



wwPDB EM Validation Summary Report ⓘ

Aug 14, 2023 – 04:13 pm BST

PDB ID : 8PUG
EMDB ID : EMD-17942
Title : Structure of the immature HTLV-1 CA lattice from full-length Gag VLPs:
CA-NTD refinement
Authors : Obr, M.; Percipalle, M.; Chernikova, D.; Yang, H.; Thader, A.; Pinke, G.;
Porley, D.; Mansky, L.M.; Dick, R.A.; Schur, F.K.M.
Deposited on : 2023-07-17
Resolution : 5.90 Å (reported)
Based on initial model : .

This is a wwPDB EM Validation Summary Report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<https://www.wwpdb.org/validation/2017/EMValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

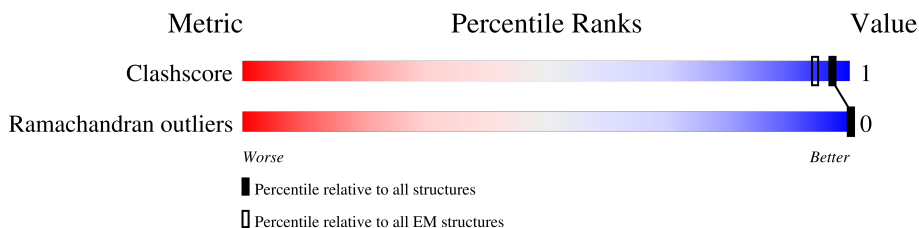
EMDB validation analysis : 0.0.1.dev50
MolProbity : 4.02b-467
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ : 1.9.9
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.35

1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:
ELECTRON MICROSCOPY

The reported resolution of this entry is 5.90 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	158937	4297
Ramachandran outliers	154571	4023

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments of the bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions $\leq 5\%$. The upper red bar (where present) indicates the fraction of residues that have poor fit to the EM map (all-atom inclusion $< 40\%$). The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	429	26% 74%
1	B	429	26% 74%
1	C	429	26% 74%

2 Entry composition [i](#)

There is only 1 type of molecule in this entry. The entry contains 1356 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Gag polyprotein.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
1	A	113	452	226	113	113	0	0
1	B	113	452	226	113	113	0	0
1	C	113	452	226	113	113	0	0

CYS
PRO
THR
ILE
TYR
GLU
PRO
GLU
PRO
GLU
GLU
ASP
ALA
LEU
LEU
LEU
LEU
ASP
PRO
PRO
ALA
ASP
ILE
PRO
HIS
TRP
HIS
PRO
LYS
ASN
SER
SER
ILE
GLY
GLY
GLU
VAL

● Molecule 1: Gag polyprotein



MET
GLY
GLN
ILE
THR
SER
SER
ARG
SER
SER
ALA
SER
SER
ILE
LEU
PRO
PRO
ARG
LEU
LEU
LEU
GLY
GLY
VAL
ALA
ALA
HIS
HIS
HIS
TRP
LEU
HIS
ASN
PHE
LEU
ILE
GLN
GLN
THR
ALA
ALA
TYR
ARG
LEU
GLU
GLU
PRO
PRO
GLY
PRO
SER
SER
TYR
PRO
ASP
PHE
HIS
SER
GLN
LEU
LEU
LYS
LYS
PHE
LEU
LEU
LYS
ILE
GLU
THR
PRO
ALA
PRO
ARG
ILE
PRO

CYS
PRO
ILE
ASN
SER
TYR
LEU
LEU
ALA
LEU
PRO
LEU
LEU
PRO
LYS
MET
HIS
PRO
HIS
GLY
GLY
ALA
PRO
PRO
ASN
HIS
HIS
R13
AG0
G33
A101
P125
GLY
SER
ALA
LYS
ASP
PRO
SER
TRP
ALA
SER
ILE
LEU
GLN
GLY
LEU
GLU
GLU
PRO
PRO
TYR
HIS
ALA
PHE
VAL
GLU
ARG
LEU
ASN
ILE
PRO

TYR
VAL
GLU
PRO
THR
ALA
PRO
GLN
VAL
LEU
LEU
VAL
MET
HIS
PRO
HIS
GLY
GLY
ALA
PRO
PRO
ASN
HIS
HIS
R13
AG0
G33
A101
P125
GLY
SER
ALA
LYS
ASP
PRO
SER
TRP
ALA
SER
ILE
LEU
GLN
GLY
LEU
GLU
GLU
PRO
PRO
TYR
HIS
ALA
PHE
VAL
GLU
ARG
LEU
ASN
ILE
ALA

LEU
ASP
ASN
GLY
LEU
PRO
GLU
GLY
THR
PRO
LYS
PRO
ASP
PRO
ILE
LEU
SER
ARG
GLY
LEU
LEU
ALA
TYR
SER
ASN
ALA
ASN
LYS
GLU
CYS
GLN
LYS
LEU
LEU
GLN
ALA
ARG
GLY
HIS
HIS
THR
ASN
SER
ASP
CYS
THR
GLN
PRO
ARG
PRO
PRO
PRO
GLY
PRO
CYS
PRO
LEU
CYS
GLN
ASP
LEU
THR
HIS
TRP
LYS
ARG
ALA
CYS
GLN
THR
ARG
LEU
THR
LYS
PRO
LYS
THR
ILE
PRO
GLU
VAL
LEU

VAL
VAL
GLN
PRO
LYS
PRO
PRO
ASN
GLN
PRO
CYS
PHE
ARG
CYS
GLY
LEU
LEU
ALA
GLY
HIS
TRP
SER
ARG
ASP
CYS
THR
GLN
PRO
ARG
PRO
PRO
GLY
PRO
CYS
PRO
LEU
CYS
THR
HIS
TRP
LYS
ARG
ALA
CYS
GLN
THR
ARG
LEU
THR
LYS
PRO
THR
ILE
PRO
GLU
VAL
GLU

PRO
GLU
GLU
ASP
ALA
LEU
LEU
LEU
ASP
PRO
HIS
HIS
PRO
LYS
ASN
SER
SER
GLY
VAL

4 Experimental information

Property	Value	Source
EM reconstruction method	SUBTOMOGRAM AVERAGING	Depositor
Imposed symmetry	POINT, C2	Depositor
Number of subtomograms used	132000	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{\AA}^2$)	143.5	Depositor
Minimum defocus (nm)	1250	Depositor
Maximum defocus (nm)	4000	Depositor
Magnification	80000	Depositor
Image detector	GATAN K3 BIOQUANTUM (6k x 4k)	Depositor
Maximum map value	0.013	Depositor
Minimum map value	-0.008	Depositor
Average map value	0.000	Depositor
Map value standard deviation	0.000	Depositor
Recommended contour level	0.0018	Depositor
Map size (\AA)	441.92, 441.92, 441.92	wwPDB
Map dimensions	320, 320, 320	wwPDB
Map angles ($^\circ$)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (\AA)	1.381, 1.381, 1.381	Depositor

5 Model quality [i](#)

5.1 Standard geometry [i](#)

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z >5	RMSZ	# Z >5
1	A	0.24	0/451	0.47	0/562
1	B	0.24	0/451	0.46	0/562
1	C	0.24	0/451	0.44	0/562
All	All	0.24	0/1353	0.46	0/1686

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts [i](#)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	452	0	114	0	0
1	B	452	0	114	1	0
1	C	452	0	114	0	0
All	All	1356	0	342	1	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 1.

