



## wwPDB EM Validation Summary Report ⓘ

Sep 23, 2025 – 06:11 PM EDT

PDB ID : 9NHB / pdb\_00009nhb  
EMDB ID : EMD-49407  
Title : Cryo-EM structure of SIWI-piRNA-target(25-nt)-GTSF1 complex  
Authors : Sarkar, S.; Gebert, L.F.R.; MacRae, I.J.  
Deposited on : 2025-02-24  
Resolution : 3.70 Å(reported)  
Based on initial models : 5GUH, .

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

We welcome your comments at [validation@mail.wwpdb.org](mailto: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:

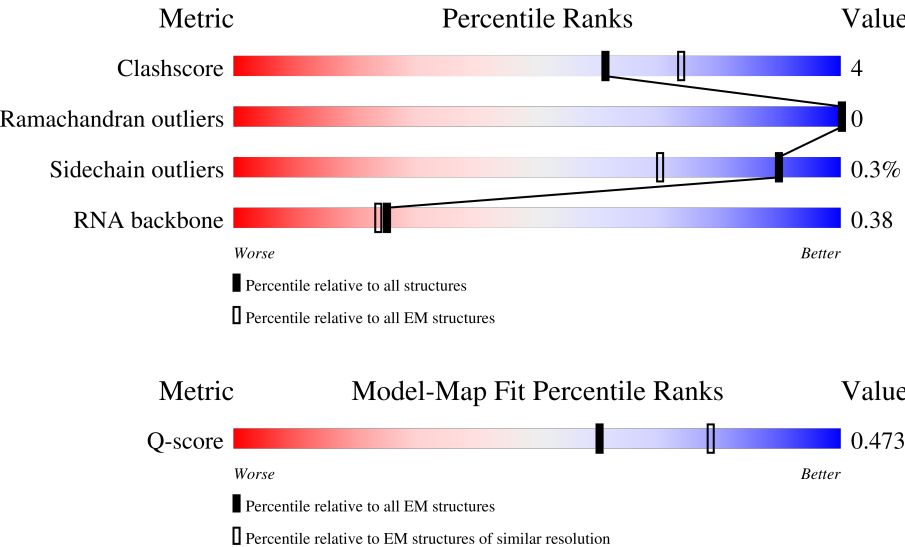
EMDB validation analysis : 0.0.1.dev129  
MolProbity : 4-5-2 with Phenix2.0  
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)  
EM percentile statistics : 202505.v01 (Using data in the EMDB archive up until May 2025)  
MapQ : 1.9.13  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.46

# 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 3.70 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.



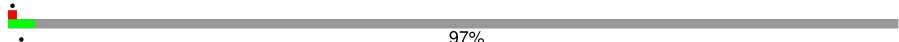
Metric	Whole archive (#Entries)	EM structures (#Entries)	Similar EM resolution (#Entries, resolution range(Å))
Clashscore	210492	15764	-
Ramachandran outliers	207382	16835	-
Sidechain outliers	206894	16415	-
RNA backbone	6643	2191	-
Q-score	-	25397	11569 ( 3.20 - 4.20 )

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  $\geq 3$ , 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  $\leq 5\%$ . 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	899	<div> <div>76%</div> <div>7%</div> <div>17%</div> </div>
2	B	25	<div> <div>40%</div> <div>32%</div> <div>8%</div> <div>12%</div> <div>8%</div> </div>
3	C	27	<div> <div>37%</div> <div>19%</div> <div>22%</div> <div>7%</div> <div>15%</div> </div>

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
4	E	149	 97%

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 7057 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 Piwi-like protein Siwi.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	748	Total	C	N	O	S	0	0
			6040	3850	1055	1107	28		

- Molecule 2 is a RNA chain called piRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	23	Total	C	N	O	P	0	0
			498	221	89	165	23		

- Molecule 3 is a RNA chain called target RNA (25-nt).

Mol	Chain	Residues	Atoms					AltConf	Trace
3	C	23	Total	C	N	O	P	0	0
			479	216	82	158	23		

- Molecule 4 is a protein called GTSF1.

Mol	Chain	Residues	Atoms				AltConf	Trace
4	E	4	Total	C	N	O	0	0
			39	24	6	9		

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

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



ILE	LEU	MET
ILE	ASP	GLU
THR	PHE	VAL
LYS	HIS	THR
ALA	VAL	ILE
LYS	THR	ALA
GLY	VAL	ASN
THR	CYS	PRO
THR	PRO	LYS
ARG	LYS	PRO
SER	ARG	ASN
GLU	HIS	GLN
GLN	MET	MET
SEN	LEU	ILE
SER	ASP	THR
LEU	THR	CYS
ARG	GLN	PRO
LYS	LEU	TYR
GLU	TYR	GLU
LEU	ILE	LYS
ILE	THR	ALA
LYS	ASP	HIS
THR	ASP	ILE
TYR	GLU	VAL
ARG	TYR	GLU
PRO	ARG	HIS
LEU	PRO	TYR
GLU	THR	ARG
	VAL	MET
	GLU	HIS
	VAL	ILE
	HIS	HIS
	ALA	LEU
	THR	GLN
	PRO	LYS
	VAL	CYS
	LEU	ARG
	PRO	LYS
	SER	GLN
	SER	HIS
	ASP	PRO
	E101	ALA
	D104	CYS
	ASP	ASN
	GLU	LYS
	THR	VAL
	SER	GLN
	THR	CYS
	SER	PRO
	TYR	PHE
	ILE	ASP
	PRO	ALA
	ASP	THR
	PRO	HIS
	SER	VAL
	LYS	ASN
	GLY	ASP
	ALA	VAL
	THR	THR

## 4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	161717	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TALOS ARCTICA	Depositor
Voltage (kV)	200	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	50	Depositor
Minimum defocus (nm)	1000	Depositor
Maximum defocus (nm)	3000	Depositor
Magnification	Not provided	
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor
Maximum map value	3.495	Depositor
Minimum map value	-2.447	Depositor
Average map value	0.001	Depositor
Map value standard deviation	0.071	Depositor
Recommended contour level	0.4	Depositor
Map size (Å)	273.0, 273.0, 273.0	wwPDB
Map dimensions	300, 300, 300	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	0.91, 0.91, 0.91	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	A	0.58	0/6164	1.02	3/8338 (0.0%)
2	B	0.60	0/557	1.20	6/866 (0.7%)
3	C	0.64	0/533	1.22	7/825 (0.8%)
4	E	0.72	0/40	1.19	0/54
All	All	0.58	0/7294	1.05	16/10083 (0.2%)

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
1	A	0	1
2	B	0	1
All	All	0	2

There are no bond length outliers.

