PAPER

Image processing tools for the validation of CryoEM maps†

C. O. S. Sorzano, [iD] *[a] J. L. Vilas,[a] E. Ramírez-Aportela,[a] J. Krieger, [iD] [a]
D. del Hoyo,[a] D. Herreros,[a] E. Fernandez-Giménez,[a] D. Marchán, [iD] [a]
J. R. Macías, [iD] [a] I. Sánchez, [iD] [a] L. del Caño,[a] Y. Fonseca-Reyna,[a]
P. Conesa,[a] A. García-Mena,[a] J. Burguet, [iD] [b] J. García Condado, [iD] [c]
J. Méndez García,[d] M. Martínez, [iD] [a] A. Muñoz-Barrutia,[e] R. Marabini,[f]
J. Vargas[b] and J. M. Carazo[a]

The number of maps deposited in public databases (Electron Microscopy Data Bank, EMDB) determined by cryo-electron microscopy has quickly grown in recent years. With this rapid growth, it is critical to guarantee their quality. So far, map validation has primarily focused on the agreement between maps and models. From the image processing perspective, the validation has been mostly restricted to using two half-maps and the measurement of their internal consistency. In this article, we suggest that map validation can be taken much further from the point of view of image processing if 2D classes, particles, angles, coordinates, defoci, and micrographs are also provided. We present a progressive validation scheme that qualifies a result validation status from 0 to 5 and offers three optional qualifiers (A, W, and O) that can be added. The simplest validation state is 0, while the most complete would be 5AWO. This scheme has been implemented in a website https://biocomp.cnb.csic.es/EMValidationService/ to which reconstructed maps and their ESI can be uploaded.

Cryo-electron microscopy is currently one of the most active techniques in structural biology. The number of maps deposited at the Electron Microscopy Data Bank is rapidly growing every year[1,2] and keeping the quality of the submitted maps is essential to maintain the scientific quality of the field. Additionally, all

[a]*Natl. Center of Biotechnology, CSIC, c/Darwin, 3, 28049, Madrid, Spain. E-mail: coss@cnb.csic.es; Fax: +34 91 585 4506; Tel: +34 91 585 45 10*

[b]*Depto. de Óptica, Univ. Complutense de Madrid, Pl. Ciencias, 1, 28040, Madrid, Spain*

[c]*Biocruces Bizkaia Instituto Investigación Sanitaria, Cruces Plaza, 48903, Barakaldo, Bizkaia, Spain*

[d]*Univ. Tecnológica de Pereira, 660003, Pereira, Colombia*

[e]*Univ. Carlos III de Madrid, Avda. de la Universidad 30, 28911Leganés, Madrid, Spain*

[f]*Escuela Politécnica Superior, Univ. Autónoma de Madrid, CSIC, C. Francisco Tomás y Valiente, 11, 28049, Madrid, Spain*

† Electronic supplementary information (ESI) available. See https://doi.org/10.1039/d2fd00059h

scientific domains have recognized open science as a way to accelerate research.[3] In this way, disclosing sufficient information in order to understand the limitations and strengths of a CryoEM map is crucial for a better use of the map. Ideally, reproducibility of the results should be achieved, and the possibility of depositing the raw data at the EMPIAR (Electron Microscopy Public Image Archive, Iudin *et al.*[4]) has certainly been a huge step forward. However, as a community we are still far from having generally adopted this reproducibility goal. For instance, from the 950 EMDB entries labelled as SARS-CoV-2, only 23 (less than 2.5%) are deposited at EMPIAR (as of March 7th, 2022). The availability of the raw data could be complemented with the availability of the image processing workflow and decisions taken to go from the raw acquisition to the final map. Despite the fact that scientific articles are generally trusted by other scientists and the general public, severe concerns about a reproducibility crisis in science have been raised.[5]

The ultimate quality measure is the consistency of the map and an atomic model.[6] However, this is only possible for high-resolution maps. Alternatively, the standard map validation practice has mostly been restricted to the internal consistency of two half-maps calculated from independent halves of the whole dataset.[7] This internal consistency is essential, and it is a good measure of the presence of random fluctuations. Its main drawback is that it is not immune to systematic biases,[8] that is, systematic mistakes committed in both halves would be rewarded in terms of the Fourier shell correlation.

Over the years, there have been many suggestions about validation measures of CryoEM maps.[9] Unfortunately, most of these measures are not currently used due to their spread across multiple software tools and the associated difficulty accessing them. To alleviate this problem, we present a validation grading system and its public availability through a web server that qualifies the CryoEM map depending on the information available to assess it. This system grades a structure map at six different levels. In this way, a map could be validated at level 0 (the deposited map), 1 (two half maps), 2 (2D classes), 3 (particles), 4 (…+angular assignment), 5 (…+micrographs and coordinates). For each of these levels, we list algorithms that can be employed. In addition to this grading system, we have three optional qualifiers: A (…+atomic model), W (…+image processing workflow), and O (…+other techniques). Those depositions wanting to achieve the highest level of validation should deposit a relatively small number of particles (in the order of 10k randomly chosen from the final set of particles) along with their micrographs and all alignment parameters. They should also include a detailed description of the image processing workflow so that the final result can be fully understood (and ideally reproduced) and some extra validation by other experimental data. This high-quality standard is ideal and will not be accomplished shortly for all deposited structures. Nevertheless, it is essential to have it as a compass to direct our community efforts.

The web server is publicly available at **https://biocomp.cnb.csic.es/ EMValidationService/**. It returns a PDF report that evaluates the correctness of the submitted map from multiple figures of merit. Obviously, map submission to a public database cannot require a compulsory deposition of all these elements. Still, a validation grading system could be adopted in which, if they are given, the consistency of the map with the different aspects that give rise to it can be assessed.

The server is aimed at structures determined by single particle analysis (SPA). Maps determined by subtomogram averaging (STA) in cryo-electron tomography can share the levels 0, 1, A, and W with the maps coming from SPA. However, more specific analyses could be developed for STA maps.

We hope that the validation server proposed in this paper will help improve the understanding of CryoEM maps and reduce the reproducibility crisis, especially if the image processing workflow is also disclosed.

1 Validation methods

The following sections describe the different methods at one's disposal to validate a CryoEM map. The availability of these methods usually depends on the accessibility of extra information, like the set of particles supporting it, their angular assignment, *etc.* The level of map validation is defined as the highest consecutive number up to which there is information available. The three extra qualifiers can be added to any of the levels. The highest degree of validation would be 5AWO. Still, we could have, for instance, a map whose validation is 4W, meaning that it has reached level 4 and a detailed description of the image processing workflow is available. At present, the highest validation that a typical map from EMDB can achieve is 1A (map + half maps + atomic model). Interestingly, since February 2022, the deposition of half maps at the EMDB has been compulsory, meaning that we have moved from validation level 0 to level 1.

1.0   Level 0: map

The first level of validation is performed when just the reconstructed map is available along with a visualization threshold. At this point, several methods can evaluate the local resolution of the map and its hand.

0.a   Center analysis. Centering of the mass and extra space available to correct for the contrast transfer function (CTF). There should be at least 30–40 Å on each side for a proper correction.

0.b   Mask analysis. At the threshold value specified by the user, most of the mass should be collected in a single connected component.

0.c   Background analysis. If we analyze the gray values outside the mask, they should not have too negative values (*e.g.*, values below five times the standard deviation of the background noise).

0.d   *B*-factor analysis. The *B*-factor line,[10] fitted between 15 Å and the resolution reported, should have a slope that is between 0 and 300 Å^2.

0.e   DeepRes.[11] This method is based on a deep learning algorithm that assesses the similarity of the texture features present in the map to the texture features observed in atomic structures.

0.f   LocBfactor.[12] This method estimates a local resolution *B*-factor by decomposing the input map into a local magnitude and phase term using the spiral transform.

0.g   LocOccupancy.[12] This method estimates the occupancy of a voxel by the macromolecule.

0.h   DeepHand.[13] This method determines for maps whose resolution is higher than 5 Å whether the map has the right hand or, on the contrary, it is the mirrored version of the correct map.

Although not yet implemented, it would be possible to detect preferential orientations or artifacts by analyzing the macromolecule's local texture and noise.

1.1   Level 1: …+half maps

If independent half maps are available, then we can further assess the local resolution with different means:

1.a   Global resolution.[14] The Fourier shell correlation (FSC) between the two half maps is the most standard method to determine the global resolution of a map. However, other measures exist, such as the spectral signal-to-noise ratio and the differential phase residual. There is a long debate about the correct thresholds for these measures. Probably the clearest threshold is the one of the SSNR (SSNR = 1). For the DPR, we have chosen 103.9° (ref. 14) and for the FSC, the standard 0.143.

1.b   Permutation test FSC.[15] This method calculates a global resolution by formulating a hypothesis test in which the distribution of the FSC of noise is calculated from the two maps.

1.c   BlocRes.[16] This method computes a local Fourier shell correlation (FSC) between the two half maps.

1.d   Resmap.[17] This method is based on a test hypothesis testing the superiority of signal over noise at different frequencies.

1.e   MonoRes.[18] This method evaluates the local energy of a point to the distribution of energy in the noise. This comparison is performed at multiple frequencies, and for each one, the monogenic transformation separates the amplitude and phase of the input map.

1.f   MonoDir.[19] This method extends the concept of local resolution to local and directional resolution by changing the shape of the filter applied to the input map. The directional analysis can reveal image alignment problems.

1.g   FSO. This method calculates the anisotropy of the energy distribution in Fourier shells. It is an indirect measure of the anisotropy of the angular distribution or the presence of heterogeneity.

1.h   FSC Directional.[20] This method analyzes the FSC in different directions and evaluates its homogeneity through the sphericity of the FSC surface.

1.2   Level 2: …+2D classes

If 2D classes of the particles used for the reconstruction are available, then the following method can be applied:

2.a   Reprojection consistency. The 2D classes can be aligned against the reconstructed map, then the correlation between reprojections of the map and the 2D classes can be analyzed. Also, analyzing the residuals (2D class minus the corresponding reprojection) can reveal systematic differences.

1.3   Level 3: …+particles

If a random subset of the particles is provided, the following actions can be performed:

3.a   Outlier detection. The set of particles is classified into the input set of 2D classes of level 2. The number of particles considered to be outliers in those classes is reported. A particle is an outlier if its Mahalanobis distance to the centroid of the class is larger than 3.[21]

3.b   2D classification internal consistency. The input particles are classified in 2D clusters. The quality of the 2D clusters is assessed through Fourier ring correlation.

3.c   2D classification external consistency. We measure the overlap between the subspace spanned by the classes in level 2 and the classes of level 3.

1.4   Level 4: …+angular assignment

If the angular assignment of the particles in level 3 is available, then the following methods can be applied:

4.a   Similarity criteria. Analysis of the distribution of the similarity between the input particles and the reprojection from the same angular orientation by different scores.

4.b   Alignability smoothness.[22] This algorithm analyzes the smoothness of the correlation function over the projection sphere and the stability of its maximum.

4.c   Alignability precision and accuracy. The precision[23] analyzes the orientation distribution of the best matching reprojections from the reference volume. If the high values are clustered around the same orientation, the precision is close to 1. Otherwise, it is closer to -1. Below 0.5, the best directions tend to be scattered. The alignability accuracy[24] compares the final angular assignment with the result of a new angular assignment. The similarity between both is again encoded between -1 and 1.

4.d   Angular error distribution. Angular error distribution between the provided angles and an independent angular assignment performed with state-of-the-art algorithms.

4.e   Classification without alignment. 3D classification of the input particles without angular refinement.

4.f   Detection of overfitting.[25] This method compares the resolution achieved by subsets of images of increasing size and by subsets of noise images of the same size.

4.g   Angular distribution efficiency.[26] This method evaluates the ability of the angular distribution to fill the Fourier space.

4.h   Sampling compensation factor.[27] This method is another way of measuring the ability of the angular distribution to fill the Fourier space.

4.i   CFT stability. Analysis of the stability of the defocus parameters. For this purpose, the defocus, B-factor, astigmatism, and phase shift can be estimated from the given particles, and these refined parameters' deviations are reported. Ideally, the differences in defoci cannot be larger than the ice thickness. The same can be done with local magnification offsets (which should be around 0) and the B-factor.

1.5   Level 5: …+micrographs and coordinates

If a random subset of micrographs and their corresponding coordinates are available, then the following measure can be taken:

5.a   Micrograph cleaner.[28] This method assigns a score between 0 and 1, reflecting the probability that the coordinate is outside a region with aggregations, ice crystals, carbon edges, *etc.*

1.6   Level A: …+atomic model

If a fitted atomic model is available, then we may apply the following validation methods:

A.a   Map-Q.[29] This method computes the local correlation between the map and each one of its atoms assumed to have a Gaussian shape.

A.b   FSC-Q.[30] This method compares the local FSC between the map and the atomic model to the local FSC of the two half maps.

A.c   Model ambiguity by molecular dynamics.[31] This method estimates the ambiguity of the atomic model in each region of the CryoEM map due to the different local resolutions or local heterogeneity.

A.d   Guinier plot of model and map.[32] This method compares the falloff in Fourier space between the map and atomic model.

A.e   Phenix CryoEM validation tools.[33] Phenix provides several tools to assess the agreement between the experimental map and its atomic model. Two large clusters of these measurements are: (1) different ways of measuring the cross-correlation between the map and model, and (2) different ways of measuring the resolution between the map and model.

A.f   EMRinger.[34] This algorithm compares the side chains of the atomic model to the CryoEM map.

A.g   DAQ.[35] This algorithm uses deep learning that can estimate the residue-wise local quality for protein models from cryo-electron microscopy (EM) maps. The method calculates the likelihood that a given density feature corresponds to an amino acid, atom, and secondary structure. These likelihoods are combined into a score that ranges from -1 (bad quality) to 1 (good quality).

We are aware that the report generated at the submission to EMDB/PDB is relatively rich in this area, including methods such as Molprobity[36] or TEMPy2.[37] Our goal is to complement this analysis with alternative tools. Some of these analyses (0.a, 0.b, 0.c, 2.a, 3.b, 3.c, 4.a, 4.d, 4.e, and 4.i) have been newly developed for this server.

1.7   Qualifier W: …+workflow

Another level of validation purely based on image processing is a detailed description of each of the image processing steps used to produce the final map so that an external user can understand the results of each step and has all the information to reproduce the whole pipeline. This information is produced, for example, by Scipion[38] in the form of a JSON that can be submitted to EMPIAR[1] or Scipion's workflow repository (**https://nolan.cnb.csic.es/cryoemworkflowviewer**) and visualized with a JavaScript viewer integrated into these databases. Alternative methods would also be valid, specifically tailored to each one of the different image processing packages.

1.8   Qualifier O: …+other techniques

Finally, if extra experimental work is available, then we may apply the following techniques.

O.a   Mass spectroscopy.[39] This method uses information from cross- and mono-links to validate the atomic model.

O.b   SAXS.[40] This method compares the expected energy profile from the reconstructed map to the one obtained by a SAXS experiment.

O.c   Tilt pair validation.[41] This method is capable of experimentally validating the hand of the reconstructed map by comparing the angular assignment of two sets of particles related by a single-axis tilt.

1.9   Availability

We have created a web server located at **https://biocomp.cnb.csic.es/ EMValidationService/**. The server has a web interface that guides the user through the different steps. We have not made the levels to be compulsorily progressive. For instance, one could have validation levels 0 and 2 without having the information for level 1. Although possible, this option is discouraged. Method A.c is sensible, but this method takes much time to execute due to the molecular dynamics underneath (about 6 hours per constructed model). For this reason, we recommend not to run it regularly to not saturate the server.

Once a job is submitted to the server, the execution time varies from 20 minutes to 16 hours if A.c is not executed or 3 days if it is. The execution time depends on the number of validations to perform, the size of the reconstructed map, and the number of images provided for validation. During its execution, the user gets a token so that they can check whether the job has already been finished.

The server asks for a URL where the image processing workflow can be visualized for validation. If Scipion has created the workflow, it is included in the report. However, the validation information truly lies on the given URL. For submitting the Scipion workflow, the user may use the scipion-em-datamanager plugin that submits the current project to the workflow repository at **https:// nolan.cnb.csic.es/cryoemworkflowviewer/entries**.   An   example   of   such a workflow can be seen in Fig. 1.
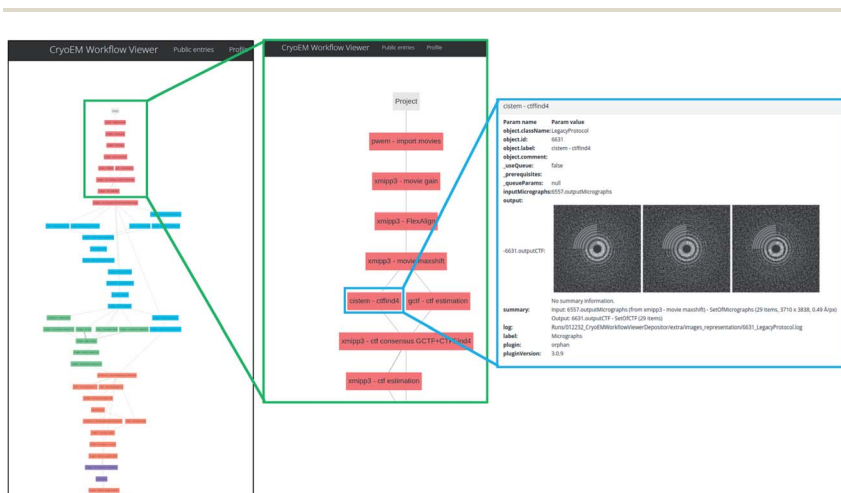
Fig. 1 An example of workflow visualization at Scipion's workflow repository. Each of the colored boxes represent a protocol executed within Scipion. The last panel on the right shows how the input parameters, data and results can be inspected using a web browser.

The server automatically constructs a Scipion workflow[38] based on the input data. Depending on the amount of data available (levels 0, 1, 2, …), the appropriate protocols are instanced and interconnected. The source code of the server is available at **https://github.com/I2PC/scipion-em-validation**. The program creates the Scipion workflow, automatically analyzes the results, and constructs a Latex document which is later compiled into a PDF. This report is handed to the user.

2   Results

To show the usefulness of the validation reports, we have applied this methodology to multiple datasets. The first one comes from the tutorial of Scipion,[38] while all others are examples from the EMDB. In the first one, we have information for evaluating all validation levels (5AWO), while for the EMDB, the validation reports are of levels 0A or 1A.

2.1   Full report

ESI 1† shows the validation report for Scipion's tutorial. The data corresponds to 30 micrographs of the apoferritin EMPIAR 10248 dataset.[42] The tutorial used 1457
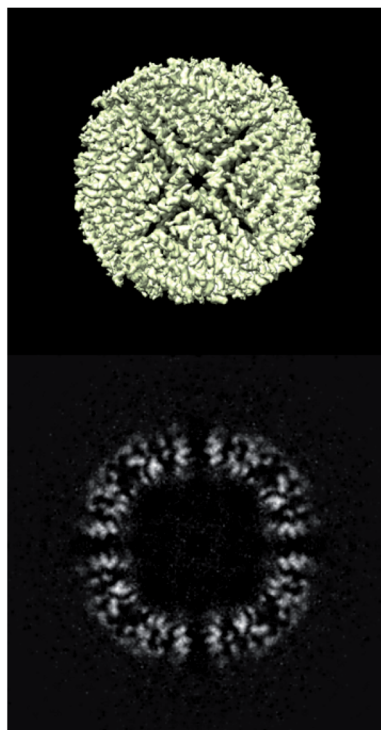
Fig. 2   Isosurface and central slice of the Scipion's tutorial dataset, apoferritin.

Abstract

The map seems to be well centered. There is no problem with the suggested threshold. There seems to be a problem with the map's background (see Sec. 2.3). The resolution does not seem to be uniform in all directions (see Sec. 4.6). The 2D classes provided by the user do not seem to correlate well with the reprojections of the map (see Sec. 6.1). It seems that the input particles cannot be easily aligned (see Sec. 9.2). It seems that the angular assignment given by the user does not match with the one produced by CryoSparc (see Sec. 9.5). It seems that the angular assignment produced by Relion does not match with the one produced by Cryosparc (see Sec. 9.6). This is probably a sign of the difficulty to align these particles. It seems that there is some problem with the CTF (see Sec. 9.11). According to phenix, it seems that there might be some mismatch between the map and its model (see Sec. 13.5). The EMRinger score is negative, it seems that the model side chains do not match the map (see Sec. 13.6). DAQ detects some mismatch between the map and its model (see Sec. 13.7).

The average resolution of the map estimated by various methods goes from 2.0Å to 4.6Å with an average of 3.3Å. The resolution provided by the user was 2.6Å. The resolution reported by the user may be overestimated.

The overall score (passing tests) of this report is 26 out of 36 evaluable items.

Fig. 3 Example of the abstract of the full report generated for Scipion's tutorial data.

experimental projections to construct a map whose nominal resolution is 2.6 Å. Although generally, the map is good (see Fig. 2) and reasonably agrees with its atomic model by visual inspection, the validation report shows that it has several problems:

(1) The background outside the map does not have zero mean, and it contains a significant amount of intensity outliers.

(2) The reported resolution seems to be overestimated according to several local resolution algorithms (DeepRes, Resmap, and MonoDir) and Phenix's comparison to the atomic model. The average resolution seems to be more around 4.4 Å, rather than 2.6 Å.

(3) Images are difficult to align, as reported by the alignability smoothness, CryoSparc alignment, and the comparison between CryoSparc and Relion alignments. This uncertainty in the alignment was also shown in MonoDir radial plots.

(4) Relion could not reliably determine the scaling factor of the images.

(5) The resolvability of the side chains is not good, as reported by EMRinger and DAQ.

None of these errors is terrible, and it is the expected result for a tutorial reconstruction from only 30 micrographs. However, this contrasts with the resolution reported by the FSC, which gives the false impression of having achieved a better map than the one obtained.

Fig. 3–5 show the summary of the report generated for this dataset. In a single glimpse, the main problems can be easily identified.

2.2 A comparison of three SARS-CoV 2 spike structures

With the Covid-19 pandemic, many structural studies have addressed all the proteins amenable to CryoEM. We have chosen three reconstructions of the spike (EMDB entries 11 337,[43] 22 301,[44] and 22 838 [45]) which are supposed to be

0.a Mass analysis	Sec. 2.1	OK
0.b Mask analysis	Sec. 2.2	OK
0.c Background analysis	Sec. 2.3	2 warnings
0.d B-factor analysis	Sec. 2.4	OK
0.e DeepRes	Sec. 2.5	1 warnings
0.f LocBfactor	Sec. 2.6	OK
0.g LocOccupancy	Sec. 2.7	OK
0.h DeepHand	Sec. 2.8	OK
1.a Global resolution	Sec. 4.1	OK
1.b FSC permutation	Sec. 4.2	OK
1.c Blocres	Sec. 4.3	OK
1.d Resmap	Sec. 4.4	1 warnings
1.e MonoRes	Sec. 4.5	OK
1.f MonoDir	Sec. 4.6	1 warnings
1.g FSO	Sec. 4.7	OK
1.h FSC3D	Sec. 6.1	OK
2.a Reprojection consistency	Sec. 6.1	OK
3.a Outlier detection	Sec. 9.1	OK
3.b 2D Classification internal consistency	Sec. 8.2	Cannot be automated
3.c 2D Classification external consistency	Sec. 8.3	OK
4.a Similarity criteria	Sec. 9.1	Cannot be automated
4.b Alignability smoothness	Sec. 9.2	1 warnings
4.c Alignability precision and accuracy	Sec. 9.3	OK
4.d1 Relion alignment	Sec. 9.4	OK
4.d2 CryoSparc alignment	Sec. 9.5	1 warnings
4.d3 Relion/CryoSparc alignments	Sec. 9.6	1 warnings
4.e Classification without alignment	Sec. 9.8	OK
4.f Overfitting detection	Sec. 9.8	OK
4.g Angular distribution efficiency	Sec. 9.9	OK
4.h SCF	Sec. 9.10	OK
4.i CTF stability	Sec. 9.11	1 warnings
5.a Micrograph cleaner	Sec. 11.1	OK
A.a MapQ	Sec. 13.1	OK
A.b FSC-Q	Sec. 13.2	OK
A.c Multimodel	Sec. 13.3	OK
A.d Map-Model Guinier	Sec. 13.4	OK
A.e Phenix validation	Sec. 13.5	1 warnings
A.f EMRinger	Sec. 13.6	1 warnings
A.g DAQ	Sec. 13.7	1 warnings
W Workflow	Sec. 14	Cannot be automated
O.b SAXS	Sec. 15.1	Cannot be automated

3

Fig. 4 Example of the index of the full report generated for Scipion's tutorial data.

reconstructed at similar resolutions (3.3, 3.7 and 3.84 Å, respectively; see Fig. 6). The first entry could be validated at level 1A, while the other two could only be validated at level 0A. The validation reports for these structures can be seen in ESI 2.†

The average resolutions measured by several methods for the different structures were 4.9, 8.2, and 3.8 Å, respectively. This points out that the internal resolution variability is much higher than the one reported by the FSC. We have empirically verified that very often, the resolution based on the FSC is at the lower extreme of the distribution of resolutions reported by many local resolution methods.[11] For this reason, this FSC resolution must be understood as "*there is a region in the map whose local resolution is this number*".

In this analysis, we can identify the following problems:

(1) The backgrounds of the three structures have problems as they are not equal to 0 and contain significant outliers.

Summary of the warnings across sections.
If it is empty below this point, it means that there are no warnings.

Section 2.3 (0.c Background analysis)
1. The null hypothesis that the background mean is 0 has been rejected because the p-value of the comparison is smaller than 0.001
2. There is a significant proportion of outlier values in the background (cdf5 ratio=2031.06)

Section 2.5 (0.e DeepRes)
1. The reported resolution, 2.60 Å, is particularly with respect to the local resolution distribution. It occupies the 0.00 percentile

Section 4.4 (1.d Resmap)
1. The reported resolution, 2.60 Å, is particularly with respect to the local resolution distribution. It occupies the 0.00 percentile

Section 4.6 (1.f MonoDir)
1. The distribution of best resolution is not uniform in all directions. The associated p-value is 0.000000.

Section 9.2 (4.b Alignability smoothness)
1. The percentage of images whose angular assignment is significantly away from the smoothed maximum is too high, 50.2%

Section 9.5 (4.d2 CryoSparc alignment)
1. The percentage of images with uncertain shift is larger than 20%

Section 9.6 (4.d3 Relion/CryoSparc alignments)
1. The percentage of images with uncertain shift is larger than 20%

Section 9.11 (4.i CTF stability)
1. The 95% confidence interval of scale factor is not centered.

Section 13.5 (A.e Phenix validation)
1. The resolution reported by the user, 2.6 Å, is significantly smaller than the resolution estimated between map and model (FSC=0.5), 4.4 Å

Section 13.6 (A.f EMRinger)
1. The EMRinger score is smaller than 1, it is 0.892.

Section 13.7 (A.g DAQ)

5

Fig. 5 Example of the warnings abstract of the full report generated for Scipion's tutorial data.

(2) Two of the maps, EMDB11337 and EMDB22838, have a reported resolution that is particularly good compared to the one computed by DeepRes (level 0) or other local resolution methods (level 1).

(3) EMDB22838 and EMDB22301 have either too low a resolution or a severe problem of anisotropic resolution, probably caused by the attraction of particles towards specific directions. Both issues are translated into the uncertainty of DeepHand to determine the handedness of the structure.

(4) EMDB11337 has problems fitting to the atomic map according to FSC-Q and Phenix. EMDB22301 has issues in this fitting according to MapQ, Phenix, and DAQ. Finally, EMDB 22838 has problems according to MapQ, EMRinger, and DAQ.
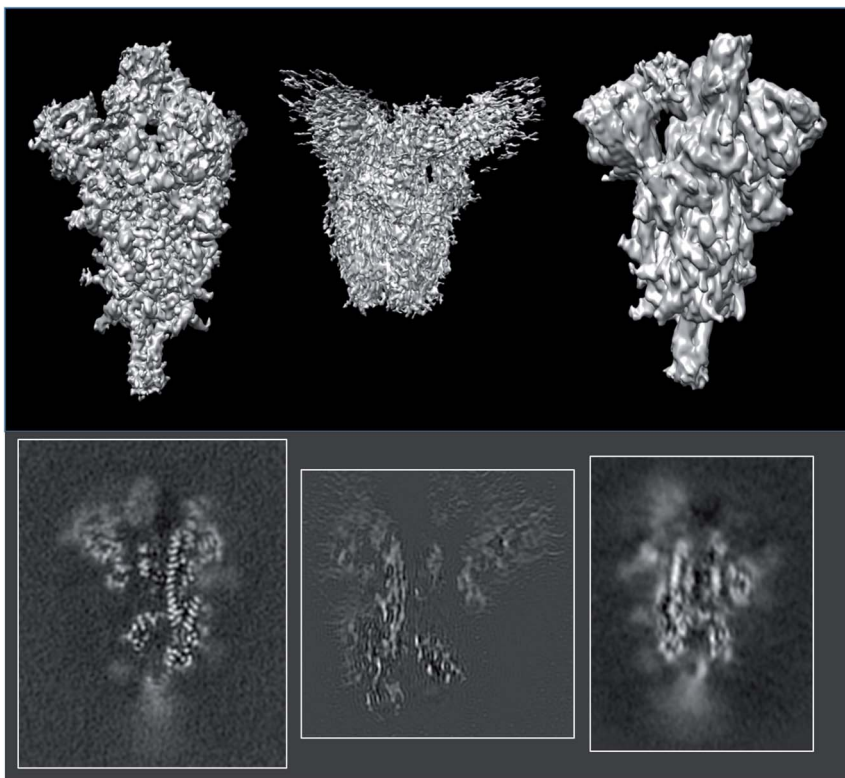
Fig. 6 Isosurfaces and central slices of EMDB 11337 (left), EMDB 22301 (middle), and EMDB 22838 (right), all of which are SARS–CoV2 spikes.

2.3   An analysis of a 1.15 Å resolution apoferritin

EMDB11668 is currently the map with the best resolution in the EMDB reconstructed by single particle analysis (see Fig. 7). It is the reconstruction of human apoferritin in a Titan Krios with a second-generation spherical aberration correction.[46] An analysis of the deposited structure reveals that the recommended threshold, 0.15, is too high and causes many mass and mask problems (see ESI 3 Report 11668_015†). In the following, we will analyze the map at a threshold of 0.05 (see ESI 3 Report 11668_005†).

In this analysis, we can identify the following problems:

(1) The mask of the map is very fragmented, with 478 289 connected components. The largest component only takes 42% of the mass. This is due to the boosting of the high frequencies caused by the B-factor correction.

(2) The mean of the background is not 0, and there is a significant amount of outlier values.

(3) The map has been B-factor corrected, resulting in a boost of noise at high frequency and a disagreement between the falloffs of the map and the model.

(4) There is a significant disagreement between the map and its model according to MapQ, Phenix, and DAQ in some regions. In Fig. 8, we show a region of the map and model in Coot. Although the model is excellent in many regions, there are portions in which the map and model do not match, as highlighted in the figure.
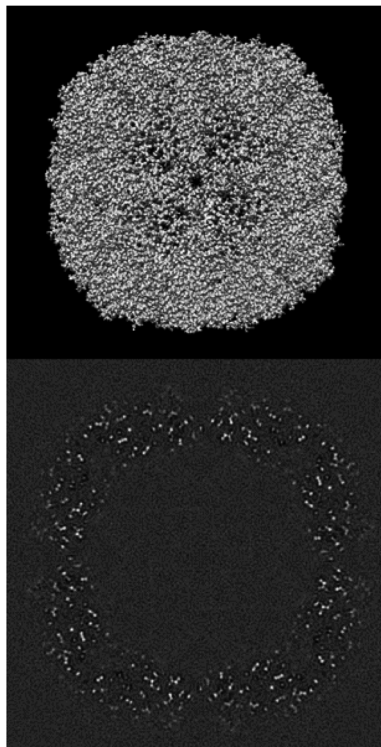
Fig. 7   Isosurface and central slice of EMDB 11668, apoferritin.
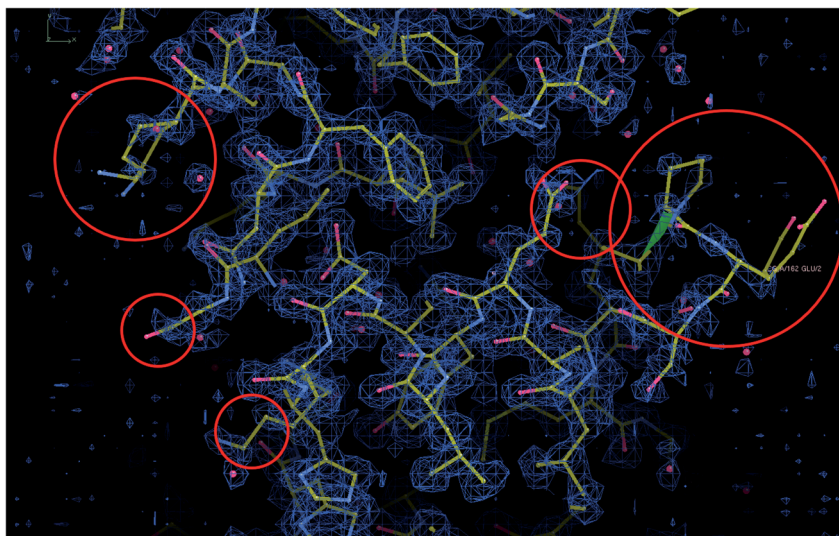
Fig. 8   Map and model representations of apoferritin with some mismatching areas highlighted.

3   Discussion

The determination of a three-dimensional map compatible with the CryoEM measurements of a macromolecule is full of decisions along the image processing pipeline (whether to keep or not this micrograph according to its contrast transfer function or beam-induced alignment, whether a region of that micrograph represents a centered projection of the macromolecule, which is the orientation of that projection with respect to a reference map and which is the exact conformation of that macromolecule, *etc.*). All these decisions involve parameter estimates. Some of these parameters are continuous, such as the micrograph defocus or the in-plane shifts of a given projection with respect to a reference map. Some others are discrete, such as whether a given projection is in one conformation or another. From this point of view, the reconstructed map is a signal (another parameter) to be estimated in a N^3-dimensional space, with N^3 being the number of map voxels.

All algorithms that have to estimate a parameter can commit an error (false positives or negatives if the parameter is binary, or a residual if the parameter is continuous). Our final map is a mixture of our good and bad decisions for all the particles involved in the reconstruction. To simplify the analysis, let us consider that we perfectly estimate all parameters of a given projection with probability p, and we do not, with probability $1 - p$. The fraction of well-estimated projections will provide us with a perfect reconstruction, while the fraction of incorrectly estimated parameters will result in an imperfect reconstruction:

$$\begin{aligned} V_{\text{reconstruction}} &= p V_{\text{perfect}} + (1 - p) V_{\text{imperfect}} \\ &= V_{\text{perfect}} + (1 - p)\left(V_{\text{imperfect}} - V_{\text{perfect}}\right) \end{aligned}$$

In this way, we see that our reconstruction is a perfect map plus bias that depends on the fraction of incorrectly estimated parameters, $1 - p$, and the difference between the imperfect and perfect reconstructions (if we make a mistake in the estimation of the projection direction of an experimental projection, but the map from the wrong direction is very similar to the map from the correct direction, the mistake will have a negligible effect). It is this bias that we call overfitting in our field. Hence, overfitting is directly linked to incorrect parameter estimates. In Sorzano *et al.*,[8] we discuss the different sources of map bias and how we can design image processing workflows that try to identify the incorrectly estimated parameters. Ultimately, this identification requires multiple estimates of the same parameter by, ideally, different algorithms. In this way, we can determine the reliability of any particular value.

The use of multiple algorithms to discard incorrectly estimated parameters or average those that agree more is seldom seen in our applied papers. However, this practice of estimating parameters only once leaves us in a position in which we cannot know whether the estimated parameter is stable or not.

Map and model validations have become a relevant research line in the field, responding to the need to assess the reliability of the reconstructed maps. Despite the multiple methods available to perform this validation, the most common is the reconstruction of two half-maps and their subsequent comparison through the Fourier shell correlation. This practice is known in the field as the gold standard. It is a cheap substitute for the multiple estimations of the same

parameter: we use the estimate of the N^3-dimensional signal as a proxy of the various estimations at a cost that involves estimating the parameters only once for each of the experimental images. If the parameter estimation errors are independent, zero-mean processes, this is a good practice. However, it does not protect us if the estimation errors are dependent or do not have a zero mean. For instance, a wrong estimate of the defocus of a micrograph will systematically affect the CTF correction of many particles. Suppose we have an attraction problem[8] (a projection direction has a larger signal-to-noise ratio, SNR, than its surrounding or a 3D class has a larger SNR than the alternatives). In that case, the parameters would also be systematically misestimated.

In Sorzano *et al.*,[8] we show that splitting the data into two halves is not a sufficient nor necessary condition for avoiding overfitting. It is not sufficient because we can make systematic mistakes in estimating the parameters in both halves. It is not necessary because not splitting the data into two halves does not necessarily lead to an incorrect estimation of the underlying parameters. The most common measure to assess the correction of a reconstructed map is the Fourier shell correlation between the two half maps. This measure is excellent when the difference between the two maps is independent and randomly distributed with zero mean. However, it is misled by systematic errors appearing in both halves.[8]

Despite the plethora of alternative validation methods, these are seldom used due to the difficulty of accessing them conveniently and smoothly. Scipion[38] is an appropriate platform for this evaluation as it integrates all the methods described in this paper, a total of 37, and provides effective ways of allowing them to interoperate. The server is open to new approaches, and any new validation tool is welcome to be incorporated into the validation report.

4   Conclusions

In this work, we have presented a validation server that assesses CryoEM maps' correction from multiple perspectives. The validation can be more profound as more data is available about the structure, the experimental images supporting it, and the image processing workflow followed to achieve it. In this way, we provide a progressive validation level that evaluates the map from different perspectives: (0) the map itself, (1) its two-halves, (2) the 2D classes of the particles, (3) the particles themselves, (4) the angular assignment of those particles, (5) the coordinates of those particles in the micrographs, (A) an atomic model of the structure, (W) the image processing workflow leading to the final result, and (O) validation through alternative experiments.

We believe that disclosing as much information as possible about the supporting experimental evidence leading to a CryoEM map will help the user of that structure better understand the strengths and weaknesses of the map at hand. And finally, it will ultimately contribute to more open and reproducible science.

Author contributions

COSS conceptualized the project, implemented its calculation engine, coordinated the integration of the different methods and performed the experiments. COSS and JMC analyzed the application and wrote the article. IS and JRM

implemented the webserver. All others helped to make the methods available through Scipion plugins. All authors have revised the manuscript.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

Notes and references

1 A. Patwardhan, *Acta Crystallogr., Sect. D: Struct. Biol.*, 2017, **73**, 503–508.

2 C. O. S. Sorzano and J. M. Carazo, Cryo-Electron Microscopy: the field of 1000+ methods, *J. Struct. Biol.*, 2022, **214**, 107861.

3 M. Woelfle, P. Olliaro and M. H. Todd, *Nat. Chem.*, 2011, **3**, 745–748.

4 A. Iudin, P. K. Korir, J. Salavert-Torres, G. J. Kleywegt and A. Patwardhan, *Nat. Methods*, 2016, **13**, 387–388.

5 M. Baker, *Nature*, 2016, **533**, 452–454.

6 C. L. Lawson, A. Kryshtafovych, P. D. Adams, P. V. Afonine, M. L. Baker, B. A. Barad, P. Bond, T. Burnley, R. Cao, J. Cheng, *et al.*, *Nat. Methods*, 2021, **18**, 156–164.

7 R. Henderson, A. Sali, M. L. Baker, B. Carragher, B. Devkota, K. H. Downing, E. H. Egelman, Z. Feng, J. Frank, N. Grigorieff, W. Jiang, S. J. Ludtke, O. Medalia, P. A. Penczek, P. B. Rosenthal, M. G. Rossmann, M. F. Schmid, G. F. Schröder, A. C. Steven, D. L. Stokes, J. D. Westbrook, W. Wriggers,

H. Yang, J. Young, H. M. Berman, W. Chiu, G. J. Kleywegt and C. L. Lawson, *Structure*, 2012, **20**, 205–214.

8 C. O. S. Sorzano, A. Jimenez-Moreno, D. Maluenda, M. Martinez, E. Ramirez-Aportela, R. Melero, A. Cuervo, J. Conesa, J. Filipovic, P. Conesa, L. del Caño, Y. C. Fonseca, J. Jiménez-de la Morena, P. Losana, R. Sáchez-García, D. Strelak, E. Fernandez-Gimenez, F. de Isidro, D. Herreros, J. L. Vilas, R. Marabini and J. M. Carazo, On bias, variance, overfitting, gold standard and consensus in Single Particle Analysis by Cryo-electron microscopy, *Acta Crystallogr., Sect. D: Struct. Biol.*, 2022, **D78**, 410–423.

9 J. L. Vilas, N. Tabassum, J. Mota, D. Maluenda, A. Jiménez-Moreno, T. Majtner, J. M. Carazo, S. T. Acton and C. O. S. Sorzano, *Curr. Opin. Struct. Biol.*, 2018, **52**, 127–145.

10 P. B. Rosenthal and R. Henderson, *J. Mol. Biol.*, 2003, **333**, 721–745.

11 E. Ramírez-Aportela, J. Mota, P. Conesa, J. M. Carazo and C. O. S. Sorzano, *IUCrJ*, 2019, **6**, 1054–1063.

12 S. Kaur, J. Gomez-Blanco, A. A. Khalifa, S. Adinarayanan, R. Sanchez-Garcia, D. Wrapp, J. S. McLellan, K. H. Bui and J. Vargas, *Nat. Commun.*, 2021, **12**, 1.

13 J. García Condado, A. Muñoz-Barrutia and C. O. S. Sorzano, *bioRxiv*, 2022, DOI: **10.1101/2022.03.01.482513**.

14 C. O. S. Sorzano, J. Vargas, J. Oton, V. Abrishami, J. M. de la Rosa-Trevin, J. Gomez-Blanco, J. L. Vilas, R. Marabini and J. M. Carazo, *Prog. Biophys. Mol. Biol.*, 2017, **124**, 1–30.

15 M. Beckers and C. Sachse, *J. Struct. Biol.*, 2020, **212**, 107579.

16 G. Cardone, J. B. Heymann and A. C. Steven, *J. Struct. Biol.*, 2013, **184**, 226–236.

17 A. Kucukelbir, F. J. Sigworth and H. D. Tagare, *Nat. Methods*, 2014, **11**, 63–65.

18 J. L. Vilas, J. Gómez-Blanco, P. Conesa, R. Melero, J. M. de la Rosa Trevín, J. Otón, J. Cuenca, R. Marabini, J. M. Carazo, J. Vargas and C. O. S. Sorzano, *Structure*, 2018, **26**, 337–344.

19 J. L. Vilas, H. D. Tagare, J. Vargas, J. M. Carazo and C. O. S. Sorzano, *Nat. Commun.*, 2020, **11**, 55.

20 Y. Z. Tan, P. R. Baldwin, J. H. Davis, J. R. Williamson, C. S. Potter, B. Carragher and D. Lyumkis, *Nat. Methods*, 2017, **14**, 793–796.

21 C. O. S. Sorzano, J. Vargas, J. M. de la Rosa-Trevín, A. Zaldívar-Peraza, J. Otón, V. Abrishami, I. Foche, R. Marabini, G. Caffarena and J. M. Carazo, *Proc. Intl. Work-Conference on Bioinformatics and Biomedical Engineering, IWBBIO*, 2014, p. 950.

22 J. Méndez, E. Garduño, J. M. Carazo and C. O. S. Sorzano, *J. Struct. Biol.*, 2021, **213**, 107771.

23 J. Vargas, J. Otón, R. Marabini, J. M. Carazo and C. O. S. Sorzano, *Sci. Rep.*, 2016, **6**, 21626.

24 J. Vargas, R. Melero, J. Gómez-Blanco, J. M. Carazo and C. O. S. Sorzano, *Sci. Rep.*, 2017, **7**, 6307.

25 B. Heymann, *AIMS Biophys.*, 2015, **2**, 21–35.

26 K. Naydenova and C. J. Russo, *Nat. Commun.*, 2017, **8**, 629.

27 P. R. Baldwin and D. Lyumkis, *Prog. Biophys. Mol. Biol.*, 2020, **150**, 160–183.

28 R. Sanchez-Garcia, J. Segura, D. Maluenda, C. O. S. Sorzano and J. M. Carazo, *J. Struct. Biol.*, 2020, **210**, 107498.

29 G. Pintilie and W. Chiu, *Acta Crystallogr., Sect. D: Struct. Biol.*, 2021, **77**, 1142–1152.

30 E. Ramírez-Aportela, D. Maluenda, Y. C. Fonseca, P. Conesa, R. Marabini, J. B. Heymann, J. M. Carazo and C. O. S. Sorzano, *Nat. Commun.*, 2021, **12**, 42.

31 M. A. Herzik, J. S. Fraser and G. C. Lander, *Structure*, 2019, **27**, 344–358.

32 J. L. Vilas, J. Vargas, M. Martínez, E. Ramírez-Aportela, R. Melero, A. Jiménez-Moreno, E. Garduño, P. Conesa, R. Marabini, D. Maluenda, J. M. Carazo and C. O. S. Sorzano, *J. Struct. Biol.*, 2020, **209**, 107447.

33 P. V. Afonine, B. P. Klaholz, N. W. Moriarty, B. K. Poon, O. V. Sobolev, T. C. Terwilliger, P. D. Adams and A. Urzhumtsev, *Acta Crystallogr., Sect. D: Struct. Biol.*, 2018, **74**, 814–840.

34 B. A. Barad, N. Echols, R. Y.-R. Wang, Y. Cheng, F. DiMaio, P. D. Adams and J. S. Fraser, *Nat. Methods*, 2015, **12**, 943–946.

35 G. Terashi, X. Wang, S. R. M. V. Subramaniya, J. J. G. Tesmer and D. Kihara, 2022, submitted, **https://colab.research.google.com/drive/1Q-Dj42QjVO8TCOLXMQBJlvm1zInxPkOu?usp=sharing**.

36 C. J. Williams, J. J. Headd, N. W. Moriarty, M. G. Prisant, L. L. Videau, L. N. Deis, V. Verma, D. A. Keedy, B. J. Hintze, V. B. Chen, S. Jain, S. M. Lewis, W. B. Arendall, J. Snoeyink, P. D. Adams, S. C. Lovell, J. S. Richardson and D. C. Richardson, *Protein Sci.*, 2018, **27**, 293–315.

37 T. Cragnolini, H. Sahota, A. P. Joseph, A. Sweeney, S. Malhotra, D. Vasishtan and M. Topf, *Acta Crystallogr., Sect. D: Struct. Biol.*, 2021, **77**, 41–47.

38 J. M. de la Rosa-Trevín, A. Quintana, L. Del Cano, A. Zaldívar, I. Foche, J. Gutiérrez, J. Gómez-Blanco, J. Burguet-Castell, J. Cuenca-Alba, V. Abrishami, J. Vargas, J. Otón, G. Sharov, J. L. Vilas, J. Navas, P. Conesa, M. Kazemi, R. Marabini, C. O. S. Sorzano and J. M. Carazo, *J. Struct. Biol.*, 2016, **195**, 93–99.

39 M. Sinnott, S. Malhotra, M. S. Madhusudhan, K. Thalassinos and M. Topf, *Structure*, 2020, **28**, 1061–1070.

40 A. Jiménez, S. Jonic, T. Majtner, J. Oón, J. L. Vilas, D. Maluenda, J. Mota, E. Ramírez-Aportela, M. Martínez, Y. Rancel, J. Segura, R. Sánchez-García, R. Melero, L. Del Caño, P. Conesa, L. Skjaerven, R. Marabini, J. M. Carazo and C. O. S. Sorzano, *Bioinformatics*, 2019, **35**, 2427–2433.

41 R. Henderson, S. Chen, J. Z. Chen, N. Grigorieff, L. A. Passmore, L. Ciccarelli, J. L. Rubinstein, R. A. Crowther, P. L. Stewart and P. B. Rosenthal, *J. Mol. Biol.*, 2011, **413**, 1028–1046.

42 T. Kato, F. Makino, T. Nakane, N. Terahara, T. Kaneko, Y. Shimizu, S. Motoki, I. Ishikawa, K. Yonekura and K. Namba, *Microsc. Microanal.*, 2019, **25**, 998–999.

43 R. Melero, C. O. S. Sorzano, B. Foster, J. L. Vilas, M. Martínez, R. Marabini, E. Ramírez-Aportela, R. Sanchez-Garcia, D. Herreros, L. Del Caño, P. Losana, Y. Fonseca-Reyna, P. Conesa, D. Wrapp, P. Chacón, J. S. McLellan, H. D. Tagare and J. M. Carazo, *IUCrJ*, 2020, **7**, 1059–1069.

44 L. Yurkovetskiy, X. Wang, K. E. Pascal, C. Tomkins-Tinch, T. P. Nyalile, Y. Wang, A. Baum, W. E. Diehl, A. Dauphin, C. Carbone, K. Veinotte, S. B. Egri, S. F. Schaffner, J. E. Lemieux, J. B. Munro, A. Rafique, A. Barve, P. C. Sabeti, C. A. Kyratsous, N. V. Dudkina, K. Shen and J. Luban, *Cell*, 2020, **183**, 739–751.

45 S. M.-C. Gobeil, K. Janowska, S. McDowell, K. Mansouri, R. Parks, K. Manne, V. Stalls, M. F. Kopp, R. Henderson, R. J. Edwards, B. F. Haynes and P. Acharya, *Cell Rep.*, 2021, **34**, 108630.

46 K. M. Yip, N. Fischer, E. Paknia, A. Chari and H. Stark, *Nature*, 2020, **587**, 157–161.

Validation report of Level(s)
0, 1, 2, 3, 4, 5, A, W, O

I^{2}PC Validation server

February 25, 2022
4:07pm

Abstract

The map seems to be well centered. There is no problem with the suggested threshold. There seems to be a problem with the map's background (see Sec. 2.3). The resolution does not seem to be uniform in all directions (see Sec. 4.6). The 2D classes provided by the user do not seem to correlate well with the reprojections of the map (see Sec. 6.1). It seems that the input particles cannot be easily aligned (see Sec. 9.2). It seems that the angular assignment given by the user does not match with the one produced by CryoSparc (see Sec. 9.5). It seems that the angular assignment produced by Relion does not match with the one produced by Cryosparc (see Sec. 9.6). This is probably a sign of the difficulty to align these particles. It seems that there is some problem with the CTF (see Sec. 9.11). According to phenix, it seems that there might be some mismatch between the map and its model (see Sec. 13.5). The EMRinger score is negative, it seems that the model side chains do not match the map (see Sec. 13.6). DAQ detects some mismatch between the map and its model (see Sec. 13.7).

The average resolution of the map estimated by various methods goes from 2.0Å to 4.6Å with an average of 3.3Å. The resolution provided by the user was 2.6Å. The resolution reported by the user may be overestimated.

The overall score (passing tests) of this report is 26 out of 36 evaluable items.

0.a Mass analysis	Sec. 2.1	OK
0.b Mask analysis	Sec. 2.2	OK
0.c Background analysis	Sec. 2.3	2 warnings
0.d B-factor analysis	Sec. 2.4	OK
0.e DeepRes	Sec. 2.5	1 warnings
0.f LocBfactor	Sec. 2.6	OK
0.g LocOccupancy	Sec. 2.7	OK
0.h DeepHand	Sec. 2.8	OK
1.a Global resolution	Sec. 4.1	OK
1.b FSC permutation	Sec. 4.2	OK
1.c Blocres	Sec. 4.3	OK
1.d Resmap	Sec. 4.4	1 warnings
1.e MonoRes	Sec. 4.5	OK
1.f MonoDir	Sec. 4.6	1 warnings
1.g FSO	Sec. 4.7	OK
1.h FSC3D	Sec. 6.1	OK
2.a Reprojection consistency	Sec. 6.1	OK
3.a Outlier detection	Sec. 9.1	OK
3.b 2D Classification internal consistency	Sec. 8.2	Cannot be automated
3.c 2D Classification external consistency	Sec. 8.3	OK
4.a Similarity criteria	Sec. 9.1	Cannot be automated
4.b Alignability smoothness	Sec. 9.2	1 warnings
4.c Alignability precision and accuracy	Sec. 9.3	OK
4.d1 Relion alignment	Sec. 9.4	OK
4.d2 CryoSparc alignment	Sec. 9.5	1 warnings
4.d3 Relion/CryoSparc alignments	Sec. 9.6	1 warnings
4.e Classification without alignment	Sec. 9.8	OK
4.f Overfitting detection	Sec. 9.8	OK
4.g Angular distribution efficiency	Sec. 9.9	OK
4.h SCF	Sec. 9.10	OK
4.i CTF stability	Sec. 9.11	1 warnings
5.a Micrograph cleaner	Sec. 11.1	OK
A.a MapQ	Sec. 13.1	OK
A.b FSC-Q	Sec. 13.2	OK
A.c Multimodel	Sec. 13.3	OK
A.d Map-Model Guinier	Sec. 13.4	OK
A.e Phenix validation	Sec. 13.5	1 warnings
A.f EMRinger	Sec. 13.6	1 warnings
A.g DAQ	Sec. 13.7	1 warnings
W Workflow	Sec. 14	Cannot be automated
O.b SAXS	Sec. 15.1	Cannot be automated

Summary of the warnings across sections.
If it is empty below this point, it means that there are no warnings.

Section 2.3 (0.c Background analysis)
1. **The null hypothesis that the background mean is 0 has been rejected because the p-value of the comparison is smaller than 0.001**
2. **There is a significant proportion of outlier values in the background (cdf5 ratio=2031.06)**

Section 2.5 (0.e DeepRes)
1. **The reported resolution, 2.60 Å, is particularly with respect to the local resolution distribution. It occupies the 0.00 percentile**

Section 4.4 (1.d Resmap)
1. **The reported resolution, 2.60 Å, is particularly with respect to the local resolution distribution. It occupies the 0.00 percentile**

Section 4.6 (1.f MonoDir)
1. **The distribution of best resolution is not uniform in all directions. The associated p-value is 0.000000.**

Section 9.2 (4.b Alignability smoothness)
1. **The percentage of images whose angular assignment is significantly away from the smoothed maximum is too high, 50.2%**

Section 9.5 (4.d2 CryoSparc alignment)
1. **The percentage of images with uncertain shift is larger than 20%**

Section 9.6 (4.d3 Relion/CryoSparc alignments)
1. **The percentage of images with uncertain shift is larger than 20%**

Section 9.11 (4.i CTF stability)
1. **The 95% confidence interval of scale factor is not centered.**

Section 13.5 (A.e Phenix validation)
1. **The resolution reported by the user, 2.6 Å, is significantly smaller than the resolution estimated between map and model (FSC=0.5), 4.4 Å**

Section 13.6 (A.f EMRinger)
1. **The EMRinger score is smaller than 1, it is 0.892.**

Section 13.7 (A.g DAQ)

1. **The average DAQ is smaller than 0.5.**

Contents

1 Input data

Input map: /home/coss/ScipionUserData/projects/Example_10248_Scipion3/-
Runs/010948_XmippProtLocSharp/extra/sharpenedMap_1.mrc
SHA256 hash: 58f4d24dafbbf69aad9790730d8910d2fbbe0de545b17f28850bc2a1b4b5230b
Voxel size: 0.740000 (Å)
Visualization threshold: 0.002500
Resolution estimated by user: 2.600000

Orthogonal slices of the input map
Explanation:
In the orthogonal slices of the map, the noise outside the protein should not
have any structure (stripes going out, small blobs, particularly high or low
densities, ...)

Results:
See Fig. 1.

(a) X Slice 125   (b) Y Slice 125   (c) Z Slice 125

Figure 1: Central slices of the input map in the three dimensions

Orthogonal slices of maximum variance of the input map
Results:
See Fig. 2.

(a) X Slice 123   (b) Y Slice 123   (c) Z Slice 123

Figure 2: Slices of maximum variation in the three dimensions

Orthogonal projections of the input map

Explanation:

In the projections there should not be stripes (this is an indication of directional overweighting, or angular attraction), and there should not be a dark halo around or inside the structure (this is an indication of incorrect CTF correction or the reconstruction of a biased map).

Results:

See Fig. 3.

(a) X Projection   (b) Y Projection   (c) Z Projection

Figure 3: Projections in the three dimensions

Isosurface views of the input map

Explanation:

An isosurface is the surface of all points that have the same gray value. In these views there should not be many artifacts or noise blobs around the map.

Results:
See Fig. 4.

(a) View 1   (b) View 2   (c) View 3

Figure 4: Isosurface at threshold=0.002500. Views generated by ChimeraX at a the following X, Y, Z angles: View 1 (0,0,0), View 2 (90, 0, 0), View 3 (0, 90, 0).

Orthogonal slices of maximum variance of the mask
Explanation:
The mask has been calculated at the suggested threshold 0.002500, the largest connected component was selected, and then dilated by 2Å.
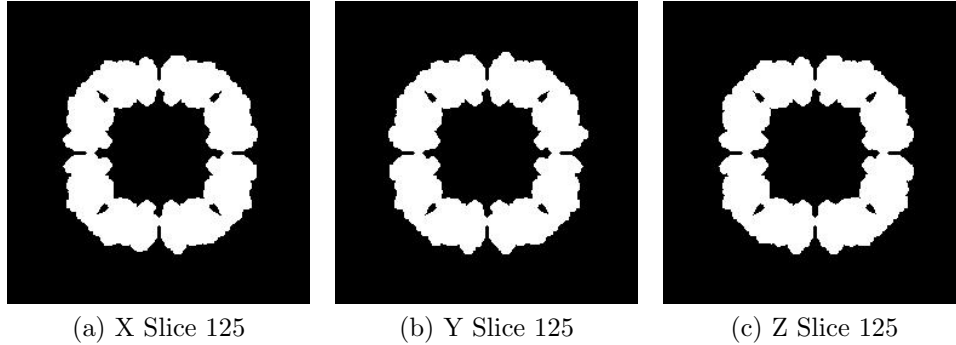
Results:
See Fig. 5.

| (a) X Slice 125 | (b) Y Slice 125 | (c) Z Slice 125 |

Figure 5: Slices of maximum variation in the three dimensions of the mask

2 Level 0 analysis

2.1 Level 0.a Mass analysis

Explanation:
The reconstructed map must be relatively well centered in the box, and there should be at least 30Å (the exact size depends on the CTF) on each side to make sure that the CTF can be appropriately corrected.

Results:
The space from the left and right in X are 31.08 and 31.08 Å, respectively. There is a decentering ratio (abs(Right-Left)/Size)% of 0.00%

The space from the left and right in Y are 33.30 and 31.08 Å, respectively. There is a decentering ratio (abs(Right-Left)/Size)% of 1.20%

The space from the left and right in Z are 32.56 and 31.08 Å, respectively. There is a decentering ratio (abs(Right-Left)/Size)% of 0.80%

The center of mass is at (x,y,z)=(125.05,125.02,124.99). The decentering of the center of mass (abs(Center)/Size)% is 0.02, 0.01, and 0.01, respectively.%

Automatic criteria: The validation is OK if 1) the decentering and

12

center of mass less than 20% of the map dimensions in all directions, and 2) the extra space on each direction is more than 20% of the map dimensions.

STATUS: OK

2.2 Level 0.b Mask analysis

Explanation:
The map at the suggested threshold should have most of its mass concentrated in a single connected component. It is normal that after thresholding there are a few thousands of very small, disconnected noise blobs. However, there total mass should not exceed 10%. The raw mask (just thresholding) and the mask constructed for the analysis (thresholding + largest connected component + dilation) should significantly overlap. Overlap is defined by the overlapping coefficient (size(Raw AND Constructed)/size(Raw)) that is a number between 0 and 1, the closer to 1, the more they agree.

Results:

Raw mask: At threshold 0.002500, there are 2172 connected components with a total number of voxels of 357214 and a volume of 144751.69 Å^3 (see Fig. 6). The size and percentage of the total number of voxels for the raw mask are listed below (up to 95% of the mass), the list contains (No. voxels (volume in Å^3), percentage, cumulatedPercentage):

(353875 (143398.64), 99.07, 99.07)

Number of components to reach 95% of the mass: 1

The average size of the remaining 2171 components is 1.54 voxels (0.41 Å^3). Their size go from 72 voxels (29.18 Å^3) to 1 voxels (0.41 Å^3).
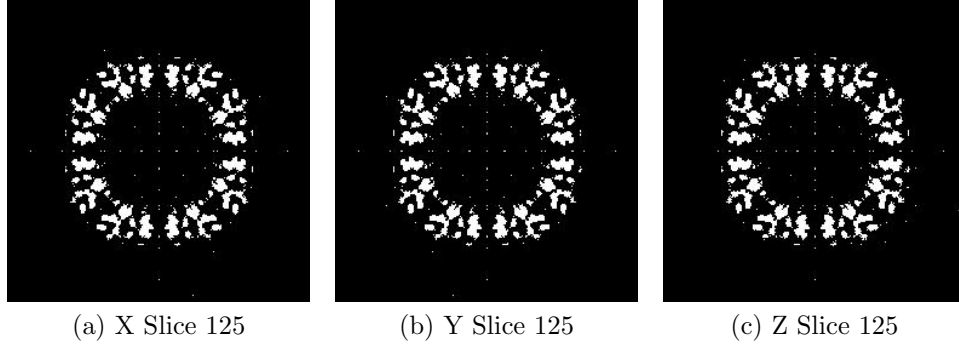
The slices of the raw mask can be seen in Fig. 6.

(a) X Slice 125	(b) Y Slice 125	(c) Z Slice 125

Figure 6: Maximum variance slices in the three dimensions of the raw mask

The following table shows the variation of the mass enclosed at different thresholds (see Fig. 7):

Threshold	Voxel mass	Molecular mass(kDa)	# Aminoacids
0.0006	934328.00	313.68	2851.64
0.0013	600420.00	201.58	1832.53
0.0019	448905.00	150.71	1370.09
0.0025	350635.00	117.72	1070.16
0.0032	279330.00	93.78	852.54
0.0038	223199.00	74.93	681.22
0.0045	179094.00	60.13	546.61
0.0051	143532.00	48.19	438.07
0.0057	114177.00	38.33	348.48
0.0064	89515.00	30.05	273.21
0.0070	69341.00	23.28	211.63
0.0076	52372.00	17.58	159.84
0.0083	38167.00	12.81	116.49
0.0089	26921.00	9.04	82.16
0.0096	18314.00	6.15	55.90
0.0102	12779.00	4.29	39.00
0.0108	7836.00	2.63	23.92
0.0115	4494.00	1.51	13.72
0.0121	2108.00	0.71	6.43
0.0127	1032.00	0.35	3.15
0.0134	534.00	0.18	1.63
0.0140	241.00	0.08	0.74
0.0147	96.00	0.03	0.29
0.0153	30.00	0.01	0.09
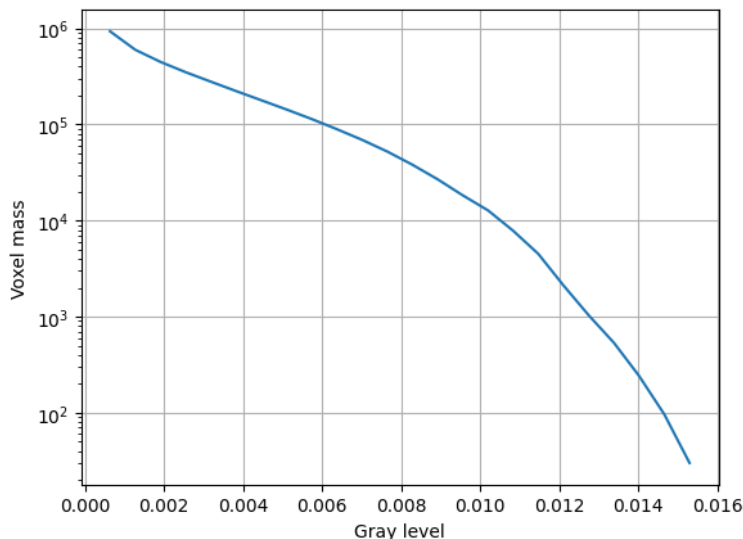
Figure 7: Voxel mass as a function of the gray level.

Constructed mask: After keeping the largest component of the previous mask and dilating it by 2Å, there is a total number of voxels of 1730415 and a volume of 701205.69 Å^3. The overlap between the raw and constructed mask is 1.00.

Automatic criteria: The validation is OK if 1) to keep 95% of the mass we need to keep at most 5 connected components; and 2) the average volume of the blobs outside the given threshold has a size smaller than 5Å^3; and 3) the overlap between the raw mask and the mask constructed for the analysis is larger than 75%.

STATUS: OK

2.3 Level 0.c Background analysis

Explanation:
Background is defined as the region outside the macromolecule mask. The background mean should be zero, and the number of voxels with a very low or very high value (below 5 standard deviations of the noise) should be very small and they should be randomly distributed without any specific structure.

15

Sometimes, you can see some structure due to the symmetry of the structure.

Results:

The null hypothesis that the background mean is 0 was tested with a one-sample Student's t-test. The resulting t-statistic and p-value were -666.91 and 0.000000, respectively.

The mean and standard deviation of the background were -0.000050 and 0.000279. The percentage of background voxels whose absolute value is larger than 5 times the standard deviation is 0.12 % (see Fig. 8). The same percentage from a Gaussian would be 0.000057% (ratio between the two percentages: 2031.055410).

Slices of the background beyond 5*sigma can be seen in Fig. 8.
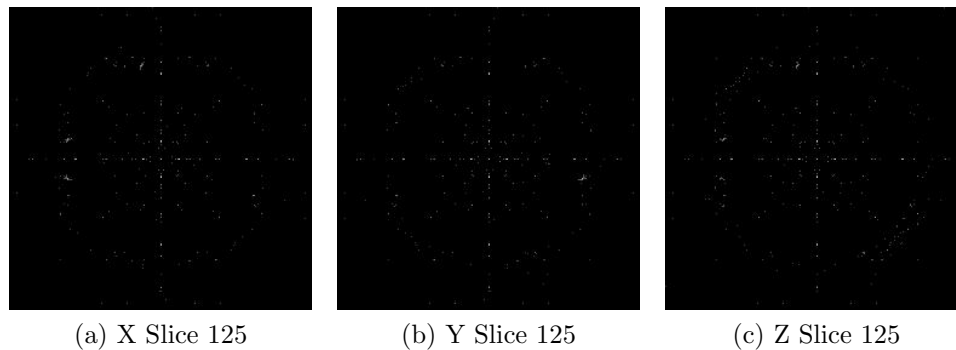
| (a) X Slice 125 | (b) Y Slice 125 | (c) Z Slice 125 |

Figure 8: Maximum variance slices in the three dimensions of the parts of the background beyond 5*sigma

Automatic criteria: The validation is OK if 1) the p-value of the null hypothesis that the background has 0 mean is larger than 0.001; and 2) the number of voxels above or below 5 sigma is smaller than 20 times the amount expected for a Gaussian with the same standard deviation whose mean is 0.

