



# wwPDB EM Validation Summary Report ⓘ

Sep 28, 2024 – 09:00 PM EDT

PDB ID : 8TXC  
EMDB ID : EMD-41680  
Title : Characterization of the Chlamydomonas Flagellar Mastigoneme Filament Subunit MST1 Structure at 3.9 angstrom  
Authors : Yue, W.; Kai, Z.  
Deposited on : 2023-08-23  
Resolution : 3.90 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](mailto: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.dev113  
MolProbity : 4.02b-467  
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)  
MapQ : 1.9.13  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.39

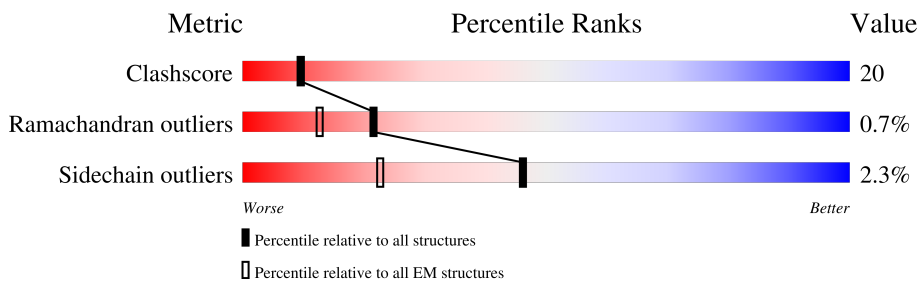
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 3.90 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	210492	15764
Ramachandran outliers	207382	16835
Sidechain outliers	206894	16415

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  $\geq 3$ , 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	1987	

## 2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 13687 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 Mastigoneme-like protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	1894	13687	8643	2234	2727	83	0	0

There are 48 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	141	LEU	VAL	conflict	UNP Q8LRM7
A	142	LEU	THR	conflict	UNP Q8LRM7
A	143	ALA	GLY	conflict	UNP Q8LRM7
A	144	SER	LEU	conflict	UNP Q8LRM7
A	145	LYS	GLU	conflict	UNP Q8LRM7
A	146	THR	ASP	conflict	UNP Q8LRM7
A	147	VAL	GLY	conflict	UNP Q8LRM7
A	149	ILE	HIS	conflict	UNP Q8LRM7
A	150	TYR	LEU	conflict	UNP Q8LRM7
A	151	VAL	CYS	conflict	UNP Q8LRM7
A	517	ARG	LYS	conflict	UNP Q8LRM7
A	530	GLU	GLY	conflict	UNP Q8LRM7
A	619	THR	ALA	conflict	UNP Q8LRM7
A	800	SER	THR	conflict	UNP Q8LRM7
A	820	SER	PHE	conflict	UNP Q8LRM7
A	?	-	GLY	deletion	UNP Q8LRM7
A	?	-	THR	deletion	UNP Q8LRM7
A	?	-	PRO	deletion	UNP Q8LRM7
A	?	-	GLY	deletion	UNP Q8LRM7
A	?	-	PRO	deletion	UNP Q8LRM7
A	?	-	TYR	deletion	UNP Q8LRM7
A	?	-	PHE	deletion	UNP Q8LRM7
A	?	-	LEU	deletion	UNP Q8LRM7
A	1399	LYS	ARG	conflict	UNP Q8LRM7
A	?	-	PRO	deletion	UNP Q8LRM7
A	?	-	GLU	deletion	UNP Q8LRM7
A	1868	PRO	ALA	conflict	UNP Q8LRM7
A	1897	PRO	GLN	conflict	UNP Q8LRM7

*Continued on next page...*

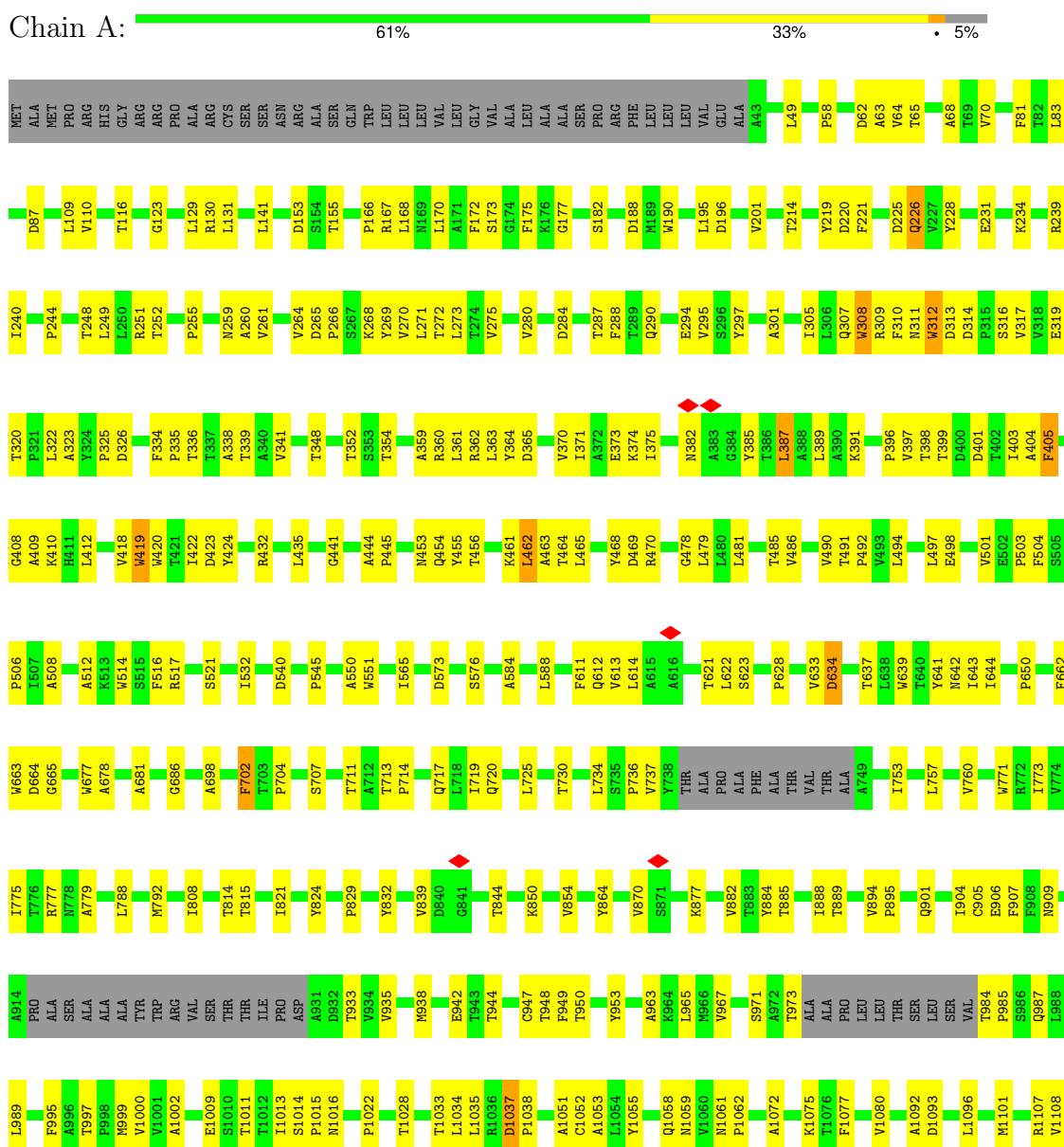
*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	1914	PRO	ARG	conflict	UNP Q8LRM7
A	1915	PRO	ARG	conflict	UNP Q8LRM7
A	1917	PRO	HIS	conflict	UNP Q8LRM7
A	1919	SER	ALA	conflict	UNP Q8LRM7
A	1920	PRO	ARG	conflict	UNP Q8LRM7
A	1921	PRO	ARG	conflict	UNP Q8LRM7
A	1924	ASN	THR	conflict	UNP Q8LRM7
A	1925	ARG	ALA	conflict	UNP Q8LRM7
A	1926	SER	LEU	conflict	UNP Q8LRM7
A	1935	SER	PRO	conflict	UNP Q8LRM7
A	1978	ASP	-	expression tag	UNP Q8LRM7
A	1979	ALA	-	expression tag	UNP Q8LRM7
A	1980	GLU	-	expression tag	UNP Q8LRM7
A	1981	MET	-	expression tag	UNP Q8LRM7
A	1982	GLN	-	expression tag	UNP Q8LRM7
A	1983	PRO	-	expression tag	UNP Q8LRM7
A	1984	GLN	-	expression tag	UNP Q8LRM7
A	1985	ASP	-	expression tag	UNP Q8LRM7
A	1986	ASP	-	expression tag	UNP Q8LRM7
A	1987	GLU	-	expression tag	UNP Q8LRM7

