wwPDB EM Validation Summary Report

Oct 6, 2024 – 07:57 pm BST

PDB ID	:	7AOF
EMDB ID	:	EMD-11843
Title	:	Atomic structure of the poxvirus transcription late pre-initiation complex
Authors	:	Grimm, C.; Bartuli, J.; Fischer, U.
Deposited on	:	2020-10-14
Resolution	:	2.98 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 ⓘ 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 ⓘ) were used in the production of this report:

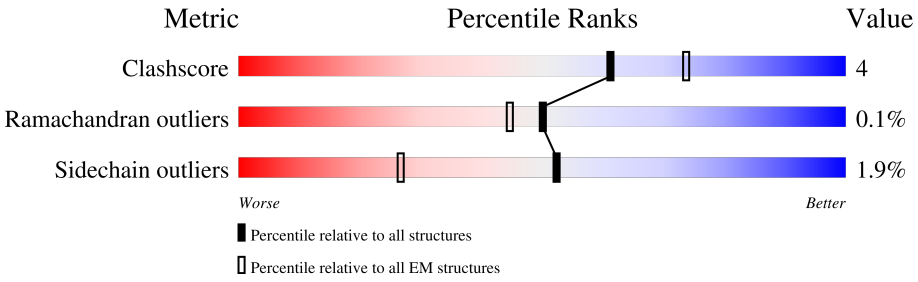
EMDB validation analysis	:	0.0.1.dev113
Mogul	:	1.8.4, CSD as541be (2020)
MolProbity	:	4.02b-467
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
MapQ	:	1.9.13
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

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 2.98 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	210492	15764
Ramachandran outliers	207382	16835
Sidechain outliers	206894	16415

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments of the bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the EM map (all-atom inclusion < 40%). The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	1286	87% 11% ••
2	B	1164	85% 12% •
3	C	305	88% 11% •
4	E	185	85% 14% ••
5	F	164	57% 5% 37%
6	G	161	82% 12% • 5%
7	I	795	73% 10% 16%
8	J	63	76% 17% • •

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
9	S	259	6% 57% 8% 36%

2 Entry composition (i)

There are 11 unique types of molecules in this entry. The entry contains 32727 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 147 kDa polypeptide.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	1268	10188	6556	1679	1908	45	0	0

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

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	B	1129	9091	5794	1554	1695	48	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	6	ASN	ASP	conflict	UNP Q49PH2

- Molecule 3 is a protein called DNA-directed RNA polymerase 35 kDa subunit.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	C	304	2484	1608	399	464	13	0	0

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

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	E	184	1495	966	248	276	5	0	0

- Molecule 5 is a protein called DNA-directed RNA polymerase 19 kDa subunit.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
5	F	103	849	545	148	153	3	0	0

- Molecule 6 is a protein called DNA-directed RNA polymerase 18 kDa subunit.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	G	153	Total	C	N	O	S	0	0
			1192	753	198	235	6		

- Molecule 7 is a protein called Protein H4.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	I	665	Total	C	N	O	S	0	0
			5573	3649	891	1012	21		

- Molecule 8 is a protein called DNA-directed RNA polymerase 7 kDa subunit.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	J	61	Total	C	N	O	S	0	0
			490	310	88	88	4		

- Molecule 9 is a protein called DNA-directed RNA polymerase 30 kDa polypeptide.

Mol	Chain	Residues	Atoms						AltConf	Trace
9	S	167	Total	C	N	O	P	S	0	0
			1360	849	219	285	3	4		

- Molecule 10 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		AltConf
10	A	1	Total	Mg	0
			1	1	

- Molecule 11 is ZINC ION (three-letter code: ZN) (formula: Zn).

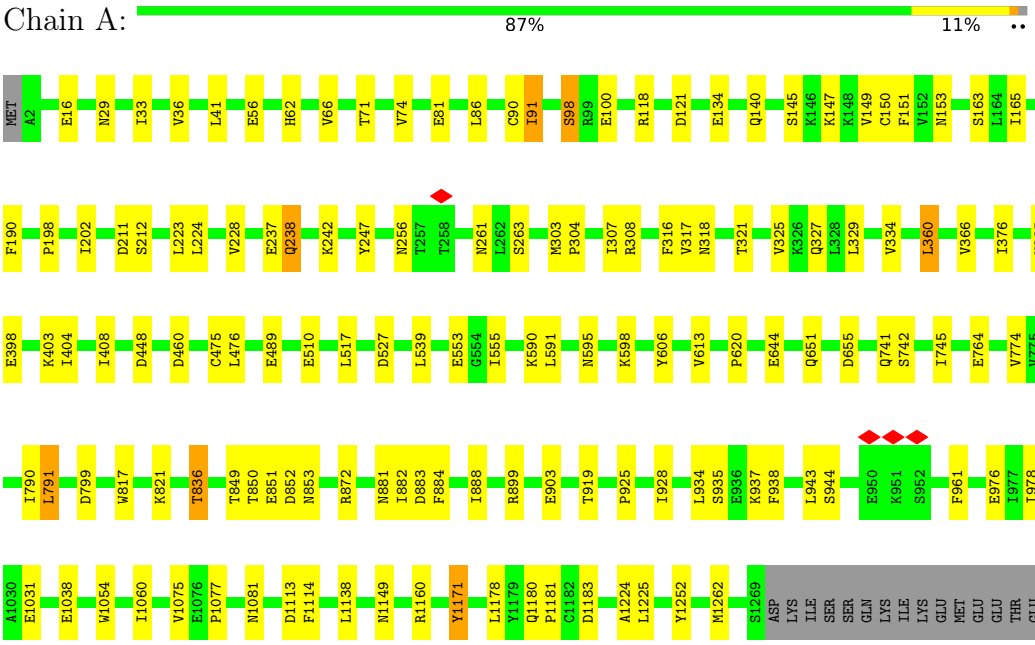
Mol	Chain	Residues	Atoms		AltConf
11	A	2	Total	Zn	0
			2	2	
11	B	1	Total	Zn	0
			1	1	
11	I	1	Total	Zn	0
			1	1	

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.

