Public Data-based Report

Validation Report Service

Cryo-EM Map Validation Report

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

This report has been generated based on data publicly available at EMDB.

Basic Entry Information:

EMDB ID: EMD-25771
PDB ID: 7tad
Title: CryoEM structure of the (NPR1)2-(TGA3)2 complex
Authors: See EMDB entry link
Deposited on: 2021-12-20T00:00:00
Reported Resolution: 3.6 Å

Contact Us:

Instruct Image Processing Center (I^2PC)
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: **September 14, 2026, 12:12am**

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). According to phenix, it seems that there might be some mismatch between the map and its model (see Sec. 4.1). DAQ detects some mismatch between the map and its model (see Sec. 4.5).

The average resolution of the map estimated by various methods goes from 1.5Å to 6.4Å with an average of 3.7Å. The resolution reported by the user was 3.6Å.

The overall score (passing tests) of this report is 5 out of 9 evaluable items.

0.a Mass analysis	Sec. 2.1	3 warnings
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
A.a Phenix validation	Sec. 4.1	2 warnings
A.d Map-Model Guinier	Sec. 4.2	OK
A.e MapQ	Sec. 4.3	OK
A.f EMRinger	Sec. 4.4	OK
A.g DAQ	Sec. 4.5	1 warnings

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=4588.45)**

Section 4.1 (A.a Phenix validation)

1. **The percentage of residues that have a cross-correlation below 0.5 is 16.4, that is larger than 10%**
2. **The resolution reported by the user, 3.6 Å, is significantly smaller than the resolution estimated between map and model (FSC=0.5), 6.4 Å**

Section 4.5 (A.g DAQ)

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

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 18
2.4 Level 0.d B-factor analysis 20

3 Atomic model **23**

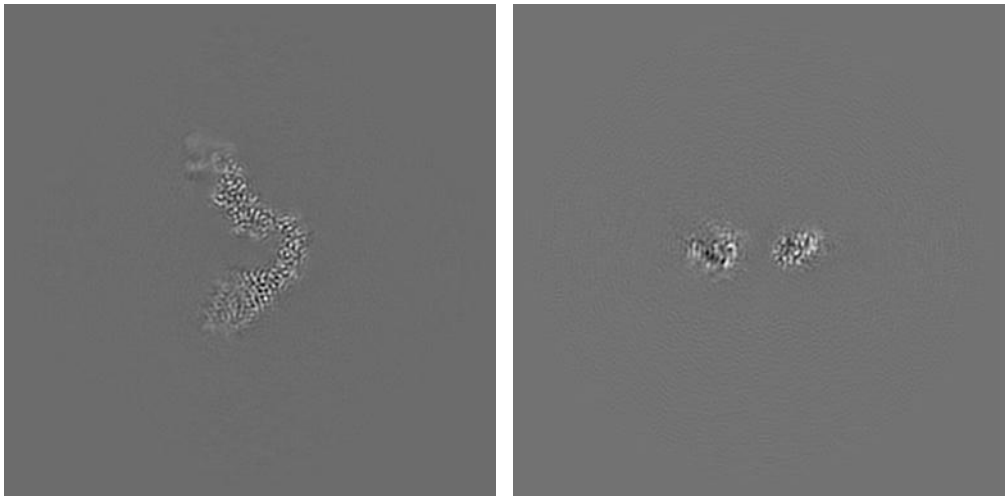
4 Level A Analysis **23**
4.1 Level A.a Phenix validation 24
4.2 Level A.d Map-Model Guinier analysis 28
4.3 Level A.e MapQ 29
4.4 Level A.f EMRinger validation 31
4.5 Level A.g DAQ validation 35

1 Input data

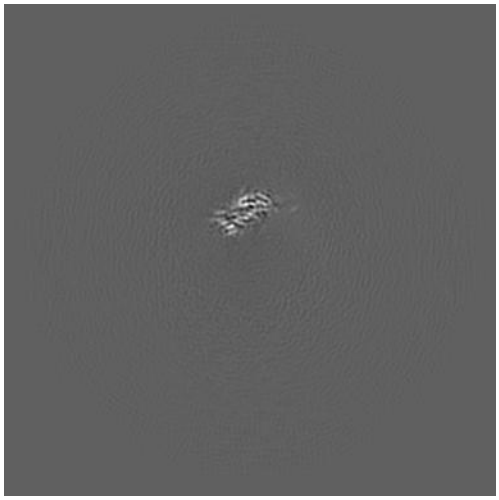
Input map: emd_25771.map
SHA256 hash: 131692672e2197842869be921221055a79d68c1d3fd53934a793a5b4ced70736
Voxel size: 1.066000 (Å)
Visualization threshold: 4.200000
Resolution estimated by user: 3.6

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 192 (b) Y Slice 192



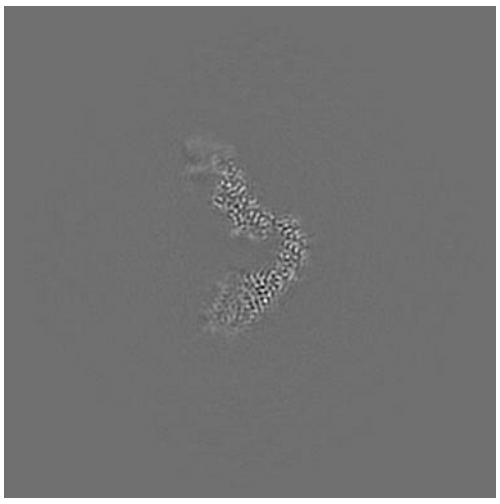
(c) Z Slice 192

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

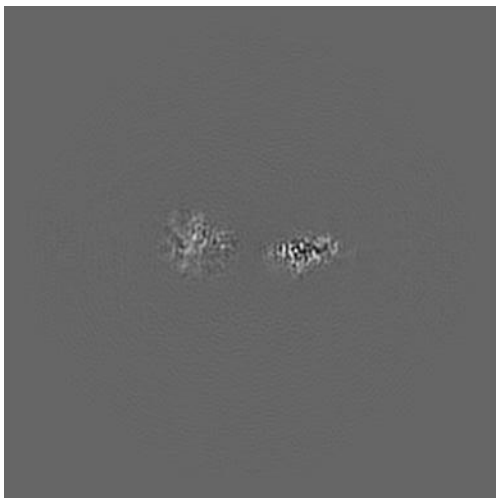
Orthogonal slices of maximum variance of the input map

Results:

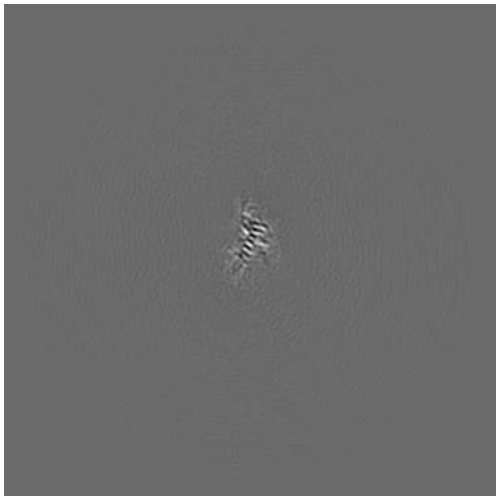
See Fig. 2.



