nature portfolio

Peer Review File

Distance-AF Improves Predicted Protein Structure Models by AlphaFold2 with User-Specified Distance Constraints

Corresponding Author: Professor Daisuke Kihara

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
This manuscript reports a new deep learning method to adjust AlphaFold2-predicted protein tertiary structures to fit user-provided residue-residue distances. This kind of method is needed because in some cases AlphaFold2 fails to predict the orientation and relative distances between domains of multi-domain proteins even though the structures of individual domains may be correctly predicted. The method is also different from other methods by directly optimizing a loss function based on the users-provided distance restraints as well as some existing loss function of AlphaFold2. Moreover, it works substantially better than two related methods, AlphaFoldLink and Rosetta, that address the similar problem on a test dataset. Moreover, the authors convincingly demonstrate how the method can be used to adjust AlphaFold2-predicted structures to fit cryo-EM density maps better, to correctly predict active and inactive states of G protein and generate dynamic conformations from NMR data. Overall, the method is very interesting and useful.

Minor comments for revision:
(1) The loss function of the deep network depends on the definition of domains. Can you discuss how sensitive the performance of the method is to the domain definition in the discussion section?
(2) Some labels of subfigures in Figure 2 are not consistent how the subfigures are cited in the text, even though readers may be able to figure out which sub-figures are referred from the textual context.

Reviewer #2

(Remarks to the Author)
Abstract:
1- Please ensure consistent use of 'AlphaFold2' throughout the manuscript (not 'Alphafold2').
2- The abstract could be strengthened by including specific details about the model's accuracy before and after using the approach and how it compares to existing methods.

Main Text:
1- Given the release of AlphaFold3, the authors should acknowledge its capabilities and provide a clear rationale for focusing on AlphaFold2. This could include reasons such as AlphaFold2's continued relevance in certain applications, its accessibility, or specific advantages pertinent to the study's objectives. Providing this context will help readers understand the study's positioning within the current landscape of protein structure prediction research.

Results:
Overview of Distance-AF
1- "Distance-AF builds on the AF2 network architecture to predict protein structures while incorporating distance constraints." Please add a citation for AlphaFold2 (AF2) in this sentence, and ensure that all key methods and tools are properly cited throughout the paper.
2- The manuscript states that "Users can input a few distance constraints between specified domains," but it does not clearly describe how this is done in practice. For reproducibility and usability, the authors should clarify the following:
a) What method is used to input these constraints (e.g., command line arguments, configuration file, graphical interface)?
b) What is the required format and unit of the input (e.g., residue indices, distances in Ångströms)?

c) Are there any examples, templates, or documentation provided to guide users?
d) Most importantly, how can users define/choose and implement these constraints in the workflow? A step-by-step explanation or illustrative example would significantly improve clarity.
3- In Fig. 1a, could you please clearly highlight both the original AF2 architecture and the proposed Distance-AF? Since this section states that Distance-AF is built on the AF2 network, it would be clearer if the figure visually distinguished between the two architectures.
4- “As default, a structure is optimized under Eq. 2 for 30,000 iterations.” How was the number of iterations determined? Please clarify the rationale or criteria used for selecting this value.
5- “In Fig. 1b and 1c, we show an example…..”. It would be helpful to show the structural accuracy in this example both before and after applying the proposed method. Please ensure this information is clearly described in the main text—not just in the figure caption—and indicate how the accuracy was estimated, including citations for any tools or metrics used.

Performance on incorrect AF2 models
6- “From them, we selected AF2 models which have an RMSD over 10 Å to the
corresponding PDB structure and have an average pLDDT of over 0.75.”. How the structures with RMSD > 10Å and pLDDT over 0.75 were selected for this analysis? Does the AlphaFold Database provide any filtering options for these metrics, or were they computed separately? Please add more details in the manuscript.
7- “Models were eliminated if they have long disordered regions, domain-swapping, and local domain structures which are entirely incorrect.” Similarly, more details needed regarding this filtering process. For example, how structure with long disordered regions are determined?
8- Please ensure that details are provided on how structural accuracy was evaluated before and after refinement using the 4 metrics, including the specific metrics used and how they were computed.