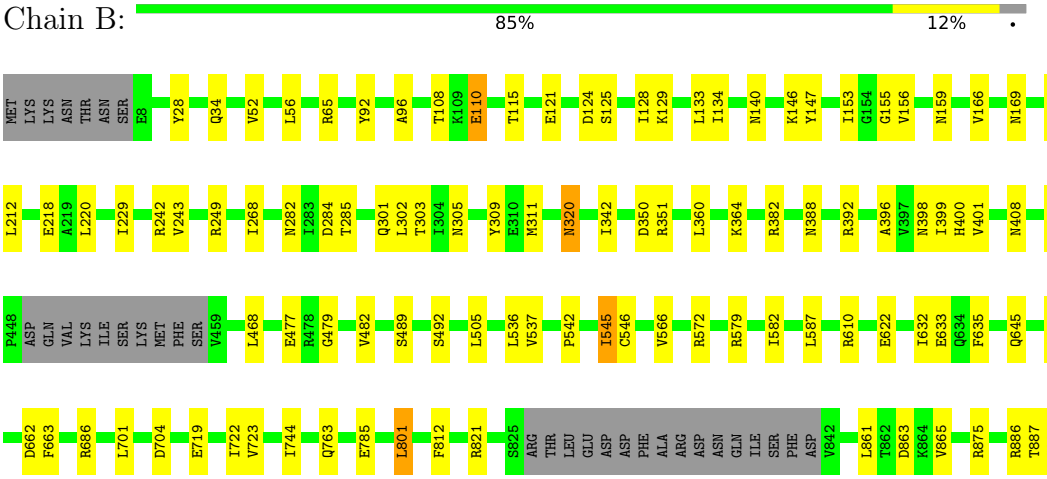
• Molecule 1: DNA-directed RNA polymerase 147 kDa polypeptide

Chain A:



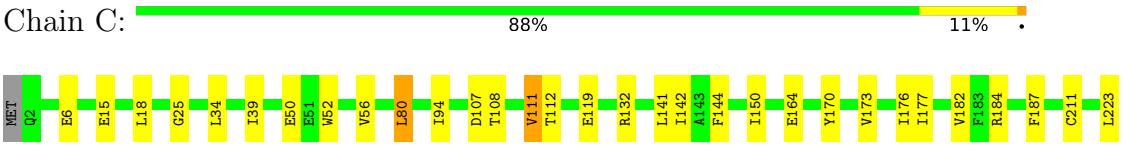
• Molecule 2: DNA-directed RNA polymerase

Chain B:



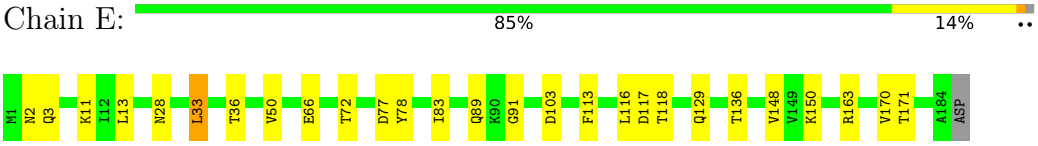
• Molecule 3: DNA-directed RNA polymerase 35 kDa subunit

Chain C:



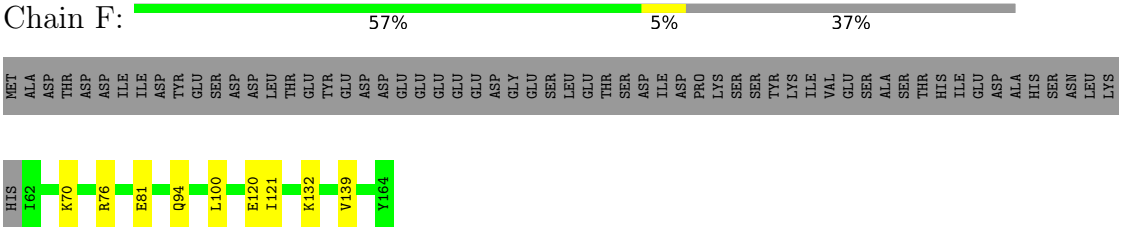
• Molecule 4: DNA-directed RNA polymerase subunit

Chain E:



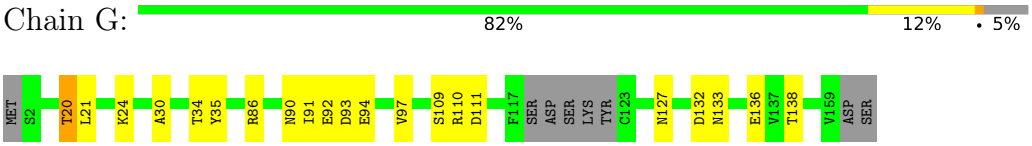
• Molecule 5: DNA-directed RNA polymerase 19 kDa subunit

Chain F:



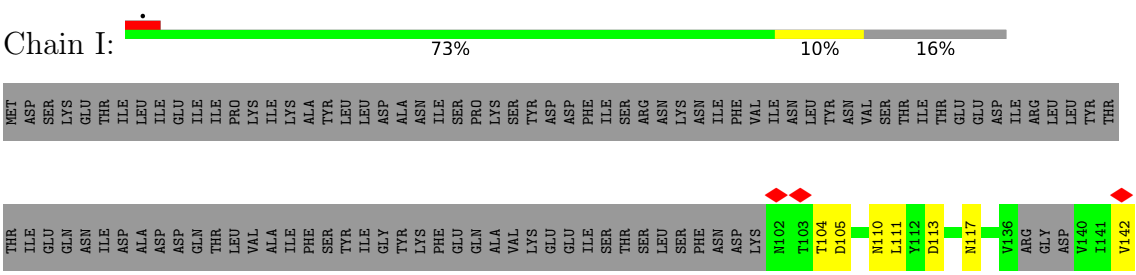
• Molecule 6: DNA-directed RNA polymerase 18 kDa subunit

