wwPDB EM Validation Summary Report (i)

Nov 29, 2022 – 09:34 PM EST

PDB ID	:	8E6X
EMDB ID	:	EMD-27930
Title	:	Escherichia coli Rho-dependent transcription pre-termination complex containing 18 nt long RNA spacer, lambda-tR1 rut RNA, Mg-ADP-BeF3, and NusG; TEC part
Authors	:	Molodtsov, V.; Wang, C.
Deposited on	:	2022-08-23
Resolution	:	4.27 Å(reported)

This is a wwPDB EM Validation Summary Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/EMValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

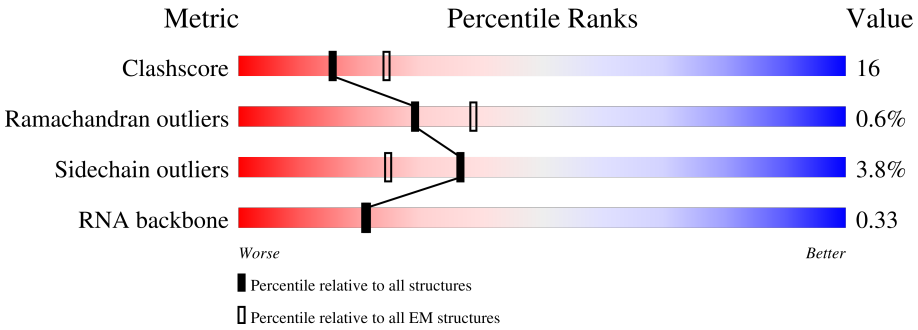
EMDB validation analysis	:	0.0.1.dev43
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.9
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.31.2

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
ELECTRON MICROSCOPY

The reported resolution of this entry is 4.27 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.



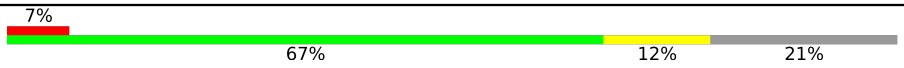
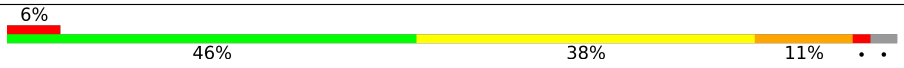
Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826
RNA backbone	4643	859

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	5	60	17% 30% 8% 45%
2	6	60	22% 35% 5% 38%
3	7	35	31% 51% 23% 23%
4	A	1342	81% 18%
5	B	1407	66% 25% 5%
6	C	329	60% 7% 33%
6	D	329	54% 12% 34%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
7	E	91	
8	F	181	

2 Entry composition (i)

There are 10 unique types of molecules in this entry. The entry contains 39532 atoms, of which 11239 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 NT DNA.

Mol	Chain	Residues	Atoms						AltConf	Trace
			Total	C	H	N	O	P		
1	5	33	913	322	237	125	196	33	0	0

- Molecule 2 is a DNA chain called T DNA.

Mol	Chain	Residues	Atoms						AltConf	Trace
			Total	C	H	N	O	P		
2	6	37	1029	357	281	126	228	37	0	0

- Molecule 3 is a RNA chain called RNA with 18 nt long spacer.

Mol	Chain	Residues	Atoms						AltConf	Trace
			Total	C	H	N	O	P		
3	7	27	676	253	109	84	203	27	0	0

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

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	A	1340	10567	6631	1841	2052	43	0	0

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

Mol	Chain	Residues	Atoms						AltConf	Trace
			Total	C	H	N	O	S		
5	B	1335	21000	6526	10612	1854	1958	50	0	0

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

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
6	C	221	1698	1060	299	333	6	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					AltConf	Trace
6	D	218	Total 1677	C 1048	N 297	O 326	S 6	0	0

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

Mol	Chain	Residues	Atoms					AltConf	Trace
7	E	72	Total 577	C 352	N 110	O 114	S 1	0	0

- Molecule 8 is a protein called Transcription termination/antitermination protein NusG.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	F	175	Total 1392	C 881	N 249	O 255	S 7	0	0

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

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

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

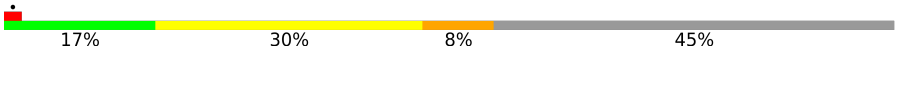
Mol	Chain	Residues	Atoms		AltConf
10	B	2	Total 2	Zn 2	0

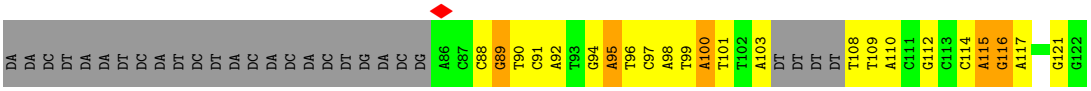
3 Residue-property plots (i)

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: NT DNA

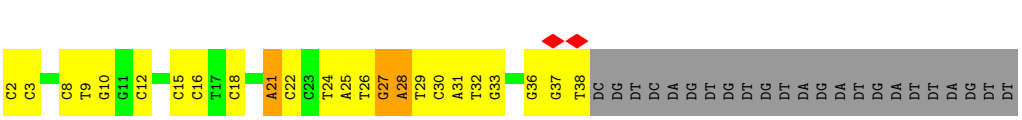
Chain 5:





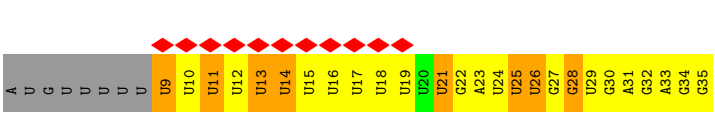
• Molecule 2: T DNA

Chain 6:



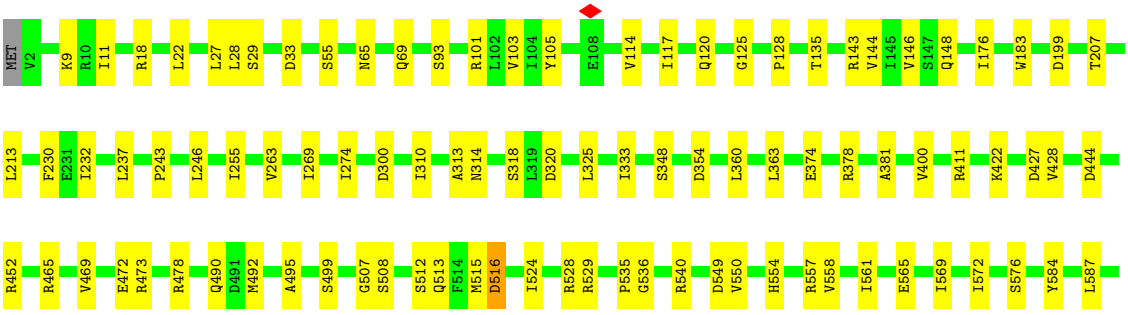
• Molecule 3: RNA with 18 nt long spacer

