Nucleic Acids Research, 2005, Vol. 33, Web Server issue ***W363–W367***
doi:10.1093/nar/gki481

PatchDock and SymmDock: servers for rigid and symmetric docking

Dina Schneidman-Duhovny, Yuval Inbar, Ruth Nussinov[1,2] and Haim J. Wolfson*

School of Computer Science, Raymond and Beverly Sackler Faculty of Exact Sciences and [1]Sackler Institute of Molecular Medicine, Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv 69978, Israel and [2]Basic Research Program, SAIC-Frederick Inc., Laboratory of Experimental and Computational Biology NCI-Frederick, Building 469, Room 151, Frederick, MD 21702, USA

Received February 14, 2005; Revised March 22, 2005; Accepted April 18, 2005

ABSTRACT

Here, we describe two freely available web servers for molecular docking. The PatchDock method performs structure prediction of protein–protein and protein–small molecule complexes. The SymmDock method predicts the structure of a homomultimer with cyclic symmetry given the structure of the monomeric unit. The inputs to the servers are either protein PDB codes or uploaded protein structures. The services are available at http://bioinfo3d.cs.tau.ac.il. The methods behind the servers are very efficient, allowing large-scale docking experiments.

INTRODUCTION

Automated prediction of protein–protein interactions and protein–small molecule interactions is one of the most challenging problems in structural biology. Many biological studies, both in academia and in industry, may benefit from credible high-accuracy interaction predictions. In the docking problem the goal is to find the correct association of two interacting molecules given a structural representation for each molecule separately. In the case of protein–protein docking an accurate prediction will point out most of the residue–residue contacts involved in the target interaction.

Many docking algorithms (1–18) have been developed in the last three decades. However, only a few algorithms are currently available as a free web service [the pioneer was the ClusPro server (19)]. The algorithms mostly differ in the method for searching the six-dimensional transformation space that they apply, and in their evaluation of the resolved complexes. The majority of these algorithms are computationally too heavy for large-scale experiments. We have developed PatchDock (20), a very efficient algorithm for protein–small ligand and protein–protein docking. The algorithm was verified on enzyme–inhibitor and antibody–antigen complexes from benchmark 0.0 (21), where it successfully found near-native solutions for most of the cases. The algorithm was also successfully tested in the last three rounds (3–5) of the Critical Assessment of PRediction of Interactions (CAPRI) (22).

Motivated by the fact that many homomultimers are assembled in a cyclically symmetric arrangement, we have developed SymmDock (23), a docking algorithm for the prediction of cyclically symmetric homomultimers. Although other docking algorithms that predict special cases of symmetric interactions (D_2) of homomultimers have already been developed (24,25), SymmDock is the first one that a priori restricts its transformational search space only to symmetric transformations, and thus gains both in efficiency and performance.

Here, we describe the web servers that we have developed for docking using the PatchDock and SymmDock algorithms. We briefly discuss the methods and focus on user interface issues.

PATCHDOCK: MOLECULAR DOCKING ALGORITHM BASED ON SHAPE COMPLEMENTARITY PRINCIPLES

PatchDock (20) is a geometry-based molecular docking algorithm. It is aimed at finding docking transformations that yield good molecular shape complementarity. Such transformations, when applied, induce both wide interface areas and small amounts of steric clashes. A wide interface is ensured to include several matched local features of the docked molecules that have complementary characteristics. The PatchDock algorithm divides the Connolly dot surface representation (26,27) of the molecules into concave, convex and flat patches. Then, complementary patches are matched in order to generate candidate transformations. Each candidate transformation is further evaluated by a scoring function that considers both geometric fit and atomic desolvation

*To whom correspondence should be addressed. Tel/Fax: +972 3 640 6476; Email: wolfson@tau.ac.il
Correspondence may also be addressed to Dina Schneidman-Duhovny. Tel: +972 3 640 5395; Fax: +972 3 640 6476; Email: duhovka@tau.ac.il

energy (28). Finally, an RMSD (root mean square deviation) clustering is applied to the candidate solutions to discard redundant solutions. The main reason behind PatchDock's high efficiency is its fast transformational search, which is driven by local feature matching rather than brute force searching of the six-dimensional transformation space. It further speeds up the computational processing time by utilizing advanced data structures and spatial pattern detection techniques, such as geometric hashing and pose clustering, that were originally developed in the field of computer vision as detailed in Ref. (20). The run time of PatchDock for two input proteins of average size (about 300 amino acids) is <10 min on a single 1.0 GHz PC processor under the Linux operating system.