(a) X Slice 193



(b) Y Slice 180



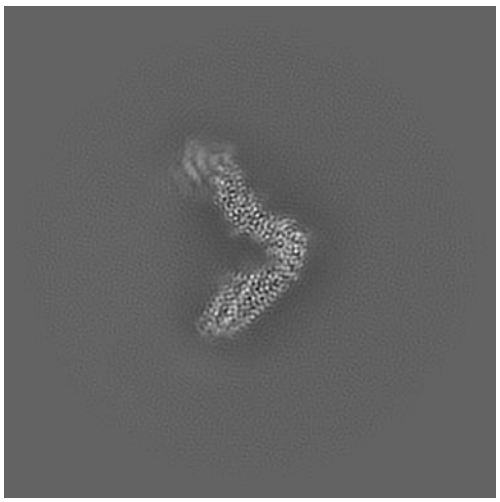
(c) Z Slice 169

Figure 2: Slices of maximum variation in the three dimensions

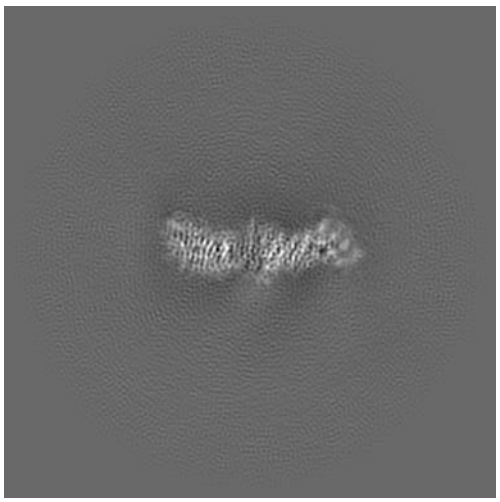
Orthogonal projections of the input map

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

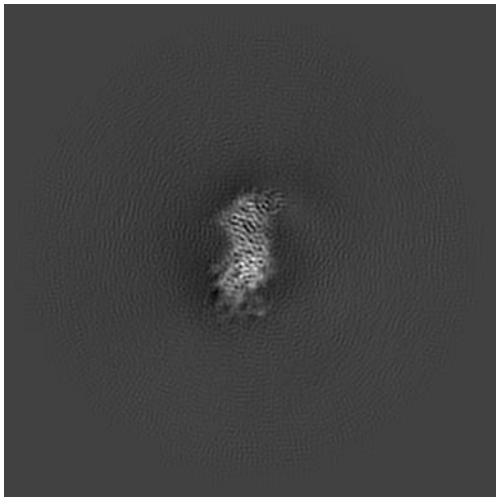
Results:
See Fig. 3.



(a) X Projection



(b) Y Projection



(c) Z Projection

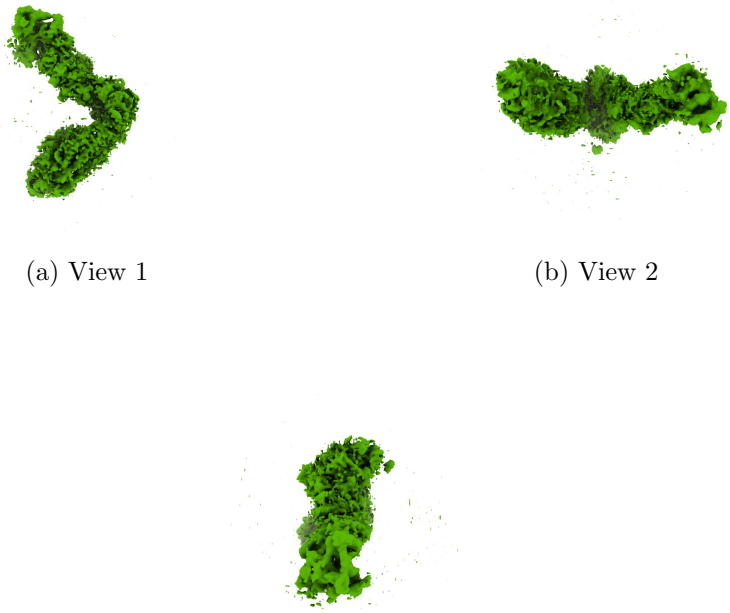
Figure 3: Projections in the three dimensions

Isosurface views of the input map

Explanation:

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

Results:
See Fig. 4.



(a) View 1

(b) View 2

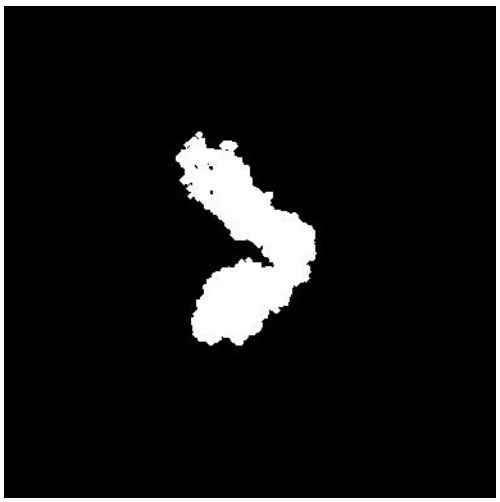
(c) View 3

Figure 4: Isosurface at threshold=4.200000. 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).

