WORLDWIDE PDB PROTEIN DATA BANK

wwPDB EM Validation Summary Report (i)

Mar 12, 2024 – 02:12 PM JST

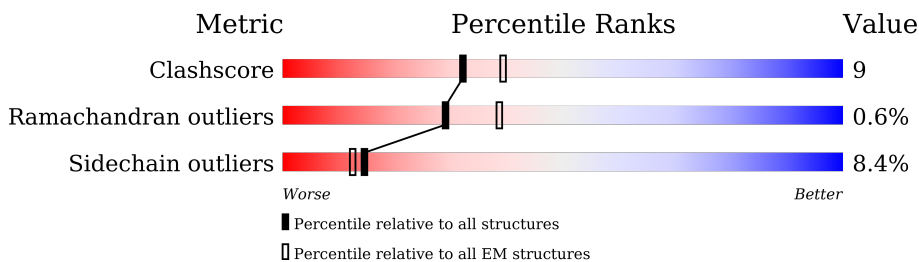|		
PDB ID	:	8JI0
EMDB ID	:	EMD-36303
Title	:	Cryo-EM structure of the TcsH-CROP in complex with TMPRSS2
Authors	:	Zhou, R.; Tao, L.; Zhan, X.
Deposited on	:	2023-05-25
Resolution	:	3.00 Å(reported)

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 3.00 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	Percentile Ranks	Value
Clashscore		9
Ramachandran outliers		0.6%
Sidechain outliers		8.4%

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	B	424	55% 22% • • 20%
2	A	786	17% • 80%

WORLDWIDE
wwPDB
PROTEIN DATA BANK

2　Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 3934 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 Transmembrane protease serine 2.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	B	341	Total 2669	C 1688	N 462	O 496	S 23	0	0

There are 38 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	69	MET	-	initiating methionine	UNP O15393
B	70	GLY	-	expression tag	UNP O15393
B	71	ILE	-	expression tag	UNP O15393
B	72	LEU	-	expression tag	UNP O15393
B	73	PRO	-	expression tag	UNP O15393
B	74	SER	-	expression tag	UNP O15393
B	75	PRO	-	expression tag	UNP O15393
B	76	GLY	-	expression tag	UNP O15393
B	77	MET	-	expression tag	UNP O15393
B	78	PRO	-	expression tag	UNP O15393
B	79	ALA	-	expression tag	UNP O15393
B	80	LEU	-	expression tag	UNP O15393
B	81	LEU	-	expression tag	UNP O15393
B	82	SER	-	expression tag	UNP O15393
B	83	LEU	-	expression tag	UNP O15393
B	84	VAL	-	expression tag	UNP O15393
B	85	SER	-	expression tag	UNP O15393
B	86	LEU	-	expression tag	UNP O15393
B	87	LEU	-	expression tag	UNP O15393
B	88	SER	-	expression tag	UNP O15393
B	89	VAL	-	expression tag	UNP O15393
B	90	LEU	-	expression tag	UNP O15393
B	91	LEU	-	expression tag	UNP O15393
B	92	MET	-	expression tag	UNP O15393
B	93	GLY	-	expression tag	UNP O15393
B	94	CYS	-	expression tag	UNP O15393
B	95	VAL	-	expression tag	UNP O15393
B	96	ALA	-	expression tag	UNP O15393

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	97	GLU	-	expression tag	UNP O15393
B	98	THR	-	expression tag	UNP O15393
B	99	GLY	-	expression tag	UNP O15393
B	100	HIS	-	expression tag	UNP O15393
B	101	HIS	-	expression tag	UNP O15393
B	102	HIS	-	expression tag	UNP O15393
B	103	HIS	-	expression tag	UNP O15393
B	104	HIS	-	expression tag	UNP O15393
B	105	HIS	-	expression tag	UNP O15393
B	255	GLN	ARG	conflict	UNP O15393

- Molecule 2 is a protein called Maltose/maltodextrin-binding periplasmic protein,Hemorrhagic toxin.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	A	157	Total 1265	C 824	N 203	O 237	S 1	0	0

There are 37 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1833	MET	-	initiating methionine	UNP P0AEX9
A	1834	ALA	-	expression tag	UNP P0AEX9
A	1835	SER	-	expression tag	UNP P0AEX9
A	1836	MET	-	expression tag	UNP P0AEX9
A	1837	THR	-	expression tag	UNP P0AEX9
A	1838	GLY	-	expression tag	UNP P0AEX9
A	1839	GLY	-	expression tag	UNP P0AEX9
A	1840	GLN	-	expression tag	UNP P0AEX9
A	1841	GLN	-	expression tag	UNP P0AEX9
A	1842	MET	-	expression tag	UNP P0AEX9
A	1843	GLY	-	expression tag	UNP P0AEX9
A	1844	ARG	-	expression tag	UNP P0AEX9
A	1845	GLY	-	expression tag	UNP P0AEX9
A	1846	SER	-	expression tag	UNP P0AEX9
A	1847	HIS	-	expression tag	UNP P0AEX9
A	1848	HIS	-	expression tag	UNP P0AEX9
A	1849	HIS	-	expression tag	UNP P0AEX9
A	1850	HIS	-	expression tag	UNP P0AEX9
A	1851	HIS	-	expression tag	UNP P0AEX9
A	1852	HIS	-	expression tag	UNP P0AEX9
A	1853	HIS	-	expression tag	UNP P0AEX9

WORLDWIDE
PDB
PROTEIN DATA BANK

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	1854	HIS	-	expression tag	UNP P0AEX9
A	1855	MET	-	expression tag	UNP P0AEX9
A	2222	GLY	-	linker	UNP P0AEX9
A	2223	SER	-	linker	UNP P0AEX9
A	2224	SER	-	linker	UNP P0AEX9
A	2225	SER	-	linker	UNP P0AEX9
A	2226	LEU	-	linker	UNP P0AEX9
A	2227	GLU	-	linker	UNP P0AEX9
A	2228	VAL	-	linker	UNP P0AEX9
A	2229	LEU	-	linker	UNP P0AEX9
A	2230	PHE	-	linker	UNP P0AEX9
A	2231	GLN	-	linker	UNP P0AEX9
A	2232	GLY	-	linker	UNP P0AEX9
A	2233	PRO	-	linker	UNP P0AEX9
A	2234	GLU	-	linker	UNP P0AEX9
A	2235	PHE	-	linker	UNP P0AEX9