### 3 Residue-property plots

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: Mastigoneme-like protein



G1833	T1721	G1616	G1541	G1488	C1392	S1306	A1221	S1109
S1834	Y1722	S1617	K1542	T1469	E1393	G1307	T1222	C1110
C1837	S1723	F1618	S1543	F1470	D1394	F1308	R1227	F1111
K1838	R1724	A1619	E1544	A1471	G1395	V1309	C1228	F1112
P1839	C1733	S1620	T1548	P1474	T1396	S1310	I1229	P1124
C1840	G1736	L1624	N1549	G1475	W1397	A1314	P1230	K1126
S1841	T1737	C1627	L1550	A1476	S1398	T1315	G1231	F1126
L1842	V1738	A1630	Y1553	C1479	K1399	G1316	P1232	S1134
S1844	A1739	A1630	Y1554	Q1480	A1400	C1317	Y1235	V1143
K1845	S1740	G1633	T1555	Q1481	C1405	T1318	W1236	V1143
P1846	K1741	G1633	Q1563	C1482	C1408	C1320	S1237	A1144
K1847	S1744	F1635	L1564	K1483	P1409	E1322	G1240	D1145
T1848	G1753	T1636	P1565	P1484	P1410	E1322	G1240	T1148
R1871	G1753	F1636	P1565	G1485	P1410	G1323	Y1244	S1149
P1877	C1768	A1642	P1568	N1487	G1411	G1324	C1244	L1150
P1878	P1769	M1643	G1569	S1488	T1412	Y1325	C1247	P1169
R1881	Y1772	F1646	T1570	L1489	Y1413	H1326	Q1250	P1175
P1920	Y1773	M1647	F1571	M1490	R1414	G1329	T1251	K1176
P1921	G1774	M1647	G1491	G1491	M1415	T1332	I1252	G1177
S1926	P1775	L1651	D1492	R1493	T1416	T1333	A1253	F1178
P1927	G1778	I1652	T1576	T1494	Q1421	T1334	A1256	R1179
S1934	F1788	L1656	M1580	Q1495	L1422	P1335	A1257	C1182
S1935	R1792	F1670	L1581	Q1496	P1425	T1343	A1257	G1185
P1939	G1793	F1671	L1582	M1497	F1426	T1344	L1258	E1187
Q1945	K1796	Y1670	L1586	F1508	A1435	Y1351	I1189	E1187
ASN	G1797	F1671	V1588	P1509	T1436	P1355	P1268	L1188
GLY	K1797	Q1672	D1589	A1510	E1440	M1356	T1269	L1188
ASP	P1798	S1673	G1590	L1511	E1440	T1357	T1270	I1190
PRO	Y1799	A1675	T1591	R1512	E1440	C1358	V1274	S1191
VAL	D1800	E1676	G1592	A1513	E1440	R1359	T1275	M1192
GLY	F1801	T1677	C1593	Y1514	E1440	Q1360	H1276	D1193
HIS	F1802	T1678	Y1594	T1515	E1440	C1361	L1277	N1194
ARG	F1805	C1680	T1594	T1516	E1440	P1362	A1278	Y1195
ARG	P1806	T1681	Y1598	G1518	E1440	M1364	S1283	R1200
ALA	G1807	A1682	G1599	V1520	E1440	T1365	V1201	V1201
L1956	V1807	C1683	T1600	Y1525	E1440	Y1366	A1202	A1202
L1957	V1808	G1686	F1601	A1526	E1440	L1367	V1203	V1203
S1958	R1809	G1687	N1602	A1454	E1440	P1368	A1204	A1205
L1959	T1812	T1687	D1603	A1454	E1440	L1369	A1204	A1205
M1960	A1813	Y1688	E1604	P1455	E1440	R1370	Y1291	Y1208
Q1962	A1813	A1689	F1605	F1457	E1440	G1371	F1292	G1211
PRO	L1817	P1699	S1606	G1458	E1440	Q1372	Q1293	Q1212
GLN	L1817	P1699	Q1607	M1458	E1440	A1374	P1294	Q1212
ASP	Q1827	Q1706	P1608	V1461	E1440	S1377	G1298	R1213
ASP	C1828	Q1706	V1609	V1461	E1440	S1377	C1298	C1213
T1829	T1829	K1715	C1610	C1462	E1440	S1385	K1284	R1214
E1830	E1830	K1715	M1536	I1463	E1440	A1386	K1286	V1300
D1831	D1831	M1719	F1537	P1464	E1440	A1386	K1286	C1216
L1832	L1832	G1720	F1538	P1466	E1440	T1387	K1286	P1217
			M1539	P1466	E1440	P1388	K1286	G1218
			S1615	A1467	E1440	C1389	K1286	G1218
					E1440		K1286	C1213
					E1440		K1286	R1214
					E1440		K1286	V1300
					E1440		K1286	C1301
					E1440		K1286	C1216
					E1440		K1286	P1217
					E1440		K1286	G1218
					E1440		K1286	G1219
					E1440		K1286	P1303
					E1440		K1286	C1304
					E1440		K1286	T1220

## 4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	70074	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	NONE	Depositor
Microscope	TFS GLACIOS	Depositor
Voltage (kV)	200	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	40	Depositor
Minimum defocus (nm)	-900	Depositor
Maximum defocus (nm)	2500	Depositor
Magnification	Not provided	
Image detector	GATAN K2 QUANTUM (4k x 4k)	Depositor
Maximum map value	0.786	Depositor
Minimum map value	-0.222	Depositor
Average map value	0.007	Depositor
Map value standard deviation	0.026	Depositor
Recommended contour level	0.0728	Depositor
Map size ( $\text{\AA}$ )	588.288, 588.288, 588.288	wwPDB
Map dimensions	512, 512, 512	wwPDB
Map angles ( $^\circ$ )	90.0, 90.0, 90.0	wwPDB
Pixel spacing ( $\text{\AA}$ )	1.149, 1.149, 1.149	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	0.32	2/14067 (0.0%)	0.58	9/19393 (0.0%)

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	2

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	1509	PRO	CG-CD	-12.95	1.07	1.50
1	A	1509	PRO	CB-CG	5.21	1.75	1.50

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	1509	PRO	N-CD-CG	-11.93	85.31	103.20
1	A	1509	PRO	CA-N-CD	-11.31	95.66	111.50
1	A	1335	PRO	CA-N-CD	-7.93	100.40	111.50
1	A	1509	PRO	CA-CB-CG	-7.65	89.47	104.00
1	A	1169	PRO	CA-N-CD	-7.09	101.57	111.50

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	1334	THR	Peptide
1	A	1527	LYS	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	13687	0	13241	529	0
All	All	13687	0	13241	529	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 20.

