W O R L D W I D E

wwPDB

PROTEIN DATA BANK

wwPDB EM Validation Summary Report (i)

Oct 28, 2024 – 07:43 AM EDT

PDB ID	:	7R87
EMDB ID	:	EMD-24310
Title	:	The structure of human ABCG5-WT/ABCG8-I419E
Authors	:	Sun, Y.; Li, X.; Long, T.
Deposited on	:	2021-06-26
Resolution	:	3.40 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 3.40 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.
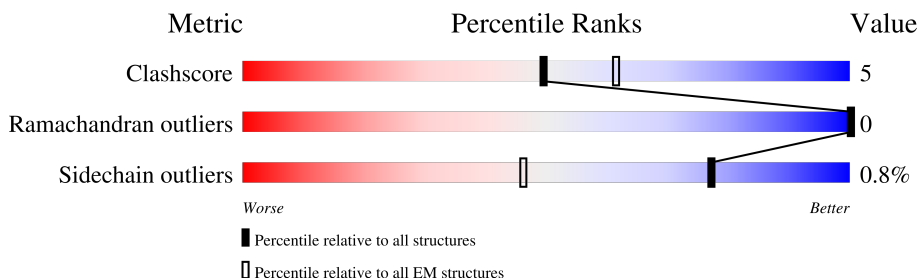
Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	210492	15764
Ramachandran outliers	207382	16835
Sidechain outliers	206894	16415

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 <=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	666	
2	B	715	
3	C	245	
4	D	234	

2 Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 10793 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 ATP-binding cassette sub-family G member 5.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	586	Total 4608	C 2984	N 781	O 814	S 29	0	0

There are 15 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	652	GLY	-	expression tag	UNP Q9H222
A	653	SER	-	expression tag	UNP Q9H222
A	654	HIS	-	expression tag	UNP Q9H222
A	655	HIS	-	expression tag	UNP Q9H222
A	656	HIS	-	expression tag	UNP Q9H222
A	657	HIS	-	expression tag	UNP Q9H222
A	658	HIS	-	expression tag	UNP Q9H222
A	659	HIS	-	expression tag	UNP Q9H222
A	660	GLY	-	expression tag	UNP Q9H222
A	661	HIS	-	expression tag	UNP Q9H222
A	662	HIS	-	expression tag	UNP Q9H222
A	663	HIS	-	expression tag	UNP Q9H222
A	664	HIS	-	expression tag	UNP Q9H222
A	665	HIS	-	expression tag	UNP Q9H222
A	666	HIS	-	expression tag	UNP Q9H222

- Molecule 2 is a protein called ATP-binding cassette sub-family G member 8.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	564	Total 4481	C 2908	N 759	O 785	S 29	0	0

There are 43 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	419	GLU	ILE	engineered mutation	UNP Q9H221
B	674	ALA	-	expression tag	UNP Q9H221

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	675	SER	-	expression tag	UNP Q9H221
B	676	ASN	-	expression tag	UNP Q9H221
B	677	SER	-	expression tag	UNP Q9H221
B	678	LEU	-	expression tag	UNP Q9H221
B	679	GLU	-	expression tag	UNP Q9H221
B	680	VAL	-	expression tag	UNP Q9H221
B	681	LEU	-	expression tag	UNP Q9H221
B	682	PHE	-	expression tag	UNP Q9H221
B	683	GLN	-	expression tag	UNP Q9H221
B	684	GLY	-	expression tag	UNP Q9H221
B	685	PRO	-	expression tag	UNP Q9H221
B	686	ASN	-	expression tag	UNP Q9H221
B	687	VAL	-	expression tag	UNP Q9H221
B	688	ASP	-	expression tag	UNP Q9H221
B	689	SER	-	expression tag	UNP Q9H221
B	690	LYS	-	expression tag	UNP Q9H221
B	691	ARG	-	expression tag	UNP Q9H221
B	692	ARG	-	expression tag	UNP Q9H221
B	693	TRP	-	expression tag	UNP Q9H221
B	694	LYS	-	expression tag	UNP Q9H221
B	695	LYS	-	expression tag	UNP Q9H221
B	696	ASN	-	expression tag	UNP Q9H221
B	697	PHE	-	expression tag	UNP Q9H221
B	698	ILE	-	expression tag	UNP Q9H221
B	699	ALA	-	expression tag	UNP Q9H221
B	700	VAL	-	expression tag	UNP Q9H221
B	701	SER	-	expression tag	UNP Q9H221
B	702	ALA	-	expression tag	UNP Q9H221
B	703	ALA	-	expression tag	UNP Q9H221
B	704	ASN	-	expression tag	UNP Q9H221
B	705	ARG	-	expression tag	UNP Q9H221
B	706	PHE	-	expression tag	UNP Q9H221
B	707	LYS	-	expression tag	UNP Q9H221
B	708	LYS	-	expression tag	UNP Q9H221
B	709	ILE	-	expression tag	UNP Q9H221
B	710	SER	-	expression tag	UNP Q9H221
B	711	SER	-	expression tag	UNP Q9H221
B	712	SER	-	expression tag	UNP Q9H221
B	713	GLY	-	expression tag	UNP Q9H221
B	714	ALA	-	expression tag	UNP Q9H221
B	715	LEU	-	expression tag	UNP Q9H221

- Molecule 3 is a protein called 2C7 Fab heavy chain.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	C	114	Total	C	N	O	S	0	0
			901	571	154	171	5		

- Molecule 4 is a protein called 2C7 Fab light chain.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	D	105	Total	C	N	O	S	0	0
			803	503	133	163	4		

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: ATP-binding cassette sub-family G member 5

• Molecule 2: ATP-binding cassette sub-family G member 8

LEU TRP LYS ALA GLU THR LYS ASP LEU ASP GLU ASP THR CYS VAL GLU SER SER VAL THR PRO LEU ASP THR ASN CYS LEU PRO SER PRO THR LYS M392 P393 F399 L402 N409 F445 L452 F461 I464 I468 R475 A476 M477 E483 T488 P491

Y492 C505 M513 P514 R522 W536 L537 V538 V539 F540 I544 M545 A546 A550 L553 M558 A559 G575 I578 N579 L580 L583 W584 P587 R597 W598 C599 L603 Q607 Y613 K614 M615 PRO LEU GLY ASN LEU THR ILE ALA VAL SER G626

I629 V632 M633 E634 I643 K668 P669 SER GLN ASP TRP ALA SER ASN SER LEU GLU VAL LEU PHE GLN GLY PRO ASN VAL ASP SER LYS ARG ARG TRP LYS LYS ASN PHE ILE ALA VAL SER ALA ALA ASN ARG LYS LYS ILE SER SER SER GLY ALA LEU

• Molecule 3: 2C7 Fab heavy chain

Chain C:

40% 7% 53%

MET GLY TRP SER CYS ILE ILE LEU PHE LEU VAL ALA THR ALA THR GLY VAL HIS SER GLU V21 K22 L30 G34 C41 A42 T43 F48 M53 E54 R57 K62 S70 R71 N75 D76 G87 R93 S96 Q97 N98 I99 L100 L107 R108 A109 E110

D111 T112 A113 I114 Y115 A118 V134 THR VAL SER SER ALA SER THR LYS GLY PRO SER VAL PHE PRO LEU ALA PRO SER SER LYS SER THR SER GLY GLY THR ALA ALA LEU GLY CYS LEU VAL LYS ASP TYR PHE PRO GLU PRO VAL THR VAL SER TRP ASN SER GLY ALA LEU THR

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

• Molecule 4: 2C7 Fab light chain

Chain D:

40% 5% 55%

MET GLY TRP SER CYS ILE ILE LEU PHE LEU VAL ALA THR ALA ARG THR GLY VAL HIS SER ASP I22 S34 V39 S40 L41 E48 K80 S83 R86 T94 I95 L103 T122 K123 L124 E125 I126 LYS ARG THR VAL ALA ALA PRO SER VAL PHE ILE PHE PRO

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

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

4　Experimental information　ⓘ

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	500000	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 $(e^-/Å^2)$	60	Depositor
Minimum defocus (nm)	Not provided	
Maximum defocus (nm)	Not provided	
Magnification	Not provided	
Image detector	GATAN K3 (6k x 4k)	Depositor
Maximum map value	0.082	Depositor
Minimum map value	-0.044	Depositor
Average map value	0.000	Depositor
Map value standard deviation	0.002	Depositor
Recommended contour level	0.015	Depositor
Map size (Å)	246.56801, 246.56801, 246.56801	wwPDB
Map dimensions	296, 296, 296	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	0.833, 0.833, 0.833	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	0.24	0/4699	0.38	0/6357
2	B	0.24	0/4584	0.38	0/6205
3	C	0.24	0/923	0.44	0/1251
4	D	0.24	0/819	0.42	0/1107
All	All	0.24	0/11025	0.39	0/14920

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	4608	0	4752	58	0
2	B	4481	0	4543	50	0
3	C	901	0	856	12	0
4	D	803	0	777	6	0
All	All	10793	0	10928	119	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 119 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:C:41:CYS:HB3	3:C:100:LEU:HB3	1.79	0.64
3:C:54:GLU:HB2	3:C:118:ALA:HB3	1.80	0.62
2:B:464:ILE:HD13	2:B:546:ALA:HB2	1.82	0.61
1:A:498:ARG:HA	1:A:578:ASN:HD21	1.66	0.61
1:A:467:TYR:HA	1:A:470:HIS:HD2	1.64	0.61

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	580/666 (87%)	563 (97%)	17 (3%)	0	100	100
2	B	554/715 (78%)	537 (97%)	17 (3%)	0	100	100
3	C	112/245 (46%)	106 (95%)	6 (5%)	0	100	100
4	D	103/234 (44%)	100 (97%)	3 (3%)	0	100	100
All	All	1349/1860 (72%)	1306 (97%)	43 (3%)	0	100	100

There are no Ramachandran outliers to report.

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	512/581 (88%)	507 (99%)	5 (1%)	73	83
2	B	485/616 (79%)	481 (99%)	4 (1%)	79	87

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
3	C	94/207 (45%)	94 (100%)	0	100	100
4	D	90/204 (44%)	90 (100%)	0	100	100
All	All	1181/1608 (73%)	1172 (99%)	9 (1%)	77	87

5 of 9 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	584	TRP
2	B	597	ARG
1	A	624	PHE
1	A	628	TYR
2	B	213	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 33 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	607	GLN
3	C	58	GLN
4	D	65	GLN
1	A	462	GLN
1	A	440	ASN

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 oligosaccharides 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 ⓘ

This section contains visualisations of the EMDB entry EMD-24310. 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 ⓘ
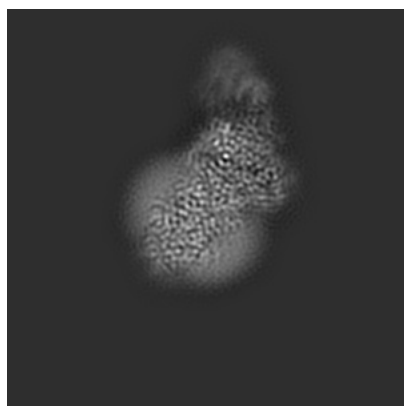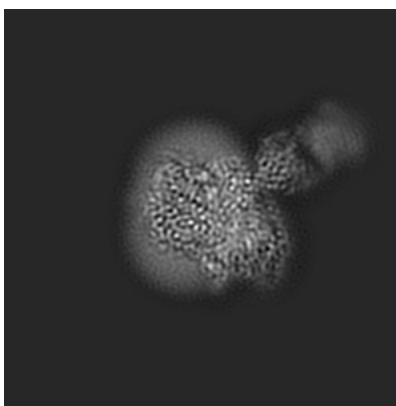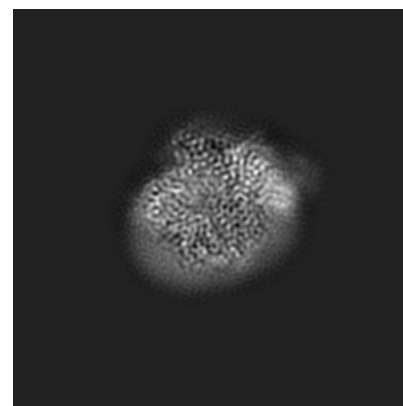
6.1.1 Primary map

| X | Y | Z |

6.1.2 Raw map

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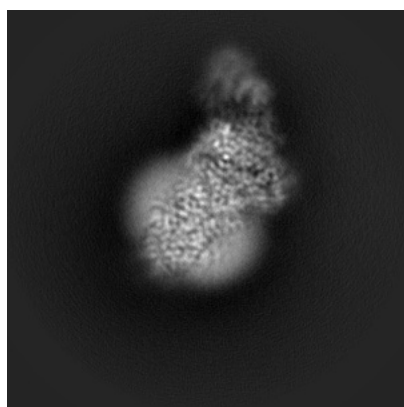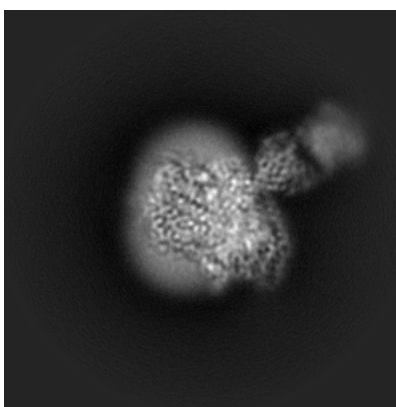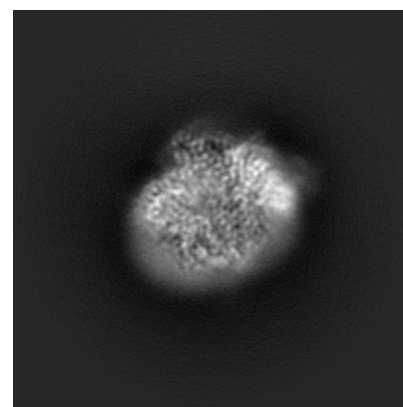
The images above show the map projected in three orthogonal directions.

