Published in final edited form as:
Proteins. 2010 November 15; 78(15): 3197–3204. doi:10.1002/prot.22790.

An Integrated Suite of Fast Docking Algorithms

Efrat Mashiach[1], **Dina Schneidman-Duhovny**[1,2], **Aviyah Peri**[1], **Yoli Shavit**[1], **Ruth Nussinov**[3,4,1], and **Haim J. Wolfson**[1,2]

[1]Blavatnik School of Computer Science Raymond and Beverly Sackler Faculty of Exact Sciences Tel Aviv University, Tel Aviv 69978, Israel

[3]Basic Research Program, SAIC-Frederick, Inc. Center for Cancer Research Nanobiology Program NCI - Frederick Frederick, MD 21702, USA

[4]Department of Human Genetics and Molecular Medicine Sackler Faculty of Medicine Tel Aviv University, Tel Aviv 69978, Israel

Abstract

The CAPRI experiment (Critical Assessment of Predicted Interactions) simulates realistic and diverse docking challenges, each case having specific properties that may be exploited by docking algorithms. Motivated by the different CAPRI challenges, we developed and implemented a comprehensive suite of docking algorithms. These were incorporated into a dynamic docking protocol, consisting of four main stages: (1) Biological and bioinformatics research aiming to predict the binding site residues, to define distance constraints between interface atoms and to analyze the flexibility of molecules; (2) Rigid or flexible docking, performed by the PatchDock or FlexDock method, which utilizes the information gathered in the previous step. Symmetric complexes are predicted by the SymmDock method; (3) Flexible refinement and re-ranking of the rigid docking solution candidates, performed by FiberDock; and finally, (4) clustering and filtering the results based on energy funnels. We analyzed the performance of our docking protocol on a large benchmark and on recent CAPRI targets. The analysis has demonstrated the importance of biological information gathering prior to docking, which significantly increased the docking success rate, and of the refinement and re-scoring stage that significantly improved the ranking of the rigid docking solutions. Our failures were mostly a result of mishandling backbone flexibility, inaccurate homology modeling, or incorrect biological assumptions. Most of the methods are available at http://bioinfo3d.cs.tau.ac.il/.

INTRODUCTION

Protein-protein interactions play a major role in cellular function. Thus, revealing the three-dimensional structure of a protein-protein complex can help understand how the complex functions in the cell, and guide the design of drugs that can either prevent the formation of the complex or increase its stability. Computational docking methods aim to predict the atomic resolution three-dimensional (3D) structure of a complex, given the coordinates of the unbound conformations of the molecules from which it is assembled. The CAPRI experiment simulates realistic docking challenges and reveals strengths and weaknesses of current docking methods[1-4]. Motivated by the different CAPRI challenges, we have

[2]Address correspondence to School of Computer Science, Tel Aviv University, Tel Aviv 69978, Israel. TEL: +972-3-640-8268, FAX: +972-3-640-6476, wolfson@tau.ac.il. .
[2]Current address: Department of Bioengineering and Therapeutic Sciences, Department of Pharmaceutical Chemistry, University of California, San Francisco San Francisco, CA 94158, USA

developed over the years a comprehensive suite of docking algorithms. These algorithms were incorporated into a dynamic docking protocol that can be used for predicting the structure of many different types of molecular complexes. In this study, we examined our docking protocol performance on CAPRI targets in rounds 13-19 and on a large docking benchmark. The analysis demonstrated the importance of collecting reliable biological information on the binding site residues and contacts, the significant contribution of flexible refinement methods and the significant difficulty that still exists in docking homology models and flexible proteins.

METHODS

Over the years our group has developed a set of efficient and practical docking algorithms. These methods were integrated into a comprehensive docking suite that can be used for predicting many types of molecular complexes with different properties (hinge motion, flexible loops, symmetric interactions, etc.) and restraints (biological information about the binding site location, distance constraints, etc.). Our docking protocol consists of four main stages, detailed below.

1. Biological and bioinformatics research of the interacting proteins

The goal of this preliminary stage is to define restraints that will reduce the search space of the docking. The methods we use in the following stage can receive as an input potential binding site residues and pair-wise atomic distance constraints. Reduction of the search space can be achieved by analysis of the biological function of the interaction. For example, if one of the proteins performs a modification in a certain site of the interacting molecule, then a distance constraint can be defined between the active site and the modified site. Information about the binding site can be obtained from different sources, such as mutations that decrease the binding affinity, sites that are known not to be in the interface (eg. functional sites that are active during the interaction) and bioinformatics conservation analysis[5], which predicts conserved surface patches that often imply the binding site location. Multiple sequence and structure alignment (e.g. BLAST[6], MultiProt[7]) often help in finding homologous proteins that might form similar interactions.

