Received: 30 August 2022 | Revised: 28 November 2022 | Accepted: 1 December 2022

DOI: 10.1002/pro.4530

TOOLS FOR PROTEIN SCIENCE

Evaluation of AlphaFold2 structures as docking targets

Matthew Holcomb[1] | **Ya-Ting Chang**[1] | **David S. Goodsell**[1,2,3,4] | **Stefano Forli**[1]

[1]Department of Integrative Structural and Computational Biology, The Scripps Research Institute, La Jolla, California, USA

[2]Research Collaboratory for Structural Bioinformatics Protein Data Bank, Rutgers, The State University of New Jersey, Piscataway, New Jersey, USA

[3]Institute for Quantitative Biomedicine, Rutgers, The State University of New Jersey, Piscataway, New Jersey, USA

[4]Rutgers Cancer Institute of New Jersey, Rutgers, The State University of New Jersey, New Brunswick, New Jersey, USA

Correspondence
Stefano Forli, Department of Integrative Structural and Computational Biology, The Scripps Research Institute, 10550 N. Torrey Pines Road, La Jolla, CA 92037, USA.
Email: forli@scripps.edu

Funding information
National Institute of General Medical Sciences, Grant/Award Number: GM069832

Review Editor: Nir Ben-Tal

Abstract

AlphaFold2 is a promising new tool for researchers to predict protein structures and generate high-quality models, with low backbone and global root-mean-square deviation (RMSD) when compared with experimental structures. However, it is unclear if the structures predicted by AlphaFold2 will be valuable targets of docking. To address this question, we redocked ligands in the PDBbind datasets against the experimental co-crystallized receptor structures and against the AlphaFold2 structures using AutoDock-GPU. We find that the quality measure provided during structure prediction is not a good predictor of docking performance, despite accurately reflecting the quality of the alpha carbon alignment with experimental structures. Removing low-confidence regions of the predicted structure and making side chains flexible improves the docking outcomes. Overall, despite high-quality prediction of backbone conformation, fine structural details limit the naive application of AlphaFold2 models as docking targets.

KEYWORDS

AlphaFold2, AutoDock protein structure prediction, computational docking, computer-aided drug design, drug design and development, virtual screening

1 | INTRODUCTION

Computer-aided design is an essential component of modern drug discovery and development. As presented in a recent literature review (Sabe et al., 2021), computational methods have been acknowledged in the discovery process for at least 70 commercialized drugs. The study documents the use of ~80 methods for virtual screening, with programs of the AutoDock suite (Forli et al., 2016; Morris et al., 2009) as the most used. The use of these virtual screening methods with homology modeling as targets is less clear—in the review of these 70 drugs, homology modeling was only noted in four. However, docking experiments with computationally derived target models have a long history.

Template-based homology models have long been used in docking experiments with some significant successes. Several survey studies have provided a few guidelines (Bordogna et al., 2011; Fan et al., 2009; McGovern & Shoichet, 2003). As might be expected, docking results are only as good as the model, and in particular, the exact conformation of the active site, so in

general, structure of *holo* complexes provide better results than *apo* target structures, which are in turn generally better than modeled structures. Modeling some degree of receptor flexibility can provide some respite for small problems with active site geometry. A full review of docking with homology models is beyond the scope of this report, but an overall rule-of-thumb has emerged that homology models based on templates with >50% sequence identity can be expected to be accurate enough for docking and screening (Bordogna et al., 2011). The emergence of machine-learning methods for protein structure prediction promises to lift this limitation.

The recent advance in structure prediction with AlphaFold2 (AF2) expanded the scope of protein structures available for docking to cover the entire human proteome (Jumper et al., 2021). Recently these predictions were expanded to cover virtually every known protein sequence (Callaway, 2022; Varadi et al., 2022). This will enable molecular biology and medicinal chemistry programs to apply structure-based methods, such as virtual screening, to otherwise uncharacterized targets. However, it remains an open question whether the metrics used to validate structures generated to ensure they will be high-quality targets for docking.

Use of computationally derived models of proteins in docking simulations poses multiple challenges. Perhaps most importantly, docking methods are typically quite sensitive to the local details of loop flexibility and side-chain conformation. A variety of methods are used to address this challenge, including docking approaches that include target flexibility within the docking experiment and methods that dock to an ensemble of conformations derived from molecular dynamics of the target. Additionally, methods such as AF2 are currently most effective for monomeric proteins that have defined, folded native structures. We might expect that this limitation will be lifted as the method is improved to predict functional oligomeric assemblies.

In this study, we seek to benchmark the performance of AutoDock using AF2-determined structures relative to experimentally resolved structures in redocking exercises. We are adopting the structure and standards of the CASF competition (Su et al., 2019), in which comparisons between docking methodologies are usually made, to instead evaluate the source of the target structure.

