wwPDB EM Validation Summary Report ⓘ

Mar 20, 2024 – 02:09 PM JST

PDB ID	:	7D7C
EMDB ID	:	EMD-30604
Title	:	CryoEM structure of gp55-dependent RNA polymerase-promoter open complex
Authors	:	Shi, J.; Wen, A.; Jin, S.; Feng, Y.
Deposited on	:	2020-10-03
Resolution	:	3.60 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 3.60 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		17
Ramachandran outliers		0.1%
Sidechain outliers		0.4%

Worse   *Better*

∎ Percentile relative to all structures
▯ Percentile relative to all EM structures
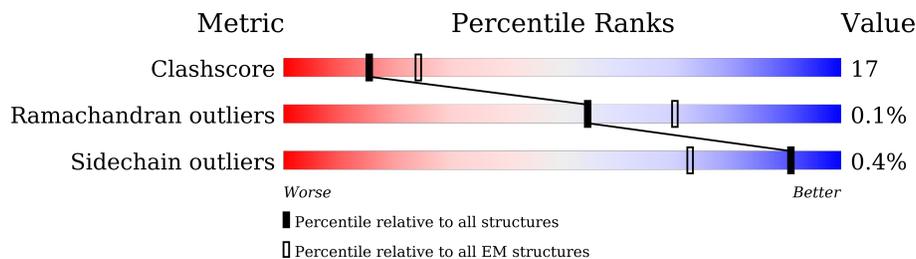
Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826

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	329	41% 25% • 33%
1	B	329	41% 26% 34%
2	C	1342	63% 34% •
3	D	1407	61% 33% 5%
4	T	59	8% 31% 61%
5	F	185	43% 31% • 26%
6	N	59	8% 34% 58%

2 Entity composition ⓘ

There are 8 unique types of molecules in this entry. The entry contains 26249 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 alpha.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	219	Total	C	N	O	S	0	0
			1686	1056	298	326	6		
1	B	218	Total	C	N	O	S	0	0
			1681	1050	297	328	6		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
2	C	1316	Total	C	N	O	S	0	0
			10381	6514	1810	2014	43		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
3	D	1335	Total	C	N	O	S	0	0
			10384	6524	1851	1959	50		

- Molecule 4 is a DNA chain called DNA (template strand).

Mol	Chain	Residues	Atoms					AltConf	Trace
4	T	23	Total	C	N	O	P	0	0
			467	225	75	144	23		

- Molecule 5 is a protein called gp55.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	F	137	Total	C	N	O	S	0	0
			1128	720	189	214	5		

- Molecule 6 is a DNA chain called DNA (nontemplate strand).

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
6	N	25	519	246	105	143	25	0	0

- Molecule 7 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		AltConf
			Total	Mg	
7	D	1	1	1	0

- Molecule 8 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		AltConf
			Total	Zn	
8	D	2	2	2	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 alpha

Chain A:
41% 25% • 33%

Chain B:
41% 26% 34%

- Molecule 1: DNA-directed RNA polymerase subunit alpha

- Molecule 2: DNA-directed RNA polymerase subunit beta

Chain C:

• Molecule 3: DNA-directed RNA polymerase subunit beta'

Chain D:
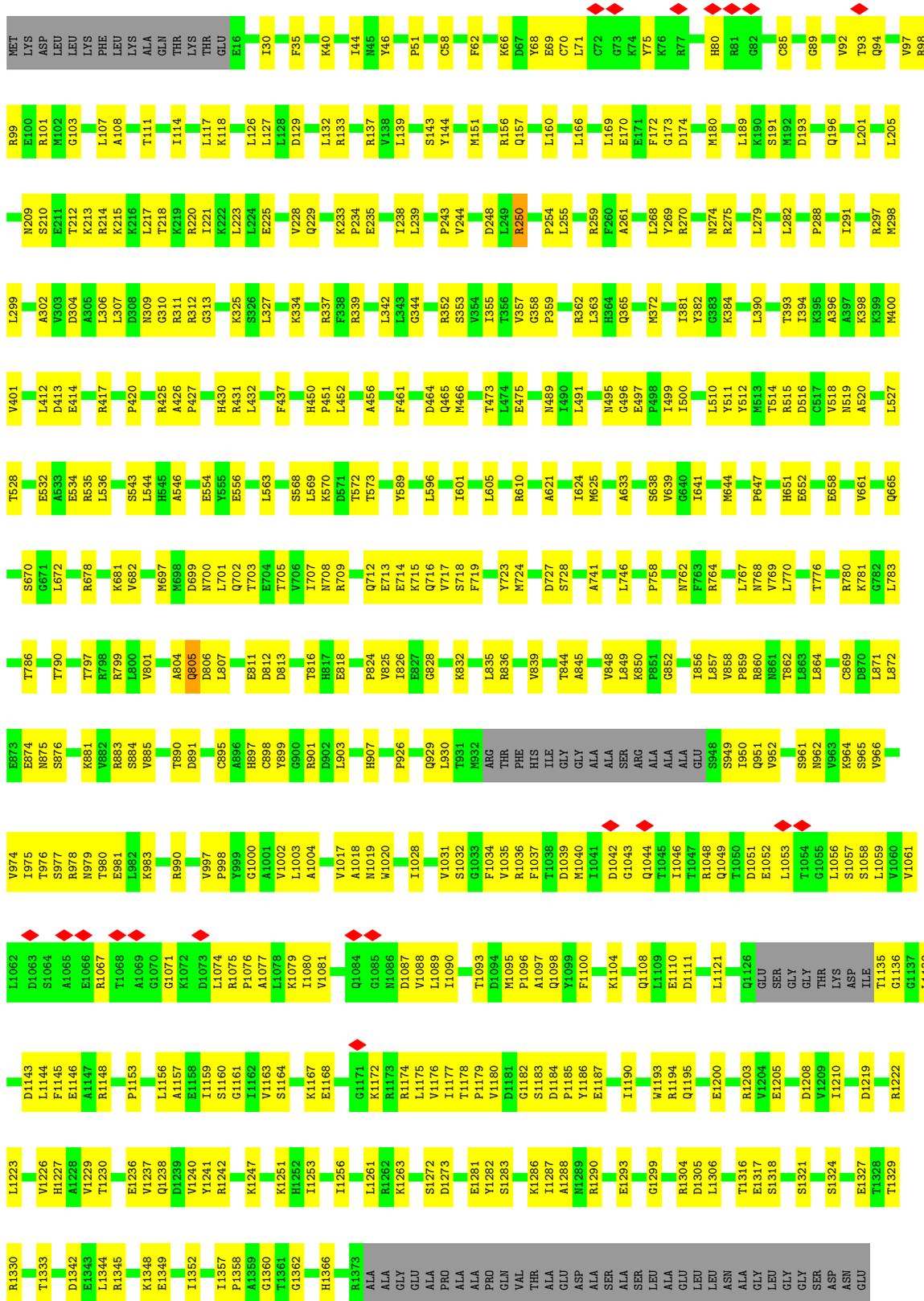
• Molecule 4: DNA (template strand)