Flexibility analysis determines the docking strategy. Hinges can be indentified using HingeProt[8] and flexible loops can be recognized by B-factors, NMR and structural comparison of different X-ray structures of the same or homologous proteins. Search space reduction dramatically increases the docking success rate. However, if incorrect information is used, or the results of the research are misinterpreted, the docking is highly likely to fail. Therefore, only high confidence information should be used.

2. Rigid or hinge bent flexible docking

Rigid docking is performed by the PatchDock method[9,10], which is an efficient, geometry-based technique. The method can explicitly reduce the search space based on the data collected in the previous step, eg. interface and non-interface residues and distance constraints. Hinge movement is handled by the FlexDock algorithm[11] and if the target complex is a symmetric multimer, the SymmDock method[9,12] is used.

Our rigid and flexible docking methods are very efficient and are probably the fastest techniques available. Thousands of candidate complexes can be generated within minutes on a standard PC computer. To deal with unbound real-life cases a small amount of steric clashes is allowed at the interface. In order to rank the candidate complexes with energy based scoring function, the interface is optimized and the clashes are removed by a refinement method.

3. Flexible refinement and re-ranking

In this stage the top 1000-5000 solutions from the rigid/flexible docking stage are refined and re-scored. Here we use the FiberDock method[13] that optimizes the side-chain conformations in the interface, models backbone movements, minimizes the relative rigid-body orientations of the molecules, and ranks the refined solutions by a binding-energy scoring function. The FiberDock method is based on our previously developed FireDock method[14,15] that also models side-chain movements, but keeps the backbone rigid. In the CAPRI experiment, we have used FireDock up to target 40, from which we started to use FiberDock in the refinement stage. The refinement process results in a drastic improvement in the ranking and accuracy of the predicted models (see the RESULTS section).

For refining models of symmetric complexes, we developed a new refinement method called SymmRef. Similarly to FireDock[15], this method optimizes the side-chain conformations in the interfaces and minimizes the rigid-body orientations of the symmetric units. However, unlike FireDock, here the refinement preserves the symmetry of the complex.

4. Clustering and filtering

In the final stage we try to identify the near native solutions from the top 50-100 refined models. First, we cluster the top ranked models and leave only the lowest energy model from each cluster as a representative. Then, we search for energy-score funnels around the remaining candidate docking solutions. Here we use our recently developed ValiDock server (http://bioinfo3d.cs.tau.ac.il/ValiDock/) that randomly samples rigid-body perturbations around each candidate solution, refines each perturbation by FiberDock and draws a graph of energy-score vs. RMSD from the candidate solution. Energy funnels are known to be a reliable indicator of near-native docking solutions[16,17], although in many cases energy funnels are also found in false docking solutions.

CAPRI participation—In the CAPRI challenge we participated as predictors, using three different docking protocols: **(1) PatchDock webserver**[9] (http://bioinfo3d.cs.tau.ac.il/PatchDock/) – rigid docking without human intervention. In cases of symmetric docking we used the SymmDock webserver[9] instead (http://bioinfo3d.cs.tau.ac.il/SymmDock/) **(2) FireDock**[14]**/FiberDock webserver** (http://bioinfo3d.cs.tau.ac.il/FireDock/, http://bioinfo3d.cs.tau.ac.il/FiberDock/) – rigid docking by the PatchDock server followed by refinement of the top 1000 solutions by the FireDock/FiberDock server. (3) **Human prediction** – here we performed the full, four stage docking protocol described above. In addition, we participated in the scoring challenges[1], where we used FireDock/FiberDock to refine and re-rank the uploaded models.

RESULTS

In recent CAPRI rounds (13-19) we were among the best performing predictor teams by submitting near native models in 6 out of the 13 targets. In 4 targets, a near native model was also submitted by our automatic webservers (PatchDock, SymmDock, FireDock or FiberDock). In this section we briefly describe our successes and failures in these CAPRI rounds. For 6 of the targets, in which we failed, we identified failure causes, repeated the interactions prediction and in 3 of them we were able to obtain a near native model in the top 10 solutions (Table 1). Additionally, we tested the performance of our docking protocol in “blind” docking on protein-protein docking benchmark 3.0[18] cases with limited flexibility. Our analysis demonstrates the significant contribution of the refinement and re-ranking stage in the docking protocol.

