wwPDB EM Validation Summary Report (i)

Jul 7, 2021 – 11:23 am BST

PDB ID	:	7OI0
EMDB ID	:	EMD-12916
Title	:	E.coli delta rbfA pre-30S ribosomal subunit class D
Authors	:	Maksimova, E.; Korepanov, A.; Baymukhametov, T.; Kravchenko, O.; Stolboushkina, E.
Deposited on	:	2021-05-11
Resolution	:	2.76 Å(reported)
Based on initial model	:	4V4Q

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 (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

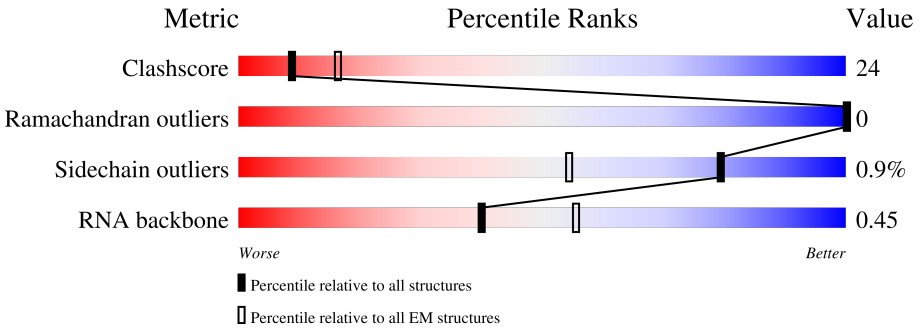
EMDB validation analysis	:	0.0.0.dev84
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.22

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 2.76 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.



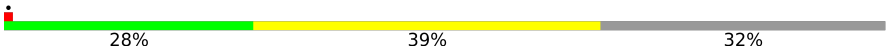
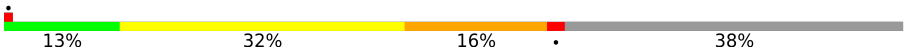
Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826
RNA backbone	4643	859

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	D	205	59% 40%
2	F	135	10% 33% 41% 26%
3	H	129	61% 39%
4	K	128	49% 39% 34% 26%
5	L	123	59% 41%
6	O	89	48% 51%
7	P	82	63% 35%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
8	Q	83	
9	R	74	
10	T	86	
11	A	1542	

2 Entry composition

There are 12 unique types of molecules in this entry. The entry contains 29255 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 30S ribosomal protein S4.

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

- Molecule 2 is a protein called 30S ribosomal protein S6, fully modified isoform.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	F	100	Total	C	N	O	S	0	0
			817	515	148	148	6		

- Molecule 3 is a protein called 30S ribosomal protein S8.

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

- Molecule 4 is a protein called 30S ribosomal protein S11.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	K	95	Total	C	N	O	S	0	0
			702	433	137	130	2		

- Molecule 5 is a protein called 30S ribosomal protein S12.

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

- Molecule 6 is a protein called 30S ribosomal protein S15.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	O	88	Total	C	N	O	S	0	0
			716	440	146	129	1		

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
O	79	ARG	GLN	conflict	UNP A0A4S5B232

- Molecule 7 is a protein called 30S ribosomal protein S16.

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

- Molecule 8 is a protein called 30S ribosomal protein S17.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	Q	80	Total	C	N	O	S	0	0
			648	411	121	113	3		

- Molecule 9 is a protein called 30S ribosomal protein S18.

Mol	Chain	Residues	Atoms				AltConf	Trace
9	R	50	Total	C	N	O	0	0
			407	259	76	72		

- Molecule 10 is a protein called 30S ribosomal protein S20.

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

- Molecule 11 is a RNA chain called 16S rRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
11	A	963	Total	C	N	O	P	0	0
			20700	9228	3823	6686	963		

- Molecule 12 is water.

Mol	Chain	Residues	Atoms		AltConf
12	D	6	Total	O	0
			6	6	
12	L	4	Total	O	0
			4	4	
12	P	18	Total	O	0
			18	18	

Continued on next page...

Continued from previous page...

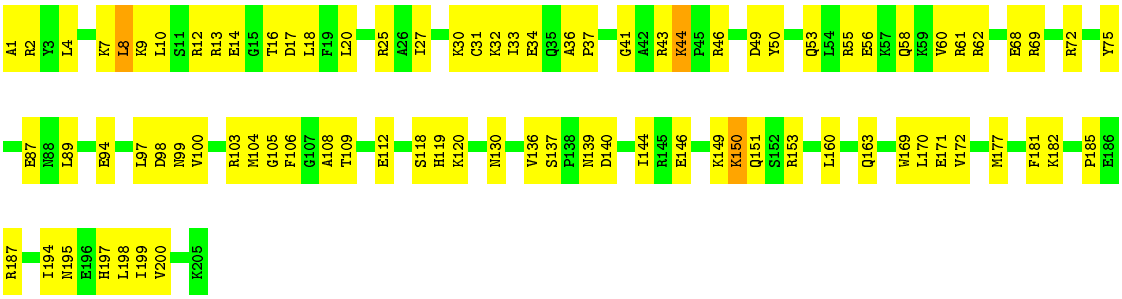
Mol	Chain	Residues	Atoms	AltConf
12	Q	3	Total O 3 3	0
12	T	6	Total O 6 6	0
12	A	337	Total O 337 337	0

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.

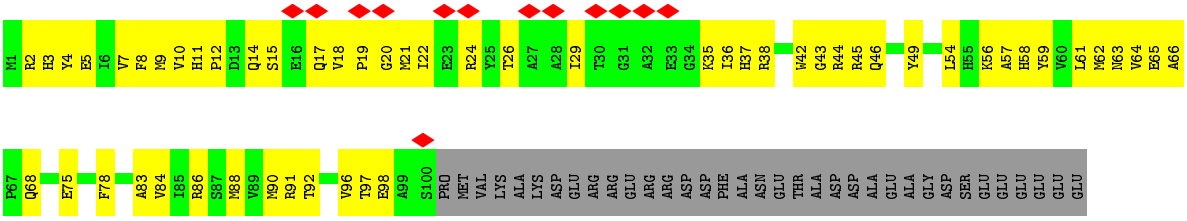
- Molecule 1: 30S ribosomal protein S4

Chain D:



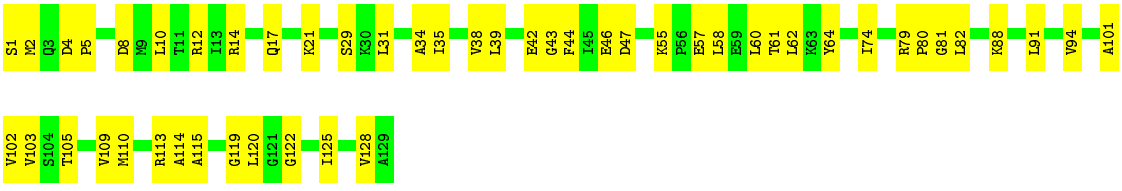
- Molecule 2: 30S ribosomal protein S6, fully modified isoform

Chain F:



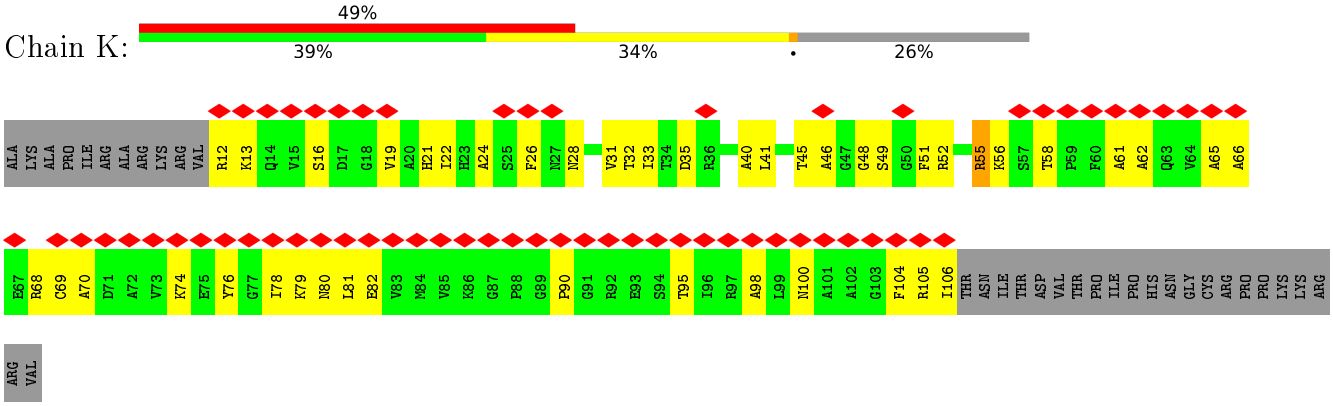
- Molecule 3: 30S ribosomal protein S8

Chain H:



- Molecule 4: 30S ribosomal protein S11

Chain K:

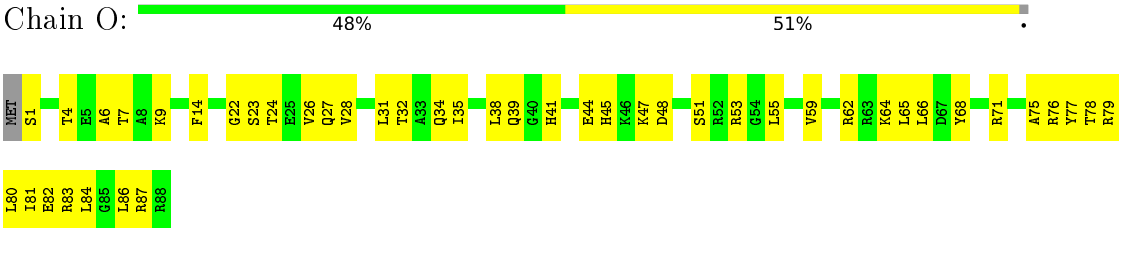


- Molecule 5: 30S ribosomal protein S12

Chain L:

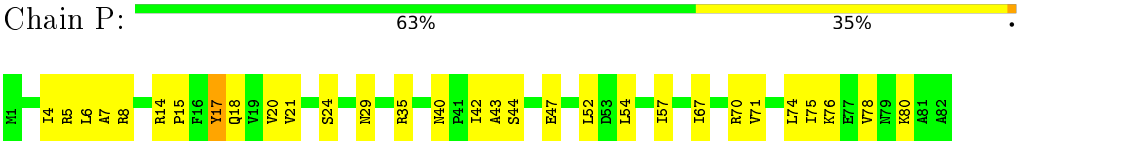
- Molecule 6: 30S ribosomal protein S15

Chain O:



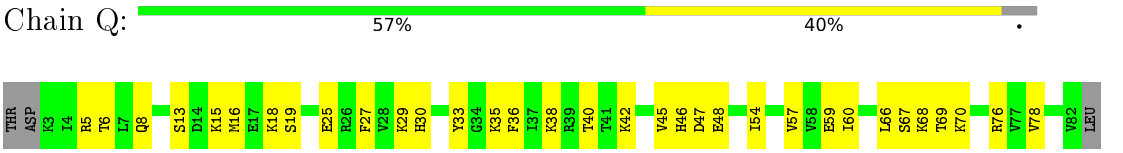
- Molecule 7: 30S ribosomal protein S16

Chain P:



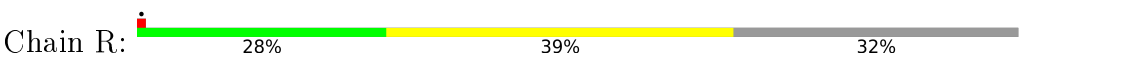
- Molecule 8: 30S ribosomal protein S17

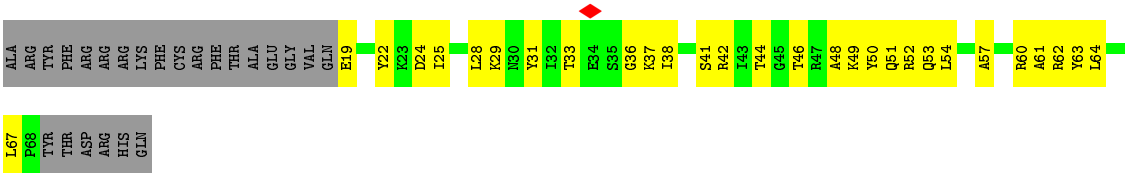
Chain Q:



- Molecule 9: 30S ribosomal protein S18

Chain R:



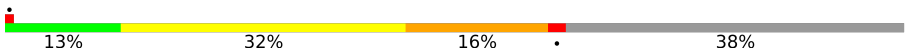


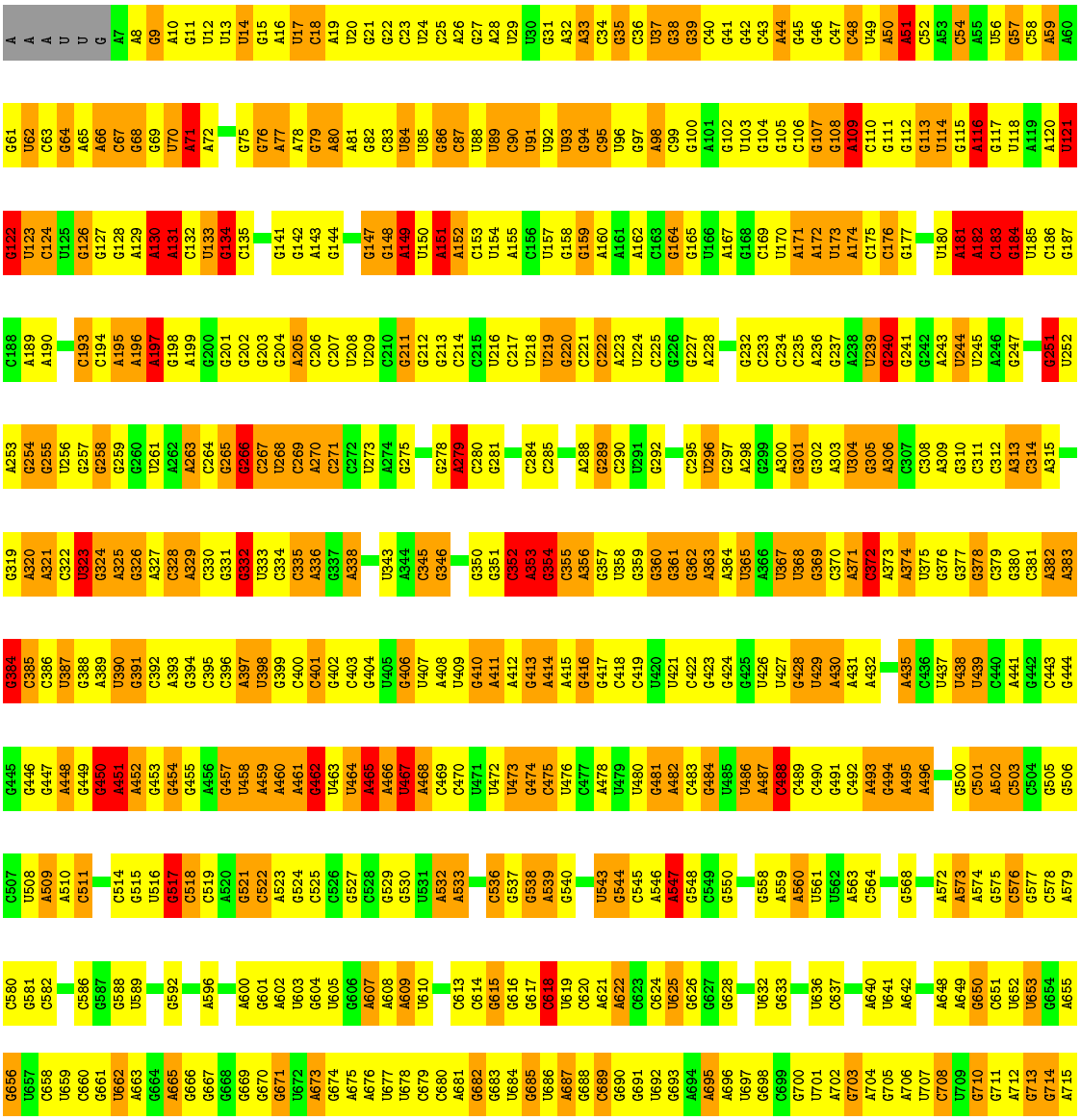
• Molecule 10: 30S ribosomal protein S20