Examples of models built by Distance-AF
9- “This 258-amino-acid protein consists of two domains: the N-terminal domain…”. It would be helpful to label the N-terminal and C-terminal on the structures in the figures for better clarity.

Application to structure fitting in cryo-EM maps
10- “In structure modeling for a cryo-EM map, it is common to fit predicted models by AF2 to the map density48”. Fitting AF2 structures to cryo-EM maps is an important and actively evolving topic. It would strengthen the paper to cite more recent and relevant studies in this area—for example, recent advances or applications published in the last couple of years:

• Chen, Jason, et al. "Protein Structure Refinement via DeepTracer and AlphaFold2." bioRxiv (2023): 2023-08.

• Gao, Jerry, et al. "DomainFit: Identification of protein domains in cryo-EM maps at intermediate resolution using AlphaFold2-predicted models." Structure 32.8 (2024): 1248-1259.

• Alshammari, Maytha, Jing He, and Willy Wriggers. "Flexible fitting of AlphaFold2-predicted models to cryo-EM density maps using elastic network models: a methodical affirmation." Bioinformatics Advances 5.1 (2025): vbae181.

Application to local structural modification
11- In addition to presenting and explaining the results, the authors could strengthen the manuscript by further analyzing the significance and impact of the observed enhancements. Please consider applying this approach consistently across all sections of the Results.

Discussion:
1- To strengthen the discussion, consider more explicitly tying your key findings back to the methods used and the results obtained. Highlight how your methodological approach contributed to improved outcomes, and clearly articulate how your results compare with existing approaches in the field. Emphasizing the novelty and advantages of your method, supported by the comparative performance metrics, will help reinforce the significance of your contribution. Additionally, reflection on potential applications and how your method advances the field would leave the reader with a stronger appreciation of the study’s broader impact.

References:
1- Please check the references for consistency in formatting, accuracy, and completeness. Ensure all cited works are included in the reference list and that the citation style follows the journal’s guidelines.

Methods:
1- Since this is a deep learning-based method, it would be helpful to include information on the number of training, validation, and testing sets used in the study. Additionally, please provide details on the data splitting strategy (e.g., cross-validation, stratified sampling) and any techniques employed to avoid overfitting, such as regularization or dropout. Including these details will help readers assess the robustness, generalizability, and performance of the model.
2- It would strengthen the manuscript to include a brief justification for the chosen method over alternative approaches, especially if there are known trade-offs. Doing so will enhance clarity and allow readers to better assess the rigor and rationale behind the study design.

Reviewer #3

(Remarks to the Author)
This manuscript proposes Distance-AF, a post-processing method designed to enhance AlphaFold2-predicted protein structures by introducing user-defined inter-residue distance constraints. The method begins with an AlphaFold2-generated structure and iteratively modifies it to better satisfy specified distance relationships. The authors demonstrate its application in several scenarios, including cryo-EM map fitting, GPCR conformational modeling, and NMR ensemble generation. While the method is promising and addresses meaningful challenges in current structure prediction workflows, several areas require clarification or additional detail. These include methodological transparency, dataset selection, and comparative evaluation. I outline specific points below that, if addressed, would help enhance the clarity and rigor of the manuscript.

1. There is a lack of clarification on whether training or parameter tuning was performed using a separate training set or validation data. While the method appears to rely solely on test-time loss minimization of AlphaFold2 outputs, it remains unclear whether any data was used for hyperparameter selection (e.g., loss weighting scheme). For better methodological transparency and reproducibility, it is recommended that the authors clarify the details of the training and parameter tuning.

