W O R L D W I D E

wwPDB

P R O T E I N   D A T A   B A N K

wwPDB EM Validation Summary Report ⓘ

Mar 3, 2024 – 07:54 AM EST

PDB ID	:	6C6S
EMDB ID	:	EMD-7349
Title	:	CryoEM structure of E.coli RNA polymerase elongation complex bound with RfaH
Authors	:	Kang, J.Y.; Artsimovitch, I.; Landick, R.; Darst, S.A.
Deposited on	:	2018-01-19
Resolution	:	3.70 Å(reported)

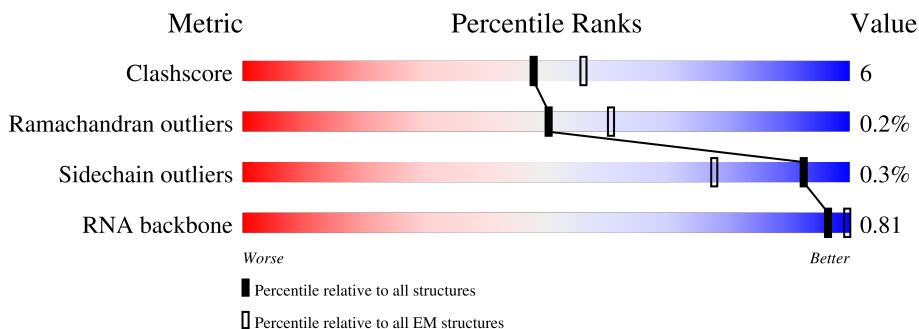The following versions of software and data (see references ⓘ) were used in the production of this report:

EMDB validation analysis	:	0.0.1.dev70
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.13
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 3.70 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.
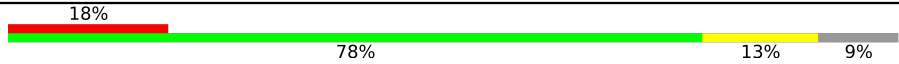
Metric	Percentile Ranks	Value
Clashscore		6
Ramachandran outliers		0.2%
Sidechain outliers		0.3%
RNA backbone		0.81

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
RNA backbone	4643	859

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	29	· 69% 31%
2	B	29	· 79% 21%
3	R	20	25% 20% 5% 50%
4	G	239	· 77% 15% 8%
4	H	239	· 70% 21% 8%
5	I	1342	9% 84% 16%
6	J	1407	9% 77% 18% 5%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
7	K	91	18% ▬▬▬ 78% ▬▬▬▬▬▬ 13% ▬ 9% ▬
8	D	162	33% ▬▬▬▬ 74% ▬▬▬▬▬▬ 25% ▬ •

2 Entry composition ⓘ

There are 10 unique types of molecules in this entry. The entry contains 27691 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 DNA chain called DNA (29-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	29	Total 598	C 282	N 117	O 171	P 28	0	0

- Molecule 2 is a DNA chain called DNA (29-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	29	Total 585	C 279	N 102	O 176	P 28	0	0

- Molecule 3 is a RNA chain called RNA (5'-R(*GP*CP*AP*UP*UP*CP*AP*AP*AP*GP* CP*CP*GP*AP*GP*AP*GP*GP*UP*A)-3').

Mol	Chain	Residues	Atoms					AltConf	Trace
3	R	10	Total 218	C 97	N 43	O 68	P 10	0	0

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

Mol	Chain	Residues	Atoms					AltConf	Trace
4	G	221	Total 1694	C 1061	N 296	O 331	S 6	0	0
4	H	219	Total 1687	C 1053	N 298	O 330	S 6	0	0

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
G	235	GLU	-	expression tag	UNP P0A7Z4
G	236	VAL	-	expression tag	UNP P0A7Z4
G	237	LEU	-	expression tag	UNP P0A7Z4
G	238	PHE	-	expression tag	UNP P0A7Z4
G	239	GLN	-	expression tag	UNP P0A7Z4
H	235	GLU	-	expression tag	UNP P0A7Z4

Continued on next page...

W O R L D W I D E
wwPDB
PROTEIN DATA BANK

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
H	236	VAL	-	expression tag	UNP P0A7Z4
H	237	LEU	-	expression tag	UNP P0A7Z4
H	238	PHE	-	expression tag	UNP P0A7Z4
H	239	GLN	-	expression tag	UNP P0A7Z4

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

Mol	Chain	Residues	Atoms					AltConf	Trace
5	I	1341	Total 10577	C 6636	N 1842	O 2056	S 43	0	0

- Molecule 6 is a protein called DNA-directed RNA polymerase subunit beta'.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	J	1335	Total 10388	C 6526	N 1854	O 1958	S 50	0	0

- Molecule 7 is a protein called DNA-directed RNA polymerase subunit omega.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	K	83	Total 655	C 399	N 123	O 132	S 1	0	0

- Molecule 8 is a protein called Transcription antitermination protein RfaH.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	D	161	Total 1286	C 828	N 222	O 232	S 4	0	0

- Molecule 9 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		AltConf
9	R	1	Total 1	Mg 1	0

- Molecule 10 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		AltConf
10	J	2	Total 2	Zn 2	0

3   Residue-property plots (i)

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 (29-MER)

Chain A:
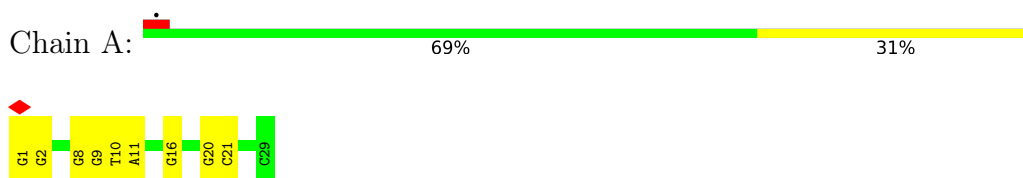
- Molecule 2: DNA (29-MER)

