wwPDB EM Validation Summary Report (i)

Jun 10, 2024 – 12:23 PM EDT

PDB ID	:	8GHU
EMDB ID	:	EMD-40051
Title	:	Methyltransferase RmtC bound to the 30S ribosomal subunit
Authors	:	Srinivas, P.; Conn, G.L.; Dunham, C.M.
Deposited on	:	2023-03-12
Resolution	:	3.00 Å(reported)

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 3.00 Å.

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

Metric	Percentile Ranks	Value
Clashscore		8
Ramachandran outliers		0.1%
Sidechain outliers		4.4%
RNA backbone		0.57

Worse   *Better*

▮ Percentile relative to all structures

▯ Percentile relative to all EM structures
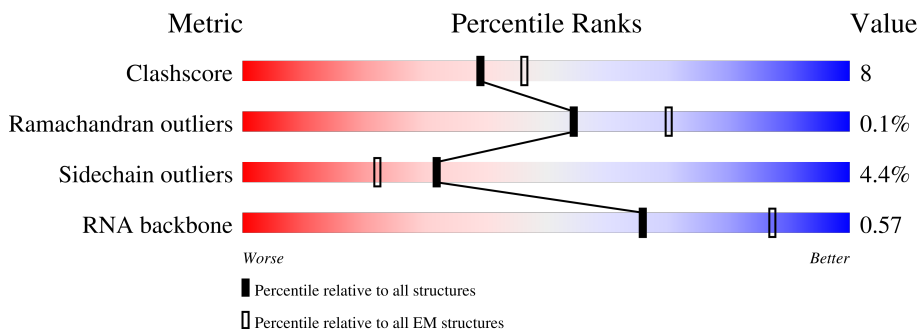
Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826
RNA backbone	4643	859

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	281	48% 48%
2	a	1532	79% 21%
3	c	206	98%
4	d	205	98%
5	e	150	98%
6	f	100	97%
7	g	151	97%

Continued from previous page...

Mol	Chain	Length	Quality of chain
8	h	129	98%
9	l	123	98%
10	m	93	99%
11	o	88	98%
12	p	82	95% · 5%
13	q	80	98%
14	r	55	7% · 100%
15	t	85	98%

2 Entry composition ⓘ

There are 16 unique types of molecules in this entry. The entry contains 47408 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 16S rRNA (guanine(1405)-N(7))-methyltransferase.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	281	Total	C	N	O	S	0	0
			2260	1445	373	433	9		

- Molecule 2 is a RNA chain called 16S rRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	a	1532	Total	C	N	O	P	0	0
			32900	14678	6040	10651	1531		

- Molecule 3 is a protein called 30S ribosomal protein S3.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	c	206	Total	C	N	O	S	0	0
			1625	1028	305	289	3		

- Molecule 4 is a protein called 30S ribosomal protein S4.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	d	205	Total	C	N	O	S	0	0
			1643	1026	315	298	4		

- Molecule 5 is a protein called 30S ribosomal protein S5.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	e	150	Total	C	N	O	S	0	0
			1106	687	211	202	6		

- Molecule 6 is a protein called 30S ribosomal protein S6, non-modified isoform.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	f	100	Total	C	N	O	S	0	0
			818	515	148	149	6		

- Molecule 7 is a protein called 30S ribosomal protein S7.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	g	151	Total	C	N	O	S	0	0
			1182	735	227	216	4		

- Molecule 8 is a protein called 30S ribosomal protein S8.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	h	129	Total	C	N	O	S	0	0
			979	616	173	184	6		

- Molecule 9 is a protein called 30S ribosomal protein S12.

Mol	Chain	Residues	Atoms					AltConf	Trace
9	l	123	Total	C	N	O	S	0	0
			955	590	196	165	4		

- Molecule 10 is a protein called 30S ribosomal protein S13.

Mol	Chain	Residues	Atoms					AltConf	Trace
10	m	93	Total	C	N	O	S	0	0
			724	451	141	129	3		

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
m	46	LYS	SER	conflict	UNP C3SR52
m	94	ARG	GLY	conflict	UNP C3SR52

- Molecule 11 is a protein called 30S ribosomal protein S15.

Mol	Chain	Residues	Atoms					AltConf	Trace
11	o	88	Total	C	N	O	S	0	0
			714	439	144	130	1		

- Molecule 12 is a protein called 30S ribosomal protein S16.

Mol	Chain	Residues	Atoms					AltConf	Trace
12	p	82	Total	C	N	O	S	0	0
			649	406	128	114	1		

- Molecule 13 is a protein called 30S ribosomal protein S17.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
13	q	80	649	411	121	114	3	0	0

- Molecule 14 is a protein called 30S ribosomal protein S18.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
14	r	55	456	288	86	82	0	0

- Molecule 15 is a protein called 30S ribosomal protein S20.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
15	t	85	665	411	137	114	3	0	0

- Molecule 16 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		AltConf
			Total	Mg	
16	a	79	79	79	0
16	d	1	1	1	0
16	e	1	1	1	0
16	o	1	1	1	0
16	t	1	1	1	0

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: 16S rRNA (guanine(1405)-N(7))-methyltransferase
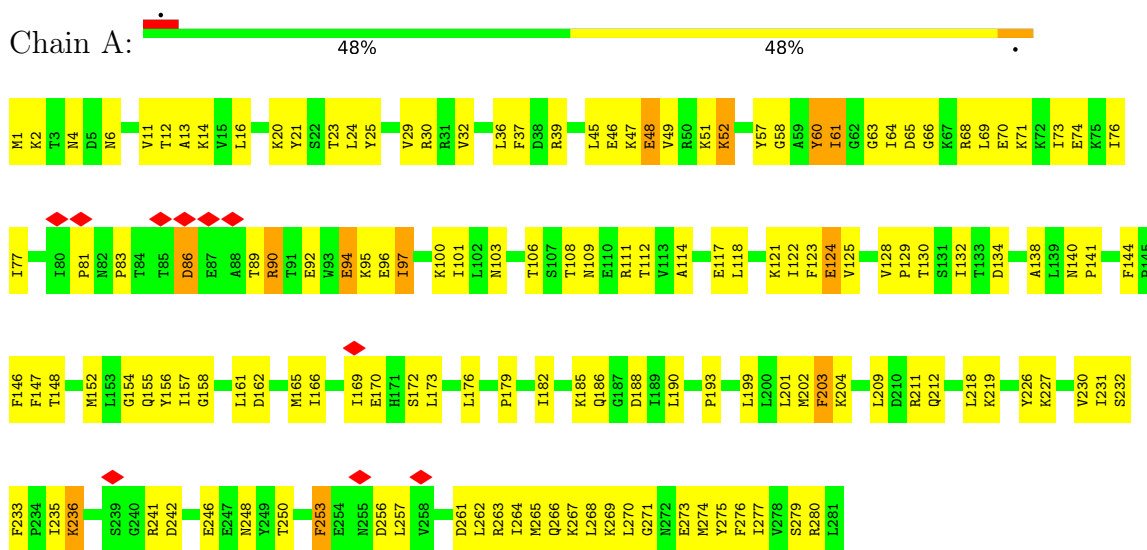
Chain A:

