wwPDB EM Validation Summary Report (i)

Dec 12, 2022 – 06:08 am GMT

PDB ID	:	6XZD
EMDB ID	:	EMD-10659
Title	:	Influenza C virus polymerase complex without chicken ANP32A - Subclass 2
Authors	:	Keown, J.R.; Carrique, L.; Fan, H.; Grimes, J.M.; Fodor, E.
Deposited on	:	2020-02-04
Resolution	:	3.40 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:

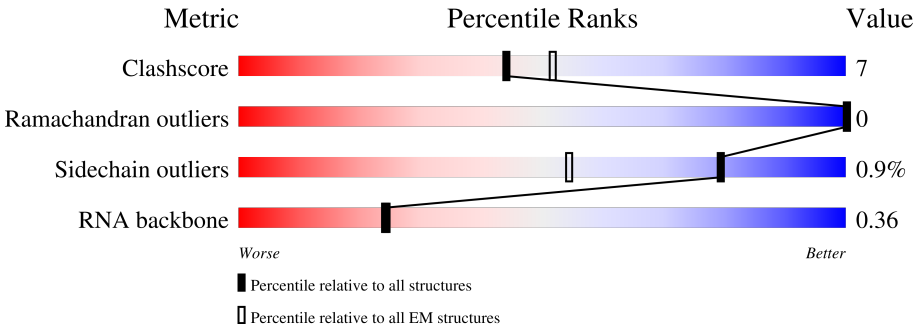
EMDB validation analysis	:	0.0.1.dev43
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.9
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.31.3

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 3.40 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	IN1	47	17% 17% 6% 57%
2	AP1	709	87% 11%
2	DP1	709	62% 12% 26%
3	BP1	754	82% 12% 6%
3	EP1	754	63% 18% 18%
4	CP1	774	84% 16%
4	FP1	774	9% 60% 17% 23%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 63749 atoms, of which 31961 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(*AP*GP*UP*AP*GP*AP*AP*AP*CP*AP*AP*GP*GP*GP*CP*CP*CP*UP*GP*C)-3').

Mol	Chain	Residues	Atoms						AltConf	Trace
1	IN1	20	Total	C	H	N	O	P	0	0
			650	193	221	84	133	19		

- Molecule 2 is a protein called Polymerase acidic protein.

Mol	Chain	Residues	Atoms						AltConf	Trace
2	AP1	696	Total	C	H	N	O	S	0	0
			11291	3599	5645	955	1049	43		
2	DP1	525	Total	C	H	N	O	S	0	0
			8558	2710	4306	729	779	34		

- Molecule 3 is a protein called RNA-directed RNA polymerase catalytic subunit.

Mol	Chain	Residues	Atoms						AltConf	Trace
3	BP1	712	Total	C	H	N	O	S	0	0
			11444	3602	5761	960	1068	53		
3	EP1	615	Total	C	H	N	O	S	0	0
			9855	3116	4961	822	911	45		

- Molecule 4 is a protein called Polymerase basic protein 2.

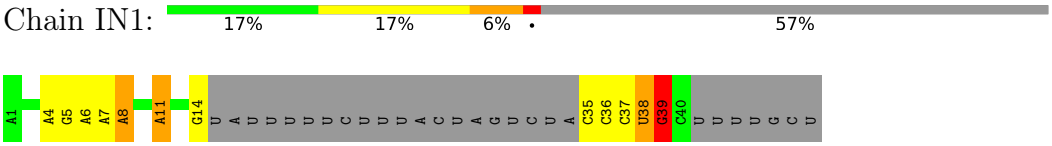
Mol	Chain	Residues	Atoms						AltConf	Trace
4	CP1	772	Total	C	H	N	O	S	0	0
			12409	3888	6261	1080	1142	38		
4	FP1	599	Total	C	H	N	O	S	0	0
			9542	3004	4806	820	884	28		

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.

