AlphaFold2 Predicts Whether Proteins Interact Amidst Confounding Structural Compatibility

Juliette Martin

AlphaFold2 predicts whether proteins interact amidst confounding structural compatibility

Juliette Martin[1,2]

1: Univ Lyon, CNRS, UMR 5086 MMSB, 7 passage du Vercors F-69367, Lyon, France
2: Laboratory of Biology and Modeling of the Cell, Ecole Normale Supérieure de Lyon, CNRS UMR 5239, Inserm U1293, University Claude Bernard Lyon 1, Lyon, France.
juliette.martin@ens-lyon.fr

Abstract

Predicting whether two proteins physically interact is one of the holy grails of computational biology, galvanized by rapid advancements in deep learning. AlphaFold2, although not developed with this goal, seems promising in this respect. Here, I test the prediction capability of AlphaFold2 on a very challenging data set, where proteins are structurally compatible, even when they do not interact. AlphaFold2 achieves high discrimination between interacting and non-interacting proteins, and the cases of misclassifications can either be rescued by revisiting the input sequences or can suggest false positives and negatives in the data set. Alphafold2 is thus not impaired by the compatibility between protein structures and has the potential to be applied at large scale.

Introduction

Prediction of whether proteins interact has profound implications to suggest functions for uncharacterized proteins, understand protein activity and regulation at the molecular level, and more generally, highlight protein functions in the context of global interactomes. Numerous computational methods have been developed to predict whether or not two proteins physically interact, based on their sequences and 3D structures, see the following references for review [1–6].

The formidable capability of AlphaFold2 (AF2) to predict protein 3D structures [7] has stimulated the creativity of the scientific community to evaluate what are the application range and limits of AF2 predictions [8–20] and extend the tool beyond its initial prediction task [18,21–38]. Prediction of protein-protein complex structures, a task traditionally addressed by protein-protein docking, has rapidly been tackled by modification in the input for the AF2 monomer pipeline [24]. A specific model for protein-protein complexes is now available, with breakthrough prediction results [39]. Note that in this case, predictions are made with the prior knowledge that the proteins physically interact.

Logically then, the capability of AlphaFold2 to predict whether proteins interact has recently been explored, by using the predicted quality of modeled interfaces as prediction criterion, with very encouraging discrimination capability [28,35,40–45].

In this short article, I challenge AF2 on a particular, presumably difficult data set in which non-interacting protein pairs are special cases, in which the two (non-interacting) proteins are structurally similar to available experimental complexes [46]. This feature should challenge AF2, since the proteins are compatible in terms of structures. Using the ipTM score of AF2, I found that AF2 is very accurate at discriminating interacting from non-interacting pairs, even in this challenging context, attending an AUC value of 0.93. Interestingly, model recycling did not improve the discriminative power. The analysis of the few misclassified cases provides suggestions to further improve the discrimination and how to use AF2 for pair screening.

Material and Methods

Data set

Protein pairs from *S cerevisiae* are taken from our previous study [46], details about these data set can be found in [46] and are briefly summarized below.

Interacting protein pairs

The initial data set of interacting protein pairs was extracted from three sources: high confidence physical interactions from BioGrid [47], direct interactions from the KUPS resource [48], and high confidence physical interactions detected by yeast-two-hybrid from Ito et al [49].

Non-interacting protein pairs

The initial negative data set was extracted from three sources : the negative data set used by Yu et al which are simply sampled from pairs without experimental evidence of interaction [50], pairs of protein from the KUPS resource [48], which have no evidence of interaction and also distant GO annotations, and the negative data set built by Trabuco et al [51] from the Ito data set, where the yeast-two-hybrid data set is used to select proteins pairs that do not physically interact but correctly detected in the experiment .

In our previous study, those pairs were compared with known structures. We had screened homology models of S. *cerevisiae* proteins against a non-redundant database of experimentally known dimers and we had selected pairs where the monomers structurally matched with the experimental dimers (TM score >0.8), and, once superimposed on those dimers, could form an interface of reasonable size (> 20 residues) and without extensive clashes (less than 3 between Cαs). Using these criteria resulted in a data set of 22 non-interacting and 222 interacting proteins. In this work, I use the data set of 22 non-interacting proteins and a random sample of 22 interacting pairs, see Table S1.

AlphaFold2 Models

AlphaFold2 (AF2) predictions are computed using LocalColabFold (https://github.com/YoshitakaMo/localcolabfold), a local installation of ColabFold [39,52]. ColabFold replaces the time-consuming step of multiple sequence alignment (MSA) creation by an ultra-fast step with MMseqs2 [53]. The version used is colabfold version 1.5.2 with model alphafold2_multimer_v3. No templates are used (default); models are not

minimized (default); I tested both with (default) and without model recycling (--num-recycle 0), and different modes of sequence pairing for the MSA: unpaired+paired (default), paired only (--pair-mode paired), and unpaired only (--pair-mode unpaired). The resulting 5 models are ranked according to the ipTM score (interface predicted TM score) computed by AF2. The ipTM score was introduced in AlphaFold-Multimer [39]. It is computed in the structure module of AF2, by a network that takes as input the pair representation of the MSA produced by the Evoformer blocks [7]. This network has been trained to predict a non-symmetric matrix of predicted aligned errors (PAE) values, that captures the error in the position of the Cα atom of residue j when the predicted and true structures are aligned using the backbone frame of residue i [7]. In AlphaFold-Multimer, the ipTM only incorporates the PAE values of pairs where residues i and j come from different chains [39]. During inference, the ipTM is thus predicted only from the MSA information and does not explicitly use the predicted interface.

With the unpaired+paired scoring scheme and no recycling, the prediction took on average 8 minutes per pair on a GPU *RTX 3080 Lite Hash Rate.*

VoroMQA rescoring

For comparison purposes, the VoroQMA scoring [54] was also tested for the discrimination between interacting and non-interacting pairs. This method was primarily developed for quality assessment of monomeric models [55]; it is based on knowledge-based statistical potentials and Voronoi tessellation to quantify the contact area between atoms. In the case of protein-protein complexes, VoroMQA outputs an *interface score* that is derived from the local scores of atoms that participate in inter-chain contacts and an *interface energy* which is the sum of the interface contact areas multiplied by the corresponding pseudo-energy values from the statistical potential [56]. Both the interface score and the interface energy values were considered.

Assessing classification performance

The separation between scores of interacting and non-interacting pairs is measured by the AUC value obtained when using the scores to predict whether the proteins interact. Two prediction scores are considered: the AF2 ipTM score and the pDockQ score, recently introduced by Bryant et al [40], which is derived from the plDDT scores of interface residues. Statistical significance between AUC values is assessed using the non-parametric DeLong's test [57] implemented in the pROC package [58].

The classification performance is measured by the accuracy, i.e., percentage of correctly classified pairs. Statistical difference between accuracies is assessed using the MacNemar test [59].