The worst 5 of 16 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	22	G	P-O3'-C3'	-12.03	102.15	120.20
2	B	21	U	P-O3'-C3'	-9.02	106.68	120.20
2	B	20	U	P-O3'-C3'	-8.73	107.11	120.20
3	C	20	A	O4'-C4'-C3'	8.11	112.11	104.00
3	C	22	U	P-O3'-C3'	-7.78	108.53	120.20

There are no chirality outliers.

All (2) planarity outliers are listed below:



Mol	Chain	Res	Type	Group
1	A	739	ARG	Sidechain
2	B	7	G	Sidechain

## 5.2 Too-close contacts [i](#)

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	6040	0	6130	38	0
2	B	498	0	246	10	0
3	C	479	0	249	16	0
4	E	39	0	25	0	0
5	A	1	0	0	0	0
All	All	7057	0	6650	56	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 4.

The worst 5 of 56 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:C:21:C:H2'	3:C:22:U:C6	2.23	0.73
2:B:18:G:C2	3:C:9:A:C2	2.81	0.69
1:A:526:ALA:HB3	1:A:653:SER:H	1.58	0.68
1:A:762:LEU:HD22	1:A:766:TYR:HE1	1.60	0.66
1:A:534:VAL:HG21	1:A:586:VAL:HG11	1.79	0.64

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	740/899 (82%)	709 (96%)	31 (4%)	0	100	100
4	E	2/149 (1%)	2 (100%)	0	0	100	100
All	All	742/1048 (71%)	711 (96%)	31 (4%)	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	A	669/783 (85%)	667 (100%)	2 (0%)	91	94
4	E	4/140 (3%)	4 (100%)	0	100	100
All	All	673/923 (73%)	671 (100%)	2 (0%)	90	94

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

Mol	Chain	Res	Type
1	A	701	ASN
1	A	708	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	541	GLN
1	A	602	ASN
1	A	716	ASN
1	A	376	ASN
1	A	170	GLN

### 5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	B	23/25 (92%)	6 (26%)	1 (4%)
3	C	22/27 (81%)	6 (27%)	0
All	All	45/52 (86%)	12 (26%)	1 (2%)

5 of 12 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
2	B	2	G
2	B	3	G
2	B	19	U
2	B	21	U
2	B	22	G

All (1) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
2	B	1	U

## 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.

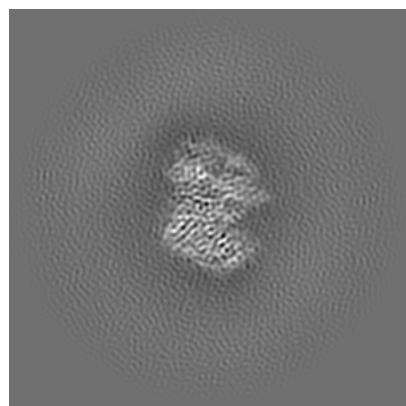
## 6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-49407. These allow visual inspection of the internal detail of the map and identification of artifacts.

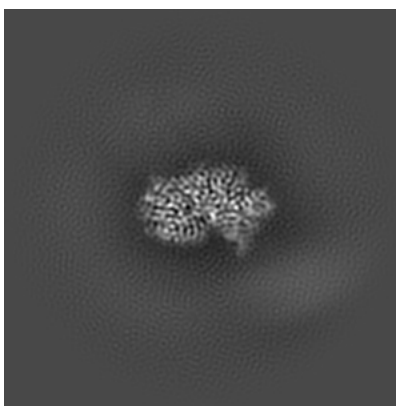
Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