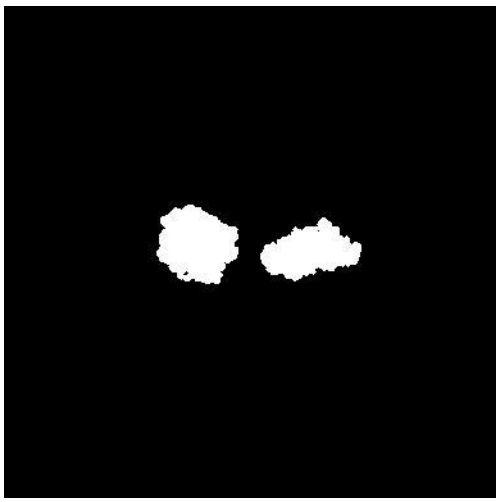
Orthogonal slices of maximum variance of the mask with hard borders

Explanation:
The mask with hard borders has been calculated at the suggested threshold 4.200000, the largest connected component was selected, and then dilated by 2Å.

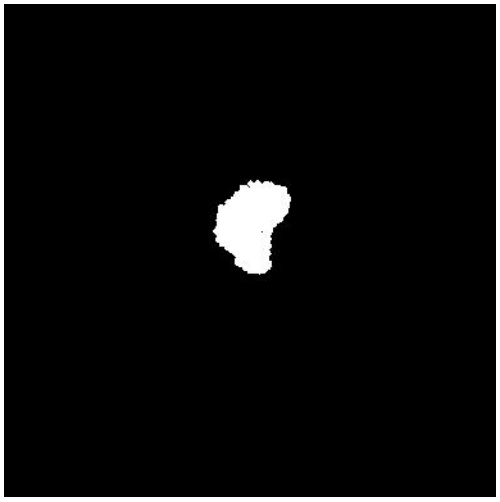
Results:
See Fig. 5.



(a) X Slice 186



(b) Y Slice 177



(c) Z Slice 206

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 4.200000, the largest connected component was selected, and then dilated by

2Å.

Results:
See Fig. 6.

(a) X Slice 187 (b) Y Slice 177

(c) Z Slice 206

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 30Å (the exact size depends on the CTF) on each side to make sure that the CTF can be appropriately corrected.

Results:
The space from the left and right in X are 164.16 and 170.56 Å, respectively. There is a decentering ratio (abs(Right-Left)/Size)% of 1.56%

The space from the left and right in Y are 141.78 and 147.11 Å, respectively. There is a decentering ratio (abs(Right-Left)/Size)% of 1.30%

The space from the left and right in Z are 123.66 and 105.53 Å, respectively. There is a decentering ratio (abs(Right-Left)/Size)% of 4.43%

The center of mass is at (x,y,z)=(894.13,-1028.64,-42.26). The decentering of the center of mass (abs(Center)/Size)% is 182.85, 317.87, and 61.01, 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, there total mass should not exceed 10%. The raw mask (just thresholding) and the mask constructed for the analysis (thresholding + largest connected component + dilation) should significantly overlap. Overlap is defined by the overlapping coefficient (size(Raw AND Constructed)/size(Raw)) that is a number between 0 and 1, the closer to 1, the more they agree.

Results:

Raw mask: At threshold 4.200000, there are 1143 connected components

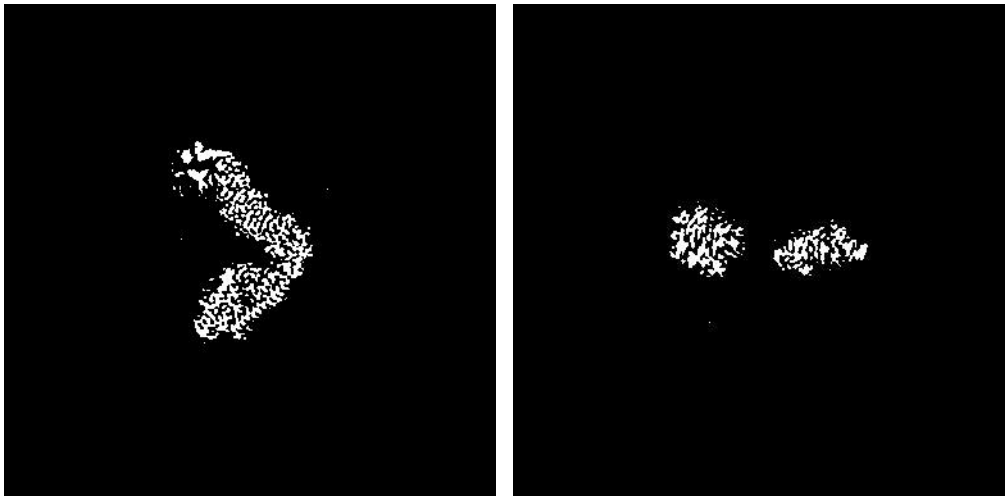
with a total number of voxels of 113349 and a volume of 137305.93 Å^3 (see Fig. 7). The size and percentage of the total number of voxels for the raw mask are listed below (up to 95% of the mass or the first 100 clusters, whatever happens first), the list contains (No. voxels (volume in Å^3), percentage, cumulated percentage):

(109665 (132843.30), 96.75, 96.75)

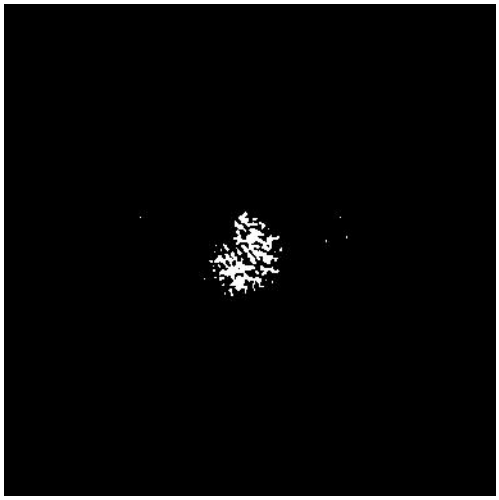
Number of components to reach 95% of the mass: 1

The average size of the remaining 1142 components is 3.23 voxels (1.21 Å^3). Their size go from 109665 voxels (132843.30 Å^3) to 1 voxels (1.21 Å^3).

