



Full wwPDB EM Validation Report ⓘ

Apr 20, 2024 – 06:52 PM EDT

PDB ID : 8W22
EMDB ID : EMD-43737
Title : Umb1 umbrella toxin particle (local refinement of UmbB1 bound ALF of UmbC1 and UmbA1)
Authors : Park, Y.J.; Zhao, Q.; Seattle Structural Genomics Center for Infectious Disease (SSGCID); DiMaio, F.; Mougous, J.D.; Veesler, D.
Deposited on : 2024-02-19
Resolution : 4.00 Å(reported)

This is a Full wwPDB EM Validation 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.dev92
MolProbity : 4.02b-467
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ : 1.9.13
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.36.2

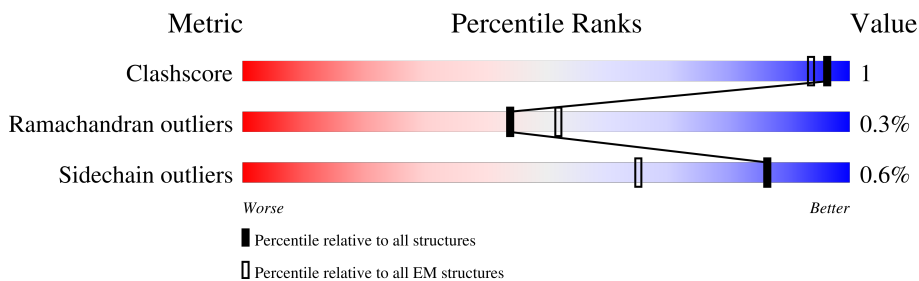
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 4.00 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	1368	7% 93%
2	B	166	74% .. 25%
3	C	515	8% 78% . 21%

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 3750 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 Intein C-terminal splicing domain-containing protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	99	685	426	134	123	2	0	0

There are 32 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	initiating methionine	UNP Q9ACV2
A	2	ARG	-	expression tag	UNP Q9ACV2
A	3	ARG	-	expression tag	UNP Q9ACV2
A	4	ARG	-	expression tag	UNP Q9ACV2
A	5	ILE	-	expression tag	UNP Q9ACV2
A	6	PRO	-	expression tag	UNP Q9ACV2
A	7	SER	-	expression tag	UNP Q9ACV2
A	8	ARG	-	expression tag	UNP Q9ACV2
A	9	THR	-	expression tag	UNP Q9ACV2
A	10	PRO	-	expression tag	UNP Q9ACV2
A	11	GLY	-	expression tag	UNP Q9ACV2
A	12	SER	-	expression tag	UNP Q9ACV2
A	13	GLY	-	expression tag	UNP Q9ACV2
A	14	ALA	-	expression tag	UNP Q9ACV2
A	15	LYS	-	expression tag	UNP Q9ACV2
A	16	GLN	-	expression tag	UNP Q9ACV2
A	17	LYS	-	expression tag	UNP Q9ACV2
A	18	SER	-	expression tag	UNP Q9ACV2
A	19	TRP	-	expression tag	UNP Q9ACV2
A	20	PHE	-	expression tag	UNP Q9ACV2
A	21	PRO	-	expression tag	UNP Q9ACV2
A	22	ARG	-	expression tag	UNP Q9ACV2
A	23	ARG	-	expression tag	UNP Q9ACV2
A	24	SER	-	expression tag	UNP Q9ACV2
A	25	LEU	-	expression tag	UNP Q9ACV2
A	26	GLN	-	expression tag	UNP Q9ACV2
A	27	VAL	-	expression tag	UNP Q9ACV2
A	28	LEU	-	expression tag	UNP Q9ACV2

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	29	LEU	-	expression tag	UNP Q9ACV2
A	30	SER	-	expression tag	UNP Q9ACV2
A	31	ALA	-	expression tag	UNP Q9ACV2
A	32	GLY	-	expression tag	UNP Q9ACV2

- Molecule 2 is a protein called Secreted protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	B	125	805	517	143	139	6	0	0

- Molecule 3 is a protein called Secreted esterase.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	C	405	2260	1390	446	419	5	0	0

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	506	GLY	-	expression tag	UNP Q9ACV4
C	507	SER	-	expression tag	UNP Q9ACV4
C	508	HIS	-	expression tag	UNP Q9ACV4
C	509	HIS	-	expression tag	UNP Q9ACV4
C	510	HIS	-	expression tag	UNP Q9ACV4
C	511	HIS	-	expression tag	UNP Q9ACV4
C	512	HIS	-	expression tag	UNP Q9ACV4
C	513	HIS	-	expression tag	UNP Q9ACV4
C	514	HIS	-	expression tag	UNP Q9ACV4
C	515	HIS	-	expression tag	UNP Q9ACV4

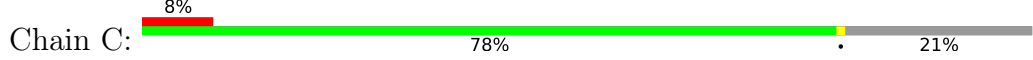
ALA	ALA	ASP	GLU	ALA	SER	ALA	GLY	LEU	ALA	ALA	THR	ALA	ALA	ALA	ASP	GLU	VAL	LEU	THR
ASP	GLU	ASP	ALA	GLY	LEU	GLY	GLY	LEU	ASP	ASP	ASP	ASP	THR	LEU	ALA	ALA	ASP	GLU	VAL
MET	PRO	CYS	ALA	ALA	GLY	SER	PRO	PRO	ASP	ASP	ASP	THR	ASP	LEU	ALA	GLY	ASP	THR	LEU
LEU	VAL	SER	LEU	LEU	LEU	ASP	ASP	VAL	VAL	VAL	GLY	THR	THR	THR	ASP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