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: Transmembrane protease serine 2
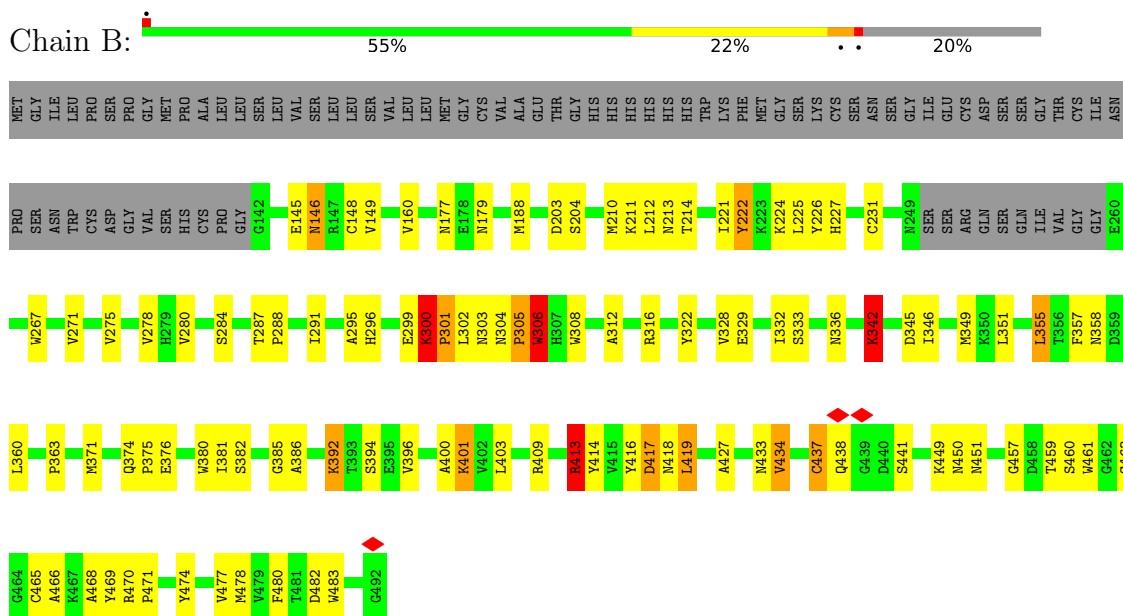
- Molecule 2: Maltose/maltodextrin-binding periplasmic protein,Hemorrhagic toxin

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

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

GLY LYS LYS TYR TYR PHE ASP ASN LYS SER LYS ALA VAL THR GLY TRP GLN THR ILE ASP GLY LYS LYS TYR TYR PHE ASN PRO ASN THR ALA ILE ALA ALA MET GLY TRP GLN ALA ILE ASP GLY LYS LYS TYR TYR PHE ASN PRO ASN THR ALA ILE ALA THR GLY TRP GLN THR

ILE ASP GLY LYS LYS TYR TYR PHE ASN ASN PRO ASN THR ALA ILE ALA THR GLY TRP GLN THR ILE ASP GLY LYS LYS TYR TYR PHE ASN PRO ASN THR ALA ILE ALA THR GLY TRP GLN THR ILE ASN SER LYS ASN PHE ASN ASP ILE MET LEU GLY VAL

PHE LYS GLY PRO ASP GLY PHE GLU TYR PHE ALA PRO ALA ASN THR HIS ASN ASN ASN GLU GLY GLN GLN SER ILE THR TYR   Q2460 S2477 K2478 A2479 V2480 R2484 H2489 R2490 F2491 P2495 I2499 G2500 T2506 L2507 N2511 L2518 V2523 G2526 G2529 F2530 E2531

S2547 Y2550 F2554 R2561 Y2562 D2565 T2581 D2600 I2603 I2608 D2609 Q2613 I2616 TYR GLY

4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	760140	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	NONE	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose $(e^-/\text{Å}^2)$	50	Depositor
Minimum defocus (nm)	1500	Depositor
Maximum defocus (nm)	2000	Depositor
Magnification	Not provided	
Image detector	GATAN K3 (6k x 4k)	Depositor
Maximum map value	7.515	Depositor
Minimum map value	-5.729	Depositor
Average map value	-0.001	Depositor
Map value standard deviation	0.100	Depositor
Recommended contour level	0.3	Depositor
Map size (Å)	278.272, 278.272, 278.272	wwPDB
Map dimensions	256, 256, 256	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.087, 1.087, 1.087	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	B	0.55	8/2740 (0.3%)	0.73	13/3724 (0.3%)
2	A	0.32	0/1307	0.51	1/1771 (0.1%)
All	All	0.49	8/4047 (0.2%)	0.67	14/5495 (0.3%)

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	B	0	2

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	304	ASN	C-O	-9.70	1.04	1.23
1	B	306	TRP	C-O	-5.90	1.12	1.23
1	B	413	ARG	C-O	-5.83	1.12	1.23
1	B	305	PRO	C-O	-5.83	1.11	1.23
1	B	303	ASN	C-O	-5.65	1.12	1.23

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	342	LYS	CB-CG-CD	-7.67	91.65	111.60
1	B	305	PRO	N-CD-CG	-7.53	91.91	103.20
1	B	306	TRP	N-CA-CB	-7.40	97.28	110.60
1	B	300	LYS	CB-CA-C	-6.99	96.41	110.40
1	B	305	PRO	CB-CA-C	-6.73	95.17	112.00

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	300	LYS	Mainchain
1	B	306	TRP	Mainchain

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	B	2669	0	2561	60	0
2	A	1265	0	1159	13	0
All	All	3934	0	3720	72	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 9.

The worst 5 of 72 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:B:300:LYS:HB3	1:B:301:PRO:CD	1.94	0.98
1:B:300:LYS:O	1:B:300:LYS:NZ	2.03	0.90
1:B:300:LYS:HB3	1:B:301:PRO:HD3	1.54	0.90
1:B:419:LEU:HD23	1:B:461:TRP:CH2	2.06	0.90
1:B:371:MET:O	1:B:449:LYS:NZ	2.07	0.88

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	B	337/424 (80%)	300 (89%)	34 (10%)	3 (1%)	17	55
2	A	155/786 (20%)	145 (94%)	10 (6%)	0	100	100
All	All	492/1210 (41%)	445 (90%)	44 (9%)	3 (1%)	29	64

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	300	LYS
1	B	145	GLU
1	B	305	PRO

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	B	290/360 (81%)	259 (89%)	31 (11%)	6	26
2	A	128/637 (20%)	124 (97%)	4 (3%)	40	75
All	All	418/997 (42%)	383 (92%)	35 (8%)	14	38

5 of 35 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	460	SER
1	B	470	ARG
2	A	2565	ASP
1	B	333	SER
1	B	332	ILE

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

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 ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

