity). Again, a model based on a template with that level of homology can deteriorate docking results. Target 54 (unbound/model) was in principle straightforward, involving the modeling of Rep16 based on the template with PDB code 3LTJ (88% sequence identity), but the submitted solutions were incorrect for us as well as for the majority of participants. Indeed, despite the scoring set contained several acceptable models, no group was able to identify them (see Table 4.1).

4.1.2 Sixth CAPRI edition

4.1.2.1 Successful predictions

4.1.2.1.1 Targets T60-64 (unbound/peptide models)

This set of targets consisted in the interaction between mouse importin-α and five different nuclear signaling peptides: Gu-α (T60), a28 (T61), a58 (T62), b6 (T63), and b141 (T64). These five complexes were evaluated as three targets: (i) the

major binding site; (ii) the minor binding site; and (iii) the six central peptide residues in the minor binding site. Coordinates of importin-α were taken from PDB 1EJL, bound to a large T antigen seven-residue peptide. The target peptides were modeled using a dual strategy, based on *ab initio* molecular dynamics or homologous templates (see section 3.2.7). For each submission, five of the protein–peptide models were built by template-based homology modeling, and the other five by docking, using as input structures the models generated by molecular dynamics. Overall, we obtained excellent results. More in detail, as predictors we submitted one high- and four medium-accuracy models for the major binding site in the five peptide complexes; one acceptable model for the complete minor binding site of one of the peptide complexes (T63); and five medium models for the six-residue minor binding site of the five peptide complexes. Interestingly, all the correct protein–peptide models submitted as predictors were directly built based on homologous templates, which shows that the use of unrestrained molecular dynamics to build the conformations of the peptides did not yield suitable input structures for docking.

4.1.2.1.2 Targets T65-66 (unbound/peptide models)

These two targets consisted in the Ct peptide of ssDNA binding protein (SBB-Ct) in complex with RNaseH (T65) or DNA helicase (T66). The structure of RNaseH was available as unbound (PDB 2RN2), while the coordinates of the DNA helicase were provided as unbound by the CAPRI organizers (now available as PDB 4NL4). The structure of the peptide was not available. As predictors, we applied an ad hoc template-based homology modeling procedure (see section 3.2.7). We submitted one acceptable model as predictors. However, for the T65 complex (PDB 4Z0U) we were not able to submit any correct model, either as predictors or as scorers. This case was highly difficult for the majority of participants, as there was only one successful group out of > 40 participants. Indeed, RNaseH binding to Ct peptide involved a large conformational change: RNaseH interface atoms, defined as those within 5 Å distance from Ct peptide in the complex structure, showed 4.2 Å RMSD between the unbound and bound structures.

4.1.2.1.3 Target 67 (unbound/peptide model)

Target T67 was the interaction between Nedd4 WW3 domain and the PPxY motif of ARRDC3. The organizers provided the unbound structure for the protein, and the peptide structure was modeled. As predictors, we applied an ad hoc template-based homology modeling procedure (see section 3.2.7). We submitted acceptable models as predictors.

4.1.2.1.4 Targets T96-97 (model/model)

These targets consisted in the interaction between eGFP and the artificial a-repeat eGFP-binder A (T96) or C (T97). We modeled eGFP based on a FRET-optimized cerulean fluorescent protein (PDB 4EN1, 92% sequence identity). The structures of eGFP-binder A and C were modeled based on a homologous template (PDB 3LTJ) with 82% and 74% sequence identity, respectively. We submitted acceptable models for T97 as servers and as scorers, while we failed to submit any correct model for T96. The main reason for the different performance of these two targets could be related to the larger deviation of the modeled a-repeat eGFP-binder A protein with respect to the bound structure in T96 as compared to that of the a-repeat eGFP-binder C protein in T97 (BB-IntRMSD 5 Å and 2 Å, respectively). In the case of scorers, we also obtained better results for T97.

4.1.2.1.5 Target 103 (model/model)

Target T103 consisted in the Ube2Z protein in complex with Fat10. The structures of the Ube2Z and Fat10 proteins were modeled based on homologous templates (PDB 3CEG and 3U30, with 43% and 32% sequence identity, respectively). We submitted acceptable models only as scorers.

4.1.2.1.6 Targets 104-105 (model/model)

In these two targets, in addition to providing near-native models, the organizers asked to predict the interfacial water positions. Target T104 consisted in the interaction between pyoAP41 and ImAP41 proteins. As none of these structures were available, they were modeled based on homologous templates, colicin E9 (48% sequence identity) and Im9 immunity protein (46% sequence identity), respectively. Both template structures were extracted from the PDB 1BXI. Target T105 consisted in the interaction between pyoS2 and ImS2 proteins, whose structures were modeled based on colicin E2 DNase (52% sequence identity) and Im2 immunity protein (59% sequence identity), as found in PDB 3U43 chains B and A, respectively. We applied distance restraints after pyDock protocol by selecting those docking poses in which two key contacting residues, pyoAP41 Y81 and ImAP41 F59 (equivalent to colicin E9 F86 and Im9 Y54), or pyoS2 Y85 and ImS2 Y55 (equivalent to colicin E2 F86 and Im2 Y54), were within an arbitrary distance of 6 Å. We submitted acceptable (or better quality) predictions for complex structure and water positions in the two targets, both as predictors and as scorers. There was a clear correlation between the quality of our predictions for the complex structure and that of the interfacial water positions.

4.1.2.2 Unsuccessful predictions

In the following cases, we were not able to submit any correct model, either as predictors or as scorers. These cases seemed to be difficult for the majority of the participant groups.

4.1.2.2.1 T59 (unbound/model)

Target T59 consisted in the interaction between the LSm domain of Edc3 protein and the ribosomal protein Rps28b. The NMR structure of Edc3 was available both as unbound (PDB 4A53) and in complex with a short helical leucine-rich motif (HLM) from Dcp2m mRNA Decapping Complex (PDB 4A54). The structure of the Rps28b was not available and had to be modeled. We decided to use the cryo-EM structure PDB 3IZB (superseded by 4V6I) (SI 85%).

4.1.2.2.2 T95 (unbound/unbound)

Target T95 consisted in the interaction between PRC1 ubiquitylation module and the nucleosome core particle, whose coordinates were available as unbound structures (PDB 3RPG and 3LZ0, respectively). Once the complex structure was released (PDB 4R8P), we found that the molecules did not show large conformational changes upon binding (RMSD of < 1 Å to the bound conformation calculated on all the Cα atoms and roughly 1.5 Å on all the DNA atoms with respect to the complex structure). Nevertheless, this was a challenging case in which only three participants submitted acceptable models. Most likely, the presence of DNA made docking and scoring extremely difficult.

4.1.2.2.3 T98-101 (unbound/model)

Targets T98–101 consisted in the interaction between the Ubiquitin Carboxyl-terminal hydrolases L5 or L5Ub (with ubiquitin covalently bound) proteins and either RPN13 activator or Ino80G inhibitor. The structures for these complexes were later released with the following PDB codes: 4UEM, 4UEL, 4UF6, and 4UF5, respectively. These cases were highly difficult for all participants, as there was not found a single acceptable model among all the participants. The main challenges in these cases were the large conformational changes of both the interacting proteins upon binding, the inaccuracy in the modeling of Ubiquitin carboxyl-terminal hydrolase L5 interface, as well as the small interface area between the docking partners.

4.1.2.2.4 T107 (unbound/unbound)

Target T107 consisted in the interaction between the hemopexin binding protein and hemopexin. This case was highly challenging for all the groups, as there was not a single acceptable model among all participants. The main reason for the

difficulty of this target lies on the large conformational changes of the hemopexin binding protein upon hemopexin binding, especially involving a large loop (residues 707–730) located within the complex interface (unbound-to-bound Cα-RMSD 16.2 Å). Another potential reason for the target difficulty could be the large size of the hemopexin binding protein, composed of around 800 amino acid residues, for which our methodology cannot provide enough sampling, as it was found during the last CAPRI edition (Pallara et al. 2013).

4.1.2.3 CASP-CAPRI experiment

CAPRI round 30, the first joint CASP–CAPRI experiment, consisted in 25 targets of homo- and hetero-oligomers from the CASP11 2014 round (targets T68–94, excluding T76 and T86, which were canceled). We submitted at least one acceptable model in 11 out of the 12 easy homodimer targets, either as predictors or as scorers. Also, as scorers, we successfully predicted two out of the six difficult homodimer targets, and one out of the two hetero-complex targets. On the contrary, we did not submit any successful model for any of the five tetrameric targets, where the inaccuracy of the homology-built subunit models and the smaller pair-wise interfaces severely limited the ability to derive the correct assembly mode. Globally, pyDock predictions were placed among the top 10 ranked groups out of about 25 predictors, and among the top 5 ranked groups out of about 12 scorers participating in this experiment.

4.2 pyDockLite: A modified pyDock scoring function for fast exploration of docking energy landscapes

As shown in the previous section, the docking method pyDock, developed in our laboratory, has been successfully applied to a large variety of cases in the assessment experiment CAPRI, and it has been found as one of the most competitive protein-protein docking protocols worldwide. Under the rigid-body approximation, it has reached a good balance between computational requirements and performance. However, its scoring function is too heavy to tackle computationally demanding problems such as those that involve explicit modeling of flexibility. Here, we have developed the new pyDockLite scoring function, with the goal of making pyDock more efficient without losing too much performance.

4.2.1 pyDockLite description and performance in protein-protein benchmark 4

4.2.1.1 pyDockLite desolvation term

As described in section 3.1.2, in order to compute pyDock desolvation term we need to calculate the BSAs based on the differences between the total ASA of the complex or docking model and that of the unbound components (see equations Eq. 3.3 and Eq. 3.5). The computation of ASA is performed with Naccess, and it is quite demanding. For a set of generated rigid-body docking poses, the ASA of the unbound components can be computed only once, but the ASA of the different docking models need to be computed each time, constituting the most demanding part of the pyDock evaluation of rigid-body docking solutions. Therefore, we explored whether the BSA between two atoms could be somehow estimated as a function of their unbound ASA values and their atomic distance.

Figure 4.1 plots the BSA computed with Naccess, using a water sphere of radius 1.4 Å, for a toy system composed of two carbon atoms of 1.8 Å radius. The x-axis represents the distance between the centers of mass of the atoms, and the y-axis the computed BSA. When the distance between the centers of mass is equal or greater than 6.4 Å, the value of BSA is zero, i.e., the total ASA of the atoms remain constant. That distance, 6.4 Å, is the diameter of the atom carbon, 3.6 Å, plus the diameter of a water molecule used in Naccess calculations, 2.8 Å. Below 6.4 Å, BSA linearly increases when the distance decreases. Incidentally, a BSA value is calculated for the purposes of this test even when the atoms are clashing, i.e., when their distance is below 3.6 Å. Interestingly, Figure 4.1 suggests that, for a system formed by two atoms, we can estimate the BSA value from the distance between the carbon atoms.

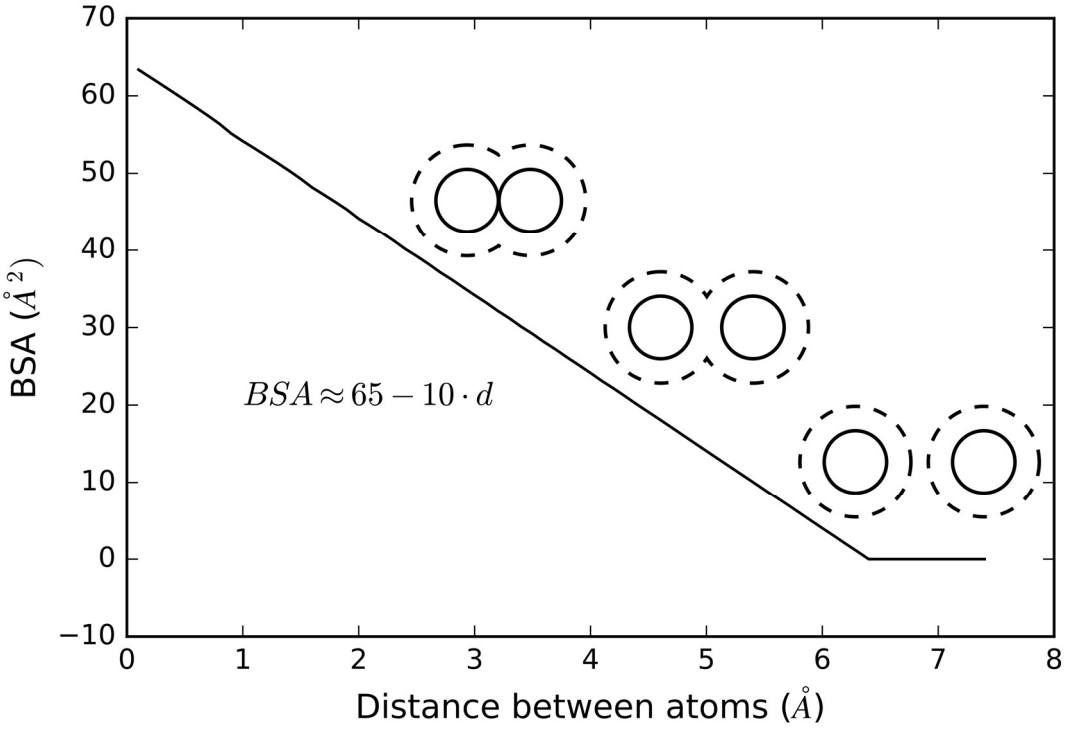

Figure 4.1. BSA as a function of the distance between two carbon atoms of 1.8 Å radius. BSA values have been computed with the software Naccess.

For a system formed by many atoms, e.g., a docking pose, the situation is more complicated. For simplicity, we could extrapolate the above-mentioned observation to all the atoms of a molecule and express the BSA of a given atom as a function of the distance with its closest atom of the other subunit. Thus, we could get an estimated BSA value with the following procedure:

- Compute with Naccess the ASA for all the atoms in the unbound subunits: $ASA_i^{unbound}$
- In the docked conformation, for each atom i, compute BSA_i according to Eq. 4.1, where $d_{closest}$ is the distance to the nearest atom of the complementary subunit.
- Check that BSA is not greater than ASA in the unbound conformation, i.e., it is not possible to bury more surface area than the initially available (Eq. 4.2).

$$BSA_i(d_{closest}) = \begin{cases} 64 - 10 \cdot d_{closest}, & if\ 0 < d_{closest} \leq 6.4 \\ 0, & otherwise \end{cases} \quad \text{Eq. 4.1}$$

$$if\ BSA_i > ASA_i^{unbound} \rightarrow BSA_i = ASA_i^{unbound} \quad \text{Eq. 4.2}$$

Two big approximations are assumed in this approach. First, the BSA of a given atom is only attributed to the nearest atom of the complementary subunit (see Figure 4.2-A). Most of the times, this may not be the case, since other atoms could also contribute to the BSA (even though to a lesser extent than the closest atom). A possible variant would be to include in Eq. 4.1 all neighbor atoms of the complementary subunit, but in that case, the error could be even more significant. Unlike other magnitudes BSA does not comply with the additive property, i.e., the combined BSA due to two given atoms on a third one is not always equal to the sum of the individually induced BSAs (see Figure 4.2-B). The second approximation is considering a single atom type, in this case, carbon, in our model to approximate Naccess computed BSA. We made this decision to simplify the model and reduce the computation times as much as possible. In any case, BSA computation is a prerequisite to estimating the desolvation energy term, which is always calculated using the corresponding atomic solvation parameters for each atom type (see equations Eq. 3.3 and Eq. 3.5). Figure 4.3 shows that despite the approximations assumed by the model, the correlation with pyDock ASA-based desolvation energy is high.

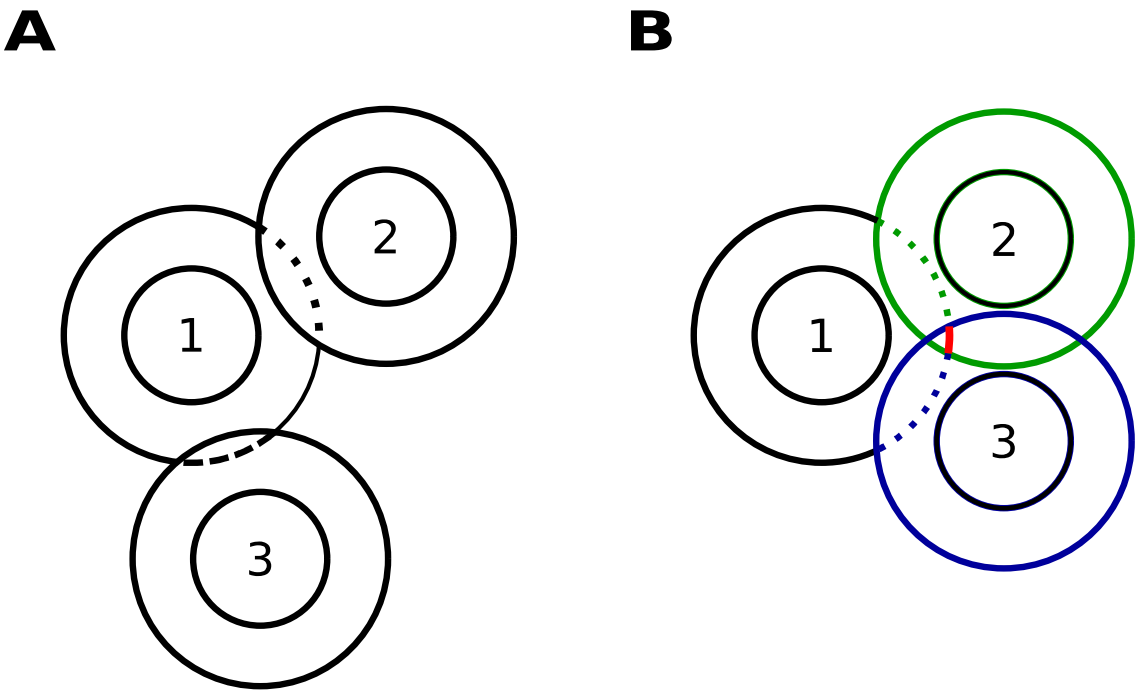

Figure 4.2. Schematic view of BSA computing approximations. A) pyDockLite solvation implementation only considers the BSA of the closest atom. In the case of atom 1, it will only take into account the effect of atom 2 (dotted line) but will ignore the effect of atom 3 (dashed line). B) BSA is not additive: The BSAs due to atoms 2 (green) and 3 (blue) over atom 1 superimposed in a particular region (red), which will be computed only once.

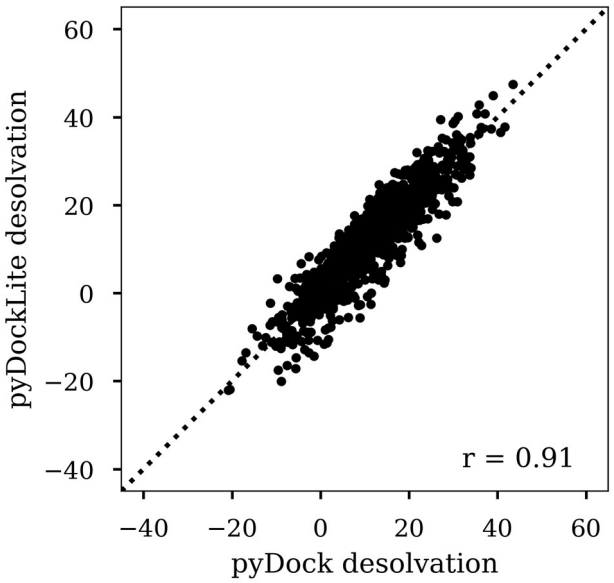

Figure 4.3. Correlation plot between pyDock desolvation (x-axis) and pyDockLite desolvation (y-axis) for the 10,000 docking poses generated by FTDock to model trypsin/CMTI-1 squash inhibitor (PDB 1PPE) from the unbound structures of Trypsin (PDB code 1BTP) and CMTI-1 squash inhibitor (PDB code 1LU0). Only 1000 random docking poses are shown for the sake of clarity, but the Pearson correlation coefficient is computed over the 10,000 evaluated poses.

4.2.1.2 pyDockLite electrostatic term

We adopted a basic truncation technique to reduce the computational cost associated with the calculation of electrostatics and van der Waals energies (Eq. 3.2 and Eq. 3.6), by defining a distance cutoff beyond which the energy value is set to 0. Figure 4.4 shows the correlation between the standard and the truncated electrostatic energies for different cutoffs. We obtained Pearson correlation coefficients equal to 0.86, 0.99 and 1.00 for cutoff distances of 10, 30 and 50 Å, respectively. From these results, a cutoff of 30 Å seems to define a good balance between accuracy and speed.

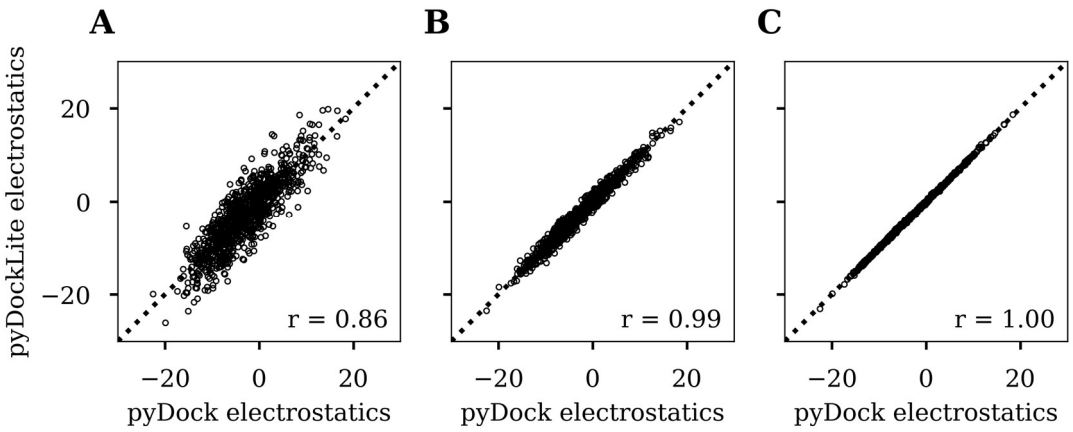

Figure 4.4. Correlation plot between pyDock (x-axis) and pyDockLite (y-axis) electrostatics for cutoff distances of 10 Å (A), 30 Å (B) and 50 Å (C). Data corresponding to 10,000 docking poses generated by FTDock to model trypsin/CMTI-1 squash inhibitor complex (PDB code 1PPE) from their unbound components. Only 1000 random docking poses are shown for the sake of clarity, but the Pearson correlation coefficient is computed over the 10,000 evaluated poses.

4.2.1.3 pyDockLite van der Waals term

We followed the same approach to define the van der Waals distance cutoff beyond which the energy value is set to 0. As van der Waals interactions quickly decay with distance, the Pearson correlation coefficient between the standard and the truncated energies is already 1.00 for a cutoff distance of 10 Å (see Figure 4.5). However, in the first pyDockLite implementation van der Waals cutoff distance was set to 30 Å, the same value as the electrostatic cutoff. There was a reason for this counter-intuitive decision. It was quicker to calculate the atoms for which the electrostatic and van der Waals interactions were to be computed only once, for a single distance, than to do the calculation of the atoms twice, for two different distances, even though we would be computing "unnecessary" van der Waals interactions between atoms more than 10 Å apart. In a second pyDockLite implementation we modified the way by which atoms were selected, allowing to define the cutoffs independently without performance losses. All pyDockLite results presented in this thesis were obtained before the second version of pyDockLite was implemented, i.e., with a common cutoff distance of 30 Å for both, electrostatics and van der Waals terms.

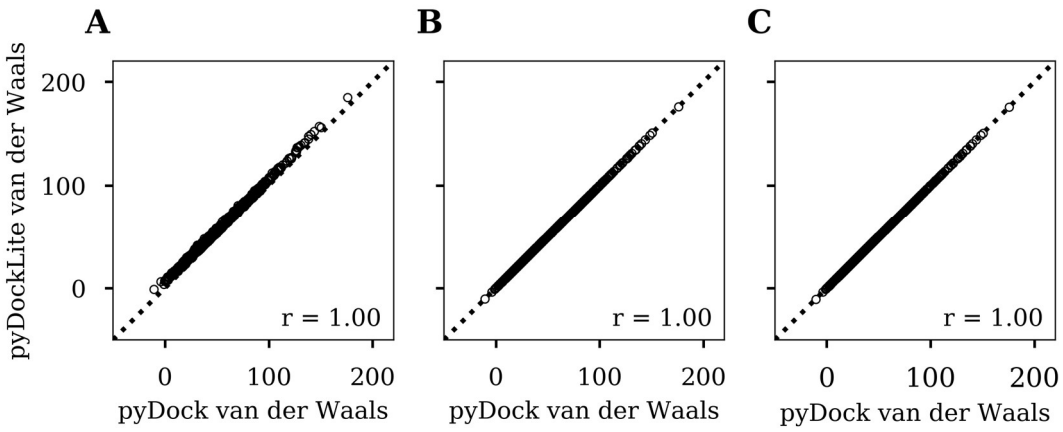

Figure 4.5. Correlation plot between pyDock (x-axis) and pyDockLite (y-axis) van der Waals for cutoff distances of 10 Å (A), 30 Å (B) and 50 Å (C). Data corresponding to 10,000 docking poses generated by FTDock to model Trypsin/CMTI-1 squash inhibitor complex (PDB code 1PPE). Only 1000 random docking poses are shown for the sake of clarity, but the Pearson correlation coefficient is computed over the 10,000 evaluated poses.

4.2.2 pyDock *vs.* pyDockLite performance

After defining the individual terms of the new pyDockLite scoring function, as above described, we compared its performance with that of pyDock. Figure 4.6 shows the correlation between pyDockLite and pyDock values for all cases of the protein-protein docking benchmark 4. We see that correlation between both functions is high, with a Pearson correlation coefficient of 0.92.

We also evaluated the docking success rates of pyDockLite scoring function over the 176 cases of protein-protein docking benchmark 4. As Figure 4.7 depicts, pyDockLite and pyDock performances were very similar. Top 10 success rates were 17% for pyDockLite and 18% for pyDock, while top 1 and top 100 were 3% and 31% for pyDockLite and 5% and 36% for pyDock, respectively.

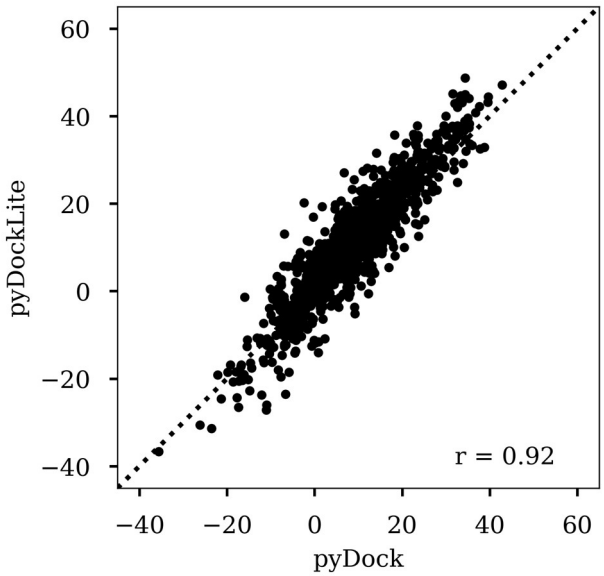

Figure 4.6. Correlation plot between pyDock (x-axis) and pyDockLite (y-axis) scoring values for the docking poses generated by FTDock for the 176 cases of protein-protein docking benchmark 4. Only 1000 random docking poses are shown for the sake of clarity, but the Pearson correlation coefficient is computed over the 1,760,000 evaluated poses.

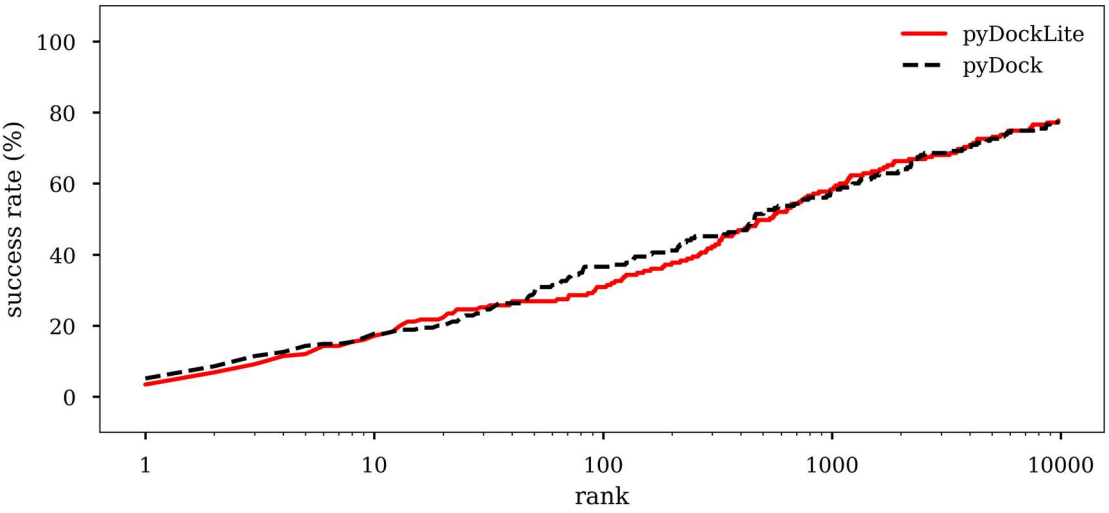

Figure 4.7. Docking success rates for pyDock (dashed-black) and pyDockLite (solid-red) over the 176 cases of protein-protein docking benchmark 4 for the top N docking models selected by the respective scoring function.

We compared pyDock and pyDockLite computational speed for several cases of different sizes. Table 4.3 shows the computation times, in seconds, required to compute 100 docking poses by applying pyDock in standard conditions, as well as pyDockLite with three different distance cutoffs. Two complexes of different sizes are included in the analysis. Complex with PDB ID 1PPE may be considered a small complex, with receptor and ligand docking subunits of 221 and 29 residues,

respectively. Complex with PDB ID 2VIS is bigger, with receptor and ligand docking subunits of 960 and 431, respectively. Regardless of the size of the complexes, even the slower version of pyDockLite was up to 10 times faster than pyDock. Therefore, pyDockLite may be considered a good proxy to pyDock when a fast scoring function is required.

Case	Size[a]	pyDock	pyDockLite (10 Å)	pyDockLite (20 Å)	pyDockLite (30 Å)
1PPE	221/29	58 s	2 s	4 s	6 s
2VIS	960/431	604 s	27 s	36 s	57 s

Table 4.3. Computation time, in seconds, required to score 100 docking poses by applying pyDock in standard conditions, and pyDockLite with distance cutoffs of 10, 20 or 30 Å. [a] Number of residues of the docking receptor/ligand subunits.

4.2.3 pyDockLite integración in LightDock for fast exploration of docking landscapes based on GSO algorithm

pyDockLite scoring function was used for the exploration of docking energy landscapes in LightDock (Jiménez-García et al. 2018b), a multi-scale protein-protein docking framework that integrates Glowworm Swarm Optimization (GSO) (Krishnanand and Ghose 2009) and ANM (Atilgan et al. 2001; Doruker, Atilgan, and Bahar 2000) representation of proteins to efficiently sample the rotational and translational space, including protein flexibility. Unlike most protein-protein docking algorithms that separate between the sampling and the scoring phase, LightDock integrates scoring and sampling while exploring the energetic landscape. In a standard LightDock run the energy function is evaluated millions of times. Therefore, a fast scoring function like pyDockLite was an excellent candidate to be included among the functions implemented within LightDock. Currently, LightDock implements nine different scoring functions. It even allows to combine them as a linear combination of terms with user-defined weights.

The predictive capabilities of LightDock were assessed on the 230 complexes of protein-protein docking benchmark 5. Two different scoring functions were used during the sampling phase: DFIRE (Zhou and Zhou 2002) (named LightDock-DFIRE) and pyDockLite (named LightDock-pyDockLite). Additionally, all the models generated with LightDock-DFIRE and LightDock-pyDockLite were merged and rescored in the end with pyDock scoring function. We called these scoring variants LightDock-DFIRE/pyDock and LightDock-pyDockLite/pyDock,

respectively. As shown in Figure 4.8-A, LightDock-pyDockLite top 10 success rate was better than that achieved by LightDock-DFIRE, and only slightly worse than the combination of FTDock and pyDock (FTDock/pyDock). pyDock rescoring improved the performance of both methods. The improvement was greater for LightDock-DFIRE, indicating that the differences between LightDock-pyDockLite and LightDock-DFIRE results depended on the scoring of the resulting models, given that the sampling with DFIRE provided good models that were only identified after rescoring them with pyDock. By combining the models generated by LightDock-pyDockLite and LightDock-DFIRE and rescoring them with pyDock, we achieved the best overall performance, with success rates of 19% and 44% for the top 10 and top 100, respectively.

We further investigate if LightDock performance depended on the flexibility of the interacting proteins. We classified the 230 cases according to the RMSD of the interface Cα atoms (Cα-IntRMSD) between the unbound and bound conformations, as defined by protein-protein docking benchmark 5, in five different categories: rigid (Cα-IntRMSD $<$ 0.5 Å), low-flexible (0.5 Å $<$ Cα-IntRMSD $<$ 1.0 Å), medium-flexible (1.0 Å $<$ I- Cα-IntRMSD $<$ 2.0 Å), flexible (2.0 Å $<$ Cα-IntRMSD $<$ 3.0 Å) and highly-flexible (Cα-IntRMSD $>$ 3.0 Å). As could be expected, FTDock/pyDock performance is higher for rigid cases, while LightDock-pyDockLite performs better in the low-flexible cases (see Figure 4.8-B). These results suggest that even though the ANM modeling is probably improving the results for flexible cases, it is also decreasing the performance for rigid cases, introducing flexibility when it is not required. LightDock-DFIRE achieved worse performance than LightDock-pyDockLite for the low-flexible cases, suggesting that, due to its coarse-grained nature, DFIRE scoring cannot take advantage of ANM modeling this type of cases.

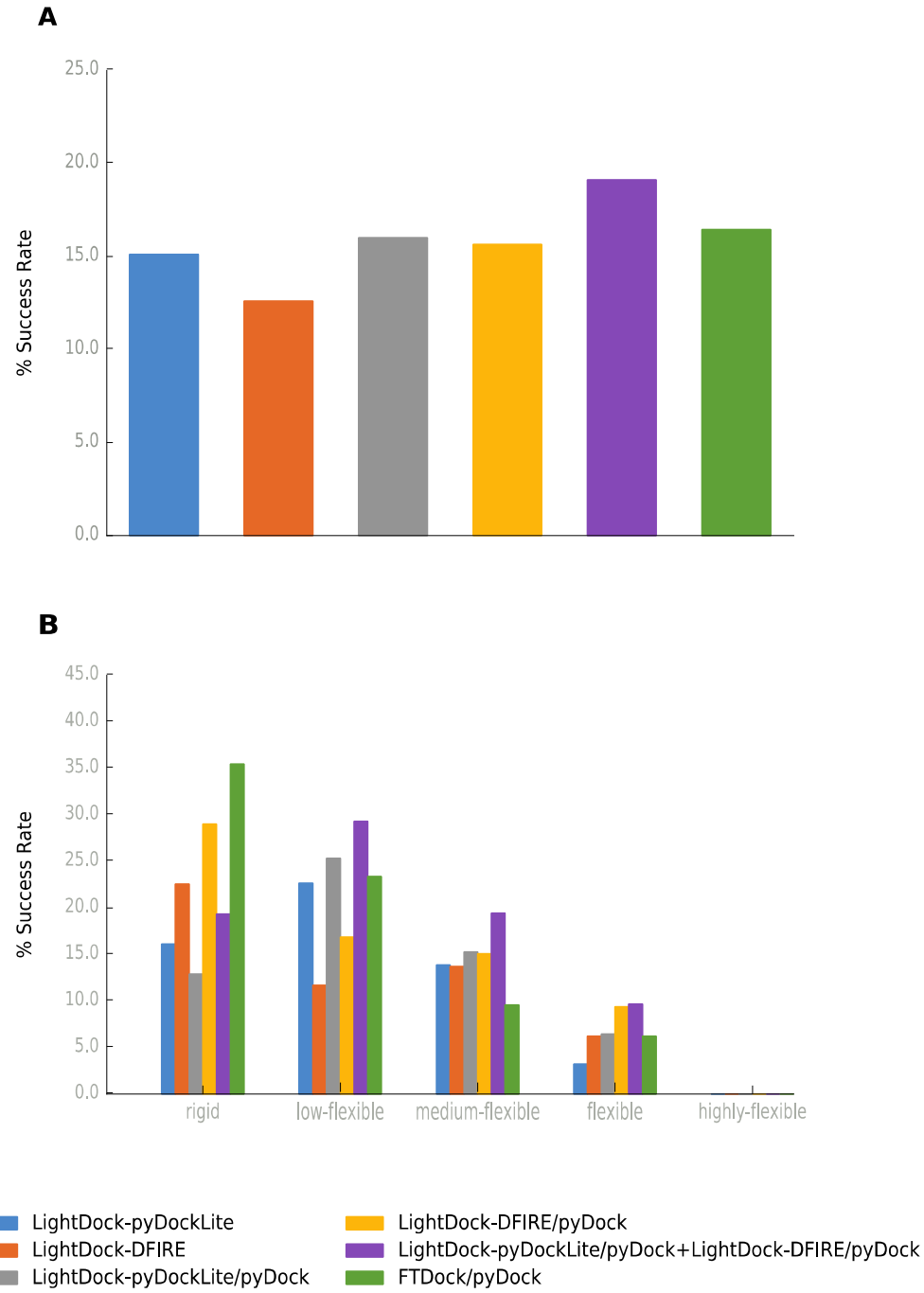

Figure 4.8. Predictive success rates for LightDock on 230 cases of the protein-protein docking benchmark 5. (A) Success rates for the top 10 docking models for: LightDock-pyDockLite (blue), LightDock-DFIRE (orange), LightDock-pyDockLite/pyDock (grey), LightDock-DFIRE/pyDock (yellow), combination of LightDock-pyDockLite/pyDock and LightDock-DFIRE/pyDock (purple). For comparison, the performance of the standard protein-protein docking protocols FTDock/pyDock (green) and ZDock 3.0.2 (red) are shown. (B) Top 10 success rates according to unbound-to-bound conformational changes measured by the RMSD of the interface Cα atoms (Cα-IntRMSD), as defined in the protein-protein docking benchmark 5. Complexes are classified as rigid (Cα-IntRMSD $<$ 0.5 Å), low-flexible (0.5 Å $<$ Cα-IntRMSD $<$ 1.0 Å), medium-flexible (1.0 Å $<$ Cα-IntRMSD $<$ 2.0 Å), flexible (2.0 Å $<$ Cα-IntRMSD $<$ 3.0 Å) and highly-flexible (Cα-IntRMSD $>$ 3.0 Å).