The worst 5 of 529 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:1509:PRO:CG	1:A:1509:PRO:CB	1.76	1.36
1:A:1189:ILE:HD11	1:A:1214:ARG:HB2	1.51	0.90
1:A:1227:ARG:HE	1:A:1229:ILE:HD13	1.36	0.90
1:A:1421:GLN:HB2	1:A:1436:THR:HB	1.56	0.88
1:A:1957:LEU:HA	1:A:1960:MET:HE3	1.56	0.87

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	1884/1987 (95%)	1693 (90%)	178 (9%)	13 (1%)	<b>19</b> 54

5 of 13 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1062	PRO
1	A	1077	PHE
1	A	1201	VAL
1	A	1528	PRO
1	A	1799	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	1488/1571 (95%)	1454 (98%)	34 (2%)	45 64

5 of 34 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1722	TYR
1	A	1724	ARG
1	A	1809	ARG
1	A	662	PHE
1	A	634	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	1487	ASN
1	A	1563	GLN
1	A	1580	ASN
1	A	1421	GLN
1	A	1061	ASN

### 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 oligosaccharides 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.

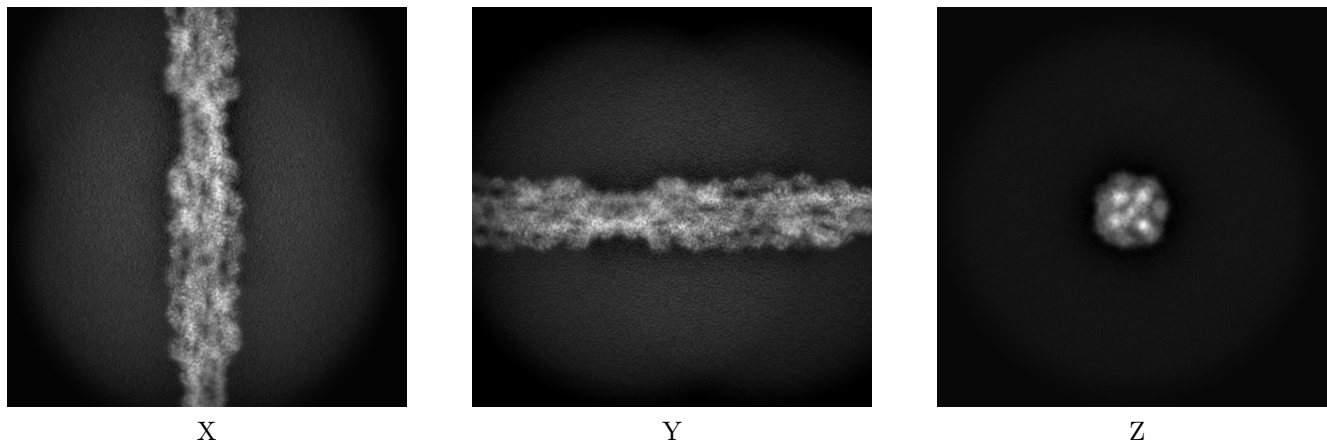
## 6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-41680. These allow visual inspection of the internal detail of the map and identification of artifacts.

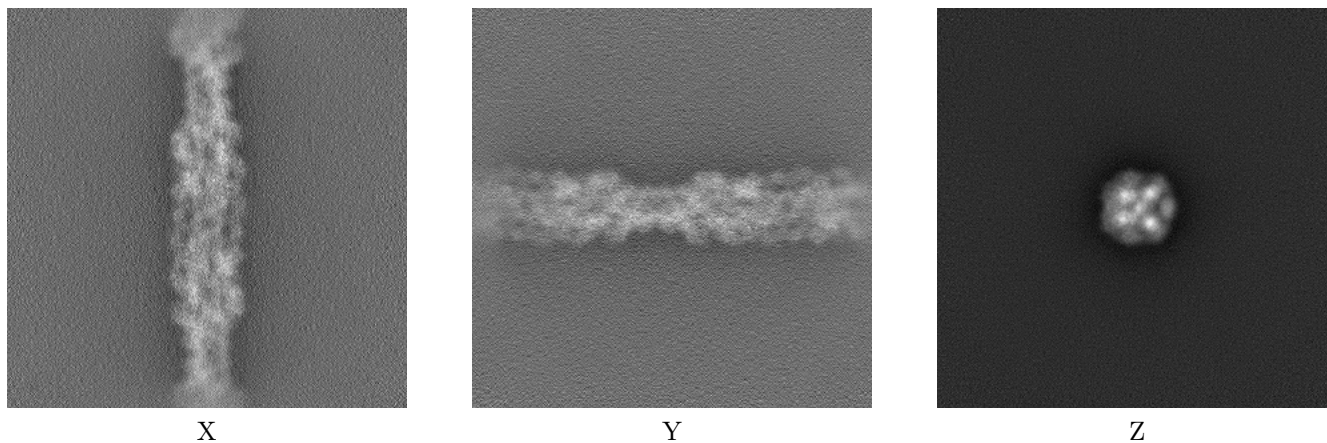
Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

### 6.1 Orthogonal projections [i](#)

#### 6.1.1 Primary map



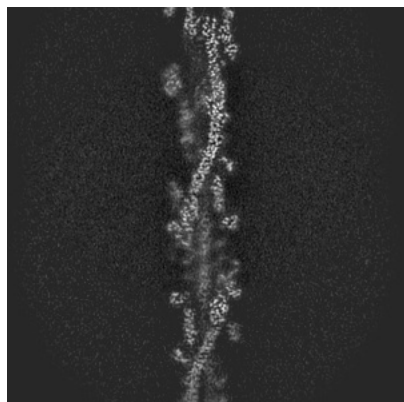
#### 6.1.2 Raw map



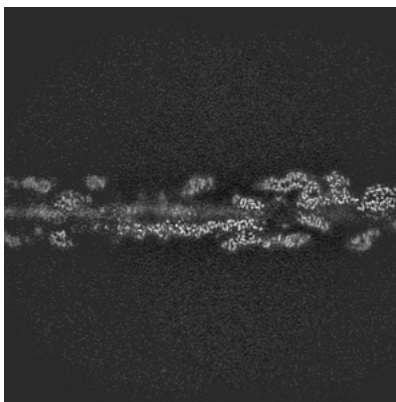
The images above show the map projected in three orthogonal directions.

