



wwPDB EM Validation Summary Report ⓘ

Dec 14, 2021 – 06:19 am GMT

PDB ID : 7OBB
EMDB ID : EMD-12797
Title : Cryo-EM structure of human RNA Polymerase I Open Complex
Authors : Misiaszek, A.D.; Girbig, M.; Mueller, C.W.
Deposited on : 2021-04-21
Resolution : 3.30 Å (reported)
Based on initial model : 7AEI

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
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) : 2.24

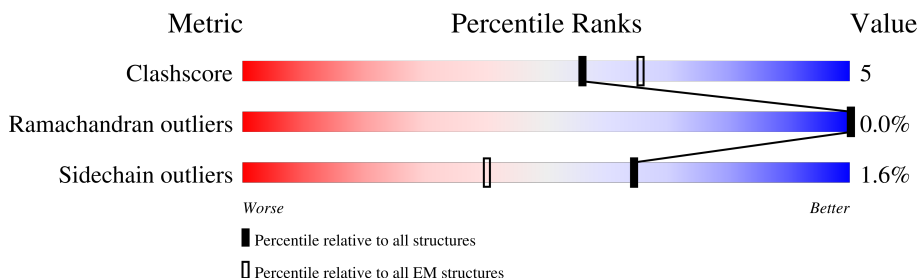
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 3.30 Å.

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









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

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments of the bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions $\leq 5\%$. 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	1720	
2	B	1135	
3	C	346	
4	E	210	
5	F	127	
6	G	338	
7	H	150	
8	I	126	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
9	J	67	 84% 12%
10	K	133	 70% 11% 20%
11	L	58	 57% 22% 21%
12	M	419	 22% 75%
13	N	510	 25% 71%
14	T	43	 9% 19% 79%
15	S	43	 19% 19% 81%

2 Entry composition

There are 16 unique types of molecules in this entry. The entry contains 33048 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 I subunit RPA1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	1510	12050	7664	2114	2192	80	0	0

- Molecule 2 is a protein called DNA-directed RNA polymerase I subunit RPA2.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	B	1130	8962	5740	1527	1622	73	0	0

- Molecule 3 is a protein called DNA-directed RNA polymerases I and III subunit RPAC1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	C	309	2474	1561	440	462	11	0	0

- Molecule 4 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	E	210	1728	1094	301	324	9	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	44	PHE	SER	conflict	UNP P19388

- Molecule 5 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC2.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
5	F	77	619	398	105	111	5	0	0

- Molecule 6 is a protein called DNA-directed RNA polymerase I subunit RPA43.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
6	G	153	1183	749	209	218	7	0	0

- Molecule 7 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC3.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
7	H	149	1197	759	195	238	5	0	0

- Molecule 8 is a protein called DNA-directed RNA polymerase I subunit RPA12.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
8	I	104	809	494	145	159	11	0	0

- Molecule 9 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC5.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
9	J	64	507	328	86	87	6	0	0

- Molecule 10 is a protein called DNA-directed RNA polymerases I and III subunit RPAC2.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
10	K	107	856	531	153	165	7	0	0

- Molecule 11 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC4.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
11	L	46	388	241	75	66	6	0	0

- Molecule 12 is a protein called DNA-directed RNA polymerase I subunit RPA49.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
12	M	105	837	522	154	155	6	0	0

- Molecule 13 is a protein called DNA-directed RNA polymerase I subunit RPA34.

Mol	Chain	Residues	Atoms					AltConf	Trace
13	N	149	Total	C	N	O	S	0	0
			1083	684	193	201	5		

- Molecule 14 is a DNA chain called DNA template strand.

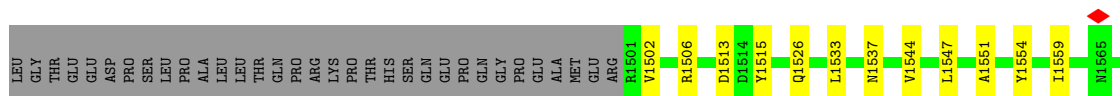
Mol	Chain	Residues	Atoms					AltConf	Trace
14	T	9	Total	C	N	O	P	0	0
			185	89	34	53	9		

- Molecule 15 is a DNA chain called DNA non-template strand.

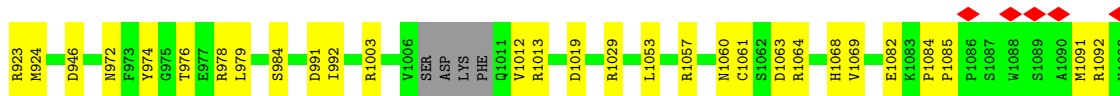
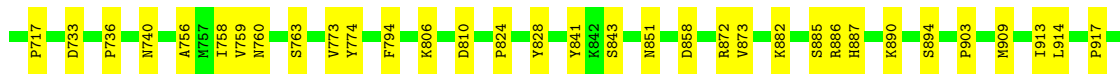
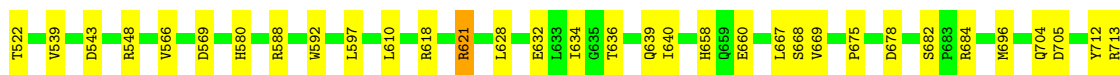
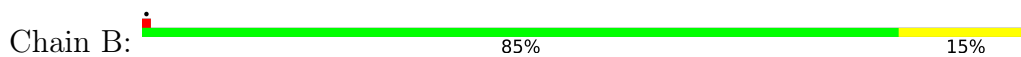
Mol	Chain	Residues	Atoms					AltConf	Trace
15	S	8	Total	C	N	O	P	0	0
			163	79	26	50	8		

- Molecule 16 is ZINC ION (three-letter code: ZN) (formula: Zn).

