



wwPDB EM Validation Summary Report ⓘ

Oct 28, 2024 – 07:59 PM EDT

PDB ID : 8FWM
EMDB ID : EMD-29512
Title : Structure of tail-neck junction of Agrobacterium phage Milano
Authors : Sonani, R.R.; Wang, F.; Esteves, N.C.; Kelly, R.J.; Sebastian, A.; Kreutzberger, M.A.B.; Leiman, P.G.; Scharf, B.E.; Egelman, E.H.
Deposited on : 2023-01-23
Resolution : 3.49 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) : 2.39

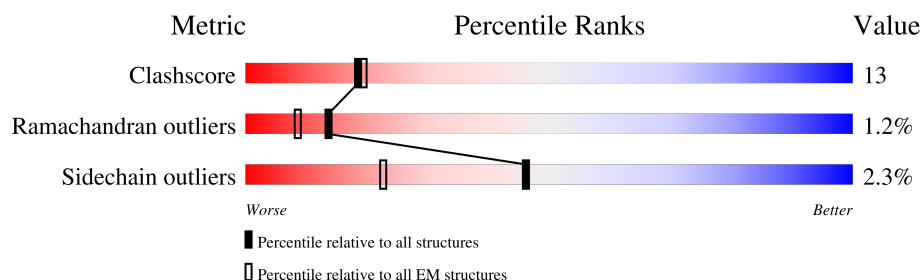
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 3.49 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.





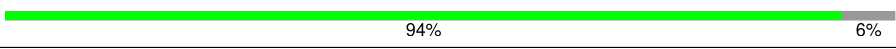
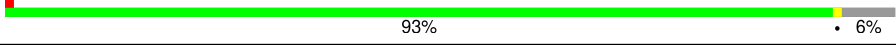
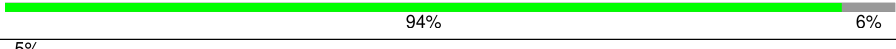

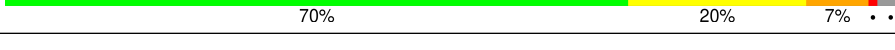
Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	210492	15764
Ramachandran outliers	207382	16835
Sidechain outliers	206894	16415

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	3	230	
1	D	230	
1	E	230	
1	F	230	
1	G	230	
2	p	503	
2	q	503	
2	r	503	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	s	503	
3	AV	136	
3	Aa	136	
3	Ad	136	
3	Ae	136	
4	AS	178	
4	AT	178	

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 26669 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 Collar sheath protein, gp13.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	3	230	Total	C	N	O	S	0	0
			1723	1090	287	337	9		
1	D	230	Total	C	N	O	S	0	0
			1723	1090	287	337	9		
1	E	230	Total	C	N	O	S	0	0
			1723	1090	287	337	9		
1	F	225	Total	C	N	O	S	0	0
			1690	1072	281	328	9		
1	G	230	Total	C	N	O	S	0	0
			1723	1090	287	337	9		

- Molecule 2 is a protein called Tail sheath protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	p	369	Total	C	N	O	S	0	0
			2842	1785	487	551	19		
2	q	373	Total	C	N	O	S	0	0
			2871	1804	491	557	19		
2	r	369	Total	C	N	O	S	0	0
			2842	1785	487	551	19		
2	s	372	Total	C	N	O	S	0	0
			2863	1800	490	554	19		

- Molecule 3 is a protein called Tail-tube, gp21.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	AV	128	Total	C	N	O	S	0	0
			977	606	162	202	7		
3	Aa	128	Total	C	N	O	S	0	0
			977	606	162	202	7		
3	Ae	128	Total	C	N	O	S	0	0
			977	606	162	202	7		
3	Ad	128	Total	C	N	O	S	0	0
			977	606	162	202	7		

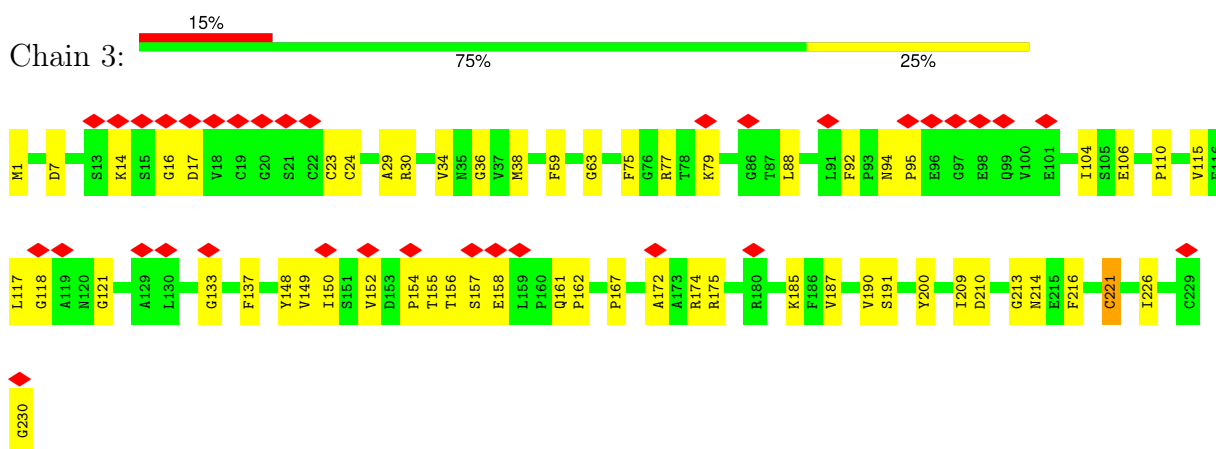
- Molecule 4 is a protein called Tail-terminator protein, gp18.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	AS	171	Total	C	N	O	S	0	0
			1376	874	230	269	3		
4	AT	172	Total	C	N	O	S	0	0
			1385	879	232	271	3		

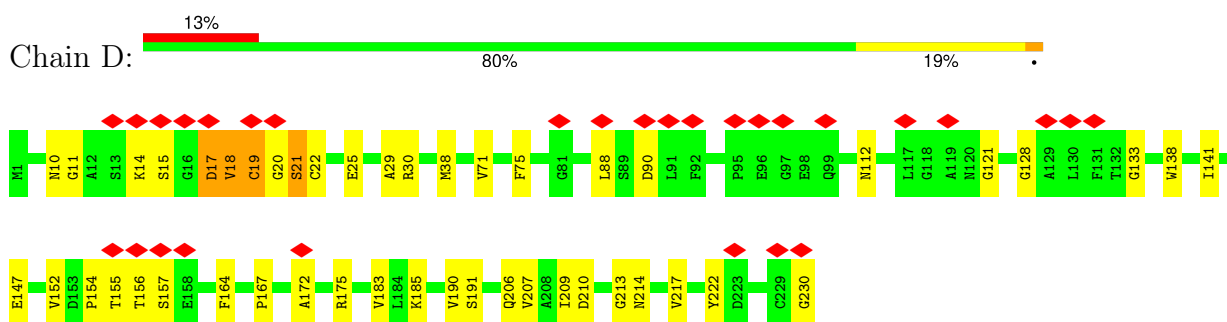
3 Residue-property plots

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.