● Molecule 2: Secreted protein



MET	ALA	ASN	THR	SER	ARG	THR	ARG	THR	ARG	GLN	ALA	LEU	LEU	TRP	ILE	ALA	VAL	LEU	THR
E42	Y147	G146	G166																

● Molecule 3: Secreted esterase



MET	PHE	ARG	MET	PRO	ARG	PRO	PRO	ILE	ARG	GLN	ALA	LEU	LEU	ALA	ALA	VAL	LEU	LEU	THR
T30	L46	G146	N205	G206	S256	A262	K287	PRO	ASP	GLY	THR	L292	R296	E297	LEU	TRP	ASN	ASN	THR
M307	G312	G313	D314	S317	D316	G319	N320	A403	G321	I322	ALA	ALA	ALA	VAL	VAL	ARG	ASP	GLY	THR
L382	G337	T338	A339	L343	H351	ASP	THR	SER	TRP	K356	R363	F366	N367	G368	R369	ASP	G371	A374	G377
L404	L410	T411	M417	ALA	ASP	GLY	ARG	ASP	D423	V427	G431	S432	L433	Y436	A437	GLY	ASN	THR	LYS
G498	A504	G473	D474	I475	A477	Y487	THR	GLY	THR	GLY	GLY	LEU	LEU	ASP	ALA	A447	D453	T457	P458
G464	GLY	ASP	PHE	ASN	GLY	THR	ASP	GLY	ASN	THR	THR	THR	THR	THR	THR	THR	THR	THR	THR
A504	G498	ALA	ALA	ALA	THR	THR	THR	THR	THR	THR	THR	THR	THR	THR	THR	THR	THR	THR	THR

4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	386275	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 ($e^-/\text{\AA}^2$)	60	Depositor
Minimum defocus (nm)	200	Depositor
Maximum defocus (nm)	3000	Depositor
Magnification	Not provided	
Image detector	GATAN K3 (6k x 4k)	Depositor
Maximum map value	9.149	Depositor
Minimum map value	-6.592	Depositor
Average map value	0.000	Depositor
Map value standard deviation	0.072	Depositor
Recommended contour level	0.9	Depositor
Map size (\AA)	512.544, 512.544, 512.544	wwPDB
Map dimensions	304, 304, 304	wwPDB
Map angles ($^\circ$)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (\AA)	1.686, 1.686, 1.686	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	0.59	0/692	0.48	0/936
2	B	0.46	0/822	0.53	0/1126
3	C	0.39	0/2282	0.48	0/3133
All	All	0.45	0/3796	0.49	0/5195

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	685	0	637	2	0
2	B	805	0	678	1	0
3	C	2260	0	1487	2	0
All	All	3750	0	2802	5	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 1.

All (5) 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:140:GLU:N	1:A:141:PRO:CD	2.66	0.58

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:147:TYR:O	2:B:148:GLY:C	2.50	0.49
3:C:457:THR:N	3:C:458:PRO:CD	2.78	0.47
3:C:457:THR:N	3:C:458:PRO:HD2	2.35	0.41
1:A:140:GLU:N	1:A:141:PRO:HD2	2.36	0.40

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	97/1368 (7%)	94 (97%)	2 (2%)	1 (1%)	15	53
2	B	123/166 (74%)	115 (94%)	7 (6%)	1 (1%)	19	58
3	C	379/515 (74%)	371 (98%)	8 (2%)	0	100	100
All	All	599/2049 (29%)	580 (97%)	17 (3%)	2 (0%)	44	75

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	148	GLY
1	A	121	ASP

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	53/961 (6%)	53 (100%)	0	100	100
2	B	53/123 (43%)	53 (100%)	0	100	100
3	C	74/395 (19%)	73 (99%)	1 (1%)	67	81
All	All	180/1479 (12%)	179 (99%)	1 (1%)	86	92

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	C	46	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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 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](#)

There are no chain breaks in this entry.

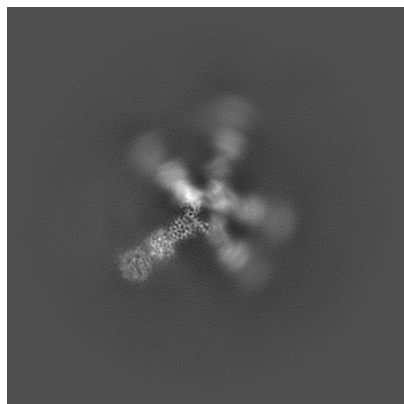
6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-43737. These allow visual inspection of the internal detail of the map and identification of artifacts.

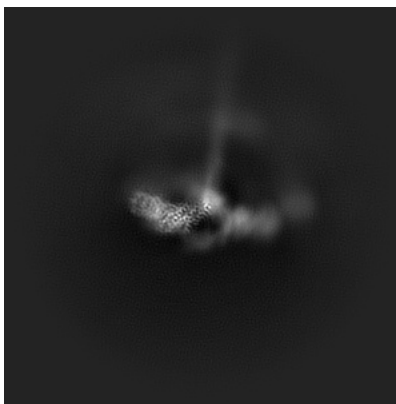
Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

