



wwPDB EM Validation Summary Report ⓘ

Nov 12, 2022 – 09:42 AM EST

PDB ID : 6PIJ
EMDB ID : EMD-20351
Title : Target DNA-bound V. cholerae ThiQ-Cascade complex, closed conformation
Authors : Halpin-Healy, T.; Klompe, S.; Sternberg, S.H.
Deposited on : 2019-06-26
Resolution : 2.90 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

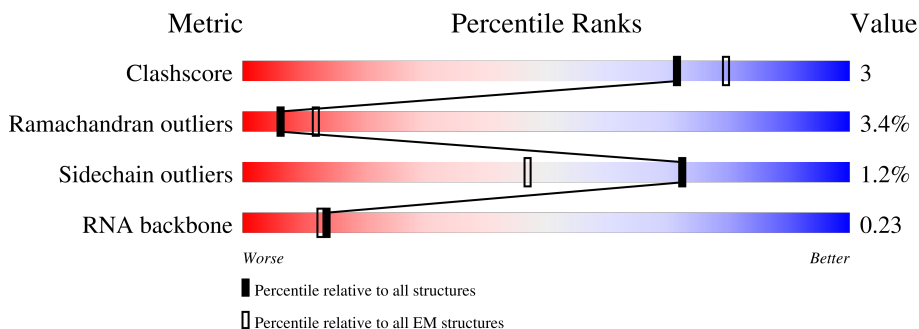
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 2.90 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 ≥ 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	351	<div> <div>6%</div> <div>87%</div> <div>13%</div> <div>.</div> </div>
1	B	351	<div> <div>85%</div> <div>11%</div> <div>.</div> </div>
1	C	351	<div> <div>89%</div> <div>8%</div> <div>.</div> </div>
1	D	351	<div> <div>88%</div> <div>8%</div> <div>.</div> </div>
1	E	351	<div> <div>85%</div> <div>10%</div> <div>.</div> </div>
1	F	351	<div> <div>77%</div> <div>11%</div> <div>12%</div> <div>.</div> </div>
2	G	511	<div> <div>10%</div> <div>85%</div> <div>14%</div> <div>.</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	H	197	
4	I	358	
5	J	369	
6	1	60	
7	2	27	
8	3	9	

2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 29671 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 cas7 type I-F CRISPR-associated protein Csy3.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	351	Total	C	N	O	S	0	0
			2782	1766	479	523	14		
1	B	339	Total	C	N	O	S	0	0
			2714	1727	467	506	14		
1	C	340	Total	C	N	O	S	0	0
			2715	1727	467	507	14		
1	D	338	Total	C	N	O	S	0	0
			2705	1721	465	505	14		
1	E	339	Total	C	N	O	S	0	0
			2710	1724	466	506	14		
1	F	310	Total	C	N	O	S	0	0
			2509	1603	431	461	14		

- Molecule 2 is a protein called cas5_8 naturally occurring fusion protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	G	511	Total	C	N	O	S	0	0
			4009	2533	699	757	20		

- Molecule 3 is a protein called type I-F CRISPR-associated endoribonuclease Cas6/Csy4.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	H	197	Total	C	N	O	S	0	0
			1606	1022	285	292	7		

- Molecule 4 is a protein called TniQ monomer I.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	I	358	Total	C	N	O	S	0	0
			2928	1878	509	526	15		

- Molecule 5 is a protein called TniQ monomer 2.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	J	369	Total	C	N	O	S	0	0
			2983	1912	516	539	16		

- Molecule 6 is a RNA chain called guide RNA.

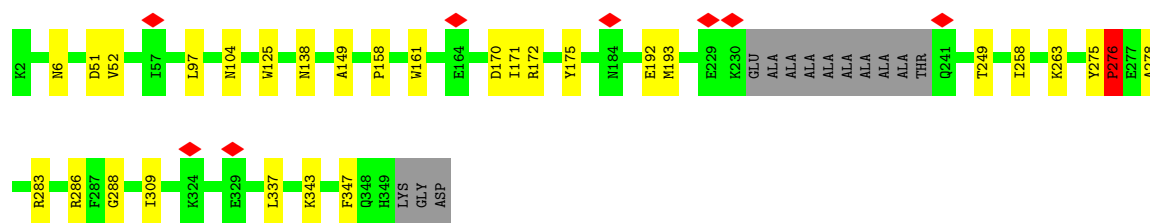
Mol	Chain	Residues	Atoms					AltConf	Trace
6	1	60	Total	C	N	O	P	0	0
			1271	569	219	424	59		

- Molecule 7 is a DNA chain called Targeting strand ssDNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	2	27	Total	C	N	O	P	0	0
			561	264	111	159	27		

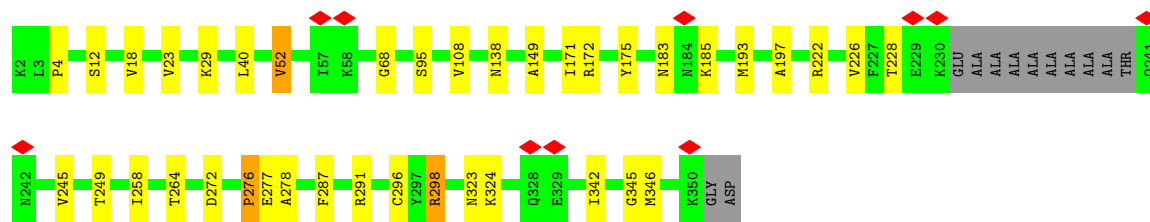
- Molecule 8 is a DNA chain called Non-targeting strand ssDNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	3	9	Total	C	N	O	P	0	0
			178	85	29	55	9		



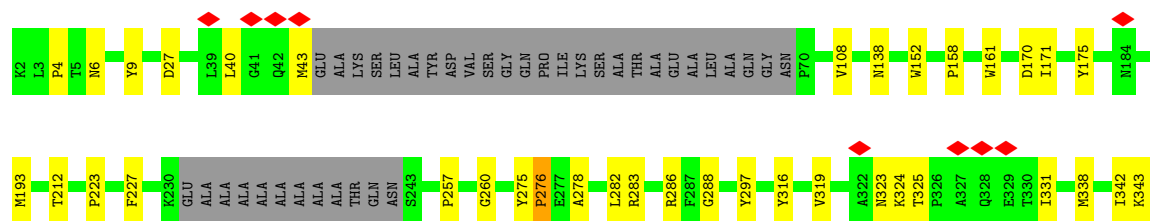
- Molecule 1: cas7 type I-F CRISPR-associated protein Csy3

Chain E: 85% 10% . .



