Journal of Structural Biology 213 (2021) 107780

Contents lists available at ScienceDirect

Journal of Structural Biology

journal homepage: www.elsevier.com/locate/yjsbi

Cryo-EM density maps adjustment for subtraction, consensus and sharpening

E. Fernández-Giménez [a,b], M. Martínez [a], R. Sánchez-García [a], R. Marabini [b], E. Ramírez-Aportela [a], P. Conesa [a], J.M. Carazo [a], C.O.S. Sorzano [a,c,*]

[a] Centro Nac. Biotecnología (CSIC), c/Darwin, 3, 28049 Cantoblanco, Madrid, Spain
[b] Univ. Autónoma de Madrid, 28049 Cantoblanco, Madrid, Spain
[c] Univ. San Pablo – CEU, Campus Urb. Montepríncipe, 28668 Boadilla del Monte, Madrid, Spain

ARTICLE INFO

Keywords:
Subtraction
Sharpening
Map fusion
SPA
Subtomogram averaging
Cryo-EM

ABSTRACT

Electron cryomicroscopy (cryo-EM) has emerged as a powerful structural biology instrument to solve near-atomic three-dimensional structures. Despite the fast growth in the number of density maps generated from cryo-EM data, comparison tools among these reconstructions are still lacking. Current proposals to compare cryo-EM data derived volumes perform map subtraction based on adjustment of each volume grey level to the same scale. We present here a more sophisticated way of adjusting the volumes before comparing, which implies adjustment of grey level scale and spectrum energy, but keeping phases intact inside a mask and imposing the results to be strictly positive. The adjustment that we propose leaves the volumes in the same numeric frame, allowing to perform operations among the adjusted volumes in a more reliable way. This adjustment can be a preliminary step for several applications such as comparison through subtraction, map sharpening, or combination of volumes through a consensus that selects the best resolved parts of each input map. Our development might also be used as a sharpening method using an atomic model as a reference. We illustrate the applicability of this algorithm with the reconstructions derived of several experimental examples. This algorithm is implemented in Xmipp software package and its applications are user-friendly accessible through the cryo-EM image processing framework Scipion.

1. Introduction

Cryo-EM is becoming a widely used technique for the determination of the atomic structure of proteins and macromolecular complexes. The number of density maps reconstructed from cryo-EM data is increasing both in 'Single Particle Analysis' (SPA) and 'Subtomogram Averaging' (StA). However, the comparison between these reconstructions is still an open problem in the field and usually implies volume subtraction.

To compare reconstructions, the volumes must be the same size and aligned, but also they must be in a common numerical frame and they should have comparable energy, both in real and Fourier space. There are recent proposals for comparing reconstructed volumes which perform volume subtraction as TemPy:DiffMap (Joseph et al., 2020) by making an amplitude scaling in Fourier space moving a small window along the two maps to be compared. This amplitude scaling can be performed globally, as was suggested by Terwilliger et al. (2020). Structure subtraction is also a key step for some angular alignment approaches, notably focused classification (Bai et al., 2015; Punjani et al., 2017).

In this article, we propose an enhancement of this approach to adjust the numerical values of the two volumes before subtracting and thus, we expect better subtraction results, and consequently, more accurate differences in densities. This difference can be performed between two cryoEM maps, a cryoEM map and an atomic model which will be converted internally into a density map, or between an SPA and an StA maps. Moreover, the proposed adjustment can be used for other applications. First, if we invert the roles of the reference (now the atomic model) and adjusted volume (now the SPA map), the adjustment operator acts as a map sharpener. Second, if we have several reconstructions of the same macromolecule (for example obtained from different reconstruction algorithms), we have observed that usually each of these reconstructions have parts better resolved than others, while other parts are worse. We may use the adjustment procedure to make sure that all maps are in the same numerical framework and then apply a map fusion

* Corresponding author at: Centro Nac. Biotecnología (CSIC), c/Darwin, 3, 28049 Cantoblanco, Madrid, Spain.
E-mail address: coss@cnb.csic.es (C.O.S. Sorzano).

https://doi.org/10.1016/j.jsb.2021.107780
Received 13 April 2021; Received in revised form 4 August 2021; Accepted 9 August 2021
Available online 29 August 2021

technique based on wavelets. In this way, we have a volume that is a consensus of the input reconstructions, trying to keep the best quality parts of the different inputs. To the best of our knowledge, this is the first attempt to combine multiple reconstructions of the same structure obtained by different methods into a single map.

2. Methods

2.1. Numerical adjustment of volumes

We define the set of operations necessary to assimilate the values of one volume to another without losing the structural information as numerical adjustment of volumes. If this adjustment is not performed, the different gray scales and energies of the two volumes make a quantitative comparison impossible. Thus, to get comparable maps, we have developed an adjustment algorithm based on projectors onto sets, specifically being all of them, except for the first one, projectors onto convex sets (POCS) (Madisetti and Williams, 1999).

Lets note the two volumes to be adjusted as V_1 and V_2. V_1 will be chosen as the "reference" volume, while V_2 will be modified to be numerically adjusted to the reference volume. To perform this adjustment, we use five projectors in an iterative way, concentrating their effects to a given region in real space, to transform a map V_2 into another map whose numerical values are as close as possible to those of a reference map V_1, previously restricted by a non-negative constraint.

The first projector applies the Fourier amplitude of the reference volume to the input volume (which will be adjusted), in order to have both volumes with comparable energies. The amplitudes of the reference volume are not applied directly, but they are modulated by the quotient of the radial averages (denoted by a over line) of the Fourier magnitudes of the reference volume and the input volume. Given the Fourier transform of a map, $\widehat{V}(\boldsymbol{\omega})$, and the Fourier transform of the reference, $\widehat{V}_1(\boldsymbol{\omega})$, the proposed projector is

$$\mathscr{P}_1\left(\widehat{V}\left(\boldsymbol{\omega}\right)\right) = \widehat{V}\left(\boldsymbol{\omega}\right)\frac{\overline{\left|\widehat{V}_1\right|}\left(\omega\right)}{\overline{\left|\widehat{V}\right|}\left(\omega\right)} \tag{1}$$

This projector is similar to the approaches used in Jakobi et al. (2017) and Joseph et al. (2020).

The second projector limits the minimum and maximum gray value of the input volume in real space, $V(\mathbf{r})$, to take the minimum, m_1, and maximum, M_1 values of the reference volume.

$$\mathscr{P}_2(V(\mathbf{r})) = \min(\max(V(\mathbf{r}), m_1), M_1) \tag{2}$$

This second projector ensures that the range of gray of the adjusted map does not exceed that of the reference map.

The third projector considers two binary masks in real space, one for the reference map and another for the input map ($W_1(\mathbf{r})$ and $W_2(\mathbf{r})$), so masks must be the same size as the maps:

$$\mathscr{P}_3(V(\mathbf{r})) = W_1(\mathbf{r})W_2(\mathbf{r})V(\mathbf{r}) \tag{3}$$

This projector computes the intersection of the masks and apply the resultant mask to the modified volume. Thus, it concentrates the action of the rest of the projectors to a particular region in space, i. e. the region composed by the intersection of the masks.

The fourth projector takes the Fourier phase of the original V_2 and applies it to the current estimate of the adjusted map:

$$\mathscr{P}_4\left(\widehat{V}\left(\boldsymbol{\omega}\right)\right) = |\widehat{V}\left(\boldsymbol{\omega}\right)|\frac{\widehat{V}_2\left(\boldsymbol{\omega}\right)}{\left|\widehat{V}_2\left(\boldsymbol{\omega}\right)\right|} \tag{4}$$

The goal of this projector is to preserve the structural information of V_2 as faithfully as possible, as the phase of a map contains most of the three-dimensional information of the macromolecule.

Finally, the last projector imposes non-negativity to the adjusted map:

$$\mathscr{P}_5(V(\mathbf{r})) = \max(V(\mathbf{r}), 0) \tag{5}$$

The reason is that macromolecules should not have negative density values, and all the operations performed in Fourier space may induce some artificial ringing that should be eliminated. Although it is true that negative values can also be caused by other sources (as actual densities lower than the solvent density or imperfect CTF correction), at the point of making two structures as similar as possible while still preserving the structural details conveyed by the second structure, we cannot identify all these effects. Thus, we decided to clip negative values as they normally do not correspond to structural details of the macromolecule of interest if the background has been normalized to zero. This latter choice is not physically correct (as the Coulomb potential of the ice is not zero), but it is generally adopted as a way to prevent 3D reconstruction artifacts due to a non-zero background.

