Integrating AlphaFold and deep learning for atomistic interpretation of cryo-EM maps

Xin Dai[1,*], Longlong Wu[2], Shinjae Yoo[1], and Qun Liu[3,*]

[1]Computational Science Initiative, Brookhaven National Laboratory, Upton, NY, USA
[2]Condensed Matter Physics and Materials Science Department, Brookhaven National Laboratory, Upton, NY, USA
[3]Biology Department, Brookhaven National Laboratory, Upton, NY, USA
[*]corresponding to: xdai@bnl.gov or qunliu@bnl.gov

ABSTRACT

Interpretation of cryo-electron microscopy (cryo-EM) maps requires building and fitting 3-D atomic models of biological molecules. AlphaFold-predicted models generate initial 3-D coordinates; however, model inaccuracy and conformational heterogeneity often necessitate labor-intensive manual model building and fitting into cryo-EM maps. In this work, we designed a protein model-building workflow, which combines a deep-learning cryo-EM map enhancement tool, ResEM(**Res**olution **E**nhance**M**ent) and AlphaFold. A benchmark test using 37 cryo-EM maps shows that ResEM achieves state-of-the-art performance in optimizing real space model-map correlations between ResEM-enhanced maps and ground-truth models. Furthermore, in a subset of 17 datasets where the initial AlphaFold predictions are less accurate, the workflow significantly improves their model accuracy. Our work demonstrates that the integration of modern deep learning image enhancement and AlphaFold may lead to automated model building and fitting for the atomistic interpretation of cryo-EM maps.

Introduction

The resolution revolution in cryo-electron microscopy (cryo-EM) has greatly accelerated the experimental determination of protein structures[1]. The reconstructed cryo-EM maps have reached near-atomic resolutions. However, to make use of cryo-EM maps to understand protein function, atomistic protein structures, including chain tracing, amino-acid residue registration, side-chain conformation, and 3-D coordinates, have to be modeled and refined.

Building atomic models into cryo-EM maps is non-trivial. When the resolution of cryo-EM maps is at 3.5 Å or higher, PHENIX[2] or BUCCANEER[3] is used to build a partial model, and then time-consuming, iterative model adjustment in programs such as *Coot*[4] and PHENIX.Refine[5] is performed to improve the model quality and its fitness to the cryo-EM map. In model building, the first step is to identify secondary structure features such as α helices and β strands in a cryo-EM map. To facilitate the second structure recognition, machine learning approaches are employed to automatically detect secondary structures as implemented in programs such as SSELearner[6], γ-TEMPy[7], EMNUSS[8], Emap2sec[9]. Recently, deep learning (DL) methods based on 3D convolutional neural networks, such as UNet[10], were also employed to predict protein backbone structure[11,12] or model protein complex structures[13–15].

A protein's sequence determines its structure and function. Recent advances in deep learning algorithms have made a breakthrough in protein structure prediction with examples of programs, e.g., AlphaFold[16], RoseTTAFold[17], and ESMFold[18]. These programs predict high-quality atomic models that can be used to support experimental data for understanding biological complexity. Among these programs, AlphaFold produces the most accurate atomic models[18]. AlphaFold takes the sequence of a protein to produce a 3-D atomic model that can serve as a starting point for the interpretation of cryo-EM maps through model building and refinement. However, the AlphaFold model represents a still structure with knowledge learned from the experimental structures deposited to the Protein Data Bank[19](PDB). The AlphaFold model sometimes has poor quality and does not accommodate conformational changes. It is thus highly desirable to have a model-building workflow that leverages AlphaFold to facilitate atomistic interpretations of cryo-EM maps.

AlphaFold can use structure templates to improve its prediction accuracy. Recently, Terwilliger *et. al*[20] proposed an intuitive solution, which first fits and refines the AlphaFold model against its cryo-EM map, then uses the refined model as a template to guide AlphaFold prediction. They demonstrate that refined templates may help to improve the predicted model quality. However, most of the proteins tested in Ref.[20] are relatively small, averaging 311 residues. For several examples, their strategy resulted in model overfitting.

Cryo-EM maps are averaged reconstructions from thousands to millions of particles. Due to radiation damage, heterogeneity, and intrinsic molecular movement, the high-resolution contrast in cryo-EM densities was lost, causing blurring. Therefore, a

contrast-restoration process using a global B-factor sharpening has been routinely used to improve map interpretability and reduce overfitting[2,21,22]. However, a global B factor can not accommodate local radiation damage, protein heterogeneity, and flexibility in cryo-EM maps. Consequently, local B-factor sharpening methods were developed to post-process cryo-EM maps as implemented in programs Locscale[23], localDeblur[24], and Locspiral[25]. By learning structural features from atomic coordinates, deep learning methods were also deployed to perform local B-factor sharpening for cryo-EM [26–29]. Among these deep-learning methods, DeepEMhancer[29], which is based on a 3D UNet structure to improve the cryo-EM density map features, has shown good performance, hence is becoming one of the most popular tools to perform local B-factor sharpening to enhance cryo-EM maps.