All (1) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:22:ALA:O	1:B:25:GLN:N	2.50	0.44

There are no symmetry-related clashes.

5.3 Torsion angles [i](#)

5.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all EM entries.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	111/429 (26%)	109 (98%)	2 (2%)	0	100	100
1	B	111/429 (26%)	107 (96%)	4 (4%)	0	100	100
1	C	111/429 (26%)	105 (95%)	6 (5%)	0	100	100
All	All	333/1287 (26%)	321 (96%)	12 (4%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains [i](#)

There are no protein residues with a non-rotameric sidechain to report in this entry.

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

5.6 Ligand geometry [i](#)

There are no ligands in this entry.

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

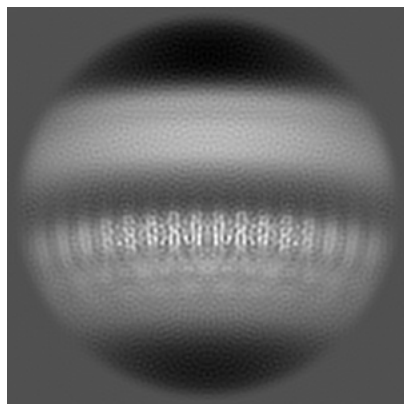
6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-17942. These allow visual inspection of the internal detail of the map and identification of artifacts.

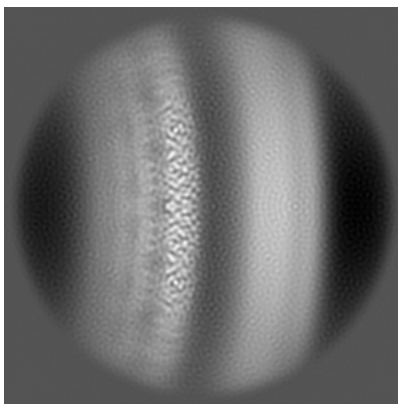
Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

6.1 Orthogonal projections [i](#)

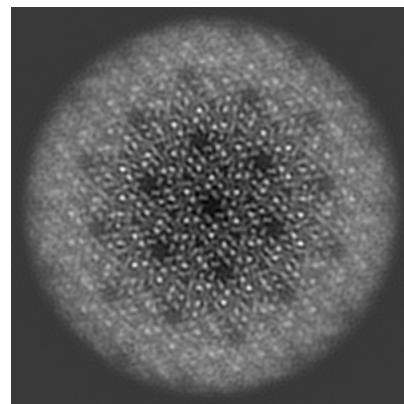
6.1.1 Primary map



X

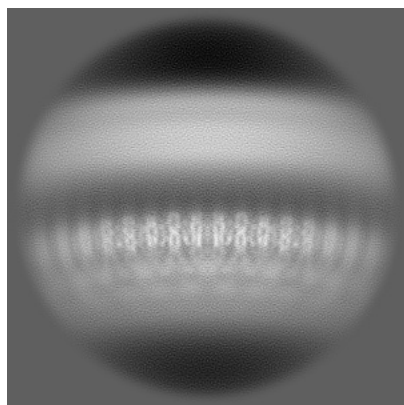


Y

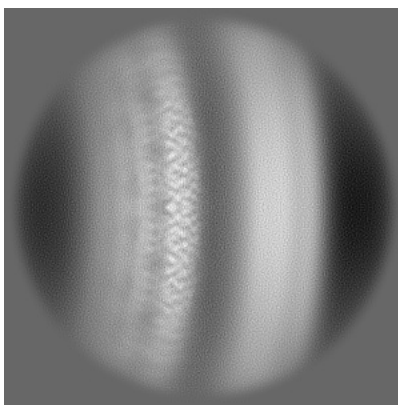


Z

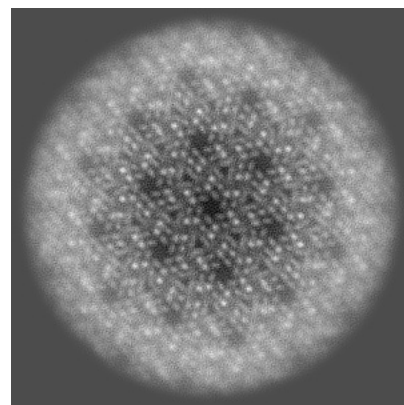
6.1.2 Raw map



X



Y

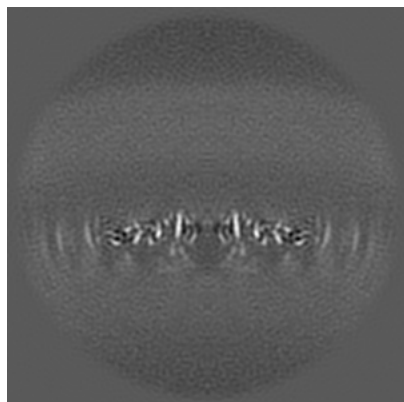


Z

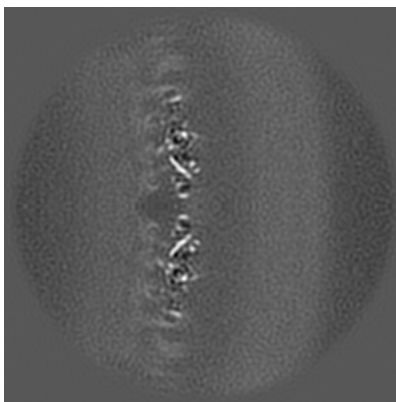
The images above show the map projected in three orthogonal directions.