## 6.2 Central slices [i](#)

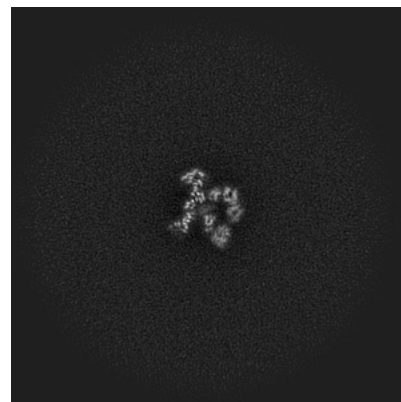
### 6.2.1 Primary map



X Index: 256

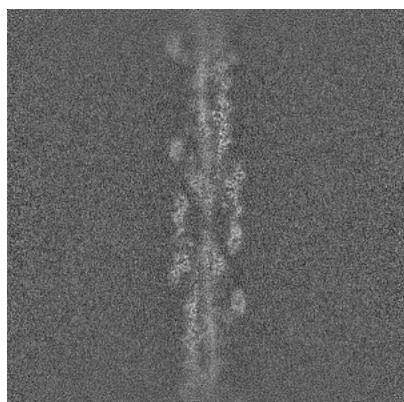


Y Index: 256

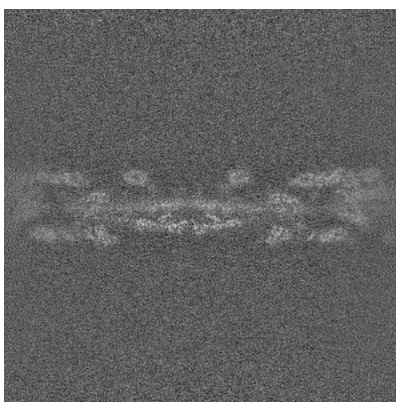


Z Index: 256

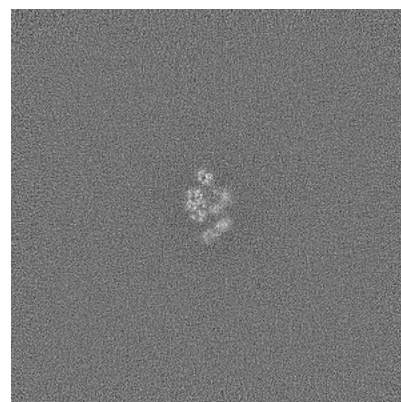
### 6.2.2 Raw map



X Index: 256



Y Index: 256



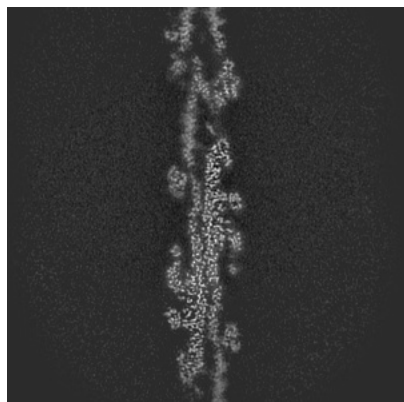
Z Index: 256

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

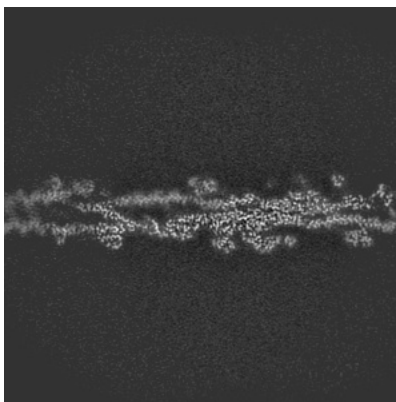


## 6.3 Largest variance slices [i](#)

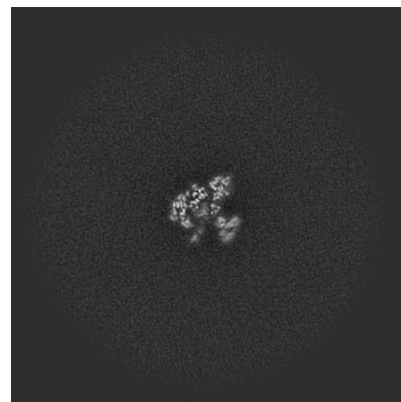
### 6.3.1 Primary map



X Index: 231

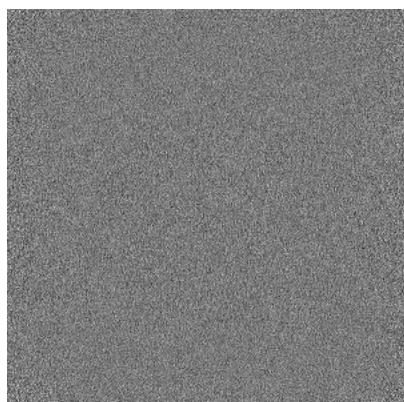


Y Index: 268

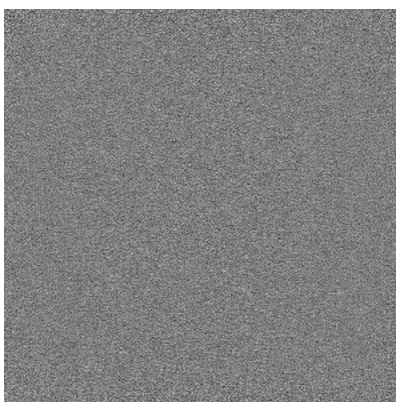


Z Index: 307

