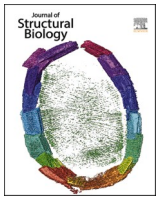
Journal of Structural Biology 215 (2023) 108024

Contents lists available at ScienceDirect

Journal of Structural Biology

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



Research Article

A new algorithm for particle weighted subtraction to decrease signals from unwanted components in single particle analysis

E. Fernández-Giménez [a,b], M.M. Martínez [a], R. Marabini [a,b], D. Strelak [c], R. Sánchez-García [d,e], J.M. Carazo [a], C.O.S. Sorzano [a,*]

[a] Centro Nac. Biotecnología (CSIC), c/Darwin, 3, 28049 Cantoblanco, Madrid, Spain
[b] Univ. Autonoma de Madrid, 28049 Cantoblanco, Madrid, Spain
[c] Institute of Computer Science, Masaryk University, Botanická 68a, 60200 Brno, Czech Republic
[d] Department of Statistics, University of Oxford, 24-29 St Giles', Oxford OX1 3LB, United Kingdom
[e] Astex Pharmaceuticals, 436 Cambridge Science Park, Cambridge CB4 0QA, UK

ARTICLE INFO

Keywords:
Projection subtraction
Nanodisc
Ligand
SPA
Cryo-EM

ABSTRACT

Single particle analysis (SPA) in cryo-electron microscopy (cryo-EM) is highly used to obtain the near-atomic structure of biological macromolecules. The current methods allow users to produce high-resolution maps from many samples. However, there are still challenging cases that require extra processing to obtain high resolution. This is the case when the macromolecule of the sample is composed of different components and we want to focus just on one of them. For example, if the macromolecule is composed of several flexible subunits and we are interested in a specific one, if it is embedded in a viral capsid environment, or if it has additional components to stabilize it, such as nanodiscs. The signal from these components, which in principle we are not interested in, can be removed from the particles using a projection subtraction method. Currently, there are two projection subtraction methods used in practice and both have some limitations. In fact, after evaluating their results, we consider that the problem is still open to new solutions, as they do not fully remove the signal of the components that are not of interest. Our aim is to develop a new and more precise projection subtraction method, improving the performance of state-of-the-art methods. We tested our algorithm with data from public databases and an in–house data set. In this work, we show that the performance of our algorithm improves the results obtained by others, including the localization of small ligands, such as drugs, whose binding location is unknown *a priori*.

1. Introduction

Single particle analysis (SPA) in cryo-electron microscopy (cryo-EM) has emerged as a reliable method for elucidating the atomic structure of macromolecules and biological complexes, owing to its remarkable ability to produce high-resolution electronic density maps (below 3 Å) (Neumann et al., 2018).

Nevertheless, certain challenging samples require additional processing to achieve high resolution. For example, macromolecules that consist of flexible multiple subunits often face alignment difficulties, with the larger subunit dominating the alignment and leading to poorer resolution for smaller subunits. In other cases, macromolecules may be surrounded by additional proteins (such as a specific protein within a virus capsid) or other molecules (such as a nanodisc). The presence of these surrounding molecules or domains complicates image processing, as they typically impact classifications and alignments, resulting in the molecule of interest being under-resolved.

To tackle this problem, the logical approximation is to remove the signal of the components we are not interested in from the processing. We can perform it by subtracting directly in the volume with the methods developed for this purpose in ChimeraX (Pettersen et al., 2021) or Xmipp (Fernandez-Gimenez et al., 2021). However, while volume subtraction is computationally more efficient, it only cleans the view of the final structure of interest, but it has no impact on its resolution.

Alternatively, it is possible to mask or remove the unwanted signal at the particle level. The goal is to obtain a new set of 2D projections containing solely the signal from the 3D components of interest, which can then be used for iteratively classifying, reconstructing, and refining

* Corresponding author.
E-mail address: coss@cnb.csic.es (C.O.S. Sorzano).

https://doi.org/10.1016/j.jsb.2023.108024
Received 4 July 2023; Received in revised form 22 August 2023; Accepted 4 September 2023
Available online 11 September 2023

without the unwanted signal, which may lead to an improvement of the final resolution. Once again, there are several methods available to achieve this task. For instance, Relion (Kimanius et al., 2021) offers focus refinement capabilities through masking and projection subtraction, and CryoSPARC (Punjani et al., 2017) provides a program for projection subtraction.

However, this task is not entirely solved in practice, as many undesired signals remain in the subtracted particles. In this article, we present a new method for projection subtraction, developed within the Xmipp (Strelak et al., 2021; de la Rosa-Trevin et al., 2013; Sorzano et al., 2004) software package, that improves the results of the state-of-the-art methods, as we show in the Results section. This new algorithm is also available in the Scipion (Jimenez-Moreno et al., 2021; de la Rosa-Trevin et al., 2016) framework for cryo-EM image processing.

2. Methods

The basic idea behind a projection subtraction algorithm is to take every input particle, compute the projection of the reference volume corresponding to the pose of the input particle, and subtract it from the original image, normally operating in real space. Usually, we do not want to subtract the projection of the whole particle, but only of a specific region. This region will be determined by a volume mask m, either defining the region to keep (as required by Relion) or the region to subtract (as required by CryoSPARC). See an example of the masks in Fig. 1. In the case of our algorithm, the user can choose if the input mask defines the region to keep or subtract. The complete projection subtraction workflow of our algorithm is reflected in Fig. 2:

However, the real complexity of the algorithm resides in preparing the projection before subtracting, as a projection of a reference volume and a real particle have different characteristics, and they usually have different ranges of values. Thus, instead of subtracting the projection, we will modify it previously to have an adjusted version suitable for subtraction:

$$i_s(\mathbf{s}) = i(\mathbf{s}) - p'(\mathbf{s}) \tag{1}$$

being $(\mathbf{s})$ the spatial 2D coordinate, i_s the subtracted particle, i the input particle, and p' the adjusted projection. Let $p(\mathbf{s})$ be the projection of the map to be subtracted along the same direction and the same in-plane shift as the image i.