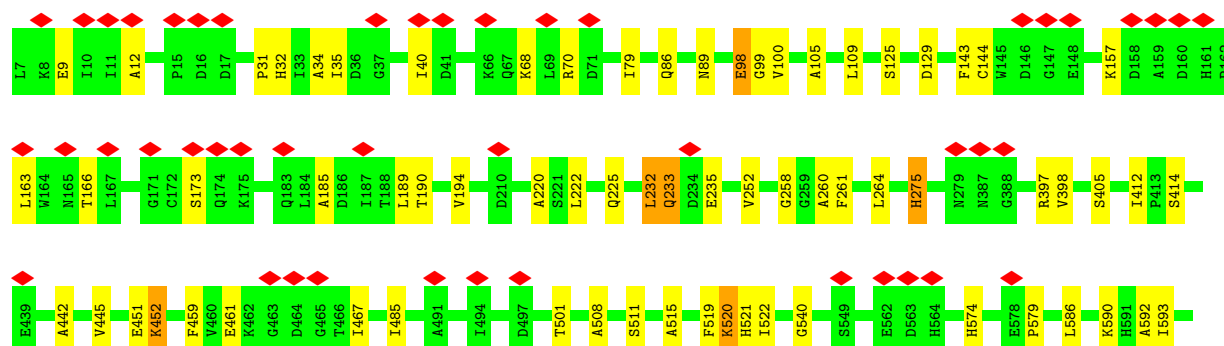
- Molecule 1: cas7 type I-F CRISPR-associated protein Csy3

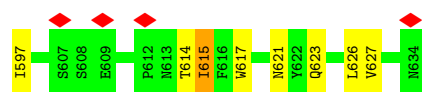
Chain F: 77% 11% 12%



- Molecule 2: cas5_8 naturally occurring fusion protein

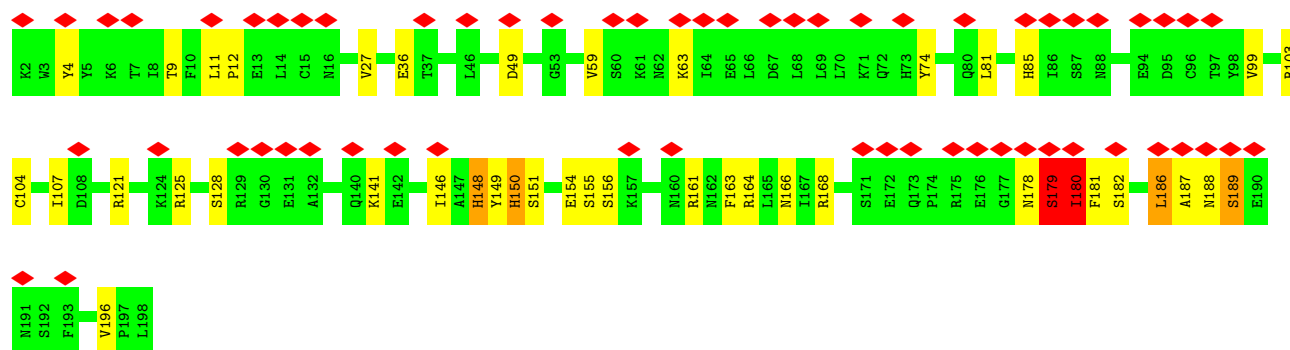
Chain G: 10% 85% 14%





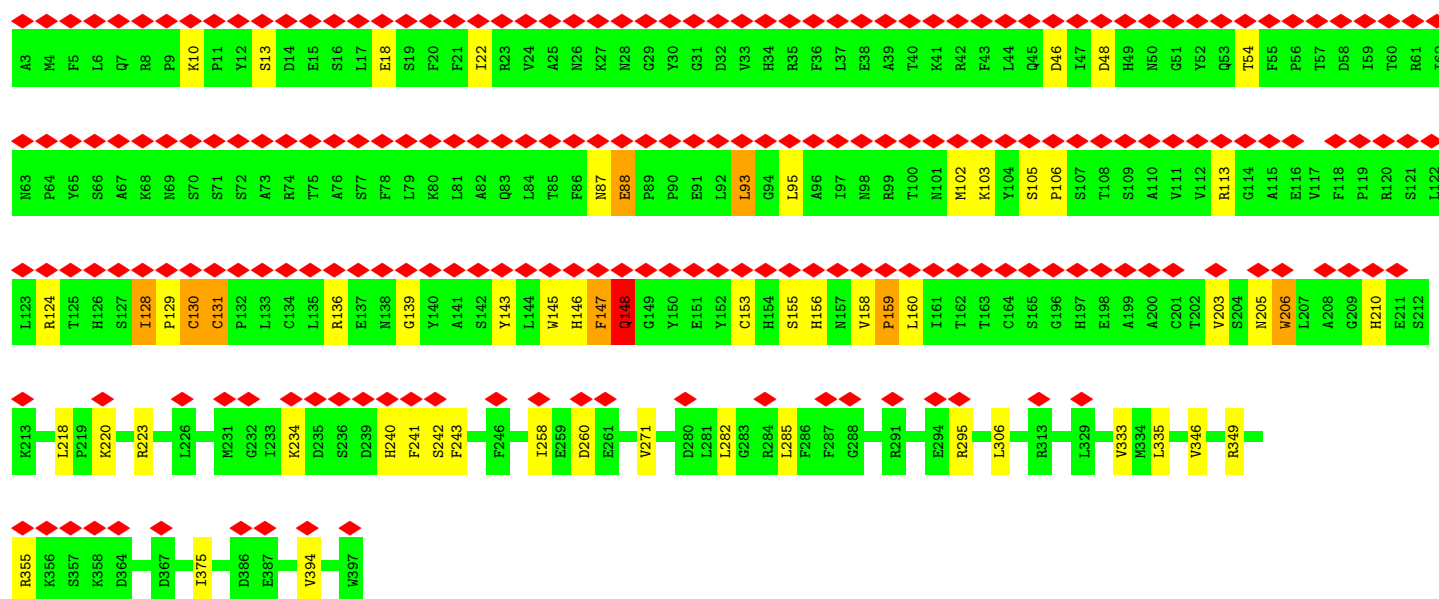
- Molecule 3: type I-F CRISPR-associated endoribonuclease Cas6/Csy4

Chain H: 30% 78% 19% ..



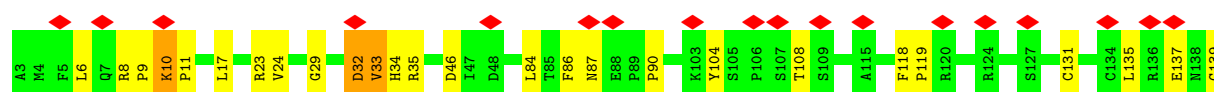
- Molecule 4: ThiQ monomer I

Chain I: 59% 83% 14% .



- Molecule 5: ThiQ monomer 2

Chain J: 57% 85% 14% .