WARNINGS: 2 warnings

1. **The null hypothesis that the background mean is 0 has been rejected because the p-value of the comparison is smaller than 0.001**
2. **There is a significant proportion of outlier values in the background (cdf5 ratio=2031.06)**

2.4   Level 0.d B-factor analysis

Explanation:
The B-factor line [Rosenthal and Henderson, 2003] fitted between 15Åand the resolution reported should have a slope that is between 0 and 300 Å^2.

Results:
Fig. 9 shows the logarithm (in natural units) of the structure factor (the module squared of the Fourier transform) of the experimental map, its fitted line, and the corrected map. The estimated B-factor was -93.3. The fitted line was $\log(|F|^2) = -23.3/R^2 + (-13.2)$.

Figure 9: Guinier plot. The X-axis is the square of the inverse of the resolution in Å.

17

| (a) X Slice 128 | (b) Y Slice 128 | (c) Z Slice 122 |

Figure 10: Slices of maximum variation in the three dimensions of the B-factor corrected map

Automatic criteria: The validation is OK if the B-factor is in the range [-300,0].

STATUS: OK

2.5 Level 0.e Local resolution with DeepRes

Explanation:
DeepRes [Ramírez-Aportela et al., 2019] measures the local resolution using a neural network that has been trained on the appearance of atomic structures at different resolutions. Then, by comparing the local appearance of the input map to the appearance of the atomic structures a local resolution label can be assigned.

Results:

Fig. 11 shows the histogram of the local resolution according to DeepRes. Some representative percentiles are:

Percentile	Resolution(Å)
2.5%	3.41
25%	3.84
50%	4.09
75%	4.35
97.5%	4.90

The reported resolution, 2.60 Å, is at the percentile 0.0. Fig. 12 shows some representative views of the local resolution.
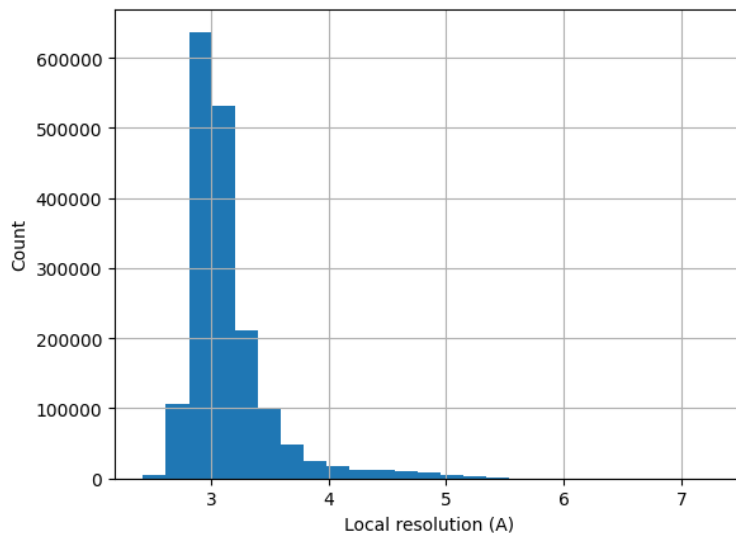

Figure 11: Histogram of the local resolution according to deepres.
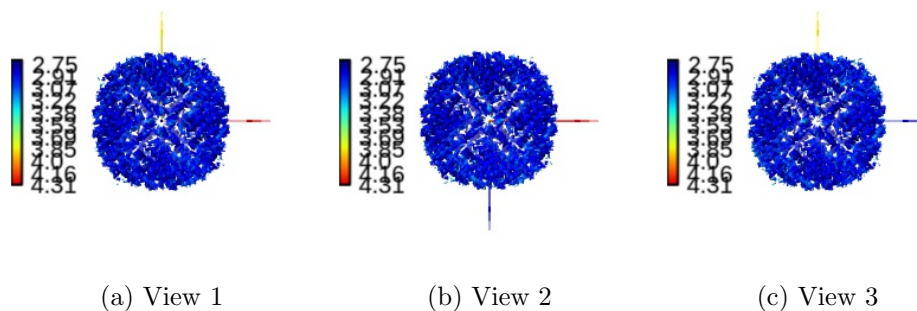
(a) View 1 (b) View 2 (c) View 3

Figure 12: Local resolution according to DeepRes. Views generated by ChimeraX at a the following X, Y, Z angles: View 1 (0,0,0), View 2 (90, 0, 0), View 3 (0, 90, 0).

Automatic criteria: The validation is OK if the percentile of the user provided resolution is larger than 0.1% of the percentile of the local resolution as estimated by DeepRes.

WARNINGS: 1 warnings

1. **The reported resolution, 2.60 Å, is particularly with respect to the local resolution distribution. It occupies the 0.00 percentile**

2.6 Level 0.f Local B-factor

Explanation:
LocBfactor [Kaur et al., 2021] estimates a local resolution B-factor by decomposing the input map into a local magnitude and phase term using the spiral transform.

Results:

Fig. 13 shows the histogram of the local B-factor according to LocBfactor. Some representative percentiles are:

Percentile	Local B-factor (Å^{-2})
2.5%	-184.43
25%	-159.09
50%	-146.41
75%	-133.15
97.5%	-106.90

Fig. 14 shows some representative views of the local B-factor.

Figure 13: Histogram of the local B-factor according to LocBfactor.

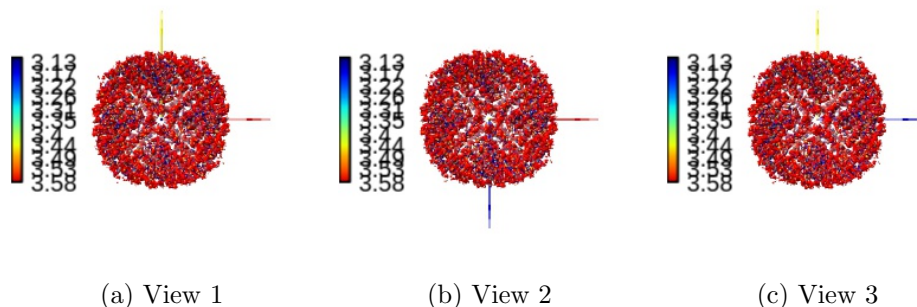|		
(a) View 1	(b) View 2	(c) View 3

Figure 14: Local B-factor according to LocBfactor. Views generated by ChimeraX at a the following X, Y, Z angles: View 1 (0,0,0), View 2 (90, 0, 0), View 3 (0, 90, 0).

Automatic criteria: The validation is OK if the median B-factor is in the range [-300,0].

STATUS: OK

2.7 Level 0.g Local Occupancy

Explanation:
LocOccupancy [Kaur et al., 2021] estimates the occupancy of a voxel by the macromolecule.

Results:

Fig. 15 shows the histogram of the local occupancy according to LocOccupancy. Some representative percentiles are:

Percentile	Local Occupancy [0-1]
2.5%	0.08
25%	0.58
50%	0.83
75%	1.00
97.5%	1.00

Fig. 16 shows some representative views of the local occupancy.

Figure 15: Histogram of the local occupancy according to LocOccupancy.

| (a) View 1 | (b) View 2 | (c) View 3 |

Figure 16: Local occupancy according to LocOccupancy. Views generated by ChimeraX at a the following X, Y, Z angles: View 1 (0,0,0), View 2 (90, 0, 0), View 3 (0, 90, 0).

Automatic criteria: The validation is OK if the median occupancy is larger than 50%.

23

STATUS: OK

2.8 Level 0.h Hand correction

Explanation:
Deep Hand determines the correction of the hand for those maps with a resolution smaller than 5Å. The method calculates a value between 0 (correct hand) and 1 (incorrect hand) using a neural network to assign its hand.

Results:

Deep hand assigns a score of 0.085 to the input volume.
Automatic criteria: The validation is OK if the deep hand score is smaller than 0.5.

STATUS: OK

3 Half maps

Half map 1: /home/coss/ScipionUserData/projects/Example_10248_Scipion3/-Runs/010450_XmippProtReconstructHighRes/extra/Iter001/volume01.vol
SHA256 hash: a8d09c9ee945f8eeae5704ed042fceba0de6d128d80dab51ea23f2c046a91993

Half map 2: /home/coss/ScipionUserData/projects/Example_10248_Scipion3/-Runs/010450_XmippProtReconstructHighRes/extra/Iter001/volume02.vol
SHA256 hash: a93a1717f03b3c901f4800f1e0fe7fe3824aa71b0abc131dcf8897aa863eb6e7

Slices of the first half map can be seen in Fig. 17.
Slices of the second half map can be seen in Fig. 18.
Slices of the difference between both maps can be seen in Fig. 19. There should not be any structure in this difference. Sometimes some patterns are seen if the map is symmetric.

(a) X Slice 125   (b) Y Slice 125   (c) Z Slice 125

Figure 17: Slices of maximum variation in the three dimensions of Half 1

(a) X Slice 125   (b) Y Slice 125   (c) Z Slice 125

Figure 18: Slices of maximum variation in the three dimensions of Half 2

(a) X Slice 125   (b) Y Slice 125   (c) Z Slice 125

Figure 19: Slices of maximum variation in the three dimensions of the difference Half1-Half2.

4 Level 1 analysis

4.1 Level 1.a Global resolution

Explanation: The Fourier Shell Correlation (FSC) between the two half maps is the most standard method to determine the global resolution of a map. However, other measures exist such as the Spectral Signal-to-Noise Ratio and the Differential Phase Residual. There is a long debate about the right thresholds for these measures. Probably, the most clear threshold is the one of the SSNR (SSNR=1). For the DPR we have chosen 103.9° and for the FSC, the standard 0.143. For a deep discussion of all these thresholds, see [Sorzano et al., 2017]. Note that these thresholds typically result in resolution values that are at the lower extreme of the local resolution range, meaning that this resolution is normally in the first quarter. It should not be understood as the average resolution of the map.

Except for the noise, the FSC and DPR should be approximately monotonic. They should not have any "coming back" behavior. If they have, this is typically due to the presence of a mask in real space or non-linear processing.

Results:
Fig. 20 shows the FSC and the 0.143 threshold. The resolution according to the FSC is 3.09Å. The map information is well preserved (FSC>0.9) up to

4.75Å.

Fig. 21 shows the DPR and the 103.9° threshold. The resolution according to the DPR is 2.69Å.

Fig. 22 shows the SSNR and the SSNR=1 threshold. The resolution according to the SSNR is 2.92Å.

The mean resolution between the three methods is 2.90Å and its range is within the interval [2.69, 3.09]Å.
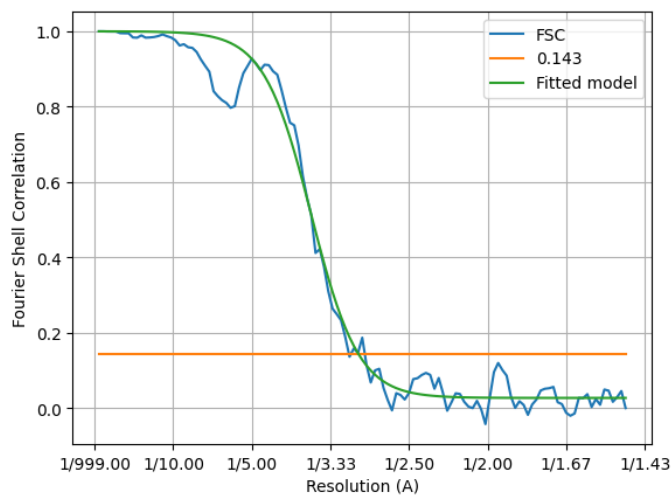
Figure 20: Fourier Shell correlation between the two halves.
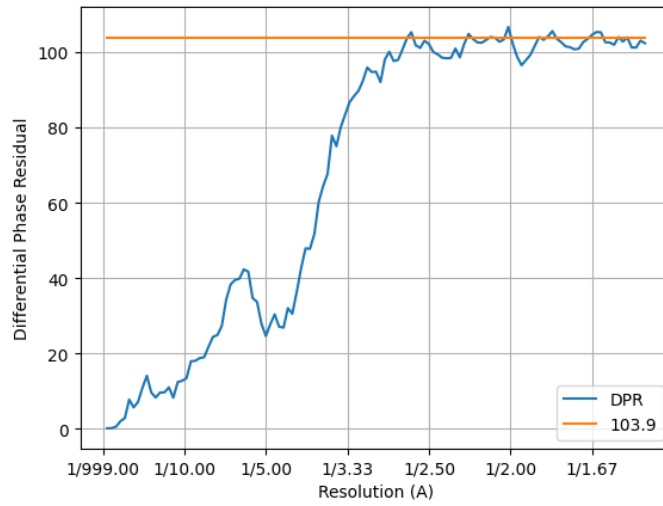
27

Figure 21: Differential Phase Residual between the two halves.

Figure 22: Spectral Signal-to-Noise Ratio estimated from the two halves.

Automatic criteria: The validation is OK if the user provided resolution is larger than 0.8 times the resolution estimated by 1) FSC, 2) DPR, and 3) SSNR.

28

4.2   Level 1.b FSC permutation

Explanation:
This method [Beckers and Sachse, 2020] calculates a global resolution by formulating a hypothesis test in which the distribution of the FSC of noise is calculated from the two maps.

Results:

The resolution at 1% of FDR was 2.7. The estimated B-factor was -85.7. Fig. 23 shows the estimated FSC and resolution.
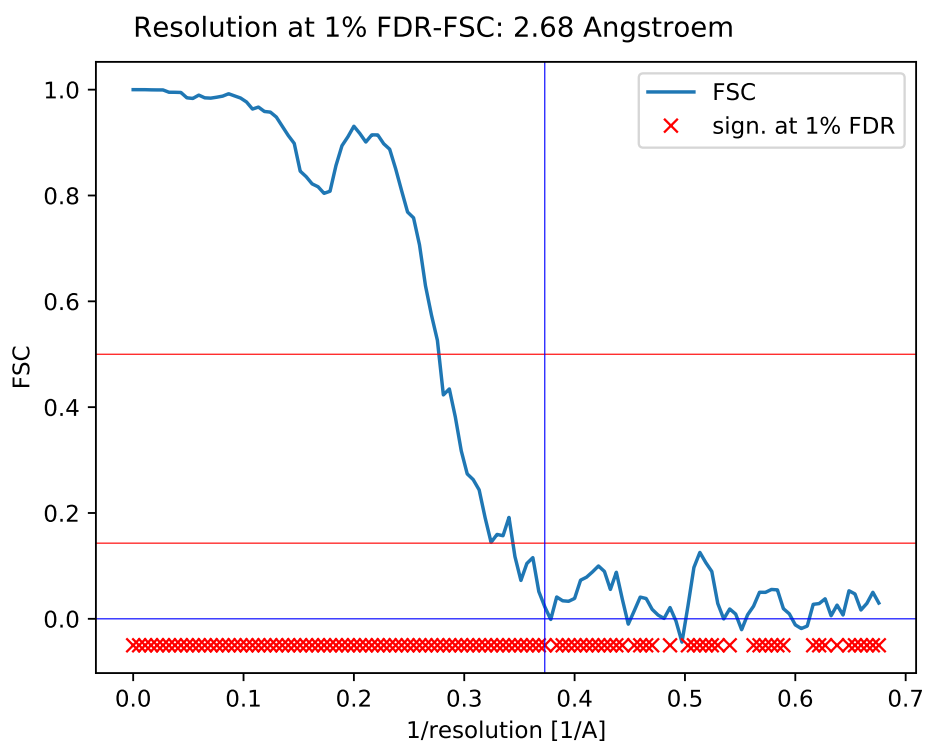
Resolution at 1% FDR-FSC: 2.68 Angstroem

Figure 23: FSC and resolution estimated by a permutation test.

Automatic criteria: The validation is OK if the user provided resolution is larger than 0.8 times the resolution estimated by FSC permutation.

STATUS: OK

4.3 Level 1.c Local resolution with Blocres

Explanation:
This method [Cardone et al., 2013] computes a local Fourier Shell Correlation (FSC) between the two half maps.

Results:

Fig. 24 shows the histogram of the local resolution according to Blocres. Some representative percentiles are:

Percentile	Resolution(Å)
2.5%	2.75
25%	2.91
50%	3.04
75%	3.21
97.5%	4.31

The reported resolution, 2.60 Å, is at the percentile 0.2. Fig. 25 shows some representative views of the local resolution.

Figure 24: Histogram of the local resolution according to blocres.

(a) View 1   (b) View 2   (c) View 3

Figure 25: Local resolution according to Blocres. Views generated by ChimeraX at a the following X, Y, Z angles: View 1 (0,0,0), View 2 (90, 0, 0), View 3 (0, 90, 0).

Automatic criteria: The validation is OK if the percentile of the user provided resolution is larger than 0.1% of the percentile of the local resolution as estimated by BlocRes.

STATUS: OK

31

4.4   Level 1.d Local resolution with Resmap

Explanation:
This method [Kucukelbir et al., 2014] is based on a test hypothesis testing of the superiority of signal over noise at different frequencies.

Results:

Fig. 26 shows the histogram of the local resolution according to Resmap. Some representative percentiles are:

Percentile	Resolution(Å)
2.5%	3.13
25%	3.45
50%	3.52
75%	3.55
97.5%	3.58

The reported resolution, 2.60 Å, is at the percentile 0. Fig. 27 shows some representative views of the local resolution.
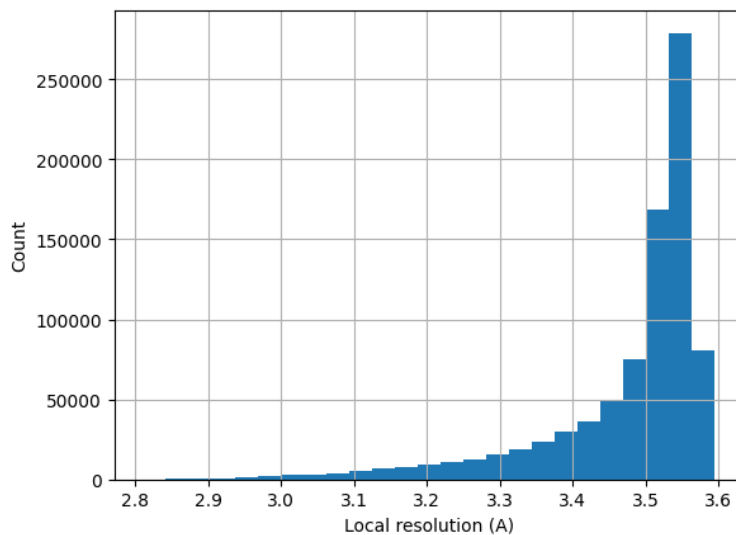
Figure 26: Histogram of the local resolution according to Resmap.

(a) View 1   (b) View 2   (c) View 3

Figure 27: Local resolution according to Resmap. Views generated by ChimeraX at a the following X, Y, Z angles: View 1 (0,0,0), View 2 (90, 0, 0), View 3 (0, 90, 0).

Automatic criteria: The validation is OK if the percentile of the user provided resolution is larger than 0.1% of the percentile of the local resolution as estimated by Resmap.

WARNINGS: 1 warnings

1. **The reported resolution, 2.60 Å, is particularly with respect to the local resolution distribution. It occupies the 0.00 percentile**

4.5 Level 1.e Local resolution with MonoRes

Explanation:
MonoRes [Vilas et al., 2018] evaluates the local energy of a point with respect to the distribution of energy in the noise. This comparison is performed at multiple frequencies and for each one, the monogenic transformation separates the amplitude and phase of the input map. Then the energy of the amplitude within the map is compared to the amplitude distribution observed in the noise, and a hypothesis test is run for every voxel to check if its energy is signficantly above the level of noise.

Results:

Fig. 28 shows the histogram of the local resolution according to MonoRes. Some representative percentiles are:

Percentile	Resolution(Å)
2.5%	1.74
25%	3.97
50%	4.58
75%	6.39
97.5%	9.08

The reported resolution, 2.60 Å, is at the percentile 5.6. Fig. 29 shows some representative views of the local resolution
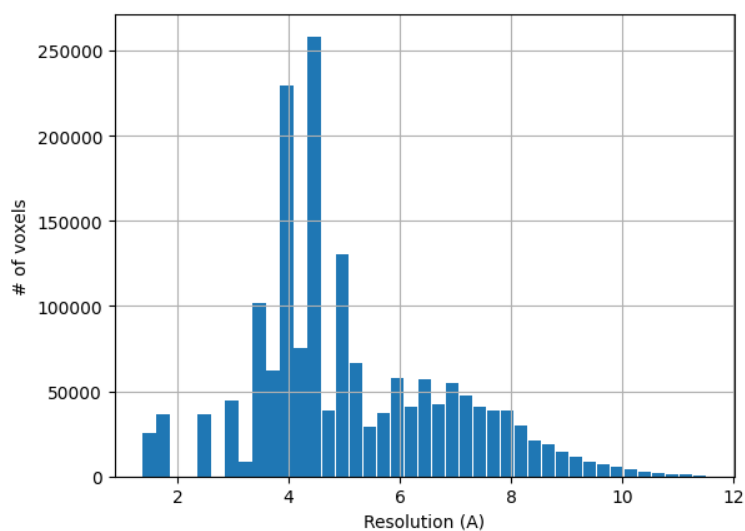
Figure 28: Histogram of the local resolution according to MonoRes.

(a) View 1   (b) View 2   (c) View 3
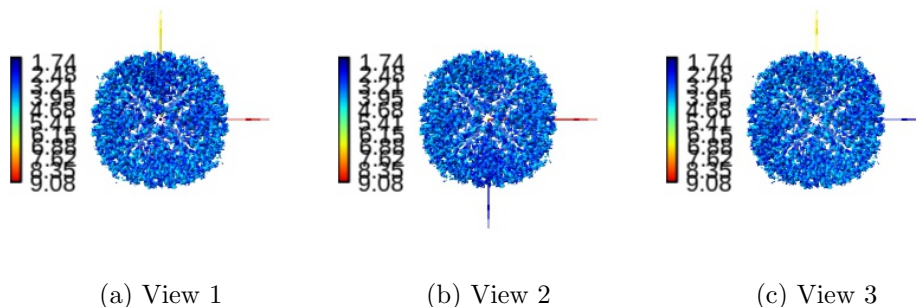
Figure 29: Local resolution according to Monores. Views generated by ChimeraX at a the following X, Y, Z angles: View 1 (0,0,0), View 2 (90, 0, 0), View 3 (0, 90, 0).

Automatic criteria: The validation is OK if the percentile of the user provided resolution is larger than 0.1% of the percentile of the local resolution as estimated by MonoRes.

STATUS: OK

4.6 Level 1.f Local and directional resolution with MonoDir

Explanation:
MonoDir [Vilas et al., 2020] extends the concept of local resolution to local and directional resolution by changing the shape of the filter applied to the input map. The directional analysis can reveal image alignment problems.

The histogram of best resolution voxels per direction (Directional Histogram 1D) shows how many voxels in the volume have their maximum resolution in that direction. Directions are arbitrarily numbered from 1 to N. This histogram should be relatively flat. We perform a Kolmogorov-Smirnov test to check its uniformity. If the null hypothesis is rejected, then the directional resolution is not uniform. It does not mean that it is wrong, and it could be caused by several reasons: 1) the angular distribution is not uniform, 2) there are missing directions, 3) there is some anisotropy in the data (including some preferential directional movement).

Ideally, the radial average of the minimum, maximum, and average res-

olution at each voxel (note that these are spatial radial averages) should be flat and as low as possible. If they show some slope, this is associated with inaccuracies in the angular assignment. These averages make sense when the shells are fully contained within the protein. As the shells approach the outside of the protein, these radial averages make less sense.

Results:

Fig. 30 shows the 1D directional histogram and Fig. 31 the 2D directional histogram. We compared the 1D directional histogram to a uniform distribution using a Kolmogorov-Smirnov test. The D statistic was 0.064887, and the p-value of the null hypothesis 0.000000.

The radial average of the minimum, maximum and average resolution at each voxel is shown in Fig. 32. The overall mean of the directional resolution is 2.03
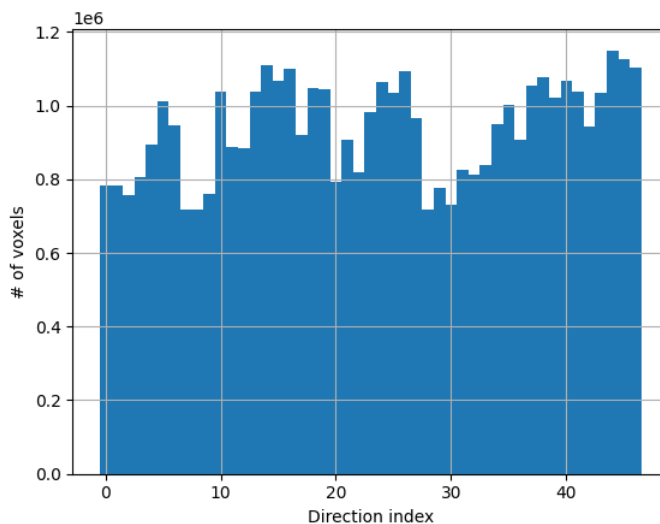
Figure 30: Histogram 1D of the best direction at each voxel.
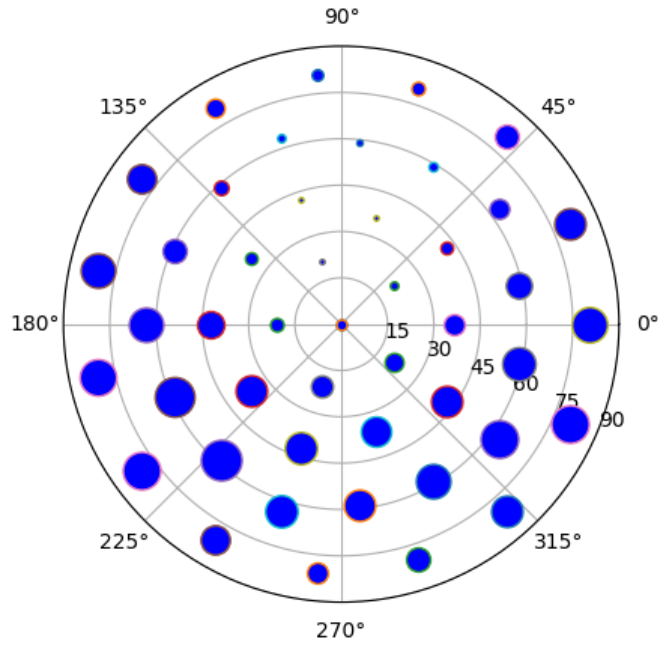
Figure 31: Histogram 2D of the best direction at each voxel. The azimuthal rotation is circular, while the tilt angle is the radius. The size of the point is proportional to the number of voxels whose maximum resolution is in that direction (this count can be seen in Fig. 30.
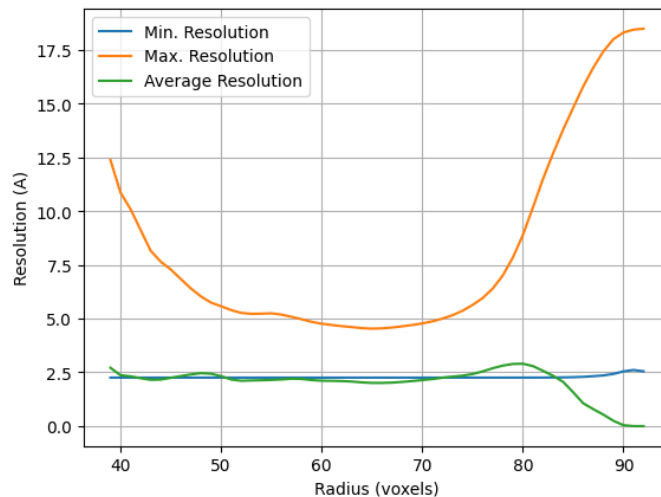
Figure 32: Radial averages (in space) of the minimum, maximum and average resolution at each voxel.

Automatic criteria: The validation is OK if 1) the null hypothesis that the directional resolution is not uniform is not rejected with a threshold of 0.001 for the p-value, and 2) the resolution provided by the user is not smaller than 0.8 times the average directional resolution.

WARNINGS: 1 warnings

1. **The distribution of best resolution is not uniform in all directions. The associated p-value is 0.000000.**

4.7 Level 1.g Fourier Shell Occupancy

Explanation:
This method calculates the anisotropy of the energy distribution in Fourier shells. This is an indirect measure of anisotropy of the angular distribution or the presence of heterogeneity. A natural threshold for this measure is 0.5. However, 0.9 and 0.1 are also interesting values that define the frequency at which the occupancy is 90% and 10%, respectively. This region is shaded in the plot.
Results:

Fig. 33 shows the Fourier Shell Occupancy and its anisotropy. The directional resolution is shown in Fig. 34. The resolution according to the FSO is 2.85Å. Fourier shells are occupied at between 90 and than 10% in the range [2.91, 2.27]Å.
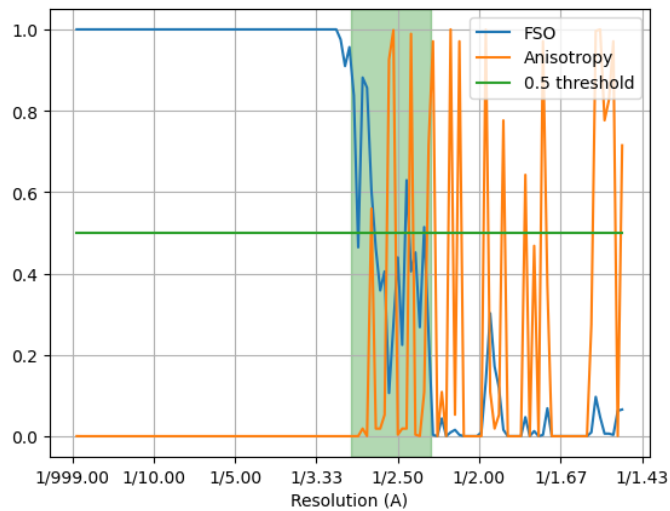
Figure 33: FSO and anisotropy.
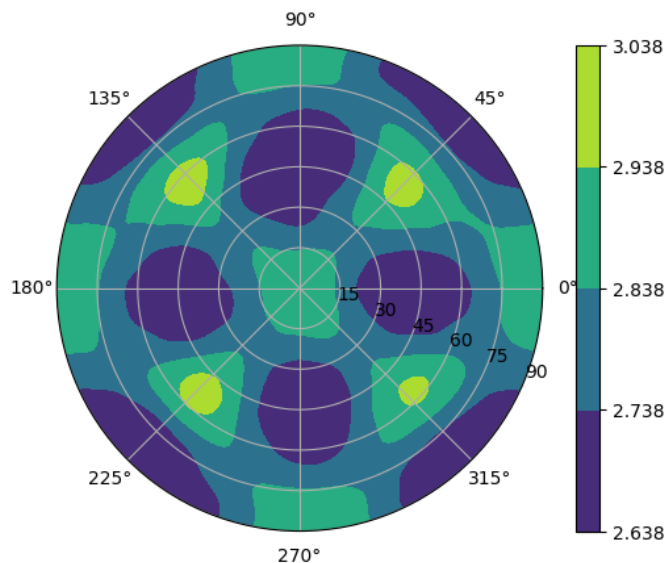
Figure 34: Directional resolution in the projection sphere.

Automatic criteria: The validation is OK if the resolution provided by the user is not smaller than 0.8 times the resolution estimated by the first cross of FSO below 0.5.

STATUS: OK

4.8 Level 1.h Fourier Shell Correlation 3D

Explanation:
This method analyzes the FSC in different directions and evaluates its homogeneity.
Results:

Fig. 35 shows the FSCs in X, Y, Z, and the global FSC. Fig. 36 shows the global FSC and the histogram of the directional FSC. Finally, Fig. 37 shows the rotational average of the map power in Fourier space. The FSC 3D resolutions at a 0.143 threshold in X, Y, and Z are 2.64, 2.66, and 2.66 Å, respectively. The global resolution at the same threshold is 2.60 Å. The resolution range is [2.60, 2.66]Å.
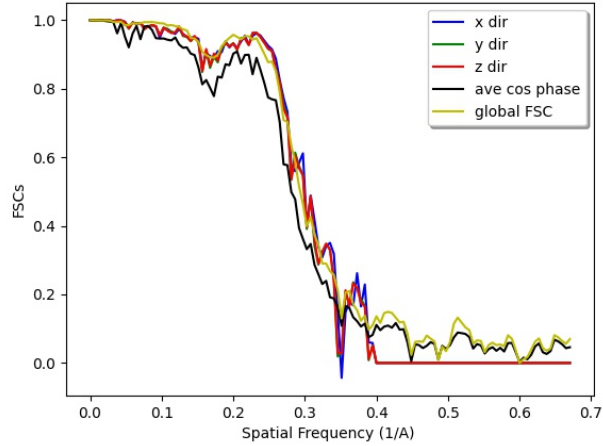
Figure 35: FSC in X, Y, Z, the global FSC, and the Average Cosine Phase.
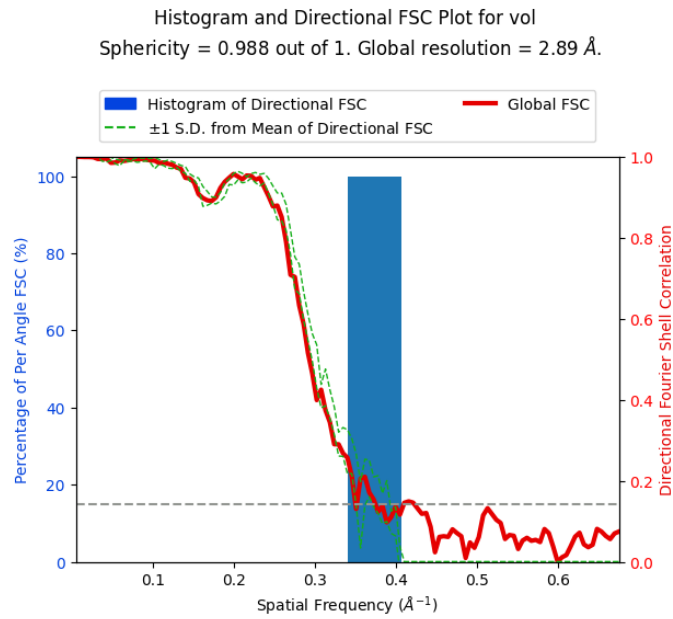
Figure 36: Global FSC and histogram of the directional FSC.
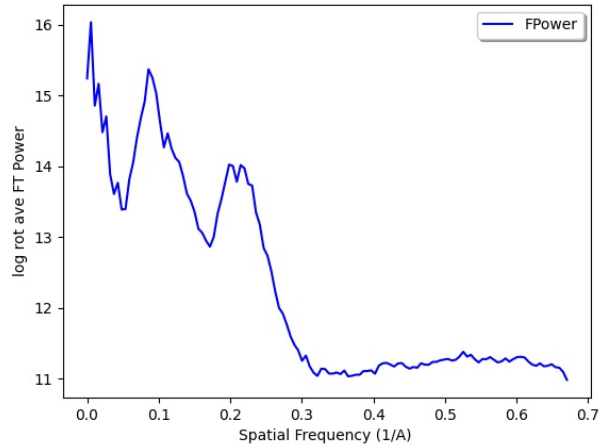
Figure 37: Logarithm of the radial average of the input map power in Fourier space.

Automatic criteria: The validation is OK if the resolution provided by the user is not smaller than 0.8 the resolution estimated by the first cross of the global directional FSC below 0.143.

STATUS: OK

5 2D Classes

Set of 2D classes: /home/coss/ScipionUserData/projects/Example_10248_Scipion3/-Runs/012458_XmippProtCropResizeParticles/extra/output_images.stk

The classes can be seen in Fig. 38.
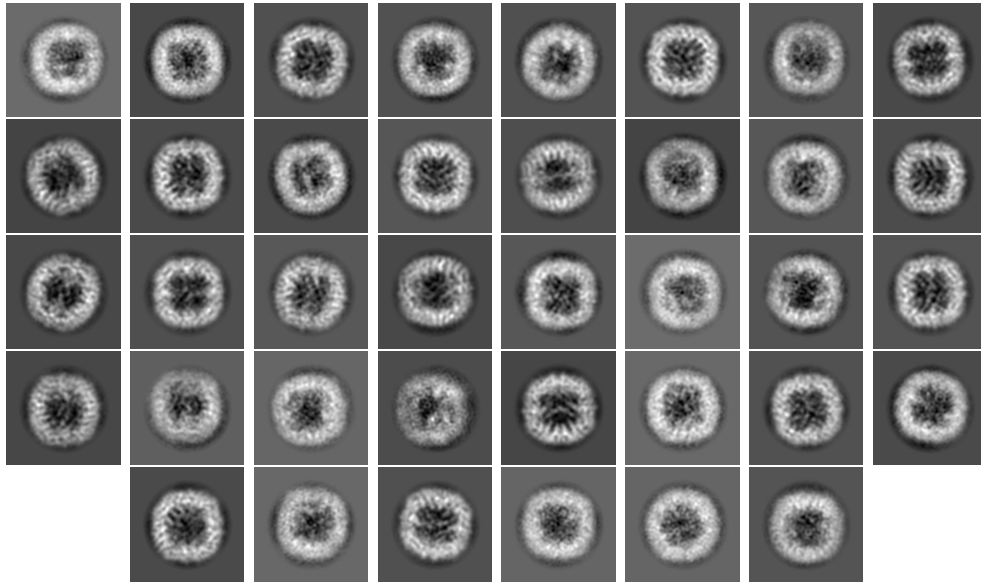
Figure 38: Set of 2D classes provided by the user

6 Level 2 analysis

6.1 Level 2.a Reprojection consistency

Explanation:
The 2D classes can be aligned against the reconstructed map, then the correlation between reprojections of the map and the 2D classes can be analyzed. Also, analyzing the residuals (2D class minus the corresponding reprojection) can reveal systematic differences between them.

Results:

Fig. 39 shows the histogram of the cross-correlation between the 2D classes and the map reprojections. The average correlation is 0.843624, and its range is [0.717260,0.906353]. Now we show the 2D classes, the corresponding reprojection, the difference between both (residual), the covariance matrix of the residual image, and the correlation between the 2D class and the reprojection. For a perfect match, the residual would be just noise, and its covariance matrix should be a diagonal. Rows have been sorted by correlation so that

the worse correlating images are displayed at the beginning.

Also, 2D classes of a high-resolution map should also be of high resolution. This cannot, for the moment, be automatically assessed. But a visual inspection should confirm that the resolution of the 2D classes match the reported resolution of the map.
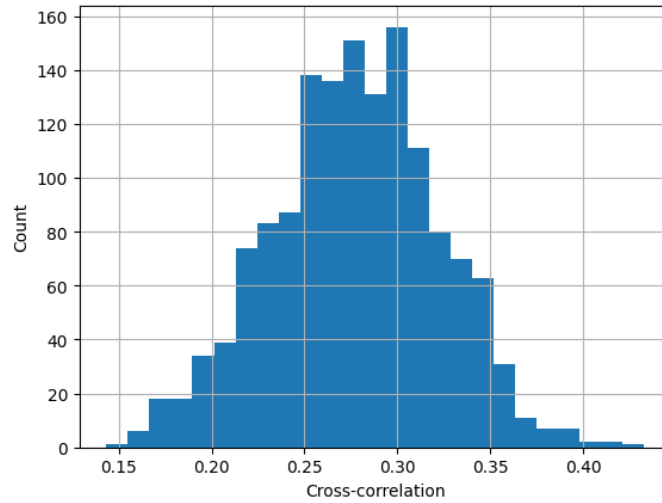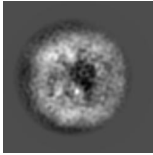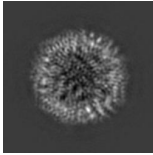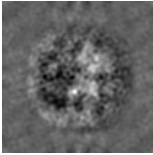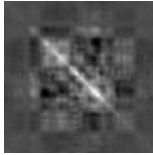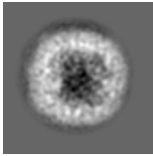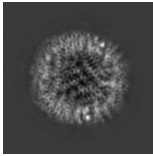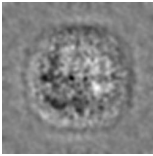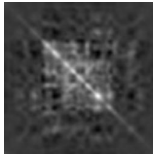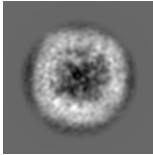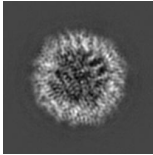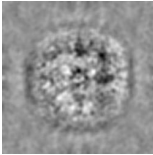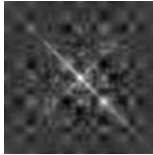
Figure 39: Histogram of the correlation coefficient between the 2D classes provided by the user and the corresponding reprojections.

2D Class	Reprojection	Residual	Covariance	Correlation
				0.72
				0.78
				0.79

44

0.81

0.81

0.81

0.81

0.81

0.82

0.82

0.83

0.83

46

0.88

0.88

0.88

0.89

0.89

0.89

0.89

0.89

0.89

0.90

0.90

0.90

0.90

0.90

0.90

0.90

0.90

0.91
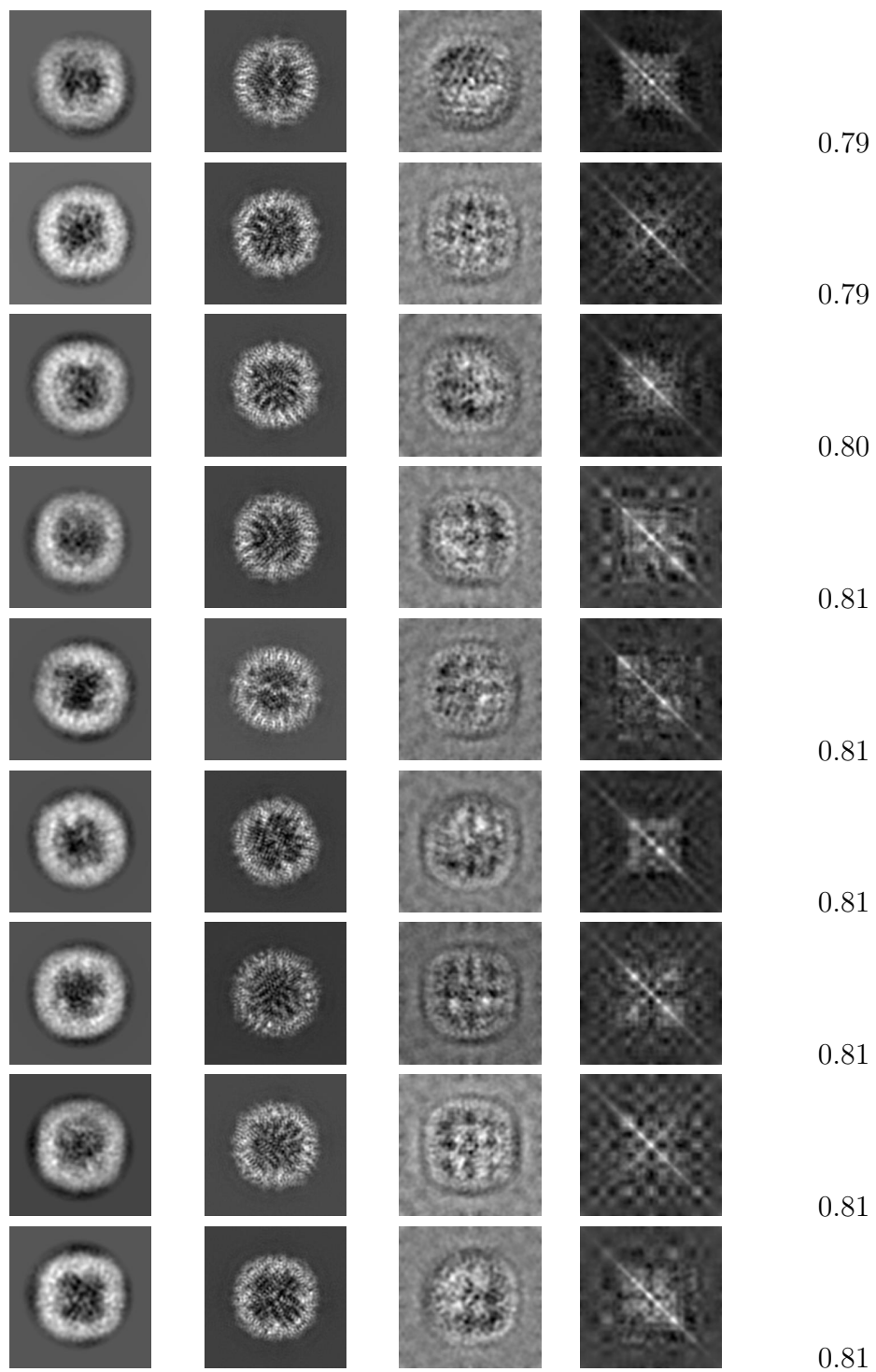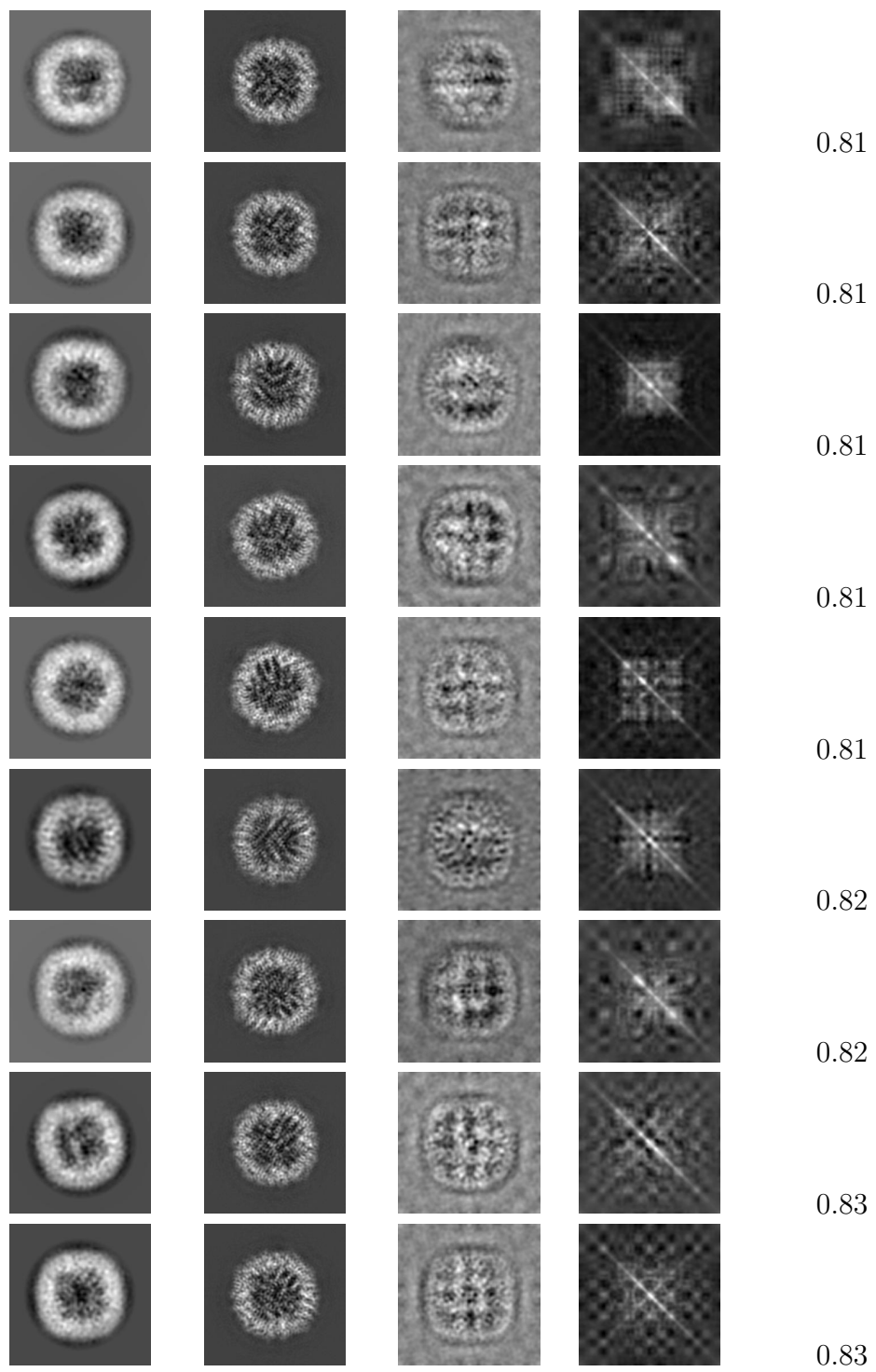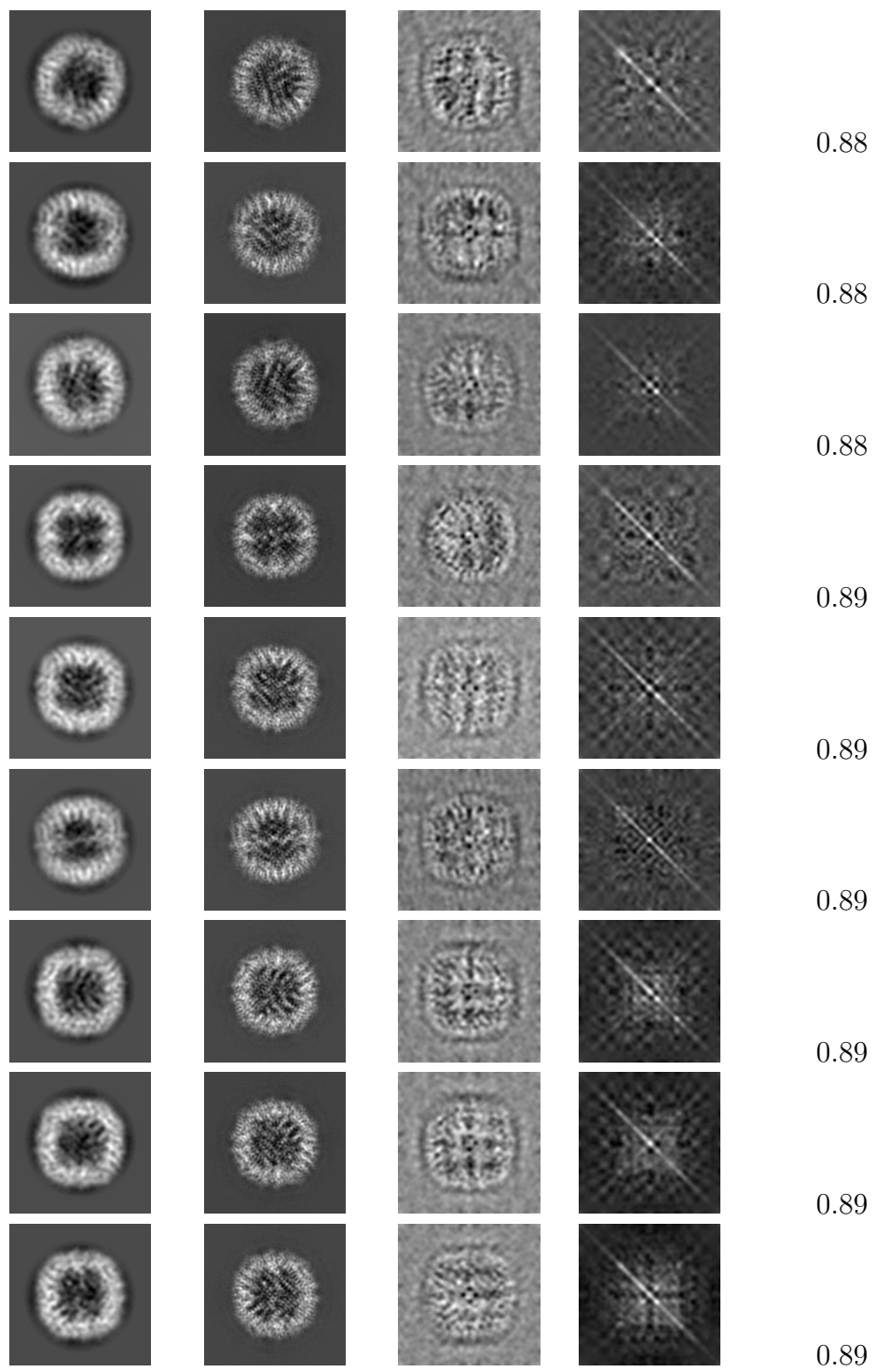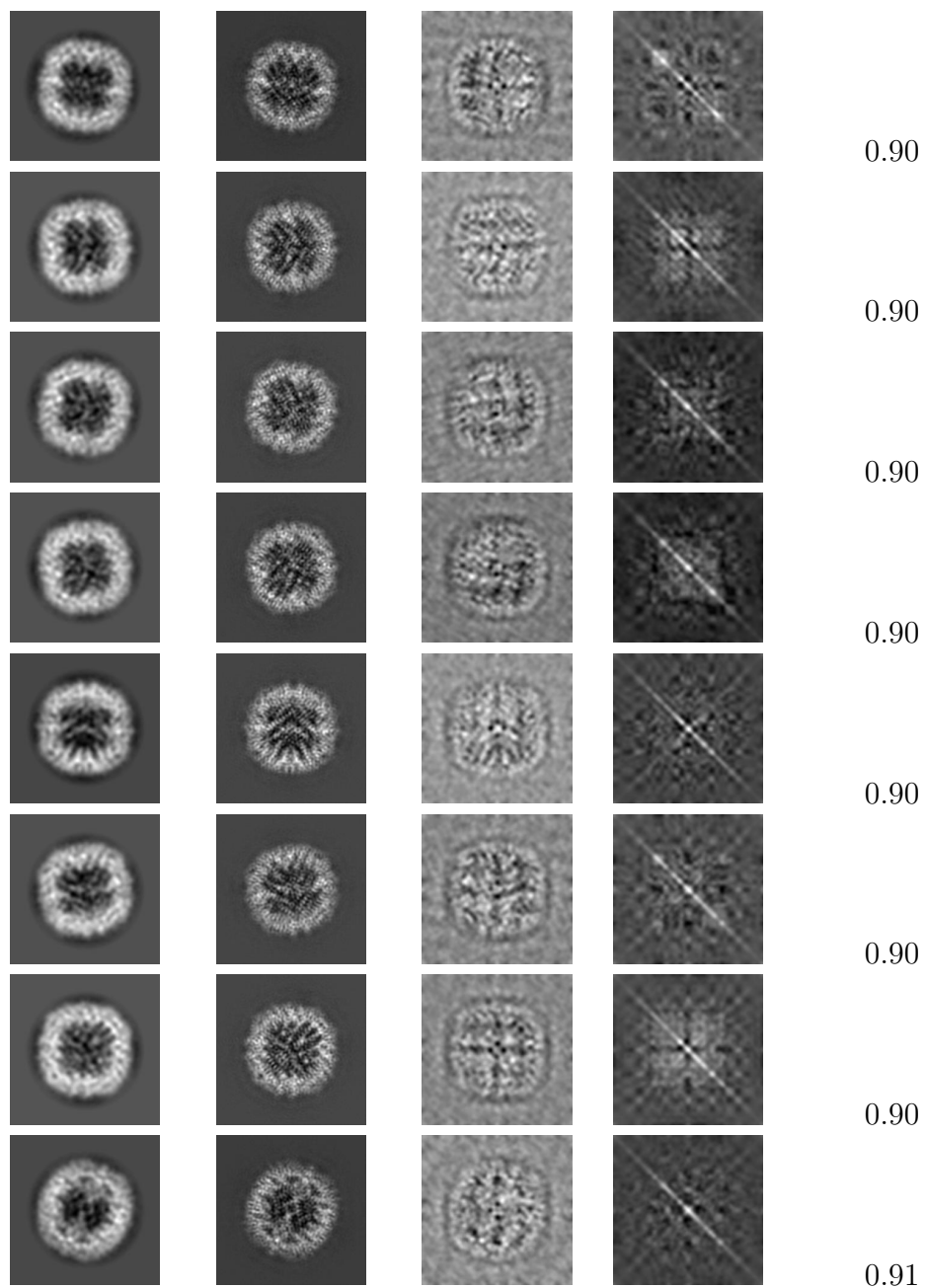
Automatic criteria: The validation is OK if the proportion of classes for which the correlation is below 0.7 is smaller than 20%.

48

7   Particles

Set of Particles: /home/coss/ScipionUserData/projects/Example_10248_Scipion3/-Runs/010450_XmippProtReconstructHighRes/particles.sqlite

1457 images were provided by the user. The first 32 can be seen in Fig. 40.
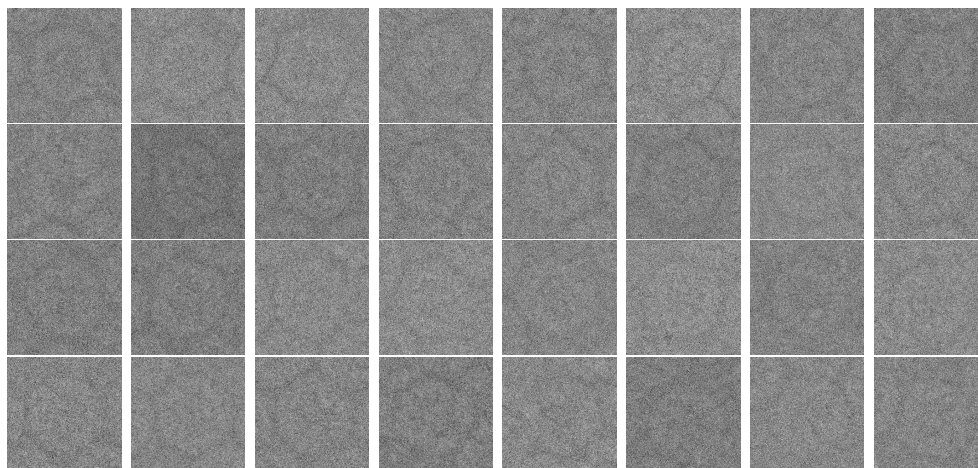
Figure 40: First particles of the set of particles provided by the user

8   Level 3 analysis

This analysis compares the experimental images provided to the 2D classes provided of Level 2.

8.1   Level 3.a Outlier detection
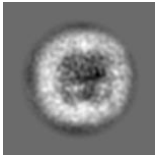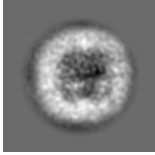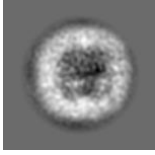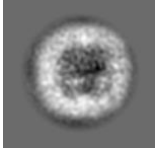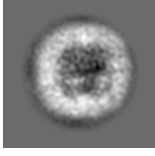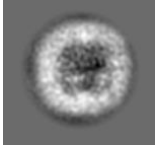
Explanation:
The set of particles is classified into the input set of 2D classes of Level 2. The number of particles that are considered to be outliers in those classes is

reported. A particle is an outlier if its Mahalanobis distance to the centroid of the class is larger than 3 [Sorzano et al., 2014]. This distance takes into account the covariance of the images assigned to that class.

Results:

The following table shows the input classes, the number of particles assigned to them, and the fraction of these particles that are considered to be part of the core (the closer to 1, the better).

2D Class	No. Particles	Core fraction
	17	0.810
	10	0.833
	18	0.857
	35	0.875
	14	0.875
	22	0.880
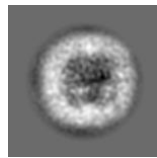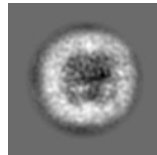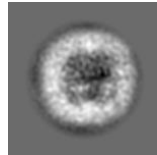
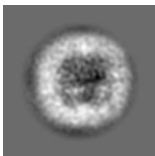|	
23	0.885
24	0.889
25	0.893
37	0.902
28	0.903
19	0.905
48	0.906
125	0.906
49	0.907
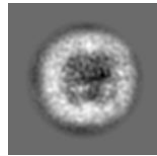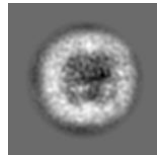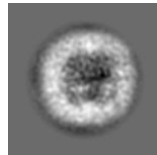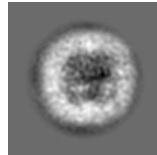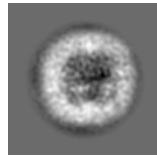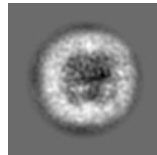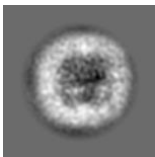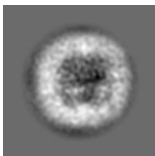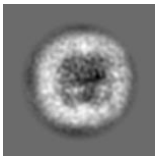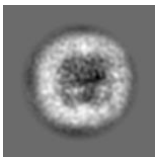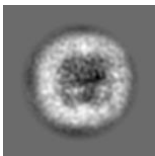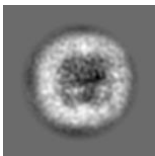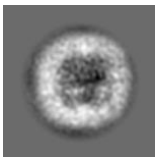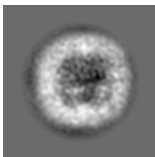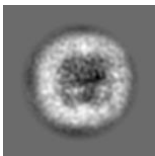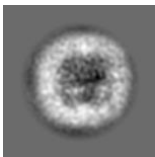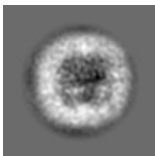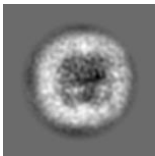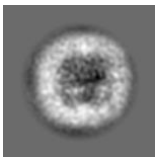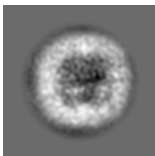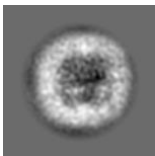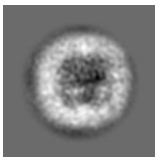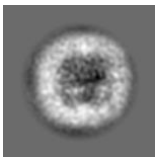
20   0.909

42   0.913

32   0.914

33   0.917

46   0.920

36   0.923

40   0.930

82   0.932

102   0.936

	44	0.936
	64	0.941
	81	0.942
	88	0.946
	65	0.956
	25	0.962
	10	1.000
	5	1.000
	6	1.000

9   1.000

1   1.000

4   1.000

10   1.000

2   1.000
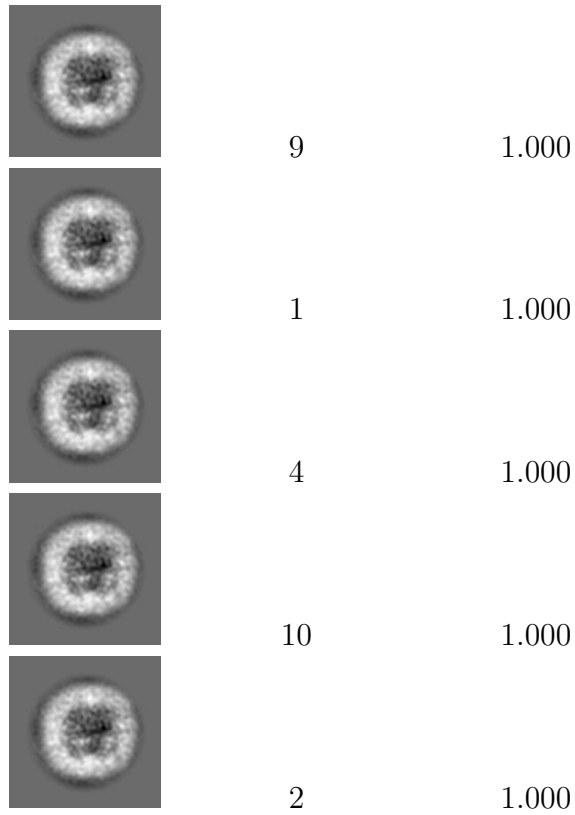
Fig. 41 shows the histogram of the core fraction of the classes. Fig. 42 shows the histogram of the size of the classes.
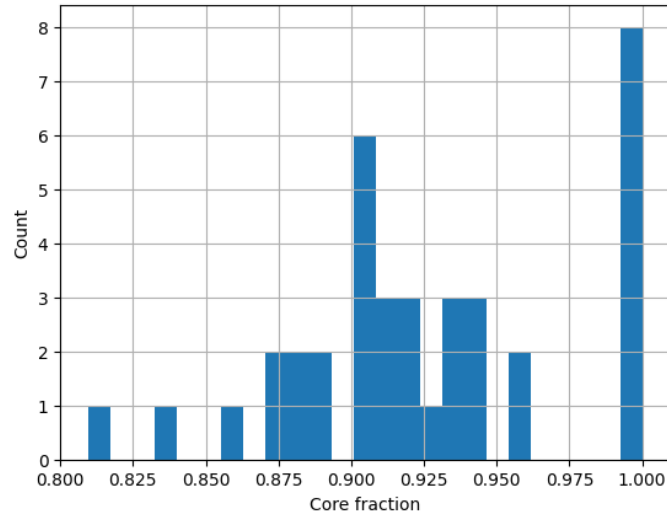
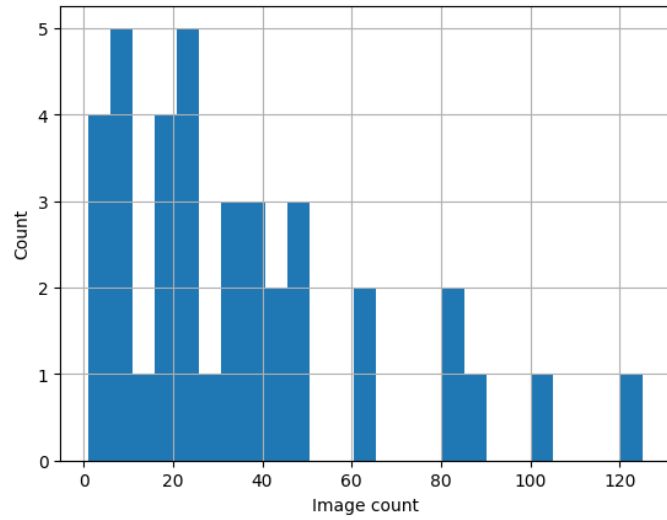Figure 41: Histogram of the core fraction of the 2D classes.

Figure 42: Histogram of the number of particles assigned to the 2D classes.

Automatic criteria: The validation is OK if the number of classes whose core is smaller than 70% of the size of the class is smaller than 20%.

8.2 Level 3.b Classification internal consistency

Explanation:
The input particles are classified in 2D clusters. The quality of the 2D clusters is assessed through Fourier Ring Correlation.

