Structure
Article

A Multidomain Flexible Docking Approach to Deal with Large Conformational Changes in the Modeling of Biomolecular Complexes

Ezgi Karaca[1] and Alexandre M.J.J. Bonvin[1,*]
[1]Bijvoet Center for Biomolecular Research, Faculty of Science, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands
*Correspondence: a.m.j.j.bonvin@uu.nl
DOI 10.1016/j.str.2011.01.014

SUMMARY

Binding-induced backbone and large-scale conformational changes represent one of the major challenges in the modeling of biomolecular complexes by docking. To address this challenge, we have developed a flexible multidomain docking protocol that follows a "divide-and-conquer" approach to model both large-scale domain motions and small- to medium-scale interfacial rearrangements: the flexible binding partner is treated as an assembly of subparts/domains that are docked simultaneously making use of HADDOCK's multidomain docking ability. For this, the flexible molecules are cut at hinge regions predicted using an elastic network model. The performance of this approach is demonstrated on a benchmark covering an unprecedented range of conformational changes of 1.5 to 19.5 Å. We show from a statistical survey of known complexes that the cumulative sum of eigenvalues obtained from the elastic network has some predictive power to indicate the extent of the conformational change to be expected.

INTRODUCTION

Proteins play a vital role in all kinds of biological processes through their complex interactions with other biomolecules and small ligands. Revealing all the functional steps in the lifetime of a protein requires 3D atomic-level information about the complexes it forms. This information can be obtained from classical experimental techniques such as X-ray crystallography and NMR, although, considering the enormous number of complexes, the need for accurate and complementary computational methods like docking is evident (Bonvin, 2006; Lensink and Mendez, 2008; Moreira et al., 2010; Pons et al., 2010; Ritchie, 2008; Zacharias, 2010). Docking is defined as the modeling of the structure of a complex (the bound conformation) starting from the free forms (unbound conformations) of the interaction partners. This becomes an extremely challenging task when the conformation of a protein changes significantly upon complex formation (Andrusier et al., 2008; Bonvin, 2006; Lensink and Mendez, 2008; Zacharias, 2010).

Various mechanisms have been proposed over the years to describe the binding process and its associated conformational changes. Fischer suggested that the active sites of the interacting molecules are complementary to each other, obeying a lock-and-key mechanism (Fischer, 1894). Koshland extended this model by proposing that the binding follows an induced fit rule, where the molecules adopt different conformations in order to fit into the active site of their interaction partner (Koshland, 1958). Later on, Kumar et al. reevaluated this issue from a statistical mechanics point of view and proposed the conformational selection model (Kumar et al., 2000): the native state of a protein exists in an ensemble of conformations sampling various states and environmental effects shift the equilibrium toward the bound conformation. Subsequent to this hypothesis, Grunberg et al. (2006) provided a consensus model, treating the concept of recognition and binding as two different processes with conformational selection at the basis of the recognition process, starting with the formation of an encounter complex.

Conformational changes occurring upon binding can be classified into four categories: local rearrangements, collective global motions, a mixture of the two, and binding-induced folding events. Local changes cover loop rearrangements (Andrusier et al., 2008), secondary structure element alterations and motions, etc. Global changes refer to large-scale domain motions like hinge and shear (Andrusier et al., 2008; Gerstein and Krebs, 1998; Gerstein and Echols, 2004). Binding-induced folding events are observed in specific recognition when the folding of one of (partially) unfolded partner occurs only upon binding (Bonvin, 2006; Mittag et al., 2010; Zacharias, 2010).

Various strategies have been followed to model these different types of conformational changes (Andrusier et al., 2008; Bonvin, 2006; Lensink and Mendez, 2008; Moreira et al., 2010; Zacharias, 2010). For the lock-and-key cases, namely, when shape complementarily drives the interaction, rigid body algorithms (e.g., geometric hashing [Bachar et al., 1993] and Fast Fourier Transform [Katchalski-Katzir et al., 1992]) provide fast and accurate solutions. Small- to medium-scale backbone conformational changes (in the range of 1–2 Å) can be modeled through algorithms that address the induced fit mechanism (mostly based on Molecular Dynamics and Monte Carlo simulations (Dominguez et al., 2003; Gray et al., 2003)). For some types of conformational changes, e.g., loop rearrangements or allosteric effects, this may not provide satisfactory results. Under such conditions, docking can be started from ensembles of structures (coming from NMR, Molecular Dynamics or Monte Carlo Simulations, Normal Mode Analysis, graph theoretic approaches …),

following thus the conformational selection model (Chaudhury and Gray, 2008; Dominguez et al., 2003; Krol et al., 2007). As an alternative method, a multicopy mean-field approach (Bastard et al., 2003; May and Zacharias, 2007) can also be used. The final class of algorithms is based on the consensus binding model introduced by Grunberg et al. (2006): structures are docked starting from an ensemble of conformations and then refined allowing for small conformational changes to take place (Chaudhury and Gray, 2008; Dominguez et al., 2003). This class of algorithms can be useful if different types of conformational changes are observed at the same time. Modeling of the last class of conformational changes, namely, folding upon binding should be considered separately from regular docking as it requires incorporation of protein structure prediction within the context of docking. As a first step to deal with this problem, Baker's group recently developed a *fold-and-dock* protocol within Rosetta for modeling symmetric homo-oligomers (Das et al., 2009).

All of the above-mentioned algorithms work reasonably well for small- to medium-scale local conformational changes, but they usually fail to model large loop rearrangements, secondary structure changes, and collective domain motions where the backbone rmsd of a protein changes by more than 2–3 Å (Bonvin, 2006; Dobbins et al., 2008; Lensink and Mendez, 2008). This is mainly due to the complexity of the conformational space to be sampled and the difficulty of predicting a priori such type of conformational changes (Bonvin, 2006; Dobbins et al., 2008; Lensink and Mendez, 2008). Hybrid methodologies that combine different approaches are thus needed in order to address this problem. Until now five hybrid methods have been proposed to model large-scale conformational changes: ATTRACT (May and Zacharias, 2007, 2008), SwarmDock (Li et al., 2010; Moal and Bates, 2010), multistage docking procedure of MolFit (Ben-Zeev et al., 2005), FlexDock (Schneidman-Duhovny et al., 2007), and fold-tree representation of Rosetta (Wang et al., 2007). ATTRACT can incorporate soft harmonic low frequency modes into the docking procedure and provide fast relaxation/adaptation of the structure on a global scale (May and Zacharias, 2008). SwarmDock also uses normal modes in the docking procedure but, in contrast to ATTRACT, it takes a linear combination of both low and high frequency normal modes into account for addressing high-frequency thermal motions occurring at the interface (Moal and Bates, 2010). MolFit treats the individual domains of the molecules as soft rigid objects and then docks them in a sequential multistage two-body docking protocol (Ben-Zeev et al., 2005). The method of FlexDock exploits a similar methodology; it dissects the flexible protein into rigid domains and performs a pairwise docking of the separate domains using PatchDock (Schneidman-Duhovny et al., 2005) followed by assembly of the resulting models. In Rosetta, the protein is represented as a fold-tree, which provides a mean for defining flexible regions between centers of the rigid molecules and thus allowing domains to move with respect to each other during rigid-body optimization (Wang et al., 2007). Next to these docking approaches, there is a recently published flexible refinement server, FiberDock constructed to deal with large conformational changes. FiberDock deforms the backbone in the direction of the selected normal modes (Mashiach et al., 2010). The normal mode selection is based on the correlation of that particular mode with the van der Waals forces. All these approaches are quite promising, but they also have some limitations, especially in treating interfacial induced backbone and side-chain conformational changes at the same time.

Our in-house docking software HADDOCK has already proved its ability to deal with small- to medium-scale induced conformational changes by combining docking from ensembles of starting structures with flexible refinement of both side chains and backbone (de Vries et al., 2007, 2010; Dominguez et al., 2003). Moreover, we have recently demonstrated HADDOCK's ability to deal with multiple molecules (>2) to build macromolecular assemblies (Karaca et al., 2010). Here, we combine all these aspects and propose a straightforward and easy-to-apply docking protocol that can deal with large conformational changes while accounting for local changes at the same time. Following a "divide-and-conquer" approach, our flexible multidomain docking (FMD) protocol within HADDOCK partitions the flexible molecule into rigid bodies with connectivity restraints between them and performs a simultaneous multibody docking of all components (Karaca et al., 2010). The proteins are dissected into domains by cutting them at hinge regions predicted from an Elastic Network Model (Emekli et al., 2008). This allows modeling of global scale changes at the rigid body docking stage. The resulting models are subsequently subjected to a flexible refinement involving both side chains and backbone motions to deal with small- to medium-scale induced conformational changes. Our FMD protocol was tested against a benchmark of eleven protein-protein complexes that are experiencing domain motions, spanning a range of conformational changes from 1.5 Å to as much as 19.5 Å. This new FMD protocol is shown to be an excellent approach to model conformational changes as large as 19.5 Å.