4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	74366	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI POLARA 300	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{\AA}^2$)	50	Depositor
Minimum defocus (nm)	Not provided	
Maximum defocus (nm)	Not provided	
Magnification	Not provided	
Image detector	GATAN K3 (6k x 4k)	Depositor
Maximum map value	0.179	Depositor
Minimum map value	-0.078	Depositor
Average map value	-0.000	Depositor
Map value standard deviation	0.005	Depositor
Recommended contour level	0.025	Depositor
Map size (Å)	424.2, 424.2, 424.2	wwPDB
Map dimensions	400, 400, 400	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.0605, 1.0605, 1.0605	Depositor

5 Model quality

5.1 Standard geometry

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	0.68	0/2854	0.79	0/3880
1	B	0.68	0/2785	0.78	0/3783
1	C	0.68	0/2786	0.78	0/3786
1	D	0.68	0/2776	0.77	0/3772
1	E	0.68	0/2781	0.78	0/3779
1	F	0.68	0/2577	0.79	0/3498
2	G	0.70	0/4098	0.80	0/5563
3	H	0.68	0/1643	0.82	0/2216
4	I	0.69	0/3010	0.82	0/4078
5	J	0.69	0/3065	0.81	0/4157
6	1	0.23	0/1419	0.68	0/2209
7	2	0.29	0/631	0.78	0/973
8	3	0.28	0/197	0.79	0/300
All	All	0.66	0/30622	0.79	0/41994

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
1	A	0	3
1	B	0	4
1	D	0	1
1	E	0	1
2	G	0	9
3	H	0	4
4	I	0	9
5	J	0	5
All	All	0	36

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 36 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	235	ALA	Peptide
1	A	25	PRO	Peptide
1	A	348	GLN	Peptide
1	B	183	ASN	Peptide
1	B	57	ILE	Peptide

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	A	2782	0	2704	19	0
1	B	2714	0	2643	13	0
1	C	2715	0	2637	12	0
1	D	2705	0	2630	11	0
1	E	2710	0	2632	19	0
1	F	2509	0	2447	21	0
2	G	4009	0	3977	30	0
3	H	1606	0	1594	37	0
4	I	2928	0	2839	20	0
5	J	2983	0	2891	20	0
6	1	1271	0	643	13	0
7	2	561	0	302	4	0
8	3	178	0	102	0	0
All	All	29671	0	28041	200	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 200 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)
3:H:156:SER:N	3:H:163:PHE:HE1	1.60	0.98
3:H:155:SER:C	3:H:163:PHE:CE1	2.43	0.92
3:H:156:SER:N	3:H:163:PHE:CE1	2.42	0.87

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:G:574:HIS:HB2	2:G:597:ILE:HD11	1.60	0.83
3:H:151:SER:HB2	3:H:164:ARG:HG2	1.72	0.72

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	349/351 (99%)	305 (87%)	35 (10%)	9 (3%)	5	20
1	B	335/351 (95%)	293 (88%)	31 (9%)	11 (3%)	4	15
1	C	336/351 (96%)	287 (85%)	41 (12%)	8 (2%)	6	22
1	D	334/351 (95%)	297 (89%)	32 (10%)	5 (2%)	10	34
1	E	335/351 (95%)	293 (88%)	36 (11%)	6 (2%)	8	29
1	F	304/351 (87%)	265 (87%)	30 (10%)	9 (3%)	4	17
2	G	505/511 (99%)	389 (77%)	89 (18%)	27 (5%)	2	6
3	H	195/197 (99%)	148 (76%)	39 (20%)	8 (4%)	3	11
4	I	350/358 (98%)	254 (73%)	77 (22%)	19 (5%)	2	6
5	J	363/369 (98%)	274 (76%)	74 (20%)	15 (4%)	3	11
All	All	3406/3541 (96%)	2805 (82%)	484 (14%)	117 (3%)	6	15

5 of 117 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	17	ASP
1	A	26	ASP
1	A	349	HIS
1	B	53	SER
1	B	57	ILE

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	A	300/301 (100%)	296 (99%)	4 (1%)	69	90
1	B	297/301 (99%)	293 (99%)	4 (1%)	69	90
1	C	296/301 (98%)	294 (99%)	2 (1%)	84	95
1	D	296/301 (98%)	291 (98%)	5 (2%)	60	86
1	E	296/301 (98%)	293 (99%)	3 (1%)	76	92
1	F	277/301 (92%)	273 (99%)	4 (1%)	67	89
2	G	451/451 (100%)	449 (100%)	2 (0%)	91	97
3	H	177/177 (100%)	175 (99%)	2 (1%)	73	92
4	I	322/322 (100%)	315 (98%)	7 (2%)	52	81
5	J	323/323 (100%)	319 (99%)	4 (1%)	71	91
All	All	3035/3079 (99%)	2998 (99%)	37 (1%)	72	91

5 of 37 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	I	130	CYS
5	J	131	CYS
4	I	220	LYS
4	I	355	ARG
1	D	283	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 29 such sidechains are listed below:

Mol	Chain	Res	Type
1	E	153	ASN
5	J	157	ASN
1	F	349	HIS
4	I	210	HIS
1	F	183	ASN

5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
6	1	59/60 (98%)	27 (45%)	2 (3%)

5 of 27 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
6	1	3	G
6	1	8	C
6	1	9	U
6	1	10	U
6	1	14	G

All (2) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
6	1	28	C
6	1	45	A

5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

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

5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

5.6 Ligand geometry [i](#)

There are no ligands in this entry.

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
4	I	3
5	J	2
2	G	2

The worst 5 of 7 chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	J	175:SER	C	198:GLU	N	32.21
1	I	165:SER	C	196:GLY	N	27.25
1	G	279:ASN	C	387:ASN	N	23.34
1	G	52:LYS	C	63:ALA	N	13.43
1	I	358:LYS	C	364:ASP	N	11.88

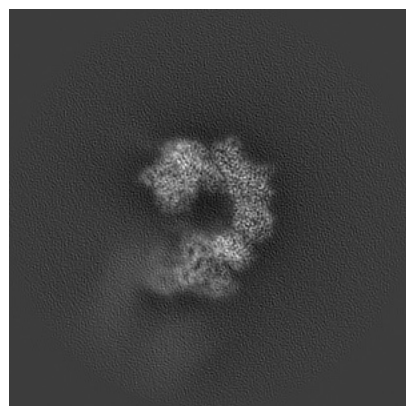
6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-20351. These allow visual inspection of the internal detail of the map and identification of artifacts.

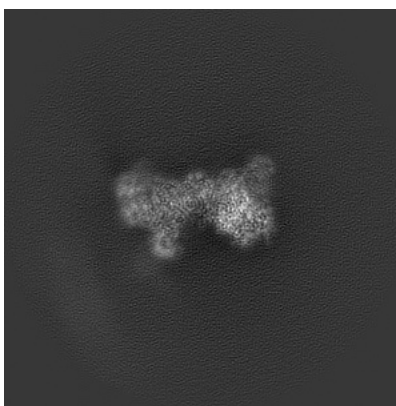
Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

