




DeepEMhancer: a deep learning solution for cryo-EM volume post-processing

Ruben Sanchez-Garcia ^{1,4}, Josue Gomez-Blanco^{2,3}, Ana Cuervo¹, Jose Maria Carazo ¹, Carlos Oscar S. Sorzano ¹✉ & Javier Vargas^{2,3}✉

Cryo-EM maps are valuable sources of information for protein structure modeling. However, due to the loss of contrast at high frequencies, they generally need to be post-processed to improve their interpretability. Most popular approaches, based on global B-factor correction, suffer from limitations. For instance, they ignore the heterogeneity in the map local quality that reconstructions tend to exhibit. Aiming to overcome these problems, we present DeepEMhancer, a deep learning approach designed to perform automatic post-processing of cryo-EM maps. Trained on a dataset of pairs of experimental maps and maps sharpened using their respective atomic models, DeepEMhancer has learned how to post-process experimental maps performing masking-like and sharpening-like operations in a single step. DeepEMhancer was evaluated on a testing set of 20 different experimental maps, showing its ability to reduce noise levels and obtain more detailed versions of the experimental maps. Additionally, we illustrated the benefits of DeepEMhancer on the structure of the SARS-CoV-2 RNA polymerase.

¹Biocomputing Unit, Centro Nacional de Biotecnología-CSIC, Madrid, Spain. ²Department of Anatomy and Cell Biology, McGill University, Montréal, QC, Canada. ³Departamento de Óptica, Universidad Complutense de Madrid, Madrid, Spain. ⁴Present address: Department of Statistics, University of Oxford, Oxford, UK. ✉email: coss@cnb.csic.es; jvargas@fis.ucm.es

Almost one decade after the beginning of the so-called “resolution revolution”, cryogenic electron microscopy (cryo-EM) has become one of the most versatile tools in the field of structural biology. Beginning from thousands of single-particle projection images, cryo-EM workflows are capable of obtaining three-dimensional (3D) reconstructions of many macromolecules at “near-atomic” resolution levels. However, the ultimate goal of the cryo-EM single-particle analysis is not the obtention of 3D maps but the detailed atomic understanding through the derivation of atomic models.

During the atomic model building process, raw 3D maps are rarely employed, as they suffer from loss of contrast at high resolution¹ that makes difficult the detection and interpretability of residues and secondary structure. Fortunately, loss of contrast can be alleviated using different contrast restoration algorithms, which are usually known as sharpening methods. The first sharpening approach for cryo-EM maps was introduced by Rosenthal and Henderson¹ and their formulation, based on the global *B*-factor correction, is still at the basis of the most commonly employed sharpening methods, including RELION postprocessing^{2,3} or Phenix AutoSharpen⁴. The principle behind these algorithms consists in the correction of the raw maps by boosting the amplitude of their high-frequency Fourier components. The strength of the amplitude boost at each frequency depends on the frequency itself and on a single number, the *B*-factor, that measures the global loss of contrast. Thus, although the different global *B*-factor-based methods differ in the procedures employed to determine the *B*-factor that is applied, they modify the volume globally in a similar manner.

Despite being widely used, global *B*-factor-based approaches present an important limitation: they do not consider the differences in quality that different parts of the map may present and they produce density maps that do not correspond to the scattering properties of biological macromolecules⁵. Consequently, for the case of maps that exhibit heterogeneous local resolution, some regions could be undersharpened whereas others could be oversharpened. Recently, local sharpening algorithms that alleviate this shortcoming, have been proposed. Thus, the LocScale⁶ algorithm uses the information contained in an atomic model to locally scale up a map. Such transformation is achieved by means of a sliding window approach in which the amplitudes of the map region that lay inside the window are scaled up to agree with the atomic model provided. Following a totally different strategy, the LocalDeblur⁷ algorithm employs a Wiener filtering approach that performs local deblurring with a strength proportional to an estimation of the local resolution, that has to be pre-computed. Similarly, LocSpiral⁸ employs the spiral phase transformation to factorize the volume and then perform a local enhancement based on the normalization and thresholding of the amplitudes.

Despite their benefits, current local sharpening approaches present some drawbacks. Thus, both LocSpiral and LocalDeblur depend on masks to distinguish the macromolecule from the noise and LocalDeblur requires also an estimation of the local resolution of the map. On the other hand, the main strength of LocScale, its ability to employ the structural information of atomic models, could also be regarded as its main weakness since the availability of atomic models limits its applicability.

With the aim of overcoming these shortcomings, in this work, we present Deep cryo-EM Map Enhancer (DeepEMhancer), a fully automatic deep learning-based approach that performs cryo-EM volume post-processing. Deep learning has revolutionized the field of artificial intelligence and its impact has been felt in many others including cryo-EM. Deep learning in cryo-EM was first applied to the problem of particle picking^{9–11} and since then, it has evolved to deal with other questions such as map

reconstruction^{12,13}, map segmentation^{14,15}, or local resolution determination^{16,17}. As in most of those methods, our approach relies on a convolutional neural network (CNN) that is trained on massive quantities of data. Particularly, our development, which follows a simple image super-resolution setup¹⁸, exploits the vast amount of structural information that is contained in the Electron Microscopy Data Bank (EMDB) database¹⁹ in order to mimic the local sharpening effect of the LocScale algorithm. However, DeepEMhancer does not require any atomic model to function and, contrary to previous methods, it also performs automatic (tight) masking of input maps. Our results show that DeepEMhancer, which works in a fully automatic manner, is able to largely improve the interpretability of the maps contained in our benchmark, performing better than classical global *B*-factor approaches.

Results

DeepEMhancer is based on an end-to-end U-net architecture²⁰ trained in a supervised manner. Particularly, we implemented a 3D U-net consisting of three downsampling blocks and three upsampling blocks that process cubic chunks of the input map (see Supplementary Table 1 for more details). Training was performed using pairs of input maps and target maps, consisting of experimental cryo-EM maps and tightly masked LocScale post-processed maps. Despite other possible alternatives (e.g., LocalDeblur, etc.) LocScale was chosen as the method to produce targets because it makes use of atomic model information, which tends to produce high-quality results. For a complete description of the data preparation, training, and evaluation processes see the “Methods” section.

DeepEMhancer performance on the testing set. In order to assess the quality of DeepEMhancer predictions, we first compared them against the target maps generated by LocScale. Thus, for DeepEMhancer maps, we measured a median correlation coefficient of 0.9 against LocScale maps in contrast to 0.6 for input maps (see Supplementary Fig. 1). Such an important increase in the correlation coefficient implies that DeepEMhancer has learned to accurately reproduce the effect of LocScale sharpening with one important advantage: no atomic models are required to employ DeepEMhancer.

Although reproducing the LocScale-sharpening effect was our main objective, the ultimate goal of map post-processing is to simplify the process of atomic model building. With the aim of studying if DeepEMhancer also contributes to that purpose, we next explored whether DeepEMhancer post-processed maps were more similar to the actual atomic models. To do so, we computed, for all the maps included in the testing set, the Fourier shell correlation coefficient (FSC) resolution between the input (half maps average) and post-processed maps against the reference maps obtained from the atomic models. As it is shown in Fig. 1, for all the examples included in the testing set, the application of DeepEMhancer increased the similarity of the input maps with respect to the references (blue and green bars). Particularly, the median improvement achieved by DeepEMhancer was $\sim 0.6 \text{ \AA}$ ($\sim 14\%$ in the frequency domain). Such an important improvement confirms that the maps computed by DeepEMhancer are more similar to the target maps.

DeepEMhancer post-processing operation performs a non-linear transformation of the experimental volume that produces a set of effects that could be broadly classified as masking/denoising and sharpening-like features enhancement. In order to disentangle the contribution of the different effects, we have also computed the FSC of the input and post-processed maps using a tight mask derived from the atomic model. As it can be observed

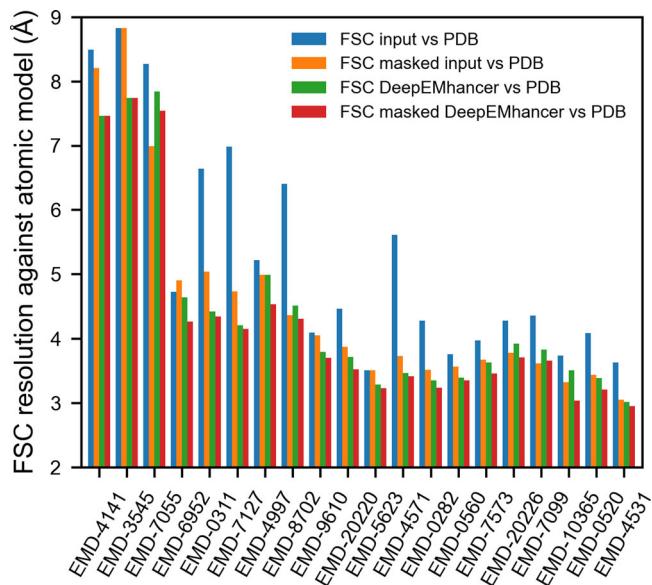


Fig. 1 DeepEMhancer produces maps that are more similar to the atomic models. Resolution (determined by Fourier shell correlation coefficient, FSC) between the reference maps obtained from the atomic model and (1) the input maps (blue), (2) the input maps tightly masked (orange), (3) the post-processed maps by DeepEMhancer (green) and (4) the post-processed maps by DeepEMhancer tightly masked (red). EMD entries are sorted by published global resolution.

in Fig. 1, the FSC resolution obtained for the post-processed maps tends to be better than the values computed for the input independently of the mask application (green and red bars vs orange bar), which implies that the masking effect is of high-quality, as the resolutions for the unmasked DeepEMhancer results tend to be better than the ones for the masked input maps.

Comparison with other methods. With the aim of comparing DeepEMhancer with the commonly employed global B -factor-based sharpening methods, we repeated the same experiments using the post-processed maps obtained with the Relion post-processing algorithm^{2,3}. Before it is important to notice that contrary to DeepEMhancer, Relion automatic masking is a simple process, and thus, in order to make the comparison more interesting, we used instead the masks derived from the atomic models.

Still, when we evaluated the FSC for the masked regions, only a few maps improved, while many others worsened, leading to a median improvement that was negligible (<0.05 Å) for both FSC and median DeepRes resolution (see Figs. 2 and 3).

Similarly, and, although it is true that the trend is not as strong as in the previous experiment, DeepEMhancer also tends to improve the resolution of the masked regions (Fig. 1, orange vs. red bars), which supposes an enhancement of the map features. Leaving aside some problematic examples such as EMD-7055²¹, that will be discussed in Supplementary Note 1 and Supplementary Fig. 2, most of the evaluated maps exhibit a non-negligible improvement in resolution, especially notable when compared to B -factor-based results (see next section), with a median value of ~ 0.3 Å.

Alternatively, with the aim of obtaining a complementary measurement of improvement, we computed the DeepRes local resolution for the input and post-processed maps. As can be appreciated in Fig. 2, all test cases treated with DeepEMhancer improved in terms of DeepRes local resolution, with dramatic improvements of more than 0.7 Å and a median improvement of

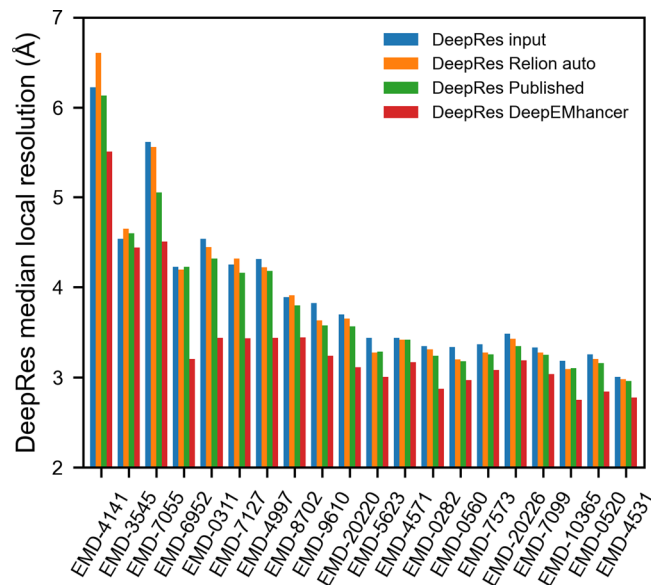


Fig. 2 DeepEMhancer produces better quality maps. DeepRes median local resolution estimation for (1) the input maps (blue), (2) the post-processed maps obtained with Relion postprocessing automatic B -factor (orange), (3) the post-processed maps deposited in EMDB (green) and (4) the post-processed maps obtained with DeepEMhancer (Red). EMD entries are sorted by published global resolution.

~ 0.4 Å. Again, those figures, consistent with the FSC-based measurements, point out that DeepEMhancer is improving the interpretability of the maps.

We acknowledge that the automatic determination of the B -factor can lead to less accurate results than if it were manually selected and it may be the reason behind the poor observed performance. Thus, we have also included in the comparison the post-processed maps deposited in EMDB in which the estimation of B -factor was carried out by the authors. In this case, the improvement in resolution, with median values of ~ 0.15 and ~ 0.1 Å for DeepRes and FSC, respectively, although closer to the values obtained using DeepEMhancer, are still considerably inferior (see Figs. 2 and 3). Such a difference in performance can be partially explained by the ability of local sharpening methods to deal better with low-quality regions of input maps as is shown in Supplementary Figure 3 and discussed in Supplementary Note 2.

