wwPDB EM Validation Summary Report

Nov 7, 2024 – 12:39 PM EST

PDB ID	:	9DYG
EMDB ID	:	EMD-47303
Title	:	iSAT-PNA-RA20, RA20-B-c class
Authors	:	Sheng, K.; Dong, X.; Lee, J.
Deposited on	:	2024-10-14
Resolution	:	5.27 Å(reported)
Based on initial models	:	4YBB, 8RPY

This is a wwPDB EM Validation Summary Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/EMValidationReportHelp
with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

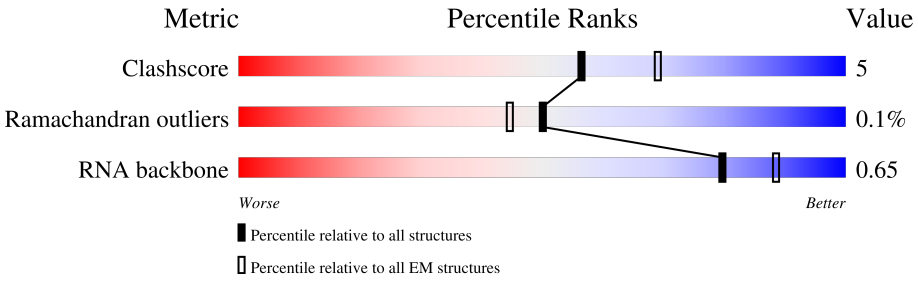
EMDB validation analysis	:	0.0.1.dev113
Mogul	:	2022.3.0, CSD as543be (2022)
MolProbity	:	4.02b-467
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
MapQ	:	1.9.13
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

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 5.27 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.



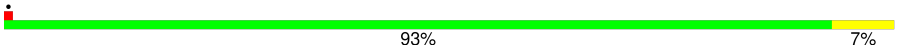
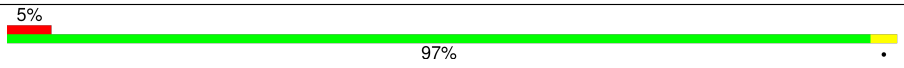
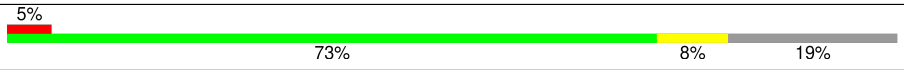
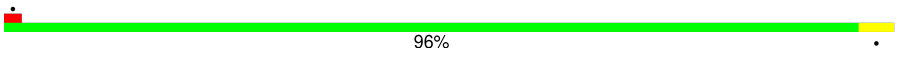
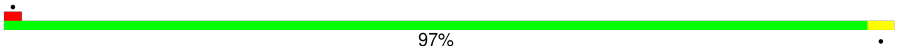
Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	210492	15764
Ramachandran outliers	207382	16835
RNA backbone	6643	2191

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	CA	1742	7% 75% 23%
2	2	37	97%
3	D	204	19% 66% 29%
4	E	201	5% 81% 17%
5	J	142	97%
6	K	110	15% 57% 42%
7	N	120	7% 88% 5% 7%
8	P	113	19% 91% 6%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
9	Q	117	
10	R	103	
11	S	110	
12	T	93	
13	U	102	
14	Y	63	

2 Entry composition (i)

There are 14 unique types of molecules in this entry. The entry contains 42760 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 RNA chain called 23S rRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
1	CA	1742	37444	16703	6941	12058	1742	0	0

- Molecule 2 is a protein called Large ribosomal subunit protein bL34.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
2	2	37	148	74	37	37	0	0

- Molecule 3 is a protein called Large ribosomal subunit protein uL3.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
3	D	144	576	288	144	144	0	0

- Molecule 4 is a protein called Large ribosomal subunit protein uL4.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
4	E	167	668	334	167	167	0	0

- Molecule 5 is a protein called Large ribosomal subunit protein uL13.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
5	J	142	568	284	142	142	0	0

- Molecule 6 is a protein called Large ribosomal subunit protein uL14.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
6	K	64	256	128	64	64	0	0

- Molecule 7 is a protein called Large ribosomal subunit protein bL17.

Mol	Chain	Residues	Atoms				AltConf	Trace
7	N	112	Total 448	C 224	N 112	O 112	0	0

- Molecule 8 is a protein called Large ribosomal subunit protein bL19.

Mol	Chain	Residues	Atoms				AltConf	Trace
8	P	106	Total 424	C 212	N 106	O 106	0	0

- Molecule 9 is a protein called Large ribosomal subunit protein bL20.

Mol	Chain	Residues	Atoms				AltConf	Trace
9	Q	117	Total 468	C 234	N 117	O 117	0	0

- Molecule 10 is a protein called Large ribosomal subunit protein bL21.

Mol	Chain	Residues	Atoms				AltConf	Trace
10	R	103	Total 412	C 206	N 103	O 103	0	0

- Molecule 11 is a protein called Large ribosomal subunit protein uL22.

Mol	Chain	Residues	Atoms				AltConf	Trace
11	S	97	Total 388	C 194	N 97	O 97	0	0

- Molecule 12 is a protein called Large ribosomal subunit protein uL23.

Mol	Chain	Residues	Atoms				AltConf	Trace
12	T	75	Total 300	C 150	N 75	O 75	0	0

- Molecule 13 is a protein called Large ribosomal subunit protein uL24.

Mol	Chain	Residues	Atoms				AltConf	Trace
13	U	102	Total 408	C 204	N 102	O 102	0	0

- Molecule 14 is a protein called Large ribosomal subunit protein uL29.

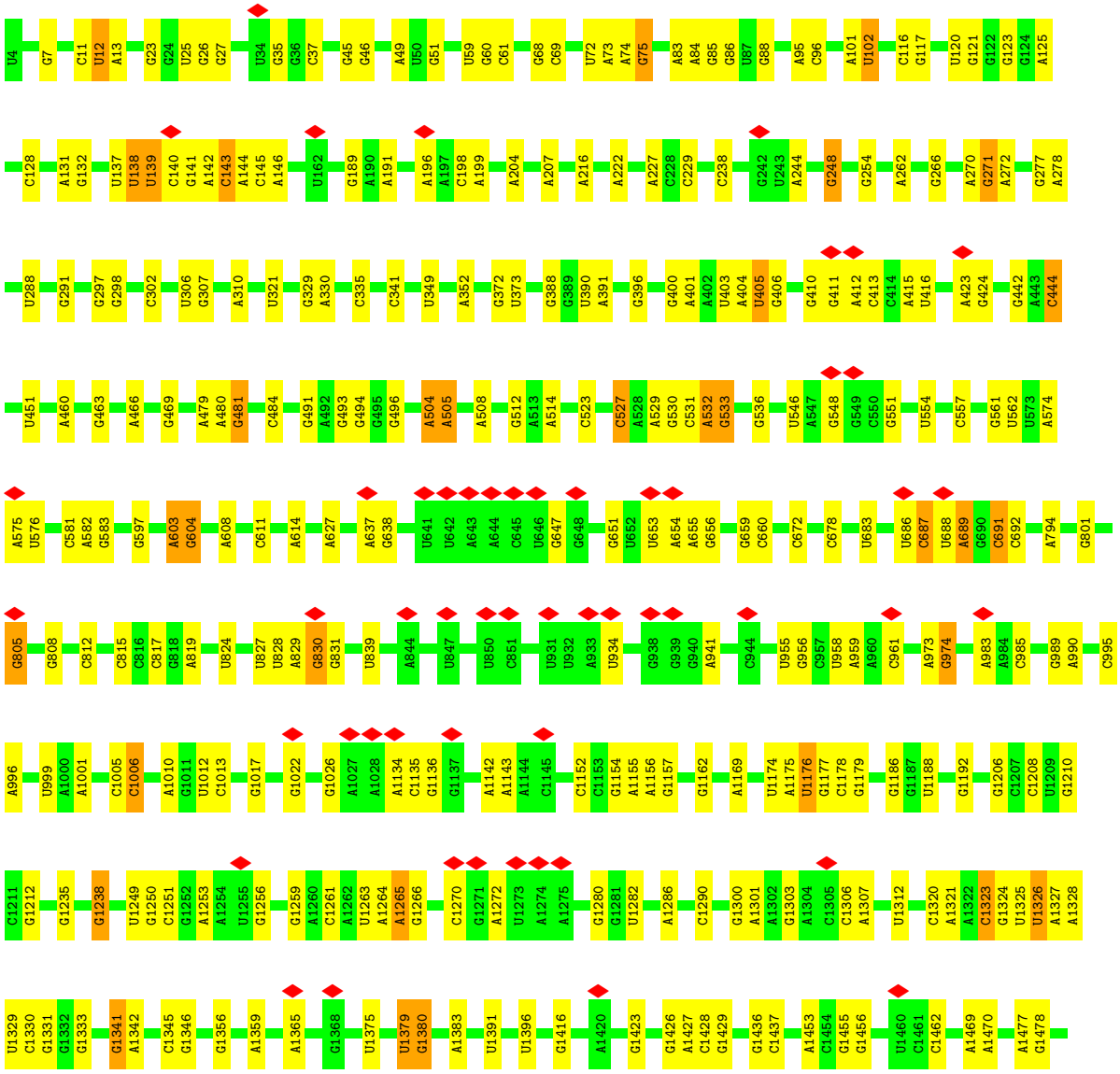
Mol	Chain	Residues	Atoms				AltConf	Trace
14	Y	63	Total 252	C 126	N 63	O 63	0	0