Target 29: Trm8/Trm82 tRNA guanin-N(7)-methyltransferase

The docking challenge in target 29 was to predict the structure of Trm8 complexed with Trm82 tRNA guanin-N(7)-methyltransferase (PDB 2VDV and 2VDU). Trm82 was given in the bound conformation and Trm8 in the unbound conformation. In this challenge we failed to predict the structure of the complex but succeeded in the scoring stage, where we identified two acceptable solutions with FireDock. The input structure of Trm8 was truncated at the N-terminal (residue 73). In the rigid-docking stage of our protocol, we blocked the site of this residue, since we assumed that the missing N-terminal segment is occupying this area. This was a wrong assumption since in the native structure of the complex[19] residue 73 is interacting with Trm8. This assumption prevented PatchDock from generating any near native solution. In addition, Trm8 has a flexible loop in the interface that could have been easily identified by the fact that part of it was missing in the X-ray structure. We repeated the docking experiment and ran PatchDock without blocking any site and after removing the flexible loop from the Trm8 structure. The first acceptable solution of PatchDock was ranked 2123 (LRMSD of 4.88Å and IRMSD of 2.60Å). After refining and re-ranking the top 5000 PatchDock solutions by FiberDock the same solution was ranked in the 3rd place and its LRMSD decreased to 4.33Å and IRMSD to 1.90Å.

Target 30: Rnd1-GTP bound to RBD dimer

In this target the challenge was to predict the structure of the complex of Rnd1-GTP and RBD (PDB 2R20), given their unbound structure. Based on the literature we correctly identified residues in the interface for both proteins[20-22]. However, in the CAPRI prediction we used only the suspected binding site of Rnd1 and not of RBD and we did not get an acceptable solution in the top 10 models of FireDock. This was probably due to the flexible loop in the binding site of Rnd1 that decreased the shape complementarity of near native solutions. We repeated the experiment using the binding site information for both proteins and obtained an acceptable solution in the 5th place.

Target 32: Protease savinase bound to Bi-functional inhibitor BASI

In this challenge, the goal was to predict the structure of Protease savinase bound to its inhibitor[23] (PDB 3BX1) given their unbound structures. In the literature we found a description of a homologous complex of proteinase K and its inhibitor (the coordinates of the complex were not available)[24]. Based on this description, we identified the interface of Protease savinase and the active loop of the inhibitor. This information was given to PatchDock and the results were refined and re-ranked by FireDock. The active loop of the inhibitor was suspected to be flexible. Therefore, we re-modeled it in the top solutions of FireDock using ModLoop[25] and refined these solutions again by FireDock. This process resulted in two high accuracy, two medium and two acceptable models in our 10 submitted predictions.

Targets 33-34: Methyl transferase bound to its RNA substrate

Target 33 presented a challenge of modeling protein-RNA interaction. The unbound structure of a methyl-transferase protein (structure not yet published) was given, while for the RNA molecule only a homologous structure was available. The RNA molecule contains three hairpins. We predicted hinge-movements between the hairpins, and therefore tried to dock them separately. First, we docked the third hairpin with a distance constraint between the active site of the methyl-transferase and the methylated nucleotide (distance of up to 6Å). Then we docked the remaining two hairpins together to the top 10 results from the previous step. We were not able to produce a correct model for this target. An alignment between the modeled and the bound structure of the RNA molecule revealed a significant

difference in the conformation of the third hairpin. This structural difference prevented PatchDock from generating a near native model in the first step of the docking process.

In target 34 we were given the bound structure of the RNA molecule. Here we used the same distance constraint and obtained two acceptable solutions in the top 10 models of the PatchDock server. However, they were disqualified due to low sequence identity that resulted from the fact that we docked the homologue protein of the methyl-transferase and not its model. As human predictors we suspected that the active-site of the protein can slightly open in the direction of the first normal mode. Therefore, we created an "open" form of the protein and used it in addition to the original "closed" form in cross-docking with PatchDock and FireDock and got 4 acceptable solutions in the top 10.

Targets 35-36: Xylanase Xyn10B

In target 35[26] (PDB 2W5F) the goal was to predict the structure of a covalently linked molecule with two domains: Polysaccharide binding module CBM22 and the catalytic module GH10. The two domains had to be modeled by homologues. In target 36 the bound structure of CBM22 was given. Like the vast majority of the predictors, we failed to predict the structure of the two-domain protein in both targets, due to inaccurate homology modeling.