In the light of these results, we can state that DeepEMhancer maps tend to be more similar to the atomic models than the ones obtained using global B -factor-based methods and thus, more useful for the process of model building. Finally, for the sake of completeness, we also computed FSC curves to compare our approach with other state-of-the-art sharpening approaches, showing that our fully automatic approach produces competing if no better results for many cases (see Supplementary Note 3 and Supplementary Figs. 4–9).

Visual inspection of testing maps. The purpose of this section is to further explore the results obtained with DeepEMhancer for some of the maps included in the testing set with the aim of illustrating how the improvements in global quality measurements translate to tangible improvements in the quality of the maps.

EMD-7099. The EMD-7099²² is a high-resolution volume (global resolution 3.1 Å) of a multidrug resistance ATP-driven pump.

EMD-7099 presents 17 transmembrane helices and, although the overall quality of the map is excellent, visualizing the transmembrane regions is challenging because of the signal that comes from the lipids. As a result, important parts of the protein are not traced. Due to the fact that DeepEMhancer was trained to ignore the signal coming from lipidic layers, this example illustrates the unique characteristics of DeepEMhancer when applied to membrane proteins. Thus, as can be observed in Fig. 4a–d, DeepEMhancer has been able to suppress the signal coming from the lipid layer in a much more simple and effective way than diminishing the threshold in the raw map or the *B*-factor-based sharpened maps. The noise suppression effect simplifies the process of model building, as the researchers do not have to deal with masks or larger thresholds that make the visualization of near-to-noise level features more difficult. Yet not only DeepEMhancer produces a noise reduction effect, but also it is able to enhance some parts of the map that under *B*-factor-based sharpening seem noisy and disconnected. Such improvement, although observed in several regions of the map, is more noticeable at the transmembrane region. Thus, the most important enhancement is depicted in Fig. 4e, f, in which an important part of the backbone of the protein has been de novo traced thanks to DeepEMhancer enhancement, that has restored the densities corresponding to residues A195 to I203 in chain A of PDB 6bhu. Although it is true that this region was present in the raw data map, its intensity range was so close to one of the lipidic layers that after conventional *B*-factor post-processing, the region was so damaged that modeling was not possible. On the contrary, not only DeepEMhancer was able to suppress most of the signal coming from the lipid layer but also it was able to restore the density of the region so that it looks smooth and continuous.

EMD-4997. The EMD-4997²³ is a medium-high resolution volume (4.0 Å) for a murine epithelial anion transporter. As in the previous example, the overall quality of the map is quite good, yet it presents lower quality regions. Figure 5a shows an overview of the published map, displayed at the recommended threshold, and the map obtained with DeepEMhancer. Although it is true that both the published map and the post-processed map look very similar, it is also true that there exist important differences. Firstly, the map processed with DeepEMhancer is cleaner than the published one. Serve as an example the removal of the artifacts that the published map presents near the elbow of the complex (see Fig. 5a, red box). More importantly, there can also be found many regions for which the DeepEMhancer post-processed volume resolves better the different residues of the regions. One such example can be found near the N-terminal end of the protein complex. Thus, as it is shown in Fig. 5b, the densities that correspond to the strands of the β -sheet are better separated than in the published volume. It is important to notice that this better separation is not a consequence of the employed thresholds, as it is proven by the fact that rising the threshold makes the densities corresponding to the backbone discontinuous before the densities for the two strands separate (see Fig. 5b). As a result, we can affirm that the quality of this region has been improved by the usage of DeepEMhancer.

Another similar example is displayed in Fig. 5c. In this case, two non-contiguous aromatic residues, Y361 and H121, seem connected in the published map. However, when DeepEMhancer is applied, the densities corresponding to the two residues look separated while the backbone remains continuous.

Use case EMD-30178 from SARS-CoV-2 RNA-dependent RNA polymerase. In order to further explore the benefits of the DeepEMhancer algorithm, we analyzed more deeply the post-

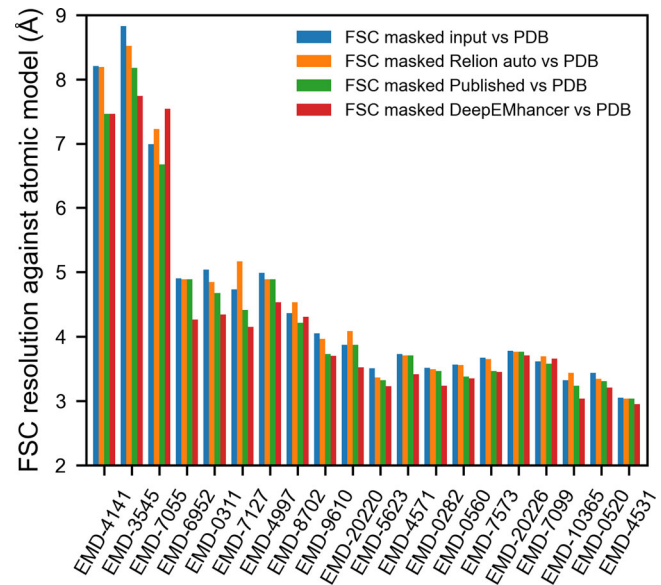


Fig. 3 DeepEMhancer produces better results than global *B*-factor-based methods. Resolution (determined by Fourier shell correlation coefficient, FSC) between the reference maps obtained from the atomic model and (1) the input maps (blue), (2) the post-processed maps obtained with Relion postprocessing automatic *B*-factor (orange), (3) the post-processed maps deposited in EMDb (green), and (4) the post-processed maps obtained with DeepEMhancer (red). EMDb entries are sorted by published global resolution.

processing of EMD-30178 map from Gao²⁴, corresponding to the SARS-CoV-2 RNA-dependent RNA polymerase. The published map presents detailed structure up to 2.9 Å resolution, however, as is often the case in cryo-EM, the resolution of the map is highly heterogeneous. We have chosen this map not only for the importance of this structure in current days but also because of the fact that the heterogeneous quality of the map density presents an ideal case for DeepEMhancer software. As it is shown in Fig. 6a, the application of the algorithm reduces the noise and improves the consistency and depiction of the map. To better illustrate these differences, we have chosen two different regions in chains A and D where the differences between the published and the DeepEMhancer map can be appreciated (Fig. 6b and c). While the density in the published map looks noisy or discontinuous depending on the displayed threshold (Fig. 6b and c, left and middle panel), the application of the DeepEMhancer software results in a well-defined continuous density where the side chains are nicely depicted (Fig. 6b and c, right panel). This improvement in the map density allowed us to close the loop between residues in the β -sheet V115 to I132 from chain D tracing three new residues that were not traced in the published structure (Fig. 6b). The improvement of the density is not only applicable to the edges of the map but it can be also appreciated in its core. Residues H362–L366 in chain A, traced on the published map were positioned more accurately on the density after map post-processing (Fig. 6c).

Discussion

The number of deposited high-resolution cryo-EM maps has soared since the beginning of the ‘resolution revolution’. As a result, there is an increasing number of atomic models that are being built using cryo-EM as the primary source of information. However, building atomic models directly from the raw maps is

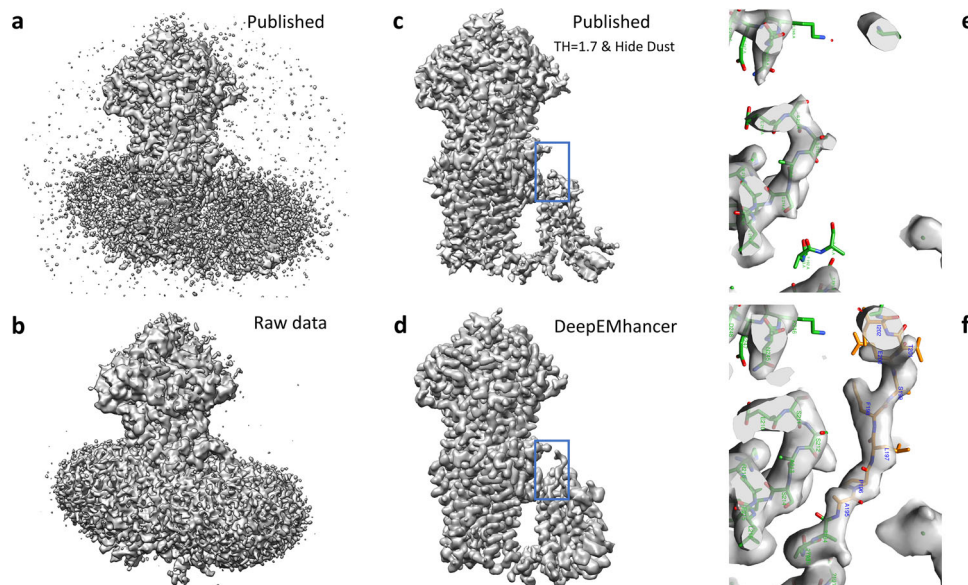


Fig. 4 DeepEMhancer results on testing map EMD-7099. **a** Lateral view of the published map (B -factor sharpened, shown at the threshold recommended by the authors). **b** Lateral view of the raw data map obtained from the half maps that was used as input for DeepEMhancer. **c** Lateral view of the published map after rising the threshold and removing the small connected components so that the signal coming from the lipids was suppressed. As a collateral consequence, some densities corresponding to the protein were also lost. **d** Lateral view of the map obtained with DeepEMhancer. **e** Zoom-in of the region marked with a blue box in **c**. **f** Zoom-in of the region marked with a blue box in **d**, in which DeepEMhancer post-processed map, contrary to the published map, shows the densities corresponding to a missing loop in PDB 6bhu chain A. As a result, the residues A195 to I203 have been de novo modeled (new residues depicted in yellow, published in green).

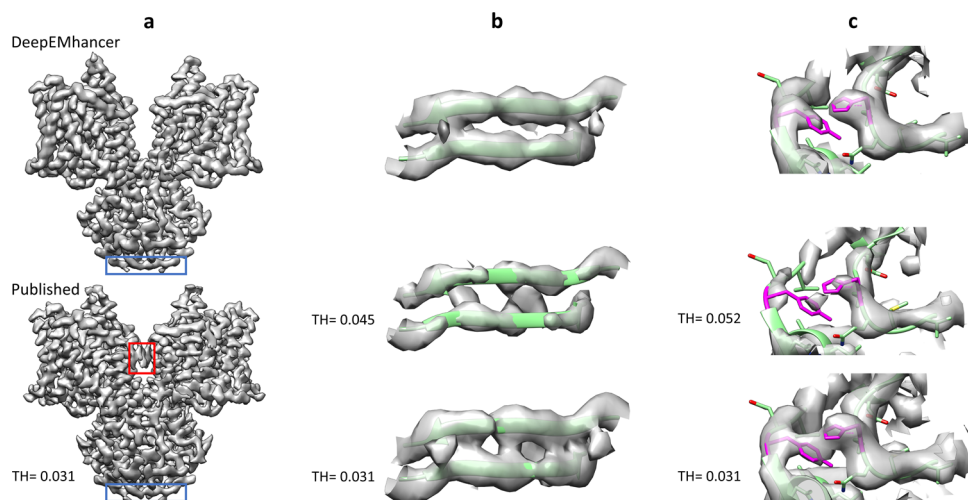


Fig. 5 DeepEMhancer results on testing map EMD-4997. **a** Overview of the published map (B -factor sharpened, shown at the threshold recommended by the authors), bottom, and the map obtained with DeepEMhancer, top. Red box highlights an artifact that has been automatically removed by DeepEMhancer. Blue box delimits the region shown in **b**. **b** Zoom-in of the region marked with a blue box that contains the β -sheet R7-A10, chains A and B. The published volume is shown at the recommended threshold and at the threshold at which the backbone begins to look discontinuous. As it can be appreciated, the DeepEMhancer solution resolves better than the published map of the two strands of the sheet. **c** Zoom-in of the region centered at chain B residues H121 and Y361 (colored in magenta). The published volume is shown at the recommended threshold and at the smaller threshold at which the density that connects the two residues disappears. As it can be appreciated, DeepEMhancer post-processed map resolves better than the published map of the two residues.

generally not possible. Instead, maps are post-processed in order to enhance the contrast of their high-resolution features.

In this work, we have presented DeepEMhancer, a map post-processing method based on deep learning. Trained on pairs of experimental cryo-EM maps and post-processed maps constructed with LocScale using atomic models, DeepEMhancer has learned how to perform a high-quality post-processing operation

that reproduces the effects of masking and local sharpening in an automatic fashion.