Comparison with other methods

The performance of AF2 to predict whether proteins interact was compared with two other methods. The first method, DeepTrio [60], is a deep learning method based on single sequences. The sequences were submitted to the DeepTrio server (http://bis.zju.edu.cn/deeptrio), and predictions were run with the yeast model. The second method is based on explicit co-evolution signal detection. Following the work of Green et al [61], the co-evolution signal between proteins was computed using the EVcouplings python package (https://github.com/debbiemarkslab/EVcouplings) and the average value of the top-ten inter protein residue pairs was used as predictor. The prediction took on average 4 hours per pair when running on 2 cores on CPU Intel Xeon w5-3423.

Data availability

All the input sequences, input parameters and resulting models can be found in the Zenodo archive https://zenodo.org/doi/10.5281/zenodo.10118918.

Results

Although AF2 has recently been used to discriminate interacting from non-interacting pairs with promising results [28,35,40,41], it is always worthy of pushing the system to the limits to better know its applicability range. Here, I propose to further test AF2 prediction capability in extreme conditions. I submitted to AF2 prediction a particularly challenging data set from a previous study [46]. In this data set, all the pairs are supported by structural data: interacting but also non-interacting pairs are compatible in shape, as assessed by their high similarity to experimental dimers, as explained in the Methods section.

Classification performance

Table 1- AUC values obtained with ipTM score used as a predictor, and different Alphafold2 options. The differences between AUC values are not statistically significant between the best models obtained with different MSA pairing and recycling modes (within the last line), or between models generated by different networks (within each column).

		paired+unpaired MSA		paired MSA	unpaired MSA
		With recycling	No recycling	No recycling	No recycling
AlphaFold network	1	0.9215	0.8905	0.7893	0.9081
	2	0.9298	0.8698	0.8202	0.8812
	3	0.9132	0.906	0.7593	0.8244
	4	0.8967	0.8709	0.7665	0.9514
	5	0.8843	0.9112	0.874	0.9339
	Best	0.8636	0.9256	0.8595	0.9143

AlphaFold2 models were predicted for 22 interacting pairs and 22 non-interacting pairs, with or without recycling, and various MSA pairing schemes. The ipTM scores were then used to predict whether the proteins of a given pair interact. The results are shown in Table 1. As can be seen in Table 1, the discrimination between interacting and non-interacting pairs is very high, attaining AUC values between 0.86 and 0.93, and accuracy

values between 80 and 86% (see Table S1). Unfortunately, the limited size of the data set does not allow to statistically differentiate those different settings (DeLong test p-values>0.05). Notably, the model recycling did not significantly improve the prediction, and even produced a lower AUC value: 0.92 without recycling *versus* 0.86 with recycling.

Effect of model recycling

In order to explore why the recycling did not improve the prediction, I monitored the evolution of the ipTM scores of the models during the recycling process, see Figure 1.

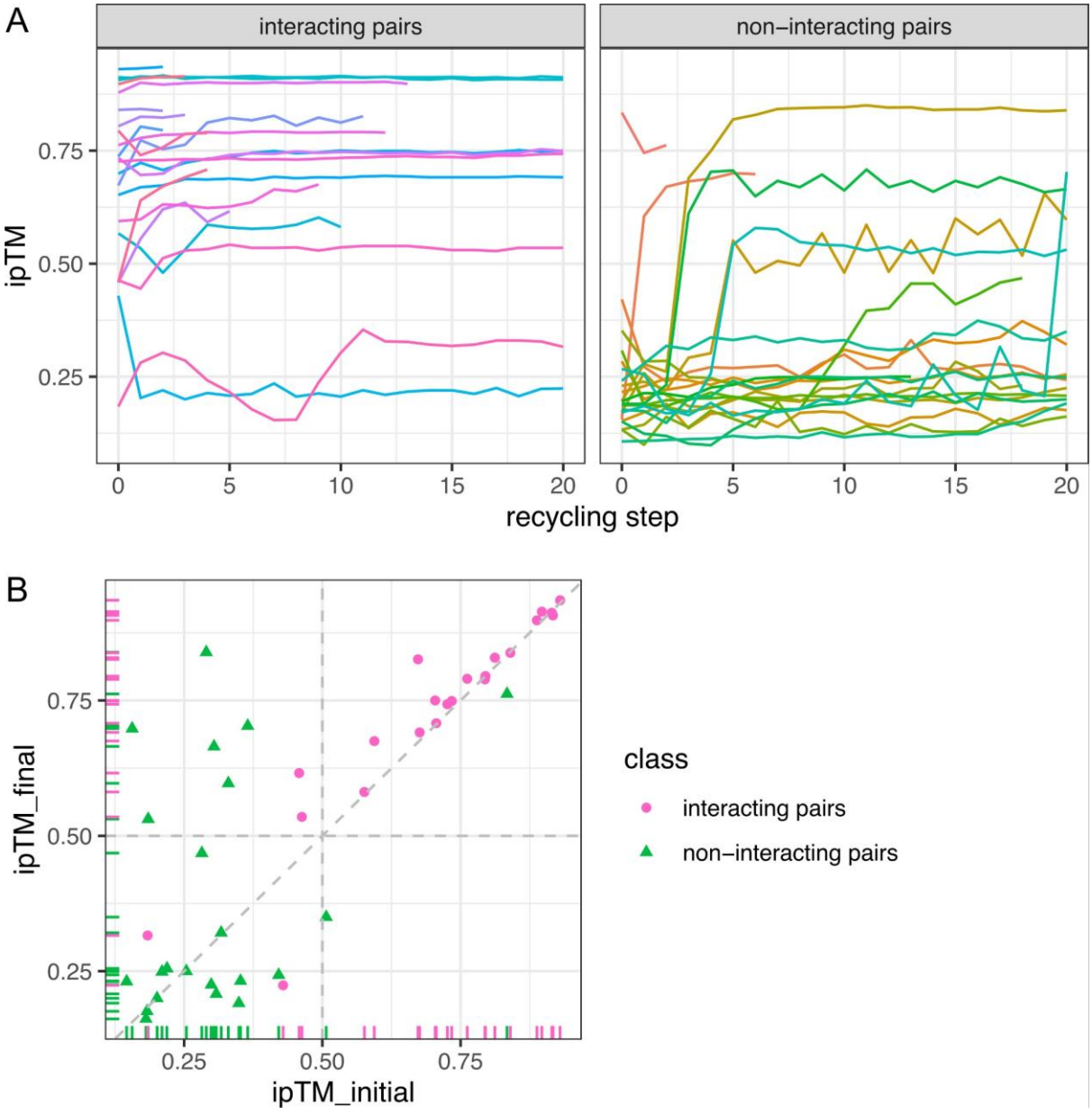

Figure 1. Effect of recycling on the ipTM scores (MSA mode paire_unpaired). A: ipTM scores of models at each recycling step. B: comparison between ipTM scores at the beginning and the end of the recycling.

As can be seen in Figure 1A, for a majority of models, the ipTM remains in the same range during recycling: many interacting pairs have high ipTM scores from the beginning, and a majority of non-interacting pairs have low initial ipTM scores cannot be improved by the recycling.

