



wwPDB EM Validation Summary Report ⓘ

Jun 26, 2023 – 07:28 pm BST

PDB ID : 8C9M
EMDB ID : EMD-16511
Title : HERV-K Gag immature lattice
Authors : Krebs, A.-S.; Liu, H.-F.; Zhou, Y.; Rey, J.S.; Levintov, L.; Perilla, J.R.; Bartesaghi, A.; Zhang, P.
Deposited on : 2023-01-23
Resolution : 3.20 Å(reported)

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) : 2.33

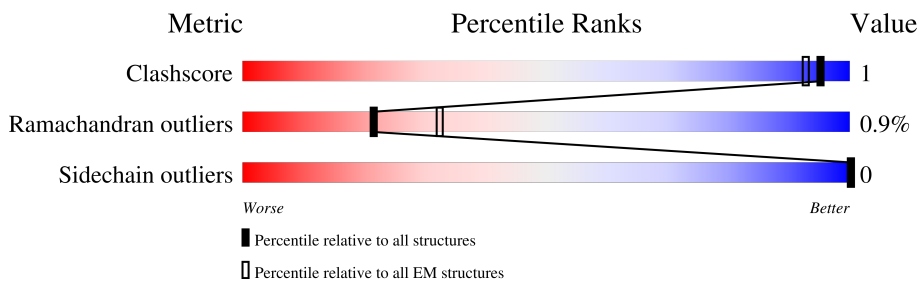
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 3.20 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
Sidechain outliers	154315	3826

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	666	
1	B	666	
1	C	666	
1	D	666	
1	E	666	
1	F	666	

2 Entry composition [i](#)

There is only 1 type of molecule in this entry. The entry contains 21480 atoms, of which 10752 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 protein.

Mol	Chain	Residues	Atoms						AltConf	Trace
			Total	C	H	N	O	S		
1	A	228	3580	1127	1792	315	334	12	0	0
1	B	228	3580	1127	1792	315	334	12	0	0
1	C	228	3580	1127	1792	315	334	12	0	0
1	D	228	3580	1127	1792	315	334	12	0	0
1	E	228	3580	1127	1792	315	334	12	0	0
1	F	228	3580	1127	1792	315	334	12	0	0

There are 78 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-191	ILE	LEU	conflict	UNP Q96897
A	-185	ASN	LYS	conflict	UNP Q96897
A	-174	GLY	SER	conflict	UNP Q96897
A	-57	PRO	SER	conflict	UNP Q96897
A	38	HIS	TYR	conflict	UNP Q96897
A	120	CYS	VAL	conflict	UNP Q96897
A	133	THR	ALA	conflict	UNP Q96897
A	160	GLN	GLU	conflict	UNP Q96897
A	191	PRO	SER	conflict	UNP Q96897
A	209	ALA	VAL	conflict	UNP Q96897
A	255	ASN	SER	conflict	UNP Q96897
A	291	GLN	LEU	conflict	UNP Q96897
A	356	PRO	LEU	conflict	UNP Q96897
B	-191	ILE	LEU	conflict	UNP Q96897
B	-185	ASN	LYS	conflict	UNP Q96897
B	-174	GLY	SER	conflict	UNP Q96897
B	-57	PRO	SER	conflict	UNP Q96897
B	38	HIS	TYR	conflict	UNP Q96897

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	120	CYS	VAL	conflict	UNP Q96897
B	133	THR	ALA	conflict	UNP Q96897
B	160	GLN	GLU	conflict	UNP Q96897
B	191	PRO	SER	conflict	UNP Q96897
B	209	ALA	VAL	conflict	UNP Q96897
B	255	ASN	SER	conflict	UNP Q96897
B	291	GLN	LEU	conflict	UNP Q96897
B	356	PRO	LEU	conflict	UNP Q96897
C	-191	ILE	LEU	conflict	UNP Q96897
C	-185	ASN	LYS	conflict	UNP Q96897
C	-174	GLY	SER	conflict	UNP Q96897
C	-57	PRO	SER	conflict	UNP Q96897
C	38	HIS	TYR	conflict	UNP Q96897
C	120	CYS	VAL	conflict	UNP Q96897
C	133	THR	ALA	conflict	UNP Q96897
C	160	GLN	GLU	conflict	UNP Q96897
C	191	PRO	SER	conflict	UNP Q96897
C	209	ALA	VAL	conflict	UNP Q96897
C	255	ASN	SER	conflict	UNP Q96897
C	291	GLN	LEU	conflict	UNP Q96897
C	356	PRO	LEU	conflict	UNP Q96897
D	-191	ILE	LEU	conflict	UNP Q96897
D	-185	ASN	LYS	conflict	UNP Q96897
D	-174	GLY	SER	conflict	UNP Q96897
D	-57	PRO	SER	conflict	UNP Q96897
D	38	HIS	TYR	conflict	UNP Q96897
D	120	CYS	VAL	conflict	UNP Q96897
D	133	THR	ALA	conflict	UNP Q96897
D	160	GLN	GLU	conflict	UNP Q96897
D	191	PRO	SER	conflict	UNP Q96897
D	209	ALA	VAL	conflict	UNP Q96897
D	255	ASN	SER	conflict	UNP Q96897
D	291	GLN	LEU	conflict	UNP Q96897
D	356	PRO	LEU	conflict	UNP Q96897
E	-191	ILE	LEU	conflict	UNP Q96897
E	-185	ASN	LYS	conflict	UNP Q96897
E	-174	GLY	SER	conflict	UNP Q96897
E	-57	PRO	SER	conflict	UNP Q96897
E	38	HIS	TYR	conflict	UNP Q96897
E	120	CYS	VAL	conflict	UNP Q96897
E	133	THR	ALA	conflict	UNP Q96897
E	160	GLN	GLU	conflict	UNP Q96897

Continued on next page...

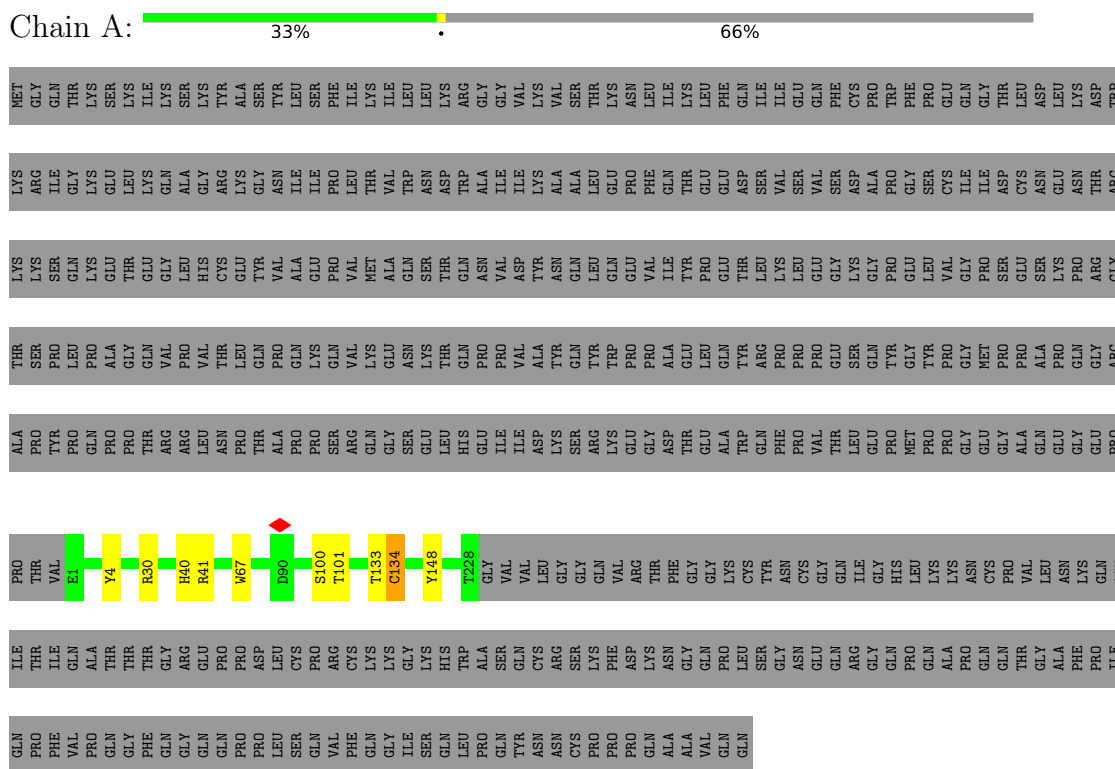
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	191	PRO	SER	conflict	UNP Q96897
E	209	ALA	VAL	conflict	UNP Q96897
E	255	ASN	SER	conflict	UNP Q96897
E	291	GLN	LEU	conflict	UNP Q96897
E	356	PRO	LEU	conflict	UNP Q96897
F	-191	ILE	LEU	conflict	UNP Q96897
F	-185	ASN	LYS	conflict	UNP Q96897
F	-174	GLY	SER	conflict	UNP Q96897
F	-57	PRO	SER	conflict	UNP Q96897
F	38	HIS	TYR	conflict	UNP Q96897
F	120	CYS	VAL	conflict	UNP Q96897
F	133	THR	ALA	conflict	UNP Q96897
F	160	GLN	GLU	conflict	UNP Q96897
F	191	PRO	SER	conflict	UNP Q96897
F	209	ALA	VAL	conflict	UNP Q96897
F	255	ASN	SER	conflict	UNP Q96897
F	291	GLN	LEU	conflict	UNP Q96897
F	356	PRO	LEU	conflict	UNP Q96897

