



wwPDB EM Validation Summary Report ⓘ

Dec 31, 2024 – 06:50 AM EST

PDB ID : 8HQO
EMDB ID : EMD-34952
Title : Neck of DT57C bacteriophage in the full state
Authors : Ayala, R.; Moiseenko, A.V.; Chen, T.H.; Kulikov, E.E.; Golomidova, A.K.;
Orekhov, P.S.; Street, M.A.; Sokolova, O.S.; Letarov, A.V.; Wolf, M.
Deposited on : 2022-12-13
Resolution : 3.20 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.dev113
MolProbity : 4.02b-467
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)
MapQ : 1.9.13
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.40

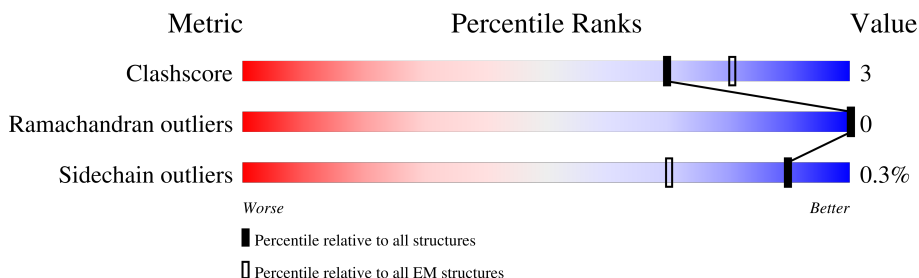
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 3.20 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.



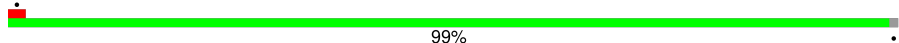
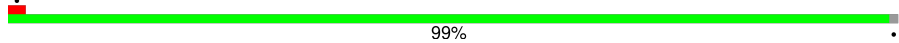
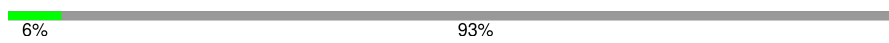
Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	210492	15764
Ramachandran outliers	207382	16835
Sidechain outliers	206894	16415

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	B	405	<div> <div>7%</div> <div>84%</div> <div>7%</div> <div>9%</div> </div>
1	C	405	<div> <div>8%</div> <div>85%</div> <div>6%</div> <div>9%</div> </div>
1	D	405	<div> <div>6%</div> <div>85%</div> <div>6%</div> <div>9%</div> </div>
1	E	405	<div> <div>7%</div> <div>84%</div> <div>6%</div> <div>9%</div> </div>
2	P	170	<div> <div>94%</div> <div>6%</div> </div>
2	Q	170	<div> <div>95%</div> <div>5%</div> </div>
2	R	170	<div> <div>94%</div> <div>6%</div> </div>
2	S	170	<div> <div>92%</div> <div>8%</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	a	161	 99%
3	b	161	 99%
4	x	1227	 6%93%

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 20871 atoms, of which 634 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 Portal protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	D	368	Total	C	N	O	S	0	0
			2925	1858	494	559	14		
1	E	368	Total	C	N	O	S	0	0
			2925	1858	494	559	14		
1	B	368	Total	C	N	O	S	0	0
			2925	1858	494	559	14		
1	C	368	Total	C	N	O	S	0	0
			2925	1858	494	559	14		

- Molecule 2 is a protein called Head completion protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	R	170	Total	C	N	O	S	0	0
			1345	853	222	262	8		
2	S	170	Total	C	N	O	S	0	0
			1345	853	222	262	8		
2	P	170	Total	C	N	O	S	0	0
			1345	853	222	262	8		
2	Q	170	Total	C	N	O	S	0	0
			1345	853	222	262	8		

- Molecule 3 is a protein called Tail terminator protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	a	160	Total	C	N	O	S	0	0
			1279	805	215	255	4		
3	b	160	Total	C	N	O	S	0	0
			1279	805	215	255	4		

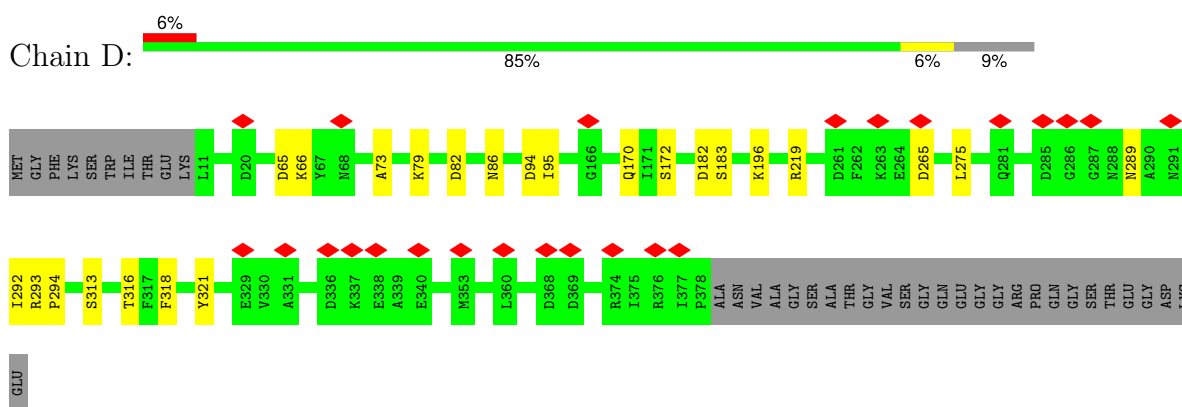
- Molecule 4 is a protein called Tape measure protein.

Mol	Chain	Residues	Atoms						AltConf	Trace
4	x	80	Total	C	H	N	O	S	0	0
			1233	366	634	109	120	4		

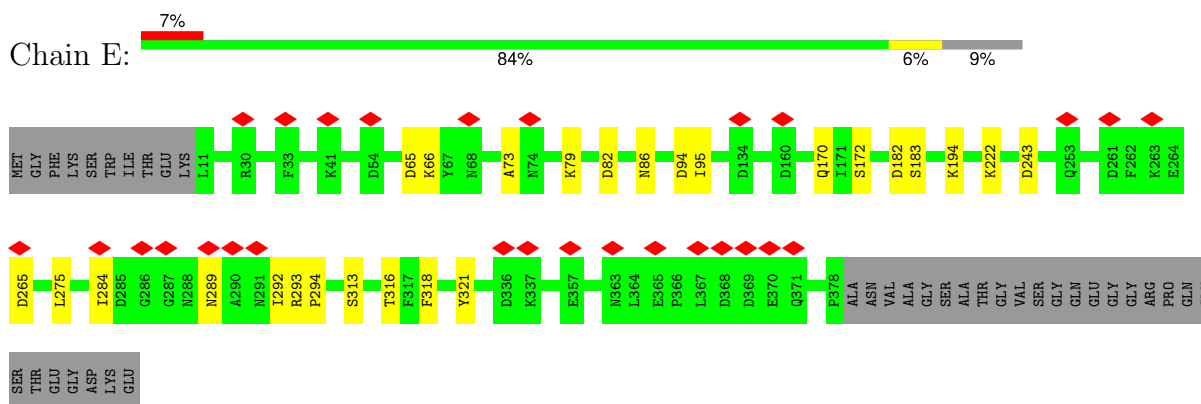
3 Residue-property plots [i](#)

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.