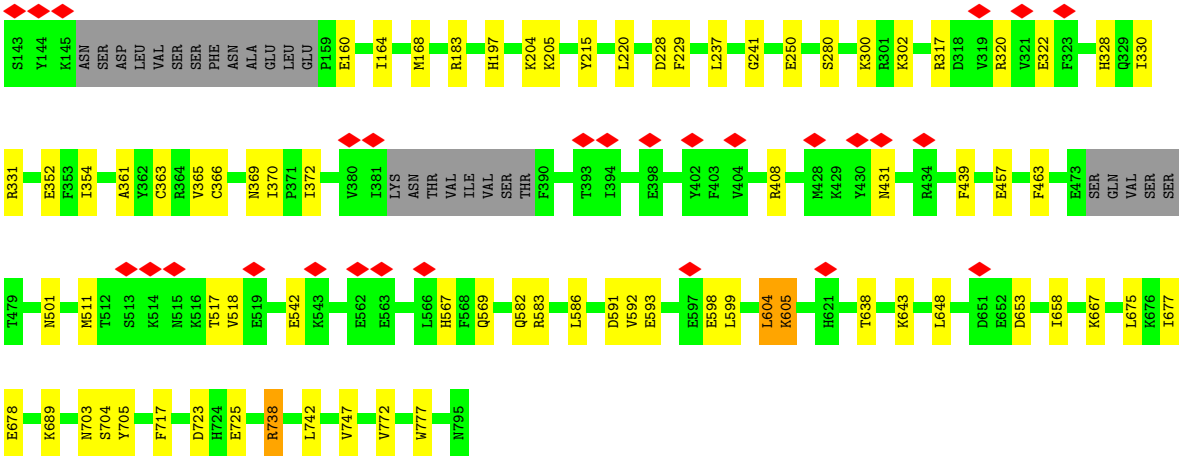
Chain G:



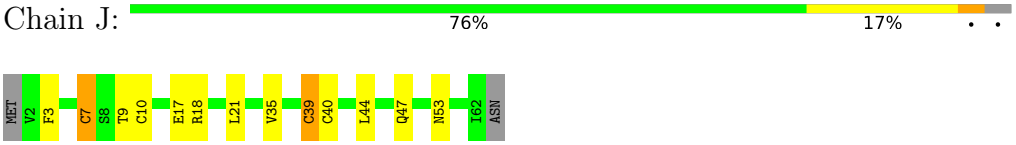
• Molecule 7: Protein H4

Chain I:

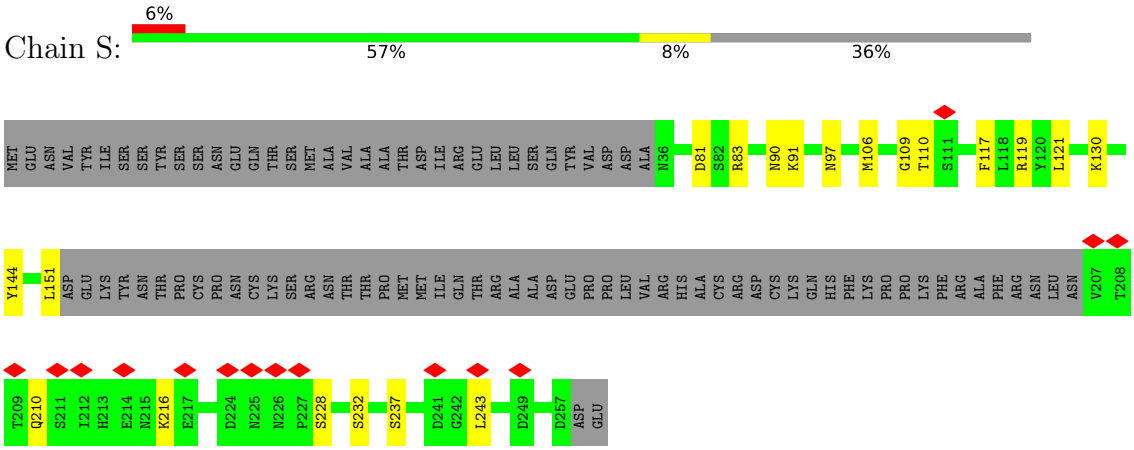




• Molecule 8: DNA-directed RNA polymerase 7 kDa subunit



• Molecule 9: DNA-directed RNA polymerase 30 kDa polypeptide



4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	88828	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 ($e^-/Å^2$)	78.90	Depositor
Minimum defocus (nm)	Not provided	
Maximum defocus (nm)	Not provided	
Magnification	Not provided	
Image detector	FEI FALCON III (4k x 4k)	Depositor
Maximum map value	0.121	Depositor
Minimum map value	-0.040	Depositor
Average map value	0.000	Depositor
Map value standard deviation	0.004	Depositor
Recommended contour level	0.0163	Depositor
Map size (Å)	340.32, 340.32, 340.32	wwPDB
Map dimensions	320, 320, 320	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.0635, 1.0635, 1.0635	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, SEP

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	0/10394	0.41	0/14052
2	B	0.25	0/9281	0.41	0/12537
3	C	0.26	0/2540	0.44	1/3440 (0.0%)
4	E	0.25	0/1522	0.47	1/2069 (0.0%)
5	F	0.24	0/863	0.38	0/1158
6	G	0.25	0/1209	0.44	0/1639
7	I	0.25	0/5703	0.39	0/7712
8	J	0.23	0/494	0.40	0/663
9	S	0.24	0/1352	0.40	0/1817
All	All	0.25	0/33358	0.41	2/45087 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	E	170	VAL	CG1-CB-CG2	6.94	122.00	110.90
3	C	111	VAL	CG1-CB-CG2	6.70	121.62	110.90

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	10188	0	10312	94	0
2	B	9091	0	9146	81	0
3	C	2484	0	2470	24	0
4	E	1495	0	1548	16	0
5	F	849	0	874	7	0
6	G	1192	0	1181	14	0
7	I	5573	0	5618	46	0
8	J	490	0	530	7	0
9	S	1360	0	1316	12	0
10	A	1	0	0	0	0
11	A	2	0	0	0	0
11	B	1	0	0	0	0
11	I	1	0	0	0	0
All	All	32727	0	32995	274	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 4.