Chain B:
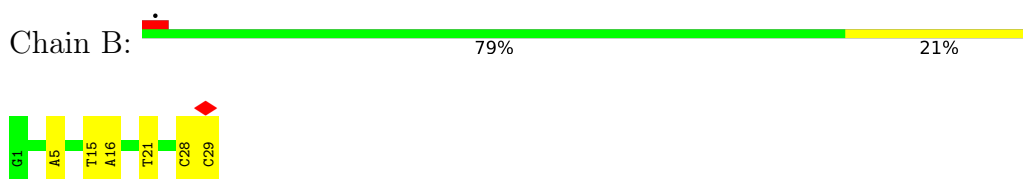
- Molecule 3: RNA (5'-R(*GP*CP*AP*UP*UP*CP*AP*AP*AP*GP*CP*CP*GP*AP*GP*AP *GP*GP*UP*A)-3')
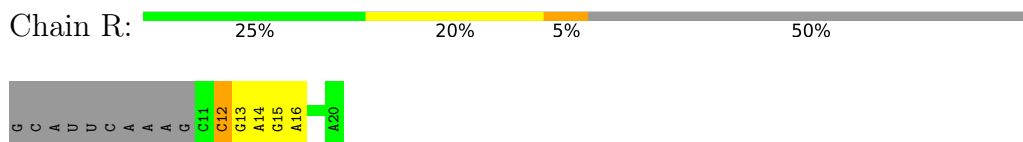
Chain R:

- Molecule 4: DNA-directed RNA polymerase subunit alpha

Chain G:

- Molecule 4: DNA-directed RNA polymerase subunit alpha

Chain H:
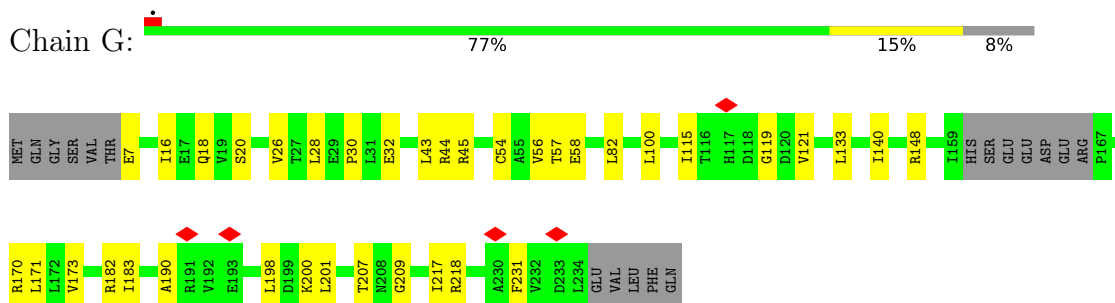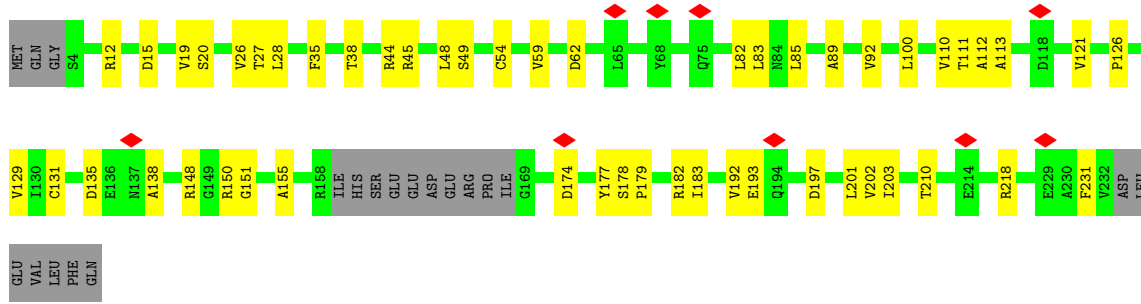
• Molecule 5: DNA-directed RNA polymerase subunit beta
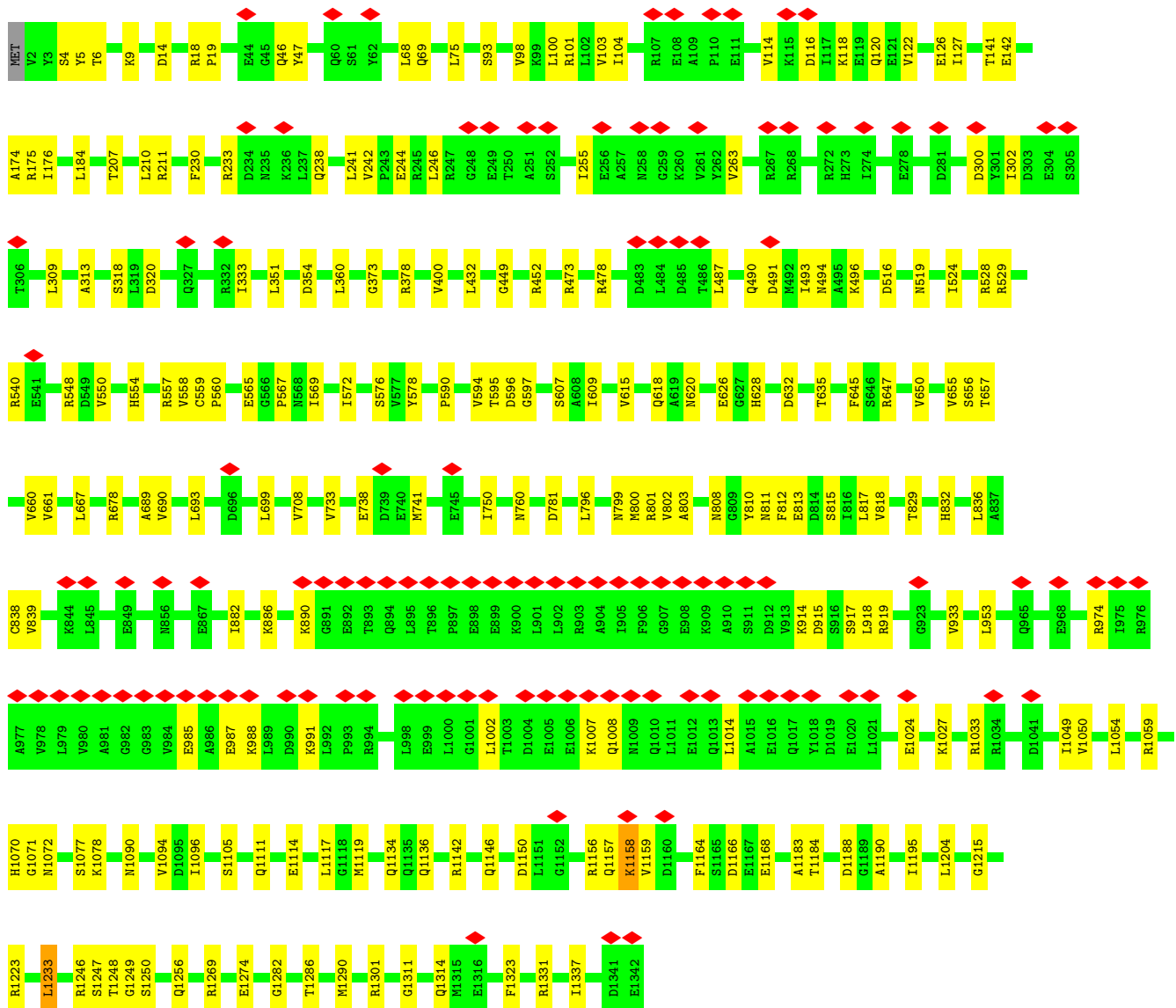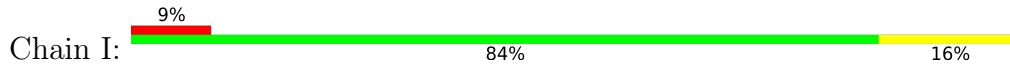
• Molecule 6: DNA-directed RNA polymerase subunit beta'