6　Map visualisation (i)

This section contains visualisations of the EMDB entry EMD-36303. 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
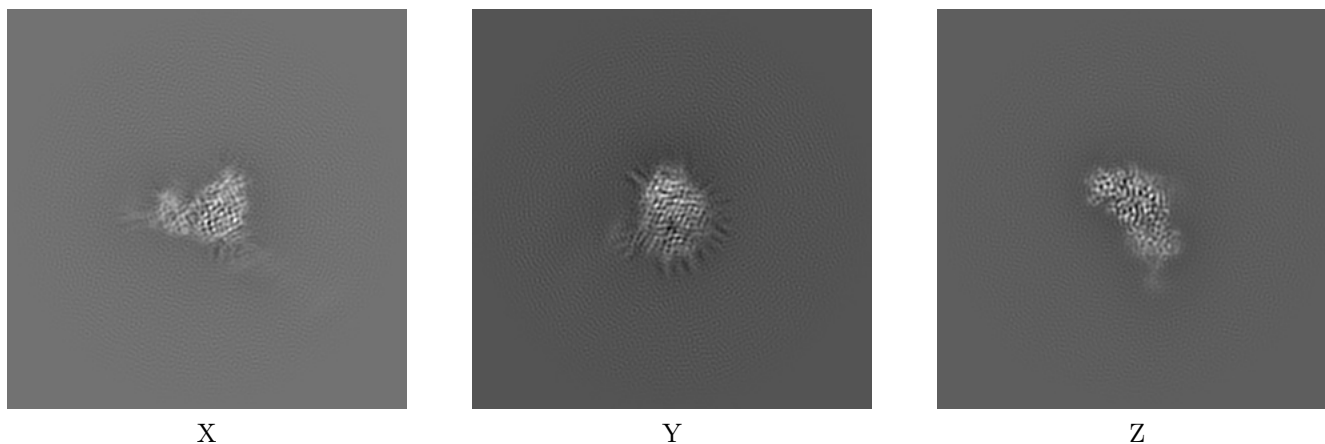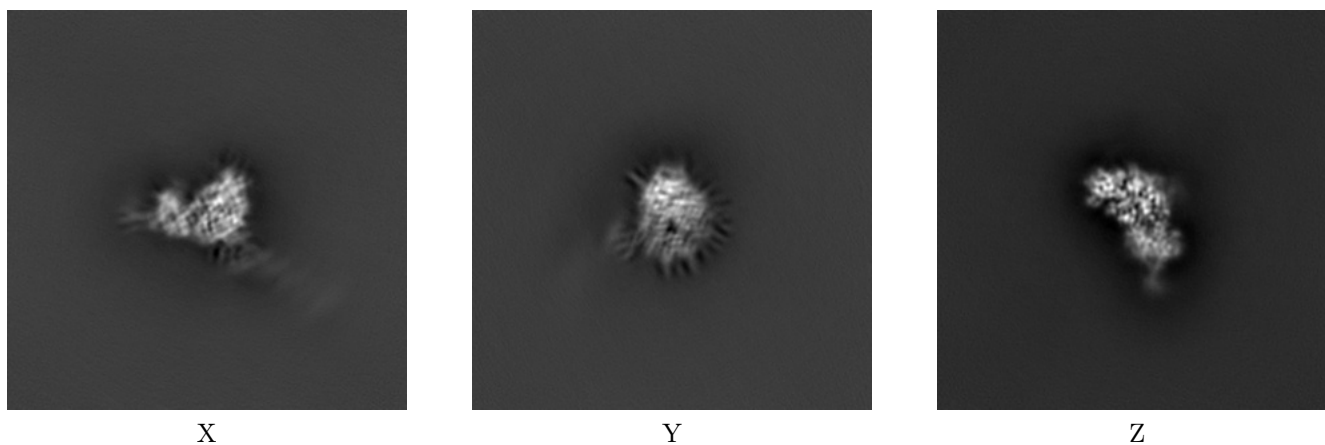
The images above show the map projected in three orthogonal directions.

6.2 Central slices (i)

6.2.1 Primary map
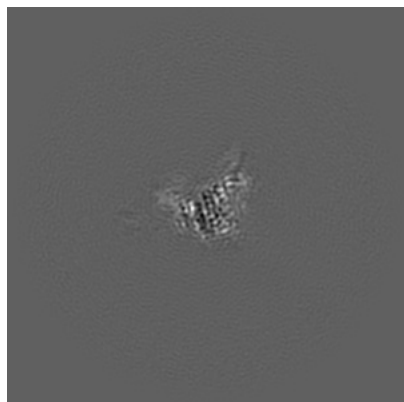
X Index: 128
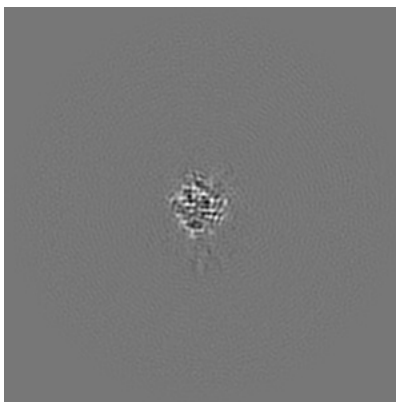
Y Index: 128
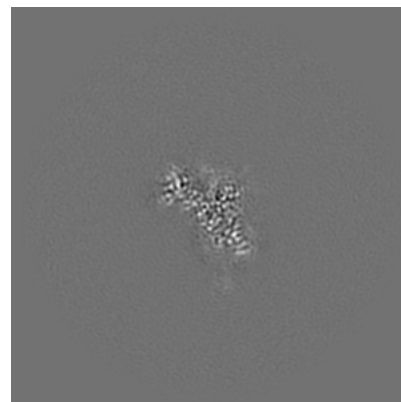
Z Index: 128

6.2.2 Raw map
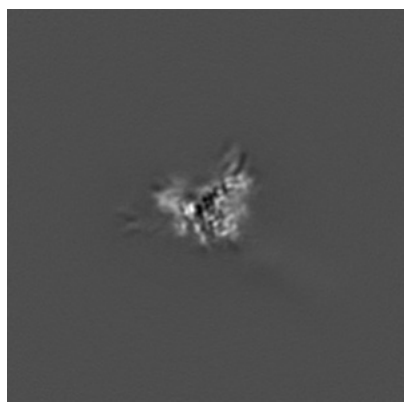
X Index: 128
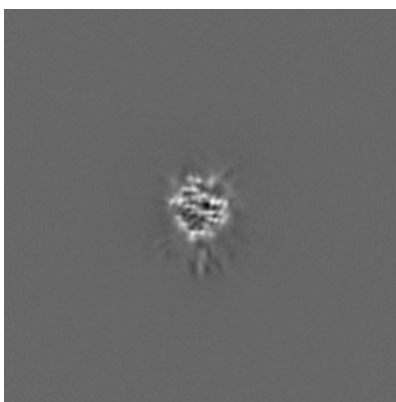
Y Index: 128
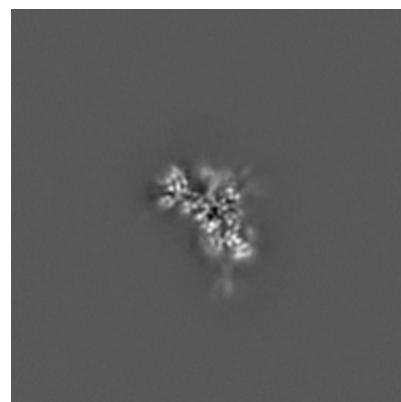
Z Index: 128