The worst 5 of 274 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)
1:A:527:ASP:O	9:S:210:GLN:NE2	2.16	0.78
2:B:542:PRO:HA	2:B:545:ILE:HD13	1.66	0.77
4:E:11:LYS:NZ	4:E:103:ASP:OD2	2.18	0.77
7:I:320:ARG:NH1	7:I:322:GLU:OE1	2.17	0.76
1:A:976:GLU:OE2	1:A:1149:ASN:ND2	2.18	0.76

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	
1	A	1266/1286 (98%)	1211 (96%)	55 (4%)	0	100	100
2	B	1123/1164 (96%)	1075 (96%)	48 (4%)	0	100	100
3	C	302/305 (99%)	284 (94%)	18 (6%)	0	100	100
4	E	182/185 (98%)	171 (94%)	11 (6%)	0	100	100
5	F	101/164 (62%)	98 (97%)	3 (3%)	0	100	100
6	G	149/161 (92%)	141 (95%)	8 (5%)	0	100	100
7	I	655/795 (82%)	619 (94%)	34 (5%)	2 (0%)	37	68
8	J	59/63 (94%)	53 (90%)	6 (10%)	0	100	100
9	S	160/259 (62%)	147 (92%)	13 (8%)	0	100	100
All	All	3997/4382 (91%)	3799 (95%)	196 (5%)	2 (0%)	50	79

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
7	I	605	LYS
7	I	604	LEU

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	
1	A	1139/1157 (98%)	1122 (98%)	17 (2%)	60	82
2	B	1030/1064 (97%)	1013 (98%)	17 (2%)	56	80
3	C	286/287 (100%)	277 (97%)	9 (3%)	35	66
4	E	174/175 (99%)	171 (98%)	3 (2%)	56	80
5	F	94/151 (62%)	94 (100%)	0	100	100
6	G	136/144 (94%)	135 (99%)	1 (1%)	81	91
7	I	633/755 (84%)	618 (98%)	15 (2%)	44	72
8	J	60/62 (97%)	55 (92%)	5 (8%)	9	32
9	S	152/237 (64%)	150 (99%)	2 (1%)	65	84

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	3704/4032 (92%)	3635 (98%)	69 (2%)	52 77

5 of 69 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
7	I	569	GLN
7	I	643	LYS
8	J	21	LEU
2	B	159	ASN
2	B	153	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

3 non-standard protein/DNA/RNA residues are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
9	SEP	S	228	9	8,9,10	1.55	1 (12%)	8,12,14	1.57	2 (25%)
9	SEP	S	232	9	8,9,10	1.55	1 (12%)	8,12,14	1.50	2 (25%)
9	SEP	S	237	9	8,9,10	1.57	1 (12%)	8,12,14	1.66	2 (25%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
9	SEP	S	228	9	-	3/5/8/10	-
9	SEP	S	232	9	-	4/5/8/10	-
9	SEP	S	237	9	-	2/5/8/10	-

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
9	S	237	SEP	P-O1P	3.42	1.61	1.50
9	S	232	SEP	P-O1P	3.37	1.61	1.50
9	S	228	SEP	P-O1P	3.36	1.61	1.50

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
9	S	237	SEP	P-OG-CB	-3.04	109.93	118.30
9	S	237	SEP	OG-CB-CA	3.03	111.09	108.14
9	S	232	SEP	OG-CB-CA	2.85	110.92	108.14
9	S	228	SEP	P-OG-CB	-2.83	110.51	118.30
9	S	228	SEP	OG-CB-CA	2.82	110.89	108.14

There are no chirality outliers.

5 of 9 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
9	S	228	SEP	CB-OG-P-O3P
9	S	232	SEP	CA-CB-OG-P
9	S	232	SEP	CB-OG-P-O2P
9	S	232	SEP	CB-OG-P-O1P
9	S	228	SEP	CB-OG-P-O2P

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates (i)

There are no oligosaccharides in this entry.

5.6 Ligand geometry (i)

Of 5 ligands modelled in this entry, 5 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.

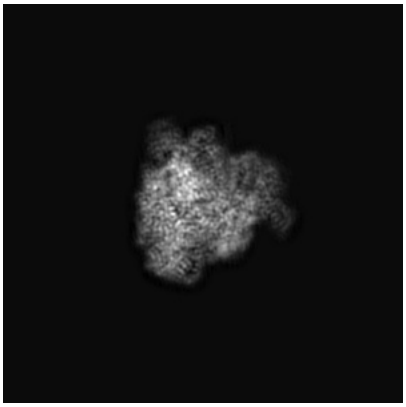
6 Map visualisation

This section contains visualisations of the EMDB entry EMD-11843. These allow visual inspection of the internal detail of the map and identification of artifacts.

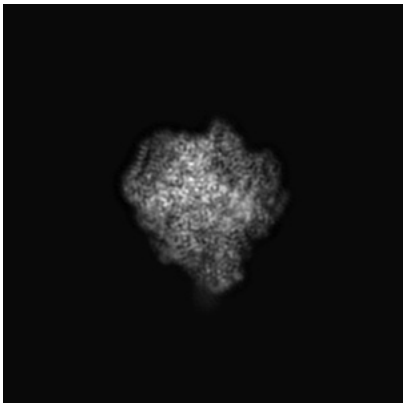
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections

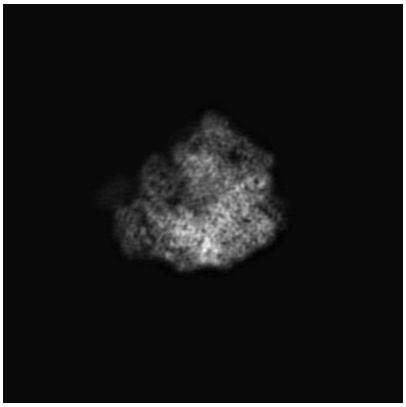
6.1.1 Primary map



X



Y

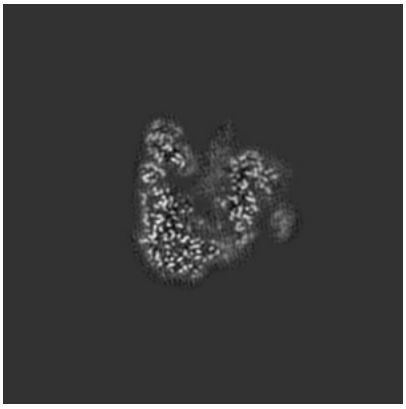


Z

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

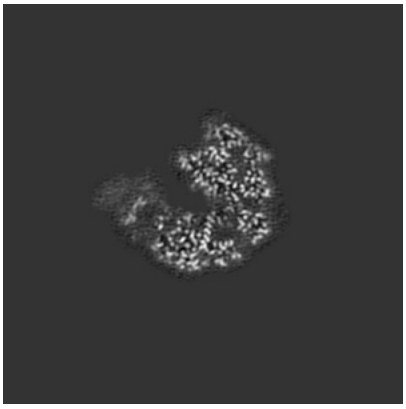
6.2 Central slices

6.2.1 Primary map



X Index: 160

Y Index: 160

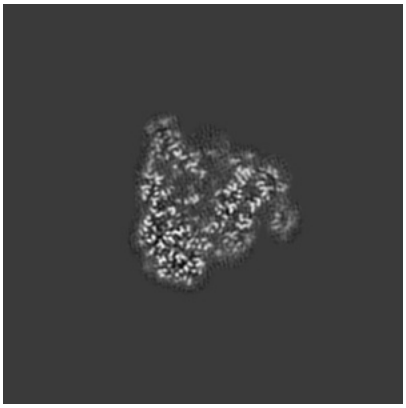


Z Index: 160

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

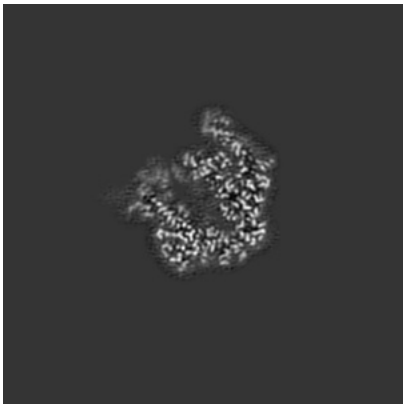
6.3 Largest variance slices (i)

6.3.1 Primary map



X Index: 167

Y Index: 136

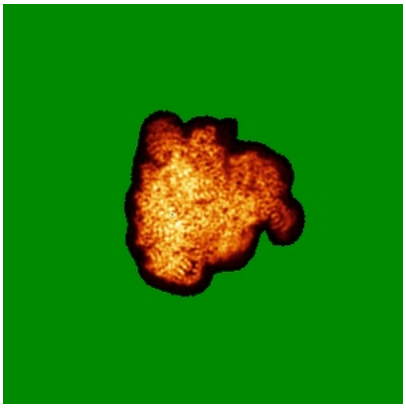


Z Index: 151

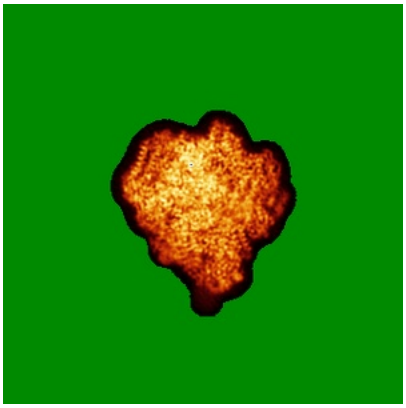
The images above show the largest variance slices of the map in three orthogonal directions.

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

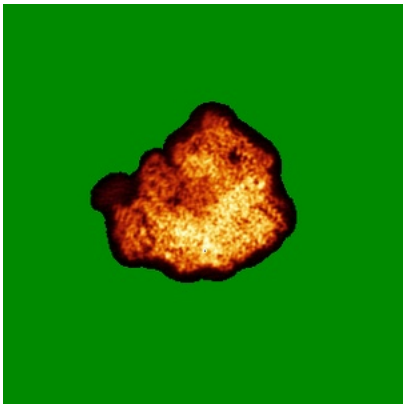
6.4.1 Primary map



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 (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 0.0163. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

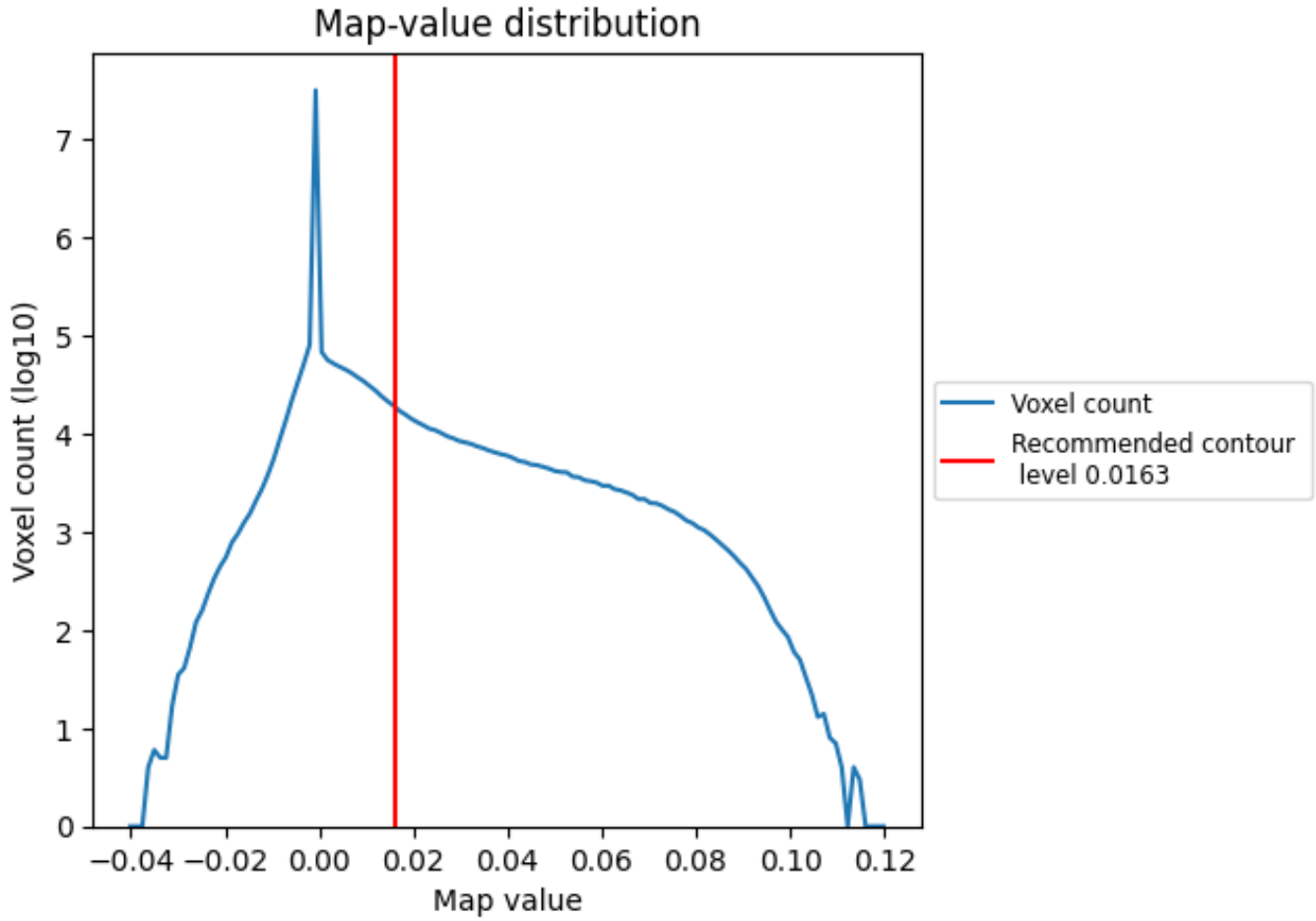
6.6 Mask visualisation (i)

