wwPDB EM Validation Summary Report (i)

Mar 9, 2026 – 07:17 PM UTC

PDB ID	:	9EAK / pdb_00009eak
EMDB ID	:	EMD-47834
Title	:	SpCas9 with 17-bp R-loop containing 2 terminal mismatches (State I - pre-activation)
Authors	:	Kiernan, K.A.; Taylor, D.W.
Deposited on	:	2024-11-11
Resolution	:	2.76 Å(reported)
Based on initial model	:	6O0Z

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:

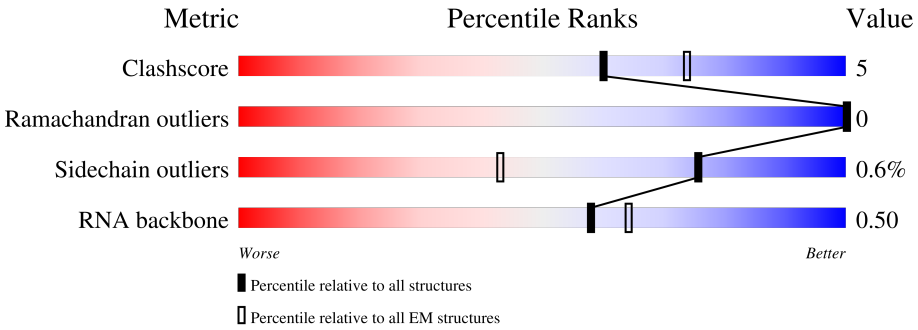
EMDB validation analysis	:	0.0.1.dev132
MolProbity	:	4-5-2 with Phenix2.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
EM percentile statistics	:	202505.v01 (Using data in the EMDB archive up until May 2025)
MapQ	:	1.9.13
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

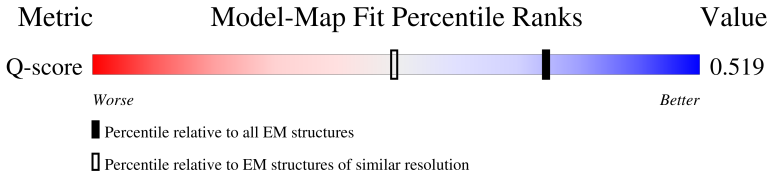
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:
ELECTRON MICROSCOPY

The reported resolution of this entry is 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.





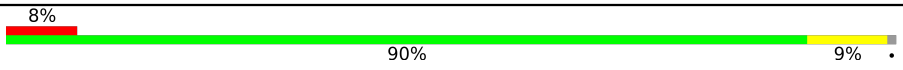
Metric	Whole archive (#Entries)	EM structures (#Entries)	Similar EM resolution (#Entries, resolution range(Å))
Clashscore	229148	23984	-
Ramachandran outliers	224038	23583	-
Sidechain outliers	223484	23102	-
RNA backbone	8273	3508	-
Q-score	-	25397	10642 (2.26 - 3.26)

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	B	97	62% 31% 5%
2	C	55	44% 15% 42%
3	D	55	13% 38% 60%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	A	1368	

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 12879 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 sgRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
1	B	95	2038	912	376	655	95	0	0

- Molecule 2 is a DNA chain called DNA (Target Strand).

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
2	C	32	648	311	106	199	32	0	0

- Molecule 3 is a DNA chain called DNA (Non-target strand).

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
3	D	22	457	216	90	129	22	0	0

- Molecule 4 is a protein called CRISPR-associated endonuclease Cas9/Csn1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	A	1356	9735	6209	1736	1772	18	0	0

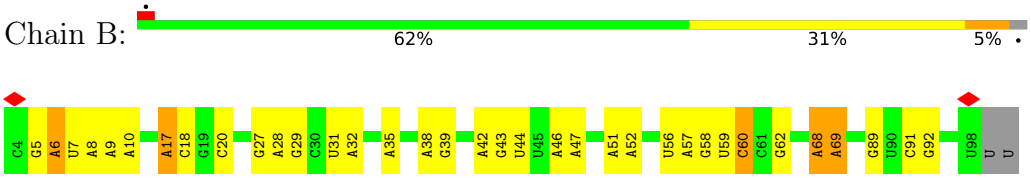
- Molecule 5 is MAGNESIUM ION (CCD ID: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		AltConf
			Total	Mg	
5	A	1	1	1	0

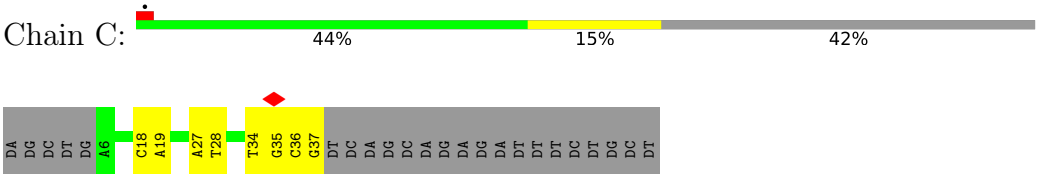
3 Residue-property plots (i)

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.

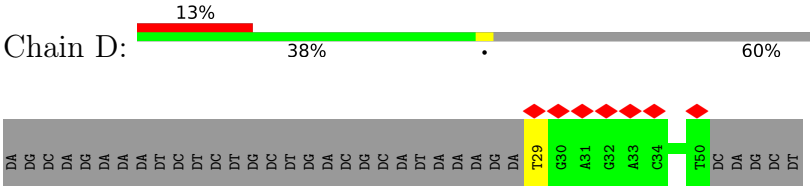
• Molecule 1: sgRNA



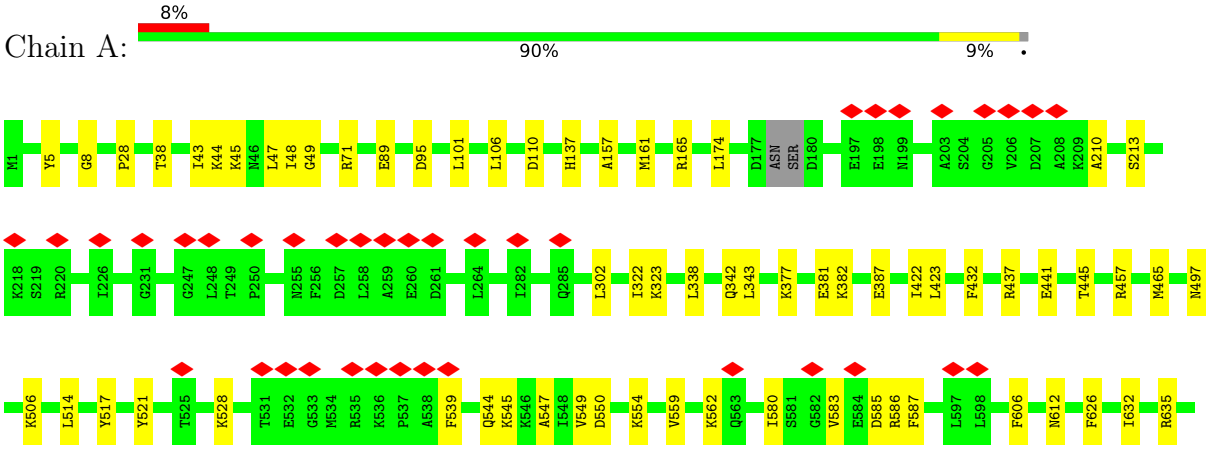
• Molecule 2: DNA (Target Strand)

