Published online 12 May 2021 *Nucleic Acids Research, 2021, Vol. 49, Web Server issue* ***W277–W284***
https://doi.org/10.1093/nar/gkab358

InterEvDock3: a combined template-based and free docking server with increased performance through explicit modeling of complex homologs and integration of covariation-based contact maps

Chloé Quignot[1], Guillaume Postic[2], Hélène Bret[1], Julien Rey[2], Pierre Granger[1], Samuel Murail[2], Pablo Chacón[3], Jessica Andreani[1,*], Pierre Tufféry[2,*] and Raphaël Guerois[1,*]

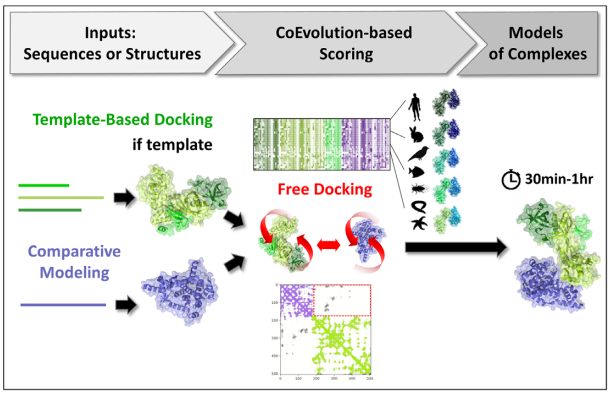
[1]Université Paris-Saclay, CEA, CNRS, Institute for Integrative Biology of the Cell (I2BC), 91198 Gif-sur-Yvette, France, [2]Université de Paris, CNRS UMR 8251, INSERM U1133, RPBS, Paris 75205, France and [3]Department of Biological Physical Chemistry, Rocasolano Institute of Physical Chemistry C.S.I.C, Madrid, Spain

Received March 13, 2021; Revised April 09, 2021; Editorial Decision April 21, 2021; Accepted April 23, 2021

ABSTRACT

The InterEvDock3 protein docking server exploits the constraints of evolution by multiple means to generate structural models of protein assemblies. The server takes as input either several sequences or 3D structures of proteins known to interact. It returns a set of 10 consensus candidate complexes, together with interface predictions to guide further experimental validation interactively. Three key novelties were implemented in InterEvDock3 to help obtain more reliable models: users can (i) generate template-based structural models of assemblies using close and remote homologs of known 3D structure, detected through an automated search protocol, (ii) select the assembly models most consistent with contact maps from external methods that implement covariation-based contact prediction with or without deep learning and (iii) exploit a novel coevolution-based scoring scheme at atomic level, which leads to significantly higher free docking success rates. The performance of the server was validated on two large free docking benchmark databases, containing respectively 230 unbound targets (Weng dataset) and 812 models of unbound targets (PPI4DOCK dataset). Its effectiveness has also been proven on a number of challenging examples. The InterEvDock3 web interface is available at http://bioserv.rpbs.univ-paris-diderot.fr/services/InterEvDock3/.

GRAPHICAL ABSTRACT



INTRODUCTION

Physical interactions between proteins are driving the structural and dynamical organization of cells and have played fundamental roles in the emergence and evolution of complex cellular functions (1). To elucidate the functioning of cell machines and unravel their cross-talk, it is essential to characterize the molecular bases underlying their specific association into large and dynamic assemblies (2). Computational methods for predicting the structure of protein assemblies are needed to exploit the enormous amount of physical interactions discovered by proteomics technologies and make the most of 3D structures already available (3).

In numerous cases, sequence conservation and coevolution bring major insights into the constraints that were imposed during evolution at the interface of protein complexes (4). InterEvDock3 is a unique protein docking server

*To whom correspondence should be addressed. Tel: +33 1 69 08 67 17; Fax: +33 1 69 08 47 12; Email: raphael.guerois@cea.fr
Correspondence may also be addressed to Pierre Tufféry. Tel: +33 1 57 27 83 86; Fax: +33 1 57 27 83 72; Email: pierre.tuffery@univ-paris-diderot.fr
Correspondence may also be addressed to Jessica Andreani. Tel: +33 1 69 08 67 17; Fax: +33 1 69 08 47 12; Email: jessica.andreani@cea.fr

that exploits the constraints of evolution by multiple means to generate models of protein structural assemblies. InterEvDock3 integrates into a single server three major procedures that can be used by biologists to get reliable models from the sequences of their proteins of interest: (i) Template-based docking; (ii) Free-template docking using covariation-inferred contact maps for cases with large co-alignments (joint multiple sequence alignments, coMSAs in short, of two interacting partners); (iii) Free docking using explicit modeling of interologs and scoring at atomic level for cases with coMSAs as small as 10 sequences.

Critical assessment of strategies for modeling protein interactions, as achieved by CAPRI (Critical Assessment of PRedicted Interactions) experiments (5,6), continuously rate template-based modeling as a reliable approach provided that a template can be unambiguously identified. Useful servers for that purpose were developed in recent years, either for pairs of chains or multiple sequence inputs such as SnapDock (7), SWISS-MODEL Oligo (8,9) and HDOCK (10). Search for templates hierarchically clustered can be further achieved with PPI3D (11) and KBDOCK (12). If no template can be found, free docking approaches become powerful alternatives in which the orientation between two binding partners can be extensively sampled and scored. Many template-free docking servers have been developed using a global rigid-body docking approach, sometimes followed by rescoring such as PatchDock (13), FireDock (14), ZDOCK (15), FRODOCK (16), pyDockWEB (17), ClusPro (18) and GRAMM-X (19). Servers and methods such as RosettaDock integrated into the ROSIE framework (20,21), HADDOCK (22,23), LightDock (24) or PyDOCKweb are appropriate for focused sampling around an initial guess. A few servers also have the specificity to include significant degrees of flexibility as SwarmDock (25), ATTRACT (26) or HADDOCK. In complement to these free docking approaches, seminal works highlighted the interest of template-based modeling (27,28) and of using homology models for docking (29–32). Among all these docking servers and methods, HDOCK has the particularity of running a hybrid template-based and free docking strategy restricted to dimeric cases when only sequences are provided.