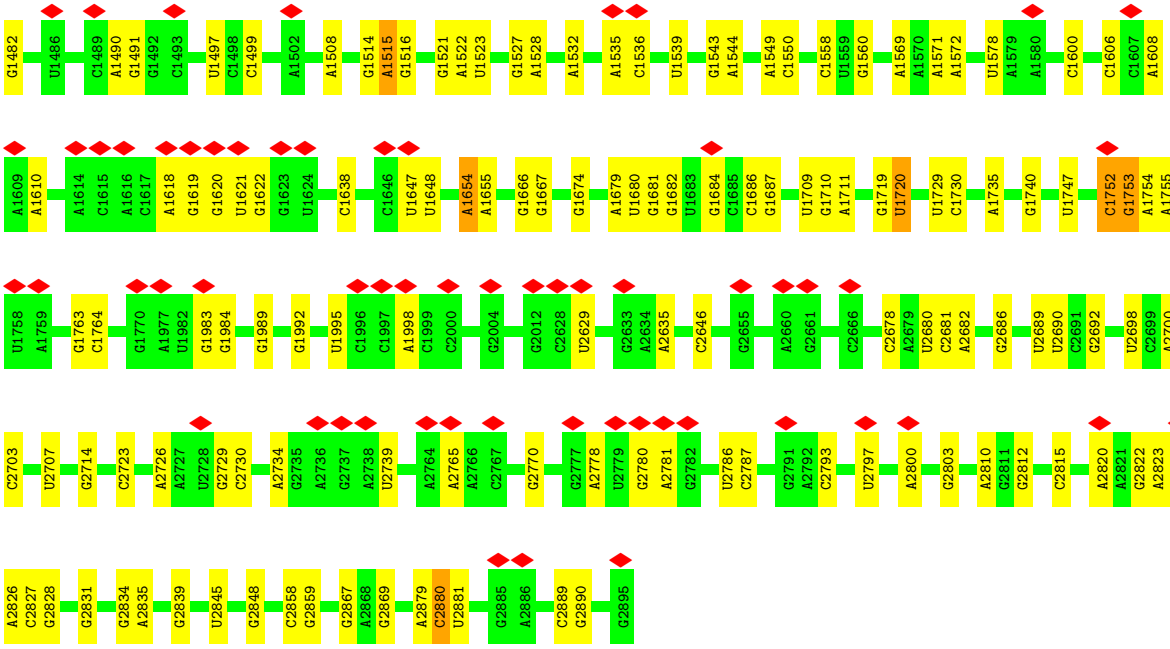
3 Residue-property plots (i)

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: 23S rRNA

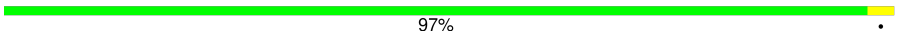
Chain CA:





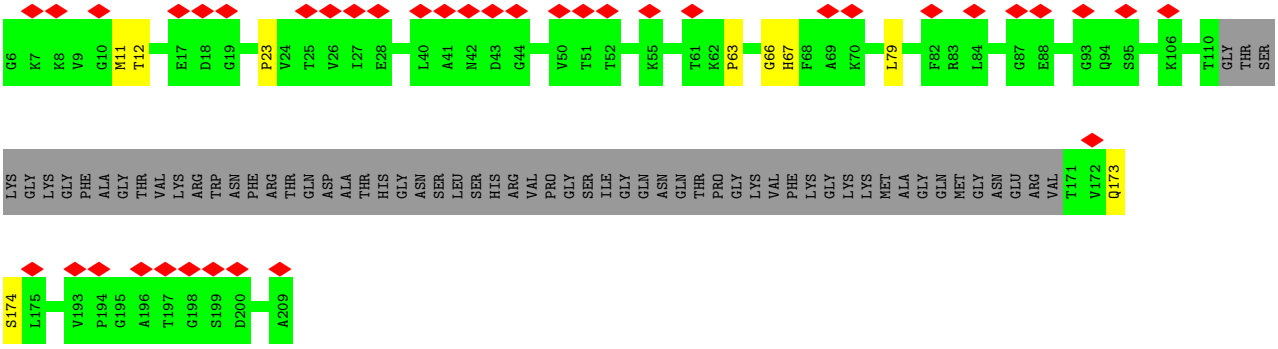
• Molecule 2: Large ribosomal subunit protein bL34

Chain 2:



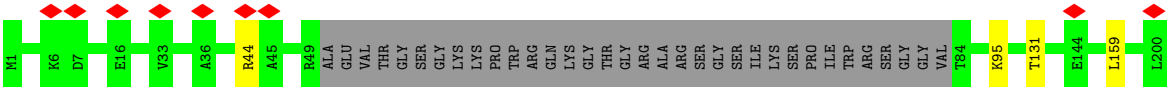
• Molecule 3: Large ribosomal subunit protein uL3

Chain D:



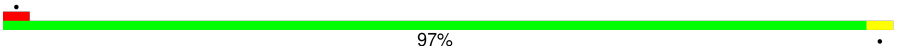
• Molecule 4: Large ribosomal subunit protein uL4

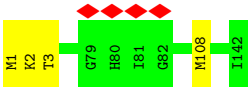
Chain E:



• Molecule 5: Large ribosomal subunit protein uL13

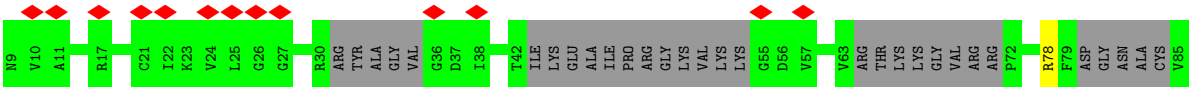
Chain J:

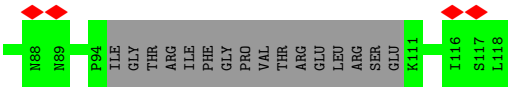




• Molecule 6: Large ribosomal subunit protein uL14

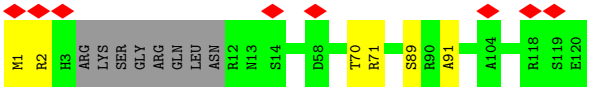
Chain K:





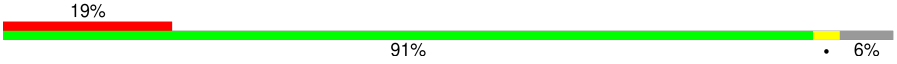
• Molecule 7: Large ribosomal subunit protein bL17

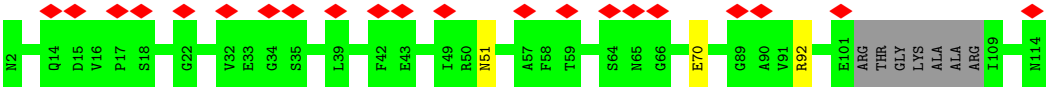
Chain N:



• Molecule 8: Large ribosomal subunit protein bL19

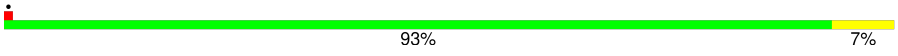
Chain P:

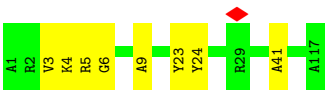




• Molecule 9: Large ribosomal subunit protein bL20

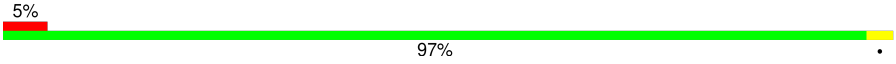
Chain Q:

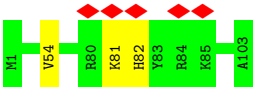




• Molecule 10: Large ribosomal subunit protein bL21

Chain R:



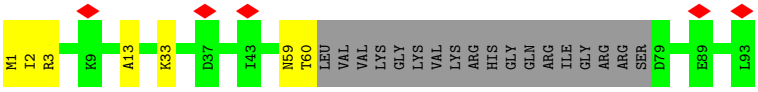


• Molecule 11: Large ribosomal subunit protein uL22

Chain S:

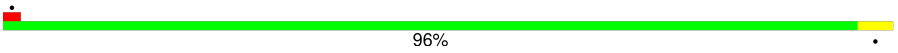
• Molecule 12: Large ribosomal subunit protein uL23

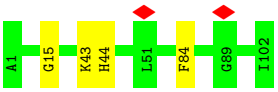
Chain T:



• Molecule 13: Large ribosomal subunit protein uL24

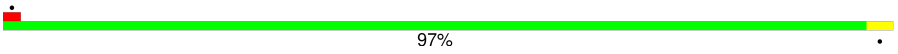
Chain U:

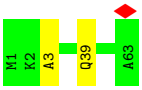




• Molecule 14: Large ribosomal subunit protein uL29

Chain Y:





4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	9000	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	NONE; CTF correction were performed by Patch CTF Estimation of cryoSPARC	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/Å^2$)	25.4	Depositor
Minimum defocus (nm)	1000	Depositor
Maximum defocus (nm)	2000	Depositor
Magnification	Not provided	
Image detector	GATAN K3 BIOQUANTUM (6k x 4k)	Depositor
Maximum map value	0.031	Depositor
Minimum map value	-0.009	Depositor
Average map value	0.000	Depositor
Map value standard deviation	0.001	Depositor
Recommended contour level	0.00642	Depositor
Map size (Å)	419.99997, 419.99997, 419.99997	wwPDB
Map dimensions	400, 400, 400	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.05, 1.05, 1.05	Depositor

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: 6MZ, PSU

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	CA	0.10	0/41892	0.64	0/65337
2	2	0.22	0/147	0.43	0/182
3	D	0.26	0/574	0.52	0/714
4	E	0.24	0/666	0.47	0/829
5	J	0.25	0/567	0.46	0/707
6	K	0.24	0/250	0.52	0/302
7	N	0.24	0/446	0.49	0/554
8	P	0.25	0/422	0.49	0/524
9	Q	0.23	0/467	0.41	0/582
10	R	0.26	0/411	0.55	0/512
11	S	0.25	0/386	0.48	0/479
12	T	0.25	0/298	0.47	0/369
13	U	0.25	0/407	0.53	0/507
14	Y	0.24	0/251	0.41	0/312
All	All	0.13	0/47184	0.63	0/71910

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

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	CA	37444	0	18839	244	0
2	2	148	0	42	1	0
3	D	576	0	166	12	0
4	E	668	0	177	7	0
5	J	568	0	162	13	0
6	K	256	0	69	1	0
7	N	448	0	121	15	0
8	P	424	0	115	3	0
9	Q	468	0	128	25	0
10	R	412	0	128	6	0
11	S	388	0	103	14	0
12	T	300	0	77	25	0
13	U	408	0	121	17	0
14	Y	252	0	69	2	0
All	All	42760	0	20317	247	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 247 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:CA:995:C:H41	5:J:2:LYS:C	1.06	1.57
1:CA:143:C:H4'	12:T:2:ILE:C	1.32	1.48
1:CA:137:U:H1'	12:T:1:MET:N	1.44	1.32
1:CA:973:A:H5"	10:R:81:LYS:CA	1.61	1.28
1:CA:1188:U:C5'	10:R:82:HIS:O	1.79	1.28

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	
2	2	35/37 (95%)	35 (100%)	0	0	100	100
3	D	140/204 (69%)	135 (96%)	5 (4%)	0	100	100
4	E	163/201 (81%)	160 (98%)	3 (2%)	0	100	100
5	J	140/142 (99%)	135 (96%)	5 (4%)	0	100	100
6	K	52/110 (47%)	49 (94%)	3 (6%)	0	100	100
7	N	108/120 (90%)	100 (93%)	8 (7%)	0	100	100
8	P	102/113 (90%)	98 (96%)	4 (4%)	0	100	100
9	Q	115/117 (98%)	112 (97%)	3 (3%)	0	100	100
10	R	101/103 (98%)	95 (94%)	5 (5%)	1 (1%)	13	49
11	S	93/110 (84%)	89 (96%)	4 (4%)	0	100	100
12	T	71/93 (76%)	65 (92%)	6 (8%)	0	100	100
13	U	100/102 (98%)	89 (89%)	11 (11%)	0	100	100
14	Y	61/63 (97%)	59 (97%)	2 (3%)	0	100	100
All	All	1281/1515 (85%)	1221 (95%)	59 (5%)	1 (0%)	50	83

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
10	R	54	VAL

5.3.2 Protein sidechains (i)

There are no protein residues with a non-rotameric sidechain to report in this entry.

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	CA	1729/1742 (99%)	230 (13%)	9 (0%)

5 of 230 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	CA	7	G
1	CA	12	U
1	CA	13	A
1	CA	35	G
1	CA	49	A

5 of 9 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
1	CA	2680	U
1	CA	2858	C
1	CA	1379	U
1	CA	1621	U
1	CA	1680	U

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

2 non-standard protein/DNA/RNA residues are modelled in this entry.

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	$\#\|Z\| > 2$
1	PSU	CA	955	1	18,21,22	1.15	1 (5%)	21,30,33	1.91	5 (23%)
1	6MZ	CA	1618	1	17,25,26	1.34	2 (11%)	15,36,39	2.87	5 (33%)

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. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
1	PSU	CA	955	1	-	0/7/25/26	0/2/2/2
1	6MZ	CA	1618	1	-	5/5/27/28	0/3/3/3

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	CA	1618	6MZ	C6-C5	-4.24	1.38	1.44
1	CA	955	PSU	C6-C5	3.88	1.39	1.35
1	CA	1618	6MZ	C2-N3	2.33	1.35	1.32

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	CA	1618	6MZ	C1'-N9-C4	-6.63	115.00	126.64
1	CA	1618	6MZ	N3-C2-N1	-6.35	120.05	128.67
1	CA	955	PSU	N1-C2-N3	4.74	120.17	115.17
1	CA	955	PSU	C4-N3-C2	-4.72	119.88	126.37
1	CA	1618	6MZ	C2-N1-C6	4.35	119.98	116.60

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	CA	1618	6MZ	C5-C6-N6-C9
1	CA	1618	6MZ	N1-C6-N6-C9
1	CA	1618	6MZ	O4'-C4'-C5'-O5'
1	CA	1618	6MZ	C3'-C4'-C5'-O5'
1	CA	1618	6MZ	C4'-C5'-O5'-P

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates (i)

There are no oligosaccharides 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)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	CA	10

The worst 5 of 10 chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	CA	2012:G	O3'	2628:C	P	60.73
1	CA	1028:A	O3'	1134:A	P	23.55
1	CA	2742:G	O3'	2763:G	P	19.29
1	CA	851:C	O3'	926:G	P	17.03
1	CA	1977:A	O3'	1982:U	P	13.42

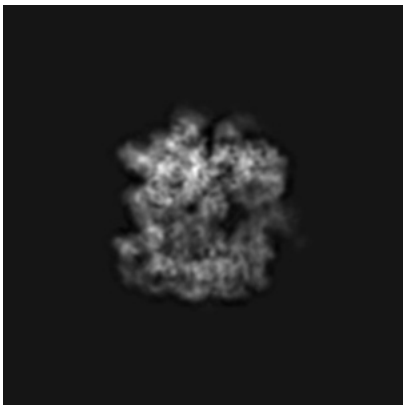
6 Map visualisation

This section contains visualisations of the EMDB entry EMD-47303. These allow visual inspection of the internal detail of the map and identification of artifacts.