In this study, we designed a deep learning model, ResEM(**Res**olution **E**nhance**M**ent), which can effectively enhance the quality of the cryo-EM maps. We developed a workflow to integrate the ResEM and AlphaFold to facilitate the model-building and refinement processes. We show that our ResEM outperforms DeepEMhancer by generating less noisy density maps that have higher correlations with the associated protein models. In addition, our workflow allows the refinement of larger proteins (chain length of 1017 residues on average) and significantly improves the interpretability of cryo-EM maps where their initial AlphaFold predictions are less accurate.

Results

Map to model workflow

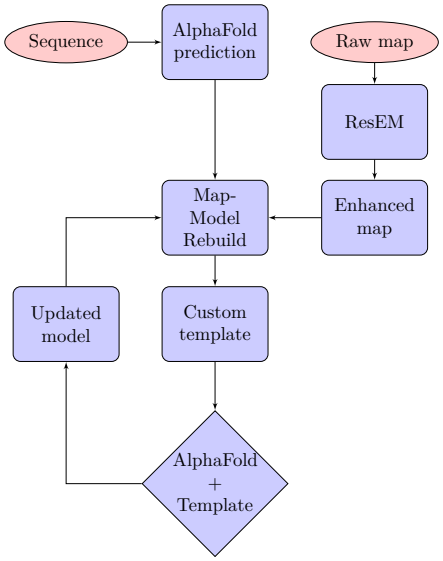

Figure 1. Schematics of the map to model workflow. After getting an initial model from AlphaFold and the enhanced cryo-EM map from our deep learning model ResEM, we use *phenix.dock_and_build* to rebuild the model with the ResEM-enhanced density map. The resulting model serves as a template for the next round of AlphaFold prediction. Then we again invoke *phenix.dock_and_build* to refine the new AlphaFold model against the ResEM-enhanced map, completing a full cycle of the workflow.

Our protein model-building workflow is illustrated in Fig. 1. The input data are the target sequence and its associated raw cryo-EM map. The output is the atomic coordinates that have been modeled and fitted into the ResEM-enhanced cryo-EM map. Our model-building strategy follows that of Ref.[20], but we used the ResEM-enhanced map for rebuilding AlphaFold models. The workflow can be iterated multiple times, although the gain will likely be marginal after the first cycle[20].

Deep Learning Model

Fig. 2 illustrates the structure of ResEM, which takes a raw cryo-EM map as input and outputs an enhanced map of the same dimension. The model has a U-Net[10] structure, consisting of encoding (down-sampling), bottleneck and decoding (up-sampling) blocks. The convolutional U-Net and its variants have shown good performances in image denoising tasks[29–31]. The residual connections between the encoding and decoding blocks alleviate the vanishing gradients problems and allow the model to scale up to learn more features[32]. The LeakyReLU (or ReLU) activation function after the final convolutional layer is to increase the output contrast further.

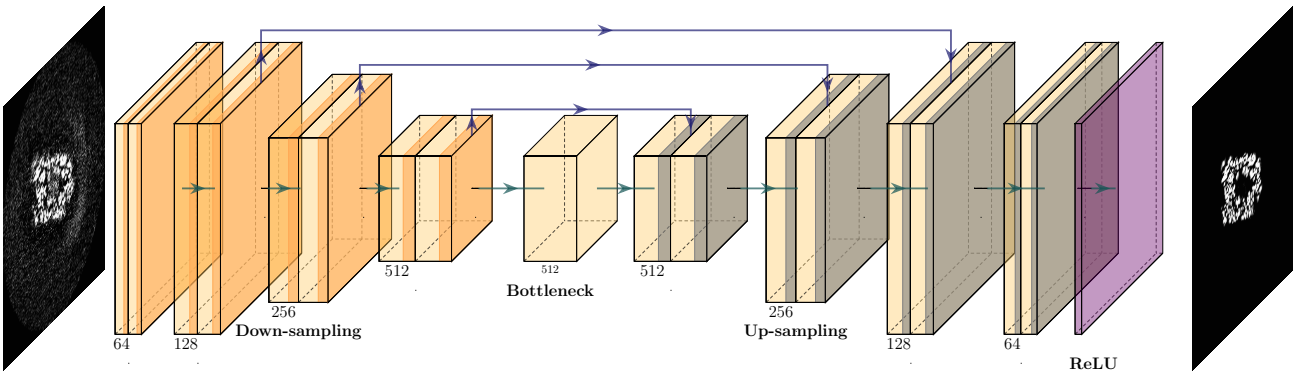

Figure 2. ResEM model structure. The number below each layer denotes the number of $3 \times 3 \times 3$ convolutional filters. The dark blue arrows stand for skip connections. We use strided convolution and transposed convolution for the down- and up-sampling[30], respectively. All the down-sampling, bottleneck, and up-sampling blocks are repeated twice.