6.2 Central slices ⓘ
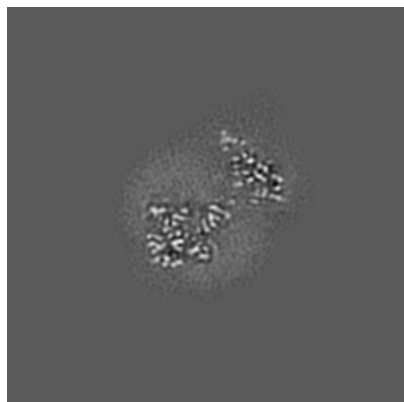
6.2.1 Primary map
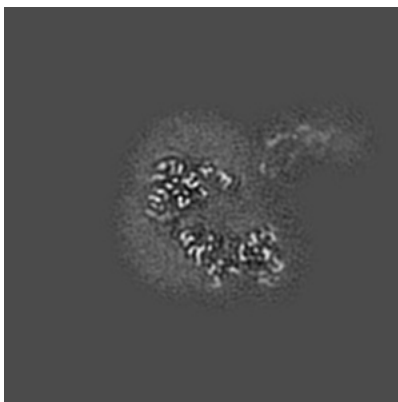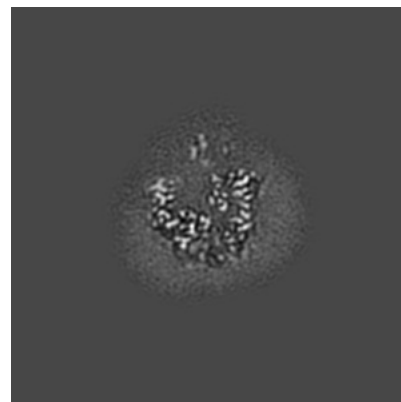
X Index: 148   Y Index: 148   Z Index: 148

6.2.2 Raw map
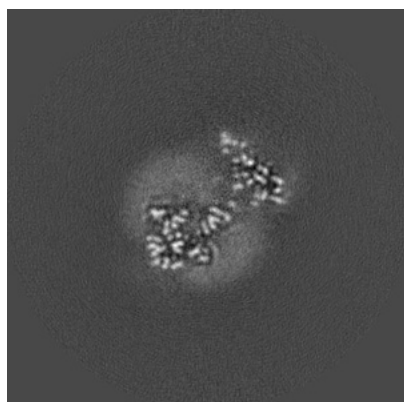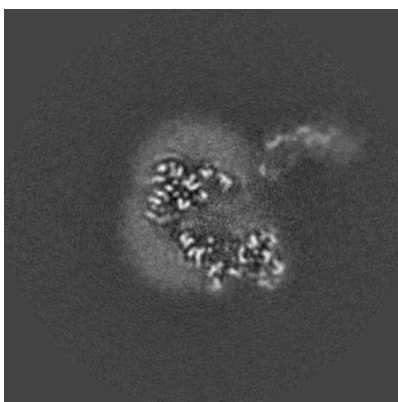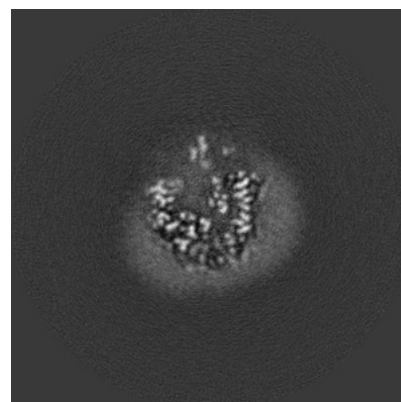
X Index: 148   Y Index: 148   Z Index: 148

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

6.3 Largest variance slices (i)

6.3.1 Primary map
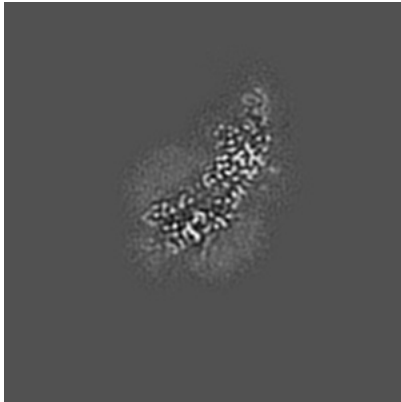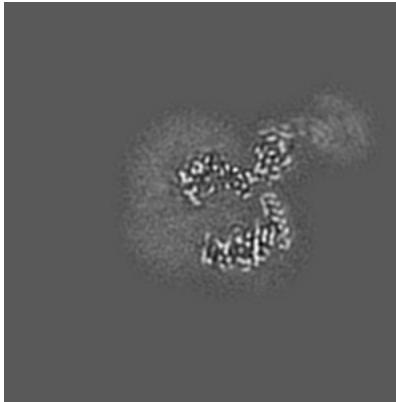
| X Index: 168 | Y Index: 162 | Z Index: 135 |
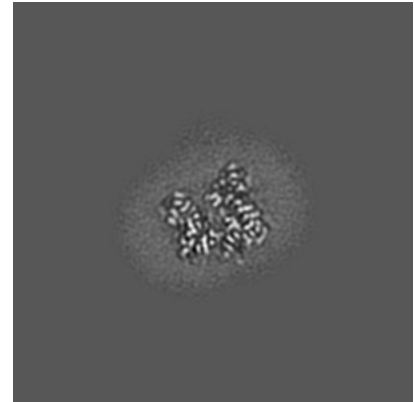
6.3.2 Raw map
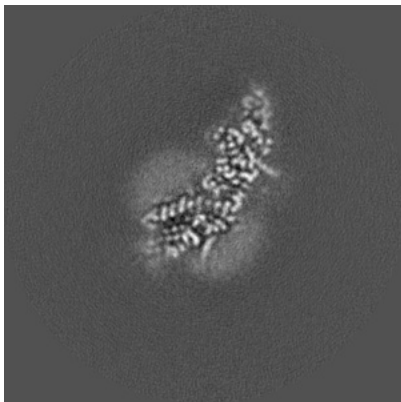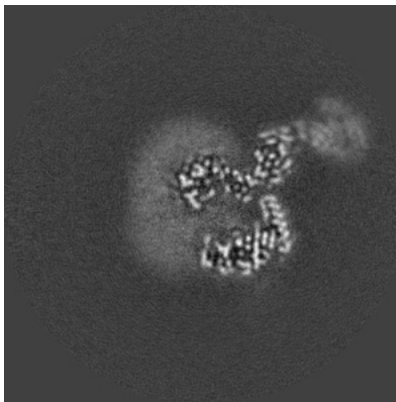
| X Index: 169 | Y Index: 162 | Z Index: 152 |
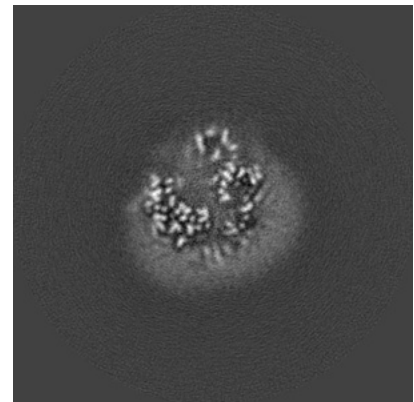
The images above show the largest variance slices of the map in three orthogonal directions.

