



# wwPDB EM Validation Summary Report ⓘ

Mar 27, 2023 – 10:05 AM EDT

PDB ID : 8FA1  
EMDB ID : EMD-28947  
Title : Cryo-EM structure of the SARS-CoV-2 HR1HR2 fusion core complex with N969K mutation  
Authors : Yang, K.; Brunger, A.T.  
Deposited on : 2022-11-25  
Resolution : 2.51 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](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.dev50  
MolProbity : 4.02b-467  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
MapQ : 1.9.9  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.32.2

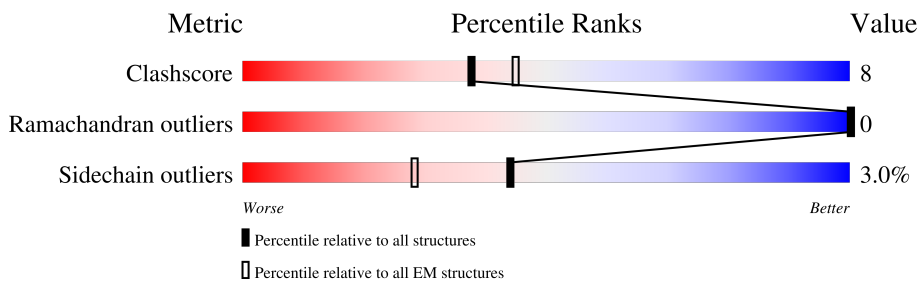
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 2.51 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



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

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	257	
1	B	257	
1	C	257	
2	D	45	
2	E	45	
2	F	45	

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 2553 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 Ferritin, Dps family protein and Spike protein S2' chimera.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
1	A	71	532	329	94	109	0	0
1	B	71	532	329	94	109	0	0
1	C	71	532	329	94	109	0	0

There are 36 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	732	MET	-	initiating methionine	UNP B2J981
A	733	SER	-	expression tag	UNP B2J981
A	734	HIS	-	expression tag	UNP B2J981
A	735	HIS	-	expression tag	UNP B2J981
A	736	HIS	-	expression tag	UNP B2J981
A	737	HIS	-	expression tag	UNP B2J981
A	738	HIS	-	expression tag	UNP B2J981
A	739	HIS	-	expression tag	UNP B2J981
A	740	GLY	-	expression tag	UNP B2J981
A	741	SER	THR	conflict	UNP B2J981
A	916	ALA	-	linker	UNP B2J981
A	969	LYS	ASN	engineered mutation	UNP A0A8B6RK57
B	732	MET	-	initiating methionine	UNP B2J981
B	733	SER	-	expression tag	UNP B2J981
B	734	HIS	-	expression tag	UNP B2J981
B	735	HIS	-	expression tag	UNP B2J981
B	736	HIS	-	expression tag	UNP B2J981
B	737	HIS	-	expression tag	UNP B2J981
B	738	HIS	-	expression tag	UNP B2J981
B	739	HIS	-	expression tag	UNP B2J981
B	740	GLY	-	expression tag	UNP B2J981
B	741	SER	THR	conflict	UNP B2J981
B	916	ALA	-	linker	UNP B2J981
B	969	LYS	ASN	engineered mutation	UNP A0A8B6RK57

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	732	MET	-	initiating methionine	UNP B2J981
C	733	SER	-	expression tag	UNP B2J981
C	734	HIS	-	expression tag	UNP B2J981
C	735	HIS	-	expression tag	UNP B2J981
C	736	HIS	-	expression tag	UNP B2J981
C	737	HIS	-	expression tag	UNP B2J981
C	738	HIS	-	expression tag	UNP B2J981
C	739	HIS	-	expression tag	UNP B2J981
C	740	GLY	-	expression tag	UNP B2J981
C	741	SER	THR	conflict	UNP B2J981
C	916	ALA	-	linker	UNP B2J981
C	969	LYS	ASN	engineered mutation	UNP A0A8B6RKS7

- Molecule 2 is a protein called Spike protein S2' HR2.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
2	D	42	319	195	55	69	0	0
2	E	42	319	195	55	69	0	0
2	F	42	319	195	55	69	0	0

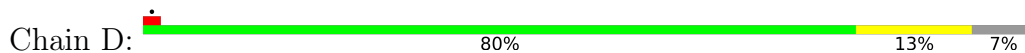


GLU GLY  
GLN SER  
GLU PHE  
TYR TYR  
SER LEU  
HIS HIS  
GLU GLU  
PHE PHE  
PHE PHE  
ASN ASN  
GLU GLU  
ILE ILE  
SER SER  
GLY TYR  
VAL ASN  
Gln Gln  
VAL VAL  
GLN GLN  
ASP ASP  
HIS HIS  
ILE ILE  
HIS HIS  
GLU GLU  
ILE ILE  
GLY GLY  
SER SER  
GLY ARG  
LEU LEU  
ASP ASP  
GLY GLY  
THR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

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

THR LEU  
GLY PHE  
ALA ALA  
TYR TYR  
E918  
Q935  
L938  
L948  
N960  
V963  
Q964  
Q965  
L966  
K969  
A972  
V976  
L977  
N978  
L981  
S982  
R983  
L984  
D985  
E986

● Molecule 2: Spike protein S2' HR2



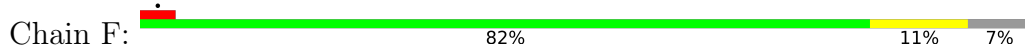
LYS ASN  
H1159  
D1163  
I1169  
V1177  
N1178  
I1179  
L1186  
A1190  
L1200  
GLN

● Molecule 2: Spike protein S2' HR2



LYS ASN  
H1159  
I1160  
D1163  
G1167  
D1168  
I1169  
S1170  
G1171  
V1177  
N1178  
I1179  
L1186  
A1190  
D1199  
L1200  
GLN