• Molecule 7: DNA-directed RNA polymerase subunit omega

• Molecule 8: Transcription antitermination protein RfaH

Chain D:

33%

74%

25%

MET Q2 C9 Q13 L14 Q18 L29 T34 L35 E36 K37 I38 V39 R40 G41 T44 S47 L50 F51 P52 L55 F56 H65 T66 T67 N70 S76 A83 V88 Y99 K100 P101 K102 D103 I104 V105 D106 P107 A108 T109 P110 Y111 P112 G113 D114

K115 V116 I117 I118 T119 E120 G121 A122 F123 E124 G125 F126 Q127 A128 I129 F130 T131 E132 P133 D134 G135 E136 A137 R138 S139 M140 L141 L142 L143 N144 L145 I146 N147 K148 E149 I150 K151 H152 S153 V154 K155 N156 T157 E158 F159 R160 K161 L162

4 Experimental information ⓘ

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	107500	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 $(e^-/\text{Å}^2)$	71.0	Depositor
Minimum defocus (nm)	500	Depositor
Maximum defocus (nm)	3000	Depositor
Magnification	38462	Depositor
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor
Maximum map value	0.121	Depositor
Minimum map value	-0.072	Depositor
Average map value	0.000	Depositor
Map value standard deviation	0.005	Depositor
Recommended contour level	0.02	Depositor
Map size (Å)	332.8, 332.8, 332.8	wwPDB
Map dimensions	256, 256, 256	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.3, 1.3, 1.3	Depositor

5   Model quality ⓘ

5.1   Standard geometry ⓘ

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	0.96	0/672	0.98	2/1037 (0.2%)
2	B	1.08	0/653	1.02	0/1005
3	R	0.94	0/244	0.92	0/379
4	G	0.41	0/1714	0.61	0/2323
4	H	0.37	0/1706	0.62	0/2311
5	I	0.46	1/10746 (0.0%)	0.61	4/14499 (0.0%)
6	J	0.44	0/10545	0.62	2/14236 (0.0%)
7	K	0.37	0/657	0.62	0/886
8	D	0.39	0/1317	0.62	0/1786
All	All	0.49	1/28254 (0.0%)	0.64	8/38462 (0.0%)

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
5	I	0	4
6	J	0	4
7	K	0	1
All	All	0	9

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	I	550	VAL	C-N	-5.64	1.21	1.34

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	I	550	VAL	C-N-CA	6.87	138.86	121.70
6	J	903	LEU	C-N-CA	5.74	136.05	121.70
5	I	1233	LEU	CA-CB-CG	5.49	127.93	115.30
1	A	8	DG	P-O3'-C3'	5.32	126.08	119.70
1	A	11	DA	P-O3'-C3'	5.29	126.04	119.70

There are no chirality outliers.

5 of 9 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
5	I	1157	GLN	Peptide
5	I	1158	LYS	Peptide
5	I	594	VAL	Peptide
5	I	595	THR	Peptide
6	J	313	GLY	Peptide

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	598	0	325	6	0
2	B	585	0	328	8	0
3	R	218	0	109	3	0
4	G	1694	0	1722	24	0
4	H	1687	0	1722	36	0
5	I	10577	0	10590	124	0
6	J	10388	0	10611	147	0
7	K	655	0	663	8	0
8	D	1286	0	1303	28	0
9	R	1	0	0	0	0
10	J	2	0	0	0	0
All	All	27691	0	27373	345	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 6.

The worst 5 of 345 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)
6:J:416:ILE:HG13	6:J:441:LEU:HD11	1.72	0.71
4:H:48:LEU:HB2	4:H:183:ILE:HD11	1.71	0.71
5:I:69:GLN:HE21	5:I:101:ARG:HD2	1.58	0.68
8:D:110:PRO:HG3	8:D:161:LYS:HB2	1.77	0.66
2:B:16:DA:H2"	6:J:352:ARG:HH12	1.61	0.65

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	
4	G	217/239 (91%)	201 (93%)	15 (7%)	1 (0%)	29	66
4	H	215/239 (90%)	201 (94%)	14 (6%)	0	100	100
5	I	1339/1342 (100%)	1229 (92%)	107 (8%)	3 (0%)	47	78
6	J	1329/1407 (94%)	1215 (91%)	113 (8%)	1 (0%)	51	83
7	K	81/91 (89%)	75 (93%)	6 (7%)	0	100	100
8	D	159/162 (98%)	135 (85%)	23 (14%)	1 (1%)	25	62
All	All	3340/3480 (96%)	3056 (92%)	278 (8%)	6 (0%)	50	78

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
5	I	596	ASP
5	I	1158	LYS
6	J	805	GLN
4	G	119	GLY
8	D	52	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	
4	G	186/206 (90%)	186 (100%)	0	100	100
4	H	187/206 (91%)	186 (100%)	1 (0%)	88	94
5	I	1156/1157 (100%)	1154 (100%)	2 (0%)	93	97
6	J	1120/1168 (96%)	1114 (100%)	6 (0%)	88	94
7	K	70/75 (93%)	70 (100%)	0	100	100
8	D	141/142 (99%)	141 (100%)	0	100	100
All	All	2860/2954 (97%)	2851 (100%)	9 (0%)	92	96

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

Mol	Chain	Res	Type
6	J	1268	ASN
6	J	1373	ARG
6	J	53	ARG
6	J	60	ARG
6	J	709	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 19 such sidechains are listed below:

Mol	Chain	Res	Type
6	J	777	HIS
6	J	1326	GLN
8	D	20	HIS
6	J	1268	ASN
5	I	1313	HIS

5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
3	R	9/20 (45%)	1 (11%)	0

All (1) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
3	R	12	C

There are no RNA pucker outliers to report.

