wwPDB EM Validation Summary Report (i)

Jul 3, 2024 – 06:58 pm BST

PDB ID	:	9EO9
EMDB ID	:	EMD-19849
Title	:	PHF type tau filament from V337M mutant
Authors	:	Qi, C.; Scheres, S.H.W.; Michel, G.
Deposited on	:	2024-03-14
Resolution	:	3.30 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 (i) 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 (i)) were used in the production of this report:

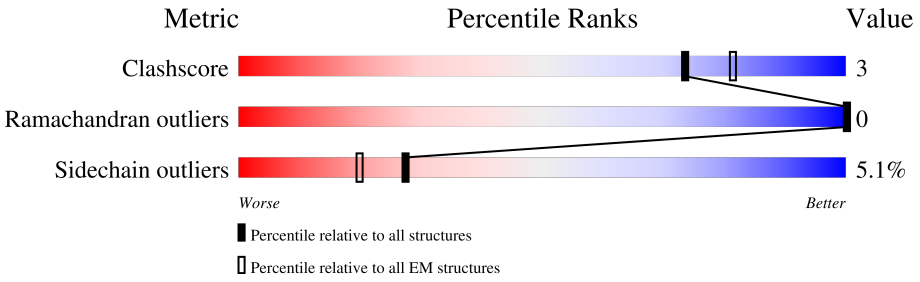
EMDB validation analysis	:	0.0.1.dev92
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.13
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

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 3.30 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 <=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	442	16% 82%
1	B	442	16% 82%

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 1190 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 Isoform Tau-F of Microtubule-associated protein tau.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	77	595	372	109	112	2	1	0
1	B	77	595	372	109	112	2	1	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	337	MET	VAL	variant	UNP P10636
B	337	MET	VAL	variant	UNP P10636

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: Isoform Tau-F of Microtubule-associated protein tau

Chain A:

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

SER GLU THR SER ASP ALA LYS SER THR PRO THR ALA GLU ASP VAL THR ALA PRO LEU VAL ASP GLU GLY ALA PRO GLY LYS GLN LYS ALA ALA ALA GLN PRO HIS THR THR GLU ILE PRO GLU GLY THR THR ALA GLU GLU ALA GLY ILE GLY ASP THR PRO SER LEU GLU ASP GLU ALA ALA GLY

HIS VAL THR GLN ALA ARG MET VAL SER LYS SER LYS ASP GLY THR GLY SER ASP ASP LYS LYS ALA LYS GLY ALA ASP GLY LYS THR LYS ILE ALA THR PRO ARG GLY ALA ALA PRO PRO GLY GLN LYS GLY GLN ALA ASN ALA THR ARG ILE PRO ALA LYS THR PRO PRO ALA PRO LYS

THR PRO PRO SER SER GLY GLU PRO PRO PRO LYS SER GLY ASP ARG SER GLY TYR SER SER PRO GLY SER PRO GLY THR PRO GLY SER ARG THR PRO SER LEU PRO THR PRO PRO THR ARG GLU PRO LYS LYS VAL ALA VAL VAL ARG THR PRO PRO LYS SER PRO SER SER ALA LYS

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

PRO GLY GLY G304 V318 S324 E342 K343 L344 D345 F346 K347 V350 I371 E372 E380 ASN ALA LYS ALA LYS THR ASP HIS GLY ALA GLU ILE VAL TYR LYS SER PRO VAL VAL SER GLY ASP THR SER PRO ARG HIS LEU SER ASN VAL SER SER THR GLY SER ILE ASP

MET VAL ASP SER PRO GLN LEU ALA THR LEU ALA ASP GLU VAL SER ALA SER LEU ALA LYS GLN GLY LEU

• Molecule 1: Isoform Tau-F of Microtubule-associated protein tau

Chain B:

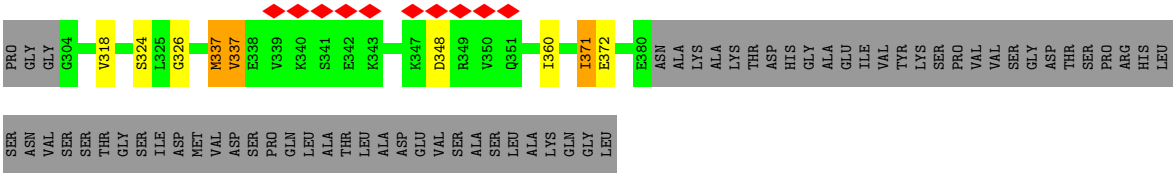
MET ALA GLU PRO ARG GLN GLU PHE GLU VAL MET GLU ASP HIS ALA GLY THR TYR GLY LEU GLY ASP ARG LYS ASP GLN GLY GLY TYR THR MET HIS GLN ASP GLN GLU GLY ASP THR ASP ALA GLY LEU LYS GLU SER PRO LEU GLN THR PRO THR GLU ASP GLY SER GLU GLU PRO GLY

SER GLU THR SER ASP ALA LYS SER THR PRO THR ALA GLU ASP VAL THR ALA PRO LEU VAL ASP GLU GLY ALA PRO GLY LYS GLN LYS ALA ALA ALA GLN PRO HIS THR THR GLU ILE PRO GLU GLY THR THR ALA GLU GLU ALA GLY ILE GLY ASP THR PRO SER LEU GLU ASP GLU ALA ALA GLY

HIS VAL THR GLN ALA ARG MET VAL SER LYS SER LYS ASP GLY THR GLY SER ASP ASP LYS LYS ALA LYS GLY ALA ASP GLY LYS THR LYS ILE ALA THR PRO ARG GLY ALA ALA PRO PRO GLY GLN LYS GLY GLN ALA ASN ALA THR ARG ILE PRO ALA LYS THR PRO PRO ALA PRO LYS

THR PRO PRO SER SER GLY GLU PRO PRO PRO LYS SER GLY ASP ARG SER GLY TYR SER SER PRO GLY SER PRO GLY THR PRO GLY SER ARG THR PRO SER LEU PRO THR PRO PRO THR ARG GLU PRO LYS LYS VAL ALA VAL VAL ARG THR PRO PRO LYS SER PRO SER SER ALA LYS

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


PRO
GLY
GLY
G304
V318
S324
L325
G326
M337
V337
E338
V339
K340
S341
E342
K343
K347
D348
R349
V350
Q351
I360
I371
E372
E380
ASN
ALA
LYS
ALA
LYS
THR
ASP
HIS
GLY
ALA
GLU
ILE
VAL
TYR
LYS
SER
PRO
VAL
VAL
SER
GLY
ASP
THR
SER
PRO
ARG
HIS
LEU
SER
ASN
VAL
SER
SER
THR
GLY
SER
ILE
ASP
MET
VAL
ASP
SER
PRO
GLN
LEU
ALA
THR
LEU
ALA
ASP
GLU
VAL
SER
ALA
SER
LEU
ALA
LYS
GLN
GLY
LEU