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

6.4.1 Primary map
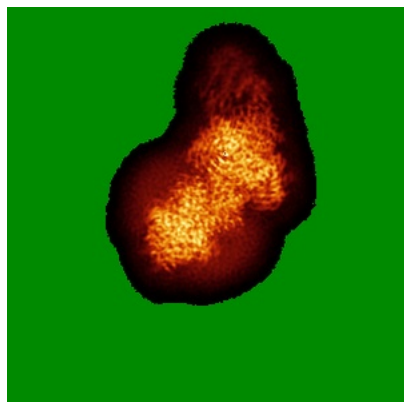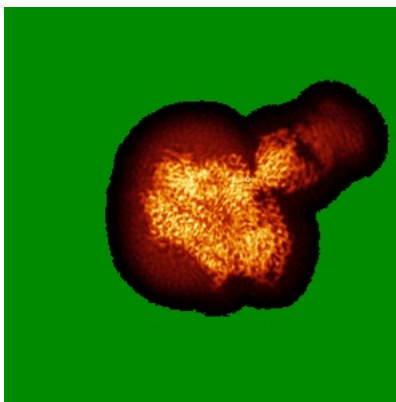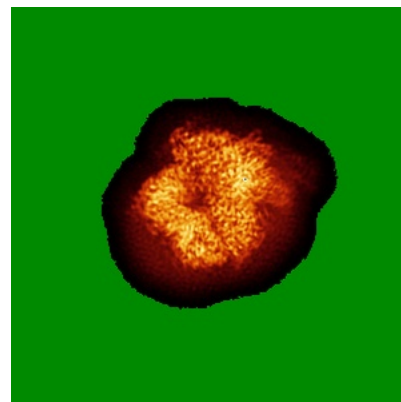
X   Y   Z

6.4.2 Raw map

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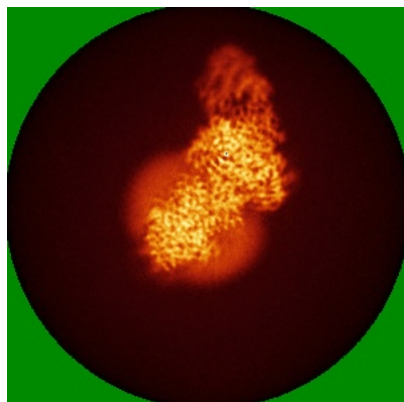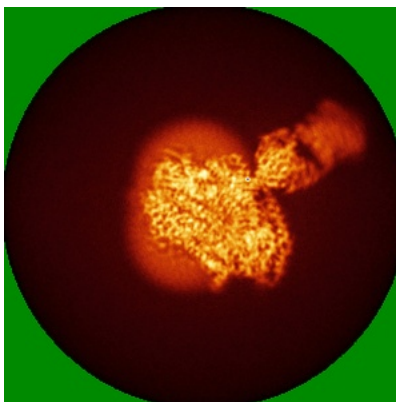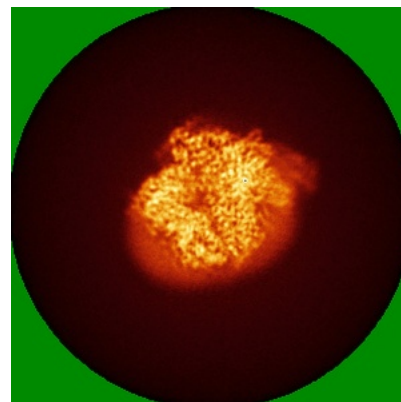
The images above show the map standard deviation projections with false color in three orthogonal directions. Minimum values are shown in green, max in blue, and dark to light orange shades represent small to large values respectively.

6.5   Orthogonal surface views ⓘ

6.5.1   Primary map

X	Y	Z
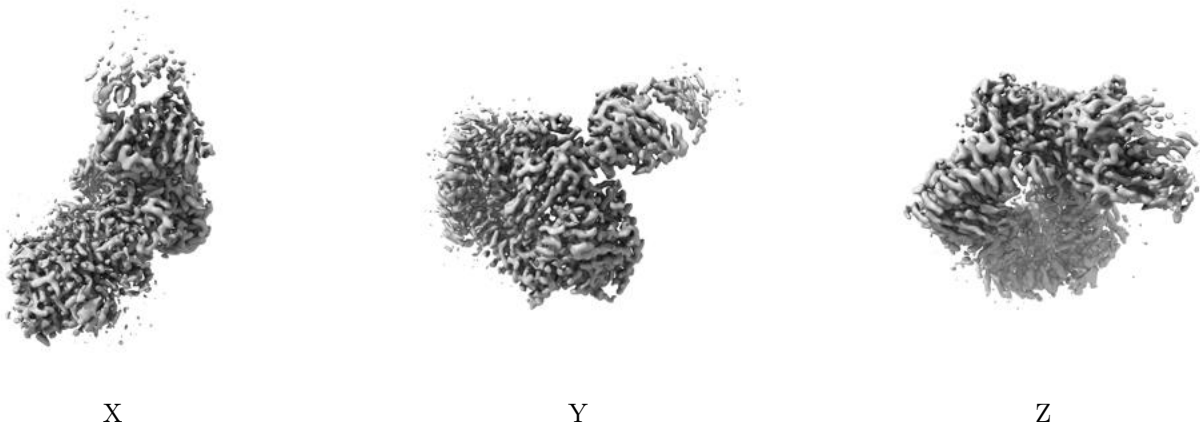
The images above show the 3D surface view of the map at the recommended contour level 0.015. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

