



Public Data-based Report



Validation Report Service

Cryo-EM Map Validation Report

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

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

Basic Entry Information:

EMDB ID: [EMD-21702](#)

Title: Empty phage G cryoEM capsid structure at 9 Angstrom resolution

Authors: [See EMDB entry link](#)

Deposited on: 2020-04-14T00:00:00

Reported Resolution: 9.0 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, 7:47pm**

Context

Cryo-electron microscopy is currently one of the most active techniques in Structural Biology. The number of maps deposited at the [Electron Microscopy Data Bank](#) is rapidly growing every year and keeping the quality of the submitted maps is essential to maintain the scientific quality of the field. The ultimate quality measure is the consistency of the map and an atomic model. However, this is only possible for high resolution maps. Over the years there have been many suggestions about validation measures of 3DEM maps. Unfortunately, most of these measures are not currently in use for their spread in multiple software tools and the associated difficulty to access them. To alleviate this problem, we made available a validation grading system that evaluate the information provided to assess the map.

This system grades a map from 0 to 5 depending on the amount of information available. In this way, a map could be validated at Level 0 (the deposited map), 1 (two half maps), 2 (2D classes), 3 (particles), 4 (... + angular assignment), 5 (... + micrographs and coordinates). In addition, we can have three optional qualifiers: A (... + atomic model), W (... + image processing workflow), and O (... + other techniques). To know more about this service read this [paper](#)

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



Summarized overall quality

The map seems to be well centered. There 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 1 out of 3 evaluable items.

0.a Mass analysis	Sec. 2.1	OK
0.b Mask analysis	Sec. 2.2	3 warnings
0.c Background analysis	Sec. 2.3	2 warnings
0.d B-factor analysis	Sec. 2.4	Does not apply

Summary of the warnings across sections.

Section 2.2 (0.b Mask analysis)

1. **There might be a problem of connectivity at this threshold because more than 5 connected components are needed to reach 95% of the total mask. Probably a smaller threshold will not cause this issue.**
2. **There might be a problem with noise and artifacts, because the average noise blob has a volume of 24.389008 Å³.**
3. **There might be a problem in the construction of the mask, because the overlap is smaller than 0.75. A common reason is that the suggested threshold causes too many disconnected components.**

Section 2.3 (0.c Background analysis)

1. **The null hypothesis that the background mean is 0 has been rejected because the p-value of the comparison is smaller than 0.001**
2. **There is a significant proportion of outlier values in the background (cdf5 ratio=8061.52)**

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	19
2.4	Level 0.d B-factor analysis	21
3	Half maps	21

1 Input data

Input map: emd_21702.map

SHA256 hash: cb753eaf2bf777c1b545a8672108ecd80600180dce02ff3eaf970520a16216a

Voxel size: 2.900000 (Å)

Visualization threshold: 0.421000

Resolution estimated by user: 9.0

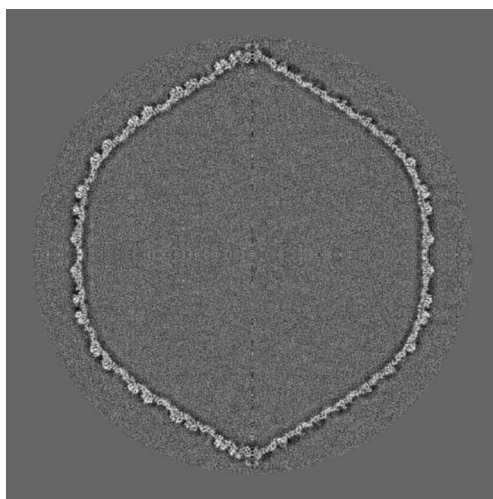
Orthogonal slices of the input map

Explanation:

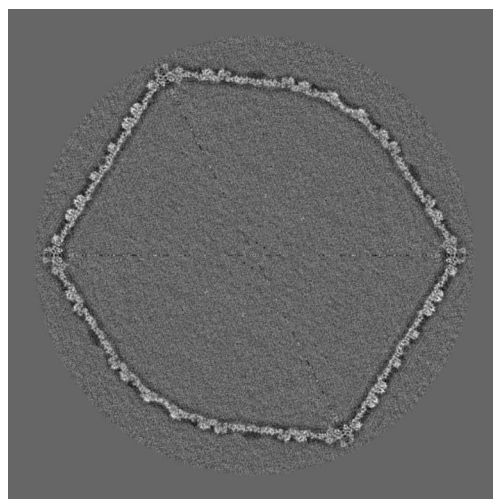
In the orthogonal slices of the map, the noise outside the protein should not have any structure (stripes going out, small blobs, particularly high or low densities, ...)

Results:

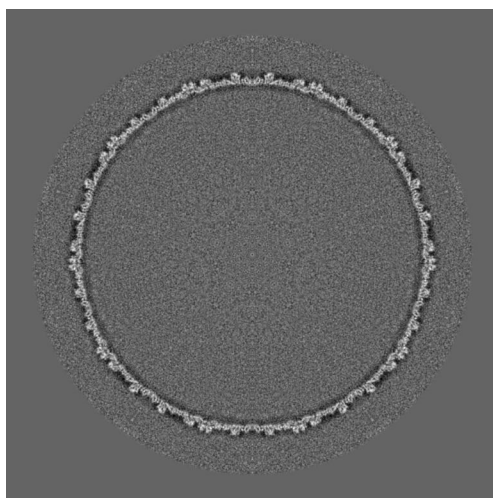
See Fig. 1.