4 Experimental information (i)

Property	Value	Source
EM reconstruction method	HELICAL	Depositor
Imposed symmetry	HELICAL, twist=-1.07°, rise=4.76 Å, axial sym=C1	Depositor
Number of segments used	75617	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	NONE	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/Å^2$)	40	Depositor
Minimum defocus (nm)	1000	Depositor
Maximum defocus (nm)	2000	Depositor
Magnification	Not provided	
Image detector	FEI FALCON IV (4k x 4k)	Depositor
Maximum map value	0.040	Depositor
Minimum map value	-0.024	Depositor
Average map value	0.000	Depositor
Map value standard deviation	0.002	Depositor
Recommended contour level	0.008	Depositor
Map size (Å)	329.6, 329.6, 329.6	wwPDB
Map dimensions	400, 400, 400	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	0.824, 0.824, 0.824	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	0.36	0/603	0.72	0/804
1	B	0.45	1/603 (0.2%)	0.78	0/804
All	All	0.41	1/1206 (0.1%)	0.75	0/1608

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1
1	B	0	2
All	All	0	3

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	337[B]	VAL	C-N	5.82	1.47	1.34

There are no bond angle outliers.

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	318	VAL	Peptide
1	B	318	VAL	Peptide
1	B	326	GLY	Peptide

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	595	0	613	3	0
1	B	595	0	613	3	0
All	All	1190	0	1226	6	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 3.

The worst 5 of 6 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:A:346:PHE:HD1	1:A:350:VAL:HG11	1.71	0.56
1:B:360:ILE:HD12	1:B:360:ILE:N	2.26	0.51
1:A:371:ILE:HD12	1:A:371:ILE:C	2.33	0.49
1:B:371:ILE:C	1:B:371:ILE:HD12	2.35	0.47
1:B:371:ILE:HD12	1:B:372:GLU:N	2.29	0.47

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	74/442 (17%)	68 (92%)	6 (8%)	0	100	100
1	B	74/442 (17%)	69 (93%)	5 (7%)	0	100	100
All	All	148/884 (17%)	137 (93%)	11 (7%)	0	100	100

There are no Ramachandran outliers to report.

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	68/359 (19%)	65 (96%)	3 (4%)	28	59
1	B	68/359 (19%)	64 (94%)	4 (6%)	19	49
All	All	136/718 (19%)	129 (95%)	7 (5%)	27	54

5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	324	SER
1	B	337[A]	MET
1	B	371	ILE
1	B	348	ASP
1	A	371	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	362	HIS
1	B	362	HIS

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)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	B	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	B	336:GLN	C	337[B]:VAL	N	1.79

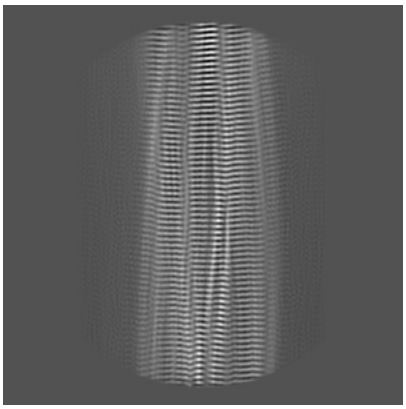
6 Map visualisation (i)

This section contains visualisations of the EMDB entry EMD-19849. These allow visual inspection of the internal detail of the map and identification of artifacts.

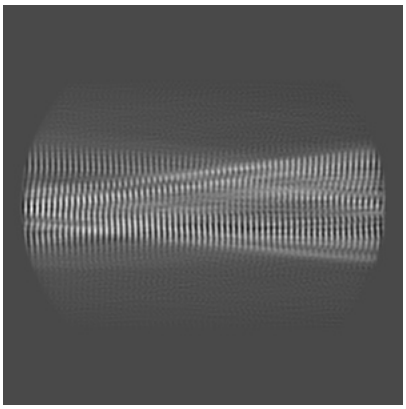
Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

