



wwPDB EM Validation Summary Report ⓘ

Nov 6, 2022 – 08:21 AM EST

PDB ID : 6BM0
EMDB ID : EMD-7114
Title : Cryo-EM structure of human CPSF-160-WDR33 complex at 3.8 Å resolution
Authors : Sun, Y.; Zhang, Y.; Hamilton, K.; Walz, T.; Tong, L.
Deposited on : 2017-11-12
Resolution : 3.80 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

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

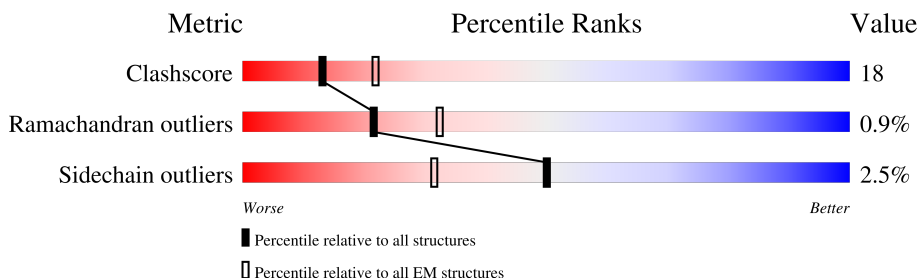
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 3.80 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 ≥ 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	

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 12244 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	1175	Total	C	N	O	S	0	0
			9318	5991	1595	1677	55		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	363	Total	C	N	O	S	0	0
			2926	1850	532	525	19		

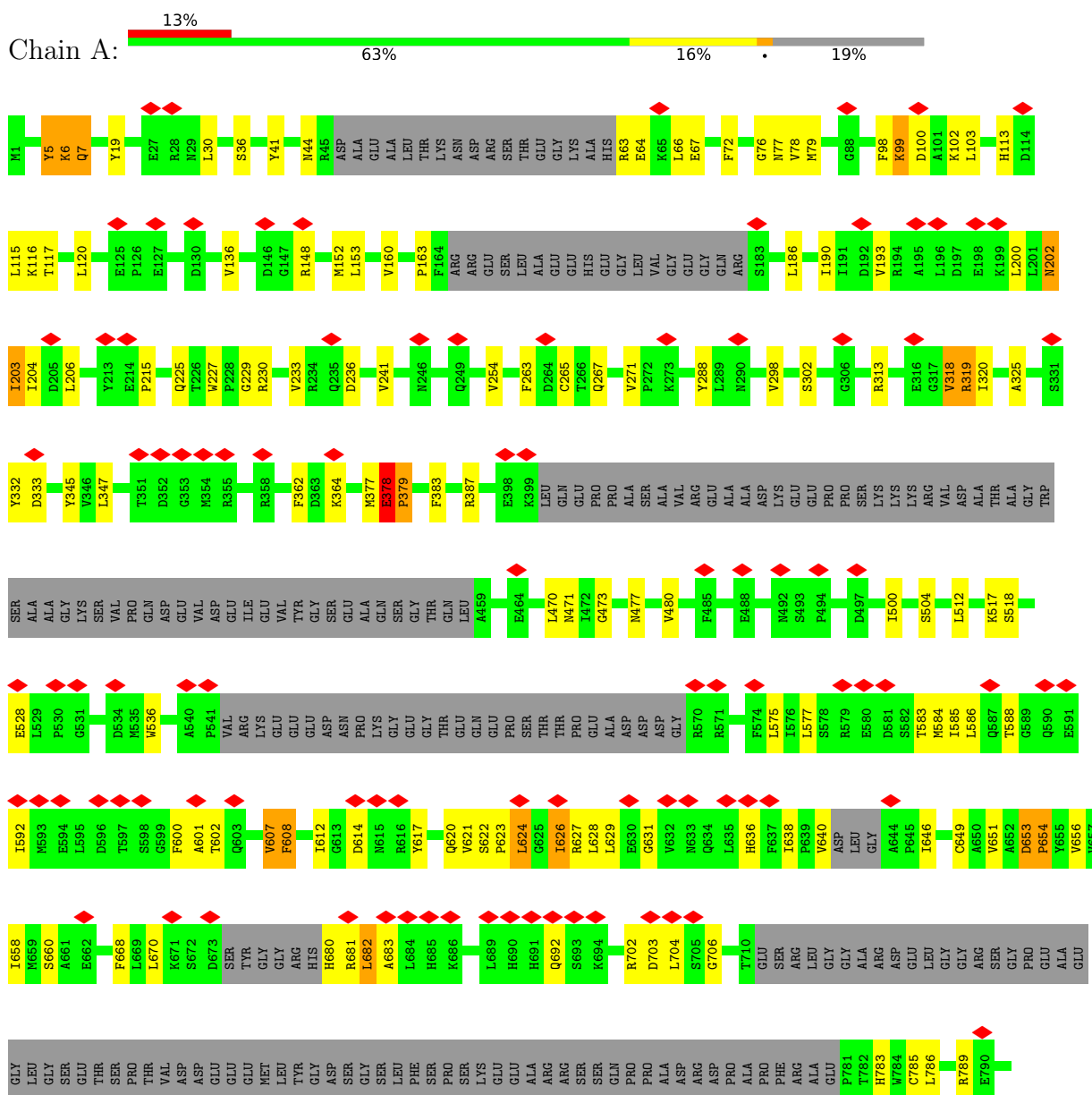
There are 15 discrepancies between the modelled and reference sequences:

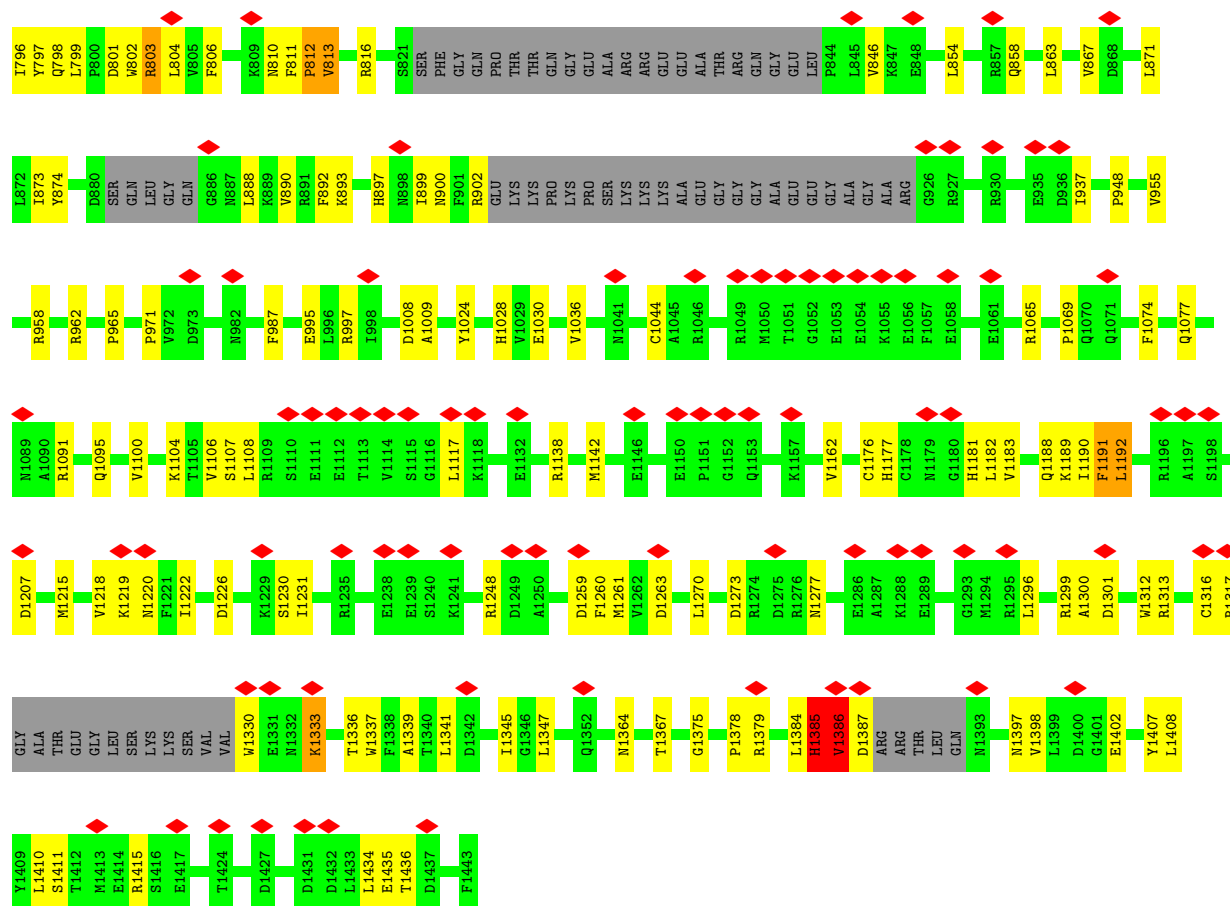
Chain	Residue	Modelled	Actual	Comment	Reference
B	-14	MET	-	expression tag	UNP Q9C0J8
B	-13	GLY	-	expression tag	UNP Q9C0J8
B	-12	SER	-	expression tag	UNP Q9C0J8
B	-11	SER	-	expression tag	UNP Q9C0J8
B	-10	HIS	-	expression tag	UNP Q9C0J8
B	-9	HIS	-	expression tag	UNP Q9C0J8
B	-8	HIS	-	expression tag	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

