



HAL
open science

FTDMP: a framework for protein-protein, protein-DNA and protein-RNA docking and scoring

Kliment Olechnovič, Rita Banciul, Justas Dapkūnas, Česlovas Venclovas

► To cite this version:

Kliment Olechnovič, Rita Banciul, Justas Dapkūnas, Česlovas Venclovas. FTDMP: a framework for protein-protein, protein-DNA and protein-RNA docking and scoring. *Proteins - Structure, Function and Bioinformatics*, In press, 10.1002/prot.26792 . hal-04879411

HAL Id: hal-04879411

<https://hal.science/hal-04879411v1>

Submitted on 10 Jan 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

FTDMP: a framework for protein-protein, protein-DNA and protein-RNA docking and scoring

Kliment Olechnovič^{1,2,*}, Rita Banciul^{1,*}, Justas Dapkūnas¹ and Česlovas Venclovas¹

¹*Institute of Biotechnology, Life Sciences Center, Vilnius University, Saulėtekio 7, LT-10257 Vilnius, Lithuania*

²*Univ. Grenoble Alpes, CNRS, Grenoble INP, LJK, 38000 Grenoble, France*

* Contributed equally

Short title: FTDMP: a framework for biomolecular docking

Keywords: Protein-protein interactions, protein complexes, docking, scoring, model quality assessment

Corresponding author:

Česlovas Venclovas
Institute of Biotechnology, Life Sciences Center
Vilnius University
Saulėtekio 7,
LT-10257 Vilnius, Lithuania

Phone: +370-5-223-4368

E-mail: ceslovas.venclovas@bti.vu.lt

Abstract

FTDMP is a software framework for biomolecular docking and scoring. It can perform docking of subunits containing one or more protein, DNA or RNA chains, followed by subsequent scoring of the resulting models. FTDMP can also be used for ranking of user-provided models of biomolecular complexes, generated by any structure prediction method. FTDMP evaluates models according to the consensus-based method VoroIF-jury, which combines individual scores derived from the Voronoi tessellation of biomolecular structures. In addition to the default scoring mode, FTDMP can easily adopt additional scores, thus it may be used as a tool to assess newly developed scoring functions. FTDMP was evaluated during blind testing in recent CAPRI experiments and using protein-protein, protein-DNA and protein-RNA docking benchmarks. It proved to be a useful tool for different research tasks, related to modeling biomolecular interactions. The software, cleaned docking benchmarks and benchmarking results are available at <https://bioinformatics.lt/software/ftdmp/>.

Introduction

To perform their functions, proteins often interact with various biomolecules, including other proteins, peptides, and nucleic acids. Knowledge of the three-dimensional (3D) structures of the resulting complexes is a prerequisite for understanding their molecular mechanism. In recent years, these structures are more and more often predicted using computational methods. In particular, the field was revolutionized by the advent of deep learning-based structure prediction methods, AlphaFold¹ and AlphaFold-Multimer², offering unprecedented accuracy. Most recently, newly developed AlphaFold 3³, RoseTTAFold-NA⁴ and RoseTTAFold All-Atom⁵ introduced the features to also predict protein complexes with nucleic acids and small molecules.

Despite these advancements, significant challenges remain. CASP^{6,7} and CAPRI⁸ experiments, which have monitored progress in predicting protein interactions for many years, highlight persistent difficulties in accurately predicting certain interactions. Even with the introduction of AlphaFold, modeling of antibody-antigen or host-pathogen interactions remains difficult^{2,7-9}. The accuracy of AlphaFold deteriorates with the increasing number of chains and residues in the complex^{7,10}. Protein-nucleic acid complexes are in general modeled with lower success rates compared to protein-protein complexes³. In some cases, these limitations may be complemented by alternative methods, such as docking. Docking methods can help to tackle cases where AlphaFold and similar methods encounter difficulties. Additionally, docking can be used to generate synthetic data, for example, by generating structure models of varying accuracy for training machine learning methods¹¹.

Molecular docking consists of two major stages: (1) construction of structural models and (2) estimating their accuracy (scoring). It should be emphasized that the scoring stage is important not only for docking. While most of the modern modeling methods provide accuracy self-estimation, these evaluations are often inconsistent across different methods

and cannot be directly compared. Independent scoring methods are therefore necessary to reliably select the best models from a pool of diverse predictions.

Our group ('Venclovas') has participated in CAPRI experiments since 2016, initially relying on template-based methods to predict the structures of protein complexes^{12,13}. For subsequent CAPRI Rounds (47-52, see Supplementary Table S1) we continued using homology modeling, sometimes in combination with docking. However, the release of AlphaFold-Multimer represented a clear watershed in our methodology. We observed that AlphaFold-Multimer significantly outperforms template-based approaches, and can accurately predict structures even when templates are not available¹⁴. Thus, during the post-AlphaFold CAPRI Rounds (53-55) we focused our efforts on the development of solutions for the cases when AlphaFold and other similar methods encounter difficulties.

Here we present FTDMP (Flexible Toolkit Dedicated to Multimeric Predictions), a software framework for biomolecular docking and scoring. FTDMP can be used to perform docking and subsequent scoring, or to rank user-provided models from various sources. Developed and tested during the CASP15-CAPRI experiment¹⁴, FTDMP enabled our group to achieve top results in both multimeric structure prediction^{7,8} and CAPRI scoring categories⁸. Besides its testing in recent CASP and CAPRI rounds, we evaluated FTDMP using protein-protein^{15,16}, protein-DNA¹⁷ and protein-RNA¹⁸⁻²¹ docking benchmarks. In addition, we tested FTDMP as a framework for evaluation of new scoring functions.

Materials and methods

Description of the FTDMP framework

FTDMP is a framework for biomolecular docking and scoring. Its default pipeline (Fig. 1) consists of:

- generating a large number of docking models from given subunit structures²²,
- filtering them using fast interface-focused scoring²³,
- ranking the remaining models using a consensus method called VoroIF-jury¹⁴,
- relaxing a small number of models to get rid of clashes and improve interface interactions²⁴,
- re-ranking the relaxed models by VoroIF-jury again.