In our algorithm, we perform this adjustment in Fourier space, previous to the subtraction in real space. Moreover, we will adjust each projection to its particle individually. For calculating the adjusted projection, let us introduce some notation. Let $I(\boldsymbol{\omega})$ and $P(\boldsymbol{\omega})$ be the Fourier transform of the i and p images, respectively. Note that these Fourier transforms are complex-valued vectors. Let us consider a set of frequencies on which we will perform the subtraction, Ω. We define the $\mathrm{vec}_\Omega\{\cdot\}$ operator that acts on the Fourier transform of an image as

$$\mathrm{vec}_\Omega\left\{I\right\} = \begin{pmatrix} \mathrm{Re}\{I(\boldsymbol{\omega}_0)\} \\ \mathrm{Im}\{I(\boldsymbol{\omega}_0)\} \\ \mathrm{Re}\{I(\boldsymbol{\omega}_1)\} \\ \mathrm{Im}\{I(\boldsymbol{\omega}_1)\} \\ \cdots \end{pmatrix} \tag{2}$$

where Re and Im extract the real and imaginary parts of a complex number, and $\boldsymbol{\omega}_i$ goes over all frequencies in the set Ω. This vector is called $\mathbf{I}$, and its k-th component is I_k. We also define the corresponding frequency vector as

$$\mathbf{w} = \begin{pmatrix} |\boldsymbol{\omega}_0| \\ |\boldsymbol{\omega}_0| \\ |\boldsymbol{\omega}_1| \\ |\boldsymbol{\omega}_1| \\ \cdots \end{pmatrix} \tag{3}$$

and we refer to its k-th component as w_k. We also define a diagonal matrix W whose main diagonal is this vector.

Let us define H as a diagonal matrix representing the microscope Contrast Transfer Function (CTF). We now consider two scaling transformations, T_0 and T_1 that will minimize the Euclidean distance between $\mathbf{I}$ and $H\mathbf{P}$

$$\epsilon^2 = \left\|\mathbf{I} - T\left(H\mathbf{P}, \mathbf{w}\right)\right\|^2 \tag{4}$$

where T is any of the transformations (T_0 or T_1). The two transformations are given by a linear model of order zero and order one respectively

$$\begin{aligned} T_0(H\mathbf{P}, \mathbf{w}) &= \beta_{00} H\mathbf{P} \\ T_1(H\mathbf{P}, \mathbf{w}) &= (\beta_{01} + \beta_{11} W) H\mathbf{P} \end{aligned} \tag{5}$$

where β_{ij} are constants we must find to minimize the above Euclidean distance. Note that the first transformation, T_0, is a grayscale adjustment, and T_1 is a grayscale adjustment and a projection sharpening or dampening to adjust the map projection to subtract from the experimental image. We compute both models even though T_0 is contained in T_1, because if T_0 fits well enough it is not worth adding a new variable in order to avoid overfitting, as happens generally in linear regression.

The β coefficients are found by standard linear regression:

$$\begin{aligned} \beta_{00} &= \left(\left\|H\mathbf{P}\right\|^2\right)^{-1} \langle H\mathbf{P}, \mathbf{I}\rangle \\ \begin{pmatrix} \beta_{01} \\ \beta_{11} \end{pmatrix} &= \begin{pmatrix} \|H\mathbf{P}\|^2 & \|H\mathbf{P}\|_W^2 \\ \|H\mathbf{P}\|_W^2 & \|H\mathbf{P}\|_{W^2}^2 \end{pmatrix}^{-1} \begin{pmatrix} \langle H\mathbf{P}, \mathbf{I}\rangle \\ \langle H\mathbf{P}, \mathbf{I}\rangle_W \end{pmatrix} \end{aligned} \tag{6}$$

In the expressions above we have made use of the weighted inner product with definition $\langle \mathbf{x}, \mathbf{y}\rangle_A = \mathbf{x}^T A \mathbf{y}$, and its associated norm $\|\mathbf{x}\|_A^2 = \langle \mathbf{x}, \mathbf{x}\rangle_A$.

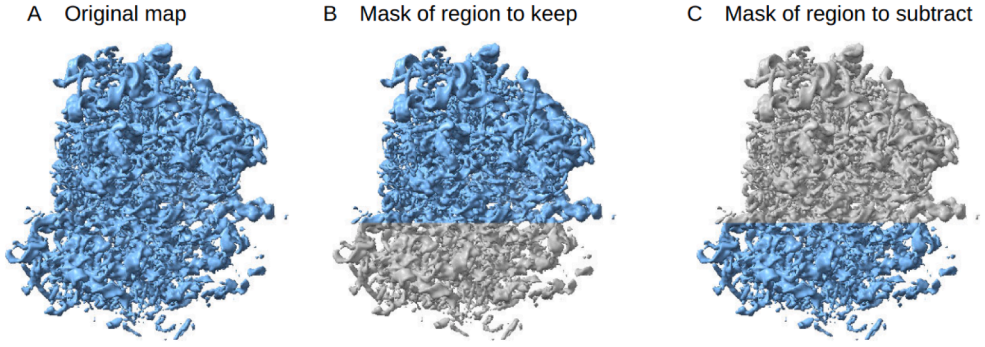


Fig. 1. Example of masks defining the region to keep or the region to subtract. The original map (A) shows a complete ribosome, from which we want to subtract the large subunit in order to keep the small one, thus, the mask defining the part to keep (required by Relion) is colored in grey in (B), and the mask defining the part to subtract (required by CryoSPARC) is colored in grey in (C). Note that the mask to keep is complementary to the mask to subtract.

