Full wwPDB EM Validation Report (i)

May 18, 2024 – 08:28 pm BST

PDB ID	:	6TJX
EMDB ID	:	EMD-10514
Title	:	Cryo-EM structure of TypeII tau filaments extracted from the brains of individuals with Corticobasal degeneration
Authors	:	Zhang, W.; Murzin, A.G.; Falcon, B.; Shi, Y.; Goedert, M.; Scheres, S.H.W.
Deposited on	:	2019-11-27
Resolution	:	3.00 Å(reported)

This is a Full wwPDB EM Validation 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:

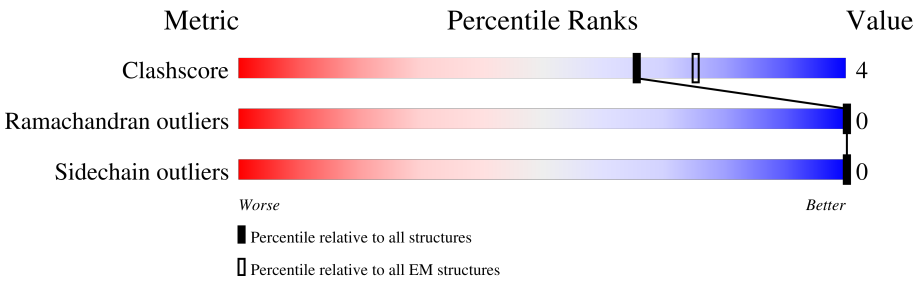
EMDB validation analysis	:	0.0.1.dev92
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.13
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

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 3.00 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	441	22% • 76%
1	B	441	22% • 76%
1	C	441	22% • 76%
1	D	441	22% • 76%
1	E	441	22% • 76%
1	F	441	22% • 76%

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 4878 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 Microtubule-associated protein tau.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	107	813	507	151	153	2	0	0
1	B	107	813	507	151	153	2	0	0
1	C	107	813	507	151	153	2	0	0
1	D	107	813	507	151	153	2	0	0
1	E	107	813	507	151	153	2	0	0
1	F	107	813	507	151	153	2	0	0

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: 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 ALA ALA ALA GLN PRO HIS 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 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 K274 K281 L282 D295 K298 S320 I328 P332 K347 D348 G365 G366 G367 K370 I371 E372 T373 H374

K375 L376 T377 F378 R379 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: Microtubule-associated protein tau

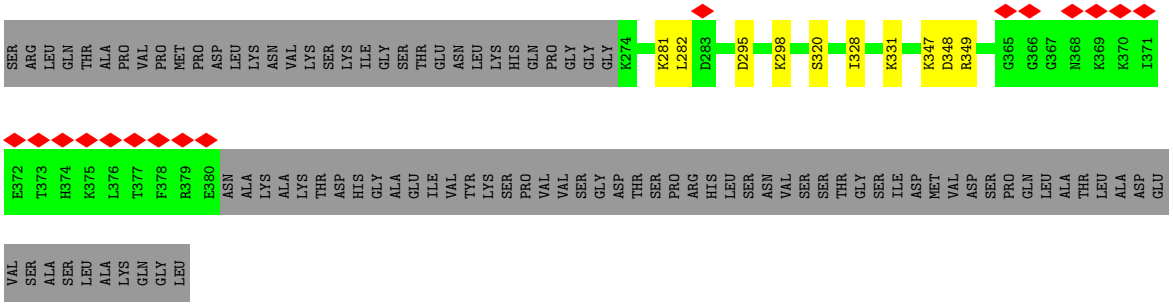
Chain B:

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 ALA ALA ALA GLN PRO HIS 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

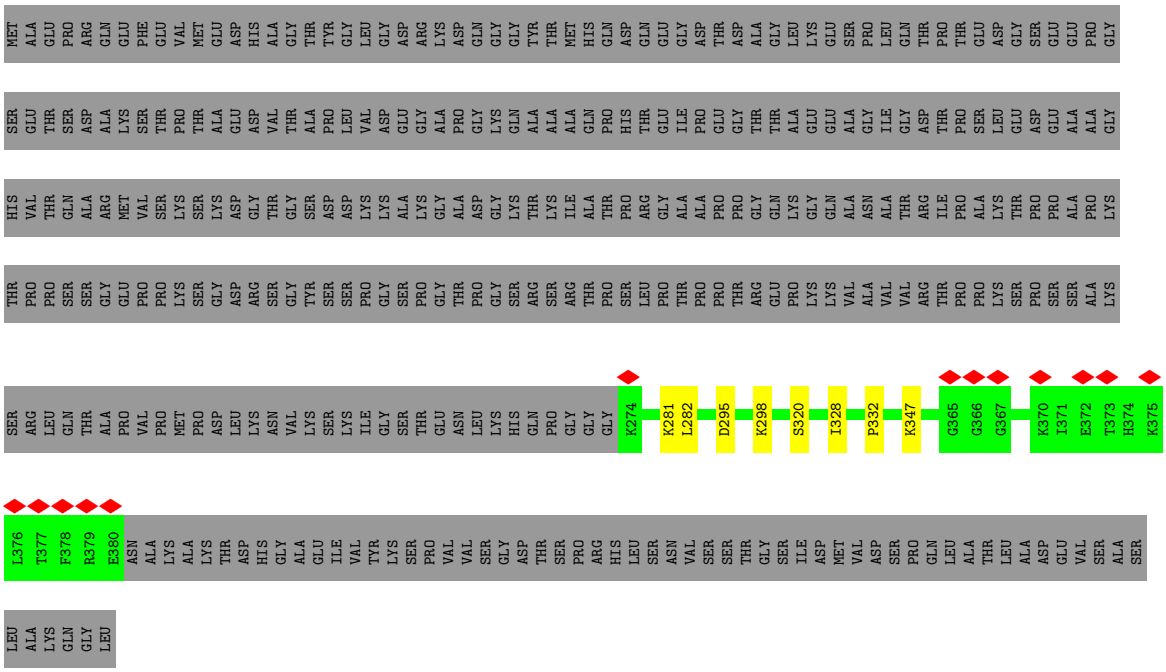
HIS VAL THR GLN ALA ARG MET VAL SER LYS SER LYS ASP GLY THR GLY SER ASP ASP LYS LYS ALA LYS GLY ALA ASP GLY LYS THR LYS ILE ALA THR PRO ARG GLY ALA ALA PRO PRO GLY GLN LYS GLY GLN ALA ASN ALA THR ARG ILE PRO ALA LYS THR PRO PRO ALA PRO LYS

THR PRO PRO SER SER GLY GLU PRO PRO 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



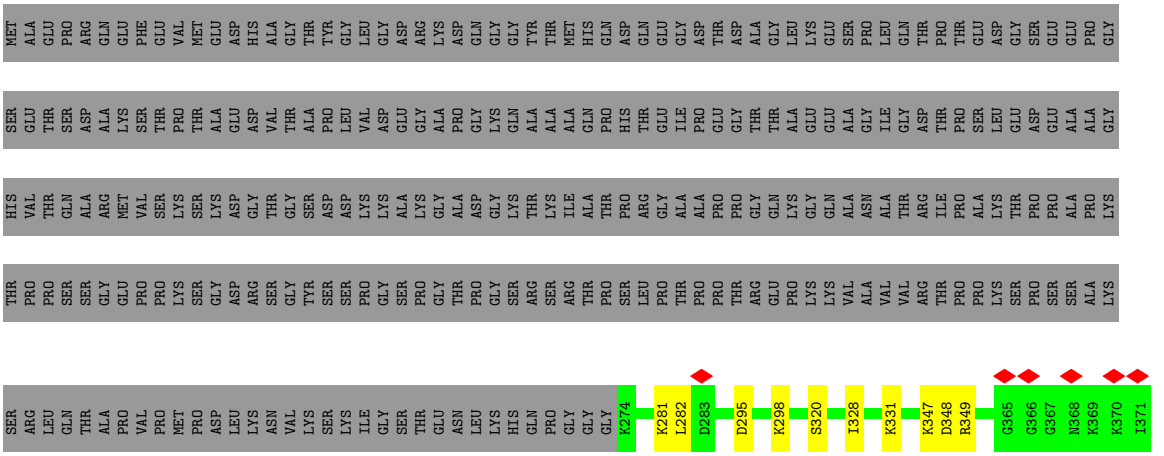
• Molecule 1: Microtubule-associated protein tau