Although it is true that DeepEMhancer could have been trained on other targets, for instance, the simulated maps obtained directly from the atomic models, we discarded this alternative for two reasons. The first reason is that we wanted to reproduce the state-of-the-art local sharpening effect and not a

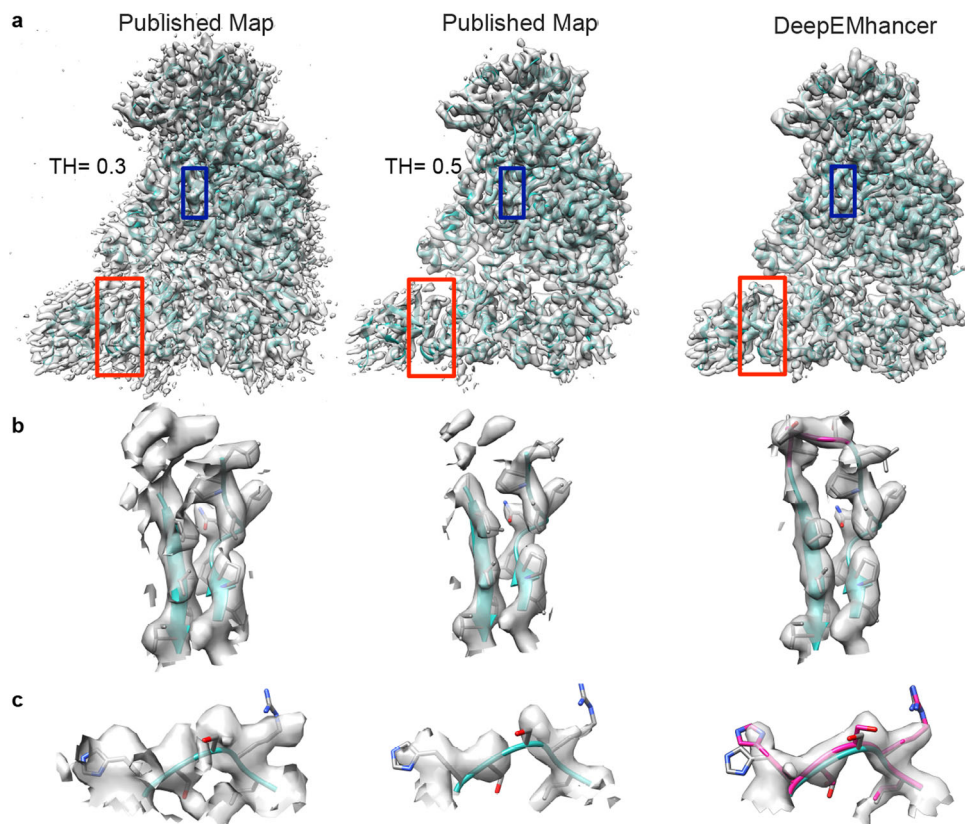


Fig. 6 Use case EMD-30178 from SARS-CoV-2 RNA-dependent RNA polymerase. a Overview of the published map displayed with two different thresholds 0.3 (recommended by the authors, left) and 0.5 (middle panel) and processed with DeepEMhancer (right). PDB 7btf is shown in ribbon, red squares designated the zoomed areas in **b** panel and blue squares the zoomed areas in **c**. **b** Zoom-in and extraction of the density from the 3D reconstruction of the published map at different thresholds and DeepEMhancer map corresponding to the red squares in **a**, chain D from residues V115–I132. Newly traced residues in the DeepEMhancer map are shown in pink. **c** Zoom-in and extraction of the density from the 3D reconstruction of the published map at different thresholds and DeepEMhancer map corresponding to the blue boxes.

new type of post-processing that could not be compatible with downstream atomic modeling tools. The other one is empirical: we obtained better results when targets were produced with LocScale than when the targets were directly obtained from the atomic models. As it is discussed in Supplementary Note 4 and illustrated in Supplementary Figs. 10–13, our neural network tends to suffer from underfitting when trained on maps derived from atomic models and thus, the results are blurrier than the ones obtained when using LocScale maps as a target. One possible explanation for such behavior could be the fact that, when using LocScale, the input and target maps, although different, still share some similar properties such as intensity ranges or local quality, which are not necessarily preserved when using simulated maps as targets. As a consequence, it is reasonable to believe that as the input and target maps become more similar, the training process should also become easier. For these reasons, we expect that super-resolution approaches trained on maps derived from atomic models will only be possible when more powerful models will be employed at the cost of more powerful computational resources and larger datasets.

The performance of our algorithm has been assessed using a testing set of 20 experimental maps that were not used for training nor during the trial-and-error process required for its implementation. In all cases, the similarity between the maps obtained from the atomic models and the experimental maps improved after the application of DeepEMhancer. Additionally, we evaluated in detail the performance of DeepEMhancer on two of those maps, showing that, not only DeepEMhancer facilitates

the visualization of cryo-EM maps, but also that DeepEMhancer can unveil some details that are not easily recognizable in the raw maps.

Nevertheless, it is important to highlight that DeepEMhancer is not the ultimate solution and that different examples will benefit from considering simultaneously different post-processing techniques. This is of especial importance for some of the cases in which DeepEMhancer, by dataset scarcity, presents limitations, for instance, when dealing with uncommon posttranslational modifications (see Supplementary Note 5 and Supplementary Figs. 14 and 15).

Another important caveat that all methods intended to enhance maps need to face is the problem of model validation. Although the results here presented have been validated using as ground truth the published models, in real-world scenarios such ground truth models are not available, and thus, the goodness of the results should be addressed by the users. To that end, we recommend trying and comparing different approaches since orthogonal methods should reveal inconsistencies. On the contrary, we discourage users from trying to estimate the resolution of post-processed maps, as there is no obvious way of doing it without ground-truth and even in those cases, masking effects could be challenging (see Supplementary Note 3).

Finally, with the aim of illustrating how beneficial DeepEMhancer could be in real-world scenarios, we have employed it on a map of the RNA polymerase of the SARS-CoV 2 virus, improving its quality of the map and the quality of the associated atomic model.

Methods

Raw data collection. DeepEMhancer has been trained and evaluated using as input a subset of cryo-EM maps obtained from the EMDB¹⁹ that meet the following requirements: (1) resolution better than 7 Å; (2) have one and only one atomic model associated; (3) correlation between the atomic model and the map better than 0.6; and (4) half maps available. As a result, an original list of 415 maps was compiled. However, this initial list is highly redundant and, in order to avoid biases in both the training and evaluation procedures, this list was further filtered to reduce its redundancy (see subsection “Redundancy control”). Finally, after a visual inspection aimed at removing problematic cases that survived the automatic filtering procedure, a total amount of 147 maps, with an average reported resolution of 3.8 Å, were selected.

Since the main objective of DeepEMhancer is to perform a sharpening-like post-processing transformation, it is important to ensure that the maps used in this study were not previously sharpened. Given the fact that most of the maps deposited in EMDB are sharpened and many are also masked, we decided to employ only the half-maps available in EMDB (condition number 4). Due to the lack of an appropriate searching tool in EMDB and a file name convention, we had to analyze all the map file names included in the database looking for the substring “half” to recover the half maps. Full maps were obtained averaging respective half maps.

As learning targets, we employed the output generated by LocScale using as input the aforementioned maps and their associated atomic models. Additionally, the output maps were tightly masked using as masks the maps simulated from the atomic models after a thresholding operation (see Supplementary Note 6 and Supplementary Fig. 16).

Data preparation. Due to the fact that the monomers (amino acids, nucleotides, etc.) that compose the macromolecules have fixed size but the deposited maps vary in voxel size, both the input and the target maps were resampled to 1 Å/voxel size with the aim of facilitating the learning process. After that, the intensity of each volume was normalized using the classical cryo-EM approach by which the map noise statistics are forced to adopt a fixed mean and standard deviation (0 and 0.1, respectively). Finally, due to GPU memory limitations, the maps were chunked into 64 × 64 × 64 cubes, the maximum size that our computing systems were able to efficiently manage. As a result, more than 70k volume cubes, including both signal cubes and noise-only cubes were used for training.

Redundancy control. In order to perform the train/test/validation split used to develop and evaluate our method, it is important to consider that the universe of proteins is highly redundant and that the EMDB entries are even more redundant. Serve as an example the case of the ribosome, which supposes ~10% of all EMDB entries. Thus, in order to avoid an over-optimistic performance estimation, we have ensured that the train, test, and validation sets are mutually exclusive in the sense that their intersections are empty under a certain equivalence criterion. Particularly, we consider that two EMDB entries are equivalent if they share one sequence that belongs to the same 30% sequence identity cluster. Similarly, with the aim of eliminating potential bias in the evaluation, we have guaranteed that only one member per cluster is included in testing and validation sets. On the contrary, we have relaxed our quite strict redundancy control policy in the training set allowing up to five cluster representatives in an attempt to increase the size of this set. This decision is founded on the fact that even maps of the same exact protein may present different statistics due to the intrinsic variability of cryo-EM reconstruction workflows and thus, limiting their presence in the training set may be difficult for the generalization of the neural network.

As a result, a list of 107, 21, and 20 maps were used for training, validation, and testing, respectively. The full list of the EMDB entries used can be found in Supplementary Note 7 and Supplementary Data 1.

Neural network architecture. We have employed a 3D U-net-like neural network²⁰ as a regression model for the estimation of post-processed maps. Our neural network consists of three downsampling blocks and three upsampling blocks with skip connections. Each block contains three convolutional layers followed by group normalization²⁵ and PRelu activation²⁶. The number of filters for each block is 3 × 32, 3 × 64, and 3 × 128, respectively. Downsampling is carried out using strided convolution and upsampling is performed via transposed convolution. See Supplementary Table 1 for additional details.

Neural network training. Our neural network was trained using stochastic gradient descent with a batch size of 8 cubes. Initial learning rate was set to 10⁻³ and decreased by a factor of 0.5 when the validation loss did not improve during 5 epochs. As a loss function, a mean absolute error was employed. Data augmentation, consisting of random 90° rotations, gaussian blurring, and patch corruption was applied to the training data.

Neural network inference. In order to perform volume post-processing, the input volume is pre-processed as described in the “Data preparation” subsection. Then, the resized and normalized volume is chunked into overlapping cubes of size 64 × 64 × 64 with strides of 16 voxels. Each cube is individually processed by the trained neural network, yielding post-processed cubes. After that, the post-processed cubes

are re-assembled into the final volume averaging the overlapping parts. Finally, the processed volume is resized to the size of the original volume, thus, showing the correct sampling rate value.

Evaluation. With the aim of guiding the cross-validation process, we computed the correlation coefficient between the maps produced by DeepEMhancer and the maps used as learning targets (masked LocScale post-processed maps). Once the final model was selected, the quality of DeepEMhancer predictions was assessed comparing the input and processed maps against the reference maps obtained from the atomic models. Specifically, we computed the FSC between them and we estimated the resolution using 0.5 as the threshold. Due to the fact that DeepEMhancer performs a non-conventional post-processing operation, including masking and enhancement operations, in order to disentangle the two effects, the FSC was also computed after masking the maps to compare with a tight mask derived from the atomic model.

As a complementary metric, we also applied DeepRes¹⁷ over the input and processed maps. DeepRes is a deep learning-based local resolution method that, contrary to others, is sensitive to the sharpening process and thus, it can provide an alternative estimation of the post-processing effect.

Finally, for comparison purposes, we repeated the FSC and DeepRes experiments using the Relion postprocessing program^{2,3}. As Relion automatic masking is very simple, in order to make the comparison more interesting, we decided to execute the postprocessing algorithm using the mask derived from the atomic models. Similarly, since the automatic determination of the *B*-factor can produce worse results than a manually selected one, in addition to the maps computed using an automatically determined *B*-factor by Relion, we also considered the sharpened map deposited in EMDB.

EMD-30178 map evaluation and atomic model modification. DeepEMhancer was applied to the half maps deposited in EMDB entry EMD-30178. The published and post-processed maps were visually inspected using Coot²⁷ and Chimera²⁸, and chosen regions on the 7btf PDB were newly built or modified using Coot.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

All training and testing examples used in this work can be found in the EMDB and PDB databases. Accessions codes are included in Supplementary Note 7 and Supplementary Data 1. Post-processed map examples and trained models are freely available at <http://campins.cnb.csic.es/deepEMhancer/examples>. Data used during figure preparation is available in Supplementary Data 2 and 3. All other data are available from the corresponding authors upon reasonable request.

Code availability

DeepEMhancer is freely available at <https://github.com/rsanchezgarc/deepEMhancer> and as an Xmipp protocol for Scipion v3 (<https://github.com/I2PC/scipion-em-xmipp>).

Received: 20 August 2020; Accepted: 17 June 2021;

Published online: 15 July 2021

References

1. Rosenthal, P. B. & Henderson, R. Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *J. Mol. Biol.* **333**, 721–745 (2003).
2. Kimanius, D., Forsberg, B. O., Scheres, S. H. & Lindahl, E. Accelerated cryo-EM structure determination with parallelisation using GPUs in RELION-2. *Elife* **5**, e18722 (2016).
3. Zivanov, J. et al. New tools for automated high-resolution cryo-EM structure determination in RELION-3. *Elife* **7**, e42166 (2018).
4. Terwilliger, T. C., Sobolev, O. V., Afonine, P. V. & Adams, P. D. Automated map sharpening by maximization of detail and connectivity. *Acta Crystallogr. Sect. D* **74**, 545–559 (2018).
5. Vilas, J. L. et al. Re-examining the spectra of macromolecules. Current practice of spectral quasi *B*-factor flattening. *J. Struct. Biol.* **209**, 107447 (2020).
6. Jakobi, A. J., Wilmanns, M. & Sachse, C. Model-based local density sharpening of cryo-EM maps. *Elife* **6**, e27131 (2017).
7. Ramirez-Aportela, E. et al. Automatic local resolution-based sharpening of cryo-EM maps. *Bioinformatics* **36**, 765–772 (2020).
8. Kaur, S. et al. Local computational methods to improve the interpretability and analysis of cryo-EM maps. *Nat. Commun.* **12**, 1240 (2021).
9. Wagner, T. et al. SPHIRE-crYOLO is a fast and accurate fully automated particle picker for cryo-EM. *Commun. Biol.* **2**, 218 (2019).

