W O R L D W I D E
wwPDB
P R O T E I N   D A T A   B A N K

wwPDB EM Validation Summary Report (i)

Mar 4, 2024 – 04:55 pm GMT

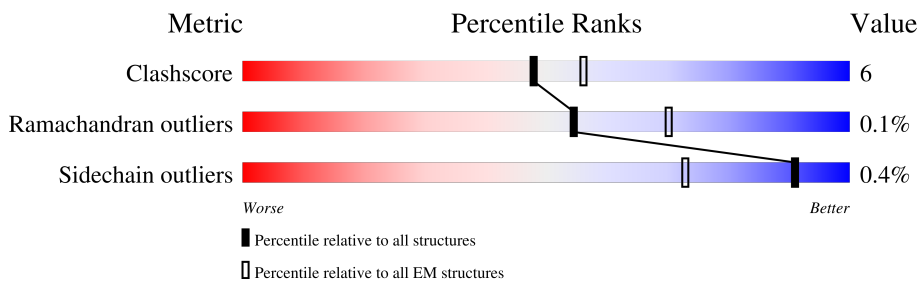|		
PDB ID	:	8OK2
EMDB ID	:	EMD-16916
Title	:	Bipartite interaction of TOPBP1 with the GINS complex
Authors	:	Day, M.; Oliver, A.W.; Pearl, L.H.
Deposited on	:	2023-03-26
Resolution	:	4.10 Å(reported)

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 4.10 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	Percentile Ranks	Value
Clashscore		6
Ramachandran outliers		0.1%
Sidechain outliers		0.4%

Worse .... *Better*

▮ Percentile relative to all structures
▯ Percentile relative to all EM structures

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 <=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	151	77% .... 19% •
2	B	185	78% 16% 5%
3	C	216	76% 14% 10%
4	D	262	63% 14% 23%
5	E	470	38% 5% • 56%

2   Entry composition ⓘ

There are 5 unique types of molecules in this entry. The entry contains 7430 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 DNA replication complex GINS protein PSF1.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	145	Total	C	N	O	S	0	0
			1188	744	217	217	10		

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	97	ILE	VAL	engineered mutation	UNP Q14691

- Molecule 2 is a protein called DNA replication complex GINS protein PSF2.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	175	Total	C	N	O	S	0	0
			1424	911	241	263	9		

- Molecule 3 is a protein called DNA replication complex GINS protein PSF3.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	C	195	Total	C	N	O	S	0	0
			1551	980	268	296	7		

- Molecule 4 is a protein called DNA replication complex GINS protein SLD5.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	D	203	Total	C	N	O	S	0	0
			1678	1065	290	313	10		

There are 39 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	-38	MET	-	initiating methionine	UNP Q9BRT9
D	-37	HIS	-	expression tag	UNP Q9BRT9
D	-36	HIS	-	expression tag	UNP Q9BRT9

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	-35	HIS	-	expression tag	UNP Q9BRT9
D	-34	HIS	-	expression tag	UNP Q9BRT9
D	-33	HIS	-	expression tag	UNP Q9BRT9
D	-32	HIS	-	expression tag	UNP Q9BRT9
D	-31	GLY	-	expression tag	UNP Q9BRT9
D	-30	ARG	-	expression tag	UNP Q9BRT9
D	-29	MET	-	expression tag	UNP Q9BRT9
D	-28	ASP	-	expression tag	UNP Q9BRT9
D	-27	TYR	-	expression tag	UNP Q9BRT9
D	-26	LYS	-	expression tag	UNP Q9BRT9
D	-25	ASP	-	expression tag	UNP Q9BRT9
D	-24	ASP	-	expression tag	UNP Q9BRT9
D	-23	ASP	-	expression tag	UNP Q9BRT9
D	-22	ASP	-	expression tag	UNP Q9BRT9
D	-21	LYS	-	expression tag	UNP Q9BRT9
D	-20	ALA	-	expression tag	UNP Q9BRT9
D	-19	ASP	-	expression tag	UNP Q9BRT9
D	-18	TYR	-	expression tag	UNP Q9BRT9
D	-17	LYS	-	expression tag	UNP Q9BRT9
D	-16	ASP	-	expression tag	UNP Q9BRT9
D	-15	ASP	-	expression tag	UNP Q9BRT9
D	-14	ASP	-	expression tag	UNP Q9BRT9
D	-13	ASP	-	expression tag	UNP Q9BRT9
D	-12	LYS	-	expression tag	UNP Q9BRT9
D	-11	ALA	-	expression tag	UNP Q9BRT9
D	-10	ASP	-	expression tag	UNP Q9BRT9
D	-9	TYR	-	expression tag	UNP Q9BRT9
D	-8	LYS	-	expression tag	UNP Q9BRT9
D	-7	ASP	-	expression tag	UNP Q9BRT9
D	-6	ASP	-	expression tag	UNP Q9BRT9
D	-5	ASP	-	expression tag	UNP Q9BRT9
D	-4	ASP	-	expression tag	UNP Q9BRT9
D	-3	LYS	-	expression tag	UNP Q9BRT9
D	-2	GLY	-	expression tag	UNP Q9BRT9
D	-1	ARG	-	expression tag	UNP Q9BRT9
D	0	PRO	-	expression tag	UNP Q9BRT9

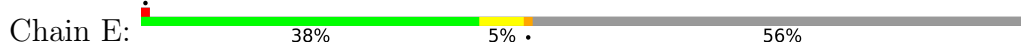- Molecule 5 is a protein called Topoisomerase (DNA) II binding protein 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
5	E	206	1589	1010	259	312	8	0	0

There are 34 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	308	MET	-	initiating methionine	UNP A7E2X7
E	309	GLY	-	expression tag	UNP A7E2X7
E	310	SER	-	expression tag	UNP A7E2X7
E	311	HIS	-	expression tag	UNP A7E2X7
E	312	HIS	-	expression tag	UNP A7E2X7
E	313	HIS	-	expression tag	UNP A7E2X7
E	314	HIS	-	expression tag	UNP A7E2X7
E	315	HIS	-	expression tag	UNP A7E2X7
E	316	HIS	-	expression tag	UNP A7E2X7
E	317	GLY	-	expression tag	UNP A7E2X7
E	318	SER	-	expression tag	UNP A7E2X7
E	319	LEU	-	expression tag	UNP A7E2X7
E	320	GLU	-	expression tag	UNP A7E2X7
E	321	VAL	-	expression tag	UNP A7E2X7
E	322	LEU	-	expression tag	UNP A7E2X7
E	323	PHE	-	expression tag	UNP A7E2X7
E	324	GLN	-	expression tag	UNP A7E2X7
E	325	GLY	-	expression tag	UNP A7E2X7
E	326	PRO	-	expression tag	UNP A7E2X7
E	327	HIS	-	expression tag	UNP A7E2X7
E	328	MET	-	expression tag	UNP A7E2X7
E	329	ALA	-	expression tag	UNP A7E2X7
E	330	SER	-	expression tag	UNP A7E2X7
E	767	ARG	-	expression tag	UNP A7E2X7
E	768	ARG	-	expression tag	UNP A7E2X7
E	769	ALA	-	expression tag	UNP A7E2X7
E	770	TRP	-	expression tag	UNP A7E2X7
E	771	SER	-	expression tag	UNP A7E2X7
E	772	HIS	-	expression tag	UNP A7E2X7
E	773	PRO	-	expression tag	UNP A7E2X7
E	774	GLN	-	expression tag	UNP A7E2X7
E	775	PHE	-	expression tag	UNP A7E2X7
E	776	GLU	-	expression tag	UNP A7E2X7
E	777	LYS	-	expression tag	UNP A7E2X7