2. AlphaFold3 has recently been released. The authors should discuss how their method compares to AlphaFold3. In addition, the comparison with other SOTA structure refinement or multi-conformation generation methods is lacking and should be supplemented.

3. The current evaluation is based on 25 AlphaFold2-predicted structures with domain-level errors. However, whether this selection adequately covers a diverse range of protein classes is unclear. There may be an implicit dataset bias toward specific protein types. To better assess the generalizability of Distance-AF, the authors are encouraged to extend the evaluation to a broader and more diverse test set.

4. The manuscript claims that Distance-AF is efficient in time and memory usage. However, no quantitative comparisons are provided. Please specify the average runtime per protein. How does this compare to AlphaFold2 inference or Rosetta refinement?

5. The manuscript does not clearly explain how user-defined distance constraints are generated. Are they manually annotated from experimental data (e.g., crosslinking, Cryo-EM maps, NMR NOEs) or predicted from AI-based methods? Please clarify the format users should follow for providing constraints and whether the method accommodates uncertainty or noise in those constraints.

6. In cases where multiple user-specified distance constraints may conflict with one another (e.g., forming geometric loops or exceeding possible backbone stretch), how does the method handle constraint prioritization or infeasibility?

7. Figure 1 does not sufficiently explain how structures evolve over optimization steps. Please enhance the figure caption with descriptions of key changes (e.g., movement of domains, change in inter-residue distances, color mapping).

8. While Distance-AF achieves the target distances, it is unclear whether the resulting models maintain acceptable stereochemistry. For example, do the final structures exhibit violations in bond angles, steric clashes, or Ramachandran outliers? Such checks should be included to confirm biological plausibility.

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
The authors addressed my comments well.

Reviewer #2

(Remarks to the Author)
I have reviewed the revised manuscript, and I'm pleased to see that the author has thoroughly addressed all of the comments and suggestions I previously provided. The revisions have significantly improved the clarity and quality of the work.

Everything looks good, and I have no further concerns. I recommend the manuscript for acceptance.

Reviewer #3

(Remarks to the Author)
The authors addressed all my comments.

Responses to Comments by Reviewer #1

This manuscript reports a new deep learning method to adjust AlphaFold2-predicted protein tertiary structures to fit user-provided residue-residue distances. This kind of method is needed because in some cases AlphaFold2 fails to predict the orientation and relative distances between domains of multi-domain proteins even though the structures of individual domains may be correctly predicted. The method is also different from other methods by directly optimizing a loss function based on the users-provided distance restraints as well as some existing loss function of AlphaFold2. Moreover, it works substantially better than two related methods, AlphaFoldLink and Rosetta, that address the similar problem on a test dataset. Moreover, the authors convincingly demonstrate how the method can be used to adjust AlphaFold2-predicted structures to fit cryo-EM density maps better, to correctly predict active and inactive states of G protein and generate dynamic conformations from NMR data. Overall, the method is very interesting and useful.

Minor comments for revision:
(1) The loss function of the deep network depends on the definition of domains. Can you discuss how sensitive the performance of the method is to the domain definition in the discussion section?

As the additional input of Distance-AF is the distances between domains, or more precisely, residue pairs (that are usually taken across domains), the modeling accuracy is in principle influenced by the accuracy of the input information. Since the provided information is actual distances, it would be more accurate to discuss how the performance is affected by the accuracy of the distances instead of domain definition. And actually, we did investigate how the structure modeling accuracy is affected by perturbating the distance information in Fig. 2b. The results was discussed on page 13.

We further added a following paragraph in the discussion section:

> Distance-AF works best when specified distance constraints are between two structured domains and linear move of domains are required to reach the target conformation. The modeling results are in principle affected by the accuracy of the provided distance constraints, but as shown in Fig. 2b, the results were reasonably robust to perturbations of the distances. Modelling is challenging if pulling domains cause severe atom clashes or substantial rotation is needed for achieving the target conformation.