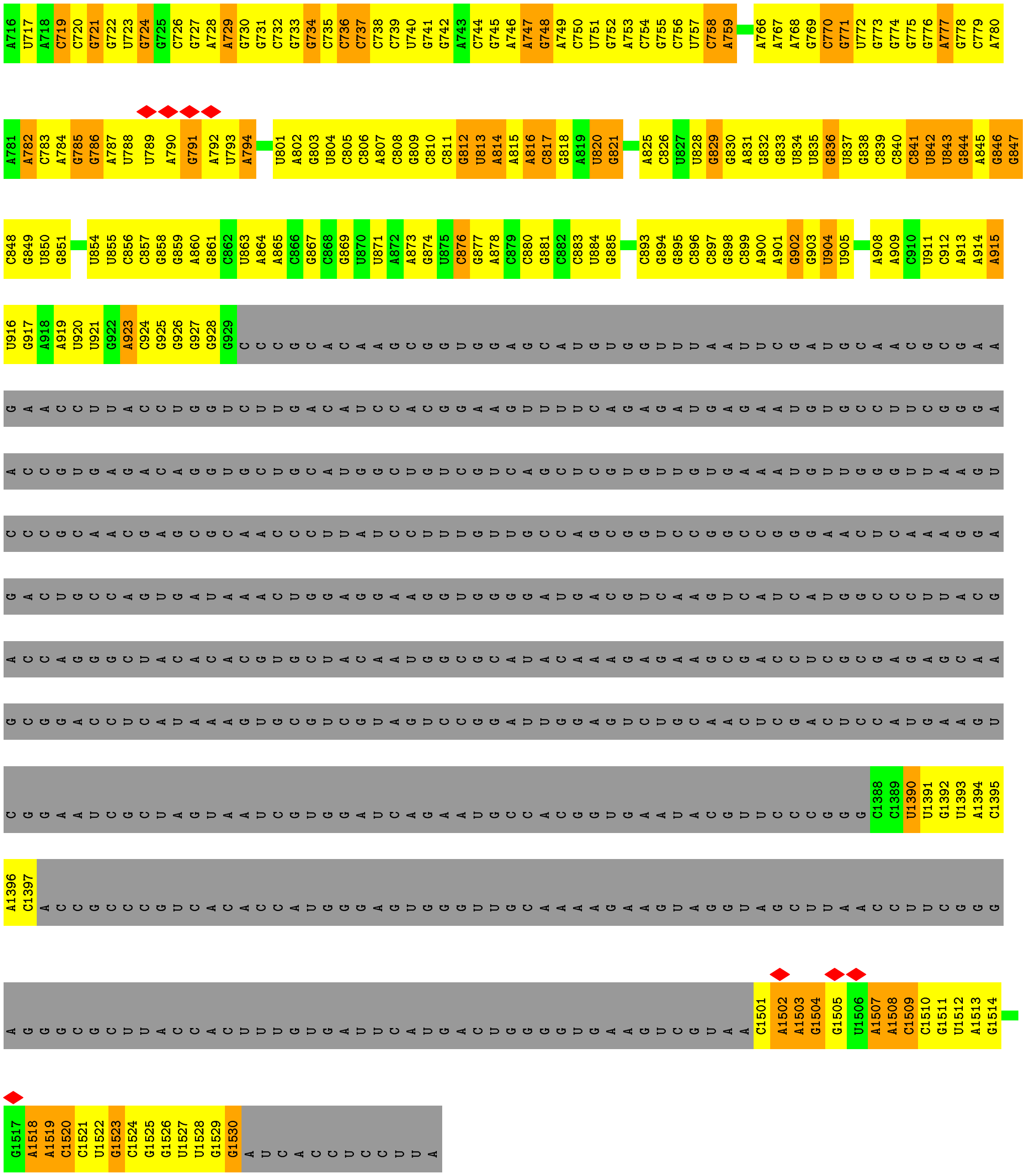
Chain T:

• Molecule 11: 16S rRNA

Chain A:






A716 U717 A718 C719 C720 G721 G722 U723 G724 G725 C726 G727 A728 A729 G730 G731 C732 G733 G734 C735 C736 C737 C738 C739 U740 G741 G742 A743 C744 G745 A746 A747 G748 A749 C750 U751 G752 A753 C754 G755 C756 U757 C758 A759 A766 A767 A768 G769 C770 G771 U772 G773 G774 G775 G776 A777 G778 C779 A780
A781 A782 C783 A784 G785 G786 A787 U788 U789 A790 G791 A792 U793 A794 U801 A802 G803 U804 C805 C806 A807 C808 G809 C810 C811 G812 U813 A814 A815 A816 C817 G818 A819 U820 G821 A825 C826 U827 U828 G829 G830 A831 G832 G833 U834 U835 G836 U837 G838 C839 C840 C841 U842 U843 G844 A845 G846 G847
C848 G849 U850 G851 U854 U855 C856 C857 G858 G859 A860 G861 C862 U863 A864 A865 C866 G867 C868 G869 U870 U871 A872 A873 G874 U875 C876 G877 A878 C879 C880 G881 C882 C883 U884 G885 C893 G894 G895 C896 C897 G898 C899 A900 A901 G902 G903 U904 U905 A908 A909 C910 U911 C912 A913 A914 A915
U916 G917 A918 A919 U920 U921 G922 A923 C924 G925 G926 G927 G928 G929
C1388 C1389 U1390 U1391 G1392 U1393 A1394 C1395
A1396 C1397
C1501 A1502 A1503 G1504 G1505 U1506 A1507 A1508 C1509 C1510 G1511 U1512 A1513 G1514
G1517 A1518 A1519 C1520 C1521 U1522 G1523 C1524 G1525 G1526 U1527 U1528 G1529 G1530

4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	106319	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING ONLY	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/Å^2$)	2.5	Depositor
Minimum defocus (nm)	Not provided	
Maximum defocus (nm)	Not provided	
Magnification	Not provided	
Image detector	FEI FALCON II (4k x 4k)	Depositor
Maximum map value	5.611	Depositor
Minimum map value	-2.640	Depositor
Average map value	0.007	Depositor
Map value standard deviation	0.083	Depositor
Recommended contour level	0.218	Depositor
Map size (Å)	378.4, 378.4, 378.4	wwPDB
Map dimensions	440, 440, 440	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	0.86, 0.86, 0.86	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	D	0.62	0/1665	0.67	0/2227
2	F	0.29	0/835	0.48	0/1128
3	H	0.47	0/989	0.61	0/1326
4	K	0.26	0/713	0.46	0/960
5	L	0.68	0/969	0.77	0/1300
6	O	0.41	0/724	0.53	0/966
7	P	1.03	2/659 (0.3%)	0.79	0/884
8	Q	0.80	0/657	0.68	0/881
9	R	0.25	0/412	0.46	0/553
10	T	0.66	0/671	0.69	0/888
11	A	1.71	381/23184 (1.6%)	1.20	132/36171 (0.4%)
All	All	1.50	383/31478 (1.2%)	1.10	132/47284 (0.3%)