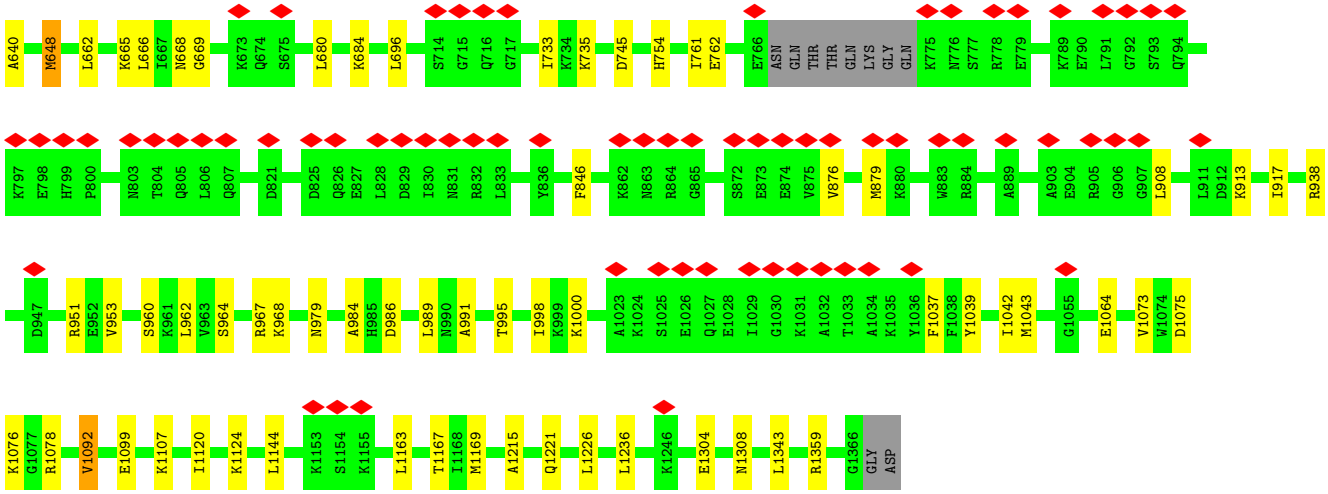


• Molecule 3: DNA (Non-target strand)



• Molecule 4: CRISPR-associated endonuclease Cas9/Csn1




A640 M648 L662 K665 L666 I667 N668 G669 K673 Q674 S675 L680 K684 L696 S714 G715 Q716 G717 I733 K734 K735 D745 H754 I761 E762 E766 ASN GLN THR THR GLN LYS GLY GLN K775 N776 S777 R778 E779 K789 E790 L791 G792 S793 Q794
K797 E798 H799 P800 N803 T804 Q805 L806 Q807 D821 D825 Q826 E827 L828 D829 I830 N831 R832 L833 Y836 F846 K862 N863 R864 G865 S872 E873 E874 V875 V876 M879 K880 W883 R884 A889 A903 E904 R905 G906 G907 L908 L911 D912 K913 I917 R938
D947 R951 E952 V953 S960 K961 L962 V963 S964 R967 K968 N979 A984 H985 D986 L989 N990 A991 T995 I998 K999 K1000 A1023 K1024 S1025 E1026 Q1027 E1028 I1029 G1030 K1031 A1032 T1033 A1034 K1035 Y1036 F1037 F1038 Y1039 I1042 M1043 G1055 E1064 V1073 W1074 D1075
K1076 G1077 R1078 V1092 E1099 K1107 I1120 K1124 L1144 K1153 S1154 K1155 L1163 T1167 I1168 M1169 A1215 Q1221 L1226 L1236 K1246 E1304 N1308 L1343 R1359 G1366 GLY ASP

4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	112156	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	TFS KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/Å^2$)	80	Depositor
Minimum defocus (nm)	1200	Depositor
Maximum defocus (nm)	2200	Depositor
Magnification	Not provided	
Image detector	GATAN K3 BIOCONTINUUM (6k x 4k)	Depositor
Maximum map value	0.943	Depositor
Minimum map value	-0.346	Depositor
Average map value	-0.000	Depositor
Map value standard deviation	0.012	Depositor
Recommended contour level	0.11	Depositor
Map size (Å)	426.5984, 426.5984, 426.5984	wwPDB
Map dimensions	512, 512, 512	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	0.8332, 0.8332, 0.8332	Depositor

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: MG

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	B	0.24	0/2284	0.40	0/3559
2	C	0.27	0/722	0.50	0/1111
3	D	0.23	0/514	0.43	0/792
4	A	0.26	0/9910	0.50	0/13460
All	All	0.26	0/13430	0.48	0/18922

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
4	A	0	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
4	A	951	ARG	Sidechain

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	B	2038	0	1023	16	0
2	C	648	0	365	5	0
3	D	457	0	247	1	0
4	A	9735	0	8739	88	0
5	A	1	0	0	0	0
All	All	12879	0	10374	104	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 104 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)
4:A:662:LEU:HD12	4:A:666:LEU:CD2	1.77	1.13
4:A:662:LEU:HD12	4:A:666:LEU:HD23	1.05	1.03
4:A:106:LEU:HD13	4:A:110:ASP:HB2	1.44	0.94
4:A:640:ALA:HA	4:A:648:MET:HE3	1.50	0.93
4:A:662:LEU:CD1	4:A:666:LEU:HD23	1.97	0.92

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	
4	A	1350/1368 (99%)	1288 (95%)	62 (5%)	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
4	A	826/1227 (67%)	821 (99%)	5 (1%)	78 88

