WORLDWIDE wwPDB PROTEIN DATA BANK

Full wwPDB EM Validation Report (i)

May 20, 2024 – 07:00 PM EDT

PDB ID	:	7UM0
EMDB ID	:	EMD-24763
Title	:	Structure of the phage AR9 non-virion RNA polymerase holoenzyme in complex with two DNA oligonucleotides containing the AR9 P077 promoter as determined by cryo-EM
Authors	:	Leiman, P.G.; Fraser, A.; Sokolova, M.L.
Deposited on	:	2022-04-05
Resolution	:	3.80 Å(reported)
Based on initial model	:	7S01

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 3.80 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
Ramachandran outliers		0
Sidechain outliers		1.9%

Worse   *Better*

▮ Percentile relative to all structures
▯ Percentile relative to all EM structures
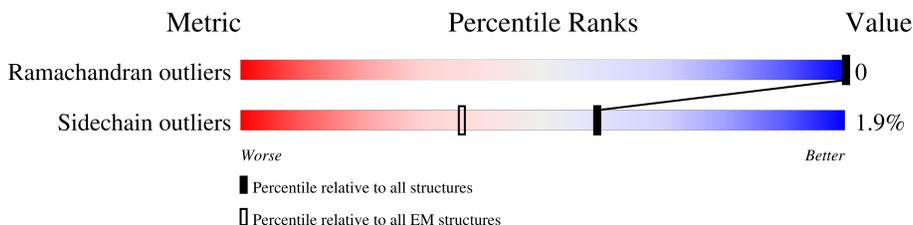
Metric	Whole archive (#Entries)	EM structures (#Entries)
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826

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	464	25% 97% ••
2	d	448	6% 94% • 5%
3	c	496	12% 95% ••
4	D	631	11% 98% •
5	C	665	• 96% ••
6	B	3	100%

2　Entry composition ⓘ

There are 7 unique types of molecules in this entry. The entry contains 21760 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-directed RNA polymerase subunit.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	459	Total 3814	C 2457	N 616	O 721	S 20	0	0

- Molecule 2 is a protein called DNA-directed RNA polymerase beta' subunit.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	d	426	Total 3488	C 2256	N 561	O 663	S 8	0	0

There are 22 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
d	-21	MET	-	expression tag	UNP A0A172JIH0
d	-20	GLY	-	expression tag	UNP A0A172JIH0
d	-19	SER	-	expression tag	UNP A0A172JIH0
d	-18	SER	-	expression tag	UNP A0A172JIH0
d	-17	HIS	-	expression tag	UNP A0A172JIH0
d	-16	HIS	-	expression tag	UNP A0A172JIH0
d	-15	HIS	-	expression tag	UNP A0A172JIH0
d	-14	HIS	-	expression tag	UNP A0A172JIH0
d	-13	HIS	-	expression tag	UNP A0A172JIH0
d	-12	HIS	-	expression tag	UNP A0A172JIH0
d	-11	SER	-	expression tag	UNP A0A172JIH0
d	-10	SER	-	expression tag	UNP A0A172JIH0
d	-9	GLY	-	expression tag	UNP A0A172JIH0
d	-8	GLU	-	expression tag	UNP A0A172JIH0
d	-7	ASN	-	expression tag	UNP A0A172JIH0
d	-6	LEU	-	expression tag	UNP A0A172JIH0
d	-5	TYR	-	expression tag	UNP A0A172JIH0
d	-4	PHE	-	expression tag	UNP A0A172JIH0
d	-3	GLN	-	expression tag	UNP A0A172JIH0
d	-2	GLY	-	expression tag	UNP A0A172JIH0
d	-1	HIS	-	expression tag	UNP A0A172JIH0

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
d	0	HIS	-	expression tag	UNP A0A172JIH0

- Molecule 3 is a protein called DNA-directed RNA polymerase beta subunit.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	c	484	Total 4003	C 2580	N 658	O 754	S 11	0	0

- Molecule 4 is a protein called DNA-directed RNA polymerase.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	D	629	Total 5123	C 3266	N 835	O 1005	S 17	0	0

- Molecule 5 is a protein called DNA-directed RNA polymerase.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	C	649	Total 5271	C 3337	N 868	O 1041	S 25	0	0

- Molecule 6 is a DNA chain called DNA (5'-D(P*GP*UP*U)-3').

Mol	Chain	Residues	Atoms					AltConf	Trace
6	B	3	Total 60	C 28	N 9	O 20	P 3	0	0

- Molecule 7 is ZINC ION (three-letter code: ZN) (formula: Zn).

| Mol	Chain	Residues	Atoms		AltConf
7	D	1	Total 1	Zn 1	0

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-directed RNA polymerase subunit

Chain A:
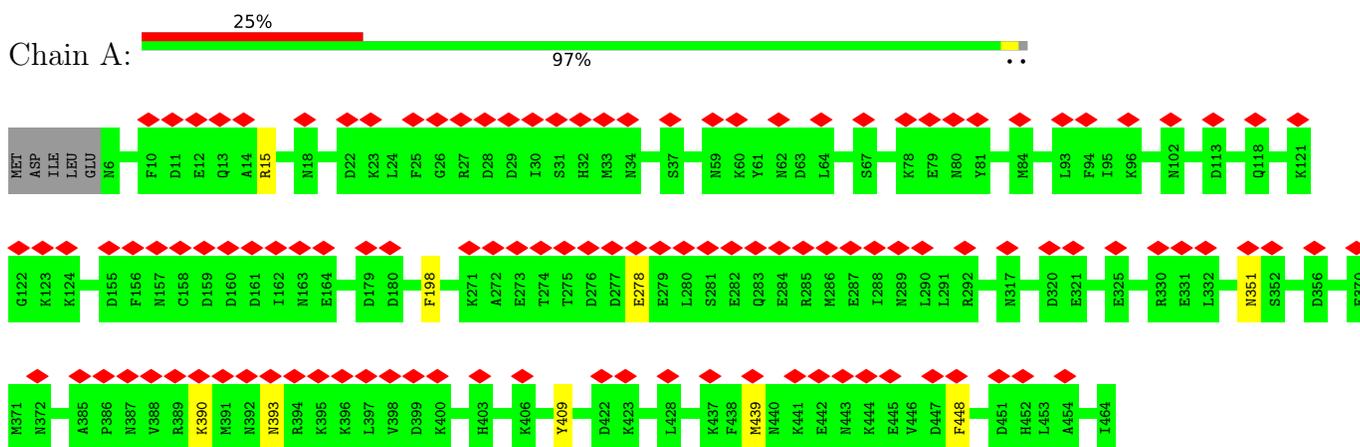
• Molecule 2: DNA-directed RNA polymerase beta' subunit
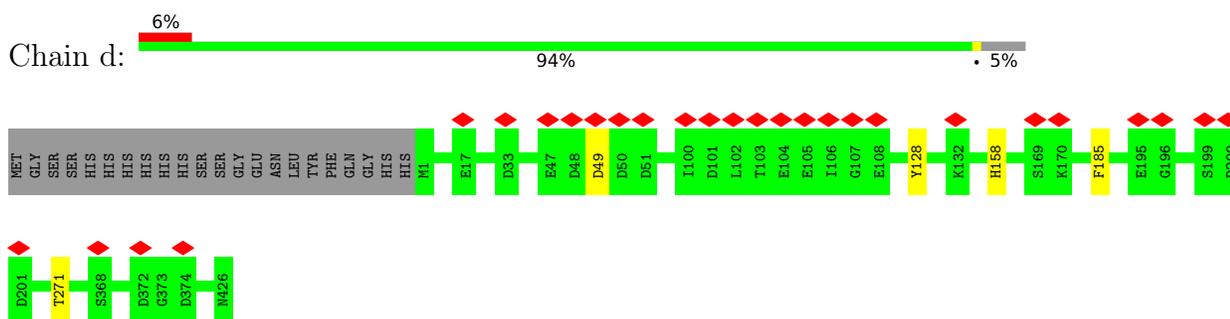
Chain d:

• Molecule 3: DNA-directed RNA polymerase beta subunit
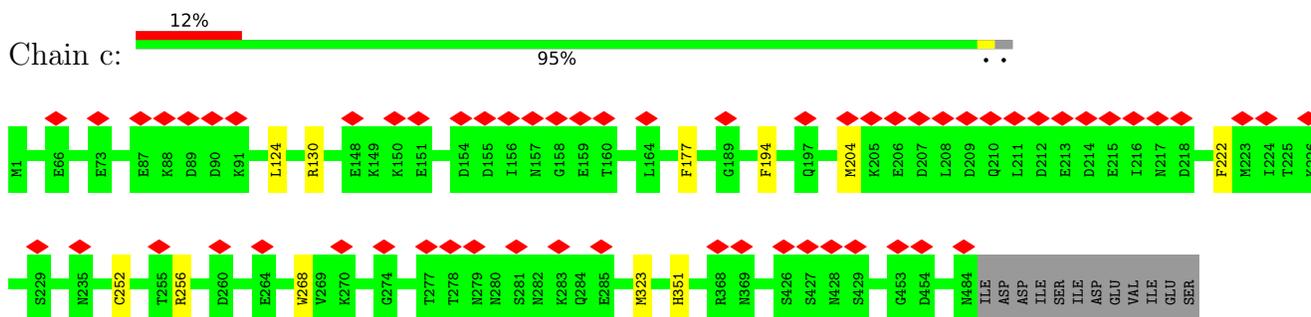
Chain c:

• Molecule 4: DNA-directed RNA polymerase
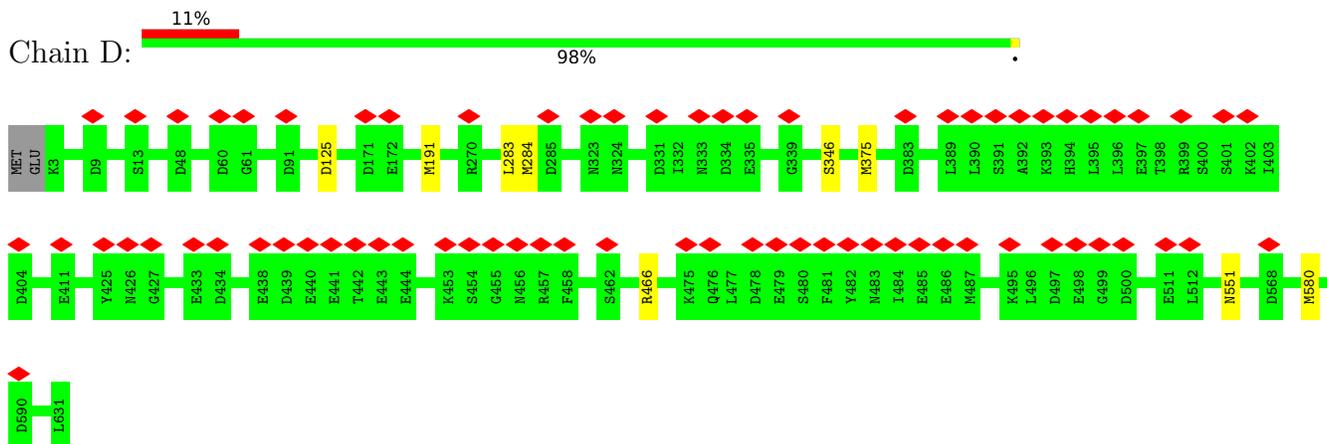
Chain D:

• Molecule 5: DNA-directed RNA polymerase
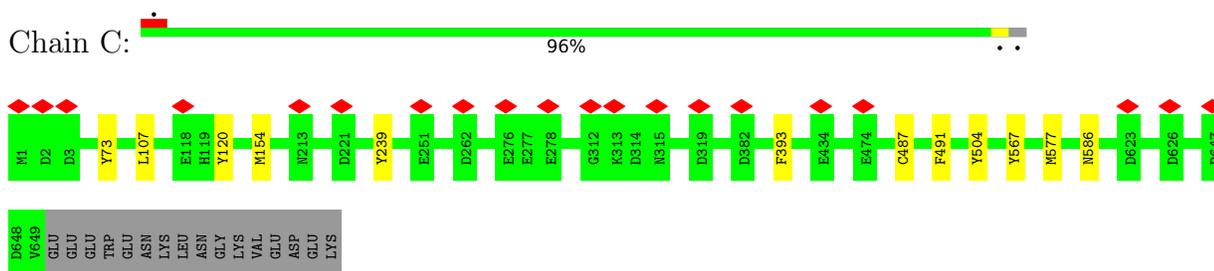
Chain C:

• Molecule 6: DNA (5'-D(P*GP*UP*U)-3')

Chain B:

There are no outlier residues recorded for this chain.

4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	106867	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 $(e^-/\text{Å}^2)$	44	Depositor
Minimum defocus (nm)	1000	Depositor
Maximum defocus (nm)	4000	Depositor
Magnification	Not provided	
Image detector	GATAN K3 BIOQUANTUM (6k x 4k)	Depositor
Maximum map value	17.809	Depositor
Minimum map value	-6.673	Depositor
Average map value	0.000	Depositor
Map value standard deviation	1.000	Depositor
Recommended contour level	4.0	Depositor
Map size (Å)	216.00002, 216.00002, 216.00002	wwPDB
Map dimensions	200, 200, 200	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.08, 1.08, 1.08	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: ZN

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	0.25	0/3876	0.45	0/5214
2	d	0.25	0/3558	0.43	0/4804
3	c	0.25	0/4082	0.45	0/5502
4	D	0.24	0/5205	0.44	0/7016
5	C	0.25	0/5370	0.45	0/7241
6	B	0.72	0/66	0.97	0/99
All	All	0.25	0/22157	0.45	0/29876

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2   Too-close contacts (i)

Due to software issues we are unable to calculate clashes - this section is therefore empty.

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	457/464 (98%)	441 (96%)	16 (4%)	0	100	100
2	d	424/448 (95%)	414 (98%)	10 (2%)	0	100	100
3	c	482/496 (97%)	470 (98%)	12 (2%)	0	100	100
4	D	627/631 (99%)	617 (98%)	10 (2%)	0	100	100
5	C	647/665 (97%)	622 (96%)	25 (4%)	0	100	100
All	All	2637/2704 (98%)	2564 (97%)	73 (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	440/445 (99%)	431 (98%)	9 (2%)	55	75
2	d	390/409 (95%)	385 (99%)	5 (1%)	69	82
3	c	457/469 (97%)	446 (98%)	11 (2%)	49	71
4	D	588/590 (100%)	579 (98%)	9 (2%)	65	81
5	C	593/608 (98%)	581 (98%)	12 (2%)	55	75
All	All	2468/2521 (98%)	2422 (98%)	46 (2%)	59	76

All (46) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	15	ARG
1	A	198	PHE
1	A	278	GLU
1	A	351	ASN
1	A	390	LYS
1	A	393	ASN
1	A	409	TYR
1	A	439	MET
1	A	448	PHE
2	d	49	ASP

