wwPDB EM Validation Summary Report (i)

Feb 21, 2022 – 07:32 pm GMT

PDB ID	:	7ORI
EMDB ID	:	EMD-13038
Title	:	La Crosse virus polymerase at replication late-elongation stage
Authors	:	Arragain, B.; Durieux Trouilleton, Q.; Baudin, F.; Cusack, S.; Schoehn, G.; Malet, H.
Deposited on	:	2021-06-06
Resolution	:	3.90 Å(reported)
Based on initial model	:	6Z8K

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

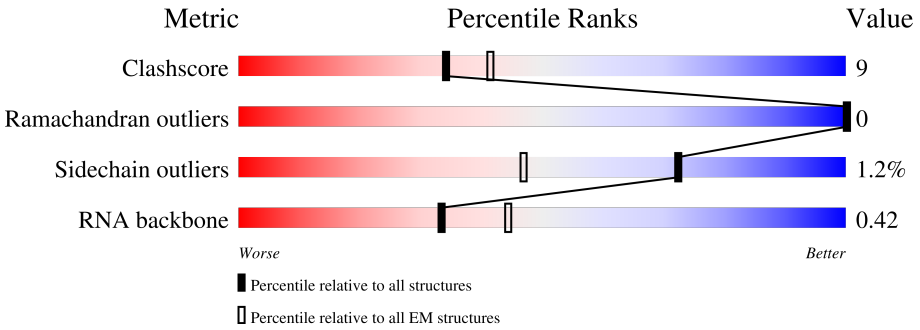
EMDB validation analysis	:	0.0.0.dev97
Mogul	:	1.8.4, CSD as541be (2020)
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.26

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 3.90 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	H	17	18% 35% 12% 35%
2	T	30	13% 47% 10% 30%
3	P	30	10% 27% • 60%
4	A	2276	69% 19% 12%

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 17326 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 RNA chain called RNA (5'-R(P*AP*CP*GP*AP*GP*UP*GP*UP*CP*GP *U)-3').

Mol	Chain	Residues	Atoms					AltConf	Trace
1	H	11	Total	C	N	O	P	0	0
			237	105	42	79	11		

- Molecule 2 is a RNA chain called RNA (5'-R(P*AP*AP*CP*GP*UP*UP*AP*UP*CP*UP *AP*UP*AP*AP*CP*AP*CP*UP*AP*CP*U)-3').

Mol	Chain	Residues	Atoms					AltConf	Trace
2	T	21	Total	C	N	O	P	0	0
			439	198	74	146	21		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
3	P	12	Total	C	N	O	P	0	0
			256	115	46	83	12		

- Molecule 4 is a protein called La Crosse virus polymerase.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	A	2007	Total	C	N	O	S	0	0
			16383	10499	2717	3059	108		

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	34	LYS	HIS	engineered mutation	UNP A5HC98
A	1028G	SER	-	insertion	UNP A5HC98
A	1028H	GLY	-	insertion	UNP A5HC98
A	1028I	TRP	-	insertion	UNP A5HC98
A	1028J	SER	-	insertion	UNP A5HC98
A	1028K	HIS	-	insertion	UNP A5HC98

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	1028L	PRO	-	insertion	UNP A5HC98
A	1028M	GLN	-	insertion	UNP A5HC98
A	1028N	PHE	-	insertion	UNP A5HC98
A	1028O	GLU	-	insertion	UNP A5HC98
A	1028P	LYS	-	insertion	UNP A5HC98
A	1028Q	GLY	-	insertion	UNP A5HC98
A	1028R	SER	-	insertion	UNP A5HC98
A	1028S	GLY	-	insertion	UNP A5HC98

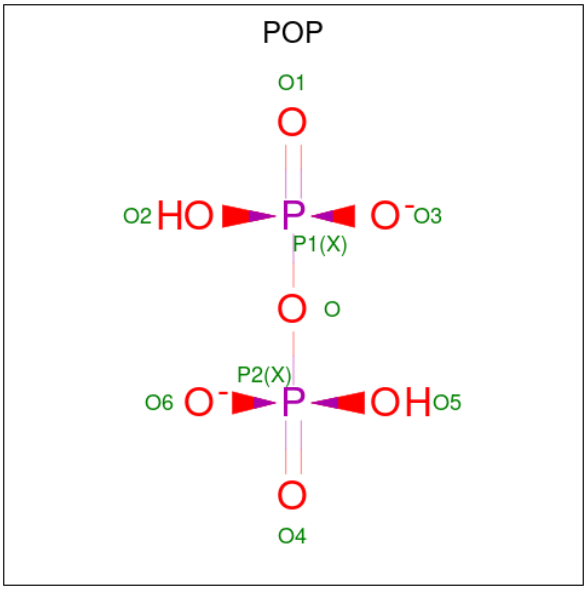
- Molecule 5 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	AltConf
5	A	1	Total Zn 1 1	0

- Molecule 6 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

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

- Molecule 7 is PYROPHOSPHATE 2- (three-letter code: POP) (formula: $H_2O_7P_2$).



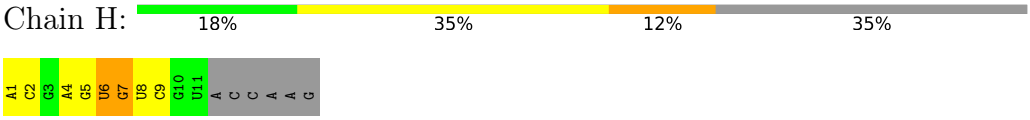
Mol	Chain	Residues	Atoms	AltConf
7	A	1	Total O P 9 7 2	0

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.

