



## wwPDB EM Validation Summary Report ⓘ

Nov 6, 2022 – 07:56 AM EST

PDB ID : 6DNH  
EMDB ID : EMD-7112  
Title : Cryo-EM structure of human CPSF-160-WDR33-CPSF-30-PAS RNA complex at 3.4 Å resolution  
Authors : Sun, Y.; Zhang, Y.; Hamilton, K.; Walz, T.; Tong, L.  
Deposited on : 2018-06-06  
Resolution : 3.40 Å (reported)  
Based on initial model : 2RHK

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.dev43  
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) : 2.31.2

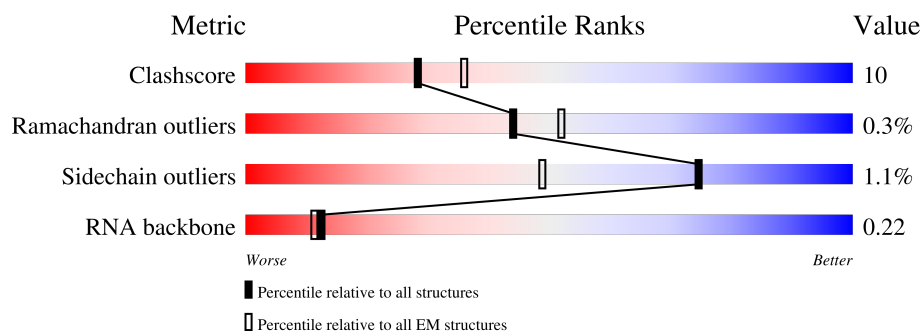
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 3.40 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	A	1443	
2	B	587	
3	C	245	
4	E	17	

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 13442 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 Cleavage and polyadenylation specificity factor subunit 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	1177	Total	C	N	O	S	0	0
			9337	6003	1600	1679	55		

- Molecule 2 is a protein called pre-mRNA 3' end processing protein WDR33.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	376	Total	C	N	O	S	0	0
			3035	1916	557	542	20		

There are 15 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-14	MET	-	expression tag	UNP Q9C0J8
B	-13	GLY	-	expression tag	UNP Q9C0J8
B	-12	SER	-	expression tag	UNP Q9C0J8
B	-11	SER	-	expression tag	UNP Q9C0J8
B	-10	HIS	-	expression tag	UNP Q9C0J8
B	-9	HIS	-	expression tag	UNP Q9C0J8
B	-8	HIS	-	expression tag	UNP Q9C0J8
B	-7	HIS	-	expression tag	UNP Q9C0J8
B	-6	HIS	-	expression tag	UNP Q9C0J8
B	-5	HIS	-	expression tag	UNP Q9C0J8
B	-4	SER	-	expression tag	UNP Q9C0J8
B	-3	SER	-	expression tag	UNP Q9C0J8
B	-2	GLY	-	expression tag	UNP Q9C0J8
B	-1	LEU	-	expression tag	UNP Q9C0J8
B	0	VAL	-	expression tag	UNP Q9C0J8

- Molecule 3 is a protein called Cleavage and polyadenylation specificity factor subunit 4.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	C	117	Total	C	N	O	S	0	0
			917	587	151	165	14		

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	0	GLY	-	expression tag	UNP O95639

- Molecule 4 is a RNA chain called RNA (5'-R(P\*AP\*AP\*UP\*AP\*AP\*AP\*C)-3').