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	d	128	TYR
2	d	158	HIS
2	d	185	PHE
2	d	271	THR
3	c	124	LEU
3	c	130	ARG
3	c	177	PHE
3	c	194	PHE
3	c	204	MET
3	c	222	PHE
3	c	252	CYS
3	c	256	ARG
3	c	268	TRP
3	c	323	MET
3	c	351	HIS
4	D	125	ASP
4	D	191	MET
4	D	283	LEU
4	D	284	MET
4	D	346	SER
4	D	375	MET
4	D	466	ARG
4	D	551	ASN
4	D	580	MET
5	C	73	TYR
5	C	107	LEU
5	C	120	TYR
5	C	154	MET
5	C	239	TYR
5	C	393	PHE
5	C	487	CYS
5	C	491	PHE
5	C	504	TYR
5	C	567	TYR
5	C	577	MET
5	C	586	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (32) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	117	ASN
1	A	157	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	163	ASN
1	A	372	ASN
1	A	393	ASN
2	d	24	ASN
2	d	99	ASN
2	d	120	ASN
2	d	158	HIS
2	d	282	ASN
3	c	14	GLN
3	c	279	ASN
3	c	369	ASN
3	c	474	HIS
4	D	8	ASN
4	D	177	ASN
4	D	239	ASN
4	D	299	ASN
4	D	412	ASN
4	D	494	ASN
4	D	551	ASN
4	D	574	GLN
5	C	16	HIS
5	C	29	ASN
5	C	50	GLN
5	C	153	ASN
5	C	203	ASN
5	C	208	ASN
5	C	412	ASN
5	C	465	ASN
5	C	488	GLN
5	C	559	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 monosaccharides in this entry.

5.6   Ligand geometry Ⓘ

Of 1 ligands modelled in this entry, 1 is monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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 ⓘ

This section contains visualisations of the EMDB entry EMD-24763. 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 ⓘ
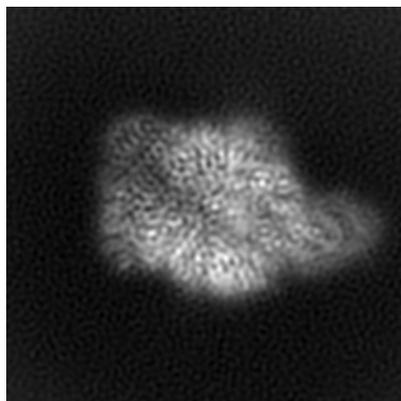
6.1.1 Primary map

X
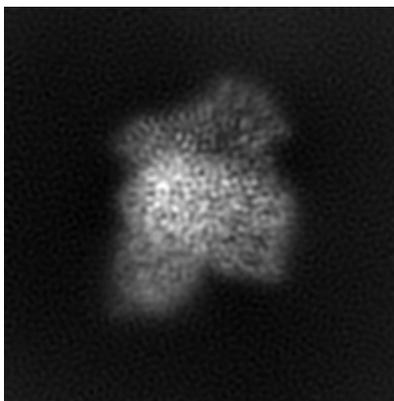
Y
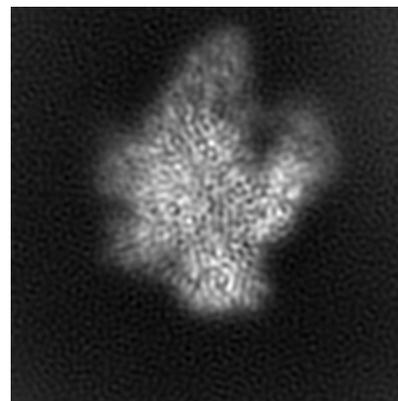
Z

6.1.2 Raw map

X

Y

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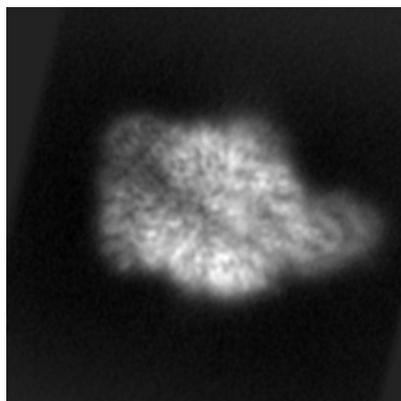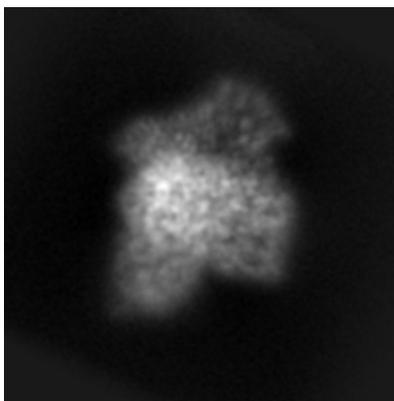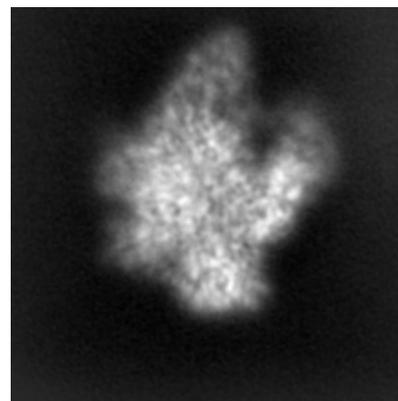
The images above show the map projected in three orthogonal directions.

6.2　Central slices ⓘ
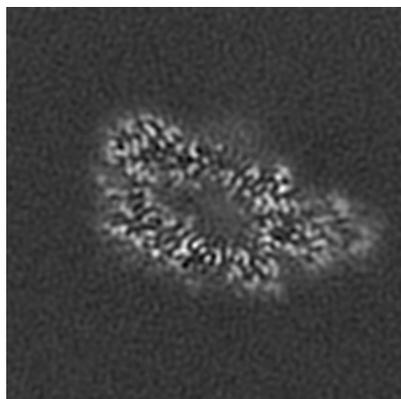
6.2.1　Primary map
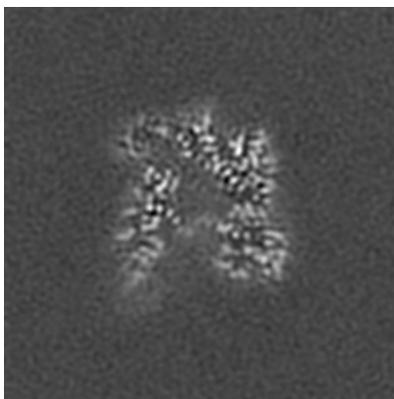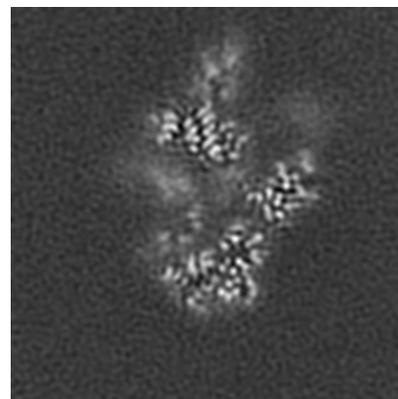
X Index: 100　　　　Y Index: 100　　　　Z Index: 100

