ELSEVIER

Flexible protein–protein docking

Alexandre MJJ Bonvin

Predicting the structure of protein–protein complexes using docking approaches is a difficult problem whose major challenges include identifying correct solutions, and properly dealing with molecular flexibility and conformational changes. Flexibility can be addressed at several levels: implicitly, by smoothing the protein surfaces or allowing some degree of interpenetration (soft docking) or by performing multiple docking runs from various conformations (cross or ensemble docking); or explicitly, by allowing sidechain and/or backbone flexibility. Although significant improvements have been achieved in the modeling of sidechains, methods for the explicit inclusion of backbone flexibility in docking are still being developed. A few novel approaches have emerged involving collective degrees of motion, multicopy representations and multibody docking, which should allow larger conformational changes to be modeled.

Addresses
Bijvoet Center for Biomolecular Research, Science Faculty, Utrecht University, NL-3584 CH, Utrecht, The Netherlands

Corresponding author: Bonvin, Alexandre MJJ
(a.m.j.j.bonvin@chem.uu.nl)
http://www.nmr.chem.uu.nl/~abonvin

Current Opinion in Structural Biology 2006, **16**:194–200

This review comes from a themed issue on
Theory and simulation
Edited by Joel Janin and Michael Levitt

Available online 17th February 2006

0959-440X/$ – see front matter

DOI 10.1016/j.sbi.2006.02.002

Introduction

Given the increased focus on interactions in the current post-genomic era, structural knowledge of complexes is required to understand how the various biomolecular units work together to fulfill their tasks. The number of expected biomolecular complexes will, however, exceed the number of proteins in a proteome by at least one order of magnitude; a significant fraction of these will be extremely difficult to study using classical structural methods such as NMR and X-ray crystallography. Therefore, the importance of computational approaches such as docking, the process of predicting the three-dimensional structure of a complex based on its known constituents, is evident. Unfortunately, predicting the structure of protein–protein complexes is a difficult problem, with major challenges that include identifying correct solutions, and properly dealing with flexibility and conformational changes. In this review, recent progress in the latter area will be addressed.

To monitor the performance of current docking methods, CAPRI (Critical Assessment of Predicted Interactions), a community-wide blind docking experiment, has been established (http://capri.ebi.ac.uk). The recent CAPRI results [1•] indicate that, although for 'easy' targets that show only small backbone conformational changes, excellent predictions can be obtained by the modeling community as whole, targets for which conformational changes take place upon binding are extremely challenging (even for backbone RMSD changes as small as 2 Å!). Initially, most protein–protein docking approaches have been developed based on rigid-body docking algorithms, thus ignoring any conformational change that might occur upon binding. However, the realization of the importance of flexibility in docking is leading to new developments. Flexibility can be introduced at several levels: implicitly, by smoothing the protein surfaces or allowing some degree of interpenetration (soft docking) or by performing multiple docking runs from various conformations (cross or ensemble docking); or explicitly, by allowing sidechain and/or backbone flexibility, either during docking or in a refinement step. A few novel approaches are emerging that involve collective degrees of motion or multicomponent rigid-body docking with flexible hinges. One common denominator here is that some *a priori* knowledge of flexible and/or hinge regions is often required. I will therefore first discuss recent progress in the prediction of flexibility and conformational changes in biomolecules before reviewing the various implementations of flexibility in docking. Note that flexibility also plays a major role in small-molecule docking; however, because of the smaller molecular sizes, the problem is more tractable (but not less challenging!) and many of the developments that I report here have been previously implemented in protein–ligand docking (for an overview, see [2,3]).

Analyzing and predicting protein flexibility

Introducing flexibility in protein–protein docking is facilitated if knowledge of flexible regions and possible conformational changes is available. Such knowledge can be derived *in silico* from molecular dynamics (MD) simulations in combination with principal components analysis. These simulations are, however, computationally demanding and limited in terms of amplitudes of motions, which makes them less suited to large-scale application and/or large systems. Fast and simpler alternatives are now emerging within the context of docking. Provided that several structures of a protein and/or its

homologues have been solved, information about flexible and hinge regions can be derived from three-dimensional structure similarity analysis [4]. Analysis of collective motions (e.g. by normal mode or principal components analysis) based on simplified descriptions of the molecules, such as graph theory [5] and elastic network [6] models, provides a computationally cheap alternative to MD simulations [5,7–10]. For identifying flexible and hinge regions, methods that limit the analysis to backbone or Cα atoms are usually sufficient [8,9]. However, if the aim is to generate starting conformations for docking, all (heavy) atom models are required, and several tools and web servers are now available for this purpose [11–13].