6.1 Orthogonal projections (i)

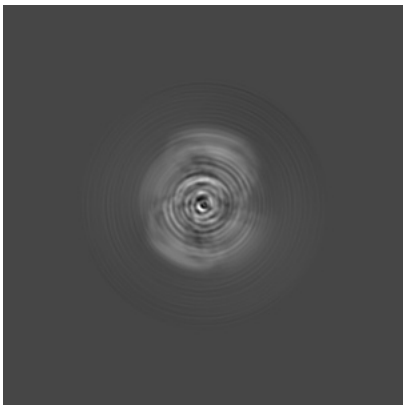
6.1.1 Primary map



X

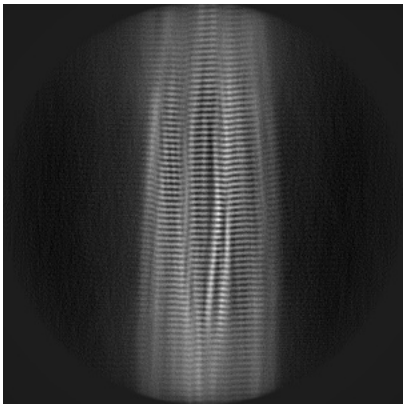


Y

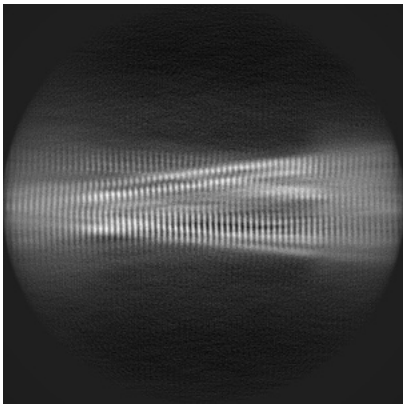


Z

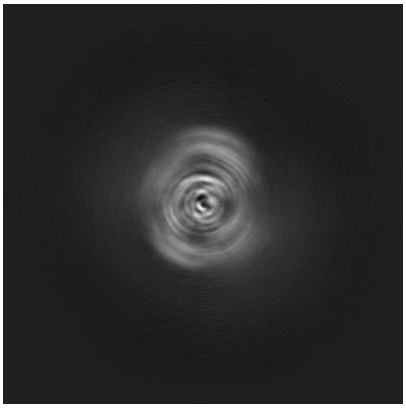
6.1.2 Raw map



X



Y

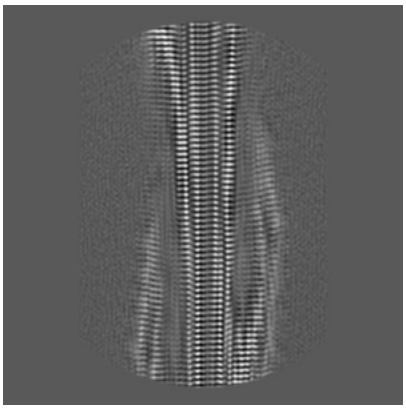


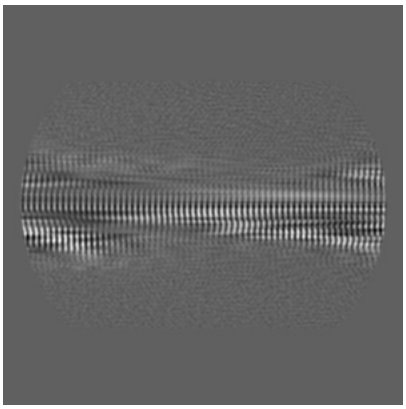
Z

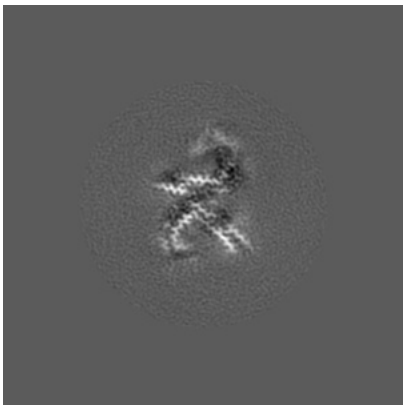
The images above show the map projected in three orthogonal directions.

6.2 Central slices (i)

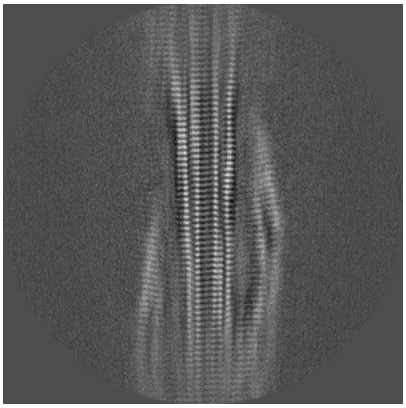
6.2.1 Primary map

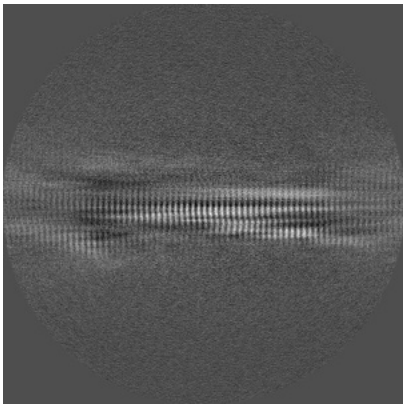

X Index: 200

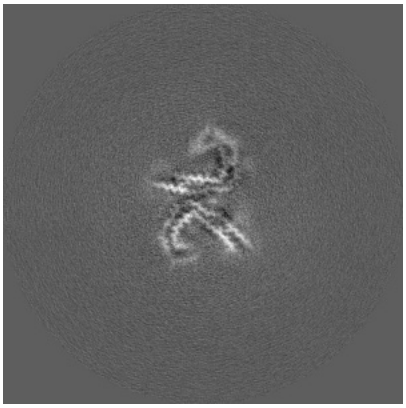

Y Index: 200


Z Index: 200

6.2.2 Raw map


X Index: 200

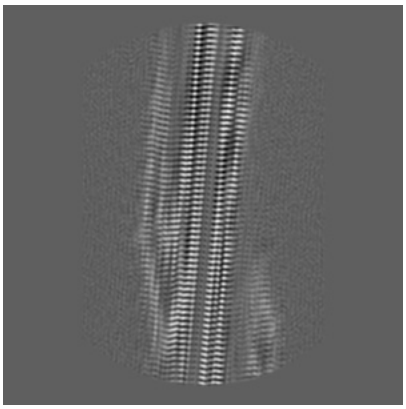

Y Index: 200

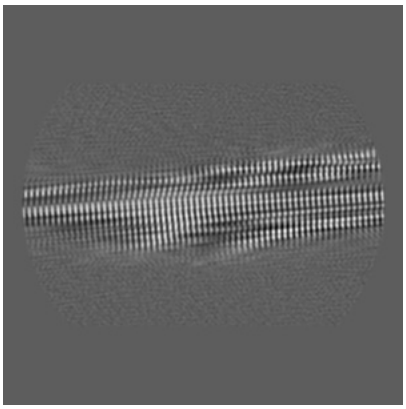

Z Index: 200

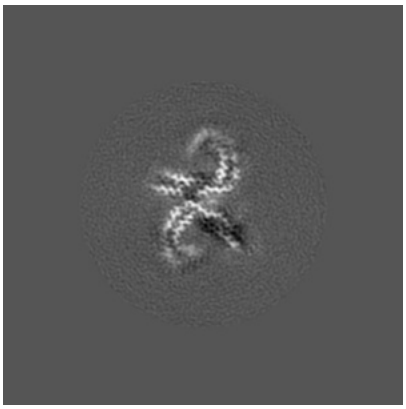
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

