



wwPDB EM Validation Summary Report ⓘ

Oct 6, 2021 – 08:24 am BST

PDB ID : 7OK0
EMDB ID : EMD-12960
Title : Cryo-EM structure of the Sulfolobus acidocaldarius RNA polymerase at 2.88 Å
Authors : Pilotto, S.; Fouqueau, T.; Lukoyanova, N.; Sheppard, C.; Lucas-Staat, S.; Diaz-Santin, L.M.; Matelska, D.; Prangishvili, D.; Cheung, A.C.M.; Werner, F.
Deposited on : 2021-05-17
Resolution : 2.90 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

EMDB validation analysis : 0.0.0.dev97
Mogul : 1.8.5 (274361), CSD as541be (2020)
MolProbity : 4.02b-467
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.23.2

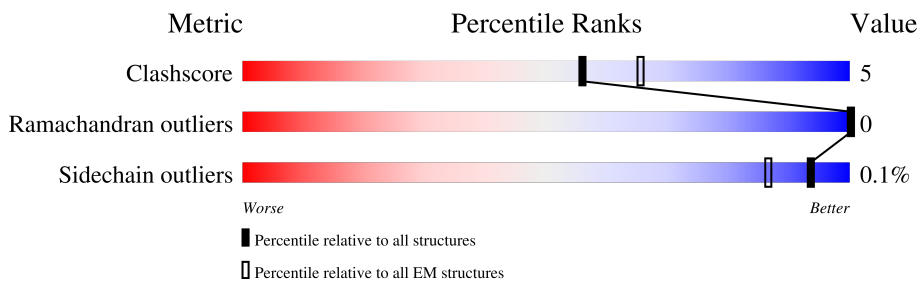
1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 2.90 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.




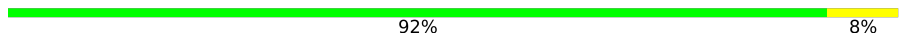



Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826

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 $\leq 5\%$. 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	880	
2	B	1126	
3	C	393	
4	D	264	
5	E	183	
6	F	114	
7	G	130	
8	H	84	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
9	K	89	
10	L	90	
11	N	66	
12	P	48	
13	Q	105	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
16	F3S	D	301	-	-	X	-

2 Entry composition i

There are 16 unique types of molecules in this entry. The entry contains 25636 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 DNA-directed RNA polymerase subunit A'.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	880	7023	4468	1237	1288	30	0	0

- Molecule 2 is a protein called DNA-directed RNA polymerase subunit B.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	B	1109	8725	5531	1533	1628	33	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	772	ILE	VAL	variant	UNP P11513

- Molecule 3 is a protein called DNA-directed RNA polymerase subunit A''.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	C	393	3087	1942	533	598	14	0	0

- Molecule 4 is a protein called DNA-directed RNA polymerase subunit D.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	D	261	2057	1313	342	389	13	0	0

- Molecule 5 is a protein called DNA-directed RNA polymerase subunit E.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
5	E	51	402	262	62	78	0	0

- Molecule 6 is a protein called DNA-directed RNA polymerase, subunit F.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
6	F	7	56	36	8	11	1	0	0

- Molecule 7 is a protein called DNA-directed RNA polymerase, subunit G.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
7	G	122	991	637	155	190	9	0	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
G	125	HIS	-	expression tag	UNP Q4JAY4
G	126	HIS	-	expression tag	UNP Q4JAY4
G	127	HIS	-	expression tag	UNP Q4JAY4
G	128	HIS	-	expression tag	UNP Q4JAY4
G	129	HIS	-	expression tag	UNP Q4JAY4
G	130	HIS	-	expression tag	UNP Q4JAY4

- Molecule 8 is a protein called DNA-directed RNA polymerase subunit H.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
8	H	78	617	399	104	114		0	0

- Molecule 9 is a protein called DNA-directed RNA polymerase subunit K.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
9	K	86	690	441	120	127	2	0	0

- Molecule 10 is a protein called DNA-directed RNA polymerase subunit L.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
10	L	90	707	459	112	134	2	0	0

- Molecule 11 is a protein called DNA-directed RNA polymerase subunit N.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
11	N	66	534	339	97	91	7	0	0

- Molecule 12 is a protein called DNA-directed RNA polymerase subunit P.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
12	P	47	376	241	67	63	5	0	0

- Molecule 13 is a protein called Conserved protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
13	Q	44	357	226	57	72	2	0	0

- Molecule 14 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		AltConf
			Total	Mg	
14	A	1	1	1	0

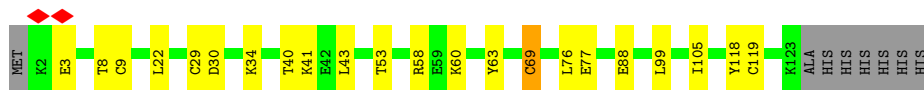
- Molecule 15 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		AltConf
			Total	Zn	
15	A	3	3	3	0
15	B	1	1	1	0
15	N	1	1	1	0
15	P	1	1	1	0

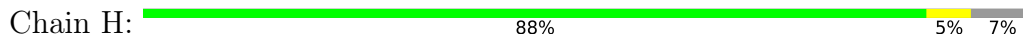
- Molecule 16 is FE3-S4 CLUSTER (three-letter code: F3S) (formula: Fe₃S₄).



