wwPDB EM Validation Summary Report (i)

May 13, 2025 – 01:06 PM EDT

PDB ID	:	8VFY / pdb_00008vfy
EMDB ID	:	EMD-43194
Title	:	Cryo-EM structure of FoxA1 in complex with ALBN1 nucleosome (class 1)
Authors	:	Zhou, B.R.; Bai, Y.
Deposited on	:	2023-12-22
Resolution	:	2.89 Å(reported)

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

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/EMValidationReportHelp
with specific help available everywhere you see the (i) 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 (i)) were used in the production of this report:

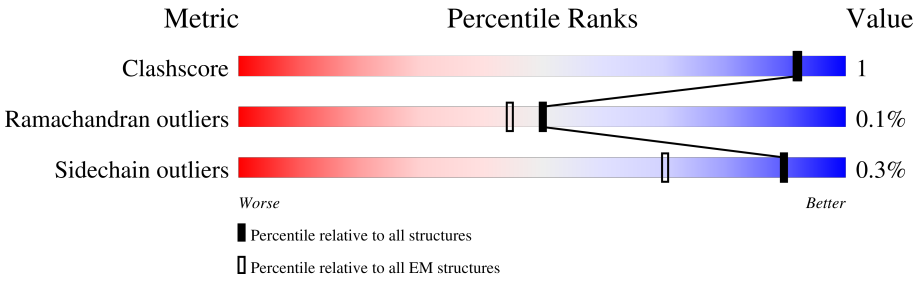
EMDB validation analysis	:	0.0.1.dev118
MolProbity	:	4-5-2 with Phenix2.0rc1
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
MapQ	:	1.9.13
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.43.1

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 2.89 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.



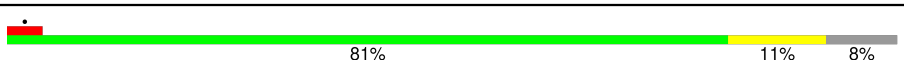
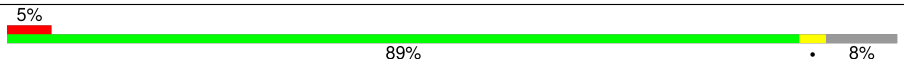
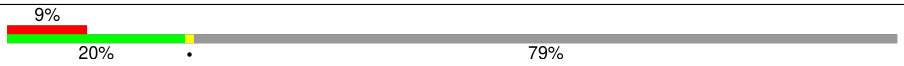
Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	210492	15764
Ramachandran outliers	207382	16835
Sidechain outliers	206894	16415

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	136	70% · 29%
1	E	136	68% · 29%
2	B	103	79% · 19%
2	F	103	77% · 19%
3	C	130	78% 6% 16%
3	G	130	88% · 8%
4	D	126	69% 6% 25%
4	H	126	71% 5% 25%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
5	I	186	
6	J	186	
7	O	478	

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 14046 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 Histone H3.1.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	97	Total	C	N	O	S	0	0
			799	505	153	137	4		
1	E	97	Total	C	N	O	S	0	0
			799	505	153	137	4		

- Molecule 2 is a protein called Histone H4.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	83	Total	C	N	O	S	0	0
			662	418	129	114	1		
2	F	83	Total	C	N	O	S	0	0
			662	418	129	114	1		

- Molecule 3 is a protein called Histone H2A type 1-B/E.

Mol	Chain	Residues	Atoms				AltConf	Trace
3	C	109	Total	C	N	O	0	0
			840	529	166	145		
3	G	120	Total	C	N	O	0	0
			927	582	185	160		

- Molecule 4 is a protein called Histone H2B type 1-J.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	D	95	Total	C	N	O	S	0	0
			745	468	136	139	2		
4	H	95	Total	C	N	O	S	0	0
			745	468	136	139	2		

- Molecule 5 is a DNA chain called DNA (171-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
5	I	171	Total	C	N	O	P	0	0
			3501	1673	619	1038	171		

- Molecule 6 is a DNA chain called DNA (171-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
6	J	171	Total	C	N	O	P	0	0
			3510	1672	653	1014	171		

- Molecule 7 is a protein called Hepatocyte nuclear factor 3-alpha.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	O	102	Total	C	N	O	S	0	0
			856	548	152	149	7		

There are 6 discrepancies between the modelled and reference sequences:

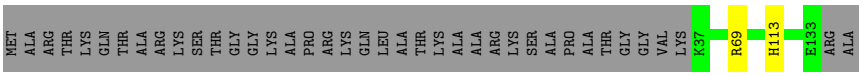
Chain	Residue	Modelled	Actual	Comment	Reference
O	473	HIS	-	expression tag	UNP P55317
O	474	HIS	-	expression tag	UNP P55317
O	475	HIS	-	expression tag	UNP P55317
O	476	HIS	-	expression tag	UNP P55317
O	477	HIS	-	expression tag	UNP P55317
O	478	HIS	-	expression tag	UNP P55317

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: Histone H3.1

Chain A:

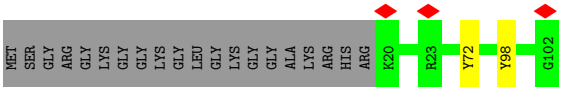


- Molecule 1: Histone H3.1

Chain E:

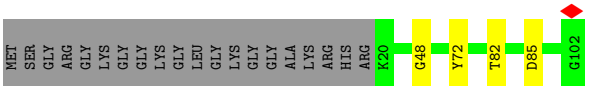
- Molecule 2: Histone H4

Chain B:



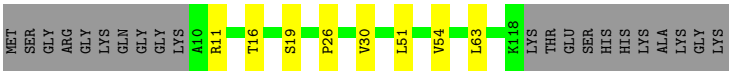
- Molecule 2: Histone H4

Chain F:



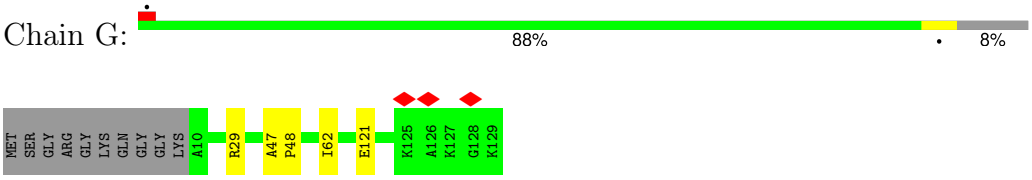
- Molecule 3: Histone H2A type 1-B/E

Chain C:



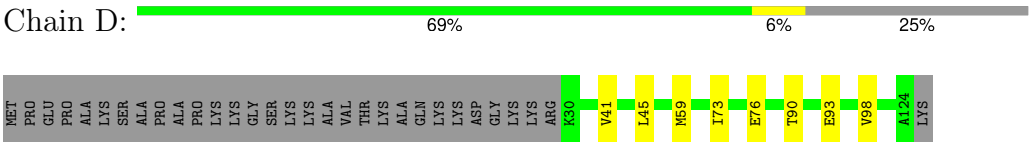
- Molecule 3: Histone H2A type 1-B/E

Chain G:



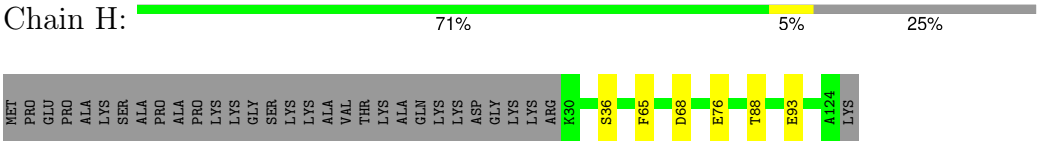
• Molecule 4: Histone H2B type 1-J

Chain D:



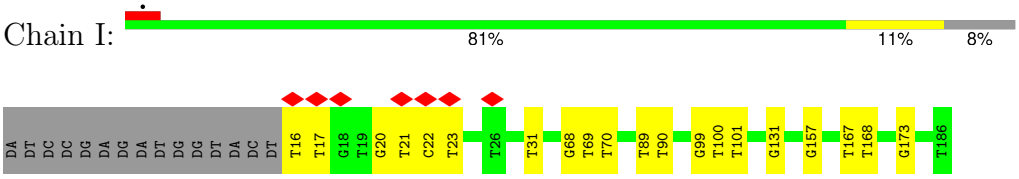
• Molecule 4: Histone H2B type 1-J

Chain H:



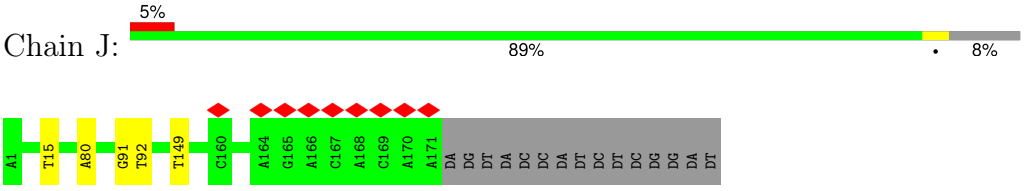
• Molecule 5: DNA (171-MER)

Chain I:



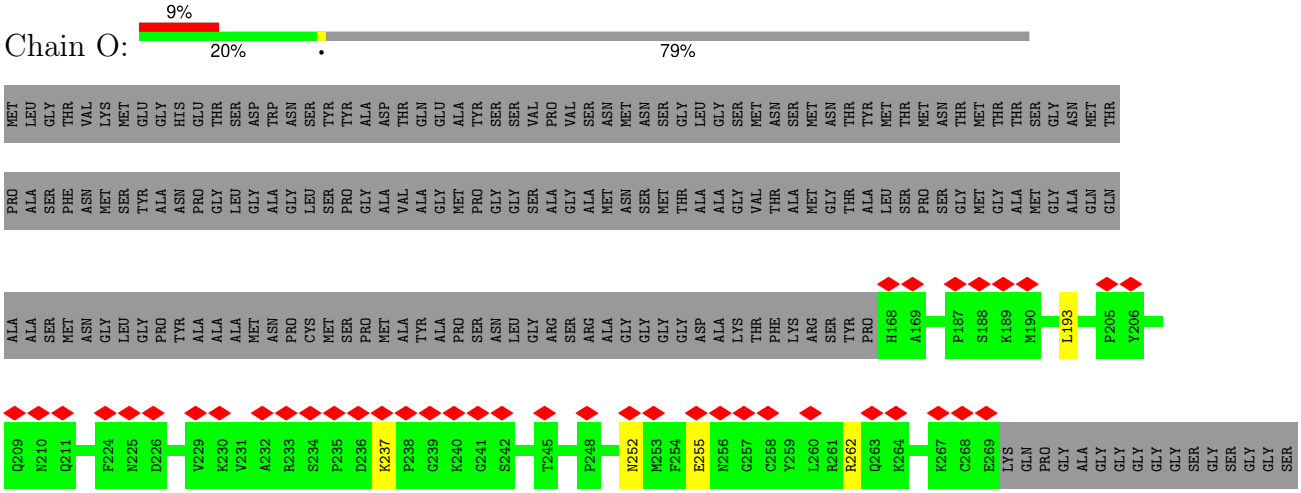
• Molecule 6: DNA (171-MER)

Chain J:



• Molecule 7: Hepatocyte nuclear factor 3-alpha

Chain O:



GLY ALA LYS GLY GLY PRO GLU SER ARG LYS ASP PRO SER GLY ALA SER ASN PRO SER ALA ASP SER PRO LEU HIS ARG GLY VAL HIS GLY LYS THR GLY GLN LEU GLU GLY ALA PRO ALA PRO GLY PRO ALA ALA SER PRO GLN THR LEU ASP HIS SER GLY ALA THR ALA THR GLY GLY

ALA SER GLU LEU LYS THR PRO ALA SER SER THR ALA PRO PRO ILE SER SER GLY PRO GLY ALA LEU ALA SER VAL PRO ALA SER HIS PRO ALA HIS GLY LEU ALA PRO HIS GLU SER GLN LEU HIS LEU LYS GLY ASP PRO HIS TYR SER PHE ASN HIS PRO PHE SER ILE ASN ASN LEU

MET SER SER SER GLU GLN GLN HIS LYS LEU ASP PHE LYS ALA TYR GLU GLN ALA LEU GLN TYR SER PRO TYR GLY SER THR LEU PRO ALA SER LEU PRO LEU GLY SER ALA SER VAL THR THR ARG SER PRO ILE GLU PRO SER ALA LEU GLU PRO ALA TYR TYR GLN GLY VAL TYR SER

ARG PRO VAL LEU ASN THR SER HIS HIS HIS HIS HIS HIS

4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	231331	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	NONE	Depositor
Microscope	TFS KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/Å^2$)	1	Depositor
Minimum defocus (nm)	800	Depositor
Maximum defocus (nm)	1800	Depositor
Magnification	Not provided	
Image detector	GATAN K3 (6k x 4k)	Depositor
Maximum map value	0.057	Depositor
Minimum map value	-0.023	Depositor
Average map value	0.000	Depositor
Map value standard deviation	0.001	Depositor
Recommended contour level	0.007	Depositor
Map size (Å)	337.92, 337.92, 337.92	wwPDB
Map dimensions	256, 256, 256	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.32, 1.32, 1.32	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	A	0.17	0/811	0.31	0/1088
1	E	0.17	0/811	0.30	0/1088
2	B	0.18	0/669	0.33	0/894
2	F	0.16	0/669	0.30	0/894
3	C	0.16	0/850	0.28	0/1146
3	G	0.15	0/939	0.29	0/1262
4	D	0.16	0/756	0.29	0/1015
4	H	0.15	0/756	0.26	0/1015
5	I	0.25	0/3922	0.52	0/6052
6	J	0.25	0/3942	0.49	0/6081
7	O	0.16	0/883	0.37	0/1192
All	All	0.21	0/15008	0.43	0/21727

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	A	799	0	838	2	0
1	E	799	0	838	2	0
2	B	662	0	709	2	0
2	F	662	0	709	3	0
3	C	840	0	902	6	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	G	927	0	994	4	0
4	D	745	0	771	7	0
4	H	745	0	771	5	0
5	I	3501	0	1936	13	0
6	J	3510	0	1924	4	0
7	O	856	0	832	3	0
All	All	14046	0	11224	36	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 1.