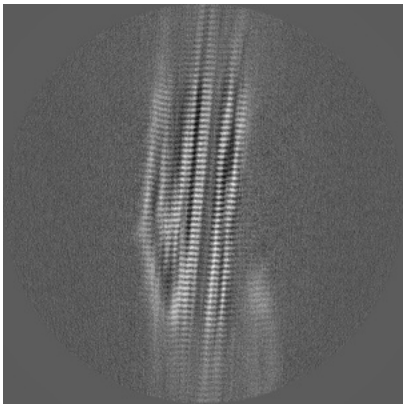
6.3.1 Primary map

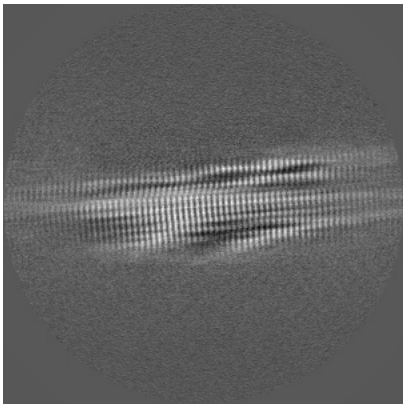

X Index: 175

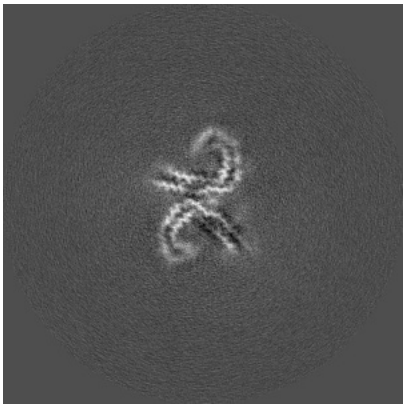

Y Index: 224


Z Index: 190

6.3.2 Raw map


X Index: 174

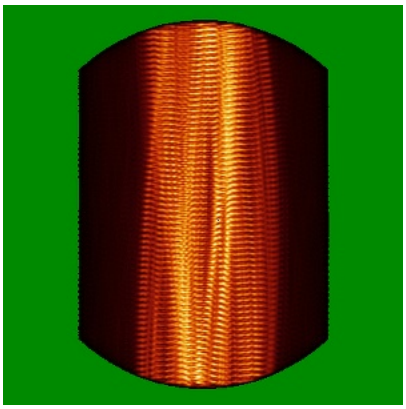

Y Index: 224

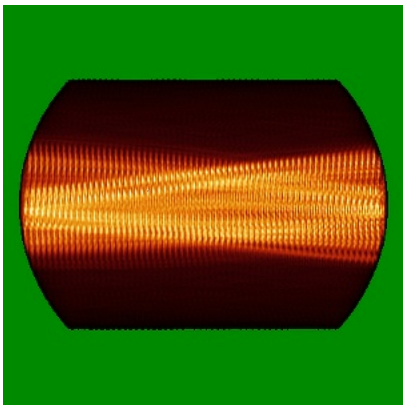

Z Index: 213

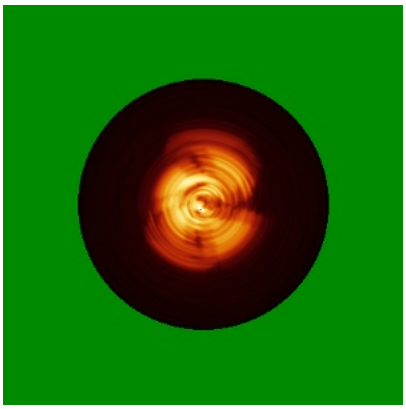
The images above show the largest variance slices of the map in three orthogonal directions.

6.4 Orthogonal standard-deviation projections (False-color) (i)

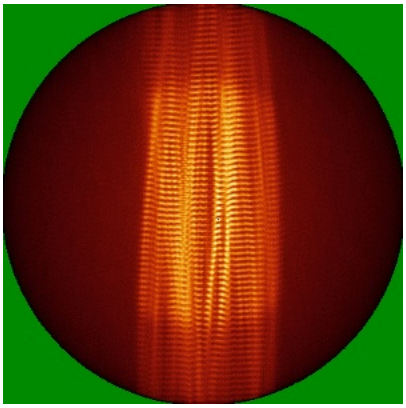
6.4.1 Primary map

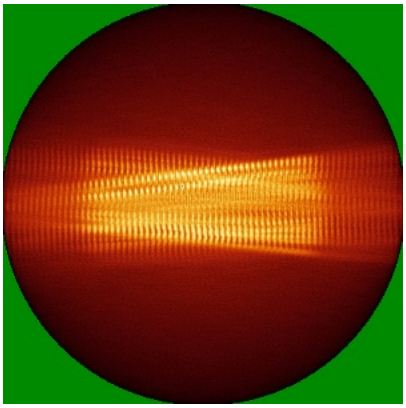

X

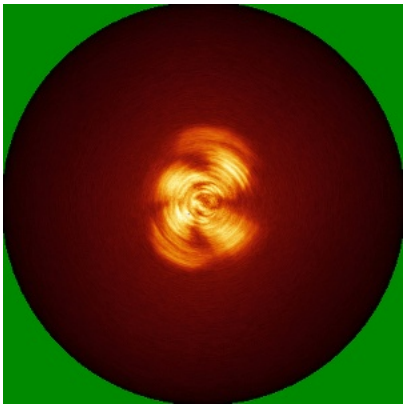

Y


Z

6.4.2 Raw map


X

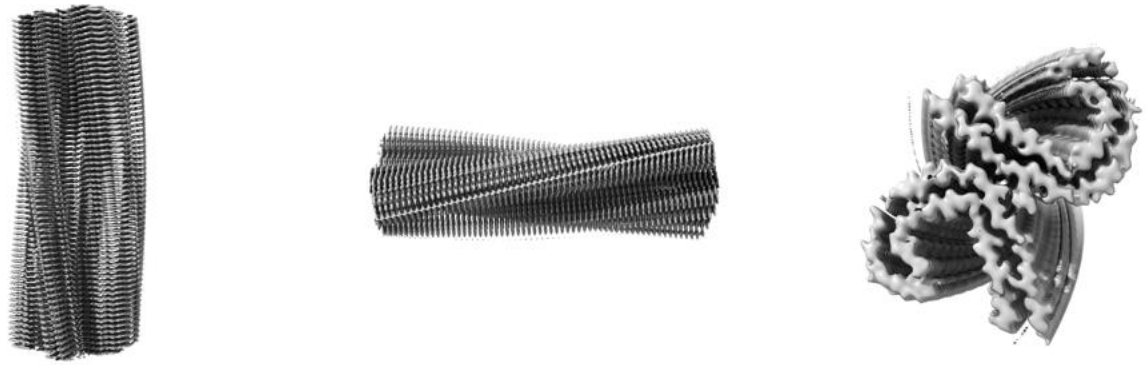

Y


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)

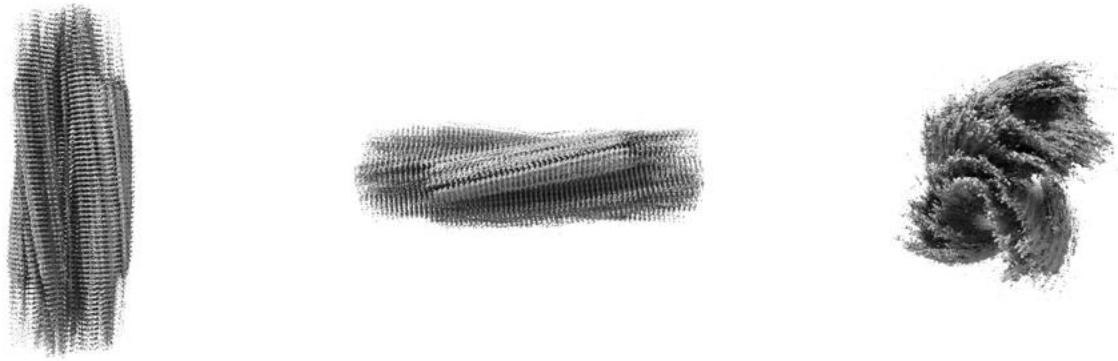
6.5.1 Primary map



X Y Z

The images above show the 3D surface view of the map at the recommended contour level 0.008. 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



X Y Z

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.