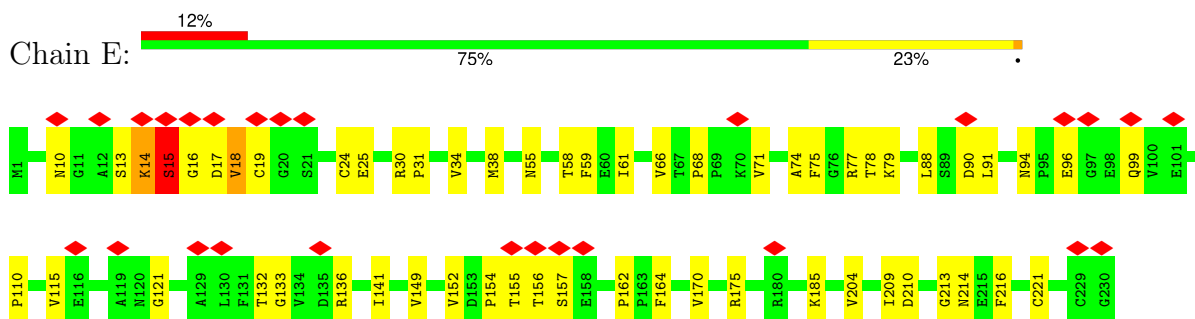
- Molecule 1: Collar sheath protein, gp13



- Molecule 1: Collar sheath protein, gp13



- Molecule 1: Collar sheath protein, gp13



Chain F:

10% 80% 18%

1 2 9 12 13 14 15 16 17 18 19 20 21 22 23 37 38 50 55 59 62 68 69 70 79 88 89 90 95 96 97 98 99 107 117 118 119 120 121 128 129 130 133 136 137 138 147 148 149 152 153 154 155 156 157 158 159 166 167 168 169 172 175 185 191 198 201 202 209 210 211 212 213 221 229 230

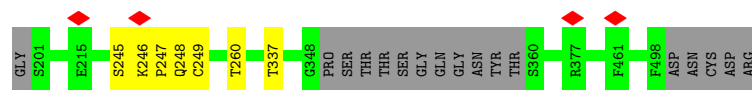
Chain G:

13% 79% 20%

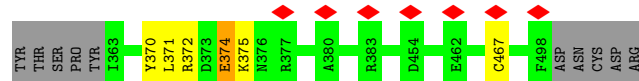
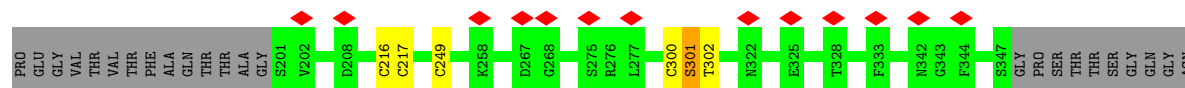
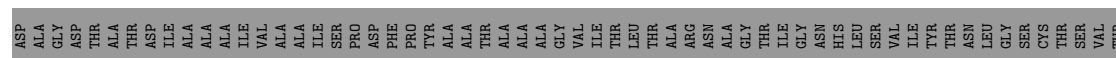
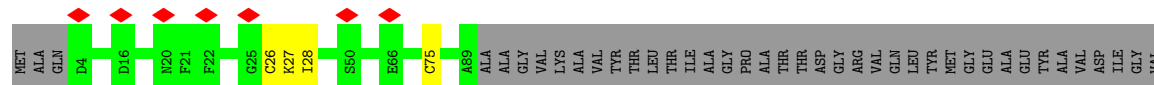
W138 V149 I150 S151 P154 T155 T156 S157 E158 L159 P160 P171 A172 R175 R180 K185 F186 V187 I209 D210 G213 N214 E215 F216 V217 Y222 C229 G230

R1 R9 S13 K14 S15 G16 D17 V18 C19 G20 S21 C22 C23 C24 E25 R30 M38 E60 A74 F75 K79 V80 E85 L88 S89 D90 L91 P95 E98 Q99 V100 E101 M108 G118 A119 M120 G121 A129 L130 F131 T132 G133 V134 D135 R136

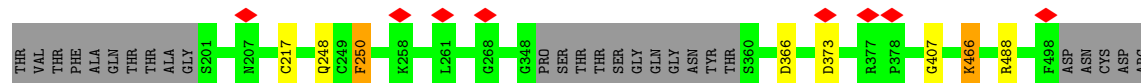
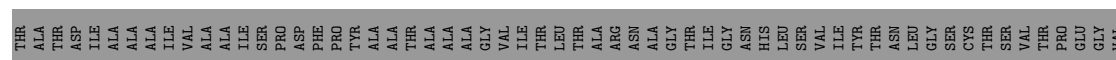
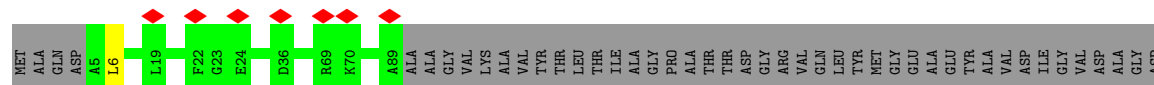
[illegible][illegible]



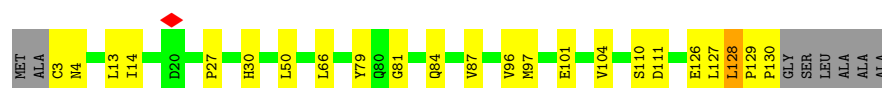
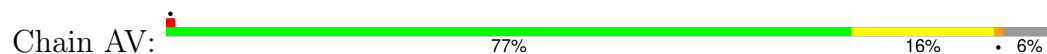
- Molecule 2: Tail sheath protein



- Molecule 2: Tail sheath protein



- Molecule 3: Tail-tube, gp21



- Molecule 3: Tail-tube, gp21



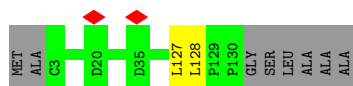
- Molecule 3: Tail-tube, gp21

Chain Ae:  94% 6%



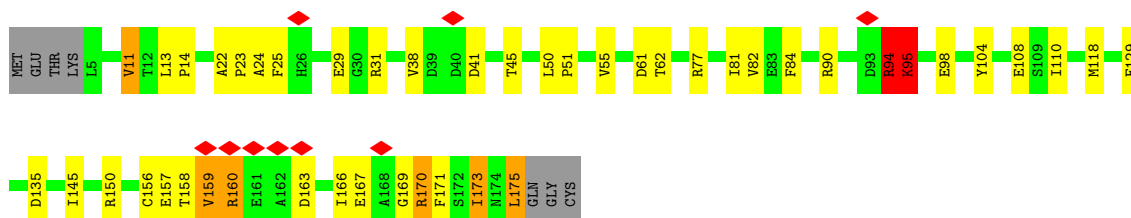
- Molecule 3: Tail-tube, gp21

Chain Ad:  93% 6%



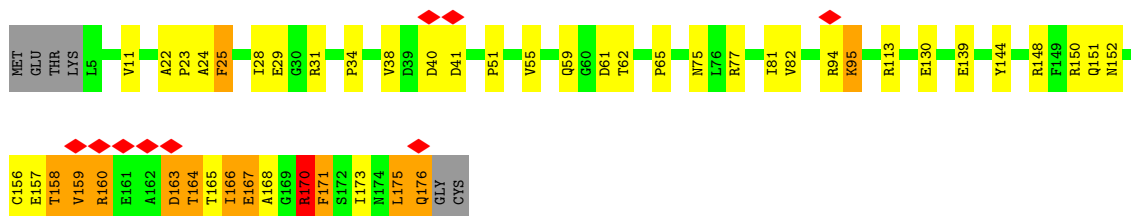
- Molecule 4: Tail-terminator protein, gp18

Chain AS:  5% 70% 21% . . .



- Molecule 4: Tail-terminator protein, gp18

Chain AT:  5% 70% 20% 7% . . .



4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C3	Depositor
Number of particles used	10216	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	TFS KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{\AA}^2$)	50	Depositor
Minimum defocus (nm)	1200	Depositor
Maximum defocus (nm)	2200	Depositor
Magnification	Not provided	
Image detector	GATAN K3 (6k x 4k)	Depositor
Maximum map value	0.923	Depositor
Minimum map value	-0.527	Depositor
Average map value	-0.000	Depositor
Map value standard deviation	0.043	Depositor
Recommended contour level	0.15	Depositor
Map size (Å)	483.84003, 483.84003, 483.84003	wwPDB
Map dimensions	448, 448, 448	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.08, 1.08, 1.08	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	3	0.30	0/1768	0.49	0/2414
1	D	0.29	0/1768	0.47	0/2414
1	E	0.29	0/1768	0.48	0/2414
1	F	0.27	0/1734	0.46	0/2368
1	G	0.27	0/1768	0.47	0/2414
2	p	0.29	0/2900	0.50	0/3946
2	q	0.30	0/2931	0.50	0/3989
2	r	0.29	0/2900	0.51	0/3946
2	s	0.27	0/2923	0.51	1/3978 (0.0%)
3	AV	0.31	0/993	0.50	0/1358
3	Aa	0.29	0/993	0.49	0/1358
3	Ad	0.30	0/993	0.50	0/1358
3	Ae	0.29	0/993	0.49	0/1358
4	AS	0.47	0/1407	0.57	1/1904 (0.1%)
4	AT	0.47	0/1416	0.51	0/1916
All	All	0.31	0/27255	0.50	2/37135 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	s	6	LEU	CA-CB-CG	5.58	128.14	115.30
4	AS	163	ASP	CB-CG-OD2	5.12	122.91	118.30

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	3	1723	0	1654	35	0
1	D	1723	0	1654	30	0
1	E	1723	0	1653	38	0
1	F	1690	0	1623	30	0
1	G	1723	0	1654	35	0
2	p	2842	0	2743	0	0
2	q	2871	0	2768	0	0
2	r	2842	0	2743	0	0
2	s	2863	0	2763	0	0
3	AV	977	0	950	12	0
3	Aa	977	0	949	0	0
3	Ad	977	0	950	0	0
3	Ae	977	0	949	0	0
4	AS	1376	0	1311	30	0
4	AT	1385	0	1319	29	0
All	All	26669	0	25683	228	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 13.