Chain T: 8%   31%   61%

G1 G2 C3 T4 G5 C6 T7 T8 C9 A10 G11 T12 A13 DT DC DA DG DG DA DG DT DA T22 T23 T24 A25 T26 A27 C28 T29 C30 T31 DC DA DG DT DA DA DT DA DG DT DG DC DT DG DA DG DC DT DC DT DT DT DA DT DT DA DG

• Molecule 5: gp55

Chain F: 43%   31%   •   26%

MET SER GLU THR LYS PRO LYS TYR ASN Y10 V11 N12 N13 K14 L17 Q18 A19 I20 I21 K24 N29 ASN LYS ASP PRO ASN LYS VAL V37 L49 L54 S55 K56 R57 F58 N59 F60 S61 G62 Y63 T64 K68 Q69 E70 M71 G75 I76 E77 N85 F86 D87 E88

T89 K90 Y91 Y97 I98 T99 Q100 A101 N104 A105 F106 V107 Q108 R109 I110 K111 K112 E113 R114 K115 A118 K119 V125 H126 N127 V128 Y129 D130 S131 R132 D133 V137 A138 L139 F144 I145 D150 K151 M152 T153 HIS TYR GLU GLU SER THR TYR ARG THR PRO GLY ALA GLU

LYS LYS SER VAL VAL ASP ASP SER PRO SER LEU ASP PHE LEU TYR GLU ALA ASN ASP

• Molecule 6: DNA (nontemplate strand)

Chain N: 8%   34%   58%

DC DT DA DA DT DA DA DA DG DA DG DC DT DC DA DG DC DA DC DT DA DT DT DA DC DT DG A22 G23 A24 G25 T26 A27 T28 A29 A30 A31 DT DA DC DT DC DC DT G38 A39 T40 A41 C42 T43 G44 A45 A46 G47 C51 C52

4 Experimental information ⓘ

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	57776	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 $(e^-/\text{Å}^2)$	59	Depositor
Minimum defocus (nm)	Not provided	
Maximum defocus (nm)	Not provided	
Magnification	Not provided	
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor
Maximum map value	0.080	Depositor
Minimum map value	-0.035	Depositor
Average map value	0.000	Depositor
Map value standard deviation	0.003	Depositor
Recommended contour level	0.01	Depositor
Map size (Å)	334.592, 334.592, 334.592	wwPDB
Map dimensions	256, 256, 256	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.307, 1.307, 1.307	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: MG, 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	0.40	0/1706	0.51	0/2312
1	B	0.35	0/1700	0.49	0/2304
2	C	0.42	1/10547 (0.0%)	0.50	0/14232
3	D	0.38	0/10541	0.49	0/14232
4	T	0.67	0/519	1.08	0/796
5	F	0.37	0/1149	0.48	0/1546
6	N	0.71	0/584	0.90	0/897
All	All	0.41	1/26746 (0.0%)	0.53	0/36319

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	1084	ASP	C-N	-5.29	1.21	1.34

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1686	0	1726	58	0
1	B	1681	0	1714	66	0
2	C	10381	0	10392	356	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	D	10384	0	10608	355	0
4	T	467	0	265	14	0
5	F	1128	0	1105	59	0
6	N	519	0	281	28	0
7	D	1	0	0	0	0
8	D	2	0	0	0	0
All	All	26249	0	26091	862	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 17.