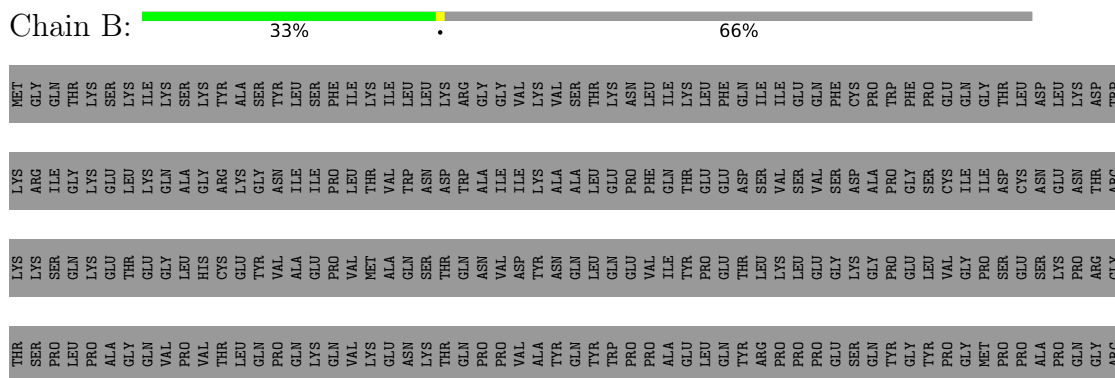
3 Residue-property plots [i](#)

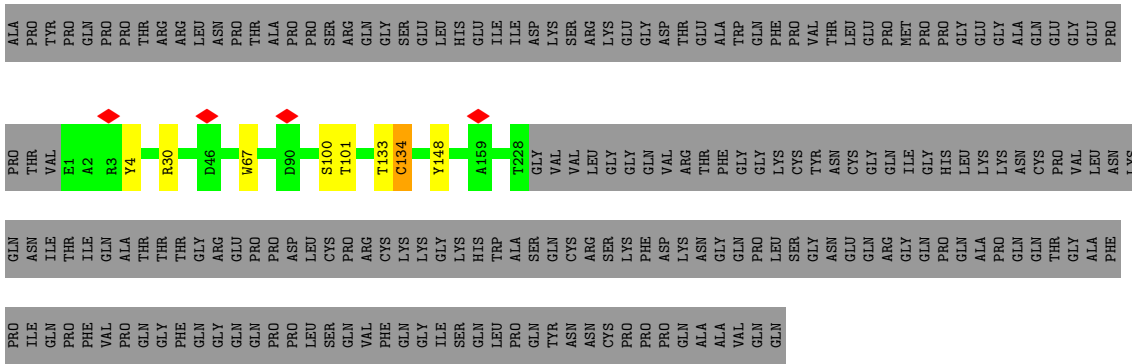
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and atom inclusion in map density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red diamond above a residue indicates a poor fit to the EM map for this residue (all-atom inclusion < 40%). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

- Molecule 1: Gag protein

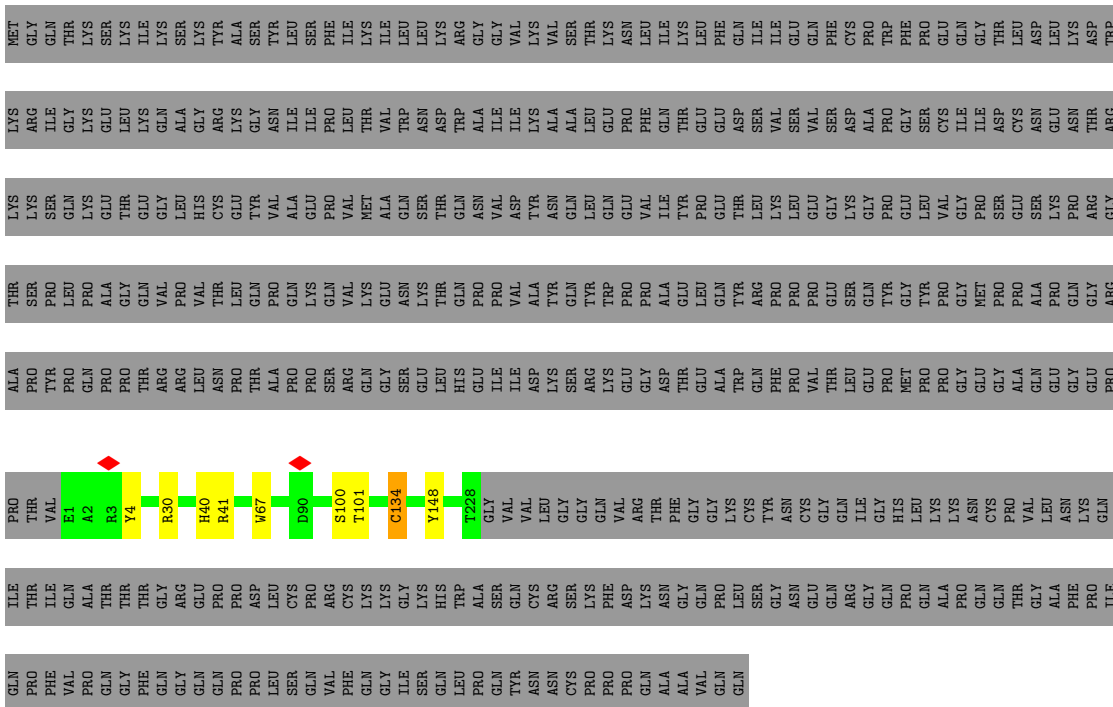


- Molecule 1: Gag protein

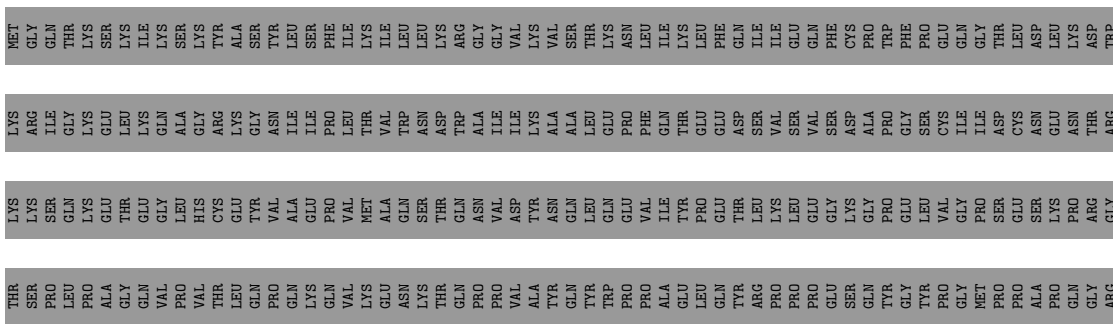




Molecule 1: Gag protein



Molecule 1: Gag protein



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

PRO
THR
VAL
E1
A2
R3
Y4
H40
R41
M67
D90
S100
T101
T133
C134
Y148
T228
GLY
VAL
VAL
CYS
LEU
GLY
GLY
GLN
VAL
ASP
THR
PHE
GLY
GLY
PRO
LEU
CYS
TYR
ASN
CYS
GLY
GLN
ILE
GLY
HIS
PRO
LEU
LYS
LYS
ASN
CYS
PRO
VAL
THR
GLY
ALA
PHE
PRO
ILE
GLN