InterEvDock3 is unique in integrating template-based and free docking with a flexible processing of inputs (single- or multi-protein, structures or sequences) and using evolutionary information to rank the generated docking models. In comparison to the previous versions of the server, three novel prediction strategies have been developed. From an input containing multiple sequences, the server will first explore whether any homologous complex is available as a template to predict the structure of an assembly by comparative modeling techniques (Figure 1). Only a subset of the query proteins may fulfill these conditions and InterEvDock3 will build template-based models for the largest such subset. For interfaces that cannot be modeled by homology, InterEvDock3 proposes two complementary free docking approaches. The first one is particularly well adapted when large coMSAs can be generated for the subunits of interest. It exploits the covariation-inferred contact maps that can be generated from servers and methods that implement Direct Coupling Analyses such as EvComplex (33) or that combine this analysis with deep learning techniques such as ComplexContact (34) or trRosetta (35). The third modeling solution proposed by InterEvDock3 is a new strategy we developed for atomic-based scoring integrating interface coevolution (36). For each model sampled by FRODOCK (16), atomic models of interfaces of coevolved homologs are scored using a successful consensus strategy exploiting the complementarity of scores such as InterEvScore (37), SOAP-PP (38) and Rosetta interface score (39). Moving from a coarse-grained integration of evolutionary information to an atomic representation, the success rates of InterEvDock3 were increased by more than 20%.

In addition to these three novelties, the InterEvDock3 server inherits many features of previous versions that allowed our group to rank among the best predictors on past CAPRI challenges (40–42). Such useful features include automatic generation of co-alignments for multi-subunit inputs, handling of user-specific distance constraints, interactive viewers for rapid visual inspection complemented with tailored PyMOL scripts to facilitate the comprehensive analysis of the docking results, and the possibility to use a previous job id session for hot restarts. These functionalities have been adapted and extended to the new methodological improvements in order to facilitate their use.

The InterEvDock3 pipeline was extensively tested and the free docking part was benchmarked on 230 unbound partners of the Weng database (43) and 812 unbound homology models from the PPI4DOCK database (30). 40% and 35% of the 230 and 812 cases, respectively, have an acceptable or better solution among the top 10 consensus models returned by InterEvDock3. InterEvDock3 also outputs a list of the 10 residues most likely involved in the interface and at least one residue was correctly predicted in 91% of the 812 benchmark cases. We also present several challenging cases to illustrate the extended applicability reached by the incorporation of novel evolution-based docking strategies.

THE INTEREVDOCK3 SERVER

Presentation of the InterEvDock3 web interface and of the molecular docking procedure

Figure 1 illustrates the new features implemented in the InterEvDock3 pipeline. The server can process two or several proteins provided either in sequence or structure format in the 'Partner A' and/or 'Partner B' sections of the web form. In InterEvDock3, there are overall three major tracks.

Template-based docking protocols

First, a template-based modeling protocol can be activated by submitting protein sequences as inputs (Figure 1.1). The user can provide several sequences in the 'Partner A' section to explore all the complexes that contain homologs of this set of sequences. When at least one homologous complex is detected, the server will automatically build a 3D model using a template-based docking strategy. If one or several sequences are provided in both sections A and B, the server will first generate 3D models for both partners individually and then run a free docking protocol between them. The search for a structural complex template is performed

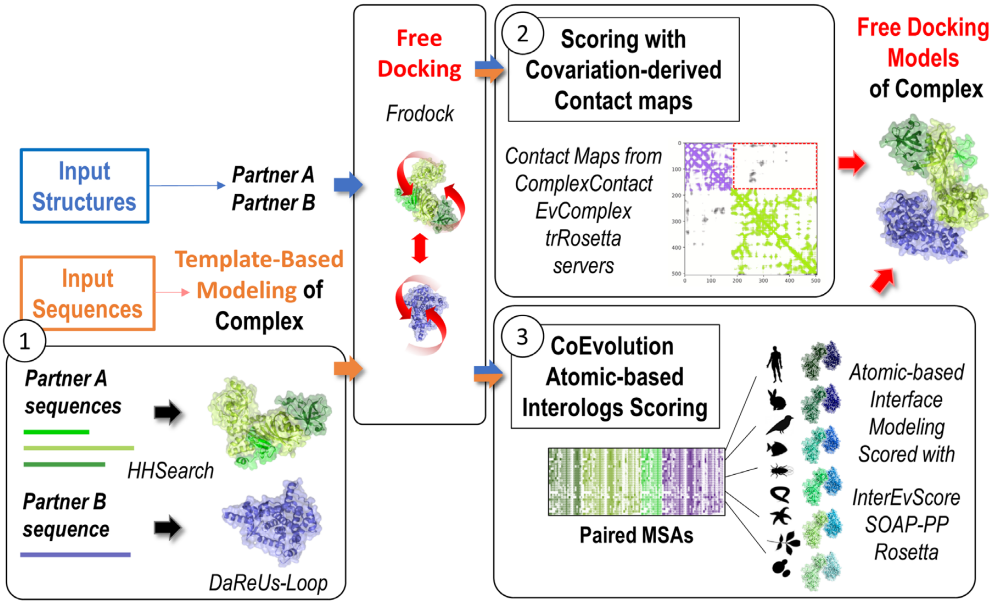


Figure 1. General pipeline highlighting the 3 novel approaches implemented in the InterEvDock3 server. Mode 1 uses sequences as input and runs a template-based modeling protocol search for close and remote homologs with HHsearch in an exhaustive manner to generate models of homomeric and heteromeric assemblies. Mode 2 uses 3D structures of monomers or homomeric complexes to run a free docking approach trying to satisfy the contacts predicted in the contact map provided as input and obtained from servers such as ComplexContact (34), EvComplex (52) or trRosetta (35). Mode 3 uses 3D structures of monomers or multimeric complexes (possibly modeled from sequences in Mode 1) and implements a new strategy for scoring interfaces with coevolution information using an atomic-based scoring of interface models for 10 to 40 pairs of interolog sequences using scores such as InterEvScore (37), SOAP-PP (38) or Rosetta scoring function (39). A single complex model is returned after mode 1 while 10 best models are returned for every score calculated in modes 2 and 3. Mode 3 also returns a consensus selection of 10 best models and a prediction of residues likely to be in the interface. Template-based modeling (mode 1) is in general more reliable than free docking approaches and should be favored if possible. In case large multiple sequence alignments can be obtained, it is advised to run first the coevolution-derived contact mode (mode 2) which can strongly restrict the relative orientations between docked partners. Mode 3 should be privileged in cases where only shallow co-MSAs can be obtained.

