



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-34412](#)

**PDB ID:** [8h0i](#)

**Title:** Cryo-EM structure of APOBEC3G-Vif complex

**Authors:** [See EMDB entry link](#)

**Deposited on:** 2022-09-29T00:00:00

**Reported Resolution:** 2.8 Å

---

### Contact Us:

Instruct Image Processing Center ([I<sup>2</sup>PC](#))  
Biocomputing Unit ([BCU](#))  
[i2pc@cnb.csic.es](mailto: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, 3:48am**

## 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 have some problem in its centering or extra space (see Sec. 2.1). 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). It seems that the Guinier plot of the map and its model do not match (see Sec. 6.3). 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 1.9Å to 7.5Å with an average of 3.4Å. The resolution reported by the user was 2.8Å.

**The overall score (passing tests) of this report is 16 out of 22 evaluable items.**

|                         |                          |                            |
|-------------------------|--------------------------|----------------------------|
| 0.a Mass analysis       | Sec. <a href="#">2.1</a> | <a href="#">3 warnings</a> |
| 0.b Mask analysis       | Sec. <a href="#">2.2</a> | <a href="#">OK</a>         |
| 0.c Background analysis | Sec. <a href="#">2.3</a> | <a href="#">2 warnings</a> |
| 0.d B-factor analysis   | Sec. <a href="#">2.4</a> | <a href="#">OK</a>         |
| 0.e DeepRes             | Sec. <a href="#">2.5</a> | <a href="#">1 warnings</a> |
| 0.f LocBfactor          | Sec. <a href="#">2.6</a> | <a href="#">OK</a>         |
| 0.g LocOccupancy        | Sec. <a href="#">2.7</a> | <a href="#">OK</a>         |
| 0.h DeepHand            | Sec. <a href="#">2.8</a> | <a href="#">OK</a>         |
| 1.a Global resolution   | Sec. <a href="#">4.1</a> | <a href="#">OK</a>         |
| 1.b FSC permutation     | Sec. <a href="#">4.2</a> | <a href="#">OK</a>         |
| 1.c Blocres             | Sec. <a href="#">4.3</a> | <a href="#">OK</a>         |
| 1.d Resmap              | Sec. <a href="#">4.4</a> | <a href="#">OK</a>         |
| 1.e MonoRes             | Sec. <a href="#">4.5</a> | <a href="#">OK</a>         |
| 1.f MonoDir             | Sec. <a href="#">4.6</a> | <a href="#">1 warnings</a> |
| 1.g FSO                 | Sec. <a href="#">4.7</a> | <a href="#">OK</a>         |
| 1.h FSC3D               | Sec. <a href="#">4.8</a> | <a href="#">OK</a>         |
| A.a Phenix validation   | Sec. <a href="#">6.1</a> | <a href="#">OK</a>         |
| A.b FSC-Q               | Sec. <a href="#">6.2</a> | <a href="#">OK</a>         |
| A.d Map-Model Guinier   | Sec. <a href="#">6.3</a> | <a href="#">1 warnings</a> |
| A.e MapQ                | Sec. <a href="#">6.4</a> | <a href="#">1 warnings</a> |
| A.f EMRinger            | Sec. <a href="#">6.5</a> | <a href="#">OK</a>         |
| A.g DAQ                 | Sec. <a href="#">6.6</a> | <a href="#">OK</a>         |

## Summary of the warnings across sections.

### Section 2.1 (0.a Mass analysis)

1. **The center of mass in X may be significantly shifted. This is common when the refinement is applied exclusively to one protein region.**
2. **The center of mass in Y may be significantly shifted. This is common when the refinement is applied exclusively to one protein region.**
3. **The center of mass in Z may be significantly shifted. This is common when the refinement is applied exclusively to one protein region.**

### 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=3131.06)**

### Section 2.5 (0.e DeepRes)

1. **The reported resolution, 2.80 Å, is particularly high with respect to the local resolution distribution. It occupies the 0.02 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 6.3 (A.d Map-Model Guinier)

1. **The correlation is smaller than 0.5, it is 0.420.**

### Section 6.4 (A.e MapQ)

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

# Contents

|          |                                                                   |           |
|----------|-------------------------------------------------------------------|-----------|
| <b>1</b> | <b>Input data</b>                                                 | <b>6</b>  |
| <b>2</b> | <b>Level 0 analysis</b>                                           | <b>13</b> |
| 2.1      | Level 0.a Mass analysis . . . . .                                 | 13        |
| 2.2      | Level 0.b Mask analysis . . . . .                                 | 14        |
| 2.3      | Level 0.c Background analysis . . . . .                           | 18        |
| 2.4      | Level 0.d B-factor analysis . . . . .                             | 20        |
| 2.5      | Level 0.e Local resolution with DeepRes . . . . .                 | 23        |
| 2.6      | Level 0.f Local B-factor . . . . .                                | 26        |
| 2.7      | Level 0.g Local Occupancy . . . . .                               | 29        |
| 2.8      | Level 0.h Hand correction . . . . .                               | 32        |
| <b>3</b> | <b>Half maps</b>                                                  | <b>32</b> |
| <b>4</b> | <b>Level 1 analysis</b>                                           | <b>36</b> |
| 4.1      | Level 1.a Global resolution . . . . .                             | 36        |
| 4.2      | Level 1.b FSC permutation . . . . .                               | 38        |
| 4.3      | Level 1.c Local resolution with Blocres . . . . .                 | 39        |
| 4.4      | Level 1.d Local resolution with Resmap . . . . .                  | 42        |
| 4.5      | Level 1.e Local resolution with MonoRes . . . . .                 | 45        |
| 4.6      | Level 1.f Local and directional resolution with MonoDir . . . . . | 48        |
| 4.7      | Level 1.g Fourier Shell Occupancy . . . . .                       | 51        |
| 4.8      | Level 1.h Fourier Shell Correlation 3D . . . . .                  | 53        |
| <b>5</b> | <b>Atomic model</b>                                               | <b>55</b> |
| <b>6</b> | <b>Level A Analysis</b>                                           | <b>56</b> |
| 6.1      | Level A.a Phenix validation . . . . .                             | 56        |
| 6.2      | Level A.b FSC-Q . . . . .                                         | 62        |
| 6.3      | Level A.d Map-Model Guinier analysis . . . . .                    | 64        |
| 6.4      | Level A.e MapQ . . . . .                                          | 65        |
| 6.5      | Level A.f EMRinger validation . . . . .                           | 68        |
| 6.6      | Level A.g DAQ validation . . . . .                                | 72        |

# 1 Input data

Input map: emd\_34412.map

SHA256 hash: 7f52f056002cc5ac34d9d4698a067b7377d95c945fc6c5364f91d55fc59b3c05

Voxel size: 0.840000 (Å)

Visualization threshold: 4.000000

Resolution estimated by user: 2.8

## Orthogonal slices of the input map

### **Explanation:**

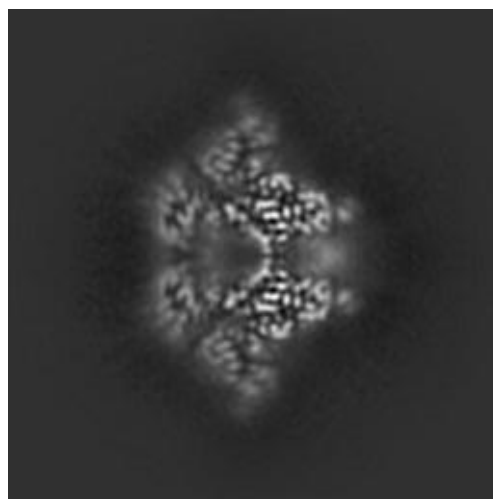
In the orthogonal slices of the map, the noise outside the protein should not have any structure (stripes going out, small blobs, particularly high or low densities, ...)

### **Results:**

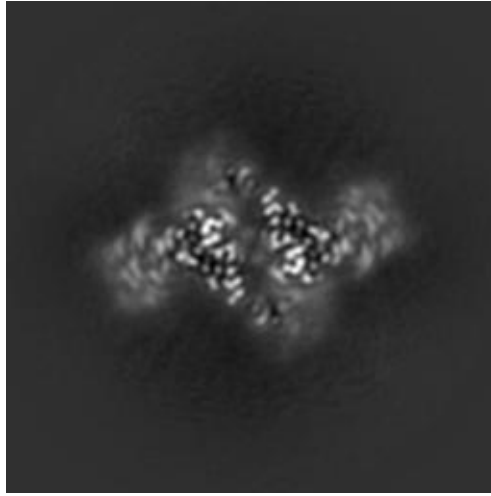
See Fig. 1.



(a) X Slice 128



(b) Y Slice 128



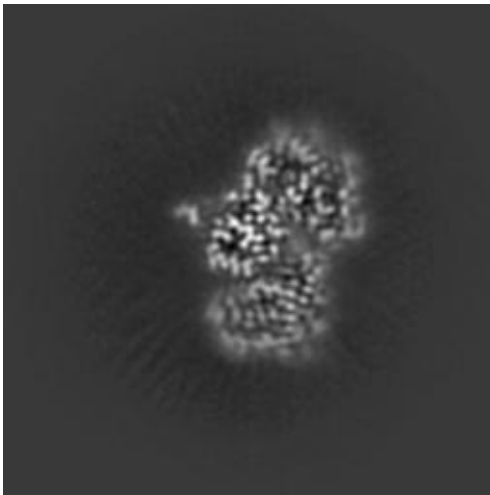
(c) Z Slice 128

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

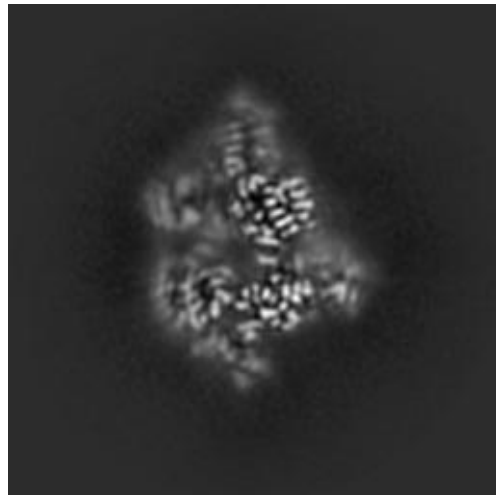
Orthogonal slices of maximum variance of the input map

**Results:**

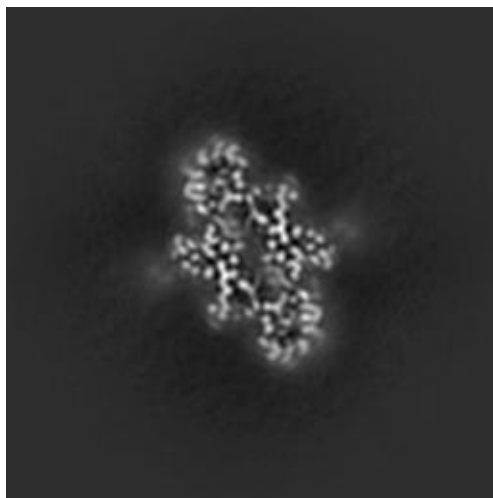
See Fig. 2.



(a) X Slice 107



(b) Y Slice 122



(c) Z Slice 145

Figure 2: Slices of maximum variation in the three dimensions

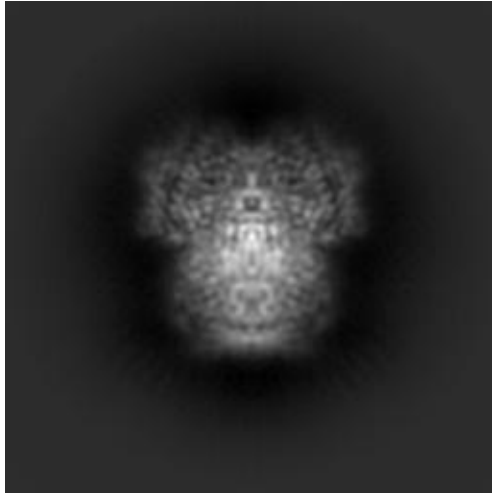
### Orthogonal projections of the input map

#### **Explanation:**

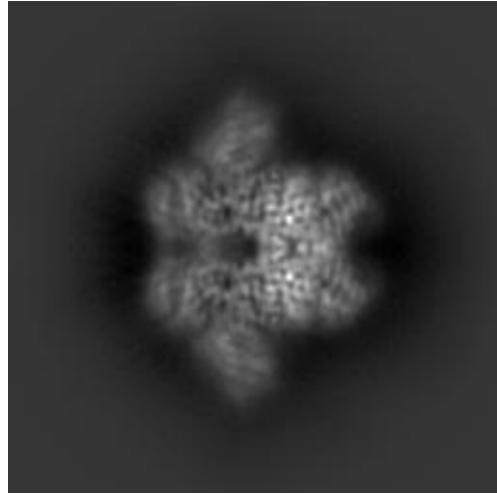
In the projections there should not be stripes (this is an indication of directional overweighting, or angular attraction), and there should not be a dark halo around or inside the structure (this is an indication of incorrect CTF correction or the reconstruction of a biased map).

#### **Results:**

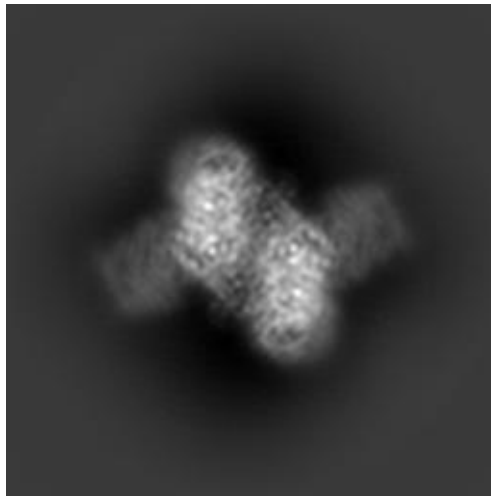
See Fig. [3](#).



(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.