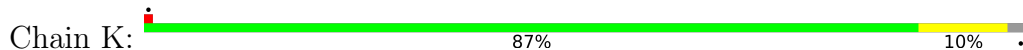
Mol	Chain	Residues	Atoms			AltConf
			Total	Fe	S	
16	D	1	7	3	4	0



- Molecule 8: DNA-directed RNA polymerase subunit H



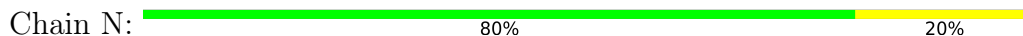
- Molecule 9: DNA-directed RNA polymerase subunit K



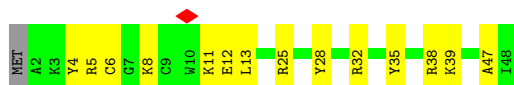
- Molecule 10: DNA-directed RNA polymerase subunit L



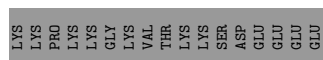
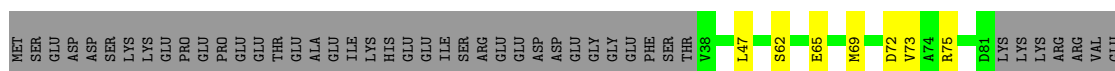
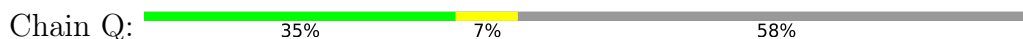
- Molecule 11: DNA-directed RNA polymerase subunit N



- Molecule 12: DNA-directed RNA polymerase subunit P



- Molecule 13: Conserved protein



4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	423157	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION; CTF aplitude was further corrected on selected particles after 3D reconstruction to correct the 30 degrees tilt used during data collection	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{\AA}^2$)	44.46	Depositor
Minimum defocus (nm)	1300	Depositor
Maximum defocus (nm)	2500	Depositor
Magnification	81000	Depositor
Image detector	GATAN K3 (6k x 4k)	Depositor
Maximum map value	54.865	Depositor
Minimum map value	-22.123	Depositor
Average map value	-0.000	Depositor
Map value standard deviation	1.000	Depositor
Recommended contour level	4.0	Depositor
Map size (Å)	360.22, 360.22, 360.22	wwPDB
Map dimensions	332, 332, 332	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.085, 1.085, 1.085	Depositor

5 Model quality [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: MG, F3S, ZN

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	0.26	0/7173	0.51	0/9699
2	B	0.26	0/8885	0.51	0/12016
3	C	0.24	0/3116	0.49	0/4193
4	D	0.26	0/2090	0.48	0/2824
5	E	0.26	0/410	0.43	0/551
6	F	0.25	0/56	0.49	0/73
7	G	0.31	0/1006	0.53	1/1342 (0.1%)
8	H	0.26	0/630	0.45	0/851
9	K	0.26	0/697	0.48	0/938
10	L	0.26	0/716	0.44	0/964
11	N	0.27	0/545	0.53	0/735
12	P	0.27	0/382	0.61	0/510
13	Q	0.22	0/360	0.44	0/482
All	All	0.26	0/26066	0.50	1/35178 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
7	G	69	CYS	CA-CB-SG	8.44	129.20	114.00

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts [i](#)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen

atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	7023	0	7109	89	0
2	B	8725	0	8862	93	0
3	C	3087	0	3190	34	0
4	D	2057	0	2110	19	0
5	E	402	0	382	7	0
6	F	56	0	61	0	0
7	G	991	0	996	15	0
8	H	617	0	644	3	0
9	K	690	0	745	7	0
10	L	707	0	751	6	0
11	N	534	0	545	8	0
12	P	376	0	399	11	0
13	Q	357	0	360	5	0
14	A	1	0	0	0	0
15	A	3	0	0	0	0
15	B	1	0	0	0	0
15	N	1	0	0	0	0
15	P	1	0	0	0	0
16	D	7	0	0	4	0
All	All	25636	0	26154	267	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 5.

The worst 5 of 267 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)
4:D:209:CYS:HB3	16:D:301:F3S:S2	2.14	0.87
7:G:69:CYS:HB3	7:G:119:CYS:HA	1.59	0.85
11:N:24:ARG:NH2	11:N:37:GLU:OE2	2.20	0.75
7:G:69:CYS:CB	7:G:119:CYS:HA	2.18	0.74
1:A:98:CYS:SG	1:A:153:GLN:NE2	2.60	0.74

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	878/880 (100%)	851 (97%)	27 (3%)	0	100	100
2	B	1105/1126 (98%)	1055 (96%)	50 (4%)	0	100	100
3	C	391/393 (100%)	381 (97%)	10 (3%)	0	100	100
4	D	257/264 (97%)	244 (95%)	13 (5%)	0	100	100
5	E	47/183 (26%)	47 (100%)	0	0	100	100
6	F	5/114 (4%)	5 (100%)	0	0	100	100
7	G	120/130 (92%)	117 (98%)	3 (2%)	0	100	100
8	H	76/84 (90%)	75 (99%)	1 (1%)	0	100	100
9	K	84/89 (94%)	82 (98%)	2 (2%)	0	100	100
10	L	88/90 (98%)	87 (99%)	1 (1%)	0	100	100
11	N	64/66 (97%)	59 (92%)	5 (8%)	0	100	100
12	P	45/48 (94%)	38 (84%)	7 (16%)	0	100	100
13	Q	42/105 (40%)	41 (98%)	1 (2%)	0	100	100
All	All	3202/3572 (90%)	3082 (96%)	120 (4%)	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	767/768 (100%)	765 (100%)	2 (0%)	92	98

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	B	948/972 (98%)	948 (100%)	0	100	100
3	C	341/348 (98%)	341 (100%)	0	100	100
4	D	234/239 (98%)	234 (100%)	0	100	100
5	E	40/157 (26%)	40 (100%)	0	100	100
6	F	7/109 (6%)	7 (100%)	0	100	100
7	G	113/121 (93%)	113 (100%)	0	100	100
8	H	68/75 (91%)	68 (100%)	0	100	100
9	K	78/81 (96%)	78 (100%)	0	100	100
10	L	78/78 (100%)	78 (100%)	0	100	100
11	N	60/60 (100%)	60 (100%)	0	100	100
12	P	41/43 (95%)	41 (100%)	0	100	100
13	Q	40/98 (41%)	40 (100%)	0	100	100
All	All	2815/3149 (89%)	2813 (100%)	2 (0%)	93	98

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	573	ARG
1	A	576	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
2	B	921	GLN

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](#)

Of 8 ligands modelled in this entry, 7 are monoatomic - leaving 1 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
16	F3S	D	301	4	0,9,9	-	-	-		

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '2' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
16	F3S	D	301	4	-	-	0/3/3/3

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
16	D	301	F3S	4	0

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

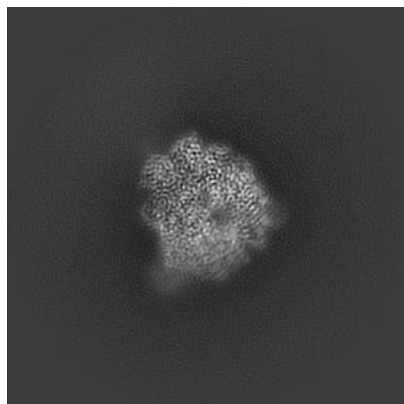
6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-12960. These allow visual inspection of the internal detail of the map and identification of artifacts.