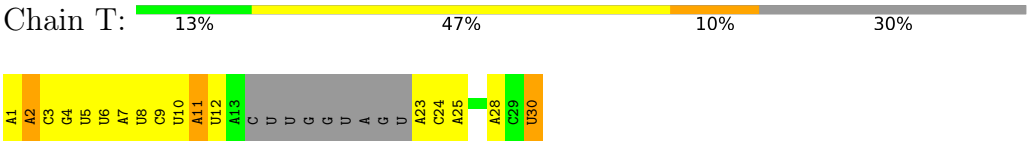
• Molecule 1: RNA (5'-R(P*AP*CP*GP*AP*GP*UP*GP*UP*CP*GP*U)-3')

Chain H:



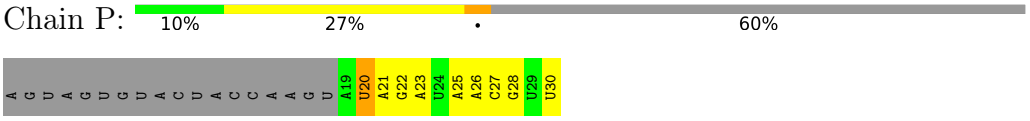
• Molecule 2: RNA (5'-R(P*AP*AP*CP*GP*UP*UP*AP*UP*CP*UP*AP*UP*AP*AP*CP*AP*CP*UP*AP*CP*U)-3')

Chain T:



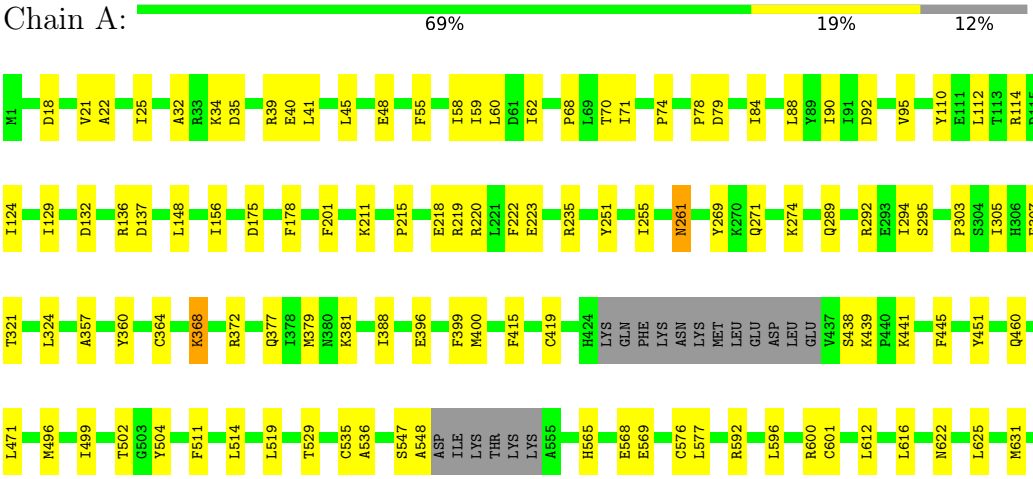
• Molecule 3: RNA (5'-R(P*AP*UP*AP*GP*AP*UP*AP*AP*CP*GP*UP*U)-3')

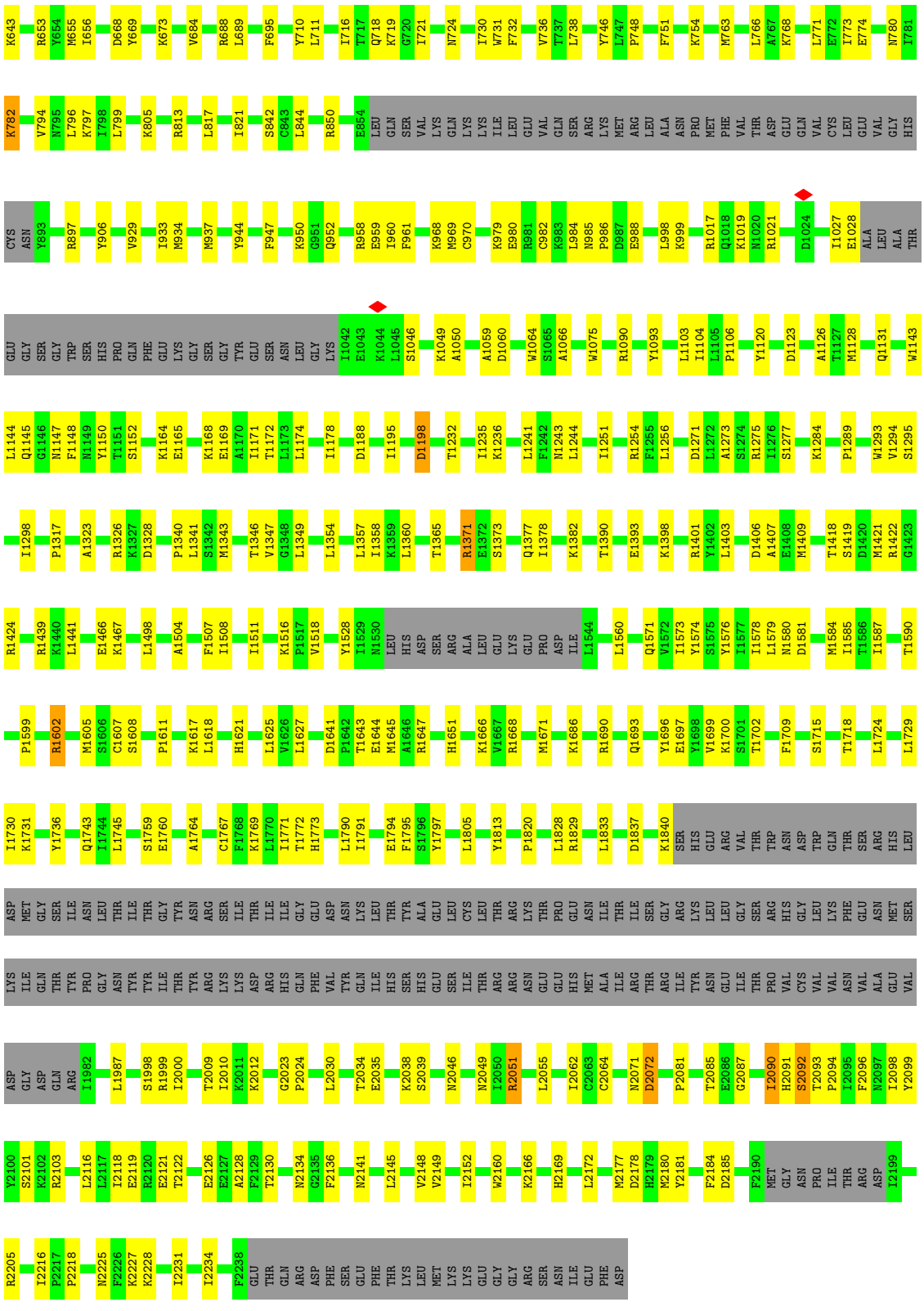
Chain P:



• Molecule 4: La Crosse virus polymerase

Chain A:





4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	24151	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	TFS GLACIOS	Depositor
Voltage (kV)	200	Depositor
Electron dose ($e^-/Å^2$)	60	Depositor
Minimum defocus (nm)	800	Depositor
Maximum defocus (nm)	2000	Depositor
Magnification	36000	Depositor
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor
Maximum map value	0.028	Depositor
Minimum map value	-0.012	Depositor
Average map value	0.000	Depositor
Map value standard deviation	0.001	Depositor
Recommended contour level	0.005	Depositor
Map size (Å)	297.69998, 297.69998, 297.69998	wwPDB
Map dimensions	260, 260, 260	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.145, 1.145, 1.145	Depositor

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: POP, 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	H	0.69	1/264 (0.4%)	0.67	0/408
2	T	0.22	0/488	0.79	0/753
3	P	0.23	0/286	0.72	0/443
4	A	0.27	0/16721	0.45	0/22547
All	All	0.28	1/17759 (0.0%)	0.48	0/24151

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	H	1	A	OP3-P	-10.66	1.48	1.61

There are no bond angle outliers.

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	H	237	0	119	4	0
2	T	439	0	226	14	0
3	P	256	0	129	10	0
4	A	16383	0	16426	289	0
5	A	1	0	0	0	0
6	A	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	A	9	0	0	0	0
All	All	17326	0	16900	305	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 9.

The worst 5 of 305 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)
4:A:813:ARG:NH1	4:A:986:PRO:HD3	1.67	1.09
4:A:2064:CYS:HB2	4:A:2169:HIS:HE1	1.15	1.03
4:A:2064:CYS:CB	4:A:2169:HIS:HE1	1.76	0.98
4:A:2064:CYS:HB2	4:A:2169:HIS:CE1	1.97	0.98
4:A:813:ARG:HH12	4:A:986:PRO:HD3	1.35	0.89

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	1991/2276 (88%)	1923 (97%)	68 (3%)	0	100	100

There are no Ramachandran outliers to report.

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	1852/2097 (88%)	1829 (99%)	23 (1%)	71 83

5 of 23 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	A	1390	THR
4	A	2051	ARG
4	A	1602	ARG
4	A	2072	ASP
4	A	655	MET

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
4	A	622	ASN
4	A	760	HIS
4	A	1580	ASN
4	A	1704	HIS
4	A	1743	GLN

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	H	10/17 (58%)	4 (40%)	0
2	T	19/30 (63%)	9 (47%)	0
3	P	11/30 (36%)	1 (9%)	0
All	All	40/77 (51%)	14 (35%)	0

5 of 14 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	H	6	U
1	H	7	G
1	H	8	U
1	H	9	C
2	T	2	A

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, 2 are monoatomic - leaving 1 for Mogul analysis.

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	$\#\|Z\| > 2$
7	POP	A	2303	-	6,8,8	0.80	0	13,13,13	0.74	0

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
7	POP	A	2303	-	-	0/6/6/6	-

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.

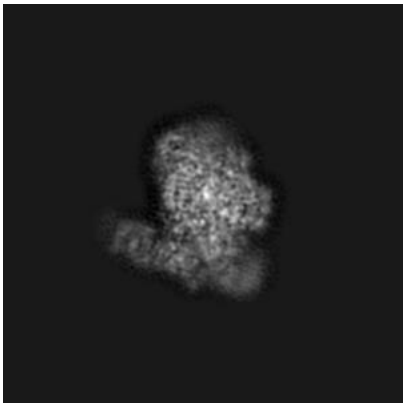
6 Map visualisation

This section contains visualisations of the EMDB entry EMD-13038. These allow visual inspection of the internal detail of the map and identification of artifacts.