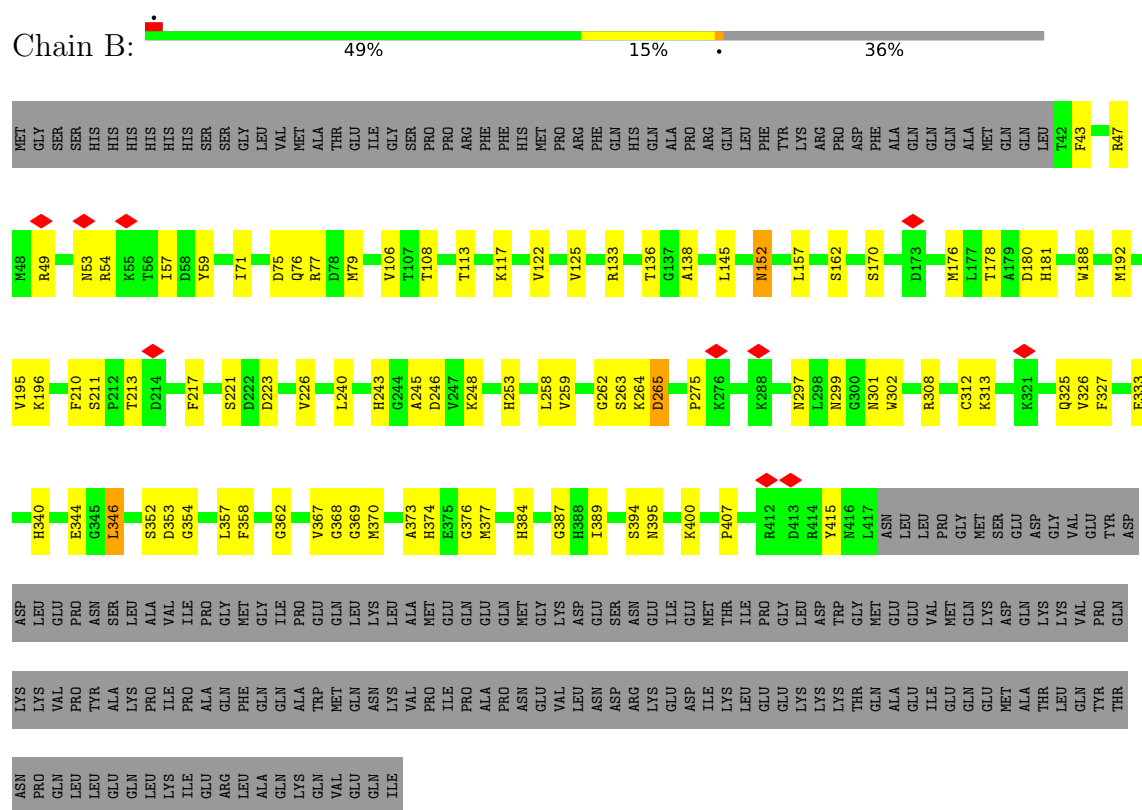
Mol	Chain	Residues	Atoms					AltConf	Trace
4	E	7	Total	C	N	O	P	0	0
			150	68	30	45	7		

- Molecule 5 is ZINC ION (three-letter code: ZN) (formula: Zn).

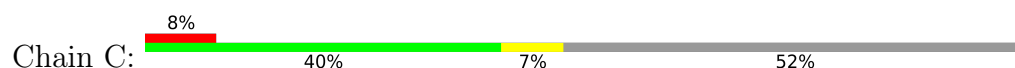
Mol	Chain	Residues	Atoms		AltConf
5	C	3	Total	Zn	0
			3	3	



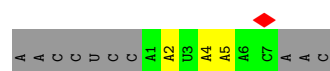
- Molecule 2: pre-mRNA 3' end processing protein WDR33



- Molecule 3: Cleavage and polyadenylation specificity factor subunit 4



- Molecule 4: RNA (5'-R(P\*AP\*AP\*UP\*AP\*AP\*AP\*C)-3')



## 4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	173632	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	70	Depositor
Minimum defocus (nm)	900	Depositor
Maximum defocus (nm)	2800	Depositor
Magnification	46729	Depositor
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor
Maximum map value	0.315	Depositor
Minimum map value	-0.177	Depositor
Average map value	0.000	Depositor
Map value standard deviation	0.014	Depositor
Recommended contour level	0.06	Depositor
Map size (Å)	205.44, 205.44, 205.44	wwPDB
Map dimensions	192, 192, 192	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.07, 1.07, 1.07	Depositor



## 5 Model quality [i](#)

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

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	A	0.47	0/9553	0.62	0/12967
2	B	0.47	0/3120	0.62	0/4224
3	C	0.40	0/941	0.57	0/1261
4	E	0.46	0/168	0.79	0/259
All	All	0.47	0/13782	0.62	0/18711

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	9337	0	9390	187	0
2	B	3035	0	2949	82	0
3	C	917	0	889	18	0
4	E	150	0	77	2	0
5	C	3	0	0	0	0
All	All	13442	0	13305	280	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 10.

The worst 5 of 280 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:B:245:ALA:HB2	2:B:265:ASP:HB2	1.27	1.16
1:A:653:ASP:HB3	1:A:654:PRO:CD	1.76	1.15
1:A:653:ASP:HB3	1:A:654:PRO:HD3	1.24	1.13
2:B:245:ALA:HB2	2:B:265:ASP:CB	1.84	1.07
1:A:1273:ASP:O	1:A:1275:ASP:N	1.88	1.05

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	1151/1443 (80%)	1056 (92%)	92 (8%)	3 (0%)	41	72
2	B	374/587 (64%)	342 (91%)	31 (8%)	1 (0%)	41	72
3	C	115/245 (47%)	98 (85%)	16 (14%)	1 (1%)	17	49
All	All	1640/2275 (72%)	1496 (91%)	139 (8%)	5 (0%)	44	72

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	653	ASP
1	A	1274	ARG
3	C	2	GLN
1	A	201	LEU
2	B	152	ASN

### 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	1030/1235 (83%)	1017 (99%)	13 (1%)	69	84
2	B	326/514 (63%)	323 (99%)	3 (1%)	78	90
3	C	99/211 (47%)	99 (100%)	0	100	100
All	All	1455/1960 (74%)	1439 (99%)	16 (1%)	74	86

5 of 16 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	265	ASP
2	B	49	ARG
1	A	1235	ARG
1	A	1317	ARG
1	A	1233	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	343	HIS
2	B	416	ASN
1	A	1177	HIS
1	A	1332	ASN
2	B	152	ASN

### 5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
4	E	6/17 (35%)	2 (33%)	0

All (2) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
4	E	4	A
4	E	5	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 3 ligands modelled in this entry, 3 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 [i](#)

There are no chain breaks in this entry.

