



## wwPDB EM Validation Summary Report ⓘ

Mar 23, 2022 – 08:03 pm GMT

PDB ID : 7Q3L  
EMDB ID : EMD-13793  
Title : Human 17S U2 snRNP 5' domain  
Authors : Tholen, J.; Galej, W.P.  
Deposited on : 2021-10-28  
Resolution : 2.30 Å(reported)  
Based on initial models : 6Y5Q, 6EN4

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

EMDB validation analysis : 0.0.0.dev97  
MolProbity : 4.02b-467  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.27

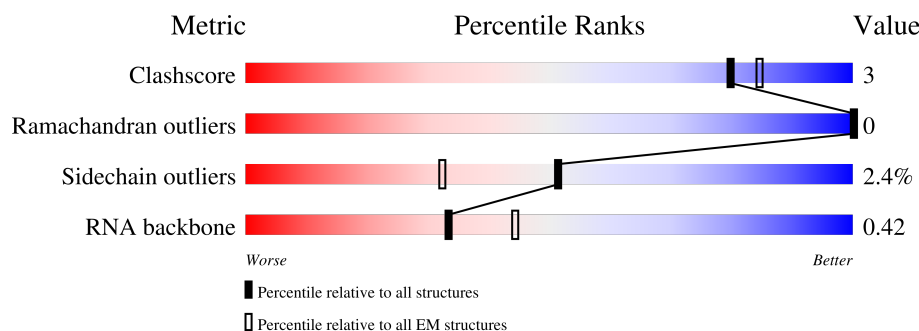
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 2.30 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
RNA backbone	4643	859

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	B	895	
2	C	1217	
3	E	86	
4	G	110	
5	2	188	
6	9	501	
7	q	803	

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
8	A	1304	
9	p	1031	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
10	ZN	9	601	-	-	X	-

## 2 Entry composition

There are 11 unique types of molecules in this entry. The entry contains 20470 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 Splicing factor 3B subunit 2.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	B	110	Total	C	N	O	S	0	0
			940	613	170	154	3		

- Molecule 2 is a protein called Splicing factor 3B subunit 3.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	C	1189	Total	C	N	O	S	1	0
			9336	5925	1586	1779	46		

- Molecule 3 is a protein called Splicing factor 3B subunit 5.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	E	70	Total	C	N	O	S	0	0
			575	364	100	106	5		

- Molecule 4 is a protein called PHD finger-like domain-containing protein 5A.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	G	81	Total	C	N	O	S	0	0
			613	376	109	115	13		

- Molecule 5 is a RNA chain called U2 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	2	41	Total	C	N	O	P	0	0
			865	388	144	292	41		

- Molecule 6 is a protein called Splicing factor 3A subunit 3.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	9	92	Total	C	N	O	S	0	0
			769	493	137	135	4		

- Molecule 7 is a protein called HIV Tat-specific factor 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	q	66	Total	C	N	O	S	0	0
			537	347	86	99	5		

There are 49 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
q	-47	GLY	-	expression tag	UNP O43719
q	-46	PRO	-	expression tag	UNP O43719
q	-45	GLY	-	expression tag	UNP O43719
q	-44	SER	-	expression tag	UNP O43719
q	-43	SER	-	expression tag	UNP O43719
q	-42	GLY	-	expression tag	UNP O43719
q	-41	MET	-	expression tag	UNP O43719
q	-40	ASP	-	expression tag	UNP O43719
q	-39	GLU	-	expression tag	UNP O43719
q	-38	LYS	-	expression tag	UNP O43719
q	-37	THR	-	expression tag	UNP O43719
q	-36	THR	-	expression tag	UNP O43719
q	-35	GLY	-	expression tag	UNP O43719
q	-34	TRP	-	expression tag	UNP O43719
q	-33	ARG	-	expression tag	UNP O43719
q	-32	GLY	-	expression tag	UNP O43719
q	-31	GLY	-	expression tag	UNP O43719
q	-30	HIS	-	expression tag	UNP O43719
q	-29	VAL	-	expression tag	UNP O43719
q	-28	VAL	-	expression tag	UNP O43719
q	-27	GLU	-	expression tag	UNP O43719
q	-26	GLY	-	expression tag	UNP O43719
q	-25	LEU	-	expression tag	UNP O43719
q	-24	ALA	-	expression tag	UNP O43719
q	-23	GLY	-	expression tag	UNP O43719
q	-22	GLU	-	expression tag	UNP O43719
q	-21	LEU	-	expression tag	UNP O43719
q	-20	GLU	-	expression tag	UNP O43719
q	-19	GLN	-	expression tag	UNP O43719
q	-18	LEU	-	expression tag	UNP O43719
q	-17	ARG	-	expression tag	UNP O43719
q	-16	ALA	-	expression tag	UNP O43719
q	-15	ARG	-	expression tag	UNP O43719
q	-14	LEU	-	expression tag	UNP O43719
q	-13	GLU	-	expression tag	UNP O43719
q	-12	HIS	-	expression tag	UNP O43719

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
q	-11	HIS	-	expression tag	UNP O43719
q	-10	PRO	-	expression tag	UNP O43719
q	-9	GLN	-	expression tag	UNP O43719
q	-8	GLY	-	expression tag	UNP O43719
q	-7	GLN	-	expression tag	UNP O43719
q	-6	ARG	-	expression tag	UNP O43719
q	-5	GLU	-	expression tag	UNP O43719
q	-4	PRO	-	expression tag	UNP O43719
q	-3	GLY	-	expression tag	UNP O43719
q	-2	GLY	-	expression tag	UNP O43719
q	-1	GLY	-	expression tag	UNP O43719
q	0	GLY	-	expression tag	UNP O43719
q	1	SER	-	expression tag	UNP O43719

- Molecule 8 is a protein called Splicing factor 3B subunit 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	A	814	Total	C	N	O	S	0	0
			6479	4158	1120	1162	39		

- Molecule 9 is a protein called Probable ATP-dependent RNA helicase DDX46.

Mol	Chain	Residues	Atoms					AltConf	Trace
9	p	18	Total	C	N	O	S	0	0
			172	103	35	33	1		

- Molecule 10 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		AltConf
10	G	3	Total	Zn	0
			3	3	
10	9	1	Total	Zn	0
			1	1	

- Molecule 11 is water.