The worst 5 of 36 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)
3:G:121:GLU:N	3:G:121:GLU:OE1	2.35	0.59
4:D:90:THR:OG1	4:D:93:GLU:OE1	2.20	0.58
1:A:69:ARG:NH1	5:I:131:DG:OP2	2.36	0.58
1:E:116:ARG:NH1	1:E:118:THR:O	2.38	0.57
3:C:16:THR:O	3:C:19:SER:OG	2.21	0.55

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all EM entries.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	95/136 (70%)	94 (99%)	1 (1%)	0	100	100
1	E	95/136 (70%)	94 (99%)	1 (1%)	0	100	100
2	B	81/103 (79%)	81 (100%)	0	0	100	100
2	F	81/103 (79%)	80 (99%)	1 (1%)	0	100	100
3	C	107/130 (82%)	106 (99%)	1 (1%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
3	G	118/130 (91%)	117 (99%)	1 (1%)	0	100 100
4	D	93/126 (74%)	92 (99%)	1 (1%)	0	100 100
4	H	93/126 (74%)	92 (99%)	1 (1%)	0	100 100
7	O	100/478 (21%)	91 (91%)	8 (8%)	1 (1%)	13 40
All	All	863/1468 (59%)	847 (98%)	15 (2%)	1 (0%)	50 77

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
7	O	237	LYS

5.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all EM entries.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	85/111 (77%)	85 (100%)	0	100 100
1	E	85/111 (77%)	84 (99%)	1 (1%)	67 89
2	B	68/79 (86%)	68 (100%)	0	100 100
2	F	68/79 (86%)	68 (100%)	0	100 100
3	C	85/100 (85%)	85 (100%)	0	100 100
3	G	94/100 (94%)	94 (100%)	0	100 100
4	D	81/105 (77%)	81 (100%)	0	100 100
4	H	81/105 (77%)	80 (99%)	1 (1%)	67 89
7	O	94/369 (26%)	94 (100%)	0	100 100
All	All	741/1159 (64%)	739 (100%)	2 (0%)	90 97

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

Mol	Chain	Res	Type
1	E	63	ARG
4	H	88	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	76	GLN
1	A	85	GLN
4	D	63	ASN
1	E	108	ASN
4	H	82	HIS

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates (i)

There are no 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)

There are no chain breaks in this entry.

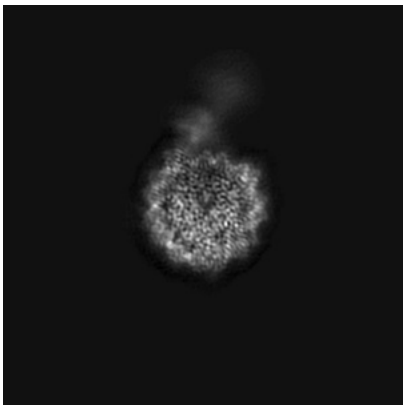
6 Map visualisation

This section contains visualisations of the EMDB entry EMD-43194. These allow visual inspection of the internal detail of the map and identification of artifacts.

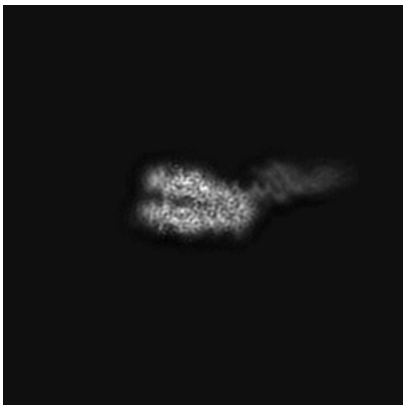
Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