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-7349. 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 |
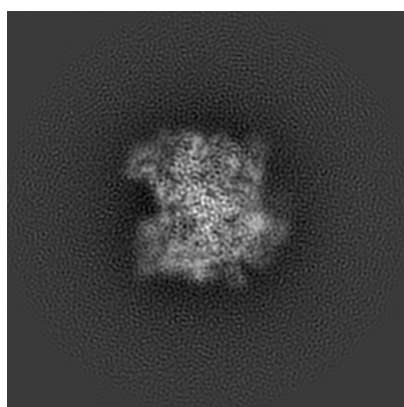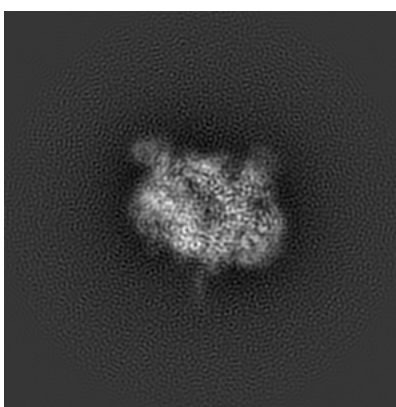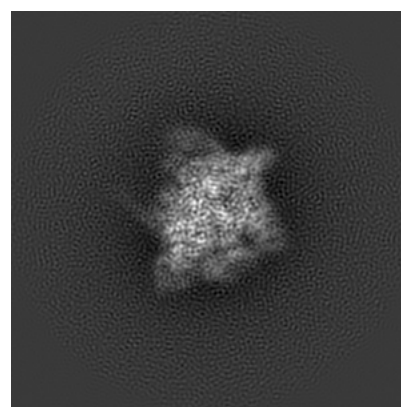
The images above show the map projected in three orthogonal directions.

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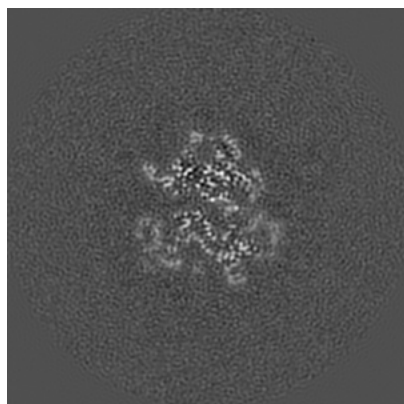
6.2.1 Primary map
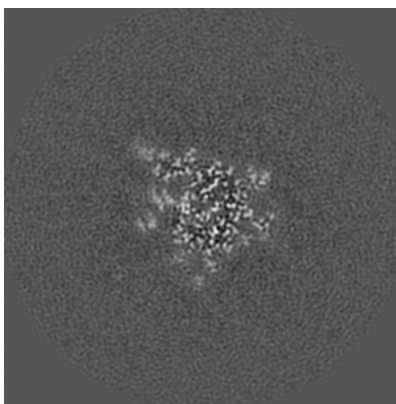
| X Index: 128 | Y Index: 128 | Z Index: 128 |
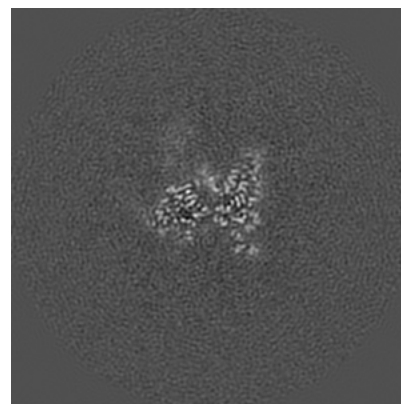
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

6.3.1 Primary map
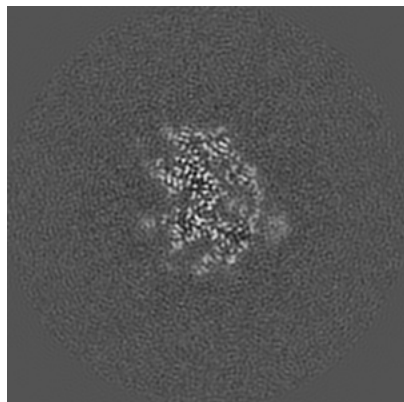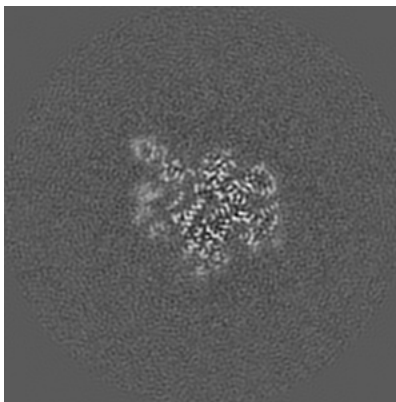
| X Index: 120 | Y Index: 122 | Z Index: 136 |
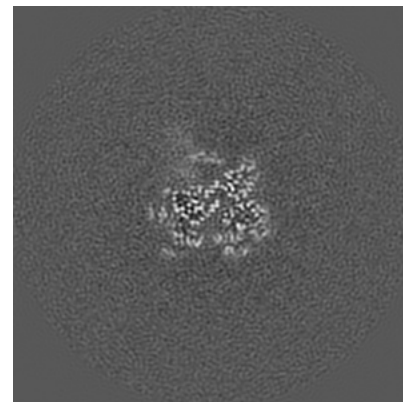
The images above show the largest variance slices of the map in three orthogonal directions.