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: DNA replication complex GINS protein PSF1

Chain A:

- Molecule 2: DNA replication complex GINS protein PSF2

Chain B:

- Molecule 3: DNA replication complex GINS protein PSF3

Chain C:

- Molecule 4: DNA replication complex GINS protein SLD5

Chain D:

R192 V195 I196 Q203 Q221 L222 I223

• Molecule 5: Topoisomerase (DNA) II binding protein 1

Chain E:

38% 5% • 56%

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

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

LEU GLU CYS PHE SER LYS GLY TYR MET LEU SER GLU GLU PRO TYR ILE HIS ALA ASN TYR GLN PRO VAL GLU ILE PRO VAL SER HIS GLN PRO GLU SER LYS ALA ALA LEU LEU LYS LYS LYS ASN SER SER PHE SER LYS LYS ASP PHE ALA PRO SER GLU LYS HIS E484 Q485 A486 D487

E488 D489 Q493 Y494 E495 N496 G497 SER SER THR VAL VAL GLU ALA LYS THR SER GLU ALA ARG PRO PHE ASN ASP SER THR HIS ALA GLU PRO LEU ASN ASP SER THR HIS ILE SER LEU GLN GLU GLU ASN GLN SER SER VAL SER VAL SER HIS CYS VAL PRO ASP VAL SER THR GLU

E550 S554 Q555 K556 S569 N573 N578 I582 R588 D592 Y593 A594 T606 V607 V610 N613 I620 L633 P645 L646 R681 R700 W711 W720 A731 E739 N740 S741 THR LYS GLU GLU ARG SER LEU GLU GLU THR GLU ILE THR ASN

GLY ILE ASN LEU ASN SER ASP THR ALA GLU HIS PRO ARG ARG ALA TRP SER HIS PRO GLN PHE GLU LYS

4 Experimental information ⓘ

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	154278	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 $(e^-/\text{Å}^2)$	39.69290657	Depositor
Minimum defocus (nm)	1200	Depositor
Maximum defocus (nm)	2200	Depositor
Magnification	Not provided	
Image detector	FEI FALCON IV (4k x 4k)	Depositor
Maximum map value	0.038	Depositor
Minimum map value	-0.021	Depositor
Average map value	0.000	Depositor
Map value standard deviation	0.001	Depositor
Recommended contour level	0.003	Depositor
Map size (Å)	207.2, 207.2, 207.2	wwPDB
Map dimensions	280, 280, 280	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	0.74, 0.74, 0.74	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	0.26	0/1211	0.50	0/1626
2	B	0.24	0/1454	0.49	0/1969
3	C	0.25	0/1586	0.47	0/2140
4	D	0.25	0/1710	0.50	0/2305
5	E	0.25	0/1617	0.47	0/2190
All	All	0.25	0/7578	0.49	0/10230

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   ⓘ

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	1188	0	1180	25	0
2	B	1424	0	1449	20	0
3	C	1551	0	1502	22	0
4	D	1678	0	1700	27	0
5	E	1589	0	1577	24	0
All	All	7430	0	7408	93	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 6.

The worst 5 of 93 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:91:ARG:HH22	5:E:606:THR:HG22	1.50	0.77
1:A:97:ILE:HD11	5:E:610:VAL:HG11	1.66	0.75
2:B:1:MET:HB2	4:D:38:LYS:HD3	1.75	0.69
2:B:1:MET:HB3	2:B:154:ASN:HD21	1.57	0.68
2:B:44:VAL:HB	2:B:48:LEU:HD23	1.77	0.67

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	143/151 (95%)	140 (98%)	3 (2%)	0	100	100
2	B	173/185 (94%)	167 (96%)	6 (4%)	0	100	100
3	C	193/216 (89%)	187 (97%)	6 (3%)	0	100	100
4	D	201/262 (77%)	191 (95%)	10 (5%)	0	100	100
5	E	202/470 (43%)	199 (98%)	2 (1%)	1 (0%)	29	67
All	All	912/1284 (71%)	884 (97%)	27 (3%)	1 (0%)	54	84

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
5	E	554	SER

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	126/132 (96%)	126 (100%)	0	100	100
2	B	159/169 (94%)	159 (100%)	0	100	100
3	C	166/186 (89%)	166 (100%)	0	100	100
4	D	188/240 (78%)	188 (100%)	0	100	100
5	E	175/412 (42%)	172 (98%)	3 (2%)	60	78
All	All	814/1139 (72%)	811 (100%)	3 (0%)	91	94

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
5	E	554	SER
5	E	555	GLN
5	E	556	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 13 such sidechains are listed below:

Mol	Chain	Res	Type
3	C	154	GLN
4	D	32	GLN
5	E	496	ASN
4	D	143	ASN
4	D	154	GLN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

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

5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

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-16916. 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
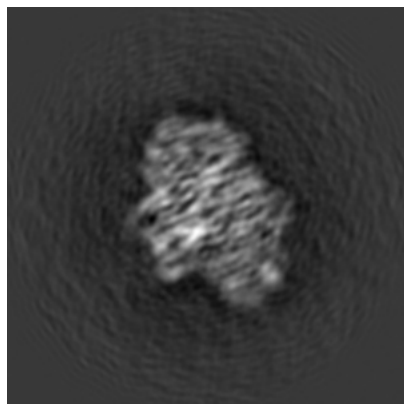
X
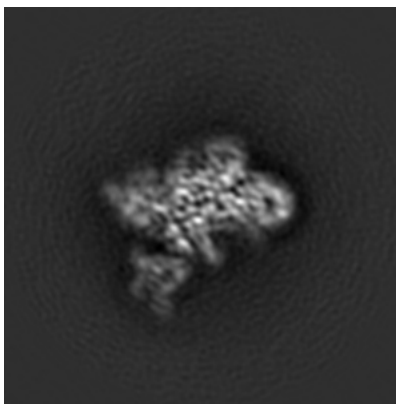
Y
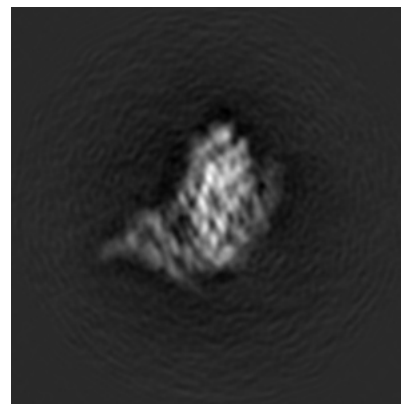
Z