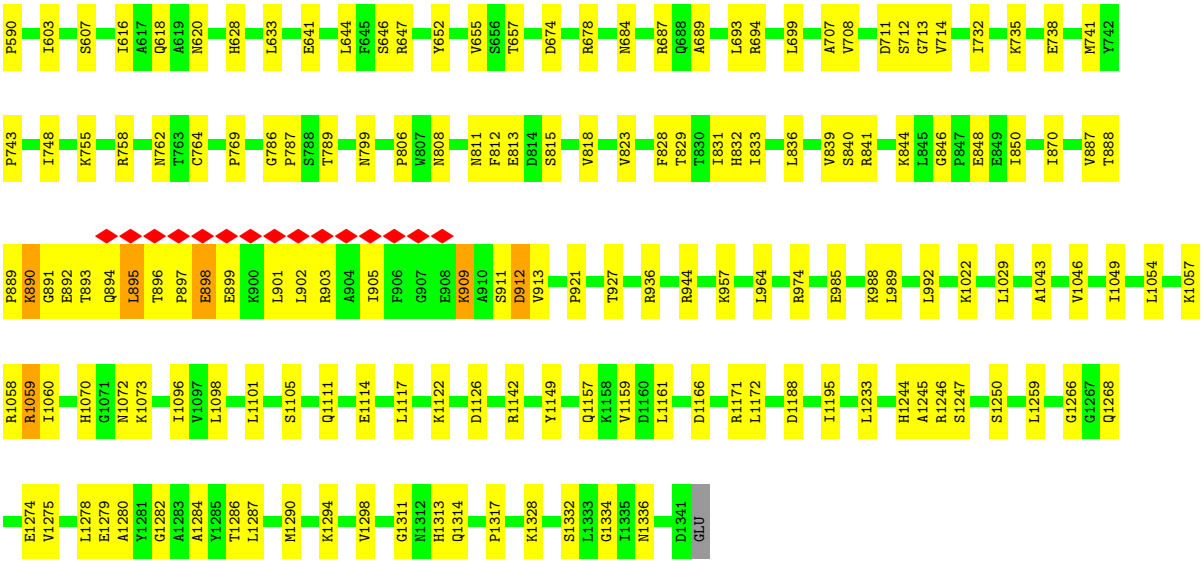
Chain 7:



• Molecule 4: DNA-directed RNA polymerase subunit beta

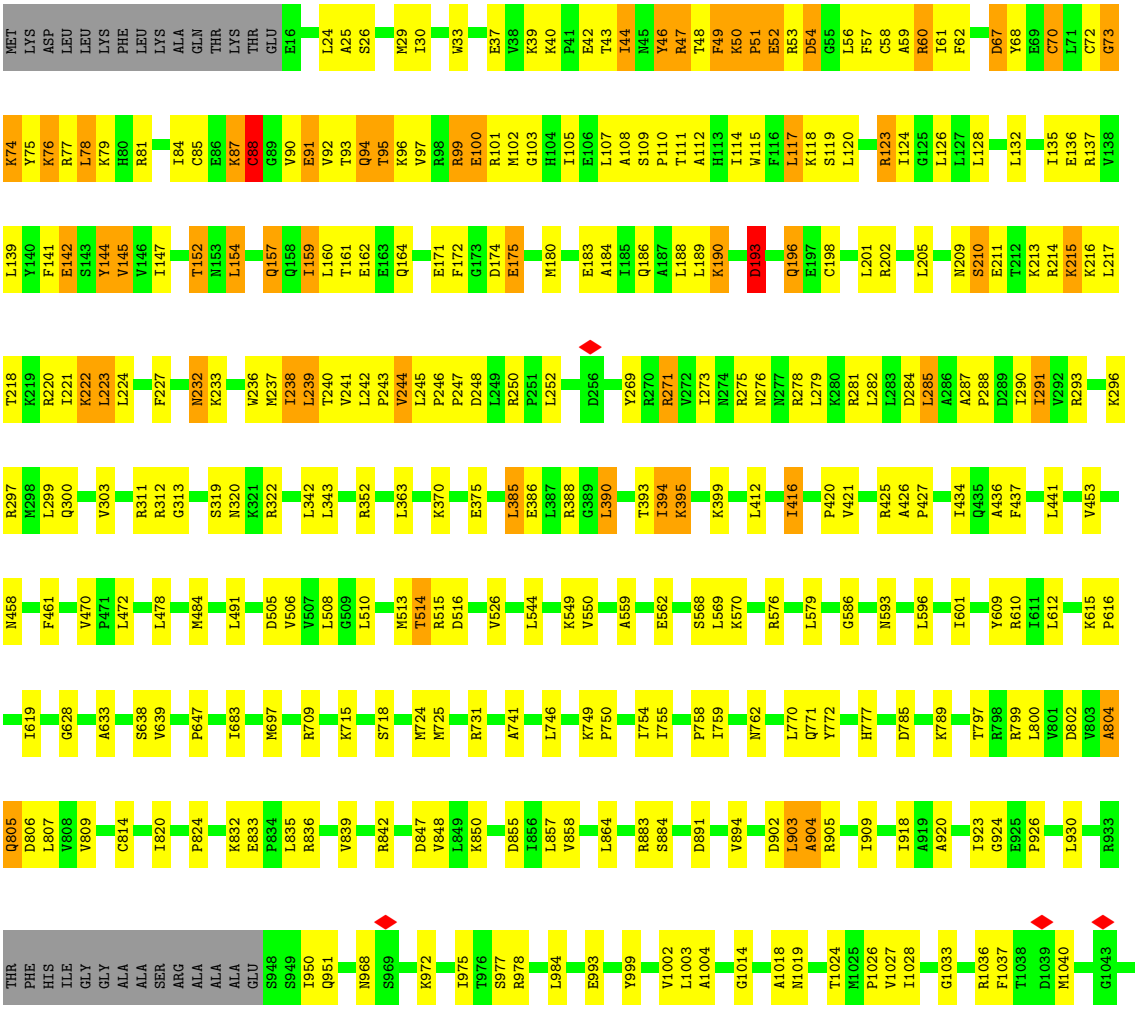
Chain A:

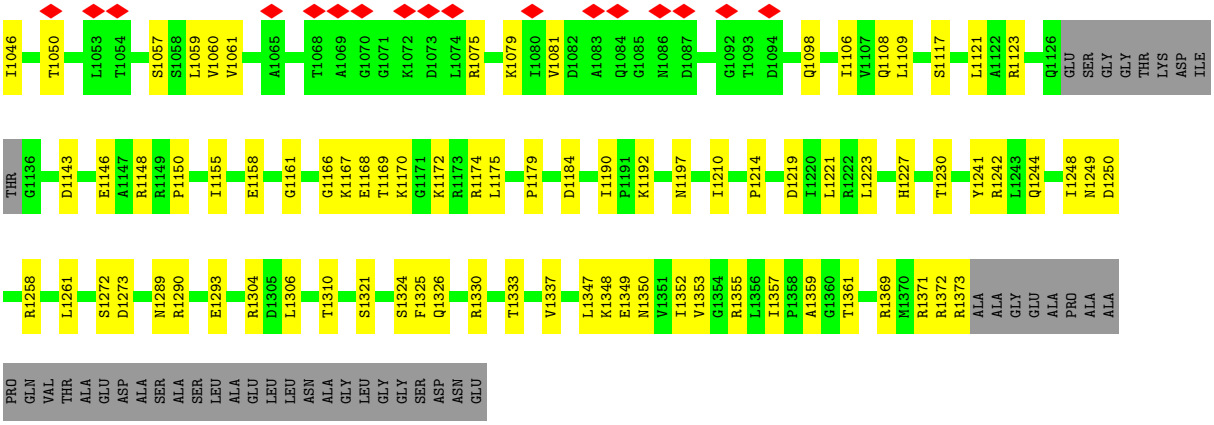




• Molecule 5: DNA-directed RNA polymerase subunit beta'

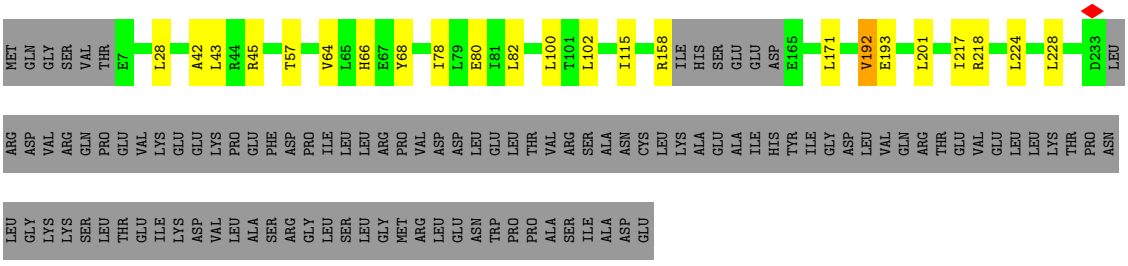
Chain B:





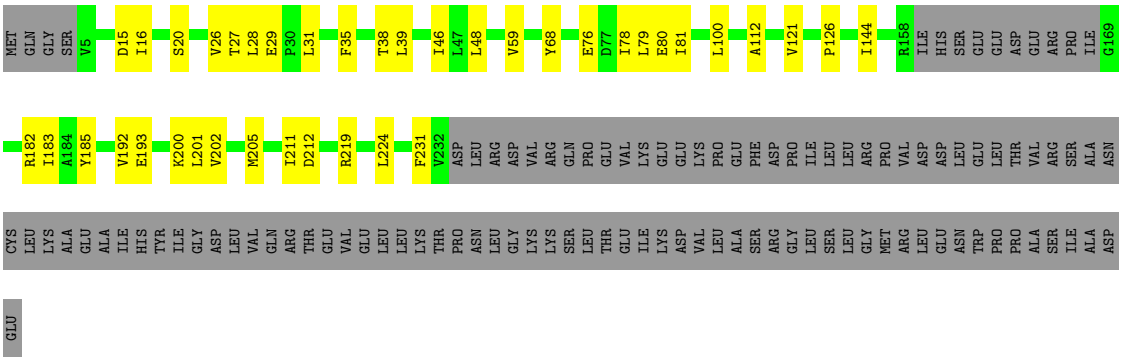
• Molecule 6: DNA-directed RNA polymerase subunit alpha

Chain C:



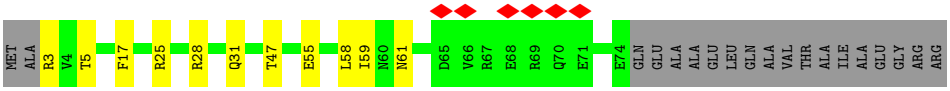
• Molecule 6: DNA-directed RNA polymerase subunit alpha

Chain D:



• Molecule 7: DNA-directed RNA polymerase subunit omega

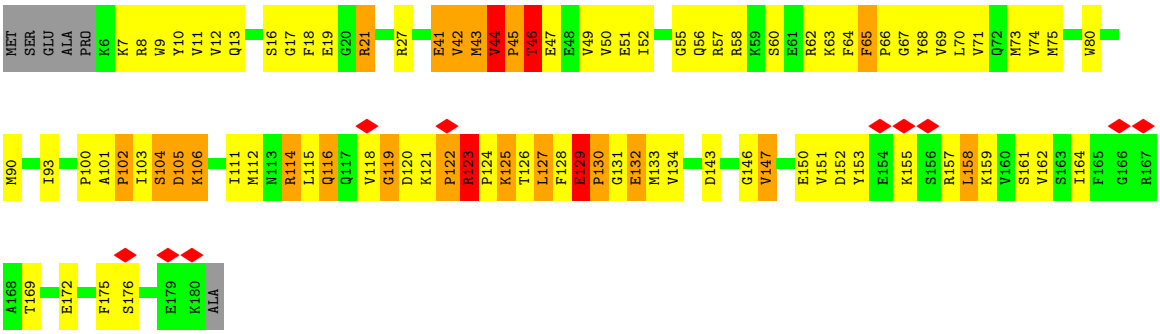
Chain E:



• Molecule 8: Transcription termination/antitermination protein NusG

Chain F:

6%
46%
38%
11%


MET SER GLU ALA PRO K6 K7 R8 W9 Y10 V11 V12 Q13 S16 G17 F18 E19 G20 R21 R27 E41 V42 M43 V44 P45 T46 E47 E48 V49 V50 E51 I52 G55 Q56 R57 R58 K59 S60 E61 R62 K63 F64 F65 P66 G67 Y68 V69 L70 V71 Q72 M73 V74 M75 W80
M90 I93 P100 A101 P102 I103 S104 D105 K106 I111 M112 N113 R114 L115 Q116 Q117 V118 G119 D120 K121 P122 R123 P124 K125 T126 L127 F128 E129 P130 G131 E132 M133 V134 D143 G146 V147 E150 V151 D152 Y153 E154 K155 S156 R157 L158 K159 V160 S161 V162 S163 I164 F165 G166 R167
A168 T169 E172 F175 S176 E179 K180 ALA

4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	69316	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 ($e^-/Å^2$)	52	Depositor
Minimum defocus (nm)	1250	Depositor
Maximum defocus (nm)	2500	Depositor
Magnification	Not provided	
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor
Maximum map value	0.031	Depositor
Minimum map value	-0.011	Depositor
Average map value	-0.000	Depositor
Map value standard deviation	0.001	Depositor
Recommended contour level	0.0035	Depositor
Map size (Å)	410.496, 410.496, 410.496	wwPDB
Map dimensions	384, 384, 384	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.069, 1.069, 1.069	Depositor

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: ZN, MG

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	5	1.12	8/757 (1.1%)	0.95	1/1163 (0.1%)
2	6	1.09	4/834 (0.5%)	1.44	7/1283 (0.5%)
3	7	0.66	1/630 (0.2%)	1.24	10/978 (1.0%)
4	A	0.43	1/10736 (0.0%)	0.60	1/14487 (0.0%)
5	B	0.52	3/10545 (0.0%)	0.66	5/14236 (0.0%)
6	C	0.41	0/1718	0.62	0/2328
6	D	0.36	0/1696	0.62	0/2298
7	E	0.34	0/579	0.59	0/779
8	F	0.44	2/1421 (0.1%)	0.67	6/1914 (0.3%)
All	All	0.52	19/28916 (0.1%)	0.70	30/39466 (0.1%)

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	B	0	5
6	C	0	1
6	D	0	1
All	All	0	7

The worst 5 of 19 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	B	88	CYS	CB-SG	-10.16	1.65	1.82
1	5	109	DT	O3'-P	8.68	1.71	1.61
4	A	828	PHE	C-N	8.65	1.53	1.34
2	6	10	DG	C1'-N9	-8.10	1.35	1.47
8	F	102	PRO	C-N	8.06	1.52	1.34

The worst 5 of 30 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	6	27	DG	O3'-P-O5'	23.17	148.02	104.00
2	6	27	DG	OP1-P-O3'	-20.72	59.62	105.20
2	6	28	DA	P-O3'-C3'	13.79	136.24	119.70
2	6	28	DA	O3'-P-O5'	11.61	126.05	104.00
3	7	9	U	C2-N3-C4	-11.35	120.19	127.00

There are no chirality outliers.

5 of 7 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
5	B	1184	ASP	Peptide
5	B	1326	GLN	Peptide
5	B	313	GLY	Peptide
5	B	416	ILE	Peptide
5	B	804	ALA	Peptide

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	5	676	237	373	52	0
2	6	748	281	419	52	0
3	7	567	109	280	55	0
4	A	10567	0	10585	230	0
5	B	10388	10612	10611	388	0
6	C	1698	0	1718	15	0
6	D	1677	0	1713	23	0
7	E	577	0	588	10	0
8	F	1392	0	1369	251	0
9	B	1	0	0	0	0
10	B	2	0	0	0	0
All	All	28293	11239	27656	915	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 16.

The worst 5 of 915 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)
8:F:44:VAL:HG21	8:F:69:VAL:CG1	1.26	1.64
8:F:52:ILE:HD12	8:F:57:ARG:CD	1.27	1.57
8:F:44:VAL:CG2	8:F:69:VAL:HG13	1.09	1.55
8:F:52:ILE:CD1	8:F:57:ARG:HD2	1.25	1.54
8:F:51:GLU:HB2	8:F:147:VAL:CG2	1.37	1.48

There are no symmetry-related clashes.