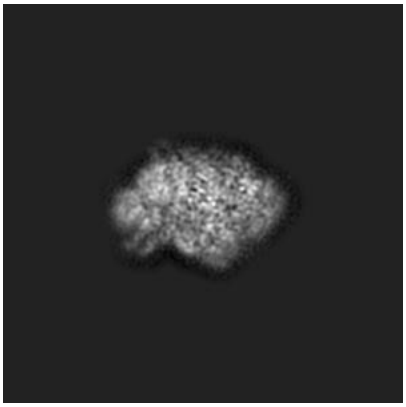
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections

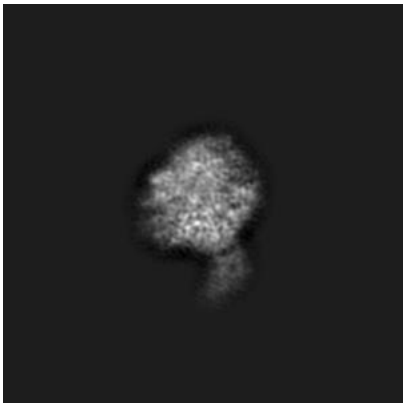
6.1.1 Primary map



X



Y

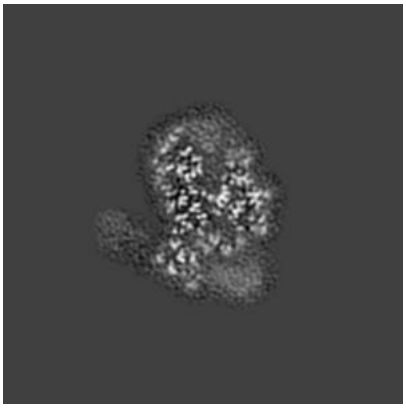


Z

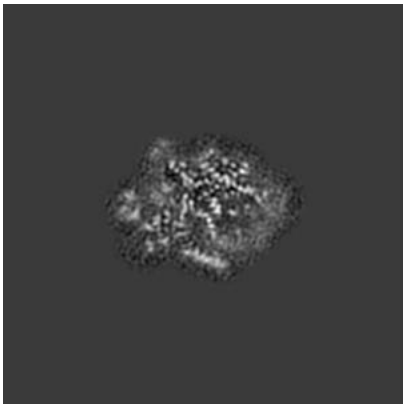
The images above show the map projected in three orthogonal directions.

6.2 Central slices

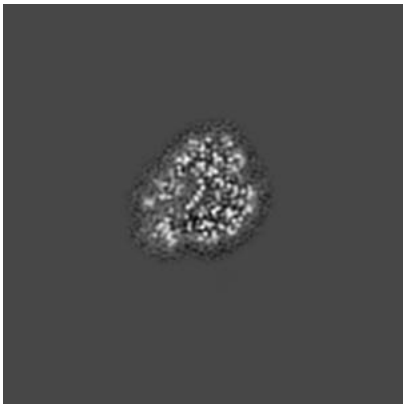
6.2.1 Primary map



X Index: 130



Y Index: 130

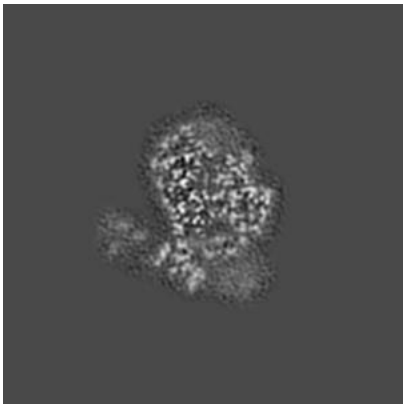


Z Index: 130

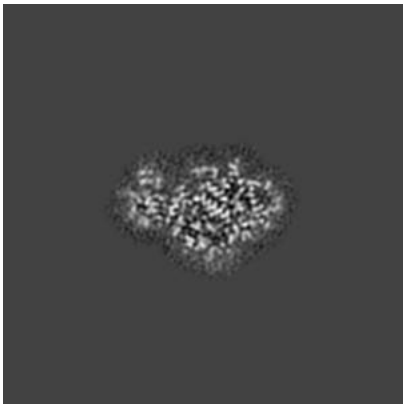
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