By default FTDMP uses a number of scoring methods based on Voronoi tessellation to rank the models¹⁴, and also allows easy introduction of custom scoring functions.

In addition to the docking and scoring pipelines (ftdmp-all and ftdmp-qa-all, respectively), FTDMP offers useful helper scripts, e.g., to split multi-model PDB files to files containing separate structures, to renumber biomolecular structures in PDB format according to given sequences, to generate structure files from docking results, to relax structures using OpenMM, etc.

Docking methods

FTDMP accepts two subunits for docking, static and mobile, which can be composed of one or several chains in PDB format. The mobile subunit is docked to the static one, and a large diversity of binding poses is generated aiming to exhaustively explore the possible conformational space. By default, FTDMP uses FTDock software for the conformational sampling²². FTDock was modified by adding parallelization to employ multiple processors. Initial filtering of conformations was removed, and all generated conformations are recorded

for subsequent scoring. In addition to FTDock, using other docking software, such as Hex²⁵ or SAM²⁶ is also possible with FTDMP.

Scoring methods

FTDMP uses scoring methods, mostly based on VoroMQA²³ and VoroIF-GNN¹¹, as described previously¹⁴. VoroIF-GNN is used only for the protein-protein interfaces. VoroMQA was initially developed for assessment of protein structures. It has been adapted for scoring protein-nucleic acid interfaces as follows. First, protein and nucleic acid atom types were mapped to a residue-agnostic set of atom types, used in Knodle²⁷ (Supplementary Table S2). The original VoroMQA was derived from the sums of contact areas of interactions between protein atom types, computed from a non-redundant set of PDB protein structures²³. For the modified (“generalized”) VoroMQA potential, we used the same input structures, but summed the contact areas by more general types of atom-atom interactions described using the Knodle atom types. The obtained sums were then converted to contact potential values using the same algorithm as described in the original VoroMQA publication²³. The resulting “generalized” VoroMQA is then able to evaluate structures, consisting of proteins, DNA and RNA.

Model scoring and ranking is done by the VoroIF-jury procedure¹⁴ which is based on the consensus between rankings according to multiple scoring functions. VoroIF-jury first computes rankings using several methods. Then it selects top models according to each ranking, and calculates the average interface CAD-score²⁸ value for every model with all the other selected models as a consensus measure – the VoroIF-jury score. If multiple scoring methods favor similar models, these models will get a high consensus score. Therefore, the models are finally ranked according to the VoroIF-jury score. For every model in the final ranking, its maximum similarity (interface CAD-score) to the higher ranked models is

recorded. This enables removing redundant entries from the ranking list by applying a threshold to these maximum similarity values, effectively performing a greedy clustering of models. In the FTDMP software, the maximum similarity threshold is adjustable, with the default value of 0.9. Full details of the VoroIF-jury algorithm are reported in our previous publication¹⁴.

Docking benchmarks

To test the performance of FTDMP in protein-protein docking, we used the most recent protein-protein benchmark v.5.5^{15,16}, which contains 252 cases (158 rigid-body, 59 medium difficulty, 35 difficult), 67 of which are antibody-antigen complexes. The protein-DNA docking benchmark created by the Bonvin group and containing 47 complexes was used to evaluate protein-DNA docking¹⁷. The dataset for protein-RNA docking was taken from three different benchmarks¹⁸⁻²⁰, which were combined excluding redundant cases²¹. We selected only the complexes that had unbound both protein and RNA structures in the Protein Data Bank (PDB)²⁹. This yielded a total of 42 cases.

Only the complexes' PDB IDs and chain IDs were taken from the original benchmarks. The structures of complexes from all the benchmarks were downloaded directly from the PDB²⁹ and only the chains given in the benchmarks' tables were selected. To facilitate model evaluation, all the structures, including the complex, constituting bound and unbound monomers were renumbered according to the sequences in the FASTA file of the complex obtained from PDB, since DockQ, RMSD, and CAD-score calculations (see below) require identical model and target residue numbering. For several cases, we checked if there are alternative binding modes using the PPI3D web server³⁰. When alternative interfaces are present, docking models need to be evaluated against all of them, selecting the highest

evaluation scores. The cases with the alternatives together with a few corrected inconsistencies and errors are listed in Supplementary Table S3.

The B-form DNA structure for the protein-DNA docking was modelled using the ChimeraX software³¹. Modified nucleotides in the DNA sequence were changed to the corresponding canonical nucleotides before ChimeraX modeling, since ChimeraX generates the DNA without modifications and the default scoring methods in FTDMP cannot handle modified nucleotides. The modified nucleotides of both bound and unbound RNA structures were also changed to the corresponding canonical nucleotides.

Evaluation of model accuracy

Protein-protein docking results were evaluated by comparing resulting structure models to corresponding experimental structures and classifying the models as incorrect, acceptable, medium, or high accuracy, according to the CAPRI criteria³², estimated using the DockQ software³³. The protein-nucleic acid docking results were evaluated using ligand RMSD (L-RMSD), where receptor was the protein, and ligand was DNA or RNA. A model having L-RMSD ≤ 10 Å was considered as acceptable. The success rate was defined as the percent of complexes for which at least one prediction of acceptable or higher accuracy was found among the top N models.

The interface and binding site CAD-score values were calculated for all docking models²⁸. Interface CAD-score values were used to test if the pre-filtering of docking poses filters out any accurate models, and for an alternative definition of docking success, where models having interface CAD-score > 0.1 were considered as acceptable. Binding site CAD-score was employed to evaluate the predictions of epitopes in the cases of antibody-antigen complexes, and DNA and RNA binding sites. The binding site (epitope) prediction was considered acceptable when the binding site CAD-score was above 0.25.

Evaluation of FTDMP using docking benchmarks

Both unbound-unbound docking and bound-bound docking was performed using the docking benchmarks. This allowed us to evaluate the scoring of models without the limitations of rigid-body docking in cases where conformational changes occur during complex formation. Comparing unbound-unbound and bound-bound docking results also provides information on the effect that conformational changes have on identifying a good model.