Implicit treatment of flexibility in docking

Soft docking

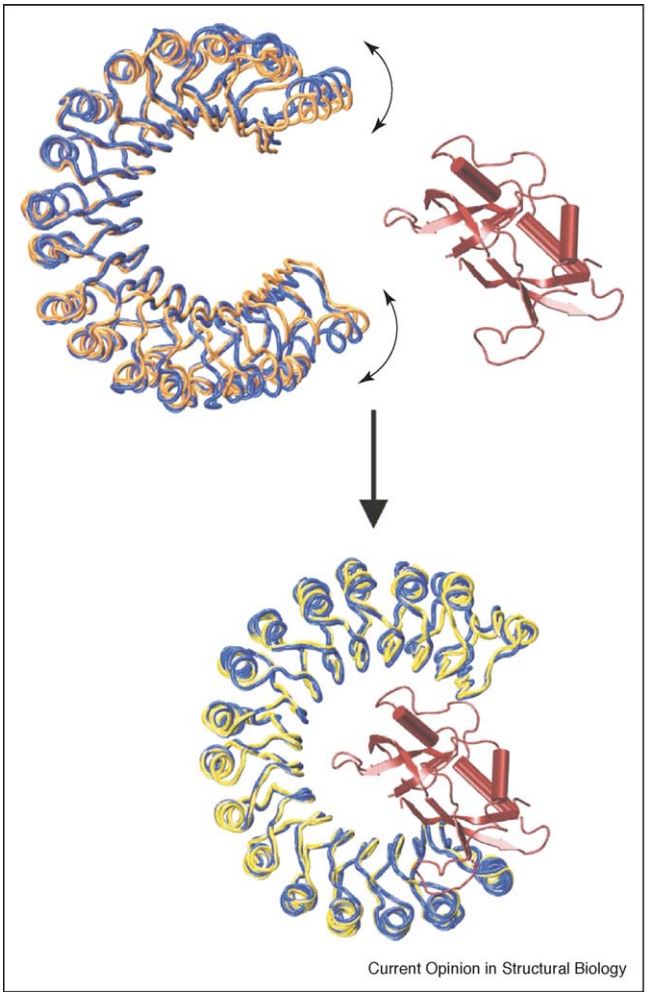
Many docking methods map the structures of the receptor and/or ligand onto a three-dimensional grid to simplify the search problem. In such representations, sidechain (and small backbone) rearrangements can be modeled by allowing some degree of interpenetration of the surfaces to be matched, the so-called soft-core approach [14], or by trimming long sidechains [15]. Snapshots of an MD simulation have recently been used to build grids for docking in which only grid points consistently occupied by all conformations are considered, thus excluding mobile regions from the construction of the grid [16]. This resulted in an increased hit rate and prediction accuracy for one of the recent CAPRI targets [16]. Note that surface softening does not always lead to improved results [17•]. One of the drawbacks of this kind of flexibility treatment is that severe steric clashes are frequently introduced when returning to a full-atom representation, thus requiring further optimization.

Ensemble docking

Another, rather different, implicit treatment of flexibility can be achieved by performing (rigid-body) docking of ensembles of conformations taken from, for example, NMR structures, MD simulations or any other conformational sampling method. In such cases, the docking process is repeated from various combinations of starting structures (ensemble or cross docking), which can lead to a significant increase in computing time, depending on the implementation. The conformations can span various degrees of flexibility, from small mainly sidechain rearrangements to large-scale global backbone motions.

Even small sidechain rearrangements from short MD simulations can be beneficial for improving docking predictions. It has been shown that residues important for molecular recognition will usually sample bound conformations during MD simulations of the unbound protein [18•,19]. Similarly, exposed sidechains in an ensemble of NMR structures usually sample various conformations; their use in data-driven docking was shown to increase both accuracy and hit rate [20]. Two rather systematic studies have been published that investigate the use of MD structures in cross docking [17•,21•]. Both studies indicate that ensemble docking improves the performance in terms of the number of (near) native solutions, but also leads to increased numbers of false positives (wrong solutions with high scores) and therefore makes the scoring more difficult. In addition, no clear correlation was found between RMSD from the bound form and success rate. Several groups have made use of MD snapshots as the starting point for docking in the latest CAPRI rounds [22,23,24•].

Figure 1



Normal-mode-based docking of RNase A to the porcine RNase A inhibitor (PDB code 1DFJ) using ATTRACT [41••]. Simultaneous docking minimization of RNase A (red cartoon representation) in translational and rotational degrees of freedom, and against the softest global mode of the unbound inhibitor yielded results in closer agreement with the bound form (blue) than simple rigid-body docking. The flexible global modes of the receptor (inhibitor) were calculated using the approximate normal mode method of Hinsen *et al*. [43]. The unbound inhibitor (yellow coil, top) and the docked conformation after optimization against soft global modes (orange coil, bottom) are compared to the bound crystal form (blue coils).

