Public Data-based Report

Validation Report Service

Cryo-EM Map Validation Report

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

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

Basic Entry Information:

EMDB ID: EMD-37176
Title: Apoferritin (4K counting mode from 32 Falcon 4 micrographs)
Authors: See EMDB entry link
Deposited on: 2023-08-13T00:00:00
Reported Resolution: 3.824 Å

Contact Us:

Instruct Image Processing Center (I²PC)
Biocomputing Unit (BCU)
i2pc@cnb.csic.es
VRS Website

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

Last update: **August 21, 2026, 10:24am**

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.

SCIPION

Summarized overall quality

The map seems to be well centered. There seems to be a problem with the suggested threshold (see Sec. 2.2). There seems to be a problem with the map's background (see Sec. 2.3).

The overall score (passing tests) of this report is 2 out of 4 evaluable items.

0.a Mass analysis	Sec. 2.1	OK
0.b Mask analysis	Sec. 2.2	1 warnings
0.c Background analysis	Sec. 2.3	2 warnings
0.d B-factor analysis	Sec. 2.4	OK

Summary of the warnings across sections.

Section 2.2 (0.b Mask analysis)

1. **There might be a problem with noise and artifacts, because the average noise blob has a volume of 6.989783 Å^3.**

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=6793.21)**

Contents

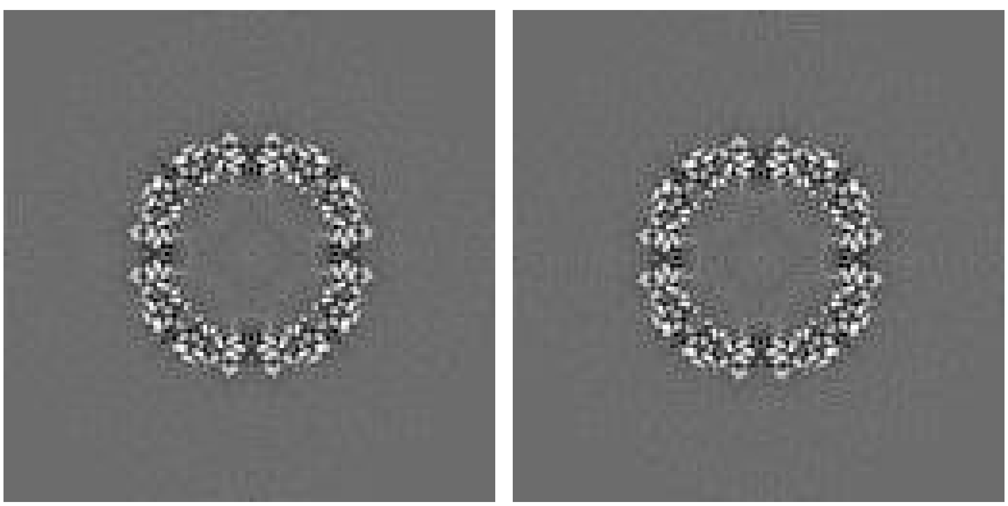
1 Input data

Input map: emd_37176.map
SHA256 hash: 3489ce4dfd7219a678e35d6926c914975a11d92f8dd926f829363cda4270d8ce
Voxel size: 1.912000 (Å)
Visualization threshold: 0.073000
Resolution estimated by user: 3.824