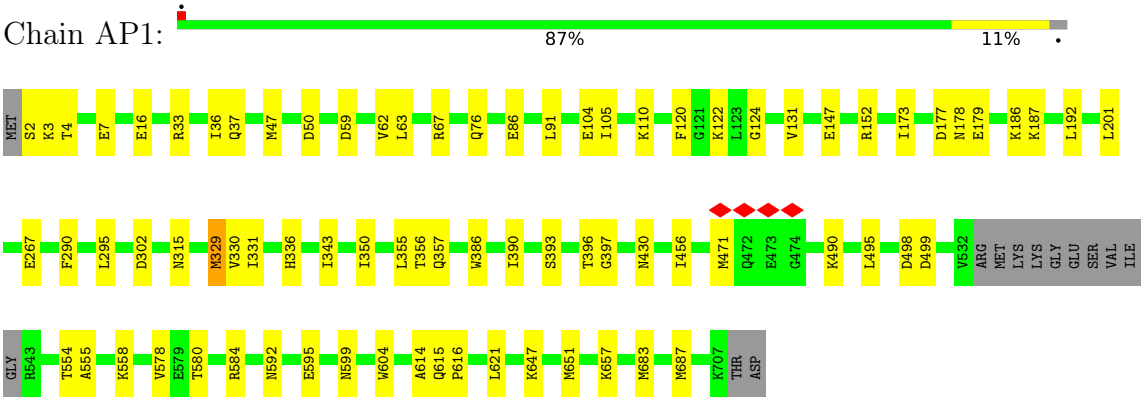
• Molecule 1: RNA (5'-R(*AP*GP*UP*AP*GP*AP*AP*AP*CP*AP*AP*GP*GP*GP*CP*CP *CP*UP*GP*C)-3')

Chain IN1:



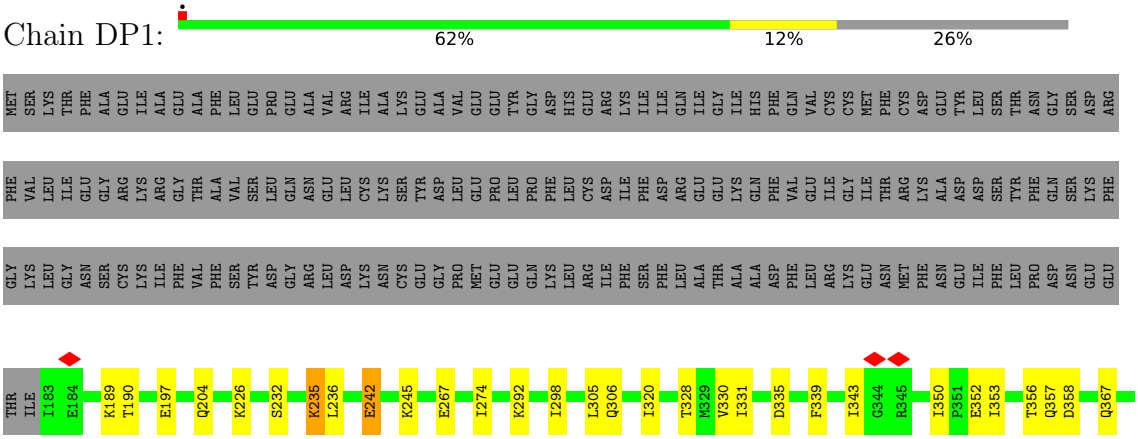
• Molecule 2: Polymerase acidic protein

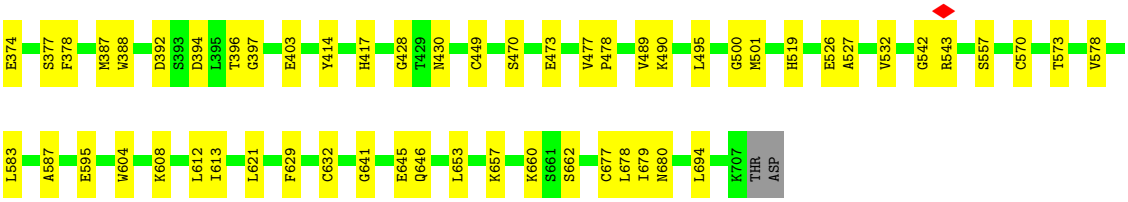
Chain AP1:



• Molecule 2: Polymerase acidic protein

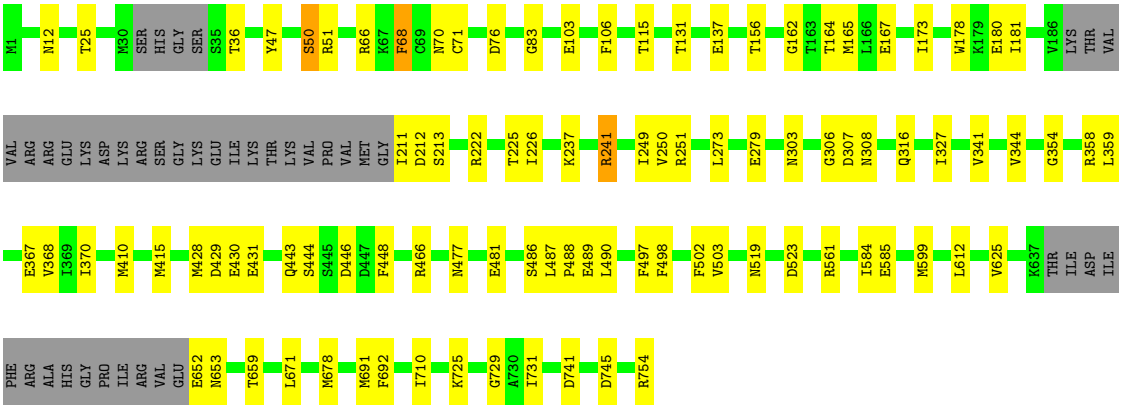
Chain DP1:





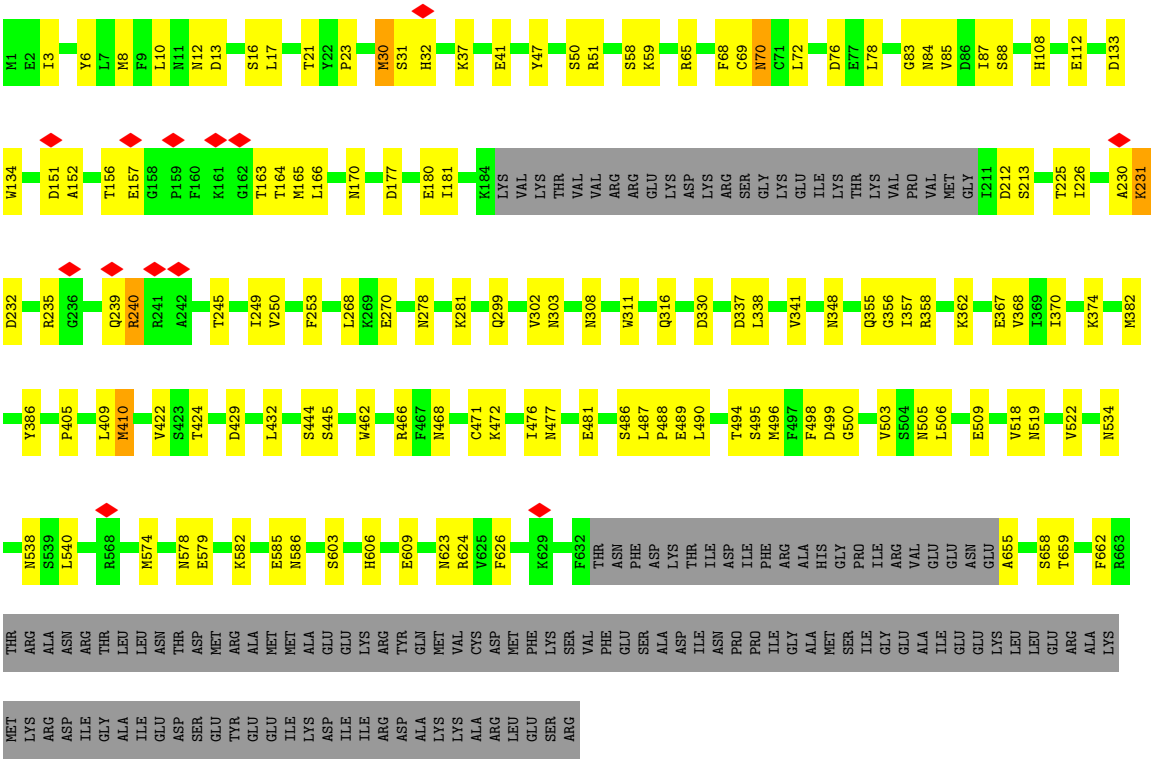
• Molecule 3: RNA-directed RNA polymerase catalytic subunit