• Molecule 2: 16S rRNA
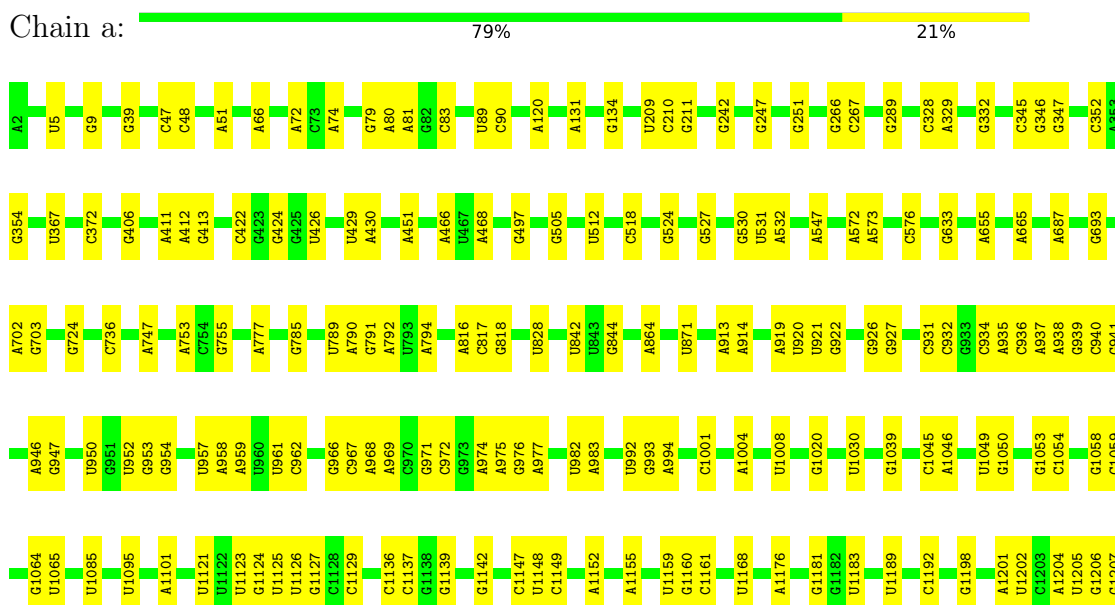
Chain a:

• Molecule 3: 30S ribosomal protein S3

Chain c:

98%
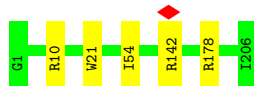
• Molecule 4: 30S ribosomal protein S4

Chain d:

98%
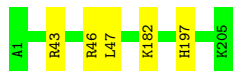
• Molecule 5: 30S ribosomal protein S5

Chain e:

98%

• Molecule 6: 30S ribosomal protein S6, non-modified isoform

Chain f:

97%

• Molecule 7: 30S ribosomal protein S7

Chain g:

97%

• Molecule 8: 30S ribosomal protein S8

Chain h:

98%

S1 E41 F44 K55 A129

• Molecule 9: 30S ribosomal protein S12

Chain l:

98%

A1 R55 R93 R120 A123

• Molecule 10: 30S ribosomal protein S13

Chain m:

99%

A2 R29 R94

• Molecule 11: 30S ribosomal protein S15

Chain o:

98%

S2 K48 R54 R72 R89
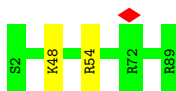
• Molecule 12: 30S ribosomal protein S16

Chain p:

95%   5%

M1 R5 I42 K46 A82

• Molecule 13: 30S ribosomal protein S17

Chain q:

98%

K3 R26 E48 V82

• Molecule 14: 30S ribosomal protein S18

7%

Chain r:

100%

E19 I43 T44 G45 T46 R47 A48 H73
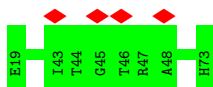
• Molecule 15: 30S ribosomal protein S20

Chain t:

98%
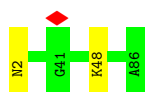
4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	129736	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 $(e^-/\text{Å}^2)$	51	Depositor
Minimum defocus (nm)	500	Depositor
Maximum defocus (nm)	3500	Depositor
Magnification	Not provided	
Image detector	GATAN K3 BIOQUANTUM (6k x 4k)	Depositor
Maximum map value	26.880	Depositor
Minimum map value	-11.218	Depositor
Average map value	0.000	Depositor
Map value standard deviation	1.000	Depositor
Recommended contour level	0.1	Depositor
Map size (Å)	319.36, 319.36, 319.36	wwPDB
Map dimensions	400, 400, 400	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	0.7984, 0.7984, 0.7984	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, ZIV

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	0.52	0/2303	0.75	0/3107
2	a	0.32	0/36782	0.70	0/57379
3	c	0.39	0/1652	0.49	0/2225
4	d	0.39	0/1665	0.48	0/2227
5	e	0.43	0/1119	0.53	0/1504
6	f	0.37	0/836	0.47	0/1128
7	g	0.51	0/1196	0.62	0/1602
8	h	0.33	0/989	0.46	0/1326
9	l	0.40	0/969	0.55	0/1300
10	m	0.37	0/730	0.48	0/975
11	o	0.38	0/722	0.46	0/964
12	p	0.38	0/659	0.49	0/884
13	q	0.36	0/658	0.48	0/881
14	r	0.30	0/463	0.44	0/621
15	t	0.34	0/671	0.43	0/888
All	All	0.35	0/51414	0.66	0/77011

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2   Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2260	0	2286	156	0
2	a	32900	0	16532	0	0
3	c	1625	0	1699	0	0
4	d	1643	0	1710	0	0
5	e	1106	0	1148	0	0
6	f	818	0	808	0	0
7	g	1182	0	1240	0	0
8	h	979	0	1034	0	0
9	l	955	0	1019	0	0
10	m	724	0	768	0	0
11	o	714	0	734	0	0
12	p	649	0	666	0	0
13	q	649	0	691	0	0
14	r	456	0	478	0	0
15	t	665	0	714	0	0
16	a	79	0	0	0	0
16	d	1	0	0	0	0
16	e	1	0	0	0	0
16	o	1	0	0	0	0
16	t	1	0	0	0	0
All	All	47408	0	31527	156	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 8.

The worst 5 of 156 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:257:LEU:O	1:A:257:LEU:HD23	1.39	1.19
1:A:48:GLU:HG3	1:A:51:LYS:CE	1.73	1.17
1:A:23:THR:C	1:A:24:LEU:HD12	1.66	1.13
1:A:48:GLU:HG3	1:A:51:LYS:HE2	1.09	1.08
1:A:73:ILE:HA	1:A:76:ILE:HG22	1.32	1.08

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	279/281 (99%)	227 (81%)	51 (18%)	1 (0%)	34	72
3	c	204/206 (99%)	183 (90%)	21 (10%)	0	100	100
4	d	203/205 (99%)	188 (93%)	15 (7%)	0	100	100
5	e	148/150 (99%)	131 (88%)	17 (12%)	0	100	100
6	f	98/100 (98%)	88 (90%)	10 (10%)	0	100	100
7	g	149/151 (99%)	127 (85%)	22 (15%)	0	100	100
8	h	127/129 (98%)	118 (93%)	9 (7%)	0	100	100
9	l	121/123 (98%)	105 (87%)	16 (13%)	0	100	100
10	m	91/93 (98%)	83 (91%)	8 (9%)	0	100	100
11	o	86/88 (98%)	80 (93%)	6 (7%)	0	100	100
12	p	80/82 (98%)	74 (92%)	6 (8%)	0	100	100
13	q	78/80 (98%)	71 (91%)	7 (9%)	0	100	100
14	r	53/55 (96%)	46 (87%)	7 (13%)	0	100	100
15	t	83/85 (98%)	79 (95%)	4 (5%)	0	100	100
All	All	1800/1828 (98%)	1600 (89%)	199 (11%)	1 (0%)	54	85