6.1.2 Raw map

X

Y

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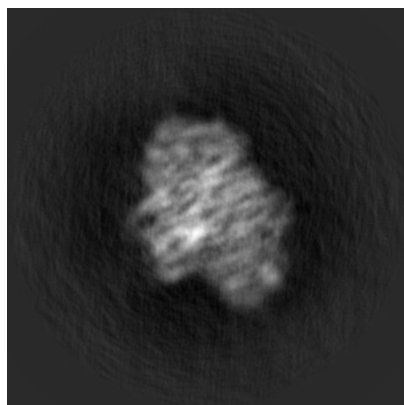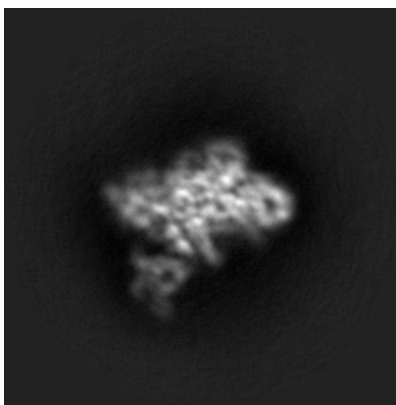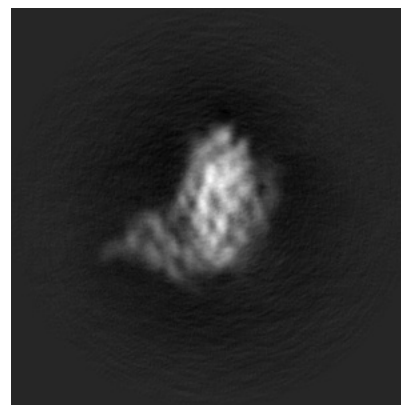
The images above show the map projected in three orthogonal directions.

6.2   Central slices ⓘ
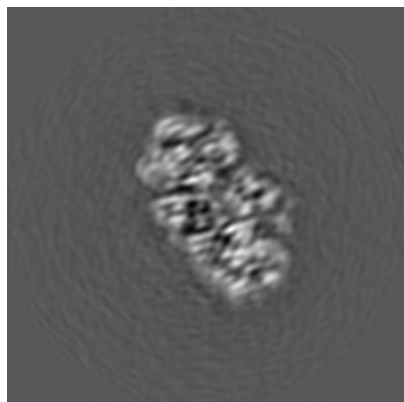
6.2.1   Primary map
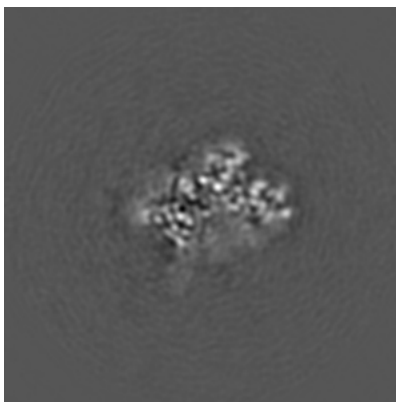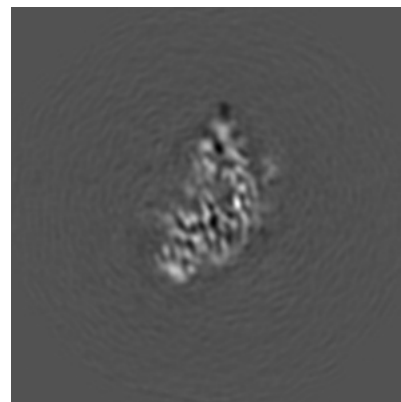
X Index: 140   Y Index: 140   Z Index: 140

6.2.2   Raw map
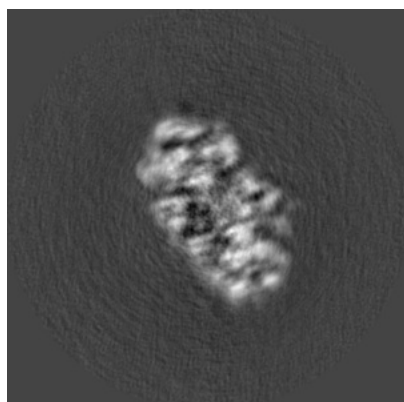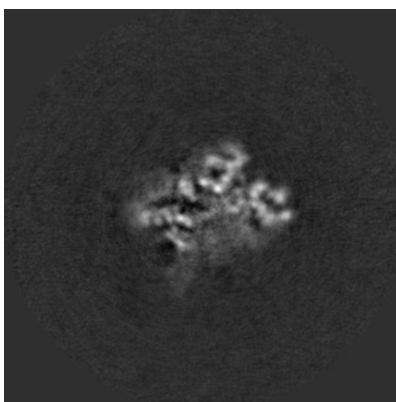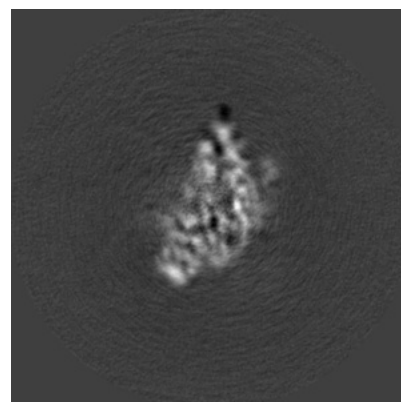
X Index: 140   Y Index: 140   Z Index: 140

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

6.3 Largest variance slices ⓘ

6.3.1 Primary map
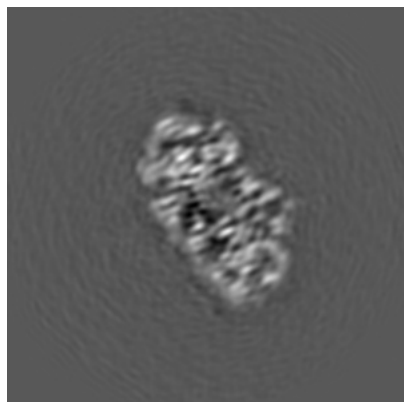
X Index: 142
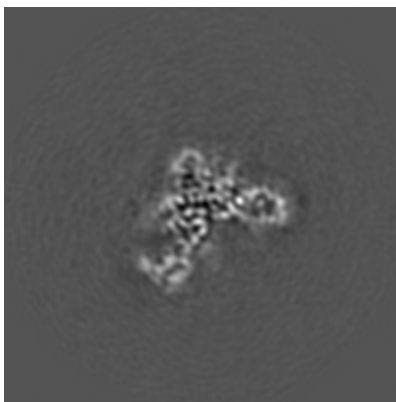
Y Index: 129
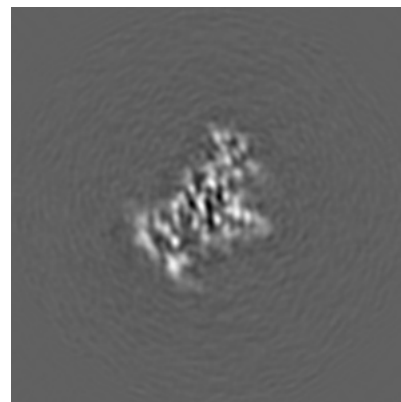
Z Index: 124

