



wwPDB EM Validation Summary Report ⓘ

Sep 10, 2024 – 12:10 PM JST

PDB ID : 8K17
EMDB ID : EMD-36787
Title : Human collagen prolyl processing enzyme complex, P3H1/CRTAP/PPIB heterotrimer, bound to collagen alpha-1(I) chain
Authors : Li, W.; Peng, J.; Yao, D.; Rao, B.; Xia, Y.; Wang, Q.; Li, S.; Cao, M.; Shen, Y.; Ma, P.; Liao, R.; Qin, A.; Zhao, J.; Cao, Y.
Deposited on : 2023-07-10
Resolution : 3.18 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.dev112
Mogul : 1.8.5 (274361), CSD as541be (2020)
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) : 2.38.2

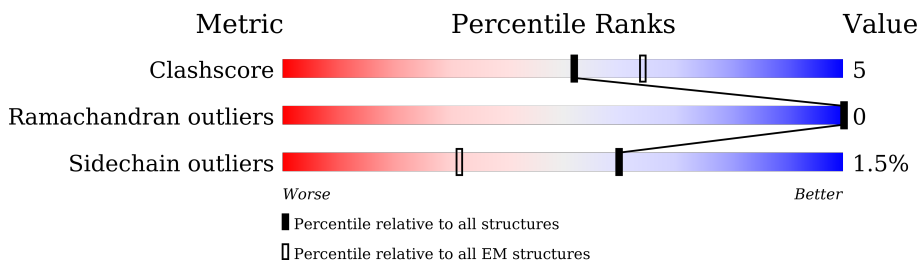
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 3.18 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	A	736	 5% 75% 12% 12%
2	B	442	 5% 73% 8% 18%
3	C	256	 61% 11% 28%
4	E	21	 5% 48% 52%

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 9748 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 Prolyl 3-hydroxylase 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	646	5193	3289	889	985	30	0	0

- Molecule 2 is a protein called Cartilage-associated protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	B	362	2994	1920	506	551	17	0	0

There are 41 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	402	ALA	-	expression tag	UNP O75718
B	403	ALA	-	expression tag	UNP O75718
B	404	ALA	-	expression tag	UNP O75718
B	405	LEU	-	expression tag	UNP O75718
B	406	GLU	-	expression tag	UNP O75718
B	407	VAL	-	expression tag	UNP O75718
B	408	LEU	-	expression tag	UNP O75718
B	409	PHE	-	expression tag	UNP O75718
B	410	GLN	-	expression tag	UNP O75718
B	411	GLY	-	expression tag	UNP O75718
B	412	PRO	-	expression tag	UNP O75718
B	413	SER	-	expression tag	UNP O75718
B	414	ALA	-	expression tag	UNP O75718
B	415	TRP	-	expression tag	UNP O75718
B	416	SER	-	expression tag	UNP O75718
B	417	HIS	-	expression tag	UNP O75718
B	418	PRO	-	expression tag	UNP O75718
B	419	GLN	-	expression tag	UNP O75718
B	420	PHE	-	expression tag	UNP O75718
B	421	GLU	-	expression tag	UNP O75718
B	422	LYS	-	expression tag	UNP O75718

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	423	GLY	-	expression tag	UNP O75718
B	424	GLY	-	expression tag	UNP O75718
B	425	GLY	-	expression tag	UNP O75718
B	426	SER	-	expression tag	UNP O75718
B	427	GLY	-	expression tag	UNP O75718
B	428	GLY	-	expression tag	UNP O75718
B	429	GLY	-	expression tag	UNP O75718
B	430	SER	-	expression tag	UNP O75718
B	431	GLY	-	expression tag	UNP O75718
B	432	GLY	-	expression tag	UNP O75718
B	433	SER	-	expression tag	UNP O75718
B	434	ALA	-	expression tag	UNP O75718
B	435	TRP	-	expression tag	UNP O75718
B	436	SER	-	expression tag	UNP O75718
B	437	HIS	-	expression tag	UNP O75718
B	438	PRO	-	expression tag	UNP O75718
B	439	GLN	-	expression tag	UNP O75718
B	440	PHE	-	expression tag	UNP O75718
B	441	GLU	-	expression tag	UNP O75718
B	442	LYS	-	expression tag	UNP O75718

- Molecule 3 is a protein called Peptidyl-prolyl cis-trans isomerase B.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	C	185	1438	923	246	265	4	0	0

There are 40 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	185	ALA	-	expression tag	UNP P23284
C	186	ALA	-	expression tag	UNP P23284
C	187	ALA	-	expression tag	UNP P23284
C	188	GLU	-	expression tag	UNP P23284
C	189	ASN	-	expression tag	UNP P23284
C	190	LEU	-	expression tag	UNP P23284
C	191	TYR	-	expression tag	UNP P23284
C	192	PHE	-	expression tag	UNP P23284
C	193	GLN	-	expression tag	UNP P23284
C	194	GLY	-	expression tag	UNP P23284
C	195	ASP	-	expression tag	UNP P23284
C	196	TYR	-	expression tag	UNP P23284

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	197	LYS	-	expression tag	UNP P23284
C	198	ASP	-	expression tag	UNP P23284
C	199	HIS	-	expression tag	UNP P23284
C	200	ASP	-	expression tag	UNP P23284
C	201	GLY	-	expression tag	UNP P23284
C	202	ASP	-	expression tag	UNP P23284
C	203	TYR	-	expression tag	UNP P23284
C	204	LYS	-	expression tag	UNP P23284
C	205	ASP	-	expression tag	UNP P23284
C	206	HIS	-	expression tag	UNP P23284
C	207	ASP	-	expression tag	UNP P23284
C	208	ILE	-	expression tag	UNP P23284
C	209	ASP	-	expression tag	UNP P23284
C	210	TYR	-	expression tag	UNP P23284
C	211	LYS	-	expression tag	UNP P23284
C	212	ASP	-	expression tag	UNP P23284
C	213	ASP	-	expression tag	UNP P23284
C	214	ASP	-	expression tag	UNP P23284
C	215	ASP	-	expression tag	UNP P23284
C	216	LYS	-	expression tag	UNP P23284
C	217	HIS	-	expression tag	UNP P23284
C	218	HIS	-	expression tag	UNP P23284
C	219	HIS	-	expression tag	UNP P23284
C	220	HIS	-	expression tag	UNP P23284
C	221	HIS	-	expression tag	UNP P23284
C	222	HIS	-	expression tag	UNP P23284
C	223	HIS	-	expression tag	UNP P23284
C	224	HIS	-	expression tag	UNP P23284

- Molecule 4 is a protein called synthetic substrate.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
4	E	10	66	43	13	10	0	0