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	193	PRO

5.3.2 Protein sidechains ⓘ

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	250/250 (100%)	221 (88%)	29 (12%)	5	23
3	c	170/170 (100%)	165 (97%)	5 (3%)	42	76
4	d	172/172 (100%)	167 (97%)	5 (3%)	42	76
5	e	113/113 (100%)	110 (97%)	3 (3%)	44	77
6	f	87/87 (100%)	84 (97%)	3 (3%)	37	72
7	g	124/124 (100%)	119 (96%)	5 (4%)	31	68
8	h	104/104 (100%)	101 (97%)	3 (3%)	42	76
9	l	103/103 (100%)	100 (97%)	3 (3%)	42	76
10	m	75/75 (100%)	74 (99%)	1 (1%)	69	89
11	o	76/76 (100%)	74 (97%)	2 (3%)	46	78
12	p	65/65 (100%)	61 (94%)	4 (6%)	18	52
13	q	74/74 (100%)	72 (97%)	2 (3%)	44	77
14	r	48/48 (100%)	48 (100%)	0	100	100
15	t	65/65 (100%)	63 (97%)	2 (3%)	40	75
All	All	1526/1526 (100%)	1459 (96%)	67 (4%)	32	65

5 of 67 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
11	o	48	LYS
12	p	1	MET
15	t	2	ASN
1	A	227	LYS
1	A	219	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 22 such sidechains are listed below:

Mol	Chain	Res	Type
7	g	96	ASN
13	q	30	HIS
10	m	91	HIS
14	r	51	GLN
3	c	184	ASN

5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	a	1529/1532 (99%)	320 (20%)	0

5 of 320 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
2	a	5	U
2	a	9	G
2	a	39	G
2	a	47	C
2	a	48	C

There are no RNA pucker outliers to report.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 83 ligands modelled in this entry, 83 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 ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

6 Map visualisation ⓘ

This section contains visualisations of the EMDB entry EMD-40051. 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 ⓘ
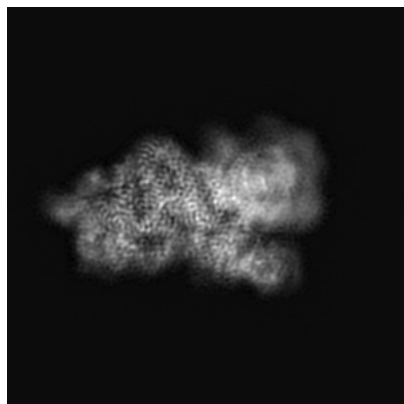
6.1.1 Primary map

X
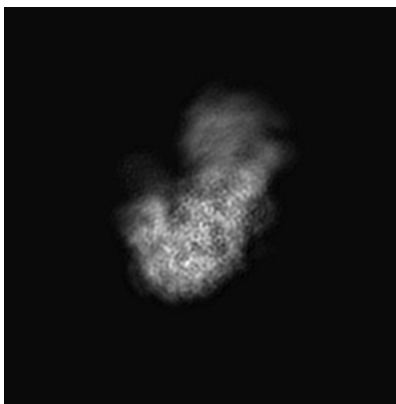
Y
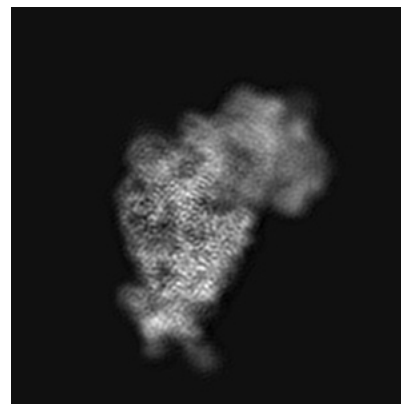
Z

6.1.2 Raw map

X

Y

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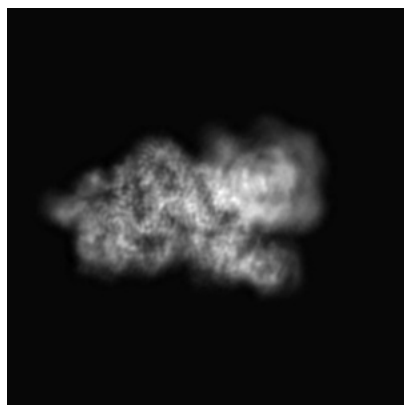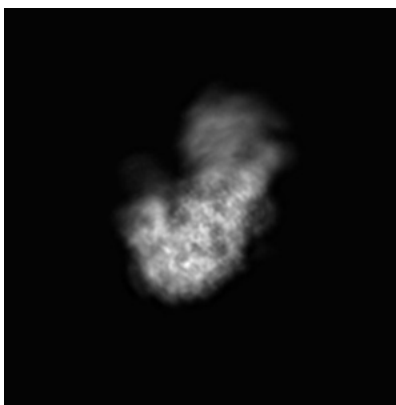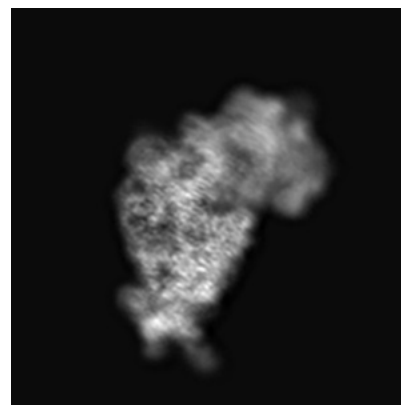
The images above show the map projected in three orthogonal directions.

6.2 Central slices ⓘ
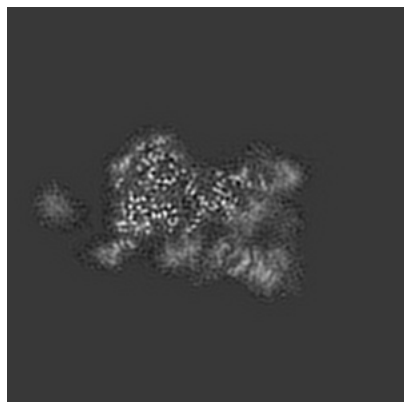
6.2.1 Primary map
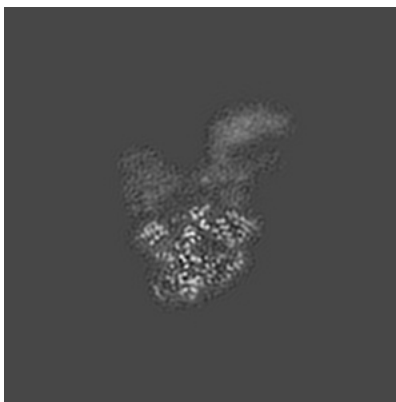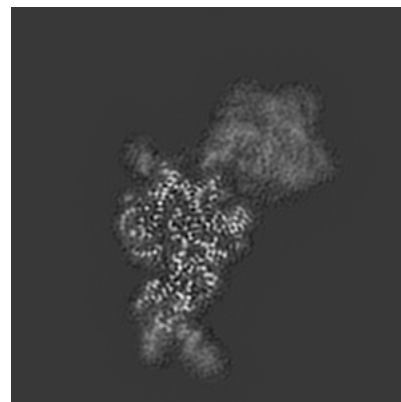
X Index: 200	Y Index: 200	Z Index: 200