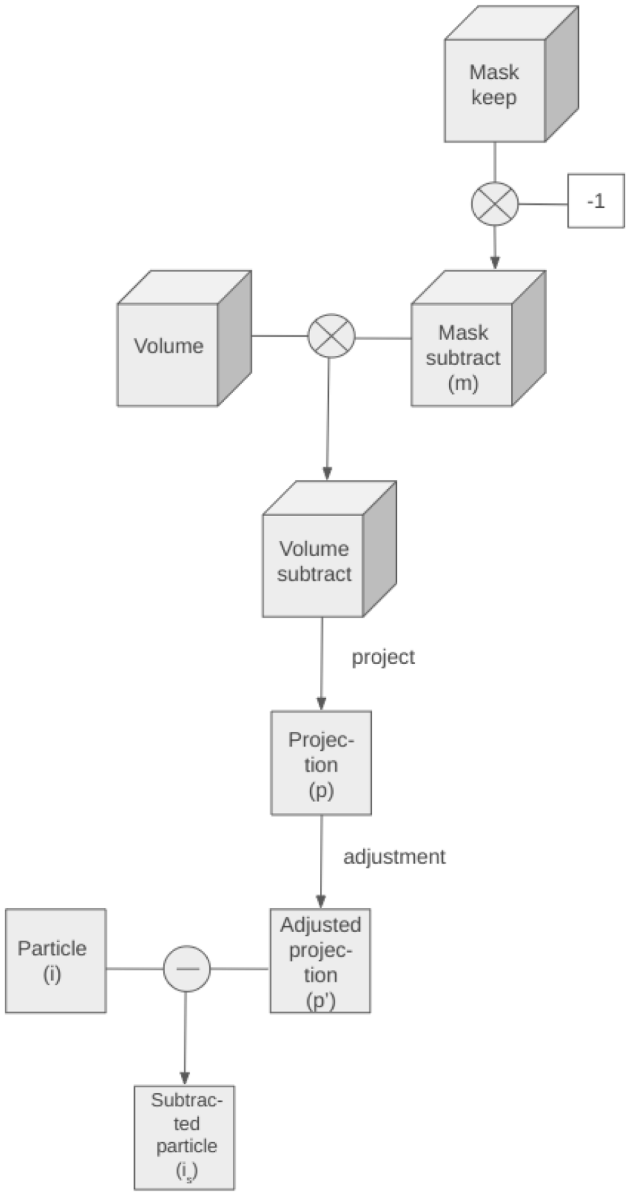


Fig. 2. Subtraction process schema. The input volume is masked with a mask that defines the region to subtract (m). However, the user can input a mask of the region to keep and its inverse will be automatically computed in order to obtain m. The subtraction volume is then projected generating p, which will be adjusted (p') in order to be subtracted to the input particle i, obtaining as a result the subtracted particle i_s.

For any experimental image $\mathbf{I}$, we choose the transformation that maximizes the determination coefficient

$$T\left(\mathbf{P},\mathbf{w}\right) = \begin{cases} T_0(\mathbf{P},\mathbf{w}) & R^2_{0,adj} > R^2_{1,adj} \\ T_1(\mathbf{P},\mathbf{w}) & R^2_{0,adj} < R^2_{1,adj} \end{cases} \tag{7}$$

where the adjusted coefficient of determination (R^2) of the i-th model is calculated as

$$\begin{aligned} R^2_{i,adj} &= 1 - \left(1 - R^2_i\right)\frac{N-1}{N-k-1} \\ R^2_i &= 1 - \frac{\|\mathbf{I} - T_i(\mathbf{P},\mathbf{w})\|^2}{\|\mathbf{I} - \bar{I}\|^2} \end{aligned} \tag{8}$$

where N is the dimension of $\mathbf{I}$, k is the degree of the polynomial in the T_i transformation, and $\bar{I}$ the average value of the vector $\mathbf{I}$.

The adjusted determination coefficient ($R^2_{i,adj}$) is reported for each particle and it is also used to rank them according to their quality, so the user can discard the worst particles.

The subtracted image is finally

$$p' = FT^{-1}\{T(HP)\}m_c \tag{9}$$

being FT^{-1} the inverse Fourier Transform and m_c a circular mask in real space whose diameter equals the particle box size. This mask is applied in order to avoid edge and corner artifacts due to the adjustment process.

Note that T_0 is similar to Relion adjustment, as Relion also multiples each particle by a constant, which is referred to as a"scale factor" in their metadata. However, Relion estimates that constant during the refinement step, which then has the same value for all the particles in the same micrograph. In our case, we estimate a different constant for each particle, which may increase our precision. We cannot evaluate the similarities or differences of our algorithm in comparison to the one of CryoSPARC as their subtraction algorithm is not published and its code is not open source. Another important difference is that we can manage masks that define the part to keep or the part to subtract, which is a useful feature depending on the application of subtraction and/or the difficulty of creating the mask. However, Relion only manages input masks of the region to keep, and CryoSPARC only manages masks of the region to subtract, which is a drawback depending on the application of the subtraction and the availability of masks, as will show in the Results section.

3. Results and discussion

To validate our algorithm, we have compared its performance with the ones obtained with the state-of-the-art projection subtraction methods: Relion and CryoSPARC. We have chosen three different scenarios: 1) focused refinement, 2) subtraction of unwanted signals, and 3) ligand discovery. As a general result of all the experiments, it has turned out that the majority of the projections have been modified just by a grayscale adjustment (T_0).

3.1. Focused refinement

3.1.1. Ribosomal small subunit

We have chosen the data set from the EMPIAR entry 10028, the *Plasmodium falciparum* 80S ribosome, with the aim of improving the resolution of the small subunit (whose molecular mass is about 1.2MDa).

As the data set already contains the extracted particles, a first reconstruction and refinement of the ribosome has been performed using CryoSPARC non-uniform refinement (Punjani et al., 2017), Fig. 3 (A). As can be seen, the quality of the map in the region of the small subunit (bottom) shows more unconnected densities than in the region of the big subunit (top). This was expected due to the fact that alignment was driven by the big subunit, as it has a larger signal than the small subunit.

In Fig. 3, the results of focal refinement by Relion (B) and subtract projections by CryoSPARC (C), Xmipp (E) and Relion (F) are shown. Fig. 3 (D) shows the mask for the region that we want to keep in subtractions and focus on focal refinement (the small subunit). As can be appreciated qualitatively in regions highlighted by red squares, Relion focal refinement and CryoSPARC are able to slightly improve the original reconstruction. However, Xmipp and Relion subtraction are able to improve it more. Additionally, CryoSPARC does not remove completely the signal from the large subunit. In this case, the results of Relion and Xmipp are comparable.