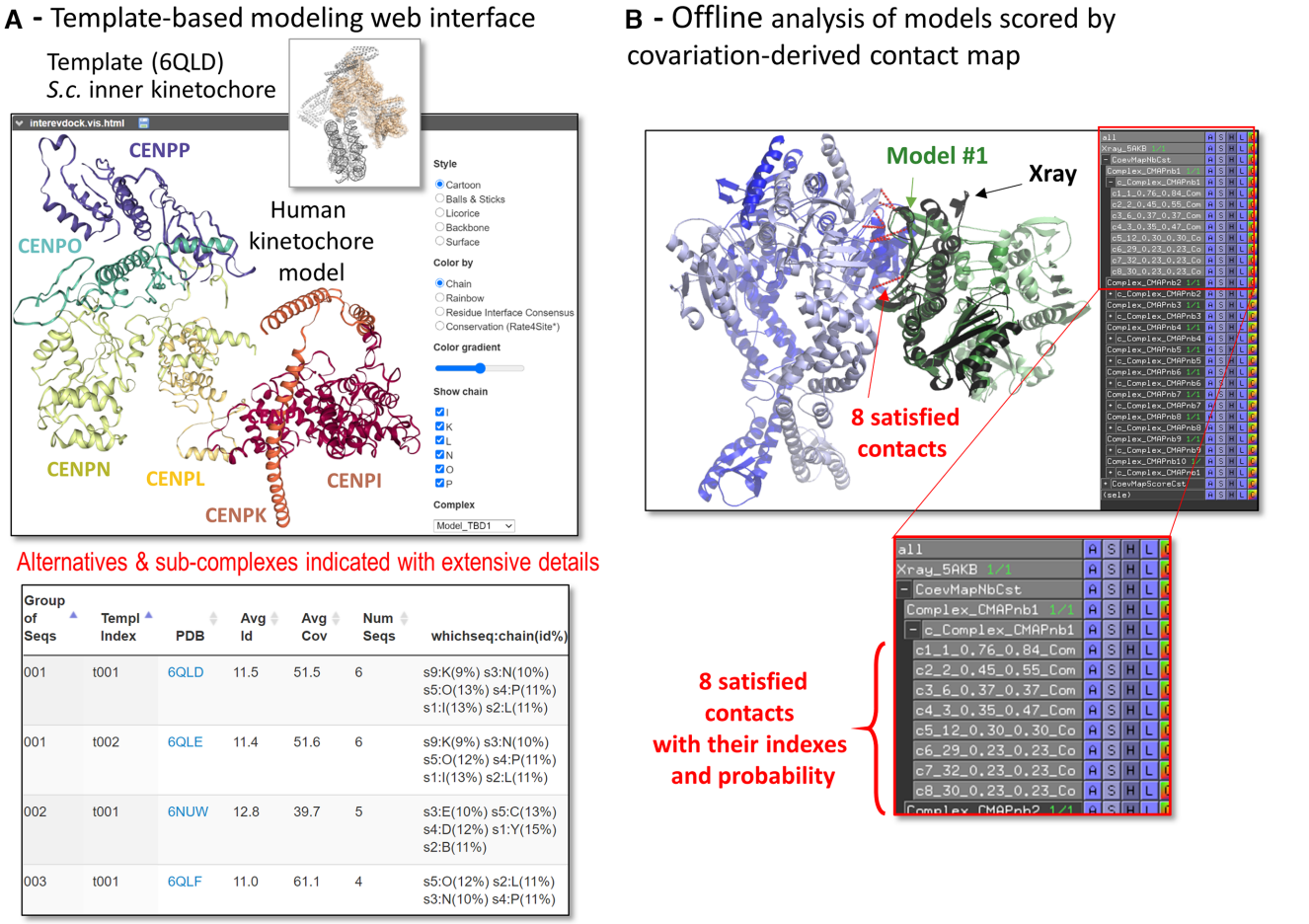
using a protocol we recently developed to identify, among a list of interacting proteins, those which have homologs structurally characterized in complex with each other (44). The latter protocol was also included in the Proteo3Dnet server (45). Close and remote homologous complexes are identified using HHsearch (46) and the computed alignments are used to create a structural assembly by threading the maximum number of input subunits possible. Side-chains are modeled using OSCAR-star (47). Insertions and deletions are subsequently modeled by homology using our DaReUS-Loop program adapted to work on inputs with multiple chains (48,49). A table provided in the output of the automatic procedure lists a number of possible templates which can help select alternative structures (Figure 2A). More options are available through the 'advanced options' menu allowing, for instance, to change the maximum authorized loop length in template-based modeling or to specify a different template. When structure inputs are provided instead of sequences, the homology building step is logically bypassed.

Free docking using coevolution information for scoring

Once the structures of Partners A and B are available, either through 3D homology modeling or directly from the user inputs, an exhaustive rigid-body sampling is performed between them using the FRODOCK algorithm. A notable feature of InterEvDock3 is to use evolutionary information, when available, to score the models sampled by exhaustive rigid-body search. In InterEvDock3, this evolutionary information is integrated into two novel and complementary ways.