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

PRO
PHE
VAL
PRO
GLY
PHE
GLY
GLN
GLN
PRO
PRO
LEU
SER
GLN
VAL
PHE
GLN
GLY
ILE
SER
GLN
LEU
GLN
TYR
ASN
ASN
CYS
PRO
PRO
GLN
ALA
VAL
VAL
GLN

4 Experimental information

Property	Value	Source
EM reconstruction method	SUBTOMOGRAM AVERAGING	Depositor
Imposed symmetry	POINT, C6	Depositor
Number of subtomograms used	188111	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 ($e^-/\text{\AA}^2$)	127.5	Depositor
Minimum defocus (nm)	2000	Depositor
Maximum defocus (nm)	5000	Depositor
Magnification	Not provided	
Image detector	FEI FALCON IV (4k x 4k)	Depositor
Maximum map value	1.198	Depositor
Minimum map value	-0.932	Depositor
Average map value	0.000	Depositor
Map value standard deviation	0.031	Depositor
Recommended contour level	0.0736	Depositor
Map size (\AA)	384.768, 384.768, 384.768	wwPDB
Map dimensions	256, 256, 256	wwPDB
Map angles ($^\circ$)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (\AA)	1.503, 1.503, 1.503	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	1.11	1/1825 (0.1%)	0.69	2/2467 (0.1%)
1	B	1.12	1/1825 (0.1%)	0.69	2/2467 (0.1%)
1	C	1.13	1/1825 (0.1%)	0.69	2/2467 (0.1%)
1	D	1.12	1/1825 (0.1%)	0.70	2/2467 (0.1%)
1	E	1.12	1/1825 (0.1%)	0.70	2/2467 (0.1%)
1	F	1.11	1/1825 (0.1%)	0.70	1/2467 (0.0%)
All	All	1.12	6/10950 (0.1%)	0.69	11/14802 (0.1%)

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	F	67	TRP	CA-C	-5.35	1.39	1.52
1	B	67	TRP	CA-C	-5.33	1.39	1.52
1	D	67	TRP	CA-C	-5.33	1.39	1.52
1	A	67	TRP	CA-C	-5.33	1.39	1.52
1	C	67	TRP	CA-C	-5.33	1.39	1.52

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	148	TYR	CB-CG-CD1	-6.45	117.13	121.00
1	D	148	TYR	CB-CG-CD1	-6.15	117.31	121.00
1	F	148	TYR	CB-CG-CD1	-6.11	117.33	121.00
1	E	148	TYR	CB-CG-CD1	-6.09	117.34	121.00
1	A	148	TYR	CB-CG-CD1	-6.06	117.36	121.00

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	1788	1792	1794	4	0
1	B	1788	1792	1794	4	0
1	C	1788	1792	1794	3	0
1	D	1788	1792	1794	3	0
1	E	1788	1792	1794	3	0
1	F	1788	1792	1794	4	0
All	All	10728	10752	10764	21	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.

The worst 5 of 21 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:E:100:SER:OG	1:E:101:THR:N	2.43	0.51
1:F:100:SER:OG	1:F:101:THR:N	2.43	0.51
1:A:100:SER:OG	1:A:101:THR:N	2.43	0.51
1:D:100:SER:OG	1:D:101:THR:N	2.43	0.50
1:B:100:SER:OG	1:B:101:THR:N	2.43	0.50

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	226/666 (34%)	218 (96%)	6 (3%)	2 (1%)	17 56

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	226/666 (34%)	218 (96%)	6 (3%)	2 (1%)	17	56
1	C	226/666 (34%)	218 (96%)	6 (3%)	2 (1%)	17	56
1	D	226/666 (34%)	218 (96%)	6 (3%)	2 (1%)	17	56
1	E	226/666 (34%)	218 (96%)	6 (3%)	2 (1%)	17	56
1	F	226/666 (34%)	218 (96%)	6 (3%)	2 (1%)	17	56
All	All	1356/3996 (34%)	1308 (96%)	36 (3%)	12 (1%)	21	56

5 of 12 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	4	TYR
1	B	4	TYR
1	C	4	TYR
1	D	4	TYR
1	E	4	TYR

5.3.2 Protein sidechains [i](#)

In the following table, the Percentiles column shows the percent sidechain 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 sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	192/570 (34%)	192 (100%)	0	100	100
1	B	192/570 (34%)	192 (100%)	0	100	100
1	C	192/570 (34%)	192 (100%)	0	100	100
1	D	192/570 (34%)	192 (100%)	0	100	100
1	E	192/570 (34%)	192 (100%)	0	100	100
1	F	192/570 (34%)	192 (100%)	0	100	100
All	All	1152/3420 (34%)	1152 (100%)	0	100	100

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

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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.

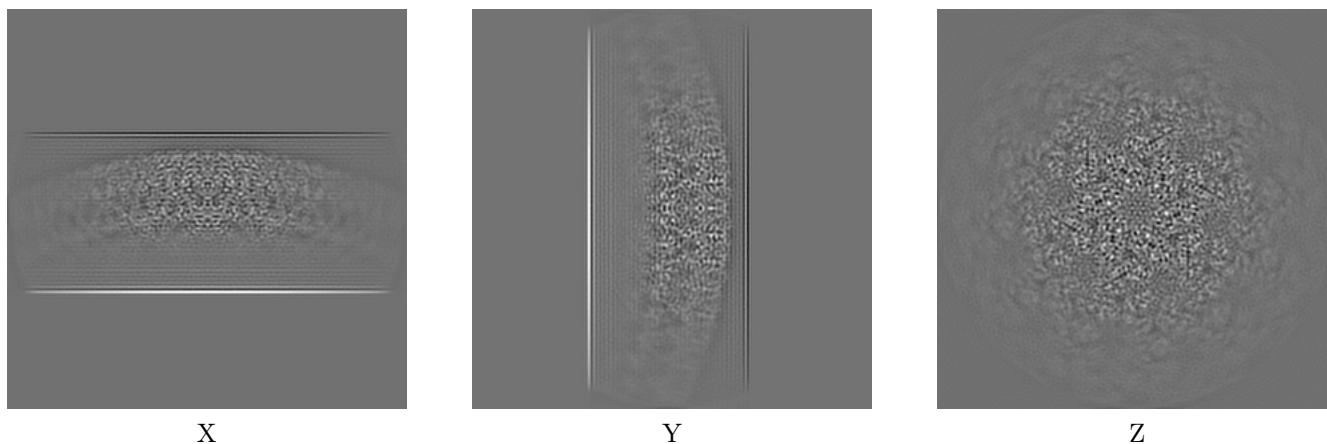
6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-16511. These allow visual inspection of the internal detail of the map and identification of artifacts.

