



Public Data-based Report



Validation Report Service

Cryo-EM Map Validation Report

Report to assess Cryo-EM Volume Map at Level(s) A, 0, 1

This report has been generated based on data publicly available at [EMDB](#).

Basic Entry Information:

EMDB ID: [EMD-34397](#)

PDB ID: [8gzg](#)

Title: Cryo-EM structure of Synechocystis sp. PCC 6803 RPitc

Authors: [See EMDB entry link](#)

Deposited on: 2022-09-27T00:00:00

Reported Resolution: 3.13 Å

Contact Us:

Instruct Image Processing Center ([I²PC](#))

Biocomputing Unit ([BCU](#))

i2pc@cnb.csic.es

[VRS Website](#)

National Center for Biotechnology (CNB)
St/ Darwin, 3 (Autonomous University of Madrid)
28049 Cantoblanco, Madrid (Spain)

Last update: **August 21, 2025, 5:28am**

Context

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 and keeping the quality of the submitted maps is essential to maintain the scientific quality of the field. The ultimate quality measure is the consistency of the map and an atomic model. However, this is only possible for high resolution maps. Over the years there have been many suggestions about validation measures of 3DEM maps. Unfortunately, most of these measures are not currently in use for their spread in multiple software tools and the associated difficulty to access them. To alleviate this problem, we made available a validation grading system that evaluate the information provided to assess the map.

This system grades a map from 0 to 5 depending on the amount of information available. 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). In addition, we can have three optional qualifiers: A (... + atomic model), W (... + image processing workflow), and O (... + other techniques). To know more about this service read this [paper](#)

This Validation Report Service uses Scipion (see this [link](#) for more detail) as workflow engine and ChimeraX (see this [link](#) for more detail) to generate the 3D views. For more information about the different methods and softwares used for this report, see the references [here](#).



Summarized overall quality

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](#)). According to phenix, it seems that there might be some mismatch between the map and its model (see Sec. [6.1](#)). There seems to be a problem with its MapQ scores (see Sec. [6.4](#)).

The average resolution of the map estimated by various methods goes from 1.5Å to 7.1Å with an average of 3.6Å. The resolution reported by the user was 3.1Å.

The overall score (passing tests) of this report is 18 out of 21 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	OK
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	OK
1.e MonoRes	Sec. 4.5	OK
1.f MonoDir	Sec. 4.6	Could not be measured
1.g FSO	Sec. 4.7	OK
1.h FSC3D	Sec. 4.8	OK
A.a Phenix validation	Sec. 6.1	1 warnings
A.b FSC-Q	Sec. 6.2	OK
A.d Map-Model Guinier	Sec. 6.3	OK
A.e MapQ	Sec. 6.4	2 warnings
A.f EMRinger	Sec. 6.5	OK
A.g DAQ	Sec. 6.6	OK

Summary of the warnings across sections.

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=8269.56)**

Section 6.1 (A.a Phenix validation)

1. **The percentage of residues that have a cross-correlation below 0.5 is 10.2, that is larger than 10%**

Section 6.4 (A.e MapQ)

1. **Program output seems to be too homogeneous. There might be some program issues analyzing the data.**
2. **The median Q-score is less than 0.1.**

Contents

1	Input data	6
2	Level 0 analysis	13
2.1	Level 0.a Mass analysis	13
2.2	Level 0.b Mask analysis	14
2.3	Level 0.c Background analysis	17
2.4	Level 0.d B-factor analysis	19
2.5	Level 0.e Local resolution with DeepRes	22
2.6	Level 0.f Local B-factor	25
2.7	Level 0.g Local Occupancy	28
2.8	Level 0.h Hand correction	31
3	Half maps	31
4	Level 1 analysis	35
4.1	Level 1.a Global resolution	35
4.2	Level 1.b FSC permutation	37
4.3	Level 1.c Local resolution with Blocres	38
4.4	Level 1.d Local resolution with Resmap	41
4.5	Level 1.e Local resolution with MonoRes	44
4.6	Level 1.f Local and directional resolution with MonoDir	47
4.7	Level 1.g Fourier Shell Occupancy	48
4.8	Level 1.h Fourier Shell Correlation 3D	50
5	Atomic model	52
6	Level A Analysis	53
6.1	Level A.a Phenix validation	53
6.2	Level A.b FSC-Q	58
6.3	Level A.d Map-Model Guinier analysis	61
6.4	Level A.e MapQ	62
6.5	Level A.f EMRinger validation	64
6.6	Level A.g DAQ validation	68

1 Input data

Input map: emd_34397.map

SHA256 hash: 5a7e191c433e52210e4b6564d193c5a1974c2b4098a79396c797ee715e593bdd

Voxel size: 1.014000 (Å)

Visualization threshold: 0.010600

Resolution estimated by user: 3.13

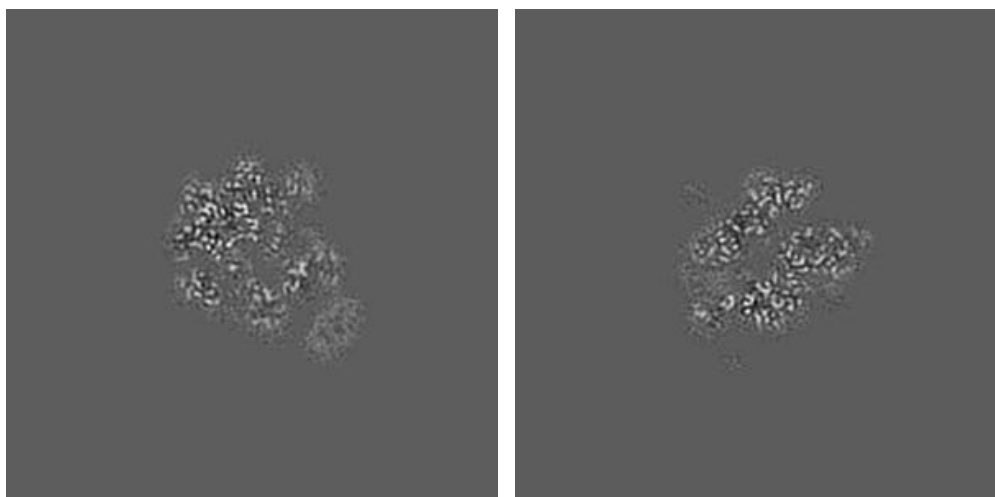
Orthogonal slices of the input map

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, ...)

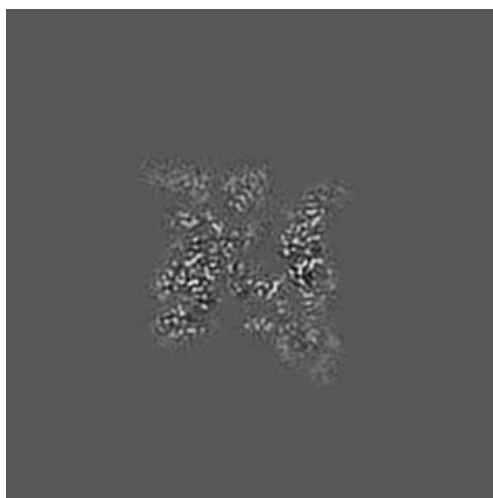
Results:

See Fig. 1.



(a) X Slice 150

(b) Y Slice 150



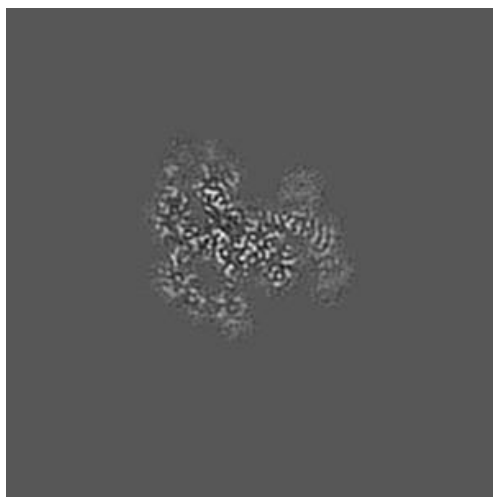
(c) Z Slice 150

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

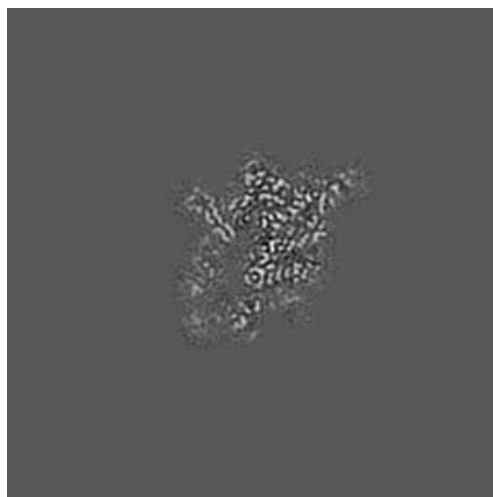
Orthogonal slices of maximum variance of the input map

Results:

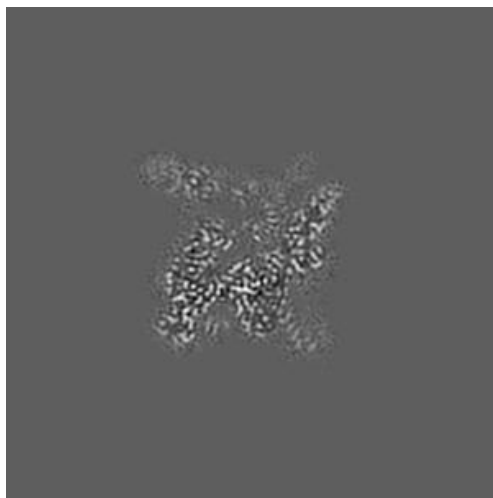
See Fig. 2.



(a) X Slice 115



(b) Y Slice 130



(c) Z Slice 157

Figure 2: Slices of maximum variation in the three dimensions

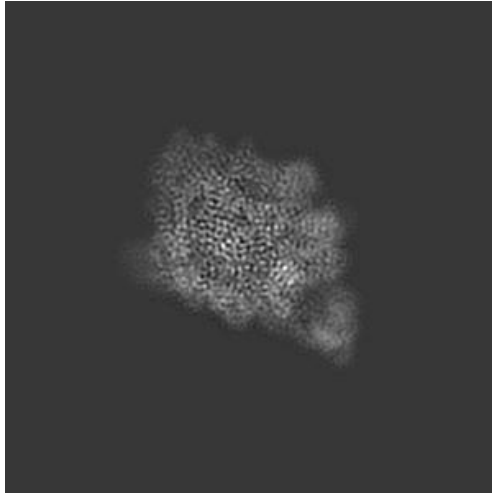
Orthogonal projections of the input map

Explanation:

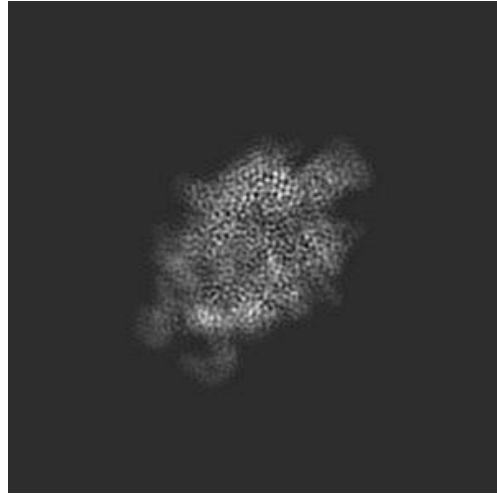
In the projections there should not be stripes (this is an indication of directional overweighting, or angular attraction), and there should not be a dark halo around or inside the structure (this is an indication of incorrect CTF correction or the reconstruction of a biased map).

Results:

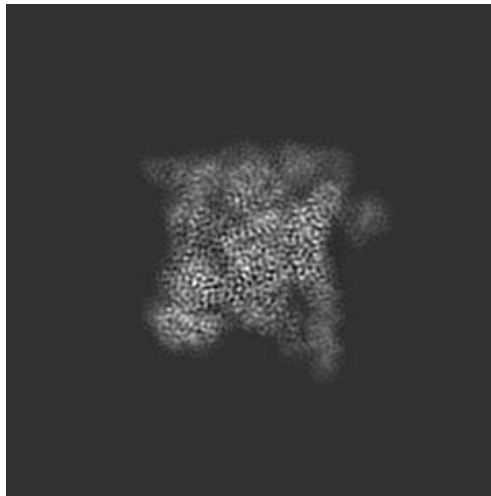
See Fig. [3](#).



(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.

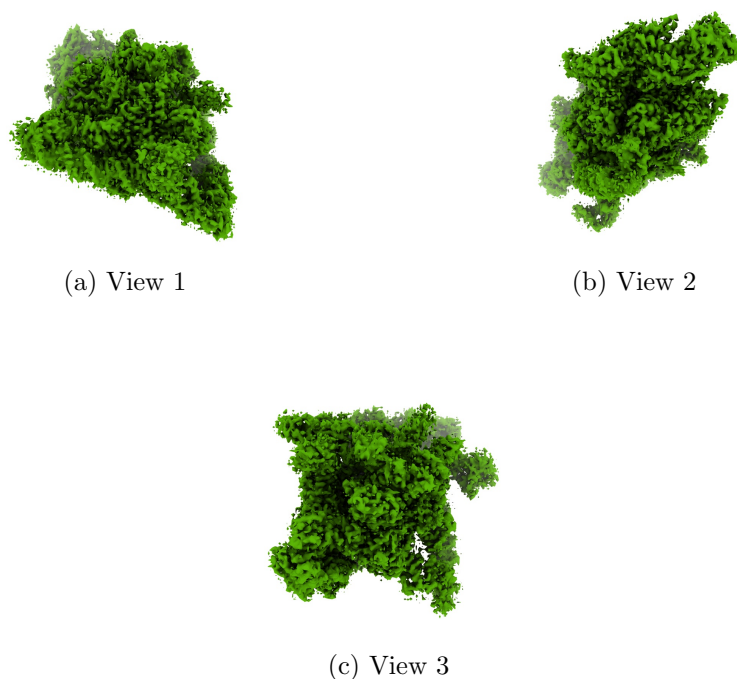


Figure 4: Isosurface at threshold=0.010600. Views generated by ChimeraX at a the following X, Y, Z angles: View 1 (0, -90, -90), View 2 (-90, 0, -90), View 3 (0, 0, 0).