6.1 Orthogonal projections [i](#)

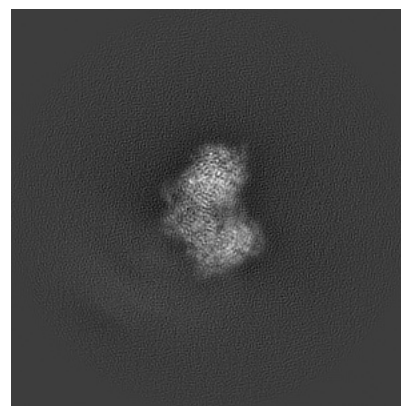
6.1.1 Primary map



X

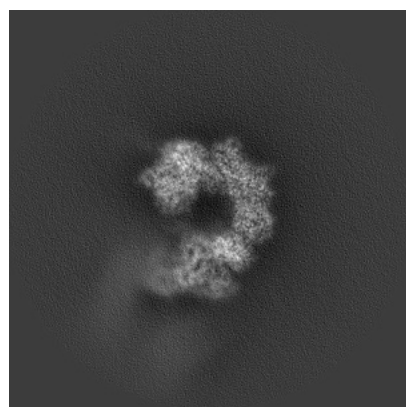


Y

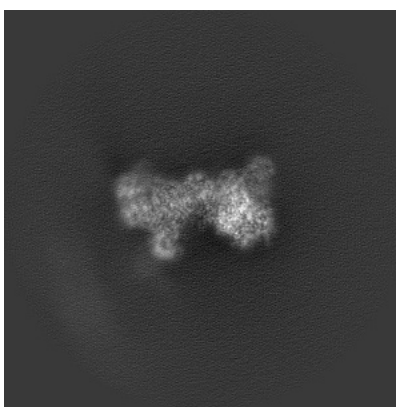


Z

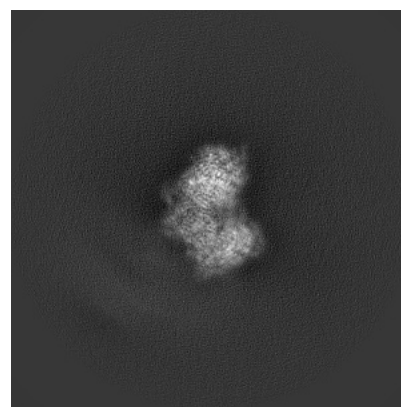
6.1.2 Raw map



X



Y

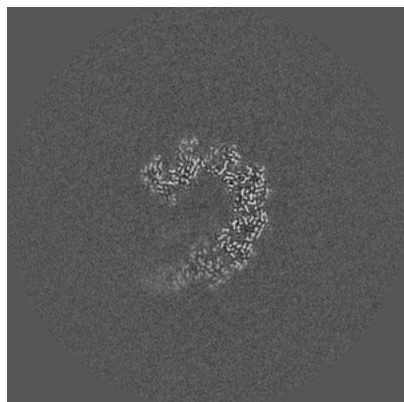


Z

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

6.2 Central slices [i](#)

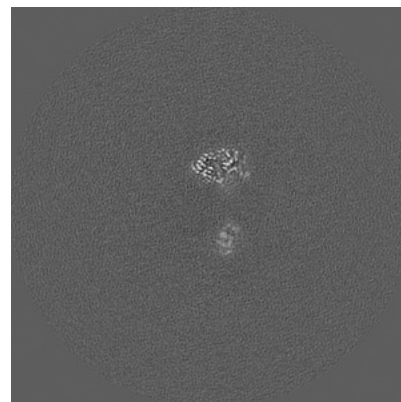
6.2.1 Primary map



X Index: 200

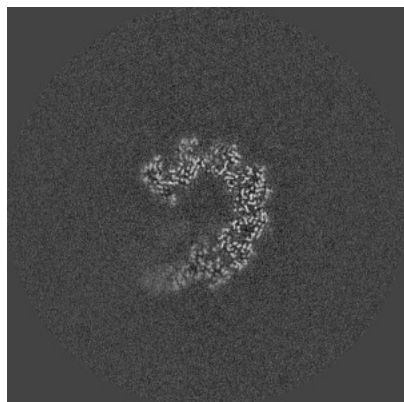


Y Index: 200



Z Index: 200

6.2.2 Raw map



X Index: 200



Y Index: 200

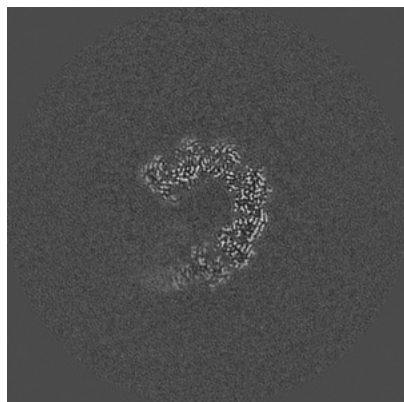


Z Index: 200

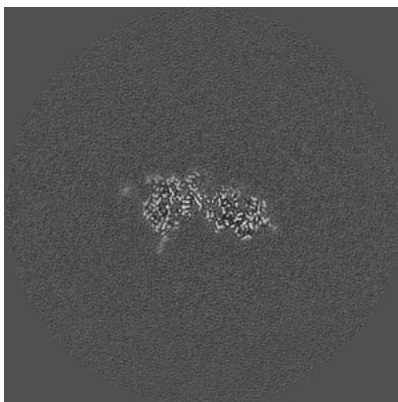
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices [i](#)

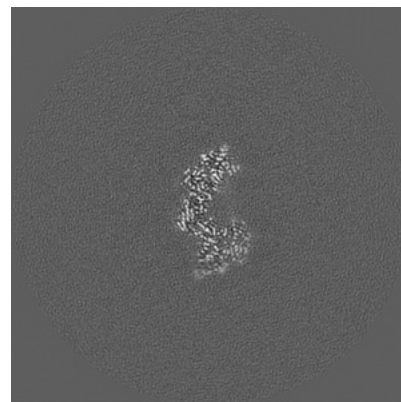
6.3.1 Primary map



X Index: 199

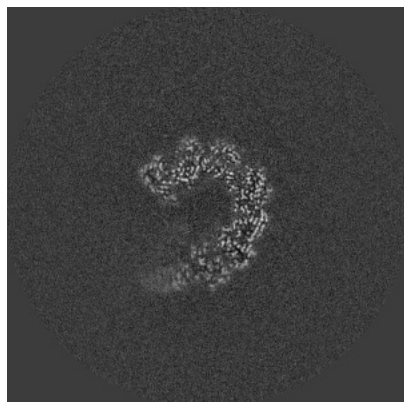


Y Index: 227

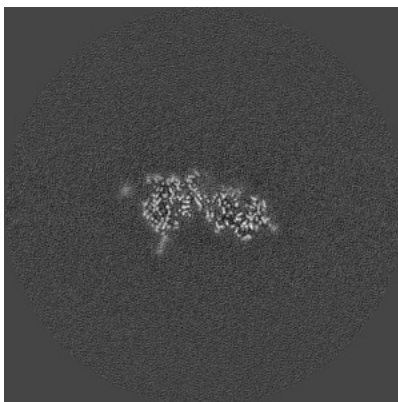


Z Index: 231

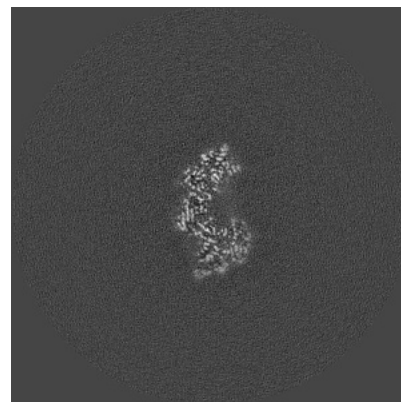
6.3.2 Raw map



X Index: 199



Y Index: 227



Z Index: 231

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 0.025. 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



X



Y



Z

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.