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	A	648	MET
4	A	696	LEU
4	A	953	VAL
4	A	1037	PHE
4	A	1092	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
4	A	501	ASN
4	A	758	ASN
4	A	980	ASN
4	A	920	GLN
4	A	420	HIS

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	B	94/97 (96%)	20 (21%)	4 (4%)

5 of 20 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	B	5	G
1	B	6	A
1	B	17	A
1	B	20	C
1	B	28	A

All (4) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
1	B	5	G
1	B	38	A
1	B	42	A
1	B	68	A

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)

Of 1 ligands modelled in this entry, 1 is monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

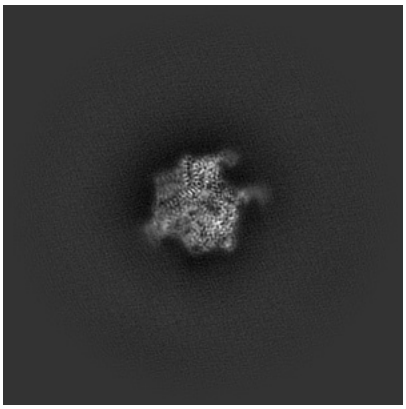
6 Map visualisation

This section contains visualisations of the EMDB entry EMD-47834. These allow visual inspection of the internal detail of the map and identification of artifacts.

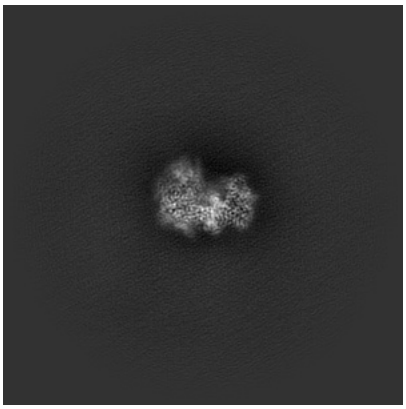
Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

6.1 Orthogonal projections

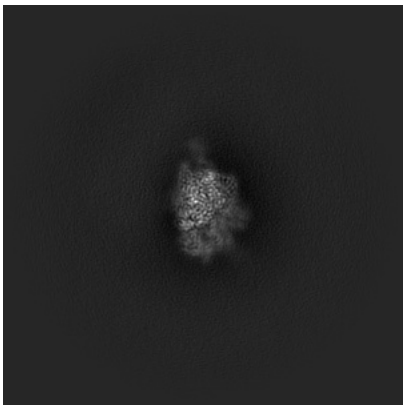
6.1.1 Primary map



X

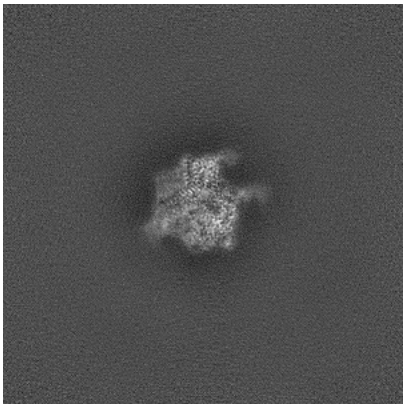


Y

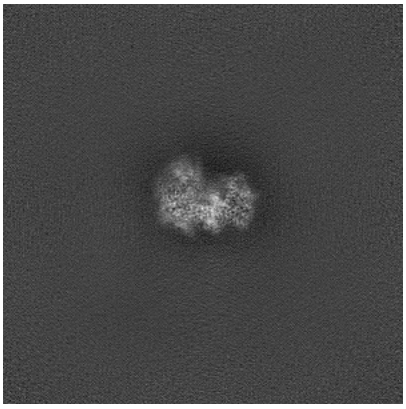


Z

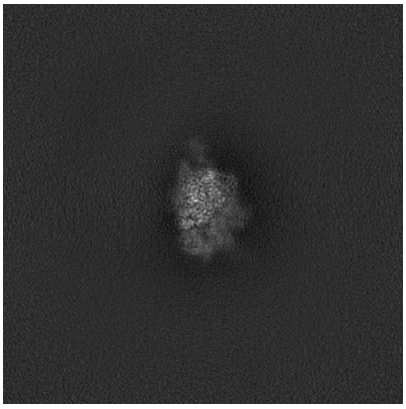
6.1.2 Raw map



X



Y

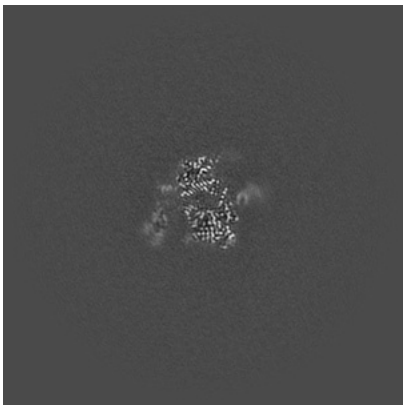


Z

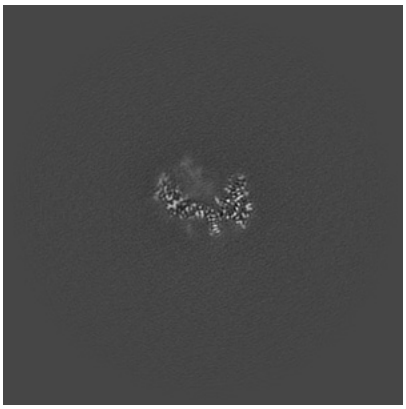
The images above show the map projected in three orthogonal directions.