Results:

Fig. 43 shows the histogram of the resolution of each one of the classes. This resolution strongly depends on the number of particles assigned to the class, and this server only sees a small fraction of the particles. Fig. 44 shows a scatter plot of the resolution (in Å^{-1}) in the classes versus the number of particles as measured by FRC=0.5.
The following table shows each class, the number of particles assigned to it, and its resolution as measured by FRC=0.5.
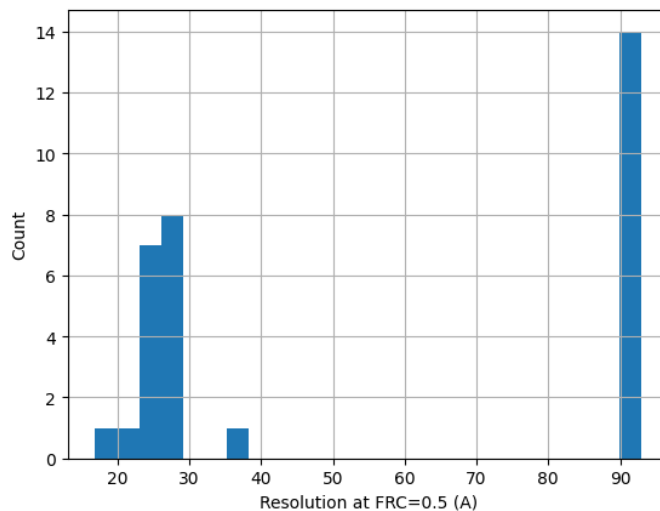

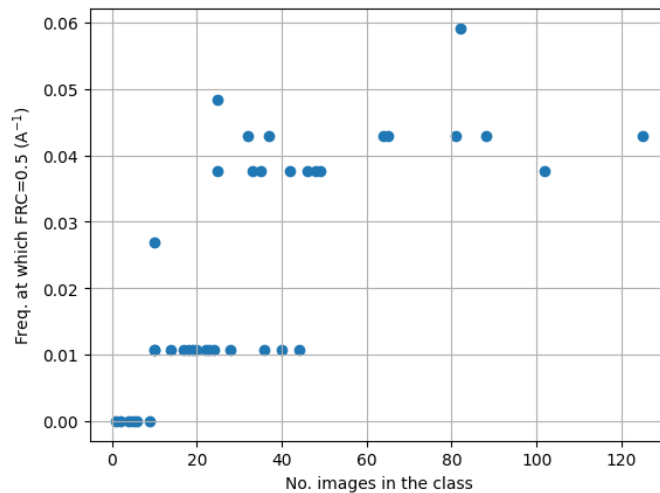Figure 43: Histogram of the resolution at FRC=0.5 of the different classes.

Figure 44: Scatter plot of the frequency at which FRC=0.5 (Å^{-1}) vs the number of particles assigned to each class.

STATUS: Cannot be automatically evaluated

8.3 Level 3.c Classification external consistency

Explanation:
The input particles were classified with CryoSparc [Punjani et al., 2017] using the same number of classes as the ones provided by the user. Except for the difference in number of particles between the original classification and the number of particles available to the server, the new classes should resemble the old ones.

Results:

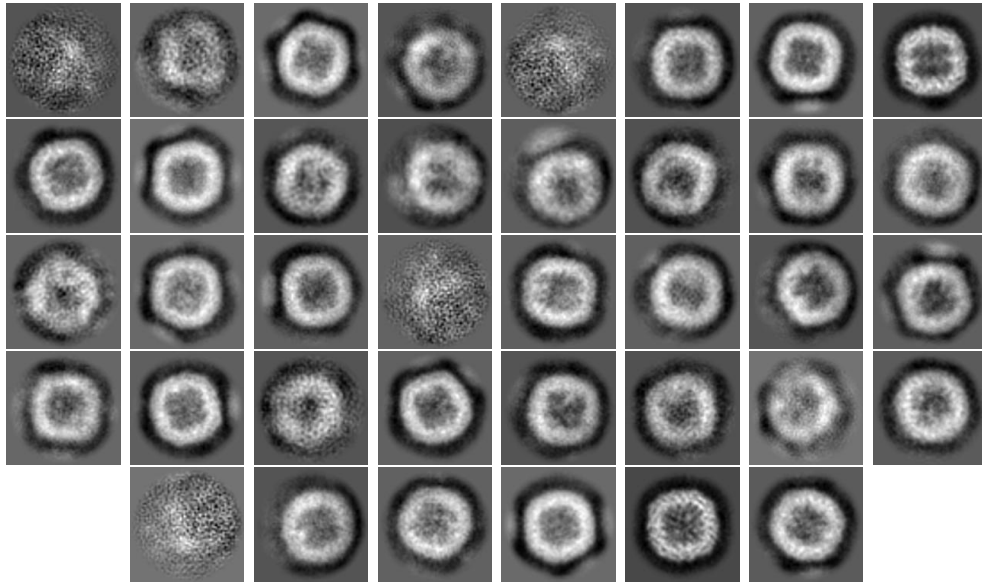Fig. 45 shows the new classification. The classification provided by the user is in Fig. 38.

Figure 45: Set of 2D classes calculated by CryoSparc. These should be compared to those in Fig. 38.

Fig. 46 shows the probability density function of the correlation of the user classes compared to the newly computed and vice versa. Ideally, these two distributions should be similar. We compared these two distributions with a Kolmogorov-Smirnov (KS) two-sample test. The KS statistic was 0.236842 and the p-value 0.238941.
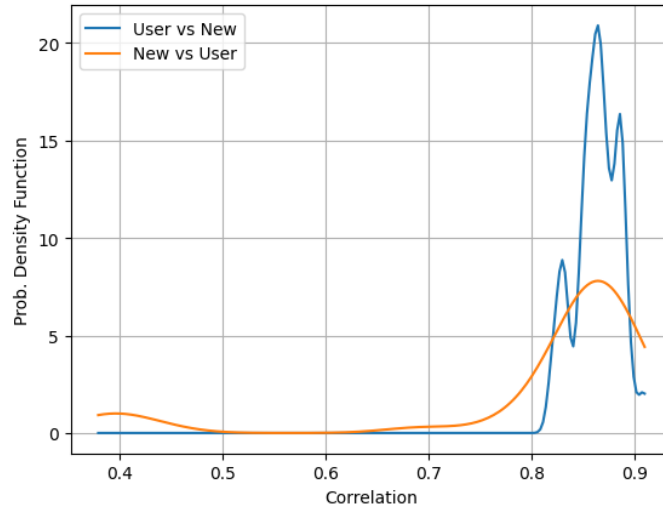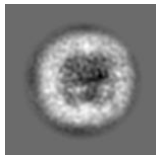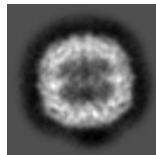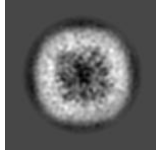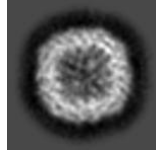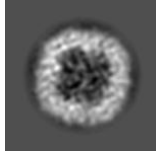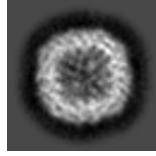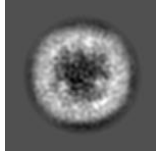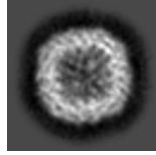
Figure 46: Probability density function of the correlation of the user classes compared to the newly computed classes and vice versa.

The following table shows for each class in the User set which is the best match in the New set and its correlation coefficient.

User class	New class	Correlation
		0.856
		0.887
		0.853
		0.866

0.873

0.909

0.851

0.875

0.862

0.886

0.888

0.865

0.868

0.885

0.875

0.867

0.848

0.846

0.866

0.851

0.863

0.858

0.830

0.885

0.863

0.820

0.830

0.831

0.886

0.833

0.875

0.887

0.882

0.823

0.896

0.856

0.856

0.865
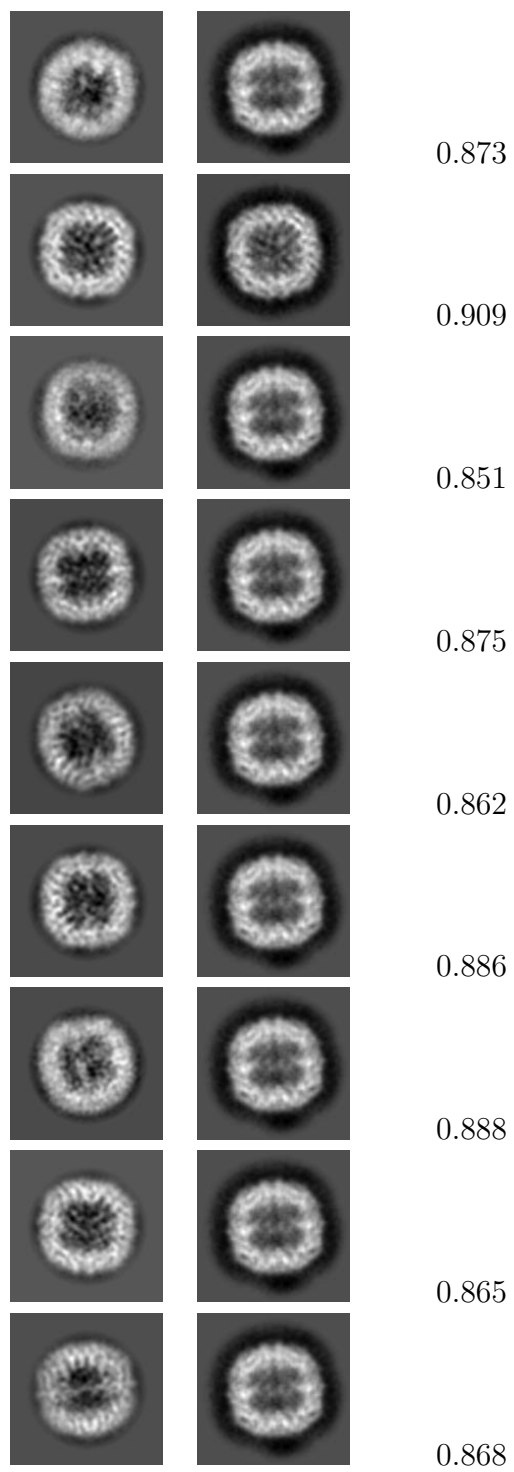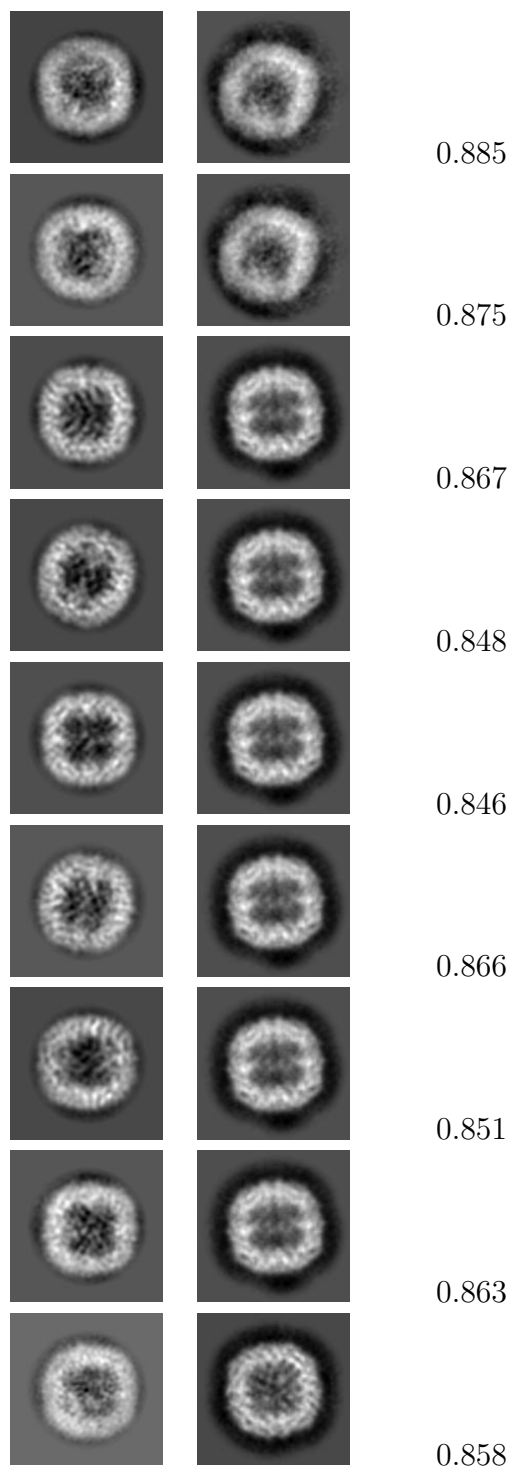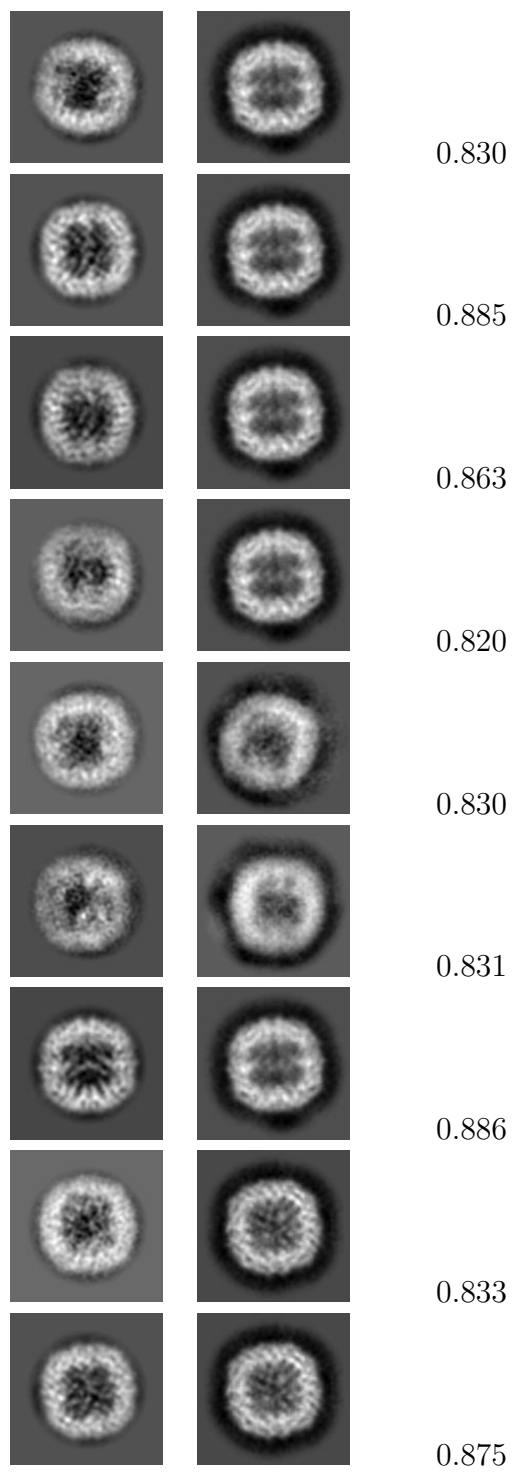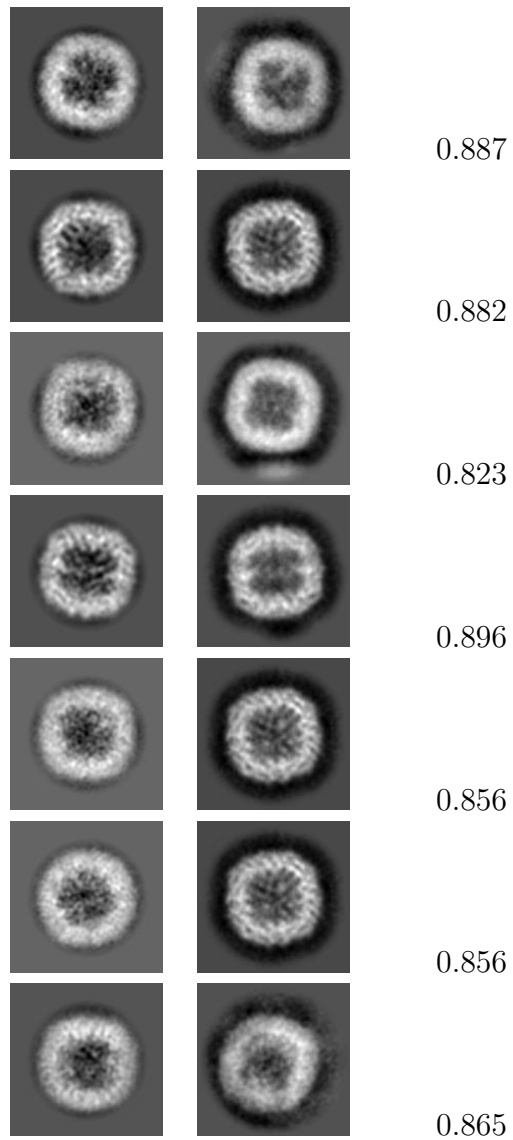
Automatic criteria: The validation is OK if 1) no class from the user correlates less than 0.8 with the newly computed classes, and 2) the equality of the two correlation distributions (user vs new, new vs user) cannot be rejected with a threshold for the p-value of 0.001.

STATUS: OK

9 Level 4 analysis

This analysis compares the experimental images provided along with their angular assignment to the reconstructed map.

9.1 Level 4.a Similarity criteria

Explanation:
We measured the similarity between the experimental images, with the angles and shifts provided by the user, and reprojections of the input map along the same direction. We measured the correlation and IMED (IMage Euclidean Distance, which is a generalized measure of the standard Euclidean Distance in which nearby pixels also contribute to the calculation of the final distance between the image at a given point) [Sorzano et al., 2015] between both sets of images. If the set of particles is properly assigned there should be a single peak in the 1D histograms of these two similarity measures, and a single cloud in their joint scatter plot. The presence of multiple peaks could reveal a mixture of different conformations, the presence of misaligned particles or contaminations, or the difference between isolated particles and particles with other objects around.

It must be noted that there is a dependence between similarity metrics and defocus. Typically this dependence is such that lower defoci imply lower similarity due to the smallest contrast. You have to be sure that the groups seen in the similarity measures are not caused by defocus groups.

Results:

Fig. 47 shows the histograms of the cross-correlation and IMED, a joint scatter plot and the dependence of the cross-correlation with the defocus.
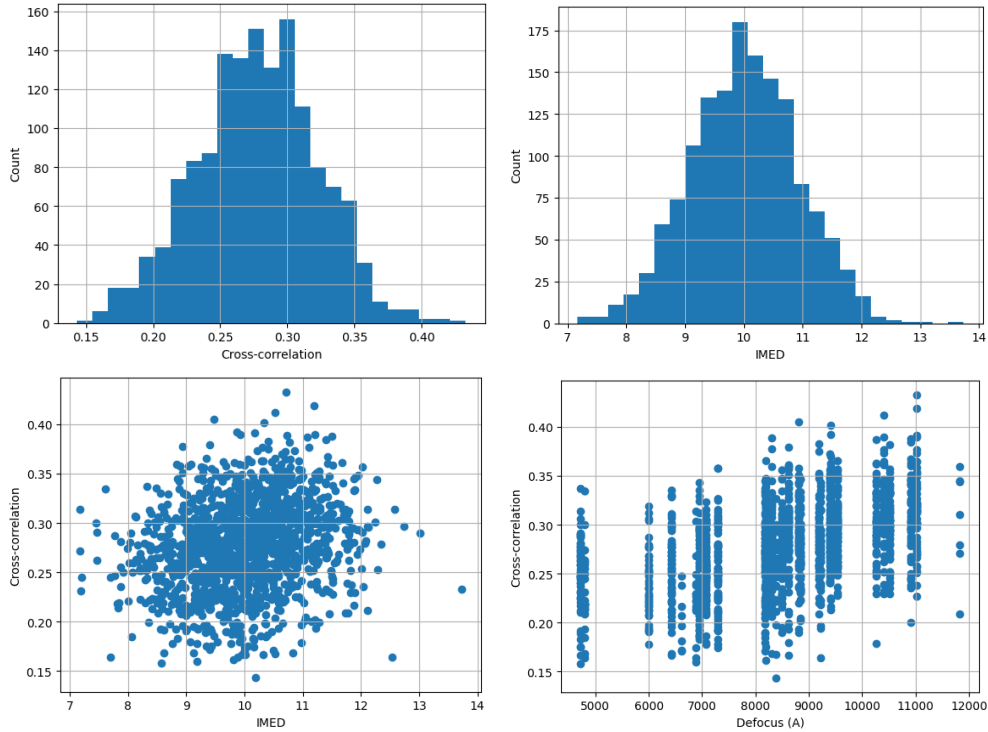
Figure 47: Top: Histogram of the cross-correlation (CC) and IMED between the experimental images and their corresponding reprojections. Bottom: Scatter plots of CC vs IMED, and CC vs defocus.

STATUS: Cannot be automatically evaluated

9.2 Level 4.b Alignability smoothness

Explanation:
This algorithm [Méndez et al., 2021] analyzes the smoothness of the correlation function over the projection sphere and the stability of its maximum. Ideally, the angular assignment given by the user should coincide with the maximum of the smoothed cross-correlation landscape.

Results:

Fig. 48 shows the histogram of the angular distance between the angular

assignment given by the user and the maximum of the smoothed landscape of cross-correlations. plot. The average angular distance 11.946. The percentage of images whose distance is larger than 10 is 50.2%.
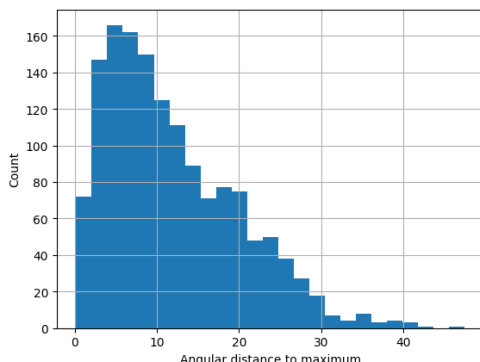

Figure 48: Histogram of the angular distance between the angular assignment given by the user and the maximum of the smoothed landscape of cross-correlation.

Automatic criteria: The validation is OK if less than 30% of the images have their angular assignment is further than 10 degrees from the smoothed cross-correlation maximum.

WARNINGS: 1 warnings

1. **The percentage of images whose angular assignment is significantly away from the smoothed maximum is too high, 50.2%**

9.3   Level 4.c Alignability precision and accuracy

Explanation:
The precision [Vargas et al., 2016] analyzes the orientation distribution of the best matching reprojections from the reference volume. If the high values are clustered around the same orientation, then the precision is close to 1. Otherwise, it is closer to -1. Below 0.5 the best directions tend to be scattered. The alignability accuracy [Vargas et al., 2017] compares the final angular assignment with the result of a new angular assignment. The similarity between both is again encoded between -1 and 1.

Results:

Fig. 49 shows the histograms of the accuracy and precision, and a joint scatter plot. The average accuracy was 0.701 and the average precision 0.789. The percentage of images whose accuracy is below 0.5 is 18.3%, and the percentage of images whose precision is below 0.5 is 8.1%.
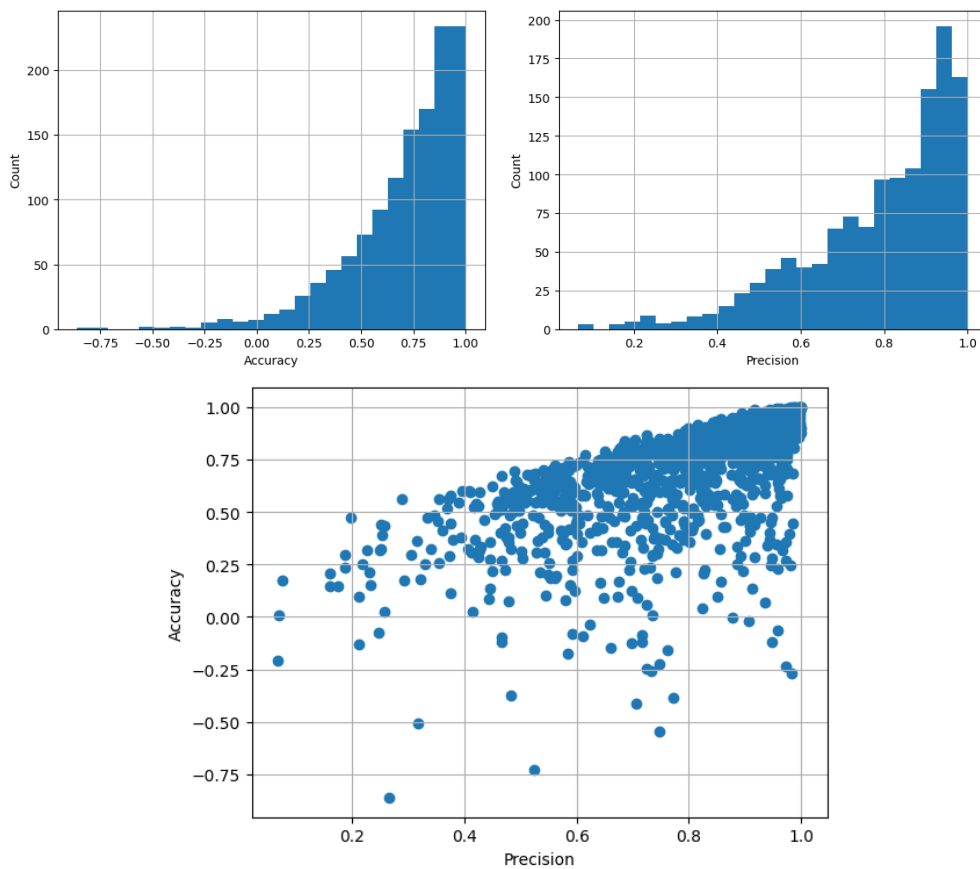

Figure 49: Top: Histogram of the accuracy and precision. Bottom: Scatter plot of both measures.

Automatic criteria: The validation is OK if less than 30% of the images have an 1) accuracy and 2) precision smaller than 0.5.

9.4   Level 4.d1 Relion alignment

Explanation:

We have performed an independent angular assignment using Relion autorefine [Scheres, 2012]. Images were downsampled to a pixel size of 3Å. Then, we measured the difference between the angular assignment of the particles given by the user and the one done by Relion.

Results:

Fig. 50 shows some representative slices of the reconstruction performed by Relion for checking its correctness.

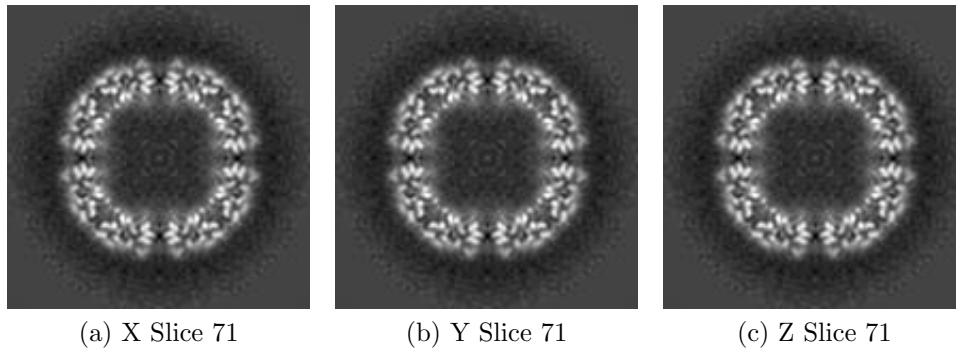| (a) X Slice 71 | (b) Y Slice 71 | (c) Z Slice 71 |

Figure 50: Slices of maximum variation in the three dimensions of the map reconstructed by Relion

Fig. 51 shows the shift and angular difference between the alignment given by the user and the one calculated by Relion. The median shift difference was 0.3Å, and the median angular difference 0.5 degrees. Their corresponding median absolute deviations were 0.2 and 0.3, respectively. Particles with a shift difference larger than 5Å or an angular difference larger than 5 degrees would be considered as incorrectly assigned in one of the two assignments (the user's or the new one). 0.1% of particles were considered to have an uncertain shift, and 0.5% of particles were considered to have an uncertain
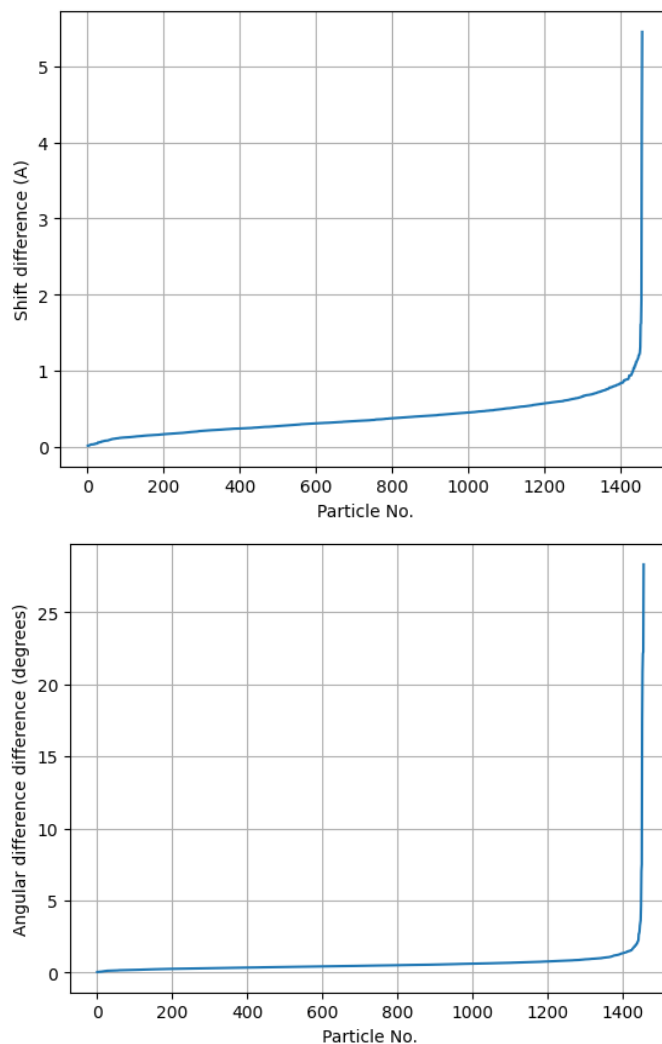
alignment.

Figure 51: Top: Shift difference between the alignment given by the user and the one calculated by Relion. Bottom: Angular difference. The X-axis represents all particles sorted by their difference.

Automatic criteria: The validation is OK if less than 20% of the images have 1) a shift difference larger than 5Å, and 2) an angular difference larger than 5 degrees.

STATUS: OK

9.5   Level 4.d2 CryoSparc alignment

Explanation:
We have performed an independent angular assignment using CryoSparc non-homogeneous refinement [Punjani et al., 2020]. Images were downsampled to a pixel size of 3Å. Then, we measured the difference between the angular assignment of the particles given by the user and the one done by CryoSparc.

Results:

Fig. 52 shows some representative slices of the reconstruction performed by Cryosparc for checking its correctness.

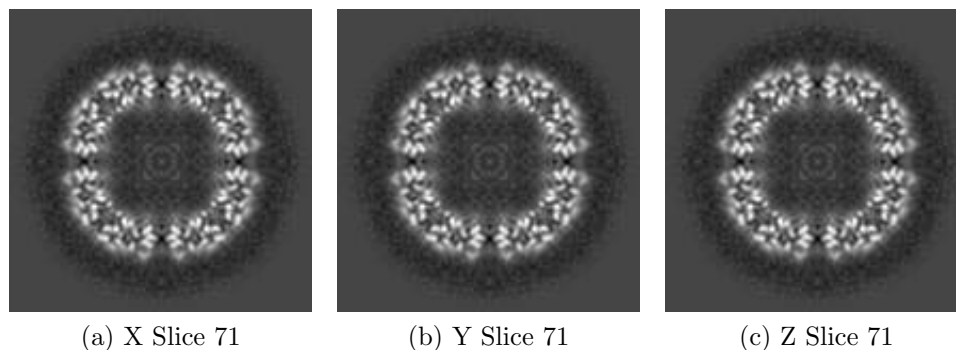(a) X Slice 71   (b) Y Slice 71   (c) Z Slice 71

Figure 52: Slices of maximum variation in the three dimensions of the map reconstructed by Cryosparc

Fig. 53 shows the shift and angular difference between the alignment given by the user and the one calculated by CryoSparc. The median shift difference was 3.2Å, and the median angular difference 0.5 degrees. Their corresponding median absolute deviations were 1.9 and 0.2, respectively. Particles with a shift difference larger than 5Å or an angular difference larger than 5 degrees would be considered as incorrectly assigned in one of the two assignments (the user's or the new one). 23.1% of particles were considered

70

to have an uncertain shift, and 0.1% of particles were considered to have an uncertain alignment.
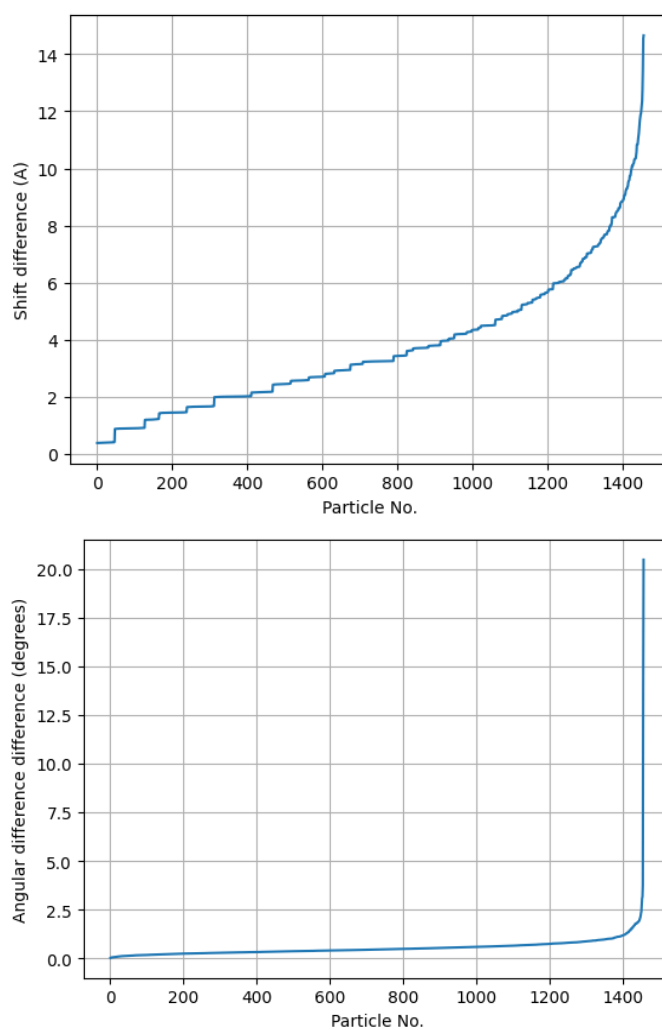

Figure 53: Top: Shift difference between the alignment given by the user and the one calculated by CryoSparc. Bottom: Angular difference. The X-axis represents all particles sorted by their difference.

Automatic criteria: The validation is OK if less than 20% of the images have 1) a shift difference larger than 5Å, and 2) an angular difference

71

larger than 5 degrees.

WARNINGS: 1 warnings

1. **The percentage of images with uncertain shift is larger than 20%**

9.6 Level 4.d3 Relion/CryoSparc alignments

Explanation:
In Secs. 9.4 and 9.5 we compared the angular assignment given by the user to the angular assignment of Relion and CryoSparc, respectively. We now compare these two alignments as a way to measure the "intrinsic" uncertainty in the angular assignment. This comparison gives an estimate of the alignability of the input images.

Results:

Fig. 54 shows the shift and angular difference between the alignment given by Relion and the one calculated by CryoSparc. The median shift difference was 3.2Å, and the median angular difference 0.3 degrees. Their corresponding median absolute deviations were 2.0 and 0.2, respectively. Particles with a shift difference larger than 5Å or an angular difference larger than 5 degrees would be considered as incorrectly assigned in one of the two assignments (the user's or the new one). 23.7% of particles were considered to have an uncertain shift, and 0.6% of particles were considered to have an uncertain alignment.
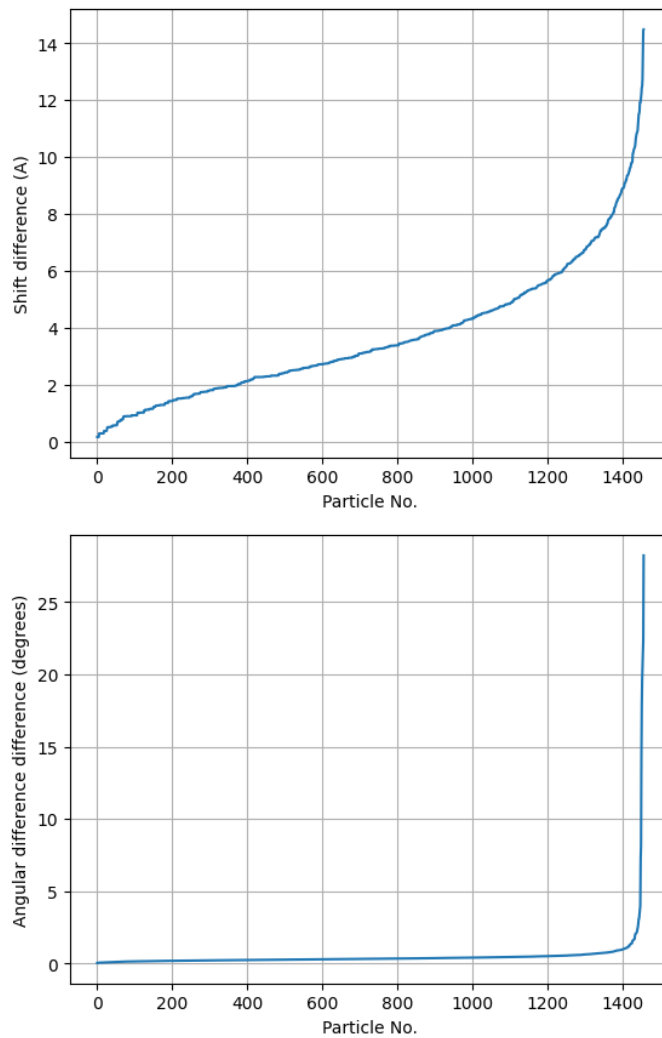
Figure 54: Top: Shift difference between the alignment given by Relion and the one calculated by CryoSparc. Bottom: Angular difference. The X-axis represents all particles sorted by their difference.

Automatic criteria: The validation is OK if less than 20% of the images have 1) a shift difference larger than 5Å, and 2) an angular difference larger than 5 degrees.

WARNINGS: 1 warnings

73

1. **The percentage of images with uncertain shift is larger than 20%**

9.7 Level 4.e Classification without alignment

Explanation:
We have performed a 3D classification of the input particles in two classes without aligning them using Relion [Scheres, 2012] to confirm they belong to a single state. Images were downsampled to a pixel size of 3Å. A valid result would be: 1) a class attracting most particles and an almost empty class, 2) two classes with an arbitrary number of images in each one, but without any significant structural difference between the two.

Results:

The classification converged to a single class with 1457 out of 1457 images in it.
Automatic criteria: The validation is OK if the classification converged to a single class.

STATUS: OK

9.8 Level 4.f Overfitting detection

Explanation:
The detection of overfitting can be performed through a series of 5 reconstructions with an increasing number of experimental particles and the same number of pure noise particles [Heymann, 2015]. The resolution of the reconstructions with experimental particles should always be better than those from noise.

Results:

We tested with subsets of 21, 72, 218, 546 and 728 particles. Fig. 55 shows the inverse of the resolution as a function of the number of particles.
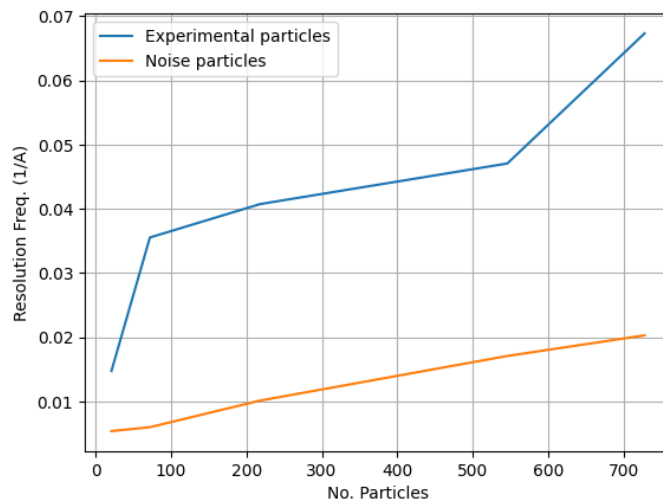
Figure 55: Inverse of the resolution as a function of the number of particles.

Automatic criteria: The validation is OK if the resolution of noise particles is never better than the resolution of true particles.

STATUS: OK

9.9 Level 4.g Angular distribution efficiency

Explanation:
This method [Naydenova and Russo, 2017] evaluates the ability of the angular distribution to fill the Fourier space. It determines a resolution per direction based on the number of particles in each direction and reports the distribution efficiency, a number between 0 (inefficient) and 1 (total efficiency).

Results:

Fig. 56 shows the histogram of the measured resolutions per direction. The average resolution was 2.6 Å, and its range [1.5, 3.4].
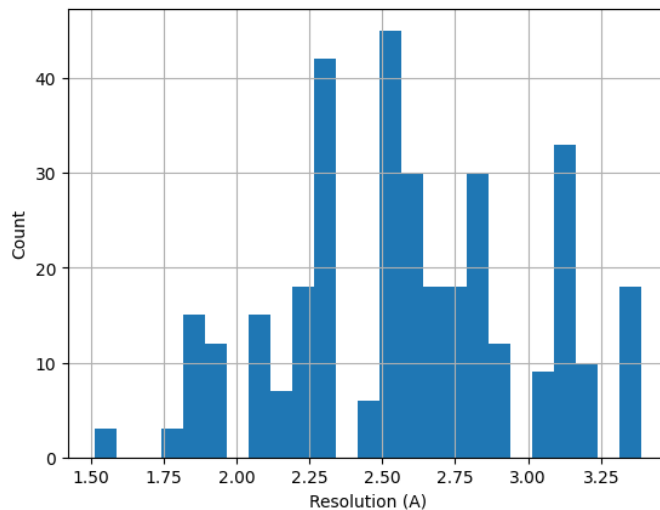
Figure 56: Histogram of the directional resolution according to the angular distribution efficiency.

Automatic criteria: The validation is OK if the resolution reported by the user is larger than 0.8 times the average directional resolution.

STATUS: OK

9.10   Level 4.h Sampling compensation factor

Explanation:
SCF [Baldwin and Lyumkis, 2020] measures the ability of the angular distribution to fill the Fourier space.

Results:

The results of the SCF analysis was:
Tilt= 0.0000
Number of zeros= 0.0000
Fraction of zeros= 0.0000
QkoverPk= 0.0000
SCF= 0.9498

SCFStar= 0.9498

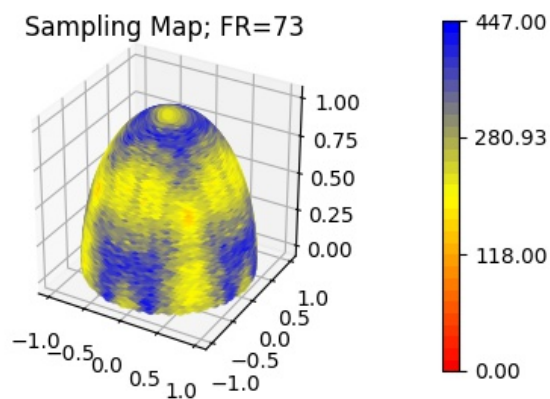Fig. 57 shows the SCF plot for this angular distribution.

Figure 57: SCF plot.

Automatic criteria: The validation is OK if the SCF*>0.5.

STATUS: OK

9.11 Level 4.i CTF stability

Explanation:
We estimated the per-particle defocus, B-factor, astigmatism, and phase shift using Relion's ctf refine. Ideally, the differences in defoci cannot be larger than the ice thickness. We also estimated the local magnification offsets (which should be around 0) and the B-factor.

Results:

The following list shows the median, confidence intervals and links to the histograms for the refined parameters. Ideally these should all concentrate around 0, except for the defocus and the phase shift that must be centered around their true values.

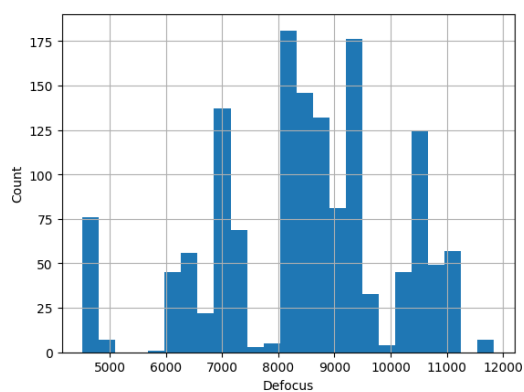| Parameter	Median	95% Confidence interval	Histogram
Defocus (Å)	8596.51	[4707.2,11021.6]	Fig. 58
Astigmatism (Å)	336.38	[63.0,587.7]	Fig. 59
Defocus difference (Å)	73.58	[-4515.4,3342.9]	Fig. 60
Astigmatism difference (Å)	-17.92	[-430.6,327.5]	Fig. 61
Phase shift (degrees)	-5.06	[-20.7, 9.0]	Fig. 62
B-factor (Å^2)	0.19	[-1.3, 3.4]	Fig. 63
Scale factor	0.052	[0.037,0.065]	Fig. 64

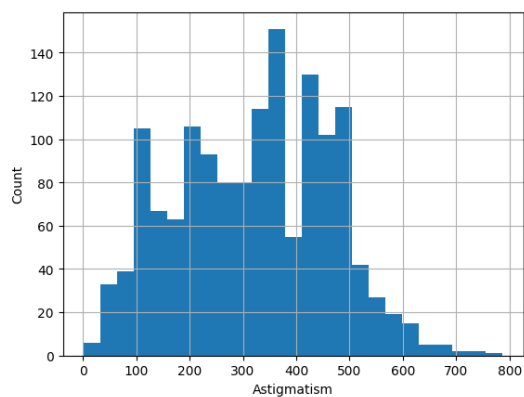Figure 58: Histogram of the defocus after local refinement (Å).

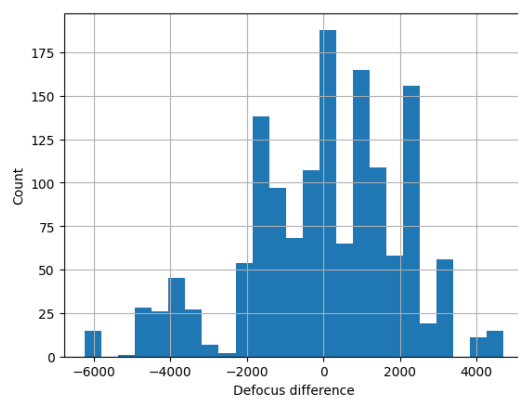Figure 59: Histogram of the astigmatism after local refinement (Å).

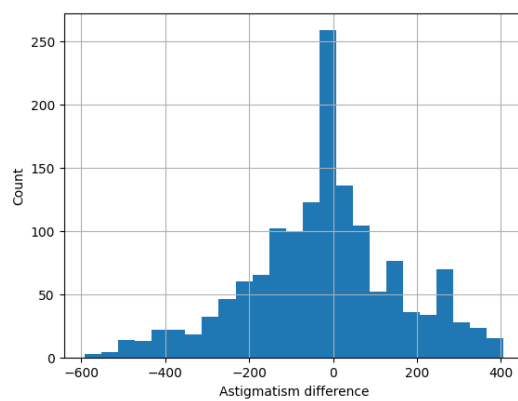Figure 60: Histogram of the difference in defocus after local refinement (Å).

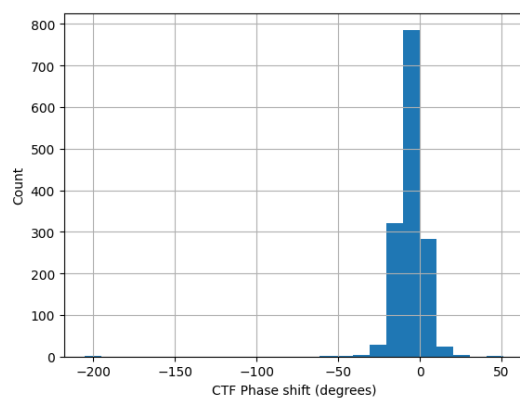Figure 61: Histogram of the difference in astigmatism after local refinement (Å).
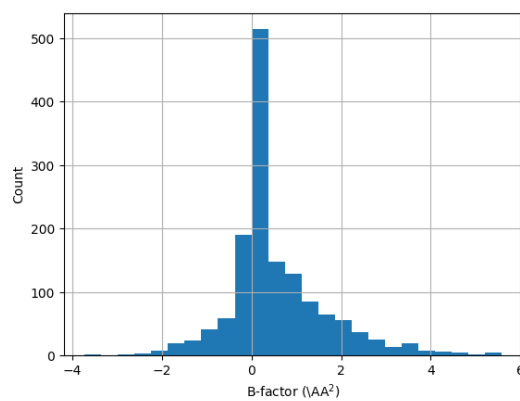
Figure 62: Histogram of the CTF phase shift (degrees).

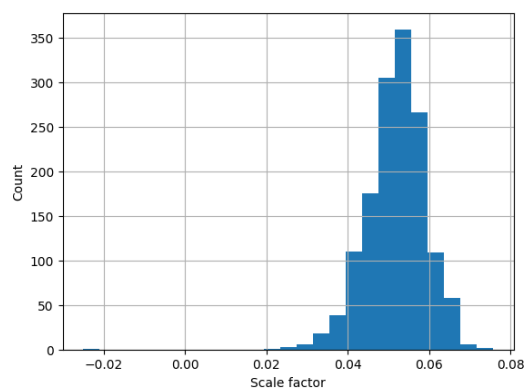Figure 63: Histogram of the B-factor ($\mathring{A}^2$).

Figure 64: Histogram of the Scale factor.

Automatic criteria: The validation is OK if 1) astigmatism is between -5000 and 5000; 2) the defocus difference is between -5000 and 5000; 3) the astigmatism difference is between -5000 and 5000; 4) the B-factor is between -5 and 5; and 5) the scale factor is between -0.1 and 0.1.

WARNINGS: 1 warnings

1. **The 95% confidence interval of scale factor is not centered.**

10   Micrographs

Set of Micrographs: /home/coss/ScipionUserData/projects/Example_10248_Scipion3/-Runs/006458_XmippProtMovieCorr/extra/*mic.mrc

30 micrographs were provided by the user. The first 2 can be seen in Fig. 65.
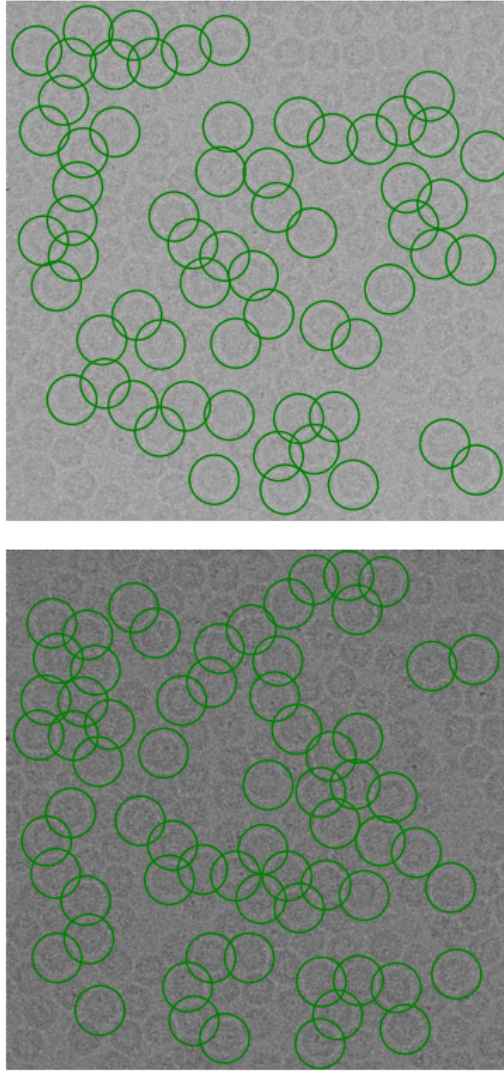
Figure 65: Two example micrographs with their coordinates.

11 Level 5 analysis

11.1 Level 5.a Micrograph cleaner

Explanation:
This method assigns a score between 0 (bad coordinate) and 1 (good coordi-

nate) reflecting the probability that the coordinate is outside a region with aggregations, ice crystals, carbon edges, etc. [Sanchez-Garcia et al., 2020]

Results:

0 coordinates out of 1457 (0.0 %) were scored below 0.9 by Micrograph-Cleaner.

   Automatic criteria: The validation is OK if less than 20% of the coordinates are suspected to lie in aggregations, contaminations, ice crystals, etc.

   STATUS: OK

12   Atomic model

Atomic model: /home/coss/ScipionUserData/projects/Example_10248_Scipion3/-centered4V1W.pdb

See Fig. 66.

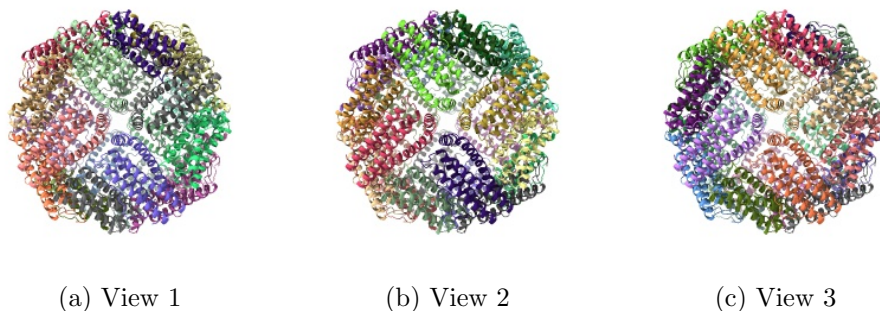(a) View 1   (b) View 2   (c) View 3

Figure 66: Input atomic model Views generated by ChimeraX at a the following X, Y, Z angles: View 1 (0,0,0), View 2 (90, 0, 0), View 3 (0, 90, 0).

13 Level A analysis

13.1 Level A.a MapQ

Explanation:
MapQ [?] computes the local correlation between the map and each one of its atoms assumed to have a Gaussian shape.

Results:

Fig. 67 shows the histogram of the Q-score according calculated by MapQ. Some representative percentiles are:

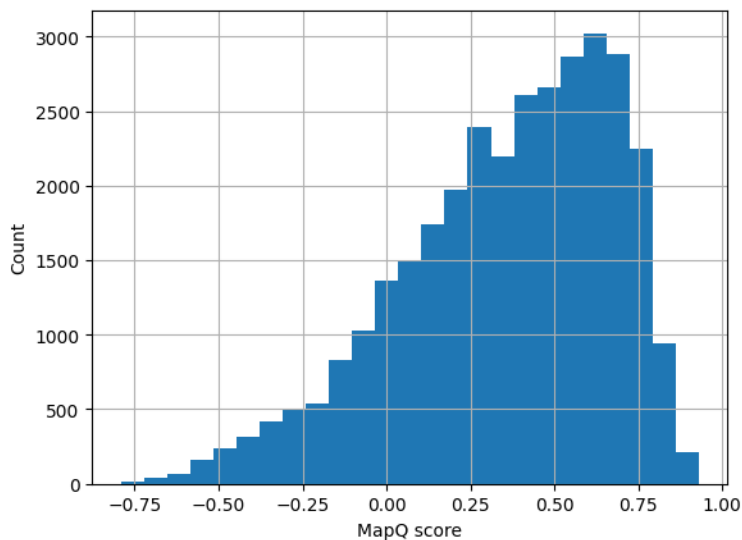| Percentile	MapQ score [0-1]
2.5%	-0.38
25%	0.15
50%	0.40
75%	0.61
97.5%	0.81

Figure 67: Histogram of the Q-score.

The following table shows the average Q score and estimated resolution for each chain.

Chain	Average Q score [0-1]	Estimated Resol. (Å)
A	0.36	4.3
B	0.36	4.3
C	0.37	4.2
D	0.36	4.3
E	0.36	4.3
F	0.36	4.3
G	0.35	4.3
H	0.36	4.3
I	0.35	4.3
J	0.36	4.3
K	0.36	4.3
L	0.36	4.3
M	0.35	4.3
N	0.35	4.3
O	0.36	4.3
P	0.35	4.3
Q	0.35	4.3
R	0.35	4.3
S	0.35	4.3
T	0.35	4.3
U	0.35	4.3
V	0.35	4.3
W	0.35	4.3
X	0.36	4.3

Automatic criteria: The validation is OK if the median Q-score is larger than 0.1.

STATUS: OK

13.2   Level A.b FSC-Q

Explanation:
FSC-Q [Ramírez-Aportela et al., 2021] compares the local FSC between the

85

map and the atomic model to the local FSC of the two half maps. FSC-Qr is the normalized version of FSC-Q to facilitate comparisons. Typically, FSC-Qr should take values between -1.5 and 1.5, being 0 an indicator of good matching between map and model.

Results:

Fig. 68 shows the histogram of FSC-Qr and Fig. 69 the colored isosurface of the atomic model converted to map. The average FSC-Qr is 0.53, its 95% confidence interval is [-1.00, 2.39]. The percentage of values whose FSC-Qr absolute value is beyond 1.5 is 7.2 %.
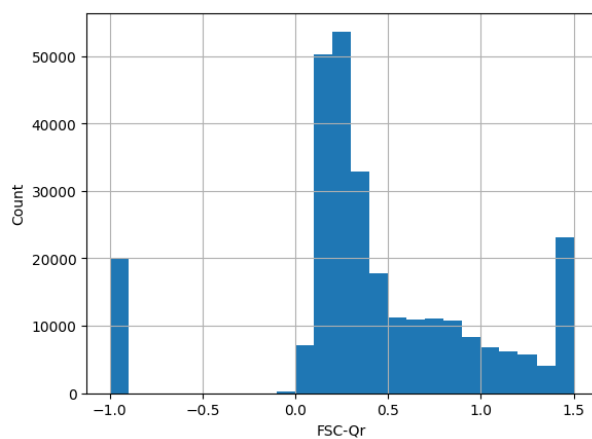

Figure 68: Histogram of the FSC-Qr limited to -1.5 and 1.5.

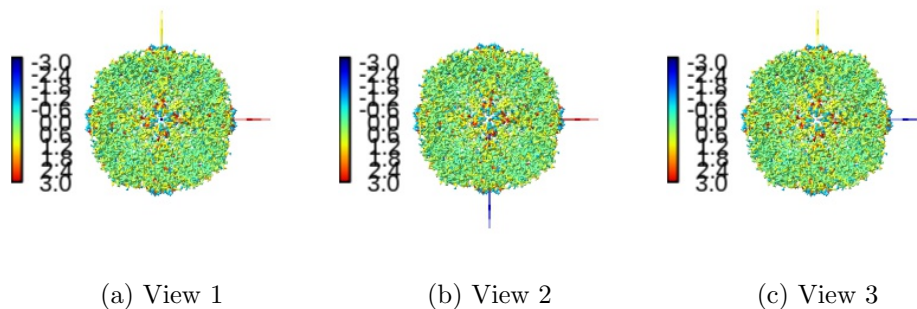|		
(a) View 1	(b) View 2	(c) View 3

Figure 69: Isosurface of the atomic model colored by FSC-Qr between -1.5 and 1.5 Views generated by ChimeraX at a the following X, Y, Z angles: View 1 (0,0,0), View 2 (90, 0, 0), View 3 (0, 90, 0).

Automatic criteria: The validation is OK if the percentage of residues whose FSC-Q is larger than 1.5 in absolute value is smaller than 10%.

STATUS: OK

13.3 Level A.c Multimodel stability

Explanation:
The method of [Herzik et al., 2019] estimates the ambiguity of the atomic model in each region of the CryoEM map due to the different local resolutions or local heterogeneity.

Results:

Fig. 70 shows the histogram of the RMSD of the different models. The average RMSD between models is 0.45 Å. Fig. 71 shows the atomic model colored by RMSD.
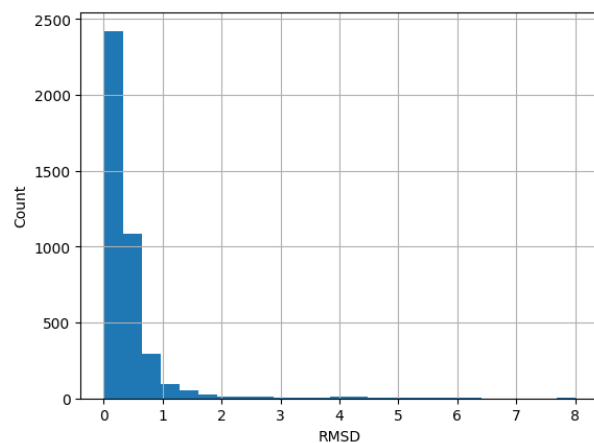
Figure 70: Histogram of RMSD of the different atoms of the multiple models.

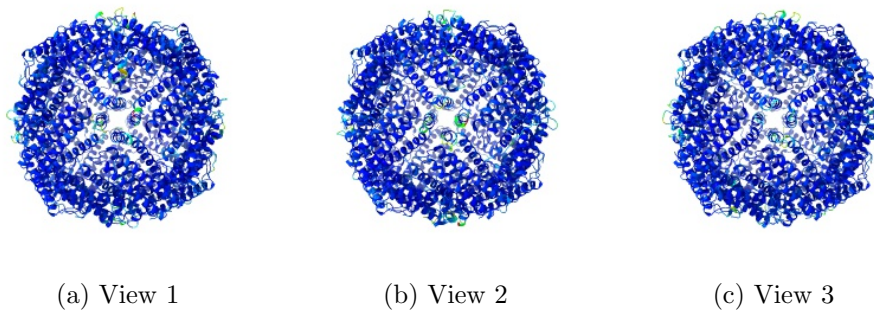(a) View 1

(b) View 2

(c) View 3

Figure 71: Atomic model colored by RMSD Views generated by ChimeraX at a the following X, Y, Z angles: View 1 (0,0,0), View 2 (90, 0, 0), View 3 (0, 90, 0).

Automatic criteria: The validation is OK if the average RMSD is smaller than 2Å.

STATUS: OK

13.4 Level A.d Map-Model Guinier analysis

Explanation:

We compared the Guinier plot [Rosenthal and Henderson, 2003] of the atomic model and the experimental map. We made the mean of both profiles to be equal (and equal to the mean of the atomic model) to make sure that they had comparable scales.

Results:

Fig. 72 shows the logarithm (in natural units) of the structure factor (the module squared of the Fourier transform) of the atom model and the experimental map. The correlation between the two profiles was 0.977.
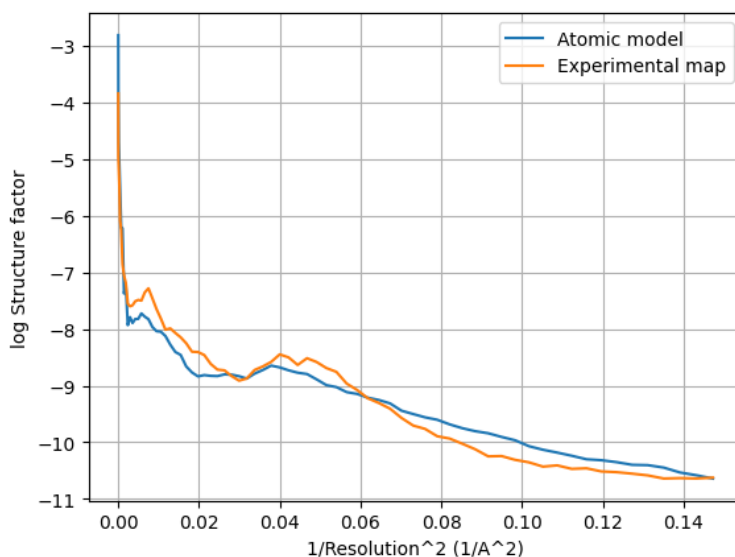

Figure 72: Guinier plot of the atom model and experimental map. The X-axis is the square of the inverse of the resolution in Å.

Automatic criteria: The validation is OK if the correlation between the two Guinier profiles is larger than 0.5.

STATUS: OK

13.5 Level A.e Phenix validation

Explanation:
Phenix provides a number of tools to assess the agreement between the experimental map and its atomic model [Afonine et al., 2018]. There are several cross-correlations to assess the quality of the fitting:

- CC (mask): Model map vs. experimental map correlation coefficient calculated considering map values inside a mask calculated around the macromolecule.
- CC (box): Model map vs. experimental map correlation coefficient calculated considering all grid points of the box.
- CC (volume) and CC (peaks) compare only map regions with the highest density values and regions below a certain contouring threshold level are ignored. CC (volume): The map region considered is defined by the N highest points inside the molecular mask. CC (peaks): In this case, calculations consider the union of regions defined by the N highest peaks in the model-calculated map and the N highest peaks in the experimental map.
- Local real-space correlation coefficients CC (main chain) and CC (side chain) involve the main skeleton chain and side chains, respectively.

There are also multiple ways of measuring the resolution:

- d99: Resolution cutoff beyond which Fourier map coefficients are negligibly small. Calculated from the full map.
- d_model: Resolution cutoff at which the model map is the most similar to the target (experimental) map. For d_model to be meaningful, the model is expected to fit the map as well as possible. d_model (B factors = 0) tries to avoid the blurring of the map.
- d_FSC_model; Resolution cutoff up to which the model and map Fourier coefficients are similar at FSC values of 0, 0.143, 0.5.