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

7 Map analysis

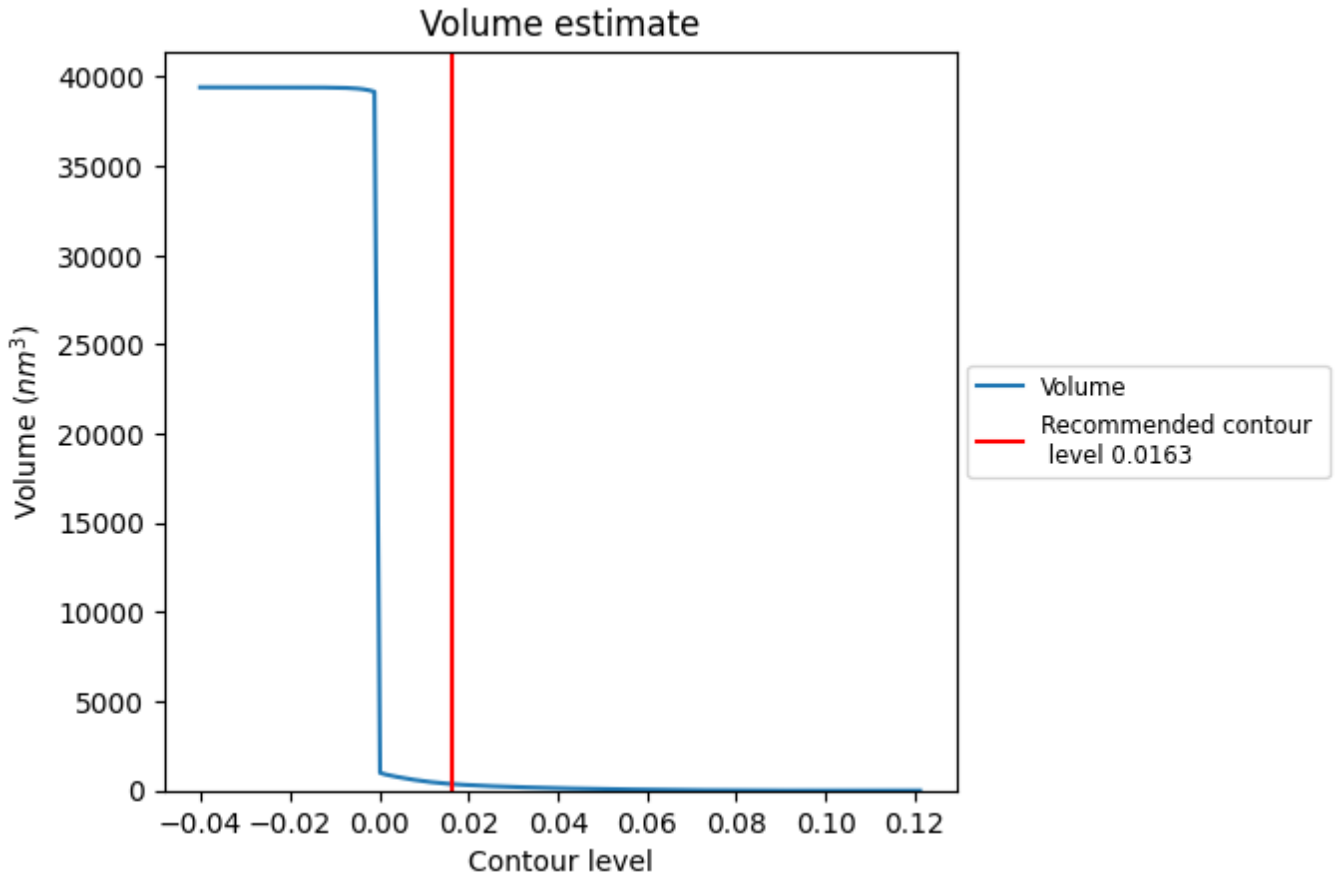
This section contains the results of statistical analysis of the map.

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.