(a) X Slice 384



(b) Y Slice 384



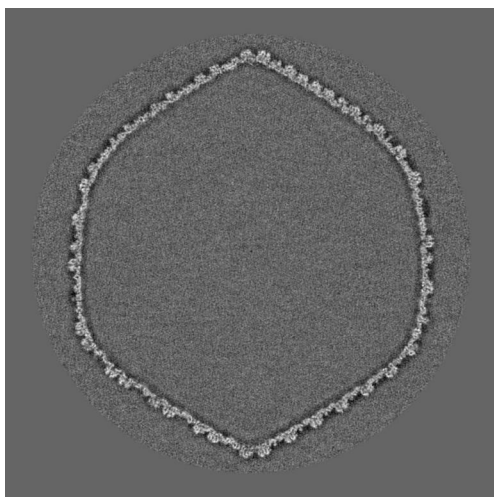
(c) Z Slice 384

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

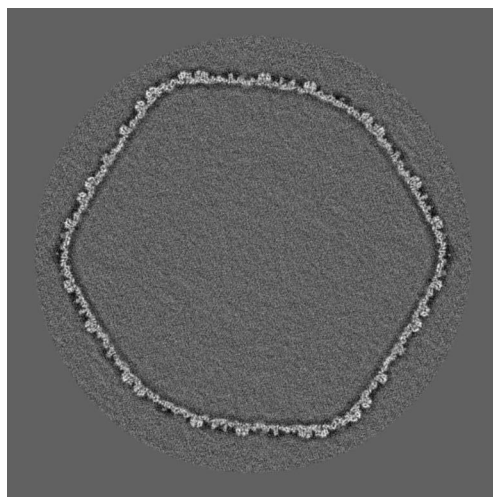
Orthogonal slices of maximum variance of the input map

Results:

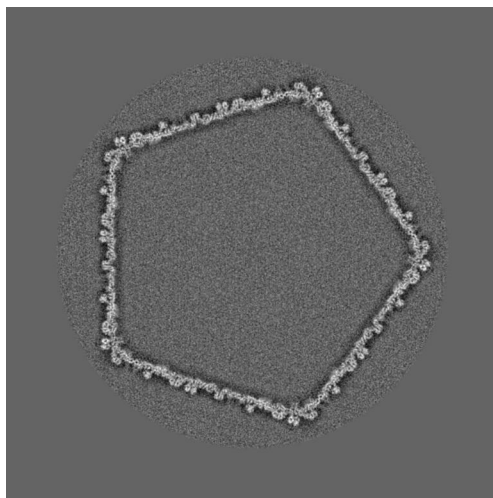
See Fig. 2.



(a) X Slice 365



(b) Y Slice 347



(c) Z Slice 541

Figure 2: Slices of maximum variation in the three dimensions

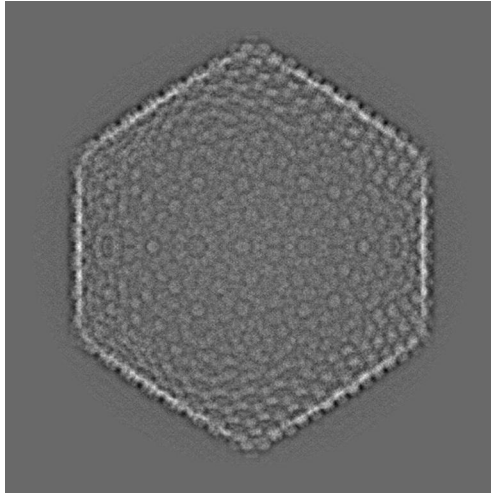
Orthogonal projections of the input map

Explanation:

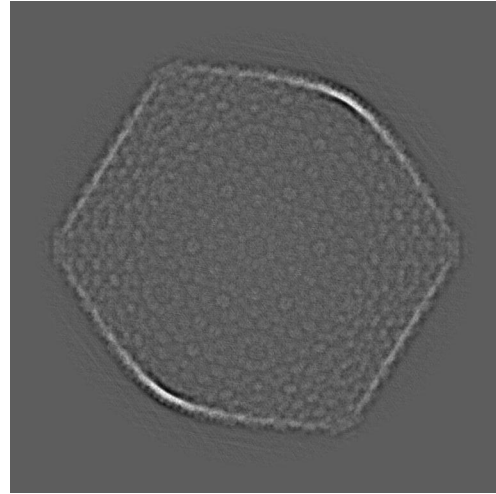
In the projections there should not be stripes (this is an indication of directional overweighting, or angular attraction), and there should not be a dark halo around or inside the structure (this is an indication of incorrect CTF correction or the reconstruction of a biased map).