The worst 5 of 383 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
11	A	452	A	N9-C4	-10.64	1.31	1.37
11	A	353	A	N9-C4	-10.25	1.31	1.37
11	A	355	C	N3-C4	-8.73	1.27	1.33
11	A	113	G	C6-N1	-8.63	1.33	1.39
11	A	394	G	C5-C4	-7.84	1.32	1.38

The worst 5 of 132 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
11	A	108	G	C4-C5-N7	15.57	117.03	110.80
11	A	108	G	C5-N7-C8	-12.91	97.84	104.30
11	A	108	G	C6-C5-N7	-12.81	122.71	130.40
11	A	452	A	C2-N3-C4	-10.97	105.11	110.60
11	A	452	A	N3-C4-C5	10.44	134.11	126.80

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	D	1643	0	1710	120	0
2	F	817	0	808	62	0
3	H	979	0	1034	50	0
4	K	702	0	702	41	0
5	L	955	0	1019	61	0
6	O	716	0	742	50	0
7	P	649	0	666	25	0
8	Q	648	0	691	33	0
9	R	407	0	438	26	0
10	T	665	0	714	23	0
11	A	20700	0	10405	711	0
12	A	337	0	0	22	0
12	D	6	0	0	0	0
12	L	4	0	0	0	0
12	P	18	0	0	3	0
12	Q	3	0	0	1	0
12	T	6	0	0	1	0
All	All	29255	0	18929	1094	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 24.

The worst 5 of 1094 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:D:31:CYS:SG	11:A:429:U:H5"	1.50	1.52
11:A:26:A:N6	11:A:558:G:N3	1.91	1.17
1:D:8:LEU:CD2	11:A:429:U:H3'	1.73	1.16
1:D:31:CYS:SG	11:A:429:U:C5'	2.36	1.12
11:A:411:A:C2	11:A:430:A:N6	2.17	1.11

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	D	203/205 (99%)	177 (87%)	26 (13%)	0	100	100
2	F	98/135 (73%)	92 (94%)	6 (6%)	0	100	100
3	H	127/129 (98%)	117 (92%)	10 (8%)	0	100	100
4	K	93/128 (73%)	83 (89%)	10 (11%)	0	100	100
5	L	121/123 (98%)	102 (84%)	19 (16%)	0	100	100
6	O	86/89 (97%)	82 (95%)	4 (5%)	0	100	100
7	P	80/82 (98%)	70 (88%)	10 (12%)	0	100	100
8	Q	78/83 (94%)	68 (87%)	10 (13%)	0	100	100
9	R	48/74 (65%)	41 (85%)	7 (15%)	0	100	100
10	T	83/86 (96%)	81 (98%)	2 (2%)	0	100	100
All	All	1017/1134 (90%)	913 (90%)	104 (10%)	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	D	172/172 (100%)	169 (98%)	3 (2%)	60	76
2	F	87/116 (75%)	86 (99%)	1 (1%)	73	84
3	H	104/104 (100%)	102 (98%)	2 (2%)	57	73
4	K	69/98 (70%)	68 (99%)	1 (1%)	67	79

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
5	L	103/103 (100%)	102 (99%)	1 (1%)	76 85
6	O	76/77 (99%)	76 (100%)	0	100 100
7	P	65/65 (100%)	65 (100%)	0	100 100
8	Q	74/77 (96%)	74 (100%)	0	100 100
9	R	43/64 (67%)	43 (100%)	0	100 100
10	T	65/65 (100%)	65 (100%)	0	100 100
All	All	858/941 (91%)	850 (99%)	8 (1%)	79 87

5 of 8 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
5	L	95	HIS
4	K	55	ARG
3	H	88	LYS
2	F	35	LYS
3	H	113	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
8	Q	8	GLN
10	T	69	ASN
4	K	100	ASN
5	L	71	HIS
6	O	41	HIS

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
11	A	960/1542 (62%)	259 (26%)	24 (2%)

5 of 259 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
11	A	9	G
11	A	14	U
11	A	18	C
11	A	20	U

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
11	A	25	C

5 of 24 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
11	A	429	U
11	A	776	G
11	A	532	A
11	A	812	G
11	A	266	G

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.

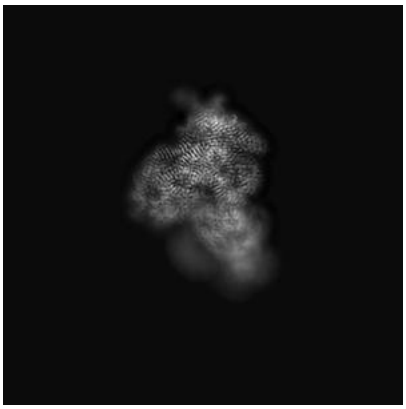
6 Map visualisation

This section contains visualisations of the EMDB entry EMD-12916. These allow visual inspection of the internal detail of the map and identification of artifacts.

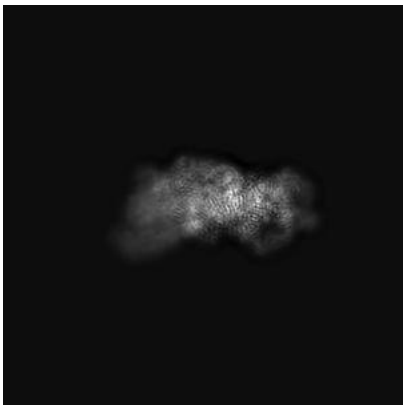
Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