6.1 Orthogonal projections

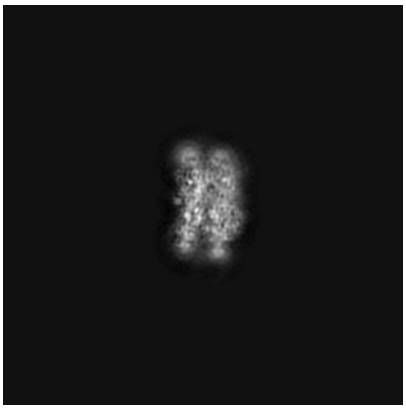
6.1.1 Primary map



X

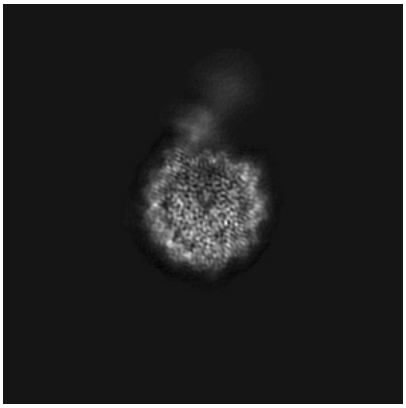


Y

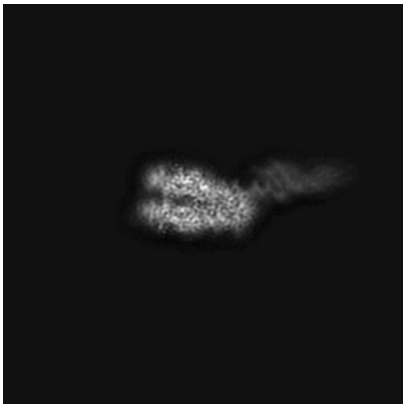


Z

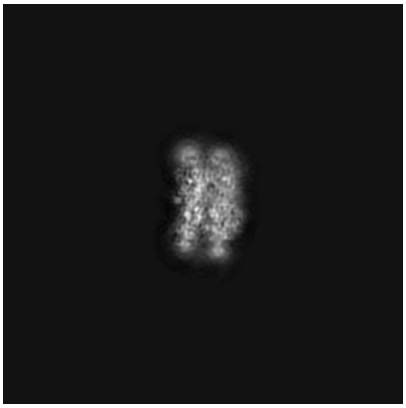
6.1.2 Raw map



X



Y

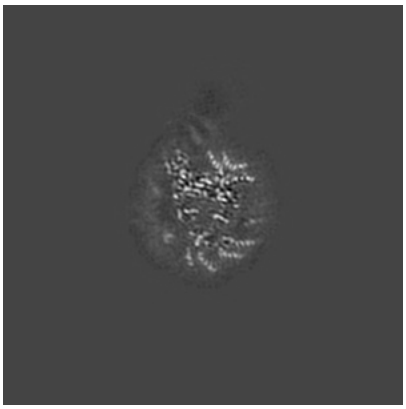


Z

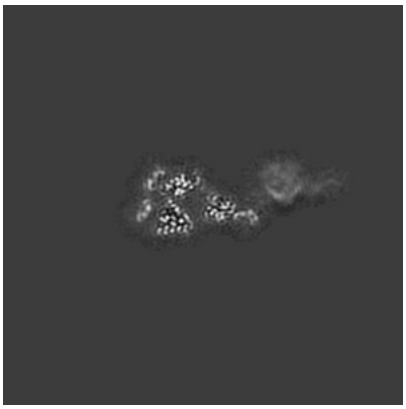
The images above show the map projected in three orthogonal directions.

6.2 Central slices (i)

6.2.1 Primary map



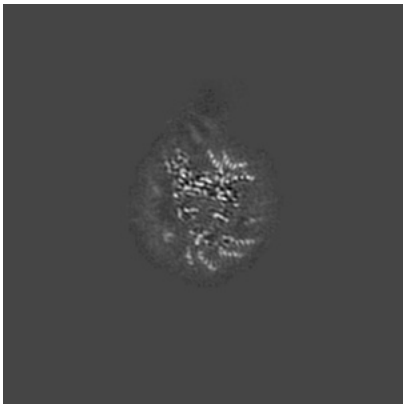
X Index: 128



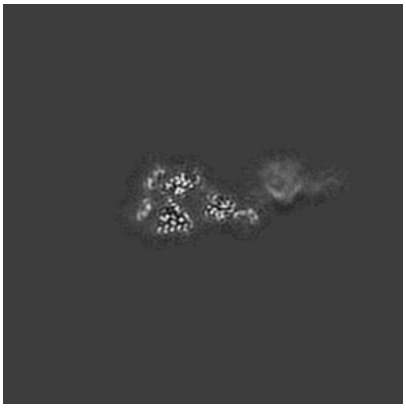
Y Index: 128

Z Index: 128

6.2.2 Raw map



X Index: 128



Y Index: 128

Z Index: 128

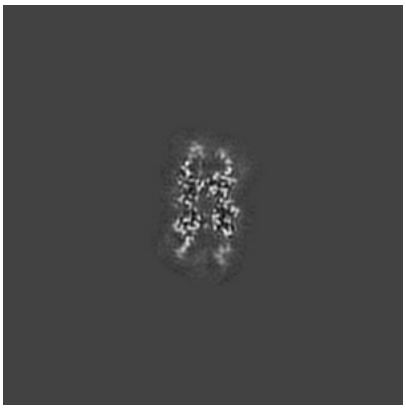
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

6.3.1 Primary map

X Index: 123

Y Index: 120

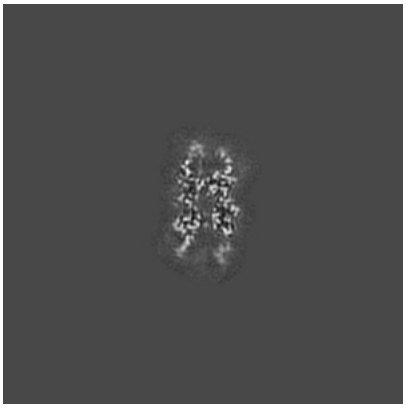


Z Index: 118

6.3.2 Raw map

X Index: 123

Y Index: 120

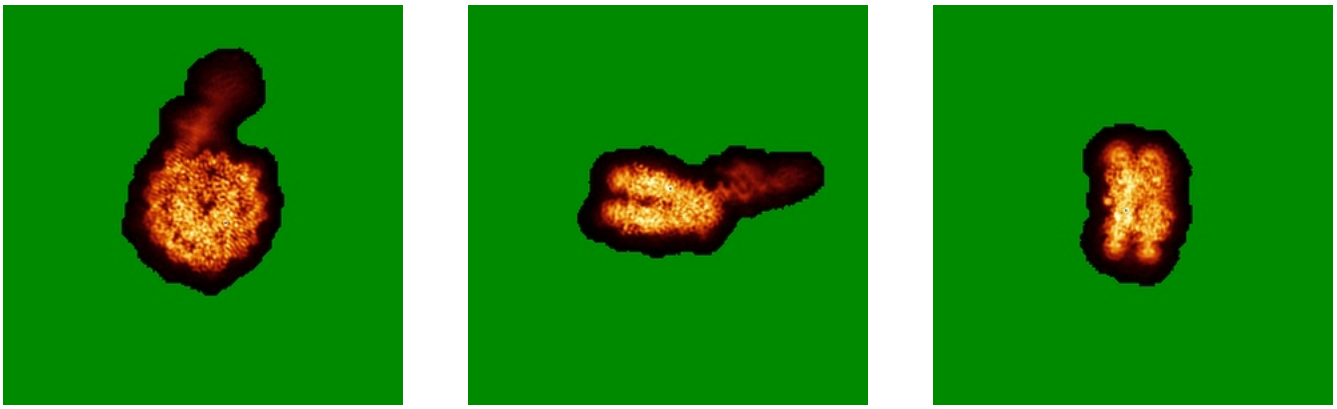


Z Index: 118

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

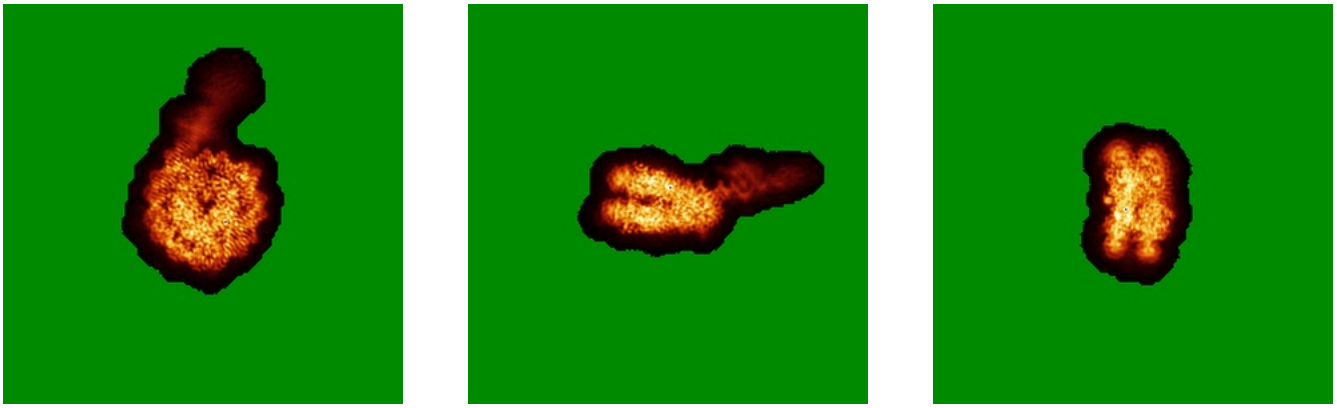
6.4 Orthogonal standard-deviation projections (False-color) (i)

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.