Results:

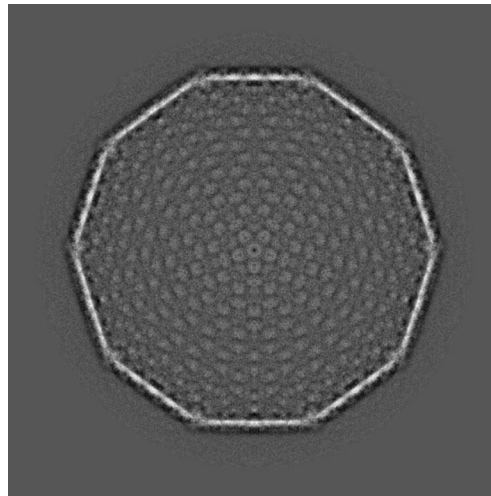
See Fig. [3](#).



(a) X Projection



(b) Y Projection



(c) Z Projection

Figure 3: Projections in the three dimensions

Isosurface views of the input map

Explanation:

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

Results:

See Fig. 4.

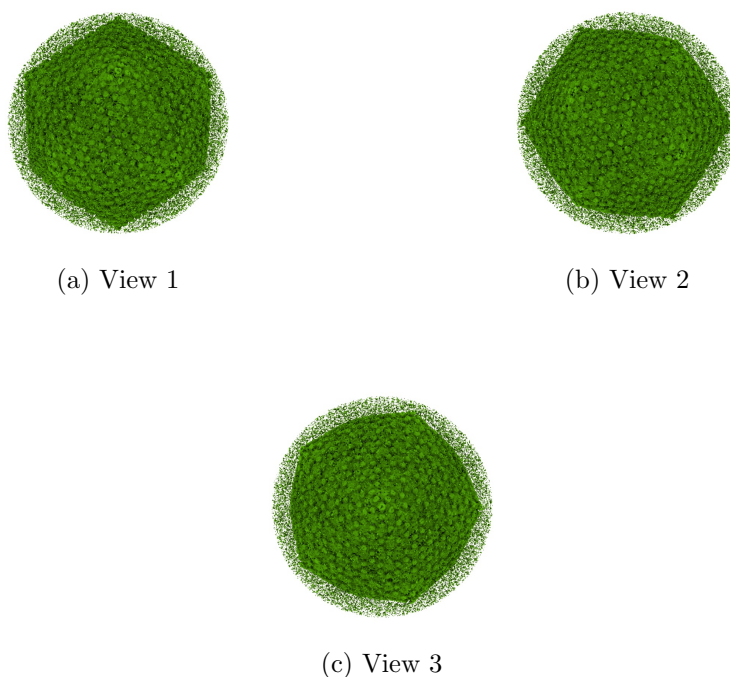


Figure 4: Isosurface at threshold=0.421000. Views generated by ChimeraX at a the following X, Y, Z angles: View 1 (0, -90, -90), View 2 (-90, 0, -90), View 3 (0, 0, 0).

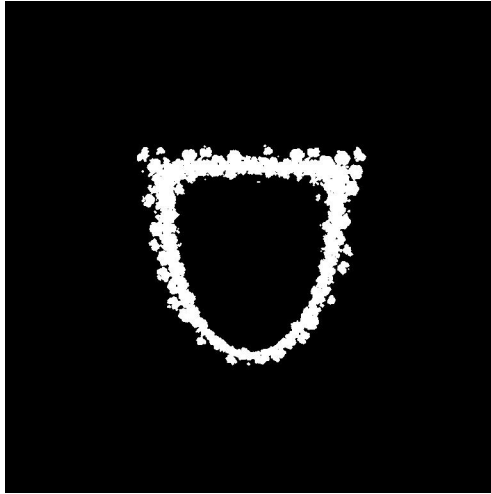
Orthogonal slices of maximum variance of the mask with hard borders

Explanation:

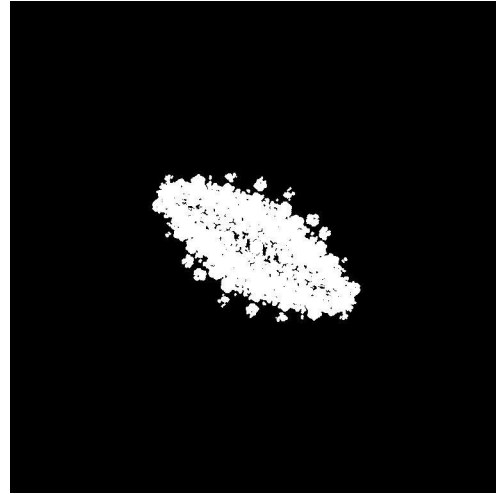
The mask with hard borders has been calculated at the suggested threshold 0.421000, the largest connected component was selected, and then dilated by 2Å.

Results:

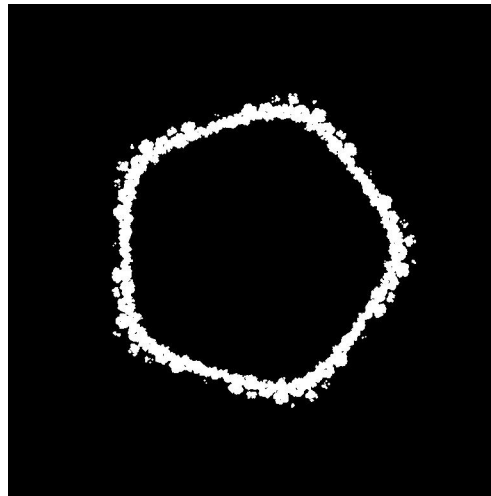
See Fig. 5.