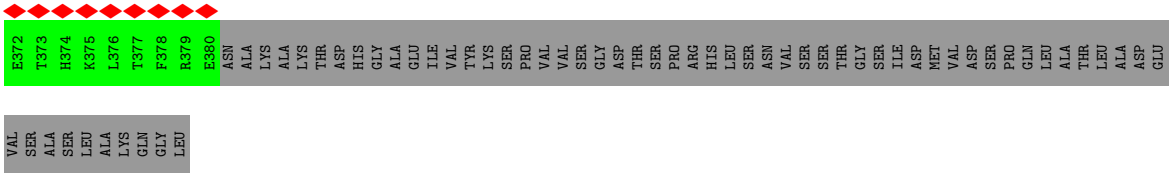
Chain C:



• Molecule 1: Microtubule-associated protein tau

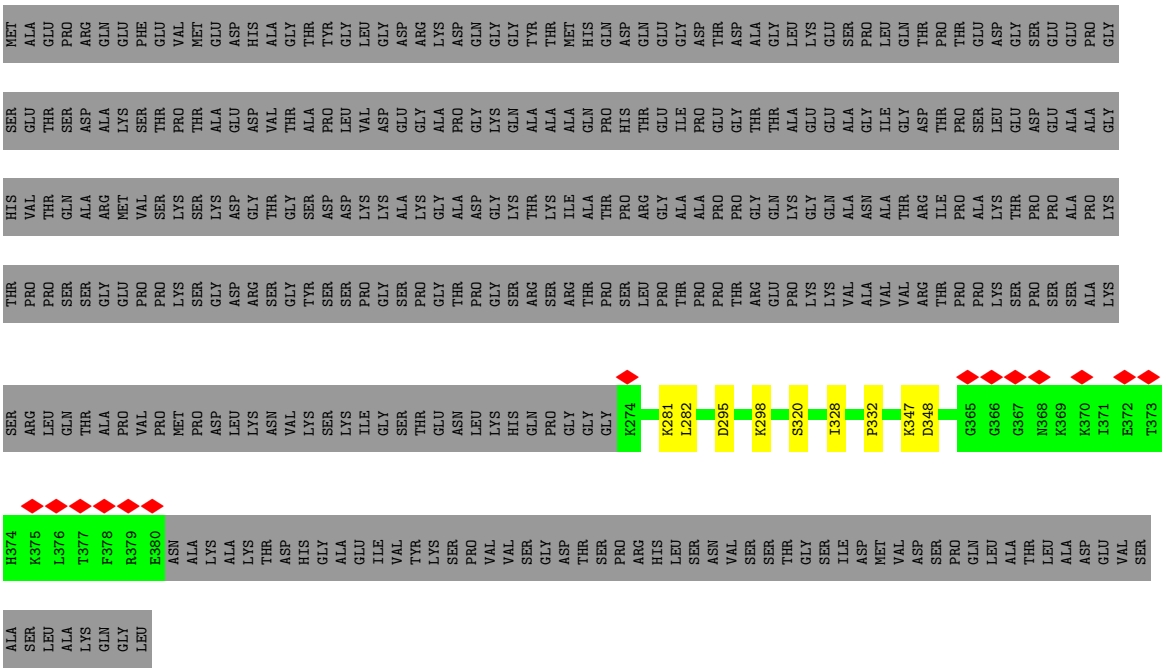
Chain D:





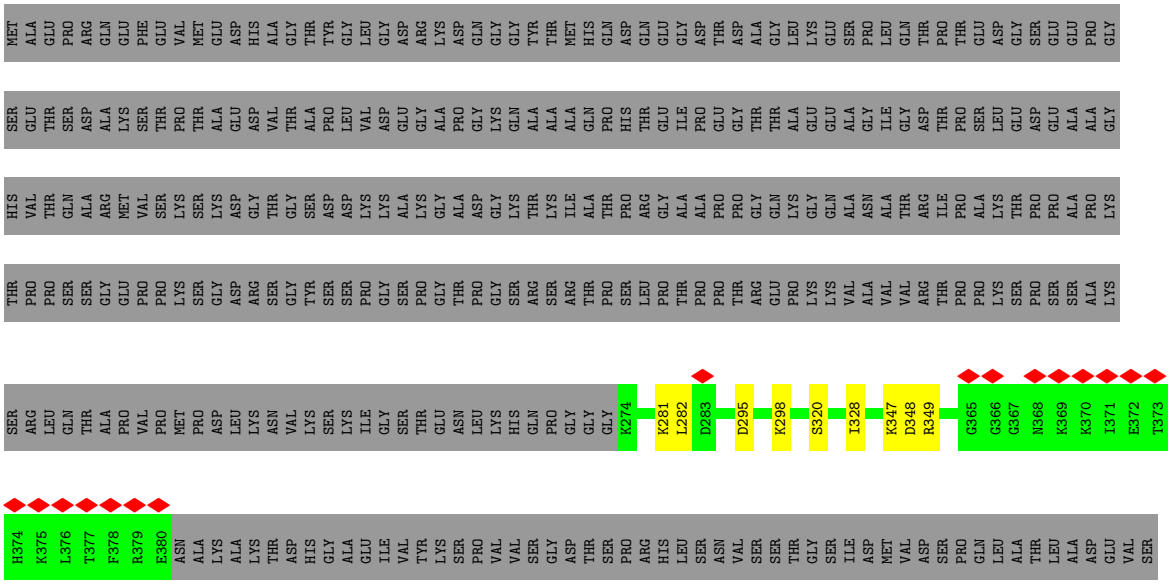
• Molecule 1: Microtubule-associated protein tau

Chain E:



• Molecule 1: Microtubule-associated protein tau

Chain F:



ALA
SER
LEU
ALA
LYS
GLN
GLY
LEU

4 Experimental information (i)

Property	Value	Source
EM reconstruction method	HELICAL	Depositor
Imposed symmetry	HELICAL, twist=-0.61°, rise=4.786 Å, axial sym=C2	Depositor
Number of segments used	20752	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION; Aligned, non-dose-weighted micrographs were used to estimate the contrast transfer function (CTF) using CTFFIND4.1	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/Å^2$)	1.346	Depositor
Minimum defocus (nm)	1700	Depositor
Maximum defocus (nm)	2800	Depositor
Magnification	Not provided	
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor
Maximum map value	0.097	Depositor
Minimum map value	-0.043	Depositor
Average map value	0.001	Depositor
Map value standard deviation	0.004	Depositor
Recommended contour level	0.0166	Depositor
Map size (Å)	379.5, 379.5, 379.5	wwPDB
Map dimensions	330, 330, 330	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.15, 1.15, 1.15	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	0.50	0/824	0.65	0/1101
1	B	0.50	0/824	0.65	0/1101
1	C	0.50	0/824	0.65	0/1101
1	D	0.50	0/824	0.65	0/1101
1	E	0.49	0/824	0.65	0/1101
1	F	0.50	0/824	0.65	0/1101
All	All	0.50	0/4944	0.65	0/6606

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	813	0	851	11	0
1	B	813	0	851	12	0
1	C	813	0	851	7	0
1	D	813	0	851	8	0
1	E	813	0	851	8	0
1	F	813	0	851	9	0
All	All	4878	0	5106	41	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 4.