6.3.2 Raw map
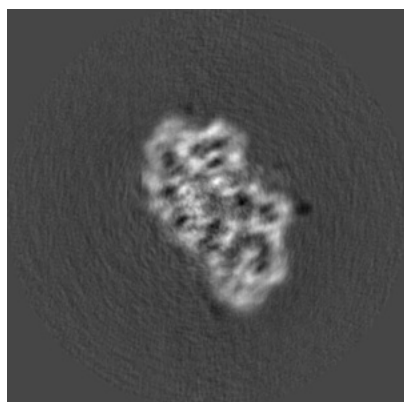
X Index: 147
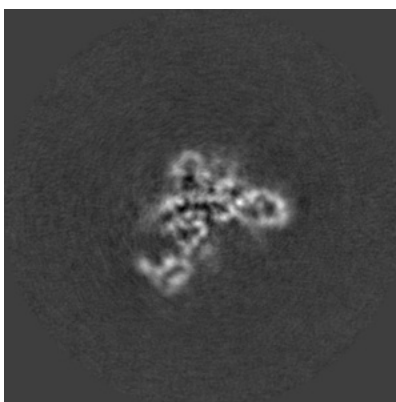
Y Index: 129
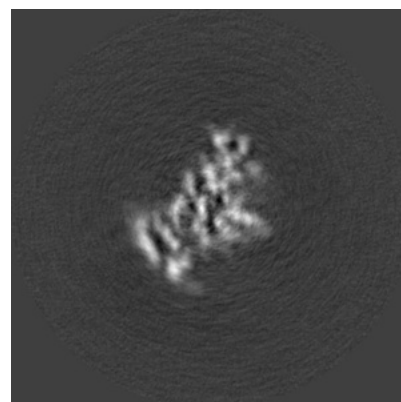
Z Index: 123

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
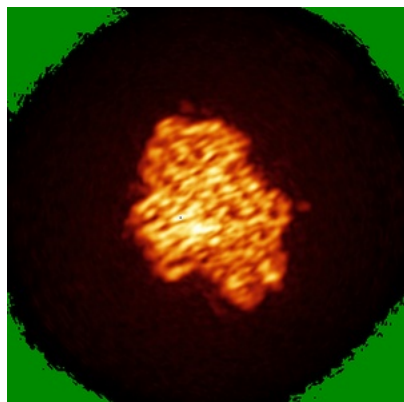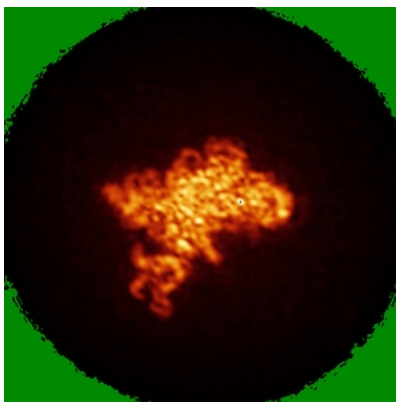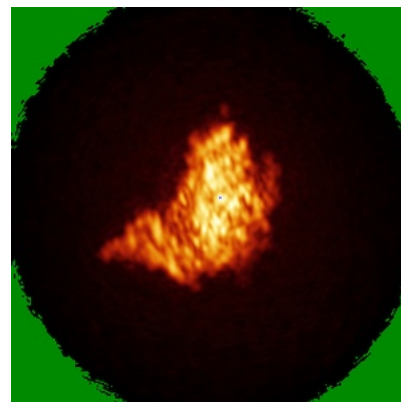
| X | Y | Z |

6.4.2 Raw map

| X | Y | Z |
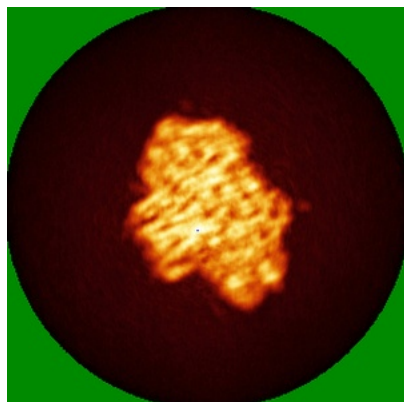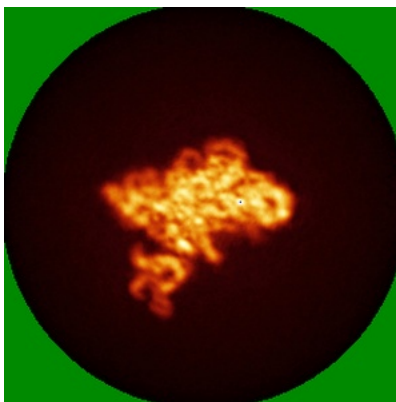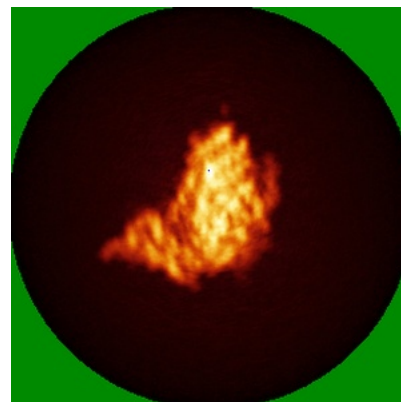
The images above show the map standard deviation projections with false color in three orthogonal directions. Minimum values are shown in green, max in blue, and dark to light orange shades represent small to large values respectively.