We start the iterations restricting the analysis to the common region defined by the masks W_1 and W_2 $V^{(0)} = \mathscr{P}_3(V_2(\mathbf{r}))$. Note that these masks are optional and if they are not provided, they are assumed to cover the whole input maps. Then, we sequentially apply the projectors described above. When needed we also include the Fourier and inverse Fourier transform operators ($\mathscr{F}$ and $\mathscr{F}^{-1}$). Given the current adjusted map at iteration k, $V^{(k)}(\mathbf{r})$ we produce the $k+1$-th iteration as

$$V^{(k+1)}(\mathbf{r}) = \left(\mathscr{P}_5\circ\mathscr{F}^{-1}\circ\mathscr{P}_4\circ\mathscr{F}\circ\mathscr{P}_3\circ\mathscr{P}_2\circ\mathscr{F}^{-1}\circ\mathscr{P}_1\circ\mathscr{F}\right)\left(V^{(k)}(\mathbf{r})\right) \tag{6}$$

where $\circ$ denotes operator composition.

In summary, this is an iterative method which try to look for a non-negative volume that has the energy of V_1 in Fourier space, the phases of V_2, and whose minimum and maximum do not exceed the ones of V_1. All this search is performed in a region constrained by the masks W_1 and W_2.

Since the method usually converges after five iterations, this number of iterations has been selected by default. However, the number of iterations can be modified by the user if it is observed that convergence is not reached or if it is reached in less iterations. In order to know the degree of convergence after every projector, as well as after each iteration, the difference in terms of energy is estimated between the previous result and the current one. Since this value is reported right away, the user can estimate whether the process has converged or it needs some more iterations. At the end of the iterations, let us refer to the numerically adjusted volume as $\widetilde{V}_2(\mathbf{r})$.

2.2. Applications

In the following, we illustrate three possible uses of the numerical adjustment procedure described above. Each one differs in the problem to solve and the inputs to the procedure. In all cases, it is extremely important that the two maps to adjust have been spatially registered so that both are at the same location. The numerical adjustment procedure described above has been implemented in Xmipp (de la Rosa-Trevín et al., 2013), and the different applications described below are available through the cryo-EM image processing framework Scipion (de la Rosa-Trevín et al., 2016).

2.2.1. Volume subtraction

Volume subtraction is the first of the applications. In this problem we have two volumes V_1 and V_2 and we want to see where the differences between the two volumes are. Volume subtraction has been extensively used in the field to identify small proteins in viral capsids, factors, or ligands bound to a given macromolecule. They can also be used to describe conformational changes between two different states of the same macromolecule or the same structure solved by different techniques like single particle analysis and subtomogram averaging, or single particle analysis and X-ray diffraction.

We may perform the numerical adjustment at some desired resolution (for instance, we may limit the operations to the resolution of V_1, V_2, the minimum of both, or any other resolution of interest). Let us refer as $\widetilde{V}_1$ to the lowpass filtered V_1 volume, and as $\widetilde{V}_2$ to the numerically adjusted V_2 volume filtered to the same resolution as V_1. Then, the volume subtraction is performed as

$$\Delta V(\mathbf{r}) = V_1(\mathbf{r})\left(1 - W_1(\mathbf{r})W_2(\mathbf{r})\right) + \max\left(\widetilde{V}_1(\mathbf{r}) - \widetilde{V}_2(\mathbf{r}), 0\right)W_1(\mathbf{r})W_2(\mathbf{r}) \tag{7}$$

whose interpretation is "keep the original V_1 in those regions in which the two volumes have not been adjusted ($1 - W_1(\mathbf{r})W_2(\mathbf{r})$), and compute the difference between the adjusted volumes in the remaining regions ($W_1(\mathbf{r})W_2(\mathbf{r})$)". Note that the masks W_1 and W_2 do not need to be hard, and soft masks make perfect sense.

2.2.2. Map sharpening

The numerical adjustment procedure described above can be used as sharpening method. If V_1 is an atomic model converted to a density map (Sorzano et al., 2015), and V_2 is the cryoEM map, then we may use the numerical adjustment to push the cryoEM map to have the same amplitude spectrum as the reference volume V_1, while keeping its original information, which is mostly stored in the phases of the Fourier coefficients of V_2. The result can be low pass filtered to the input resolution if the user suspect that over-sharpening is being produced. This approach of amplitude scaling, although similar to the one proposed by Terwilliger et al. (2020) and Jakobi et al. (2017), add some restraints like constraining the range of the sharpened map and the spatial region in which the adjustment is performed.

2.2.3. Map consensus

Very often, we have multiple 3D reconstructions of the same structure obtained from the same data, but using different 3D reconstruction methods like CryoSparc (Punjani et al., 2017), Relion (Scheres, 2012), or Xmipp (de la Rosa-Trevín et al., 2013; Sorzano et al., 2018). In general, none of the maps is superior to all others in all regions. Normally, we observe that some parts of the maps are better reconstructed in one of the maps, while some other parts are better preserved in some other map. A common upgraded map could include the best domains selected from all maps. In this work, we have addressed the fusion of the optimal parts of the maps into a single one based on the measure of their local quality. The local quality of the map may be estimated by local resolution (Vilas et al., 2018), the local similarity between the map and its atomic model (Ramírez-Aportela et al., 2021), or the local energy of its wavelet coefficients (Pajares and de la Cruz, 2004).

In any case, we may fuse the different maps into a single one using any measure of their local quality. In this work, we propose to do that by cherry-picking the coefficients of the wavelet transform of the volume depending on the local quality measure. This is a well-known image fusion technique (Pajares and de la Cruz, 2004). Let us consider a set of input volumes $V_1(\mathbf{r}), V_2(\mathbf{r}), \ldots, V_N(\mathbf{r})$. We refer to the wavelet transform at scale s of the input volume $V_n(\mathbf{r})$ as $\widehat{V}_n^s(\mathbf{r})$. Then, the fusion is performed by constructing a new wavelet transform that, at every location and scale, takes the wavelet coefficient from the volume with better local properties

$$\widehat{V}^s_{\text{consensus}}(\mathbf{r}) = \widehat{V}^s_{\Phi(s,\mathbf{r})}(\mathbf{r}) \tag{8}$$

where $\Phi(s,\mathbf{r})$ is a function that considers the local quality of the N input volumes at that scale, and returns the index (1, 2, …, N) from which the wavelet coefficient must be taken from. A typical operator used in image fusion, and the one used here for our experiments, is simply

$$\Phi(s,\mathbf{r}) = \underset{n}{argmax}\left|\widehat{V}_n^s(\mathbf{r})\right| \tag{9}$$

Finally the consensus map is obtained by merely inverting the 3D wavelet transform using the consensus coefficients.

In addition to being able to keep the best coefficients from all the input volumes, we may also compute the local disagreement of the different wavelet transforms so that we have an estimate of those regions where most maps agree as well as the regions in which they disagree. We propose to do so by the following qualifier:

$$\Delta(\mathbf{r}) = \max_s\left(\max_n\left|\widehat{V}_n^s(\mathbf{r})\right| - \min_n\left|\widehat{V}_n^s(\mathbf{r})\right|\right) \tag{10}$$

Note that for this fusion to work and obtain as result a quantitative fusion map, we need that all the input maps have similar numerical values, otherwise their wavelet coefficients cannot be freely combined into a single consensus wavelet transform as they are not comparable. The numerical adjustment procedure proposed in this paper has allowed us to successfully devise this map consensus operator as illustrated in Section 3.3.

3. Results and discussion

We note that our algorithm tries to match two input signals to have similar numerical values. For this reason, along this section we make a special emphasis on showing the slices of the resulting volumes in order to see the details in the signal itself (values of the pixels). In the isosurface representation of the volumes the signal is thresholded and, consequently, we cannot appreciate its internal details.

3.1. Subtraction

We have selected three examples of subtraction to illustrate the utility and versatility of the method, being the first one the subtraction between a cryo-EM map and a converted atomic model, the second example the subtraction between two different density maps, and the third example the difference between a converted atomic model and a map. In all cases, the second volume is adjusted to the first one, as described in Section 2.2.3 previously to the subtraction.