Ensembles of structures generated from collective-motion-based methods (see above) have also been applied to protein–protein docking. These usually sample larger conformational changes involving backbone atoms. Structures obtained from flexibility analysis based on the graph theory algorithm FIRST [5], coupled with conformational sampling using ROCK [11], were shown to improve the quality of results for protein docking with both small molecules and peptides [25•]. Similarly, conformations generated from essential dynamics analysis [26] of an ensemble of structures calculated with CONCOORD [7] have been used in CAPRI [27].

Explicit treatment of flexibility in docking

In the final refinement step of most docking approaches, a limited degree of flexibility is introduced by performing a (usually) short energy minimization (EM). This typically does not lead to any significant improvement in RMSD from the target structure, but rather is meant to remove clashes and improve the energetics, which can have a significant impact on the scoring performance [28]. In the past few years, flexibility has been explicitly introduced in the docking process. This is, however, only possible when molecules are explicitly represented (rather than, for example, on a grid).

Sidechain flexibility

In ICM-DISCO, docking against a soft grid is followed by Monte Carlo (MC) optimization of the ligand sidechains [29]. This procedure is quite successful in reproducing induced changes in surface sidechains as long as no large backbone rearrangements take place. In HADDOCK [30], sidechains of both receptor and ligand are treated as being flexible during the first MD simulated annealing refinement stage. Several other methods describe sidechain flexibility based on rotamer libraries. ATTRACT [31,32] uses a reduced protein representation together with multiple sidechain copies: switching between rotamers is performed at various stages during docking and the best conformation is selected for the next few subsequent EM steps. In RosettaDock [33], after the initial low-resolution search, the sidechains are repacked and further optimized in an MC search that includes rigid-body displacements. Recent developments include the sampling of off-rotamer sidechain conformations [34•] and a new solvated rotamer library [35]. Rather impressive results on the accuracy of sidechain positioning were obtained for those CAPRI targets that exhibit only minor backbone conformational changes upon binding [36,37].

Backbone flexibility

Dealing with backbone flexibility in protein–protein docking is still an open challenge. For protein–peptide docking, an MC-based flexible approach was reported that explicitly samples protein sidechain and ligand backbone and sidechain rotations [38]. Of all the approaches represented in CAPRI (see Table III in [1•]), only one, HADDOCK [30], explicitly allows backbone flexibility during an MD simulated annealing refinement stage in torsion angle space. Guided docking [39], which allows some degree of backbone rearrangement, was also

Figure 2

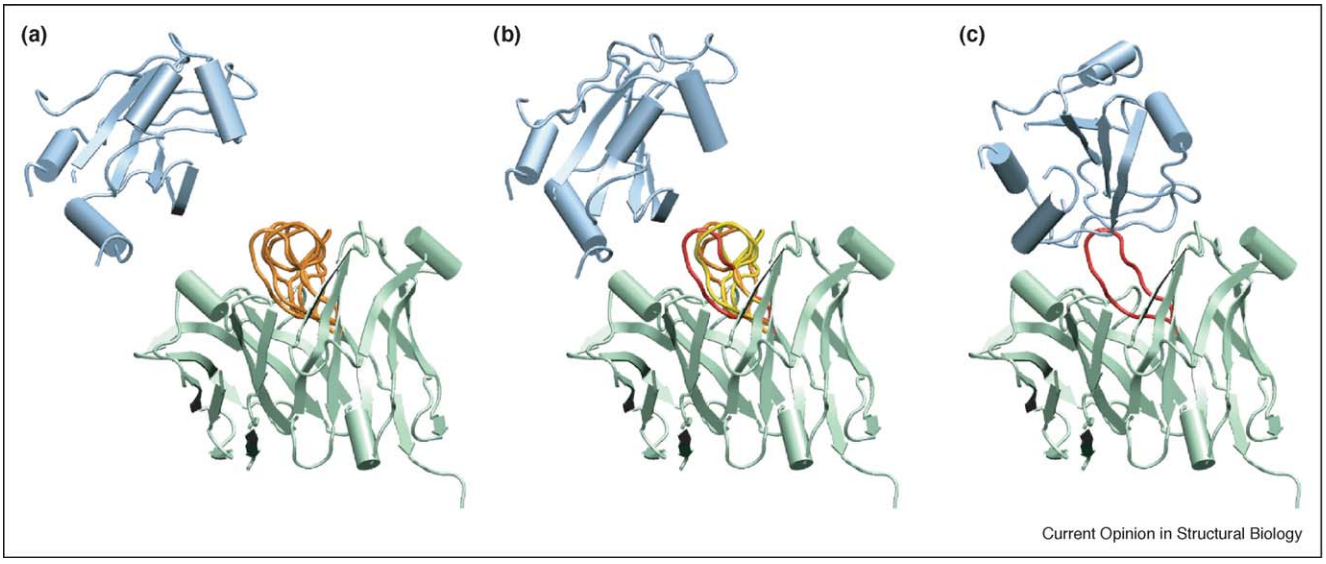


Illustration of the treatment of loop flexibility in the docking of the IgE Fv Spe7–thioredoxin complex (PDB code 1OAZ) using the mean-field multicopy approach [46•]. Cα traces of the receptor and ligand are indicated in green and blue, respectively. **(a)** At the start, each loop copy is assigned a weight [between 1 (red) and 0 (white)], according to its interaction energy with the ligand. **(b)** During docking, the most favorable loop conformation obtains a high weight and dominates the minimization. **(c)** At the end, the most favorable loop conformation has a weight close to 1.