(2) Some labels of subfigures in Figure 2 are not consistent how the subfigures are cited

in the text, even though readers may be able to figure out which sub-figures are referred from the textual context.

Thank you for pointing it out. We have fixed this problem.

Responses to Comments by Reviewer #2:

Abstract:
1- Please ensure consistent use of 'AlphaFold2' throughout the manuscript (not 'Alphafold2').

Thank you for pointing out. We corrected them.

2- The abstract could be strengthened by including specific details about the model's accuracy before and after using the approach and how it compares to existing methods.

We added sentences describing results in Abstract, while keeping the length to 150 words as required in the author guideline.

Main Text:
1- Given the release of AlphaFold3, the authors should acknowledge its capabilities and provide a clear rationale for focusing on AlphaFold2. This could include reasons such as AlphaFold2's continued relevance in certain applications, its accessibility, or specific advantages pertinent to the study's objectives. Providing this context will help readers understand the study's positioning within the current landscape of protein structure prediction research.

AlphaFold3 uses diffusion model as the network architecture used to generate the 3D structure of target proteins. The diffusion model algorithm has randomness in nature, and, although possible, it is not straightforward to apply fixed values of distance constraints in the structure modeling process. On the other hand, AlphaFold2 is a supervised learning-based method, which is aimed at optimizing (minimizing) a loss function, and it is more suitable and understandable to add distance constraints as one of the loss function terms and balance it with other terms in the original loss function.

We added this explanation on page 4.

Results:
Overview of Distance-AF
1- “Distance-AF builds on the AF2 network architecture to predict protein structures while incorporating distance constraints.” Please add a citation for AlphaFold2 (AF2) in this sentence, and ensure that all key methods and tools are properly cited throughout the paper.

AlphaFold2 was cited in earlier place, page 3, but I agree that it is good to cite in the Result section also. We cited it in page 5 as suggested, and also page 8, when FAPE loss is first mentioned, and in page 34. In page 34 we cited Rosetta and AlphaLink paper, because it was the first place they appear in the Method section, although they are cited in an earlier page.

2- The manuscript states that “Users can input a few distance constraints between specified domains,” but it does not clearly describe how this is done in practice. For reproducibility and usability, the authors should clarify the following:

a) What method is used to input these constraints (e.g., command line arguments, configuration file, graphical interface)?

Thank you for the comment. The example files are provided in the GitHub repository, https://github.com/kiharalab/Distance-AF, as mentioned in the Code availability. And example files and instruction is given there.

To provide the explanation in the paper, now we added a new section, “Running Distance-AF” in the Method section, a section before the “Running existing methods”. In page 5, where you mentioned, we added sentences to guide readers to refer to the “Running Distance-AF” section.

To answer the specific question a), distance constraints will be provided in a configuration file. We mentioned it in the “Running Distance-AF” section (page 34).

b) What is the required format and unit of the input (e.g., residue indices, distances in Ångströms)?

As you wrote, the format is residue indices, and the distance between them. We added this information in the Running Distance-AF section (page 34). An example file is provided in the Github.

c) Are there any examples, templates, or documentation provided to guide users?

Yes, in the Github repository, example input files and an output file are provided in the Example/ folder. On the top page of Github, we wrote the instruction in the "Usage" section. We mentioned this in the Running Distance-AF section (page 34), and briefly explained the steps.

d) Most importantly, how can users define/choose and implement these constraints in the workflow? A step-by-step explanation or illustrative example would significantly improve clarity.

Constraints users want to specify need to be written in the distance constraint file. The example of the file is provided in the Github repository, https://github.com/kiharalab/Distance-AF/tree/main/Example/1IXCA. It is dist_constraint.txt. As also answered for the question a) above, the format is the indices of the two residues, and the distance between them.