6.5 Orthogonal surface views ⓘ

6.5.1 Primary map

| X | Y | Z |

The images above show the 3D surface view of the map at the recommended contour level 0.003. 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 |
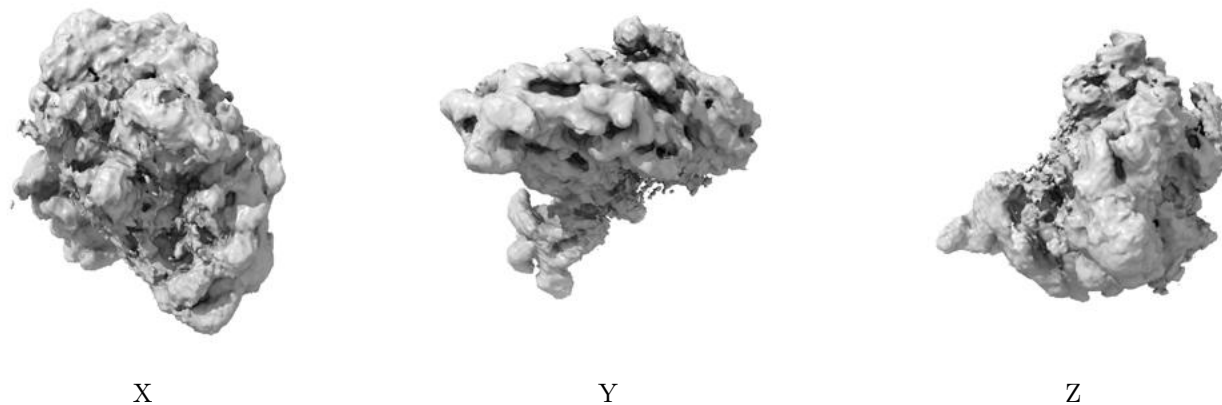
These images show the 3D surface of the raw map. The raw map's contour level was selected so that its surface encloses the same volume as the primary map does at its recommended contour level.

6.6 Mask visualisation ⓘ

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

7 Map analysis ⓘ

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

7.1 Map-value distribution ⓘ
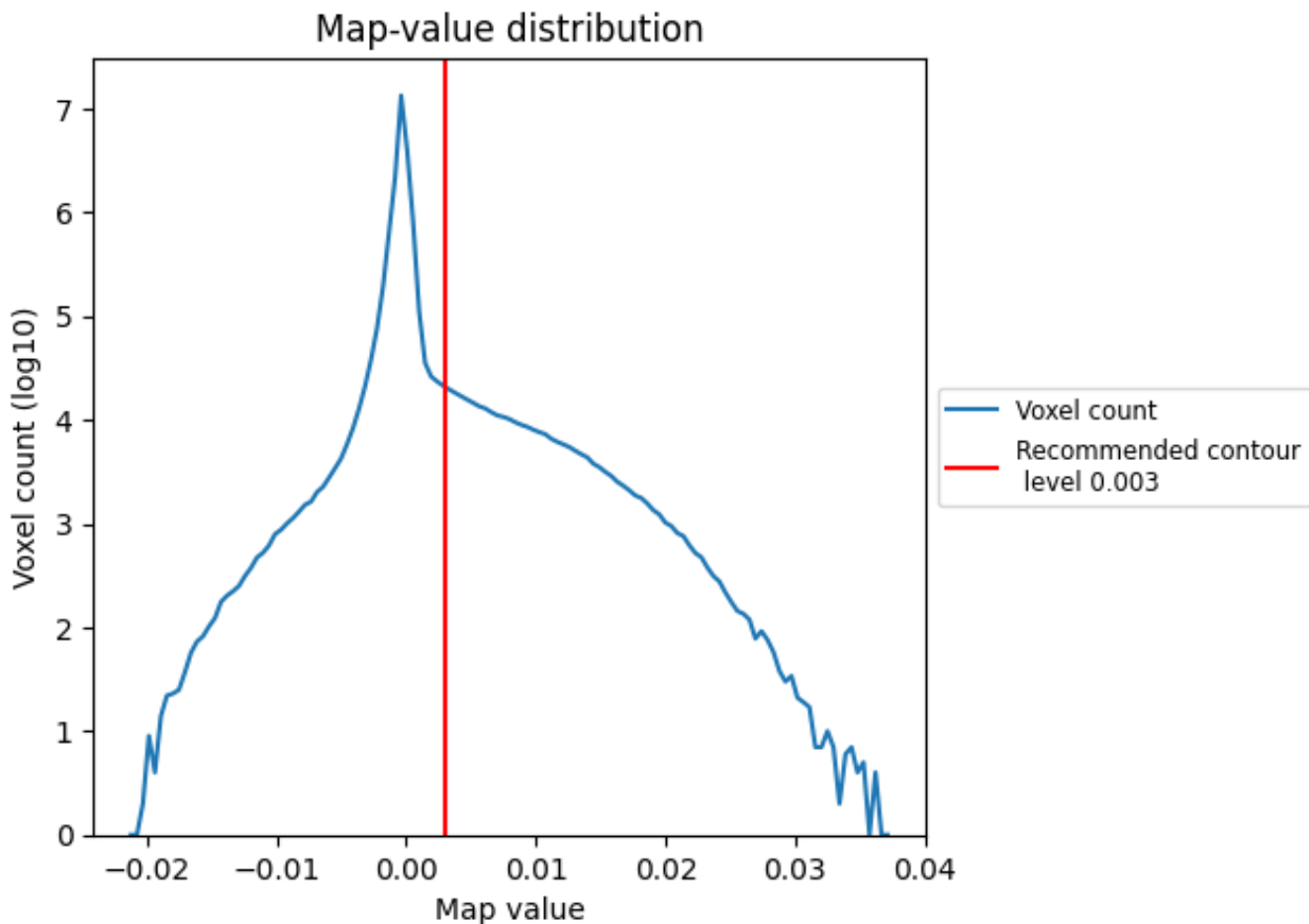
The map-value distribution is plotted in 128 intervals along the x-axis. The y-axis is logarithmic. A spike in this graph at zero usually indicates that the volume has been masked.