Mol	Chain	Residues	Atoms		AltConf
11	B	1	Total	O	0
			1	1	
11	C	137	Total	O	0
			137	137	

*Continued on next page...*

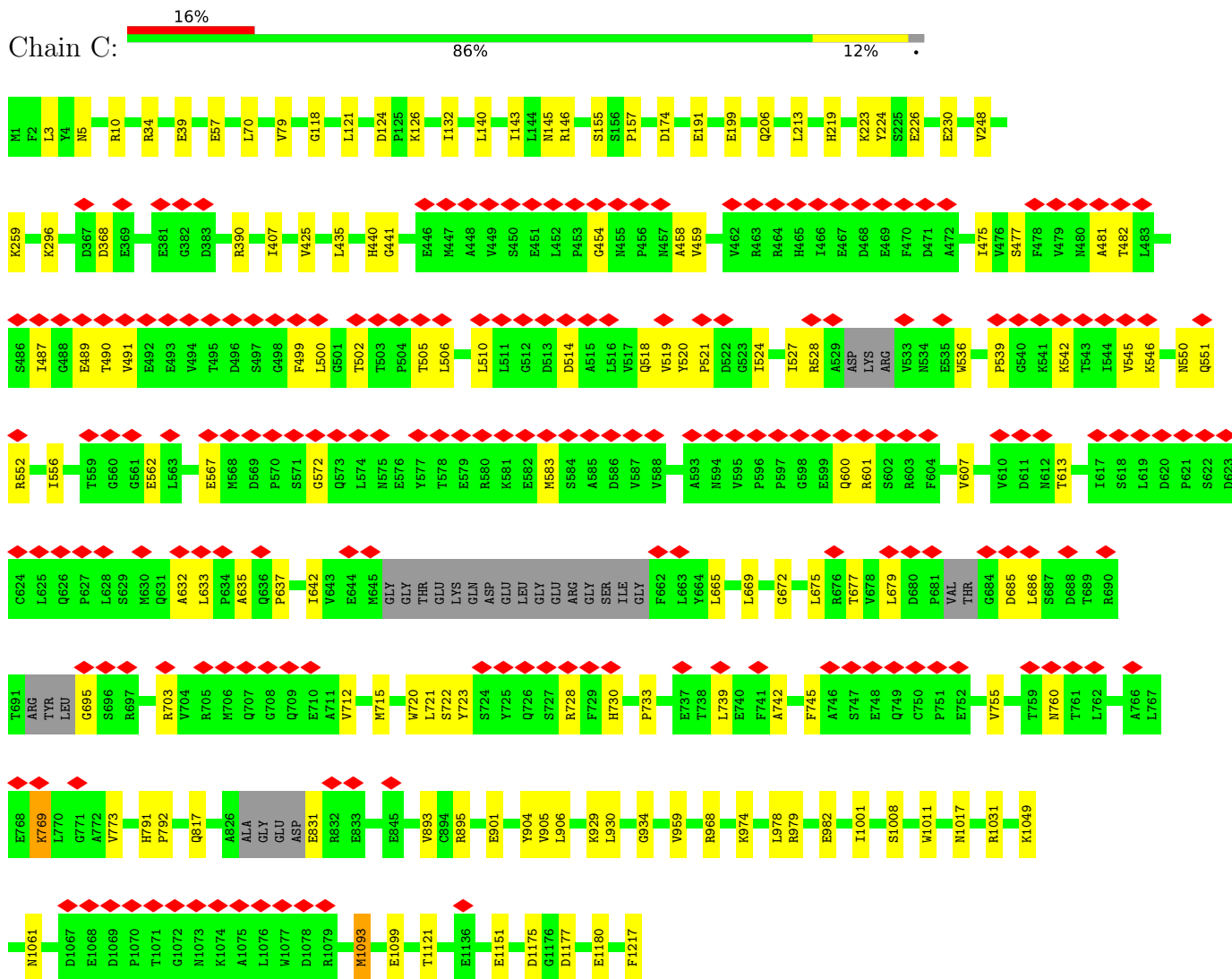
*Continued from previous page...*

Mol	Chain	Residues	Atoms		AltConf
11	E	24	Total 24	O 24	0
11	G	3	Total 3	O 3	0
11	A	15	Total 15	O 15	0

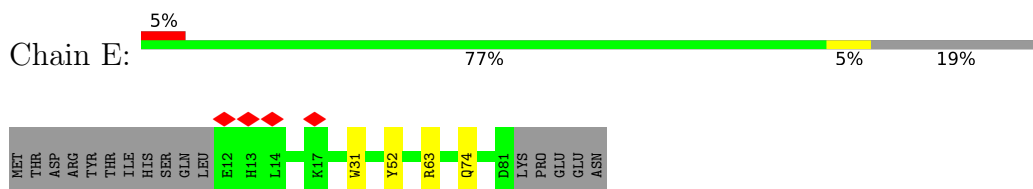




- Molecule 2: Splicing factor 3B subunit 3



- Molecule 3: Splicing factor 3B subunit 5



- Molecule 4: PHD finger-like domain-containing protein 5A









[illegible]

## 4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	225934	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 ( $e^-/\text{\AA}^2$ )	53.45	Depositor
Minimum defocus (nm)	800	Depositor
Maximum defocus (nm)	1800	Depositor
Magnification	130000	Depositor
Image detector	GATAN K3 BIOQUANTUM (6k x 4k)	Depositor
Maximum map value	0.048	Depositor
Minimum map value	-0.016	Depositor
Average map value	-0.000	Depositor
Map value standard deviation	0.001	Depositor
Recommended contour level	0.0062	Depositor
Map size (Å)	409.59998, 409.59998, 409.59998	wwPDB
Map dimensions	640, 640, 640	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	0.64, 0.64, 0.64	Depositor

## 5 Model quality

### 5.1 Standard geometry

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

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	0.38	0/965	0.85	0/1296
2	C	0.33	0/9527	0.69	0/12925
3	E	0.36	0/592	0.85	1/800 (0.1%)
4	G	0.54	0/621	0.88	2/833 (0.2%)
5	2	0.83	0/964	1.16	16/1495 (1.1%)
6	9	0.40	0/793	0.88	1/1076 (0.1%)
7	q	0.38	0/542	0.77	0/724
8	A	0.33	0/6601	0.76	3/8934 (0.0%)
9	p	0.42	0/173	1.03	0/225
All	All	0.38	0/20778	0.77	23/28308 (0.1%)

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
2	C	0	2
3	E	0	1
4	G	0	1
6	9	0	1
8	A	0	1
All	All	0	6

There are no bond length outliers.

The worst 5 of 23 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	G	82	ARG	NE-CZ-NH2	-8.39	116.10	120.30
5	2	52	G	P-O3'-C3'	-6.38	112.04	119.70

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	9	479	TYR	CB-CA-C	6.36	123.11	110.40
5	2	31	G	P-O3'-C3'	-6.33	112.10	119.70
5	2	47	U	P-O3'-C3'	-6.27	112.18	119.70

There are no chirality outliers.

5 of 6 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
6	9	479	TYR	Sidechain
2	C	5	ASN	Mainchain
2	C	817	GLN	Mainchain
3	E	31	TRP	Mainchain
4	G	82	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	940	0	955	2	0
2	C	9336	0	9252	80	0
3	E	575	0	537	1	0
4	G	613	0	595	2	0
5	2	865	0	436	1	0
6	9	769	0	726	6	0
7	q	537	0	545	0	0
8	A	6479	0	6694	23	0
9	p	172	0	168	0	0
10	9	1	0	0	2	0
10	G	3	0	0	0	0
11	A	15	0	0	0	0
11	B	1	0	0	0	0
11	C	137	0	0	0	0
11	E	24	0	0	0	0
11	G	3	0	0	0	0
All	All	20470	0	19908	109	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 3.