6.1 Orthogonal projections [i](#)

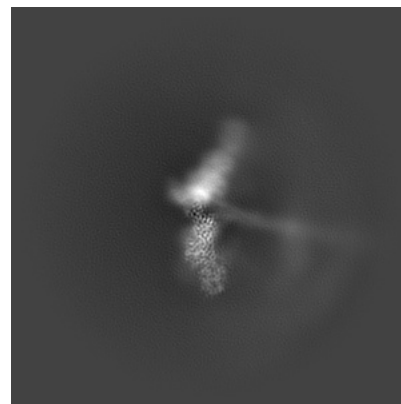
6.1.1 Primary map



X

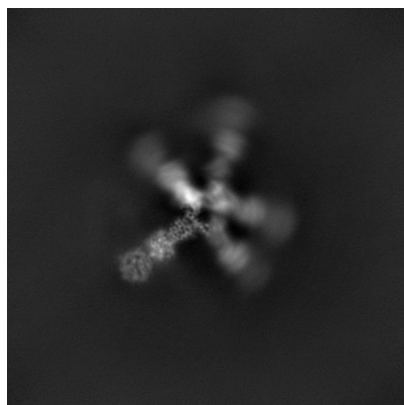


Y

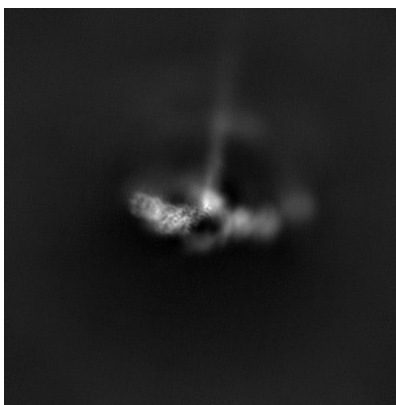


Z

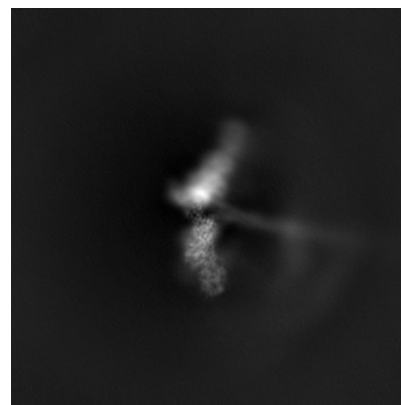
6.1.2 Raw map



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: 152

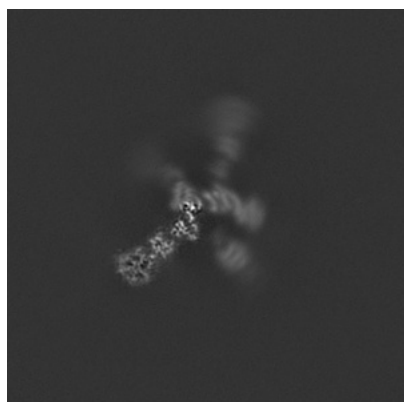


Y Index: 152

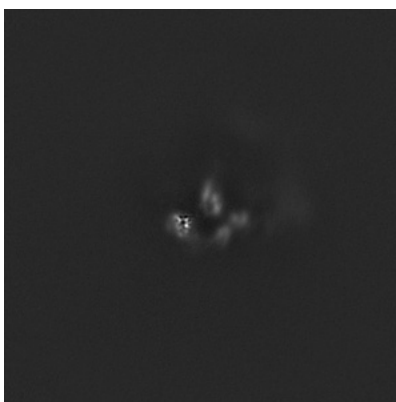


Z Index: 152

6.2.2 Raw map



X Index: 152



Y Index: 152

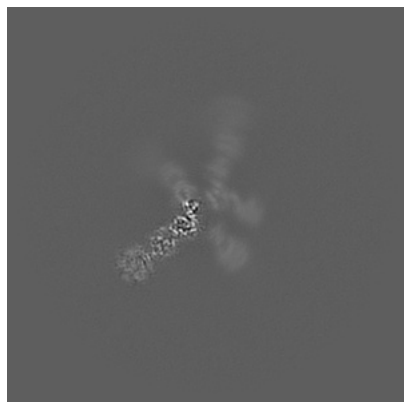


Z Index: 152

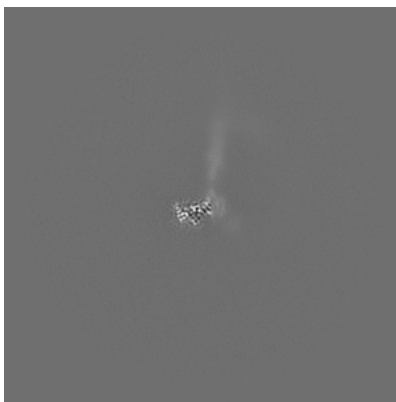
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices [i](#)