Orthogonal slices of maximum variance of the mask with hard borders

Explanation:

The mask with hard borders has been calculated at the suggested threshold 0.010600, the largest connected component was selected, and then dilated by 2Å.

Results:

See Fig. 5.

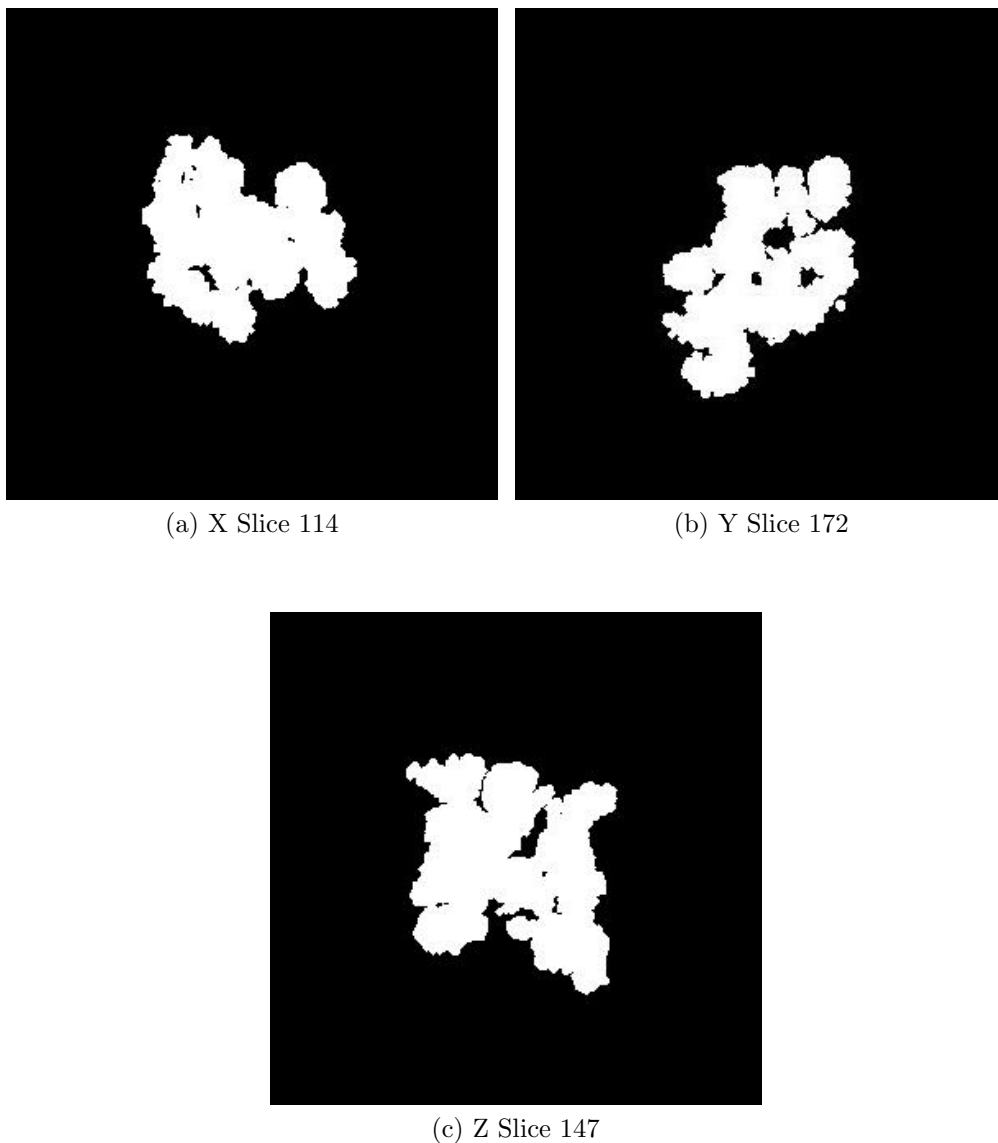


Figure 5: Slices of maximum variation in the three dimensions of the mask with hard borders

Orthogonal slices of maximum variance of the mask with soft borders

Explanation:

The mask with soft borders has been calculated at the suggested threshold 0.010600, the largest connected component was selected, and then dilated by

2Å.

Results:
See Fig. 6.



(a) X Slice 114



(b) Y Slice 171



(c) Z Slice 148

Figure 6: Slices of maximum variation in the three dimensions of the mask with soft borders

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 $30\text{\AA}$ (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 80.11 and 63.88 $\text{\AA}$, respectively. There is a decentering ratio $(\text{abs}(\text{Right-Left})/\text{Size})\%$ of 5.33%

The space from the left and right in Y are 67.94 and 80.11 $\text{\AA}$, respectively. There is a decentering ratio $(\text{abs}(\text{Right-Left})/\text{Size})\%$ of 4.00%

The space from the left and right in Z are 74.02 and 75.04 $\text{\AA}$, respectively. There is a decentering ratio $(\text{abs}(\text{Right-Left})/\text{Size})\%$ of 0.33%

The center of mass is at $(x,y,z)=(149.28,150.02,149.26)$. The decentering of the center of mass $(\text{abs}(\text{Center})/\text{Size})\%$ is 0.24, 0.01, and 0.25, 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. For local reconstruction, focused refinement, or similar, warnings are expected.

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 $(\text{size}(\text{Raw AND Constructed})/\text{size}(\text{Raw}))$ that is a number between 0 and 1, the closer to 1, the more they agree.

Results:

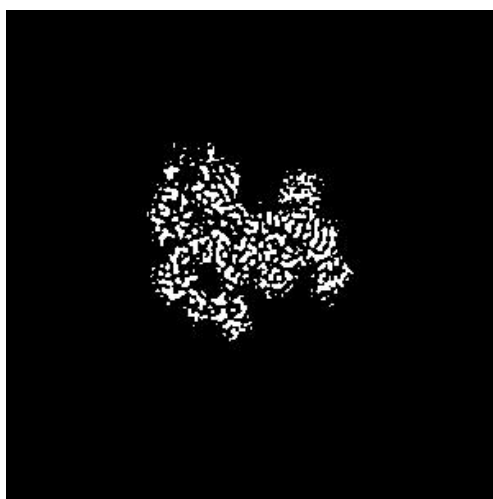
Raw mask: At threshold 0.010600, there are 2026 connected components with a total number of voxels of 272763 and a volume of $284380.18 \text{ \AA}^3$ (see Fig. 7). The size and percentage of the total number of voxels for the raw mask are listed below (up to 95% of the mass or the first 100 clusters, whatever happens first), the list contains (No. voxels (volume in $\text{\AA}^3$), percentage, cumulated percentage):

(264849 (276129.12), 97.10, 97.10)

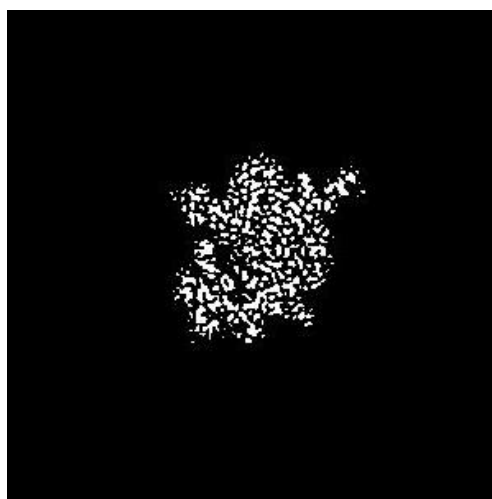
Number of components to reach 95% of the mass: 1

The average size of the remaining 2025 components is 3.91 voxels (1.04 $\text{\AA}^3$). Their size go from 264849 voxels (276129.12 $\text{\AA}^3$) to 1 voxels (1.04 $\text{\AA}^3$).

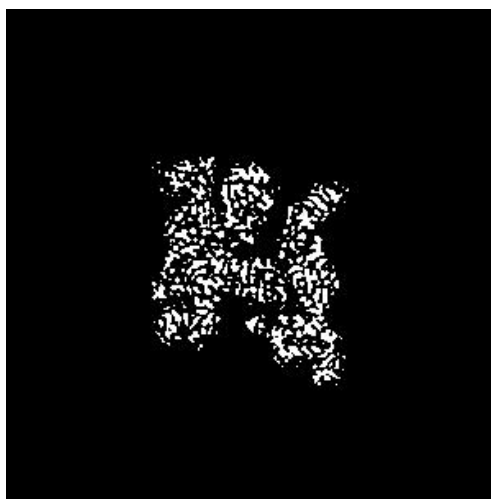
The slices of the raw mask can be seen in Fig. 7.



(a) X Slice 115



(b) Y Slice 133



(c) Z Slice 147

Figure 7: 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. 8):

Threshold	Voxel mass	Molecular mass(kDa)	# Aminoacids
0.0056	422154.00	364.65	3315.01
0.0112	258412.00	223.21	2029.21
0.0169	167981.00	145.10	1319.09
0.0225	117838.00	101.79	925.34
0.0281	87262.00	75.38	685.23
0.0337	66310.00	57.28	520.71
0.0394	50908.00	43.97	399.76
0.0450	38938.00	33.63	305.76
0.0506	29486.00	25.47	231.54
0.0562	21850.00	18.87	171.58
0.0619	15955.00	13.78	125.29
0.0675	11226.00	9.70	88.15
0.0731	7528.00	6.50	59.11
0.0787	4848.00	4.19	38.07
0.0844	3067.00	2.65	24.08
0.0900	1831.00	1.58	14.38
0.0956	1013.00	0.88	7.95
0.1012	525.00	0.45	4.12
0.1069	274.00	0.24	2.15
0.1125	122.00	0.11	0.96
0.1181	52.00	0.04	0.41
0.1237	20.00	0.02	0.16
0.1294	7.00	0.01	0.05
0.1350	3.00	0.00	0.02

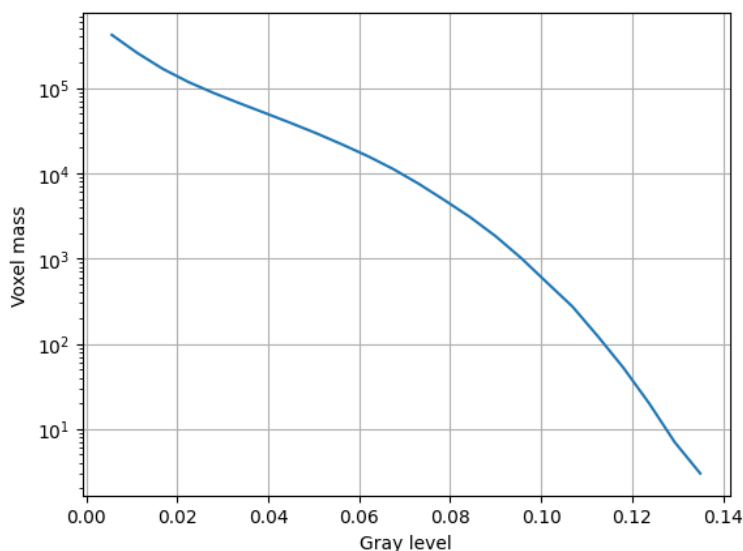


Figure 8: Voxel mass as a function of the gray level.

Constructed mask: After keeping the largest component of the previous mask and dilating it by $2\text{\AA}$, there is a total number of voxels of 980818 and a volume of $1022591.77 \text{ \AA}^3$. The overlap between the raw and constructed mask is 0.98.

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 $5\text{\AA}^3$; 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:

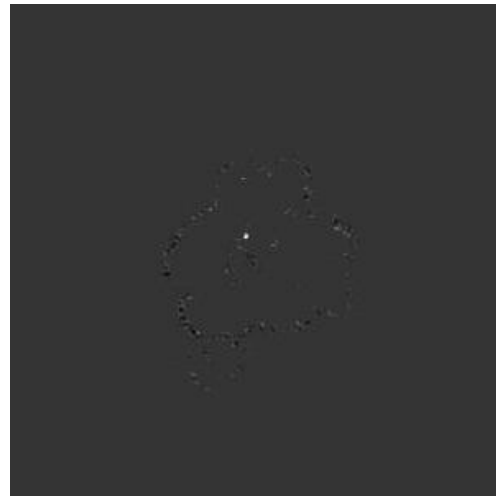
The null hypothesis that the background mean is 0 was tested with a one-sample Student's t-test. The resulting t-statistic and p-value were 41.59 and 0.000000, respectively.

The mean and standard deviation (σ) of the background were 0.000003 and 0.000385. The percentage of background voxels whose absolute value is larger than 5 times the standard deviation is 0.47 % (see Fig. 9). The same percentage from a Gaussian would be 0.000057% (ratio between the two percentages: 8269.558353).

Slices of the background beyond 5σ can be seen in Fig. 9.



(a) X Slice 101



(b) Y Slice 159



(c) Z Slice 143

Figure 9: 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=8269.56)**

2.4 Level 0.d B-factor analysis

Explanation:

The B-factor line (see this [link](#) for more details) fitted between 15Å and the resolution reported should have a slope that is between 0 and 300 Å².

Results:

Fig. 10 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 -23.5. The fitted line was $\log(|F|^2) = -5.9/R^2 + (-13.2)$.