The worst 5 of 228 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:AV:30:HIS:HE2	3:AV:79:TYR:HH	1.12	0.93
4:AS:94:ARG:HH21	4:AS:94:ARG:H	1.22	0.88
1:G:38:MET:HG2	1:G:185:LYS:HG2	1.63	0.78
1:F:79:LYS:HD3	1:F:153:ASP:HB2	1.66	0.76
4:AS:25:PHE:HZ	4:AS:110:ILE:HA	1.50	0.76

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	3	228/230 (99%)	205 (90%)	21 (9%)	2 (1%)	14	49
1	D	228/230 (99%)	201 (88%)	23 (10%)	4 (2%)	7	35
1	E	228/230 (99%)	206 (90%)	19 (8%)	3 (1%)	10	41
1	F	221/230 (96%)	201 (91%)	20 (9%)	0	100	100
1	G	228/230 (99%)	204 (90%)	23 (10%)	1 (0%)	30	64
2	p	363/503 (72%)	288 (79%)	69 (19%)	6 (2%)	7	36
2	q	367/503 (73%)	292 (80%)	71 (19%)	4 (1%)	12	45
2	r	363/503 (72%)	282 (78%)	76 (21%)	5 (1%)	9	40
2	s	366/503 (73%)	307 (84%)	56 (15%)	3 (1%)	16	51
3	AV	126/136 (93%)	115 (91%)	11 (9%)	0	100	100
3	Aa	126/136 (93%)	115 (91%)	11 (9%)	0	100	100
3	Ad	126/136 (93%)	112 (89%)	14 (11%)	0	100	100
3	Ae	126/136 (93%)	117 (93%)	9 (7%)	0	100	100
4	AS	169/178 (95%)	135 (80%)	28 (17%)	6 (4%)	3	22
4	AT	170/178 (96%)	142 (84%)	22 (13%)	6 (4%)	3	24
All	All	3435/4062 (85%)	2922 (85%)	473 (14%)	40 (1%)	14	43

5 of 40 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	p	55	ILE
2	q	247	PRO
2	s	466	LYS
4	AS	95	LYS
4	AT	164	THR

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	3	191/191 (100%)	188 (98%)	3 (2%)	58	76
1	D	191/191 (100%)	188 (98%)	3 (2%)	58	76
1	E	191/191 (100%)	187 (98%)	4 (2%)	48	71
1	F	187/191 (98%)	186 (100%)	1 (0%)	86	93
1	G	191/191 (100%)	188 (98%)	3 (2%)	58	76
2	p	313/408 (77%)	311 (99%)	2 (1%)	84	91
2	q	316/408 (78%)	311 (98%)	5 (2%)	58	76
2	r	313/408 (77%)	300 (96%)	13 (4%)	25	54
2	s	315/408 (77%)	308 (98%)	7 (2%)	47	70
3	AV	113/116 (97%)	111 (98%)	2 (2%)	54	74
3	Aa	113/116 (97%)	113 (100%)	0	100	100
3	Ad	113/116 (97%)	111 (98%)	2 (2%)	54	74
3	Ae	113/116 (97%)	113 (100%)	0	100	100
4	AS	147/153 (96%)	137 (93%)	10 (7%)	13	39
4	AT	148/153 (97%)	135 (91%)	13 (9%)	8	31
All	All	2955/3357 (88%)	2887 (98%)	68 (2%)	46	69

5 of 68 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	AT	95	LYS
4	AT	157	GLU
4	AT	173	ILE
2	r	75	CYS
2	r	28	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 35 such sidechains are listed below:

Mol	Chain	Res	Type
2	s	470	GLN
2	s	475	HIS
4	AT	75	ASN
2	p	452	ASN
2	p	365	ASN

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.

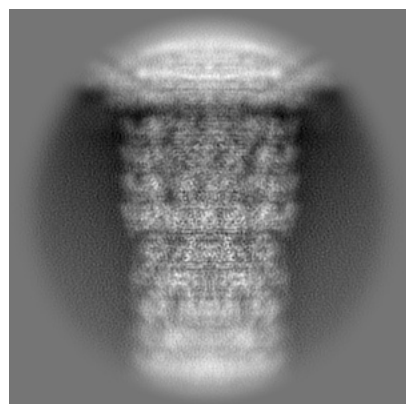
6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-29512. These allow visual inspection of the internal detail of the map and identification of artifacts.

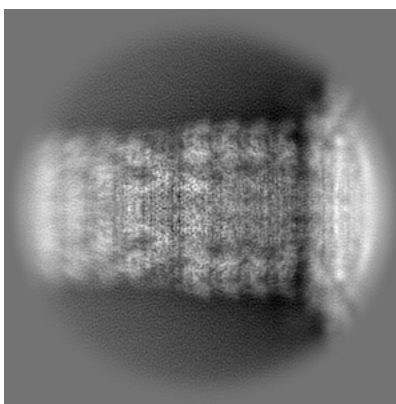
Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