3.1.2. Crowded/Viral environment: Monomer B of hexon 1 of human Adenovirus

To compare Xmipp and Relion projection subtraction results in more

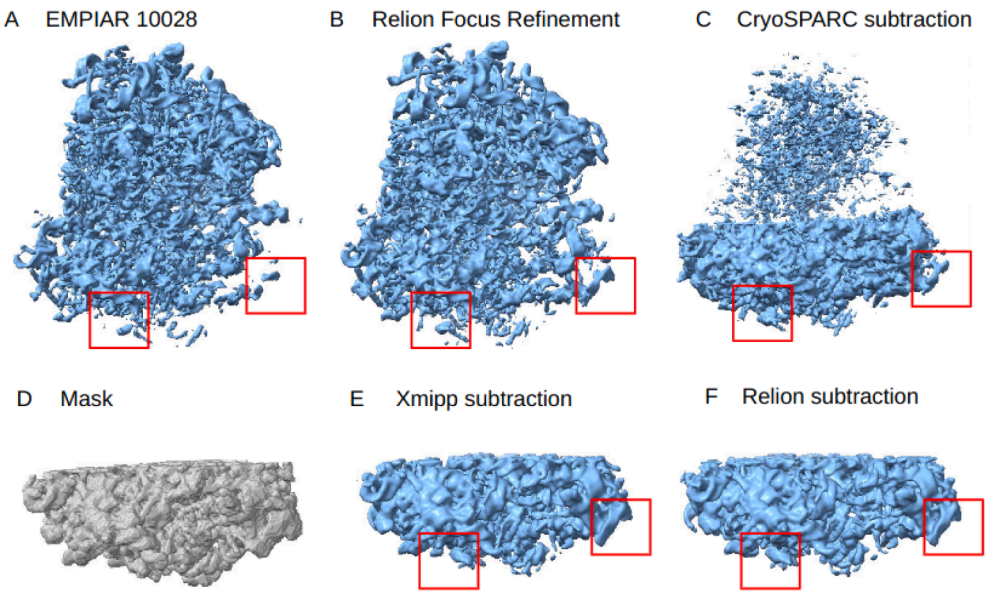


Fig. 3. (A) Refined map from particles in EMPIAR 10028. (B) Focal refinement of small subunit by using Relion focus refinement program. (C) Refined map from subtracted particles by CryoSPARC. (D) Mask of small subunit used in subtractions (region to keep) and focal refinement. (E) Refined map from subtracted particles by Xmipp. (F) Refined map from subtracted particles by Relion. Parts with remarkable differences among the maps are squared in red.

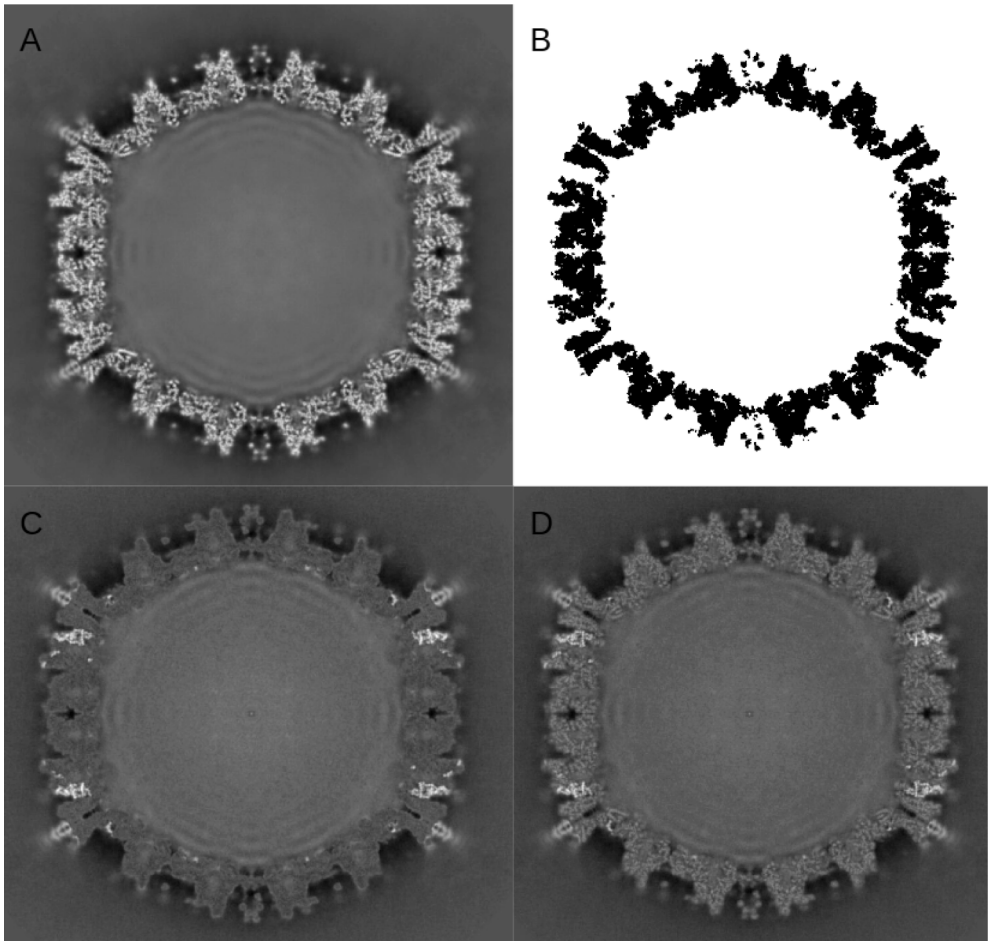


Fig. 4. Human Adenovirus (central slice) (A) Reconstruction without subtraction (B) Mask of the capsid (in black is the region to subtract) (C) Map reconstruction of subtracted particles by Xmipp. (D) Map reconstruction of subtracted particles by Relion.

detail, we have chosen a smaller target in a crowded environment: the monomer B of the hexon 1 of the human Adenovirus (molecular mass 108KDa) from an in–house data set of the Adenovirus capsid. Central slice of the complete Adenovirus previous to subtraction is shown in Fig. 4 (A). In this case, we wanted to subtract the whole capsid except for the monomer B of hexon 1, therefore a mask of the region that we wanted to keep (thus, the capsid is in black) is built and its central slice is shown in Fig. 4 (B). Central slices of the volumes reconstructed after the subtractions with Xmipp and Relion are shown in Figs. (C) and (D) respectively.