A small number of protein pairs however, see their ipTM score increasing with recycling, but since it is the case for both interacting and non-interacting pairs, this results in the decrease of the AUC value. As shown in Figure 1B, two interacting and 8 non-interacting protein pairs indeed start with a low ipTM (<0.5) and have a high ipTM (>0.5) score at the end of the recycling process. There are also a few cases of models that have slightly lower ipTM after recycling, but this is observed mainly for non-interacting pairs, and the changes are limited.

Misclassified protein pairs

In this section, I considered models obtained without recycling, and the default MSA pairing option. The distribution of ipTM scores of the models is presented in Figure 2 for interacting and non-interacting pairs. As said before, this distribution results in an AUC equal to 0.93, indicating a good separation between score distributions. However, in order to perform a prediction from these scores, a prediction cutoff has to be chosen. Classically, the prediction cutoff is derived from the ROC curve, e.g., by maximizing the vertical distance to the diagonal or by setting an accepted false positive rate. In the present situation, the limited size of the data set produced a stepped ROC curve, with high variance on estimated metrics [62]. For this reason, it seemed more reasonable to choose the prediction cutoff based on external reasons instead. Since the cutoff of 0.5 for the TM score has been shown to be reliable for protein topology classification [63], this cutoff was chosen for the ipTM score.

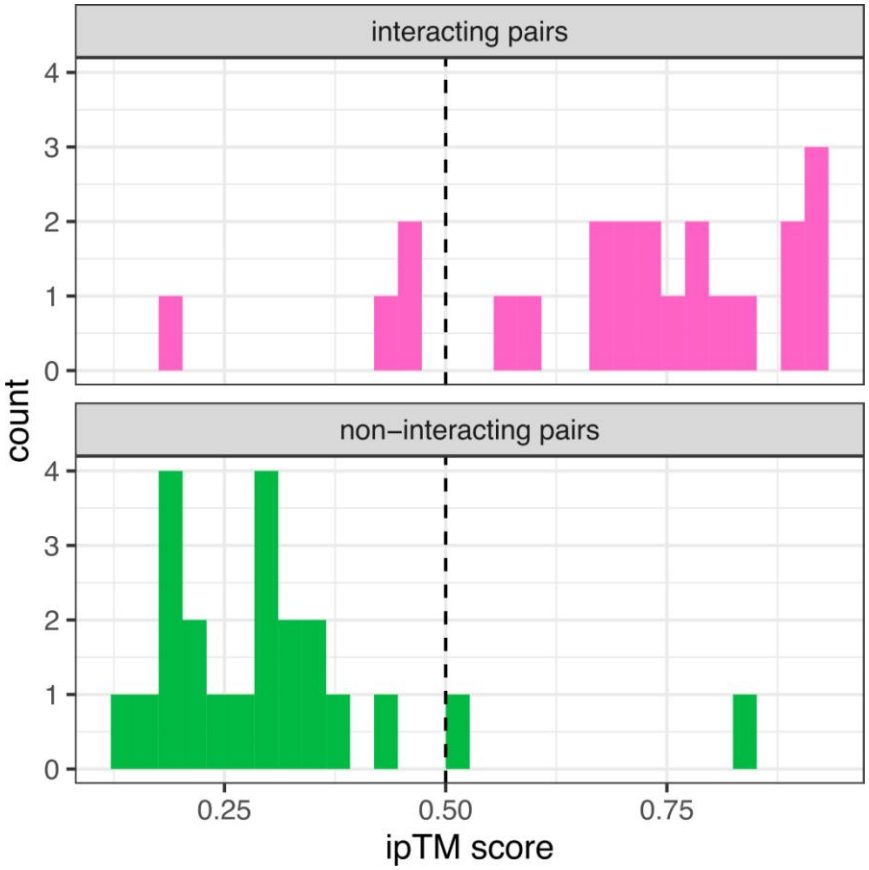

Figure 2. Distribution of ipTM scores of AlphaFold2 models of interacting and non-interacting pairs (no recycling, default MSA pairing). The vertical dashed line indicates the prediction cutoff (0.5).

With an ipTM cutoff set to 0.5, 6 pairs are misclassified, corresponding to an accuracy equal to 86%, see Figure 2. The misclassified cases are shown in Figures 3 and 4 and discussed below.