2 | RESULTS

2.1 | Docking using PDBbind structures and AlphaFold2 models

To assess the performance of AutoDock simulations using AlphaFold2 structures, we started with the refined set of protein-ligand complexes in the PDBbind database (Wang et al., 2005). For docking calculations, we used AutoDock-GPU v1.5.3 (Santos-Martins et al., 2021), the current version of AutoDock that leverages acceleration by Graphics Processing Units. Of the 5316 complexes in the refined set of the PDBbind database, 2907 are human proteins and had structures predicted by AF2 in January of 2022. Of this set, 2474 were parsed without issues in ligand valence and protonation states which could not be resolved by OpenBabel (O'Boyle et al., 2011), and were used for docking. Redocking against the crystal structure with AutoDock-GPU led to a 41% success rate, measured by the RMSD of the highest-ranked docked pose of less than 2 Å. This is comparable to a previously published performance of AutoDock 4 of 45% (Huey et al., 2007). By contrast, docking against the AF2-predicted structures led to a 17% success rate. The docking success rate was not substantially impacted by the oligomeric nature of the complex, and both docking tasks were hindered by the absence of a cofactor during docking otherwise present in the crystal complex (Table 1).

A comparison of the docking RMSD against the predicted and cocrystal structures shows two lines along which most of the complexes fall (Figure 1a). In the complexes lying along the diagonal, the docking performance was comparable on the two sets of structures, indicating that poor RMSD in this region is due to inherent challenges in predicting the complex rather than features of the AF2 predicted structures. By contrast, the group of complexes lying in the low RMSD region for redocking against the crystal structure but have varying RMSDs against the predicted structure reflect complexes where the use of an AF2 predicted structure introduced a new challenge in the docking. This increased RMSD in the AF2 prediction is not reflective of poor alignment between the structures, which is not predictive of docking performance (Figure 1b). However, this alignment was correlated with the "predicted Local Distance Test" (pLDDT) used as a confidence measure for the AF2 predictions (Figure 1c). This reflects a disparity between the structural metrics used to assess prediction quality and docking performance. There is a bimodal distribution in the docking RMSDs for both docking tasks, which is more pronounced for the predicted structures (Figure 1a). Visual inspection suggests the first peak, centered at about 1.5 Å in both histograms, is associated with binding to the correct pocket, in an accurate or nearly accurate orientation, and the second peak, a broad peak at about 6 Å in both histograms, is associated with binding to a nearby pocket, or to the correct pocket with a highly inaccurate binding mode. The first peak roughly corresponds to the 2 Å RMSD cutoff defined as a successful docking.

TABLE 1 Summary of docking results

		<2 Å RMSD		2–5 Å RMSD		>5 Å RMSD	
Structures	**Number**	**PDB (%)**	**AF2 (%)**	**PDB (%)**	**AF2 (%)**	**PDB (%)**	**AF2 (%)**
All	2474	41	17	25	24	34	60
Monomer	1797	40	17	24	24	36	59
Oligomer	677	44	17	28	22	28	60
No cofactor	1821	47	18	24	22	30	60
Cofactor	653	25	15	31	28	44	57