Cryo-EM map enhancement

After training the ResEM model, we conducted a benchmark test on 37 cryo-EM maps, encompassing 20 maps used by DeepEMhancer. The reported resolution range is 2.8-6.7 Å, and the median resolution is 3.45 Å. Table 1 summarizes the average performance in real space. In addition to two common metrics for image quality assessment, i.e., the structural similarity index measure (SSIM)[33] and peak signal-to-noise ratio (PSNR), which are calculated using input maps against target maps, we used 4 different kinds of model-map correlation coefficients defined in[34]. Compared with the raw and DeepEMhancer maps, our method achieves higher SSIM, PSNR, and real-space correlation coefficients, indicating that ResEM excels DeepEMhancer in cryo-EM map enhancement while preserving structural information. Besides the metrics in real space, another important evaluation criterion is the resolution as estimated by using the Fourier Shell Correlation (FSC) curve. Table 2 shows median resolutions following different definitions (mask & cutoff) of the same 37 cryo-EM maps. Again ResEM-enhanced maps show higher resolutions in all scenarios. Note that for ResEM, the improvement of resolutions determined at a lower cutoff (0.143) is more significant than that at a higher cutoff (0.5). Unlike raw maps, the resolution differences between masked and unmasked cases for ResEM-enhanced maps are negligible. These observations indicate the exceptional capability of our method to restore high-frequency structure details of cryo-EM maps.

Metric	SSIM	PSNR	CC_{box}	CC_{mask}	CC_{volume}	CC_{peak}
Raw	1.06	13.04	0.737	0.724	0.727	0.616
DeepEMhancer	98.29	37.22	0.704	0.681	0.689	0.656
ResEM	**98.97**	**43.27**	**0.810**	**0.739**	**0.745**	**0.704**

Table 1. Benchmark: Real space results. CC_{box} is the whole map correlation, CC_{mask} and CC_{volume} are correlations between model-calculated and experimental density around atomic centers with slightly different definitions. CC_{peak} denotes the correlation coefficients between the highest values in the models and target maps[34].

Resolution	Masked (0.143)	Unmasked (0.143)	Masked (0.5)	Unmasked (0.5)
Raw	3.29	3.68	3.72	4.51
DeepEMhancer	2.41	2.42	3.96	3.98
ResEM	**2.36**	**2.38**	**3.68**	**3.68**

Table 2. Benchmark: Median resolution estimated by FSC curves under different conditions. The masks are generated using the deposited PDB models.

Map-Model building and Refinement

We have demonstrated that ResEM can uncover high-resolution features of cryo-EM maps. In this session, we will test whether the improved map quality can help protein model building and refinement. We will primarily use C_α score (see Methods) to measure the similarity between the predicted model and the deposited PDB model.

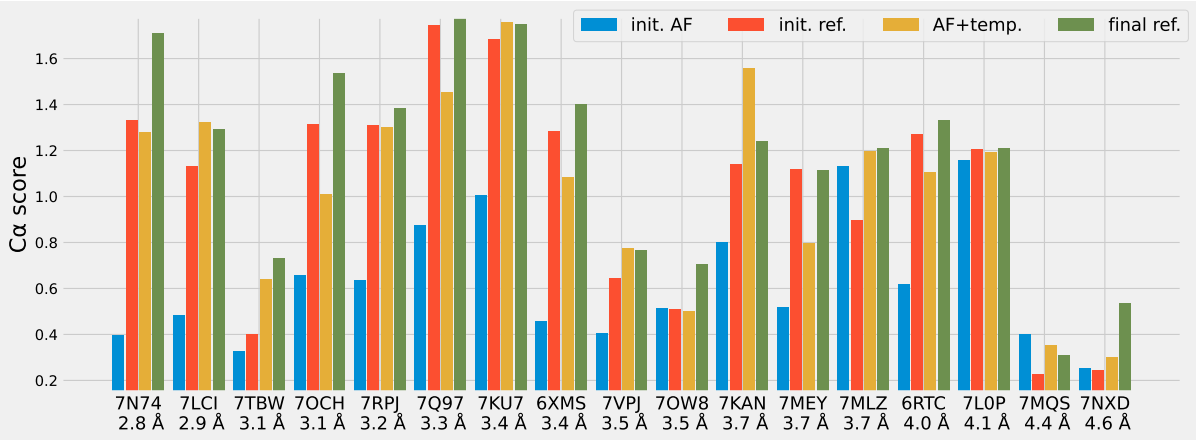

Figure 3. Progress of model accuracy per iteration.