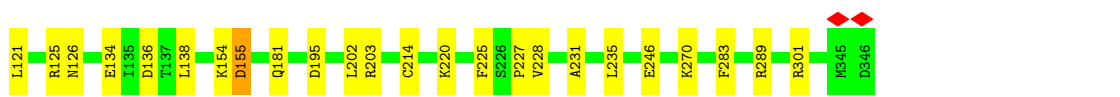
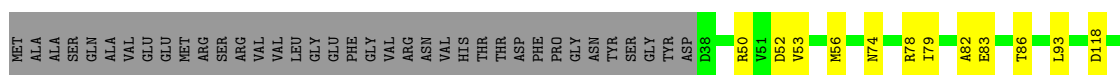
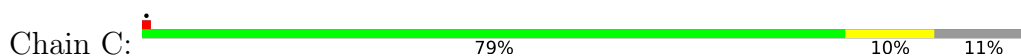
Mol	Chain	Residues	Atoms		AltConf
16	A	2	Total	Zn	0
			2	2	
16	B	1	Total	Zn	0
			1	1	
16	I	2	Total	Zn	0
			2	2	
16	J	1	Total	Zn	0
			1	1	
16	L	1	Total	Zn	0
			1	1	



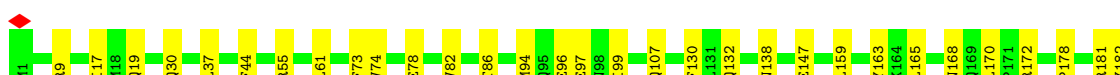
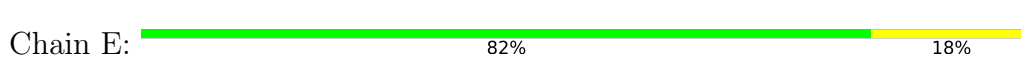
• Molecule 2: DNA-directed RNA polymerase I subunit RPA2



• Molecule 3: DNA-directed RNA polymerases I and III subunit RPAC1



• Molecule 4: DNA-directed RNA polymerases I, II, and III subunit RPABC1



4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	175912	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{\AA}^2$)	41.2	Depositor
Minimum defocus (nm)	750	Depositor
Maximum defocus (nm)	2250	Depositor
Magnification	105000	Depositor
Image detector	GATAN K3 (6k x 4k)	Depositor
Maximum map value	0.026	Depositor
Minimum map value	-0.009	Depositor
Average map value	0.000	Depositor
Map value standard deviation	0.002	Depositor
Recommended contour level	0.00386	Depositor
Map size (Å)	236.73601, 236.73601, 236.73601	wwPDB
Map dimensions	288, 288, 288	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	0.822, 0.822, 0.822	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:
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	0.42	0/12296	0.64	3/16600 (0.0%)
2	B	0.53	0/9181	0.64	5/12426 (0.0%)
3	C	0.47	0/2522	0.66	2/3418 (0.1%)
4	E	0.36	0/1759	0.63	2/2376 (0.1%)
5	F	0.36	0/629	0.60	0/850
6	G	0.30	0/1207	0.66	0/1634
7	H	0.45	0/1219	0.61	1/1644 (0.1%)
8	I	0.42	0/823	0.60	0/1107
9	J	0.60	0/516	0.69	0/696
10	K	0.48	0/871	0.62	0/1174
11	L	0.51	0/394	0.59	0/524
12	M	0.33	0/852	0.57	0/1146
13	N	0.34	0/1118	0.58	0/1532
14	T	0.61	0/207	0.98	0/317
15	S	0.57	0/181	1.08	0/277
All	All	0.46	0/33775	0.64	13/45721 (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	3
2	B	0	1
6	G	0	1
13	N	0	2
All	All	0	7

There are no bond length outliers.

The worst 5 of 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	195	ASP	CB-CG-OD1	6.38	124.04	118.30
2	B	16	LEU	CA-CB-CG	5.82	128.67	115.30
4	E	37	LEU	CA-CB-CG	5.76	128.56	115.30
1	A	663	ASP	CB-CG-OD1	5.61	123.35	118.30
2	B	426	LEU	CA-CB-CG	5.59	128.17	115.30

There are no chirality outliers.

5 of 7 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	387	LEU	Mainchain
1	A	458	ILE	Mainchain
1	A	491	HIS	Peptide
2	B	457	ALA	Peptide
6	G	86	LEU	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	12050	0	12184	152	0
2	B	8962	0	8944	101	0
3	C	2474	0	2469	23	0
4	E	1728	0	1749	25	0
5	F	619	0	655	9	0
6	G	1183	0	1172	14	0
7	H	1197	0	1156	18	0
8	I	809	0	752	12	0
9	J	507	0	524	7	0
10	K	856	0	840	11	0
11	L	388	0	393	6	0
12	M	837	0	819	13	0
13	N	1083	0	1079	10	0
14	T	185	0	103	1	0
15	S	163	0	93	0	0
16	A	2	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
16	B	1	0	0	0	0
16	I	2	0	0	0	0
16	J	1	0	0	0	0
16	L	1	0	0	0	0
All	All	33048	0	32932	343	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 343 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:463:ALA:HB2	1:A:568:ILE:HD12	1.65	0.77
1:A:1353:SER:OG	1:A:1547:LEU:HD21	1.84	0.76
2:B:312:LEU:HD21	2:B:327:LEU:HB3	1.72	0.72
1:A:468:TYR:N	1:A:540:ARG:O	2.24	0.69
1:A:1559:ILE:HA	1:A:1582:GLY:HA3	1.76	0.67

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	1490/1720 (87%)	1363 (92%)	127 (8%)	0	100	100
2	B	1126/1135 (99%)	1045 (93%)	81 (7%)	0	100	100
3	C	307/346 (89%)	290 (94%)	16 (5%)	1 (0%)	41	71
4	E	208/210 (99%)	191 (92%)	17 (8%)	0	100	100
5	F	75/127 (59%)	73 (97%)	2 (3%)	0	100	100
6	G	149/338 (44%)	126 (85%)	23 (15%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
7	H	147/150 (98%)	137 (93%)	10 (7%)	0	100	100
8	I	100/126 (79%)	94 (94%)	6 (6%)	0	100	100
9	J	62/67 (92%)	58 (94%)	4 (6%)	0	100	100
10	K	105/133 (79%)	98 (93%)	7 (7%)	0	100	100
11	L	44/58 (76%)	39 (89%)	5 (11%)	0	100	100
12	M	101/419 (24%)	92 (91%)	9 (9%)	0	100	100
13	N	147/510 (29%)	139 (95%)	8 (5%)	0	100	100
All	All	4061/5339 (76%)	3745 (92%)	315 (8%)	1 (0%)	100	100

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	C	155	ASP

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	1331/1504 (88%)	1296 (97%)	35 (3%)	46	71
2	B	987/992 (100%)	977 (99%)	10 (1%)	76	86
3	C	272/302 (90%)	271 (100%)	1 (0%)	91	95
4	E	192/192 (100%)	192 (100%)	0	100	100
5	F	67/111 (60%)	66 (98%)	1 (2%)	65	81
6	G	128/288 (44%)	127 (99%)	1 (1%)	81	89
7	H	130/131 (99%)	129 (99%)	1 (1%)	81	89
8	I	93/111 (84%)	92 (99%)	1 (1%)	73	85
9	J	53/56 (95%)	53 (100%)	0	100	100
10	K	96/119 (81%)	96 (100%)	0	100	100
11	L	43/55 (78%)	39 (91%)	4 (9%)	9	30

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
12	M	91/366 (25%)	91 (100%)	0	100	100
13	N	118/427 (28%)	116 (98%)	2 (2%)	60	78
All	All	3601/4654 (77%)	3545 (98%)	56 (2%)	64	79

5 of 56 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1101	LYS
13	N	72	LYS
2	B	313	ASN
13	N	21	LYS
8	I	103	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 49 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	643	ASN
3	C	160	ASN
2	B	715	GLN
2	B	918	HIS
3	C	305	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](#)

Of 7 ligands modelled in this entry, 7 are monoatomic - leaving 0 for Mogul analysis.