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
4	A	1338/1342 (100%)	1221 (91%)	116 (9%)	1 (0%)	51 85
5	B	1329/1407 (94%)	1200 (90%)	120 (9%)	9 (1%)	22 62
6	C	217/329 (66%)	203 (94%)	12 (6%)	2 (1%)	17 56
6	D	214/329 (65%)	198 (92%)	16 (8%)	0	100 100
7	E	70/91 (77%)	67 (96%)	3 (4%)	0	100 100
8	F	173/181 (96%)	142 (82%)	22 (13%)	9 (5%)	2 22
All	All	3341/3679 (91%)	3031 (91%)	289 (9%)	21 (1%)	29 65

5 of 21 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
8	F	42	VAL
8	F	44	VAL
8	F	122	PRO
8	F	123	ARG
5	B	175	GLU

5.3.2 Protein sidechains (i)

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	A	1155/1157 (100%)	1143 (99%)	12 (1%)	76 86
5	B	1120/1168 (96%)	1049 (94%)	71 (6%)	18 45
6	C	186/286 (65%)	186 (100%)	0	100 100
6	D	185/286 (65%)	185 (100%)	0	100 100
7	E	63/75 (84%)	63 (100%)	0	100 100
8	F	150/158 (95%)	123 (82%)	27 (18%)	1 11
All	All	2859/3130 (91%)	2749 (96%)	110 (4%)	36 58

5 of 110 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
5	B	215	LYS
5	B	291	ILE
8	F	176	SER
8	F	143	ASP
5	B	222	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 18 such sidechains are listed below:

Mol	Chain	Res	Type
6	D	117	HIS
8	F	116	GLN
7	E	31	GLN
4	A	894	GLN
6	D	66	HIS

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
3	7	26/35 (74%)	12 (46%)	0

5 of 12 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
3	7	13	U
3	7	14	U
3	7	15	U
3	7	17	U
3	7	19	U

There are no RNA pucker outliers to report.

5.4 Non-standard residues in protein, DNA, RNA chains (i)

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

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 (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

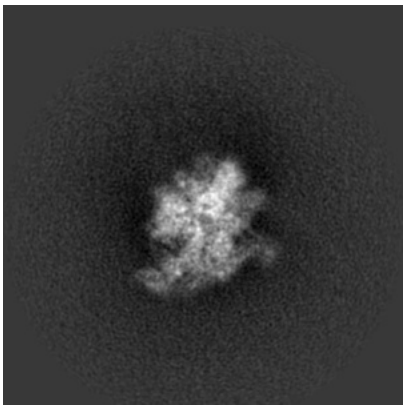
6 Map visualisation

This section contains visualisations of the EMDB entry EMD-27930. These allow visual inspection of the internal detail of the map and identification of artifacts.

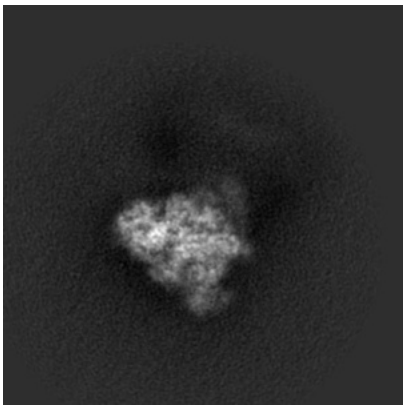
Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

6.1 Orthogonal projections

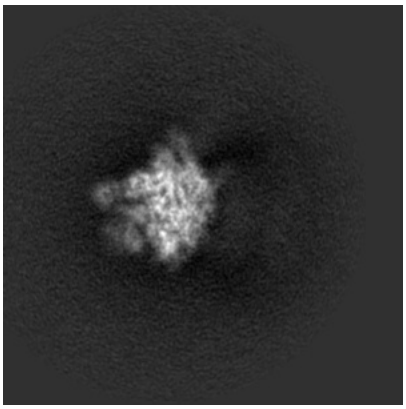
6.1.1 Primary map



X

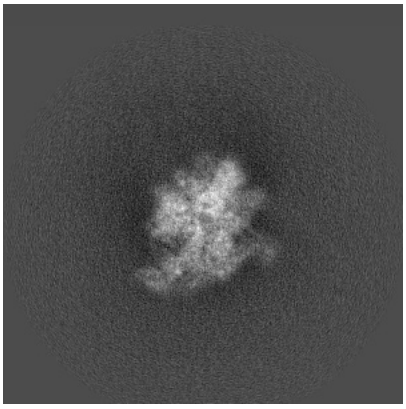


Y

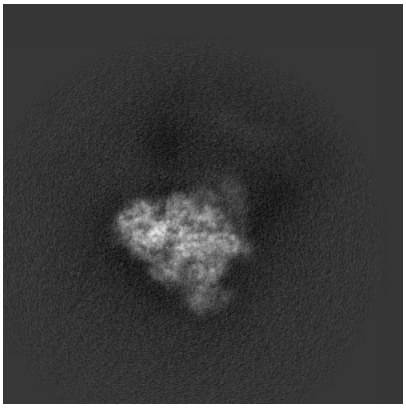


Z

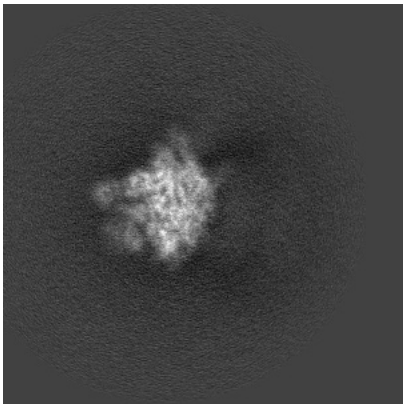
6.1.2 Raw map



X



Y

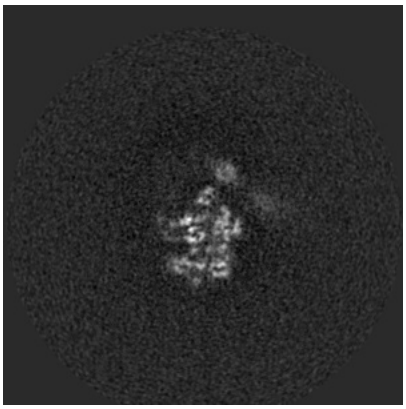


Z

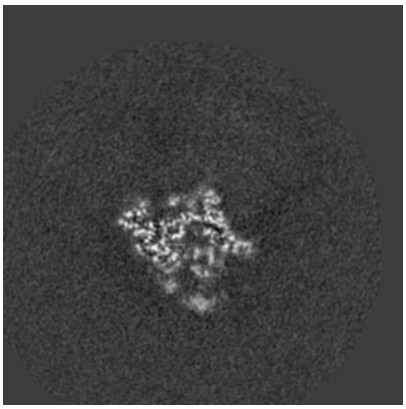
The images above show the map projected in three orthogonal directions.