6.2 Central slices (i)

6.2.1 Primary map



X Index: 256

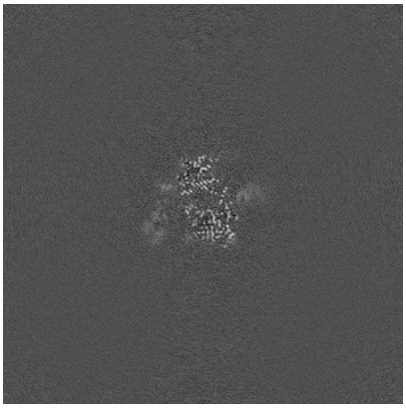


Y Index: 256

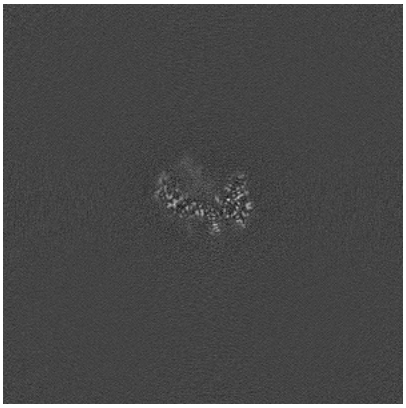


Z Index: 256

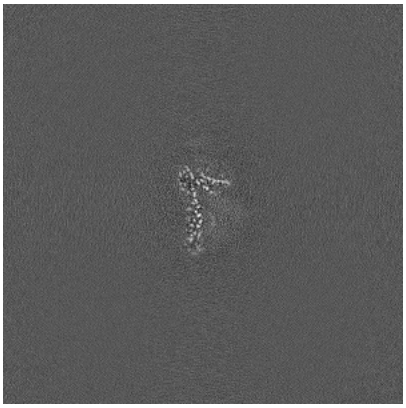
6.2.2 Raw map



X Index: 256



Y Index: 256



Z Index: 256

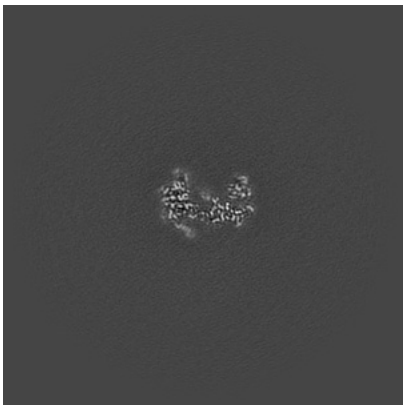
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

6.3.1 Primary map



X Index: 242

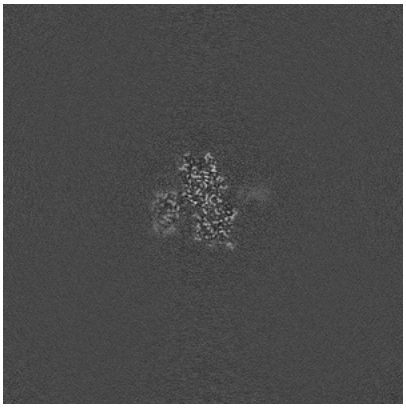


Y Index: 269



Z Index: 268

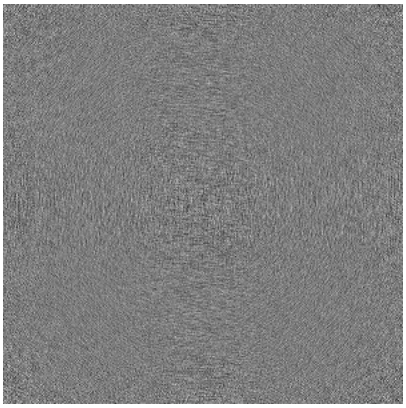
6.3.2 Raw map



X Index: 246



Y Index: 259

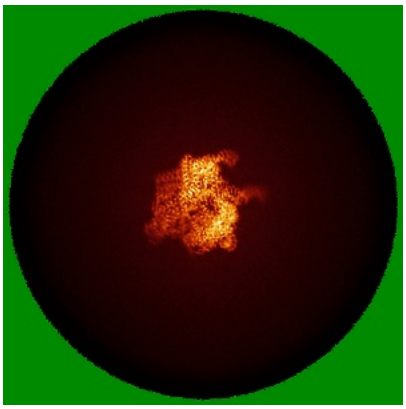


Z Index: 0

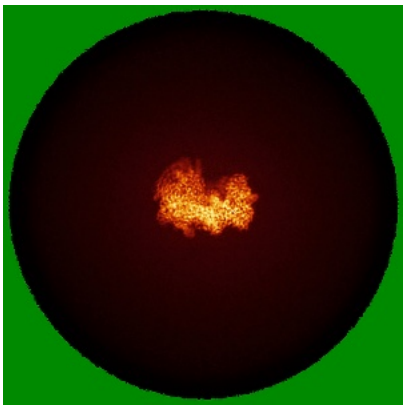
The images above show the largest variance slices of the map in three orthogonal directions.

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

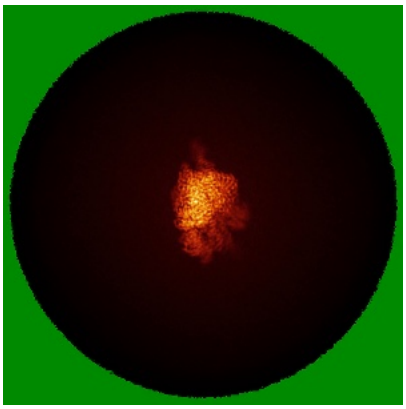
6.4.1 Primary map



X

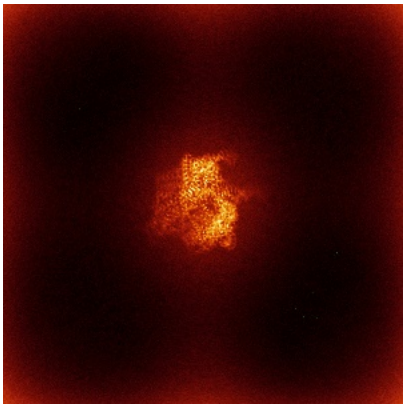


Y

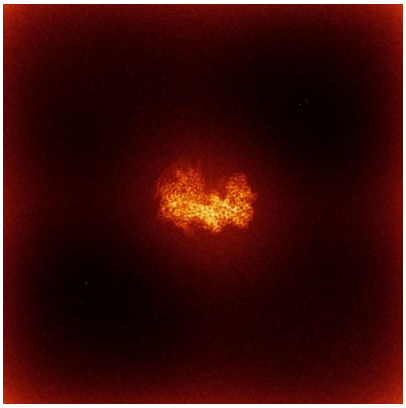


Z

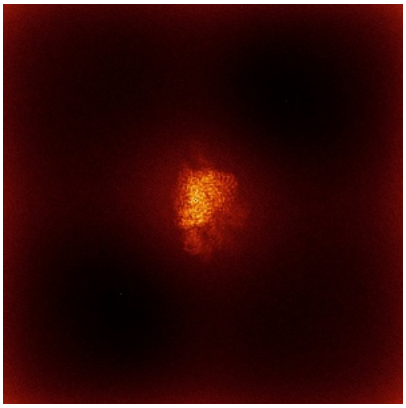
6.4.2 Raw map



X



Y

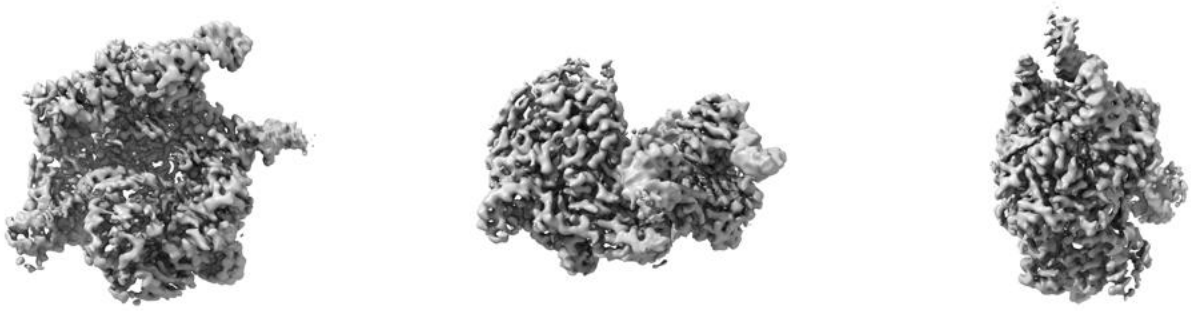


Z

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 0.11. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