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

6.4.1 Primary map

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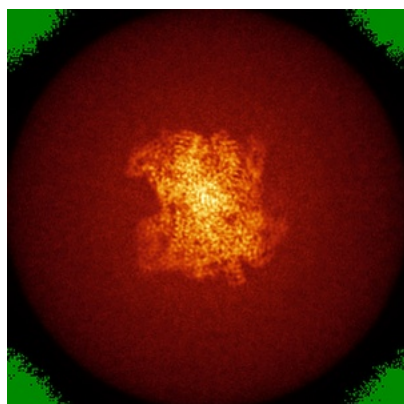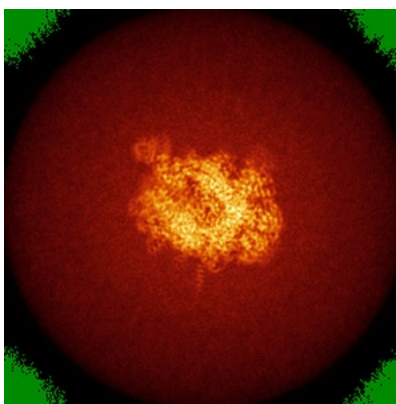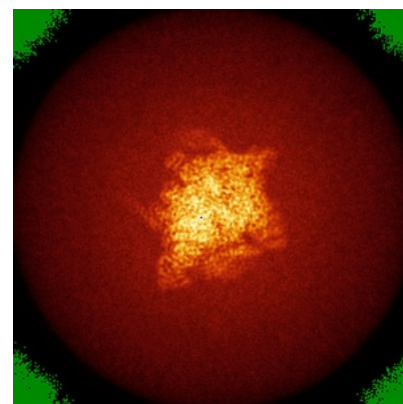
The images above show the map standard deviation projections with false color in three orthogonal directions. Minimum values are shown in green, max in blue, and dark to light orange shades represent small to large values respectively.

6.5　Orthogonal surface views ⓘ

6.5.1　Primary map

| X | Y | Z |
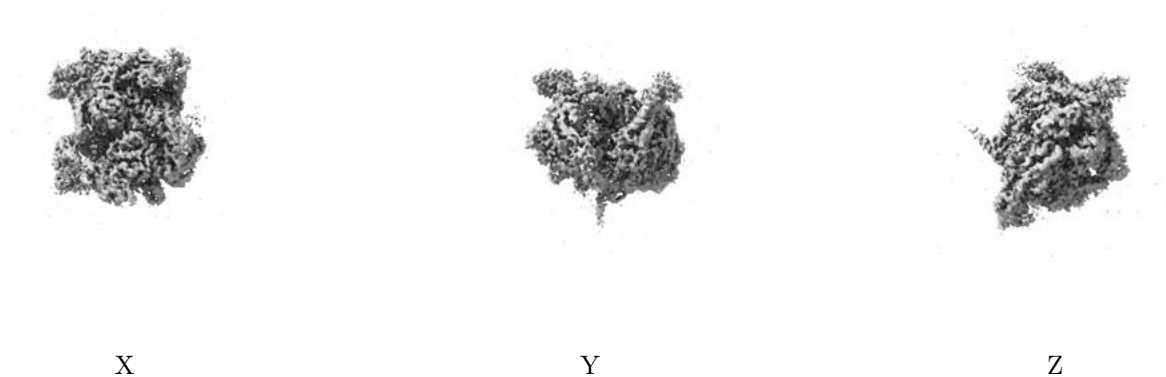
The images above show the 3D surface view of the map at the recommended contour level 0.02. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

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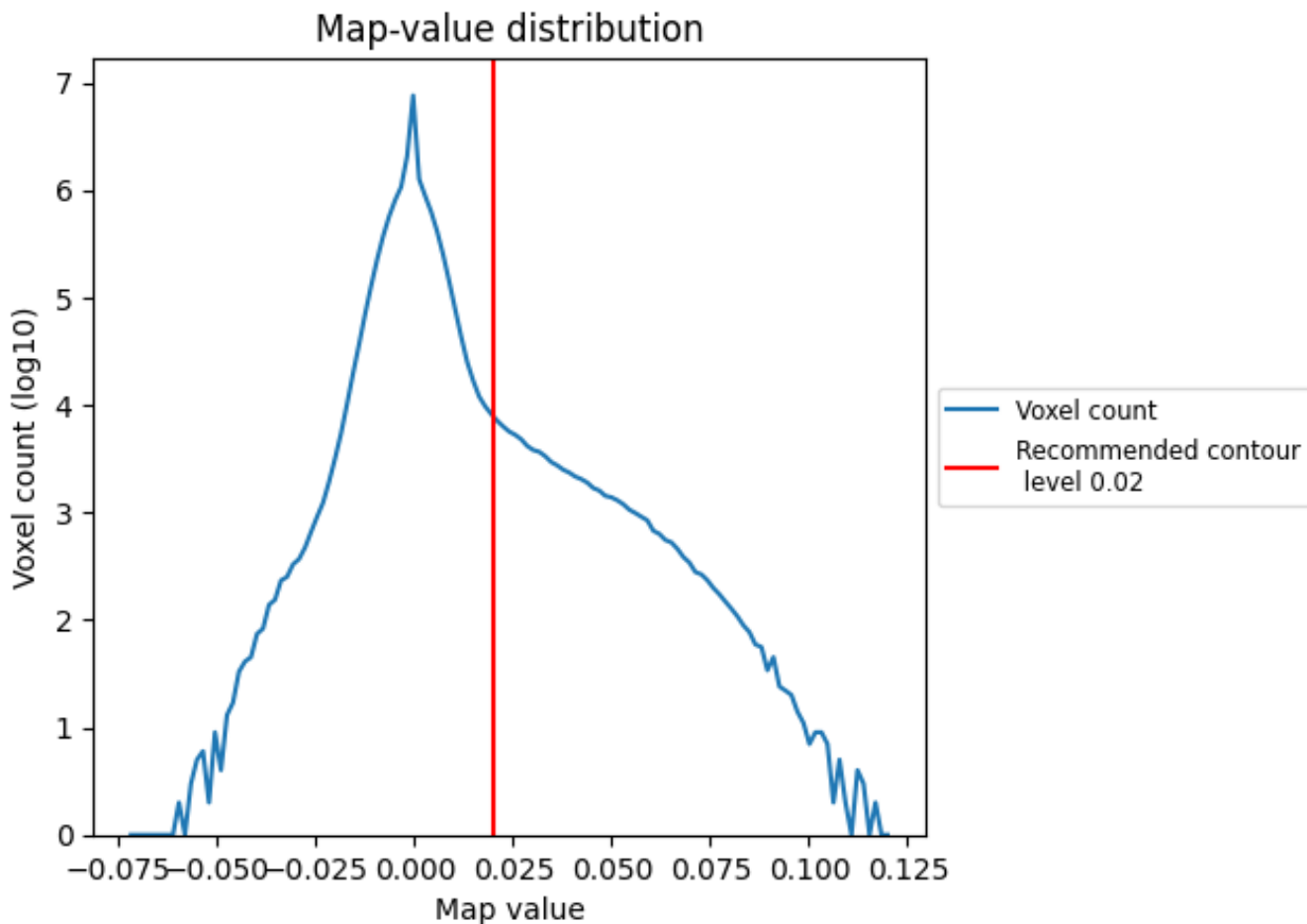
The map-value distribution is plotted in 128 intervals along the x-axis. The y-axis is logarithmic. A spike in this graph at zero usually indicates that the volume has been masked.