The worst 5 of 109 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)
2:C:728:ARG:NH2	2:C:1180:GLU:OE1	1.88	1.06
6:9:408:CYS:HG	10:9:601:ZN:ZN	0.72	1.01
2:C:145:ASN:OD1	8:A:677:CYS:SG	2.49	0.70
2:C:930:LEU:HG	2:C:934:GLY:HA2	1.74	0.70
2:C:454:GLY:O	2:C:760:ASN:ND2	2.27	0.67

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	B	102/895 (11%)	99 (97%)	3 (3%)	0	100	100
2	C	1177/1217 (97%)	1124 (96%)	53 (4%)	0	100	100
3	E	68/86 (79%)	68 (100%)	0	0	100	100
4	G	79/110 (72%)	75 (95%)	4 (5%)	0	100	100
6	9	88/501 (18%)	84 (96%)	4 (4%)	0	100	100
7	q	56/803 (7%)	56 (100%)	0	0	100	100
8	A	812/1304 (62%)	799 (98%)	13 (2%)	0	100	100
9	p	16/1031 (2%)	16 (100%)	0	0	100	100
All	All	2398/5947 (40%)	2321 (97%)	77 (3%)	0	100	100

There are no Ramachandran outliers to report.

### 5.3.2 Protein sidechains ⓘ

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	B	101/776 (13%)	97 (96%)	4 (4%)	31	44
2	C	1031/1051 (98%)	1012 (98%)	19 (2%)	59	75
3	E	61/77 (79%)	60 (98%)	1 (2%)	62	78
4	G	69/95 (73%)	65 (94%)	4 (6%)	20	27
6	9	79/446 (18%)	77 (98%)	2 (2%)	47	65
7	q	62/694 (9%)	62 (100%)	0	100	100
8	A	700/1104 (63%)	681 (97%)	19 (3%)	44	61
9	p	18/892 (2%)	17 (94%)	1 (6%)	21	29
All	All	2121/5135 (41%)	2071 (98%)	50 (2%)	51	66

5 of 50 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
6	9	442	ASN
8	A	885	ASP
9	p	171	ARG
6	9	461	LYS
8	A	739	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
3	E	35	GLN
4	G	50	ASN
8	A	550	HIS
6	9	488	ASN
2	C	775	ASN

### 5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
5	2	39/188 (20%)	6 (15%)	0

5 of 6 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
5	2	28	C
5	2	38	A
5	2	39	U
5	2	41	U
5	2	48	A

There are no RNA pucker outliers to report.

## 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](#)

Of 4 ligands modelled in this entry, 4 are 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 ⓘ

There are no chain breaks in this entry.

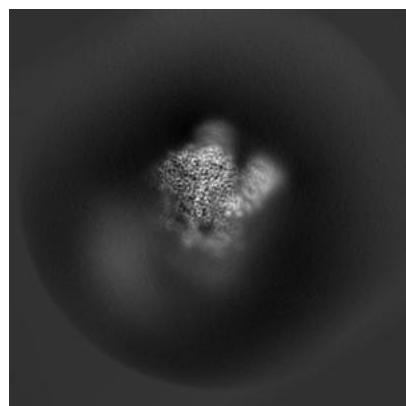
## 6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-13793. These allow visual inspection of the internal detail of the map and identification of artifacts.

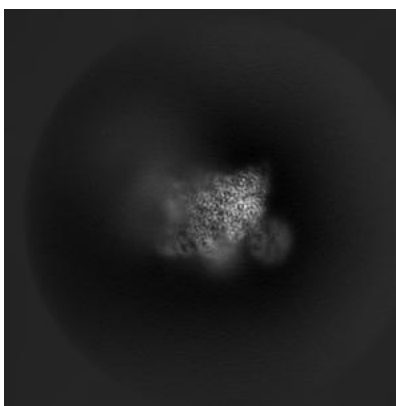
Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

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

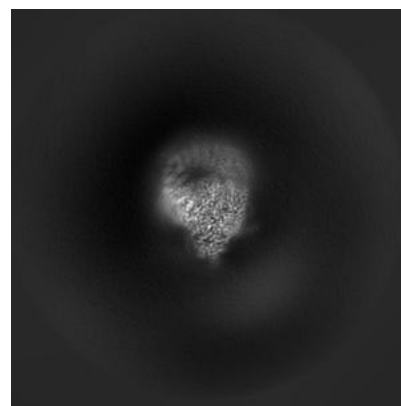
#### 6.1.1 Primary map



X

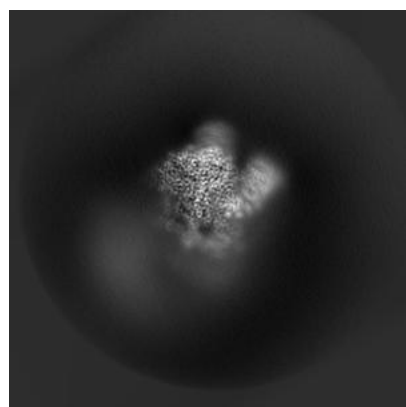


Y

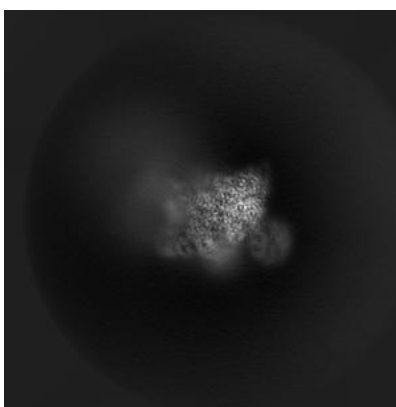


Z

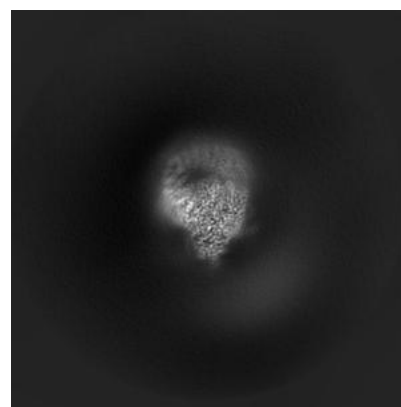
#### 6.1.2 Raw map



X



Y

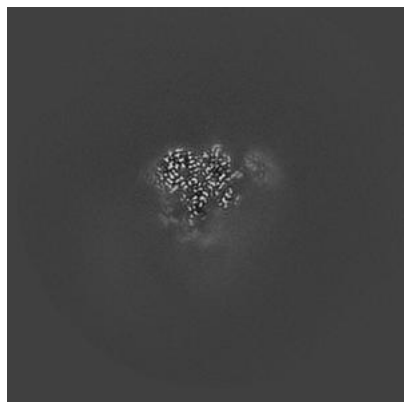


Z

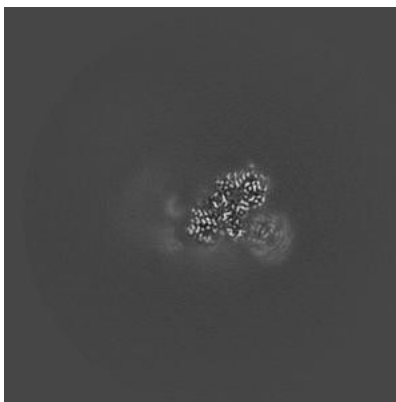
The images above show the map projected in three orthogonal directions.