6.5.2   Raw map

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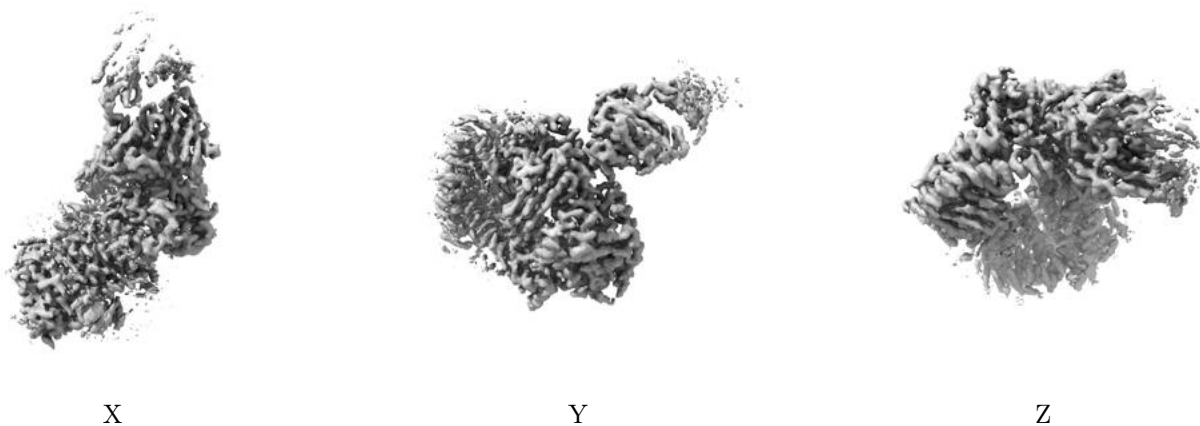
These images show the 3D surface of the raw map. The raw map's contour level was selected so that its surface encloses the same volume as the primary map does at its recommended contour level.

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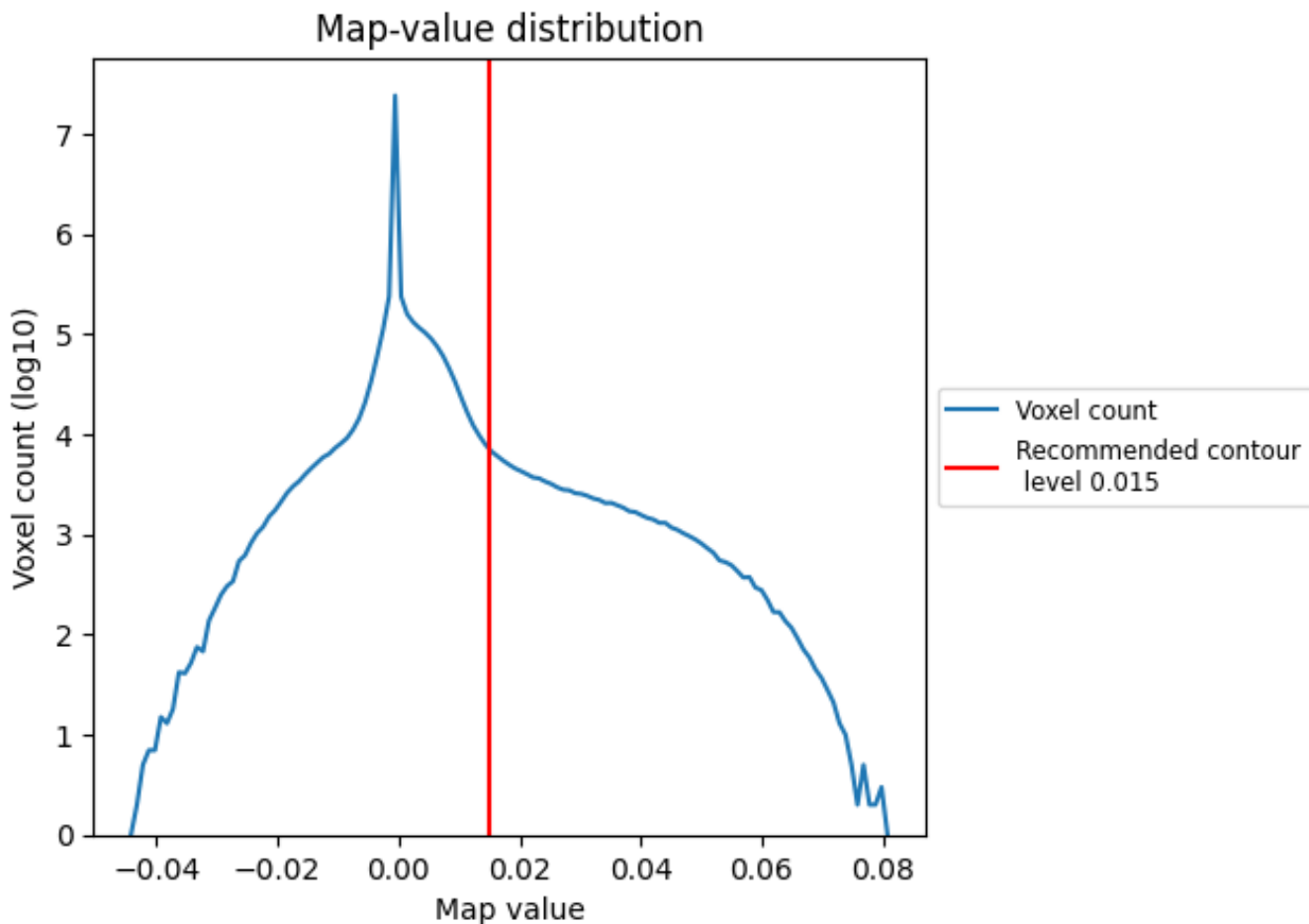
The map-value distribution is plotted in 128 intervals along the x-axis. The y-axis is logarithmic. A spike in this graph at zero usually indicates that the volume has been masked.