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.

No monomer is involved in short contacts.

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.

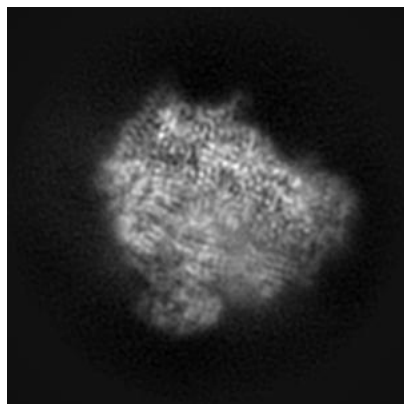
6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-12797. These allow visual inspection of the internal detail of the map and identification of artifacts.

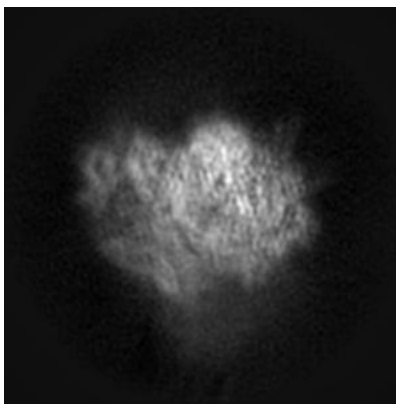
Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

6.1 Orthogonal projections [i](#)

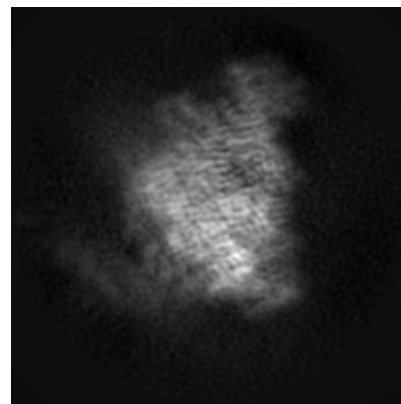
6.1.1 Primary map



X

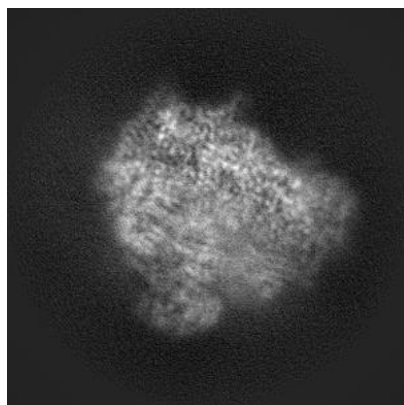


Y

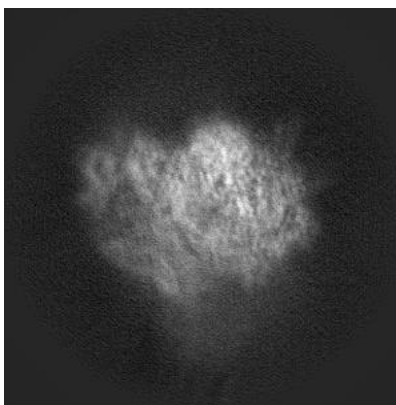


Z

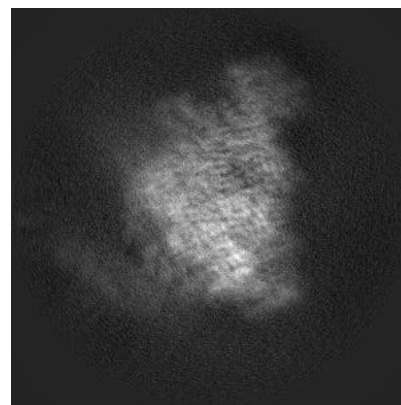
6.1.2 Raw map



X



Y

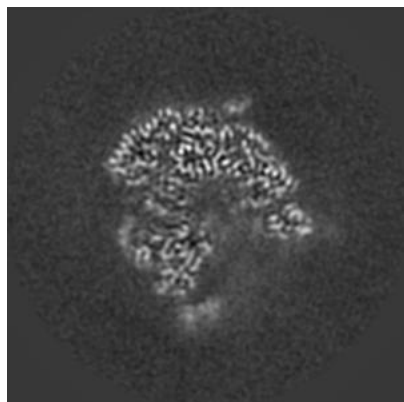


Z

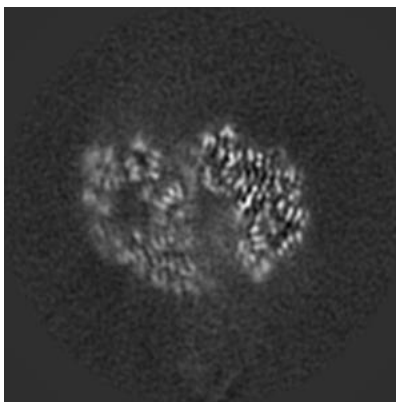
The images above show the map projected in three orthogonal directions.