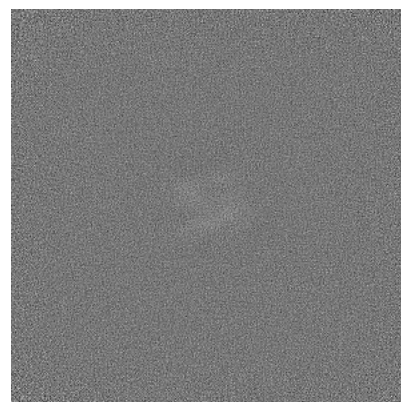
### 6.3.2 Raw map



X Index: 0



Y Index: 0

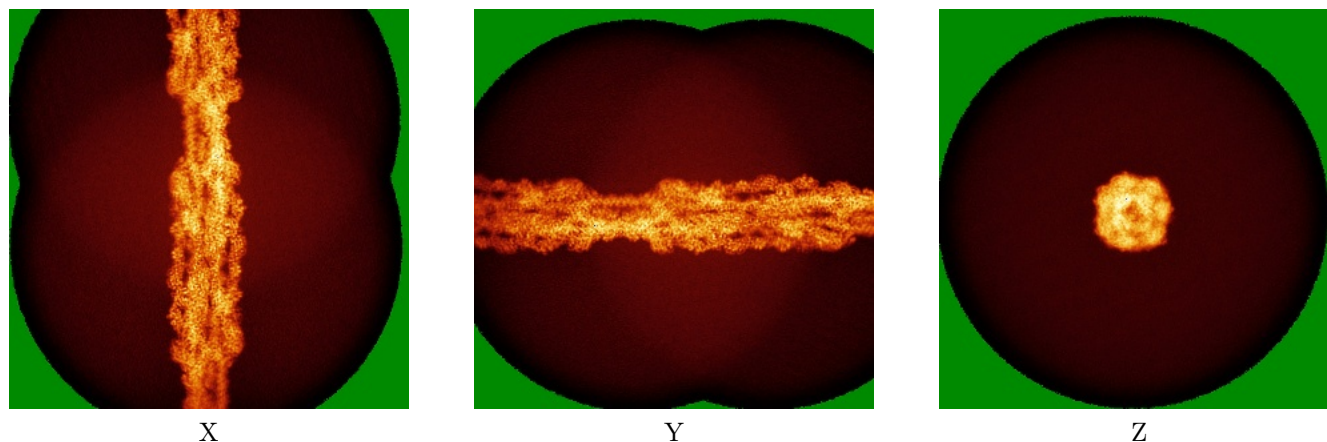


Z Index: 0

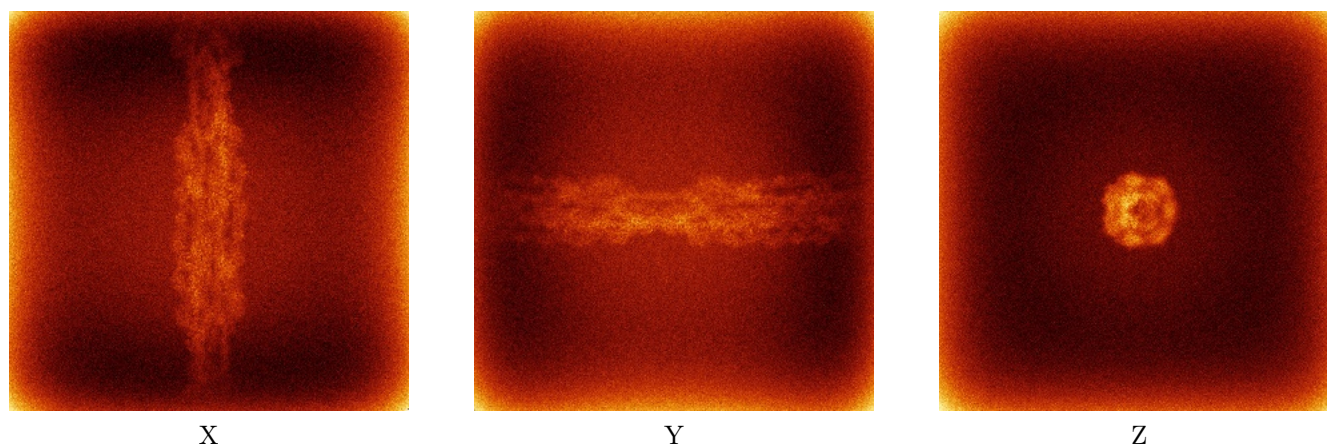
The images above show the largest variance slices of the map in three orthogonal directions.

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

### 6.4.1 Primary map



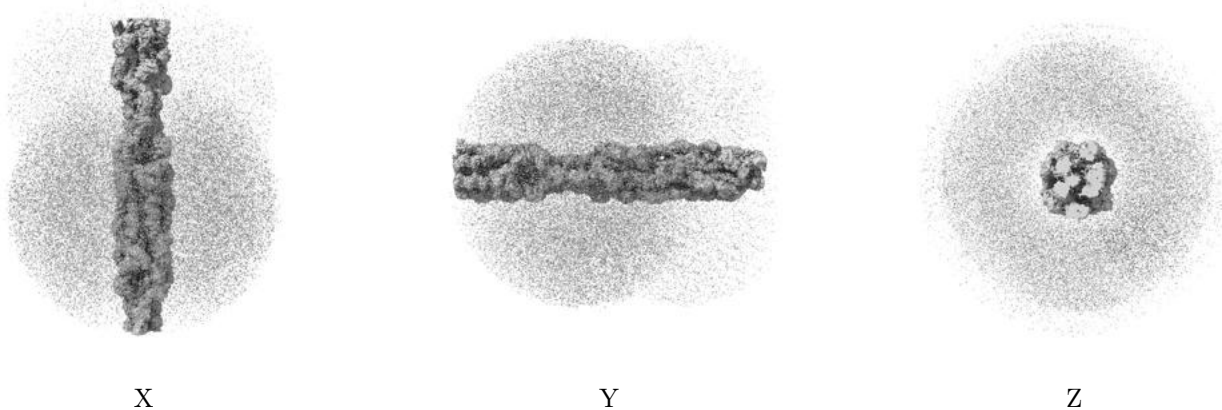
### 6.4.2 Raw map



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.

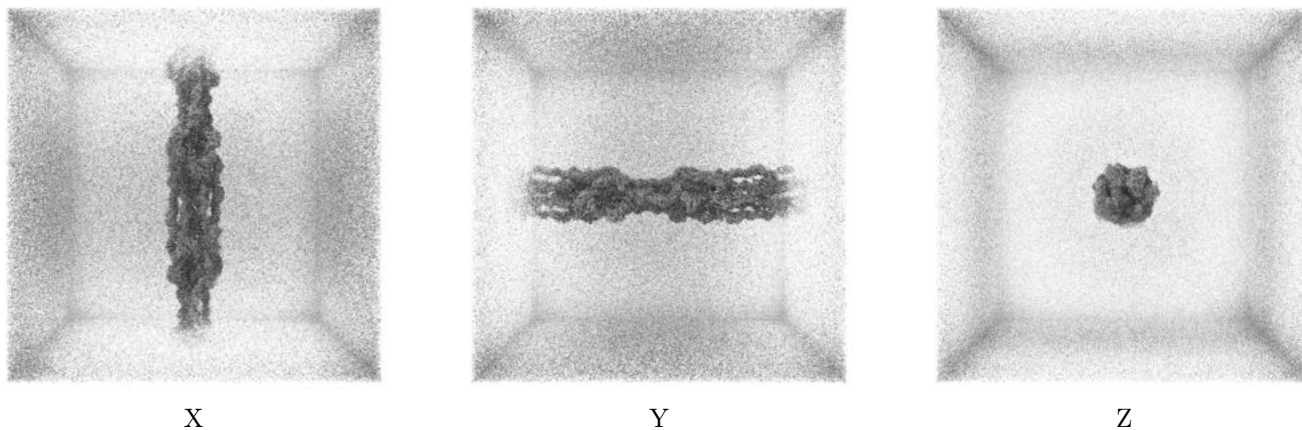
## 6.5 Orthogonal surface views [i](#)

### 6.5.1 Primary map



The images above show the 3D surface view of the map at the recommended contour level 0.0728. 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](#)