Chain BP1:



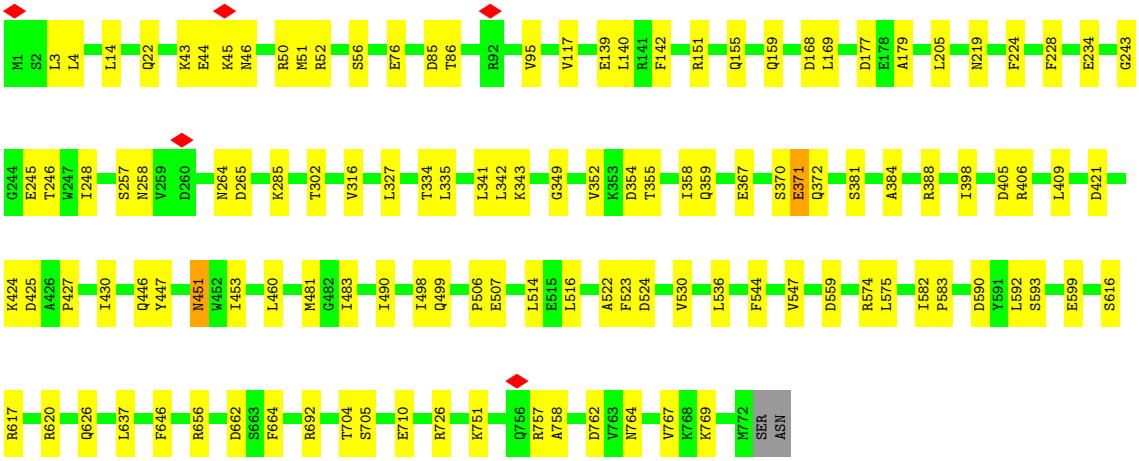
• Molecule 3: RNA-directed RNA polymerase catalytic subunit

Chain EP1:



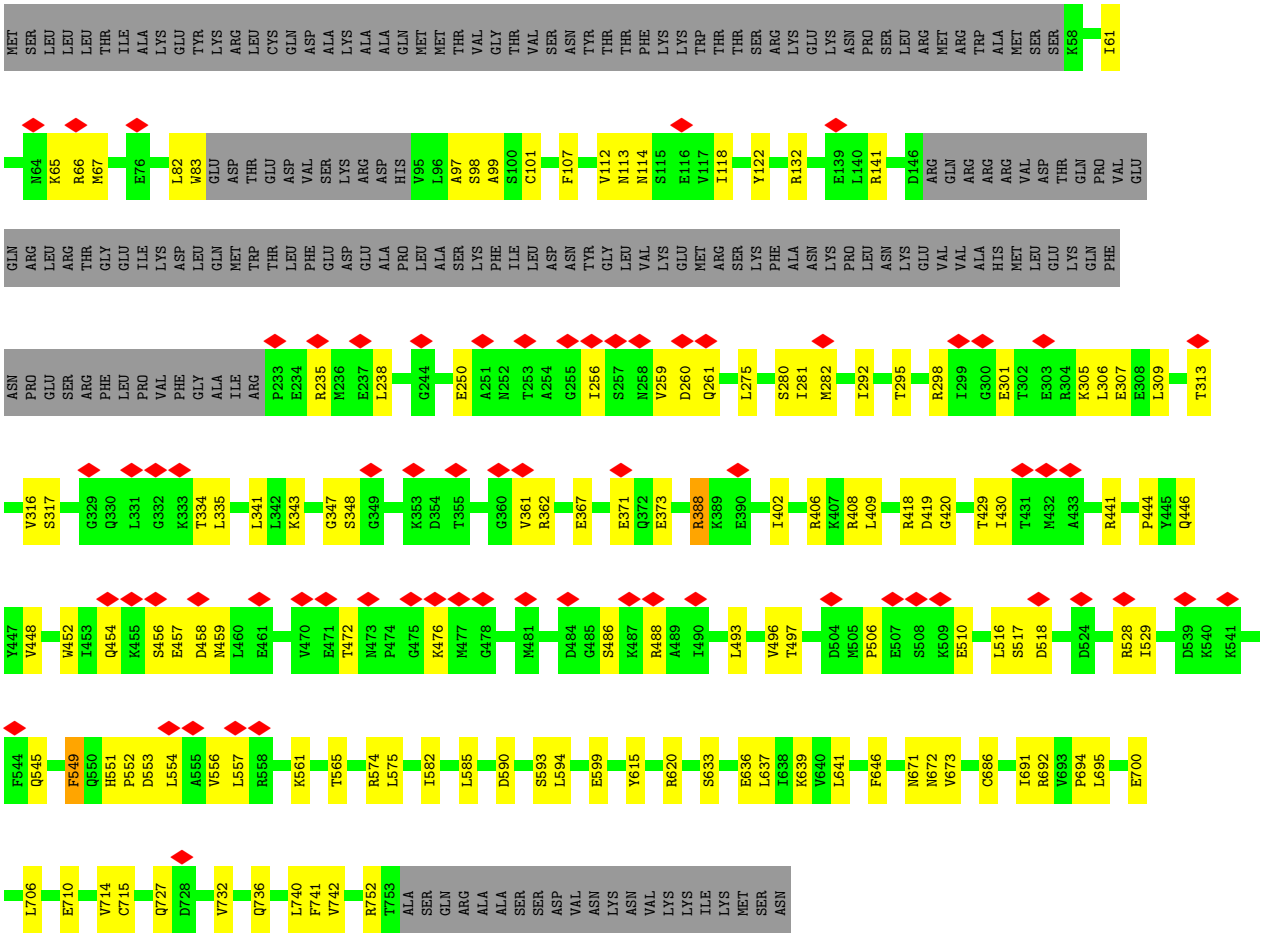
• Molecule 4: Polymerase basic protein 2