6.2 Central slices [i](#)

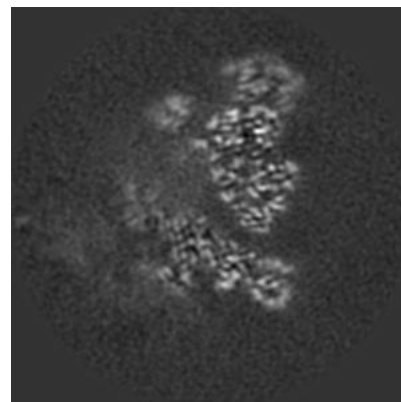
6.2.1 Primary map



X Index: 144

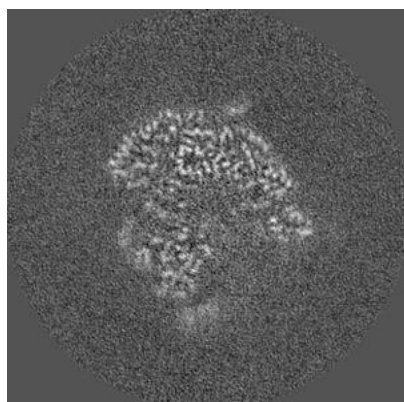


Y Index: 144

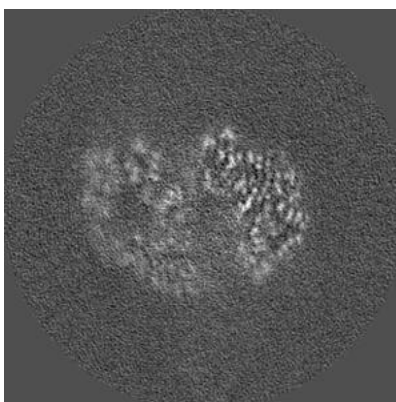


Z Index: 144

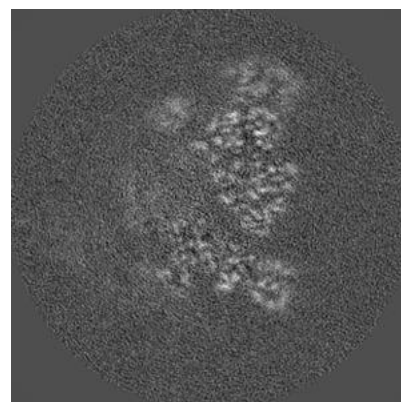
6.2.2 Raw map



X Index: 144



Y Index: 144

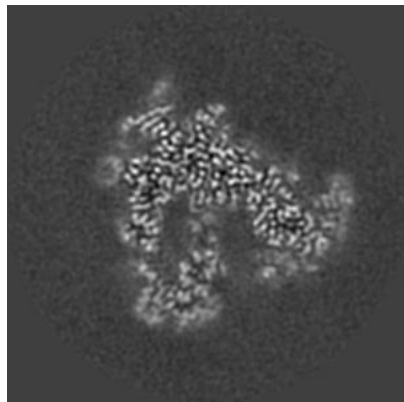


Z Index: 144

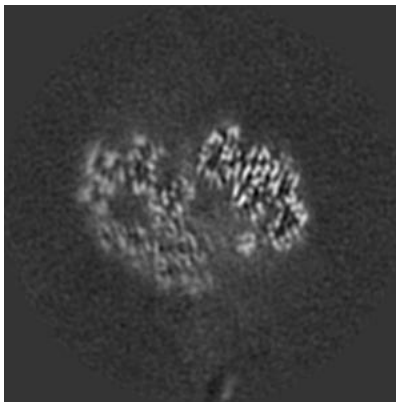
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices [i](#)