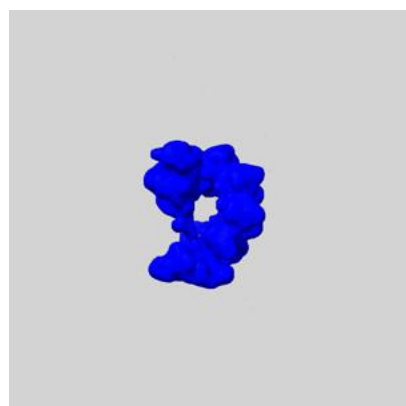
6.5 Mask visualisation [i](#)

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

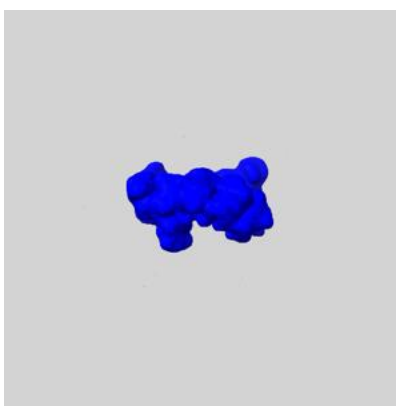
A mask typically either:

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

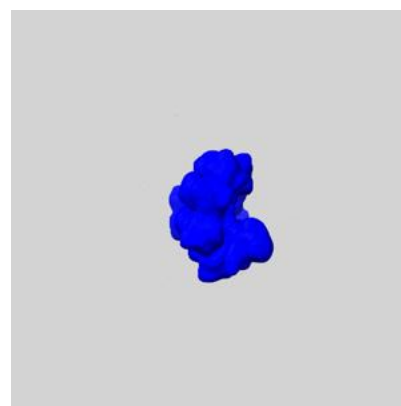
6.5.1 emd_20351_msk_1.map [i](#)



