wwPDB EM Validation Summary Report

Oct 28, 2023 – 12:47 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*
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:

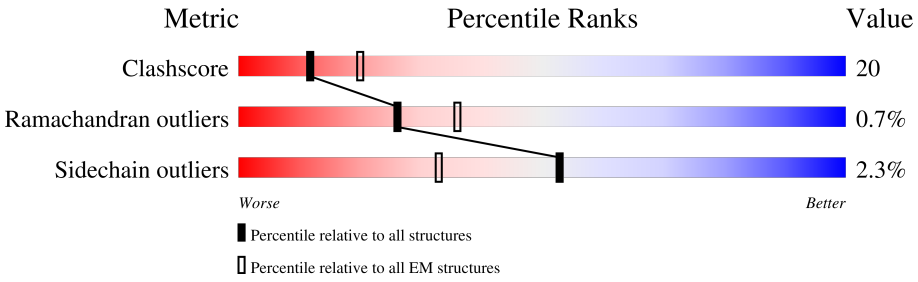
EMDB validation analysis	:	0.0.1.dev70
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.9
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 3.90 Å.

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



Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	158937	4297
Ramachandran outliers	154571	4023
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	1987	61% 33% • 5%

2 Entry composition (i)

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...

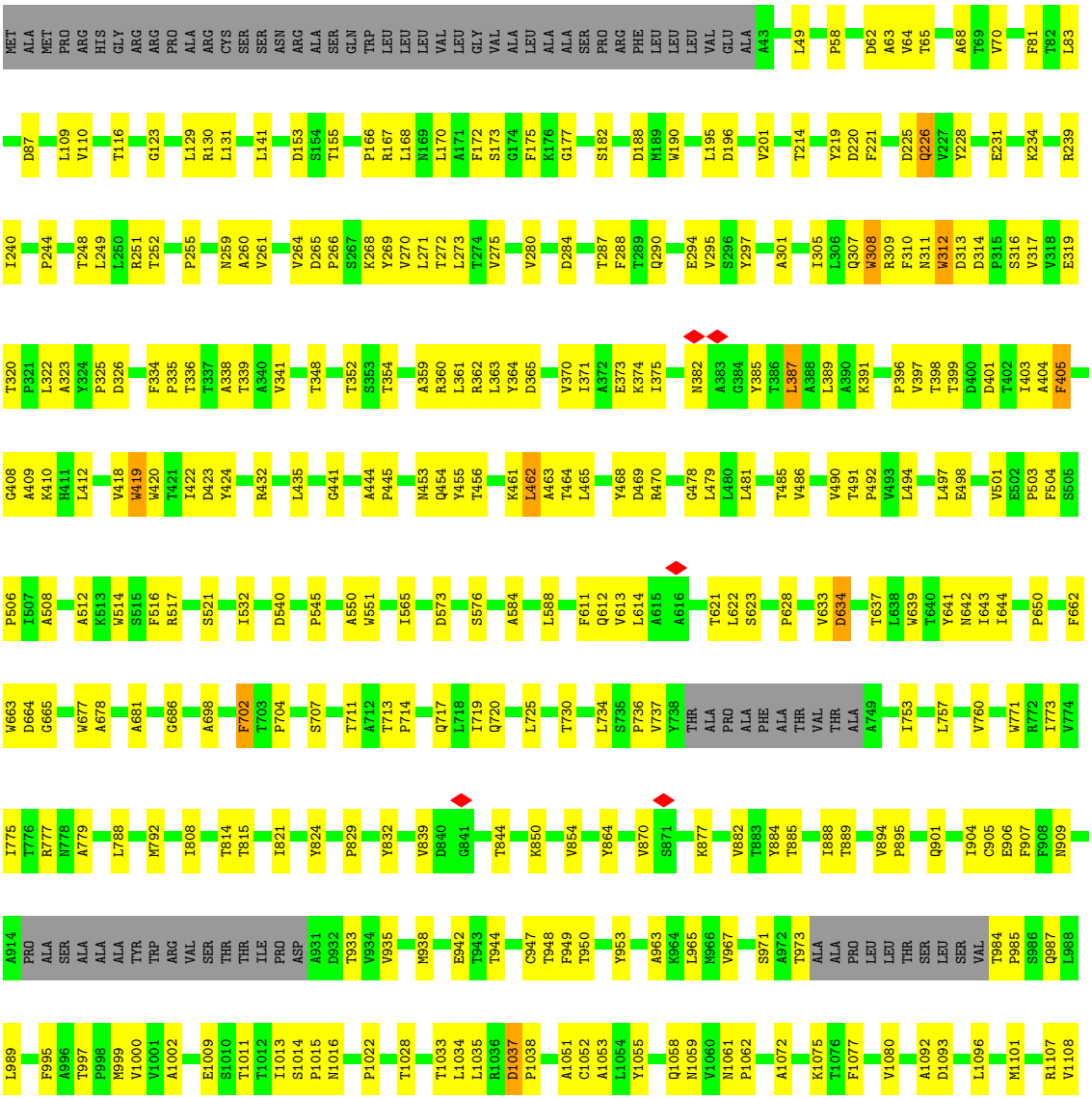
Chain	Residue	Modelled	Actual	Comment	Reference
A	1914	PRO	ARG	conflict	UNP Q8LRM7
A	1915	PRO	ARG	conflict	UNP Q8LRM7
A	1917	PRO	HIS	conflict	UNP Q8LRM7
A	1919	SER	ALA	conflict	UNP Q8LRM7
A	1920	PRO	ARG	conflict	UNP Q8LRM7
A	1921	PRO	ARG	conflict	UNP Q8LRM7
A	1924	ASN	THR	conflict	UNP Q8LRM7
A	1925	ARG	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