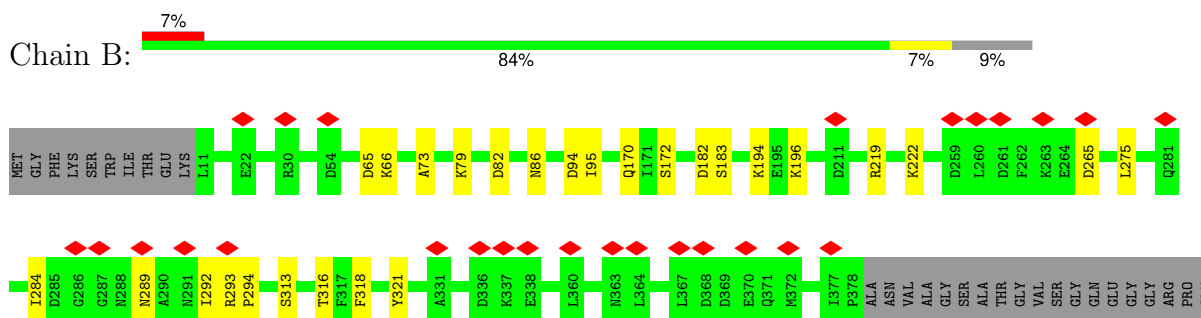
- Molecule 1: Portal protein



- Molecule 1: Portal protein

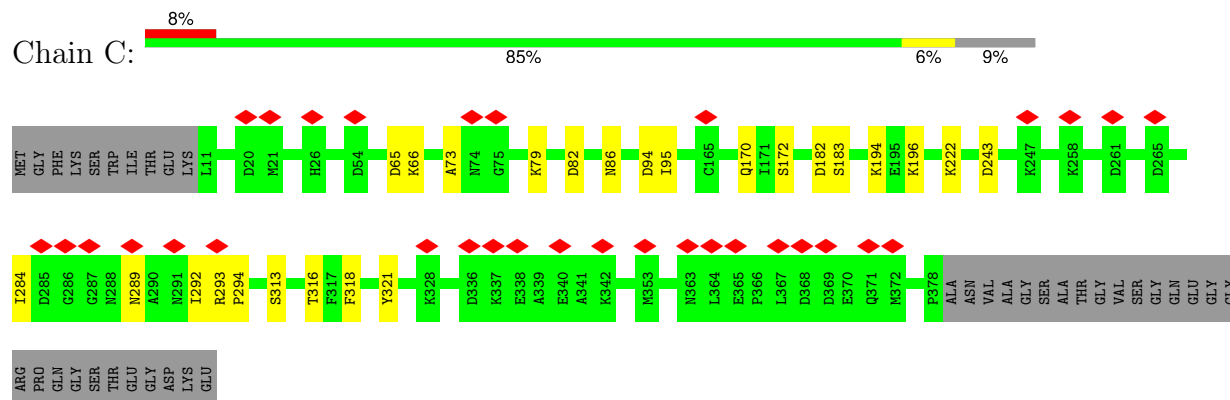


- Molecule 1: Portal protein

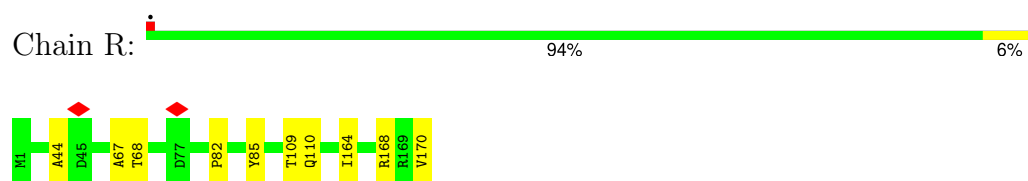


GLY
SER
THR
GLU
GLY
ASP
LYS
GLU

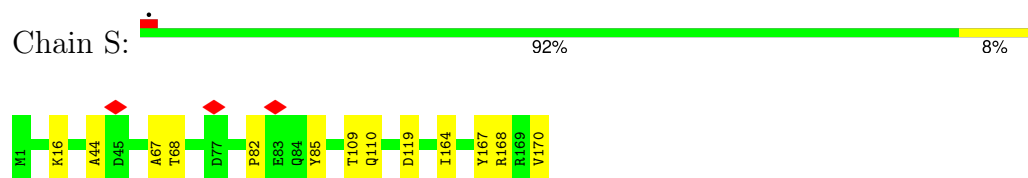
• Molecule 1: Portal protein



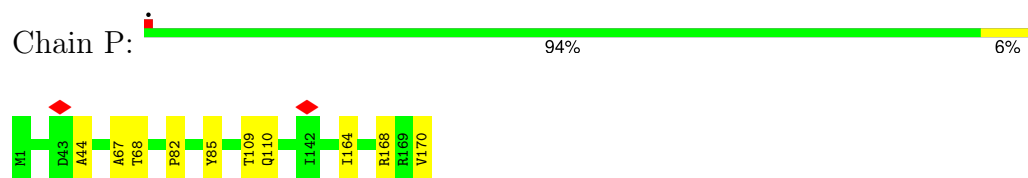
• Molecule 2: Head completion protein



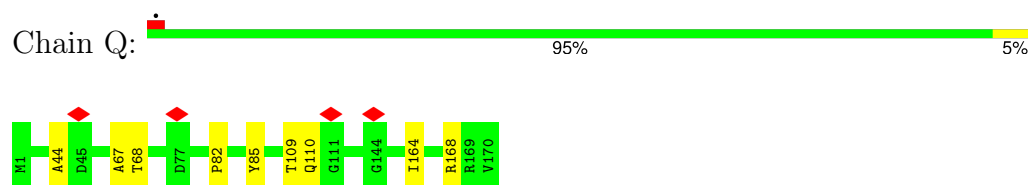
• Molecule 2: Head completion protein



• Molecule 2: Head completion protein



• Molecule 2: Head completion protein



• Molecule 3: Tail terminator protein



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

4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	34102	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 ($e^-/\text{\AA}^2$)	67	Depositor
Minimum defocus (nm)	1000	Depositor
Maximum defocus (nm)	4000	Depositor
Magnification	Not provided	
Image detector	FEI FALCON III (4k x 4k)	Depositor
Maximum map value	31.963	Depositor
Minimum map value	-9.475	Depositor
Average map value	-0.017	Depositor
Map value standard deviation	0.972	Depositor
Recommended contour level	3.1	Depositor
Map size (Å)	420.0, 420.0, 420.0	wwPDB
Map dimensions	300, 300, 300	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.4, 1.4, 1.4	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	B	0.36	0/2976	0.54	0/4024
1	C	0.36	0/2976	0.54	0/4024
1	D	0.36	0/2976	0.54	0/4024
1	E	0.36	0/2976	0.54	0/4024
2	P	0.38	0/1371	0.55	0/1856
2	Q	0.38	0/1371	0.55	0/1856
2	R	0.38	0/1371	0.55	0/1856
2	S	0.38	0/1371	0.55	0/1856
3	a	0.33	0/1305	0.55	0/1773
3	b	0.33	0/1305	0.55	0/1773
4	x	0.24	0/600	0.52	0/802
All	All	0.36	0/20598	0.54	0/27868

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	B	2925	0	2950	20	0
1	C	2925	0	2950	17	0
1	D	2925	0	2950	16	0
1	E	2925	0	2950	19	0
2	P	1345	0	1347	6	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	Q	1345	0	1347	5	0
2	R	1345	0	1347	6	0
2	S	1345	0	1347	8	0
3	a	1279	0	1244	0	0
3	b	1279	0	1244	0	0
4	x	599	634	634	0	0
All	All	20237	634	20310	83	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 83 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:B:82:ASP:OD1	1:B:86:ASN:ND2	2.13	0.81
1:C:82:ASP:OD1	1:C:86:ASN:ND2	2.13	0.81
1:D:82:ASP:OD1	1:D:86:ASN:ND2	2.13	0.80
1:E:82:ASP:OD1	1:E:86:ASN:ND2	2.13	0.80
2:S:44:ALA:O	2:S:110:GLN:NE2	2.30	0.62

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	366/405 (90%)	353 (96%)	13 (4%)	0	100	100
1	C	366/405 (90%)	353 (96%)	13 (4%)	0	100	100
1	D	366/405 (90%)	353 (96%)	13 (4%)	0	100	100
1	E	366/405 (90%)	353 (96%)	13 (4%)	0	100	100
2	P	168/170 (99%)	168 (100%)	0	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	Q	168/170 (99%)	168 (100%)	0	0	100	100
2	R	168/170 (99%)	168 (100%)	0	0	100	100
2	S	168/170 (99%)	168 (100%)	0	0	100	100
3	a	158/161 (98%)	152 (96%)	6 (4%)	0	100	100
3	b	158/161 (98%)	152 (96%)	6 (4%)	0	100	100
4	x	78/1227 (6%)	75 (96%)	3 (4%)	0	100	100
All	All	2530/3849 (66%)	2463 (97%)	67 (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	328/354 (93%)	327 (100%)	1 (0%)	91	96
1	C	328/354 (93%)	327 (100%)	1 (0%)	91	96
1	D	328/354 (93%)	327 (100%)	1 (0%)	91	96
1	E	328/354 (93%)	327 (100%)	1 (0%)	91	96
2	P	149/149 (100%)	149 (100%)	0	100	100
2	Q	149/149 (100%)	149 (100%)	0	100	100
2	R	149/149 (100%)	149 (100%)	0	100	100
2	S	149/149 (100%)	149 (100%)	0	100	100
3	a	142/143 (99%)	142 (100%)	0	100	100
3	b	142/143 (99%)	142 (100%)	0	100	100
4	x	63/967 (6%)	61 (97%)	2 (3%)	34	65
All	All	2255/3265 (69%)	2249 (100%)	6 (0%)	90	96

5 of 6 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	79	LYS
4	x	4	LYS
4	x	62	ARG
1	E	79	LYS
1	D	79	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 17 such sidechains are listed below:

Mol	Chain	Res	Type
3	b	18	GLN
3	b	153	GLN
3	a	18	GLN
3	a	23	GLN
3	a	153	GLN

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 oligosaccharides 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.