### 6.1 Orthogonal projections [i](#)

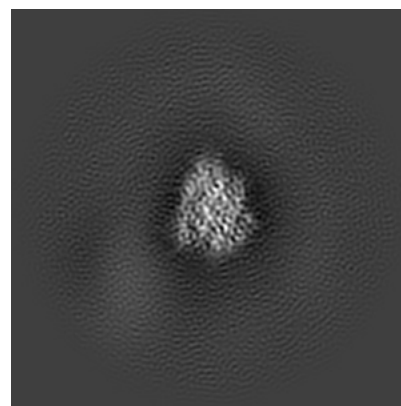
#### 6.1.1 Primary map



X

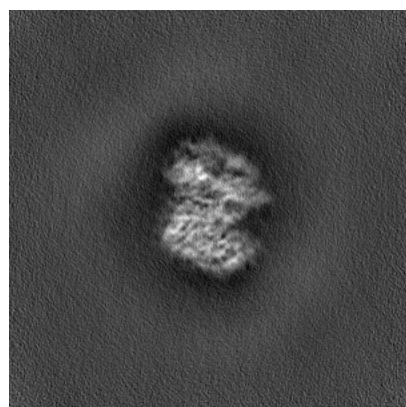


Y

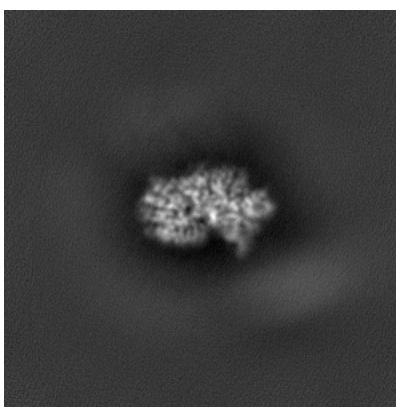


Z

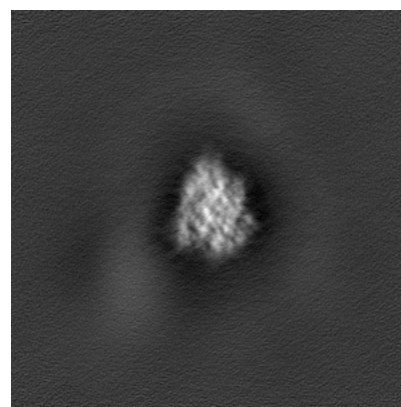
#### 6.1.2 Raw map



X



Y

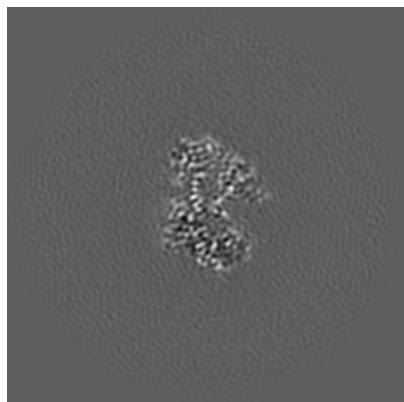


Z

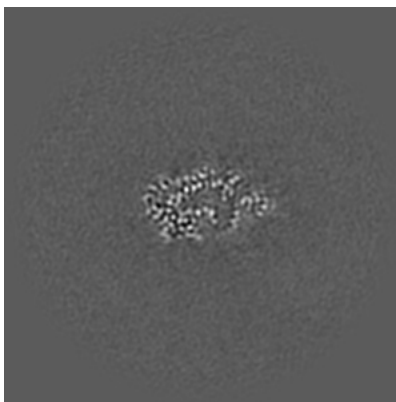
The images above show the map projected in three orthogonal directions.

## 6.2 Central slices [i](#)

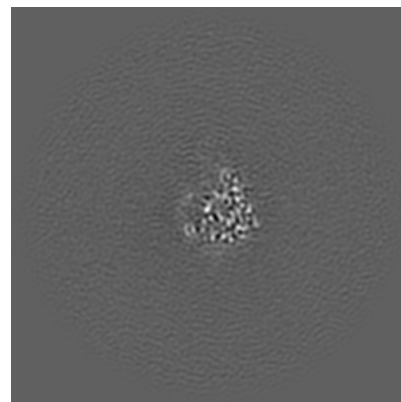
### 6.2.1 Primary map



X Index: 150

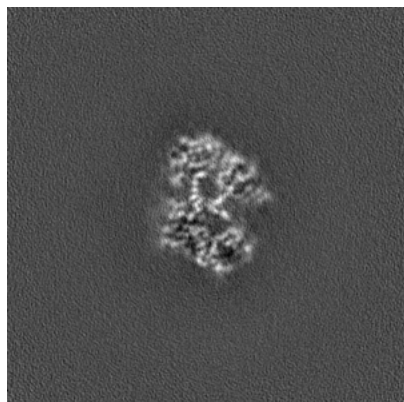


Y Index: 150

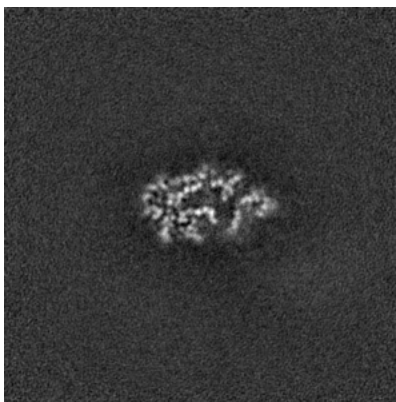


Z Index: 150

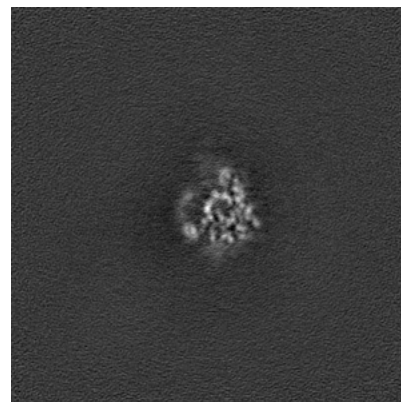
### 6.2.2 Raw map



X Index: 150



Y Index: 150

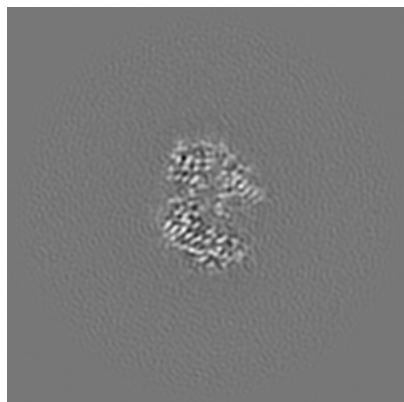


Z Index: 150

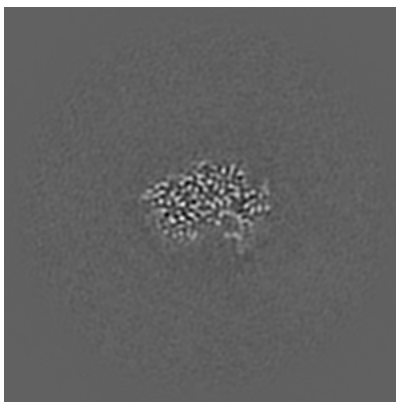
The images above show central slices of the map in three orthogonal directions.