7.2 Volume estimate ⓘ
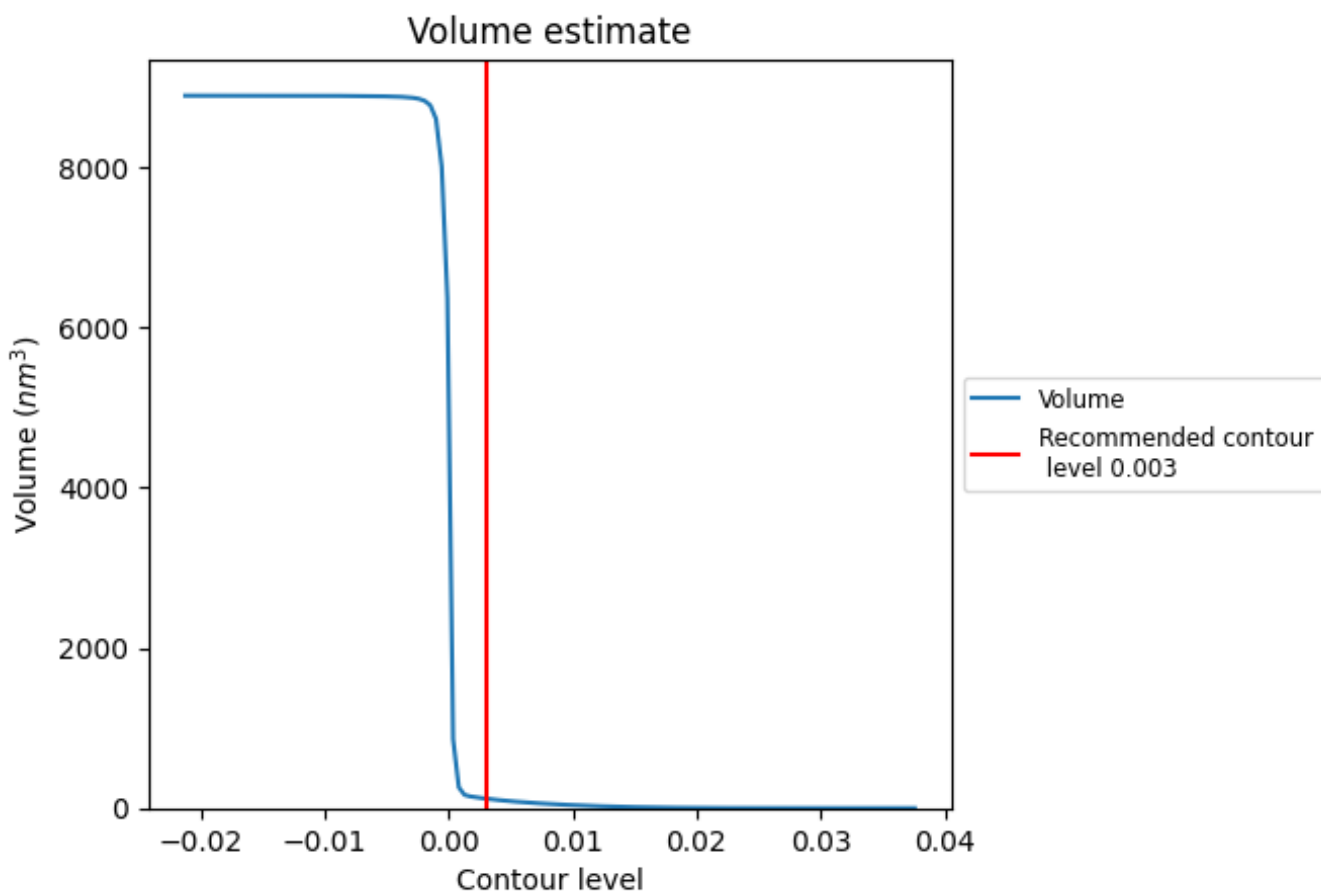
The volume at the recommended contour level is 119 nm^3; this corresponds to an approximate mass of 107 kDa.

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 ⓘ
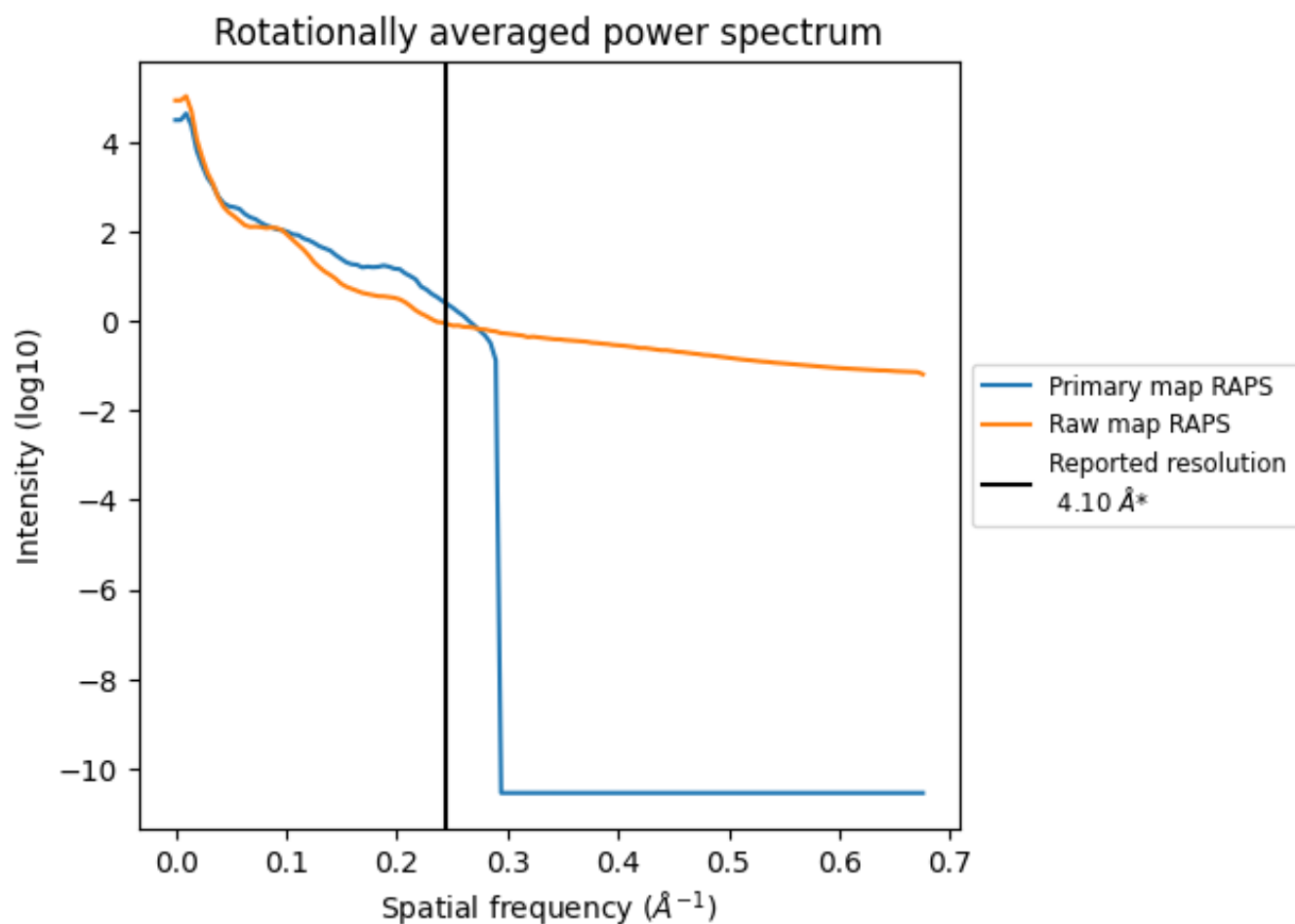
*Reported resolution corresponds to spatial frequency of 0.244 Å^{-1}