6.2.2　Raw map
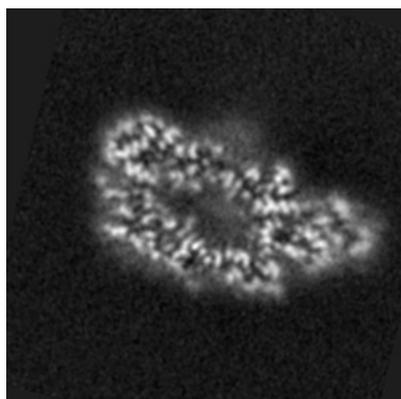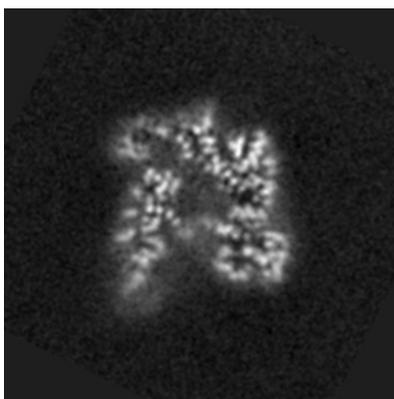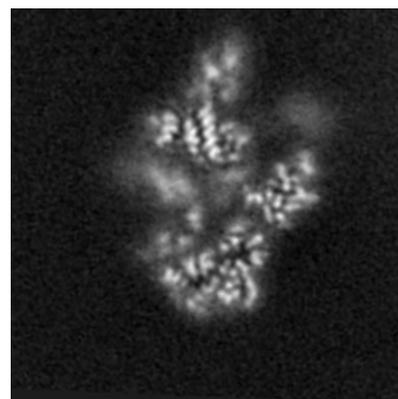
X Index: 100　　　　Y Index: 100　　　　Z Index: 100

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

6.3 Largest variance slices ⓘ
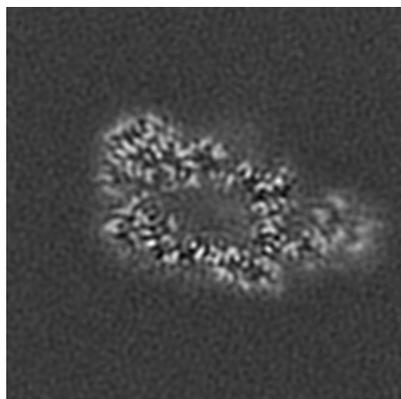
6.3.1 Primary map
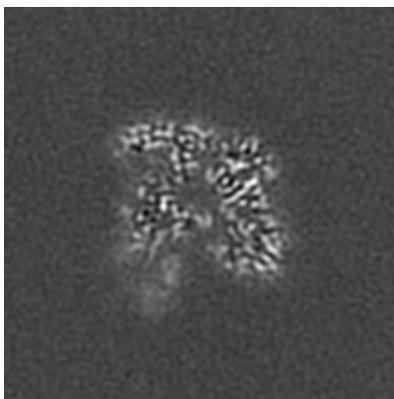
| X Index: 102 | Y Index: 94 | Z Index: 77 |
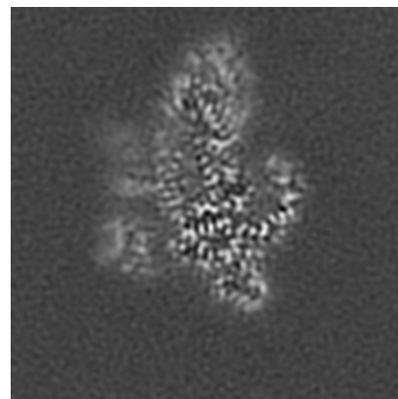
6.3.2 Raw map
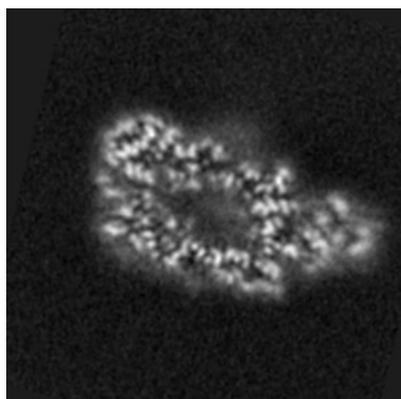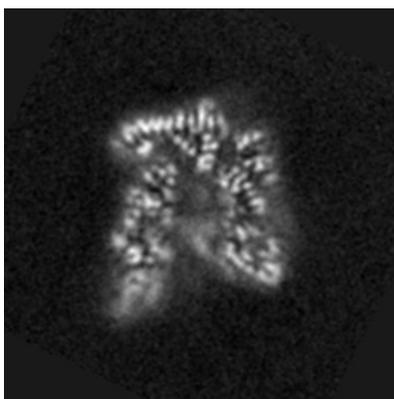
| X Index: 101 | Y Index: 105 | Z Index: 78 |
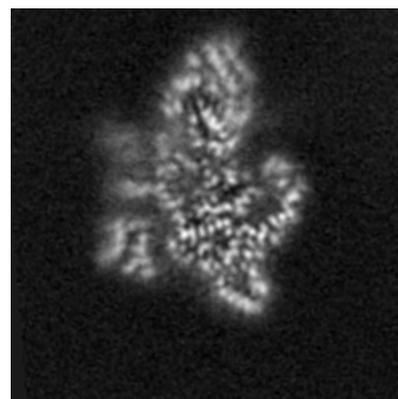
The images above show the largest variance slices of the map in three orthogonal directions.

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

6.4.1 Primary map
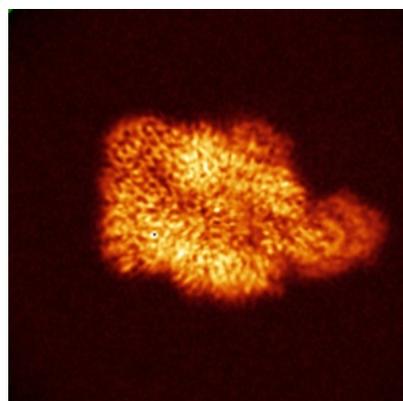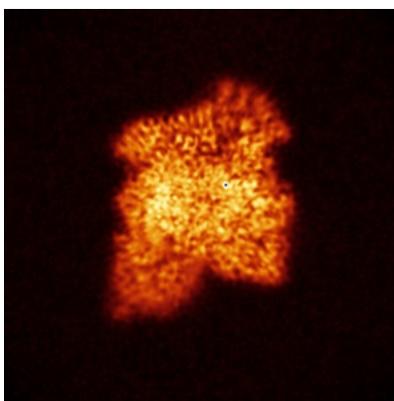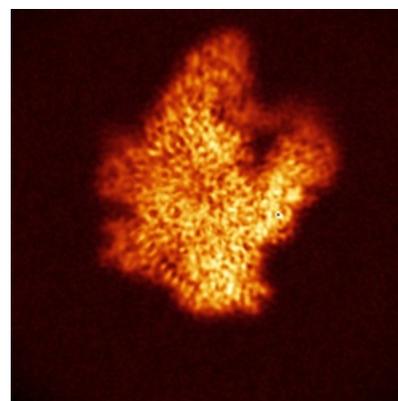
| X | Y | Z |

6.4.2 Raw map

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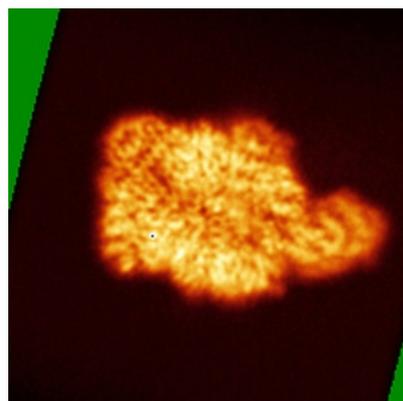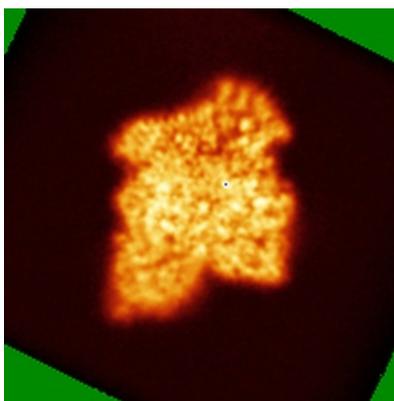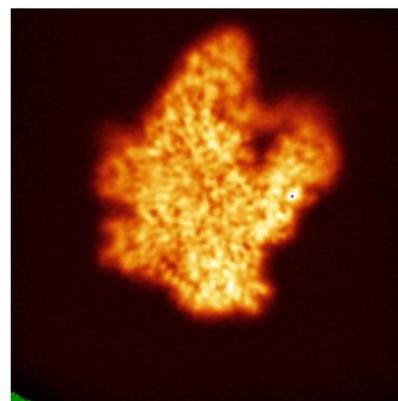
The images above show the map standard deviation projections with false color in three orthogonal directions. Minimum values are shown in green, max in blue, and dark to light orange shades represent small to large values respectively.