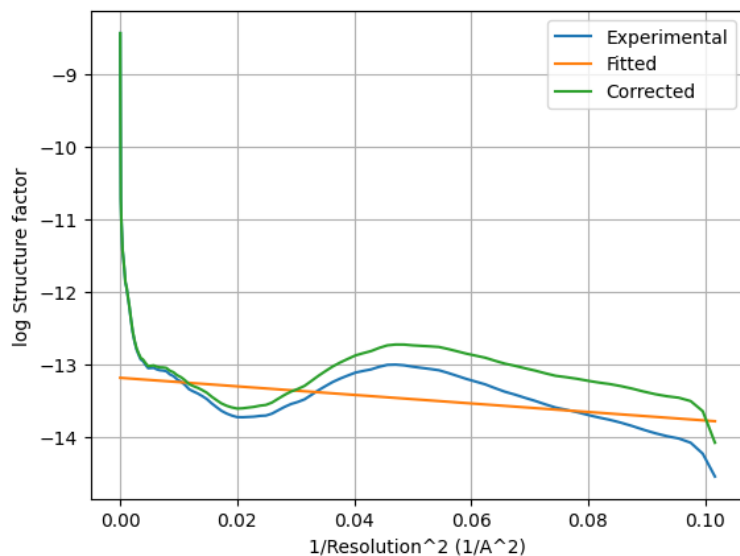


Figure 10: Guinier plot. The X-axis is the square of the inverse of the resolution in Å.

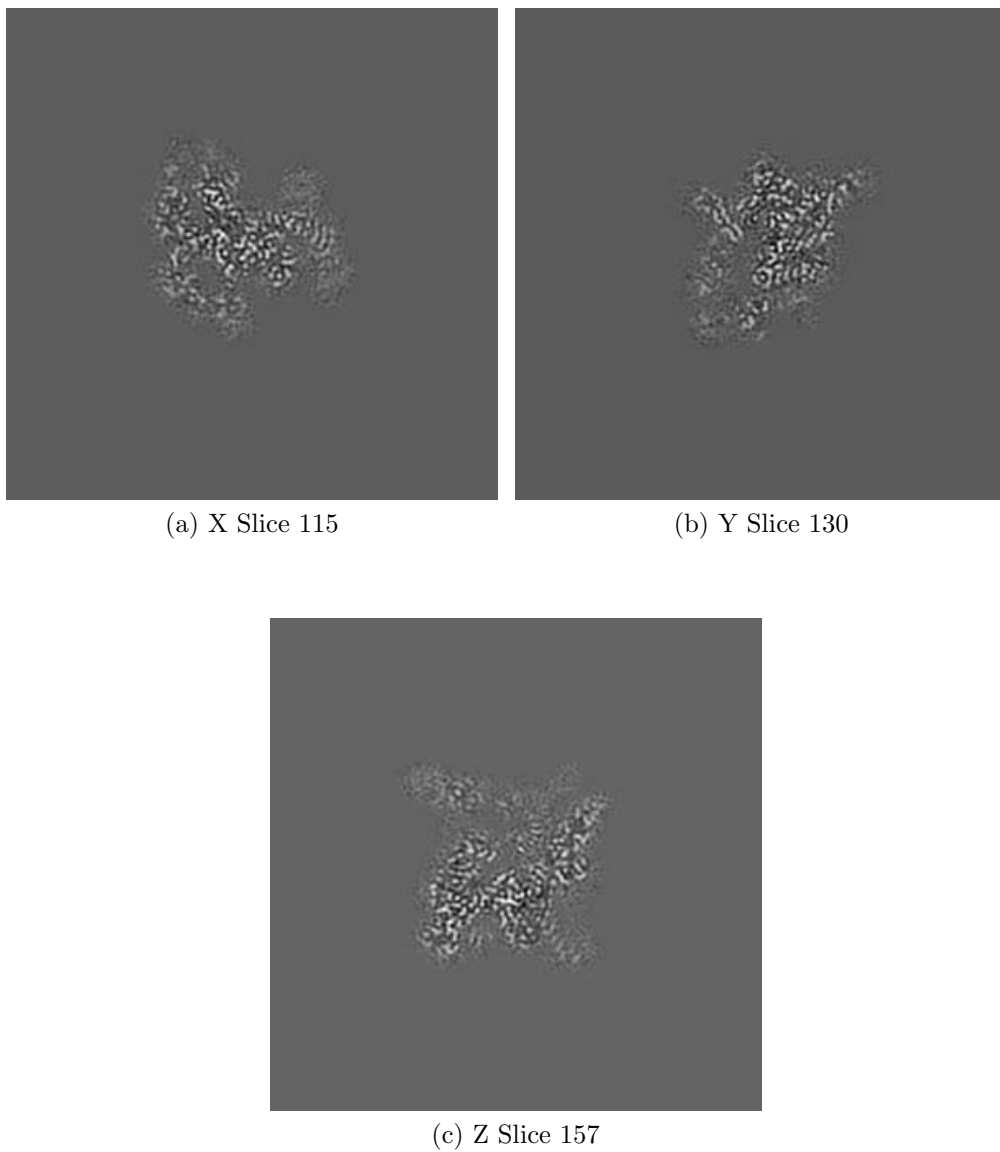


Figure 11: Slices of maximum variation in the three dimensions of the B-factor corrected map

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 (see this [link](#) for more details) 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. [12](#) shows the histogram of the local resolution according to DeepRes. Some representative percentiles are:

Percentile	Resolution($\text{\AA}$)
2.5%	2.50
25%	2.64
50%	3.07
75%	3.52
97.5%	4.26

The reported resolution, 3.13 $\text{\AA}$, is at the percentile 53.5. Fig. [13](#) shows some representative views of the local resolution.

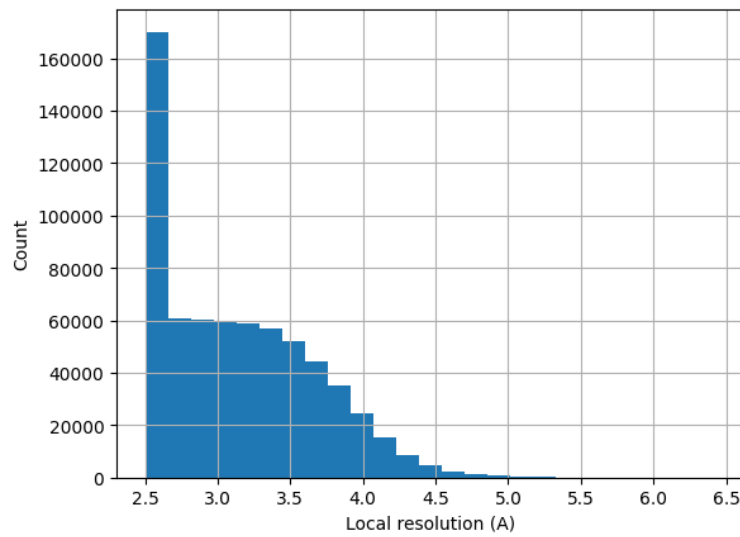
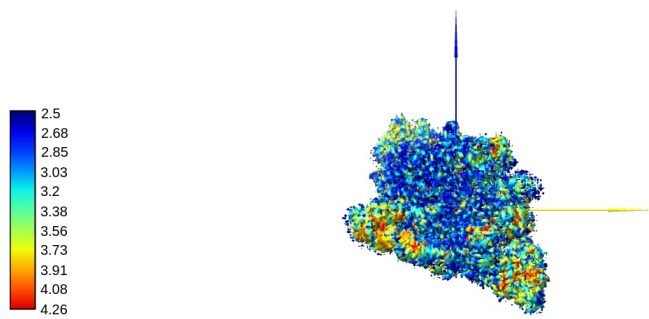
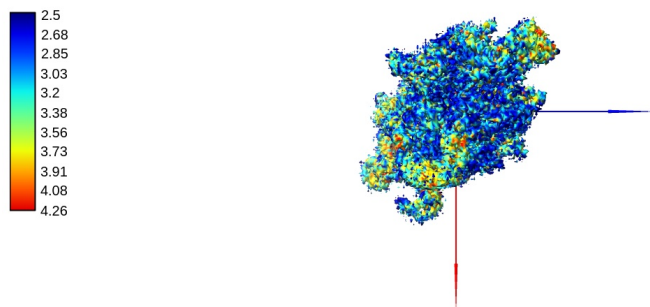


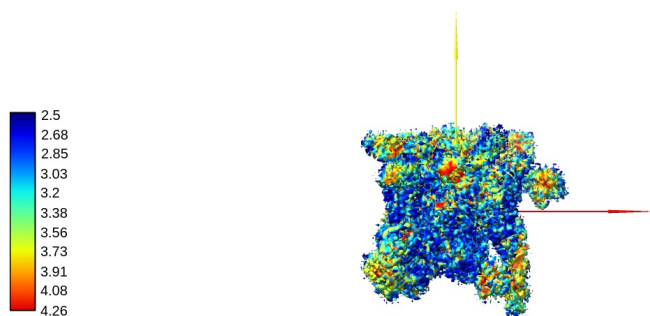
Figure 12: Histogram of the local resolution according to deepres.



(a) View 1



(b) View 2



(c) View 3

Figure 13: Local resolution according to DeepRes. Views generated by ChimeraX at the following X, Y, Z angles: View 1 (0, -90, -90), View 2 (-90, 0, -90), View 3 (0, 0, 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 (see this [link](#) for more details) estimates a local resolution B-factor by decomposing the input map into a local magnitude and phase term using the spiral transform.

Results:

Fig. [14](#) shows the histogram of the local B-factor according to LocBfactor. Some representative percentiles are:

Percentile	Local B-factor ($\text{\AA}^{-2}$)
2.5%	-135.39
25%	-105.38
50%	-89.67
75%	-73.59
97.5%	-39.77

Fig. [15](#) shows some representative views of the local B-factor.

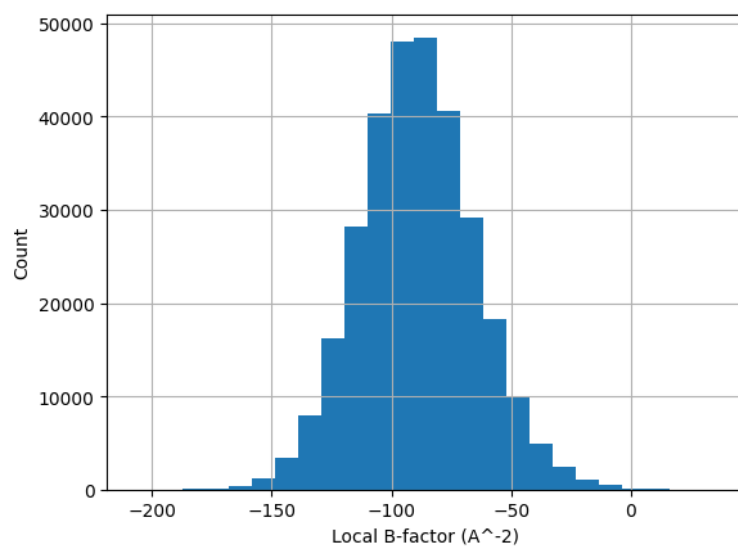
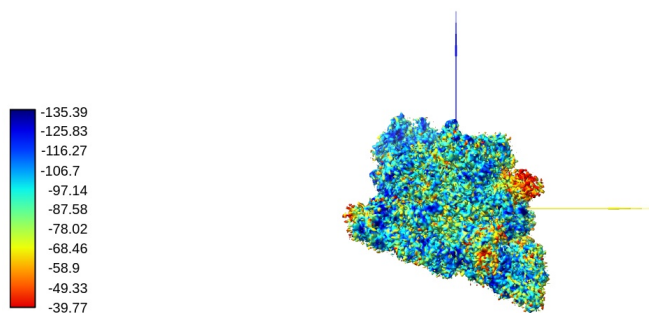
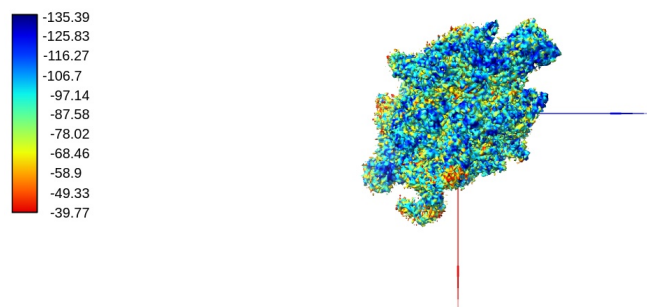


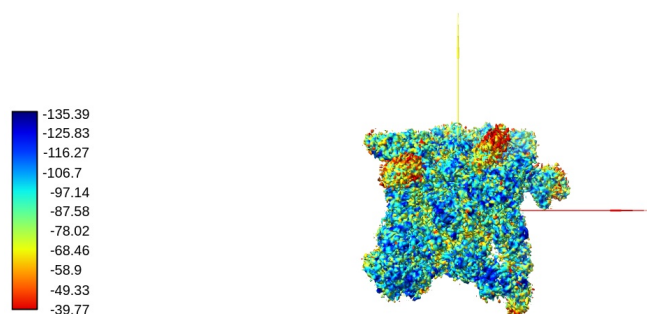
Figure 14: Histogram of the local B-factor according to LocBfactor.



(a) View 1



(b) View 2



(c) View 3

Figure 15: Local B-factor according to LocBfactor. Views generated by ChimeraX at the following X, Y, Z angles: View 1 (0, -90, -90), View 2 (-90, 0, -90), View 3 (0, 0, 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 (see this [link](#) for more details) estimates the occupancy of a voxel by the macromolecule.

Results:

Fig. [16](#) shows the histogram of the local occupancy according to LocOccupancy. Some representative percentiles are:

Percentile	Local Occupancy [0-1]
2.5%	0.00
25%	0.60
50%	0.87
75%	1.00
97.5%	1.00

Fig. [17](#) shows some representative views of the local occupancy.