3.1.1. Map – Model: Asymmetric unit of human adenovirus 41

To generate the starting reference map, we extracted the map fraction used to model the asymmetric unit (ASU) of the atomic structure of the capsid of the human adenovirus 41 (HAdV-F41) (Pérez-Illana et al., 2021) (see Fig. 1A). Then, we fitted the ASU atomic model to the reference map (Fig. 1(B)). This fitted atomic model was then converted into a density map using electron atomic scattering factors (Sorzano et al., 2015). To have both maps in the same position, we assigned the origin of the reference map to the model-derived map which will be adjusted and subtracted from the reference map. After that, a binary mask was computed for each of the maps.

Once we have both maps in registration and the corresponding masks, we have applied our algorithm for adjustment and subtraction (described in Sections 2.1 and 2.2.1). The adjustment has been performed with five iterations. The reference and adjusted volumes have been low pass filtered at 4 Å of resolution previously to the subtraction, as this is the resolution of the input map and smallest differences will not be reliable. The low pass filter was implemented as a raised-cosine filter whose amplitude was 1 up to 4 Å and smoothly decays (as a cosine) to 0, at 2.9 Å.

The subtraction result is showed in Fig. 1C. As can be appreciated, most of the map region where the ASU was fitted has been removed. Fig. 1D shows the central slice of the reference volume, while Fig. 1F shows the central slice of the volume resulting from the subtraction performed by the proposed algorithm. The central slice of the result obtained with UCSF ChimeraX v1.1 (Goddard et al., 2017) is showed in Fig. 1E. It can be seen that the result of the proposed algorithm in the subtraction region is less noisy and the remaining density is clearly

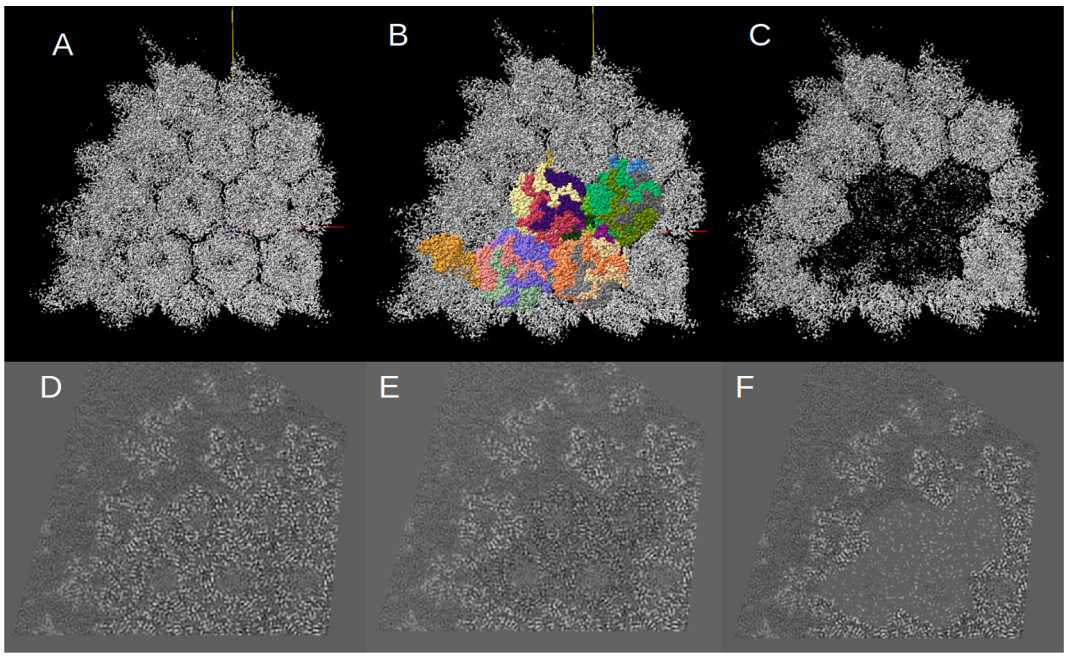

Fig. 1. A) Reference map extracted from the capsid of the human adenovirus 41 (HAdV-F41) (EMD-10768). B) AdV-F41 ASU atomic structure (PDB ID6YBA) fitted to the reference map. C) Subtraction result of reference map in A minus the map derived from the conversion of atomic model in B with the proposed algorithm. D) Central slice of reference map. E) central slice of the subtraction computed by ChimeraX. F) Central slice of C.

visible. In contrast, the result of ChimeraX is more noisy in the subtracted area and it is difficult to distinguish between remaining density and the noise.

In Fig. 2 it can also be observed that the result of ChimeraX subtraction is noisier (left column) as more small unconnected densities remains than in the result of the proposed algorithm (left column), even though all maps have been dust filtered. In the bottom part of Fig. 2 a low pass filter have been applied in order to strengthen the signal. To check if the signal is enough in the results, we focus in protein core V of the adenovirus (Rafie et al., 2021.), which is present in the reference map, but it is not traced in the atomic model of the ASU from which we derive the map to subtract. We have remarked this protein with a red rectangle in Fig. 2. As can be seen in the figure, signal is a bit higher in the case of ChimeraX low pass filtered map (C) in comparison with low pass filtered version of the difference map computed by proposed algorithm (D). Nevertheless, in the case of proposed algorithm the signal is enough to identify the lack of this protein in the reference map, easily in the low pass filtered version, but also in the original result. However, in the original result of ChimeraX even though there is more density in the region where the protein is, it is very difficult to know which densities correspond really to the protein and which are noise.

3.1.2. *Map – Map: Hepatitis-B viral capsid*

In this second experiment, we apply the subtraction between two reconstructed density maps. In order to highlight the versatility of the adjustment and subtraction algorithm, we have chosen a density map that comes from SPA reconstruction of the Hepatitis-B viral capsid as reference map and a second map of the same structure to subtract but, in this case, the map is a subtomogram average (StA) reconstruction.

For the SPA map we have chosen entry 21653 from EMDB, which has a box of 640 × 640 × 640, a voxel size of 0.65 Å and a resolution of 4.6 Å (see Fig. 3 left). The StA map comes from EMDB entry 3015, with a box of 240 × 240 × 240, a voxel size of 2.17 Å, and a resolution of 8.1 Å (see Fig. 3 right).

We have adjusted the StA map to the SPA map, as the SPA map has better resolution. The StA map was resized to have same box and pixel size than the SPA map. Then, the resized StA map was aligned to the SPA map. Once the maps were aligned, a binary tight mask was computed for each one. The adjustment was performed with five iterations. The subtraction was performed up to a resolution of 8.1 Å (the resolution of the StA map), meaning that adjusted volume has been low pass filtered at mentioned resolution previously to the subtraction. As can be seen in Fig. 4, there are differences all around the capsid, due to the difference of resolution in input maps, corroborating a lack of detail in StA map in comparison with SPA map. However, the StA reconstruction could show interesting differences in conformation due to the nature of the technique.

Thus, using this method it is possible to compare two maps of the same macromolecule obtained by different techniques and get reliable results in terms of densities as they have been numerically adjusted, which it is not the case if the subtraction is performed directly between the original SPA and StA maps as the pixel values of the maps probably will not be comparable. In Fig. 5 most notably differences between SPA and StA maps of the Hepatitis-B viral capsid are pointed by red arrows. This differences correspond to high frequency details and thus can be caused by the lack of the side chain densities in StA map due to the low resolution of the map.

3.1.3. *Model – Map*

In this section, we carried out the subtraction of the atomic model (which has been previously converted into a density map using a method based on Electron Atomic Scattering Factors, Sorzano et al. (2015)) of the envelope trimer protein of HIV BG505 in complex with the rabbit antibody E70 Fab (PDB ID6P62) from the experimental cryo-EM map (EMD-20259) from which the atomic structure was modeled.