Docking generates around 40,000 models with default FTDMP parameters. Prefiltering of these models is then performed by selecting top 3,000 models according to VoromQA interaction interface energy score²³. These top models are then ranked using different interface-focused scoring functions. For each scoring function, the top 200 models are selected, then the union of all the selected models and all their ranks is passed to the VoroIF-jury algorithm¹⁴. Then the top 200 models ranked by VoroIF-jury are relaxed using OpenMM²⁴. The relaxed models are then re-evaluated again with VoroIF-jury to generate the final ranked list of the best models.

Evaluation of FTDMP scoring

Model selection by FTDMP using the VoroIF-jury algorithm and default scoring functions was evaluated for the CAPRI target T188 (Round 53), during the CASP15-CAPRI experiment (CAPRI Round 54), and in the subsequent CAPRI Round 55. During these CAPRI rounds we had to select the best model from sets of diverse structure models generated by different methods.

In addition to the default scoring functions, implemented in FTDMP, the framework can use custom scoring methods. We tested this feature for protein-RNA docking using

specialized DARS-RNP and QUASI-RNP potentials³⁴ and 42 protein-RNA docking benchmark cases.

Results and Discussion

Protein-protein docking

Protein-protein docking capabilities of FTDMP were evaluated using 252 protein complexes from protein-protein docking benchmark v5.5^{15,16}, which includes 67 antibody-antigen complexes. The results are presented in Fig. 2, A-B, and Supplementary Fig. S1, A-B. Considering only the top prediction, according to the CAPRI criteria the success rate for unbound docking was 12%, compared to 80% for bound docking. The success rate is considerably higher when top 10 or top 100 predictions are taken into account (for unbound-unbound docking 26% and 60%, for bound-bound docking 60% and 95%). A slightly lower success rate was observed for antibody-antigen complexes: 4 (6%), 14 (21%), and 32 (48%) acceptable or better models for the top 1, top 10, and top 100 predictions, respectively. Interestingly, the success of bound-bound antibody-antigen docking was similar to the success rate in the whole protein-protein docking benchmark. This difference might be attributed to the antibody and antigen conformational changes upon binding, which are quite frequent¹⁶. The overall observed difference between unbound-unbound and bound-bound docking illustrates that even the slightest conformational changes upon binding have impact on rigid-body docking result.

To compare the FTDMP results with other docking programs, we used the previously reported results^{15,16}. For 52 benchmark v5 cases FTDMP performs better than the rigid-body *ab initio* docking methods pyDock and ZDOCK (Supplementary Table S4)¹⁵. FTDMP also obtained more accurate predictions than the information-driven docking method HADDOCK, which used predicted interface residues as restraints. Notably, a failure in predicting interface residues lead to incorrect HADDOCK models in several cases¹⁵. FTDMP is slightly outperformed by SwarmDock, which uses a flexible docking algorithm (Supplementary Table S4)¹⁵. For 67 antibody-antigen docking cases, FTDMP results are again better than

ZDOCK's, but slightly lower compared to ClusPro (Supplementary Table S4)¹⁶. Yet, for both ZDOCK and ClusPro, the docking was restrained to favor models with antibody CDR loops at the interface, whereas no restraints were applied during docking with FTDMP.

Protein-DNA docking

The FTDMP framework was evaluated on the protein-DNA docking benchmark containing 47 cases, of which 13 are classified as easy, 22 as intermediate and 12 as difficult¹⁷. The results are shown in Fig. 2, C, and Supplementary Fig. S1, C. The unbound-unbound docking had success rates of 9% (4) and 13% (6) for the top 1 and top 10 models, respectively. Compared to pyDockDNA³⁵, FTDMP success rate is higher than for top 1 (6%), but lower for the top 10 (16.4%) predictions (Supplementary Table S4). When comparing the *ab initio* docking results with no prior knowledge about the interface residues, FTDMP outperforms the HDOCK server which has success rates of 4.3% for top 1 and 8.7% for top 10 models³⁶.

To analyze how detrimental the conformational changes are to the protein-DNA rigid-body docking, we performed bound-bound docking for the same 47 cases. The bound-bound docking success rates were 83% and 94% for the top 1 and top 10 models, respectively, which is almost ten times higher than the unbound-unbound docking results. The high success of bound-bound docking demonstrates the ability of the FTDMP scoring to identify good models of protein-DNA complexes if they are present in the model set. Moreover, the huge difference between bound and unbound docking exemplifies the importance of flexibility for the protein-DNA docking. Similar results were obtained previously using HADDOCK. While the unbound-unbound protein-DNA docking results were similar to our rigid-body docking results, significant improvement was observed when flexibility was incorporated into the docking protocols^{37, 38}.

Protein-RNA docking

Protein-RNA docking was tested using 42 protein-RNA complexes (21 rigid-body, 15 medium difficulty, 6 difficult)¹⁸⁻²⁰. Both unbound-unbound and bound-bound docking was performed and the results are shown in Figure 2, D, and Supplementary Figure S1, D. The unbound docking success rate was 24 % and 26%, the bound docking success was 62% and 74% for the top 1 and top 10 models, respectively.

We compared these results to two RNA docking servers (Supplementary Table S4). Compared to P3DOCK²¹, FTDMP selected better top models, but the test set contained only 12 overlapping cases. Out of 32 protein-RNA complexes that match our benchmark set and considering top 3 predictions, NPDock³⁹ generates accurate models for 16 (bound docking) and 15 (unbound docking) cases, while success rate of FTDMP is 22 for bound and 11 for unbound docking. Thus, FTDMP is better at selecting bound-bound models and a bit worse at selecting unbound-unbound models.

The overall protein-RNA docking results demonstrate that protein-RNA docking is a challenging task. The FTDMP framework can generate reasonably accurate predictions for complexes in which RNA and protein undergo negligible conformational changes upon binding. However, it is almost impossible to dock protein and RNA if medium or large conformational changes occur.

Prediction of binding sites

While the results of structure prediction for biomolecular complexes by docking are at best satisfactory, it has been observed that the binding sites are predicted reasonably well in the docking models⁴⁰⁻⁴³. Oftentimes a model with well-predicted binding sites might be sufficient for practical applications. Therefore, we decided to check if FTDMP can predict

epitopes of the antibody-antigen interactions, as well as DNA and RNA binding sites in the protein molecules. The success for binding site prediction was defined as having a binding site CAD-score >0.25 in the protein (antigen) molecule. The results are presented in Fig. 3, and detailed results are given in Supplementary Table S5.