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



(a) X Slice 185 (b) Y Slice 177



(c) Z Slice 157

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
4.7494	103661.00	104.04	945.77
9.4988	58306.00	58.52	531.97
14.2483	37494.00	37.63	342.08
18.9977	25531.00	25.62	232.94
23.7471	17929.00	17.99	163.58
28.4965	12682.00	12.73	115.71
33.2460	8819.00	8.85	80.46
37.9954	6156.00	6.18	56.17
42.7448	4255.00	4.27	38.82
47.4942	2871.00	2.88	26.19
52.2437	1970.00	1.98	17.97
56.9931	1339.00	1.34	12.22
61.7425	868.00	0.87	7.92
66.4919	561.00	0.56	5.12
71.2414	344.00	0.35	3.14
75.9908	207.00	0.21	1.89
80.7402	121.00	0.12	1.10
85.4896	81.00	0.08	0.74
90.2391	55.00	0.06	0.50
94.9885	37.00	0.04	0.34
99.7379	21.00	0.02	0.19
104.4873	6.00	0.01	0.05
109.2368	3.00	0.00	0.03
113.9862	2.00	0.00	0.02

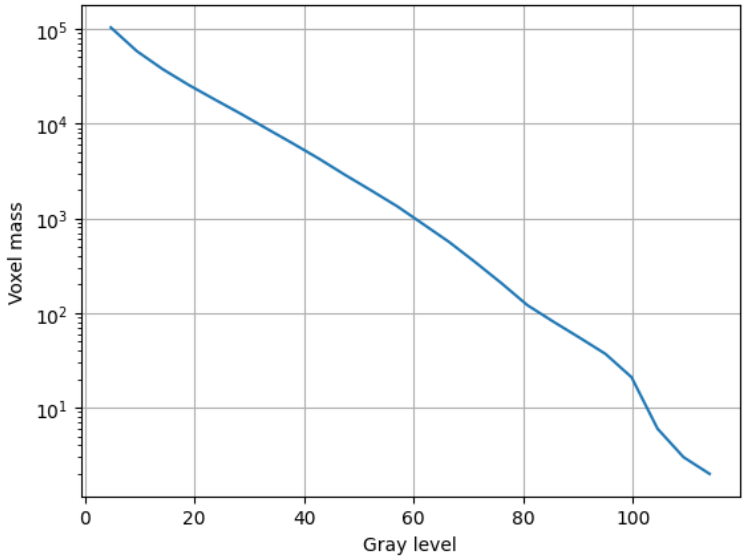


Figure 8: Voxel mass as a function of the gray level.

Constructed mask: After keeping the largest component of the previous mask and dilating it by 2Å, there is a total number of voxels of 325800 and a volume of 394659.62 Å^3. The overlap between the raw and constructed mask is 0.97.

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

STATUS: OK

2.3 Level 0.c Background analysis

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

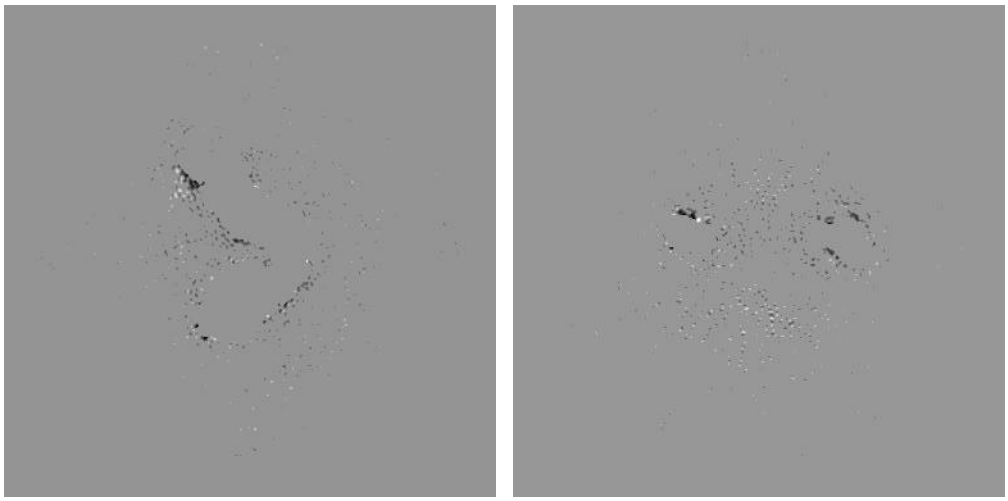
small and they should be randomly distributed without any specific structure. Sometimes, you can see some structure due to the symmetry of the structure.

Results:

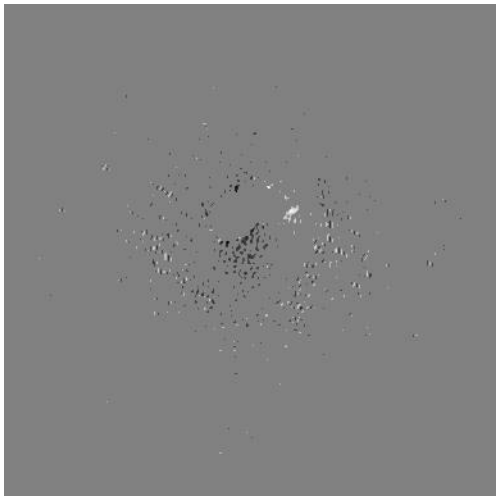
The null hypothesis that the background mean is 0 was tested with a one-sample Student's t-test. The resulting t-statistic and p-value were -236.27 and 0.000000, respectively.

The mean and standard deviation (sigma) of the background were -0.017541 and 0.557047. The percentage of background voxels whose absolute value is larger than 5 times the standard deviation is 0.26 % (see Fig. 9). The same percentage from a Gaussian would be 0.000057% (ratio between the two percentages: 4588.445130).

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



(a) X Slice 189 (b) Y Slice 156



(c) Z Slice 194

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=4588.45)**

2.4 Level 0.d B-factor analysis

Explanation:
The B-factor line (see this link for more details) fitted between 15Åand the resolution reported should have a slope that is between 0 and 300 $Å^2$.
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 -18.4. The fitted line was $\log(|F|^2) = -4.6/R^2 + (-7.8)$.