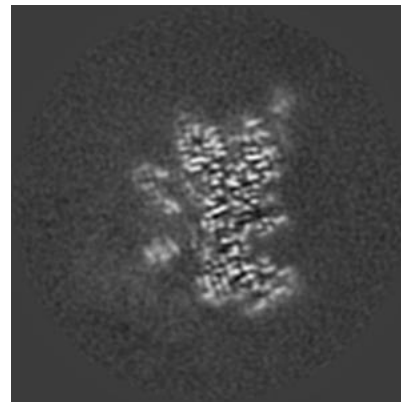
6.3.1 Primary map



X Index: 165

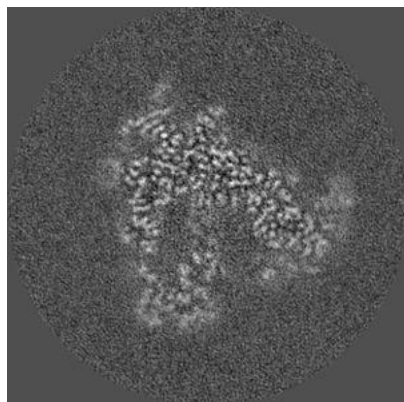


Y Index: 139

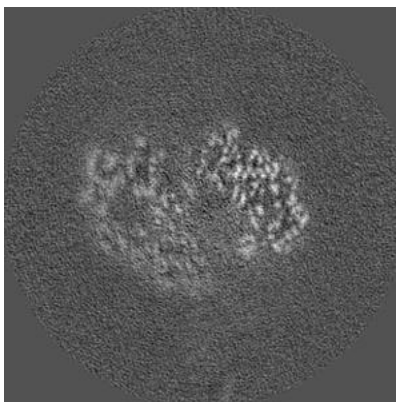


Z Index: 171

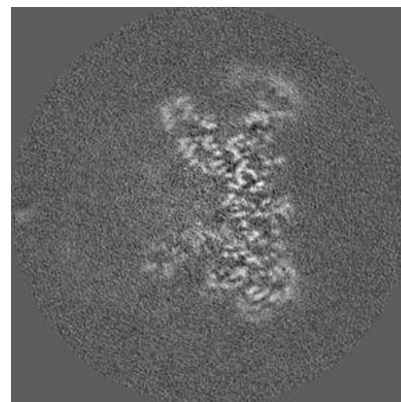
6.3.2 Raw map



X Index: 165



Y Index: 140

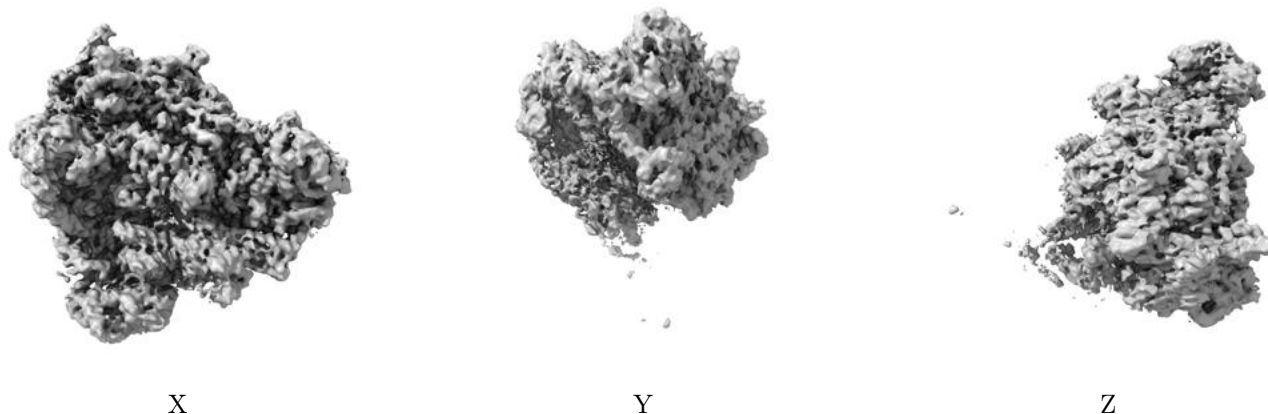


Z Index: 161

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

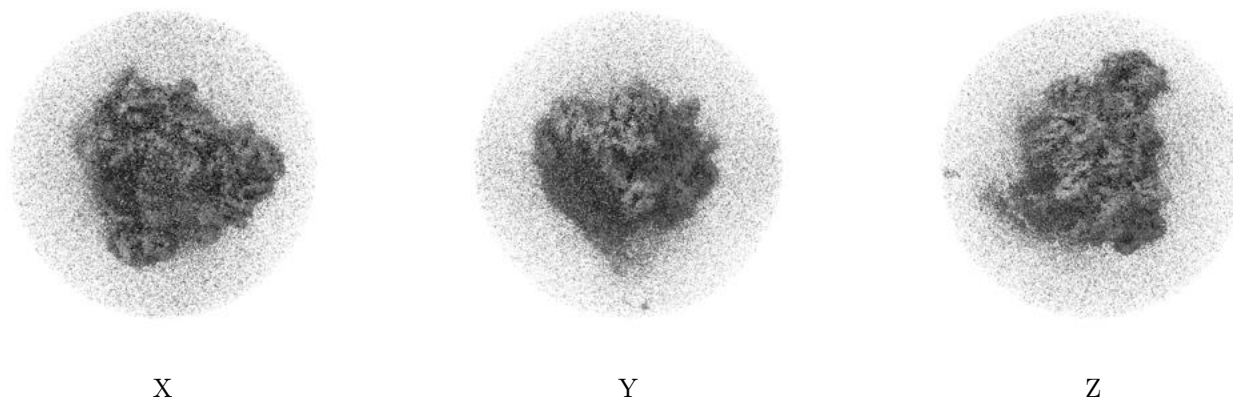
6.4 Orthogonal surface views [i](#)

6.4.1 Primary map



The images above show the 3D surface view of the map at the recommended contour level 0.00386. 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.

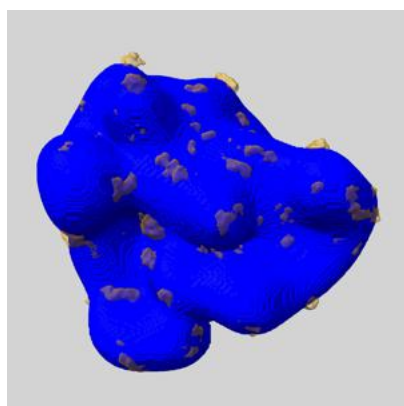
6.5 Mask visualisation [i](#)

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

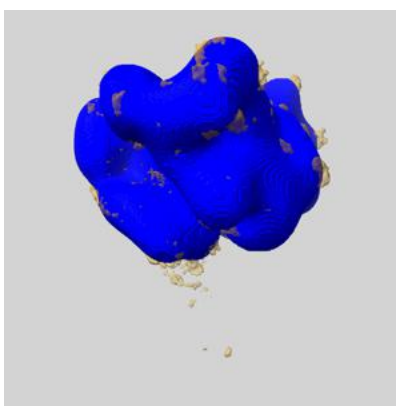
A mask typically either:

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

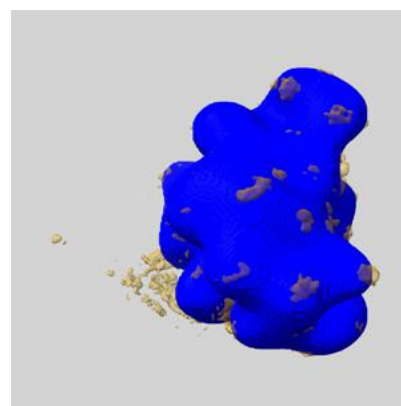
6.5.1 emd_12797_msk_1.map [i](#)