6.5 Orthogonal surface views ⓘ

6.5.1 Primary map

X   Y   Z
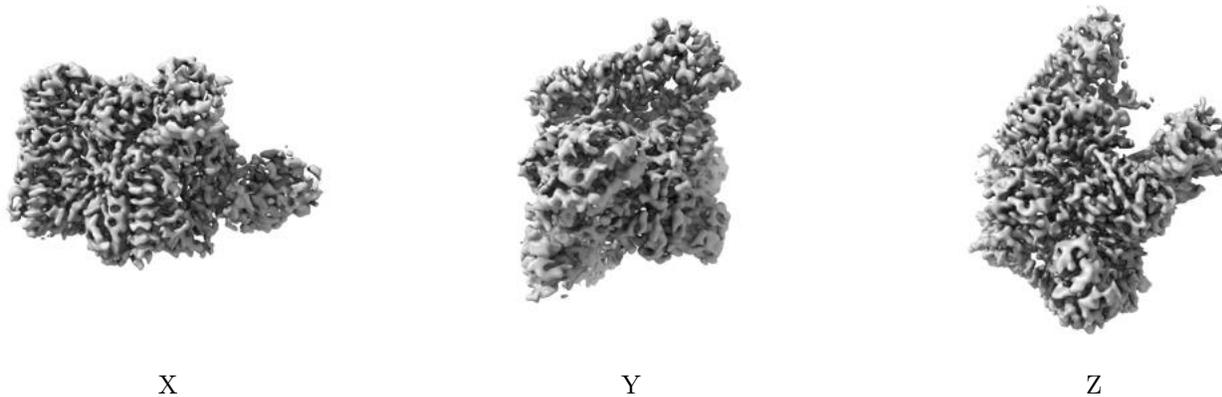
The images above show the 3D surface view of the map at the recommended contour level 4.0. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

6.5.2 Raw map

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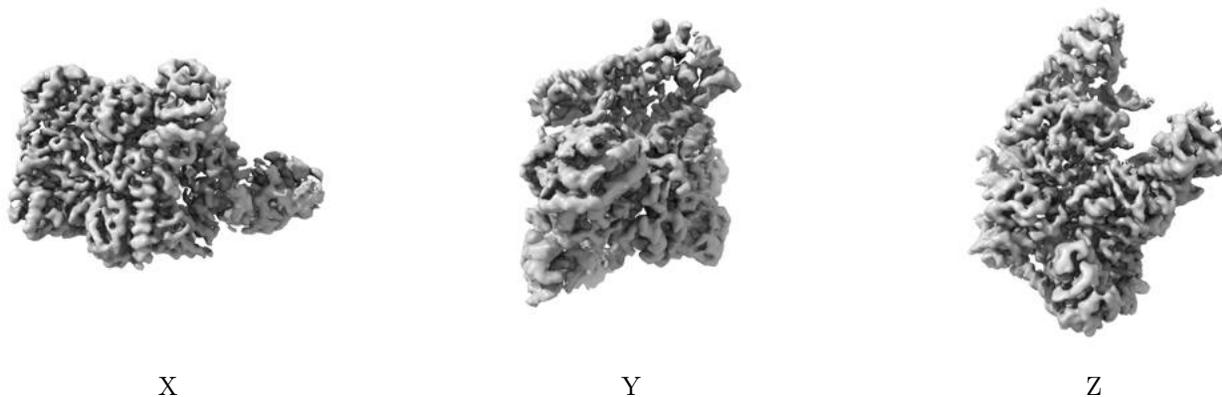
These images show the 3D surface of the raw map. The raw map's contour level was selected so that its surface encloses the same volume as the primary map does at its recommended contour level.

6.6 Mask visualisation ⓘ

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

7 Map analysis (i)

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

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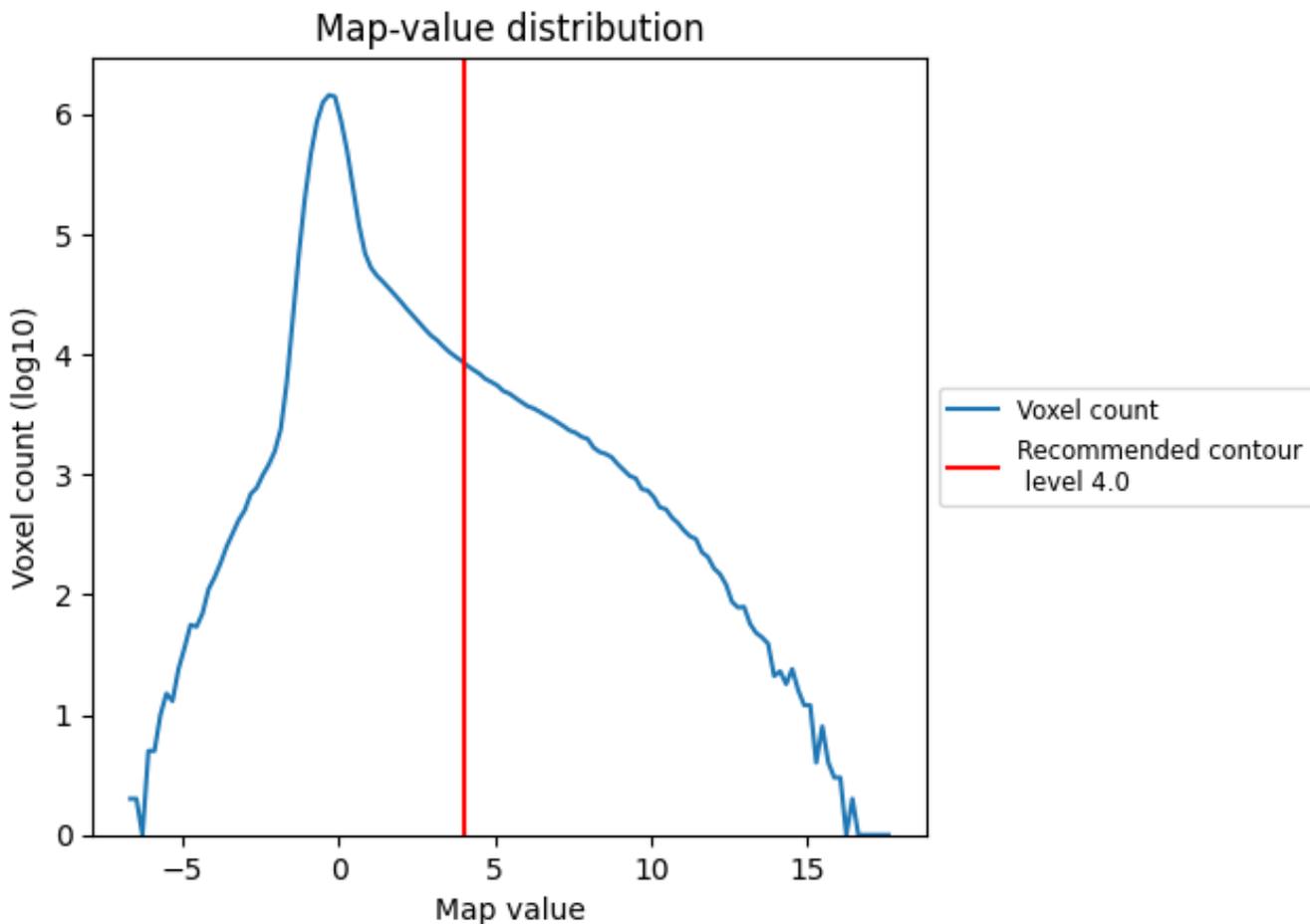
The map-value distribution is plotted in 128 intervals along the x-axis. The y-axis is logarithmic. A spike in this graph at zero usually indicates that the volume has been masked.