6.1 Orthogonal projections [i](#)

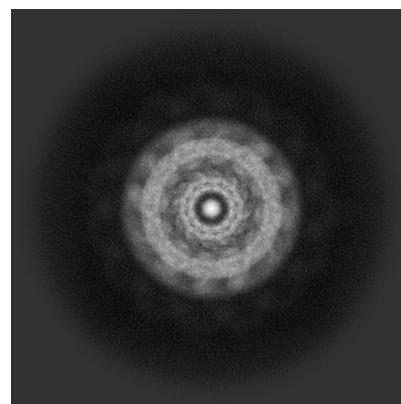
6.1.1 Primary map



X

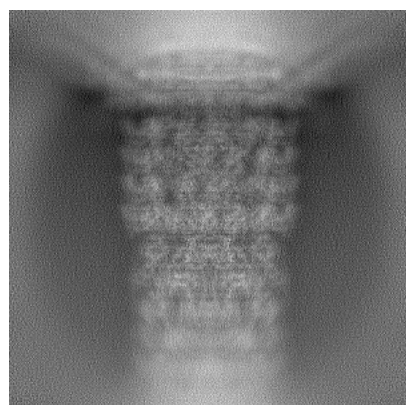


Y

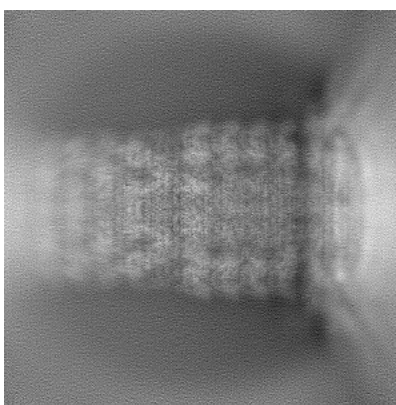


Z

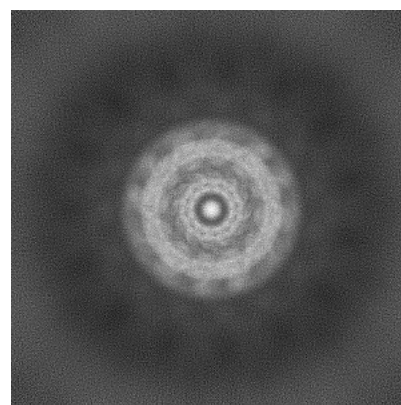
6.1.2 Raw map



X



Y

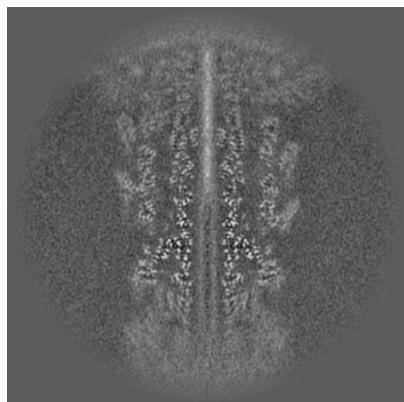


Z

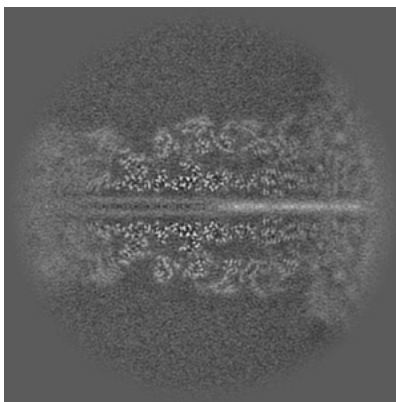
The images above show the map projected in three orthogonal directions.

6.2 Central slices [i](#)

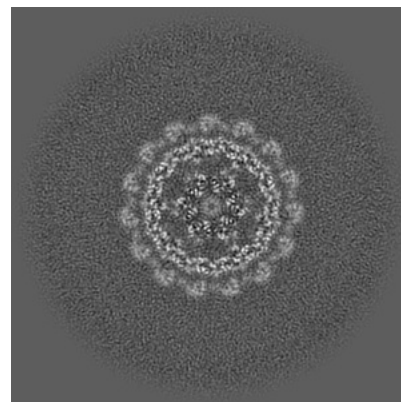
6.2.1 Primary map



X Index: 224

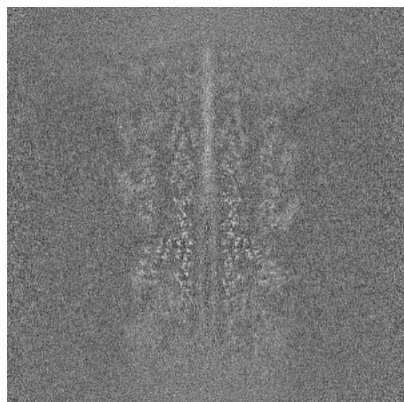


Y Index: 224

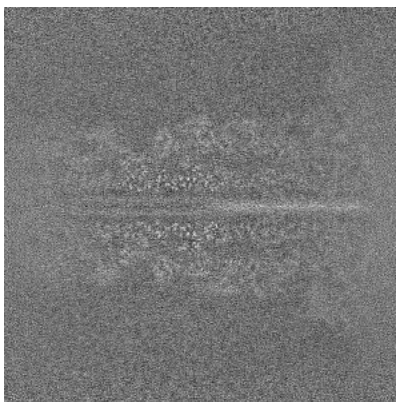


Z Index: 224

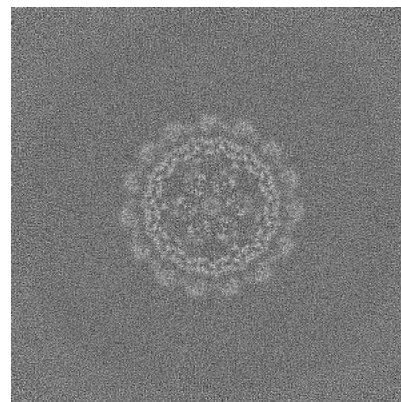
6.2.2 Raw map



X Index: 224



Y Index: 224

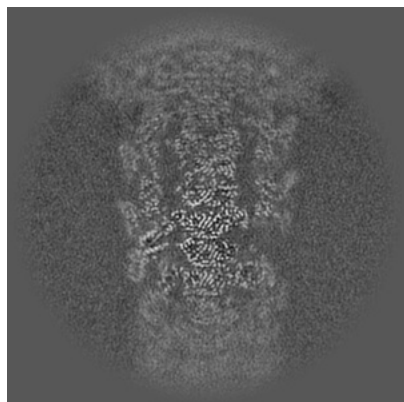


Z Index: 224

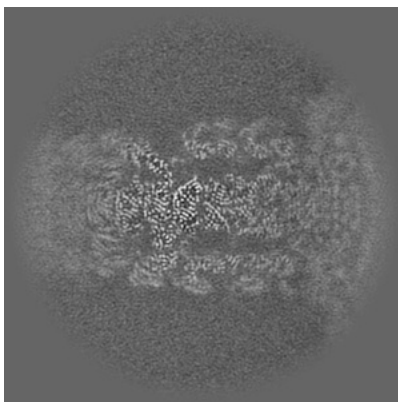
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices [i](#)

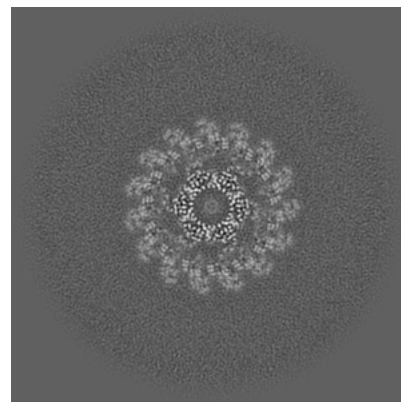
6.3.1 Primary map



X Index: 246

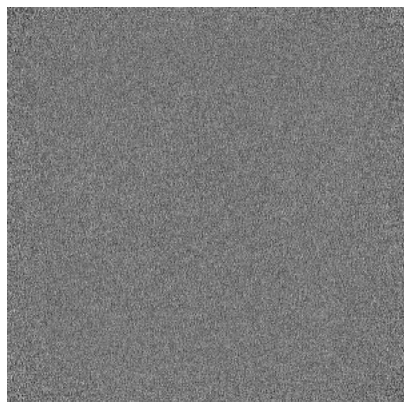


Y Index: 246

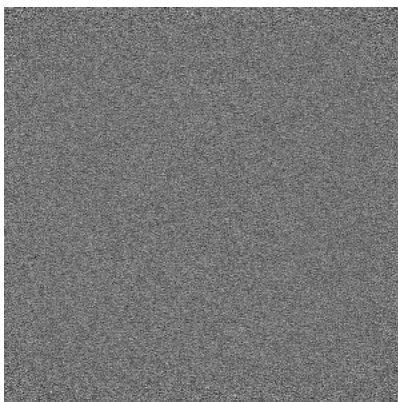


Z Index: 212

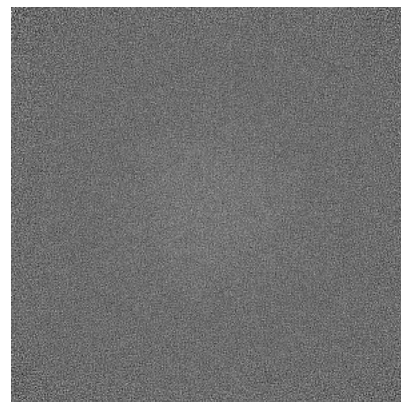
6.3.2 Raw map



X Index: 0



Y Index: 0

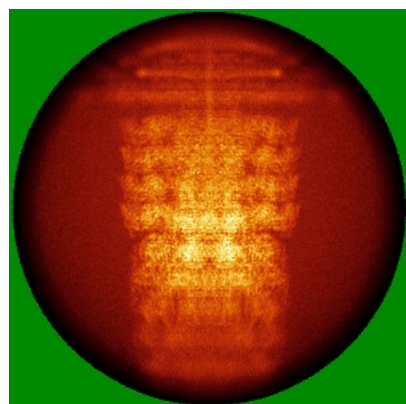


Z Index: 0

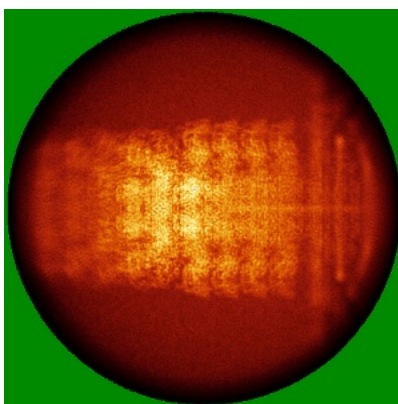
The images above show the largest variance slices of the map in three orthogonal directions.