(a) X Slice 144



(b) Y Slice 118



(c) Z Slice 564

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

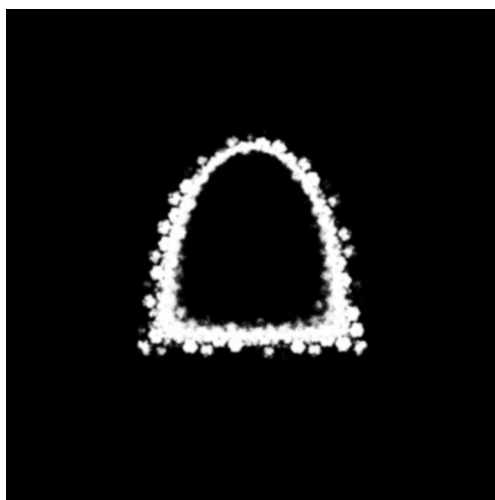
Orthogonal slices of maximum variance of the mask with soft borders

Explanation:

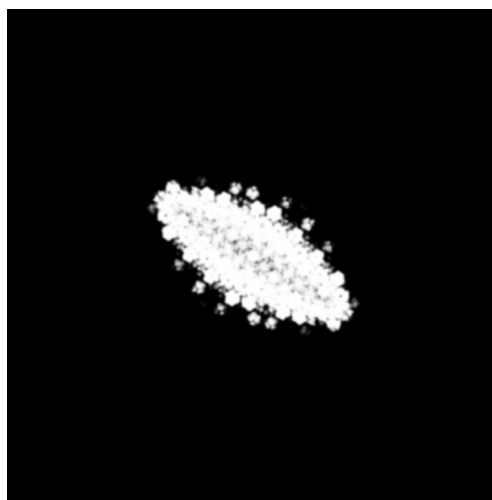
The mask with soft borders has been calculated at the suggested threshold 0.421000, the largest connected component was selected, and then dilated by

2Å.

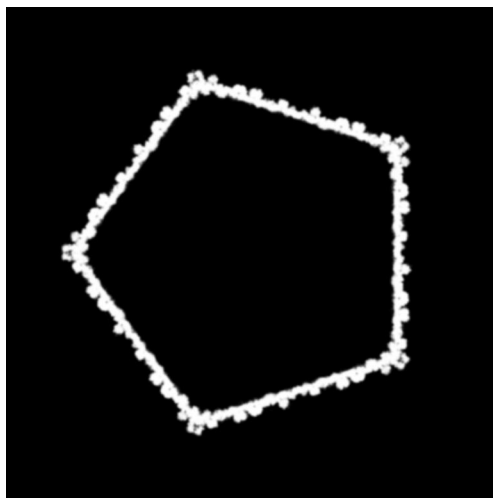
Results:
See Fig. 6.



(a) X Slice 624



(b) Y Slice 651



(c) Z Slice 237

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 243.60 and $243.60\text{ \AA}$, respectively. There is a decentering ratio $(\text{abs}(\text{Right-Left})/\text{Size})\%$ of 0.00%

The space from the left and right in Y are 269.70 and $272.60\text{ \AA}$, respectively. There is a decentering ratio $(\text{abs}(\text{Right-Left})/\text{Size})\%$ of 0.13%

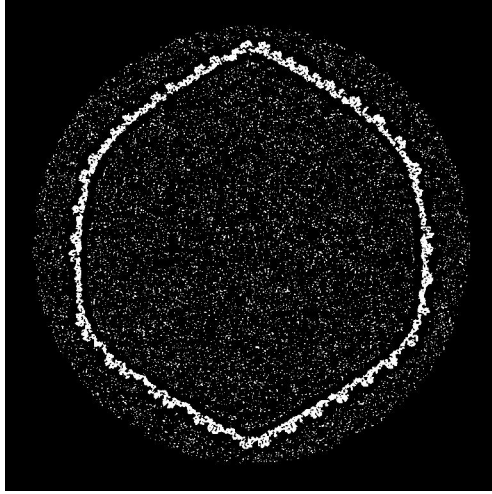
The space from the left and right in Z are 153.70 and $153.70\text{ \AA}$, respectively. There is a decentering ratio $(\text{abs}(\text{Right-Left})/\text{Size})\%$ of 0.00%