THE PATCHDOCK WEB SERVER: INPUT, OUTPUT AND USER INTERFACE

We have developed a very simple and intuitive web interface for the PatchDock algorithm available at http://bioinfo3d.cs.tau.ac.il/PatchDock/. The basic usage needs almost no explanation. The more advanced optional features are also very easy to learn. Once the docking request is submitted, the PatchDock algorithm starts the prediction process. The user is notified when the results are ready by an email message that contains a link to a web page where the predictions are presented. On this page the user can both view specific predictions and download a compressed file of the top scoring solutions (see Figure 1).

Input

The input is two molecules in PDB format. The molecules are either uploaded to the server or retrieved from the Protein Data Bank. In the second case the user has only to enter the PDB code. In order to dock a certain chain or chains, the user should specify the desired chain ID or IDs. The only other non-optional field in the docking request form is the user email for result notification.

In addition, there are four optional fields in the docking request form.

(i) *Clustering RMSD* — a positive number that specifies the radius of the RMSD clustering in angstroms. This value is used in the final clustering stage of the algorithm. It ensures that the distance between any two output solutions will be at least the specified clustering RMSD value. The default value for this parameter is 4 Å.

(ii) *Complex Type*. PatchDock has different sets of parameters, optimized for different types of complexes. If this field is not specified, the program will use a default configuration. In the case of the enzyme–inhibitor complex type, the algorithm restricts the search space to the cavities of the enzyme molecule. In the case of the antibody–antigen complex type, the algorithm automatically detects the complementarity-determining regions (CDRs) of the antibody and restricts the search to these regions (note: the antibody should be specified as 'receptor molecule'). In the case of protein–small ligand docking, the algorithm uses a parameter set optimized for small-size molecules.

(iii) and (iv) *Potential binding sites* of the receptor and the ligand, respectively. When biological data that point to specific residues belonging to a potential binding site is available, the user can incorporate such information in the algorithm. This information is supplied as an uploaded file that lists the residues of the potential binding site or sites. The file should be in the following format: in each line the residue index and its chain ID as in the PDB file of the receptor or the ligand, separated by a space. An example is shown on the server help page http://bioinfo3d.cs.tau.ac.il/PatchDock/help.html.

Output

A web page that presents the top 20 solutions is automatically generated. The user receives an email message with the URL of this page (a web link). The solutions are presented in a table, a row for each solution. The geometric score, the desolvation energy (28), the interface area size and the actual rigid transformation of the solution are shown. A link to a PDB file that presents the docking solution is also available in each line. The user may view or download it. There is also an option to view additional, lower ranking solutions by pressing the 'next 20 solutions' button at the lower right corner of the table. In the solutions page an option to download the top scoring solutions is available. The solutions are downloaded as a compressed file in ZIP format. This compressed file contains the PDB files of the top scoring solutions. The number of solutions is determined by the user, though we have limited it to 100. Our experience in the CAPRI (22) and other benchmarks (21) indicates that a near-native result is found among the top 100, and very often among the top 10 solutions.

SYMMDOCK: PREDICTION OF COMPLEXES WITH C_n SYMMETRY USING GEOMETRY-BASED DOCKING

SymmDock is a geometry-based docking algorithm for the prediction of a cyclically symmetric complex given the structure of its asymmetric unit. Like PatchDock it exploits local feature matching to produce the candidate set of transformations. Whereas PatchDock detects transformations with high shape complementarity, SymmDock restricts its search to symmetric cyclic transformations of a given order n. The algorithm takes advantage of the special characteristics of cyclically symmetric transformations in both its search and clustering methodologies. Details of the algorithm and some experimental results are given in Ref. (23).