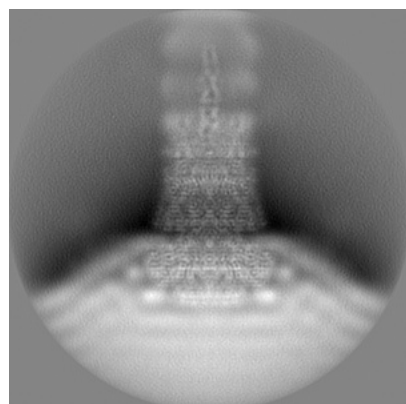
6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-34952. These allow visual inspection of the internal detail of the map and identification of artifacts.

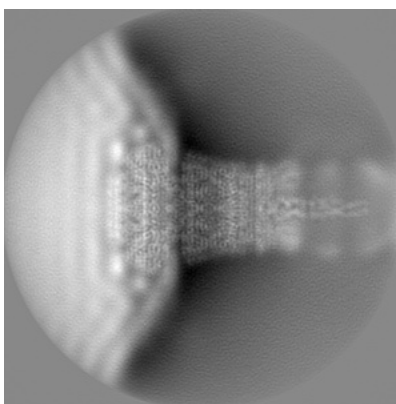
Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

6.1 Orthogonal projections [i](#)

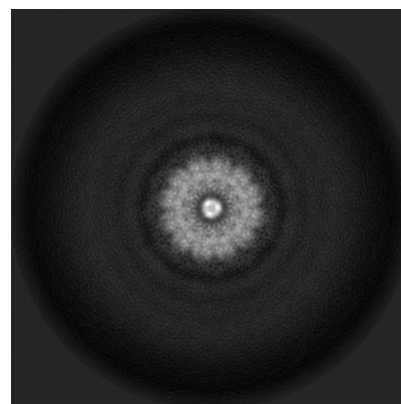
6.1.1 Primary map



X

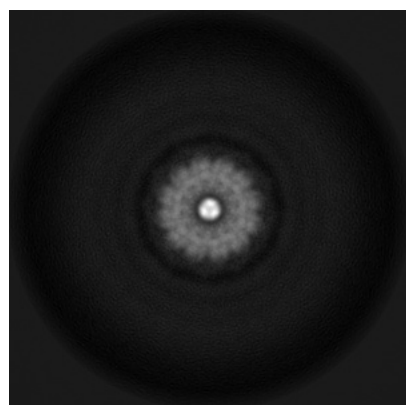


Y

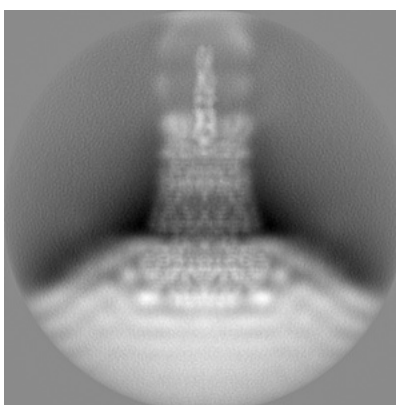


Z

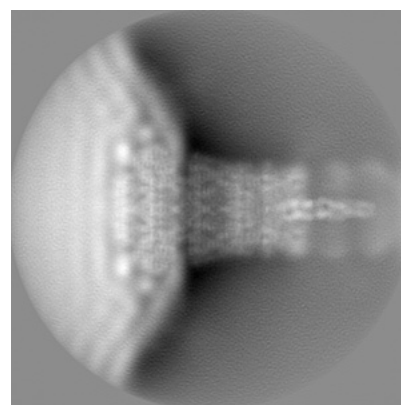
6.1.2 Raw map



X



Y

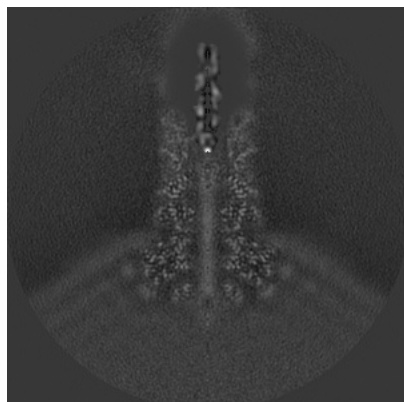


Z

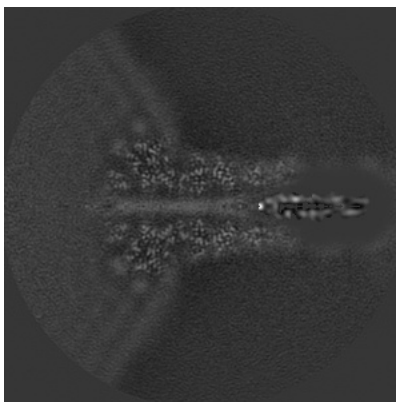
The images above show the map projected in three orthogonal directions.