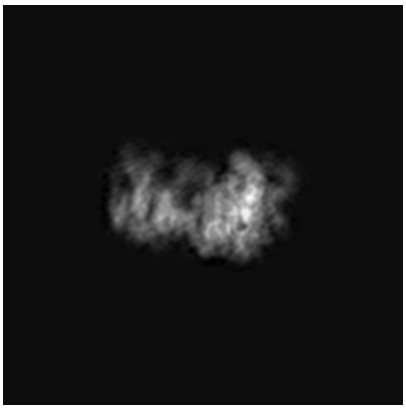
Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

6.1 Orthogonal projections

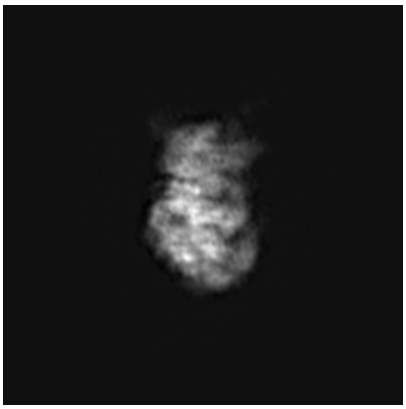
6.1.1 Primary map



X

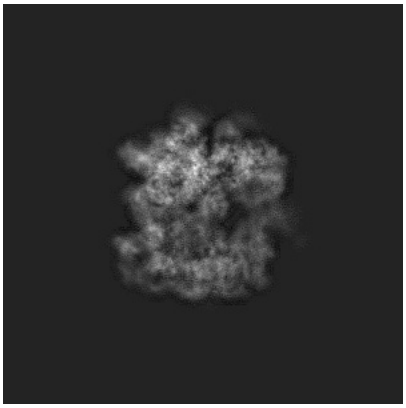


Y

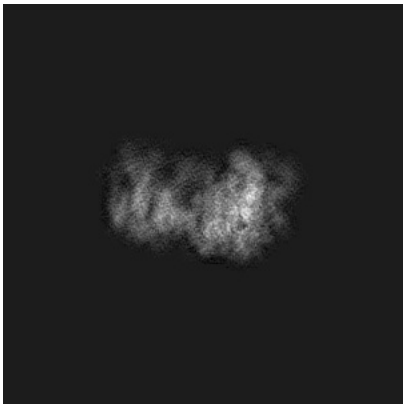


Z

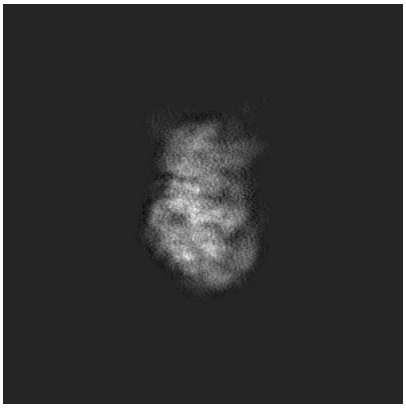
6.1.2 Raw map



X



Y

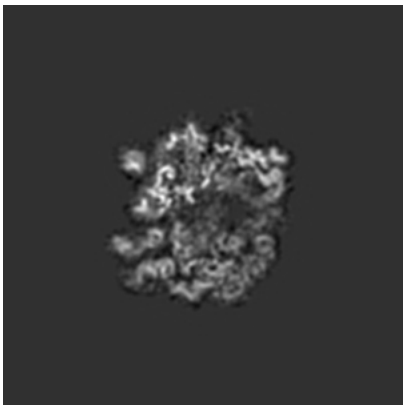


Z

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

6.2 Central slices (i)

6.2.1 Primary map

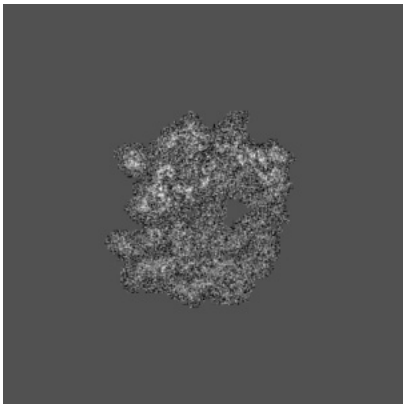


X Index: 200

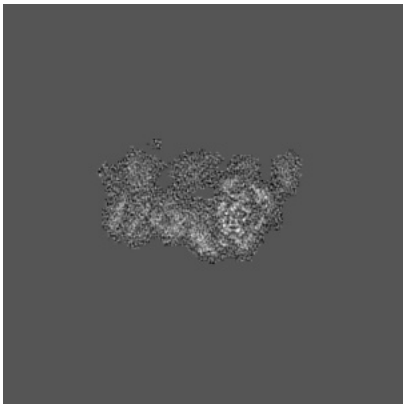
Y Index: 200

Z Index: 200

6.2.2 Raw map



X Index: 200



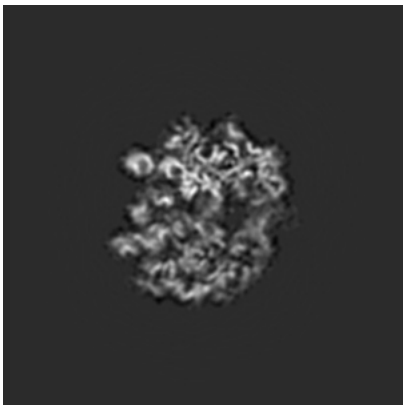
Y Index: 200

Z Index: 200

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

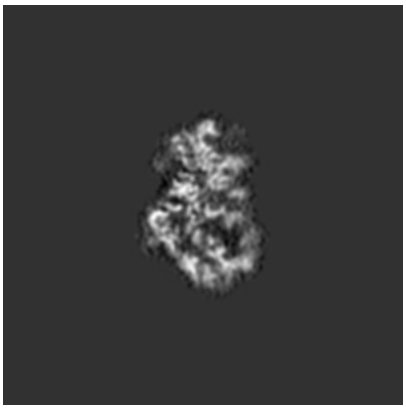
6.3 Largest variance slices (i)