6.2.2 Raw map
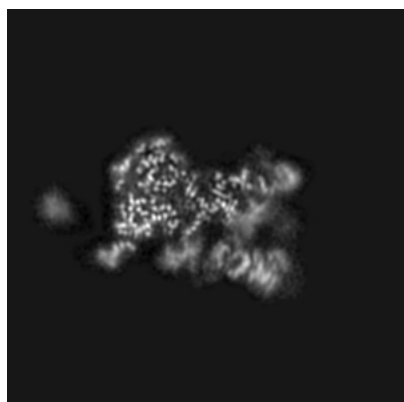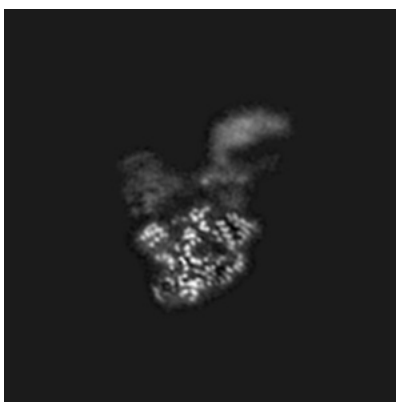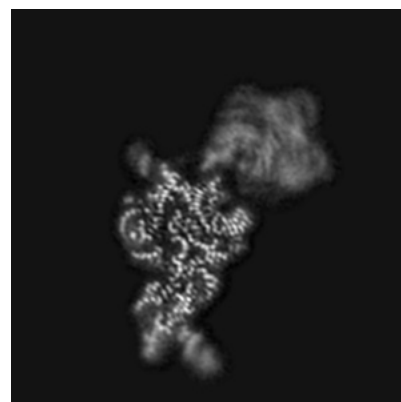
X Index: 200	Y Index: 200	Z Index: 200

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

6.3 Largest variance slices (i)

6.3.1 Primary map
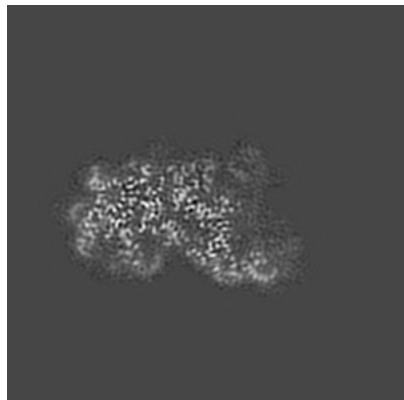
X Index: 168
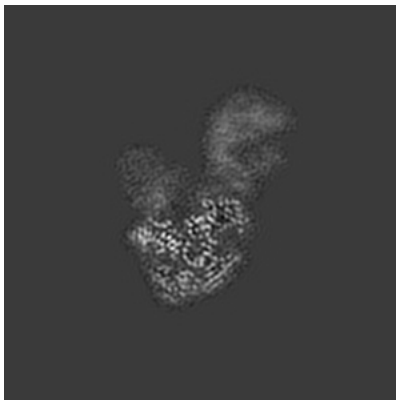
Y Index: 209
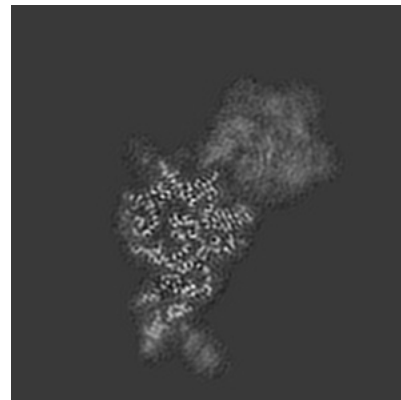
Z Index: 203

6.3.2 Raw map
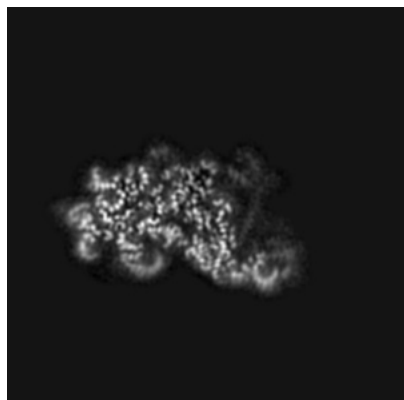
X Index: 171
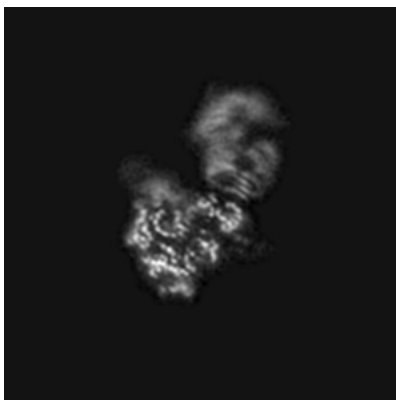
Y Index: 226
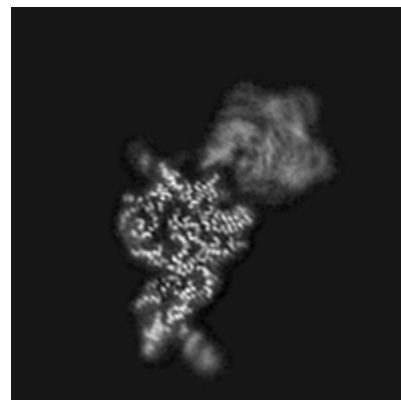
Z Index: 202

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

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

6.4.1   Primary map
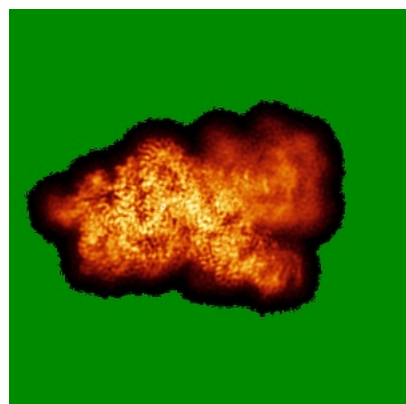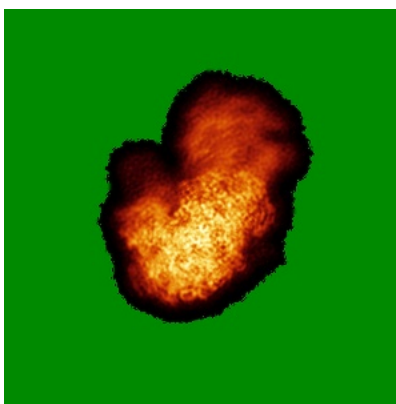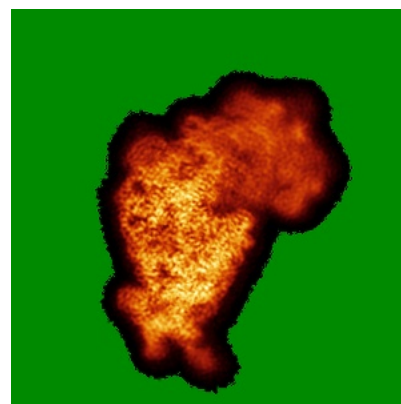
| X | Y | Z |

6.4.2   Raw map

| X | Y | Z |
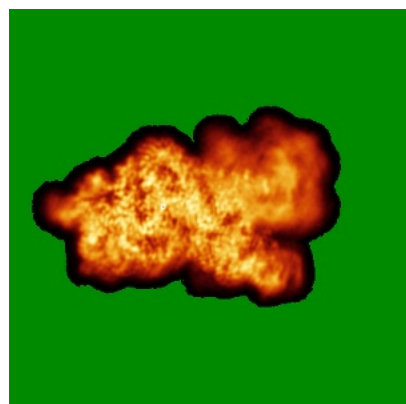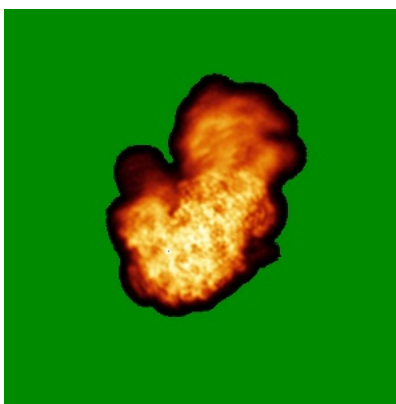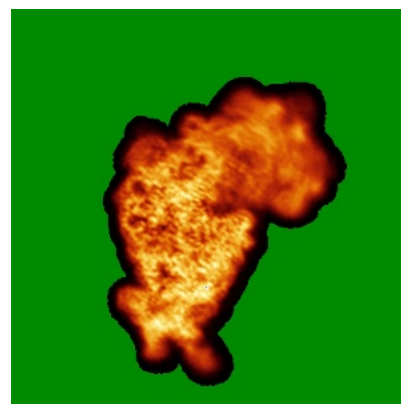
The images above show the map standard deviation projections with false color in three orthogonal directions. Minimum values are shown in green, max in blue, and dark to light orange shades represent small to large values respectively.