RESULTS

Benchmark Compilation

The FMD benchmark was compiled according to three major criteria: (1) One of the partners should only undergo a small conformational change, of less than 2.0 Å, in order to decrease the level of complexity, (2) the other partner should experience conformational change emanating from hinge motions (ideally from one hinge position) and the hinge motion should be functionally involved in the binding process, (3) both the bound and unbound conformations of the partners should be available. Eleven protein-protein complexes were found that fit all the above-mentioned requirements, except 1NPE, a former Critical Assessment of Prediction of Interactions (CAPRI, Mendez et al., 2003) target, with only the bound conformation of the ligand available (Table 1). The other complexes were taken from the Protein-Protein Docking Benchmark 4.0 (Hwang et al., 2010), eight of which belonged to the difficult category. A vast range of conformational changes is covered, from 1.5 to 19.5 Å, with 1IRA, 1H1V, 1Y64 1F6M, and 1FAK being particularly challenging (Table 1 and Figure 1).

The Workflow of Flexible Multidomain Docking

Dissecting Proteins into Domains

Hinges were predicted using the HingeProt server (Emekli et al., 2008). They were then filtered to guarantee the structural

Table 1. Selected Complexes for the Flexible Multidomain Docking Benchmark

Complex ID	Receptor ID[a]	Ligand ID[a]	Backbone Conformational Change Range (Å)	
			Receptor	Ligand
1IRA[b] (Schreuder et al., 1997)	1G0Y_R (Vigers et al., 2000)	1ILR_1 (Schreuder et al., 1995)	19.5	0.7
1H1V[b] (Choe et al., 2002)	1D0N_B (Burtnick et al., 1997)	1IJJ_B (Bubb et al., 2002)	13.9	1.6
1Y64[b] (Otomo et al., 2005)	1UX5_A (Xu et al., 2004)	2FXU_A (Rizvi et al., 2006)	10.3	1.1
1F6M[b] (Lennon et al., 2000)	1CL0_A (Lennon et al., 1999)	2TIR_A (Nikkola et al., 1993)	7.3	0.9
1FAK[b] (Zhang et al., 1999)	1QFK_HL (Pike et al., 1999)	1TFH_B (Huang et al., 1998)	6.0	1.0
1ZLI[b] (Arolas et al., 2005)	2JTO_A (Pantoja-Uceda et al., 2008)	1KWM_A (Barbosa Pereira et al., 2002)	3.8	0.6
1E4K[b] (Sondermann et al., 2000)	2DTQ_AB (Matsumiya et al., 2007)	1FNL_A (Zhang et al., 2000)	2.9	1.7
1IBR[b] (Vetter et al., 1999)	1F59_A (Bayliss et al., 2000)	1QG4_A (Kent et al., 1999)	2.9	1.1
1KKL[b] (Fieulaine et al., 2002)	1JB1_AB (Fieulaine et al., 2001)	2HPR (Liao and Herzberg, 1994)	2.6	0.5
1NPE[c] (Takagi et al., 2003)	1KLO_A (Stetefeld et al., 1996)	1NPE_A (Takagi et al., 2003)	1.8	—
1DFJ[b] (Kobe and Deisenhofer, 1995)	2BNH_A (Kobe and Deisenhofer, 1996)	9RSA_B (Nachman et al., 1990)	1.5	0.7

[a] The PDB and chain ID's are indicated as PDB ID_chain ID.
[b] Docking was performed starting from the unbound conformations of both receptor and ligand.
[c] For this particular protein, as the unbound conformation of the ligand was not available, the docking was performed from the unbound conformation of the receptor and the bound conformation of the ligand.

integrity and dissect the flexible molecule in as few components as possible in order to maintain the compactness of the expected solution. This filtering is based on the domain motions study of Hayward in which hinges were found to be more likely at the termini of α helices, β sheets, and in loops that connect domains, less likely in the center of interdomain α helices, and not likely within β sheets (Hayward, 1999). In addition, we also discarded hinge predictions close to N- and C-terminal residues, as it is highly improbable that such residues can have a hinge-like behavior. Subsequently, the monomer was cut at the first peptide bond following the predicted hinge residue. Short explanations for the basis of the hinge selection for each protein can be found in Table 2.

Defining Ambiguous Interaction and Connectivity Restraints

The Ambiguous Interaction Restraints (AIRs) were generated via a new HADDOCK server interface especially built to allow fine-tuning of restraints in the case of multidomain docking (http://haddock.chem.uu.nl/services/GenTBL/) (de Vries et al., 2010): (1) Restraints for the interface of the rigid molecule were kept ambiguous for the distinct interfaces of the separated flexible molecule, (2) No AIRs were defined between the separated domains of the flexible molecule (see Figure 2). The AIRs were defined based on interface residues identified from the crystal structure of the complexes. By choosing an ideal definition of the interface (but not of the contacts made) we focus on the problem of dealing with conformational changes. This represents thus a best-case scenario. To maintain the connectivity between the separated chains, two connectivity restraint files consisting of an unambiguous distance restraint between the C and N termini of the separated domains were prepared. The first file was used during the rigid body energy minimization (*it0*) and semiflexible refinement in torsion angle space (*it1*), to enforce a chain separation of no more than 10 Å (upper distance restraint). The second one was imposed during the final refinement in explicit solvent (*water*) in order to restore the connectivity (to the real peptide distance, 1.3 Å). As the N and C termini of the cut domains are created artificially, they were kept uncharged. Three residues on both sides of the separated domains were defined as fully flexible to allow for more flexibility of the linker region. All other HADDOCK parameters were left to their default values. The summary of the workflow is illustrated in Figure 3.

Performance of Flexible Multidomain Docking

Both standard semiflexible two-body docking and the new FMD protocol were applied to the benchmark for comparison. Two-body docking could generate good solutions (two stars) for 1DFJ and 1NPE and acceptable ones (one star) for 1E4K and 1KKL. The best solutions were ranked at the top for the latter three (Table 3). For the other cases, two-body docking failed completely in generating any acceptable solutions. When the new FMD protocol was applied, acceptable or better solutions were obtained for each benchmark case (Table 3). The improvement is impressive particularly for the seven cases (1IRA, 1H1V, 1Y64, 1F6M, 1FAK, 1ZLI, 1IBR) for which two-body docking was failing: the best l-rmsd values of the top four challenging cases decreased from 38.6 to 7.5 Å for 1IRA, from 31.1 to 9.6 Å for 1H1V, from 3.2 to 9.5 Å for 1Y64 and from 39.4 to 8.7 Å for 1F6M (note also the very high fraction of native contacts, Table 3). Besides providing acceptable-to-good solutions for all of the cases, FMD-HADDOCK could rank them at the top using the standard scoring scheme in HADDOCK consisting of a weighted sum of van der Waals, electrostatic energies, and an empirical desolvation term (Fernandez-Recio et al., 2004) ($HADDOCK_{score} = 1.0\ E_{vdw} + 0.2\ E_{elec} + 1.0\ E_{desol}$). Furthermore, the quality of the models improved for 1E4K, 1KKL, 1NPE, and 1DFJ, for which two-body docking already provided reasonable solutions. 1E4K's and 1KKL's best models, ranked among the top five, are now a two star (good) solution and 1NPE's top-ranking structure is almost a three star (high accuracy)