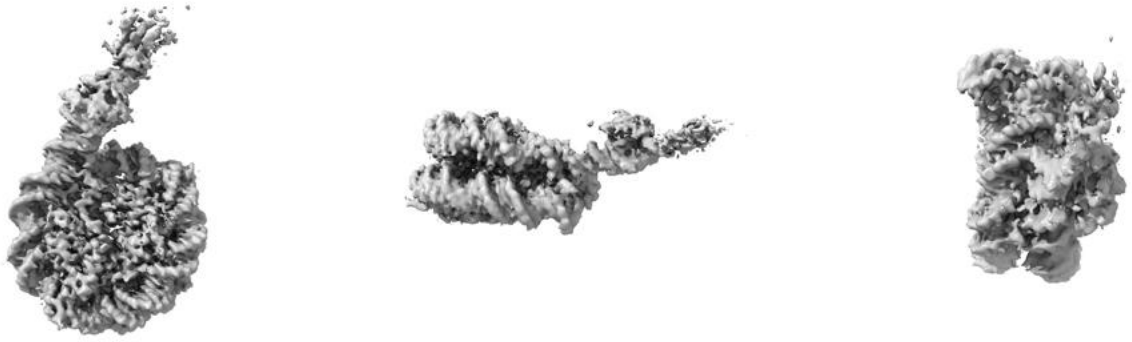
6.5 Orthogonal surface views (i)

6.5.1 Primary map

X Y Z

The images above show the 3D surface view of the map at the recommended contour level 0.007. 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.

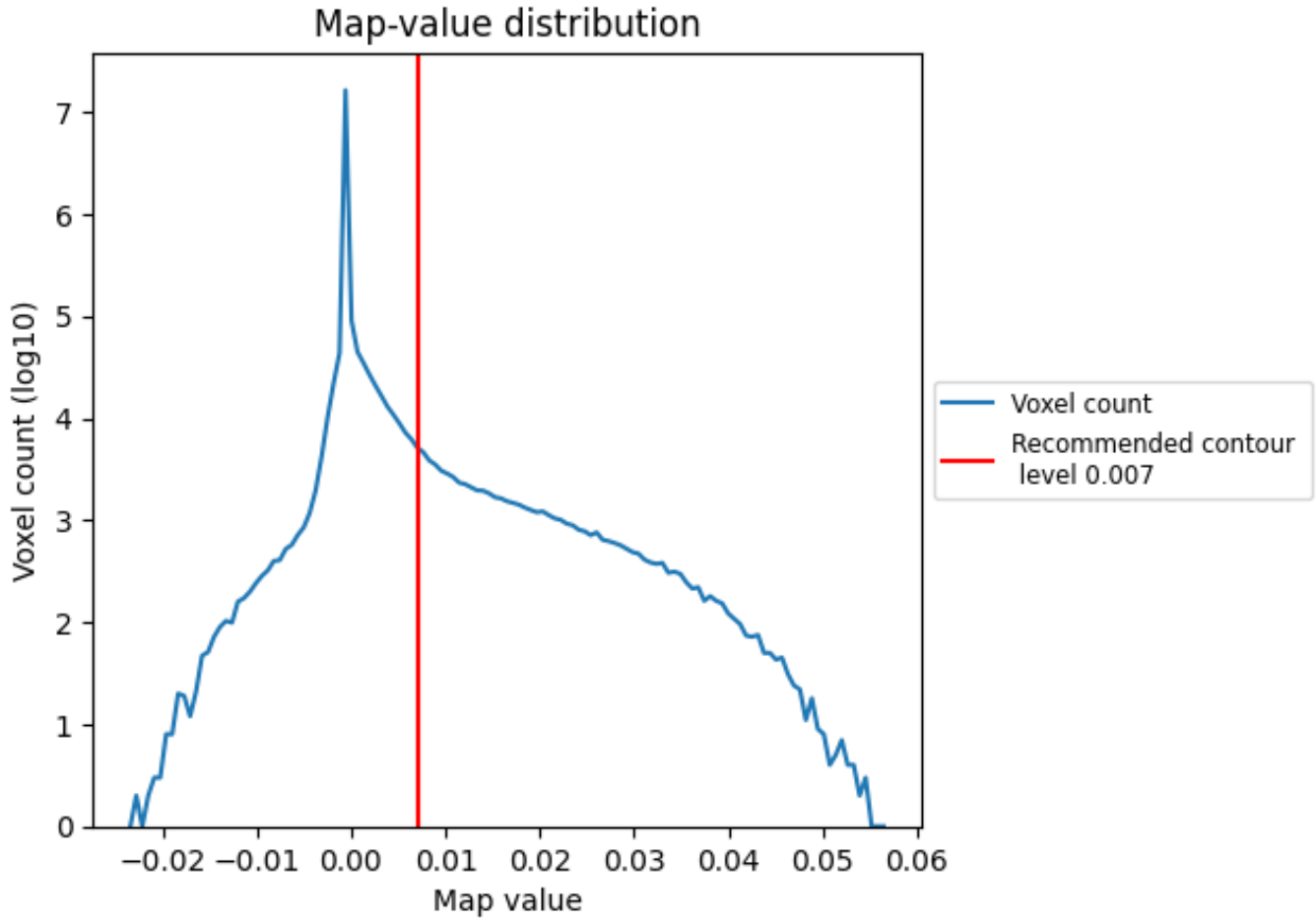
6.6 Mask visualisation (i)