Fig. 3 shows the overall progression of C_α score for 17 samples after one complete iteration of our workflow. We choose samples with 1) low initial C_α scores, i.e., samples where the AlphaFold predictions are less accurate, and 2) the reported resolutions spread evenly from 2.8 Å to 4.6 Å. Moreover, we selected the longest chain to model whenever possible for each PDB entry of a multi-protein complex. The average number of protein residues is 1017 (Fig. 3). In most cases, the final refined models achieve significant improvement over their initial AlphaFold models. The biggest improvement typically happens in the model-map rebuilding step, with the PDB entry 7MQS as an exception, to which we attribute its failure of improvement to a combination of low resolution (4.4 Å) and a relatively poor initial AlphaFold model (C_α score 0.4).

To further test the utility of the workflow, we down-selected 10 samples with the lowest initial C_α scores, which represent "hard" cases for model building, and conducted benchmark tests by replacing ResEM-enhanced maps with raw maps in the workflow. The results are shown in Fig. 4. Note that when using the raw maps, the initial map-model rebuilding failed for PDB entries 6XMS and 7NXD. Hence we showed the C_α score from the previous step (AlphaFold + template) instead. In all 10 "hard" cases, the workflow using ResEM-enhanced maps achieves higher C_α scores than that of using raw maps. On average, ResEM boosts the C_α score by 26% (excludes PDB entry 6XMS). So from Fig. 3 and 4 we conclude that ResEM-enhanced cryo-EM maps result in improved model quality.

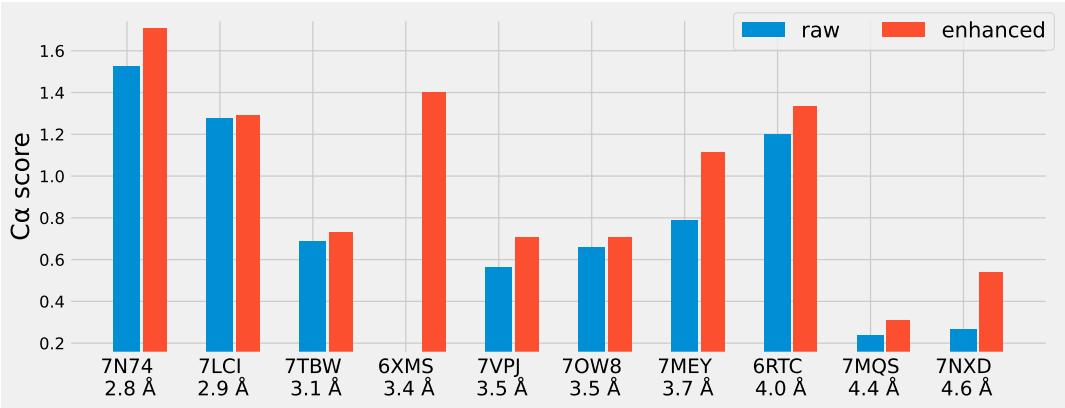

Figure 4. Benchmark with raw maps. C_α scores represent the final refinement outcome after a full cycle.

To visualize the performance of our workflow, finally, we present three samples, each with a different resolution (Fig. 5). We focus on protein regions where the initial AlphaFold models (white) have substantial deviations from the deposited PDB models (magenta). In the first two cases (PDB entries 7N74 and 7VPJ), the model building and refinement managed to correct