In addition to these resolution measurements the overall isotropic B factor is another indirect measure of the quality of the map.
Results:

To avoid ringing in Fourier space a smooth mask with a radius of 6.0 Å has been applied.
Overall correlation coefficients:

$$
\begin{aligned}
\text{CC (mask)} &= & 0.711 \\
\text{CC (box)} &= & 0.665 \\
\text{CC (volume)} &= & 0.714 \\
\text{CC (peaks)} &= & 0.578 \\
\text{CC (main chain)} &= & 0.684 \\
\text{CC (side chain)} &= & 0.654
\end{aligned}
$$

Correlation coefficients per chain:

Chain	Cross-correlation
A	0.685993
B	0.683896
C	0.687664
D	0.686794
E	0.683483
F	0.685114
G	0.685297
H	0.686324
I	0.684185
J	0.687296
K	0.686834
L	0.686390
M	0.684727
N	0.686484
O	0.683226
P	0.684781
Q	0.684891
R	0.683841
S	0.684050
T	0.687113
U	0.686038
V	0.686198
W	0.686749
X	0.684988

We now show the correlation profiles of the different chain per residue.
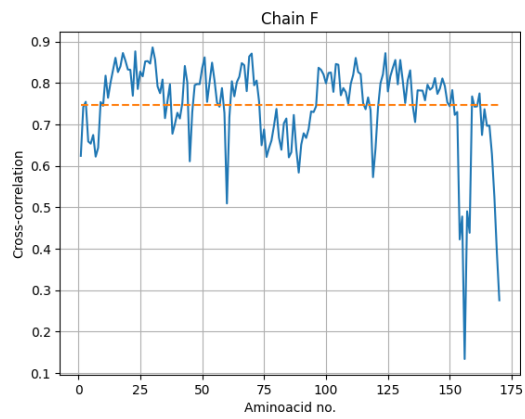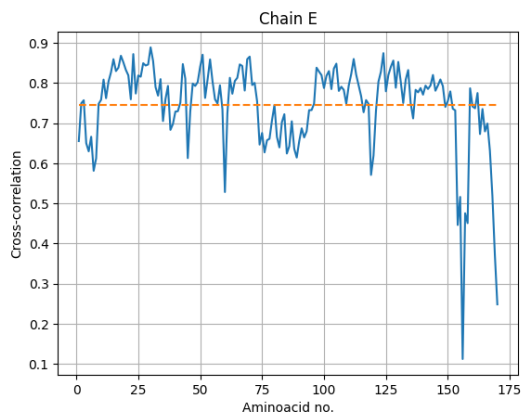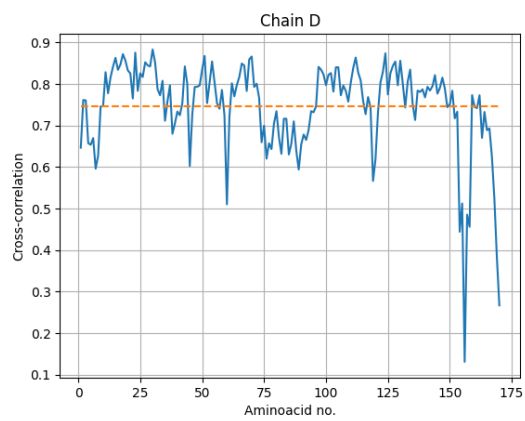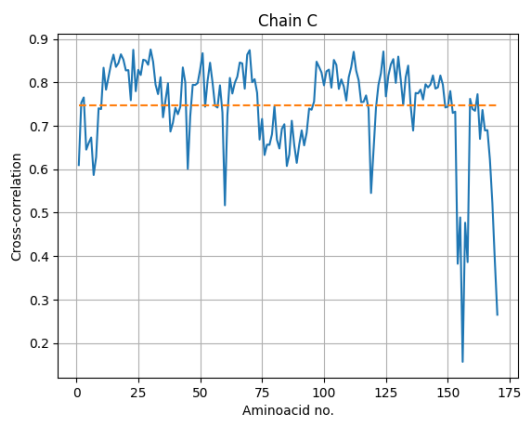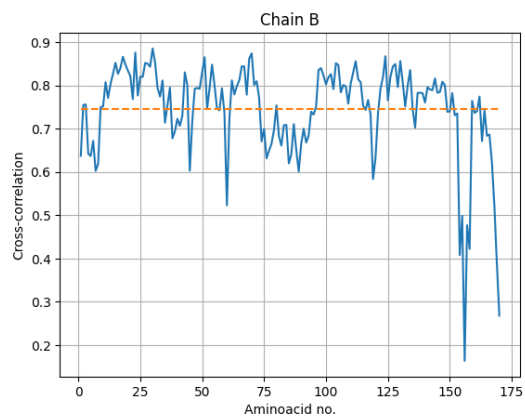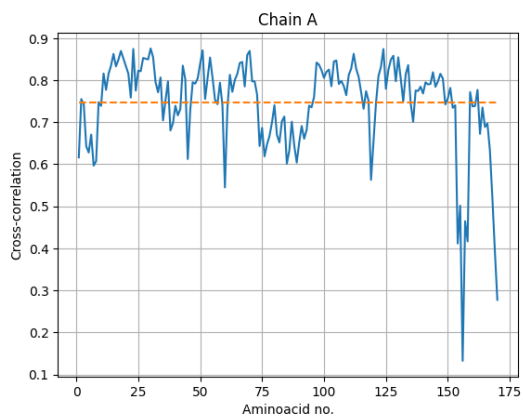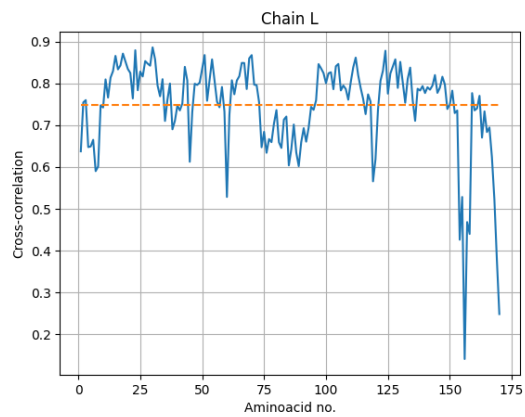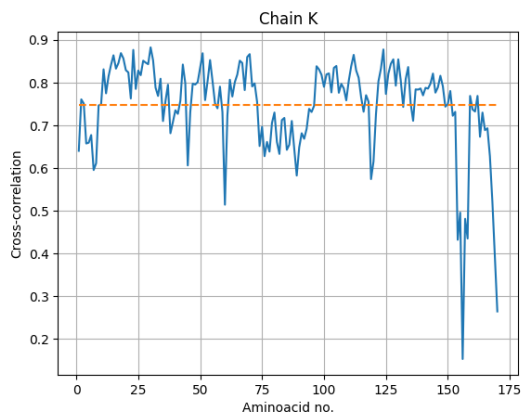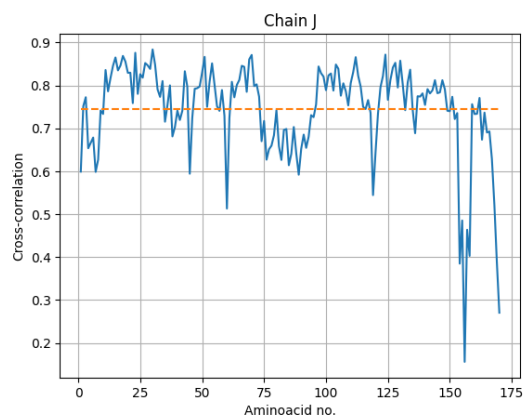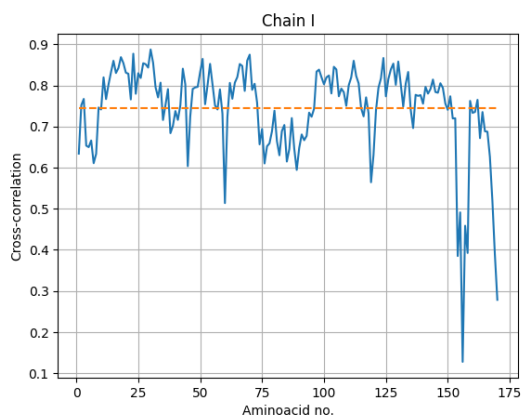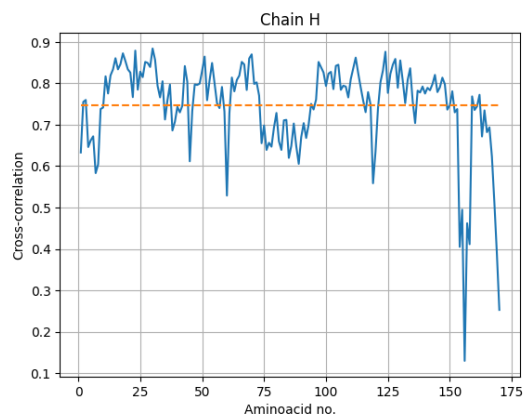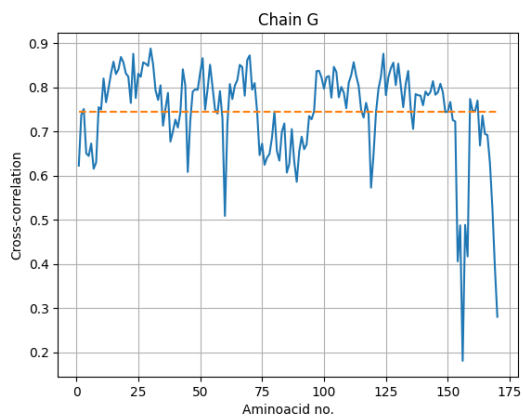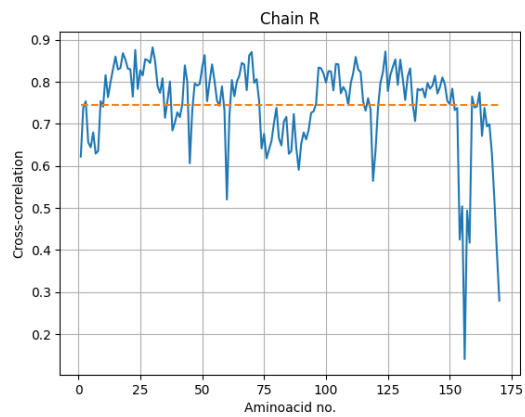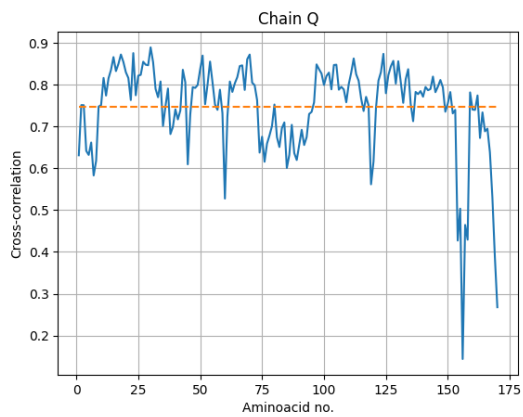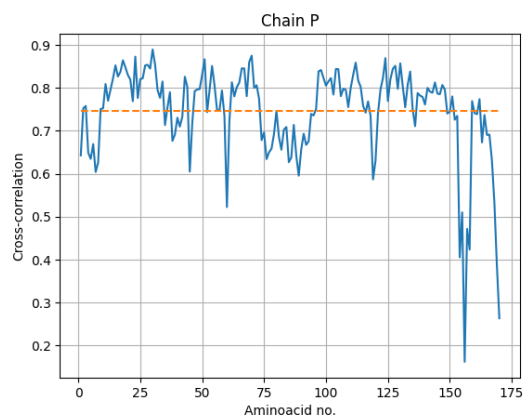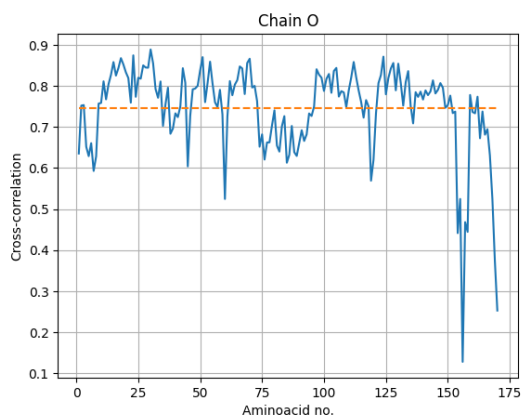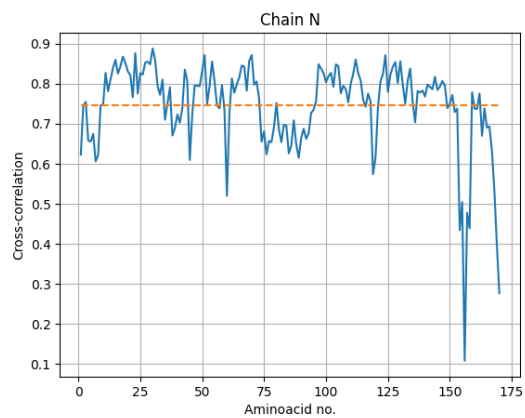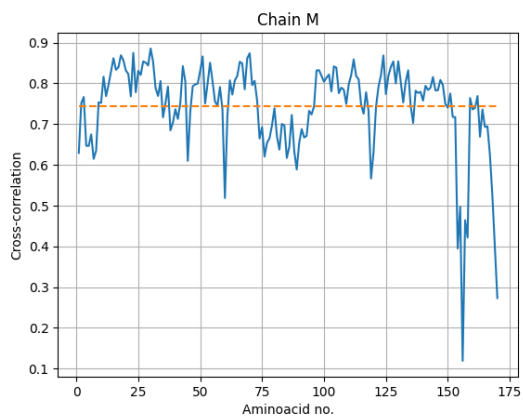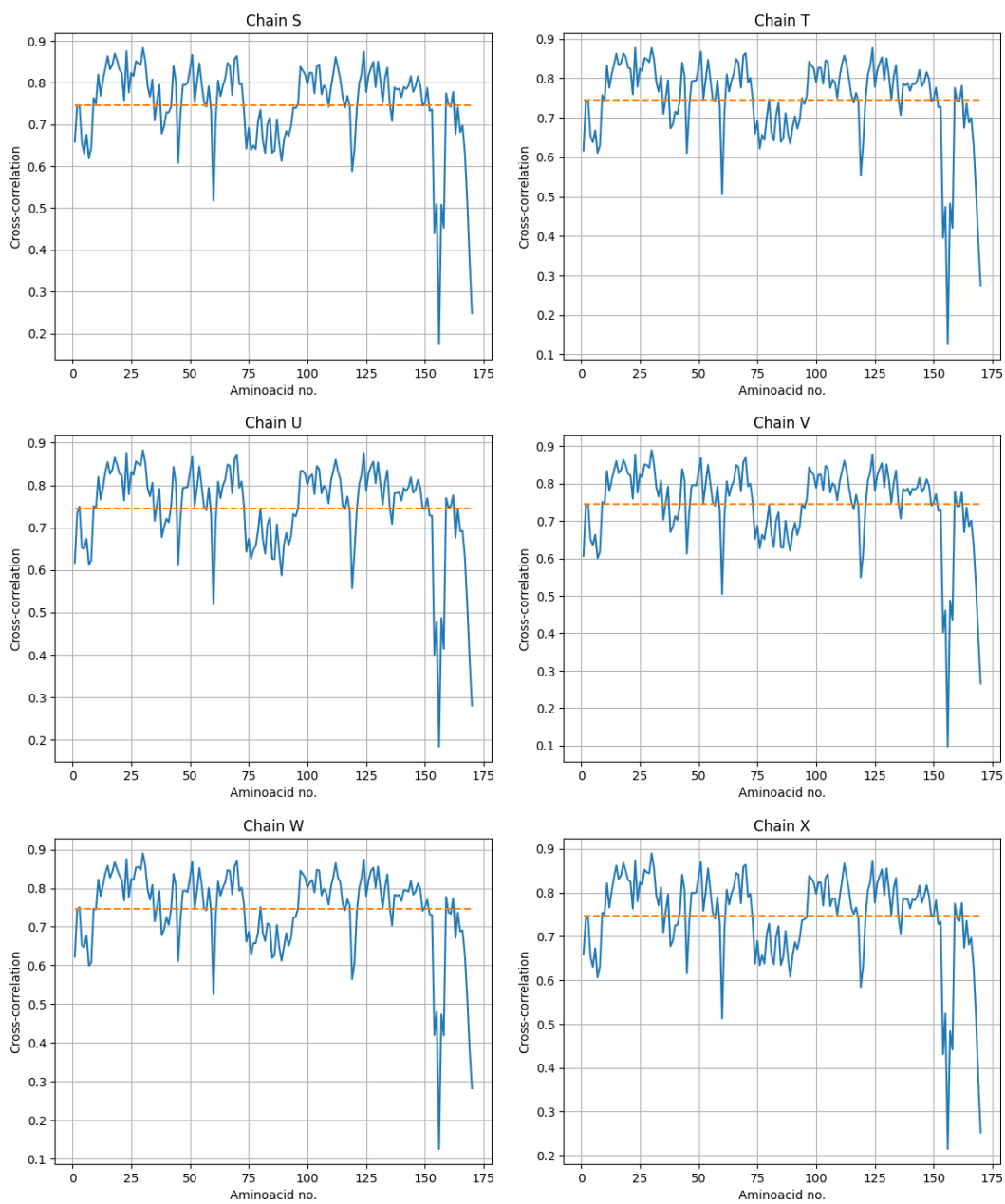
Fig. 73 shows the histogram of all cross-correlations evaluated at the residues. The percentage of residues whose correlation is below 0.5 is 3.8 %.
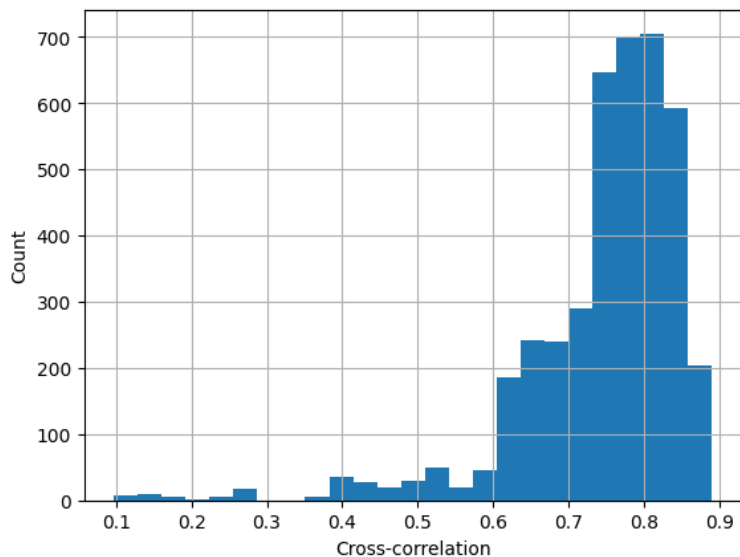
Figure 73: Histogram of the cross-correlation between the map and model evaluated for all residues.

Resolutions estimated from the model:

Resolution (Å)	Masked	Unmasked
d99	1.7	1.6
d_model	3.8	3.8
d_model (B-factor=0)	4.3	4.3
FSC_model=0	3.0	2.9
FSC_model=0.143	3.4	3.4
FSC_model=0.5	4.4	4.5

Overall isotropic B factor:

B factor	Masked	Unmasked
Overall B-iso	85.0	85.0

Fig. 74 shows the FSC between the input map and the model.

96
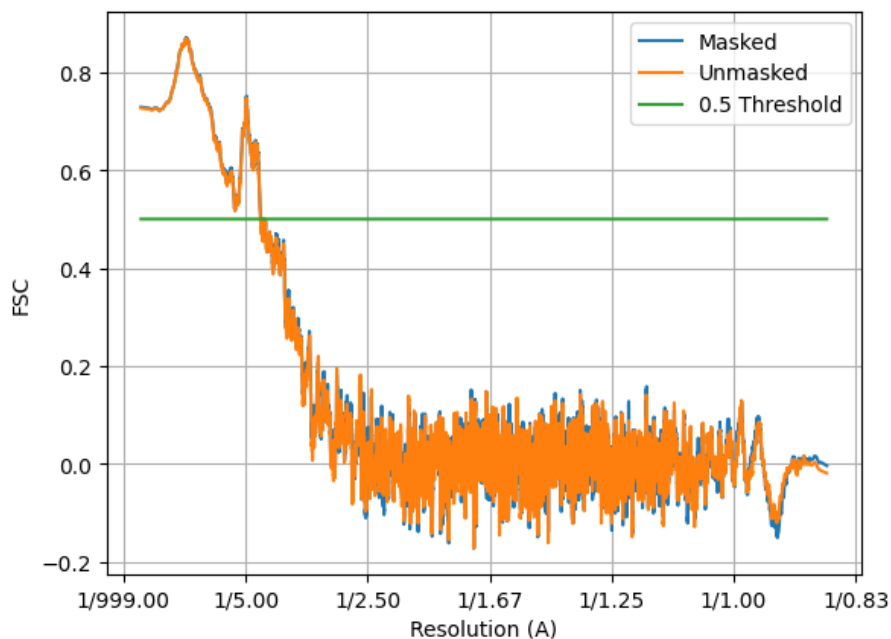
Figure 74: FSC between the input map and model with and without a mask constructed from the model. The X-axis is the square of the inverse of the resolution in Å.

Automatic criteria: The validation is OK if 1) the percentage of residues whose correlation is smaller than 0.5 is smaller than 10%, and 2) the resolution reported by the user is larger than 0.8 times the resolution estimated between the map and model at FSC=0.5.

WARNINGS: 1 warnings

1. **The resolution reported by the user, 2.6 Å, is significantly smaller than the resolution estimated between map and model (FSC=0.5), 4.4 Å**

13.6 Level A.f EMRinger validation

Explanation:
EMringer [Barad et al., 2015] compares the side chains of the atomic model to the CryoEM map. The following features are reported:

- Optimal Threshold: Electron potential map cutoff value at which the maximum EMRinger score was obtained.
- Rotamer Ratio: Fraction of rotameric residues at the Optimal threshold value.
- Max Zscore: Z-score computed to determine the significance of the distribution at the Optimal threshold value.
- Model Length: Total of non-gamma-branched, non-proline aminoacids with a non-H gamma atom used in global EMRinger score computation.
- EMRinger Score: Maximum EMRinger score calculated at the Optimal Threshold.

A rotameric residue is one in which EMRinger peaks that fall within defined rotamers based on chi1, this often suggests a problem with the modelling of the backbone. In general, the user should look at the profiles and identify regions that may need improvement.

Results:

General results:

Optimal threshold	0.005050
Rotamer ratio	0.649
Max. Zscore	4.87
Model length	2976
EMRinger Score	0.892

Fig. 75 shows the EMRinger score and fraction of rotameric residues as a function of the map threshold. The optimal threshold was selected looking for the maximum EMRinger score in this plot.
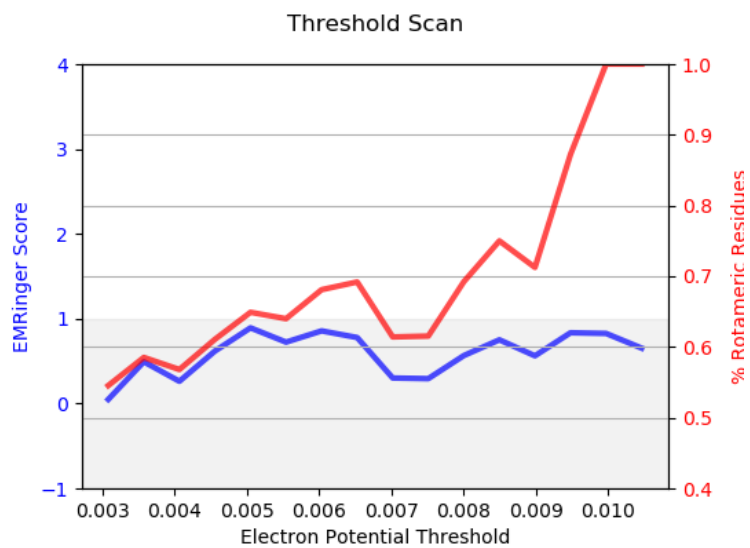
Figure 75: EMRinger score and fraction of rotameric residues as a function of the map threshold.

Fig. 76 shows the histogram for rotameric (blue) and non-rotameric (red) residues at the optimal threshold.
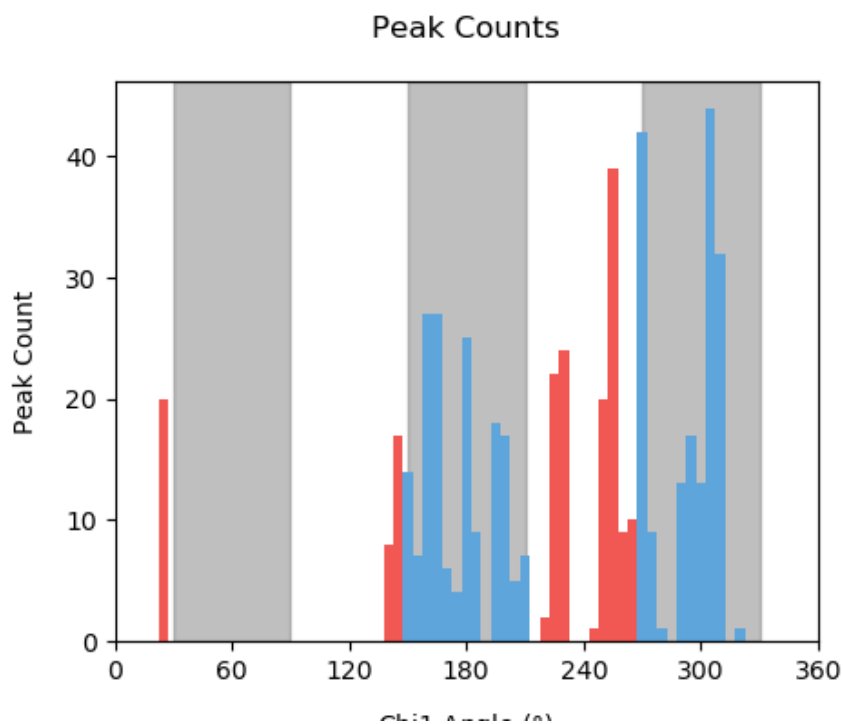
Figure 76: Histogram for rotameric (blue) and non-rotameric (red) residues at the optimal threshold as a function of the angle Chi1.

The following plots show the rolling window EMRinger analysis of the different chains to distinguish regions of improved model quality. This analysis was performed on rolling sliding 21-residue windows along the primary sequence of the protein chains.

100

Rolling window - Chain C, Threshold 0.000000

Rolling window - Chain D, Threshold 0.000000

Rolling window - Chain E, Threshold 0.000000

Rolling window - Chain F, Threshold 0.000000

Rolling window - Chain G, Threshold 0.000000

Rolling window - Chain H, Threshold 0.000000

101

103

Rolling window - Chain U, Threshold 0.000000

Rolling window - Chain V, Threshold 0.000000

Rolling window - Chain W, Threshold 0.000000

Rolling window - Chain X, Threshold 0.000000
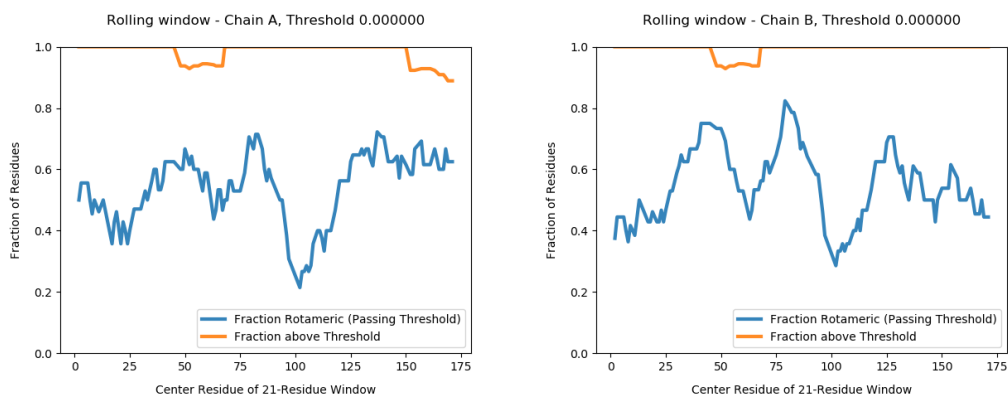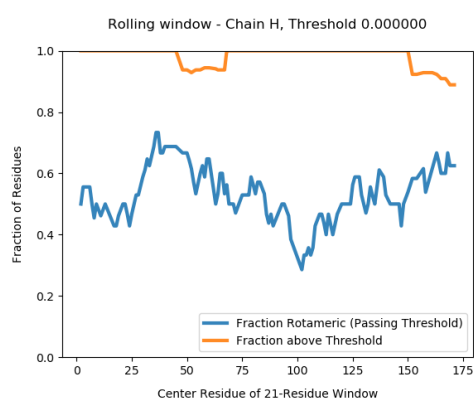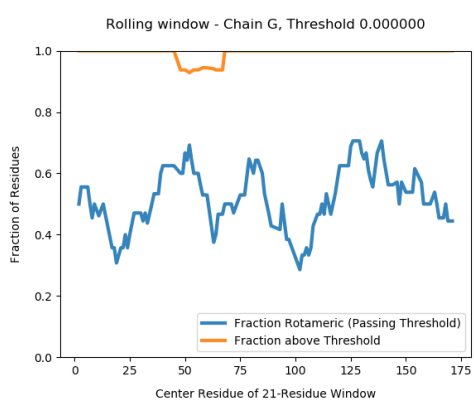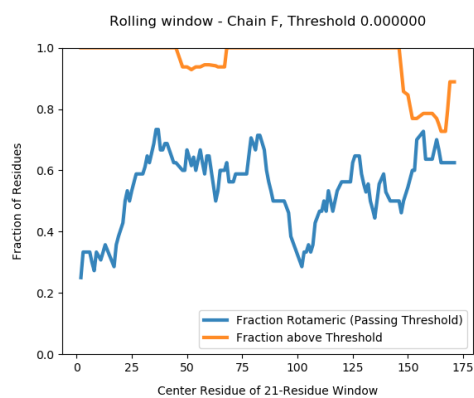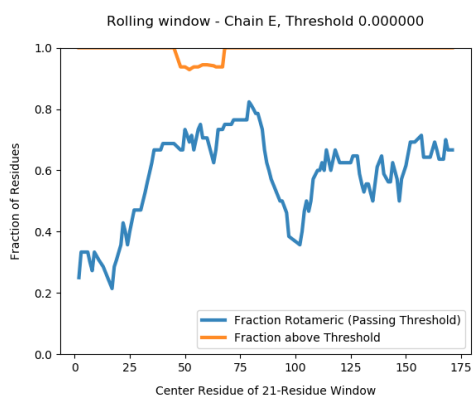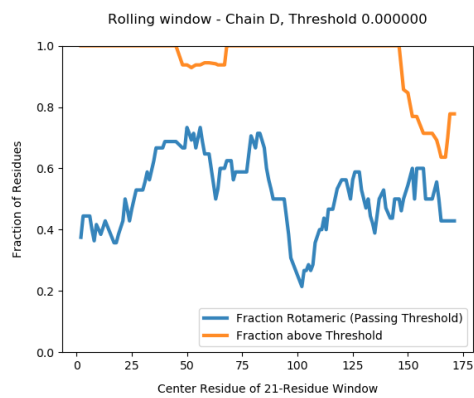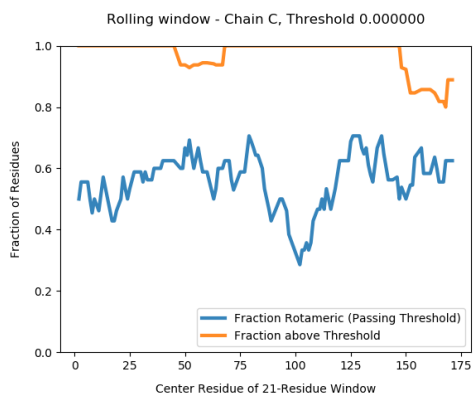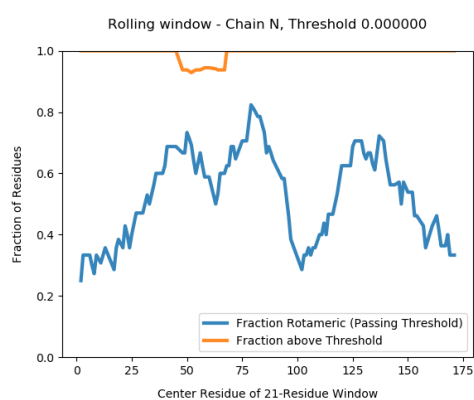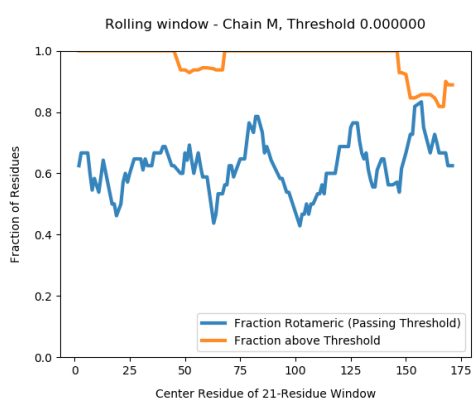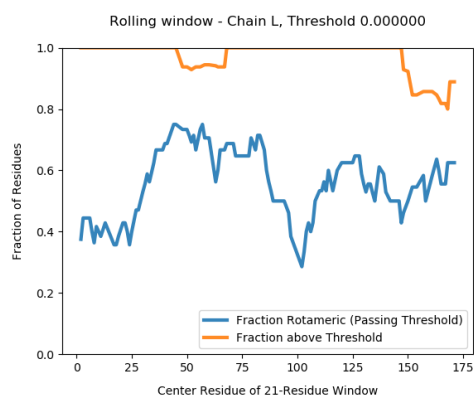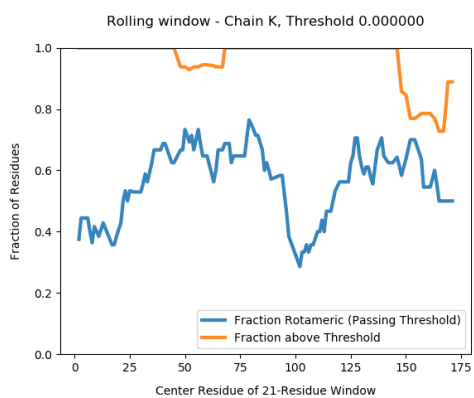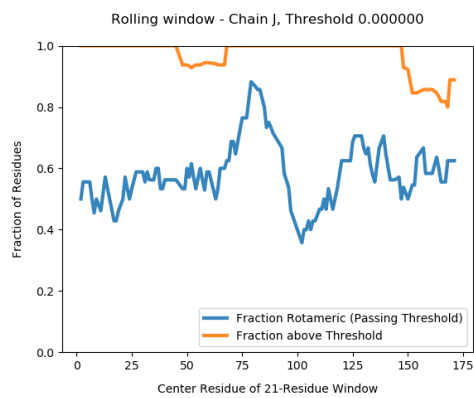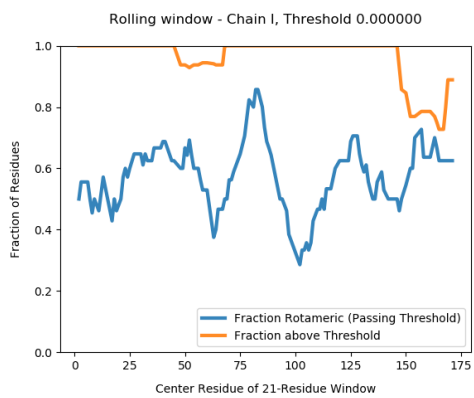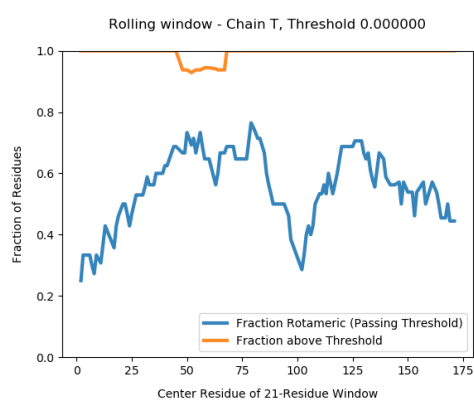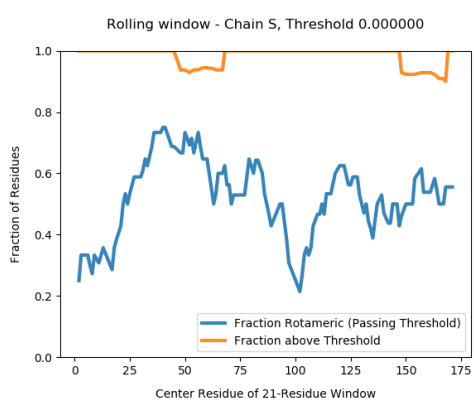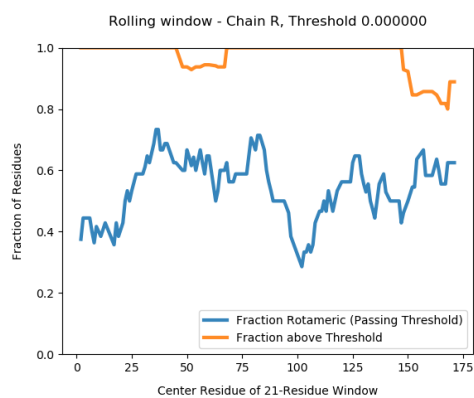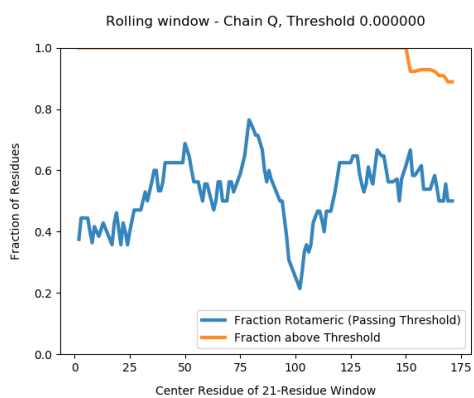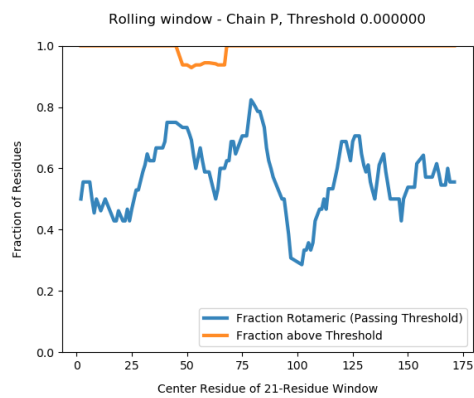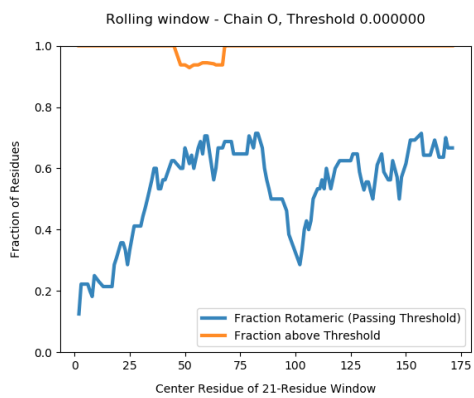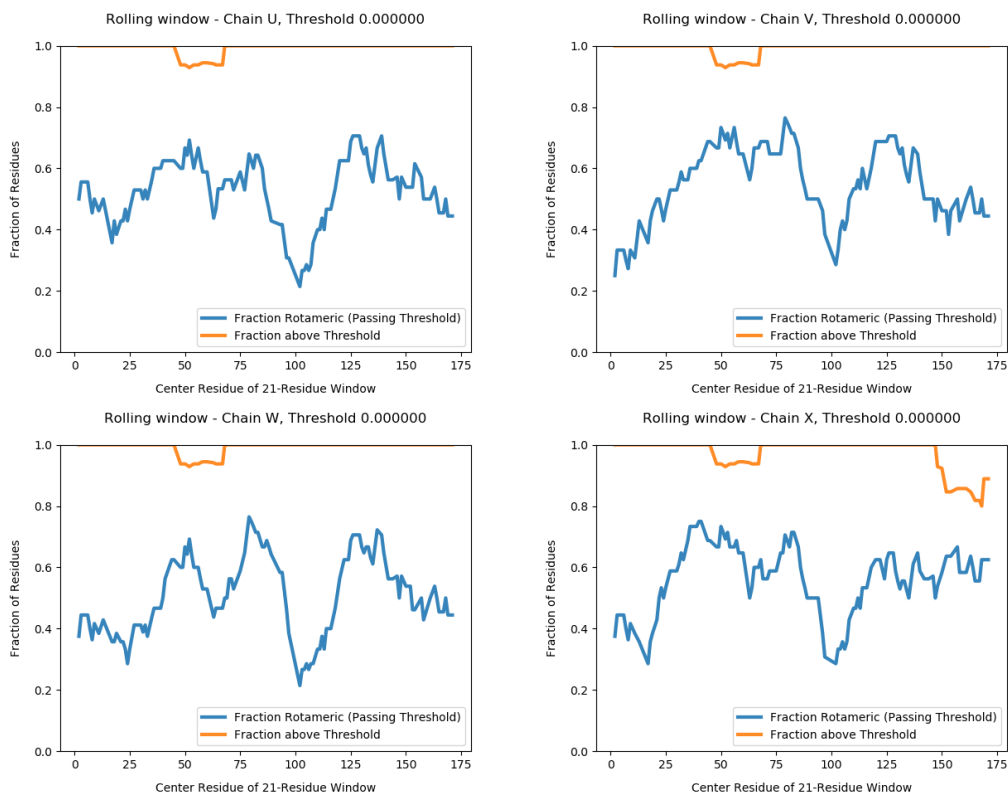
Automatic criteria: The validation is OK if the EMRinger score and Max. Zscore are larger than 1.

WARNINGS: 1 warnings

1. **The EMRinger score is smaller than 1, it is 0.892.**

13.7   Level A.g DAQ validation

Explanation:
DAQ [Terashi et al., 2022] is a computational tool using deep learning that can estimate the residue-wise local quality for protein models from cryo-Electron Microscopy maps. The method calculates the likelihood that a given density feature corresponds to an aminoacid, atom, and secondary structure. These likelihoods are combined into a score that ranges from -1 (bad quality) to 1 (good quality).

104

Results:

Fig. 77 shows the histogram of the DAQ values. The mean and standard deviation were -0.0 and 0.3, respectively.
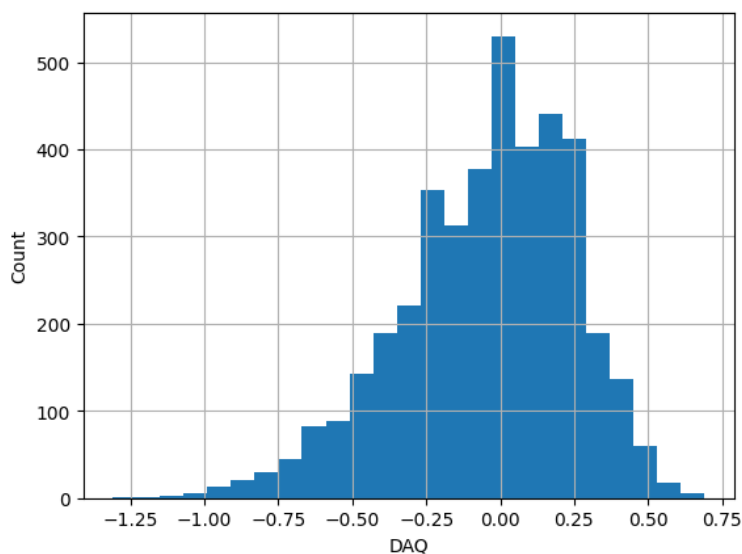
Figure 77: Histogram of the DAQ values.

The atomic model colored by DAQ can be seen in Fig. 78.

(a) View 1   (b) View 2   (c) View 3
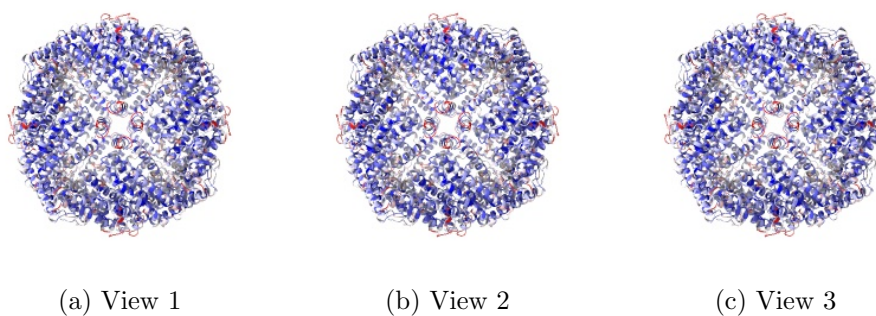
Figure 78: Atomic model colored by DAQ Views generated by ChimeraX at a the following X, Y, Z angles: View 1 (0,0,0), View 2 (90, 0, 0), View 3 (0, 90, 0).

Automatic criteria: The validation is OK if the average DAQ score is larger than 0.5.

WARNINGS: 1 warnings

1. **The average DAQ is smaller than 0.5.**

14 Workflow

Workflow file: `http://nolan.cnb.csic.es/cryoemworkflowviewer/workflow/637ca2bbcd57e45e88f6fabb7f6b1095a3ca0de6`
SHA256 hash: 5d8c5ff8948f4ac986f5d43f819515e25668bfdcd954b8fb8c41d15cdf00fda2

Fig. 79 shows the image processing workflow followed in Scipion to achieve these results.

Figure 79: Image processing workflow in Scipion to achieve these results.

STATUS: Cannot be automatically evaluated

15 Other experimental techniques

15.1 O.b SAXS

SAXS file: /home/coss/ScipionUserData/projects/Example_10248_Scipion3/-
SASDE55.dat
SHA256 hash: 69241b1c720a4d6568943daa7d762d9923c2375c75cd9b1ec9bc128e5a3bfce2

Explanation:

The method in [Jiménez et al., 2019] compares the expected energy profile from the reconstructed map to the one obtained by a SAXS experiment.

Results:

The radius of gyration was 51.0 Å. The χ^2 between the simulated curve and the experimental one was 27.6. Fig. 80 shows the two SAXS profiles for comparison.
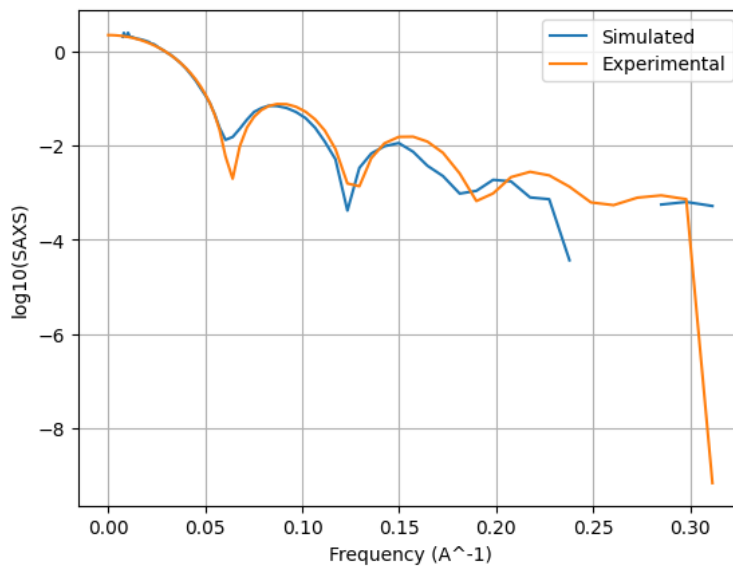
Figure 80: Simulated and experimental SAXS curves.

STATUS: Cannot be automatically evaluated

References

[Rosenthal and Henderson, 2003] Rosenthal, P. B. and Henderson, R. (2003). Optimal determination of particle orientation, absolute hand, and contrast loss in single particle electron-cryomicroscopy. *J. Molecular Biology*, 333:721–745.

[Kaur et al., 2021] Kaur, S., Gomez-Blanco, J., Khalifa, A. A., Adi-narayanan, S., Sanchez-Garcia, R., Wrapp, D., McLellan, J. S., Bui, K. H., and Vargas, J. (2021). Local computational methods to improve the interpretability and analysis of cryo-EM maps. *Nature Communications*, 12(1):1–12.

[Ramírez-Aportela et al., 2019] Ramírez-Aportela, E., Mota, J., Conesa, P., Carazo, J. M., and Sorzano, C. O. S. (2019). DeepRes: a new deep-learning- and aspect-based local resolution method for electron-microscopy maps. *IUCRj*, 6:1054–1063.

[Sorzano et al., 2017] Sorzano, C. O. S., Vargas, J., Oton, J., Abrishami, V., de la Rosa-Trevin, J. M., Gomez-Blanco, J., Vilas, J. L., Marabini, R., and Carazo, J. M. (2017). A review of resolution measures and related aspects in 3D electron microscopy. *Progress in biophysics and molecular biology*, 124:1–30.

[Beckers and Sachse, 2020] Beckers, M. and Sachse, C. (2020). Permutation testing of fourier shell correlation for resolution estimation of cryo-em maps. *J. Structural Biology*, 212(1):107579.

[Cardone et al., 2013] Cardone, G., Heymann, J. B., and Steven, A. C. (2013). One number does not fit all: Mapping local variations in resolution in cryo-em reconstructions. *J. Structural Biology*, 184:226–236.

[Kucukelbir et al., 2014] Kucukelbir, A., Sigworth, F. J., and Tagare, H. D. (2014). Quantifying the local resolution of cryo-EM density maps. *Nature Methods*, 11:63–65.

[Vilas et al., 2018] 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., and Sorzano, C. O. S. (2018). MonoRes: automatic and unbiased estimation of local resolution for electron microscopy maps. *Structure*, 26:337–344.

[Vilas et al., 2020] Vilas, J. L., Tagare, H. D., Vargas, J., Carazo, J. M., and Sorzano, C. O. S. (2020). Measuring local-directional resolution and local anisotropy in cryo-EM maps. *Nature communications*, 11:55.

[Sorzano et al., 2014] Sorzano, C. O. S., Vargas, J., de la Rosa-Trevín, J. M., Zaldívar-Peraza, A., Otón, J., Abrishami, V., Foche, I., Marabini, R., Caffarena, G., and Carazo, J. M. (2014). Outlier detection for single particle analysis in electron microscopy. In *Proc. Intl. Work-Conference on Bioinformatics and Biomedical Engineering, IWBBIO*, page 950.

[Punjani et al., 2017] Punjani, A., Brubaker, M. A., and Fleet, D. J. (2017). Building proteins in a day: Efficient 3D molecular structure estimation with electron cryomicroscopy. *IEEE Trans. Pattern Analysis & Machine Intelligence*, 39:706–718.

[Sorzano et al., 2015] Sorzano, C. O. S., Vargas, J., de la Rosa-Trevín, J. M., Otón, J., Álvarez-Cabrera, A. L., Abrishami, V., Sesmero, E., Marabini, R., and Carazo, J. M. (2015). A statistical approach to the initial volume problem in single particle analysis by electron microscopy. *J. Structural Biology*, 189:213–219.

[Méndez et al., 2021] Méndez, J., Garduño, E., Carazo, J. M., and Sorzano, C. O. S. (2021). Identification of incorrectly oriented particles in Cryo-EM single particle analysis. *J. Structural Biology*, 213:107771.

[Vargas et al., 2017] Vargas, J., Melero, R., Gómez-Blanco, J., Carazo, J. M., and Sorzano, C. O. S. (2017). Quantitative analysis of 3D alignment quality: its impact on soft-validation, particle pruning and homogeneity analysis. *Scientific Reports*, 7:6307.

[Vargas et al., 2016] Vargas, J., Otón, J., Marabini, R., Carazo, J. M., and Sorzano, C. O. S. (2016). Particle alignment reliability in single particle electron cryomicroscopy: a general approach. *Scientific Reports*, 6:21626.

[Scheres, 2012] Scheres, S. H. W. (2012). A Bayesian view on cryo-EM structure determination. *J. Molecular Biology*, 415:406–418.

[Punjani et al., 2020] Punjani, A., Zhang, H., and Fleet, D. J. (2020). Non-uniform refinement: adaptive regularization improves single-particle cryo-EM reconstruction. *Nature Methods*, 17(12):1214–1221.

[Heymann, 2015] Heymann, B. (2015). Validation of 3DEM reconstructions: The phantom in the noise. *AIMS Biophysics*, 2:21–35.

[Naydenova and Russo, 2017] Naydenova, K. and Russo, C. J. (2017). Measuring the effects of particle orientation to improve the efficiency of electron cryomicroscopy. *Nature communications*, 8:629.

[Baldwin and Lyumkis, 2020] Baldwin, P. R. and Lyumkis, D. (2020). Nonuniformity of projection distributions attenuates resolution in Cryo-EM. *Progress in Biophysics and Molecular Biology*, 150:160–183.

[Sanchez-Garcia et al., 2020] Sanchez-Garcia, R., Segura, J., Maluenda, D., Sorzano, C. O. S., and Carazo, J. M. (2020). MicrographCleaner: A python package for cryo-EM micrograph cleaning using deep learning. *J. Structural Biology*, 210:107498.

[Pintilie et al., 2020] Pintilie, G., Zhang, K., Su, Z., Li, S., Schmid, M. F., and Chiu, W. (2020). Measurement of atom resolvability in cryo-em maps with q-scores. *Nature methods*, 17(3):328–334.

[Ramírez-Aportela et al., 2021] Ramírez-Aportela, E., Maluenda, D., Fonseca, Y. C., Conesa, P., Marabini, R., Heymann, J. B., Carazo, J. M., and Sorzano, C. O. 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.

[Herzik et al., 2019] Herzik, M. A., Fraser, J. S., and Lander, G. C. (2019). A multi-model approach to assessing local and global cryo-EM map quality. *Structure*, 27(2):344–358.e3.

[Afonine et al., 2018] Afonine, P. V., Klaholz, B. P., Moriarty, N. W., Poon, B. K., Sobolev, O. V., Terwilliger, T. C., Adams, P. D., and Urzhumtsev, A. (2018). New tools for the analysis and validation of cryo-EM maps and atomic models. *Acta Crystallographica D, Struct. Biol.*, 74:814–840.

[Barad et al., 2015] Barad, B. A., Echols, N., Wang, R. Y.-R., Cheng, Y., DiMaio, F., Adams, P. D., and Fraser, J. S. (2015). EMRinger: side chain-directed model and map validation for 3D cryo-electron microscopy. *Nature Methods*, 12(10):943–946.

[Terashi et al., 2022] Terashi, G., Wang, X., Subramaniya, S.R.M.V., Tesmer, J.J.G. and Kihara, D. (2022). Residue-Wise Local Quality Estimation for Protein Models from Cryo-EM Maps. (submitted).

[Jiménez et al., 2019] Jiménez, A., Jonic, S., Majtner, T., Oón, J., Vilas, J. L., Maluenda, D., Mota, J., Ramírez-Aportela, E., Martínez, M., Rancel, Y., Segura, J., Sánchez-García, R., Melero, R., Del Caño, L., Conesa, P., Skjaerven, L., Marabini, R., Carazo, J. M., and Sorzano, C. O. S. (2019). Validation of electron microscopy initial models via small angle X-ray scattering curves. *Bioinformatics*, 35:2427–2433.

[Henderson et al., 2011] Henderson, R., Chen, S., Chen, J. Z., Grigorieff, N., Passmore, L. A., Ciccarelli, L., Rubinstein, J. L., Crowther, R. A., Stewart, P. L., and Rosenthal, P. B. (2011). Tilt-pair analysis of images from a range of different specimens in single-particle electron cryomicroscopy. *J. Molecular Biology*, 413(5):1028–1046.

Validation report of Level(s) 0, 1, A

I²PC Validation server

February 25, 2022
8:40pm

Abstract

The map seems to be well centered. There is no problem with the suggested threshold. There seems to be a problem with the map's background (see Sec. 2.3). The resolution does not seem to be uniform in all directions (see Sec. 4.6). According to FSC-Q, it seems that there is a mismatch between the map and its model (see Sec. 6.2). According to phenix, it seems that there might be some mismatch between the map and its model (see Sec. 6.4).

The average resolution of the map estimated by various methods goes from 0.3Å to 10.2Å with an average of 4.9Å. The resolution provided by the user was 3.3Å. The resolution reported by the user may be overestimated.

The overall score (passing tests) of this report is 14 out of 20 evaluable items.

0.a Mass analysis	Sec. 2.1	OK
0.b Mask analysis	Sec. 2.2	OK
0.c Background analysis	Sec. 2.3	2 warnings
0.d B-factor analysis	Sec. 2.4	OK
0.e DeepRes	Sec. 2.5	1 warnings
0.f LocBfactor	Sec. 2.6	OK
0.g LocOccupancy	Sec. 2.7	OK
0.h DeepHand	Sec. 2.8	OK
1.a Global resolution	Sec. 4.1	1 warnings
1.b FSC permutation	Sec. 4.2	OK
1.c Blocres	Sec. 4.3	OK
1.d Resmap	Sec. 4.4	Could not be measured
1.e MonoRes	Sec. 4.5	OK
1.f MonoDir	Sec. 4.6	2 warnings
1.g FSO	Sec. 4.7	OK
1.h FSC3D	Sec. 4.8	Could not be measured
A.a MapQ	Sec. 6.1	OK
A.b FSC-Q	Sec. 6.2	1 warnings
A.d Map-Model Guinier	Sec. 6.3	OK
A.e Phenix validation	Sec. 6.4	1 warnings
A.f EMRinger	Sec. 6.5	OK
A.g DAQ	Sec. 6.6	OK

Summary of the warnings across sections.
If it is empty below this point, it means that there are no warnings.

Section 2.3 (0.c Background analysis)

1. **The null hypothesis that the background mean is 0 has been rejected because the p-value of the comparison is smaller than 0.001**

2. **There is a significant proportion of outlier values in the background (cdf5 ratio=4280.03)**

Section 2.5 (0.e DeepRes)

1. **The reported resolution, 3.30 Å, is particularly with respect to the local resolution distribution. It occupies the 0.03 percentile**

Section 4.1 (1.a Global resolution)

1. **The reported resolution, 3.30 Å, is particularly high with respect to the resolution calculated by the FSC, 7.54 Å**

Section 4.6 (1.f MonoDir)

1. **The distribution of best resolution is not uniform in all directions. The associated p-value is 0.000000.**

2. **The resolution reported by the user, 3.30Å, is at least 80% smaller than the average directional resolution, 7.31 Å.**

Section 6.2 (A.b FSC-Q)

1. **The percentage of voxels that have a FSC-Qr larger than 1.5 in absolute value is 10.2, that is larger than 10%**

Section 6.4 (A.e Phenix validation)

1. **The percentage of residues that have a cross-correlation below 0.5 is 20.8, that is larger than 10%**

Contents

1 Input data

Input map: /home/coss/data/Dropbox/Aplicaciones/ShareLaTeX/MapValidation/-
EMDB11337/emd_11337.map
SHA256 hash: d969dcfa8853ce92e8d9932e578ff2db49e7d00d1cfa3921607bcbbd4ff3cc23
Voxel size: 1.047000 (Å)
Visualization threshold: 0.165000
Resolution estimated by user: 3.300000

Orthogonal slices of the input map
Explanation:
In the orthogonal slices of the map, the noise outside the protein should not
have any structure (stripes going out, small blobs, particularly high or low
densities, ...)

Results:
See Fig. 1.

(a) X Slice 216   (b) Y Slice 216   (c) Z Slice 216

Figure 1: Central slices of the input map in the three dimensions

Orthogonal slices of maximum variance of the input map
Results:
See Fig. 2.

(a) X Slice 231   (b) Y Slice 225   (c) Z Slice 213

Figure 2: Slices of maximum variation in the three dimensions

Orthogonal projections of the input map

Explanation:

In the projections there should not be stripes (this is an indication of directional overweighting, or angular attraction), and there should not be a dark halo around or inside the structure (this is an indication of incorrect CTF correction or the reconstruction of a biased map).

Results:

See Fig. 3.

(a) X Projection   (b) Y Projection   (c) Z Projection

Figure 3: Projections in the three dimensions

Isosurface views of the input map

Explanation:

An isosurface is the surface of all points that have the same gray value. In these views there should not be many artifacts or noise blobs around the map.

Results:
See Fig. 4.

(a) View 1　　　　　　(b) View 2　　　　　　(c) View 3

Figure 4: Isosurface at threshold=0.165000. Views generated by ChimeraX at a the following X, Y, Z angles: View 1 (0,0,0), View 2 (90, 0, 0), View 3 (0, 90, 0).

Orthogonal slices of maximum variance of the mask
Explanation:
The mask has been calculated at the suggested threshold 0.165000, the largest connected component was selected, and then dilated by 2Å.

Results:
See Fig. 5.

(a) X Slice 216   (b) Y Slice 216   (c) Z Slice 216

Figure 5: Slices of maximum variation in the three dimensions of the mask

2 Level 0 analysis

2.1 Level 0.a Mass analysis

Explanation:

The reconstructed map must be relatively well centered in the box, and there should be at least 30Å (the exact size depends on the CTF) on each side to make sure that the CTF can be appropriately corrected.

Results:

The space from the left and right in X are 91.09 and 136.11 Å, respectively. There is a decentering ratio (abs(Right-Left)/Size)% of 9.95%

The space from the left and right in Y are 145.53 and 151.81 Å, respectively. There is a decentering ratio (abs(Right-Left)/Size)% of 1.39%

The space from the left and right in Z are 146.58 and 162.28 Å, respectively. There is a decentering ratio (abs(Right-Left)/Size)% of 3.47%

The center of mass is at (x,y,z)=(212.02,228.50,203.12). The decentering of the center of mass (abs(Center)/Size)% is 0.92, 2.89, and 2.98, respectively.%

Automatic criteria: The validation is OK if 1) the decentering and

center of mass less than 20% of the map dimensions in all directions, and 2) the extra space on each direction is more than 20% of the map dimensions.

STATUS: OK

2.2   Level 0.b Mask analysis

Explanation:
The map at the suggested threshold should have most of its mass concentrated in a single connected component. It is normal that after thresholding there are a few thousands of very small, disconnected noise blobs. However, there total mass should not exceed 10%. The raw mask (just thresholding) and the mask constructed for the analysis (thresholding + largest connected component + dilation) should significantly overlap. Overlap is defined by the overlapping coefficient (size(Raw AND Constructed)/size(Raw)) that is a number between 0 and 1, the closer to 1, the more they agree.

Results:

Raw mask: At threshold 0.165000, there are 874 connected components with a total number of voxels of 341566 and a volume of 392025.83 Å^3 (see Fig. 6). The size and percentage of the total number of voxels for the raw mask are listed below (up to 95% of the mass), the list contains (No. voxels (volume in Å^3), percentage, cumulatedPercentage):

(338472 (388474.75), 99.09, 99.09)

Number of components to reach 95% of the mass: 1

The average size of the remaining 873 components is 3.54 voxels (1.15 Å^3). Their size go from 248 voxels (284.64 Å^3) to 1 voxels (1.15 Å^3).
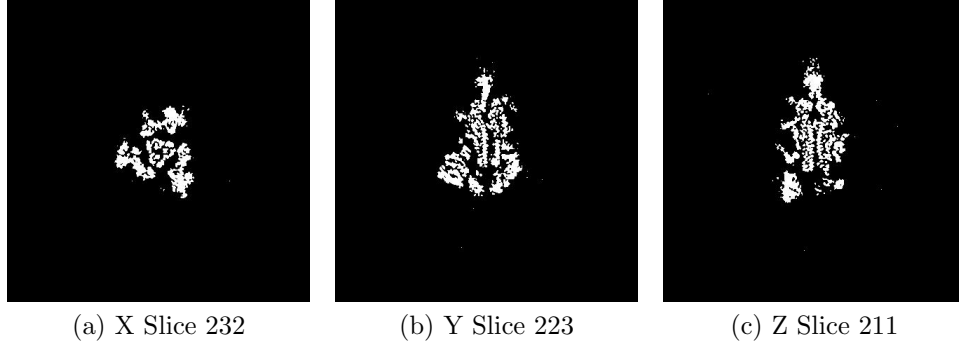
The slices of the raw mask can be seen in Fig. 6.

9

(a) X Slice 232   (b) Y Slice 223   (c) Z Slice 211

Figure 6: Maximum variance slices in the three dimensions of the raw mask

The following table shows the variation of the mass enclosed at different thresholds (see Fig. 7):

Threshold	Voxel mass	Molecular mass(kDa)	# Aminoacids
0.0622	1664392.00	1582.66	14387.84
0.1245	456945.00	434.51	3950.06
0.1867	300097.00	285.36	2594.19
0.2490	211046.00	200.68	1824.39
0.3112	149710.00	142.36	1294.17
0.3734	107653.00	102.37	930.61
0.4357	79118.00	75.23	683.94
0.4979	58725.00	55.84	507.65
0.5601	43849.00	41.70	379.05
0.6224	32674.00	31.07	282.45
0.6846	24175.00	22.99	208.98
0.7469	17504.00	16.64	151.31
0.8091	12522.00	11.91	108.25
0.8713	8618.00	8.19	74.50
0.9336	5718.00	5.44	49.43
0.9958	3462.00	3.29	29.93
1.0580	1918.00	1.82	16.58
1.1203	931.00	0.89	8.05
1.1825	384.00	0.37	3.32
1.2448	167.00	0.16	1.44
1.3070	58.00	0.06	0.50
1.3692	20.00	0.02	0.17
1.4315	6.00	0.01	0.05
1.4937	3.00	0.00	0.03
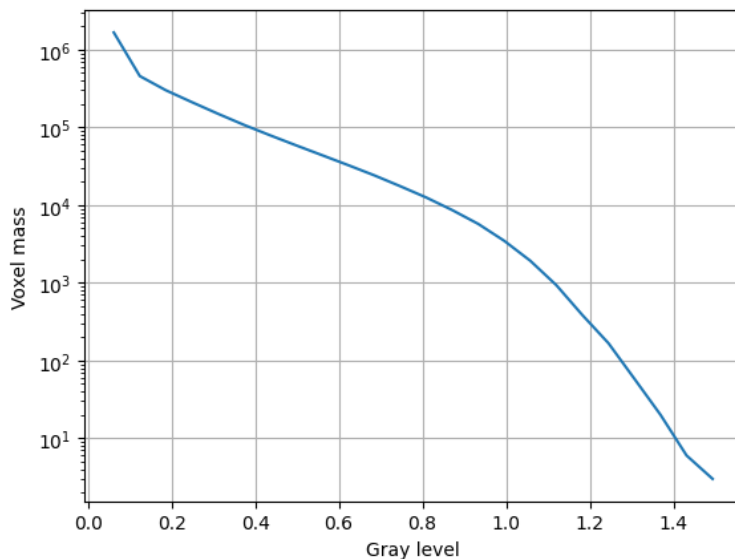
Figure 7: Voxel mass as a function of the gray level.

Constructed mask: After keeping the largest component of the previous mask and dilating it by 2Å, there is a total number of voxels of 799000 and a volume of 917036.93 Å^3. The overlap between the raw and constructed mask is 0.99.

Automatic criteria: The validation is OK if 1) to keep 95% of the mass we need to keep at most 5 connected components; and 2) the average volume of the blobs outside the given threshold has a size smaller than 5Å^3; and 3) the overlap between the raw mask and the mask constructed for the analysis is larger than 75%.

STATUS: OK

2.3 Level 0.c Background analysis

Explanation:
Background is defined as the region outside the macromolecule mask. The background mean should be zero, and the number of voxels with a very low or very high value (below 5 standard deviations of the noise) should be very

small and they should be randomly distributed without any specific structure. Sometimes, you can see some structure due to the symmetry of the structure.

Results:

The null hypothesis that the background mean is 0 was tested with a one-sample Student's t-test. The resulting t-statistic and p-value were -839.68 and 0.000000, respectively.