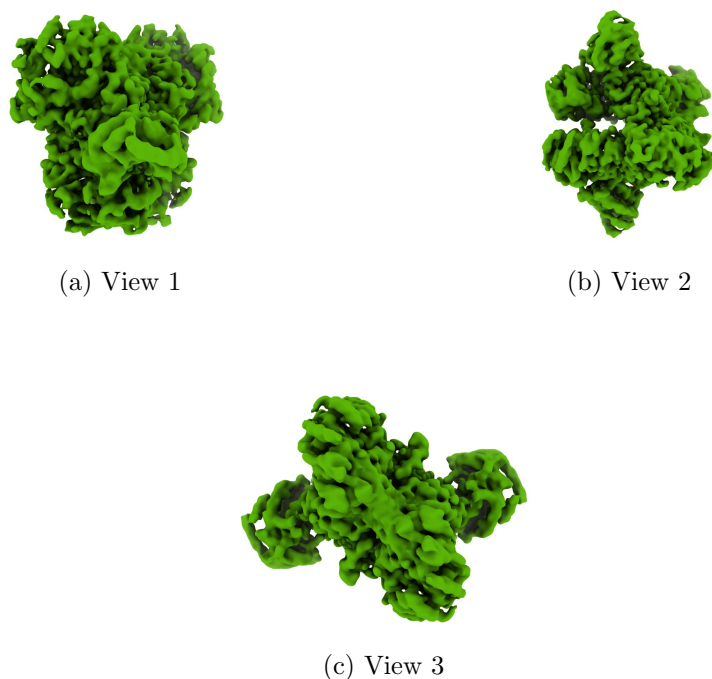


Figure 4: Isosurface at threshold=4.000000. Views generated by ChimeraX at a the following X, Y, Z angles: View 1 (0, -90, -90), View 2 (-90, 0, -90), View 3 (0, 0, 0).

#### Orthogonal slices of maximum variance of the mask with hard borders

##### **Explanation:**

The mask with hard borders has been calculated at the suggested threshold 4.000000, the largest connected component was selected, and then dilated by 2Å.

**Results:**  
See Fig. 5.

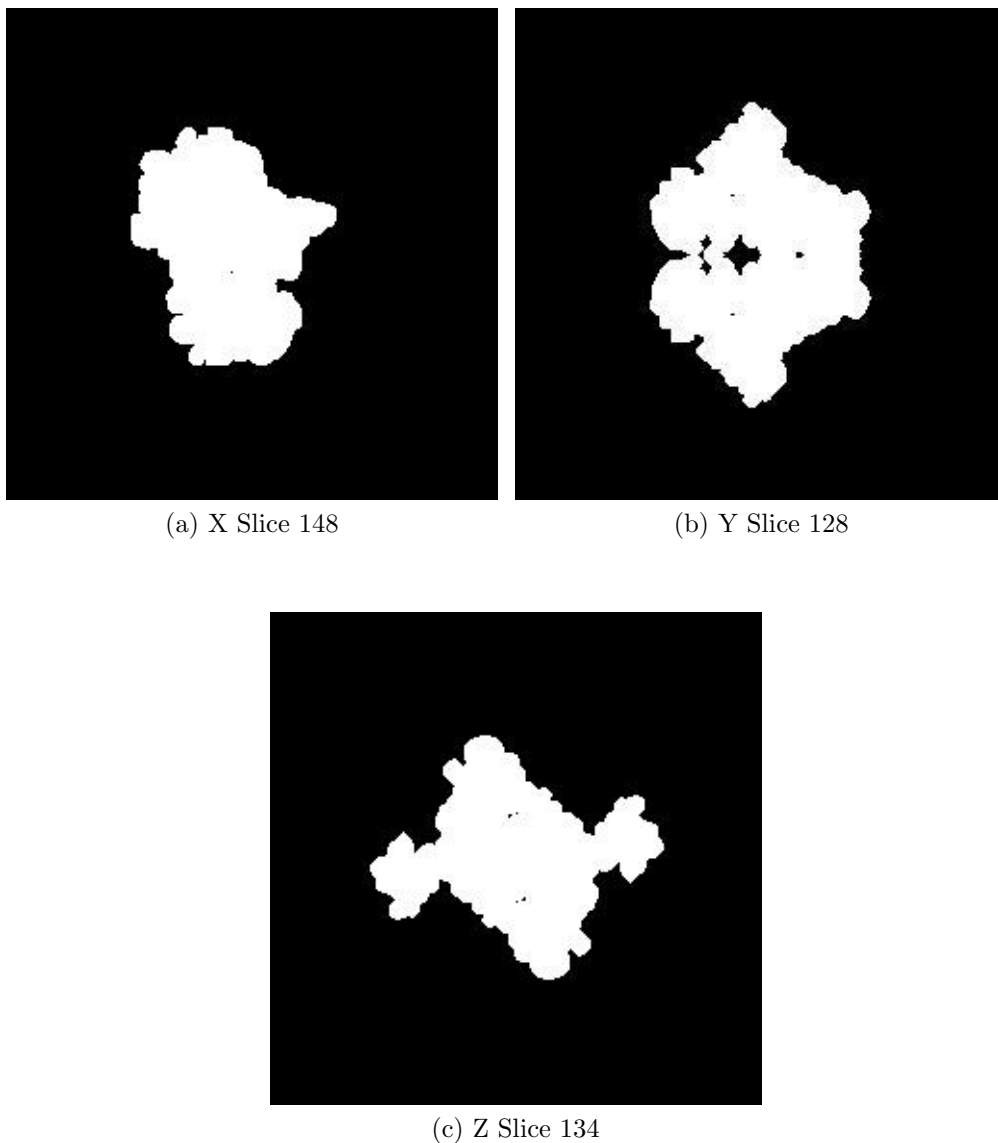


Figure 5: Slices of maximum variation in the three dimensions of the mask with hard borders

#### Orthogonal slices of maximum variance of the mask with soft borders

##### **Explanation:**

The mask with soft borders has been calculated at the suggested threshold 4.000000, the largest connected component was selected, and then dilated by

2Å.

**Results:**  
See Fig. 6.



(a) X Slice 148



(b) Y Slice 128



(c) Z Slice 134

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 37.80 and 37.80  $\text{\AA}$ , respectively. There is a decentering ratio  $(\text{abs}(\text{Right-Left})/\text{Size})\%$  of 0.00%

The space from the left and right in Y are 54.60 and 54.60  $\text{\AA}$ , respectively. There is a decentering ratio  $(\text{abs}(\text{Right-Left})/\text{Size})\%$  of 0.00%

The space from the left and right in Z are 57.96 and 52.92  $\text{\AA}$ , respectively. There is a decentering ratio  $(\text{abs}(\text{Right-Left})/\text{Size})\%$  of 2.34%

The center of mass is at (x,y,z)=(-860754.01,-927314.32,-70522227.18). The decentering of the center of mass (abs(Center)/Size)% is 336282.04, 362282.16, and 27547794.99, 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.

**WARNINGS:** 3 warnings

1. **The center of mass in X may be significantly shifted. This is common when the refinement is applied exclusively to one protein region.**
2. **The center of mass in Y may be significantly shifted. This is common when the refinement is applied exclusively to one protein region.**
3. **The center of mass in Z may be significantly shifted. This is common when the refinement is applied exclusively to one protein region.**

## 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, their 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 4.000000, there are 9 connected components with

a total number of voxels of 234256 and a volume of  $138844.47 \text{ \AA}^3$  (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):

(234174 (138795.87), 99.96, 99.96)

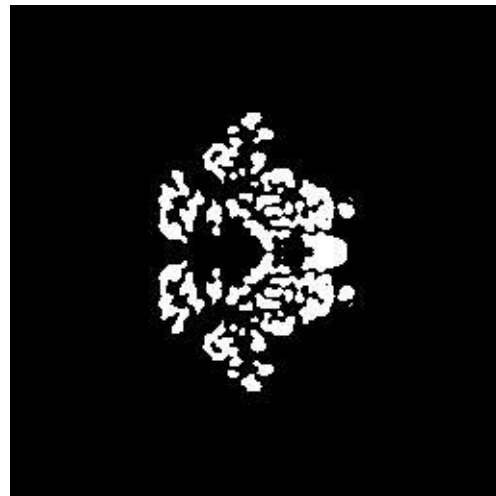
Number of components to reach 95% of the mass: 1

The average size of the remaining 8 components is 10.25 voxels (  $0.59 \text{ \AA}^3$ ). Their size go from 234174 voxels ( $138795.87 \text{ \AA}^3$ ) to 6 voxels (  $3.56 \text{ \AA}^3$ ).

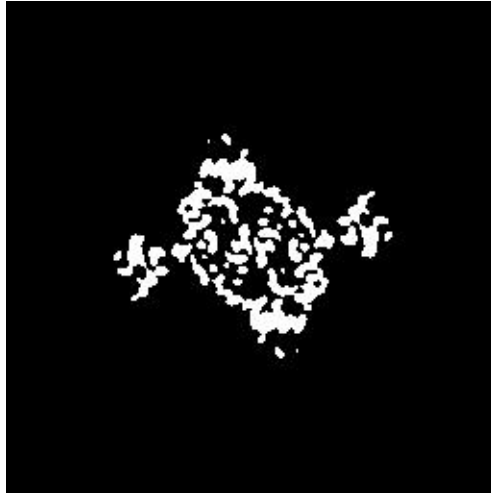
The slices of the raw mask can be seen in Fig. 7.



(a) X Slice 106



(b) Y Slice 128



(c) Z Slice 133

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 |
|-----------|------------|---------------------|--------------|
| 0.8316    | 717526.00  | 352.35              | 3203.14      |
| 1.6632    | 542235.00  | 266.27              | 2420.61      |
| 2.4947    | 402595.00  | 197.70              | 1797.24      |
| 3.3263    | 297942.00  | 146.31              | 1330.06      |
| 4.1579    | 221233.00  | 108.64              | 987.62       |
| 4.9895    | 163703.00  | 80.39               | 730.79       |
| 5.8211    | 120587.00  | 59.21               | 538.32       |
| 6.6526    | 89688.00   | 44.04               | 400.38       |
| 7.4842    | 67285.00   | 33.04               | 300.37       |
| 8.3158    | 50845.00   | 24.97               | 226.98       |
| 9.1474    | 38344.00   | 18.83               | 171.17       |
| 9.9789    | 28376.00   | 13.93               | 126.67       |
| 10.8105   | 20720.00   | 10.17               | 92.50        |
| 11.6421   | 14502.00   | 7.12                | 64.74        |
| 12.4737   | 9694.00    | 4.76                | 43.28        |
| 13.3053   | 6342.00    | 3.11                | 28.31        |
| 14.1368   | 3980.00    | 1.95                | 17.77        |
| 14.9684   | 2254.00    | 1.11                | 10.06        |
| 15.8000   | 1244.00    | 0.61                | 5.55         |
| 16.6316   | 638.00     | 0.31                | 2.85         |
| 17.4632   | 294.00     | 0.14                | 1.31         |
| 18.2947   | 100.00     | 0.05                | 0.45         |
| 19.1263   | 30.00      | 0.01                | 0.13         |
| 19.9579   | 8.00       | 0.00                | 0.04         |

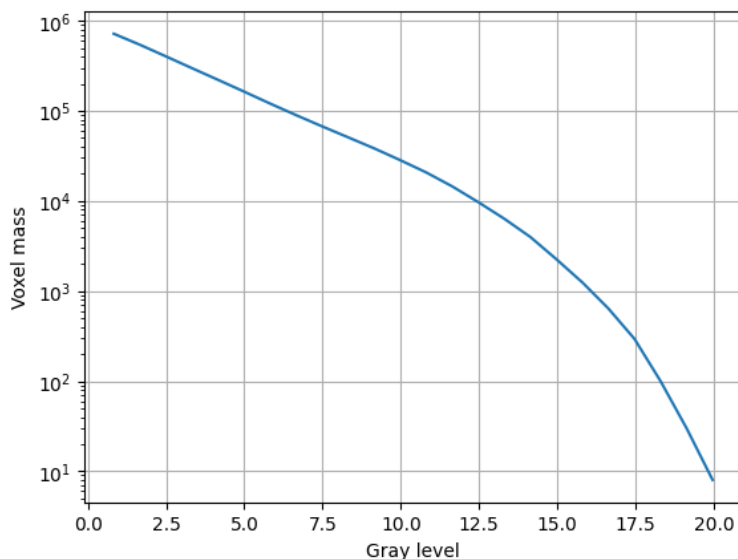


Figure 8: Voxel mass as a function of the gray level.

**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 684590 and a volume of  $405759.23 \text{ \AA}^3$ . 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  $5\text{\AA}^3$ ; 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

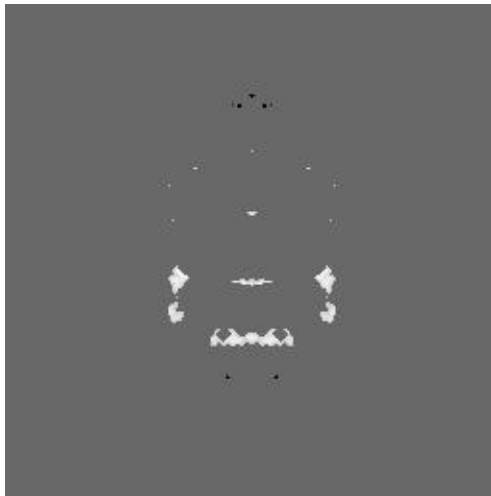
small and they should be randomly distributed without any specific structure. Sometimes, you can see some structure due to the symmetry of the structure.

### Results:

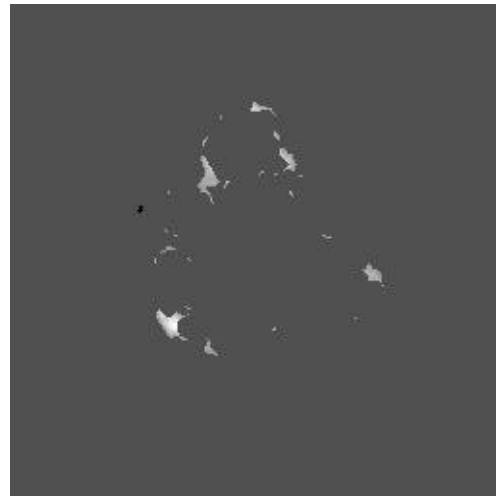
The null hypothesis that the background mean is 0 was tested with a one-sample Student's t-test. The resulting t-statistic and p-value were -1806.90 and 0.000000, respectively.

The mean and standard deviation ( $\sigma$ ) of the background were -0.152762 and 0.339152. The percentage of background voxels whose absolute value is larger than 5 times the standard deviation is 0.18 % (see Fig. 9). The same percentage from a Gaussian would be 0.000057% (ratio between the two percentages: 3131.058630).