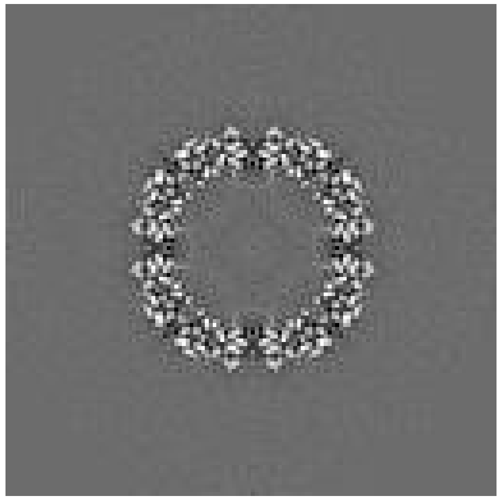
Orthogonal slices of the input map

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 64 (b) Y Slice 64

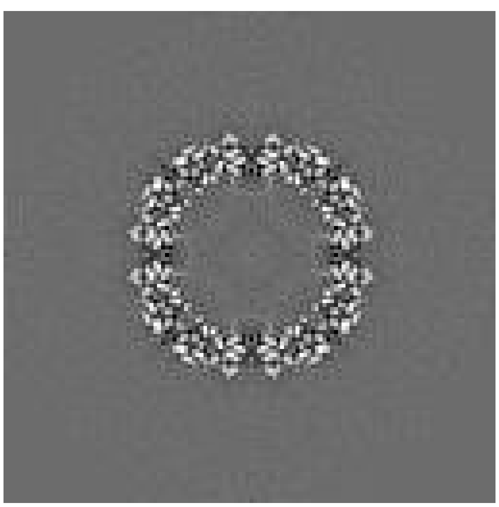


(c) Z Slice 64

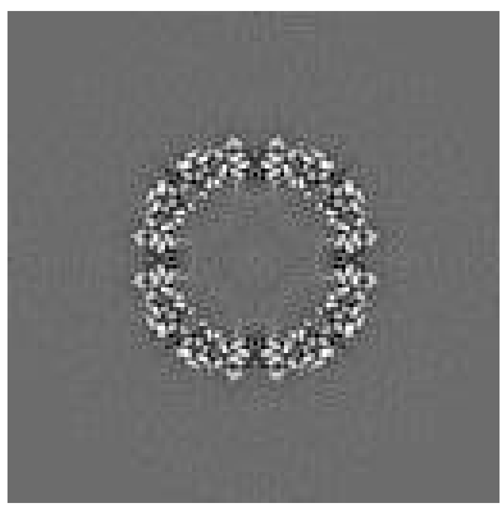
Figure 1: Central slices of the input map in the three dimensions

Orthogonal slices of maximum variance of the input map

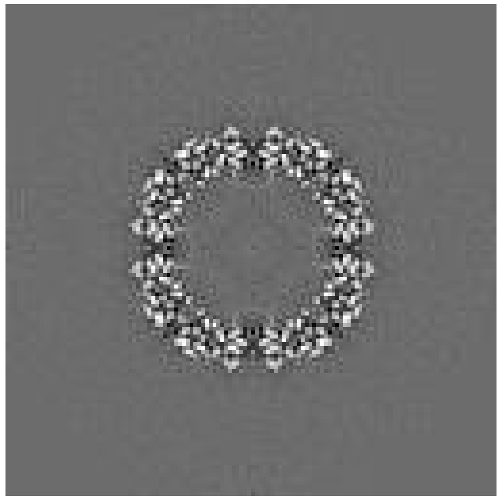
Results:
See Fig. 2.



(a) X Slice 64



(b) Y Slice 64



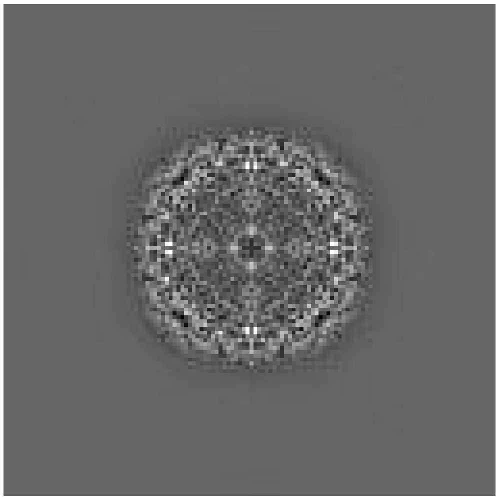
(c) Z Slice 64

Figure 2: Slices of maximum variation in the three dimensions

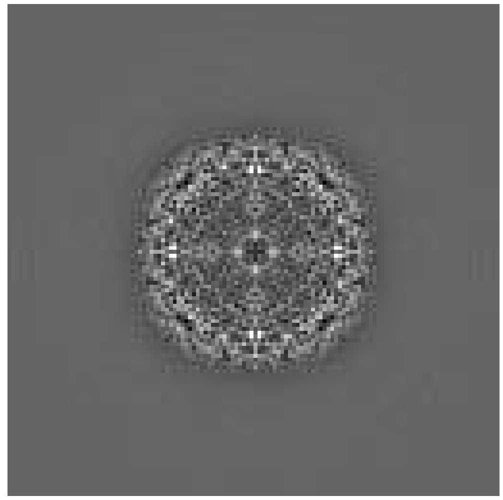
Orthogonal projections of the input map

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

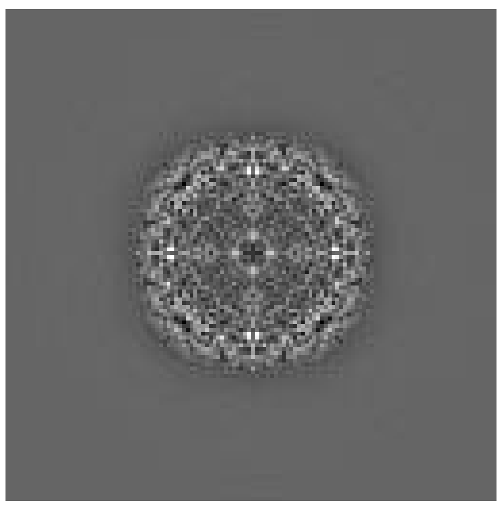
Results:
See Fig. 3.