Users are expected to obtain such distance information from experiments, such as crosslinking, or by fitting the current structure model to an EM map, from literature, or from an idea by the users based on their expert knowledge of the protein they want to model. Distance-AF is a tool to make the structure with the provided distance constraints, which can be visualized and examined by the users to confirm or design biological hypotheses.

We added this explanation in the Running Distance-AF section (page 34).

3- In Fig. 1a, could you please clearly highlight both the original AF2 architecture and the proposed Distance-AF? Since this section states that Distance-AF is built on the AF2 network, it would be clearer if the figure visually distinguished between the two architectures.

There is no difference in the network model architecture between AF2 and Distance-AF. The difference between AF2 and Distance-AF lies in their loss function. Distance-AF

works as overfitting of a predicted (incorrect) model by AF2, by adding distance loss (Eq. 1). In addition, the FAPE loss, which in principle makes correct main-chain structures, was modified. The loss function is shown as a part of Fig. 1a, but we now added Supplementary Figure 2 that illustrates the loss function. Explanation of each loss terms are added in the figure caption of Suppl. Fig. 2. It is referred to in page 8.

4- “As default, a structure is optimized under Eq. 2 for 30,000 iterations.” How was the number of iterations determined? Please clarify the rationale or criteria used for selecting this value.

Optimization needs to be performed until the overall loss function (Eq. 2), particularly the distance loss (Eq. 1) is saturated. For example, in Fig. 1b and 1c, we discussed that the loss was almost converged at iteration 1000, and after iteration 1000, Distance-AF continues refining local structures until the end of iterations (page 9). The iteration needed for convergence is larger for larger proteins. Based on such large proteins, we set 30,000 as a default max iterations, although now as mentioned in the text, usually the iteration is terminated much earlier if we see the convergence of the distance loss. We added the explanation in page 9.

5- “In Fig. 1b and 1c, we show an example…..”. It would be helpful to show the structural accuracy in this example both before and after applying the proposed method. Please ensure this information is clearly described in the main text—not just in the figure caption—and indicate how the accuracy was estimated, including citations for any tools or metrics used.

We added the accuracy data in the main text (page 9): “In this example, RMSD improved from 20.65 Å to 4.27 Å by applying Distance-AF in comparison with the native structure (PDB ID: 6P66). “In caption, the resulting accuracy was already written. The tool used to compute RMSD is TM-align. We added it in page 9.

Performance on incorrect AF2 models

6- “From them, we selected AF2 models which have an RMSD over 10 Å to the corresponding PDB structure and have an average pLDDT of over 0.75.”. How the structures with RMSD > 10Å and pLDDT over 0.75 were selected for this analysis? Does the AlphaFold Database provide any filtering options for these metrics, or were they computed separately? Please add more details in the manuscript.

AlphaFold Database does not provide such an option. We computed RMSD between AlphaFold Database (AFDB) entries (predicted models by AlphaFold2) with corresponding PDB entry by ourselves. We identified AFDB entries that have a corresponding PDB structure (not all the AFDB entries have experimentally known structures in PDB). Then, we computed RMSD between them. Particularly, we provided details about how we selected protein models that have two domains. We added more explanations in page 10.

7- “Models were eliminated if they have long disordered regions, domain-swapping, and local domain structures which are entirely incorrect.” Similarly, more details needed regarding this filtering process. For example, how structure with long disordered regions are determined?

We added more explanations in page 9. The detailed structural criteria, including the long disordered regions, were visually examined using Pymol and counted if the length of disordered regions are more than about 30 residues.

8- Please ensure that details are provided on how structural accuracy was evaluated before and after refinement using the 4 metrics, including the specific metrics used and how they were computed.

We reported the four metrics, RMSD, TM-score, GDT-TS, and GDT-HA, and the software we used, TM-align and Pymol in page 9.

Examples of models built by Distance-AF

9- “This 258-amino-acid protein consists of two domains: the N-terminal domain…”. It would be helpful to label the N-terminal and C-terminal on the structures in the figures for better clarity.