X



Y

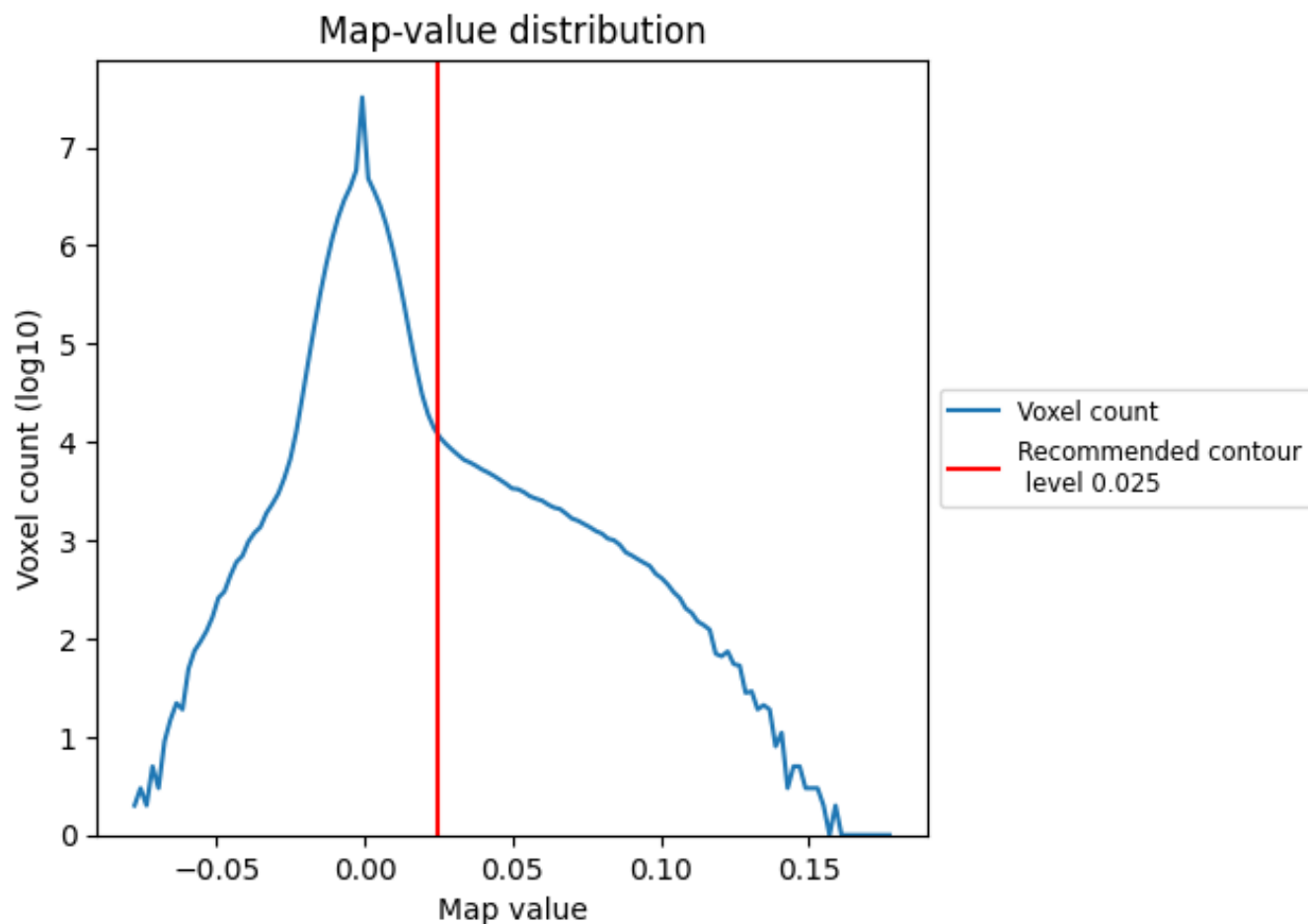


Z

7 Map analysis [i](#)

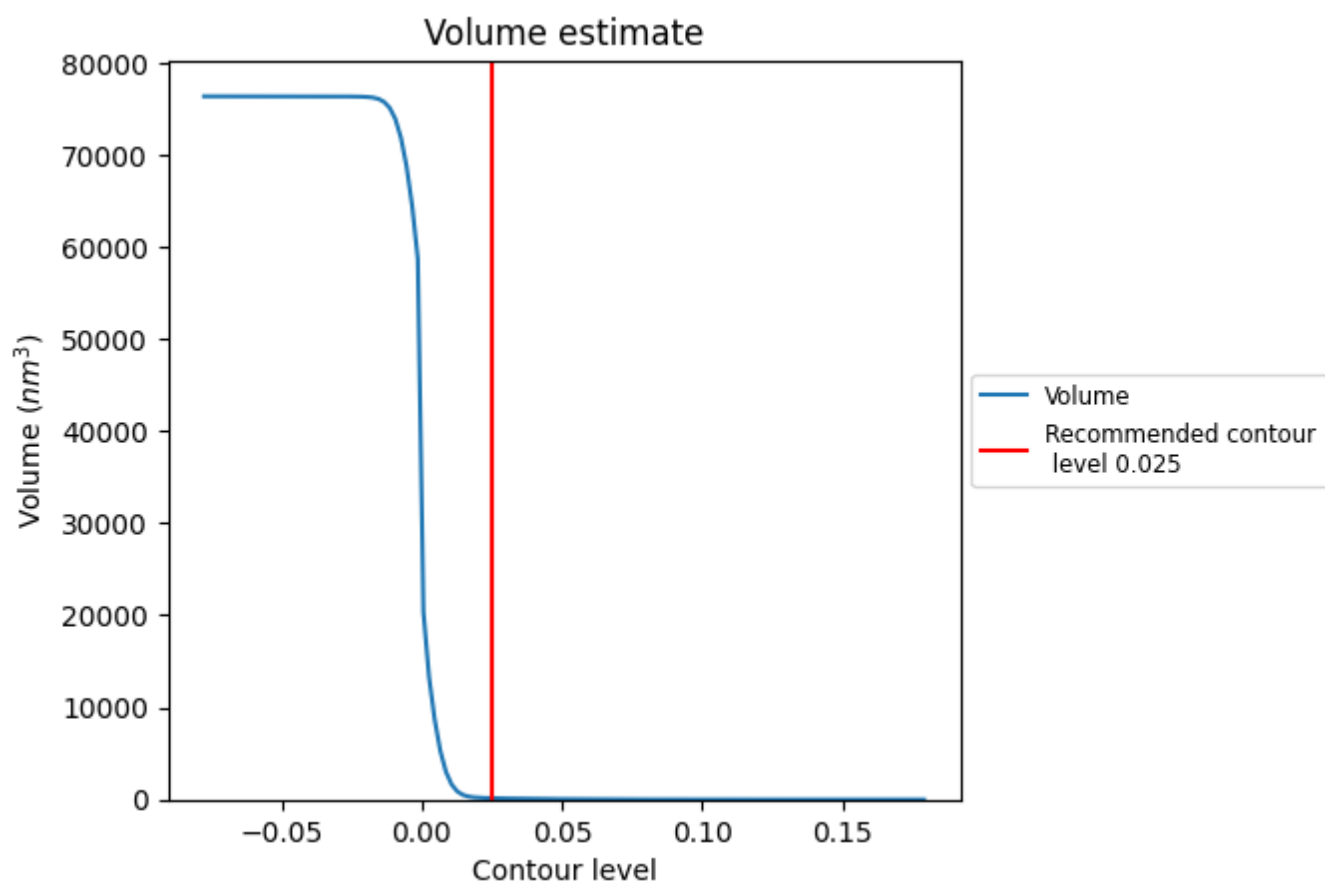
This section contains the results of statistical analysis of the map.

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.

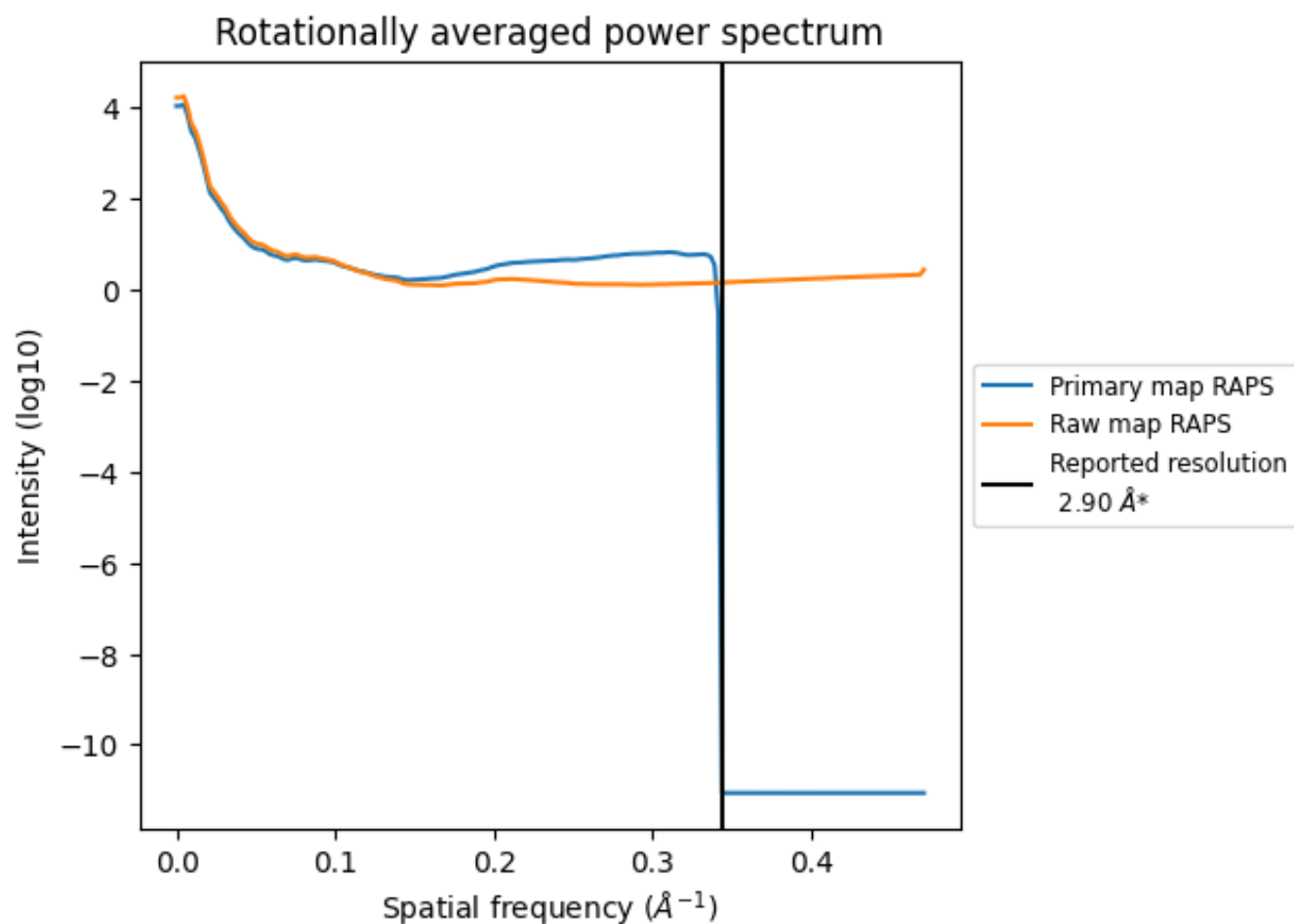
7.2 Volume estimate [i](#)



The volume at the recommended contour level is 149 nm³; this corresponds to an approximate mass of 134 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 ⓘ