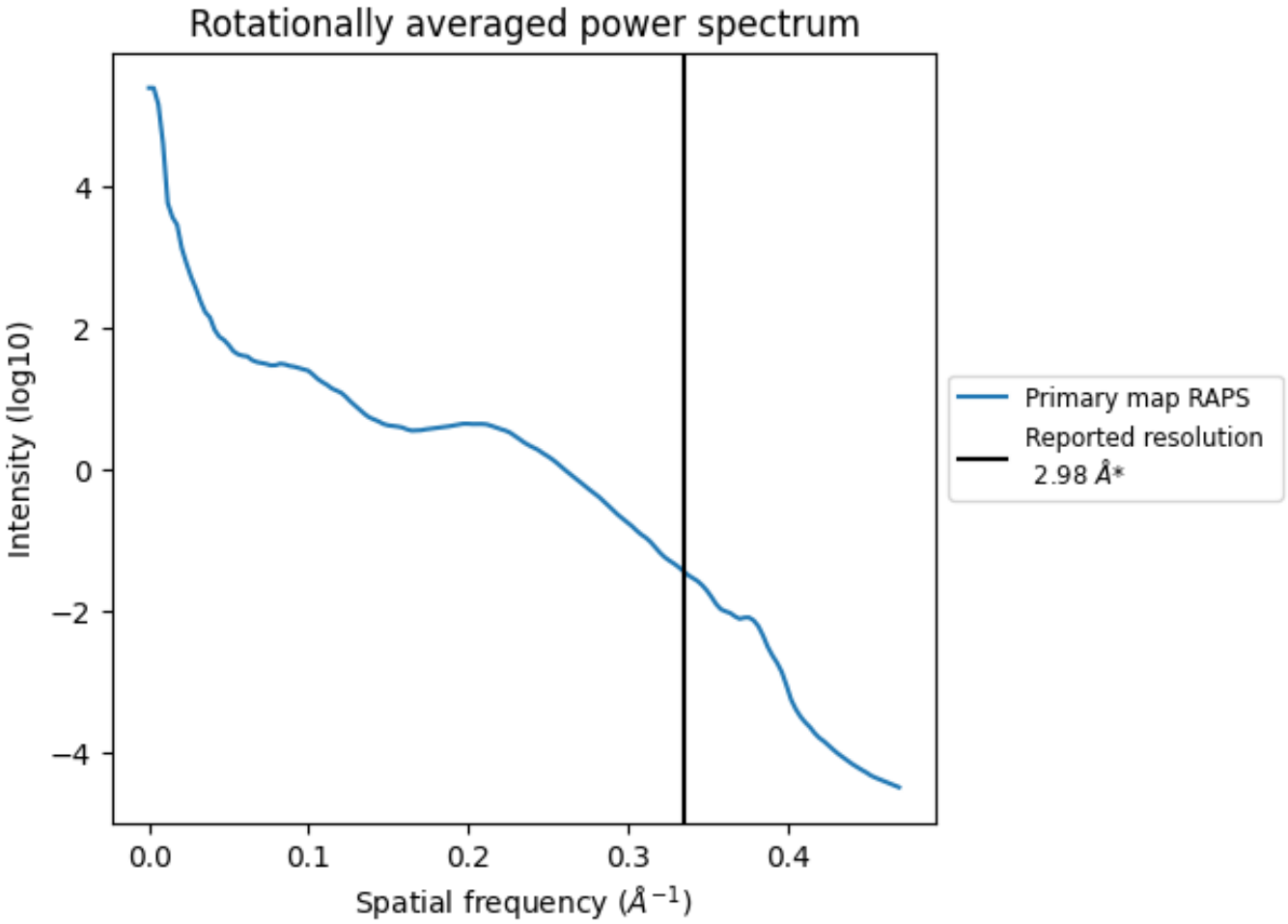
7.2 Volume estimate (i)



The volume at the recommended contour level is 367 nm^3; this corresponds to an approximate mass of 331 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 0.336 Å^{-1}

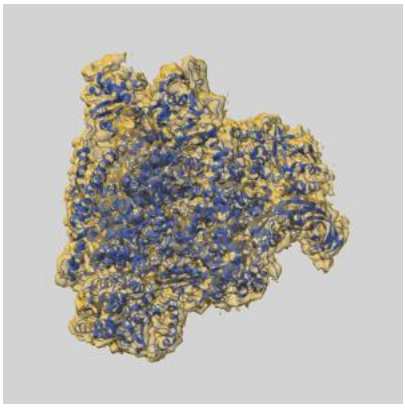
8 Fourier-Shell correlation

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

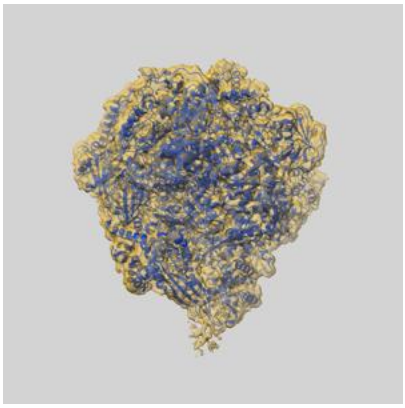
9 Map-model fit

This section contains information regarding the fit between EMDB map EMD-11843 and PDB model 7AOF. Per-residue inclusion information can be found in section 3 on page 6.

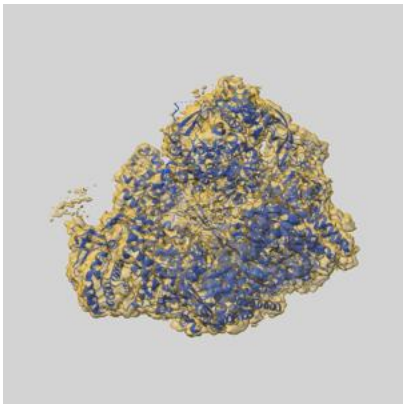
9.1 Map-model overlay



X



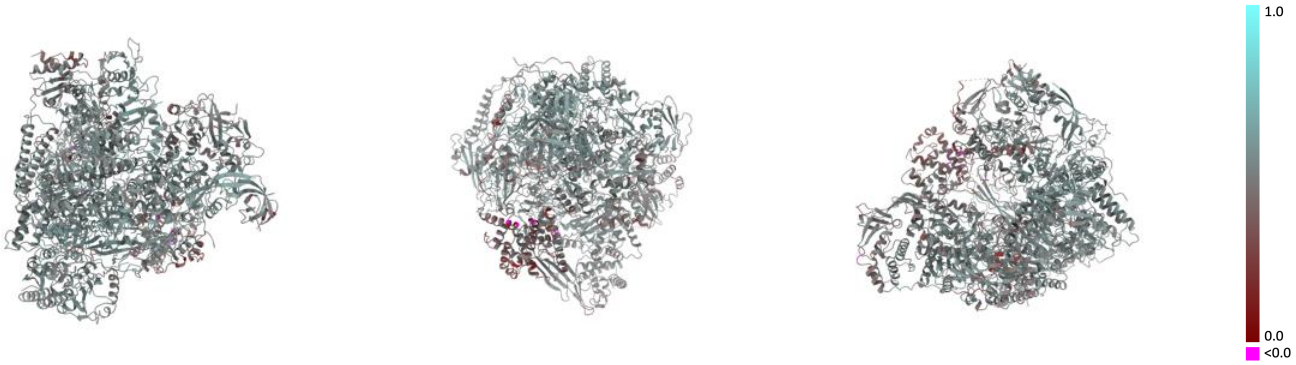
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.0163 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.