4.3 pyDockLite applied to fast minimization of docking poses

4.3.1 Rigid-body minimization

As explained in the methods section, pyDock standard docking protocol evaluates a set of given conformations generated in the sampling phase by FTDock. FTDock is a high-speed algorithm based on the discretization of the interacting molecules and a global scan of the translational and rotational spaces, propelled by Fast Fourier Transform algorithms, in search of conformations with maximum surface complementarity between their subunits. By discretizing the molecules and the rotational and translational space in its search, FTDock sampling of the energy landscape is rather coarse, and the resulting docking solutions, when converted from grid-representation to atomic coordinates (maintaining the internal relative atomic positions in the rigid-body approach), might not be in the optimal geometric position for a full-atom energy-based scoring function like pyDock. We reasoned it was necessary to optimize the position of each rigid-body docking pose in the new energy-based pyDock landscape. Therefore, we have developed here a methodology to search for local energy minima from a set of rigid-body sampled poses. We first developed pyDockLite, a new force field derived from pyDock scoring function in which all energetic terms, i.e., electrostatics, desolvation and van der Waals, have been modified to reduce its computational cost without losing accuracy (see section 4.2). Next, we conceived a rigid-body optimization algorithm, based on the work of Mirzaei, et al. (Mirzaei et al. 2012), that performs the minimization on the 6D manifold, space which locally resembles a Euclidean space, of the rigid, affine transformations of the ligand. The core of the method consists in using the exponential parametrization to locally map the rigid body transformation manifold onto a Euclidean space, where standard optimization algorithms can be applied (see section 3.3.1).

To test the new methodology, we performed the following protocol: 1) run a protein-protein docking with pyDock over the 176 cases of protein-protein docking benchmark 4 structures, 2) sort the final docking models according to their docking energy, 3) select the first 100 ranked scored poses and optimize them to get the final 'minimized' models. We ran this protocol in optimal conditions, i.e., docking the bound structures of the complex subunits, and realistic conditions, i.e., docking the unbound structures of complex subunits. We also tested two different weighting factors for the contribution of the van der Waals energy to the total scoring. In the standard pyDock protein-protein docking protocol the

default weighting factor for the van der Waals energy is set 0.1 to allow some degree of overlapping of the interacting atoms in the rigid-body models resulting from unbound docking and thus reduce the noise of the scoring function. As the goal of minimization is to improve the geometry of the previously generated docking solutions, we wanted to check whether increasing the contribution of the van der Waals term by setting a weighting factor equal to 1.0 could be beneficial for the scoring of such minimized docking models.

4.3.1.1 Minimization improves the geometry of near-native poses from bound docking

We analyzed whether minimization is able to improve the structural quality of the docking models. One of the most common criteria to assess the quality of a docking model is the Cα-LigRMSD with respect to the native complex structure (see section 3.1.4). Figure 4.9 shows the comparison of the Cα-LigRMSD of the docking models before and after minimization for the set of near-native poses, i.e., models with a Cα-LigRMSD less than 10 Å with respect the protein complex, generated by FTDock for the 176 cases of protein-protein docking benchmark 4, starting from the bound or the unbound structures of the interacting proteins. The minimization algorithm, with a van der Waals weight set to 1.0, performed exceptionally well when applied to docking models generated from bound structures, bringing most of the near-native poses (Cα-LigRMSD $<$ 10 Å before minimization) to minimized models close to 1 Å Cα-LigRMSD, thus efficiently converging to the native structure. The minimization is much less efficient when using a van der Waals weight of 0.1. In this case, most poses improve Cα-LigRMSD, but they do not generally converge to the native structure. Performance of the algorithm was not so good when minimizing poses generated from unbound structures regardless of the van der Waals weight used. In these conditions, the algorithm reduced only slightly the Cα-LigRMSD, and for a small number of poses.

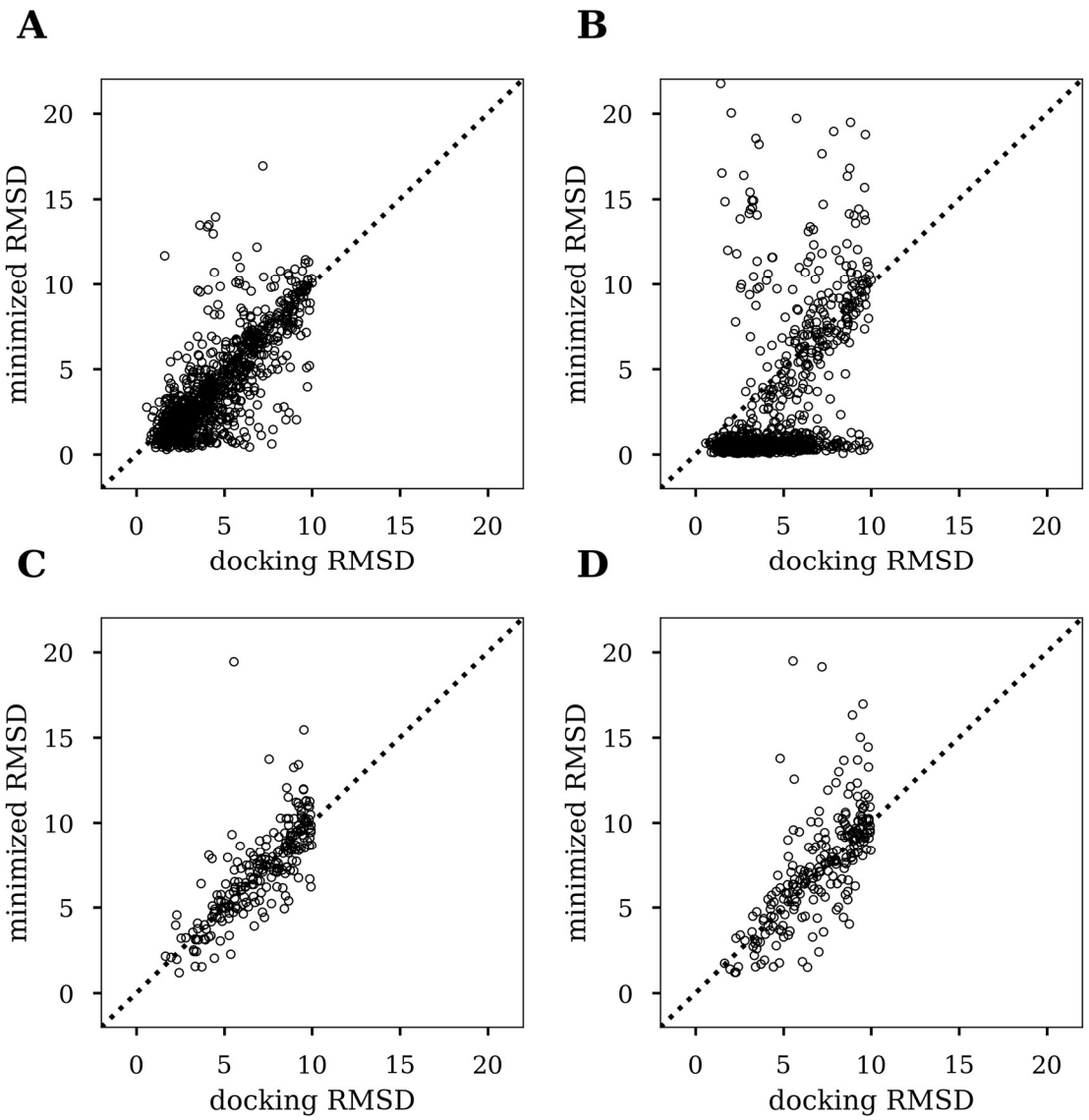

Figure 4.9. RMSD before and after rigid-body minimization of near-native docking models, i.e., with a Cα-LigRMSD ≤ 10 Å with respect the reference complex structure generated with FTDock for cases of protein-protein docking benchmark 4. Rigid-body minimizations have been performed as described in Methods by using pyDockLite with van der Waals (VDW) weight set to 0.1 or 1.0. (A) Docking models generated from bound subunit structures, minimized by pyDockLite with VDW weight set to 0.1; (B) Bound docking models, minimized by pyDockLite with VDW weight set to 1.0; (C) Docking models generated from unbound subunit structures, minimized with VDW weight set to 0.1; (D) Unbound docking models, minimized with VDW weight set to 1.0.

4.3.1.2 Minimization improves the description of docking energy landscapes

One of the objectives we pursue with the rigid-body minimization is improving the description of the docking energy landscape, with multiple minima ideally explored by the different docking models. Figure 4.10 shows the different docking energy landscapes before and after minimization, for the interaction between HISF

protein (PDB code 1THF) and amidotransferase HISH (PDB code 1K9V), represented by the distribution of pyDockLite values as a function of Cα-LigRMSD with respect to the native structure (PDB code 1GPW). When starting from bound structures, FTDock generates several near-native solutions with Cα-LigRMSD ranging from 2 to 10 Å (Figure 4.10-A, Figure 4.10-C). Most of these solutions were correctly identified by pyDockLite scoring function since they showed the lowest docking energies. When the scoring function used a van der Waals weight of 1.0, the minimization algorithm improved the description of the docking energy landscapes, for instance reducing the Cα-LigRMSD of seven docking poses to around 0.5 Å, with docking energies almost equal to those of the native complex (Figure 4.10-D and Table 4.4). Improvement of docking energy landscapes was also evident when using van der Waals weight of 0.1 (Figure 4.10-B and Table 4.4). When starting from unbound structures, FTDock could also generate near-native solutions. A few of them were ranked by pyDockLite within the top 10 poses, regardless of whether the van der Waals weight was 0.1 or 1.0. Rigid-body minimization with van der Waals weight set to 0.1 did not generally improve Cα-LigRMSD of the near-native solutions, although, it improved the Cα-LigRMSD of a non-near-native pose from 16 to 6 Å (see Table 4.5). Additionally, it kept one of the near-native poses within its top 10 ranked poses. In contrast, when minimization was performed with a van der Waals weight of 1.0, the optimization decreased the Cα-LigRMSD of most of the near-native solutions. Unfortunately, none of them were ranked within the top 10 poses (see Table 4.5).

In summary, the minimization algorithm improved the docking energy of the poses regardless of whether the minimized structures were bound or unbound. However, it could only bring the docking poses to the bottom of the energy funnel when the docking models were generated from the bound structures, that is when the docking models had the backbone and side-chains in the same conformation as the complex (Figure 4.10-B, Figure 4.10-D). These plots reveal that flexibility is essential for protein binding and a correct description of the docking landscape. It is through conformational changes that interacting proteins can reach the bottom of the energy funnel that represents the bound conformation. These plots reflect that proteins would follow the principle of minimum frustration introduce by Bryngelson and Wolynes (Bryngelson et al. 1995) not only when they fold, but also when they interact: by developing a smooth funnel-like energy landscape proteins would ensure a resilient and fast binding, avoiding local minima traps.

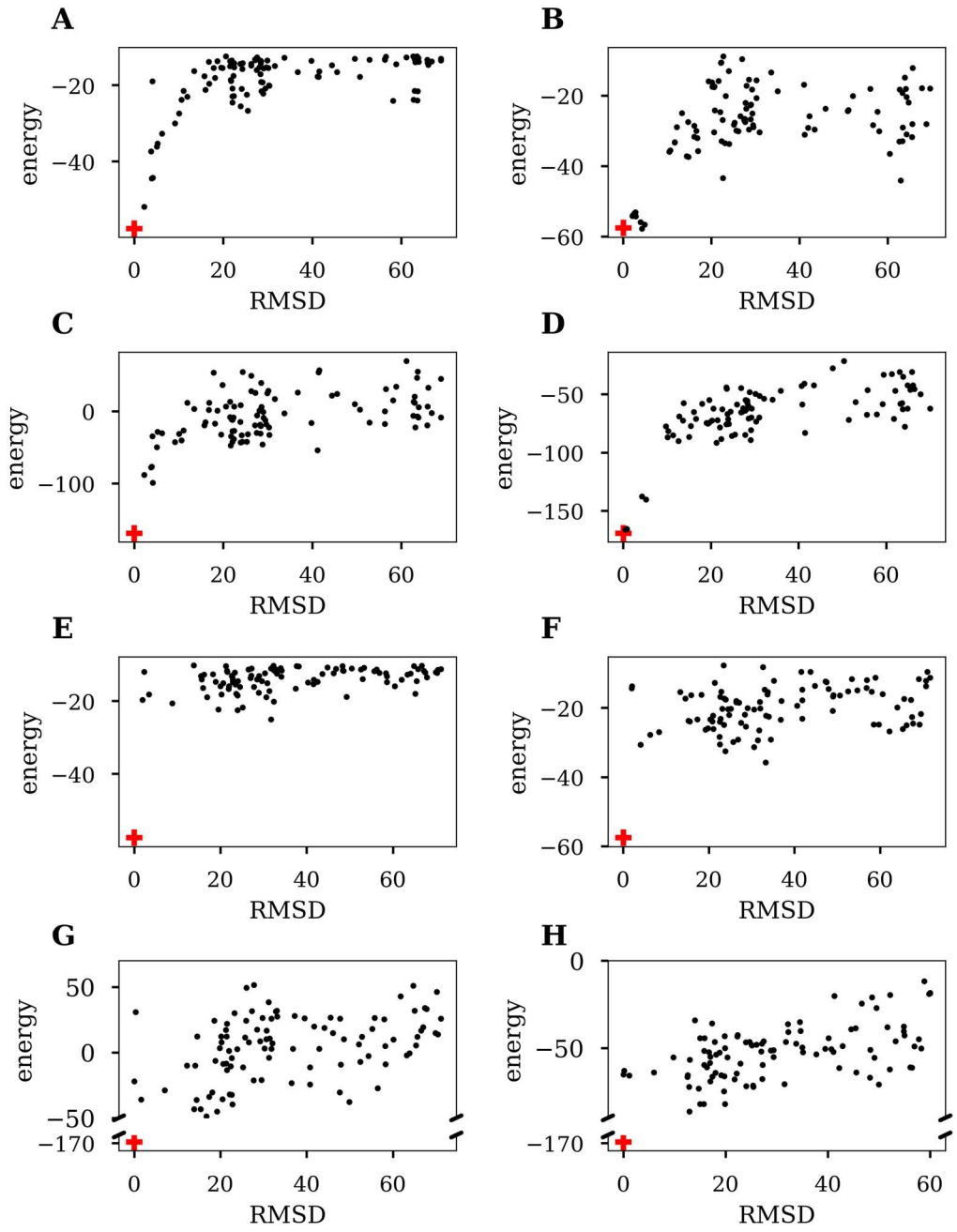

Figure 4.10. Docking energy landscapes before and after minimization for the interaction between HISF protein (PDB code 1THF) and Amidotransferase HISH (PDB code 1K9V). The plot shows, for each docking model, its docking energy calculated with pyDockLite vs. its Cα-LigRMSD with respect to the reference complex (PDB code 1GPW). The plot represents pyDock top 100 docking poses from (A) bound docking models before minimization, with pyDockLite van der Waals (VDW) weight set to 0.1, (B) bound docking models after minimization, pyDockLite VDW weight set to 0.1, (C) bound docking models before minimization, pyDockLite VDW weight set to 1.0, (D) bound docking models after minimization, pyDockLite VDW weight set to 1.0, (E) unbound docking models before minimization, pyDockLite VDW weight set to 0.1, (F) unbound docking models after minimization, pyDockLite VDW weight set to 0.1, (G) unbound docking models before minimization, pyDockLite VDW weight set to 1.0, and (H) unbound docking models after minimization, pyDockLite VDW weight set to 1.0. The native complex is shown for comparison (red cross).

	Before rigid-body minimization			After rigid-body minimization			
Pose	RMSD (Å)	Rank 0.1vdw[a]	Rank 1.0vdw[b]	RMSD[c] (Å)	Rank 0.1vdw[a]	RMSD[d] (Å)	Rank 1.0vdw[b]
824	2.2	1	2	2.8	4	0.8	7
5671	3.8	4	3	2.1	6	0.5	4
2139	3.9	2	4	2.7	9	0.6	6
8922	4.1	35	16	2.5	7	0.5	3
1000	4.1	3	1	2.6	8	0.6	5
535	5.0	5	6	4.2	1	0.5	2
218	5.2	6	26	4.8	2	4.2	9
3734	6.2	7	24	3.9	3	5.2	8
3479	9.1	8	10	2.0	5	0.5	1
5642	19.7	54	90	16.9	16	9.7	25

Table 4.4. Cα-LigRMSD and rank of near-native poses for the complex with PDB code 1GPW, between HISF protein and Amidotransferase HISH, before and after rigid-body minimization of docking models generated from bound structures. We have tested the performance of the minimization algorithm with pyDockLite as optimization function with two different van der Waals weights: 0.1 and 1.0. [a] Rank of poses with pyDockLite VDW weight set to 0.1. [b] Rank of poses with pyDockLite VDW weight set to 1.0. [c] Cα-LigRMSD of minimized poses with pyDockLite VDW weight set to 0.1. [d] Cα-LigRMSD of minimized poses with pyDockLite VDW weight set to 1.0.

	Before rigid-body minimization			After rigid-body minimization			
Pose	RMSD (Å)	Rank 0.1vdw[a]	Rank 1.0vdw[b]	RMSD[c] (Å)	Rank 0.1vdw[a]	RMSD[d] (Å)	Rank 1.0vdw[b]
5721	1.9	7	19	2.0	82	1.4	28
904	2.2	69	88	1.9	80	1.2	23
2227	3.4	12	8	4.0	4	2.6	20
3362	8.8	5	15	8.3	12	8.1	27
8418	16.1	54	62	6.2	11	16.1	8

Table 4.5. Cα-LigRMSD and rank of near-native poses for the complex with PDB code 1GPW, between HISF protein (PDB code 1THF) and Amidotransferase HISH (PDB code 1K9V), before and after rigid-body minimization of docking models generated from unbound structures. We have tested the performance of the minimization algorithm with pyDockLite as optimization function with two different VDW weights: 0.1 and 1.0. [a] Rank with pyDockLite VDW weight set to 0.1. [b] Rank with pyDockLite VDW weight set to 1.0. [c] Cα-LigRMSD of minimized poses with pyDockLite VDW weight set to 0.1. [d] Cα-LigRMSD of minimized poses with pyDockLite VDW weight set to 1.0.

4.3.1.3 Minimization improves docking success rates of bound docking

We further investigated whether the minimization protocol could improve identification of near-native docking solutions and therefore, the predictive success rates of docking. For this, we re-ranked the docking poses according to the

minimized energy and computed the success rates. When docking models were generated from bound structures, the top 10 success rates of the minimized energies were 74% and 52%, using the van der Waals weight of 1.0 and 0.1, respectively. These performance differences depending on the van der Waals weight are remarkable, especially if we consider that pyDock success rates were 60% regardless of the van der Waals weight. This is indicative of the relevance of van der Waals forces and surface complementarity when the interacting subunits are very close to the native structures. After minimization with pyDockLite van der Waals weight of 1.0, a total of 32 cases found a near-native solution within the top 10 predictions, while 3 cases lost their near-native solution within the top 10 predictions. Contrarily, when the van der Waals weight was set to 0.1, 18 cases lost their near-native solution within the top 10 predictions, and only 5 found a new near-native solution after minimization. When docking models were generated from unbound structures, the top 10 success rates obtained with the pyDockLite optimized structures were 18% and 17%, with van der Waals weight set to 0.1 and 1.0, respectively. These values were similar to the success rates obtained with pyDock: 18% and 11%, when using van der Waals weights of 0.1 and 1.0, respectively.

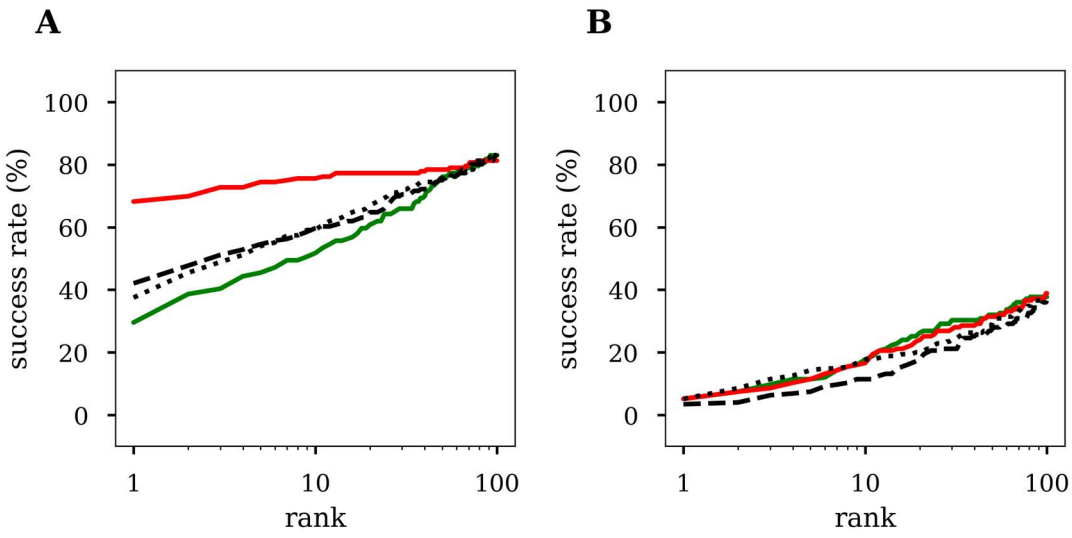

Figure 4.11. Predictive success rates for top 100 docking models generated by pyDock from (A) bound structures, and (B) unbound structures of protein-protein docking benchmark 4, before minimization and scored by pyDockLite with van der Waals weight of 0.1 (dotted-black); before minimization and scored by pyDockLite with van der Waals weight of 1.0 (dashed-black); after minimization, with pyDockLite van der Waals weight of 0.1 (solid-green); and after minimization, with pyDockLite van der Waals weight of 1.0 (solid-red).

4.3.2 Flexible minimization: Minimization with normal modes

In section 4.3.1 we showed that protein flexibility is essential for protein binding and should be considered if we want to reach the bottom of the characteristic energy funnels of protein-protein interaction landscapes. Several studies indicate that we can use a reduced set of normal modes to reproduce important biological motions (Ma, 2005) correctly. Inspired by these studies, we analyzed whether we could apply normal modes to improve the scoring and the quality of the models generated with our docking protocol. We abandoned the rigid body model of the previous section and chose the anisotropic network model (ANM) (Atilgan et al. 2001) to represent protein structures. Within the ANM framework, molecules are represented as elastic networks of beats and springs. Each residue is represented by its Cα atom as a beat, while springs represent interactions between beats (see Figure 1.2). Only interactions within a cutoff distance are considered. In our case, the cutoff was set to 15 Å. The system can be decomposed into a set of normal modes, i.e., eigenvectors, which can describe the motion and states of molecules. The linear combination of normal modes can be used to generate different molecule structures. Normal modes for receptor and ligand were pre-calculated with ProDy (Bakan, Meireles, and Bahar 2011). The five non-trivial, lowest frequency normal modes were selected to account for protein flexibility during the minimization stage as in Zacharias et al. (May and Zacharias 2008). Thus, a total of sixteen degrees of freedom, i.e., three related to the rotation matrix, three to the translation vector and ten to the subunits normal modes were optimized. It should be noted that our approach used normal modes to model backbone flexibility. Side chains were treated as rigid bodies attached to the corresponding backbone atom. We tested this new NM-based minimization by running a similar protocol to that executed with the rigid body minimization framework, i.e., 1) run a protein-protein docking with pyDock over the 176 cases of benchmark 4 structures, 2) sort the final docking models according to their docking energy, 3) select the first 100 ranked scored poses and optimize them to get the final 'minimized' models. This time, we only ran the protocol in realistic conditions, i.e., docking the unbound structures of complex subunits, and tested two van der Waals weight values: 0.1 and 1.0.

4.3.2.1 Minimization effect on pose's RMSD

Figure 4.12 shows the Cα-LigRMSD of the docking models before and after minimization for the set of near-native poses of the 176 cases of benchmark 4. We see that including global protein flexibility through normal modes does not

systematically reduce the Cα-LigRMSD of the docking models. Results seem to be slightly better with van der Waals weight set to 1.0. However, they are still poor results that do not improve those obtained with the rigid body minimization of unbound structures (see Figure 4.9-C, Figure 4.9-D).

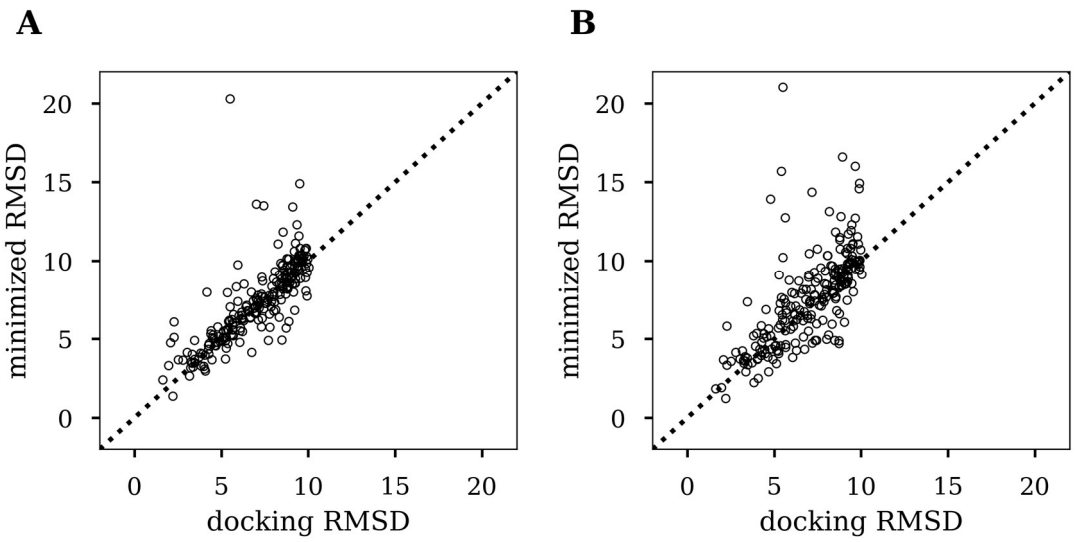

Figure 4.12. Efficiency of normal modes (NM) based minimization of near-native docking models from the protein-protein docking benchmark 4, in terms of change in Cα-LigRMSD with respect to the corresponding reference native structure, before and after minimization. Docking models were generated with FTDock starting from unbound subunit structures. NM-based minimization was performed by using pyDockLite as the function to optimize with van der Waals weight set to (A) 0.1, or (B) 1.0.

4.3.2.2 Energy landscape with and without minimization

Figure 4.13 represents the Cα-LigRMSD and energy behavior for the first 100 docked poses of case 1GPW after minimization with normal modes. Minimization with normal modes improved the docking energy, but not the Cα-LigRMSD of the docking poses. All 4 near-native solutions generated by FTDock remained with Cα-LigRMSD below 10 Å after minimization regardless of the van der Waals weight value of the scoring function. When the weight is set to 0.1, a new near-native solution is found with pose 8418, which Cα-LigRMSD decreases from 16 to 8 Å after minimization. However, none of the near-native structures are ranked within the top 10 ranked poses. On the contrary, most of the ranks of near-native solutions improved with a van der Waals weight of 1.0, with 3 near-native poses within the top 10 solutions (see Table 4.6). None of the optimized docking poses end up with docking energy closed to that of the native complex. This could indicate that in addition to backbone flexibility, we should consider side-chain flexibility when modeling protein conformational changes upon binding.

	Before NM-based minimization			After NM-based minimization			
Pose	RMSD (Å)	Rank 0.1vdw[a]	Rank 1.0vdw[b]	RMSD[c] (Å)	Rank 0.1vdw[a]	RMSD[d] (Å)	Rank 1.0vdw[b]
5721	1.9	7	19	3.3	14	1.9	5
904	2.2	69	88	6.0	51	5.8	33
2227	3.4	12	8	4.9	11	3.7	3
3362	8.8	5	15	8.1	12	8.6	10
8418	16.1	54	62	7.0	24	15.7	23

Table 4.6. Cα-LigRMSD and ranks of near-native poses for the complex with PDB code 1GPW, between HISF protein (PDB code 1THF) and Amidotransferase HISH (PDB code 1K9V], before and after NM-based minimization of docking models generated from unbound structures. We have tested the performance of the minimization algorithm with pyDockLite as optimization function with two different van der Waals (VDW) weights: 0.1 and 1.0. [a] Rank with pyDockLite VDW weight set to 0.1. [b] Rank with pyDockLite VDW weight set to 1.0. [c] Cα-LigRMSD of minimized poses with pyDockLite VDW weight set to 0.1. [d] Cα-LigRMSD of minimized poses with pyDockLite VDW weight set to 1.0.

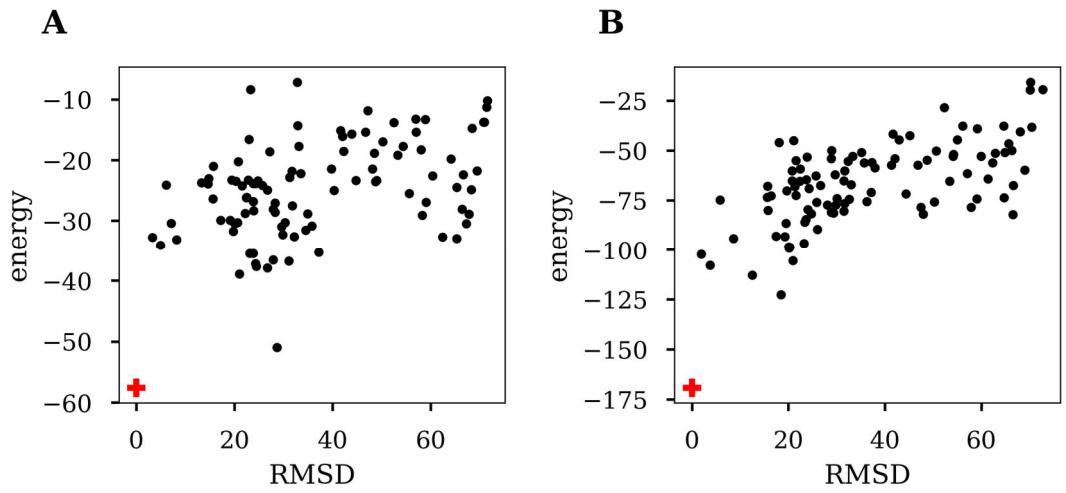

Figure 4.13. Cα-LigRMSD *vs.* docking energy for the complex with PDB code 1GPW, between HISF protein (PDB code 1THF) and Amidotransferase HISH (PDB code 1K9V). The plot represents pyDock top 100 docking poses from (A) unbound structures after NM-based minimization, with pyDockLite VDW weight set to 0.1, (B) unbound structures after NM-based minimization, with pyDockLite VDW weight set to 1.0. The native complex is shown for comparison (red cross).

4.3.2.3 Flexible minimization: little effect on docking success rates

Even though the minimization with normal modes does not improve the ligand Cα-LigRMSD of the complex models, Figure 4.14 shows that the success rates slightly improved. Top 10 success rates of the minimized scoring functions were 19% and 20% with the van der Waals weight set to 0.1 and 1.0, respectively.

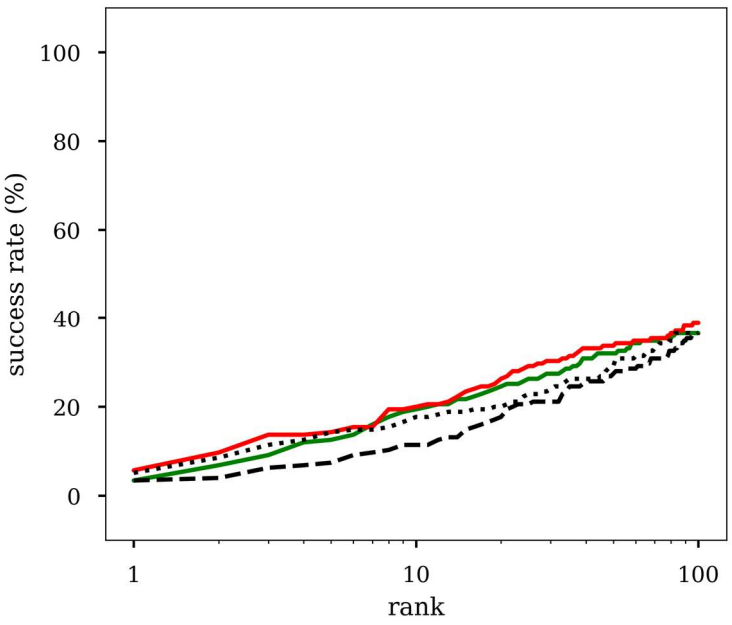

Figure 4.14. Predictive success rates for top 100 docking models generated by pyDock on unbound cases of protein-protein docking benchmark 4, after: scoring by pyDockLite with van der Waals weight set to 0.1 (dotted-black), scoring by pyDockLite with van der Waals weight set to 1.0 (dashed-black), NM-based minimization with pyDockLite van der Waals weight set to 0.1 (solid-green), and NM-based minimization with pyDockLite van der Waals weight set to 1.0 (solid-red).

We checked if improvement of the success rates could depend on the flexibility of the cases. We classified the cases according to the conformational variation of its subunits, measured as the interface RMSD of the Cα atoms (Cα-IntRMSD) in the bound and unbound conformations as defined in the protein-protein docking benchmark 4 (Hwang et al. 2010). The resulting categories were the following: rigid (Cα-IntRMSD $<$ 0.5 Å), low-flexible (0.5 Å $<$ Cα-IntRMSD $<$ 1.0 Å), medium-flexible (1.0 Å $<$ Cα-IntRMSD $<$ 2.0 Å), flexible (2.0 Å $<$ Cα-IntRMSD $<$ 3.0 Å) and highly-flexible (Cα-IntRMSD $>$ 3.0 Å). Success rates slightly improved after minimization only for the medium-flexible and flexible groups (see Figure 4.15).

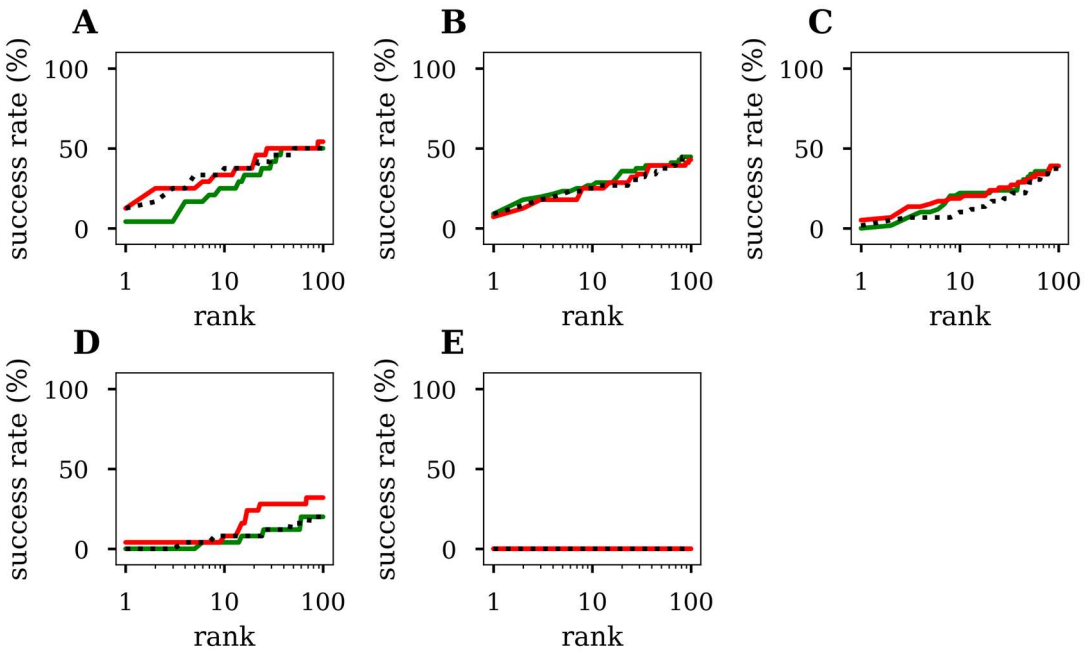

Figure 4.15. Predictive success rates for top 100 docking models generated by pyDock on unbound cases of protein-protein docking benchmark 4, after: scoring by pyDockLite with van der Waals (VDW) weight set to 0.1 (dotted-black), scoring by pyDockLite with VDW weight set to 1.0 (dashed-black), NM-based minimization with pyDockLite VDW weight set to 0.1 (solid-green), and NM-based minimization with with pyDockLite VDW weight set to 1.0 (solid-red) for cases identified according to the Cα-IntRMSD as (A) rigid, (B) low-flexible, (C) medium-flexible, (D) flexible, and (E) highly-flexible cases.

4.3.3 pyDock and SCWRL: Side-chain optimization

In section 4.3.2 we described the inclusion of normal modes into our minimization algorithm to model backbone flexibility and improve the performance of our docking protocol. In that approach, side chains were treated as rigid bodies fixed to the backbone atoms. In this section, we will add side-chain flexibility in our docking procedures. Under the rigid-body paradigm, the docking poses generated in the sampling phase keep the side chain conformations of the unbound subunits. Many groups have developed a refinement protocol to optimize the side chain positions within the context of the complex model. We adopted a similar strategy and combined FTDock, pyDock and SCWRL (Krivov, Shapovalov, and Dunbrack 2009). Specifically, the steps we followed for a given case were: 1) generate 10,000 docking poses with FTDock, 2) for each docking pose, optimize the side chains of the interface residues with SCWRL4. We defined as interface residues all residues with at least one non-hydrogen atom within 10 Å of any non-hydrogen atom of the complementary subunit, 3) evaluate the optimized docking poses with pyDock scoring function. It is worth noting that backbone atoms will remain fixed during the process. Therefore, the Cα-LigRMSD of the poses will not change, but we

could expect an improvement in the scoring by removing false positives and false negatives. We ran the protocol over a subset of cases of protein-protein docking benchmark 4. We filtered out all those complexes with more than two chains, ending up with a total of 118 cases. Figure 4.16 shows the results we obtained. The optimized protocol did not increase pyDock top 10 success rate, but improved top 100 success rate achieving a 46%, while pyDock top 100 success rate was 38%.