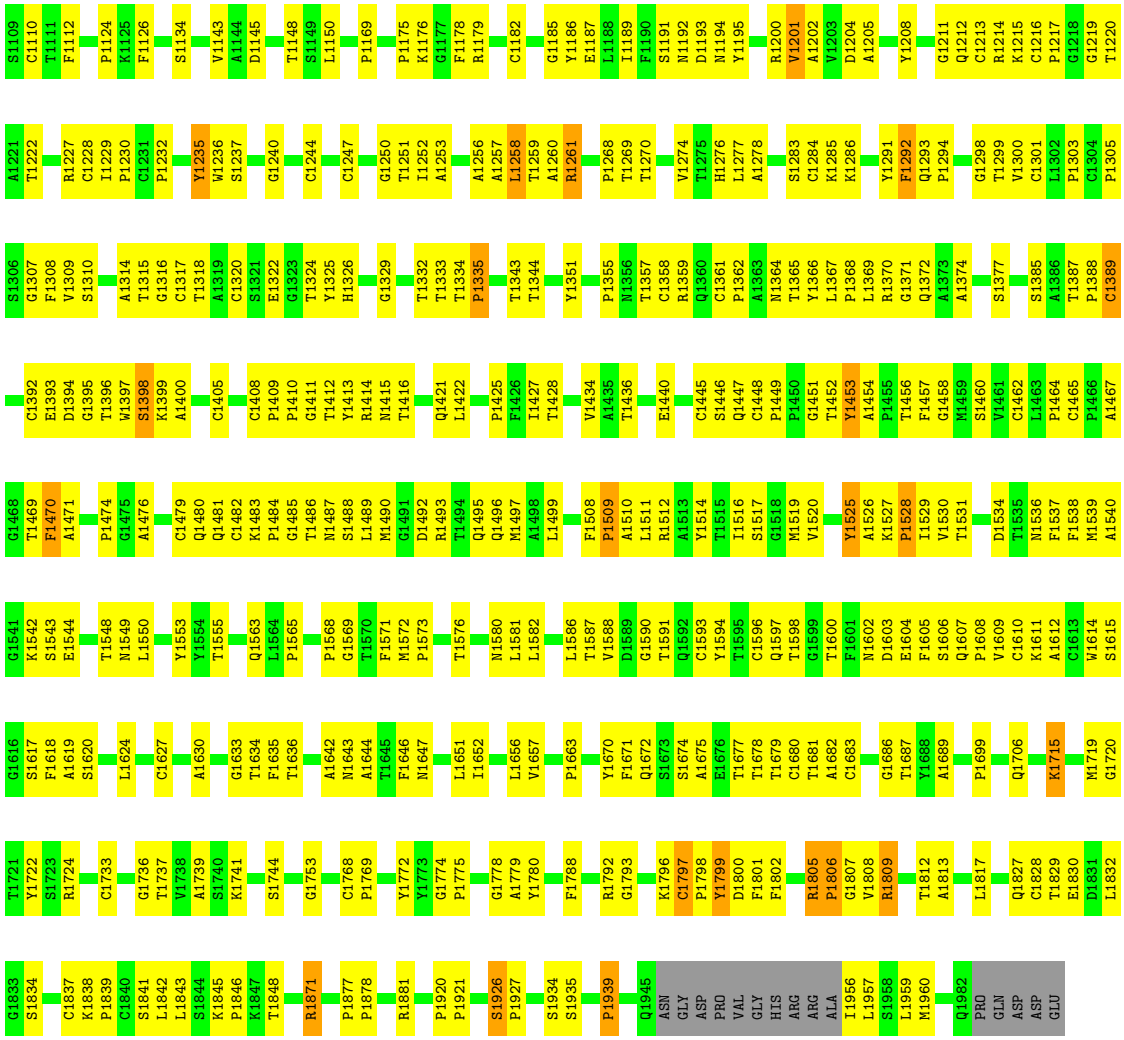
3 Residue-property plots (i)

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: Mastigoneme-like protein

Chain A:




S1109 C1110 T1111 F1112 P1124 K1125 F1126 S1134 V1143 A1144 D1145 T1148 S1149 L1150 P1169 P1175 K1176 G1177 F1178 R1179 C1182 G1185 Y1186 E1187 L1188 I1189 F1190 S1191 N1192 D1193 N1194 Y1195 R1200 V1201 A1202 V1203 D1204 A1205 Y1208 G1211 Q1212 C1213 R1214 K1215 C1216 P1217 G1218 G1219 T1220
A1221 T1222 R1227 C1228 I1229 P1230 C1231 P1232 Y1235 W1236 S1237 G1240 C1244 C1247 G1250 T1251 I1252 A1253 A1256 A1257 L1258 T1259 A1260 R1261 P1268 T1269 T1270 V1274 T1275 H1276 L1277 A1278 S1283 C1284 K1285 K1286 Y1291 F1292 Q1293 P1294 G1298 T1299 V1300 C1301 L1302 P1303 C1304 P1305
S1306 G1307 F1308 V1309 S1310 A1314 T1315 G1316 C1317 T1318 A1319 C1320 S1321 E1322 G1323 T1324 Y1325 H1326 G1329 T1332 T1333 T1334 P1335 T1343 T1344 Y1351 P1355 N1386 T1357 C1358 R1359 Q1360 C1361 P1362 A1363 N1364 T1365 Y1366 L1367 P1368 L1369 R1370 G1371 Q1372 A1373 A1374 S1377 S1385 A1386 T1387 P1388 C1389
C1392 E1393 D1394 G1395 T1396 W1397 S1398 K1399 A1400 C1405 C1408 P1409 P1410 G1411 T1412 Y1413 R1414 N1415 T1416 Q1421 L1422 P1425 F1426 I1427 T1428 V1434 A1435 T1436 E1440 C1445 S1446 Q1447 C1448 P1449 P1450 G1451 T1452 Y1453 A1454 P1455 T1456 F1457 G1458 M1459 S1460 V1461 C1462 L1463 P1464 C1465 P1466 A1467
G1468 T1469 F1470 A1471 P1474 G1475 A1476 C1479 Q1480 Q1481 C1482 K1483 P1484 G1485 T1486 N1487 S1488 L1489 M1490 G1491 D1492 R1493 T1494 Q1495 Q1496 M1497 A1498 L1499 F1508 P1509 A1510 L1511 R1512 A1513 Y1514 T1515 I1516 S1517 G1518 M1519 V1520 Y1525 A1526 K1527 P1528 I1529 V1530 T1531 D1534 T1535 N1536 F1537 F1538 M1539 A1540
G1541 K1542 S1543 E1544 T1548 N1549 L1550 Y1553 Y1554 T1555 Q1563 L1564 P1565 P1568 G1569 T1570 F1571 M1572 P1573 T1576 N1580 L1581 L1582 L1586 T1587 V1588 D1589 G1590 T1591 Q1592 C1593 Y1594 T1595 C1596 Q1597 T1598 G1599 T1600 F1601 N1602 D1603 E1604 F1605 S1606 Q1607 P1608 V1609 C1610 K1611 A1612 C1613 W1614 S1615
G1616 S1617 F1618 A1619 S1620 L1624 C1627 A1630 G1633 T1634 F1635 T1636 A1642 N1643 A1644 T1645 F1646 N1647 L1661 I1662 L1666 V1667 P1663 Y1670 F1671 Q1672 S1673 S1674 A1675 E1676 T1677 T1678 T1679 C1680 T1681 A1682 C1683 G1686 T1687 Y1688 A1689 P1699 Q1706 K1715 M1719 G1720
T1721 Y1722 S1723 R1724 C1733 G1736 T1737 V1738 A1739 S1740 K1741 S1744 G1753 C1768 P1769 Y1772 Y1773 G1774 P1775 G1778 A1779 Y1780 F1788 R1792 G1793 K1796 C1797 P1798 Y1799 D1800 F1801 F1802 R1805 P1806 G1807 V1808 R1809 T1812 A1813 L1817 Q1827 C1828 T1829 E1830 D1831 L1832
G1833 S1834 C1837 K1838 P1839 C1840 S1841 L1842 L1843 S1844 K1845 P1846 K1847 T1848 R1871 P1877 P1878 R1881 P1920 P1921 S1926 P1927 S1934 S1935 P1939 Q1945 ASN GLY ASP PRO VAL GLY HIS ARG ARG ALA I1956 L1957 S1958 L1959 M1960 Q1982 PRO GLN ASP ASP GLU

4 Experimental information (i)

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 ($e^-/Å^2$)	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 (Å)	588.288, 588.288, 588.288	wwPDB
Map dimensions	512, 512, 512	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.149, 1.149, 1.149	Depositor

5 Model quality

5.1 Standard geometry

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(^{o})	Ideal(^{o})
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%)	22 60

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%)	50 71

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 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.