6.3.1 Primary map



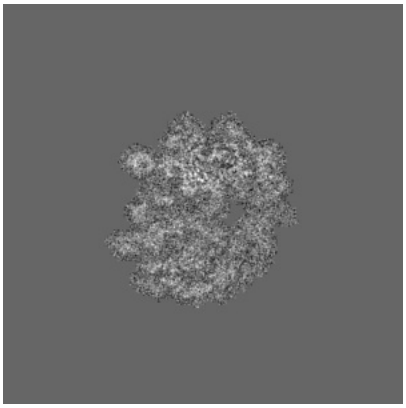
X Index: 192

Y Index: 157



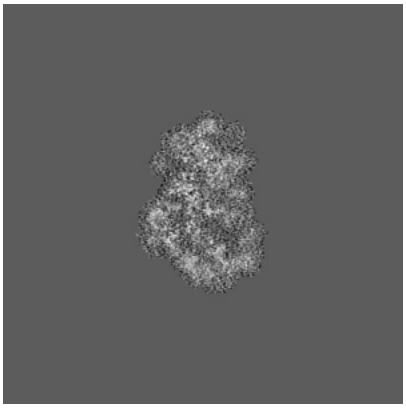
Z Index: 245

6.3.2 Raw map



X Index: 191

Y Index: 192

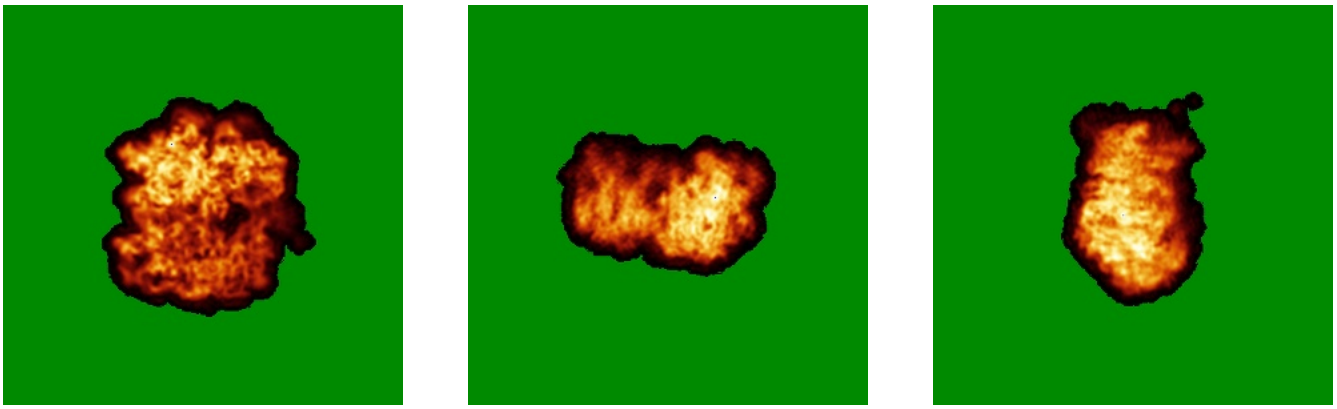


Z Index: 244

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

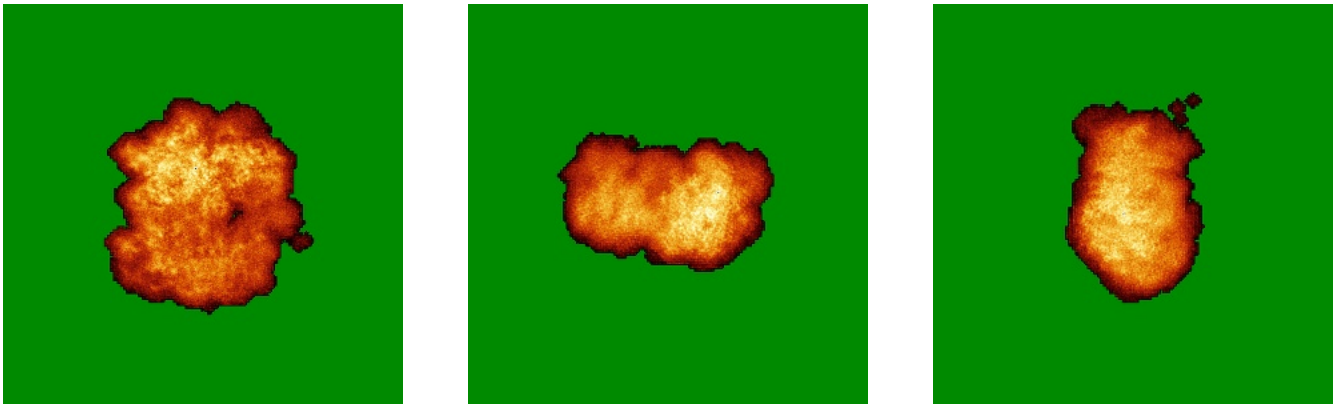
6.4 Orthogonal standard-deviation projections (False-color) (i)

6.4.1 Primary map



X Y Z

6.4.2 Raw map

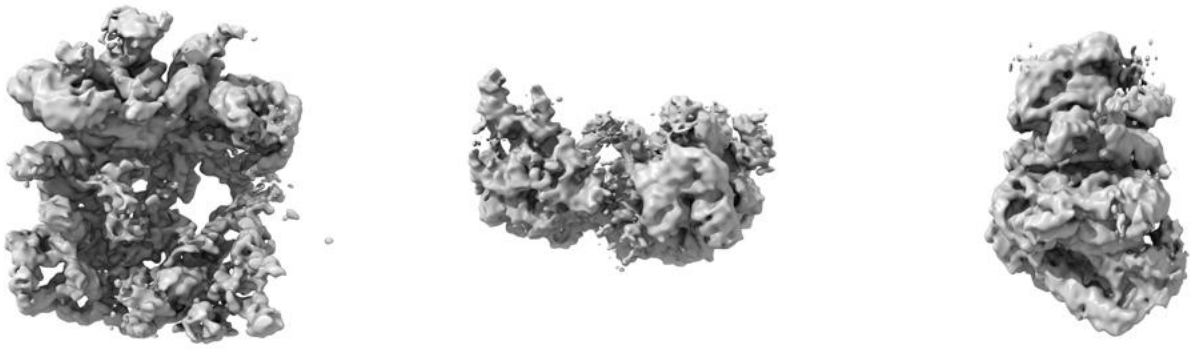


X Y Z

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

6.5 Orthogonal surface views (i)

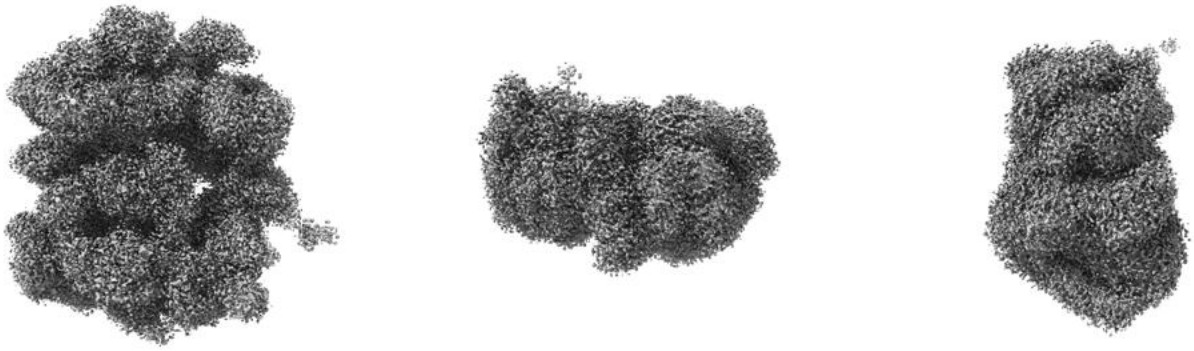
6.5.1 Primary map



X Y Z

The images above show the 3D surface view of the map at the recommended contour level 0.00642. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