3 Residue-property plots

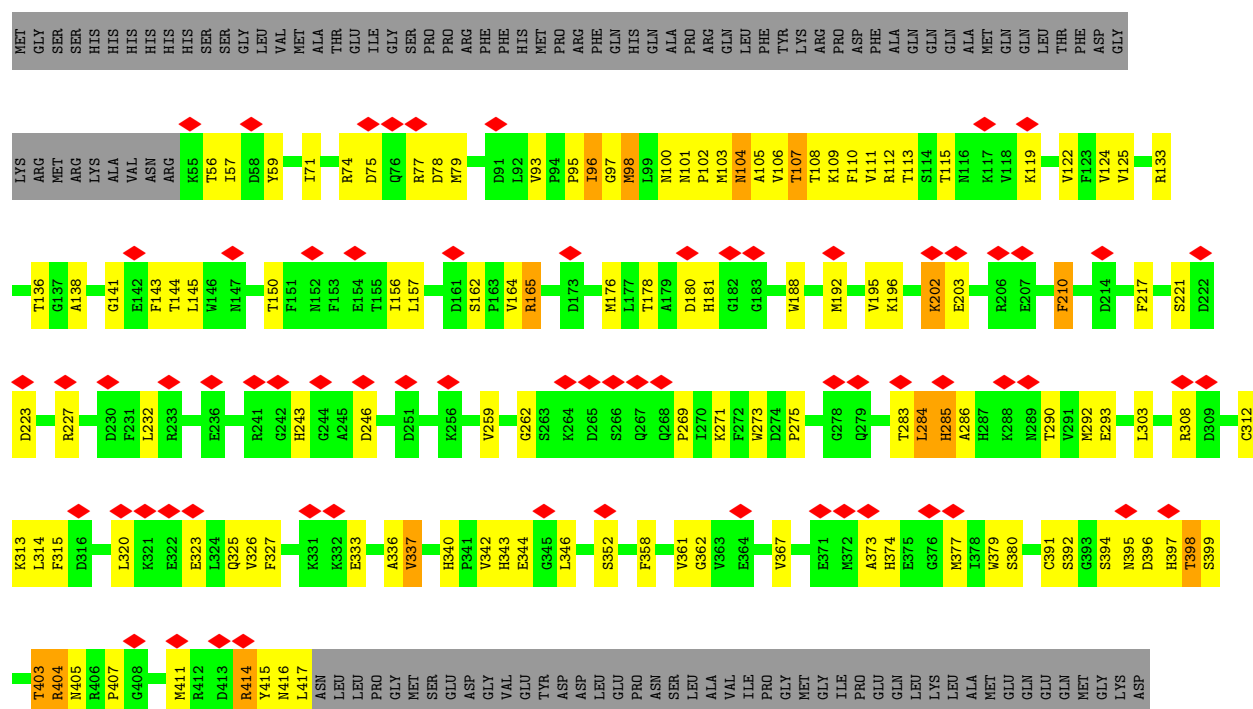
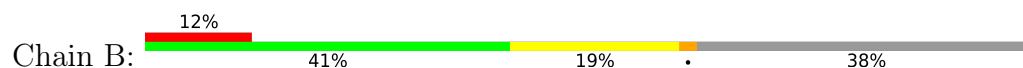
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and atom inclusion 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: Cleavage and polyadenylation specificity factor subunit 1





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



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

4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	205373	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 ($e^-/\text{\AA}^2$)	47	Depositor
Minimum defocus (nm)	800	Depositor
Maximum defocus (nm)	3000	Depositor
Magnification	38462	Depositor
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor
Maximum map value	0.257	Depositor
Minimum map value	-0.136	Depositor
Average map value	-0.000	Depositor
Map value standard deviation	0.010	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](#)

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	0.35	0/9534	0.58	0/12942
2	B	0.37	0/3010	0.61	0/4080
All	All	0.36	0/12544	0.58	0/17022

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	9318	0	9366	285	0
2	B	2926	0	2832	168	0
All	All	12244	0	12198	446	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 18.

The worst 5 of 446 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:A:628:LEU:HD11	1:A:682:LEU:CD2	1.26	1.61
1:A:318:VAL:HG21	1:A:362:PHE:CD2	1.58	1.38

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:626:ILE:HG12	1:A:668:PHE:CE2	1.58	1.36
1:A:617:TYR:HE2	1:A:680:HIS:ND1	1.33	1.26
1:A:816:ARG:HH21	1:A:1410:LEU:CD2	1.52	1.22

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	1149/1443 (80%)	1045 (91%)	95 (8%)	9 (1%)	19	57
2	B	361/587 (62%)	316 (88%)	40 (11%)	5 (1%)	11	46
All	All	1510/2030 (74%)	1361 (90%)	135 (9%)	14 (1%)	21	54

5 of 14 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	378	GLU
1	A	379	PRO
1	A	654	PRO
1	A	812	PRO
1	A	813	VAL

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	1028/1235 (83%)	1005 (98%)	23 (2%)	52	72
2	B	315/514 (61%)	304 (96%)	11 (4%)	36	64
All	All	1343/1749 (77%)	1309 (98%)	34 (2%)	50	70

5 of 34 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	284	LEU
2	B	285	HIS
2	B	404	ARG
1	A	608	PHE
1	A	607	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 15 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	1188	GLN
2	B	397	HIS
1	A	1332	ASN
2	B	416	ASN
2	B	285	HIS

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

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.

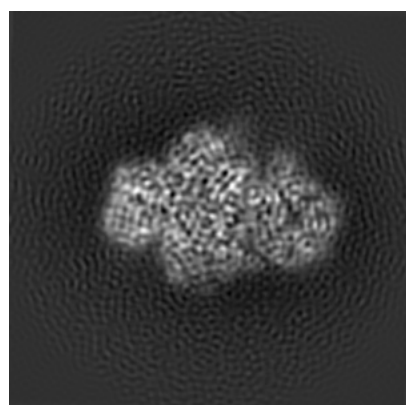
6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-7114. These allow visual inspection of the internal detail of the map and identification of artifacts.