X



Y

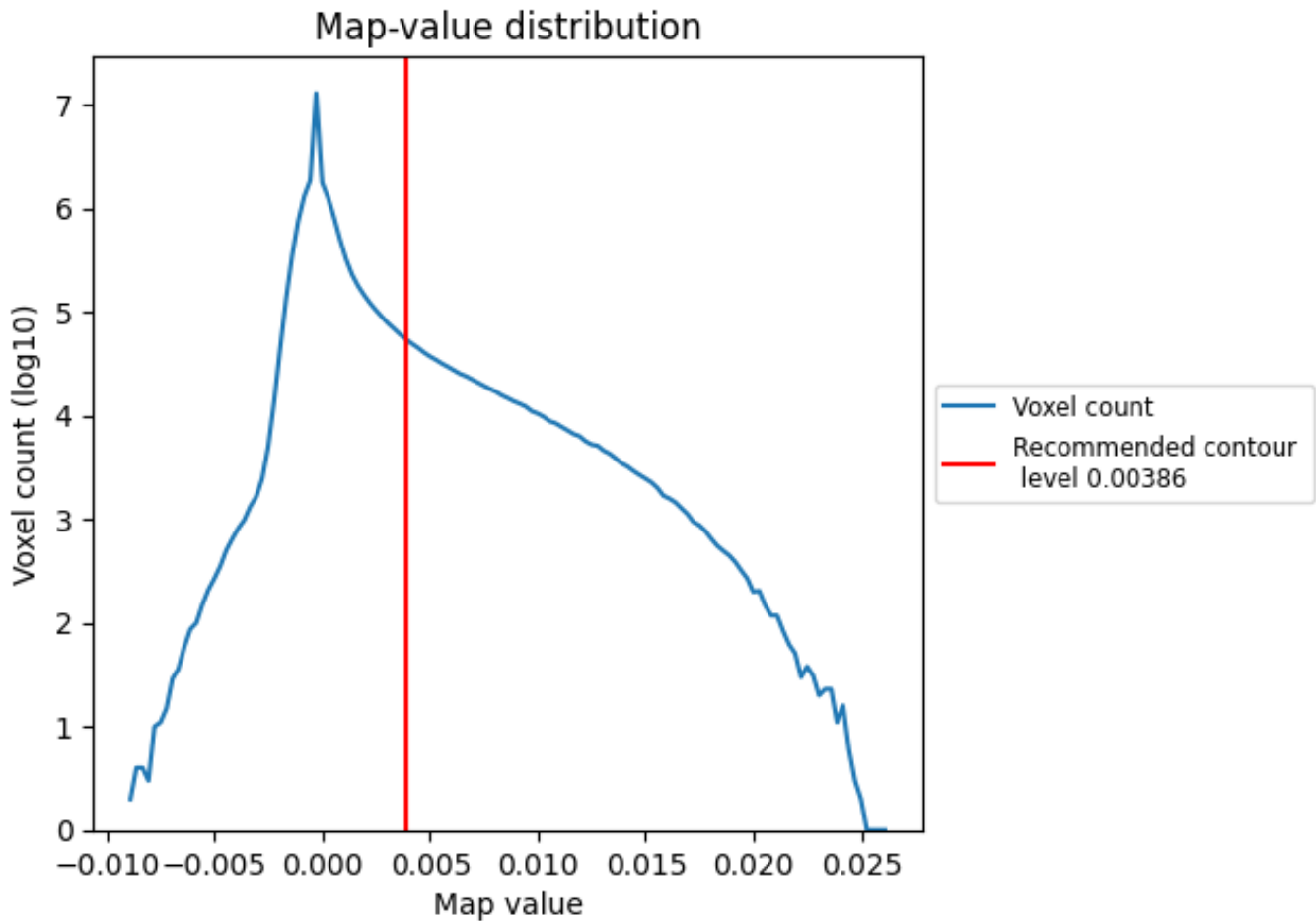


Z

7 Map analysis [i](#)

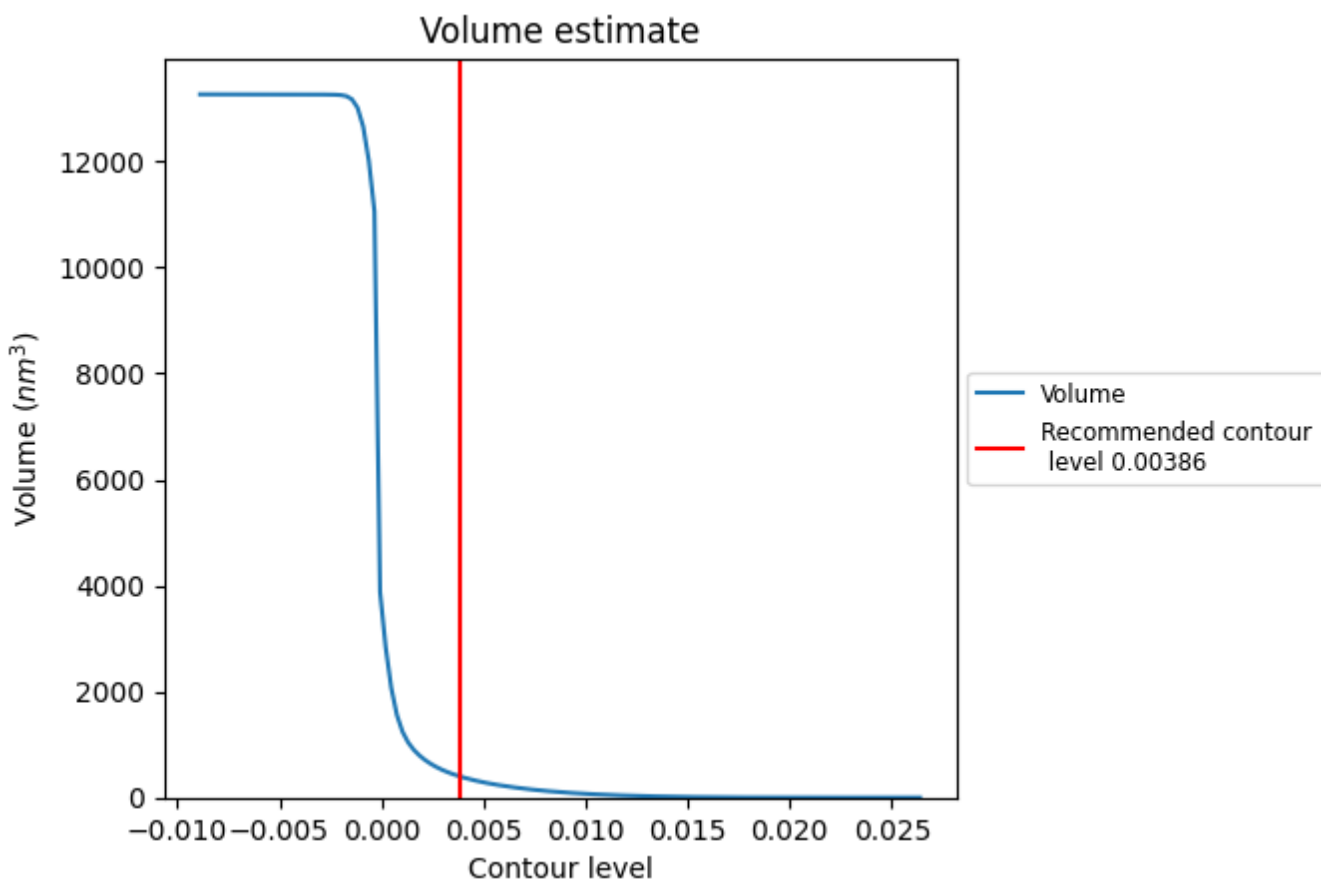
This section contains the results of statistical analysis of the map.

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.