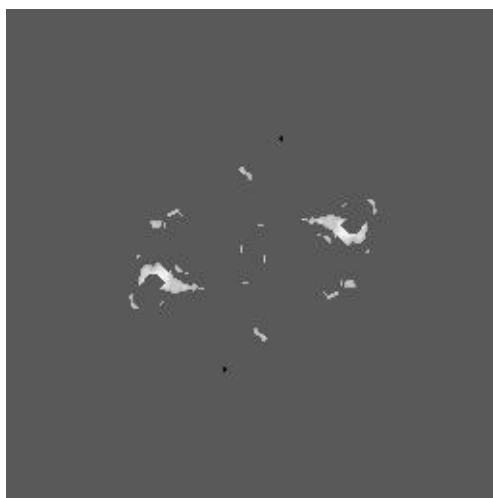
Slices of the background beyond  $5\sigma$  can be seen in Fig. 9.



(a) X Slice 128



(b) Y Slice 112



(c) Z Slice 102

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=3131.06)**

## 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 Å<sup>2</sup>.

### 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 -152.5. The fitted line was  $\log(|F|^2) = -38.1/R^2 + (-6.3)$ .

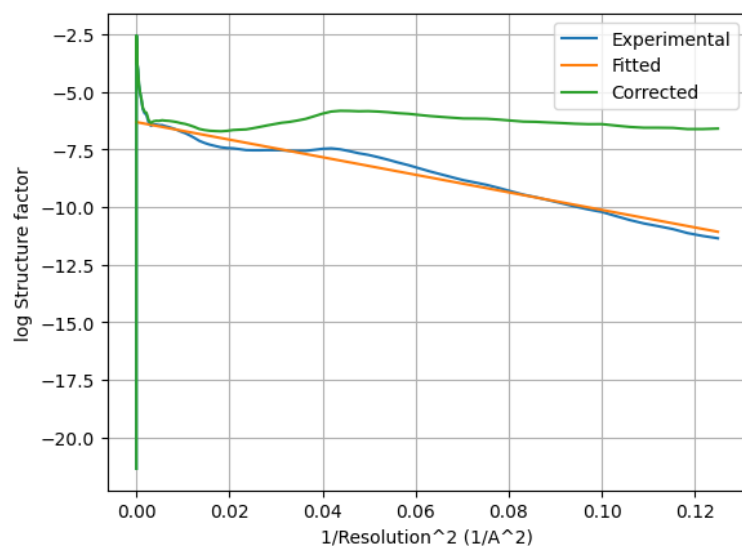


Figure 10: Guinier plot. The X-axis is the square of the inverse of the resolution in Å.

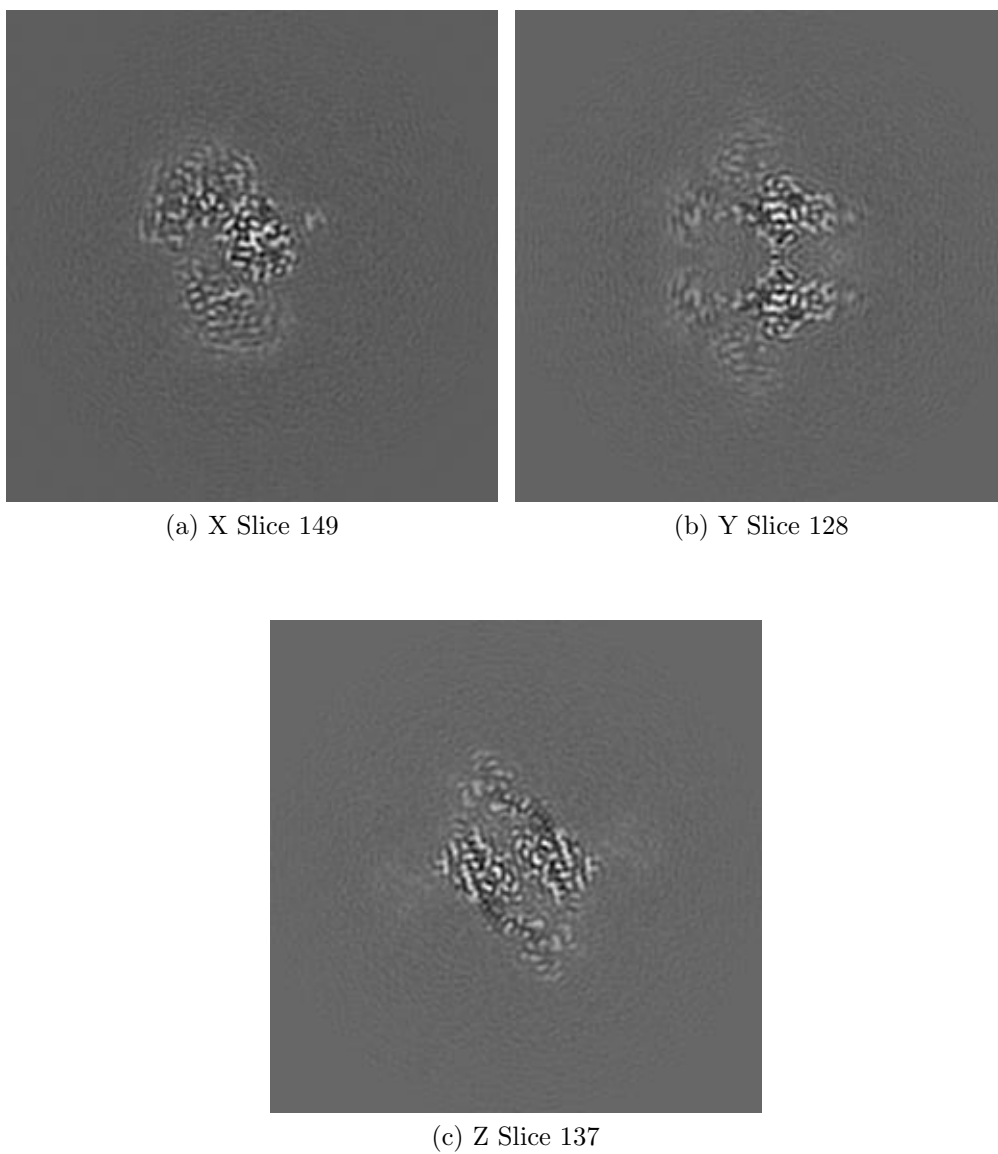


Figure 11: Slices of maximum variation in the three dimensions of the B-factor corrected map

**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%       | 3.49                       |
| 25%        | 4.06                       |
| 50%        | 4.46                       |
| 75%        | 5.42                       |
| 97.5%      | 7.05                       |

The reported resolution, 2.80  $\text{\AA}$ , is at the percentile 0.0. Fig. [13](#) shows some representative views of the local resolution.

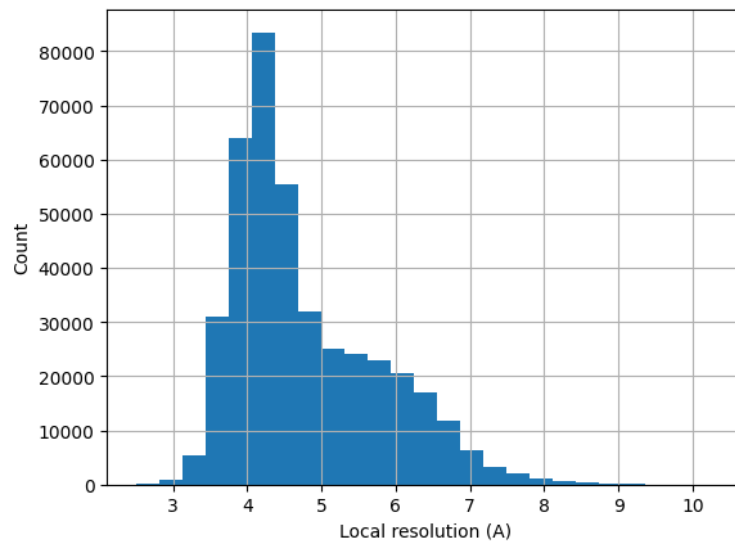
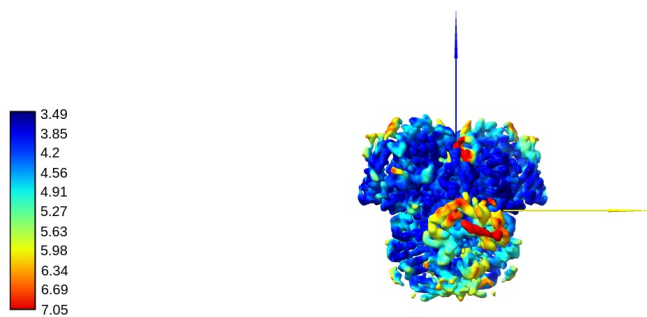
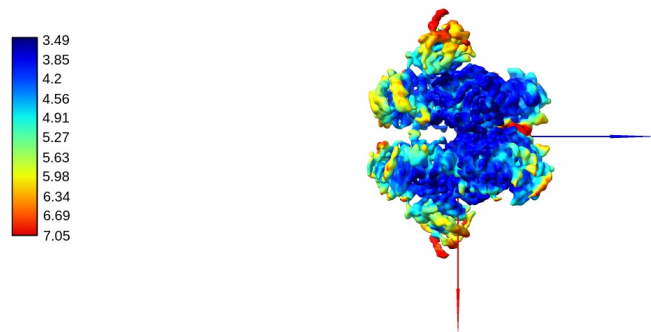


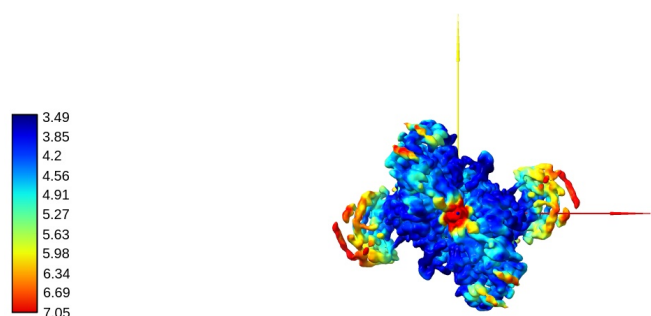
Figure 12: Histogram of the local resolution according to deepres.



(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.

**WARNINGS:** 1 warnings

1. **The reported resolution, 2.80 Å, is particularly high with respect to the local resolution distribution. It occupies the 0.02 percentile**

## 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%       | -282.41                              |
| 25%        | -247.69                              |
| 50%        | -229.90                              |
| 75%        | -212.02                              |
| 97.5%      | -179.32                              |

Fig. 15 shows some representative views of the local B-factor.

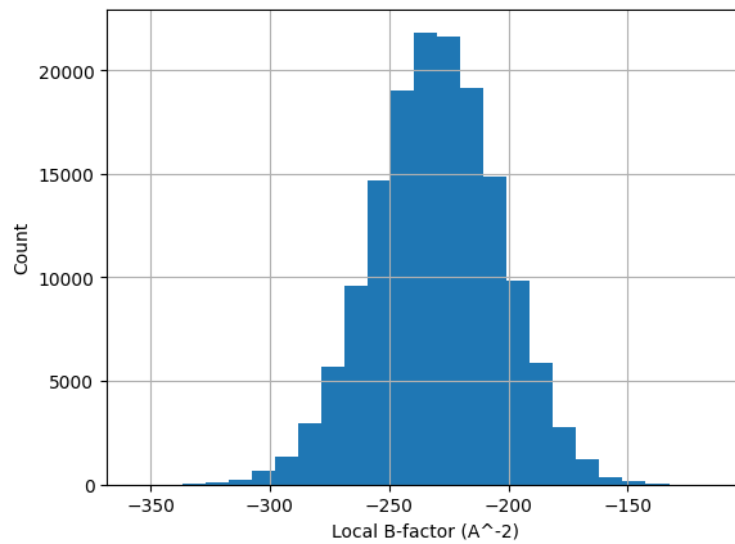
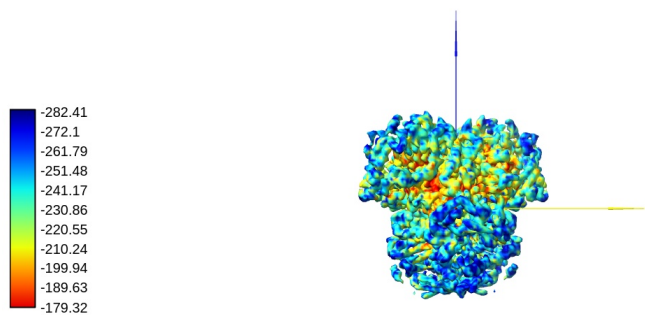
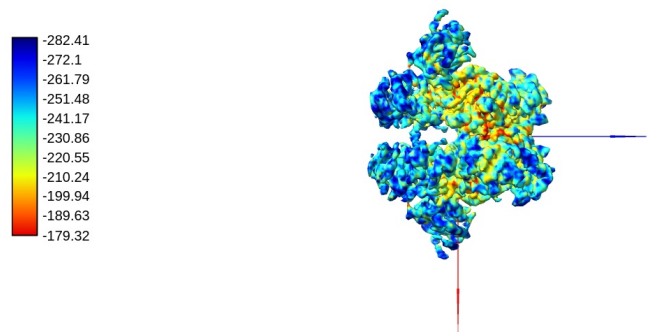


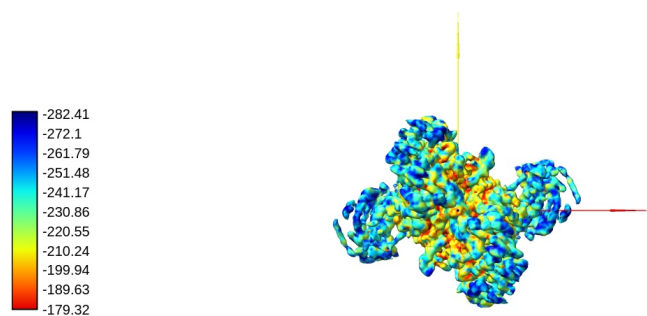
Figure 14: Histogram of the local B-factor according to LocBfactor.



(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%       | 0.08                  |
| 25%        | 0.58                  |
| 50%        | 0.83                  |
| 75%        | 1.00                  |
| 97.5%      | 1.00                  |

Fig. 17 shows some representative views of the local occupancy.