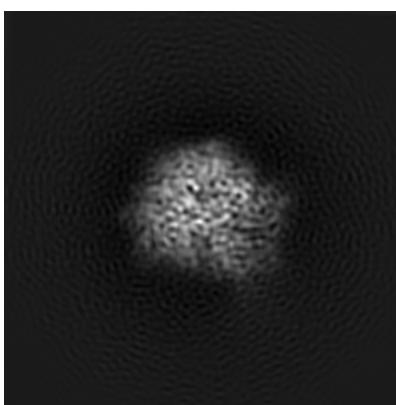
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections [i](#)

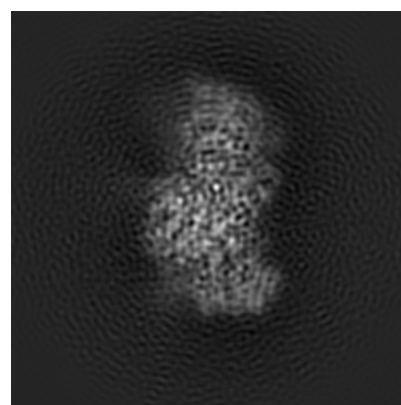
6.1.1 Primary map



X



Y

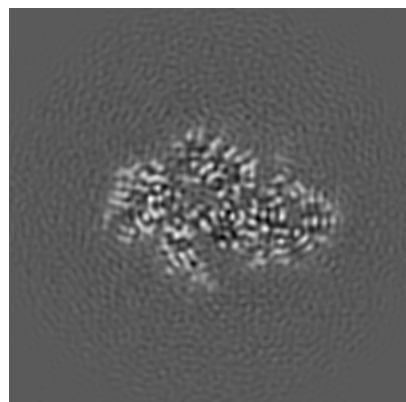


Z

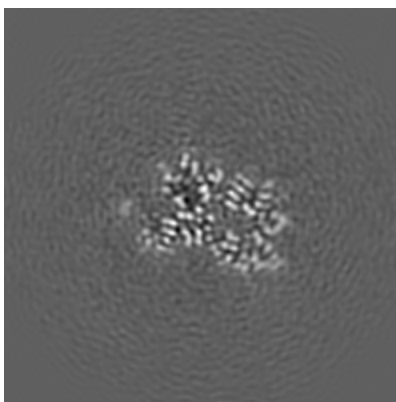
The images above show the map projected in three orthogonal directions.

6.2 Central slices [i](#)

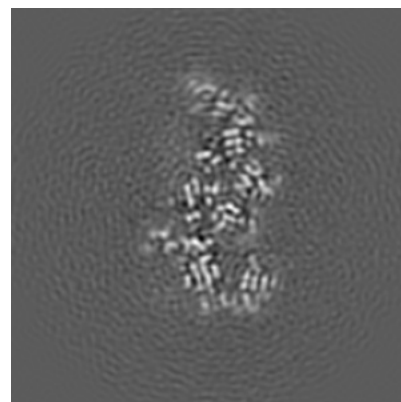
6.2.1 Primary map



X Index: 96



Y Index: 96

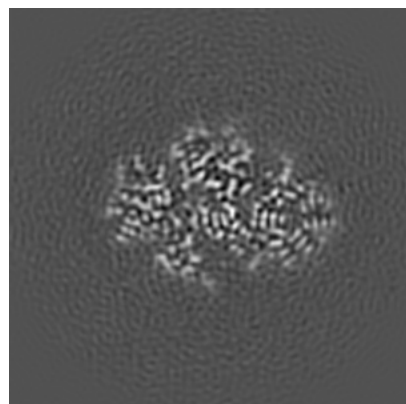


Z Index: 96

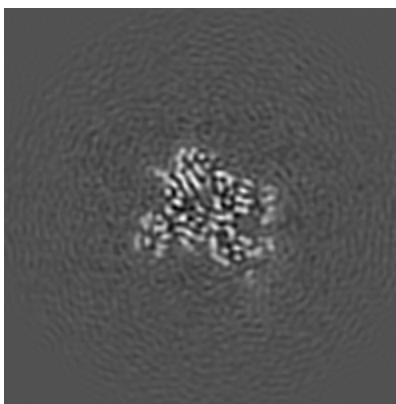
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices [i](#)

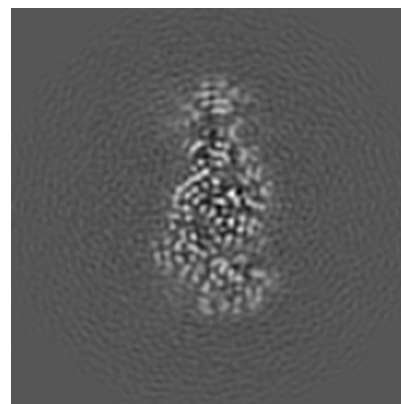
6.3.1 Primary map



X Index: 98



Y Index: 101

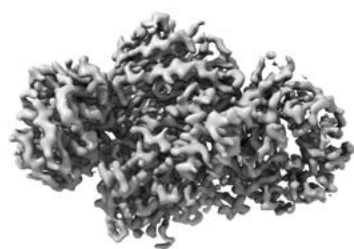


Z Index: 87

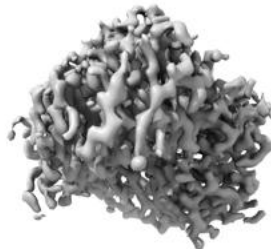
The images above show the largest variance slices of the map in three orthogonal directions.

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.