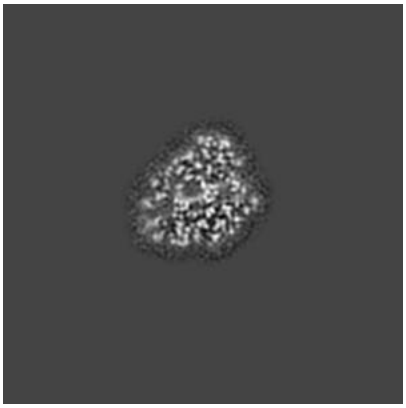
6.3.1 Primary map



X Index: 133



Y Index: 116



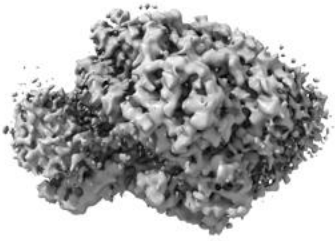
Z Index: 134

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

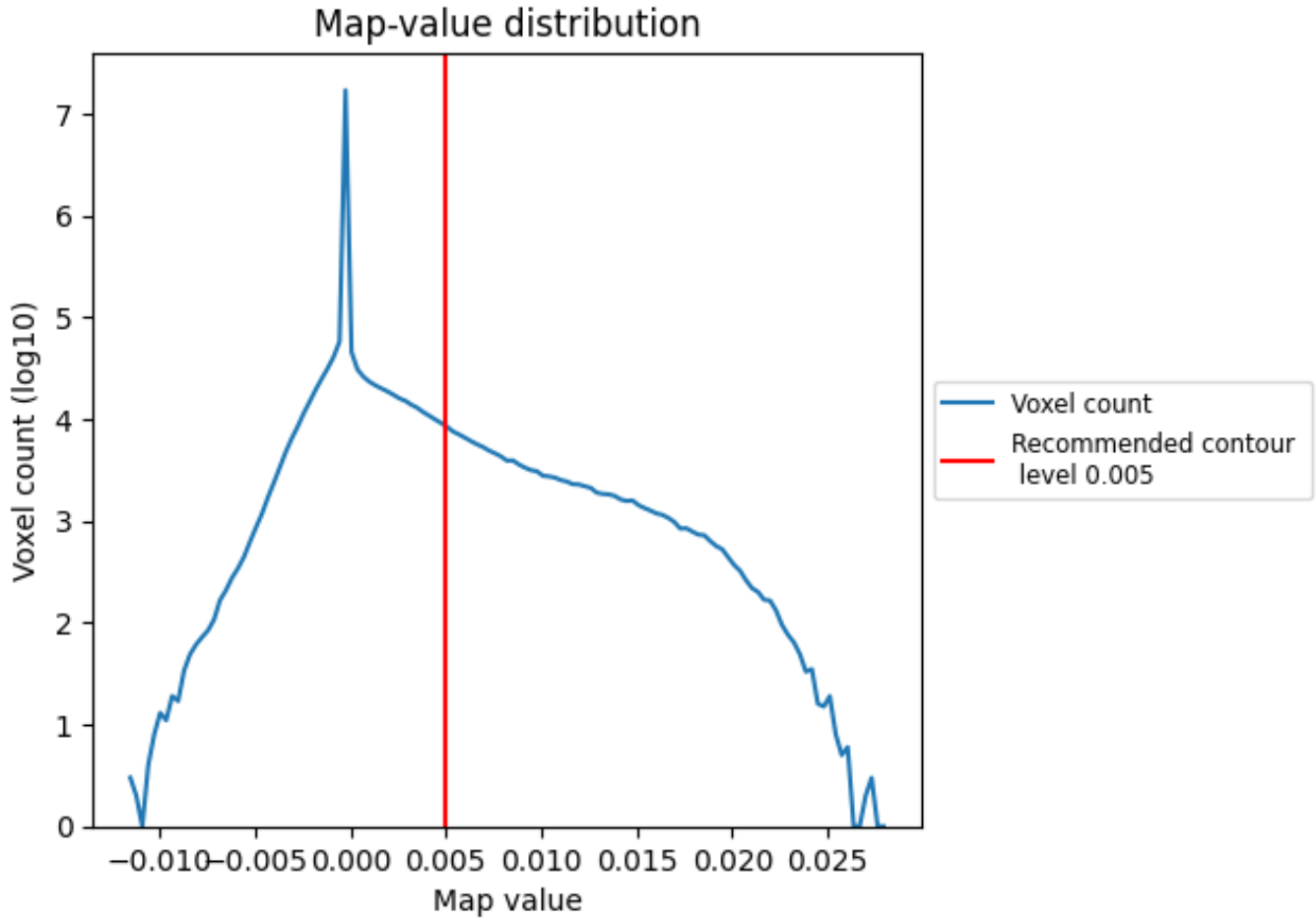
6.5 Mask visualisation (i)