6.2 Central slices [i](#)

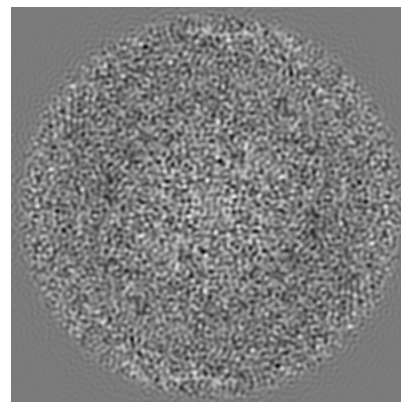
6.2.1 Primary map



X Index: 160

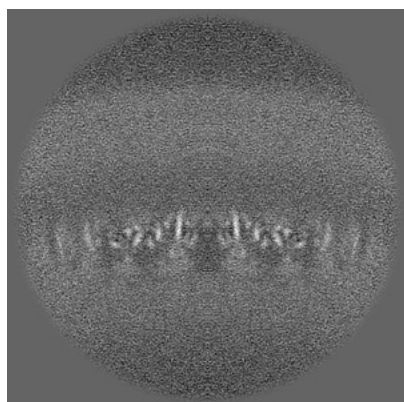


Y Index: 160

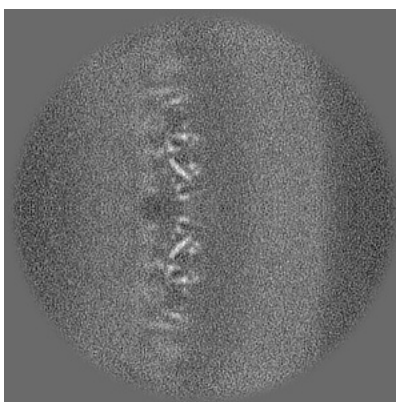


Z Index: 160

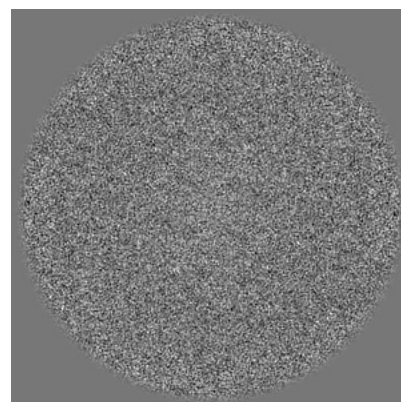
6.2.2 Raw map



X Index: 160



Y Index: 160

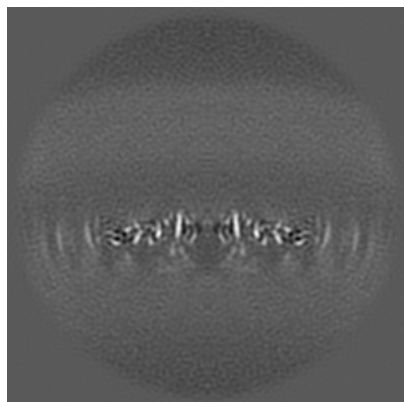


Z Index: 160

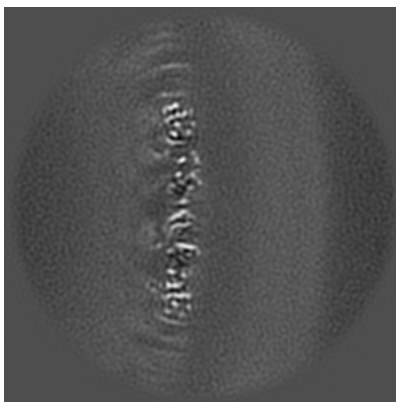
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices [i](#)