applied to a few targets [22]. For the successful predictions obtained with HADDOCK, explicit inclusion of backbone and sidechain flexibility increased the number of acceptable solutions and improved the scoring of the resulting complexes [24•]. The effect on backbone RMSD, however, was less clear, improving the solutions by up to approximately 1.5 Å when the unbound forms were structurally relatively distant, but leading to some deterioration in quality when the differences between unbound and bound forms were minimal. The major impact on the docking results was to increase the fraction of native contacts by 10–35% [24•]. Similar conclusions were drawn from an MD study of the association of barnase and barstar from short distances [40].

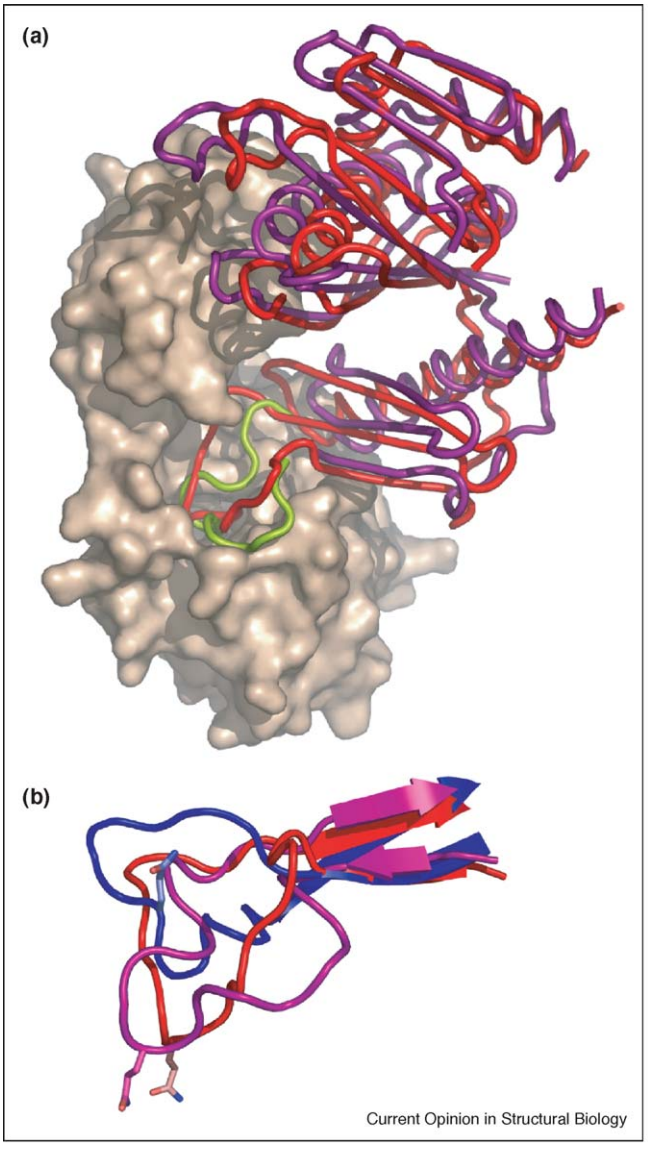
Describing large conformational changes in docking

Several promising approaches have recently been reported that allow the sampling of large conformational changes, either directly during docking or as a refinement step.

The concept of using collective motions for the analysis and prediction of motions has now been incorporated directly into the modeling process. Deformations along principal components are treated as additional degrees of freedom in ATTRACT, allowing the structures to deform along soft harmonic modes to facilitate the binding process [32]. This 'normal-modes-based' docking approach was shown to improve the results of both protein–ligand [32] and protein–protein docking [41••] (Figure 1). Large deformations can, however, lead to sidechain distortions and thus there is need for a sidechain rebuilding step. Normal modes have also been applied to the optimization of complexes against electron density maps [42,43] and the refinement of protein–DNA models [44]. Such approaches should lead to improved docking results, provided that the identified modes are relevant to the binding process; this is again related to our ability to predict motions.