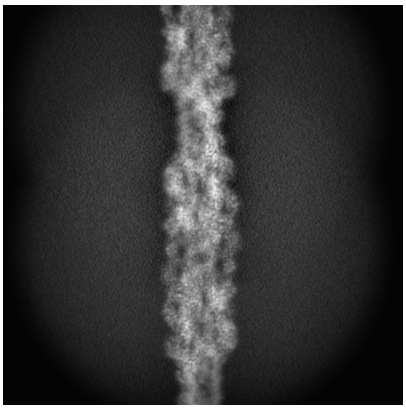
6 Map visualisation

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.

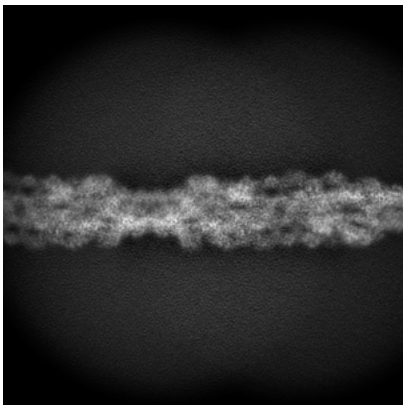
Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

6.1 Orthogonal projections

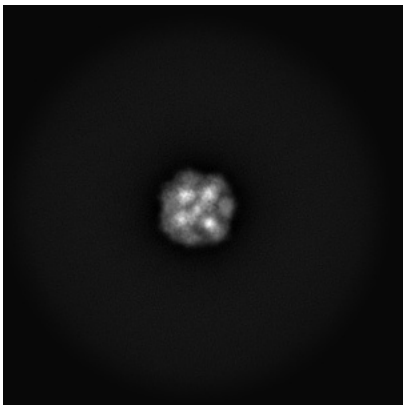
6.1.1 Primary map



X

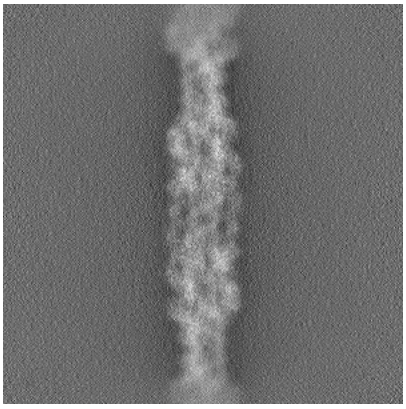


Y

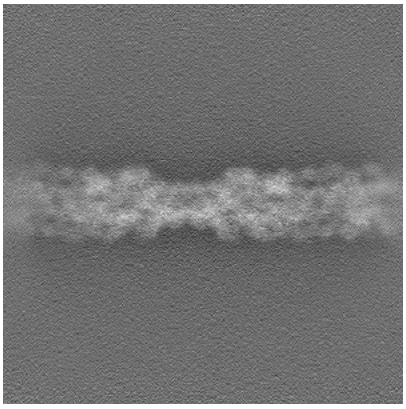


Z

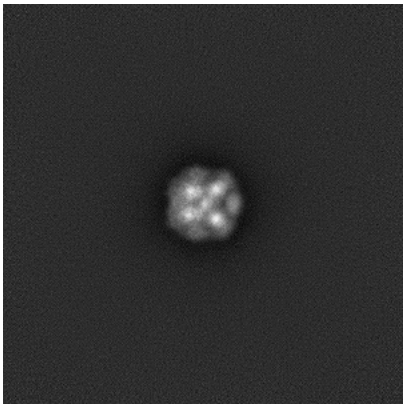
6.1.2 Raw map



X



Y

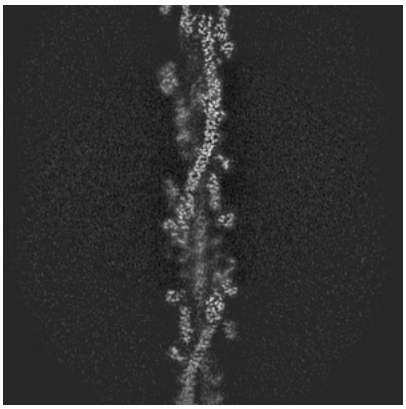


Z

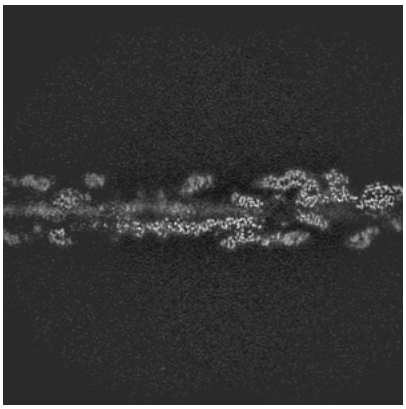
The images above show the map projected in three orthogonal directions.