7.2 Volume estimate ⓘ
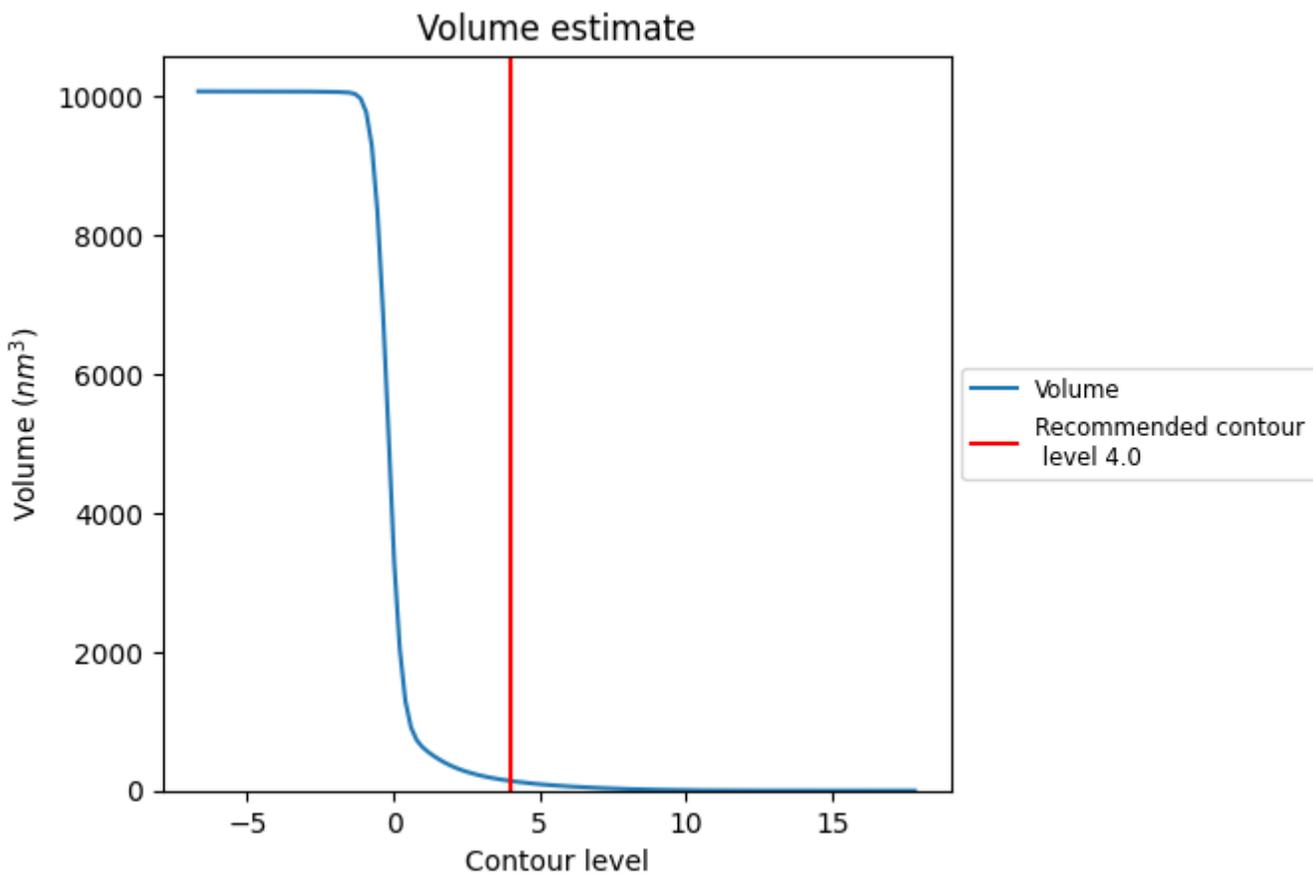
The volume at the recommended contour level is 139 nm^3; this corresponds to an approximate mass of 126 kDa.

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 ⓘ
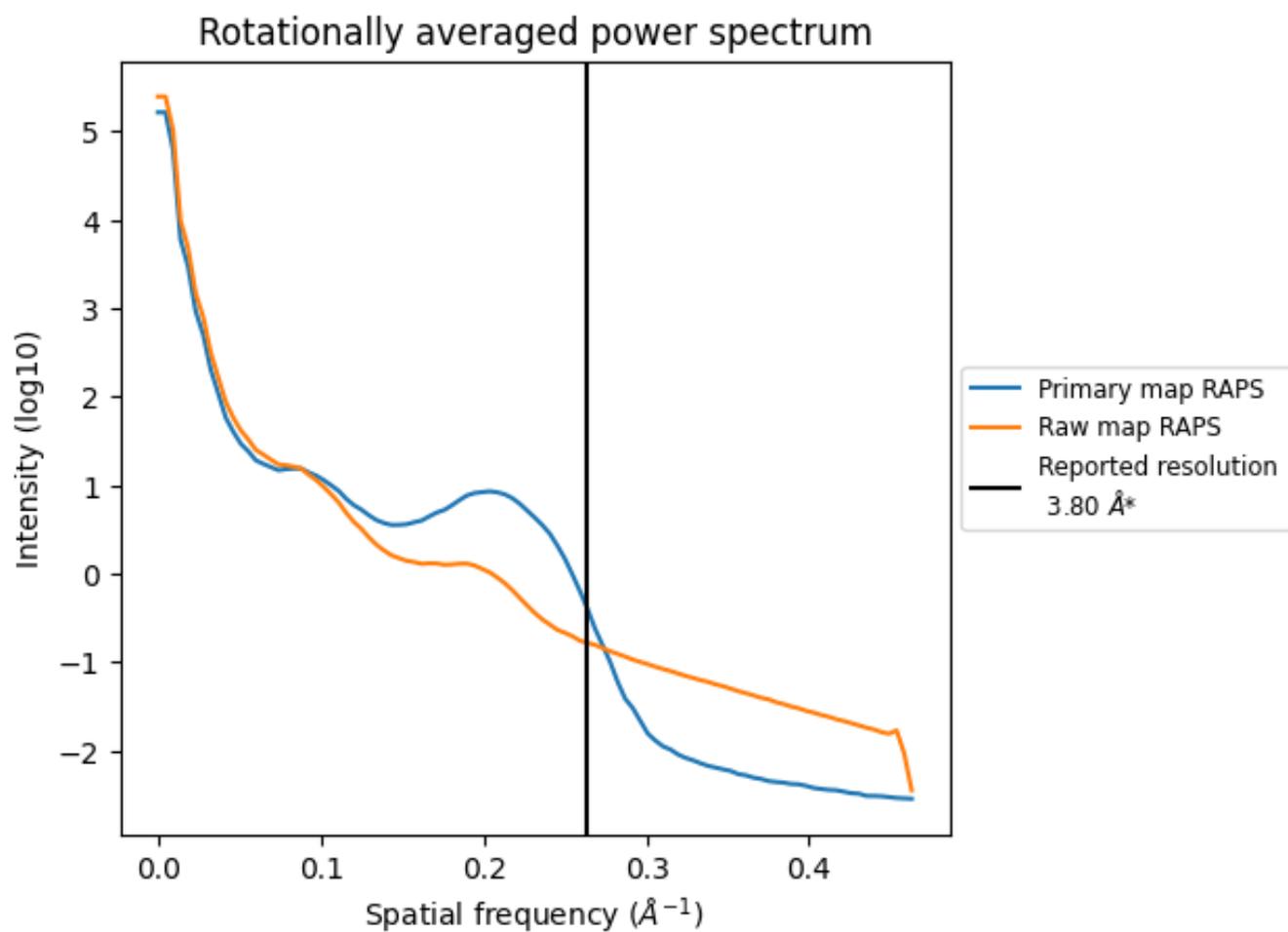
*Reported resolution corresponds to spatial frequency of 0.263 Å^{-1}