The worst 5 of 862 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:D:1161:GLY:HA3	3:D:1178:THR:O	1.43	1.15
3:D:556:GLU:O	3:D:563:LEU:HA	1.64	0.98
3:D:1087:ASP:HB3	3:D:1096:PRO:HG2	1.50	0.93
2:C:589:THR:HG23	2:C:591:TYR:CE2	2.07	0.90
2:C:105:TYR:HA	2:C:114:VAL:HA	1.62	0.82

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	215/329 (65%)	193 (90%)	22 (10%)	0	100	100
1	B	214/329 (65%)	193 (90%)	21 (10%)	0	100	100
2	C	1312/1342 (98%)	1140 (87%)	170 (13%)	2 (0%)	47	79
3	D	1329/1407 (94%)	1181 (89%)	147 (11%)	1 (0%)	51	83
5	F	133/185 (72%)	118 (89%)	15 (11%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
All	All	3203/3592 (89%)	2825 (88%)	375 (12%)	3 (0%)	54	83

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	C	45	GLY
2	C	50	GLU
3	D	805	GLN

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	186/286 (65%)	184 (99%)	2 (1%)	73	88
1	B	186/286 (65%)	186 (100%)	0	100	100
2	C	1135/1157 (98%)	1130 (100%)	5 (0%)	91	97
3	D	1120/1168 (96%)	1118 (100%)	2 (0%)	93	98
5	F	121/166 (73%)	118 (98%)	3 (2%)	47	75
All	All	2748/3063 (90%)	2736 (100%)	12 (0%)	91	97

5 of 12 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	D	250	ARG
3	D	1345	ARG
5	F	130	ASP
5	F	128	VAL
2	C	528	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 27 such sidechains are listed below:

Mol	Chain	Res	Type
2	C	1013	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	C	1136	GLN
3	D	805	GLN
2	C	1116	HIS
2	C	1157	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 ⓘ

Of 3 ligands modelled in this entry, 3 are 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-30604. These allow visual inspection of the internal detail of the map and identification of artifacts.

No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1   Orthogonal projections ⓘ

6.1.1   Primary map

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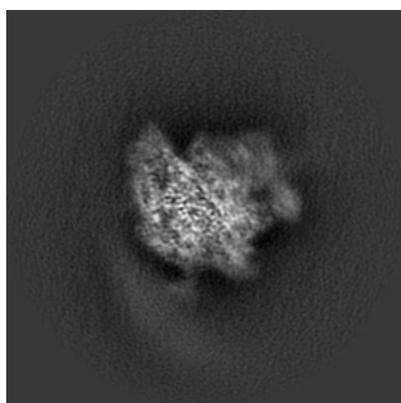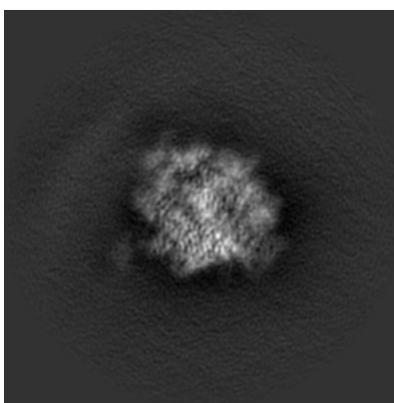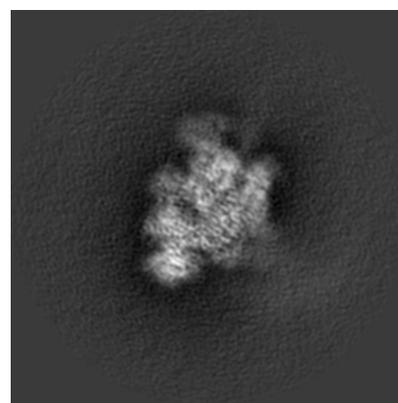
The images above show the map projected in three orthogonal directions.

6.2   Central slices ⓘ
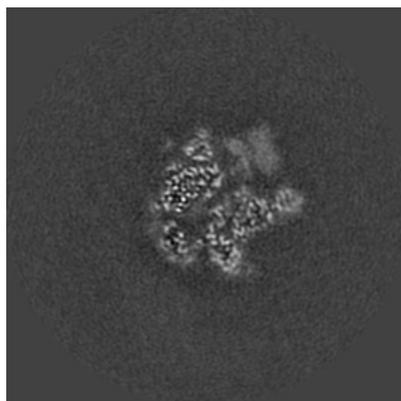
6.2.1   Primary map
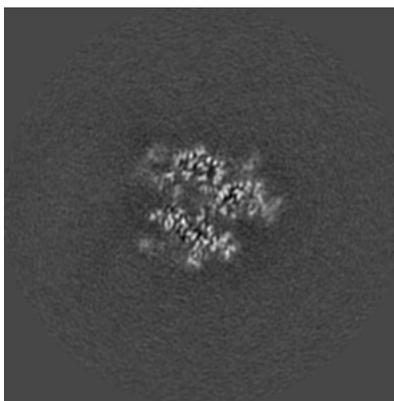
X Index: 128   Y Index: 128   Z Index: 128
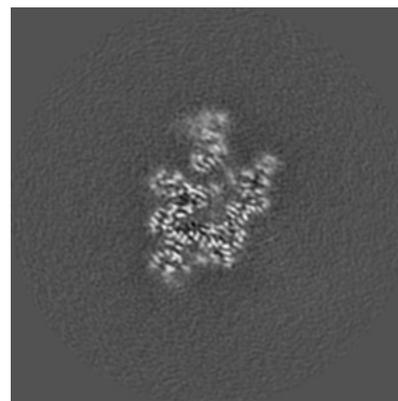
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices ⓘ
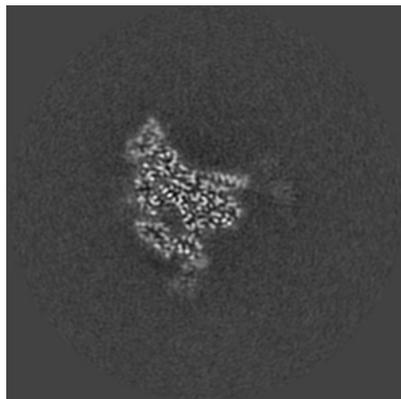
6.3.1 Primary map
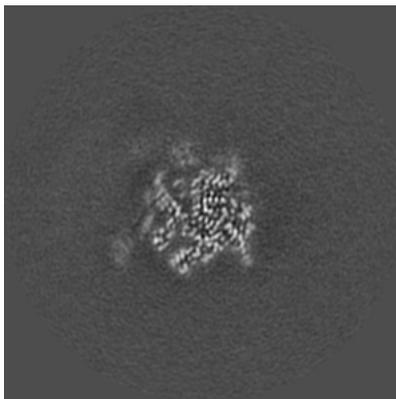
| X Index: 105 | Y Index: 112 | Z Index: 129 |
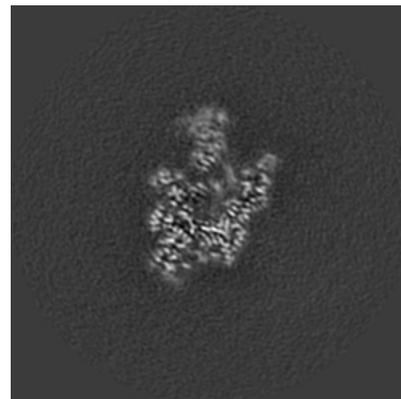
The images above show the largest variance slices of the map in three orthogonal directions.