6.5.2 Raw map



X Y Z

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

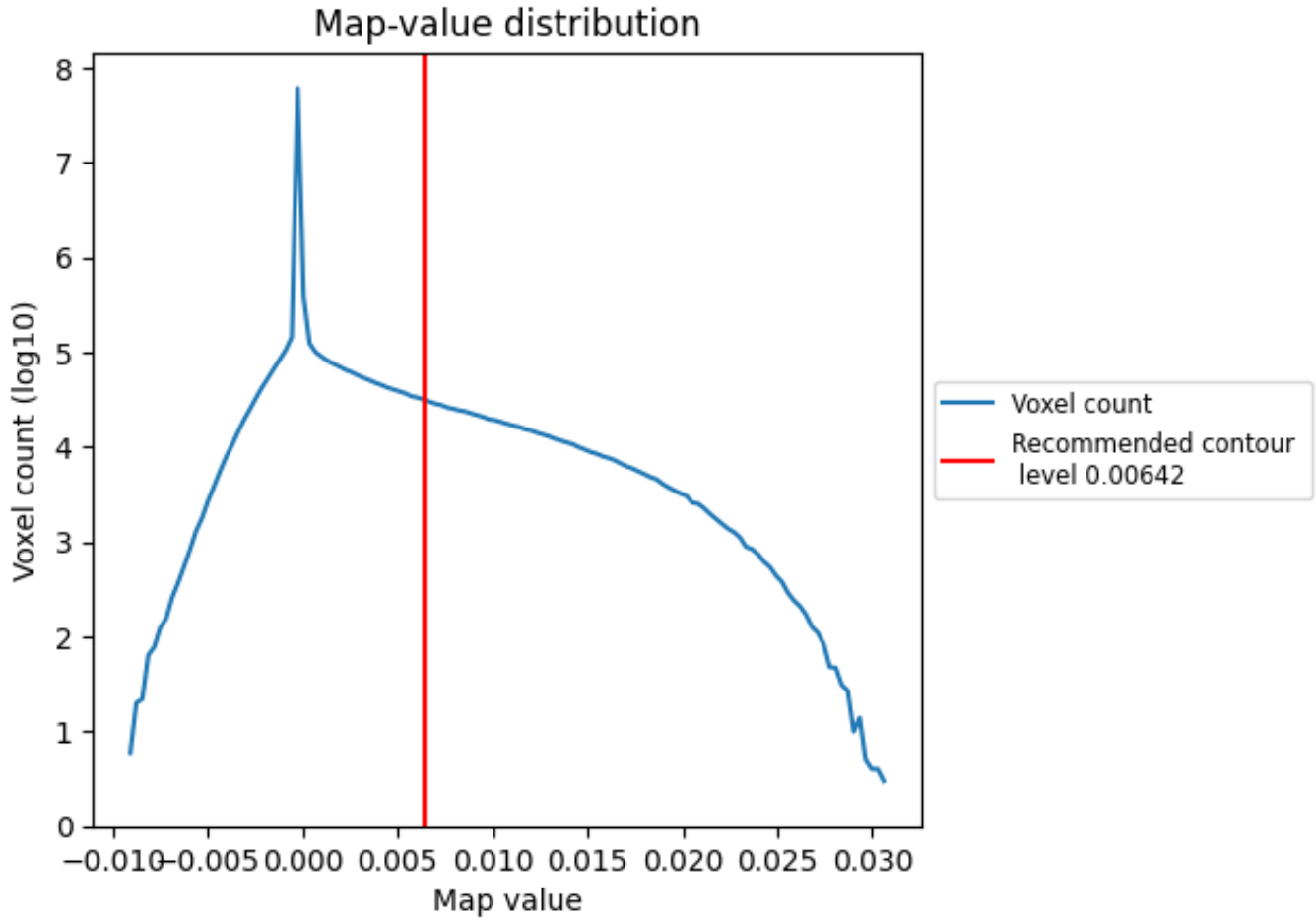
6.6 Mask visualisation (i)

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

7 Map analysis

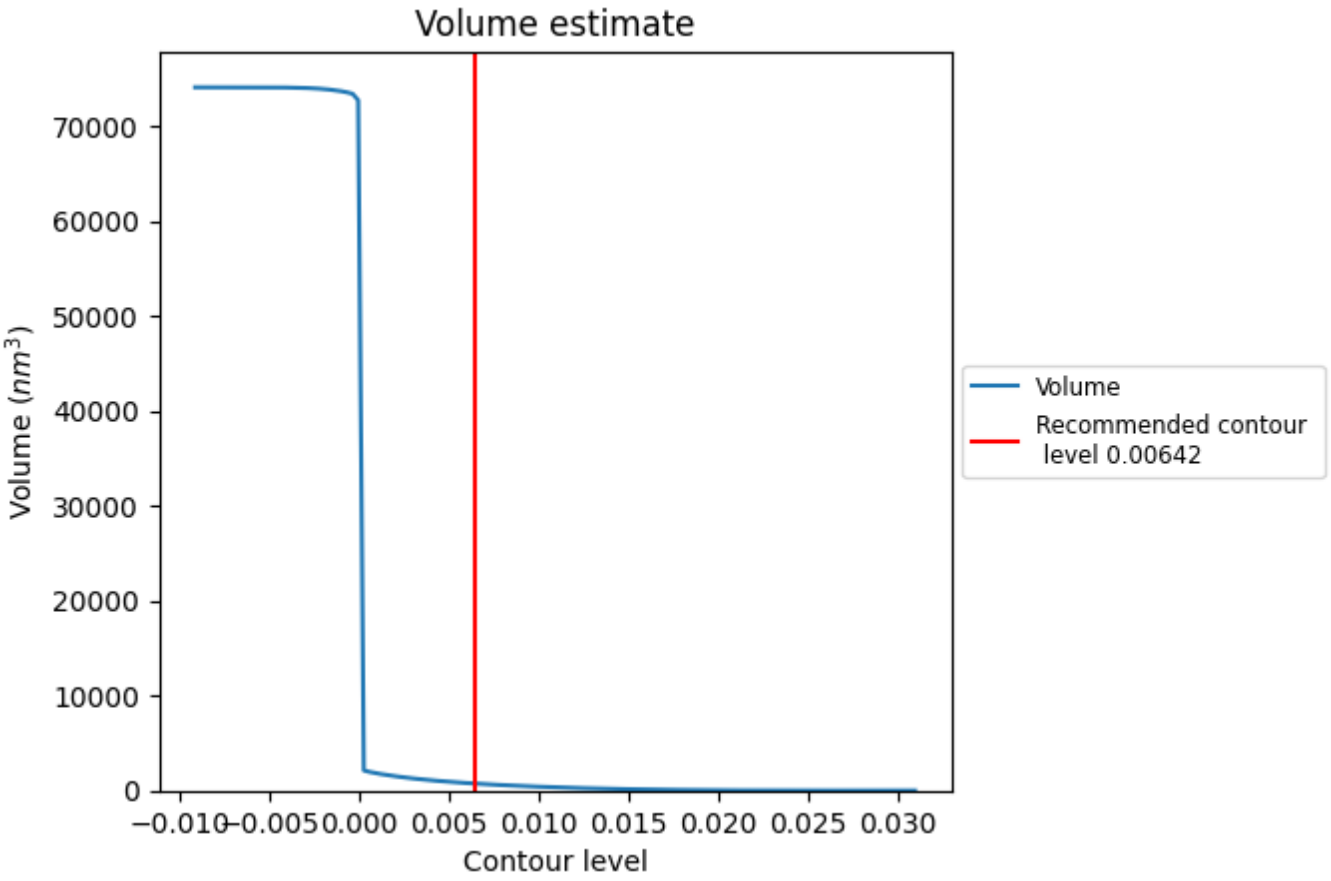
This section contains the results of statistical analysis of the map.