6.2 Central slices (i)

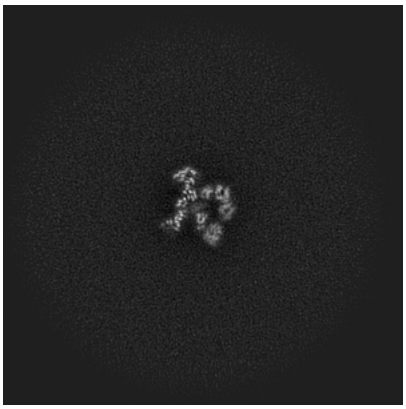
6.2.1 Primary map



X Index: 256

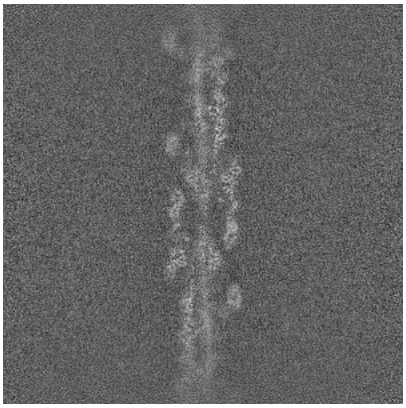


Y Index: 256

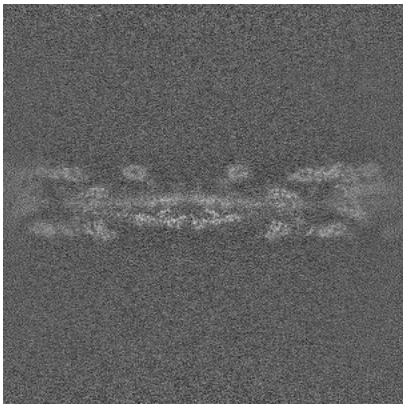


Z Index: 256

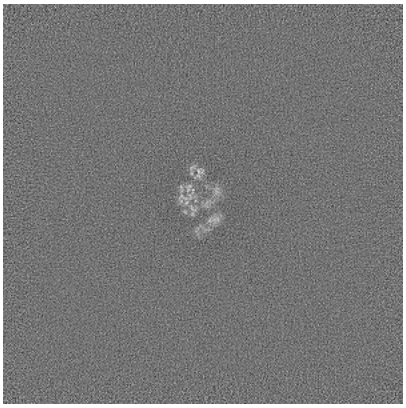
6.2.2 Raw map



X Index: 256



Y Index: 256

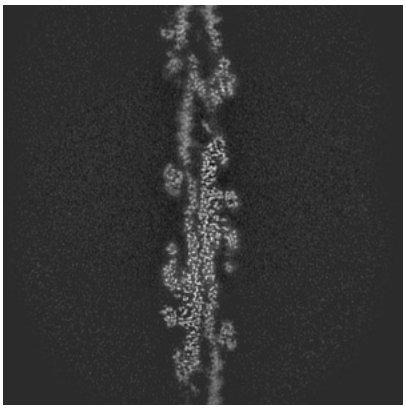


Z Index: 256

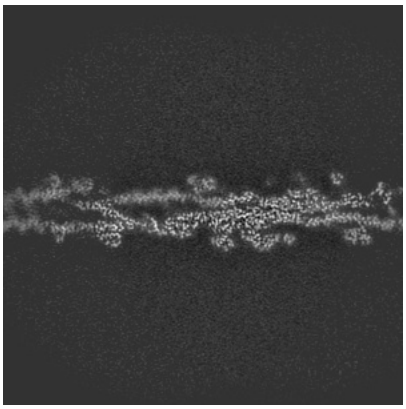
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

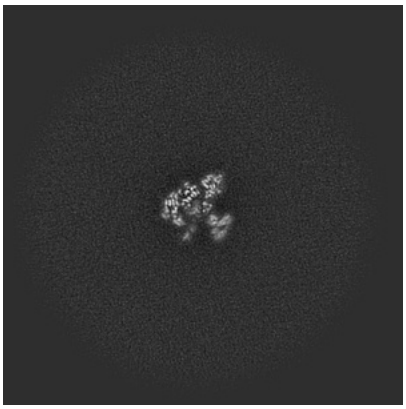
6.3.1 Primary map



X Index: 231

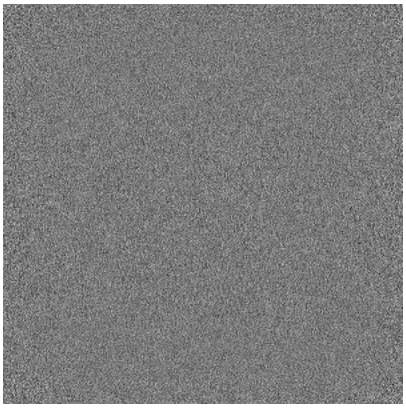


Y Index: 268

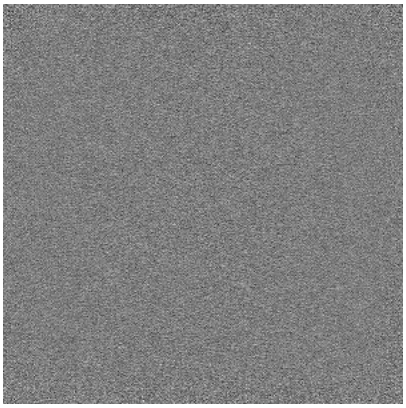


Z Index: 307

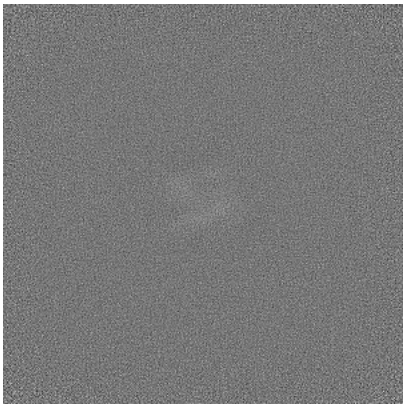
6.3.2 Raw map



X Index: 0



Y Index: 0

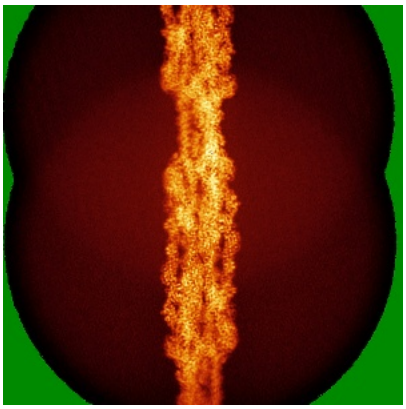


Z Index: 0

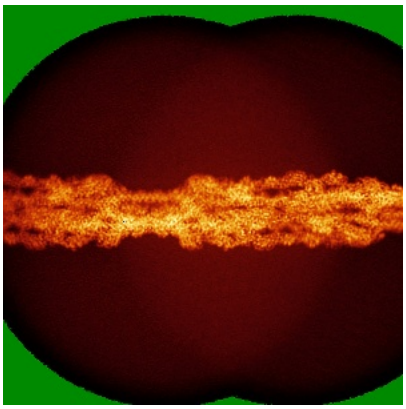
The images above show the largest variance slices of the map in three orthogonal directions.