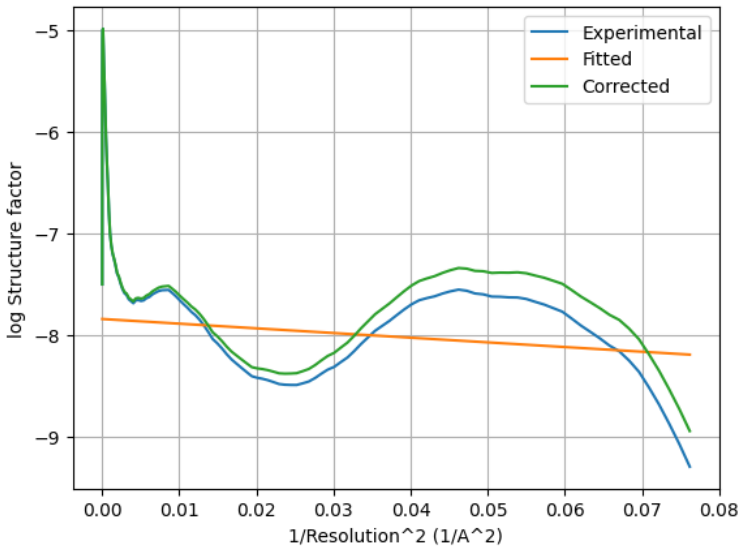
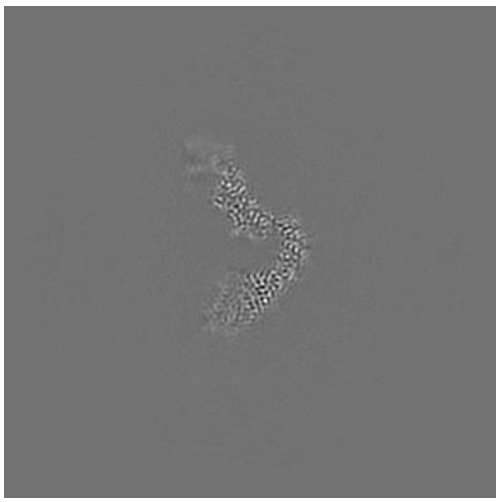
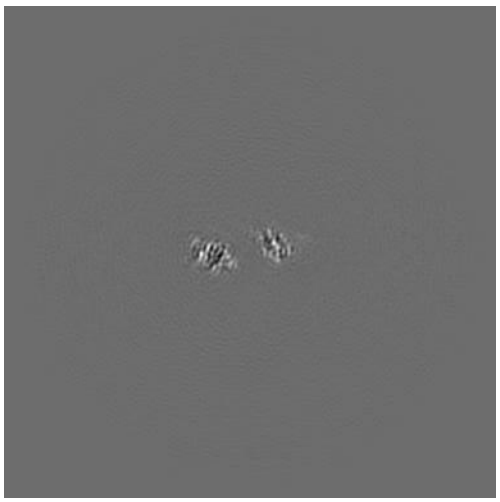


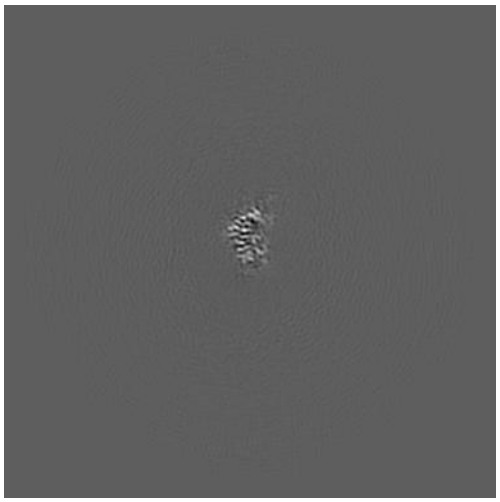
Figure 10: Guinier plot. The X-axis is the square of the inverse of the resolution in Å.



(a) X Slice 193



(b) Y Slice 203



(c) Z Slice 212

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].

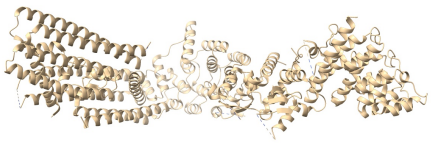
STATUS: OK

3 Atomic model

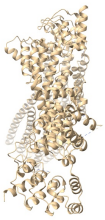
Atomic model: 7tad.cif

See Fig. 12.

(a) View 1



(b) View 2



(c) View 3

Figure 12: 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).

4 Level A Analysis

Map and model seem to be properly aligned.

4.1 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:

There are undefined residues (label_comp_id = N) and also hetero atoms which were deleted in order to perform this analysis.

To avoid ringing in Fourier space a smooth mask with a radius of 7.2 Å has been applied.

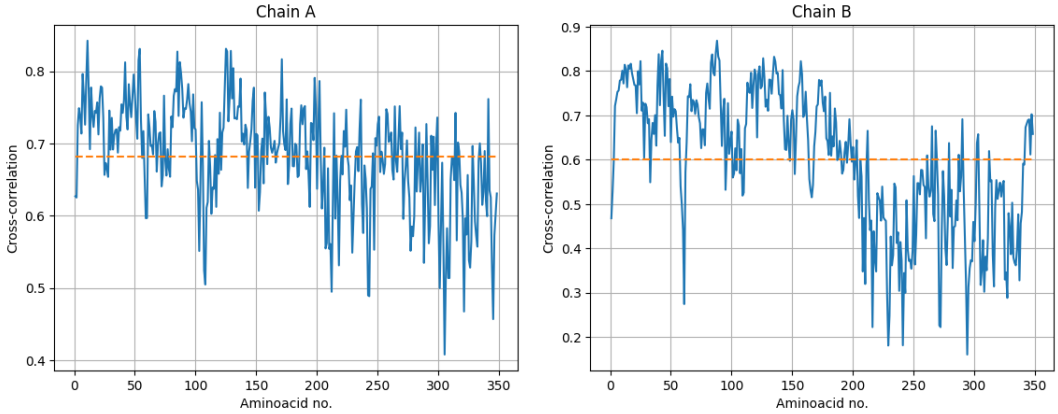
Overall correlation coefficients:

CC (mask) =	0.584
CC (box) =	0.600
CC (volume) =	0.604
CC (peaks) =	0.511
CC (main chain) =	0.586
CC (side chain) =	0.580

Correlation coefficients per chain:

Chain	**Cross-correlation**
A	0.645737
B	0.567134
C	0.587201
D	0.483789

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



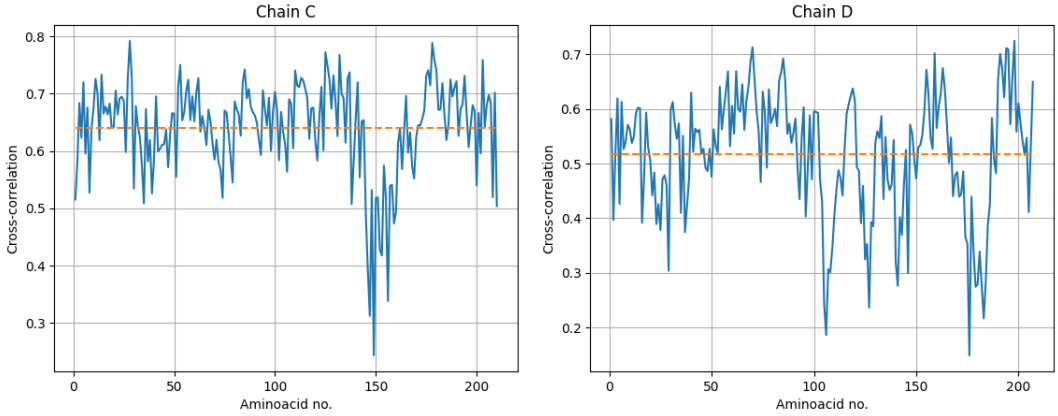


Fig. 13 shows the histogram of all cross-correlations evaluated at the residues. The percentage of residues whose correlation is below 0.5 is 16.4 %.

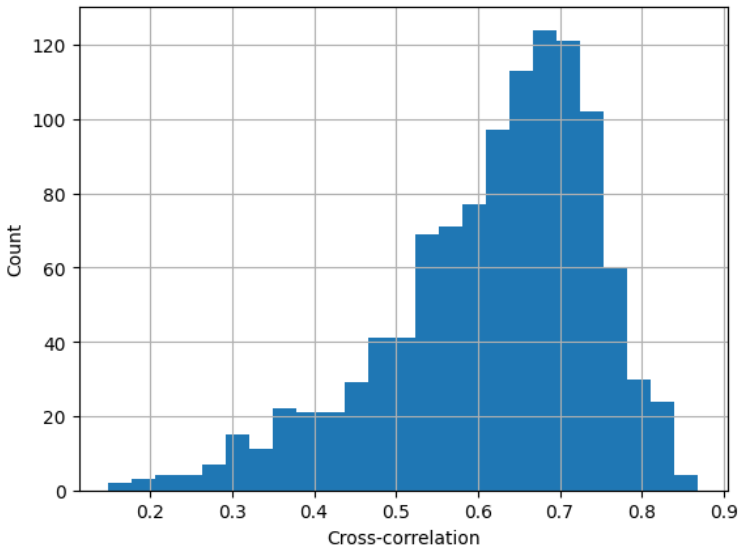


Figure 13: Histogram of the cross-correlation between the map and model evaluated for all residues.

Resolutions estimated from the model:

Resolution (Å)	**Masked**	**Unmasked**
d99	3.8	3.8
d_model	3.8	3.8
d_model (B-factor=0)	3.9	3.9
FSC_model=0	3.2	3.4
FSC_model=0.143	3.4	3.4
FSC_model=0.5	6.4	6.6

Overall isotropic B factor:

B factor	**Masked**	**Unmasked**
Overall B-iso	30.0	30.0

Fig. 14 shows the FSC between the input map and the model.

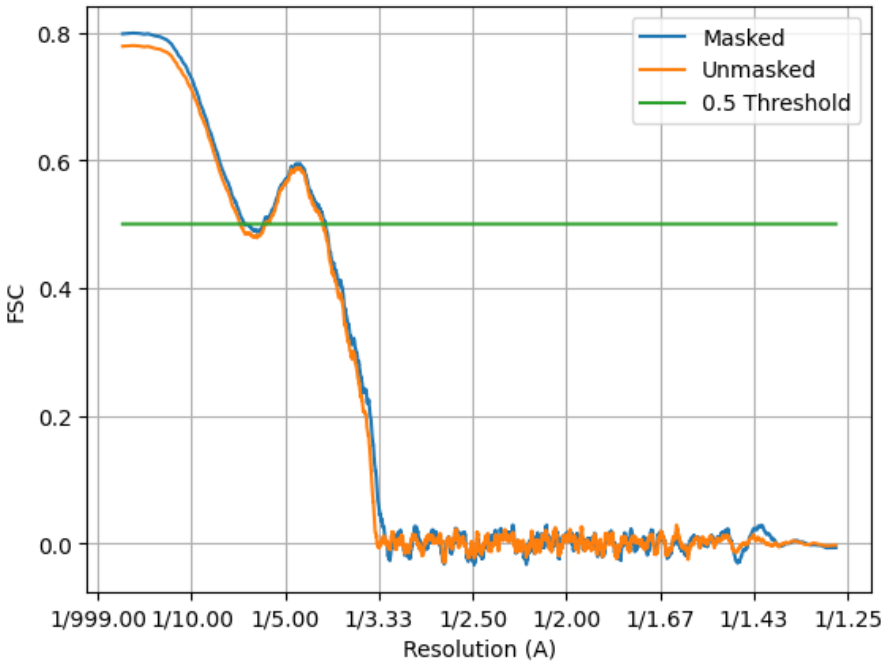


Figure 14: FSC between the input map and model with and without a mask constructed from the model. The X-axis is the square of the inverse of the resolution in Å.

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 reso-

lution reported by the user is larger than 0.8 times the resolution estimated between the map and model at FSC=0.5.

WARNINGS: 2 warnings

1. **The percentage of residues that have a cross-correlation below 0.5 is 16.4, that is larger than 10%**
2. **The resolution reported by the user, 3.6 Å, is significantly smaller than the resolution estimated between map and model (FSC=0.5), 6.4 Å**

4.2 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. 15 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 0.898.