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

7 Map analysis

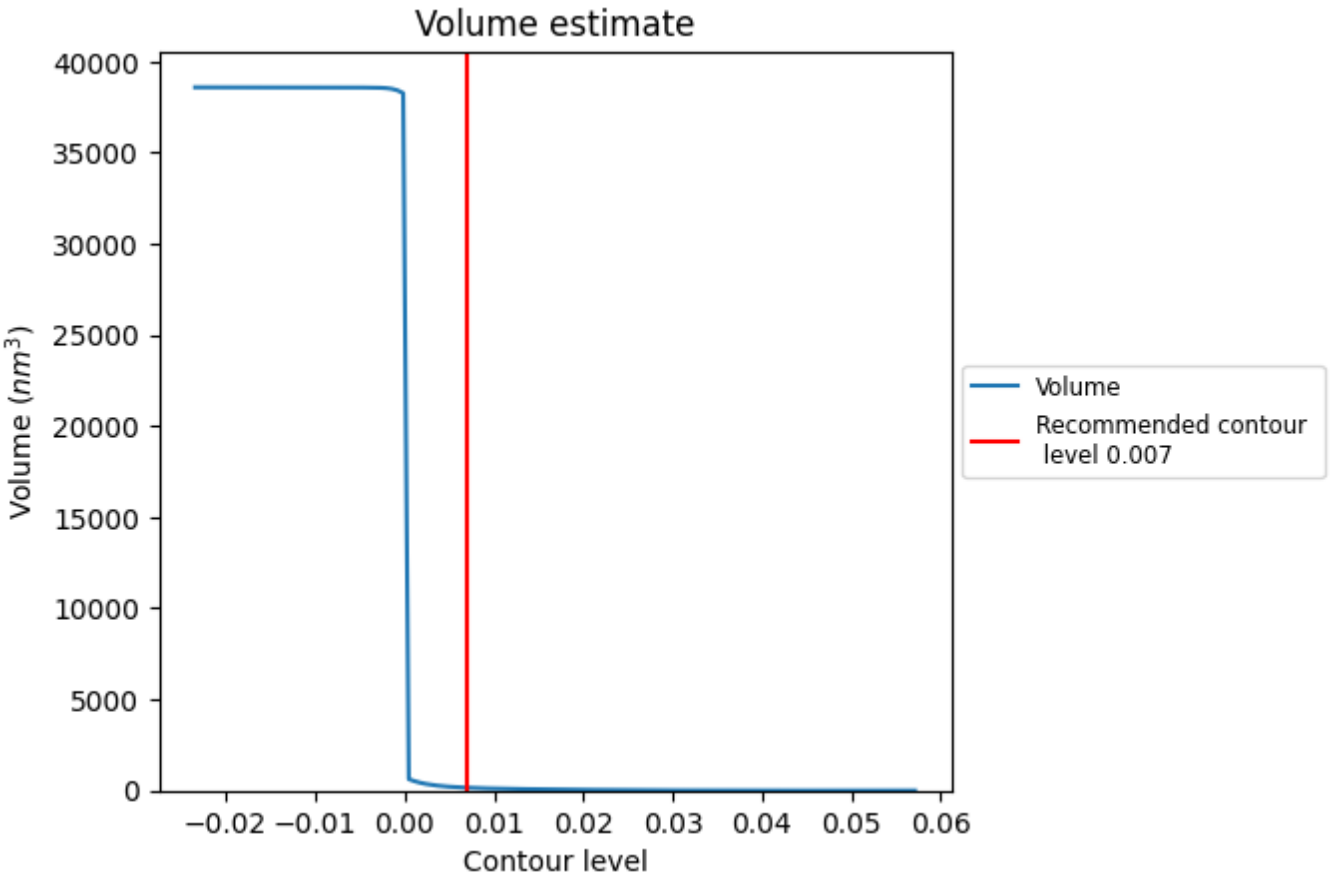
This section contains the results of statistical analysis of the map.

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.

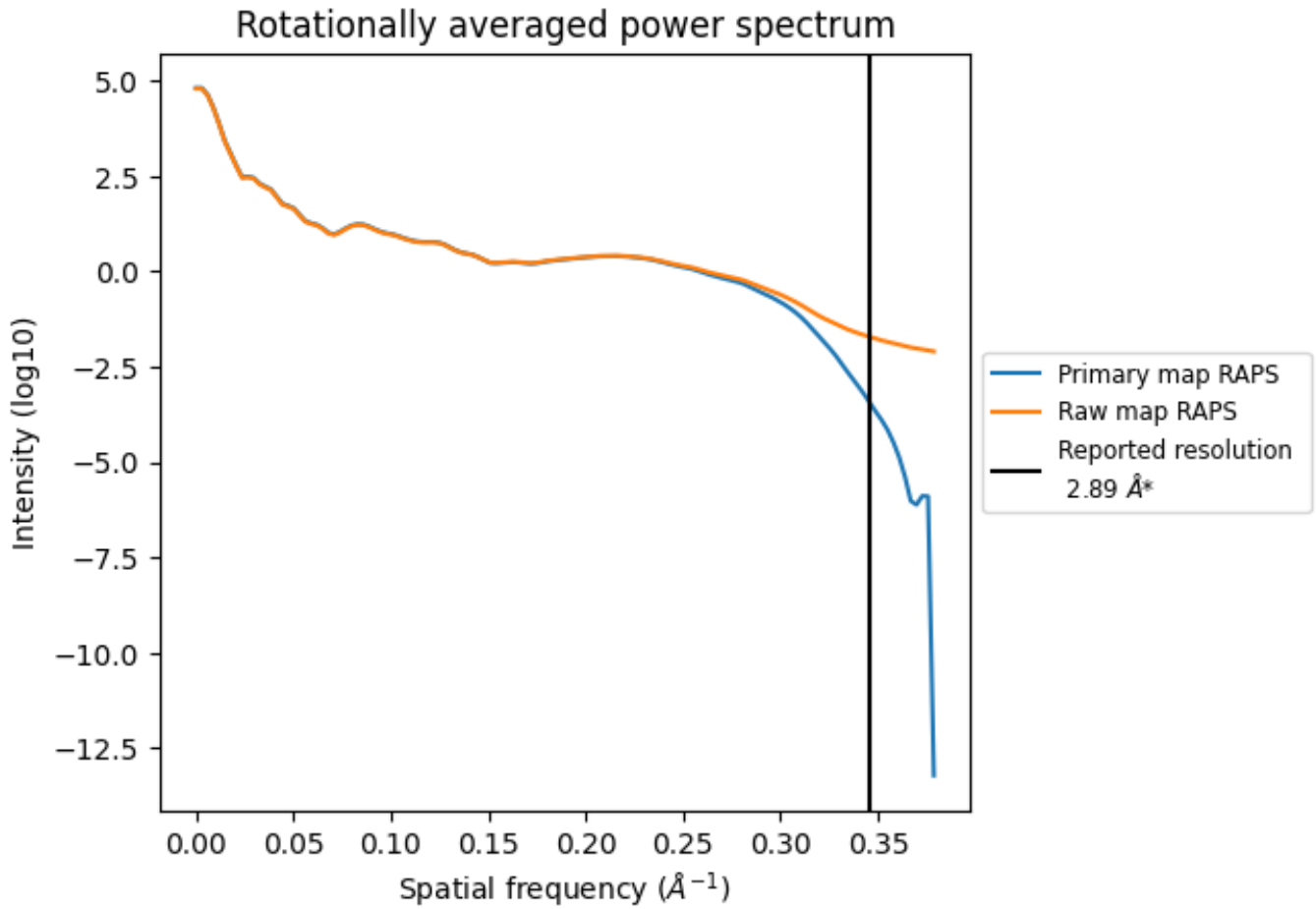
7.2 Volume estimate (i)