6.1 Orthogonal projections

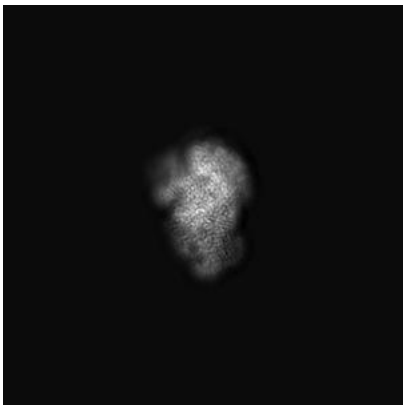
6.1.1 Primary map



X

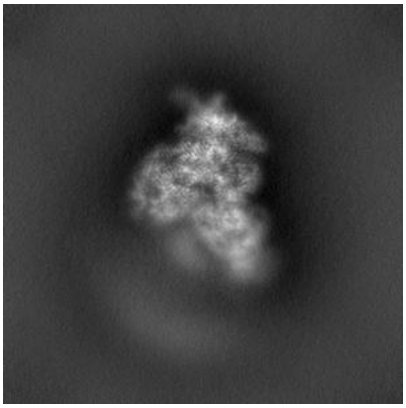


Y

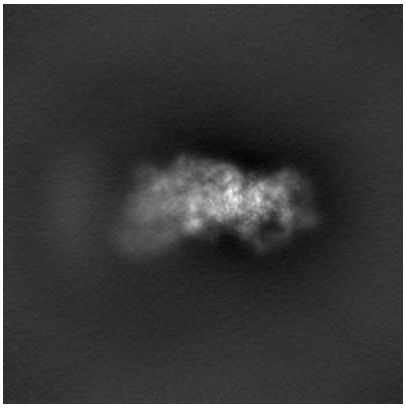


Z

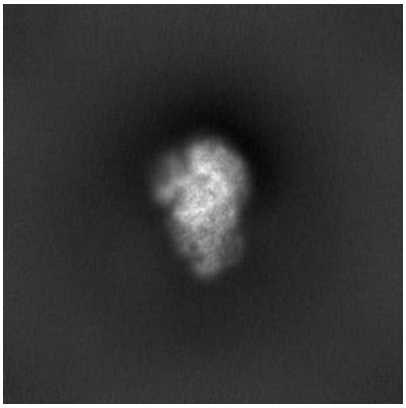
6.1.2 Raw map



X



Y

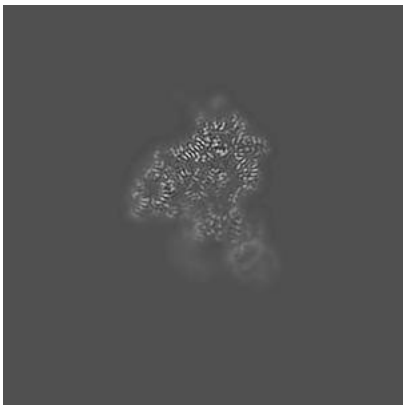


Z

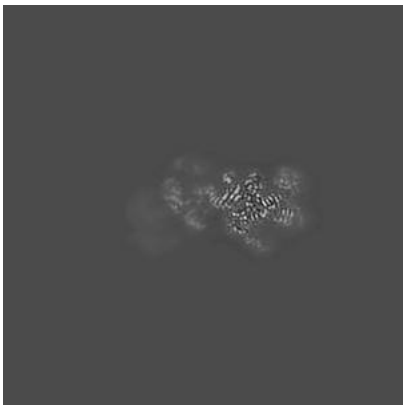
The images above show the map projected in three orthogonal directions.

6.2 Central slices (i)

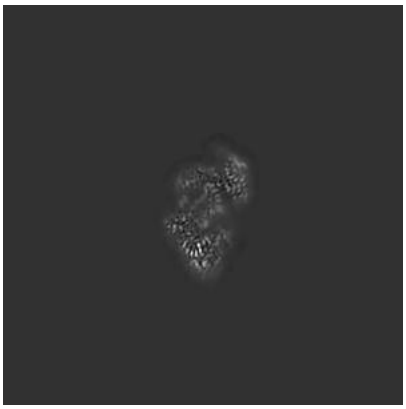
6.2.1 Primary map



X Index: 220

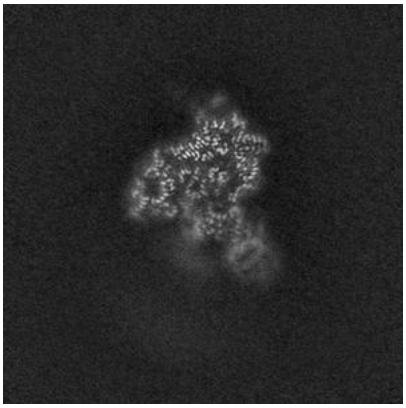


Y Index: 220

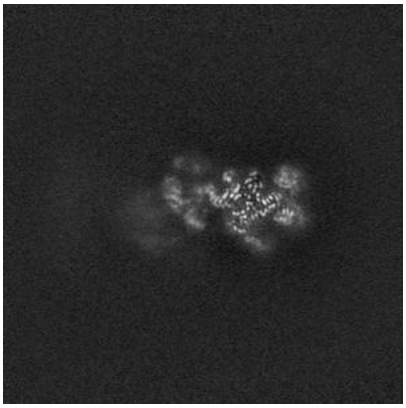


Z Index: 220

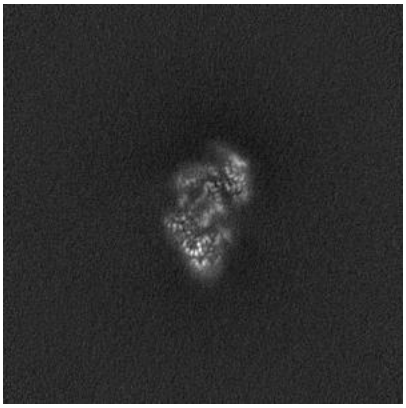
6.2.2 Raw map



X Index: 220



Y Index: 220

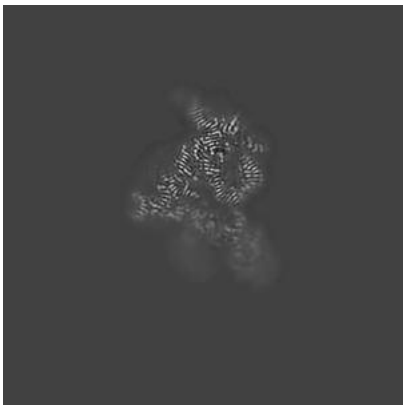


Z Index: 220

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

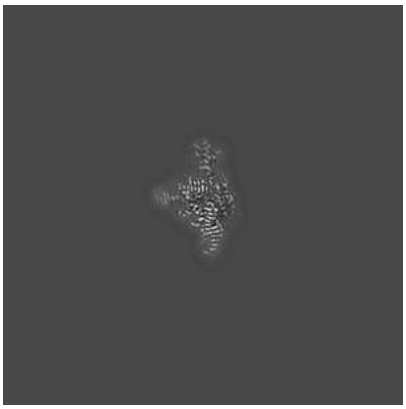
6.3 Largest variance slices (i)