6.2 Central slices (i)

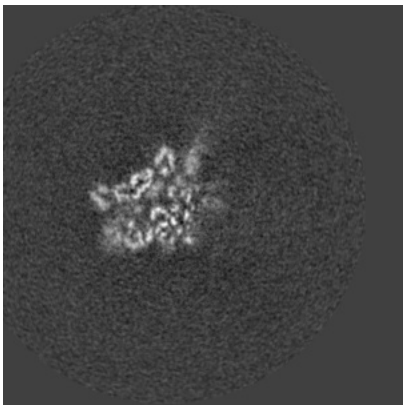
6.2.1 Primary map



X Index: 192

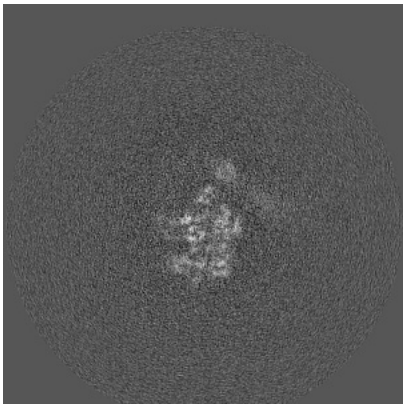


Y Index: 192

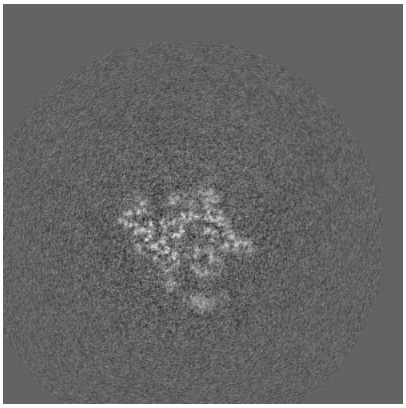


Z Index: 192

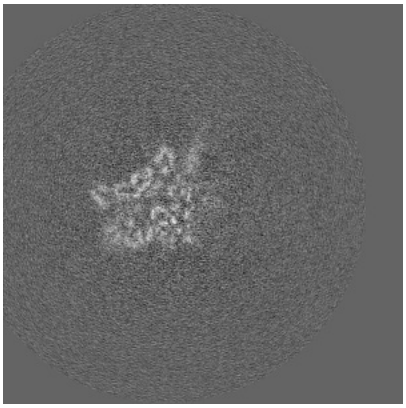
6.2.2 Raw map



X Index: 192



Y Index: 192

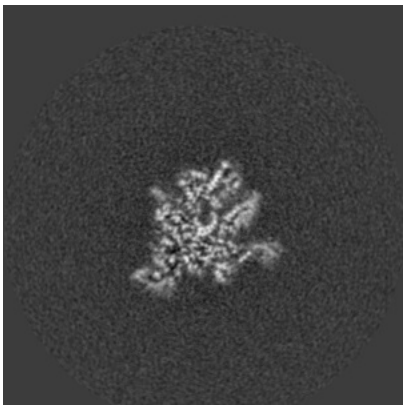


Z Index: 192

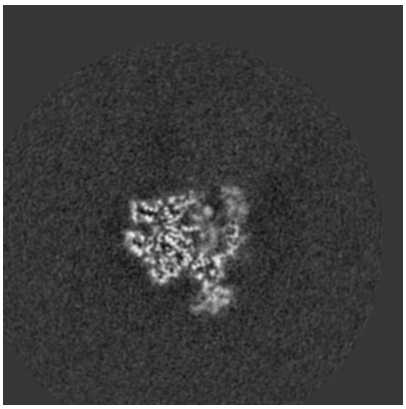
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

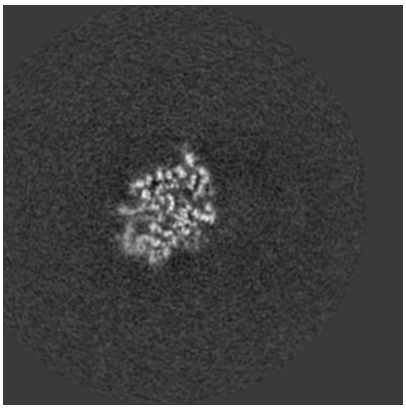
6.3.1 Primary map



X Index: 161

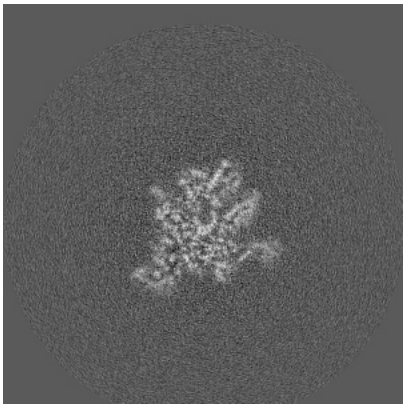


Y Index: 209

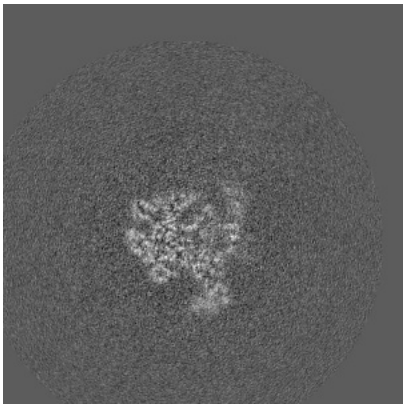


Z Index: 167

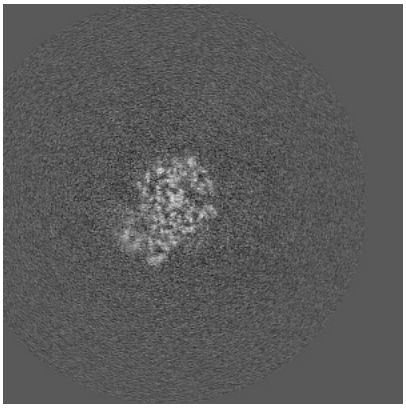
6.3.2 Raw map



X Index: 161



Y Index: 207



Z Index: 170

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

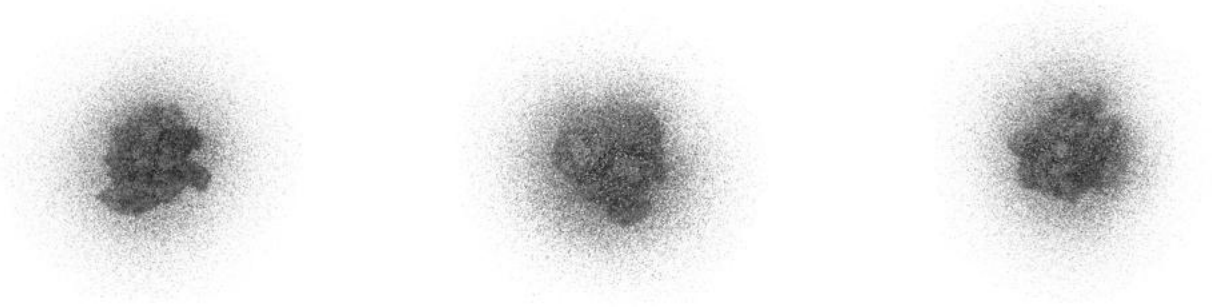
6.4 Orthogonal surface views (i)

6.4.1 Primary map

X Y Z

The images above show the 3D surface view of the map at the recommended contour level 0.0035. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

6.4.2 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.