7.2　Volume estimate (i)
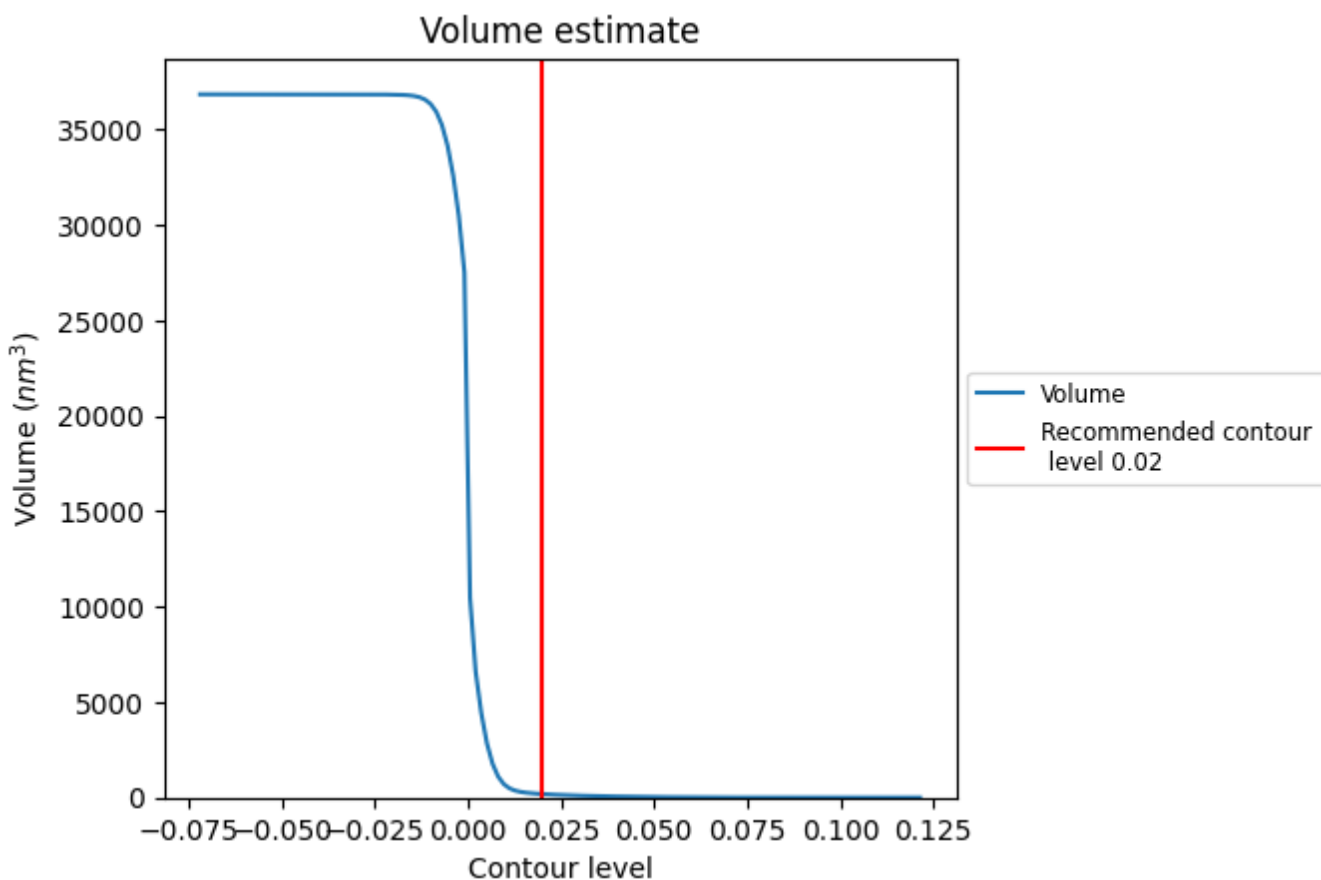
The volume at the recommended contour level is 184 nm^3; this corresponds to an approximate mass of 166 kDa.