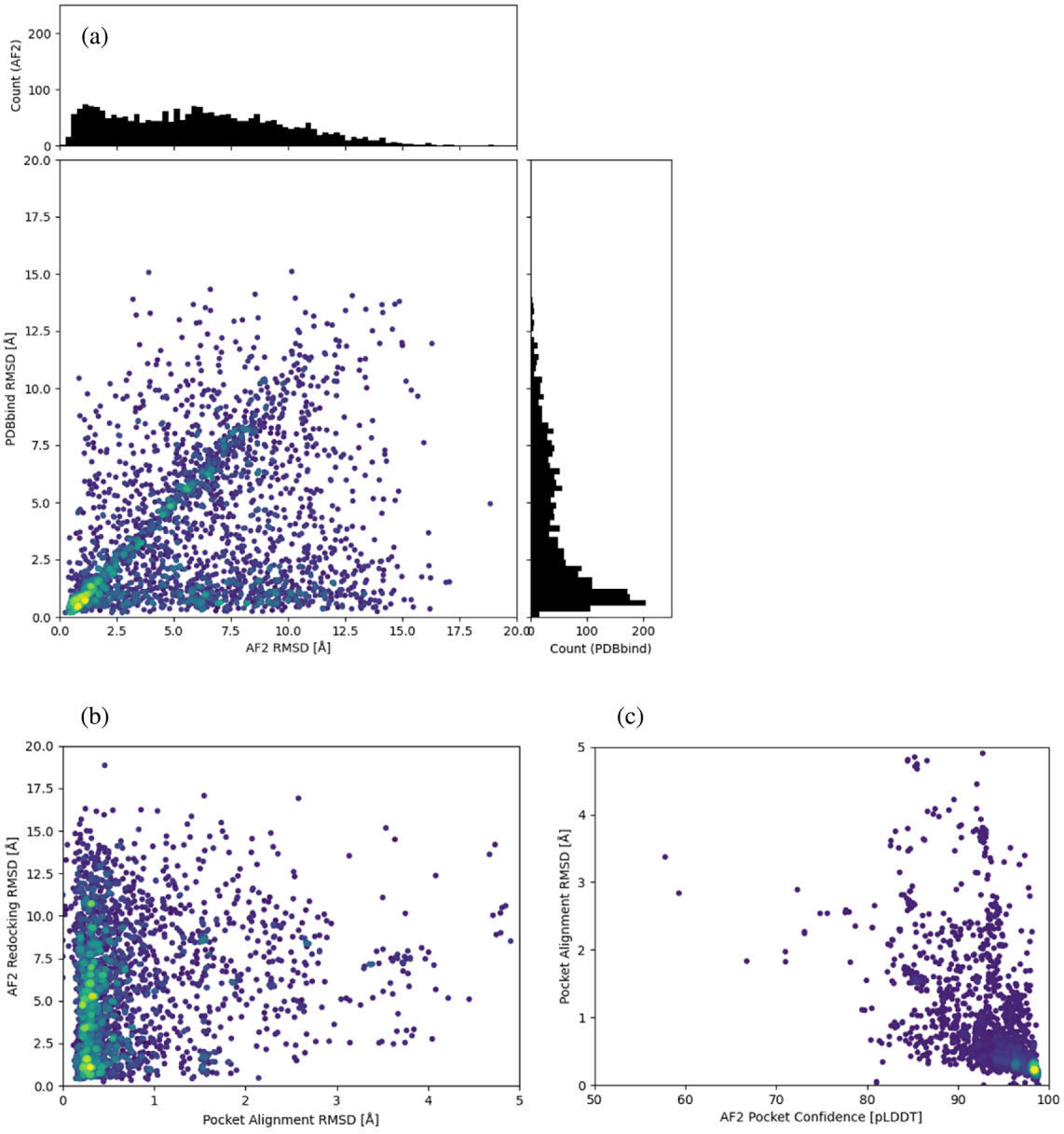

FIGURE 1 Redocking and alignment RMSD statistics. (a) Docking RMSD for first ranked pose for the crystal structure versus the AF2 predicted structure; (b) docking RMSD for first ranked pose against the AF2 predicted structure versus the RMSD between C_α of the aligned AF2 structure and the crystal structure for pocket residues; (c) pocket alignment RMSD versus mean AF2 predicted confidence for pocket residues. Points are colored by relative density of the plotted data (blue: low density, yellow: high density). Some data lies off the shown axes.

2.2 | Flexible receptor docking

AF2 is trained and tested against the accuracy of its C_α predictions, while side chains are minimized using a forcefield and are not considered in the quality of the model. In light of this, we expected there would be cases where the secondary structural characteristics of the pocket were predicted with high fidelity, but that small variations in side chain conformations would lead to inaccurate docking results. Indeed, in the cases of PDB ID 4ufl and 5hz6 the backbone alignment in the binding pocket was excellent (0.30 and 0.43 Å RMSD respectively), but side chain clashes led to inaccurate redocking (3.00 and 6.29 Å RMSD respectively; Figure 2). In AutoDock it is possible to simulate specific side chains as flexible by modeling all bonds as rotatable from the C_α on (Morris et al., 2009). Improved RMSD docking results (0.48 and 1.86 Å respectively) were obtained by making selected clashing side chains flexible during the docking.

2.3 | Low-confidence strands

Low-confidence regions of the AF2 predictions may represent barriers to docking when they are predicted to be located in close proximity to the binding site. For example, the predicted structure of 3qkd includes an alpha-helical segment that is disordered in the crystallographic structure, but is placed within the active site by AF2 (Figure 3). The active site pocket was well aligned (0.69 Å RMSD) but the docking result was inaccurate (9.70 Å RMSD). Deleting the interfering low-confidence alpha helix and making two residues flexible led to a substantially improved result (2.89 Å RMSD). While this docking RMSD would not be low enough to be counted as a success in our above statistics, the pseudosymmetry of multiple motifs in the ligand means this docked model still contains substantial information regarding the position and orientation of the ligand in the pocket.