For the antibody-antigen complexes, FTDMP managed to correctly predict the epitope for 25% of the cases, considering only the top 1 prediction (52% and 93% of cases for top 10 and top 100 predictions, respectively). The results are slightly better for protein-nucleic acid binding sites. For protein-RNA complexes, 40% of top 1 and 57% of top 10 predictions were successful. Moreover, the binding site was correctly identified for 62% of easy protein-RNA docking cases in the top docking model. The protein-DNA binding sites are predicted well in 30% of the top ranked models and in 55% cases when considering top 10 models. Fig. 4 demonstrates some of the most successful examples, where the nucleic acid molecule is bound to the actual binding site although its orientation is incorrect. Interestingly, such a situation was also observed for CAPRI target T188 (Fig. 4, Supplementary Table S6). For this CAPRI target we used FTDMP to dock the provided bound conformation of DNA to the AlphaFold-Multimer model of the protein dimer.

Performance of FTDMP in the CAPRI scoring challenge

We used FTDMP in the scoring challenge of the post-AlphaFold CAPRI Rounds 53-55, including the CASP15-CAPRI experiment, in which our group obtained the best scoring results⁸. During these rounds, we selected models of acceptable quality for 26 of 41 targets using FTDMP and VoroIF-jury (Supplementary Table S7). A slight relationship between the target difficulty and success of FTDMP scoring can be observed. For example, 9 of 15 failed cases were antibody-antigen targets, which are harder to predict even using AlphaFold-Multimer². Moreover, when we tried out FTDMP scoring on the targets of pre-AlphaFold

CAPRI rounds 47-49, it usually failed to choose models of at least acceptable quality. However, for these targets there were at best just a few good models in the scoring sets. VoroIF-jury depends on the consensus similarity among models that are ranked highly by multiple scoring functions. Limited availability of accurate models reduces the likelihood of achieving a strong consensus among the top-ranked models in multiple rankings.

Testing new scoring functions using FTDMP

In addition to the default scoring functions, the FTDMP framework can also utilize additional scoring methods. In such case, a script that provides one or more scores for a structure can be given for FTDMP, which will then rank the models based not only on the default scores, but also on the newly provided scores. In addition to the usage of the individual scores, the consensus of all the scores might be estimated by VoroIF-jury.

We have tested this FTDMP option with the freely available DARS-RNP and QUASI-RNP potentials for protein-RNA complexes³⁴. DARS-RNP and QUASI-RNP were used for ranking models of the 42 protein-RNA docking benchmark cases in two ways: together with FTDMP built-in scoring functions and without them. The results are given in Fig. 5 and Supplementary Fig. S2 for both unbound-unbound and bound-bound docking. The DARS-RNP and QUASI-RNP alone demonstrate lower success rates than FTDMP scoring. However, adding DARS-RNP and QUASI-RNP scoring functions to the VoroIF-jury slightly increases success rates for the bound-bound docking, especially when considering only the top result (from 62% to 67%). Adding these additional scoring methods did not noticeably influence the success rate of the unbound-unbound docking, but this might be attributed to the limitations of rigid-body docking, not scoring. Thus, we demonstrated that including additional model accuracy estimation methods to the FTDMP scoring may have a positive impact on the model selection.

The ability to easily introduce additional scoring methods makes FTDMP framework a useful tool to test newly developed scoring functions. To facilitate the testing, we provide all the FTDMP docking results for the docking benchmarks with tens to hundreds of thousands of models presented as translation and rotation vectors in a tabular format. Based on these tables, structure models can be generated and used to test scoring functions, skipping the docking step and saving valuable computing time.

When preparing the models, we optimized the docking parameters aiming to generate a reasonable number of structures: small enough to not substantially increase the computing time, but large enough to include accurate models. We tried to make sure that all docking sets contain at least one high accuracy model. To test this, we calculated interface CAD-score value²⁸ for every docking model. The bound-bound docking tables include at least one high accuracy model with interface CAD-score >0.5 for every complex with negligible exceptions, where medium accuracy models are still available. As for the unbound-unbound docking tables, they contain less accurate models, which is expected because of conformational changes. 227 of 252 protein-protein complexes have acceptable models (interface CAD-score > 0.1). The docking results tables of 39 of 42 protein-RNA targets and 43 of 47 protein-DNA targets include at least one acceptable model. The fact, that for unbound-unbound docking some of the benchmark cases have only models of acceptable accuracy or even all incorrect models, is an indication that not all complexes can be predicted by rigid-body docking due to the substantial conformational changes upon binding. Nonetheless, the FTDMP docking tables contain sufficiently accurate models to use them for testing new scoring functions.

Conclusions

FTDMP is a software framework for biomolecular docking and scoring. It can be used for protein-protein, protein-DNA or protein-RNA docking, as well as for scoring structural

models of biomolecular complexes generated by diverse methods. In addition to that, FTDMP provides an easy-to-use environment for testing new scoring functions for interaction interfaces. The capabilities of FTDMP were showcased both during recent CAPRI rounds and through extensive testing on docking benchmarks, demonstrating its usefulness for various tasks associated with modeling biomolecular interactions.

Software availability

FTDMP software is available at <https://github.com/kliment-olechnovic/ftdmp>.

Docking benchmarks and docking tables are available in Zenodo:

<https://doi.org/10.5281/zenodo.10517524>.

The results of FTDMP testing using docking benchmarks are available in Zenodo:

<https://doi.org/10.5281/zenodo.12804208>.

Acknowledgments

We would like to thank Marc Lensink for the help with CAPRI scoring sets.

Funding

Research Council of Lithuania (grants S-MIP-21-35 and S-IMPRESSU-24-5).