● Molecule 2: Spike protein S2' HR2



LYS ASN  
H1159  
D1163  
V1177  
N1178  
I1179  
L1186  
A1190  
D1199  
L1200  
GLN

## 4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	883152	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	NONE	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	55	Depositor
Minimum defocus (nm)	300	Depositor
Maximum defocus (nm)	2000	Depositor
Magnification	Not provided	
Image detector	GATAN K3 (6k x 4k)	Depositor
Maximum map value	1.664	Depositor
Minimum map value	-0.002	Depositor
Average map value	0.001	Depositor
Map value standard deviation	0.020	Depositor
Recommended contour level	0.05	Depositor
Map size ( $\text{\AA}$ )	205.824, 205.824, 205.824	wwPDB
Map dimensions	256, 256, 256	wwPDB
Map angles ( $^\circ$ )	90.0, 90.0, 90.0	wwPDB
Pixel spacing ( $\text{\AA}$ )	0.804, 0.804, 0.804	Depositor

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

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	0.29	0/533	0.46	0/717
1	B	0.29	0/533	0.46	0/717
1	C	0.29	0/533	0.47	0/717
2	D	0.28	0/320	0.44	0/434
2	E	0.28	0/320	0.44	0/434
2	F	0.28	0/320	0.44	0/434
All	All	0.29	0/2559	0.45	0/3453

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 [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	532	0	552	15	0
1	B	532	0	552	16	0
1	C	532	0	552	12	0
2	D	319	0	317	7	0
2	E	319	0	317	7	0
2	F	319	0	317	6	0
All	All	2553	0	2607	42	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 8.



The worst 5 of 42 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:C:978:ASN:ND2	2:F:1163:ASP:OD1	2.29	0.66
1:B:978:ASN:ND2	2:E:1163:ASP:OD1	2.33	0.60
1:A:948:LEU:HD23	2:F:1177:VAL:HG21	1.84	0.59
1:C:948:LEU:HD23	2:E:1177:VAL:HG21	1.84	0.58
1:B:948:LEU:HD23	2:D:1177:VAL:HG21	1.84	0.58

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	69/257 (27%)	69 (100%)	0	0	100	100
1	B	69/257 (27%)	69 (100%)	0	0	100	100
1	C	69/257 (27%)	69 (100%)	0	0	100	100
2	D	40/45 (89%)	35 (88%)	5 (12%)	0	100	100
2	E	40/45 (89%)	35 (88%)	5 (12%)	0	100	100
2	F	40/45 (89%)	35 (88%)	5 (12%)	0	100	100
All	All	327/906 (36%)	312 (95%)	15 (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	
1	A	61/220 (28%)	58 (95%)	3 (5%)	25	47
1	B	61/220 (28%)	58 (95%)	3 (5%)	25	47
1	C	61/220 (28%)	58 (95%)	3 (5%)	25	47
2	D	38/41 (93%)	38 (100%)	0	100	100
2	E	38/41 (93%)	38 (100%)	0	100	100
2	F	38/41 (93%)	38 (100%)	0	100	100
All	All	297/783 (38%)	288 (97%)	9 (3%)	44	68

5 of 9 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	978	ASN
1	C	983	ARG
1	B	969	LYS
1	B	978	ASN
1	B	983	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
2	E	1159	HIS
2	F	1159	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 monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

There are no ligands in this entry.

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-28947. These allow visual inspection of the internal detail of the map and identification of artifacts.

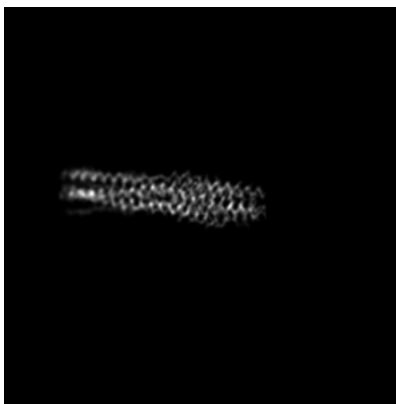
Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