The mean and standard deviation (sigma) of the background were -0.002207 and 0.023486. The percentage of background voxels whose absolute value is larger than 5 times the standard deviation is 0.25 % (see Fig. 8). The same percentage from a Gaussian would be 0.000057% (ratio between the two percentages: 4280.030049).

Slices of the background beyond 5*sigma can be seen in Fig. 8.

(a) X Slice 240   (b) Y Slice 218   (c) Z Slice 220

Figure 8: Maximum variance slices in the three dimensions of the parts of the background beyond 5*sigma

Automatic criteria: The validation is OK if 1) the p-value of the null hypothesis that the background has 0 mean is larger than 0.001; and 2) the number of voxels above or below 5 sigma is smaller than 20 times the amount expected for a Gaussian with the same standard deviation whose mean is 0.

WARNINGS: 2 warnings

12

2.4 Level 0.d B-factor analysis

Explanation:
The B-factor line [Rosenthal and Henderson, 2003] fitted between 15Åand the resolution reported should have a slope that is between 0 and 300 Å^2.

Results:
Fig. 9 shows the logarithm (in natural units) of the structure factor (the module squared of the Fourier transform) of the experimental map, its fitted line, and the corrected map. The estimated B-factor was -89.2. The fitted line was $\log(|F|^2) = -22.3/R^2 + (-10.9)$.

Figure 9: Guinier plot. The X-axis is the square of the inverse of the resolution in Å.

(a) X Slice 165   (b) Y Slice 225   (c) Z Slice 211

Figure 10: Slices of maximum variation in the three dimensions of the B-factor corrected map

Automatic criteria: The validation is OK if the B-factor is in the range [-300,0].

STATUS: OK

2.5 Level 0.e Local resolution with DeepRes

Explanation:
DeepRes [Ramírez-Aportela et al., 2019] measures the local resolution using a neural network that has been trained on the appearance of atomic structures at different resolutions. Then, by comparing the local appearance of the input map to the appearance of the atomic structures a local resolution label can be assigned.

Results:

Fig. 11 shows the histogram of the local resolution according to DeepRes. Some representative percentiles are:

14

Percentile	Resolution(Å)
2.5%	3.88
25%	4.34
50%	4.64
75%	5.02
97.5%	6.03

The reported resolution, 3.30 Å, is at the percentile 0.0. Fig. 12 shows some representative views of the local resolution.

Figure 11: Histogram of the local resolution according to deepres.

(a) View 1   (b) View 2   (c) View 3

Figure 12: Local resolution according to DeepRes. Views generated by ChimeraX at a the following X, Y, Z angles: View 1 (0,0,0), View 2 (90, 0, 0), View 3 (0, 90, 0).

Automatic criteria: The validation is OK if the percentile of the user provided resolution is larger than 0.1% of the percentile of the local resolution as estimated by DeepRes.

WARNINGS: 1 warnings

1. **The reported resolution, 3.30 Å, is particularly with respect to the local resolution distribution. It occupies the 0.03 percentile**

2.6 Level 0.f Local B-factor

Explanation:
LocBfactor [Kaur et al., 2021] estimates a local resolution B-factor by decomposing the input map into a local magnitude and phase term using the spiral transform.

Results:

Fig. 13 shows the histogram of the local B-factor according to LocBfactor. Some representative percentiles are:

Percentile	Local B-factor (Å^{-2})
2.5%	-307.82
25%	-262.94
50%	-239.10
75%	-212.06
97.5%	-149.18

Fig. 14 shows some representative views of the local B-factor.

Figure 13: Histogram of the local B-factor according to LocBfactor.

(a) View 1   (b) View 2   (c) View 3

Figure 14: Local B-factor according to LocBfactor. Views generated by ChimeraX at a the following X, Y, Z angles: View 1 (0,0,0), View 2 (90, 0, 0), View 3 (0, 90, 0).
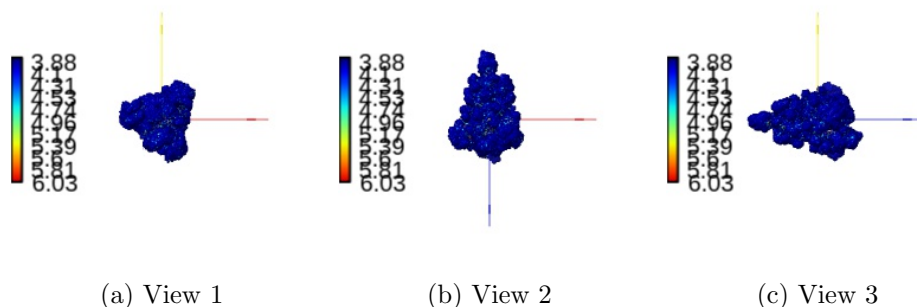
Automatic criteria: The validation is OK if the median B-factor is in the range [-300,0].

STATUS: OK

2.7   Level 0.g Local Occupancy

Explanation:
LocOccupancy [Kaur et al., 2021] estimates the occupancy of a voxel by the macromolecule.

Results:

Fig. 15 shows the histogram of the local occupancy according to LocOccupancy. Some representative percentiles are:

Percentile	Local Occupancy [0-1]
2.5%	0.23
25%	0.55
50%	0.86
75%	0.95
97.5%	1.00
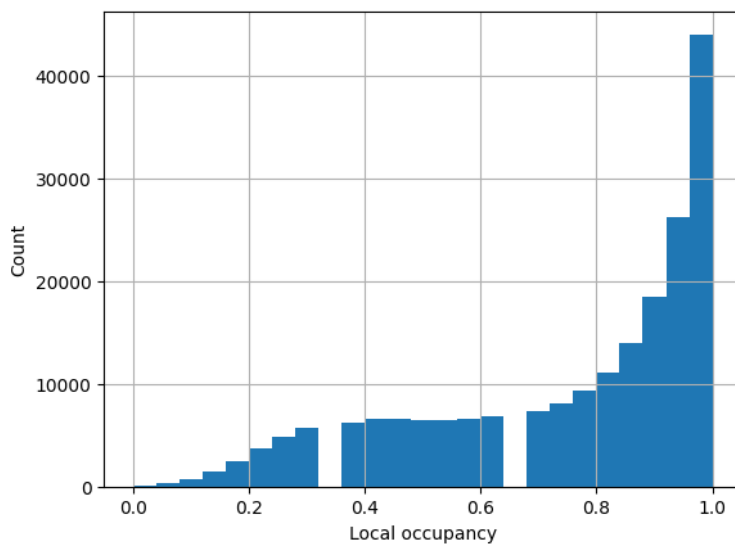
18

Fig. 16 shows some representative views of the local occupancy.

Figure 15: Histogram of the local occupancy according to LocOccupancy.

|(a) View 1|(b) View 2|(c) View 3|

Figure 16: Local occupancy according to LocOccupancy. Views generated by ChimeraX at a the following X, Y, Z angles: View 1 (0,0,0), View 2 (90, 0, 0), View 3 (0, 90, 0).

Automatic criteria: The validation is OK if the median occupancy is larger than 50%.

STATUS: OK

2.8   Level 0.h Hand correction

Explanation:
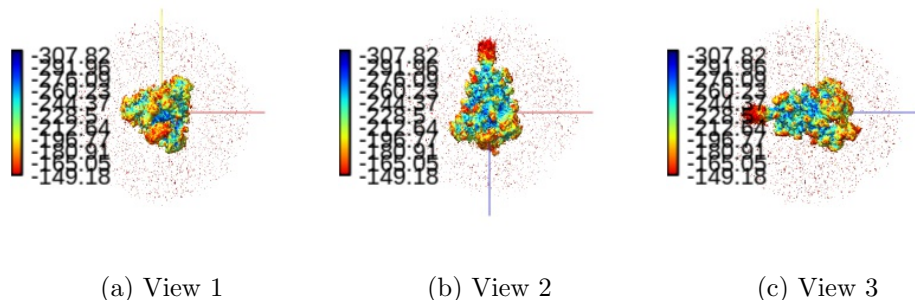Deep Hand determines the correction of the hand for those maps with a resolution smaller than 5Å. The method calculates a value between 0 (correct hand) and 1 (incorrect hand) using a neural network to assign its hand.

Results:

Deep hand assigns a score of 0.254 to the input volume.
Automatic criteria: The validation is OK if the deep hand score is smaller than 0.5.

STATUS: OK

3   Half maps

Half map 1: /home/coss/data/Dropbox/Aplicaciones/ShareLaTeX/MapValidation/-EMDB11337/emd_11337_half_map_1.map
SHA256 hash: 17945f1afcb4d373fe3b41d2a904ea0721dc63f78fa3956d012dda9e537390a6

Half map 2: /home/coss/data/Dropbox/Aplicaciones/ShareLaTeX/MapValidation/-EMDB11337/emd_11337_half_map_2.map
SHA256 hash: ded62a6026f96d71c7c487135471c7443aca06d1d32f2afdf909d3152f7f79e2

Slices of the first half map can be seen in Fig. 17.
Slices of the second half map can be seen in Fig. 18.
Slices of the difference between both maps can be seen in Fig. 19. There should not be any structure in this difference. Sometimes some patterns are seen if the map is symmetric.

(a) X Slice 386   (b) Y Slice 407   (c) Z Slice 30

Figure 17: Slices of maximum variation in the three dimensions of Half 1
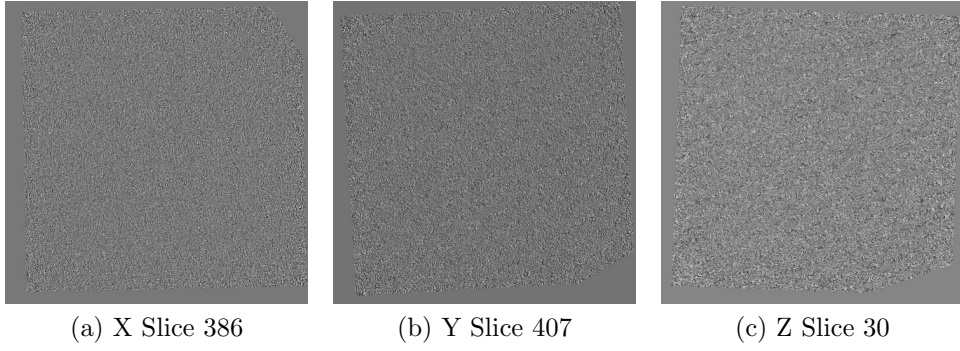

(a) X Slice 222   (b) Y Slice 224   (c) Z Slice 31

Figure 18: Slices of maximum variation in the three dimensions of Half 2

(a) X Slice 386　　　　(b) Y Slice 407　　　　(c) Z Slice 31

Figure 19: Slices of maximum variation in the three dimensions of the difference Half1-Half2.

4　Level 1 analysis

4.1　Level 1.a Global resolution

Explanation: The Fourier Shell Correlation (FSC) between the two half maps is the most standard method to determine the global resolution of a map. However, other measures exist such as the Spectral Signal-to-Noise Ratio and the Differential Phase Residual. There is a long debate about the right thresholds for these measures. Probably, the most clear threshold is the one of the SSNR (SSNR=1). For the DPR we have chosen 103.9° and for the FSC, the standard 0.143. For a deep discussion of all these thresholds, see [Sorzano et al., 2017]. Note that these thresholds typically result in resolution values that are at the lower extreme of the local resolution range, meaning that this resolution is normally in the first quarter. It should not be understood as the average resolution of the map.

Except for the noise, the FSC and DPR should be approximately monotonic. They should not have any "coming back" behavior. If they have, this is typically due to the presence of a mask in real space or non-linear processing.

Results:
Fig. 20 shows the FSC and the 0.143 threshold. The resolution according to the FSC is 7.54Å. The map information is well preserved (FSC>0.9) up to

41.58Å.

Fig. 21 shows the DPR and the 103.9° threshold. The resolution according to the DPR is 3.82Å.

Fig. 22 shows the SSNR and the SSNR=1 threshold. The resolution according to the SSNR is 3.79Å.

The mean resolution between the three methods is 5.05Å and its range is within the interval [3.79, 7.54]Å.
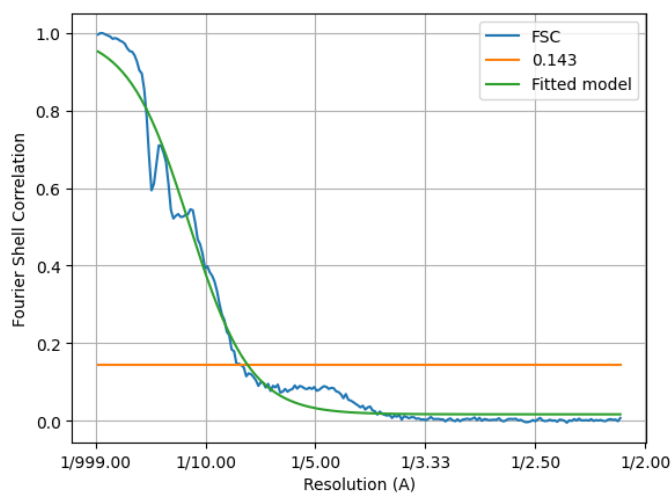
Figure 20: Fourier Shell correlation between the two halves.
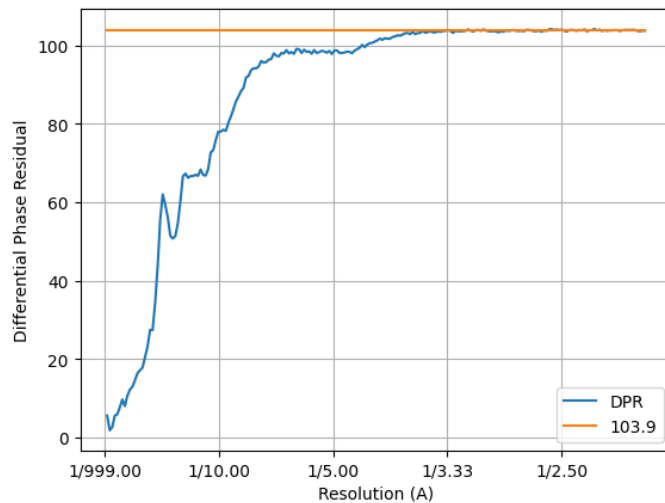
Figure 21: Differential Phase Residual between the two halves.
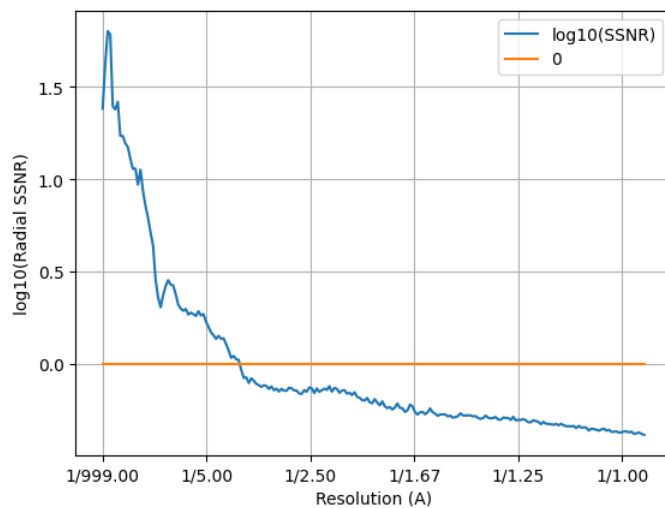
Figure 22: Spectral Signal-to-Noise Ratio estimated from the two halves.

Automatic criteria: The validation is OK if the user provided resolution is larger than 0.8 times the resolution estimated by 1) FSC, 2) DPR, and 3) SSNR.

24

WARNINGS: 1 warnings

1. **The reported resolution, 3.30 Å, is particularly high with respect to the resolution calculated by the FSC, 7.54 Å**

4.2   Level 1.b FSC permutation

Explanation:
This method [Beckers and Sachse, 2020] calculates a global resolution by formulating a hypothesis test in which the distribution of the FSC of noise is calculated from the two maps.

Results:

The resolution at 1% of FDR was 3.4. The estimated B-factor was -36.4. Fig. 23 shows the estimated FSC and resolution.

Figure 23: FSC and resolution estimated by a permutation test.

Automatic criteria: The validation is OK if the user provided resolution is larger than 0.8 times the resolution estimated by FSC permutation.

STATUS: OK

4.3 Level 1.c Local resolution with Blocres

Explanation:
This method [Cardone et al., 2013] computes a local Fourier Shell Correlation (FSC) between the two half maps.

Results:

26

Fig. 24 shows the histogram of the local resolution according to Blocres. Some representative percentiles are:

Percentile	Resolution(Å)
2.5%	3.19
25%	3.79
50%	5.34
75%	7.22
97.5%	9.38

The reported resolution, 3.30 Å, is at the percentile 7.0. Fig. 25 shows some representative views of the local resolution.

Figure 24: Histogram of the local resolution according to blocres.

(a) View 1　　　　　　(b) View 2　　　　　　(c) View 3

Figure 25: Local resolution according to Blocres. Views generated by ChimeraX at a the following X, Y, Z angles: View 1 (0,0,0), View 2 (90, 0, 0), View 3 (0, 90, 0).

Automatic criteria: The validation is OK if the percentile of the user provided resolution is larger than 0.1% of the percentile of the local resolution as estimated by BlocRes.

STATUS: OK

4.4   Level 1.d Local resolution with Resmap

Explanation:
This method [Kucukelbir et al., 2014] is based on a test hypothesis testing of the superiority of signal over noise at different frequencies.

Results:

ERROR: The protocol failed.

4.5   Level 1.e Local resolution with MonoRes

Explanation:
MonoRes [Vilas et al., 2018] evaluates the local energy of a point with respect to the distribution of energy in the noise. This comparison is performed at

multiple frequencies and for each one, the monogenic transformation separates the amplitude and phase of the input map. Then the energy of the amplitude within the map is compared to the amplitude distribution observed in the noise, and a hypothesis test is run for every voxel to check if its energy is signficantly above the level of noise.

Results:

Fig. 26 shows the histogram of the local resolution according to MonoRes. Some representative percentiles are:

Percentile	Resolution(Å)
2.5%	3.38
25%	5.09
50%	10.17
75%	14.74
97.5%	16.00

The reported resolution, 3.30 Å, is at the percentile 2.5. Fig. 27 shows some representative views of the local resolution
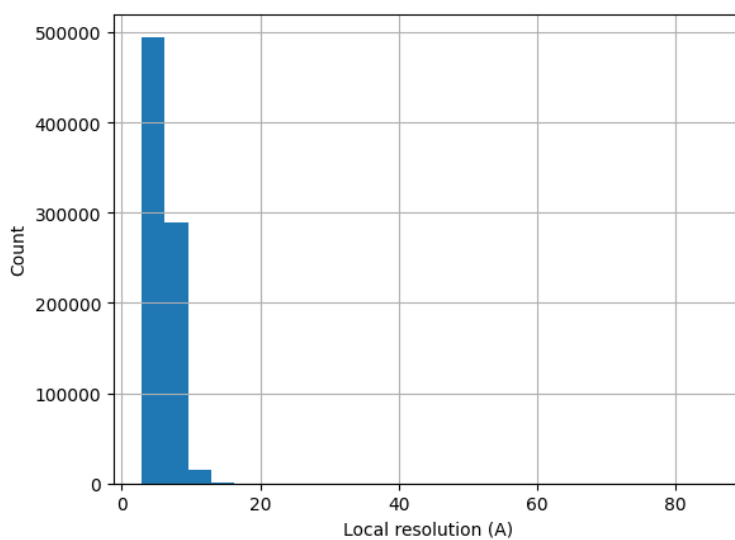

Figure 26: Histogram of the local resolution according to MonoRes.

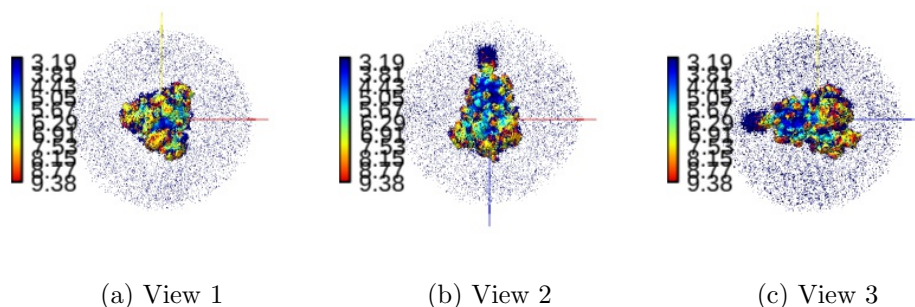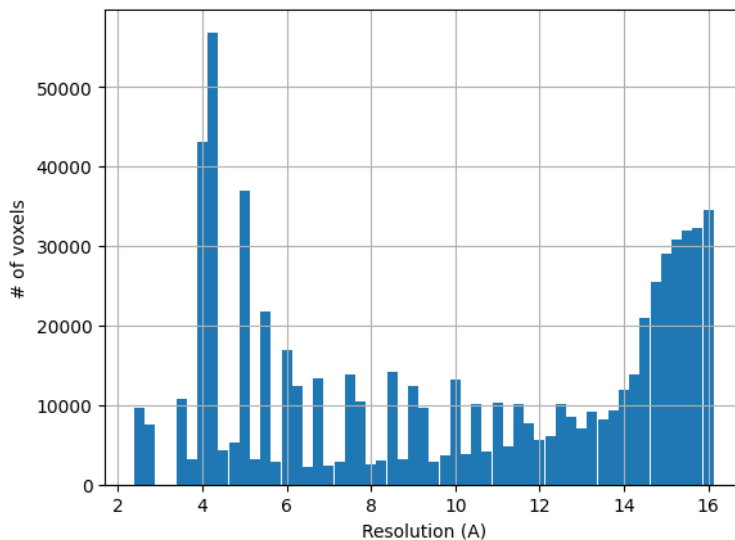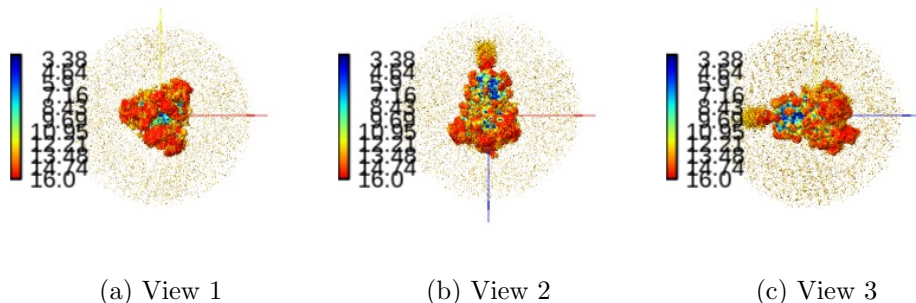(a) View 1   (b) View 2   (c) View 3

Figure 27: Local resolution according to Monores. Views generated by ChimeraX at a the following X, Y, Z angles: View 1 (0,0,0), View 2 (90, 0, 0), View 3 (0, 90, 0).

Automatic criteria: The validation is OK if the percentile of the user provided resolution is larger than 0.1% of the percentile of the local resolution as estimated by MonoRes.

STATUS: OK

4.6 Level 1.f Local and directional resolution with MonoDir

Explanation:
MonoDir [Vilas et al., 2020] extends the concept of local resolution to local and directional resolution by changing the shape of the filter applied to the input map. The directional analysis can reveal image alignment problems.

The histogram of best resolution voxels per direction (Directional Histogram 1D) shows how many voxels in the volume have their maximum resolution in that direction. Directions are arbitrarily numbered from 1 to N. This histogram should be relatively flat. We perform a Kolmogorov-Smirnov test to check its uniformity. If the null hypothesis is rejected, then the directional resolution is not uniform. It does not mean that it is wrong, and it could be caused by several reasons: 1) the angular distribution is not uniform, 2) there are missing directions, 3) there is some anisotropy in the data (including some preferential directional movement).

Ideally, the radial average of the minimum, maximum, and average res-

olution at each voxel (note that these are spatial radial averages) should be flat and as low as possible. If they show some slope, this is associated with inaccuracies in the angular assignment. These averages make sense when the shells are fully contained within the protein. As the shells approach the outside of the protein, these radial averages make less sense.

Results:

Fig. 28 shows the 1D directional histogram and Fig. 29 the 2D directional histogram. We compared the 1D directional histogram to a uniform distribution using a Kolmogorov-Smirnov test. The D statistic was 0.048545, and the p-value of the null hypothesis 0.000000.

The radial average of the minimum, maximum and average resolution at each voxel is shown in Fig. 30. The overall mean of the directional resolution is 7.31
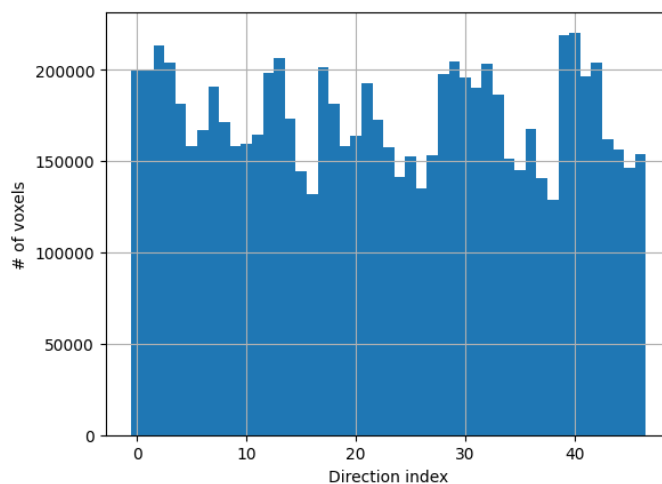

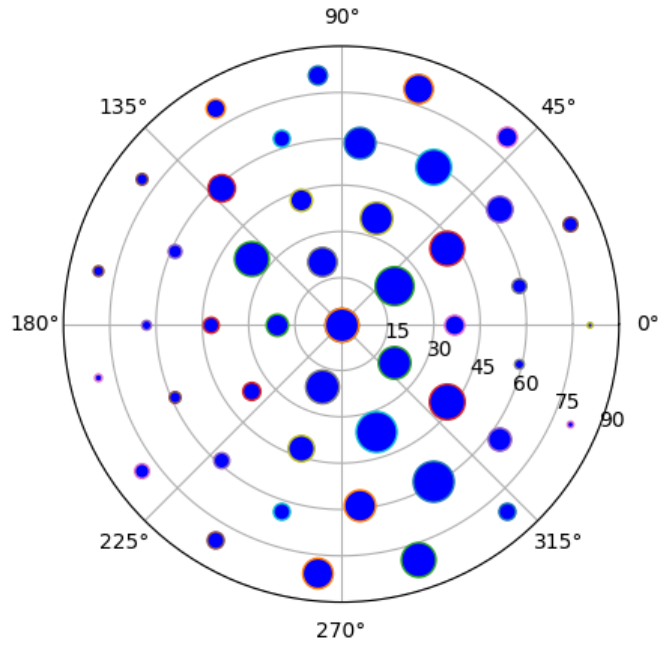Figure 28: Histogram 1D of the best direction at each voxel.

Figure 29: Histogram 2D of the best direction at each voxel. The azimuthal rotation is circular, while the tilt angle is the radius. The size of the point is proportional to the number of voxels whose maximum resolution is in that direction (this count can be seen in Fig. 28.

32
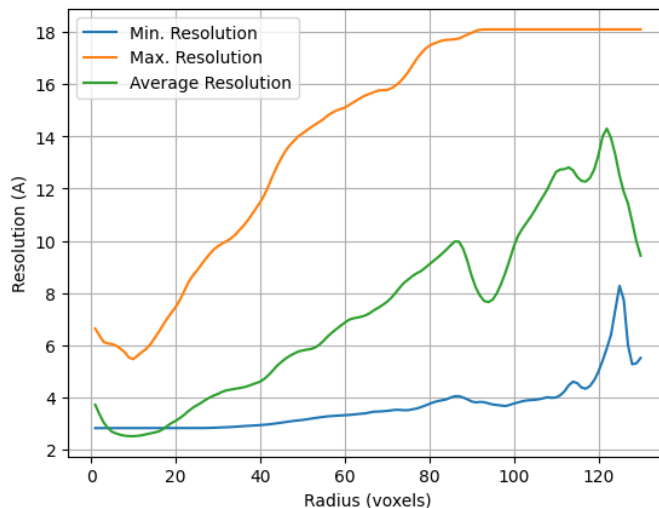
Figure 30: Radial averages (in space) of the minimum, maximum and average resolution at each voxel.

Automatic criteria: The validation is OK if 1) the null hypothesis that the directional resolution is not uniform is not rejected with a threshold of 0.001 for the p-value, and 2) the resolution provided by the user is not smaller than 0.8 times the average directional resolution.

WARNINGS: 2 warnings

1. **The distribution of best resolution is not uniform in all directions. The associated p-value is 0.000000.**
2. **The resolution reported by the user, 3.30Å, is at least 80% smaller than the average directional resolution, 7.31 Å.**

4.7   Level 1.g Fourier Shell Occupancy

Explanation:
This method calculates the anisotropy of the energy distribution in Fourier shells. This is an indirect measure of anisotropy of the angular distribution or the presence of heterogeneity. A natural threshold for this measure is 0.5. However, 0.9 and 0.1 are also interesting values that define the frequency at which the occupancy is 90% and 10%, respectively. This region is shaded in

the plot.
Results:

Fig. 31 shows the Fourier Shell Occupancy and its anisotropy. The directional resolution is shown in Fig. 32. The resolution according to the FSO is 3.39Å. Fourier shells are occupied at between 90 and than 10% in the range [3.58, 3.29]Å.
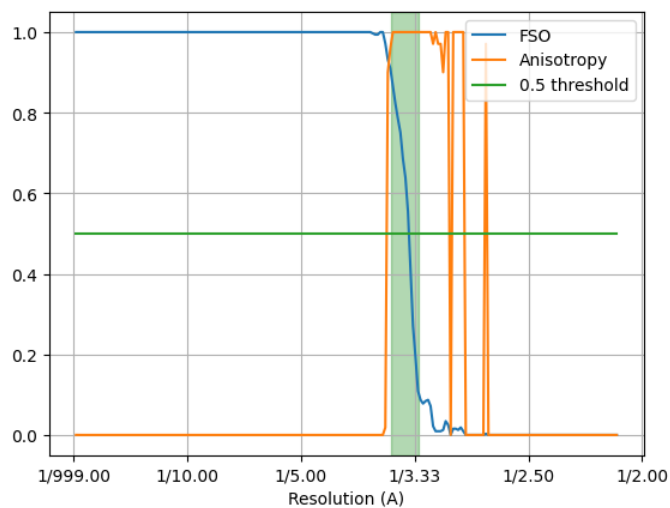
Figure 31: FSO and anisotropy.
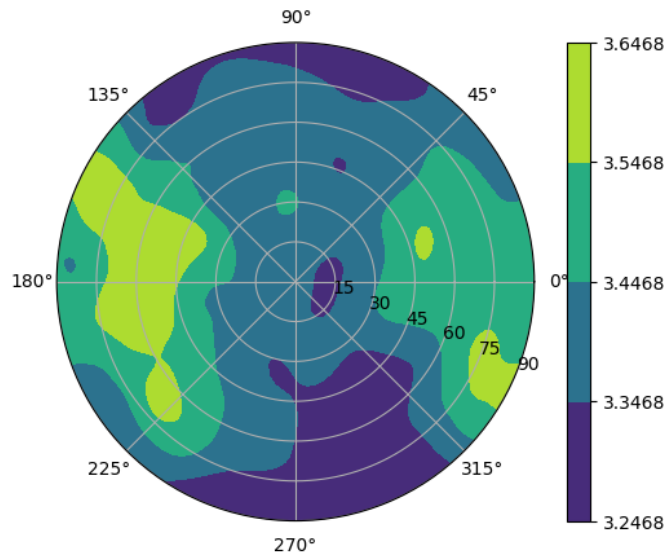
Figure 32: Directional resolution in the projection sphere.

Automatic criteria: The validation is OK if the resolution provided by the user is not smaller than 0.8 times the resolution estimated by the first cross of FSO below 0.5.

STATUS: OK

4.8 Level 1.h Fourier Shell Correlation 3D

Explanation:
This method analyzes the FSC in different directions and evaluates its homogeneity.
Results:

ERROR: The protocol failed.

5 Atomic model

Atomic model: /home/coss/data/Dropbox/Aplicaciones/ShareLaTeX/MapValidation/-EMDB11337/6zp7_updated_centered.pdb

See Fig. 33.

(a) View 1   (b) View 2   (c) View 3
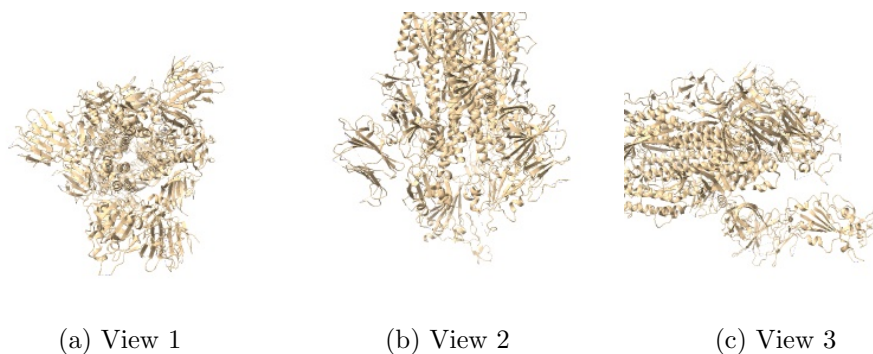
Figure 33: Input atomic model Views generated by ChimeraX at a the following X, Y, Z angles: View 1 (0,0,0), View 2 (90, 0, 0), View 3 (0, 90, 0).

6 Level A analysis

6.1 Level A.a MapQ

Explanation:
MapQ [**?**] computes the local correlation between the map and each one of its atoms assumed to have a Gaussian shape.

Results:

Fig. 34 shows the histogram of the Q-score according calculated by MapQ. Some representative percentiles are:

Percentile	MapQ score [0-1]
2.5%	-0.54
25%	-0.14
50%	0.10
75%	0.33
97.5%	0.67
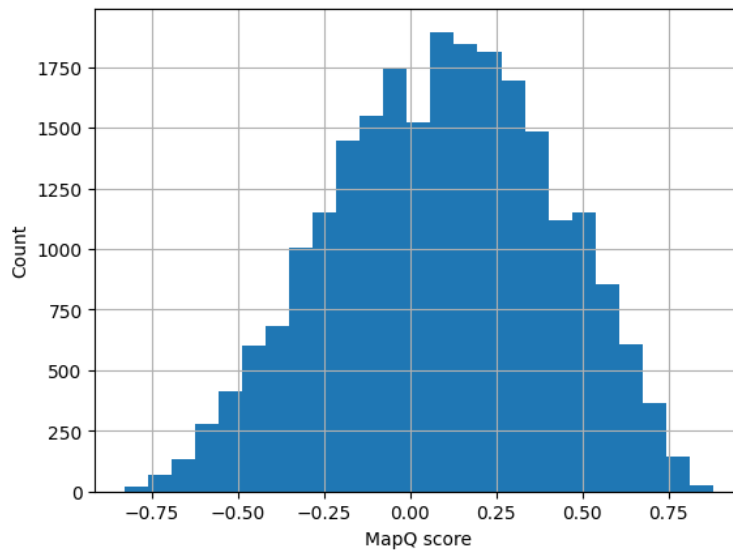
Figure 34: Histogram of the Q-score.

The following table shows the average Q score and estimated resolution for each chain.

Chain	Average Q score [0-1]	Estimated Resol. (Å)
A	0.15	5.4
A	0.16	0.0
A	0.12	0.0
A	0.09	0.0
B	0.04	6.1
B	0.01	0.0
B	0.19	0.0
C	0.07	5.9
C	0.17	0.0
D	0.03	0.0
E	0.29	0.0
E	0.37	0.0
E	0.32	0.0
F	0.06	0.0
F	0.24	0.0
F	-0.03	0.0
G	0.21	0.0
H	0.20	0.0
I	0.17	0.0
J	0.04	0.0
K	0.36	0.0
L	0.09	0.0
M	0.16	0.0
N	0.32	0.0
N	0.29	0.0
O	0.07	0.0
O	0.24	0.0
P	0.26	0.0
Q	0.37	0.0
Q	0.28	0.0
R	0.33	0.0
R	0.15	0.0
R	0.10	0.0
S	0.10	0.0
T	0.30	0.0
U	0.14	0.0
U	0.05	0.0
U	0.05	0.0
V	0.15	0.0
W	0.18	0.0
W	-0.05	0.0
X	0.04	0.0
Y	0.17	0.0
Y	0.14	0.0
Z	-0.09	0.0

Automatic criteria: The validation is OK if the median Q-score is larger than 0.1.

STATUS: OK

6.2   Level A.b FSC-Q

Explanation:
FSC-Q [Ramírez-Aportela et al., 2021] compares the local FSC between the map and the atomic model to the local FSC of the two half maps. FSC-Qr is the normalized version of FSC-Q to facilitate comparisons. Typically, FSC-Qr should take values between -1.5 and 1.5, being 0 an indicator of good matching between map and model.

Results:

Fig. 35 shows the histogram of FSC-Qr and Fig. 36 the colored isosurface of the atomic model converted to map. The average FSC-Qr is 0.82, its 95% confidence interval is [-1.00, 2.91]. The percentage of values whose FSC-Qr absolute value is beyond 1.5 is 10.2 %.
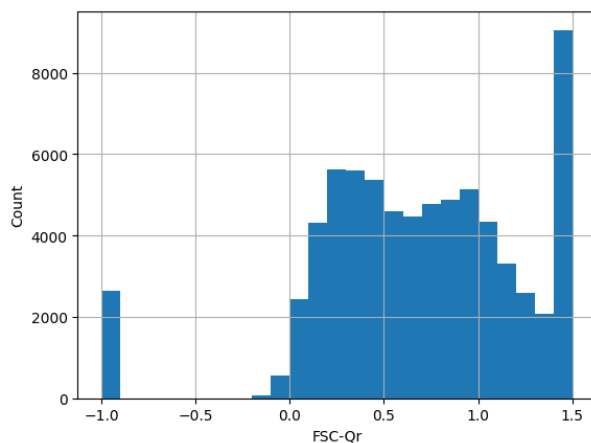

Figure 35: Histogram of the FSC-Qr limited to -1.5 and 1.5.
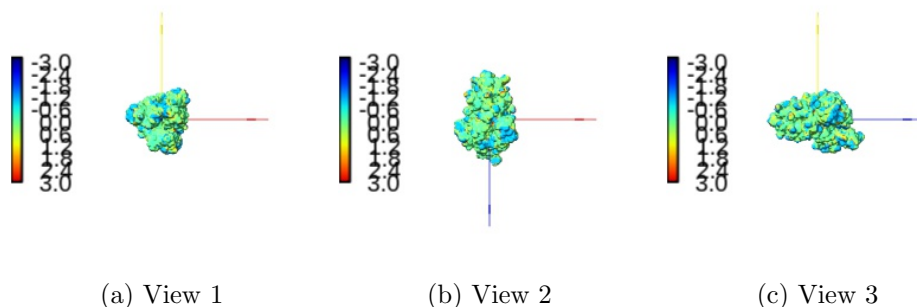
|(a) View 1|(b) View 2|(c) View 3|

Figure 36: Isosurface of the atomic model colored by FSC-Qr between -1.5 and 1.5 Views generated by ChimeraX at a the following X, Y, Z angles: View 1 (0,0,0), View 2 (90, 0, 0), View 3 (0, 90, 0).

Automatic criteria: The validation is OK if the percentage of residues whose FSC-Q is larger than 1.5 in absolute value is smaller than 10%.

WARNINGS: 1 warnings

1. **The percentage of voxels that have a FSC-Qr larger than 1.5 in absolute value is 10.2, that is larger than 10%**

6.3   Level A.d Map-Model Guinier analysis

Explanation:
We compared the Guinier plot [Rosenthal and Henderson, 2003] of the atomic model and the experimental map. We made the mean of both profiles to be equal (and equal to the mean of the atomic model) to make sure that they had comparable scales.

Results:
Fig. 37 shows the logarithm (in natural units) of the structure factor (the module squared of the Fourier transform) of the atom model and the experimental map. The correlation between the two profiles was 0.970.
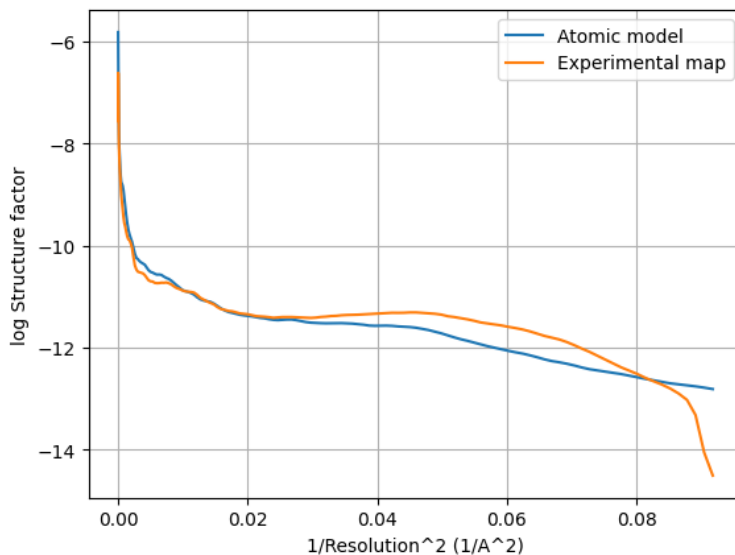
Figure 37: Guinier plot of the atom model and experimental map. The X-axis is the square of the inverse of the resolution in Å.

Automatic criteria: The validation is OK if the correlation between the two Guinier profiles is larger than 0.5.

STATUS: OK

6.4 Level A.e Phenix validation

Explanation:
Phenix provides a number of tools to assess the agreement between the experimental map and its atomic model [Afonine et al., 2018]. There are several cross-correlations to assess the quality of the fitting:

- CC (mask): Model map vs. experimental map correlation coefficient calculated considering map values inside a mask calculated around the macromolecule.
- CC (box): Model map vs. experimental map correlation coefficient calculated considering all grid points of the box.

41

- CC (volume) and CC (peaks) compare only map regions with the highest density values and regions below a certain contouring threshold level are ignored. CC (volume): The map region considered is defined by the N highest points inside the molecular mask. CC (peaks): In this case, calculations consider the union of regions defined by the N highest peaks in the model-calculated map and the N highest peaks in the experimental map.
- Local real-space correlation coefficients CC (main chain) and CC (side chain) involve the main skeleton chain and side chains, respectively.

There are also multiple ways of measuring the resolution:

- d99: Resolution cutoff beyond which Fourier map coefficients are negligibly small. Calculated from the full map.
- d_model: Resolution cutoff at which the model map is the most similar to the target (experimental) map. For d_model to be meaningful, the model is expected to fit the map as well as possible. d_model (B factors = 0) tries to avoid the blurring of the map.
- d_FSC_model; Resolution cutoff up to which the model and map Fourier coefficients are similar at FSC values of 0, 0.143, 0.5.

In addition to these resolution measurements the overall isotropic B factor is another indirect measure of the quality of the map.

Results:

To avoid ringing in Fourier space a smooth mask with a radius of 6.6 Å has been applied.

Overall correlation coefficients:

$$
\begin{aligned}
\text{CC (mask)} &= 0.592 \\
\text{CC (box)} &= 0.643 \\
\text{CC (volume)} &= 0.671 \\
\text{CC (peaks)} &= 0.559 \\
\text{CC (main chain)} &= 0.611 \\
\text{CC (side chain)} &= 0.598
\end{aligned}
$$

Correlation coefficients per chain:

Chain	Cross-correlation
A	0.585339
C	0.627294
B	0.633729
D	0.492978
E	0.629175
F	0.536286
G	0.653338
H	0.679420
I	0.621025
J	0.467733
K	0.622740
L	0.614352
M	0.652587
N	0.658026
O	0.506773
P	0.670467
Q	0.610106
R	0.653794
S	0.446171
T	0.390683
U	0.552064
V	0.576380
W	0.502179
X	0.599782
Y	0.607647
Z	0.604608
a	0.507833
b	0.374672
c	0.414481
d	0.294232

We now show the correlation profiles of the different chain per residue.
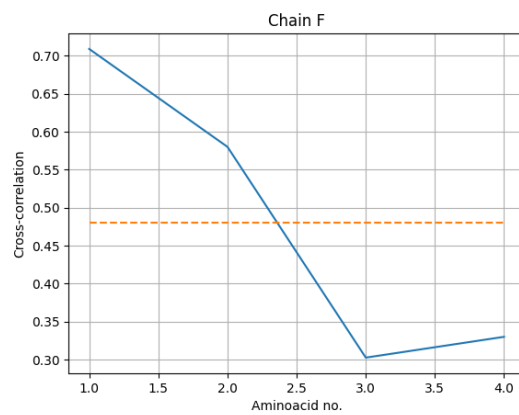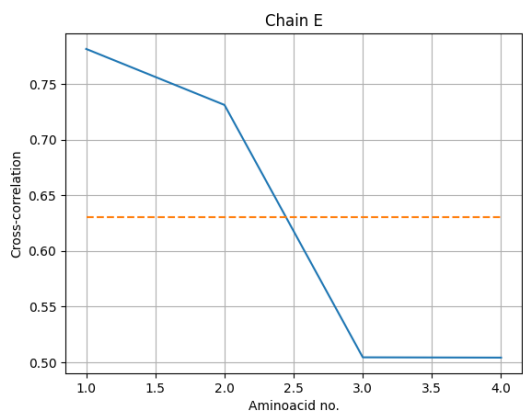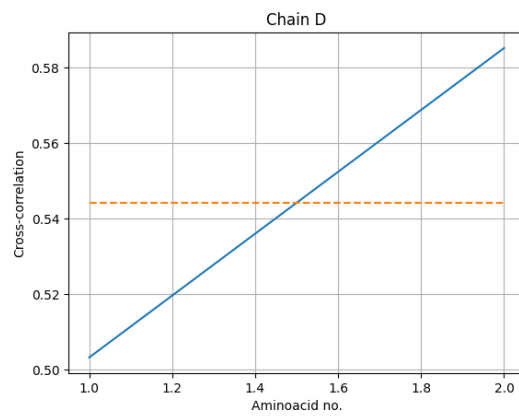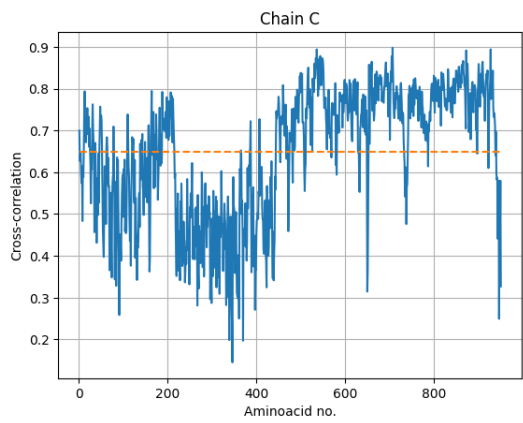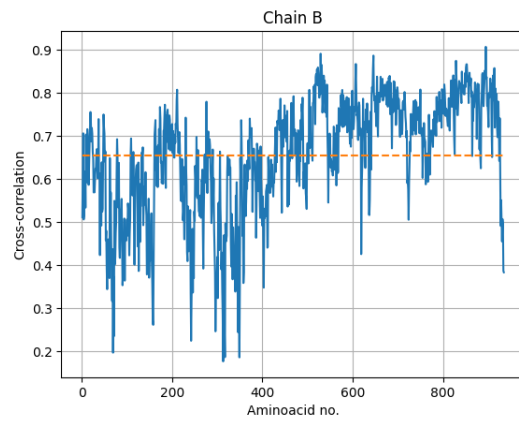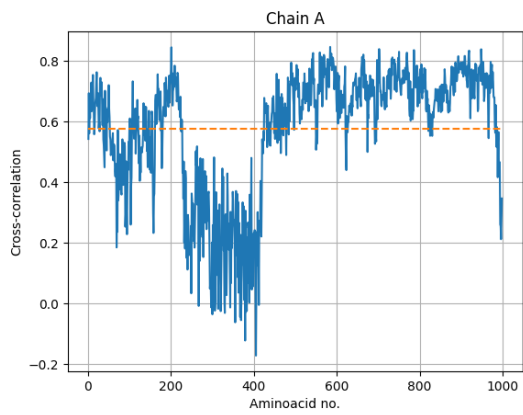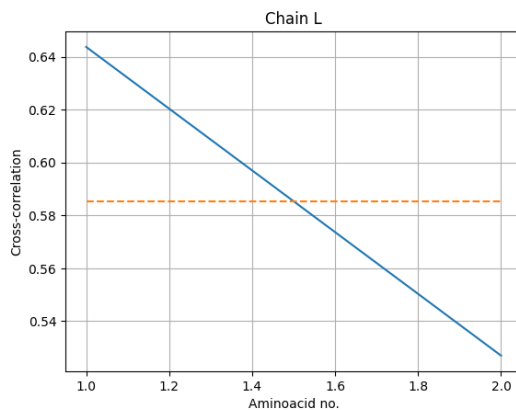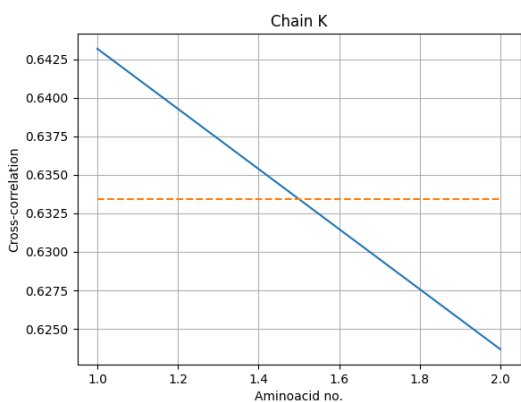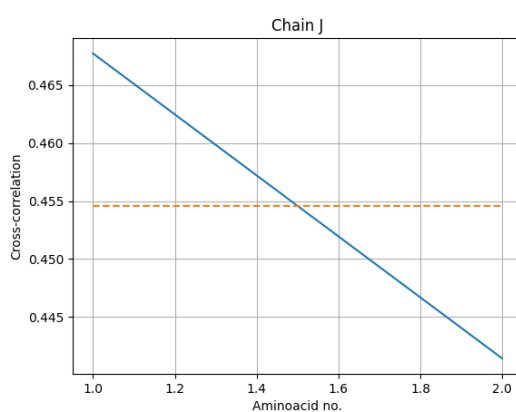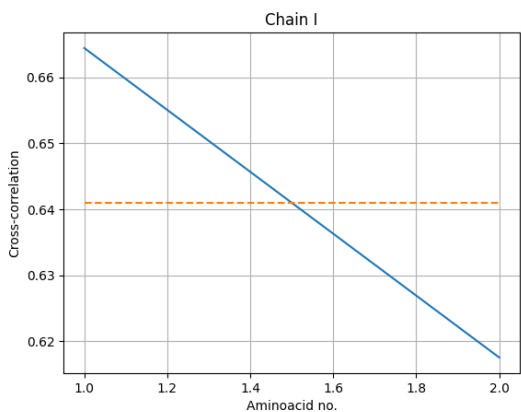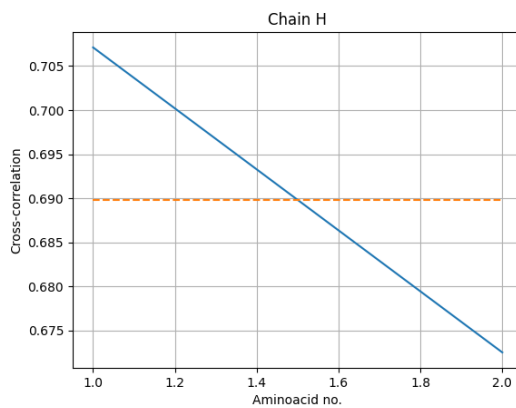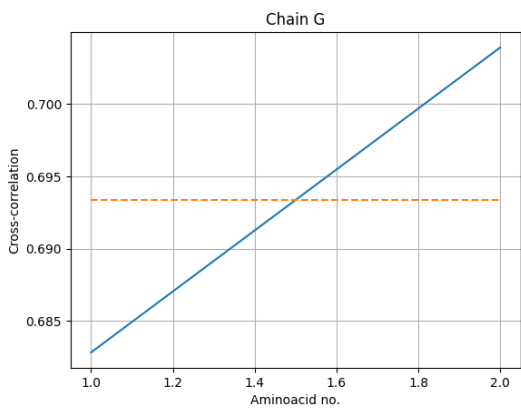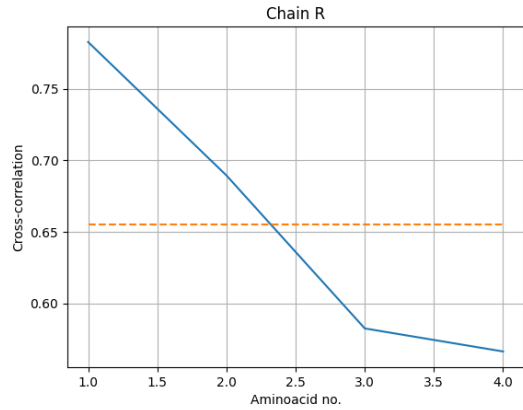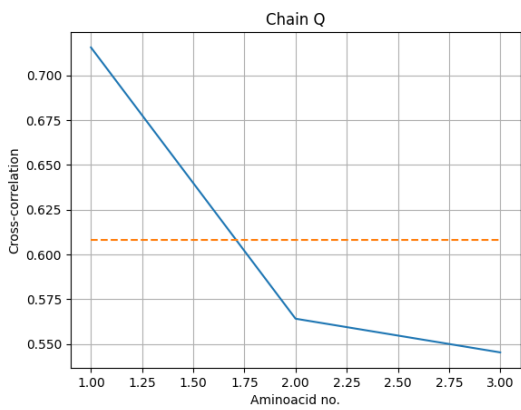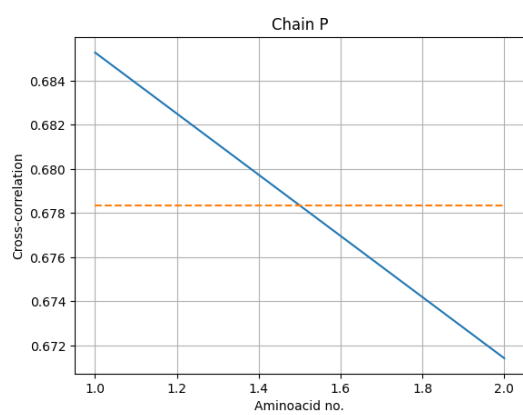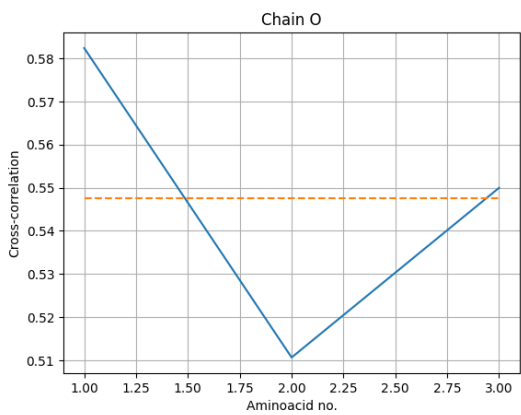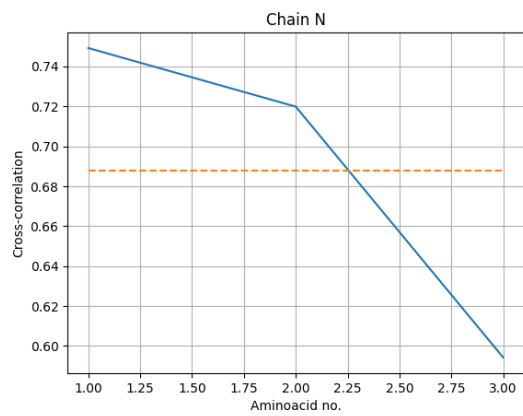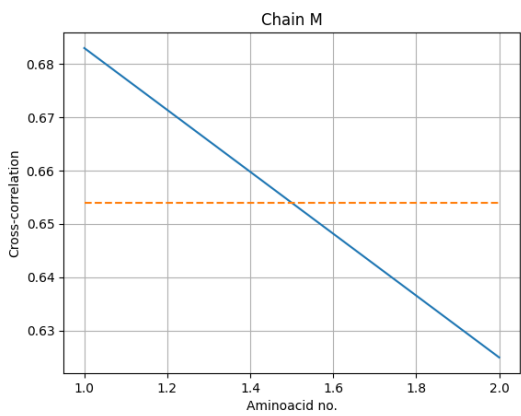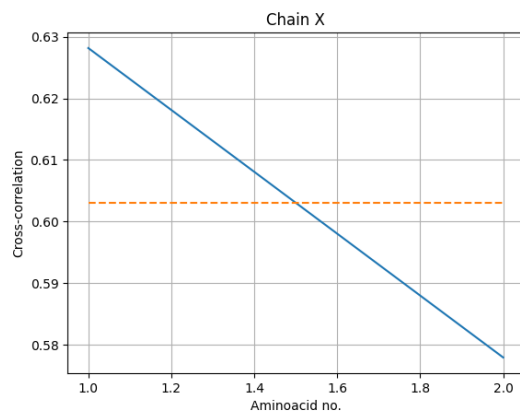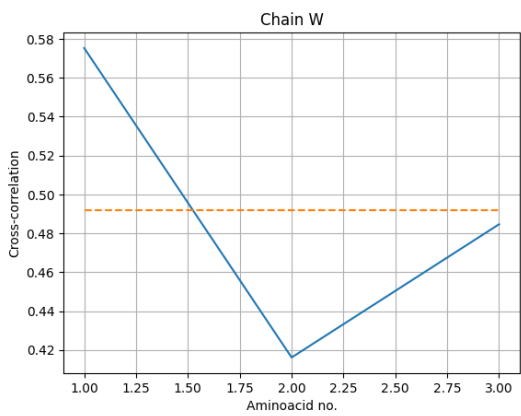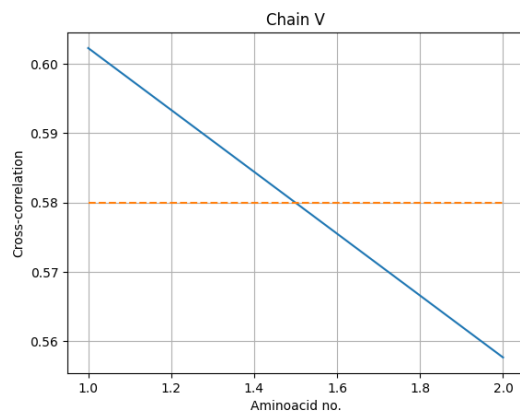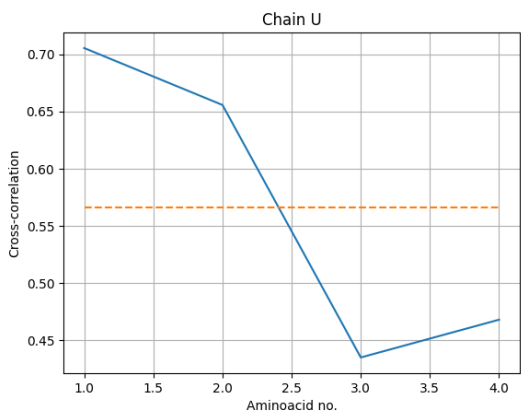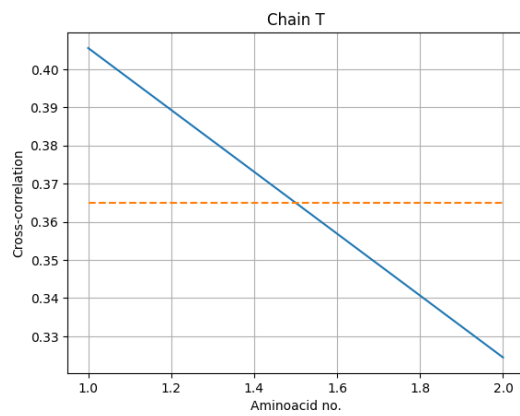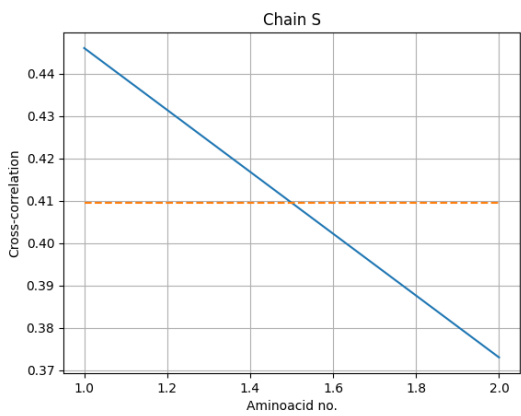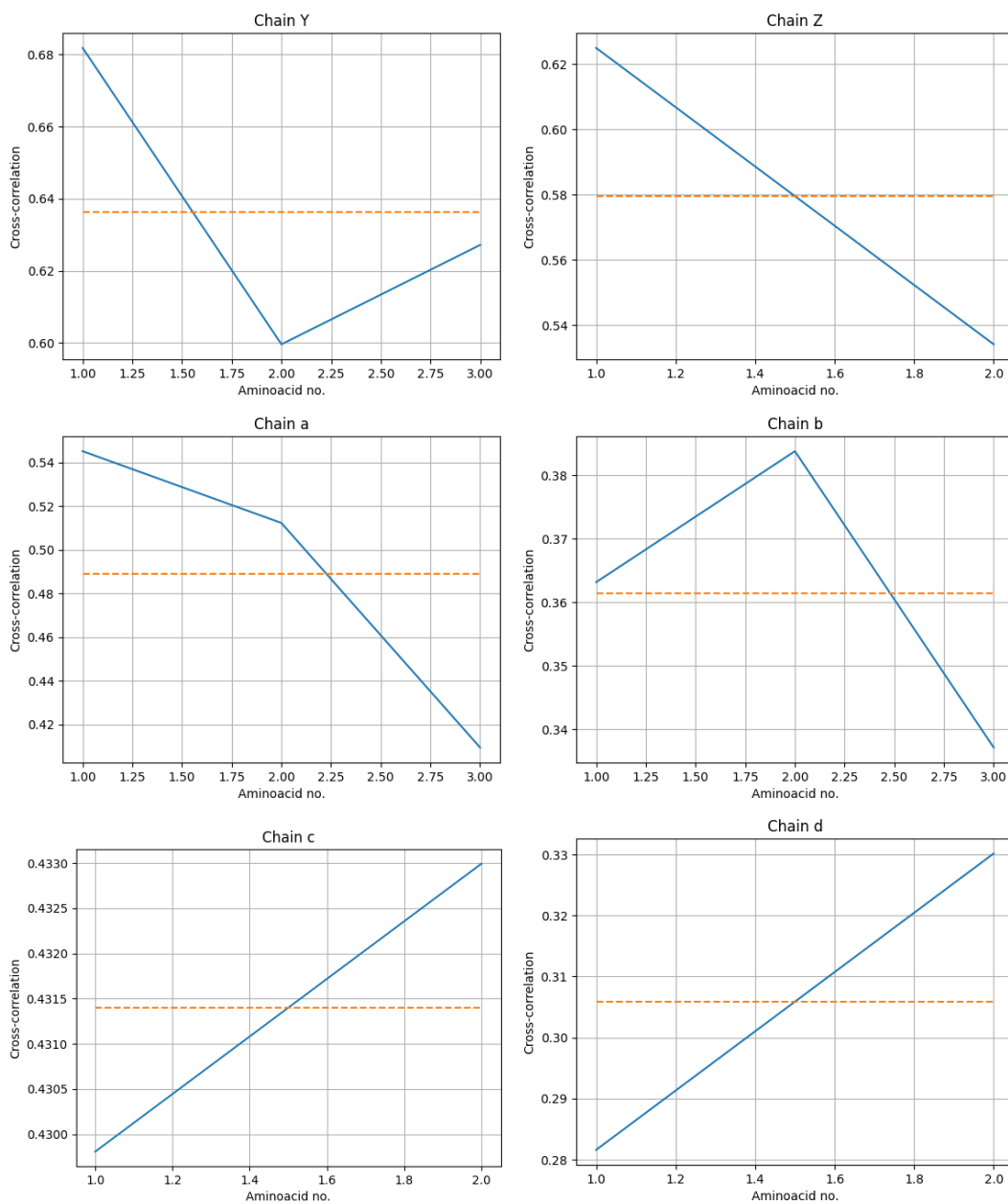
Fig. 38 shows the histogram of all cross-correlations evaluated at the residues. The percentage of residues whose correlation is below 0.5 is 20.8 %.

48
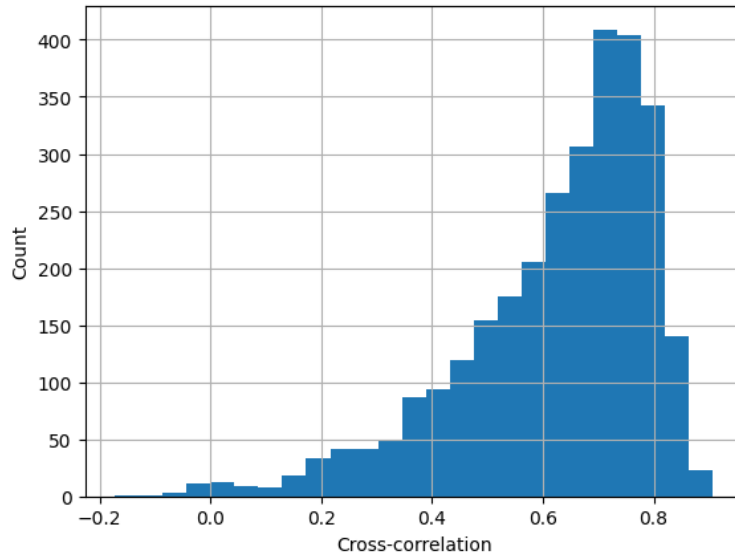
Figure 38: Histogram of the cross-correlation between the map and model evaluated for all residues.

Resolutions estimated from the model:

Resolution (Å)	Masked	Unmasked
d99	4.0	3.9
d_model	3.7	3.7
d_model (B-factor=0)	6.9	7.1
FSC_model=0	3.3	3.3
FSC_model=0.143	3.4	3.4
FSC_model=0.5	4.0	4.3

Overall isotropic B factor:

B factor	Masked	Unmasked
Overall B-iso	110.0	115.0

Fig. 39 shows the FSC between the input map and the model.