## 6.2 Central slices [i](#)

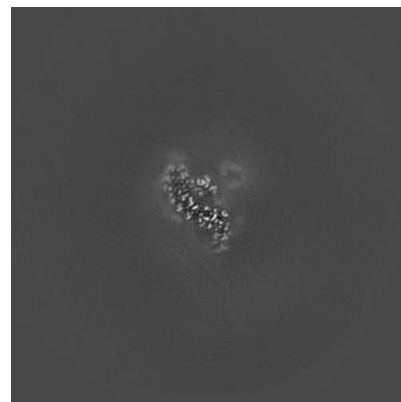
### 6.2.1 Primary map



X Index: 320

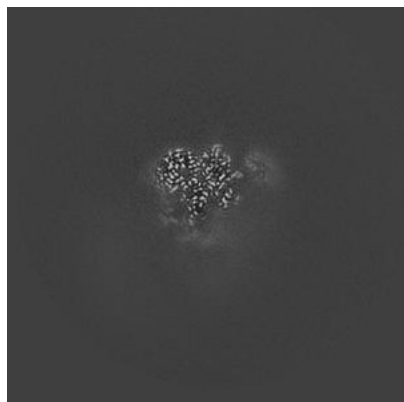


Y Index: 320

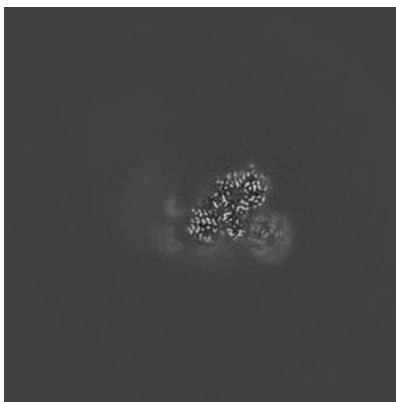


Z Index: 320

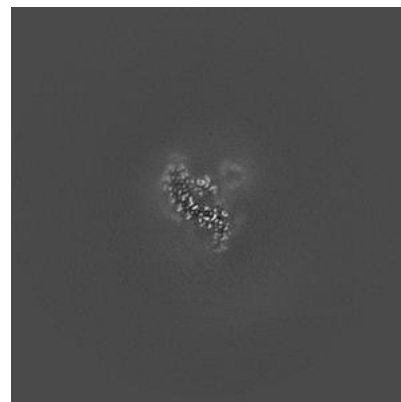
### 6.2.2 Raw map



X Index: 320



Y Index: 320

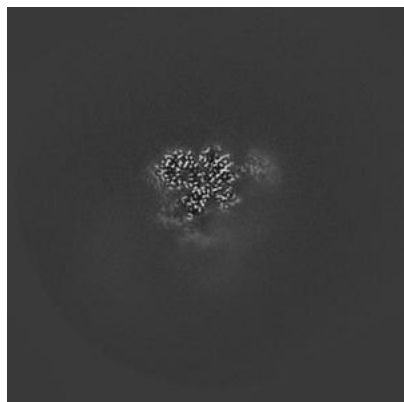


Z Index: 320

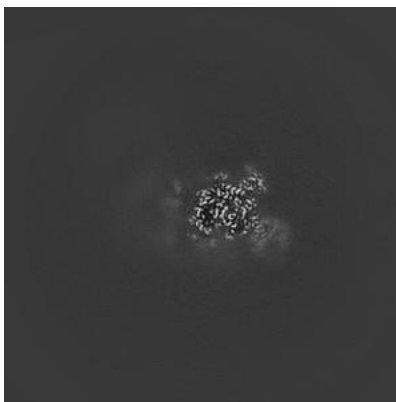
The images above show central slices of the map in three orthogonal directions.