10. Wang, F. et al. DeepPicker: A deep learning approach for fully automated particle picking in cryo-EM. *J. Struct. Biol.* **195**, 325–336 (2016).
11. Zhu, Y., Ouyang, Q. & Mao, Y. A deep convolutional neural network approach to single-particle recognition in cryo-electron microscopy. *BMC Bioinforma.* **18**, 348 (2017).
12. Gupta, H., McCann, M. T., Donati, L. & Unser, M. CryoGAN: a new reconstruction paradigm for single-particle cryo-EM via deep adversarial learning. Preprint at *bioRxiv* <https://doi.org/10.1101/2020.03.20.001016> (2020).
13. Zhong, E. D., Bepler, T., Davis, J. H. & Berger, B. Reconstructing continuous distributions of 3D protein structure from cryo-EM images. Preprint at *arXiv* <https://arxiv.org/abs/1909.05215v3> (2019).
14. 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).
15. Si, D. et al. Deep learning to predict protein backbone structure from high-resolution cryo-EM density maps. *Sci. Rep.* **10**, 1–22 (2020).
16. Avramov, T. et al. Deep learning for validating and estimating resolution of cryo-electron microscopy density maps †. *Molecules* **24**, 1181 (2019).
17. Ramírez-Aportela, E., Mota, J., Conesa, P., Carazo, J. M. & Sorzano, C. O. S. DeepRes: a new deep-learning- and aspect-based local resolution method for electron-microscopy maps. *IUCrJ* **6**, 1054–1063 (2019).
18. Yang, W. et al. Deep learning for single image super-resolution: a brief review. *IEEE Trans. Multimed.* **21**, 3106–3121 (2019).
19. Lawson, C. L. et al. EMDatabank unified data resource for 3DEM. *Nucleic Acids Res.* **44**, D396–D403 (2015).
20. Ronneberger, O., Fischer, P. & Brox, T. U-net: convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention-MICCAI*, Vol. 9351, (eds Navab, N., Hornegger, J., Wells, W. M., Frangi, A. F.) 234–241 (2015).
21. Tentorey, J. L. et al. The structural basis of flagellin detection by NAIP5: a strategy to limit pathogen immune evasion. *Science (80-)* **358**, 888–893 (2017).
22. Johnson, Z. L. & Chen, J. ATP Binding enables substrate release from multidrug resistance protein 1. *Cell* **172**, 81–89e10 (2018).
23. Walter, J. D., Sawicka, M. & Dutzler, R. Cryo-EM structures and functional characterization of murine Slc26a9 reveal mechanism of uncoupled chloride transport. *Elife* **8**, e46986 (2019).
24. Gao, Y. et al. Structure of the RNA-dependent RNA polymerase from COVID-19 virus. *Science (80-)* **368**, 779–782 (2020).
25. Wu, Y. & He, K. Group normalization. *Int. J. Comput. Vis.* **128**, 742–755 (2020).
26. He, K., Zhang, X., Ren, S. & Sun, J. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In *Proc. 2015 International IEEE International Conference on Computer Vision*, 1026–1034 (2015).
27. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. Sect. D* **60**, 2126–2132 (2004).
28. Pettersen, E. F. et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).

Acknowledgements

The authors would like to acknowledge economical support from: The Spanish Ministry of Science and Innovation through Grants: Proyectos de I+D+i - RTI Tipo A PID2019-108850RA-I00, SEV 2017-0712, PID2019-104757RB-I00/ AEI/10.13039/501100011033; the “Comunidad Autónoma de Madrid” through Grant S2017/BMD-3817; CSIC: PIE/ COVID-19 number 202020E079; European Union (EU) and Horizon 2020 through grants EOSC Life (INFRAEOSC-04-2018, Proposal: 824087) and HighResCells (ERC - 2018- SyG, Proposal: 810057). J.V. acknowledges economical support from the Ramón y Cajal 2018 program (RYC2018-024087-I).

Author contributions

Conceptualization: J.V., C.O.S.S., R.S.-G.; Methodology: R.S.-G., J.V., and C.O.S.S.; Software implementation: R.S.-G., J.G.; Evaluation: R.S.-G., J.V., A.C.; Writing: R.S.-G., A.C., J.V., C.O.S.S. and J.M.C.; Supervision: J.V., C.O.S.S.; Funding acquisition: J.M.C., J.V.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s42003-021-02399-1>.

Correspondence and requests for materials should be addressed to C.O.S.S. or J.V.

Peer review information *Communications Biology* thanks the anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available. Primary handling editors: Jung-Eun Lee, Christina Karlsson Rosenthal, George Inglis.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021

DeepEMhancer: a deep learning solution for cryo-EM volume post-processing

Ruben Sanchez-Garcia, Josue Gomez-Blanco, Ana Cuervo, Jose Maria Carazo, Carlos Oscar S. Sorzano, Javier Vargas

Content

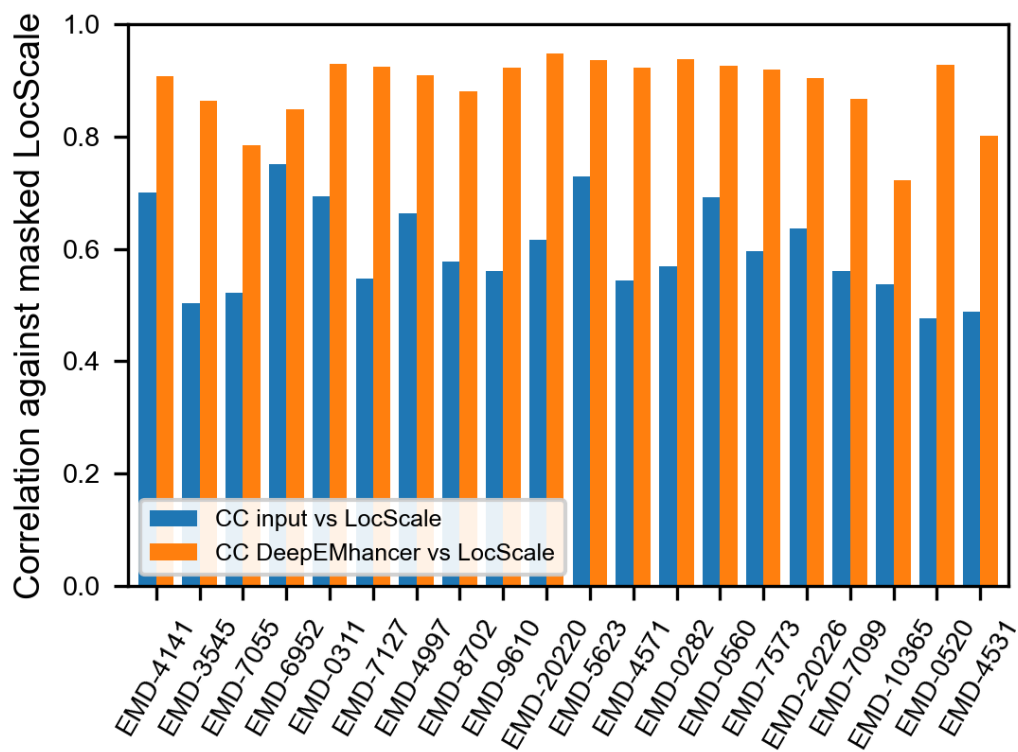
Supplementary Table 1: Neural network architecture	2
Supplementary Figure 1	4
Supplementary Note 1: Visual inspection of testing maps	5
Supplementary Figure 2	5
Supplementary Note 2: Map local quality impact on performance	6
Supplementary Figure 3	7
Supplementary Note 3: FSC curves of the studied maps	8
Supplementary Figures 4-9	9
Supplementary Note 4: Target selection impact in model performance	13
Supplementary Figure 10	13
Supplementary Figures 11-13	14
Supplementary Figures 14 and 15	17
Supplementary Note 6: Training targets generation workflow	19
Supplementary Figure 16	19
Supplementary Note 7: List of EMDB entries used in this work	20
Supplementary references	23

Supplementary Table 1. Neural network architecture

Layer	Type	Parents	#Kernels	Stride	Kernel size	Output shape
1	Conv3d+GN+PReLU	Input	32	1	5	64x64x64x32
2	Conv3d+GN+PReLU	1	32	1	5	64x64x64x32
3	Conv3d+GN+PReLU	2	32	2	5	32x32x32x32
4	Conv3d+GN+PReLU	3	64	1	5	32x32x32x64
5	Conv3d+GN+PReLU	4	64	1	5	32x32x32x64
6	Conv3d+GN+PReLU	5	64	2	5	16x16x16x64
7	Conv3d+GN+PReLU	6	128	1	5	16x16x16x128
8	Conv3d+GN+PReLU	7	128	1	5	16x16x16x128
9	Conv3d+GN+PReLU	8	128	2	5	8x8x8x128
10	Conv3d+GN+PReLU	9	128	1	5	8x8x8x128
11	Concat +Conv3d_trans	10	128	2	5	16x16x16x128
12	Conv3d+GN+PReLU	10 & 7	128	1	5	16x16x16x128
13	Conv3d+GN+PReLU	12	128	1	5	16x16x16x128
14	Conv3d+GN+PReLU	13	128	1	5	16x16x16x128
15	Concat +Conv3d_trans	14 & 4	64	2	5	32x32x32x64
16	Conv3d+GN+PReLU	15	64	1	5	32x32x32x64
17	Conv3d+GN+PReLU	16	64	1	5	32x32x32x64
18	Conv3d+GN+PReLU	17	64	1	5	32x32x32x64
19	Concat +Conv3d_trans	18 & 1	32	2	5	64x64x64x32
20	Conv3d+GN+PReLU	19	32	1	5	64x64x64x32
21	Conv3d+GN+PReLU	20	32	1	5	64x64x64x32
22	Conv3d+GN+PReLU	21	32	1	5	64x64x64x32

23	Conv3d_trans	22	16	2	5	128x128x128x16
24	Conv3d+GN+PReLU	23	8	2	5	64x64x64x8
25	Conv3d	24	1	1	5	64x64x64x1

Total number of parameters: 51,119,889



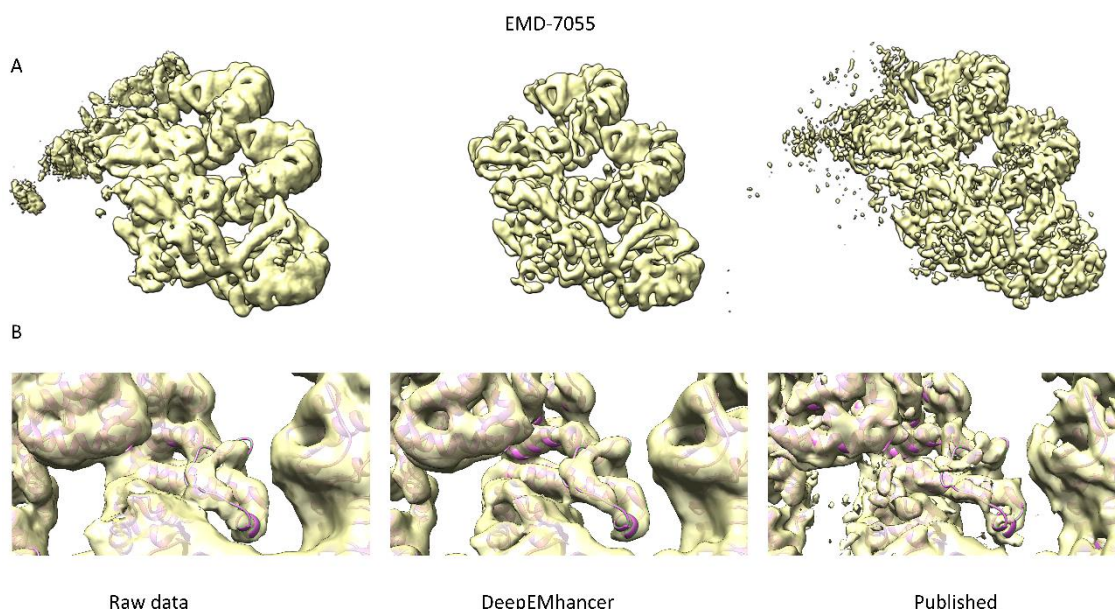
Supplementary Figure 1. Correlation coefficient for the input maps of the testing set before (blue) and after (orange) the treatment with DeepEMhancer compared against LocScale processed maps.

Supplementary Note 1: Visual inspection of testing maps

EMD-7055

The EMD-7055¹ is a medium resolution volume of the NAIP5-NLRC4-flagellin inflammasome. For the purposes of this article, the main interesting aspect of this volume is the apparent poor performance of DeepEMhancer according to Main Text Figure 1. One of the reasons behind this behaviour is the fact that a mask was applied to only three subunits at the last stages of the refinement process. As a result, the volume contains signal for both masked and unmasked subunits although their intensity levels vary severely. Consequently, our neural network has tried to restore both the originally masked and unmasked regions, and thus, the results are not as good as in the other cases. However, when the volume is carefully pre-processed in order to remove those unmasked subunits while preserving the normalization constraints, non-negligible improvements were observed. Secondly, another important reason for the poor measured metrics is the fact that the atomic model (PDB 6b5b) was obtained by means of rigid body fitting of an homology model instead of being traced, thus, the agreement between the atomic model and the density map is far from being perfect. As a result, the resolution estimates computed using the atomic model as reference are not too accurate.

Supplementary Figure 2 shows the overall aspect of the post-processed volume compared to the raw and the B-factor-sharpened ones. As can be appreciated in Supplementary Figure 2 panel A, the map produced by DeepEMhancer is much cleaner than the B-factor processed map. More importantly, although the level of detail in the core of the protein is similar, in the outer part of the protein, the B-factor sharpened map presents broken densities that look continuous in the map obtained with DeepEMhancer, thus facilitating the map interpretation.