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)
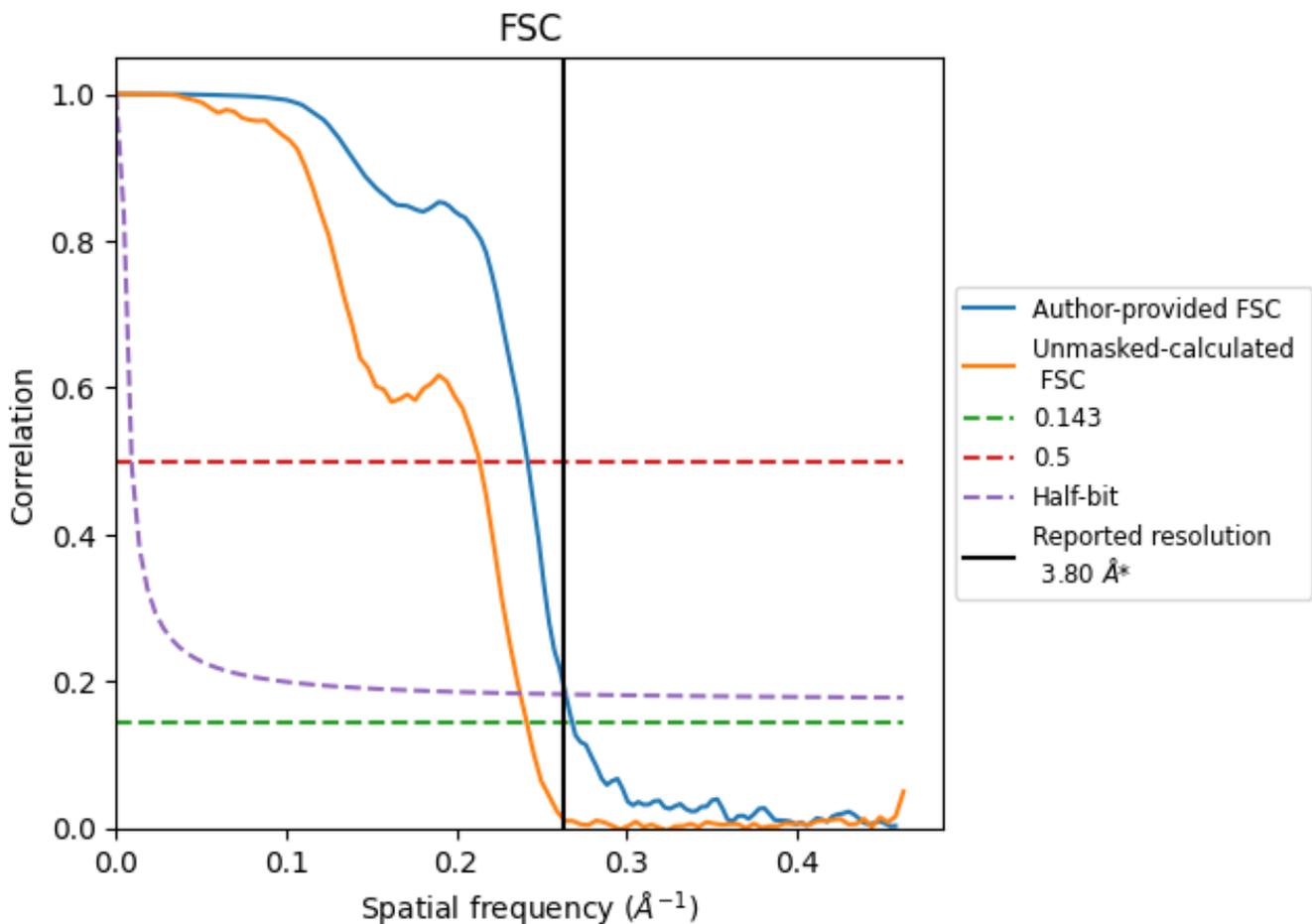
*Reported resolution corresponds to spatial frequency of 0.263 Å^{-1}

8.2 Resolution estimates (i)

| Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	3.80	-	-
Author-provided FSC curve	3.73	4.13	3.78
Unmasked-calculated*	4.14	4.68	4.21

*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-24763 and PDB model 7UM0. Per-residue inclusion information can be found in section 3 on page 5.

9.1 Map-model overlay ⓘ
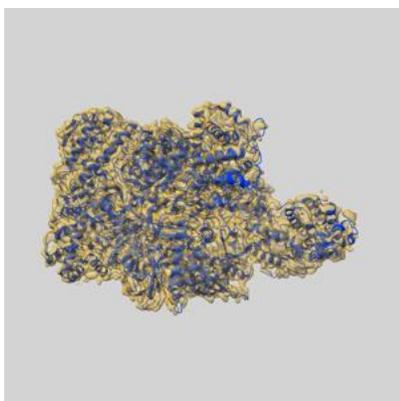
X　　　　　　　　　　　Y　　　　　　　　　　　Z
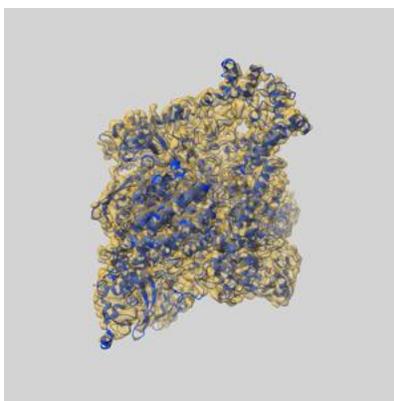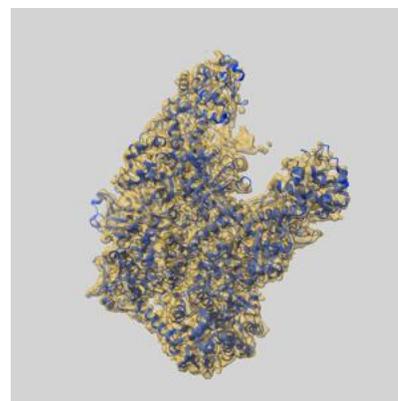
The images above show the 3D surface view of the map at the recommended contour level 4.0 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