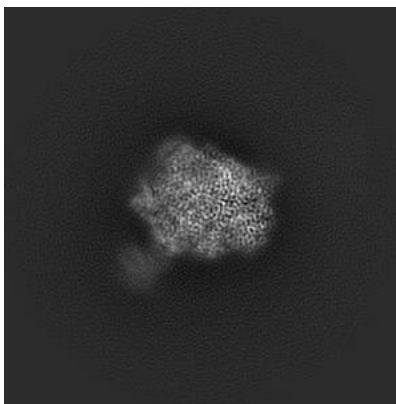
Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

6.1 Orthogonal projections [i](#)

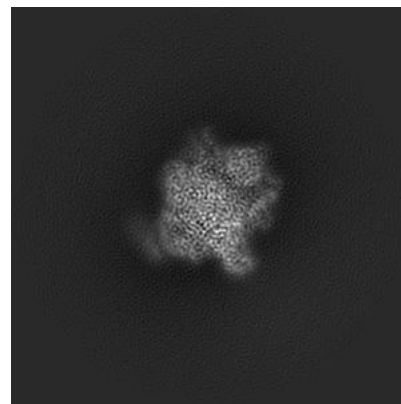
6.1.1 Primary map



X

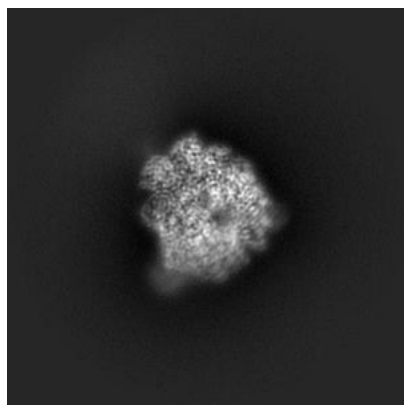


Y

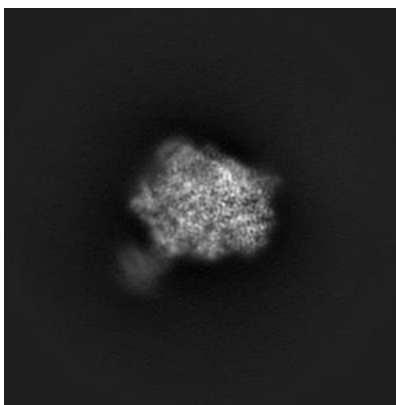


Z

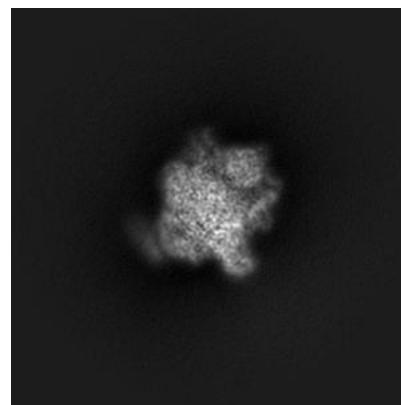
6.1.2 Raw map



X



Y

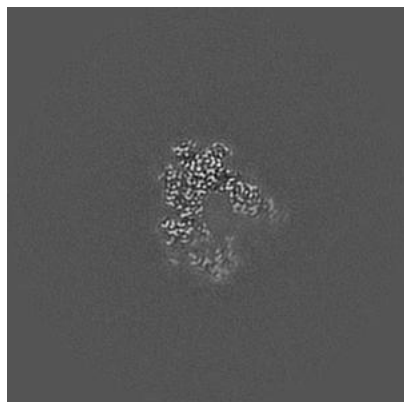


Z

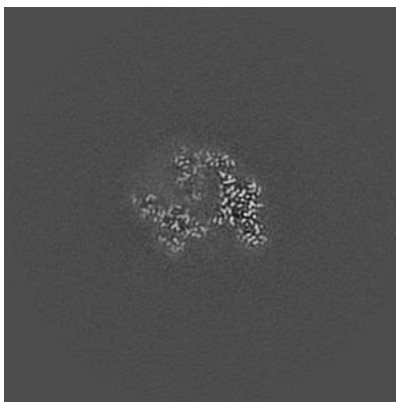
The images above show the map projected in three orthogonal directions.