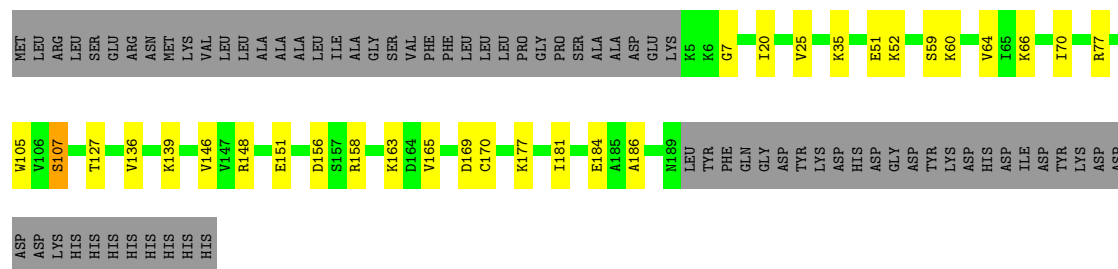
- Molecule 5 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: C₈H₁₅NO₆).



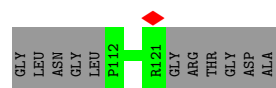
Mol	Chain	Residues	Atoms				AltConf
			Total	C	N	O	
5	A	1	14	8	1	5	0
5	A	1	14	8	1	5	0
5	B	1	14	8	1	5	0
5	B	1	14	8	1	5	0

- Molecule 6 is FE (III) ION (three-letter code: FE) (formula: Fe) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms		AltConf
			Total	Fe	
6	A	1	1	1	0



• Molecule 4: synthetic substrate



4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	1779834	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 ($e^-/\text{\AA}^2$)	50	Depositor
Minimum defocus (nm)	1500	Depositor
Maximum defocus (nm)	2600	Depositor
Magnification	Not provided	
Image detector	GATAN K3 (6k x 4k)	Depositor
Maximum map value	0.305	Depositor
Minimum map value	-0.166	Depositor
Average map value	0.000	Depositor
Map value standard deviation	0.008	Depositor
Recommended contour level	0.025	Depositor
Map size (\AA)	220.0, 220.0, 220.0	wwPDB
Map dimensions	200, 200, 200	wwPDB
Map angles ($^\circ$)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (\AA)	1.1, 1.1, 1.1	Depositor

5 Model quality [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: FE, NAG

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	0.44	1/5316 (0.0%)	0.57	3/7203 (0.0%)
2	B	0.42	0/3077	0.51	0/4160
3	C	0.46	0/1468	0.53	0/1969
4	E	0.30	0/70	0.47	0/96
All	All	0.44	1/9931 (0.0%)	0.55	3/13428 (0.0%)

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	1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	123	CYS	CB-SG	-6.06	1.72	1.82

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	123	CYS	CA-CB-SG	11.24	134.23	114.00
1	A	621	ASP	CB-CG-OD2	6.26	123.93	118.30
1	A	123	CYS	CB-CA-C	-5.57	99.27	110.40

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	123	CYS	Peptide

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	5193	0	5005	57	0
2	B	2994	0	2870	29	0
3	C	1438	0	1449	17	0
4	E	66	0	68	0	0
5	A	28	0	26	0	0
5	B	28	0	26	1	0
6	A	1	0	0	0	0
All	All	9748	0	9444	99	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 5.

The worst 5 of 99 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)
2:B:27:GLN:HE21	2:B:27:GLN:N	1.27	1.30
2:B:27:GLN:N	2:B:27:GLN:NE2	2.11	0.97
1:A:253:ASP:O	1:A:256:ASN:ND2	2.14	0.80
1:A:482:GLU:HG3	1:A:537:LEU:HD21	1.70	0.71
1:A:94:SER:HB2	1:A:97:GLN:HB2	1.72	0.71

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	640/736 (87%)	617 (96%)	23 (4%)	0	100	100
2	B	360/442 (81%)	346 (96%)	14 (4%)	0	100	100
3	C	183/256 (72%)	163 (89%)	20 (11%)	0	100	100
4	E	8/21 (38%)	7 (88%)	1 (12%)	0	100	100
All	All	1191/1455 (82%)	1133 (95%)	58 (5%)	0	100	100

There are no Ramachandran outliers to report.

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	554/628 (88%)	547 (99%)	7 (1%)	65	82
2	B	315/370 (85%)	311 (99%)	4 (1%)	65	82
3	C	151/212 (71%)	147 (97%)	4 (3%)	41	68
4	E	7/13 (54%)	7 (100%)	0	100	100
All	All	1027/1223 (84%)	1012 (98%)	15 (2%)	60	79

5 of 15 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	27	GLN
3	C	177	LYS
2	B	115	ARG
3	C	184	GLU
3	C	77	ARG

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 oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 5 ligands modelled in this entry, 1 is monoatomic - leaving 4 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z > 2$	Counts	RMSZ	$\# Z > 2$
5	NAG	B	501	2	14,14,15	0.47	0	17,19,21	0.54	0
5	NAG	B	502	2	14,14,15	0.29	0	17,19,21	0.40	0
5	NAG	A	801	1	14,14,15	0.45	0	17,19,21	0.37	0
5	NAG	A	802	1	14,14,15	0.20	0	17,19,21	0.46	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	NAG	B	501	2	-	2/6/23/26	0/1/1/1
5	NAG	B	502	2	-	4/6/23/26	0/1/1/1
5	NAG	A	801	1	-	2/6/23/26	0/1/1/1
5	NAG	A	802	1	-	2/6/23/26	0/1/1/1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 10 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	B	501	NAG	C4-C5-C6-O6
5	B	501	NAG	O5-C5-C6-O6
5	B	502	NAG	C4-C5-C6-O6
5	A	802	NAG	O5-C5-C6-O6
5	A	801	NAG	C8-C7-N2-C2

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	B	501	NAG	1	0

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.