A and B of Fig. 6 shows the central slices of the atomic model converted into density map and the experimental cryo-em map. From C to F of the same figure the results for ChimeraX, the proposed algorithm, TemPy:DiffMap method (Joseph et al., 2020) using local and global modes are shown respectively. As can be seen, the global approach of Tempy:Diff as well as ChimeraX have very noisy structures with many negative values (black pixels) and a very noisy background. For this reason, we focus the comparison of the results to the local approach of TemPy:DiffMap. The local DiffMap is noisier than the result of proposed algorithm, hindering what could be considered as real differences between the inputs and noise, also due to the presence of negative pixels

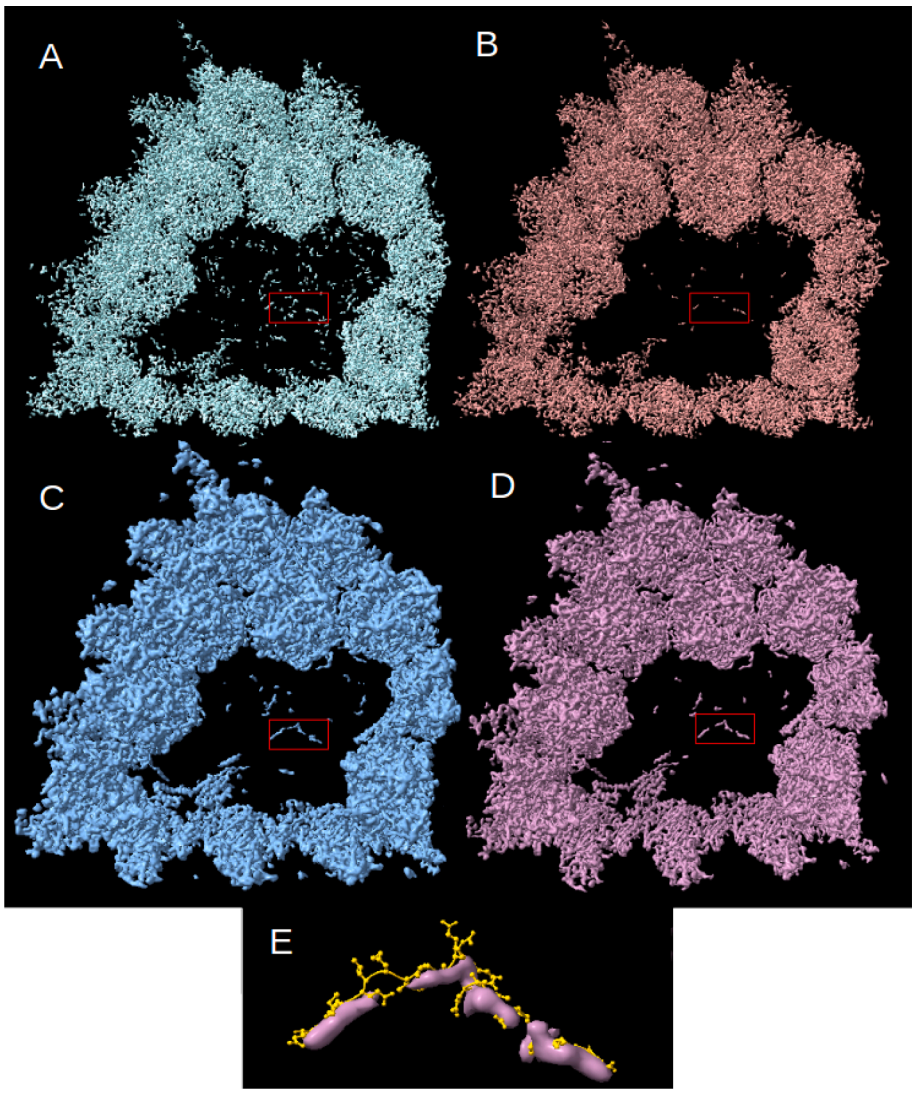

Fig. 2. Subtraction result of reference map minus map derived from ASU model performed by: A) ChimeraX subtraction B) proposed algorithm C) low pass filtered version of map in A D) low pass filtered version of map in B. Red rectangle enclosed core protein V, which is in the reference map but it is not traced in the model. E) Atomic structure of the adenovirus core V protein fitted into the density remarked in D.

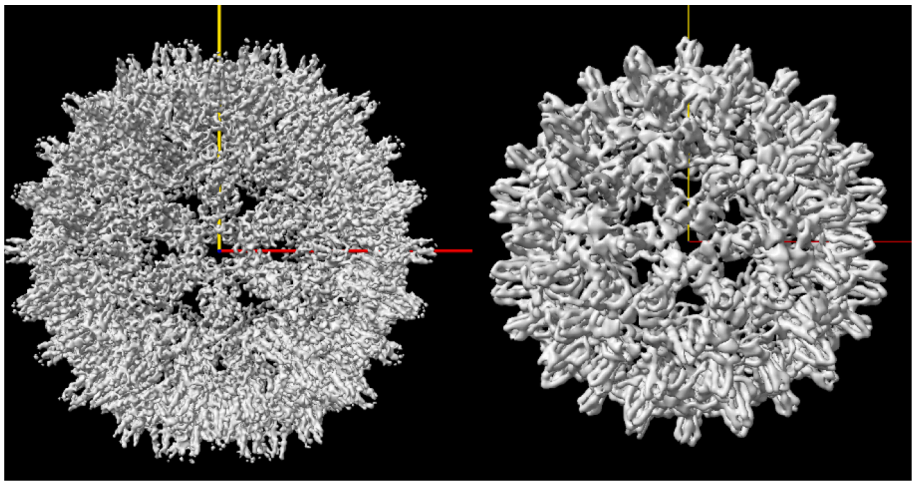

Fig. 3. Reconstruction of the Hepatitis B virus capsid using SPA (EMD-21653) (left) and StA (EMD-3015) (right).

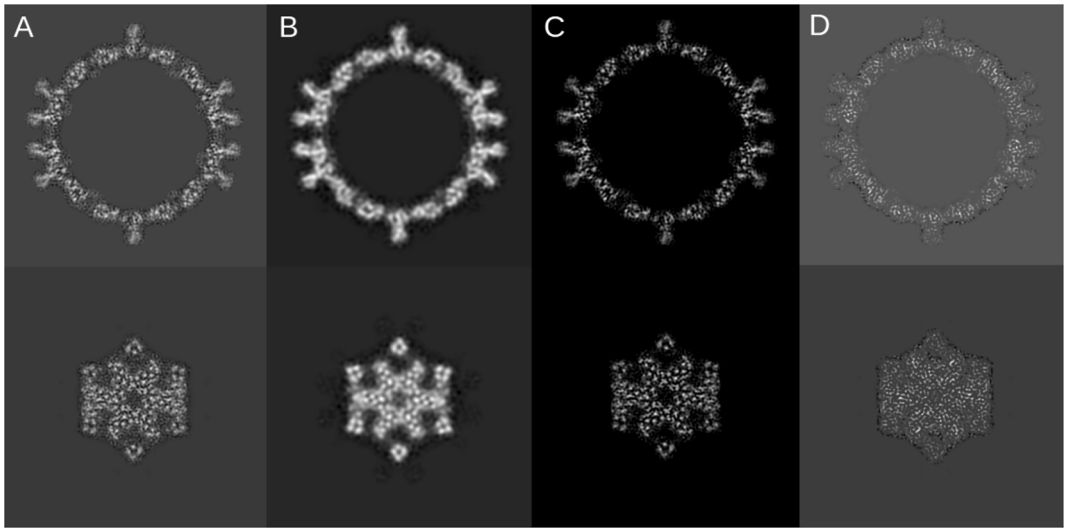

Fig. 4. Two slices from SPA Hepatitis-B viral capsid (left column), StA Hepatitis-B viral capsid (center left column), adjusted StA map to SPA map (center right column) and subtraction between both (right column).

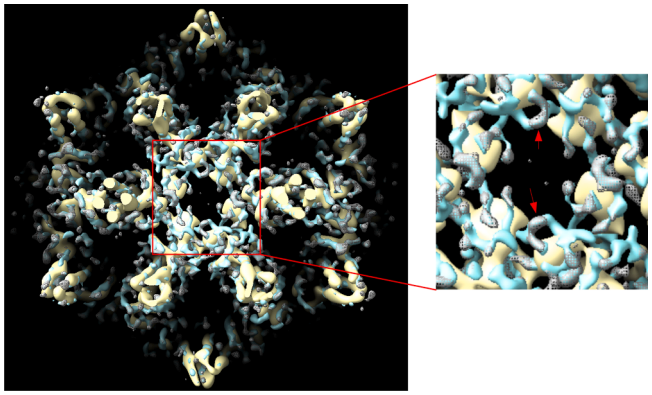

Fig. 5. Bottom view of the bottom slice in Fig. 4 of the Hepatitis-B viral capsid displayed in ChimeraX. StA volume rendered in yellow, SPA volume in light blue and difference volume in grey as a mesh. Red arrows point some remarkable differences between the subtracted volumes.