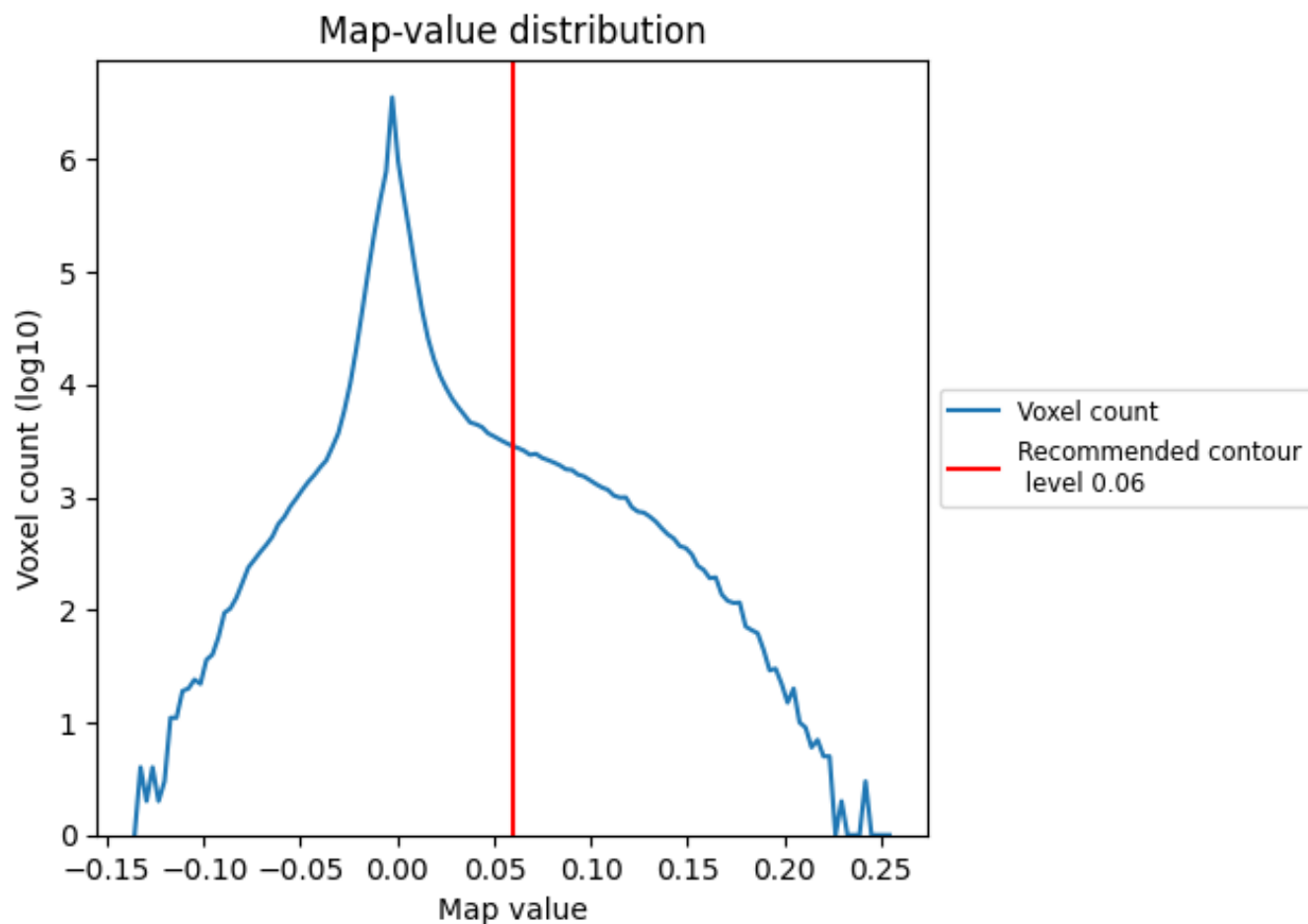
6.5 Mask visualisation

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

7 Map analysis [i](#)

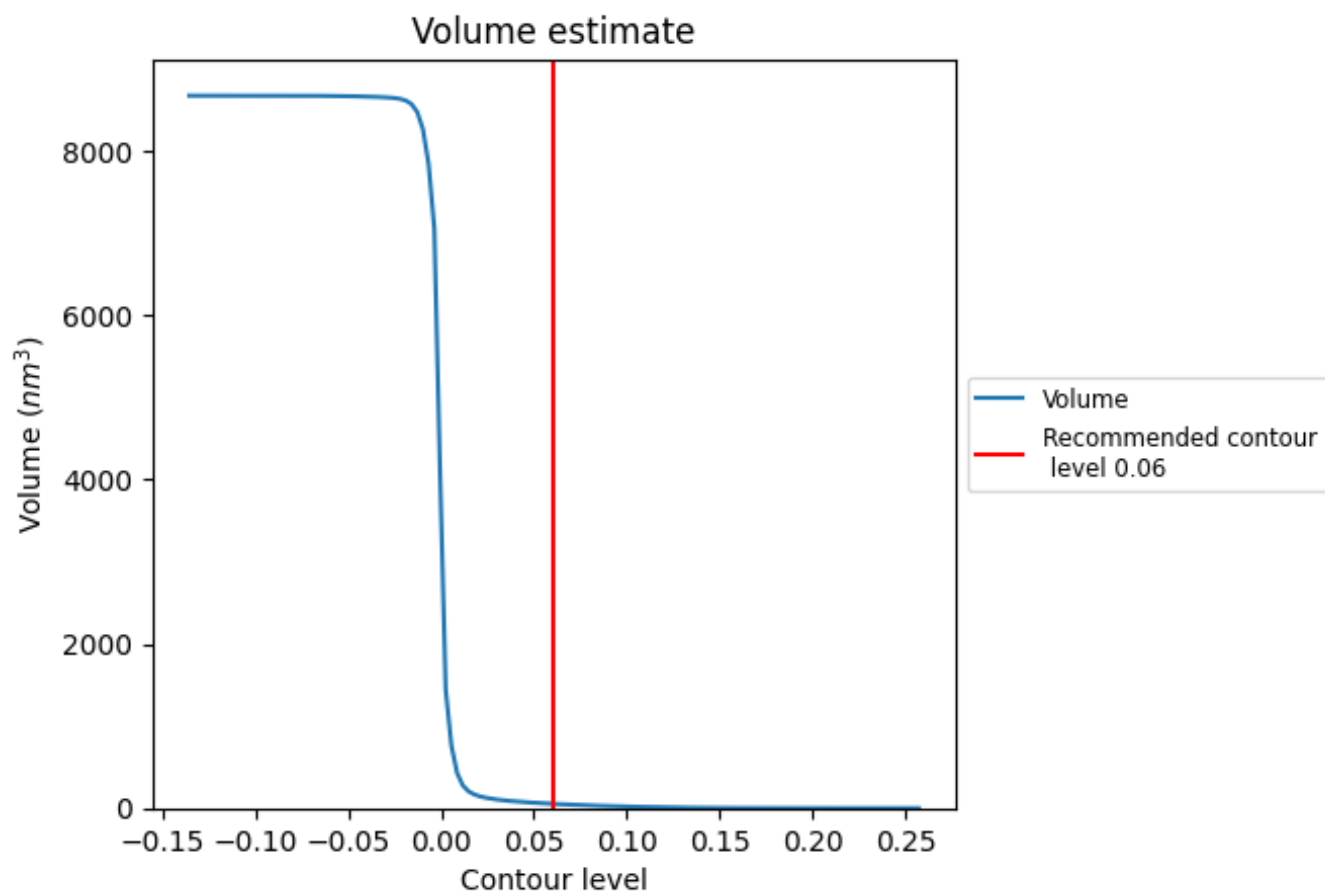