Large loop rearrangements are also difficult to model and can thus hinder docking, even when some degree of backbone flexibility is introduced. A multicopy, mean-field approach has been proposed to deal with this problem [45,46•]. Multiple loop conformations generated by systematic conformational sampling are considered; during docking, the statistical weight of each copy is adapted according to the Boltzmann criterion and the receptor experiences the mean field of the ensemble of conformations (Figure 2). However, this method can lead to a combinatorial explosion in the number of interactions to be calculated if too many flexible loops need to be considered in both the ligand and the receptor. Another, rather stringent, way of dealing with flexible loops is to simply ignore them during docking and rebuild them in the context of the complex in a loop modeling step [47]. For the HemK–RF1 complex (CAPRI target 20) [48], the only acceptable solution was obtained using this strategy. In this case, it was clear that a loop must undergo significant conformational changes to expose a glutamine sidechain for methylation by HemK. Using HADDOCK, we performed a two-stage docking: first, based on biological information, a fully flexible loop segment was

Figure 3



Result of multibody (N = 3) flexible docking of the HemK–RF1 complex (PDB code 2B3T; CAPRI target 20 [48]) using HADDOCK [30]. **(a)** Best solution (i-RMSD 1.9 Å) superimposed onto the crystal structure. HemK is shown in surface representation. The backbone of RF1 in the crystal structure is shown in red. Three-body docking was performed with RF1 cut into a flexible 16-residue loop segment (green) and a core region (magenta). **(b)** Comparison of the loop conformations of RF1 in the starting homology model (blue) (loop RMSD from the bound conformation 9.6 Å), after multibody flexible docking (purple) (RMSD 4.7 Å) and in the crystal structure (red). The glutamine sidechain that becomes methylated by HemK is rendered in sticks.

Figure 4

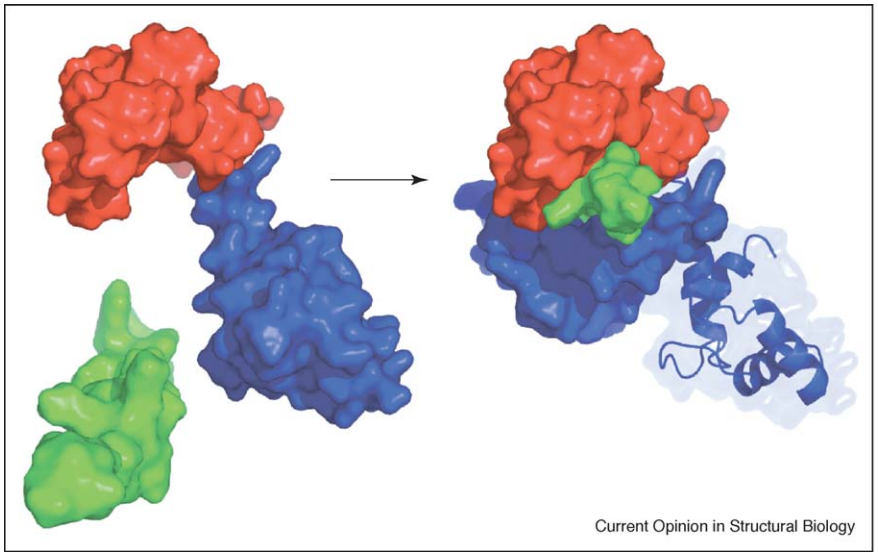


Illustration of hinge-based flexible docking of calmodulin to the myosin kinase peptide (PDB code 2BBM). The extended unbound conformation of calmodulin is represented on the left and the best solution obtained using FlexDock [49•] (2.1 Å RMSD from crystal structure) is on the right. The N- and C-terminal domains of calmodulin, which were treated as separate domains for docking, are indicated in red and blue, respectively, and the peptide is in green. To highlight the large domain movements that take place upon binding, the unbound orientation of the C-terminal domain is shown in semi-transparent blue in the docked model.

docked by MD; the resulting loop conformation was then grafted onto the RF1 core, followed by docking using the standard HADDOCK protocol (rigid-body docking followed by semi-flexible refinement), resulting in a solution with an interface RMSD (i-RMSD) of 4.2 Å and 30% native contacts. Recently, we performed a multibody (N = 3: HemK, RF1 core and RF1 loop) rigid-body docking followed by flexible refinement, leading to solutions with less than 2 Å i-RMSD and more than 50% of the native contacts (Figure 3). Although further analysis showed that acceptable solutions could be recovered at the rigid-body docking stage, flexible refinement improved the accuracy significantly (from 3.5 Å to 1.9 Å i-RMSD for our top-ranked solution).