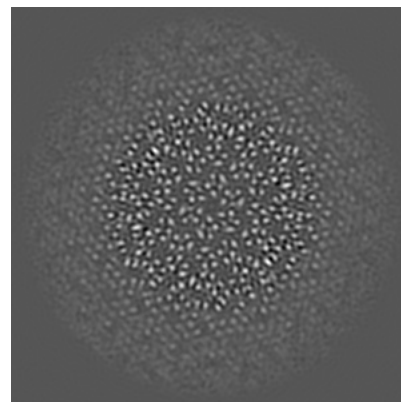
6.3.1 Primary map



X Index: 160

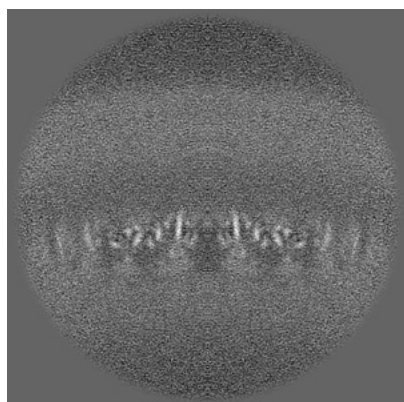


Y Index: 167

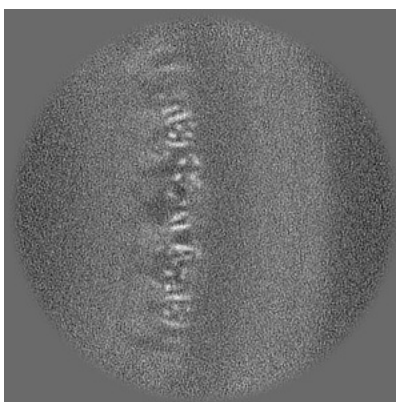


Z Index: 136

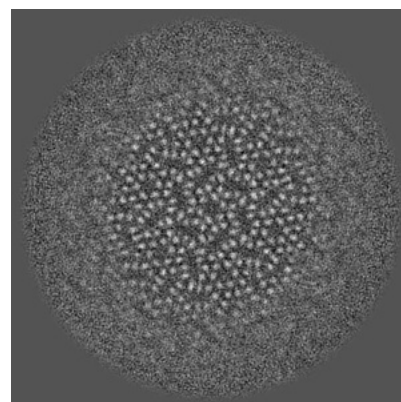
6.3.2 Raw map



X Index: 160



Y Index: 166

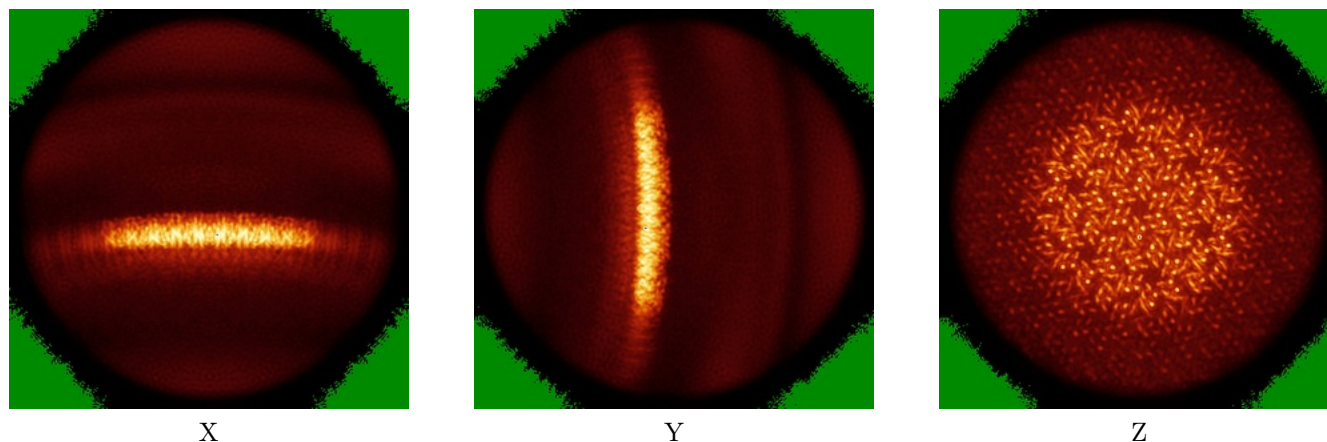


Z Index: 138

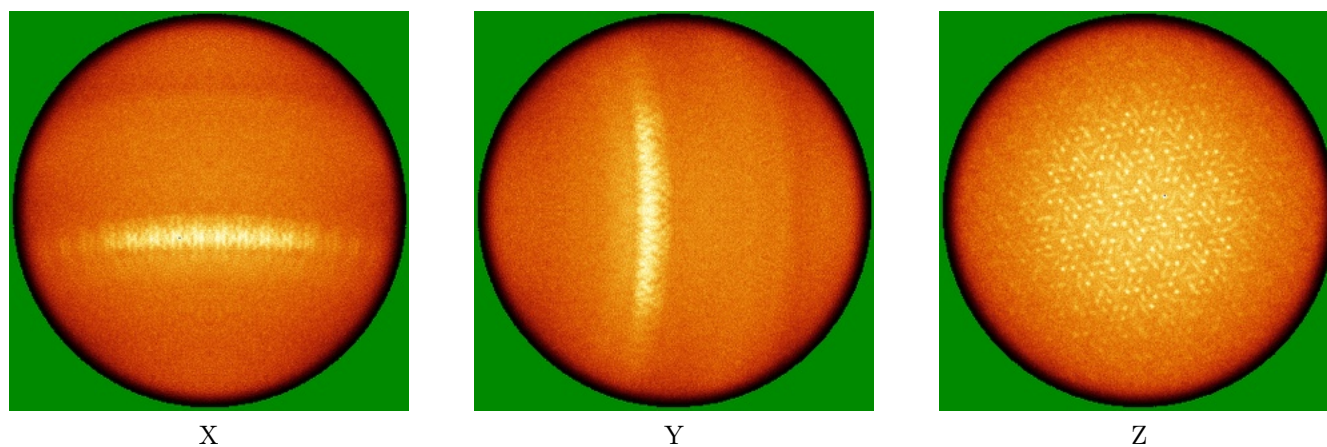
The images above show the largest variance slices of the map in three orthogonal directions.

6.4 Orthogonal standard-deviation projections (False-color) [i](#)

6.4.1 Primary map



6.4.2 Raw map



The images above show the map standard deviation projections with false color in three orthogonal directions. Minimum values are shown in green, max in blue, and dark to light orange shades represent small to large values respectively.

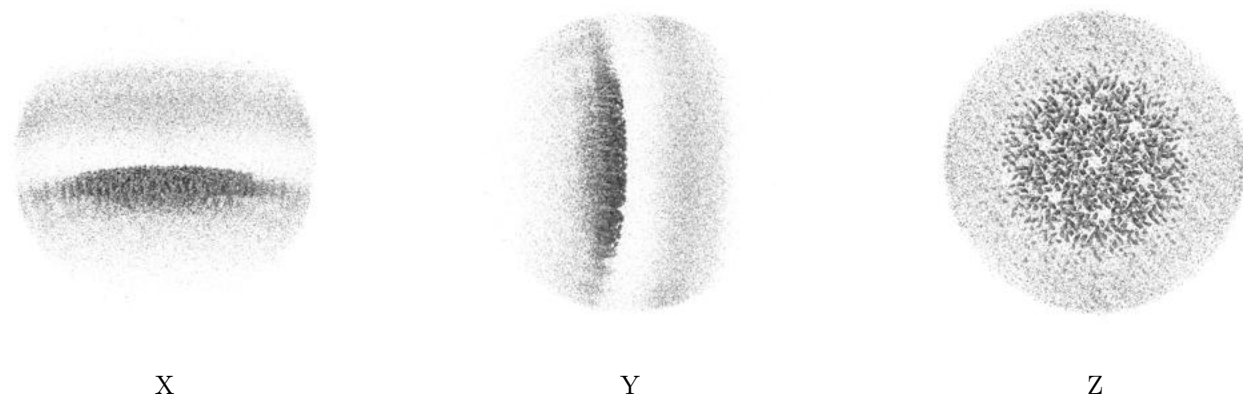
6.5 Orthogonal surface views [i](#)

6.5.1 Primary map



The images above show the 3D surface view of the map at the recommended contour level 0.0018. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

6.5.2 Raw map



These images show the 3D surface of the raw map. The raw map's contour level was selected so that its surface encloses the same volume as the primary map does at its recommended contour level.