In some cases, AF2 regions with low confidence may add additional information that can be employed when interpreting docking results. For example, in the case of 5hcy the alignment of pocket C_α positions was accurate (0.92 Å RMSD) but the docking result was inaccurate (9.85 Å RMSD). A low-confidence strand was predicted to be in the pocket, and perhaps surprisingly the tyrosine predicted to sit in the pocket by AF2 overlays very well with the crystal structure of the ligand (Figure 4). Deletion of this strand gives a structure for which flexible docking can predict the correct pose (1.20 Å RMSD), albeit in the fifth-ranked cluster by energy. However, this is the lowest energy pose which matches the pharmacophoric information given by the highlighted tyrosine, potentially allowing the identification of its relevance in a prospective study. It is worth noting that in the case of 3qkd, no side chains match pharmacophoric information of the bound ligand. This does not, however, exclude the possibility that this strand could be used to inform ligand design.

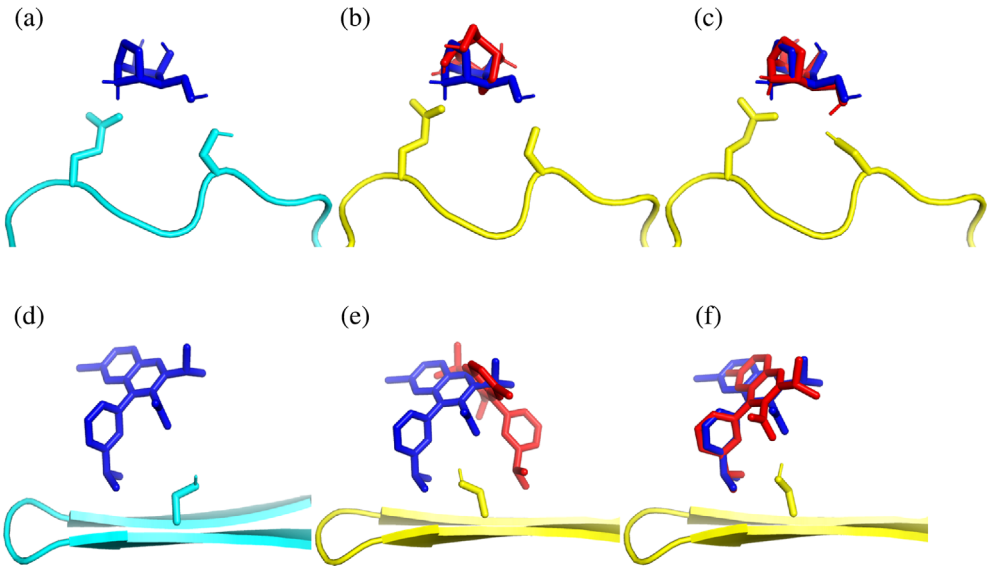

FIGURE 2 Flexible receptor docking examples (blue: ligand crystal structure, red: ligand docked to AF2 predicted receptor structure, cyan: receptor crystal structure, yellow: AF2 predicted receptor structure). (a, d) Crystal structures of complexes from PDB ID 4ufl and 5hz6; (b, e) best energy result docking against AF2 predicted structures of 4ufl and 5hz6; (c, f) best energy result docking against AF2 predicted structures of 4ufl and 5hz6 with flexible side chains. Minimal structural information of the receptor is included to give context to the flexible side chain, other secondary structure is omitted. Only side chains treated as flexible in the dockings for panels (c) and (f) are shown.

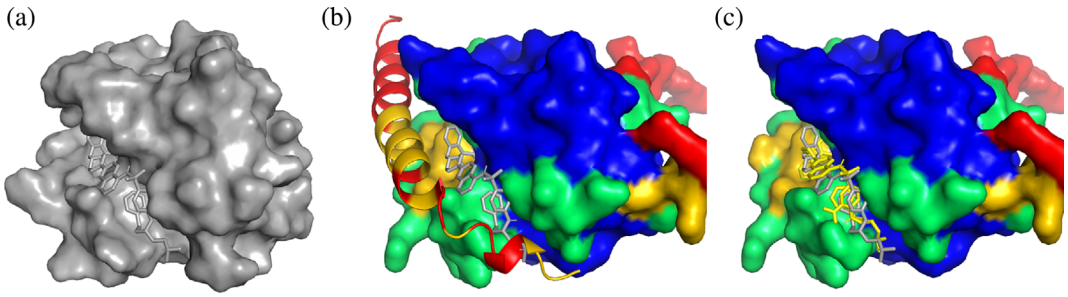

FIGURE 3 An example of a low-confidence region interfering with docking. (a) Crystal structure from PDB ID 3qkd (gray); (b) AF2 predicted structure (colored by confidence measure, blue: pLDDT > 90, green: 90 > pLDDT > 70, yellow: 70 > pLDDT > 50, red: pLDDT < 50) highlighting overlap between low-confidence region and ligand crystal structure; (c) docked pose with best score (yellow). Phe105 and Tyr195 were treated flexibly in the AF2 model but are hidden for visual clarity.