## 6.3 Largest variance slices [i](#)

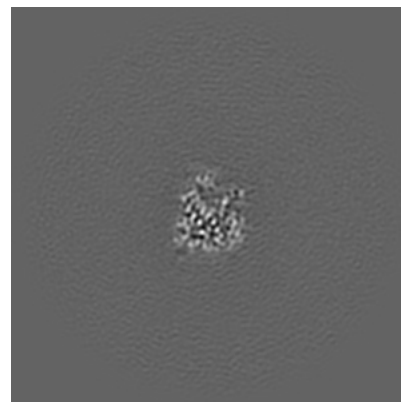
### 6.3.1 Primary map



X Index: 154

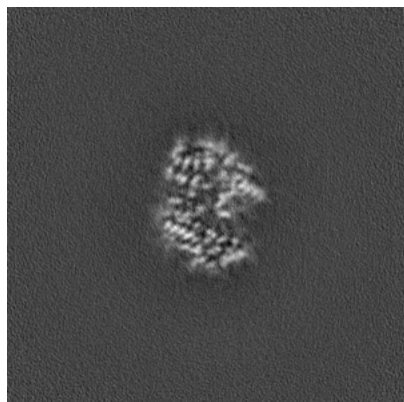


Y Index: 138

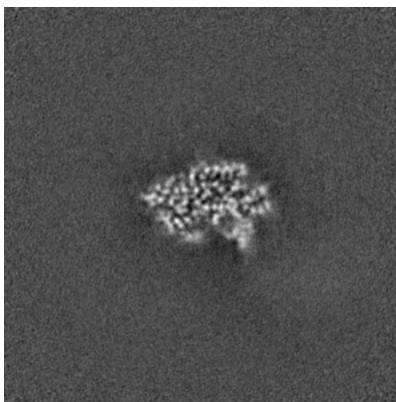


Z Index: 133

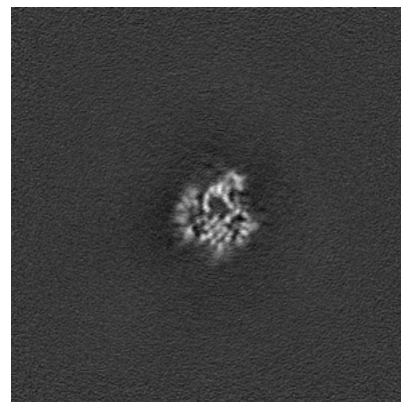
### 6.3.2 Raw map



X Index: 155



Y Index: 141



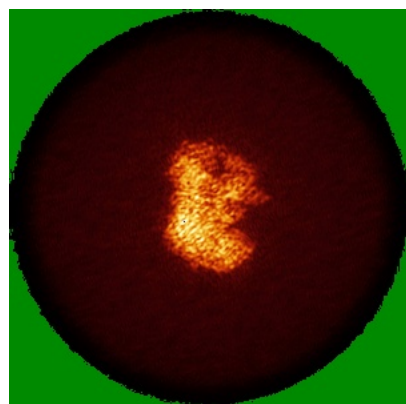
Z Index: 144

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

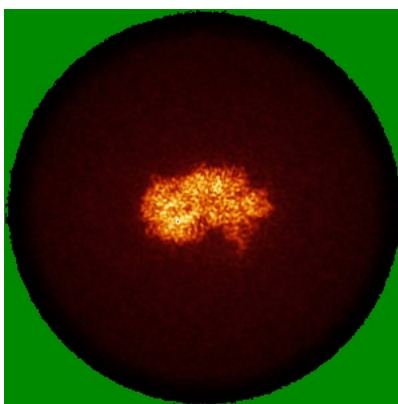


## 6.4 Orthogonal standard-deviation projections (False-color) [i](#)

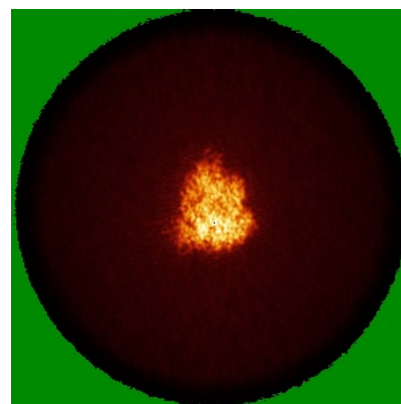
### 6.4.1 Primary map



X

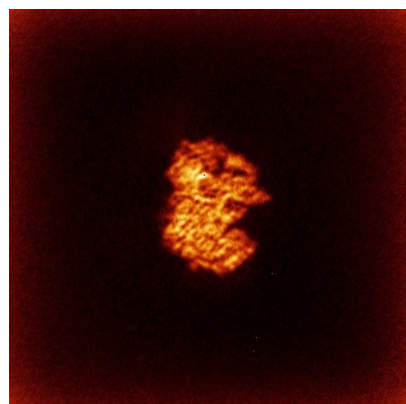


Y

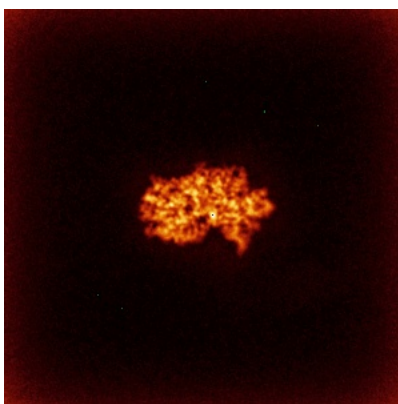


Z

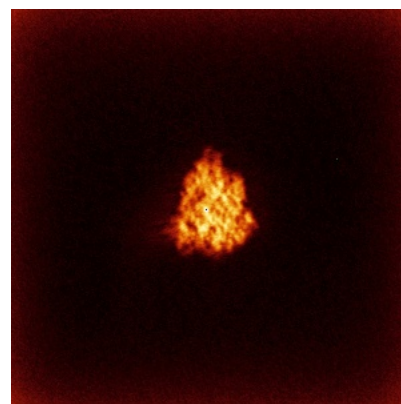
### 6.4.2 Raw map



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



The images above show the 3D surface view of the map at the recommended contour level 0.4. 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



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.