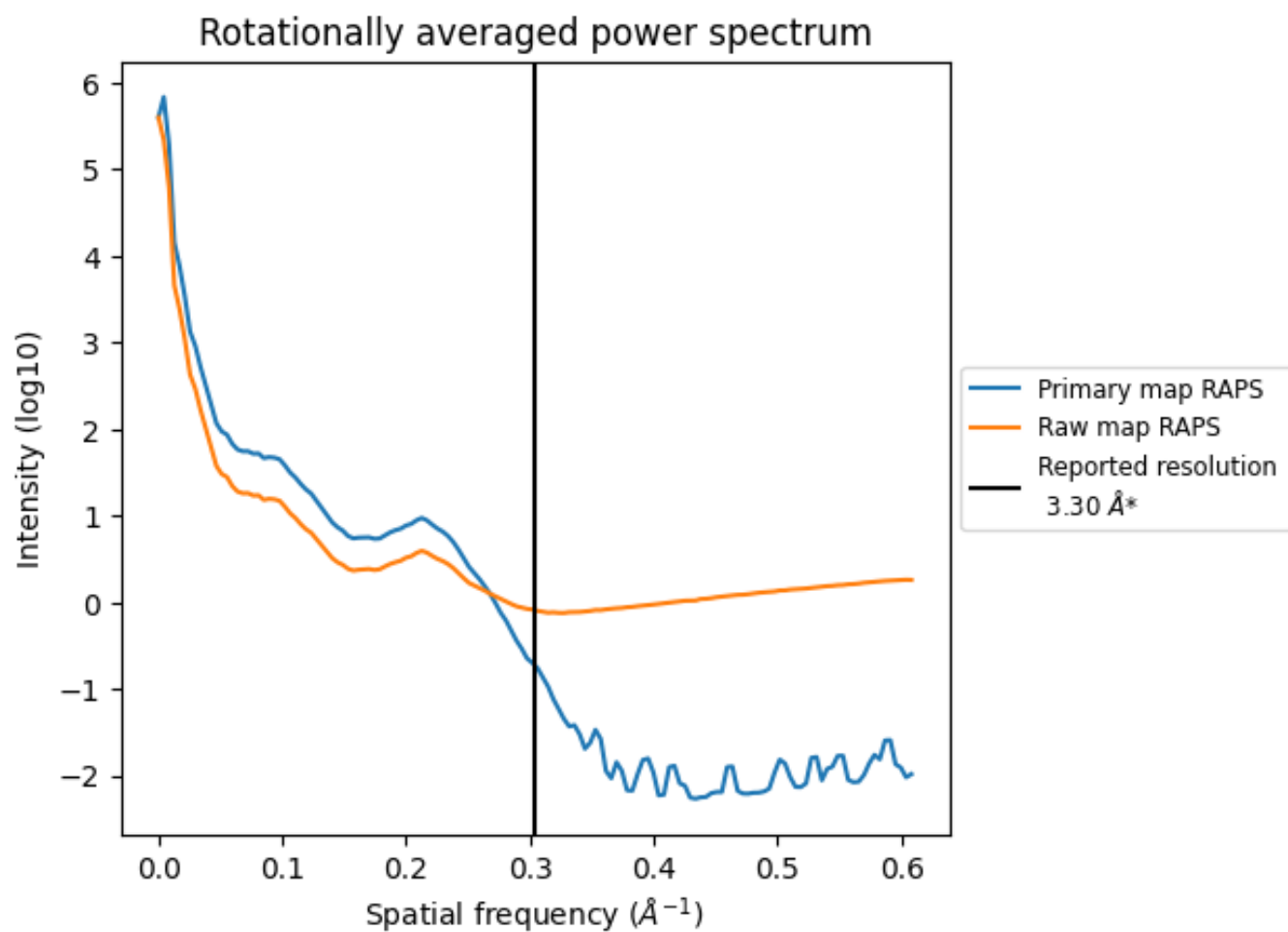
7.2 Volume estimate [i](#)



The volume at the recommended contour level is 397 nm^3 ; this corresponds to an approximate mass of 359 kDa.

The volume estimate graph shows how the enclosed volume varies with the contour level. The recommended contour level is shown as a vertical line and the intersection between the line and the curve gives the volume of the enclosed surface at the given level.

7.3 Rotationally averaged power spectrum i

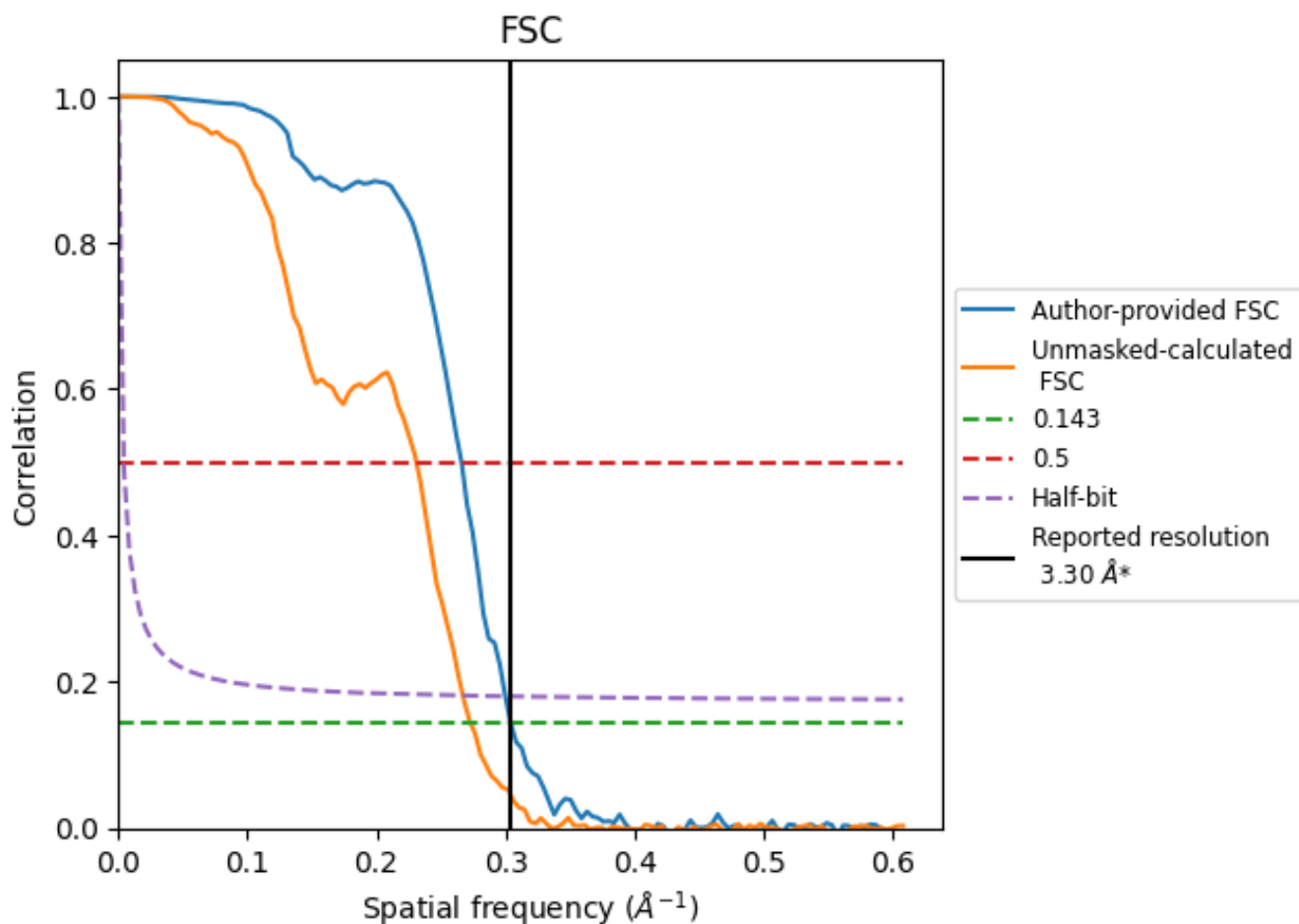


*Reported resolution corresponds to spatial frequency of 0.303 Å⁻¹

8 Fourier-Shell correlation [i](#)

Fourier-Shell Correlation (FSC) is the most commonly used method to estimate the resolution of single-particle and subtomogram-averaged maps. The shape of the curve depends on the imposed symmetry, mask and whether or not the two 3D reconstructions used were processed from a common reference. The reported resolution is shown as a black line. A curve is displayed for the half-bit criterion in addition to lines showing the 0.143 gold standard cut-off and 0.5 cut-off.