THE SYMMDOCK WEB SERVER: INPUT, OUTPUT AND USER INTERFACE

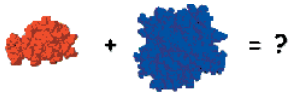
The user interface of SymmDock (http://bioinfo3d.cs.tau.ac.il/SymmDock/) is even simpler than that of PatchDock, since the input here consists of only one molecule and the symmetry order. For consistency and for the convenience of the user, the SymmDock web server has the same look-and-feel as PatchDock. The invocation and the receipt of results are also the same; namely, the result notification with the relevant web link is emailed to the user.

Input

Here, the input consists of two elements: the asymmetric unit (i.e. the monomer), and the symmetry order (2 for dimer, 3 for

A

PatchDock



Molecular Docking Algorithm Based on Shape Complementarity Principles

[About PatchDock] [Web Server] [Help] [FAQ]

NEW **SymmDock** - An Algorithm for Prediction of Complexes with C_n Symmetry

Type PDB codes of receptor and ligand molecules or upload files in PDB format

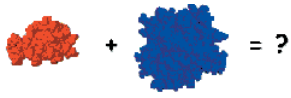
Receptor Molecule:	2kai:AB	(PDB:chainId e.g. 2kai:AB) **or** upload file: Browse...
Ligand Molecule:	2kai:I	(PDB:chainId e.g. 2kai:I) **or** upload file: Browse...
e-mail address:	inbaryuv@tau.ac.il	(the results are sent to this address)
Clustering RMSD:	4.0	
Complex Type:	Default	Be sure to give receptor and ligand in the corresponding order!
Receptor Binding Site:	Browse...	upload receptor binding site file (optional)
Ligand Binding Site:	Browse...	upload ligand binding site file (optional)

Submit Form Clear

Beta 1.2 Version
Contact: duhovka@tau.ac.il

B

PatchDock



Molecular Docking Algorithm Based on Shape Complementarity Principles

[About PatchDock] [Web Server] [Help] [FAQ]

Solution No	Score	Area	ACE	Transformation	PDB file of the complex
1	10442	1464.70	-32.68	2.77 -0.19 -1.83 51.01 38.72 13.49	result.1.pdb
2	10400	1241.90	130.16	0.18 -0.95 -2.10 65.02 15.14 34.30	result.2.pdb
3	10376	1254.00	-37.24	-1.59 0.94 -0.09 23.67 -3.83 -37.71	result.3.pdb
4	10370	1257.30	3.84	-1.50 -0.60 -1.28 24.90 31.86 22.86	result.4.pdb
5	10244	1370.80	63.53	0.85 0.32 -2.66 77.96 8.17 -5.26	result.5.pdb
6	10242	1348.60	59.74	-0.44 0.46 -0.78 13.10 24.19 -22.98	result.6.pdb
7	10136	1255.20	190.67	3.14 -0.45 2.18 58.67 -26.93 19.90	result.7.pdb
8	9898	1413.10	-184.06	3.14 0.33 -0.31 -1.47 1.05 -7.74	result.8.pdb
9	9886	1197.80	74.37	-2.17 -0.61 -0.68 5.57 14.37 18.90	result.9.pdb
10	9884	1407.00	269.66	2.06 0.53 -0.98 20.76 33.92 -19.37	result.10.pdb
11	9816	1406.60	208.16	1.77 -0.23 -0.56 6.71 20.21 21.24	result.11.pdb
12	9806	1187.80	-14.61	1.88 0.44 2.67 67.98 -23.44 -6.01	result.12.pdb
13	9726	1256.20	111.62	-0.78 -1.46 1.04 26.74 -11.76 40.37	result.13.pdb
14	9610	1114.50	-155.44	2.44 1.45 1.46 43.91 -17.57 -37.19	result.14.pdb
15	9574	1395.90	-300.28	-0.75 0.88 -2.72 59.34 2.02 -36.83	result.15.pdb
16	9522	1112.40	281.48	1.66 -0.82 -0.48 20.40 8.96 40.25	result.16.pdb
17	9408	1410.50	-26.93	1.86 0.09 1.47 34.75 -44.45 13.23	result.17.pdb
18	9294	1469.60	36.53	2.07 -0.23 -1.64 47.37 37.04 19.31	result.18.pdb
19	9274	1141.40	158.26	0.74 -0.16 -0.88 12.94 34.38 5.56	result.19.pdb
20	9190	1135.50	-73.73	-1.00 -0.37 1.08 14.81 -37.71 6.18	result.20.pdb
					show next 20 »»