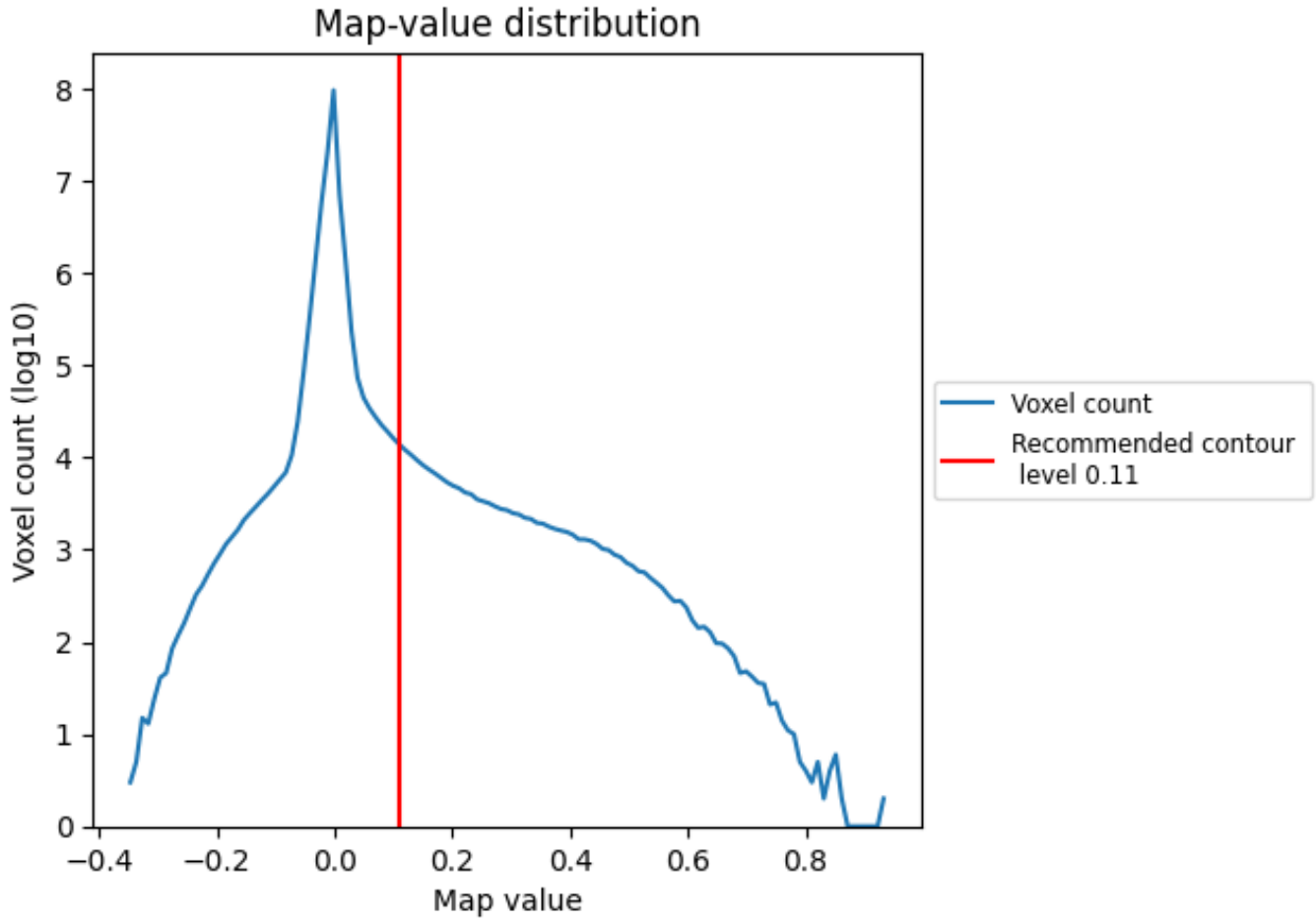
6.6 Mask visualisation (i)