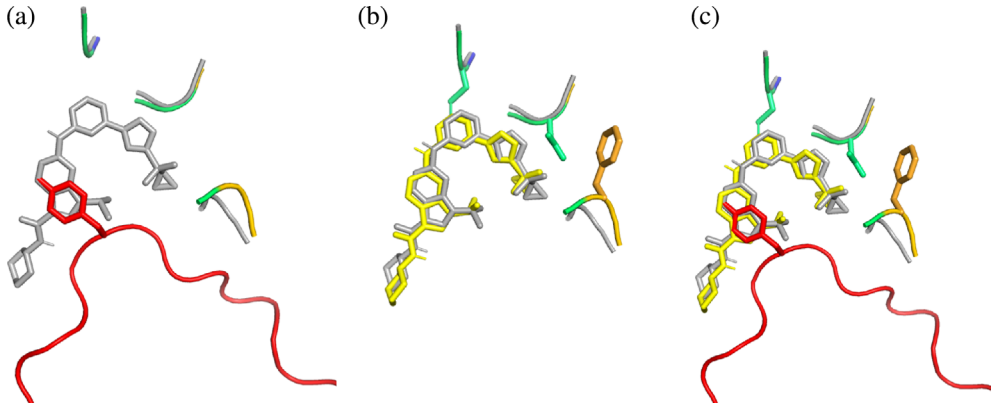

FIGURE 4 An example of a low-confidence strand with potential pharmacophoric information. (a) Crystal structure from PDB ID 5hcy (gray) and AF2 predicted structure (colored by confidence measure, blue: pLDDT > 90, green: 90 > pLDDT > 70, yellow: 70 > pLDDT > 50, red: pLDDT < 50). A tyrosine with good pharmacophoric agreement with the crystallized ligand is shown; (b) docked ligand with best agreement to crystal (yellow) and side chains made flexible shown. (c) The overlap of the tyrosine in the low-confidence strand with the ligand. The pose in panel (b) is the best pose in the fifth-ranked cluster. Minimal structural information of the receptor is included to give context to the flexible side chain, other secondary structure is omitted. Only side chains treated as flexible in the docking are shown.

2.4 | Comparison with apo structures

A reasonable point of comparison for the success rate against AF2-predicted structures might be apo structures, which also lack a ligand to induce conformational changes. Structures were extracted from the APOBind dataset which correspond to the structures examined here (Aggarwal et al., 2021), totaling 1764 of the 2474 structures included in the results above. Among this set, no significant difference was observed in performance for either the crystallographic or AF2 redocking tasks, indicating no bias was introduced by limiting the analysis to this overlap (Figure 5e). Perhaps surprisingly, the docking against apo structures performed significantly worse than against AF2 structures (10% vs. 16%). On inspecting these results, AF2 produced models that aligned better with the holo than the apo structure in many cases. For example, in the case of 1c83 a loop is in an open position in the apo structure, and closed in both holo and AF2 predicted (Figure 5a,b). The case of 1rpj is more extreme, with multiple loops and helices opening rearranged between apo and holo, with AF2 again matching the holo (Figure 5c,d). This suggests AF2 and apo structures may be complementary in docking tasks. In fact, the consideration of both, selecting the best energy pose across the two targets, leads to a comparable success rate to the AF2 alone (17%), and considering the best pose of each leads to substantially improved performance (22%; Figure 5e).

3 | DISCUSSION

In this work, we used AutoDock-GPU to evaluate structures predicted by AF2 as targets for docking. We found that the vast majority of predicted structures aligned well with the cocrystal structures, and that the pLDDTs, a measure of confidence of an AF2 model, for residues in the pocket correlated with this alignment (Figure 1c).