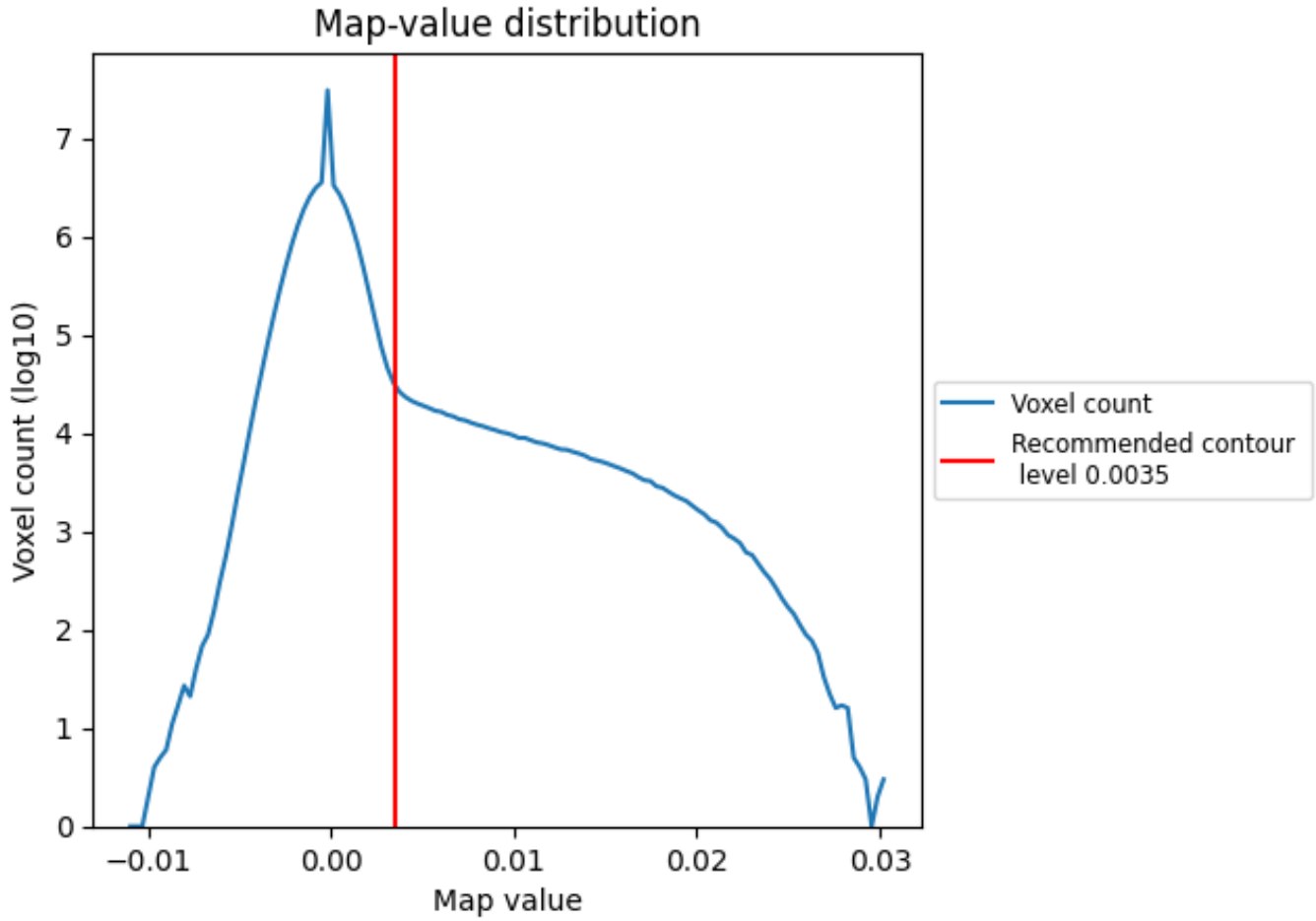
6.5 Mask visualisation (i)

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

7 Map analysis

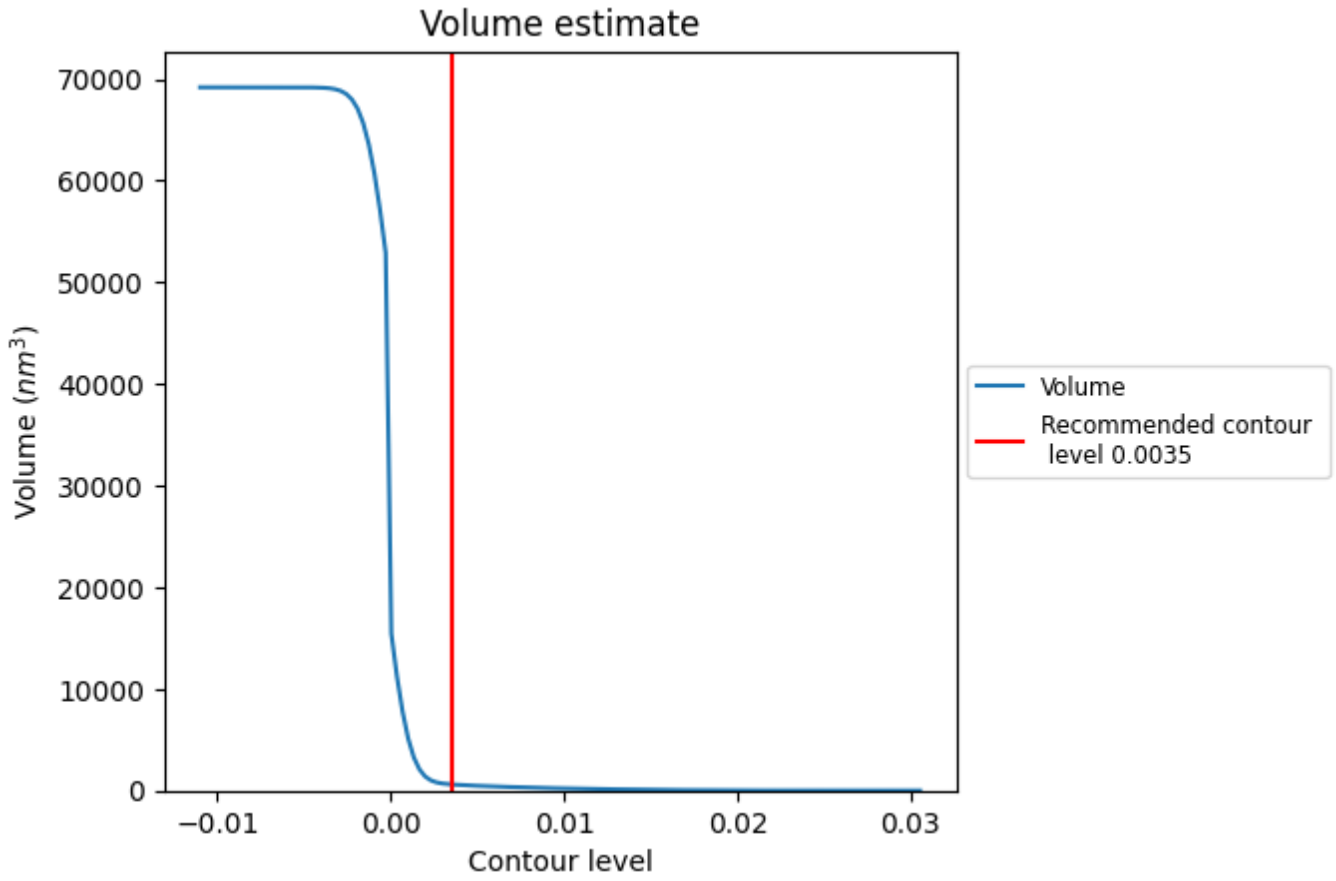
This section contains the results of statistical analysis of the map.

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.