6.3.1 Primary map



X Index: 149

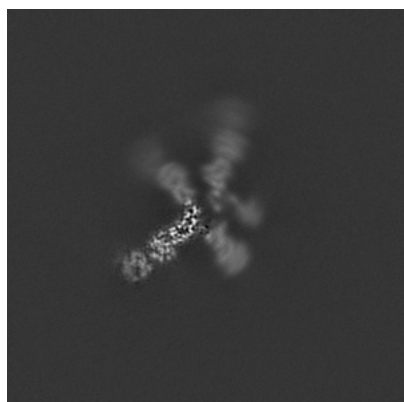


Y Index: 139

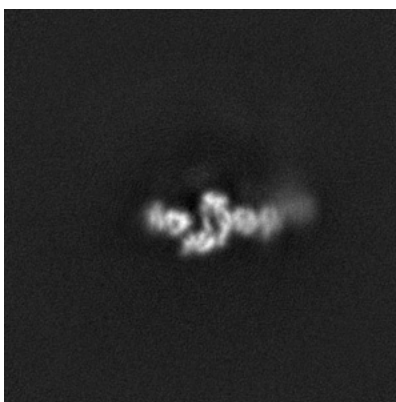


Z Index: 138

6.3.2 Raw map



X Index: 146



Y Index: 162

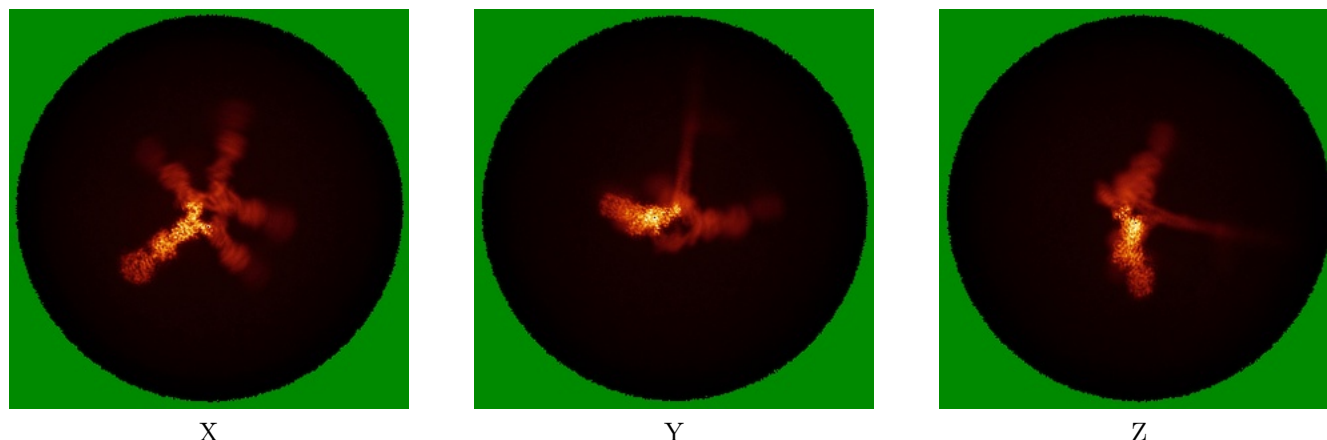


Z Index: 154

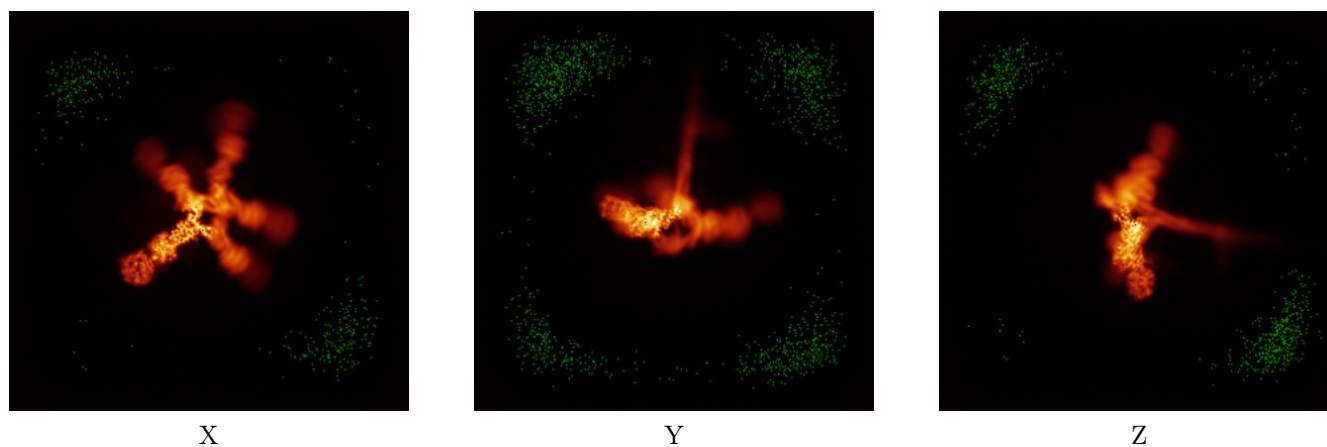
The images above show the largest variance slices of the map in three orthogonal directions.

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

6.4.1 Primary map



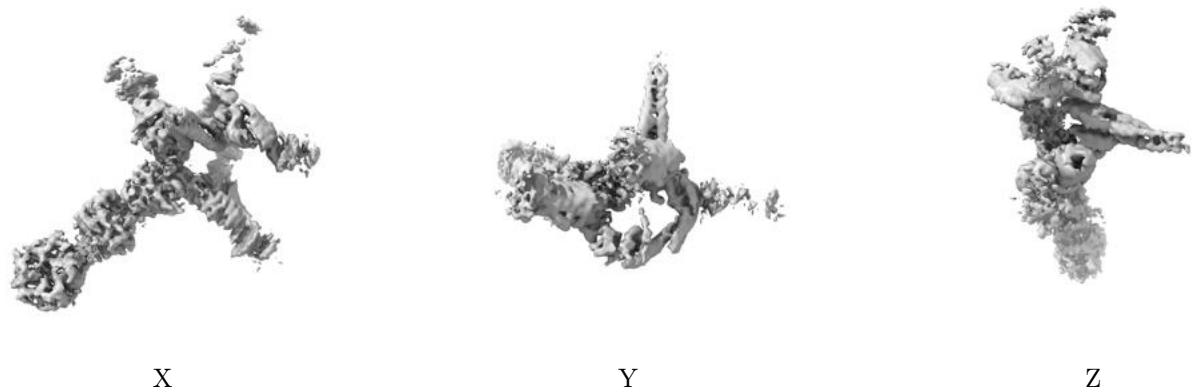
6.4.2 Raw map



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.

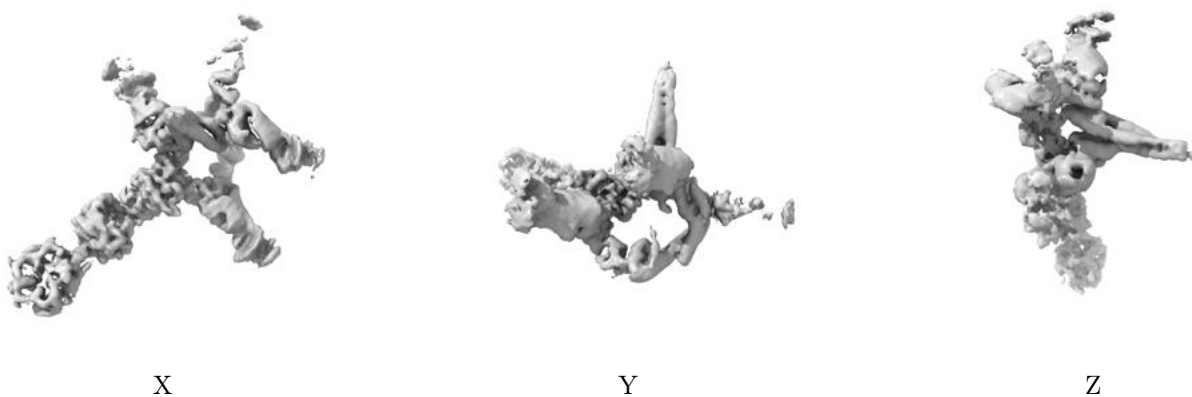
6.5 Orthogonal surface views [i](#)

6.5.1 Primary map



The images above show the 3D surface view of the map at the recommended contour level 0.9. 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.