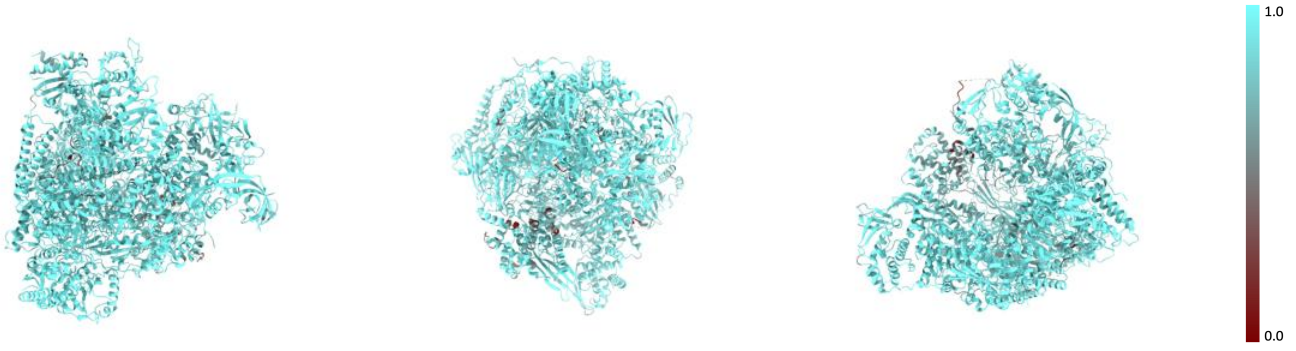
9.2 Q-score mapped to coordinate model (i)



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.

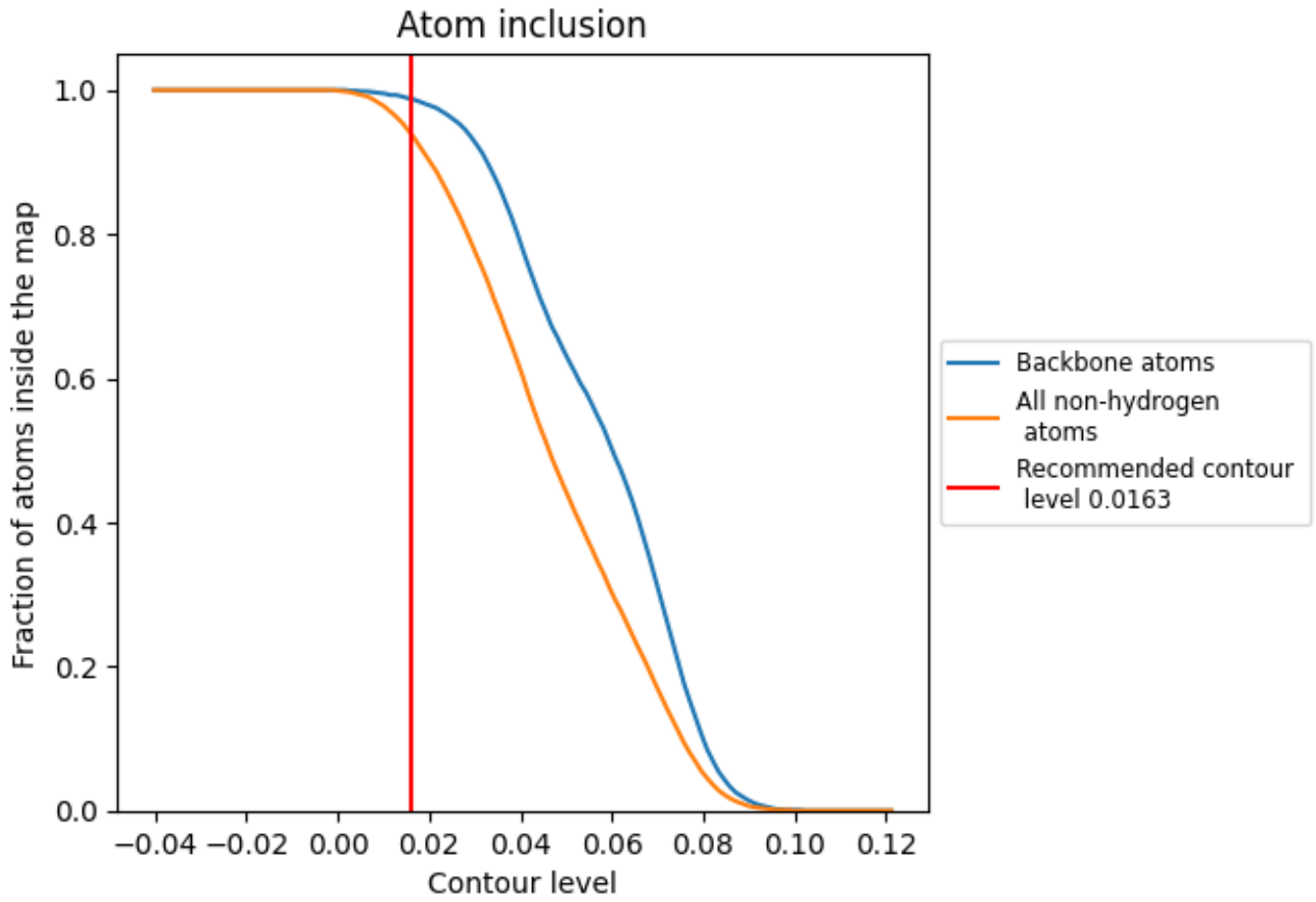
9.3 Atom inclusion mapped to coordinate model (i)



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 (0.0163).

9.4 Atom inclusion (i)



At the recommended contour level, 99% of all backbone atoms, 94% 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 (0.0163) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.9370	0.5050
A	0.9540	0.5200
B	0.9660	0.5300
C	0.9710	0.5210
E	0.9660	0.5290
F	0.9770	0.5400
G	0.9410	0.5110
I	0.8610	0.4310
J	0.9900	0.5330
S	0.7860	0.4400