(a) X Projection



(b) Y Projection



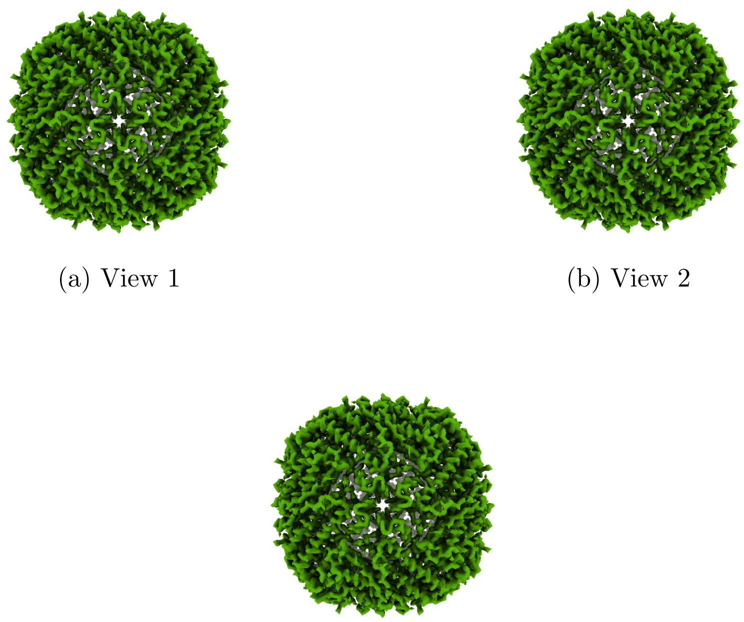
(c) Z Projection

Figure 3: Projections in the three dimensions

Isosurface views of the input map

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

Results:
See Fig. 4.



(a) View 1

(b) View 2

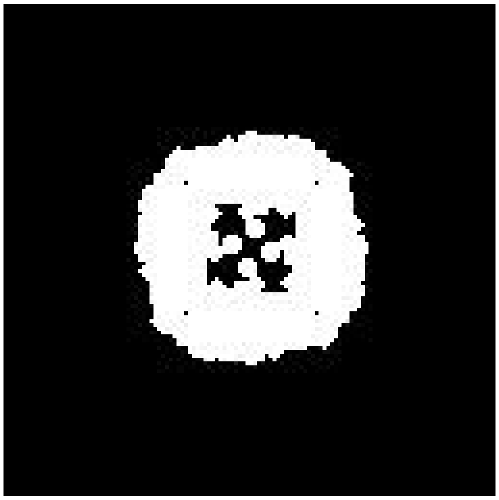
(c) View 3

Figure 4: Isosurface at threshold=0.073000. Views generated by ChimeraX at a the following X, Y, Z angles: View 1 (0, -90, -90), View 2 (-90, 0, -90), View 3 (0, 0, 0).

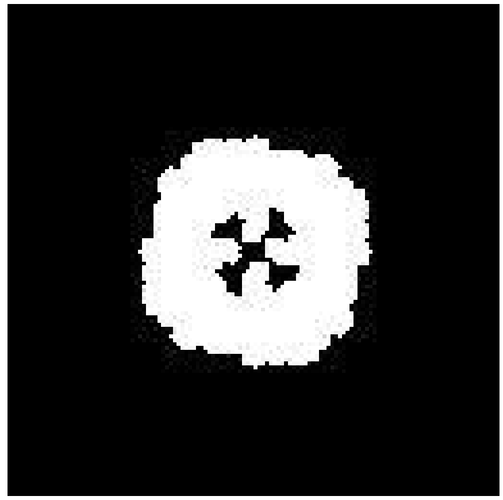
Orthogonal slices of maximum variance of the mask with hard borders

Explanation:
The mask with hard borders has been calculated at the suggested threshold 0.073000, the largest connected component was selected, and then dilated by 2Å.

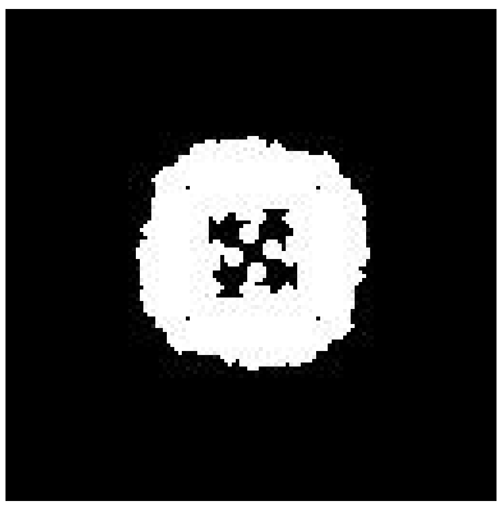
Results:
See Fig. 5.