6.3.1 Primary map



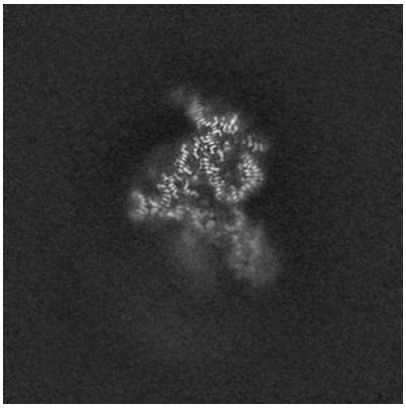
X Index: 213

Y Index: 238

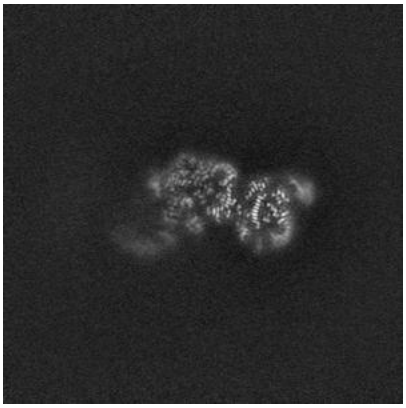


Z Index: 279

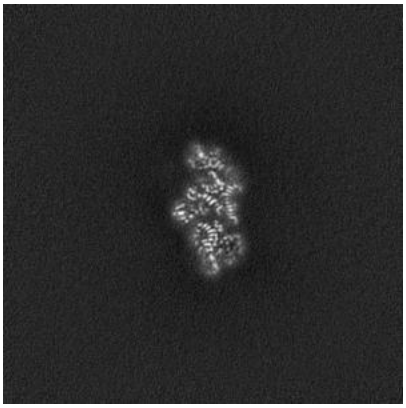
6.3.2 Raw map



X Index: 213



Y Index: 238

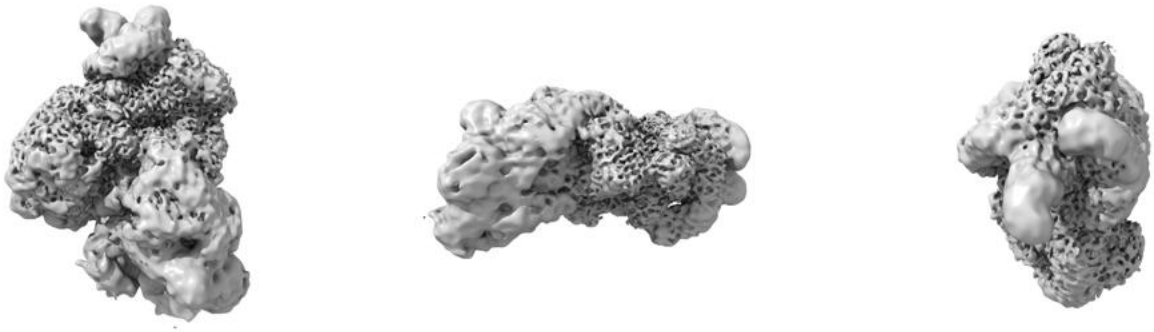


Z Index: 250

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

6.4 Orthogonal surface views (i)

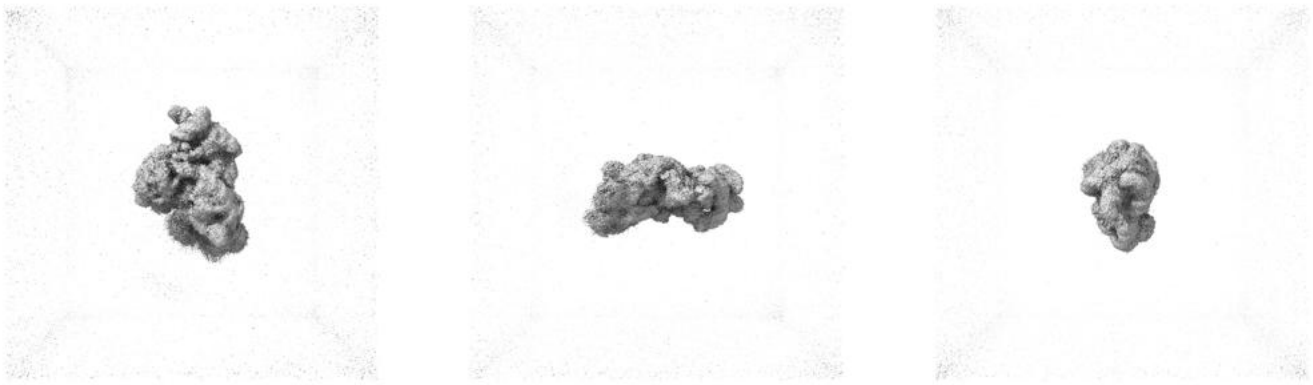
6.4.1 Primary map



X Y Z

The images above show the 3D surface view of the map at the recommended contour level 0.218. 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



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.

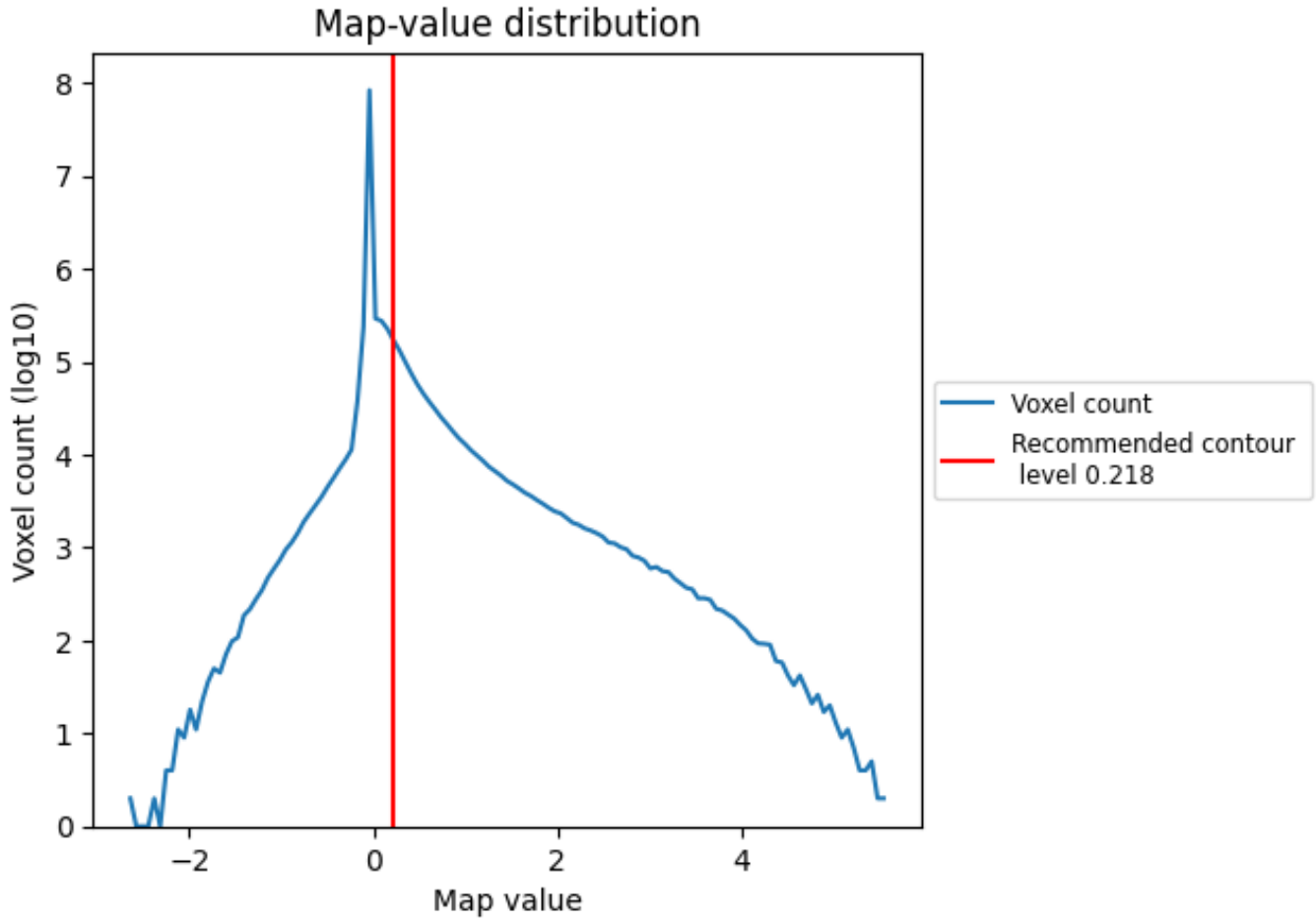
6.5 Mask visualisation (i)

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

7 Map analysis

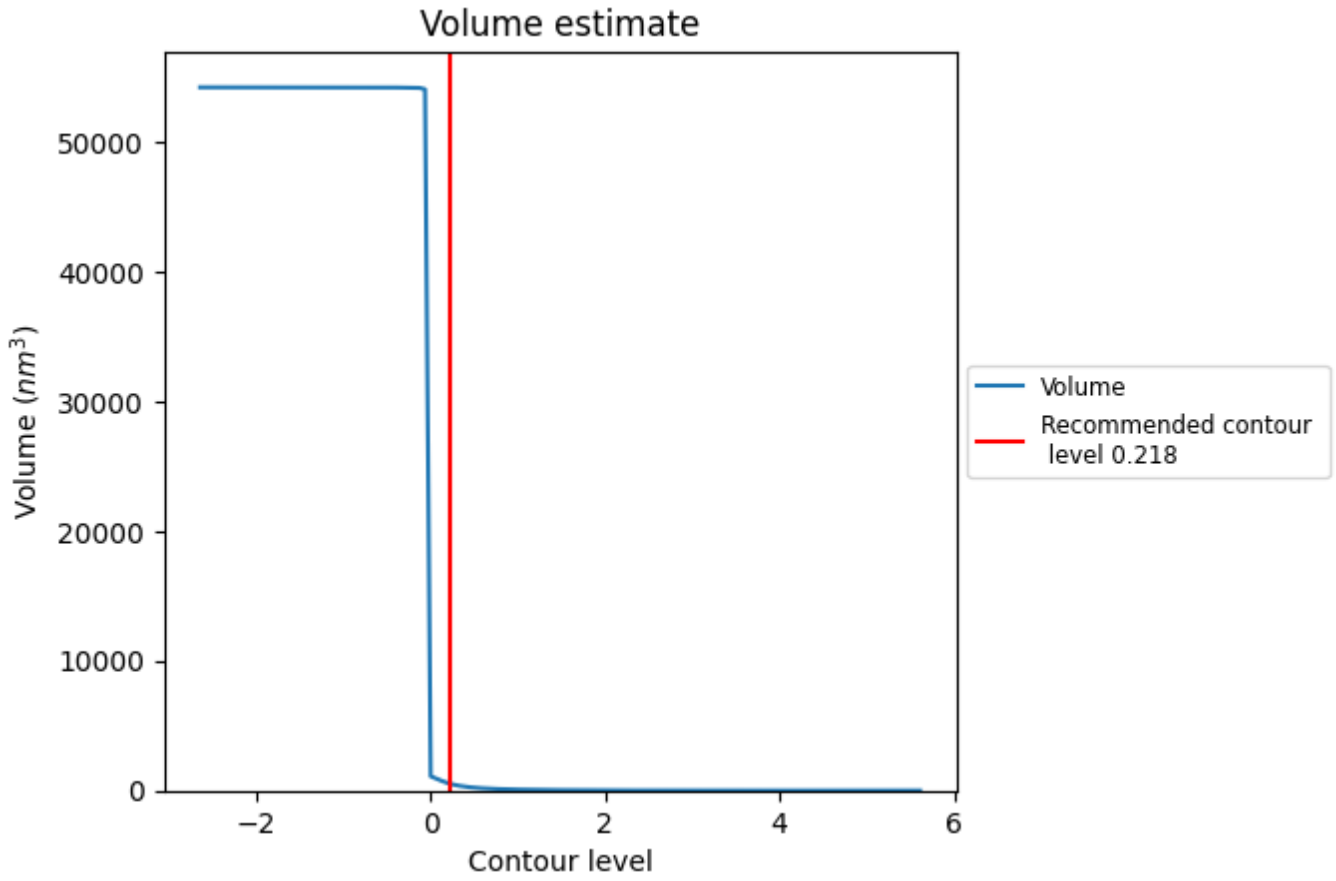
This section contains the results of statistical analysis of the map.