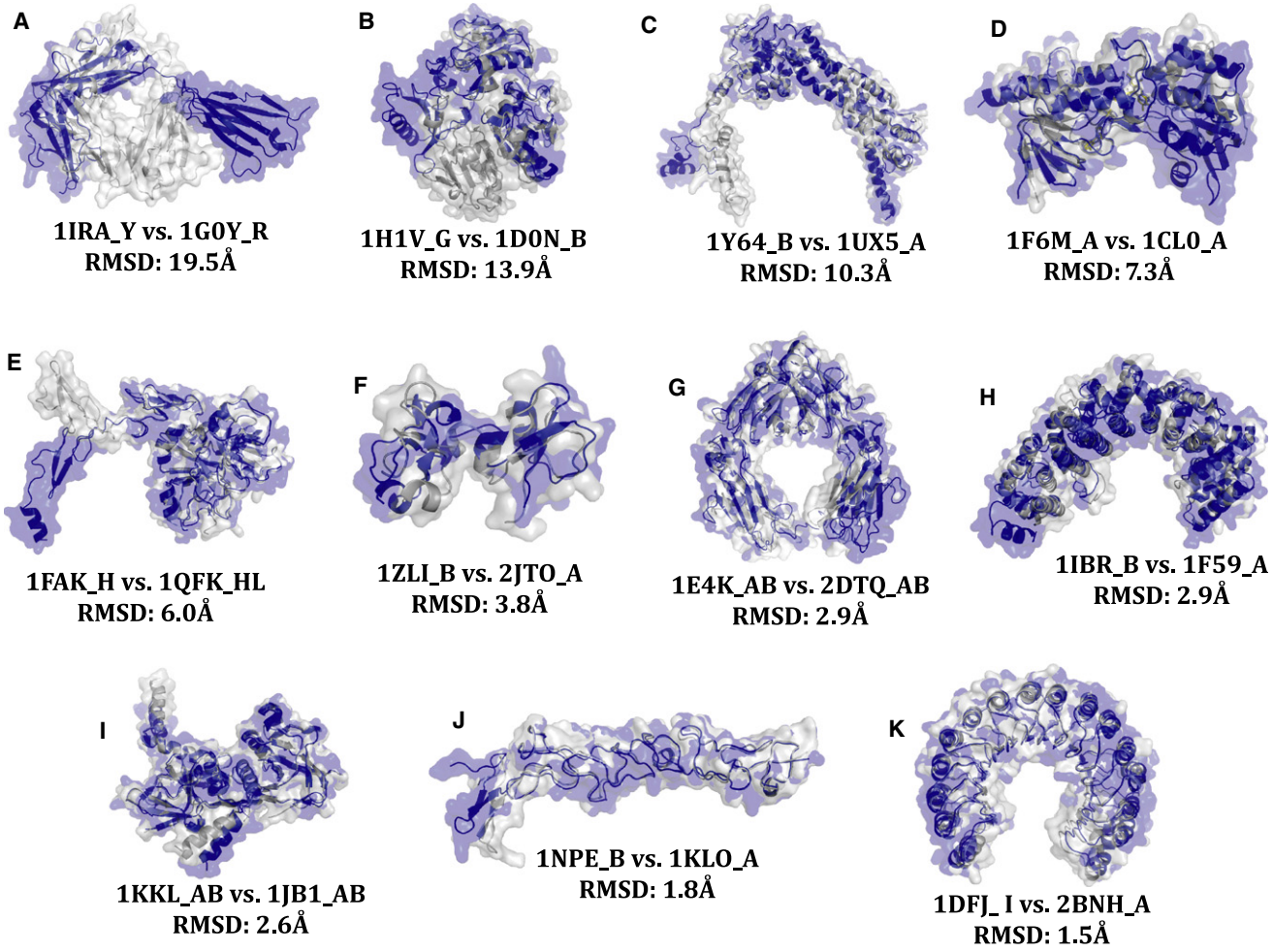

Figure 1. Comparison of the Unbound and Bound Superimposed Conformations of the Receptors in our Benchmark
(A) 1IRA, (B) 1H1V, (C) 1Y64, (D) 1F6M, (E) 1FAK, (F) 1ZLI, (G) 1IBR, (H) 1E4K, (I) 1KKL, (J) 1NPE, (K) 1DFJ. For each case, the positional backbone rmsd between the two forms is indicated. The figures were generated with Pymol (DeLano, 2002).

prediction (with i-rmsd = 1.1 Å and Fnat = 0.95). For 1DFJ the quality of the best model did not change much, but its rank improved significantly from 116 to 5 (see Table 3). The best predictions superimposed onto the reference crystal structures are shown in Figure 4.

Within the standard HADDOCK protocol the final ranking of solutions is based on cluster averages rather than individual ranks. Each of the top-ranking clusters contained at least one acceptable prediction, with an average Fnat of 0.44 for the worst and of 0.82 for the best case (Table 4). This stresses the excellent performance of our scoring scheme. We also analyzed the number of acceptable or better structures generated during each stage of the docking (Table 5). One can observe a clear decrease in the number of acceptable solutions as a function of the docking difficulty (related to the extent of the conformational changes). For the challenging cases, scoring becomes even more critical since the fraction of good solutions within the pool of sampled conformations decreases, as can be seen in the case of 1Y64: For this complex, only one acceptable solution is generated after refinement; despite this, HADDOCK ranked it at the top (Table 3). For the other cases, the FMD protocol provided a pool of near native predictions after each refinement step (*it1, water*) (Table 5).

DISCUSSION

Docking has become a popular approach to model biomolecular interactions. Current approaches are able to deal with small side-chain and backbone alterations while dealing with large conformational changes is still one of the major bottlenecks in the field. Considering the various types of conformational changes and the focus of existing docking approaches on mainly dealing with rather local small changes, it is evident that new developments are required to tackle modeling of large conformational changes occurring upon binding.

Here, we have presented a straightforward and easy-to-apply flexible multidomain docking protocol that can deal with large collective backbone conformational changes by treating the flexible partner as a collection of subdomains with connectivity restraints between them. Our results have revealed that FMD

Table 2. List of Selected Hinges and the Basis of the Selection

Complex ID	Receptor ID	Hinge Predictions[a] (selected hinge in boldface)	Basis of the Selection
1IRA_Y	1G0Y_R	13,98,**203**,307	13 and 307 are at the N and C termini, respectively. 98 and 203 are in a linker, but 203 is within a 10 residue longer linker than 98.
1H1V_G	1D0N_B	423,446,500,534,**631**	Only 534 and 631 are in a linker, where 631 is more flexible compared with 534 according to experimentally determined β-factors.
1Y64_B	1UX5_A	1427,**1403**,1544,1647,1699	1403 is in a linker while the others are in α helices.
1F6M_A	1CL0_A	11,79,110,**115**,148,**245**,283,303	This protein is cut at two places. 115 and 245 are selected as probable hinges, as they lie on a linker that connects two large domains.
1FAK_H	1QFK_L	83,**89**,130	130 is a C-terminal residue. Both 83 and 89 are in the same linker; 89 was selected arbitrarily.
1ZLI_B	2JTO_A	**37**	The only hinge prediction is selected.
1E4K_AB	2DTQ_AB	338,444(in chain A) **340**(in chain B)	The receptor of this complex is a symmetrical homodimer (composed of chain A and B). Except 444, which is at the C terminus, both predictions point out the same regions in different chains, thus just one monomer is cut at 340.
1IBR_B	1F59_A	116,**237**,322	All predictions are in the middle of α helices. Therefore 237 is selected, as it corresponds to the center of mass of the receptor.
1KKL	1JB1_ABC	156,201,218,225,262,264,281, 285,286(2 times),**291**,307	The hinge prediction is run on chains A, AB, and ABC. The receptor is a symmetric trimer, thus all the predictions are considered. The most frequently predicted hinge regions are around 289-293 linker, in which 291 is located.
1NPE_B	1KLO_A	**43**,89,122	In the literature it was already stated that this protein is composed of three modules, containing 4 disulphide bridges. Residue 122 forms a disulphide bridge. Residue 43 separates the first module, while residue 89 is in the middle of the second module (Stetefeld et al., 1996).
1DFJ_I	2BNH_A	**117**,229,326	Residue 326 is in the middle of an α helix, residue 229 is completely in the center of a barrel formed by α helices and β sheets and residue 117 is in a linker.