References

1. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Žídek A, Potapenko A, Bridgland A, Meyer C, Kohl SAA, Ballard AJ, Cowie A, Romera-Paredes B, Nikolov S, Jain R, Adler J, Back T, Petersen S, Reiman D, Clancy E, Zielinski M, Steinegger M, Pacholska M, Berghammer T, Bodenstein S, Silver D, Vinyals O, Senior AW, Kavukcuoglu K, Kohli P, Hassabis D. Highly accurate protein structure prediction with AlphaFold. *Nature* 2021;596(7873):583–589.
2. Evans R, O'Neill M, Pritzel A, Antropova N, Senior A, Green T, Žídek A, Bates R, Blackwell S, Yim J, Ronneberger O, Bodenstein S, Zielinski M, Bridgland A, Potapenko A, Cowie A, Tunyasuvunakool K, Jain R, Clancy E, Kohli P, Jumper J, Hassabis D. Protein complex prediction with AlphaFold-Multimer. *bioRxiv* 2022:2021.10.04.463034.
3. Abramson J, Adler J, Dunger J, Evans R, Green T, Pritzel A, Ronneberger O, Willmore L, Ballard AJ, Bambrick J, Bodenstein SW, Evans DA, Hung C-C, O'Neill M, Reiman D, Tunyasuvunakool K, Wu Z, Žemgulytė A, Arvaniti E, Beattie C, Bertolli O, Bridgland A, Cherepanov A, Congreve M, Cowen-Rivers AI, Cowie A, Figurnov M, Fuchs FB, Gladman H, Jain R, Khan YA, Low CMR, Perlin K, Potapenko A, Savy P, Singh S, Stecula A, Thillaisundaram A, Tong C, Yakneen S, Zhong ED, Zielinski M, Žídek A, Bapst V, Kohli P, Jaderberg M, Hassabis D, Jumper JM. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* 2024;630(8016):493–500.
4. Baek M, McHugh R, Anishchenko I, Jiang H, Baker D, DiMaio F. Accurate prediction of protein-nucleic acid complexes using RoseTTAFoldNA. *Nat Methods* 2024;21(1):117–121.
5. Krishna R, Wang J, Ahern W, Sturmfels P, Venkatesh P, Kalvet I, Lee GR, Morey-Burrows FS, Anishchenko I, Humphreys IR, McHugh R, Vafeados D, Li X, Sutherland GA, Hitchcock A, Hunter CN, Kang A, Brackenbrough E, Bera AK, Baek M, DiMaio F, Baker D. Generalized biomolecular modeling and design with RoseTTAFold All-Atom. *Science* 2024;384(6693):eadl2528.
6. Kryshchak A, Schwede T, Topf M, Fidelis K, Moult J. Critical assessment of methods of protein structure prediction (CASP)-Round XV. *Proteins* 2023;91(12):1539–1549.
7. Ozden B, Kryshchak A, Karaca E. The impact of AI-based modeling on the accuracy of protein assembly prediction: Insights from CASP15. *Proteins* 2023;91(12):1636–1657.
8. Lensink MF, Brysbaert G, Raouraoua N, Bates PA, Giulini M, Honorato RV, Noort C van, Teixeira JMC, Bonvin AMJJ, Kong R, Shi H, Lu X, Chang S, Liu J, Guo Z, Chen X, Morehead A, Roy RS, Wu T, Giri N, Quadir F, Chen C, Cheng J, Del Carpio CA, Ichiishi E, Rodriguez-Lumbreras LA, Fernandez-Recio J, Harmalkar A, Chu L-S, Canner S, Smanta R, Gray JJ, Li H, Lin P, He J, Tao H, Huang S-Y, Roel-Touris J, Jimenez-Garcia B, Christoffer CW, Jain AJ, Kagaya Y, Kannan H, Nakamura T, Terashi G, Verburgt JC, Zhang Y, Zhang Z, Fujuta H, Sekijima M, Kihara D, Khan O, Kotelnikov S, Ghani U, Padhorny D, Beglov D, Vajda S, Kozakov D, Negi SS, Ricciardelli T, Barradas-Bautista D, Cao Z, Chawla M, Cavallo L, Oliva R, Yin R, Cheung M, Guest JD, Lee J, Pierce BG, Shor B, Cohen T, Halfon M, Schneidman-Duhovny D, Zhu S, Yin R, Sun Y, Shen Y, Maszota-Zieleniak M, Bojarski KK, Lubecka EA, Marcisz M, Danielsson A, Dziadek L, Gaardlos M, Gieldon A, Liwo A, Samsonov SA, Slusarz R, Zieba K, Sieradzan AK, Czaplowski C, Kobayashi S, Miyakawa Y, Kiyota Y, Takeda-Shitaka M, Olechnovic K, Valancauskas L, Dapkunas J, Venclovas C, Wallner B, Yang L, Hou C, He X, Guo S, Jiang S, Ma X, Duan R, Qui L, Xu X, Zou X, Velankar S, Wodak SJ. Impact of AlphaFold on structure prediction of protein complexes: The CASP15-CAPRI experiment. *Proteins* 2023;91(12):1658–1683.
9. Yin R, Feng BY, Varshney A, Pierce BG. Benchmarking AlphaFold for protein complex modeling reveals accuracy determinants. *Protein Sci* 2022;31(8):e4379.