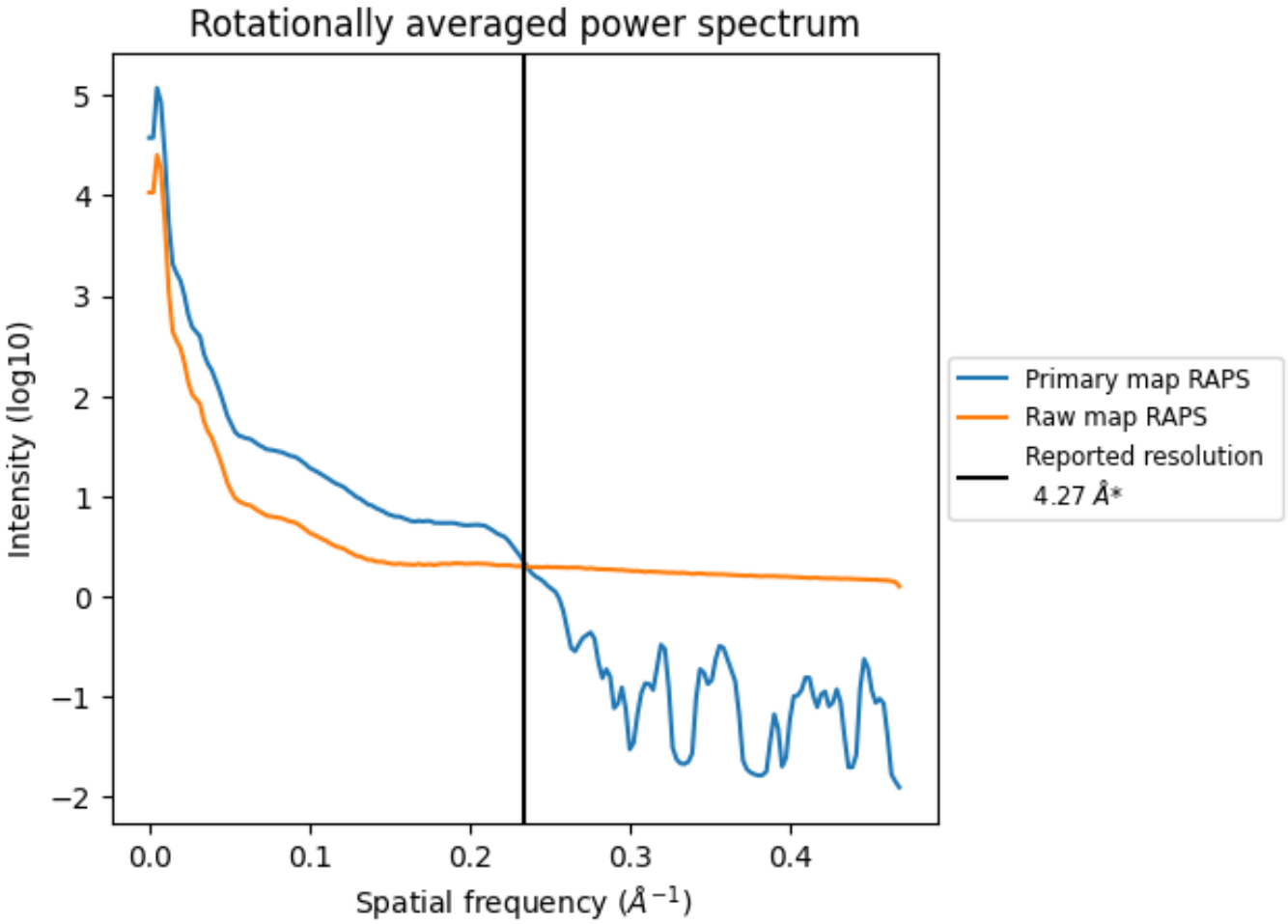
7.2 Volume estimate (i)



The volume at the recommended contour level is 606 nm^3; this corresponds to an approximate mass of 547 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)

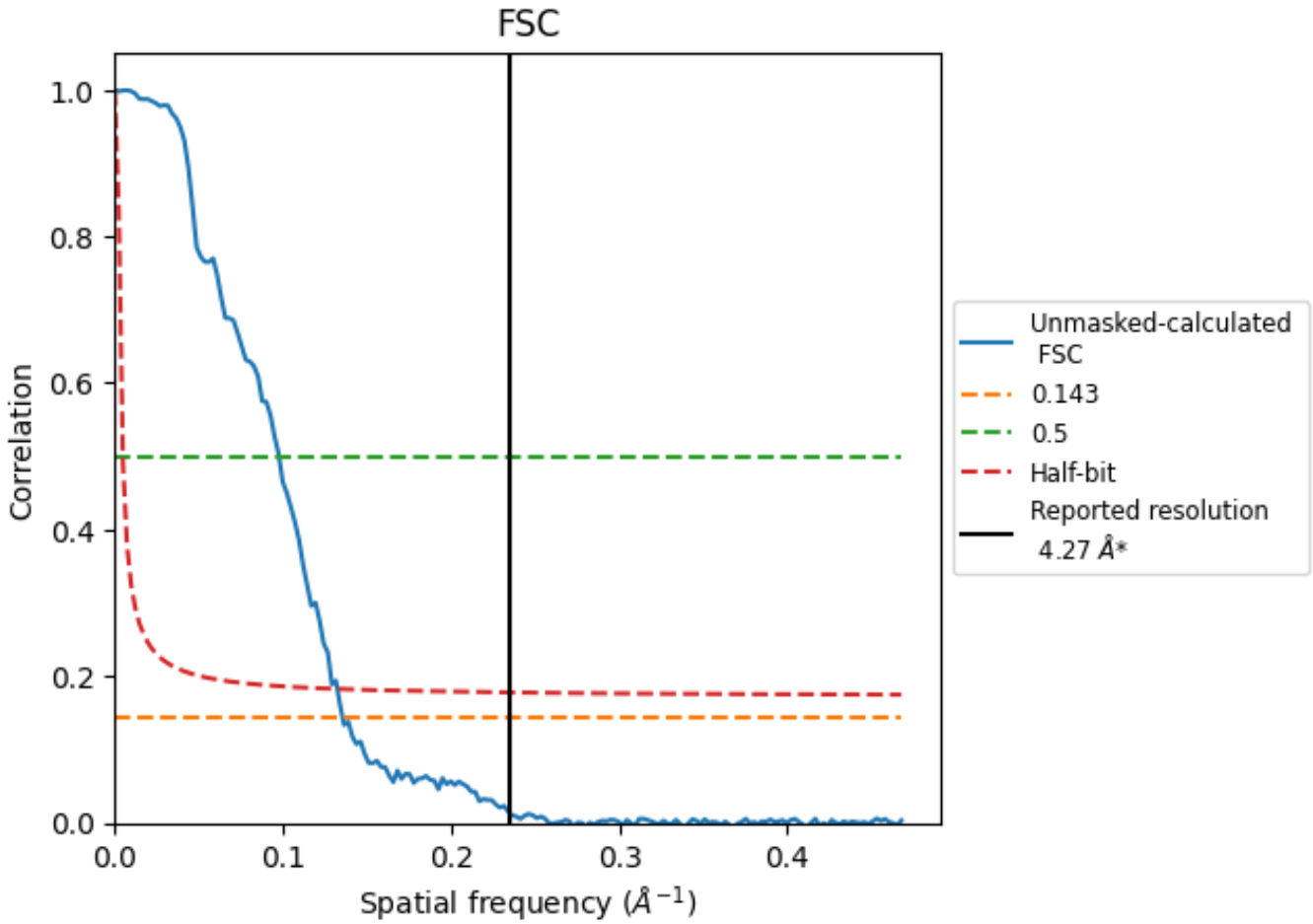


*Reported resolution corresponds to spatial frequency of 0.234 Å^{-1}

8 Fourier-Shell correlation

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



*Reported resolution corresponds to spatial frequency of 0.234 Å^{-1}

8.2 Resolution estimates (i)

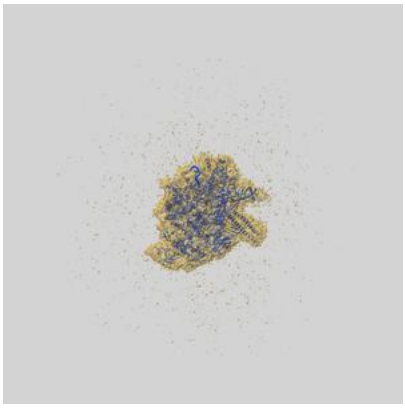
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	4.27	-	-
Author-provided FSC curve	-	-	-
Unmasked-calculated*	7.38	10.24	7.55

*Resolution estimate based on FSC curve calculated by comparison of deposited half-maps. The value from deposited half-maps intersecting FSC 0.143 CUT-OFF 7.38 differs from the reported value 4.27 by more than 10 %

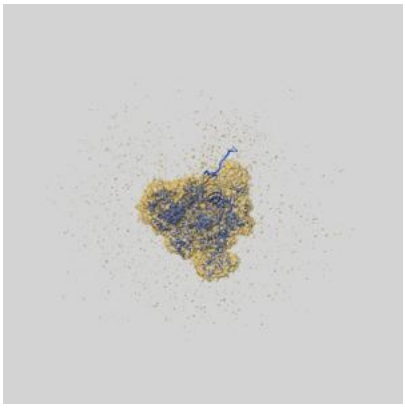
9 Map-model fit

This section contains information regarding the fit between EMDB map EMD-27930 and PDB model 8E6X. Per-residue inclusion information can be found in section 3 on page 6.