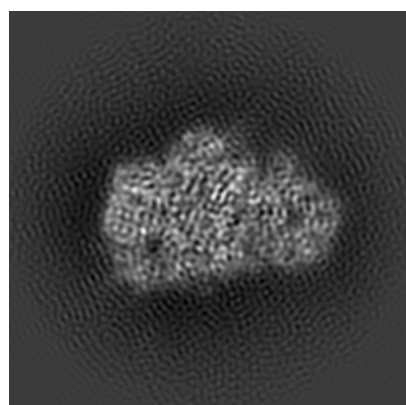
## 6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-7112. These allow visual inspection of the internal detail of the map and identification of artifacts.

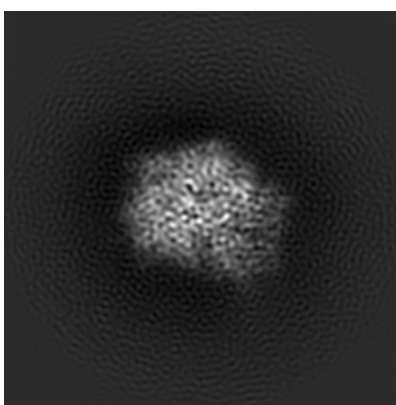
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

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

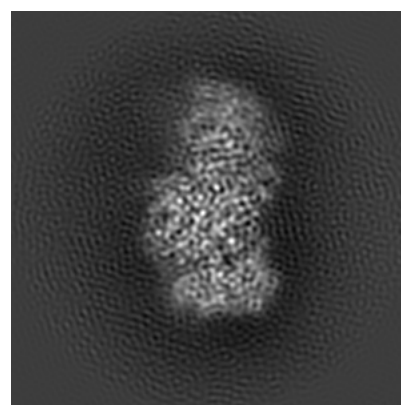
#### 6.1.1 Primary map



X



Y

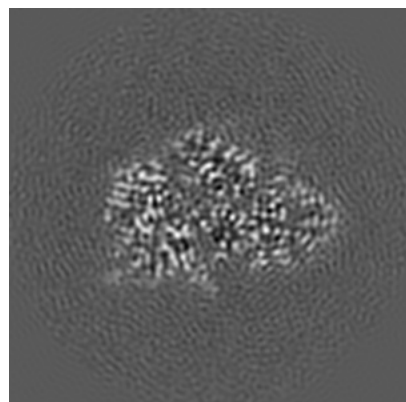


Z

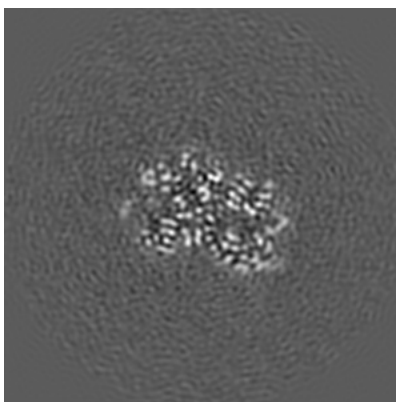
The images above show the map projected in three orthogonal directions.

### 6.2 Central slices [i](#)

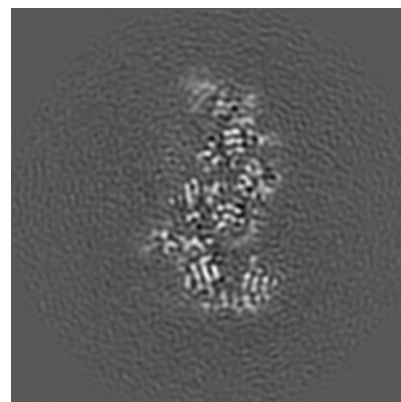
#### 6.2.1 Primary map



X Index: 96



Y Index: 96

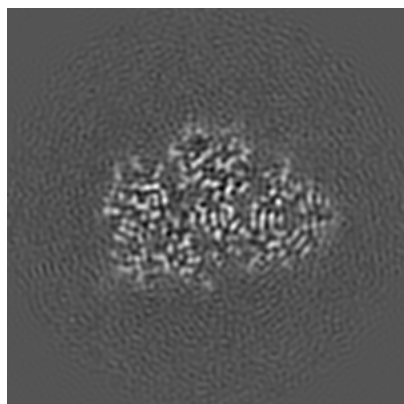


Z Index: 96

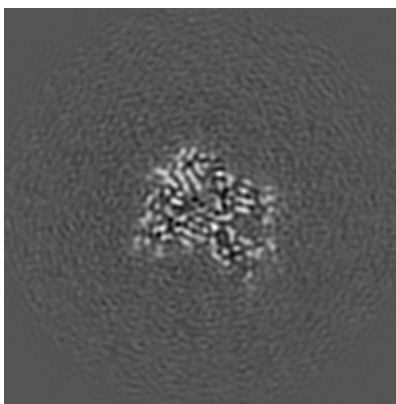
The images above show central slices of the map in three orthogonal directions.