(a) X Slice 47



(b) Y Slice 46



(c) Z Slice 81

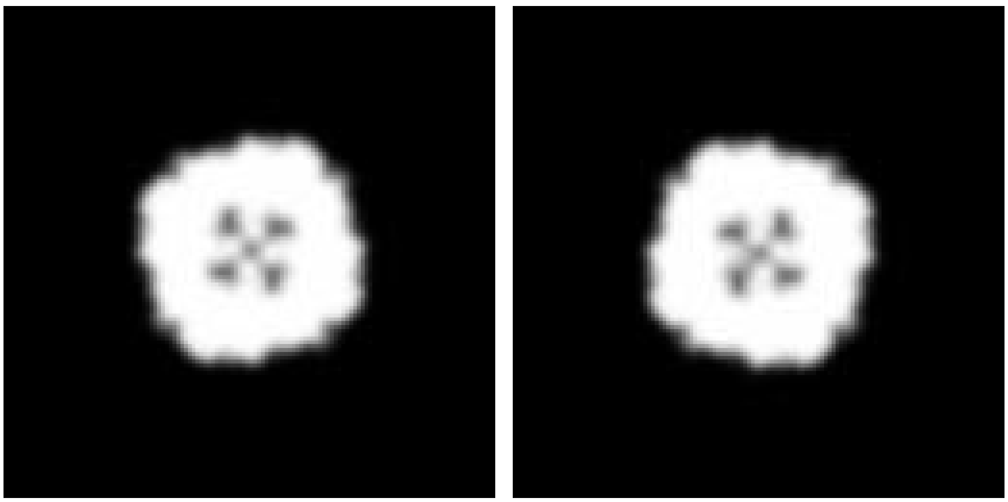
Figure 5: Slices of maximum variation in the three dimensions of the mask with hard borders

Orthogonal slices of maximum variance of the mask with soft borders

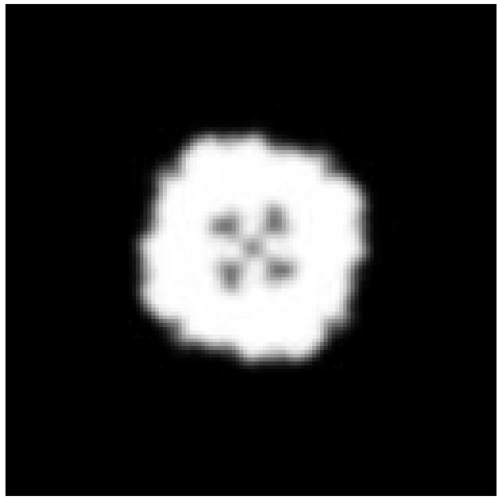
Explanation:
The mask with soft borders has been calculated at the suggested threshold 0.073000, the largest connected component was selected, and then dilated by

2Å.

Results:
See Fig. 6.



(a) X Slice 45 (b) Y Slice 45



(c) Z Slice 83

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 59.27 and 59.27 Å, respectively. There is a decentering ratio (abs(Right-Left)/Size)% of 0.00%

The space from the left and right in Y are 59.27 and 59.27 Å, respectively. There is a decentering ratio (abs(Right-Left)/Size)% of 0.00%

The space from the left and right in Z are 59.27 and 59.27 Å, respectively. There is a decentering ratio (abs(Right-Left)/Size)% of 0.00%

The center of mass is at (x,y,z)=(61.94, 61.94, 61.94). The decentering of the center of mass (abs(Center)/Size)% is 1.61, 1.61, and 1.61, respectively.

Automatic criteria: The validation is OK if 1) the decentering and center of mass less than 20% of the map dimensions in all directions, and 2) the extra space on each direction is more than 20% of the map dimensions. For local reconstruction, focused refinement, or similar, warnings are expected.

STATUS: OK

2.2 Level 0.b Mask analysis

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

Results:

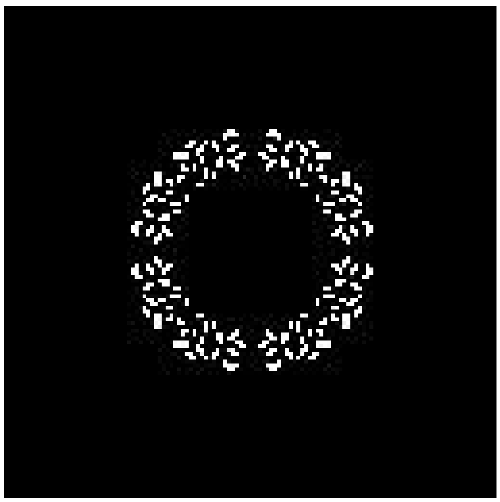
Raw mask: At threshold 0.073000, there are 13 connected components with a total number of voxels of 20484 and a volume of 143178.71 Å^3 (see Fig. 7). The size and percentage of the total number of voxels for the raw mask are listed below (up to 95% of the mass or the first 100 clusters, whatever happens first), the list contains (No. voxels (volume in Å^3), percentage, cumulated percentage):

(20448 (142927.07), 99.82, 99.82)

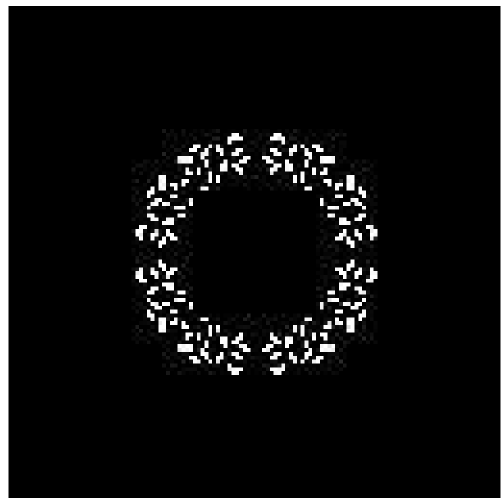
Number of components to reach 95% of the mass: 1

The average size of the remaining 12 components is 3.00 voxels (6.99 Å^3). Their size go from 20448 voxels (142927.07 Å^3) to 3 voxels (20.97 Å^3).