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

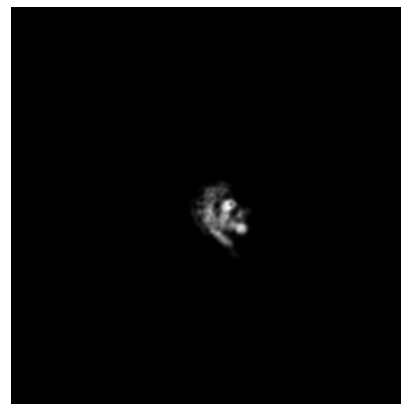
#### 6.1.1 Primary map



X

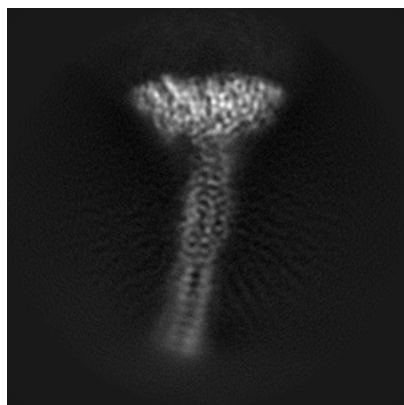


Y

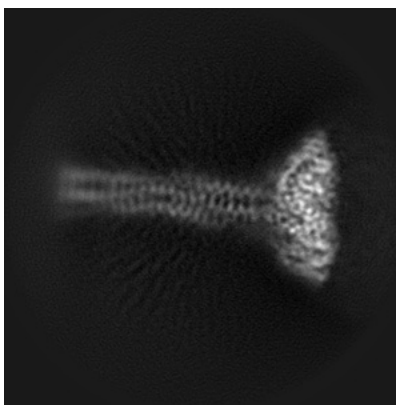


Z

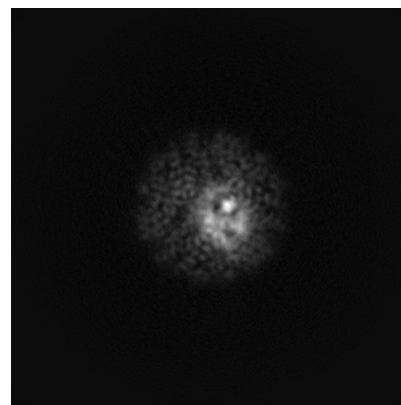
#### 6.1.2 Raw map



X



Y



Z

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

## 6.2 Central slices [i](#)

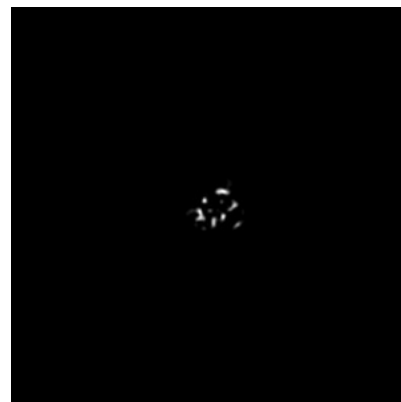
### 6.2.1 Primary map



X Index: 128

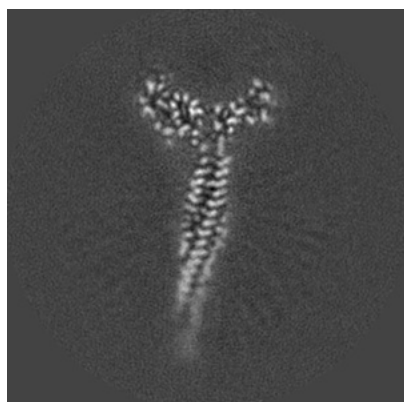


Y Index: 128

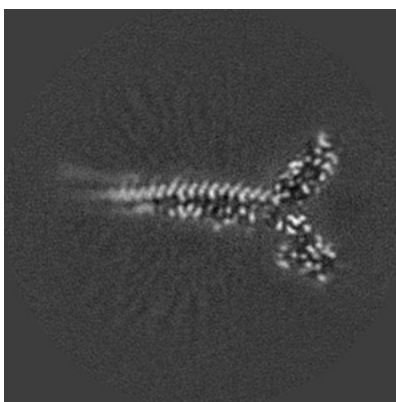


Z Index: 128

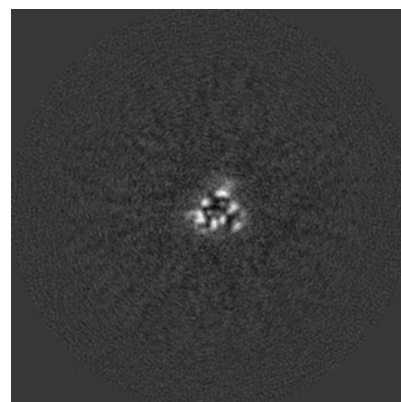
### 6.2.2 Raw map



X Index: 128



Y Index: 128

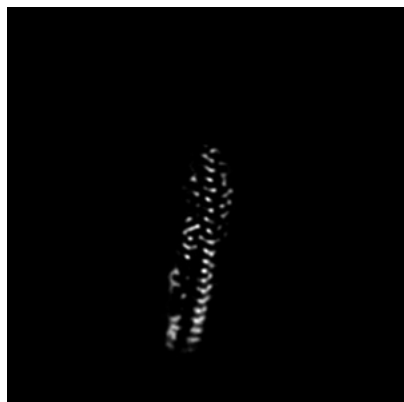


Z Index: 128

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

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

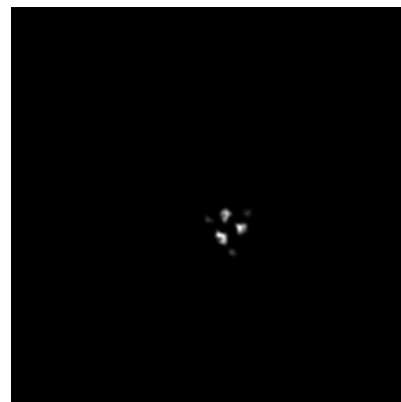
### 6.3.1 Primary map



X Index: 137



Y Index: 115

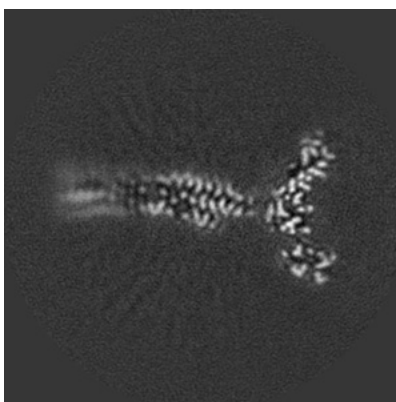


Z Index: 56

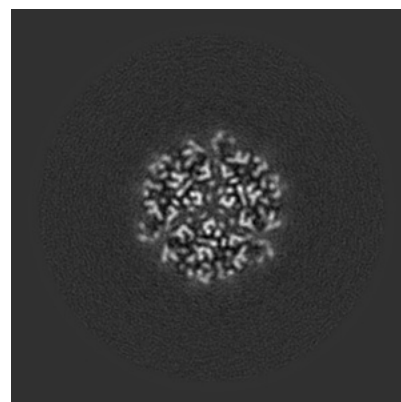
### 6.3.2 Raw map



X Index: 139



Y Index: 120

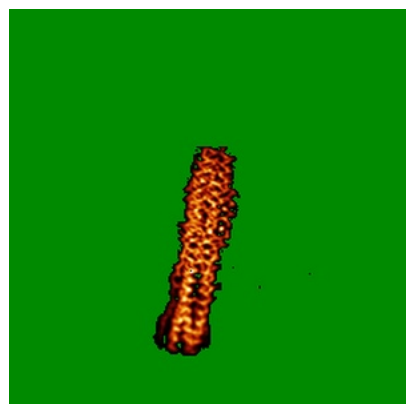


Z Index: 195

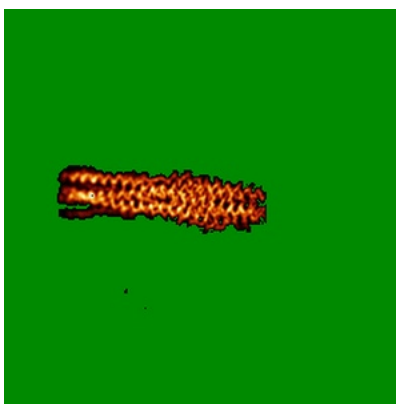
The images above show the largest variance slices of the map in three orthogonal directions.