Free docking with scoring based on covariation-inferred contact maps

The second novel approach implemented in the InterEvDock3 server is the possibility to exploit the contact maps predicted from external servers analyzing covariation with direct coupling analysis and/or deep learning techniques (Figure 1.2) (33–35). The format of the input contact map is a simple 3-column text file, such as the one generated by the ComplexContact server ('.gcnn_inter_list' extension), sorting by decreasing probability (col 3) the indexes of the predicted contacting residue pairs (col 1 and 2). For every model, InterEvDock3 considers the contacts listed in the covariation map and counts how many are satisfied at an inter-atomic threshold distance of 8Å (default value, adjustable in options). We typically advise users to submit the top 100 or top 200 predicted inter-molecular contacts with the idea that the docking step will help discriminate the correct ones and filter out contacts incompatible with the best scored docking models. A sum weighted by the probabilities of these predicted contacts is also calculated. The server returns two sets of 10 models maximizing either the number of satisfied structural contacts or the sum of satisfied contacts weighted by their probability. Because vicinal residues often bring redundant contact information in the



Group of Seqs	Templ Index	PDB	Avg Id	Avg Cov	Num Seqs	whichseq:chain(id%)
001	t001	6QLD	11.5	51.5	6	s9:K(9%) s3:N(10%) s5:O(13%) s4:P(11%) s1:I(13%) s2:L(11%)
001	t002	6QLE	11.4	51.6	6	s9:K(9%) s3:N(10%) s5:O(12%) s4:P(11%) s1:I(13%) s2:L(11%)
002	t001	6NUW	12.8	39.7	5	s3:E(10%) s5:C(13%) s4:D(12%) s1:Y(15%) s2:B(11%)
003	t001	6QLF	11.0	61.1	4	s5:O(12%) s2:L(11%) s3:N(10%) s4:P(11%)

Figure 2. (**A**) Model automatically generated using InterEvDock3 in mode 1 for 6 human proteins forming a subcomplex of the inner kinetochore display in the online web interface of InterEvDock3. Despite very low mean sequence identity between human and yeast subunits (11.5% on average), HHsearch detected orthologous subunits for CENPI (13%), CENPO (13%), CENPL (11%), CENPP (11%), CENPN (10%) and CENPK (9%) as indicated in the table of templates listing all the possible complexes and subcomplexes which could be used as templates for this set of sequences. (**B**) Structural model obtained using InterEvDock3 in mode 2 (by applying the top 100 contacts predicted by the ComplexContact server) for the complex between *E. coli* MutS homodimer and MutL homodimeric N-terminal domain (PDB:1B63). In the crystal structure of the complex, a single domain of MutL was crystallized (5AKB). A PyMol script 'start_analysis_cmap.pml' is distributed in the downloadable results archive to get all the predicted contacts in the map which were satisfied in the model as shown in the figure.

context of protein structures or structural models used as free docking inputs, InterEvDock3 uses a grouping strategy in which vicinal sequence positions are considered equivalent. By default, a central position is grouped with its two upstream and two downstream positions (this number can be adjusted in the advanced options and redundant grouping can be canceled by setting the parameter to 0). Importantly, contact map-based scoring in InterEvDock3 takes into account the ambiguities generated by the docking of homo-multimers in which several monomers may simultaneously contribute to otherwise incompatible contacts (see examples in the Results section).

Free docking using a consensus scoring approach running at the atomic level

To identify the most likely coevolved interfaces, free docking poses can be rescored using a consensus approach combining up to five different scores (Figure 1.3). The scoring protocol exploits coMSAs of both input partners, automatically and efficiently built if not provided by the user, even for oligomeric inputs. The novelty in InterEvDock3 (36) is that a subset of 10 to 40 representative couples of homologous sequences are selected from the coMSAs and are rapidly modeled at the atomic level to be scored simultaneously. The representative pairs of sequences form a group of so-called interolog sequences. For every docking pose and for each interolog, the server rapidly generates an atomic model of the interface with side-chain conformations predicted using OSCAR-star (47) without modeling insertions. The latter approximation can be tolerated because sequences share more than 30% identity with the master sequence. All these steps are parallelized allowing to score 10,000 rigid-body models in a reasonable time (see below). For every model, a fast scoring mode is calculated by summing the scores of all modeled interologs using SOAP-PP and InterEvScore. The FRODOCK score is only calculated for the master sequence. For the three scores, the top 50 models (150 models