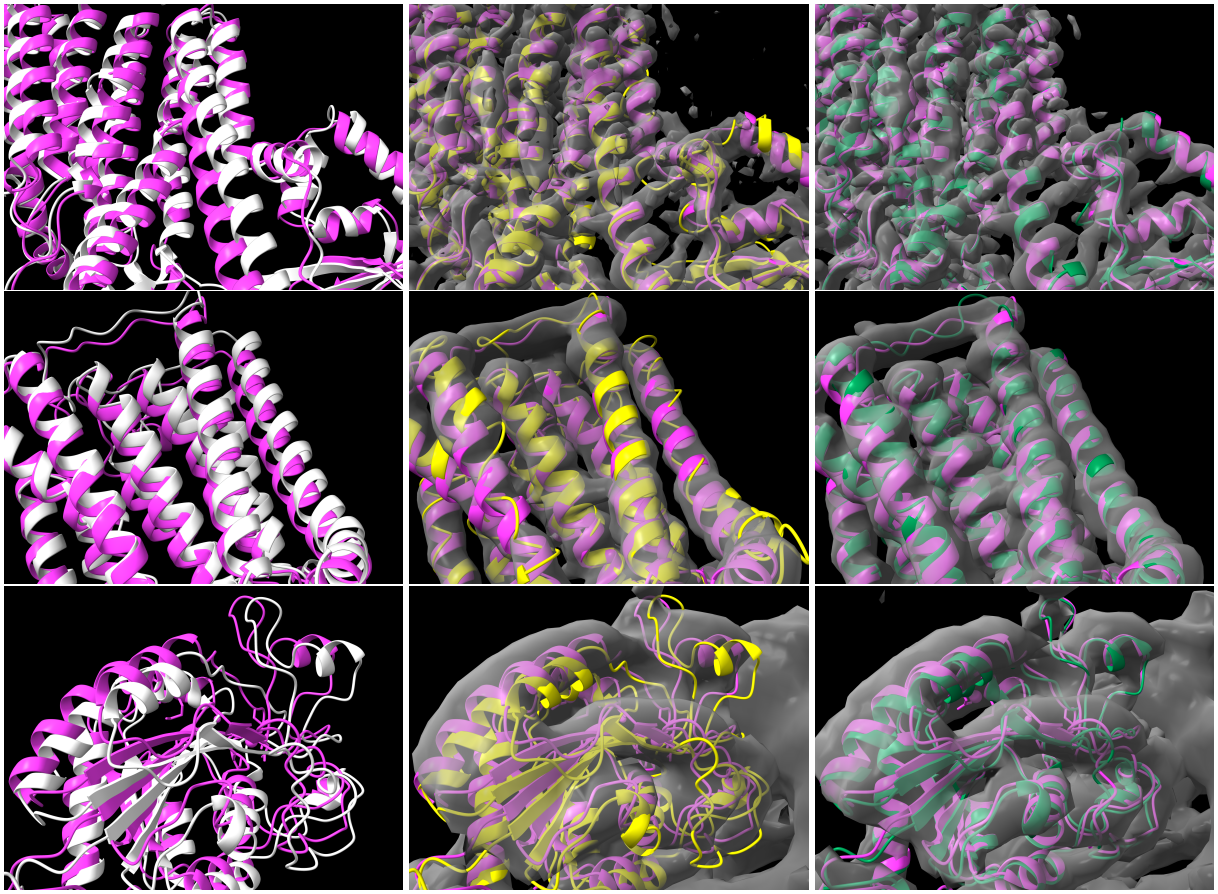

Figure 5. Comparison of model building and refinement with different cryo-EM maps. The top, middle, and bottom rows show partial views of 7N74 (2.8 Å), 7VPJ (3.5 Å), and 7NXD (4.6 Å), respectively. The original AlphaFold models (white), models refined using original maps (yellow), and models refined using ResEM-enhanced maps (green) are superimposed with the deposited models (magenta). The original (middle column) and ResEM-enhanced maps (right column) are shown in the background (transparent grey).

the deviated AlphaFold models. Subsequent refinement using the ResEM-enhanced maps shows better results than those using the raw maps. For 7NXD, where the AlphaFold model is extremely poor (C_α score 0.25) and the resolution is low (4.6 Å), refinement using the raw map failed to make any positive corrections. In contrast, using the ResEM-enhanced map allowed for the correction of the AlphaFold model.

Discussion

In this study, we have designed a map to model building and refinement workflow by combining AlphaFold, deep learning image enhancement, and PHENIX map-model building tool to facilitate cryo-EM map interpretations. We tested the refinement progress for samples with reported resolutions ranging from 2.8 to 4.6 Å (Fig. 3). In most cases, our workflow improves the protein structure model significantly after one iteration. To achieve the improvement, we developed a deep learning model, ResEM, which significantly enhances the atomic details of 3D cryo-EM maps, as evidenced by the improvement of real- and Fourier-space model-map correlations (Table 1,2). The ResEM-enhanced maps also benefit the PHENIX model building tool to refine AlphaFold predicted models (Fig. 4). Only for cryo-EM maps at low resolutions ($>$ 4.4 Å) with very poor AlphaFold models (C_α score lower than 0.4), our method failed to make a significant improvement to the structure model, which sets a current limit of applicability of our workflow.

It is worth mentioning that the current workflow may be further optimized. After a full cycle, the resulting model can serve as the starting point for additional refinement by using *phenix.real_space_refine* and *Coot* to adjust the model-map fitness while maintaining good stereochemistry. Fig. 6 shows the real-space refinement results with *phenix.real_space_refine* for the same 10 PDB entries of Fig. 4. We can see that using ResEM-enhanced cryo-EM maps also help with refinement and model building.