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

6.3 Largest variance slices ⓘ

6.3.1 Primary map
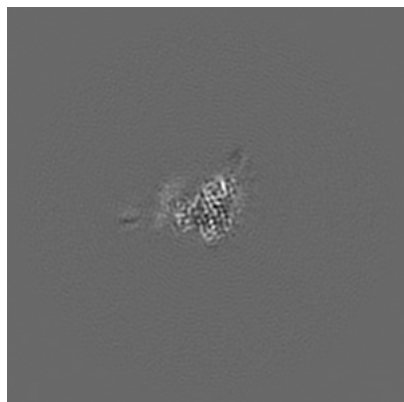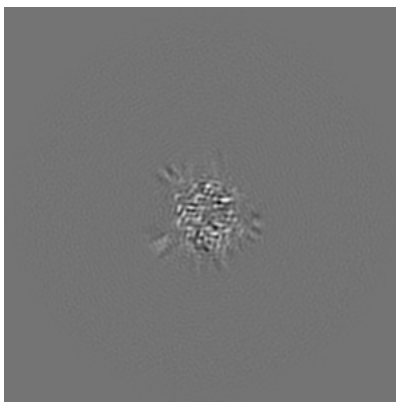
| X Index: 132 | Y Index: 138 | Z Index: 124 |
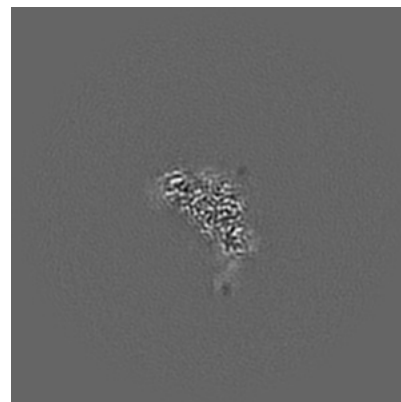
6.3.2 Raw map
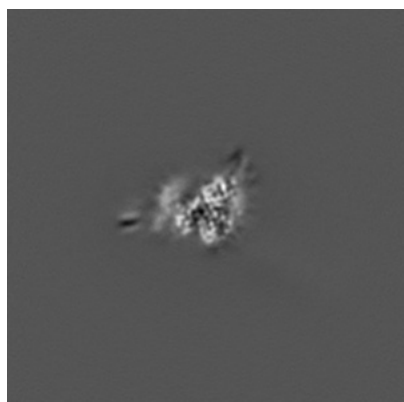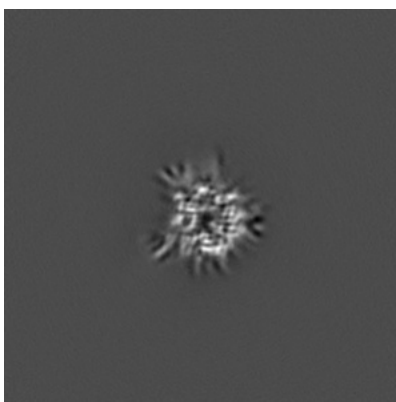
| X Index: 132 | Y Index: 140 | Z Index: 129 |
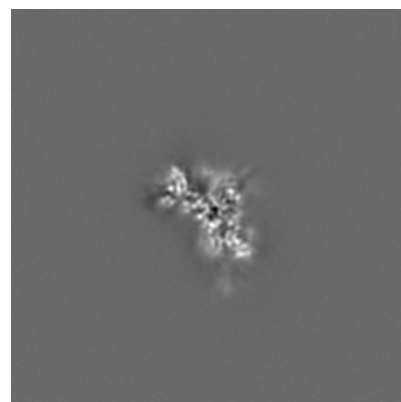
The images above show the largest variance slices of the map in three orthogonal directions.

6.4 Orthogonal standard-deviation projections (False-color) ⓘ

6.4.1 Primary map

X Y Z

6.4.2 Raw map

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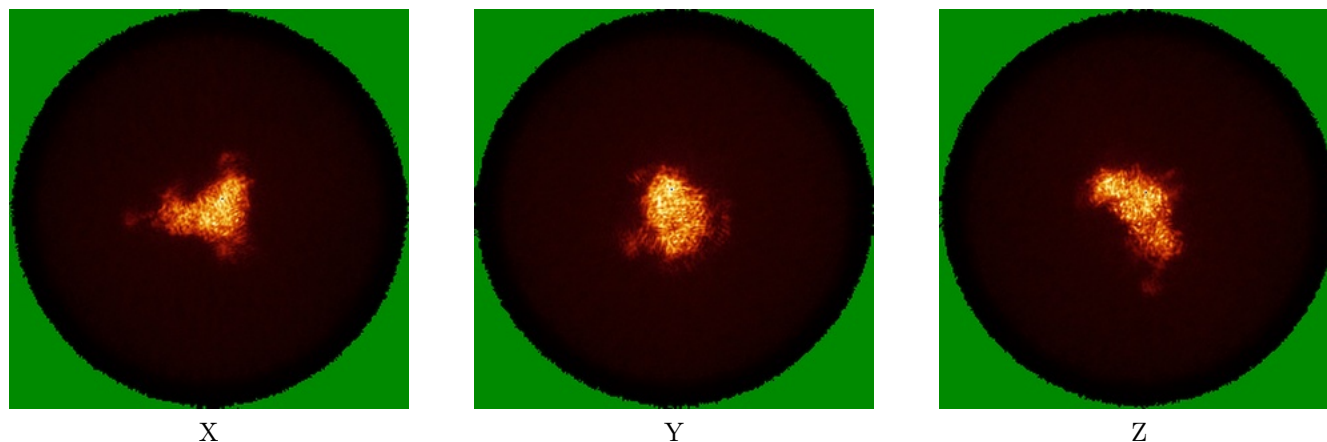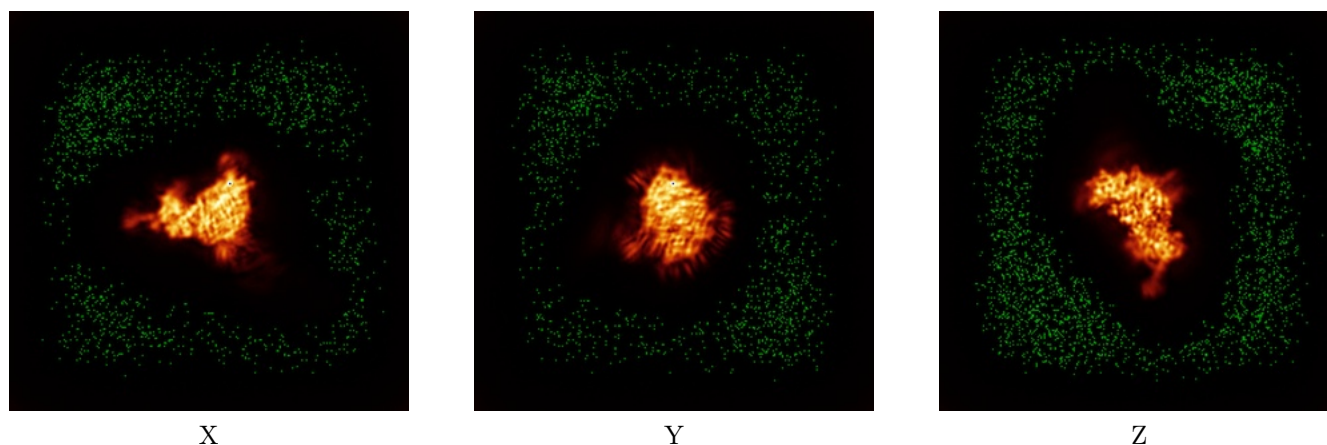
The images above show the map standard deviation projections with false color in three orthogonal directions. Minimum values are shown in green, max in blue, and dark to light orange shades represent small to large values respectively.