DOWNLOAD best solutions as a ZIP file:

how many solutions to zip? 10 (any number from 2 to 100)

GO (this takes few seconds, please wait patiently)

Figure 1. The PatchDock user interface. (**A**) The request form of PatchDock. The receptor molecule and the ligand molecule are given either by the PDB code of the molecule (chain IDs are optional) or by uploading a file in PDB format. (**B**) The solutions page presents the geometric score, interface area size and desolvation energy of the 20 top scoring solutions. The user can use the 'show next 20' button to view solutions of lower score. The user can download each solution by pressing the solution link in the rightmost column or download an archive file (ZIP format) of the best solutions using the action button at the bottom of the page.

trimer etc.) as shown in Figure 2a. The asymmetric unit may be uploaded in PDB format or specified by its PDB ID. The symmetry order may be any integer ≥ 2. Notice that SymmDock predicts only cyclic symmetry. If the arrangement of the input monomers in its native complex follows a different type of symmetry, then SymmDock is not suitable for such a prediction. It should be noted, though, that a more sophisticated repetitive application of SymmDock can handle other

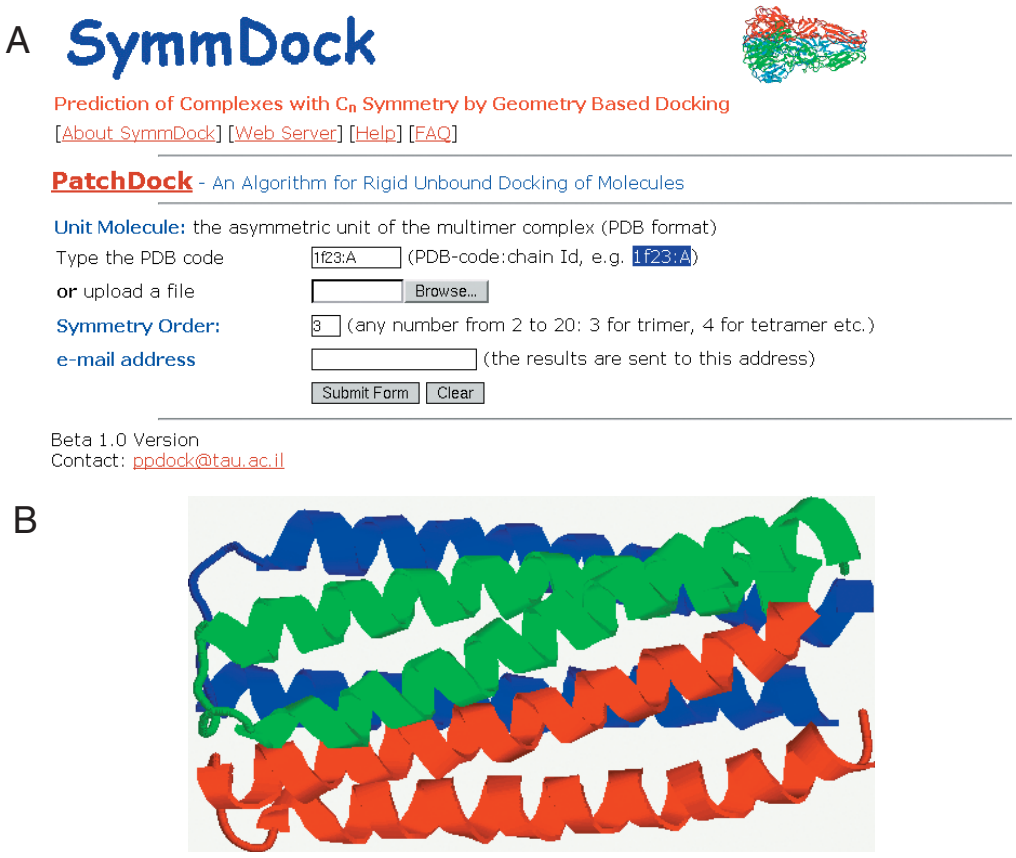


Figure 2. The SymmDock request page and an example solution. (**A**) The request form of SymmDock. The asymmetric unit (the monomer) is given either by the PDB code of the molecule (chain IDs are optional) or by uploading a file in PDB format. The order of the symmetry must also be filled in. Once the request is submitted the prediction process begins. When it is completed, an email message with a link to a solutions page is sent to the user. The solutions page follows the same format as shown for PatchDock in Figure 1b. (**B**) A cartoon representation of a solution PDB file that was predicted and generated by SymmDock for a C_3 example.

symmetries as well. Such an example is presented in Ref. (23), where we have reconstructed the SP1 dodecamer, which is a hexamer of dimers. The user's email address is also requested for notification purpose.

Output