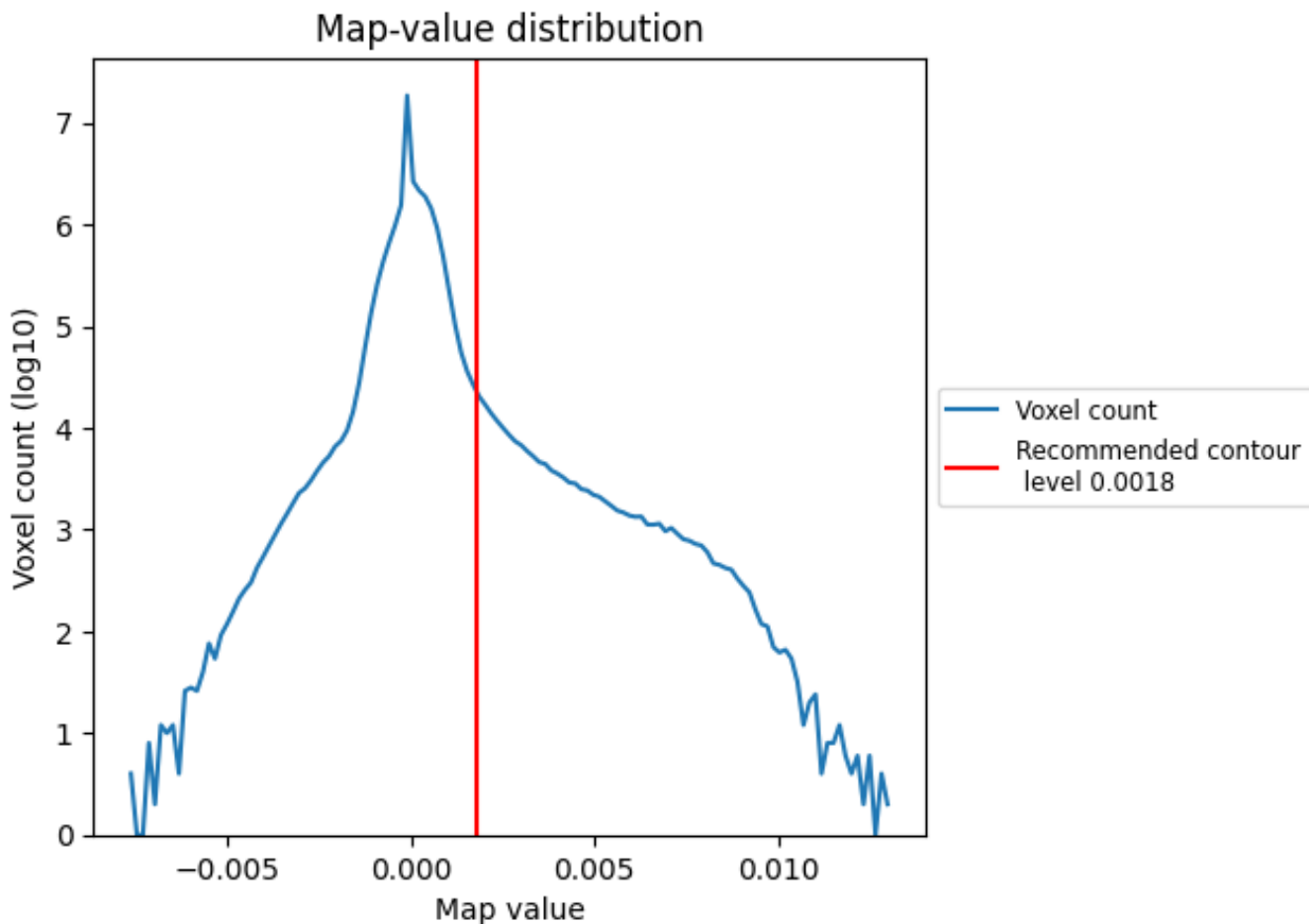
6.6 Mask visualisation [i](#)

This section was not generated. No masks/segmentation were deposited.

7 Map analysis [i](#)

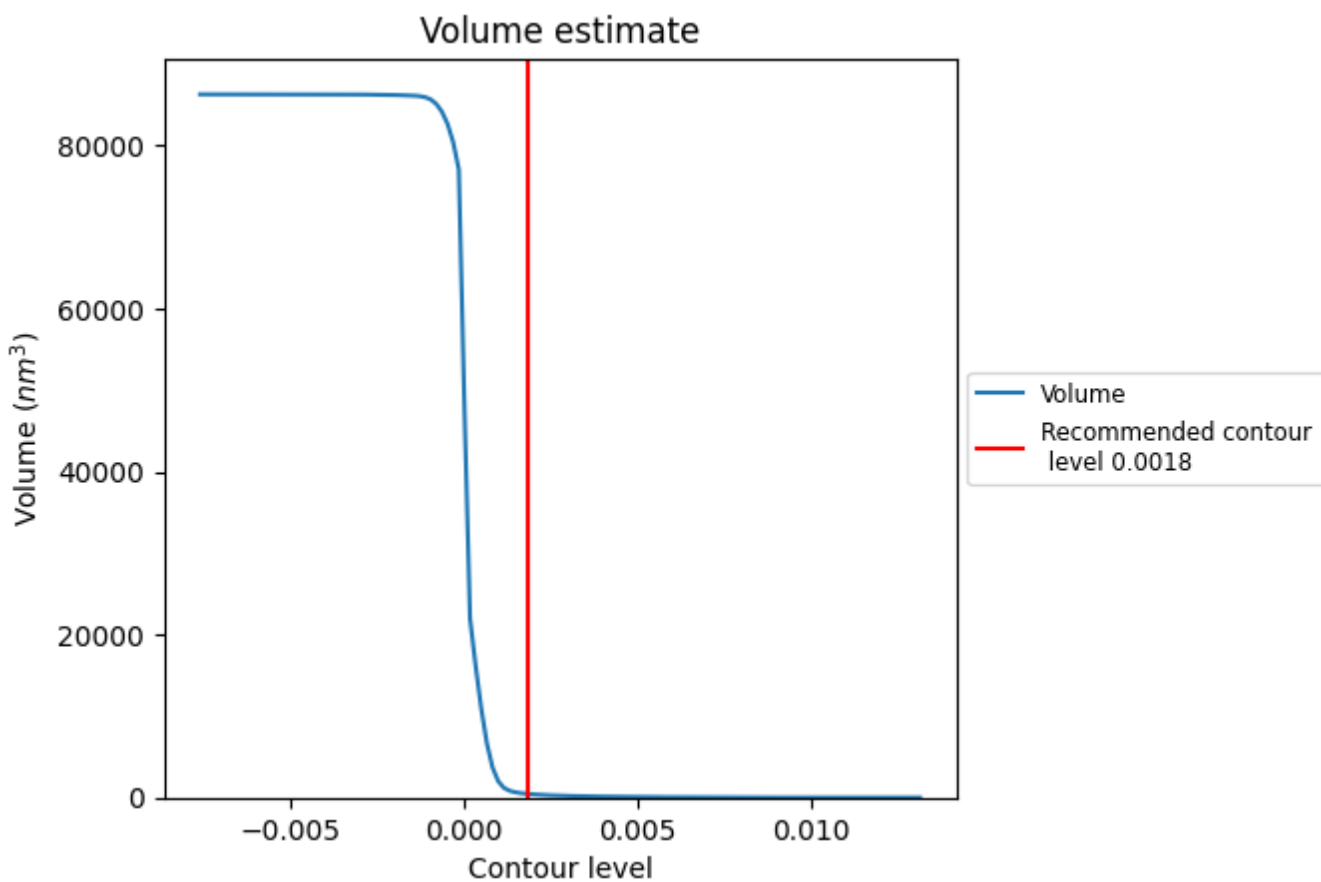
This section contains the results of statistical analysis of the map.