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

6.4.1 Primary map

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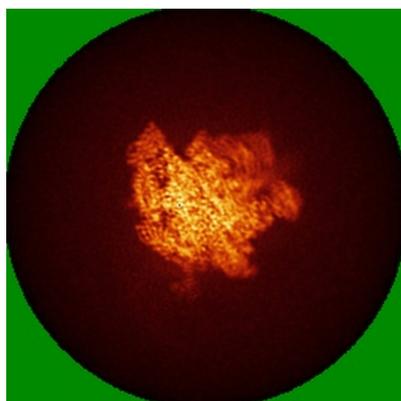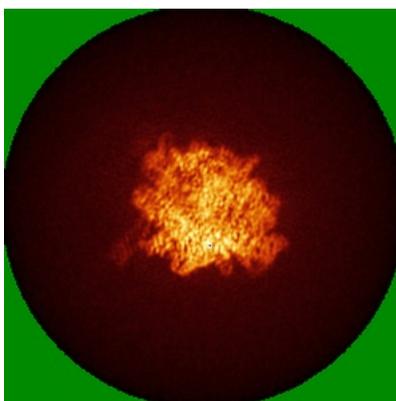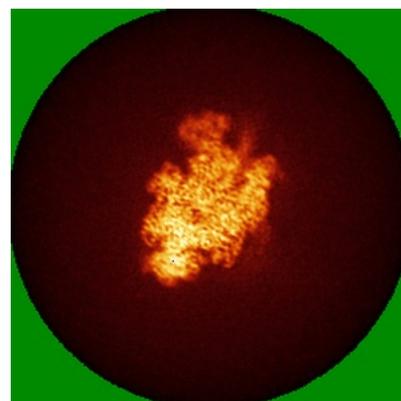
The images above show the map standard deviation projections with false color in three orthogonal directions. Minimum values are shown in green, max in blue, and dark to light orange shades represent small to large values respectively.

6.5 Orthogonal surface views (i)

6.5.1 Primary map

X   Y   Z
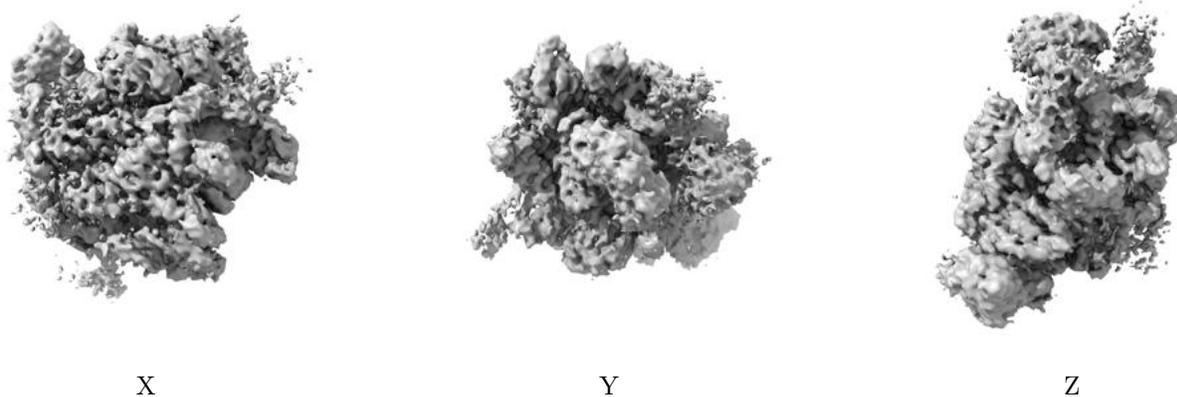
The images above show the 3D surface view of the map at the recommended contour level 0.01. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