6.2 Central slices [i](#)

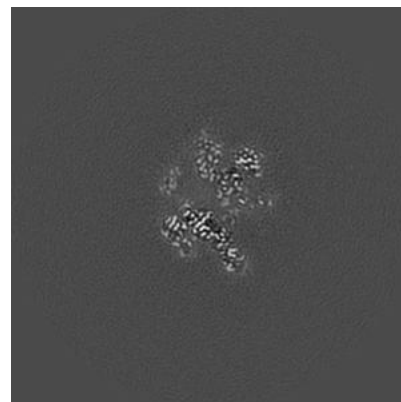
6.2.1 Primary map



X Index: 166

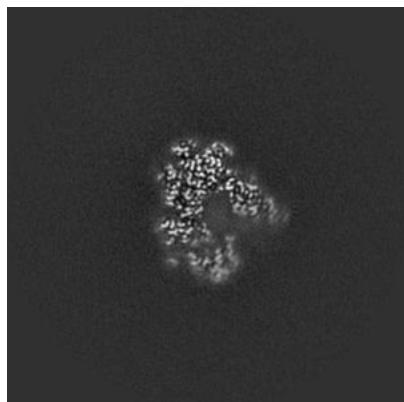


Y Index: 166

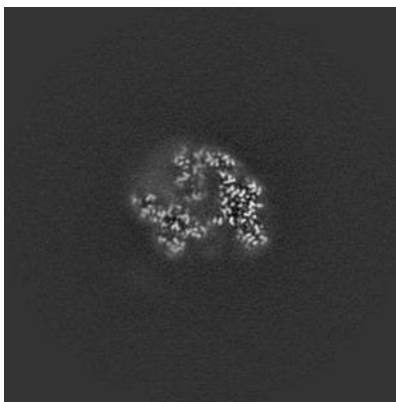


Z Index: 166

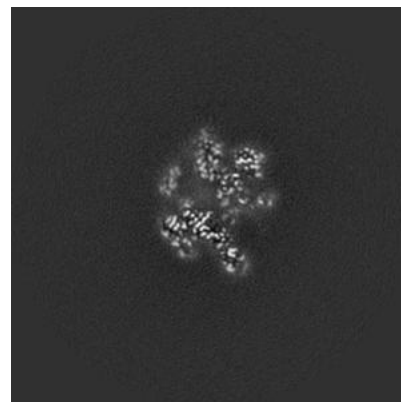
6.2.2 Raw map



X Index: 166



Y Index: 166

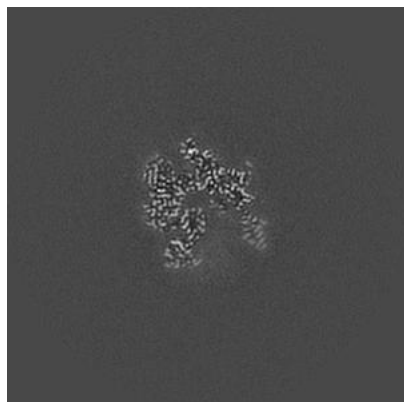


Z Index: 166

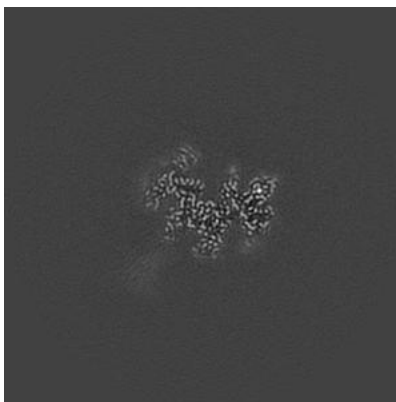
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices [i](#)