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

7 Map analysis

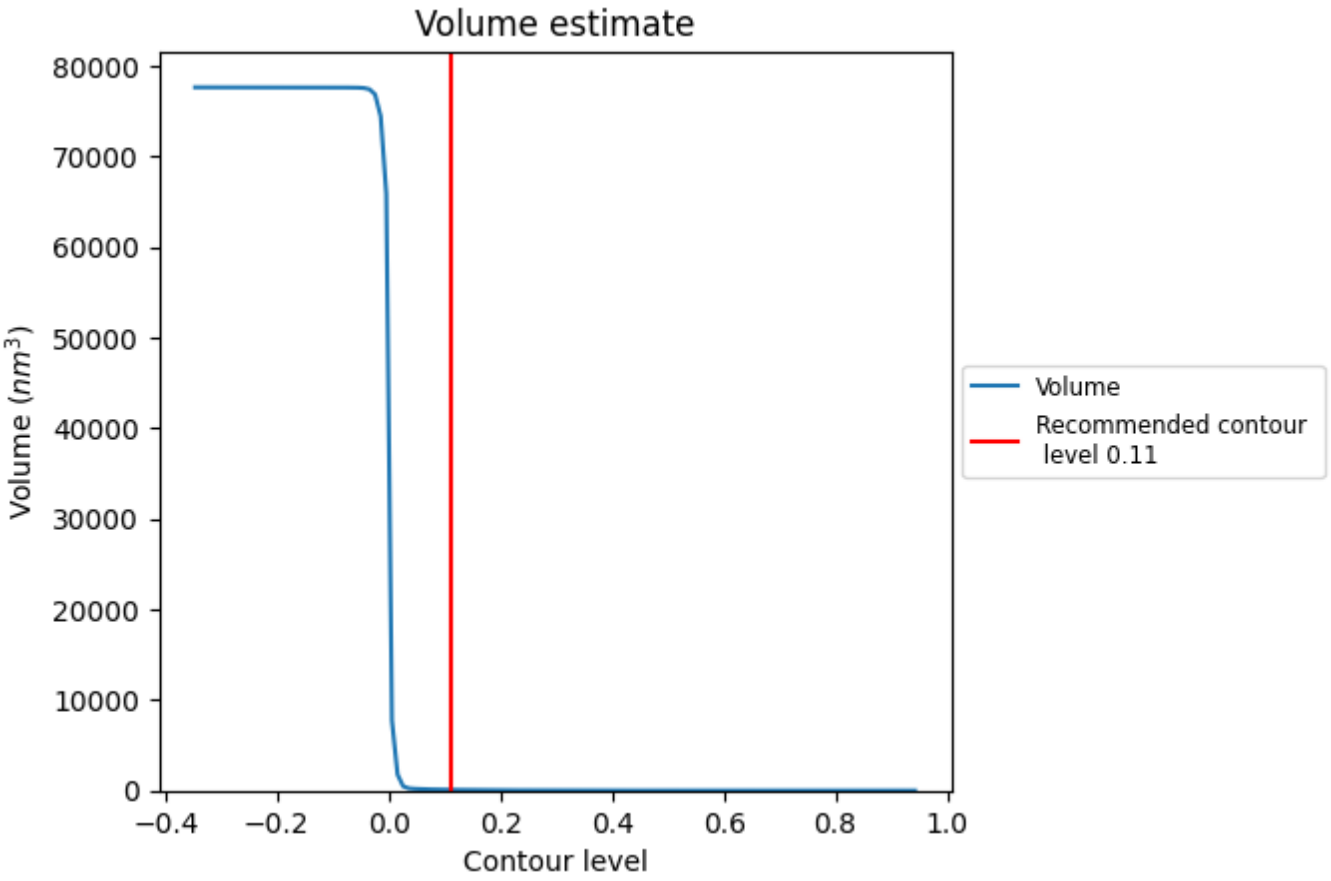
This section contains the results of statistical analysis of the map.

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.

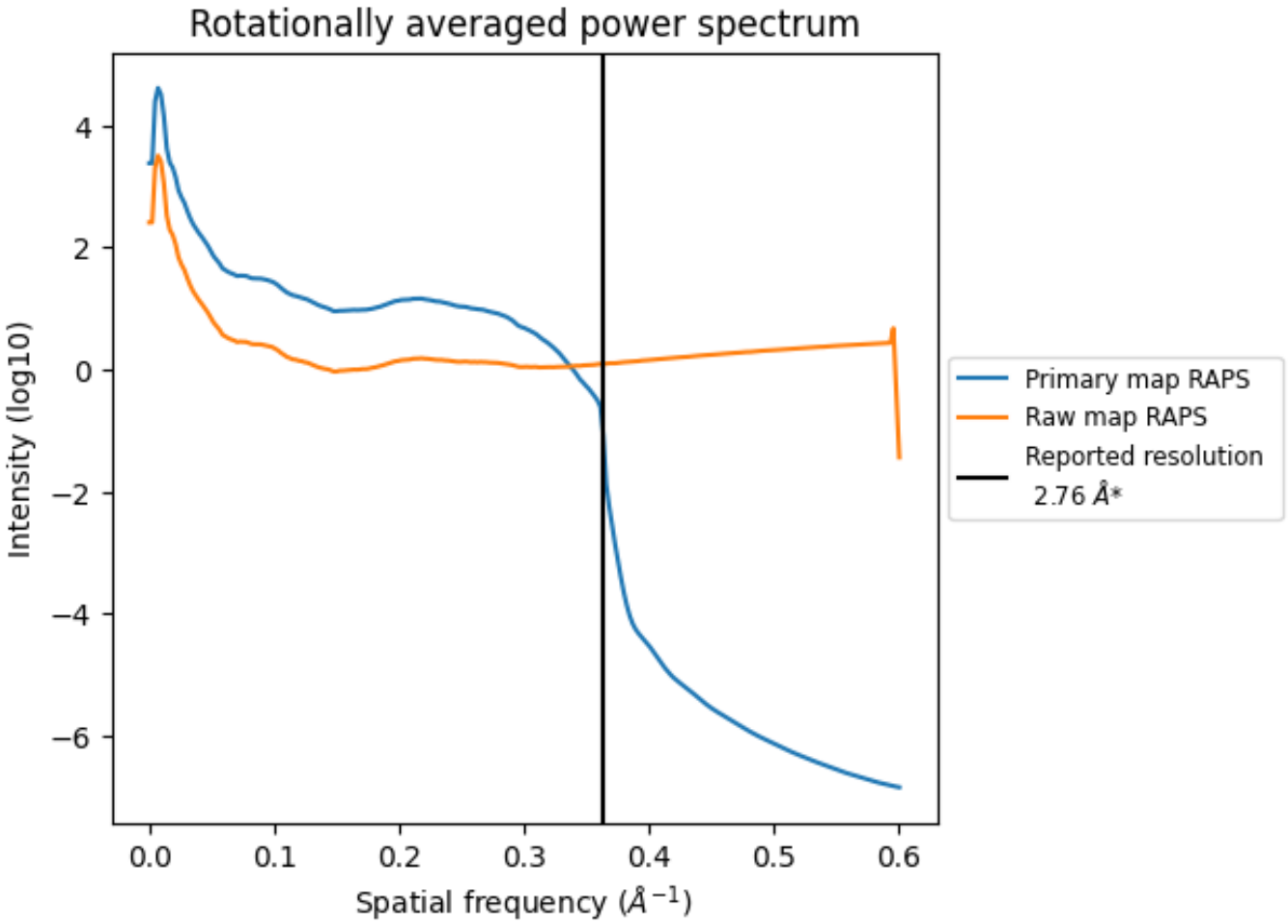
7.2 Volume estimate (i)



The volume at the recommended contour level is 88 nm^3; this corresponds to an approximate mass of 80 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)