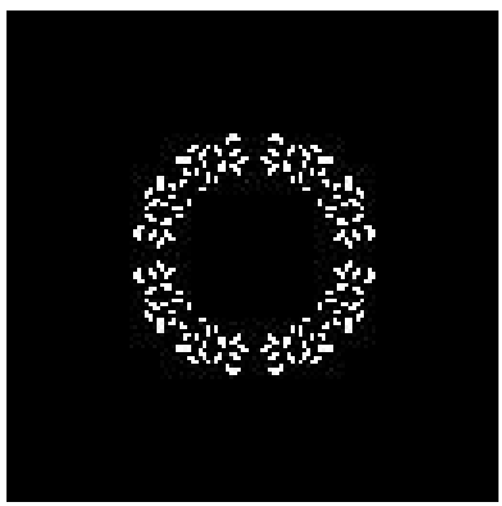
The slices of the raw mask can be seen in Fig. 7.



(a) X Slice 64



(b) Y Slice 64



(c) Z Slice 64

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**
0.0108	54676.00	316.63	2878.46
0.0216	37168.00	215.24	1956.74
0.0324	32150.00	186.18	1692.56
0.0432	28418.00	164.57	1496.09
0.0539	25508.00	147.72	1342.89
0.0647	22572.00	130.72	1188.32
0.0755	20076.00	116.26	1056.92
0.0863	18012.00	104.31	948.26
0.0971	16032.00	92.84	844.02
0.1079	14184.00	82.14	746.73
0.1187	12240.00	70.88	644.38
0.1295	10536.00	61.01	554.68
0.1402	8688.00	50.31	457.39
0.1510	7152.00	41.42	376.52
0.1618	5616.00	32.52	295.66
0.1726	4416.00	25.57	232.48
0.1834	3408.00	19.74	179.42
0.1942	2352.00	13.62	123.82
0.2050	1488.00	8.62	78.34
0.2158	840.00	4.86	44.22
0.2265	312.00	1.81	16.43
0.2373	120.00	0.69	6.32
0.2481	72.00	0.42	3.79
0.2589	24.00	0.14	1.26

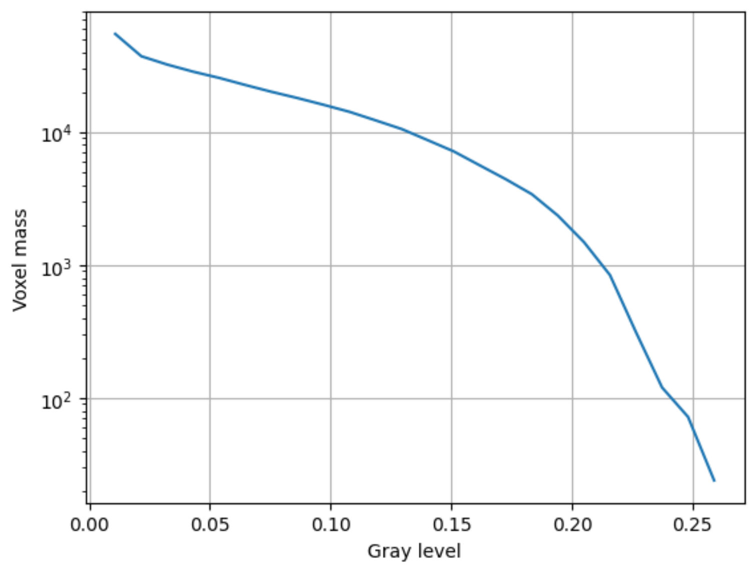


Figure 8: Voxel mass as a function of the gray level.

Constructed mask: After keeping the largest component of the previous mask and dilating it by 2Å, there is a total number of voxels of 134418 and a volume of 939552.59 Å^3. The overlap between the raw and constructed mask is 1.00.

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

WARNINGS: 1 warnings

1. **There might be a problem with noise and artifacts, because the average noise blob has a volume of 6.989783 Å^3.**

2.3 Level 0.c Background analysis

Explanation:
Background is defined as the region outside the macromolecule mask. The

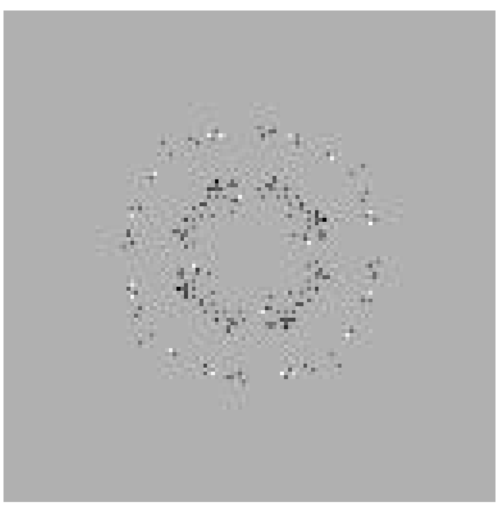
background mean should be zero, and the number of voxels with a very low or very high value (below 5 standard deviations of the noise) should be very small and they should be randomly distributed without any specific structure. Sometimes, you can see some structure due to the symmetry of the structure.