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

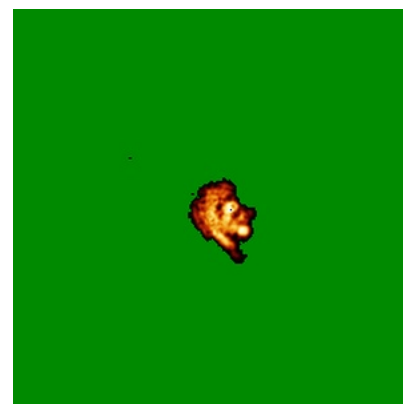
### 6.4.1 Primary map



X



Y

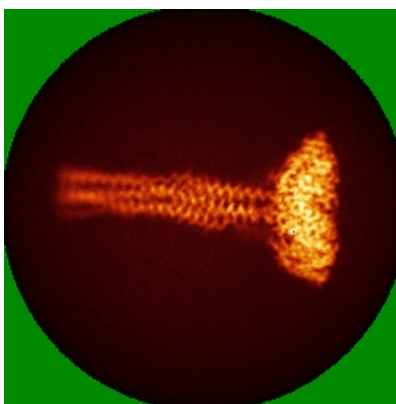


Z

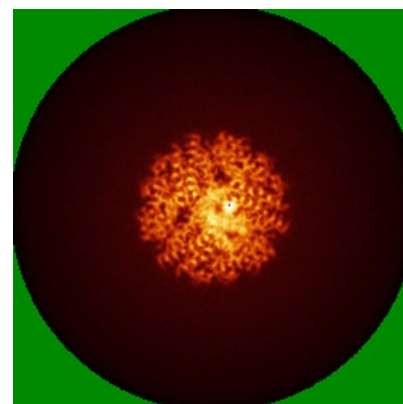
### 6.4.2 Raw map



X



Y

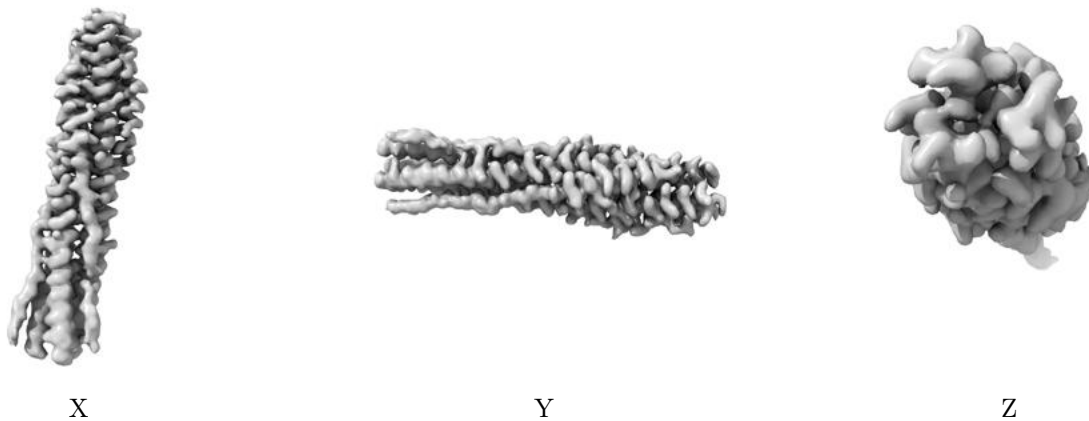


Z

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.

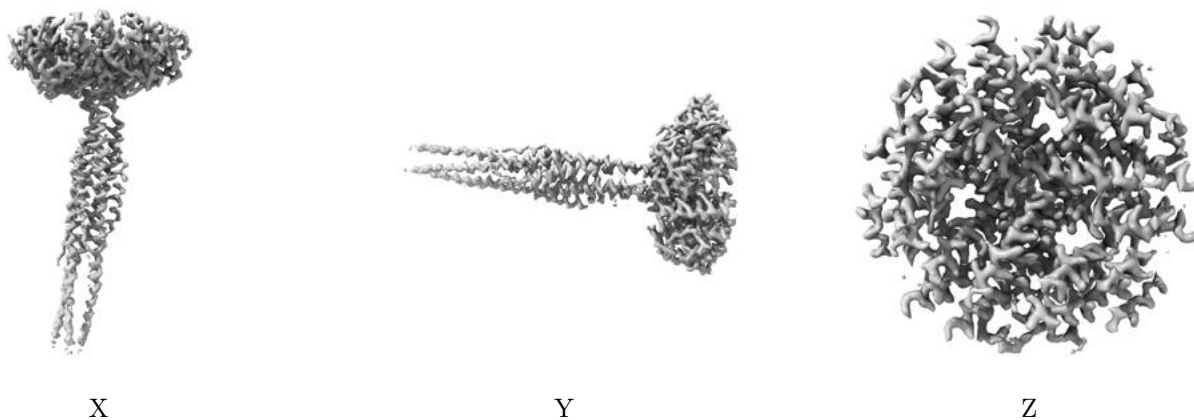
## 6.5 Orthogonal surface views [i](#)

### 6.5.1 Primary map



The images above show the 3D surface view of the map at the recommended contour level 0.05. 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](#)