This section contains the results of statistical analysis of the map.

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.

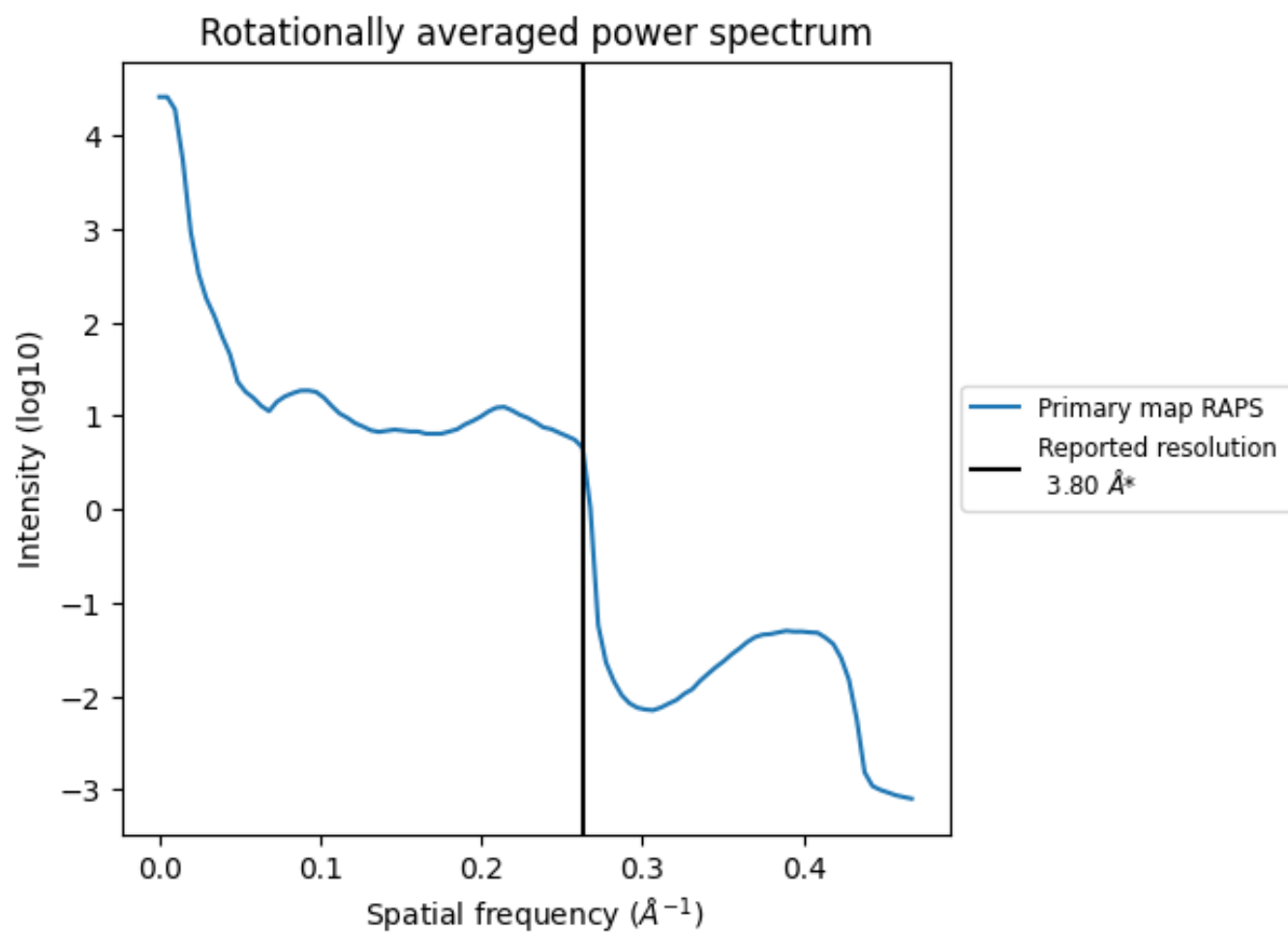
7.2 Volume estimate [i](#)