0.00, 51.56)(503 (12267.67), 0.00, 51.57)(503 (12267.67), 0.00, 51.57)(497
 (12121.34), 0.00, 51.57)(497 (12121.34), 0.00, 51.58)(432 (10536.05), 0.00,
 51.58)(432 (10536.05), 0.00, 51.58)(365 (8901.99), 0.00, 51.58)(365 (8901.99),
 0.00, 51.58)(362 (8828.82), 0.00, 51.59)(362 (8828.82), 0.00, 51.59)(338 (8243.48),
 0.00, 51.59)(338 (8243.48), 0.00, 51.59)(337 (8219.10), 0.00, 51.59)(337 (8219.10),
 0.00, 51.60)(336 (8194.71), 0.00, 51.60)(336 (8194.71), 0.00, 51.60)(336 (8194.71),
 0.00, 51.60)(336 (8194.71), 0.00, 51.60)(328 (7999.59), 0.00, 51.61)(328 (7999.59),
 0.00, 51.61)(328 (7999.59), 0.00, 51.61)(328 (7999.59), 0.00, 51.61)(324 (7902.04),
 0.00, 51.61)(324 (7902.04), 0.00, 51.61)(324 (7902.04), 0.00, 51.62)(324 (7902.04),
 0.00, 51.62)(321 (7828.87), 0.00, 51.62)(321 (7828.87), 0.00, 51.62)(321 (7828.87),
 0.00, 51.62)(321 (7828.87), 0.00, 51.63)(320 (7804.48), 0.00, 51.63)(320 (7804.48),
 0.00, 51.63)(318 (7755.70), 0.00, 51.63)(318 (7755.70), 0.00, 51.63)(310 (7560.59),
 0.00, 51.63)(310 (7560.59), 0.00, 51.64)(309 (7536.20), 0.00, 51.64)(309 (7536.20),
 0.00, 51.64)(303 (7389.87), 0.00, 51.64)(303 (7389.87), 0.00, 51.64)(302 (7365.48),
 0.00, 51.64)(302 (7365.48), 0.00, 51.65)(302 (7365.48), 0.00, 51.65)(302 (7365.48),
 0.00, 51.65)(300 (7316.70), 0.00, 51.65)(300 (7316.70), 0.00, 51.65)(296 (7219.15),
 0.00, 51.65)(296 (7219.15), 0.00, 51.66)(294 (7170.37), 0.00, 51.66)(294 (7170.37),
 0.00, 51.66)(285 (6950.87), 0.00, 51.66)(285 (6950.87), 0.00, 51.66)(282 (6877.70),
 0.00, 51.66)(282 (6877.70), 0.00, 51.67)(280 (6828.92), 0.00, 51.67)(280 (6828.92),
 0.00, 51.67)(279 (6804.53), 0.00, 51.67)(279 (6804.53), 0.00, 51.67)(278 (6780.14),
 0.00, 51.67)(278 (6780.14), 0.00, 51.68)(276 (6731.37), 0.00, 51.68)(276 (6731.37),
 0.00, 51.68)(275 (6706.98), 0.00, 51.68)(275 (6706.98), 0.00, 51.68)(273 (6658.20),
 0.00, 51.68)(273 (6658.20), 0.00, 51.68)(271 (6609.42), 0.00, 51.69)(271 (6609.42),
 0.00, 51.69)(270 (6585.03), 0.00, 51.69)(270 (6585.03), 0.00, 51.69)(269 (6560.64),
 0.00, 51.69)(269 (6560.64), 0.00, 51.69)(268 (6536.25), 0.00, 51.69)(268 (6536.25),
 0.00, 51.70)(268 (6536.25), 0.00, 51.70)(268 (6536.25), 0.00, 51.70)(266 (6487.48),
 0.00, 51.70)(266 (6487.48), 0.00, 51.70)(265 (6463.09), 0.00, 51.70)(265 (6463.09),
 0.00, 51.71)(265 (6463.09), 0.00, 51.71)(265 (6463.09), 0.00, 51.71)(264 (6438.70),
 0.00, 51.71)(264 (6438.70), 0.00, 51.71)(262 (6389.92), 0.00, 51.71)(262 (6389.92),
 0.00, 51.71)(262 (6389.92), 0.00, 51.72)(262 (6389.92), 0.00, 51.72)

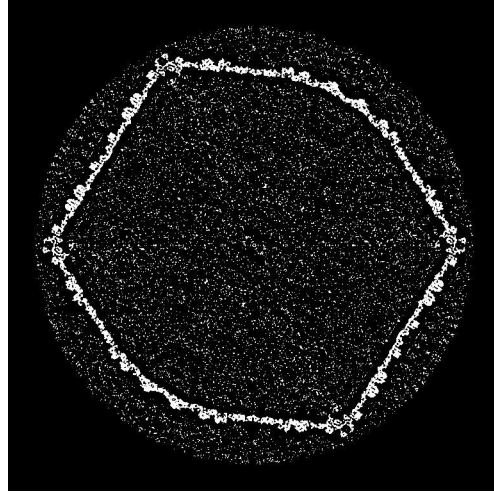
Number of components to reach 95% of the mass: 601249

The average size of the remaining 578769 components is 1.55 voxels (24.39 Å³). Their size go from 3 voxels (73.17 Å³) to 1 voxels (24.39 Å³).

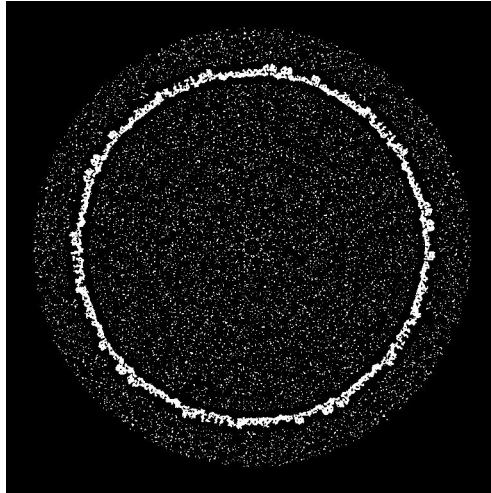
The slices of the raw mask can be seen in Fig. 7.