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

7 Map analysis

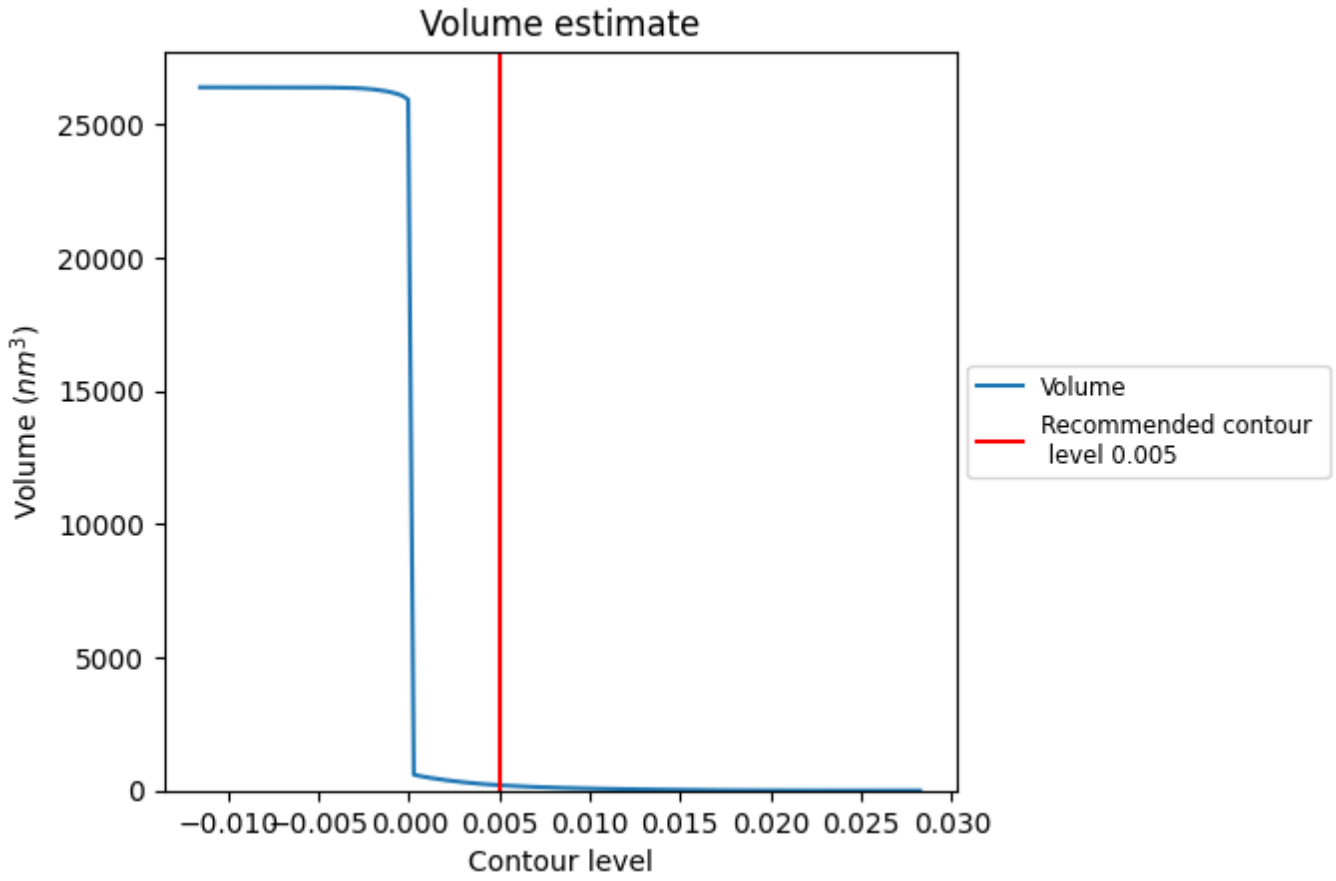
This section contains the results of statistical analysis of the map.

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.

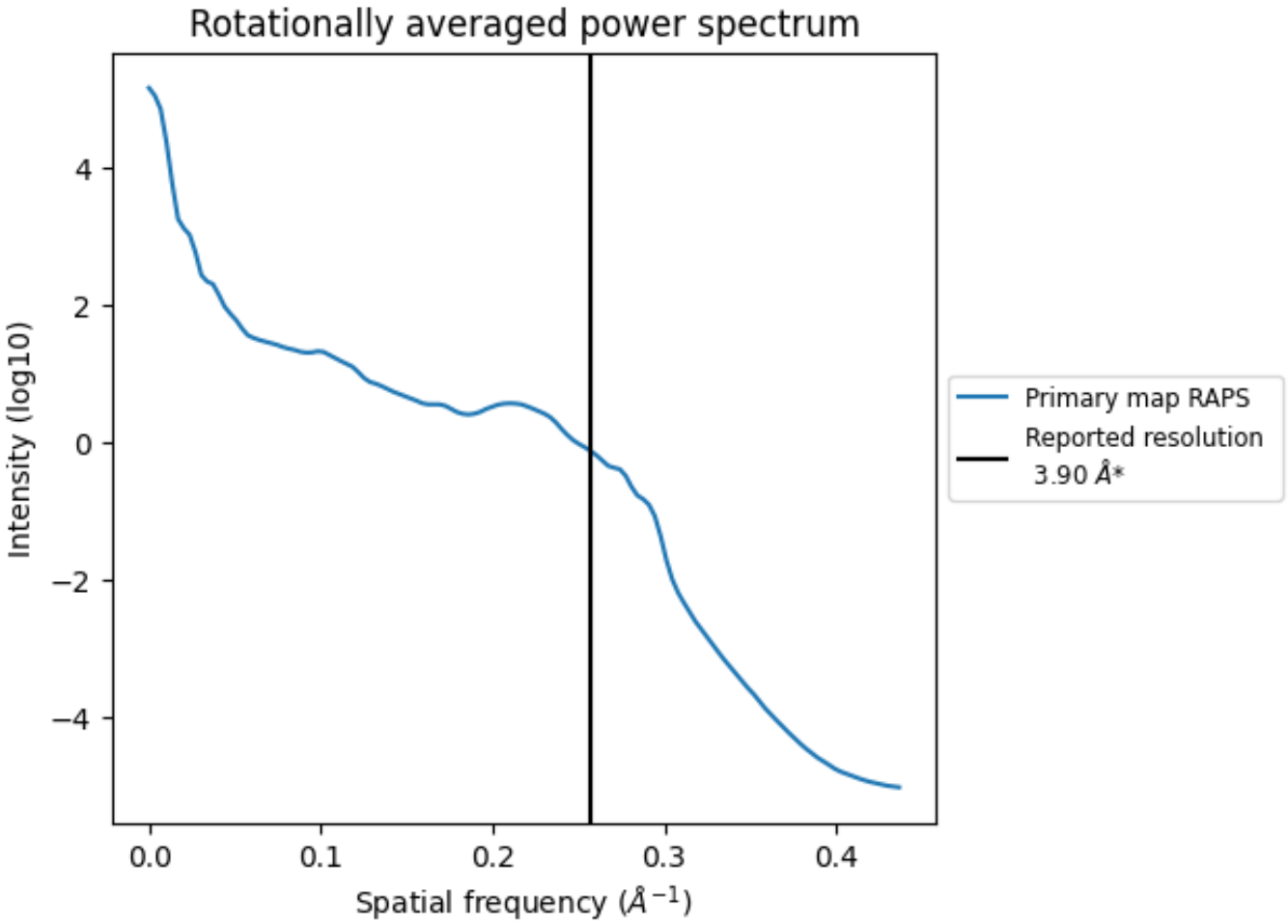
7.2 Volume estimate (i)