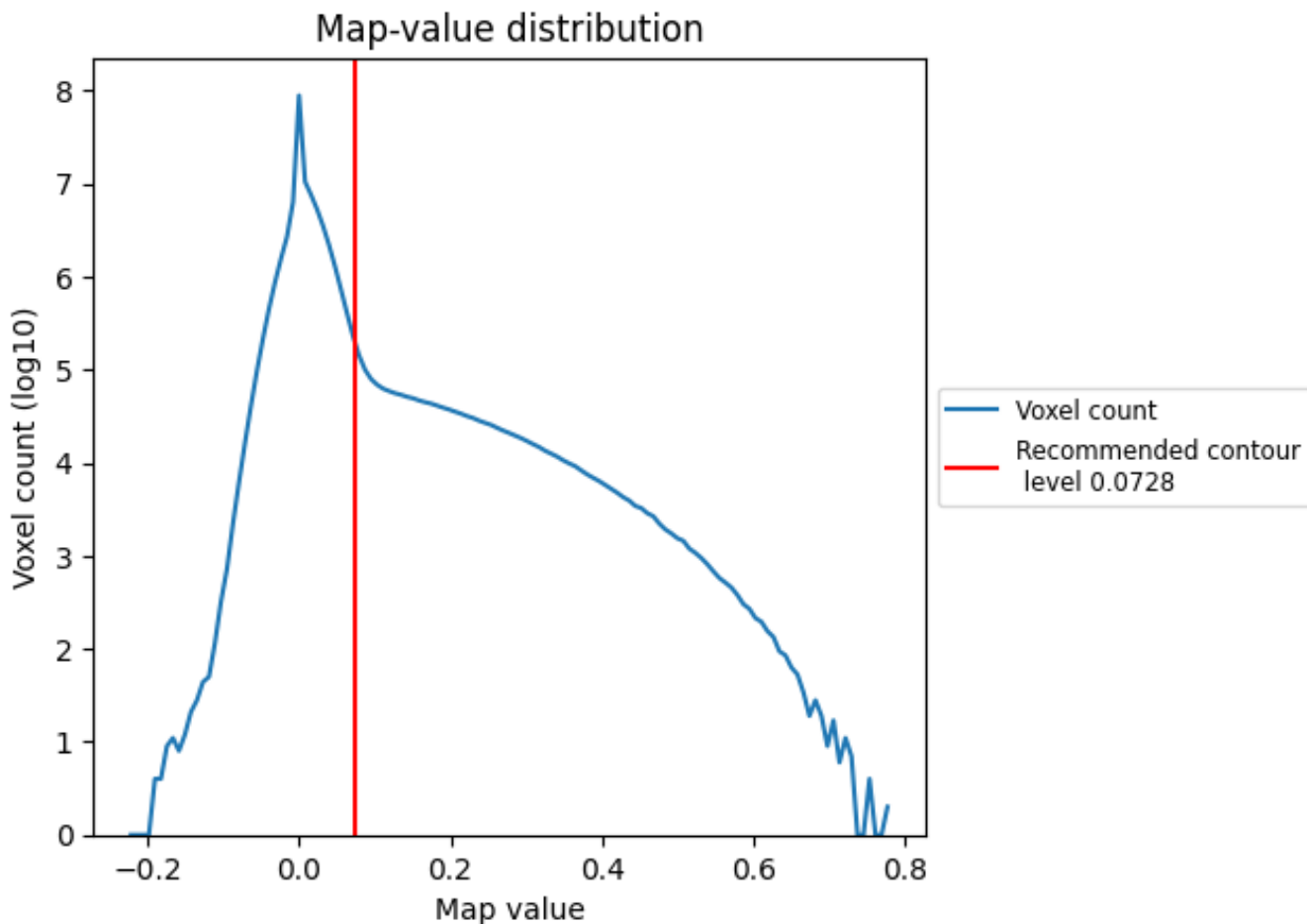
This section was not generated. No masks/segmentation were deposited.



## 7 Map analysis [i](#)

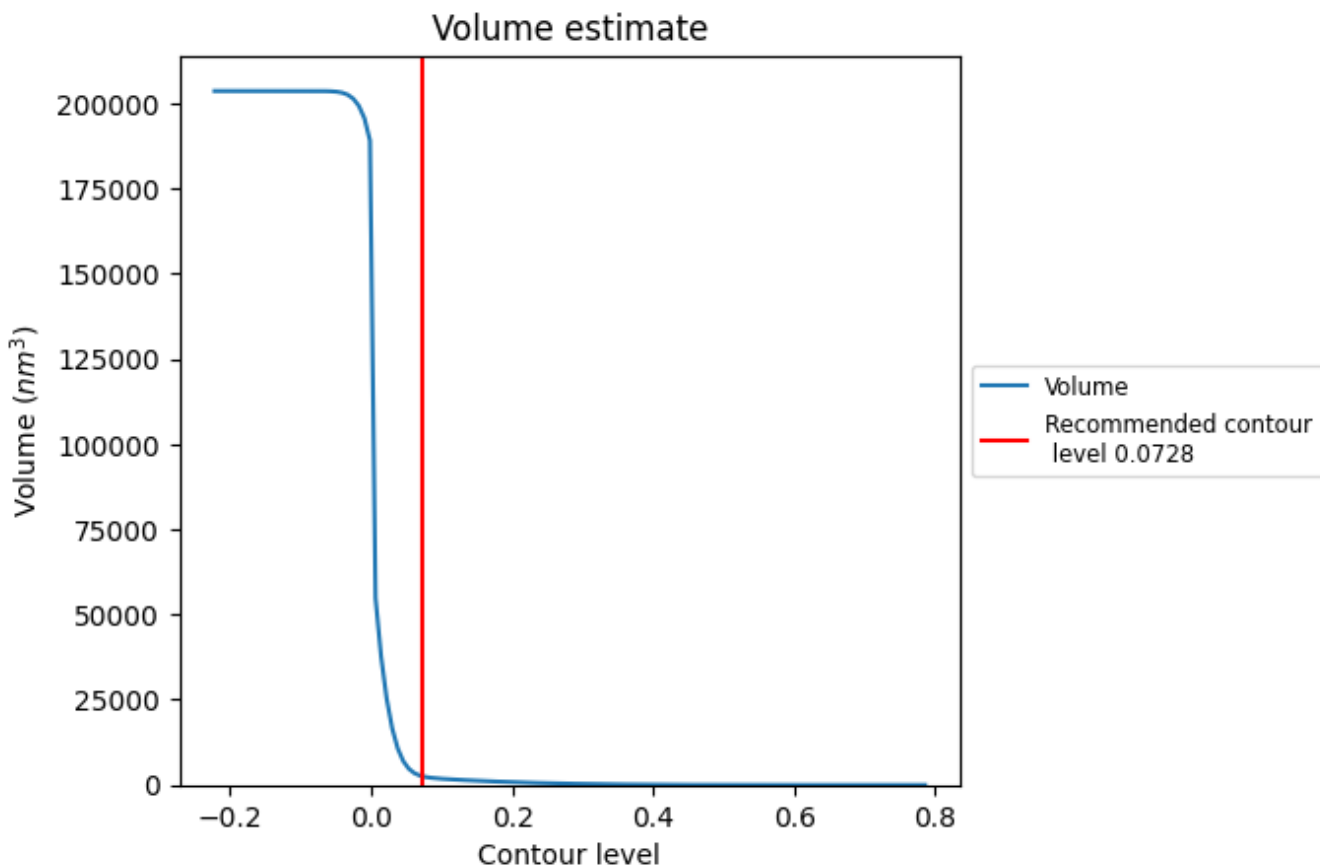
This section contains the results of statistical analysis of the map.