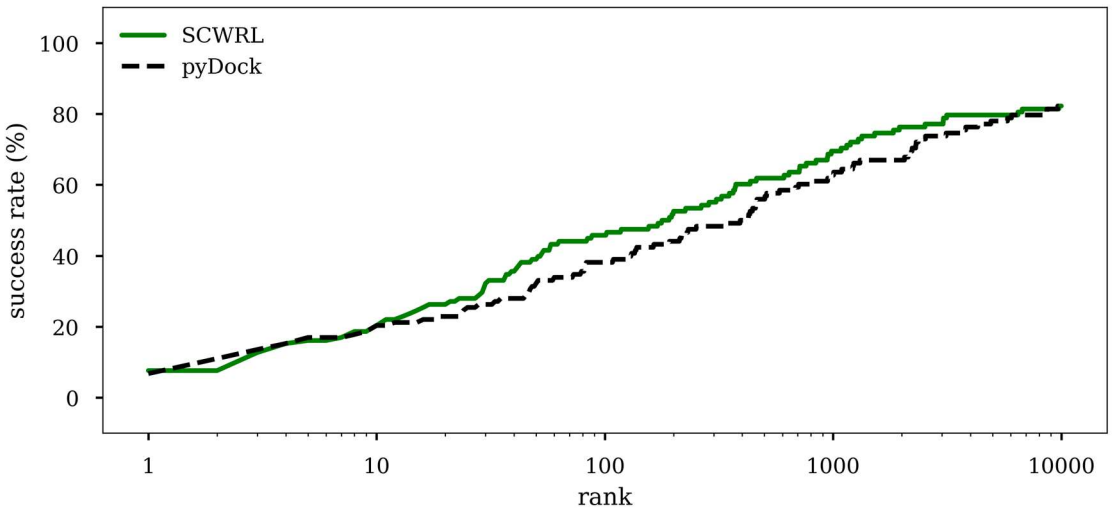

Figure 4.16. Success rates of SCWRL-based side chain optimization of docking models generated for unbound cases of protein-protein docking benchmark 4 constituted by only two chains (solid black line). For comparison, results by pyDock are also shown (green dashed line).

4.4 Ensemble-based description of docking landscapes

Most protein-protein docking protocols uses scoring functions to evaluate docking poses and discriminate between good, i.e., near-native, and bad conformations. The implicit assumption is that the different energetic minima forming the docking energy landscape are represented by single docking poses, being rigid-body or flexible conformations, which are scored individually. In this thesis, we have analyzed the concept that each energetic minima of the docking energy landscape can be formed by ensembles of docking orientations or conformations, and we have explored the consequences of scoring each minimum on the basis of such ensembles, i.e., not just relying on a single docking pose or conformation. To further develop this rationale, we present here different methodological approaches in which ensembles of poses are generated and scored. First, we describe a method that creates the ensembles by clustering the rigid-body docking poses generated in the sampling phase. In the second approach, we generate the

ensembles by applying restricted molecular dynamics to the original docking poses. Finally, we present the results we obtain when we combine both approaches.

4.4.1 Rigid-body docking ensembles by clustering docking poses

In protein-protein docking, several conditions are required to obtain near-native solutions. On the one hand, the structures of the docked subunits should be appropriate, i.e., sufficiently similar to the bound structures. On the other hand, the subunits must be correctly oriented in space. Rigid-body docking algorithms keep the structural conformation of the subunits fixed and vary their orientation in their search for protein complex models. Since in a realistic case, the docked subunits are not bound, but unbound structures, the algorithms cannot always identify correctly oriented poses as their best-scored solutions due to incorrect conformations of interface residues. Nevertheless, based on the assumption of the existence of funnel-like docking landscapes, it is reasonable to think that, on average, algorithms will score conformations around the correct orientation higher than conformations far from the correct orientation. Based on this idea, we developed a protocol that first divides the conformational space into clusters or ensembles of docking poses and later searches near-native solutions in those ensembles with better scores. A similar clustering approach applied to docking has been recently published (Pfeiffenberger et al. 2017). Here, we employed pyDock and MODELLER zDope as scoring functions and used protein-protein docking benchmark 4 to evaluate its applicability. More specifically, the steps we followed were 1) run pyDock over the 176 cases of benchmark 4 structures, 2) sort the final docking models according to their docking energy, 3) select the first 500 ranked scored poses and cluster them, 4) for each cluster of docking poses, compute the mean and best values of pyDock and zDope and use these values to score the ensembles, 5) for each ensemble select a representative pose and evaluate the performance of the protocol (see Figure 4.17).

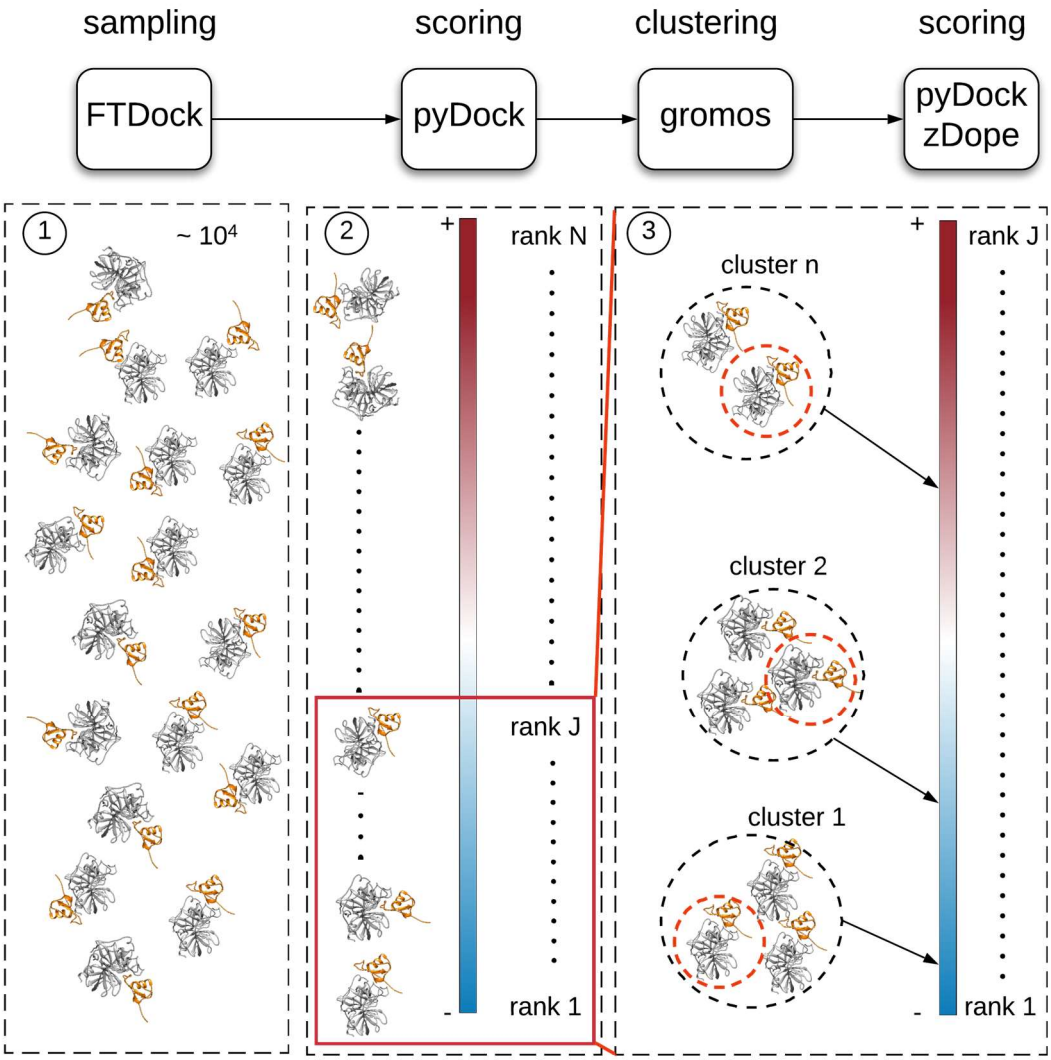

Figure 4.17. Schematic overview of the method. 1) Run FTDock to generate 10,000 docking poses. 2) Evaluate the docking poses with pyDock scoring function. Select the first 500 ranked docking poses. 3) Cluster the selected poses with gromos and a 10 Å cutoff. Rank clusters according to the 'independent' and consensus scores derived from the mean and best (minimum) values of pyDock energy and zDope per cluster. Select a cluster representative based on its pyDock energy and zDope score.

4.4.1.1 Cluster size

Clustering is a technique commonly applied in protein-protein docking to remove similar solutions and improve scoring. In this work, we have used the clustering algorithm gromos (Daura et al. 1999) to discretize the conformational space in clusters or ensembles of docking poses with Cα-LigRMSD between them below 10 Å and then rank these ensembles according to the aggregated values of pyDock and MODELLER zDope applied to its members.

The average number of clusters per case was 245±74, with a mean number of poses per cluster of 2.03±2.30. Interestingly, we observed that clusters with near-

native solutions are, on average, bigger (6.21±7.13 average size) than clusters without near-native solutions (2.00±2.18 average size), (Welch test p-value: 1.49E-145). However, the significant variability in size among clusters precludes the use of this parameter to identify the near-native clusters.

4.4.1.2 Combining several scoring functions in consensus scores improves performance

We used two different scoring functions to evaluate the docking poses in the cluster ensembles. PyDock (P) (T. M.-K. Cheng, Blundell, and Fernandez-Recio 2007) is a combination of three energetic terms, while zDope (Z) (Shen and Sali 2006) is a statistical potential developed within the MODELLER package to assess the quality of structural models. We computed the mean and best values of these scoring functions for all docking poses within a given cluster, i.e., P_{mean} and P_{best} for pyDock, Z_{mean} and Z_{best} for zDope, and used these values to rank the clusters. Combining the cluster ranks of these four independent scores, we developed consensus scores that we also applied to evaluate the clusters. For instance, the consensus score "P_{mean},P_{best}" was computed by summing up, for each cluster, its P_{mean} and P_{best} ranks. This way, we built five different consensus scores: "P_{best},Z_{best}", "P_{mean},P_{best}", "P_{mean},Z_{mean}", "Z_{mean},Z_{best}" and "P_{mean},P_{best},Z_{mean},Z_{best}". As cluster representative, we chose the model with the best pyDock energy "pose(P_{best})", the best zDope "pose(Z_{best})" or the model with the best consensus score "pose(P, Z)". Figure 4.18 shows the top10 success rates obtained for the different scoring strategies. We can compare these results with the top 10 'standard' success rates, i.e., those we obtained if we scored the docking poses, without the clustering framework, according to their pyDock energy (P_{pose}), zDope score (Z_{pose}) or consensus score (P_{pose},Z_{pose}) which are 18%, 18%, and 20%, respectively (see Figure 4.19). While P_{best} and Z_{best} slightly outperformed the standard scoring schemes P_{pose} and Z_{pose}, the clustering 'independent' scores based on computing the mean cluster values P_{mean} and Z_{mean} obtained worse success rates. However, when these mean cluster values were combined into the clustering consensus score "P_{mean},Z_{mean}" the performance significantly improved. The method applied to select the cluster representative does not seem to have as much influence in the results as the strategy to score the clusters. Selecting the pose with best zDope within each cluster appears to be the best method, but the differences with the other two methods are small.

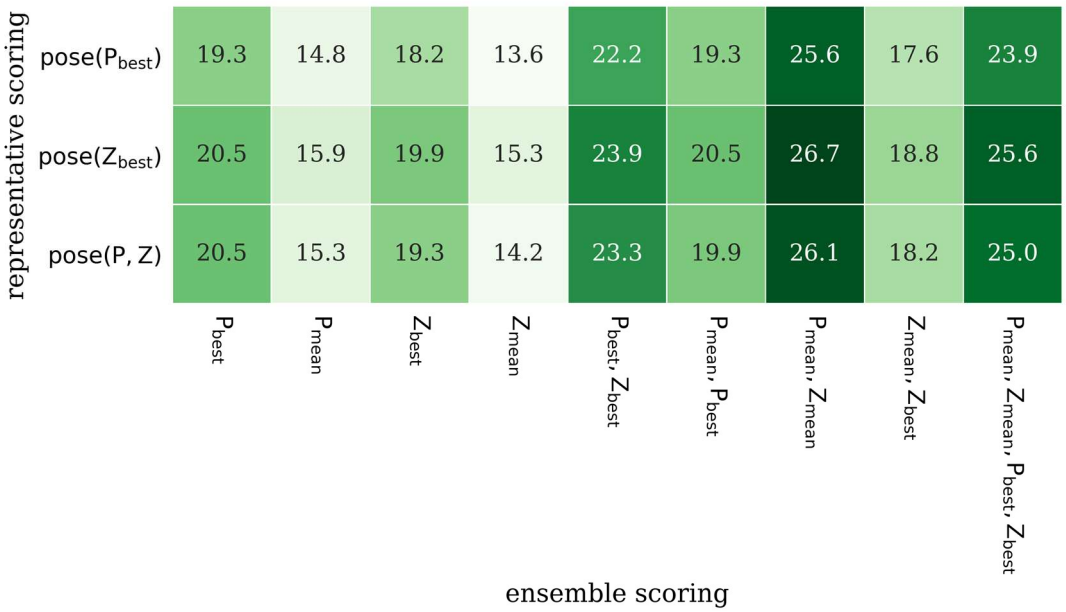

Figure 4.18. Docking success rates for the top 10 predicted models on the protein-protein docking benchmark 4 for several clustering scoring schemes. For each scoring scheme, the figure shows the strategy used to score the clusters/ensembles (x-axis), and the strategy used to pick the cluster/ensemble representative (y-axis).

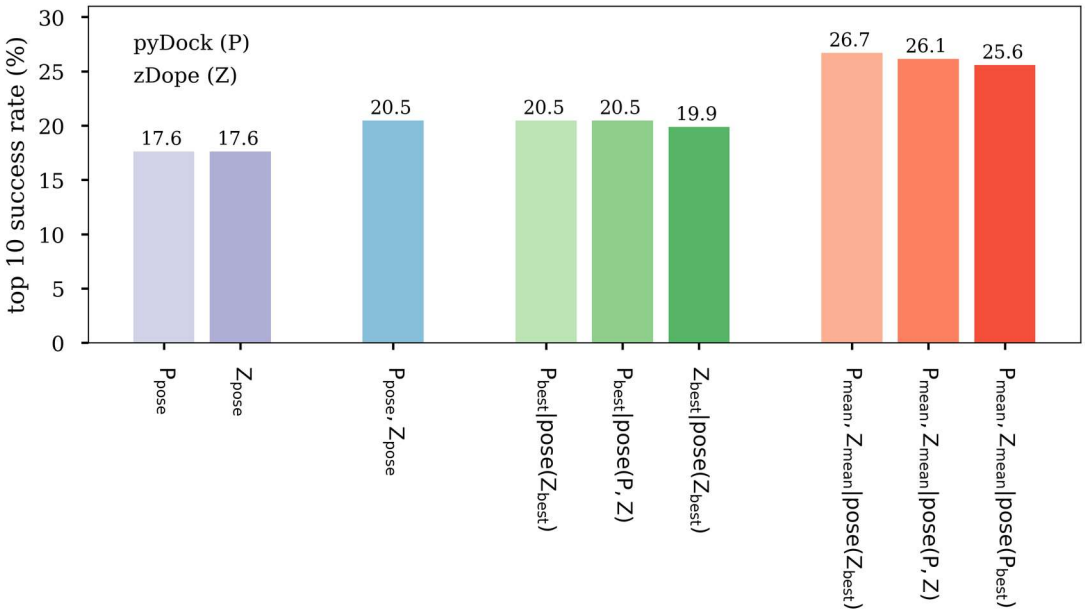

Figure 4.19. Docking success rates for the top 10 predicted models on the protein-protein docking benchmark 4 for the best-performing clustering scoring schemes, based on independent scoring (green bars) or consensus scoring (red bars). For comparison, scoring of the original FTDock docking poses according to pyDock (P) and MODELLER zDope (Z) are shown (grey bars), as well as scoring of the original FTDock docking poses according to the consensus score of pyDock and MODELLER zDope (blue bar).

4.4.2 Flexible conformational ensembles by restricted molecular dynamics

In the previous section, we generated docking ensembles by clustering rigid-body docking poses and analyzed different strategies for scoring them. Here we explore the use of docking ensembles generated by conformational variability. With this purpose, we have generated conformational ensembles with MODELLER (Sali and Blundell 1993), a software extensively applied for homology modeling of protein structures by the satisfaction of spatial restraints. Here, we used it to generate, from each pyDock docking pose, ensembles of 50 conformers. We ran MODELLER with the default parameters. Each model was first optimized with the variable target function method (VTFM) with conjugate gradients (CG) and then refined using molecular dynamics (MD) with simulated annealing (SA). We expected this method would improve the conformational and energetic description of the different minima forming the docking landscape. On the one hand, it explicitly models the flexibility of backbone and side-chains, departing from the rigid-body paradigm. On the other hand, each docking pose is going to be described by an ensemble of 50 models, instead of by a single structure.

Additionally, these conformers are generated by a protocol that includes the optimization of an energy function as one of its steps. We have already mentioned that within the rigid-body approximation near-native conformations can be wrongly discarded because of the backbone and side-chains differences between the unbound and the bound state. Here, as each docking pose generates 50 optimized models we expected both false-positive and false-negative to be identified. The protocol steps were the following: 1) run pyDock over the 176 cases of benchmark 4 structures, 2) sort the final docking models according to their docking energy, 3) select the first 500 ranked scored poses for further processing, 4) run MODELLER to generate an ensemble of 50 conformers from each selected docking pose, 5) for each ensemble compute the mean and best values of pyDock and zDope and use these values to score the ensembles, 5) for each ensemble select a representative structure and evaluate the performance of the protocol (see Figure 4.20).

We applied similar independent and consensus scores as the ones we described in the previous section, with the only difference that this time we used them to evaluate ensembles of conformers instead of clusters of docking poses. Similarly, as ensemble representative we chose the conformer with best pyDock energy "conformer(P_{best})", best zDope "conformer(Z_{best})" or the conformer with best consensus score "conformer(P_{best}, Z_{best})". Additionally, we also used as ensemble

representative the docking pose from which the ensemble had been generated "pose".

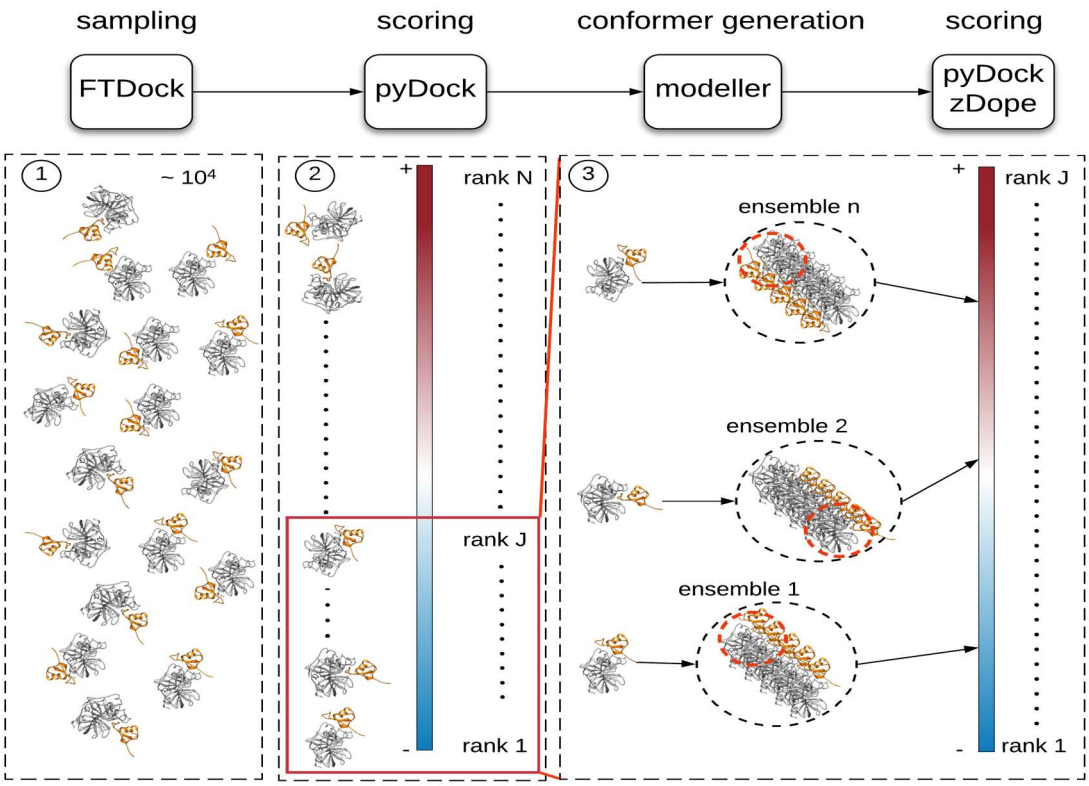

Figure 4.20. Schematic overview of the method. 1) Run FTDock to generate 10,000 docking poses. 2) Evaluate the docking poses with pyDock scoring function. Select the first 500 ranked docking poses. 3) For each selected pose, generate an ensemble of 50 conformers with MODELLER. Rank ensembles according to the independent and consensus scores derived from the mean and best (minimum) values of pyDock energy and zDope per ensemble. Select an ensemble representative based on its pyDock energy and zDope score.

As Table 4.7 shows, we did not find any significant improvement in the Cα-LigRMSD or the zDope scores of the conformers with respect the original docking poses. Conversely, the pyDock energy of the conformers improved with respect to the original docking poses. More interestingly, both the mean and the best pyDock ensemble energies showed significantly greater improvement (Welch test p-value: 1.01E-36) in ensembles from near-native solutions than in ensembles from non-near-native solutions. For example, the mean pyDock energy decreased on average -2.51±3.27 in ensembles from non-near-native solutions and -4.48±3.72 in ensembles from near-native solutions (see Table 4.7 and Figure 4.21). This could indicate that the method is able to explore the energy wells around the native solutions, which in principle should have lower energy than the non-native energy minima.

	All[a]	Near-native[b]	Non-near-native[c]
ΔligRMSD (Å)	-0.06 ± 2.52	0.11 ± 2.38	-0.07 ± 2.52
ΔzDope	0.40 ± 0.26	0.42 ± 0.24	0.40 ± 0.26
ΔpyDock	-2.52 ± 3.28	-4.48 ± 3.72	-2.51 ± 3.27

Table 4.7. Cα-LigRMSD, zDope and pyDock average changes of the conformer structures with respect to the original docking poses. Average computations have been performed for [a] all conformers, [b] conformers generated from near-native docking poses, [c] conformers generated from non-near-native docking poses.

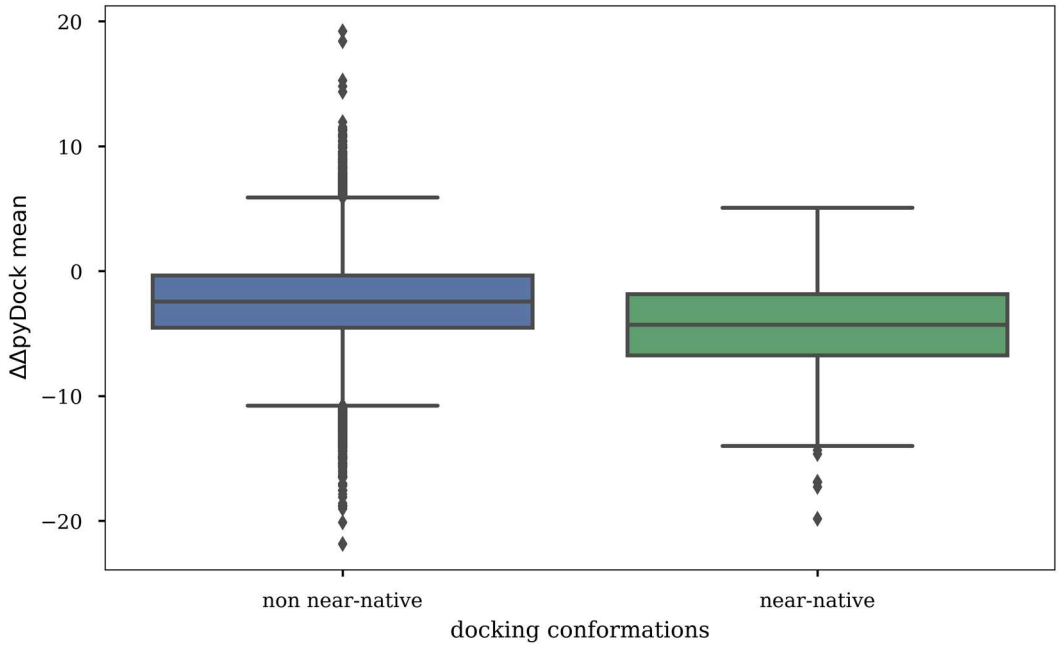

Figure 4.21. Average differences in pyDock energy of MODELLER ensembles with respect to the original docking pose, for ensembles from near-native or non-near-native docking solutions.

The results indicate that all the independent scores achieved higher top 10 success rates than the 'standard' docking (see Figure 4.22 and Figure 4.23). The independent scores based on pyDock (P_{mean} and P_{best}) obtained better results than those based on zDope (Z_{mean} and Z_{best}). Consensus scores improved the top 10 success rates as long as they included 'independent' scores based on both pyDock and zDope, i.e., the consensus scores "P_{mean},P_{best}" and "Z_{mean},Z_{best}" achieved discrete results compare to "P_{mean},Z_{mean}", "P_{best},Z_{best}" or "P_{mean},Z_{mean},P_{best},Z_{best} ".

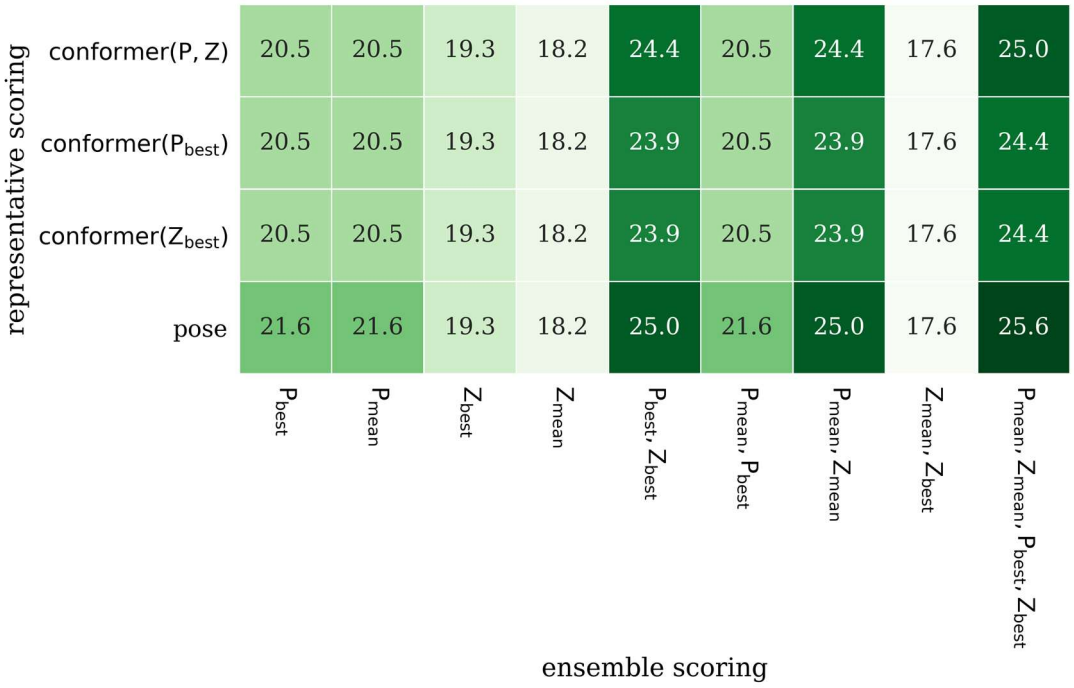

Figure 4.22. Docking success rates for the top 10 predicted models on the protein-protein docking benchmark 4 for several scoring schemes using ensembles obtained with restricted molecular dynamics: Strategy used to score the clusters/ensembles (x-axis), strategy used to pick the cluster/ensemble representative (y-axis).

As in the previous section, the method chosen to select the ensemble representative does not alter the results much. Notably, selecting as ensemble representative the pose from which the ensemble has been generated seems to be slightly better than choosing the conformers with best pyDock energy, zDope score or consensus score. This could be in line with the previously expressed idea that the Cα-LigRMSD of the conformers does not significantly improve with respect to the docking pose, but the pyDock energy description does. As a result, scoring ensembles with P_{best} or P_{mean} improve the success rates while the best representative structure for the ensembles is the rigid-body docking model before conformational sampling instead of the conformers.

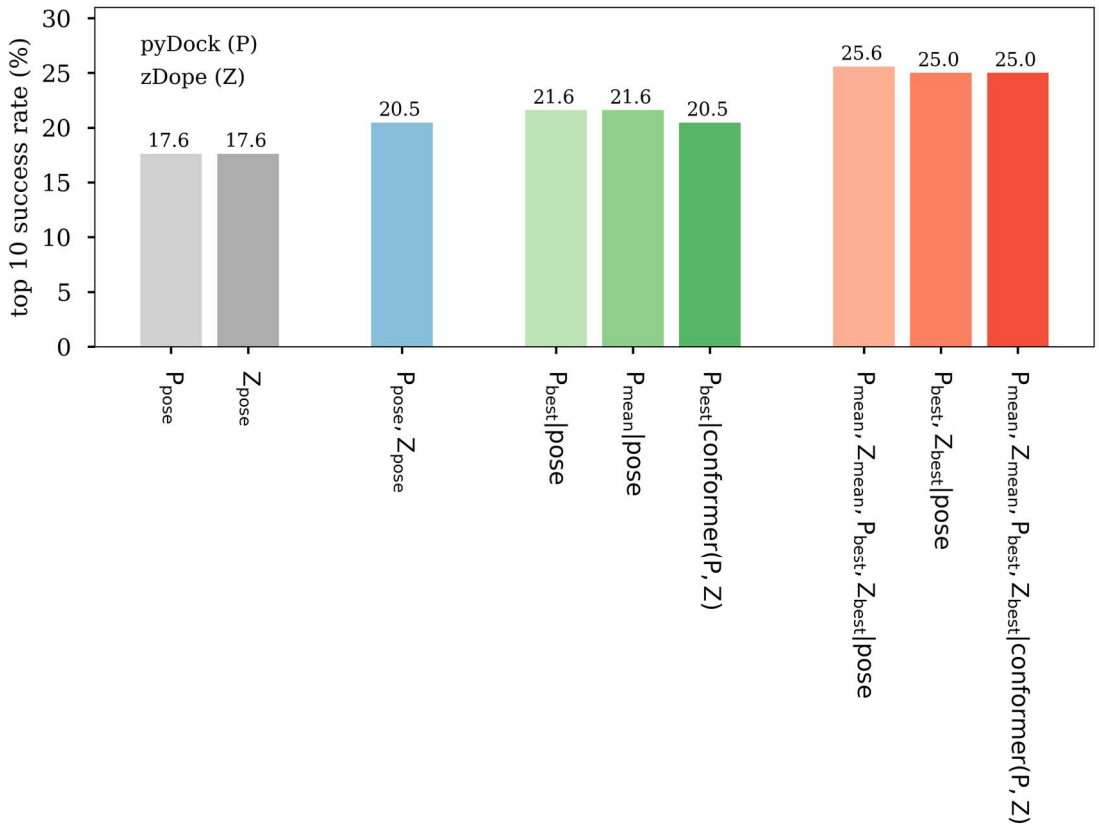

Figure 4.23. Docking success rates for the top 10 predicted models on the protein-protein docking benchmark 4 for the best-performing ensemble-based scoring schemes, based on independent scoring (green bars) or consensus scoring (red bars). For comparison, scoring of the original FTDock docking poses according to pyDock (P) and MODELLER zDope (Z) are shown (grey bars), as well as scoring of the original FTDock docking poses according to the consensus score of pyDock and MODELLER zDope (blue bar).

4.4.3 Combining clustering and flexible conformational ensembles by restricted molecular dynamics

We further investigate whether we could improve docking scoring performance by combining clustering and conformational ensembles.

There is an implicit hierarchy between the ensembles we have defined in the previous sections. Each cluster is a set of docking poses, and each pose generates an ensemble of 50 conformers. Therefore, for each cluster of poses, we can create a cluster ensemble of conformers by merging all the conformers generated from its docking poses. Then we can compute the conformers mean and best values of pyDock and zDope and use these values to score the clusters. In summary, the

approach is similar to that described in section 4.4.1 but using conformers instead of docking poses (see Figure 4.24).

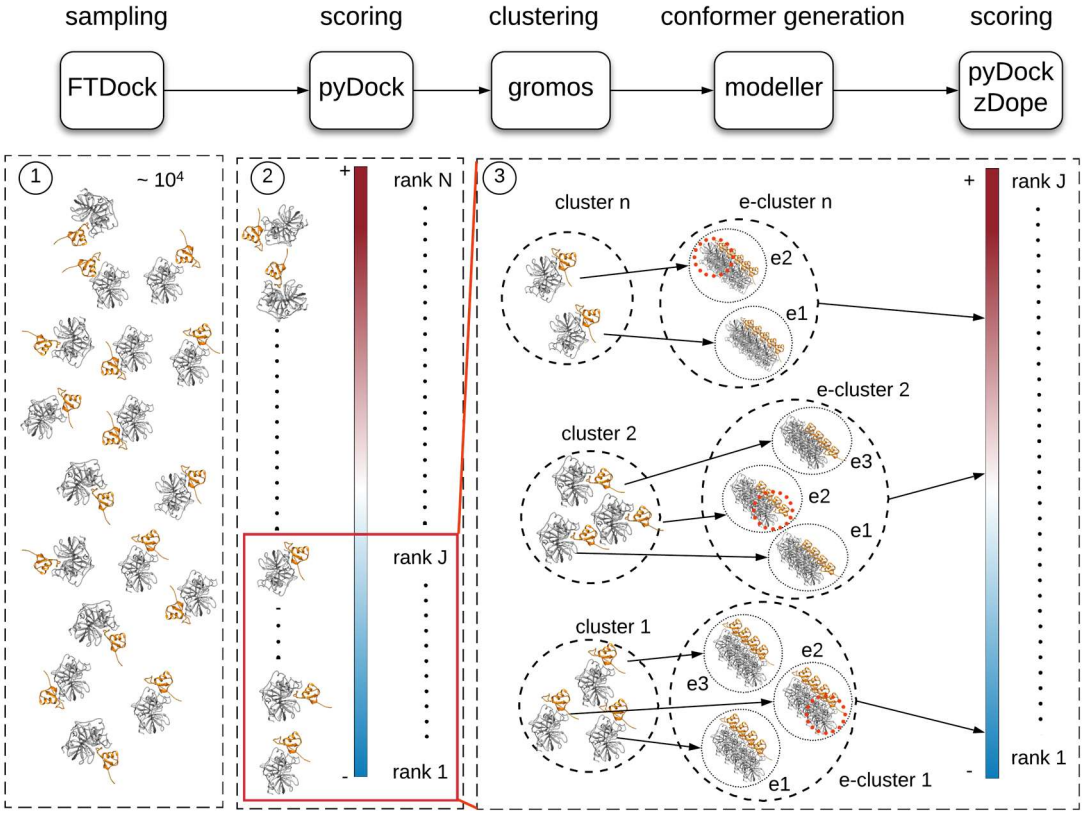

Figure 4.24. Schematic overview of the method. 1) Run FTDock to generate 10,000 docking poses. 2) Evaluate the docking poses with pyDock scoring function. Select the first 500 ranked docking poses. 3) Cluster the selected poses with gromos and a 10 Å cutoff. For each selected pose, generate an ensemble of 50 conformers with MODELLER and generate e-clusters (ensemble clusters). Rank e-clusters according to the 'independent' and consensus scores derived from the mean and best (minimum) values of pyDock energy and zDope per e-cluster. Select an e-cluster representative based on its pyDock energy and zDope score.

To score the clusters we used the same independent and consensus scores as the ones we described in the previous section, now applied to all the conformers of each cluster. As cluster representative we selected any of the following models:

- conformer within the cluster with best pyDock energy "conformer(P_{best})" or zDope score "conformer(Z_{best})",
- docking pose that generated the conformer within the cluster with best pyDock energy "pose(ensemble(P_{best}))" or zDope score "pose(ensemble(Z_{best}))",
- docking pose that generated the ensemble of conformers with better mean pyDock energy "pose(ensemble(P_{mean}))", or mean zDope score "pose(ensemble(Z_{mean}))",

- docking pose that generated the ensemble of conformers with better consensus score: "pose(ensemble(P_{mean},P_{best}))", "pose(ensemble(P_{mean},Z_{mean}))", "pose(ensemble(P_{best},Z_{best}))", "pose(ensemble(Z_{mean},Z_{best}))", "pose(ensemble(P_{mean},Z_{mean},P_{best},Z_{best}))".

This strategy slightly improved the previous approaches based separately on clustering or conformational ensembles, achieving a top 10 success rates of 28% when clusters were sorted according to "P_{mean},Z_{mean}" and the representative was the pose that generated the ensemble with the conformer with the best zDope score. This is around 10 percentage points higher than the top 10 success rate achieved by standard pyDock.