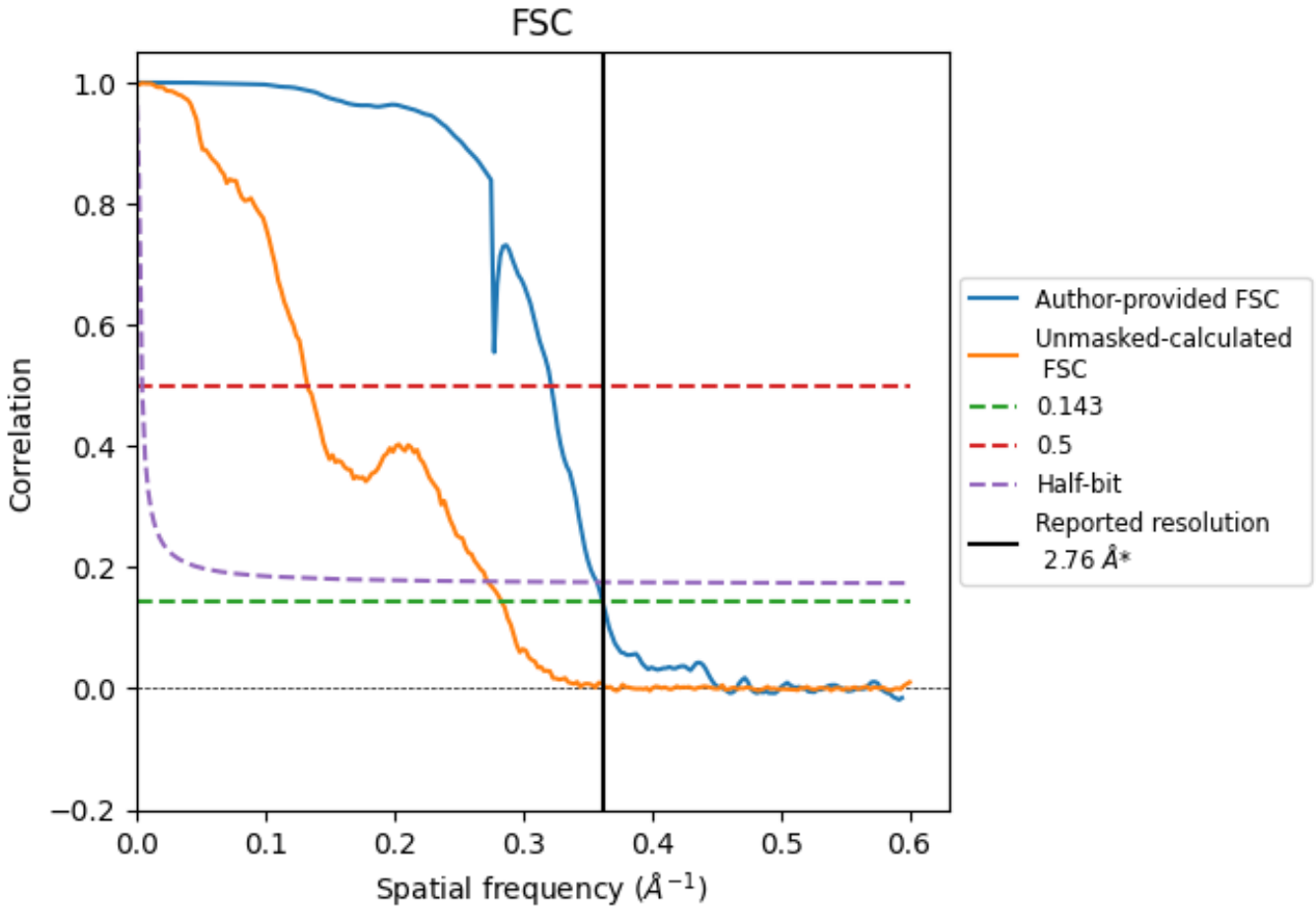


*Reported resolution corresponds to spatial frequency of 0.362 Å^{-1}

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)

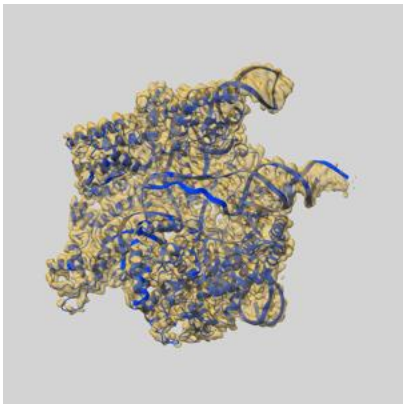
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	2.76	-	-
Author-provided FSC curve	2.76	3.11	2.80
Unmasked-calculated*	3.53	7.52	3.67

*Resolution estimate based on FSC curve calculated by comparison of deposited half-maps. The value from deposited half-maps intersecting FSC 0.143 CUT-OFF 3.53 differs from the reported value 2.76 by more than 10 %

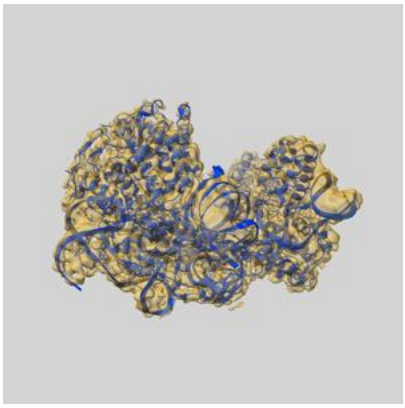
9 Map-model fit

This section contains information regarding the fit between EMDB map EMD-47834 and PDB model 9EAK. Per-residue inclusion information can be found in section 3 on page 5.

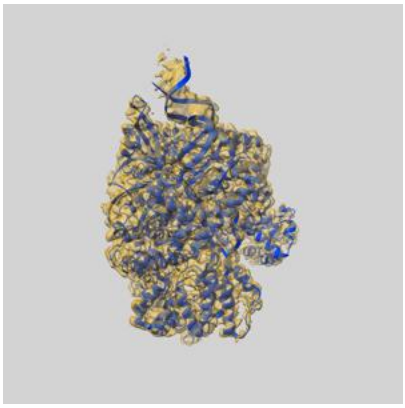
9.1 Map-model overlay



X



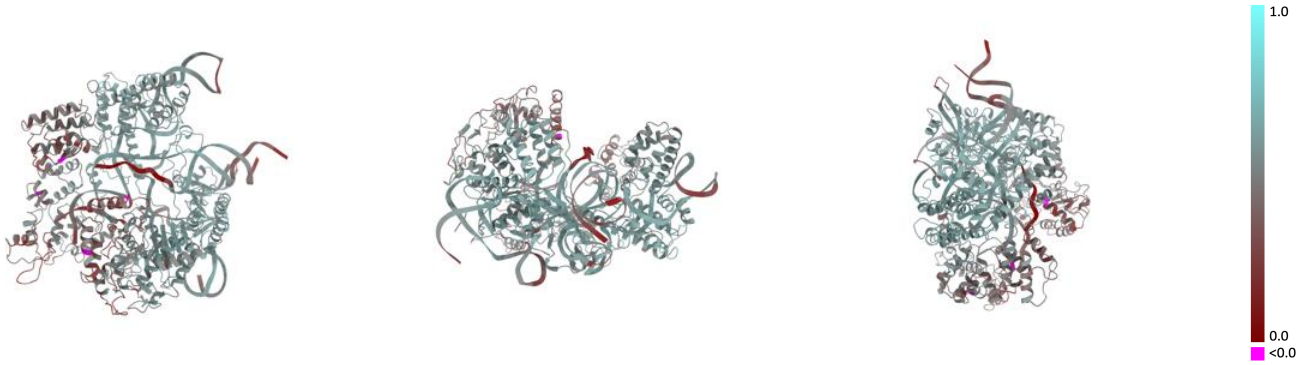
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.11 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.