(the ones that are darker than the background) in the case of DiffMap. Nevertheless, both algorithms seem to mostly agree in the bigger differences (brighter pixels), however the brighter regions appear more blurred in the result of DiffMap.

We show in Fig. 7 the correlation of the Fourier Shell Correlation (FSC) of the difference map in Fig. 6D with the input map in Fig. 6A (FSC 1, in blue) and the input map in Fig. 6B (FSC 2). As can be seen, the FSC 1 rather correlates with the FSC of the difference map as the FSC 1 comes from a model converted into density map (Fig. 6A) and thus, in it there is high frequency information that is lacking in Fig. 6B, and that high frequency information remains in the difference map (Fig. 6D). In the case of correlation with the FSC 2, (that comes from the map in Fig. 6B), ideally the correlation would be near to 0 for all the frequencies. However, in this case the FSC 2 does not correlate that much (the correlation is smaller than 0.5, actually 0.4 is the higher correlation) so thus this indicates that there is no over-subtraction at all.

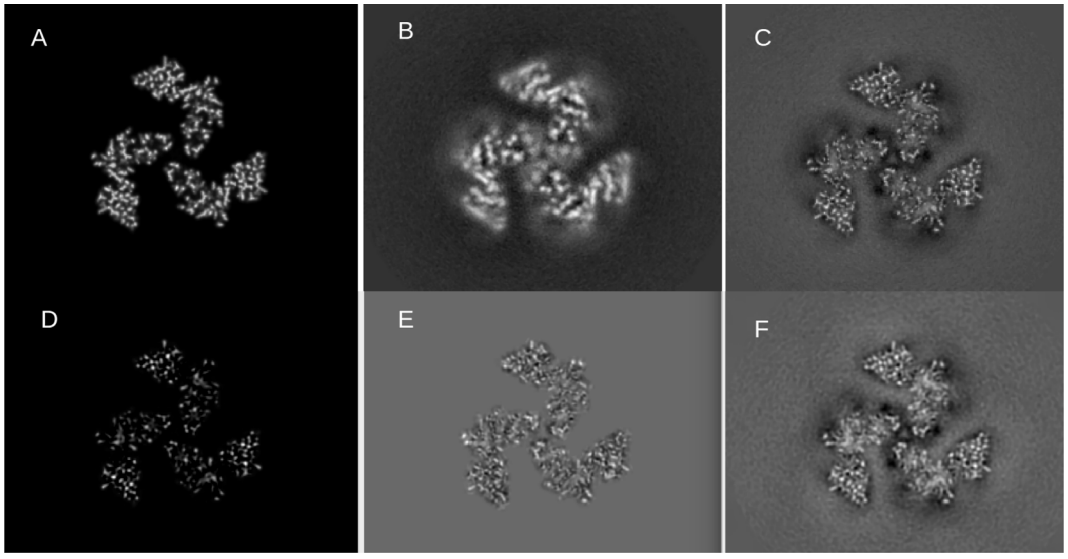

Fig. 6. Central slices of: A) PDB ID6P62 converted into density map. B) Cryo-EM density map of same structure (EMD-20259). C) Subtraction performed by ChimeraX. D) Subtraction performed by proposed algorithm. E) Local subtraction by TemPy:DiffMap. F) Global subtraction by TemPy:DiffMap.

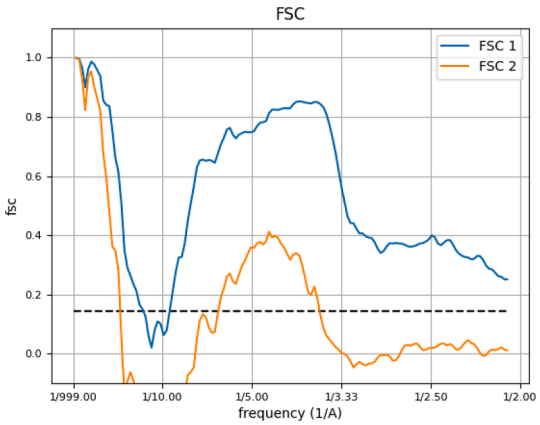

Fig. 7. Fourier Shell Correlation (FSC) of the difference map in Fig. 6D with the input map in Fig. 6A (blue) and the input map in Fig. 6B (orange).

3.2. Map sharpening

As explained in Section 2.2.2, the adjustment algorithm can be used as a post-processing step to sharpen and denoise the reconstructed map. To achieve that, a density map previously generated from an atomic model is used as reference and the experimental map is adjusted to it. To illustrate this implementation of the method, we have used an apoferritin map (EMD-11122) and its associated atomic model (PDB ID6Z9F).

The central slices of the input volume (A) and detail of the region remarked (B), reference (E) and sharpened volumes by different methods are shown in Fig. 8. We compared our result to the ones of other state-of-the-art sharpening methods: Phenix (Terwilliger et al., 2018), LocScale (Jakobi et al., 2017), DeepEMHancer (Sanchez-Garcia et al., 2020), and LocalDeBlur (Ramírez-Aportela et al., 2020).

As can be seen in Fig. 8, in this particular case, LocScale (C) has masked the structure, but the information remains very noisy and it is difficult to observe difference between the original and the sharpened map and Phenix sharpening (D) has not changed significantly the original volume. DeepEMHancer (G) denoised considerably the map, but the density remains blurred and the high frequency details are lost in some regions. This was expected as DeepEMHancer algorithm performs better in volumes whose resolution is worse than the one used here (1.56Å). In the case of LocalDeBlur (H), it has over-sharpened the result. The result of the newly proposed algorithm (F) appears denoised and sharpened and the map looks the most similar to the converted atomic model. The result is currently low pass filtered at the resolution of the original map (1.56 Å) in order to avoid over-sharpening.

In Fig. 9 we show a plot of the energy decay for the original volume, the converted atomic model and the different sharpening methods. Note that the curves are adjusted at medium frequencies (0.07 to 0.12) in order to compare the relative positions of the curves in low and high frequencies. It can be seen that LocalDeBlur (blue), LocScale (red) and Phenix (pink) have a general fall of energy similar to the one of the original map (yellow), but LocalDeBlur and LocScale move away in the high frequencies. However, the decay of energy of the sharpened results must assimilate as much as possible to the one of the converted atomic model (orange). In this case it can be seen that DeepEMHancer (green) and the proposed algorithm (purple) are the nearest. Note that the proposed algorithm actually uses the atomic model in the sharpening process, while LocalDeBlur, DeepEMHancer and Phenix do not. In the case of LocScale, it also uses the atomic model in the sharpening process but it energy decay is closer to the map in intermediate frequencies.

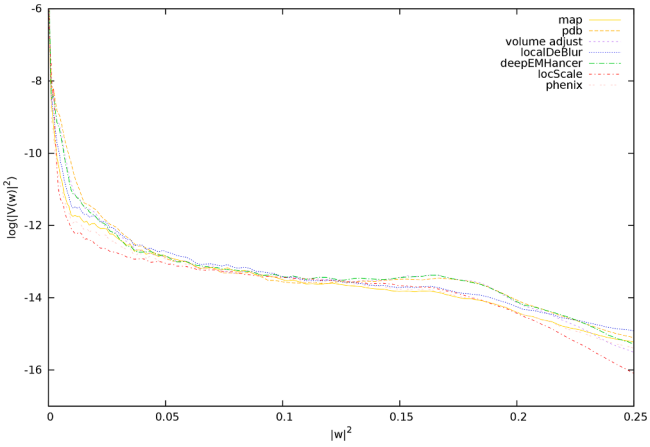

Fig. 9. Energy decay plot of original volume, converted atomic model and different sharpening algorithms. Curves are adjusted at medium frequencies.