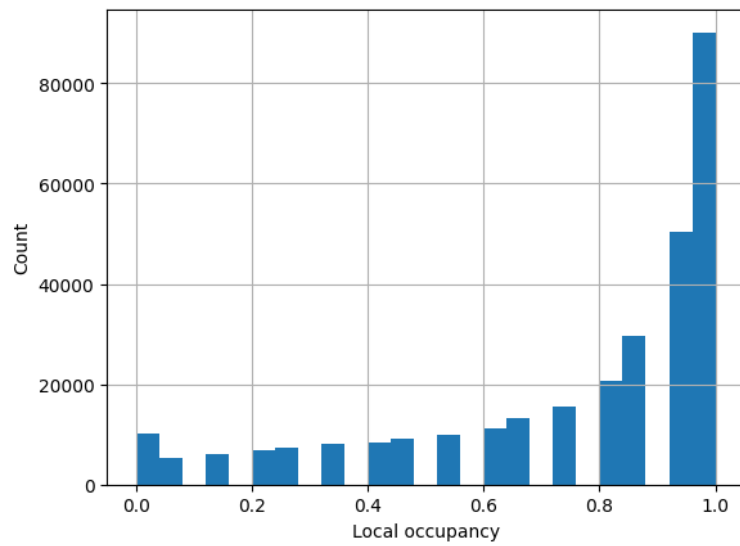
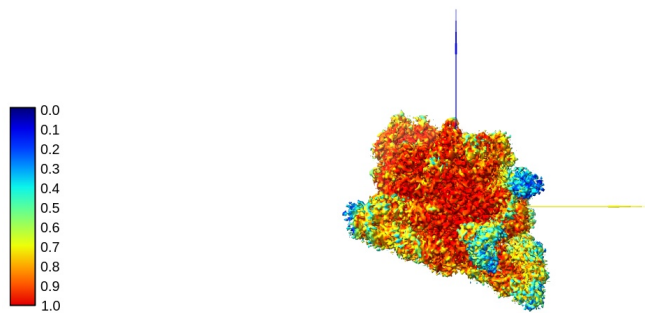
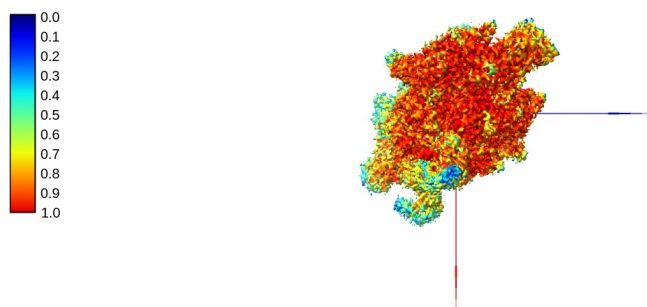


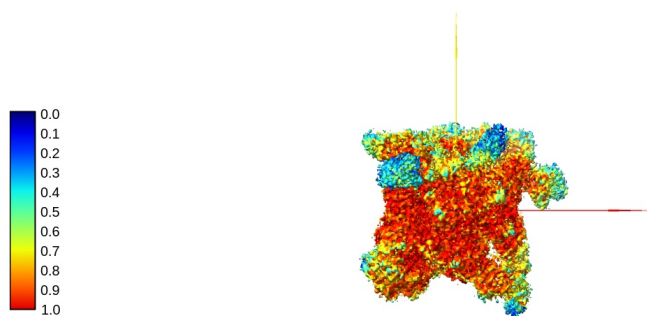
Figure 16: Histogram of the local occupancy according to LocOccupancy.



(a) View 1



(b) View 2



(c) View 3

Figure 17: Local occupancy according to LocOccupancy. Views generated by ChimeraX at a the following X, Y, Z angles: View 1 (0, -90, -90), View 2 (-90, 0, -90), View 3 (0, 0, 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 (see this [link](#) for more details) 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 0.245 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: emd_34397_half_map_2.map

SHA256 hash: 6884309e7f0fd489018dbcc0dbf2d3e2a85d48cfcb02cdaa29c3c2ab8da1209

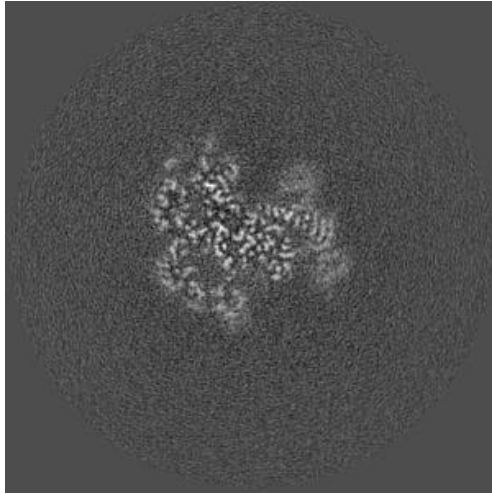
Half map 2: emd_34397_half_map_1.map

SHA256 hash: dad610e0cb43157066f987e26373e895687b9e9f2a5e76723b465b7924e08dd2

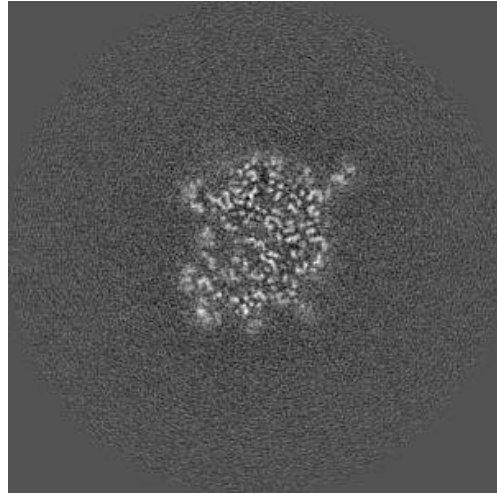
Slices of the first half map can be seen in Fig. [18](#).

Slices of the second half map can be seen in Fig. [19](#).

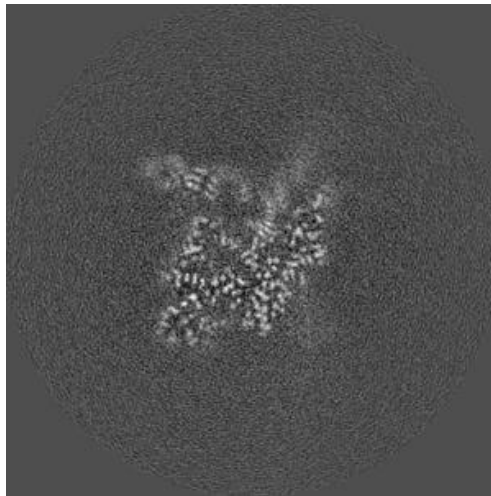
Slices of the difference between both maps can be seen in Fig. [20](#). There should not be any structure in this difference. Sometimes some patterns are seen if the map is symmetric.



(a) X Slice 115

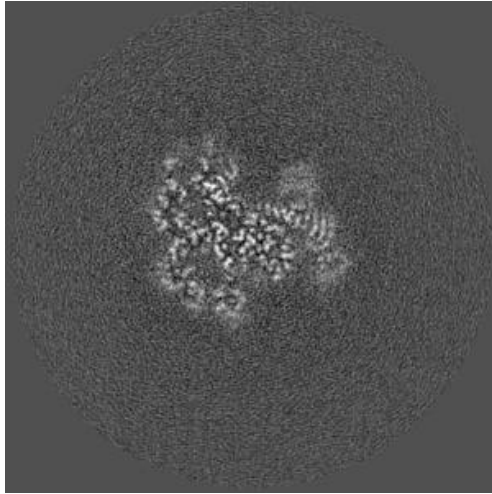


(b) Y Slice 134

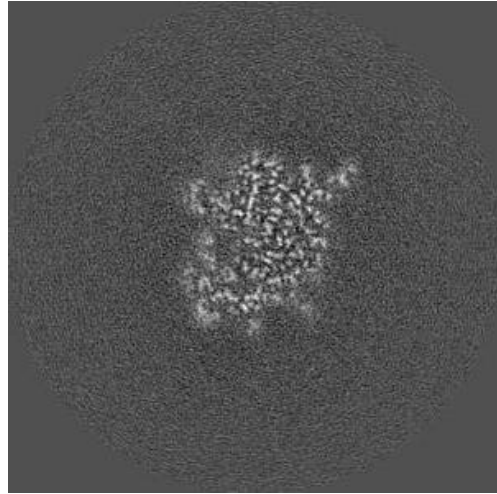


(c) Z Slice 160

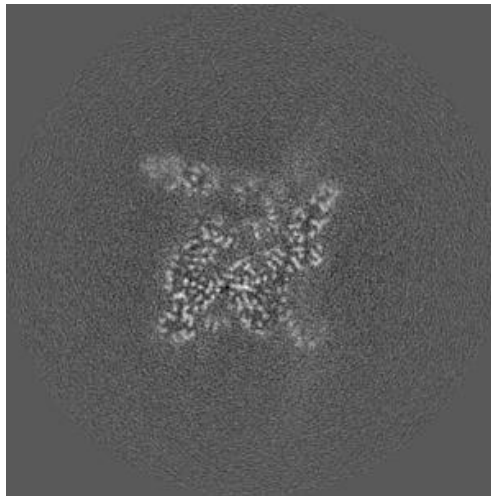
Figure 18: Slices of maximum variation in the three dimensions of Half 1



(a) X Slice 115

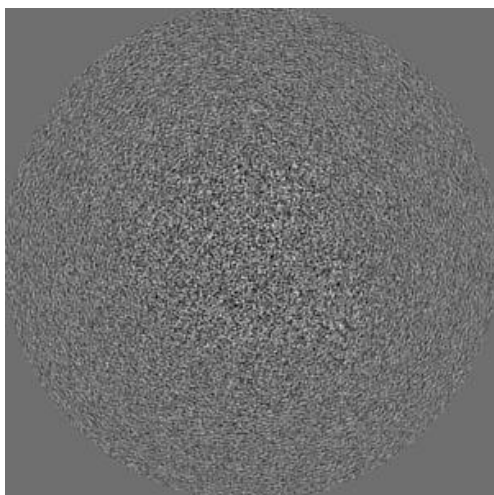


(b) Y Slice 133

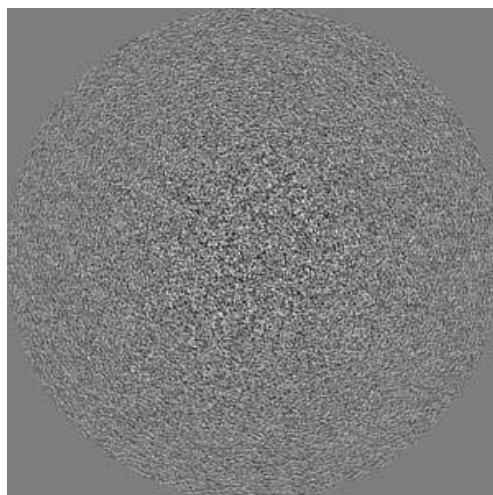


(c) Z Slice 157

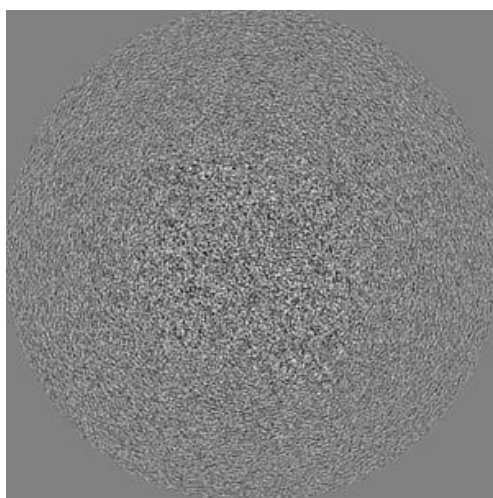
Figure 19: Slices of maximum variation in the three dimensions of Half 2



(a) X Slice 152



(b) Y Slice 140



(c) Z Slice 149

Figure 20: 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 this [link](#). 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. 21 shows the FSC and the 0.143 threshold. The resolution according to the FSC is 3.64Å. The map information is well preserved (FSC>0.9) up to 8.08Å.

Fig. 22 shows the DPR and the 103.9° threshold. The resolution according to the DPR is 3.09Å.

Fig. 23 shows the SSNR and the SSNR=1 threshold. The resolution according to the SSNR is 3.78Å.

The mean resolution between the three methods is 3.51Å and its range is within the interval [3.09, 3.78]Å.

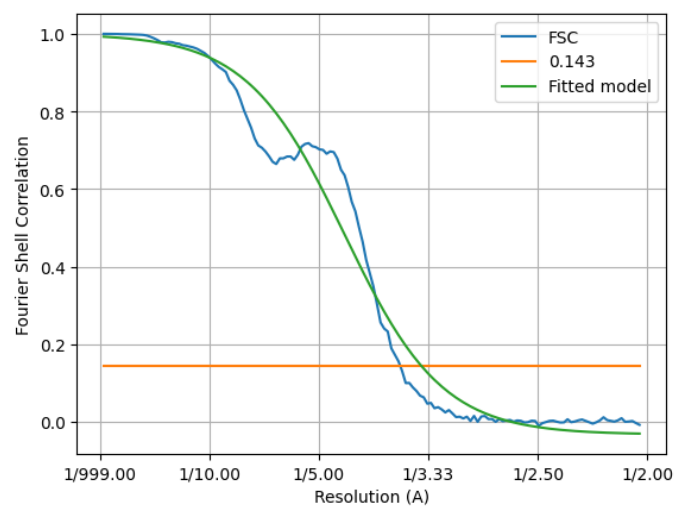


Figure 21: Fourier Shell correlation between the two halves.