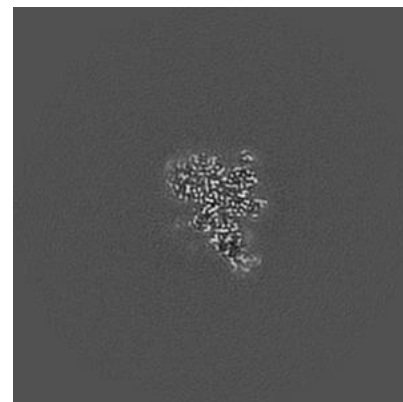
6.3.1 Primary map



X Index: 180

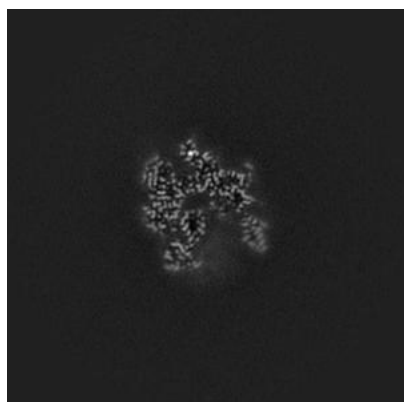


Y Index: 152

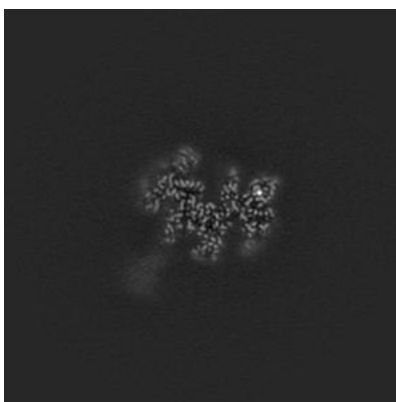


Z Index: 182

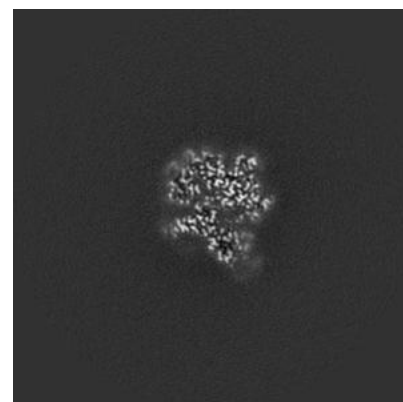
6.3.2 Raw map



X Index: 180



Y Index: 152

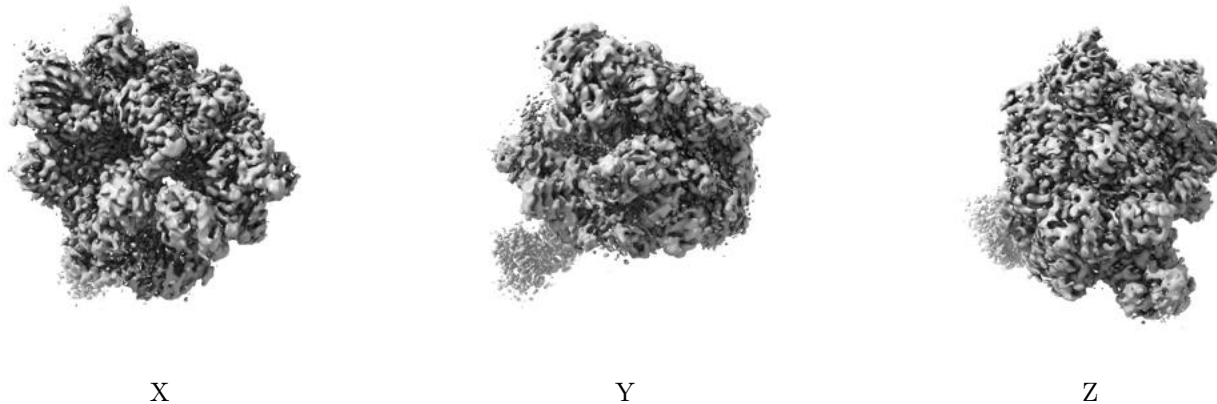


Z Index: 178

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

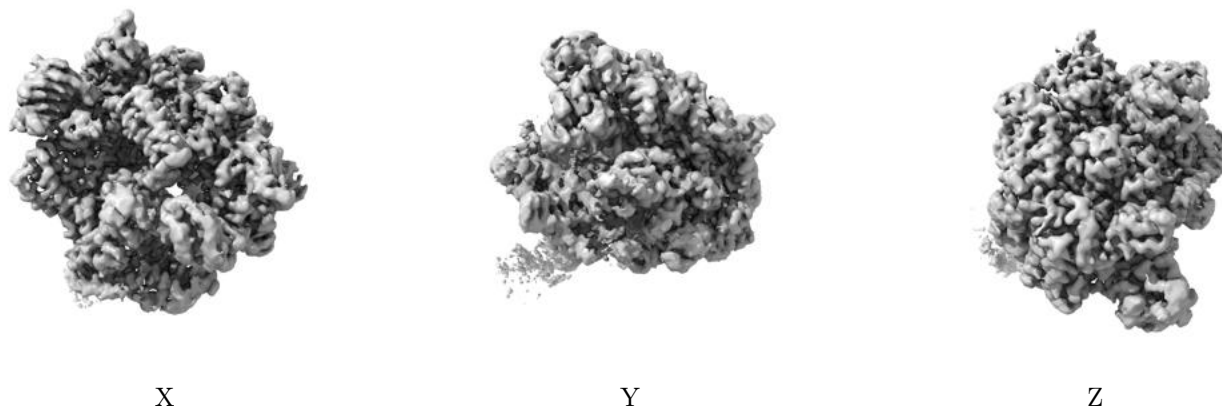
6.4 Orthogonal surface views [i](#)

6.4.1 Primary map



The images above show the 3D surface view of the map at the recommended contour level 4.0. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

6.4.2 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.

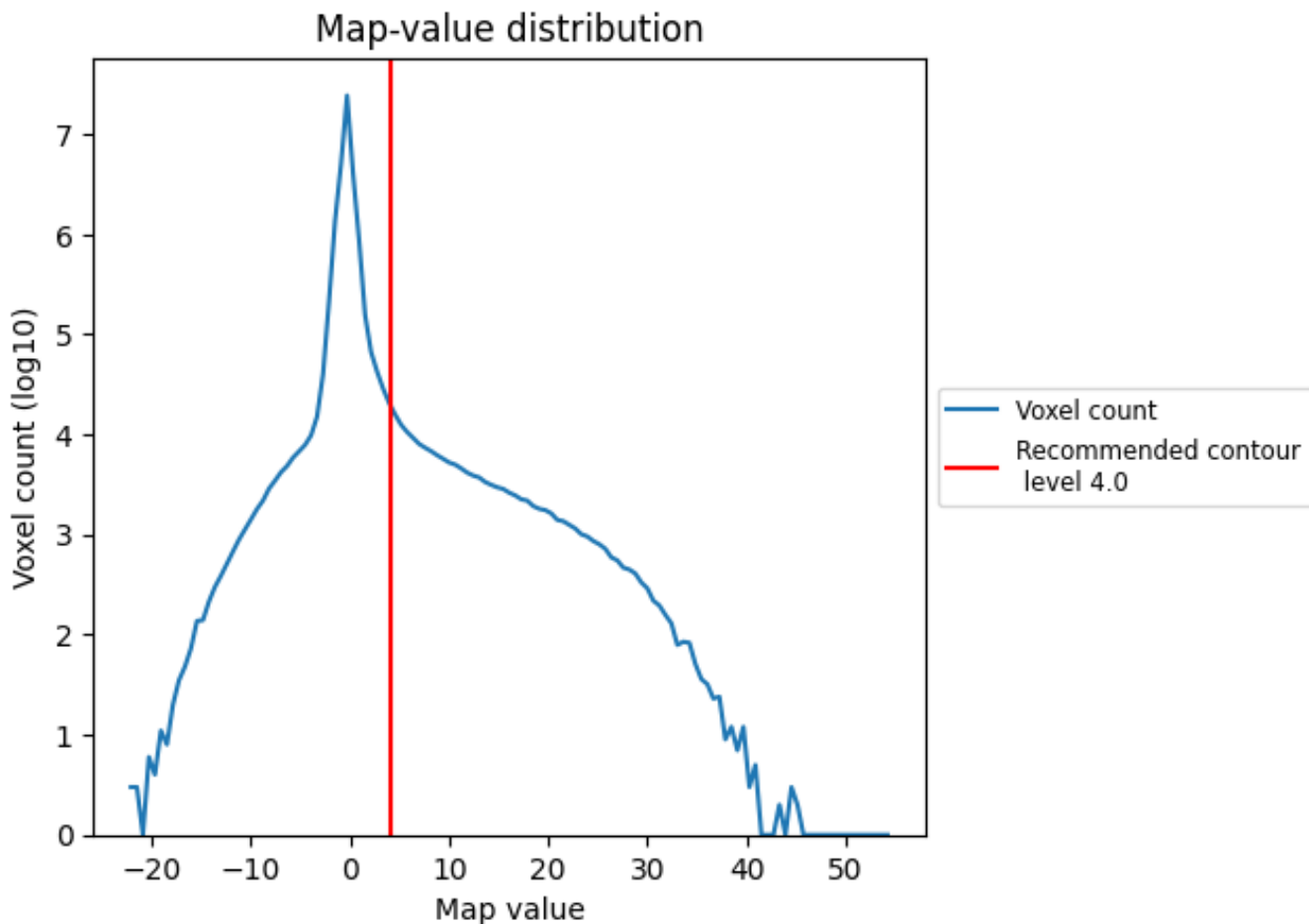
6.5 Mask visualisation [i](#)