The labels are added to clarify N-terminal and C-terminal in Figure 3a.

Application to structure fitting in cryo-EM maps

10- “In structure modeling for a cryo-EM map, it is common to fit predicted models by AF2 to the map density48”. Fitting AF2 structures to cryo-EM maps is an important and actively evolving topic. It would strengthen the paper to cite more recent and relevant studies in this area—for example, recent advances or applications published in the last couple of years:

• Chen, Jason, et al. "Protein Structure Refinement via DeepTracer and AlphaFold2." bioRxiv (2023): 2023-08.
• Gao, Jerry, et al. "DomainFit: Identification of protein domains in cryo-EM maps at intermediate resolution using AlphaFold2-predicted models." Structure 32.8 (2024): 1248-1259.
• Alshammari, Maytha, Jing He, and Willy Wriggers. "Flexible fitting of AlphaFold2-predicted models to cryo-EM density maps using elastic network models: a methodical affirmation." Bioinformatics Advances 5.1 (2025): vbae181.

Thank you, we cited these 3 papers.

Application to local structural modification
11- In addition to presenting and explaining the results, the authors could strengthen the manuscript by further analyzing the significance and impact of the observed enhancements. Please consider applying this approach consistently across all sections of the Results.

Thank you for pointing this out. GPCR is an important class of drug target, which shares over 30% of drug targets for which drugs were developed so far. It is known that the active and inactive form of the GCPR have conformational difference, which consists of changes of alpha helices, i.e. local conformational difference, rather than large global move. Here we used GPCR as examples to demonstrate that Distance AF is able to model active and inactive forms from other functional states. We added the explanation in page 22, at the beginning of this section.

We also went over other sections and modified titles or added description of purpose of the experiments in some sections.

Discussion:
1- To strengthen the discussion, consider more explicitly tying your key findings back to the methods used and the results obtained. Highlight how your methodological approach contributed to improved outcomes, and clearly articulate how your results compare with existing approaches in the field. Emphasizing the novelty and advantages of your method, supported by the comparative performance metrics, will help reinforce the significance of your contribution. Additionally, reflection on potential applications and how your method advances the field would leave the reader with a stronger appreciation of the study's broader impact.

Thank you for the comment. In discussion, we made a new paragraph to emphasize the novelty and approaches, and summarized performance comparison over other methods. Also, we added a new paragraph to summarize applications (page 28).

References:
1- Please check the references for consistency in formatting, accuracy, and completeness. Ensure all cited works are included in the reference list and that the citation style follows the journal's guidelines.

We checked the reference format.

Methods:
1- Since this is a deep learning-based method, it would be helpful to include information on the number of training, validation, and testing sets used in the study. Additionally, please provide details on the data splitting strategy (e.g., cross-validation, stratified sampling) and any techniques employed to avoid overfitting, such as regularization or dropout. Including these details will help readers assess the robustness, generalizability, and performance of the model.

As originally mentioned in page 4, Distance-AF was not trained using a training set and validation set as usually performed in supervised learning. This is very different from most of the deep learning-based methods that are trained on a training set in a supervised fashion to achieve good performance on the training set. Distance-AF is not pre-trained. The weights of the neural network is the same as the original AlphaFold2, which Distance-AF is based on. Instead of pre-training, Distance-AF adds an additional scoring term, the distant-constraint loss (Eq. 1) to the loss function, to the other loss terms as shown in Eq. 2. To answer your earlier question, these loss functions are illustrated in Supplementary Fig. 2. Thus, there is no training, validation sets used. For a given target, this loss function is optimized in an "overfitting" manner. The advantages of this approach over conventional supervised training-based approach is that the method is not biased to the training set used (because no training set is used). This is very important because for this work, there is no large-amount of existing appropriate distance data we can used for training.