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

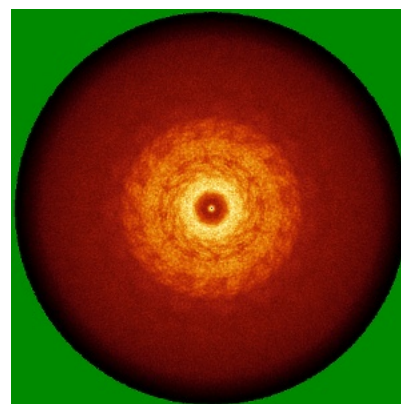
6.4.1 Primary map



X

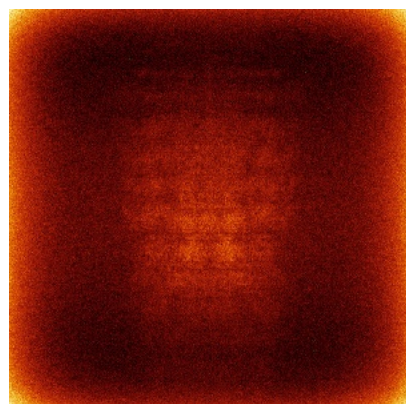


Y

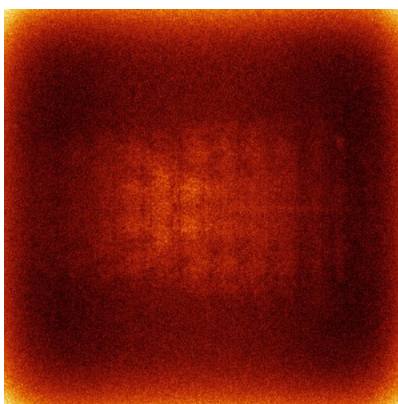


Z

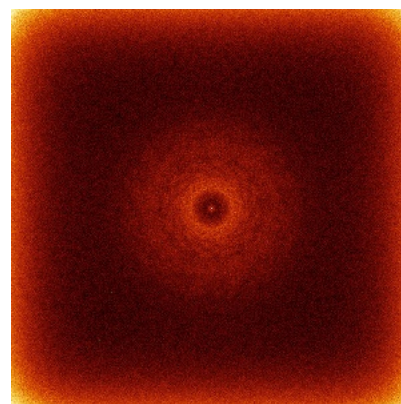
6.4.2 Raw map



X



Y

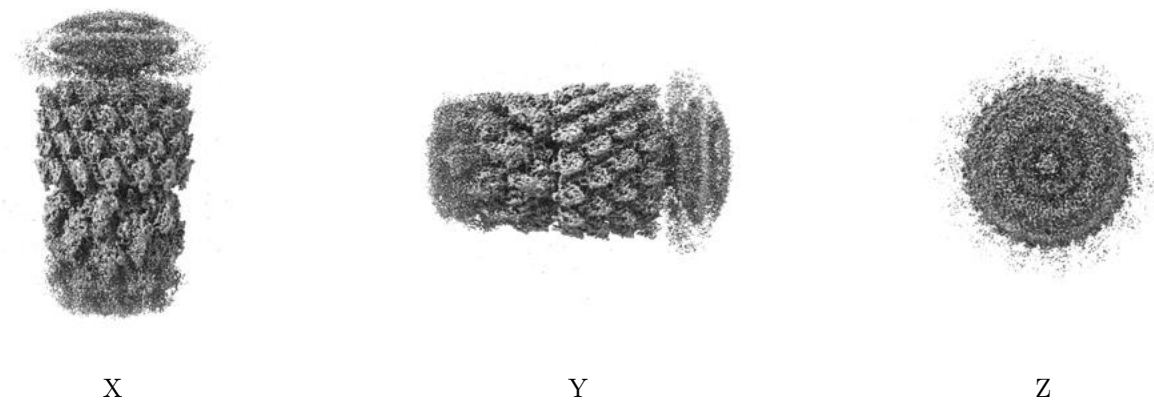


Z

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.

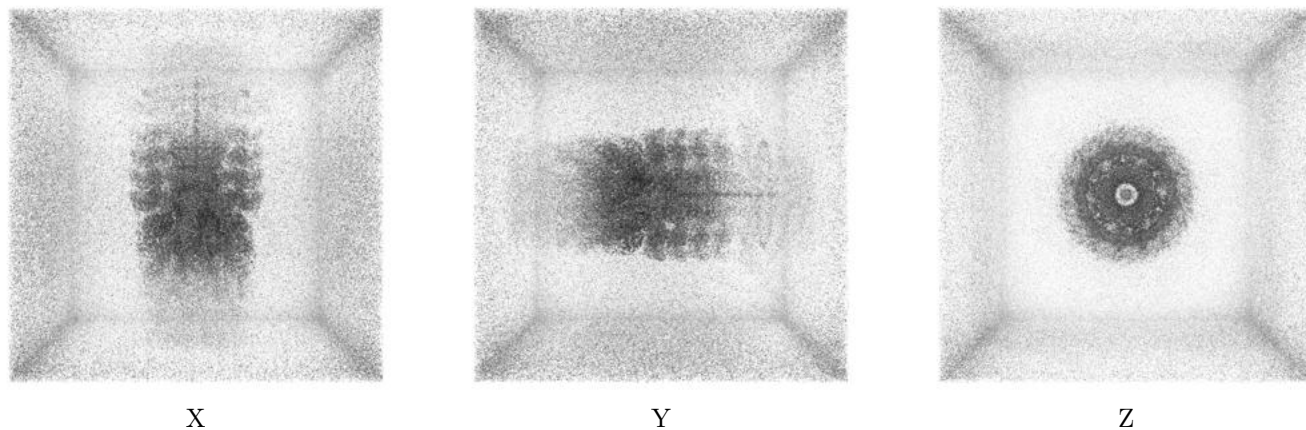
6.5 Orthogonal surface views [i](#)

6.5.1 Primary map



The images above show the 3D surface view of the map at the recommended contour level 0.15. 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.