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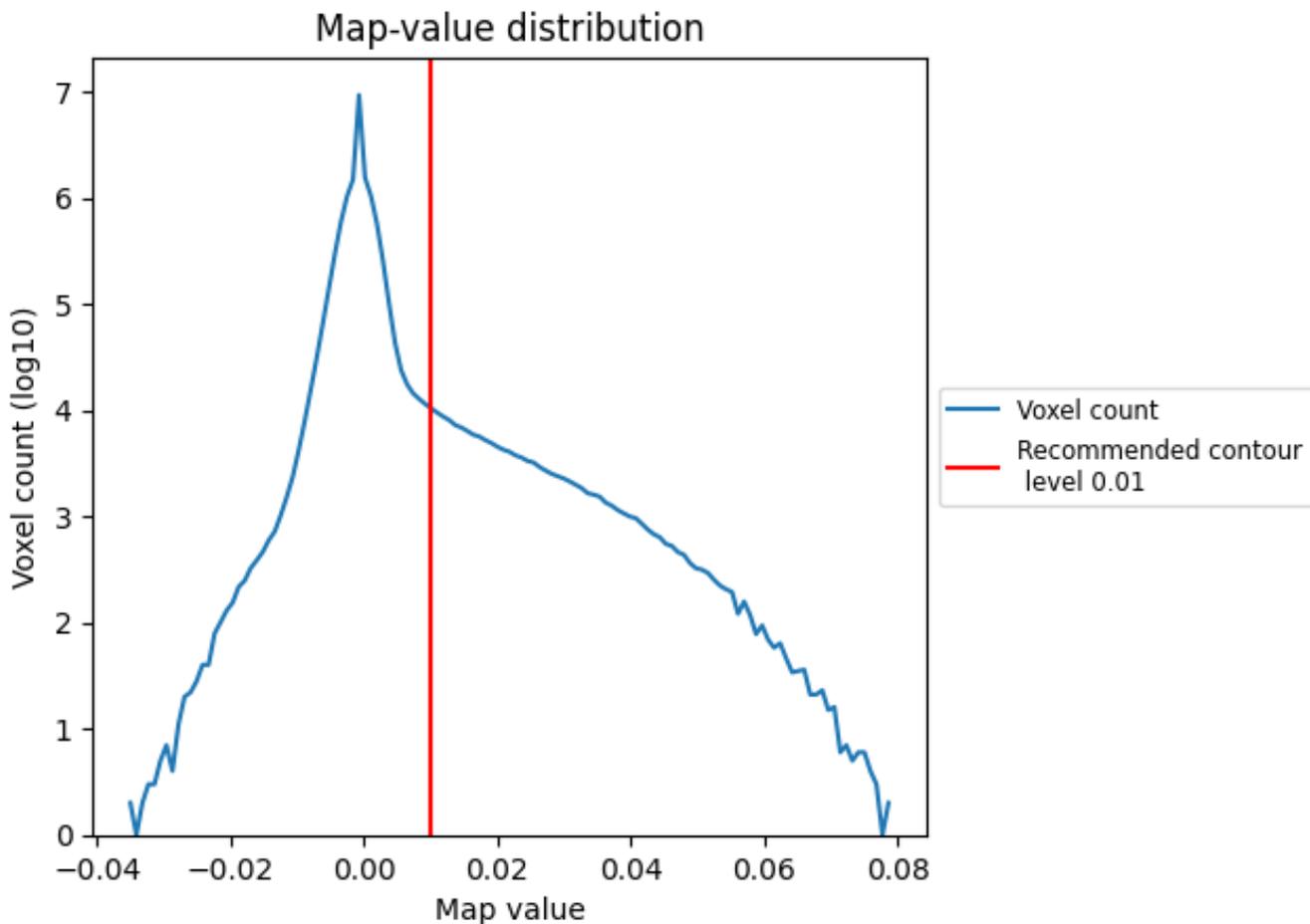
The map-value distribution is plotted in 128 intervals along the x-axis. The y-axis is logarithmic. A spike in this graph at zero usually indicates that the volume has been masked.