All (41) 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:295:ASP:OD1	1:B:295:ASP:N	2.39	0.54
1:D:295:ASP:OD1	1:D:295:ASP:N	2.39	0.54
1:F:295:ASP:OD1	1:F:295:ASP:N	2.39	0.54
1:E:295:ASP:OD1	1:E:295:ASP:N	2.39	0.52
1:A:295:ASP:OD1	1:A:295:ASP:N	2.39	0.51
1:C:295:ASP:OD1	1:C:295:ASP:N	2.39	0.51
1:A:298:LYS:HE3	1:C:298:LYS:HG2	1.94	0.50
1:A:298:LYS:HG2	1:E:298:LYS:HE3	1.96	0.48
1:B:347:LYS:HG2	1:D:347:LYS:HD2	1.96	0.47
1:B:298:LYS:HE3	1:D:298:LYS:HG2	1.96	0.47
1:B:298:LYS:HG2	1:F:298:LYS:HE3	1.97	0.47
1:A:348:ASP:OD2	1:B:347:LYS:NZ	2.31	0.45
1:E:320:SER:HB2	1:E:328:ILE:HD13	2.00	0.44
1:A:320:SER:HB2	1:A:328:ILE:HD13	2.00	0.44
1:B:320:SER:HB2	1:B:328:ILE:HD13	1.99	0.44
1:D:320:SER:HB2	1:D:328:ILE:HD13	2.00	0.44
1:E:348:ASP:OD2	1:F:347:LYS:NZ	2.32	0.44
1:C:320:SER:HB2	1:C:328:ILE:HD13	1.99	0.43
1:F:320:SER:HB2	1:F:328:ILE:HD13	2.00	0.43
1:A:347:LYS:HG2	1:C:347:LYS:HD2	2.00	0.43
1:A:347:LYS:HD2	1:E:347:LYS:HG2	2.00	0.43
1:C:320:SER:HB2	1:C:328:ILE:HG21	2.01	0.43
1:D:320:SER:HB2	1:D:328:ILE:HG21	2.01	0.43
1:A:320:SER:HB2	1:A:328:ILE:HG21	2.01	0.43
1:E:320:SER:HB2	1:E:328:ILE:HG21	2.01	0.43
1:B:320:SER:HB2	1:B:328:ILE:HG21	2.01	0.43
1:B:347:LYS:HD2	1:F:347:LYS:HG2	2.00	0.43
1:F:320:SER:HB2	1:F:328:ILE:HG21	2.01	0.43
1:F:281:LYS:HG2	1:F:282:LEU:HD12	2.01	0.43
1:B:281:LYS:HG2	1:B:282:LEU:HD12	2.01	0.42
1:D:281:LYS:HG2	1:D:282:LEU:HD12	2.01	0.42
1:A:281:LYS:HG2	1:A:282:LEU:HD12	2.01	0.42
1:E:281:LYS:HG2	1:E:282:LEU:HD12	2.01	0.42
1:C:281:LYS:HG2	1:C:282:LEU:HD12	2.01	0.42
1:B:331:LYS:HG2	1:D:331:LYS:HB2	2.01	0.42
1:F:348:ASP:HB3	1:F:349:ARG:H	1.72	0.41
1:D:348:ASP:HB3	1:D:349:ARG:H	1.72	0.41
1:B:298:LYS:HD3	1:F:298:LYS:HZ2	1.86	0.41
1:B:348:ASP:HB3	1:B:349:ARG:H	1.72	0.41
1:A:332:PRO:HD2	1:C:332:PRO:HA	2.02	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:332:PRO:HA	1:E:332:PRO:HD2	2.03	0.40

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	105/441 (24%)	100 (95%)	5 (5%)	0	100	100
1	B	105/441 (24%)	100 (95%)	5 (5%)	0	100	100
1	C	105/441 (24%)	100 (95%)	5 (5%)	0	100	100
1	D	105/441 (24%)	100 (95%)	5 (5%)	0	100	100
1	E	105/441 (24%)	100 (95%)	5 (5%)	0	100	100
1	F	105/441 (24%)	100 (95%)	5 (5%)	0	100	100
All	All	630/2646 (24%)	600 (95%)	30 (5%)	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	94/358 (26%)	94 (100%)	0	100	100
1	B	94/358 (26%)	94 (100%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	C	94/358 (26%)	94 (100%)	0	100	100
1	D	94/358 (26%)	94 (100%)	0	100	100
1	E	94/358 (26%)	94 (100%)	0	100	100
1	F	94/358 (26%)	94 (100%)	0	100	100
All	All	564/2148 (26%)	564 (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. All (13) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	299	HIS
1	A	362	HIS
1	B	299	HIS
1	B	329	HIS
1	B	362	HIS
1	C	299	HIS
1	C	362	HIS
1	D	299	HIS
1	D	362	HIS
1	E	299	HIS
1	E	362	HIS
1	F	299	HIS
1	F	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)

There are no chain breaks in this entry.

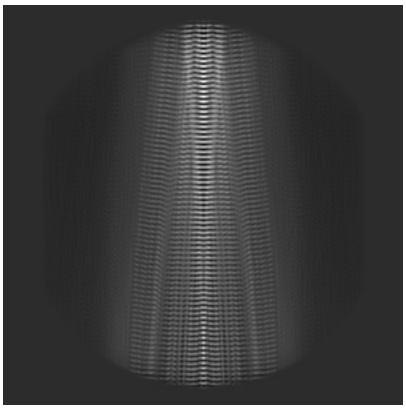
6 Map visualisation (i)

This section contains visualisations of the EMDB entry EMD-10514. These allow visual inspection of the internal detail of the map and identification of artifacts.

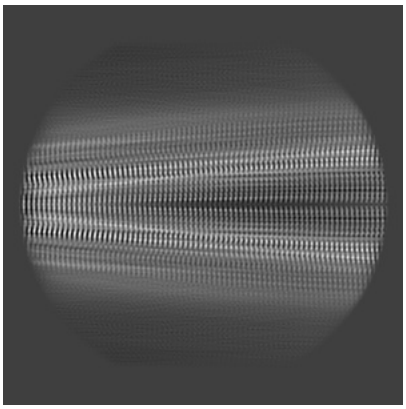
Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