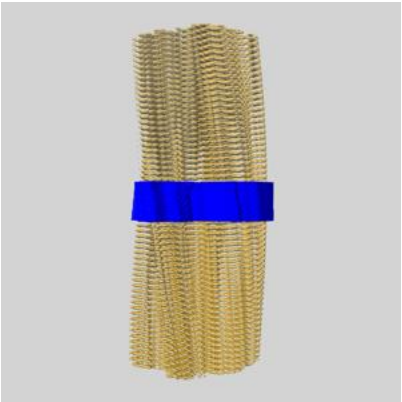
6.6 Mask visualisation (i)

This section shows the 3D surface view of the primary map at 50% transparency overlaid with the specified mask at 0% transparency

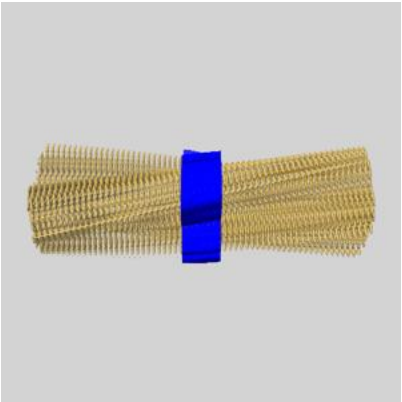
A mask typically either:

- Encompasses the whole structure
- Separates out a domain, a functional unit, a monomer or an area of interest from a larger structure

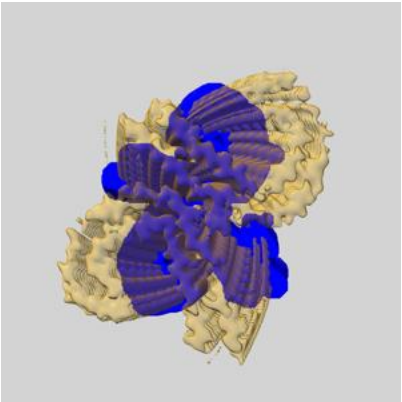
6.6.1 emd_19849_msk_1.map (i)