10. Bryant P, Pozzati G, Zhu W, Shenoy A, Kundrotas P, Elofsson A. Predicting the structure of large protein complexes using AlphaFold and Monte Carlo tree search. *Nat Commun* 2022;13(1):6028.
11. Olechnovič K, Venclovas Č. VoroIF-GNN: Voronoi tessellation-derived protein-protein interface assessment using a graph neural network. *Proteins* 2023;91(12):1879–1888.
12. Dapkūnas J, Olechnovič K, Venclovas Č. Modeling of protein complexes in CAPRI Round 37 using template-based approach combined with model selection. *Proteins* 2018;86 Suppl 1:292–301.
13. Dapkūnas J, Kairys V, Olechnovič K, Venclovas Č. Template-based modeling of diverse protein interactions in CAPRI rounds 38–45. *Proteins* 2020;88(8):939–947.
14. Olechnovič K, Valančauskas L, Dapkūnas J, Venclovas Č. Prediction of protein assemblies by structure sampling followed by interface-focused scoring. *Proteins* 2023;91(12):1724–1733.
15. Vreven T, Moal IH, Vangone A, Pierce BG, Kastritis PL, Torchala M, Chaleil R, Jiménez-García B, Bates PA, Fernandez-Recio J, Bonvin AMJJ, Weng Z. Updates to the Integrated Protein-Protein Interaction Benchmarks: Docking Benchmark Version 5 and Affinity Benchmark Version 2. *J Mol Biol* 2015;427(19):3031–3041.
16. Guest JD, Vreven T, Zhou J, Moal I, Jeliakov JR, Gray JJ, Weng Z, Pierce BG. An expanded benchmark for antibody-antigen docking and affinity prediction reveals insights into antibody recognition determinants. *Structure* 2021;29(6):606–621.e5.
17. Dijk M van, Bonvin AMJJ. A protein-DNA docking benchmark. *Nucleic Acids Res* 2008;36(14):e88.
18. Pérez-Cano L, Jiménez-García B, Fernández-Recio J. A protein-RNA docking benchmark (II): extended set from experimental and homology modeling data. *Proteins* 2012;80(7):1872–1882.
19. Huang S-Y, Zou X. A nonredundant structure dataset for benchmarking protein-RNA computational docking. *J Comput Chem* 2013;34(4):311–318.
20. Nithin C, Mukherjee S, Bahadur RP. A non-redundant protein-RNA docking benchmark version 2.0. *Proteins* 2017;85(2):256–267.
21. Zheng J, Hong X, Xie J, Tong X, Liu S. P3DOCK: a protein-RNA docking webserver based on template-based and template-free docking. *Bioinformatics* 2020;36(1):96–103.
22. Gabb HA, Jackson RM, Sternberg MJ. Modelling protein docking using shape complementarity, electrostatics and biochemical information. *J Mol Biol* 1997;272(1):106–120.
23. Olechnovič K, Venclovas Č. VoroMQA: Assessment of protein structure quality using interatomic contact areas. *Proteins* 2017;85(6):1131–1145.
24. Eastman P, Swails J, Chodera JD, McGibbon RT, Zhao Y, Beauchamp KA, Wang L-P, Simmonett AC, Harrigan MP, Stern CD, Wiewiora RP, Brooks BR, Pande VS. OpenMM 7: Rapid development of high performance algorithms for molecular dynamics. *PLoS Comput Biol* 2017;13(7):e1005659.
25. Ritchie DW, Kemp GJ. Protein docking using spherical polar Fourier correlations. *Proteins* 2000;39(2):178–194.
26. Ritchie DW, Grudinin S. Spherical polar Fourier assembly of protein complexes with arbitrary point group symmetry. *J Appl Cryst* 2016;49(1):158–167.
27. Kadukova M, Grudinin S. Knodle: A Support Vector Machines-Based Automatic Perception of Organic Molecules from 3D Coordinates. *J Chem Inf Model* 2016;56(8):1410–1419.
28. Olechnovič K, Venclovas Č. Contact Area-Based Structural Analysis of Proteins and Their Complexes Using CAD-Score. *Methods Mol Biol* 2020;2112:75–90.
29. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. The Protein Data Bank. *Nucleic Acids Res* 2000;28(1):235–242.

30. Dapkūnas J, Timinskas A, Olechnovič K, Tomkuvienė M, Venclovas Č. PPI3D: a web server for searching, analyzing and modeling protein-protein, protein-peptide and protein-nucleic acid interactions. *Nucleic Acids Res* 2024;52(W1):W264–W271.
31. Meng EC, Goddard TD, Pettersen EF, Couch GS, Pearson ZJ, Morris JH, Ferrin TE. UCSF ChimeraX: Tools for structure building and analysis. *Protein Sci* 2023;32(11):e4792.
32. Méndez R, Leplae R, De Maria L, Wodak SJ. Assessment of blind predictions of protein-protein interactions: current status of docking methods. *Proteins* 2003;52(1):51–67.
33. Basu S, Wallner B. DockQ: A Quality Measure for Protein-Protein Docking Models. *PLoS One* 2016;11(8):e0161879.
34. Tuszynska I, Bujnicki JM. DARS-RNP and QUASI-RNP: new statistical potentials for protein-RNA docking. *BMC Bioinformatics* 2011;12:348.
35. Rodríguez-Lumbreras LA, Jiménez-García B, Giménez-Santamarina S, Fernández-Recio J. pyDockDNA: A new web server for energy-based protein-DNA docking and scoring. *Front Mol Biosci* 2022;9:988996.
36. Yan Y, Zhang D, Zhou P, Li B, Huang S-Y. HDock: a web server for protein-protein and protein-DNA/RNA docking based on a hybrid strategy. *Nucleic Acids Res* 2017;45(W1):W365–W373.
37. Dijk M van, Bonvin AMJJ. Pushing the limits of what is achievable in protein-DNA docking: benchmarking HADDOCK's performance. *Nucleic Acids Res* 2010;38(17):5634–5647.
38. Kurkcuoğlu Z, Bonvin AMJJ. Pre- and post-docking sampling of conformational changes using ClustENM and HADDOCK for protein-protein and protein-DNA systems. *Proteins* 2020;88(2):292–306.
39. Tuszynska I, Magnus M, Jonak K, Dawson W, Bujnicki JM. NPDock: a web server for protein-nucleic acid docking. *Nucleic Acids Res* 2015;43(W1):W425–430.
40. Dapkūnas J, Olechnovič K, Venclovas Č. Structural modeling of protein complexes: Current capabilities and challenges. *Proteins* 2019;87(12):1222–1232.
41. Dapkūnas J, Olechnovič K, Venclovas Č. Modeling of protein complexes in CASP14 with emphasis on the interaction interface prediction. *Proteins* 2021;89(12):1834–1843.
42. Fernández-Recio J, Totrov M, Abagyan R. Identification of protein-protein interaction sites from docking energy landscapes. *J Mol Biol* 2004;335(3):843–865.
43. Viswanathan R, Fajardo E, Steinberg G, Haller M, Fiser A. Protein-protein binding supersites. *PLoS Comput Biol* 2019;15(1):e1006704.