Chain CP1:



• Molecule 4: Polymerase basic protein 2

Chain FP1:



4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	169000	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 ($e^-/Å^2$)	38.8	Depositor
Minimum defocus (nm)	2000	Depositor
Maximum defocus (nm)	3500	Depositor
Magnification	105000	Depositor
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor
Maximum map value	0.790	Depositor
Minimum map value	-0.275	Depositor
Average map value	0.001	Depositor
Map value standard deviation	0.020	Depositor
Recommended contour level	0.13	Depositor
Map size (Å)	383.6, 383.6, 383.6	wwPDB
Map dimensions	280, 280, 280	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.37, 1.37, 1.37	Depositor

5 Model quality

5.1 Standard geometry

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	IN1	0.95	0/480	1.12	3/745 (0.4%)
2	AP1	0.68	0/5764	0.59	0/7746
2	DP1	0.65	0/4342	0.56	0/5835
3	BP1	0.69	0/5780	0.63	0/7762
3	EP1	0.43	0/4986	0.51	0/6707
4	CP1	0.65	0/6259	0.60	0/8425
4	FP1	0.31	0/4821	0.46	0/6493
All	All	0.60	0/32432	0.58	3/43713 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	IN1	39	G	O4'-C1'-N9	6.47	113.38	108.20
1	IN1	5	G	O4'-C1'-N9	5.29	112.43	108.20
1	IN1	39	G	C4-N9-C1'	5.04	133.05	126.50

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	IN1	429	221	221	7	0
2	AP1	5646	5645	5645	55	0
2	DP1	4252	4306	4305	71	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	BP1	5683	5761	5759	63	0
3	EP1	4894	4961	4959	106	0
4	CP1	6148	6261	6261	83	0
4	FP1	4736	4806	4806	94	0
All	All	31788	31961	31956	439	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 7.