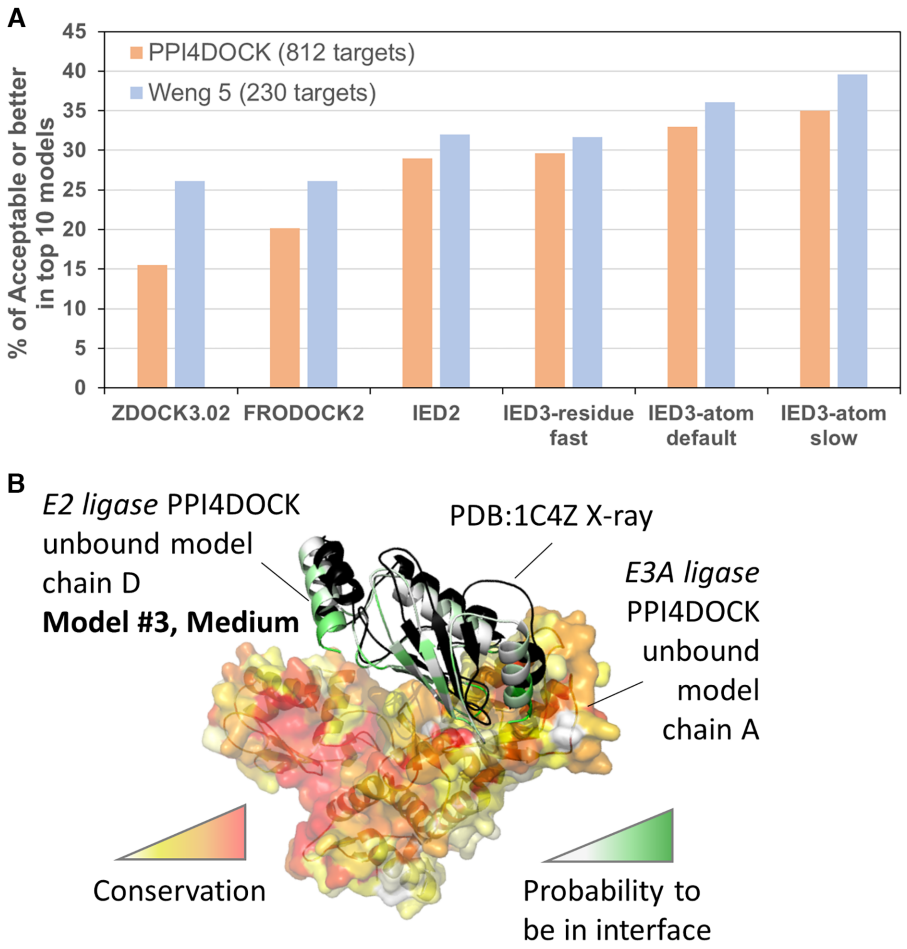


Figure 3. (**A**) Increased performance of InterEvDock3 (IED3) free docking server (Figure 1 mode 3) benchmarked on the Weng and PPI4DOCK datasets compared to the success rates obtained with the previous version of InterEvDock2 (IED2) (53). Success rates obtained for three options which can be activated in the web interface are shown. (**B**) Example of a docking model obtained for a target of the PPI4DOCK dataset involving unbound models of E2 and E3A ligases. The model representation was obtained from the PyMOL script available in the result archive from InterEvDock3. Colored representation corresponds to either the conservation (yellow-red scale) or the probability of a residue to be part of the interface (white-green scale).

in total) are returned to the user and a consensus calculation extracts a final list of 10 best models. The consensus accounts for the fact that decoys that are well ranked by at least two different scoring methods have a higher chance of being correct (36). Increased performance can be further achieved by activating the option to run Rosetta interface scoring on the 150 selected models at the cost of a longer computation time for large complexes (>500 residues). As in InterEvDock2, users can provide information on residues (or pairs of residues) involved in the interface as restraints to filter sampled solutions. When constraints are provided by the user, the output returned by the server will provide information about whether or not each constraint was used during docking (e.g. constraints involving residues that are not exposed on the surface of the protein are excluded).

Server output

The web page resulting from an InterEvDock3 submission contains information about the best ranked decoys, which can be explored interactively thanks to the NGL applet (50). Models can be minimized or taken without modifications with respect to the input structures. Detailed results are available in a downloadable archive, also containing a script for easy loading and offline visualization of the best docking solutions with PyMOL (The PyMOL Molecular Graphics System, Schrödinger, LLC). The InterEvDock3 server benefits from parallelized implementation in the dedicated infrastructure built at RPBS and from data privacy ensured in the Mobyle framework.

Several tutorial cases using sequences as input to the docking or adding contact maps are available on the InterEvDock3 submission page, presented in a video tutorial and described in the help page at https://bioserv.rpbs.univ-paris-diderot.fr/services/InterEvDock3/#examples.

Runtime

Free docking with atomic coevolution scoring (Figure 1.3) takes altogether around 30 min for 2 proteins of 100 residues each and 1 h 15 min for two proteins of 500 residues each. 50% of the time is dedicated to coMSA generation. When free docking using a covariation map (Figure 1.2), the user can score a much larger amount of decoys, considering them

all or reducing them from 300 000 decoys (1h, the default setting) to 100 000 decoys (30 min). Template search and query-template alignment steps take only a few minutes, whatever the size of the proteins (Figure 1.1). The comparative modeling step typically takes 20 min (two chains of 150 residues) to 2 h (five chains of 500 residues), also depending on the number of loops to remodel. To save time, each docking run is associated with a session identifier which can be provided as input of another run (hot restart option) to re-use previously sampled docking poses and explore alternative template, scoring and post-processing options. Session identifiers are also useful to run several fast rigid-body docking simulations and, only when results are satisfactory, apply a final minimization of the models by recalling the best run.

RESULTS

Identification of close and remote homolog multi-subunit complexes for template-based docking

A key step of our strategy (44) is that homologous templates can be identified at both high and very low sequence identity. An illustration of the limits which can be reached by template-based docking in InterEvDock3 is shown in Figure 2A for the modeling of six subunits of the human inner kinetochore based on the structure solved in yeast. Despite sequence identities ranging from 9% to 13% for all the subunits, InterEvDock3 could automatically generate the remote homologous assembly. To our knowledge, the Swiss-model server (9) is the only alternate server offering the possibility to generate template-based models of complexes at such low sequence identities. Swissmodel requires that all submitted sequences match a common template, while InterEvDock3 offers a more exploratory approach, also detecting sub-complexes and allowing to build a model for a subset of the submitted sequences (Figure 2A).

Unique to InterEvDock3 is the possibility to combine template-based docking of a set of sequences with a subsequent free docking protocol to explore how modeled assemblies interact together. An example is provided with the ComM protein, a conserved helicase involved in bacterial competence (Supplementary Figure S1).

Free docking and scoring from covariation-inferred contact maps

The use of covariation maps has been demonstrated as a powerful means to predict the assembly of heteromeric complexes (51,52). As an example, we tested the case of a challenging bacterial complex between two homomeric subunits involved in DNA mismatch repair, MutS and MutL. We used the ComplexContact server (Figure 2B) and the trRosetta method (Supplementary Figure S2) to predict contact maps based on an automatically generated co-alignment by ComplexContact with >5000 sequences. Of note, the trRosetta server folds single chain proteins rather than multi-protein assemblies. However, it is possible to download a standalone version in which joint co-alignments can be given as inputs and as a result retrieve the list of predicted inter-molecular contacts ranked by their decreasing probability. Figure 2B highlights that the best model is of Medium quality with respect to the native crystal structure as defined by CAPRI criteria. When downloading the results archive users can run a PyMol script as shown in Figure 2B which assigns dashed lines to all the satisfying contacts with their probabilities in the covariation map, allowing the user to judge the reliability of every contact. Another target proposed recently in CAPRI, target T163, could also be handled successfully by the server (Supplementary Figure S3) and a set of six examples using either sequences or structures as inputs is shown in Supplementary Figure S4.