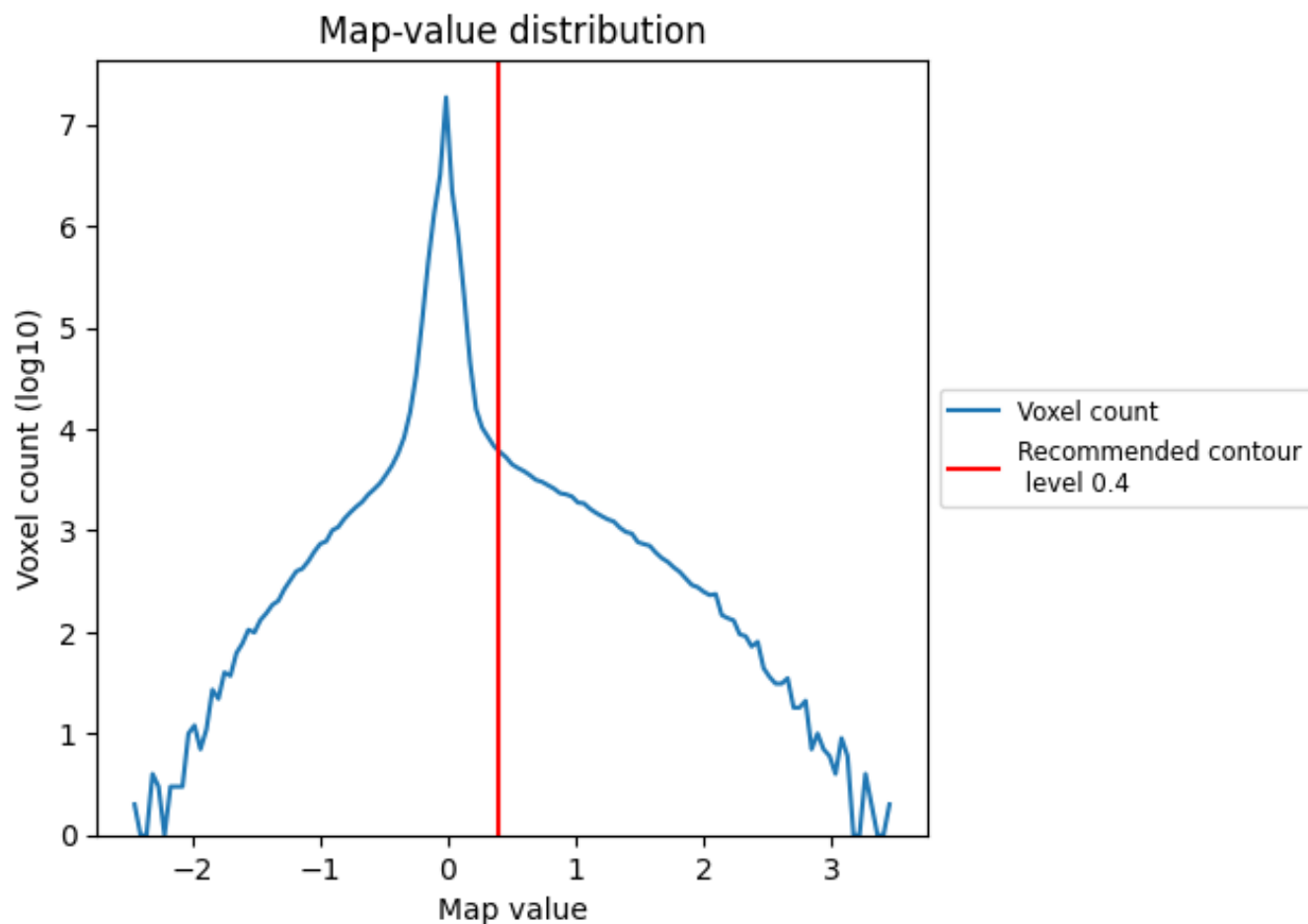
## 6.6 Mask visualisation [i](#)

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

## 7 Map analysis [i](#)

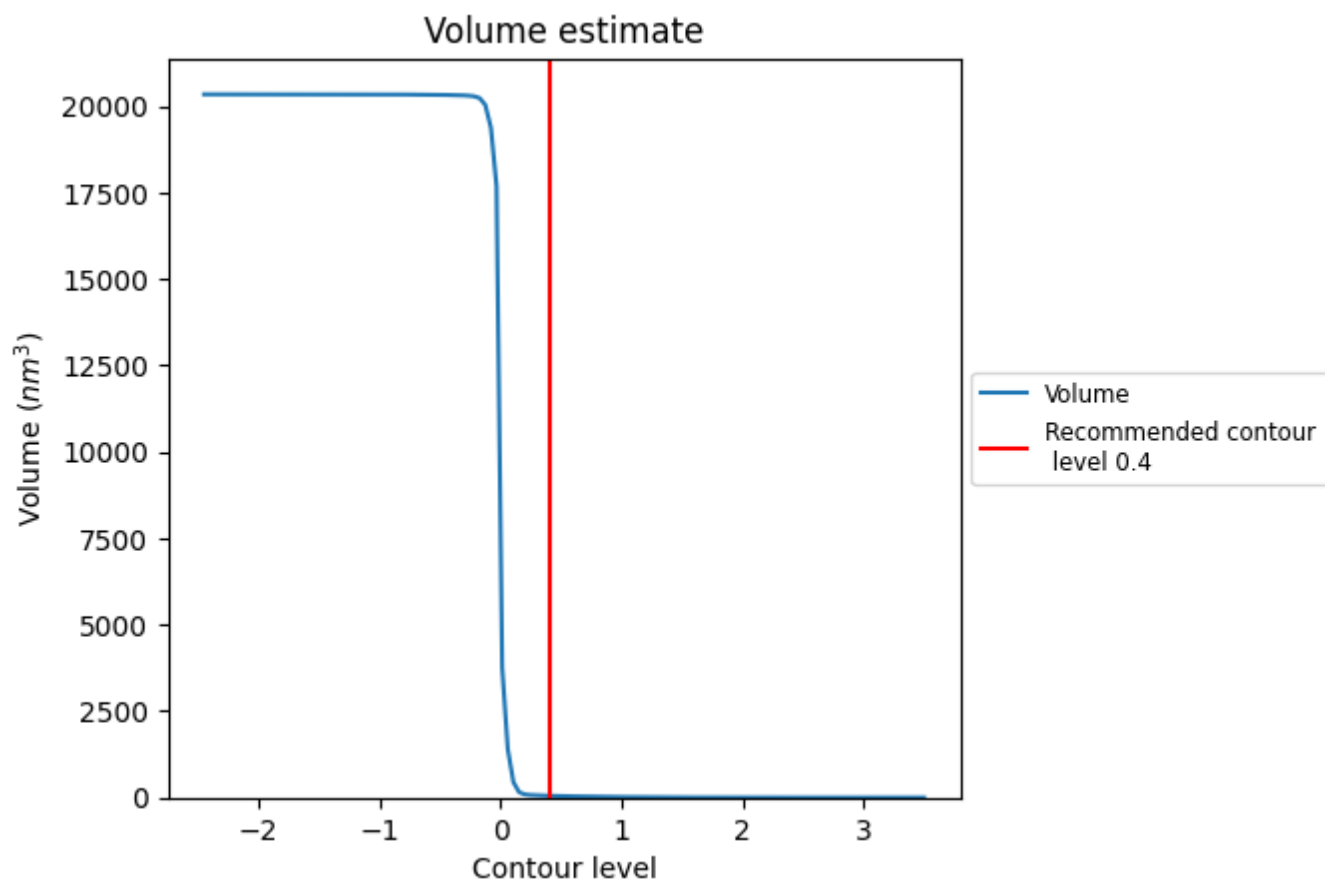
This section contains the results of statistical analysis of the map.