6.2 Central slices [i](#)

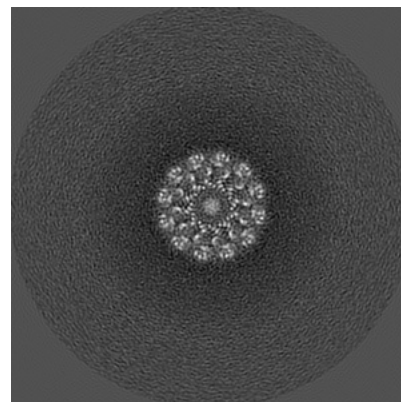
6.2.1 Primary map



X Index: 150

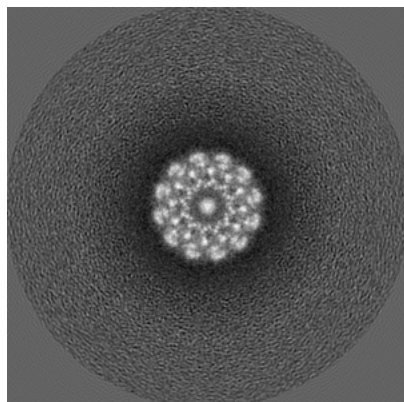


Y Index: 150

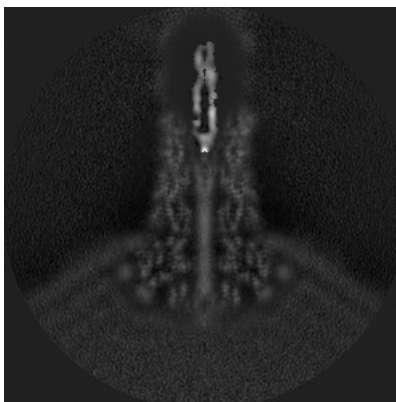


Z Index: 150

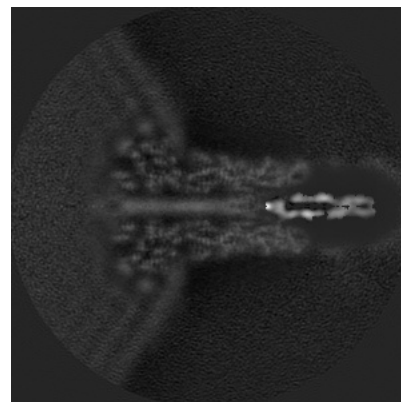
6.2.2 Raw map



X Index: 150



Y Index: 150

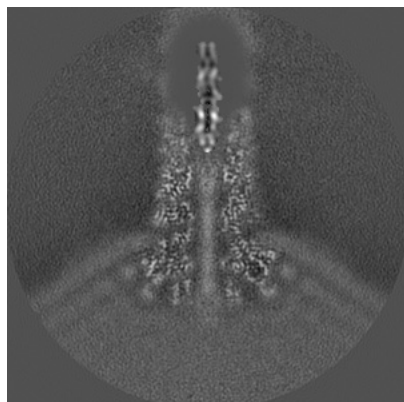


Z Index: 150

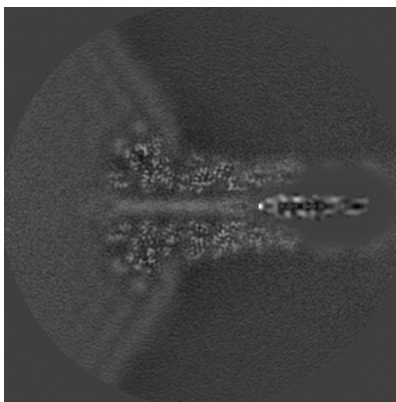
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices [i](#)

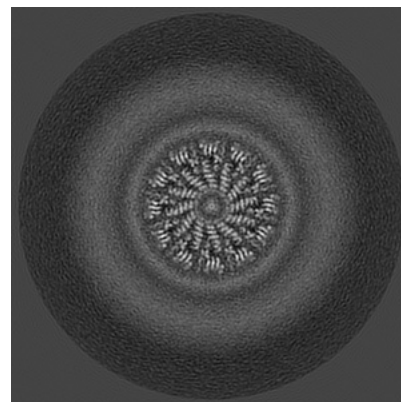
6.3.1 Primary map



X Index: 148

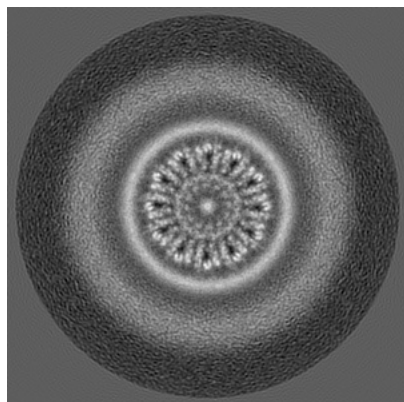


Y Index: 151

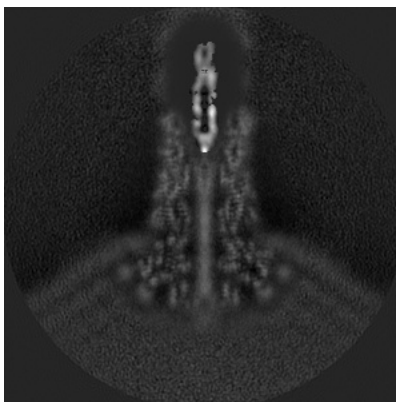


Z Index: 105

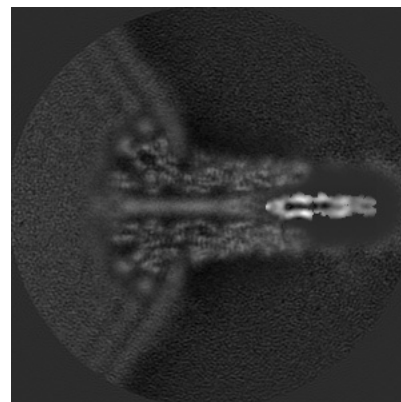
6.3.2 Raw map



X Index: 101



Y Index: 151

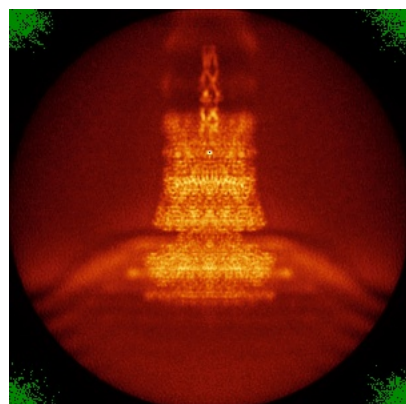


Z Index: 148

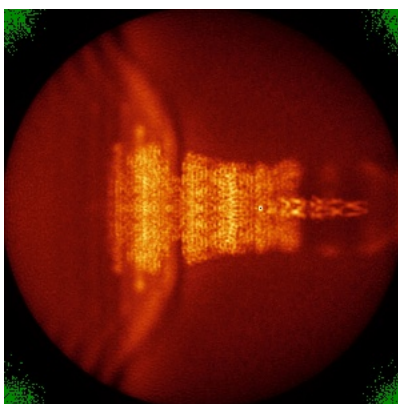
The images above show the largest variance slices of the map in three orthogonal directions.

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

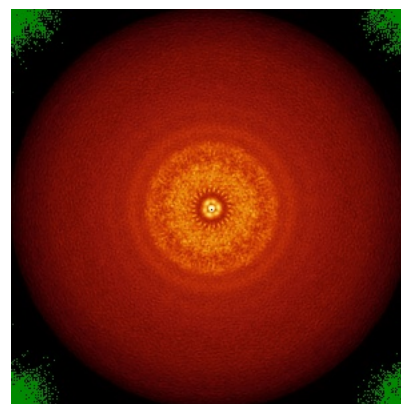
6.4.1 Primary map



X

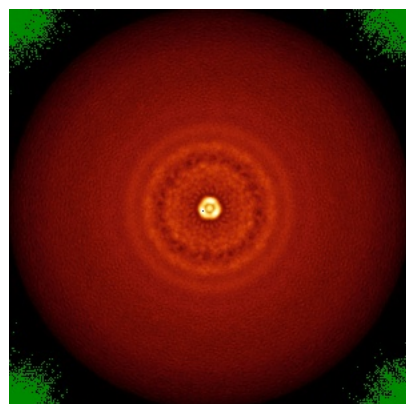


Y

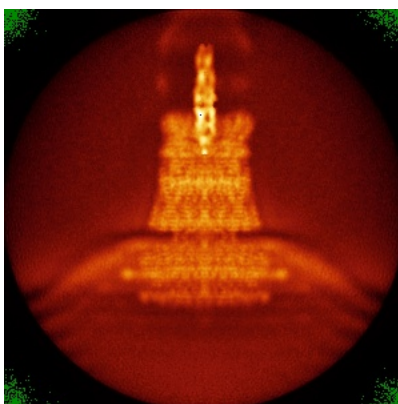


Z

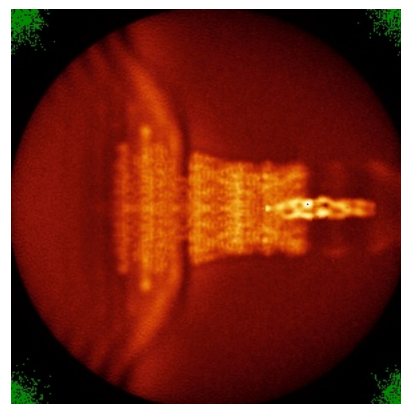
6.4.2 Raw map



X



Y

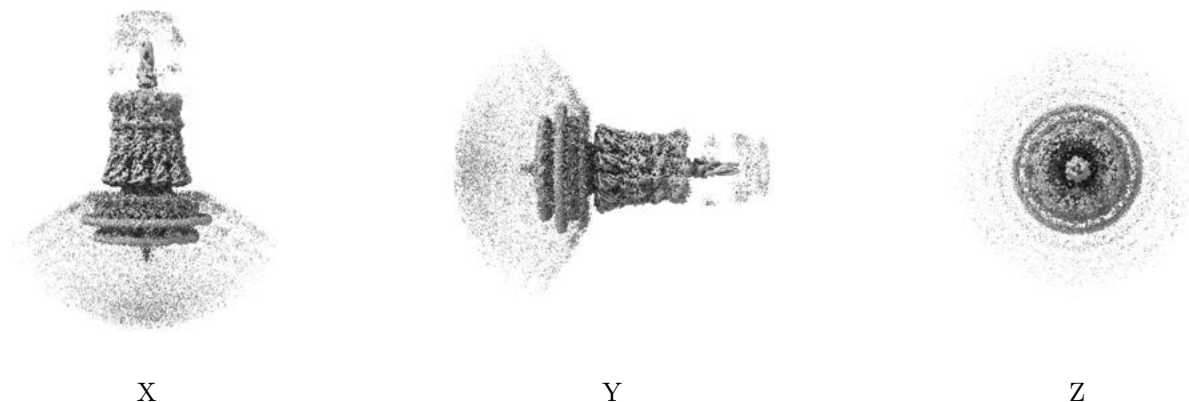


Z

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.