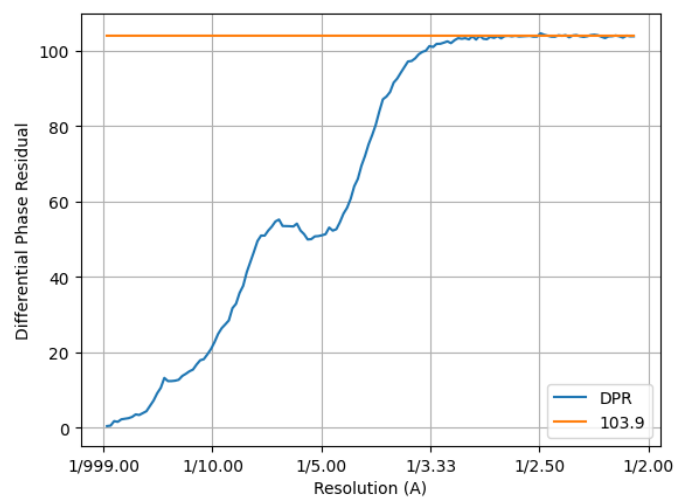


Figure 22: Differential Phase Residual between the two halves.

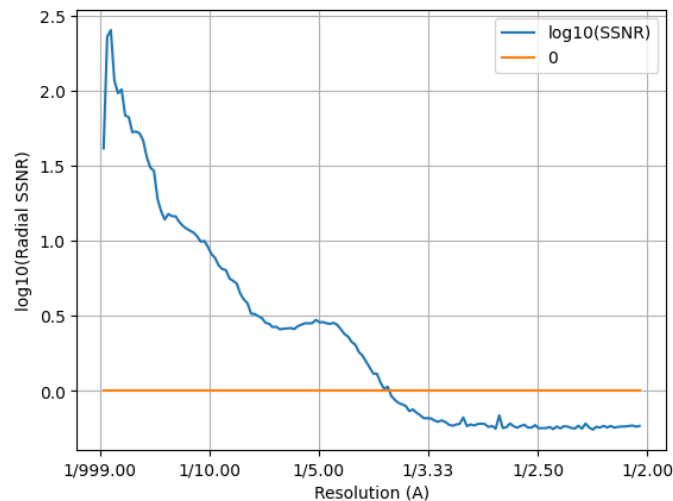


Figure 23: Spectral Signal-to-Noise Ratio estimated from the two halves.

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.

STATUS: OK

4.2 Level 1.b FSC permutation

Explanation:

This method (see this [link](#) for more details) 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 3.0. The estimated B-factor was -66.59. Fig. 24 shows the estimated FSC and resolution.

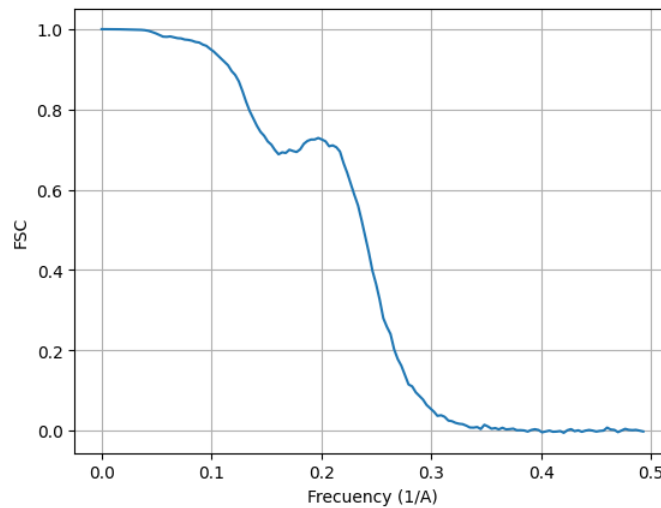


Figure 24: FSC and resolution estimated by a permutation test.

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 (see this [link](#) for more details) computes a local Fourier Shell Correlation (FSC) between the two half maps.

Results:

Fig. [25](#) shows the histogram of the local resolution according to Blocres. Some representative percentiles are:

Percentile	Resolution($\text{\AA}$)
2.5%	3.00
25%	3.30
50%	3.81
75%	5.26
97.5%	7.30

The reported resolution, 3.13 $\text{\AA}$, is at the percentile 11.1. Fig. 26 shows some representative views of the local resolution.

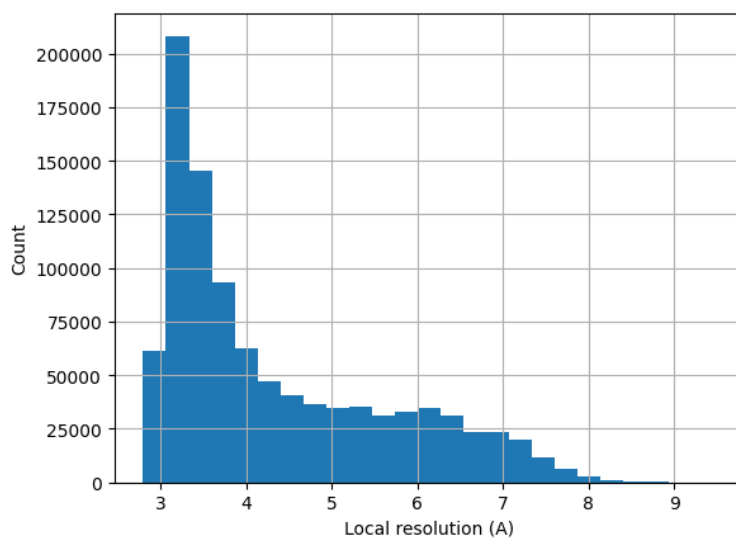
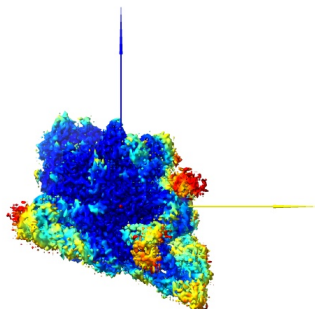
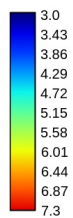
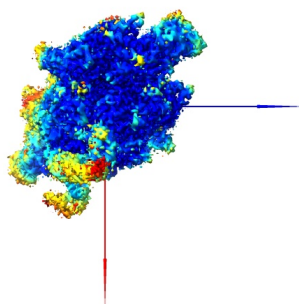
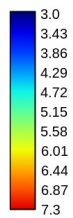


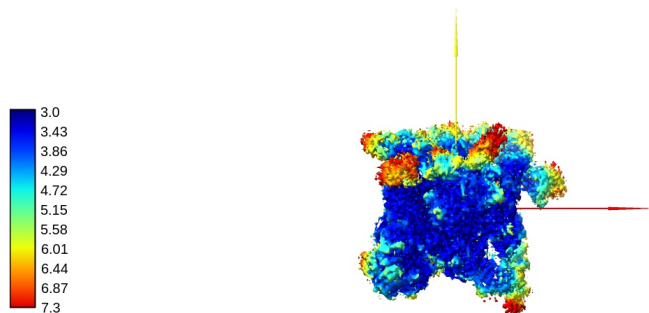
Figure 25: Histogram of the local resolution according to blocres.



(a) View 1



(b) View 2



(c) View 3

Figure 26: Local resolution according to Blocres. Views generated by ChimeraX at the following X, Y, Z angles: View 1 (0, -90, -90), View 2 (-90, 0, -90), View 3 (0, 0, 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 (see this [link](#) for more details) is based on a test hypothesis testing of the superiority of signal over noise at different frequencies.

Results:

Fig. 27 shows the histogram of the local resolution according to Resmap. Some representative percentiles are:

Percentile	Resolution($\text{\AA}$)
2.5%	2.88
25%	3.23
50%	3.56
75%	3.84
97.5%	4.06

The reported resolution, 3.13 $\text{\AA}$, is at the percentile 13. Fig. 28 shows some representative views of the local resolution.

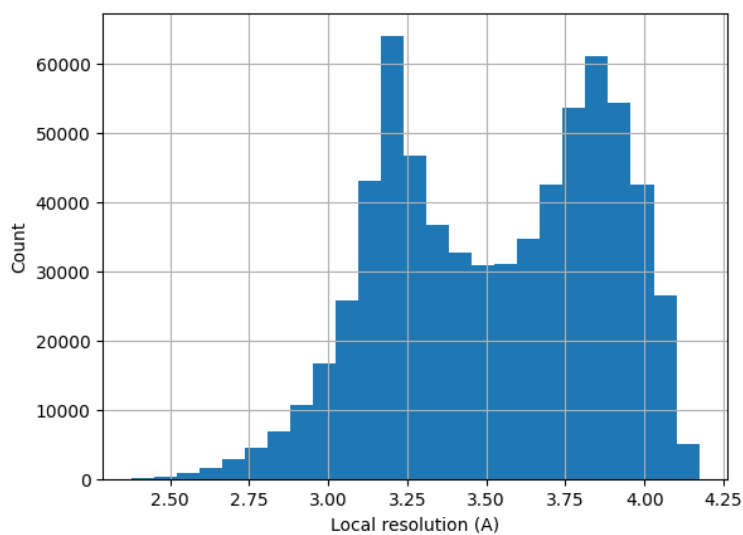
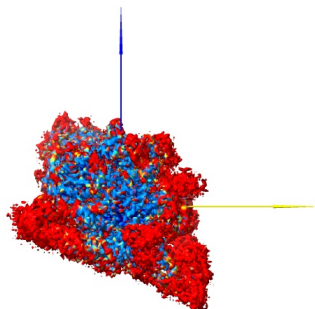
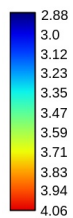
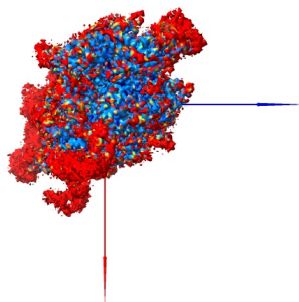
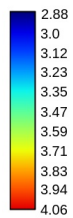


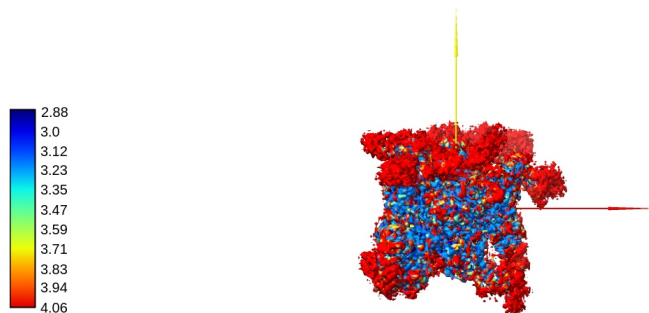
Figure 27: Histogram of the local resolution according to Resmap.



(a) View 1



(b) View 2



(c) View 3

Figure 28: Local resolution according to Resmap. Views generated by ChimeraX at a the following X, Y, Z angles: View 1 (0, -90, -90), View 2 (-90, 0, -90), View 3 (0, 0, 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.

STATUS: **OK**

4.5 Level 1.e Local resolution with MonoRes

Explanation:

MonoRes (see this [link](#) for more details) 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 significantly above the level of noise.

Results:

Fig. 29 shows the histogram of the local resolution according to MonoRes. Some representative percentiles are:

Percentile	Resolution($\text{\AA}$)
2.5%	2.54
25%	4.03
50%	6.68
75%	10.60
97.5%	14.39

The reported resolution, 3.13 $\text{\AA}$, is at the percentile 4.6. Fig. 30 shows some representative views of the local resolution

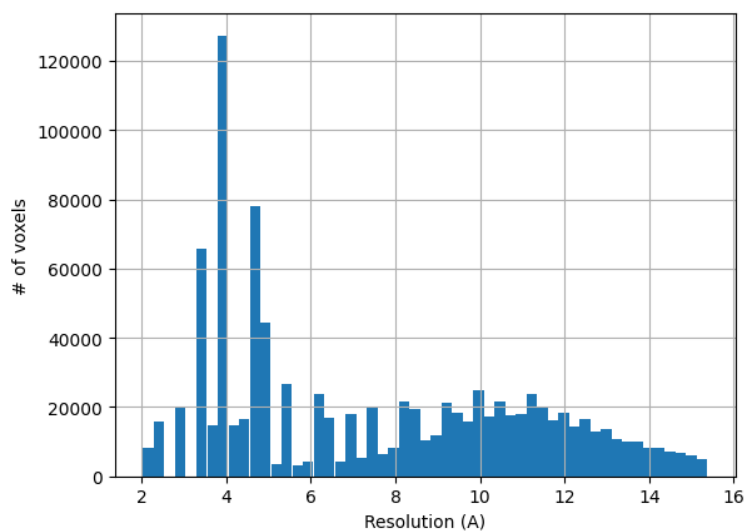
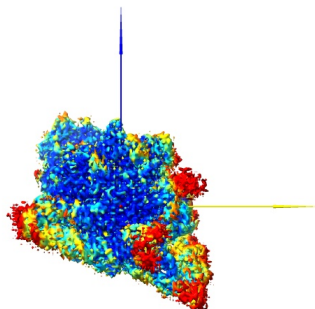
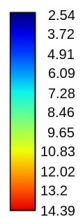
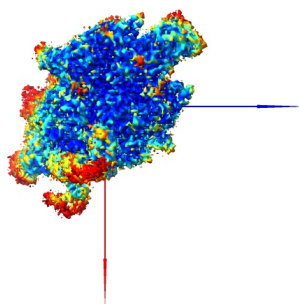
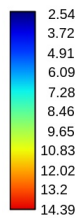


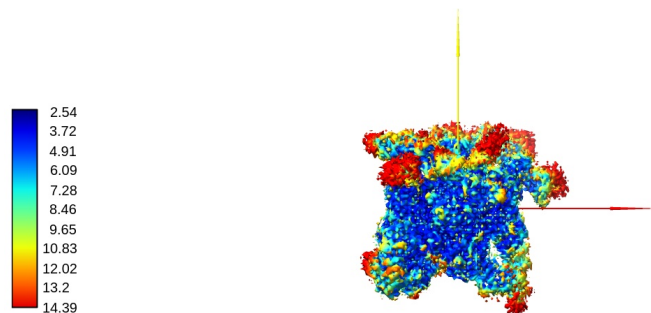
Figure 29: Histogram of the local resolution according to MonoRes.