A map of monomer B is shown in Fig. 5 (A). It has been obtained by converting PDB entry 6b1t (Dai et al., 2017) into a density map at 3Åresolution. A reconstruction of monomer B with Xmipp subtracted particles is shown in Fig. 5 (B), and the result of refining it is shown in (D). Analogously, the reconstruction with Relion subtracted particles is shown in Fig. 5 (C), and the result of its refinement is in (E). Both refinements have been done with Relion auto-refine (Kimanius et al., 2021) with the same parameters. As can be seen, on the right side of (C) and (E), there are considerable amounts of signals that do not correspond to monomer B, but they have not been removed by Relion subtraction. We can appreciate the undesired remaining signal also in the central slices, in the second row of Fig. 5. The region pointed out with arrows from (B) to (E) correspond to the adjacent monomer B, as monomer B is arranged in a pentameric form around the penton. Thus, it remains in all the cases because it was included in the mask used to define the region to keep (monomer B) in the projections subtraction. The result produced by CryoSPARC subtraction is shown in the Supplementary material Fig. S-1. We consider that it is substantially worse than the previously presented results, as it leaves more signal on the right side and loses part of the signal in the monomer region.

3.2. Subtraction of unwanted signals

As stated in the introduction, having nanodiscs in the sample is very useful for sample preparation of membrane proteins. Still, it is inconvenient when doing image processing as they may drive image alignment. Thus, subtraction of the signal generated by the nanodisc in the particles can improve the final result. To check our ability to handle this case we have used the particles from EMPIAR entry 10005, which sample is a capsaicin receptor that has been embedded in nanodiscs during the sample preparation process. We have generated a mask of the region to keep from related EMDB entry 5778, which is the capsaicin receptor without the nanodisc.

In Fig. 6 (top), we compare two different slices of the map after particle refinement as it is in the entry of EMPIAR 10005 with equivalent two slices of the map obtained after refining the same data set (with the same refinement parameters) once Xmipp has subtracted them.

We can appreciate that the nanodisc is present in the first two cases, while in the subtraction it has been removed (see red arrows in the figure). Even though the reported resolution by FSC does not improve, the local resolution improves as shown in Fig. 6 (bottom), especially in the part where the nanodisc was (bottom of the structure), as the signal from the structure is now stronger in that part. There are also important improvements at the top of the structure in the Figure, as it seems the alignment without subtraction was perturbed by the nanodisc than by the structure itself.

The result obtained by Relion subtraction (see Fig. S-2 in Supplementary materials) with the same particles and the same refinement parameters produces a structure highly degraded. For CryoSPARC subtraction, a mask of the region to subtract is needed, which in this case is the nanodisc. It is difficult to obtain a thigh mask of the nanodisc as there is no direct way to get it, however, we compute an approximate mask of the nanodisc by subtracting the volume in EMD-5778 from the reconstructed volume from the particles in EMPIAR 10005. Even though the mask represents approximately the nanodisc, the result of CryoSPARC subtraction also produces a structure highly degraded (it is shown in Fig. S-3 of Supplementary materials).

3.3. Ligand discovery

3.3.1. Simulated data

To further evaluate the precision of our subtraction method, we decided to test its performance in detecting a small molecule in the context of a much larger one. In this case, we are interested in removing the signal of the small subunit of the ribosome, keeping just a ligand that is bound to it.

In this experiment, we have used the same data set as in Section 3.1.1, as the small subunit of the 80S ribosome of the data set has a drug bounded (emetine). We have used the related PDB entry 3j7a to generate a map at a 2Åresolution without the emetine ligand by removing it from the PDB in ChimeraX and converting it to a density map. To check the performance of the subtraction, we have to avoid other sources of errors, such as alignment errors. Thus, we have not used the original particles of the data set, but we have created a set of projections from the volume in

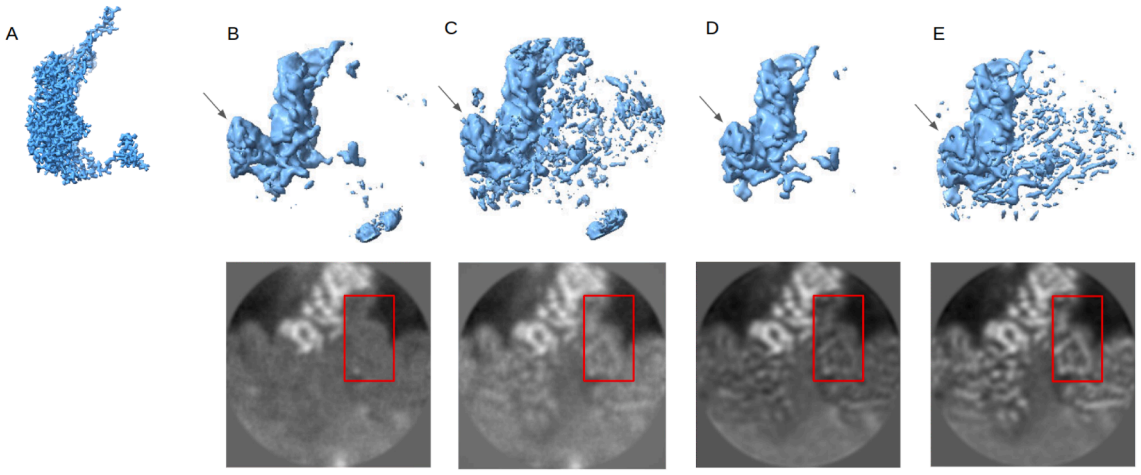


Fig. 5. Top: Monomer B of hexon 1 of the human Adenovirus. Below: a central slice of each map. (A) Converted 3Åmap from PDB entry 6b1t (Dai et al., 2017). (B) Map reconstruction of subtracted particles by Xmipp. (C) Map reconstruction of subtracted particles by Relion. (D) Map refinement of subtracted particles by Xmipp, the reported resolution by FSC is 4.0Å. (E) Map refinement of subtracted particles by Relion, the reported resolution by FSC is 4.1Å. An area showing different signal levels is marked by a red rectangle (ideally, they should have been subtracted). Arrows point to a zone that does not belong to monomer B but was included in the mask used for the subtractio.n in both cases.