The volume at the recommended contour level is 201 nm^3; this corresponds to an approximate mass of 182 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)

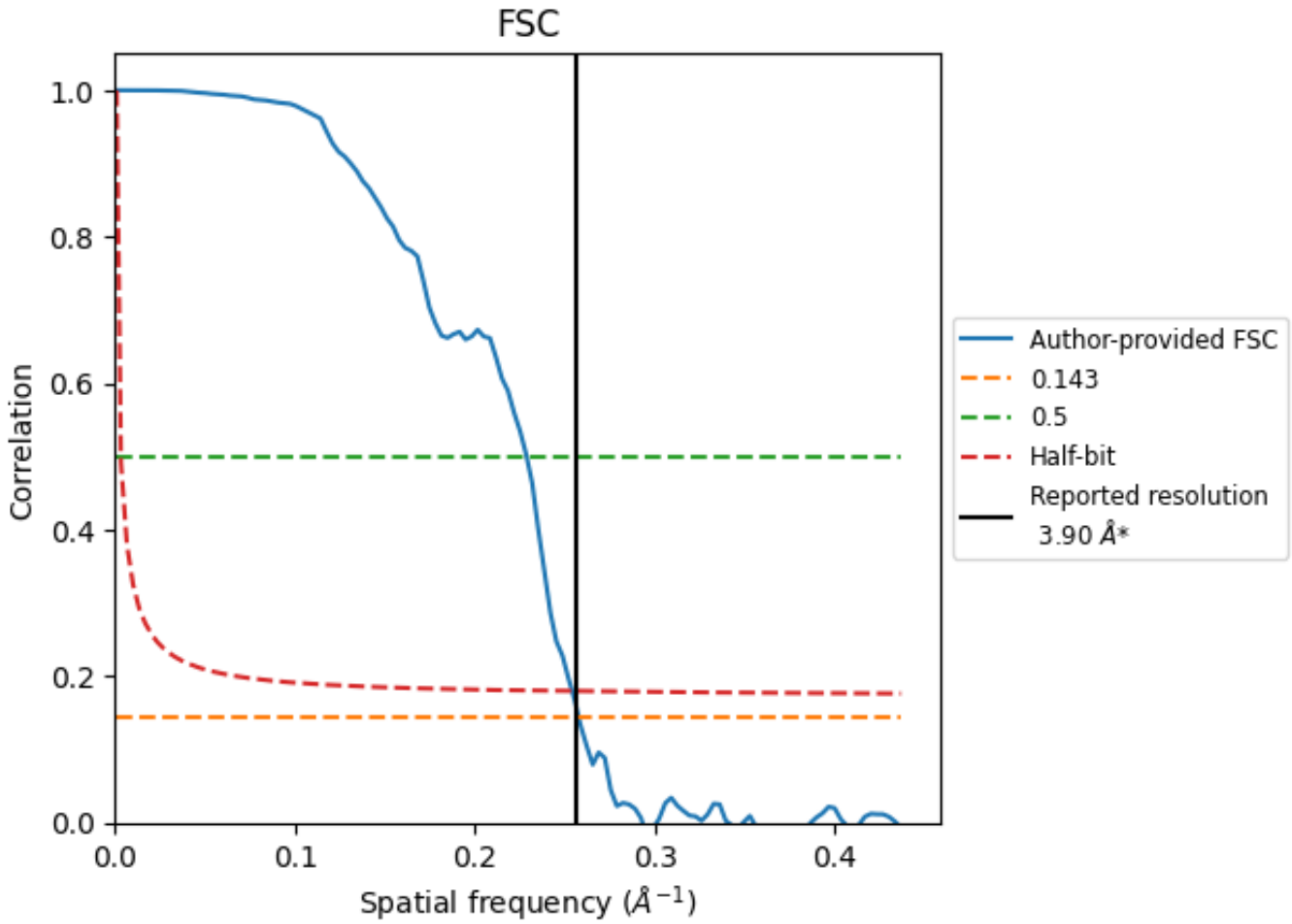


*Reported resolution corresponds to spatial frequency of 0.256 $Å^{-1}$

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 0.256 Å^{-1}

8.2 Resolution estimates (i)

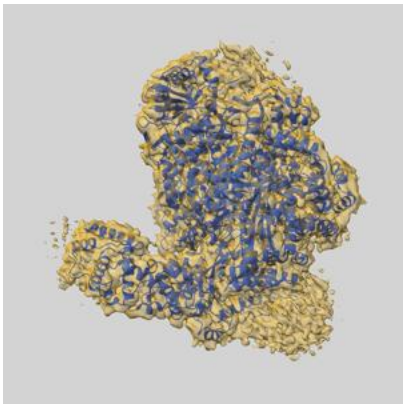
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	3.90	-	-
Author-provided FSC curve	3.88	4.37	3.94
Unmasked-calculated*	-	-	-

*Resolution estimate based on FSC curve calculated by comparison of deposited half-maps.