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.

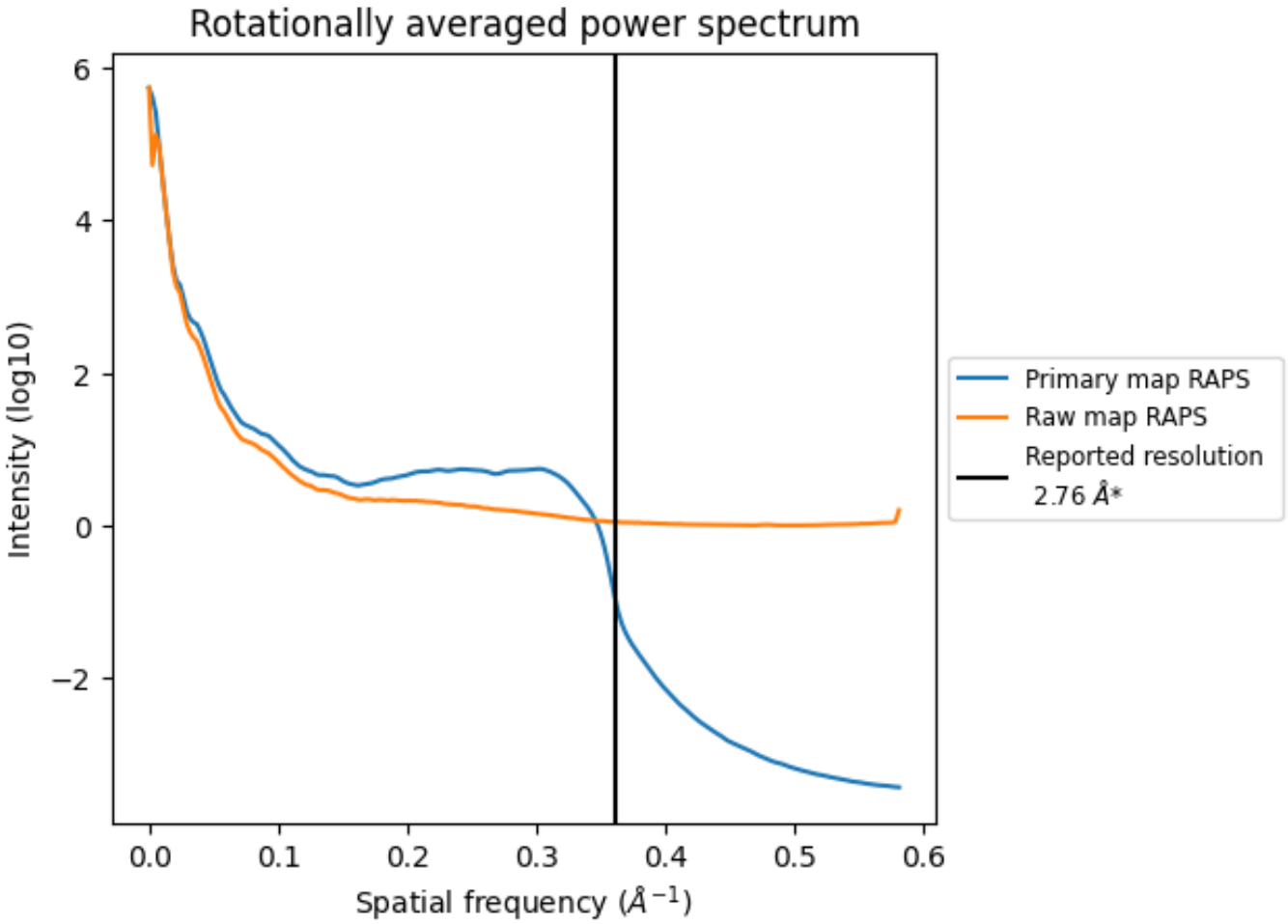
7.2 Volume estimate



The volume at the recommended contour level is 544 nm^3; this corresponds to an approximate mass of 491 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)