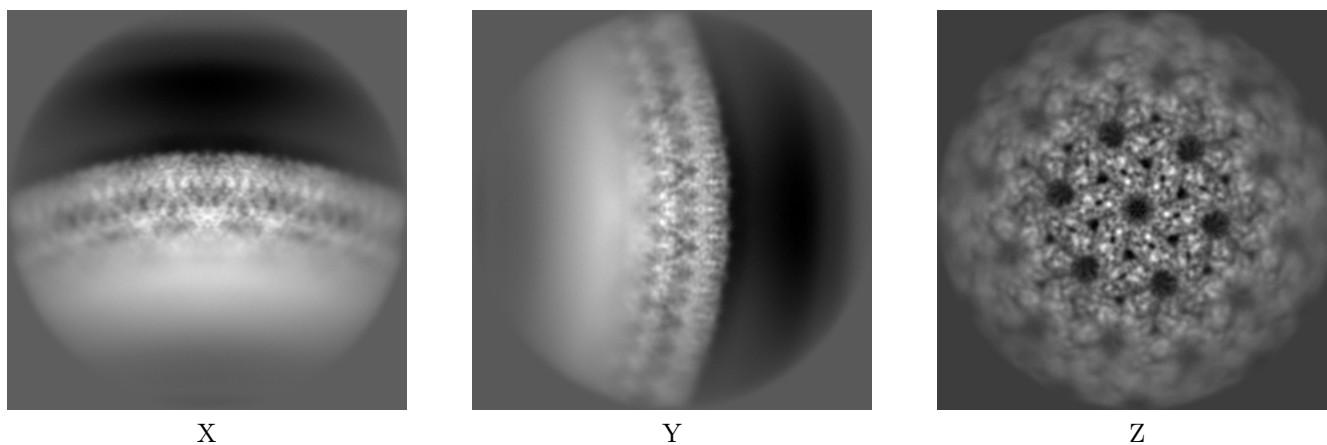
Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

6.1 Orthogonal projections [i](#)

6.1.1 Primary map



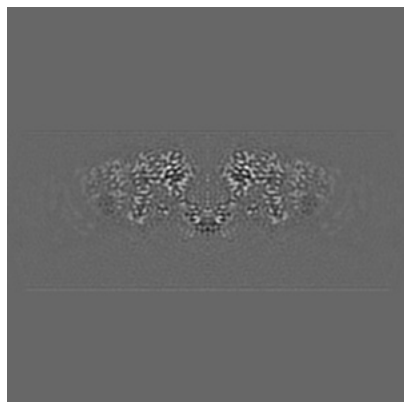
6.1.2 Raw map



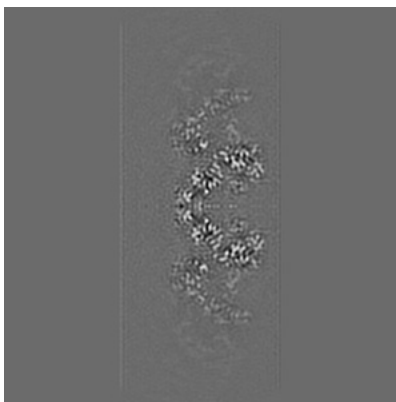
The images above show the map projected in three orthogonal directions.

6.2 Central slices [i](#)

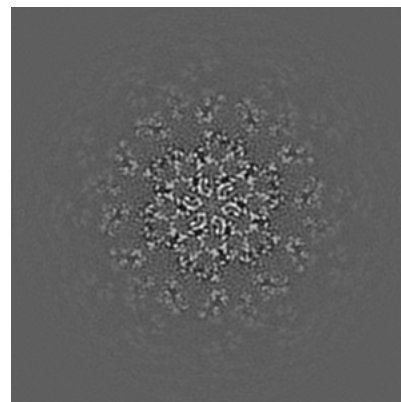
6.2.1 Primary map



X Index: 128

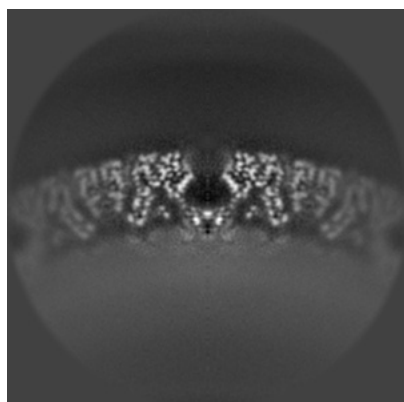


Y Index: 128

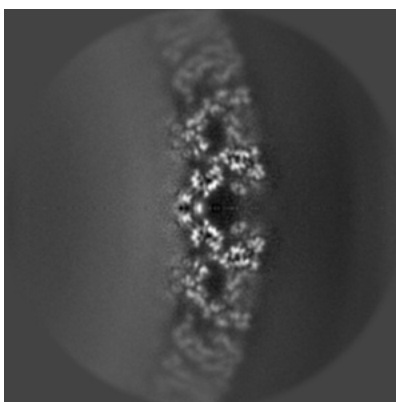


Z Index: 128

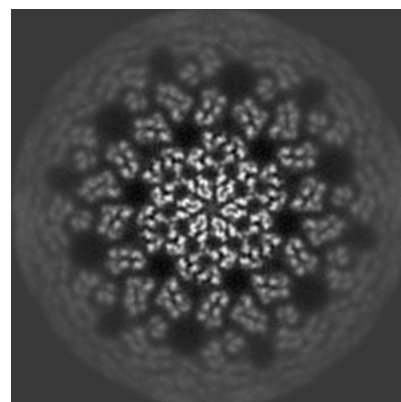
6.2.2 Raw map



X Index: 128



Y Index: 128

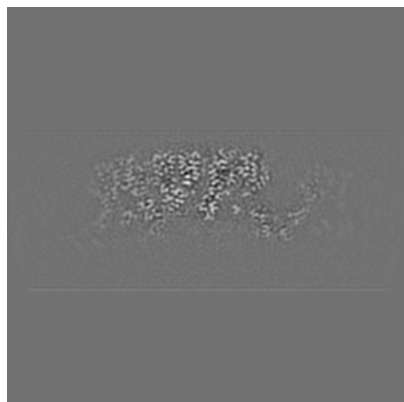


Z Index: 128

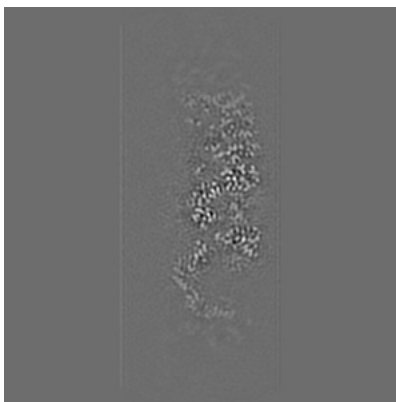
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices [i](#)

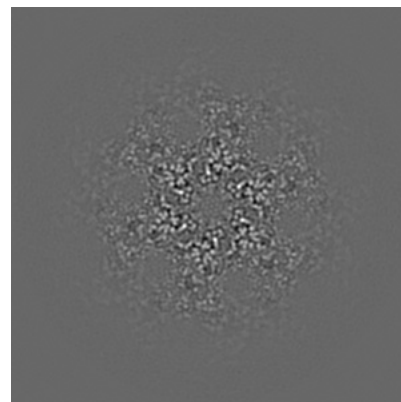
6.3.1 Primary map



X Index: 110

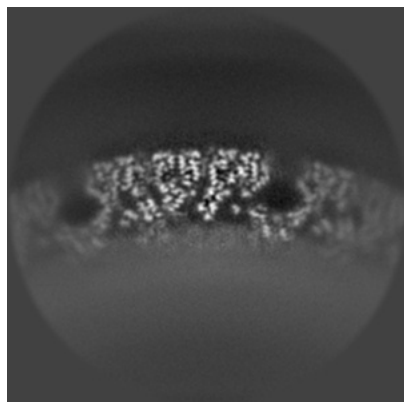


Y Index: 141

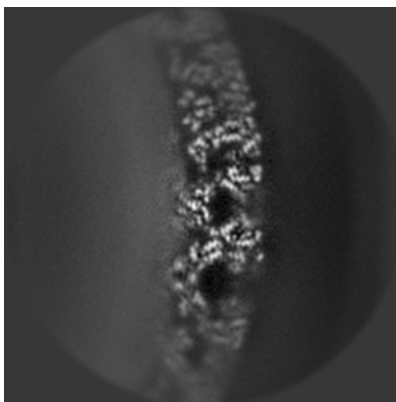


Z Index: 147

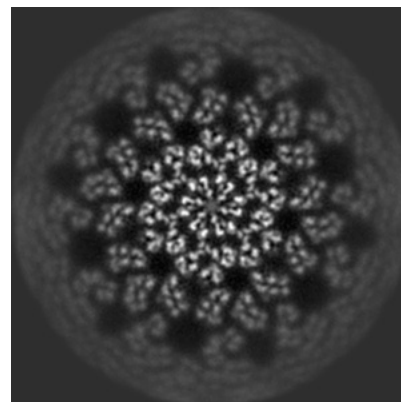
6.3.2 Raw map



X Index: 110



Y Index: 136

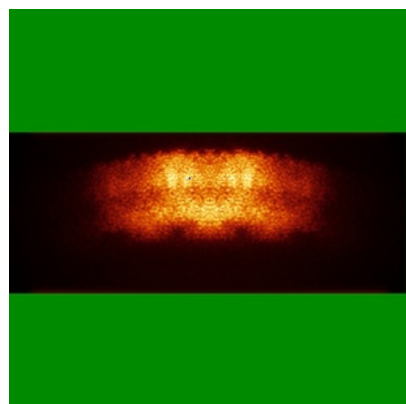


Z Index: 127

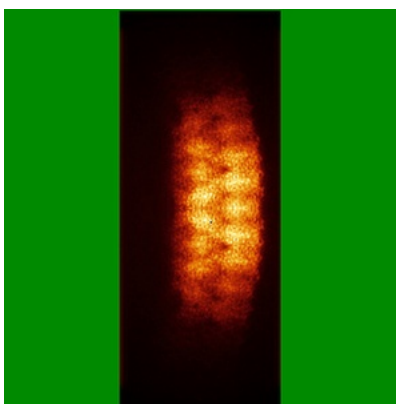
The images above show the largest variance slices of the map in three orthogonal directions.