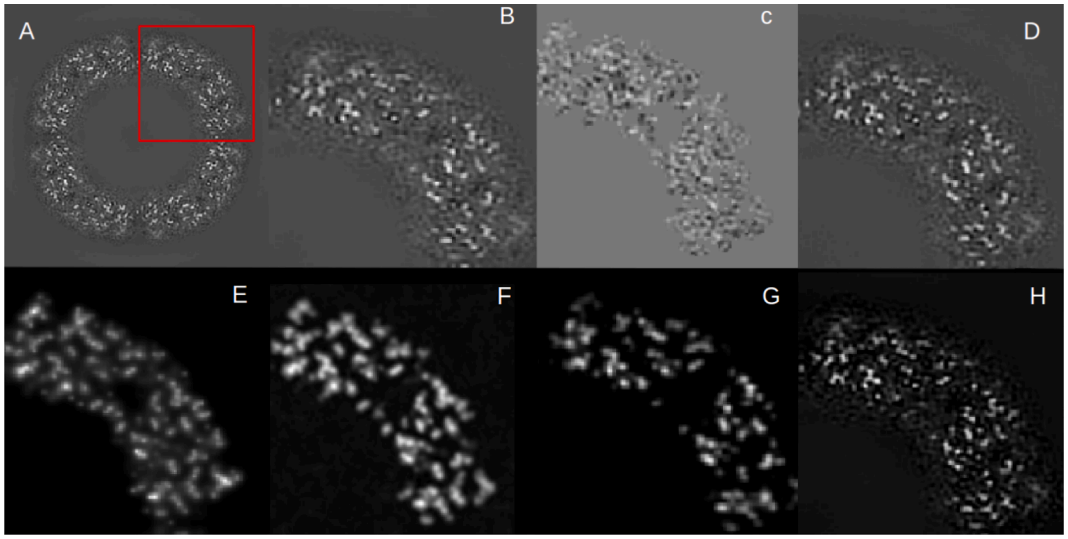

Fig. 8. Central slices of: A) Apoferritin density map EMD-11122. B) Detail of the region remarked in red in A. C) Result of LocScale. D) Result of Phenix sharpening. E) Apoferritin density map derived from atomic model PDB ID 6Z9F. F) Result of proposed algorithm as sharpening method low pass filtered at the input map resolution (1.56 Å). G) Result of DeepEMHancer. H) Result of LocalDeBlur.

3.3. Consensus

We show here two examples of consensus maps. The first example uses several apoferritin experimental maps and it is meant to show the correctness of the method in a setup in which the solution is known. The second example shows its application to several reconstructions of the Sars-CoV-2 Spike.

3.3.1. Apoferritn

As a proof of concept, we have adjusted and fused six different reconstructions of apoferritin from EMDB with different resolutions (entries 0144 with a resolution of 1.64 Å, 6800 with a resolution of 2.90 Å, 6801 with a resolution of 3.20 Å, 3854 with a resolution of 3.15 Å, 3853 with a resolution of 2.50 Å and 4213 with a resolution of 2.14 Å, see Fig. 10). The six maps have been cropped to 200 × 200 × 200, as it was the smallest input box size and their voxel size has been set to 0.81 Å, as it was the smallest input voxel size.

Then, the six volumes were aligned and adjusted having as common reference the first volume (EMD-0144), as it was the one with best resolution (1.65 Å). The first volume was also adjusted to itself, in order to be strictly positive and low pass filtered like the rest. Each map was adjusted to the reference and low pass filtered to its own resolution. Then, we used the six adjusted volumes as input for the consensus. No parameters are required for this algorithm.

The consensus volume is shown in Fig. 11. The color scale represents the degree of difference between the inputs, being dark blue the smallest differences and red the largest ones. As expected, the consensus volume is very similar to the one with best resolution. This fact confirms the correct performance of the consensus algorithm, which is keeping in each region the information of best resolution.

In Fig. 12, the central slice of consensus volume (A and B) and the central slice for the best input volume (C) are displayed together with the central slice of PDBPDB6WXWX6 converted into density map (D) as reference. It can be seen that in general all the signal from the edges of the map is reinforced in consensus result in comparison with the input map of best resolution. Red arrows in the Fig. 12 points to regions with differences between input volume (C) and consensus (B), where the signal in the consensus is reinforced and thus, is more similar to the signal in D. These are small differences but they point out that the algorithm, as designed, is taking the most energetic coefficients of any of the input volumes, and that in these regions, not all the input volumes agree.

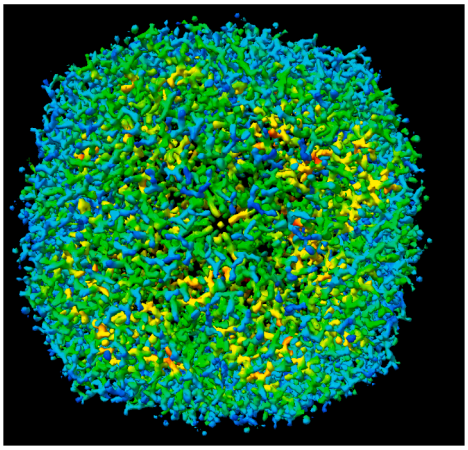

Fig. 11. Isosurface representation of the consensus of the six reconstructions of apoferritin displayed. The colors represent the degree of similarity between the signal of the input maps, being blue the smallest differences between input volumes and red the biggest differences.

3.3.2. SARS-CoV-2 spike

To better illustrate the utility of volume consensus, we have used this algorithm with different reconstructions of the spike of the SARS-CoV-2 obtained from the same dataset with different reconstruction algorithms: CryoSparc (Punjani et al., 2017), Relion (Scheres, 2012) and Xmipp HighRes (Sorzano et al., 2018). Two of the input maps correspond to the additional maps (sharpened maps) of entry EMD-11328, which processing details are described in Melero et al. (2020), while the other three are in-home reconstructions with same softwares and

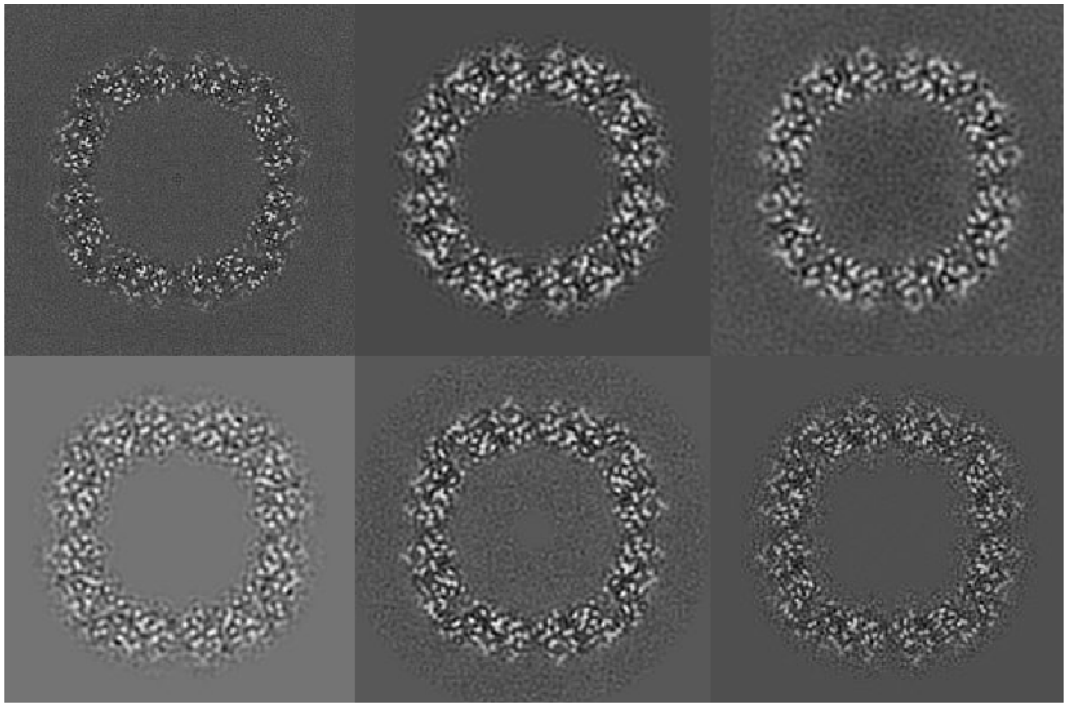

Fig. 10. Central slices of six different reconstructions of apoferritin from EMDB with different resolutions. They have been used as inputs for the consensus algorithm described in this article.