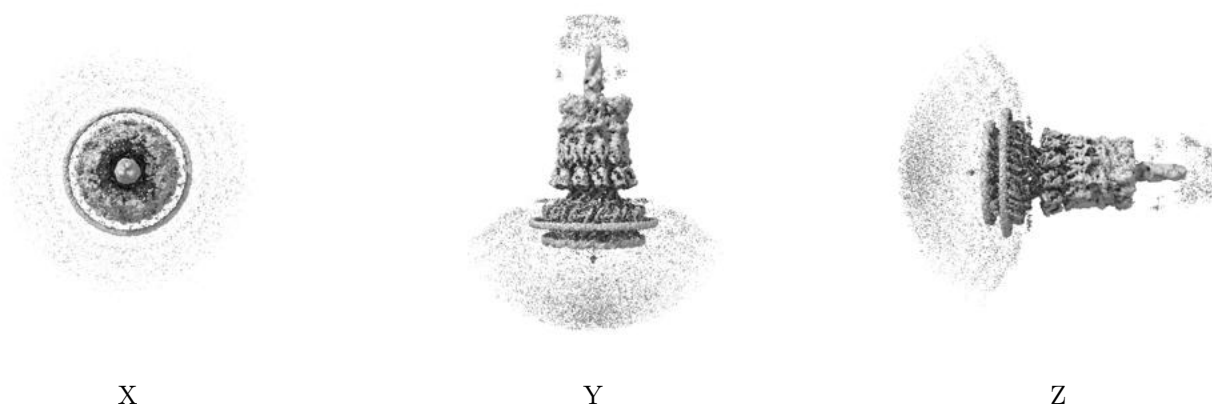
6.5 Orthogonal surface views [i](#)

6.5.1 Primary map



The images above show the 3D surface view of the map at the recommended contour level 3.1. 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.

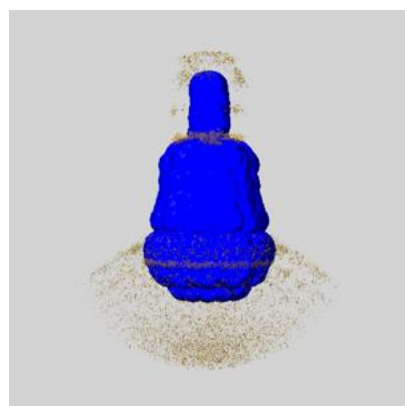
6.6 Mask visualisation [i](#)

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

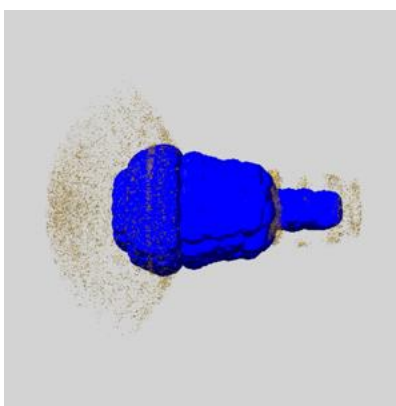
A mask typically either:

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

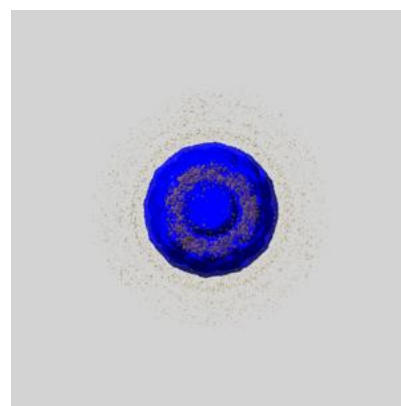
6.6.1 emd_34952_msk_1.map [i](#)