6.1 Orthogonal projections (i)

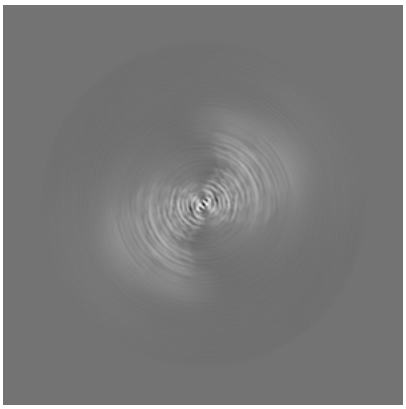
6.1.1 Primary map



X

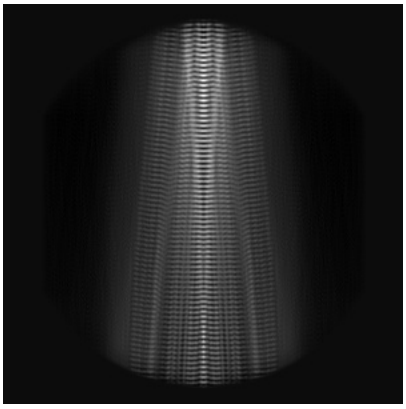


Y

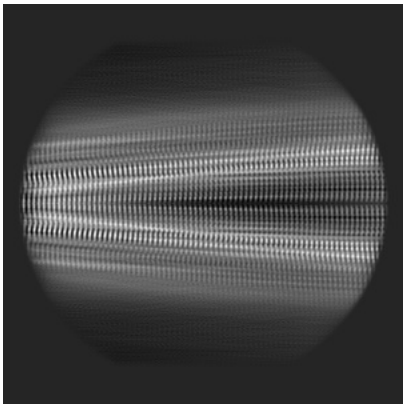


Z

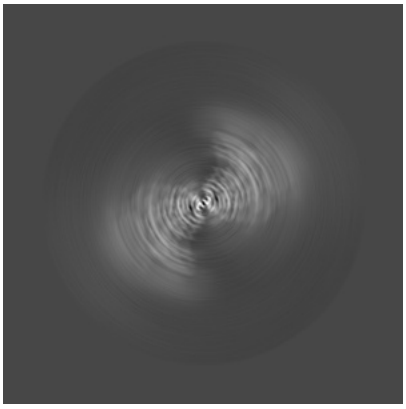
6.1.2 Raw map



X



Y

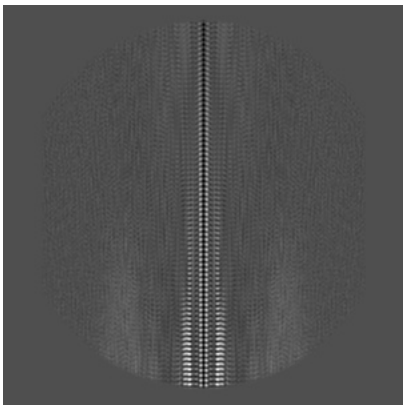


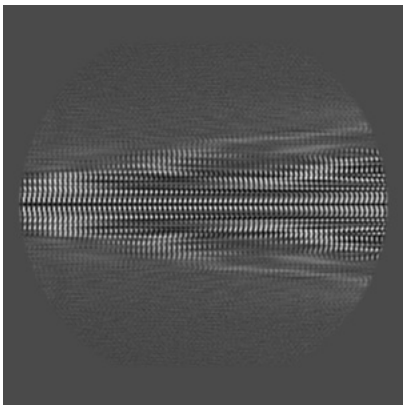
Z

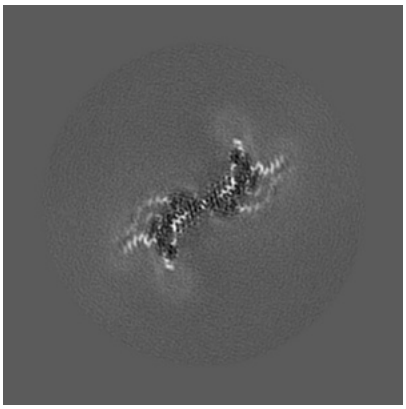
The images above show the map projected in three orthogonal directions.

6.2 Central slices (i)

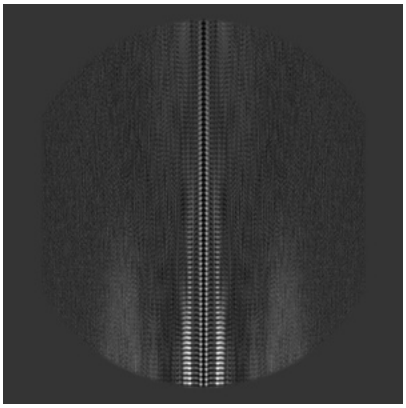
6.2.1 Primary map

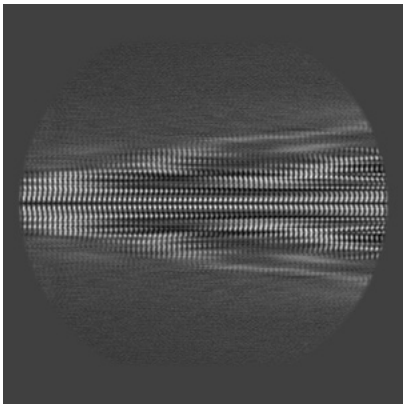

X Index: 165

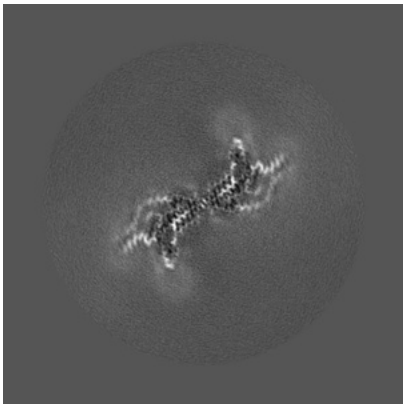

Y Index: 165


Z Index: 165

6.2.2 Raw map


X Index: 165

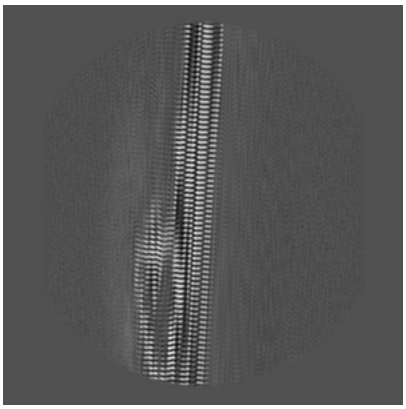

Y Index: 165

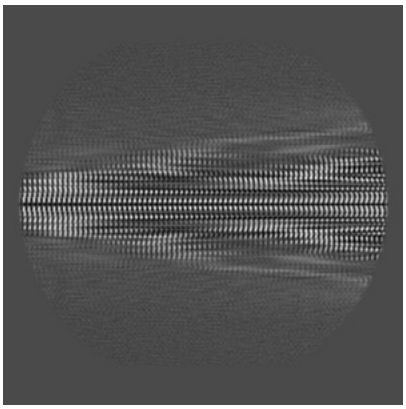

Z Index: 165

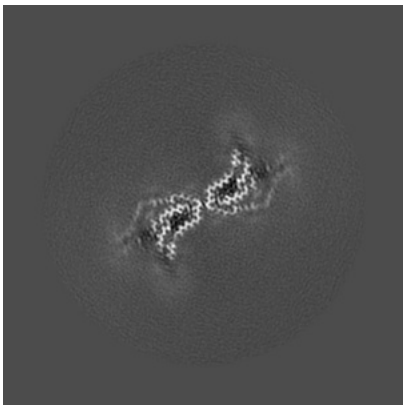
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

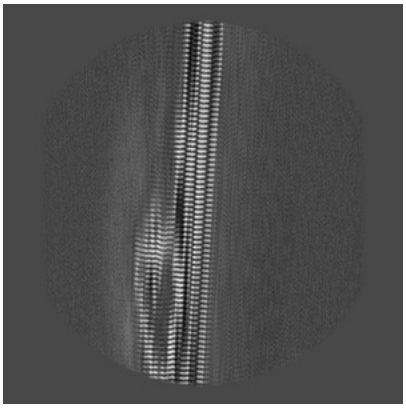
6.3.1 Primary map

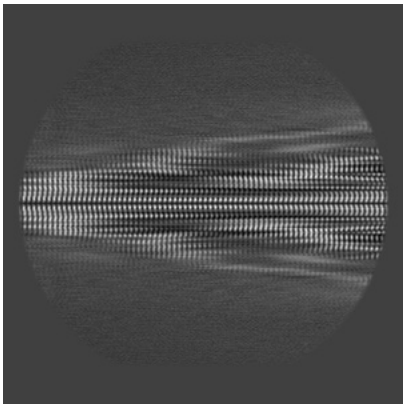

X Index: 144

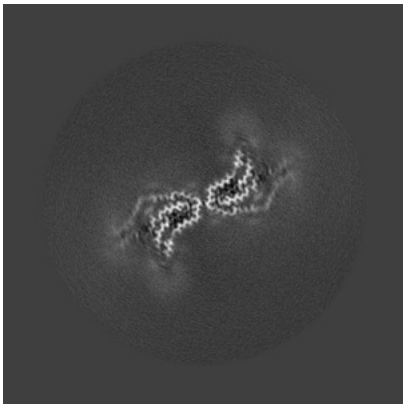

Y Index: 165


Z Index: 175

6.3.2 Raw map


X Index: 144

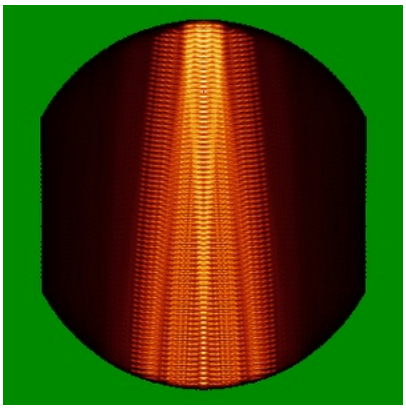

Y Index: 165

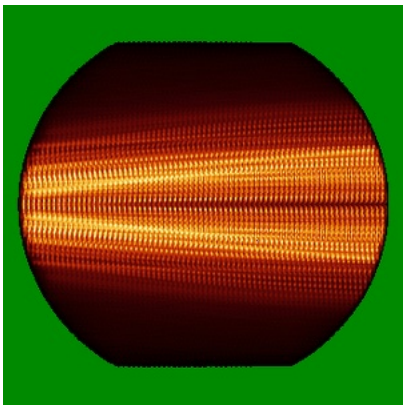

Z Index: 200

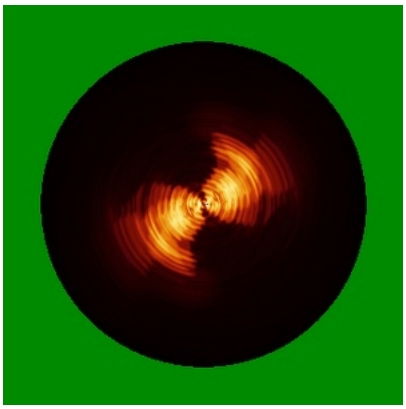
The images above show the largest variance slices of the map in three orthogonal directions.