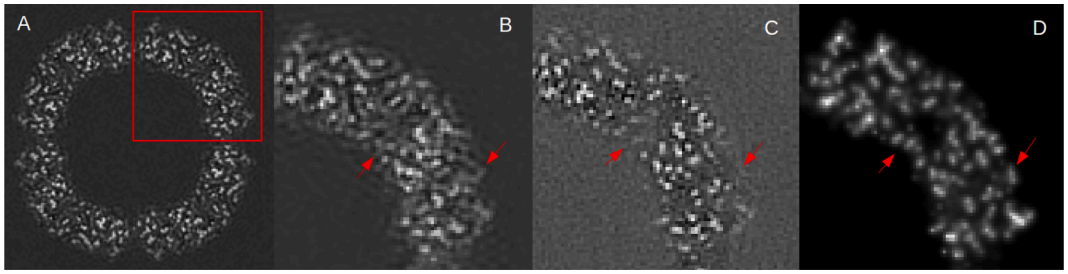

Fig. 12. A) Central slice of consensus result. B) Detail of the squared region in A. C) Same region as in B of the best resolution input apoferritin map (first map of Fig. 10) D) Same region as in B of the atomic model PDB 6WX6 converted into density map. Red arrows points to the regions where significant differences are found between the maps.

similar processing that have not been submitted to any public database. We aligned the five density maps of this structure and computed a tight mask for each one.

We remark here that these five different reconstructions of the spike are all in the same conformation, as this consensus algorithm is not intended for structures in different conformations, as the result will be a mixture of the input conformations which will make up an unreal structure.

We adjusted all of the reconstructions against the first one of the inputs, with default parameters. We arbitrarily took the first one as reference, as the five reconstructions were similar in resolution, but the result does not significantly change depending on the map selected as reference.

Then, we performed the volume consensus with the five adjusted volumes. The result is showed in Fig. 13. The colors in the figure represent the degree of similarity, being dark blue the smallest differences between input volumes and red the biggest ones. As can be appreciated, in the core of the structure there are not significant discrepancies. However, there are green to red regions at the top and bottom of the structure, indicating that these are regions with larger differences between the input reconstructions. We have used here reconstructions of the spike in the "up" conformation, however we know that the flexibility of this spike is high and the "up" conformation is not unique but there are many of them which differs in small details (Melero et al., 2020). Thus, the consensus method in this case is also useful to see in which small parts of the map the reconstructions differ due to details in similar conformations.

To explore those differences, we show the slices that correspond to that part of the structure. In Fig. 14 the same slice for the five input reconstructions and for the consensus are shown. Red arrows points to regions where there are differences between the inputs, but in the consensus appears the best input density in each region, no matter from

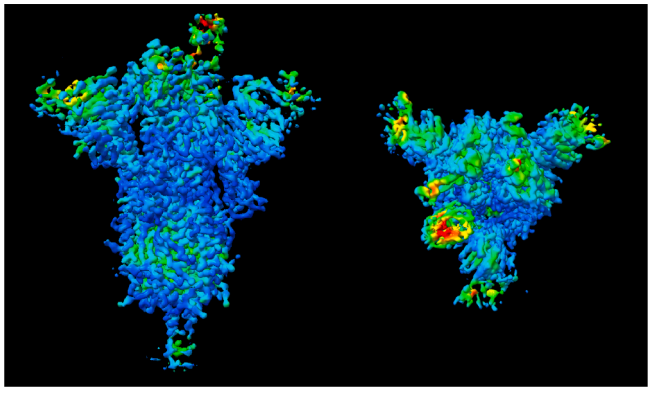

Fig. 13. Result of the consensus volume from five reconstructions of the SARS-CoV-2 Spike, side (left) and top (right) views. The colors represent the degree of similarity of the input map signal, being dark blue the smallest differences between input volumes and red the biggest differences.

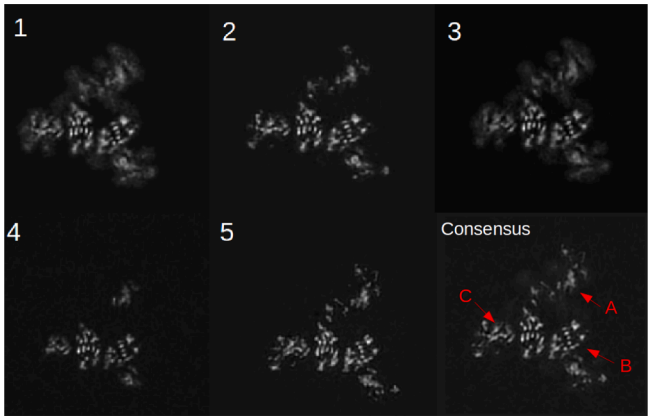

Fig. 14. Same slice of the five input reconstructions of SARS-CoV-2 Spike and the consensus result. Red arrows points to the regions where we can appreciate a fusion of the information provided by the input volumes, obtaining a better density.

what input comes from. In Fig. 15 we show the densities of each of the input volumes and the consensus volume that correspond to the regions pointed by red arrows in Fig. 14. The consensus volume is also displayed as a mesh superimposed to each of the input volumes to see the difference. As can be appreciated, in the three cases the consensus is the volume with more well-defined density in comparison with the rest, as it is build with the best parts of each of its inputs.

4. Conclusions

We have presented in this article a procedure to adjust the numerical values of two density maps. The procedure finds a trade-off between the Fourier amplitudes of one volume and the phases of another. Additionally, it imposes non-negativity, range and locality (mask) constraints to the result. We have also shown three different applications in which this operator is useful (volume subtraction, sharpening and volume consensus).

In the case of subtraction and sharpening, the results show that the proposed algorithm is at the level of the state-of-the-art methods, even improving the results in most cases. In the case of consensus algorithm, we did not find any other method among cryo-EM software packages that performs a similar task. We remark that consensus is not designed to work with structures in different conformations, but to combine different estimations of the same macromolecule on the same conformation. Moreover, in structures with high flexibility, as the case of the SARS-CoV-2 spike presented in Section 3.3.2, the consensus algorithm is not intended to generate as a result and improved map, but is a useful tool to identify the regions of the map where there are more differences among the input maps.

These methods are implemented in the Xmipp package and are user-

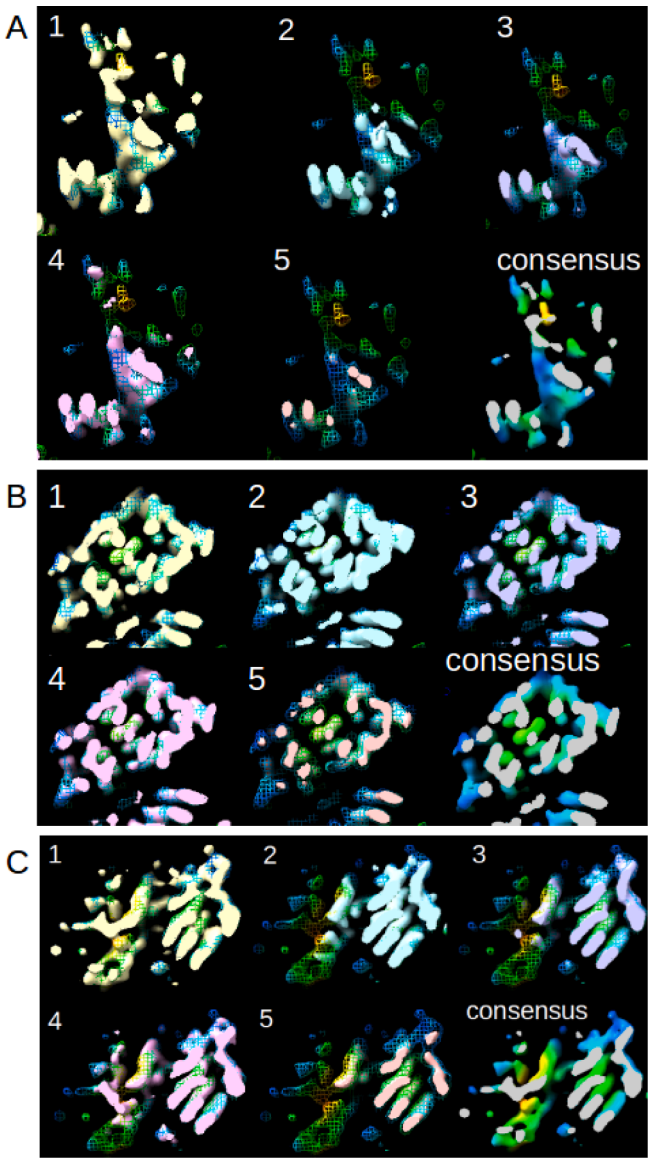

Fig. 15. Detail in 3D of the regions pointed by arrows in Fig. 14. Input volumes are displayed in different colors from 1 to 5. Consensus volume is displayed individually and superimposed to each input volume as a gradient-colored mesh.

friendly accessible through the cryo-EM workflow engine Scipion.