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

7 Map analysis [i](#)

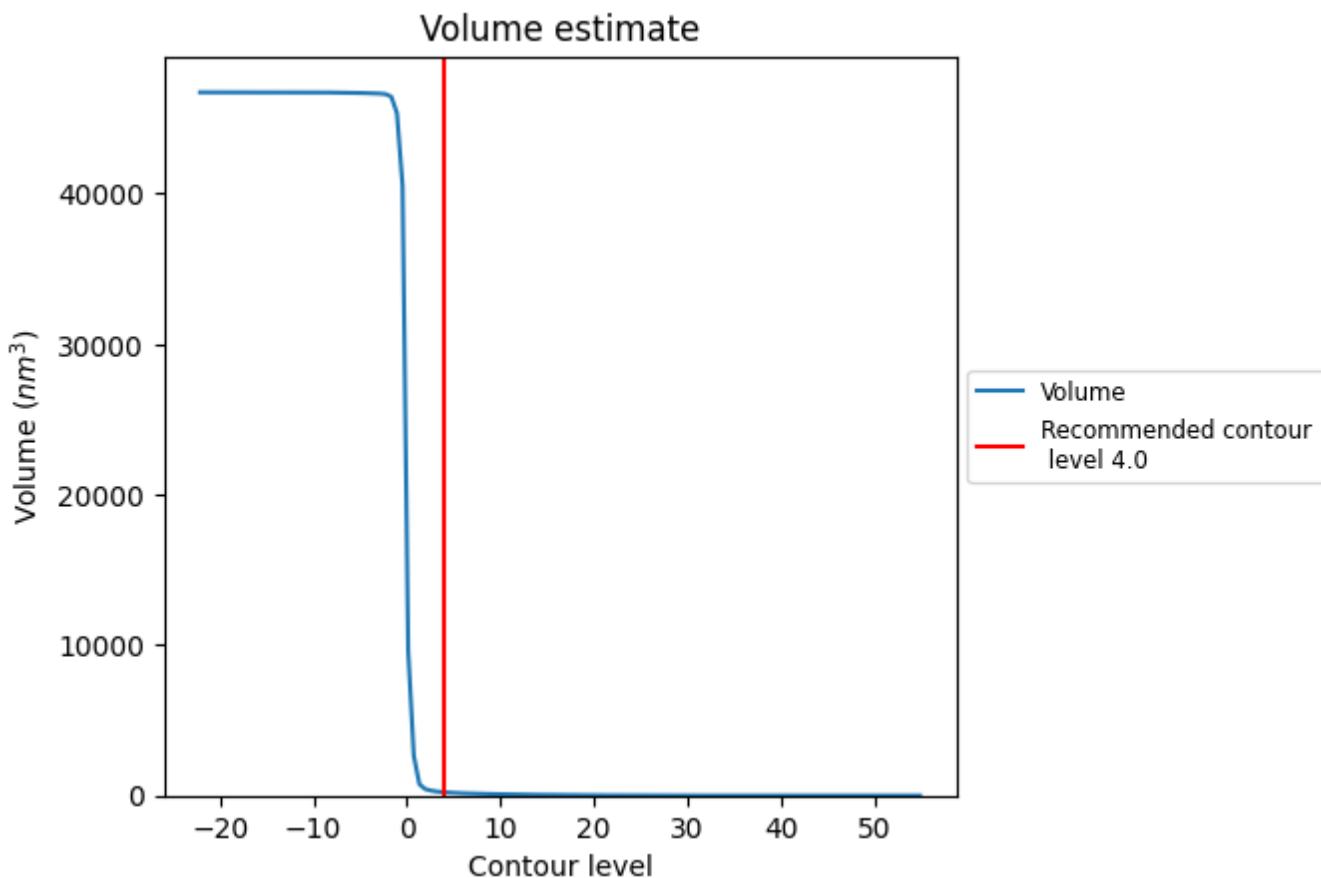
This section contains the results of statistical analysis of the map.

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.

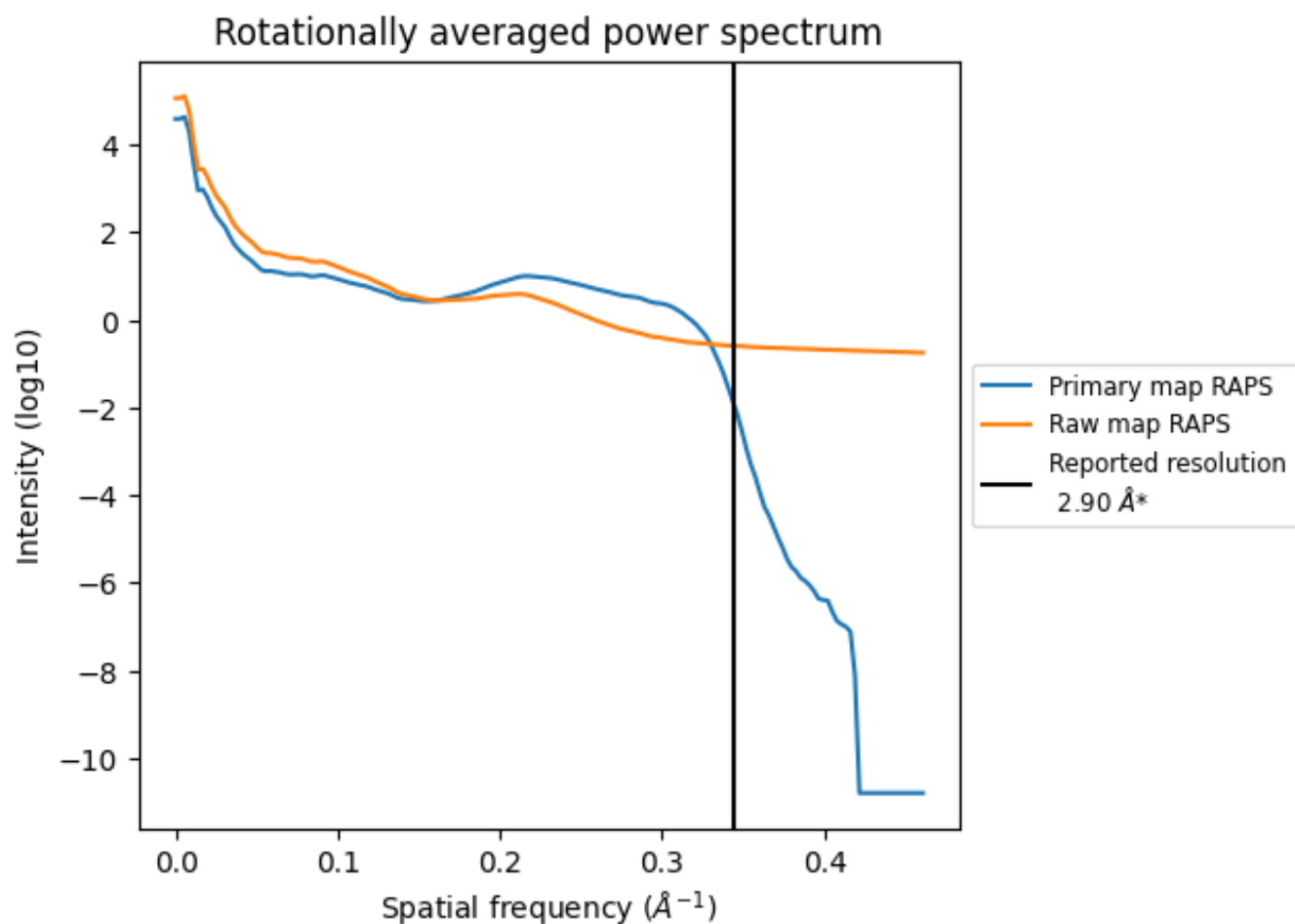
7.2 Volume estimate [i](#)