Target 37: G-protein Arf6 bound to Leucine zipper of JIP4

In this target[27] (PDB 2W83) we had to model the structure of a Leucine Zipper domain and dock it to the G-protein Arf6. In the biological analysis stage of our docking protocol, we identified a conserved surface patch on one of the proteins (Arf6), by using ConSurf[5]. We used this patch as the location of the interface. In addition, we blocked the location of the GTP binding site in Arf6. These constraints were used in the rigid docking stage. After refining and rescoring the solutions by FireDock we were able to get an acceptable solution among the 10 submitted models. In the scoring challenge we also used FireDock and got plausible results: one medium and two acceptable solutions. The submitted medium accuracy model and one of the acceptable models were disqualified due to the high number of clashes they had. These clashes could have been resolved by a second FireDock refinement.

Targets 38-39: Centaurin-α1 bound to FHA domain of KIF13B

The goal in target 38 was to predict the structure of the complex of Centaurin-α_1 and the FHA domain of KIF13B, using the unbound structure of Centaurin-α_1 and a homology model of the FHA domain. In target 39 the bound structure of the FHA domain was given. In the docking of both targets we relied on published biological experiments[28] that showed that the FHA domain of KIF13B binds to the GAP domain of Centaurin-α_1. According to the published complex structure[29] (PDB 3FM8), this information was incorrect, and hence we failed in both targets. We retried to dock the two targets without any binding site information. We used PatchDock and refined and re-ranked the top 5000 results by FiberDock. In target 38, we did not have a near native solution in the top 5000 results of PatchDock due to a flexible loop in the interface of KIF13B. However, in target 39 PatchDock produced a near-native solution, with LRMSD of 3Å and IRMSD of 2.18Å, which was ranked 3059. FiberDock brought this model to the 6th place with LRMSD of 1.71Å and IRMSD of 1.01Å.

Target 40: A complex of Trypsin and protease inhibitor

In target 40[30] the goal was to predict the structure of Bovine Trypsin bound to the protease inhibitor (PDB 3E8L). We were given the unbound structure of the Trypsin molecule and

the bound structure of the inhibitor. Our FiberDock server obtained one high accuracy solution and one acceptable solution in the top 10 models. During this round, after submitting the webserver results, information on the location of the active sites of the inhibitor was published. The active sites were located on two different loops. We used this information when running PatchDock and FireDock, and validated the solutions by the existence of binding energy funnels. In addition, we found two known structures of Trypsin with peptide inhibitors (PDB 1YF4 and 3BTG). Using these structures, we structurally aligned the active loops of our inhibitor with the peptides from these structures. The two approaches resulted in similar results of high accuracy. As human predictors we submitted three high accuracy and three acceptable models. In the scoring stage, we used our new FiberDock method and submitted five high accuracy, two medium accuracy and one acceptable models.

Target 41: Colicin E9 bound to Im2

In this target the goal was to predict the structure of Colicin E9 DNase domain in complex with the Im2 protein[31] (PDB 2WPT). Both proteins were given in their unbound structure. By running BLAST we indentified a high quality homologous complex of Colicin E9 DNase domain in complex with Im9 (PDB 1BXI). The PDB file of Im2 contained 60 NMR structures which we structurally superimposed[7] on the Im9 structure in the homologue structure. We refined these 60 models by the FiberDock webserver. The first solution was of high accuracy and the remaining nine were medium accuracy models. As human predictors we randomly sampled 100 perturbations, for each NMR structure, and refined these 6000 models by FiberDock. Now the top 10 submitted models contained seven medium accuracy and three acceptable models. This case shows that increasing the number of solution candidates is not always beneficial, since the scoring function is not accurate enough to reliably differentiate between high, medium and acceptable accuracy models.

Target 42: Designed TPR oligomer

In this target we had to predict the structure of a designed protein[32] (PDB 2WQH). We modeled the structure of the protein by a known homologous structure of TPR (PDB 1NA0). The sequence of the two proteins is identical, except for three amino acids in the TPR. Three Aspartates in the homologue were changed to Tyrosines. For the webserver submission, we ran the SymmDock server[9] and refined the results by our new SymmRef method, in order to resolve the clashes in the solutions. The submitted models included one high accuracy result. As human predictor we suspected that the mutations might be in the interface and stabilize the interaction. Hence, we repeated the SymmDock and SymmRef runs with the information that the mutated Tyrosines are in the interface. This assumption was correct, and in our top 10 submitted models we had one high accuracy model and one acceptable model.