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

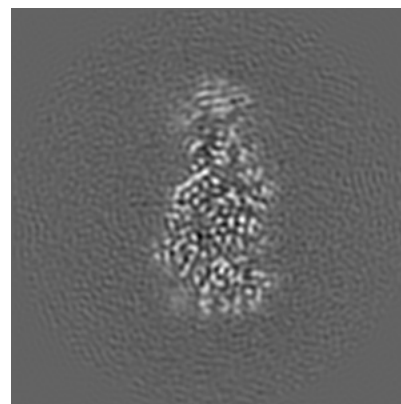
### 6.3.1 Primary map



X Index: 98



Y Index: 101

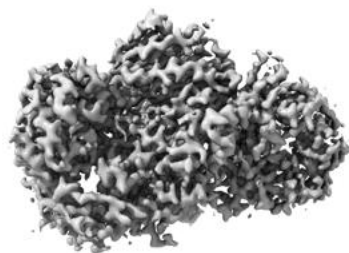


Z Index: 87

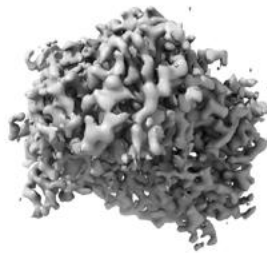
The images above show the largest variance slices of the map in three orthogonal directions.

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

### 6.4.1 Primary map



X



Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.06. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