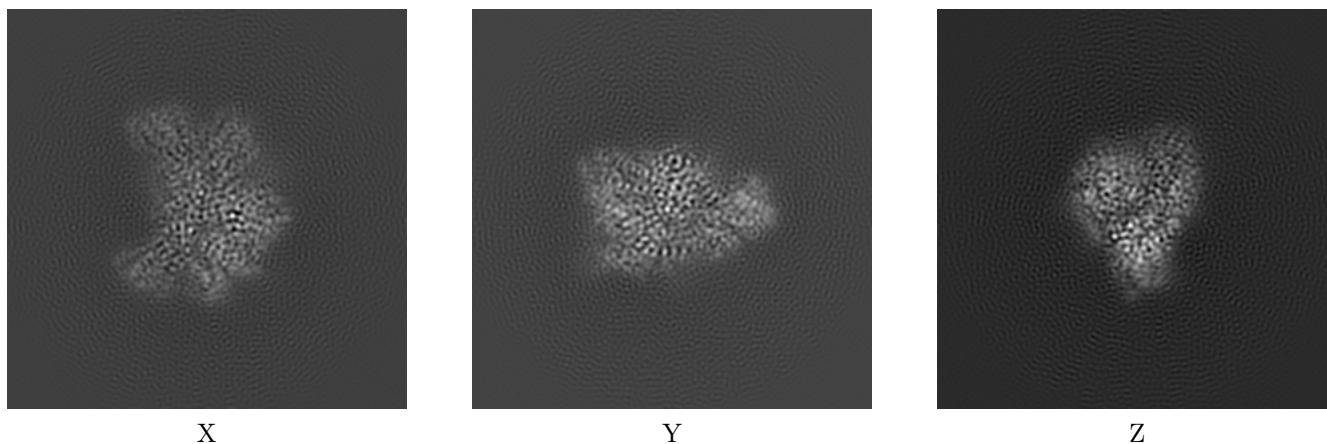
6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-36787. These allow visual inspection of the internal detail of the map and identification of artifacts.

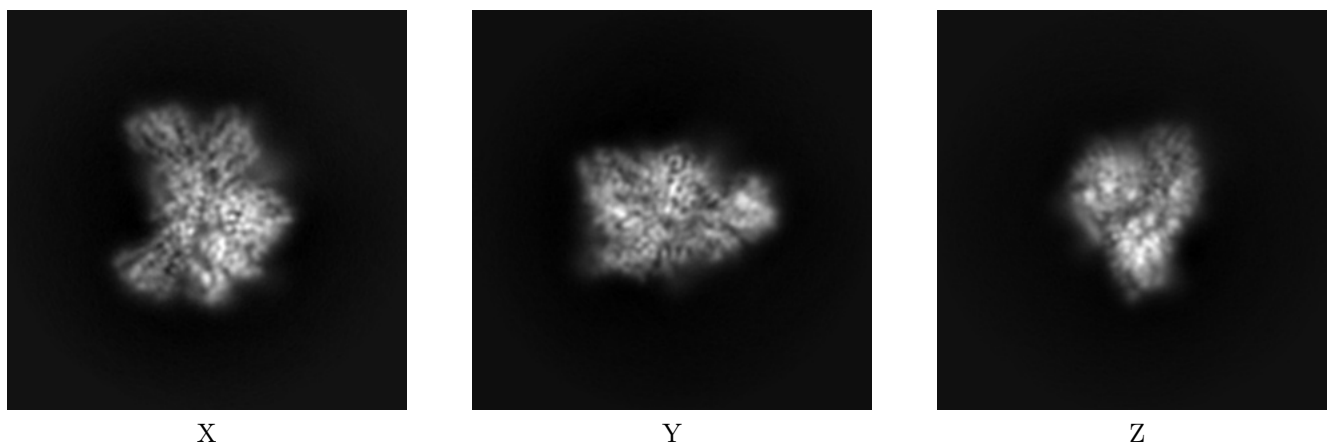
Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

6.1 Orthogonal projections [i](#)

6.1.1 Primary map



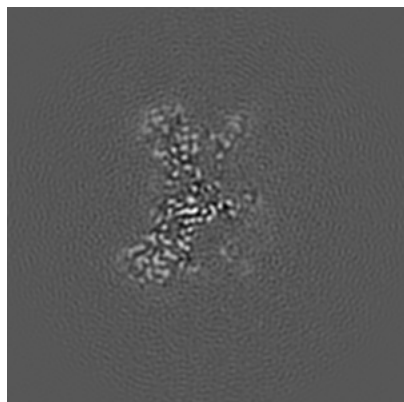
6.1.2 Raw map



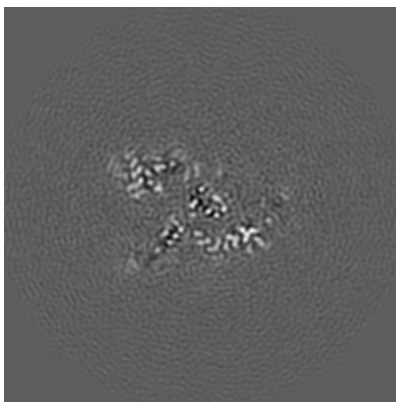
The images above show the map projected in three orthogonal directions.