In Figure 4.25 we can see how the differences between the best 'independent' score "P_{best}" and the best consensus score "P_{mean},Z_{mean}" were smaller than in previous strategies. On the contrary, the method to choose the cluster representative seems to have a more significant impact on the final results than before. In general, those scoring schemes including pyDock energetic terms behaved better than zDope-based metrics for scoring the ensemble clusters, while the latter seemed to identify the cluster representative more successfully.

Even though the combination of clustering and conformational ensembles described above achieved good results, we must be aware of the important computational cost associated with it. For each docking pose, 50 models must be generated and evaluated. To reduce this cost, we devised a second protocol. After clustering the original top 500 docking poses, we sorted the clusters by the best pose zDope, which is the metric that achieved the best top 100 success rate. Then, we select the docking poses of the first 100 clusters for further processing. From each one of the selected poses, we generated 50 conformationally different models with MODELLER and repeated the procedure described above for this reduced set of conformers.

Figure 4.26 shows the results obtained with this variant. Although the top 10 success rates virtually did not improve, the total number of conformers to evaluate was reduced by around 40%.

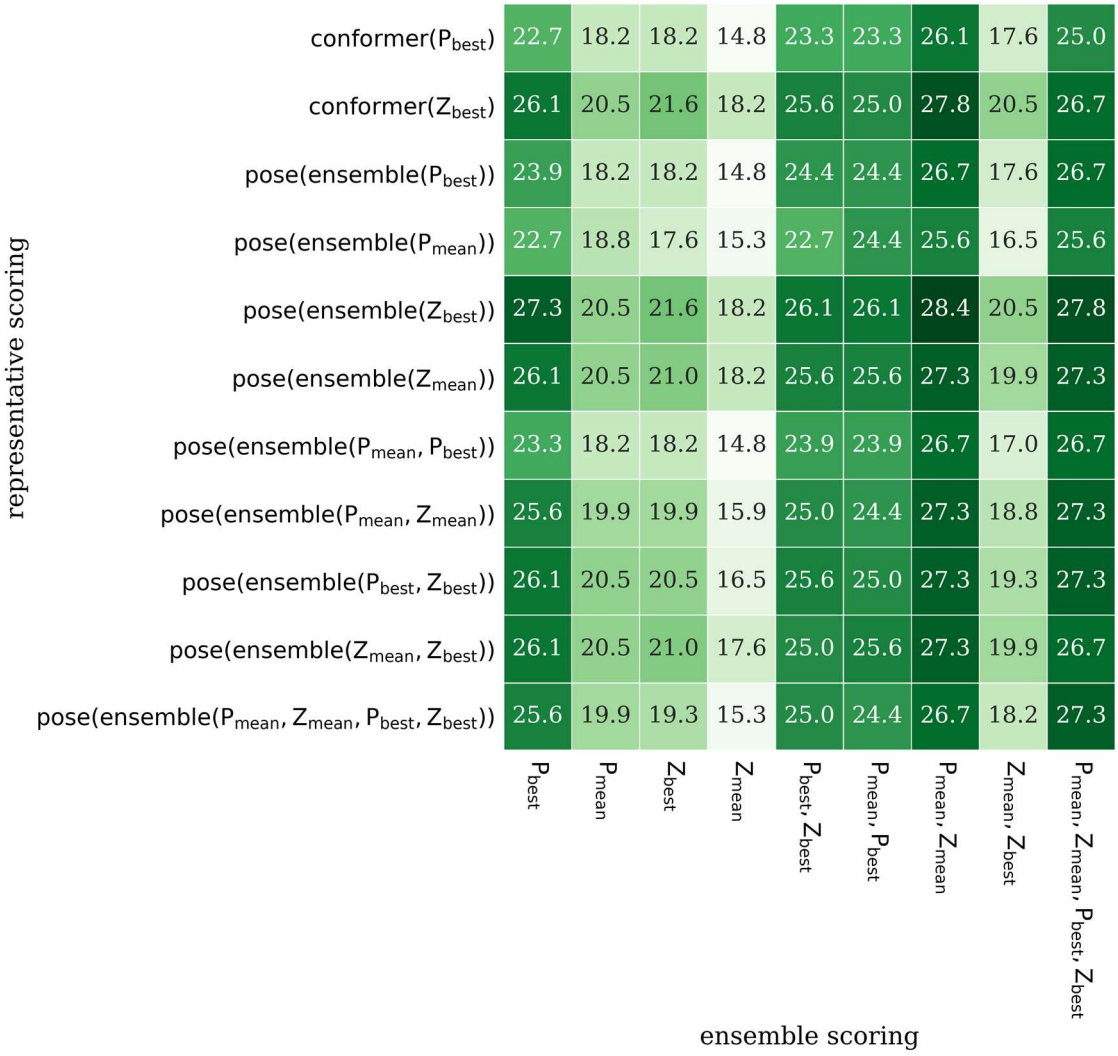

Figure 4.25. Docking success rates for the top 10 predicted models on the protein-protein docking benchmark 4 for several scoring schemes combining clustering and ensembles obtained with restricted molecular dynamics. For each scheme, the strategy used to score the clusters/ensembles (x-axis), and the strategy used to pick the cluster/ensemble representative (y-axis) are shown.

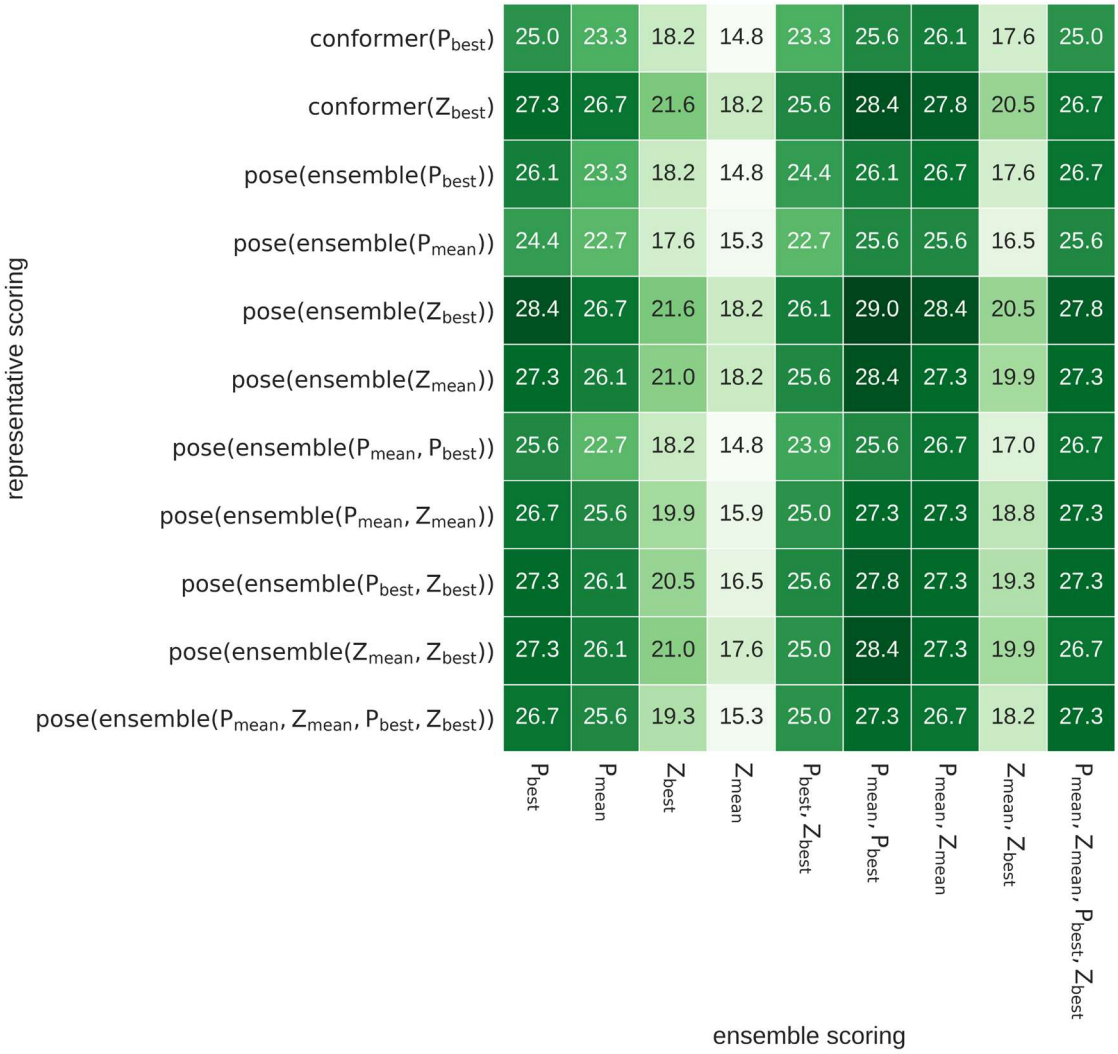

Figure 4.26. Docking success rates for the top 10 predicted models on the protein-protein docking benchmark 4 for several scoring schemes combining clustering and ensembles obtained with restricted molecular dynamics, after selecting docking poses of the first 100 clusters sorted by "best pose zDope". For each scheme, the strategy used to score the clusters/ensembles (x-axis), and the strategy used to pick the cluster/ensemble representative (y-axis) are shown.

4.5 Analysis of docking energy at the residue level

Using pyDock docking methodology, with the new scoring approaches developed here, we are able to explore the docking energy landscapes efficiently, and we can describe the different energy minima as ensembles of docking solutions. However, in some circumstances, we might want to have a more detailed description, at the level of residue or atoms, of the docking energy of the different states conforming the docking landscapes. In the following sections, we will first describe how we can partition pyDock docking energy at the residue level. Then we will use this partitioned energy to estimate changes in binding affinity upon mutation to alanine, i.e., as an *in-silico* alanine scanning mutagenesis predictor.

4.5.1 pyDock energy partition: description and examples

Next, we will explain the steps we have followed to partition the three energetic terms of pyDock scoring function: electrostatics, desolvation and van der Waals.

4.5.1.1 pyDock electrostatic term

We could rearrange equation Eq. 3.2 and express E_{ele} as a sum of the individual electrostatic interactions of the receptor atoms as

$$E_{ele} = \sum_{i}^{atom\_rec} E_{ele}^{i} = E_{ele}^{1} + E_{ele}^{2} + \cdots + E_{ele}^{atom\_rec-1} + E_{ele}^{atom\_rec} \quad \text{Eq. 4.3}$$

where the individual electrostatic energy for a receptor atom i is

$$E_{ele}^{i} = 332.0 \sum_{j}^{atom\_lig} \frac{q_i q_j}{4d_{ij}^2} \quad \text{Eq. 4.4}$$

Now, it is straightforward to regroup the terms and express the total electrostatic energy as a sum of the electrostatic energies of the residues

$$E_{ele} = \underbrace{E_{ele}^{1} + \cdots + E_{ele}^{k}}_{residue^{1}} + \cdots + \underbrace{E_{ele}^{l} + \cdots + E_{ele}^{atom\_rec}}_{residue^{res\_rec}} \quad \text{Eq. 4.5}$$

$$E_{ele} = \sum_{i}^{res\_rec} E_{ele}^{i} = E_{ele}^{1} + \cdots + E_{ele}^{res\_rec} \quad \text{Eq. 4.6}$$

where res_rec is the number of residues in the receptor.

In the same way, we could express the total electrostatic energy, E_{ele}, as a sum of the individual electrostatic interactions of the ligand atoms

$$E_{ele} = \sum_{i}^{atom\_lig} E_{ele}^{i} = E_{ele}^{1} + E_{ele}^{2} + \cdots + E_{ele}^{lig-1} + E_{ele}^{atom\_lig} \quad \text{Eq. 4.7}$$

or a sum of the individual electrostatic energies of the ligand residues

$$E_{ele} = \sum_{i}^{res\_lig} E_{ele}^{i} = E_{ele}^{1} + \cdots + E_{ele}^{res\_lig} \quad \text{Eq. 4.8}$$

4.5.1.2 pyDock van der Waals term

We can follow the same reasoning and express pyDock total van der Waals energy (see Eq. 3.6) as a sum of atomic interactions

$$E_{vdw} = \sum_{i}^{atom\_rec} E_{vdw}^{i} = E_{vdw}^{1} + E_{vdw}^{2} + \cdots + E_{vdw}^{atom\_rec-1} + E_{vdw}^{atom\_rec} \quad \text{Eq. 4.9}$$

with

$$E_{vdw}^{i} = \sum_{j}^{atom\_lig} \sqrt{e_i e_j}\left(\left(\frac{r_i + r_j}{d_{ij}}\right)^{12} - 2\left(\frac{r_i + r_j}{d_{ij}}\right)^{6}\right) \quad \text{Eq. 4.10}$$

or a sum of residue interactions

$$E_{vdw} = \underbrace{E_{vdw}^{1} + \cdots + E_{vdw}^{k}}_{residue^{1}} + \cdots + \underbrace{E_{vdw}^{l} + \cdots + E_{vdw}^{atom\_rec}}_{residue^{res\_rec}} \quad \text{Eq. 4.11}$$

$$E_{vdw} = \sum_{i}^{res\_rec} E_{vdw}^{i} = E_{vdw}^{1} + \cdots + E_{vdw}^{res\_rec} \quad \text{Eq. 4.12}$$

Moreover, the same argument applies to the atoms and residues of the ligand.

4.5.1.3 pyDock desolvation term

Partitioning desolvation energy may be trickier than partitioning electrostatics or van der Waals terms. We have previously seen that the electrostatics and var der Waals potentials of a given atom depend on an intrinsic property of the atom, namely, the charge. We will see that the desolvation energy we have defined is not only a function of the atom types upon which it is calculated but also of the *environment* of these atoms. This situation allows devising two different approaches to partition the desolvation energy.

4.5.1.3.1 Residue protein desolvation

Equation Eq. 3.3 already provides the total desolvation energy as a function of the individual desolvation energies of the atoms of receptor and ligand.

From equations Eq. 3.3 and Eq. 3.5 we can define the atom protein desolvation (atomProtDesolv) of a given atom i as

$$atomProtDesolv_i = ADP_i \cdot \Delta ASA_i = ADP_i \cdot BSA_i \qquad \text{Eq. 4.13}$$

From here, it is trivial to express the total desolvation energy as a sum of the residue protein desolvations (resProtDesolv) energies by grouping the individual atomProtDesolv.

$$\begin{aligned} E_{desolv} = & \underbrace{atomProtDesolv_1 + \cdots + atomProtDesolv_k}_{resProtDesolv_{rec1}} + \cdots + \\ & + \underbrace{atomProtDesolv_l + \cdots + atomProtDesolv_{atom\_rec}}_{resProtDesol\ _{res\_rec}} + \\ & + \underbrace{atomProtDesolv_1 + \cdots + atomProtDesolv_m}_{resProtDesolv_{lig1}} + \cdots + \\ & + \underbrace{atomProtDesolv_n + \cdots + atomProtDesolv_{ato\ \_lig}}_{resProtDesolv_{res\_lig}} = \\ = & \sum_{i}^{res\_rec} E^i_{desolv} + \sum_{j}^{res\_lig} E^j_{desolv} \end{aligned} \qquad \text{Eq. 4.14}$$

It is worth noting that in Eq. 4.14 the total desolvation energy is the sum of the partial protein desolvation energies of receptor and ligand. This is different to what we saw when partitioning electrostatics or van der Waals, where the total energy could be computed by summing up the residue energies of a single subunit, either the receptor or the ligand (see equations Eq. 4.6, Eq. 4.8 and Eq. 4.12).

4.5.1.3.2 Residue binding desolvation

We can devise a second approach for partitioning pyDock desolvation energy. As we mentioned before, the first approach described in the last subsection could reproduce pyDock desolvation energy values separately for receptor or ligand residues but did not include specific contributions from the partner residues. As we have previously mentioned, the BSA variable we defined in Eq. 3.4 is a function of the environment where atoms are located. We can clearly see this in the simple system of two atoms represented in Figure 4.1. The BSA of atoms is a function of the distance between them. While the accessible surface area of atom 1 is reduced by the proximity of atom 2, atom 1 also alters the accessible surface area of atom 2 and therefore will contribute to BSA-based desolvation energy "caused" by atom 1. Under this point of view, BSA-desolvation of a given atom would depend not only on its own type and BSA but also on that of their neighbor atoms (environment). Therefore, for a given atom we should account for its own BSA but also for the changes in BSA that its presence has in its environment,

what we can call the induced BSA of the atom in its environment. With this assumption, we could express the atom binding desolvation of an atom i as

$$atomBindDesolv_i = \underbrace{ADP_i \cdot BSA_i}_{individual} + \underbrace{\sum_j^{closest} ADP_i^j \cdot BSA_i^j}_{induced} \qquad \text{Eq. 4.15}$$

If we recall Eq. 4.13, we can rewrite this equation as

$$atomBindDesolv_i = \underbrace{atomProtDesolv_i}_{individual} + \underbrace{\sum_j^{closest} atomProtDesolv_i^j}_{induced} \qquad \text{Eq. 4.16}$$

Therefore, we can express the atomBindDesolv of a given atom i as the sum of its atomProtDesolv and the atomProtDesolv energies of the atoms of the complementary subunit that has atom i as closest atom.

Then, the total desolvation energy could be expressed as the sum over all the atomBindDesolv energies of the receptor or ligand atoms, because either set of values for receptor or ligand atoms already includes the contribution to desolvation from the atoms of the complementary subunit.

$$E_{desolv} = \sum_i^{atom\_rec} atomBindDesolv_i = \sum_j^{atom\_lig} atomBindDesolv_j \qquad \text{Eq. 4.17}$$

We could also rearrange the terms, as we have done previously, to express the total desolvation energy as the sum of the resBindDesolv energies of the residues of receptor or ligand

$$\begin{aligned} E_{desolv} = & \underbrace{atomBindDesolv_1 + \cdots + atomBindDesolv_k}_{resBindDesolv_{rec1}} + \cdots + \\ & + \underbrace{atomBindDesolv_l + \cdots + atomBindDesolv_{atom\_rec}}_{resBindDeso\ \ res\_rec} = \\ & = \sum_i^{res\_rec} resBindDesolv_i \end{aligned} \qquad \text{Eq. 4.18}$$

$$E_{desolv} = \underbrace{atomBindDesolv_1 + \cdots + atomBindDesolv_m}_{resBindDesolv_{lig1}} + \cdots + \\ + \underbrace{atomBindDesolv_n + \cdots + atomBindDesolv_{atom\_lig}}_{resBindDesolv_{res\_lig}} = \\ = \sum_{j}^{res\_lig} resBindDesolv_j \qquad \text{Eq. 4.19}$$

The definition of resBindDesolv allows us to express the total desolvation energy as the sum of the residue binding energies of either the receptor or the ligand subunits, as we did when partitioning electrostatics or van der Waals (see equations Eq. 1.6, Eq. 1.8 and Eq. 1.12).

By partitioning pyDock energy, we can estimate the contribution of each residue to the total energy. In the following sections, we will show that this new feature has applications in the identification of critical residues in the formation of complexes, i.e., hot-spots residues, or in the prediction of the effects of mutations in protein binding.

4.5.2 pyDock energy partition can identify important residues for binding affinity

We have extensively applied pyDock residue energy to identify key residues in protein-protein interactions. As an example, Figure 4.27 shows MEK1 and BRAF interface characterization by pyDock residue energy and *in-silico* alanine scanning. *In-silico* alanine scanning is considered an accurate technique. It combines molecular dynamics simulations and binding energy calculations with the MM-GBSA method (Miller et al. 2012)(see section 3.4.1). Both methods require the structure of the complex, in this case, protein complex with PDB ID 4MNE, to make their predictions. The amount of computational resources required by these methods varies a lot. At one end of the spectrum, pyDock residue energies can be computed in seconds. At the other end, *in-silico* alanine scanning may require many hours, even days, to obtain the results (see section 3.4.1). We can note in Figure 4.27 that pyDock residue energy and *in-silico* alanine scanning have similar results and agree with experimentally determined crystal structures. This is remarkable since the amount of resources required to compute pyDock residue energy are negligible compared to those required by *in-silico* alanine scanning.

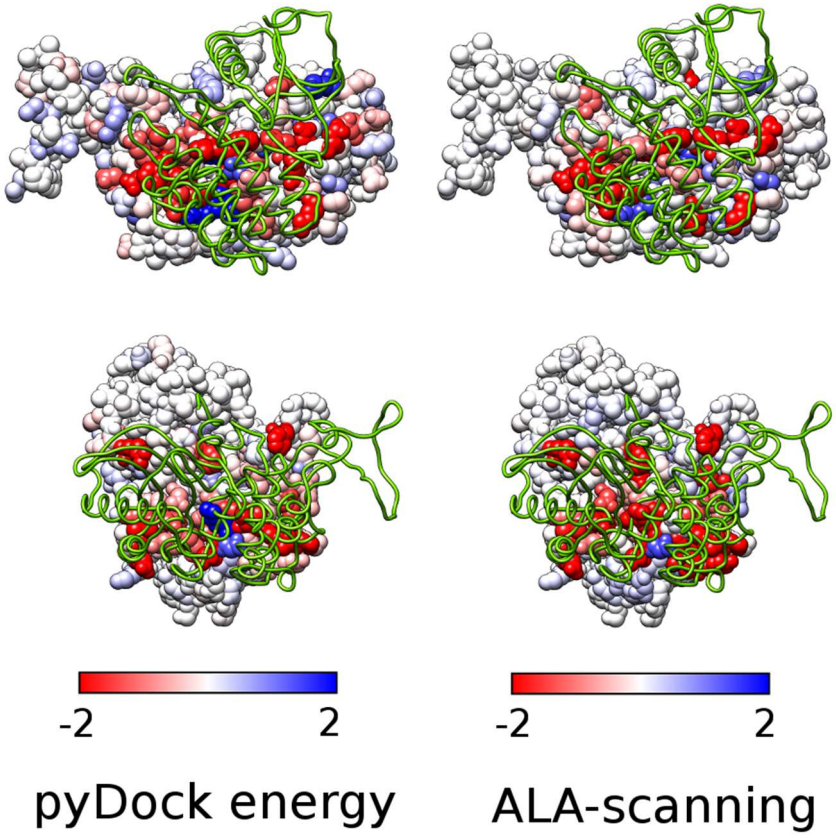

Figure 4.27. MEK1-BRAF interface characterization. MEK1 and BRAF interface characterization (first and second line, respectively) using different computational techniques: pyDock binding energy decomposition and binding free energy change (ΔΔG) estimated by *in-silico* alanine scanning.

4.5.3 Estimating side-chain contribution to binding affinity

As mentioned in the introduction, alanine scanning mutagenesis is an experimental technique widely used to determine the energetic contribution of the side chain of a specific residue to the binding energy of a given protein complex. The method aims to determine the binding affinity differences between the wild-type and alanine mutated complexes. As this is a time-consuming method many alternative, computational approaches have been developed.

In the previous section, we saw how we could decompose the docking energy of a protein complex as the sum of the energies of its atoms. Once we have this atomic-level decomposition, it is straightforward to estimate the energetic side-chain contributions of all residues of the complex. Namely, the side-chain contribution for a given residue i is

$$E^{i}_{pyDock\_sidechain} = \sum_{j}^{sidechain} E^{j}_{pyDock} \qquad \text{Eq. 4.20}$$

where the summation extends to the side-chain atoms of residue *i*. It is important to remark that to mimic alanine scanning mutagenesis we do exclude Cβ carbon, and its bonded hydrogens, from this summation.

Finally, we can easily relate the docking energy of the side-chain atoms with binding affinity changes upon mutation to alanine

$$E^{i}_{pyDock\_sidechain} = \sum_{j}^{sidechain} E^{j}_{pyDock} = -\Delta\Delta G(residue^{i}) \qquad \text{Eq. 4.21}$$

4.5.4 pyDock side-chain energy can estimate binding affinity changes in mutations to alanine

We wanted to test whether pyDock sidechain energy was a good proxy to estimate experimental changes in binding affinity upon mutation to alanine. To this end, we compared the *in-silico* predictions with experimental data contained in SKEMPI database (Moal and Fernández-Recio 2012b). First, we selected all the single mutations to alanine stored in SKEMPI. Then, we discarded mutations with i) undefined wild-type or mutant binding affinity values, ii) no detectable binding or unfolded mutants or iii) undefined experimental temperature. We also removed from the data set all mutations with Gly as wild-type residue since our approach cannot model the "missing" side-chain in a Gly to Ala mutation. We evaluated the possibility of removing mutations with Pro as wild-type residue. Even though Pro and Ala residues have both a carbon beta (Cβ), Pro residues have a distinctive cyclic side-chain that confers specific properties like high conformational rigidity compared to other amino acids. Finally, we decided to keep them assuming that their predictions should be considered with caution. We ended up with a total of 635 experimental mutations. As Figure 4.28 reflects, the wild-type residues were not uniformly distributed within the data set: mutations

from Arg, Glu or Lys were overrepresented with more than 60 items, while mutations from Cys and Met were underrepresented with less than 10.

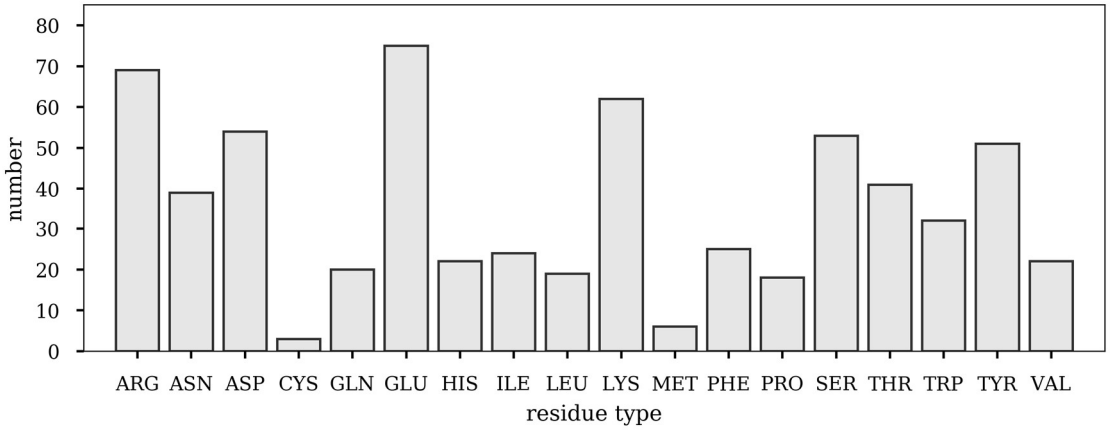

Figure 4.28. Number of experimental mutations in the data set grouped by wild-type residue.

Based on the wild-type complex structures, we evaluated the different pyDock energetic terms for the side-chain atoms of the experimentally mutated residues. In order to compare our predictions with a reference *in-silico* method, we also computed the binding affinity changes upon mutation with FoldX (Guerois, Nielsen, and Serrano 2002).

From the analysis of the results showed in Figure 4.29 we can conclude that, among the pyDock energy terms, the van der Waals component achieved the highest correlation with experimental data, with a Pearson correlation score of 0.52. On the other side was the desolvation term with a modest 0.07. We obtained this correlation value by partitioning the desolvation term following the protein desolvation approach described in section 4.5.1.3.1. We also tested the performance of the binding desolvation method described in section 4.5.1.3.2, but we obtained similar results. The discrete correlation of pyDock desolvation term had already been reported by Pallara *et al.* (Pallara et al. 2013) where it was argued that although desolvation could play a key role in rigid-body docking, maybe as a means of compensating for inaccuracies introduced by electrostatic and van der Waals contributions, it was not so crucial for binding affinity predictions computed from complex structures. Based on these results, we devised a new descriptor consisting in the sum of the electrostatics and van der Waals side-chain contributions that from now on we name pyDockSC$_{ele+vdw}$. This new descriptor outperformed the prediction capabilities of the rest of scoring terms, showing a Pearson correlation of 0.57 between predicted and experimental values. For comparison, the correlation between FoldX predictions and the experimental

values for the same set of mutations to alanine was 0.27 (see Figure 4.29 and Figure 4.30).

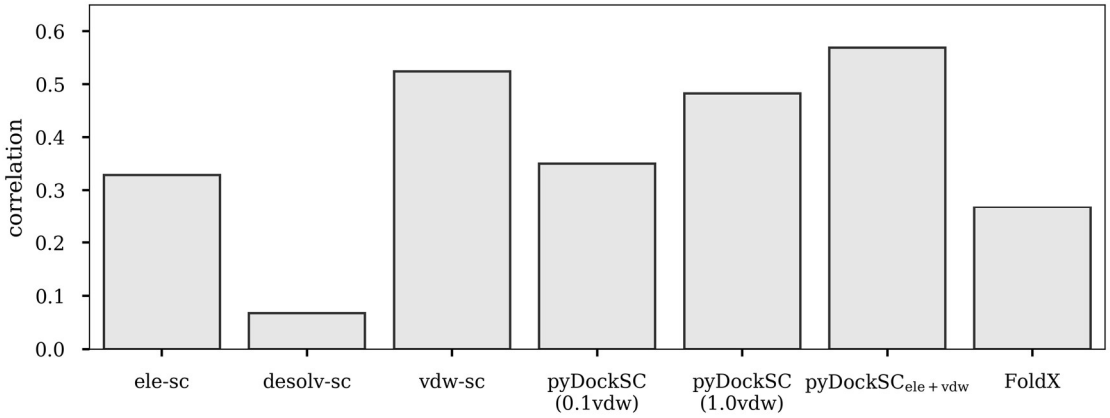

Figure 4.29. Correlation between experimental binding affinity changes upon mutation to alanine and the predicted values obtained by different scoring functions: pyDock electrostatics of side-chain atoms (ele-sc), pyDock desolvation of side-chain atoms (desolv-sc), pyDock van der Waals of side-chain atoms (vdw-sc), pyDock energy of side-chain atoms with van der Waals weight set to 0.1 (pyDockSC (0.1vdw)), pyDock energy of side-chain atoms with van der Waals weight set to 1.0 (pyDockSC (1.0vdw)), pyDock electrostatics plus pyDock van der Waals of side-chain atoms (pyDockSC$_{ele+vdw}$) and FoldX.

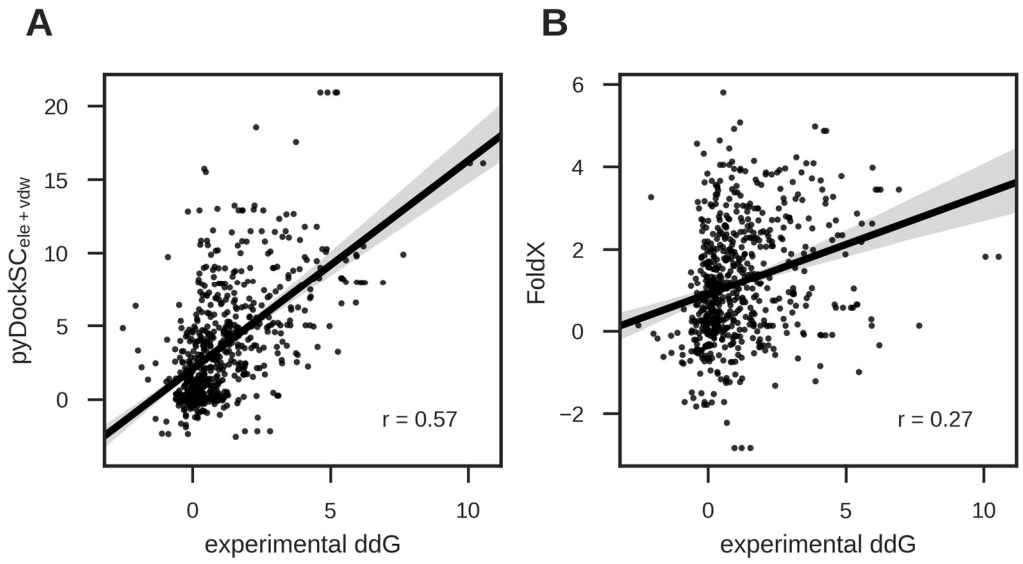

Figure 4.30. Comparison between experimental and predicted ΔΔG for mutations to alanine. Values predicted by: A) pyDockSC$_{ele+vdw}$,B) FoldX. Data from SKEMPI.

This correlation is not uniform for all residue types. As Figure 4.31 shows, Arg, Asp, His, Lys, and Thr showed the highest correlation between predicted and experimental values, with Pearson correlation coefficients of 0.63, 0.65, 0.63, 0.74 and 0.75, respectively. This suggests that our scoring function reasonably described alanine mutations of charged residues. The correlation was negative for

Met and Pro. The low correlation found for proline is not surprising, given its unique properties. In the case of methionine, we should consider the small number of methionine mutations available in the data set and be cautious when interpreting the results.

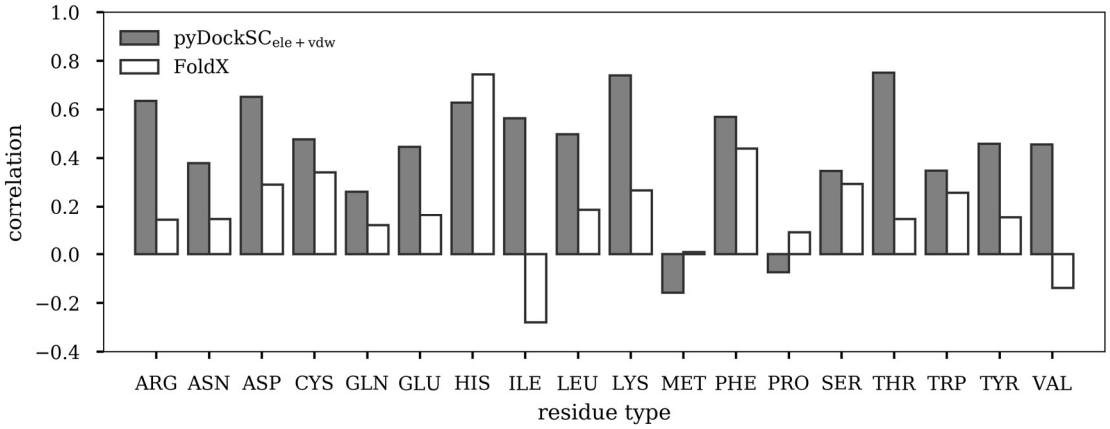

Figure 4.31. Pearson correlation between experimental ΔΔG for single mutations to alanine from SKEMPI and the predicted values obtained by combining pyDock electrostatics and van der Waals energies for side-chain atoms, pyDockSC$_{ele+vdw}$ (dark-grey), for the different types of mutated residues. For comparison, the correlation with the predicted values obtained by FoldX (white) is shown.

We have explored whether the technique used to obtain the experimental ΔΔG values has some effect on the resulting values, as it was previously found for binding affinity estimations (Kastritis and Bonvin 2010).

Figure 4.32 shows that some experimental techniques like SFPF and RA showed higher correlations regardless of the *in-silico* technique applied, i.e., pyDockSC$_{ele+vdw}$ or FoldX. Others, like ELFA, achieved discrete correlations of 0.27 with pyDockSC$_{ele+vdw}$ and 0.08 with FoldX. We should also note the particular cases of EMSA and IAGE that got acceptable correlations with pyDockSC$_{ele+vdw}$, but negative correlations with FoldX (see Table 4.8 to identify the experimental technique associated to each acronym).

Acronym	Experimental technique
ELFA	Enzyme-linked functional assay
EMSA	Electrophoretic mobility shift
FL	Fluorescence
IAFL	Fluorescence inhibition assay
IAGE	Agarose gel inhibition assay
IARA	Radioligand inhibition assay
IASP	Spectroscopic inhibition assay
ITC	Isothermal titration calorimetry
SP	Other spectroscopic methods
RA	Radioactive ligand binding
SFFL	Stopped-flow fluorescence
SFPF	Stopped-flow spectrophotometry
SPR	Surface plasmon resonance

Table 4.8. Experimental techniques.

Additionally, we can also analyze the slope values of the linear regression models between $\Delta\Delta G$ and $pyDockSC_{ele+vdw}$ for the different experimental techniques (see Figure 4.33). While the slope value is between 1 and 1.5 for 8 experimental techniques, it is between 2 and 3 for the remaining 4 techniques, which are found among the highest ones regarding correlation. In summary, this analysis indicates that not all experimental techniques may have the same degree of reliability and that they may not even share the same scale. These are important points that should be considered in the development and evaluation of *in-silico* predictors of binding affinity changes upon mutation.

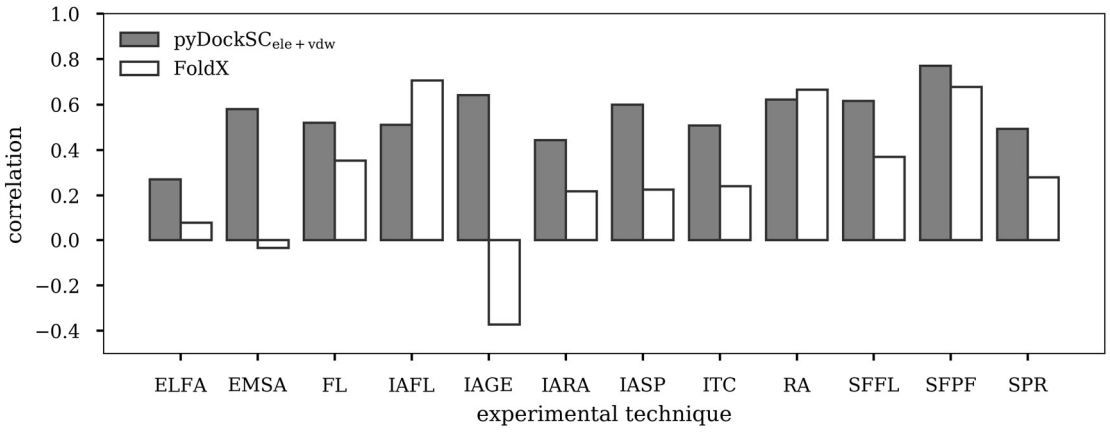

Figure 4.32. Pearson correlation between experimental ΔΔG for single mutations to alanine and the values predicted by pyDockSC$_{ele+vdw}$ (dark-grey) or FoldX (white), grouped by the experimental technique used.