7.2   Volume estimate ⓘ
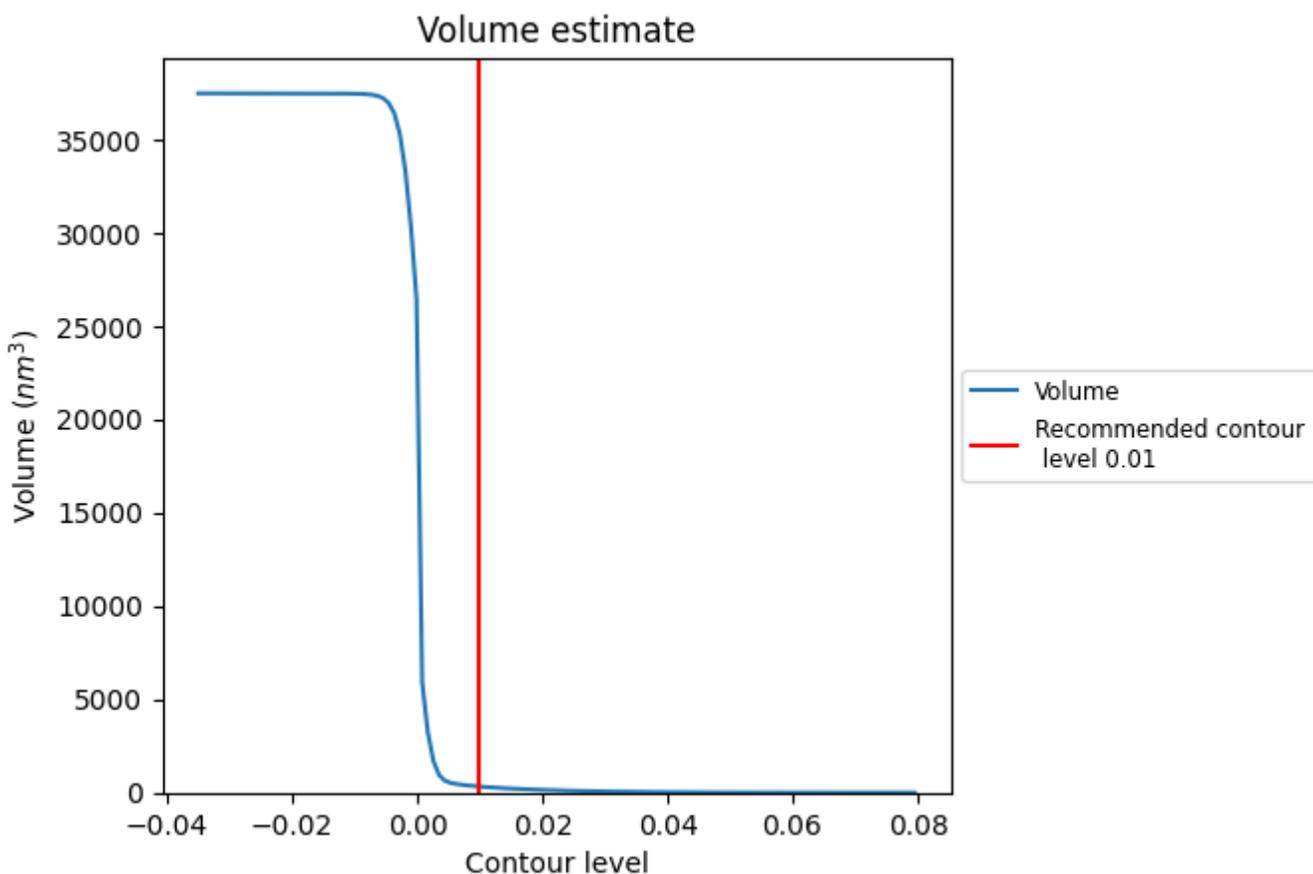
Volume estimate

The volume at the recommended contour level is 325 nm^3; this corresponds to an approximate mass of 294 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 ⓘ
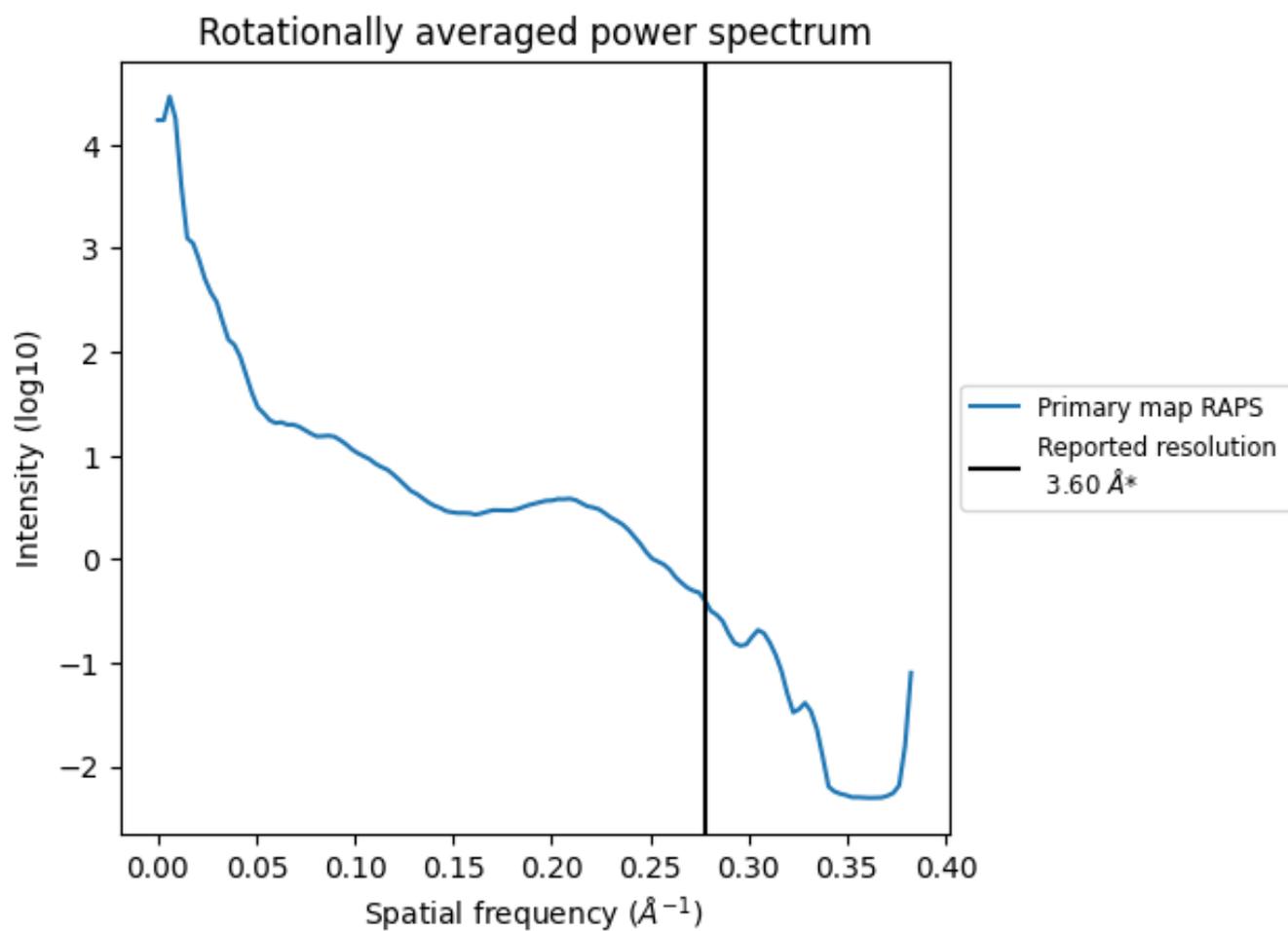
*Reported resolution corresponds to spatial frequency of 0.278 Å^{-1}