### 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.

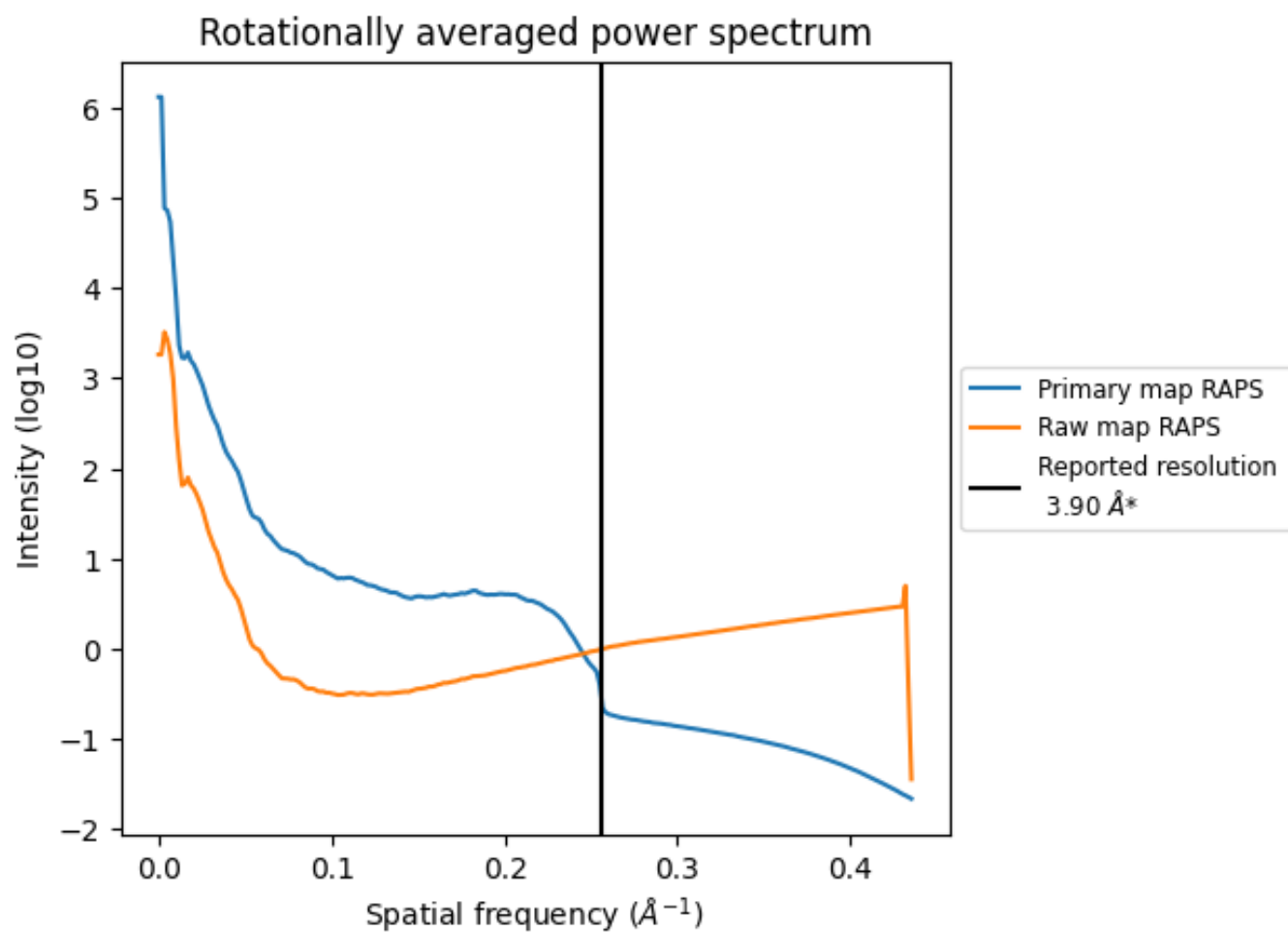
## 7.2 Volume estimate [\(i\)](#)



The volume at the recommended contour level is 2555  $\text{nm}^3$ ; this corresponds to an approximate mass of 2308 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](#)