Figure legends

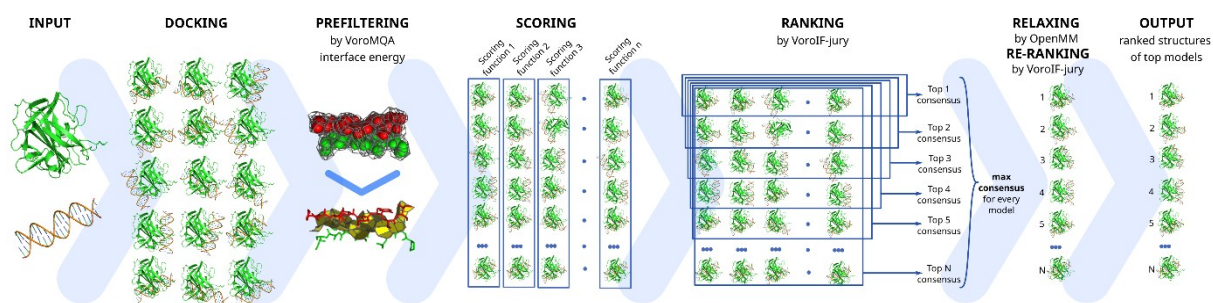


Figure 1. The workflow of docking and scoring using FTDMP

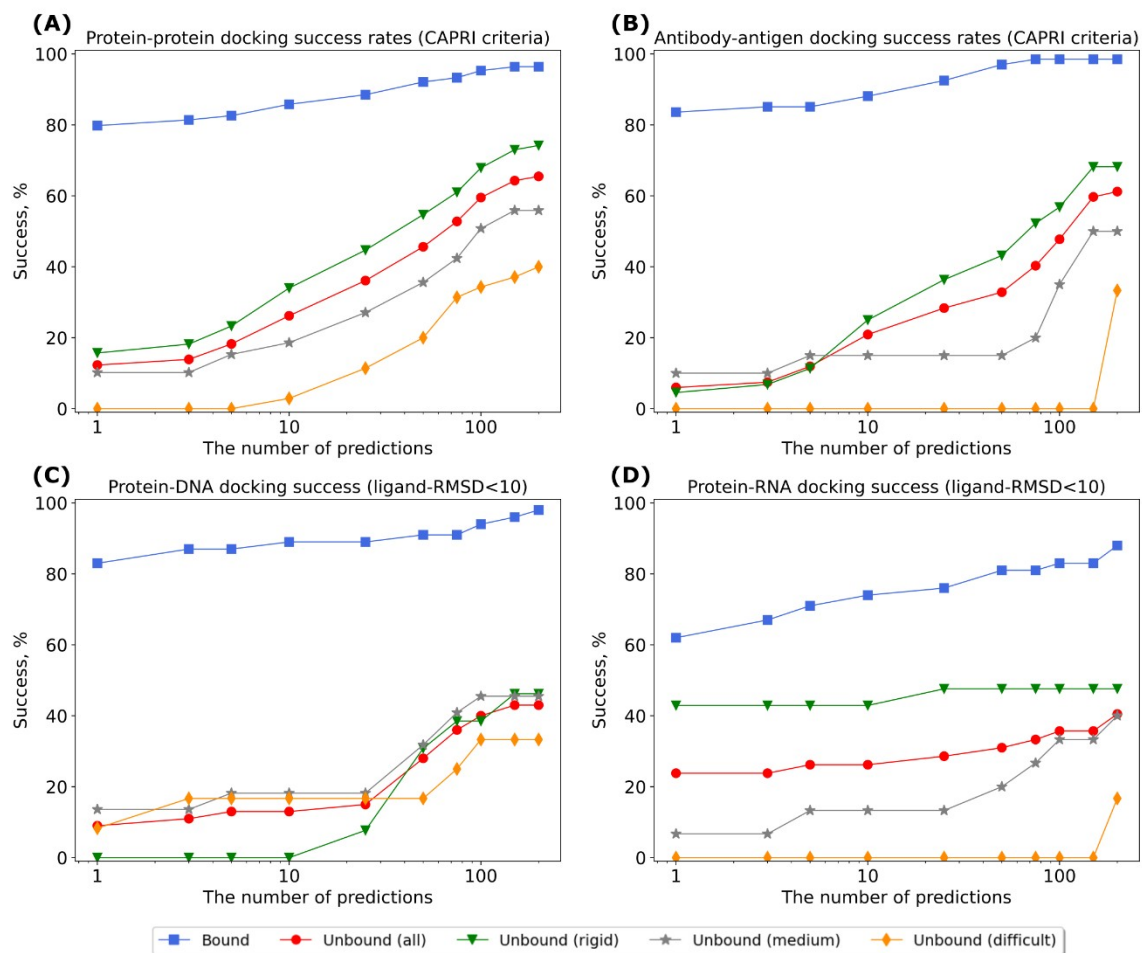


Figure 2. The summary of FTDMP docking results for the protein-protein (A), antibody-antigen (B), protein-DNA (C) and protein-RNA (D) docking benchmarks