Free docking and scoring of coevolution at atomic level

In previous versions of the InterEvDock3 server, we highlighted that coevolution information captured by InterEvScore at the coarse-grained residue level significantly increased docking success rates when combined with other atomic-based scoring functions such as FRODOCK or SOAP-PP. Because performance gains are observed even when as few as ten sequences are retrieved in the co-MSAs, this consensus-based approach is highly complementary to the use of the covariation maps which require larger co-alignments. To further increase docking performance, we developed a method to score coevolved interfaces at the atomic level rather than at the residue level (36).

We exhaustively benchmarked the performance of that new atomic-based scoring strategy in the InterEvDock3 pipeline (mode 3 in Figure 1) and compared it to the previous performance of InterEvDock2 and to other well established methods such as ZDOCK3.02 and FRODOCK 2.1. We tested 230 targets of the Weng dataset (unbound experimental monomers) together with 812 targets from the PPI4DOCK dataset (unbound homology models). As shown in Figure 3A and Supplementary Figure S5A, in all the scoring variants, a significant increase in the number of targets with Acceptable or better models among the top 10 decoys is observed reaching up to 35% and 39.6% of correct predictions on the PPI4DOCK and Weng dataset, respectively. InterEvDock3 can take advantage of several atomic-based scores in addition to SOAP-PP such as the Rosetta interface score applied either at the level of the master sequence or applied as for SOAP-PP on all the interologs (so-called IED3-atom slow mode). This strategy was thoroughly benchmarked, showing improvements with respect to several standard scoring functions on the two benchmark datasets (36). In addition to the top 10 consensus models, users can also download an output archive containing the PDB files of the top 50 models of every individual score used in the consensus together with a PyMOL script to visualize important features mapped at the surface of the models (Figure 3B). Of key interest for experimental biologists using mode 3 presented in Figure 1, InterEvDock3 predicts a list of 10 residues most likely involved in the complex interface that can be targeted for mutagenesis (5 on each partner). For these residue predictions, we reach 91% success rate, with 737 of the 812 PPI4DOCK benchmark cases having at least one of the 10 predicted residues involved in the actual interface (Supplementary Figure S5B).

CONCLUSION

The efficiency of protein–protein docking algorithms bolsters the development of increasingly sophisticated scoring functions able to better discriminate native interfaces. In InterEvDock3, we managed to design an efficient atomic based scoring of interface coevolution that significantly increases the reliability of the models. Typical server runtimes are 30 min (for proteins of around 100 residues) to 1 hour 15 min (for proteins of around 500 residues). As complementary approaches are essential to the evolution of the docking field, we also provide the opportunity to combine template-based modeling of oligomeric assemblies and use of increasingly important covariation-derived contact maps. While the server can use all these steps in a fully automated manner, it also offers a range of options at each step of the simulation to help users refine their hypotheses, as well as a user-friendly on-line and off-line display of the results.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We are very grateful to Andrej Sali (SOAP-PP), Tal Pupko (Rate4Site), Peter Rose (NGL viewer), Sjoerd de Vries (DaReUS-Loop), Kazutaka Katoh (MAFFT), Daron Standley (OSCAR-star), Johannes Söding, Milot Mirdita and Christian Roth (HHsuite) and the RosettaCommons community for making their methods available for implementation in the InterEvDock3 server.

FUNDING

French Infrastructure for Integrated Structural Biology (FRISBI) [ANR-10-INSB-05-01 to R.G.]; Agence Nationale de la Recherche through grants CHIPSET [ANR-15-CE11-0008-01 to R.G.]; ESPRINet [ANR-18-CE45–0005-01 to J.A.]; French Institute for Bioinformatics (IFB) [ANR-14–2011-IFB to G.P.]; IdEx Université de Paris [ANR-18-IDEX-0001 to P.T.], IDEX Paris-Saclay [IDI 2017 to C.Q.]; MINECO [BFU2016–76220-P to P.C.]; AEI/FEDER, UE [PID2019-109041GB-C21 to P.C.]. Funding for open access charge: FRISBI [ANR-18-CE45-0005-01].
Conflict of interest statement. None declared.