Just as in PatchDock, a web page is generated to show the predicted solutions, and a link to that page is sent to the user by email. Here, instead of showing just pairwise interactions involved in the complex, the whole multimer is generated for each solution. For example, if the user request was to predict a symmetric complex of order of 12, then each solution is a dodecamer of the asymmetric unit. (An example of a predicted trimer is shown in Figure 2b.) Viewing or downloading just one solution, viewing lower ranked solutions and the option of downloading a set of solutions are just as described for PatchDock output.

ACKNOWLEDGEMENTS

We thank our Structural Bioinformatics Group and the system team of the Tel Aviv University Computer Science School. This research has been supported in part by the Center of Excellence in Geometric Computing and its Applications funded by the Israel Science Foundation (administered by the Israel Academy of Sciences). The research of H.J.W. is partially supported by the Hermann Minkowski-Minerva Center for Geometry at Tel Aviv University. The research of Y.I. has been supported by the Eshkol Fellowship funded by the Israeli Ministry of Science. The research of R.N. has been funded in whole or in part with Federal funds from the National Cancer Institute, National Institutes of Health, under contract number NO1-CO-12400. The content of this publication does not necessarily reflect the view 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. Funding to pay the Open Access publication charges for this article was provided by SAIC-Frederick, Inc.

Conflict of interest statement. None declared.

REFERENCES

1. Katchalski-Katzir,E., Shariv,I., Eisenstein,M., Friesem,A., Aflalo,C. and Vakser,I. (1992) Molecular surface recognition: determination of geometric fit between protein and their ligands by correlation techniques. *Proc. Natl Acad. Sci. USA*, **89**, 2195–2199.
2. Kuntz,I., Blaney,J., Oatley,S., Langridge,R. and Ferrin,T. (1982) A geometric approach to macromolecule–ligand interactions. *J. Mol. Biol.*, **161**, 269–288.
3. Connolly,M.L. (1986) Shape complementarity at the hemoglobin alpha 1 beta 1 subunit interface. *Biopolymers*, **25**, 1229–1247.