X



Y

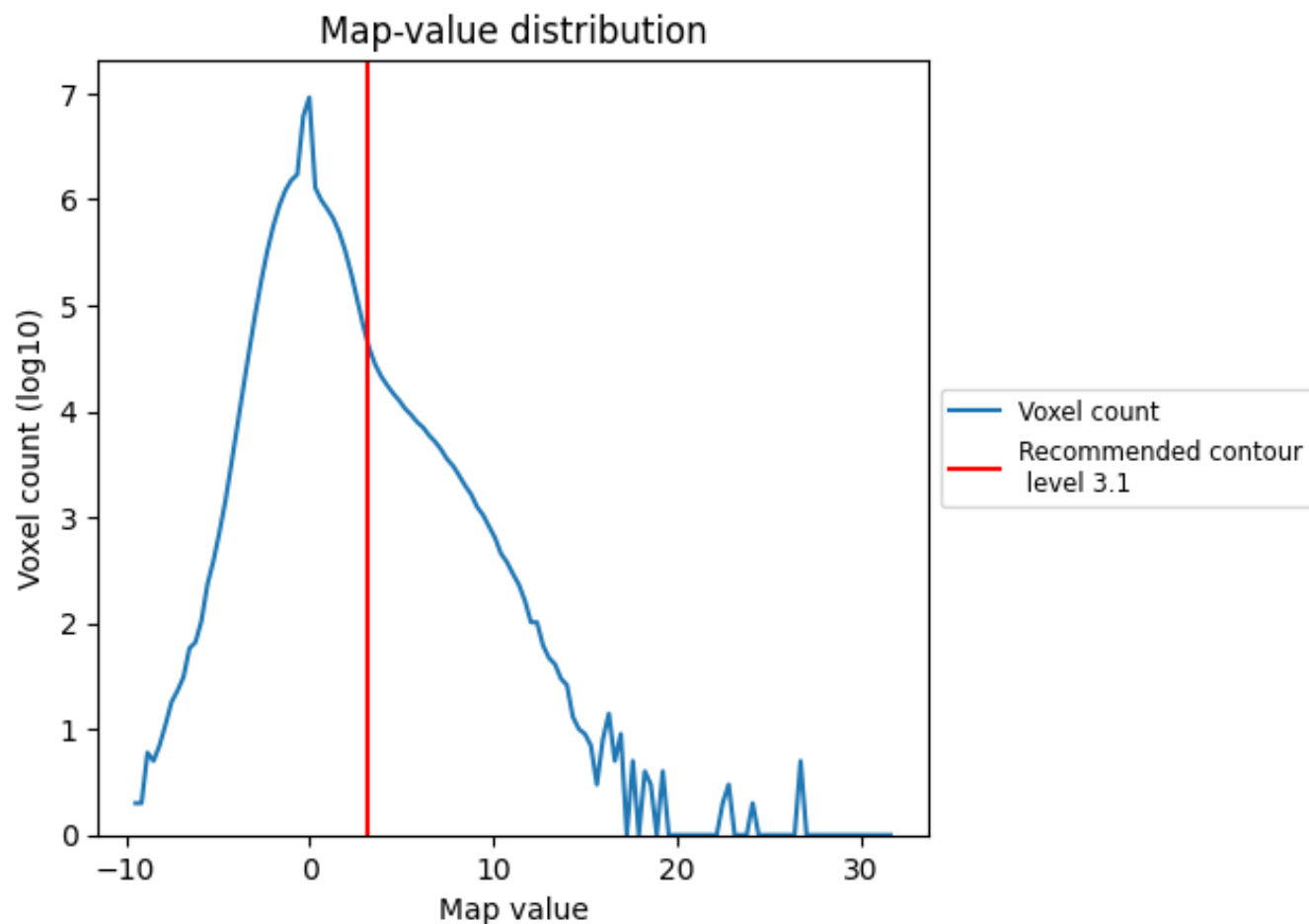


Z

7 Map analysis [i](#)

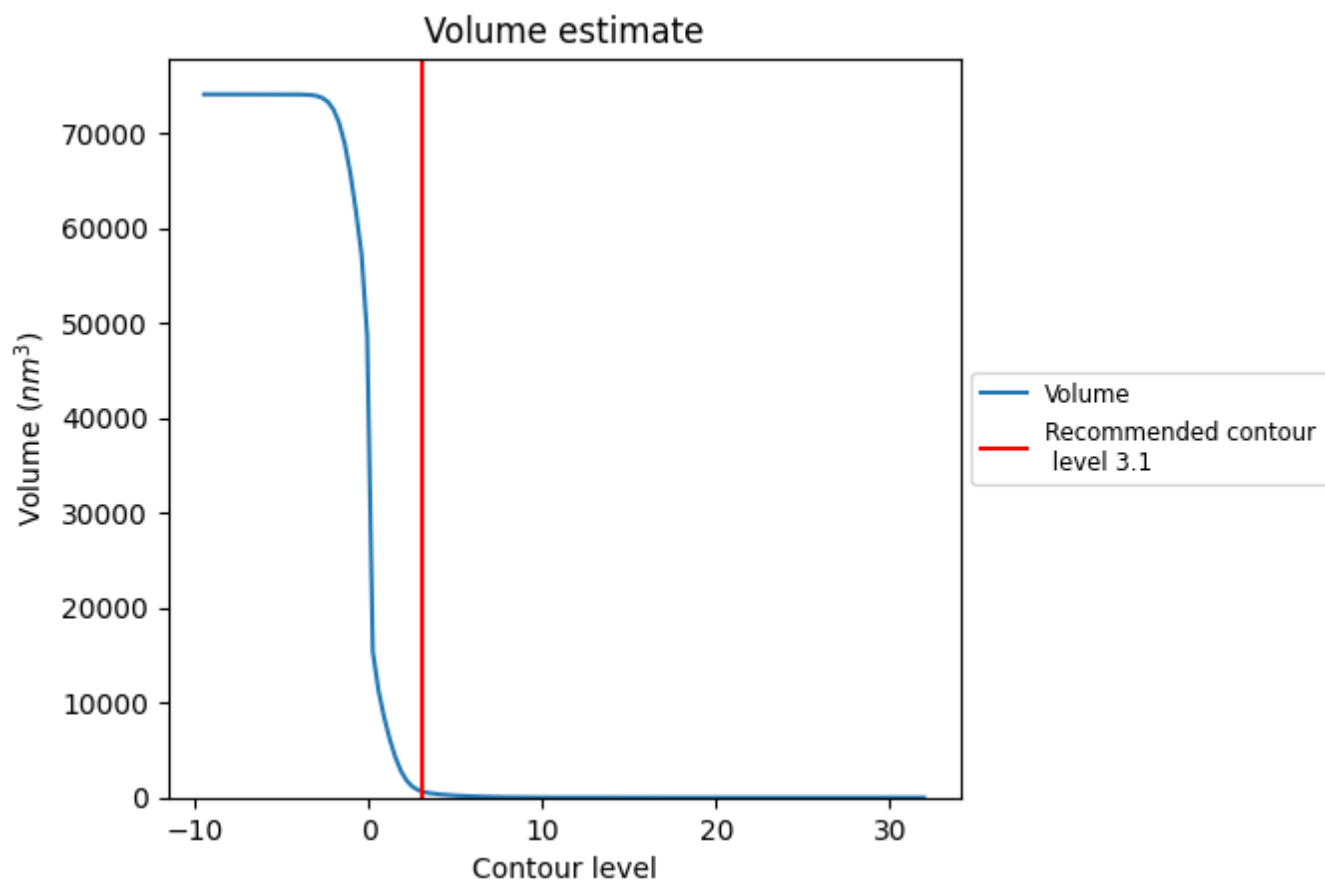
This section contains the results of statistical analysis of the map.

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.

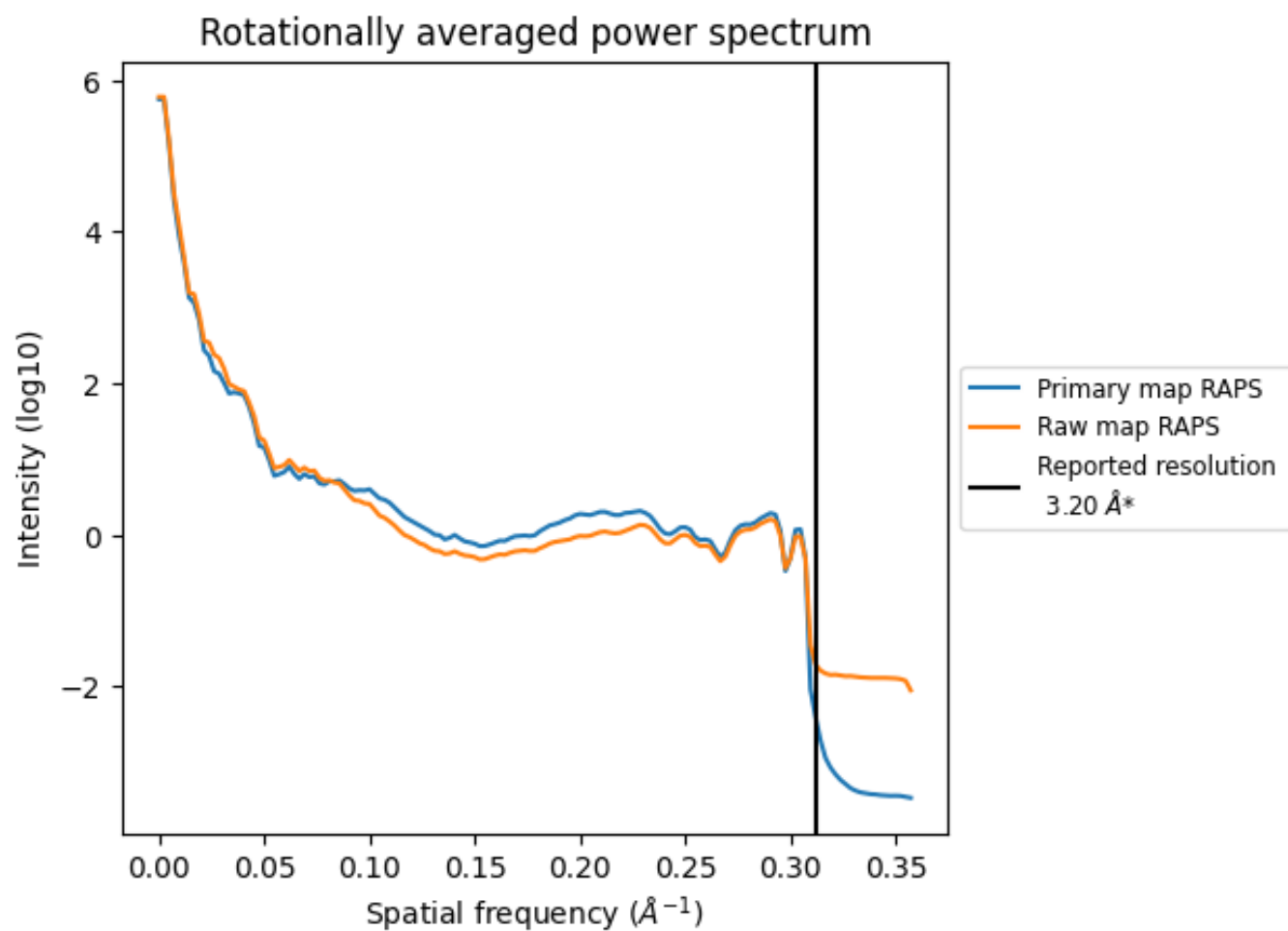
7.2 Volume estimate [i](#)



The volume at the recommended contour level is 631 nm^3 ; this corresponds to an approximate mass of 570 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 ⓘ