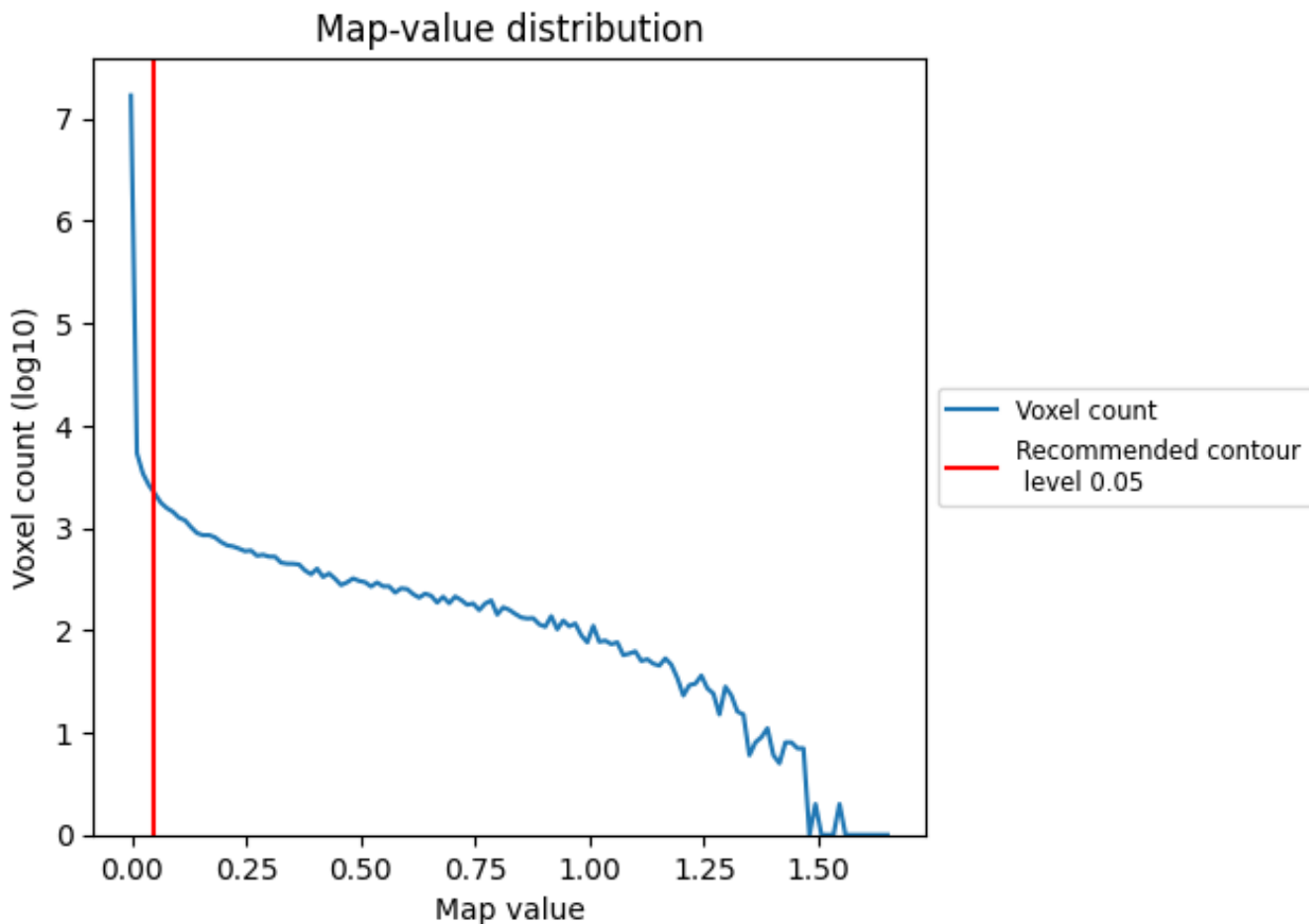
This section was not generated. No masks/segmentation were deposited.



## 7 Map analysis [i](#)

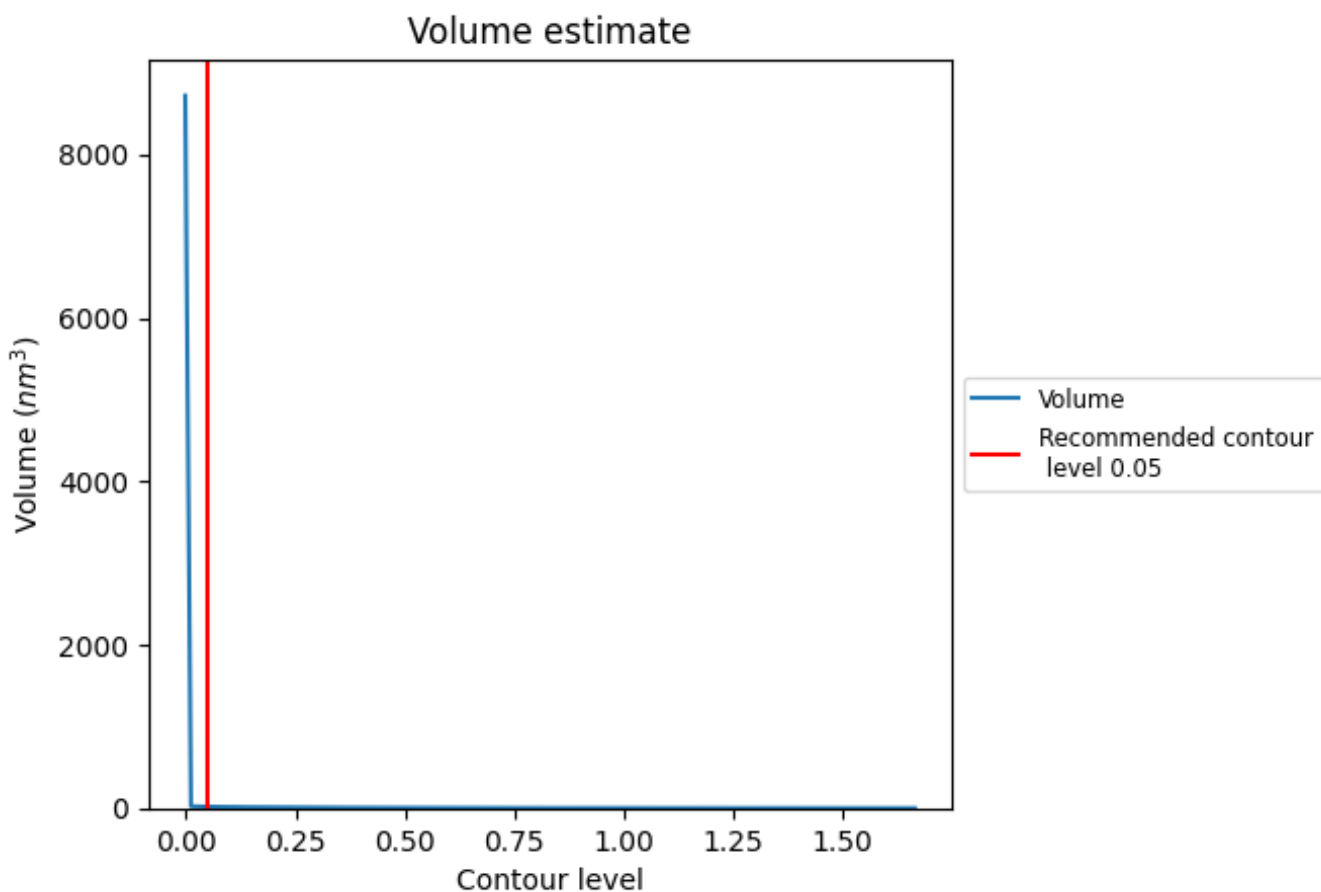
This section contains the results of statistical analysis of the map.