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

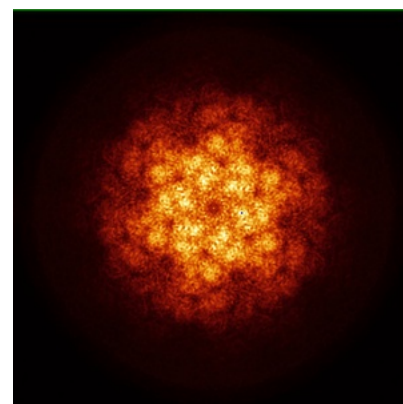
6.4.1 Primary map



X

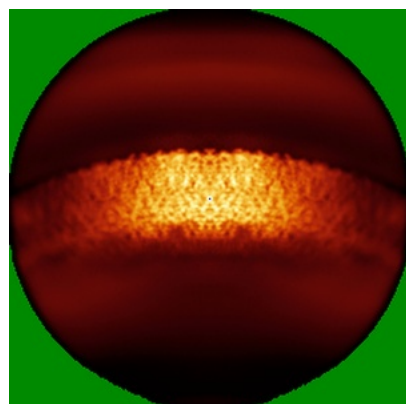


Y

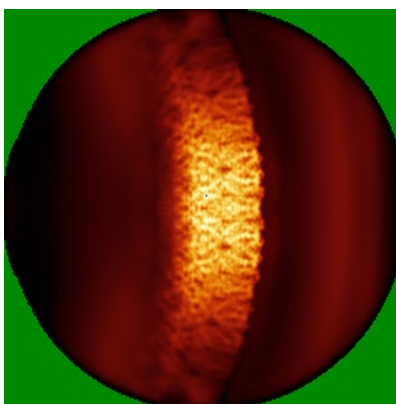


Z

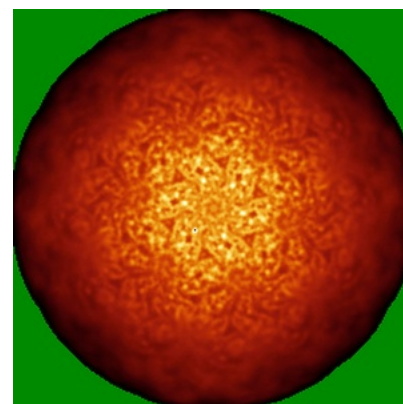
6.4.2 Raw map



X



Y

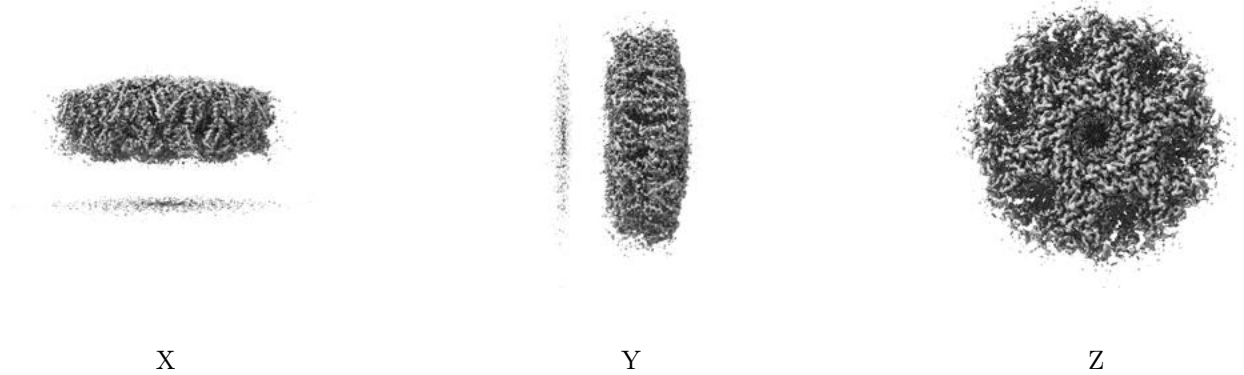


Z

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.

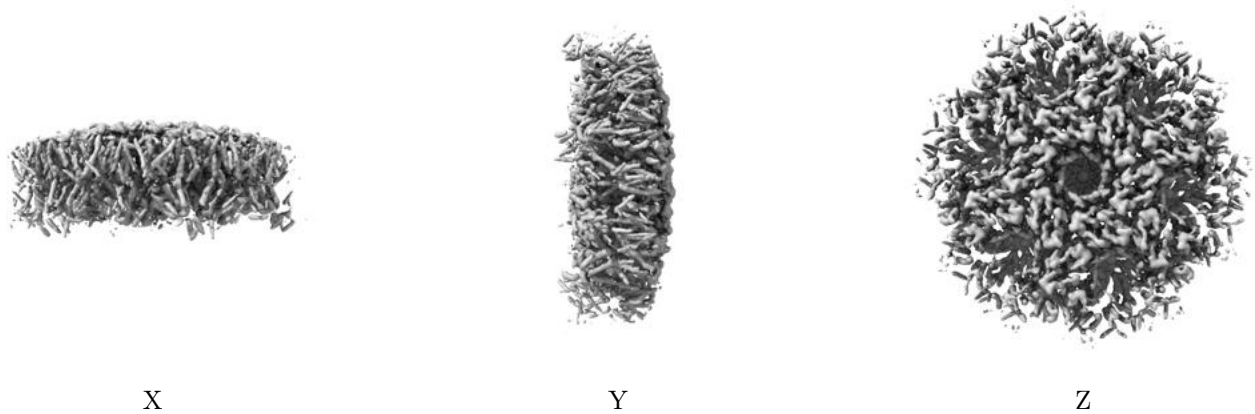
6.5 Orthogonal surface views [i](#)

6.5.1 Primary map



The images above show the 3D surface view of the map at the recommended contour level 0.0736. 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.