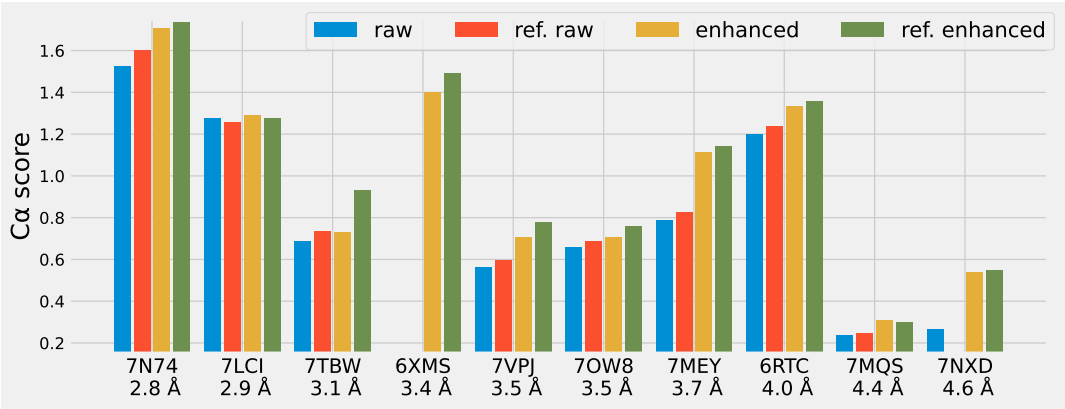

Figure 6. Model accuracy progress with *phenix.real_space_refine*. Here "raw" and "enhanced" refer to the models obtained after one full cycle of our workflow outlined in the Results session. "ref." means refinement. The real-space refinement fails for 7NXD when using the raw map.

Through our experiments, we have found that *phenix.dock_and_build* is sensitive to the map quality and the resolution limit at which it conducts the model building and refinement. Whenever using *phenix.dock_and_build* with ResEM-enhanced maps, we suggest trying to perform model building and refinement at the reported FSC resolution and 2 Å, then select the model with a better model-map correction to proceed. When using the raw cryo-EM maps for model building and refinement, we set the resolution limit as the reported resolution.

Before our work, DeepEMhancer represents the best deep learning tool for post-processing cryo-EM density maps to restore high-resolution details. By incorporating a larger amount of training data (478 for ResEM and 127 for DeeEMhancer), a careful design of the network structure and the training scheme (see Methods), plus the use of higher-resolution simulated maps as target training data, we were able to enhance cryo-EM maps to a better quality compared to DeepEMhancer (Tables 1,2), while maintaining a similar model size (49.6 M vs 51.1 M model parameters). It is also worth mentioning that ResEM is much faster in terms of post-processing. For reference, it took about 24 minutes for the ResEM to process 37 maps in the benchmark test set, while DeepEMhancer used more than 8 hours to process the same set of maps using identical hardware.

The open-sourced OpenFold offers a wide range of possibilities to incorporate additional data modalities and new model components. One promising direction is to integrate our deep learning model with AlphaFold in a more straightforward manner and retrain (or fine-tune) the conglomerated model to achieve better structure prediction and extend to downstream tasks, e.g., function predictions.

Methods

Cryo-EM maps

We collected the training and test datasets from EMDB[35] by restricting the reported map resolution range from 2.5 to 5 Å. All data used were deposited after January 2020. Only entries with valid half-maps were kept to get the "unpolished" raw maps. Then, we performed pairwise sequence alignment to exclude duplicate structures, which eliminates entries with sequence identity scores over 40%. All the EMDB entry IDs used in this study can be found in our GitHub repository[36]. The entire training set consists of 478 cryo-EM maps, from which 20% were randomly selected as the validation set to monitor the training progress. We also reserve 37 cryo-EM maps as the test set, which encompasses 20 entries used by DeepEMhancer. All examples described in the Results section (see above) were chosen from the test set only. After getting the original map by averaging two half-maps, we standardized the average map to a voxel volume of 1 $Å^3$ using ChimeraX[37]. The target map used for training was calculated from the deposited PDB structure by ChimeraX (*molmap* command) at 2 Å resolution. All input and target maps are normalized to [0, 1] using min-max normalization.

Training and validation of deep learning model

For training, the input maps are augmented by adding random Gaussian noise, random anisotropy (down-sampling in a random axis), and random blur (random-sized Gaussian filter) using TorchIO[38] to increase the generalizability. Because each cryo-EM map has different dimensions, we first split the map into chunks of size 64^3, then re-assemble it back to its original dimension for both training and testing. To avoid boundary artifacts, we adopt the approach of Ref.[14] by using only the central part (50^3) of the output in the re-assembling process.

The loss function is a combination of L-1 loss and a correlation function, i.e., loss $\propto L_1 + w(1 - \mathrm{corr}(y_{\mathrm{pred}}, y_{\mathrm{target}}))$, where $\mathrm{corr}(y_{\mathrm{pred}},\ y_{\mathrm{target}})$ is a local real space correlation function weighted by a spherically symmetric kernel[39], and the combination weight w is set to 0.5 such that two loss terms are of the comparable scale. We use the Adam optimizer[40] with a batch size of 16 (16×64^3 cubes), a weight decay of 1×10^{-5}, and an initial learning rate of 5×10^{-5}, which decreases to 2×10^{-5} after 10 epochs. The total training stopped after 30 epochs, and we selected the model with the highest real space correlation coefficient. In addition to the general training schemes (learning rate, batch size, and number of epochs), we found the most important hyperparameter to be the number of repeated convolutional blocks for each stage (up-sampling, bottleneck and down-sampling) in the model (Fig. 2), which determines the overall model scale. We experimented with 1, 2, 3 and found 2 to be commensurate with the available training data. ResEM was implemented using PyTorch v1.12[41]. All the training and testing were performed on a machine with a single Nvidia V100 GPU.