6.2 Central slices [i](#)

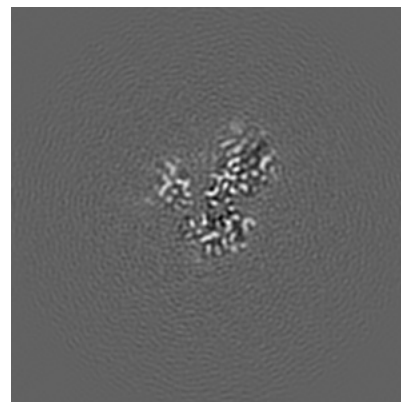
6.2.1 Primary map



X Index: 100

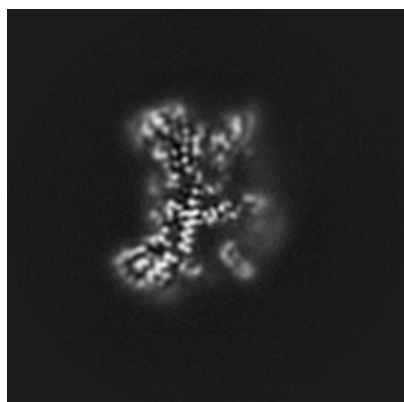


Y Index: 100

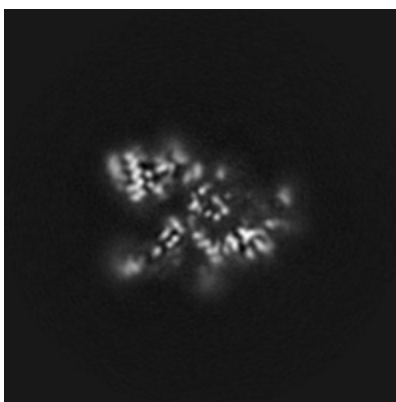


Z Index: 100

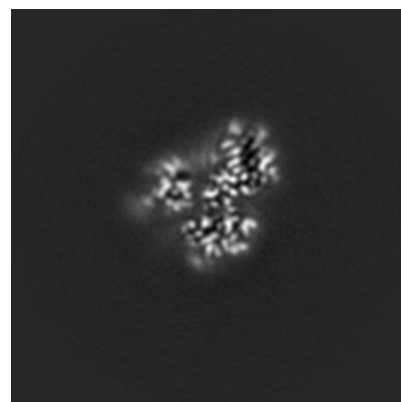
6.2.2 Raw map



X Index: 100



Y Index: 100

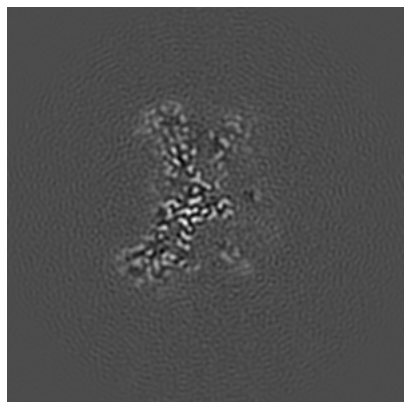


Z Index: 100

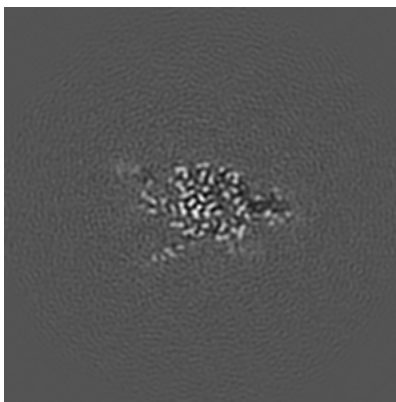
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices [i](#)

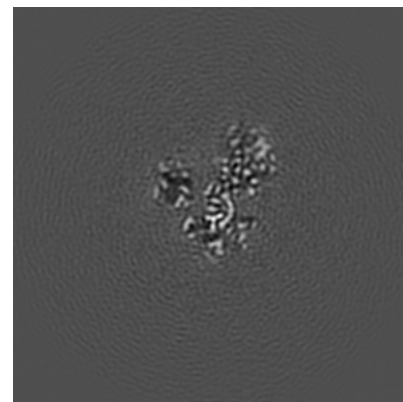
6.3.1 Primary map



X Index: 99

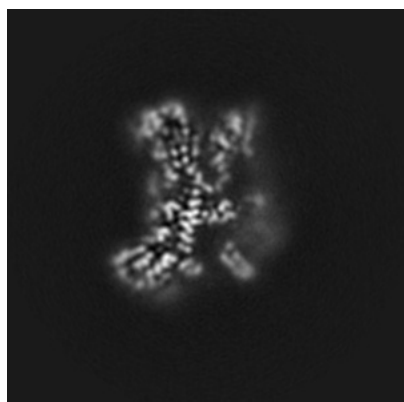


Y Index: 91

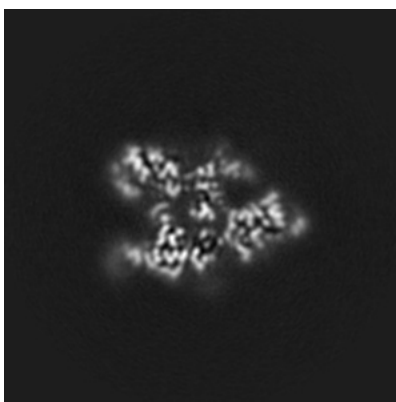


Z Index: 98

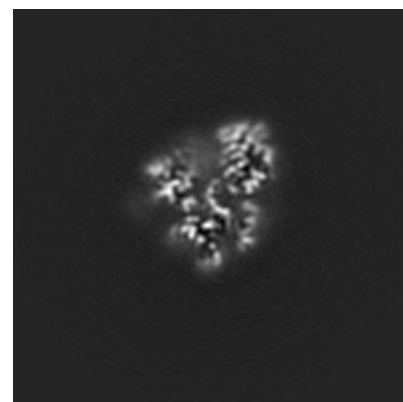
6.3.2 Raw map



X Index: 99



Y Index: 107

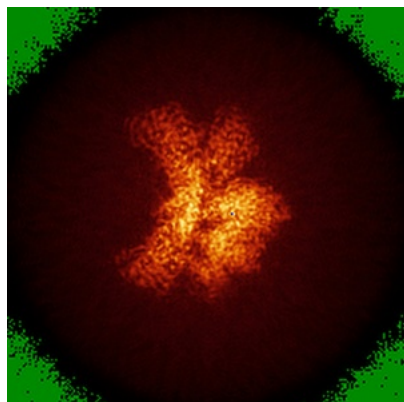


Z Index: 95

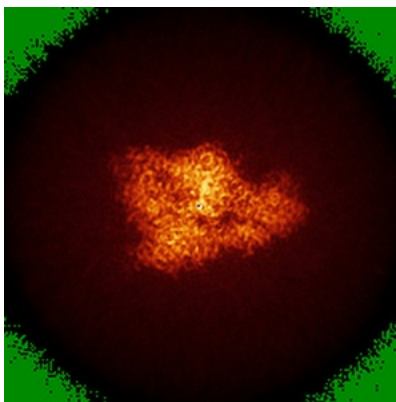
The images above show the largest variance slices of the map in three orthogonal directions.

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

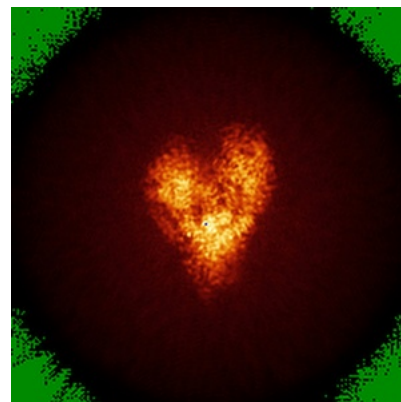
6.4.1 Primary map



X

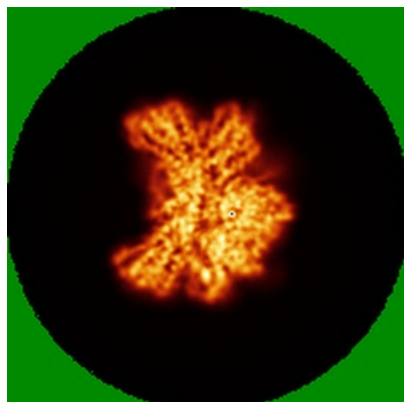


Y

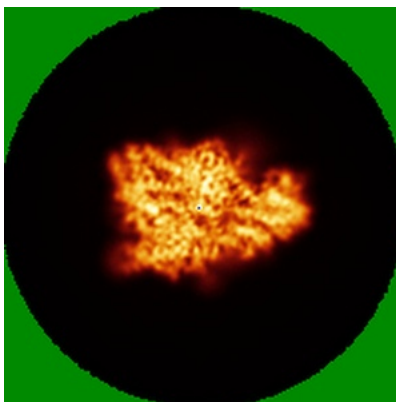


Z

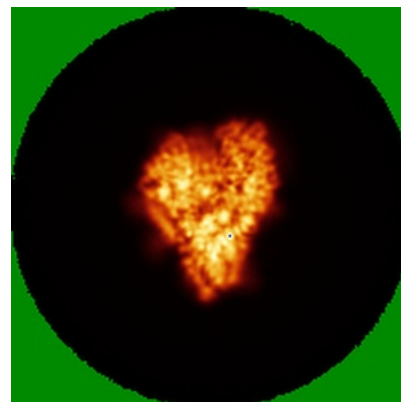
6.4.2 Raw map



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



X



Y



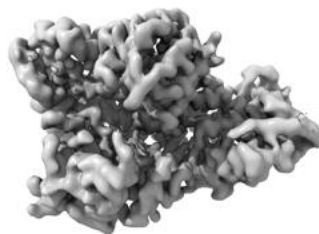
Z

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.