6.5 Orthogonal surface views (i)

6.5.1 Primary map

X   Y   Z
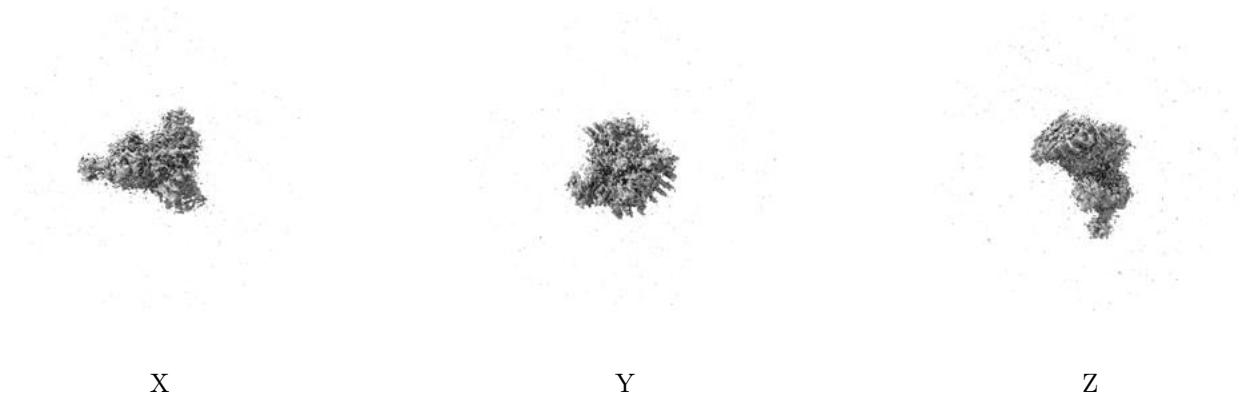
The images above show the 3D surface view of the map at the recommended contour level 0.3. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

6.5.2 Raw map

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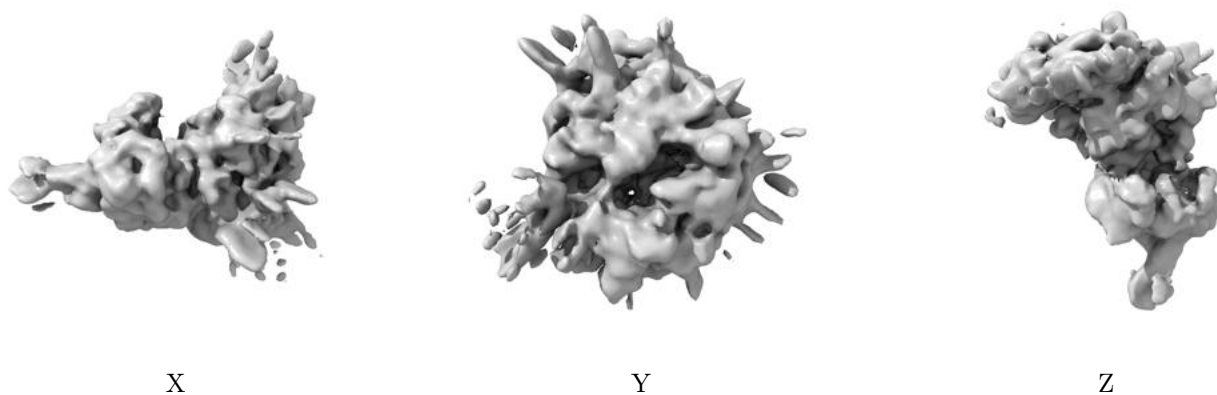
These images show the 3D surface of the raw map. The raw map's contour level was selected so that its surface encloses the same volume as the primary map does at its recommended contour level.

6.6 Mask visualisation (i)

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

7　Map analysis ⓘ

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

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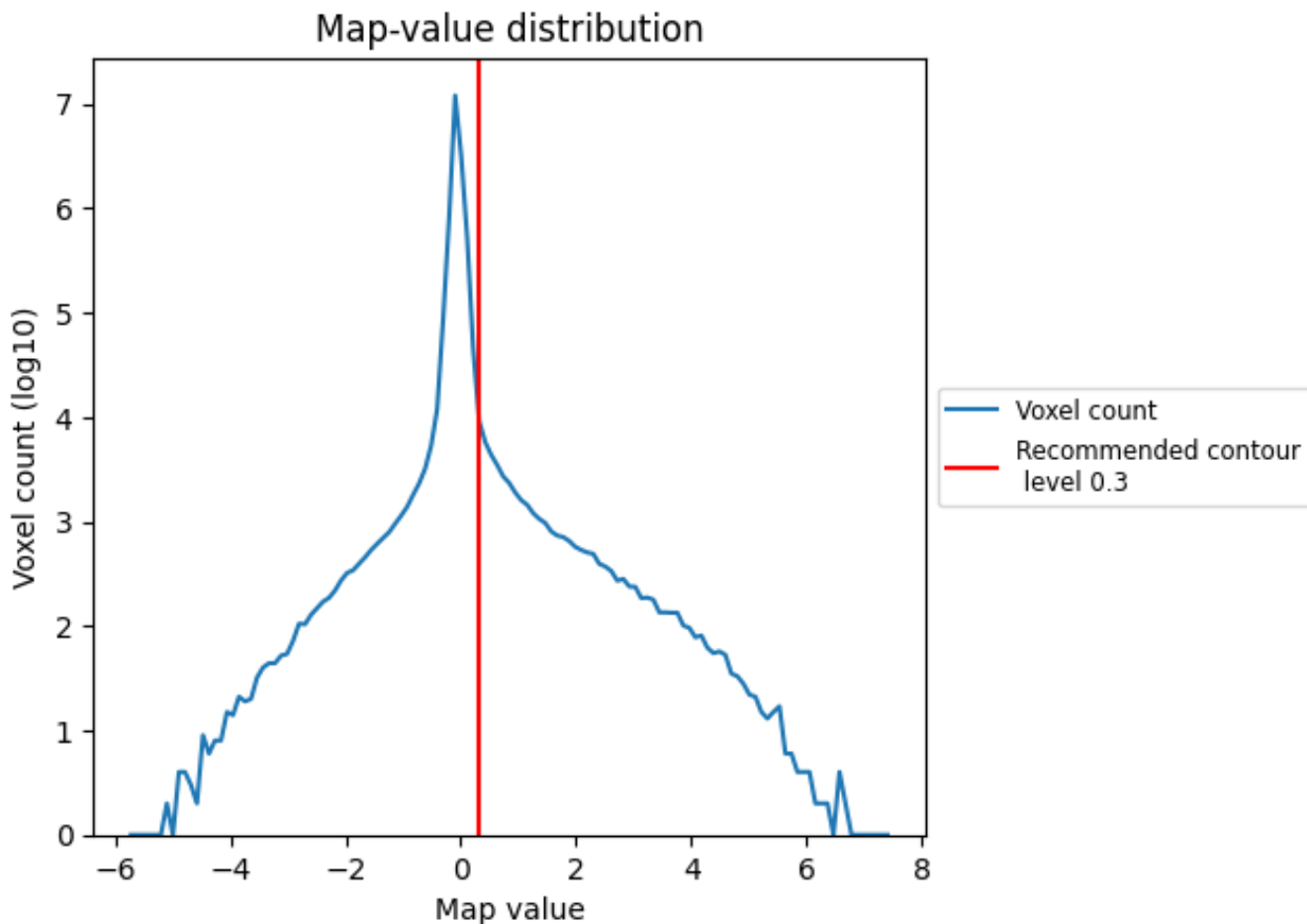
The map-value distribution is plotted in 128 intervals along the x-axis. The y-axis is logarithmic. A spike in this graph at zero usually indicates that the volume has been masked.