The volume at the recommended contour level is 156 nm^3; this corresponds to an approximate mass of 141 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)

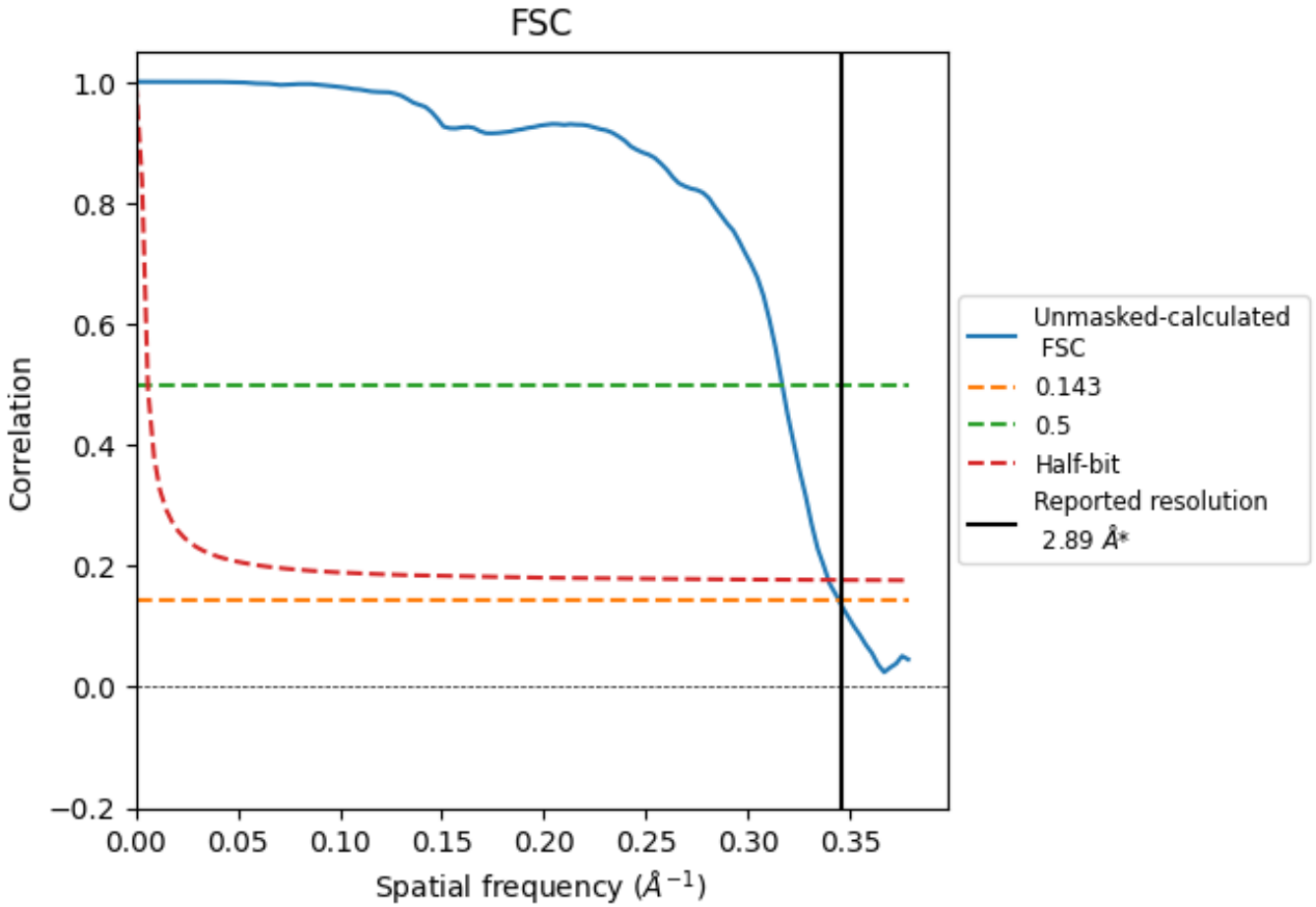


*Reported resolution corresponds to spatial frequency of 0.346 $Å^{-1}$

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 0.346 Å^{-1}

8.2 Resolution estimates (i)

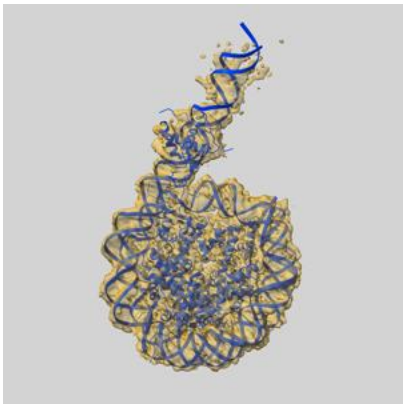
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	2.89	-	-
Author-provided FSC curve	-	-	-
Unmasked-calculated*	2.90	3.15	2.94

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

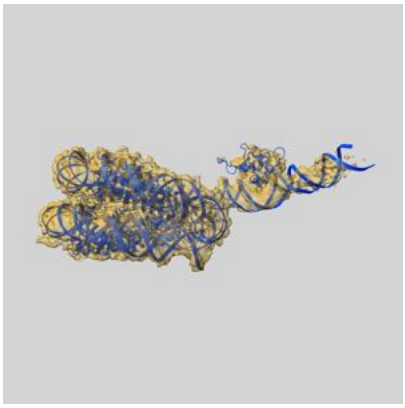
9 Map-model fit

This section contains information regarding the fit between EMDB map EMD-43194 and PDB model 8VFY. Per-residue inclusion information can be found in section 3 on page 6.

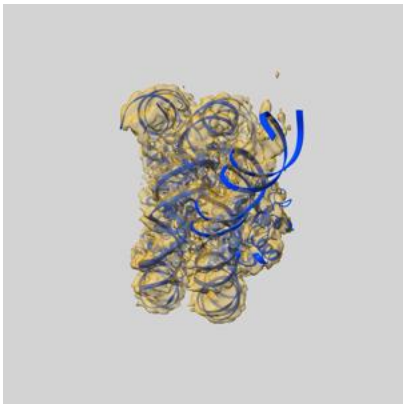
9.1 Map-model overlay



X



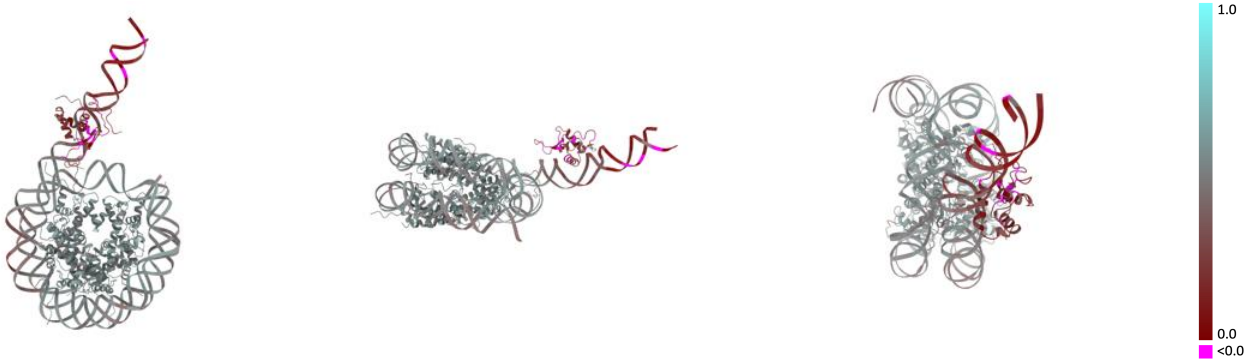
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.007 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.