8   Fourier-Shell correlation ⓘ

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 ⓘ
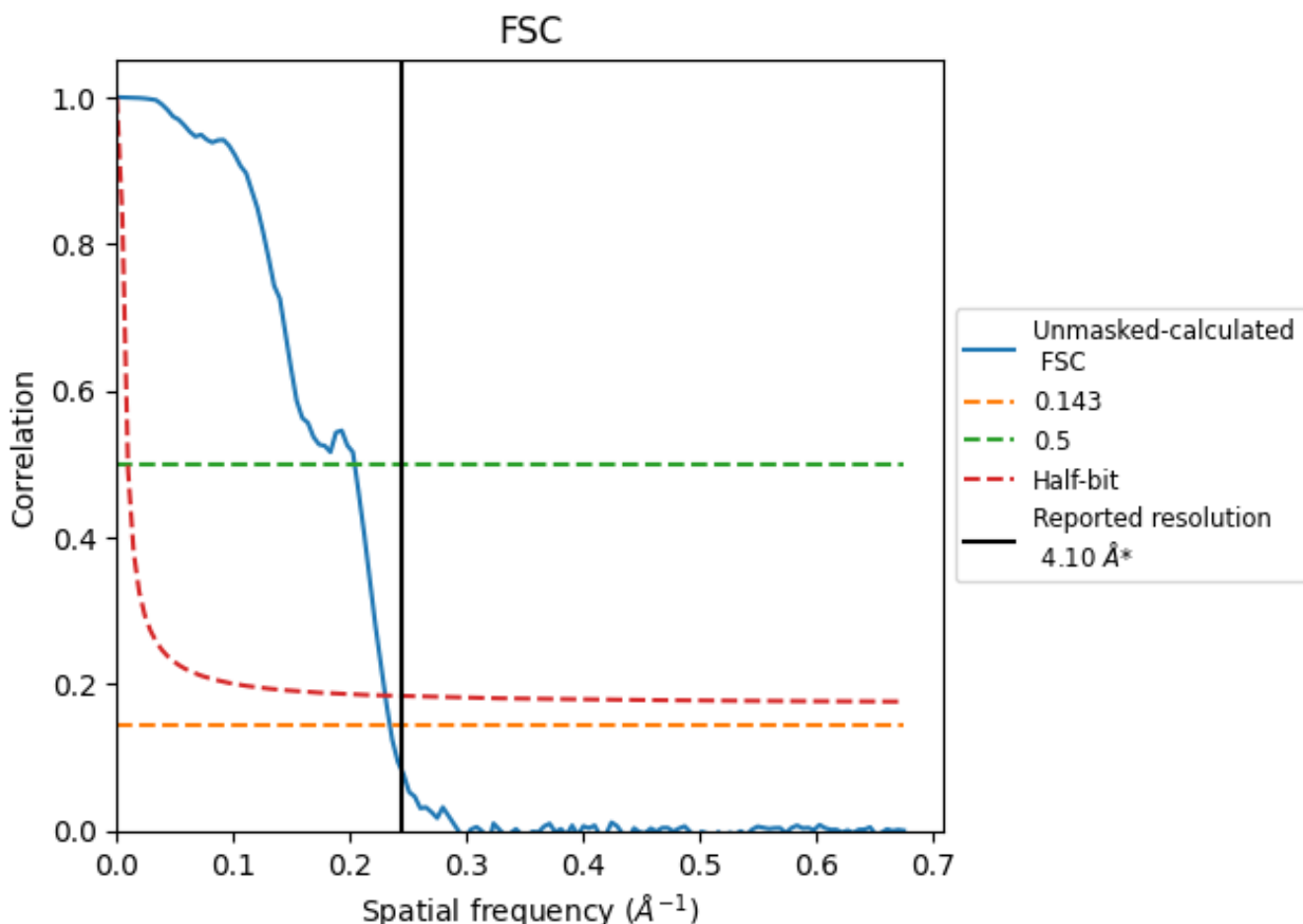
*Reported resolution corresponds to spatial frequency of 0.244 Å^{-1}

8.2   Resolution estimates (i)

Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	4.10	-	-
Author-provided FSC curve	-	-	-
Unmasked-calculated*	4.26	4.90	4.34

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

9　Map-model fit ⓘ

This section contains information regarding the fit between EMDB map EMD-16916 and PDB model 8OK2. Per-residue inclusion information can be found in section 3 on page 6.

9.1　Map-model overlay ⓘ

X　　　　　　　　　　　　　　　Y　　　　　　　　　　　　　　　Z
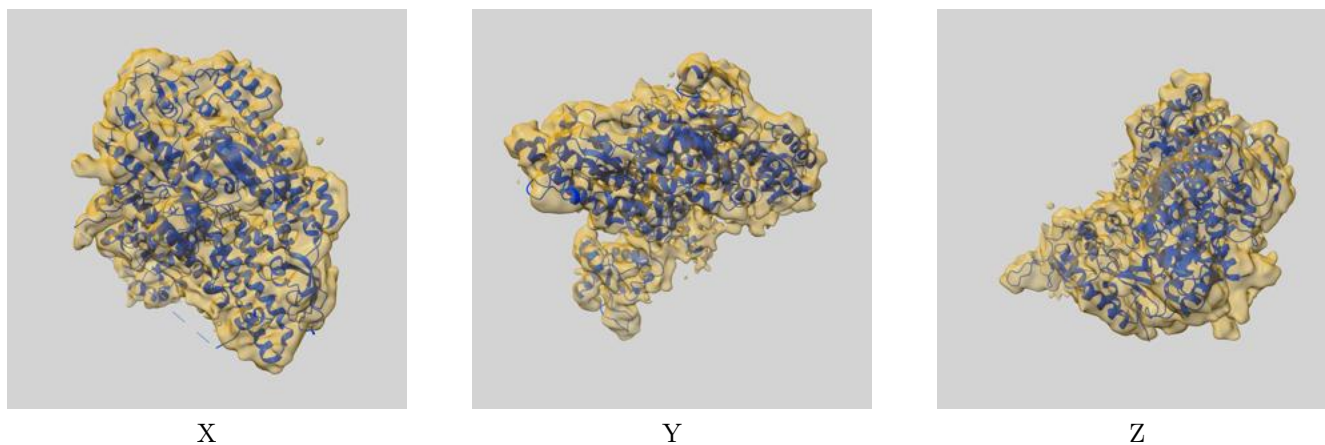
The images above show the 3D surface view of the map at the recommended contour level 0.003 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

9.2   Q-score mapped to coordinate model (i)

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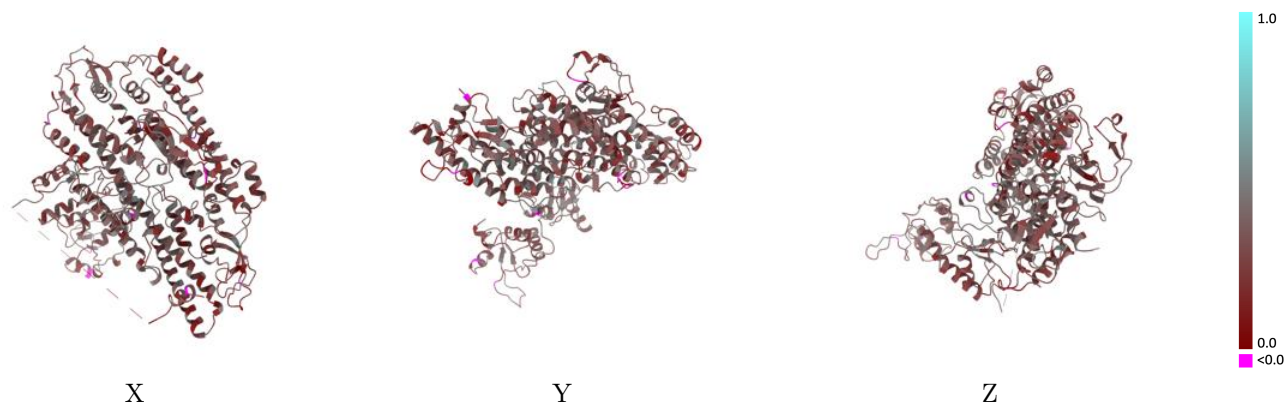
The images above show the model with each residue coloured according its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions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