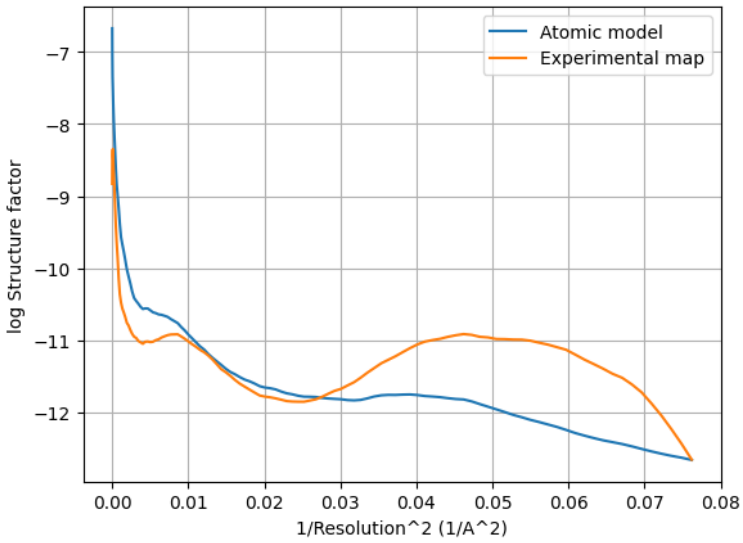


Figure 15: Guinier plot of the atom model and experimental map. The X-axis is the square of the inverse of the resolution in Å.

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

STATUS: OK

4.3 Level A.e 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. 16 shows the histogram of the calculated Q-score. Some representative percentiles are:

Percentile	MapQ score [-1,1]
2.5%	-0.24
25%	0.28
50%	0.49
75%	0.64
97.5%	0.80

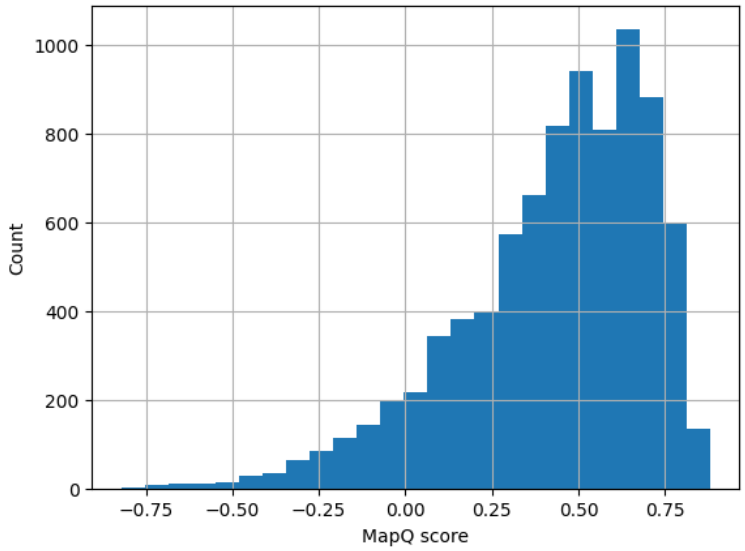
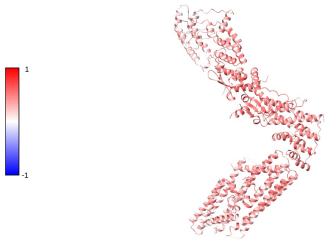


Figure 16: Histogram of the per-atom Q-score.

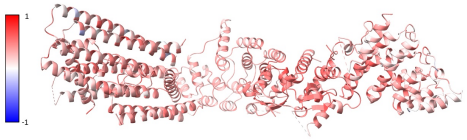
The following table shows the average Q-score and estimated resolution for each chain.

Chain	Average Q-score [-1,-1]	Estimated Resol. (Å)
A	0.50	1.5
B	0.40	1.5
C	0.44	1.5
D	0.35	1.5

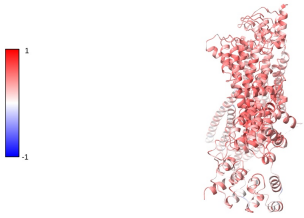
The atomic model colored by MapQ can be seen in Fig. 17.



(a) View 1



(b) View 2



(c) View 3

Figure 17: Atomic model colored by MapQ 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 Q-score is larger than 0.35, which is the threshold for the resolution of 3.6 Å.

STATUS: OK

4.4 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	57.239374
Rotamer ratio	0.857
Max. Zscore	3.75
Model length	698
EMRinger Score	1.418

Fig. 18 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.

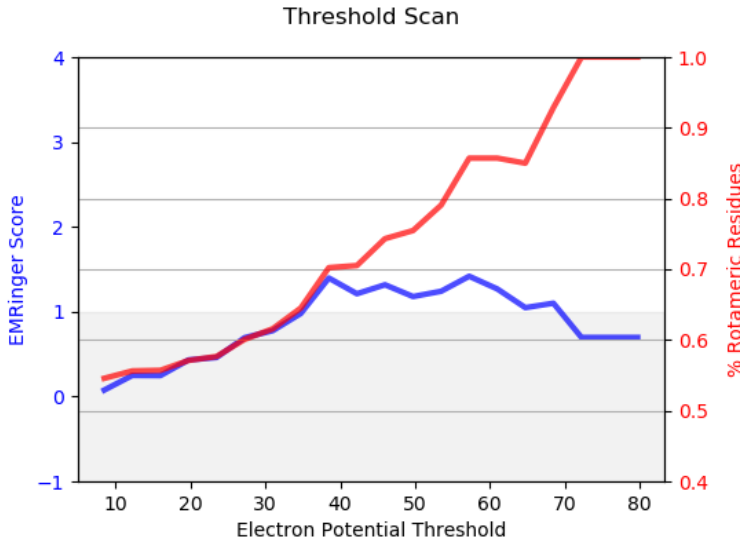


Figure 18: EMRinger score and fraction of rotameric residues as a function of the map threshold.

Fig. 19 shows the histogram for rotameric (blue) and non-rotameric (red) residues at the optimal threshold.