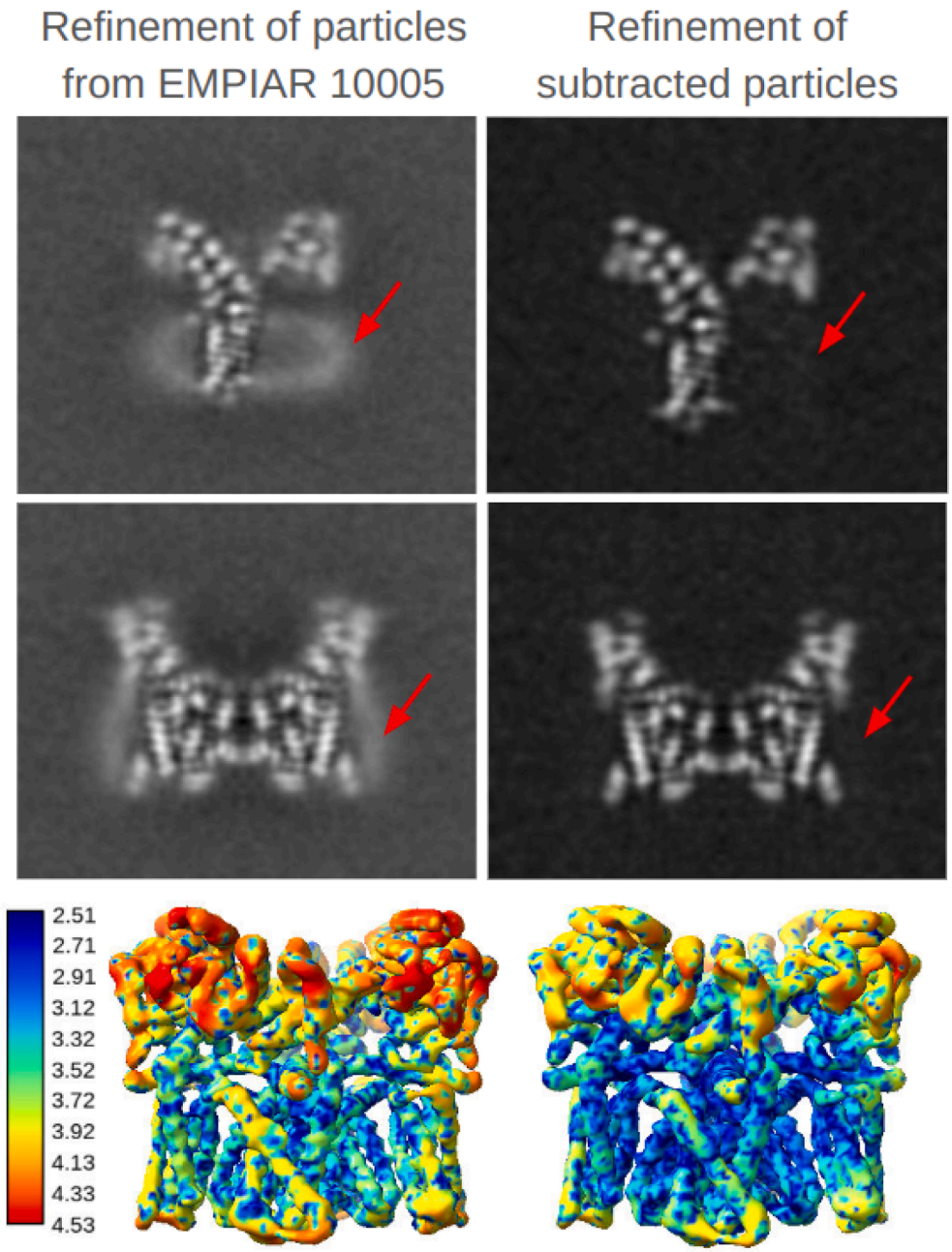


Fig. 6. (Top) The image displays two different slices from the volumes obtained through the refinement of particles in EMPIAR 10005. The left side represents the particles before undergoing subtraction by Xmipp, while the right side shows the particles after subtraction. The same refinement parameters have been applied. The red arrows highlight the signal produced by the nanodisc, which has been eliminated in the subtracted case. (Bottom) Local resolution measure with MonoRes (Vilas et al., 2018) of the refined volumes for (left) particles in EMPIAR 10005 and (right) the same particles once Xmipp has .subtracted them.

EMDB-related entry 2660 to use them as input particles. We have added to these projections a simulated CTF and Gaussian noise with a standard deviation of 50 (see Fig. S-4 in Supplementary material).

The result of reconstructing the subtracted particles with Xmipp is shown in ChimeraX in Fig. 7. The map has been filtered with a Gaussian filter with a standard deviation of one to see the density corresponding to emetine better. We have performed the same experiment with CryoSPARC subtraction, and it has not been able to remove completely the density of the small subunit of the ribosome (see the result in Supplementary material Fig. S-5). We cannot perform this experiment with Relion as it needs as input a mask of the region to keep instead of the region to subtract, requiring knowing ahead the location of the ligand.

3.3.2. *Experimental data*

In this section, we have used signal subtraction for ligand discovery, as in Section 3.3.1, but in this case, with experimental data. The data consists of the human enzyme pyruvate kinase M2 with two pairs of ligands, a sugar (1,6-di-O-phosphono-beta-D-fructofuranose, which is the one we are interested in) and an amino-acid (threonine). We have used our subtraction approach, starting from the movies available in EMPIAR 10647. As a reference volume, we have used the related PDB entry 6tth, removing the ligands in ChimeraX and converting the resulting structure (only the enzyme) into a density map at 2Åresolution by using Xmipp (Sorzano et al., 2015). Then, we created a binary tight mask for this map to be used as the mask of the region to subtract because the aim is to get just the signal from the ligands. The result is shown in Fig. 8 (volume displayed in ChimeraX with PDB as reference) and a slice of the volume in Fig. 9 (A). We have called this section "ligand discovery" as this method can be used to find a ligand that is bound to our sample, but still, we do not know where it is. Thus, as in Section 3.3.1, we cannot perform this experiment with Relion because it needs as input the mask of the region to keep. CryoSPARC subtraction cannot recover the signal from the ligand in this case (see Fig. S-6 in Supplementary Materials).