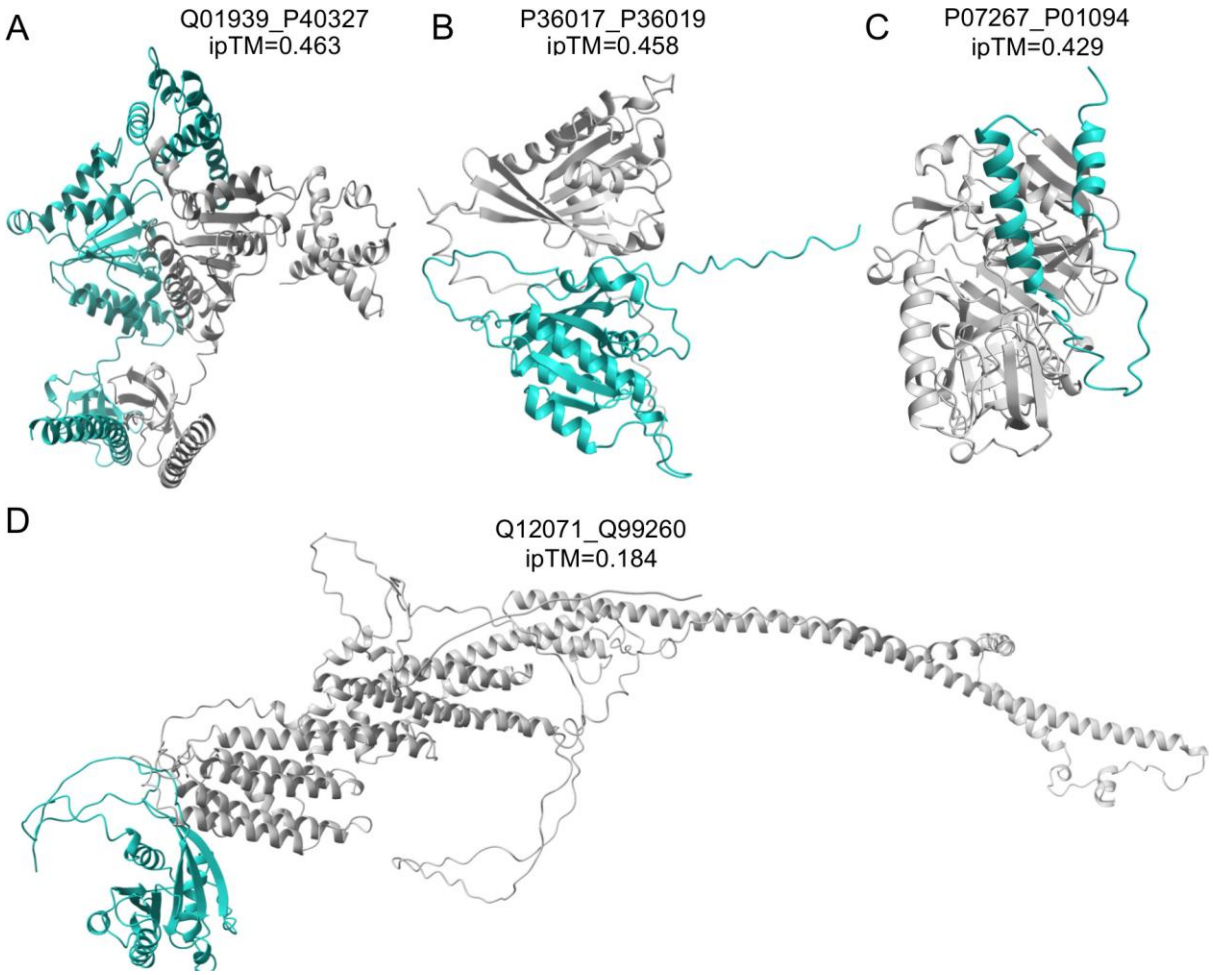

Figure 3. AF2 models of interacting pairs misclassified as non-interacting (ipTM<0.5). A: pair formed by the subunits 8 (Q01939, in blue) and 4 (P40327, in grey) of the 26s proteasome. B: pair formed by the vacuolar protein sorting-associated protein 21 (P36017 in grey) and the GTP-binding protein YPT53 (P36019, in blue). C: pair formed by saccharopepsin (P07267, in grey) and its inhibitor (P01094, in blue). D: pair formed by the vacuolar protein sorting-associated protein 54 (Q12071, in grey) and the GTP-binding protein YPT6 (Q99260, in blue).

Interacting pairs predicted as non-interacting

Four interacting pairs are incorrectly classified as non-interacting (ipTM score <0.5). The first case is the dimer formed by the regulatory subunit 8 homolog and subunit 4 homolog (Uniprot ids Q01939 and P40327) of the 26S proteasome, which is an assembly of

47 protein chains. The dimer obtains an ipTM score equal to 0.463, close to the prediction cutoff, see Figure 3A. The comparison of the AF2 model with the experimental structure of the 26S proteasome (PDB id 3JCO) reveals that the AF2 model is structurally similar to the dimer formed by the subunits 7 and 4, meaning that AF2 wrongly placed the subunit 8 in place of the subunit 7, see Figure S1. This is possible because subunits 8 and 7 are structurally similar (TM score = 0.6 between experimental structures). So, in this case, AF2 prediction was confused by the presence of another interacting chain with similar structure, highlighting a type of competition toward the same interface between partners of similar folds during the prediction.

The pair formed by the vacuolar protein sorting-associated protein 21 and the GTP-binding protein YPT53 (Uniprot ids P36017 and P36019) obtains an ipTM score equal to 0.458, close to the prediction cutoff. Furthermore, the PAE matrix displays low values between the two chains, suggesting a good confidence in the relative orientation of the two protein chains. The AF2 model, shown in Figure 3B, has an interface involving the disorder C-tail of the GTP protein, which could explain the low score. Interestingly, the model obtained with recycling has a good ipTM score (0.616) with a similar configuration but without the disordered part at the interface (see Figure S2). However, model recycling significantly increases the computation time, which is a limiting factor in the perspective of pair screening. Alternatively to model recycling, the prediction can be run with the disordered part chopped from the sequence, which produces a model similar to the recycled one, with an ipTM score equal to 0.63 (see Figure S2). So, in that case, it is possible to ‘rescue’ the prediction by chopping the sequence.

The pair formed by the saccharopepsin and its inhibitor (Uniprot ids P07267 and P01094) obtains an ipTM score equal to 0.429, see Figure 3C. The full-length sequence of the saccharopepsin contains an N-terminal propeptide of 75 residues that is cleaved upon activation of the enzyme [64]. The comparison of the AF2 model with the experimental structure (PDB id 3COJ) reveals that the binding cleft where the inhibitor is supposed to bind is occluded by the N-terminal region of the enzyme corresponding to the propeptide. Re-

running the prediction after chopping the propeptide sequence results in a model with a good ipTM score equal to 0.63 and in good agreement with the experimental structure (see Figure S3). So, in this case also, it is possible to rescue the prediction with appropriate sequence chopping.

The pair formed by the vacuolar protein sorting-associated protein 54 and the GTP-binding protein YPT6 (Uniprot ids Q12071 and Q99260) obtains a very low ipTM score equal to 0.184, see Figure 3D. The 5 models are drastically different from each other, with even drastic situations where the proteins are not in contact in two of the models (see Figure S4). There is evidence of physical interaction between these proteins, as detected by affinity purification. However, there is no evidence of direct physical interaction by two-hybrid assay. So there is the possibility that this pair is in fact a false positive case.

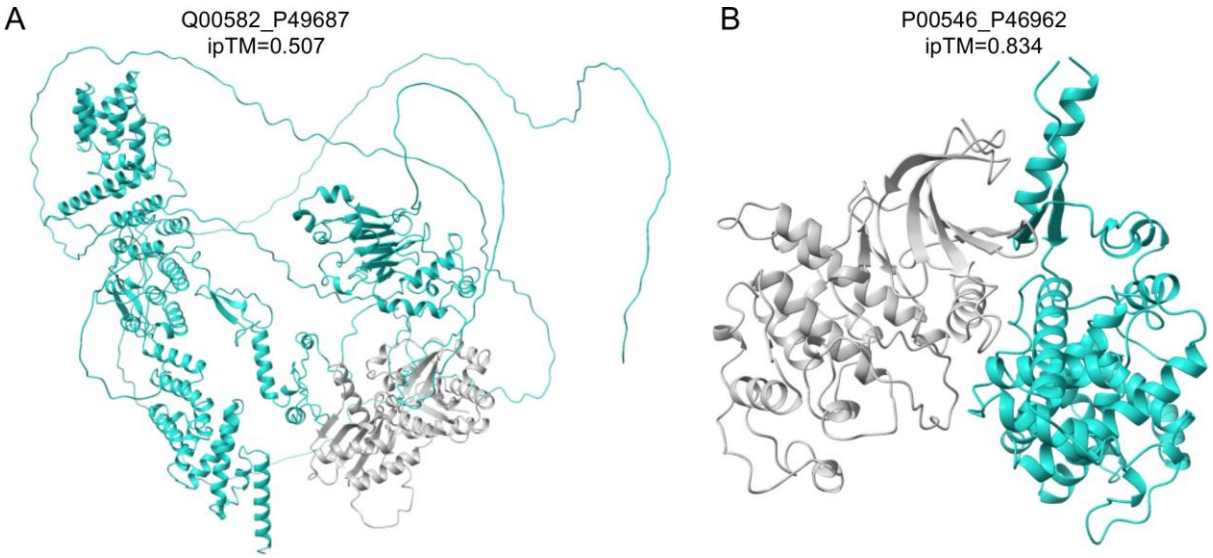

Figure 4: AF2 models of non-interacting pair misclassified as interacting (ipTM>0.5). A: pair formed by the GTP-binding protein GTR1 (Q00582, in grey) and the nucleoporin NUP145 (P49687, in blue). B: pair formed by cyclin-dependent kinase 1 (P00546, in grey) and the CTD kinase subunit beta (P46962, in blue)

Non-interacting pairs predicted as interacting

Two pairs of non-interacting proteins are incorrectly classified as interacting. The pair between GTP-binding protein GTR1 and nucleoporin NUP145, a component of the nuclear pore complex (Uniprot id Q00582 and P49687) obtains a border line ipTM score equal to 0.507. The examination of the AF2 model, see Figure 4A, reveals that the nucleoporin has a disordered N-terminal region that mediates the interaction with GTR1. After chopping the disordered parts of the nucleoporin, the best model obtains an ipTM score equal to 0.23. So in this case, the initial ipTM score was meaningless because of the disordered parts, and the pair can be excluded by adequate sequence chopping.