8 Fourier-Shell correlation (i)

This section was not generated. No FSC curve or half-maps provided.

9 Map-model fit ⓘ

This section contains information regarding the fit between EMDB map EMD-30604 and PDB model 7D7C. Per-residue inclusion information can be found in section 3 on page 5.
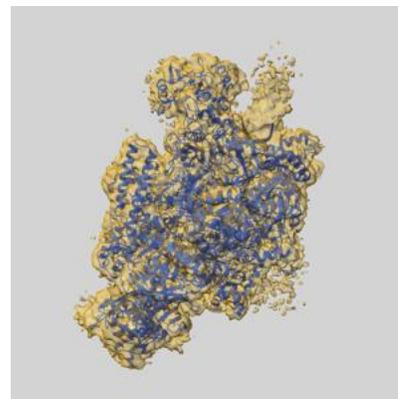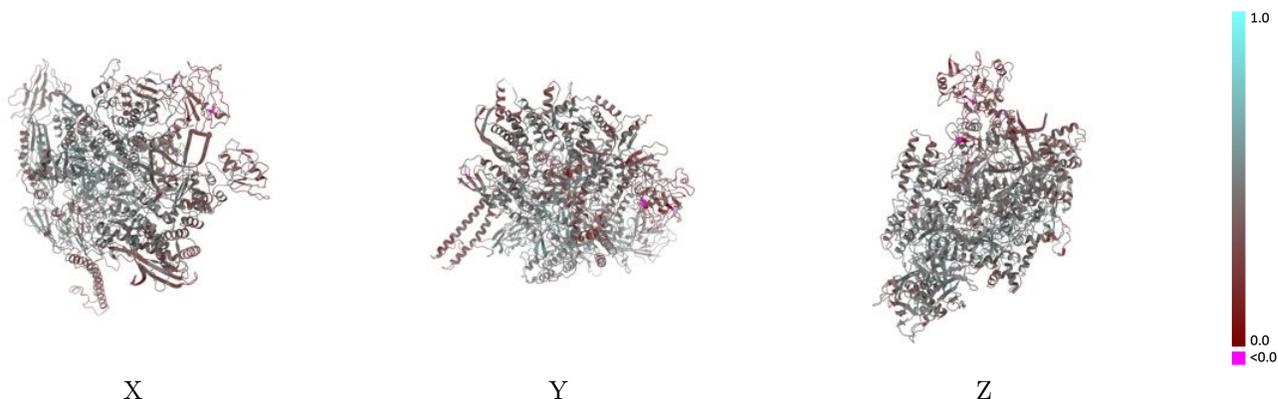
## 9.1 Map-model overlay ⓘ
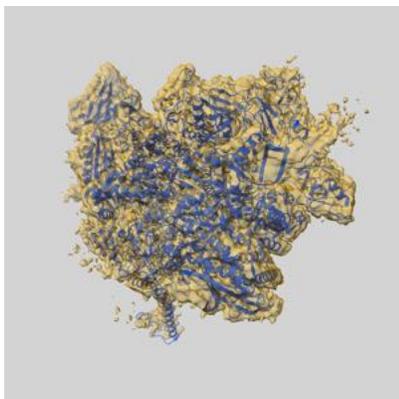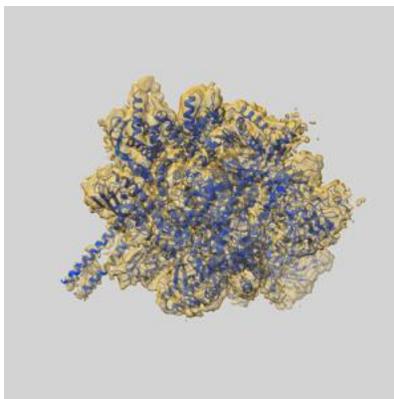
| X | Y | Z |