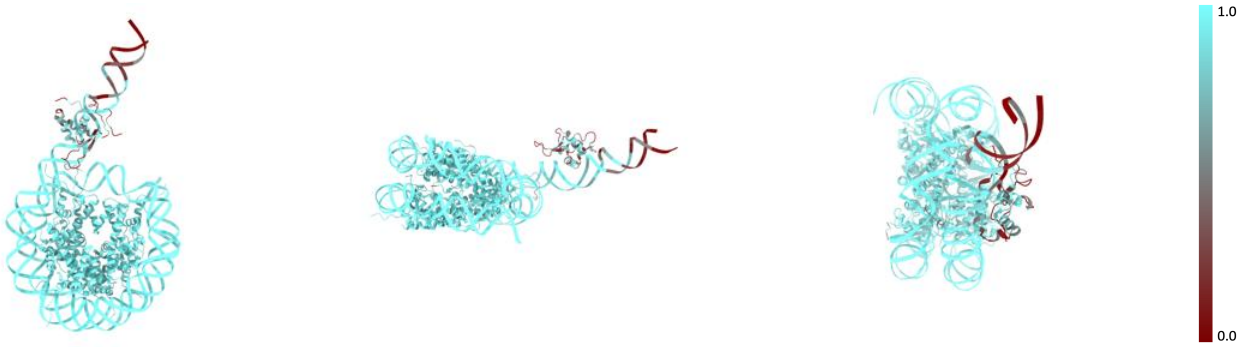
9.2 Q-score mapped to coordinate model (i)



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.

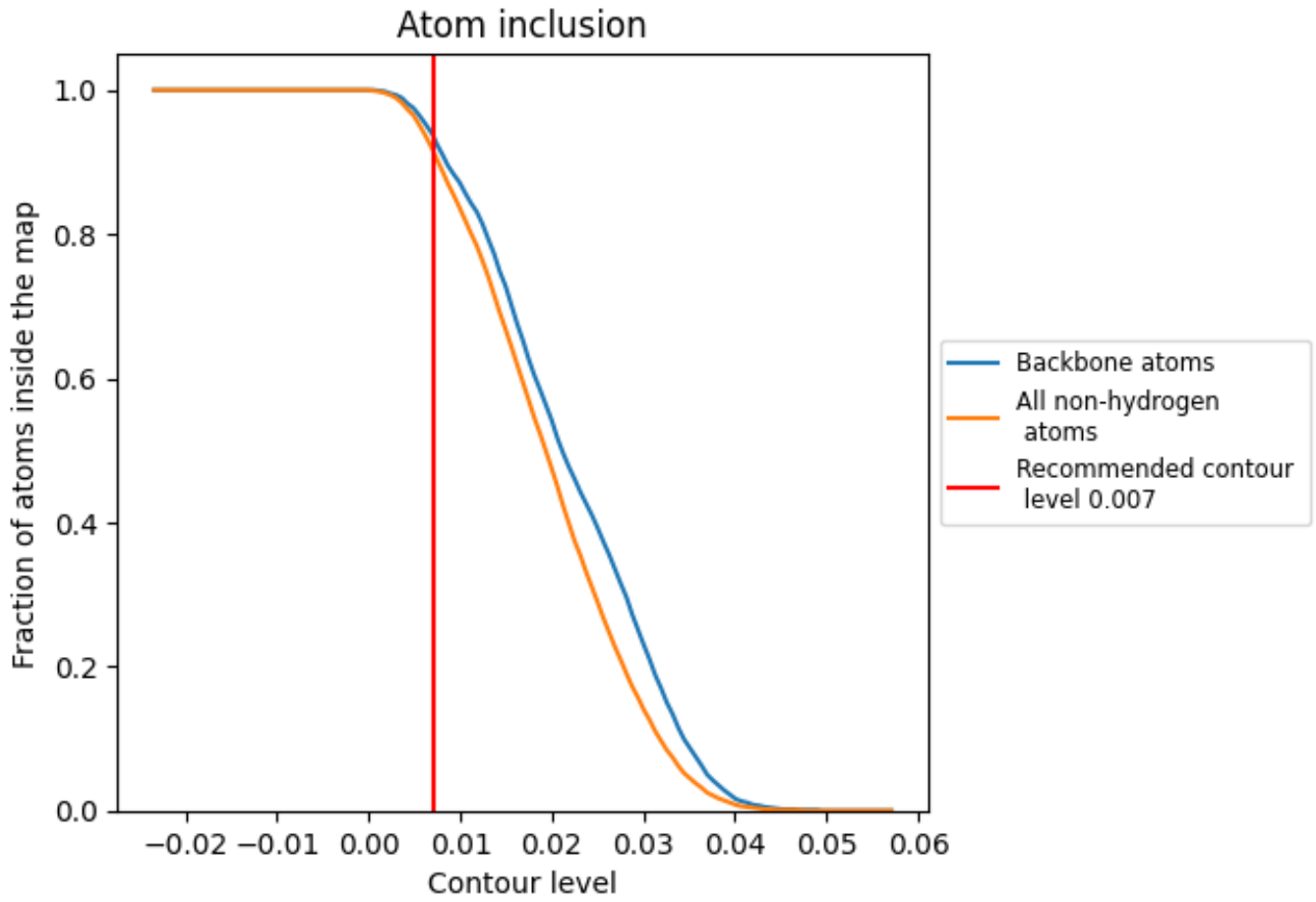
9.3 Atom inclusion mapped to coordinate model (i)



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 (0.007).

9.4 Atom inclusion (i)



At the recommended contour level, 94% of all backbone atoms, 92% 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 (0.007) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.9170	0.4630
A	0.9550	0.5270
B	0.9470	0.5240
C	0.9630	0.5340
D	0.9630	0.5170
E	0.9810	0.5420
F	0.9690	0.5390
G	0.9310	0.5150
H	0.9680	0.5270
I	0.9350	0.4450
J	0.9290	0.4360
O	0.4940	0.1710