8.1 FSC [i](#)



*Reported resolution corresponds to spatial frequency of 0.303 Å⁻¹

8.2 Resolution estimates [i](#)

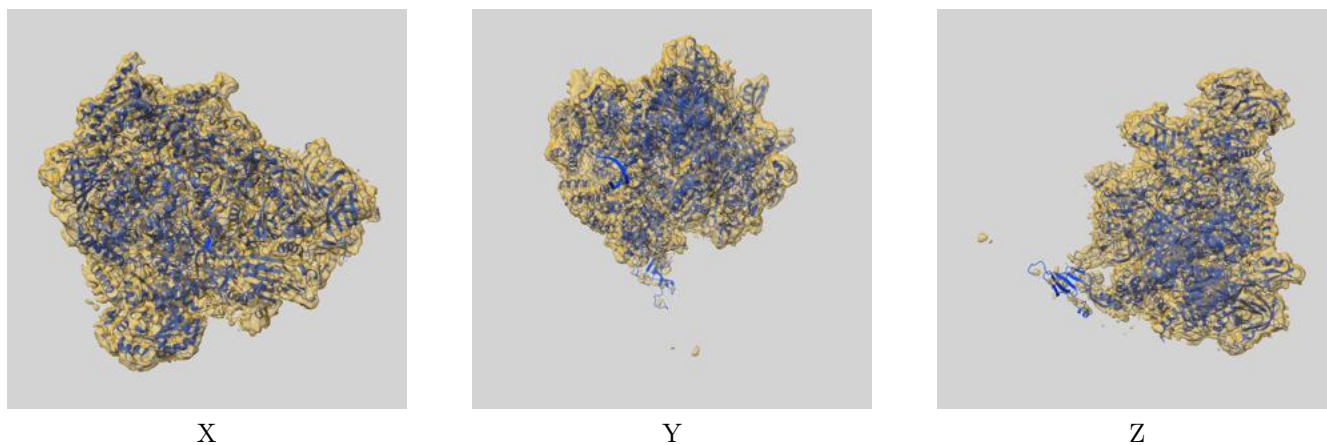
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	3.30	-	-
Author-provided FSC curve	3.29	3.76	3.33
Unmasked-calculated*	3.66	4.33	3.75

*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 3.66 differs from the reported value 3.3 by more than 10 %

9 Map-model fit [i](#)

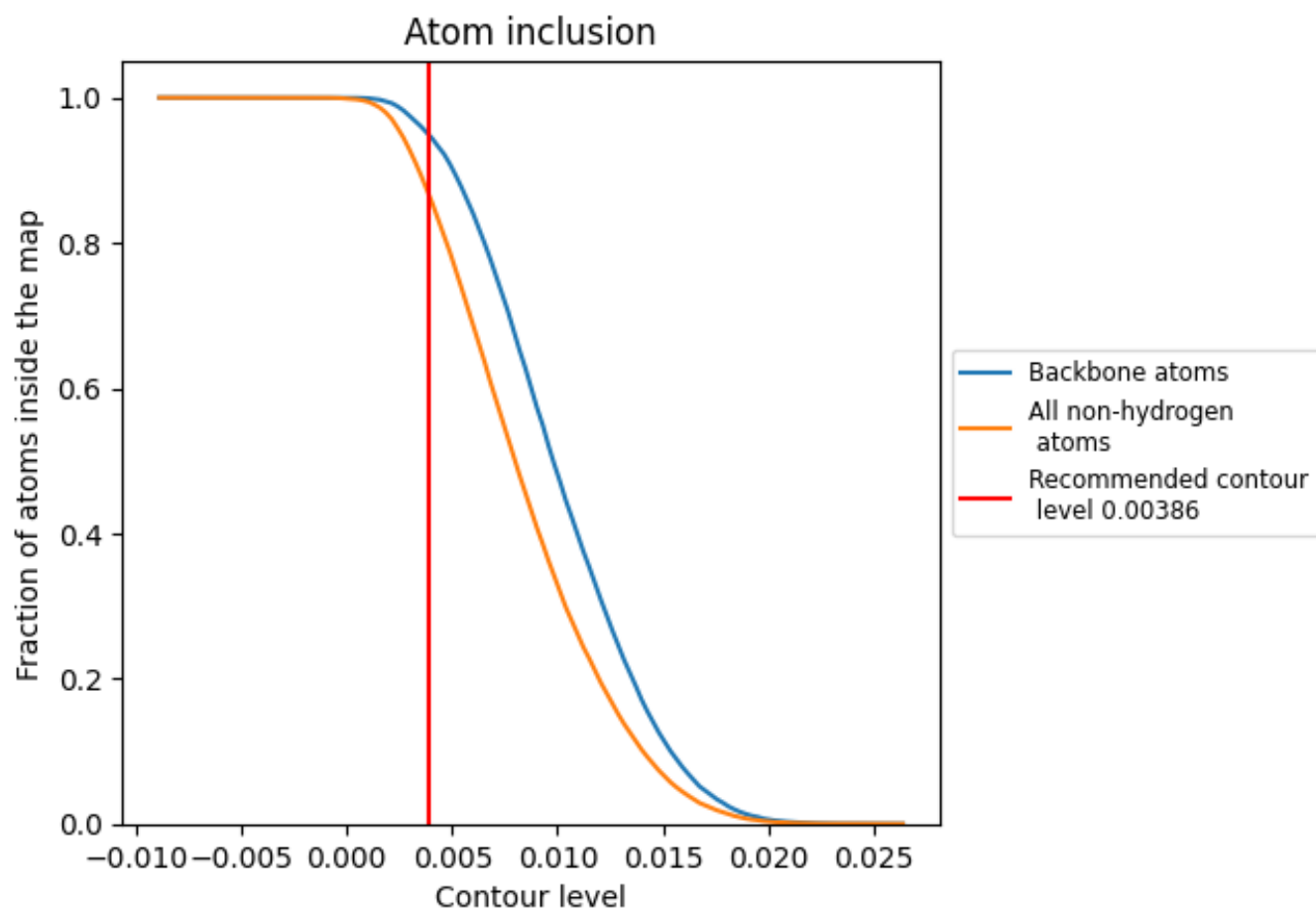
This section contains information regarding the fit between EMDB map EMD-12797 and PDB model 7OBB. Per-residue inclusion information can be found in section 3 on page 7.

9.1 Map-model overlay [i](#)



The images above show the 3D surface view of the map at the recommended contour level 0.00386 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, 95% of all backbone atoms, 87% of all non-hydrogen atoms, are inside the map.