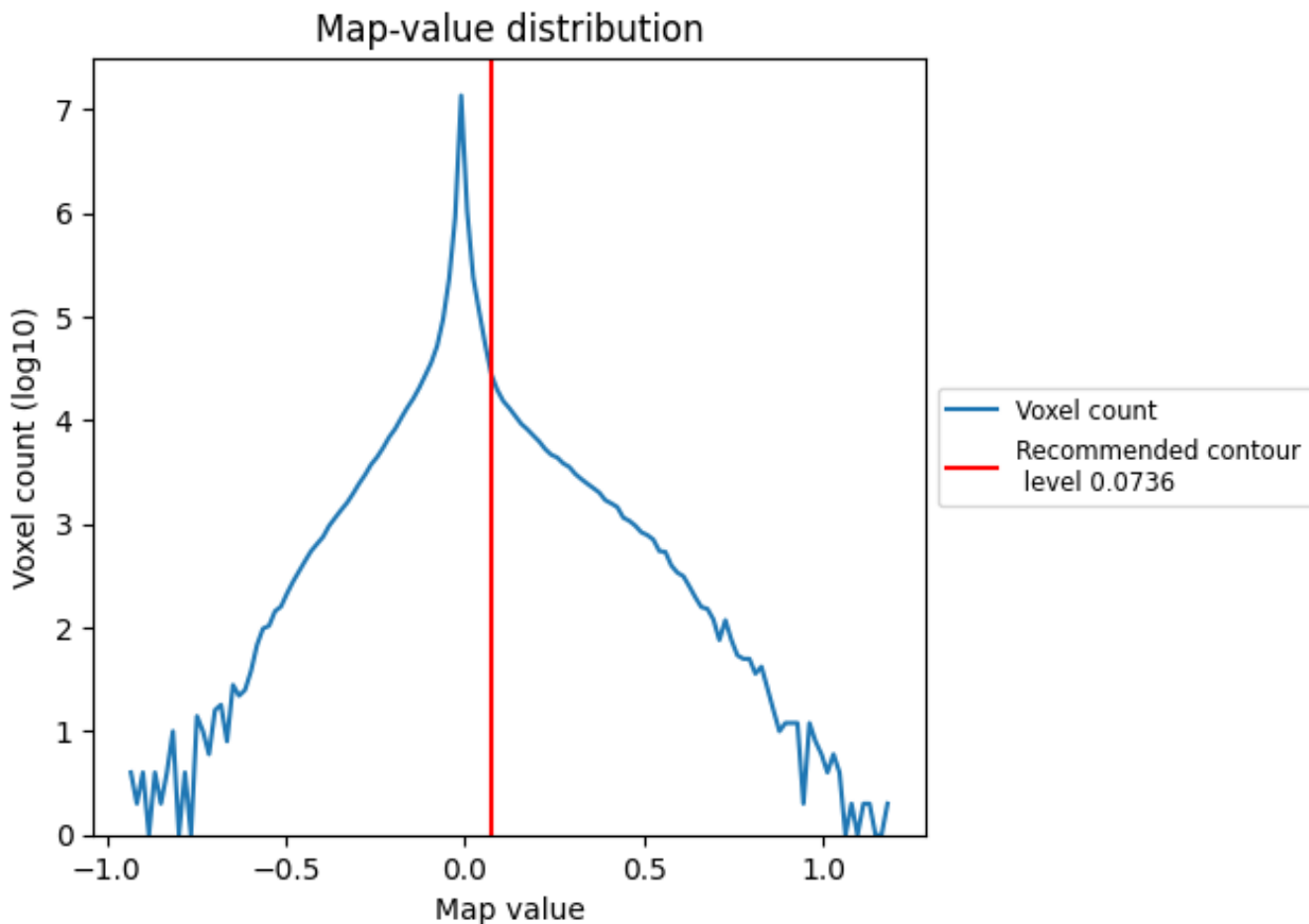
6.6 Mask visualisation [i](#)

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

7 Map analysis [i](#)

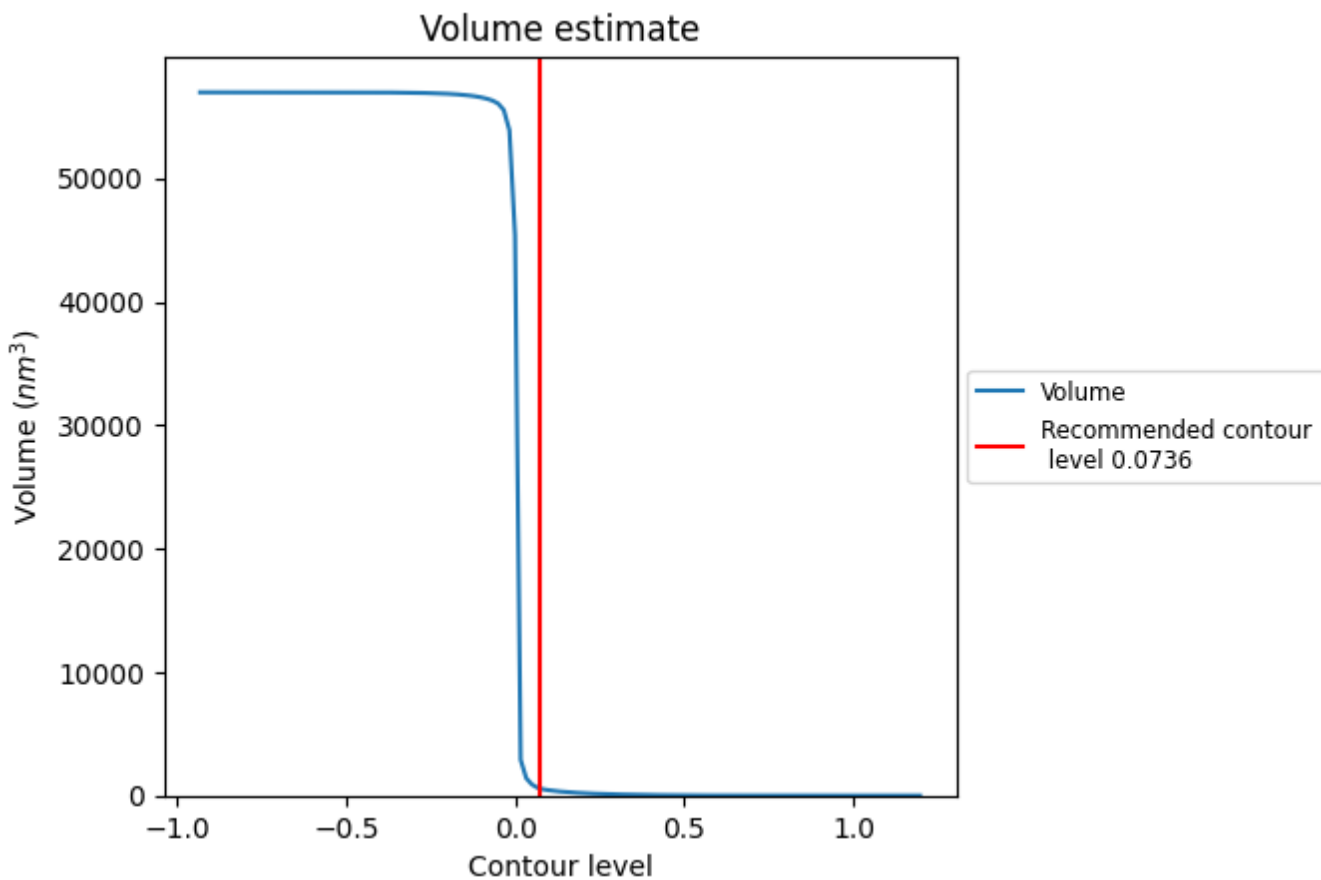
This section contains the results of statistical analysis of the map.

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.