This was originally mentioned on page 4, but we agree that this is not clearly stated. Thus, as pointed out, we added explanation in the Method section as a new paragraph

named “Structure modeling with constraints as an overfitting problem”. (page 30). We also added an explanation earlier in the main text, page 7-8.

“Settings of Computation” section in Methods describes some details of the optimization (overfitting) procedure of the loss function. We also mentioned there that L2 regularization and drop out were used (page 33).

2- It would strengthen the manuscript to include a brief justification for the chosen method over alternative approaches, especially if there are known trade-offs. Doing so will enhance clarity and allow readers to better assess the rigor and rationale behind the study design.

In Discussion (page 28), we now added the advantages of this approach over conventional supervised learning-based approach.

Responses to Comments by Reviewer #3:

This manuscript proposes Distance-AF, a post-processing method designed to enhance AlphaFold2-predicted protein structures by introducing user-defined inter-residue distance constraints. The method begins with an AlphaFold2-generated structure and iteratively modifies it to better satisfy specified distance relationships. The authors demonstrate its application in several scenarios, including cryo-EM map fitting, GPCR conformational modeling, and NMR ensemble generation. While the method is promising and addresses meaningful challenges in current structure prediction workflows, several areas require clarification or additional detail. These include methodological transparency, dataset selection, and comparative evaluation. I outline specific points below that, if addressed, would help enhance the clarity and rigor of the manuscript.

1. There is a lack of clarification on whether training or parameter tuning was performed using a separate training set or validation data. While the method appears to rely solely on test-time loss minimization of AlphaFold2 outputs, it remains unclear whether any data was used for hyperparameter selection (e.g., loss weighting scheme). For better methodological transparency and reproducibility, it is recommended that the authors clarify the details of the training and parameter tuning.

As the reviewer mentioned, this method works as loss minimization of AF2 output, and no training was performed. For clarification, we added more explanation on page 7-8, and in the Method section we added a new paragraph, “Structure modeling with constraints as an overfitting problem” in page 29. Parameter details are also provided in page 33, “Settings of Computation”. We also added illustration of loss function terms in Supplementary Fig. S2.

In Eq. 2, loss terms were weighted depending on the stage of the distance loss minimization. The idea of using different weights came from observation of the behavior of intermediate model conformations in a very early stage of the development of this method for a very small number of cases (2 - 3 cases) used at that time. We mentioned it on page 9.

2. AlphaFold3 has recently been released. The authors should discuss how their method compares to AlphaFold3. In addition, the comparison with other SOTA structure refinement or multi-conformation generation methods is lacking and should be supplemented.

Thank you for this comment. We now ran AlphaFold3 and AlphaFold2-Conformations, which is the SOTA method for multi conformation generation. With AlphaFold2-Conformations, we generated 130 conformations and reported the results by selecting the best (i.e. smallest RMSD) models and also by selecting a model by its pLDDT confidence score. For almost all the cases, AF3 and AlphaFold2-Conformations performed similarly with AF2, and Distance-AF performed substantially better than them. The results are shown in the new Supplementary Fig 4 and discussed on page 14.

3. The current evaluation is based on 25 AlphaFold2-predicted structures with domain-level errors. However, whether this selection adequately covers a diverse range of protein classes is unclear. There may be an implicit dataset bias toward specific protein types. To better assess the generalizability of Distance-AF, the authors are encouraged to extend the evaluation to a broader and more diverse test set.

First, we confirmed that the 25 targets used in Fig. 2 have less than 25% sequence identity between each other, to ensure that they are non-redundant and a diverse set. The sequence identity distribution is shown in Supplementary Fig. 3, and it is mentioned on page 10.