### 7.1 Map-value distribution [i](#)



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.

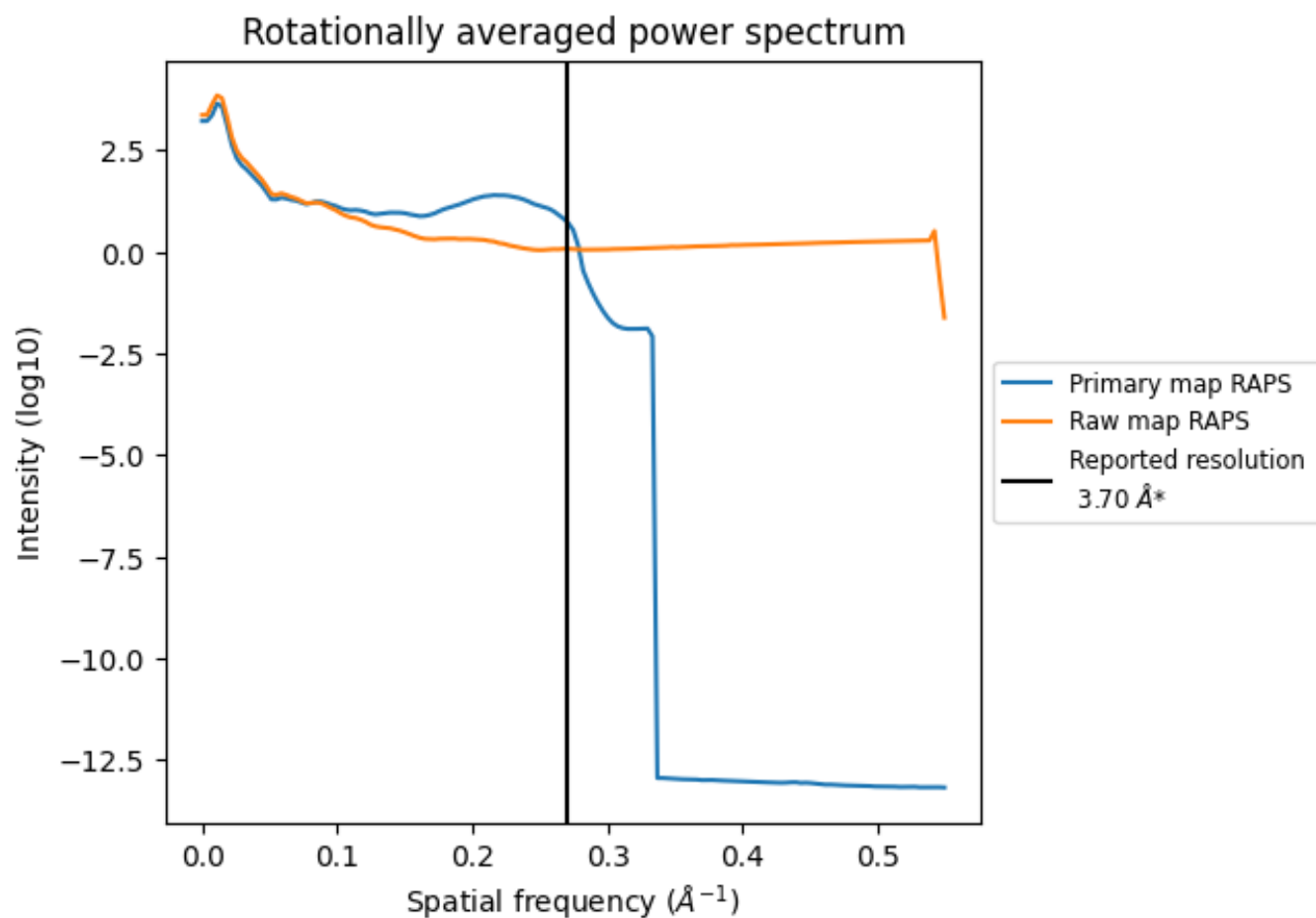
## 7.2 Volume estimate [i](#)



The volume at the recommended contour level is 51  $\text{nm}^3$ ; this corresponds to an approximate mass of 46 kDa.

The volume estimate graph shows how the enclosed volume varies with the contour level. The recommended contour level is shown as a vertical line and the intersection between the line and the curve gives the volume of the enclosed surface at the given level.