(a) X Slice 364



(b) Y Slice 384



(c) Z Slice 392

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
0.1432	52503847.00	1060908.63	9644623.95
0.2864	30563427.00	617573.86	5614307.84
0.4297	17375720.00	351099.06	3191809.65
0.5729	10840848.00	219053.46	1991395.07
0.7161	7876345.00	159151.81	1446834.66
0.8593	6318552.00	127674.58	1160677.96
1.0025	5192798.00	104927.25	953884.08
1.1458	4202899.00	84925.05	772045.91
1.2890	3300950.00	66700.00	606363.60
1.4322	2499300.00	50501.61	459105.57
1.5754	1819676.00	36768.92	334262.95
1.7186	1264836.00	25557.66	232342.36
1.8619	837691.00	16926.64	153878.53
2.0051	526632.00	10641.29	96738.96
2.1483	309422.00	6252.27	56838.86
2.2915	169288.00	3420.68	31097.13
2.4347	84268.00	1702.74	15479.50
2.5780	38942.00	786.87	7153.40
2.7212	17628.00	356.20	3238.15
2.8644	7254.00	146.58	1332.51
3.0076	2602.00	52.58	477.97
3.1508	916.00	18.51	168.26
3.2941	244.00	4.93	44.82
3.4373	42.00	0.85	7.72

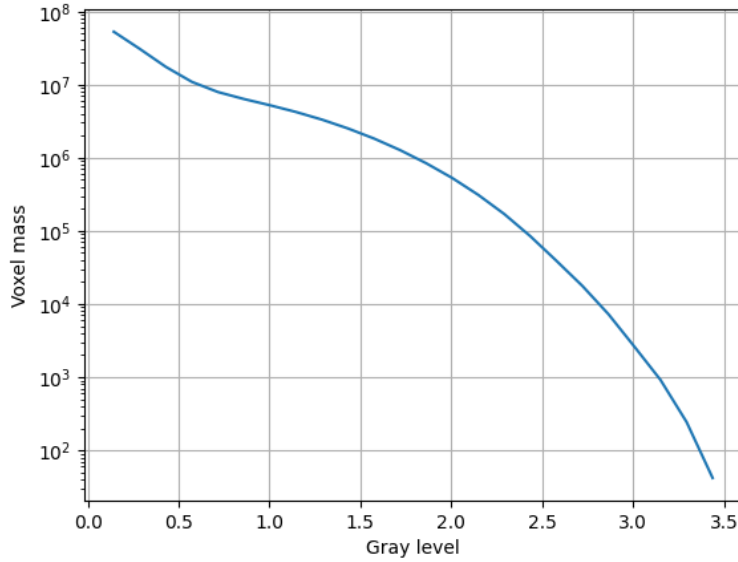


Figure 8: Voxel mass as a function of the gray level.

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 15278167 and a volume of $372619330.60 \text{ \AA}^3$. The overlap between the raw and constructed mask is 0.52.

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%.

WARNINGS: 3 warnings

1. **There might be a problem of connectivity at this threshold because more than 5 connected components are needed to reach 95% of the total mask. Probably a smaller threshold will not cause this issue.**
2. **There might be a problem with noise and artifacts, because the average noise blob has a volume of $24.389008 \text{ \AA}^3$.**
3. **There might be a problem in the construction of the mask,**

because the overlap is smaller than 0.75. A common reason is that the suggested threshold causes too many disconnected components.

2.3 Level 0.c Background analysis

Explanation:

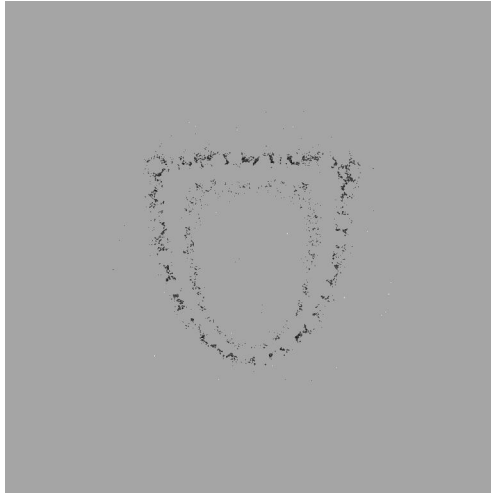
Background is defined as the region outside the macromolecule mask. The background mean should be zero, and the number of voxels with a very low or very high value (below 5 standard deviations of the noise) should be very small and they should be randomly distributed without any specific structure. Sometimes,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

Results:

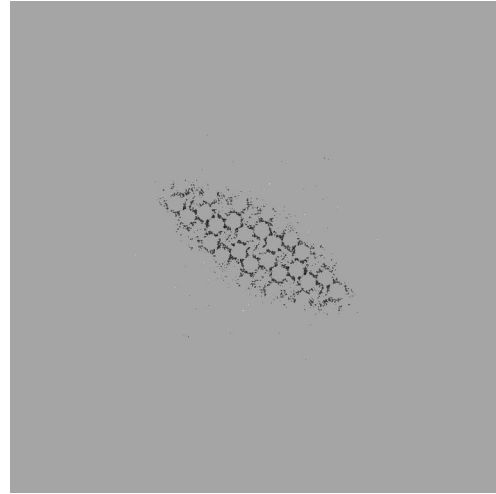
The null hypothesis that the background mean is 0 was tested with a one-sample Student's t-test. The resulting t-statistic and p-value were -2281.09 and 0.000000, respectively.