4. Jiang,F. and Kim,S. (1991) "Soft docking": matching of molecular surface cubes. *J. Mol. Biol.*, **219**, 79–102.
5. Walls,P. and Sternberg,M. (1992) New algorithms to model protein–protein recognition based on surface complementarity; applications to antibody–antigen docking. *J. Mol. Biol.*, **228**, 227–297.
6. Norel,R., Lin,S., Wolfson,H.J. and Nussinov,R. (1994) Shape complementarity at protein–protein interfaces. *Biopolymers*, **34**, 933–940.
7. Fischer,D., Lin,S., Wolfson,H.J. and Nussinov,R. (1995) A geometry-based suite of molecular docking processes. *J. Mol. Biol.*, **248**, 459–477.
8. Vakser,I. (1995) Protein docking for low-resolution structures. *Protein Eng.*, **8**, 371–377.
9. Vakser,I. (1996) Main-chain complementarity in protein–protein recognition. *Protein Eng.*, **9**, 741–744.
10. Althaus,E., Kohlbacher,O., Lenhof,H. and Müller,P. (2002) A combinatorial approach to protein docking with flexible side chains. *J. Comput. Biol.*, **9**, 597–612.
11. Jones,G., Willet,P., Glen,R.C., Leach,A.R. and Taylor,R. (1997) Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol.*, **267**, 727–748.
12. Vakser,I.A., Matar,O.G. and Lam,C.F. (1999) A systematic study of low-resolution recognition in protein–protein complexes. *Proc. Natl Acad. Sci. USA*, **96**, 8477–8482.
13. Gabb,H.A., Jackson,R.M. and Sternberg,M.J. (1997) Modelling protein docking using shape complementarity, electrostatics, and biochemical information. *J. Mol. Biol.*, **272**, 106–120.
14. Camacho,C.J., Gatchell,D.W., Kimura,S.R. and Vajda,S. (2000) Scoring docked conformations generated by rigid-body protein–protein docking. *Proteins*, **40**, 525–537.
15. Goldman,B.B. and Wipke,W.T. (2000) QSD quadratic shape descriptors. 2. Molecular docking using quadratic shape descriptors (QSDock). *Proteins*, **38**, 79–94.
16. Gardiner,E.J., Willett,P. and Artymiuk,P.J. (2001) Protein docking using a genetic algorithm. *Proteins*, **44**, 44–56.
17. Chen,R. and Weng,Z. (2002) Docking unbound proteins using shape complementarity, desolvation, and electrostatics. *Proteins*, **47**, 281–294.
18. 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.
19. Comeau,S.R., Gatchell,D.W., Vajda,S. and Camacho,C.J. (2004) ClusPro: a fully automated algorithm for protein–protein docking. *Nucleic Acids Res.*, **32**, W96–W99.
20. Duhovny,D., Nussinov,R. and Wolfson,H.J. (2002) Efficient unbound docking of rigid molecules. In Guigo,R. and Gusfield,D. (eds), *Proceedings of the Fourth International Workshop on Algorithms in Bioinformatics.* Springer-Verlag GmbH Rome, Italy, September 17–21, 2002, Vol. 2452, pp. 185–200.
21. Chen,R., Mintseris,J., Janin,J. and Weng,Z. (2003) A protein–protein docking benchmark. *Proteins*, **52**, 88–91.
22. Inbar,Y., Schneidman-Duhovny,D., Halperin,I., Oron,A., Nussinov,R. and Wolfson,H.J. (2005) Approaching the CAPRI challenge with efficient geometry based docking. *Proteins* (in press).
23. Schneidman-Duhovny,D., Inbar,Y., Nussinov,R. and Wolfson,H.J. (2005) Geometry based flexible and symmetric protein docking. *Proteins* (in press).
24. Eisenstein,M., Shariv,I., Koren,G., Friesem,A.A. and Katchalski-Katzir,E. (1997) Modeling supra-molecular helices: extension of the molecular surface recognition algorithm and application to the protein coat of the tobacco mosaic virus. *J. Mol. Biol.*, **266**, 135–143.
25. Berchanski,A. and Eisenstein,M. (2003) Construction of molecular assemblies via docking: modeling of tetramers with D2 symmetry. *Proteins*, **53**, 817–829.
26. Connolly,M.L. (1983) Solvent-accessible surfaces of proteins and nucleic acids. *Science*, **221**, 709–713.
27. Connolly,M.L. (1983) Analytical molecular surface calculation. *J. Appl. Crystallogr.*, **16**, 548–558.
28. Zhang,C., Vasmatzis,G., Cornette,J.L. and DeLisi,C. (1997) Determination of atomic desolvation energies from the structures of crystallized proteins. *J. Mol. Biol.*, **267**, 707–726.