49
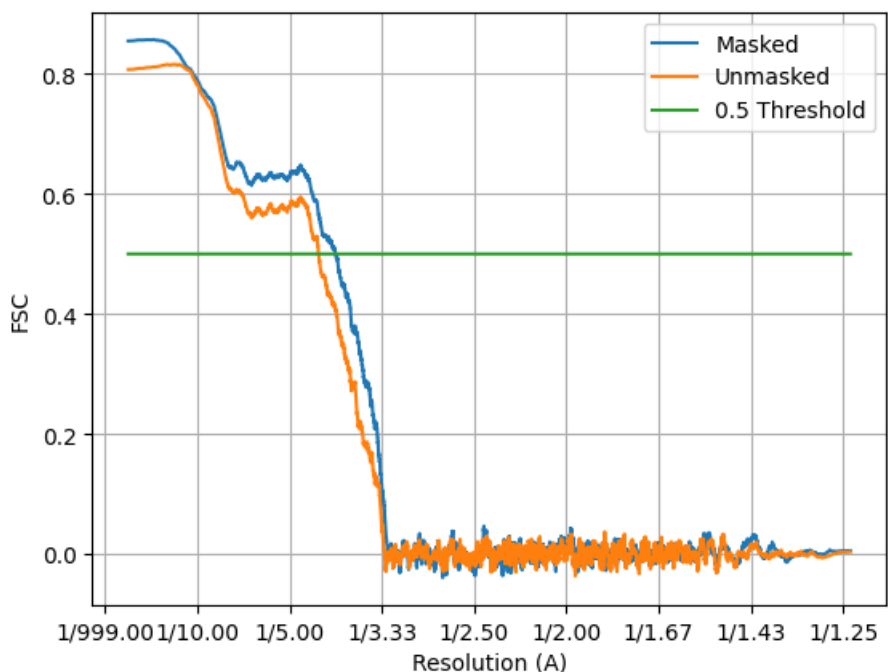
Figure 39: FSC between the input map and model with and without a mask constructed from the model. The X-axis is the square of the inverse of the resolution in Å.

Automatic criteria: The validation is OK if 1) the percentage of residues whose correlation is smaller than 0.5 is smaller than 10%, and 2) the resolution reported by the user is larger than 0.8 times the resolution estimated between the map and model at FSC=0.5.

WARNINGS: 1 warnings

1. **The percentage of residues that have a cross-correlation below 0.5 is 20.8, that is larger than 10%**

6.5   Level A.f EMRinger validation

Explanation:
EMringer [Barad et al., 2015] compares the side chains of the atomic model to the CryoEM map. The following features are reported:

- Optimal Threshold: Electron potential map cutoff value at which the maximum EMRinger score was obtained.
- Rotamer Ratio: Fraction of rotameric residues at the Optimal threshold value.
- Max Zscore: Z-score computed to determine the significance of the distribution at the Optimal threshold value.
- Model Length: Total of non-gamma-branched, non-proline aminoacids with a non-H gamma atom used in global EMRinger score computation.
- EMRinger Score: Maximum EMRinger score calculated at the Optimal Threshold.

A rotameric residue is one in which EMRinger peaks that fall within defined rotamers based on chi1, this often suggests a problem with the modelling of the backbone. In general, the user should look at the profiles and identify regions that may need improvement.

Results:

General results:

Optimal threshold	0.601310
Rotamer ratio	0.713
Max. Zscore	6.12
Model length	1723
EMRinger Score	1.474

Fig. 40 shows the EMRinger score and fraction of rotameric residues as a function of the map threshold. The optimal threshold was selected looking for the maximum EMRinger score in this plot.
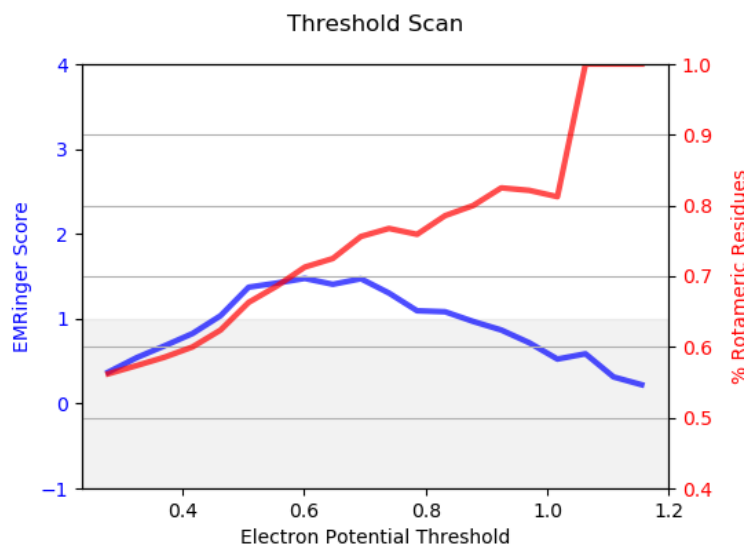
Figure 40: EMRinger score and fraction of rotameric residues as a function of the map threshold.

Fig. 41 shows the histogram for rotameric (blue) and non-rotameric (red) residues at the optimal threshold.
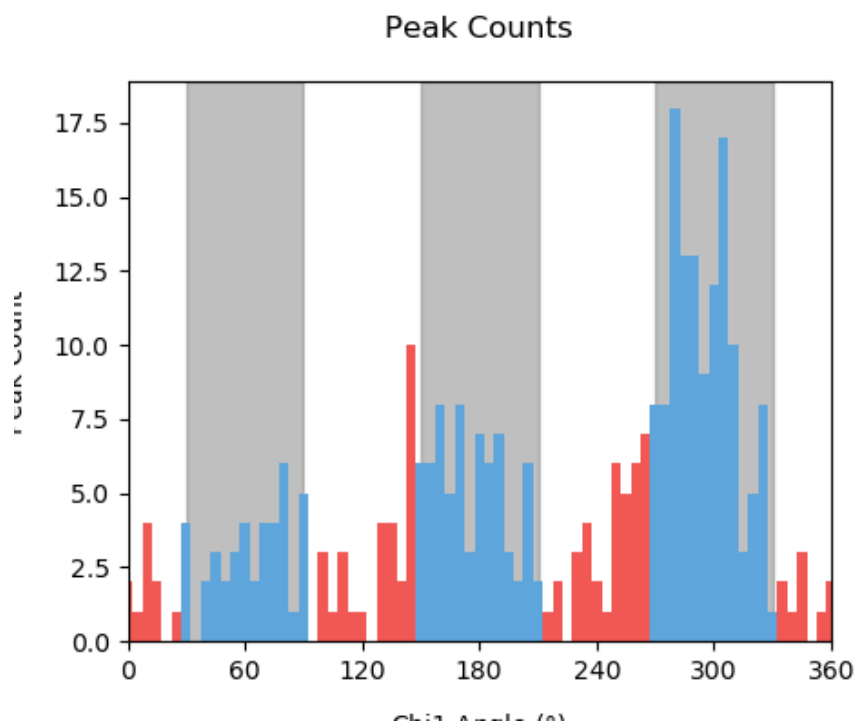
Figure 41: Histogram for rotameric (blue) and non-rotameric (red) residues at the optimal threshold as a function of the angle Chi1.

The following plots show the rolling window EMRinger analysis of the different chains to distinguish regions of improved model quality. This analysis was performed on rolling sliding 21-residue windows along the primary sequence of the protein chains.

Rolling window - Chain C, Threshold 0.000000
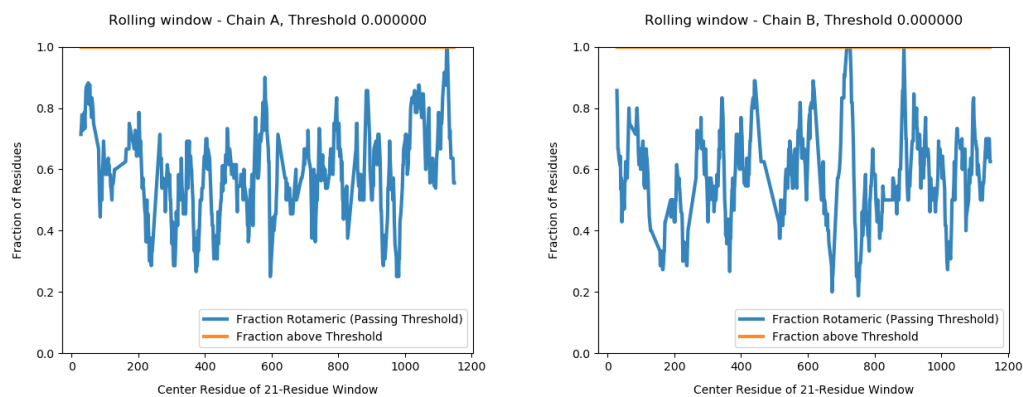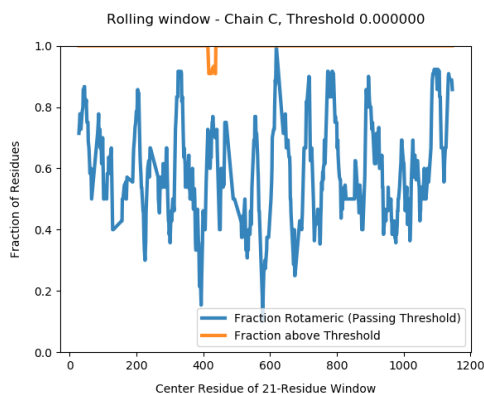
Automatic criteria: The validation is OK if the EMRinger score and Max. Zscore are larger than 1.

STATUS: OK

6.6 Level A.g DAQ validation

Explanation:

DAQ [Terashi et al., 2022] is a computational tool using deep learning that can estimate the residue-wise local quality for protein models from cryo-Electron Microscopy maps. The method calculates the likelihood that a given density feature corresponds to an aminoacid, atom, and secondary structure. These likelihoods are combined into a score that ranges from -1 (bad quality) to 1 (good quality).

Results:

Fig. 42 shows the histogram of the DAQ values. The mean and standard deviation were 0.5 and 0.4, respectively.
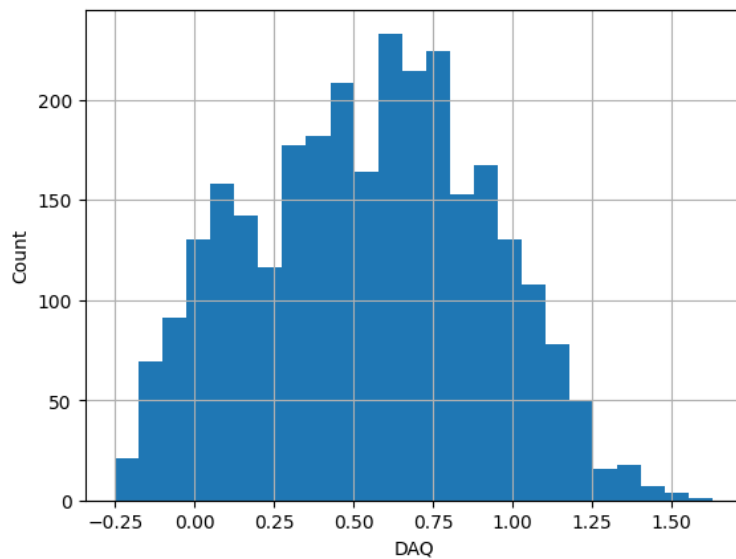
Figure 42: Histogram of the DAQ values.

The atomic model colored by DAQ can be seen in Fig. 43.

(a) View 1   (b) View 2   (c) View 3
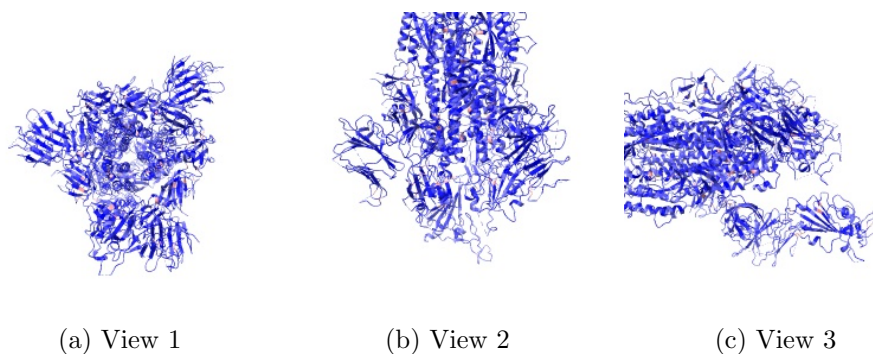
Figure 43: Atomic model colored by DAQ Views generated by ChimeraX at a the following X, Y, Z angles: View 1 (0,0,0), View 2 (90, 0, 0), View 3 (0, 90, 0).

Automatic criteria: The validation is OK if the average DAQ score is larger than 0.5.

STATUS: OK

References

[Rosenthal and Henderson, 2003] Rosenthal, P. B. and Henderson, R. (2003). Optimal determination of particle orientation, absolute hand, and contrast loss in single particle electron-cryomicroscopy. *J. Molecular Biology*, 333:721–745.

[Ramírez-Aportela et al., 2019] Ramírez-Aportela, E., Mota, J., Conesa, P., Carazo, J. M., and Sorzano, C. O. S. (2019). DeepRes: a new deep-learning- and aspect-based local resolution method for electron-microscopy maps. *IUCRj*, 6:1054–1063.

[Kaur et al., 2021] Kaur, S., Gomez-Blanco, J., Khalifa, A. A., Adinarayanan, S., Sanchez-Garcia, R., Wrapp, D., McLellan, J. S., Bui, K. H., and Vargas, J. (2021). Local computational methods to improve the interpretability and analysis of cryo-EM maps. *Nature Communications*, 12(1):1–12.

[Sorzano et al., 2017] Sorzano, C. O. S., Vargas, J., Oton, J., Abrishami, V., de la Rosa-Trevin, J. M., Gomez-Blanco, J., Vilas, J. L., Marabini, R., and Carazo, J. M. (2017). A review of resolution measures and related aspects in 3D electron microscopy. *Progress in biophysics and molecular biology*, 124:1–30.

[Beckers and Sachse, 2020] Beckers, M. and Sachse, C. (2020). Permutation testing of fourier shell correlation for resolution estimation of cryo-em maps. *J. Structural Biology*, 212(1):107579.

[Cardone et al., 2013] Cardone, G., Heymann, J. B., and Steven, A. C. (2013). One number does not fit all: Mapping local variations in resolution in cryo-em reconstructions. *J. Structural Biology*, 184:226–236.

[Kucukelbir et al., 2014] Kucukelbir, A., Sigworth, F. J., and Tagare, H. D. (2014). Quantifying the local resolution of cryo-EM density maps. *Nature Methods*, 11:63–65.

[Vilas et al., 2018] 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., and Sorzano, C. O. S. (2018). MonoRes: automatic and unbiased estimation of local resolution for electron microscopy maps. *Structure*, 26:337–344.

[Vilas et al., 2020] Vilas, J. L., Tagare, H. D., Vargas, J., Carazo, J. M., and Sorzano, C. O. S. (2020). Measuring local-directional resolution and local anisotropy in cryo-EM maps. *Nature communications*, 11:55.

[Pintilie et al., 2020] Pintilie, G., Zhang, K., Su, Z., Li, S., Schmid, M. F., and Chiu, W. (2020). Measurement of atom resolvability in cryo-em maps with q-scores. *Nature methods*, 17(3):328–334.

[Ramírez-Aportela et al., 2021] Ramírez-Aportela, E., Maluenda, D., Fonseca, Y. C., Conesa, P., Marabini, R., Heymann, J. B., Carazo, J. M., and Sorzano, C. O. 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.

[Afonine et al., 2018] Afonine, P. V., Klaholz, B. P., Moriarty, N. W., Poon, B. K., Sobolev, O. V., Terwilliger, T. C., Adams, P. D., and Urzhumtsev, A. (2018). New tools for the analysis and validation of cryo-EM maps and atomic models. *Acta Crystallographica D, Struct. Biol.*, 74:814–840.

[Barad et al., 2015] Barad, B. A., Echols, N., Wang, R. Y.-R., Cheng, Y., DiMaio, F., Adams, P. D., and Fraser, J. S. (2015). EMRinger: side chain-directed model and map validation for 3D cryo-electron microscopy. *Nature Methods*, 12(10):943–946.

[Terashi et al., 2022] Terashi, G., Wang, X., Subramaniya, S.R.M.V., Tesmer, J.J.G. and Kihara, D. (2022). Residue-Wise Local Quality Estimation for Protein Models from Cryo-EM Maps. (submitted).

Validation report of Level(s) 0, A

I^{2}PC Validation server

March 2, 2022
7:19am

Abstract

The map seems to be well centered. There seems to be a problem with the suggested threshold (see Sec. 2.2). There seems to be a problem with the map's background (see Sec. 2.3). There seems to be a problem with the map hand (see Sec. 2.8). There seems to be a problem with its MapQ scores (see Sec. 4.1). According to phenix, it seems that there might be some mismatch between the map and its model (see Sec. 4.3). DAQ detects some mismatch between the map and its model (see Sec. 4.5).

The average resolution of the map estimated by various methods goes from 3.2Å to 6.3Å with an average of 4.3Å. The resolution provided by the user was 3.7Å.

The overall score (passing tests) of this report is 7 out of 13 evaluable items.

Summary of the warnings across sections.

If it is empty below this point, it means that there are no warnings.

Section 2.2 (0.b Mask analysis)

1. **There might be a problem of connectivity at this threshold because more than 5 connected components are needed to reach 95% of the total mask.**

Section 2.3 (0.c Background analysis)

1. **The null hypothesis that the background mean is 0 has been rejected because the p-value of the comparison is smaller than 0.001**

2. **There is a significant proportion of outlier values in the background (cdf5 ratio=12398.45)**

Section 2.8 (0.h DeepHand)

1. **The orientation of the volume is uncertain.**

Section 4.1 (A.a MapQ)

1. **The median Q-score is less than 0.1.**

Section 4.3 (A.e Phenix validation)

1. **The percentage of residues that have a cross-correlation below 0.5 is 20.7, that is larger than 10%**

Section 4.5 (A.g DAQ)

1. **The average DAQ is smaller than 0.5.**

Contents

1 Input data

Input map: /home/coss/data/Dropbox/Aplicaciones/ShareLaTeX/MapValidation/-EMDB22301/emd_22301.map
SHA256 hash: 241e0d7426641a7594406084743c1ad370b5ee646474d4e9b0262133c8b7490f
Voxel size: 0.520000 (Å)
Visualization threshold: 0.100000
Resolution estimated by user: 3.700000

Orthogonal slices of the input map
Explanation:
In the orthogonal slices of the map, the noise outside the protein should not
have any structure (stripes going out, small blobs, particularly high or low
densities, ...)

Results:
See Fig. 1.

| (a) X Slice 256 | (b) Y Slice 256 | (c) Z Slice 256 |

Figure 1: Central slices of the input map in the three dimensions

Orthogonal slices of maximum variance of the input map
Results:
See Fig. 2.

(a) X Slice 176   (b) Y Slice 237   (c) Z Slice 267

Figure 2: Slices of maximum variation in the three dimensions

Orthogonal projections of the input map

Explanation:

In the projections there should not be stripes (this is an indication of directional overweighting, or angular attraction), and there should not be a dark halo around or inside the structure (this is an indication of incorrect CTF correction or the reconstruction of a biased map).

Results:

See Fig. 3.

(a) X Projection   (b) Y Projection   (c) Z Projection

Figure 3: Projections in the three dimensions

Isosurface views of the input map

Explanation:

An isosurface is the surface of all points that have the same gray value. In these views there should not be many artifacts or noise blobs around the map.

Results:
See Fig. 4.

(a) View 1   (b) View 2   (c) View 3

Figure 4: Isosurface at threshold=0.100000. Views generated by ChimeraX at a the following X, Y, Z angles: View 1 (0,0,0), View 2 (90, 0, 0), View 3 (0, 90, 0).

Orthogonal slices of maximum variance of the mask
Explanation:
The mask has been calculated at the suggested threshold 0.100000, the largest connected component was selected, and then dilated by 2Å.

Results:
See Fig. 5.

(a) X Slice 256   (b) Y Slice 256   (c) Z Slice 256
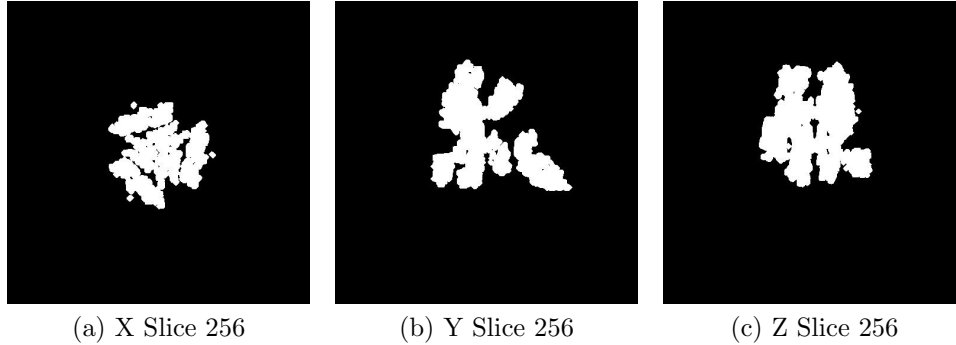
Figure 5: Slices of maximum variation in the three dimensions of the mask

2 Level 0 analysis

2.1 Level 0.a Mass analysis

Explanation:
The reconstructed map must be relatively well centered in the box, and there should be at least 30Å (the exact size depends on the CTF) on each side to make sure that the CTF can be appropriately corrected.

Results:
The space from the left and right in X are 51.48 and 98.28 Å, respectively. There is a decentering ratio (abs(Right-Left)/Size)% of 17.58%

The space from the left and right in Y are 69.16 and 66.56 Å, respectively. There is a decentering ratio (abs(Right-Left)/Size)% of 0.98%

The space from the left and right in Z are 83.20 and 58.24 Å, respectively. There is a decentering ratio (abs(Right-Left)/Size)% of 9.38%

The center of mass is at (x,y,z)=(254.51,256.07,256.16). The decentering of the center of mass (abs(Center)/Size)% is 0.29, 0.01, and 0.03, respectively.%

Automatic criteria: The validation is OK if 1) the decentering and

center of mass less than 20% of the map dimensions in all directions, and 2) the extra space on each direction is more than 20% of the map dimensions.

STATUS: OK

2.2   Level 0.b Mask analysis

Explanation:
The map at the suggested threshold should have most of its mass concentrated in a single connected component. It is normal that after thresholding there are a few thousands of very small, disconnected noise blobs. However, there total mass should not exceed 10%. The raw mask (just thresholding) and the mask constructed for the analysis (thresholding + largest connected component + dilation) should significantly overlap. Overlap is defined by the overlapping coefficient (size(Raw AND Constructed)/size(Raw)) that is a number between 0 and 1, the closer to 1, the more they agree.

Results:

Raw mask: At threshold 0.100000, there are 4117 connected components with a total number of voxels of 606329 and a volume of 85254.71 Å^3 (see Fig. 6). The size and percentage of the total number of voxels for the raw mask are listed below (up to 95% of the mass or the first 100 clusters, whatever happens first), the list contains (No. voxels (volume in Å^3), percentage, cumulatedPercentage):

, (513332 (72178.59), 84.66, 84.66), (1535 (215.83), 0.25, 84.92), (1530 (215.13), 0.25, 85.17), (973 (136.81), 0.16, 85.33), (961 (135.12), 0.16, 85.49), (955 (134.28), 0.16, 85.64), (955 (134.28), 0.16, 85.80), (934 (131.33), 0.15, 85.96), (911 (128.09), 0.15, 86.11), (881 (123.88), 0.15, 86.25), (709 (99.69), 0.12, 86.37), (639 (89.85), 0.11, 86.47), (583 (81.97), 0.10, 86.57), (470 (66.09), 0.08, 86.65), (458 (64.40), 0.08, 86.72), (451 (63.41), 0.07, 86.80), (447 (62.85), 0.07, 86.87), (444 (62.43), 0.07, 86.94), (441 (62.01), 0.07, 87.02), (428 (60.18), 0.07, 87.09), (421 (59.20), 0.07, 87.16), (421 (59.20), 0.07, 87.23), (418 (58.77), 0.07, 87.30), (389 (54.70), 0.06, 87.36), (365 (51.32), 0.06, 87.42), (364 (51.18), 0.06, 87.48), (360 (50.62), 0.06, 87.54), (359 (50.48), 0.06, 87.60), (342 (48.09), 0.06, 87.65), (342 (48.09), 0.06, 87.71),

9

(319 (44.85), 0.05, 87.76), (316 (44.43), 0.05, 87.82), (316 (44.43), 0.05, 87.87), (314 (44.15), 0.05, 87.92), (312 (43.87), 0.05, 87.97), (308 (43.31), 0.05, 88.02), (308 (43.31), 0.05, 88.07), (305 (42.89), 0.05, 88.12), (303 (42.60), 0.05, 88.17), (298 (41.90), 0.05, 88.22), (292 (41.06), 0.05, 88.27), (288 (40.50), 0.05, 88.32), (279 (39.23), 0.05, 88.36), (277 (38.95), 0.05, 88.41), (277 (38.95), 0.05, 88.46), (270 (37.96), 0.04, 88.50), (264 (37.12), 0.04, 88.54), (264 (37.12), 0.04, 88.59), (263 (36.98), 0.04, 88.63), (263 (36.98), 0.04, 88.67), (261 (36.70), 0.04, 88.72), (260 (36.56), 0.04, 88.76), (257 (36.14), 0.04, 88.80), (251 (35.29), 0.04, 88.84), (250 (35.15), 0.04, 88.88), (248 (34.87), 0.04, 88.93), (232 (32.62), 0.04, 88.96), (228 (32.06), 0.04, 89.00), (224 (31.50), 0.04, 89.04), (224 (31.50), 0.04, 89.08), (219 (30.79), 0.04, 89.11), (215 (30.23), 0.04, 89.15), (215 (30.23), 0.04, 89.18), (213 (29.95), 0.04, 89.22), (212 (29.81), 0.03, 89.25), (212 (29.81), 0.03, 89.29), (212 (29.81), 0.03, 89.32), (211 (29.67), 0.03, 89.36), (209 (29.39), 0.03, 89.39), (207 (29.11), 0.03, 89.43), (205 (28.82), 0.03, 89.46), (197 (27.70), 0.03, 89.49), (190 (26.72), 0.03, 89.52), (190 (26.72), 0.03, 89.55), (189 (26.57), 0.03, 89.59), (183 (25.73), 0.03, 89.62), (182 (25.59), 0.03, 89.65), (181 (25.45), 0.03, 89.68), (179 (25.17), 0.03, 89.71), (178 (25.03), 0.03, 89.73), (177 (24.89), 0.03, 89.76), (176 (24.75), 0.03, 89.79), (174 (24.47), 0.03, 89.82), (174 (24.47), 0.03, 89.85), (174 (24.47), 0.03, 89.88), (169 (23.76), 0.03, 89.91), (168 (23.62), 0.03, 89.93), (167 (23.48), 0.03, 89.96), (164 (23.06), 0.03, 89.99), (163 (22.92), 0.03, 90.02), (161 (22.64), 0.03, 90.04), (160 (22.50), 0.03, 90.07), (159 (22.36), 0.03, 90.10), (156 (21.93), 0.03, 90.12), (155 (21.79), 0.03, 90.15), (153 (21.51), 0.03, 90.17), (150 (21.09), 0.02, 90.20), (150 (21.09), 0.02, 90.22), (150 (21.09), 0.02, 90.25)

Number of components to reach 95% of the mass: 509

The average size of the remaining 3608 components is 8.40 voxels (0.14 Å^3). Their size go from 38 voxels (5.34 Å^3) to 1 voxels (0.14 Å^3).

The slices of the raw mask can be seen in Fig. 6.
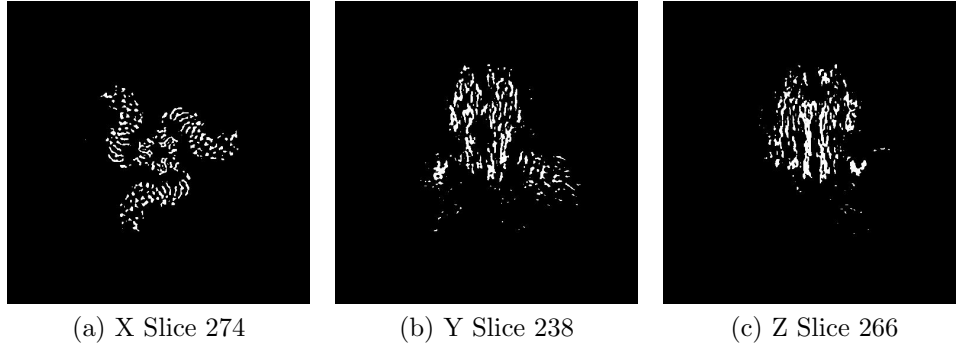
| (a) X Slice 274 | (b) Y Slice 238 | (c) Z Slice 266 |

Figure 6: Maximum variance slices in the three dimensions of the raw mask

The following table shows the variation of the mass enclosed at different thresholds (see Fig. 7):

Threshold	Voxel mass	Molecular mass(kDa)	# Aminoacids
0.0382	2348446.00	273.58	2487.09
0.0764	944917.00	110.08	1000.70
0.1146	477970.00	55.68	506.19
0.1528	280700.00	32.70	297.27
0.1910	176094.00	20.51	186.49
0.2292	112998.00	13.16	119.67
0.2674	72979.00	8.50	77.29
0.3056	46742.00	5.45	49.50
0.3438	30185.00	3.52	31.97
0.3820	19841.00	2.31	21.01
0.4202	12972.00	1.51	13.74
0.4584	8563.00	1.00	9.07
0.4966	5528.00	0.64	5.85
0.5348	3604.00	0.42	3.82
0.5730	2250.00	0.26	2.38
0.6112	1405.00	0.16	1.49
0.6494	845.00	0.10	0.89
0.6876	568.00	0.07	0.60
0.7258	356.00	0.04	0.38
0.7640	204.00	0.02	0.22
0.8022	108.00	0.01	0.11
0.8404	52.00	0.01	0.06
0.8786	17.00	0.00	0.02
0.9168	5.00	0.00	0.01
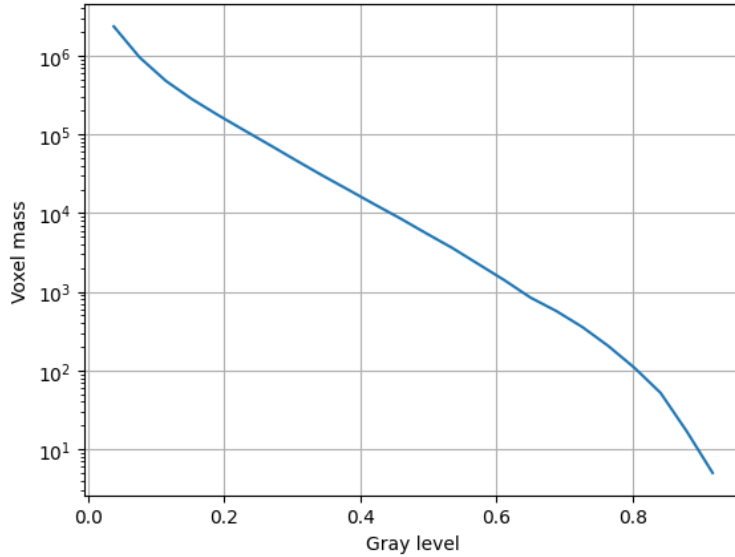
Figure 7: Voxel mass as a function of the gray level.

Constructed mask: After keeping the largest component of the previous mask and dilating it by 2Å, there is a total number of voxels of 2984355 and a volume of 419624.19 Å^3. The overlap between the raw and constructed mask is 0.88.

Automatic criteria: The validation is OK if 1) to keep 95% of the mass we need to keep at most 5 connected components; and 2) the average volume of the blobs outside the given threshold has a size smaller than 5Å^3; and 3) the overlap between the raw mask and the mask constructed for the analysis is larger than 75%.

WARNINGS: 1 warnings

1. **There might be a problem of connectivity at this threshold because more than 5 connected components are needed to reach 95% of the total mask.**

2.3 Level 0.c Background analysis

Explanation:
Background is defined as the region outside the macromolecule mask. The background mean should be zero, and the number of voxels with a very low or very high value (below 5 standard deviations of the noise) should be very small and they should be randomly distributed without any specific structure. Sometimes, you can see some structure due to the symmetry of the structure.

Results:

The null hypothesis that the background mean is 0 was tested with a one-sample Student's t-test. The resulting t-statistic and p-value were 949.55 and 0.000000, respectively.

The mean and standard deviation (sigma) of the background were 0.000721 and 0.008701. The percentage of background voxels whose absolute value is larger than 5 times the standard deviation is 0.71 % (see Fig. 8). The same percentage from a Gaussian would be 0.000057% (ratio between the two percentages: 12398.449346).

Slices of the background beyond 5*sigma can be seen in Fig. 8.
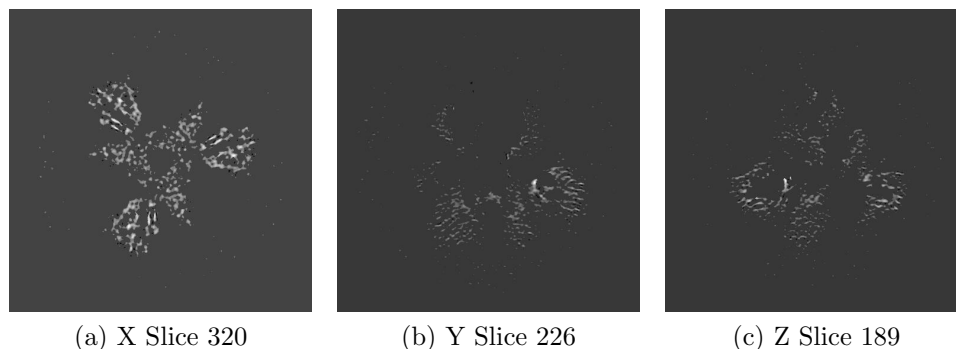
| (a) X Slice 320 | (b) Y Slice 226 | (c) Z Slice 189 |

Figure 8: Maximum variance slices in the three dimensions of the parts of the background beyond 5*sigma

Automatic criteria: The validation is OK if 1) the p-value of the null

hypothesis that the background has 0 mean is larger than 0.001; and 2) the number of voxels above or below 5 sigma is smaller than 20 times the amount expected for a Gaussian with the same standard deviation whose mean is 0.

WARNINGS: 2 warnings

1. **The null hypothesis that the background mean is 0 has been rejected because the p-value of the comparison is smaller than 0.001**
2. **There is a significant proportion of outlier values in the background (cdf5 ratio=12398.45)**

2.4   Level 0.d B-factor analysis

Explanation:
The B-factor line [Rosenthal and Henderson, 2003] fitted between 15Å and the resolution reported should have a slope that is between 0 and 300 Å^2.

Results:
Fig. 9 shows the logarithm (in natural units) of the structure factor (the module squared of the Fourier transform) of the experimental map, its fitted line, and the corrected map. The estimated B-factor was -14.3. The fitted line was $\log(|F|^2) = -3.6/R^2 + (-11.7)$.

Figure 9: Guinier plot. The X-axis is the square of the inverse of the resolution in Å.

(a) X Slice 179   (b) Y Slice 237   (c) Z Slice 267

Figure 10: Slices of maximum variation in the three dimensions of the B-factor corrected map

Automatic criteria: The validation is OK if the B-factor is in the range [-300,0].

STATUS: OK

15

2.5 Level 0.e Local resolution with DeepRes

Explanation:
DeepRes [Ramírez-Aportela et al., 2019] measures the local resolution using a neural network that has been trained on the appearance of atomic structures at different resolutions. Then, by comparing the local appearance of the input map to the appearance of the atomic structures a local resolution label can be assigned.

Results:

Fig. 11 shows the histogram of the local resolution according to DeepRes. Some representative percentiles are:

Percentile	Resolution(Å)
2.5%	2.50
25%	2.78
50%	3.21
75%	3.60
97.5%	4.26

The reported resolution, 3.70 Å, is at the percentile 80.3. Fig. 12 shows some representative views of the local resolution.
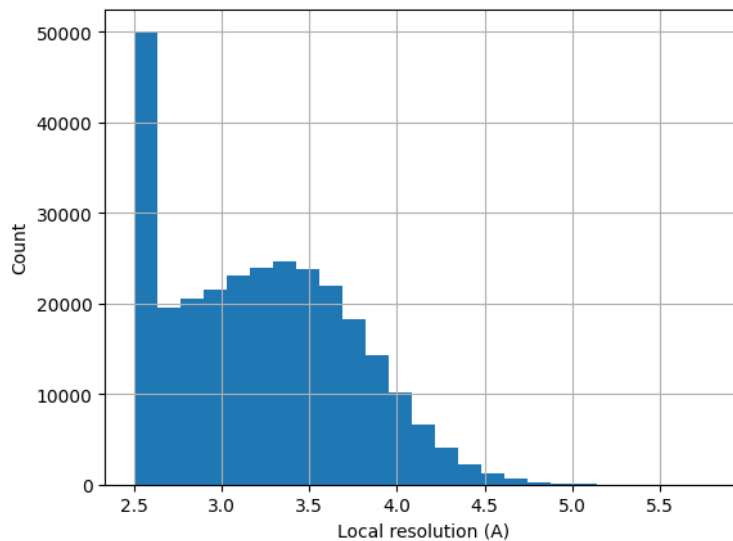
Figure 11: Histogram of the local resolution according to deepres.

(a) View 1   (b) View 2   (c) View 3

Figure 12: Local resolution according to DeepRes. Views generated by ChimeraX at a the following X, Y, Z angles: View 1 (0,0,0), View 2 (90, 0, 0), View 3 (0, 90, 0).

Automatic criteria: The validation is OK if the percentile of the user provided resolution is larger than 0.1% of the percentile of the local resolution as estimated by DeepRes.

STATUS: OK

2.6 Level 0.f Local B-factor

Explanation:
LocBfactor [Kaur et al., 2021] estimates a local resolution B-factor by decomposing the input map into a local magnitude and phase term using the spiral transform.

Results:

Fig. 13 shows the histogram of the local B-factor according to LocBfactor. Some representative percentiles are:

Percentile	Local B-factor ($\mathring{A}^{-2}$)
2.5%	-196.41
25%	-136.44
50%	-106.85
75%	-76.67
97.5%	-22.25

Fig. 14 shows some representative views of the local B-factor.

Figure 13: Histogram of the local B-factor according to LocBfactor.

(a) View 1	(b) View 2	(c) View 3

Figure 14: Local B-factor according to LocBfactor. Views generated by ChimeraX at a the following X, Y, Z angles: View 1 (0,0,0), View 2 (90, 0, 0), View 3 (0, 90, 0).

Automatic criteria: The validation is OK if the median B-factor is in the range [-300,0].

STATUS: OK

2.7 Level 0.g Local Occupancy

Explanation:
LocOccupancy [Kaur et al., 2021] estimates the occupancy of a voxel by the macromolecule.

Results:

Fig. 15 shows the histogram of the local occupancy according to LocOccupancy. Some representative percentiles are:

Percentile	Local Occupancy [0-1]
2.5%	0.18
25%	0.64
50%	0.82
75%	0.91
97.5%	1.00
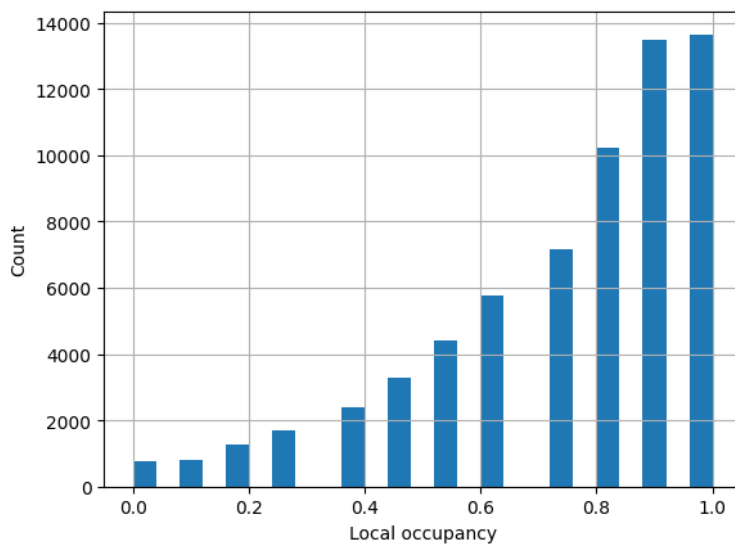
19

Fig. 16 shows some representative views of the local occupancy.

Figure 15: Histogram of the local occupancy according to LocOccupancy.

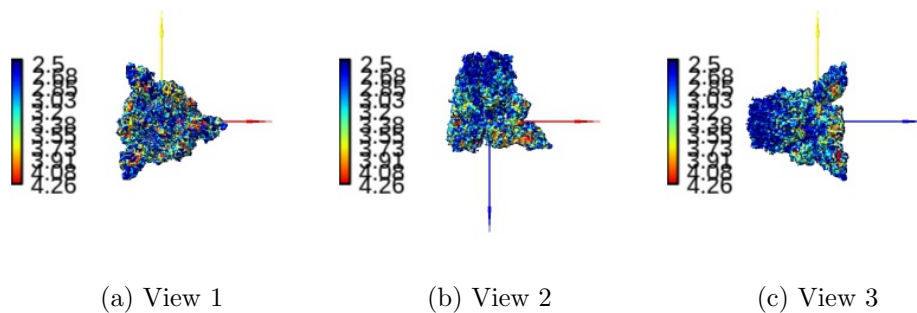| (a) View 1 | (b) View 2 | (c) View 3 |

Figure 16: Local occupancy according to LocOccupancy. Views generated by ChimeraX at a the following X, Y, Z angles: View 1 (0,0,0), View 2 (90, 0, 0), View 3 (0, 90, 0).

Automatic criteria: The validation is OK if the median occupancy is larger than 50%.

STATUS: OK

2.8   Level 0.h Hand correction

Explanation:
Deep Hand determines the correction of the hand for those maps with a resolution smaller than 5Å. The method calculates a value between 0 (correct hand) and 1 (incorrect hand) using a neural network to assign its hand.

Results:
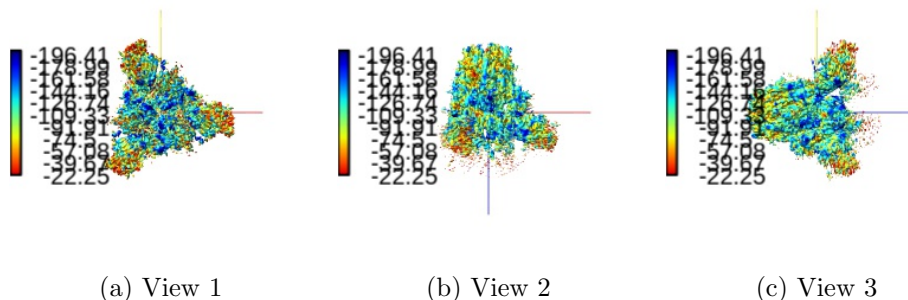
Deep hand assigns a score of 0.500 to the input volume.
Automatic criteria: The validation is OK if the deep hand score is smaller than 0.5.

WARNINGS: 1 warnings

1. **The orientation of the volume is uncertain.**

3   Atomic model

Atomic model: /home/coss/data/Dropbox/Aplicaciones/ShareLaTeX/MapValidation/-EMDB22301/6xs6_updated.cif

See Fig. 17.
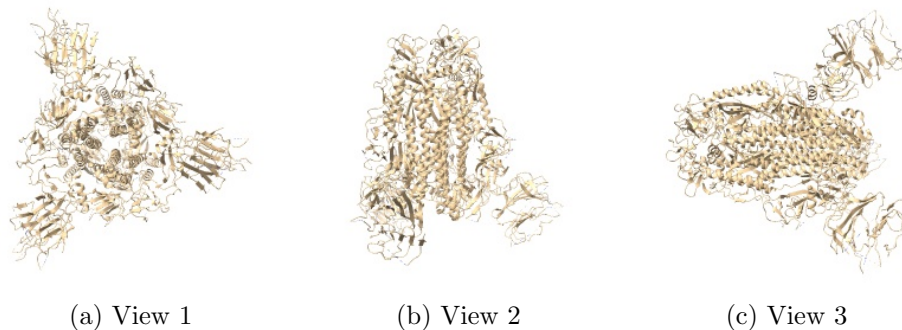
(a) View 1   (b) View 2   (c) View 3

Figure 17: Input atomic model Views generated by ChimeraX at a the following X, Y, Z angles: View 1 (0,0,0), View 2 (90, 0, 0), View 3 (0, 90, 0).

4   Level A analysis

4.1   Level A.a MapQ

Explanation:
MapQ [Pintilie et al., 2020] computes the local correlation between the map and each one of its atoms assumed to have a Gaussian shape.

Results:

Fig. 18 shows the histogram of the Q-score according calculated by MapQ. Some representative percentiles are:

Percentile	MapQ score [0-1]
2.5%	-0.38
25%	0.00
50%	0.00
75%	0.00
97.5%	0.36
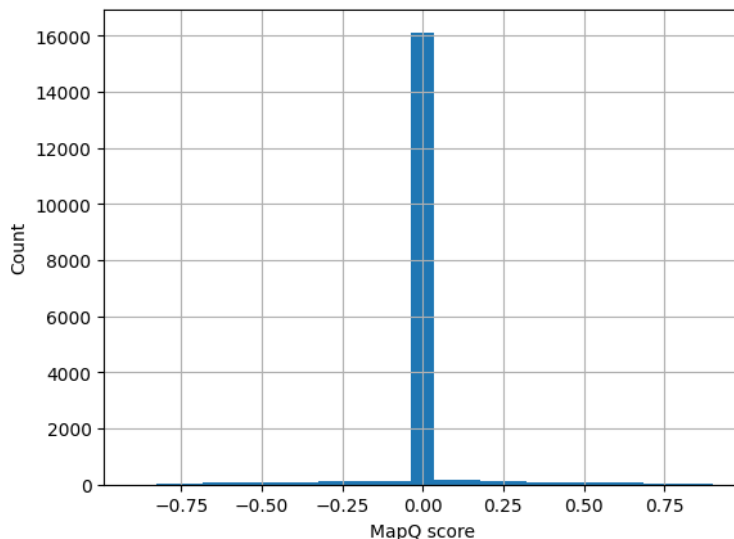
Figure 18: Histogram of the Q-score.

The following table shows the average Q score and estimated resolution for each chain.

Chain	Average Q score [0-1]	Estimated Resol. (Å)
A	0.00	6.3
B	0.00	6.3
C	-0.00	6.3

Automatic criteria: The validation is OK if the median Q-score is larger than 0.1.

WARNINGS: 1 warnings

1. **The median Q-score is less than 0.1.**

4.2 Level A.d Map-Model Guinier analysis

Explanation:
We compared the Guinier plot [Rosenthal and Henderson, 2003] of the atomic model and the experimental map. We made the mean of both profiles to be equal (and equal to the mean of the atomic model) to make sure that they had comparable scales.

Results:
Fig. 19 shows the logarithm (in natural units) of the structure factor (the module squared of the Fourier transform) of the atom model and the experimental map. The correlation between the two profiles was 0.955.

Figure 19: Guinier plot of the atom model and experimental map. The X-axis is the square of the inverse of the resolution in Å.

Automatic criteria: The validation is OK if the correlation between the two Guinier profiles is larger than 0.5.

STATUS: OK

4.3   Level A.e Phenix validation

Explanation:
Phenix provides a number of tools to assess the agreement between the experimental map and its atomic model [Afonine et al., 2018]. There are several cross-correlations to assess the quality of the fitting:

- CC (mask): Model map vs. experimental map correlation coefficient calculated considering map values inside a mask calculated around the macromolecule.
- CC (box): Model map vs. experimental map correlation coefficient calculated considering all grid points of the box.
- CC (volume) and CC (peaks) compare only map regions with the highest density values and regions below a certain contouring threshold level are ignored. CC (volume): The map region considered is defined by the N highest points inside the molecular mask. CC (peaks): In this case, calculations consider the union of regions defined by the N highest peaks in the model-calculated map and the N highest peaks in the experimental map.
- Local real-space correlation coefficients CC (main chain) and CC (side chain) involve the main skeleton chain and side chains, respectively.

There are also multiple ways of measuring the resolution:

- d99: Resolution cutoff beyond which Fourier map coefficients are negligibly small. Calculated from the full map.
- d_model: Resolution cutoff at which the model map is the most similar to the target (experimental) map. For d_model to be meaningful, the model is expected to fit the map as well as possible. d_model (B factors = 0) tries to avoid the blurring of the map.
- d_FSC_model; Resolution cutoff up to which the model and map Fourier coefficients are similar at FSC values of 0, 0.143, 0.5.

In addition to these resolution measurements the overall isotropic B factor is another indirect measure of the quality of the map.

Results:

To avoid ringing in Fourier space a smooth mask with a radius of 7.4 Å has been applied.

Overall correlation coefficients:

$$
\begin{aligned}
\text{CC (mask)} &= 0.590 \\
\text{CC (box)} &= 0.574 \\
\text{CC (volume)} &= 0.596 \\
\text{CC (peaks)} &= 0.412 \\
\text{CC (main chain)} &= 0.562 \\
\text{CC (side chain)} &= 0.558
\end{aligned}
$$

Correlation coefficients per chain:

Chain	Cross-correlation
A	0.550332
B	0.551502
C	0.551974

We now show the correlation profiles of the different chain per residue.
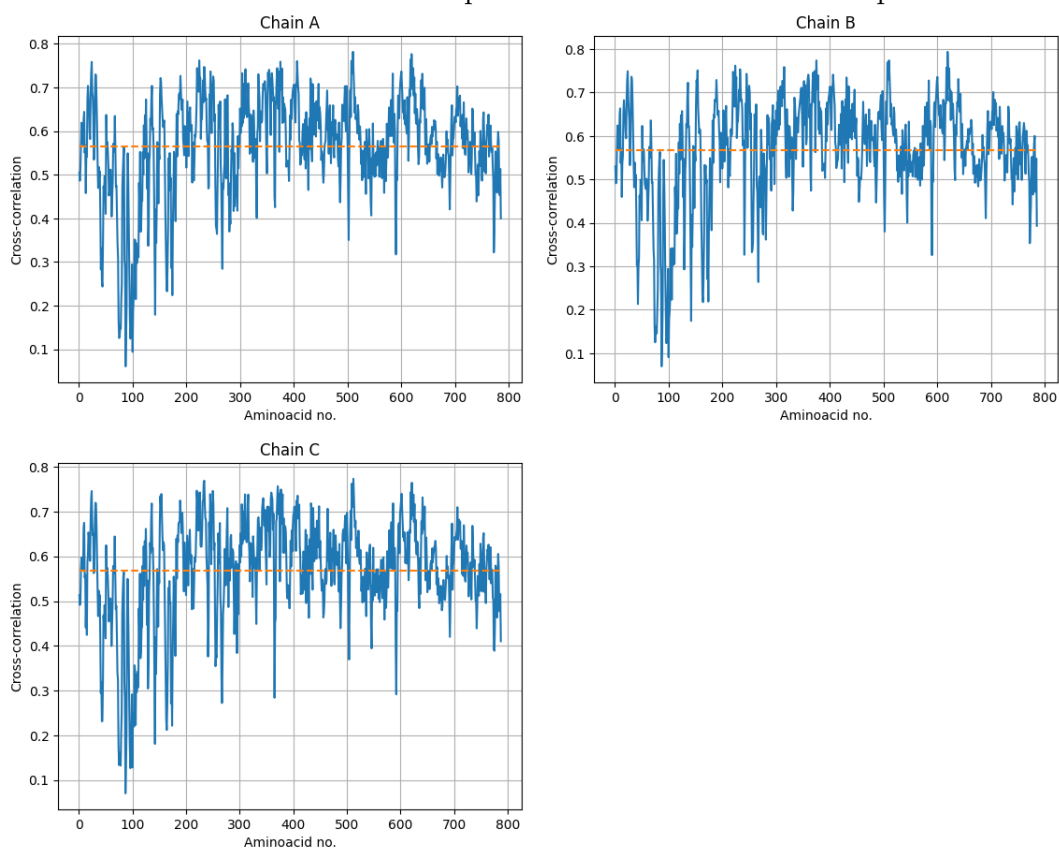
Fig. 20 shows the histogram of all cross-correlations evaluated at the residues. The percentage of residues whose correlation is below 0.5 is 20.7 %.
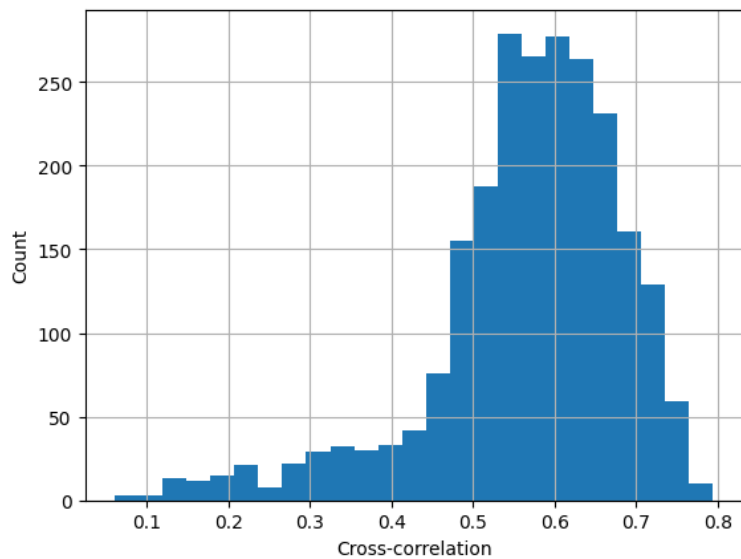
Figure 20: Histogram of the cross-correlation between the map and model evaluated for all residues.

Resolutions estimated from the model:

Resolution (Å)	Masked	Unmasked
d99	2.2	2.2
d_model	2.2	2.2
d_model (B-factor=0)	3.3	3.3
FSC_model=0	2.1	2.2
FSC_model=0.143	2.2	2.2
FSC_model=0.5	3.3	3.4

Overall isotropic B factor:

B factor	Masked	Unmasked
Overall B-iso	35.0	35.0

Fig. 21 shows the FSC between the input map and the model.

Figure 21: FSC between the input map and model with and without a mask constructed from the model. The X-axis is the square of the inverse of the resolution in Å.

Automatic criteria: The validation is OK if 1) the percentage of residues whose correlation is smaller than 0.5 is smaller than 10%, and 2) the resolution reported by the user is larger than 0.8 times the resolution estimated between the map and model at FSC=0.5.

WARNINGS: 1 warnings

1. **The percentage of residues that have a cross-correlation below 0.5 is 20.7, that is larger than 10%**

4.4 Level A.f EMRinger validation

Explanation:
EMringer [Barad et al., 2015] compares the side chains of the atomic model to the CryoEM map. The following features are reported:

- Optimal Threshold: Electron potential map cutoff value at which the maximum EMRinger score was obtained.
- Rotamer Ratio: Fraction of rotameric residues at the Optimal threshold value.
- Max Zscore: Z-score computed to determine the significance of the distribution at the Optimal threshold value.
- Model Length: Total of non-gamma-branched, non-proline aminoacids with a non-H gamma atom used in global EMRinger score computation.
- EMRinger Score: Maximum EMRinger score calculated at the Optimal Threshold.

A rotameric residue is one in which EMRinger peaks that fall within defined rotamers based on chi1, this often suggests a problem with the modelling of the backbone. In general, the user should look at the profiles and identify regions that may need improvement.

Results:

General results:

Optimal threshold	0.200989
Rotamer ratio	0.744
Max. Zscore	8.59
Model length	1394
EMRinger Score	2.301

Fig. 22 shows the EMRinger score and fraction of rotameric residues as a function of the map threshold. The optimal threshold was selected looking for the maximum EMRinger score in this plot.
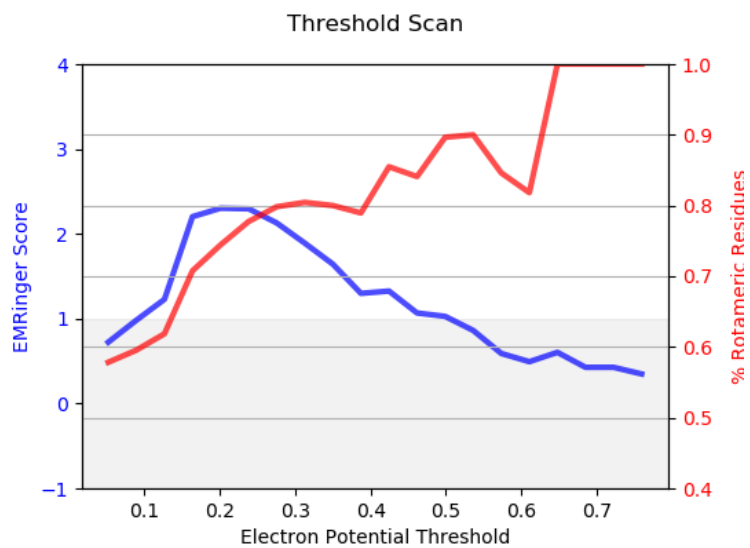
Figure 22: EMRinger score and fraction of rotameric residues as a function of the map threshold.

Fig. 23 shows the histogram for rotameric (blue) and non-rotameric (red) residues at the optimal threshold.
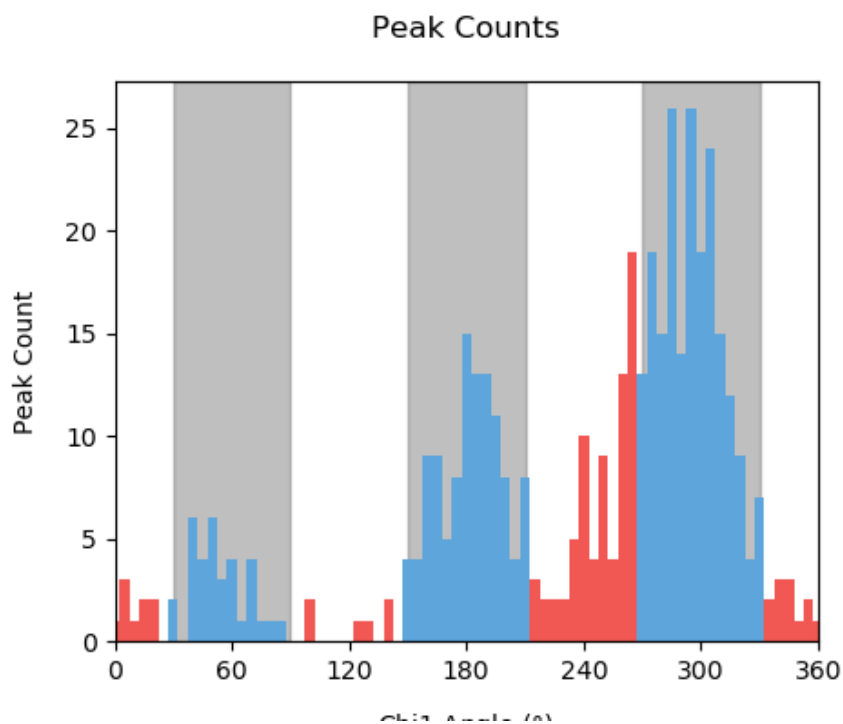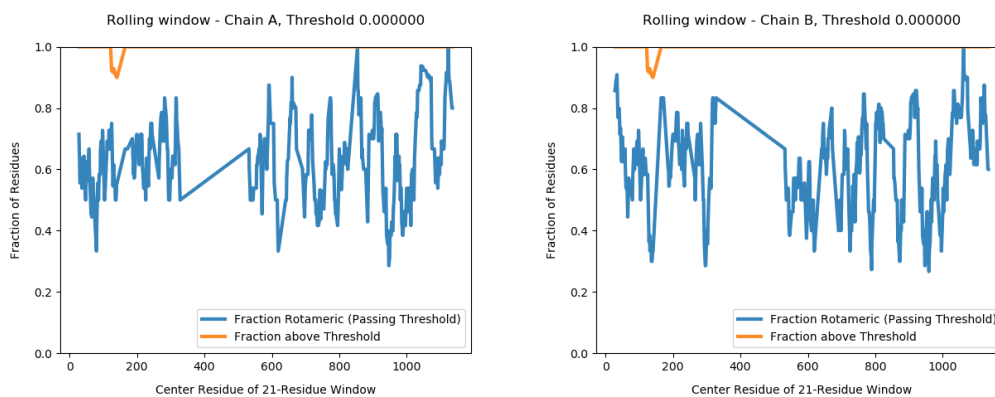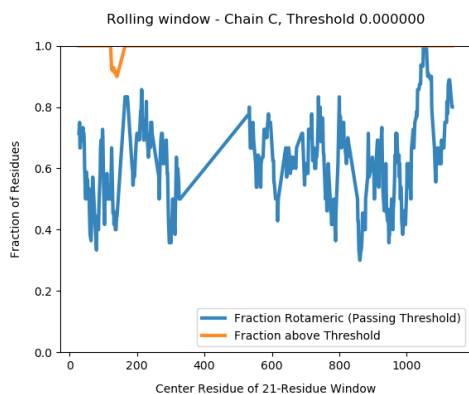
Figure 23: Histogram for rotameric (blue) and non-rotameric (red) residues at the optimal threshold as a function of the angle Chi1.

The following plots show the rolling window EMRinger analysis of the different chains to distinguish regions of improved model quality. This analysis was performed on rolling sliding 21-residue windows along the primary sequence of the protein chains.

31

Rolling window - Chain C, Threshold 0.000000

Automatic criteria: The validation is OK if the EMRinger score and Max. Zscore are larger than 1.

STATUS: OK

4.5 Level A.g DAQ validation

Explanation:
DAQ [Terashi et al., 2022] is a computational tool using deep learning that can estimate the residue-wise local quality for protein models from cryo-Electron Microscopy maps. The method calculates the likelihood that a given density feature corresponds to an aminoacid, atom, and secondary structure. These likelihoods are combined into a score that ranges from -1 (bad quality) to 1 (good quality).

Results:
Fig. 24 shows the histogram of the DAQ values. The mean and standard deviation were 0.1 and 0.2, respectively.

Figure 24: Histogram of the DAQ values.

The atomic model colored by DAQ can be seen in Fig. 25.

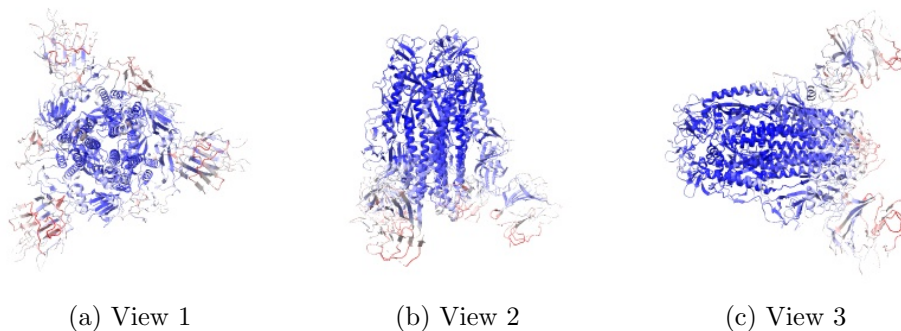(a) View 1   (b) View 2   (c) View 3

Figure 25: Atomic model colored by DAQ Views generated by ChimeraX at a the following X, Y, Z angles: View 1 (0,0,0), View 2 (90, 0, 0), View 3 (0, 90, 0).

Automatic criteria: The validation is OK if the average DAQ score is larger than 0.5.

WARNINGS: 1 warnings

1. **The average DAQ is smaller than 0.5.**

References

[Rosenthal and Henderson, 2003] Rosenthal, P. B. and Henderson, R. (2003). Optimal determination of particle orientation, absolute hand, and contrast loss in single particle electron-cryomicroscopy. *J. Molecular Biology*, 333:721–745.

[Ramírez-Aportela et al., 2019] Ramírez-Aportela, E., Mota, J., Conesa, P., Carazo, J. M., and Sorzano, C. O. S. (2019). DeepRes: a new deep-learning- and aspect-based local resolution method for electron-microscopy maps. *IUCRj*, 6:1054–1063.

[Kaur et al., 2021] Kaur, S., Gomez-Blanco, J., Khalifa, A. A., Adinarayanan, S., Sanchez-Garcia, R., Wrapp, D., McLellan, J. S., Bui, K. H., and Vargas, J. (2021). Local computational methods to improve the interpretability and analysis of cryo-EM maps. *Nature Communications*, 12(1):1–12.

[Pintilie et al., 2020] Pintilie, G., Zhang, K., Su, Z., Li, S., Schmid, M. F., and Chiu, W. (2020). Measurement of atom resolvability in cryo-em maps with q-scores. *Nature methods*, 17(3):328–334.

[Afonine et al., 2018] Afonine, P. V., Klaholz, B. P., Moriarty, N. W., Poon, B. K., Sobolev, O. V., Terwilliger, T. C., Adams, P. D., and Urzhumtsev, A. (2018). New tools for the analysis and validation of cryo-EM maps and atomic models. *Acta Crystallographica D, Struct. Biol.*, 74:814–840.

[Barad et al., 2015] Barad, B. A., Echols, N., Wang, R. Y.-R., Cheng, Y., DiMaio, F., Adams, P. D., and Fraser, J. S. (2015). EMRinger: side chain-directed model and map validation for 3D cryo-electron microscopy. *Nature Methods*, 12(10):943–946.

[Terashi et al., 2022] Terashi, G., Wang, X., Subramaniya, S.R.M.V., Tesmer, J.J.G. and Kihara, D. (2022). Residue-Wise Local Quality Estimation for Protein Models from Cryo-EM Maps. (submitted).

Validation report of Level(s) 0, A

I²PC Validation server

March 2, 2022
6:55am

Abstract

The map seems to be well centered. There is no problem with the suggested threshold. There seems to be a problem with the map's background (see Sec. 2.3). There seems to be a problem with its local B-factor (see Sec. 2.6). There seems to be a problem with the map hand (see Sec. 2.8). There seems to be a problem with its MapQ scores (see Sec. 4.1). The EMRinger score is negative, it seems that the model side chains do not match the map (see Sec. 4.4). DAQ detects some mismatch between the map and its model (see Sec. 4.5).

The average resolution of the map estimated by various methods goes from 1.1Å to 6.4Å with an average of 3.9Å. The resolution provided by the user was 3.8Å.

The overall score (passing tests) of this report is 6 out of 13 evaluable items.

Summary of the warnings across sections.

If it is empty below this point, it means that there are no warnings.