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

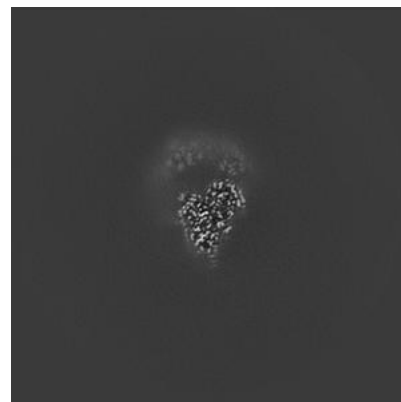
### 6.3.1 Primary map



X Index: 324

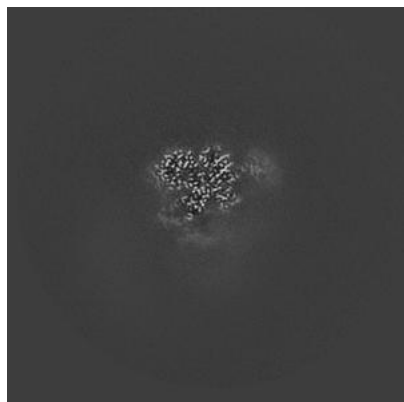


Y Index: 307

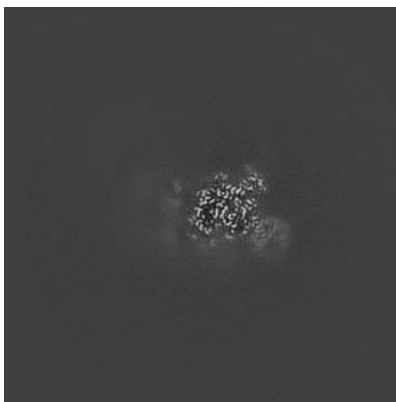


Z Index: 361

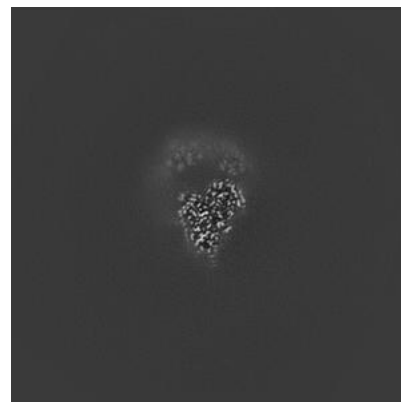
### 6.3.2 Raw map



X Index: 324



Y Index: 307

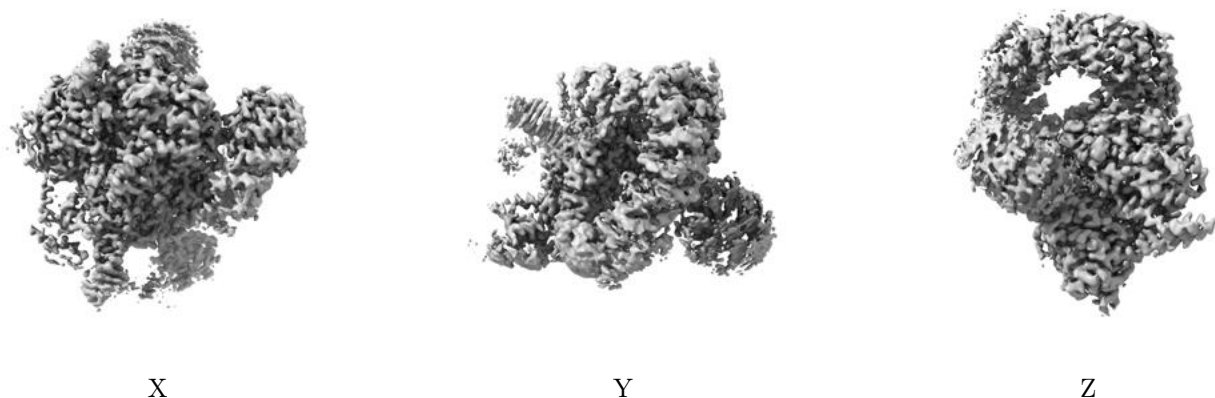


Z Index: 361

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

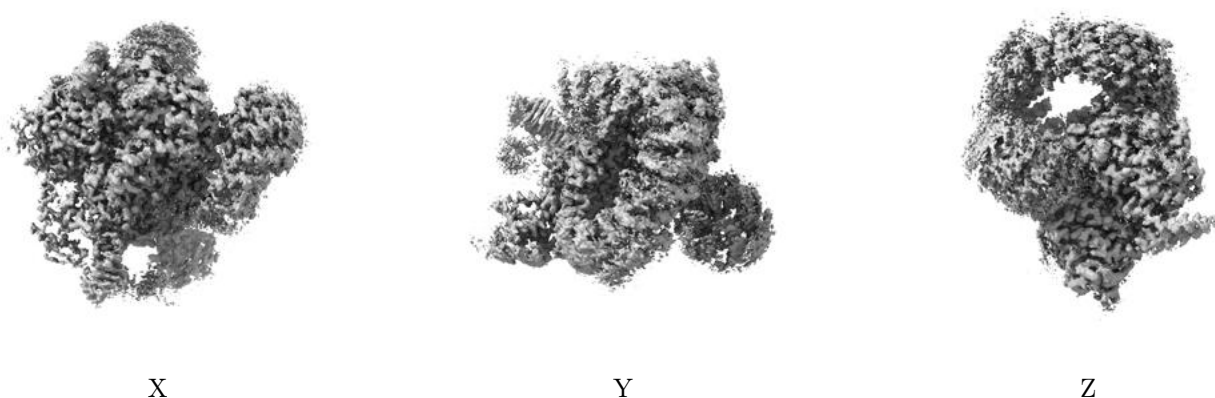
## 6.4 Orthogonal surface views [i](#)

### 6.4.1 Primary map



The images above show the 3D surface view of the map at the recommended contour level 0.0062. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

### 6.4.2 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.



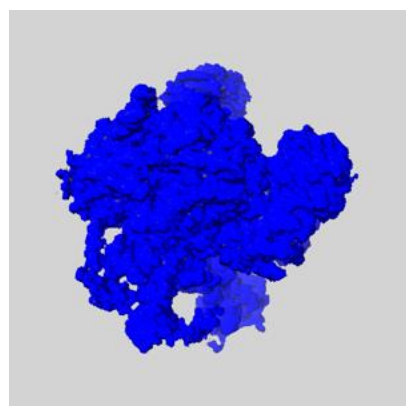
## 6.5 Mask visualisation [i](#)

This section shows the 3D surface view of the primary map at 50% transparency overlaid with the specified mask at 0% transparency

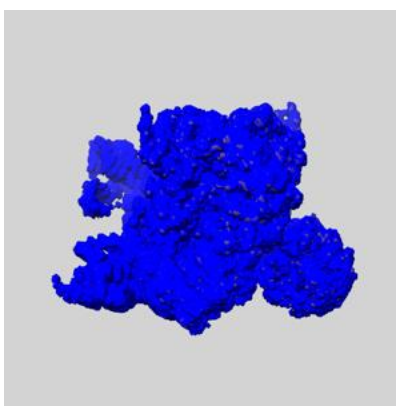
A mask typically either:

- Encompasses the whole structure
- Separates out a domain, a functional unit, a monomer or an area of interest from a larger structure

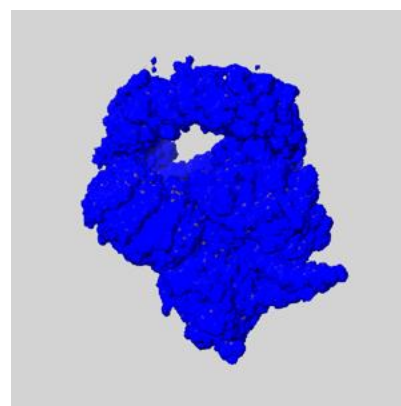
### 6.5.1 emd\_13793\_msk\_1.map [i](#)