Dissecting proteins into subdomains can thus offer a way to deal with large conformational changes in docking. This has been implemented in FlexDock [49•]: once hinge regions have been identified (or predicted), a fast two-body rigid docking is performed with the various combinations of subdomains. The resulting fragments are then assembled using a graph theory algorithm. This very promising approach was able to model the very large conformational change that occurs upon the binding of calmodulin to a target peptide (Figure 4) and produced the only acceptable solution for the LicT dimer (CAPRI target 9). A recent review by Wolfson *et al.* [50••] nicely illustrates the combination of flexibility prediction with this new 'hinge-bending' rigid-body docking approach. A similar, incremental, multibody, multistage docking strategy was successfully applied to several CAPRI targets using MolFit [51].

Conclusions and perspectives

The proper treatment of flexibility in protein–protein docking is clearly an active field of research, as evident from recent developments. It is, however, still in its infancy and remains a challenging problem, in particular when it comes to describing (large) conformational changes involving backbone atoms. Most probably, there will not be a unique solution; rather, it will be the proper combination of approaches for representing conformational changes and flexibility at several levels that will lead to success. This will be tightly coupled to our ability to properly identify flexible and hinge regions, and predict conformational changes. In addition, parallel developments in scoring schemes will be required, because the inclusion of flexibility in docking, although improving the success rate and the fraction of correctly predicted intermolecular contacts, can complicate the identification of the correct solutions. Finally, an even more challenging problem is awaiting us — the modeling of complexes in which the folding of one of the partners is coupled to binding. This will require us to address protein folding and structure prediction in the context of docking.

Acknowledgements

The kind contributions of Martin Zacharias (International University of Bremen) for Figure 1, Karine Bastard and Chantal Prévost (Institut de Biologie Physico-Chimique, Paris) for Figure 2, and Dina Schneidman-Duhovny, Yuval Inbar, Haim Wolfson and Ruth Nussinov (Tel Aviv

University) for Figure 4 are acknowledged. I also thank Victor Hsu (Oregon State University) for careful reading of this manuscript.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. • Méndez R, Leplae R, Lensink MF, Wodak SJ: **Assessment of CAPRI predictions in rounds 3-5 shows progress in docking procedures**. *Proteins* 2005, **60**:150-169.

This paper gives an overview of the current state of the art in protein–protein docking, as assessed from recent rounds of the blind docking experiment CAPRI. The characteristics (see in particular Table III) and performance of the various docking approaches are discussed, together with aspects of flexibility treatment.

2. Brooijmans N, Kuntz ID: **Molecular recognition and docking algorithms**. *Annu Rev Biophys Biomol Struct* 2003, **32**:335-373.

3. Erickson JA, Jalaie M, Robertson DH, Lewis RA, Vieth M: **Lessons in molecular recognition: the effects of ligand and protein flexibility on molecular docking accuracy**. *J Med Chem* 2004, **47**:45-55.

4. O'Hearn SD, Kusalik AJ, Angel JF: **MolCom: a method to compare protein molecules based on 3-D structural and chemical similarity**. *Protein Eng* 2003, **16**:169-178.

5. Jacobs DJ, Rader AJ, Kuhn LA, Thorpe MF: **Protein flexibility predictions using graph theory**. *Proteins* 2001, **44**:150-165.

6. Tirion MM: **Large amplitude elastic motions in proteins from a single-parameter, atomic analysis**. *Phys Rev Lett* 1996, **77**:1905-1908.

7. de Groot BL, van Aalten DM, Scheek RM, Amadei A, Vriend G, Berendsen HJ: **Prediction of protein conformational freedom from distance constraints**. *Proteins* 1997, **29**:240-251.

8. Alexandrov V, Lehnert U, Echols N, Milburn D, Engelman D, Gerstein M: **Normal modes for predicting protein motions: a comprehensive database assessment and associated Web tool**. *Protein Sci* 2005, **14**:633-643.

9. Barrett CP, Hall BA, Noble MEM: **Dynamite: a simple way to gain insight into protein motions**. *Acta Crystallogr D Biol Crystallogr* 2004, **60**:2280-2287.

10. Tai K: **Conformational sampling for the impatient**. *Biophys Chem* 2004, **107**:213-220.

11. Lei M, Zavodszky MI, Kuhn LA, Thorpe MF: **Sampling protein conformations and pathways**. *J Comput Chem* 2004, **25**:1133-1148.

12. Suhre K, Sanejouand Y-H: **ElNémo: a normal mode web server for protein movement analysis and the generation of templates for molecular replacement**. *Nucleic Acids Res* 2004, **32**:W610-W614.

13. Barrett CP, Noble MEM: **Dynamite extended: two new services to simplify protein dynamic analysis**. *Bioinformatics* 2005, **21**:3174-3175.

14. Palma PN, Krippahl L, Wampler JE, Moura JJG: **BiGGER: a new (soft) docking algorithm for predicting protein interactions**. *Proteins* 2000, **39**:372-384.

15. Heifetz A, Eisenstein M: **Effect of local shape modifications of molecular surfaces on rigid-body protein-protein docking**. *Protein Eng* 2003, **16**:179-185.

