Bioinformatics, 32(15), 2016, 2386–2388
doi: 10.1093/bioinformatics/btw141
Advance Access Publication Date: 12 March 2016
Applications Note

OXFORD

Structural bioinformatics

FRODOCK 2.0: fast protein–protein docking server

Erney Ramírez-Aportela[1,2], José Ramón López-Blanco[1] and Pablo Chacón[1,*]

[1]Department of Biological Chemical Physics, Rocasolano Physical Chemistry Institute C.S.I.C, Madrid 28006, Spain and [2]Centro de Investigaciones Biológicas, CSIC, Madrid E-28040, Spain

*To whom correspondence should be addressed.

Associate Editor: Anna Tramontano

Received on December 9, 2015; revised on February 19, 2016; accepted on March 9, 2016

Abstract

Summary: The prediction of protein–protein complexes from the structures of unbound components is a challenging and powerful strategy to decipher the mechanism of many essential biological processes. We present a user-friendly protein–protein docking server based on an improved version of FRODOCK that includes a complementary knowledge-based potential. The web interface provides a very effective tool to explore and select protein–protein models and interactively screen them against experimental distance constraints. The competitive success rates and efficiency achieved allow the retrieval of reliable potential protein–protein binding conformations that can be further refined with more computationally demanding strategies.

Availability and Implementation: The server is free and open to all users with no login requirement at http://frodock.chaconlab.org

Contact: pablo@chaconlab.org

Supplementary information: Supplementary data are available at *Bioinformatics* online.

1 Introduction

Protein–protein interactions are involved in every cellular process, and their characterization is therefore critical to understanding the underlying molecular mechanisms. The computational prediction of protein–protein complexes from the unbound 3D structures is a successful complementary alternative to ongoing structural and proteomic endeavors. A great variety of computational *ab initio* docking approaches have been developed to predict protein complexes (Huang, 2014; Ritchie, 2008). These algorithms employ a variety of efficient search strategies and energy-based scoring strategies, usually beginning with an exhaustive six-dimensional search to generate many potential predictions, followed by a detailed and more computationally demanding refining stage. However, despite the significant progress in improved scoring functions and the integrative use of biochemical and biophysical information, current algorithms are still limited. In a recent independent comprehensive assessment of current exhaustive docking programs (Huang, 2015), the best three docking approaches succeed in identifying acceptable protein–protein solutions in 21–31, 40–52 and 60–79% of the cases within the top 10 100 and 1000 first solutions, respectively. Our approach, FRODOCK (Garzon *et al.*, 2009), which combine 3D grid-based potentials with the efficiency of spherical harmonics (SH) approximations, ranked 4th of the 18 docking/scoring functions tested. Although its success rates were slightly lower than the best three methods for the top 10 predictions (19%), FRODOCK yielded competitive success rates of 46 and 76% for the top 100 and 1000 predictions, respectively. In fact, the authors of this review considered FRODOCK to be a good docking choice if more predictions can be evaluated, which is often the case in a post-docking approach. Interestingly, top ranked methods such as ZODOCK (Pierce *et al.*, 2014) include knowledge-based energy terms that FRODOCK lacks. Here, we present a further update of FRODOCK that includes an extra knowledge-based potential. This new version significantly improves the docking success rate to the levels of existing state-of-the-art approaches but achieves superior efficiency. With integration into a new web framework, users can now predict protein–protein complexes from the unbound

components in only a few minutes. The user-friendly interface permits the interactive exploration of the docking models, including additional screening filters with experimental distance constraints to enhance the success of identifying near-native docking solutions.

2 Benchmarking and implementation

The docking algorithm is detailed elsewhere (Garzon *et al.*, 2009). FRODOCK's competitive efficiency depends on the spherical harmonic (SH) formulation to accelerate the rotational part of the search. Compared with classical 3D grid-based FFT correlation algorithms, the use of SH has proven to be a faster alternative in protein–protein docking (Ritchie, 2008; Ritchie and Kemp, 2000). Moreover, linear speed-up is obtained by performing parallel efficient rotational searches of independent docking translational positions. The original docking fitness score as defined in pyDock (Cheng *et al.*, 2007) was approximated as a correlation between soft van der Waals, electrostatics, and desolvation 3D grid potential maps conveniently transformed into SH radial (volumetric) representations. Here, we improve this scoring function by adding a complementary coarse-grained knowledge-based protein-docking potential (Tobi, 2010). We employed the ADPs-II atomic potential variant derived from known protein–protein and decoy (misdocked) complexes using a linear programming technique. This two-step contact potential fitted very well to our approximation because of its efficiency and good performance, as shown in a recent survey of protein–protein scoring functions (Moal *et al.*, 2013). To optimize the relative weights of all potential terms, we followed the same procedure described in (Garzon *et al.*, 2009). We also have optimized the potential weights for the three types of interaction disclosed in the benchmark (enzyme–substrate, antigen–antibody and others). However, compared with the addition of the pairwise potential the gain was quite reduced in terms of docking success (less than 5% in the best case scenario). The new version, FRODOCK 2.0, obtained 35–45% enhancement in the top rank solution regions using the standard benchmark 3.0 (results available on the web server). Results obtained on the 176 targets of the benchmark 4.0 (Hwang *et al.*, 2010) confirm the excellent relative docking performance of FRODOCK 2.0 (see Supplementary Table S1 and Fig. S1). The success rates to find an acceptable solution based on CAPRI criteria are 10, 29, 61 and 82% when the top 1, 10, 100 and 1000 predictions are considered, respectively. The accuracy reported for ZDOCK3.02 in (Huang, 2015), is slightly higher for the top solutions (10 versus 12) but slightly lower if more predictions are considered (61% versus 52% for the top 100 and 82% versus 79% for the top 1000). Outside this independent comparison, SwarmDock (Torchala *et al.*, 2013) also reported slightly better rates for top solutions with the same benchmark (10, 36 and 65% for the top 1, 10, 100 solutions) but considering additional costly refinement procedures, including rescoring and minimization. It is important to note that, in contrast with other methods, our success rates include the docking dependence on the starting orientation as they are computed using 50 runs per case starting from distinct initial relative orientations. The method is quite robust, and we only found relative small ranking variations of the best prediction found depending of the FRODOCK's docking sensitivity.