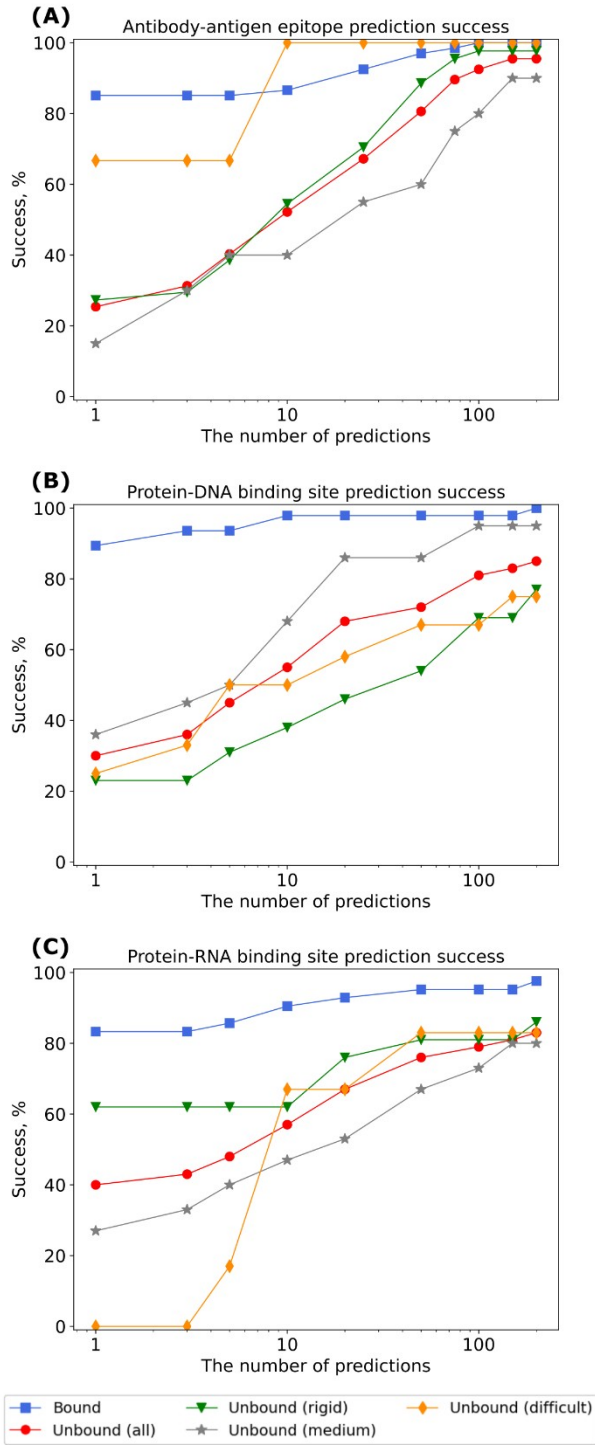
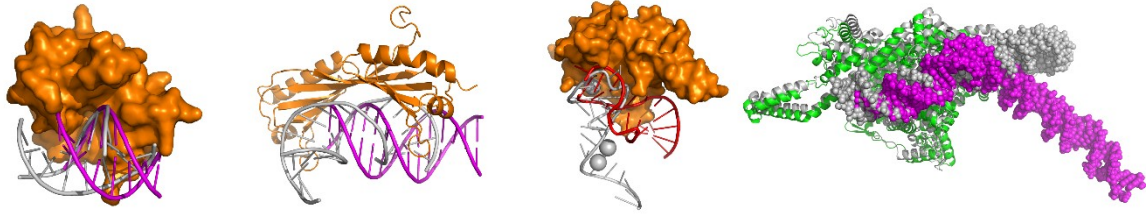


Figure 3. Epitope and nucleic acid binding site predictions using FTDMP for the antibody-antigen (A), protein-DNA (B) and protein-RNA (C) docking benchmarks

(A) 1PT3 ligand-RMSD=19.1 binding site CAD-score=0.72 1QNE ligand-RMSD=20.3 binding site CAD-score=0.52 1JBS ligand-RMSD=16.4 binding site CAD-score=0.49 CAPRI T188 (7QD5) ligand-RMSD=81.9 binding site CAD-score=0.23



(B) 1PT3 target: KFKATGKCKPVRNKKLNNAGKDLGSPVDFRANKLRKKEI... model: KFKATGKCKPVRNKKLNNAGKDLGSPVDFRANKLRKKEI...
 1QNE target: MTDGLEGSMFVLSKHPGIVPTL... model: MTDGLEGSMFVLSKHPGIVPTL...
 1JBS target: AWTCTIQQLNPKTKMFKEDKRLYSQAKAFSSIIHAPIS... model: AWTCTIQQLNPKTKMFKEDKRLYSQAKAFSSIIHAPIS...
 T188 (7QD5) target: NRRHYVELAALTELNRHNSGDI... model: NRRHYVELAALTELNRHNSGDI...

Figure 4. Examples of successful prediction of protein-nucleic acid binding sites in docking models of low accuracy at the structure (A) and sequence (B) level. All the experimental structures are gray, benchmark proteins are orange, CAPRI T188 protein model is green, model DNA is in magenta and model RNA is in red; the binding site residues are marked in the sequences

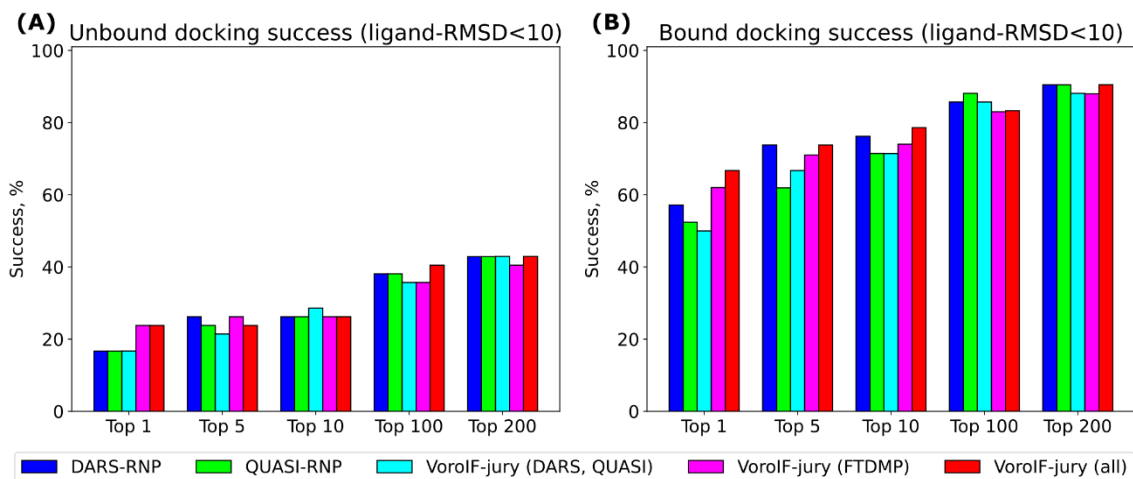


Figure 5. Protein-RNA docking results when using FTDMP with DARS-RNP and QUASI-RNP potentials as additional scoring functions for unbound-unbound (A) and bound-bound (B) docking