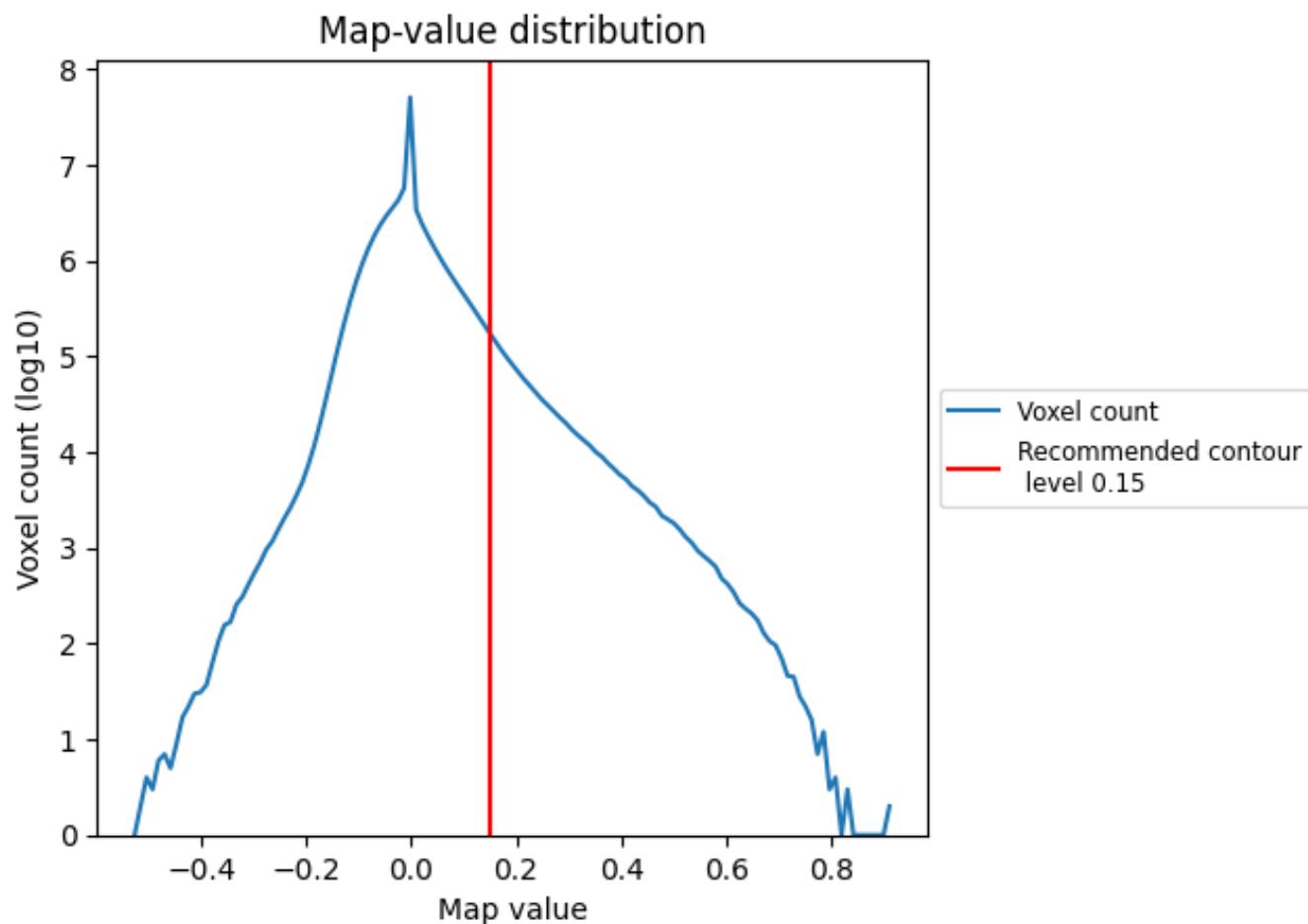
6.6 Mask visualisation [i](#)

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

7 Map analysis [i](#)

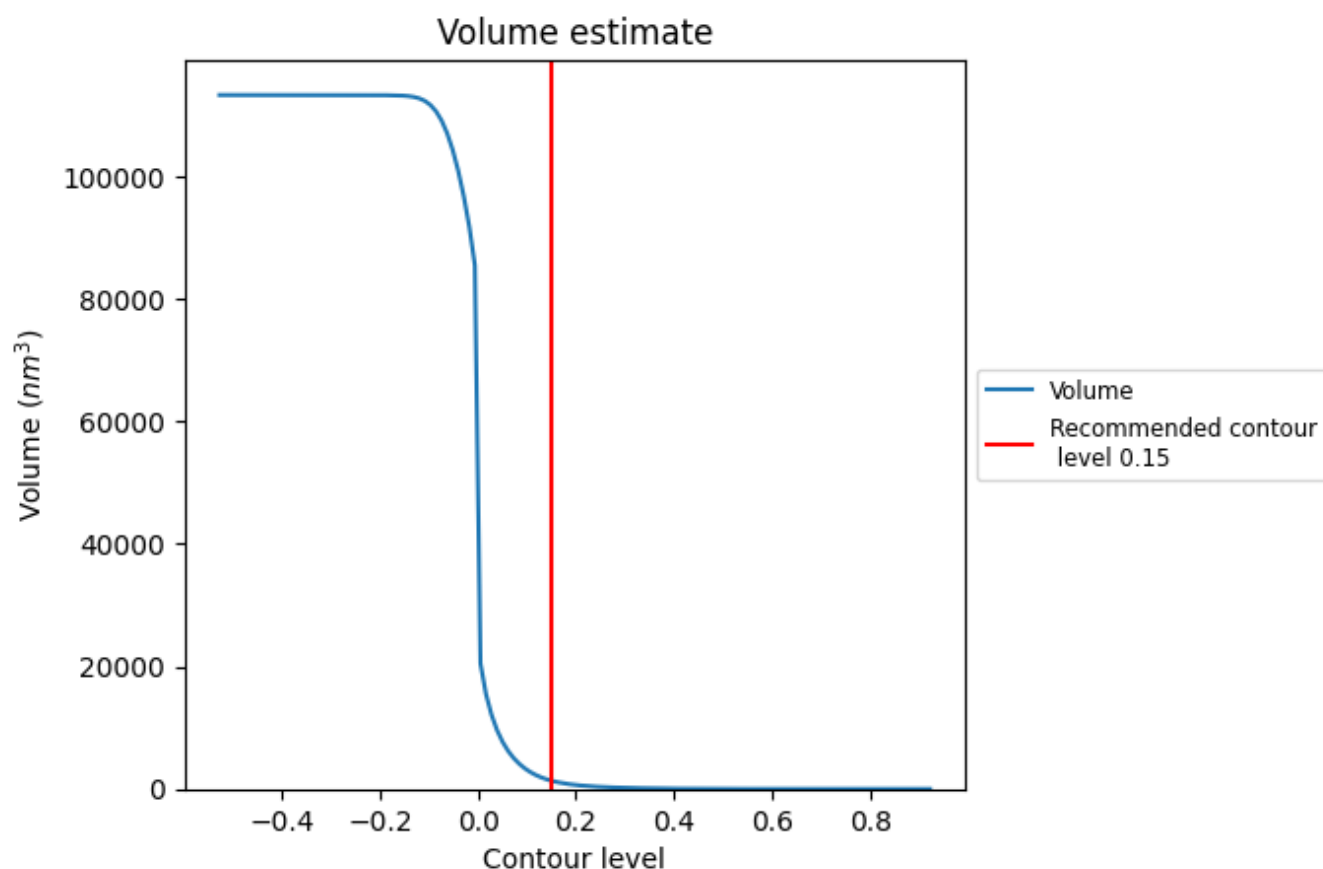
This section contains the results of statistical analysis of the map.

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.

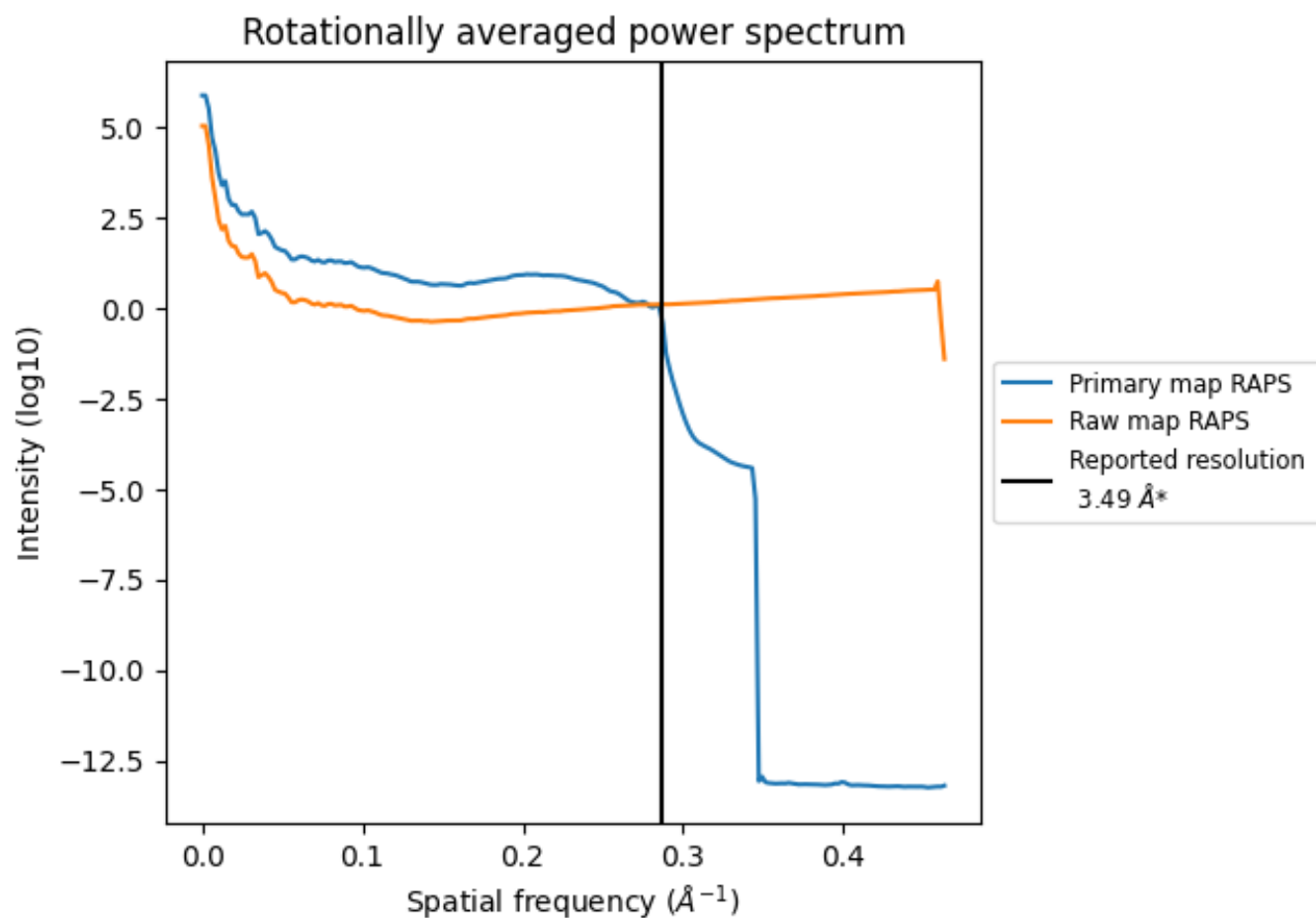
7.2 Volume estimate [i](#)