16. Ma B, Elkayam T, Wolfson H, Nussinov R: **Protein-protein interactions: structurally conserved residues distinguish between binding sites and exposed protein surfaces**. *Proc Natl Acad Sci USA* 2003, **100**:5772-5777.

17. • Smith GR, Sternberg MJE, Bates PA: **The relationship between the flexibility of proteins and their conformational states on forming protein-protein complexes with an application to protein-protein docking**. *J Mol Biol* 2005, **347**:1077-1101.

MD simulations in water of a set of 41 proteins are used to analyze the impact of flexibility on docking using 3D-Dock. Various approaches are compared: grid smoothing and ensemble docking from conformations obtained by clustering or principal components analysis. Grid smoothing does not lead to improvements, whereas ensemble docking is shown to enhance the results. An interesting observation is that interface core residues show a tendency to be less mobile than the remainder of the surface.

18. • Rajamani D, Thiel S, Vajda S, Camacho CJ: **Anchor residues in protein-protein interactions**. *Proc Natl Acad Sci USA* 2004, **101**:11287-11292.

MD simulations in water of a set of 11 proteins in their unbound form reveal that the sidechains of residues critical to complex formation, the so-called hot-spot or anchor residues, sample bound conformations during the simulations. These usually correspond to surface sidechains that bury the largest surface area upon complex formation.

19. Li X, Keskin O, Ma B, Nussinov R, Liang J: **Protein-protein interactions: hot spots and structurally conserved residues often locate in complemented pockets that pre-organized in the unbound states: implications for docking**. *J Mol Biol* 2004, **344**:781-795.

20. Dominguez C, Bonvin AMJJ, Winkler GS, van Schaik FMA, Timmers HTM, Boelens R: **Structural model of the UbcH5B/CNOT4 complex revealed by combining NMR, mutagenesis, and docking approaches**. *Structure* 2004, **12**:633-644.

21. • Grünberg R, Leckner J, Nilges M: **Complementarity of structure ensembles in protein-protein binding**. *Structure* 2004, **12**:2125-2136.

Ensembles of structures obtained from MD simulations with restraints along the principal components of motion are used for cross docking with HEX. The results for a set of 17 complexes indicate improved performance in terms of the number of near-native solutions, but a more difficult scoring; various combinations of starting structures can lead to high scoring solutions. The recognition process is discussed in terms of conformation selection.

22. Smith GR, Fitzjohn PW, Page CS, Bates PA: **Incorporation of flexibility into rigid-body docking: applications in rounds 3-5 of CAPRI**. *Proteins* 2005, **60**:263-268.

23. Camacho CJ: **Modeling side-chains using molecular dynamics improve recognition of binding region in CAPRI targets**. *Proteins* 2005, **60**:245-251.

24. • van Dijk ADJ, de Vries SJ, Dominguez C, Chen H, Zhou H-X, Bonvin AMJJ: **Data-driven docking: HADDOCK's adventures in CAPRI**. *Proteins* 2005, **60**:232-238.

Successful CAPRI predictions were obtained using a multilevel representation of flexibility: ensemble docking from conformations obtained by short MD simulations is combined with subsequent explicit sidechain and backbone flexible refinement by torsion angle MD. Flexibility is shown to have the most impact on the fraction of native contacts.

25. • Zavodszky MI, Lei M, Thorpe MF, Day AR, Kuhn LA: **Modeling correlated main-chain motions in proteins for flexible molecular recognition**. *Proteins* 2004, **57**:243-261.

The flexibility of proteins is analyzed using the geometry-based approach FIRST, combined with conformational sampling with ROCK. The resulting conformations are subsequently used in ensemble docking. This approach is applied to the docking of the cyclophilin A–cyclosporin complex.

26. Amadei A, Linssen ABM, Berendsen HJC: **Essential dynamics of proteins**. *Proteins* 1994, **17**:412-425.

27. Mustard D, Ritchie DW: **Docking essential dynamics eigenstructures**. *Proteins* 2005, **60**:269-274.

28. Li L, Weng Z: **RDOCK: refinement of rigid-body protein docking predictions**. *Proteins* 2003, **53**:693-707.

29. Fernández-Recio J, Totrov M, Abagyan R: **ICM-DISCO docking by global energy optimization with fully flexible side-chains**. *Proteins* 2003, **52**:113-117.

30. Dominguez C, Boelens R, Bonvin AMJJ: **HADDOCK: a protein-protein docking approach based on biochemical or biophysical information**. *J Am Chem Soc* 2003, **125**:1731-1737.

31. Zacharias M: **Protein-protein docking with a reduced protein model accounting for side-chain flexibility**. *Protein Sci* 2003, **12**:1271-1282.

32. Zacharias M: **Rapid protein-ligand docking using soft modes from molecular dynamics simulations to account for protein deformability: binding of FK506 to FKBP**. *Proteins* 2004, **54**:759-767.