The pair formed by the cyclin-dependent kinase 1 and the CTD kinase subunit beta (Uniprot id P00546 and P46962) achieves a very high ipTM score equal to 0.834, see Figure 4B. In this case, all models have high ipTM scores (>0.8) and are structurally similar (data not shown). Although there is no evidence of direct interaction between these two proteins, they are reported as interacting in the STRING resource [65]. Notably, a direct interaction between homologs in Drosophila was measured by yeast-two-hybrid assay [66]. This suggests that this case could be a false negative pair.

In summary, out of 6 misclassified cases, three could be corrected by sequence chopping, two can be questioned as false positive/negative, and one highlights a phenomenon of confusion between chains in interfaces in a macromolecular assembly.

Discussion

In this work, pairs of interacting and non-interacting proteins, all with compatible shapes, were submitted to AlphaFold2 to assess its capability to predict whether proteins interact, and different options were tested for the AlphaFold2 modeling. Overall, the limited size of the data set (44 pairs) did not allow to select the best modeling options, but the discrimination was very high, with AUC values between 0.86 and 0.93. This AUC range is comparable to the recent result of Bryant et al who obtained an AUC equal to 0.87 using the monomer AF2 pipeline to classify pairs of *E. coli* proteins using the pDockQ score, which is

based on the plDDT scores of interface residues [41]. In the present study, the pairs are presumably more challenging to distinguish, because they are all compatible in shape, but AlphaFold2 is still able to provide good discrimination. Interestingly, the pDockQ score achieved a separation corresponding to AUC values between 0.67 and 0.78 (see Table S3), which is significantly lower than the AUC values obtained with the ipTM score (Delong's test p-value=0.011). This suggests that in the case of this particular data set, AlphaFold is able to generate models with high plDDT scores at the interface and a reasonable number of contacts, even in the case of non-interacting proteins, but the ipTM score is able to distinguish them.

The role of recycling on this data set was investigated by following the evolution of ipTM scores with recycling. Recycling indeed improved the ipTM scores of some pairs, but this was observed not only for interacting pairs but also non-interacting pairs. The reasons why Alphafold can produce models with high ipTM scores for non-interacting protein pairs when using recycling are difficult to investigate. It is important to remember that Alphafold was not trained to predict whether proteins interact, but to provide good models for proteins and protein complexes that effectively exist. Thus, the recycling procedure will optimize the models toward this particular aim. Achieving optimal discrimination with recycling would probably require a specific training of the AlphaFold network with interacting *versus* non-interacting protein pairs.

Since the ipTM score is based on the PAE matrix predicted from the pair MSA representation, and thus, does not explicitly take into account the interface structure, it is interesting to assess the models produced by AF2 with structure-based scoring function. The VoroQMA knowledge-based statistical potential was used for this purpose, see Figure S5. It was found that the interface score of VoroQMA achieves a good level of discrimination (AUC equal to 0.74), although lower than the ipTM score (Delong's test p-value=0.04). However, the VoroQMA interface energy achieved discrimination on par with the ipTM score (AUC equal to 85%, Delong's test p-value >0.05). This indicates that the structural models

produced by AlphaFold for non-interacting and non-interacting pairs differ in terms of pairwise interface atomic contacts.

Lastly, the performance of AlphaFold was compared to other methods that are less demanding in terms of computation time. On the one hand, a deep learning online method, DeepTrio, achieved very poor discrimination, predicting most of the pairs as interacting (see Table S3). On the other hand, a method based on co-evolution signal detection, EVcouplings, also achieved poor discrimination but for the opposite reason: many of the interacting protein pairs yielded very low co-evolution signal (see Figure S6).

Conclusion

A reduced but challenging data set was submitted to AF2 in order to discriminate interacting from non-interacting pairs, resulting in very high prediction accuracy. Several misclassified cases could be rescued by appropriate sequence chopping, and some others are suggestive of incorrect annotations (false positive or false negative). A potential limitation of AF2 was observed in a case where several protein chains with structural similarity form a supra-molecular assembly. The fact that no recycling is required opens the possibility to apply this procedure at large scale. To conclude, AF2 seems a promising technology for predicting whether proteins interact, even capable of discriminating interacting from non-interacting pairs in presence of confounding structural compatibility.

Data and Software Availability statement

The sequences used in that study are listed in Table S1, and the models generated by AlphaFold are available under Zenodo https://doi.org/10.5281/zenodo.10118919.

Supporting Information

Additional extended results : accuracy and AUC values for discrimination by AF2, 3D structures of models predicted by AF2, discrimination with other methods and comparison with VoroQMA scores.

References

(1) Hu, X.; Feng, C.; Ling, T.; Chen, M. Deep Learning Frameworks for Protein–Protein Interaction Prediction. *Comput. Struct. Biotechnol. J.* **2022**, *20*, 3223–3233. https://doi.org/10.1016/j.csbj.2022.06.025.

(2) Dunham, B.; Ganapathiraju, M. K. Benchmark Evaluation of Protein-Protein Interaction Prediction Algorithms. *Mol. Basel Switz.* **2021**, *27* (1), 41. https://doi.org/10.3390/molecules27010041.

(3) Casadio, R.; Martelli, P. L.; Savojardo, C. Machine Learning Solutions for Predicting Protein–Protein Interactions. *WIREs Comput. Mol. Sci. n/a* (n/a), e1618. https://doi.org/10.1002/wcms.1618.

(4) Zahiri, J.; Bozorgmehr, J. H.; Masoudi-Nejad, A. Computational Prediction of Protein–Protein Interaction Networks: Algo-Rithms and Resources. *Curr. Genomics* **2013**, *14* (6), 397–414. https://doi.org/10.2174/1389202911314060004.

(5) Shoemaker, B. A.; Panchenko, A. R. Deciphering Protein–Protein Interactions. Part II. Computational Methods to Predict Protein and Domain Interaction Partners. *PLoS Comput Biol* **2007**, *3* (4), e43. https://doi.org/10.1371/journal.pcbi.0030043.