X



Y

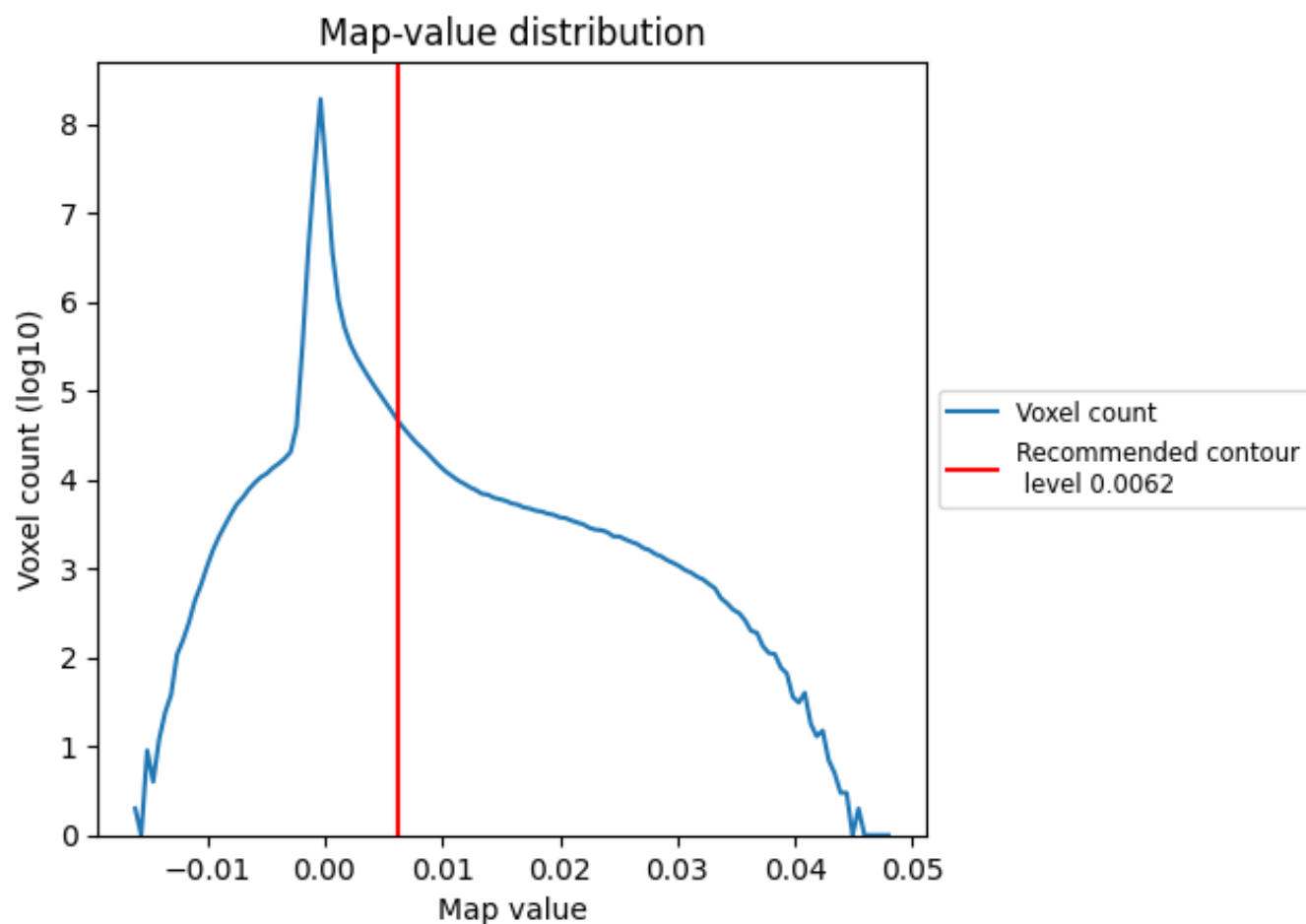


Z

## 7 Map analysis [i](#)

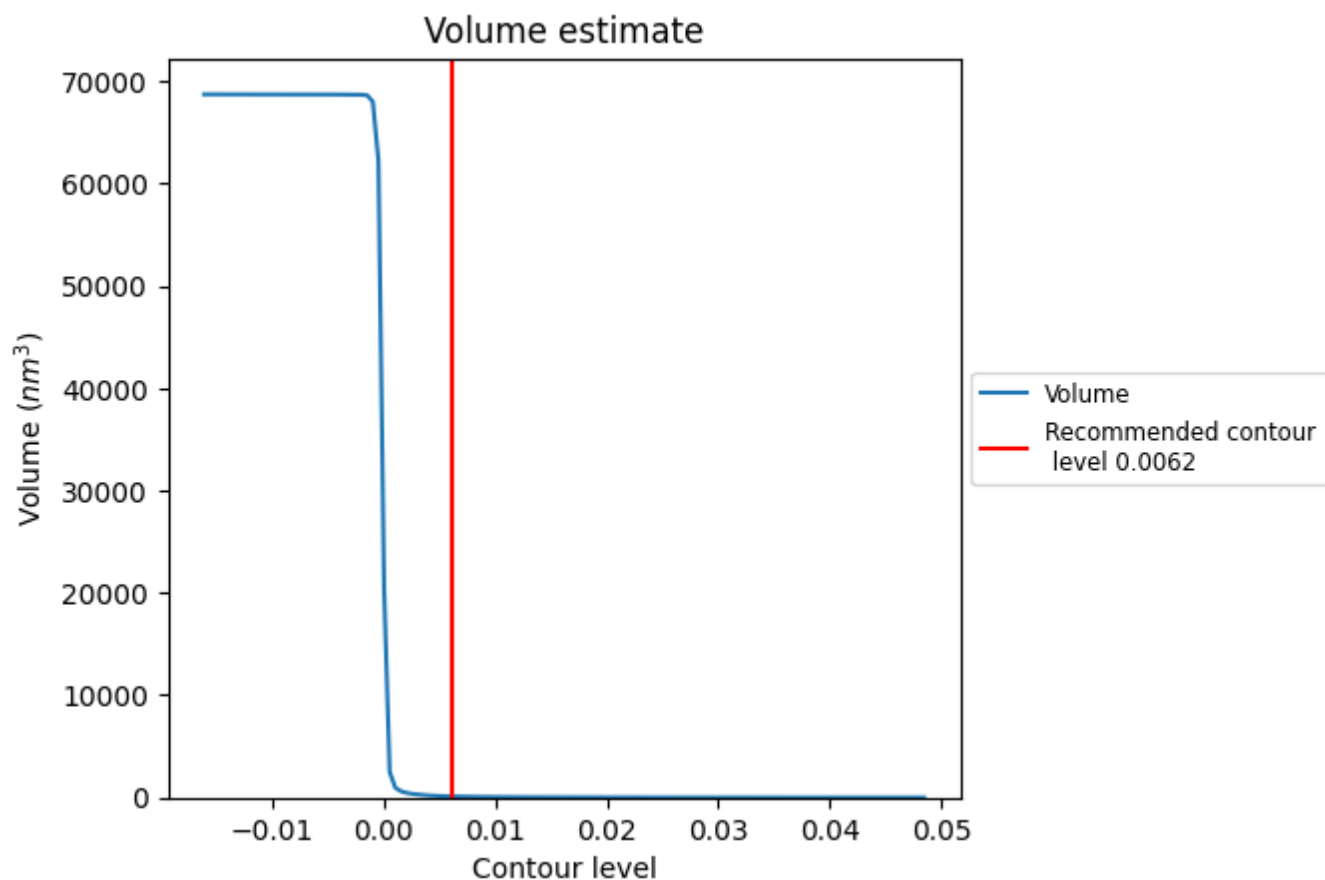
This section contains the results of statistical analysis of the map.