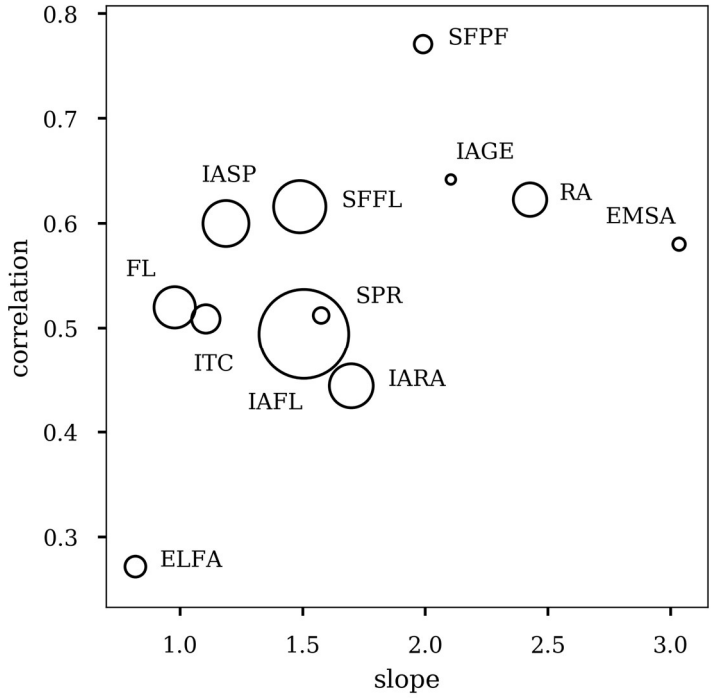

Figure 4.33. Slope and Pearson correlation coefficient of the linear regression model between pyDockSC$_{ele+vdw}$ and experimental ΔΔG for the different experimental techniques applied to SKEMPI single mutations to alanine. The size of the experimental technique circles is proportional to the number of mutations analyzed.

It is also interesting to analyze the predictive capabilities of our function case by case. Figure 4.34 shows the correlation plot between pyDockSC$_{ele+vdw}$ predictions and experimental values of changes in binding affinity upon mutation to alanine for seven complexes with the largest number of entries of single mutations to alanine in SKEMPI database.

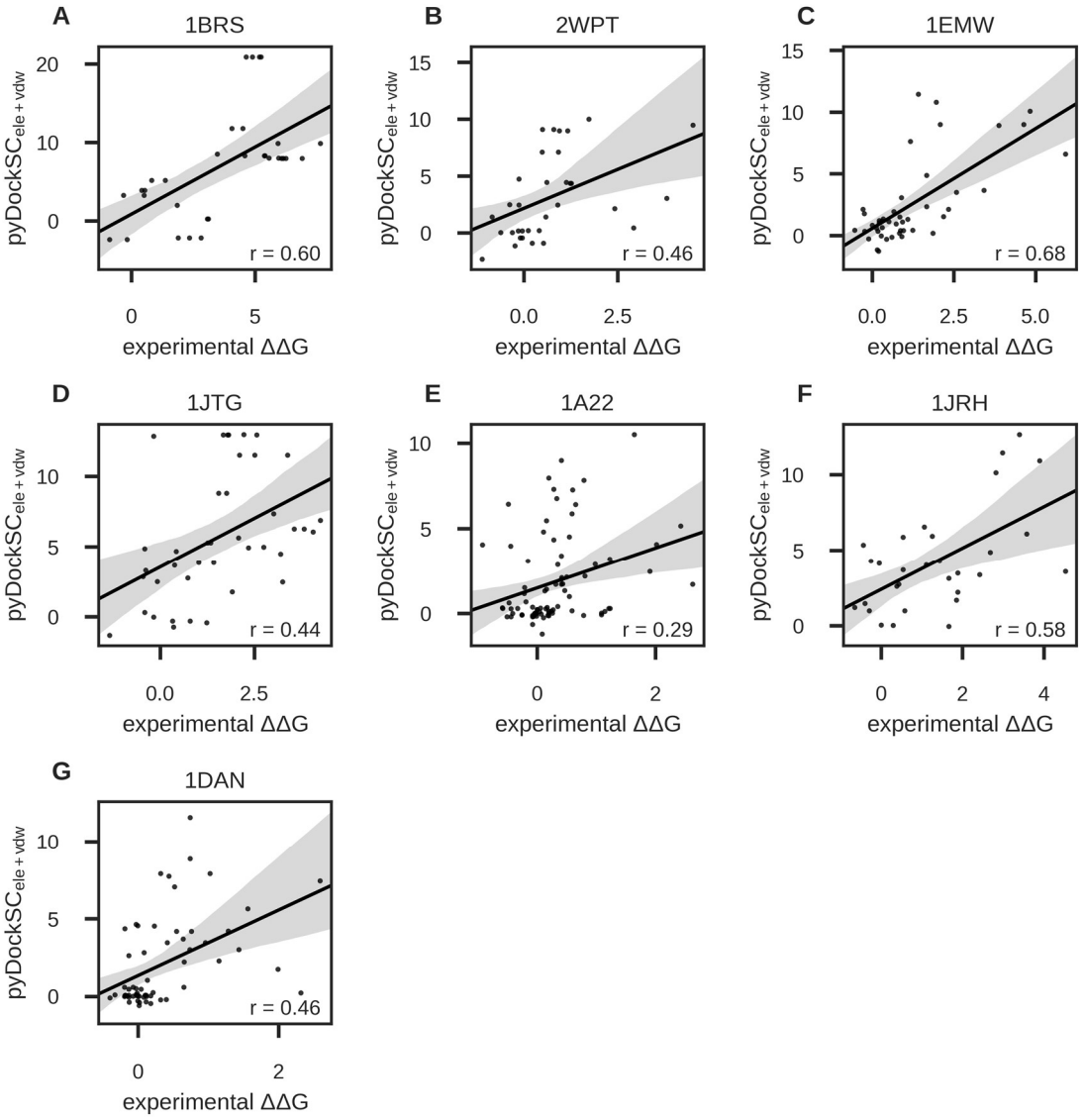

Figure 4.34. Experimental ddG *vs.* pyDockSC$_{ele+vdw}$ prediction for the seven cases with the largest number of experimental binding affinity values for mutations to alanine in SKEMPI. The cases shown correspond to the complexes with the following PDB IDs: A) 1BRS, B) 2WPT, C) 1EMW, D) 1JTG, E) 1A22, F) 1JRH and G) 1DAN.

The correlation values vary a lot depending on the case. This could be due to uncertainties within our model, but it might also be explained by the intrinsic nature of the protein binding mechanism, highly dependent on the particular system (Kastritis and Bonvin 2010). These results invite us to reflect on whether it makes sense to evaluate ΔΔG prediction methods using global tests that include many cases and a substantial heterogeneity in experimental techniques, different binding mechanisms and types of complexes, or if on the contrary, predictive models should be evaluated case by case considering more carefully all possible experimental and/or biological conditions.

4.5.5 Modeling binding affinity changes in mutations to other residues

4.5.5.1 Prediction of binding affinity changes upon mutation: an initial exploration in CAPRI

In recent years, the protein-protein docking community is increasingly interested in applying existing docking methods to understand the thermodynamics of protein-protein interactions, and as a consequence to predict protein-protein binding affinity. As a good example of this, in CAPRI round 26, participants were asked to predict the effects of mutations on protein-protein interaction affinity.

In this CAPRI round, the experimental data set was generated by Baker's group as part of their research for developing influenza inhibitors (Whitehead et al. 2012). Starting from two de-novo designed binders (HB36 and HB80) to influenza hemagglutinin (HA), single point mutant libraries were created, corresponding to all 20 amino acids in 53 and 45 different positions of HB36.4 and HB80.3, respectively. After using yeast display and fluorescence-activated cell sorting (FACS), the pre-sort and enriched libraries were sequenced with an Illumina GA-II sequencer. The enrichment value, an estimation of the effect of a given mutation on binding, was computed for each mutation as the base 2 logarithm of the ratio of the amino acid frequencies in the enriched library to that of the unenriched library.

Two prediction rounds were defined. In the first one, participants were provided with the structures of the binders HB36 (PDB ID 3R2X) and HB80 (PDB ID 4EEF), the positions at which mutations had been made and a description of how the experimental data had been generated. In the second round, participants were also provided with experimental data of the enrichment values for half of the mutations, corresponding to 9 randomly selected amino acids at each mutated position plus the wild-type residue. With this information, participants had to classify each mutation as beneficial, neutral or deleterious and rank them according to a score ranging from 0 (most deleterious) to 1 (most beneficial). Experimental data was also classified as beneficial (enrichment ratio $> +2$), deleterious (enrichment ratio $<$ -2) or neutral (-2 $\leq$ enrichment ratio $\leq +2$).

The quality of the predictions significantly varies along the 22 groups that participated in the experiment. The best-performing groups achieved a precision of around 0.1 with a coverage between 0.25 and 0.40. Clearly, these results show that there is considerable room for improvement. Nevertheless, they were three times better than what could be expected by random. Deleterious mutations were

easier to predict than beneficial ones. It was also observed that the affinity changes were predicted better for apolar to polar mutations than for polar to polar and polar to apolar mutations, which were overestimated. This could suggest the incorrect treatment of electrostatics at the interfaces. Most of the top performing groups optimized the side-chains and backbones of mutant structures, including the residues surrounding the mutated amino-acid.

Interestingly, in the first predicted round, three different approaches were applied by the top 3 performing groups: machine learning techniques (Bates group and our own group), atom-level energy functions (Weng group), and coarse-grained models (Dehouck group). However, in the second predicted round, machine learning approaches achieved the best results, being those that better exploit the additional experimental information. Regardless of the overall strategy applied, packing metrics seemed to be the best performing features. For example, the highest performing groups included terms such as Lennard-Jones potentials, statistical contact and distance scores like OPUS_PSP group potential or Tobi coarse-grained potential. Additionally, most of the top performing groups also included electrostatics and desolvation terms in their scoring functions.

Our group was one of the top performing groups in both rounds. Our approach was purely based on machine learning. For both rounds, mutant structures were generated with FoldX (Guerois, Nielsen, and Serrano 2002). In the first round, a database of 930 experimental changes in binding affinity upon mutation, later made publicly available as SKEMPI (Moal and Fernández-Recio 2012b), was used to train six regression models using 85 descriptors including contact and distance potentials, H-bonding potentials, desolvation models, entropy and folding stability scores and terms of several scoring functions like FoldX (Guerois, Nielsen, and Serrano 2002), PyRosetta (Chaudhury, Lyskov, and Gray 2010), FireDock (Andrusier, Nussinov, and Wolfson 2007), PyDock (T. M.-K. Cheng, Blundell, and Fernandez-Recio 2007), SIPPER (Pons et al. 2011), CHARMM (Brooks et al. 2009) and NIP/NSC (Mitra and Pal 2010). The regression predictions and the descriptors were subsequently used to train five classifiers: a random forest, a decision table construction algorithm, a Bayesian net, logistic regression and an alternating decision tree. The final predictions were computed as the mean confidence values of these five classifiers applied to the mutations of HB36 and HB80. The approach in the second round was similar. This time the regression predictions and the descriptors were used as features to train log2 ratio models. Several polynomial descriptors including the residue type of the mutated amino acid and the physical and chemical characterization of the wild-type and mutant amino acids were added to the features set. Three learners based on random forest

(Matlab RFreg), decision trees (WEKA M5Rules) and linear regression (WEKE Wlinreg) were first trained over the set of mutations with enrichment values provided by the organizers and then applied to the mutations with unknown log2 ratios. For each mutation, the predicted mutation type was computed as the mean prediction value of the three learners with two thresholds set to 0.0 and -1.5 to define beneficial (50) and neutral (203) mutations, respectively. The remaining 727 mutations were classified as deleterious.

4.5.5.2 Use of flexible conformational ensembles

In a previous section, we have described a new approach to estimate the binding affinity changes upon mutation to alanine by partitioning pyDock docking energy. As it cannot model the mutation at the atomic level, this method cannot compute $\Delta\Delta G$ of mutations to residues other than alanine, or mutations from glycine to alanine. In this section, we will present our new approach to try to overcome these limitations. The method is based on pyDock scoring function to estimate binding energies of complexes and uses MODELLER to model mutant structures. For a given mutation, the method applies restricted molecular dynamics with MODELLER and the algorithm described in Feyfant *et al.* (Feyfant, Sali, and Fiser 2007) to generate an ensemble of 144 mutant models. Each of these models is evaluated with pyDock scoring function and MODELLER zDope. Then, the mean pyDock energy of the 144-mutant ensemble is computed. The $\Delta\Delta G$ value is estimated from the difference between the mean ensemble pyDock energy and the pyDock energy of the wild-type complex. We explored the possibility of computing the energy of the wild-type complex from an ensemble of conformers as we did with the mutant, but we found the results were better if we consider the native wild-type complex structure instead. We tested the protocol by comparing the *in-silico* predictions with experimental data contained in the SKEMPI database. We selected all the single mutations stored in SKEMPI. Then, we discarded mutations with i) undefined wild-type or mutant binding affinity, ii) no detectable binding or unfolded mutants or iii) undefined experimental temperature, iv) entries associated to the trypsin-BPTI complex whose binding affinity data have been questioned (Krowarsch et al. 1999). We evaluate 1416 mutations in total. The global Pearson correlation coefficient between the predicted and the experimental $\Delta\Delta G$s was 0.38 (see Figure 4.35). If we compare the predicted and the experimental values for the subset of mutations to alanine, i.e., the mutations that we could estimate with pyDock$SC_{ele+vdw}$, the Pearson correlation coefficient raises to 0.46 (see Figure 4.36). This value improves to 0.53 if we drop the desolvation term from the scoring function leaving the electrostatics and van der Waals terms alone. In any case, both results are below the correlation

values obtained with pyDock$SC_{ele+vdw}$. One possibility is that MODELLER does not correctly model the structural changes introduced by the mutations. Then, we could test different modeling methods and check if their performance improves that of MODELLER. The other possibility is that MODELLER structures are accurate enough but pyDock scoring function does not evaluate them properly. In that case, including additional, subtler energetic terms like hydrogen bonding, entropic contributions, etc. could be the way to go to benefit from the structural improvements.

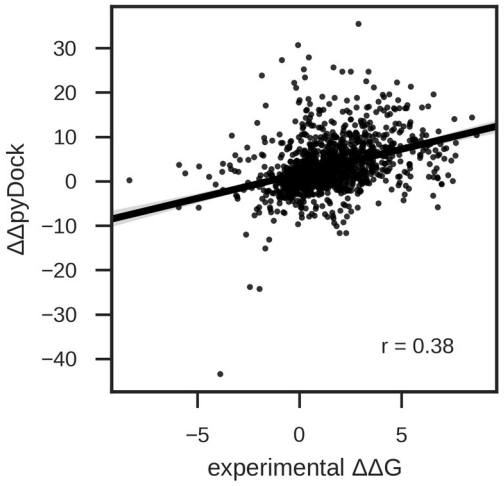

Figure 4.35. Correlation plot between experimental ΔΔG (x-axis) and ΔΔpyDock (y-axis) for 1416 mutations from SKEMPI database. PyDock energy was computed with van der Waals weight set to 1.0.

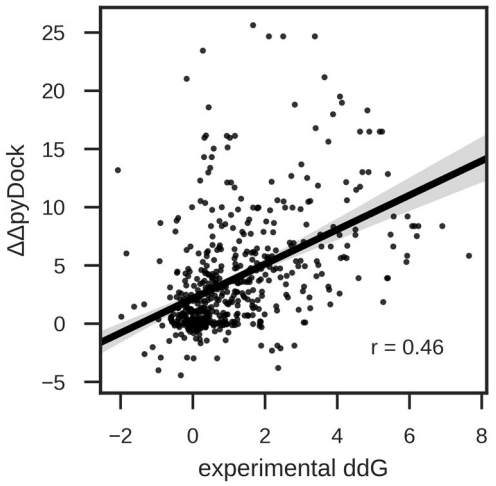

Figure 4.36. Correlation plot between experimental ΔΔG (x-axis) and ΔΔpyDock (y-axis) for the subset of single mutations to alanine in SKEMPI. PyDock energy was computed with van der Waals weight set to 1.0.

Figure 4.37 represents the Pearson correlation coefficients between the different scoring function predictions and experimental ΔΔG for single mutations to alanine, discriminating by wild-type residue type. We can check that, in general, pyDock based scoring functions behave better than FoldX. This is especially clear when the wild-type residue is isoleucine, threonine, tyrosine, and valine. The exception are mutations from methionine to alanine, for which pyDock based scoring functions obtained correlations close to 0, while FoldX predictions had a correlation of 0.61. However, we must keep in mind that the number of mutations from methionine to alanine in the dataset was small. Therefore, we should probably extend the analysis to more cases to confirm this observation.

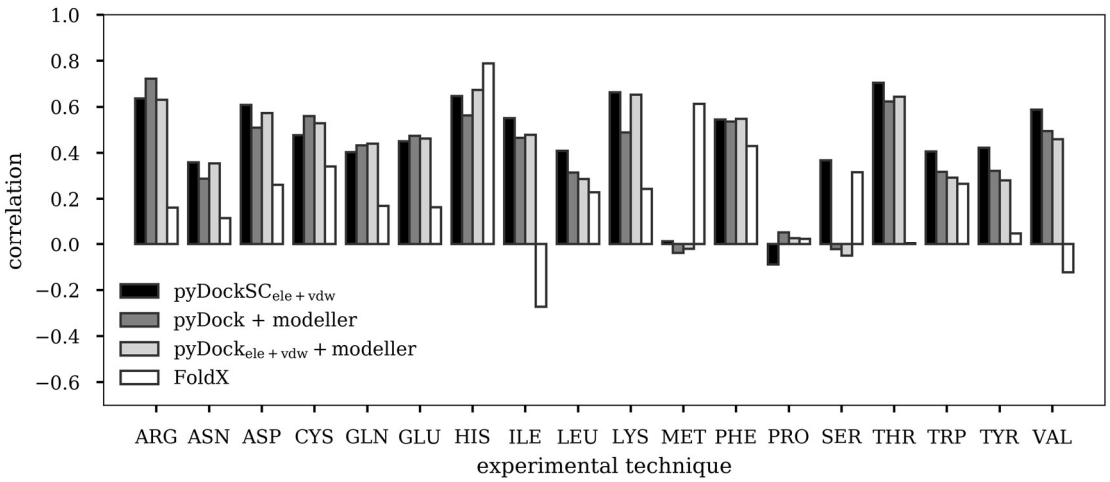

Figure 4.37. Pearson correlation coefficients between experimental ΔΔG and predictions from several scoring functions for mutations to alanine, as a function of the different types of mutated residues. The results of the following prediction functions are shown: pyDockSC$_{ele+vdw}$ (black), modeling mutation with MODELLER and evaluation with pyDock with van der Waals weight set to 1.0 (dark grey), modeling mutation with MODELLER and evaluation with pyDock electrostatics and van der Waals terms (light grey), FoldX (white).

Regarding the two approaches based on pyDock, it is not clear which one achieves the best results. For some residues like tryptophan, tyrosine or valine, it pays to apply pyDockSC$_{ele+vdw}$, for others like cysteine or glutamine the evaluation of MODELLER structures obtains slightly better results. For mutations from serine to alanine the methods that combined MODELLER and pyDock got negative correlations, closed to 0, while pyDockSC$_{ele+vdw}$ achieved a Pearson correlation of 0.37. Interestingly, the FoldX Pearson correlation for mutations from serine to alanine was 0.32, pointing out that FoldX may be modeling this type of mutations better than MODELLER.

4.6 Protein-RNA docking

Our interest to improve the structural and energetic description of protein-protein docking landscapes have been extended to other protein interactions, such as those between proteins and RNA molecules. We have explored here how to extend the application of pyDock to model protein-RNA complexes. Given the structural differences between proteins and RNA, we wanted to determine if rigid body sampling was a valid method to study protein-RNA interactions. We also checked the validity of the different pyDock energetic terms to score protein-RNA docking models. We used the data compiled in a protein-RNA benchmark developed within our group (Pérez-Cano, Jiménez-García, and Fernández-Recio 2012) to perform this research. This comparative study of protein-protein and protein-RNA interactions allowed us to identify shared and specific characteristics of each type of complexes.

4.6.1 Rigid body sampling in bound and unbound protein-RNA docking

The overall performance of a docking algorithm strongly depends on the ability to generate near-native solutions in the sampling stage. Therefore, we checked the number and quality of the near-native solutions generated by the sampling method, FTDock 2.0. The analysis was conducted in optimal conditions, i.e., using the bound structures of protein and RNA molecules, and in realistic conditions, i.e., with the unbound structures instead.

In optimal conditions, out of the 106 cases of protein-RNA docking benchmark v1.0 (Pérez-Cano, Jiménez-García, and Fernández-Recio 2012), FTDock was able to generate acceptable solutions in 95% of the cases and high-accuracy solutions in 84% of the cases. Within the set of difficult cases without high-accuracy solutions, we found that 53% of them had a buried surface area (BSA) smaller than 800 $Å^2$. On the contrary, this percentage reduced to 18% in the cases with high-accuracy solutions. We hypothesized the difficulties in finding high-accuracy solutions could be related to smaller interface sizes, which are generally associated with lower binding affinity or less specific binding. Finally, we tested that increasing the sampling, generating up to 100,000 docking poses by running FTDock with different initial random rotations for the interacting molecules, did not help to find high-accuracy solutions for the problematic cases.

In realistic conditions, tested on the smaller set of 'unbound' cases of the benchmark, FTDock generated acceptable solutions in 44% of the cases, corresponding to 3 easy (out of 6), 6 medium (out of 13) and 2 difficult (out of 6)

cases. It should be highlighted that FTDock was unable to find any high-quality solution within the set, reflecting the fact that conformational sampling could be one of the most challenging aspects in blind protein-RNA docking.

4.6.2 Scoring capabilities in bound protein-RNA docking

We further studied the performance of different energy terms and functions in the scoring of docking poses (see Figure 4.38-A). We limited our analysis to those bound docking cases for which FTDock had generated at least one high-accuracy solution. The desolvation energy term (see section 3.5.1) showed the lowest performance, followed by the pairwise propensities term (see section 3.5.2) with only around 16% cases with a predicted high-accuracy solution within the top 10 predictions. On the contrary, the terms related to shape or structural complementarity, i.e., FTDock scoring function (SCscore) and van der Waals energy term, obtained the best results, with 54% ad 67% successful cases, respectively. We should also note that electrostatics scoring was significantly poorer than FTDock and van der Waals, with 40% cases with a predicted high-accuracy solution within the top 10 predictions.

In the light of these results, we devised a new protein-RNA scoring function combining the scoring terms with best results, i.e., FTDock SCscore, van der Waals and electrostatics (Eq. 3.13). The new scoring function obtained the highest success rate, with 72% successful cases within the top 10 predictions. Interestingly, the combined function outperformed pyDock's protein-protein scoring function, which is a combination of electrostatics, desolvation and a small contribution of van der Waals energy (see Eq. 3.1).

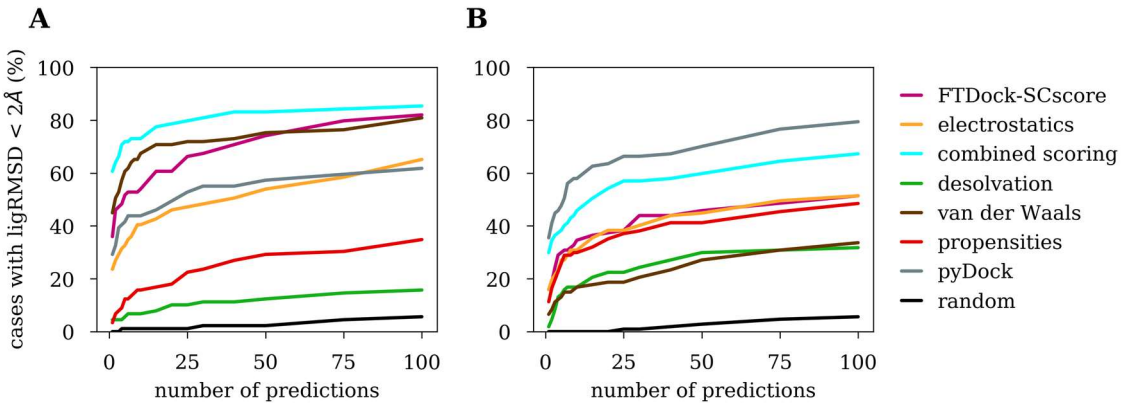

Figure 4.38. Predictive success rates for different scoring functions on docking poses generated for bound protein-RNA (A), and protein-protein (B) benchmarks data sets. For each scoring function, the percentage of cases with high-accuracy docking solutions within the top N number of predictions is shown. The following scoring functions are shown: FTDock-SCscore (magenta), residue-ribonucleotide statistical propensities (red), electrostatics (yellow), desolvation (green), van der Waals (brown), combined scoring (cyan), pyDock (grey), random (black). Only those cases in which FTDock generated at least one high-accuracy solution were considered.

4.6.3 Relationship between scoring performance and quality of sampling

The performance of a docking scoring function strongly depends on the ability of the sampling algorithm to produce near-native conformations. We analyzed the relationship between the success rate of our combined scoring function and the quality of the best near-native solutions generated by FTDock. We found that the scoring success rate of the function strongly depended on the presence of high-quality models within the set of 10.000 protein-RNA models produced by FTDock (see Figure 4.39-A). The combined scoring function identified a near-native solution within the top 10 predictions in around 80% of the cases with excellent-accuracy docking poses within the docking pool. The performance remained at good levels when the best near-native solutions were high-accuracy poses. However, the success rate drastically decreased to 20% in those cases where the best quality solutions were classified as acceptable.

We also evaluated the role of the number of high-accuracy solutions generated by FTDock. As Figure 4.39-B shows, we found that the success rates drastically decreased in those cases in which a single high-accuracy solution was produced.

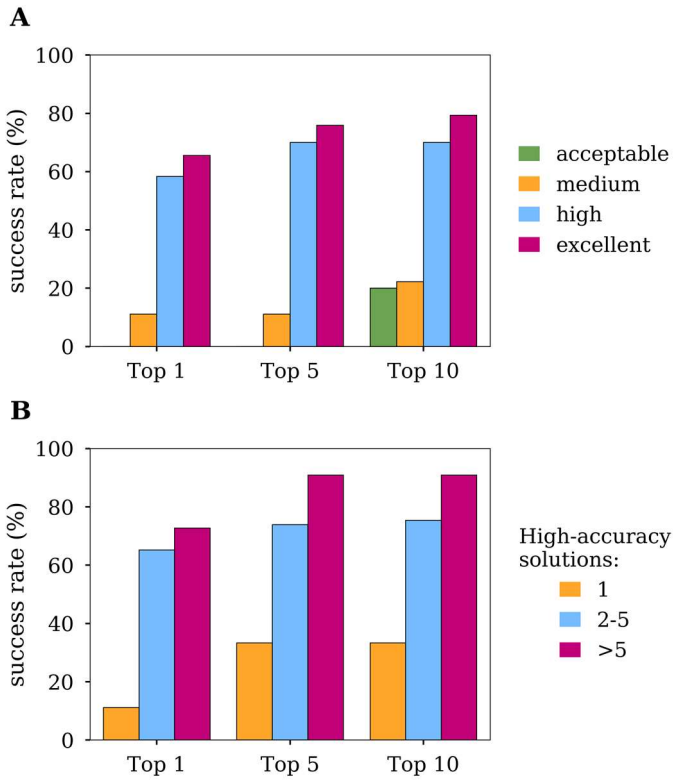

Figure 4.39. Success rates of the combined scoring function for protein-RNA docking poses as a function of the quality of the near-native docking models contained in the docking sets. (A) Percentage of cases in which a near-native solution is found within the top 1, 5 or 10 predictions considering only the cases with excellent accuracy docking solutions (magenta), high-accuracy or worse (blue), medium-accuracy (yellow) or acceptable (green). (B) Percentage of cases in which a high-accuracy docking solution is found within the top 1, 5 or 10 predictions among the cases containing a total of one (yellow), between one and five (blue), or more than five (magenta) high-accuracy docking solutions.

4.6.4 Relationship between scoring performance and protein flexibility

We analyzed the relationship between the performance of the scoring functions and the flexibility of the RNA-binding proteins (see Figure 4.40). To this end, we classified the 89 benchmark cases with at least one high-accuracy docking solution generated by FTDock in three groups, according to the Cα-RMSD between the unbound and bound protein structures, i.e., the rigid set was composed of cases with Cα-RMSD below 2.5 Å, the medium set contained cases with Cα-RMSD between 2.5 and 5 Å, and the flexible set included cases with Cα-RMSD greater than 5 Å. Unexpectedly, the combined scoring function achieved better results in the medium flexible group than in the rigid and highly flexible cases. To study these results in more detail, we analyzed the performance of each of the energetic terms of the combined scoring function depending on protein flexibility. We found that the electrostatics term showed higher success rates for highly flexible cases

(see Figure 4.40-B). Interestingly, FTDock SCscore success rates were worse for highly flexible cases than for rigid or medium cases (see Figure 4.40-C). Van der Waals energy term showed similar behavior, although it was not as dependent on protein flexibility as FTDock SCscore. We could not perform a similar analysis on RNA flexibility since we only had the structure, or a good model, of the unbound RNA for 25 of the 106 benchmark cases.

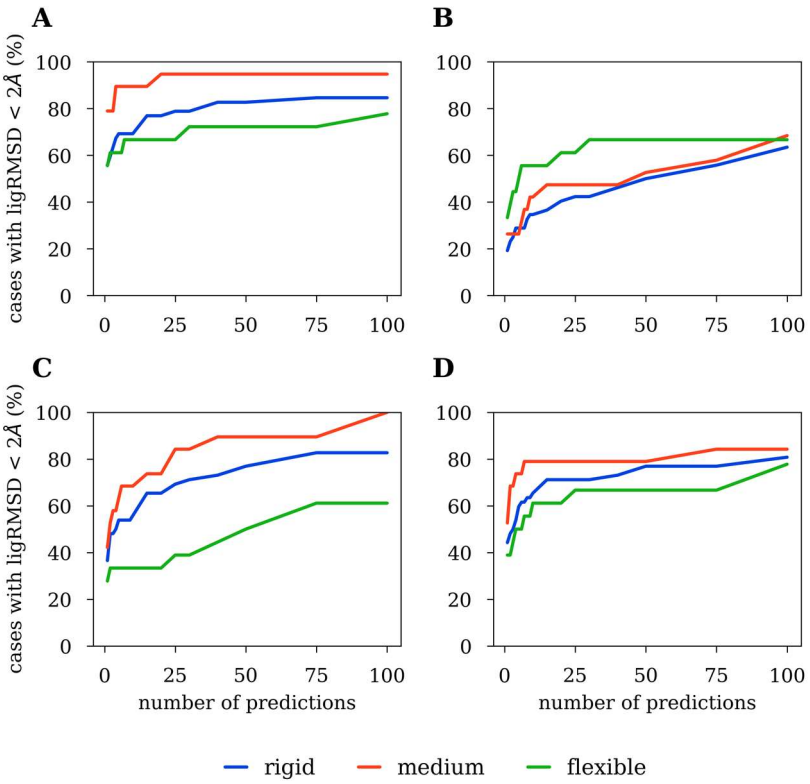

Figure 4.40. Success rates for different scoring functions on the benchmark cases as grouped according to the unbound-to-bound conformational flexibility of the protein. (A) Combined score. (B) Electrostatics. (C) FTDock SCscore. (D) Van der Waals. Results for rigid proteins (blue), medium flexible proteins (red), and highly flexible proteins (green) are shown. Only those cases in which FTDock generated at least one high-accuracy solution were considered.

4.6.5 Scoring performance on the unbound docking set

When tested on the 25 cases of the unbound docking set, FTDock generated acceptable solutions in 11 of the cases and medium-accuracy solutions in 2 cases. For its part, the combined scoring function could identify an acceptable solution within the top 10 predictions only in one case. This case (PDB 1EFW) could be considered as the easiest one, according to the unbound-to-bound conformational flexibility of the interacting molecules, since it was the only case in which both, the protein and RNA molecules had unbound-to-bound RMSDs below 2 Å.

Also, we tested NPDock (Tuszynska et al. 2015), a protein-RNA docking server, on the 25 cases of the unbound docking set. In line with the bad results we had previously obtained, NPDock was unable to identify a near-native solution within the final (up to 3) docking solutions provided by the server. Overall, these results reflect the inherent difficulties of the unbound protein-RNA docking problem.

5 Discussion

5.1 Rigid-body docking: state-of-the-art and current limitations

Protein-protein docking algorithms have experienced a significant development in the last years. One of the elements that has proved to be most useful in this development is CAPRI, the Critical Assessment of Predicted Interactions (Janin et al. 2003). CAPRI has been instrumental in setting-up a docking community and reinforcing the communication channels between the different docking groups spread around the world. It has also served as a valuable tool to gauge the state-of-the-art in the protein-protein docking field, identifying achievements and challenges that still need to be solved. Finally, it has suggested new research directions that the docking community can follow.

In the last years, a great variety of targets have been presented to CAPRI participants. Besides regular protein-protein complex targets, new challenges have been proposed such as interface water molecule prediction (Lensink et al. 2014), protein-peptide and protein-carbohydrate complex modeling and binding affinity estimations (Fleishman, Whitehead, Strauch, et al. 2011; Moretti et al. 2013; Lensink and Wodak 2013b). We should also highlight the first join CASP-CAPRI

prediction experiment that took place in the summer of 2014, bringing together the two communities in an effort to integrate different computational approaches for modeling macromolecular complexes (Lensink et al. 2016).

We can derive several discussion points from the global results obtained in the latest CAPRI rounds in which we have participated, as part of this thesis work. Following, we explain, with a little more detail, the most relevant ones.

5.1.1 Protein-protein docking performance depends on the complex to model

Despite the general improvement experienced in the last years, protein-protein docking performance highly depends on the specific complex to resolve. The performance of protein-protein docking protocols is good when the size of the complex interface is between 1000 and 1500 $Å^2$, there is no need to model the complex subunits, and there are no significant changes between the bound and unbound subunit structures. We could assume that in these easy cases proteins interact as rigid-body or by involving small induced-fit binding. Conversely, protein-protein docking performance drops in cases where interfaces are small, proteins are flexible, any of the subunits is required to be modeled, or the complex to reproduce is a higher order oligomer.

It also seems that protein-protein docking performance is higher for homodimers than for heterocomplexes. For example, in previous CAPRI rounds, about 10 to 15% of the heterocomplex submitted models were correct, while in the first CASP-CAPRI round 25% of the homodimer targets were successfully predicted. It has been reported that homodimer interfaces are often larger and more hydrophobic than those of heterocomplexes. This characteristic could explain why they are easier to predict.

5.1.2 Protein-protein docking and higher order oligomers

As we mentioned above, protein-protein docking performance drops in the case of higher order oligomers. On the one hand, higher order oligomers like tetramers often involve small interfaces, which protein-protein docking algorithms find challenging to model. Additionally, it is very complicated to limit the propagation of errors that may arise from small inaccuracies in the structures of the individual subunits and increase uncertainty during the modeling procedure when successively adding the different subunits forming the complex.

Despite the discrete performance on high order oligomers, current protein-protein docking methods have proved to be useful tools for discriminating the correct

oligomer state of protein complexes. As an example, several CASP-CAPRI targets like T70 and T74, identified initially as tetramers by computational tools like PISA (Krissinel and Henrick 2007) were later reclassified by more accurate experimental techniques. Interestingly, docking results pointed in the same direction, already suggesting that the initial oligomer state was wrongly assigned.

5.1.3 CAPRI scorers performance is better than predictors performance

It has been reported the higher performance of CAPRI scorers over predictors (Lensink and Wodak 2010a, 2013a). This is interesting since scorer predictions are based on a subset of predictors models. Therefore, scorers seem to identify proper solutions from the pools of models generated by predictors more often than predictors themselves. It has been reasoned that since scorers have to evaluate a much smaller number of structures (in the order of a few thousands) than predictors (hundreds of thousands), it is easier for them to discriminate between good and wrong solutions (Feliu and Oliva 2010; Lensink et al. 2016).

Nevertheless, the differences between the scoring schemes applied by both groups must also play a role. In this regard, meta-analysis strategies frequently used by scorers such as the clustering of docking conformations, or the refinement of selected poses may explain part of the improvements in scorer performances. Therefore, if the use of sampling algorithms that generate good complex models is undoubtedly important, it is also necessary to rely on scoring functions and meta-analysis strategies capable of discriminating those good models from the majority of incorrect ones.

5.1.4 The quality of subunit models affects protein-protein docking performance

We have already mentioned that quite often the unbound structures of the complex subunits are not available, and they need to be modeled. Not surprisingly, the accuracy of the final complex model will depend on the quality of the individual subunit models. Interestingly, groups that used several subunit models for a given target had, on average, better results than groups that consider a single subunit model per target. Some groups created the ensemble of models by considering up to five different templates. Others started from a single template that was lately modified by loop optimization and energy refinement. Notably, the schemes that obtained better results were those that perform the optimization in the context of the highest ranked models. In summary, it seems

that those groups that consider ensembles of conformations get better results than groups that work with single models. This observation has been determinant when exploring this strategy in our docking developments.

5.1.5 Integrating additional information improves performance

It has been observed that using additional biological information coming from oligomeric templates or homologous complex templates in the PDB was another critical element contributing to improving the prediction performance. For instance, in the CAPRI-CASP context, even though most of the predictor groups performed *ab-initio* protein-protein docking, two of the top three ranking groups (Seok and Guerois) mostly run template-based modeling strategies, pointing out the potential of using comparative modeling in the docking field. Moreover, the groups that performed *ab-initio* docking guided their algorithms or filtered their solutions based on structural information from homologous oligomers. That said, it has also been observed the notable improvement experienced by automatic servers, some of which obtained results close to the human groups, which generally integrate additional biological information while the servers do not.