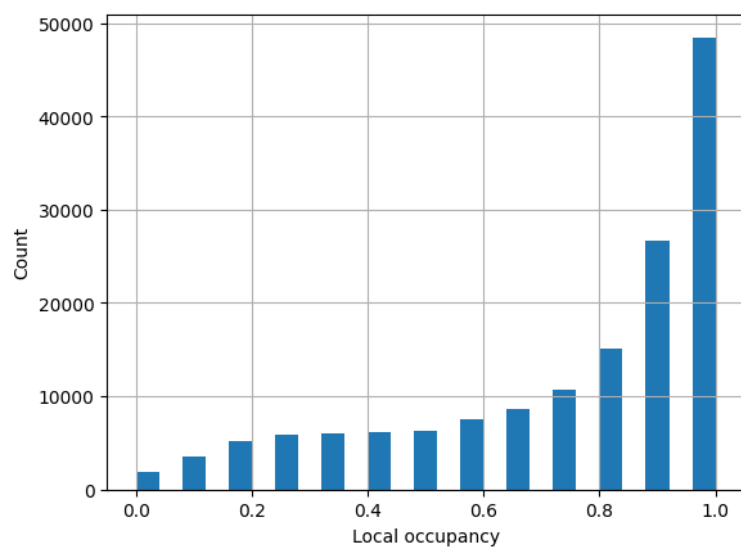
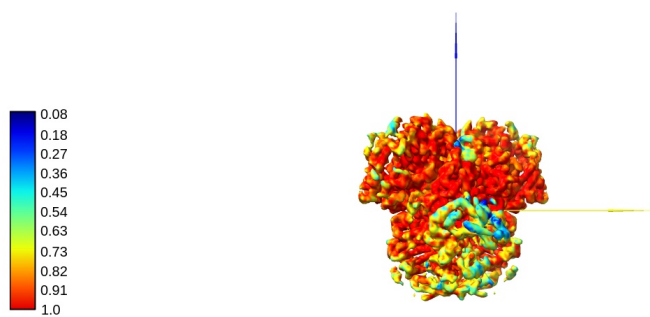
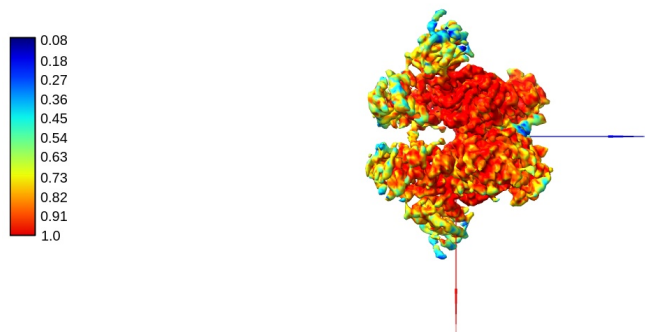


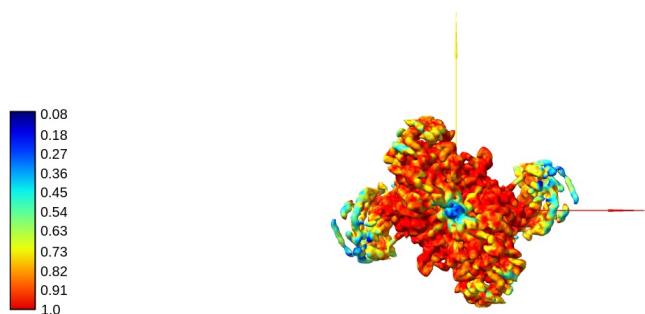
Figure 16: Histogram of the local occupancy according to LocOccupancy.



(a) View 1



(b) View 2



(c) View 3

Figure 17: Local occupancy according to LocOccupancy. Views generated by ChimeraX at 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 0.228 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\_34412\_half\_map\_1.map

SHA256 hash: 1cd493cb93de4a1db1c186cdfe9ba5873c854bd05a38b73b2ccc53b1c2fbf372

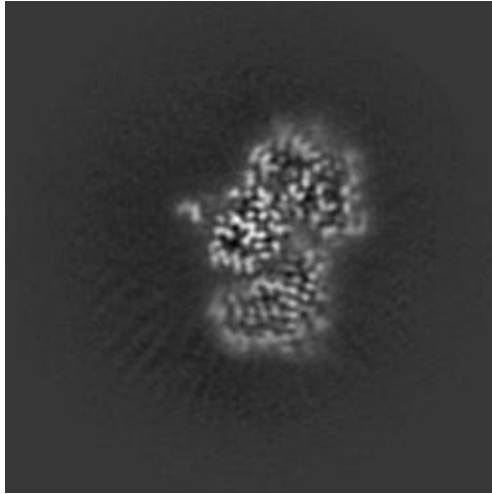
Half map 2: emd\_34412\_half\_map\_2.map

SHA256 hash: f3b03fd02eec717ba979750f6cc67dbc60392625b23a9a1239068c5c163d4d23

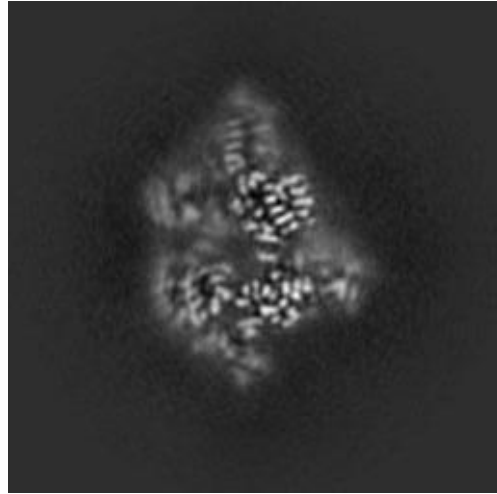
Slices of the first half map can be seen in Fig. [18](#).

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

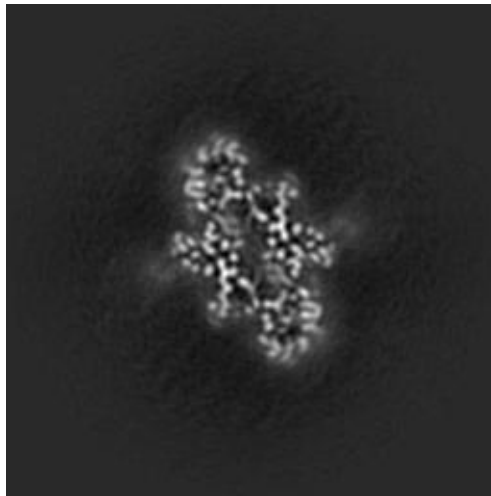
Slices of the difference between both maps can be seen in Fig. [20](#). There should not be any structure in this difference. Sometimes some patterns are seen if the map is symmetric.



(a) X Slice 107

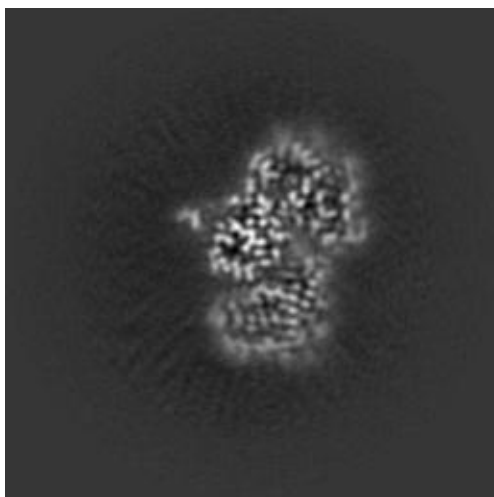


(b) Y Slice 122

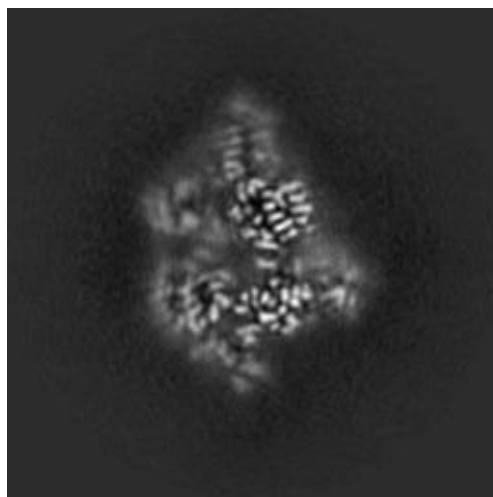


(c) Z Slice 145

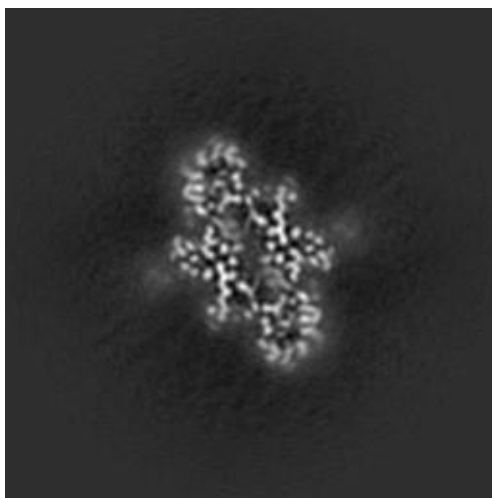
Figure 18: Slices of maximum variation in the three dimensions of Half 1



(a) X Slice 107

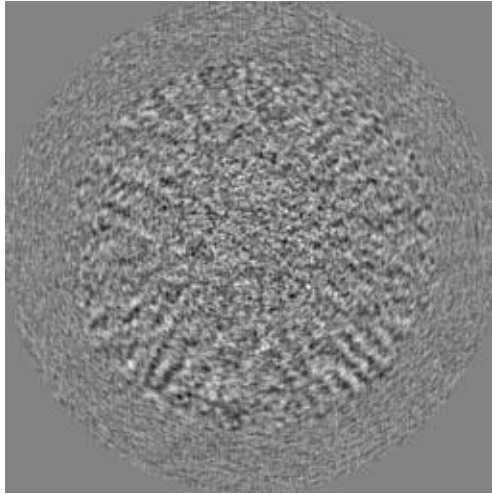


(b) Y Slice 122

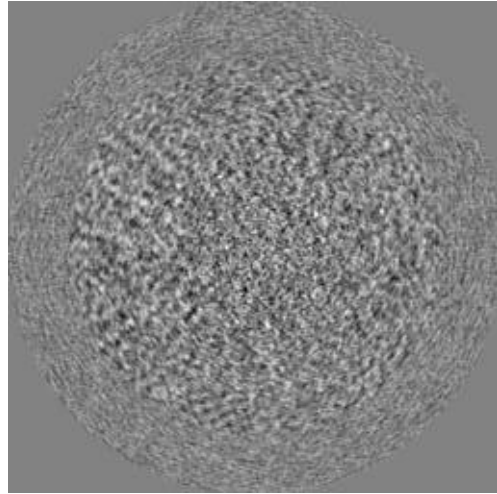


(c) Z Slice 145

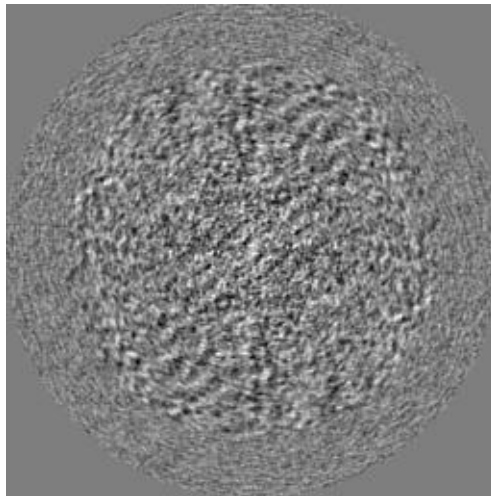
Figure 19: Slices of maximum variation in the three dimensions of Half 2



(a) X Slice 122



(b) Y Slice 118



(c) Z Slice 132

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  $103.9^\circ$  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 2.98Å. The map information is well preserved (FSC>0.9) up to 4.11Å.

Fig. 22 shows the DPR and the  $103.9^\circ$  threshold. The resolution according to the DPR is 2.66Å.

Fig. 23 shows the SSNR and the SSNR=1 threshold. The resolution according to the SSNR is 3.09Å.

The mean resolution between the three methods is 2.91Å and its range is within the interval [ 2.66, 3.09]Å.

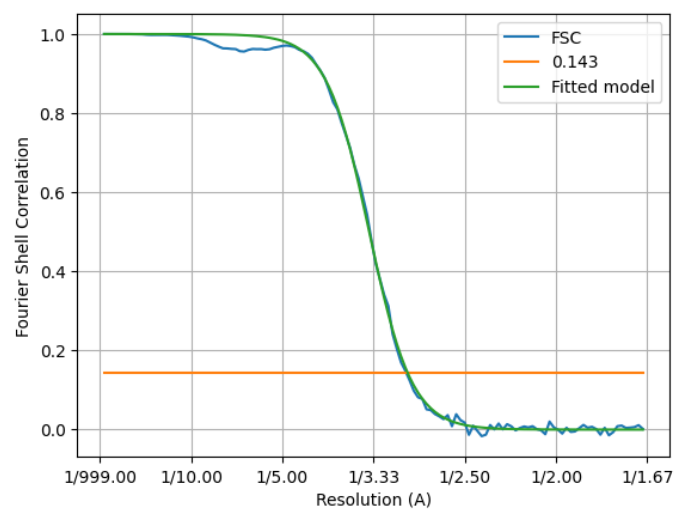


Figure 21: Fourier Shell correlation between the two halves.

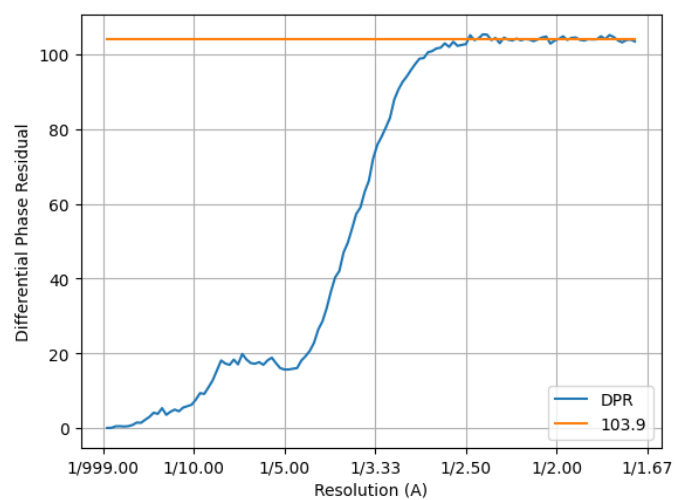


Figure 22: Differential Phase Residual between the two halves.