7.2   Volume estimate ⓘ
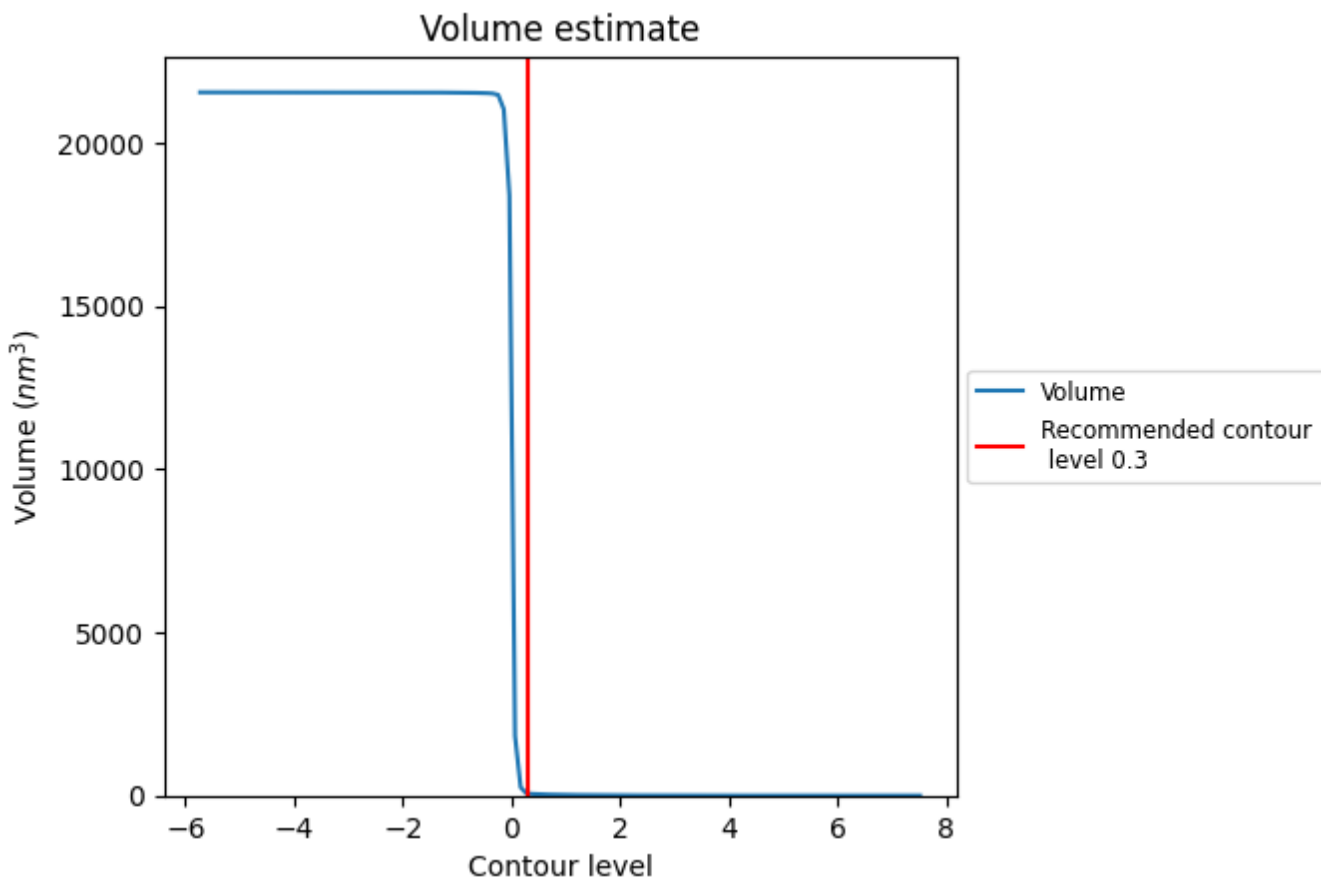
The volume at the recommended contour level is 64 nm^3; this corresponds to an approximate mass of 58 kDa.

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)
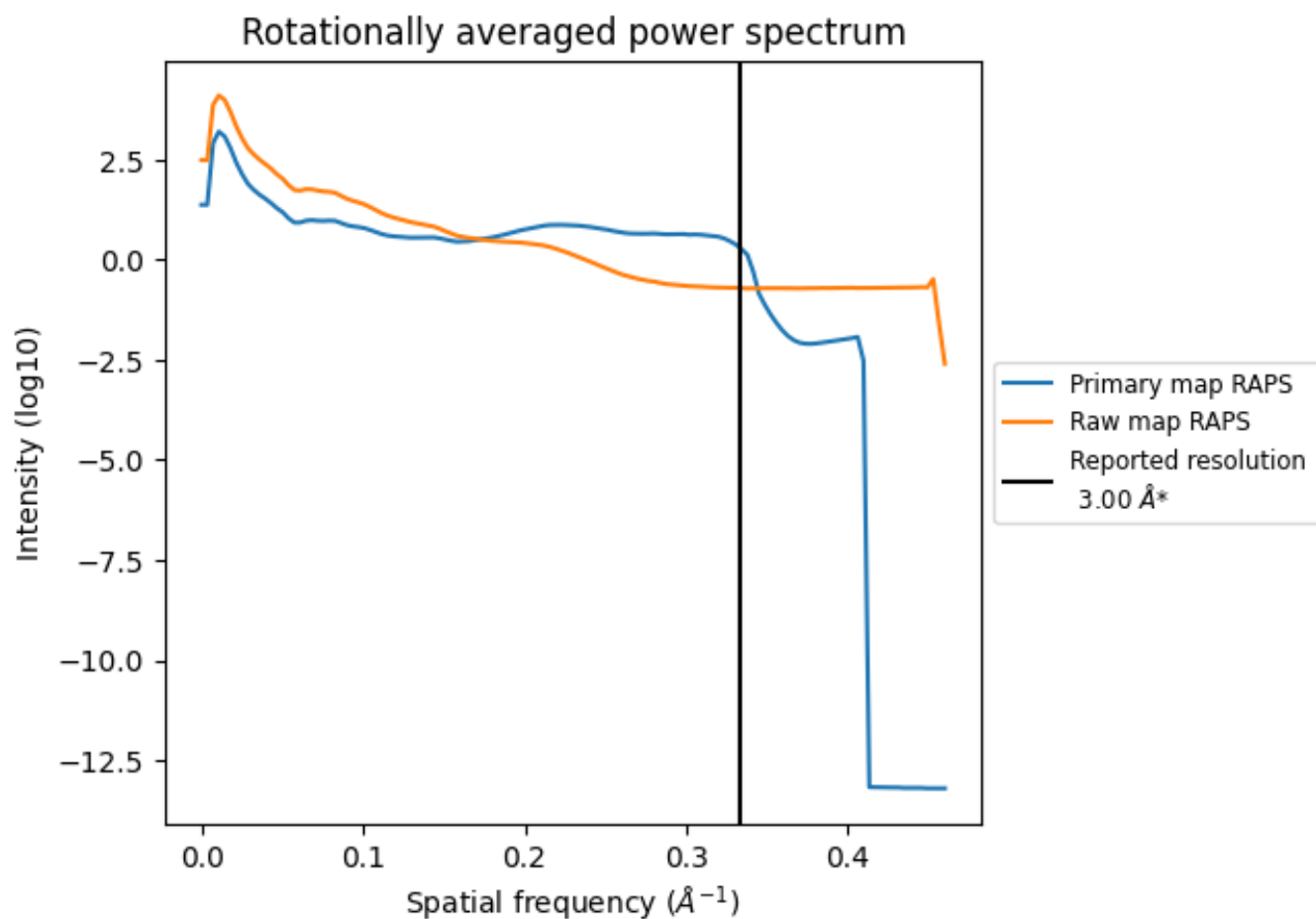
*Reported resolution corresponds to spatial frequency of 0.333 Å^{-1}