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)
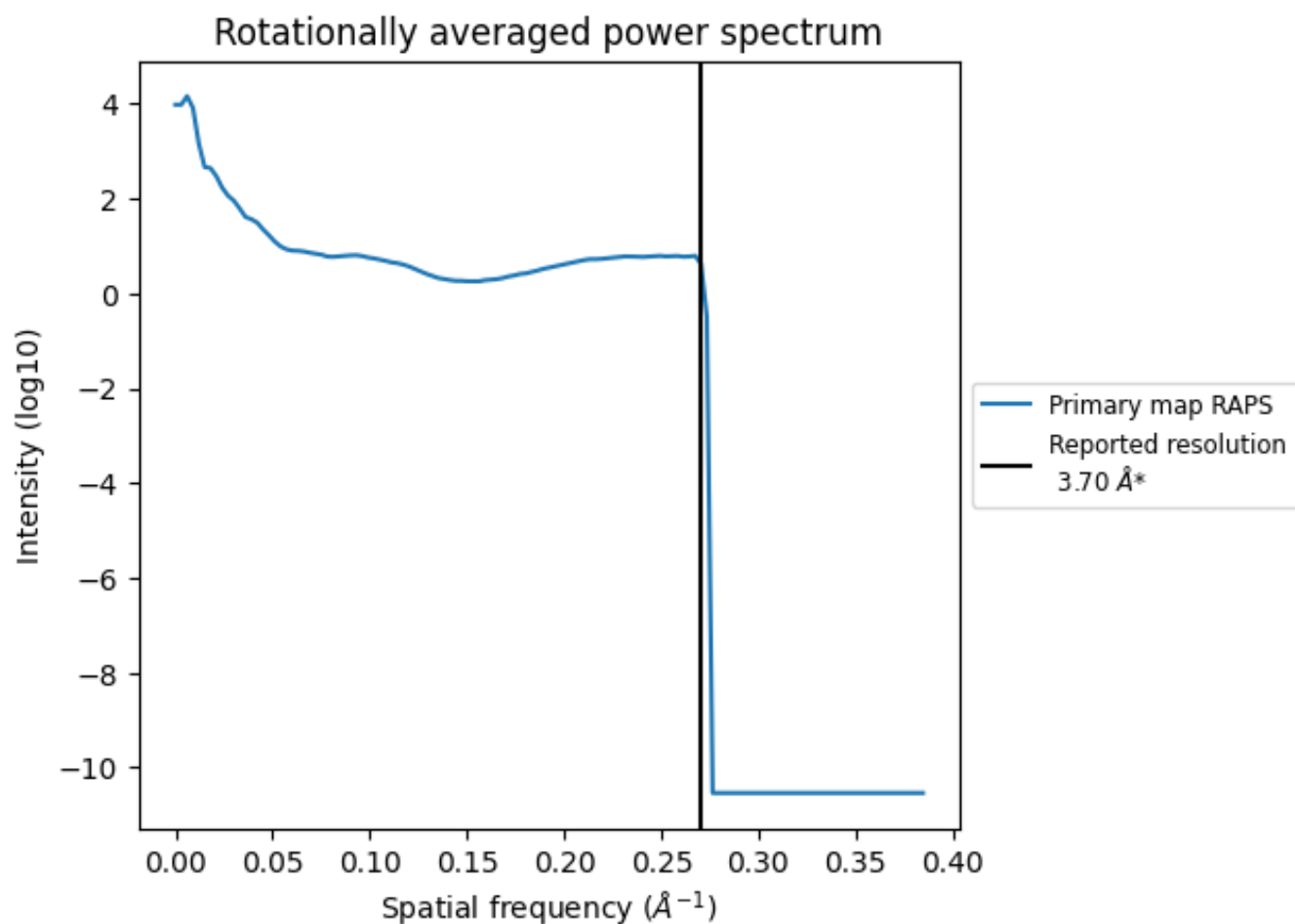
*Reported resolution corresponds to spatial frequency of 0.270 Å^{-1}

8 Fourier-Shell correlation ⓘ

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-7349 and PDB model 6C6S. Per-residue inclusion information can be found in section 3 on page 6.

9.1 Map-model overlay ⓘ

X   Y   Z
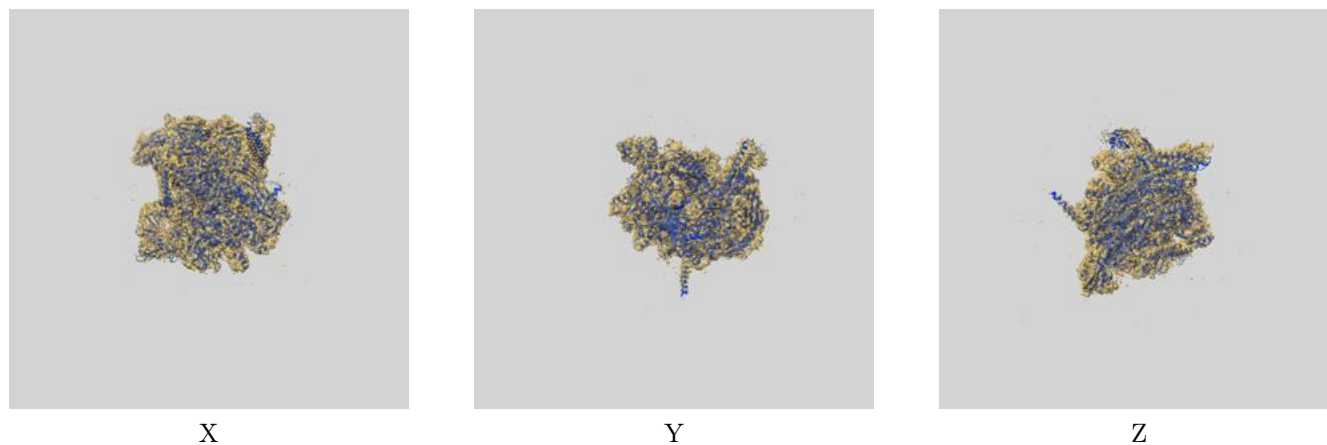
The images above show the 3D surface view of the map at the recommended contour level 0.02 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