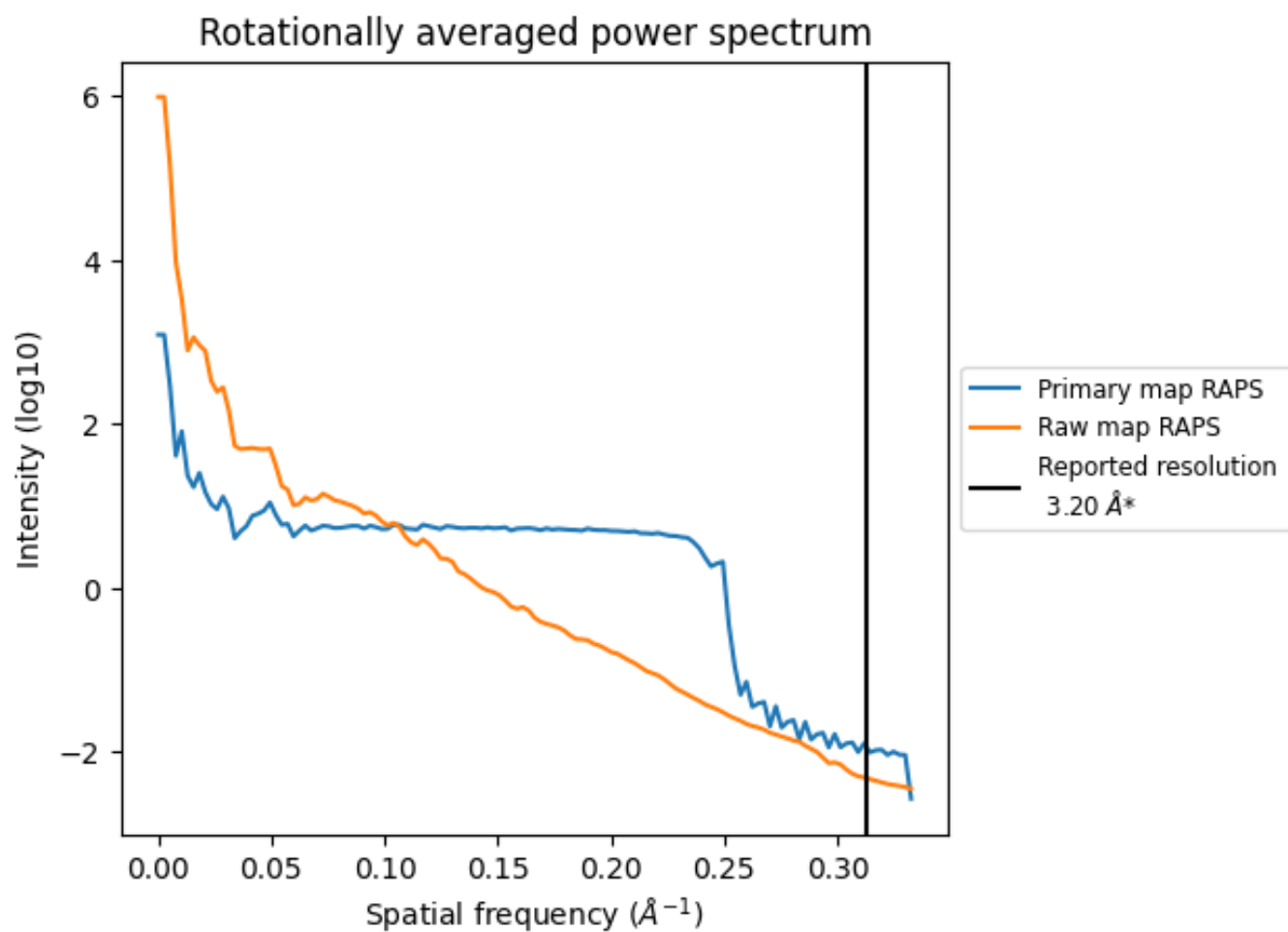
7.2 Volume estimate [i](#)



The volume at the recommended contour level is 579 nm³; this corresponds to an approximate mass of 523 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

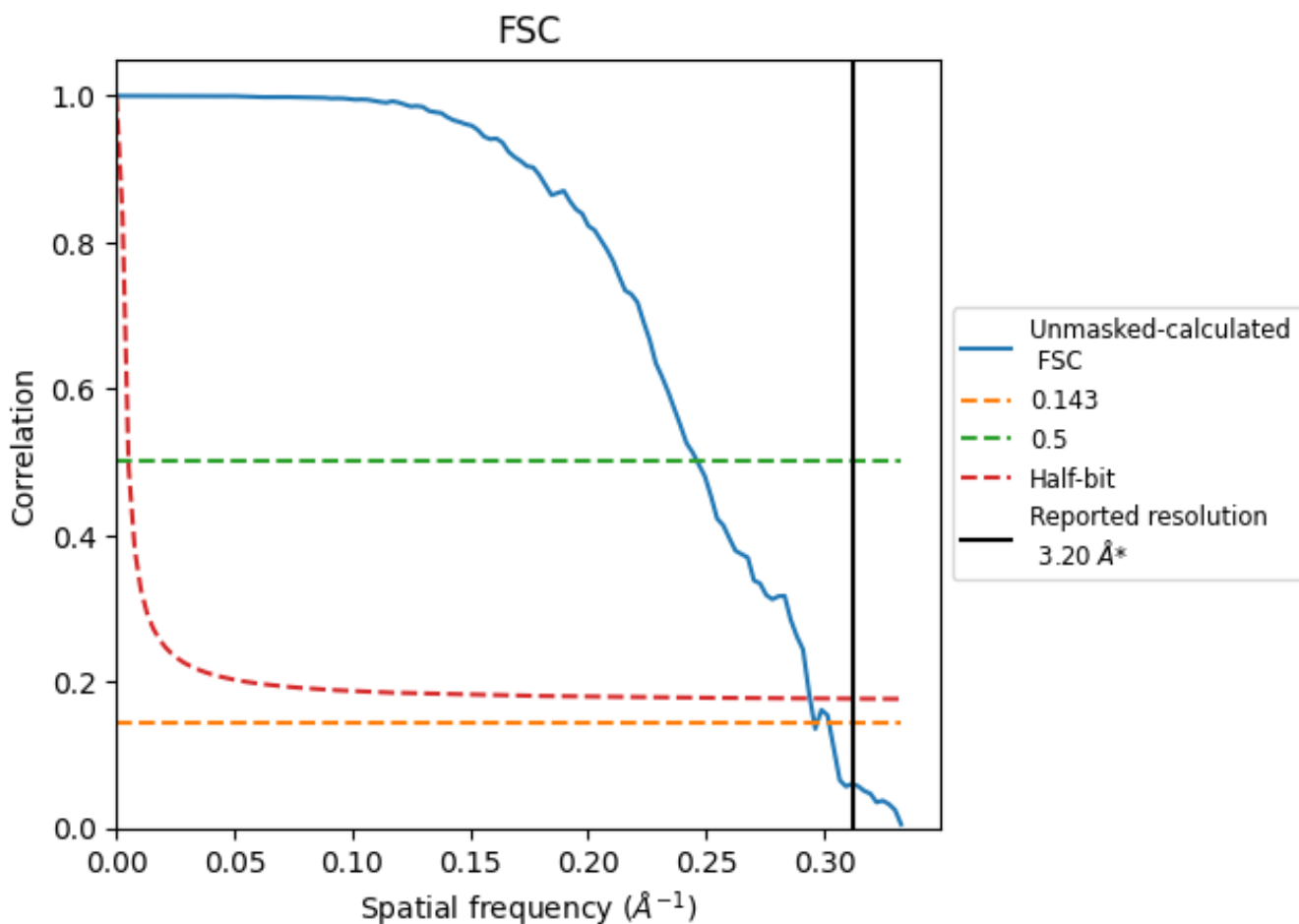


*Reported resolution corresponds to spatial frequency of 0.312 Å⁻¹

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 0.312 \AA^{-1}

8.2 Resolution estimates [i](#)

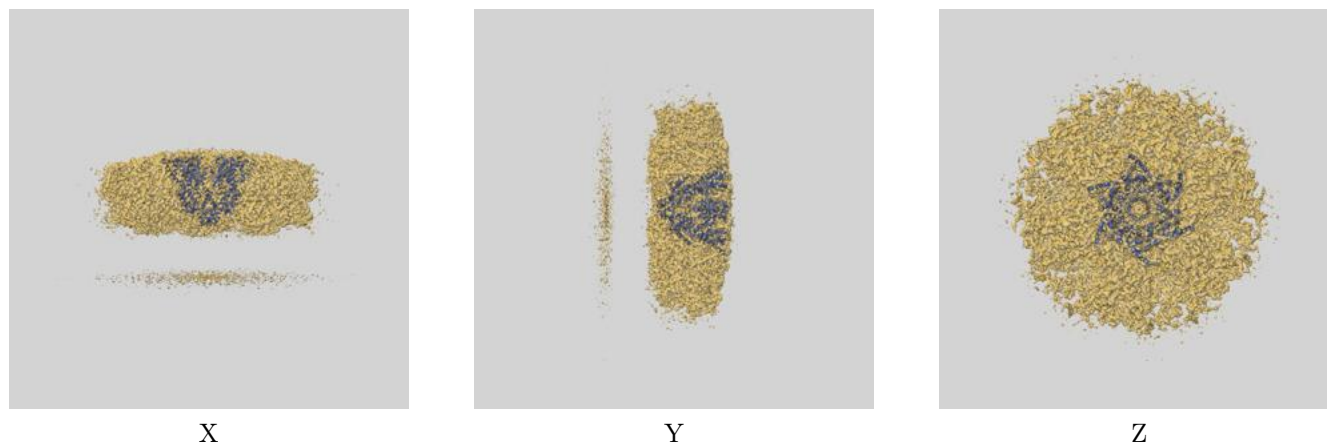
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	3.20	-	-
Author-provided FSC curve	-	-	-
Unmasked-calculated*	3.38	4.06	3.40