REFERENCES

1. Marsh,J.A. and Teichmann,S.A. (2015) Structure, dynamics, assembly, and evolution of protein complexes. *Annu. Rev. Biochem.*, **84**, 551–575.
2. Rosell,M. and Fernández-Recio,J. (2020) Docking approaches for modeling multi-molecular assemblies. *Curr. Opin. Struct. Biol.*, **64**, 59–65.
3. Vakser,I.A. (2020) Challenges in protein docking. *Curr. Opin. Struct. Biol.*, **64**, 160–165.
4. Andreani,J., Quignot,C. and Guerois,R. (2020) Structural prediction of protein interactions and docking using conservation and coevolution. *WIREs Comput. Mol. Sci.*, **10**, e1470.
5. Lensink,M.F., Velankar,S. and Wodak,S.J. (2017) Modeling protein-protein and protein-peptide complexes: CAPRI 6th edition. *Proteins*, **85**, 359–377.
6. Lensink,M.F., Nadzirin,N., Velankar,S. and Wodak,S.J. (2020) Modeling protein-protein, protein-peptide, and protein-oligosaccharide complexes: CAPRI 7th edition. *Proteins*, **88**, 916–938.
7. Estrin,M. and Wolfson,H.J. (2017) SnapDock-template-based docking by geometric hashing. *Bioinformatics*, **33**, i30–i36.
8. Bertoni,M., Kiefer,F., Biasini,M., Bordoli,L. and Schwede,T. (2017) Modeling protein quaternary structure of homo- and hetero-oligomers beyond binary interactions by homology. *Sci. Rep.*, **7**, 10480.
9. Waterhouse,A., Bertoni,M., Bienert,S., Studer,G., Tauriello,G., Gumienny,R., Heer,F.T., de Beer,T.A.P., Rempfer,C., Bordoli,L. *et al.* (2018) SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.*, **46**, W296–W303.
10. Yan,Y., Tao,H., He,J. and Huang,S.Y. (2020) The HDOCK server for integrated protein-protein docking. *Nat. Protoc.*, **15**, 1829–1852.
11. Dapkunas,J., Timinskas,A., Olechnovic,K., Margelevicius,M., Diciunas,R. and Venclovas,C. (2017) The PPI3D web server for searching, analyzing and modeling protein-protein interactions in the context of 3D structures. *Bioinformatics*, **33**, 935–937.
12. Ghoorah,A.W., Devignes,M.D., Smaïl-Tabbone,M. and Ritchie,D.W. (2014) KBDOCK 2013: a spatial classification of 3D protein domain family interactions. *Nucleic Acids Res.*, **42**, D389–D395.
13. Schneidman-Duhovny,D., Inbar,Y., Nussinov,R. and Wolfson,H.J. (2005) PatchDock and SymmDock: servers for rigid and symmetric docking. *Nucleic Acids Res.*, **33**, W363–W367.
14. Mashiach,E., Schneidman-Duhovny,D., Andrusier,N., Nussinov,R. and Wolfson,H.J. (2008) FireDock: a web server for fast interaction refinement in molecular docking. *Nucleic Acids Res.*, **36**, W229–W232.
15. Pierce,B.G., Wiehe,K., Hwang,H., Kim,B.H., Vreven,T. and Weng,Z. (2014) ZDOCK server: interactive docking prediction of protein-protein complexes and symmetric multimers. *Bioinformatics*, **30**, 1771–1773.
16. Ramírez-Aportela,E., López-Blanco,J.R. and Chacón,P. (2016) FRODOCK 2.0: fast protein-protein docking server. *Bioinformatics*, **32**, 2386–2388.
17. Jiménez-García,B., Pons,C. and Fernández-Recio,J. (2013) pyDockWEB: a web server for rigid-body protein-protein docking using electrostatics and desolvation scoring. *Bioinformatics*, **29**, 1698–1699.
18. Kozakov,D., Hall,D.R., Xia,B., Porter,K.A., Padhorny,D., Yueh,C., Beglov,D. and Vajda,S. (2017) The ClusPro web server for protein-protein docking. *Nat. Protoc.*, **12**, 255–278.
19. Tovchigrechko,A. and Vakser,I.A. (2006) GRAMM-X public web server for protein-protein docking. *Nucleic Acids Res.*, **34**, W310–W314.
20. Lyskov,S. and Gray,J.J. (2008) The RosettaDock server for local protein-protein docking. *Nucleic Acids Res.*, **36**, W233–W238.
21. Moretti,R., Lyskov,S., Das,R., Meiler,J. and Gray,J.J. (2018) Web-accessible molecular modeling with Rosetta: The Rosetta Online Server that Includes Everyone (ROSIE). *Protein Sci.*, **27**, 259–268.
22. de Vries,S.J., van Dijk,M. and Bonvin,A.M. (2010) The HADDOCK web server for data-driven biomolecular docking. *Nat. Protoc.*, **5**, 883–897.
23. van Zundert,G.C.P., Rodrigues,J., Trellet,M., Schmitz,C., Kastritis,P.L., Karaca,E., Melquiond,A.S.J., van Dijk,M., de Vries,S.J. and Bonvin,A. (2016) The HADDOCK2.2 web server: user-friendly integrative modeling of biomolecular complexes. *J. Mol. Biol.*, **428**, 720–725.
24. Roel-Touris,J., Bonvin,A. and Jiménez-García,B. (2020) LightDock goes information-driven. *Bioinformatics*, **36**, 950–952.
25. Torchala,M., Moal,I.H., Chaleil,R.A., Fernandez-Recio,J. and Bates,P.A. (2013) SwarmDock: a server for flexible protein-protein docking. *Bioinformatics*, **29**, 807–809.
26. de Vries,S.J., Schindler,C.E., Chauvot de Beauchêne,I. and Zacharias,M. (2015) A web interface for easy flexible protein-protein docking with ATTRACT. *Biophys. J.*, **108**, 462–465.
27. Xue,L.C., Rodrigues,J., Dobbs,D., Honavar,V. and Bonvin,A. (2017) Template-based protein-protein docking exploiting pairwise interfacial residue restraints. *Brief. Bioinform.*, **18**, 458–466.
28. Kundrotas,P.J., Zhu,Z., Janin,J. and Vakser,I.A. (2012) Templates are available to model nearly all complexes of structurally characterized proteins. *PNAS*, **109**, 9438–9441.