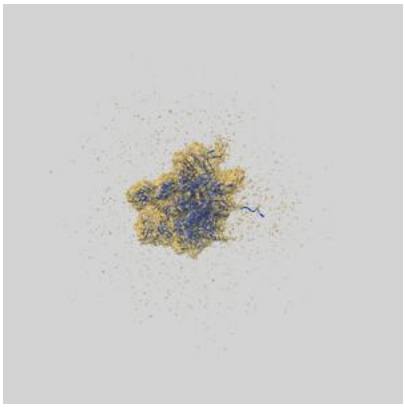
9.1 Map-model overlay



X



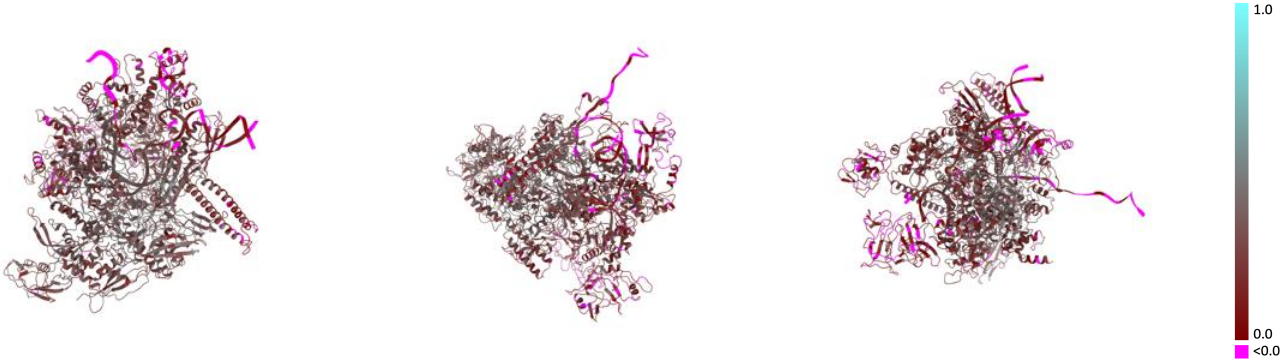
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.0035 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.

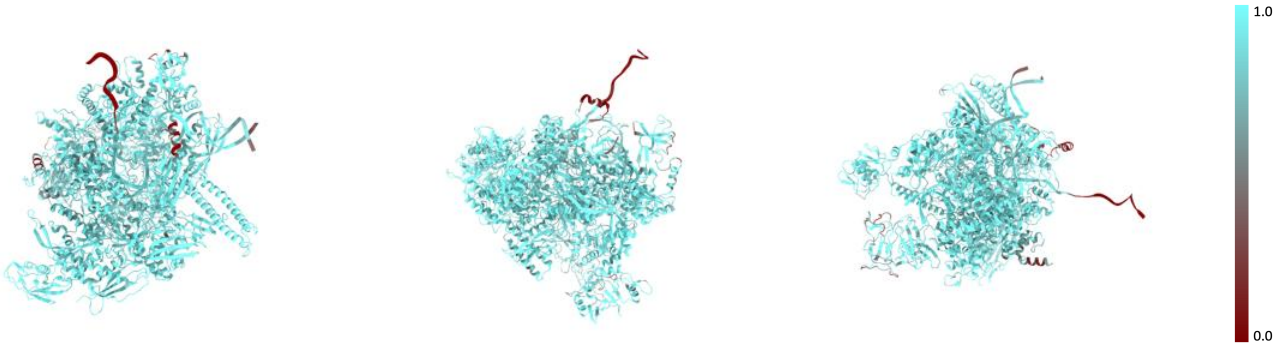
9.2 Q-score mapped to coordinate model (i)



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.

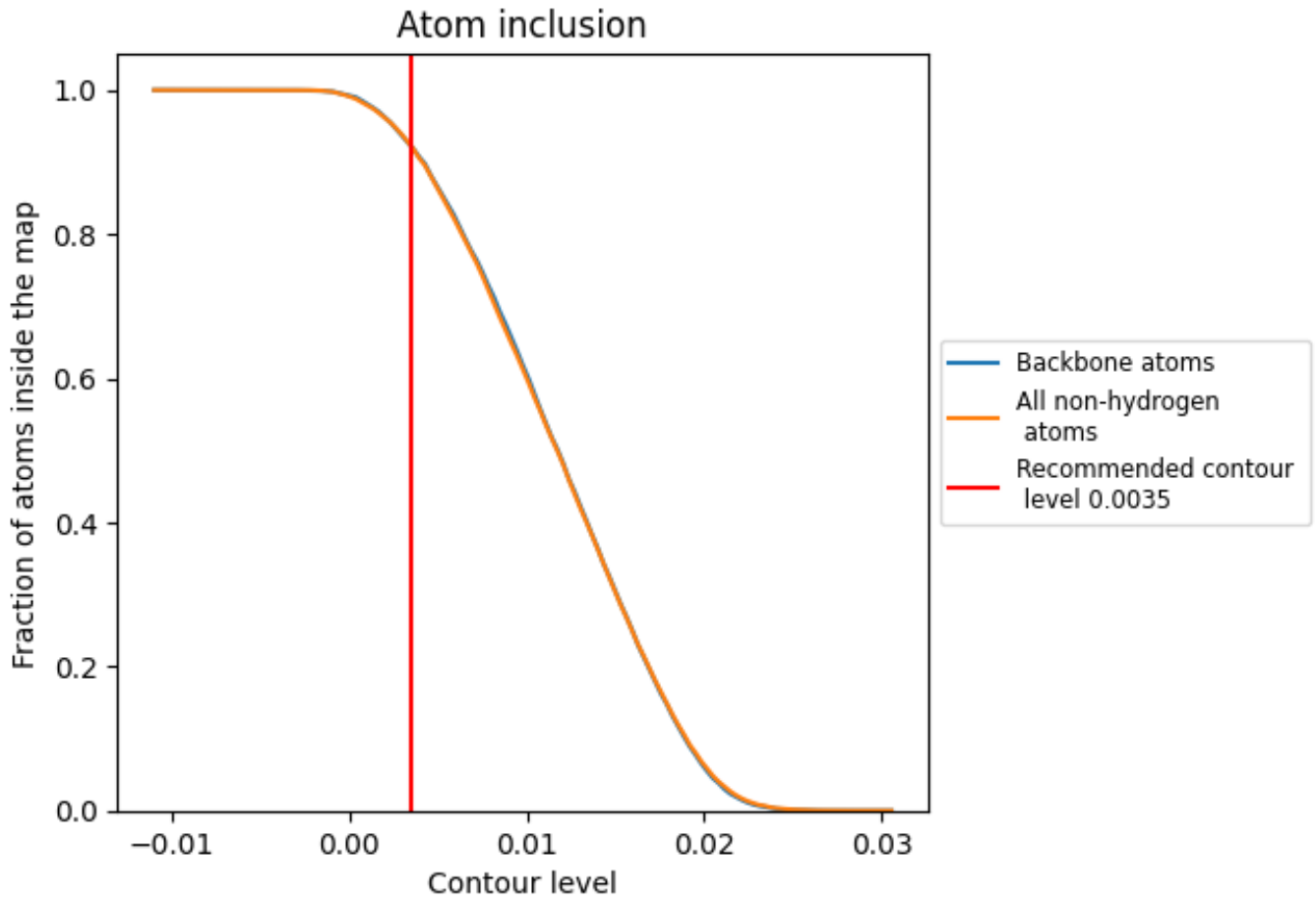
9.3 Atom inclusion mapped to coordinate model (i)



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 (0.0035).

9.4 Atom inclusion (i)



At the recommended contour level, 92% 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 (0.0035) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.9217	0.2640
5	0.9098	0.1460
6	0.9051	0.2030
7	0.5908	0.1390
A	0.9423	0.2910
B	0.9232	0.2550
C	0.9549	0.3360
D	0.9513	0.3050
E	0.6827	0.2190
F	0.8924	0.1370