(6) Wass, M. N.; David, A.; Sternberg, M. J. Challenges for the Prediction of Macromolecular Interactions. *Curr. Opin. Struct. Biol.* **2011**. https://doi.org/10.1016/j.sbi.2011.03.013.

(7) 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, S. A. A.; Ballard, A. J.; 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, A. W.; Kavukcuoglu, K.; Kohli, P.; Hassabis, D. Highly Accurate Protein Structure Prediction with AlphaFold. *Nature* **2021**, *596* (7873), 583–589. https://doi.org/10.1038/s41586-021-03819-2.

(8) Johansson-Åkhe, I.; Wallner, B. *Benchmarking Peptide-Protein Docking and Interaction Prediction with AlphaFold-Multimer*; 2021; p 2021.11.16.468810. https://doi.org/10.1101/2021.11.16.468810.

(9) Akdel, M.; Pires, D. E. V.; Pardo, E. P.; Jänes, J.; Zalevsky, A. O.; Mészáros, B.; Bryant, P.; Good, L. L.; Laskowski, R. A.; Pozzati, G.; Shenoy, A.; Zhu, W.; Kundrotas, P.; Serra, V. R.; Rodrigues, C. H. M.; Dunham, A. S.; Burke, D.; Borkakoti, N.; Velankar, S.; Frost, A.; Basquin, J.; Lindorff-Larsen, K.; Bateman, A.; Kajava, A. V.; Valencia, A.; Ovchinnikov, S.; Durairaj, J.; Ascher, D. B.; Thornton, J. M.; Davey, N. E.; Stein, A.; Elofsson, A.; Croll, T. I.; Beltrao, P. A Structural Biology Community Assessment of AlphaFold2 Applications. *Nat. Struct. Mol. Biol.* **2022**, *29* (11), 1056–1067. https://doi.org/10.1038/s41594-022-00849-w.

(10) Pozzati, G.; Zhu, W.; Bassot, C.; Lamb, J.; Kundrotas, P.; Elofsson, A. Limits and Potential of Combined Folding and Docking. *Bioinformatics* **2021**, btab760. https://doi.org/10.1093/bioinformatics/btab760.

(11) Saldaño, T.; Escobedo, N.; Marchetti, J.; Zea, D. J.; Donagh, J. M.; Rueda, A. J. V.; Gonik, E.; Melani, A. G.; Nechcoff, J. N.; Salas, M. N.; Peters, T.; Demitroff, N.; Alberti, S. F.; Palopoli, N.; Fornasari, M. S.; Parisi, G. *Impact of Protein Conformational Diversity on AlphaFold Predictions*; 2021; p 2021.10.27.466189. https://doi.org/10.1101/2021.10.27.466189.

(12) Fowler, N. J.; Williamson, M. P. The Accuracy of Protein Structures in Solution Determined by AlphaFold and NMR. *Structure* **2022**, *30* (7), 925-933.e2. https://doi.org/10.1016/j.str.2022.04.005.

(13) Chakravarty, D.; Porter, L. L. AlphaFold2 Fails to Predict Protein Fold Switching. bioRxiv March 8, 2022, p 2022.03.08.483439. https://doi.org/10.1101/2022.03.08.483439.

(14) Sawicki, L. R.; Benitez, G.; Carletti, M.; Palopoli, N.; Fornasari, M. S.; Parisi, G. Conformational Epistasis Impairs AlphaFold Structural Predictions. bioRxiv November 17, 2022, p 2022.11.15.516638. https://doi.org/10.1101/2022.11.15.516638.
(15) Wong, F.; Krishnan, A.; Zheng, E. J.; Stärk, H.; Manson, A. L.; Earl, A. M.; Jaakkola, T.; Collins, J. J. Benchmarking AlphaFold- enabled Molecular Docking Predictions for Antibiotic Discovery. *Mol. Syst. Biol.* **2022**, *18* (9), e11081. https://doi.org/10.15252/msb.202211081.
(16) Yin, R.; Feng, B. Y.; Varshney, A.; Pierce, B. G. Benchmarking AlphaFold for Protein Complex Modeling Reveals Accuracy Determinants. *Protein Sci.* **2022**, *31* (8), e4379. https://doi.org/10.1002/pro.4379.
(17) Holcomb, M.; Chang, Y.-T.; Goodsell, D. S.; Forli, S. Evaluation of AlphaFold2 Structures as Docking Targets. *Protein Sci. Publ. Protein Soc.* **2023**, *32* (1), e4530. https://doi.org/10.1002/pro.4530.
(18) Bryant, P.; Pozzati, G.; Zhu, W.; Shenoy, A.; Kundrotas, P.; Elofsson, A. Predicting the Structure of Large Protein Complexes Using AlphaFold and Sequential Assembly. bioRxiv March 12, 2022, p 2022.03.12.484089. https://doi.org/10.1101/2022.03.12.484089.
(19) McCafferty, C. L.; Pennington, E. L.; Papoulas, O.; Taylor, D. W.; Marcotte, E. M. Does AlphaFold2 Model Proteins' Intracellular Conformations? An Experimental Test Using Cross-Linking Mass Spectrometry of Endogenous Ciliary Proteins. bioRxiv August 26, 2022, p 2022.08.25.505345. https://doi.org/10.1101/2022.08.25.505345.
(20) Martin, J. When Alphafold2 Predictions Go Wrong for Protein–Protein Complexes, Is There Something to Be Learnt? *Q. Rev. Biophys.* **2022**, *55*, e6. https://doi.org/10.1017/S0033583522000051.
(21) Bruley, A.; Bitard-Feildel, T.; Callebaut, I.; Duprat, E. A Sequence-Based Foldability Score Combined with AlphaFold2 Predictions to Disentangle the Protein Order/Disorder Continuum. *Proteins Struct. Funct. Bioinforma.* **2022**, *n/a* (n/a). https://doi.org/10.1002/prot.26441.
(22) Jussupow, A.; Kaila, V. R. I. Effective Molecular Dynamics from Neural- Network Based Structure Prediction Models. **2022**, 37.
(23) del Alamo, D.; Sala, D.; Mchaourab, H. S.; Meiler, J. Sampling Alternative Conformational States of Transporters and Receptors with AlphaFold2. *eLife* **2022**, *11*, e75751. https://doi.org/10.7554/eLife.75751.
(24) Ghani, U.; Desta, I.; Jindal, A.; Khan, O.; Jones, G.; Kotelnikov, S.; Padhorny, D.; Vajda, S.; Kozakov, D. Improved Docking of Protein Models by a Combination of Alphafold2 and ClusPro. *bioRxiv* **2021**, 2021.09.07.459290. https://doi.org/10.1101/2021.09.07.459290.
(25) Roney, J. P.; Ovchinnikov, S. State-of-the-Art Estimation of Protein Model Accuracy Using AlphaFold. bioRxiv June 19, 2022, p 2022.03.11.484043. https://doi.org/10.1101/2022.03.11.484043.
(26) Sala, D.; Hildebrand, P. W.; Meiler, J. Biasing AlphaFold2 to Predict GPCRs and Kinases with User-Defined Functional or Structural Properties. *Front. Mol. Biosci.* **2023**, *10*.
(27) Stein, R. A.; Mchaourab, H. *Modeling Alternate Conformations with Alphafold2 via Modification of the Multiple Sequence Alignment*; 2021; p 2021.11.29.470469. https://doi.org/10.1101/2021.11.29.470469.
(28) Yu, D.; Chojnowski, G.; Rosenthal, M.; Kosinski, J. AlphaPulldown – a Python Package for Protein-Protein Interaction Screens Using AlphaFold-Multimer. bioRxiv August 6, 2022, p 2022.08.05.502961. https://doi.org/10.1101/2022.08.05.502961.
(29) Andorf, C. M.; Sen, S.; Hayford, R. K.; Portwood, J. L.; Cannon, E. K.; Harper, L. C.; Gardiner, J. M.; Sen, T. Z.; Woodhouse, M. R. FASSO: An AlphaFold Based Method to Assign Functional Annotations by Combining Sequence and Structure Orthology. bioRxiv November 15, 2022, p 2022.11.10.516002. https://doi.org/10.1101/2022.11.10.516002.
(30) Monzon, V.; Paysan-Lafosse, T.; Wood, V.; Bateman, A. Reciprocal Best Structure