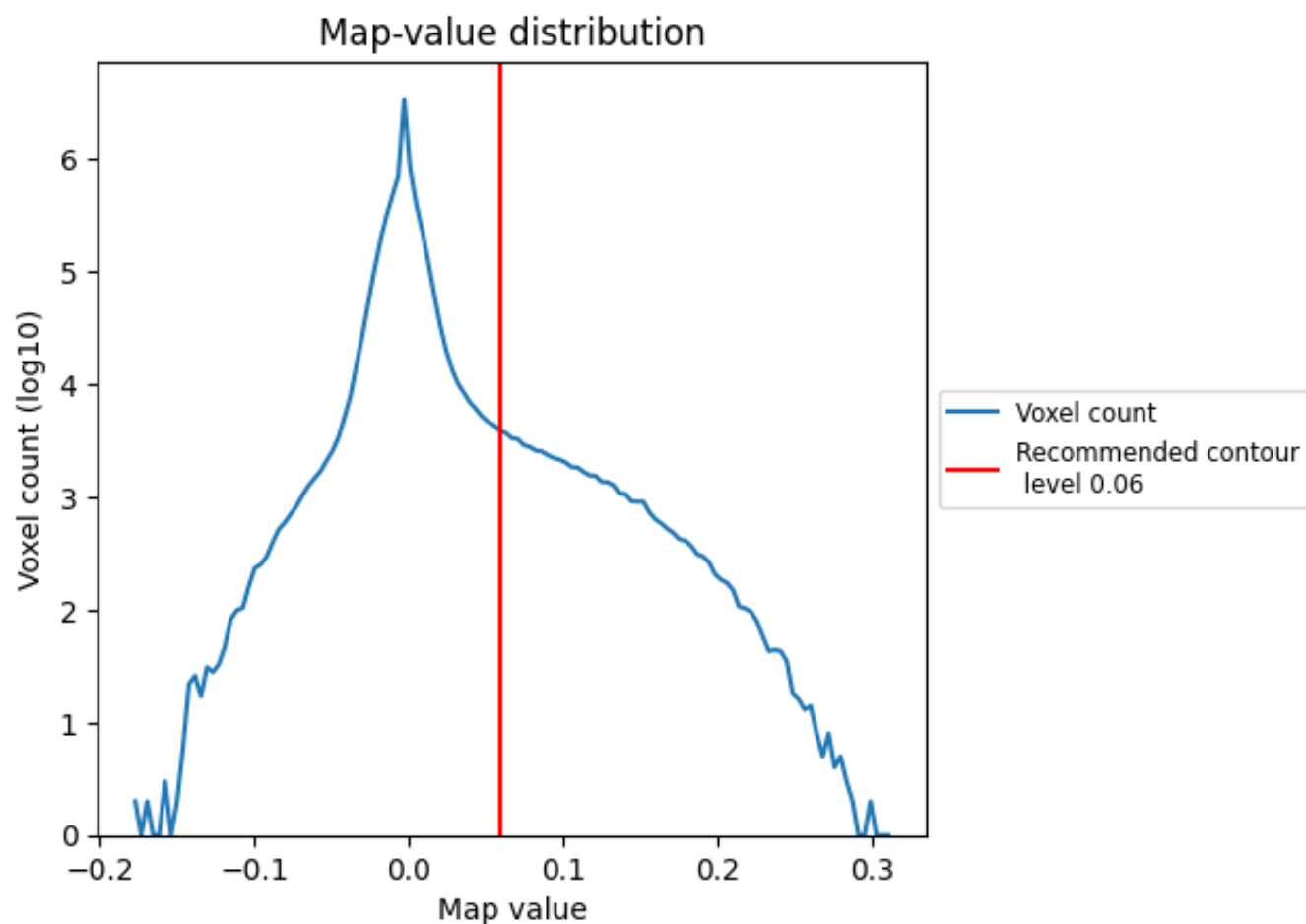
## 6.5 Mask visualisation

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.

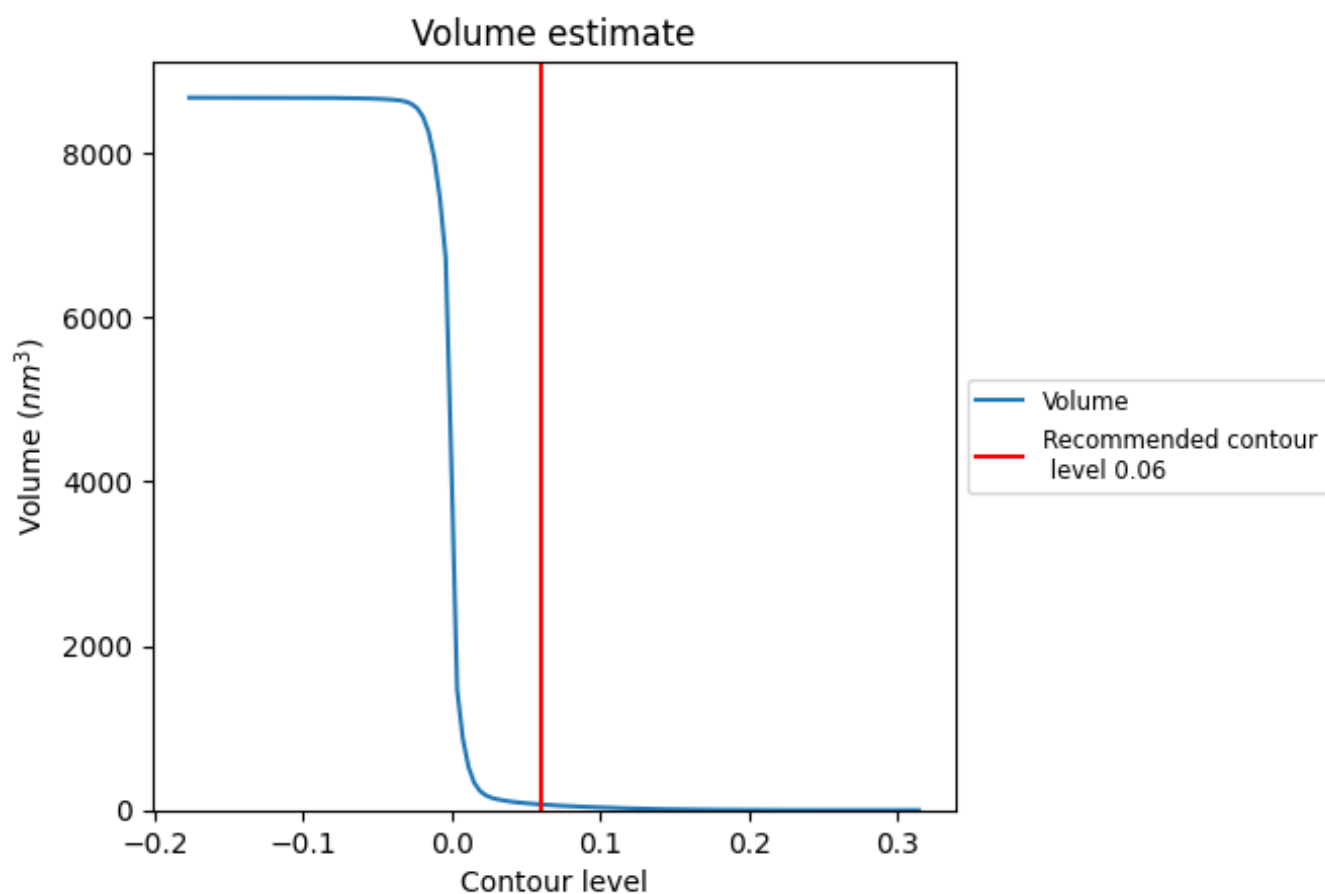
### 7.1 Map-value distribution [i](#)



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.