The volume at the recommended contour level is 1352 nm³; this corresponds to an approximate mass of 1221 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 ⓘ

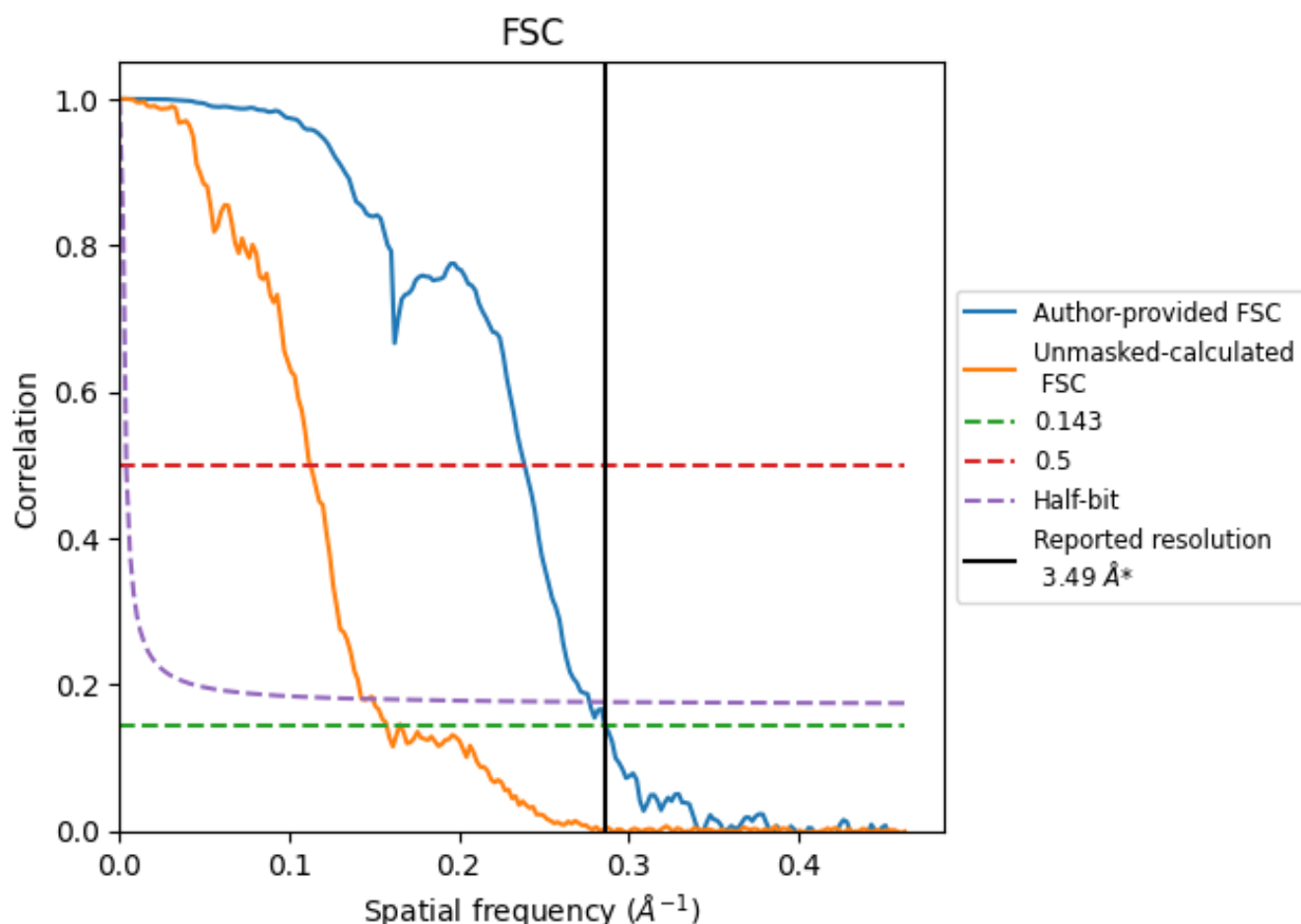


*Reported resolution corresponds to spatial frequency of 0.287 Å⁻¹

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 0.287 Å⁻¹

8.2 Resolution estimates [i](#)

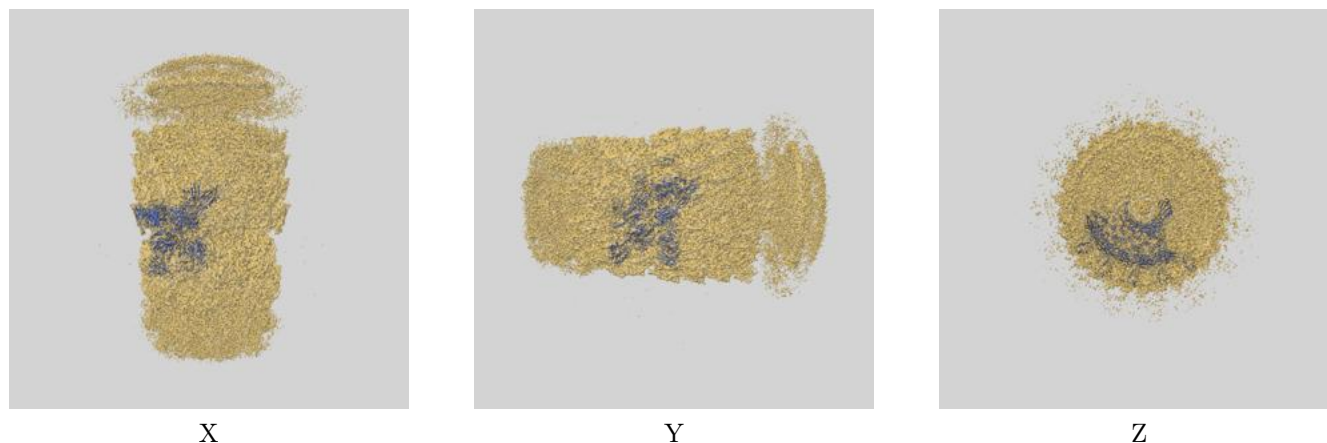
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	3.49	-	-
Author-provided FSC curve	3.49	4.19	3.61
Unmasked-calculated*	6.35	8.90	6.98

*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 6.35 differs from the reported value 3.49 by more than 10 %