7.1 Map-value distribution



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

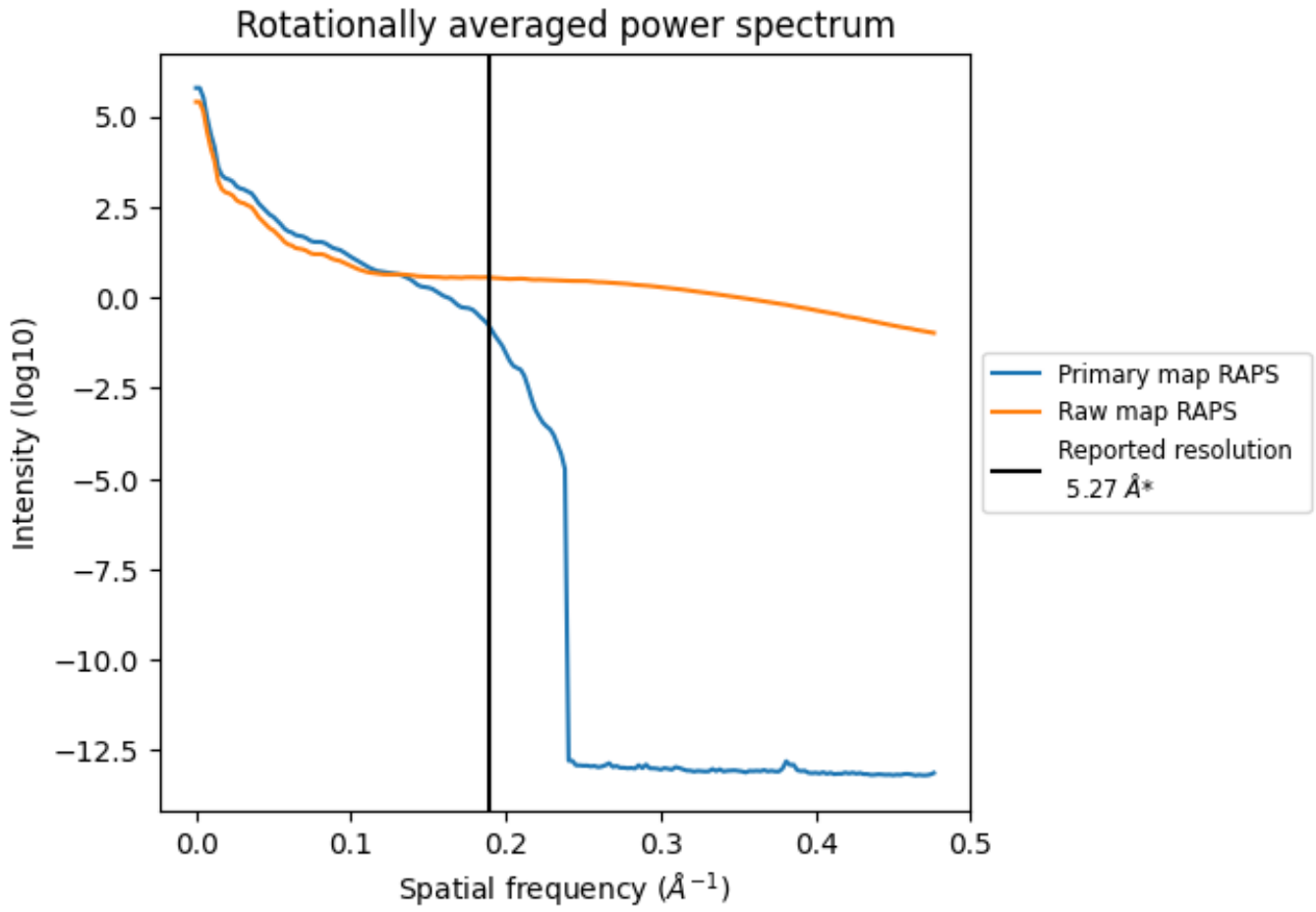
7.2 Volume estimate (i)



The volume at the recommended contour level is 740 nm^3; this corresponds to an approximate mass of 669 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)

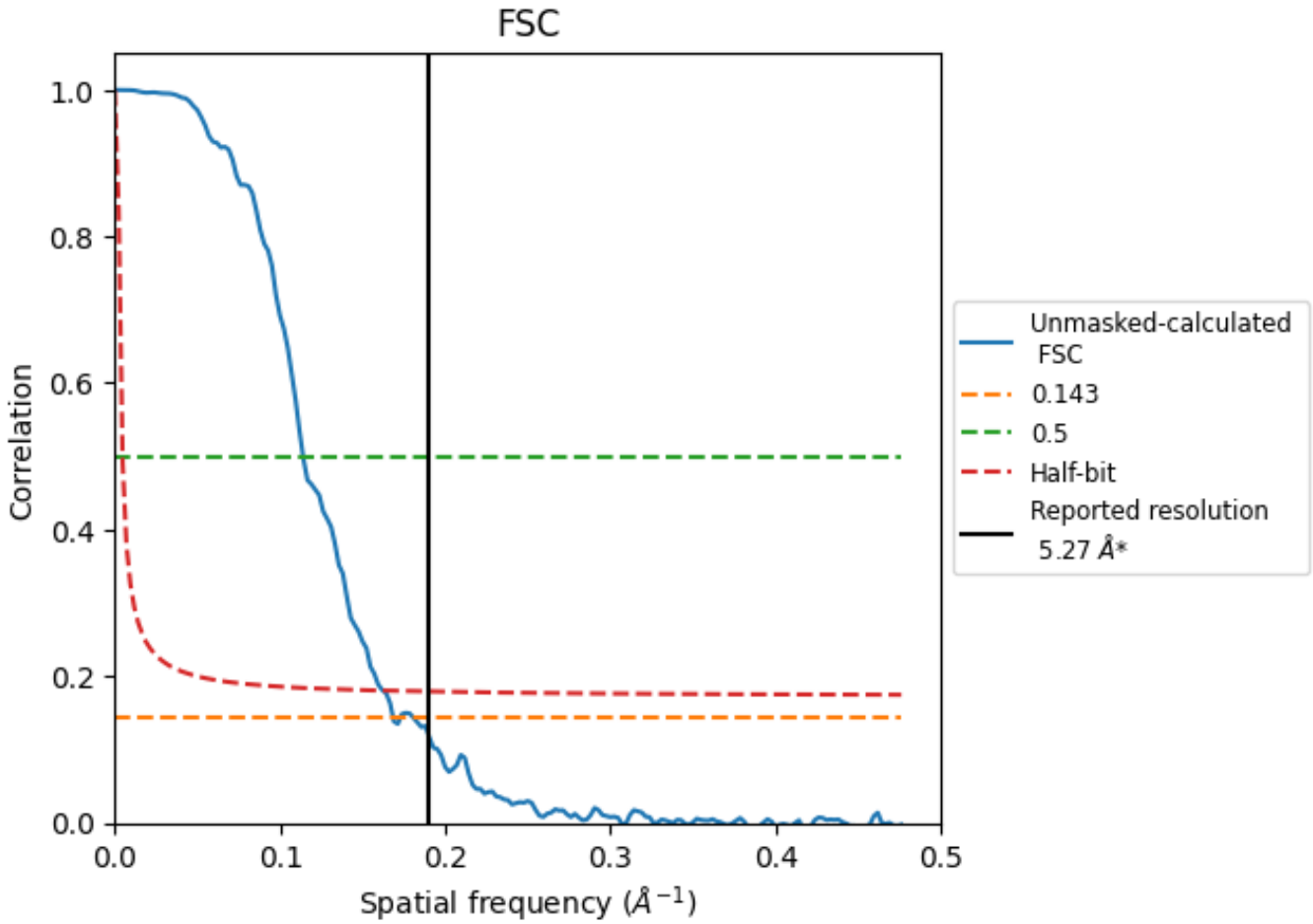


*Reported resolution corresponds to spatial frequency of 0.190 $Å^{-1}$

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 0.190 Å^{-1}

8.2 Resolution estimates (i)

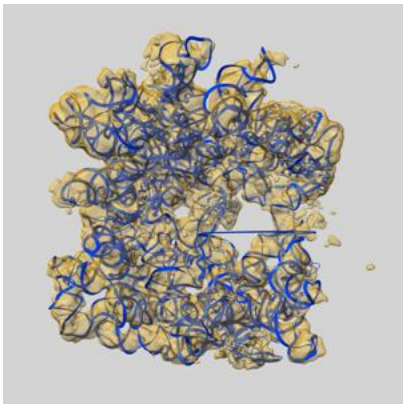
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	5.27	-	-
Author-provided FSC curve	-	-	-
Unmasked-calculated*	5.93	8.77	6.16

*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 5.93 differs from the reported value 5.27 by more than 10 %

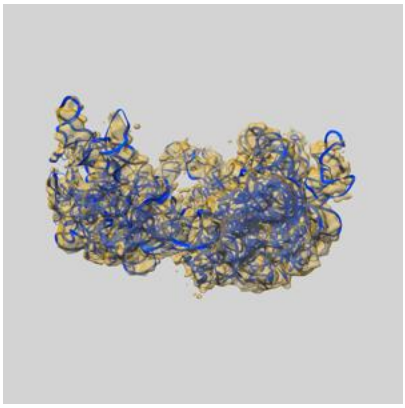
9 Map-model fit

This section contains information regarding the fit between EMDB map EMD-47303 and PDB model 9DYG. Per-residue inclusion information can be found in section 3 on page 7.