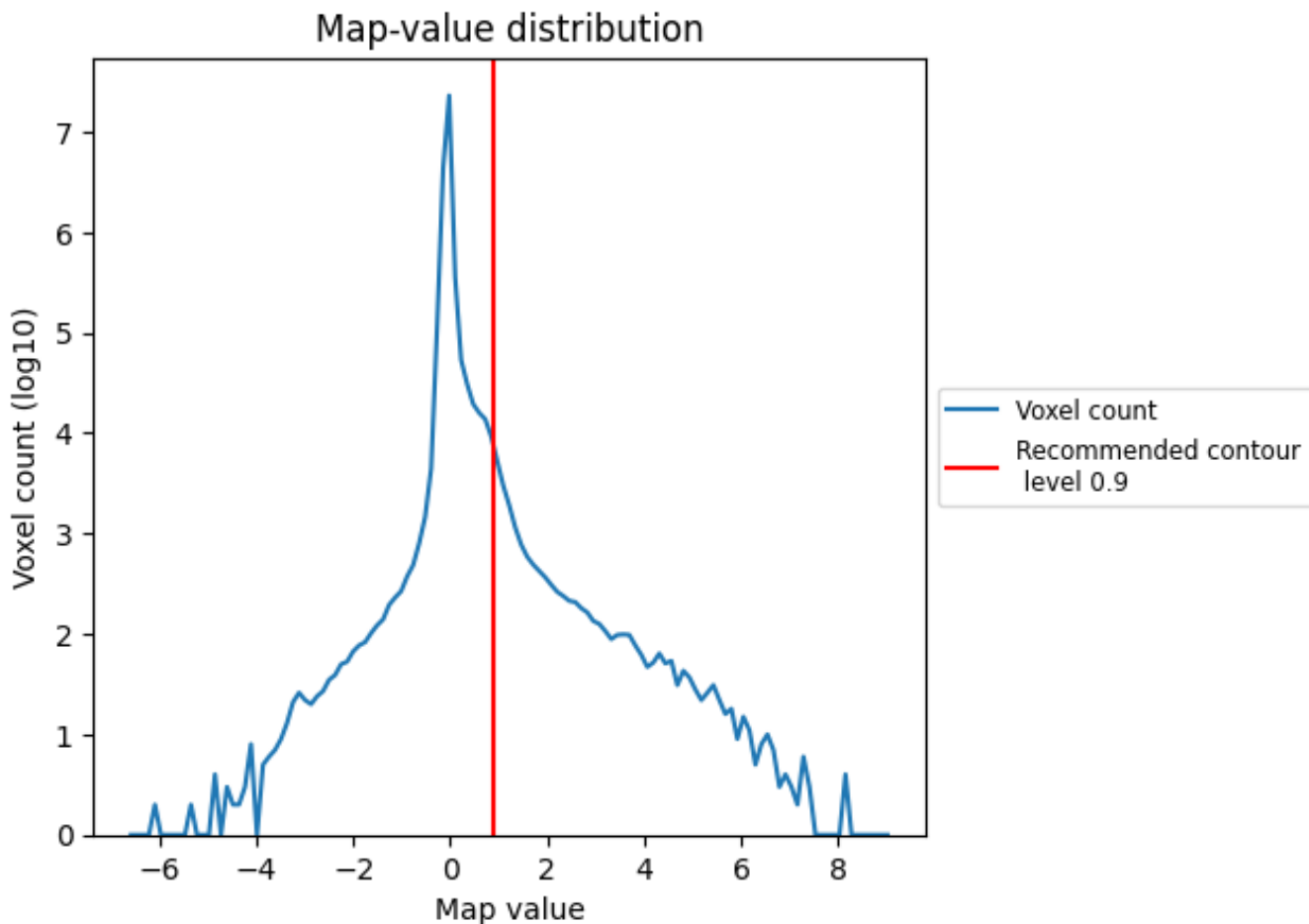
6.6 Mask visualisation [i](#)

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

7 Map analysis [i](#)

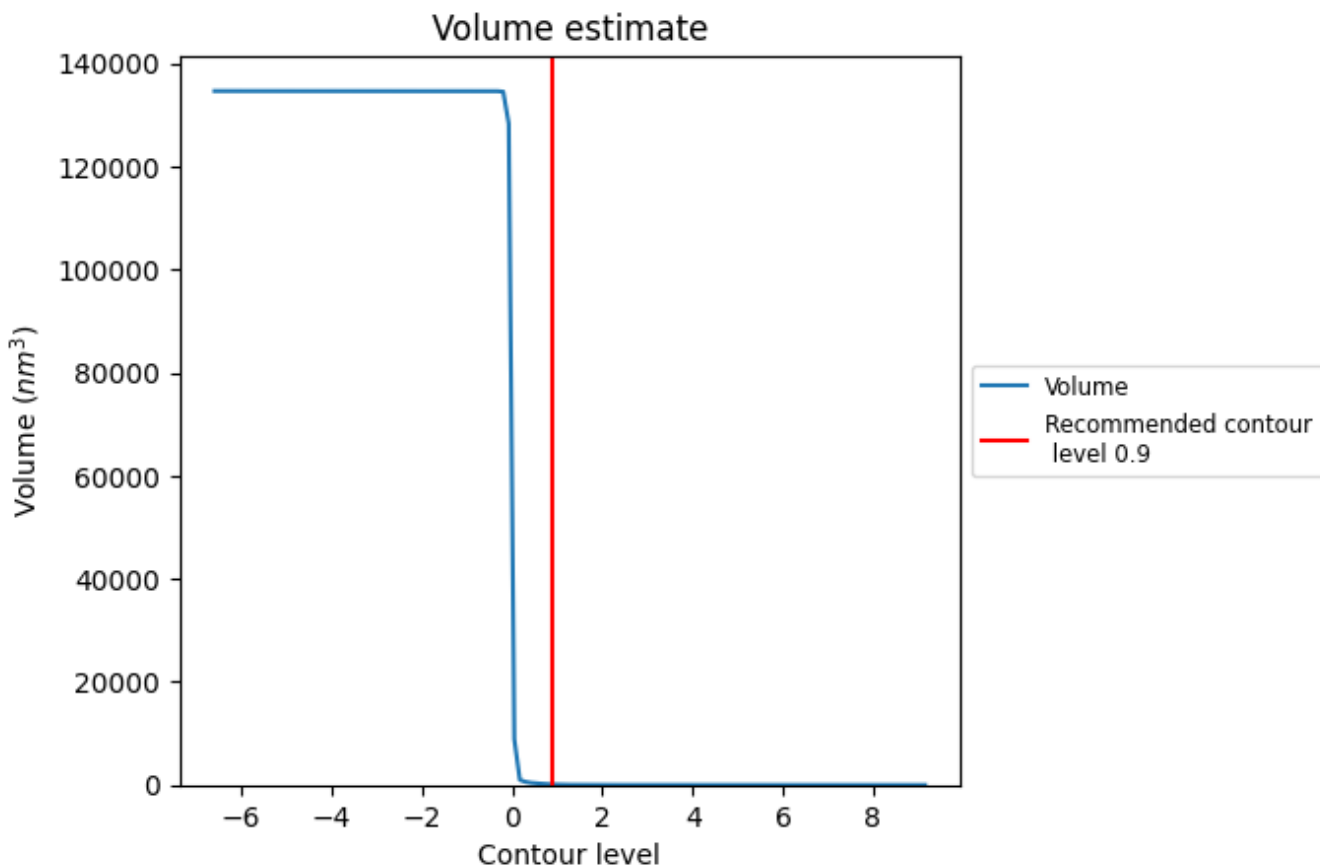
This section contains the results of statistical analysis of the map.