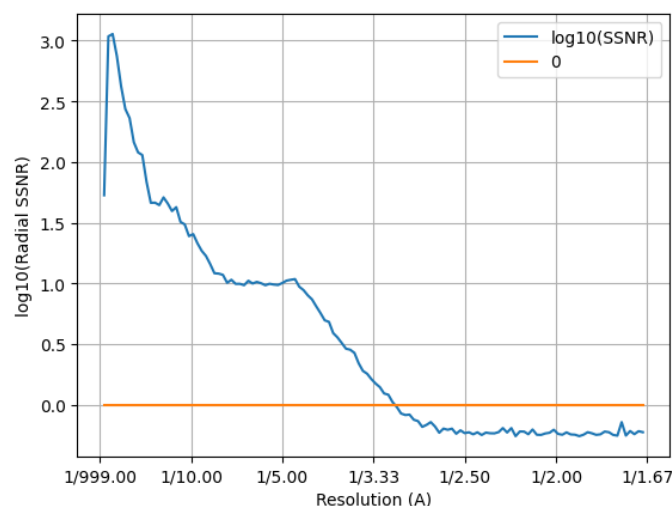


Figure 23: Spectral Signal-to-Noise Ratio estimated from the two halves.

**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 2.5. The estimated B-factor was -130.61. Fig. 24 shows the estimated FSC and resolution.

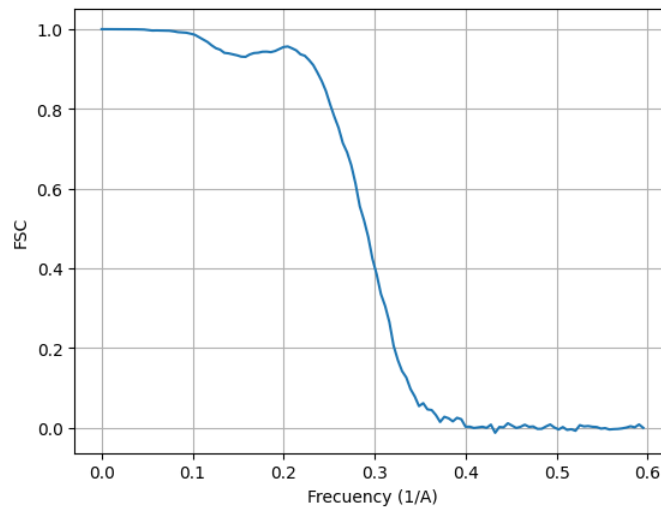


Figure 24: FSC and resolution estimated by a permutation test.

**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%       | 2.47                       |
| 25%        | 2.67                       |
| 50%        | 2.87                       |
| 75%        | 3.17                       |
| 97.5%      | 4.58                       |

The reported resolution, 2.80  $\text{\AA}$ , is at the percentile 42.3. Fig. 26 shows some representative views of the local resolution.

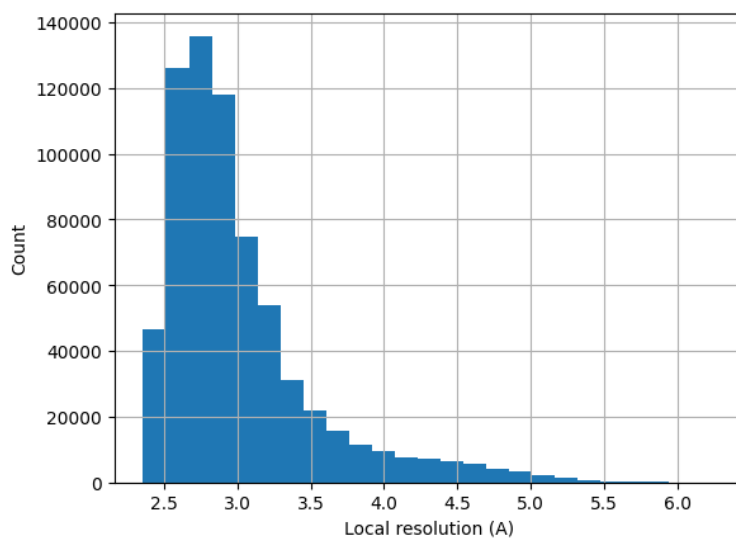
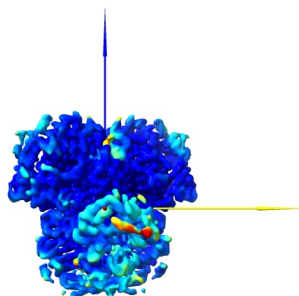
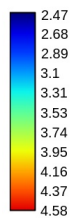
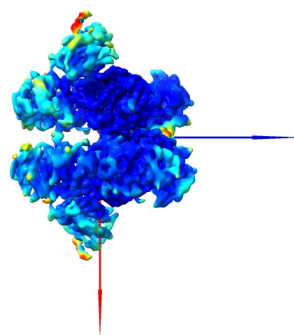
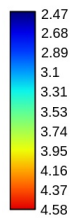


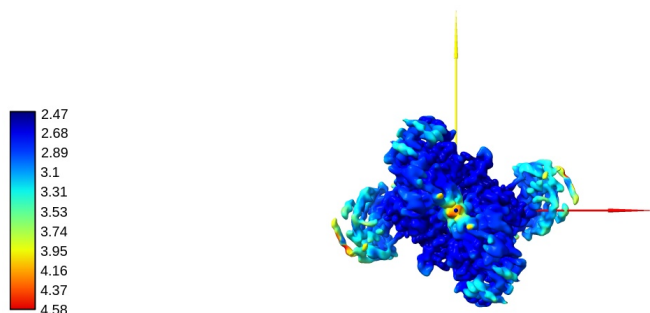
Figure 25: Histogram of the local resolution according to blocres.



(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%       | 1.80                       |
| 25%        | 2.01                       |
| 50%        | 2.66                       |
| 75%        | 3.03                       |
| 97.5%      | 3.53                       |

The reported resolution, 2.80  $\text{\AA}$ , is at the percentile 58. Fig. 28 shows some representative views of the local resolution.

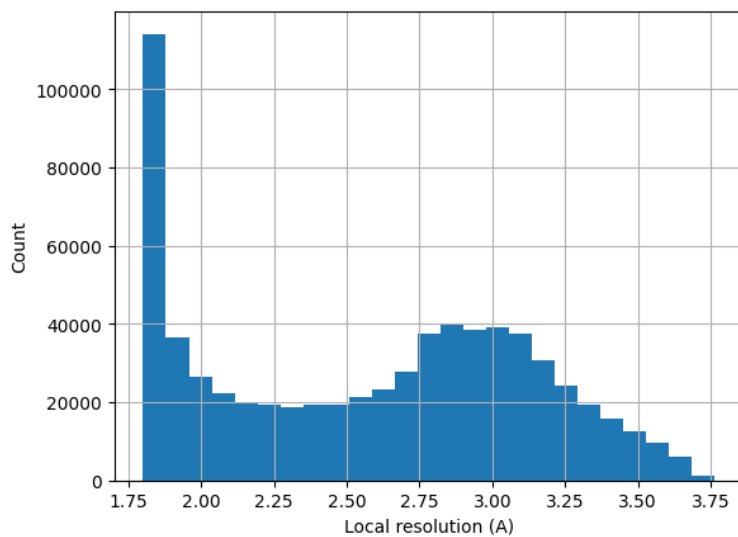
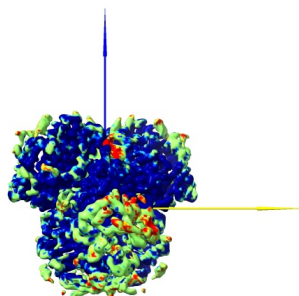
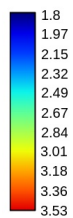
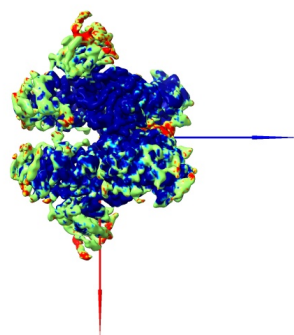
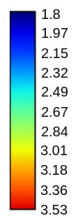


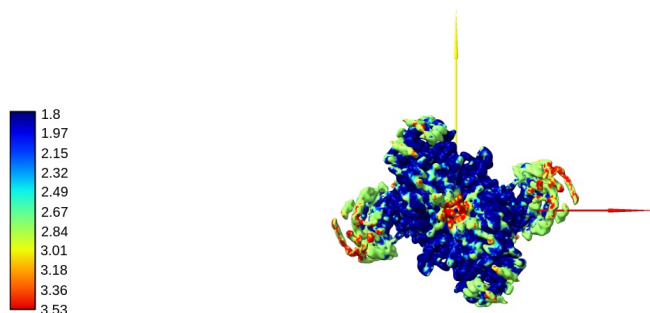
Figure 27: Histogram of the local resolution according to Resmap.



(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%       | 2.38                       |
| 25%        | 3.27                       |
| 50%        | 3.61                       |
| 75%        | 4.15                       |
| 97.5%      | 5.62                       |

The reported resolution, 2.80  $\text{\AA}$ , is at the percentile 7.1. Fig. 30 shows some representative views of the local resolution

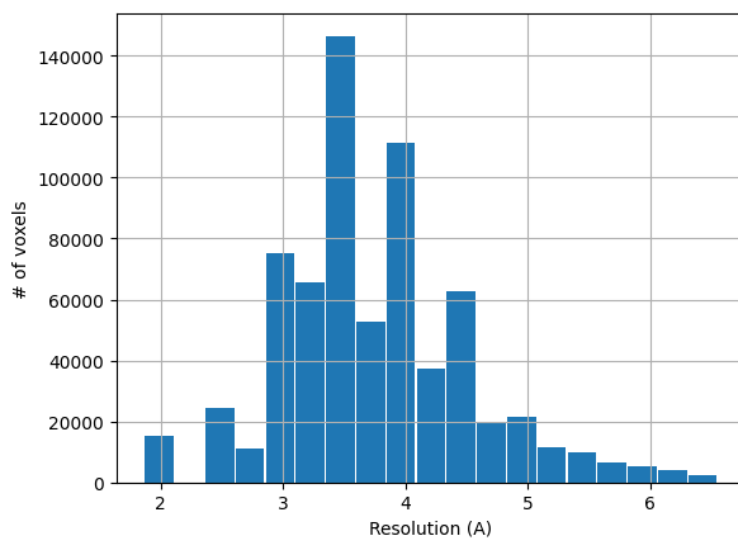
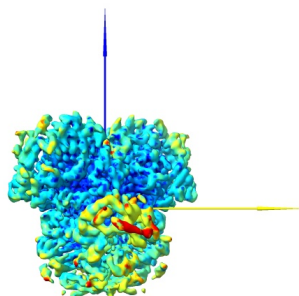
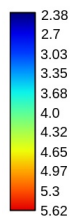
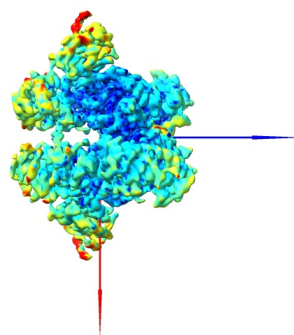
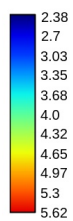


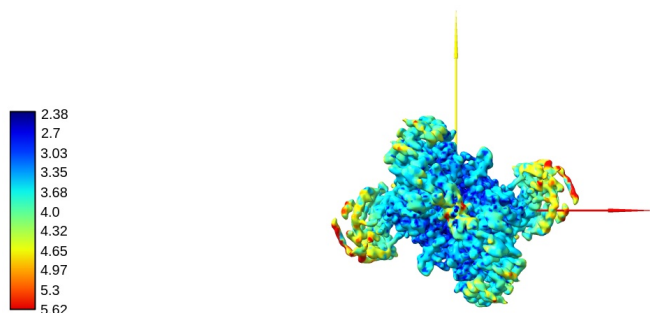
Figure 29: Histogram of the local resolution according to MonoRes.



(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:

Fig. 31 shows the 1D directional histogram and Fig. 32 the 2D directional histogram. We compared the 1D directional histogram to a uniform distribution using a Kolmogorov-Smirnov test. The D statistic was 0.056522, 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. 33. The overall mean of the directional resolution is 3.05

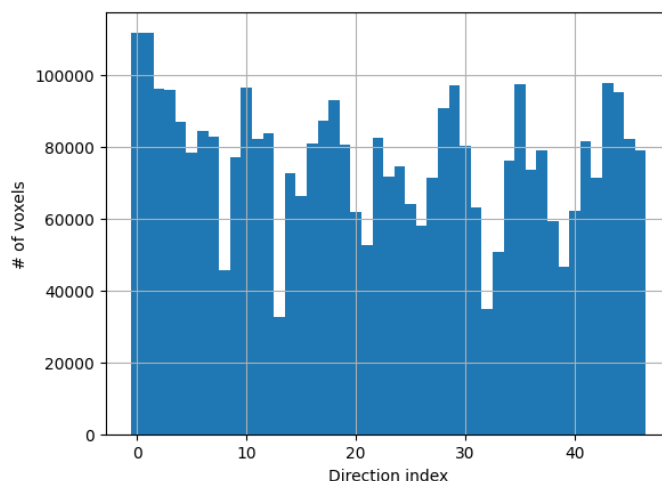


Figure 31: Histogram 1D of the best direction at each voxel.

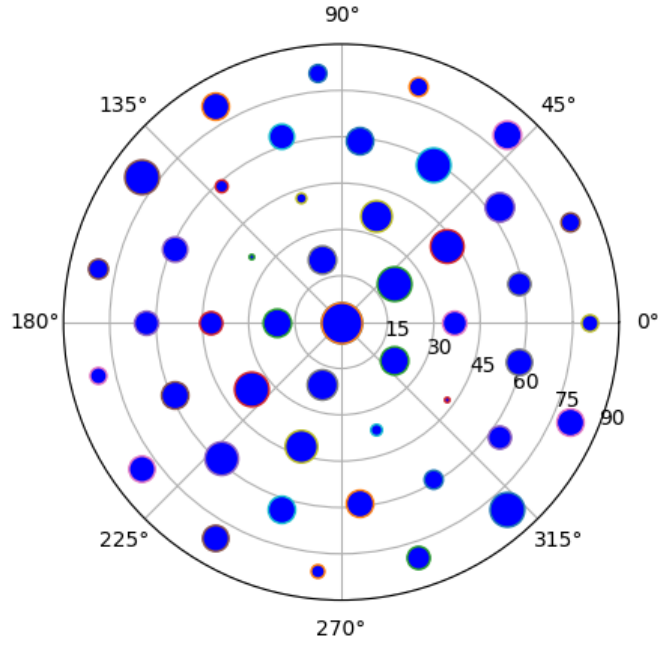


Figure 32: Histogram 2D of the best direction at each voxel. The azimuthal rotation is circular, while the tilt angle is the radius. The size of the point is proportional to the number of voxels whose maximum resolution is in that direction (this count can be seen in Fig. 31).