Distance-AF is not in the same category as general protein structure prediction methods, such as AlphaFold, which are meant to be applicable for any proteins in general. Distance-AF is a specialized tool that can be used in diverse research scenarios when distance constraints can help improving model accuracy. In the original manuscript, we showed 4 scenarios; correcting AF models, cryo-EM, functional states of GPCR, and NMR to show examples of diverse scenarios. In addition, we now also added another scenario when AF2 typically has problem in modeling, which is disordered regions. In page 19, we showed cases where disordered tail of a protein structure is substantially misplaced in AF2 models in comparison with the experimental structure. Structures are shown in the new Figure 4. These are more extreme cases compared to the previous section with Fig. 3, because in these cases, the tail is in the experimentally determined structure, but an intermediate loop region is even missing in the experimental structure due to flexible, disordered nature of the proteins. These terminal regions are predicted in a very different place by AF2, but Distance-AF was able to put it in the right place.

4. The manuscript claims that Distance-AF is efficient in time and memory usage. However, no quantitative comparisons are provided. Please specify the average runtime per protein. How does this compare to AlphaFold2 inference or Rosetta refinement?

We provided computational time and GPU memory of proteins with different lengths in Supplementary Table 11. It is mentioned on page 34.

5. The manuscript does not clearly explain how user-defined distance constraints are generated. Are they manually annotated from experimental data (e.g., crosslinking, Cryo-EM maps, NMR NOEs) or predicted from AI-based methods? Please clarify the format users should follow for providing constraints and whether the method accommodates uncertainty or noise in those constraints.

As you mentioned, primarily we expect distance information is collected from experimental data, such as crosslinking, Cryo-EM maps, NMR NOEs, or residue interactions. Or it can be also come from biological hypothesis users have, and Distance-AF will build a structure model based on such hypothetical distance information, from which users can further deepen or reject their hypothesis. We added the explanation on page 34, and also on page 4.

Regarding the format of distance constraints: the format is residue indices, and the distance between them. An example file is provided in the Github. We added this information in the Running Distance-AF section (page 34).

Regarding the noise in the distance constraints, we did have the data in Fig. 2b and described on page 14. Here, we perturbated distance constraints randomly by up to 5 Angstroms. Despite these large perturbations, modeling showed robustness; For example, 62.8% of the models remained within an RMSD of 6.0 Å, retaining the correct overall fold, even when the average perturbation was up to 2.0 Å.

6. In cases where multiple user-specified distance constraints may conflict with one another (e.g., forming geometric loops or exceeding possible backbone stretch), how does the method handle constraint prioritization or infeasibility?

It is a good point. Currently, all the distances are considered equally because the sum of the distance difference is considered as the loss function as shown in Eq. 1 (Page 8). So if some inconsistent distances are included in the constraints, the structure would optimize mainly a part of them. If each distance information has a different reliability level, we can easily add a weight to emphasize some distance over less confident ones. In the discussion section, we added this as a limitation of the method and mentioned it as a future improvement. We added

"In the current implementation, provided distance constraints are considered as a whole by the sum of the distance difference in a current structure (Eq. 1). Instead, we plan to give a weight to each distance to allow prioritizing some of the distances over others, depending on the importance or confidence." (page 28)

7. Figure 1 does not sufficiently explain how structures evolve over optimization steps. Please enhance the figure caption with descriptions of key changes (e.g., movement of domains, change in inter-residue distances, color mapping).

In page 10 and 11, we expanded the description of Fig. 1b (changes of individual loss terms) in correlation of the structure image in Fig. 1c.

8. While Distance-AF achieves the target distances, it is unclear whether the resulting models maintain acceptable stereochemistry. For example, do the final structures exhibit violations in bond angles, steric clashes, or Ramachandran outliers? Such checks should be included to confirm biological plausibility.

That is a good point. Since the loss function of Distance-AF includes the violation loss (Eq. 2) which penalizes such non-physical conformations and atom positions, there are no severe violations. We now added MolProbity score and the atom clash scores of MolProbity in Supplementary Fig. 5. We further showed that molecular dynamics relaxation even makes the scores better than the starting AF2 models with almost no change in the global conformations. We mentioned this in a new paragraph on page 14.