5.2 The challenge of considering protein flexibility in docking

The final goal of protein-protein docking algorithms is to reproduce the structure of a complex target from the structures of its components. To achieve their goal, docking algorithms must find the proper structural conformation of the subunits and their correct orientation. In Figure 4.9 we saw how the rigid-body minimization algorithm that we developed could find the native structure, the global energy well, when the subunits were in their bound state but was unable to do it when the subunits were in their unbound state. We can draw conclusions from this and apply it to the protein-protein binding process. Namely, protein flexibility plays a key role in protein-protein binding and should be included in protein-protein docking protocols to improve their performance. To this goal, in this thesis work, we have followed several approaches. For instance, the development of pyDockLite (section 4.2), a simplified scoring function derived from pyDock, allowed us to introduce normal modes in our minimization protocol to model backbone flexibility (section 4.3.2). We also implemented pyDockLite as one of the scoring functions of LightDock (section 4.2.3), a docking framework based on the Glowworm Swarm Optimization (GSO) algorithm recently

developed within our group. Even though the scheme based on the normal modes minimization did not systematically improve the quality of the original docking conformations (in terms of Cα-LigRMSD from the native reference structure) we observed a small improvement in the general predictive success rates. Moreover, when we grouped the docking cases according to their flexibility, the top 10 success rates mostly improved for the medium-flexible cases, for which the Cα-IntRMSD between the bound and unbound conformations is in the range of 1 to 2 Å. Top 20 success rate also improved for flexible cases for which Cα-IntRMSD ranged between 2 and 3 Å. These results are in line with those obtained with LightDock-pyDockLite on protein-protein benchmark 5 that outperformed pyDock for the medium-flexible cases (see Figure 4.8). That is, we got similar performances using different optimization algorithms and a common force field. On the other hand, LightDock-pyDockLite reported worse results than pyDock for the rigid cases. Interestingly, the normal modes minimization also had worse results than pyDock when the van der Waals weight of pyDockLite was set to 0.1 but achieved performance comparable to that of pyDock when using a van der Waals weight of 1.0. Given that in the pyDockLite scoring function implemented in LightDock (LightDock-pyDockLite) the van der Waals weight is set to 0.1, it would be interesting to evaluate the results of LightDock-pyDockLite with the van der Waals weight set to 1.0. This suggests that it might be better if the force-fields used in docking and minimization algorithms are different.

In addition to studying backbone flexibility, we have also modeled side-chain flexibility through rotamer optimization with SCWRL. The combination of SCWRL and pyDock matched the top 10 success rate obtained with pyDock alone but improved the performance for higher rank thresholds. For example, the top 100 success rate was 46% if we modeled side-chain flexibility compared to 38% if proteins were treated as rigid bodies.

Overall, these results are promising and indicate that by considering the flexibility of proteins we may significantly increase the performance of our docking algorithms. However, there is still plenty of room for improvement. Coupling backbone and side-chain flexibility and developing new functions to be optimized during the minimization are two possible paths to explore.

5.3 Exploring the docking landscape: regions *vs.* single points

As above discussed, we know that proteins are in continuous movement, sampling its conformational space. However, the performance of protein-protein docking

algorithms is usually assessed by comparing the resulting models to a single structure of the target complex, which has been obtained experimentally, usually by x-ray crystallography. Therefore, when we measure the quality of the docking algorithms, we are ignoring protein flexibility. By omitting the dynamic component of proteins, we are not only neglecting an essential aspect of its nature, but we may also be contributing to decreasing the effectiveness of our docking protocols and limiting the amount of useful information they can provide.

pyDock standard output consists of a list of 10,000 docking poses, their coordinates in the rotational and translational space, and their computed docking energies. Namely, pyDock provides information about *points* of the docking landscapes. The ensemble-based description that we have developed allows us to analyze *regions* of the docking landscapes instead. We have shown that by scoring ensembles/regions instead of docking poses/points the results of our docking protocols improved. It is worth considering the docking output as a distribution of related docking poses, instead of a list of independent conformations. The ensembles/clusters of rigid-body docking poses explore regions in the rotational-translational space, while the ensembles of MODELLER conformers explore regions in the proteins conformational space. We have not detected significant differences in the performances of these two ensemble methods. In both cases, we observed that pyDock scoring yielded better results than zDope. For example, P_{best} scoring obtained top 10 success rates of 20% and 22% with the clustering and conformational ensembles, respectively. However, the results really improved when using the combination of pyDock and zDope in consensus scores, with top 10 success rates of 27% and 26% for the clustering and conformational ensembles, respectively. These values are several percentage points higher than those obtained when scoring docking poses with standard pyDock (18%), zDope (18%) or their consensus score (20%). Table 4.5 shows that MODELLER conformers, on average, have not lower Cα-LigRMSD with respect the complex structure than the original docking poses. However, their docking energies improve those of the original docking poses and, remarkably, this improvement is greater for near-native conformers than non-near-native conformers. Interestingly, in a previous study, MODELLER-based conformational ensembles generated from the unbound subunits before docking (Pallara et al. 2016) did not find a significant correlation between the docking energy of the conformers in the native orientation and their structural similarity with the bound complexes. They even found that, in some cases, the conformers with best docking energy were even farther from the bound structure than the unbound one. They suggested that this could be due to MODELLER performing a limited sampling, insufficient to explore the vicinity

of the bound state. In those circumstances, they argue, small approaches toward the bound state are not related to improvements in binding energy. We ran the same default protocol to generate MODELLER conformers as Pallara and coworkers. Table 4.5 confirms that MODELLER sampling may be quite limited. Therefore, we believe that the results obtained here could be improved if we increased the conformational variability of our sampling, either with new tools or by modifying the minimization protocol we have used with MODELLER. In this regard, it is worth to recall that we achieved the highest top 10 success rates when we increased the conformational variability of the ensembles by merging the rotational-translational ensembles with the conformational ensembles.

5.4 An ensemble-based description of docking energy landscapes

As above discussed, the ensemble-based description of docking energy landscapes has advantages in the identification of the near-native docking regions. For example, the ensemble methods we presented in section 4.4 are developed to generate as output a sorted list of ensemble representatives. For that, the scoring is divided into two phases. In the first phase, the clusters/ensembles are sorted according to the aggregate value, i.e., mean, minimum, of a given metric like pyDock energy of MODELLER zDope. In the second phase, the structure representative of each cluster/ensemble is selected. Many times, the structure selected as cluster representative has not the best Cα-LigRMSD of the cluster. In fact, we may select a not near-native structure as representative of a cluster with near-native solutions. These mistakes decrease the measured success rates of our docking algorithms. Figure 5.1 shows the top 10 success rates that the different ensemble methods would obtain if we were able to assign as ensemble representative, the structure of the ensemble with lowest Cα-LigRMSD. According to these values, we may be losing 3-4 percentage points due to the incorrect identification of the ensembles representatives.

Additionally, when we limit the output of our algorithms to the cluster representatives, we are removing information that could be useful to understand the binding mechanism, propose new experiments or interpret experimental results. We have seen that by using ensembles of conformers instead of single conformations we can improve the performance of our docking method due to a better representation of the docking energy landscapes. Similarly, we think that this ensemble-centric view could be exploited for other applications like interface prediction, identification of hot-spots residues, etc.

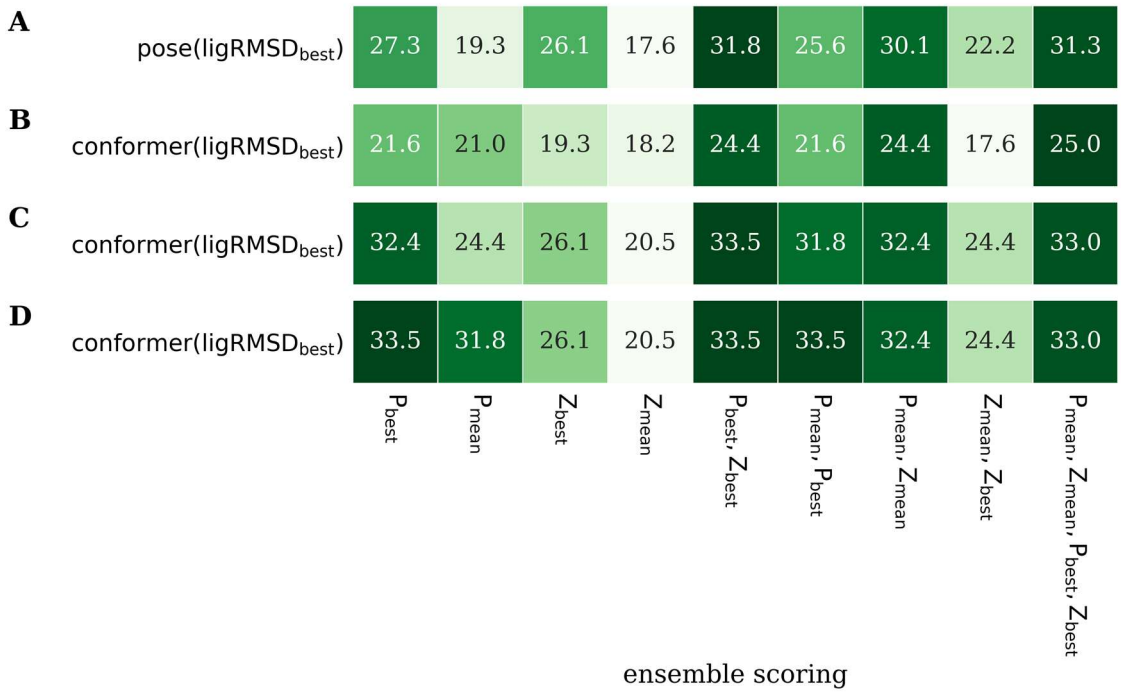

Figure 5.1. Optimal docking success rates for the top 10 predicted models on the protein-protein docking benchmark 4 for several scoring schemes using: A) clustering ensembles, B) conformational ensembles obtained with restricted molecular dynamics, C) combining clustering and conformational ensembles and D) combining clustering and conformational ensembles from docking poses of the first 100 clusters sorted by "best pose zDope". These success rates are theoretical, in the best possible conditions, that is, assuming that we are able to identify as clusters/ensemble representative the structure (pose, conformer) of lowest Cα-LigRMSD.

5.5 Docking energy description and binding affinity

5.5.1 The role of the different energetic terms in pyDock scoring function

pyDock scoring function heavily undercuts the van der Waals interaction term to avoid penalizations of near-native models due to clashes. This decision is justifiable when the docking involves unbound structures. However, we have verified that its use with bound structures can be very beneficial. For example, when we performed the rigid-body minimization with bound structures in section 4.3.1 the term that contributed the most to identify near-native solutions correctly was van der Waals.

Additionally, in the minimization including normal modes (section 4.3.2), we observed that, in general, when we changed the van der Waals weighting factor from 0.1 to 1.0 the scoring function could improve the ranking of near-native solutions. This could open the way to the development of scoring functions in

which the van der Waals weight is variable, depending on whether the atoms might have potential clashes or not. Furthermore, there are other energetic terms, like hydrogen bonding, whose inclusion does not seem necessary when performing rigid-body docking of unbound conformations but could have a positive contribution in flexible docking or when the structures of the subunits are similar to the bound conformations.

The development of pyDockLite has shown that pyDock performance does not seem to be especially sensitive to the range of action of the energetic terms. We have increased pyDock computation speed by decreasing the distance at which the effects of the energetic terms are noticeable, reducing the number of atoms that are included in the computations. Despite this reduction, that in the case of the desolvation term restricts the interaction to the closest atom, the performance of pyDockLite does not significantly worsen the results obtained with pyDock but increases the computation speed tenfold. This improvement will allow studies that until now were not feasible for pyDock like cross-docking, docking of higher order oligomers and a much more detailed exploration of the docking landscape.

5.5.2 Binding affinity

We have discussed above the different sampling and scoring schemes developed in this thesis to improve the identification of the near-native docking structures. Protein complex structures are defined by the interaction energies among the atoms of its subunits, and thus, any docking scoring function aiming to identify near-native solutions must include an implicit description of the interaction energy landscape. In this context, we have followed the trends of the docking community, which has shown a recent interest in exploring whether the docking functions can estimate, for example, the experimental binding affinity of the complex or changes in such binding affinity upon mutation. We have developed two different methods to estimate changes in binding energy upon mutation ($\Delta\Delta G$). The first method was based on partitioning pyDock energy at its atomic and residual level, which could describe mutations to alanine. The second approach combined MODELLER and pyDock to model the structural changes induced by the mutation. None of them has the accuracy required to make quantitative predictions of $\Delta\Delta G$ values. However, they can be applied to obtain a qualitative description of residue energies and identify, for instance, hot-spots residues (see section 4.5.2). We must keep in mind that pyDock scoring function has been designed as part of a rigid-body protein-protein docking algorithm, with requirements that are not always in line with those of a $\Delta\Delta G$ predictor, i.e., resilience to clashes, limited use of computational resources, etc. Nevertheless, the

results we obtained are on a par with those of other tools specifically focused on free energy calculations. That said, we should also notice that both problems, i.e., protein-protein docking and free energy calculations, should ideally be solved by a common scoring function. That is, if we had a scoring function capable of estimating free energy reliably, it is very likely that we could also use it to correctly identify the near-native structures from a pool of docking poses. Including additional terms such as hydrogen bonding, salt bridges and entropy could improve the results achieved by pyDock-derived scoring functions. All these developments should be in parallel to improvements of the sampling algorithms to obtain better quality models. In this regard, we should make efforts to enhance current methods to model the structural changes upon mutation. The results we have obtained suggest that it is beneficial to represent the different minima of the docking energy landscape as ensembles of models instead of single models. We have studied how the correlation between experimental and predicted $\Delta\Delta G$ values depend on the number of models included on the ensembles. We also tested two different strategies to select the ensemble models among the 144 conformers originally generated with MODELLER (see section 4.5.5.2): i) select the best-ranked models according to their MODELLER zDope value or ii) select the models randomly. As Figure 5.2 shows, the pyDock+MODELLER method achieved only slightly higher correlation values if we selected the best zDope models instead of choosing them randomly. More interestingly, the correlation values reached a plateau soon, with a small number of models. This suggests that we could significantly decrease the computational cost of the pyDock+MODELLER protocol by considering ensembles of 10 conformers instead of the 144 that were initially computed. The correlation values of FoldX, by contrast, are almost constant regardless of the number of models we use.

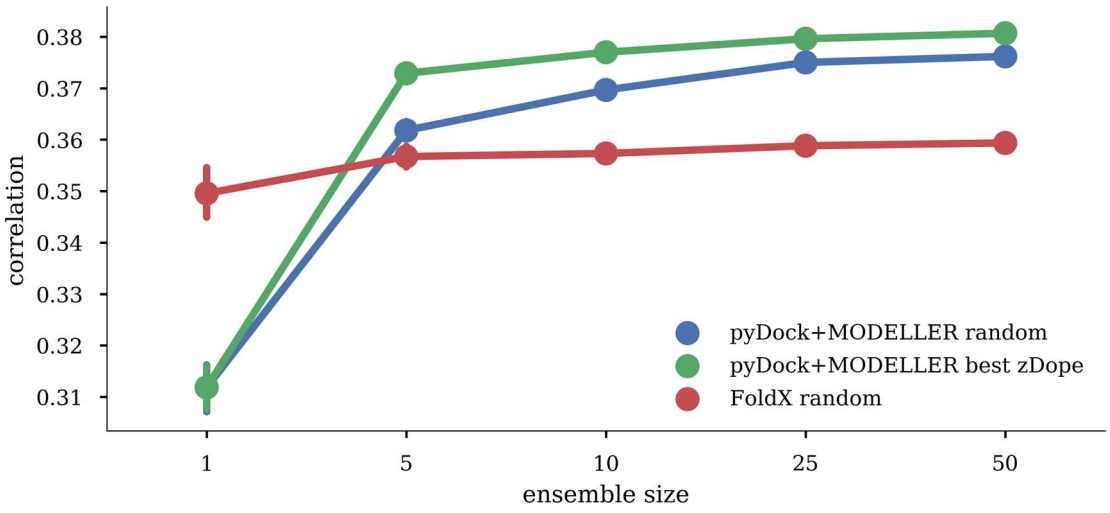

Figure 5.2. Correlation between experimental and predicted ΔΔG as a function of the number of models used to represent the mutated complex. (Blue) Predictions by pyDock+MODELLER when choosing the models randomly, (green) predictions by pyDock+MODELLER choosing the models with best zDope, (red) predictions by FoldX choosing the models randomly.

To improve the performance of our ΔΔG predictors, we must also increase the quality of the structural and affinity data we have available. Experimental data is essential for training and testing our models and predictors. Nowadays, SKEMPI is probably the largest database of experimental binding free energy changes upon mutation and has become an invaluable tool in the development of new methods. Since SKEMPI data have been collected from the literature, it reflects the interests of the experimentalists who obtained them. Therefore, the data show bias towards specific residues, types of mutations, spatial locations, binding sites, proteins and protein families (Moal and Fernández-Recio 2012b). Moal and co-workers identified these biases and proposed a cross-validation scheme based on simultaneously holding out interactions at the same, or homologous, binding sites to avoid overestimating the predictive power of the developed methods. One of the best ways to check for overfitting is external validation by which the models are tested against an independent dataset that has not been used for training the models and tuning their hyperparameters. We performed an external validation of our method based on combining MODELLER and pyDock and a second well-known predictor called mCSM (Pires, Ascher, and Blundell 2014). mCSM is a machine-learning based method that reported a Pearson correlation coefficient of 0.80 with a standard error of 1.25 Kcal/mol, after applying 10-fold cross-validation over 2317 single-point mutations (150 different proteins) taken from SKEMPI. We built the external validation dataset from the mutations reported in Gärdsvoll *et al.* (Gårdsvoll et al. 2006) and Kiel *et al.* (Kiel and Serrano 2014). We ended up with 140 mutations corresponding

to 5 different PDB structures that were neither protease-inhibitor nor antigen-antibody complexes, two of the more overrepresented complex families in SKEMPI. The Pearson correlation coefficients for MODELLER+pyDock and mCSM on the external dataset were 0.30 and 0.32, respectively. These values represent a decrement from the reported correlations when the methods were tested with SKEMPI data, 0.38 for MODELLER+pyDock, and 0.80 for mCSM. The decrease is much more pronounced in the case of mCSM, suggesting possible overfitting and generalization issues. The above example shows that to take advantage of SKEMPI, it is critical to be aware of the biases contained in its data and adopt the validation schemes proposed by its authors, or similar ones. Only this way we can avoid our models to learn patterns in the data that are originating from noise.

The experimental ΔΔG data we have available is, in fact, very noisy. Figure 4.32 already pointed out that not all experimental techniques seem to be equally reliable. Variables like temperature, pH or salt concentrations may have an important influence on the experimental values obtained. As the noise in experimental data may contribute to decreasing the performance of the models we develop it would be essential to develop a program to collect ΔΔG data *systematically*. The program should define a set of standards, i.e., experimental methods and conditions, as well as a list of systems that accurately represent the diversity of protein interactions. This effort would foster improvement of the performance of ΔΔG predictors by increasing the quality and quantity of experimental data at their disposal, currently too scarce, heterogeneous and redundant.

5.6 Protein-RNA docking

5.6.1 Comparison of the different energetic terms for protein-RNA and protein-protein docking

We performed a comparative analysis of the behavior of several energetic terms in protein-RNA and protein-protein docking. Interestingly, our results indicate an essential and distinctive role of shape and structural complementarity in protein-RNA association. Concretely, our results suggest that the FTDock score and van der Waals energy, when considered individually, had significantly better predictive rates for protein-RNA bound docking than for protein-protein docking (see Figure 4.38). Indeed, the FTDock score capabilities for protein-RNA docking were previously suggested on a much more limited set of cases (Pérez-Cano et al.

2010). On the other hand, electrostatics had slightly better predictive rates for the scoring of protein-RNA docking poses, while the desolvation energy term showed much better success rate for protein-protein docking. We also observed that the set of protein-RNA pairwise propensities developed within our group previously (Pérez-Cano et al. 2009) had a very low predictive value even in bound docking conditions (see Figure 4.38). One possible explanation is that for a given amino acid type, the binding propensities were not significantly different among the four ribonucleotides types (see Figure 3.2). Therefore, while pairwise propensities can be used to identify the RNA binding sites on proteins correctly, they are too noisy to distinguish different orientations of RNA molecules that are bound to the correct binding site. By contrast, residue-residue pairwise propensities achieved a better scoring performance in protein-protein docking, indicating a higher specificity in the residue-residue contacts than in the residue-ribonucleotide ones.

5.6.2 The role of electrostatics in protein-RNA binding depends on protein flexibility

We found that the role of electrostatics in protein-RNA bound docking strongly depends on the unbound-to-bound conformational flexibility of the RNA-binding proteins. As shown in Figure 4.40-B electrostatics was especially successful for the highly flexible cases. We hypothesized that in highly flexible cases the electrostatics contribution to binding affinity should be much more important than in rigid or medium flexible cases. This would be consistent with a situation in which highly flexible proteins would have a higher enthalpic contribution to compensate for the larger conformational entropy penalization. Electrostatics could provide this general enthalpic gain since it is more tolerant to conformational flexibility as compared to other terms like van der Waals or hydrogen bonding. To confirm this, we analyzed the number and type of charged residues at the RNA-binding sites in proteins, according to their flexibility. We defined the interface net charge as the difference between the number of positively-charged Arg/Lys and negatively-charged Asp/Glu protein residues that are found within 5 Å from the RNA molecule. Figure 5.3 shows the distribution of the interface net charge values for the different groups of proteins according to unbound-to-bound flexibility. As much as 81% of the highly flexible proteins have RNA-binding interfaces with larger positive net charge (i.e.,$> +5$), as compared to 52% of the medium flexible or rigid proteins. Interestingly, although rigid proteins have less positive interfaces in average, the range of the distribution is broader, with a few extreme cases, from the most positively charged (1C9S and

2GIC, both involving large protein oligomers) to the more negatively charged ones (1EIY, a tRNA-synthetase from Thermus thermophilus). The higher percentage of positively charged interfaces in the flexible proteins were not caused by a larger interface size since flexible proteins had in general smaller interfaces than the rigid ones (see section 5.6.3). All these findings are in line with the above-mentioned hypothesis that protein-RNA interfaces involving flexible proteins are more electrostatic.

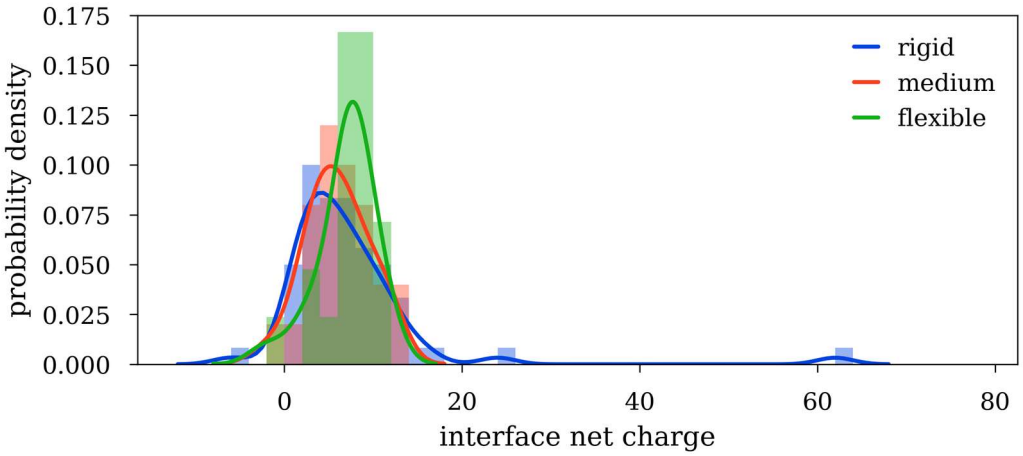

Figure 5.3. Distribution of interface net charge of native protein-RNA interfaces according to the unbound-to-bound conformational flexibility of the protein. The histogram bar values show the normalized population for each range of interface net charge values. A smoothed curve representing the values has been added for a clearer visualization.

Figure 5.4 shows two examples of protein-RNA complexes involving rigid and highly flexible proteins. The case with greater flexibility has a higher proportion of Arg/Lys residues, which yields a more positive interface and more favorable electrostatics.

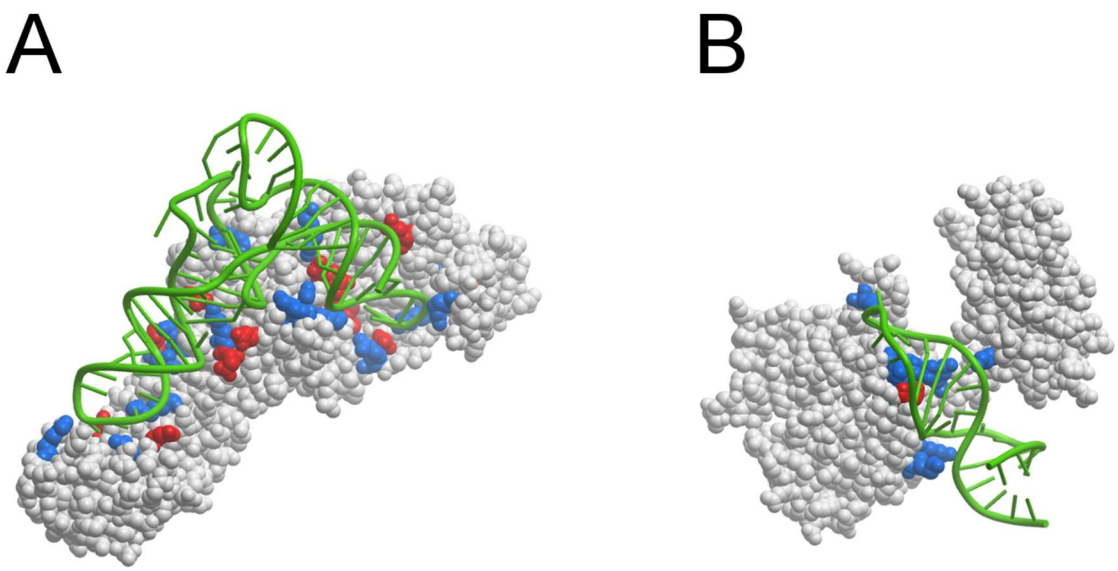

Figure 5.4. Location of positively (blue) and negatively (red) charged residues at protein-RNA interfaces in two examples. (A) A rigid case: PDB code 1N78 (interface net charge +2). (B) A highly flexible case: PDB code 1OOA (interface net charge +8).

5.6.3 RNA-binding sites of flexible proteins are smaller and have less surface complementarity

We also found that the role of structural complementarity (as defined by SCscore) in protein-RNA bound docking is strongly dependent on the unbound-to-bound conformational flexibility of the RNA-binding proteins. Indeed, the SCscore yielded much worse scoring for the highly flexible cases than for the rigid or medium flexible cases, despite using the bound coordinates of the protein and RNA components (see Figure 4.40-C). Perhaps highly flexible proteins, which rely more on long-range electrostatics interactions as above described, do not need to form highly complementary interfaces, which would involve a significant entropy penalization. Indeed, the SCscore values obtained for the native interfaces of protein-RNA complexes involving highly flexible proteins were less favorable in average (246.4 $\pm$ 66.2 a.u.) than those of the complexes involving medium flexible (270.4 $\pm$ 86.1 a.u.) or rigid proteins (285.6 $\pm$ 163.2 a.u.). This could indicate that protein-RNA interfaces involving highly flexible proteins were in average less packed than those of rigid or medium flexible proteins. Consistently, we found that none of the highly flexible proteins had a large interface (i.e. BSA $>$ 1500 Å^2). On the contrary, 12% of the medium flexible proteins, and 28% of the rigid ones had such large interfaces, with a few rigid proteins showing the largest interface size among all cases (e.g., 1C9S and 2GIC, both involving large protein oligomers). Accordingly, scoring by van der Waals energy alone was also worse for cases involving highly flexible proteins (see Figure 4.40-D). However, the

difference in performance according to unbound-to-bound conformational flexibility was not as significant as that yielded by SCscore scoring (Figure 4.40-C), perhaps due to the noisier character of the van der Waals scoring.

5.6.4 Present and future challenges in protein-RNA docking

We found that a combination of structural complementarity and electrostatic parameters is successful for the scoring of easy protein-RNA docking cases. However, given that most of the known protein-RNA complexes involve significant local and global conformational changes (Dominguez et al. 2011; Hyeon, Dima, and Thirumalai 2006; Ke and Doudna 2004), further improvement in the sampling and scoring methodology is much needed. One possibility is that scoring functions could integrate more coarse-grained parameters to deal with inaccuracies derived from sub-optimal conformational sampling in unbound protein-RNA docking. Although we have found here that the use of low-resolution (or residue- level) pairwise propensities are noisy, other previously reported medium resolution propensities could be more effective (Setny and Zacharias 2011; Tuszynska and Bujnicki 2011) and it remains to be seen whether they could complement the energy terms described in this work.

We also found that scoring (either atomic or coarse-grained) becomes extremely challenging for evaluating solutions that differ more than 5 Å RMSD from the bound structures. One of the future major challenges for protein-RNA docking will be modeling the flexibility of the protein or RNA molecules. The challenge is even greater if we consider that modeling the flexibility of RNA molecules requires different parameterization than that used for proteins (Fulle and Gohlke 2008) and involves two main difficulties: the non-ergodic behavior of RNA and the strong influence of ion molecules on the dynamic processes (Al-Hashimi and Walter 2008).

6 Conclusions

1. The analysis of our participation in CAPRI shows that protein-protein docking is still challenging when interacting proteins are flexible or need to be modeled, or when the complex to reproduce is a higher order oligomer. For these challenges, more accurate docking and faster scoring tools are needed.
2. We have developed pyDockLite, a simplified scoring function derived from pyDock which is up to 10 times faster at comparable performance. The new distance-based desolvation term therein drastically reduces the computation time required to calculate the desolvation contribution to pyDock docking energy.
3. Based on pyDockLite, we have developed a fast rigid-body minimization algorithm, which is very efficient when the subunits are in their bound conformation.
4. The introduction of normal modes in the minimization algorithm to model backbone flexibility showed an improvement in the success rates, especially for the medium-flexible and flexible cases.
5. A new ensemble-based description of the docking landscapes, integrating clustering, conformational sampling and consensus scoring, can improve docking performance. The observed improvement was related to a better energetic description rather than to increase the structural similarity with the bound state.

6. A new method to compute pyDock docking energy per residue can be used to identify energetically relevant residues in the binding process (hot-spots) and to estimate binding affinity changes upon mutation to alanine.
7. Regarding mutations to other residues, we have developed a new method to predict binding affinity changes upon mutation by combining MODELLER and pyDock. Results are in line with previous methods when tested on an external validation dataset.
8. A new scoring function combining FTDock score and pyDock electrostatics and van der Waals energy terms can be used to evaluate docking models of protein-RNA complexes.
9. The energetic analysis of protein-RNA complexes suggests a dependence of RNA recognition mechanism on the protein flexibility. Electrostatics may play a major role in highly flexible RNA-binding proteins, while structural complementarity is more important in rigid RNA-binding proteins.

7 Bibliography

Abagyan, R., and M. Totrov. 1994. "Biased Probability Monte Carlo Conformational Searches and Electrostatic Calculations for Peptides and Proteins." *Journal of Molecular Biology* 235 (3): 983–1002. https://doi.org/10.1006/jmbi.1994.1052.

Al-Hashimi, Hashim M, and Nils G Walter. 2008. "RNA Dynamics: It Is about Time." *Curr. Opin. Struct. Biol.* 18 (3): 321–29. https://doi.org/10.1016/j.sbi.2008.04.004.

Aloy, Patrick, and Robert B. Russell. 2004. "Ten Thousand Interactions for the Molecular Biologist." *Nature Biotechnology* 22 (10): 1317–21. https://doi.org/10.1038/nbt1018.

Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. "Basic Local Alignment Search Tool." *Journal of Molecular Biology* 215 (3): 403–10. https://doi.org/10.1016/S0022-2836(05)80360-2.

Amunts, Alexey, Alan Brown, Jaan Toots, Sjors H. W. Scheres, and V. Ramakrishnan. 2015. "The Structure of the Human Mitochondrial Ribosome." *Science* 348 (6230): 95–98. https://doi.org/10.1126/science.aaa1193.

Andreani, Jessica, Guilhem Faure, and Raphael Guerois. 2013. "InterEvScore: A Novel Coarse-Grained Interface Scoring Function Using a Multi-Body Statistical Potential Coupled to Evolution." *Bioinformatics* 29 (14): 1742–49. https://doi.org/10.1093/bioinformatics/btt260.

Andrusier, Nelly, Efrat Mashiach, Ruth Nussinov, and Haim J Wolfson. 2008. "Principles of Flexible Protein–Protein Docking." *Proteins* 73 (2): 271–89. https://doi.org/10.1002/prot.22170.

Andrusier, Nelly, Ruth Nussinov, and Haim J. Wolfson. 2007. "FireDock: Fast Interaction Refinement in Molecular Docking." *Proteins: Structure, Function, and Bioinformatics* 69 (1): 139–59. https://doi.org/10.1002/prot.21495.

Anfinsen, C. B., E. Haber, M. Sela, and F. H. White. 1961. "The Kinetics of Formation of Native Ribonuclease during Oxidation of the Reduced Polypeptide Chain." *Proceedings of the National Academy of Sciences of the United States of America* 47 (September): 1309–14.

Antal, Miklós A., Csaba Böde, and Peter Csermely. 2009. "Perturbation Waves in Proteins and Protein Networks: Applications of Percolation and Game Theories in Signaling and Drug Design." *Current Protein & Peptide Science* 10 (2): 161–72.

Assi, Salam A., Tomoyuki Tanaka, Terence H. Rabbitts, and Narcis Fernandez-Fuentes. 2010. "PCRPi: Presaging Critical Residues in Protein Interfaces,

a New Computational Tool to Chart Hot Spots in Protein Interfaces." *Nucleic Acids Research* 38 (6): e86–e86. https://doi.org/10.1093/nar/gkp1158.

Atilgan, A R, S R Durell, R L Jernigan, M C Demirel, O Keskin, and I Bahar. 2001. "Anisotropy of Fluctuation Dynamics of Proteins with an Elastic Network Model." *Biophysical Journal* 80 (1): 505–15.

Bahadur, R. P., and M. Zacharias. 2008. "The Interface of Protein-Protein Complexes: Analysis of Contacts and Prediction of Interactions." *Cellular and Molecular Life Sciences* 65 (7–8): 1059–72. https://doi.org/10.1007/s00018-007-7451-x.

Bahadur, Ranjit Prasad, Pinak Chakrabarti, Francis Rodier, and Joël Janin. 2004. "A Dissection of Specific and Non-Specific Protein-Protein Interfaces." *Journal of Molecular Biology* 336 (4): 943–55.

Bahar, Ivet, Ali Rana Atilgan, and Burak Erman. 1997. "Direct Evaluation of Thermal Fluctuations in Proteins Using a Single-Parameter Harmonic Potential." *Folding and Design* 2 (3): 173–81. https://doi.org/10.1016/S1359-0278(97)00024-2.

Bailey, Lucas J., Jason G. McCoy, George N. Phillips, and Brian G. Fox. 2008. "Structural Consequences of Effector Protein Complex Formation in a Diiron Hydroxylase." *Proceedings of the National Academy of Sciences of the United States of America* 105 (49): 19194–98. https://doi.org/10.1073/pnas.0807948105.

Bakan, Ahmet, Lidio M. Meireles, and Ivet Bahar. 2011. "ProDy: Protein Dynamics Inferred from Theory and Experiments." *Bioinformatics (Oxford, England)* 27 (11): 1575–77. https://doi.org/10.1093/bioinformatics/btr168.

Baker, B. M., and K. P. Murphy. 1998. "Prediction of Binding Energetics from Structure Using Empirical Parameterization." *Methods in Enzymology* 295: 294–315.

Barik, Amita, Nithin C, Manasa P, and Ranjit Prasad Bahadur. 2012. "A Protein-RNA Docking Benchmark (I): Nonredundant Cases." *Proteins* 80 (7): 1866–71. https://doi.org/10.1002/prot.24083.

Bastard, Karine, Aurélien Thureau, Richard Lavery, and Chantal Prévost. 2003. "Docking Macromolecules with Flexible Segments." *Journal of Computational Chemistry* 24 (15): 1910–20. https://doi.org/10.1002/jcc.10329.

Beauchene, Isaure Chauvot de, Sjoerd J. de Vries, and Martin Zacharias. 2016. "Binding Site Identification and Flexible Docking of Single Stranded RNA to Proteins Using a Fragment-Based Approach." *PLOS Computational Biology* 12 (1): e1004697. https://doi.org/10.1371/journal.pcbi.1004697.

Ben-Shimon, Avraham, and Miriam Eisenstein. 2010. "Computational Mapping of Anchoring Spots on Protein Surfaces." *Journal of Molecular Biology* 402 (1): 259–77. https://doi.org/10.1016/j.jmb.2010.07.021.

Berg, Jeremy M., John L. Tymoczko, Lubert Stryer, and Lubert Stryer. 2007. *Biochemistry*. 6th ed. New York: W.H. Freeman.

Bernstein, F C, T F Koetzle, G J Williams, E F Meyer Jr, M D Brice, J R Rodgers, O Kennard, T Shimanouchi, and M Tasumi. 1977. "The Protein Data Bank: A Computer-Based Archival File for Macromolecular Structures." *J. Mol. Biol.* 112 (3): 535–42.

Boehr, David D., Ruth Nussinov, and Peter E. Wright. 2009. "The Role of Dynamic Conformational Ensembles in Biomolecular Recognition." *Nature Chemical Biology* 5 (11): 789–96. https://doi.org/10.1038/nchembio.232.

Bogan, Andrew A, and Kurt S Thorn. 1998. "Anatomy of Hot Spots in Protein Interfaces." *Journal of Molecular Biology* 280 (1): 1–9. https://doi.org/10.1006/jmbi.1998.1843.

Bonvin, Alexandre MJJ. 2006. "Flexible Protein–Protein Docking." *Current Opinion in Structural Biology* 16 (2): 194–200. https://doi.org/10.1016/j.sbi.2006.02.002.

Bosshard, H. R. 2001. "Molecular Recognition by Induced Fit: How Fit Is the Concept?" *News in Physiological Sciences: An International Journal of Physiology Produced Jointly by the International Union of Physiological Sciences and the American Physiological Society* 16 (August): 171–73.

Brooks, B. R., C. L. Brooks, A. D. Mackerell, L. Nilsson, R. J. Petrella, B. Roux, Y. Won, et al. 2009. "CHARMM: The Biomolecular Simulation Program." *Journal of Computational Chemistry* 30 (10): 1545–1614. https://doi.org/10.1002/jcc.21287.