29. Rodrigues,J.P., Melquiond,A.S., Karaca,E., Trellet,M., van Dijk,M., van Zundert,G.C., Schmitz,C., de Vries,S.J., Bordogna,A., Bonati,L. *et al.* (2013) Defining the limits of homology modeling in information-driven protein docking. *Proteins*, **81**, 2119–2128.
30. Yu,J. and Guerois,R. (2016) PPI4DOCK: large scale assessment of the use of homology models in free docking over more than 1000 realistic targets. *Bioinformatics*, **32**, 3760–3767.
31. Bohnuud,T., Luo,L., Wodak,S.J., Bonvin,A.M., Weng,Z., Vajda,S., Schueler-Furman,O. and Kozakov,D. (2017) A benchmark testing ground for integrating homology modeling and protein docking. *Proteins*, **85**, 10–16.
32. Kundrotas,P.J., Anishchenko,I., Dauzhenka,T., Kotthoff,I., Mnevets,D., Copeland,M.M. and Vakser,I.A. (2018) Dockground: a comprehensive data resource for modeling of protein complexes. *Protein Sci.*, **27**, 172–181.
33. Hopf,T.A., Schärfe,C.P., Rodrigues,J.P., Green,A.G., Kohlbacher,O., Sander,C., Bonvin,A.M. and Marks,D.S. (2014) Sequence co-evolution gives 3D contacts and structures of protein complexes. *eLife*, **3**, e03430.
34. Zeng,H., Wang,S., Zhou,T., Zhao,F., Li,X., Wu,Q. and Xu,J. (2018) ComplexContact: a web server for inter-protein contact prediction using deep learning. *Nucleic Acids Res.*, **46**, W432–W437.
35. Yang,J., Anishchenko,I., Park,H., Peng,Z., Ovchinnikov,S. and Baker,D. (2020) Improved protein structure prediction using predicted interresidue orientations. *PNAS*, **117**, 1496–1503.
36. Quignot,C., Granger,P., Chacón,P., Guerois,R. and Andreani,J. (2021) Atomic-level evolutionary information improves protein-protein interface scoring. *Bioinformatics*, doi:10.1093/bioinformatics/btab254.
37. Andreani,J., Faure,G. and Guerois,R. (2013) InterEvScore: a novel coarse-grained interface scoring function using a multi-body statistical potential coupled to evolution. *Bioinformatics*, **29**, 1742–1749.
38. Dong,G.Q., Fan,H., Schneidman-Duhovny,D., Webb,B. and Sali,A. (2013) Optimized atomic statistical potentials: assessment of protein interfaces and loops. *Bioinformatics*, **29**, 3158–3166.
39. Alford,R.F., Leaver-Fay,A., Jeliazkov,J.R., O'Meara,M.J., DiMaio,F.P., Park,H., Shapovalov,M.V., Renfrew,P.D., Mulligan,V.K., Kappel,K. *et al.* (2017) The Rosetta all-atom energy function for macromolecular modeling and design. *J. Chem. Theory Comput.*, **13**, 3031–3048.
40. Lensink,M.F., Velankar,S., Kryshtafovych,A., Huang,S.Y., Schneidman-Duhovny,D., Sali,A., Segura,J., Fernandez-Fuentes,N., Viswanath,S., Elber,R. *et al.* (2016) Prediction of homoprotein and heteroprotein complexes by protein docking and template-based modeling: a CASP-CAPRI experiment. *Proteins*, **84**(Suppl. 1), 323–348.
41. Yu,J., Andreani,J., Ochsenbein,F. and Guerois,R. (2017) Lessons from (co-)evolution in the docking of proteins and peptides for CAPRI Rounds 28–35. *Proteins*, **85**, 378–390.
42. Nadaradjane,A.A., Quignot,C., Traoré,S., Andreani,J. and Guerois,R. (2020) Docking proteins and peptides under evolutionary constraints in critical assessment of prediction of interactions rounds 38 to 45. *Proteins*, **88**, 986–998.
43. Vreven,T., Moal,I.H., Vangone,A., Pierce,B.G., Kastritis,P.L., Torchala,M., Chaleil,R., Jiménez-García,B., Bates,P.A., Fernandez-Recio,J. *et al.* (2015) Updates to the integrated protein-protein interaction benchmarks: docking benchmark version 5 and affinity benchmark version 2. *J. Mol. Biol.*, **427**, 3031–3041.
44. Postic,G., Marcoux,J., Reys,V., Andreani,J., Vandenbrouck,Y., Bousquet,M.P., Mouton-Barbosa,E., Cianférani,S., Burlet-Schiltz,O., Guerois,R. *et al.* (2020) Probing protein interaction networks by combining MS-based proteomics and structural data integration. *J. Proteome Res.*, **19**, 2807–2820.
45. Postic,G., Andreani,J., Marcoux,J., Reys,V., Guerois,R., Rey,J., Mouton-Barbosa,E., Vandenbrouck,Y., Cianferani,S., Burlet-Schiltz,O. *et al.* (2021) Proteo3Dnet: a web server for the integration of structural information with interactomics data. *Nucleic Acids Res.*, gkab332.
46. Steinegger,M., Meier,M., Mirdita,M., Vöhringer,H., Haunsberger,S.J. and Söding,J. (2019) HH-suite3 for fast remote homology detection and deep protein annotation. *BMC Bioinformatics*, **20**, 473.
47. Liang,S., Zheng,D., Zhang,C. and Standley,D.M. (2011) Fast and accurate prediction of protein side-chain conformations. *Bioinformatics*, **27**, 2913–2914.
48. Karami,Y., Guyon,F., De Vries,S. and Tufféry,P. (2018) DaReUS-Loop: accurate loop modeling using fragments from remote or unrelated proteins. *Sci. Rep.*, **8**, 13673.
49. Karami,Y., Rey,J., Postic,G., Murail,S., Tufféry,P. and de Vries,S.J. (2019) DaReUS-Loop: a web server to model multiple loops in homology models. *Nucleic Acids Res.*, **47**, W423–W428.
50. Rose,A.S., Bradley,A.R., Valasatava,Y., Duarte,J.M., Prlic,A. and Rose,P.W. (2018) NGL viewer: web-based molecular graphics for large complexes. *Bioinformatics*, **34**, 3755–3758.
51. Cong,Q., Anishchenko,I., Ovchinnikov,S. and Baker,D. (2019) Protein interaction networks revealed by proteome coevolution. *Science*, **365**, 185–189.
52. Green,A.G., Elhabashy,H., Brock,K.P., Maddamsetti,R., Kohlbacher,O. and Marks,D.S. (2021) Large-scale discovery of protein interactions at residue resolution using co-evolution calculated from genomic sequences. *Nat. Commun.*, **12**, 1396.
53. Quignot,C., Rey,J., Yu,J., Tufféry,P., Guerois,R. and Andreani,J. (2018) InterEvDock2: an expanded server for protein docking using evolutionary and biological information from homology models and multimeric inputs. *Nucleic Acids Res.*, **46**, W408–W416.