(a) View 1



(b) View 2



(c) View 3

Figure 30: Local resolution according to MonoRes. Views generated by ChimeraX at a the following X, Y, Z angles: View 1 (0, -90, -90), View 2 (-90, 0, -90), View 3 (0, 0, 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 (see this [link](#) for more details) 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 di-

rectional 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 resolution 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:

ERROR: The protocol failed.

STATUS: Could not be measured

REASON: The map does not verify the requirements for MonoDir.

Possible cause:

1. **The protein is masked (lack of noise).**
2. **The protein has a large radius, thus, the algorithm cannot find a shell of noise.**

4.7 Level 1.g Fourier Shell Occupancy

Explanation:

This method (see this [link](#) for more details) 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 3.09Å. Fourier shells are occupied at between 90 and than 10% in the range [3.21, 2.99]Å.

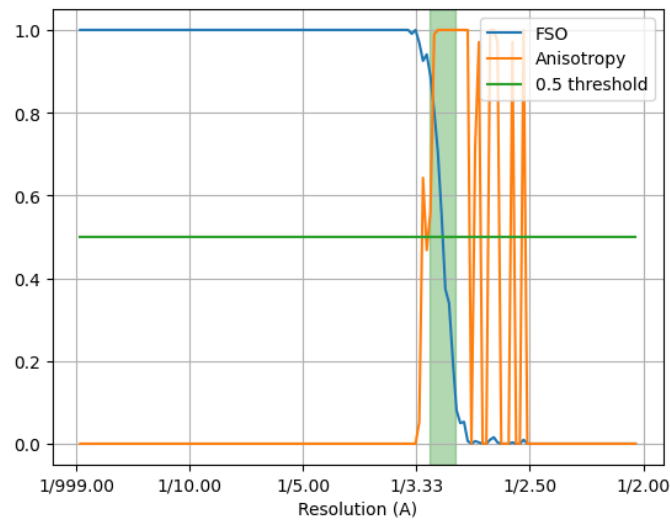


Figure 31: FSO and anisotropy.

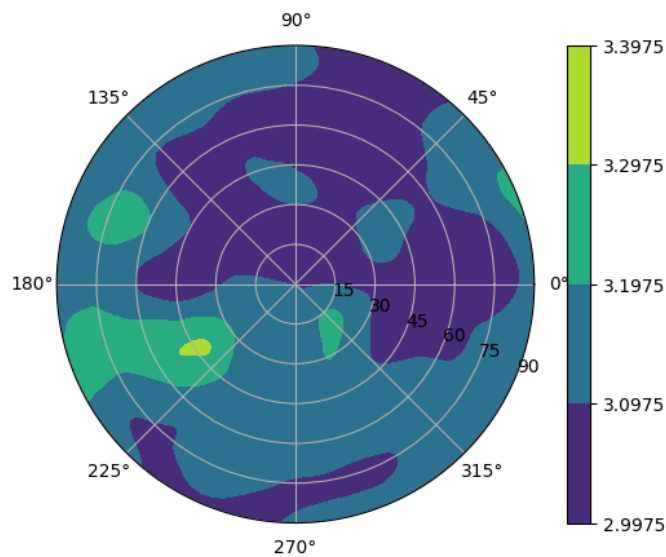


Figure 32: Directional resolution in the projection sphere.

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 (see this [link](#) for more details) analyzes the FSC in different directions and evaluates its homogeneity.

Results:

Fig. 33 shows the FSCs in X, Y, Z, and the global FSC. Fig. 34 shows the global FSC and the histogram of the directional FSC. Finally, Fig. 35 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 1.51, 1.50, and 1.49 Å, respectively. The global resolution at the same threshold is 1.51 Å. The resolution range is [1.49, 1.51]Å.

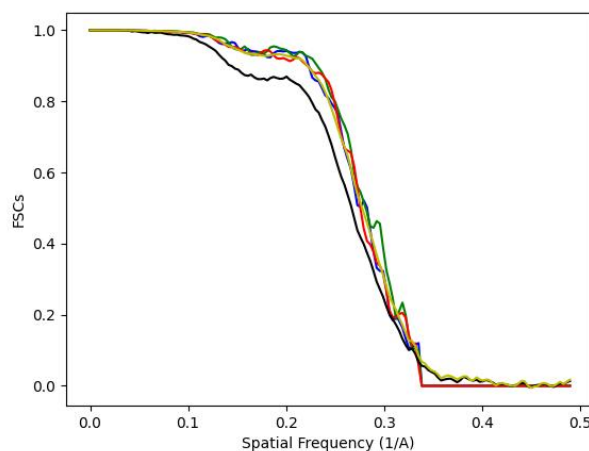


Figure 33: FSC in X, Y, Z, the global FSC, and the Average Cosine Phase.

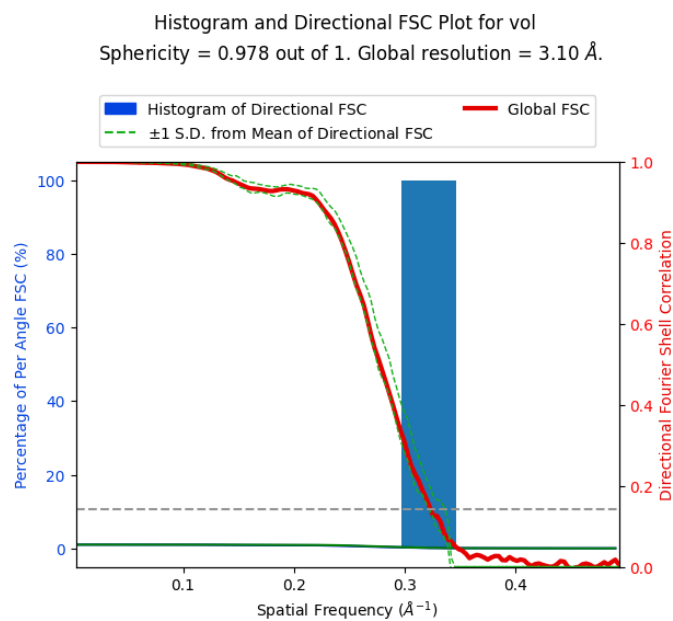


Figure 34: Global FSC and histogram of the directional FSC.

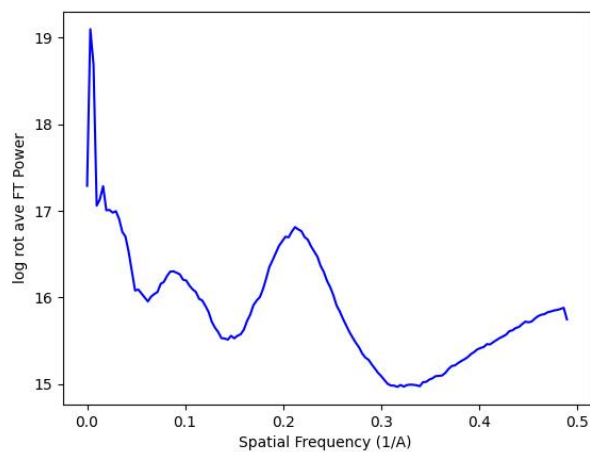


Figure 35: Logarithm of the radial average of the input map power in Fourier space.

Automatic criteria: The validation is OK if the resolution provided by

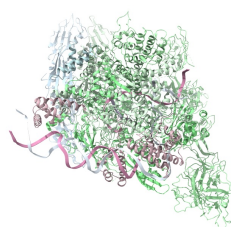
the user is not smaller than 0.8 the resolution estimated by the first cross of the global directional FSC below 0.143.

STATUS: [OK](#)

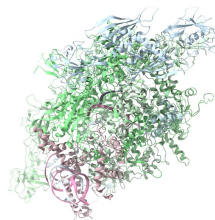
5 Atomic model

Atomic model: 8gzg.cif

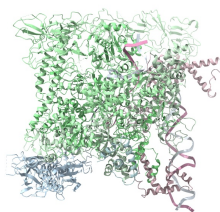
See Fig. [36](#).



(a) View 1



(b) View 2



(c) View 3

Figure 36: Input atomic model Views generated by ChimeraX at a the following X, Y, Z angles: View 1 (0, -90, -90), View 2 (-90, 0, -90), View 3 (0, 0, 0).

6 Level A Analysis

Map and model seem to be properly aligned.

6.1 Level A.a Phenix validation

Explanation:

Phenix provides a number of tools to assess the agreement between the experimental map and its atomic model (see this [link](#) for more details). 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 6.3 Å has been applied.

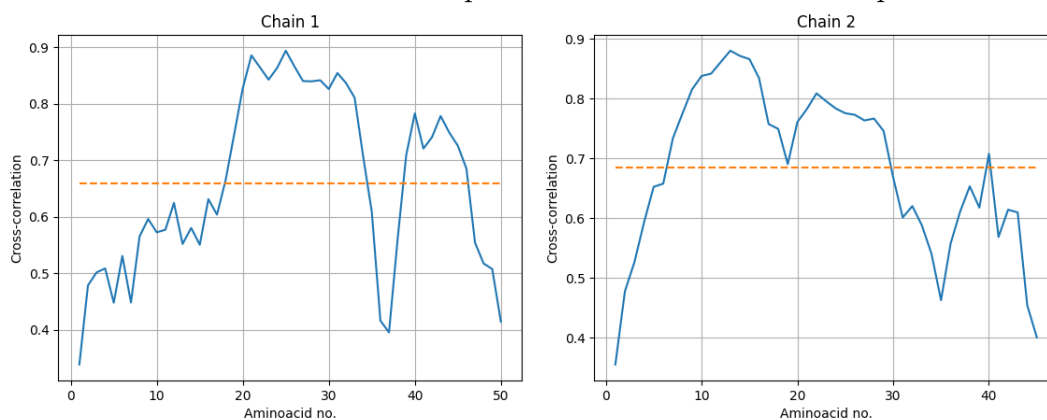
Overall correlation coefficients:

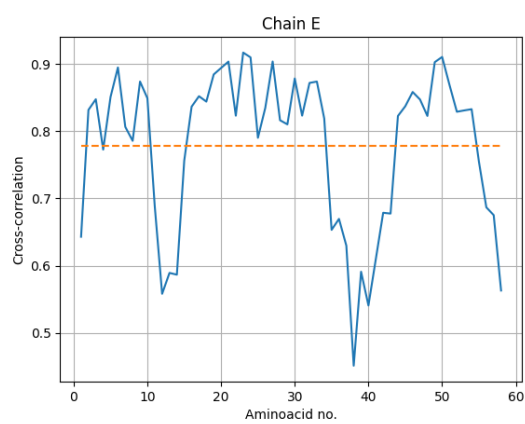
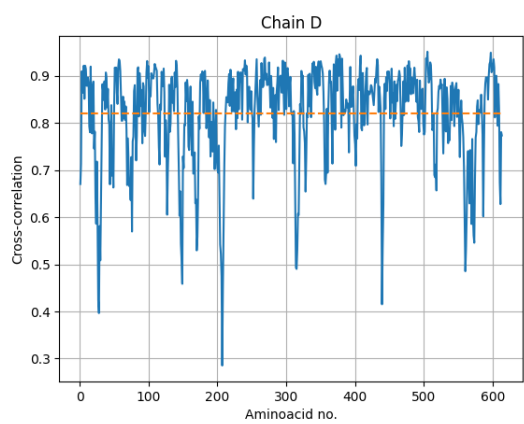
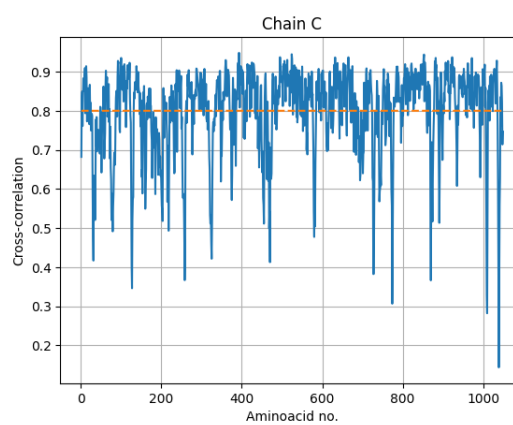
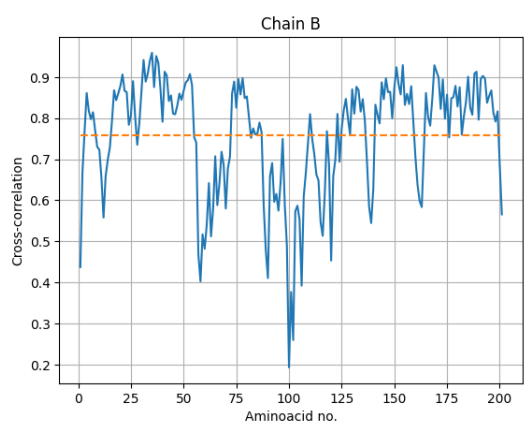
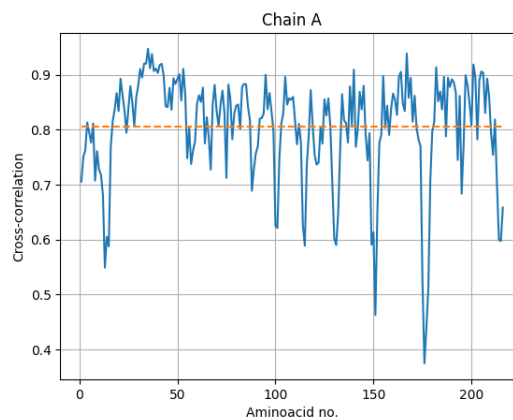
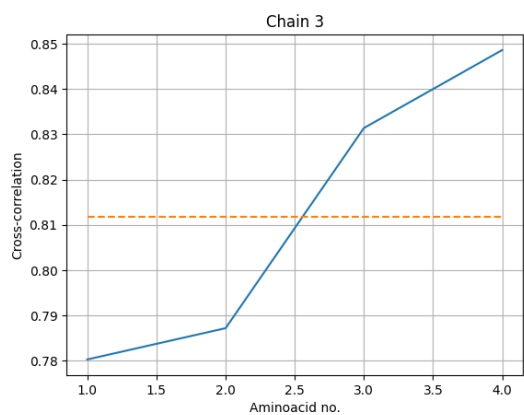
CC (mask) = 0.774
 CC (box) = 0.741
 CC (volume) = 0.752
 CC (peaks) = 0.695
 CC (main chain) = 0.777
 CC (side chain) = 0.762

Correlation coefficients per chain:

Chain	Cross-correlation
1	0.747760
2	0.754268
3	0.817903
A	0.778805
B	0.750599
C	0.775314
D	0.789713
E	0.752673
F	0.745486
Z	0.713823

We now show the correlation profiles of the different chain per residue.