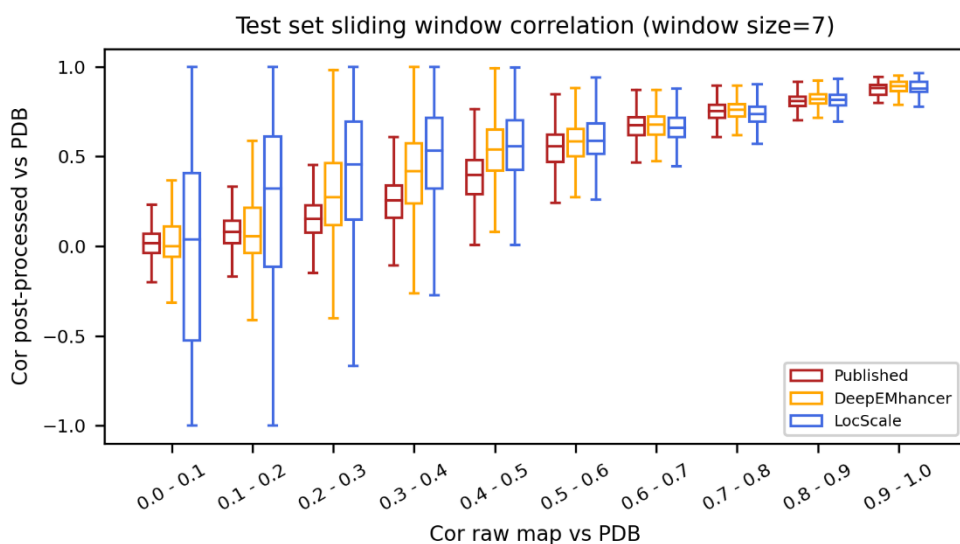
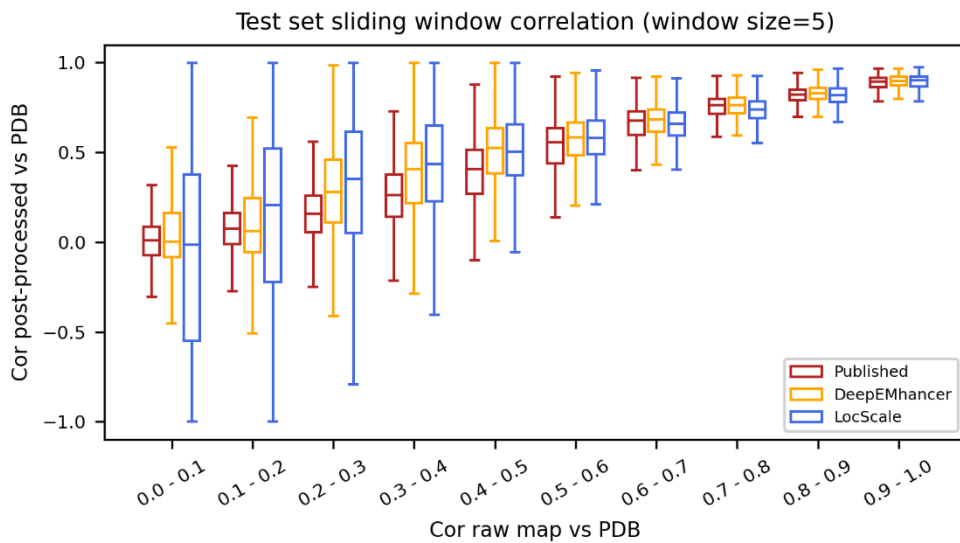


Supplementary Figure 2. DeepEMhancer post-processed volume for EMD-7055. A, overview of the raw data map, the post-processed map and the B-factor corrected map. B, zoom-in of a region containing the loop Q514-S533, which looks cleaner and better resolved in the DeepEMhancer post-processed map compared to the raw data and sharpened maps.

Supplementary Note 2: Map local quality impact on performance

In order to study how the local quality of the maps impacts the performance of DeepEMhancer, we have studied the local quality of the post-processed maps as a function of the local quality of the raw maps. Among all possible quality metrics that we could have computed, we used the local correlation between the post-processed map and the atomic model and between the raw map and the atomic model using a sliding window approach. Local resolution was discarded as a quality metric due to the wide variety of values found within the testing set. Local correlation was computed using a sliding window approach of 7x7x7 (or 5x5x5).

Supplementary Figure 3 displays the distribution of such correlation values for all maps included in the testing set. The maps were post-processed with LocScale (blue), DeepEMhancer (yellow) and with global B-factor correction as done in their publications (labelled as Published, red). As it is shown in Supplementary Figure 3, the three approaches tend to produce results of similar quality for high-quality input regions (correlation greater than 0.6). As expected, the quality of the post-processed regions decreases as the quality of the input data does, although such reduction is more pronounced for the global B-factor correction than for the other approaches, that seem comparable for the mid-quality range (correlation between 0.3 and 0.6). Finally, the quality of the post-processed maps is quite low for the correlation range between 0.1 and 0.3, being LocScale post-processed regions substantially better. In light of these results, we can state that the results produced by DeepEMhancer are more similar to the results produced by LocScale than the ones produced by global B-factor correction, and thus, they are better suited for maps of heterogeneous quality.

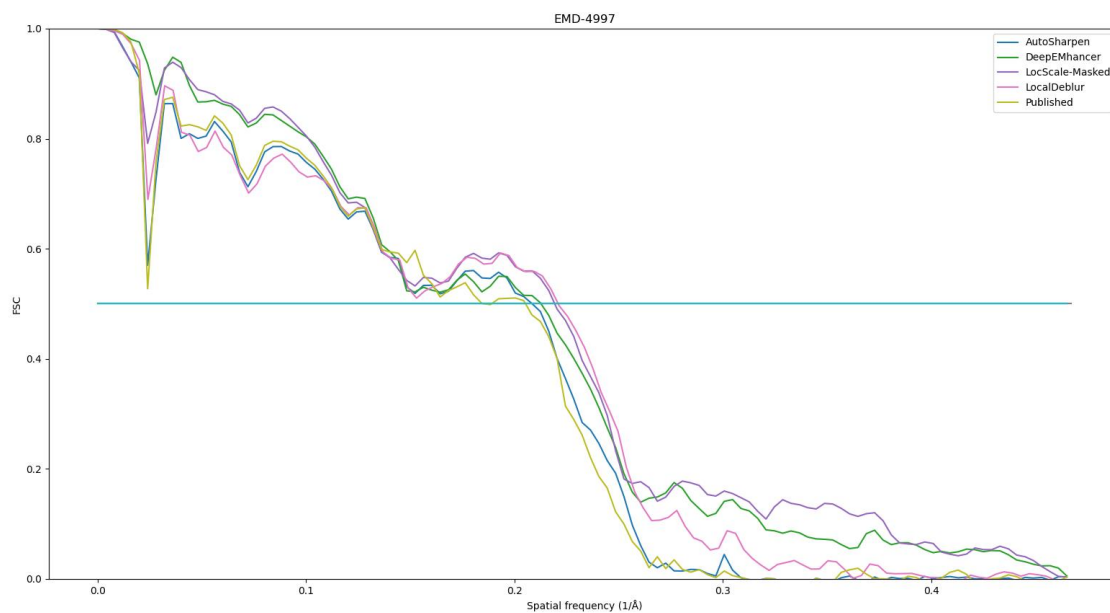


Supplementary Figure 3. Sliding window correlation between the post-processed maps and the reference (y-axis) and between the raw map and the reference (x-axis) using a window size of 5 (top) and 7 (bottom) voxels. Published maps were post-processed using global B-factor correction. Boxes enclose values between the first and third quartile (Q1 and Q3, lower and upper limit respectively). Median values are depicted as the horizontal lines within the boxes. Whiskers enclose values between $Q1 - 1.5(Q3-Q1)$ and $Q3 + 1.5(Q3-Q1)$.

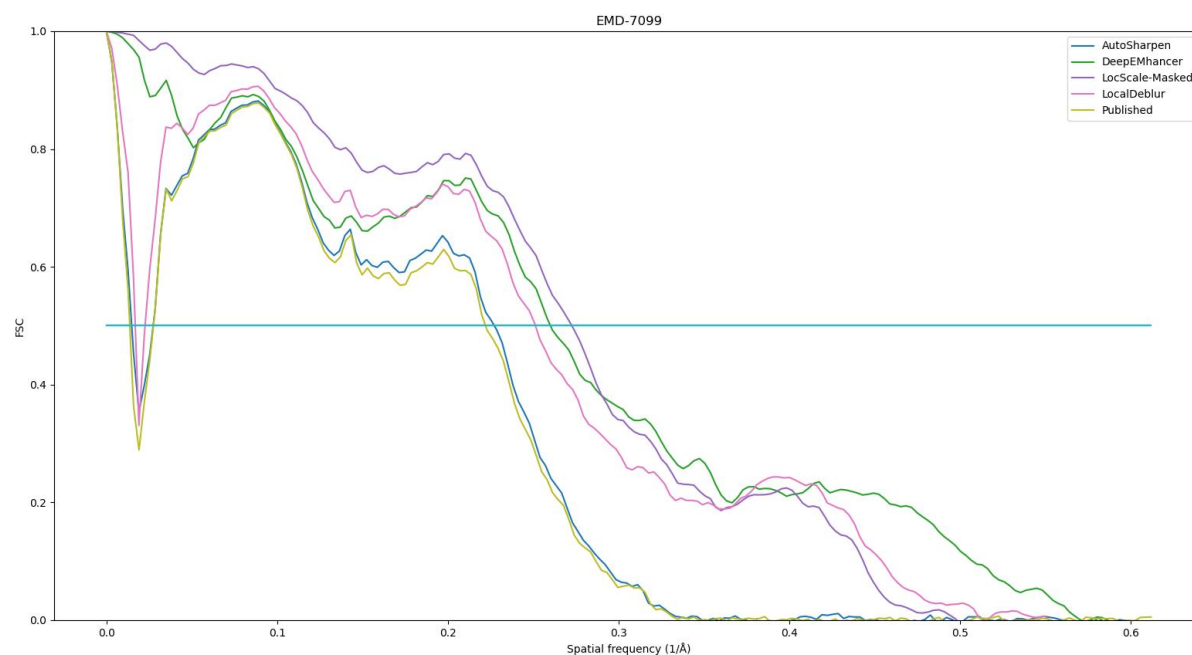
Supplementary Note 3: FSC curves of the studied maps

This section includes the FSC curves of the post-processed maps computed against the reference map derived from the atomic model for the three maps that were presented in detail in the main text: EMD-7099, EMD-4997, and EMD-30178 (Supplementary Figures 4-6). Post-processed algorithms used were LocalDeblur, LocScale, AutoSharpen and DeepEMhancer. Please notice that LocalDeblur results are making use of the atomic model information for the masks calculation and that LocalScale makes use of such information in a direct manner. Consequently, their results are expected to be among the best. Nevertheless, DeepEMhancer tends to produce results that are similar to the training targets (LocScale-Masked), being of especial quality for the EMD-7099 and EMD-30178 cases.

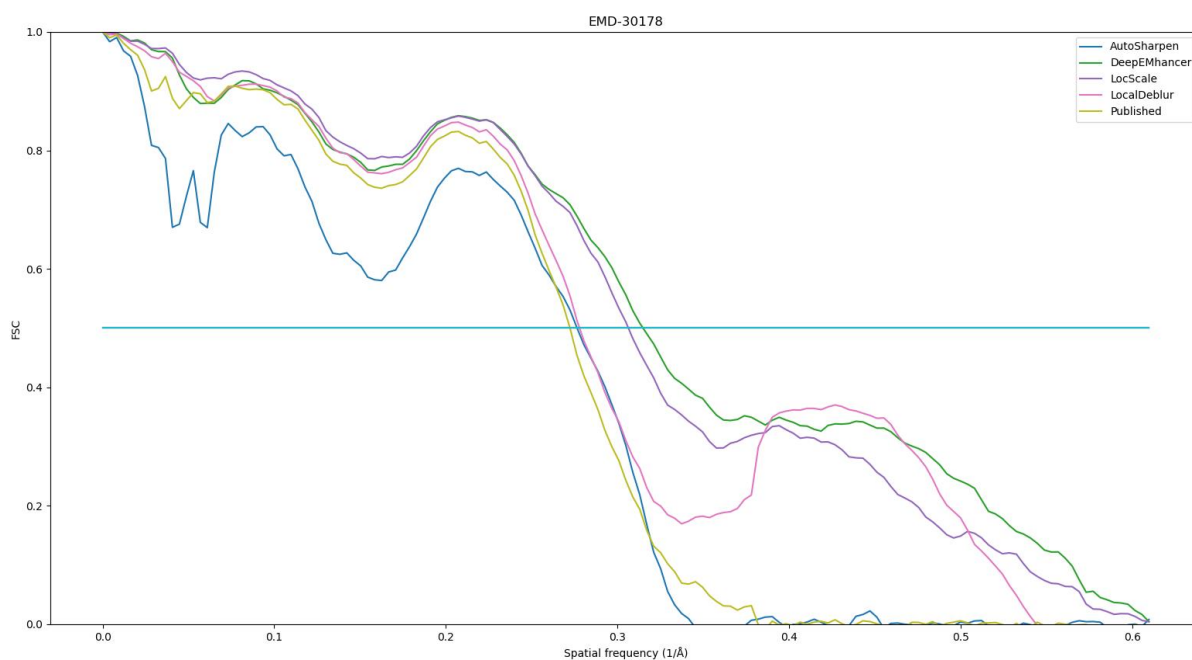
One particular caveat that can seem confusing in our plots is the bounce that the curves of LocScale-Masked, DeepEMhancer and to a lesser degree, LocalDeblur experiment. Such bounces are caused by the tight mask employed being derived from the atomic model used as reference (see Supplementary Figure 7 and 8). Concisely, the fact that the maps obtained from the atomic models introduce high frequency components due to the nature of the atomic basis functions. Masking the maps equally introduces high frequency components due to the convolution of the Fourier transform of the map with the Fourier transform of the binary mask (see Supplementary Figure 9). As a consequence, the two maps present the same kind of behaviour at high frequency thus the correlation causing the bounces in the FSC. However, we want to highlight that those bounces, that occur at FSC values <0.3 , are not relevant for the comparison against the atomic model, which should be done at threshold 0.5^4 . Nevertheless, if the bounces are desired to be removed, phase randomization could be applied to the post-processed map⁵.