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

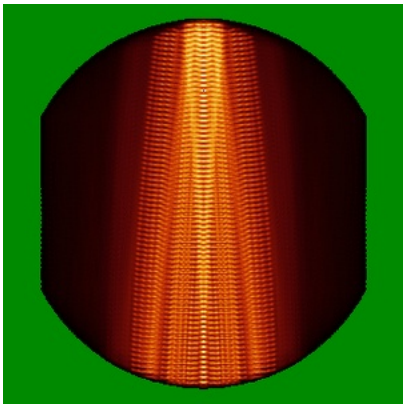
6.4.1 Primary map

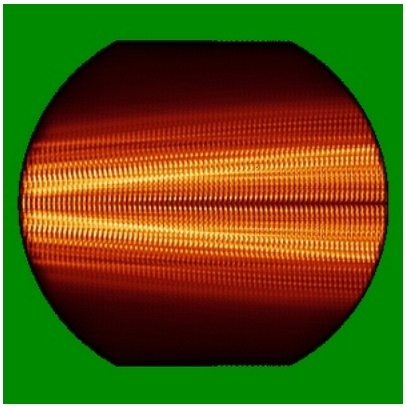

X

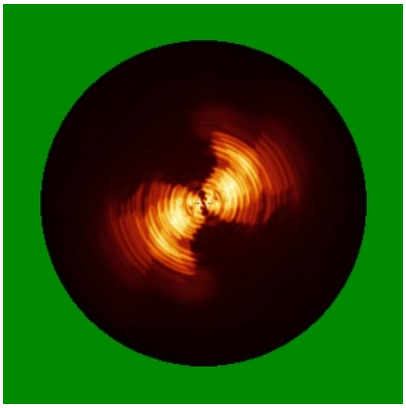

Y


Z

6.4.2 Raw map


X

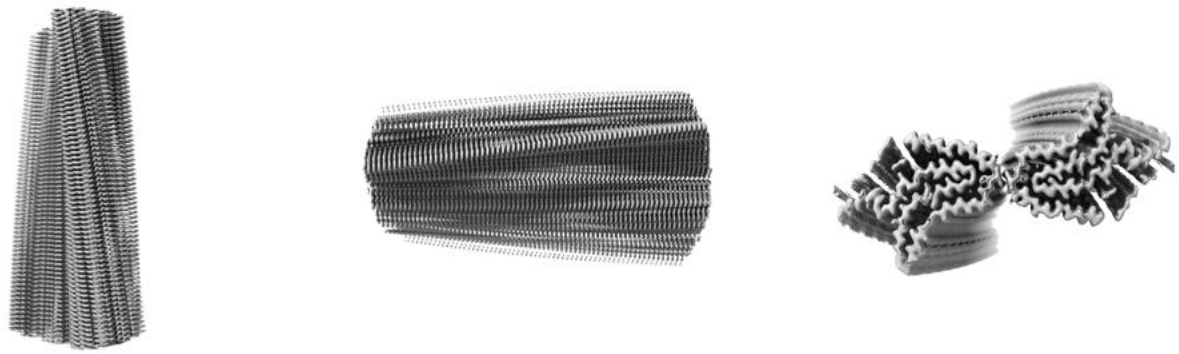

Y


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)

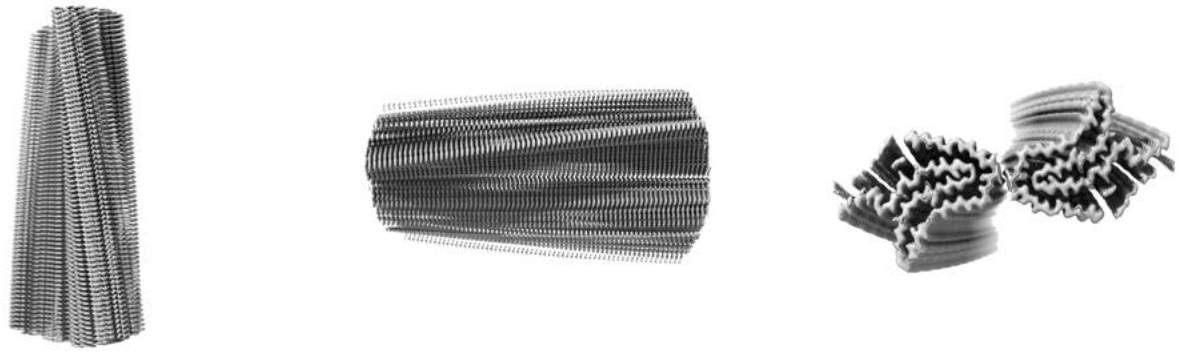
6.5.1 Primary map



X Y Z

The images above show the 3D surface view of the map at the recommended contour level 0.0166. 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.

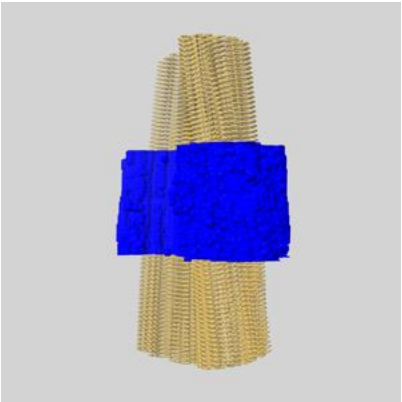
6.6 Mask visualisation (i)

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

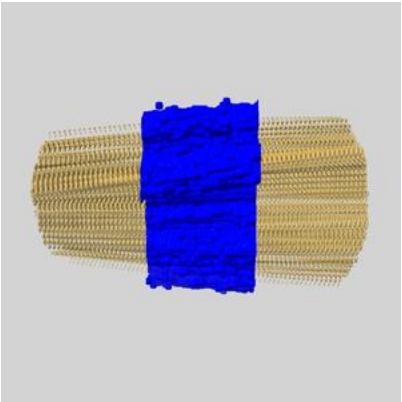
A mask typically either:

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

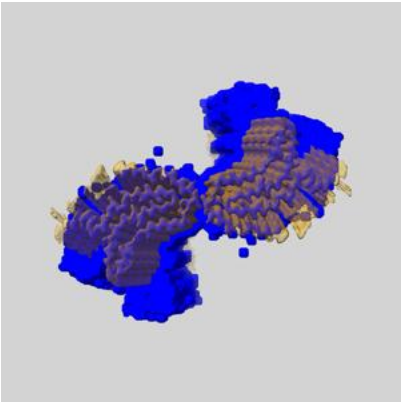
6.6.1 emd_10514_msk_1.map (i)