Hits: Using AlphaFold Models to Discover Distant Homologues. bioRxiv July 4, 2022, p 2022.07.04.498216. https://doi.org/10.1101/2022.07.04.498216.
(31) Baltzis, A.; Mansouri, L.; Jin, S.; Langer, B. E.; Erb, I.; Notredame, C. Improving Sequence Alignments with AlphaFold2 Regardless of Structural Modeling Accuracy. bioRxiv May 25, 2022, p 2022.05.24.492699. https://doi.org/10.1101/2022.05.24.492699.
(32) Bordin, N.; Sillitoe, I.; Nallapareddy, V.; Rauer, C.; Lam, S. D.; Waman, V. P.; Sen, N.; Heinzinger, M.; Littmann, M.; Kim, S.; Velankar, S.; Steinegger, M.; Rost, B.; Orengo, C. AlphaFold2 Reveals Commonalities and Novelties in Protein Structure Space for 21 Model Organisms. bioRxiv June 3, 2022, p 2022.06.02.494367. https://doi.org/10.1101/2022.06.02.494367.
(33) Brems, M. A.; Runkel, R.; Yeates, T. O.; Virnau, P. AlphaFold Predicts the Most Complex Protein Knot and Composite Protein Knots. *Protein Sci.* **2022**, *31* (8), e4380. https://doi.org/10.1002/pro.4380.
(34) Wayment-Steele, H. K.; Ovchinnikov, S.; Colwell, L.; Kern, D. Prediction of Multiple Conformational States by Combining Sequence Clustering with AlphaFold2. bioRxiv October 17, 2022, p 2022.10.17.512570. https://doi.org/10.1101/2022.10.17.512570.
(35) Gao, M.; Nakajima An, D.; Parks, J. M.; Skolnick, J. AF2Complex Predicts Direct Physical Interactions in Multimeric Proteins with Deep Learning. *Nat. Commun.* **2022**, *13* (1), 1744. https://doi.org/10.1038/s41467-022-29394-2.
(36) Chang, L.; Perez, A. Ranking Peptide Binders by Affinity with AlphaFold**. *Angew. Chem. Int. Ed.* **2023**, *62* (7), e202213362. https://doi.org/10.1002/anie.202213362.
(37) Tsaban, T.; Varga, J. K.; Avraham, O.; Ben-Aharon, Z.; Khramushin, A.; Schueler-Furman, O. Harnessing Protein Folding Neural Networks for Peptide-Protein Docking. *Nat. Commun.* **2022**, *13* (1), 176. https://doi.org/10.1038/s41467-021-27838-9.
(38) Wallner, B. AFsample: Improving Multimer Prediction with AlphaFold Using Aggressive Sampling. bioRxiv December 20, 2022, p 2022.12.20.521205. https://doi.org/10.1101/2022.12.20.521205.
(39) 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* **2021**, 2021.10.04.463034. https://doi.org/10.1101/2021.10.04.463034.
(40) Bryant, P.; Noe, F. Rapid Protein-Protein Interaction Network Creation from Multiple Sequence Alignments with Deep Learning. bioRxiv April 17, 2023, p 2023.04.15.536993. https://doi.org/10.1101/2023.04.15.536993.
(41) Bryant, P.; Pozzati, G.; Elofsson, A. Improved Prediction of Protein-Protein Interactions Using AlphaFold2. *Nat. Commun.* **2022**, *13* (1), 1265. https://doi.org/10.1038/s41467-022-28865-w.
(42) Humphreys, I. R.; Pei, J.; Baek, M.; Krishnakumar, A.; Anishchenko, I.; Ovchinnikov, S.; Zhang, J.; Ness, T. J.; Banjade, S.; Bagde, S. R.; Stancheva, V. G.; Li, X.-H.; Liu, K.; Zheng, Z.; Barrero, D. J.; Roy, U.; Kuper, J.; Femández, I. S.; Szakal, B.; Branzei, D.; Rizo, J.; Kisker, C.; Greene, E. C.; Biggins, S.; Keeney, S.; Miller, E. A.; Fromme, J. C.; Hendrickson, T. L.; Cong, Q.; Baker, D. Computed Structures of Core Eukaryotic Protein Complexes. *Science* **2021**, *374* (6573), eabm4805. https://doi.org/10.1126/science.abm4805.
(43) Burke, D. F.; Bryant, P.; Barrio-Hernandez, I.; Memon, D.; Pozzati, G.; Shenoy, A.; Zhu, W.; Dunham, A. S.; Albanese, P.; Keller, A.; Scheltema, R. A.; Bruce, J. E.; Leitner, A.; Kundrotas, P.; Beltrao, P.; Elofsson, A. Towards a Structurally Resolved Human Protein Interaction Network. *Nat. Struct. Mol. Biol.* **2023**, *30* (2), 216–225. https://doi.org/10.1038/s41594-022-00910-8.
(44) O'Reilly, F. J.; Graziadei, A.; Forbrig, C.; Bremenkamp, R.; Charles, K.; Lenz, S.; Elfmann, C.; Fischer, L.; Stülke, J.; Rappsilber, J. Protein Complexes in Cells by AI-Assisted Structural Proteomics. *Mol. Syst. Biol.* **2023**, *19* (4), e11544. https://doi.org/10.15252/msb.202311544.