AlphaFold prediction

To maximize flexibility and expandability, our workflow uses OpenFold[42], an open-source and faithful implementation of AlphaFold, to perform protein structure predictions. To get the best prediction result, we conducted two separate runs with and without multiple sequence alignment (MSA) and selected the best prediction to proceed.

Model building and refinement

C_α score is the quotient of the percentage of "close" C_α atoms and the RMSD (root mean square deviation) of those "close" C_α atoms between the query and target model. When calculating the C_α score with *phenix.chain_comparison*, we first use *phenix.superpose_and_morph* to perform secondary structure matching (SSM) to superimpose the output model with the deposited PDB model. In this study, two C_α atoms are considered close if their relative displacement is less than 2 Å after superposition. For example, if 80% of C_α atoms of a predicted model locate within 2 Å to their equivalent atoms in the deposited model, and the RMSD of those 80% C_α atoms is 1.6 Å, then the C_α score is 0.8/1.6 = 0.5.

References

1. Kühlbrandt, W. The resolution revolution. *Science* **343**, 1443–1444 (2014).
2. Liebschner, D. *et al.* Macromolecular structure determination using x-rays, neutrons and electrons: recent developments in phenix. *Acta Crystallogr. Sect. D: Struct. Biol.* **75**, 861–877 (2019).
3. Cowtan, K. The buccaneer software for automated model building. 1. tracing protein chains. *Acta crystallographica section D: biological crystallography* **62**, 1002–1011 (2006).
4. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of coot. *Acta Crystallogr. Sect. D: Biol. Crystallogr.* **66**, 486–501 (2010).
5. Afonine, P. V. *et al.* Towards automated crystallographic structure refinement with phenix. refine. *Acta Crystallogr. Sect. D: Biol. Crystallogr.* **68**, 352–367 (2012).
6. Si, D., Ji, S., Nasr, K. A. & He, J. A machine learning approach for the identification of protein secondary structure elements from electron cryo-microscopy density maps. *Biopolymers* **97**, 698–708 (2012).
7. Pandurangan, A. P., Vasishtan, D., Alber, F. & Topf, M. γ-tempy: simultaneous fitting of components in 3d-em maps of their assembly using a genetic algorithm. *Structure* **23**, 2365–2376 (2015).
8. He, J. & Huang, S.-Y. Emnuss: a deep learning framework for secondary structure annotation in cryo-em maps. *Briefings bioinformatics* **22**, bbab156 (2021).
9. Maddhuri Venkata Subramaniya, S. R., Terashi, G. & Kihara, D. Protein secondary structure detection in intermediate-resolution cryo-em maps using deep learning. *Nat. methods* **16**, 911–917 (2019).
10. Ronneberger, O., Fischer, P. & Brox, T. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, 234–241 (Springer, 2015).
11. Si, D. *et al.* Deep learning to predict protein backbone structure from high-resolution cryo-em density maps. *Sci. reports* **10**, 1–22 (2020).
12. Chen, S., Zhang, S., Li, X., Liu, Y. & Yang, Y. Segem: a fast and accurate automated protein backbone structure modeling method for cryo-em. In *2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 24–31 (IEEE, 2021).
13. He, J. & Huang, S.-Y. Full-length de novo protein structure determination from cryo-em maps using deep learning. *Bioinformatics* **37**, 3480–3490 (2021).

14. Pfab, J., Phan, N. M. & Si, D. Deeptracer for fast de novo cryo-em protein structure modeling and special studies on cov-related complexes. *Proc. Natl. Acad. Sci.* **118**, e2017525118 (2021).

15. Zhang, X., Zhang, B., Freddolino, P. L. & Zhang, Y. Cr-i-tasser: assemble protein structures from cryo-em density maps using deep convolutional neural networks. *Nat. methods* **19**, 195–204 (2022).

16. Jumper, J. *et al.* Highly accurate protein structure prediction with alphafold. *Nature* **596**, 583–589 (2021).

17. Baek, M. *et al.* Accurate prediction of protein structures and interactions using a three-track neural network. *Science* **373**, 871–876 (2021).

18. Lin, Z. *et al.* Language models of protein sequences at the scale of evolution enable accurate structure prediction. *bioRxiv* (2022).