Blind docking experiment

We tested our automated docking protocol (docking by PatchDock and refinement and re-ranking by FiberDock) on the 88 “rigid-body” cases (with IRMSD<1.5Å) from the protein-protein benchmark3.0[18]. The experiment showed the importance of the refinement and re-ranking stage of the protocol (Figure 1). For each case we ran PatchDock and refined and re-ranked the top 1000 solutions by FiberDock, which modeled backbone flexibility in the receptor and side-chain movements in both proteins. PatchDock produced a near native solution (LRMSD < 10Å or IRMSD<4Å) among the top 1000 results in 71.6% of the cases. However, after dividing the cases into three categories, Antibody-Antigen (AA), Enzyme Inhibitor (EI) and Others (O) we noticed that PatchDock produced acceptable results in the top 1000 solutions in 95.2% of the AA cases, 85.2% of the EI cases, and 50% of the other cases. In only 10.2% of the 88 cases PatchDock ranked an acceptable solution in the top 10 results. However, after refining, re-scoring and re-ranking the solutions by FiberDock, in

28.4% of the cases an acceptable solution was present in the top 10 results. When the top 50 models are considered, FiberDock achieves an acceptable solution in around 70% of the AA and EI cases. The superior results of the AA and EI cases stems from the fact that PatchDock and FiberDock were optimized to handle these types of complexes, and further improvement should be done for general protein-protein complexes.

DISCUSSION

Recent CAPRI challenges have emphasized two key issues. The first is the importance of gathering information on the specific target prior to docking, which significantly increases the success rate of the docking. The second point is our limited ability to handle docking of unbound molecules with significant backbone flexibility in the interacting area or docking of homology models, which often have inaccurate backbone conformation.

In recent CAPRI rounds (13-19) we succeeded in our docking predictions in 6 out of 13 targets. In all of these cases we used biological information and bioinformatics analysis. For targets 32 and 40 we found biological information on the interface of both interacting proteins, which drastically reduced the docking search space. In target 40, however, we also achieved a high accuracy result by our automatic FireDock server[14] without binding site information. In target 34 we used a distance constraint between the active site of the methyl transferase and the methylated nucleotide in the RNA molecule. In target 37 we detected a conserved surface patch on one of the proteins and used it as the location of the interface. On the other hand, in targets 29, 38 and 39 we used wrong biological assumptions that prevented us from predicting the correct structure of the complexes. When these wrong assumptions were eliminated we obtained correct docking solutions in two out of these three targets.

Docking targets with homology models or with significant backbone movements in the interacting molecules were the most difficult. In target 30 we indentified the correct interface prior to docking, but a flexible loop in the interface prevented us from docking the proteins correctly. Target 33 required homology modeling of an RNA molecule. Since the modeled conformation of one of its hairpins was inaccurate, the rigid-docking stage failed. In three other targets (35, 36, 38) we also failed due to inaccurate homology modeling.

Each CAPRI target reveals strengths and weaknesses of the methods we use, and guides us in developing new methods to face similar challenges more successfully. In the initial rounds of CAPRI, when our major tool was the PatchDock algorithm, we were quite successful in detecting acceptable solutions, but were less successful in detecting high and medium quality solutions[33 34]. This motivated the development of the FireDock refinement algorithm, which proved its efficacy in the currently reported rounds. As for now, the major obstacle in the docking field is to handle backbone conformational changes that occur in the interface. Motivated by this challenge, we recently developed the FiberDock method that mimics an induced-fit process and models backbone movements during the refinement of rigid docking solutions. FiberDock showed superior results over FireDock in cases with interface backbone flexibility[13], and we hope to see its contribution in the next CAPRI rounds. All our docking methods are rapid and efficient and most of them are publicly available as webservers on our website: http://bioinfo3d.cs.tau.ac.il/.

Acknowledgments

We thank Yuval Inbar, Alexandra Shulman-Peleg, Oranit Dror and Itamar Banitt for their help in the CAPRI challenges. EM is supported by the Adams Fellowship Program of the Israel Academy of Sciences and Humanities. The research of HJW was supported in part by the Israel Science Foundation [grant no. 1403/09]; and the Hermann Minkowski Minerva Geometry Center. This project has been funded in whole or in part with Federal funds from the

National Cancer Institute, National Institutes of Health, under contract number HHSN261200800001E. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government. This research was supported (in part) by the Intramural Research Program of the NIH, National Cancer Institute, Center for Cancer Research.