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

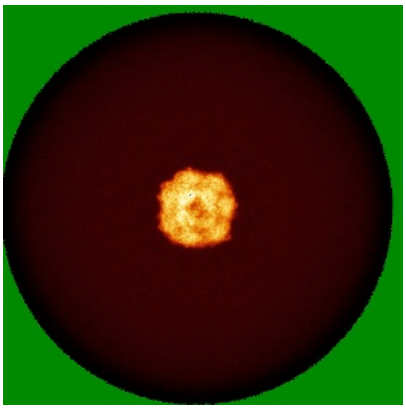
6.4.1 Primary map



X

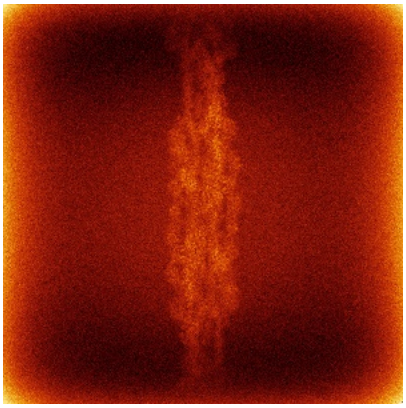


Y

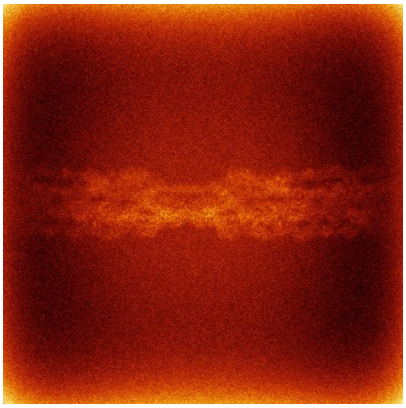


Z

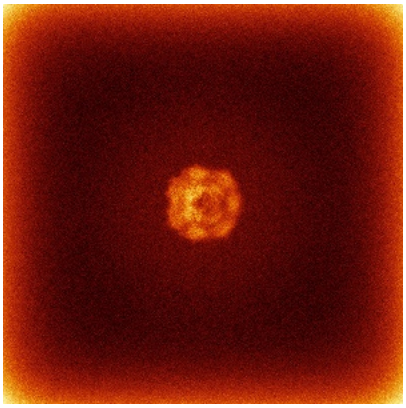
6.4.2 Raw map



X



Y

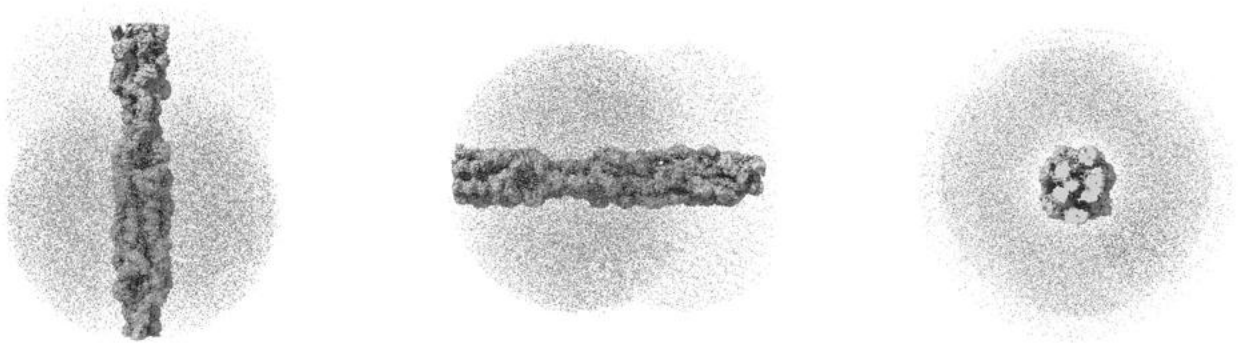


Z

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)

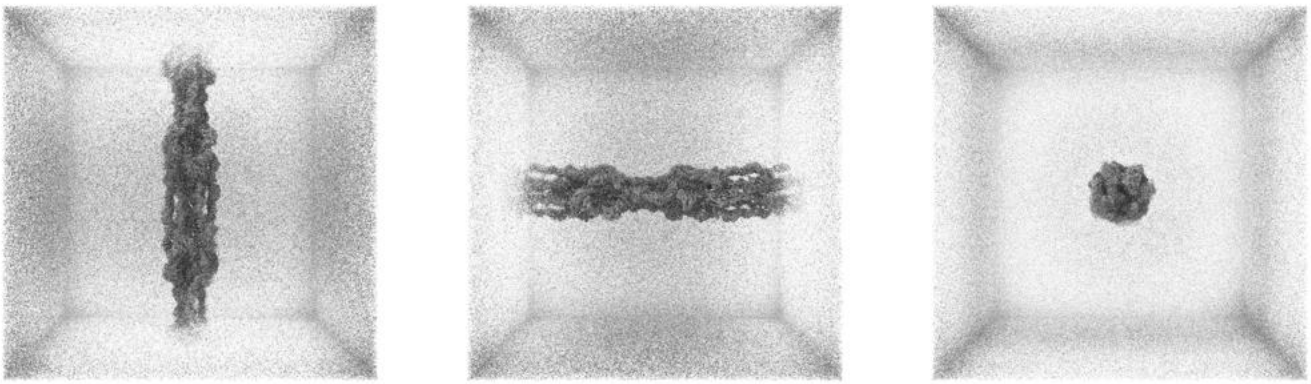
6.5.1 Primary map



X Y Z

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



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.

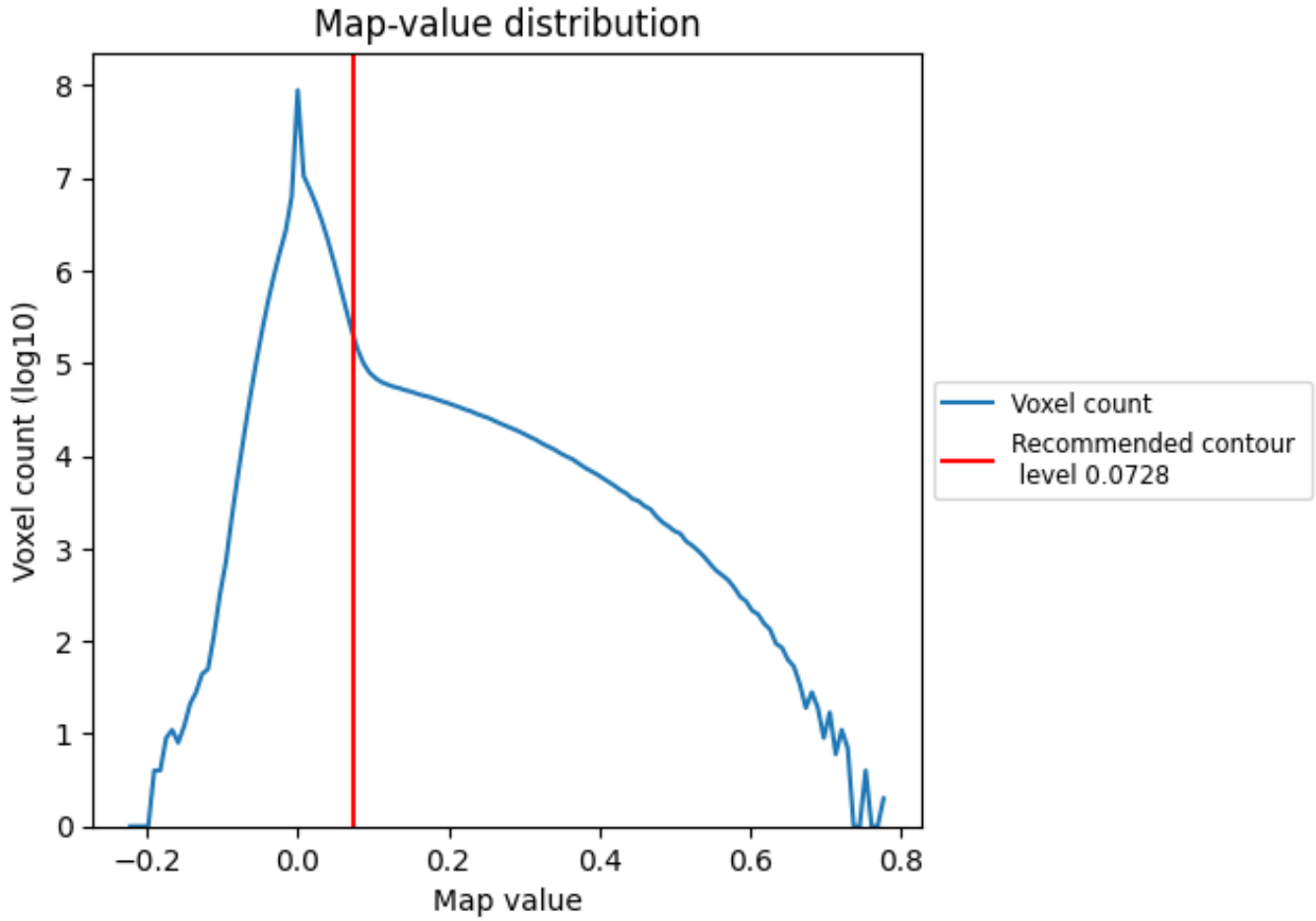
6.6 Mask visualisation (i)