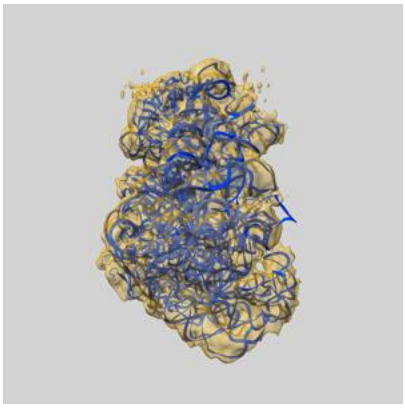
9.1 Map-model overlay



X



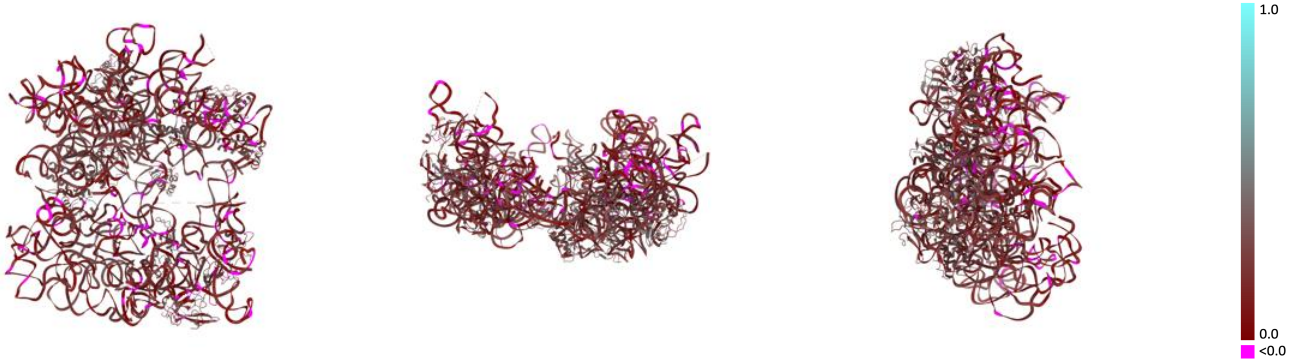
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.00642 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.

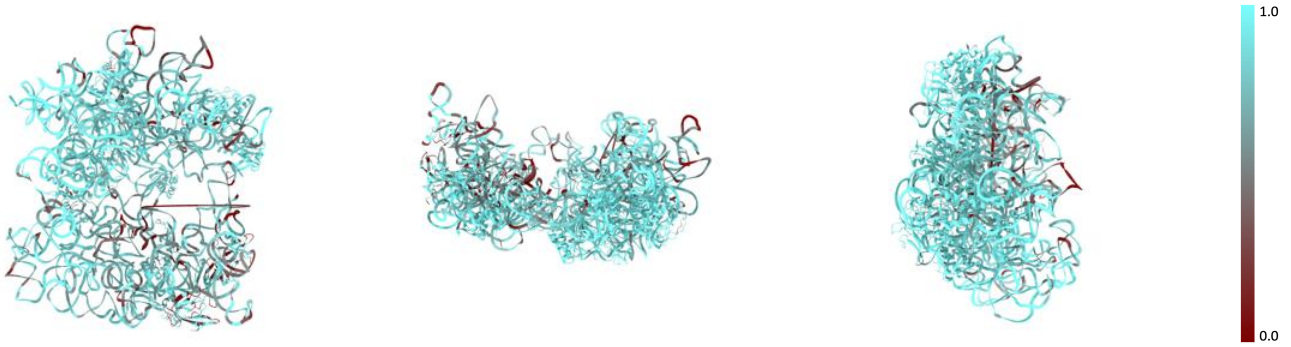
9.2 Q-score mapped to coordinate model (i)



X Y Z

The images above show the model with each residue coloured according its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions at which atoms can be resolved. Low Q-score values may therefore be expected for many entries.

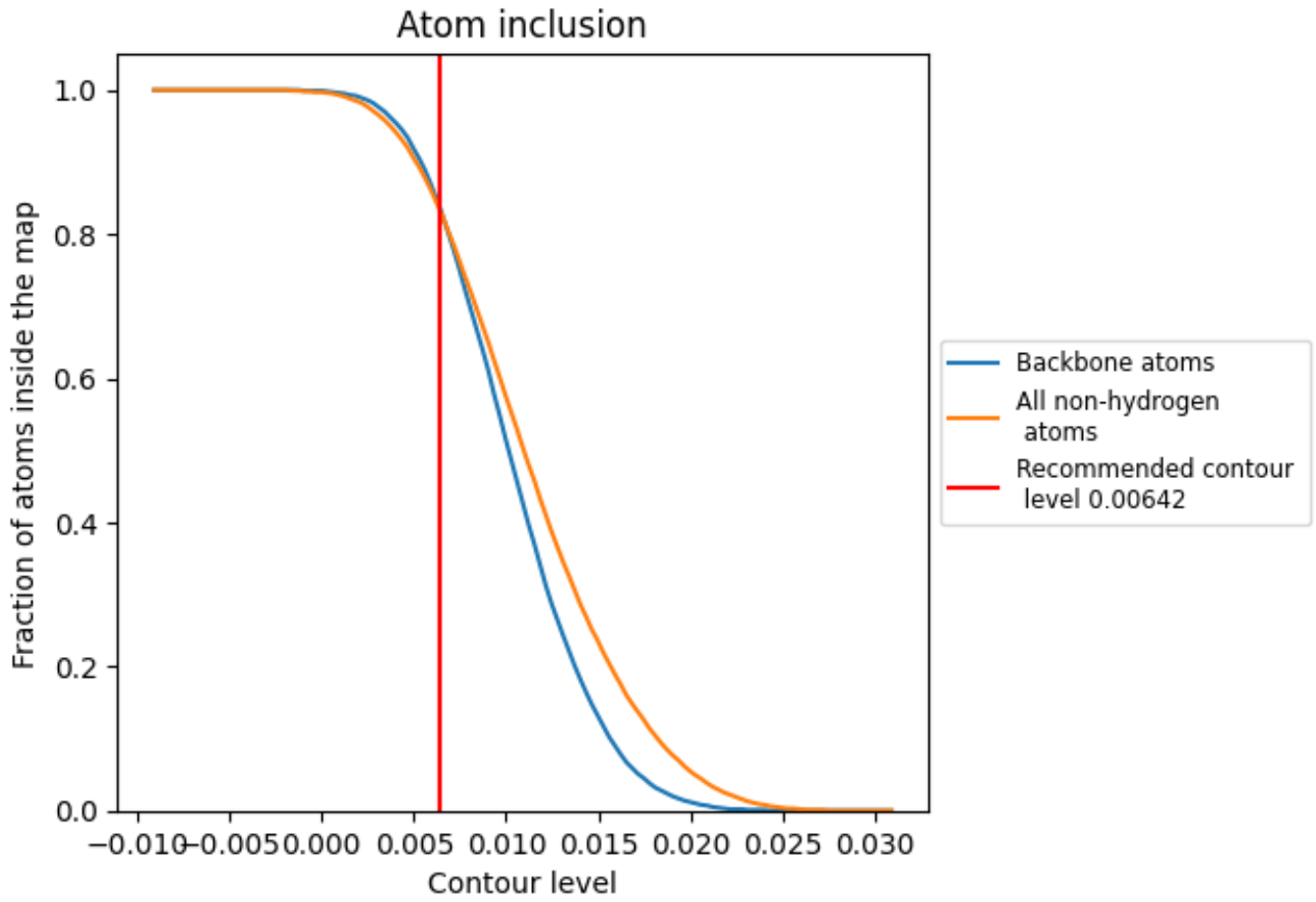
9.3 Atom inclusion mapped to coordinate model (i)



X Y Z

The images above show the model with each residue coloured according to its atom inclusion. This shows to what extent they are inside the map at the recommended contour level (0.00642).

9.4 Atom inclusion (i)



At the recommended contour level, 84% of all backbone atoms, 83% of all non-hydrogen atoms, are inside the map.

9.5 Map-model fit summary

The table lists the average atom inclusion at the recommended contour level (0.00642) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.8330	0.1730
2	0.9930	0.2960
CA	0.8270	0.1600
D	0.6540	0.2160
E	0.8920	0.2530
J	0.9440	0.2730
K	0.6090	0.2230
N	0.9020	0.2480
P	0.7360	0.2750
Q	0.9700	0.2640
R	0.9340	0.2980
S	0.9720	0.2760
T	0.9000	0.2380
U	0.9560	0.2760
Y	0.9800	0.2890