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)
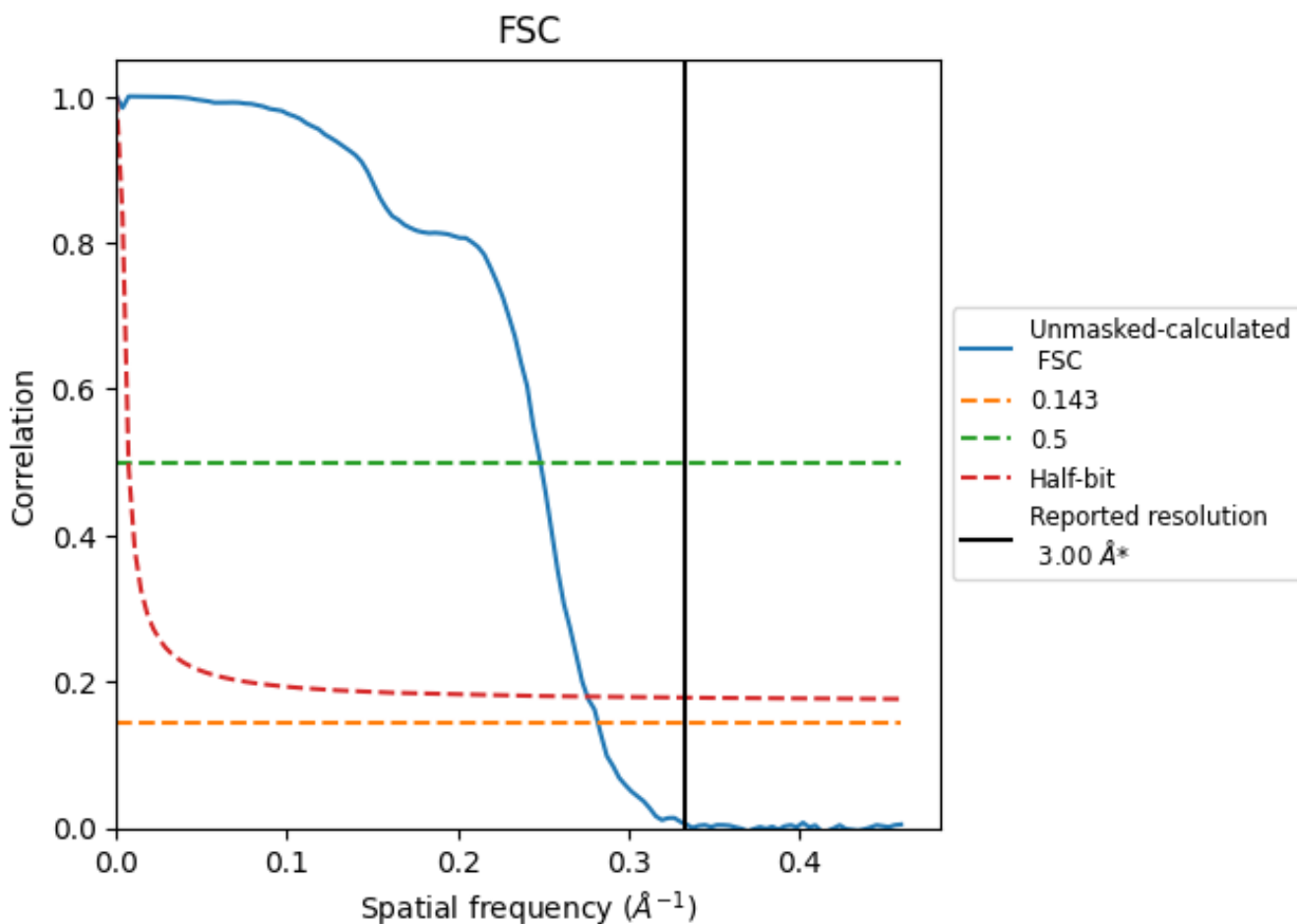
*Reported resolution corresponds to spatial frequency of 0.333 Å^{-1}

8.2   Resolution estimates ⓘ

Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	3.00	-	-
Author-provided FSC curve	-	-	-
Unmasked-calculated*	3.54	4.03	3.62

*Resolution estimate based on FSC curve calculated by comparison of deposited half-maps. The value from deposited half-maps intersecting FSC 0.143 CUT-OFF 3.54 differs from the reported value 3.0 by more than 10 %

9   Map-model fit ⓘ

This section contains information regarding the fit between EMDB map EMD-36303 and PDB model 8JI0. Per-residue inclusion information can be found in section 3 on page 6.

9.1   Map-model overlay ⓘ


<div align="center">X   Y   Z</div>
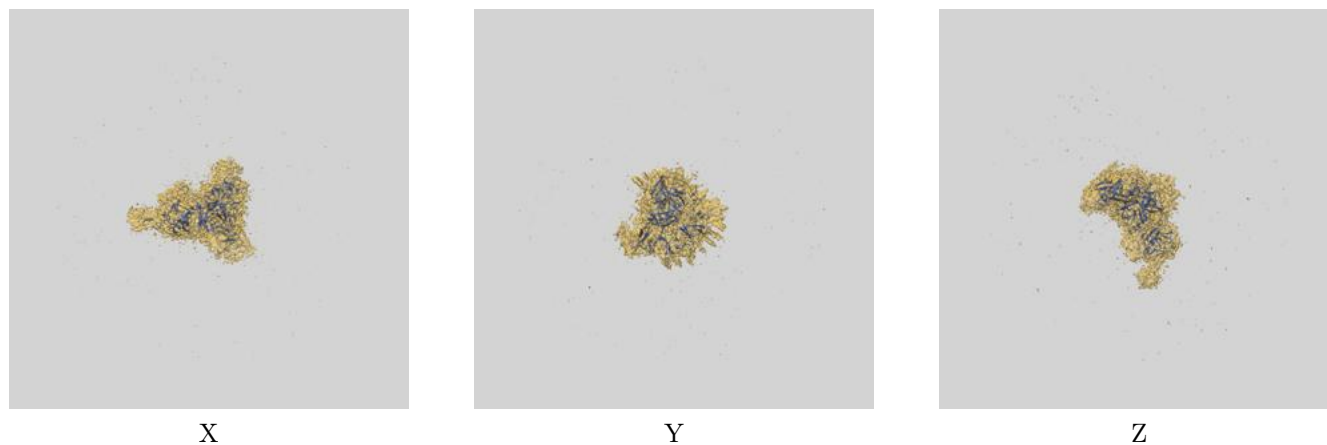
The images above show the 3D surface view of the map at the recommended contour level 0.3 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