X



Y

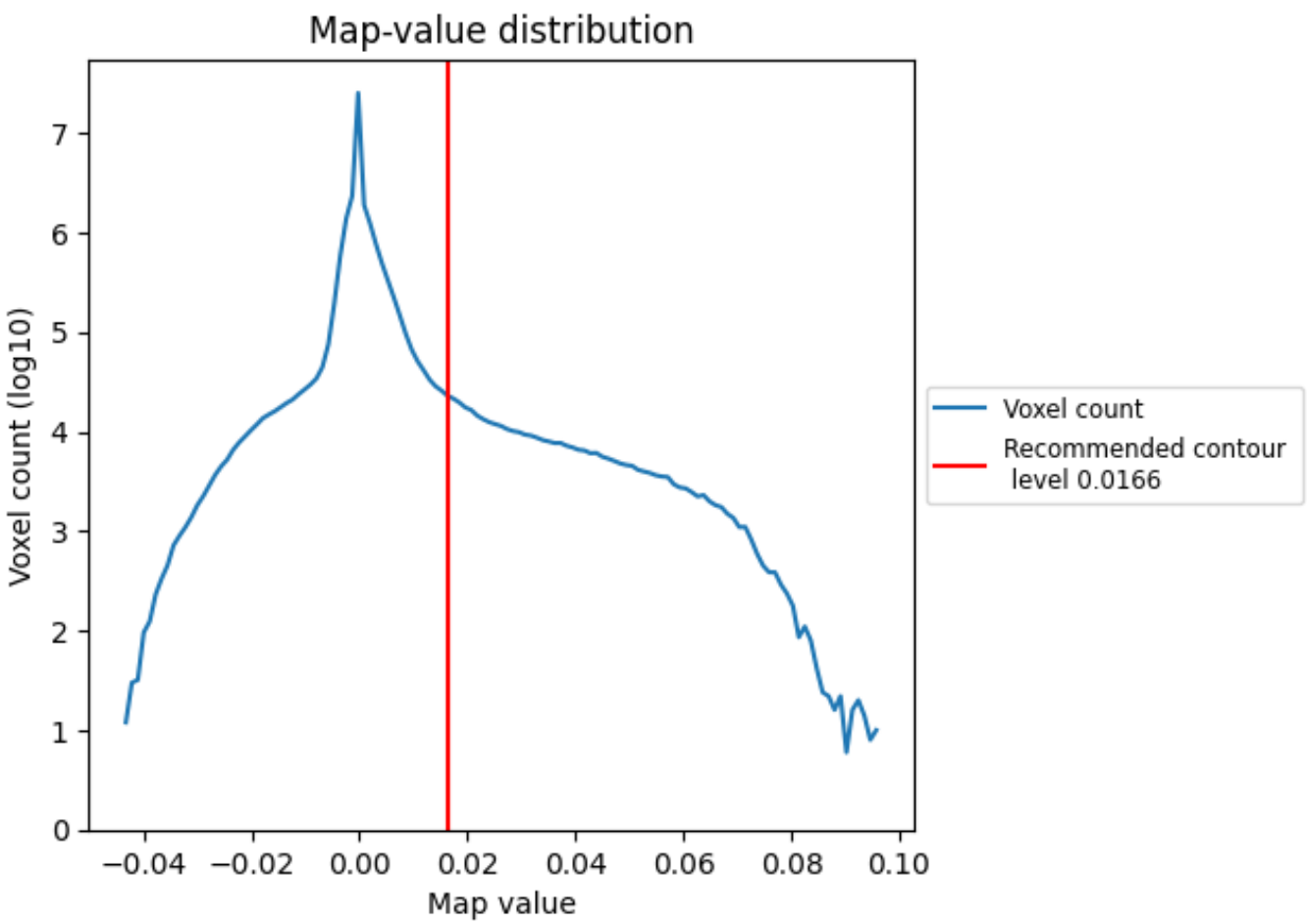


Z

7 Map analysis (i)

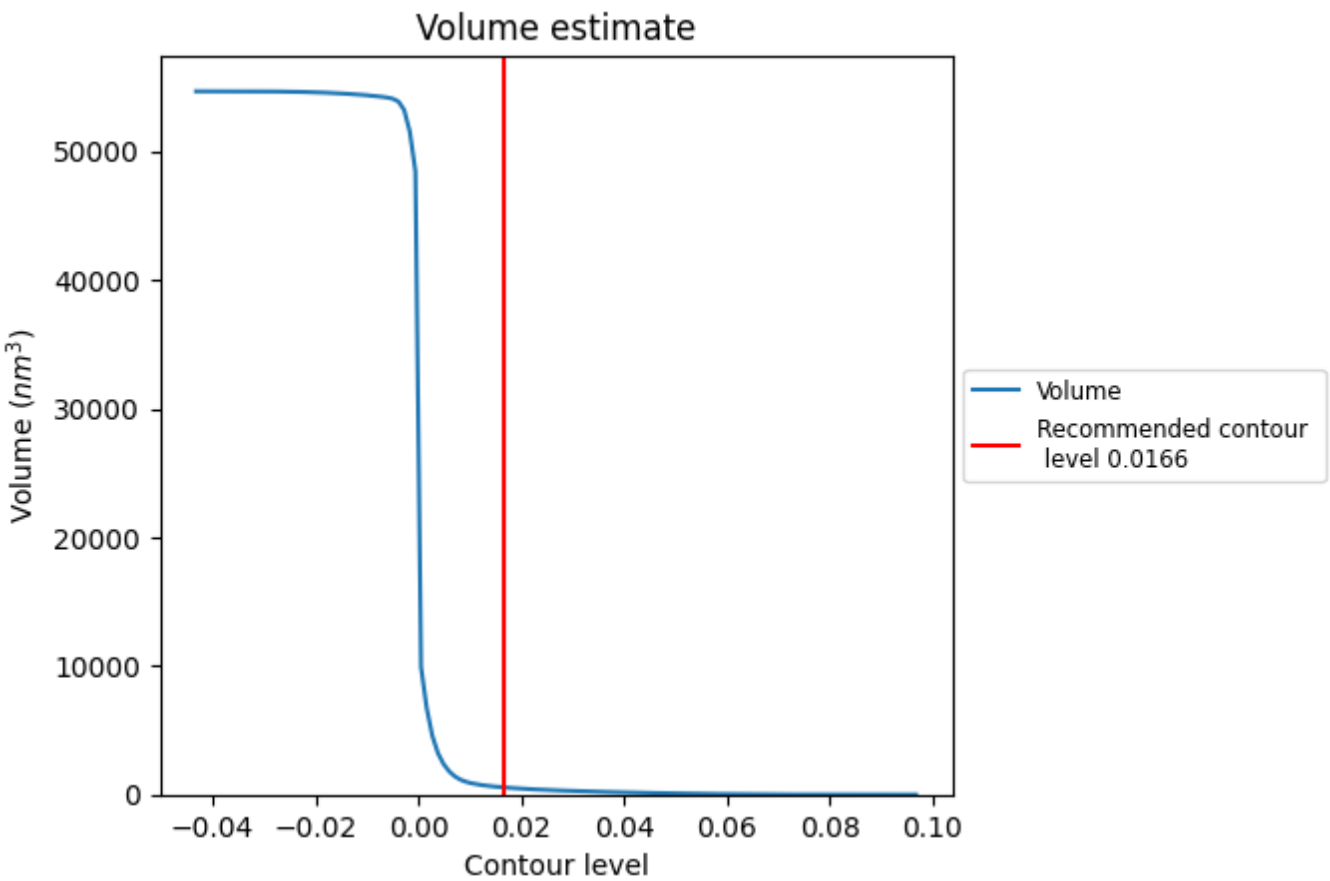
This section contains the results of statistical analysis of the map.

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.