[a] The hinge predictions were obtained with the HingeProt server (Emekli et al., 2008).

outperforms standard two-body docking especially in the cases with conformational change range ≥3 Å (1IRA, 1H1V, 1Y64, 1F6M, 1FAK, 1ZLI), something that has never been demonstrated before. Even in the case of smaller conformational changes (<3 Å) the FMD protocol improves in general both the quality and ranking of the solutions upon two-body docking (1E4K, 1IBR, 1KKL, 1NPE, 1DFJ). Interestingly two-body docking worked well for 1E4K but not for 1IBR, although both receptors experience the same extent of conformational change (2.9 Å). Possible explanations for this could be that: (1) The domain motion of 1E4K can be explained with only one hinge while more than one are needed for 1IBR and (2) the interface area of 1IBR is twice as large as that of 1E4K (1685 $Å^2$ against 810 $Å^2$). The performance of standard docking methods seems to be affected not only by the amount of conformational changes but also by their complexity and the extent of the interface to be predicted.

Other factors that may affect the docking performance are the size of the subdomains to be docked and the number of hinges at which a receptor is cut. Our results revealed that FMD is robust with respect to the size of the segments involved in docking. For 1Y64, 1ZLI, 1KKL, and 1NPE, one of the subdomains consists of less than 50 residues while the others have larger domains. Shorter segments are more prone to be flexible, thus making it harder to model the correct binding mode. Also our protocol is performing well when more than one hinge is involved in the domain motion as illustrated by the 1F6M case. FMD is quite successful to model a wide range of conformational changes (from 1.5 to 20 Å), in cases when only two domains are involved. It does not perform as well when the system complexity increases like for example for the former CAPRI target, 1TLV (Graille et al., 2005), which consists of a homodimer with each monomer experiencing a large conformational change of 12.7 Å upon activation. In this case, a four-body docking failed to generate any acceptable solution (best i-rmsd 6.2 Å and global rmsd 10.3 Å), where the multistage docking protocol of MolFit (Ben-Zeev et al., 2005) could generate a model with a global rmsd of 5.6 Å.

Another parameter that affects the docking quality is the interface information used to drive docking. In order to focus on our ability to deal with conformational changes, we assumed in this work perfect interface information; this is a best-case scenario and one can therefore expect that the performance will depend on the quality and reliability of the supplied data. We investigated this aspect by using bioinformatics interface predictions obtained from our in-house consensus predictor, CPORT, to drive docking (http://haddock.chem.uu.nl/services/CPORT/) (de Vries and Bonvin, 2011). We could only generate acceptable solutions for 1FAK, 1ZLI and 1KKL (see Table S1 available online). This is directly related to the specificity of the prediction (see Table S2).

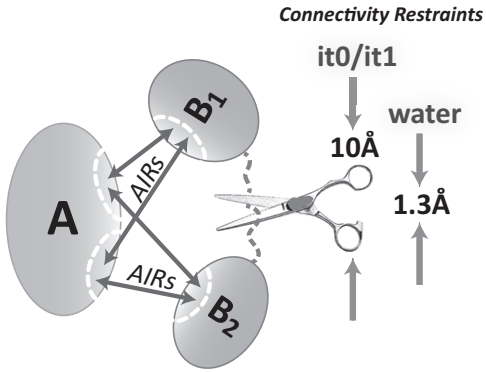

Figure 2. Schematic Representation of the Restraints Used in Docking (Ambiguous Interaction and Connectivity Restraints) and of the Dissection of the Flexible Partner into Subdomains

Only for four cases (1FAK, 1ZLI, 1KKL, and 1DFJ), the specificity of CPORT predictions was above 30% for both receptor and ligand and for three of these acceptable to medium quality solutions could be obtained. This is in line with what was observed in systematic HADDOCK runs, where CPORT was used to drive docking (de Vries and Bonvin, 2011).

In general, our FMD protocol could generate at least an acceptable solution for each case, and even medium-quality predictions for seven of them. The fraction of native contacts for the best models was above 0.5 (the CAPRI threshold for high quality predictions) for nine of them. The strength of FMD-HADDOCK resides mainly in its ability to deal with large-scale and small, induced conformational changes at the same time.

Our benchmark shares common cases with three other hybrid methodologies (see Introduction): 1NPE and 1IBR with FlexDock (see Table 2 in Schneidman-Duhovny et al., 2007); 1DFJ and 1IBR with ATTRACT (see Table 5 in May and Zacharias, 2008) and FiberDock (see Table 4 in Mashiach et al., 2010). Comparing the best l-rmsds for 1IBR, FlexDock, FiberDock and FMD produce similar somewhat better results than ATTRACT. FMD however outperforms FlexDock in scoring. A similar situation is observed for 1NPE. In the case of 1DFJ FMD, FiberDock and ATTRACT perform well in both sampling and scoring, although FiberDock's best model has a better l-rmsd.

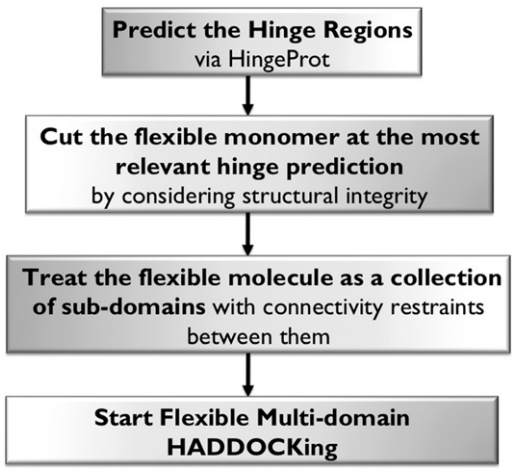

Figure 3. Workflow of Flexible Multidomain Docking in HADDOCK

Table 3. Comparison of Two-Body and Flexible Multidomain Docking Results

Flexible Multidomain Docking					
PDB ID	Quality / Rank[a]	i-rmsd (Å)	l-rmsd (Å)	Fnat	Fnonnat
1IRA	★ / 1	3.9	7.5	0.55	0.61
1H1V	★ / 11	4.6	9.6	0.49	0.67
1Y64	★ / 5	3.9	9.5	0.48	0.72
1F6M	★ / 1	3.5	8.7	0.69	0.58
1FAK	★★ / 37	2.8	4.2	0.55	0.54
	★ / 2	3.4	7.4	0.50	0.58
1ZLI	★★ / 1	2.1	3.5	0.74	0.47
1E4K	★★ / 5	2.3	4.0	0.70	0.47
	★ / 1	2.8	6.0	0.62	0.52
1IBR	★★ / 1	2.3	4.0	0.63	0.57
1KKL	★★ / 1	2.2	4.9	0.67	0.59
1NPE	★★ / 16	1.2	3.2	0.95	0.36
	★ / 1	5.5	7.1	0.74	0.51
1DFJ	★★ / 5	2.0	7.1	0.68	0.56
	★ / 1	2.3	8.3	0.60	0.61
Two-Body Docking[b]					
PDB ID	Quality / Rank	i-rmsd (Å)	l-rmsd (Å)	Fnat	Fnonnat
1IRA	- / 1	17.5	38.6	0.04	0.94
1H1V	- / 1	11.9	31.1	0.08	0.94
1Y64	- / 1	10.3	32.2	0.07	0.92
1F6M	- / 1	14.1	39.4	0.00	1.00
1FAK	- / 1	11.4	28.5	0.01	0.99
1ZLI	- / 1	14.8	25.3	0.02	0.99
1E4K	★ / 1	4.1	9.8	0.58	0.58
1IBR	- / 1	9.6	30.9	0.11	0.90
1KKL	★ / 1	3.1	12.3	0.56	0.60
1NPE	★★ / 1	1.7	6.1	0.85	0.45
1DFJ	★★ / 116	1.8	5.6	0.63	0.59
	★ / 1	2.5	7.1	0.62	0.64

The quality is expressed according to CAPRI criteria (see Experimental Procedures). The reported rank is the rank of the individual models prior to clustering. See also Tables S1–S3.

[a] The ranking is based on the HADDOCK score: 1.0 E_{vdw} + 0.2 E_{elec} + 1.0 E_{desol} (see Results).

[b] When no acceptable solution is generated, the values for the top ranked structure are reported.