REFERENCES

1. Lensink MF, Mendez R, Wodak SJ. Docking and scoring protein complexes: CAPRI 3rd Edition. Proteins. 2007; 69(4):704–718. [PubMed: 17918726]
2. Janin J, Henrick K, Moult J, Eyck LT, Sternberg MJ, Vajda S, Vakser I, Wodak SJ. CAPRI: a Critical Assessment of PRedicted Interactions. Proteins. 2003; 52(1):2–9. [PubMed: 12784359]
3. Janin J. The targets of CAPRI rounds 3-5. Proteins. 2005; 60(2):170–175. [PubMed: 15981267]
4. Janin J. Sailing the route from Gaeta, Italy, to CAPRI. Proteins. 2005; 60(2):149. [PubMed: 15981247]
5. Glaser F, Pupko T, Paz I, Bell RE, Bechor-Shental D, Martz E, Ben-Tal N. ConSurf: identification of functional regions in proteins by surface-mapping of phylogenetic information. Bioinformatics. 2003; 19(1):163–164. [PubMed: 12499312]
6. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990; 215(3):403–410. [PubMed: 2231712]
7. Shatsky M, Nussinov R, Wolfson HJ. A method for simultaneous alignment of multiple protein structures. Proteins. 2004; 56(1):143–156. [PubMed: 15162494]
8. Emekli U, Schneidman-Duhovny D, Wolfson HJ, Nussinov R, Haliloglu T. HingeProt: automated prediction of hinges in protein structures. Proteins. 2008; 70(4):1219–1227. [PubMed: 17847101]
9. Schneidman-Duhovny D, Inbar Y, Nussinov R, Wolfson HJ. PatchDock and SymmDock: servers for rigid and symmetric docking. Nucleic Acids Res. 2005; 33(Web Server issue):W363–367. [PubMed: 15980490]
10. Duhovny, D.; Nussinov, R.; Wolfson, HJ. Efficient Unbound Docking of Rigid Molecules. In: Guigó, R.; Gusfield, D., editors. Lecture Notes in Computer Science. Springer Berlin / Heidelberg; Rome, Italy: 2002. p. 185-200.(Lecture Notes in Computer Science)
11. Schneidman-Duhovny D, Nussinov R, Wolfson HJ. Automatic prediction of protein interactions with large scale motion. Proteins. 2007; 69(4):764–773. [PubMed: 17886339]
12. Schneidman-Duhovny D, Inbar Y, Nussinov R, Wolfson HJ. Geometry-based flexible and symmetric protein docking. Proteins. 2005; 60(2):224–231. [PubMed: 15981269]
13. Mashiach E, Nussinov R, Wolfson HJ. FiberDock: Flexible induced-fit backbone refinement in molecular docking. Proteins. 2010; 78(6):1503–1519. [PubMed: 20077569]
14. Mashiach E, Schneidman-Duhovny D, Andrusier N, Nussinov R, Wolfson HJ. FireDock: a web server for fast interaction refinement in molecular docking. Nucleic Acids Res. 2008; 36(Web Server issue):W229–232. [PubMed: 18424796]
15. Andrusier N, Nussinov R, Wolfson HJ. FireDock: fast interaction refinement in molecular docking. Proteins. 2007; 69(1):139–159. [PubMed: 17598144]
16. 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(2):187–194. [PubMed: 15981249]
17. Schueler-Furman O, Wang C, Bradley P, Misura K, Baker D. Progress in modeling of protein structures and interactions. Science. 2005; 310(5748):638–642. [PubMed: 16254179]
18. Hwang H, Pierce B, Mintseris J, Janin J, Weng Z. Protein-protein docking benchmark version 3.0. Proteins. 2008; 73(3):705–709. [PubMed: 18491384]
19. Leulliot N, Chaillet M, Durand D, Ulryck N, Blondeau K, van Tilbeurgh H. Structure of the yeast tRNA m7G methylation complex. Structure. 2008; 16(1):52–61. [PubMed: 18184583]
20. Oinuma I, Katoh H, Harada A, Negishi M. Direct interaction of Rnd1 with Plexin-B1 regulates PDZ-RhoGEF-mediated Rho activation by Plexin-B1 and induces cell contraction in COS-7 cells. J Biol Chem. 2003; 278(28):25671–25677. [PubMed: 12730235]