Bryngelson, Joseph D., José Nelson Onuchic, Nicholas D. Socci, and Peter G. Wolynes. 1995. "Funnels, Pathways, and the Energy Landscape of Protein Folding: A Synthesis." *Proteins: Structure, Function, and Bioinformatics* 21 (3): 167–95. https://doi.org/10.1002/prot.340210302.

Caffrey Daniel R., Somaroo Shyamal, Hughes Jason D., Mintseris Julian, and Huang Enoch S. 2009. "Are Protein–Protein Interfaces More Conserved in Sequence than the Rest of the Protein Surface?" *Protein Science* 13 (1): 190–202. https://doi.org/10.1110/ps.03323604.

Callaway, Ewen. 2015. "The Revolution Will Not Be Crystallized." *Nature* 525 (7568): 172.

Canutescu, Adrian A., Andrew A. Shelenkov, and Roland L. Dunbrack. 2003. "A Graph-Theory Algorithm for Rapid Protein Side-Chain Prediction." *Protein Science : A Publication of the Protein Society* 12 (9): 2001–14.

Case, David A. 1994. “Normal Mode Analysis of Protein Dynamics.” *Current Opinion in Structural Biology* 4 (2): 285–90. https://doi.org/10.1016/S0959-440X(94)90321-2.

Case, David A., Thomas E. Cheatham, Tom Darden, Holger Gohlke, Ray Luo, Kenneth M. Merz, Alexey Onufriev, Carlos Simmerling, Bing Wang, and Robert J. Woods. 2005. “The Amber Biomolecular Simulation Programs.” *Journal of Computational Chemistry* 26 (16): 1668–88. https://doi.org/10.1002/jcc.20290.

Chaudhury, Sidhartha, Sergey Lyskov, and Jeffrey J. Gray. 2010. “PyRosetta: A Script-Based Interface for Implementing Molecular Modeling Algorithms Using Rosetta.” *Bioinformatics* 26 (5): 689–91. https://doi.org/10.1093/bioinformatics/btq007.

Chazelle, B., C. Kingsford, and M. Singh. 2004. “A Semidefinite Programming Approach to Side Chain Positioning with New Rounding Strategies.” *Informs Journal on Computing* 16 (4): 380–92. https://doi.org/10.1287/ijoc.1040.0096.

Cheatham, T. E., P. Cieplak, and P. A. Kollman. 1999. “A Modified Version of the Cornell et Al. Force Field with Improved Sugar Pucker Phases and Helical Repeat.” *Journal of Biomolecular Structure & Dynamics* 16 (4): 845–62. https://doi.org/10.1080/07391102.1999.10508297.

Chelliah, Vijayalakshmi, Tom L. Blundell, and Juan Fernández-Recio. 2006. “Efficient Restraints for Protein-Protein Docking by Comparison of Observed Amino Acid Substitution Patterns with Those Predicted from Local Environment.” *Journal of Molecular Biology* 357 (5): 1669–82. https://doi.org/10.1016/j.jmb.2006.01.001.

Chen, Rong, Li Li, and Zhiping Weng. 2003. “ZDOCK: An Initial-Stage Protein-Docking Algorithm.” *Proteins* 52 (1): 80–87. https://doi.org/10.1002/prot.10389.

Chen, Rong, Julian Mintseris, Joël Janin, and Zhiping Weng. 2003. “A Protein–Protein Docking Benchmark.” *Proteins: Structure, Function, and Bioinformatics* 52 (1): 88–91. https://doi.org/10.1002/prot.10390.

Chen, Rong, and Zhiping Weng. 2002. “Docking Unbound Proteins Using Shape Complementarity, Desolvation, and Electrostatics.” *Proteins: Structure, Function, and Bioinformatics* 47 (3): 281–94. https://doi.org/10.1002/prot.10092.

———. 2003. “A Novel Shape Complementarity Scoring Function for Protein-Protein Docking.” *Proteins* 51 (3): 397–408. https://doi.org/10.1002/prot.10334.

Cheng, Tammy M. K., Tom L. Blundell, and Juan Fernandez-Recio. 2008. “Structural Assembly of Two-Domain Proteins by Rigid-Body Docking.” *BMC Bioinformatics* 9 (October): 441. https://doi.org/10.1186/1471-2105-9-441.

Cheng, Tammy Man-Kuang, Tom L. Blundell, and Juan Fernandez-Recio. 2007. "PyDock: Electrostatics and Desolvation for Effective Scoring of Rigid-Body Protein–Protein Docking." *Proteins: Structure, Function, and Bioinformatics* 68 (2): 503–15. https://doi.org/10.1002/prot.21419.

Chothia, C, and A M Lesk. 1986. "The Relation between the Divergence of Sequence and Structure in Proteins." *The EMBO Journal* 5 (4): 823–26.

Choura, Mouna, and Ahmed Rebaï. 2012. "Topological Features of Cancer Proteins in the Human NR-RTK Interaction Network." *Journal of Receptor and Signal Transduction Research* 32 (5): 257–62. https://doi.org/10.3109/10799893.2012.702116.

Conte, Loredana Lo, Cyrus Chothia, and Joël Janin. 1999. "The Atomic Structure of Protein-Protein Recognition Sites." *Journal of Molecular Biology* 285 (5): 2177–98. https://doi.org/10.1006/jmbi.1998.2439.

Cornell, Wendy D., Piotr Cieplak, Christopher I. Bayly, Ian R. Gould, Kenneth M. Merz, David M. Ferguson, David C. Spellmeyer, Thomas Fox, James W. Caldwell, and Peter A. Kollman. 1995. "A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules." *Journal of the American Chemical Society* 117 (19): 5179–97. https://doi.org/10.1021/ja00124a002.

Csermely, Peter, Robin Palotai, and Ruth Nussinov. 2010. "Induced Fit, Conformational Selection and Independent Dynamic Segments: An Extended View of Binding Events." *Trends in Biochemical Sciences* 35 (10): 539–46. https://doi.org/10.1016/j.tibs.2010.04.009.

Cukuroglu, Engin, H Billur Engin, Attila Gursoy, and Ozlem Keskin. 2014. "Hot Spots in Protein–Protein Interfaces: Towards Drug Discovery." *Prog. Biophys. Mol. Biol.* 116 (2): 165–73. https://doi.org/10.1016/j.pbiomolbio.2014.06.003.

Daura, Xavier, Karl Gademann, Bernhard Jaun, Dieter Seebach, Wilfred F. van Gunsteren, and Alan E. Mark. 1999. "Peptide Folding: When Simulation Meets Experiment." *Angewandte Chemie International Edition* 38 (1–2): 236–40. https://doi.org/10.1002/(SICI)1521-3773(19990115)38:1/2<236::AID-ANIE236>3.0.CO;2-M.

DeLano, Warren L. 2002. "Unraveling Hot Spots in Binding Interfaces: Progress and Challenges." *Current Opinion in Structural Biology* 12 (1): 14–20. https://doi.org/16/S0959-440X(02)00283-X.

Desmet, J., M. De Maeyer, B. Hazes, and I. Lasters. 1992. "The Dead-End Elimination Theorem and Its Use in Protein Side-Chain Positioning." *Nature* 356 (6369): 539–42.

de Vries, Sjoerd J., Isaure Chauvot de Beauchêne, Christina E. M. Schindler, and Martin Zacharias. 2016. "Cryo-EM Data Are Superior to Contact and Interface Information in Integrative Modeling." *Biophysical Journal* 110 (4): 785–97. https://doi.org/10.1016/j.bpj.2015.12.038.

Dobbins, Sara E, Victor I Lesk, and Michael J E Sternberg. 2008. "Insights into Protein Flexibility: The Relationship between Normal Modes and Conformational Change upon Protein-Protein Docking." *Proc. Natl. Acad. Sci. U. S. A.* 105 (30): 10390–95. https://doi.org/10.1073/pnas.0802496105.

Dominguez, Cyril, Rolf Boelens, and Alexandre M. J. J. Bonvin. 2003. "HADDOCK: A Protein-Protein Docking Approach Based on Biochemical or Biophysical Information." *Journal of the American Chemical Society* 125 (7): 1731–37. https://doi.org/10.1021/ja026939x.

Dominguez, Cyril, Mario Schubert, Olivier Duss, Sapna Ravindranathan, and Frédéric H-T Allain. 2011. "Structure Determination and Dynamics of Protein-RNA Complexes by NMR Spectroscopy." *Prog. Nucl. Magn. Reson. Spectrosc.* 58 (1–2): 1–61. https://doi.org/10.1016/j.pnmrs.2010.10.001.

Doruker, Pemra, Ali Rana Atilgan, and Ivet Bahar. 2000. "Dynamics of Proteins Predicted by Molecular Dynamics Simulations and Analytical Approaches: Application to α-Amylase Inhibitor." *Proteins: Structure, Function, and Bioinformatics* 40 (3): 512–24. https://doi.org/10.1002/1097-0134(20000815)40:3<512::AID-PROT180>3.0.CO;2-M.

Douguet, Dominique, Huei-Chi Chen, Andrey Tovchigrechko, and Ilya A. Vakser. 2006. "DOCKGROUND Resource for Studying Protein-Protein Interfaces." *Bioinformatics (Oxford, England)* 22 (21): 2612–18. https://doi.org/10.1093/bioinformatics/btl447.

Echols, Nathaniel, Duncan Milburn, and Mark Gerstein. 2003. "MolMovDB: Analysis and Visualization of Conformational Change and Structural Flexibility." *Nucleic Acids Research* 31 (1): 478–82.

Ekiert, Damian C., Gira Bhabha, Marc-André Elsliger, Robert H. E. Friesen, Mandy Jongeneelen, Mark Throsby, Jaap Goudsmit, and Ian A. Wilson. 2009. "Antibody Recognition of a Highly Conserved Influenza Virus Epitope." *Science (New York, N.Y.)* 324 (5924): 246–51. https://doi.org/10.1126/science.1171491.

Emekli, Ugur, Dina Schneidman-Duhovny, Haim J. Wolfson, Ruth Nussinov, and Turkan Haliloglu. 2008. "HingeProt: Automated Prediction of Hinges in Protein Structures." *Proteins* 70 (4): 1219–27. https://doi.org/10.1002/prot.21613.

Esteban-Martín, Santiago, Robert Bryn Fenwick, and Xavier Salvatella. 2012. "Synergistic Use of NMR and MD Simulations to Study the Structural Heterogeneity of Proteins." *Wiley Interdisciplinary Reviews: Computational Molecular Science* 2 (3): 466–78. https://doi.org/10.1002/wcms.1093.

Ewing, Rob M., Peter Chu, Fred Elisma, Hongyan Li, Paul Taylor, Shane Climie, Linda McBroom-Cerajewski, et al. 2007. "Large-Scale Mapping of Human Protein-Protein Interactions by Mass Spectrometry." *Molecular Systems Biology* 3: 89. https://doi.org/10.1038/msb4100134.

Feliu, Elisenda, and Baldomero Oliva. 2010. "How Different from Random Are Docking Predictions When Ranked by Scoring Functions?" *Proteins* 78 (16): 3376–85. https://doi.org/10.1002/prot.22844.

Fernández-Recio, Juan, Maxim Totrov, and Ruben Abagyan. 2003. "ICM-DISCO Docking by Global Energy Optimization with Fully Flexible Side-Chains." *Proteins* 52 (1): 113–17. https://doi.org/10.1002/prot.10383.

Feyfant, Eric, Andrej Sali, and András Fiser. 2007. "Modeling Mutations in Protein Structures." *Protein Science* 16 (9): 2030–41. https://doi.org/10.1110/ps.072855507.

Fischer, Emil. 1894. "Einfluss Der Configuration Auf Die Wirkung Der Enzyme." *Berichte Der Deutschen Chemischen Gesellschaft* 27 (3): 2985–93. https://doi.org/10.1002/cber.18940270364.

Fleishman, Sarel J., Timothy A. Whitehead, Damian C. Ekiert, Cyrille Dreyfus, Jacob E. Corn, Eva-Maria Strauch, Ian A. Wilson, and David Baker. 2011. "Computational Design of Proteins Targeting the Conserved Stem Region of Influenza Hemagglutinin." *Science (New York, N.Y.)* 332 (6031): 816–21. https://doi.org/10.1126/science.1202617.

Fleishman, Sarel J., Timothy A. Whitehead, Eva-Maria Strauch, Jacob E. Corn, Sanbo Qin, Huan-Xiang Zhou, Julie C. Mitchell, et al. 2011. "Community-Wide Assessment of Protein-Interface Modeling Suggests Improvements to Design Methodology." *Journal of Molecular Biology* 414 (2): 289–302. https://doi.org/10.1016/j.jmb.2011.09.031.

Flores, Samuel, Nathaniel Echols, Duncan Milburn, Brandon Hespenheide, Kevin Keating, Jason Lu, Stephen Wells, Eric Z. Yu, Michael Thorpe, and Mark Gerstein. 2006. "The Database of Macromolecular Motions: New Features Added at the Decade Mark." *Nucleic Acids Research* 34 (Database issue): D296-301. https://doi.org/10.1093/nar/gkj046.

Frauenfelder, H., S. G. Sligar, and P. G. Wolynes. 1991. "The Energy Landscapes and Motions of Proteins." *Science* 254 (5038): 1598–1603. https://doi.org/10.1126/science.1749933.

Fulle, Simone, and Holger Gohlke. 2008. "Analyzing the Flexibility of RNA Structures by Constraint Counting." *Biophys. J.* 94 (11): 4202–19. https://doi.org/10.1529/biophysj.107.113415.

Gabb, Henry A., Richard M. Jackson, and Michael J.E. Sternberg. 1997. "Modelling Protein Docking Using Shape Complementarity, Electrostatics and Biochemical Information." *Journal of Molecular Biology* 272 (1): 106–20. https://doi.org/10.1006/jmbi.1997.1203.

Gao, Mu, and Jeffrey Skolnick. 2010. "Structural Space of Protein-Protein Interfaces Is Degenerate, Close to Complete, and Highly Connected." *Proceedings of the National Academy of Sciences of the United States of America* 107 (52): 22517–22. https://doi.org/10.1073/pnas.1012820107.

Gao, Ying, Dominique Douguet, Andrey Tovchigrechko, and Ilya A. Vakser. 2007. "DOCKGROUND System of Databases for Protein Recognition Studies: Unbound Structures for Docking." *Proteins* 69 (4): 845–51. https://doi.org/10.1002/prot.21714.

Gårdsvoll, Henrik, Bernard Gilquin, Marie Hélène Le Du, Andre Ménèz, Thomas J. D. Jørgensen, and Michael Ploug. 2006. "Characterization of the Functional Epitope on the Urokinase Receptor. Complete Alanine Scanning Mutagenesis Supplemented by Chemical Cross-Linking." *The Journal of Biological Chemistry* 281 (28): 19260–72. https://doi.org/10.1074/jbc.M513583200.

Garma, Leonardo, Srayanta Mukherjee, Pralay Mitra, and Yang Zhang. 2012. "How Many Protein-Protein Interactions Types Exist in Nature?" *PloS One* 7 (6): e38913. https://doi.org/10.1371/journal.pone.0038913.

Grosdidier, Solene, and Juan Fernandez-Recio. 2008. "Identification of Hot-Spot Residues in Protein-Protein Interactions by Computational Docking." *BMC Bioinformatics* 9 (1): 447. https://doi.org/10.1186/1471-2105-9-447.

Grosdidier, Solène, Carles Pons, Albert Solernou, and Juan Fernández-Recio. 2007. "Prediction and Scoring of Docking Poses with PyDock." *Proteins* 69 (4): 852–58. https://doi.org/10.1002/prot.21796.

Grünberg, Raik, Johan Leckner, and Michael Nilges. 2004. "Complementarity of Structure Ensembles in Protein-Protein Binding." *Structure (London, England: 1993)* 12 (12): 2125–36. https://doi.org/10.1016/j.str.2004.09.014.

Guerler, Aysam, Brandon Govindarajoo, and Yang Zhang. 2013. "Mapping Monomeric Threading to Protein-Protein Structure Prediction." *Journal of Chemical Information and Modeling* 53 (3): 717–25. https://doi.org/10.1021/ci300579r.

Guerois, Raphael, Jens Erik Nielsen, and Luis Serrano. 2002. "Predicting Changes in the Stability of Proteins and Protein Complexes: A Study of More Than 1000 Mutations." *Journal of Molecular Biology* 320 (2): 369–87. https://doi.org/10.1016/S0022-2836(02)00442-4.

Guilhot-Gaudeffroy, Adrien, Christine Froidevaux, Jerome Aze, and Julie Bernauer. 2014. "Protein-RNA Complexes and Efficient Automatic Docking: Expanding RosettaDock Possibilities." *PloS One* 9 (9): e108928. https://doi.org/10.1371/journal.pone.0108928.

Gutmanas, Aleksandras, Younes Alhroub, Gary M. Battle, John M. Berrisford, Estelle Bochet, Matthew J. Conroy, Jose M. Dana, et al. 2014. "PDBe:

Protein Data Bank in Europe." *Nucleic Acids Research* 42 (Database issue): D285-291. https://doi.org/10.1093/nar/gkt1180.

Haliloglu, Turkan, Ivet Bahar, and Burak Erman. 1997. "Gaussian Dynamics of Folded Proteins." *Physical Review Letters* 79 (16): 3090–93. https://doi.org/10.1103/PhysRevLett.79.3090.

Halperin, Inbal, Buyong Ma, Haim Wolfson, and Ruth Nussinov. 2002. "Principles of Docking: An Overview of Search Algorithms and a Guide to Scoring Functions." *Proteins: Structure, Function, and Bioinformatics* 47 (4): 409–43. https://doi.org/10.1002/prot.10115.

Helland, Ronny, Renate Larsen, Solrun Finstad, Peter Kyomuhendo, and Atle Larsen. 2009. "Crystal Structures of G-Type Lysozyme from Atlantic Cod Shed New Light on Substrate Binding and the Catalytic Mechanism." *Cellular and Molecular Life Sciences* 66 (15): 2585–98. https://doi.org/10.1007/s00018-009-0063-x.

Henzler-Wildman, Katherine A, Ming Lei, Vu Thai, S Jordan Kerns, Martin Karplus, and Dorothee Kern. 2007. "A Hierarchy of Timescales in Protein Dynamics Is Linked to Enzyme Catalysis." *Nature* 450 (7171): 913–16. https://doi.org/10.1038/nature06407.

Henzler-Wildman, Katherine, and Dorothee Kern. 2007. "Dynamic Personalities of Proteins." *Nature* 450 (7172): 964–72. https://doi.org/10.1038/nature06522.

Honig, B., and A. Nicholls. 1995. "Classical Electrostatics in Biology and Chemistry." *Science (New York, N.Y.)* 268 (5214): 1144–49.

Horton, N., and M. Lewis. 1992. "Calculation of the Free Energy of Association for Protein Complexes." *Protein Science: A Publication of the Protein Society* 1 (1): 169–81. https://doi.org/10.1002/pro.5560010117.

Hosur, Raghavendra, Jian Peng, Arunachalam Vinayagam, Ulrich Stelzl, Jinbo Xu, Norbert Perrimon, Jadwiga Bienkowska, and Bonnie Berger. 2012. "A Computational Framework for Boosting Confidence in High-Throughput Protein-Protein Interaction Datasets." *Genome Biology* 13 (8): R76. https://doi.org/10.1186/gb-2012-13-8-r76.

Huang, Sheng-You, and Xiaoqin Zou. 2014. "A Knowledge-Based Scoring Function for Protein-RNA Interactions Derived from a Statistical Mechanics-Based Iterative Method." *Nucleic Acids Res.* 42 (7): e55. https://doi.org/10.1093/nar/gku077.

Hwang, Howook, Brian Pierce, Julian Mintseris, Joël Janin, and Zhiping Weng. 2008. "Protein-Protein Docking Benchmark Version 3.0." *Proteins* 73 (3): 705–9. https://doi.org/10.1002/prot.22106.

Hwang, Howook, Thom Vreven, Joël Janin, and Zhiping Weng. 2010. "Protein–Protein Docking Benchmark Version 4.0." *Proteins: Structure, Function, and Bioinformatics* 78 (15): 3111–3114. https://doi.org/10.1002/prot.22830.

Hyeon, Changbong, Ruxandra I Dima, and D Thirumalai. 2006. "Size, Shape, and Flexibility of RNA Structures." *J. Chem. Phys.* 125 (19): 194905. https://doi.org/10.1063/1.2364190.

Ito, T., T. Chiba, R. Ozawa, M. Yoshida, M. Hattori, and Y. Sakaki. 2001. "A Comprehensive Two-Hybrid Analysis to Explore the Yeast Protein Interactome." *Proceedings of the National Academy of Sciences of the United States of America* 98 (8): 4569–74. https://doi.org/10.1073/pnas.061034498.

Jacobs, D. J., A. J. Rader, L. A. Kuhn, and M. F. Thorpe. 2001. "Protein Flexibility Predictions Using Graph Theory." *Proteins* 44 (2): 150–65.

Janin, Joël. 2010a. "Protein–Protein Docking Tested in Blind Predictions: The CAPRI Experiment." *Molecular BioSystems* 6 (12): 2351–62. https://doi.org/10.1039/C005060C.

———. 2010b. "The Targets of CAPRI Rounds 13-19." *Proteins* 78 (15): 3067–72. https://doi.org/10.1002/prot.22774.

Janin, Joël, Kim Henrick, John Moult, Lynn Ten Eyck, Michael J. E. Sternberg, Sandor Vajda, Ilya Vakser, Shoshana J. Wodak, and Critical Assessment of PRedicted Interactions. 2003. "CAPRI: A Critical Assessment of PRedicted Interactions." *Proteins* 52 (1): 2–9. https://doi.org/10.1002/prot.10381.

Jiménez-García, Brian, Carles Pons, and Juan Fernández-Recio. 2013. "PyDockWEB: A Web Server for Rigid-Body Protein-Protein Docking Using Electrostatics and Desolvation Scoring." *Bioinformatics (Oxford, England)* 29 (13): 1698–99. https://doi.org/10.1093/bioinformatics/btt262.

Jiménez-García, Brian, Carles Pons, Dmitri I. Svergun, Pau Bernadó, and Juan Fernández-Recio. 2015. "PyDockSAXS: Protein-Protein Complex Structure by SAXS and Computational Docking." *Nucleic Acids Research*, April. https://doi.org/10.1093/nar/gkv368.

Jiménez-García, Brian, Jorge Roel-Touris, Miguel Romero-Durana, Miquel Vidal, Daniel Jiménez-González, and Juan Fernández-Recio. 2018a. "LightDock: A New Multi-Scale Approach to Protein-Protein Docking." *Bioinformatics (Oxford, England)* 34 (1): 49–55. https://doi.org/10.1093/bioinformatics/btx555.

———. 2018b. "LightDock: A New Multi-Scale Approach to Protein–Protein Docking." *Bioinformatics* 34 (1): 49–55. https://doi.org/10.1093/bioinformatics/btx555.

Jin, L, and J A Wells. 1994. "Dissecting the Energetics of an Antibody-Antigen Interface by Alanine Shaving and Molecular Grafting." *Protein Sci.* 3 (12): 2351–57. https://doi.org/10.1002/pro.5560031219.

Jones, Eric, Travis Oliphant, and Pearu Peterson. 2015. "SciPy: Open Source Scientific Tools for Python, 2001." *URL Http://Www. Scipy. Org* 73: 86.

Jones, S., and J. M. Thornton. 1996. "Principles of Protein-Protein Interactions." *Proceedings of the National Academy of Sciences* 93 (1): 13–20.

Jones, Susan. 2016. "Protein-RNA Interactions: Structural Biology and Computational Modeling Techniques." *Biophys. Rev.* 8 (4): 359–67. https://doi.org/10.1007/s12551-016-0223-9.

Jones, Susan, and Janet M. Thornton. 1997. "Analysis of Protein-Protein Interaction Sites Using Surface Patches1." *Journal of Molecular Biology* 272 (1): 121–32. https://doi.org/10.1006/jmbi.1997.1234.

Jonsson, Pall F., and Paul A. Bates. 2006. "Global Topological Features of Cancer Proteins in the Human Interactome." *Bioinformatics (Oxford, England)* 22 (18): 2291–97. https://doi.org/10.1093/bioinformatics/btl390.

Kastritis, Panagiotis L, and Alexandre M J J Bonvin. 2013. "On the Binding Affinity of Macromolecular Interactions: Daring to Ask Why Proteins Interact." *J. R. Soc. Interface* 10 (79): 20120835. https://doi.org/10.1098/rsif.2012.0835.

Kastritis, Panagiotis L, and Alexandre Mjj Bonvin. 2010. "Are Scoring Functions in Protein- Protein Docking Ready to Predict Interactomes? Clues from a Novel Binding Affinity Benchmark." *J. Proteome Res.* 9 (5): 2216–25.

Kastritis, Panagiotis L., Iain H. Moal, Howook Hwang, Zhiping Weng, Paul A. Bates, Alexandre M. J. J. Bonvin, and Joël Janin. 2011. "A Structure-Based Benchmark for Protein–Protein Binding Affinity." *Protein Science* 20 (3): 482–91. https://doi.org/10.1002/pro.580.

Katchalski-Katzir, E, I Shariv, M Eisenstein, A A Friesem, C Aflalo, and I A Vakser. 1992. "Molecular Surface Recognition: Determination of Geometric Fit between Proteins and Their Ligands by Correlation Techniques." *Proceedings of the National Academy of Sciences of the United States of America* 89 (6): 2195–99.

Ke, Ailong, and Jennifer A Doudna. 2004. "Crystallization of RNA and RNA-Protein Complexes." *Methods* 34 (3): 408–14. https://doi.org/10.1016/j.ymeth.2004.03.027.

Keskin, Ozlem, Buyong Ma, and Ruth Nussinov. 2005. "Hot Regions in Protein–Protein Interactions: The Organization and Contribution of Structurally Conserved Hot Spot Residues." *Journal of Molecular Biology* 345 (5): 1281–94. https://doi.org/10.1016/j.jmb.2004.10.077.

Keskin, Ozlem, Ruth Nussinov, and Attila Gursoy. 2008. "PRISM: Protein-Protein Interaction Prediction by Structural Matching." *Methods in Molecular Biology (Clifton, N.J.)* 484: 505–21. https://doi.org/10.1007/978-1-59745-398-1_30.

Khatter, Heena, Alexander G. Myasnikov, S. Kundhavai Natchiar, and Bruno P. Klaholz. 2015. "Structure of the Human 80S Ribosome." *Nature* 520 (7549): 640–45. https://doi.org/10.1038/nature14427.

Kiel, Christina, and Luis Serrano. 2014. "Structure-energy-based Predictions and Network Modelling of RASopathy and Cancer Missense Mutations." *Molecular Systems Biology* 10 (5): 727. https://doi.org/10.1002/msb.20145092.

Klepeis, John L., Kresten Lindorff-Larsen, Ron O. Dror, and David E. Shaw. 2009. "Long-Timescale Molecular Dynamics Simulations of Protein Structure and Function." *Current Opinion in Structural Biology* 19 (2): 120–27. https://doi.org/10.1016/j.sbi.2009.03.004.

Kollman, Peter. 1993. "Free Energy Calculations: Applications to Chemical and Biochemical Phenomena." *Chemical Reviews* 93 (7): 2395–2417. https://doi.org/10.1021/cr00023a004.

Kondrashov, Dmitry A., Qiang Cui, and George N. Phillips. 2006. "Optimization and Evaluation of a Coarse-Grained Model of Protein Motion Using x-Ray Crystal Data." *Biophysical Journal* 91 (8): 2760–67. https://doi.org/10.1529/biophysj.106.085894.

Kortemme, Tanja, and David Baker. 2004. "Computational Design of Protein-Protein Interactions." *Current Opinion in Chemical Biology* 8 (1): 91–97. https://doi.org/10.1016/j.cbpa.2003.12.008.

Koshland, D. E. 1958. "Application of a Theory of Enzyme Specificity to Protein Synthesis." *Proceedings of the National Academy of Sciences of the United States of America* 44 (2): 98–104.

Kozakov, D., R. Brenke, S. Comeau, and S. Vajda. 2006. "PIPER: An FFT-Based Protein Docking Program with Pairwise Potentials." http://arxiv.org/abs/q-bio/0605018.

Krishnan, V. V., and B. Rupp. 2012. "Macromolecular Structure Determination: Comparison of X-ray Crystallography and NMR Spectroscopy." *ELS*.

Krishnanand, K. N., and D. Ghose. 2009. "Glowworm Swarm Optimization for Simultaneous Capture of Multiple Local Optima of Multimodal Functions." *Swarm Intelligence* 3 (2): 87–124. https://doi.org/10.1007/s11721-008-0021-5.

Krissinel, Evgeny, and Kim Henrick. 2007. "Inference of Macromolecular Assemblies from Crystalline State." *Journal of Molecular Biology* 372 (3): 774–97. https://doi.org/10.1016/j.jmb.2007.05.022.

Krivov, Georgii G., Maxim V. Shapovalov, and Roland L. Dunbrack. 2009. "Improved Prediction of Protein Side-Chain Conformations with SCWRL4." *Proteins* 77 (4): 778–95. https://doi.org/10.1002/prot.22488.

Krowarsch, Daniel, Michal Dadlez, Olga Buczek, Izabela Krokoszynska, Arne O. Smalas, and Jacek Otlewski. 1999. "Interscaffolding Additivity: Binding of P1 Variants of Bovine Pancreatic Trypsin Inhibitor to Four Serine

Proteases11Edited by R. Huber." *Journal of Molecular Biology* 289 (1): 175–86. https://doi.org/10.1006/jmbi.1999.2757.

Kühlbrandt, Werner. 2014. "Cryo-EM Enters a New Era." *ELife* 3 (August): e03678. https://doi.org/10.7554/eLife.03678.

Kühlmann, U. C., A. J. Pommer, G. R. Moore, R. James, and C. Kleanthous. 2000. "Specificity in Protein-Protein Interactions: The Structural Basis for Dual Recognition in Endonuclease Colicin-Immunity Protein Complexes." *Journal of Molecular Biology* 301 (5): 1163–78. https://doi.org/10.1006/jmbi.2000.3945.

Kumar, Manoj, Sujata Sharma, Alagiri Srinivasan, Tej P. Singh, and Punit Kaur. 2011. "Structure-Based in-Silico Rational Design of a Selective Peptide Inhibitor for Thymidine Monophosphate Kinase of Mycobacterium Tuberculosis." *Journal of Molecular Modeling* 17 (5): 1173–82. https://doi.org/10.1007/s00894-010-0821-6.

Kundrotas, Petras J., Ilya A. Vakser, and Joël Janin. 2013. "Structural Templates for Modeling Homodimers." *Protein Science: A Publication of the Protein Society* 22 (11): 1655–63. https://doi.org/10.1002/pro.2361.

Kundrotas, Petras J., Zhengwei Zhu, Joel Janin, and Ilya A. Vakser. 2012. "Templates Are Available to Model Nearly All Complexes of Structurally Characterized Proteins." *Proceedings of the National Academy of Sciences of the United States of America* 109 (24): 9438–41. https://doi.org/10.1073/pnas.1200678109.

Kundu, Sibsankar, Julia S. Melton, Dan C. Sorensen, and George N. Phillips. 2002. "Dynamics of Proteins in Crystals: Comparison of Experiment with Simple Models." *Biophysical Journal* 83 (2): 723–32. https://doi.org/10.1016/S0006-3495(02)75203-X.

Kyomuhendo, Peter, Bjørnar Myrnes, Bjørn-Olav Brandsdal, Arne O. Smalås, Inge W. Nilsen, and Ronny Helland. 2010. "Thermodynamics and Structure of a Salmon Cold Active Goose-Type Lysozyme." *Comparative Biochemistry and Physiology. Part B, Biochemistry & Molecular Biology* 156 (4): 254–63. https://doi.org/10.1016/j.cbpb.2010.04.002.

Ladbury, J. E., and B. Z. Chowdhry. 1996. "Sensing the Heat: The Application of Isothermal Titration Calorimetry to Thermodynamic Studies of Biomolecular Interactions." *Chemistry & Biology* 3 (10): 791–801.

Landon, Melissa R., David R. Lancia, Jessamin Yu, Spencer C. Thiel, and Sandor Vajda. 2007. "Identification of Hot Spots within Druggable Binding Regions by Computational Solvent Mapping of Proteins." *Journal of Medicinal Chemistry* 50 (6): 1231–40. https://doi.org/10.1021/jm061134b.

Lehninger, Albert L., David L. Nelson, and Michael M. Cox. 2008. *Lehninger Principles of Biochemistry*. 5th ed. New York: W.H. Freeman.

Lensink, Marc F., Raúl Méndez, and Shoshana J. Wodak. 2007. "Docking and Scoring Protein Complexes: CAPRI 3rd Edition." *Proteins: Structure, Function, and Bioinformatics* 69 (4): 704–718. https://doi.org/10.1002/prot.21804.

Lensink, Marc F., Iain H. Moal, Paul A. Bates, Panagiotis L. Kastritis, Adrien S. J. Melquiond, Ezgi Karaca, Christophe Schmitz, et al. 2014. "Blind Prediction of Interfacial Water Positions in CAPRI." *Proteins* 82 (4): 620–32. https://doi.org/10.1002/prot.24439.

Lensink, Marc F., Sameer Velankar, Andriy Kryshtafovych, Shen-You Huang, Dina Schneidman-Duhovny, Andrej Sali, Joan Segura, et al. 2016. "Prediction of Homoprotein and Heteroprotein Complexes by Protein Docking and Template-Based Modeling: A CASP-CAPRI Experiment." *Proteins: Structure, Function, and Bioinformatics* 84 (September): 323–48. https://doi.org/10.1002/prot.25007.

Lensink, Marc F., and Shoshana J. Wodak. 2010a. "Docking and Scoring Protein Interactions: CAPRI 2009." *Proteins: Structure, Function, and Bioinformatics* 78 (15): 3073–3084. https://doi.org/10.1002/prot.22818.

———. 2010b. "Blind Predictions of Protein Interfaces by Docking Calculations in CAPRI." *Proteins: Structure, Function, and Bioinformatics* 78 (15): 3085–95. https://doi.org/10.1002/prot.22850.

———. 2013a. "Docking, Scoring, and Affinity Prediction in CAPRI." *Proteins* 81 (12): 2082–95. https://doi.org/10.1002/prot.24428.

———. 2013b. "Docking, Scoring, and Affinity Prediction in CAPRI." *Proteins* 81 (12): 2082–95. https://doi.org/10.1002/prot.24428.

Leysen, Seppe, Lise Vanderkelen, Kathleen Van Asten, Steven Vanheuverzwijn, Veronique Theuwis, Chris W. Michiels, and Sergei V. Strelkov. 2012. "Structural Characterization of the PliG Lysozyme Inhibitor Family." *Journal of Structural Biology* 180 (1): 235–42. https://doi.org/10.1016/j.jsb.2012.05.006.

Li, Xiaofan, Iain H. Moal, and Paul A. Bates. 2010. "Detection and Refinement of Encounter Complexes for Protein-Protein Docking: Taking Account of Macromolecular Crowding." *Proteins* 78 (15): 3189–96. https://doi.org/10.1002/prot.22770.

Liao, Maofu, Erhu Cao, David Julius, and Yifan Cheng. 2013. "Structure of the TRPV1 Ion Channel Determined by Electron Cryo-Microscopy." *Nature* 504 (7478): 107–12. https://doi.org/10.1038/nature12822.

Lichtarge, Olivier, Henry R. Bourne, and Fred E. Cohen. 1996. "An Evolutionary Trace Method Defines Binding Surfaces Common to Protein Families." *Journal of Molecular Biology* 257 (2): 342–58. https://doi.org/10.1006/jmbi.1996.0167.

Lindahl, Erik, and Marc Delarue. 2005. "Refinement of Docked Protein-Ligand and Protein-DNA Structures Using Low Frequency Normal Mode

Amplitude Optimization." *Nucleic Acids Research* 33 (14): 4496–4506. https://doi.org/10.1093/nar/gki730.

Liu, Shiyong, Ying Gao, and Ilya A. Vakser. 2008. "Dockground Protein–Protein Docking Decoy Set." *Bioinformatics* 24 (22): 2634–35. https://doi.org/10.1093/bioinformatics/btn497.

Liu, Song, Chi Zhang, Hongyi Zhou, and Yaoqi Zhou. 2004. "A Physical Reference State Unifies the Structure-Derived Potential of Mean Force for Protein Folding and Binding." *Proteins* 56 (1): 93–101. https://doi.org/10.1002/prot.20019.

Lockless, Steve W., and Rama Ranganathan. 1999. "Evolutionarily Conserved Pathways of Energetic Connectivity in Protein Families." *Science* 286 (5438): 295–99. https://doi.org/10.1126/science.286.5438.295.

Lodish, Harvey, Arnold Berk, S. Lawrence Zipursky, Paul Matsudaira, David Baltimore, and James Darnell. 2000. *Molecular Cell Biology*. 4th ed. W. H. Freeman.

Lovell, S. C., J. M. Word, J. S. Richardson, and D. C. Richardson. 2000. "The Penultimate Rotamer Library." *Proteins* 40 (3): 389–408.

Lu, Long, Adrian K. Arakaki, Hui Lu, and Jeffrey Skolnick. 2003. "Multimeric Threading-Based Prediction of Protein-Protein Interactions on a Genomic Scale: Application to the Saccharomyces Cerevisiae Proteome." *Genome Research* 13 (6A): 1146–54. https://doi.org/10.1101/gr.1145203.

Ma, B., S. Kumar, C.-J. Tsai, and R. Nussinov. 1999. "Folding Funnels and Binding Mechanisms." *Protein Engineering Design and Selection* 12 (9): 713–20. https://doi.org/10.1093/protein/12.9.713.

Ma, Buyong, Tal Elkayam, Haim Wolfson, and Ruth Nussinov. 2003. "Protein–Protein Interactions: Structurally Conserved Residues Distinguish between Binding Sites and Exposed Protein Surfaces." *Proceedings of the National Academy of Sciences of the United States of America* 100 (10): 5772–77. https://doi.org/10.1073/pnas.1030237100.

Ma, Jianpeng. 2004. "New Advances in Normal Mode Analysis of Supermolecular Complexes and Applications to Structural Refinement." *Current Protein & Peptide Science* 5 (2): 119–23.