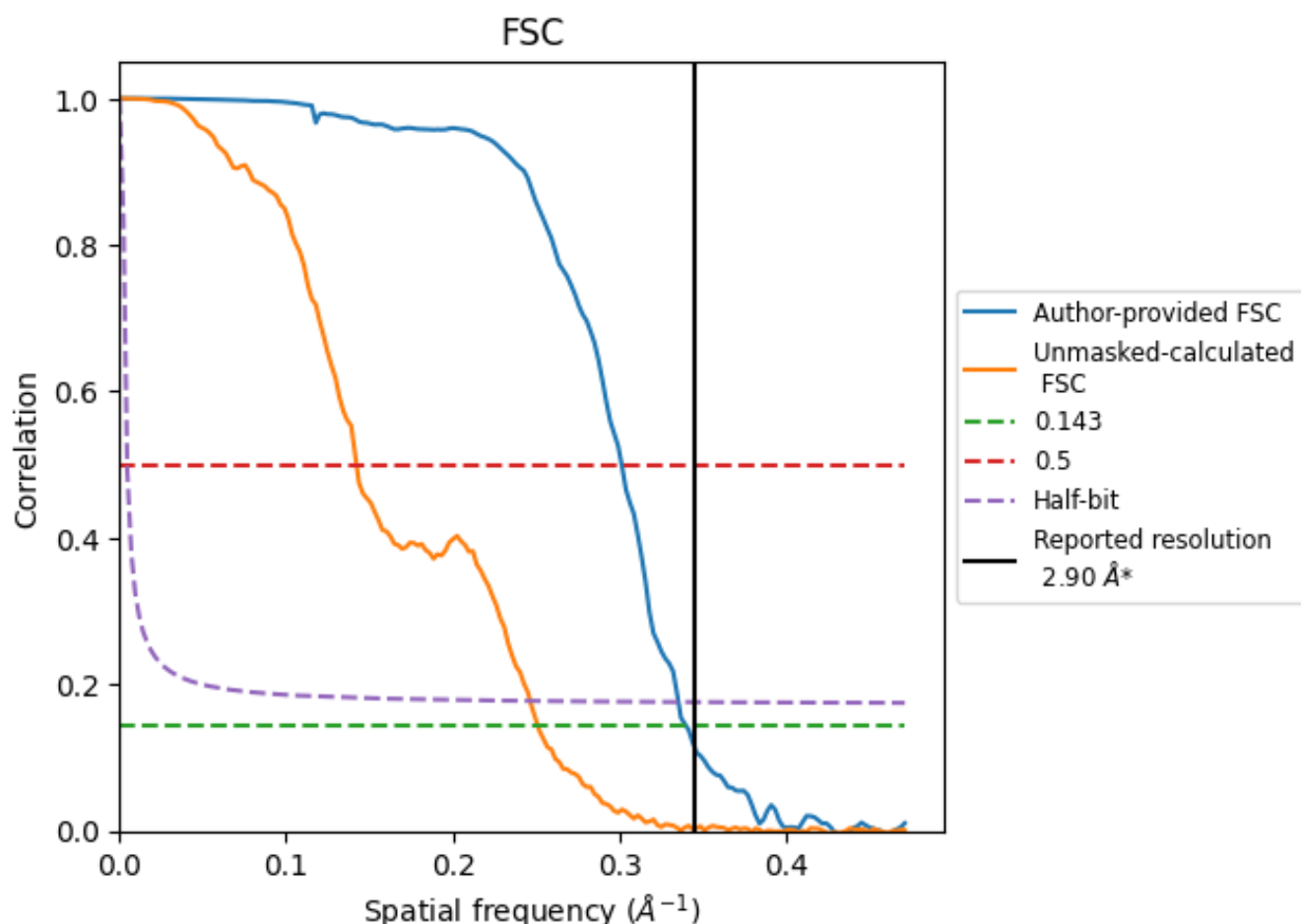


*Reported resolution corresponds to spatial frequency of 0.345 Å⁻¹

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 0.345 \AA^{-1}

8.2 Resolution estimates [i](#)

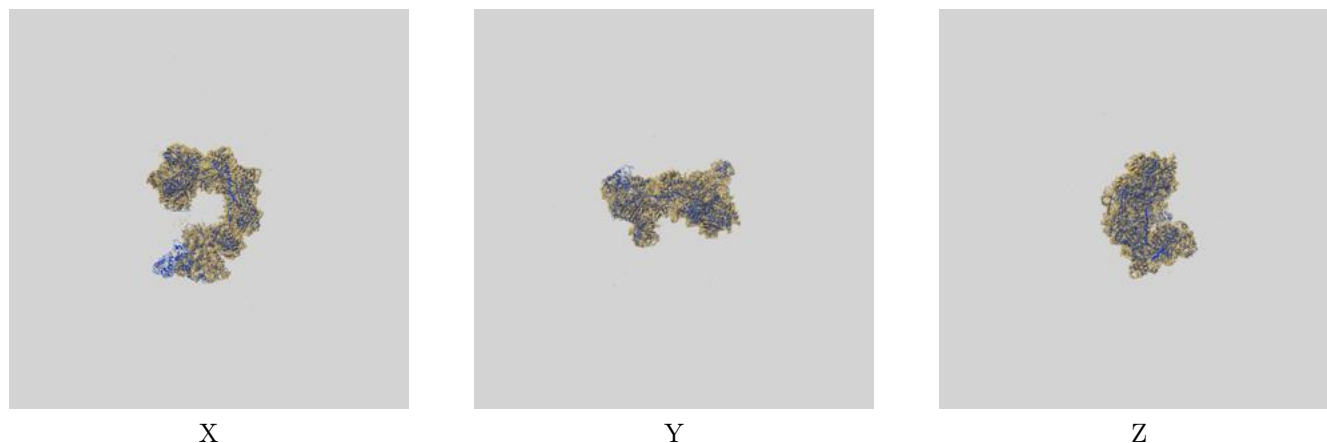
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	2.90	-	-
Author-provided FSC curve	2.94	3.32	2.98
Unmasked-calculated*	3.98	7.03	4.06

*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 3.98 differs from the reported value 2.9 by more than 10 %

9 Map-model fit [i](#)

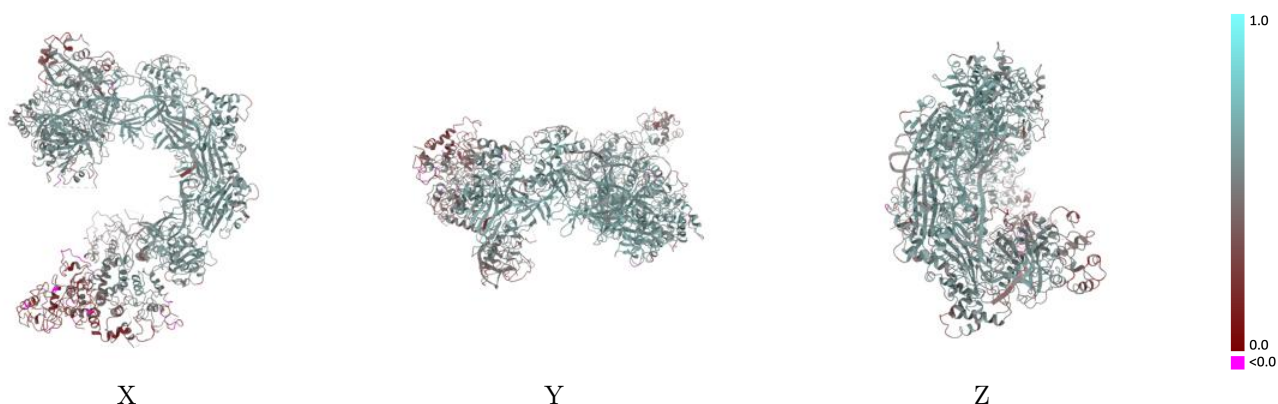
This section contains information regarding the fit between EMDB map EMD-20351 and PDB model 6PIJ. Per-residue inclusion information can be found in [section 3](#) on [page 6](#).