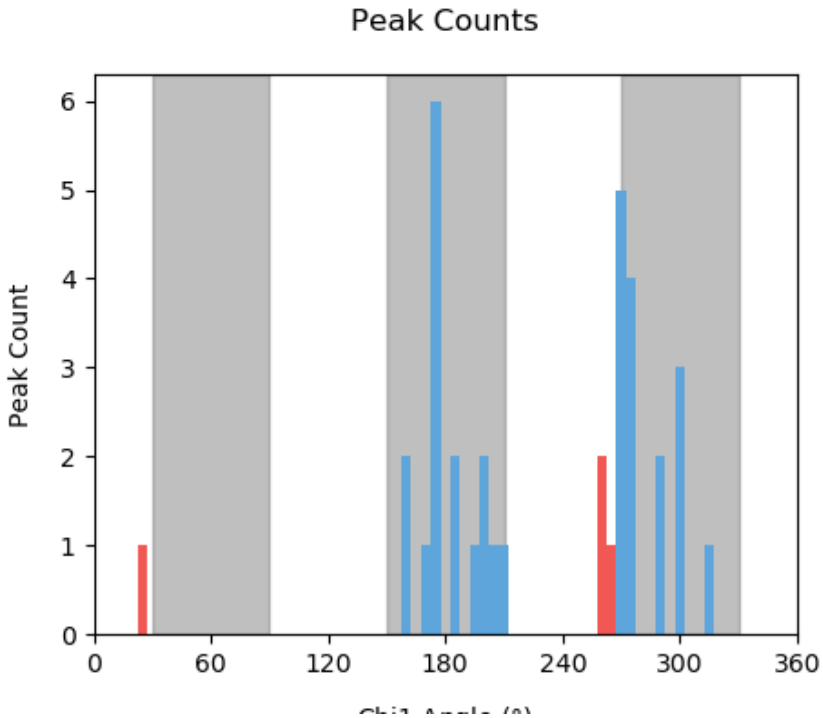
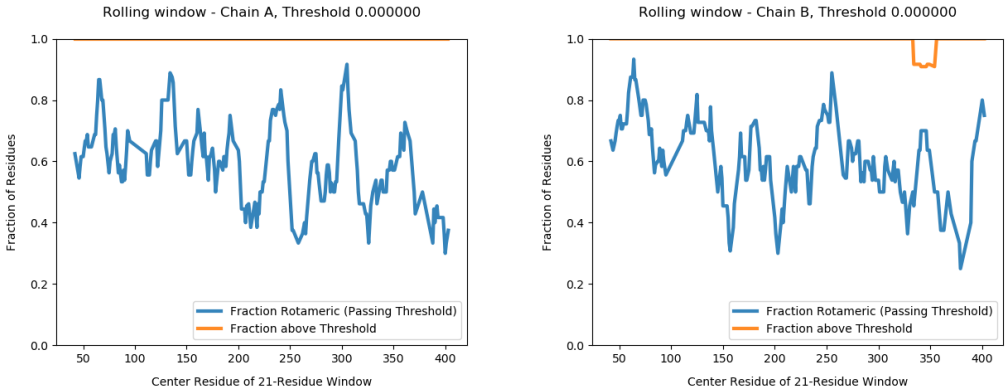
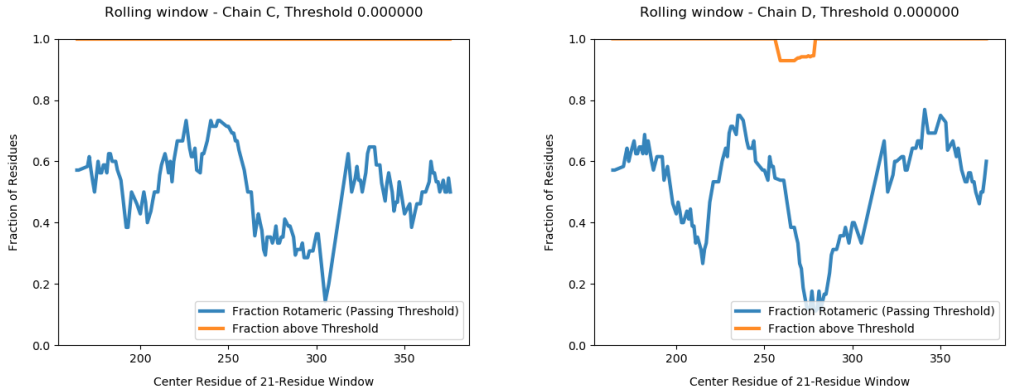


Figure 19: Histogram for rotameric (blue) and non-rotameric (red) residues at the optimal threshold as a function of the angle Chi1.

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. If straight lines are observed in the plots, this is likely due to the absence of side chains in those residues.





Automatic criteria: The validation is OK if the EMRinger score and Max. Zscore are larger than 1.

STATUS: OK

4.5 Level A.g DAQ validation

Explanation:
DAQ (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:
Fig. 20 shows the histogram of the DAQ values. The mean and standard deviation were 0.1 and 0.3, respectively.

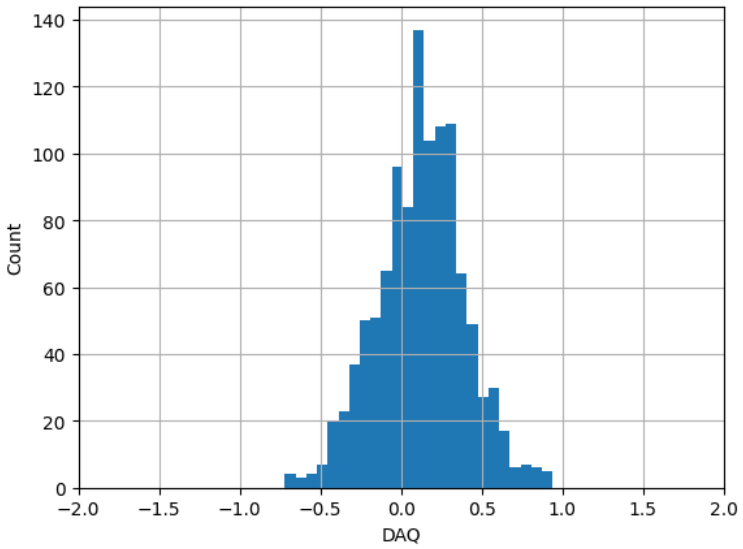
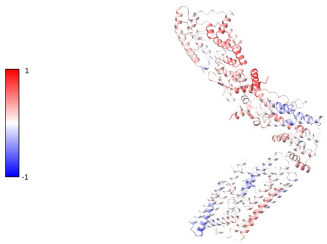
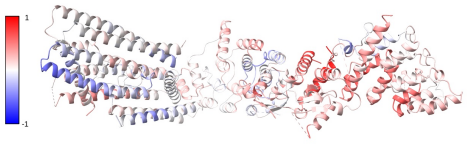


Figure 20: Histogram of the per-residue DAQ values.

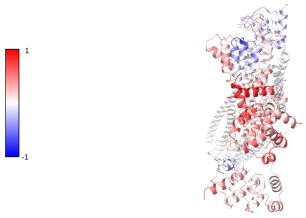
The atomic model colored by DAQ can be seen in Fig. 21.



(a) View 1



(b) View 2



(c) View 3

Figure 21: 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.

WARNINGS: 1 warnings

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