The volume at the recommended contour level is 219 nm^3 ; this corresponds to an approximate mass of 197 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

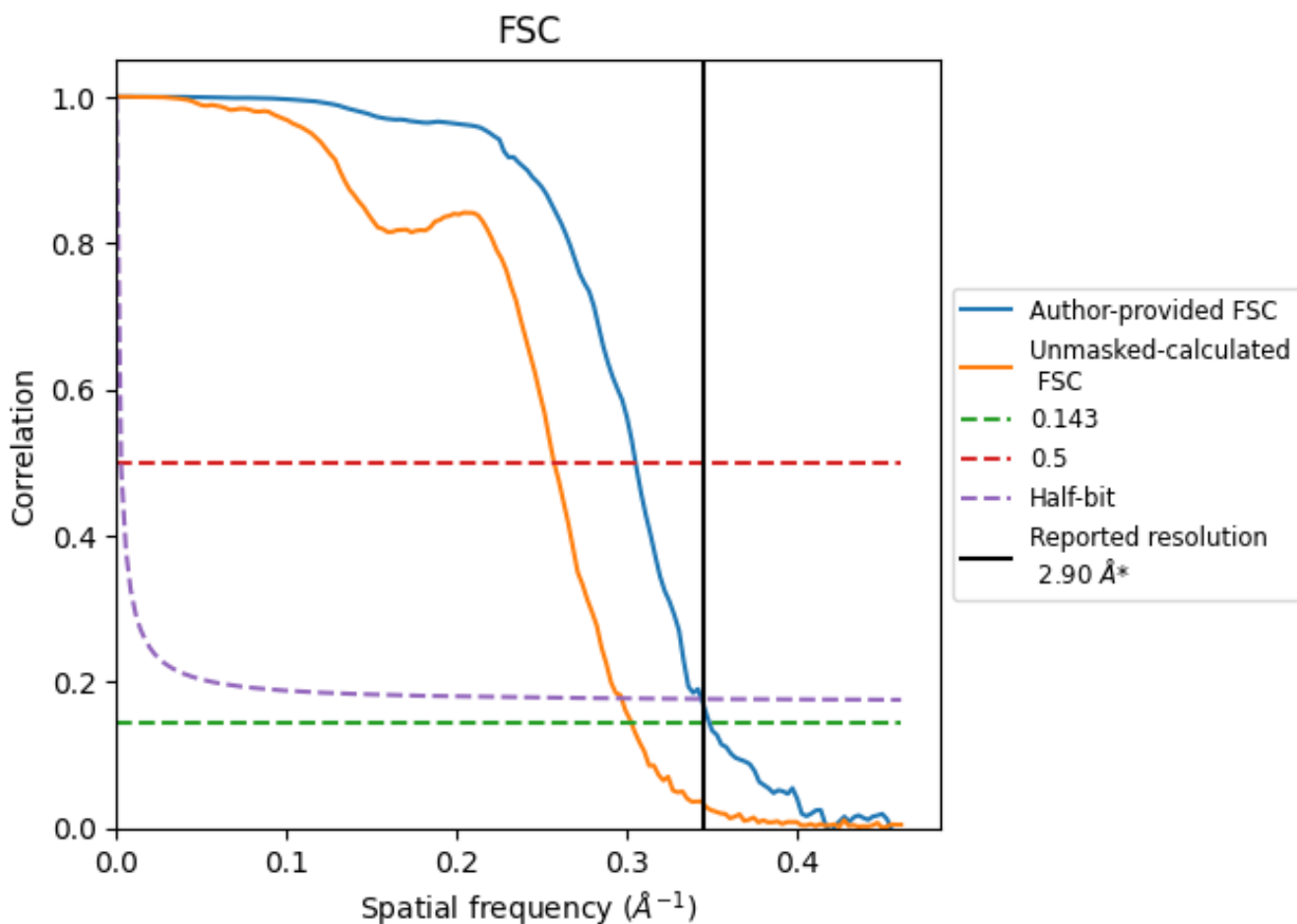


*Reported resolution corresponds to spatial frequency of 0.345 Å⁻¹

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 0.345 Å⁻¹

8.2 Resolution estimates [i](#)

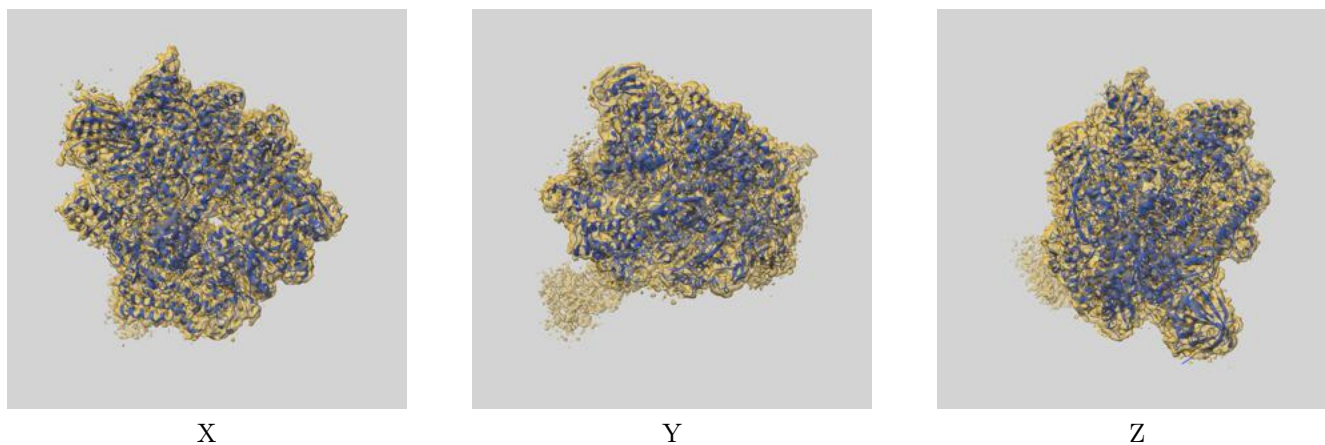
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	2.90	-	-
Author-provided FSC curve	2.87	3.28	2.91
Unmasked-calculated*	3.31	3.89	3.37