21. Tong Y, Hughes D, Placanica L, Buck M. When monomers are preferred: a strategy for the identification and disruption of weakly oligomerized proteins. Structure. 2005; 13(1):7–15. [PubMed: 15642257]
22. Vikis HG, Li W, He Z, Guan KL. The semaphorin receptor plexin-B1 specifically interacts with active Rac in a ligand-dependent manner. Proc Natl Acad Sci U S A. 2000; 97(23):12457–12462. [PubMed: 11035813]
23. Micheelsen PO, Vevodova J, De Maria L, Ostergaard PR, Friis EP, Wilson K, Skjot M. Structural and mutational analyses of the interaction between the barley alpha-amylase/subtilisin inhibitor and the subtilisin savinase reveal a novel mode of inhibition. J Mol Biol. 2008; 380(4):681–690. [PubMed: 18556023]
24. Pal GP, Kavounis CA, Jany KD, Tsernoglou D. The three-dimensional structure of the complex of proteinase K with its naturally occurring protein inhibitor, PKI3. FEBS Lett. 1994; 341(2-3):167–170. [PubMed: 8137934]
25. Fiser A, Sali A. ModLoop: automated modeling of loops in protein structures. Bioinformatics. 2003; 19(18):2500–2501. [PubMed: 14668246]
26. Najmudin, S.; Pinheiro, BA.; Romao, MJ.; Prates, JAM.; Fontes, CMGA. Putting an N-Terminal End to the Clostridium Thermocellum Xylanase Xyn10B Story: Crystallographic Structure of the Cbm22-1-Gh10 Modules Complexed with Xylohexaose. to be published
27. Isabet T, Montagnac G, Regazzoni K, Raynal B, El Khadali F, England P, Franco M, Chavrier P, Houdusse A, Menetrey J. The structural basis of Arf effector specificity: the crystal structure of ARF6 in a complex with JIP4. EMBO J. 2009; 28(18):2835–2845. [PubMed: 19644450]
28. Venkateswarlu K, Hanada T, Chishti AH. Centaurin-alpha1 interacts directly with kinesin motor protein KIF13B. J Cell Sci. 2005; 118(Pt 11):2471–2484. [PubMed: 15923660]
29. Shen, L.; Tong, Y.; Tempel, W.; MacKenzie, F.; Arrowsmith, CH.; Edwards, AM.; Bountra, C.; Weigelt, J.; Bochkarev, A.; Park, H. Crystal structure of full length centaurin alpha-1 bound with the FHA domain of KIF13B. to be published
30. Bao R, Zhou CZ, Jiang C, Lin SX, Chi CW, Chen Y. The ternary structure of the double-headed arrowhead protease inhibitor API-A complexed with two trypsins reveals a novel reactive site conformation. J Biol Chem. 2009; 284(39):26676–26684. [PubMed: 19640842]
31. Meenan NAG, Sharma A, Fleishman SJ, MacDonald C, Morel B, Boetzel R, Moore GR, Baker D, Kleanthous C. The structural and energetic basis for high selectivity in a high affinity protein-protein interaction. Proc Natl Acad Sci U S A. 2010 in press.
32. Krachler AM, Sharma A, Regan L, Kleanthous C. Self-association of TPR-domains: natural and engineered. Proteins. 2010 in press.
33. Inbar Y, Schneidman-Duhovny D, Halperin I, Oron A, Nussinov R, Wolfson HJ. Approaching the CAPRI challenge with an efficient geometry-based docking. Proteins. 2005; 60(2):217–223. [PubMed: 15981251]
34. Mendez R, Leplae R, Lensink MF, Wodak SJ. Assessment of CAPRI predictions in rounds 3-5 shows progress in docking procedures. Proteins. 2005; 60(2):150–169. [PubMed: 15981261]