9.1 Map-model overlay [i](#)



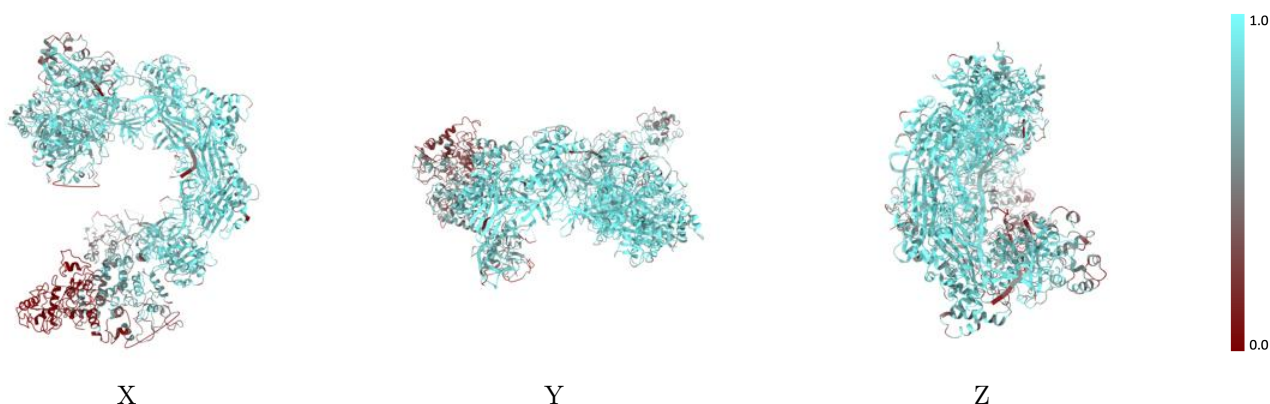
The images above show the 3D surface view of the map at the recommended contour level 0.025 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

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



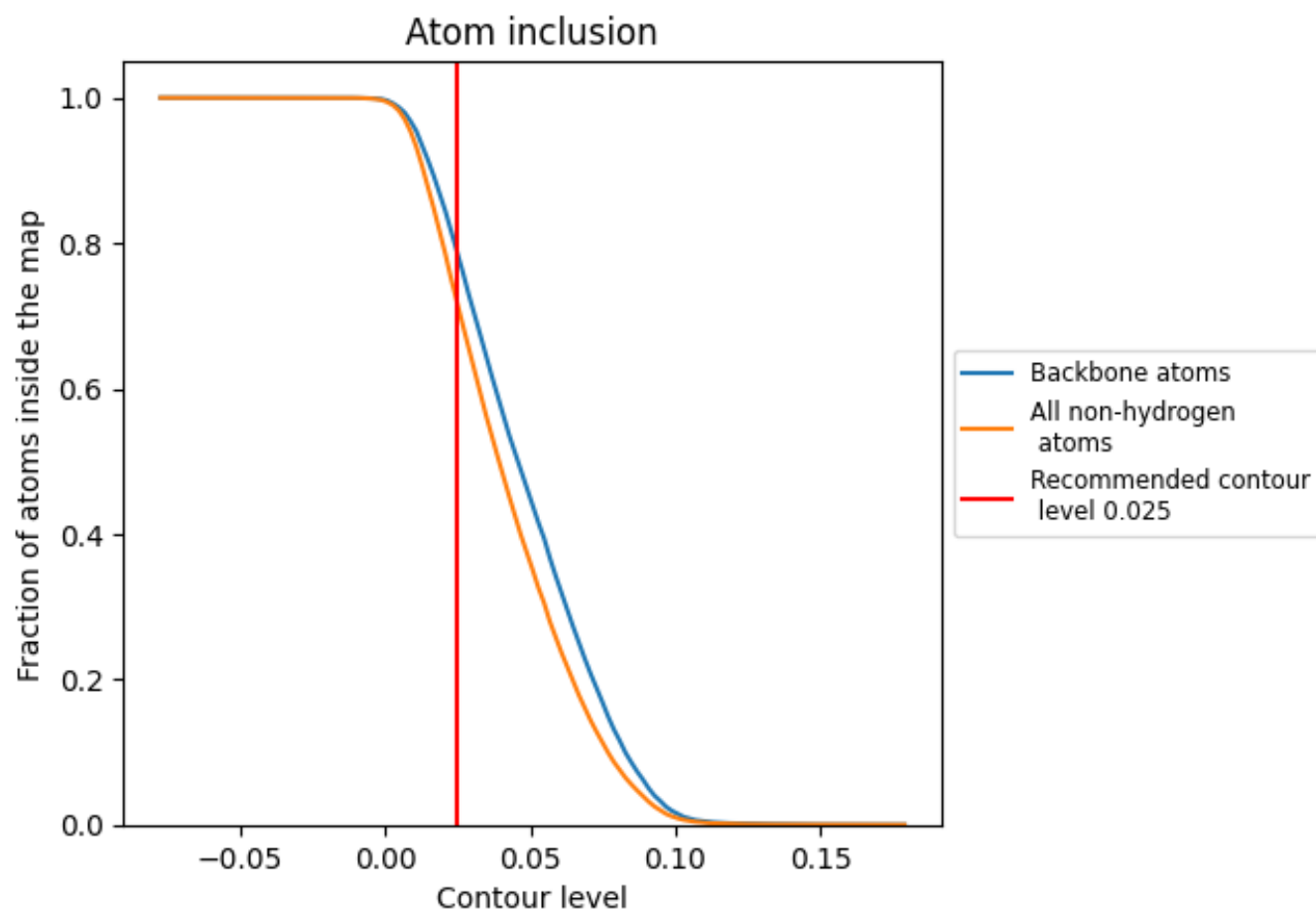
The images above show the model with each residue coloured according its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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 (0.025).

9.4 Atom inclusion [i](#)



At the recommended contour level, 78% of all backbone atoms, 71% 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 (0.025) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	<div></div> 0.7149	<div></div> 0.5080
1	<div></div> 0.8899	<div></div> 0.5490
2	<div></div> 0.7201	<div></div> 0.5360
3	<div></div> 0.5169	<div></div> 0.4700
A	<div></div> 0.8062	<div></div> 0.5430
B	<div></div> 0.8688	<div></div> 0.5720
C	<div></div> 0.8727	<div></div> 0.5800
D	<div></div> 0.8684	<div></div> 0.5780
E	<div></div> 0.8528	<div></div> 0.5690
F	<div></div> 0.8449	<div></div> 0.5680
G	<div></div> 0.7203	<div></div> 0.5000
H	<div></div> 0.5386	<div></div> 0.4210
I	<div></div> 0.3435	<div></div> 0.3650
J	<div></div> 0.3572	<div></div> 0.3580

1.0

0.0

<0.0