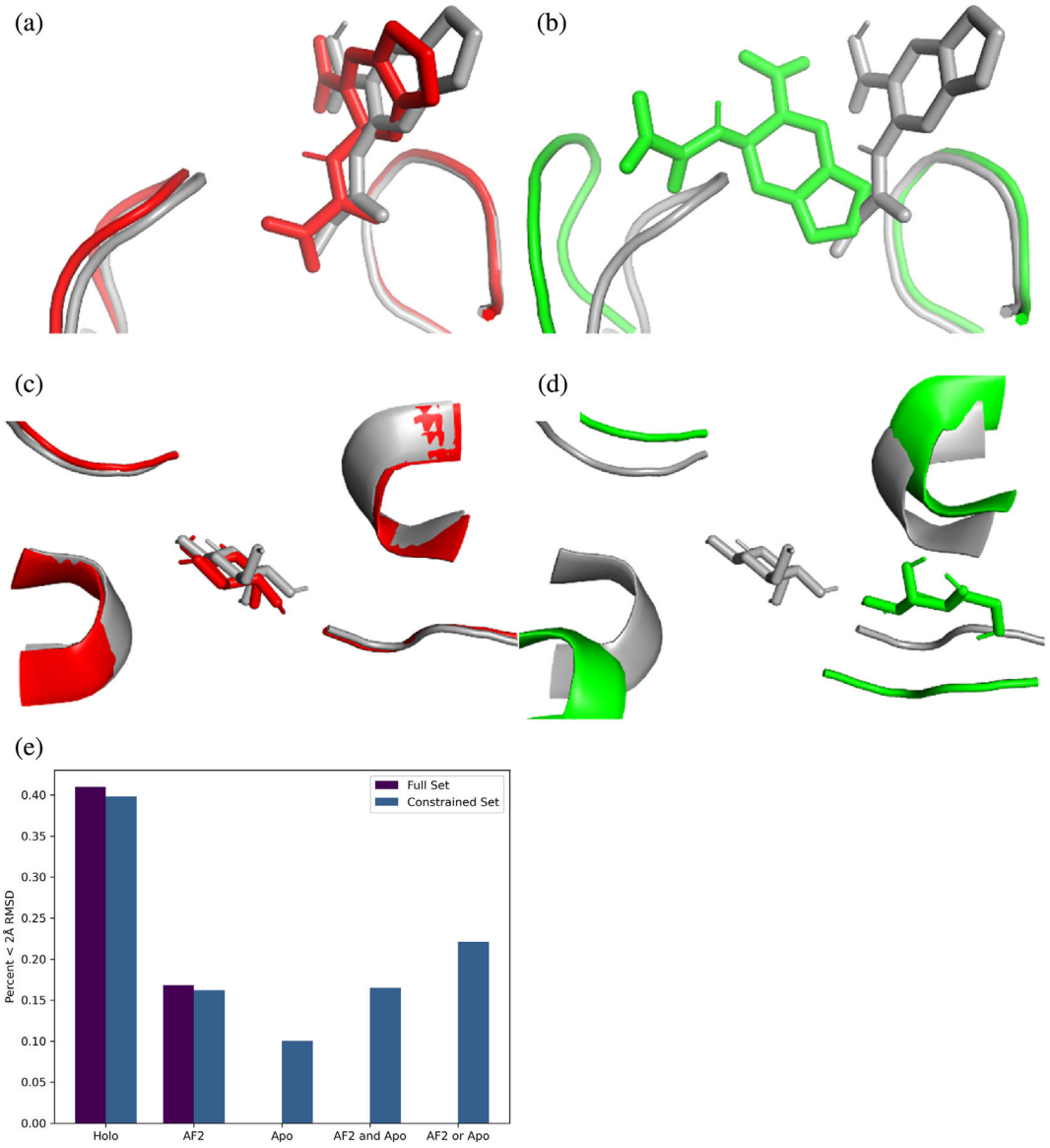

FIGURE 5 Comparison performance on AF2 predicted structures and apo structures. (a, c) Crystal structures (gray) from PDB ID 1c83 and 1rpj respectively and docking results against AF2 predicted structure (red). (b, d) Crystal structures as in (a) and (c) and docking results against apo structures (green). (e) Comparison of docking success rates against holo, AF2 predicted structures, and a mixture of AF2 predicted structures and apo structures. "AF2 and apo" refers to cases where the best-scored ligand across both docking targets is within 2 Å of the crystal structure; "AF2 or apo" refers to cases where the best-scored ligand for either target is within 2 Å. "Full set" refers to the 2474 complexes with AF2 structures, "constrained set" refers to the 1764 complexes also present in the APOBind dataset. Minimal structural information is shown for clarity.

This may be due to the fact that the PDBbind is enriched for well-folded proteins, which would be expected to perform well with AF2. As such, this work may represent an upper bound on the ability to naively apply AF2-predicted structures to computer-aided drug design. However, as well-folded proteins are generally regarded as more "druggable," this is still an important set of proteins to examine.

The loss in docking accuracy compared to redocking against the crystal structure was substantial (17% vs. 41%) and not predicted by the quality of the alignment (Table 1, Figure 1b). This indicates both that the difference in docking accuracy is not explained by poor alignment leading to inflated RMSD, and that high-quality predictions as measured by the RMSD metric used in evaluating AF2 at CASP is not indicative of the ability to use these structures for docking.

We were surprised that oligomerization played very little role in the success of the docking experiments, showing a success rate of 17% both for monomeric and oligomeric proteins in AF2 dockings. Cofactors, on the other hand, had a stronger impact on docking success,

both in the experimental redockings and in docking to AF2 structures. This is to be expected, since interactions with cofactors, and metals in particular, often play a central role in specificity of docking.