Results:

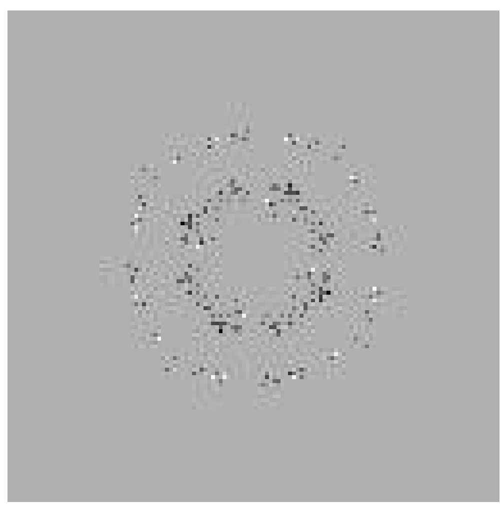
The null hypothesis that the background mean is 0 was tested with a one-sample Student's t-test. The resulting t-statistic and p-value were -116.31 and 0.000000, respectively.

The mean and standard deviation (sigma) of the background were -0.000334 and 0.004022. The percentage of background voxels whose absolute value is larger than 5 times the standard deviation is 0.39 % (see Fig. 9). The same percentage from a Gaussian would be 0.000057% (ratio between the two percentages: 6793.207953).

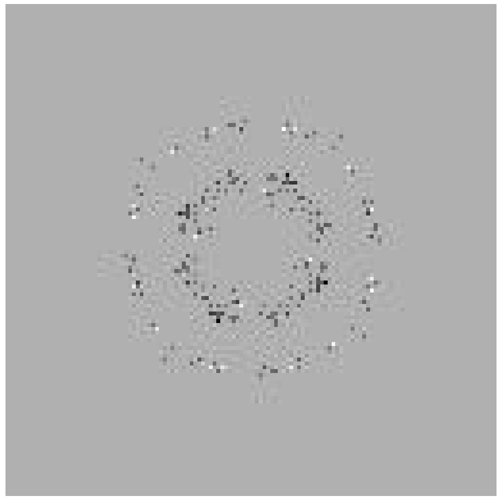
Slices of the background beyond 5*sigma can be seen in Fig. 9.



(a) X Slice 57



(b) Y Slice 57



(c) Z Slice 71

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=6793.21)**

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 -16.4. The fitted line was $\log(|F|^2) = -4.1/R^2 + (-11.0)$.

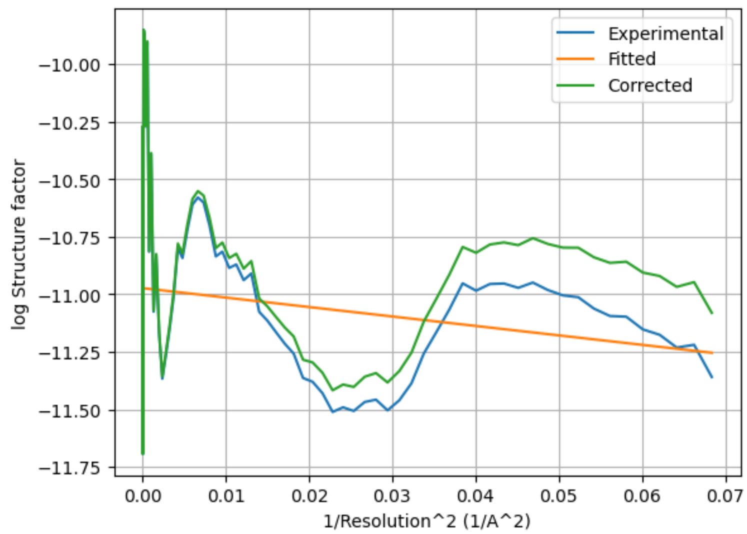
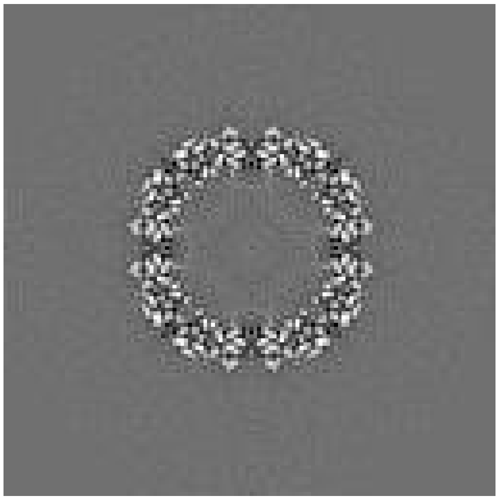
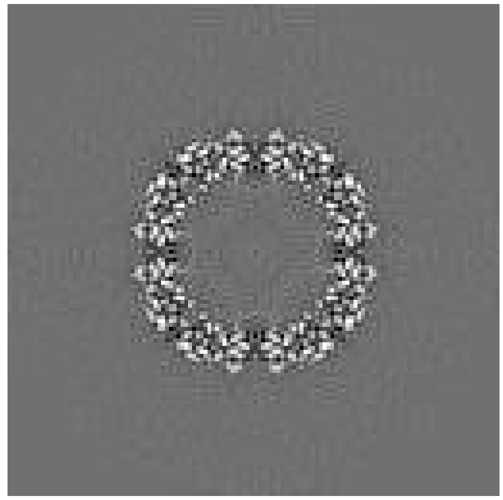