X



Y

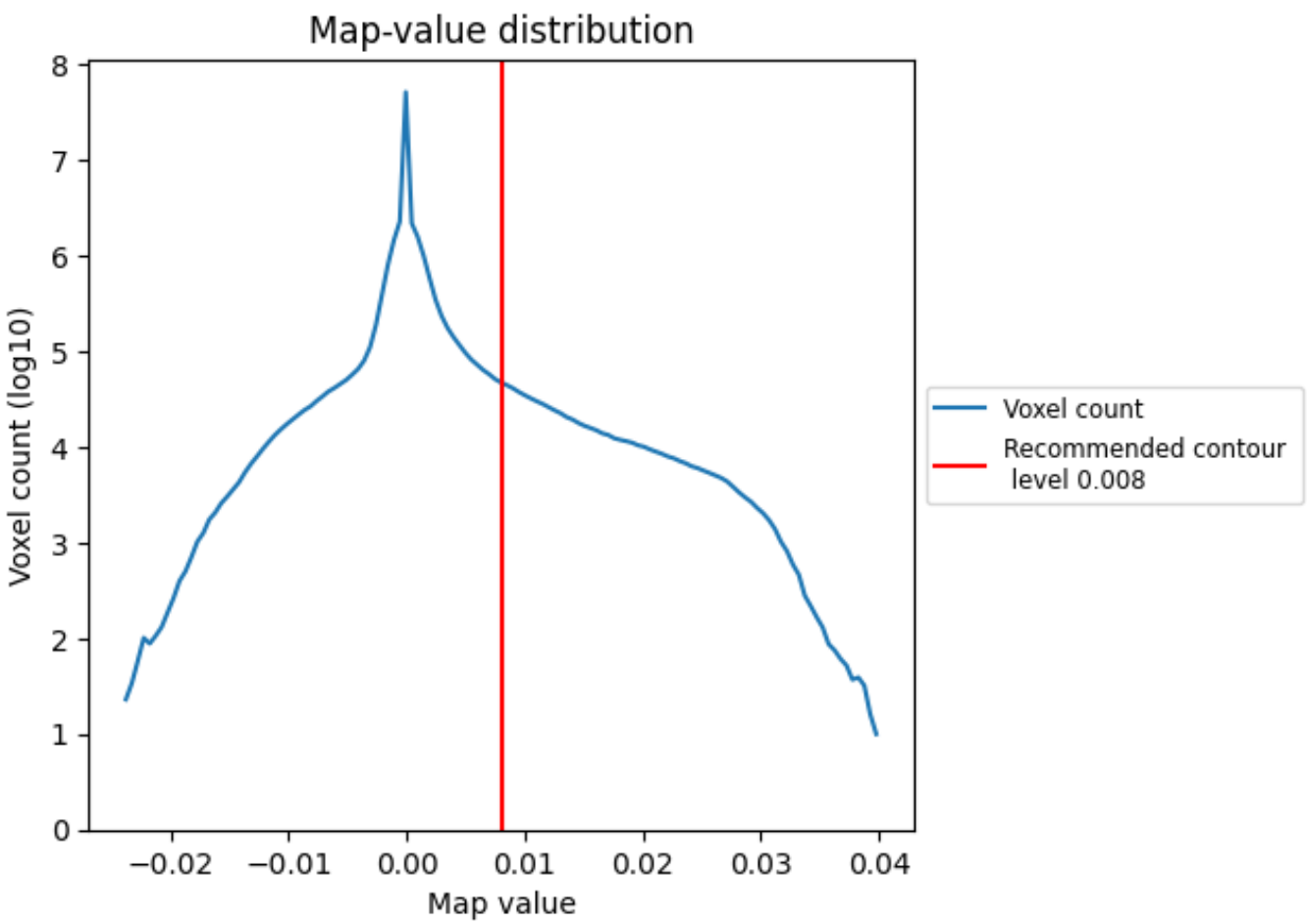


Z

7 Map analysis (i)

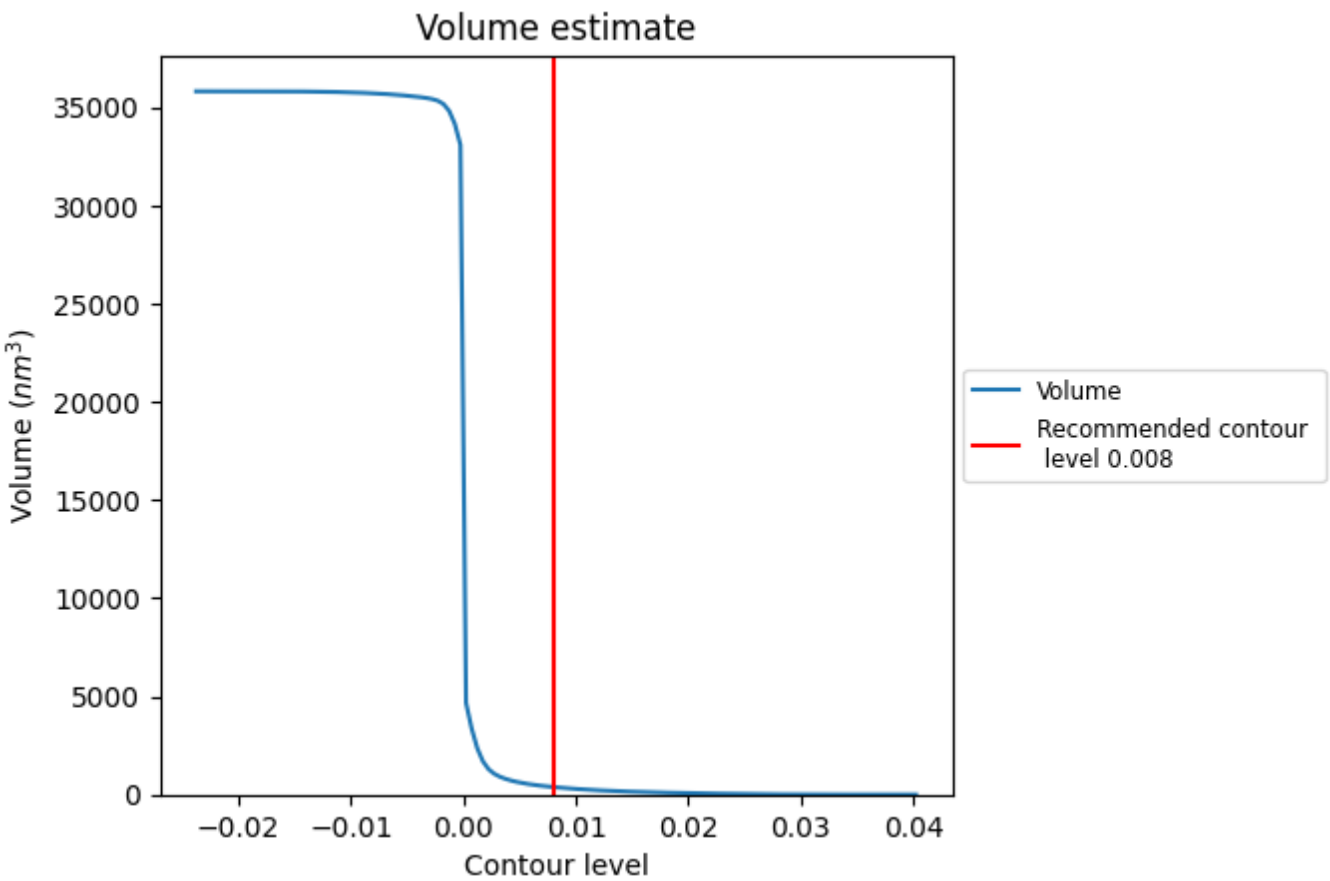
This section contains the results of statistical analysis of the map.

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.

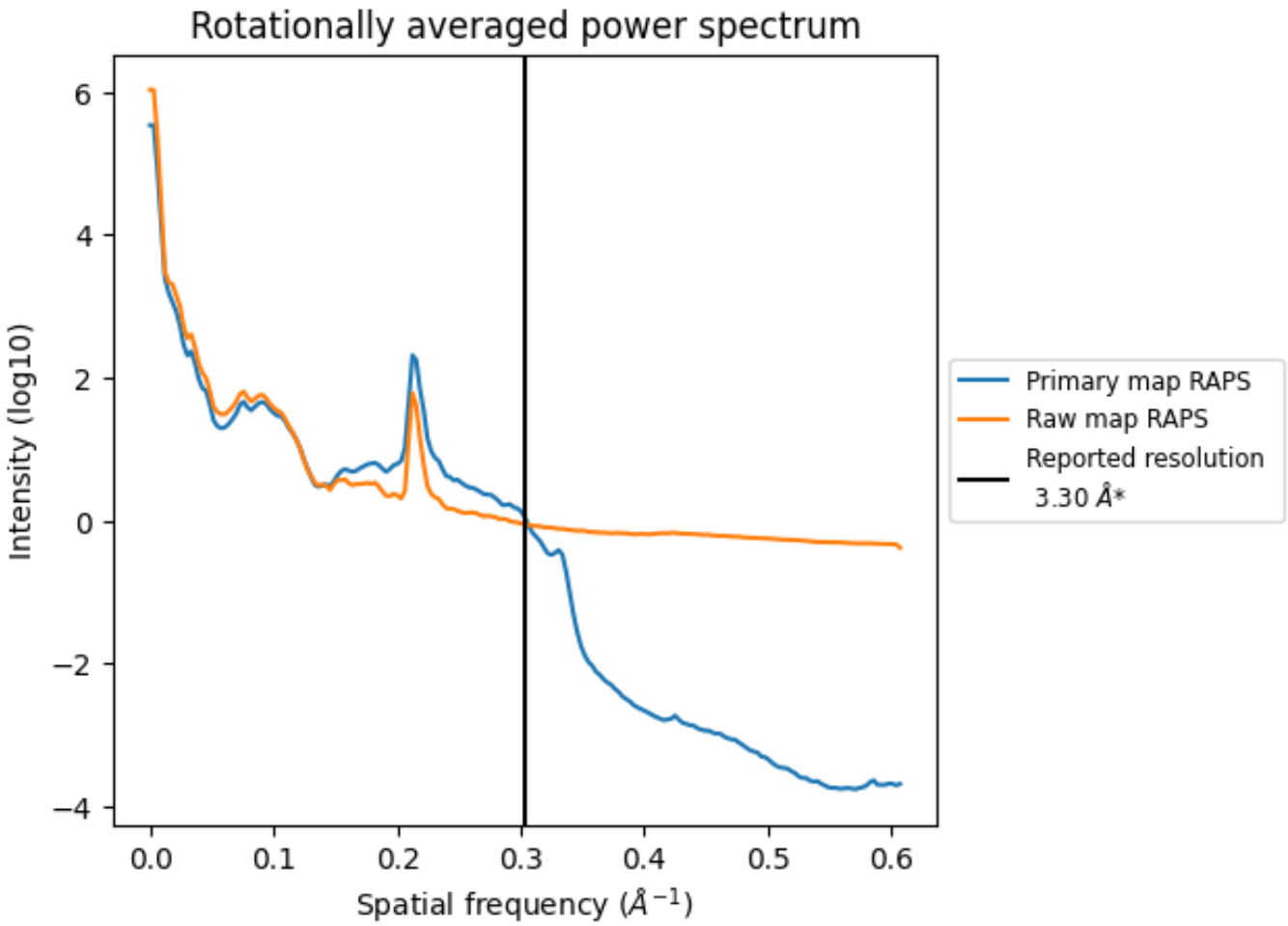
7.2 Volume estimate (i)