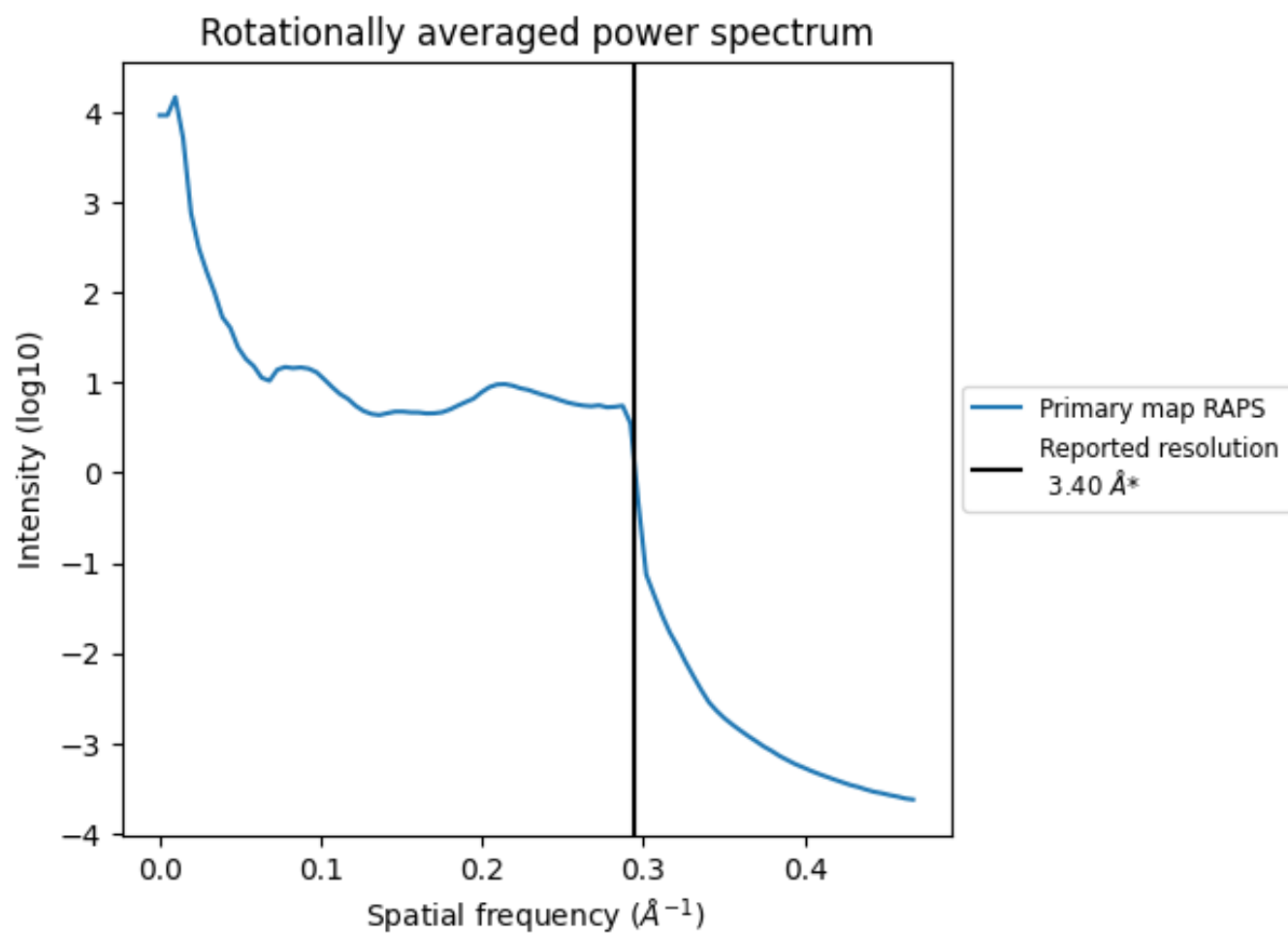
## 7.2 Volume estimate [i](#)



The volume at the recommended contour level is 70 nm<sup>3</sup>; this corresponds to an approximate mass of 63 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 0.294 Å<sup>-1</sup>

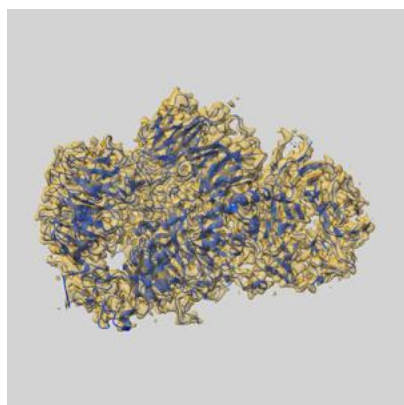
## 8 Fourier-Shell correlation

This section was not generated. No FSC curve or half-maps provided.