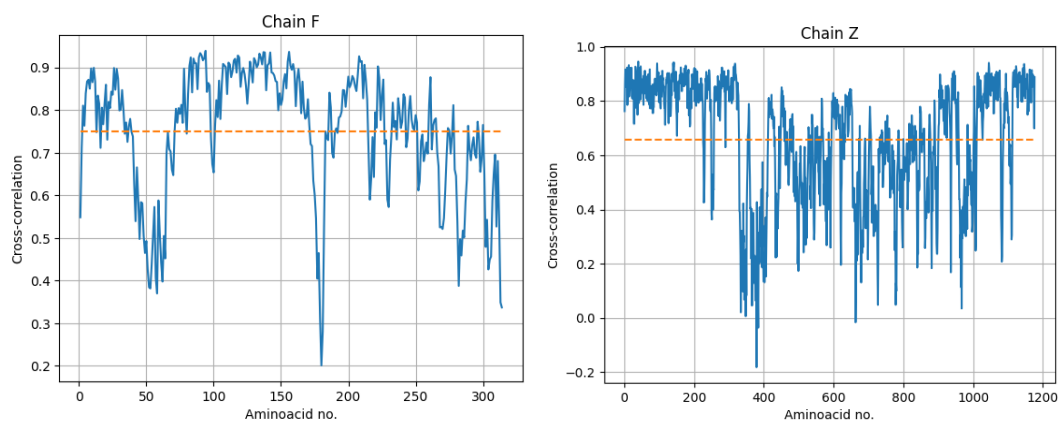


Fig. 37 shows the histogram of all cross-correlations evaluated at the residues. The percentage of residues whose correlation is below 0.5 is 10.2 %.

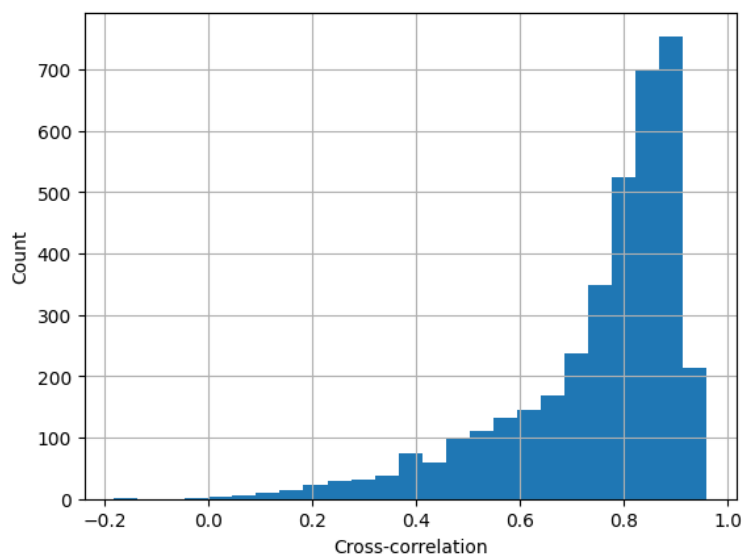


Figure 37: Histogram of the cross-correlation between the map and model evaluated for all residues.

Resolutions estimated from the model:

Resolution (Å)	Masked	Unmasked
d99	3.2	3.2
d_model	3.2	3.2
d_model (B-factor=0)	3.3	3.3
FSC_model=0	2.8	2.8
FSC_model=0.143	3.1	3.1
FSC_model=0.5	3.3	3.4

Overall isotropic B factor:

B factor	Masked	Unmasked
Overall B-iso	25.0	25.0

Fig. 38 shows the FSC between the input map and the model.

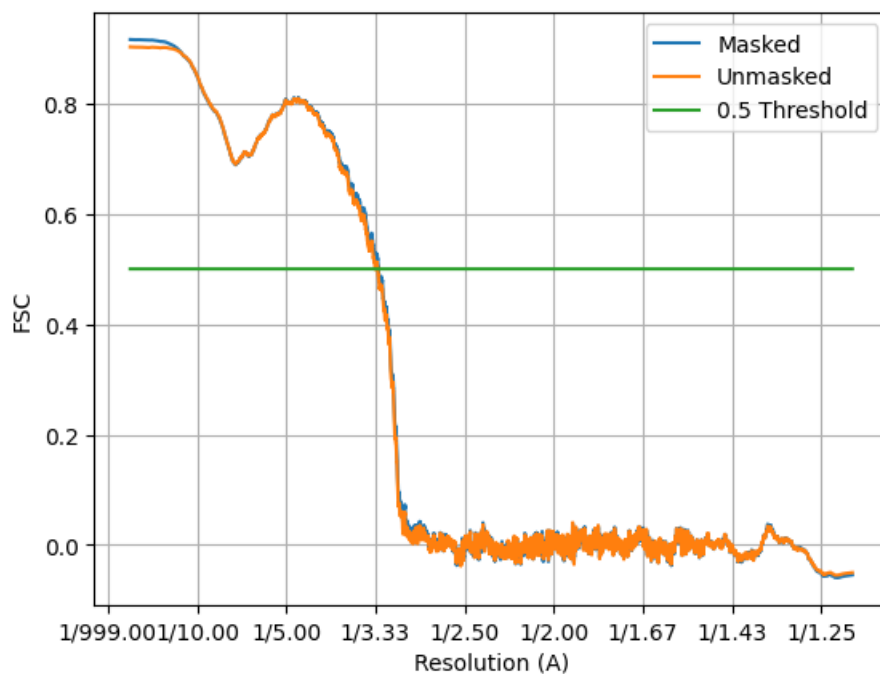


Figure 38: FSC between the input map and model with and without a mask constructed from the model. The X-axis is the square of the inverse of the resolution in Å.

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 10.2, that is larger than 10%**

6.2 Level A.b FSC-Q

Explanation:

FSC-Q (see this [link](#) for more details) 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. 39 shows the histogram of FSC-Qr and Fig. 40 the colored isosurface of the atomic model converted to map. The average FSC-Qr is 0.10, its 95% confidence interval is [-0.07, 0.80]. The percentage of values whose FSC-Qr absolute value is beyond 1.5 is 0.3 %.

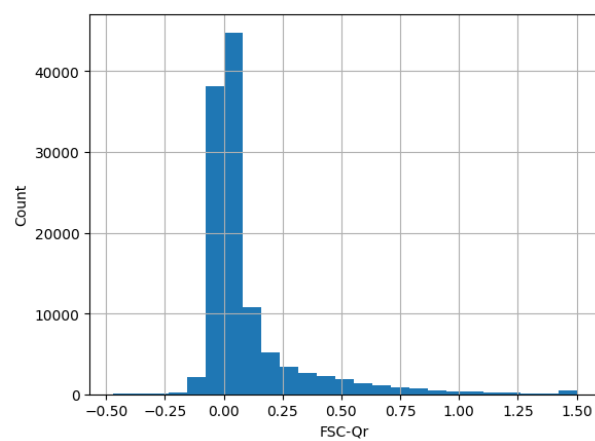
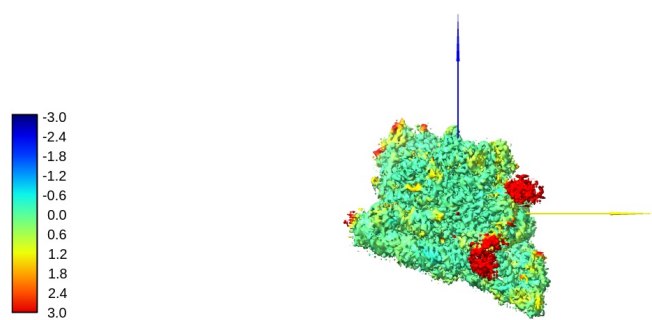
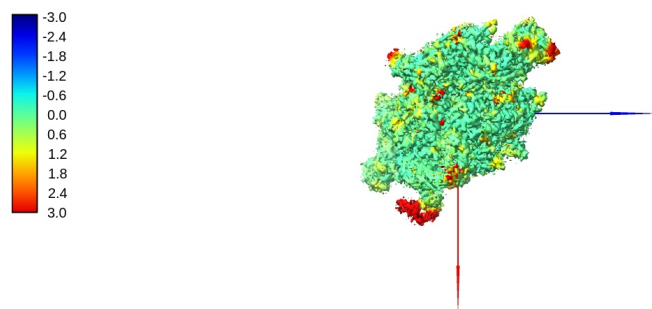


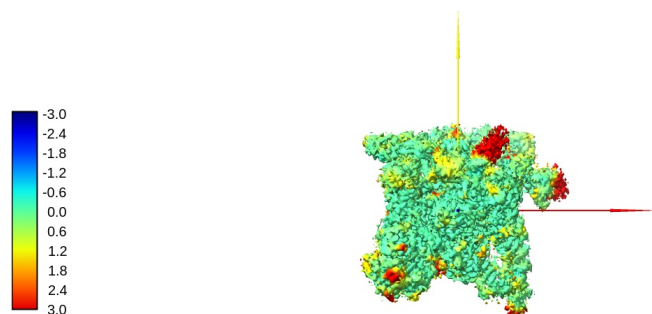
Figure 39: Histogram of the FSC-Qr limited to -1.5 and 1.5.



(a) View 1



(b) View 2



(c) View 3

Figure 40: 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, -90, -90), View 2 (-90, 0, -90), View 3 (0, 0, 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

6.3 Level A.d Map-Model Guinier analysis

Explanation:

We compared the Guinier plot (see this [link](#) for more details) 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. 41 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 0.929.

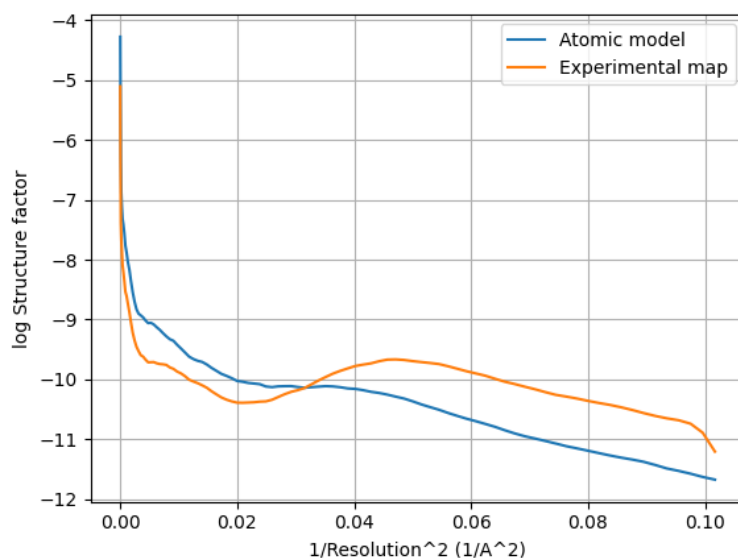


Figure 41: Guinier plot of the atom model and experimental map. The X-axis is the square of the inverse of the resolution in Å.

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 MapQ

Explanation:

MapQ (see this [link](#) for more details) computes the local correlation between the map and each one of its atoms assumed to have a Gaussian shape.

Results:

Fig. 42 shows the histogram of the calculated Q-score. Some representative percentiles are:

Percentile	MapQ score [-1,1]
2.5%	0.00
25%	0.00
50%	0.00
75%	0.00
97.5%	0.00

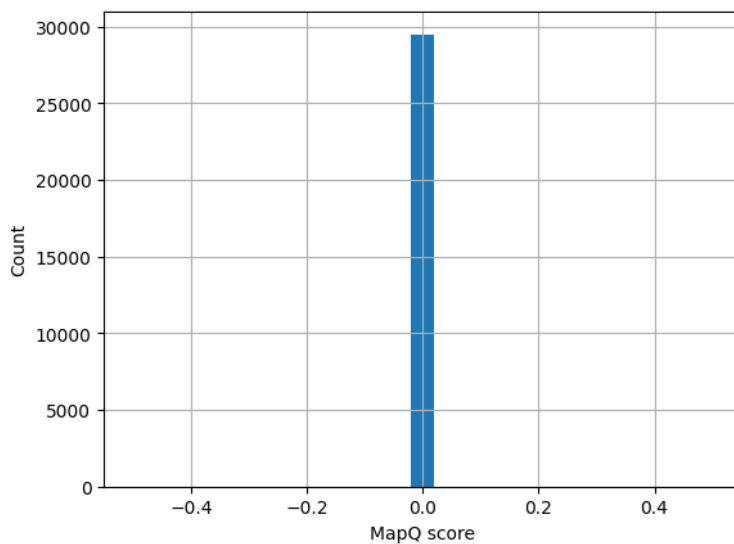
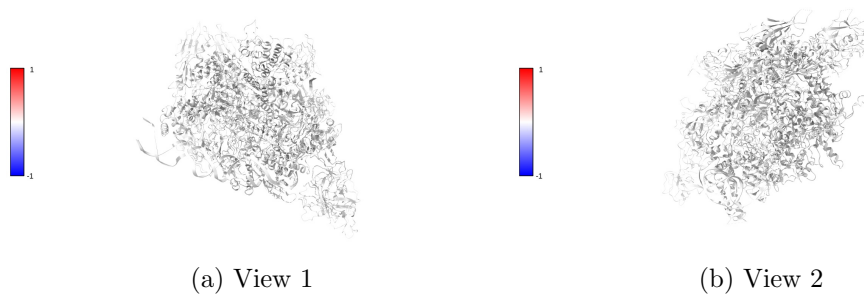


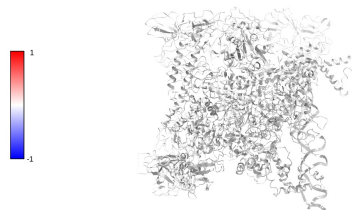
Figure 42: Histogram of the per-atom Q-score.