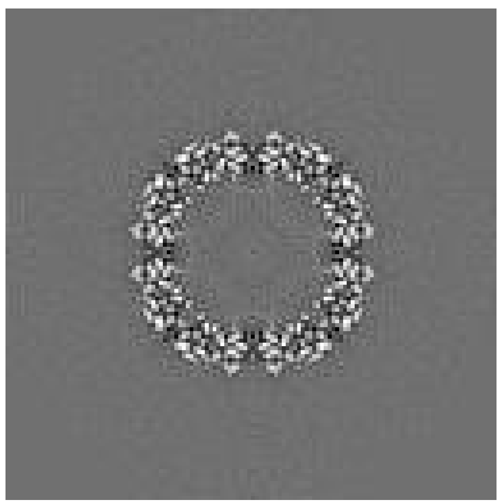
Figure 10: Guinier plot. The X-axis is the square of the inverse of the resolution in Å.



(a) X Slice 64



(b) Y Slice 64



(c) Z Slice 64

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].

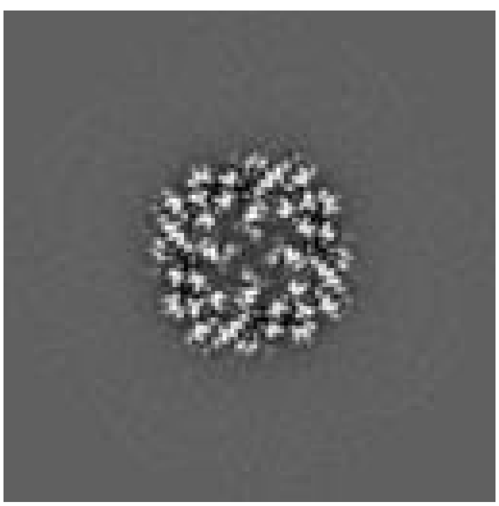
STATUS: OK

3 Half maps

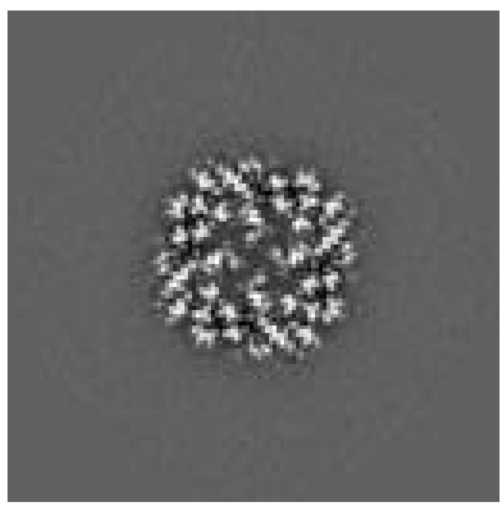
Half map 1: emd_37176_half_map_1.map
SHA256 hash: ee897580aab826cf4c3e62696115180f9da297c9818f7c1e191b3fc4358305b8

Half map 2: emd_37176_half_map_2.map
SHA256 hash: 361e606a9c26b386787ed17d382996d184d712e0e585751489cee722e0bf1309

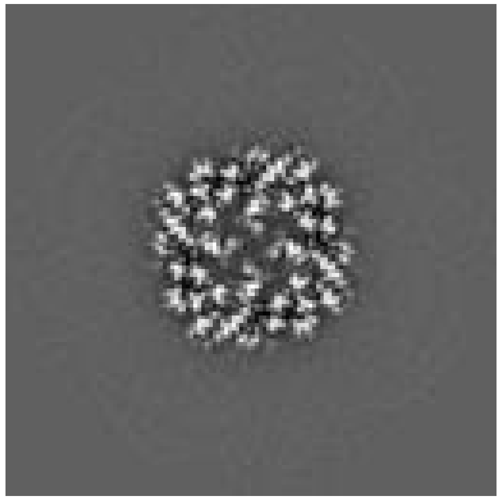
Slices of the first half map can be seen in Fig. 12.
Slices of the second half map can be seen in Fig. 13.
Slices of the difference between both maps can be seen in Fig. 14. There should not be any structure in this difference. Sometimes some patterns are seen if the map is symmetric.