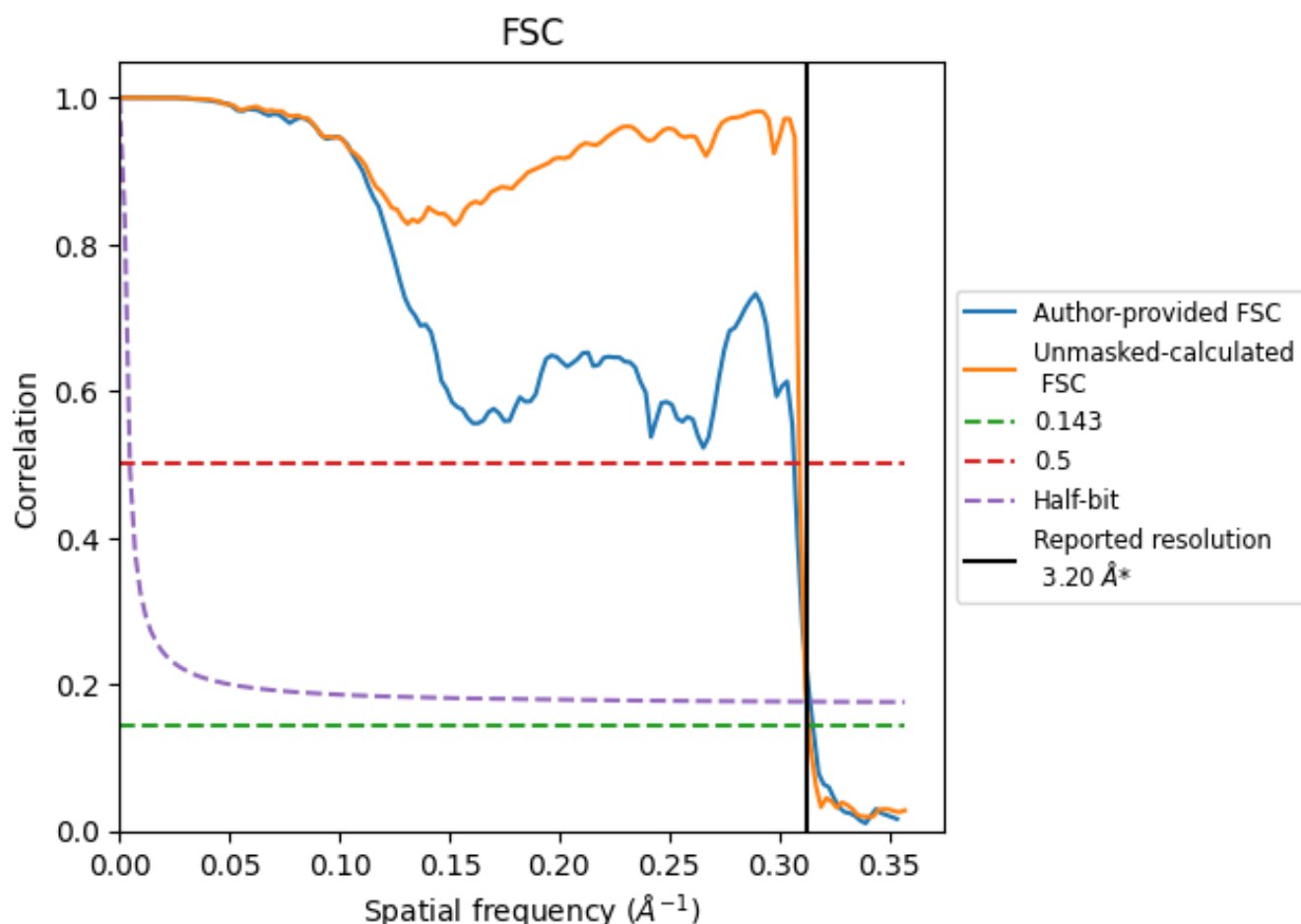


*Reported resolution corresponds to spatial frequency of 0.312 Å⁻¹

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 0.312 \AA^{-1}

8.2 Resolution estimates [i](#)

Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	3.20	-	-
Author-provided FSC curve	3.17	3.26	3.18
Unmasked-calculated*	3.19	3.23	3.20

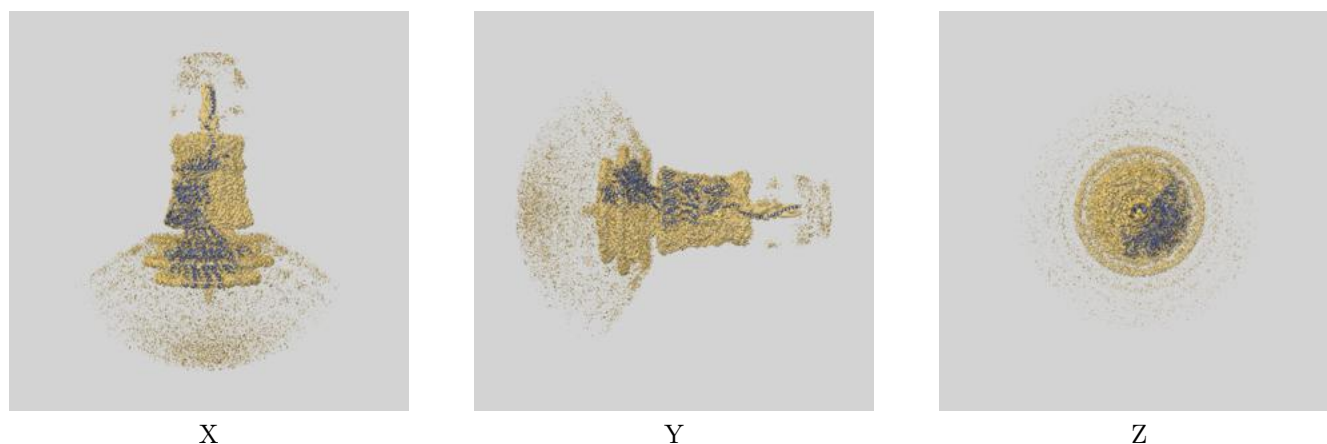
*Resolution estimate based on FSC curve calculated by comparison of deposited half-maps.

9 Map-model fit [i](#)

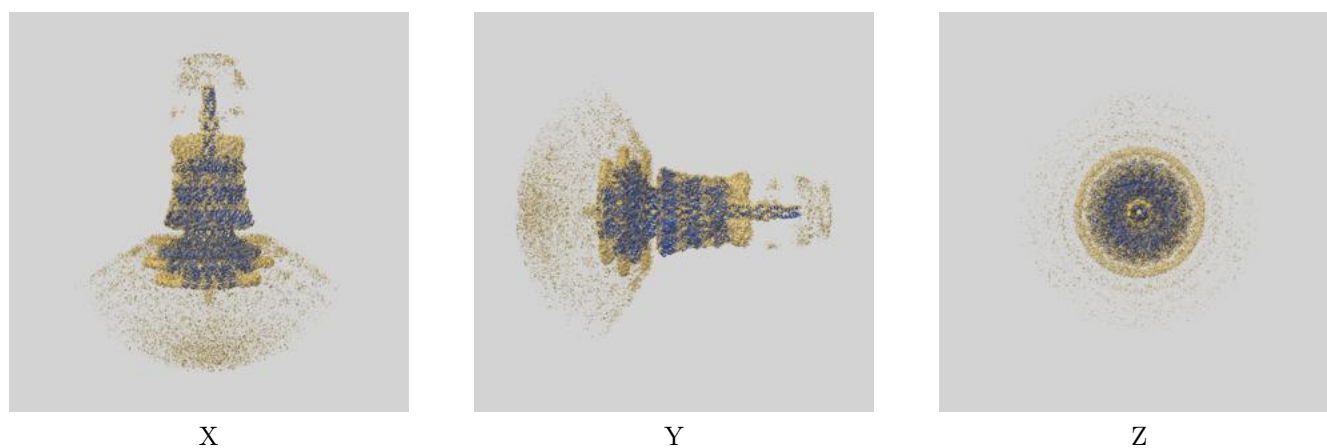
This section contains information regarding the fit between EMDB map EMD-34952 and PDB model 8HQO. Per-residue inclusion information can be found in section [3](#) on page [6](#).