*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 3.31 differs from the reported value 2.9 by more than 10 %

9 Map-model fit [i](#)

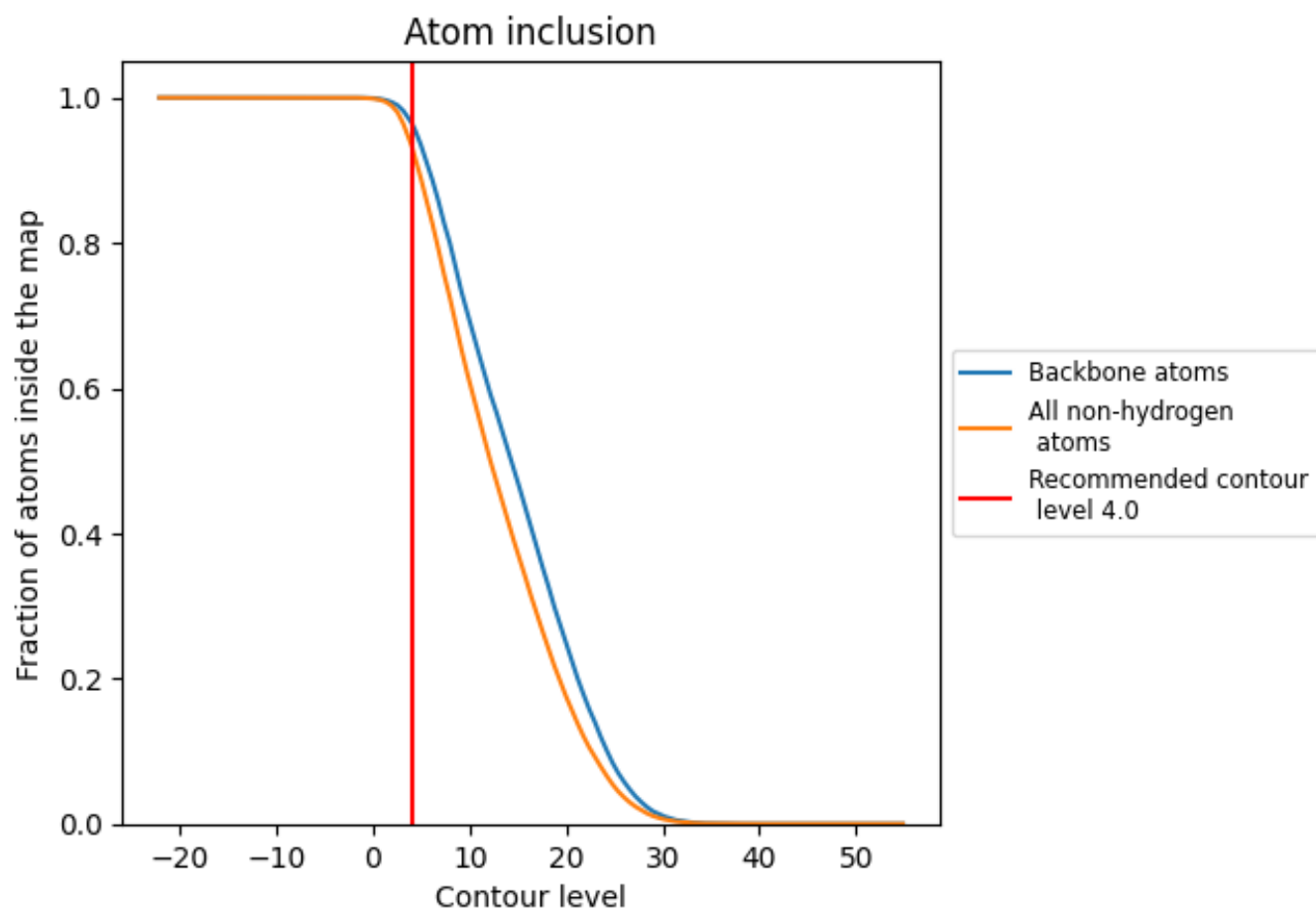
This section contains information regarding the fit between EMDB map EMD-12960 and PDB model 7OK0. Per-residue inclusion information can be found in section 3 on page 8.

9.1 Map-model overlay [i](#)



The images above show the 3D surface view of the map at the recommended contour level 4.0 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 Atom inclusion [i](#)



At the recommended contour level, 97% of all backbone atoms, 93% of all non-hydrogen atoms, are inside the map.