The volume at the recommended contour level is 53 nm³; this corresponds to an approximate mass of 48 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 0.263 Å⁻¹

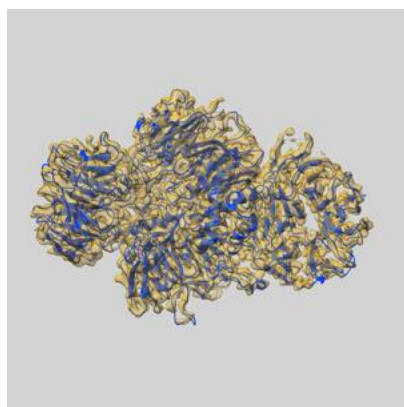
8 Fourier-Shell correlation

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

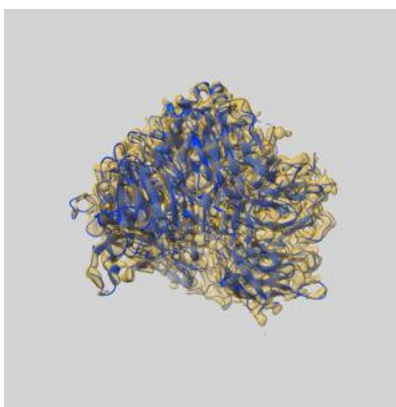
9 Map-model fit [i](#)

This section contains information regarding the fit between EMDB map EMD-7114 and PDB model 6BM0. Per-residue inclusion information can be found in [section 3](#) on [page 4](#).

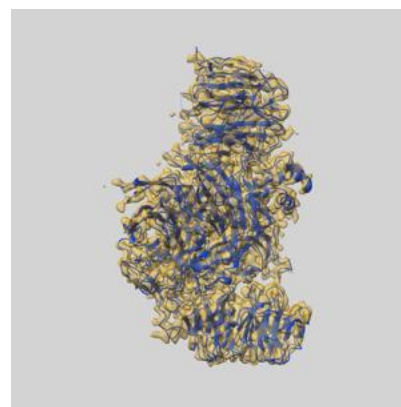
9.1 Map-model overlay [i](#)