### 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.

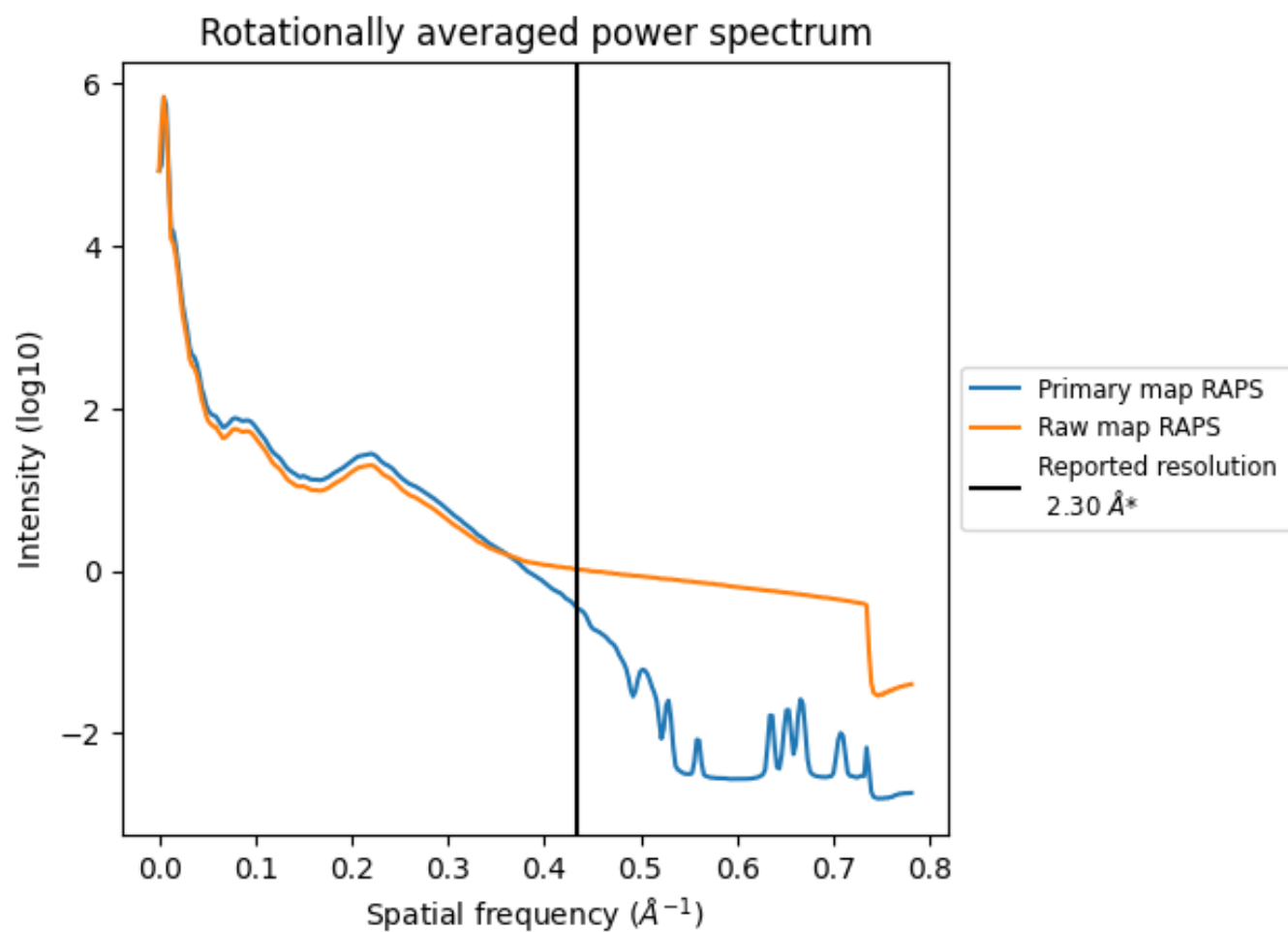
## 7.2 Volume estimate [i](#)



The volume at the recommended contour level is 106  $\text{nm}^3$ ; this corresponds to an approximate mass of 95 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](#)

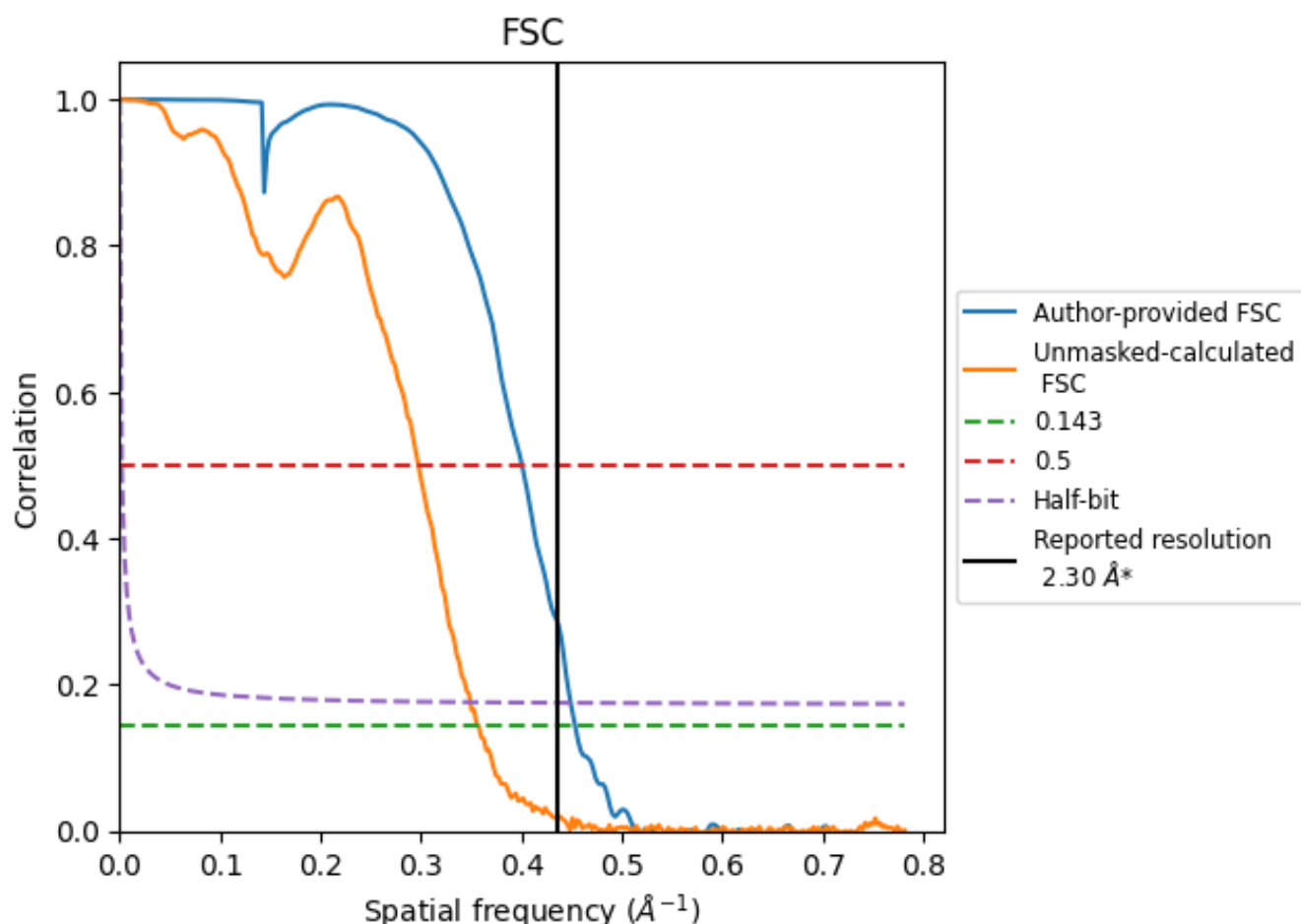


\*Reported resolution corresponds to spatial frequency of  $0.435 \text{ \AA}^{-1}$