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

7 Map analysis

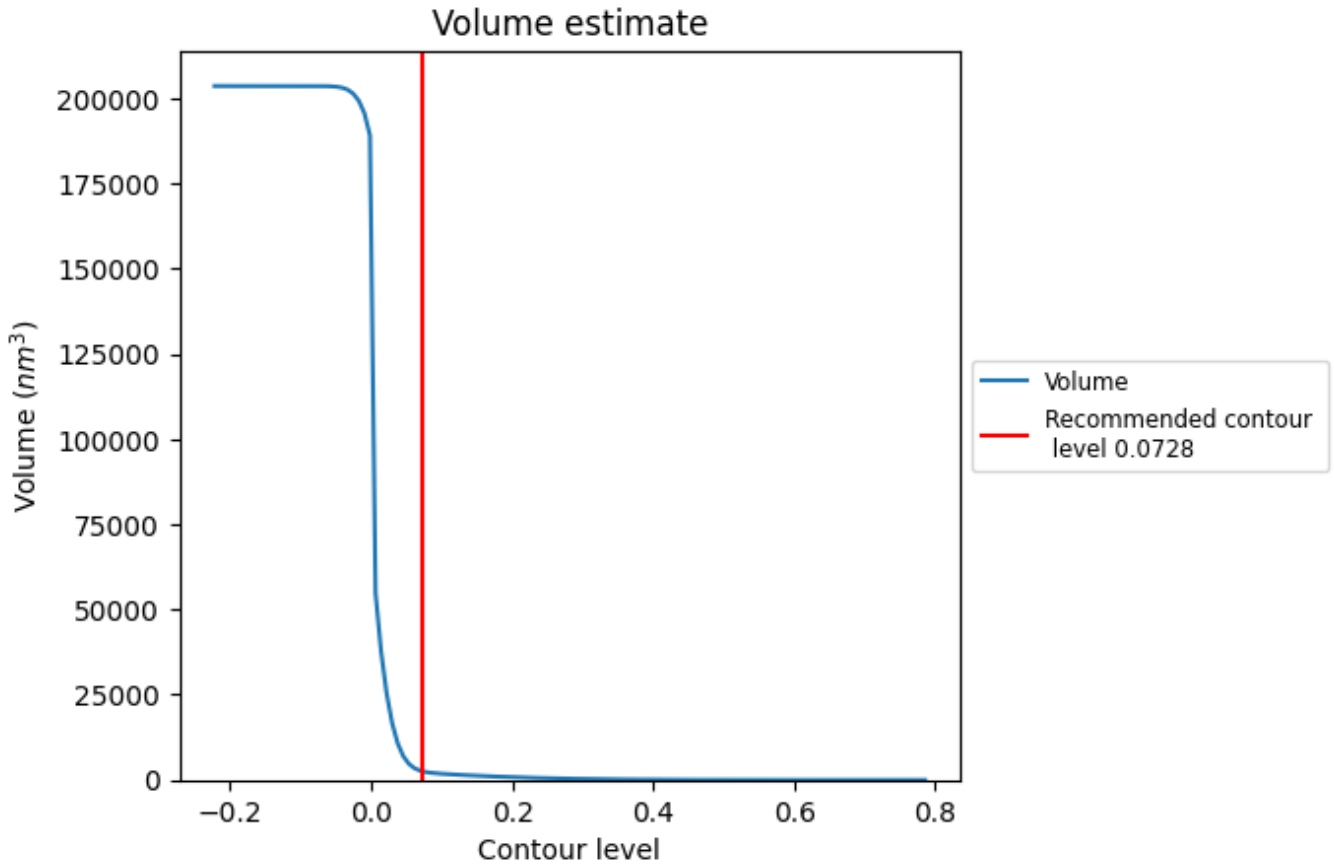
This section contains the results of statistical analysis of the map.

7.1 Map-value distribution



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.

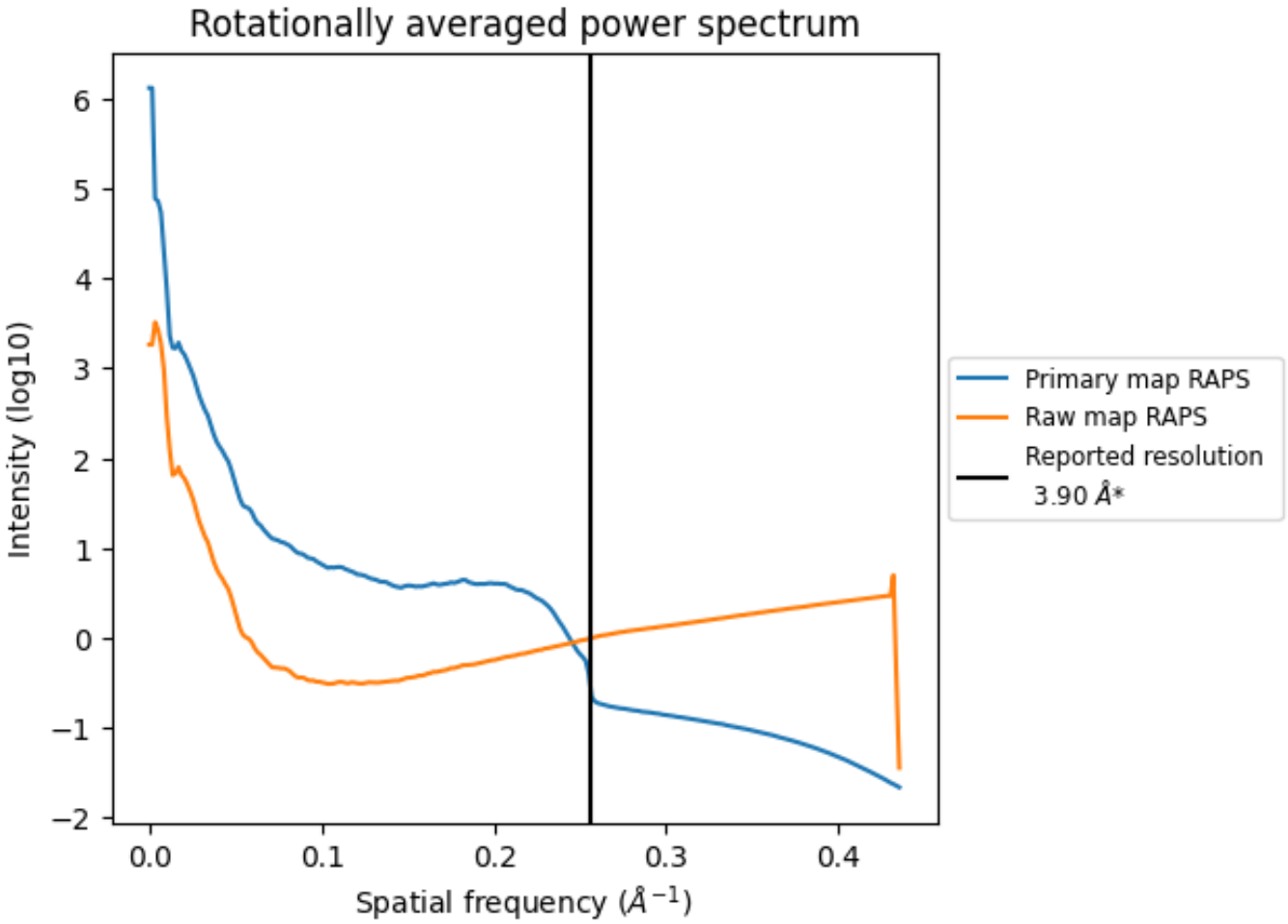
7.2 Volume estimate (i)