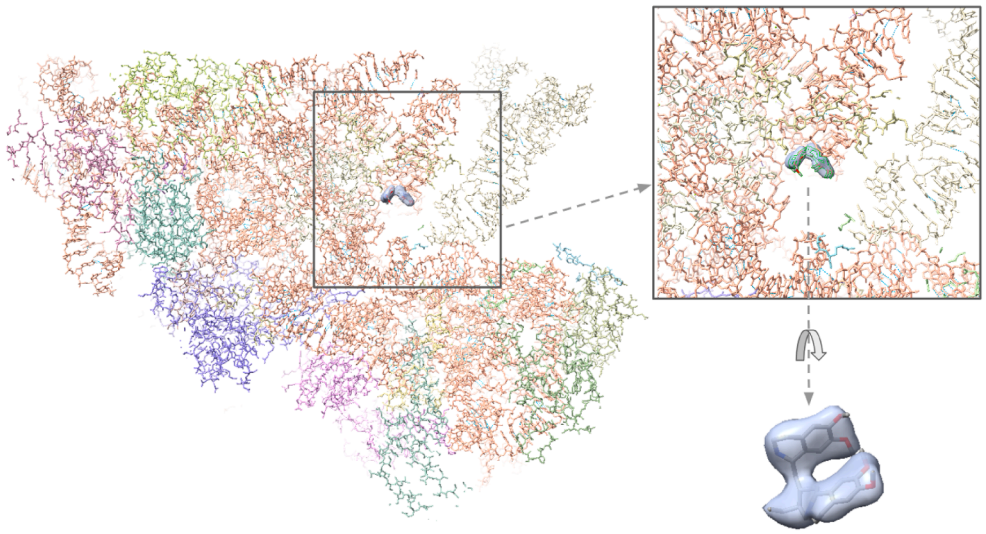


Fig. 7. Reconstructed map of subtracted particles with Xmipp (gray) over PDB PDB3j7a (small subunit of 80S ribosome bound to drug emetine). A zoomed region and detail of the region with density, where emetine fits (rotated to see the fitting better). A Gaussian filter with a standard deviation of one has been used to improve the visualization of the density of the emetine.

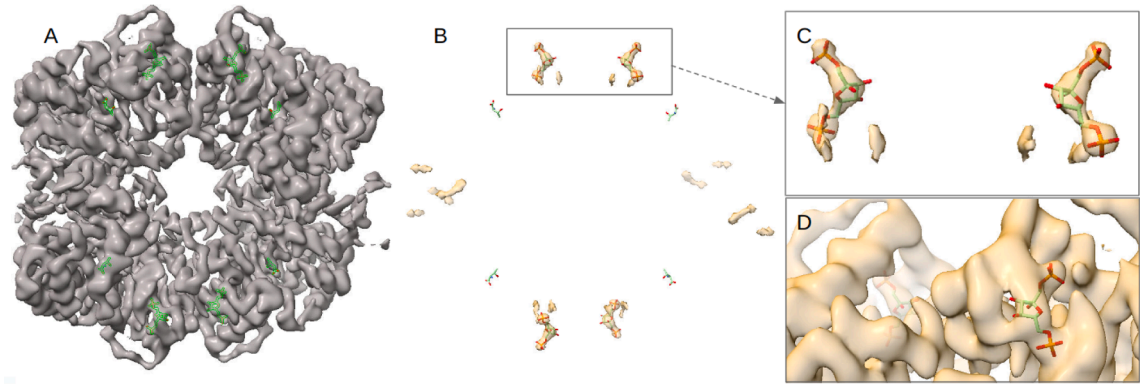


Fig. 8. (A) reconstruction of particles obtained by processing EMPIAR 10647 and (B) reconstruction of same particles subtracted by Xmipp to keep the ligand (1,6-di-O-phosphono-beta-D-fructofuranose), which PDB (6tth) is fitted and (C) zoom of the region of interest. (D) the same region from the original map (before subtraction). The subtraction map does not have any filter or further post-processing.

While volume subtraction is the preferred option because is more computationally efficient, particle subtraction is required if further image processing (such as 3D classification or refinement), is going to be needed. In Fig. 9 we compare the result of Xmipp projection subtraction (A) which implies an adjustment previous to subtraction, Xmipp volume subtraction (Fernandez-Gimenez et al., 2021)(C) which uses another kind of volume adjustment and remove negative values, (B) volume subtraction without any adjustment and, (D), volume subtraction adjusting just mean and standard deviation (D). In (E) we show the result of multiplying (A) and (C) in order to make a consensus of both results and in (F) we show the differences between (A) and (C). As can be seen in (E), the signal coming from the ligand (pointed by red arrows) gets reinforced and in (F) it disappears, showing the agreement between the results.

4. Conclusions

Complementing the now classical pipelines for cryo-EM SPA, special image processing methods are needed to achieve high-resolution information from challenging samples. This is the case of macro-molecules with flexible subunits, proteins embedded in other complex macro-molecules, such as viral capsids, and samples with external components needed for stabilization (nanodiscs, antibodies, etc.). In all these cases, performing a subtraction in the particles of the signal that is not of interest can improve the resolution of the region of interest during the reconstruction and refining process.

Currently, there are methods to perform projection subtraction in Relion and CryoSPARC. However, they do not fully solve the subtraction problem (see Supplementary materials). In this work, we have developed a new method for projection subtraction in Xmipp that improves the performance of subtraction as applied to data from several examples of challenging specimens taken from public databases (EMPIAR, EMDB, and PDB). These improvements come from the fact our approach is able to compute an adjustment for each projection to each input particle individually. Even though we propose two transformations (just grayscale adjustment or grayscale adjustment plus sharpening or dampening of frequencies) to adjust each projection, it has turned out that in most cases, a sharpening or dampening of frequencies is not justified according to our algorithm criteria. Moreover, unlike others, our method can work with the input mask of the region to keep or subtract, providing greater flexibility.