The excellent performance of the new FMD-HADDOCK protocol represents already a major step in dealing with large conformational changes occurring upon binding. One question, however, remains: How can we predict which type of conformational change should be expected upon binding so that one can choose the best suited method? Dobbins et al. (2008) pointed out that Normal Mode Analysis can provide some discrimination among the various types of conformational changes. They stated that the slowest mode of a protein experiencing significant conformational changes (C_α-rmsd > 2.0 Å) has a 2.5 times lower

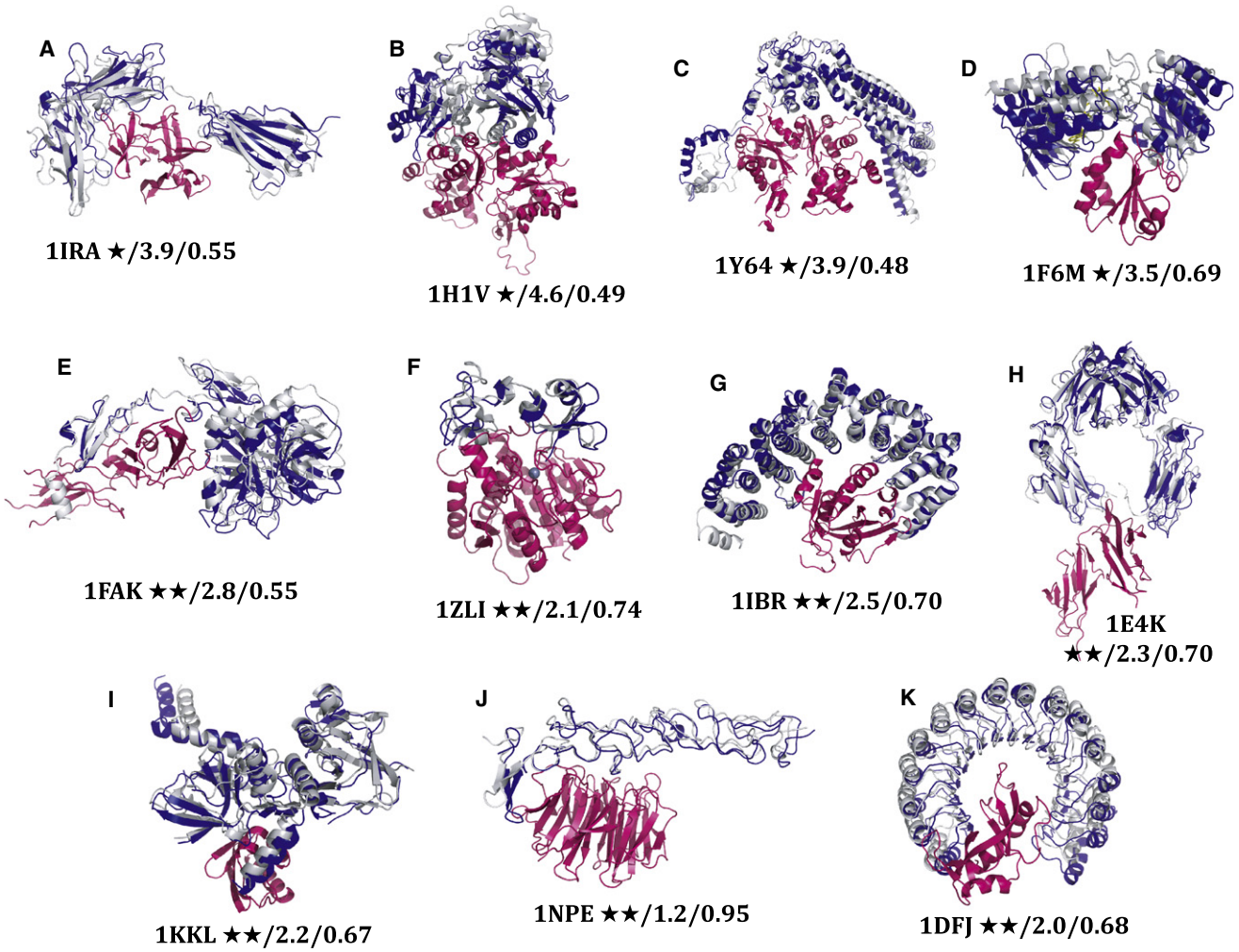

Figure 4. View of the Best HADDOCK Solutions Superimposed onto the Respective Reference Crystal Structures

The docked complexes are shown in magenta (ligand) and blue (receptor) and the reference crystal structure (only the receptor is shown) in gray: (A) 1IRA, (B) 1H1V, (C) 1Y64, (D) 1F6M, (E) 1FAK, (F) 1ZLI, (G) 1IBR, (H) 1E4K, (I) 1KKL, (J) 1NPE, (K) 1DFJ. The quality of the models is indicated as: number of stars / interface rmsd (in Å) / fraction of native contacts (see Experimental Procedures). The figures were generated with Pymol (DeLano, 2002).

frequency than for proteins undergoing limited conformational change (C_α-rmsd < 1.0 Å). To further investigate this issue, we considered a set of 268 proteins experiencing different levels of conformational changes (0.3 Å $\leq$ C_α-rmsd $\leq$ 19.5 Å). These were taken from the Docking Benchmark 4.0 (Hwang et al., 2010), excluding multimers, antibody-antigen complexes and proteins having less than 50 amino acids. For all single proteins, eigenvalues from a Gaussian Network Model (GNM) (see Experimental Procedures) (Haliloglu et al., 1997) were obtained from the HingeProt web server (Emekli et al., 2008). They were normalized and their cumulative sum (ranging from 0 to 1) was calculated and plotted as a function of the number of modes. This plot can be interpreted as the fraction of motion explained by a given number of modes. Here we assumed that proteins undergoing larger conformational changes have smaller GNM-eigenvalues, corresponding to collective low frequency motions. For these, the cumulative sum of the normalized GNM-eigenvalues should increase slower as a function of the number of modes than for rigid cases, in agreement with the observations of Dobbins et al. (2008) (Figure S1). To test this assumption, we developed a simple predictor, which predicts proteins as being rigid or flexible based on the fraction of motion explained by a given number of modes (see Supplemental Experimental Procedures). This classifier performs with the best accuracy (65 ± 16%, cross-validated accuracy) on the set of 268 proteins when a 2.6 Å rmsd cutoff is used to classify these proteins as *rigid* and *flexible* and the cutoff for the *fraction of motion explained* for the first 50 modes is taken as 0.08 (i.e., the cumulative sum of the eigenvalues for the first 50 modes should be <0.08 for the protein to be predicted as *flexible*). Our simple predictor was able to classify the receptors of the FMD benchmark with 64% accuracy (with seven out of nine flexible cases and one out of two rigid cases being correctly classified, see Figure 5). The cumulative sum of the normalized eigenvalues for a given number of modes may thus be deterministic in predicting the range of the conformational change that could be expected.

Table 4. Cluster Statistics of Flexible Multidomain Docking

PDB ID	Quality / Rank	i-rmsd (Å)	l-rmsd (Å)	Fnat	Fnonnat
1IRA	★ / 1	4.2 ± 0.4	8.0 ± 1.0	0.54 ± 0.03	0.64 ± 0.07
1H1V	★ / 2	4.8 ± 0.2	10.4 ± 1.2	0.50 ± 0.06	0.66 ± 0.06
1Y64	★ / 1	4.8 ± 0.6	11.3 ± 1.0	0.44 ± 0.06	0.74 ± 0.03
1F6M	★ / 1	3.1 ± 0.3	9.2 ± 0.8	0.49 ± 0.08	0.64 ± 0.04
1FAK	★ / 2	3.7 ± 0.5	9.8 ± 4.5	0.48 ± 0.02	0.61 ± 0.05
	★ / 1	4.2 ± 0.2	14.1 ± 1.4	0.49 ± 0.06	0.64 ± 0.07
1ZLI	★★ / 1	2.2 ± 0.2	3.8 ± 0.4	0.77 ± 0.03	0.47 ± 0.02
1E4K	★ / 1	2.5 ± 0.2	5.9 ± 0.1	0.64 ± 0.05	0.46 ± 0.04
1IBR	★★ / 1	2.7 ± 0.6	5.1 ± 1.1	0.67 ± 0.04	0.57 ± 0.03
1KKL	★★ / 1	2.4 ± 0.2	5.4 ± 0.5	0.69 ± 0.05	0.56 ± 0.03
1NPE	★★ / 2	2.2 ± 1.1	6.5 ± 4.7	0.82 ± 0.22	0.46 ± 0.06
	★ / 1	6.1 ± 0.4	8.9 ± 1.0	0.73 ± 0.03	0.46 ± 0.05
1DFJ	★ / 1	3.2 ± 0.8	9.1 ± 0.8	0.59 ± 0.02	0.67 ± 0.02

The quality is expressed according to the CAPRI criteria (see Experimental Procedures). The rank corresponds to the cluster ranking based on the average HADDOCK score (1.0 E_{vdw} + 0.2 E_{elec} + 1.0 E_{desol}) of the top four members of a cluster. Rmsd and Fnat are values reported as averages ± standard deviation calculated over the top four members of a cluster.