7.2　Volume estimate ⓘ
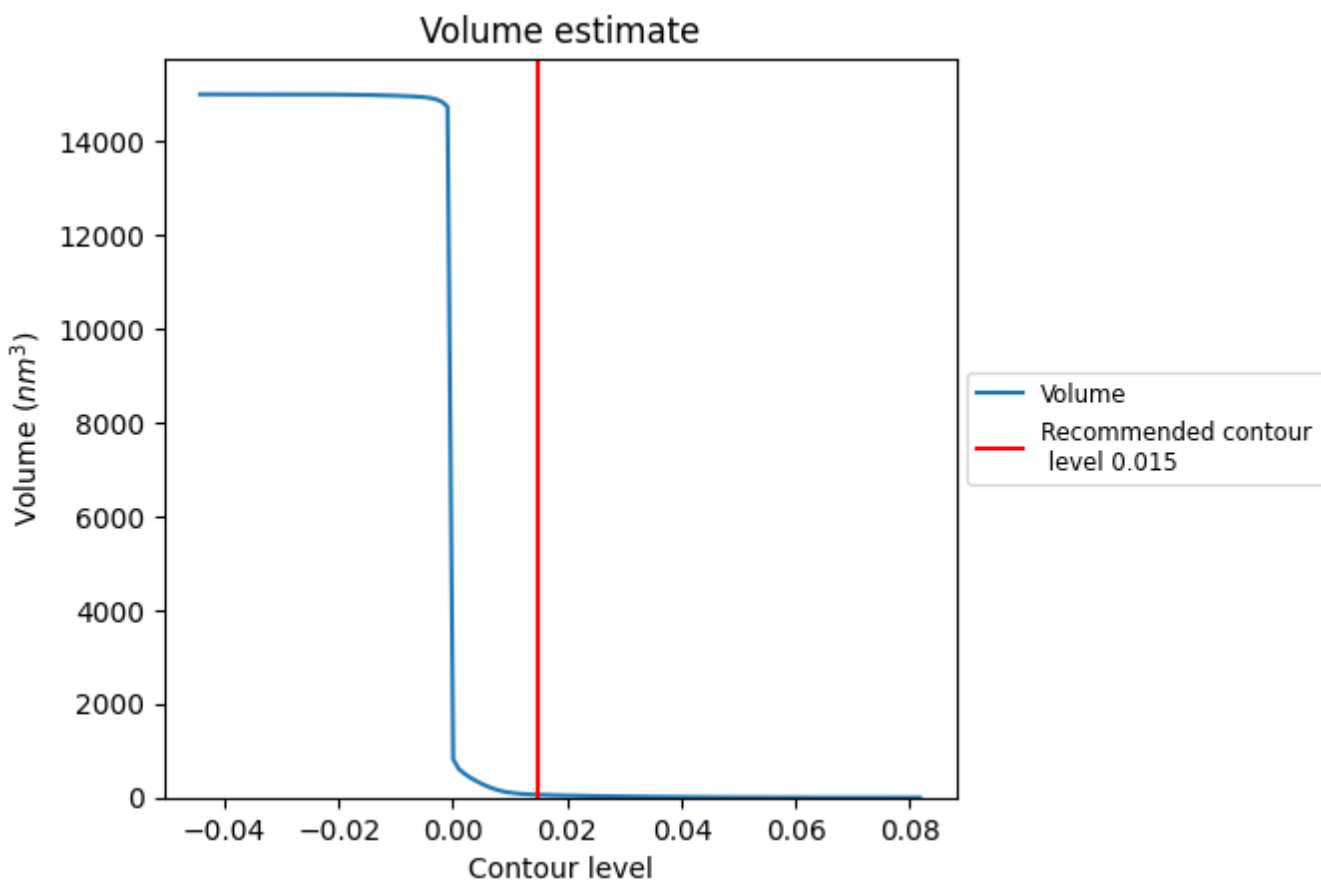
The volume at the recommended contour level is 60 nm^3; this corresponds to an approximate mass of 54 kDa.

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 ⓘ
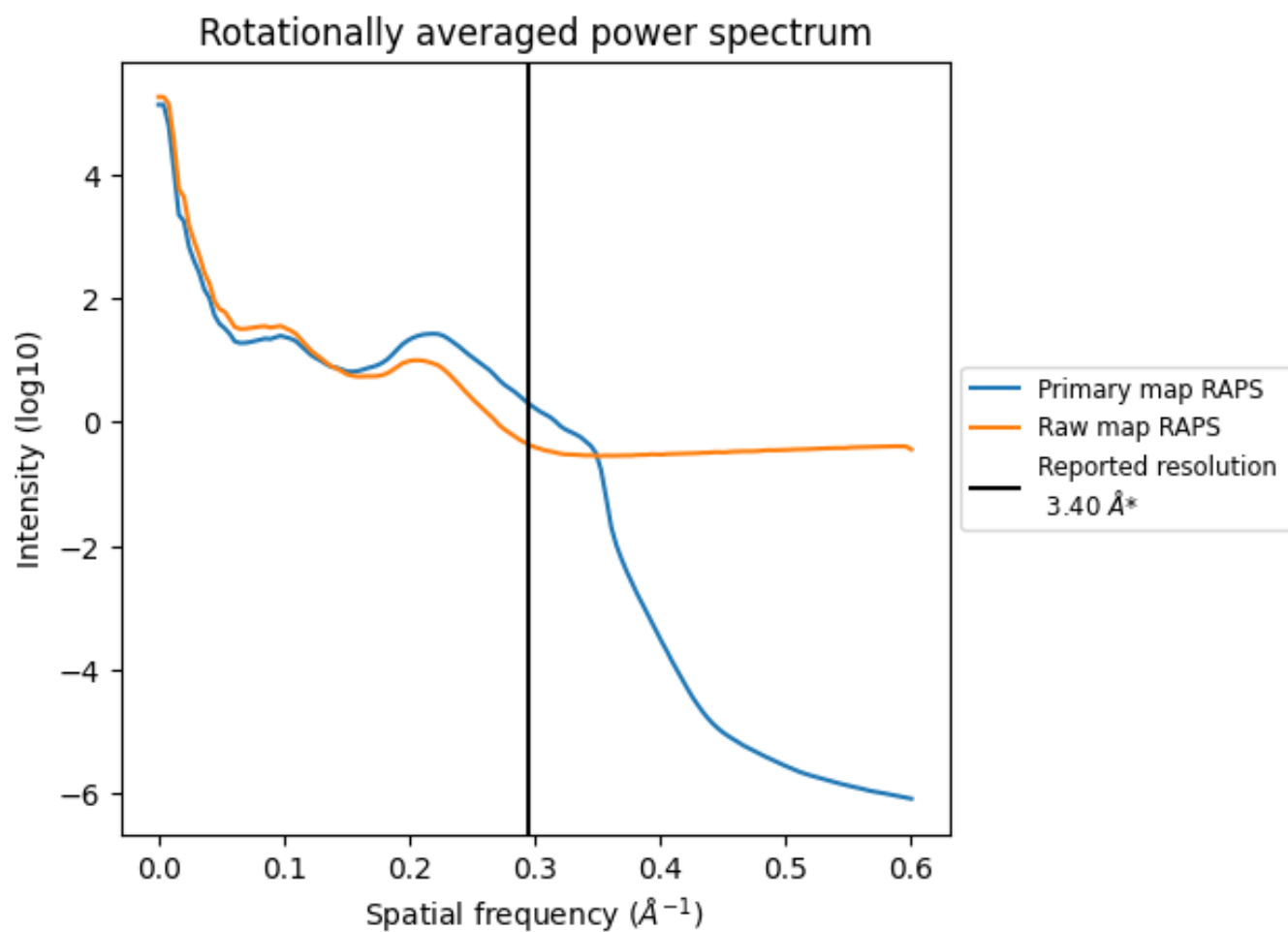
*Reported resolution corresponds to spatial frequency of 0.294 Å^{-1}