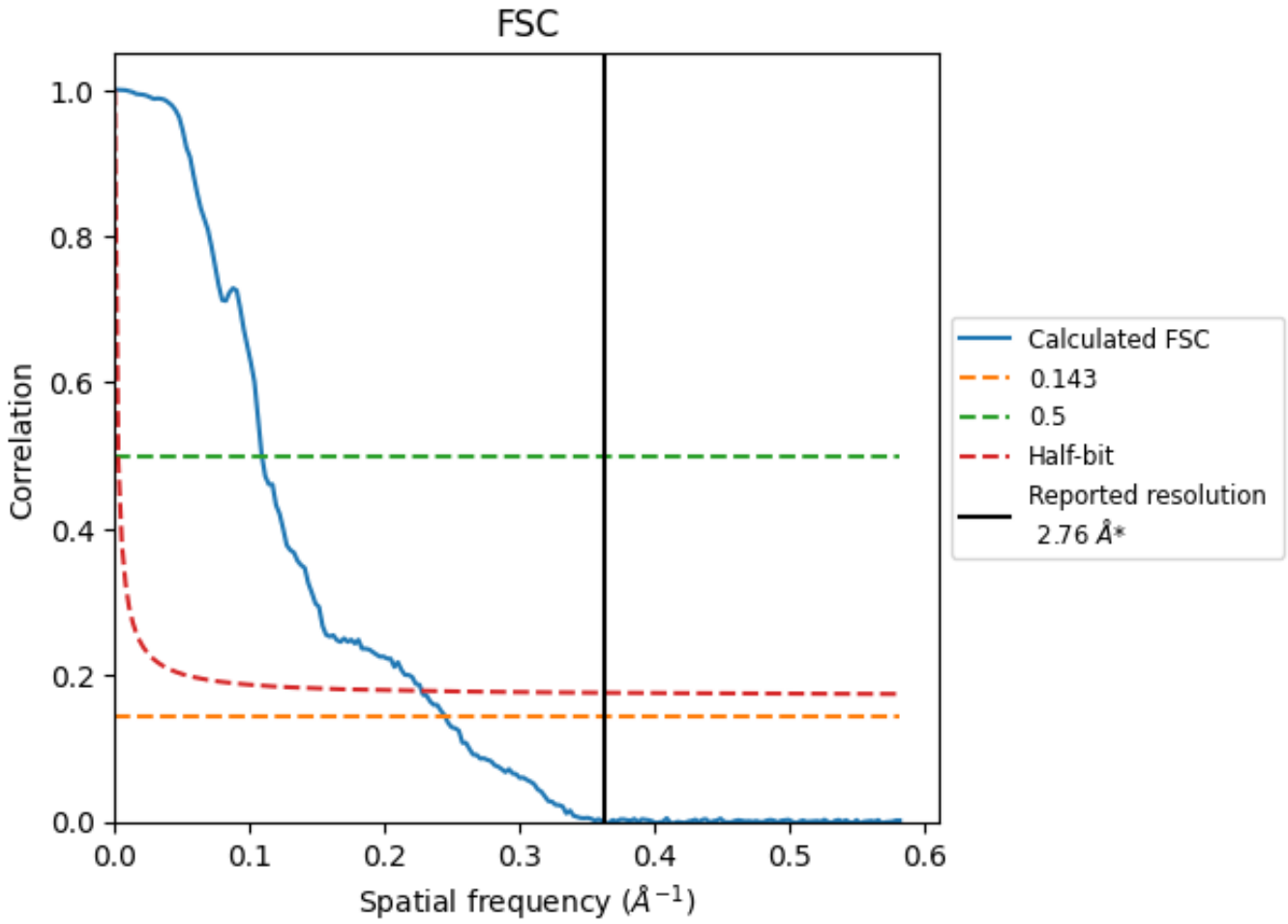


*Reported resolution corresponds to spatial frequency of 0.362 Å^{-1}

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 0.362 Å^{-1}

8.2 Resolution estimates (i)

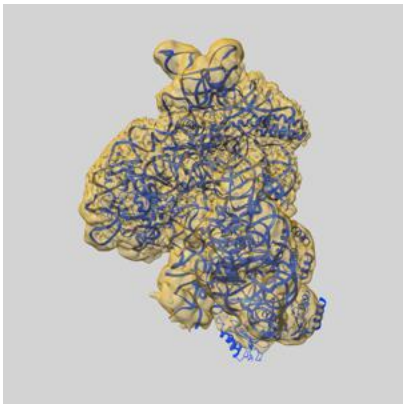
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	2.76	-	-
Author-provided FSC curve	-	-	-
Calculated*	4.08	9.17	4.40

*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 4.08 differs from the reported value 2.76 by more than 10 %

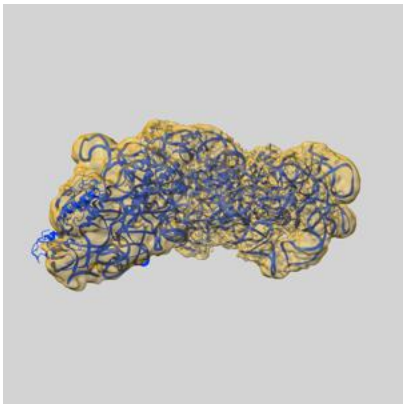
9 Map-model fit

This section contains information regarding the fit between EMDB map EMD-12916 and PDB model 7OI0. Per-residue inclusion information can be found in section 3 on page 7.

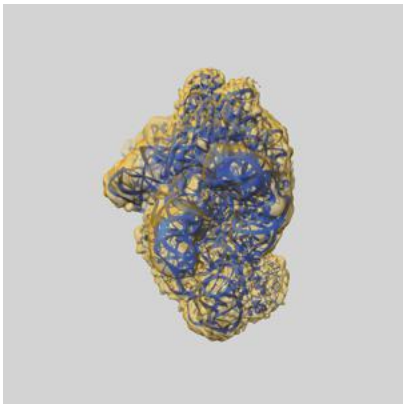
9.1 Map-model overlay



X



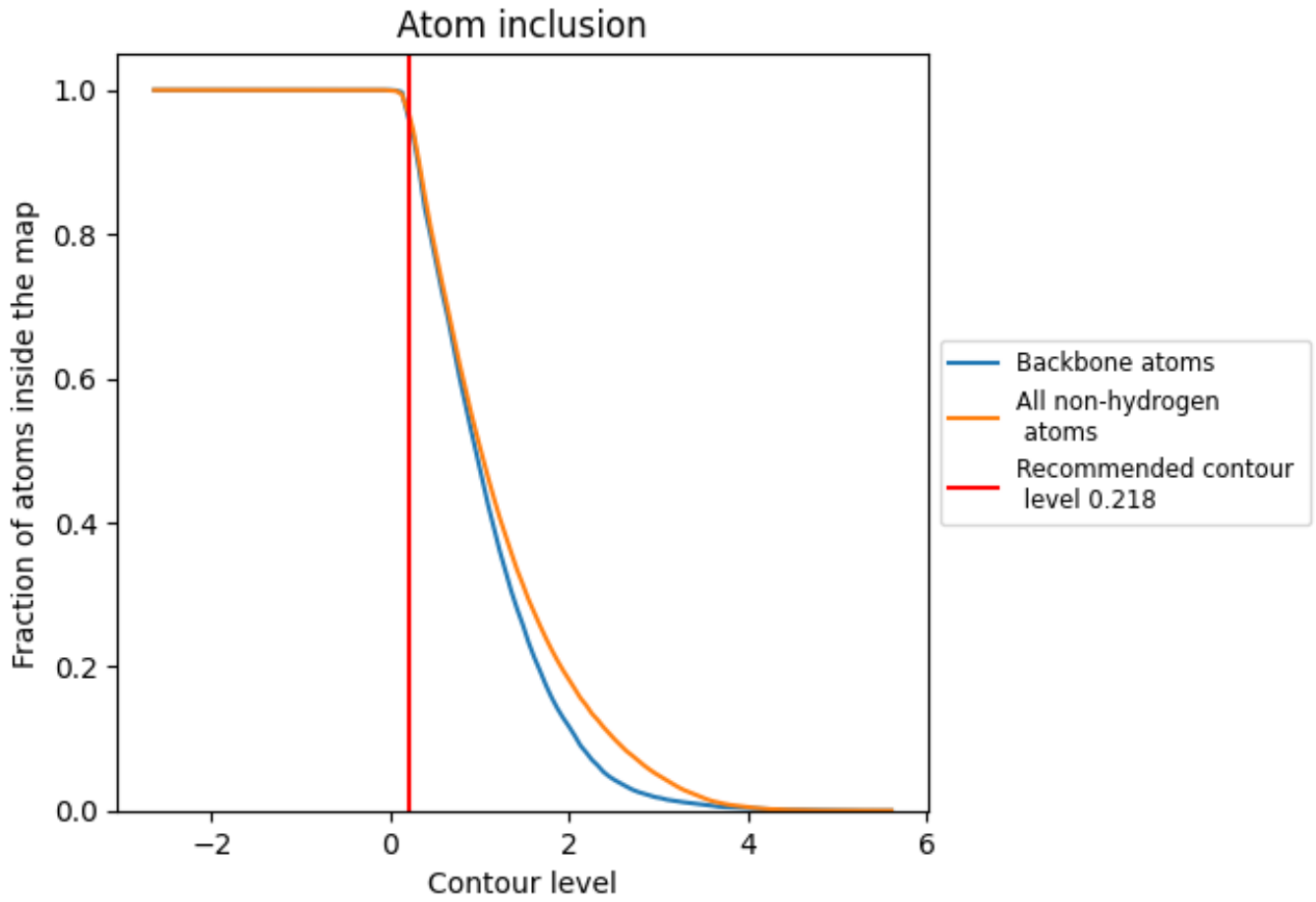
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.218 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, 96% of all backbone atoms, 96% of all non-hydrogen atoms, are inside the map.