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.

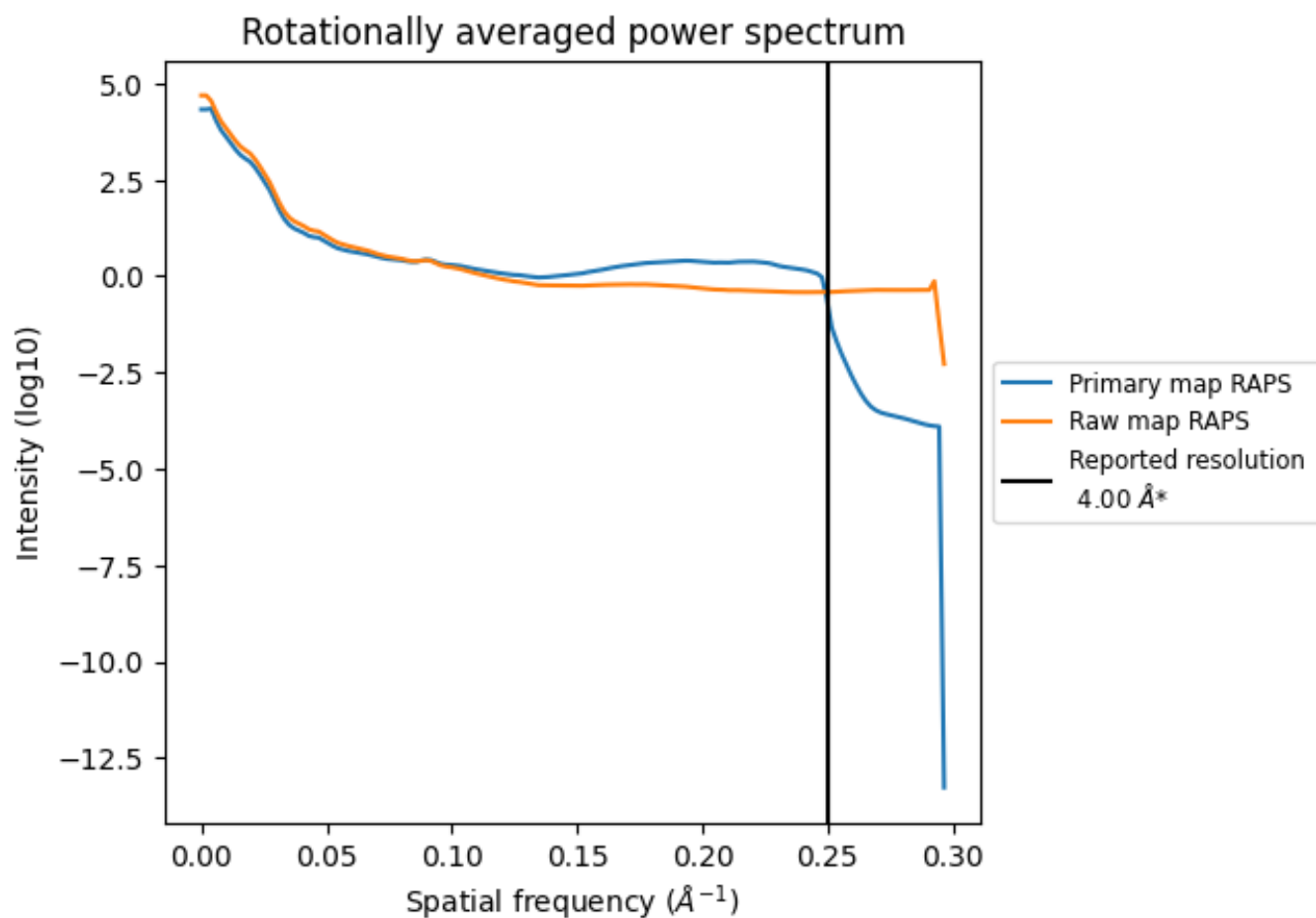
7.2 Volume estimate [\(i\)](#)



The volume at the recommended contour level is 105 nm³; this corresponds to an approximate mass of 95 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 [i](#)