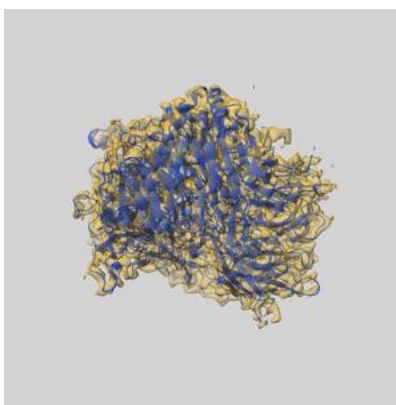
## 9 Map-model fit [i](#)

This section contains information regarding the fit between EMDB map EMD-7112 and PDB model 6DNH. Per-residue inclusion information can be found in [section 3](#) on [page 5](#).

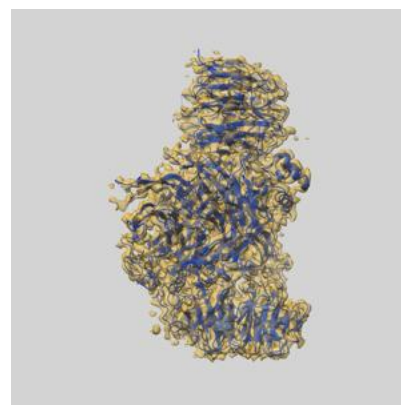
### 9.1 Map-model overlay [i](#)



X



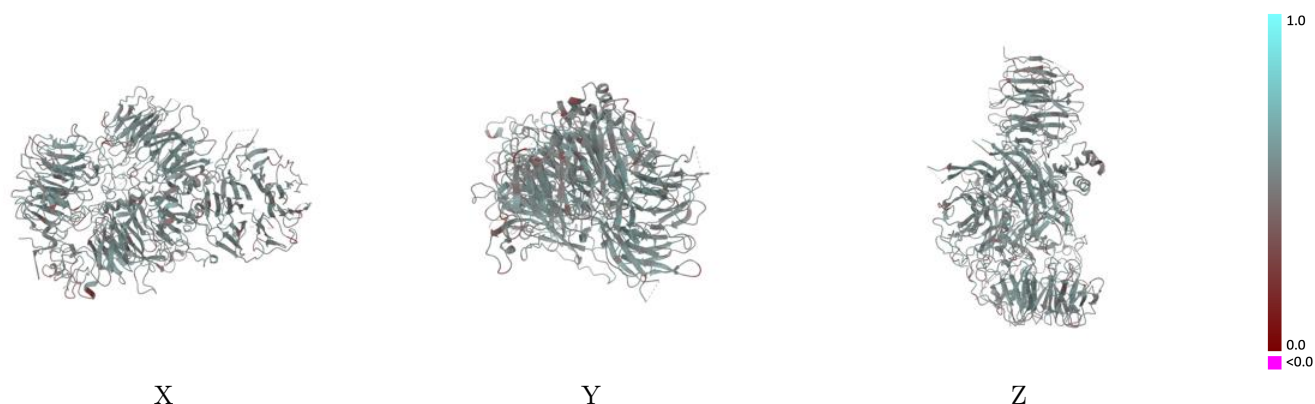
Y



Z

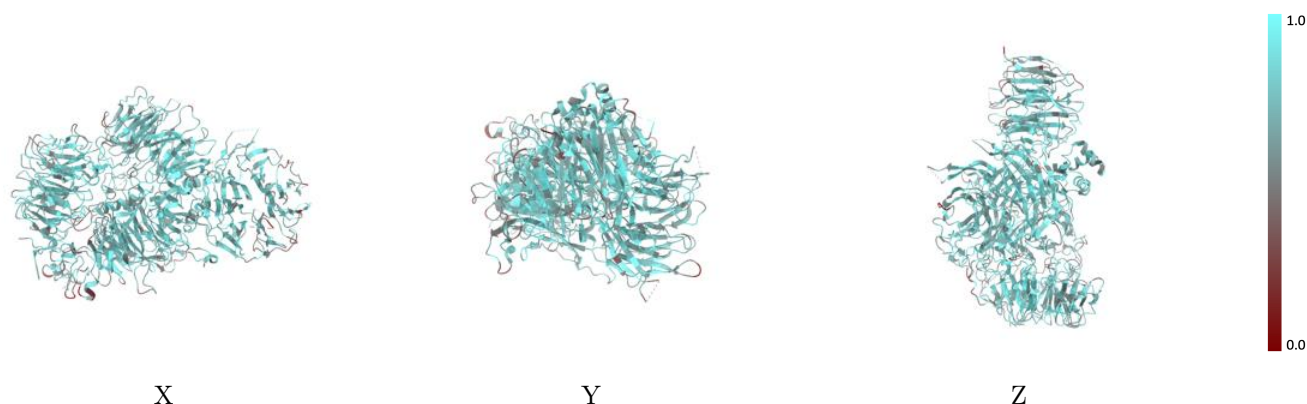
The images above show the 3D surface view of the map at the recommended contour level 0.06 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