We found that AF2 models are comparable and complementary to apo structures for use in docking. Success rate for docking against AF2 predictions exceeds that for docking against apo structures, and AutoDock-GPU is capable of identifying to correct model by docked pose energy such that these two structures may be considered jointly and still lead to an improved success rate over apo. While the score is sufficiently accurate such that success rate is not sacrificed to take the best pose from both models, it is not perfect for this discrimination and there is a higher still success rate when considering the best pose from each. In light of these results, we recommend and will adopt as best practice the use of AF2 structures alongside apo structures when predicting binding modes. This is in broad agreement with the findings of a study on the screening power of AF2-predicted structures in the context of a smaller set of 38 proteins (Zhang et al., 2022), and with a study that suggests AF2 has a tendency to predict holo structures (Saldaño et al., 2022), and in the case of GPCRs can be biased to produce the desired conformation (Heo & Feig, 2022).

We did find that in some systems where naive docking against the AF2 predicted structure was unsuccessful, the docking results can be substantially improved by straightforward manipulations of the receptor. Flexible docking was capable of resolving side chain clashes that disrupted ligand binding, and removing low-confidence regions of the predicted structure cleared occluded binding sites. While these are challenging to address in a systematic way that would mimic a prospective screen, it suggests that researchers working on an individual target, or small set, may be capable of improving on our initial performance estimation. This could include special treatment of residues in the binding site that have been identified as important in biochemical studies, and careful attention to the possible contributions of cofactors, which would be expected to reduce success rates in docking to AF2 structures.

4 | CONCLUSIONS

Docking as part of computer-guided drug design is emerging as one of the major potential applications of computed structural models like those predicted by AF2. This study reveals that naive docking to AF2 structures is currently expected to have limited success, due to challenges with local sidechain conformation and potential presence of cofactors. These challenges, however, may be ameliorated through careful application of refinements in the docking, such as addition of flexibility to key sidechains or careful manual curation of regions of low confidence in the predicted structures. Even absent these fixes, AF2 structures outperform apo structures as docking targets, and can serve as complements to them in determining binding modes. We expect that, while we did this analysis using AutoDock, the general principles are transferable to other docking programs, and more generally, to other molecular modeling tools.

4.1 | Availability

The AutoDock suite is freely available at autodock.scripps.edu.

5 | METHODS

The refined set of structures in the PDBbind database for year 2020 was downloaded (http://www.pdbbind.org.cn). Structures were retained for which corresponding AF2 structures were available from the EMBL database in January 2022, which at the time were all human proteins, based on Uniprot ID (https://alphafold.ebi.ac.uk). The predicted structures were aligned to the crystallographic reference structures using a pairwise alignment between C_α atoms in the binding pocket, as deposited in the PDBbind, using Biopython v1.79 (Cock et al., 2009). To correct for gaps in the sequence, a sequence alignment was performed using the pairwise2 module in Biopython. Two points were assigned for matched residues, one point deducted for misaligned residues, half a point deducted for opening new gaps, and two tenths of a point deducted for extending a gap. The alignment RMSDs were calculated using Biopython (Cock et al., 2009).

Waters and nonpolar hydrogens were removed. Both the crystallographic and predicted structures were processed by pdb4amber (Case et al., 2022) to correct protonation states and replace any missing atoms. Partial charges were assigned using prepare_receptor4.py within AutoDockTools (Forli et al., 2016; Morris et al., 2009). Ligands were prepared using OpenBabel v2.4.1 (O'Boyle et al., 2011) to assign protonation states and Meeko v0.3.2 (Meeko, n.d.) to assign partial charges and rotatable bonds.

AutoDock maps were generated using AutoGrid4, with 8 Å of padding around the crystallographic ligand. Docking was performed using AutoDock-GPU v1.5.3 (Santos-Martins et al., 2021) with default parameters. Briefly, for each complex 20 genetic algorithm runs were run, with the resulting conformations clustered using a soft RMSD tolerance of 2 Å. The number of evaluations was capped using a built-in heuristic based on the number of rotatable bonds (Solis-Vasquez et al., 2022), with

an asymptotic limit at 12 million evaluations. Convergence was automatically assessed by the AutoStop criterion based on the standard deviation of the energy evaluations (Solis-Vasquez et al., 2022). Default settings for AutoStop of a five-generation test rate and an energy standard deviation of 0.15 kcal/mol were used.

RMSDs relative to the crystallographic ligand position were calculated using the CalcRMS function available in rdkit v2022.03.4 (Rdkit, n.d.), which accounts for ligand symmetry.

AUTHOR CONTRIBUTIONS

Matthew Holcomb: Conceptualization (equal); data curation (lead); formal analysis (lead); visualization (equal); writing – original draft (equal); writing – review and editing (equal). **Ya-Ting Chang:** Investigation (supporting). **David S. Goodsell:** Conceptualization (lead); investigation (lead); writing – original draft (equal); writing – review and editing (equal). **Stefano Forli:** Conceptualization (lead); funding acquisition (lead); investigation (lead); methodology (lead); project administration (lead); writing – original draft (equal); writing – review and editing (equal).