Table 5. Number of Docking Models of Various Qualities for the Various Stages of the Flexible Multidomain Docking Protocol

PDB ID	Quality (★ / ★★ / ★★★) *it0*	*it1*	*water*
1IRA	27 / - / -	43 / - / -	44 / - / -
1H1V	50 / - / -	108 / - / -	116 / - / -
1Y64	- / - / -	1 / - / -	1 / - / -
1F6M	2096 / - / -	382 / - / -	382 / - / -
1FAK	151 / 3 / -	70 / 11 / -	71 / 12 / -
1ZLI	1496 / 3246 / -	5 / 395 / -	3 / 397 / -
1E4K	2569 / 1371 / -	87 / 212 / -	98 / 201 / -
1IBR	1287 / - / -	195 / 104 / -	196 / 103 / -
1KKL	521 / 1844 / -	323 / 77 / -	304 / 96 / -
1NPE	984 / 32 / 1	356 / 26 / -	353 / 30 / -
1DFJ	210 / - / -	96 / - / -	116 / 1 / -
1DFJ	210 / - / -	96 / - / -	116 / 1 / -

it0, Rigid multidomain docking; *it1,* semiflexible refinement in torsion angle space; *water,* final explicit solvent refinement.

This could provide a mean to select an appropriate method to deal with different types of expected conformational changes.

Conclusions

We have developed a new flexible multidomain docking protocol that follows a "divide-and-conquer" approach to model large-scale domain motions and small- to medium-scale rearrangements within an interface at the same time. This represents a major step in dealing with one of the major challenges and limitations in the modeling of biomolecular complexes. Our analysis has also identified indicators that might allow us to predict the extent of the conformational changes and thus select the most appropriate method to deal with them. This remains, however, a challenging problem as a variety of approaches will still have to be combined in order to properly describe simultaneously all possible types of changes. The availability of some experimental information will become crucial to drive the docking and/or for validation of the resulting models as the complexity of the problem increases. As a final remark we note that flexible multidomain docking as implemented in HADDOCK 2.1 is now also accessible via the multibody docking interface of the HADDOCK Web server (http://haddock.chem.uu.nl/services/HADDOCK/haddock.php) (de Vries et al., 2010).

EXPERIMENTAL PROCEDURES

Determination of the Hinge Regions

The hinge prediction server HingeProt (http://www.prc.boun.edu.tr/appserv/prc/hingeprot/)(Emekli et al., 2008) was used to define the hinge regions of the flexible monomers. HingeProt annotates rigid parts and possible hinge regions of the supplied protein based on two elastic network models: GNM (Haliloglu et al., 1997) and Anisotropic Network Model (Atilgan et al., 2001). In this work, two different outputs provided by the HingeProt server are used: the predicted hinge regions and the eigenvalues obtained by the decomposition of the connectivity (Kirchhoff) matrix (Haliloglu et al., 1997). The hinge predictions obtained from HingeProt are filtered in order to preserve the secondary structure integrity. Insignificant and/or structurally unreliable predictions are eliminated. The detailed procedure is described in the Results section.

Docking Protocol of HADDOCK

Both two-body docking and FMD were performed with HADDOCK 2.1 (de Vries et al., 2010; Dominguez et al., 2003; Karaca et al., 2010). The interface information used to construct the AIRs was extracted from the crystal structure of the complexes. We assumed to have an ideal definition of the interacting surfaces in order to concentrate on our ability to deal with large conformational changes. The AIR definitions are provided in Table S3. For the runs with bioinformatics predictions, we used CPORT (de Vries and Bonvin, 2011) (http://haddock.chem.uu.nl/services/CPORT/). In the FMD runs, center-of-mass restraints between the domains were turned on to ensure compactness of the solutions. Center-of-mass restraints are defined as a distance restraint between the geometric average positions of all C_α atoms within each molecule. The distance is automatically defined based on the dimension of the molecules or domains. The number of structures was increased to 5000, 400, and 400 for *it0*, *it1*, and *water*, respectively. Random removal of AIRs was turned off since ideal restraints were used. Other parameters were left to their default values. Scoring and clustering were performed according to standard HADDOCK procedures (de Vries et al., 2007; Dominguez et al., 2003).

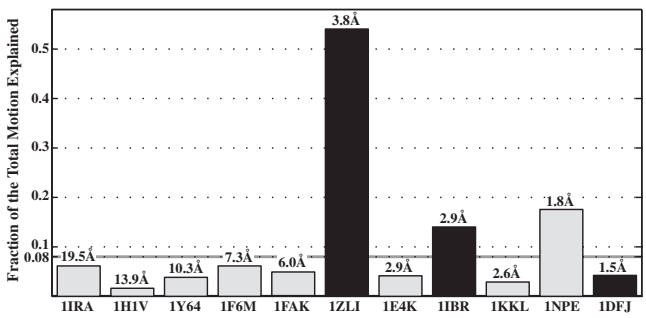

Figure 5. Fraction of Motion Explained by the First 50 GNM Modes for the 11 Receptors of the FMD Benchmark

Gray bars indicate correctly classified; black bars indicate the misclassified cases. The extent of the conformational change is indicated on top of each bar. See also Figure S1 and Supplemental Experimental Procedures.

Assessment of the Structure Quality

The docking models were evaluated according to CAPRI criteria (Mendez et al., 2003):

- Acceptable prediction (one star): i-rmsd $\leq$ 4 Å or l-rmsd $\leq$ 10 Å and Fnat $\geq$ 0.1
- Good prediction (two stars): i-rmsd $\leq$ 2 Å or l-rmsd $\leq$ 5 Å and Fnat $\geq$ 0.3
- High-quality prediction (three stars): i-rmsd $\leq$ 1 Å or l-rmsd $\leq$ 1 Å and nat $\geq$ 0.5.

i-rmsd refers to the interface rmsd, l-rmsd to the ligand rmsd, calculated over the backbone atoms of the ligand (rigid component) after fitting on the receptor, and Fnat to the fraction of native contacts. Next to Fnat we also provided Fnonnat, which is the fraction of incorrectly predicted contacts given the modeled complex. A cluster was considered of one-, two-, or three-star quality if at least one of its top four members was of the corresponding quality.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, one figure, and three tables and can be found with this article online at doi:10.1016/j.str.2011.01.014.

ACKNOWLEDGMENTS

This work was supported by the Netherlands Organization for Scientific Research (NWO) (VICI grant 700.56.442 to A.M.J.J.B.) and the European Community (FP7 FP7 e-Infrastructure "e-NMR" and "WeNMR" projects, grant numbers 213010 and 21301). Support from the national GRID Initiatives of Belgium, Italy, Germany, the Netherlands (via the Dutch BiG Grid project), Portugal, UK, South Africa, is acknowledged for the use of computing and storage facilities.

Received: October 8, 2010
Revised: January 3, 2011
Accepted: January 10, 2011
Published: April 12, 2011

REFERENCES

Andrusier, N., Mashiach, E., Nussinov, R., and Wolfson, H.J. (2008). Principles of flexible protein-protein docking. Proteins *73*, 271–289.

Arolas, J.L., Popowicz, G.M., Lorenzo, J., Sommerhoff, C.P., Huber, R., Aviles, F.X., and Holak, T.A. (2005). The three-dimensional structures of tick carboxypeptidase inhibitor in complex with A/B carboxypeptidases reveal a novel double-headed binding mode. J. Mol. Biol. *350*, 489–498.

Atilgan, A.R., Durell, S.R., Jernigan, R.L., Demirel, M.C., Keskin, O., and Bahar, I. (2001). Anisotropy of fluctuation dynamics of proteins with an elastic network model. Biophys. J. *80*, 505–515.

Bachar, O., Fischer, D., Nussinov, R., and Wolfson, H. (1993). A computer vision based technique for 3-D sequence-independent structural comparison of proteins. Protein Eng. *6*, 279–288.