Section 2.3 (0.c Background analysis)

1. **The null hypothesis that the background mean is 0 has been rejected because the p-value of the comparison is smaller than 0.001**
2. **There is a significant proportion of outlier values in the background (cdf5 ratio=8586.32)**

Section 2.5 (0.e DeepRes)

1. **The reported resolution, 3.84 Å, is particularly with respect to the local resolution distribution. It occupies the 0.09 percentile**

Section 2.6 (0.f LocBfactor)

1. **The median B-factor is out of the interval [-300,0]**

Section 2.8 (0.h DeepHand)

1. **The volume seems to be flipped.**
2. **The orientation of the volume is uncertain.**

Section 4.1 (A.a MapQ)

1. **The median Q-score is less than 0.1.**

Section 4.4 (A.f EMRinger)

1. **The EMRinger score is smaller than 1, it is 0.748.**

Section 4.5 (A.g DAQ)

1. **The average DAQ is smaller than 0.5.**

3

Contents

1 Input data

Input map: /home/coss/data/Dropbox/Aplicaciones/ShareLaTeX/MapValidation/-
EMDB22838/emd_22838.map
SHA256 hash: 2ae2d0b9aed1e9c2fab2e32a643720f604861daf3e03307d44056083415153b4
Voxel size: 1.058000 (Å)
Visualization threshold: 0.200000
Resolution estimated by user: 3.840000

Orthogonal slices of the input map
Explanation:
In the orthogonal slices of the map, the noise outside the protein should not
have any structure (stripes going out, small blobs, particularly high or low
densities, ...)
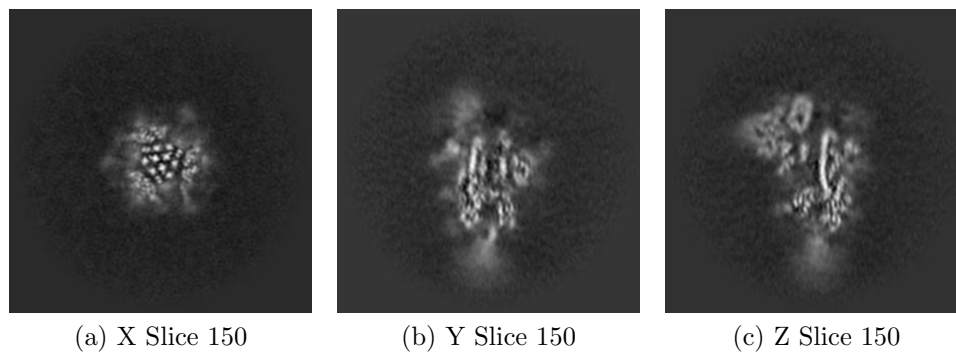
Results:
See Fig. 1.

(a) X Slice 150   (b) Y Slice 150   (c) Z Slice 150

Figure 1: Central slices of the input map in the three dimensions

Orthogonal slices of maximum variance of the input map
Results:
See Fig. 2.

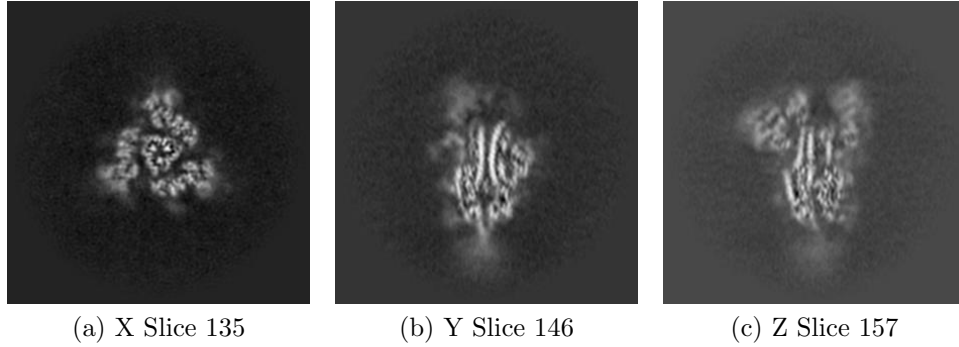(a) X Slice 135   (b) Y Slice 146   (c) Z Slice 157

Figure 2: Slices of maximum variation in the three dimensions

Orthogonal projections of the input map

Explanation:

In the projections there should not be stripes (this is an indication of directional overweighting, or angular attraction), and there should not be a dark halo around or inside the structure (this is an indication of incorrect CTF correction or the reconstruction of a biased map).
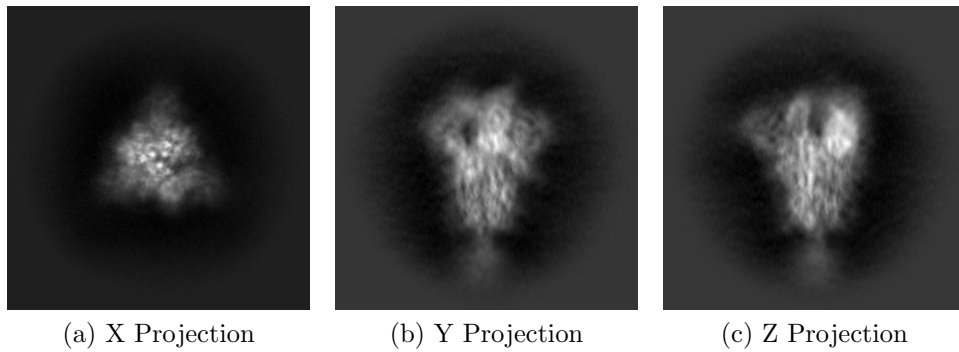
Results:

See Fig. 3.

(a) X Projection   (b) Y Projection   (c) Z Projection

Figure 3: Projections in the three dimensions

Isosurface views of the input map

Explanation:

An isosurface is the surface of all points that have the same gray value. In these views there should not be many artifacts or noise blobs around the map.

Results:
See Fig. 4.

(a) View 1   (b) View 2   (c) View 3
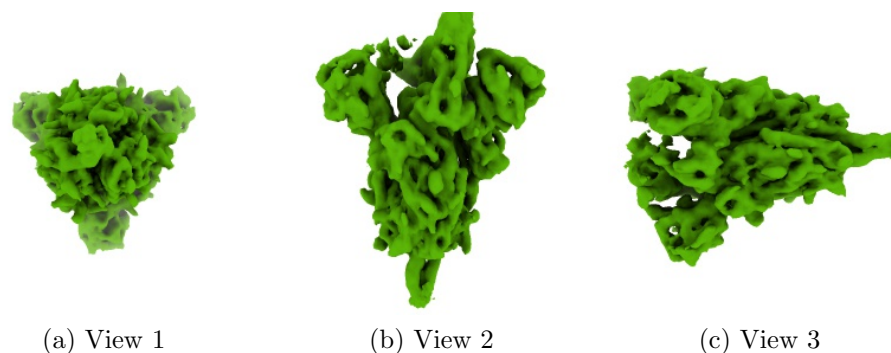
Figure 4: Isosurface at threshold=0.200000. Views generated by ChimeraX at a the following X, Y, Z angles: View 1 (0,0,0), View 2 (90, 0, 0), View 3 (0, 90, 0).

Orthogonal slices of maximum variance of the mask
Explanation:
The mask has been calculated at the suggested threshold 0.200000, the largest connected component was selected, and then dilated by 2Å.
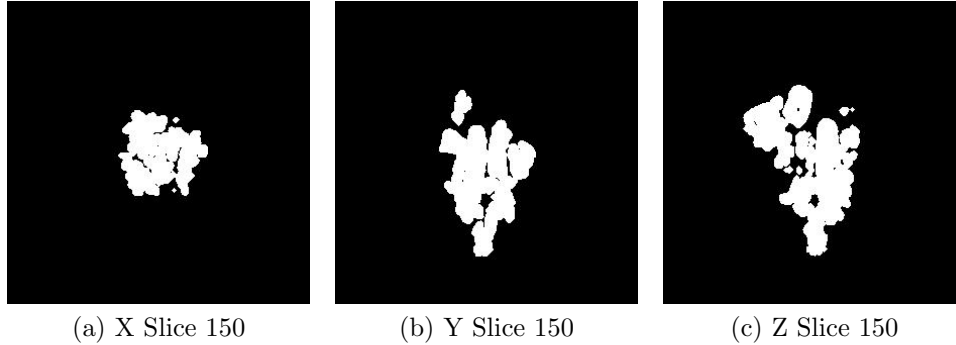
Results:
See Fig. 5.

(a) X Slice 150   (b) Y Slice 150   (c) Z Slice 150

Figure 5: Slices of maximum variation in the three dimensions of the mask

2 Level 0 analysis

2.1 Level 0.a Mass analysis

Explanation:
The reconstructed map must be relatively well centered in the box, and there should be at least 30Å (the exact size depends on the CTF) on each side to make sure that the CTF can be appropriately corrected.

Results:
The space from the left and right in X are 82.52 and 50.78 Å, respectively. There is a decentering ratio (abs(Right-Left)/Size)% of 10.00%

The space from the left and right in Y are 82.52 and 101.57 Å, respectively. There is a decentering ratio (abs(Right-Left)/Size)% of 6.00%

The space from the left and right in Z are 92.05 and 87.81 Å, respectively. There is a decentering ratio (abs(Right-Left)/Size)% of 1.33%

The center of mass is at (x,y,z)=(148.88,144.32,145.88). The decentering of the center of mass (abs(Center)/Size)% is 0.37, 1.89, and 1.37, respectively.%

Automatic criteria: The validation is OK if 1) the decentering and

8

center of mass less than 20% of the map dimensions in all directions, and 2) the extra space on each direction is more than 20% of the map dimensions.

STATUS: OK

2.2   Level 0.b Mask analysis

Explanation:
The map at the suggested threshold should have most of its mass concentrated in a single connected component. It is normal that after thresholding there are a few thousands of very small, disconnected noise blobs. However, there total mass should not exceed 10%. The raw mask (just thresholding) and the mask constructed for the analysis (thresholding + largest connected component + dilation) should significantly overlap. Overlap is defined by the overlapping coefficient (size(Raw AND Constructed)/size(Raw)) that is a number between 0 and 1, the closer to 1, the more they agree.

Results:

Raw mask: At threshold 0.200000, there are 24 connected components with a total number of voxels of 236482 and a volume of 280062.58 $Å^3$ (see Fig. 6). The size and percentage of the total number of voxels for the raw mask are listed below (up to 95% of the mass or the first 100 clusters, whatever happens first), the list contains (No. voxels (volume in $Å^3$), percentage, cumulatedPercentage):

, (236064 (279567.55), 99.82, 99.82)

Number of components to reach 95% of the mass: 2

The average size of the remaining 22 components is 19.00 voxels (1.18 $Å^3$). Their size go from 273 voxels (323.31 $Å^3$) to 1 voxels (1.18 $Å^3$).

The slices of the raw mask can be seen in Fig. 6.

9

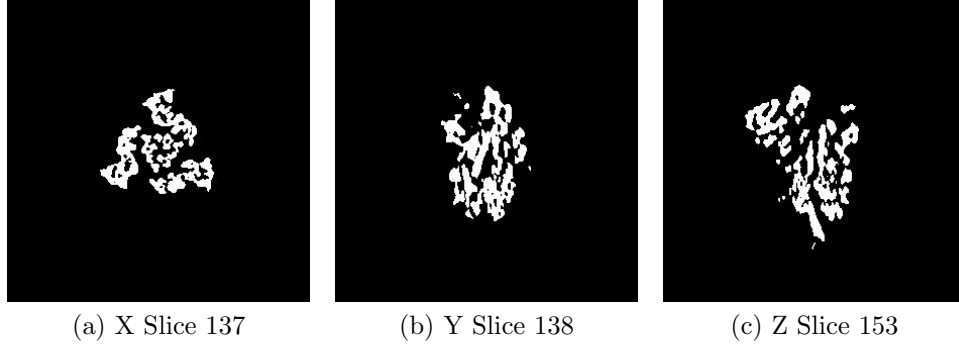(a) X Slice 137   (b) Y Slice 138   (c) Z Slice 153

Figure 6: Maximum variance slices in the three dimensions of the raw mask

The following table shows the variation of the mass enclosed at different thresholds (see Fig. 7):

| Threshold	Voxel mass	Molecular mass(kDa)	# Aminoacids
0.0333	961872.00	943.77	8579.74
0.0665	692613.00	679.58	6178.00
0.0998	518448.00	508.69	4624.47
0.1330	399920.00	392.39	3567.22
0.1663	310352.00	304.51	2768.29
0.1996	237318.00	232.85	2116.84
0.2328	178114.00	174.76	1588.75
0.2661	132228.00	129.74	1179.45
0.2994	96304.00	94.49	859.02
0.3326	69000.00	67.70	615.47
0.3659	48618.00	47.70	433.66
0.3991	33917.00	33.28	302.53
0.4324	23468.00	23.03	209.33
0.4657	16279.00	15.97	145.21
0.4989	11193.00	10.98	99.84
0.5322	7656.00	7.51	68.29
0.5654	5010.00	4.92	44.69
0.5987	3108.00	3.05	27.72
0.6320	1832.00	1.80	16.34
0.6652	917.00	0.90	8.18
0.6985	372.00	0.36	3.32
0.7318	120.00	0.12	1.07
0.7650	32.00	0.03	0.29
0.7983	6.00	0.01	0.05
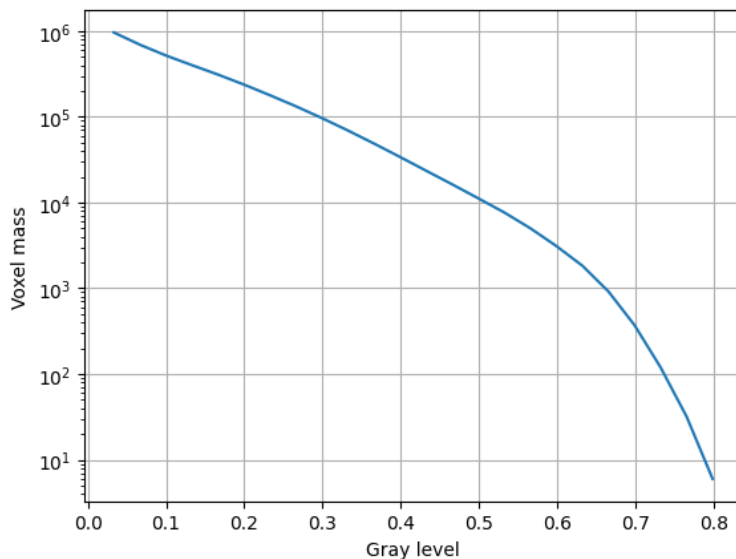
Figure 7: Voxel mass as a function of the gray level.

Constructed mask: After keeping the largest component of the previous mask and dilating it by 2Å, there is a total number of voxels of 528227 and a volume of 625572.43 Å^3. The overlap between the raw and constructed mask is 1.00.

Automatic criteria: The validation is OK if 1) to keep 95% of the mass we need to keep at most 5 connected components; and 2) the average volume of the blobs outside the given threshold has a size smaller than 5Å^3; and 3) the overlap between the raw mask and the mask constructed for the analysis is larger than 75%.

STATUS: OK

2.3 Level 0.c Background analysis

Explanation:
Background is defined as the region outside the macromolecule mask. The background mean should be zero, and the number of voxels with a very low or very high value (below 5 standard deviations of the noise) should be very

small and they should be randomly distributed without any specific structure. Sometimes, you can see some structure due to the symmetry of the structure.

Results:

The null hypothesis that the background mean is 0 was tested with a one-sample Student's t-test. The resulting t-statistic and p-value were -1661.72 and 0.000000, respectively.

The mean and standard deviation (sigma) of the background were -0.005457 and 0.016896. The percentage of background voxels whose absolute value is larger than 5 times the standard deviation is 0.49 % (see Fig. 8). The same percentage from a Gaussian would be 0.000057% (ratio between the two percentages: 8586.319777).

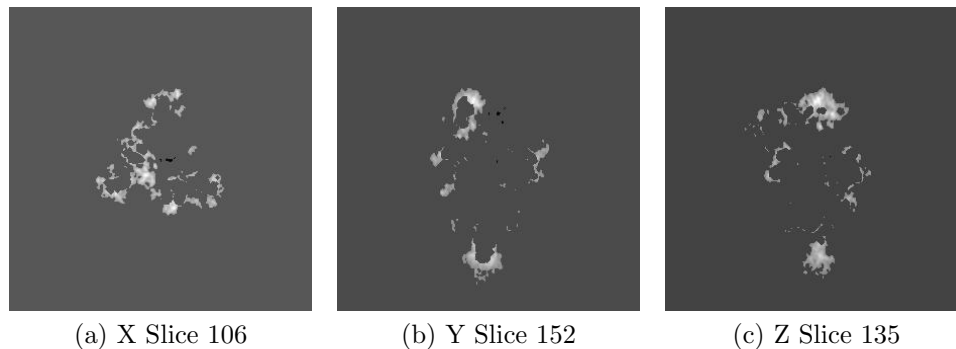Slices of the background beyond 5*sigma can be seen in Fig. 8.

| (a) X Slice 106 | (b) Y Slice 152 | (c) Z Slice 135 |

Figure 8: Maximum variance slices in the three dimensions of the parts of the background beyond 5*sigma

Automatic criteria: The validation is OK if 1) the p-value of the null hypothesis that the background has 0 mean is larger than 0.001; and 2) the number of voxels above or below 5 sigma is smaller than 20 times the amount expected for a Gaussian with the same standard deviation whose mean is 0.

WARNINGS: 2 warnings

2.4   Level 0.d B-factor analysis

Explanation:

The B-factor line [Rosenthal and Henderson, 2003] fitted between 15Åand the resolution reported should have a slope that is between 0 and 300 Å^2.

Results:

Fig. 9 shows the logarithm (in natural units) of the structure factor (the module squared of the Fourier transform) of the experimental map, its fitted line, and the corrected map. The estimated B-factor was -155.5. The fitted line was $\log(|F|^2) = -38.9/R^2 + (-10.5)$.
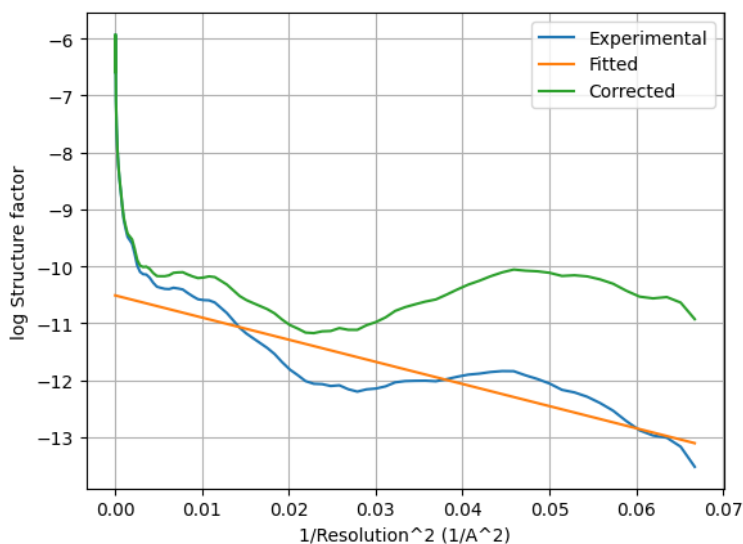

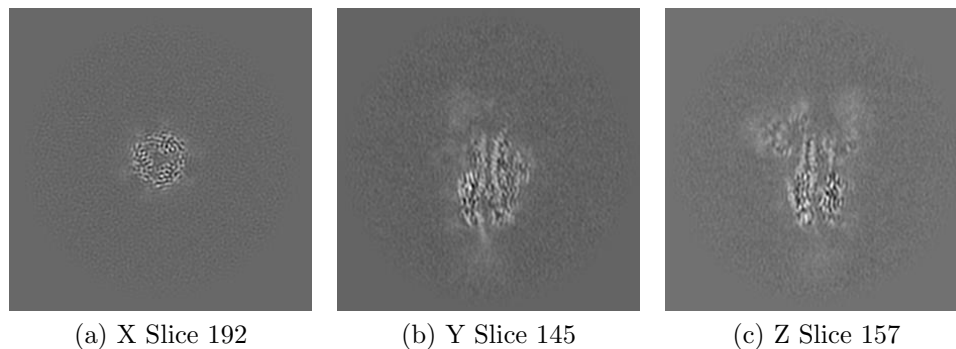Figure 9: Guinier plot. The X-axis is the square of the inverse of the resolution in Å.

13

| (a) X Slice 192 | (b) Y Slice 145 | (c) Z Slice 157 |

Figure 10: Slices of maximum variation in the three dimensions of the B-factor corrected map

Automatic criteria: The validation is OK if the B-factor is in the range [-300,0].

STATUS: OK

2.5 Level 0.e Local resolution with DeepRes

Explanation:
DeepRes [Ramírez-Aportela et al., 2019] measures the local resolution using a neural network that has been trained on the appearance of atomic structures at different resolutions. Then, by comparing the local appearance of the input map to the appearance of the atomic structures a local resolution label can be assigned.

Results:

Fig. 11 shows the histogram of the local resolution according to DeepRes. Some representative percentiles are:

Percentile	Resolution(Å)
2.5%	4.41
25%	5.54
50%	6.38
75%	7.27
97.5%	9.16

The reported resolution, 3.84 Å, is at the percentile 0.1. Fig. 12 shows some representative views of the local resolution.
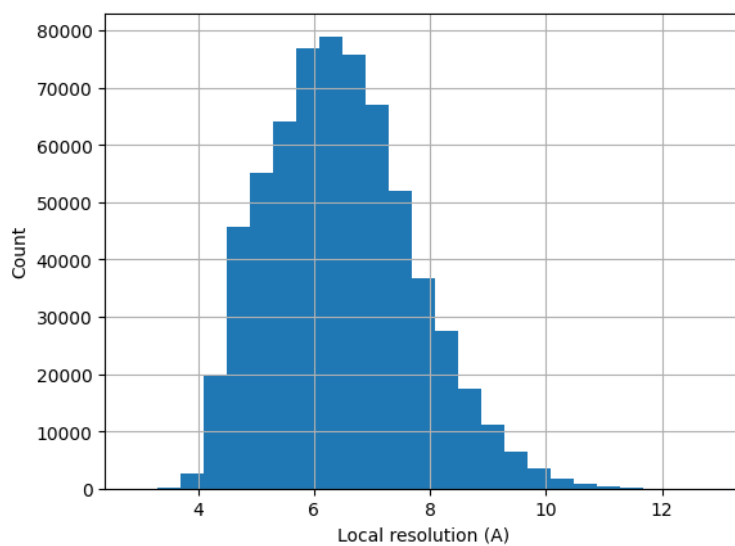

Figure 11: Histogram of the local resolution according to deepres.

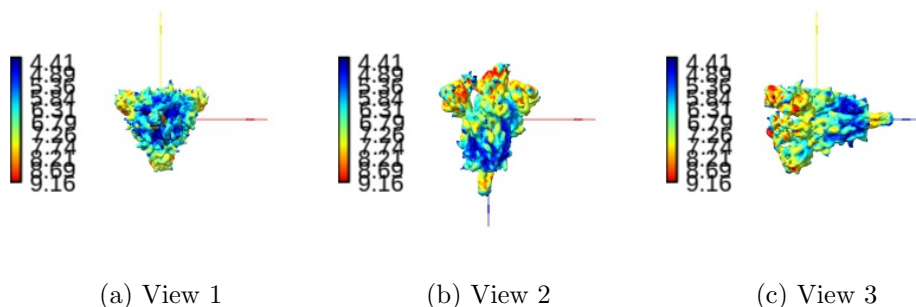(a) View 1   (b) View 2   (c) View 3

Figure 12: Local resolution according to DeepRes. Views generated by ChimeraX at a the following X, Y, Z angles: View 1 (0,0,0), View 2 (90, 0, 0), View 3 (0, 90, 0).

Automatic criteria: The validation is OK if the percentile of the user provided resolution is larger than 0.1% of the percentile of the local resolution as estimated by DeepRes.

WARNINGS: 1 warnings

1. **The reported resolution, 3.84 Å, is particularly with respect to the local resolution distribution. It occupies the 0.09 percentile**

2.6 Level 0.f Local B-factor

Explanation:
LocBfactor [Kaur et al., 2021] estimates a local resolution B-factor by decomposing the input map into a local magnitude and phase term using the spiral transform.

Results:

Fig. 13 shows the histogram of the local B-factor according to LocBfactor. Some representative percentiles are:

16

Percentile	Local B-factor (Å^{-2})
2.5%	-444.04
25%	-367.63
50%	-327.21
75%	-285.93
97.5%	-206.83

Fig. 14 shows some representative views of the local B-factor.
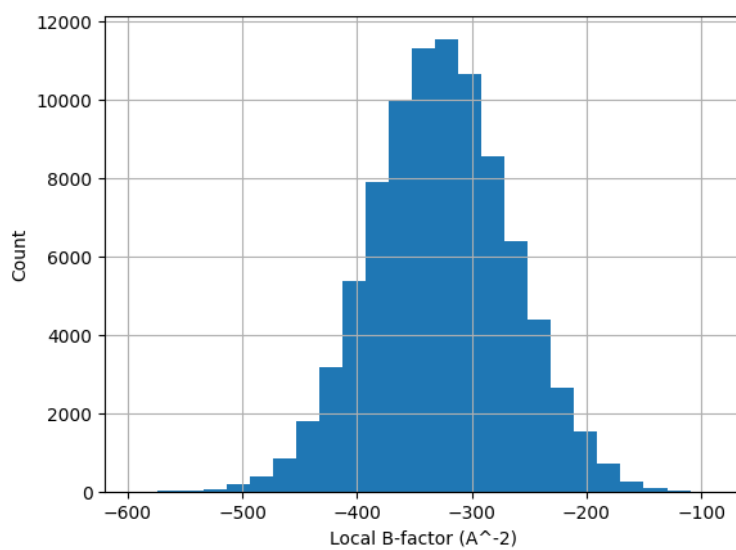
Figure 13: Histogram of the local B-factor according to LocBfactor.
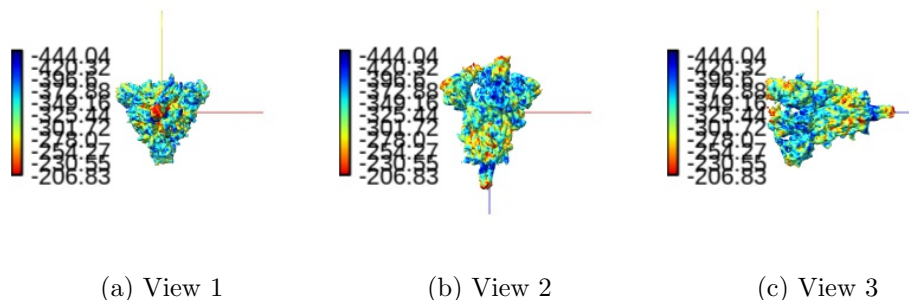
(a) View 1   (b) View 2   (c) View 3

Figure 14: Local B-factor according to LocBfactor. Views generated by ChimeraX at a the following X, Y, Z angles: View 1 (0,0,0), View 2 (90, 0, 0), View 3 (0, 90, 0).

Automatic criteria: The validation is OK if the median B-factor is in the range [-300,0].

WARNINGS: 1 warnings

1. **The median B-factor is out of the interval [-300,0]**

2.7   Level 0.g Local Occupancy

Explanation:
LocOccupancy [Kaur et al., 2021] estimates the occupancy of a voxel by the macromolecule.

Results:

Fig. 15 shows the histogram of the local occupancy according to LocOccupancy. Some representative percentiles are:

Percentile	Local Occupancy [0-1]
2.5%	0.08
25%	0.58
50%	0.83
75%	1.00
97.5%	1.00

Fig. 16 shows some representative views of the local occupancy.

Figure 15: Histogram of the local occupancy according to LocOccupancy.

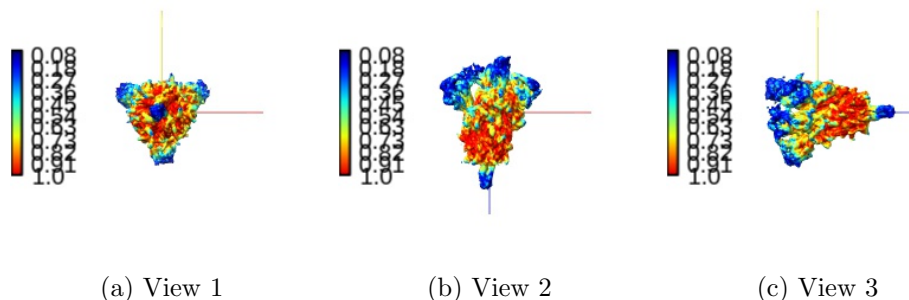|		
(a) View 1	(b) View 2	(c) View 3

Figure 16: Local occupancy according to LocOccupancy. Views generated by ChimeraX at a the following X, Y, Z angles: View 1 (0,0,0), View 2 (90, 0, 0), View 3 (0, 90, 0).

Automatic criteria: The validation is OK if the median occupancy is larger than 50%.

STATUS: OK

2.8   Level 0.h Hand correction

Explanation:
Deep Hand determines the correction of the hand for those maps with a resolution smaller than 5Å. The method calculates a value between 0 (correct hand) and 1 (incorrect hand) using a neural network to assign its hand.

Results:

Deep hand assigns a score of 0.552 to the input volume.
Automatic criteria: The validation is OK if the deep hand score is smaller than 0.5.

WARNINGS: 2 warnings

1. **The volume seems to be flipped.**
2. **The orientation of the volume is uncertain.**

3   Atomic model

Atomic model: /home/coss/data/Dropbox/Aplicaciones/ShareLaTeX/MapValidation/-EMDB22838/7kec_updated.pdb

See Fig. 17.

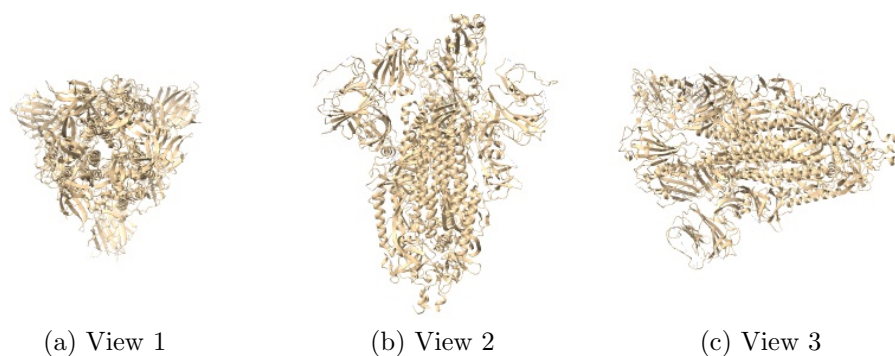(a) View 1   (b) View 2   (c) View 3

Figure 17: Input atomic model Views generated by ChimeraX at a the following X, Y, Z angles: View 1 (0,0,0), View 2 (90, 0, 0), View 3 (0, 90, 0).

4   Level A analysis

4.1   Level A.a MapQ

Explanation:
MapQ [Pintilie et al., 2020] computes the local correlation between the map and each one of its atoms assumed to have a Gaussian shape.

Results:

Fig. 18 shows the histogram of the Q-score according calculated by MapQ. Some representative percentiles are:

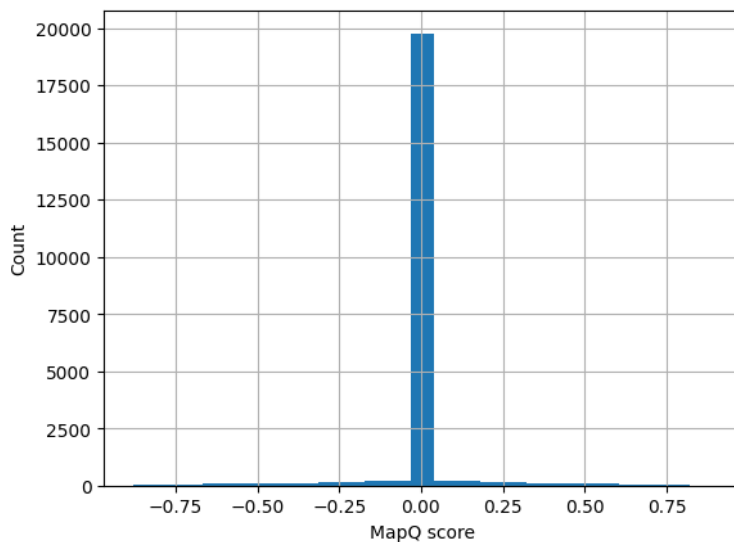| Percentile	MapQ score [0-1]
2.5%	-0.35
25%	0.00
50%	0.00
75%	0.00
97.5%	0.33

Figure 18: Histogram of the Q-score.

The following table shows the average Q score and estimated resolution for each chain.

Chain	Average Q score [0-1]	Estimated Resol. (Å)
A	-0.00	6.3
A	0.00	0.0
B	-0.00	6.3
B	-0.01	0.0
C	-0.00	6.3
C	0.00	0.0
D	0.00	0.0
E	0.00	0.0
F	0.00	0.0
G	0.00	0.0
H	0.00	0.0
I	-0.00	0.0
J	0.00	0.0
K	0.00	0.0
L	0.05	0.0
M	0.15	0.0
N	0.10	0.0
O	0.00	0.0

Automatic criteria: The validation is OK if the median Q-score is larger than 0.1.

WARNINGS: 1 warnings

1. **The median Q-score is less than 0.1.**

4.2 Level A.d Map-Model Guinier analysis

Explanation:
We compared the Guinier plot [Rosenthal and Henderson, 2003] of the atomic model and the experimental map. We made the mean of both profiles to be equal (and equal to the mean of the atomic model) to make sure that they had comparable scales.

Results:
Fig. 19 shows the logarithm (in natural units) of the structure factor (the module squared of the Fourier transform) of the atom model and the exper-

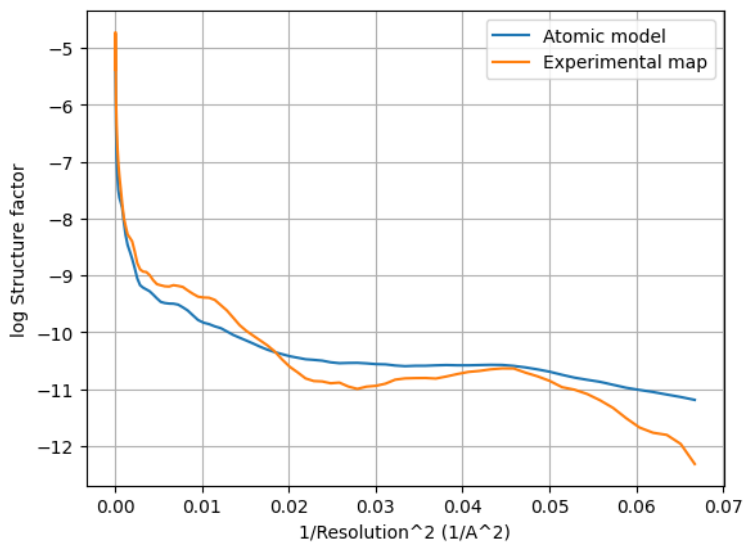imental map. The correlation between the two profiles was 0.982.

Figure 19: Guinier plot of the atom model and experimental map. The X-axis is the square of the inverse of the resolution in Å.

Automatic criteria: The validation is OK if the correlation between the two Guinier profiles is larger than 0.5.

STATUS: OK

4.3   Level A.e Phenix validation

Explanation:
Phenix provides a number of tools to assess the agreement between the experimental map and its atomic model [Afonine et al., 2018]. There are several cross-correlations to assess the quality of the fitting:

- CC (mask): Model map vs. experimental map correlation coefficient calculated considering map values inside a mask calculated around the macromolecule.
- CC (box): Model map vs. experimental map correlation coefficient calculated considering all grid points of the box.

- CC (volume) and CC (peaks) compare only map regions with the highest density values and regions below a certain contouring threshold level are ignored. CC (volume): The map region considered is defined by the N highest points inside the molecular mask. CC (peaks): In this case, calculations consider the union of regions defined by the N highest peaks in the model-calculated map and the N highest peaks in the experimental map.
- Local real-space correlation coefficients CC (main chain) and CC (side chain) involve the main skeleton chain and side chains, respectively.

There are also multiple ways of measuring the resolution:
- d99: Resolution cutoff beyond which Fourier map coefficients are negligibly small. Calculated from the full map.
- d_model: Resolution cutoff at which the model map is the most similar to the target (experimental) map. For d_model to be meaningful, the model is expected to fit the map as well as possible. d_model (B factors = 0) tries to avoid the blurring of the map.
- d_FSC_model; Resolution cutoff up to which the model and map Fourier coefficients are similar at FSC values of 0, 0.143, 0.5.

In addition to these resolution measurements the overall isotropic B factor is another indirect measure of the quality of the map.

Results:

To avoid ringing in Fourier space a smooth mask with a radius of 7.7 Å has been applied.

Overall correlation coefficients:

$$
\begin{aligned}
\text{CC (mask)} &= \quad 0.619 \\
\text{CC (box)} &= \quad 0.737 \\
\text{CC (volume)} &= \quad 0.659 \\
\text{CC (peaks)} &= \quad 0.616 \\
\text{CC (main chain)} &= \quad 0.654 \\
\text{CC (side chain)} &= \quad 0.670
\end{aligned}
$$

Correlation coefficients per chain:

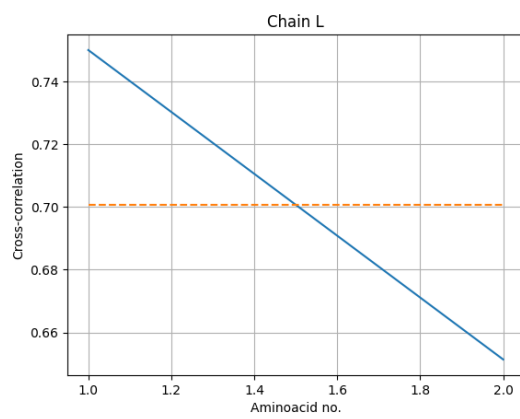| Chain	Cross-correlation
A	0.647861
B	0.674413
C	0.692197
D	0.280030
E	0.728180
F	0.746660
G	0.737342
H	0.719959
I	0.710197
J	0.709352
K	0.686594
L	0.703614
M	0.700790
N	0.698624
O	0.694403

We now show the correlation profiles of the different chain per residue.
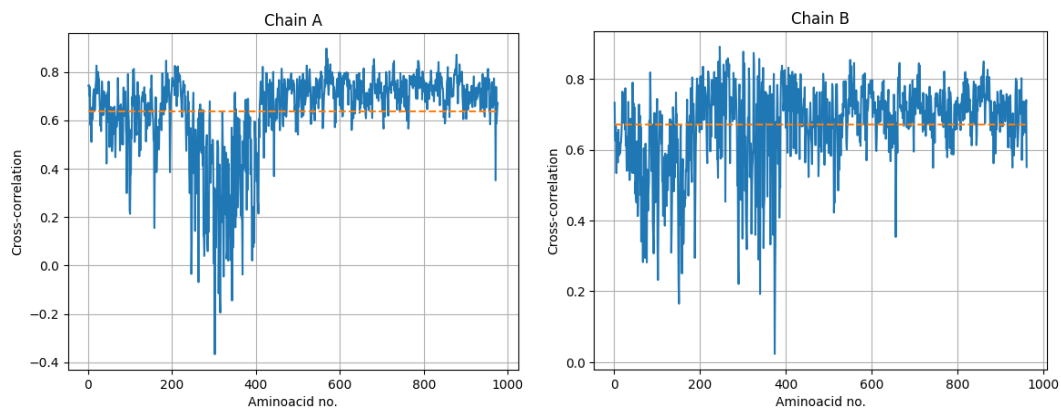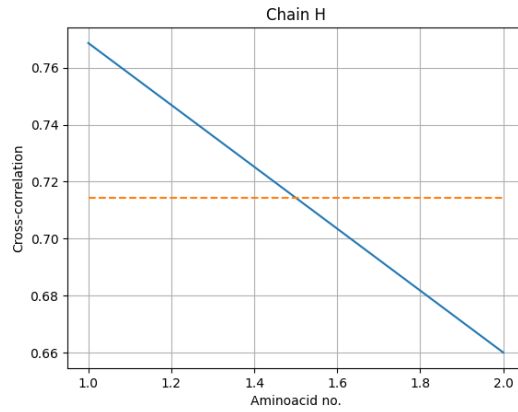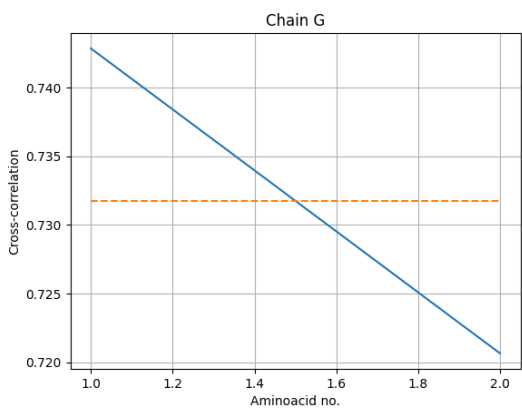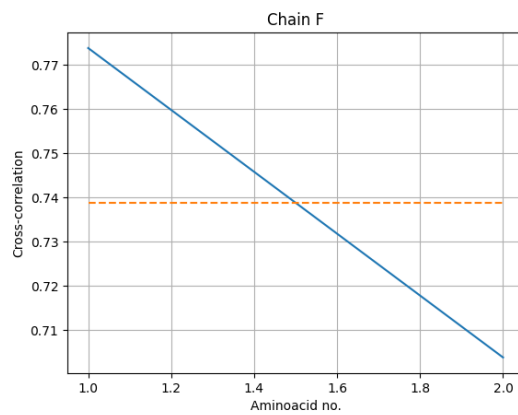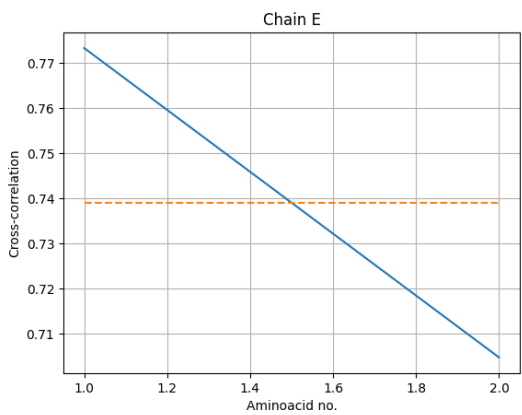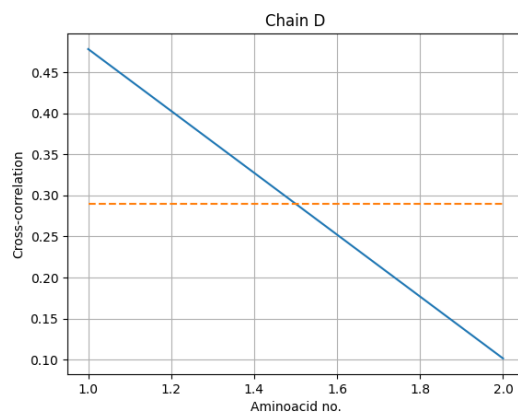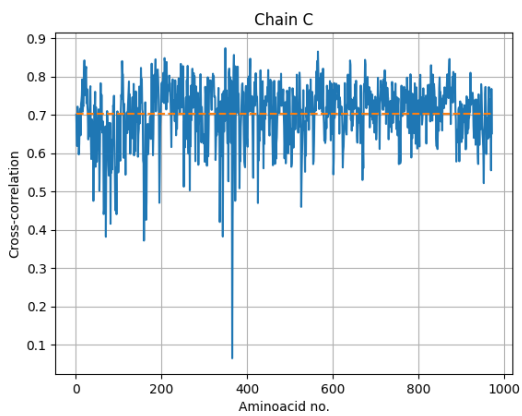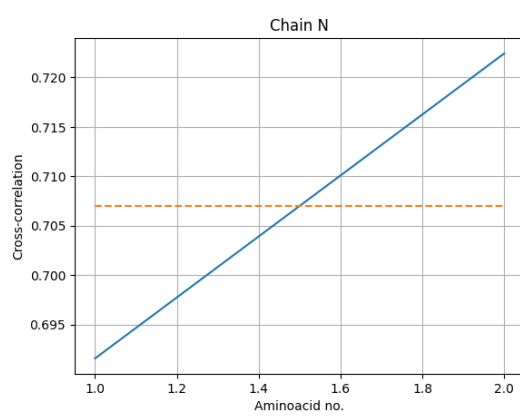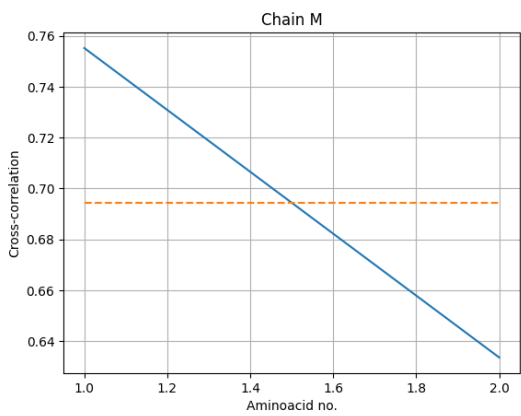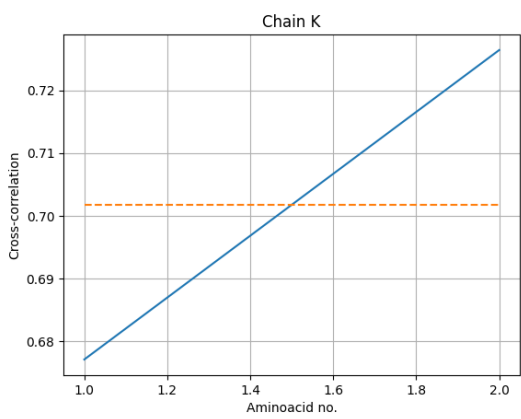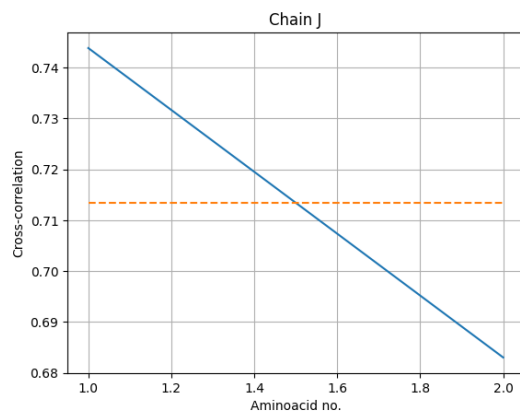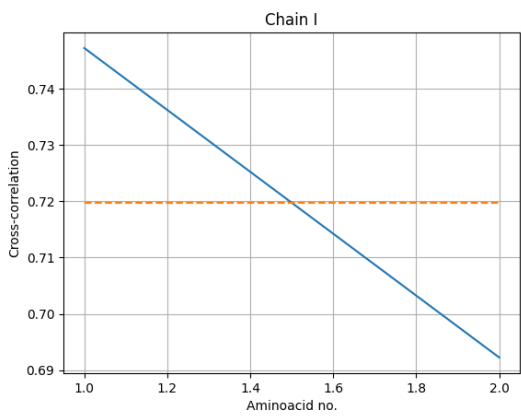
Chain O

Fig. 20 shows the histogram of all cross-correlations evaluated at the residues. The percentage of residues whose correlation is below 0.5 is 8.7 %.

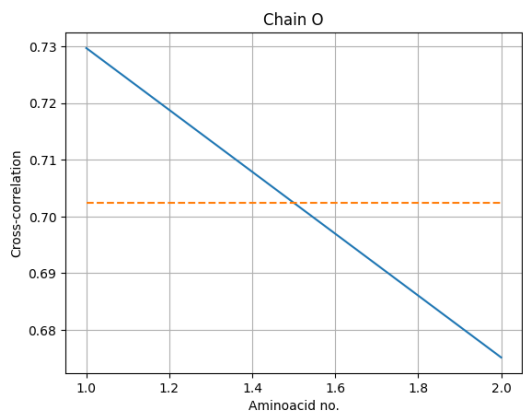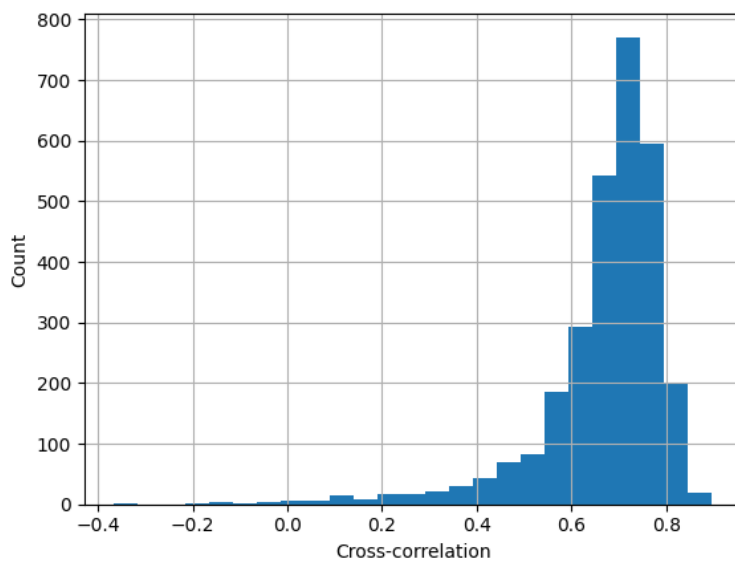Figure 20: Histogram of the cross-correlation between the map and model evaluated for all residues.

Resolutions estimated from the model:

Resolution (Å)	Masked	Unmasked
d99	4.8	4.7
d_model	3.9	3.9
d_model (B-factor=0)	8.2	8.2
FSC_model=0	3.7	3.8
FSC_model=0.143	3.8	3.8
FSC_model=0.5	4.2	4.4

Overall isotropic B factor:

B factor	Masked	Unmasked
Overall B-iso	235.0	255.0

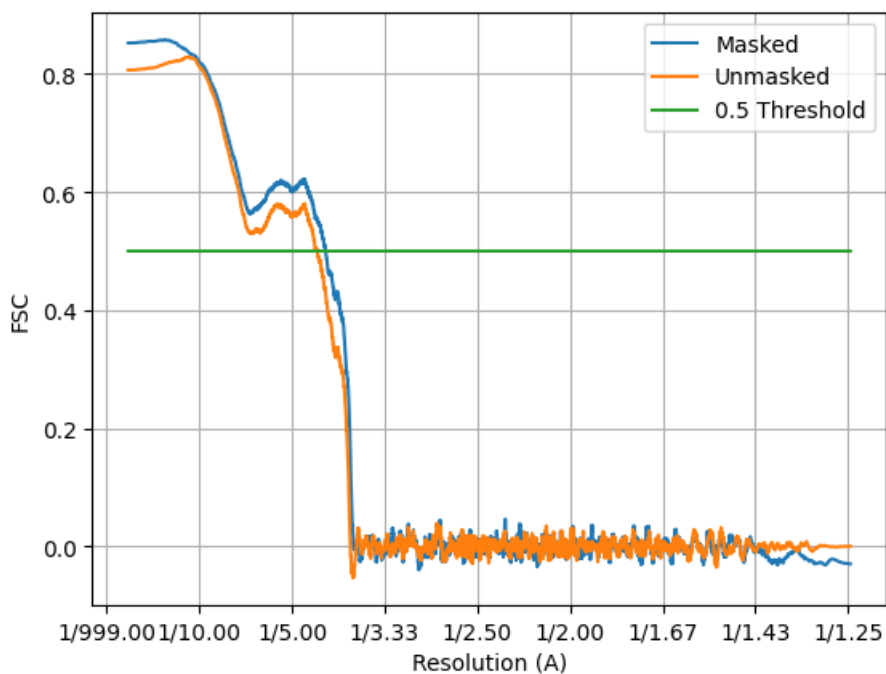Fig. 21 shows the FSC between the input map and the model.

Figure 21: FSC between the input map and model with and without a mask constructed from the model. The X-axis is the square of the inverse of the resolution in Å.

Automatic criteria: The validation is OK if 1) the percentage of residues whose correlation is smaller than 0.5 is smaller than 10%, and 2) the resolution reported by the user is larger than 0.8 times the resolution estimated between the map and model at FSC=0.5.

STATUS: OK

4.4   Level A.f EMRinger validation

Explanation:
EMringer [Barad et al., 2015] compares the side chains of the atomic model to the CryoEM map. The following features are reported:
- Optimal Threshold: Electron potential map cutoff value at which the maximum EMRinger score was obtained.
- Rotamer Ratio: Fraction of rotameric residues at the Optimal threshold value.
- Max Zscore: Z-score computed to determine the significance of the distribution at the Optimal threshold value.
- Model Length: Total of non-gamma-branched, non-proline aminoacids with a non-H gamma atom used in global EMRinger score computation.
- EMRinger Score: Maximum EMRinger score calculated at the Optimal Threshold.

A rotameric residue is one in which EMRinger peaks that fall within defined rotamers based on chi1, this often suggests a problem with the modelling of the backbone. In general, the user should look at the profiles and identify regions that may need improvement.

Results:

General results:

Optimal threshold	0.537408
Rotamer ratio	1.000
Max. Zscore	2.91
Model length	1514
EMRinger Score	0.748

Fig. 22 shows the EMRinger score and fraction of rotameric residues as

a function of the map threshold. The optimal threshold was selected looking for the maximum EMRinger score in this plot.
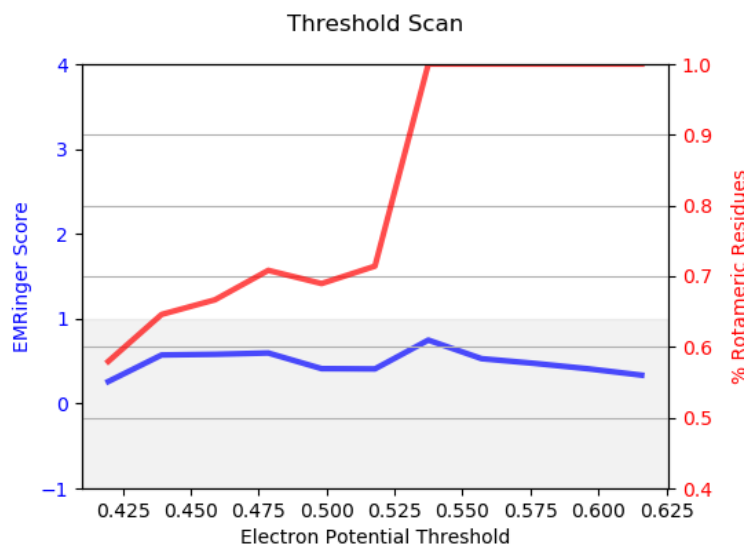

Figure 22: EMRinger score and fraction of rotameric residues as a function of the map threshold.

Fig. 23 shows the histogram for rotameric (blue) and non-rotameric (red) residues at the optimal threshold.
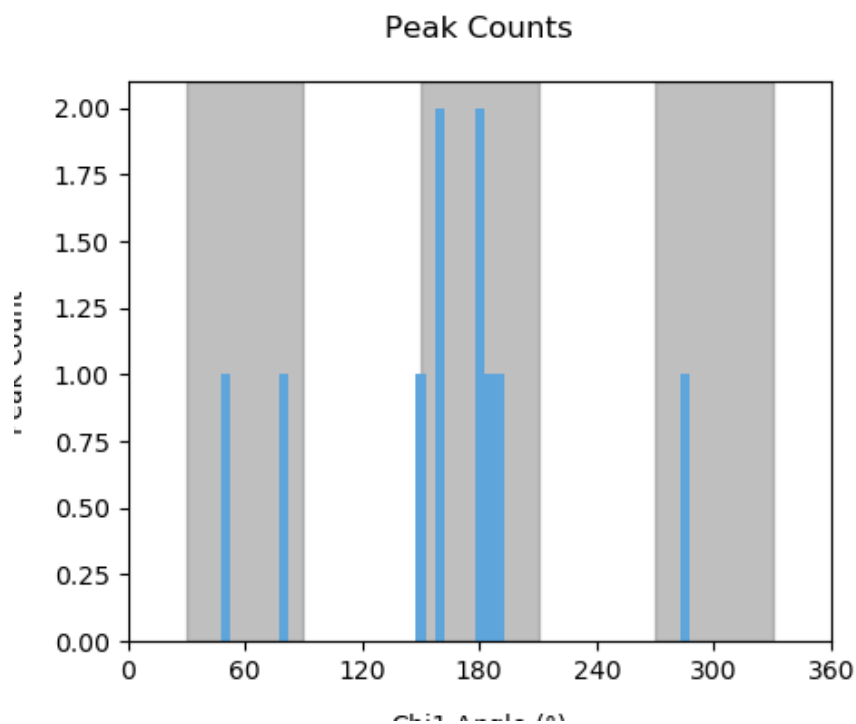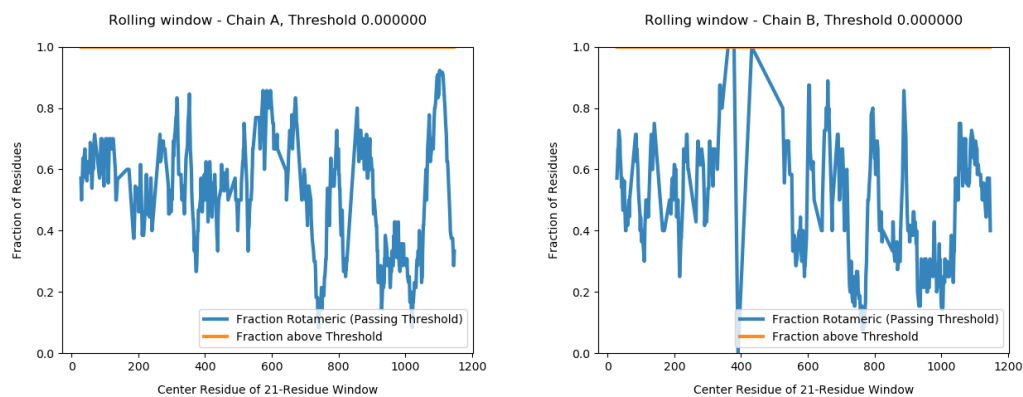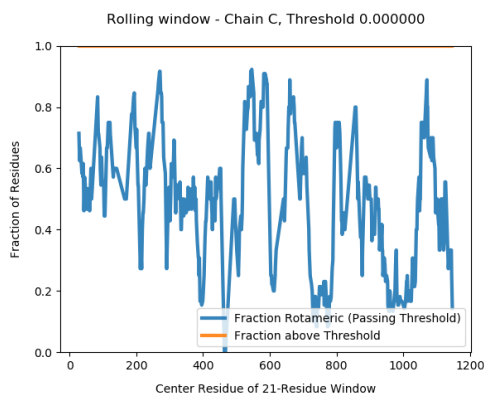
Figure 23: Histogram for rotameric (blue) and non-rotameric (red) residues at the optimal threshold as a function of the angle Chi1.

The following plots show the rolling window EMRinger analysis of the different chains to distinguish regions of improved model quality. This analysis was performed on rolling sliding 21-residue windows along the primary sequence of the protein chains.

Rolling window - Chain C, Threshold 0.000000

Automatic criteria: The validation is OK if the EMRinger score and Max. Zscore are larger than 1.

WARNINGS: 1 warnings

1. **The EMRinger score is smaller than 1, it is 0.748.**

4.5   Level A.g DAQ validation

Explanation:
DAQ [Terashi et al., 2022] is a computational tool using deep learning that can estimate the residue-wise local quality for protein models from cryo-Electron Microscopy maps. The method calculates the likelihood that a given density feature corresponds to an aminoacid, atom, and secondary structure. These likelihoods are combined into a score that ranges from -1 (bad quality) to 1 (good quality).

Results:
Fig. 24 shows the histogram of the DAQ values. The mean and standard deviation were 0.2 and 0.2, respectively.

Figure 24: Histogram of the DAQ values.

The atomic model colored by DAQ can be seen in Fig. 25.

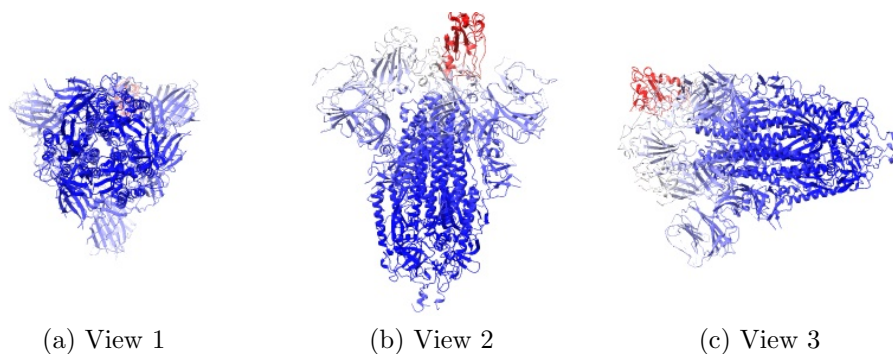(a) View 1　　　　　(b) View 2　　　　　(c) View 3

Figure 25: Atomic model colored by DAQ Views generated by ChimeraX at a the following X, Y, Z angles: View 1 (0,0,0), View 2 (90, 0, 0), View 3 (0, 90, 0).

Automatic criteria: The validation is OK if the average DAQ score is larger than 0.5.

WARNINGS: 1 warnings

1. **The average DAQ is smaller than 0.5.**

References

[Rosenthal and Henderson, 2003] Rosenthal, P. B. and Henderson, R. (2003). Optimal determination of particle orientation, absolute hand, and contrast loss in single particle electron-cryomicroscopy. *J. Molecular Biology*, 333:721–745.

[Ramírez-Aportela et al., 2019] Ramírez-Aportela, E., Mota, J., Conesa, P., Carazo, J. M., and Sorzano, C. O. S. (2019). DeepRes: a new deep-learning- and aspect-based local resolution method for electron-microscopy maps. *IUCRj*, 6:1054–1063.

[Kaur et al., 2021] Kaur, S., Gomez-Blanco, J., Khalifa, A. A., Adinarayanan, S., Sanchez-Garcia, R., Wrapp, D., McLellan, J. S., Bui, K. H., and Vargas, J. (2021). Local computational methods to improve the interpretability and analysis of cryo-EM maps. *Nature Communications*, 12(1):1–12.

[Pintilie et al., 2020] Pintilie, G., Zhang, K., Su, Z., Li, S., Schmid, M. F., and Chiu, W. (2020). Measurement of atom resolvability in cryo-em maps with q-scores. *Nature methods*, 17(3):328–334.

[Afonine et al., 2018] Afonine, P. V., Klaholz, B. P., Moriarty, N. W., Poon, B. K., Sobolev, O. V., Terwilliger, T. C., Adams, P. D., and Urzhumtsev, A. (2018). New tools for the analysis and validation of cryo-EM maps and atomic models. *Acta Crystallographica D, Struct. Biol.*, 74:814–840.

[Barad et al., 2015] Barad, B. A., Echols, N., Wang, R. Y.-R., Cheng, Y., DiMaio, F., Adams, P. D., and Fraser, J. S. (2015). EMRinger: side chain-directed model and map validation for 3D cryo-electron microscopy. *Nature Methods*, 12(10):943–946.

[Terashi et al., 2022] Terashi, G., Wang, X., Subramaniya, S.R.M.V., Tesmer, J.J.G. and Kihara, D. (2022). Residue-Wise Local Quality Estimation for Protein Models from Cryo-EM Maps. (submitted).

Validation report of Level(s) 0, A

I^{2}PC Validation server

March 2, 2022
6:09am

Abstract

The map seems to be well centered. There seems to be a problem with the suggested threshold (see Sec. 2.2). There seems to be a problem with the map's background (see Sec. 2.3). There seems to be a problem with its B-factor (see Sec. 2.4). There seems to be a problem with its MapQ scores (see Sec. 4.1). It seems that the Guinier plot of the map and its model do not match (see Sec. 4.2). According to phenix, it seems that there might be some mismatch between the map and its model (see Sec. 4.3). DAQ detects some mismatch between the map and its model (see Sec. 4.5).

The average resolution of the map estimated by various methods goes from 1.2Å to 6.6Å with an average of 3.9Å. The resolution provided by the user was 1.1Å. The resolution reported by the user may be overestimated.

The overall score (passing tests) of this report is 3 out of 10 evaluable items.

0.a Mass analysis	Sec. 2.1	OK
0.b Mask analysis	Sec. 2.2	2 warnings
0.c Background analysis	Sec. 2.3	2 warnings
0.d B-factor analysis	Sec. 2.4	1 warnings
0.e DeepRes	Sec. 2.5	Does not apply
0.f LocBfactor	Sec. 2.6	Could not be measured
0.g LocOccupancy	Sec. 2.7	Could not be measured
0.h DeepHand	Sec. 2.8	OK
A.a MapQ	Sec. 4.1	1 warnings
A.d Map-Model Guinier	Sec. 4.2	1 warnings
A.e Phenix validation	Sec. 4.3	1 warnings
A.f EMRinger	Sec. 4.4	OK
A.g DAQ	Sec. 4.5	1 warnings

Summary of the warnings across sections.

If it is empty below this point, it means that there are no warnings.

Section 2.2 (0.b Mask analysis)

1. **There might be a problem of connectivity at this threshold because more than 5 connected components are needed to reach 95% of the total mask.**
2. **There might be a problem in the construction of the mask, because the overlap is smaller than 0.75. A common reason is that the suggested threshold causes too many disconnected components.**

Section 2.3 (0.c Background analysis)

1. **The null hypothesis that the background mean is 0 has been rejected because the p-value of the comparison is smaller than 0.001**
2. **There is a significant proportion of outlier values in the background (cdf5 ratio=1081.35)**

Section 2.4 (0.d B-factor analysis)

1. **The B-factor is out of the interval [-300,0]**

Section 4.1 (A.a MapQ)

1. **The median Q-score is less than 0.1.**

Section 4.2 (A.d Map-Model Guinier)

1. **The correlation is smaller than 0.5, it is 0.497.**

Section 4.3 (A.e Phenix validation)

1. **The percentage of residues that have a cross-correlation below 0.5 is 100.0, that is larger than 10%**

Section 4.5 (A.g DAQ)

1. **The average DAQ is smaller than 0.5.**

Contents

1 Input data

Input map: /home/coss/data/Dropbox/Aplicaciones/ShareLaTeX/MapValidation/-EMDB11668/emd_11668.map
SHA256 hash: 69a72c5b39bb0573f60a4289b4e17063ebd26cee331f5f018d153aa06f184813
Voxel size: 0.492000 (Å)
Visualization threshold: 0.050000
Resolution estimated by user: 1.150000

Orthogonal slices of the input map
Explanation:
In the orthogonal slices of the map, the noise outside the protein should not have any structure (stripes going out, small blobs, particularly high or low densities, ...)
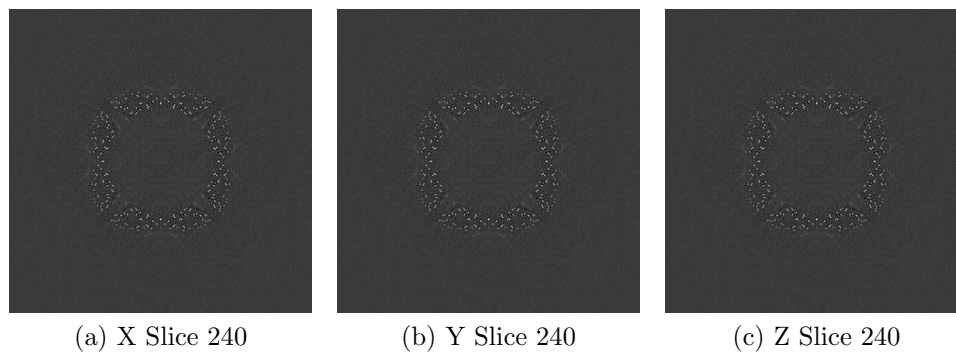
Results:
See Fig. 1.