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

34. • Wang C, Schueler-Furman O, Baker D: **Improved side-chain modeling for protein-protein docking**. *Protein Sci* 2005, **14**:1328-1339.
RosettaDock's flexible sidechain modeling based on backbone-dependent rotamer libraries is extended by off-rotamer optimization using short EM steps during the MC search. This is shown to improve both accuracy and discrimination between native and non-native solutions.

35. Jiang L, Kuhlman B, Kortemme TA, Baker D: **A ''solvated rotamer'' approach to modeling water-mediated hydrogen bonds at protein-protein interfaces**. *Proteins* 2005, **58**:893-904.

36. Schueler-Furman O, Wang C, Baker D: **Progress in protein-protein docking: atomic resolution predictions in the CAPRI experiment using RosettaDock with an improved treatment of side-chain flexibility**. *Proteins* 2005, **60**:187-194.

37. Daily MD, Masica D, Sivasubramanian A, Somarouthu S, Gray JJ: **CAPRI rounds 3-5 reveal promising successes and future challenges for RosettaDock**. *Proteins* 2005, **60**:181-186.

38. Liu Z, Dominy BN, Shakhnovich EI: **Structural mining: self-consistent design on flexible protein-peptide docking and transferable binding affinity potential**. *J Am Chem Soc* 2004, **126**:8515-8528.

39. Fitzjohn PW, Bates PA: **Guided docking: first step to locate potential binding sites**. *Proteins* 2003, **52**:28-32.

40. Ehrlich LP, Nilges M, Wade RC: **The impact of protein flexibility on protein-protein docking**. *Proteins* 2005, **58**:126-133.

41. •• May A, Zacharias M: **Accounting for global protein deformability during protein-protein and protein-ligand docking**. *Biochim Biophys Acta* 2005, **1754**:225-231.
This review addresses aspects of protein deformability in docking. It demonstrates for both protein–ligand and protein–protein docking how displacements along pre-calculated collective degrees of motion can be introduced as additional degrees of freedom during docking.

42. Tama F, Miyashita O, Brooks CL III: **Flexible multi-scale fitting of atomic structures into low-resolution electron density maps with elastic network mode analysis**. *J Mol Biol* 2004, **337**:985-999.

43. Hinsen K, Reuter N, Navaza J, Stokes DL, Lacapère J-J: **Normal mode-based fitting of atomic structure into electron density maps: application to sarcoplasmic reticulum Ca-ATPase**. *Biophys J* 2005, **88**:818-827.

44. Lindahl E, Delarue M: **Refinement of docked protein-ligand and protein-DNA structures using low frequency normal mode amplitude optimization**. *Nucleic Acids Res* 2005, **33**:4496-4506.

45. Bastard K, Thureau A, Lavery R, Prévost C: **Docking macromolecules with flexible segments**. *J Comput Chem* 2003, **24**:1910-1920.

46. • Bastard K, Prévost C, Zacharias M: **Accounting for loop flexibility during protein-protein docking**. *Proteins* 2005, in press.
The problem of loop flexibility in docking is addressed using a multicopy mean-field approach implemented in ATTRACT. The weight of each loop copy is adapted during docking by rigid-body EM based on its interaction energy following a Boltzmann criterion.

47. Rohl CA, Strauss CEM, Chivian D, Baker D: **Modeling structurally variable regions in homologous proteins with Rosetta**. *Proteins* 2004, **55**:656-677.

48. Graille M, Heurgué-Hamard V, Champ S, Mora L, Scrima N, Ulryck N, van Tilbeurgh H, Buckingham RH: **Molecular basis for bacterial class I release factor methylation by PrmC**. *Mol Cell* 2005, **20**:917-927.

49. • Schneidman-Duhovny D, Inbar Y, Nussinov R, Wolfson HJ: **Geometry-based flexible and symmetric protein docking**. *Proteins* 2005, **60**:224-231.
The challenge of dealing with large conformational changes in docking is addressed by dissecting proteins into subdomains after having identified hinge regions. Fast pairwise docking of subdomains is then performed by surface patch matching using PatchDock and the resulting solutions are combined to reconstruct the complete complex based on a graph theory algorithm. This approach, implemented in FlexDock, is extremely fast and was able to reproduce the large conformational changes that occur in calmodulin upon peptide binding.

50. •• Wolfson HJ, Shatsky M, Schneidman-Duhovny D, Dror O, Shulman-Peleg A, Ma B, Nussinov R: **From structure to function: methods and applications**. *Curr Protein Pept Sci* 2005, **6**:171-183.
This review discusses, among others, methods to identify flexible and hinge regions in proteins, and their application to hinge-based flexible docking. The described approaches are applied to the modeling of the s-Src protein kinase complexed with a proline-rich peptide.

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