(45) Lim, Y.; Tamayo-Orrego, L.; Schmid, E.; Tarnauskaite, Z.; Kochenova, O. V.; Gruar, R.; Muramatsu, S.; Lynch, L.; Schlie, A. V.; Carroll, P. L.; Chistol, G.; Reijns, M. A. M.; Kanemaki, M. T.; Jackson, A. P.; Walter, J. C. In Silico Protein Interaction Screening Uncovers DONSON's Role in Replication Initiation. *Science* **2023**, *381* (6664), eadi3448. https://doi.org/10.1126/science.adi3448.
(46) Launay, G.; Ceres, N.; Martin, J. Non-Interacting Proteins May Resemble Interacting Proteins: Prevalence and Implications. *Sci. Rep.* **2017**, *7*, 40419. https://doi.org/10.1038/srep40419.
(47) Oughtred, R.; Rust, J.; Chang, C.; Breitkreutz, B.-J.; Stark, C.; Willems, A.; Boucher, L.; Leung, G.; Kolas, N.; Zhang, F.; Dolma, S.; Coulombe-Huntington, J.; Chatr-Aryamontri, A.; Dolinski, K.; Tyers, M. The BioGRID Database: A Comprehensive Biomedical Resource of Curated Protein, Genetic, and Chemical Interactions. *Protein Sci. Publ. Protein Soc.* **2021**, *30* (1), 187–200. https://doi.org/10.1002/pro.3978.
(48) Chen, X. -w.; Jeong, J. C.; Dermyer, P. KUPS: Constructing Datasets of Interacting and Non-Interacting Protein Pairs with Associated Attributions. *Nucleic Acids Res.* **2010**, No. Database Issue, D750-754.
(49) Ito, T.; Chiba, T.; Ozawa, R.; Yoshida, M.; Hattori, M.; Sakaki, Y. A Comprehensive Two-Hybrid Analysis to Explore the Yeast Protein Interactome. *Proc. Natl. Acad. Sci.* **2001**, *98* (8), 4569–4574. https://doi.org/10.1073/pnas.061034498.
(50) Yu, J.; Guo, M.; Needham, C. J.; Huang, Y.; Cai, L.; Westhead, D. R. Simple Sequence-Based Kernels Do Not Predict Protein-Protein Interactions. *Bioinformatics* **2010**, *26* (20), 2610–2614. https://doi.org/10.1093/bioinformatics/btq483.
(51) Trabuco, L. G.; Betts, M. J.; Russell, R. B. Negative Protein-Protein Interaction Datasets Derived from Large-Scale Two-Hybrid Experiments. *Methods San Diego Calif* **2012**, *58* (4), 343–348. https://doi.org/10.1016/j.ymeth.2012.07.028.
(52) Mirdita, M.; Schütze, K.; Moriwaki, Y.; Heo, L.; Ovchinnikov, S.; Steinegger, M. ColabFold: Making Protein Folding Accessible to All. *Nat. Methods* **2022**, *19* (6), 679–682. https://doi.org/10.1038/s41592-022-01488-1.
(53) Steinegger, M.; Söding, J. MMseqs2 Enables Sensitive Protein Sequence Searching for the Analysis of Massive Data Sets. *Nat. Biotechnol.* **2017**, *35* (11), 1026–1028. https://doi.org/10.1038/nbt.3988.
(54) Olechnovič, K.; Venclovas, Č. VoroMQA Web Server for Assessing Three-Dimensional Structures of Proteins and Protein Complexes. *Nucleic Acids Res.* **2019**, *47* (W1), W437–W442. https://doi.org/10.1093/nar/gkz367.
(55) Olechnovič, K.; Venclovas, Č. VoroMQA: Assessment of Protein Structure Quality Using Interatomic Contact Areas. *Proteins Struct. Funct. Bioinforma.* **2017**, *85* (6), 1131–1145. https://doi.org/10.1002/prot.25278.
(56) Dapkūnas, J.; Olechnovič, K.; Venclovas, Č. Modeling of Protein Complexes in CAPRI Round 37 Using Template-Based Approach Combined with Model Selection. *Proteins Struct. Funct. Bioinforma.* **2018**, *86* (S1), 292–301. https://doi.org/10.1002/prot.25378.
(57) DeLong, E. R.; DeLong, D. M.; Clarke-Pearson, D. L. Comparing the Areas under Two or More Correlated Receiver Operating Characteristic Curves: A Nonparametric Approach. *Biometrics* **1988**, *44*, 837–845.
(58) Robin, X.; Turck, N.; Hainard, A.; Tiberti, N.; Lisacek, F.; Sanchez, J.-C.; Müller, M. pROC: An Open-Source Package for R and S+ to Analyze and Compare ROC Curves. *BMC Bioinformatics* **2011**, *12*, 77. https://doi.org/10.1186/1471-2105-12-77.
(59) Patrick Walters, W. Comparing Classification Models—a Practical Tutorial. *J. Comput. Aided Mol. Des.* **2022**, *36* (5), 381–389. https://doi.org/10.1007/s10822-021-00417-2.
(60) Hu, X.; Feng, C.; Zhou, Y.; Harrison, A.; Chen, M. DeepTrio: A Ternary Prediction System for Protein–Protein Interaction Using Mask Multiple Parallel Convolutional Neural Networks. *Bioinformatics* **2022**, *38* (3), 694–702. https://doi.org/10.1093/bioinformatics/btab737.
(61) Green, A. G.; Elhabashy, H.; Brock, K. P.; Maddamsetti, R.; Kohlbacher, O.; Marks, D. S. Large-Scale Discovery of Protein Interactions at Residue Resolution Using Co-Evolution Calculated from Genomic Sequences. *Nat. Commun.* **2021**, *12* (1), 1396.

https://doi.org/10.1038/s41467-021-21636-z.
(62) Hanczar, B.; Hua, J.; Sima, C.; Weinstein, J.; Bittner, M.; Dougherty, E. R. Small-Sample Precision of ROC-Related Estimates. *Bioinformatics* **2010**, *26* (6), 822–830. https://doi.org/10.1093/bioinformatics/btq037.
(63) Xu, J.; Zhang, Y. How Significant Is a Protein Structure Similarity with TM-Score = 0.5? *Bioinformatics* **2010**, *26* (7), 889. https://doi.org/10.1093/bioinformatics/btq066.
(64) Dreyer, T.; Halkier, B.; Svendsen, I.; Ottesen, M. Primary Structure of the Aspartic Proteinase A from Saccharomyces Cerevisiae. *Carlsberg Res. Commun.* **1986**, *51* (1), 27. https://doi.org/10.1007/BF02907993.
(65) Szklarczyk, D.; Gable, A. L.; Nastou, K. C.; Lyon, D.; Kirsch, R.; Pyysalo, S.; Doncheva, N. T.; Legeay, M.; Fang, T.; Bork, P.; Jensen, L. J.; von Mering, C. The STRING Database in 2021: Customizable Protein–Protein Networks, and Functional Characterization of User-Uploaded Gene/Measurement Sets. *Nucleic Acids Res.* **2021**, *49* (D1), D605–D612. https://doi.org/10.1093/nar/gkaa1074.
(66) Stanyon, C. A.; Liu, G.; Mangiola, B. A.; Patel, N.; Giot, L.; Kuang, B.; Zhang, H.; Zhong, J.; Finley, R. L. A Drosophila Protein-Interaction Map Centered on Cell-Cycle Regulators. *Genome Biol.* **2004**, *5* (12), R96. https://doi.org/10.1186/gb-2004-5-12-r96.

For Table of Contents Only