### 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.

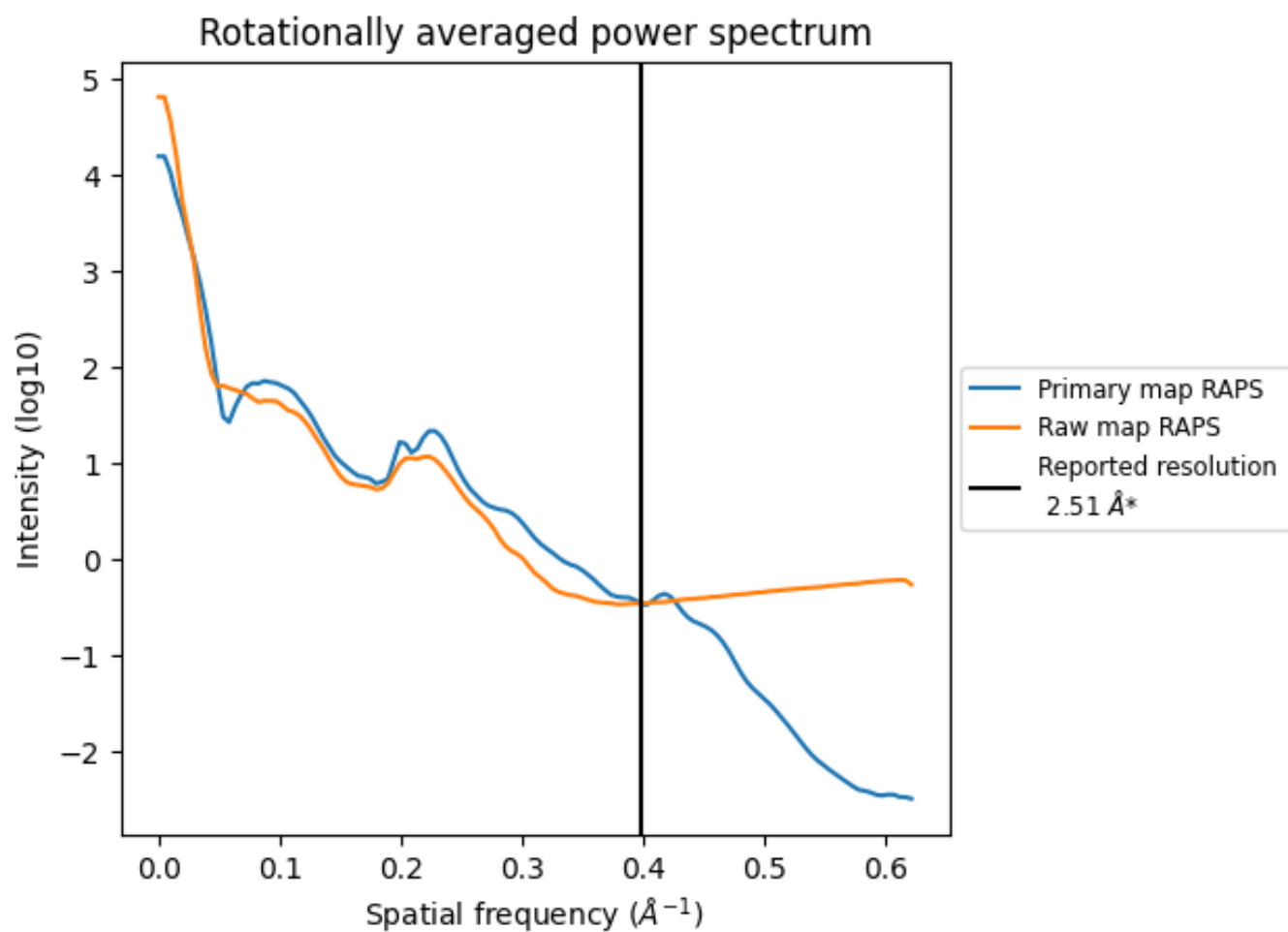
## 7.2 Volume estimate [i](#)



The volume at the recommended contour level is 17 nm<sup>3</sup>; this corresponds to an approximate mass of 16 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](#)