6.5.2 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.

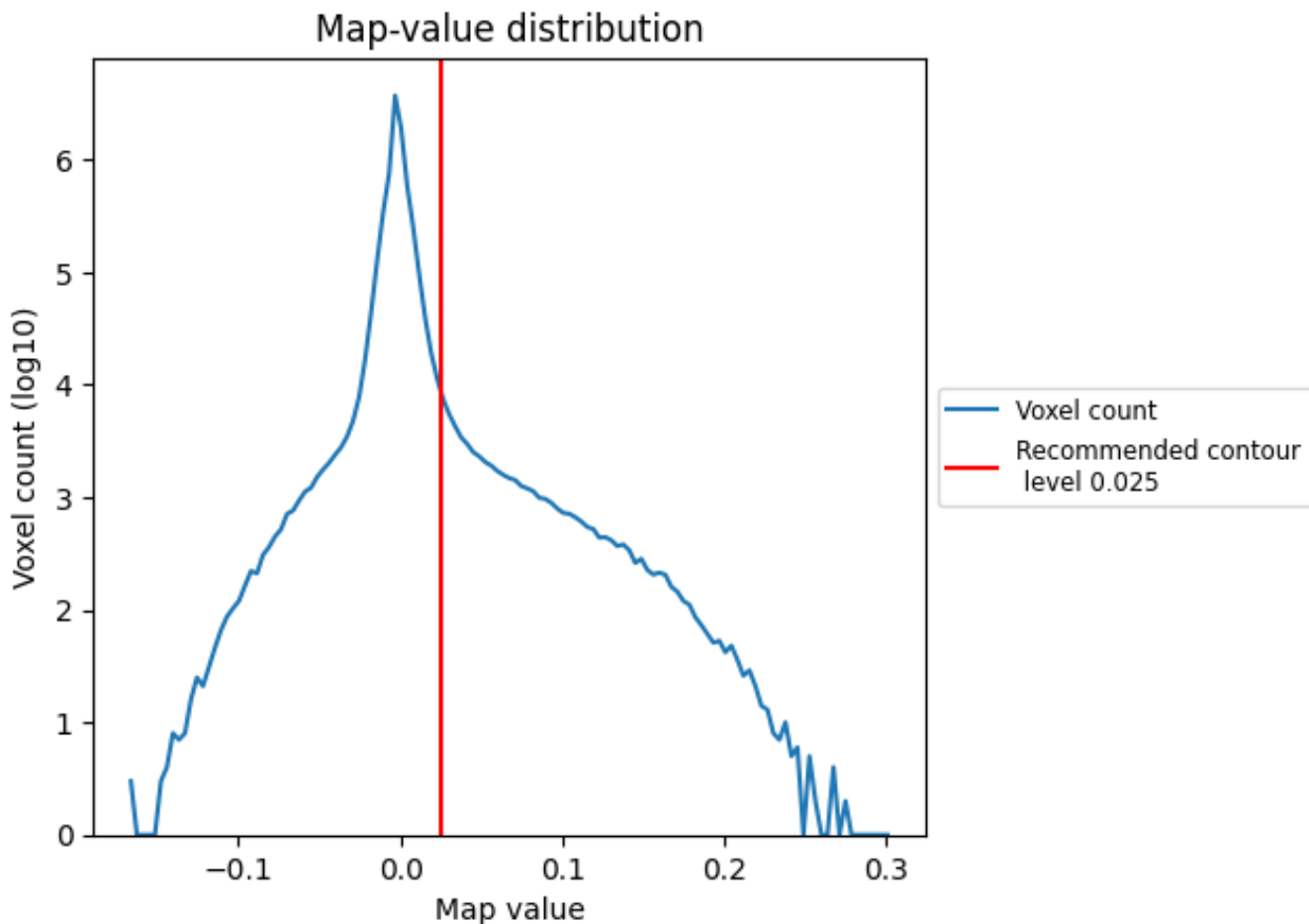
6.6 Mask visualisation [i](#)