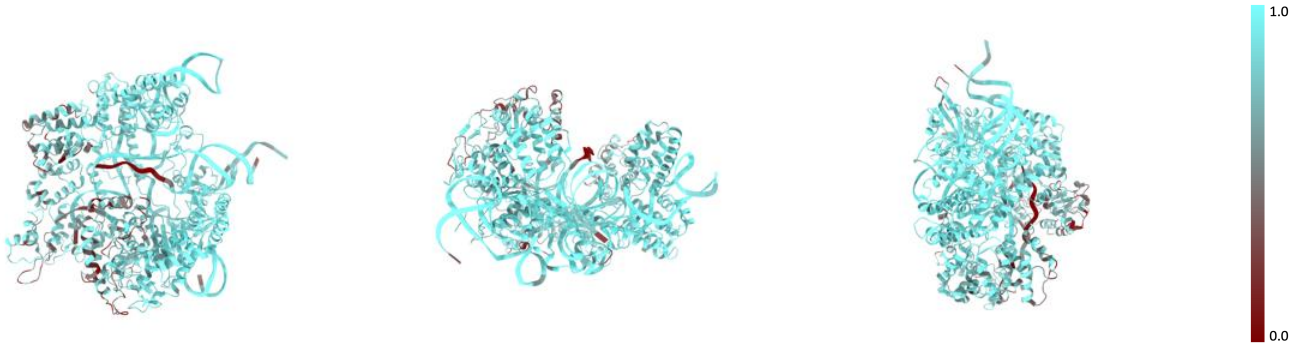
9.2 Q-score mapped to coordinate model



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.

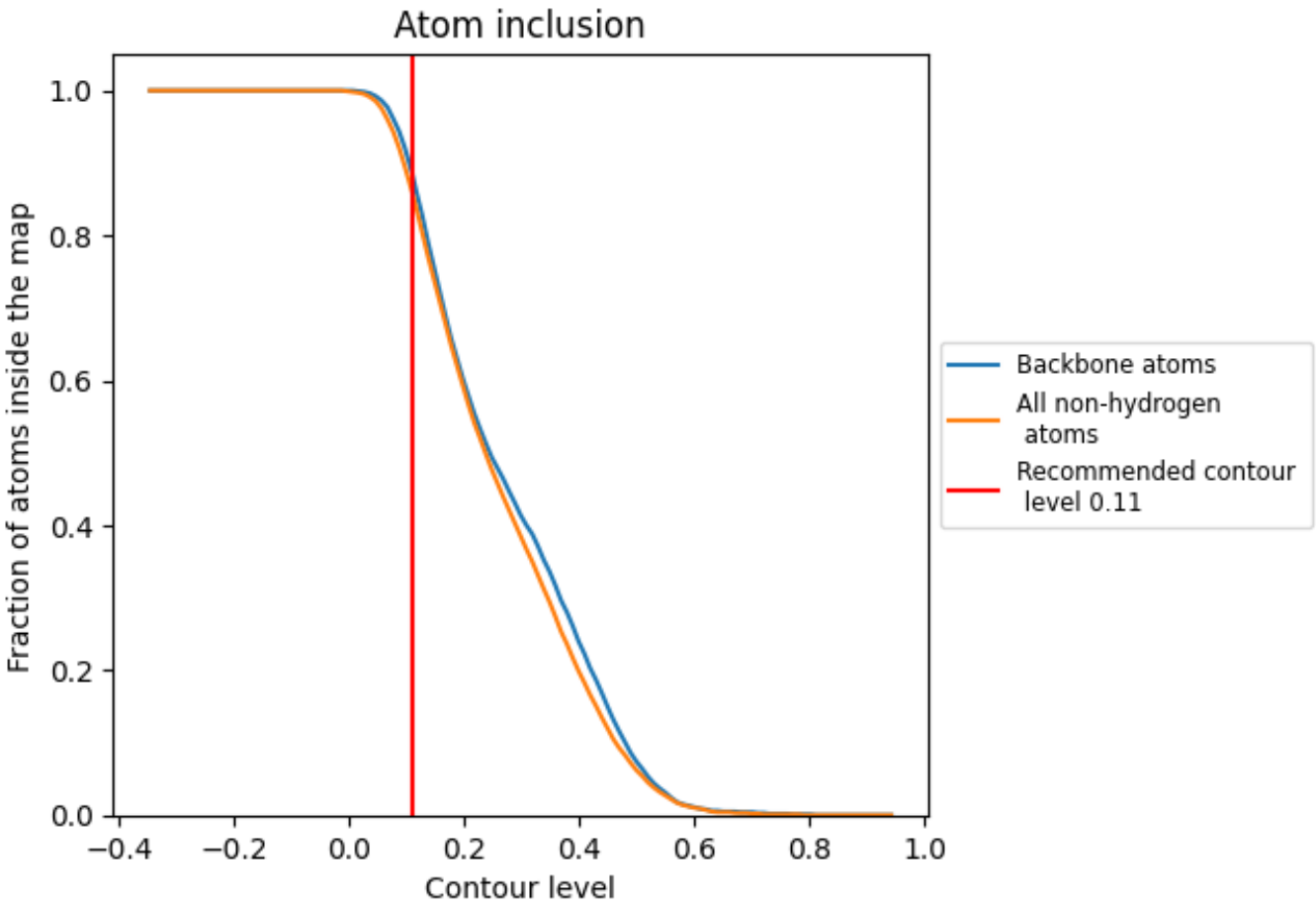
9.3 Atom inclusion mapped to coordinate model



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 (0.11).

9.4 Atom inclusion (i)



At the recommended contour level, 88% of all backbone atoms, 86% 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 (0.11) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.8580	0.5190
A	0.8520	0.5230
B	0.9400	0.5380
C	0.8500	0.4940
D	0.6320	0.3690