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)
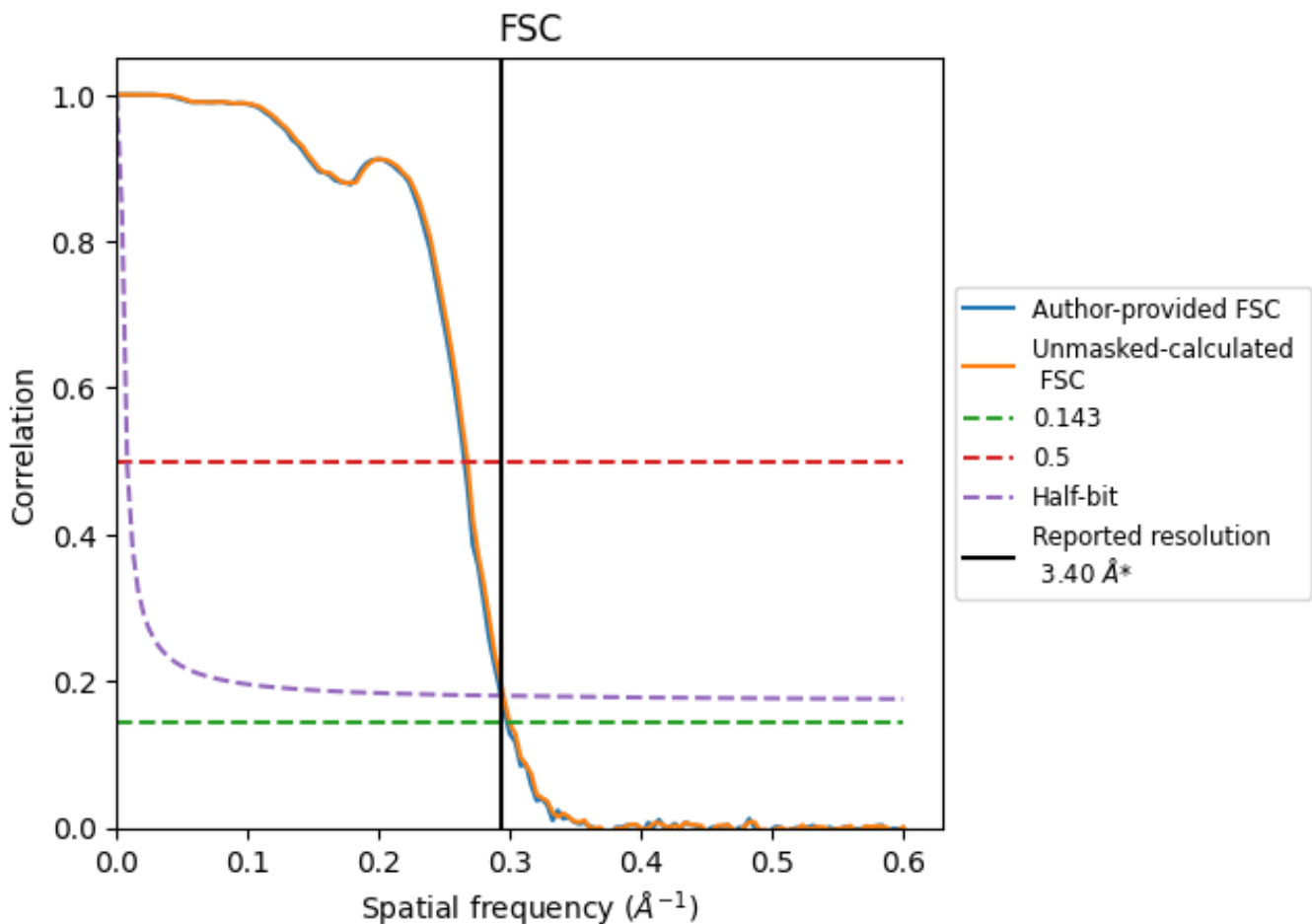
*Reported resolution corresponds to spatial frequency of 0.294 Å^{-1}

8.2 Resolution estimates (i)

Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	3.40	-	-
Author-provided FSC curve	3.36	3.77	3.41
Unmasked-calculated*	3.33	3.74	3.39

*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-24310 and PDB model 7R87. Per-residue inclusion information can be found in section 3 on page 6.

9.1 Map-model overlay ⓘ

| X | Y | Z |
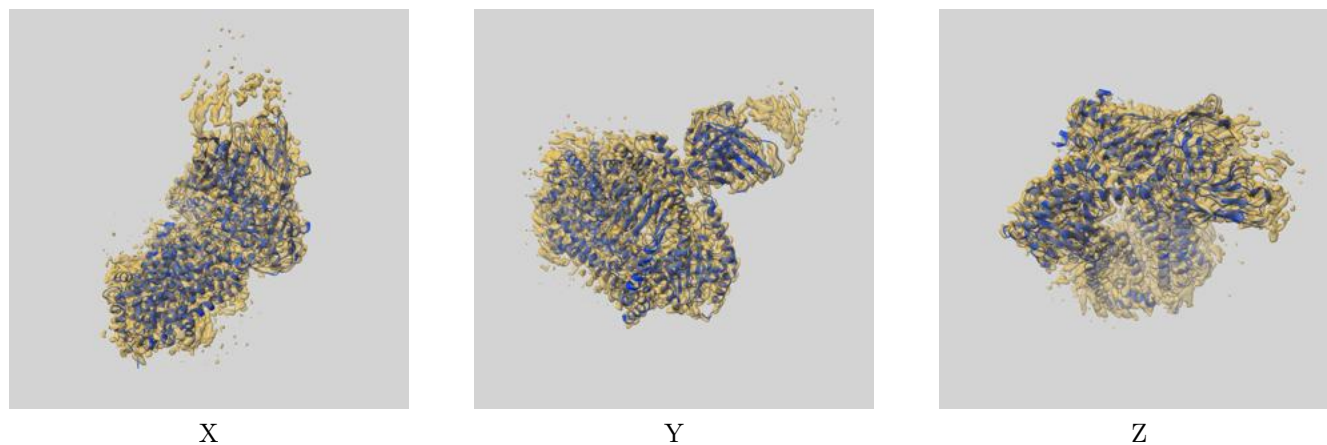
The images above show the 3D surface view of the map at the recommended contour level 0.015 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