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

7 Map analysis [i](#)

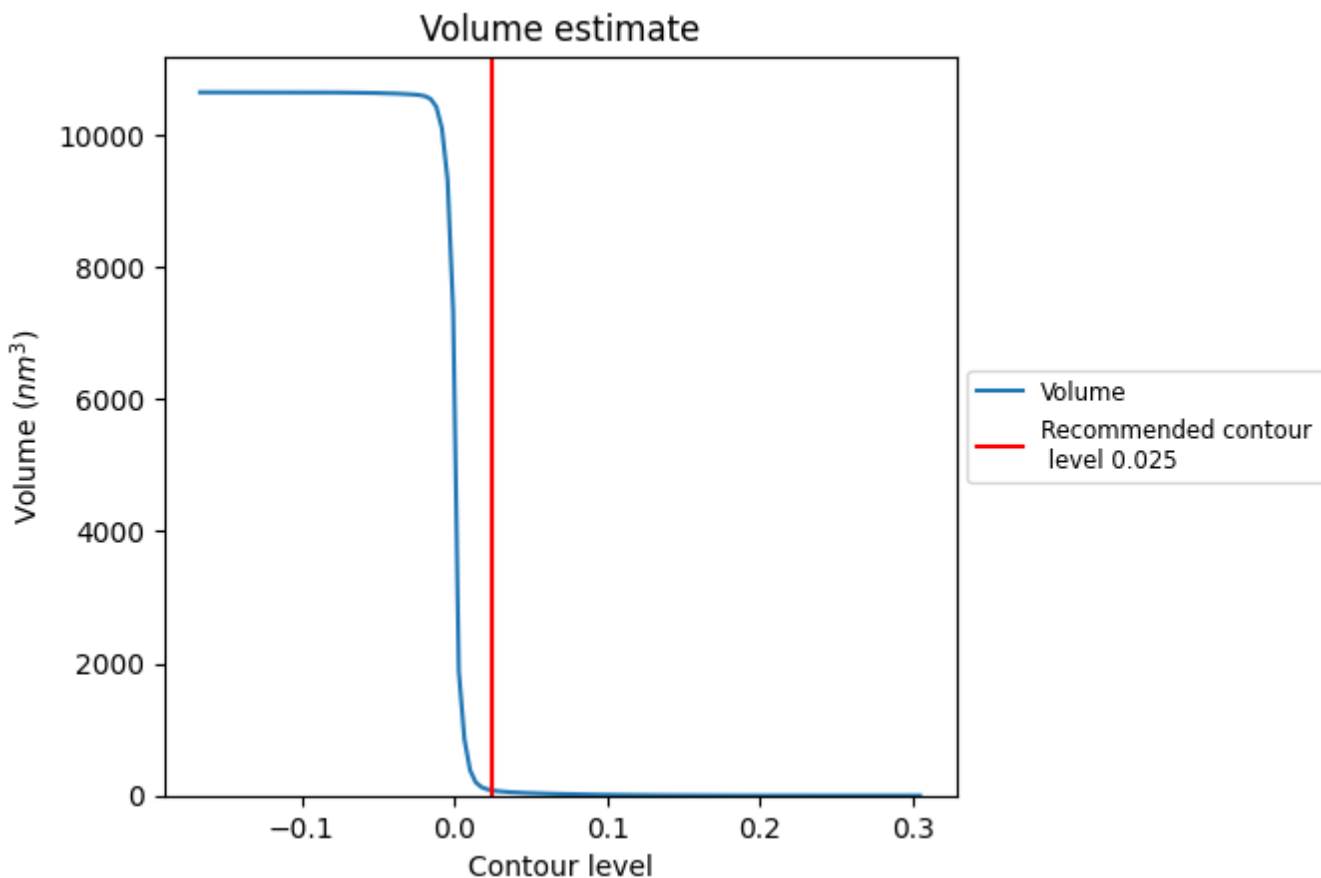
This section contains the results of statistical analysis of the map.

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.

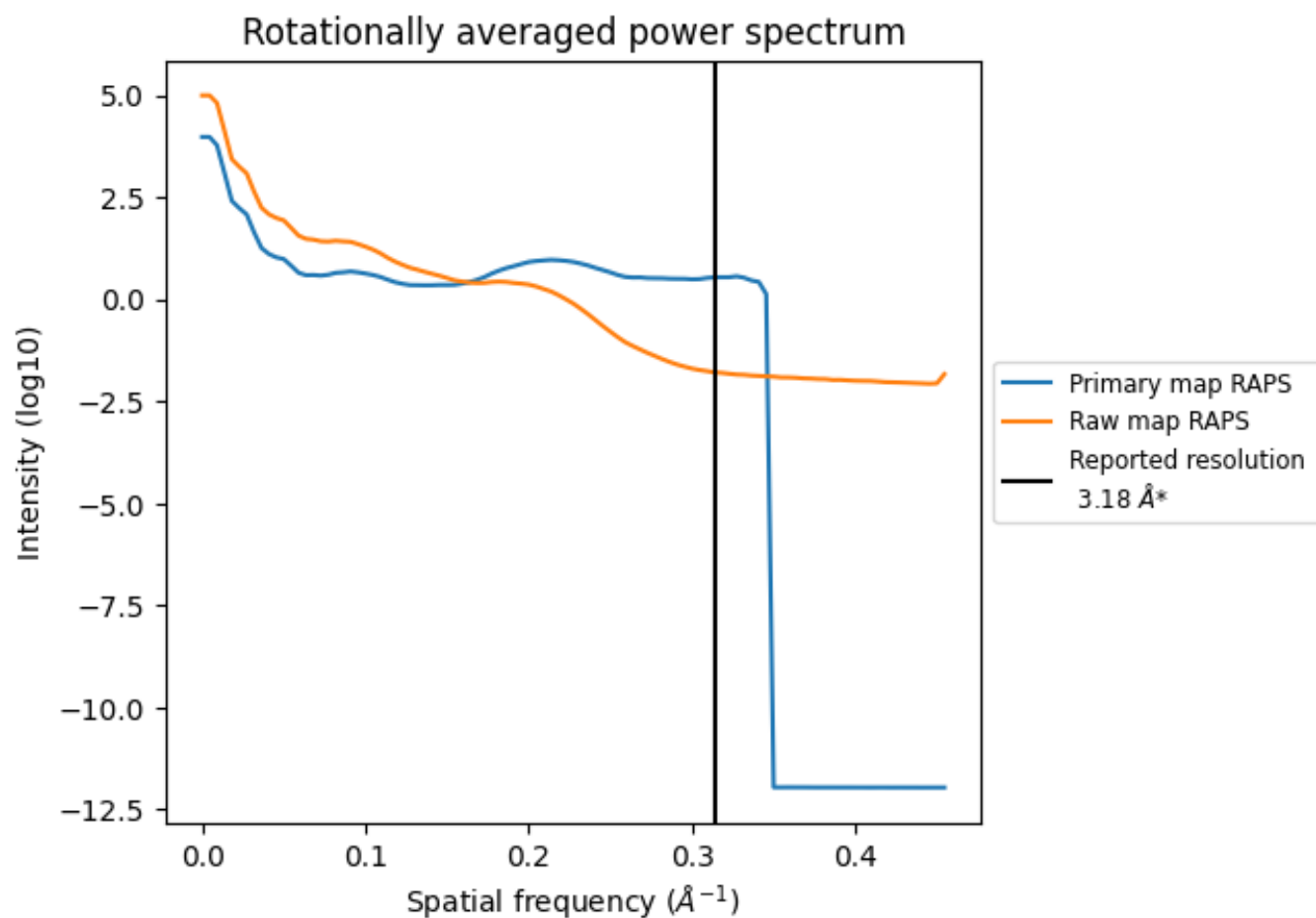
7.2 Volume estimate [i](#)



The volume at the recommended contour level is 78 nm^3 ; this corresponds to an approximate mass of 70 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](#)