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

Chain	Average Q-score [-1,-1]	Estimated Resol. (Å)
1	0.00	7.2
2	0.00	7.2
3	0.00	7.2
A	0.00	6.3
B	0.00	6.3
C	0.00	6.3
D	0.00	9.8
D	0.00	6.3
E	0.00	6.3
F	0.00	6.3
Z	0.00	9.8
Z	0.00	6.3

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





(c) View 3

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

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

WARNINGS: 2 warnings

1. **Program output seems to be too homogeneous. There might be some program issues analyzing the data.**
2. **The median Q-score is less than 0.1.**

6.5 Level A.f EMRinger validation

Explanation:

EMRinger (see this [link](#) for more details) 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	0.028484
Rotamer ratio	0.700
Max. Zscore	10.63
Model length	1941
EMRinger Score	2.414

Fig. 44 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.

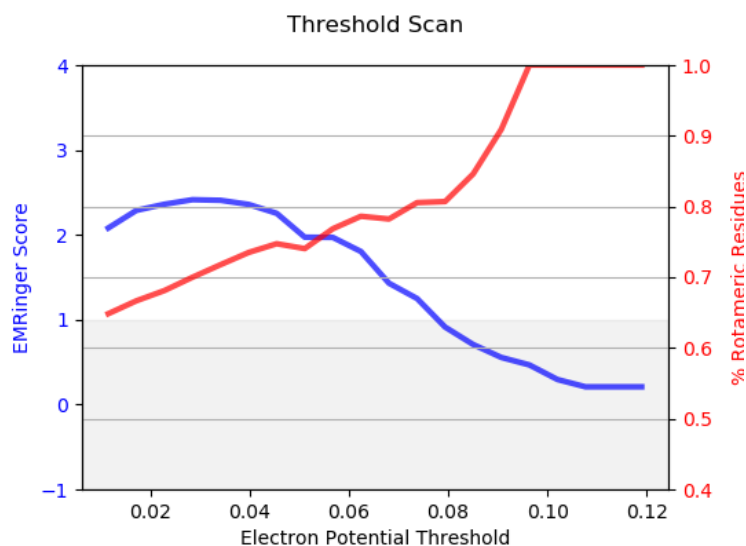


Figure 44: EMRinger score and fraction of rotameric residues as a function of the map threshold.

Fig. 45 shows the histogram for rotameric (blue) and non-rotameric (red) residues at the optimal threshold.

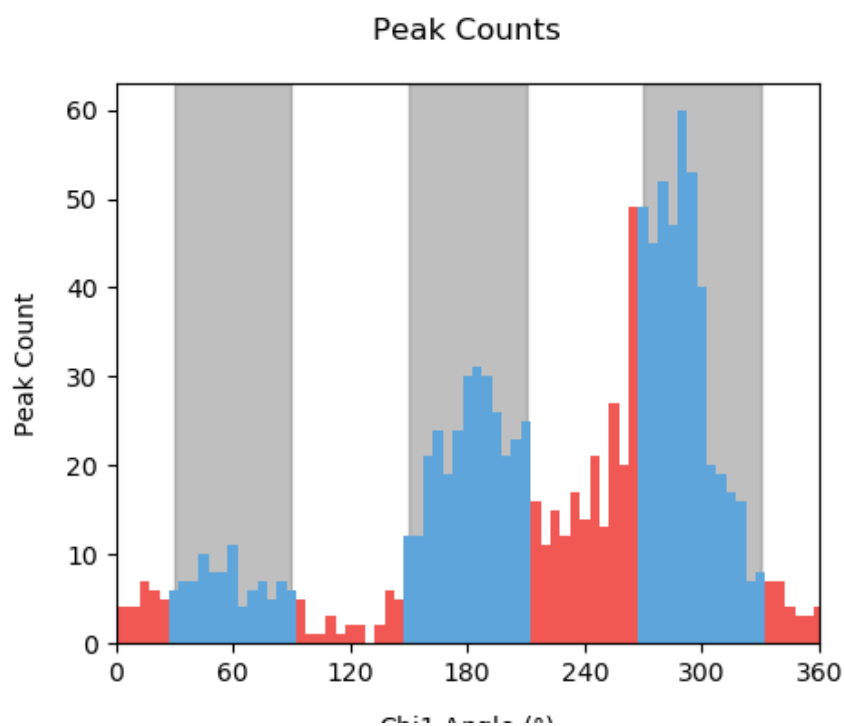
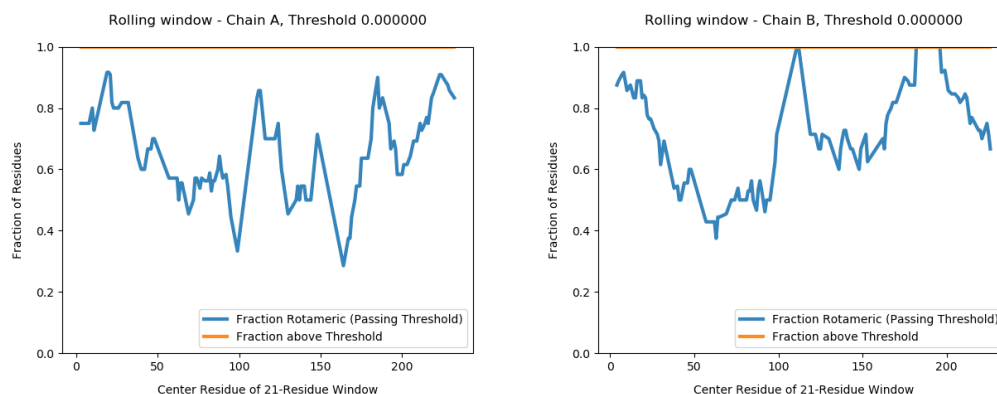
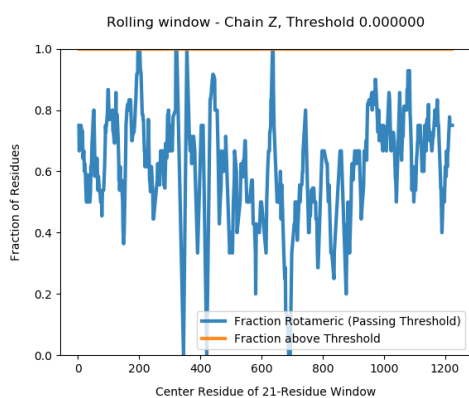
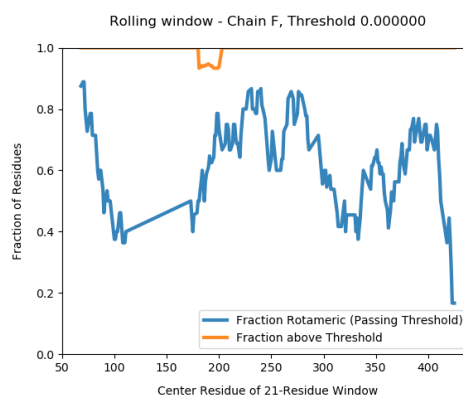
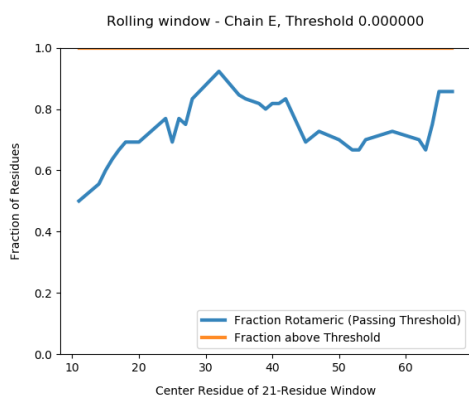
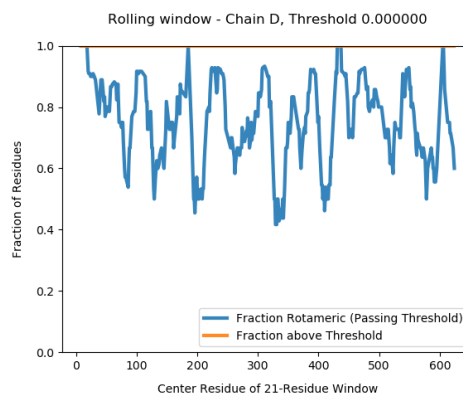
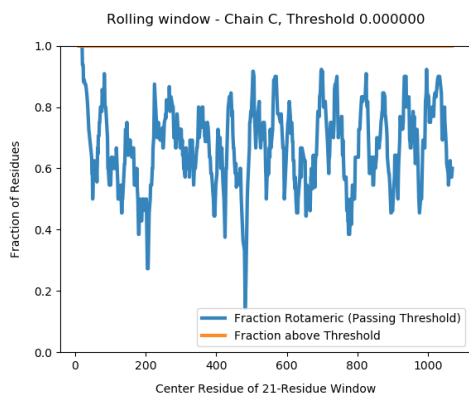


Figure 45: Histogram for rotameric (blue) and non-rotameric (red) residues at the optimal threshold as a function of the angle Chi1.

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.





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 (see this [link](#) for more details) 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:

DAQ scores obtained from DB source via 3DBionotes-WS:

<https://3dbionotes.cnb.csic.es/bws/api/emv/emd-34397/daq/>

Fig. 46 shows the histogram of the DAQ values. The mean and standard deviation were 0.7 and 0.4, respectively.

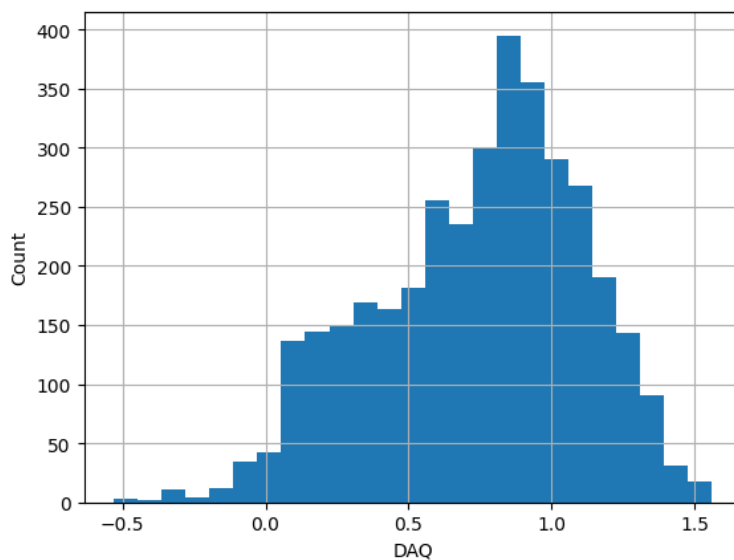


Figure 46: Histogram of the per-residue DAQ values.

The atomic model colored by DAQ can be seen in Fig. 47.

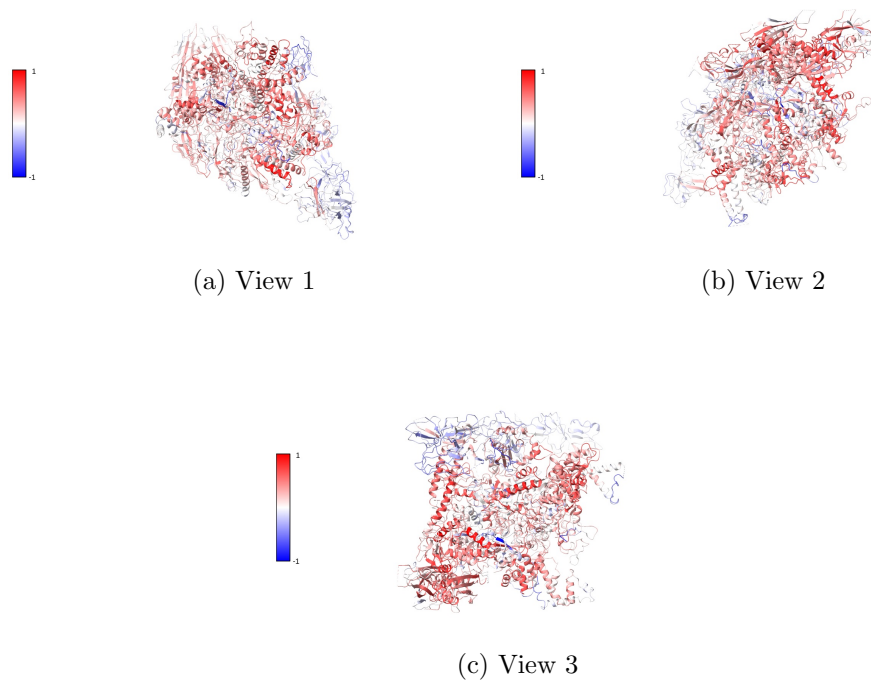


Figure 47: Atomic model colored by DAQ Views generated by ChimeraX at the following X, Y, Z angles: View 1 (0, -90, -90), View 2 (-90, 0, -90), View 3 (0, 0, 0).

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

STATUS: OK