Ma, Xiao Hui, Cun Xin Wang, Chun Hua Li, and Wei Zu Chen. 2002. "A Fast Empirical Approach to Binding Free Energy Calculations Based on Protein Interface Information." *Protein Engineering* 15 (8): 677–81.

Masi, Alessio, Riccardo Cicchi, Adolfo Carloni, Francesco Saverio Pavone, and Annarosa Arcangeli. 2010. "Optical Methods in the Study of Protein-Protein Interactions." *Advances in Experimental Medicine and Biology* 674: 33–42.

May, Andreas, and Martin Zacharias. 2008. "Energy Minimization in Low-Frequency Normal Modes to Efficiently Allow for Global Flexibility during Systematic Protein–Protein Docking." *Proteins: Structure,*

Function, and Bioinformatics 70 (3): 794–809. https://doi.org/10.1002/prot.21579.

McCammon, J. A., B. R. Gelin, and M. Karplus. 1977. "Dynamics of Folded Proteins." *Nature* 267 (5612): 585–90.

Meenan, Nicola A. G., Amit Sharma, Sarel J. Fleishman, Colin J. Macdonald, Bertrand Morel, Ruth Boetzel, Geoffrey R. Moore, David Baker, and Colin Kleanthous. 2010. "The Structural and Energetic Basis for High Selectivity in a High-Affinity Protein-Protein Interaction." *Proceedings of the National Academy of Sciences of the United States of America* 107 (22): 10080–85. https://doi.org/10.1073/pnas.0910756107.

Melo, Rita, Robert Fieldhouse, André Melo, João D. G. Correia, Maria Natália D. S. Cordeiro, Zeynep H. Gümüş, Joaquim Costa, Alexandre M. J. J. Bonvin, and Irina S. Moreira. 2016. "A Machine Learning Approach for Hot-Spot Detection at Protein-Protein Interfaces." *International Journal of Molecular Sciences* 17 (8). https://doi.org/10.3390/ijms17081215.

Méndez, Raúl, Raphaël Leplae, Leonardo De Maria, and Shoshana J. Wodak. 2003. "Assessment of Blind Predictions of Protein-Protein Interactions: Current Status of Docking Methods." *Proteins* 52 (1): 51–67. https://doi.org/10.1002/prot.10393.

Miller, Bill R., T. Dwight McGee, Jason M. Swails, Nadine Homeyer, Holger Gohlke, and Adrian E. Roitberg. 2012. "MMPBSA.Py: An Efficient Program for End-State Free Energy Calculations." *Journal of Chemical Theory and Computation* 8 (9): 3314–21. https://doi.org/10.1021/ct300418h.

Mintseris, Julian, Kevin Wiehe, Brian Pierce, Robert Anderson, Rong Chen, Joël Janin, and Zhiping Weng. 2005. "Protein-Protein Docking Benchmark 2.0: An Update." *Proteins* 60 (2): 214–16. https://doi.org/10.1002/prot.20560.

Mirzaei, Hanieh, Dmitri Beglov, Ioannis Ch. Paschalidis, Sandor Vajda, Pirooz Vakili, and Dima Kozakov. 2012. "Rigid Body Energy Minimization on Manifolds for Molecular Docking." *Journal of Chemical Theory and Computation* 8 (11): 4374–80. https://doi.org/10.1021/ct300272j.

Mitra, Pralay, and Debnath Pal. 2010. "New Measures for Estimating Surface Complementarity and Packing at Protein–Protein Interfaces." *FEBS Letters* 584 (6): 1163–68. https://doi.org/10.1016/j.febslet.2010.02.021.

Moal, Iain H, and Paul A Bates. 2010. "SwarmDock and the Use of Normal Modes in Protein-Protein Docking." *Int. J. Mol. Sci.* 11 (10): 3623–48. https://doi.org/10.3390/ijms11103623.

Moal, Iain H., and Juan Fernández-Recio. 2012a. "SKEMPI: A Structural Kinetic and Energetic Database of Mutant Protein Interactions and Its Use in Empirical Models." *Bioinformatics* 28 (20): 2600–2607. https://doi.org/10.1093/bioinformatics/bts489.

———. 2012b. “SKEMPI: A Structural Kinetic and Energetic Database of Mutant Protein Interactions and Its Use in Empirical Models.” *Bioinformatics* 28 (20): 2600–2607. https://doi.org/10.1093/bioinformatics/bts489.

Moe, Luke A., Craig A. Bingman, Gary E. Wesenberg, George N. Phillips, and Brian G. Fox. 2006. “Structure of T4moC, the Rieske-Type Ferredoxin Component of Toluene 4-Monooxygenase.” *Acta Crystallographica. Section D, Biological Crystallography* 62 (Pt 5): 476–82. https://doi.org/10.1107/S0907444906006056.

Moreira, Irina S., Pedro A. Fernandes, and Maria J. Ramos. 2007a. “Computational Alanine Scanning Mutagenesis--an Improved Methodological Approach.” *Journal of Computational Chemistry* 28 (3): 644–54. https://doi.org/10.1002/jcc.20566.

Moreira, Irina S, Pedro A Fernandes, and Maria J Ramos. 2007b. “Hot Spots--a Review of the Protein-Protein Interface Determinant Amino-Acid Residues.” *Proteins* 68 (4): 803–12. https://doi.org/10.1002/prot.21396.

Moretti, Rocco, Sarel J. Fleishman, Rudi Agius, Mieczyslaw Torchala, Paul A. Bates, Panagiotis L. Kastritis, João P. G. L. M. Rodrigues, et al. 2013. “Community-Wide Evaluation of Methods for Predicting the Effect of Mutations on Protein–Protein Interactions.” *Proteins: Structure, Function, and Bioinformatics* 81 (11): 1980–87. https://doi.org/10.1002/prot.24356.

Mosca, Roberto, Arnaud Céol, and Patrick Aloy. 2013. “Interactome3D: Adding Structural Details to Protein Networks.” *Nature Methods* 10 (1): 47–53. https://doi.org/10.1038/nmeth.2289.

Moult, John, Jan T. Pedersen, Richard Judson, and Krzysztof Fidelis. 1995. “A Large-Scale Experiment to Assess Protein Structure Prediction Methods.” *Proteins: Structure, Function, and Bioinformatics* 23 (3): ii–iv. https://doi.org/10.1002/prot.340230303.

Murzin, A. G., S. E. Brenner, T. Hubbard, and C. Chothia. 1995. “SCOP: A Structural Classification of Proteins Database for the Investigation of Sequences and Structures.” *Journal of Molecular Biology* 247 (4): 536–40. https://doi.org/10.1006/jmbi.1995.0159.

Navizet, Isabelle, Richard Lavery, and Robert L. Jernigan. 2004. “Myosin Flexibility: Structural Domains and Collective Vibrations.” *Proteins* 54 (3): 384–93. https://doi.org/10.1002/prot.10476.

Offman, Marc N., Paul W. Fitzjohn, and Paul A. Bates. 2006. “Developing a Move-Set for Protein Model Refinement.” *Bioinformatics (Oxford, England)* 22 (15): 1838–45. https://doi.org/10.1093/bioinformatics/btl192.

Ofran, Yanay, and Burkhard Rost. 2003. “Analysing Six Types of Protein-Protein Interfaces.” *Journal of Molecular Biology* 325 (2): 377–87.

———. 2007. “Protein–Protein Interaction Hotspots Carved into Sequences.” *PLoS Comput Biol* 3 (7): e119. https://doi.org/10.1371/journal.pcbi.0030119.

Orellana, Laura. 2014. “Protein Dynamics Studied by Coarse-Grained and Atomistic Theoretical Approaches.” Barcelona. Spain: University of Barcelona. http://diposit.ub.edu/dspace/handle/2445/59411.

Orellana, Laura, Manuel Rueda, Carles Ferrer-Costa, José Ramón Lopez-Blanco, Pablo Chacón, and Modesto Orozco. 2010. “Approaching Elastic Network Models to Molecular Dynamics Flexibility.” *Journal of Chemical Theory and Computation* 6 (9): 2910–23. https://doi.org/10.1021/ct100208e.

Orozco, Modesto. 2014. “A Theoretical View of Protein Dynamics.” *Chem. Soc. Rev.* 43 (14): 5051–66. https://doi.org/10.1039/c3cs60474h.

Ovchinnikov, Sergey, Hetunandan Kamisetty, and David Baker. 2014. “Robust and Accurate Prediction of Residue–Residue Interactions across Protein Interfaces Using Evolutionary Information.” *ELife* 3 (May): e02030. https://doi.org/10.7554/eLife.02030.

Pallara, Chiara, Brian Jiménez-García, Laura Pérez-Cano, Miguel Romero-Durana, Albert Solernou, Solène Grosdidier, Carles Pons, Iain H. Moal, and Juan Fernandez-Recio. 2013. “Expanding the Frontiers of Protein–Protein Modeling: From Docking and Scoring to Binding Affinity Predictions and Other Challenges.” *Proteins: Structure, Function, and Bioinformatics* 81 (12): 2192–2200. https://doi.org/10.1002/prot.24387.

Pallara, Chiara, Manuel Rueda, Ruben Abagyan, and Juan Fernández-Recio. 2016. “Conformational Heterogeneity of Unbound Proteins Enhances Recognition in Protein–Protein Encounters.” *Journal of Chemical Theory and Computation* 12 (7): 3236–49. https://doi.org/10.1021/acs.jctc.6b00204.

Pérez-Cano, Laura, and Juan Fernández-Recio. 2010. “Optimal Protein-RNA Area, OPRA: A Propensity-Based Method to Identify RNA-Binding Sites on Proteins.” *Proteins* 78 (1): 25–35. https://doi.org/10.1002/prot.22527.

Pérez-Cano, Laura, Brian Jiménez-García, and Juan Fernández-Recio. 2012. “A Protein-RNA Docking Benchmark (II): Extended Set from Experimental and Homology Modeling Data.” *Proteins* 80 (7): 1872–82. https://doi.org/10.1002/prot.24075.

Pérez-Cano, Laura, Albert Solernou, Carles Pons, and Juan Fernández-Recio. 2010. “Structural Prediction of Protein-RNA Interaction by Computational Docking with Propensity-Based Statistical Potentials.” *Pacific Symposium on Biocomputing. Pacific Symposium on Biocomputing*, 293–301.

Pérez-Cano, Laura, Albert Solernou, Carles Pons, and Juan Fern?ndez-Recio. 2009. “Structural Prediction of Protein-Rna Interaction by

Computational Docking with Propensity-Based Statistical Potentials." In *Biocomputing 2010*, 293–301. WORLD SCIENTIFIC. http://www.worldscientific.com.sire.ub.edu/doi/abs/10.1142/9789814295291_0031.

Petrey, Donald, and Barry Honig. 2003. "GRASP2: Visualization, Surface Properties, and Electrostatics of Macromolecular Structures and Sequences." *Methods in Enzymology* 374: 492–509. https://doi.org/10.1016/S0076-6879(03)74021-X.

Petrone, Paula, and Vijay S. Pande. 2006. "Can Conformational Change Be Described by Only a Few Normal Modes?" *Biophysical Journal* 90 (5): 1583–93. https://doi.org/10.1529/biophysj.105.070045.

Pfeiffenberger, Erik, Raphael A.G. Chaleil, Iain H. Moal, and Paul A. Bates. 2017. "A Machine Learning Approach for Ranking Clusters of Docked Protein-Protein Complexes by Pairwise Cluster Comparison." *Proteins: Structure, Function, and Bioinformatics* 85 (3): 528–43. https://doi.org/10.1002/prot.25218.

Pierce, Brian G., Yuichiro Hourai, and Zhiping Weng. 2011. "Accelerating Protein Docking in ZDOCK Using an Advanced 3D Convolution Library." *PloS One* 6 (9): e24657. https://doi.org/10.1371/journal.pone.0024657.

Pierce, Niles A., and Erik Winfree. 2002. "Protein Design Is NP-Hard." *Protein Engineering* 15 (10): 779–82.

Pincus, Matthew R., and Harold A. Scheraga. 1979. "Conformational Energy Calculations of Enzyme-Substrate and Enzyme-Inhibitor Complexes of Lysozyme. 2. Calculation of the Structures of Complexes with a Flexible Enzyme." *Macromolecules* 12 (4): 633–44. https://doi.org/10.1021/ma60070a016.

Pires, Douglas E V, David B Ascher, and Tom L Blundell. 2014. "MCSM: Predicting the Effects of Mutations in Proteins Using Graph-Based Signatures." *Bioinformatics* 30 (3): 335–42. https://doi.org/10.1093/bioinformatics/btt691.

Ponder, Jay W., and Frederic M. Richards. 1987. "An Efficient Newton-like Method for Molecular Mechanics Energy Minimization of Large Molecules." *Journal of Computational Chemistry* 8 (7): 1016–24. https://doi.org/10.1002/jcc.540080710.

Pons, Carles, Marco D'Abramo, Dmitri I. Svergun, Modesto Orozco, Pau Bernadó, and Juan Fernández-Recio. 2010. "Structural Characterization of Protein-Protein Complexes by Integrating Computational Docking with Small-Angle Scattering Data." *Journal of Molecular Biology* 403 (2): 217–30. https://doi.org/10.1016/j.jmb.2010.08.029.

Pons, Carles, Solène Grosdidier, Albert Solernou, Laura Pérez-Cano, and Juan Fernández-Recio. 2010. "Present and Future Challenges and Limitations

in Protein–Protein Docking." *Proteins: Structure, Function, and Bioinformatics* 78 (1): 95–108. https://doi.org/10.1002/prot.22564.

Pons, Carles, Albert Solernou, Laura Perez-Cano, Solène Grosdidier, and Juan Fernandez-Recio. 2010. "Optimization of PyDock for the New CAPRI Challenges: Docking of Homology-Based Models, Domain-Domain Assembly and Protein-RNA Binding." *Proteins* 78 (15): 3182–88. https://doi.org/10.1002/prot.22773.

Pons, Carles, David Talavera, Xavier de la Cruz, Modesto Orozco, and Juan Fernandez-Recio. 2011. "Scoring by Intermolecular Pairwise Propensities of Exposed Residues (SIPPER): A New Efficient Potential for Protein-Protein Docking." *J. Chem. Inf. Model.* 51 (2): 370–77. https://doi.org/10.1021/ci100353e.

Powell, M. J. D. 1964. "An Efficient Method for Finding the Minimum of a Function of Several Variables without Calculating Derivatives." *The Computer Journal* 7 (2): 155–62. https://doi.org/10.1093/comjnl/7.2.155.

Qiu, Di, Peter S. Shenkin, Frank P. Hollinger, and W. Clark Still. 1997. "The GB/SA Continuum Model for Solvation. A Fast Analytical Method for the Calculation of Approximate Born Radii." *The Journal of Physical Chemistry A* 101 (16): 3005–14. https://doi.org/10.1021/jp961992r.

Rao, Gita Subba, and Manoj Kumar. 2008. "Structure-Based Design of a Potent and Selective Small Peptide Inhibitor of Mycobacterium Tuberculosis 6-Hydroxymethyl-7, 8-Dihydropteroate Synthase: A Computer Modelling Approach." *Chemical Biology & Drug Design* 71 (6): 540–45. https://doi.org/10.1111/j.1747-0285.2008.00662.x.

Ravikant, D. V. S., and Ron Elber. 2010. "PIE-Efficient Filters and Coarse Grained Potentials for Unbound Protein-Protein Docking." *Proteins* 78 (2): 400–419. https://doi.org/10.1002/prot.22550.

Ritchie, David W., and Graham J.L. Kemp. 2000. "Protein Docking Using Spherical Polar Fourier Correlations." *Proteins: Structure, Function, and Bioinformatics* 39 (2): 178–94. https://doi.org/10.1002/(SICI)1097-0134(20000501)39:2<178::AID-PROT8>3.0.CO;2-6.

Rodier, Francis, Ranjit Prasad Bahadur, Pinak Chakrabarti, and Joël Janin. 2005. "Hydration of Protein-Protein Interfaces." *Proteins* 60 (1): 36–45. https://doi.org/10.1002/prot.20478.

Romo, Tod D., and Alan Grossfield. 2011. "Validating and Improving Elastic Network Models with Molecular Dynamics Simulations." *Proteins* 79 (1): 23–34. https://doi.org/10.1002/prot.22855.

Rual, Jean-François, Kavitha Venkatesan, Tong Hao, Tomoko Hirozane-Kishikawa, Amélie Dricot, Ning Li, Gabriel F. Berriz, et al. 2005. "Towards a Proteome-Scale Map of the Human Protein–Protein Interaction Network." *Nature* 437 (7062): 1173–78. https://doi.org/10.1038/nature04209.

Rueda, Manuel, Giovanni Bottegoni, and Ruben Abagyan. 2009. "Consistent Improvement of Cross-Docking Results Using Binding Site Ensembles Generated with Elastic Network Normal Modes." *J. Chem. Inf. Model.* 49 (3): 716–25. https://doi.org/10.1021/ci8003732.

Rueda, Manuel, Pablo Chacón, and Modesto Orozco. 2007. "Thorough Validation of Protein Normal Mode Analysis: A Comparative Study with Essential Dynamics." *Structure* 15 (5): 565–75. https://doi.org/10.1016/j.str.2007.03.013.

Sali, A, and T L Blundell. 1993. "Comparative Protein Modelling by Satisfaction of Spatial Restraints." *Journal of Molecular Biology* 234 (3): 779–815. https://doi.org/10.1006/jmbi.1993.1626.

Sanchez, Catherine, Corinne Lachaize, Florence Janody, Bernard Bellon, Laurence Röder, Jérôme Euzenat, François Rechenmann, and Bernard Jacq. 1999. "Grasping at Molecular Interactions and Genetic Networks in Drosophila Melanogaster Using FlyNets, an Internet Database." *Nucleic Acids Research* 27 (1): 89–94. https://doi.org/10.1093/nar/27.1.89.

Schaefer, Michael, and Martin Karplus. 1996. "A Comprehensive Analytical Treatment of Continuum Electrostatics." *The Journal of Physical Chemistry* 100 (5): 1578–99. https://doi.org/10.1021/jp9521621.

Schneidman-Duhovny, Dina, Yuval Inbar, Ruth Nussinov, and Haim J. Wolfson. 2005a. "PatchDock and SymmDock: Servers for Rigid and Symmetric Docking." *Nucleic Acids Research* 33 (Web Server issue): W363-367. https://doi.org/10.1093/nar/gki481.

———. 2005b. "Geometry-Based Flexible and Symmetric Protein Docking." *Proteins* 60 (2): 224–31. https://doi.org/10.1002/prot.20562.

Schreiber, Gideon, and Alan R. Fersht. 1995. "Energetics of Protein-Protein Interactions: Analysis Ofthe Barnase-Barstar Interface by Single Mutations and Double Mutant Cycles." *Journal of Molecular Biology* 248 (2): 478–86. https://doi.org/10.1016/S0022-2836(95)80064-6.

Setny, Piotr, and Martin Zacharias. 2011. "A Coarse-Grained Force Field for Protein-RNA Docking." *Nucleic Acids Res.* 39 (21): 9118–29. https://doi.org/10.1093/nar/gkr636.

Shapovalov, Maxim V., and Roland L. Dunbrack. 2011. "A Smoothed Backbone-Dependent Rotamer Library for Proteins Derived from Adaptive Kernel Density Estimates and Regressions." *Structure* 19 (6): 844–58. https://doi.org/10.1016/j.str.2011.03.019.

Sharabi, Oz, Chen Yanover, Ayelet Dekel, and Julia M. Shifman. 2011. "Optimizing Energy Functions for Protein-Protein Interface Design." *Journal of Computational Chemistry* 32 (1): 23–32. https://doi.org/10.1002/jcc.21594.

Shaw, David E, Ron O Dror, John K Salmon, J. P Grossman, Kenneth M Mackenzie, Joseph A Bank, Cliff Young, et al. 2009. "Millisecond-Scale

Molecular Dynamics Simulations on Anton." In *Proceedings of the Conference on High Performance Computing Networking, Storage and Analysis*, 65:1–65:11. SC '09. New York, NY, USA: ACM. https://doi.org/10.1145/1654059.1654126.

Sheinerman, F. B., R. Norel, and B. Honig. 2000. "Electrostatic Aspects of Protein-Protein Interactions." *Current Opinion in Structural Biology* 10 (2): 153–59.

Shen, Min-yi, and Andrej Sali. 2006. "Statistical Potential for Assessment and Prediction of Protein Structures." *Protein Science* 15 (11): 2507–24. https://doi.org/10.1110/ps.062416606.

Singh, Rohit, Daniel Park, Jinbo Xu, Raghavendra Hosur, and Bonnie Berger. 2010. "Struct2Net: A Web Service to Predict Protein-Protein Interactions Using a Structure-Based Approach." *Nucleic Acids Research* 38 (Web Server issue): W508-515. https://doi.org/10.1093/nar/gkq481.

Sinha, Rohita, Petras J. Kundrotas, and Ilya A. Vakser. 2010. "Docking by Structural Similarity at Protein-Protein Interfaces." *Proteins* 78 (15): 3235–41. https://doi.org/10.1002/prot.22812.

Skrabanek, Lucy, Harpreet K. Saini, Gary D. Bader, and Anton J. Enright. 2008. "Computational Prediction of Protein–Protein Interactions." *Molecular Biotechnology* 38 (1): 1–17. https://doi.org/10.1007/s12033-007-0069-2.

Smyth, M. S., and J. H. Martin. 2000. "X Ray Crystallography." *Molecular Pathology: MP* 53 (1): 8–14.

Söding, Johannes, Andreas Biegert, and Andrei N. Lupas. 2005. "The HHpred Interactive Server for Protein Homology Detection and Structure Prediction." *Nucleic Acids Research* 33 (Web Server issue): W244-248. https://doi.org/10.1093/nar/gki408.

Solernou, Albert, and Juan Fernandez-Recio. 2010. "Protein Docking by Rotation-Based Uniform Sampling (RotBUS) with Fast Computing of Intermolecular Contact Distance and Residue Desolvation." *BMC Bioinformatics* 11 (1): 352. https://doi.org/10.1186/1471-2105-11-352.

Stein, Amelie, Roberto Mosca, and Patrick Aloy. 2011. "Three-Dimensional Modeling of Protein Interactions and Complexes Is Going 'Omics." *Current Opinion in Structural Biology* 21 (2): 200–208. https://doi.org/10.1016/j.sbi.2011.01.005.

Stelzl, Ulrich, Uwe Worm, Maciej Lalowski, Christian Haenig, Felix H. Brembeck, Heike Goehler, Martin Stroedicke, et al. 2005. "A Human Protein-Protein Interaction Network: A Resource for Annotating the Proteome." *Cell* 122 (6): 957–68. https://doi.org/10.1016/j.cell.2005.08.029.

Straub, F. B., and G. Szabolcsi. 1964. “O Dinamicseszkij Aszpektah Sztukturü Fermentov. (On the Dynamic Aspects of Protein Structure).” *Molecular Biology, Problems and Perspectives*.

Stumpf, Michael P. H., Thomas Thorne, Eric de Silva, Ronald Stewart, Hyeong Jun An, Michael Lappe, and Carsten Wiuf. 2008. “Estimating the Size of the Human Interactome.” *Proceedings of the National Academy of Sciences* 105 (19): 6959–64. https://doi.org/10.1073/pnas.0708078105.

Suhre, K., and Y.-H. Sanejouand. 2004. “ElNemo: A Normal Mode Web Server for Protein Movement Analysis and the Generation of Templates for Molecular Replacement.” *Nucleic Acids Research* 32 (Web Server): W610–14. https://doi.org/10.1093/nar/gkh368.

Sun, Jingchun, and Zhongming Zhao. 2010. “A Comparative Study of Cancer Proteins in the Human Protein-Protein Interaction Network.” *BMC Genomics* 11 Suppl 3 (December): S5. https://doi.org/10.1186/1471-2164-11-S3-S5.

Szilagyi, Andras, and Yang Zhang. 2014. “Template-Based Structure Modeling of Protein-Protein Interactions.” *Current Opinion in Structural Biology* 0 (February): 10–23. https://doi.org/10.1016/j.sbi.2013.11.005.

Tama, F., and Y. H. Sanejouand. 2001. “Conformational Change of Proteins Arising from Normal Mode Calculations.” *Protein Engineering* 14 (1): 1–6.

Theodoridis, Sergios, and Konstantinos Koutroumbas. 1999. *Pattern Recognition*. London, England: Academic Press.

Thorn, K S, and A A Bogan. 2001. “ASEdb: A Database of Alanine Mutations and Their Effects on the Free Energy of Binding in Protein Interactions.” *Bioinformatics (Oxford, England)* 17 (3): 284–85.

Tidow, Henning, and Poul Nissen. 2013. “Structural Diversity of Calmodulin Binding to Its Target Sites.” *The FEBS Journal* 280 (21): 5551–65. https://doi.org/10.1111/febs.12296.

Ting, Daniel, Guoli Wang, Maxim Shapovalov, Rajib Mitra, Michael I. Jordan, and Roland L. Dunbrack Jr. 2010. “Neighbor-Dependent Ramachandran Probability Distributions of Amino Acids Developed from a Hierarchical Dirichlet Process Model.” *PLOS Computational Biology* 6 (4): e1000763. https://doi.org/10.1371/journal.pcbi.1000763.

Tirion, Monique M. 1996. “Large Amplitude Elastic Motions in Proteins from a Single-Parameter, Atomic Analysis.” *Physical Review Letters* 77 (9): 1905.

Tobi, Dror, and Ivet Bahar. 2005. “Structural Changes Involved in Protein Binding Correlate with Intrinsic Motions of Proteins in the Unbound State.” *Proceedings of the National Academy of Sciences of the United States of America* 102 (52): 18908–13. https://doi.org/10.1073/pnas.0507603102.

Torchala, Mieczyslaw, Iain H. Moal, Raphael A. G. Chaleil, Juan Fernandez-Recio, and Paul A. Bates. 2013. "SwarmDock: A Server for Flexible Protein–Protein Docking." *Bioinformatics* 29 (6): 807–9. https://doi.org/10.1093/bioinformatics/btt038.

Tsai, C. J., S. Kumar, B. Ma, and R. Nussinov. 1999. "Folding Funnels, Binding Funnels, and Protein Function." *Protein Science : A Publication of the Protein Society* 8 (6): 1181–90.

Tsai, C. J., S. L. Lin, H. J. Wolfson, and R. Nussinov. 1996. "Protein-Protein Interfaces: Architectures and Interactions in Protein-Protein Interfaces and in Protein Cores. Their Similarities and Differences." *Critical Reviews in Biochemistry and Molecular Biology* 31 (2): 127–52. https://doi.org/10.3109/10409239609106582.

———. 1997. "Studies of Protein-Protein Interfaces: A Statistical Analysis of the Hydrophobic Effect." *Protein Science: A Publication of the Protein Society* 6 (1): 53–64. https://doi.org/10.1002/pro.5560060106.

Tsai, Chung-Jung, Antonio del Sol, and Ruth Nussinov. 2008. "Allostery: Absence of a Change in Shape Does Not Imply That Allostery Is Not at Play." *Journal of Molecular Biology* 378 (1): 1–11. https://doi.org/10.1016/j.jmb.2008.02.034.

Tuncbag, N., O. Keskin, and A. Gursoy. 2010. "HotPoint: Hot Spot Prediction Server for Protein Interfaces." *Nucleic Acids Research* 38 (Web Server): W402–6. https://doi.org/10.1093/nar/gkq323.

Tuncbag, Nurcan, Attila Gursoy, and Ozlem Keskin. 2009. "Identification of Computational Hot Spots in Protein Interfaces: Combining Solvent Accessibility and Inter-Residue Potentials Improves the Accuracy." *Bioinformatics* 25 (12): 1513–20. https://doi.org/10.1093/bioinformatics/btp240.

Tuncbag, Nurcan, Attila Gursoy, Ruth Nussinov, and Ozlem Keskin. 2011. "Predicting Protein-Protein Interactions on a Proteome Scale by Matching Evolutionary and Structural Similarities at Interfaces Using PRISM." *Nature Protocols* 6 (9): 1341–54. https://doi.org/10.1038/nprot.2011.367.

Tuszynska, Irina, and Janusz M. Bujnicki. 2011. "DARS-RNP and QUASI-RNP: New Statistical Potentials for Protein-RNA Docking." *BMC Bioinformatics* 12: 348. https://doi.org/10.1186/1471-2105-12-348.

Tuszynska, Irina, Marcin Magnus, Katarzyna Jonak, Wayne Dawson, and Janusz M Bujnicki. 2015. "NPDock: A Web Server for Protein-Nucleic Acid Docking." *Nucleic Acids Res.* 43 (W1): W425-30. https://doi.org/10.1093/nar/gkv493.

Uetz, P., L. Giot, G. Cagney, T. A. Mansfield, R. S. Judson, J. R. Knight, D. Lockshon, et al. 2000. "A Comprehensive Analysis of Protein-Protein

Interactions in Saccharomyces Cerevisiae." *Nature* 403 (6770): 623–27. https://doi.org/10.1038/35001009.

Urvoas, Agathe, Asma Guellouz, Marie Valerio-Lepiniec, Marc Graille, Dominique Durand, Danielle C Desravines, Herman van Tilbeurgh, Michel Desmadril, and Philippe Minard. 2010. "Design, Production and Molecular Structure of a New Family of Artificial Alpha-Helicoidal Repeat Proteins (ARep) Based on Thermostable HEAT-like Repeats." *Journal of Molecular Biology* 404 (2): 307–27. https://doi.org/10.1016/j.jmb.2010.09.048.

Velankar, S., Y. Alhroub, C. Best, S. Caboche, M. J. Conroy, J. M. Dana, M. A. Fernandez Montecelo, et al. 2012. "PDBe: Protein Data Bank in Europe." *Nucleic Acids Research* 40 (Database issue): D445-452. https://doi.org/10.1093/nar/gkr998.

Velankar, S., C. Best, B. Beuth, C. H. Boutselakis, N. Cobley, A. W. Sousa Da Silva, D. Dimitropoulos, et al. 2010. "PDBe: Protein Data Bank in Europe." *Nucleic Acids Research* 38 (Database issue): D308-317. https://doi.org/10.1093/nar/gkp916.

Velankar, Sameer, Younes Alhroub, Anaëlle Alili, Christoph Best, Harry C. Boutselakis, Ségolène Caboche, Matthew J. Conroy, et al. 2011. "PDBe: Protein Data Bank in Europe." *Nucleic Acids Research* 39 (Database issue): D402-410. https://doi.org/10.1093/nar/gkq985.

Venkatesan, Kavitha, Jean-François Rual, Alexei Vazquez, Ulrich Stelzl, Irma Lemmens, Tomoko Hirozane-Kishikawa, Tong Hao, et al. 2009. "An Empirical Framework for Binary Interactome Mapping." *Nature Methods* 6 (1): 83–90. https://doi.org/10.1038/nmeth.1280.

Vreven, Thom, Howook Hwang, Brian G. Pierce, and Zhiping Weng. 2014. "Evaluating Template-Based and Template-Free Protein–Protein Complex Structure Prediction." *Briefings in Bioinformatics* 15 (2): 169–76. https://doi.org/10.1093/bib/bbt047.

Vreven, Thom, Iain H. Moal, Anna Vangone, Brian G. Pierce, Panagiotis L. Kastritis, Mieczyslaw Torchala, Raphael Chaleil, et al. 2015. "Updates to the Integrated Protein-Protein Interaction Benchmarks: Docking Benchmark Version 5 and Affinity Benchmark Version 2." *Journal of Molecular Biology* 427 (19): 3031–41. https://doi.org/10.1016/j.jmb.2015.07.016.

Vuignier, Karine, Julie Schappler, Jean-Luc Veuthey, Pierre-Alain Carrupt, and Sophie Martel. 2010. "Drug–Protein Binding: A Critical Review of Analytical Tools." *Analytical and Bioanalytical Chemistry* 398 (1): 53–66. https://doi.org/10.1007/s00216-010-3737-1.

Wang, Junmei, Piotr Cieplak, and Peter A. Kollman. 2000. "How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules?" *Journal of*

Computational Chemistry 21 (12): 1049–74. https://doi.org/10.1002/1096-987X(200009)21:12<1049::AID-JCC3>3.0.CO;2-F.

Wells, J A. 1991. "Systematic Mutational Analyses of Protein-Protein Interfaces." *Methods Enzymol.* 202: 390–411.

Wells, J. A. 1996. "Binding in the Growth Hormone Receptor Complex." *Proceedings of the National Academy of Sciences of the United States of America* 93 (1): 1–6.

Wells, J. A., and A. M. de Vos. 1993. "Structure and Function of Human Growth Hormone: Implications for the Hematopoietins." *Annual Review of Biophysics and Biomolecular Structure* 22: 329–51. https://doi.org/10.1146/annurev.bb.22.060193.001553.

Wells, Stephen, Scott Menor, Brandon Hespenheide, and M. F. Thorpe. 2005. "Constrained Geometric Simulation of Diffusive Motion in Proteins." *Physical Biology* 2 (4): S127-136. https://doi.org/10.1088/1478-3975/2/4/S07.

Whitehead, Timothy A, Aaron Chevalier, Yifan Song, Cyrille Dreyfus, Sarel J Fleishman, Cecilia De Mattos, Chris A Myers, et al. 2012. "Optimization of Affinity, Specificity and Function of Designed Influenza Inhibitors Using Deep Sequencing." *Nature Biotechnology* 30 (6): 543–48. https://doi.org/10.1038/nbt.2214.

Willander, M., and Safaa Al-Hilli. 2009. "Analysis of Biomolecules Using Surface Plasmons." *Methods in Molecular Biology (Clifton, N.J.)* 544: 201–29. https://doi.org/10.1007/978-1-59745-483-4_14.

Wlodarski, Tomasz, and Bojan Zagrovic. 2009. "Conformational Selection and Induced Fit Mechanism Underlie Specificity in Noncovalent Interactions with Ubiquitin." *Proceedings of the National Academy of Sciences of the United States of America* 106 (46): 19346–51. https://doi.org/10.1073/pnas.0906966106.

Wodak, S. Y. 1977. "The Structure of Cytidilyl(2',5')Adenosine When Bound to Pancreatic Ribonuclease S." *Journal of Molecular Biology* 116 (4): 855–75.

Wu, Sitao, and Yang Zhang. 2008. "MUSTER: Improving Protein Sequence Profile-Profile Alignments by Using Multiple Sources of Structure Information." *Proteins* 72 (2): 547–56. https://doi.org/10.1002/prot.21945.

Wüthrich, Kurt. 1990. "Protein Structure Determination in Solution by NMR Spectroscopy." *Journal of Biological Chemistry* 265 (36): 22059–62.

Xie, Wei, and Nikolaos V. Sahinidis. 2006. "Residue-Rotamer-Reduction Algorithm for the Protein Side-Chain Conformation Problem." *Bioinformatics* 22 (2): 188–94. https://doi.org/10.1093/bioinformatics/bti763.

Xu, D., S. L. Lin, and R. Nussinov. 1997. "Protein Binding versus Protein Folding: The Role of Hydrophilic Bridges in Protein Associations." *Journal of Molecular Biology* 265 (1): 68–84. https://doi.org/10.1006/jmbi.1996.0712.

Yang, Lee-Wei, Eran Eyal, Ivet Bahar, and Akio Kitao. 2009. "Principal Component Analysis of Native Ensembles of Biomolecular Structures (PCA_NEST): Insights into Functional Dynamics." *Bioinformatics (Oxford, England)* 25 (5): 606–14. https://doi.org/10.1093/bioinformatics/btp023.

Yang, Lee-Wei, Eran Eyal, Chakra Chennubhotla, JunGoo Jee, Angela M. Gronenborn, and Ivet Bahar. 2007. "Insights into Equilibrium Dynamics of Proteins from Comparison of NMR and X-Ray Data with Computational Predictions." *Structure (London, England: 1993)* 15 (6): 741–49. https://doi.org/10.1016/j.str.2007.04.014.

Yang, Lee-Wei, Xiong Liu, Christopher J. Jursa, Mark Holliman, A. J. Rader, Hassan A. Karimi, and Ivet Bahar. 2005. "IGNM: A Database of Protein Functional Motions Based on Gaussian Network Model." *Bioinformatics (Oxford, England)* 21 (13): 2978–87. https://doi.org/10.1093/bioinformatics/bti469.

Závodszky, P., L. V. Abaturov, and Ya M. Varshavsky. 1966. "Structure of Glyceraldehyde-3-Phosphate Dehydrogenase and Its Alteration by Coenzyme Binding." *Acta Biochim Biophys Acad Sci Hung* 1: 389–403.

Zhang, Qiangfeng Cliff, Donald Petrey, Lei Deng, Li Qiang, Yu Shi, Chan Aye Thu, Brygida Bisikirska, et al. 2012. "Structure-Based Prediction of Protein-Protein Interactions on a Genome-Wide Scale." *Nature* 490 (7421): 556–60. https://doi.org/10.1038/nature11503.

Zhang, Qiangfeng Cliff, Donald Petrey, Raquel Norel, and Barry H. Honig. 2010. "Protein Interface Conservation across Structure Space." *Proceedings of the National Academy of Sciences of the United States of America* 107 (24): 10896–901. https://doi.org/10.1073/pnas.1005894107.

Zheng, Wenjun, and Sebastian Doniach. 2003. "A Comparative Study of Motor-Protein Motions by Using a Simple Elastic-Network Model." *Proceedings of the National Academy of Sciences of the United States of America* 100 (23): 13253–58. https://doi.org/10.1073/pnas.2235686100.

Zhou, Hongyi, and Yaoqi Zhou. 2002. "Distance-Scaled, Finite Ideal-Gas Reference State Improves Structure-Derived Potentials of Mean Force for Structure Selection and Stability Prediction." *Protein Science: A Publication of the Protein Society* 11 (11): 2714–26. https://doi.org/10.1110/ps.0217002.

Zundert, G. C. P. van, J. P. G. L. M. Rodrigues, M. Trellet, C. Schmitz, P. L. Kastritis, E. Karaca, A. S. J. Melquiond, M. van Dijk, S. J. de Vries, and A. M. J. J. Bonvin. 2016. "The HADDOCK2.2 Web Server: User-

Friendly Integrative Modeling of Biomolecular Complexes." *Journal of Molecular Biology* 428 (4): 720–25. https://doi.org/10.1016/j.jmb.2015.09.014.