### 7.3 Rotationally averaged power spectrum ⓘ

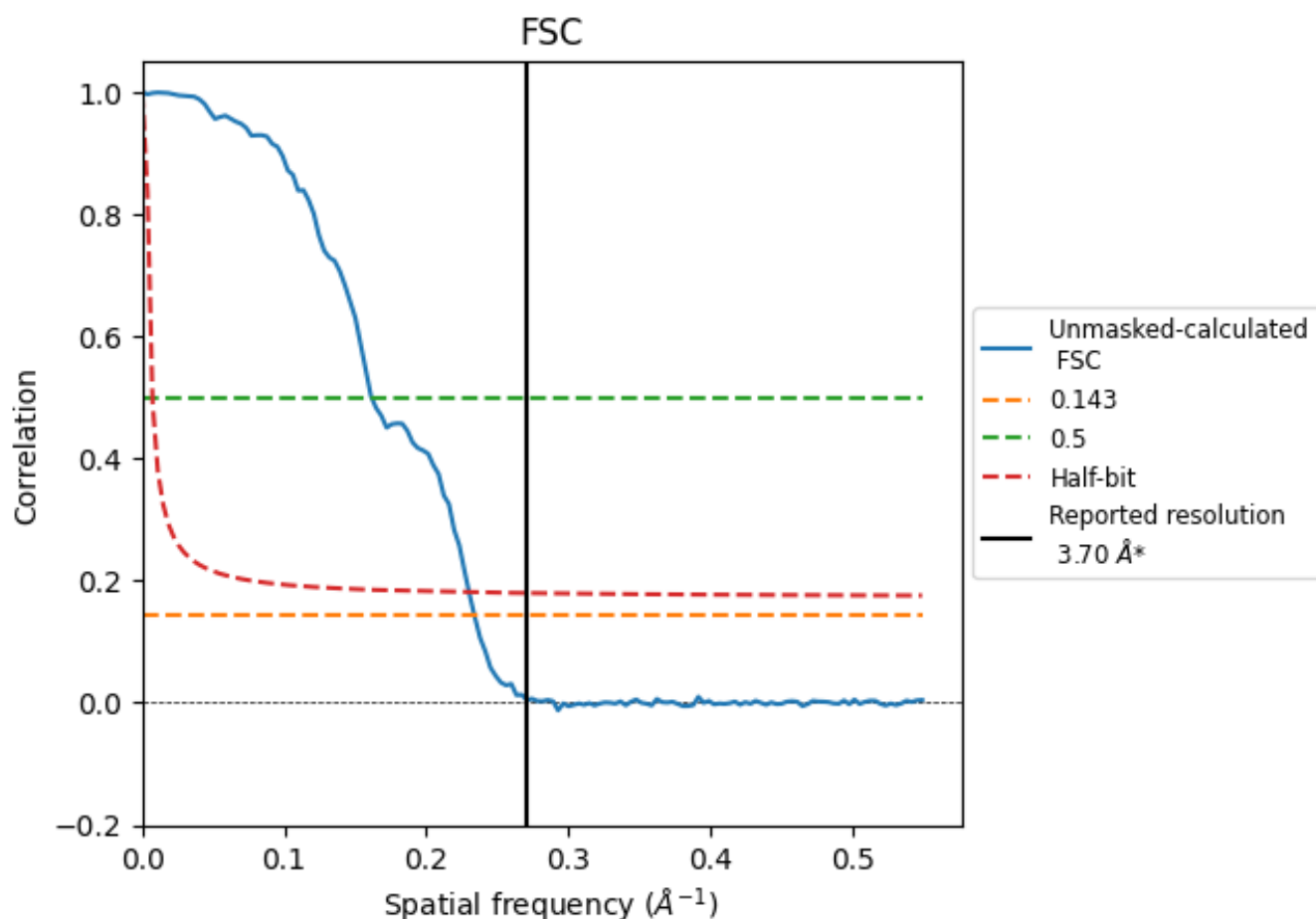


\*Reported resolution corresponds to spatial frequency of 0.270  $\text{\AA}^{-1}$

## 8 Fourier-Shell correlation [i](#)

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 [i](#)



\*Reported resolution corresponds to spatial frequency of 0.270 Å<sup>-1</sup>

## 8.2 Resolution estimates [i](#)

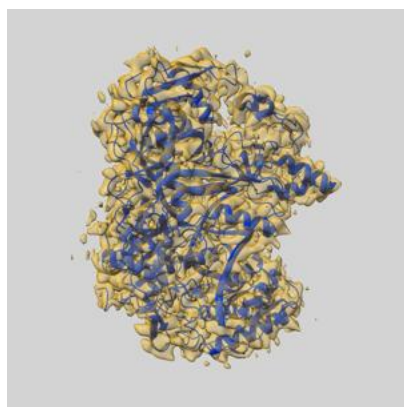
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	3.70	-	-
Author-provided FSC curve	-	-	-
Unmasked-calculated*	4.27	6.19	4.34

\*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 4.27 differs from the reported value 3.7 by more than 10 %

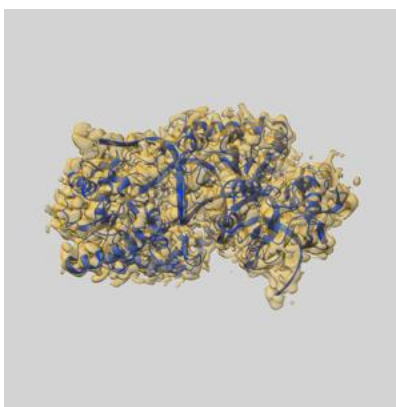
## 9 Map-model fit [i](#)

This section contains information regarding the fit between EMDB map EMD-49407 and PDB model 9NHB. Per-residue inclusion information can be found in [section 3](#) on [page 5](#).

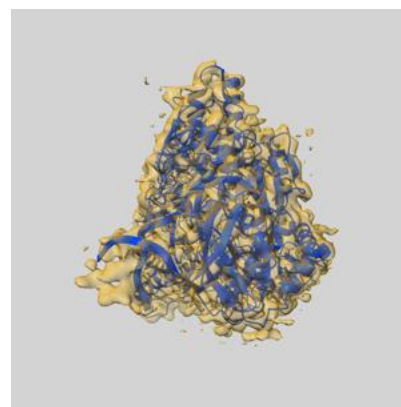
### 9.1 Map-model overlay [i](#)