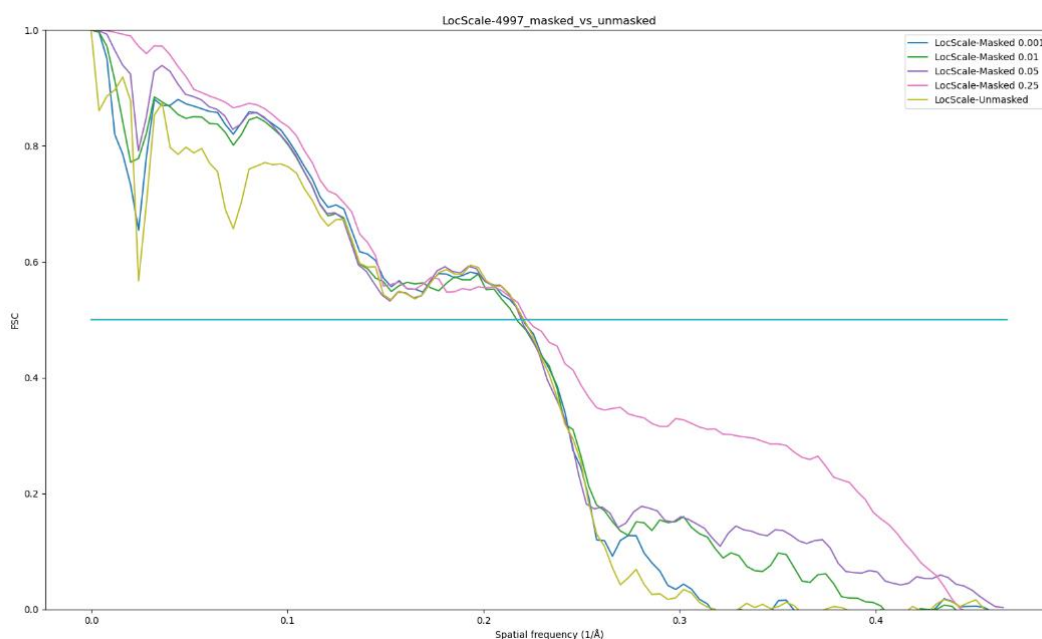
Supplementary Figure 4. FSC curves for the maps post-processed with AutoSharpen (blue), DeepEMhancer (green), masked LocScale (purple) and LocalDeblur (magenta) for the EMD-4997. FSC curves were computed using as reference the map derived from the atomic model.



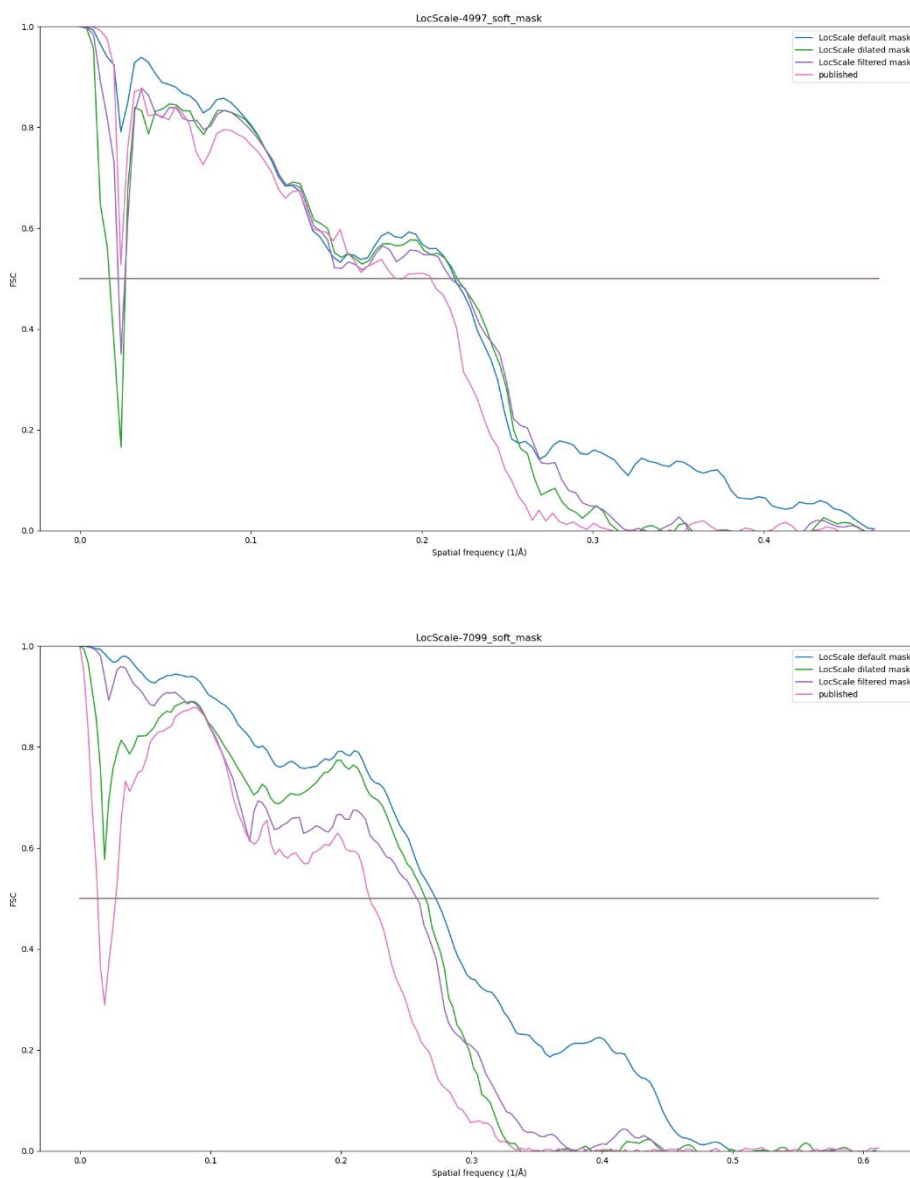
Supplementary Figure 5. FSC curves for the maps post-processed with AutoSharpen (blue), DeepEMhancer (green), masked LocScale (purple) and LocalDeblur (magenta) for the EMD-7099. FSC curves were computed using as reference the map derived from the atomic model.



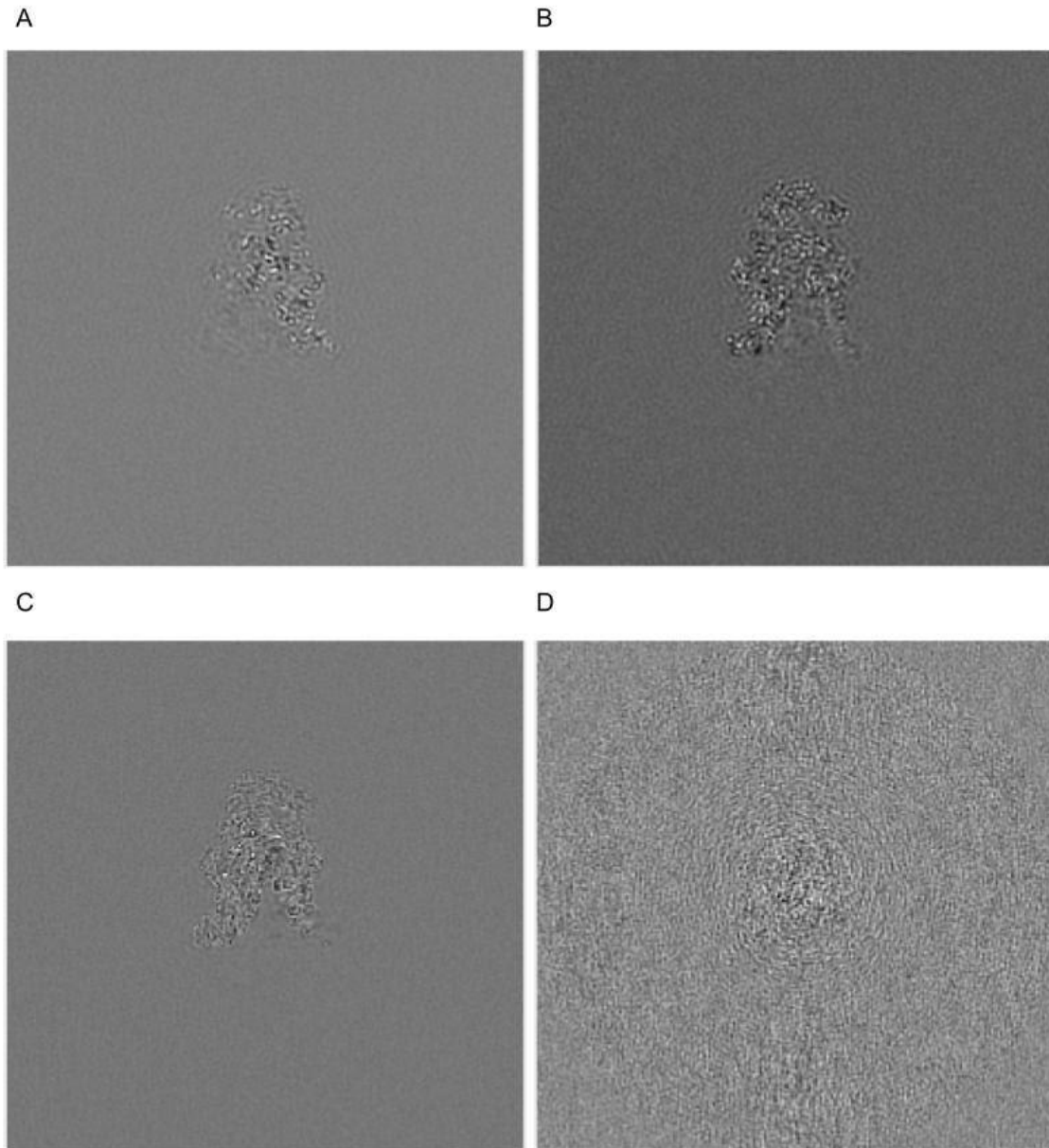
Supplementary Figure 6. FSC curves for the maps post-processed with AutoSharpen (blue), DeepEMhancer (green), masked LocScale (purple) and LocalDeblur (magenta) for the EMD-30178. FSC curves were computed using as reference the map derived from the atomic model.



Supplementary Figure 7. FSC curves for the maps post-processed with LocScale for the EMD-4997 not using (golden) the training tight mask and using masks with different degrees of tightness (blue, green, purple and magenta sorted by increasing tightness). FSC curves were computed using as reference the map derived from the atomic model. The degree of tightness is measured as the relative threshold used for mask binarization (e.g., LocScale-Masked 0.01 is computed using the binary mask obtained using as threshold the percentile 1%) for the maps post-processed with LocScale for the EMD-4997 using (blue) and not using (green) the training tight mask. FSC curves were computed using as reference the map derived from the atomic model.



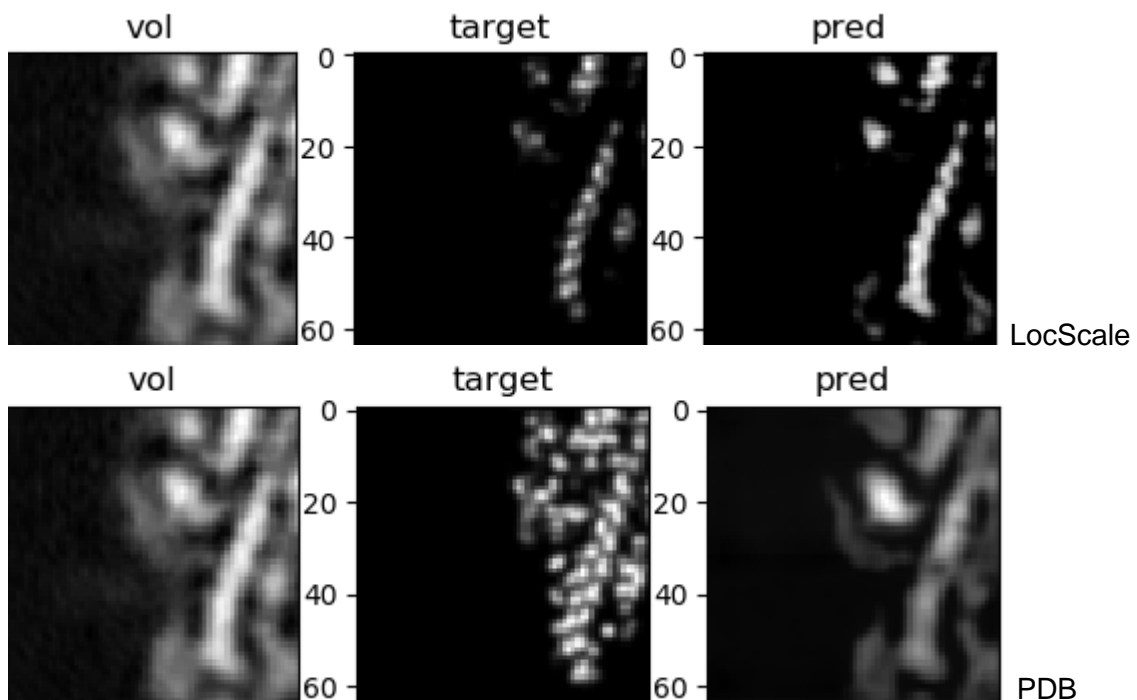
Supplementary Figure 8. FSC curves for the maps post-processed with LocScale for the EMD-4997 (top) and EMD-7099 (bottom) not using (magenta) the training tight mask and using the training mask (blue) and the train mask filtered (purple) or dilated (green). FSC curves were computed using as reference the map derived from the atomic model.