X   Y   Z
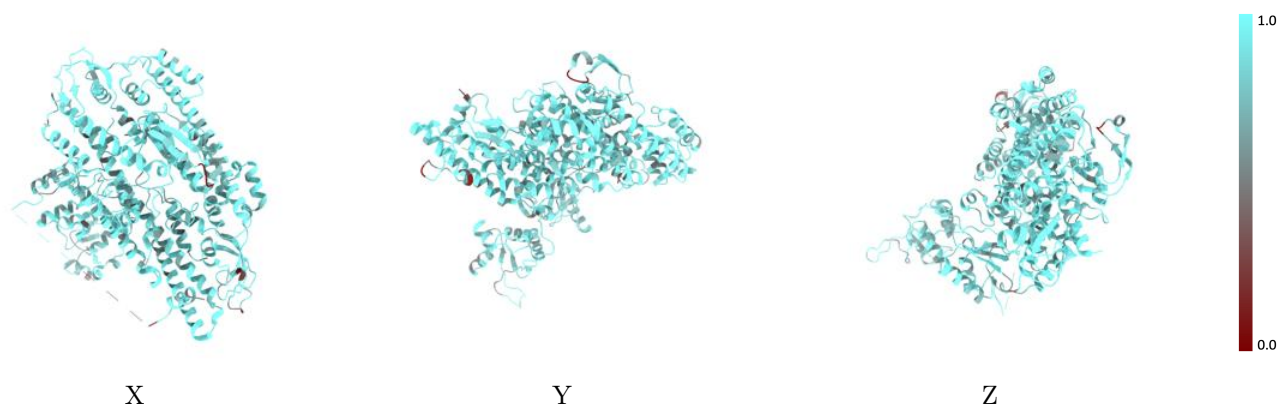
The images above show the model with each residue coloured according to its atom inclusion. This shows to what extent they are inside the map at the recommended contour level (0.003).

9.4 Atom inclusion ⓘ
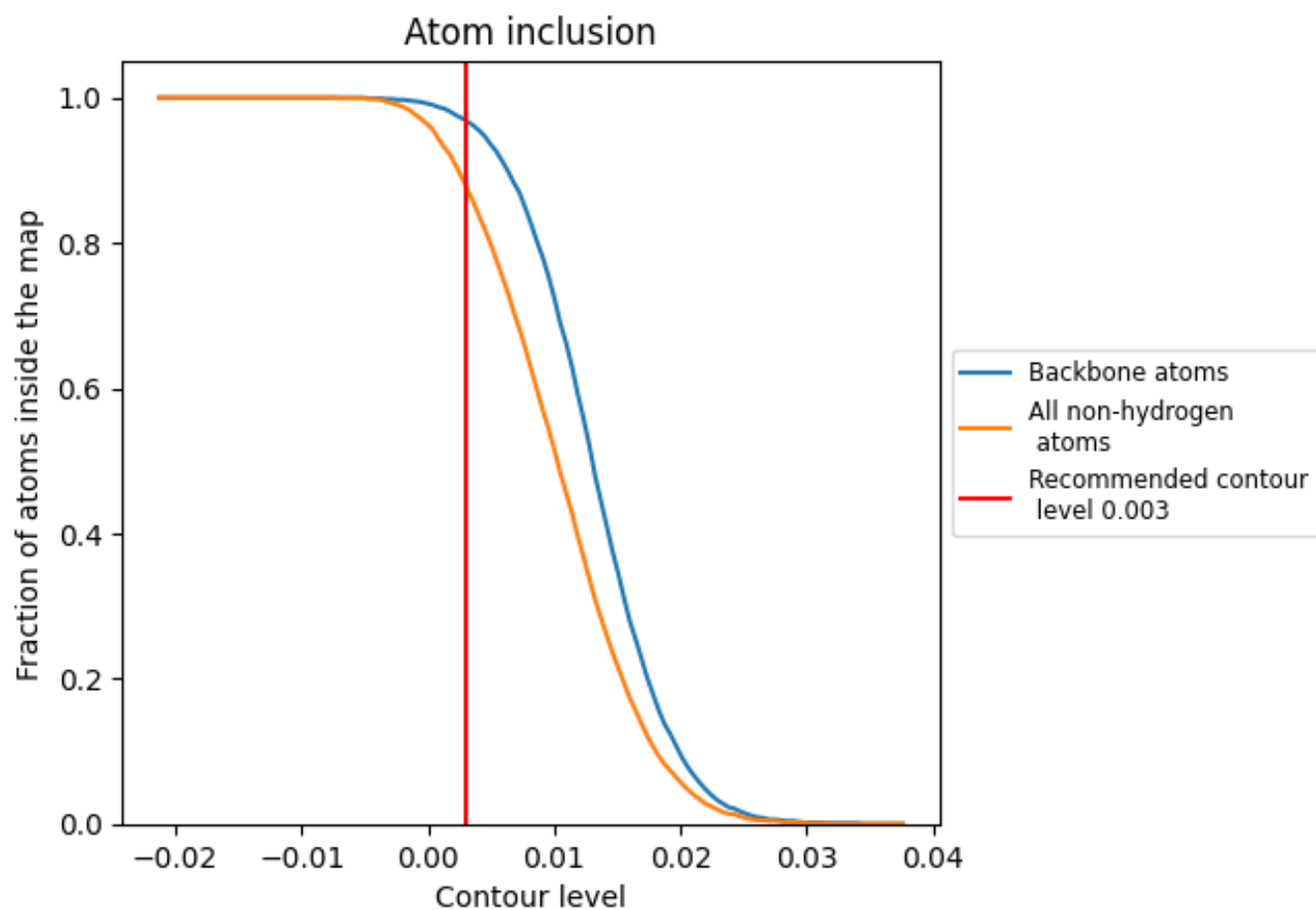
At the recommended contour level, 97% of all backbone atoms, 88% of all non-hydrogen atoms, are inside the map.

9.5   Map-model fit summary ⓘ

The table lists the average atom inclusion at the recommended contour level (0.003) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.8790	0.3130
A	0.8770	0.3330
B	0.9070	0.3260
C	0.8700	0.2950
D	0.8860	0.3250
E	0.8550	0.2890

1.0

0.0

<0.0