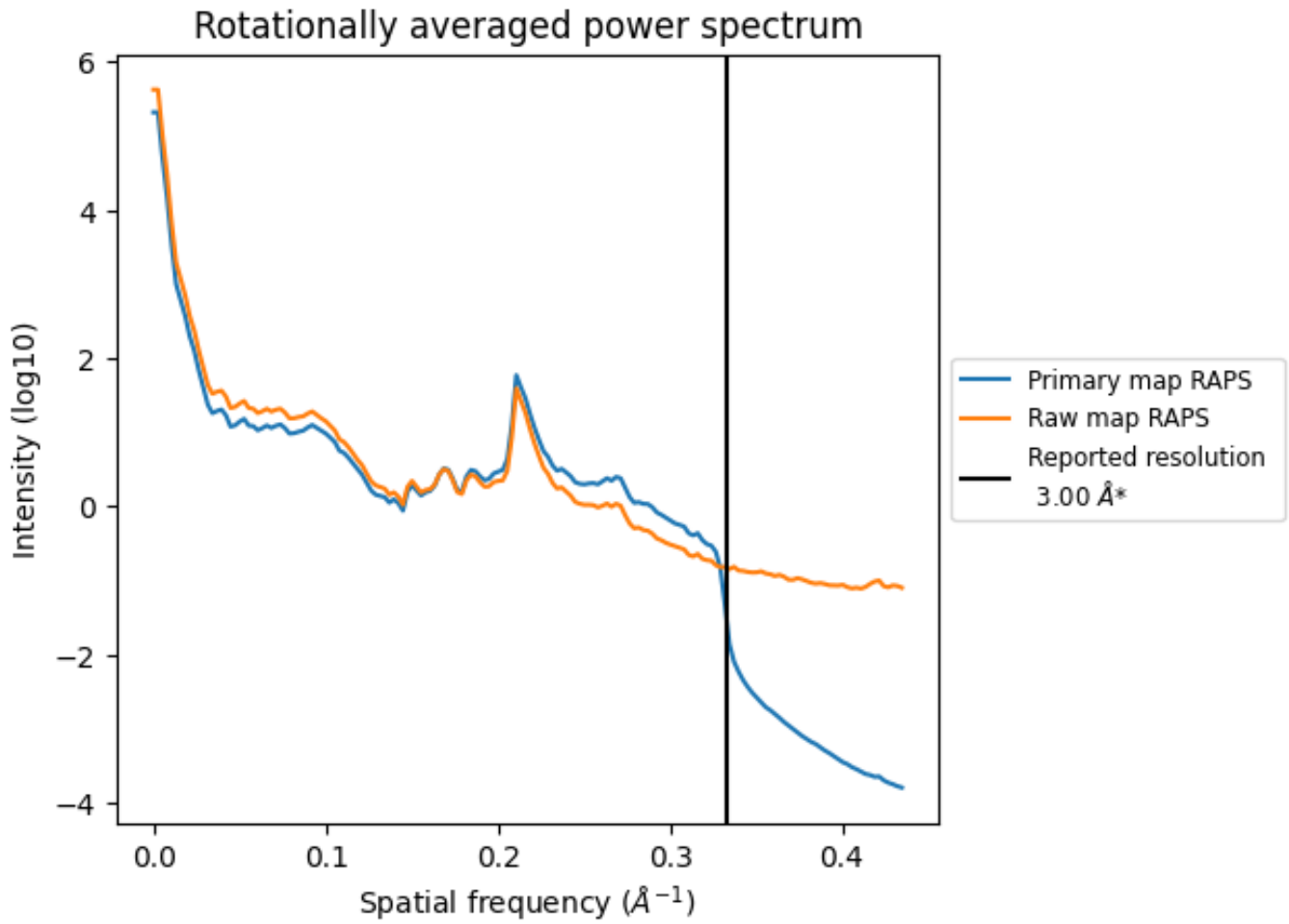
7.2 Volume estimate (i)



The volume at the recommended contour level is 558 nm^3; this corresponds to an approximate mass of 504 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)

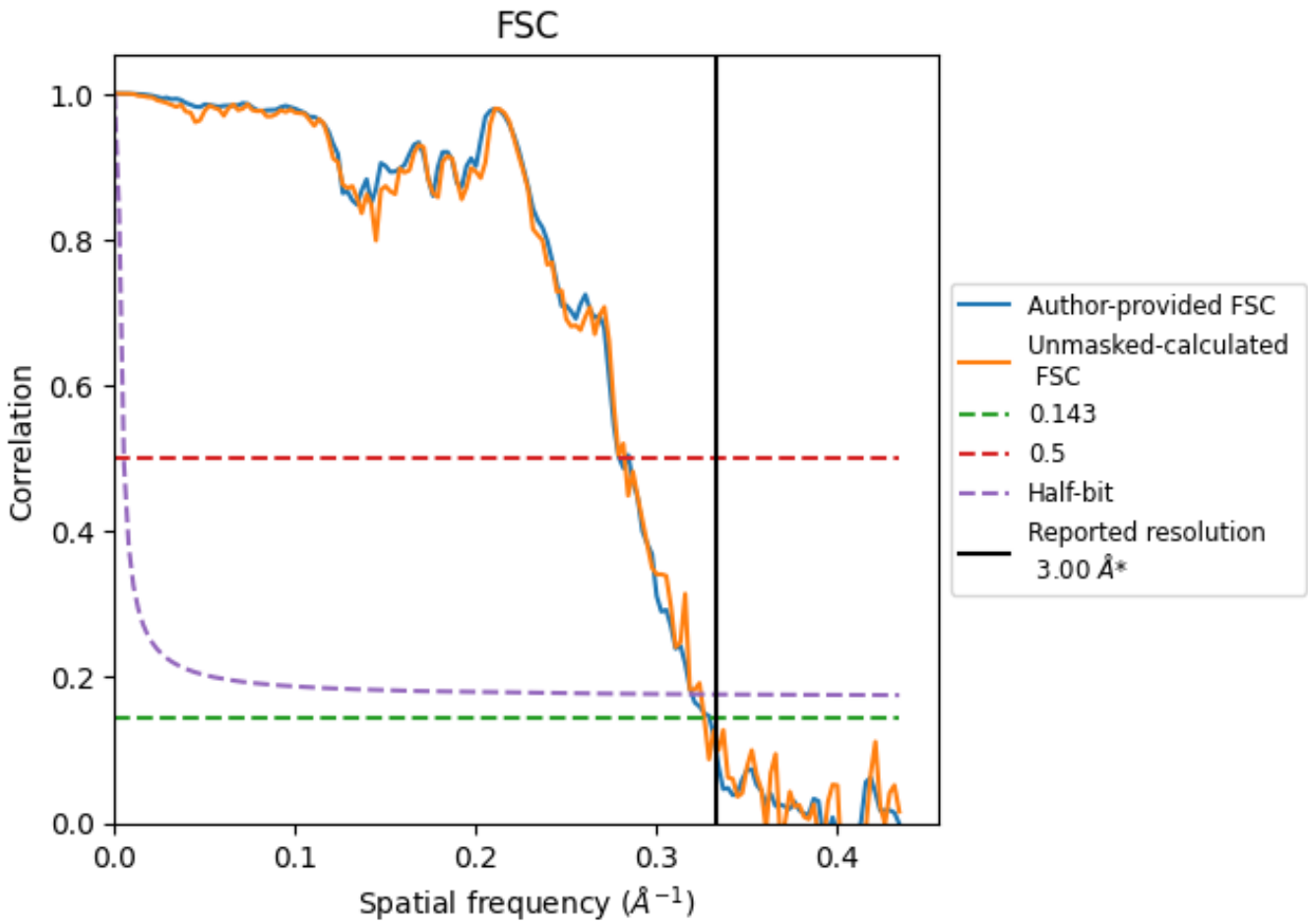


*Reported resolution corresponds to spatial frequency of 0.333 Å^{-1}

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 0.333 Å^{-1}

8.2 Resolution estimates (i)

Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	3.00	-	-
Author-provided FSC curve	3.03	3.58	3.13
Unmasked-calculated*	3.06	3.54	3.08

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

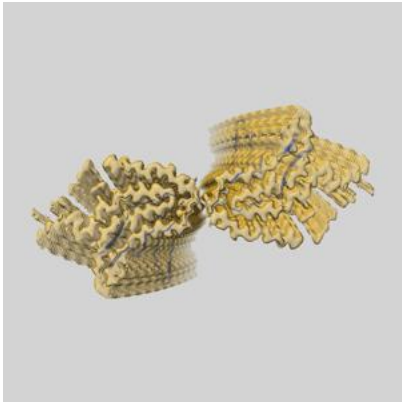
9 Map-model fit (i)

This section contains information regarding the fit between EMDB map EMD-10514 and PDB model 6TJX. Per-residue inclusion information can be found in section 3 on page 4.