9.2   Q-score mapped to coordinate model ⓘ

X   Y   Z
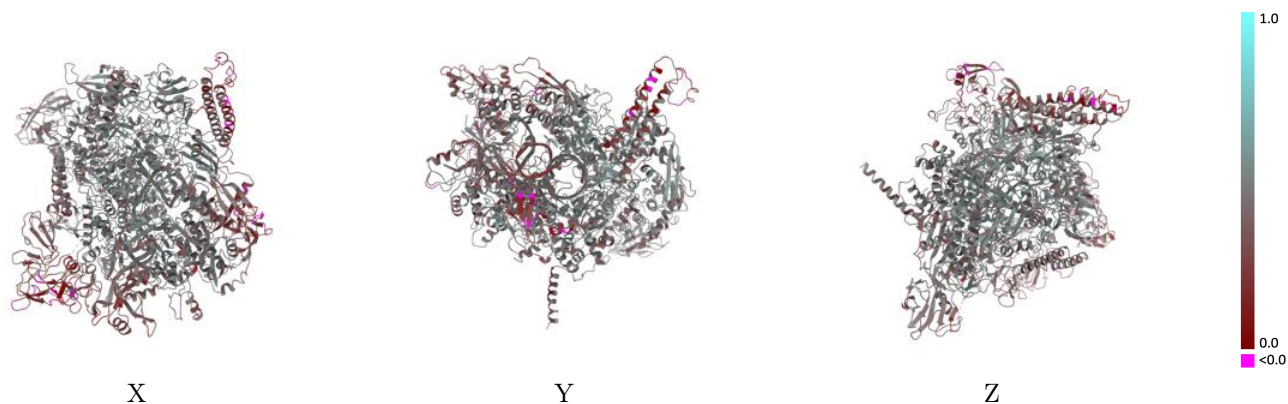
The images above show the model with each residue coloured according its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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
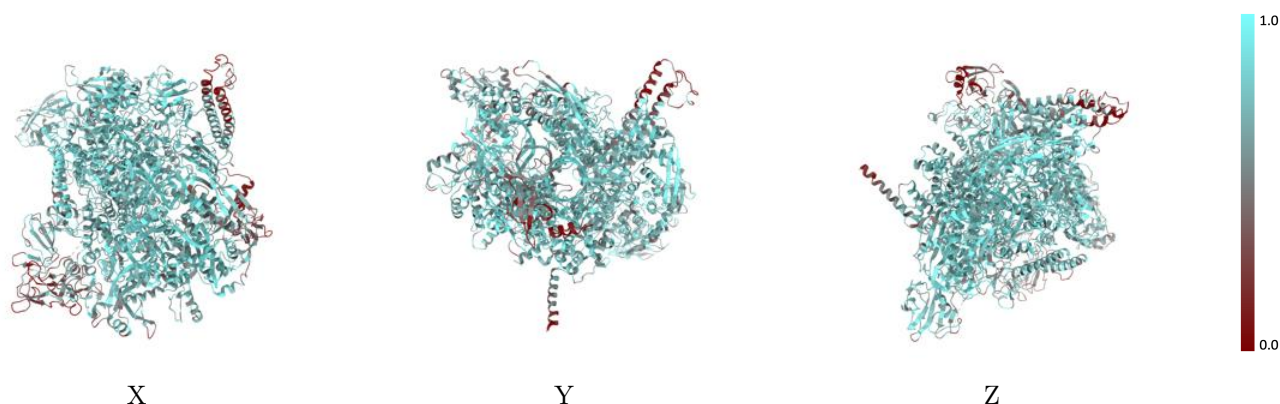
The images above show the model with each residue coloured according to its atom inclusion. This shows to what extent they are inside the map at the recommended contour level (0.02).

9.4 Atom inclusion ⓘ
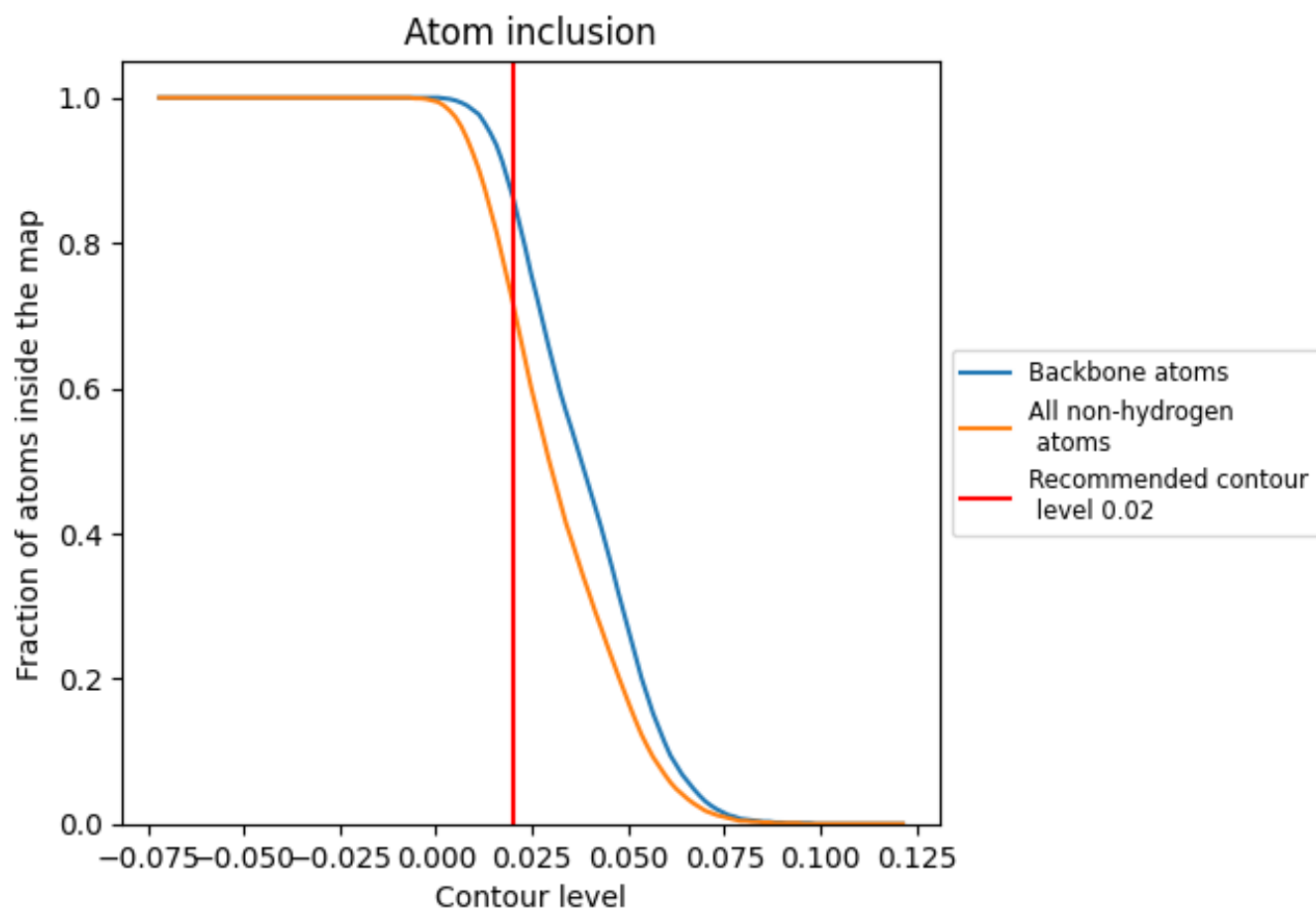
At the recommended contour level, 86% of all backbone atoms, 72% of all non-hydrogen atoms, are inside the map.

9.5 Map-model fit summary ⓘ

The table lists the average atom inclusion at the recommended contour level (0.02) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.7180	0.4260
A	0.8140	0.3870
B	0.8240	0.4030
D	0.5840	0.3520
G	0.7650	0.4660
H	0.7220	0.4230
I	0.7220	0.4340
J	0.7130	0.4240
K	0.6070	0.4210
R	0.9410	0.4700

1.0

0.0
<0.0