(a) X Slice 44

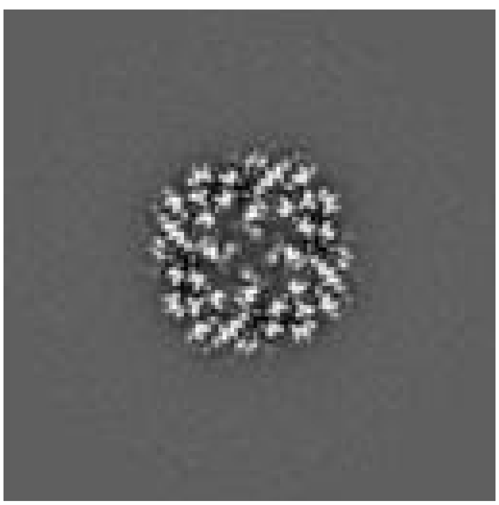


(b) Y Slice 44

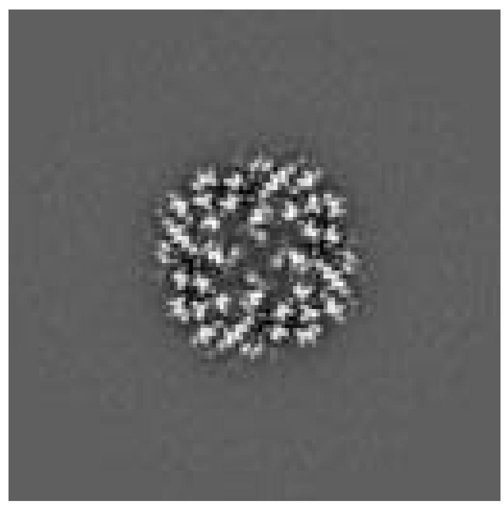


(c) Z Slice 44

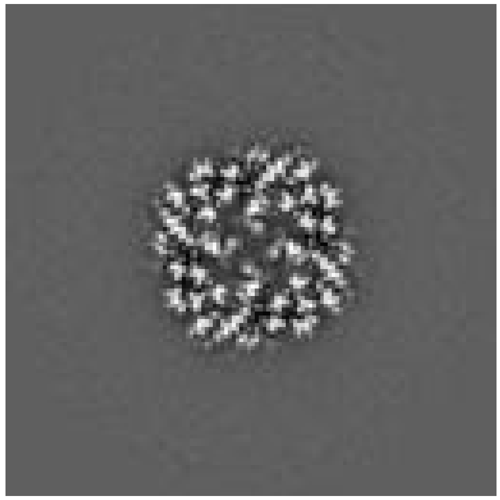
Figure 12: Slices of maximum variation in the three dimensions of Half 1



(a) X Slice 44

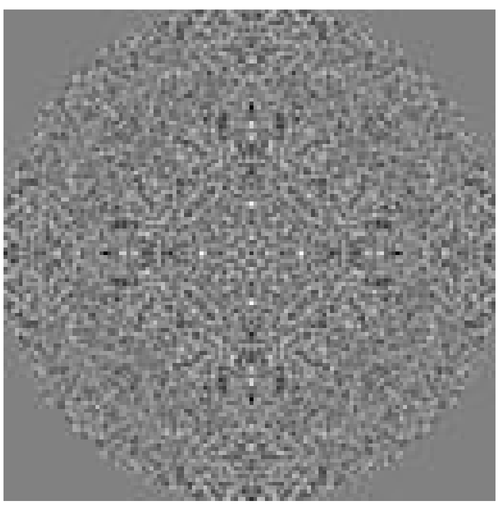


(b) Y Slice 84

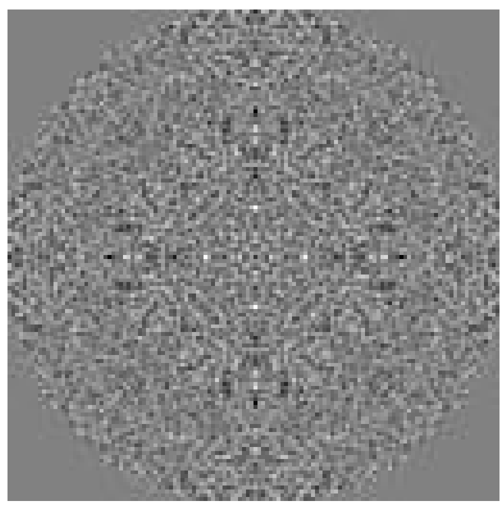


(c) Z Slice 44

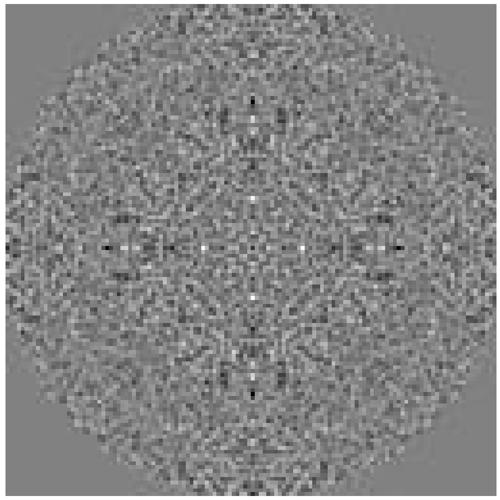
Figure 13: Slices of maximum variation in the three dimensions of Half 2



(a) X Slice 64



(b) Y Slice 64



(c) Z Slice 64

Figure 14: Slices of maximum variation in the three dimensions of the difference Half1-Half2.