The mean and standard deviation (sigma) of the background were -0.021990 and 0.201686. The percentage of background voxels whose absolute value is larger than 5 times the standard deviation is 0.46 % (see Fig. 9). The same percentage from a Gaussian would be 0.000057% (ratio between the two percentages: 8061.521595).

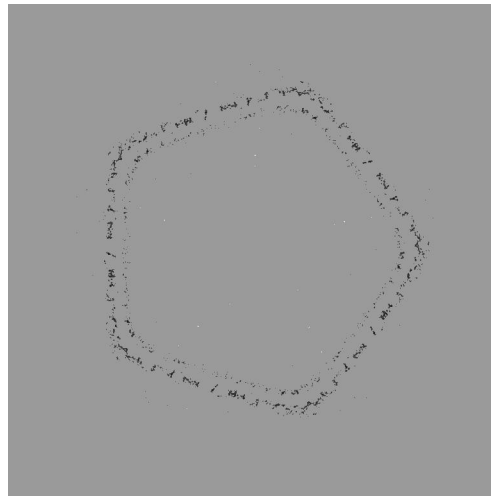
Slices of the background beyond 5*sigma can be seen in Fig. 9.



(a) X Slice 140



(b) Y Slice 107



(c) Z Slice 551

Figure 9: Maximum variance slices in the three dimensions of the parts of the background beyond $5 \cdot \sigma$

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=8061.52)**

2.4 Level 0.d B-factor analysis

Explanation:

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

Results:

This method cannot be applied to maps with a resolution worse than 8Å.

STATUS: Does not apply

3 Half maps

Half map 1: emd_21702_half_map_1.map

SHA256 hash: 7e0013774fad9963945038a1cd71ec96f604f4e609761d2ea1fcd385b11d0243

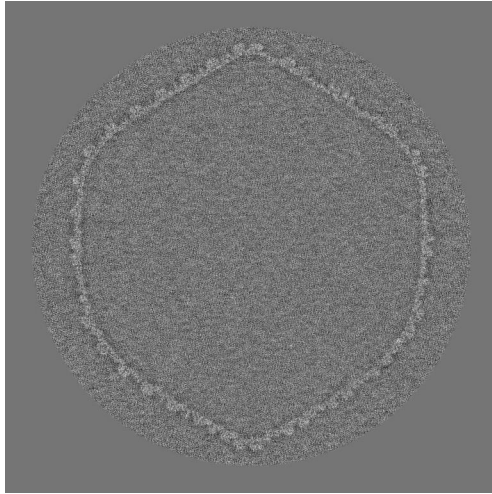
Half map 2: emd_21702_half_map_2.map

SHA256 hash: e3a15acefa4c6e3ba0dc5351dd86a2cefbaea0c27fd7c67b7a51f6210b690ede

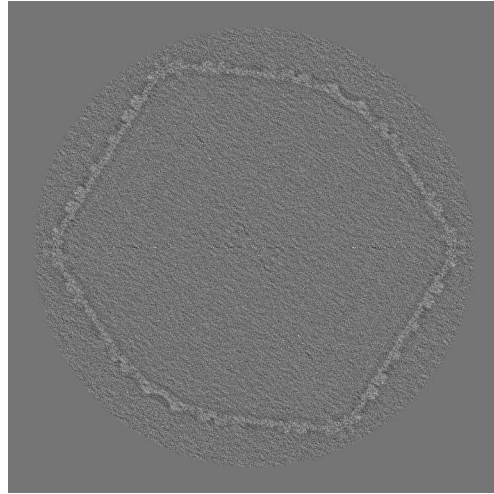
Slices of the first half map can be seen in Fig. [10](#).

Slices of the second half map can be seen in Fig. [11](#).

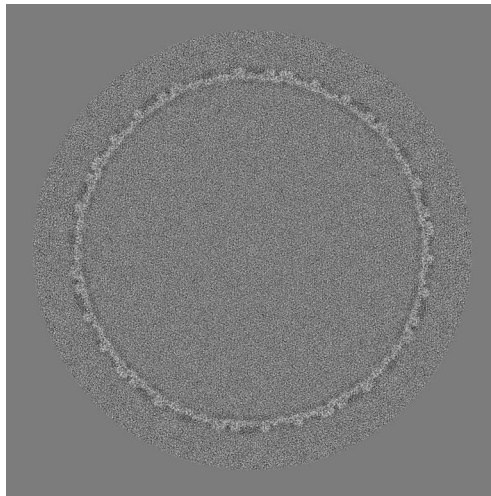
Slices of the difference between both maps can be seen in Fig. [12](#). There should not be any structure in this difference. Sometimes some patterns are seen if the map is symmetric.