3 Description of the web server

The web interface is highly intuitive and responsive to all major browsers. The server performs automatic protein–protein docking of two uploaded input protein structures (in PDB format). Alternatively the user can also select the type of interaction to slightly improve the docking success rates. In only a few minutes, a results page is shown, including a graphical interface to interactively explore the predictions or retrieve the generated models (Fig. 1). The user can inspect and compare the highly scored poses and visualize docking hot spots. Moreover, a practical interface is included to add multiple distance constraints within a user-specified tolerance, providing an effective way to narrow down the exhaustive docking search. This information is known to significantly enhance docking success rates and is often available from various biophysical and biochemical techniques such as chemical cross-linking, FRET, nuclear or electron magnetic resonances, mass spectrometry, or mutagenesis experiments. Finally, the server is completed with an extensive gallery of examples including all the benchmark test cases presented here to illustrate the quality and limitations of the approach.

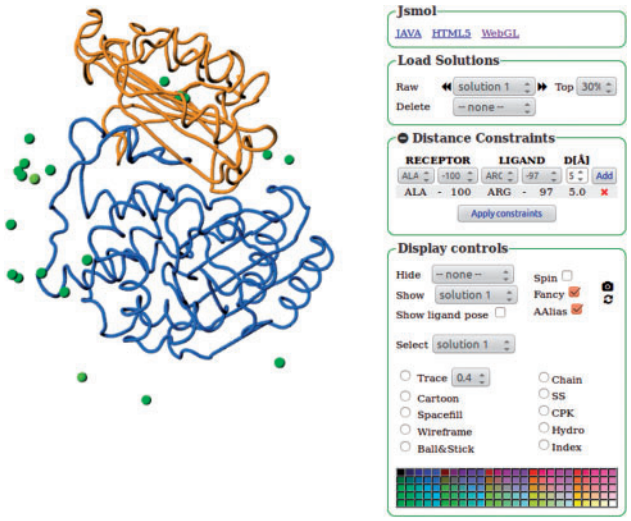

Fig. 1. Sample results page. In this example, the high score docking solution (spheres) of the ligand is displayed together with the protein receptor (dark ribbons). The spheres correspond to translational hot spots where the docking scoring was higher than 30%. The user can interactively load and represent different docking solutions with the help of JSmol (Color version of this figure is available at *Bioinformatics* online.)

One of the major advantages of the server is its efficiency. The average docking running time for the benchmark 4.0 test cases is 2 min. FRODOCK 2.0 is a faster alternative to state-of-the-art servers such as ZDOCK (Pierce *et al.*, 2014) (11 min) or SwarmDock (Torchala *et al.*, 2013) (36 h) with similar predictive performance. Moreover, the user can test different constraints and thresholds on the fly without re-running the docking protocol as in other servers because the constraint module is implemented as a post-docking procedure.

4 Conclusions

Overall, the server quickly provides reliable potential protein–protein binding conformations that can be further refined and validated by more computationally demanding post-docking procedures. The docking accuracy achieved by adding a complementary knowledge-based protein-docking potential is comparable to the results of existing global docking tools but significantly faster. Future developments include the incorporation of additional experimental constraints (e.g. low resolution information from electron microscopy maps or SAXS) and the integration with docking refinement tools.

Funding

This study was supported by Spanish grants BFU2013-44306-P, and BFU2014-51823-R from the Ministerio de Economıa y Competitividad (MINECO) and CM S2010/BMD-2353 from the Comunidad de Madrid.

Conflict of Interest: none declared.

References

Cheng,T.M.K. *et al*. (2007) PyDock: electrostatics and desolvation for effective scoring of rigid-body protein–protein docking. *Proteins: Struct. Funct. Genet*., **68**, 503–515.

Garzon,J.I. *et al*. (2009) FRODOCK: a new approach for fast rotational protein–protein docking. *Bioinformatics*, **25**, 2544–2551.

Huang,S.Y. (2014) Search strategies and evaluation in protein–protein docking: principles, advances and challenges. *Drug Discov. Today*, **19**, 1081–1096.

Huang,S.Y. (2015) Exploring the potential of global protein–protein docking: an overview and critical assessment of current programs for automatic ab initio docking. *Drug Discov. Today*, **20**, 969–977.

Hwang,H. *et al*. (2010) Protein–protein docking benchmark version 4.0. *Proteins: Struct. Funct. Bioinf*., **78**, 3111–3114.

Moal,I.H. *et al*. (2013) The scoring of poses in protein–protein docking: current capabilities and future directions. *BMC Bioinformatics*, **14**, 286–301.

Pierce,B.G. *et al*. (2014) ZDOCK server: interactive docking prediction of protein–protein complexes and symmetric multimers. *Bioinformatics*, **30**, 1771–1773.

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

Ritchie,D.W. and Kemp,G.J.L. (2000) Protein docking using spherical polar Fourier correlations. *Proteins: Struct. Funct. Genet*., **39**, 178–194.

Tobi,D. (2010) Designing coarse grained-and atom based-potentials for protein–protein docking. *BMC Struct. Biol*., **10**, 40.

Torchala,M. *et al*. (2013) SwarmDock: a server for flexible protein–protein docking. *Bioinformatics*, **29**, 807–809.