ACKNOWLEDGMENTS

Over the past 30 years, AutoDock development has been continuously supported by the National Institute of General Medical Sciences-NIH, most recently by grant GM069832 (SF).

ORCID

David S. Goodsell https://orcid.org/0000-0002-5932-2130
Stefano Forli https://orcid.org/0000-0002-5964-7111

REFERENCES

Aggarwal R, Gupta A, Priyakumar UD. 2021. Apobind: a dataset of ligand unbound protein conformations for machine learning applications in de novo drug design. ArXiv. abs/2108.09926.

Bordogna A, Pandini A, Bonati L. Predicting the accuracy of protein–ligand docking on homology models. J Comput Chem. 2011;32(1):81–98.

Callaway E. 'The entire protein universe': Ai predicts shape of nearly every known protein. Nature. 2022;608:15–6. https://doi.org/10.1038/d41586-022-02083-2

Case DA, Aktulga HM, Belfon K, Ben-Shalom IY, Berryman JT, Brozell SR, et al. Amber. San Francisco, CA: University of California; 2022.

Cock PJA, Antao T, Chang JT, Chapman BA, Cox CJ, Dalke A, et al. Biopython: freely available python tools for computational molecular biology and bioinformatics. Bioinformatics. 2009;25(11):1422–3.

Fan H, Irwin JJ, Webb BM, Klebe G, Shoichet BK, Sali A. Molecular docking screens using comparative models of proteins. J Chem Inf Model. 2009;49(11):2512–27.

Forli S, Huey R, Pique ME, Sanner MF, Goodsell DS, Olson AJ. Computational protein-ligand docking and virtual drug screening with the autodock suite. Nat Protoc. 2016;11(5):905–19.

Heo L, Feig M. Multi-state modeling of g-protein coupled receptors at experimental accuracy. Proteins. 2022;90(11):1873–85.

Huey R, Morris GM, Olson AJ, Goodsell DS. A semiempirical free energy force field with charge-based desolvation. J Comput Chem. 2007;28(6):1145–52.

Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with alphafold. Nature. 2021;596(7873):583–9.

McGovern SL, Shoichet BK. Information decay in molecular docking screens against holo, apo, and modeled conformations of enzymes. J Med Chem. 2003;46(14):2895–907.

Meeko. https://github.com/forlilab/Meeko. (n.d.). Accessed May 2022.

Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, et al. Autodock4 and autodocktools4: automated docking with selective receptor flexibility. J Comput Chem. 2009;30(16):2785–91.

O'Boyle NM, Banck M, James CA, Morley C, Vandermeersch T, Hutchison GR. Open babel: an open chemical toolbox. J Chem. 2011;3(1):33.

Rdkit. https://www.rdkit.org/. (n.d.). Accessed May 2022.

Sabe VT, Ntombela T, Jhamba LA, Maguire GEM, Govender T, Naicker T, et al. Current trends in computer aided drug design and a highlight of drugs discovered via computational techniques: a review. Eur J Med Chem. 2021;224:113705.

Saldaño T, Escobedo N, Marchetti J, Zea DJ, Mac Donagh J, Velez Rueda AJ, et al. Impact of protein conformational diversity on alphafold predictions. Bioinformatics. 2022;38(10):2742–8.

Santos-Martins D, Solis-Vasquez L, Tillack AF, Sanner MF, Koch A, Forli S. Accelerating autodock4 with gpus and gradient-based local search. J Chem Theory Comput. 2021;17(2):1060–73.

Solis-Vasquez L, Tillack AF, Santos-Martins D, Koch A, LeGrand S, Forli S. Benchmarking the performance of irregular computations in autodock-GPU molecular docking. Parallel Comput. 2022;109:102861.

Su M, Yang Q, Du Y, Feng G, Liu Z, Li Y, et al. Comparative assessment of scoring functions: the CASF-2016 update. J Chem Inf Model. 2019;59(2):895–913.

Varadi M, Anyango S, Deshpande M, Nair S, Natassia C, Yordanova G, et al. Alphafold protein structure database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. Nucleic Acids Res. 2022;50(D1):D439–44.

Wang R, Fang X, Lu Y, Yang C-Y, Wang S. The pdbbind database: methodologies and updates. J Med Chem. 2005;48(12):4111–9.

Zhang Y, Vass M, Shi D, Abualrous E, Chambers J, Chopra N, et al. 2022. Benchmarking refined and unrefined alphafold2 structures for hit discovery. ChemRxiv.

How to cite this article: Holcomb M, Chang Y-T, Goodsell DS, Forli S. Evaluation of AlphaFold2 structures as docking targets. Protein Science. 2023;32(1):e4530. https://doi.org/10.1002/pro.4530