(a) X Slice 240	(b) Y Slice 240	(c) Z Slice 240

Figure 1: Central slices of the input map in the three dimensions

Orthogonal slices of maximum variance of the input map
Results:
See Fig. 2.

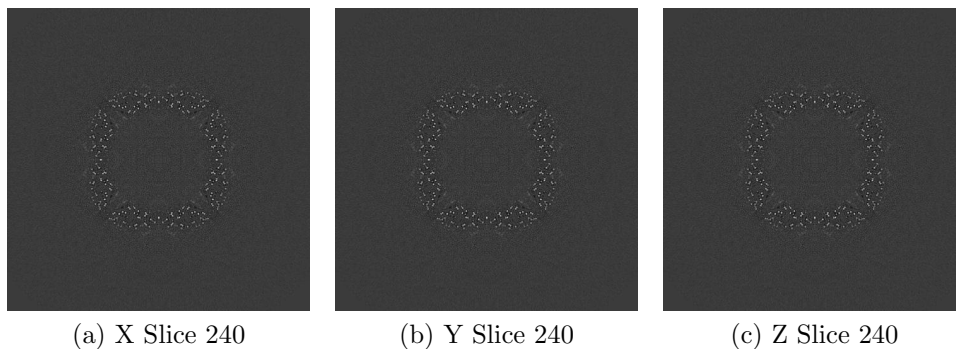(a) X Slice 240　　　　(b) Y Slice 240　　　　(c) Z Slice 240

Figure 2: Slices of maximum variation in the three dimensions

Orthogonal projections of the input map

Explanation:
In the projections there should not be stripes (this is an indication of directional overweighting, or angular attraction), and there should not be a dark halo around or inside the structure (this is an indication of incorrect CTF correction or the reconstruction of a biased map).
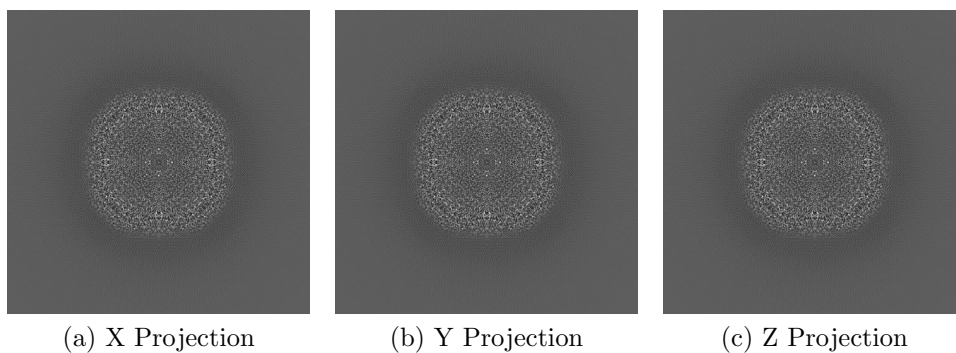
Results:
See Fig. 3.

(a) X Projection　　　　(b) Y Projection　　　　(c) Z Projection

Figure 3: Projections in the three dimensions

Isosurface views of the input map

Explanation:

An isosurface is the surface of all points that have the same gray value. In these views there should not be many artifacts or noise blobs around the map.

Results:
See Fig. 4.

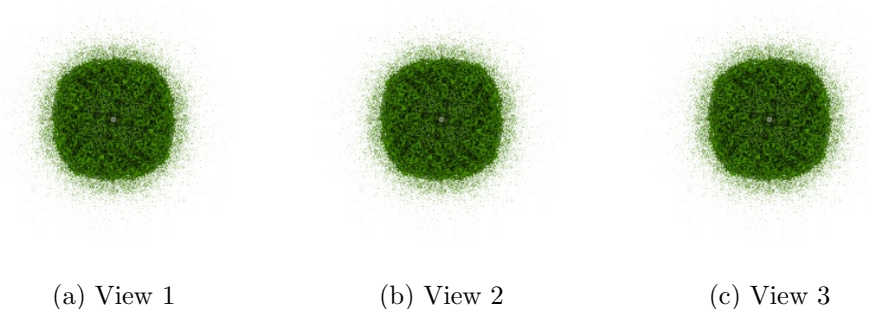(a) View 1   (b) View 2   (c) View 3

Figure 4: Isosurface at threshold=0.050000. Views generated by ChimeraX at a the following X, Y, Z angles: View 1 (0,0,0), View 2 (90, 0, 0), View 3 (0, 90, 0).

Orthogonal slices of maximum variance of the mask
Explanation:
The mask has been calculated at the suggested threshold 0.050000, the largest connected component was selected, and then dilated by 2Å.
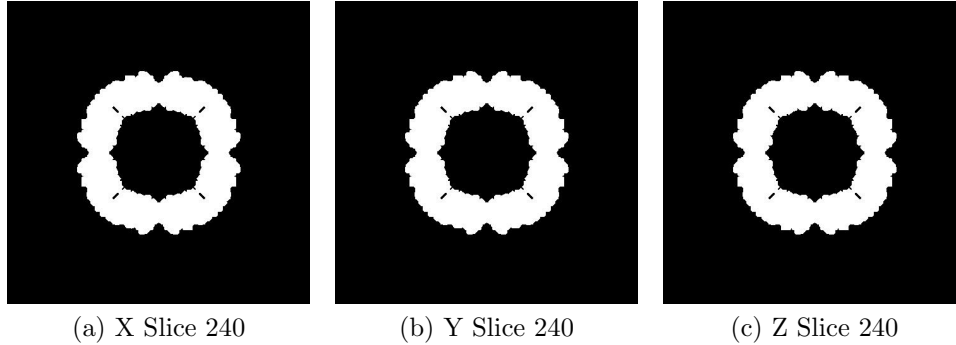
Results:
See Fig. 5.

| (a) X Slice 240 | (b) Y Slice 240 | (c) Z Slice 240 |

Figure 5: Slices of maximum variation in the three dimensions of the mask

2 Level 0 analysis

2.1 Level 0.a Mass analysis

Explanation:
The reconstructed map must be relatively well centered in the box, and there should be at least 30Å (the exact size depends on the CTF) on each side to make sure that the CTF can be appropriately corrected.

Results:
The space from the left and right in X are 54.61 and 54.61 Å, respectively. There is a decentering ratio (abs(Right-Left)/Size)% of 0.00%

The space from the left and right in Y are 54.61 and 54.61 Å, respectively. There is a decentering ratio (abs(Right-Left)/Size)% of 0.00%

The space from the left and right in Z are 54.61 and 54.61 Å, respectively. There is a decentering ratio (abs(Right-Left)/Size)% of 0.00%

The center of mass is at (x,y,z)=(240.51,240.51,240.51). The decentering of the center of mass (abs(Center)/Size)% is 0.11, 0.11, and 0.11, respectively.%

Automatic criteria: The validation is OK if 1) the decentering and

8

center of mass less than 20% of the map dimensions in all directions, and 2) the extra space on each direction is more than 20% of the map dimensions.

STATUS: OK

2.2 Level 0.b Mask analysis

Explanation:
The map at the suggested threshold should have most of its mass concentrated in a single connected component. It is normal that after thresholding there are a few thousands of very small, disconnected noise blobs. However, there total mass should not exceed 10%. The raw mask (just thresholding) and the mask constructed for the analysis (thresholding + largest connected component + dilation) should significantly overlap. Overlap is defined by the overlapping coefficient (size(Raw AND Constructed)/size(Raw)) that is a number between 0 and 1, the closer to 1, the more they agree.

Results:

Raw mask: At threshold 0.050000, there are 478289 connected components with a total number of voxels of 1531522 and a volume of 182397.36 Å^3 (see Fig. 6). The size and percentage of the total number of voxels for the raw mask are listed below (up to 95% of the mass or the first 100 clusters, whatever happens first), the list contains (No. voxels (volume in Å^3), percentage, cumulatedPercentage):

, (650384 (77457.80), 42.47, 42.47), (106 (12.62), 0.01, 42.47), (106 (12.62), 0.01, 42.48), (106 (12.62), 0.01, 42.49), (106 (12.62), 0.01, 42.49), (106 (12.62), 0.01, 42.50), (106 (12.62), 0.01, 42.51), (106 (12.62), 0.01, 42.51), (106 (12.62), 0.01, 42.52), (106 (12.62), 0.01, 42.53), (106 (12.62), 0.01, 42.54), (106 (12.62), 0.01, 42.54), (106 (12.62), 0.01, 42.55), (85 (10.12), 0.01, 42.56), (85 (10.12), 0.01, 42.56), (85 (10.12), 0.01, 42.57), (85 (10.12), 0.01, 42.57), (85 (10.12), 0.01, 42.58), (85 (10.12), 0.01, 42.58), (85 (10.12), 0.01, 42.59), (85 (10.12), 0.01, 42.59), (68 (8.10), 0.00, 42.60), (68 (8.10), 0.00, 42.60), (68 (8.10), 0.00, 42.61), (68 (8.10), 0.00, 42.61), (68 (8.10), 0.00, 42.62), (68 (8.10), 0.00, 42.62), (68 (8.10), 0.00, 42.63), (68 (8.10), 0.00, 42.63), (68 (8.10), 0.00, 42.63), (68 (8.10), 0.00, 42.64), (68 (8.10),

0.00, 42.64), (68 (8.10), 0.00, 42.65), (68 (8.10), 0.00, 42.65), (68 (8.10),
0.00, 42.66), (68 (8.10), 0.00, 42.66), (68 (8.10), 0.00, 42.67), (68 (8.10),
0.00, 42.67), (68 (8.10), 0.00, 42.67), (68 (8.10), 0.00, 42.68), (68 (8.10),
0.00, 42.68), (68 (8.10), 0.00, 42.69), (68 (8.10), 0.00, 42.69), (68 (8.10),
0.00, 42.70), (68 (8.10), 0.00, 42.70), (65 (7.74), 0.00, 42.70), (65 (7.74),
0.00, 42.71), (65 (7.74), 0.00, 42.71), (65 (7.74), 0.00, 42.72), (65 (7.74),
0.00, 42.72), (65 (7.74), 0.00, 42.73), (65 (7.74), 0.00, 42.73), (65 (7.74),
0.00, 42.73), (65 (7.74), 0.00, 42.74), (65 (7.74), 0.00, 42.74), (65 (7.74),
0.00, 42.75), (65 (7.74), 0.00, 42.75), (65 (7.74), 0.00, 42.76), (65 (7.74),
0.00, 42.76), (65 (7.74), 0.00, 42.76), (65 (7.74), 0.00, 42.77), (65 (7.74),
0.00, 42.77), (65 (7.74), 0.00, 42.78), (65 (7.74), 0.00, 42.78), (65 (7.74),
0.00, 42.79), (65 (7.74), 0.00, 42.79), (65 (7.74), 0.00, 42.79), (65 (7.74),
0.00, 42.80), (65 (7.74), 0.00, 42.80), (54 (6.43), 0.00, 42.81), (54 (6.43),
0.00, 42.81), (54 (6.43), 0.00, 42.81), (54 (6.43), 0.00, 42.82), (54 (6.43),
0.00, 42.82), (54 (6.43), 0.00, 42.82), (54 (6.43), 0.00, 42.83), (54 (6.43),
0.00, 42.83), (54 (6.43), 0.00, 42.83), (54 (6.43), 0.00, 42.84), (54 (6.43),
0.00, 42.84), (54 (6.43), 0.00, 42.84), (54 (6.43), 0.00, 42.85), (54 (6.43),
0.00, 42.85), (54 (6.43), 0.00, 42.86), (54 (6.43), 0.00, 42.86), (54 (6.43),
0.00, 42.86), (54 (6.43), 0.00, 42.87), (54 (6.43), 0.00, 42.87), (54 (6.43),
0.00, 42.87), (54 (6.43), 0.00, 42.88), (54 (6.43), 0.00, 42.88), (54 (6.43),
0.00, 42.88), (54 (6.43), 0.00, 42.89), (53 (6.31), 0.00, 42.89), (53 (6.31),
0.00, 42.89), (53 (6.31), 0.00, 42.90), (53 (6.31), 0.00, 42.90), (53 (6.31),
0.00, 42.90), (53 (6.31), 0.00, 42.91)

Number of components to reach 95% of the mass: 401714

The average size of the remaining 76575 components is 1.00 voxels (0.12
Å^3). Their size go from 1 voxels (0.12 Å^3) to 1 voxels (0.12 Å^3).

The slices of the raw mask can be seen in Fig. 6.

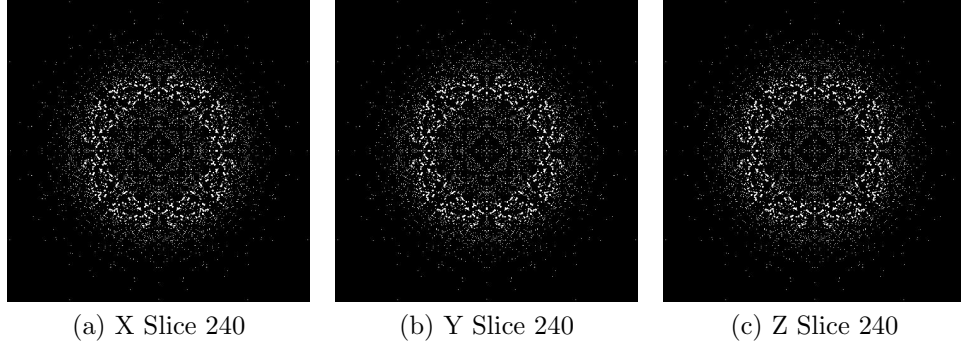| (a) X Slice 240 | (b) Y Slice 240 | (c) Z Slice 240 |

Figure 6: Maximum variance slices in the three dimensions of the raw mask

The following table shows the variation of the mass enclosed at different thresholds (see Fig. 7):

Threshold	Voxel mass	Molecular mass(kDa)	# Aminoacids
0.0422	2333120.00	230.21	2092.82
0.0845	536788.00	52.97	481.50
0.1267	352884.00	34.82	316.54
0.1690	266092.00	26.26	238.69
0.2112	202304.00	19.96	181.47
0.2535	155937.00	15.39	139.88
0.2957	117288.00	11.57	105.21
0.3380	86874.00	8.57	77.93
0.3802	63930.00	6.31	57.35
0.4224	45810.00	4.52	41.09
0.4647	32418.00	3.20	29.08
0.5069	21234.00	2.10	19.05
0.5492	13692.00	1.35	12.28
0.5914	8316.00	0.82	7.46
0.6337	5253.00	0.52	4.71
0.6759	2748.00	0.27	2.46
0.7181	1308.00	0.13	1.17
0.7604	630.00	0.06	0.57
0.8026	246.00	0.02	0.22
0.8449	102.00	0.01	0.09
0.8871	54.00	0.01	0.05
0.9294	54.00	0.01	0.05
0.9716	24.00	0.00	0.02
1.0139	24.00	0.00	0.02
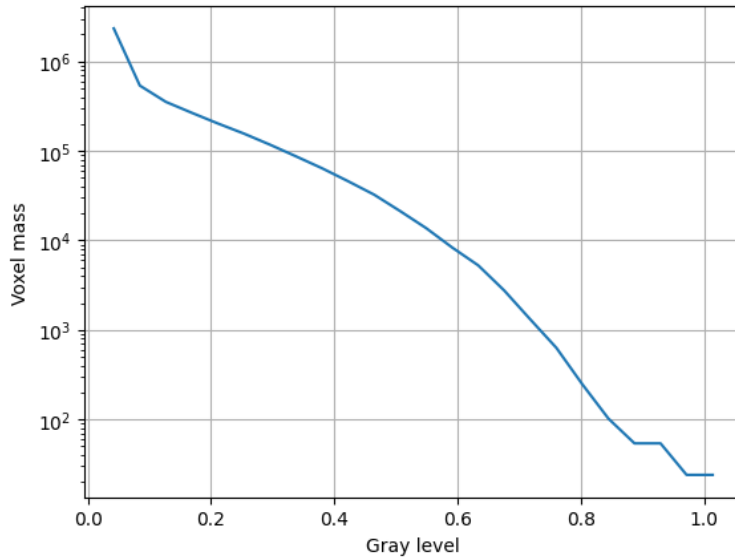
Figure 7: Voxel mass as a function of the gray level.

Constructed mask: After keeping the largest component of the previous mask and dilating it by 2Å, there is a total number of voxels of 7330938 and a volume of 873081.64 Å^3. The overlap between the raw and constructed mask is 0.60.

Automatic criteria: The validation is OK if 1) to keep 95% of the mass we need to keep at most 5 connected components; and 2) the average volume of the blobs outside the given threshold has a size smaller than 5Å^3; and 3) the overlap between the raw mask and the mask constructed for the analysis is larger than 75%.

WARNINGS: 2 warnings

1. **There might be a problem of connectivity at this threshold because more than 5 connected components are needed to reach 95% of the total mask.**
2. **There might be a problem in the construction of the mask, because the overlap is smaller than 0.75. A common reason is that the suggested threshold causes too many disconnected components.**

2.3 Level 0.c Background analysis

Explanation:
Background is defined as the region outside the macromolecule mask. The background mean should be zero, and the number of voxels with a very low or very high value (below 5 standard deviations of the noise) should be very small and they should be randomly distributed without any specific structure. Sometimes, you can see some structure due to the symmetry of the structure.

Results:

The null hypothesis that the background mean is 0 was tested with a one-sample Student's t-test. The resulting t-statistic and p-value were -203.22 and 0.000000, respectively.

The mean and standard deviation (sigma) of the background were -0.000338 and 0.016924. The percentage of background voxels whose absolute value is larger than 5 times the standard deviation is 0.06 % (see Fig. 8). The same percentage from a Gaussian would be 0.000057% (ratio between the two percentages: 1081.353334).

Slices of the background beyond 5*sigma can be seen in Fig. 8.
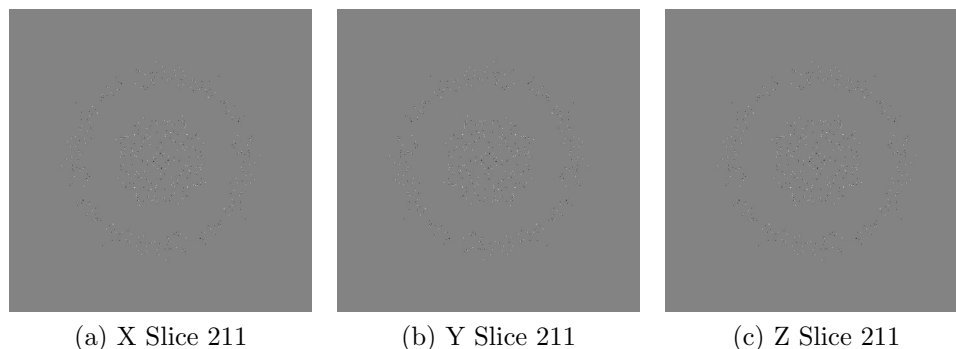
(a) X Slice 211	(b) Y Slice 211	(c) Z Slice 211

Figure 8: Maximum variance slices in the three dimensions of the parts of the background beyond 5*sigma

Automatic criteria: The validation is OK if 1) the p-value of the null

hypothesis that the background has 0 mean is larger than 0.001; and 2) the number of voxels above or below 5 sigma is smaller than 20 times the amount expected for a Gaussian with the same standard deviation whose mean is 0.

WARNINGS: 2 warnings

1. **The null hypothesis that the background mean is 0 has been rejected because the p-value of the comparison is smaller than 0.001**
2. **There is a significant proportion of outlier values in the background (cdf5 ratio=1081.35)**

2.4 Level 0.d B-factor analysis

Explanation:
The B-factor line [Rosenthal and Henderson, 2003] fitted between 15Åand the resolution reported should have a slope that is between 0 and 300 Å^2.

Results:
Fig. 9 shows the logarithm (in natural units) of the structure factor (the module squared of the Fourier transform) of the experimental map, its fitted line, and the corrected map. The estimated B-factor was 0.1. The fitted line was $\log(|F|^2) = 0.0/R^2 + (-12.6)$.
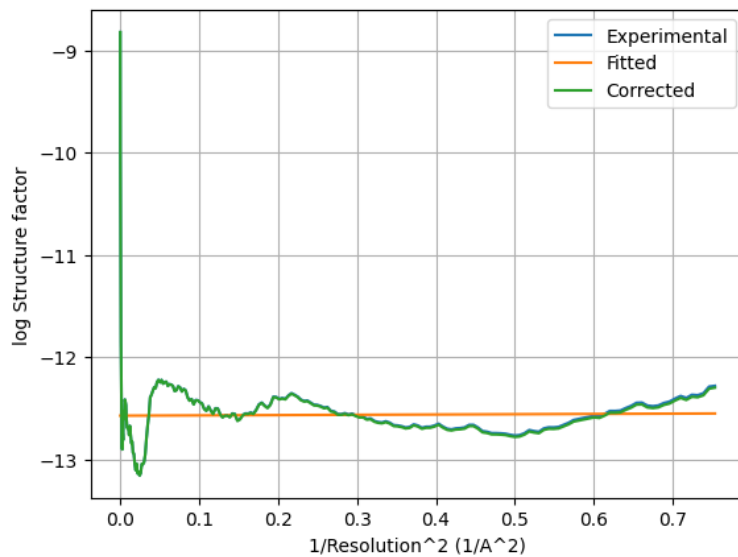
Figure 9: Guinier plot. The X-axis is the square of the inverse of the resolution in Å.
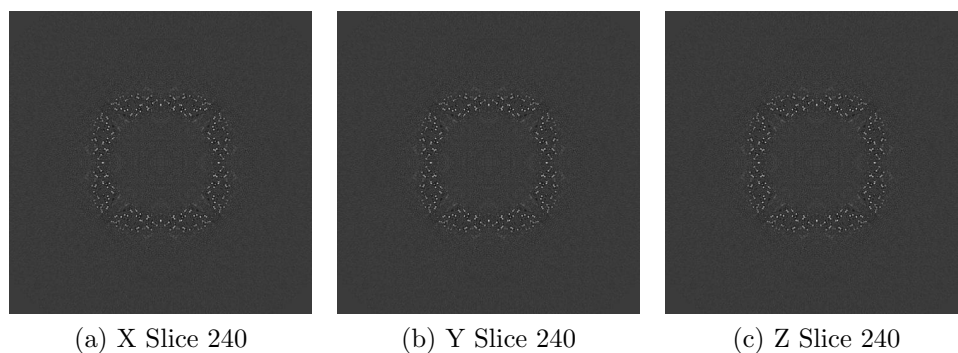
(a) X Slice 240　　　　　(b) Y Slice 240　　　　　(c) Z Slice 240

Figure 10: Slices of maximum variation in the three dimensions of the B-factor corrected map

Automatic criteria: The validation is OK if the B-factor is in the range [-300,0].

WARNINGS: 1 warnings

1. **The B-factor is out of the interval [-300,0]**

2.5 Level 0.e Local resolution with DeepRes

Explanation:
DeepRes [Ramírez-Aportela et al., 2019] measures the local resolution using a neural network that has been trained on the appearance of atomic structures at different resolutions. Then, by comparing the local appearance of the input map to the appearance of the atomic structures a local resolution label can be assigned.

Results:

This method cannot be applied to maps with a resolution better than 2Å.

2.6 Level 0.f Local B-factor

Explanation:
LocBfactor [Kaur et al., 2021] estimates a local resolution B-factor by decomposing the input map into a local magnitude and phase term using the spiral transform.

Results:

ERROR: The protocol failed.

2.7 Level 0.g Local Occupancy

Explanation:
LocOccupancy [Kaur et al., 2021] estimates the occupancy of a voxel by the macromolecule.

Results:

ERROR: The protocol failed.

2.8   Level 0.h Hand correction

Explanation:
Deep Hand determines the correction of the hand for those maps with a resolution smaller than 5Å. The method calculates a value between 0 (correct hand) and 1 (incorrect hand) using a neural network to assign its hand.

Results:

Deep hand assigns a score of 0.327 to the input volume.
Automatic criteria: The validation is OK if the deep hand score is smaller than 0.5.

STATUS: OK

3   Atomic model

Atomic model: /home/coss/data/Dropbox/Aplicaciones/ShareLaTeX/MapValidation/-EMDB11668/7a6a_updated.cif

See Fig. 11.

(a) View 1   (b) View 2   (c) View 3
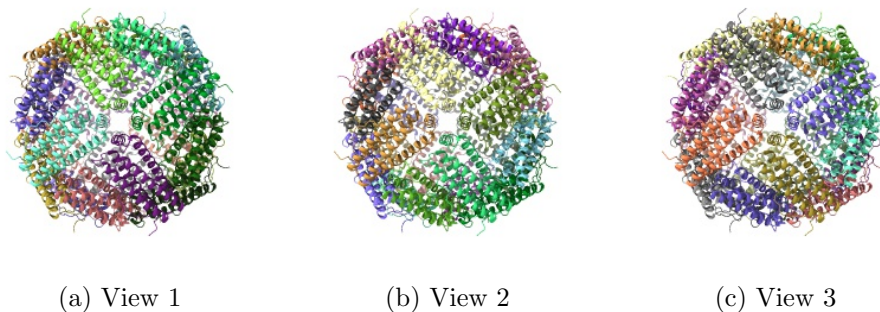
Figure 11: Input atomic model Views generated by ChimeraX at a the following X, Y, Z angles: View 1 (0,0,0), View 2 (90, 0, 0), View 3 (0, 90, 0).

4 Level A analysis

4.1 Level A.a MapQ

Explanation:
MapQ [Pintilie et al., 2020] computes the local correlation between the map and each one of its atoms assumed to have a Gaussian shape.

Results:

Fig. 12 shows the histogram of the Q-score according calculated by MapQ. Some representative percentiles are:

Percentile	MapQ score [0-1]
2.5%	-0.23
25%	0.00
50%	0.00
75%	0.00
97.5%	0.94
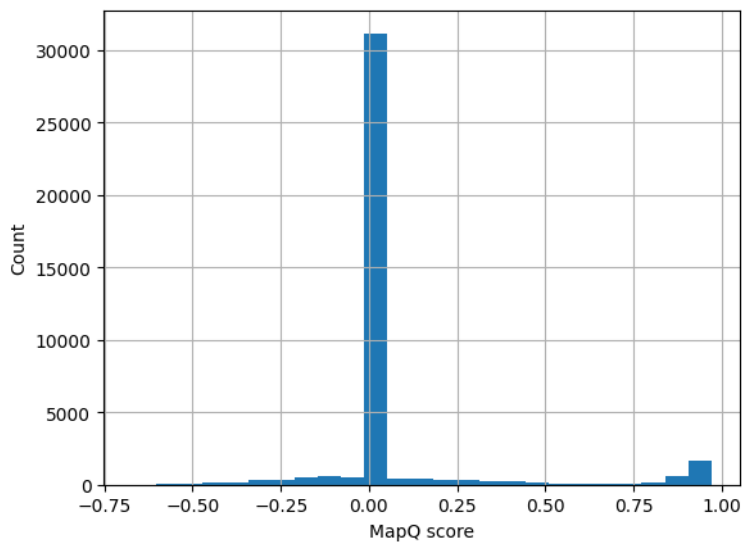
Figure 12: Histogram of the Q-score.

The following table shows the average Q score and estimated resolution for each chain.

Chain	Average Q score [0-1]	Estimated Resol. (Å)
1	0.00	11.2
1	0.00	9.8
1	0.13	5.6
1	0.00	0.0
2	0.01	11.1
2	0.05	9.4
2	0.12	5.6
2	-0.09	0.0
4	0.00	11.1
4	0.00	9.8
4	0.13	5.6
4	0.05	0.0
6	0.00	11.2
6	0.00	9.8
6	0.13	5.6
6	0.00	0.0
A	0.00	11.2
A	0.00	9.8
A	0.13	5.6
A	0.00	0.0
B	0.00	11.2
B	0.00	9.8
B	0.13	5.6
B	0.00	0.0
E	0.00	11.2
E	0.00	9.8
E	0.13	5.6
E	0.00	0.0
F	0.00	11.2
F	0.00	9.8
F	0.13	5.6
F	0.00	0.0
G	-0.00	11.2
G	0.00	9.8
G	0.13	5.6
G	0.01	0.0
H	0.00	11.2
H	0.00	9.8
H	0.13	5.6
H	0.00	0.0
I	-0.00	11.2
I	-0.05	10.3
I	0.13	5.6
I	-0.11	0.0
K	-0.00	11.2

Automatic criteria: The validation is OK if the median Q-score is larger than 0.1.

WARNINGS: 1 warnings

1. **The median Q-score is less than 0.1.**

4.2   Level A.d Map-Model Guinier analysis

Explanation:
We compared the Guinier plot [Rosenthal and Henderson, 2003] of the atomic model and the experimental map. We made the mean of both profiles to be equal (and equal to the mean of the atomic model) to make sure that they had comparable scales.

Results:
Fig. 13 shows the logarithm (in natural units) of the structure factor (the module squared of the Fourier transform) of the atom model and the experimental map. The correlation between the two profiles was 0.497.
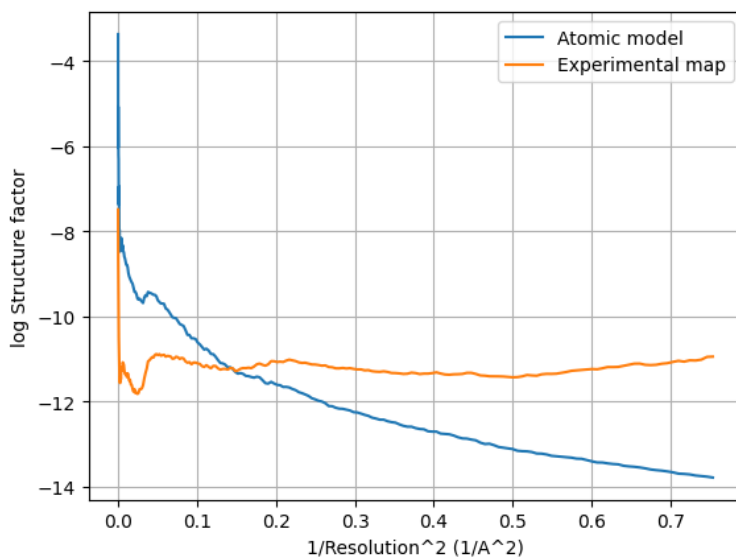
Figure 13: Guinier plot of the atom model and experimental map. The X-axis is the square of the inverse of the resolution in Å.

Automatic criteria: The validation is OK if the correlation between the two Guinier profiles is larger than 0.5.

WARNINGS: 1 warnings

1. **The correlation is smaller than 0.5, it is 0.497.**

4.3   Level A.e Phenix validation

Explanation:
Phenix provides a number of tools to assess the agreement between the experimental map and its atomic model [Afonine et al., 2018]. There are several cross-correlations to assess the quality of the fitting:

- CC (mask): Model map vs. experimental map correlation coefficient calculated considering map values inside a mask calculated around the macromolecule.
- CC (box): Model map vs. experimental map correlation coefficient calculated considering all grid points of the box.
- CC (volume) and CC (peaks) compare only map regions with the highest density values and regions below a certain contouring threshold level are ignored. CC (volume): The map region considered is defined by the N highest points inside the molecular mask. CC (peaks): In this case, calculations consider the union of regions defined by the N highest peaks in the model-calculated map and the N highest peaks in the experimental map.
- Local real-space correlation coefficients CC (main chain) and CC (side chain) involve the main skeleton chain and side chains, respectively.

There are also multiple ways of measuring the resolution:
- d99: Resolution cutoff beyond which Fourier map coefficients are negligibly small. Calculated from the full map.
- d_model: Resolution cutoff at which the model map is the most similar to the target (experimental) map. For d_model to be meaningful, the model is expected to fit the map as well as possible. d_model (B factors = 0) tries to avoid the blurring of the map.
- d_FSC_model; Resolution cutoff up to which the model and map Fourier coefficients are similar at FSC values of 0, 0.143, 0.5.

In addition to these resolution measurements the overall isotropic B factor is another indirect measure of the quality of the map.

Results:

To avoid ringing in Fourier space a smooth mask with a radius of 6.0 Å has been applied.

Overall correlation coefficients:

$$
\begin{aligned}
\text{CC (mask)} &= 0.303 \\
\text{CC (box)} &= 0.236 \\
\text{CC (volume)} &= 0.292 \\
\text{CC (peaks)} &= 0.124 \\
\text{CC (main chain)} &= 0.301 \\
\text{CC (side chain)} &= 0.291
\end{aligned}
$$

Correlation coefficients per chain:

Chain	Cross-correlation
A	0.289322
1	0.289309
K	0.289362
a	0.289360
B	0.289327
E	0.289332
e	0.289314
r	0.289360
G	0.289330
I	0.289307
M	0.289322
O	0.289368
Q	0.289324
S	0.289330
U	0.289175
W	0.289317
Y	0.289379
2	0.289310
4	0.289200
F	0.289296
H	0.289320
P	0.289311
X	0.289314
6	0.289361

We now show the correlation profiles of the different chain per residue.

24
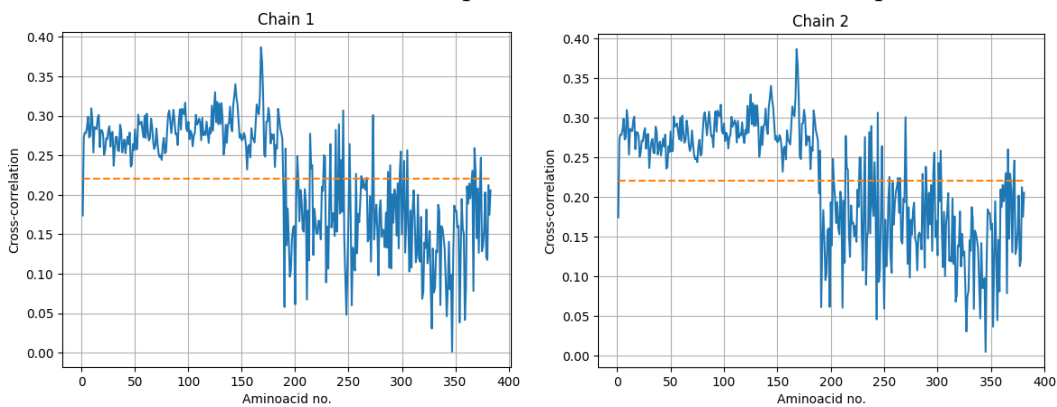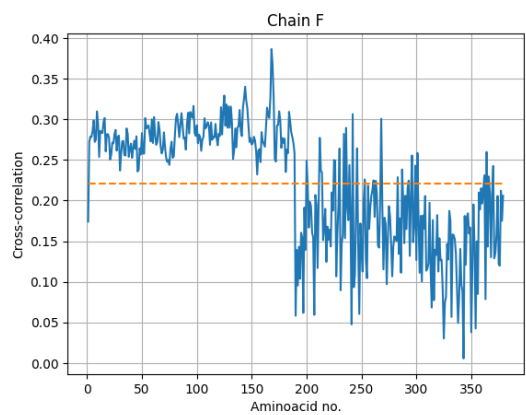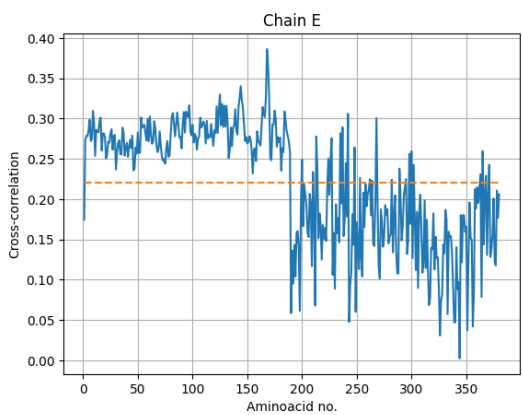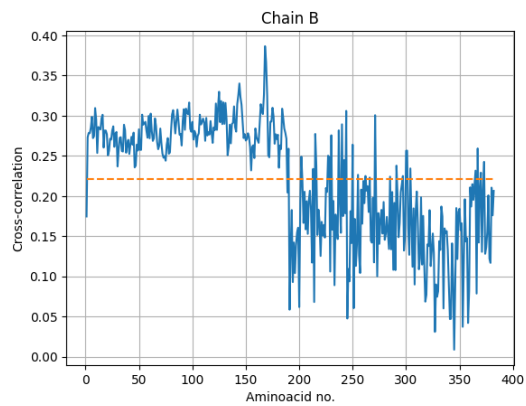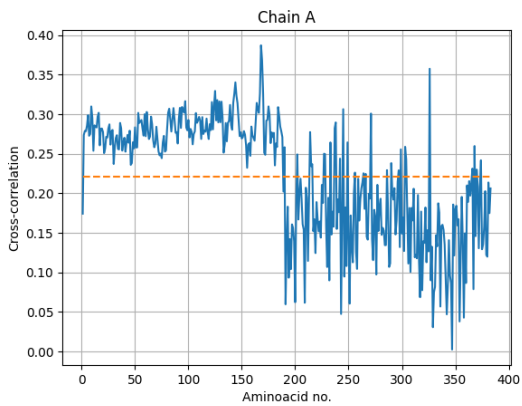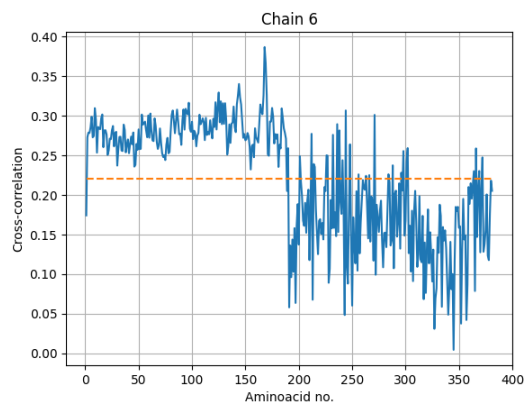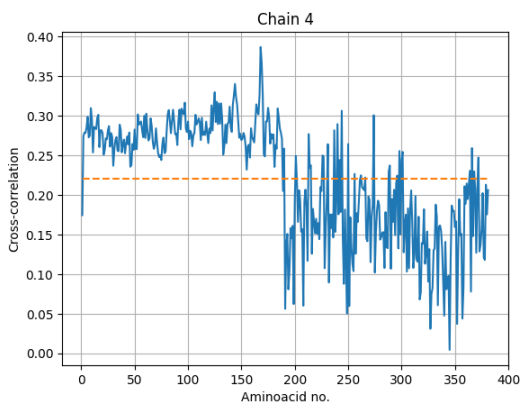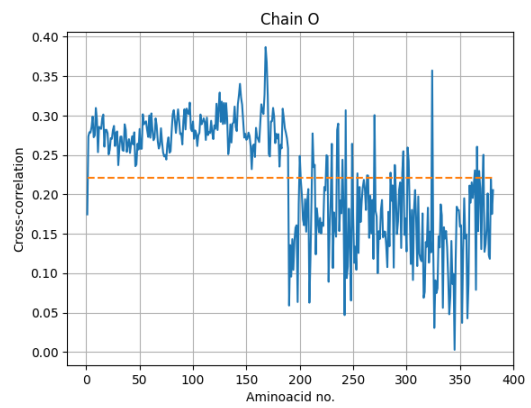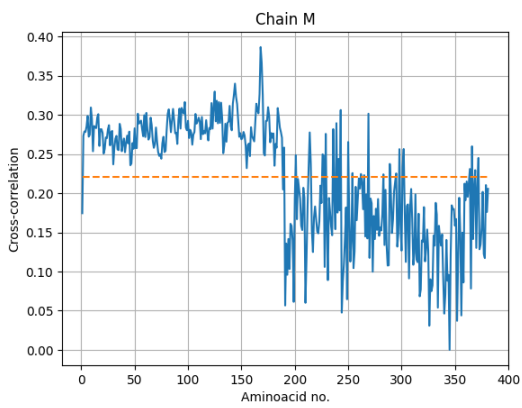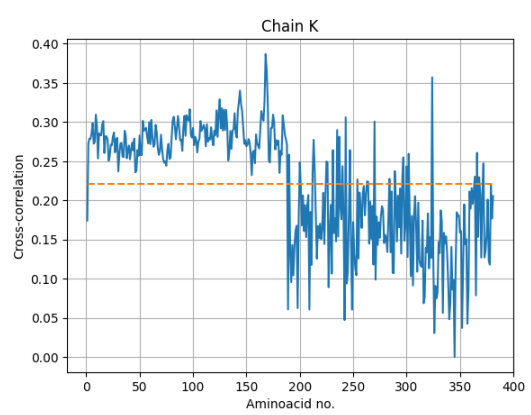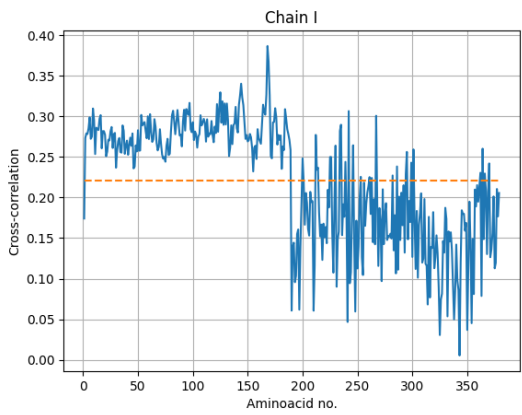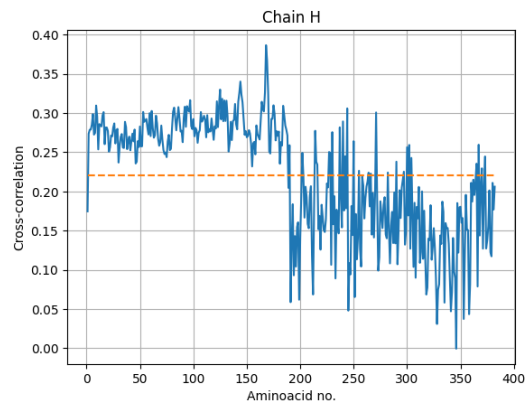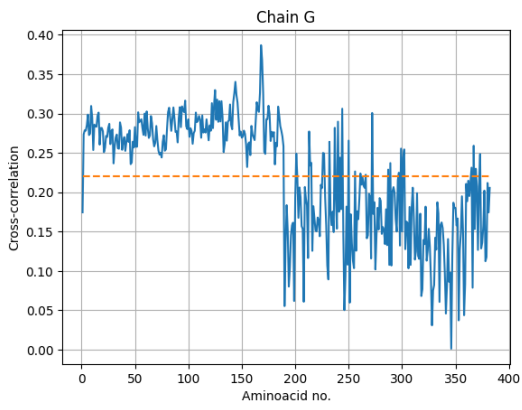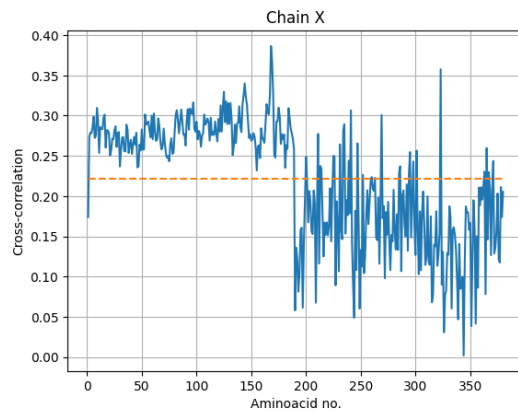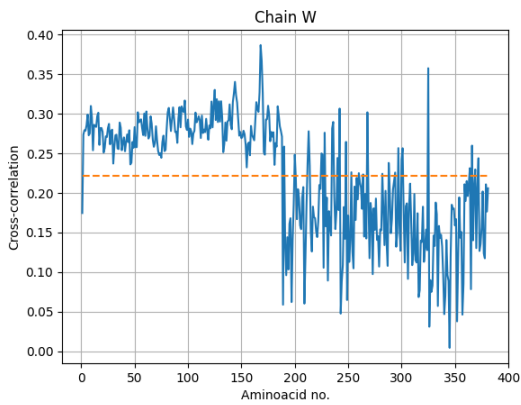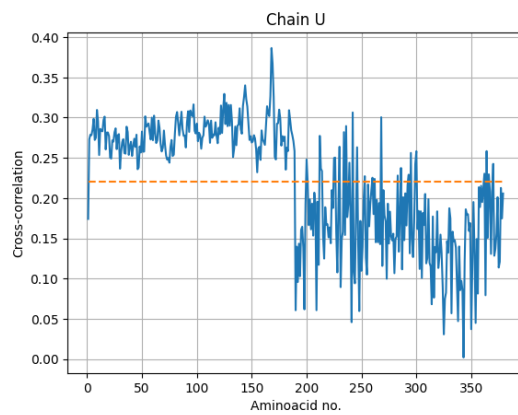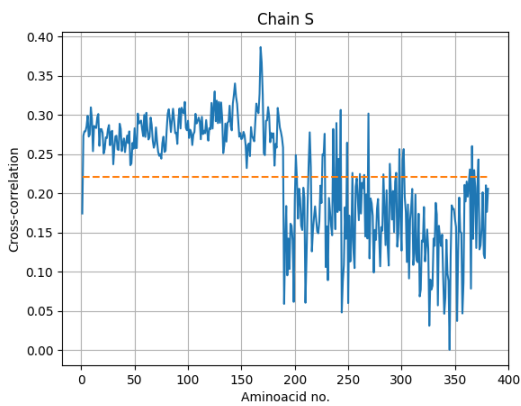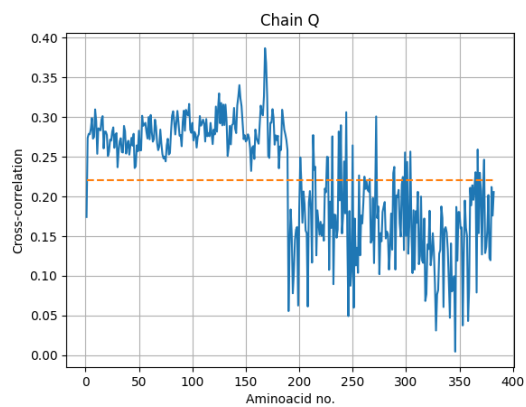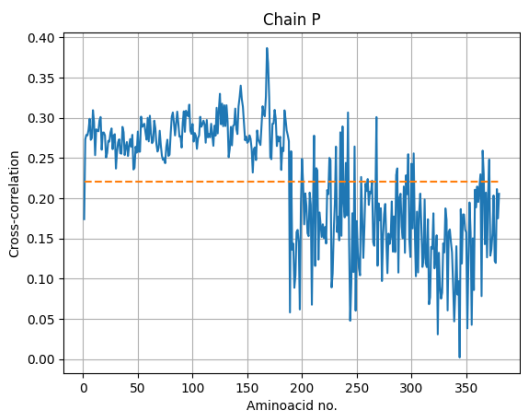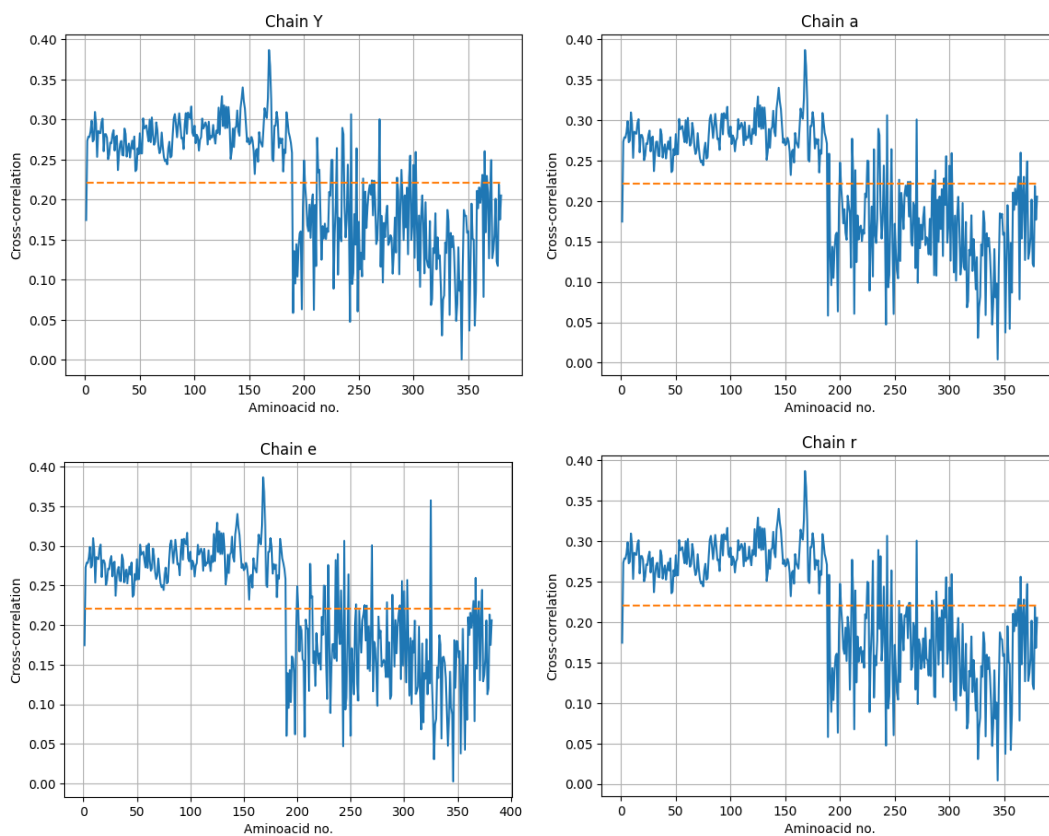
Fig. 14 shows the histogram of all cross-correlations evaluated at the residues. The percentage of residues whose correlation is below 0.5 is 100.0 %.
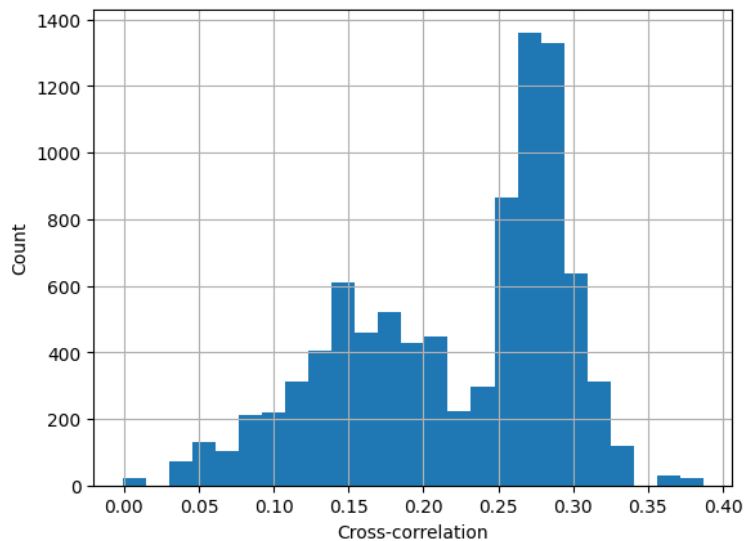
Figure 14: Histogram of the cross-correlation between the map and model evaluated for all residues.

Resolutions estimated from the model:

Resolution (Å)	Masked	Unmasked
d99	1.1	1.1
d_model	1.2	1.2
d_model (B-factor=0)	1.2	1.2
FSC_model=0	1.1	1.1
FSC_model=0.143	1.1	1.1
FSC_model=0.5	1.2	1.2

Overall isotropic B factor:

B factor	Masked	Unmasked
Overall B-iso	0.0	0.0

Fig. 15 shows the FSC between the input map and the model.

29
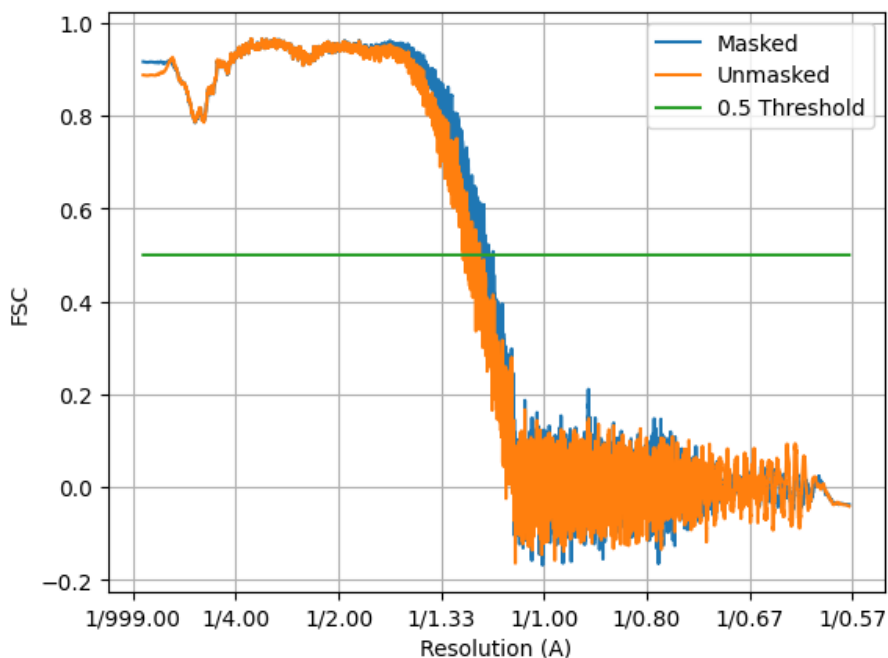
Figure 15: FSC between the input map and model with and without a mask constructed from the model. The X-axis is the square of the inverse of the resolution in Å.

Automatic criteria: The validation is OK if 1) the percentage of residues whose correlation is smaller than 0.5 is smaller than 10%, and 2) the resolution reported by the user is larger than 0.8 times the resolution estimated between the map and model at FSC=0.5.

WARNINGS: 1 warnings

1. **The percentage of residues that have a cross-correlation below 0.5 is 100.0, that is larger than 10%**

4.4 Level A.f EMRinger validation

Explanation:
EMringer [Barad et al., 2015] compares the side chains of the atomic model to the CryoEM map. The following features are reported:

- Optimal Threshold: Electron potential map cutoff value at which the maximum EMRinger score was obtained.
- Rotamer Ratio: Fraction of rotameric residues at the Optimal threshold value.
- Max Zscore: Z-score computed to determine the significance of the distribution at the Optimal threshold value.
- Model Length: Total of non-gamma-branched, non-proline aminoacids with a non-H gamma atom used in global EMRinger score computation.
- EMRinger Score: Maximum EMRinger score calculated at the Optimal Threshold.

A rotameric residue is one in which EMRinger peaks that fall within defined rotamers based on chi1, this often suggests a problem with the modelling of the backbone. In general, the user should look at the profiles and identify regions that may need improvement.

Results:

General results:

Optimal threshold	0.094601
Rotamer ratio	1.000
Max. Zscore	50.58
Model length	3144
EMRinger Score	9.021

Fig. 16 shows the EMRinger score and fraction of rotameric residues as a function of the map threshold. The optimal threshold was selected looking for the maximum EMRinger score in this plot.
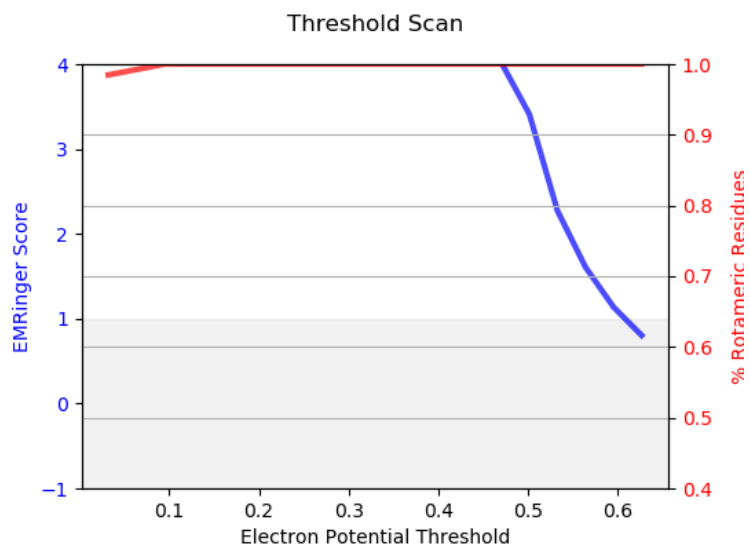
Figure 16: EMRinger score and fraction of rotameric residues as a function of the map threshold.

Fig. 17 shows the histogram for rotameric (blue) and non-rotameric (red) residues at the optimal threshold.
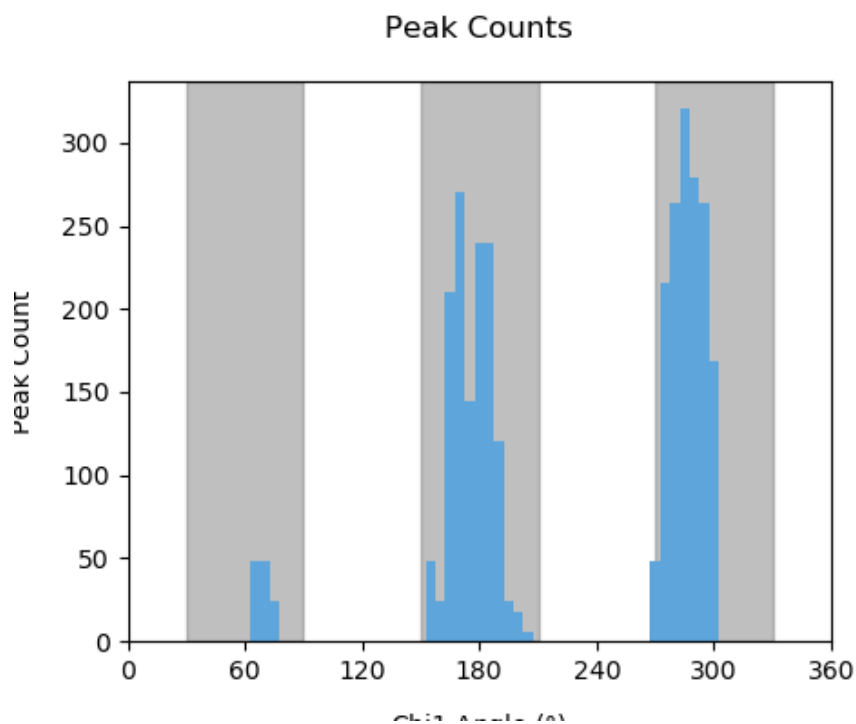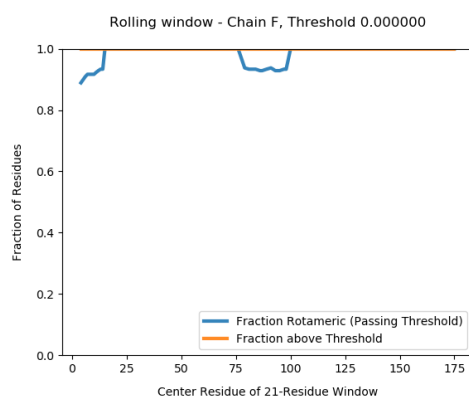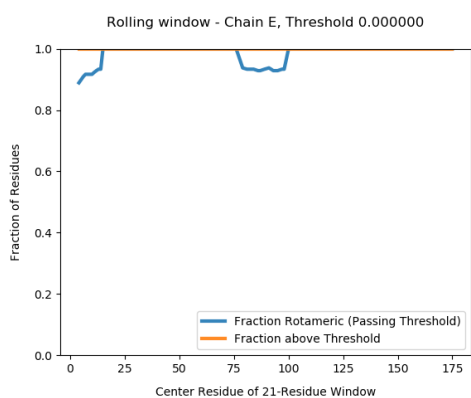
Figure 17: Histogram for rotameric (blue) and non-rotameric (red) residues at the optimal threshold as a function of the angle Chi1.

The following plots show the rolling window EMRinger analysis of the different chains to distinguish regions of improved model quality. This analysis was performed on rolling sliding 21-residue windows along the primary sequence of the protein chains.

Rolling window - Chain 4, Threshold 0.000000

Rolling window - Chain 6, Threshold 0.000000

Rolling window - Chain A, Threshold 0.000000

Rolling window - Chain B, Threshold 0.000000

Rolling window - Chain E, Threshold 0.000000

Rolling window - Chain F, Threshold 0.000000

Rolling window - Chain G, Threshold 0.000000

Rolling window - Chain H, Threshold 0.000000

Rolling window - Chain I, Threshold 0.000000

Rolling window - Chain K, Threshold 0.000000

Rolling window - Chain M, Threshold 0.000000

Rolling window - Chain O, Threshold 0.000000

Rolling window - Chain P, Threshold 0.000000

Rolling window - Chain Q, Threshold 0.000000

Rolling window - Chain S, Threshold 0.000000

Rolling window - Chain U, Threshold 0.000000

Rolling window - Chain W, Threshold 0.000000

Rolling window - Chain X, Threshold 0.000000

Rolling window - Chain Y, Threshold 0.000000

Rolling window - Chain a, Threshold 0.000000

Rolling window - Chain e, Threshold 0.000000

Rolling window - Chain r, Threshold 0.000000

Automatic criteria: The validation is OK if the EMRinger score and Max. Zscore are larger than 1.

STATUS: OK

4.5 Level A.g DAQ validation

Explanation:
DAQ [Terashi et al., 2022] is a computational tool using deep learning that can estimate the residue-wise local quality for protein models from cryo-Electron Microscopy maps. The method calculates the likelihood that a given density feature corresponds to an aminoacid, atom, and secondary structure. These likelihoods are combined into a score that ranges from -1 (bad quality) to 1 (good quality).

Results:

37

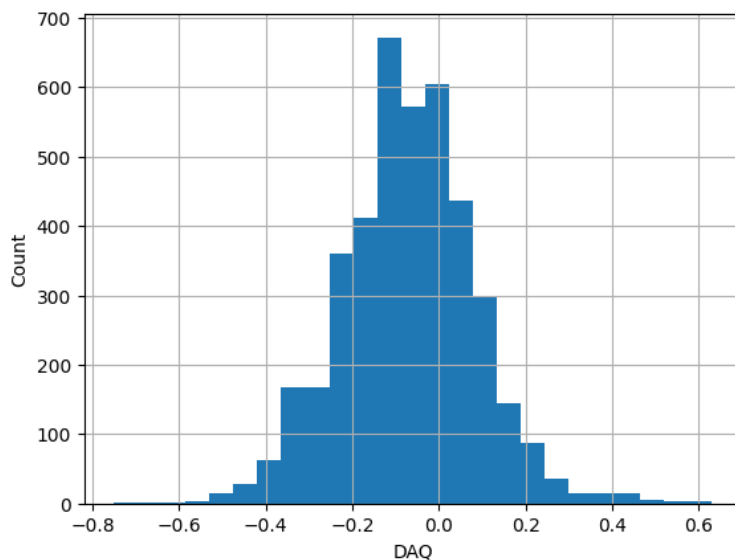Fig. 18 shows the histogram of the DAQ values. The mean and standard deviation were -0.1 and 0.2, respectively.

Figure 18: Histogram of the DAQ values.

The atomic model colored by DAQ can be seen in Fig. 19.

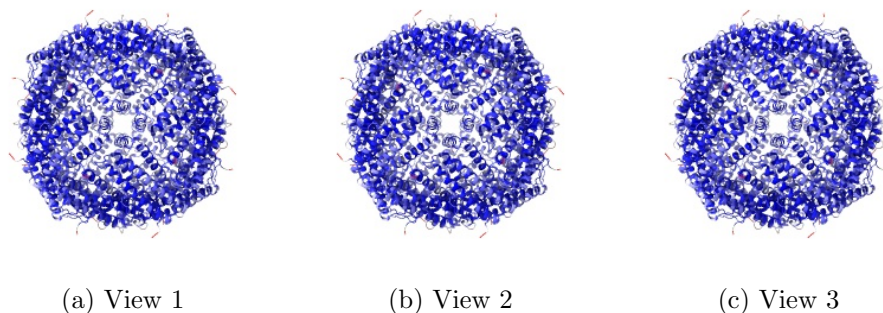(a) View 1   (b) View 2   (c) View 3

Figure 19: Atomic model colored by DAQ Views generated by ChimeraX at a the following X, Y, Z angles: View 1 (0,0,0), View 2 (90, 0, 0), View 3 (0, 90, 0).

Automatic criteria: The validation is OK if the average DAQ score is

larger than 0.5.

WARNINGS: 1 warnings

1. **The average DAQ is smaller than 0.5.**

References

[Rosenthal and Henderson, 2003] Rosenthal, P. B. and Henderson, R. (2003). Optimal determination of particle orientation, absolute hand, and contrast loss in single particle electron-cryomicroscopy. *J. Molecular Biology*, 333:721–745.

[Ramírez-Aportela et al., 2019] Ramírez-Aportela, E., Mota, J., Conesa, P., Carazo, J. M., and Sorzano, C. O. S. (2019). DeepRes: a new deep-learning- and aspect-based local resolution method for electron-microscopy maps. *IUCRj*, 6:1054–1063.

[Kaur et al., 2021] Kaur, S., Gomez-Blanco, J., Khalifa, A. A., Adinarayanan, S., Sanchez-Garcia, R., Wrapp, D., McLellan, J. S., Bui, K. H., and Vargas, J. (2021). Local computational methods to improve the interpretability and analysis of cryo-EM maps. *Nature Communications*, 12(1):1–12.

[Pintilie et al., 2020] Pintilie, G., Zhang, K., Su, Z., Li, S., Schmid, M. F., and Chiu, W. (2020). Measurement of atom resolvability in cryo-em maps with q-scores. *Nature methods*, 17(3):328–334.

[Afonine et al., 2018] Afonine, P. V., Klaholz, B. P., Moriarty, N. W., Poon, B. K., Sobolev, O. V., Terwilliger, T. C., Adams, P. D., and Urzhumtsev, A. (2018). New tools for the analysis and validation of cryo-EM maps and atomic models. *Acta Crystallographica D, Struct. Biol.*, 74:814–840.

[Barad et al., 2015] Barad, B. A., Echols, N., Wang, R. Y.-R., Cheng, Y., DiMaio, F., Adams, P. D., and Fraser, J. S. (2015). EMRinger: side chain-directed model and map validation for 3D cryo-electron microscopy. *Nature Methods*, 12(10):943–946.

[Terashi et al., 2022] Terashi, G., Wang, X., Subramaniya, S.R.M.V., Tesmer, J.J.G. and Kihara, D. (2022). Residue-Wise Local Quality Estimation for Protein Models from Cryo-EM Maps. (submitted).