The volume at the recommended contour level is 2555 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)

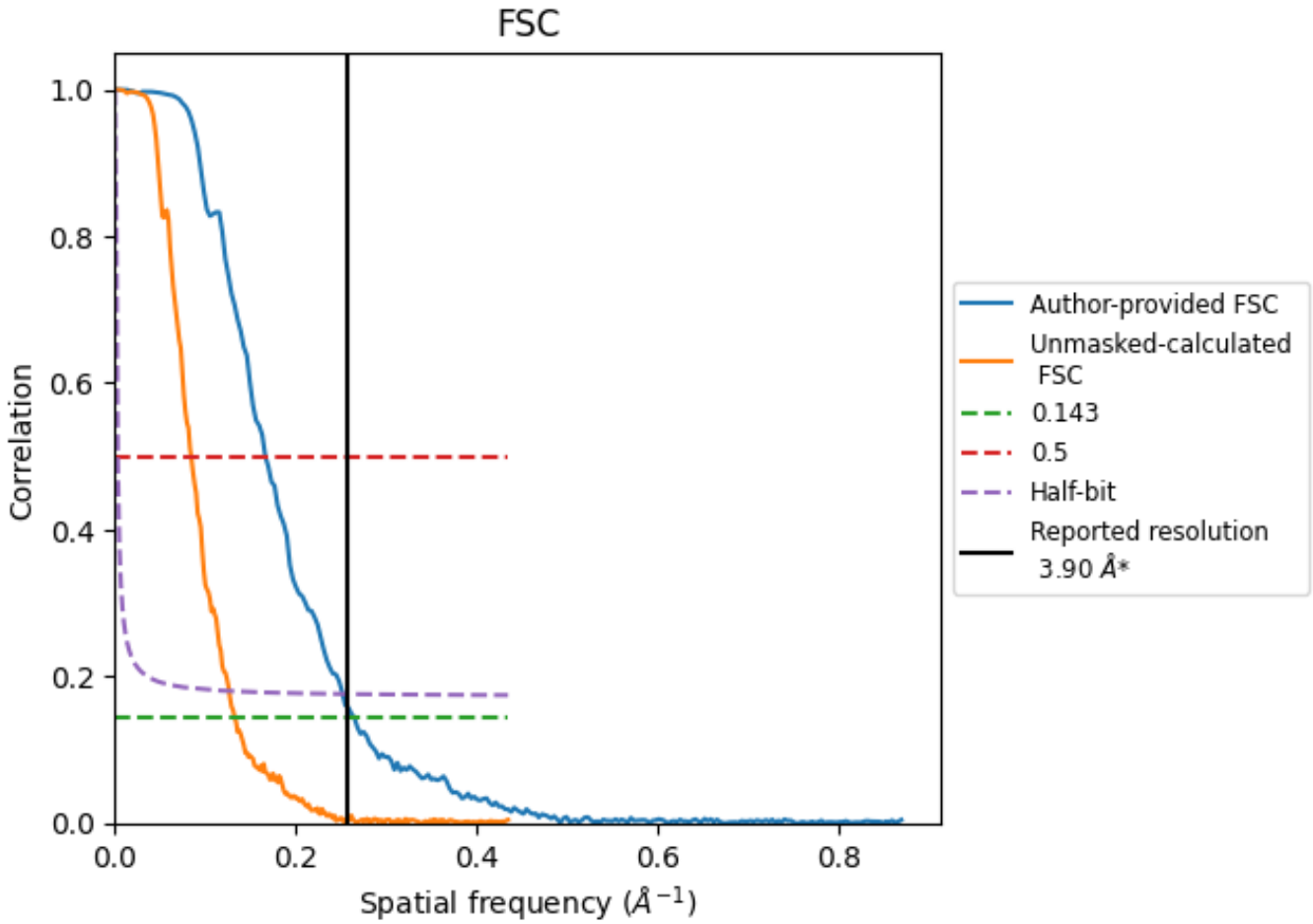


*Reported resolution corresponds to spatial frequency of 0.256 $Å^{-1}$

8 Fourier-Shell correlation

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



*Reported resolution corresponds to spatial frequency of 0.256 Å^{-1}

8.2 Resolution estimates (i)

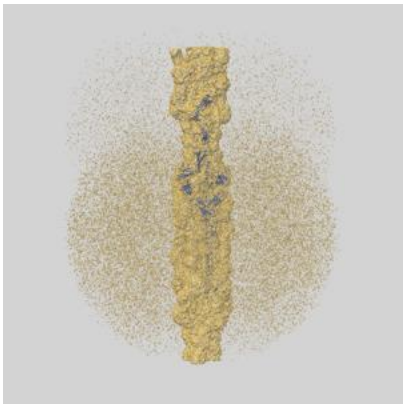
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	3.90	-	-
Author-provided FSC curve	3.77	6.00	3.98
Unmasked-calculated*	7.52	11.83	7.91

*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

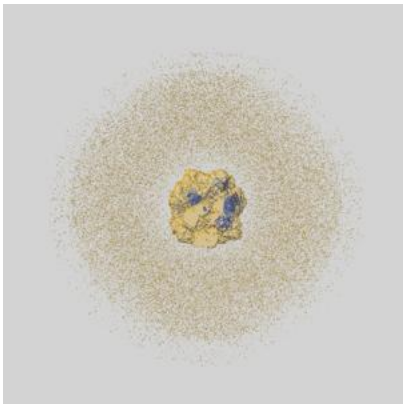
This section contains information regarding the fit between EMDB map EMD-41680 and PDB model 8TXC. Per-residue inclusion information can be found in section 3 on page 5.

