John von Neumann Institute for Computing

Verification of Protein-Protein Interactions by Use of Docking Techniques

F. Fink, R. Merkl, W. Gronwald

published in

From Computational Biophysics to Systems Biology (CBSB07),
Proceedings of the NIC Workshop 2007,
Ulrich H. E. Hansmann, Jan Meinke, Sandipan Mohanty,
Olav Zimmermann (Editors),
John von Neumann Institute for Computing, Jülich,
NIC Series, Vol. **36**, ISBN 978-3-9810843-2-0, pp. 125-127, 2007.

http://www.fz-juelich.de/nic-series/volume36

Verification of Protein-Protein Interactions by Use of Docking Techniques

Florian Fink[1], **Rainer Merkl**[2], **and Wolfram Gronwald**[1]

[1] Institute of Functional Genomics,
University of Regensburg, Josef-Engert-Strasse 9,
93053 Regensburg, Germany
E-mail: florian.fink@klinik.uni-regensburg.de,
wolfram.gronwald@biologie.uni-regensburg.de

[2] Institute for Biophysics and Physical Biochemistry,
Postfach, 93040 Regensburg, Germany
E-mail: rainer.merkl@biologie.uni-regensburg.de

For the understanding of large macromolecular complexes such as ribosomes the analysis of protein-protein interactions is essential. These intermolecular interactions are strongly dependent on the three-dimensional structures of the corresponding molecules. In case that the structures are known they can be directly used while in many other cases homology modeling techniques can be applied. We have developed a novel algorithm for this purpose that allows the combination with additional experimental data to further improve the structural models. Currently we are developing tools based on a data driven docking approach and the 3D structures of the individual molecules to investigate whether proposed intermolecular interactions can be verified or falsified. In this contribution we will show first results to demonstrate the principal applicability of our approach.

1 Introduction

Most of the various functions in a cell are mediated by large protein-protein interaction networks. For a detailed understanding of these interactions knowledge of the corresponding three-dimensional complex structures is required. However, a significant amount of these complex structures will be extremely difficult to study by conventional experimental structure determination methods. One avenue to circumvent this problem is to use the structures of the individual molecules in combination with computational docking techniques. It has been shown that reliable results can be obtained when data driven docking techniques are applied[1] and only moderate structural changes occur during complex formation. As mentioned above three-dimensional structures of the individual molecules are required for the application of docking techniques. If these structures are already available they can be directly used. In many other cases the use of homology models is applicable. We have developed for this purpose the homology modeling program PERMOL[2] that is based on restrained molecular dynamics in torsion angle space. For the case that sufficient experimental data is available to obtain low resolution structures of the single molecules we developed the ISIC[3] algorithm to improve the structural quality by combining information from different sources. The key question here is how to combine the available information ensuring that no wrong structural bias is introduced.

Next the 3D-structures of the single molecules are used for *in silico* complex formation. In this contribution we focus on the question if data driven docking techniques provide

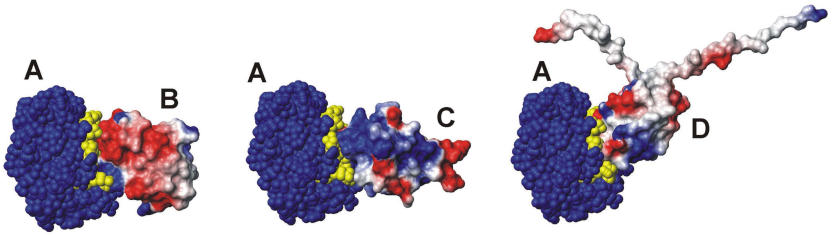

Figure 1. Example on the use of docking to query proposed intermolecular interactions. Molecules B, C and D are forced to interact with the target molecule A for which the interface (shown in yellow) is known.

Interaction A-B	Interaction A-C	Interaction A-D
E_{inter} -2512,2 kJ/mol	E_{inter} -1991,4 kJ/mol	E_{inter} -1695,6 kJ/mol

Table 1. Average interactions energies for three different trial molecules (B, C, and D) docked to the cytoplasmatic A domain (IIA(MTL)) of the mannitol transporter II (A). (B) histidine containing phospocarrier protein (HPr), (C) human cyclin dependent kinase subunit type I (CKSHS1), and (D) apo form of HMA domain of copper chaperone for superoxide dismutase.

additional information to query proposed binary protein-protein interactions. More specifically we assume that the binding-interface of one molecule is known and it is investigated which of several proposed partner molecules is the correct interaction partner.

2 Motivation

Over the last few years several high-throughput protein-protein interaction detection methods have been developed. However, as shown in the paper by von Mering et al[4] the proposed interactions of these methods usually contain many false positives. And although substantially improvements can be obtained by combing several methods it is clear that additional work is required to reduce the amount of false positives. Since intermolecular interactions are strongly dependent on the three-dimensional structures of the corresponding molecules, docking techniques should provide additional information in this regard.

3 Materials and Methods

To query proposed protein-protein interactions different trial molecules are forced to bind to the known interface of one target molecule. For this purpose the data-driven docking algorithm HADDOCK[1] that is based on the use of ambiguous interaction restraints is used. The known interface information of the target protein is provided as restraint information to the docking algorithm. This information can be obtained for example from NMR chemical shift perturbation data, mutagenesis data etc. For the proposed trial molecules

we assume that nothing is known about their interface and the complete surface of these molecules is defined as potential binding interface. Separate docking runs are performed for each of the trial molecules. Results are ordered based on the obtained interaction energies. Interaction energies are calculated based on the intermolecular van der Waals and electrostatic interface energies.

4 Results and Discussion

For testing, two molecules A and B that are known to interact are selected in the example shown in Fig. 1. Then in addition two molecules C and D, randomly selected from the PDB database are used, to perform docking runs with molecule A. For the target the binding interface shown in yellow in Fig. 1 is assumed to be known while for molecules B, C, and D the whole molecule is defined as possible interaction site. Results for this example demonstrate that the lowest interaction energies shown in Tab. 1 are obtained for the correctly interacting pair A-B while considerably higher energies are obtained for the non-interacting pairs A-C and A-D allowing to correctly discriminate between interacting and non-interacting proteins. These tests were repeated for several different test cases to investigate the general applicability of the method, where the so far obtained data show similar results (data not shown).

Therefore, in summary one can say that docking techniques can provide useful additional information to interrogate proposed protein-protein interactions.

Acknowledgments

The authors thank the bavarian genomic network for financial support.

References

1. C. Dominguez, R. Boelens and A. M. J. J. Bonvin, *HADDOCK: A Protein-Protein Docking Approach Based on Biochemical or Biophysical Information*, J. Am. Chem. Soc. **125**, 1731–1737, 2003.
2. A. Möglich, D. Weinfurtner, T. Maurer, W. Gronwald and H. R. Kalbitzer, *A Restraint Molecular Dynamics and Simulated Annealing Approach for Protein Homology Modeling Utilizing Mean angles*, BMC-Bioinformatics **6**, 91, 2005.
3. K. Brunner, W. Gronwald, J. M. Trenner, K.-P. Neidig and H. R. Kalbitzer, *A general Method for the Unbiased Improvement of Solution NMR Structures By the Use of Related X-Ray Data, the AUREMOL-ISIC Algorithm*, BMC Struc.Biol. **6**, 14, 2006.
4. C. von Mering, R. Krause, B. Snel, M. Cornell, S. G. Oliver, S. Fields and P. Bork, *Comparative Assessment of Large-Scale Data Sets of Protein-Protein Interactions*, Nature **417**, 399–403, 2002.