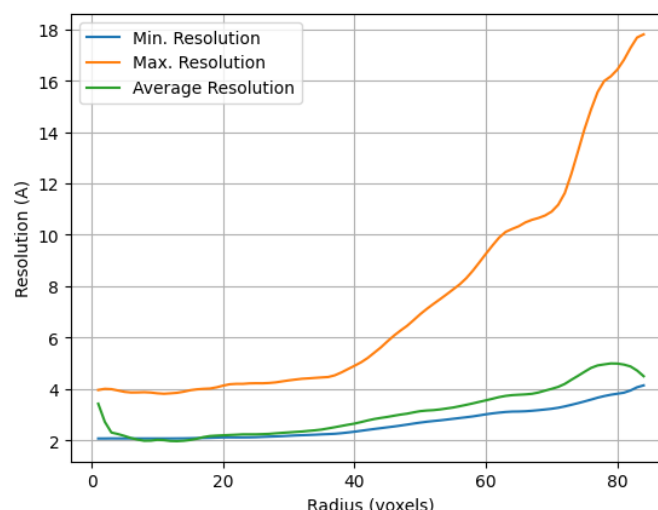


Figure 33: Radial averages (in space) of the minimum, maximum and average resolution at each voxel.

**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 (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. 34 shows the Fourier Shell Occupancy and its anisotropy. The directional resolution is shown in Fig. 35. The resolution according to the FSO is 2.69Å. Fourier shells are occupied at between 90 and than 10% in the range [ 2.85, 2.52]Å.

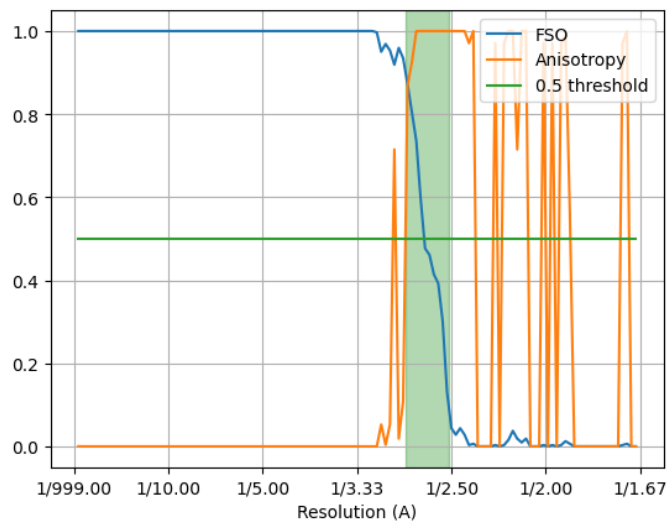


Figure 34: FSO and anisotropy.

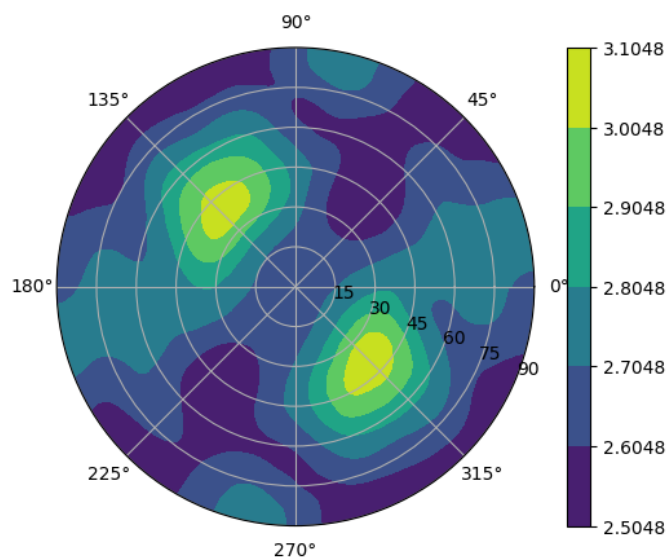


Figure 35: Directional resolution in the projection sphere.

**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. 36 shows the FSCs in X, Y, Z, and the global FSC. Fig. 37 shows the global FSC and the histogram of the directional FSC. Finally, Fig. 38 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 1.84, 1.89, and 1.91 Å, respectively. The global resolution at the same threshold is 1.92 Å. The resolution range is [ 1.84, 1.92]Å.

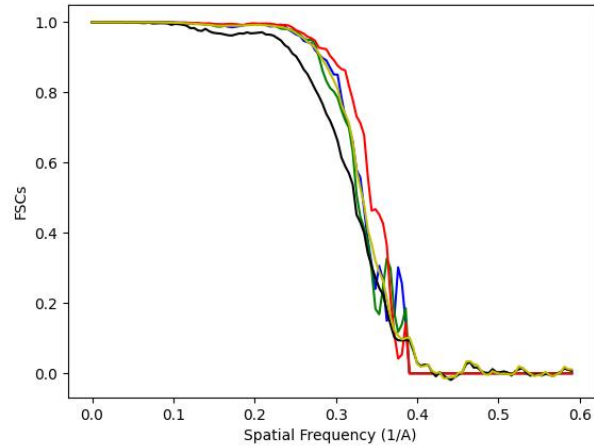


Figure 36: FSC in X, Y, Z, the global FSC, and the Average Cosine Phase.

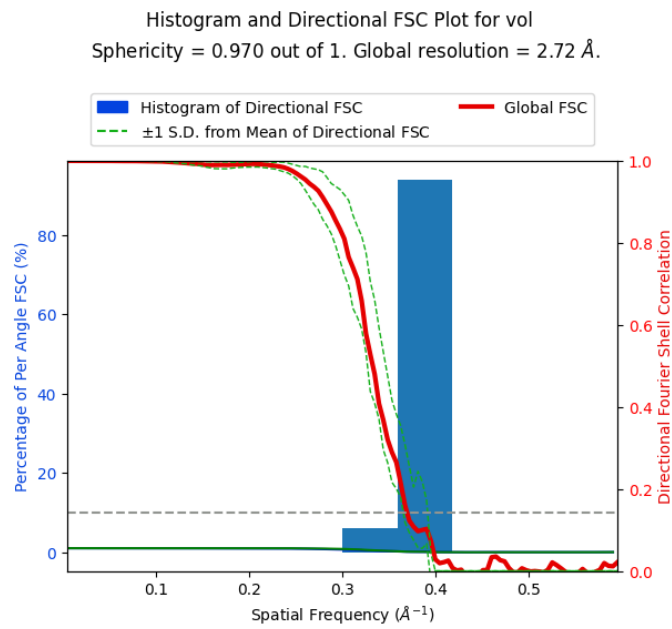


Figure 37: Global FSC and histogram of the directional FSC.

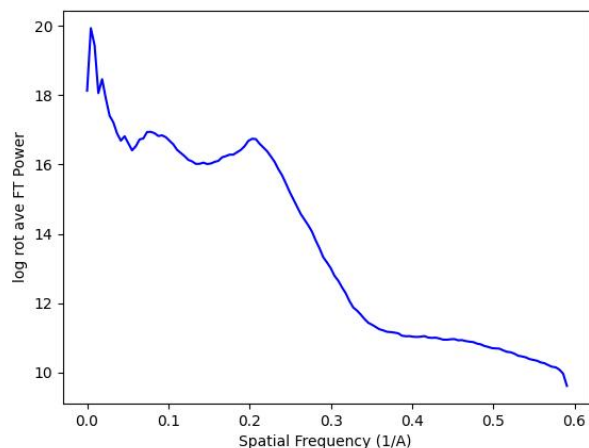


Figure 38: Logarithm of the radial average of the input map power in Fourier space.

**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: 8h0i.cif

See Fig. 39.

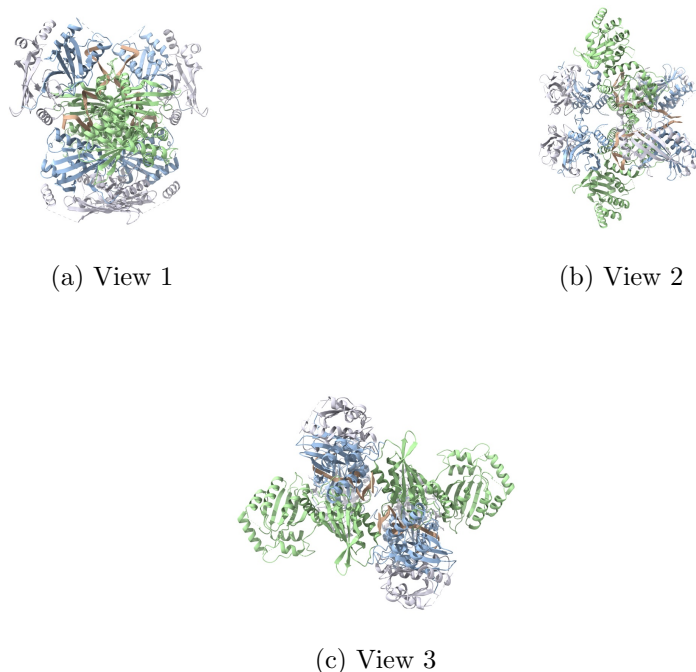


Figure 39: Input atomic model Views generated by ChimeraX at a the following X, Y, Z angles: View 1 (0, -90, -90), View 2 (-90, 0, -90), View 3 (0, 0, 0).

## 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 6.0 Å has been applied.

#### Overall correlation coefficients:

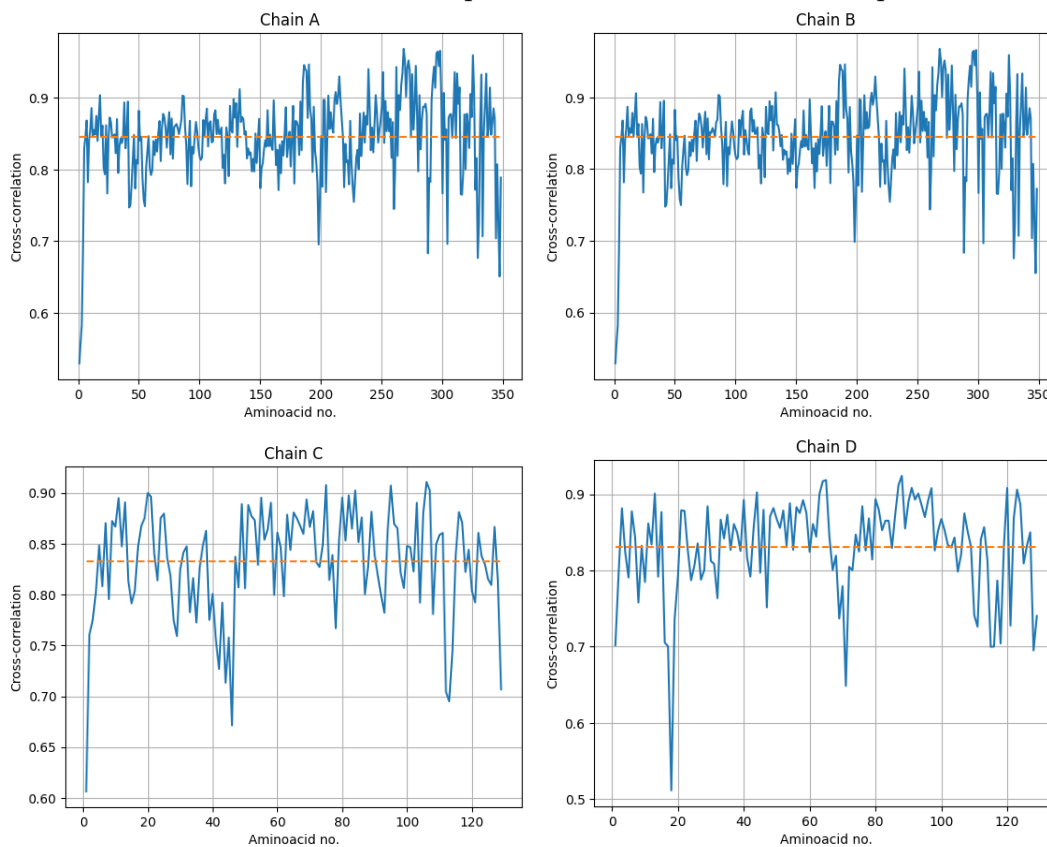
|                   |       |
|-------------------|-------|
| CC (mask) =       | 0.784 |
| CC (box) =        | 0.858 |
| CC (volume) =     | 0.784 |
| CC (peaks) =      | 0.755 |
| CC (main chain) = | 0.812 |
| CC (side chain) = | 0.783 |

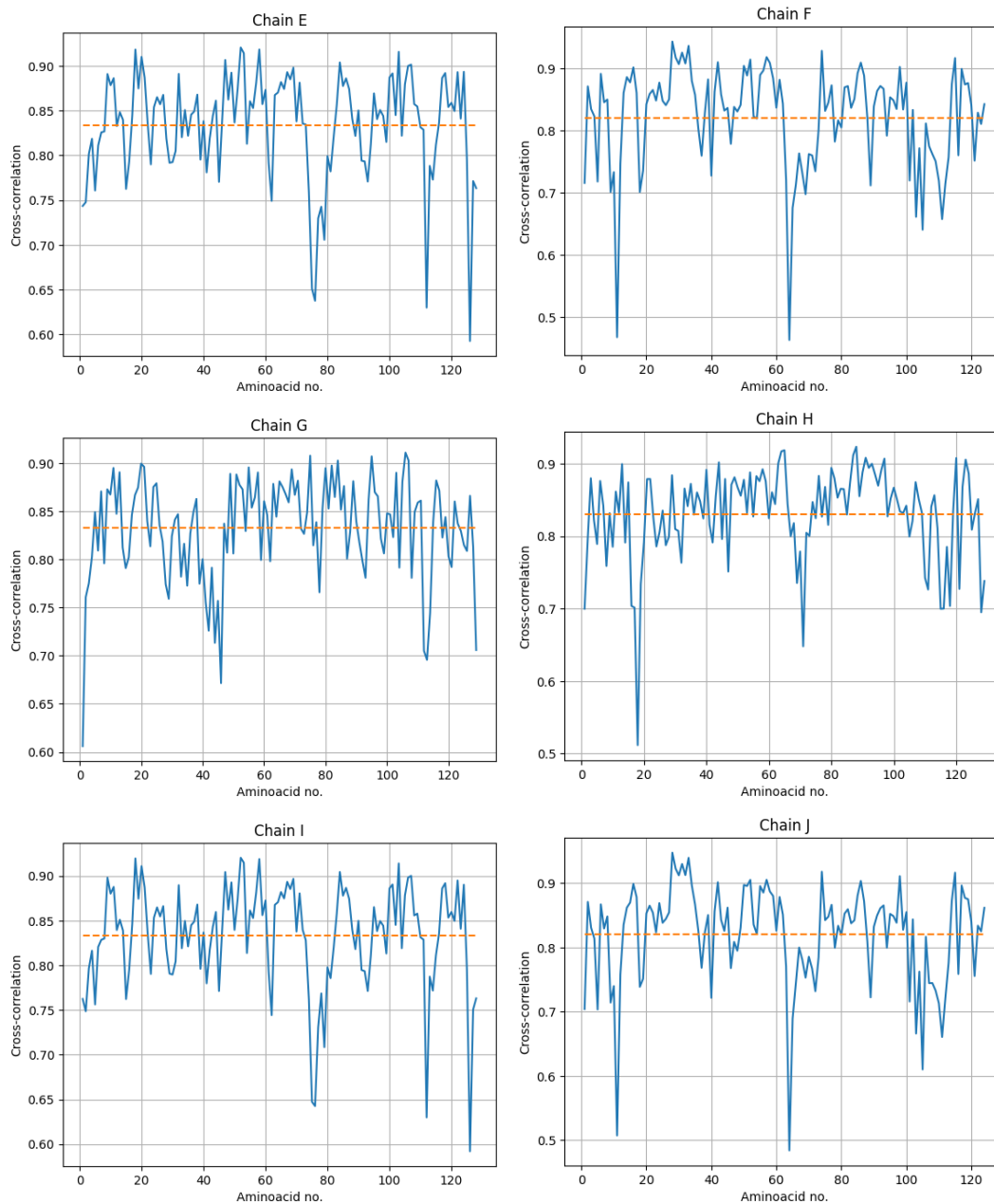
#### Correlation coefficients per chain:

# Chain Cross-correlation

|   |          |
|---|----------|
| A | 0.817320 |
| B | 0.816683 |
| C | 0.806202 |
| D | 0.798565 |
| E | 0.798529 |
| F | 0.779910 |
| G | 0.805958 |
| H | 0.798411 |
| I | 0.799218 |
| J | 0.781051 |
| X | 0.671253 |
| Y | 0.671070 |

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





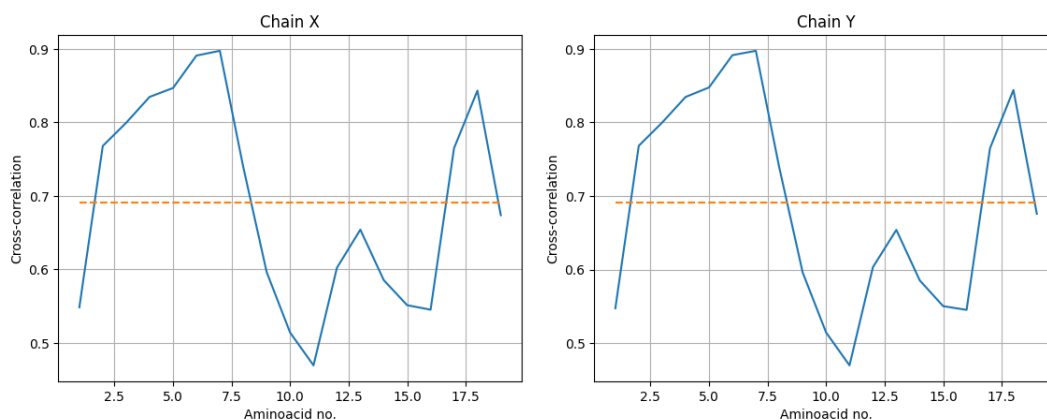


Fig. 40 shows the histogram of all cross-correlations evaluated at the residues. The percentage of residues whose correlation is below 0.5 is 0.3 %.

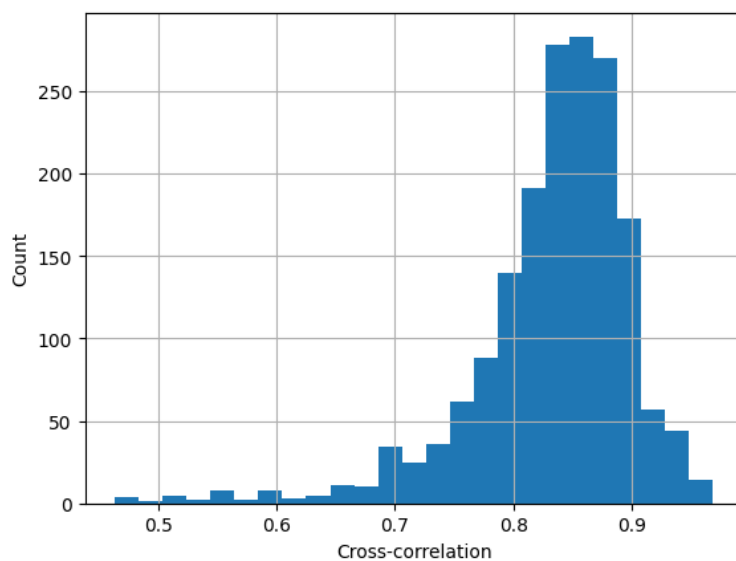


Figure 40: Histogram of the cross-correlation between the map and model evaluated for all residues.

Resolutions estimated from the model:

| Resolution (Å)       | Masked | Unmasked |
|----------------------|--------|----------|
| d99                  | 4.1    | 4.1      |
| d_model              | 3.2    | 3.1      |
| d_model (B-factor=0) | 7.5    | 8.1      |
| FSC_model=0          | 2.2    | 2.3      |
| FSC_model=0.143      | 2.5    | 2.7      |
| FSC_model=0.5        | 2.9    | 3.1      |

Overall isotropic B factor:

| B factor      | Masked | Unmasked |
|---------------|--------|----------|
| Overall B-iso | 125.0  | 140.0    |

Fig. 41 shows the FSC between the input map and the model.

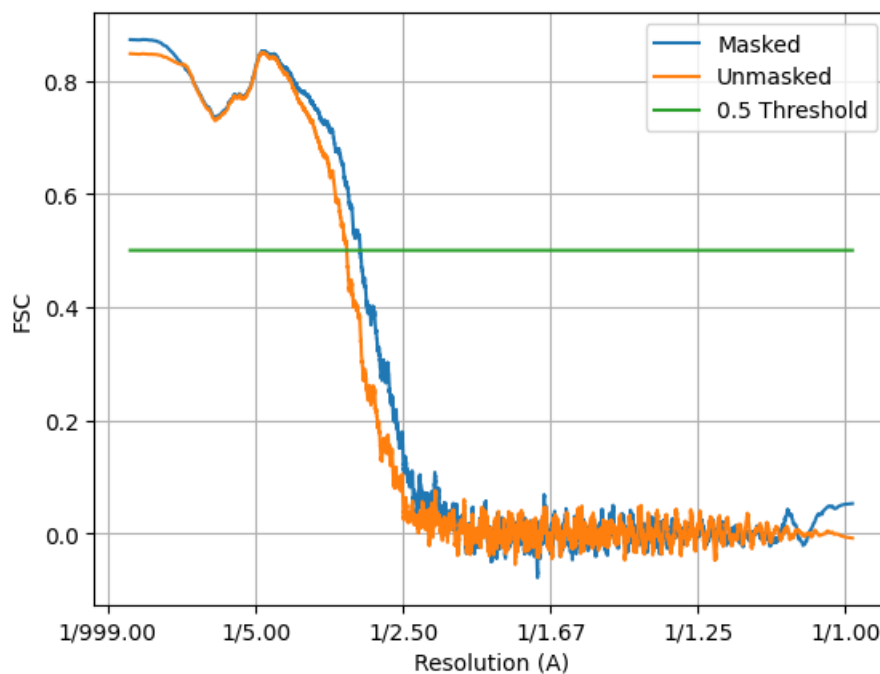


Figure 41: FSC between the input map and model with and without a mask constructed from the model. The X-axis is the square of the inverse of the resolution in Å.

**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](#)

## 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. [42](#) shows the histogram of FSC-Qr and Fig. [43](#) the colored isosurface of the atomic model converted to map. The average FSC-Qr is 0.31, its 95% confidence interval is [ 0.06, 1.25]. The percentage of values whose FSC-Qr absolute value is beyond 1.5 is 1.2 %.

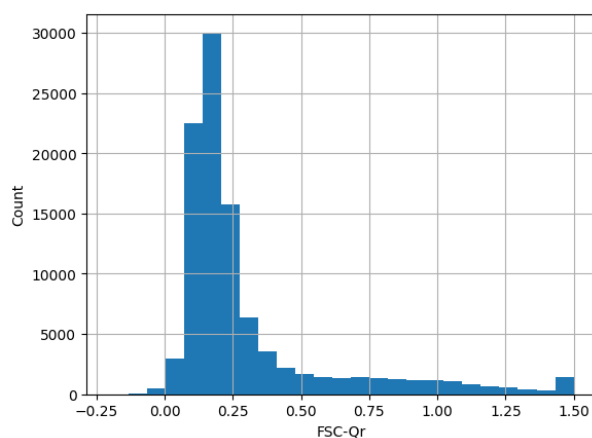
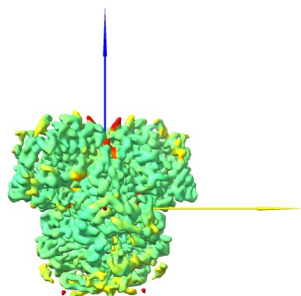
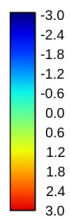
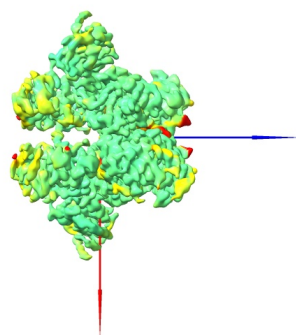
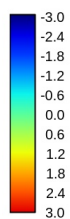


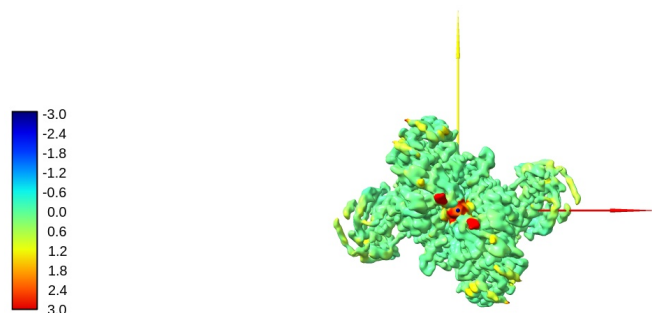
Figure 42: Histogram of the FSC-Qr limited to -1.5 and 1.5.



(a) View 1



(b) View 2



(c) View 3

Figure 43: 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. 44 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 0.420.

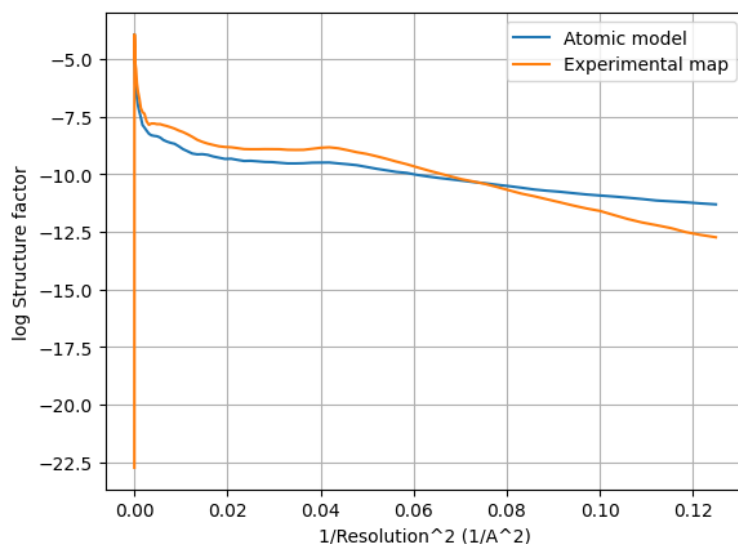


Figure 44: Guinier plot of the atom model and experimental map. The X-axis is the square of the inverse of the resolution in Å.

**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 0.420.**

## 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. 45 shows the histogram of the calculated Q-score. Some representative percentiles are:

| Percentile | MapQ score [-1,1] |
|------------|-------------------|
| 2.5%       | -0.17             |
| 25%        | 0.00              |
| 50%        | 0.00              |
| 75%        | 0.00              |
| 97.5%      | 0.19              |

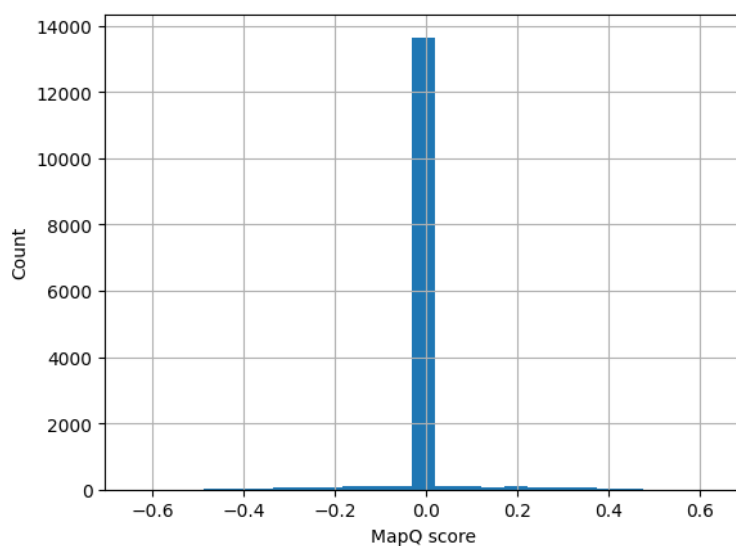
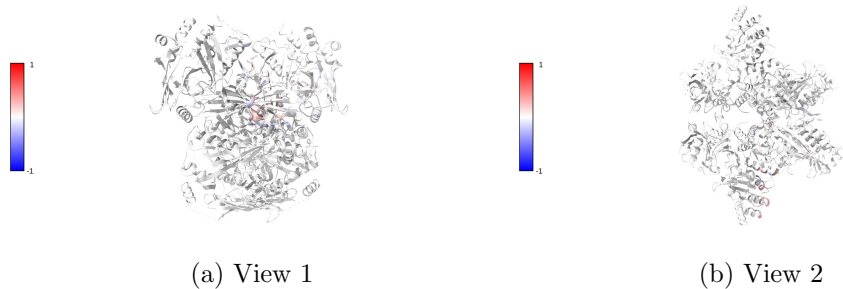


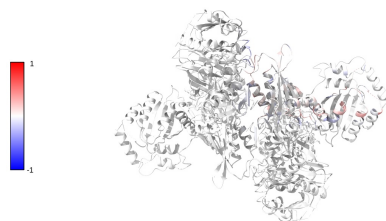
Figure 45: Histogram of the per-atom Q-score.