6.5 Orthogonal surface views (i)

6.5.1 Primary map

X	Y	Z
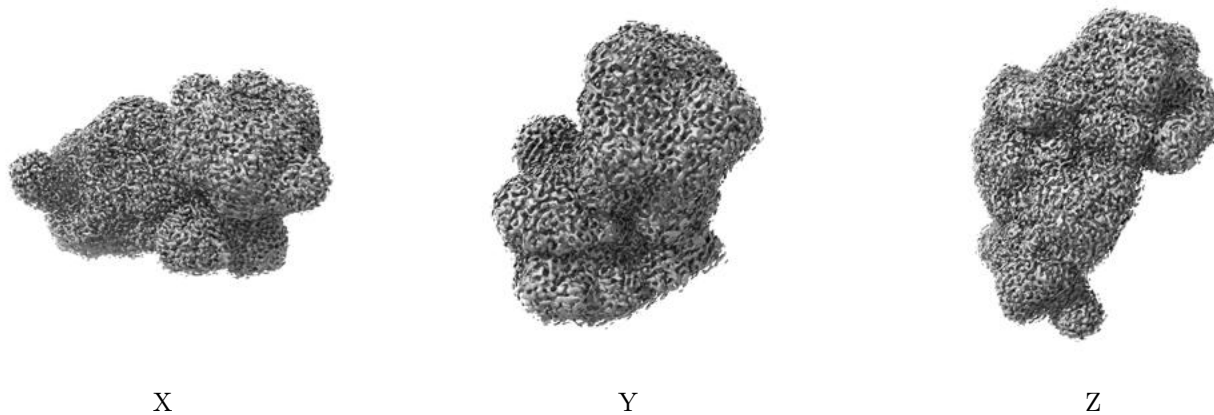
The images above show the 3D surface view of the map at the recommended contour level 0.1. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

6.5.2 Raw map

X	Y	Z
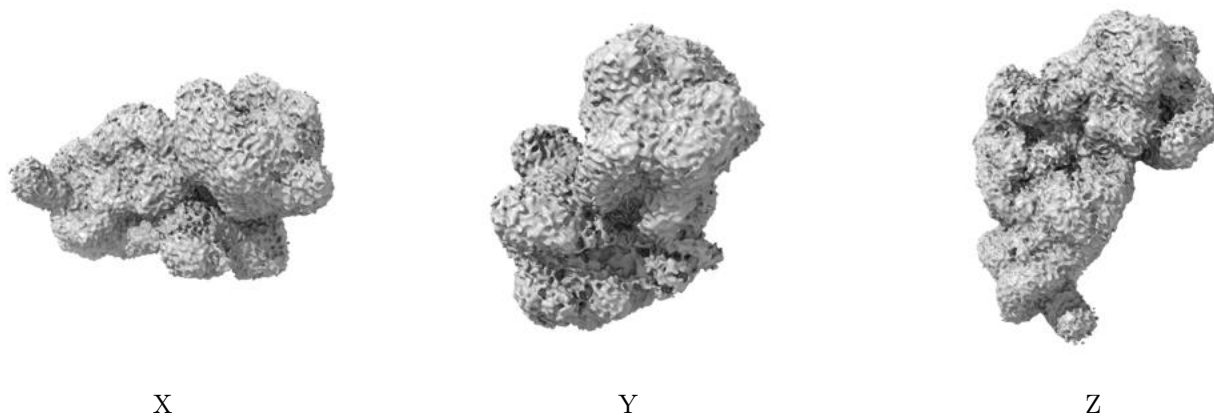
These images show the 3D surface of the raw map. The raw map's contour level was selected so that its surface encloses the same volume as the primary map does at its recommended contour level.

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 ⓘ
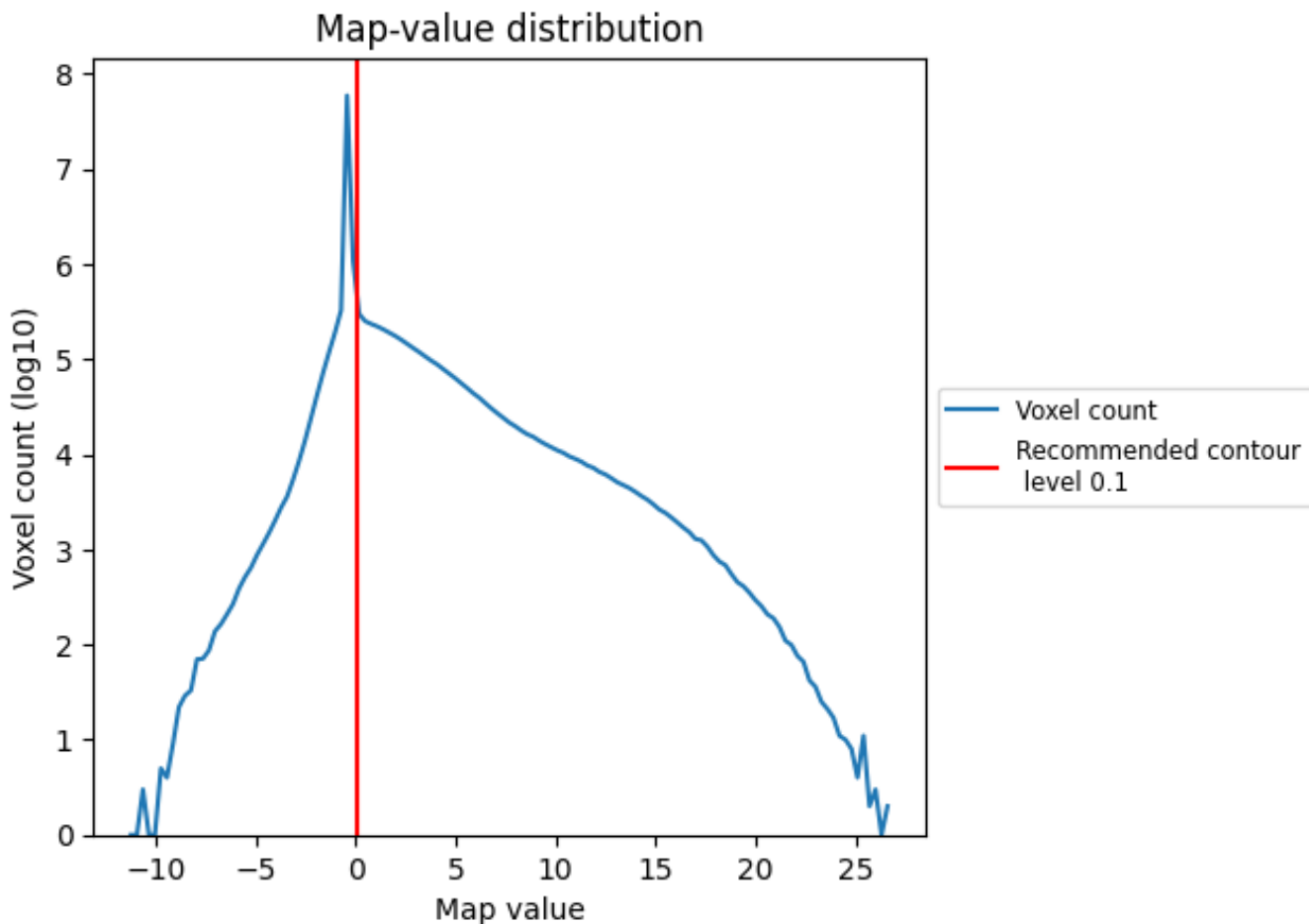
The map-value distribution is plotted in 128 intervals along the x-axis. The y-axis is logarithmic. A spike in this graph at zero usually indicates that the volume has been masked.