9 Map-model fit [i](#)

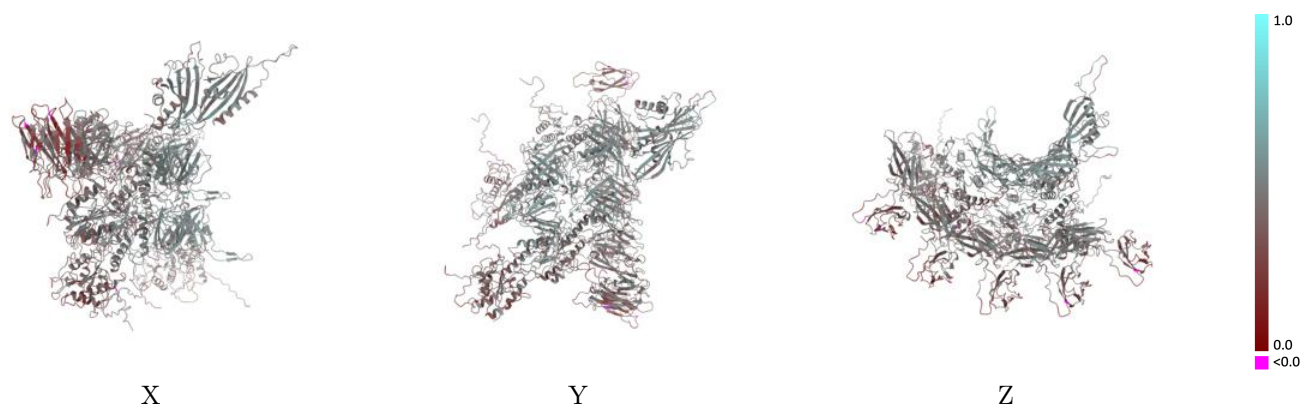
This section contains information regarding the fit between EMDB map EMD-29512 and PDB model 8FWM. Per-residue inclusion information can be found in section 3 on page 6.

9.1 Map-model overlay [i](#)



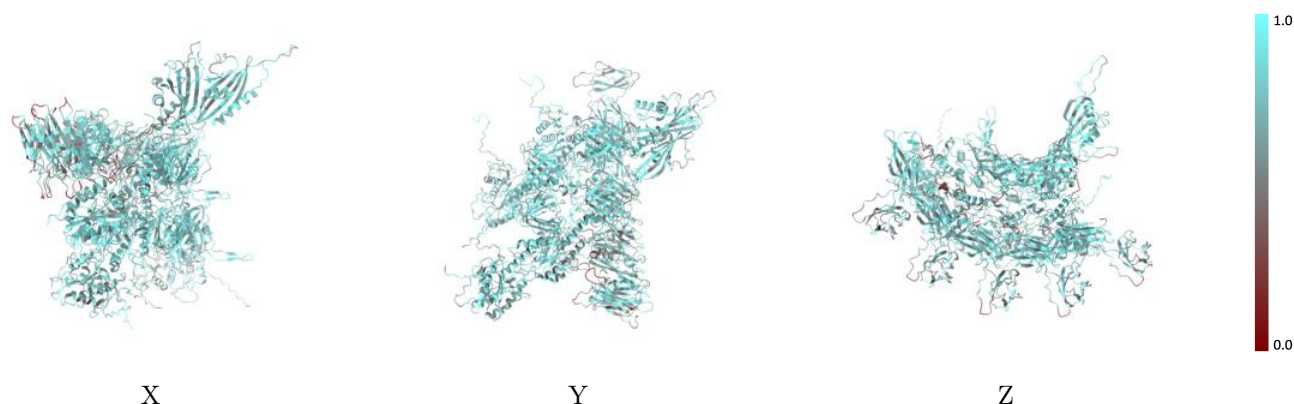
The images above show the 3D surface view of the map at the recommended contour level 0.15 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

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



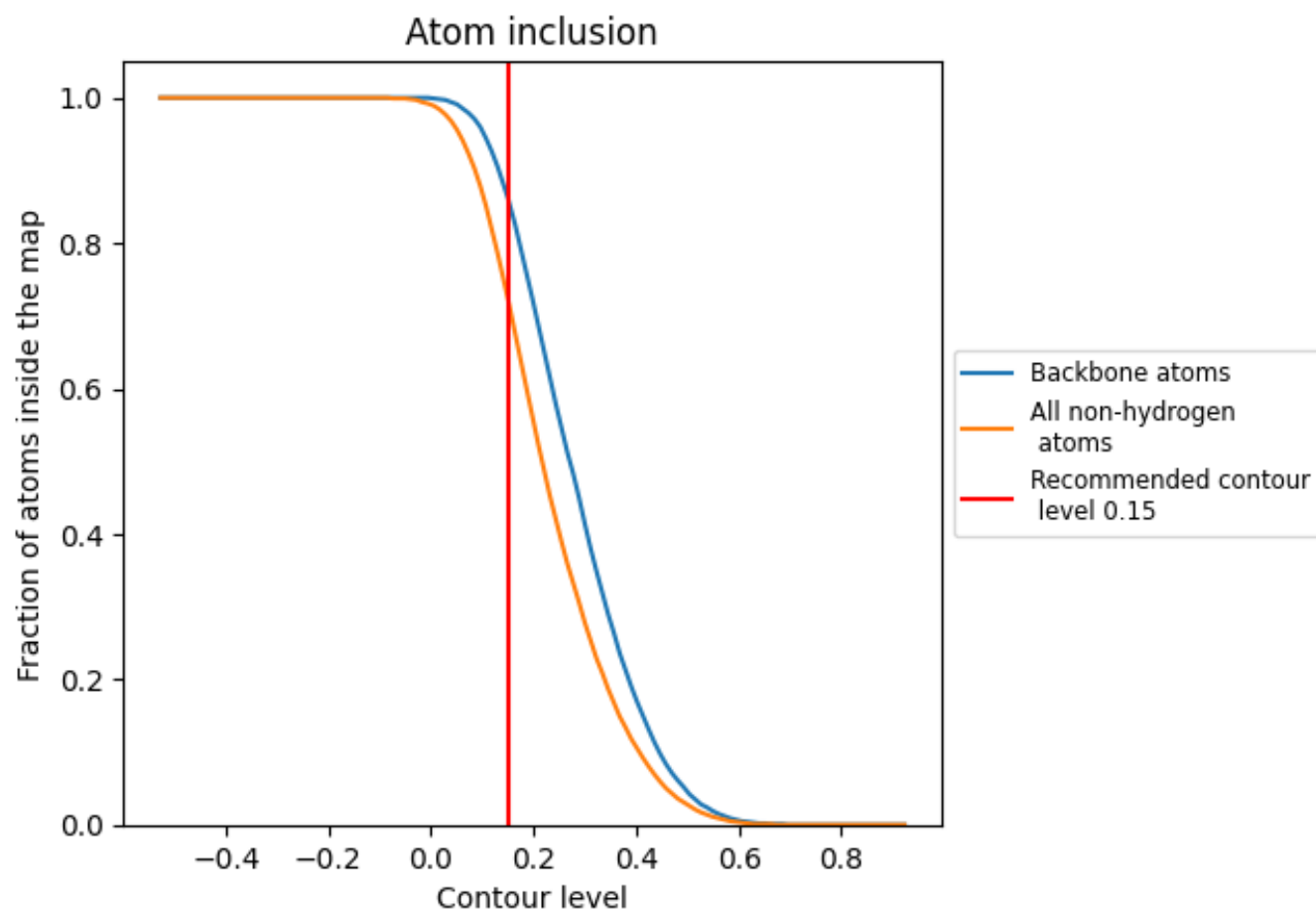
The images above show the model with each residue coloured according to its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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 (0.15).

































9.4 Atom inclusion [i](#)



At the recommended contour level, 86% of all backbone atoms, 72% 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 (0.15) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	 0.7230	 0.4270
3	 0.6670	 0.3810
AS	 0.7160	 0.4610
AT	 0.7380	 0.4850
AV	 0.7940	 0.5110
Aa	 0.7940	 0.5080
Ad	 0.7780	 0.5150
Ae	 0.7900	 0.5120
D	 0.6660	 0.3790
E	 0.6870	 0.4010
F	 0.6870	 0.3920
G	 0.6690	 0.3880
p	 0.7220	 0.4160
q	 0.7730	 0.4470
r	 0.7150	 0.3910
s	 0.7290	 0.4060