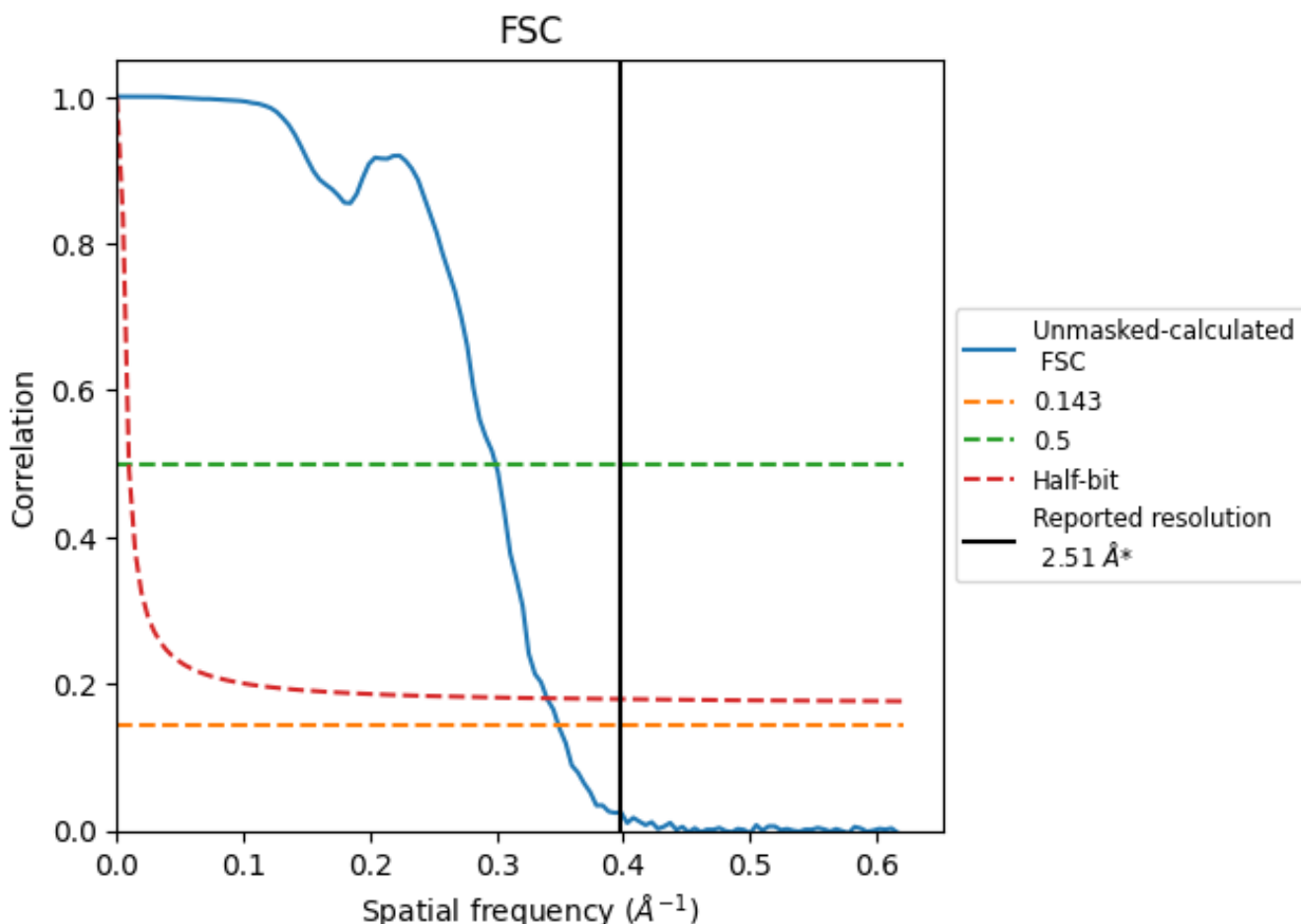


\*Reported resolution corresponds to spatial frequency of 0.398 Å<sup>-1</sup>

## 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 0.398  $\text{\AA}^{-1}$

## 8.2 Resolution estimates [i](#)

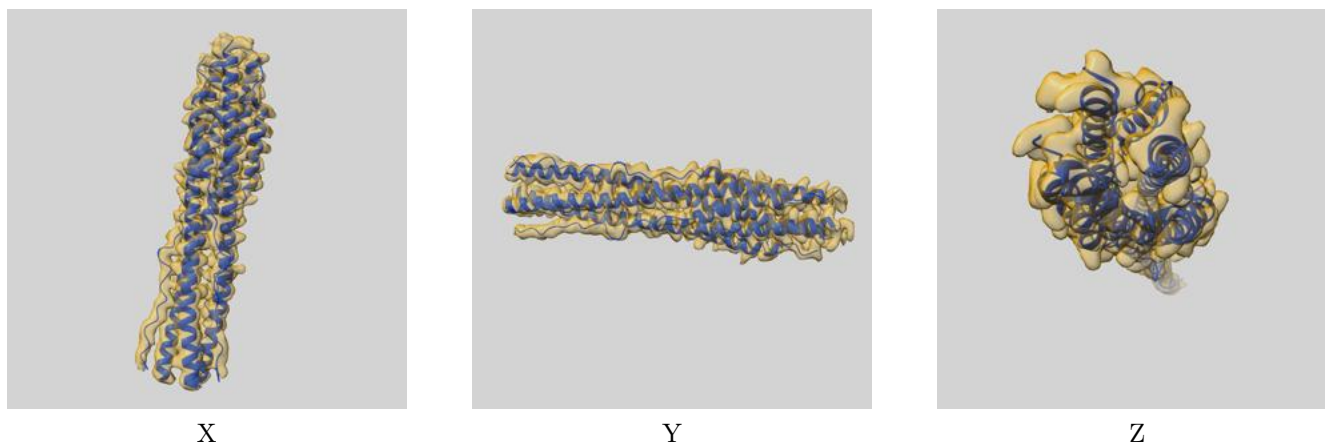
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	2.51	-	-
Author-provided FSC curve	-	-	-
Unmasked-calculated*	2.86	3.34	2.94

\*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 2.86 differs from the reported value 2.51 by more than 10 %

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

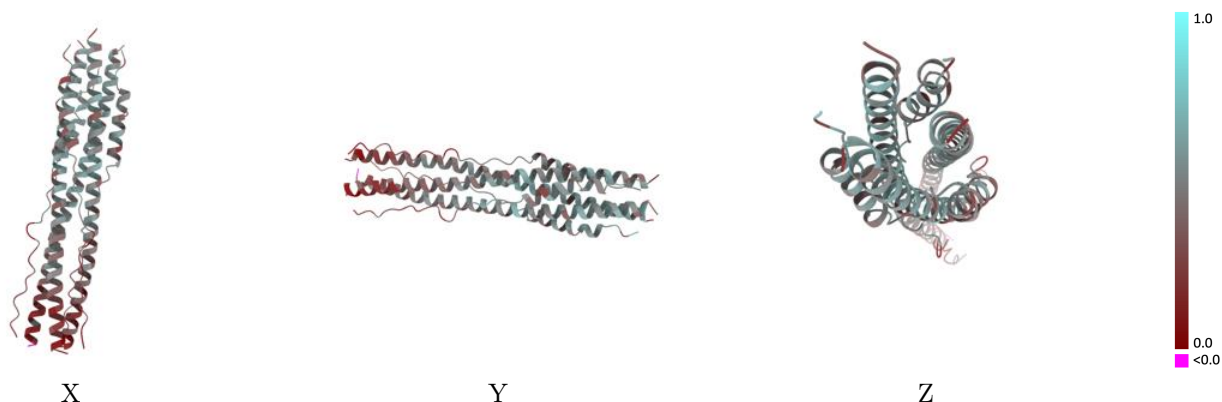
This section contains information regarding the fit between EMDB map EMD-28947 and PDB model 8FA1. Per-residue inclusion information can be found in section 3 on page 5.