9.2   Q-score mapped to coordinate model ⓘ

X   Y   Z
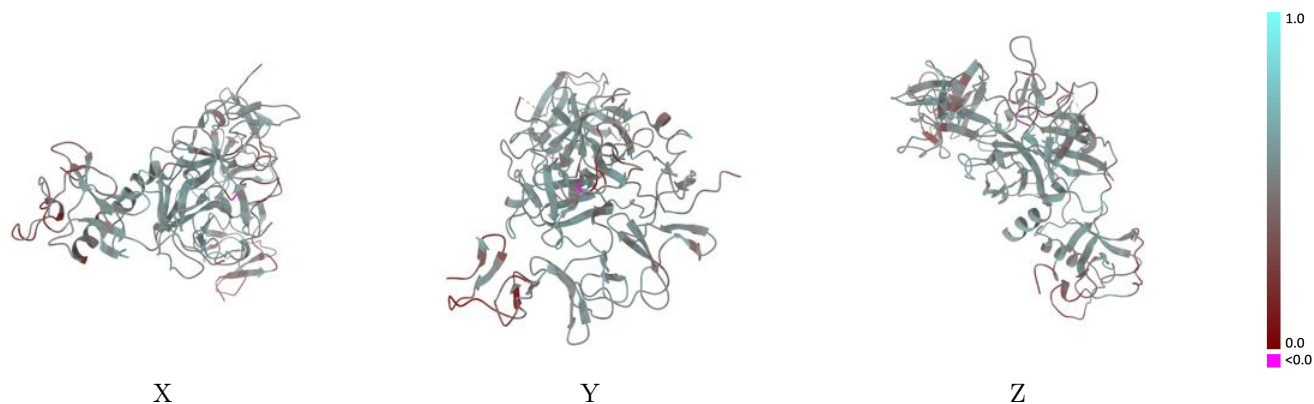
The images above show the model with each residue coloured according its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions at which atoms can be resolved. Low Q-score values may therefore be expected for many entries.
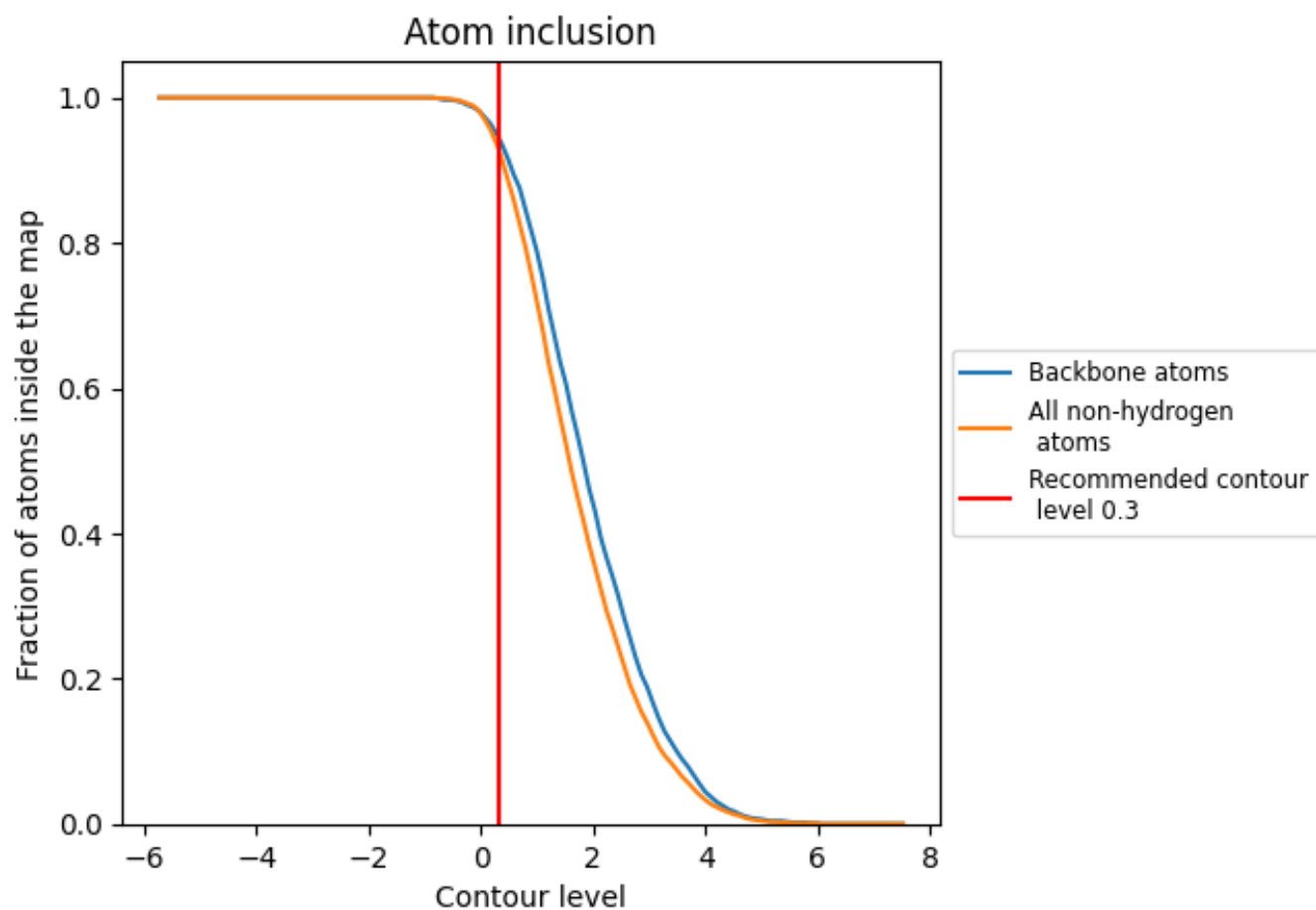
9.3   Atom inclusion mapped to coordinate model ⓘ

X   Y   Z
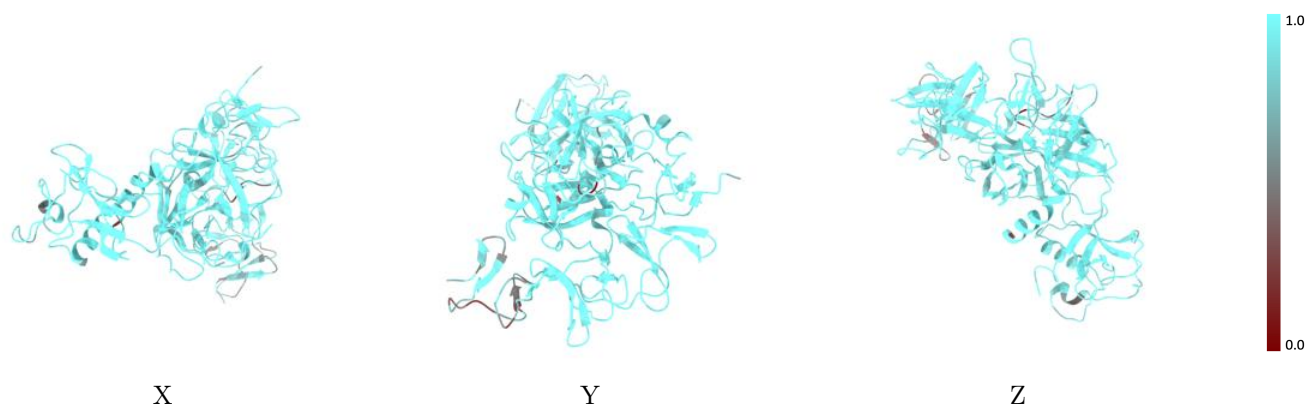
The images above show the model with each residue coloured according to its atom inclusion. This shows to what extent they are inside the map at the recommended contour level (0.3).

9.4 Atom inclusion ⓘ

At the recommended contour level, 95% of all backbone atoms, 93% 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 (0.3) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.9320	0.4900
A	0.8940	0.4700
B	0.9500	0.5000

1.0

0.0

<0.0