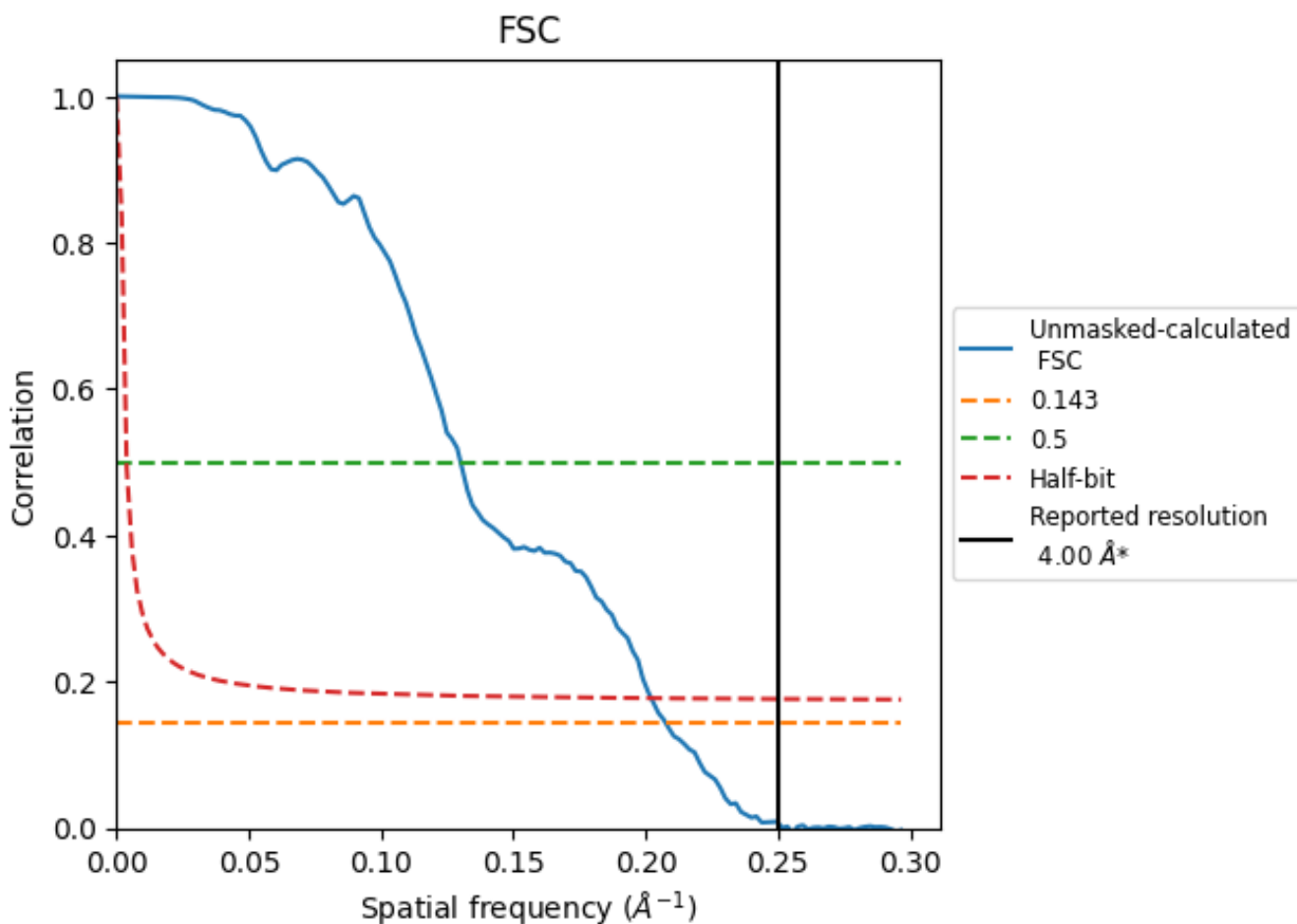


*Reported resolution corresponds to spatial frequency of 0.250 Å⁻¹

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 0.250 Å⁻¹

8.2 Resolution estimates [i](#)

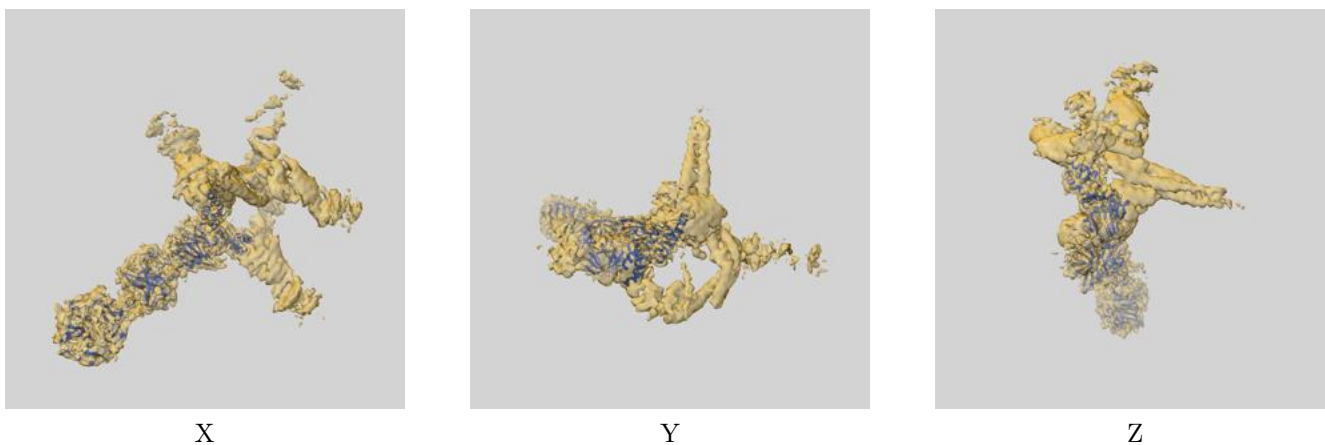
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	4.00	-	-
Author-provided FSC curve	-	-	-
Unmasked-calculated*	4.81	7.69	4.95

*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 4.81 differs from the reported value 4.0 by more than 10 %

9 Map-model fit [i](#)

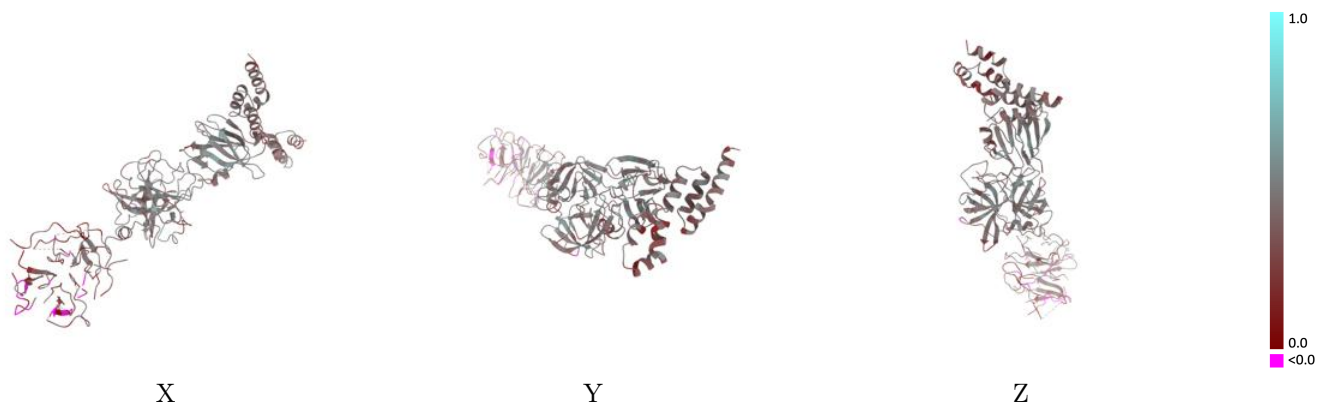
This section contains information regarding the fit between EMDB map EMD-43737 and PDB model 8W22. Per-residue inclusion information can be found in section 3 on page 5.