Barbosa Pereira, P.J., Segura-Martin, S., Oliva, B., Ferrer-Orta, C., Aviles, F.X., Coll, M., Gomis-Ruth, F.X., and Vendrell, J. (2002). Human procarboxypeptidase B: three-dimensional structure and implications for thrombin-activatable fibrinolysis inhibitor (TAFI). J. Mol. Biol. *321*, 537–547.

Bastard, K., Thureau, A., Lavery, R., and Prevost, C. (2003). Docking macromolecules with flexible segments. J. Comput. Chem. *24*, 1910–1920.

Bayliss, R., Littlewood, T., and Stewart, M. (2000). Structural basis for the interaction between FxFG nucleoporin repeats and importin-beta in nuclear trafficking. Cell *102*, 99–108.

Ben-Zeev, E., Kowalsman, N., Ben-Shimon, A., Segal, D., Atarot, T., Noivirt, O., Shay, T., and Eisenstein, M. (2005). Docking to single-domain and multiple-domain proteins: old and new challenges. Proteins *60*, 195–201.

Bonvin, A.M. (2006). Flexible protein-protein docking. Curr. Opin. Struct. Biol. *16*, 194–200.

Bubb, M.R., Govindasamy, L., Yarmola, E.G., Vorobiev, S.M., Almo, S.C., Somasundaram, T., Chapman, M.S., Agbandje-McKenna, M., and McKenna, R. (2002). Polylysine induces an antiparallel actin dimer that nucleates filament assembly: crystal structure at 3.5-A resolution. J. Biol. Chem. *277*, 20999–21006.

Burtnick, L.D., Koepf, E.K., Grimes, J., Jones, E.Y., Stuart, D.I., McLaughlin, P.J., and Robinson, R.C. (1997). The crystal structure of plasma gelsolin: implications for actin severing, capping, and nucleation. Cell *90*, 661–670.

Chaudhury, S., and Gray, J.J. (2008). Conformer selection and induced fit in flexible backbone protein-protein docking using computational and NMR ensembles. J. Mol. Biol. *381*, 1068–1087.

Choe, H., Burtnick, L.D., Mejillano, M., Yin, H.L., Robinson, R.C., and Choe, S. (2002). The calcium activation of gelsolin: insights from the 3A structure of the G4-G6/actin complex. J. Mol. Biol. *324*, 691–702.

Das, R., Andre, I., Shen, Y., Wu, Y., Lemak, A., Bansal, S., Arrowsmith, C.H., Szyperski, T., and Baker, D. (2009). Simultaneous prediction of protein folding and docking at high resolution. Proc. Natl. Acad. Sci. USA *106*, 18978–18983.

de Vries, S.J., van Dijk, A.D.J., Krzeminski, M., van Dijk, M., Thureau, A., Hsu, V., Wassenaar, T., and Bonvin, A.M.J.J. (2007). HADDOCK versus HADDOCK: new features and performance of HADDOCK2.0 on the CAPRI targets. Proteins *69*, 726–733.

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

de Vries, S.J., and Bonvin, A.M.J.J. (2011). CPORT: a consensus interface predictor and its performance in prediction-driven docking with HADDOCK. PLoS ONE, in press.

DeLano, W.L. (2002). The PyMOL Molecular Graphics System on World Wide Web (http://www.pymol.org).

Dobbins, S.E., Lesk, V.I., and Sternberg, M.J. (2008). Insights into protein flexibility: the relationship between normal modes and conformational change upon protein-protein docking. Proc. Natl. Acad. Sci. USA *105*, 10390–10395.

Dominguez, C., Boelens, R., and Bonvin, A.M.J.J. (2003). HADDOCK: a protein-protein docking approach based on biochemical or biophysical information. J. Am. Chem. Soc. *125*, 1731–1737.

Emekli, U., Schneidman-Duhovny, D., Wolfson, H.J., Nussinov, R., and Haliloglu, T. (2008). HingeProt: automated prediction of hinges in protein structures. Proteins *70*, 1219–1227.

Fernandez-Recio, J., Totrov, M., and Abagyan, R. (2004). Identification of protein-protein interaction sites from docking energy landscapes. J. Mol. Biol. *335*, 843–865.

Fieulaine, S., Morera, S., Poncet, S., Mijakovic, I., Galinier, A., Janin, J., Deutscher, J., and Nessler, S. (2002). X-ray structure of a bifunctional protein kinase in complex with its protein substrate HPr. Proc. Natl. Acad. Sci. USA *99*, 13437–13441.

Fieulaine, S., Morera, S., Poncet, S., Monedero, V., Gueguen-Chaignon, V., Galinier, A., Janin, J., Deutscher, J., and Nessler, S. (2001). X-ray structure of HPr kinase: a bacterial protein kinase with a P-loop nucleotide-binding domain. EMBO J. *20*, 3917–3927.

Fischer, E. (1894). Einfluss der Configuration auf die Wirkung der Enzyme. (Berichte der deutschen Chemischen Gesellschaft) , pp. 2985–2993.

Gerstein, M., and Echols, N. (2004). Exploring the range of protein flexibility, from a structural proteomics perspective. Curr. Opin. Chem. Biol. *8*, 14–19.

Gerstein, M., and Krebs, W. (1998). A database of macromolecular motions. Nucleic Acids Res. *26*, 4280–4290.

Graille, M., Zhou, C.Z., Receveur-Brechot, V., Collinet, B., Declerck, N., and van Tilbeurgh, H. (2005). Activation of the LicT transcriptional antiterminator involves a domain swing/lock mechanism provoking massive structural changes. J. Biol. Chem. *280*, 14780–14789.

Gray, J.J., Moughon, S., Wang, C., Schueler-Furman, O., Kuhlman, B., Rohl, C.A., and Baker, D. (2003). Protein-protein docking with simultaneous optimization of rigid-body displacement and side-chain conformations. J. Mol. Biol. *331*, 281–299.

Grunberg, R., Nilges, M., and Leckner, J. (2006). Flexibility and conformational entropy in protein-protein binding. Structure *14*, 683–693.

Haliloglu, T., Bahar, I., and Erman, B. (1997). Gaussian dynamics of folded proteins. Phys. Rev. Lett. *79*, 3090–3093.

Hayward, S. (1999). Structural principles governing domain motions in proteins. Proteins *36*, 425–435.

Huang, M., Syed, R., Stura, E.A., Stone, M.J., Stefanko, R.S., Ruf, W., Edgington, T.S., and Wilson, I.A. (1998). The mechanism of an inhibitory antibody on TF-initiated blood coagulation revealed by the crystal structures of human tissue factor, Fab 5G9 and TF.G9 complex. J. Mol. Biol. *275*, 873–894.

Hwang, H., Vreven, T., Janin, J., and Weng, Z. (2010). Protein-protein docking benchmark version 4.0. Proteins *78*, 3111–3114.

Karaca, E., Melquiond, A.S., de Vries, S.J., Kastritis, P.L., and Bonvin, A.M. (2010). Building macromolecular assemblies by information-driven docking: introducing the HADDOCK multi-body docking server. Mol. Cell. Proteomics *9*, 1784–1794.

Katchalski-Katzir, E., Shariv, I., Eisenstein, M., Friesem, A.A., Aflalo, C., and Vakser, I.A. (1992). Molecular-Surface Recognition - Determination of Geometric Fit between Proteins and Their Ligands by Correlation Techniques. Proc. Natl. Acad. Sci. USA *89*, 2195–2199.

Kent, H.M., Moore, M.S., Quimby, B.B., Baker, A.M., McCoy, A.J., Murphy, G.A., Corbett, A.H., and Stewart, M. (1999). Engineered mutants in the switch II loop of Ran define the contribution made by key residues to the interaction with nuclear transport factor 2 (NTF2) and the role of this interaction in nuclear protein import. J. Mol. Biol. *289*, 565–577.

Kobe, B., and Deisenhofer, J. (1995). A structural basis of the interactions between leucine-rich repeats and protein ligands. Nature *374*, 183–186.

Kobe, B., and Deisenhofer, J. (1996). Mechanism of ribonuclease inhibition by ribonuclease inhibitor protein based on the crystal structure of its complex with ribonuclease A. J. Mol. Biol. *264*, 1028–1043.

Koshland, D.E. (1958). Application of a theory of enzyme specificity to protein synthesis. Proc. Natl. Acad. Sci. USA *44*, 98–104.