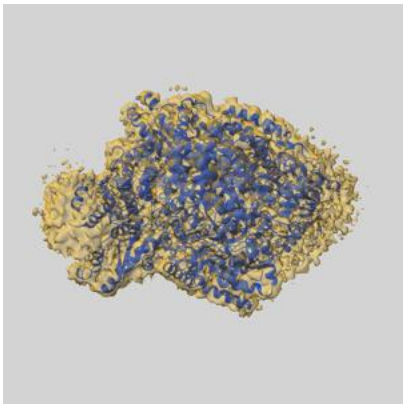
9 Map-model fit

This section contains information regarding the fit between EMDB map EMD-13038 and PDB model 7ORI. Per-residue inclusion information can be found in section 3 on page 5.

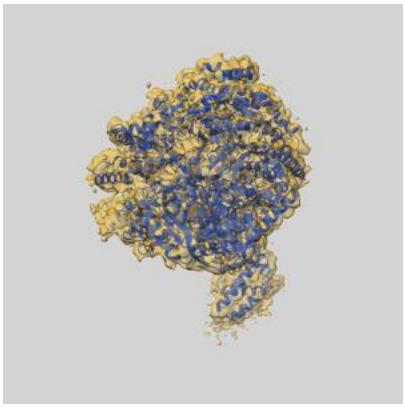
9.1 Map-model overlay



X



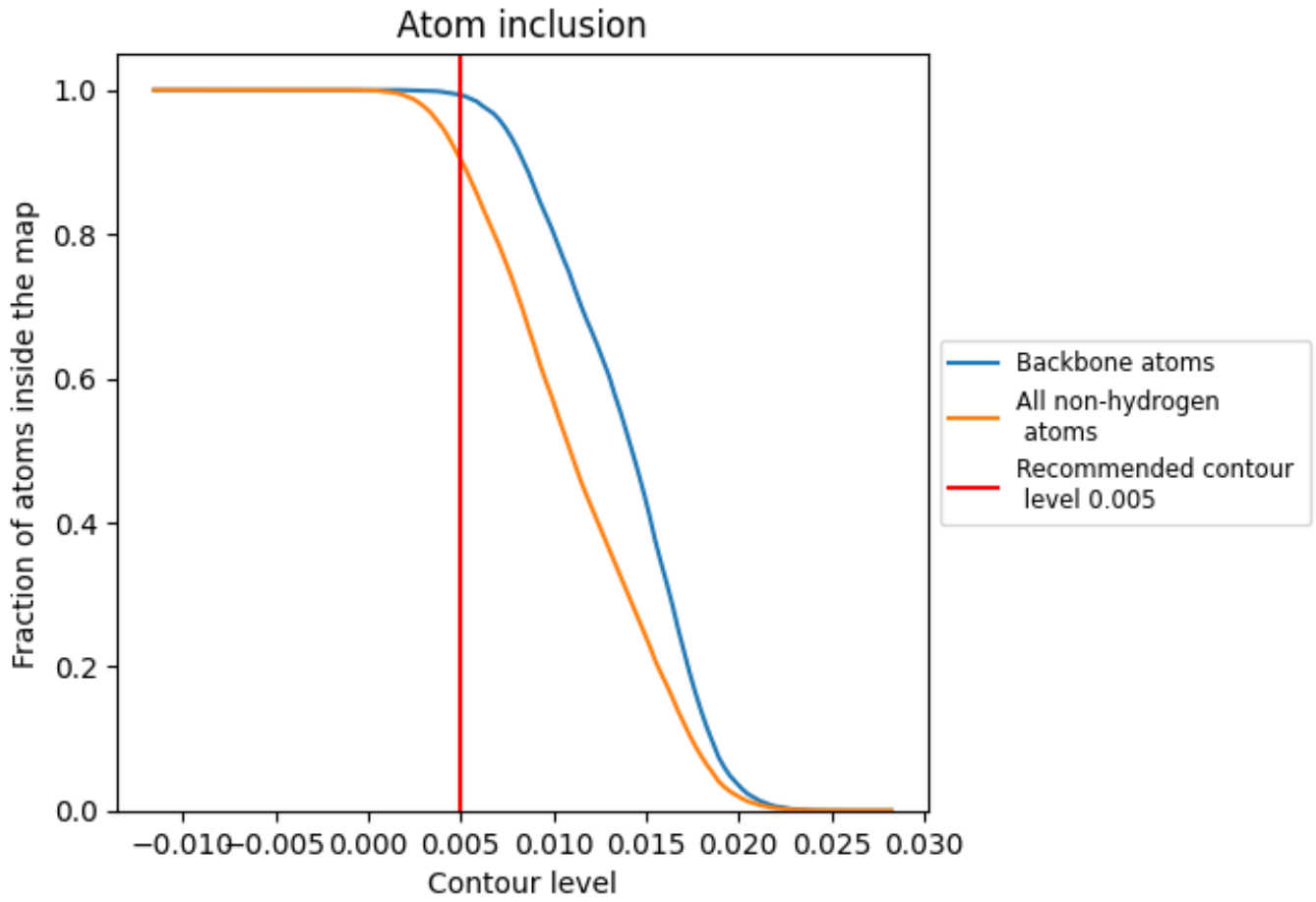
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.005 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 Atom inclusion (i)



At the recommended contour level, 99% of all backbone atoms, 90% of all non-hydrogen atoms, are inside the map.