X



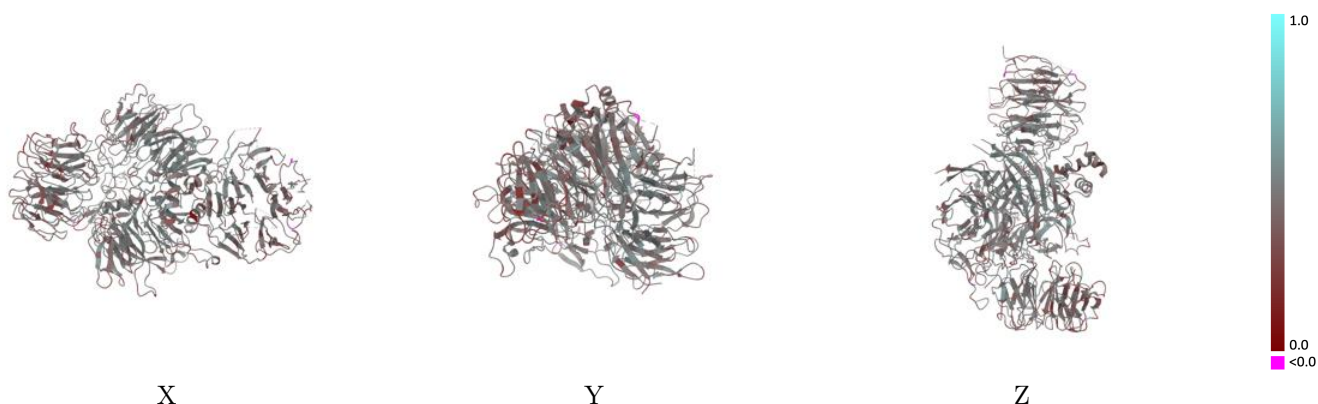
Y



Z

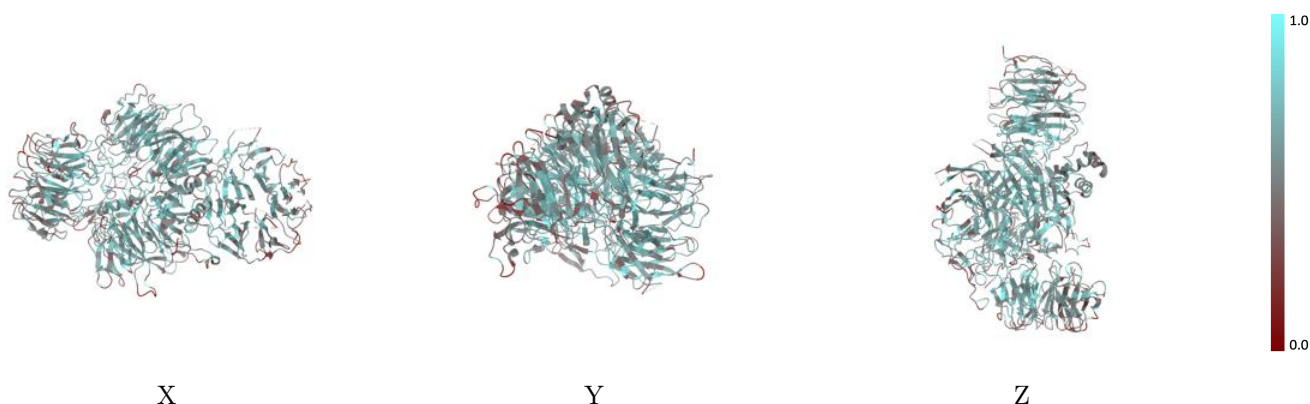
The images above show the 3D surface view of the map at the recommended contour level 0.06 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

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



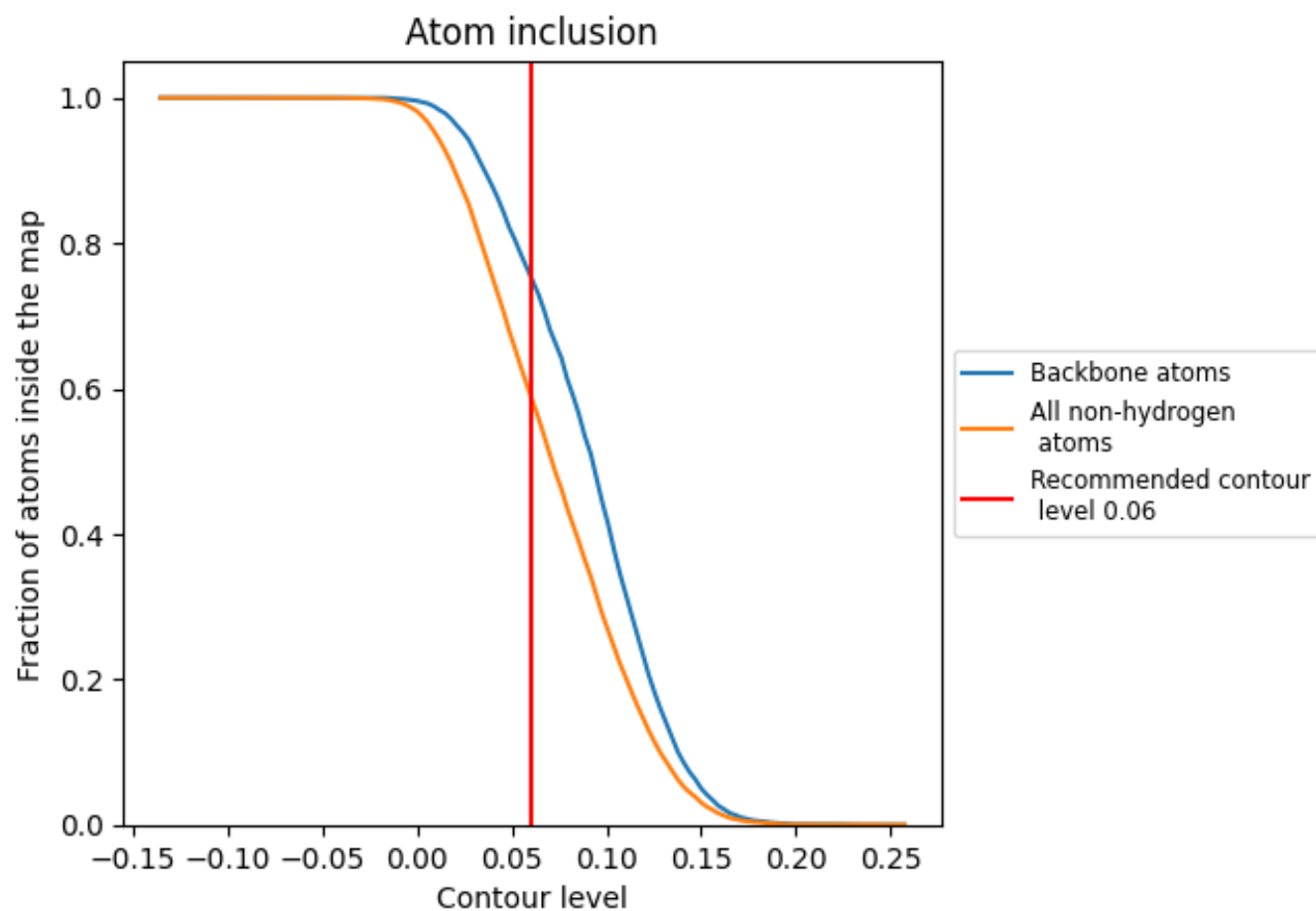
The images above show the model with each residue coloured according to its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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, 75% of all backbone atoms, 59% 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> 0.5886	<div></div> 0.4310
A	<div></div> 0.5950	<div></div> 0.4360
B	<div></div> 0.5681	<div></div> 0.4140