The volume at the recommended contour level is 382 nm^3; this corresponds to an approximate mass of 345 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)

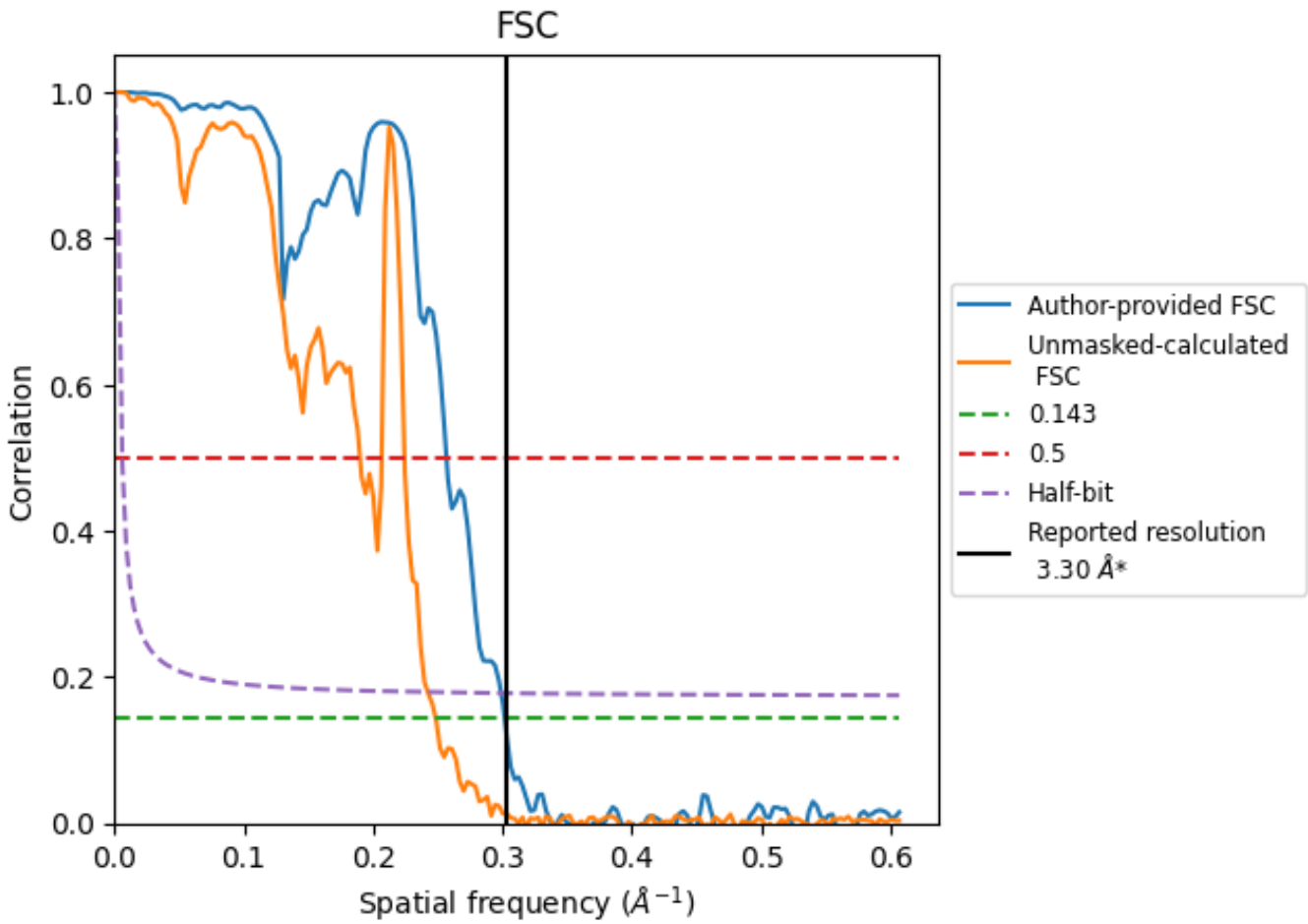


*Reported resolution corresponds to spatial frequency of 0.303 Å^{-1}

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 0.303 Å^{-1}

8.2 Resolution estimates (i)

Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	3.30	-	-
Author-provided FSC curve	3.31	3.90	3.35
Unmasked-calculated*	4.03	5.27	4.13

*Resolution estimate based on FSC curve calculated by comparison of deposited half-maps. The value from deposited half-maps intersecting FSC 0.143 CUT-OFF 4.03 differs from the reported value 3.3 by more than 10 %

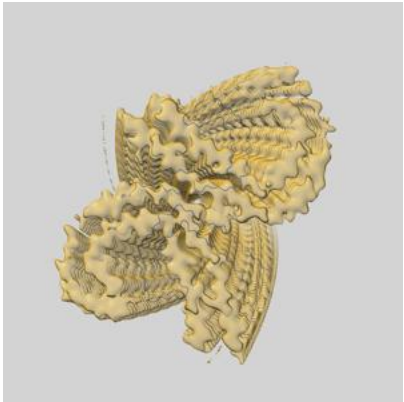
9 Map-model fit (i)

This section contains information regarding the fit between EMDB map EMD-19849 and PDB model 9EO9. Per-residue inclusion information can be found in section 3 on page 4.