CRediT authorship contribution statement

E. Fernández-Giménez: Methodology, Software, Validation, Formal analysis, Writing - original draft, Visualization. **M. Martínez:** Validation, Resources, Writing - review & editing. **R. Sánchez-García:** Software, Writing - review & editing. **R. Marabini:** Resources, Writing - review & editing. **E. Ramírez-Aportela:** Resources, Writing - review & editing. **P. Conesa:** Software. **J.M. Carazo:** Writing - review & editing, Supervision. **C.O.S. Sorzano:** Conceptualization, Methodology, Formal analysis, Writing - review & editing, Supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The Spanish Ministry of Science and Innovation through Grants: SEV 2017–0712, PID2019-104757RB-I00/ AEI/ 10.13039/501100011033, the "Comunidad Autónoma de Madrid" through Grant: S2017/BMD-3817 European Union (EU) and Horizon 2020 through grants: EOSC Life (Proposal: 824087), HighResCells (ERC – 2018 – SyG, Proposal: 810057) and iNEXT-Discovery (Proposal: 871037). The authors acknowledge the support and the use of resources of Instruct, a Landmark ESFRI project.

References

Bai, X.C., Rajendra Eeson, Yang Guanghui, Shi Yigong, Scheres Sjors H., 2015. Sampling the conformational space of the catalytic subunit of human γ-secretase. Elife, 4.

de la Rosa-Trevín, J.M., Otón, J., Marabini, R., Zaldívar, A., Vargas, J., Carazo, J.M., Sorzano C.O.S., 2013. Xmipp 3.0: an improved software suite for image processing in electron microscopy. Journal of Structural Biology, 184(2), 321–328.

de la Rosa-Trevín, J.M., Quintana, A., Del Cano, L., Zaldívar, A., Foche, I., Gutiérrez, J., Gómez-Blanco, J., Burguet-Castell, J., Cuenca-Alba, J., Abrishami, V., Vargas, J., Otón, J., Sharov, G., Vilas, J.L., Navas, J., Conesa, P., Kazemi, M., Marabini, R., Sorzano, C.O.S., Carazo, J.M., 2016. Scipion: A software framework toward integration, reproducibility and validation in 3D electron microscopy. Journal of Structural Biology 195, 93–99.

Goddard, Thomas D., Huang, Conrad C., Meng, Elaine C., Pettersen, Eric F., Couch, Gregory S., Morris, John H., Ferrin, Thomas E., 2017. Ucsf chimerax: Meeting modern challenges in visualization and analysis. Protein Science 27 (1), 14–25.

Jakobi Arjen J., Wilmanns Matthias, Sachse Carsten, 2017. Model-based local density sharpening of cryo-EM maps. eLife, 6:e27131. ISSN 2050–084X.

Joseph Agnel Praveen, Lagerstedt Ingvar, Jakobi Arjen, Burnley Tom, Patwardhan Ardan, Topf Maya, Winn Martyn, 2020. Comparing cryo-em reconstructions and validating atomic model fit using difference maps. Journal of chemical information and modeling. ISSN 1549-960X.

Madisetti, V.J., Williams, D., 1999. Digital Signal Processing Handbook. CRC Press.

Roberto Melero, Carlos Oscar S. Sorzano, Brent Foster, José-Luis Vilas, Marta Martínez, Roberto Marabini, Erney Ramírez-Aportela, Ruben Sanchez-Garcia, David Herreros, Laura del Caño, Patricia Losana, Yunior C. Fonseca-Reyna, Pablo Conesa, Daniel Wrapp, Pablo Chacon, Jason S. McLellan, Hemant D. Tagare, Jose-Maria Carazo, 2020. Continuous flexibility analysis of SARS-CoV-2 spike prefusion structures. IUCrJ, 7(6):1059–1069. doi:10.1107/S2052252520012725.

Pajares, G., de la Cruz, J.M., 2004. A wavelet-based image fusion tutorial. Patter Recognition 37, 1855–1872.

Marta Pérez-Illana, Marta Martínez, Gabriela N. Condezo, Mercedes Hernando-Pérez, Casandra Mangroo, Martha Brown, Roberto Marabini, Carmen San Martín, Cryo-em structure of enteric adenovirus hadv-f41 highlights structural variations among human adenoviruses. Science Advances, 7(9):eabd9421.

Punjani A., Rubinstein, J.L., Fleet, D.J., Brubaker, M.A., 2017. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. Nature Methods, 14:290–296. ISSN 1548-7105.

Rafie, K., Lenman, A., Fuchs, J., Rajan, A., Arnberg, N., Carlson, L.-A., 2021. The structure of enteric human adenovirus 41–a leading cause of diarrhea in children. Science Advances, 7(2):eabe0974.

Ramírez-Aportela Erney, Luis Vilas Jose, Glukhova Alisa, Melero Roberto, Conesa Pablo, Martínez Marta, Maluenda David, Mota Javier, Jiménez Amaya, Vargas Javier, Marabini Roberto, Sexton Patrick M., Carazo Jose Maria, Sorzano C.O.S., 2020. Automatic local resolution-based sharpening of cryo-EM maps. Bioinformatics, 36: 765–772.

Ramírez-Aportela Erney, Maluenda David, Fonseca Yunior C., Conesa Pablo, Marabini Roberto, Bernard Heymann J., Maria Carazo Jose, Sorzano Carlos Oscar S., 2021. Fsc-q: A cryoem map-to-atomic model quality validation based on the local fourier shell correlation. Nature Communications, 12(1), 1–7.

Sanchez-Garcia, R., Gomez-Blanco, J., Cuervo, A., Carazo, J.M., Sorzano, C.O.S., Vargas, J., 2020. Deepemhancer: a deep learning solution for cryo-em volume post-processing. bioRxiv. URL url:https://www.biorxiv.org/content/early/2020/08/17/2020.06.12.148296.

Scheres, S.H.W., 2012. Relion: implementation of a bayesian approach to cryo-EM structure determination. Journal of Structural Biology 180, 519–530.

Sorzano, C.O.S., Vargas, J., Otón, J., Abrishami, V., de la Rosa Trevín, J.M., del Riego, S., Fernández-Alderete, A., Martínez-Rey, C., Marabini, R., Carazo, J.M., 2015. Fast and accurate conversion of atomic models into electron density maps. AIMS. Biophysics 2, 8–20.

Sorzano, C.O.S., Vargas, J., de la Rosa-Trevin, J.M., Jimenez, A., Maluenda, D., Melero, R., Martinez, M., Ramirez-Aportela, E., Conesa, P., Vilas, J.L., Marabini, R., Carazo, J.M., 2018. A new algorithm for high-resolution reconstruction of single particles by electron microscopy. Journal of Structural Biology, 204:329–337. ISSN 1095-8657.

Terwilliger Thomas C., Sobolev Oleg V., Afonine Pavel V., Adams Paul D., 2018. Automated map sharpening by maximization of detail and connectivity. Acta crystallographica. Section D, Structural biology, 74:545–559. ISSN 2059-7983.

Terwilliger Thomas C., Sobolev Oleg V., Afonine Pavel V., Adams Paul D., Read Randy J., 2020. Density modification of cryo-em maps. Acta Crystallographica. Section D, Structural Biology, 76:912–925. ISSN 2059-7983.

Vilas, J.L., Gómez-Blanco, J., Conesa, P., Melero, R., de la Rosa Trevín, J.M., Otón, J., Cuenca, J., Marabini, R., Carazo, J.M., Vargas, J., Sorzano, C.O.S., 2018. MonoRes: automatic and unbiased estimation of local resolution for electron microscopy maps. Structure 26, 337–344.