9.1 Map-model overlay



X

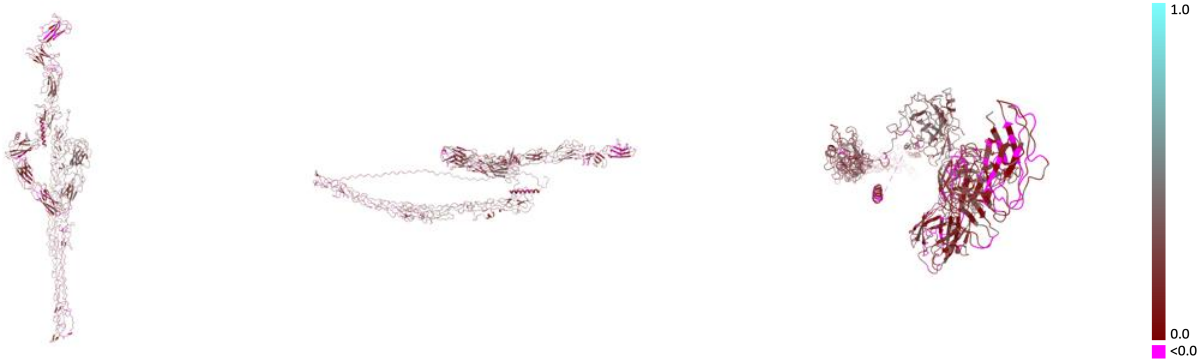
Y



Z

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.

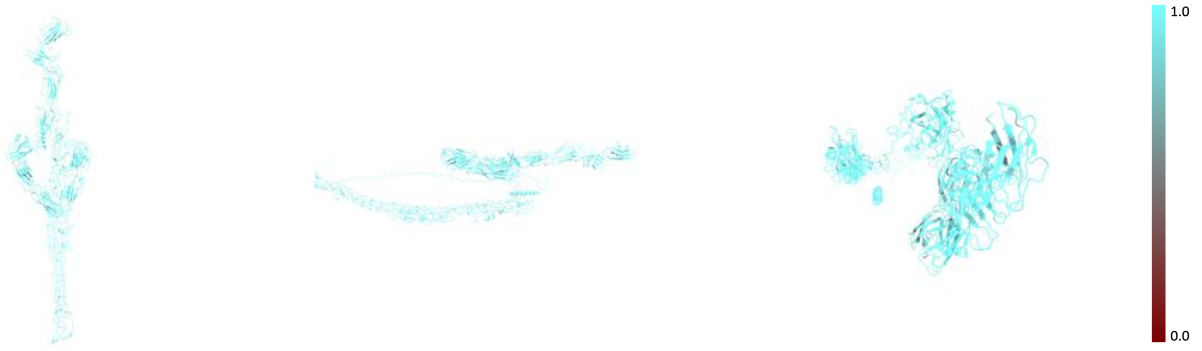
9.2 Q-score mapped to coordinate model



X Y Z

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.

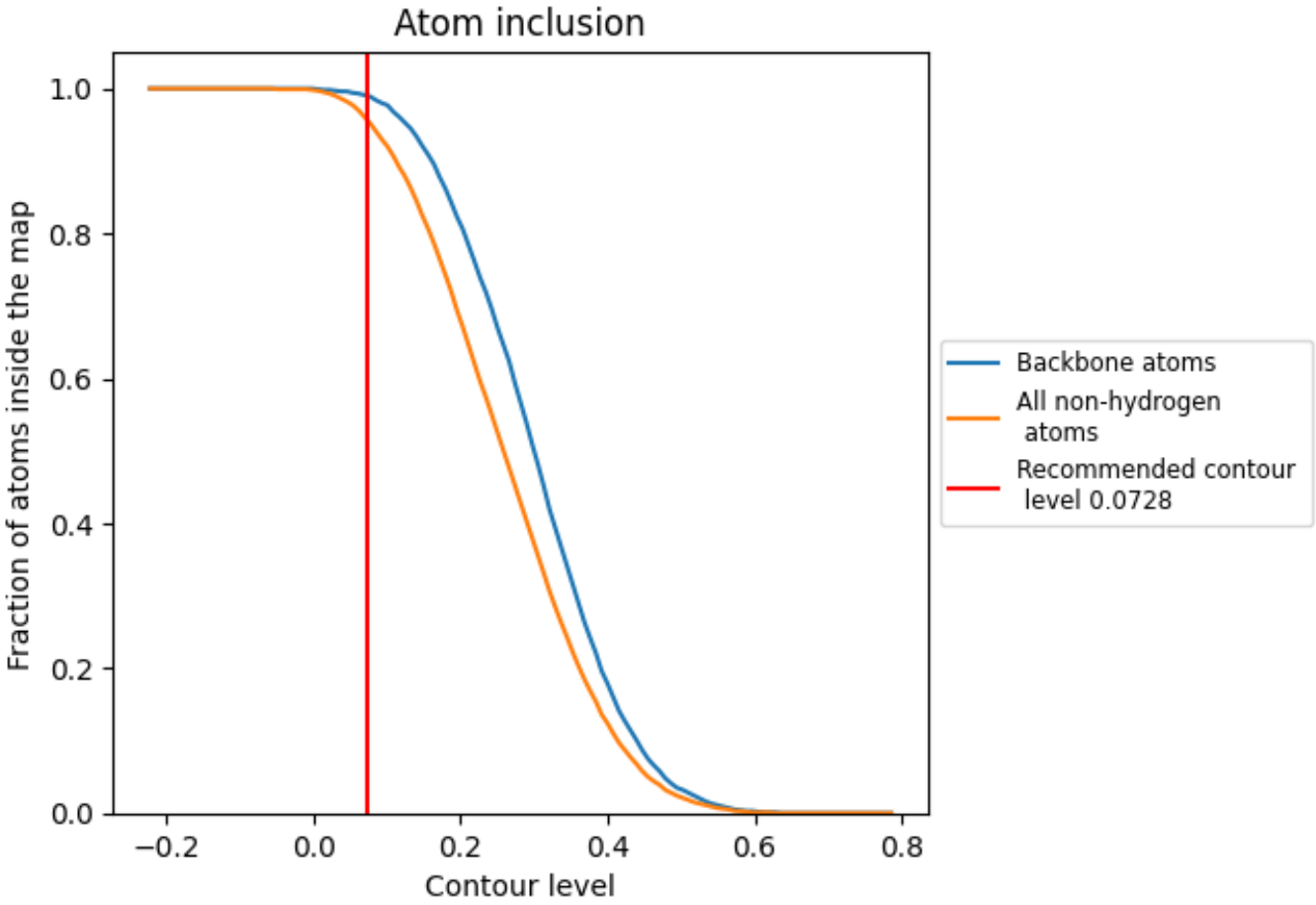
9.3 Atom inclusion mapped to coordinate model



X Y Z

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