The images above show the 3D surface view of the map at the recommended contour level 0.01 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
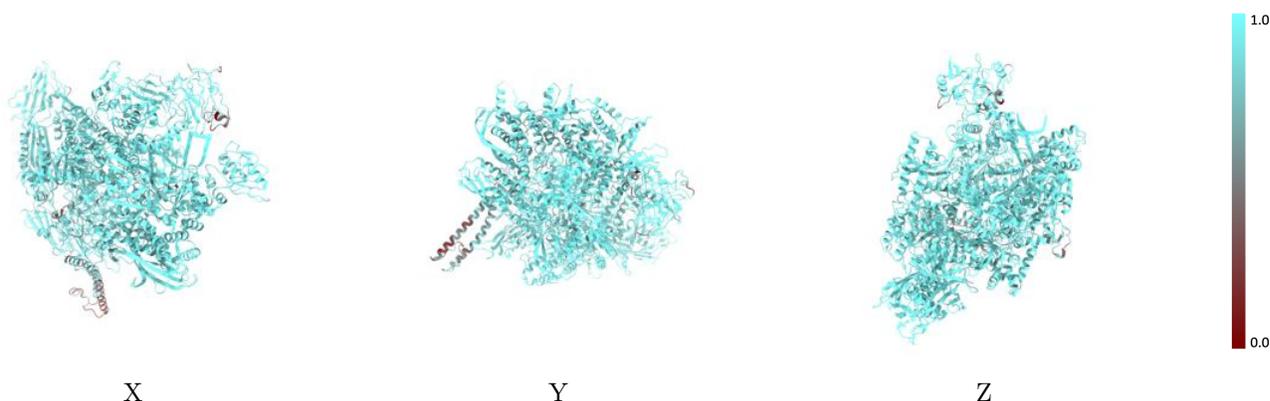
The images above show the model with each residue coloured according to its atom inclusion. This shows to what extent they are inside the map at the recommended contour level (0.01).

9.4   Atom inclusion   ⓘ
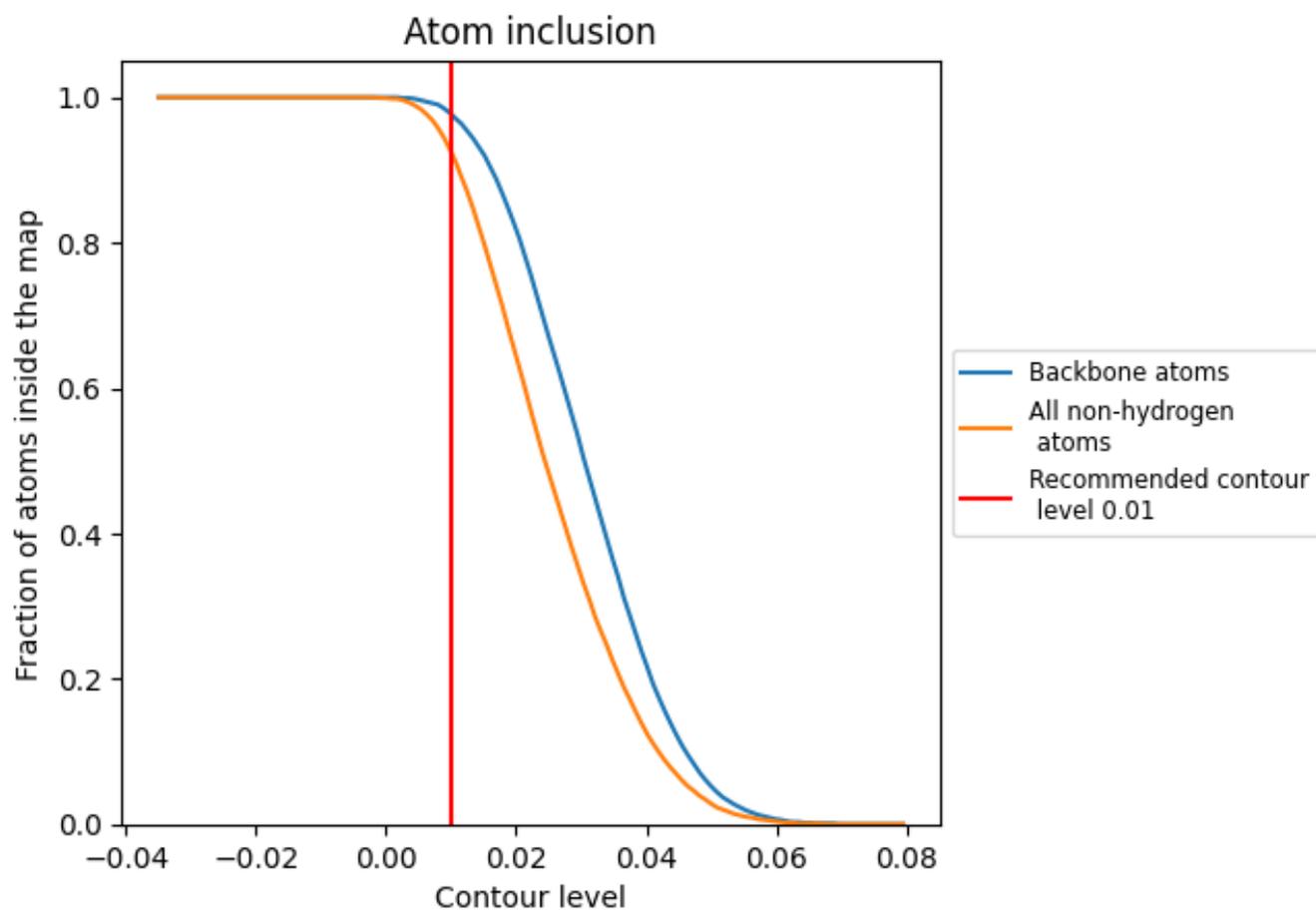
Atom inclusion

At the recommended contour level, 98% of all backbone atoms, 92% 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 (0.01) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.9250	0.4440
A	0.9500	0.4870
B	0.9610	0.4500
C	0.9150	0.4550
D	0.9250	0.4340
F	0.9010	0.4080
N	0.9540	0.3850
T	0.9530	0.3930

1.0

0.0

<0.0