7.1 Map-value distribution [i](#)



The map-value distribution is plotted in 128 intervals along the x-axis. The y-axis is logarithmic. A spike in this graph at zero usually indicates that the volume has been masked.

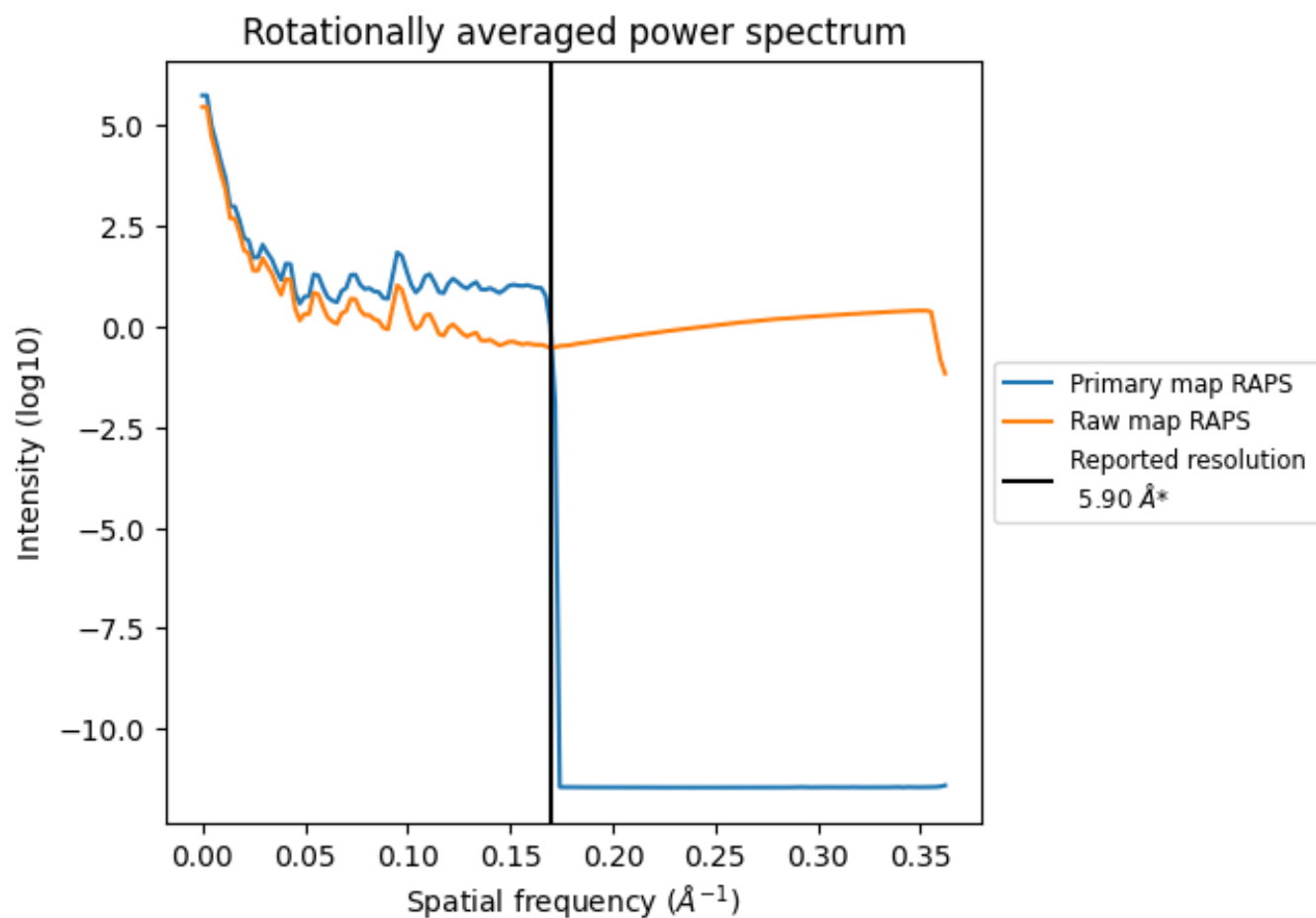
7.2 Volume estimate [\(i\)](#)



The volume at the recommended contour level is 459 nm³; this corresponds to an approximate mass of 414 kDa.

The volume estimate graph shows how the enclosed volume varies with the contour level. The recommended contour level is shown as a vertical line and the intersection between the line and the curve gives the volume of the enclosed surface at the given level.