9.2 Q-score mapped to coordinate model ⓘ

X   Y   Z
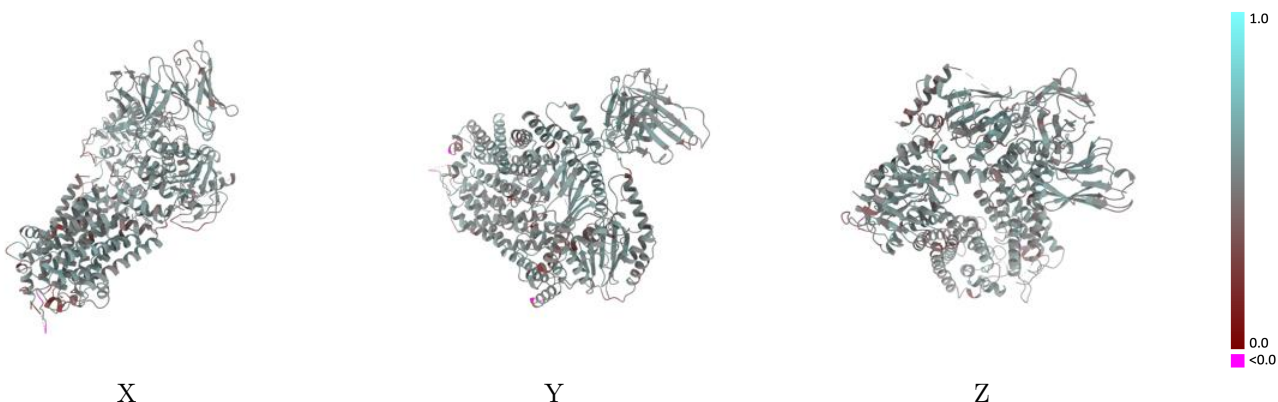
The images above show the model with each residue coloured according its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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
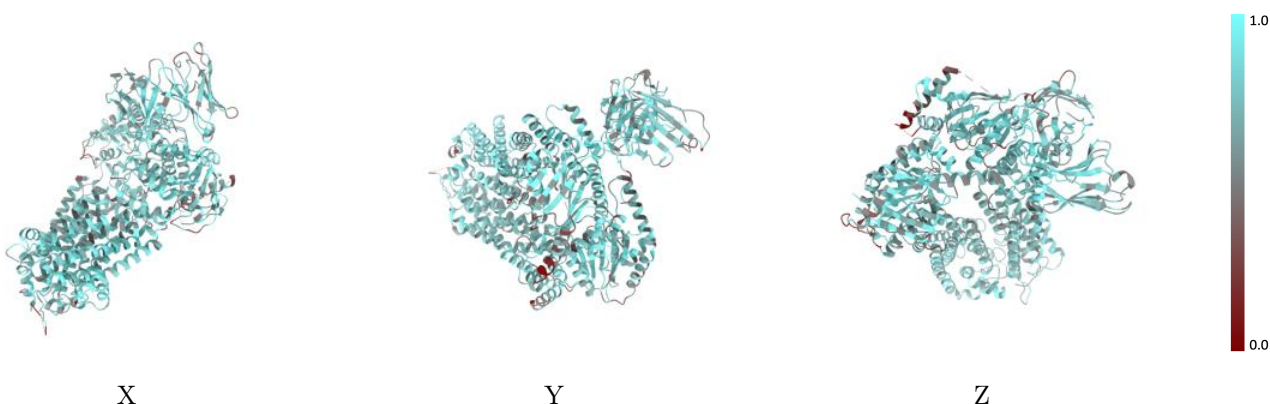
The images above show the model with each residue coloured according to its atom inclusion. This shows to what extent they are inside the map at the recommended contour level (0.015).

9.4   Atom inclusion ⓘ
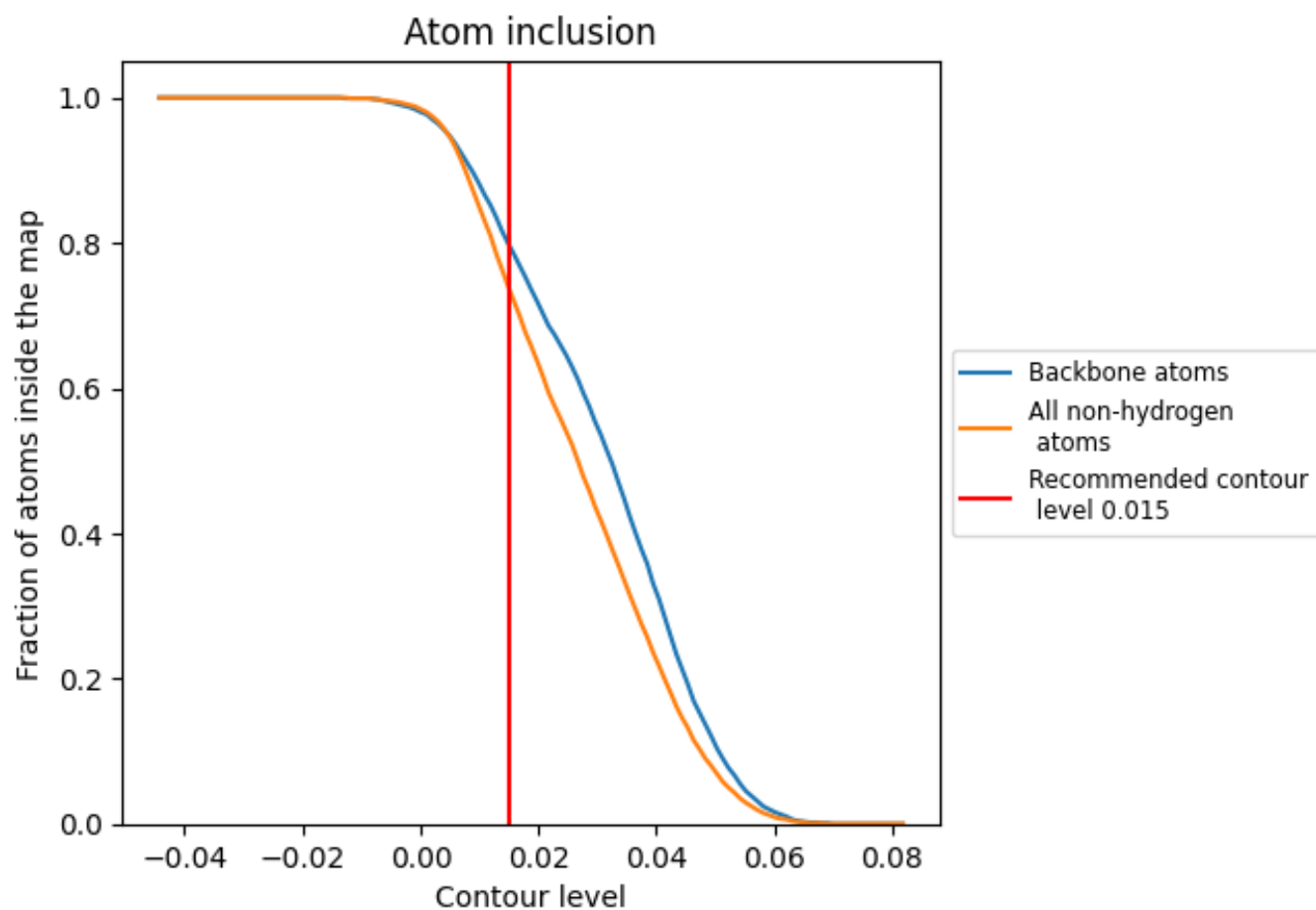
Atom inclusion

At the recommended contour level, 80% of all backbone atoms, 74% 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 (0.015) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.7370	0.5060
A	0.7380	0.5030
B	0.7450	0.5070
C	0.6910	0.5020
D	0.7310	0.5230

1.0

0.0
<0.0