Krol, M., Chaleil, R.A., Tournier, A.L., and Bates, P.A. (2007). Implicit flexibility in protein docking: cross-docking and local refinement. Proteins *69*, 750–757.

Kumar, S., Ma, B., Tsai, C.J., Sinha, N., and Nussinov, R. (2000). Folding and binding cascades: dynamic landscapes and population shifts. Protein Sci. *9*, 10–19.

Lennon, B.W., Williams, C.H., Jr., and Ludwig, M.L. (1999). Crystal structure of reduced thioredoxin reductase from Escherichia coli: structural flexibility in the isoalloxazine ring of the flavin adenine dinucleotide cofactor. Protein Sci. *8*, 2366–2379.

Lennon, B.W., Williams, C.H., Jr., and Ludwig, M.L. (2000). Twists in catalysis: alternating conformations of Escherichia coli thioredoxin reductase. Science *289*, 1190–1194.

Lensink, M.F., and Mendez, R. (2008). Recognition-induced conformational changes in protein-protein docking. Curr. Pharm. Biotechnol. *9*, 77–86.

Li, X., Moal, I.H., and Bates, P.A. (2010). Detection and refinement of encounter complexes for protein-protein docking: taking account of macromolecular crowding. Proteins *78*, 3189–3196.

Liao, D.I., and Herzberg, O. (1994). Refined structures of the active Ser83→Cys and impaired Ser46→Asp histidine-containing phosphocarrier proteins. Structure *2*, 1203–1216.

Mashiach, E., Nussinov, R., and Wolfson, H.J. (2010). FiberDock: Flexible induced-fit backbone refinement in molecular docking. Proteins *78*, 1503–1519.

Matsumiya, S., Yamaguchi, Y., Saito, J., Nagano, M., Sasakawa, H., Otaki, S., Satoh, M., Shitara, K., and Kato, K. (2007). Structural comparison of fucosylated and nonfucosylated Fc fragments of human immunoglobulin G1. J. Mol. Biol. *368*, 767–779.

May, A., and Zacharias, M. (2008). Energy minimization in low-frequency normal modes to efficiently allow for global flexibility during systematic protein-protein docking. Proteins *70*, 794–809.

May, A., and Zacharias, M. (2007). Protein-protein docking in CAPRI using ATTRACT to account for global and local flexibility. Proteins *69*, 774–780.

Mendez, R., Leplae, R., De Maria, L., and Wodak, S.J. (2003). Assessment of blind predictions of protein-protein interactions: current status of docking methods. Proteins *52*, 51–67.

Mittag, T., Kay, L.E., and Forman-Kay, J.D. (2010). Protein dynamics and conformational disorder in molecular recognition. J. Mol. Recognit. *23*, 105–116.

Moal, I.H., and Bates, P.A. (2010). Swarmdock and the use of normal modes in protein-protein docking. Int. J. Mol. Sci. *11*, 3623–3648.

Moreira, I.S., Fernandes, P.A., and Ramos, M.J. (2010). Protein-protein docking dealing with the unknown. J. Comput. Chem. *31*, 317–342.

Nachman, J., Miller, M., Gilliland, G.L., Carty, R., Pincus, M., and Wlodawer, A. (1990). Crystal structure of two covalent nucleoside derivatives of ribonuclease A. Biochemistry *29*, 928–937.

Nikkola, M., Gleason, F.K., Fuchs, J.A., and Eklund, H. (1993). Crystal structure analysis of a mutant Escherichia coli thioredoxin in which lysine 36 is replaced by glutamic acid. Biochemistry *32*, 5093–5098.

Otomo, T., Tomchick, D.R., Otomo, C., Panchal, S.C., Machius, M., and Rosen, M.K. (2005). Structural basis of actin filament nucleation and processive capping by a formin homology 2 domain. Nature *433*, 488–494.

Pantoja-Uceda, D., Arolas, J.L., Garcia, P., Lopez-Hernandez, E., Padro, D., Aviles, F.X., and Blanco, F.J. (2008). The NMR structure and dynamics of the two-domain tick carboxypeptidase inhibitor reveal flexibility in its free form and stiffness upon binding to human carboxypeptidase B. Biochemistry *47*, 7066–7078.

Pike, A.C., Brzozowski, A.M., Roberts, S.M., Olsen, O.H., and Persson, E. (1999). Structure of human factor VIIa and its implications for the triggering of blood coagulation. Proc. Natl. Acad. Sci. USA *96*, 8925–8930.

Pons, C., Grosdidier, S., Solernou, A., Perez-Cano, L., and Fernandez-Recio, J. (2010). Present and future challenges and limitations in protein-protein docking. Proteins *78*, 95–108.

Ritchie, D.W. (2008). Recent progress and future directions in protein-protein docking. Curr. Protein Pept. Sci. *9*, 1–15.

Rizvi, S.A., Tereshko, V., Kossiakoff, A.A., and Kozmin, S.A. (2006). Structure of bistramide A-actin complex at a 1.35 angstroms resolution. J. Am. Chem. Soc. *128*, 3882–3883.

Schneidman-Duhovny, D., Inbar, Y., Nussinov, R., and Wolfson, H.J. (2005). PatchDock and SymmDock: servers for rigid and symmetric docking. Nucleic Acids Res. *33*, W363–W367.

Schneidman-Duhovny, D., Nussinov, R., and Wolfson, H.J. (2007). Automatic prediction of protein interactions with large scale motion. Proteins *69*, 764–773.

Schreuder, H.A., Rondeau, J.M., Tardif, C., Soffientini, A., Sarubbi, E., Akeson, A., Bowlin, T.L., Yanofsky, S., and Barrett, R.W. (1995). Refined crystal structure of the interleukin-1 receptor antagonist. Presence of a disulfide link and a cis-proline. Eur. J. Biochem. *227*, 838–847.

Schreuder, H., Tardif, C., Trump-Kallmeyer, S., Soffientini, A., Sarubbi, E., Akeson, A., Bowlin, T., Yanofsky, S., and Barrett, R.W. (1997). A new cytokine-receptor binding mode revealed by the crystal structure of the IL-1 receptor with an antagonist. Nature *386*, 194–200.

Sondermann, P., Huber, R., Oosthuizen, V., and Jacob, U. (2000). The 3.2-A crystal structure of the human IgG1 Fc fragment-Fc gammaRIII complex. Nature *406*, 267–273.

Stetefeld, J., Mayer, U., Timpl, R., and Huber, R. (1996). Crystal structure of three consecutive laminin-type epidermal growth factor-like (LE) modules of laminin gamma1 chain harboring the nidogen binding site. J. Mol. Biol. *257*, 644–657.

Takagi, J., Yang, Y., Liu, J.H., Wang, J.H., and Springer, T.A. (2003). Complex between nidogen and laminin fragments reveals a paradigmatic beta-propeller interface. Nature *424*, 969–974.

Vetter, I.R., Arndt, A., Kutay, U., Gorlich, D., and Wittinghofer, A. (1999). Structural view of the Ran-Importin beta interaction at 2.3 A resolution. Cell *97*, 635–646.

Vigers, G.P., Dripps, D.J., Edwards, C.K., 3rd, and Brandhuber, B.J. (2000). X-ray crystal structure of a small antagonist peptide bound to interleukin-1 receptor type 1. J. Biol. Chem. *275*, 36927–36933.

Wang, C., Bradley, P., and Baker, D. (2007). Protein-protein docking with backbone flexibility. J. Mol. Biol. *373*, 503–519.

Xu, Y., Moseley, J.B., Sagot, I., Poy, F., Pellman, D., Goode, B.L., and Eck, M.J. (2004). Crystal structures of a Formin Homology-2 domain reveal a tethered dimer architecture. Cell *116*, 711–723.

Zacharias, M. (2010). Accounting for conformational changes during protein-protein docking. Curr. Opin. Struct. Biol. *20*, 180–186.

Zhang, E., St Charles, R., and Tulinsky, A. (1999). Structure of extracellular tissue factor complexed with factor VIIa inhibited with a BPTI mutant. J. Mol. Biol. *285*, 2089–2104.

Zhang, Y., Boesen, C.C., Radaev, S., Brooks, A.G., Fridman, W.H., Sautes-Fridman, C., and Sun, P.D. (2000). Crystal structure of the extracellular domain of a human Fc gamma RIII. Immunity *13*, 387–395.