7.3 Rotationally averaged power spectrum [i](#)

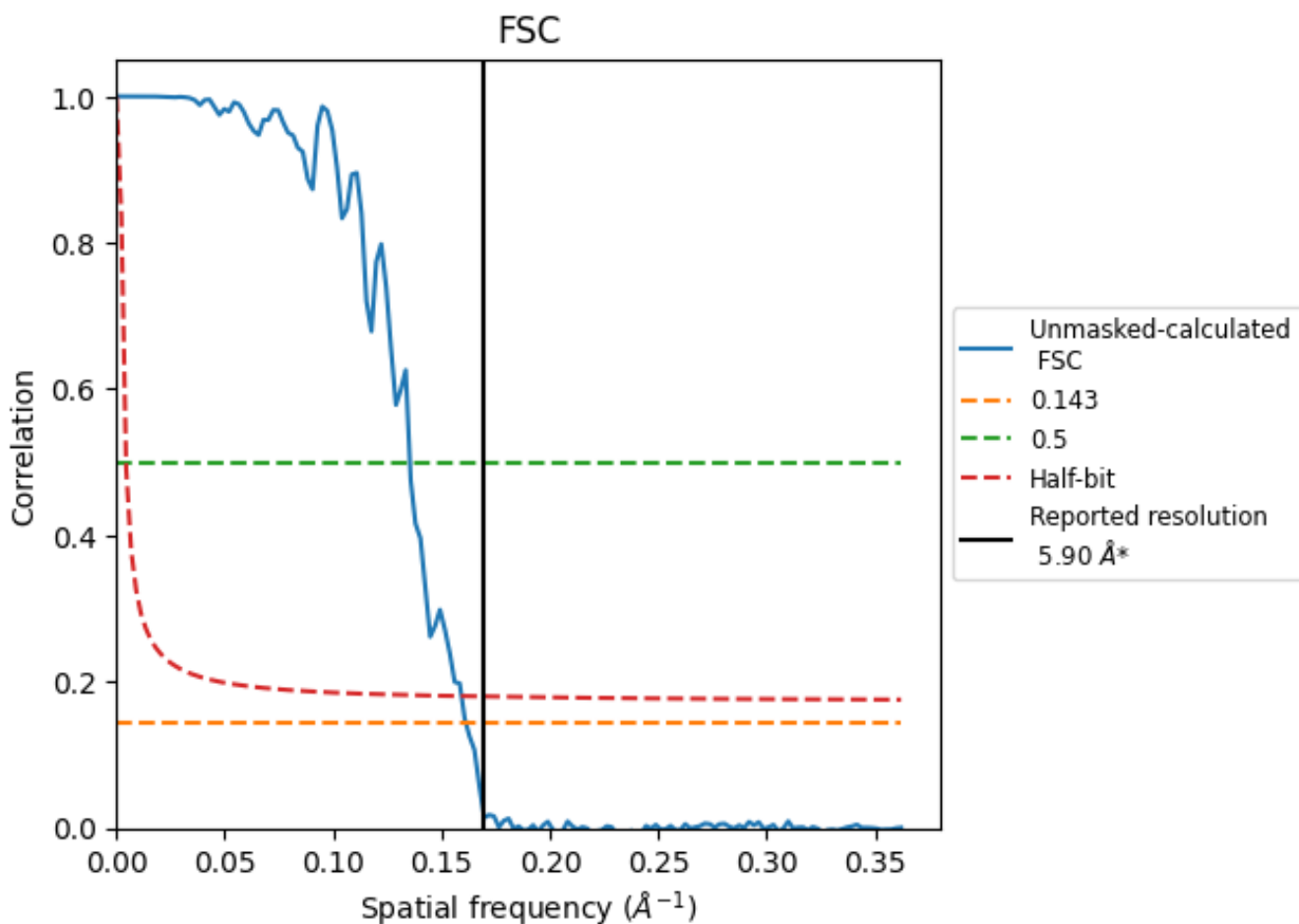


*Reported resolution corresponds to spatial frequency of 0.169 Å⁻¹

8 Fourier-Shell correlation [i](#)

Fourier-Shell Correlation (FSC) is the most commonly used method to estimate the resolution of single-particle and subtomogram-averaged maps. The shape of the curve depends on the imposed symmetry, mask and whether or not the two 3D reconstructions used were processed from a common reference. The reported resolution is shown as a black line. A curve is displayed for the half-bit criterion in addition to lines showing the 0.143 gold standard cut-off and 0.5 cut-off.

8.1 FSC [i](#)



*Reported resolution corresponds to spatial frequency of 0.169 Å⁻¹

8.2 Resolution estimates [i](#)

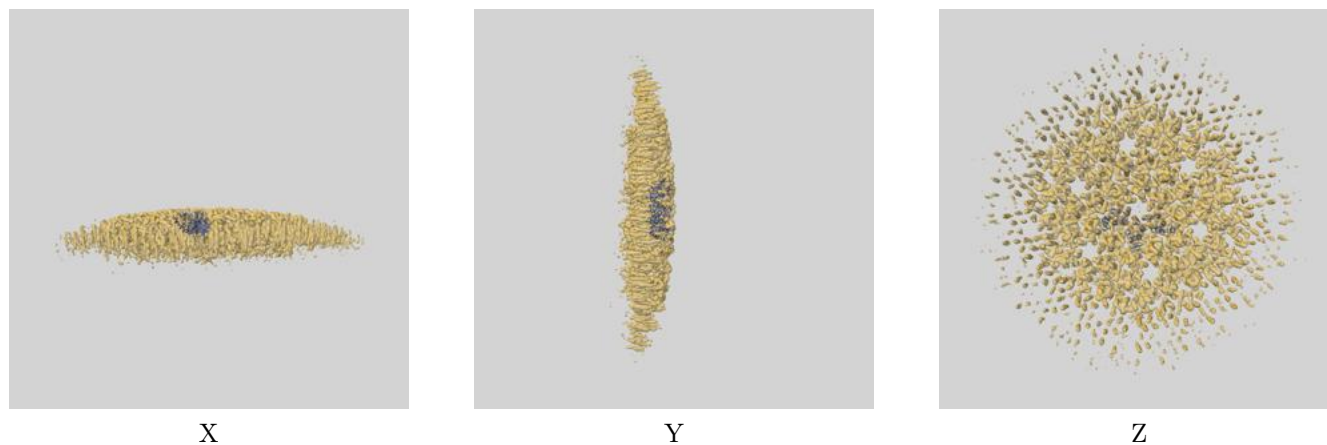
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	5.90	-	-
Author-provided FSC curve	-	-	-
Unmasked-calculated*	6.20	7.39	6.28

*Resolution estimate based on FSC curve calculated by comparison of deposited half-maps.

9 Map-model fit [i](#)

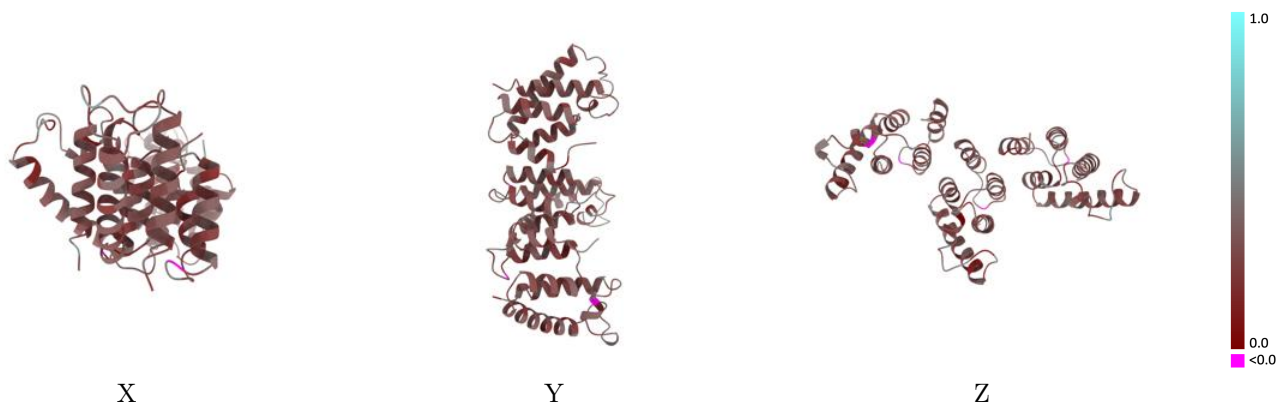
This section contains information regarding the fit between EMDB map EMD-17942 and PDB model 8PUG. Per-residue inclusion information can be found in section 3 on page 4.