Besides, we have shown that our subtraction approach is very appropriate for the task of finding a small ligand bounded to a large macro-molecule without any assumption on the location of the binding

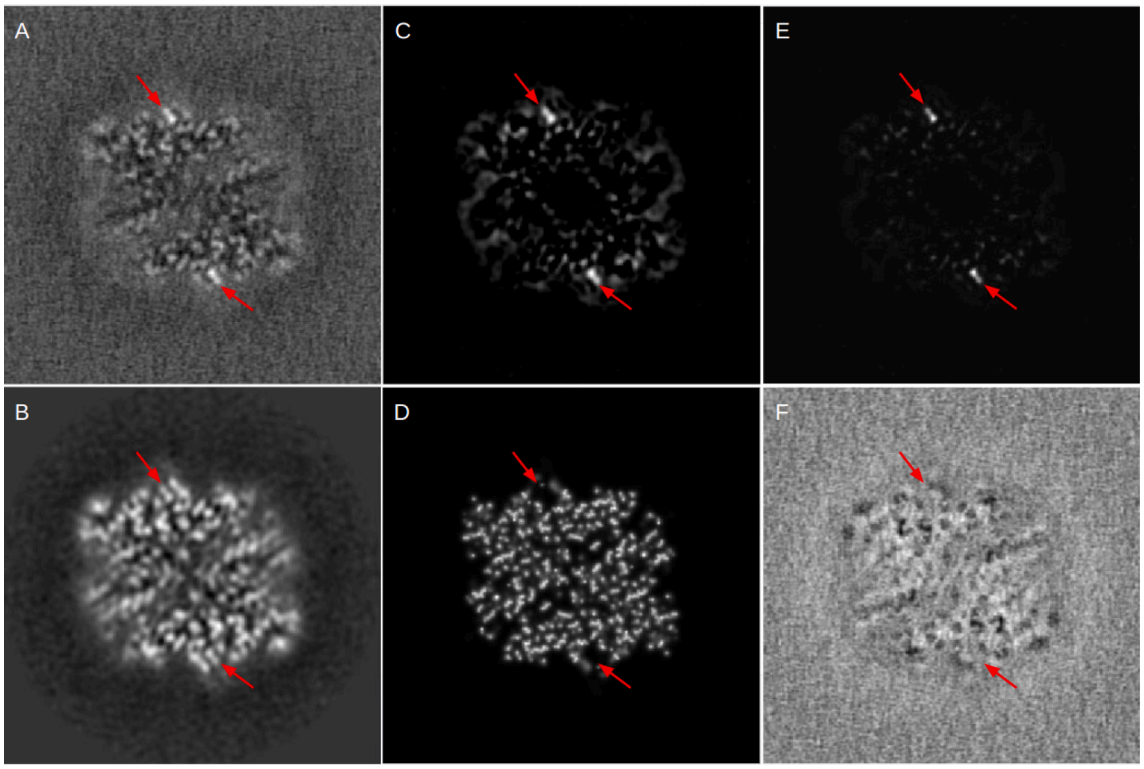


Fig. 9. Equivalent slice of the map coming from particles obtained by processing EMPIAR 10647 so as to keep the ligand (1,6-di-O-phosphono-beta-D-fructofuranose) by (A) Xmipp projection subtraction, (B) volume subtraction without adjustment, (C) Xmipp volume subtraction (Fernandez-Gimenez et al., 2021), (D) volume subtraction with just mean and standard deviation adjustment, (E) multiplication of (A) and (C) as a consensus solution which reinforce the signal from the ligand, (F) difference between (A) and (C). Red arrows point to the regio.n of the ligand.

site.

The algorithm is publicly available at https://github.com/I2PC/xmipp/blob/devel/src/xmipp/libraries/reconstruction/subtract_projection.cpp and can be used through Scipion Framework under the protocol 'subtract projection'.

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.

Data availability

No data was used for the research described in the article.

Acknowledgements

The authors acknowledge the economic support from MICIN to the Instruct Image Processing Center (I2PC) as part of the Spanish participation in Instruct-ERIC, the European Strategic Infrastructure Project (ESFRI) in Structural Biology. Grant PID2019-104757RB-I00 is funded by MCIN/AEI/ 10.13039/ 501100011033 and "ERDF A way of making Europe", by the European Union. The "Comunidad Autónoma de Madrid" through Grant S2022/BMD-7232, the European Union (EU) and Horizon 2020 through grant HighResCells (ERC-2018-SyG, Proposal: 810057). The authors also acknowledge grant PID2019-104098 GB-I00/AEI/10.13039/501100011033, cofunded by the Spanish State Research Agency and the European Regional Development and grant 2023AEP082 by Agencia Estatal CSIC. Ruben Sanchez-Garcia is funded by an Astex Pharmaceuticals Sustaining Innovation Post-Doctoral Award.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.jsb.2023.108024.

References

Dai, X., et al., 2017. Atomic structures of minor proteins VI and VII in human adenovirus. J. Virol. 91 (24), e00850–17.
de la Rosa-Trevin, J.M., et al., 2016. Scipion: A software framework toward integration, reproducibility and validation in 3D electron microscopy. J. Struct. Biol. 195, 93–99.
de la Rosa-Trevin, J.M. et al. Xmipp 3.0: An improved software suite for image processing in electron microscopy, in: Journal of Structural Biology 184.2 (2013), pp. 321–328. doi: 10.1016/j.jsb.2013.09.015.
Fernandez-Gimenez, E., et al., 2021. Cryo-EM density maps adjustment for subtraction, consensus and sharpening. J. Struct. Biol. 213 (4), 107780.
Jimenez-Moreno, A., et al., 2021. Cryo-EM and Single-Particle Analysis with Scipion. J. Visual. Exp.: JoVE.
Kimanius, D., et al., 2021. New tools for automated cryo-EM single-particle analysis in RELION-4.0. Biochem. J. 478 (24), 4169–4185.
Neumann, P., Dickmanns, A., Ficner, R., 2018. Validating resolution revolution. Structure 785–795.
Pettersen, E.f. et al., 2021 UCSF ChimeraX: Structure visualization for researchers, educators, and developers. In: Protein Sci. 30 (2021), pp. 70–82.
Punjani, A., et al., 2017. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. Nat. Methods 14, 290–296.
Sorzano, C.O.S., et al., 2015. Fast and accurate conversion of atomic models into electron density maps. AIMS Biophys. 2, 8–20.
Sorzano, C.O.S., et al., 2004. XMIPP: A new generation of an open-source image processing package for Electron Microscopy. J. Struct. Biol. 148, 194–204.
Strelak, D., et al., 2021. Advances in Xmipp for Cryo-Electron Microscopy: From Xmipp to Scipion. Molecules 26, 20.
Vilas, J.L., et al., 2018. MonoRes: automatic and unbiased estimation of Local Resolution for electron microscopy Maps. Structure 26, 337–344.