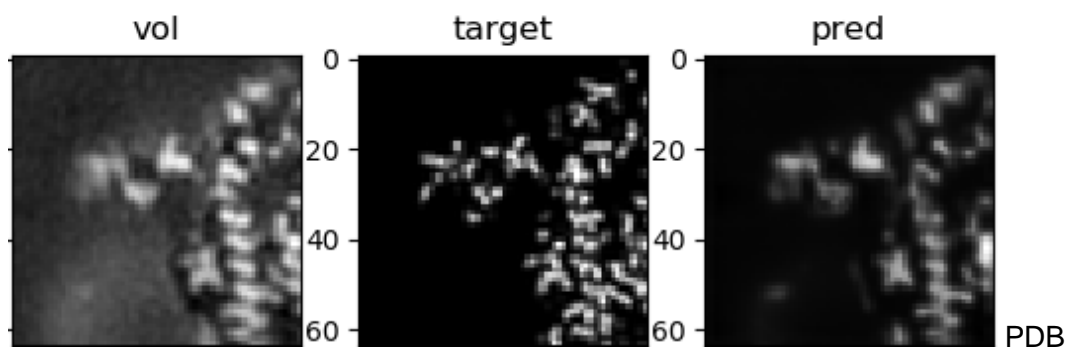
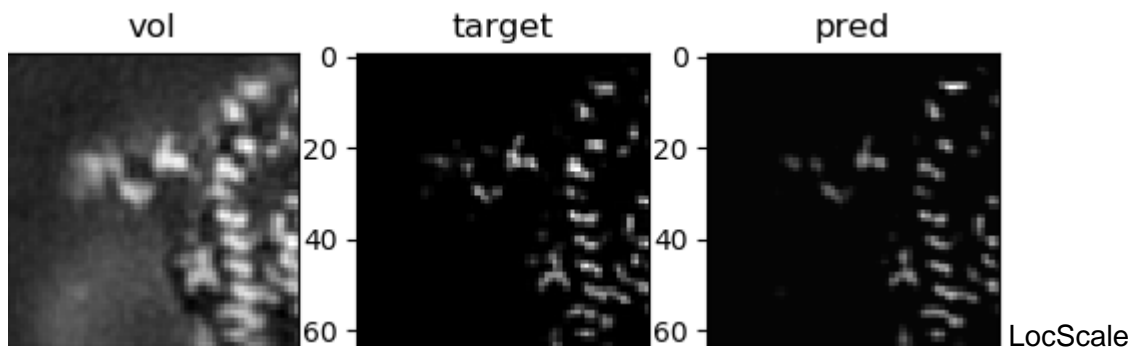
Supplementary Figure 9. Maps difference for EMD-4997 between DeepEMhancer and low-pass filtered DeepEMhancer (A), atomic model and filtered atomic model (B), LocScale-Masked and filtered LocScale Masked (C) and experimental map and filtered experimental map (D). Border induced artifacts can be observed in panels A-C.

Supplementary Note 4: Target selection impact in model performance

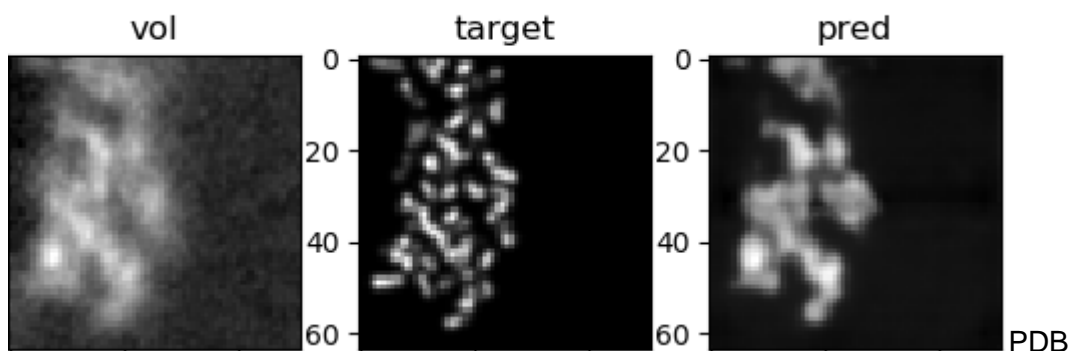
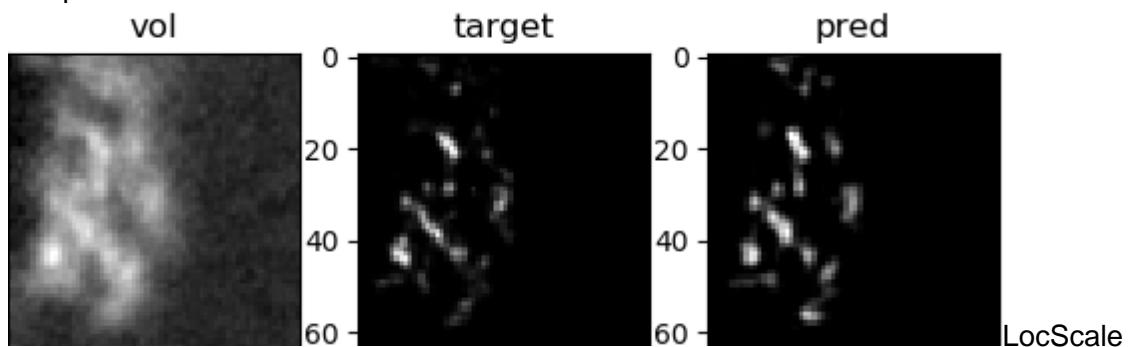
DeepEMhancer has been originally trained using as targets tightly masked volumes that were sharpened with LocScale. We also tried to train another version using as targets simulated maps derived directly from the atomic models. Although the latter option seemed to provide more accurate targets, what we found was that our implementation was not able to learn in detail how to reproduce such targets. As a consequence, the overall performance of DeepEMhancer trained using atomic models was inferior to the one trained on post-processed maps. Serve as examples the slices shown in Supplementary Figures 10-12 that illustrate how the results obtained with simulated targets look blurrier than the ones trained on post-processed maps both in training (C) and validation sets (A, B), which indicates severe underfitting. Learning curves (Supplementary Figure 13) also indicate that DeepEMhancer was not able to accurately reproduce the atomic models targets since the loss function quickly plateaus after a small reduction. In light of these evidences, as we are using the same inputs and models and the only difference between the experiments is the selected targets, we have shown that, at least for our approach, learning to reproduce atomic model targets is more difficult than post-processed maps, leading to poorer results under our available computational resources and dataset.



Supplementary Figure 10. Central slice of a 64x64x64 cube from the validation set entry EMD-6847 processed by DeepEMhancer when trained on masked LocScale targets (upper row) and simulated maps derived from PDB (lower row). vol: input volume; target: reference volume to be reproduced; pred: volume produced by DeepEMhancer.

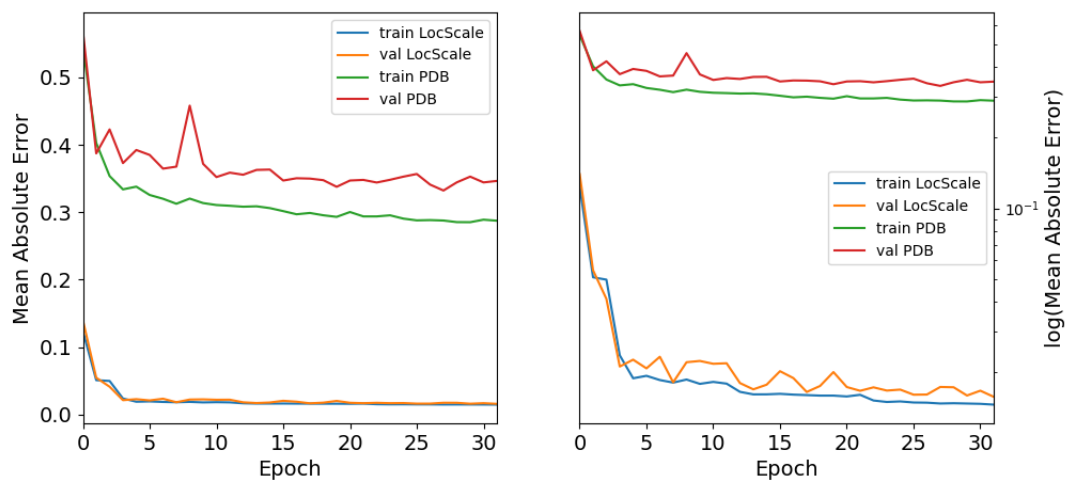


Supplementary Figure 11. Central slice of a 64x64x64 cube from the validation set entry EMD-9112 processed by DeepEMhancer when trained on masked LocScale targets (upper row) and simulated maps derived from PDB (lower row). vol: input volume; target: reference volume to be reproduced; pred: volume produced by DeepEMhancer.



Supplementary Figure 12. Central slice of a 64x64x64 cube from the training set entry EMD-20986 processed by DeepEMhancer when trained on masked LocScale targets

(upper row) and simulated maps derived from PDB (lower row). vol: input volume; target: reference volume to be reproduced; pred: volume produced by DeepEMhancer.

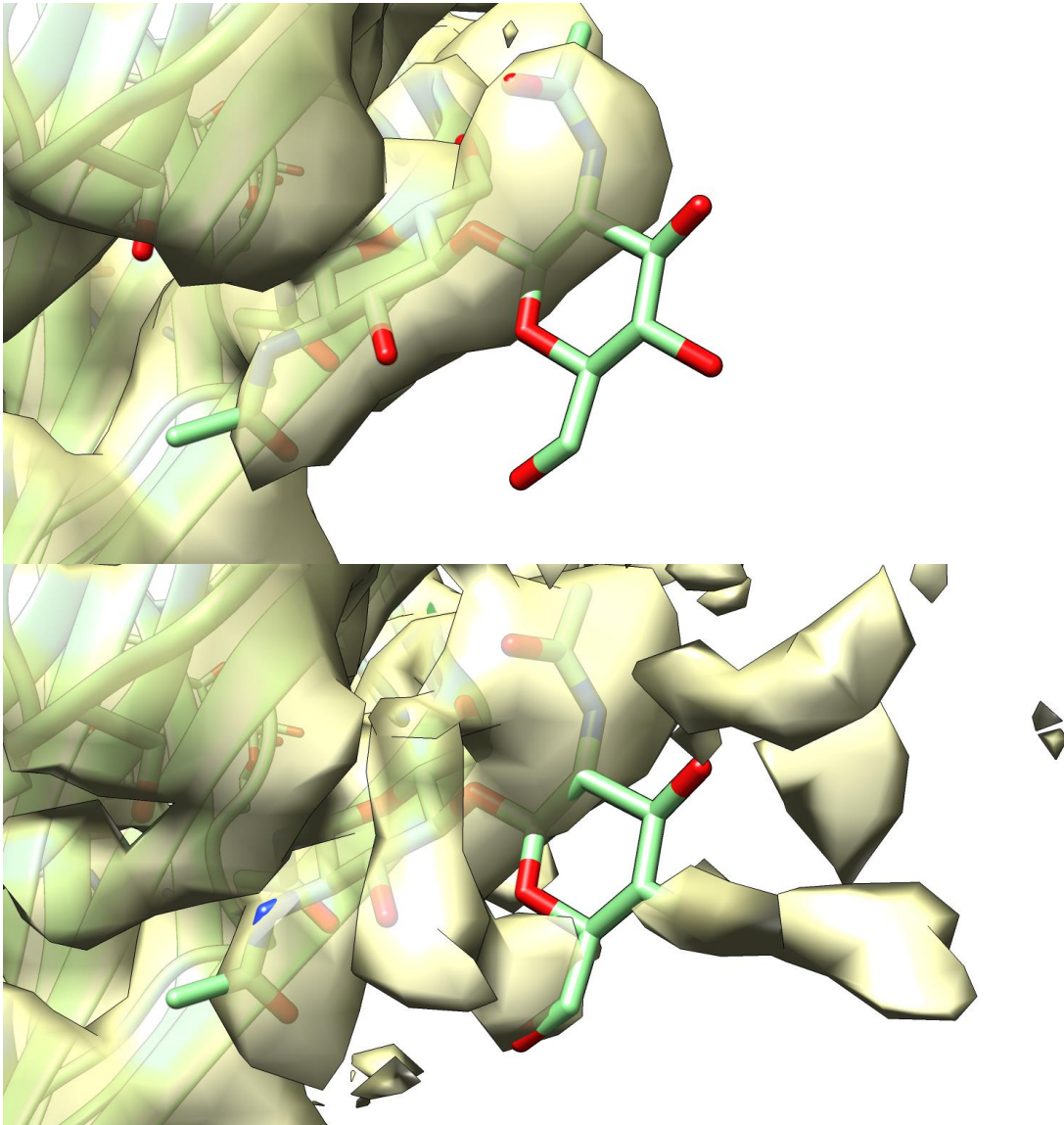


Supplementary Figure 13. Learning curves for DeepEMhancer using as targets masked LocScale post-processed volumes (blue and orange) and simulated from atomic models volumes (red and green). Both subplots differ only on the scale of the y-axis.