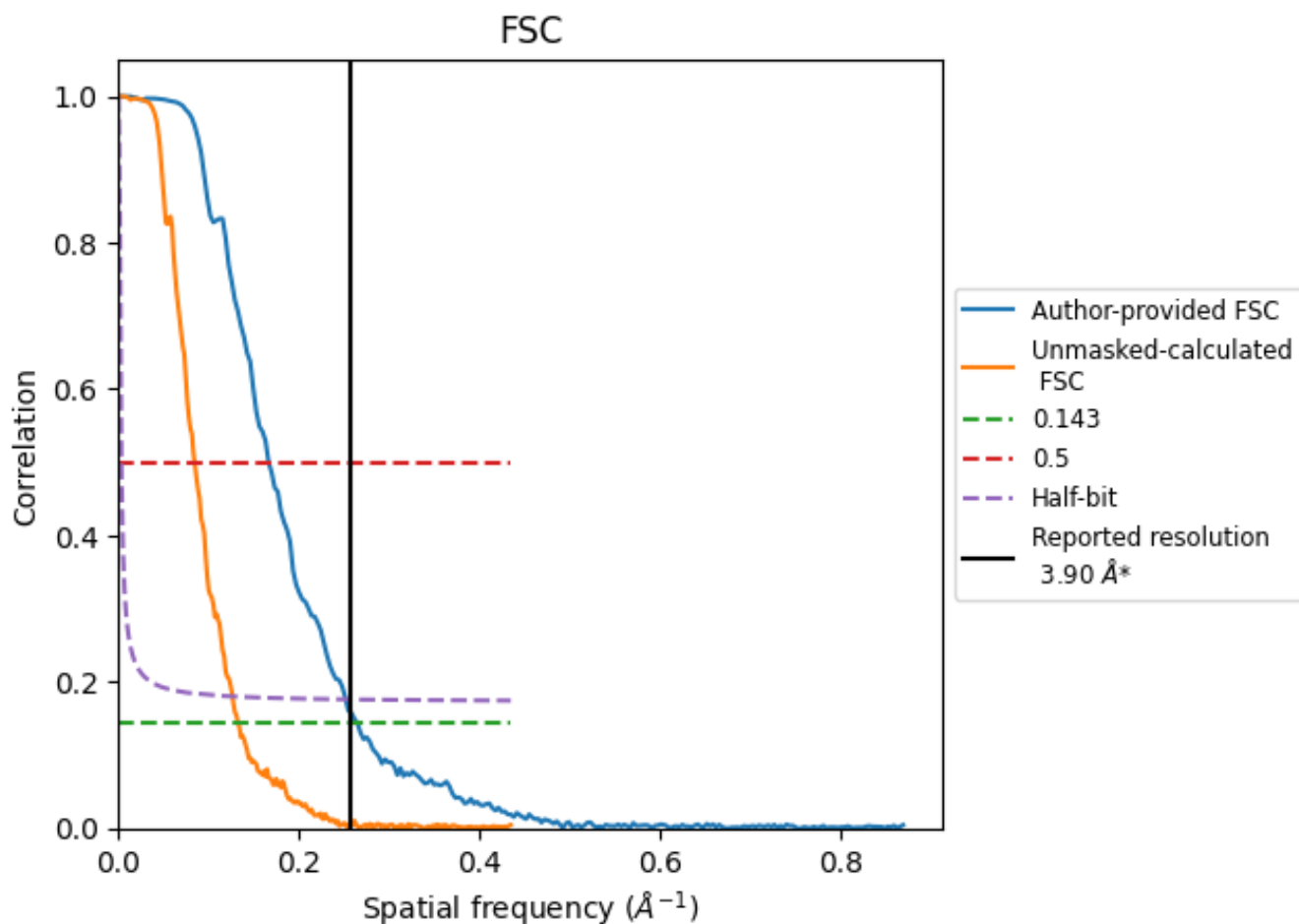


\*Reported resolution corresponds to spatial frequency of 0.256 Å<sup>-1</sup>

## 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 0.256 Å<sup>-1</sup>

## 8.2 Resolution estimates [i](#)

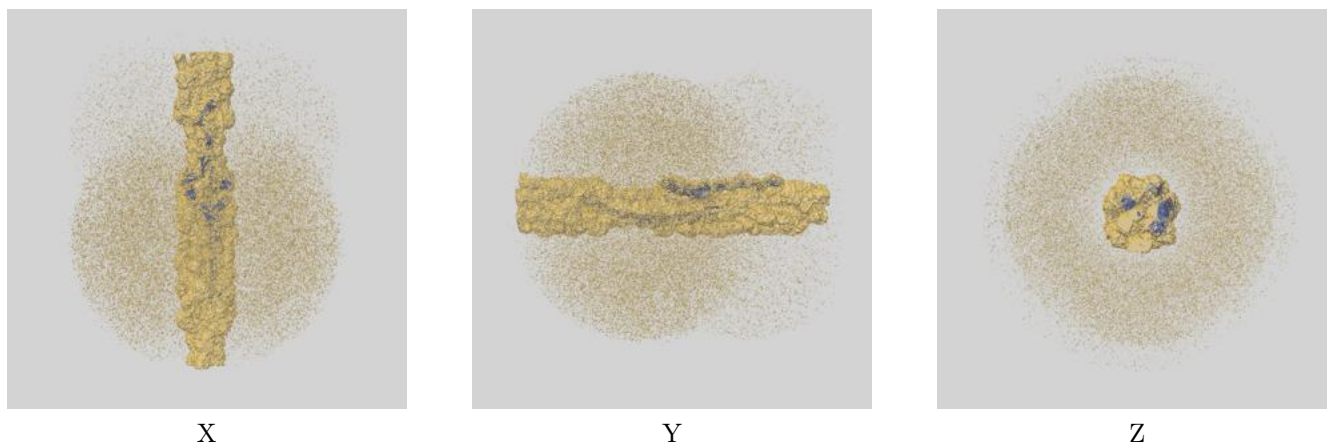
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	3.90	-	-
Author-provided FSC curve	3.77	6.00	3.98
Unmasked-calculated*	7.52	11.83	7.91

\*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 7.52 differs from the reported value 3.9 by more than 10 %

## 9 Map-model fit [i](#)

This section contains information regarding the fit between EMDB map EMD-41680 and PDB model 8TXC. Per-residue inclusion information can be found in section 3 on page 5.

### 9.1 Map-model overlay [i](#)



The images above show the 3D surface view of the map at the recommended contour level 0.0728 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 [\(i\)](#)



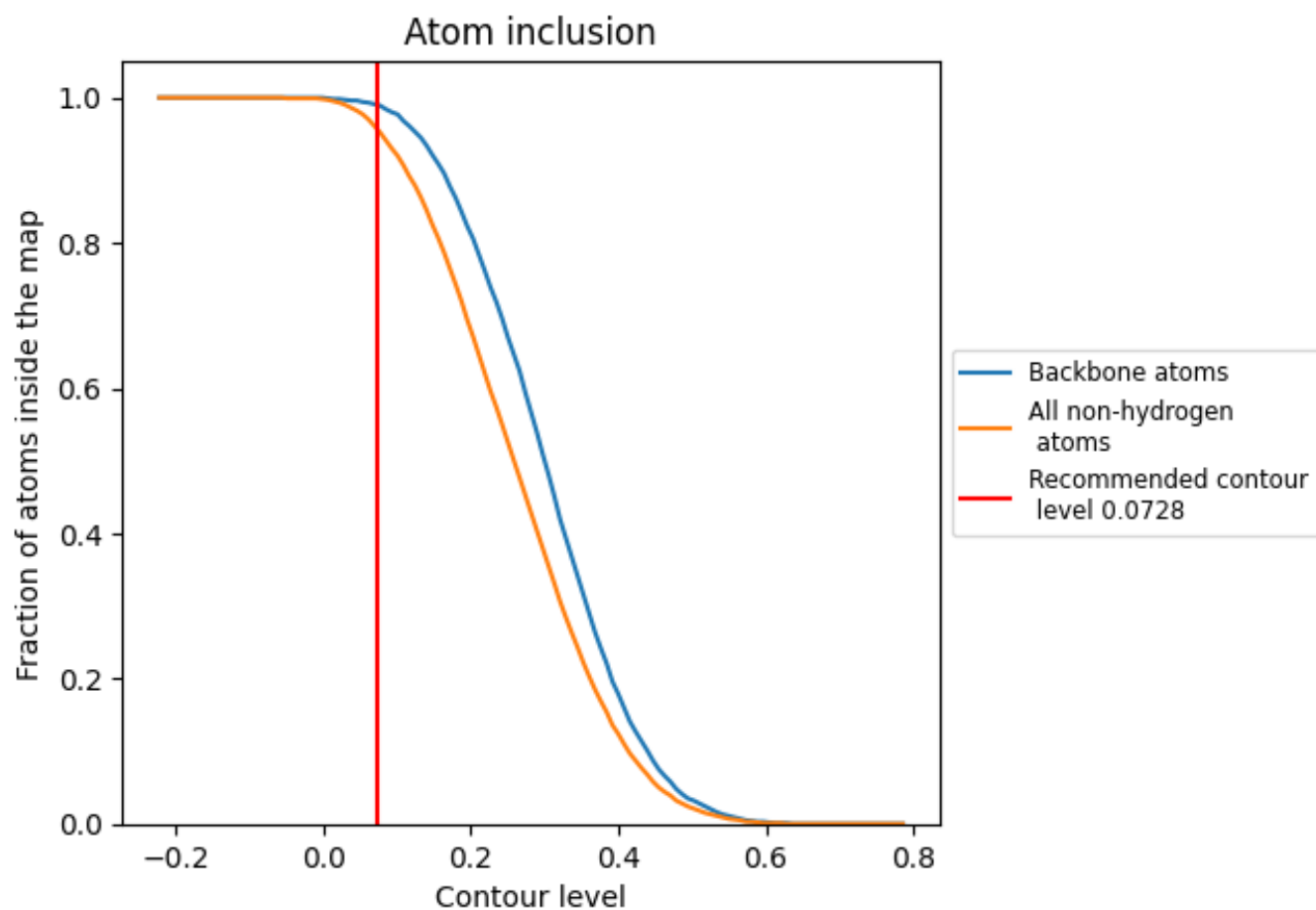
The images above show the model with each residue coloured according to its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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 (0.0728).

## 9.4 Atom inclusion [i](#)







At the recommended contour level, 99% of all backbone atoms, 96% 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 (0.0728) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	 0.9570	 0.2180
A	 0.9570	 0.2180