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



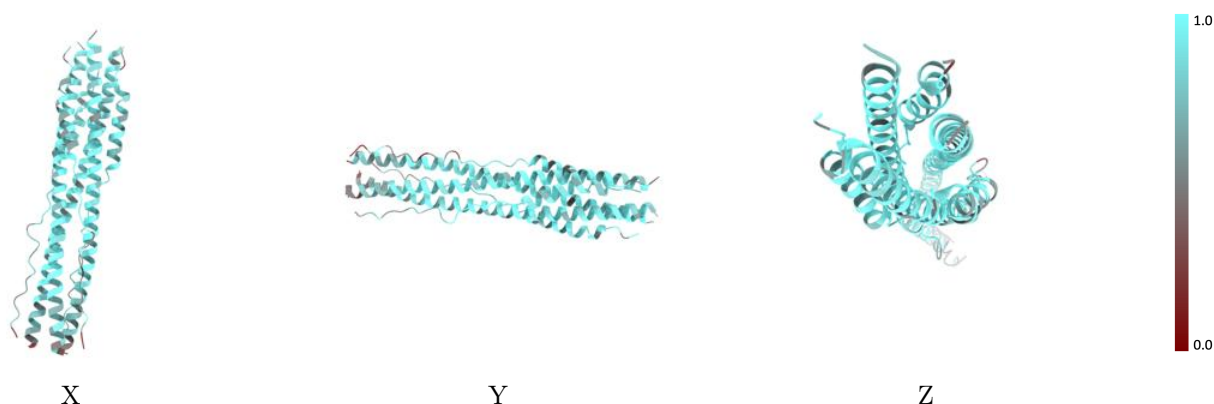
The images above show the 3D surface view of the map at the recommended contour level 0.05 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

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



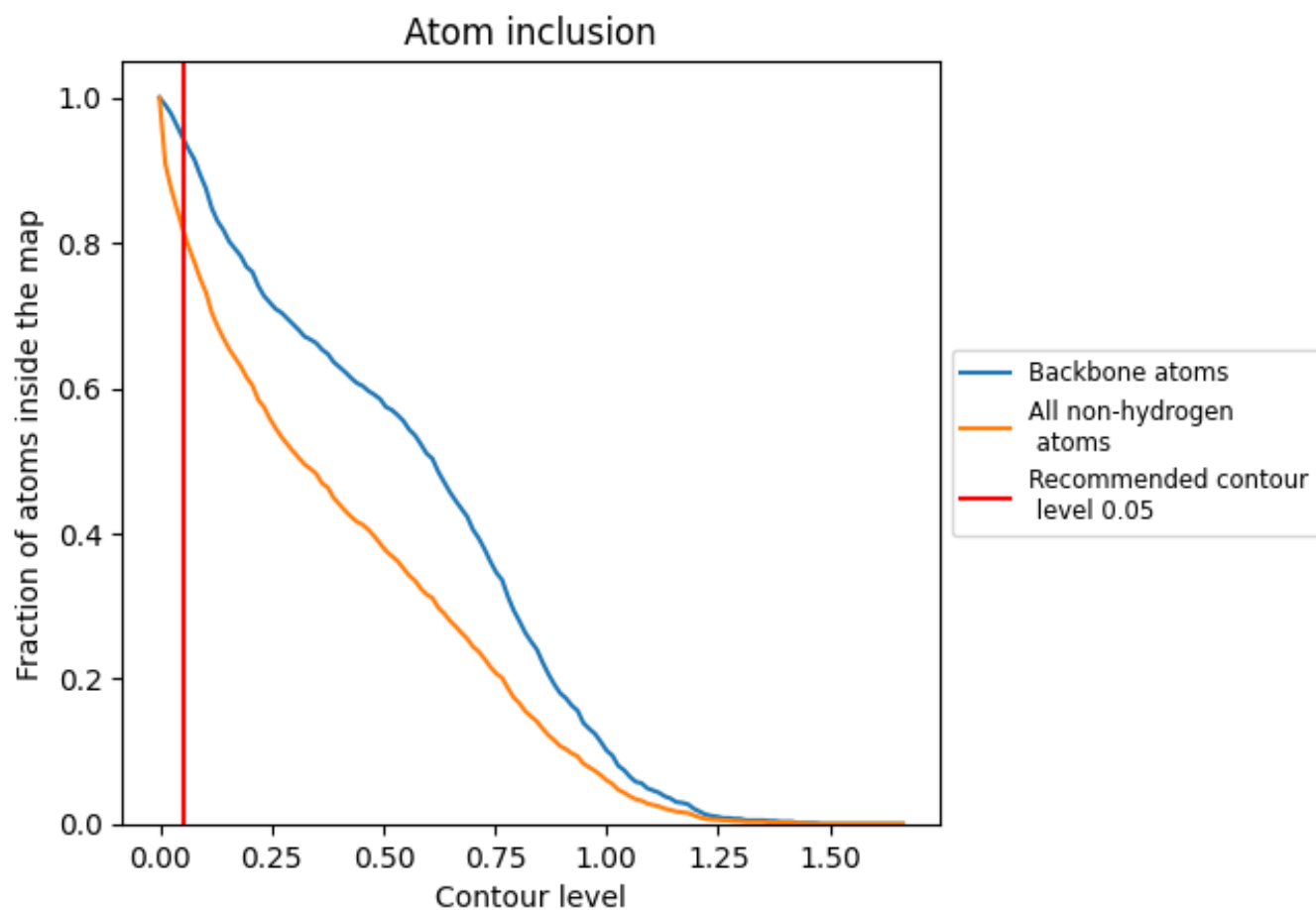
The images above show the model with each residue coloured according to its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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 (0.05).

## 9.4 Atom inclusion [i](#)

















At the recommended contour level, 94% 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 (0.05) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	 0.8190	 0.4270
A	 0.8510	 0.4470
B	 0.8590	 0.4430
C	 0.8380	 0.4240
D	 0.7980	 0.4360
E	 0.7190	 0.3770
F	 0.7890	 0.4150