9.2　Q-score mapped to coordinate model ⓘ

X　　　　　　　　　　　　　　　　Y　　　　　　　　　　　　　　　　Z
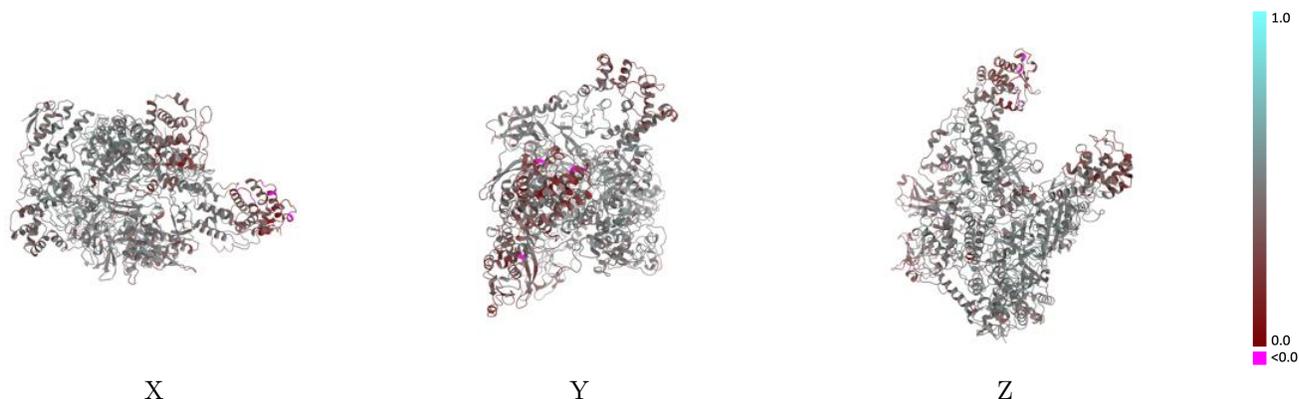
The images above show the model with each residue coloured according its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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
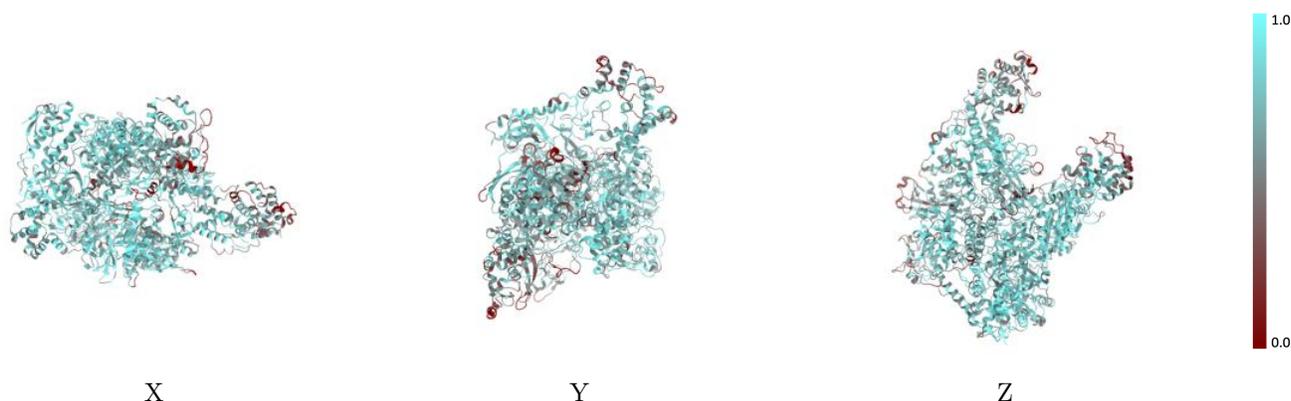
The images above show the model with each residue coloured according to its atom inclusion. This shows to what extent they are inside the map at the recommended contour level (4.0).

9.4 Atom inclusion ⓘ
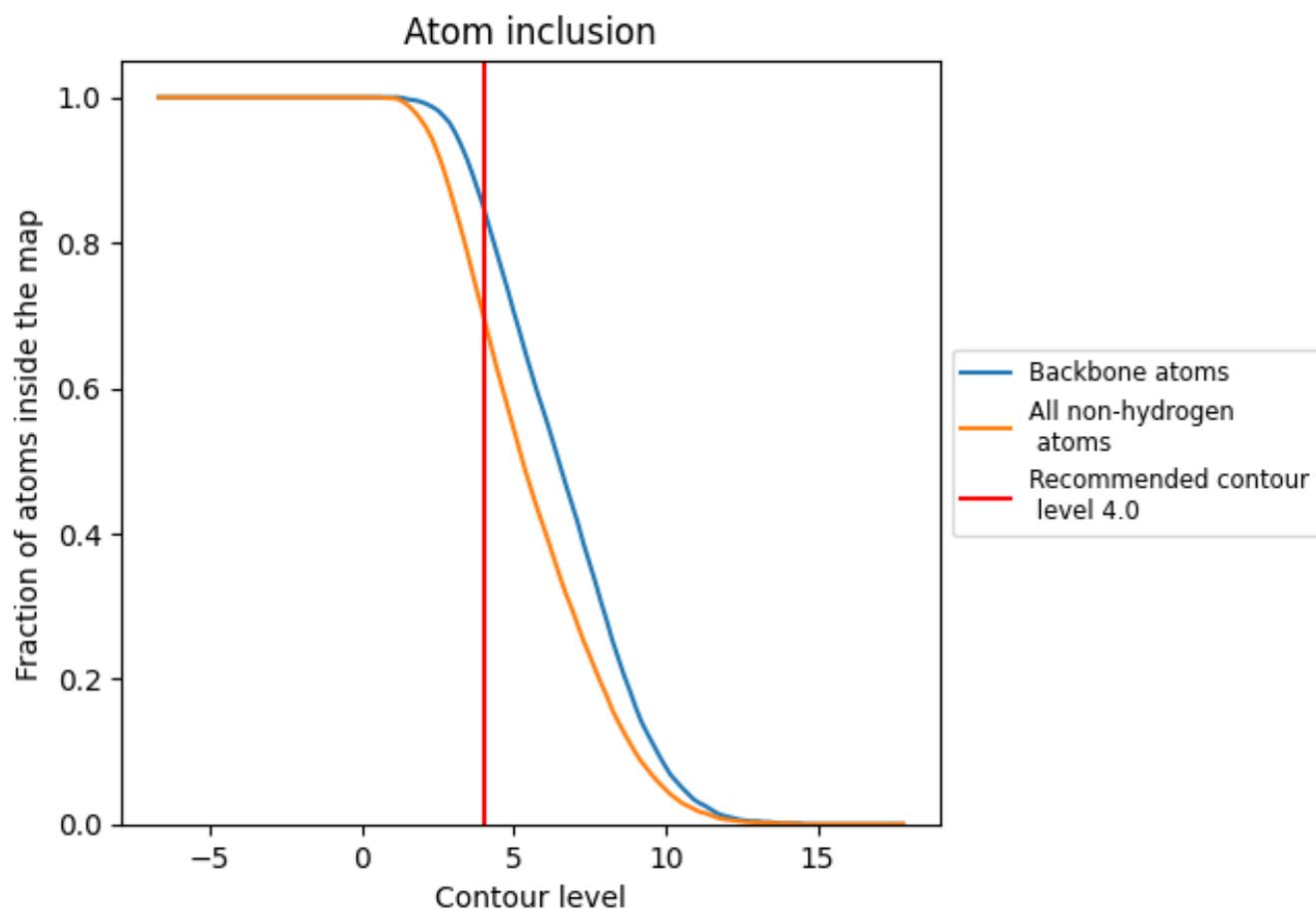
At the recommended contour level, 85% of all backbone atoms, 70% of all non-hydrogen atoms, are inside the map.

9.5　Map-model fit summary ⓘ

The table lists the average atom inclusion at the recommended contour level (4.0) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.7000	0.4370
A	0.5730	0.3520
B	0.9330	0.5030
C	0.7680	0.4730
D	0.7020	0.4410
c	0.6760	0.4300
d	0.7540	0.4740

1.0

0.0
<0.0