9.1 Map-model overlays

9.1.1 Map-model overlay [i](#)

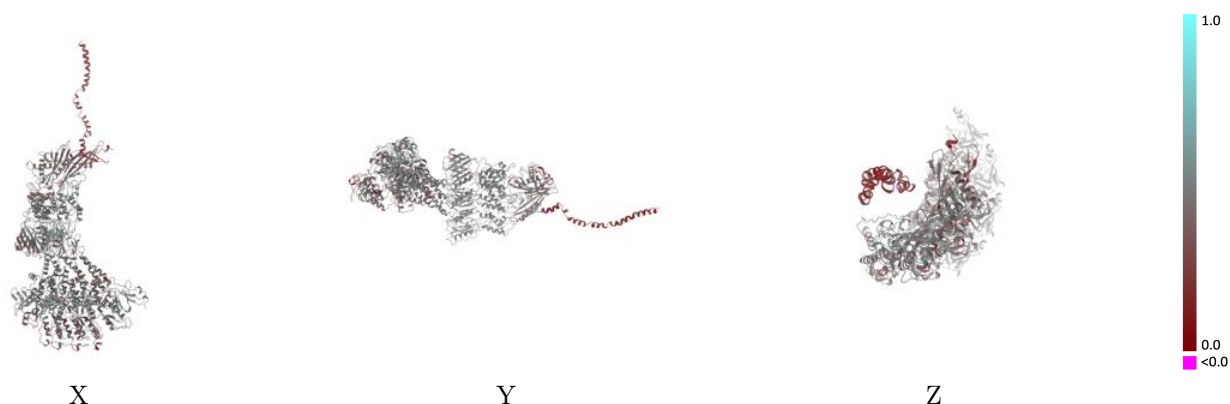


9.1.2 Map-model assembly overlay [i](#)



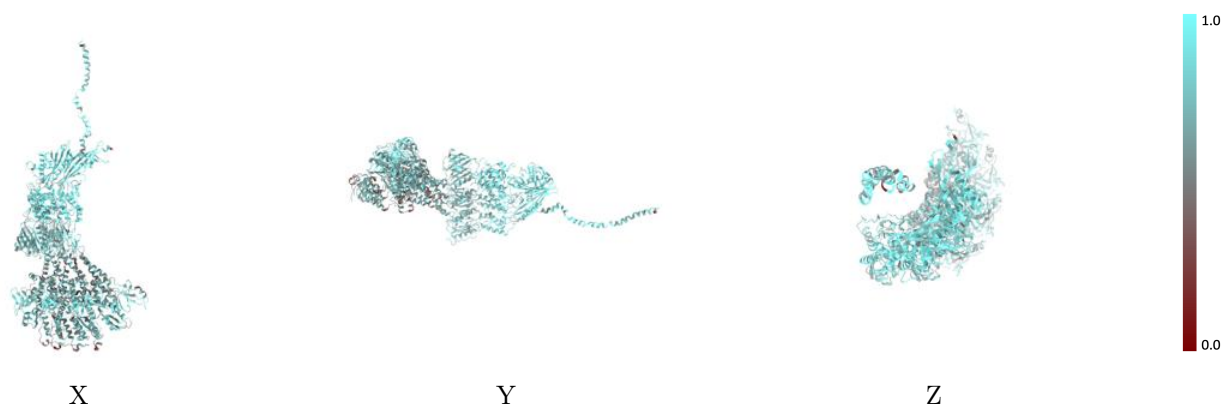
The images above show the 3D surface view of the map at the recommended contour level 3.1 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

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



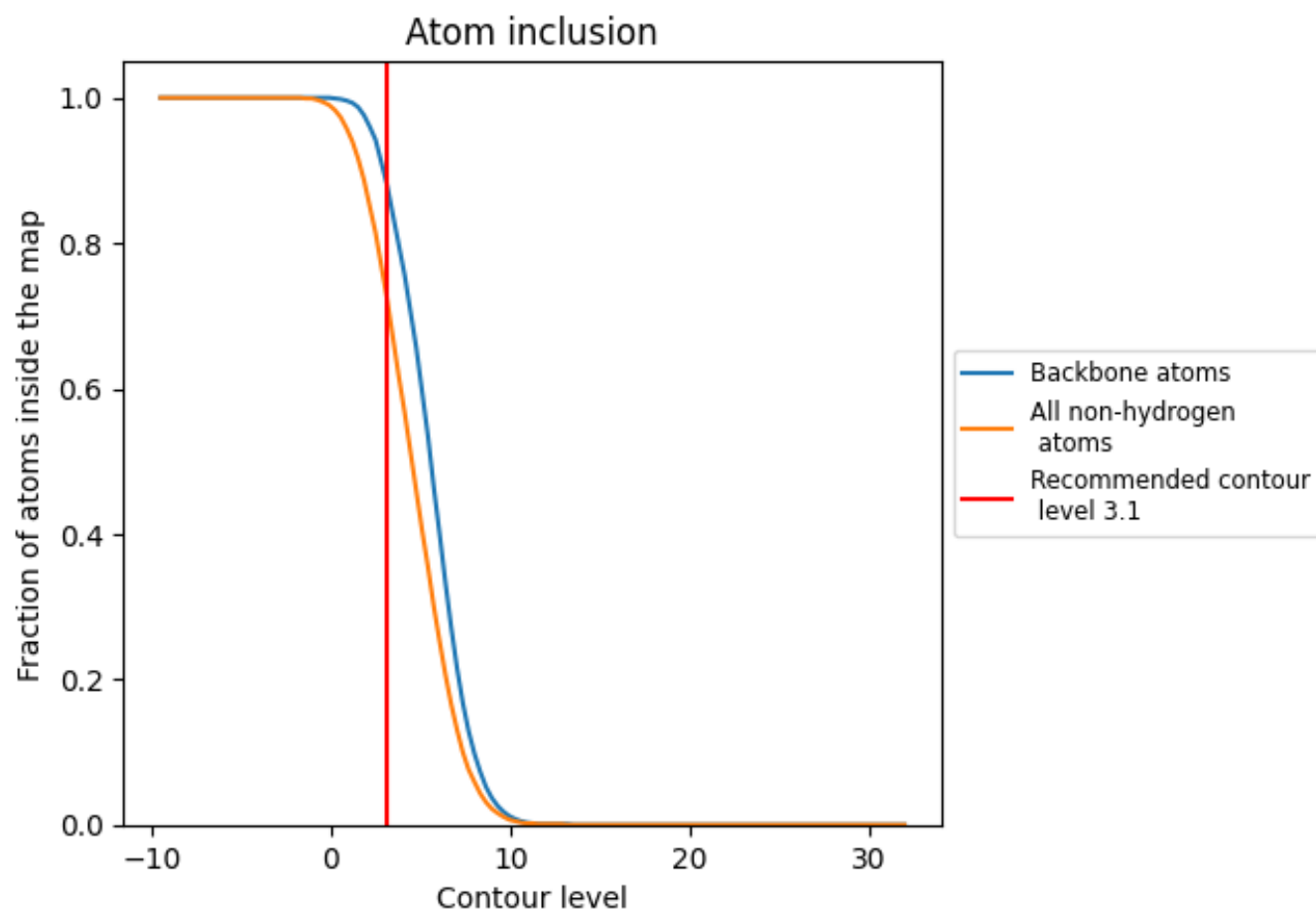
The images above show the model with each residue coloured according to its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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 (3.1).

9.4 Atom inclusion [i](#)



At the recommended contour level, 88% of all backbone atoms, 73% 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 (3.1) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	<div></div> 0.7300	<div></div> 0.4410
B	<div></div> 0.6800	<div></div> 0.4440
C	<div></div> 0.6680	<div></div> 0.4390
D	<div></div> 0.6800	<div></div> 0.4480
E	<div></div> 0.6840	<div></div> 0.4710
P	<div></div> 0.7900	<div></div> 0.4470
Q	<div></div> 0.7970	<div></div> 0.4730
R	<div></div> 0.7930	<div></div> 0.4540
S	<div></div> 0.7830	<div></div> 0.4530
a	<div></div> 0.8630	<div></div> 0.4420
b	<div></div> 0.8280	<div></div> 0.3940
x	<div></div> 0.7430	<div></div> 0.2120

1.0

0.0

<0.0