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

| Chain | Average Q-score [-1,-1] | Estimated Resol. (Å) |
|-------|-------------------------|----------------------|
| A     | 0.00                    | 9.8                  |
| A     | -0.00                   | 6.3                  |
| B     | 0.08                    | 9.1                  |
| B     | 0.01                    | 6.2                  |
| C     | 0.00                    | 6.3                  |
| D     | -0.01                   | 6.4                  |
| E     | 0.00                    | 6.3                  |
| F     | 0.00                    | 6.3                  |
| G     | -0.00                   | 6.3                  |
| H     | 0.00                    | 6.3                  |
| I     | 0.00                    | 6.3                  |
| J     | 0.00                    | 6.3                  |
| X     | -0.02                   | 7.4                  |
| Y     | -0.00                   | 7.3                  |

The atomic model colored by MapQ can be seen in Fig. 46.





(c) View 3

Figure 46: 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:** 1 warnings

1. **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 | 8.080268 |
| Rotamer ratio     | 0.883    |
| Max. Zscore       | 12.35    |
| Model length      | 1178     |
| EMRinger Score    | 3.599    |

Fig. 47 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.

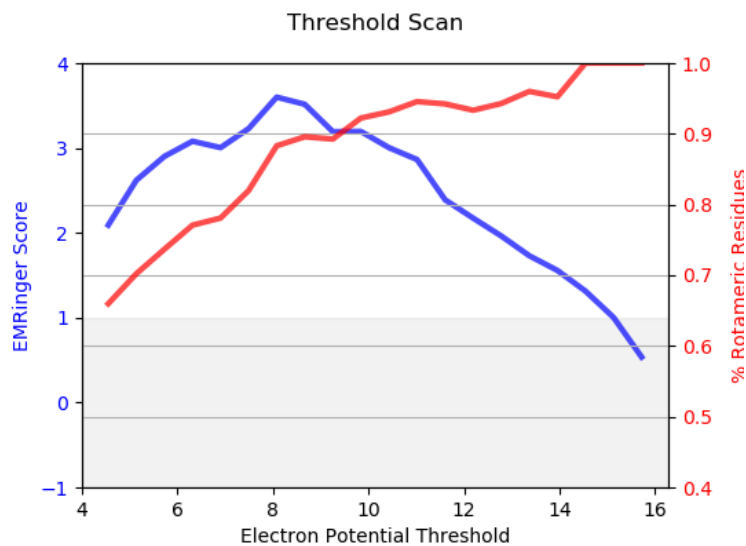


Figure 47: EMRinger score and fraction of rotameric residues as a function of the map threshold.

Fig. 48 shows the histogram for rotameric (blue) and non-rotameric (red) residues at the optimal threshold.

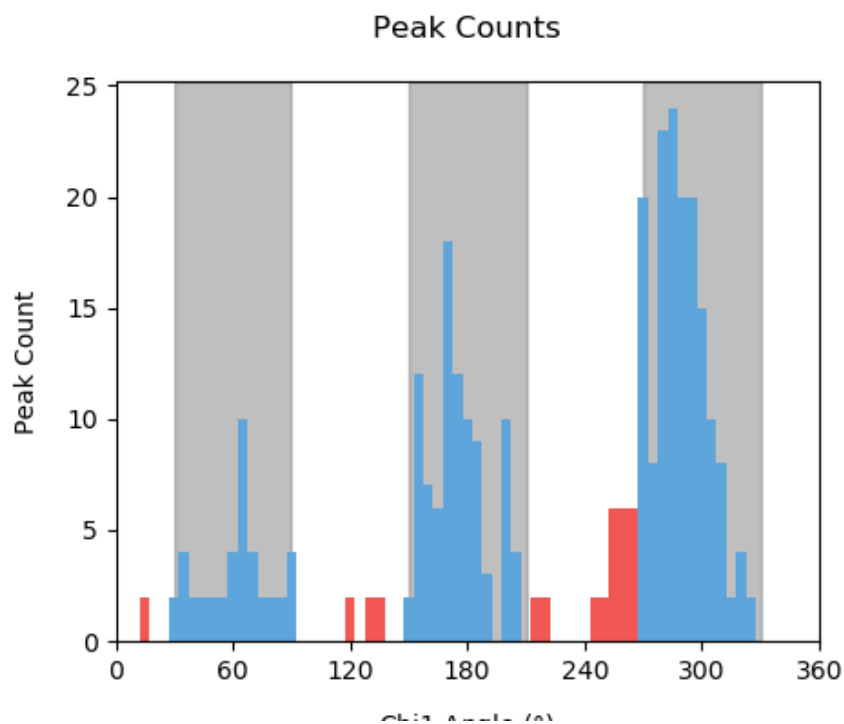
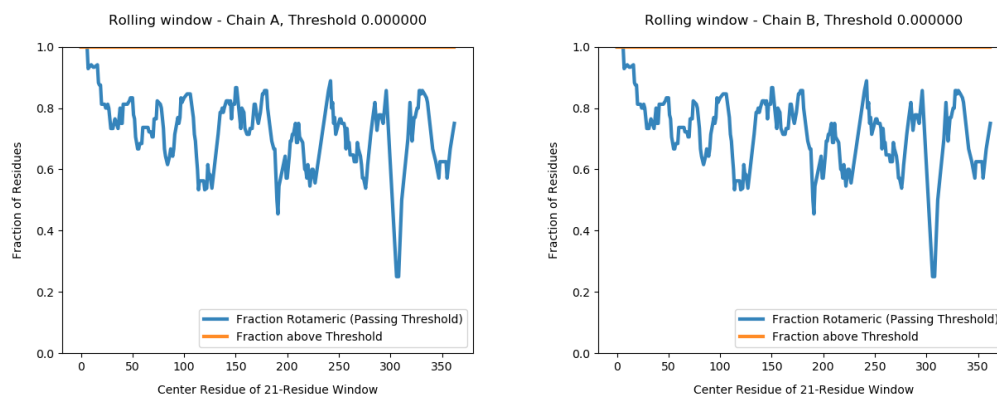
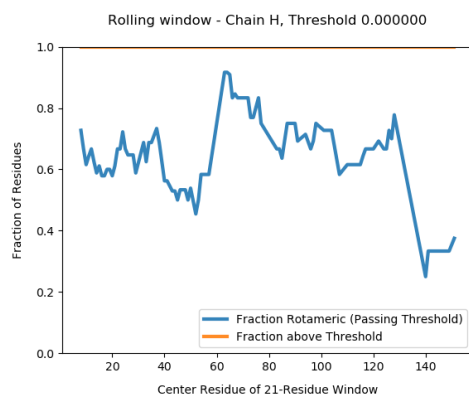
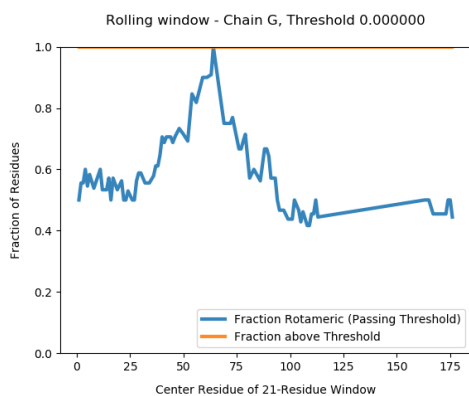
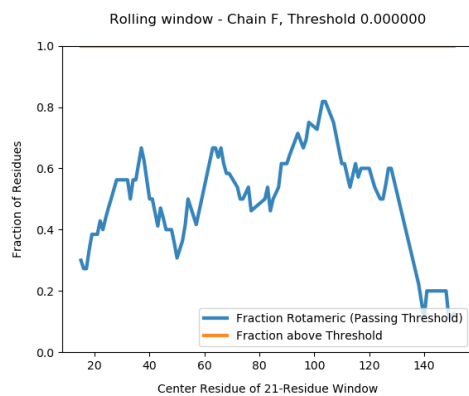
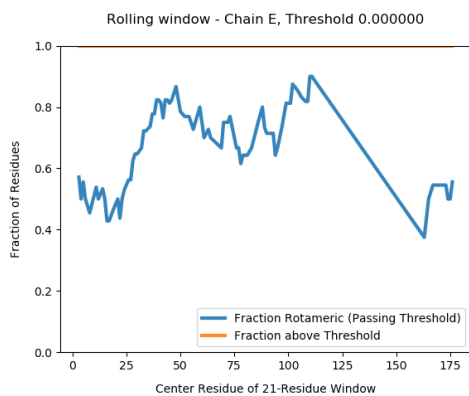
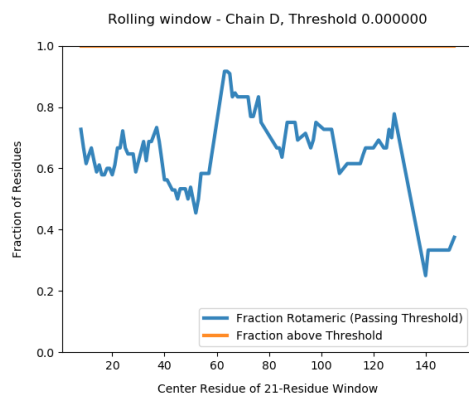
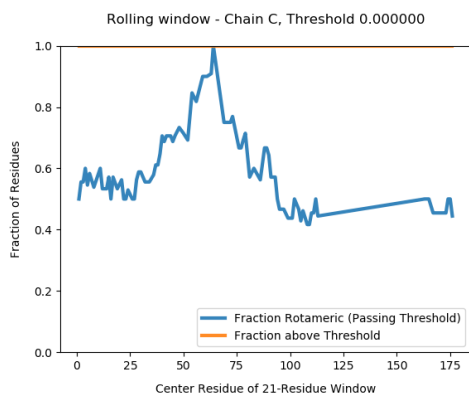
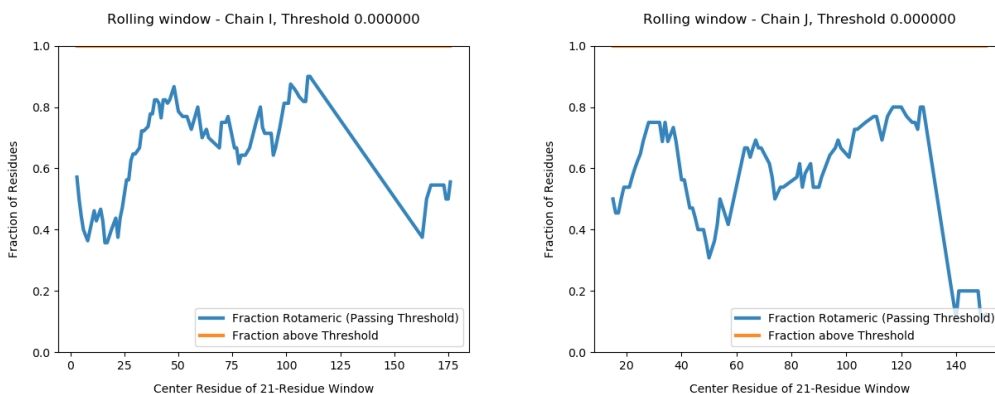


Figure 48: Histogram for rotameric (blue) and non-rotameric (red) residues at the optimal threshold as a function of the angle Chi1.

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-34412/daq/>

Fig. 49 shows the histogram of the DAQ values. The mean and standard deviation were 0.8 and 0.4, respectively.

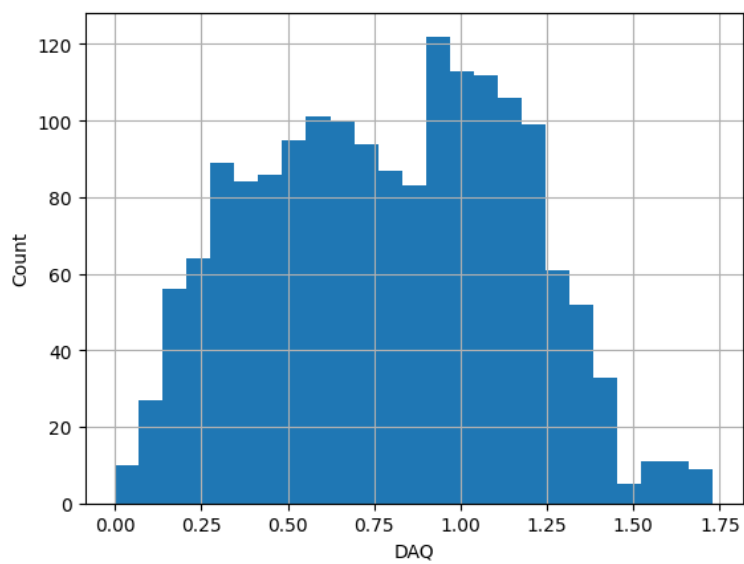
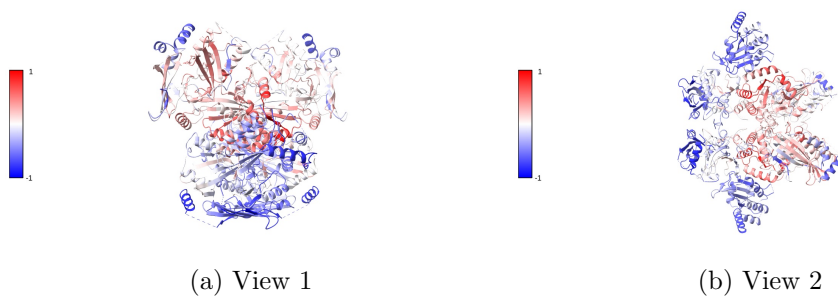
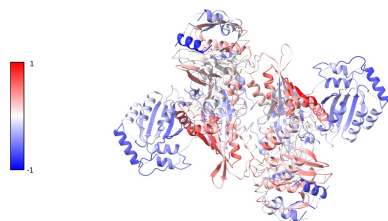


Figure 49: Histogram of the per-residue DAQ values.

The atomic model colored by DAQ can be seen in Fig. 50.





(c) View 3

Figure 50: Atomic model colored by DAQ 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 average DAQ score is larger than 0.5.

**STATUS:** OK