9.1 Map-model overlay (i)

X

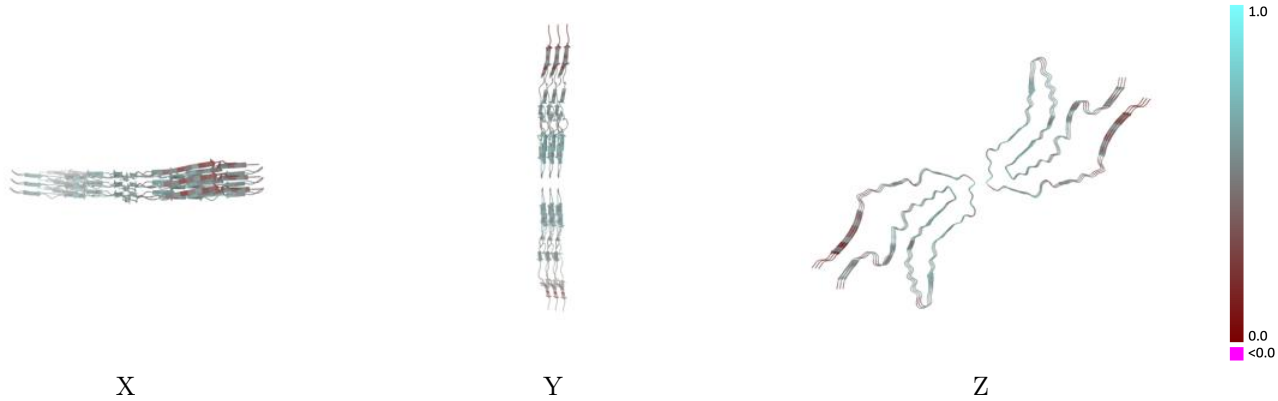
Y



Z

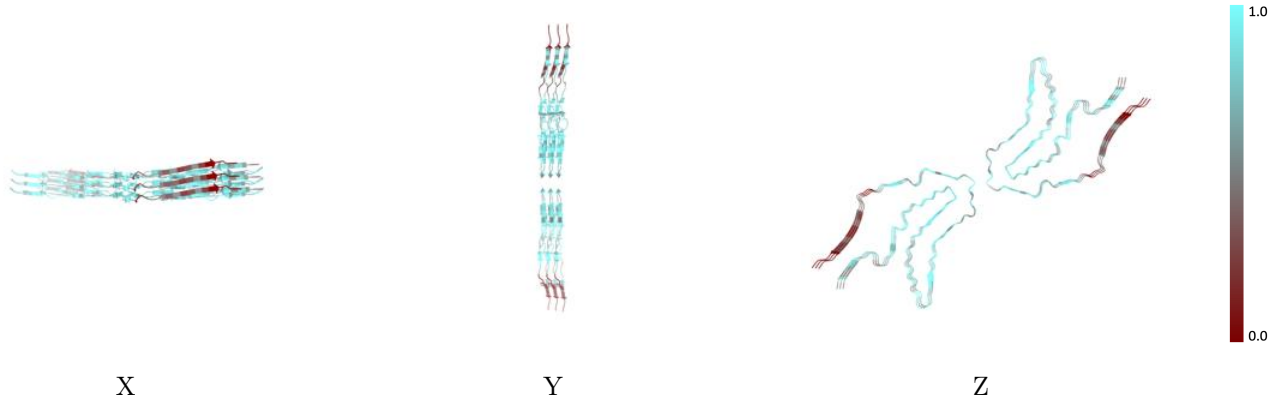
The images above show the 3D surface view of the map at the recommended contour level 0.0166 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

9.2 Q-score mapped to coordinate model (i)



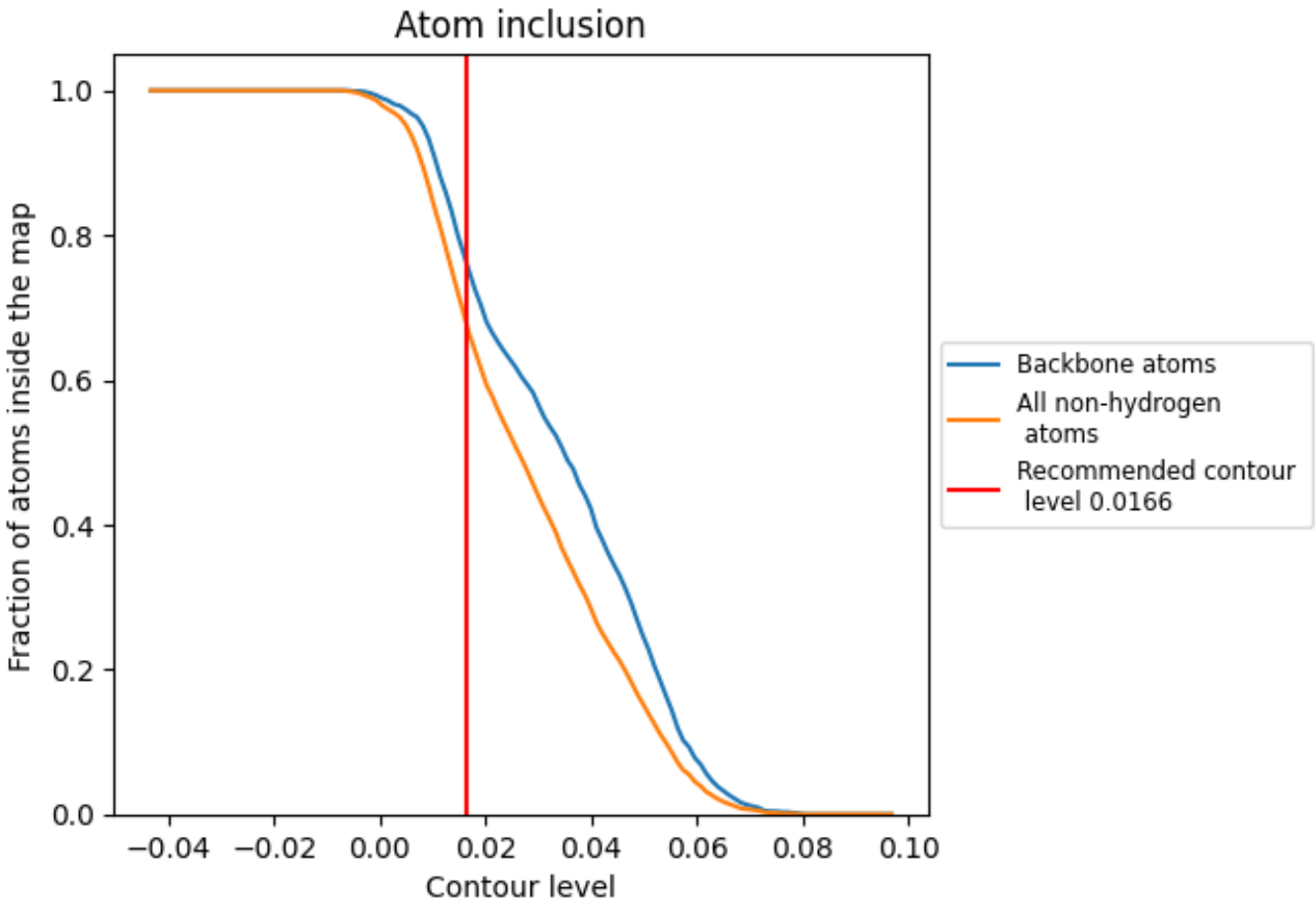
The images above show the model with each residue coloured according its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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 (0.0166).

9.4 Atom inclusion (i)



At the recommended contour level, 76% of all backbone atoms, 67% 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 (0.0166) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.6740	0.5050
A	0.6760	0.5060
B	0.6740	0.5080
C	0.6730	0.5040
D	0.6740	0.5050
E	0.6740	0.5070
F	0.6710	0.5030