9.1 Map-model overlay (i)

X

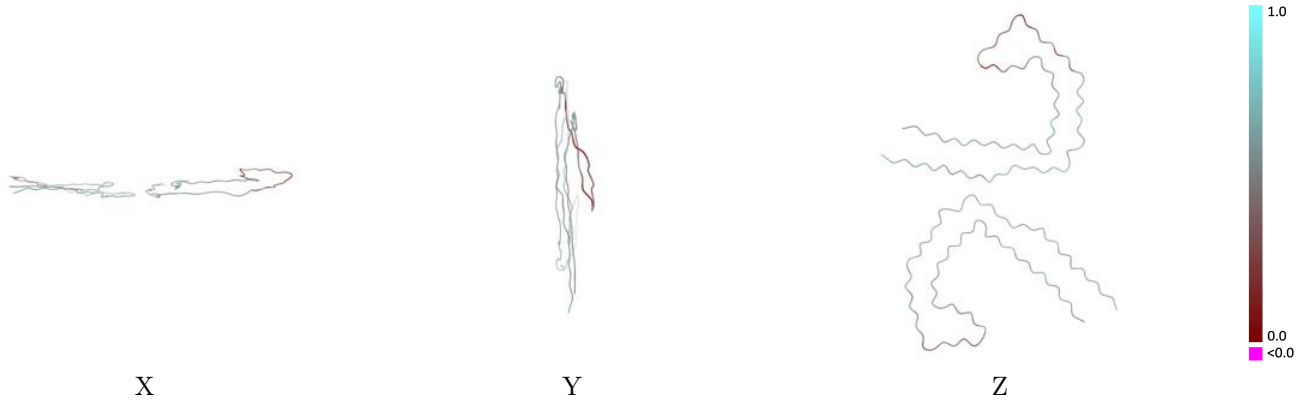
Y



Z

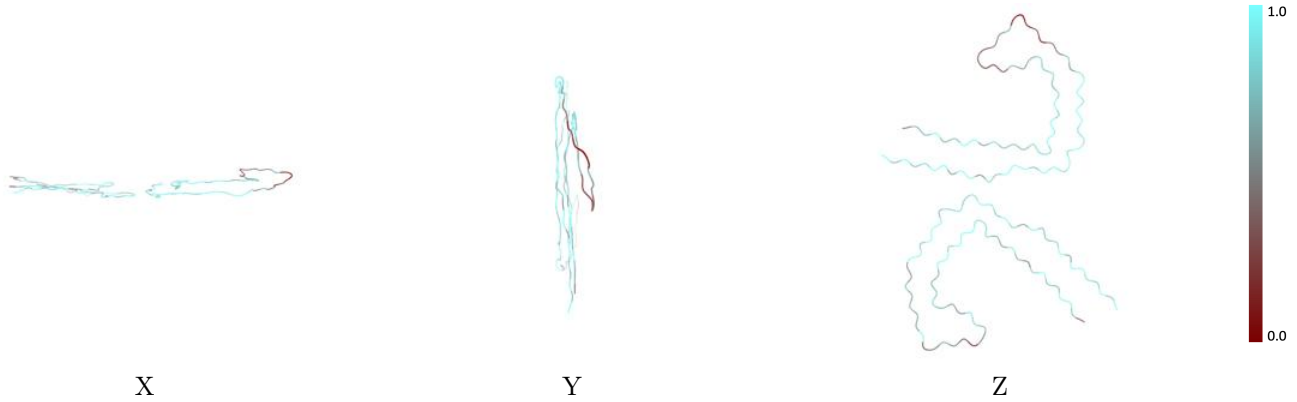
The images above show the 3D surface view of the map at the recommended contour level 0.008 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

9.2 Q-score mapped to coordinate model (i)



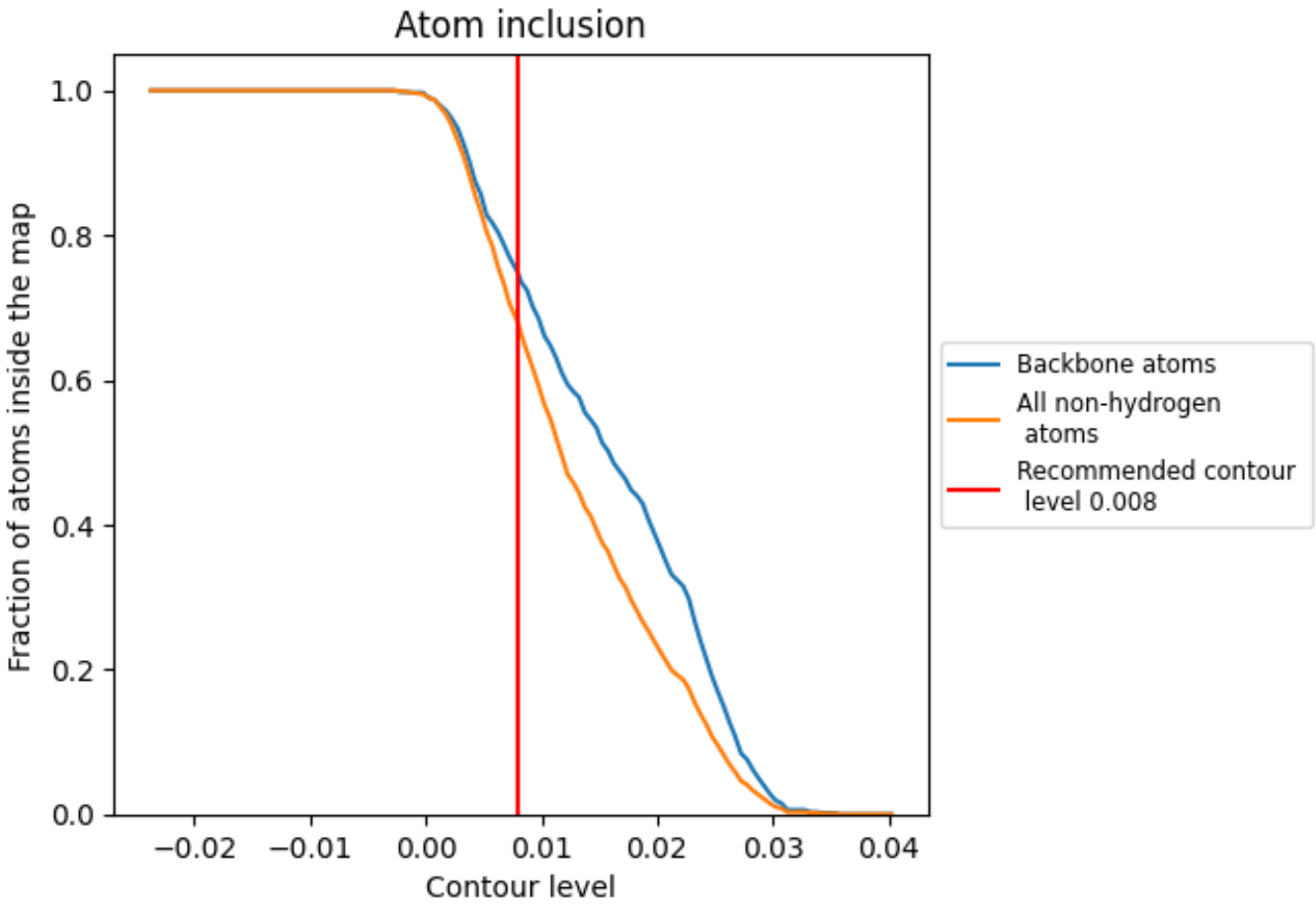
The images above show the model with each residue coloured according its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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 (0.008).

9.4 Atom inclusion (i)



At the recommended contour level, 75% of all backbone atoms, 68% 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 (0.008) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.6780	0.4860
A	0.6910	0.5000
B	0.6650	0.4720