9.1 Map-model overlay [i](#)



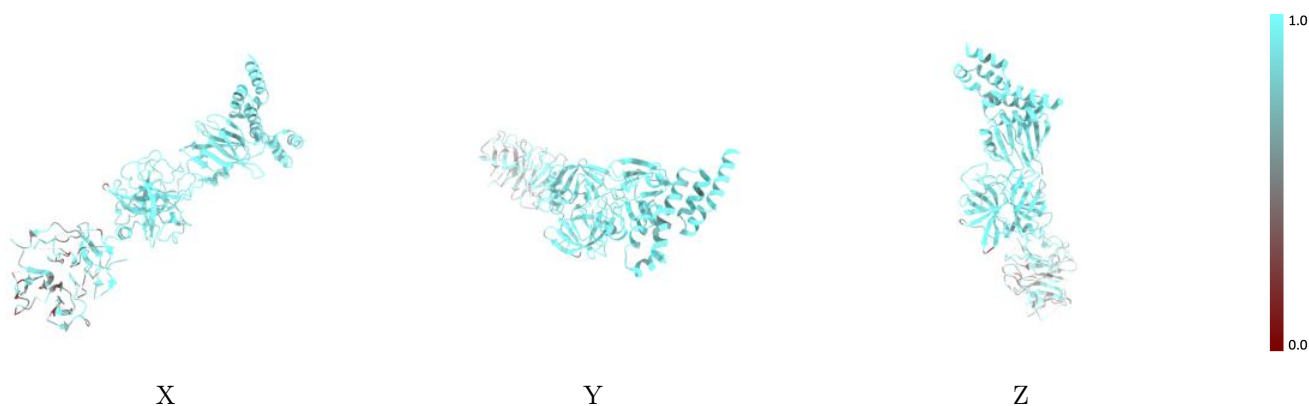
The images above show the 3D surface view of the map at the recommended contour level 0.9 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

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



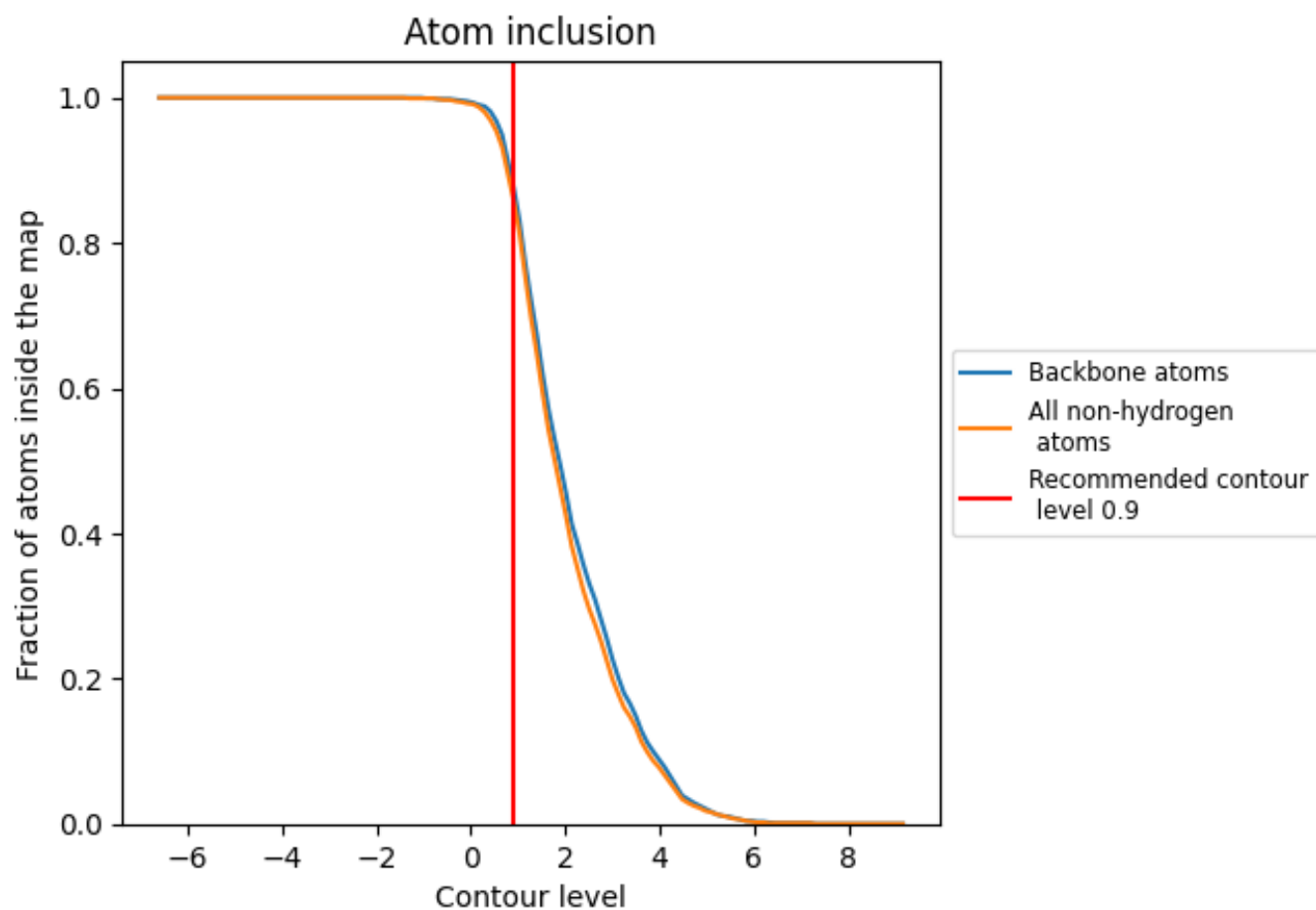
The images above show the model with each residue coloured according to its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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 (0.9).







9.4 Atom inclusion [i](#)



At the recommended contour level, 88% of all backbone atoms, 86% 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 (0.9) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	 0.8610	 0.3860
A	 0.8920	 0.3750
B	 0.9270	 0.4470
C	 0.8280	 0.3670