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

9 Map-model fit [i](#)

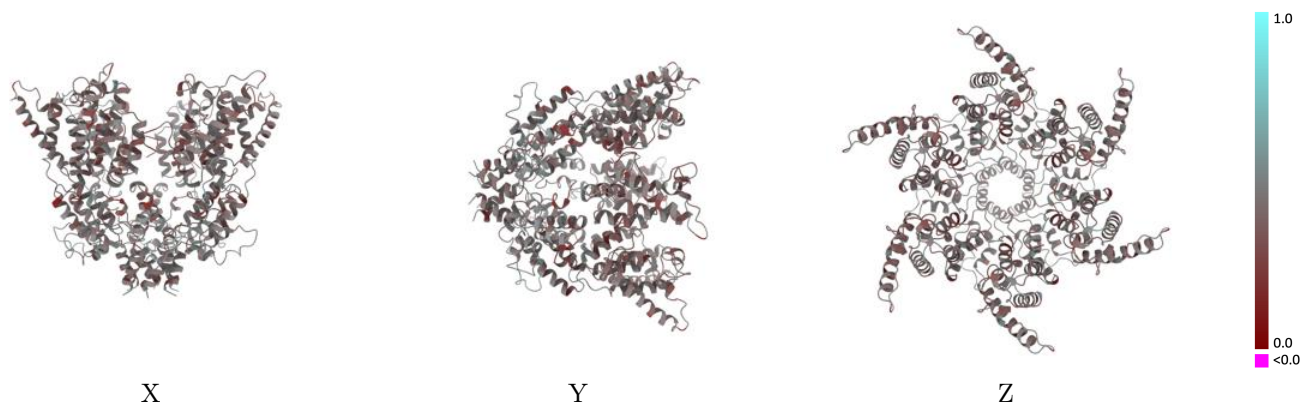
This section contains information regarding the fit between EMDB map EMD-16511 and PDB model 8C9M. Per-residue inclusion information can be found in section 3 on page 6.

9.1 Map-model overlay [i](#)



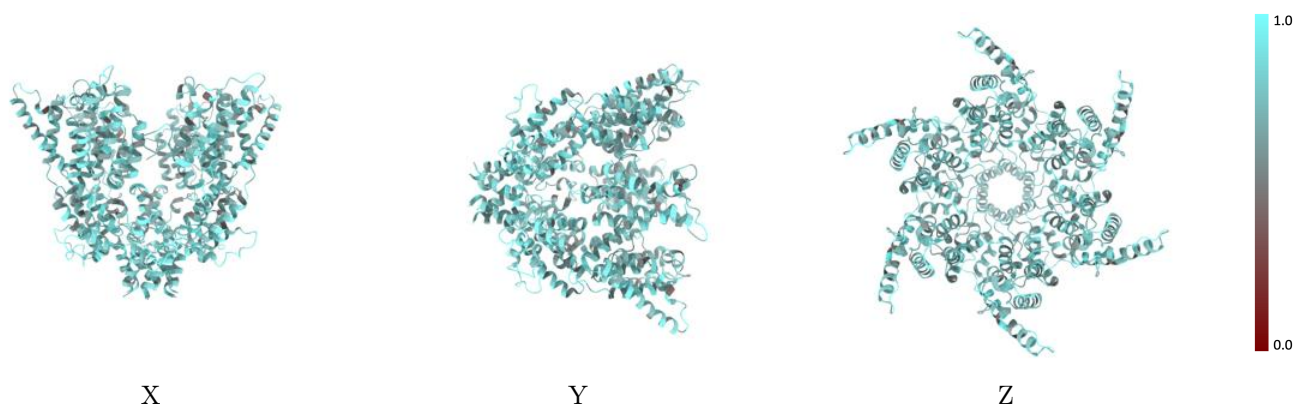
The images above show the 3D surface view of the map at the recommended contour level 0.0736 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

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



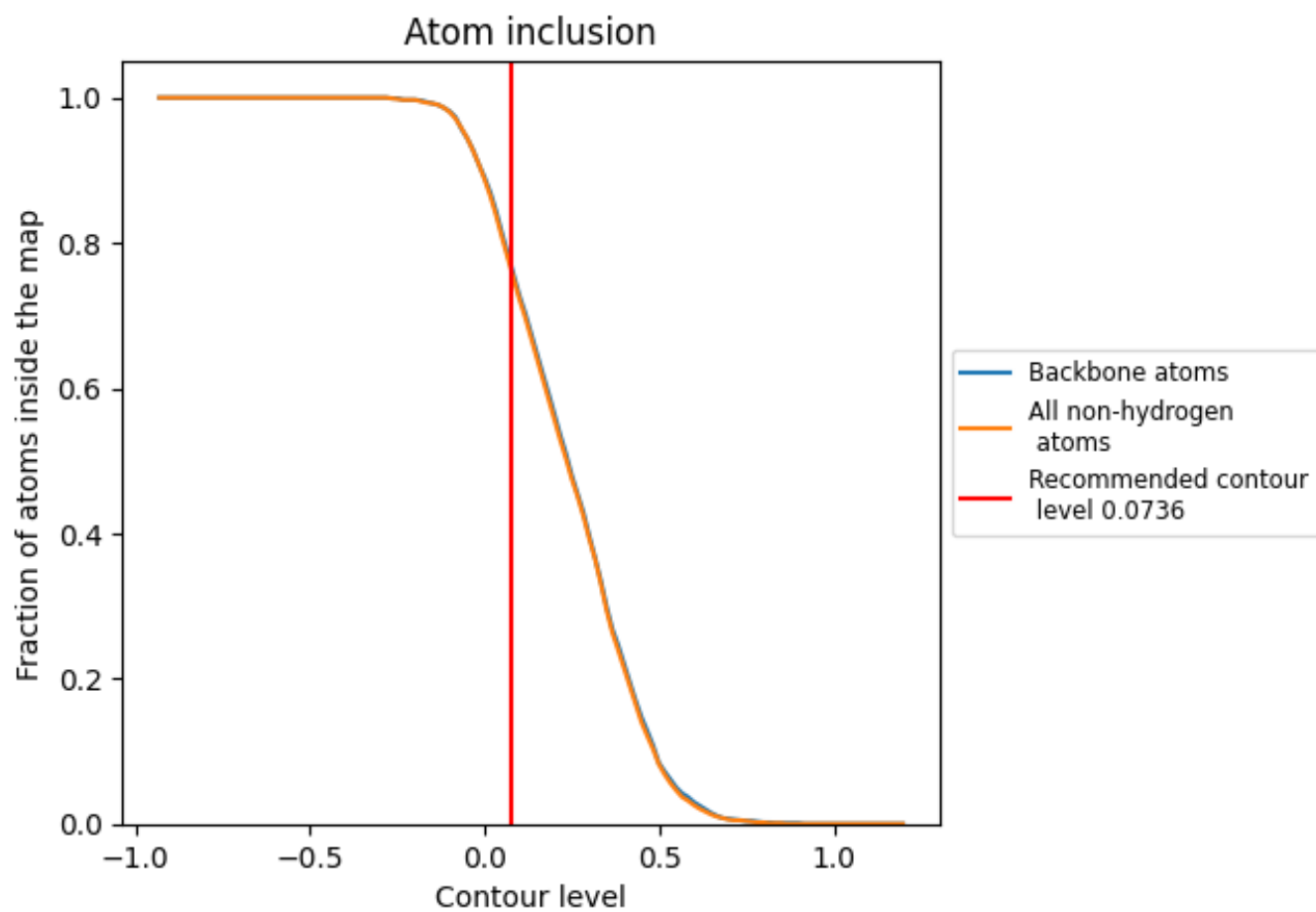
The images above show the model with each residue coloured according to its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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 (0.0736).















9.4 Atom inclusion [i](#)



At the recommended contour level, 77% of all backbone atoms, 77% of all non-hydrogen atoms, are inside the map.

9.5 Map-model fit summary [i](#)

The table lists the average atom inclusion at the recommended contour level (0.0736) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	 0.7660	 0.4140
A	 0.7710	 0.4170
B	 0.7760	 0.4160
C	 0.7760	 0.4100
D	 0.7720	 0.4090
E	 0.7730	 0.4150
F	 0.7790	 0.4150