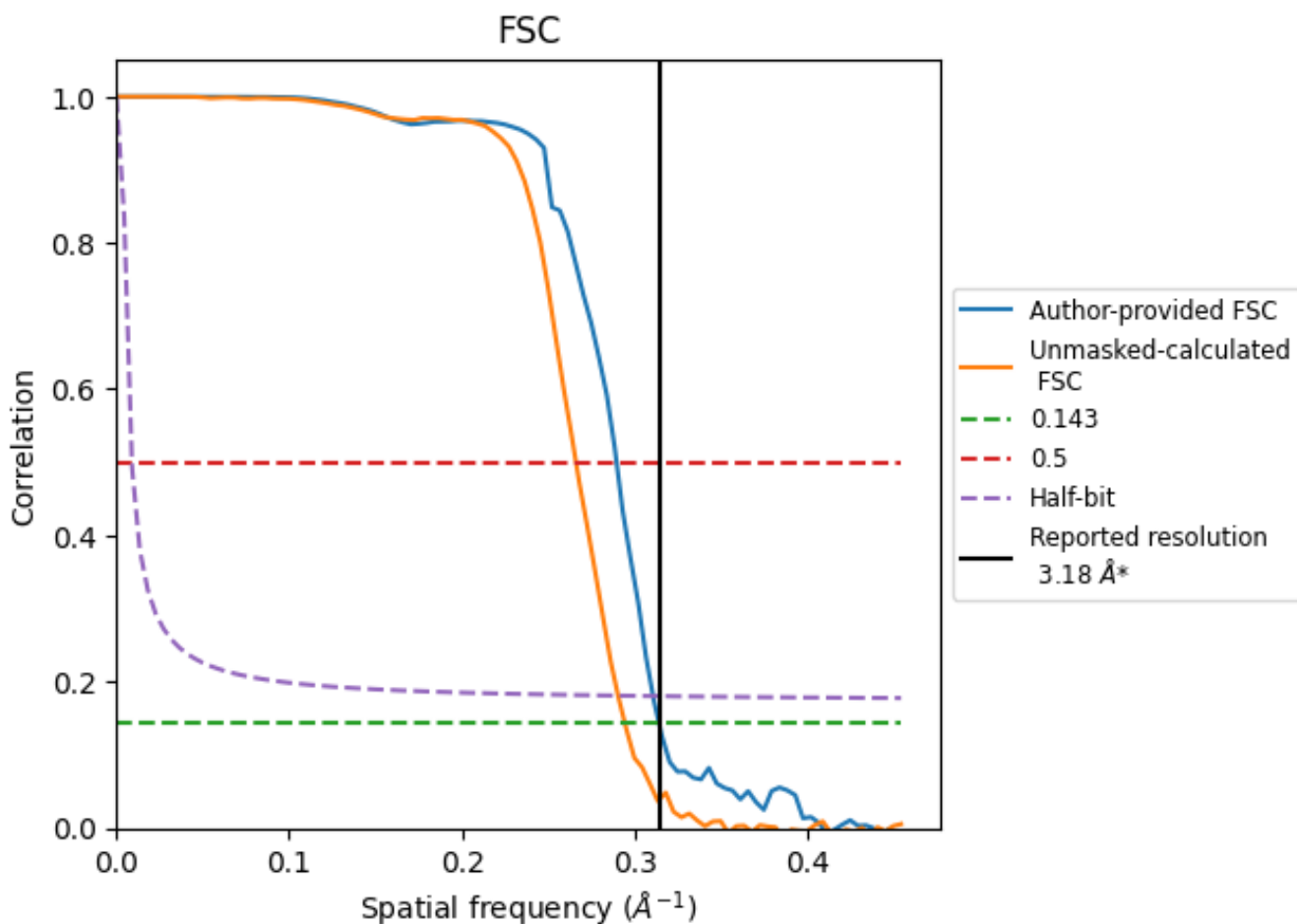


*Reported resolution corresponds to spatial frequency of 0.314 Å⁻¹

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 0.314 \AA^{-1}

8.2 Resolution estimates [i](#)

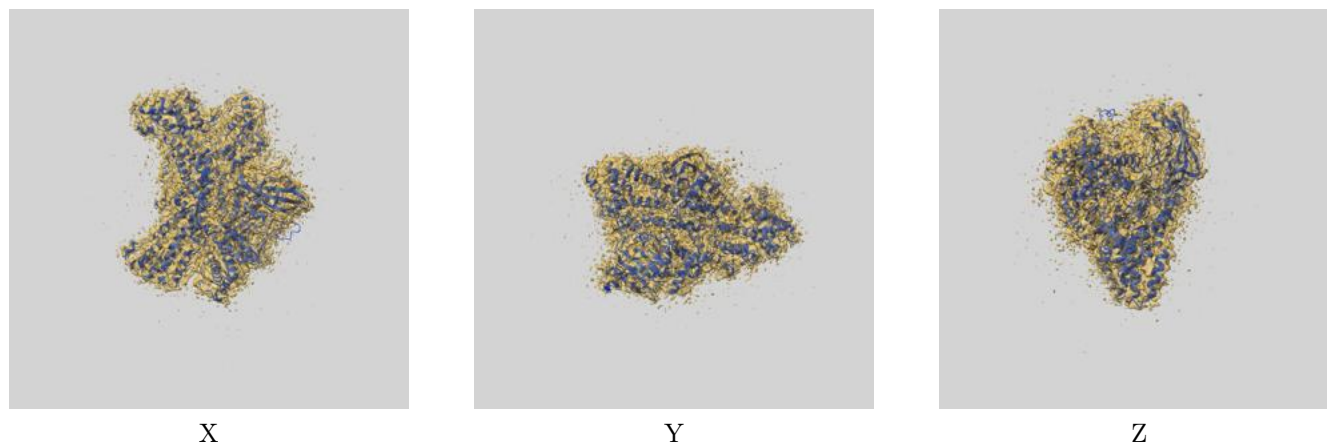
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	3.18	-	-
Author-provided FSC curve	3.18	3.45	3.22
Unmasked-calculated*	3.40	3.76	3.44

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

9 Map-model fit [i](#)

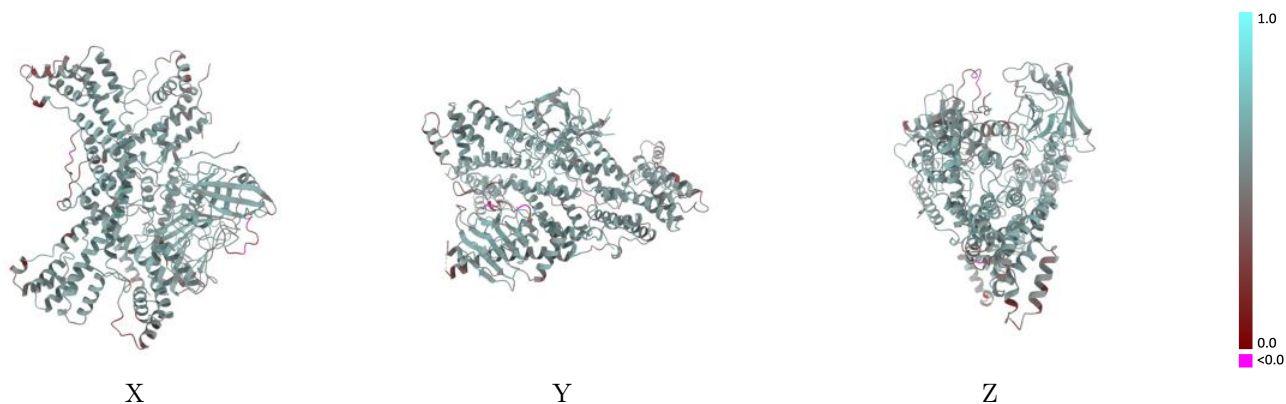
This section contains information regarding the fit between EMDB map EMD-36787 and PDB model 8K17. Per-residue inclusion information can be found in section [3](#) on page [7](#).