(a) X Slice 405

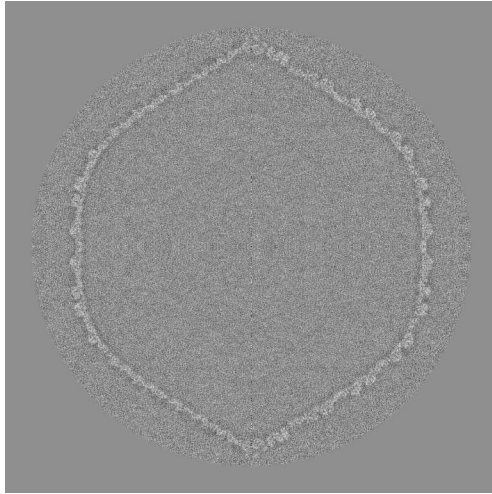


(b) Y Slice 384

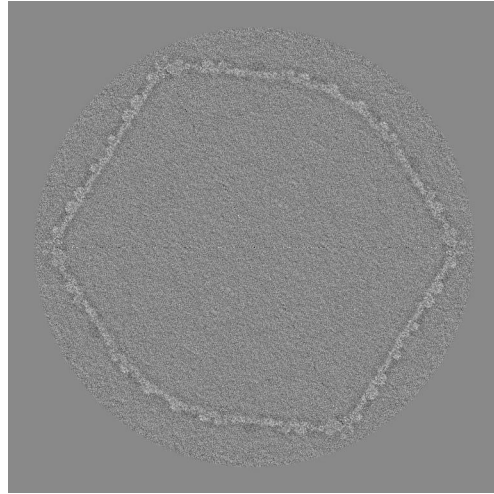


(c) Z Slice 369

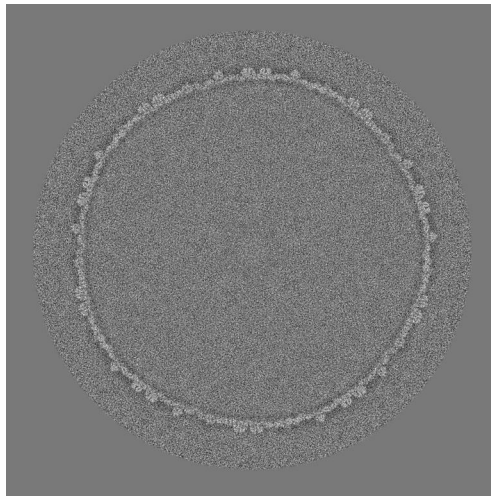
Figure 10: Slices of maximum variation in the three dimensions of Half 1



(a) X Slice 384

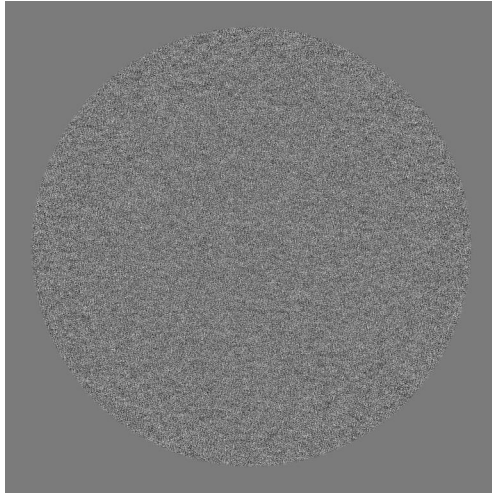


(b) Y Slice 384

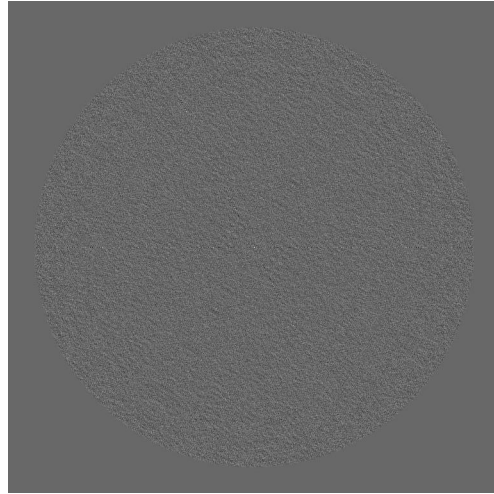


(c) Z Slice 407

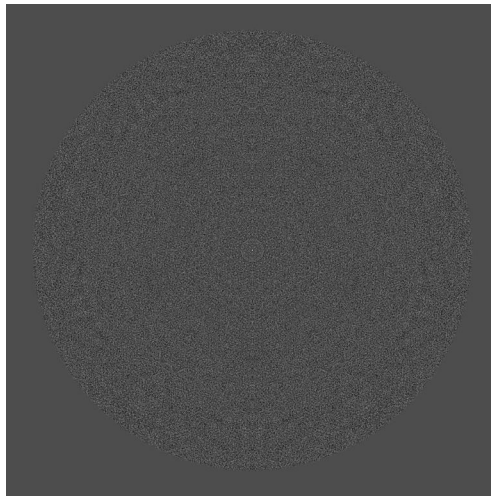
Figure 11: Slices of maximum variation in the three dimensions of Half 2



(a) X Slice 379



(b) Y Slice 384



(c) Z Slice 384

Figure 12: Slices of maximum variation in the three dimensions of the difference Half1-Half2.