19. Berman, H. M. *et al.* The protein data bank. *Nucleic acids research* **28**, 235–242 (2000).

20. Terwilliger, T. C. *et al.* Improved alphafold modeling with implicit experimental information. *Nat. methods* **19**, 1376–1382 (2022).

21. Rosenthal, P. B. & Henderson, R. Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *J. molecular biology* **333**, 721–745 (2003).

22. Zivanov, J. *et al.* New tools for automated high-resolution cryo-em structure determination in relion-3. *elife* **7**, e42166 (2018).

23. Jakobi, A. J., Wilmanns, M. & Sachse, C. Model-based local density sharpening of cryo-em maps. *Elife* **6**, e27131 (2017).

24. Ramírez-Aportela, E. *et al.* Automatic local resolution-based sharpening of cryo-em maps. *Bioinformatics* **36**, 765–772 (2020).

25. Kaur, S. *et al.* Local computational methods to improve the interpretability and analysis of cryo-em maps. *Nat. communications* **12**, 1240 (2021).

26. Luo, Z., Campos-Acevedo, A. A., Lv, L., Wang, Q. & Ma, J. Sparseness and smoothness regularized imaging for improving the resolution of cryo-em single-particle reconstruction. *Proc. Natl. Acad. Sci.* **118**, e2013756118 (2021).

27. Subramaniya, S. R. M. V., Terashi, G. & Kihara, D. Super resolution cryo-em maps with 3d deep generative networks. *Biophys. J.* **120**, 283a (2021).

28. Zehni, M., Do, M. N. & Zhao, Z. Deepsharpen: Deep-learning based sharpening of 3d reconstruction map from cryo-electron microscopy. In *2020 IEEE 17th International Symposium on Biomedical Imaging Workshops (ISBI Workshops)*, 1–4 (IEEE, 2020).

29. Sanchez-Garcia, R. *et al.* Deepemhancer: a deep learning solution for cryo-em volume post-processing. *Commun. biology* **4**, 1–8 (2021).

30. Zhang, K. *et al.* Plug-and-play image restoration with deep denoiser prior. *IEEE Transactions on Pattern Analysis Mach. Intell.* (2021).

31. Bepler, T., Kelley, K., Noble, A. J. & Berger, B. Topaz-denoise: general deep denoising models for cryoem and cryoet. *Nat. communications* **11**, 5208 (2020).

32. He, K., Zhang, X., Ren, S. & Sun, J. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 770–778 (2016).

33. Wang, Z., Bovik, A. C., Sheikh, H. R. & Simoncelli, E. P. Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing* **13**, 600–612 (2004).

34. Afonine, P. V. *et al.* New tools for the analysis and validation of cryo-em maps and atomic models. *Acta Crystallogr. Sect. D: Struct. Biol.* **74**, 814–840 (2018).

35. Lawson, C. L. *et al.* Emdatabank unified data resource for 3dem. *Nucleic acids research* **44**, D396–D403 (2016).

36. https://github.com/empyriumz/CryoEM.

37. Goddard, T. D. *et al.* Ucsf chimerax: Meeting modern challenges in visualization and analysis. *Protein Sci.* **27**, 14–25 (2018).

38. Pérez-García, F., Sparks, R. & Ourselin, S. Torchio: a python library for efficient loading, preprocessing, augmentation and patch-based sampling of medical images in deep learning. *Comput. Methods Programs Biomed.* 106236 (2021).

39. Warshamanage, R., Yamashita, K. & Murshudov, G. N. Emda: a python package for electron microscopy data analysis. *J. Struct. Biol.* **214**, 107826 (2022).

40. Kingma, D. P. & Ba, J. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980* (2014).

41. Paszke, A. *et al.* Pytorch: An imperative style, high-performance deep learning library. *Adv. neural information processing systems* **32** (2019).

42. Ahdritz, G. *et al.* Openfold: Retraining alphafold2 yields new insights into its learning mechanisms and capacity for generalization. *bioRxiv* (2022).

Acknowledgments

We would like to thank Sean McSweeney and John Shanklin for the helpful discussion. This work was supported in part by Brookhaven National Laboratory LDRD 22-008. Q. L. was supported by the U.S. Department of Energy (DOE), Office of Biological and Environmental Research as part of the Quantitative Plant Science Initiative at BNL.

Author contributions

X.D. and Q.L. designed the study and experiments. X.D. and L.W. performed the experiments. X.D., L.W., S.Y., and Q.L. analyzed the data. X.D. and Q.L. wrote the manuscript with help from other coauthors.

Competing interests

The authors declare no competing interests.

Correspondence and requests for materials

Correspondence and requests for materials should be addressed to X. D. or Q.L.

Data availability

All data used in this study are publicly available from https://www.rcsb.org (for PDB structures) and https://www.emdataresource.org for cryo-EM maps.

Code availability

ResEM is implemented as an open-source python package available at https://github.com/Structurebiology-BNL/ResEM.