## 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 0.435 Å<sup>-1</sup>

## 8.2 Resolution estimates [i](#)

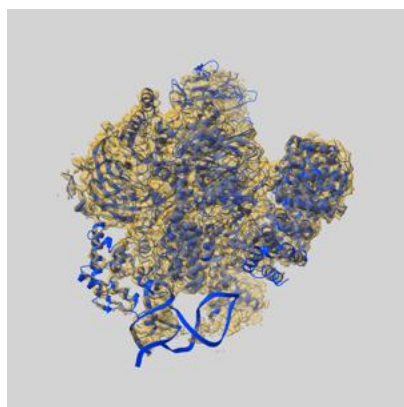
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	2.30	-	-
Author-provided FSC curve	2.21	2.50	2.23
Unmasked-calculated*	2.80	3.36	2.87

\*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 2.80 differs from the reported value 2.3 by more than 10 %

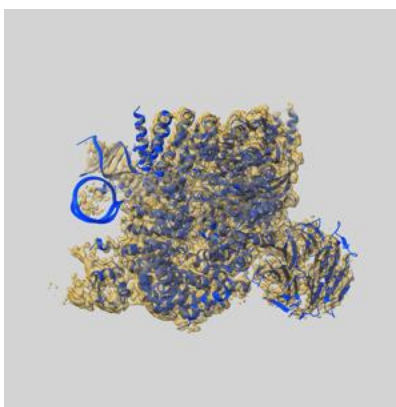
## 9 Map-model fit [i](#)

This section contains information regarding the fit between EMDB map EMD-13793 and PDB model 7Q3L. Per-residue inclusion information can be found in section [3](#) on page [8](#).

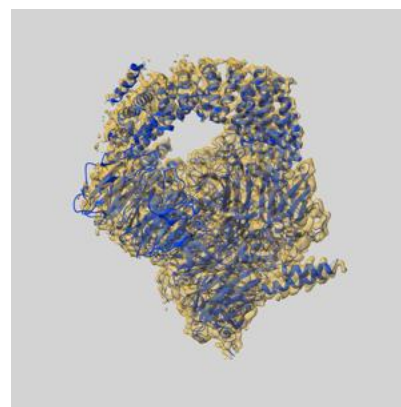
### 9.1 Map-model overlay [i](#)



X



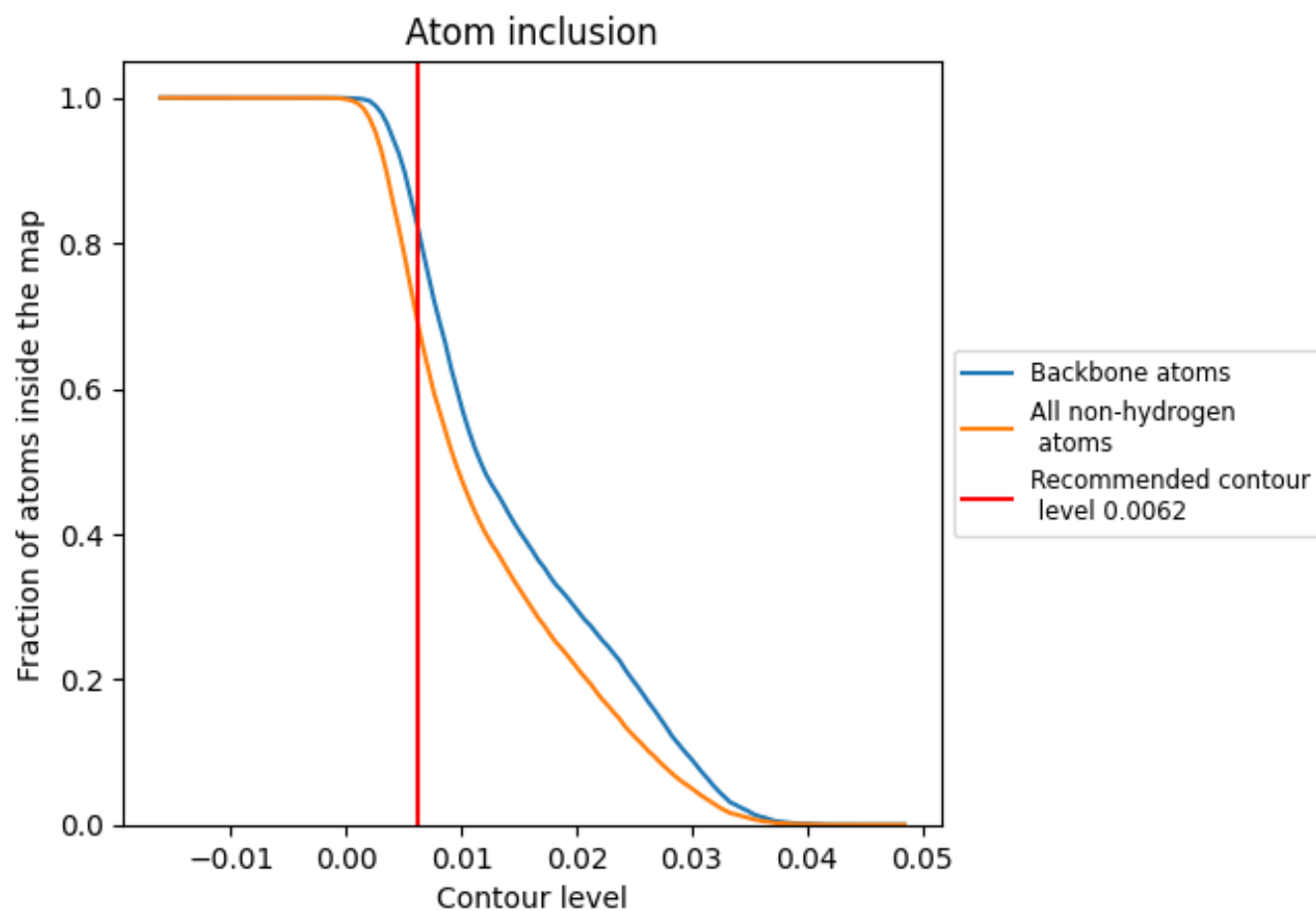
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.0062 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 Atom inclusion [i](#)



At the recommended contour level, 82% of all backbone atoms, 69% of all non-hydrogen atoms, are inside the map.