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



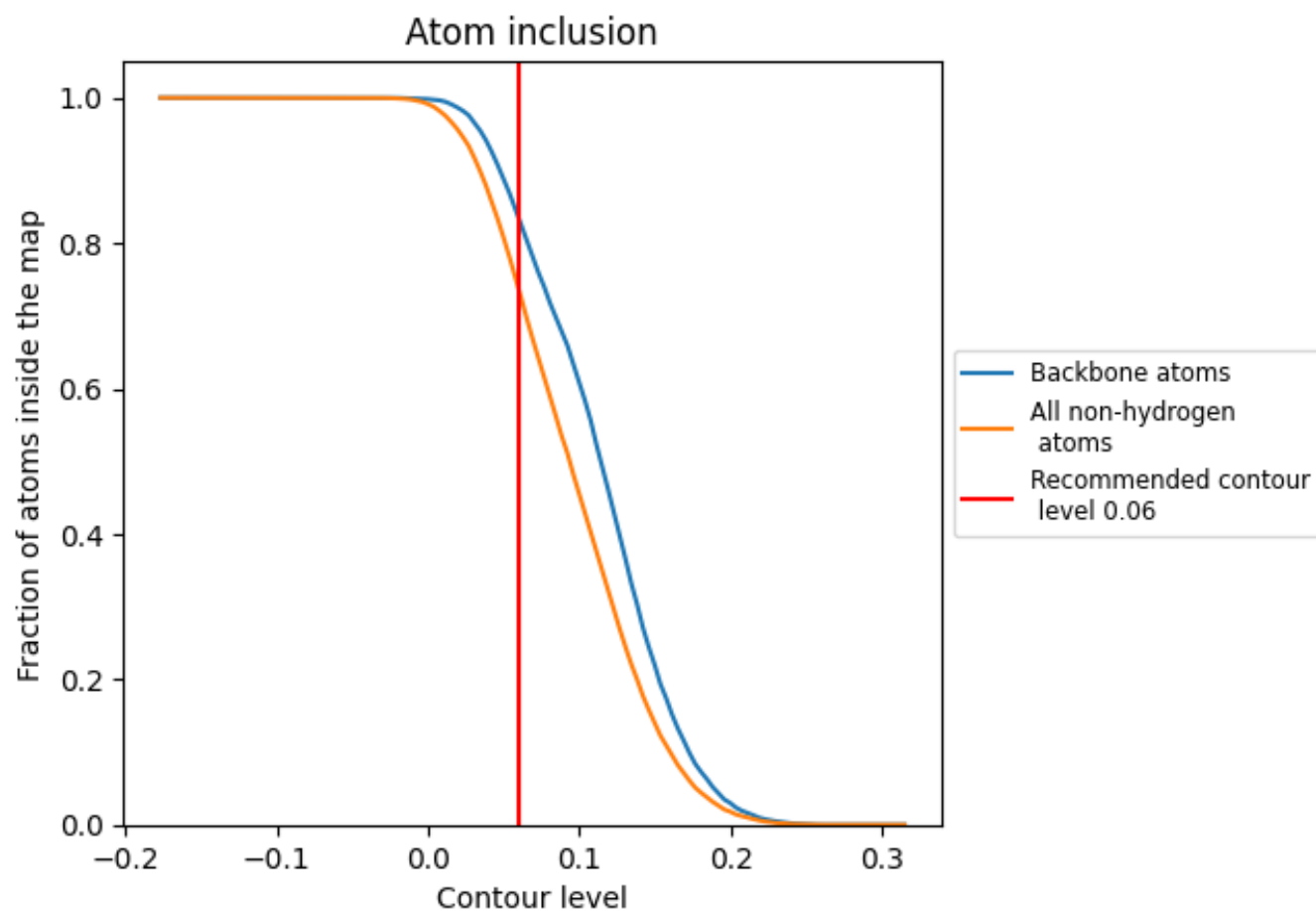
The images above show the model with each residue coloured according to its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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 (0.06).

## 9.4 Atom inclusion [i](#)



At the recommended contour level, 83% of all backbone atoms, 74% 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 (0.06) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	<div></div> 0.7364	<div></div> 0.5120
A	<div></div> 0.7447	<div></div> 0.5140
B	<div></div> 0.7454	<div></div> 0.5160
C	<div></div> 0.6228	<div></div> 0.4760
E	<div></div> 0.7467	<div></div> 0.5020