7.2　Volume estimate ⓘ
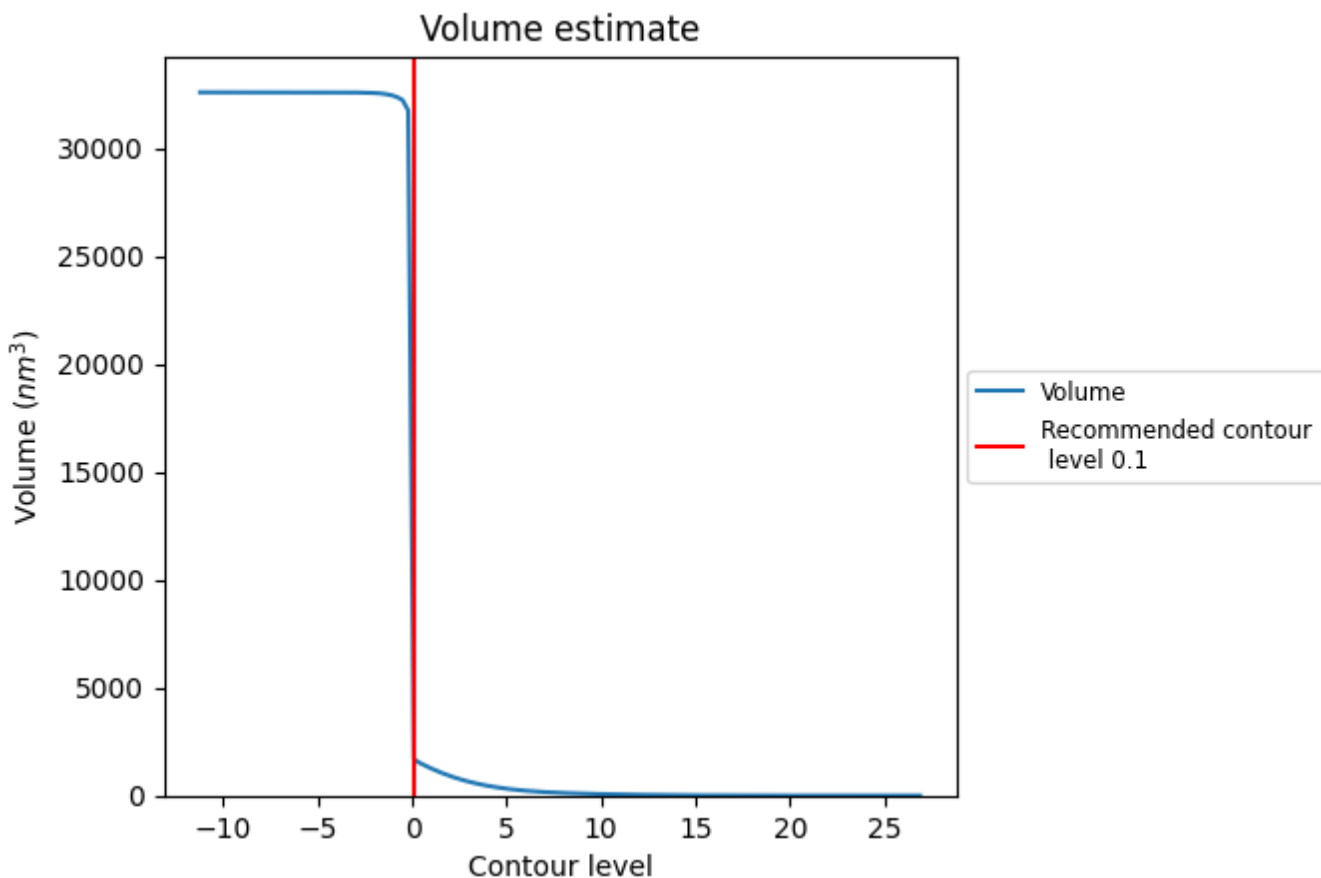
The volume at the recommended contour level is 1686 nm^3; this corresponds to an approximate mass of 1523 kDa.

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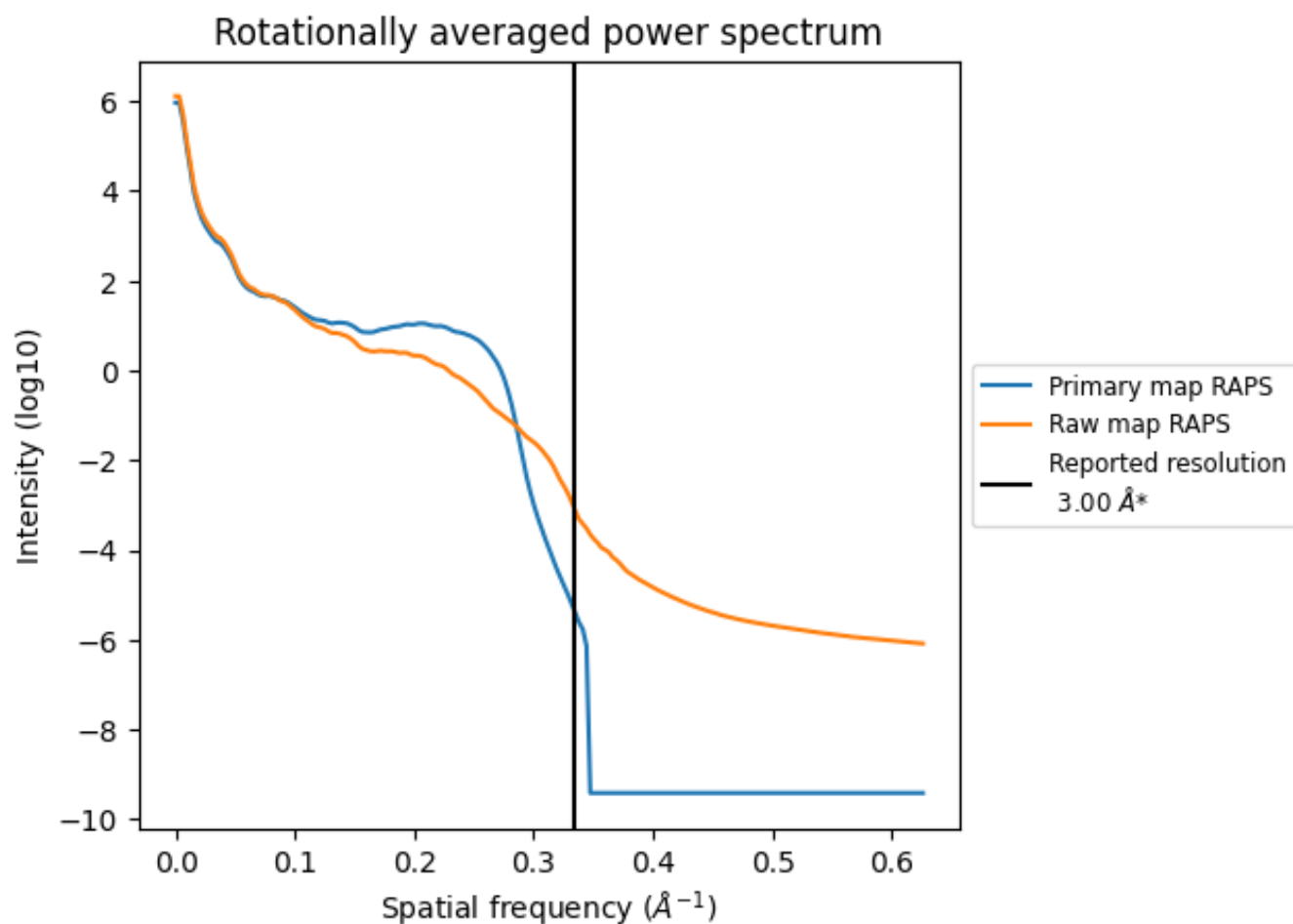
*Reported resolution corresponds to spatial frequency of 0.333 Å^{-1}