9.1 Map-model overlay [i](#)



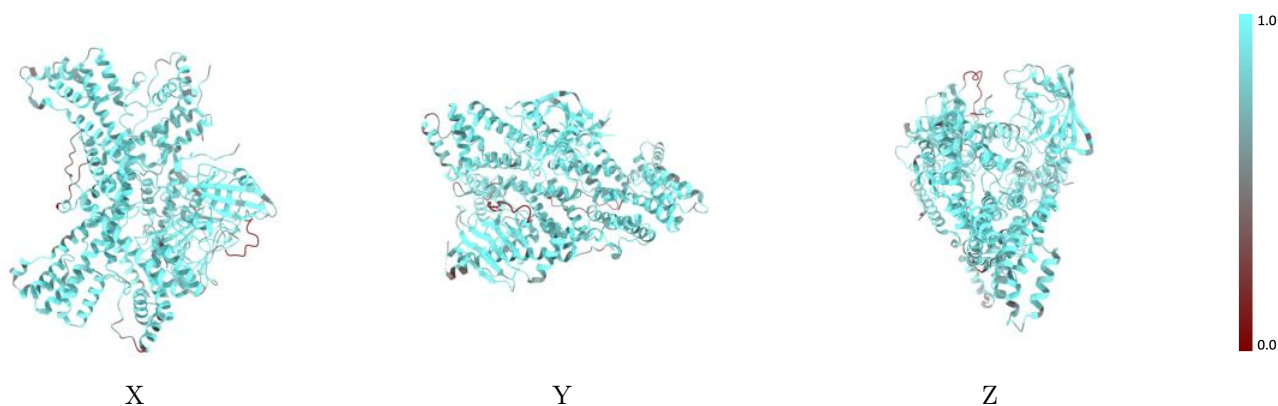
The images above show the 3D surface view of the map at the recommended contour level 0.025 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

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



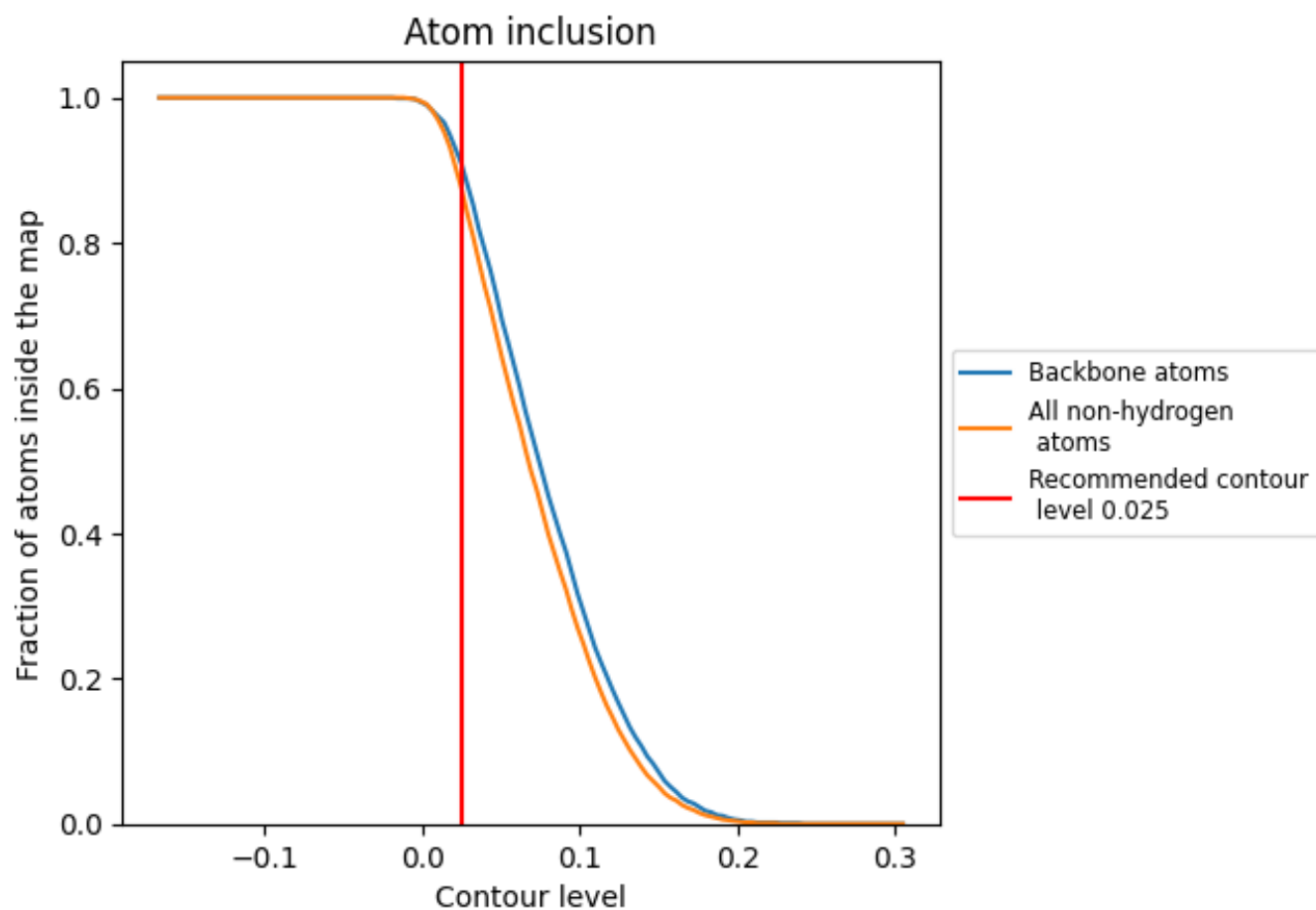
The images above show the model with each residue coloured according to its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions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










9.4 Atom inclusion [i](#)



At the recommended contour level, 91% of all backbone atoms, 88% of all non-hydrogen atoms, are inside the map.

9.5 Map-model fit summary [i](#)

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	 0.8750	 0.5540
A	 0.8650	 0.5500
B	 0.8770	 0.5500
C	 0.9110	 0.5820
E	 0.6880	 0.4840