X



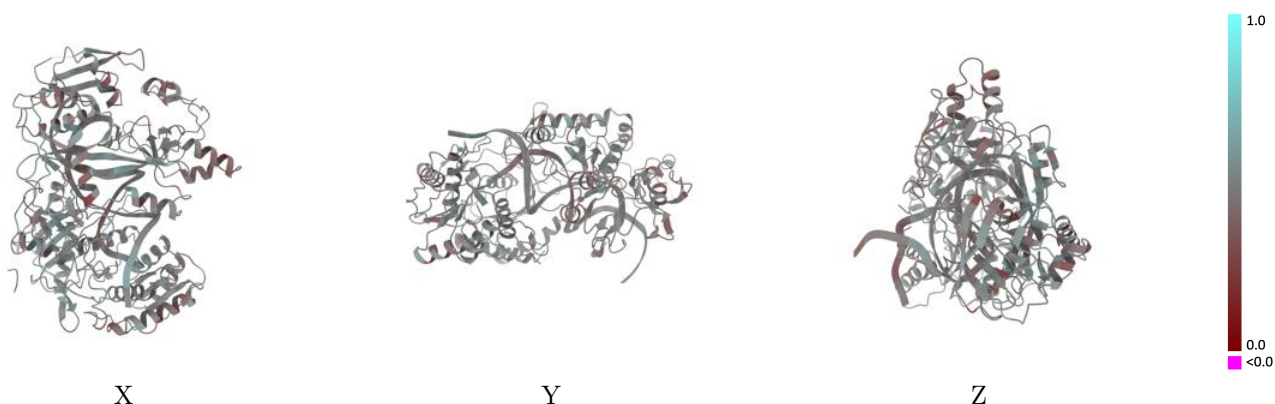
Y



Z

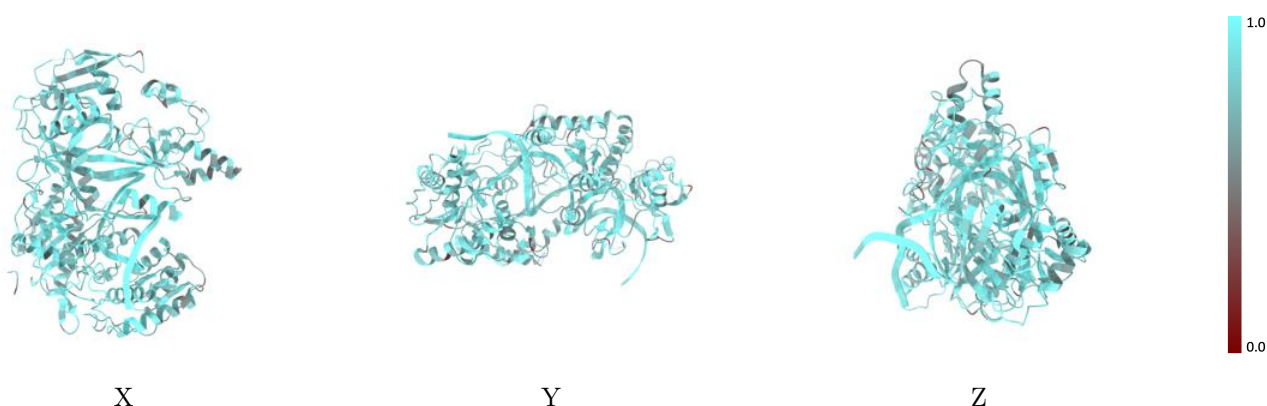
The images above show the 3D surface view of the map at the recommended contour level 0.4 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

## 9.2 Q-score mapped to coordinate model [i](#)



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.

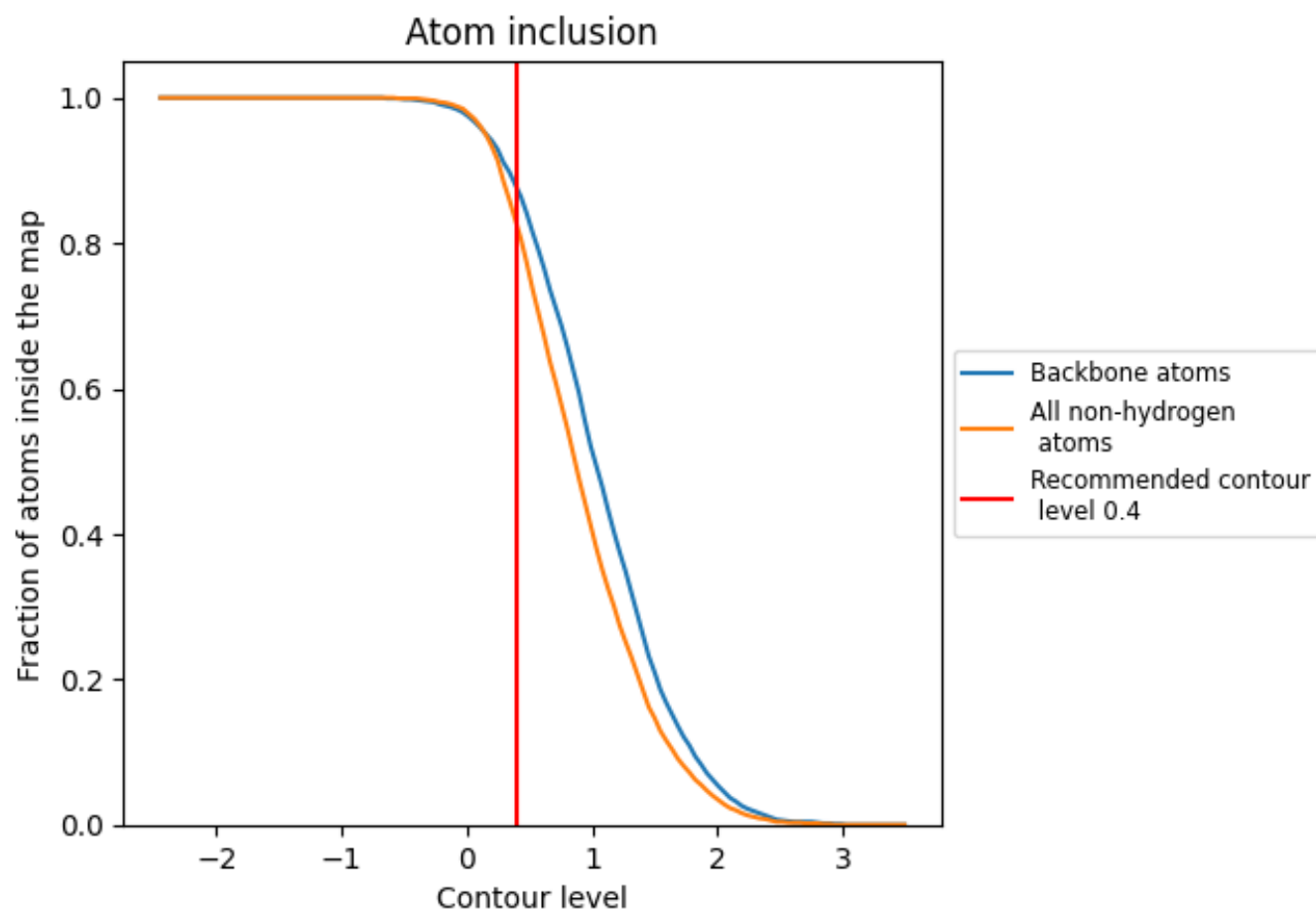
## 9.3 Atom inclusion mapped to coordinate model [i](#)



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 (0.4).



## 9.4 Atom inclusion [i](#)



At the recommended contour level, 88% of all backbone atoms, 82% 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 (0.4) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	<div></div> 0.8240	<div></div> 0.4730
A	<div></div> 0.8130	<div></div> 0.4750
B	<div></div> 0.9040	<div></div> 0.4620
C	<div></div> 0.8890	<div></div> 0.4570
E	<div></div> 0.6320	<div></div> 0.4810