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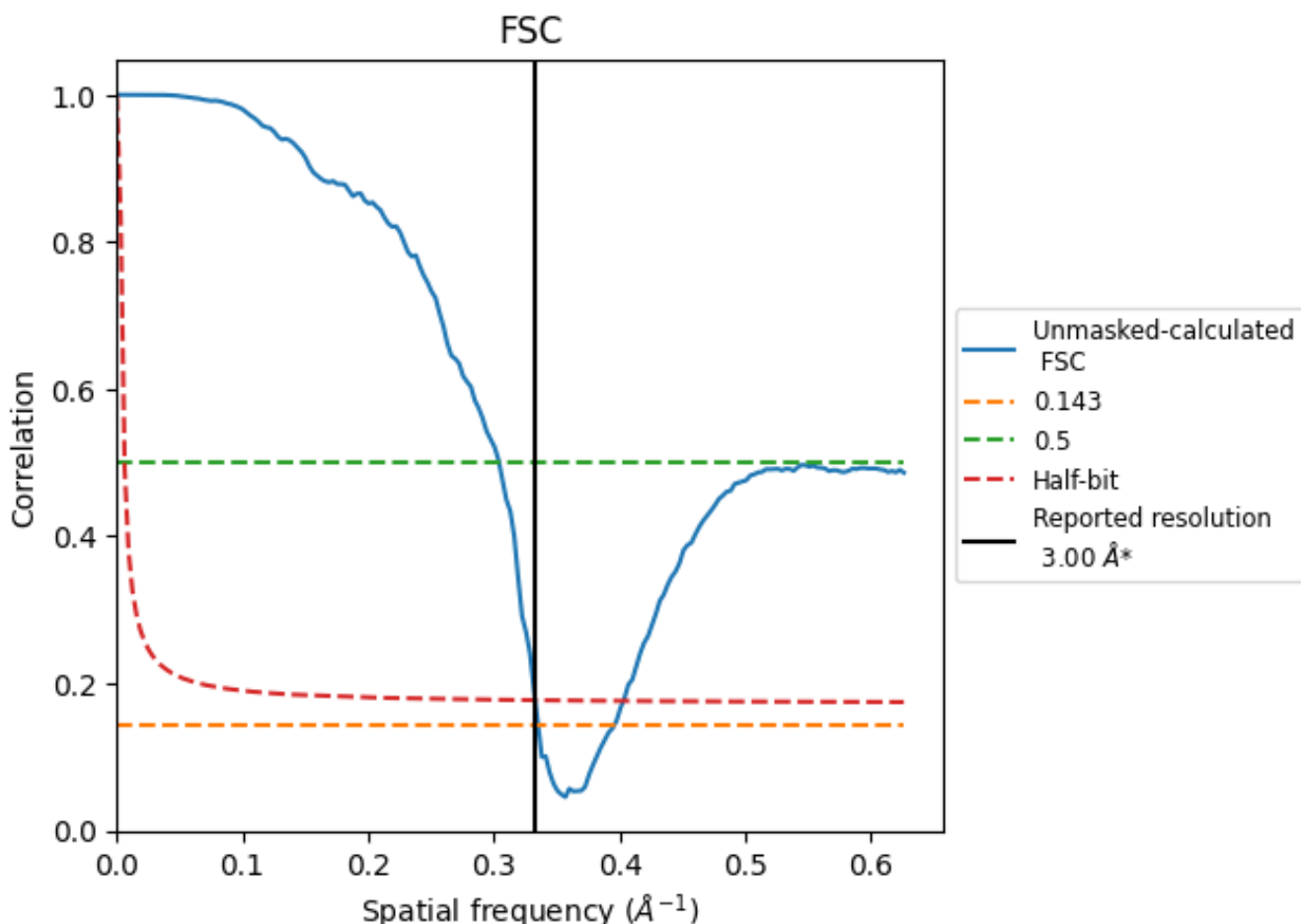
*Reported resolution corresponds to spatial frequency of 0.333 Å^{-1}

8.2　Resolution estimates ⓘ

Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	3.00	-	-
Author-provided FSC curve	-	-	-
Unmasked-calculated*	2.98	3.29	3.00

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

9   Map-model fit ⓘ

This section contains information regarding the fit between EMDB map EMD-40051 and PDB model 8GHU. Per-residue inclusion information can be found in section 3 on page 7.

9.1   Map-model overlay ⓘ

| X | Y | Z |
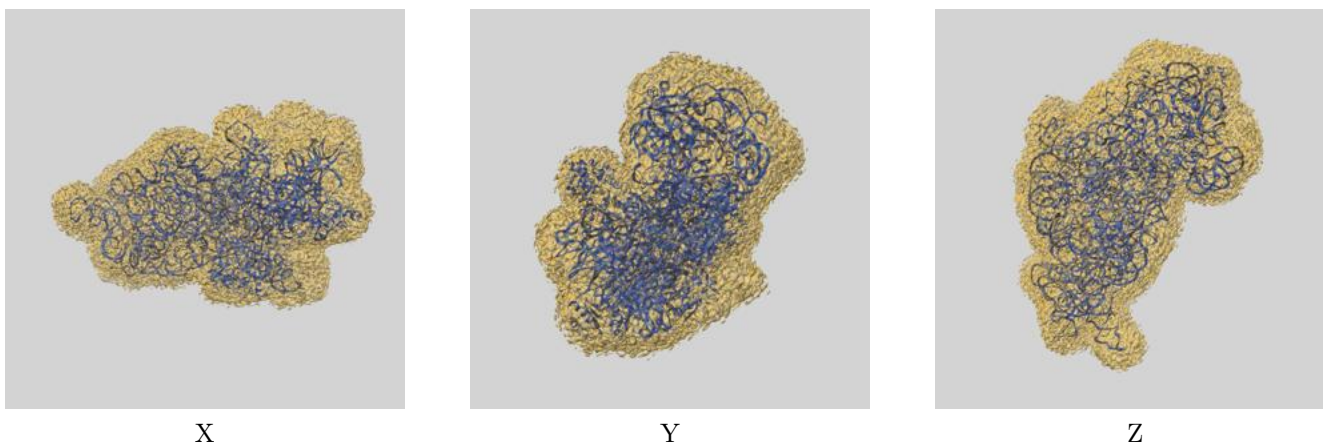
The images above show the 3D surface view of the map at the recommended contour level 0.1 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