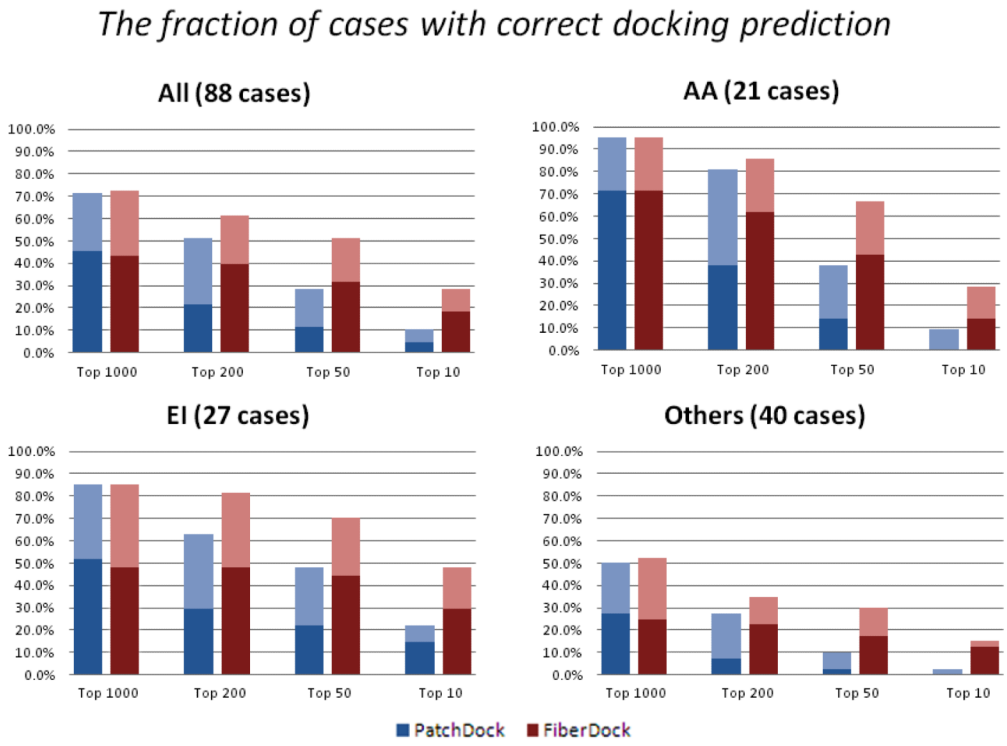

Figure 1.

Blind docking experiment. We tested our "blind" docking protocol on the 88 benchmark cases with IRMSD < 1.5Å. For each case we ran PatchDock and refined and re-ranked the top 1000 solutions by FiberDock. The graphs show the fraction of cases where a near native solution was present in the top 10, 50, 200 and 1000 solutions of PatchDock (shades of blue) and FiberDock (shades of red). The dark red/blue bars show the fraction of cases with medium accuracy solutions (LRMSD < 5Å or IRMSD < 2Å) and the light bars show the fraction of the additional cases with acceptable solutions (LRMSD < 10Å or IRMSD < 4Å). The top-left graph shows the results of all the 88 cases, the top-right graph shows the results of the Antibody-Antigen (AA) cases, the bottom left shows the Enzyme-Inhibitor (EI) cases, and the bottom right the other cases. This experiment demonstrates the importance of the refinement and re-ranking stage by FiberDock.

Table 1

CAPRI achievements. The table shows the number of near native solutions found in each CAPRI target by all the participating groups, by our automatic webservers and by our group. The solutions are divided into three accuracy levels: high (***), medium (**), and acceptable (*). Cases without any solution are marked by '-' and cases in which we did not submit any models are marked by 'X'.

Target	All groups			PatchDock/ SymmDock[a]			FireDock/FiberDock[a]			Human Prediction[b]			CAPRI retry[c]			Scoring[d]		
	***	**	*	***	**	*	***	**	*	***	**	*	***	**	*	***	**	*
T29	-	9	8	-	-	-	X	X	X	-	-	-	-	1	-	-	-	2
T30	-	-	2	-	-	-	X	X	X	-	-	-	-	-	1	-	-	-
T32	15	13	6	-	-	-	-	-	-	2	2	2	X	X	X	-	-	-
T33	-	-	-	-	-	-	-	-	-	-	-	-	X	X	X	-	-	-
T34	-	25	40	-	-	2 †	X	X	X	-	-	3† +1	X	X	X	X	X	X
T35	-	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
T36	-	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
T37	1	7	13	-	-	-	-	-	-	-	-	1	X	X	X	-	1 †	1+1 †
T38	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
T39	1	2	-	-	-	-	-	-	-	-	-	-	-	1	-	-	-	-
T40	79	54	31	-	-	-	1	-	1	3	-	3	X	X	X	5	2	1
T41	24	58	67	X	X	X	1	9	-	-	7	3	X	X	X	X	X	X
T42	9	5	6	1	-	-	X	X	X	1	-	1	X	X	X	X	X	X

[a] The achievements of our webservers: PatchDock/SymmDock (SymmDock was used only in target 42) and FireDock/FiberDock (FiberDock was first used in target 40).

[b] The achievements of our group, where we used the full docking protocol, described in the Methods section.

[c] We analyzed six cases in which we failed and retried to dock the targets. In three of these targets we had a near native result in the top 10 solutions (see the Results section for details).

[d] Our achievements in the Scoring challenge of CAPRI.

† - Accurate results that were disqualified due to clashes or bad sequence identity.