9.1 Map-model overlay [i](#)



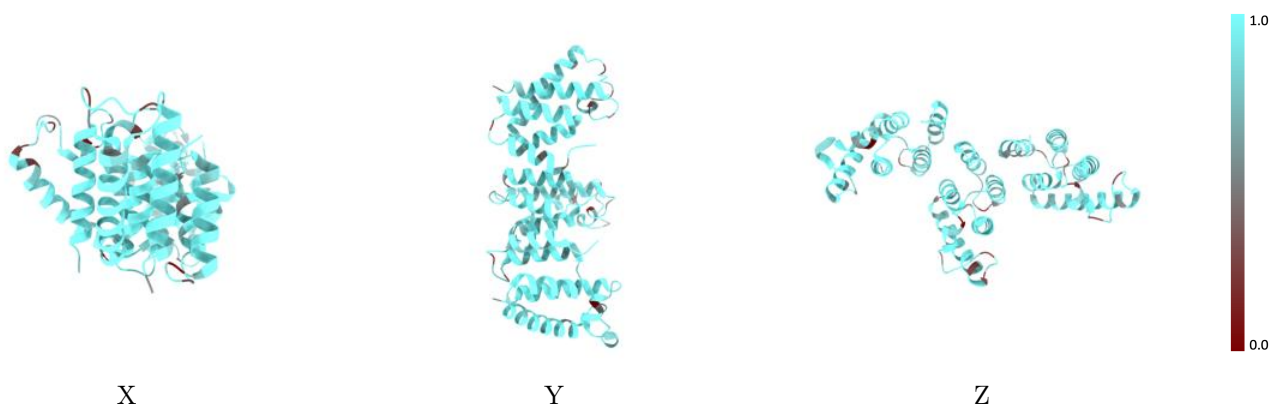
The images above show the 3D surface view of the map at the recommended contour level 0.0018 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

9.2 Q-score mapped to coordinate model [\(i\)](#)



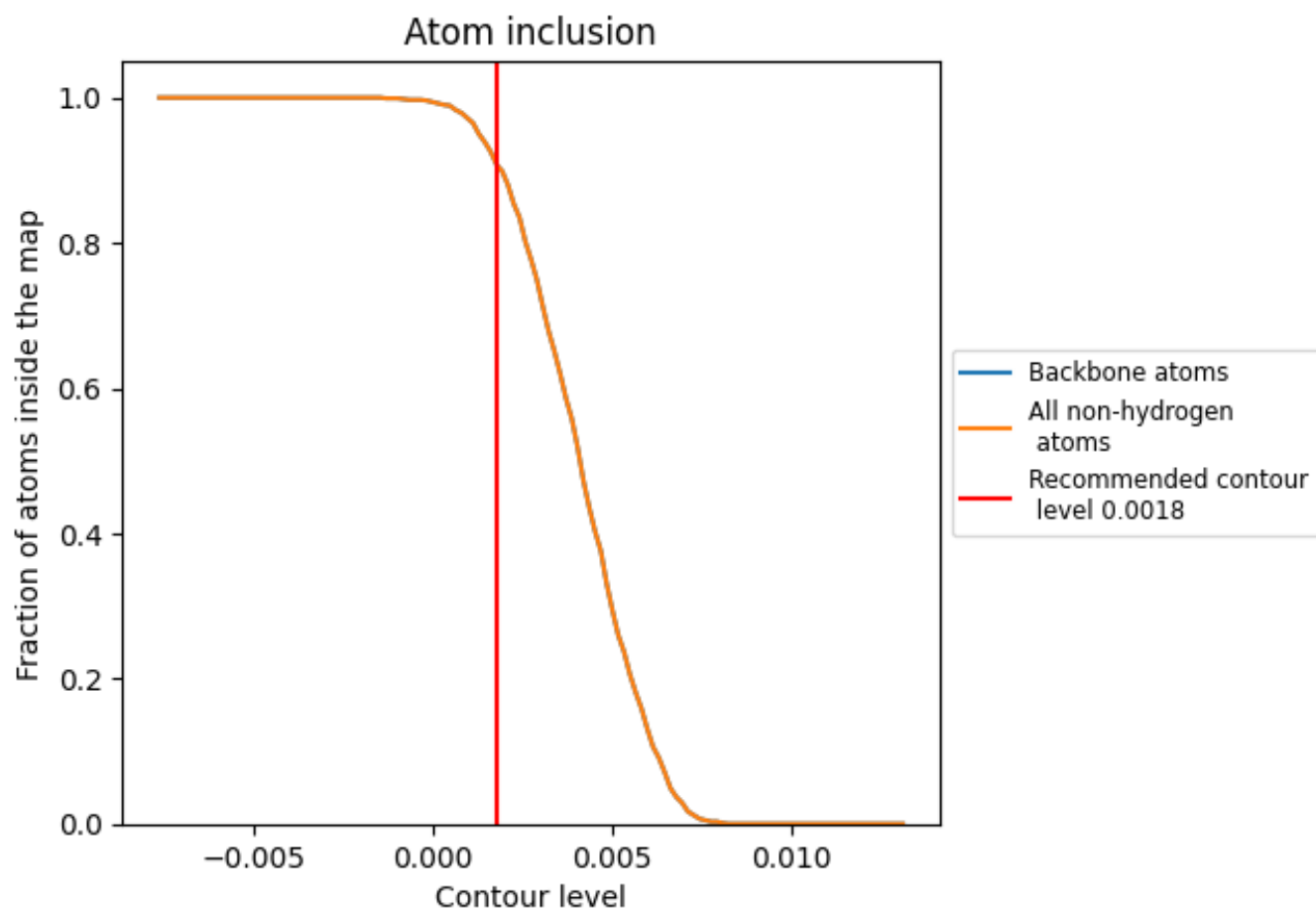
The images above show the model with each residue coloured according to its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions at which atoms can be resolved. Low Q-score values may therefore be expected for many entries.

9.3 Atom inclusion mapped to coordinate model [\(i\)](#)



The images above show the model with each residue coloured according to its atom inclusion. This shows to what extent they are inside the map at the recommended contour level (0.0018).





9.4 Atom inclusion [i](#)



At the recommended contour level, 91% of all backbone atoms, 91% of all non-hydrogen atoms, are inside the map.

9.5 Map-model fit summary

The table lists the average atom inclusion at the recommended contour level (0.0018) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	 0.9070	 0.3040
A	 0.9090	 0.3030
B	 0.9000	 0.3040
C	 0.9110	 0.3050