9.2   Q-score mapped to coordinate model ⓘ
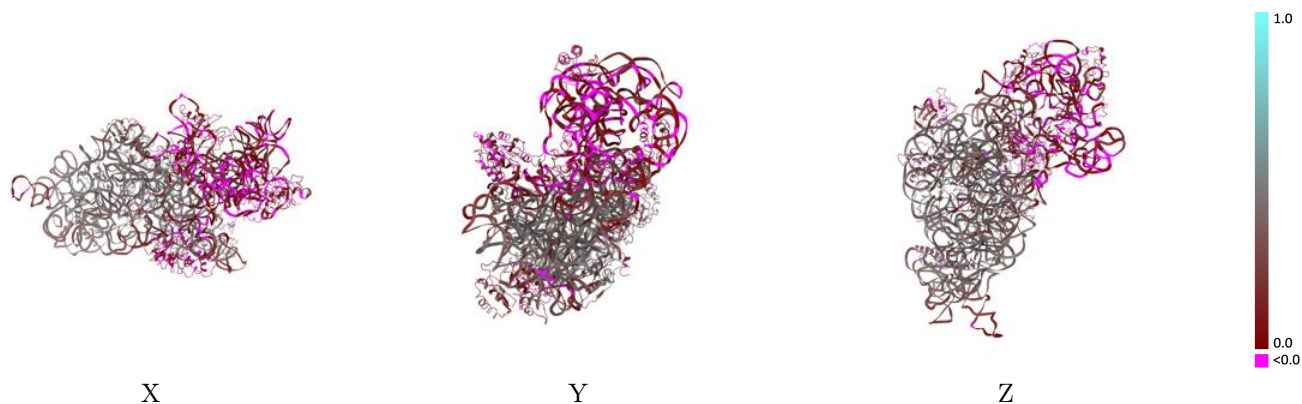
The images above show the model with each residue coloured according its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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 ⓘ
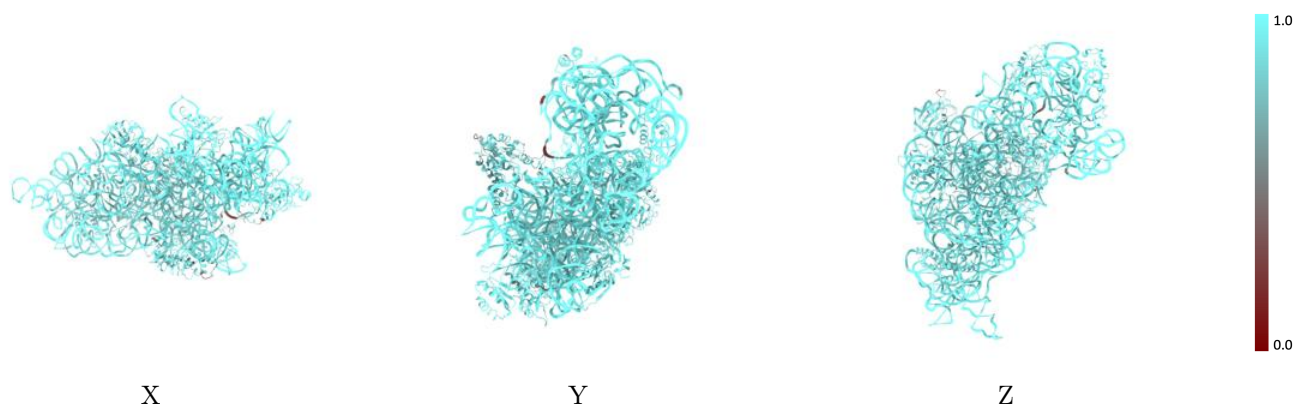
The images above show the model with each residue coloured according to its atom inclusion. This shows to what extent they are inside the map at the recommended contour level (0.1).

9.4　Atom inclusion　ⓘ
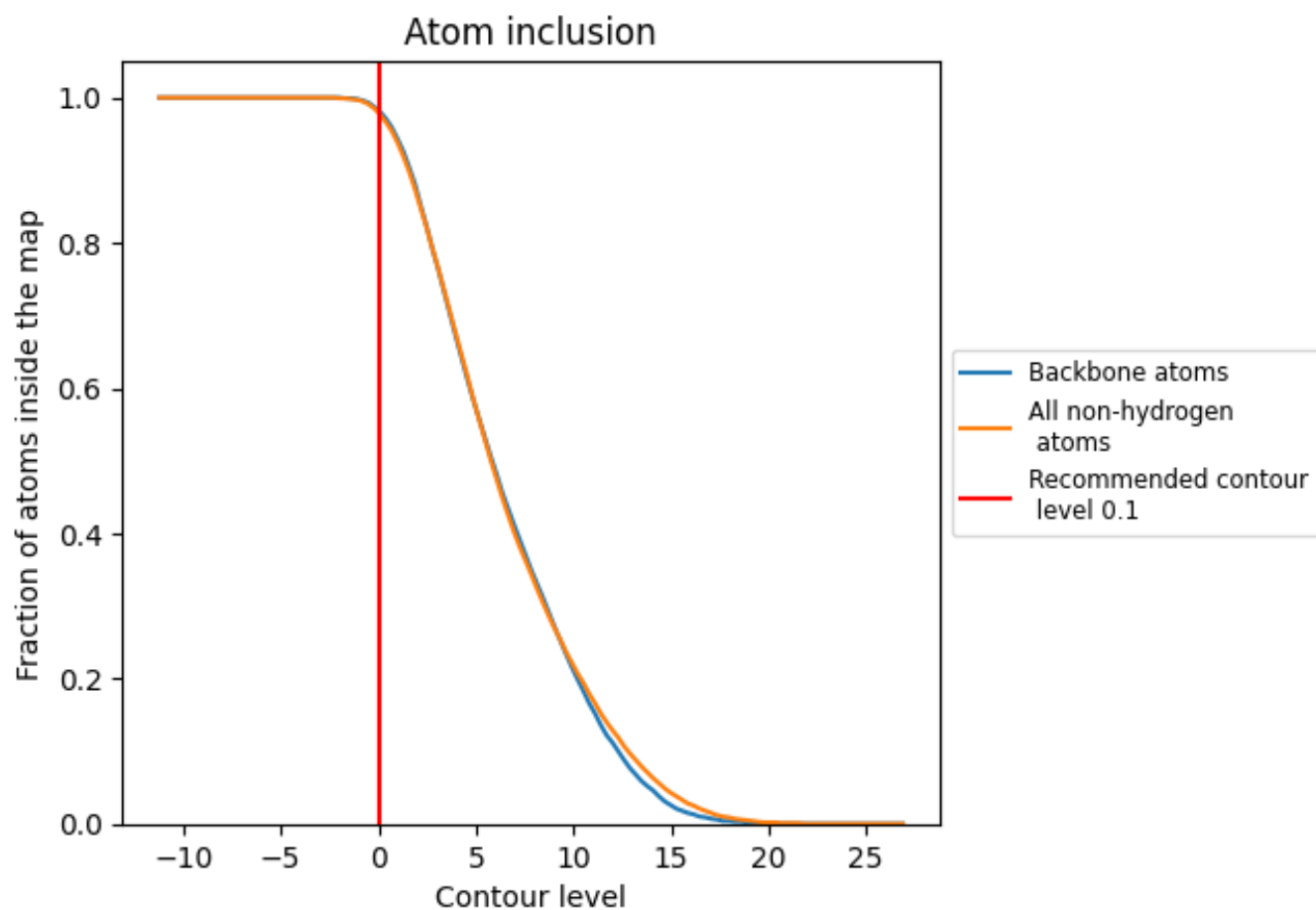
At the recommended contour level, 98% of all backbone atoms, 98% of all non-hydrogen atoms, are inside the map.

9.5 Map-model fit summary (i)

The table lists the average atom inclusion at the recommended contour level (0.1) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.9760	0.2590
A	0.8630	0.1050
a	0.9910	0.2800
c	0.9640	0.1290
d	0.9780	0.3480
e	0.9580	0.3240
f	0.9570	0.2450
g	0.9490	0.0180
h	0.9590	0.3590
l	0.9620	0.2890
m	0.9760	0.0590
o	0.9190	0.1490
p	0.9700	0.3620
q	0.9720	0.3220
r	0.8380	0.0620
t	0.9750	0.2920

1.0

0.0
<0.0