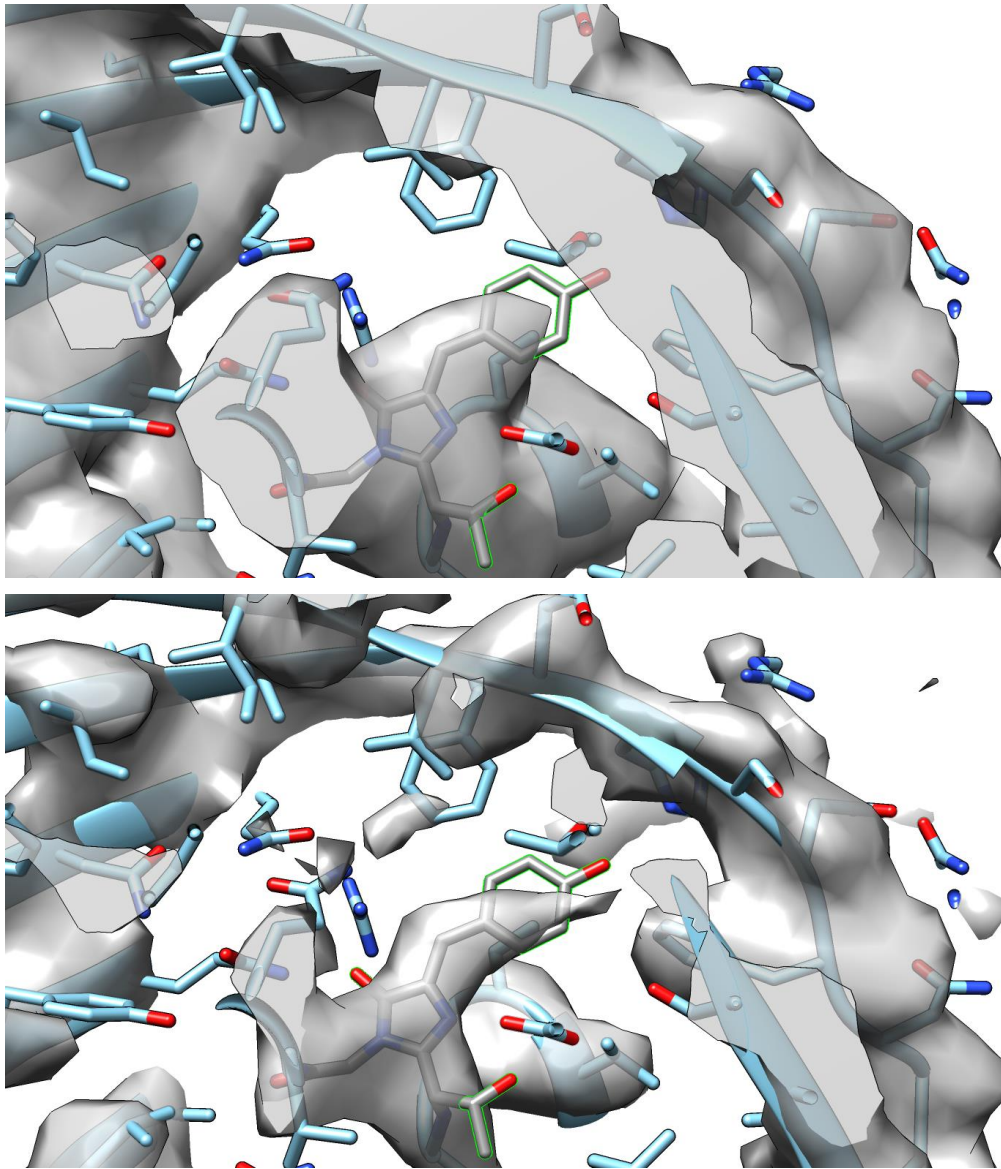
Supplementary Note 5: Dealing with post-translational modifications

Cryo-EM maps tend to exhibit heterogeneous local quality, leading to poorly defined regions in many macromolecules and consequently, unsolved regions in atomic models. This is especially true for the post-translational modifications that many residues may exhibit. Consequently, most atomic models do not include them (or only partially). Since we are making use of atomic models in the learning process, it is expected that our method will not deal well with such modifications. Indeed, we have realized that apart from a few glycans, little other examples were present in the training set. As a consequence, we have recorded a few successful examples in which glycans become more interpretable after DeepEMhancer. For instance, Supplementary Figure 14 shows one of such examples belonging to EMD-0282. Another successful instance can be found in Melero et al⁶.

On the contrary when applied to other types of modifications, worse results are expected. However, since the training set contained also ligands, which are more diverse than residues, we expect that the network will not mask out the densities corresponding to the modifications providing their intensity is strong enough. Thus, serve as an example the EMD-9374 that contains the modified residue CRO ($\{2-[(1R,2R)-1\text{-amino-}2\text{-hydroxypropyl}]-4-(4\text{-hydroxybenzylidene})-5\text{-oxo-}4,5\text{-dihydro-}1H\text{-imidazol-}1\text{-yl}\}$ acetic acid). As displayed in the Supplementary Figure 15, the density corresponding to this modified residue looks shorter in the DeepEMhancer map, and although it was not totally masked out, it is better represented in the original map. Despite this map following the trend we expected, we cannot ensure that it will be the case for any possible map and as a consequence, we recommend the users to proceed with caution in the regions they expect to find post-translational modification. Hopefully, this limitation will be reduced version after version of the program as the number of atomic models including post-translational modifications will increase in the future, when we will be able to retrain our model in a more representative dataset.



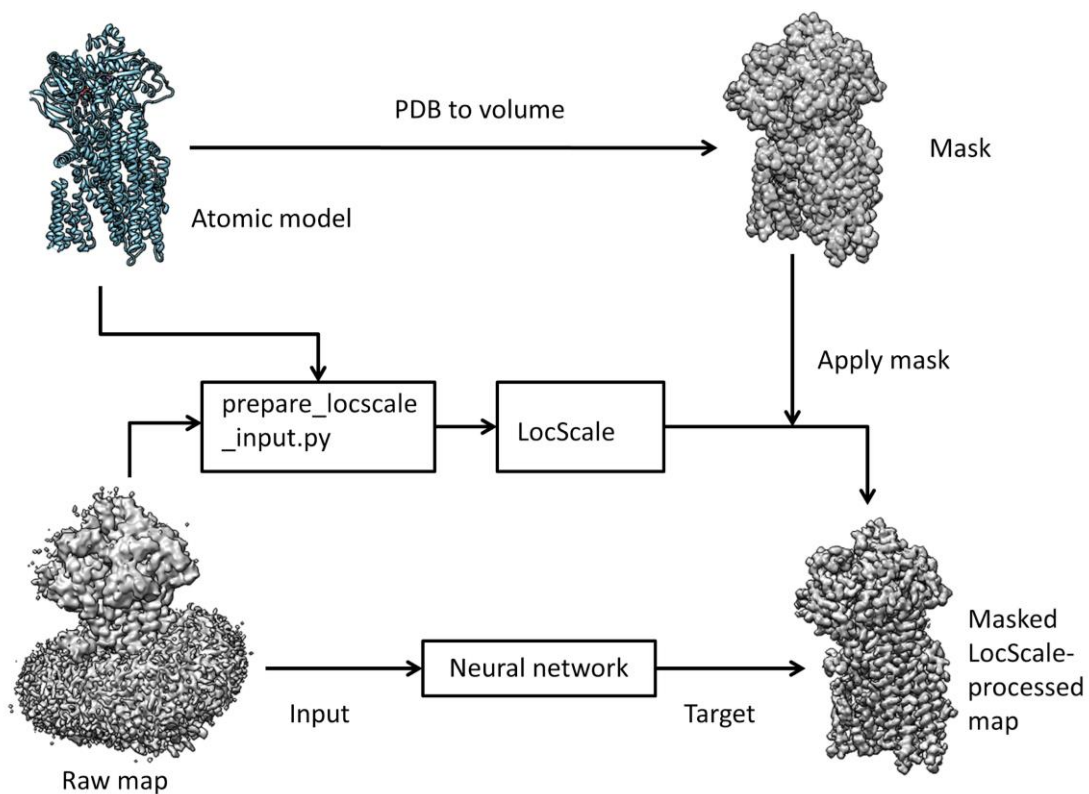
Supplementary Figure 14. Published map (bottom) and DeepEMhancer map (top) for the EMD-0282 centred at one NAG glycan.



Supplementary Figure 15. Published map (bottom) and DeepEMhancer map (top) for the EMD-9374 centered at the CRO modified residue.

Supplementary Note 6: Training targets generation workflow

Atomic models associated with each EMD entry were downloaded from the PDB database. Then, as required by LocScale², we ruled out all entries which contained non-refined atomic B-factors. For each atomic model that survived atomic B-factor filtering, we computed continuous masks normalizing the simulated volumes that were produced with the e2pdb2mrc program from the EMAN-2 suite³ using default parameters and the reported resolution. Percentile 95 was selected as threshold in order to obtain binary masks. Such tight masks were required during training in order to improve convergence since unmasked targets (that would require the network to learn how to predict random noise in order to get 0 loss) were producing worse results. Atomic models were also supplied as input, together with the average map computed from the half-maps, to the prepare_locscale_input.py program (provided in LocScale repository), that generated the corrected and reference volume required for LocScale execution. LocScale was run using as window size the recommended value of $7 * \text{average_map_resolution} / \text{pixel_size}$. Finally, the computed output was masked using the aforementioned mask. As a result, we obtained as training targets LocScale post-processed and masked versions of the input volumes.



Supplementary Figure 16. Workflow employed to generate training data for DeepEMhancer neural network.

Supplementary Note 7: List of EMDB entries used in this work

Train

EMD-0026
EMD-0038
EMD-0071
EMD-0093
EMD-0094
EMD-0132
EMD-0234
EMD-0244
EMD-0408
EMD-0415
EMD-4288
EMD-0452
EMD-0490
EMD-0500
EMD-0501
EMD-0552
EMD-0567
EMD-0589
EMD-0592
EMD-0665
EMD-0776
EMD-10049
EMD-10069
EMD-10100
EMD-10105
EMD-10106
EMD-10134
EMD-10273
EMD-10279
EMD-10324
EMD-10333
EMD-10418
EMD-10534
EMD-10585
EMD-10595
EMD-10617
EMD-20145
EMD-20146
EMD-20189
EMD-20234
EMD-20249
EMD-20254
EMD-20259
EMD-20270
EMD-20271
EMD-20352
EMD-20521
EMD-20986
EMD-21012

EMD-21107
EMD-21144
EMD-21391
EMD-3661
EMD-3662
EMD-3802
EMD-3885
EMD-3908
EMD-4032
EMD-4073
EMD-4148
EMD-4162
EMD-4192
EMD-4214
EMD-4241
EMD-4272
EMD-4401
EMD-4404
EMD-4429
EMD-4588
EMD-4589
EMD-4593
EMD-4728
EMD-4746
EMD-4748
EMD-4759
EMD-4888
EMD-4889
EMD-4890
EMD-4907
EMD-4917
EMD-4918
EMD-4941
EMD-4983
EMD-6479
EMD-7009
EMD-7041
EMD-7065
EMD-7090
EMD-7334
EMD-7335
EMD-7770
EMD-7869
EMD-8437
EMD-8438
EMD-8911
EMD-8958
EMD-8960
EMD-9111
EMD-9258
EMD-9259
EMD-9891
EMD-9931
EMD-9934
EMD-9935

EMD-9939
EMD-9941
EMD-9695

Validation

EMD-0193
EMD-0257
EMD-0264
EMD-0499
EMD-10401
EMD-20133
EMD-20449
EMD-20508
EMD-20849
EMD-4611
EMD-4646
EMD-4733
EMD-4789
EMD-6847
EMD-7133
EMD-7882
EMD-8069
EMD-9112
EMD-9298
EMD-9374
EMD-9664

Test

EMD-0282
EMD-0311
EMD-0520
EMD-0560
EMD-10365
EMD-20220
EMD-20226
EMD-3545
EMD-4141
EMD-4531
EMD-4571
EMD-4997
EMD-5623
EMD-6952
EMD-7055
EMD-7099
EMD-7127
EMD-7573
EMD-8702
EMD-9610

Supplementary references

1. Tenthorey, J. L. *et al.* The structural basis of flagellin detection by NAIP5: A strategy to limit pathogen immune evasion. *Science* (80-.). **358**, 888–893 (2017).
2. Jakobi, A. J., Wilmanns, M. & Sachse, C. Model-based local density sharpening of cryo-EM maps. *Elife* **6**, (2017).
3. Tang, G. *et al.* EMAN2: An extensible image processing suite for electron microscopy. *J. Struct. Biol.* **157**, 38–46 (2007).
4. Rosenthal, P. B. & Henderson, R. Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *J. Mol. Biol.* **333**, 721–745 (2003).
5. Chen, S. *et al.* High-resolution noise substitution to measure overfitting and validate resolution in 3D structure determination by single particle electron cryomicroscopy. *Ultramicroscopy* **135**, 24–35 (2013).
6. Melero, R. *et al.* Continuous flexibility analysis of SARS-CoV-2 spike prefusion structures. *IUCrJ* **7**, 1059–1069 (2020).