The worst 5 of 439 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:BP1:303:ASN:ND2	3:BP1:488:PRO:O	1.89	1.05
4:FP1:575:LEU:HD12	4:FP1:582:ILE:HD13	1.45	0.97
2:DP1:449:CYS:SG	2:DP1:490:LYS:NZ	2.45	0.89
3:BP1:237:LYS:NZ	3:BP1:307:ASP:OD1	2.06	0.88
4:CP1:44:GLU:OE1	4:CP1:50:ARG:NH2	2.06	0.88

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	
2	AP1	692/709 (98%)	661 (96%)	31 (4%)	0	100	100
2	DP1	523/709 (74%)	505 (97%)	18 (3%)	0	100	100
3	BP1	704/754 (93%)	661 (94%)	43 (6%)	0	100	100
3	EP1	609/754 (81%)	566 (93%)	43 (7%)	0	100	100
4	CP1	770/774 (100%)	729 (95%)	41 (5%)	0	100	100
4	FP1	593/774 (77%)	552 (93%)	41 (7%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	3891/4474 (87%)	3674 (94%)	217 (6%)	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
2	AP1	620/631 (98%)	617 (100%)	3 (0%)	88 94
2	DP1	468/631 (74%)	463 (99%)	5 (1%)	73 86
3	BP1	632/669 (94%)	626 (99%)	6 (1%)	78 90
3	EP1	547/669 (82%)	539 (98%)	8 (2%)	65 82
4	CP1	677/679 (100%)	671 (99%)	6 (1%)	78 90
4	FP1	521/679 (77%)	517 (99%)	4 (1%)	81 91
All	All	3465/3958 (88%)	3433 (99%)	32 (1%)	79 90

5 of 32 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	FP1	118	ILE
4	FP1	261	GLN
4	CP1	371	GLU
4	CP1	52	ARG
4	FP1	388	ARG

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

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	IN1	18/47 (38%)	8 (44%)	1 (5%)

5 of 8 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	IN1	6	A
1	IN1	8	A
1	IN1	11	A
1	IN1	14	G
1	IN1	36	C

All (1) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
1	IN1	38	U

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)

There are no ligands in this entry.

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.

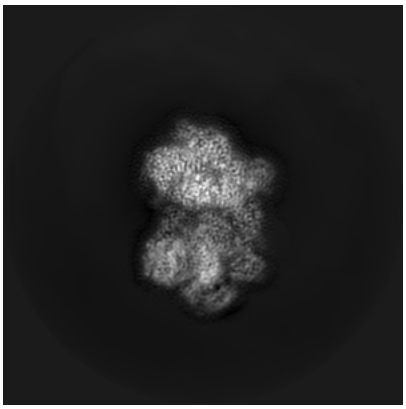
6 Map visualisation

This section contains visualisations of the EMDB entry EMD-10659. These allow visual inspection of the internal detail of the map and identification of artifacts.

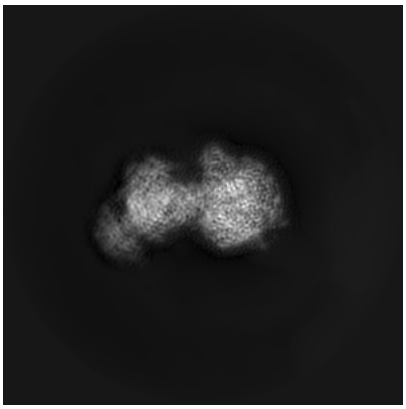
Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

6.1 Orthogonal projections

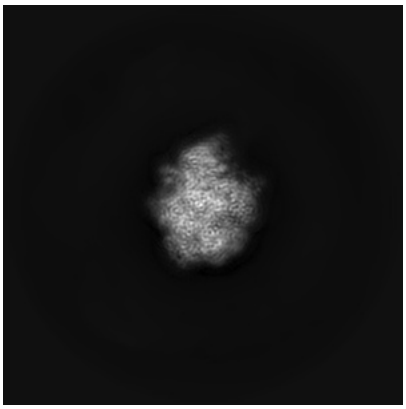
6.1.1 Primary map



X

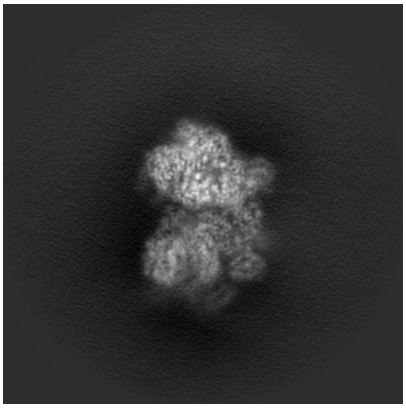


Y

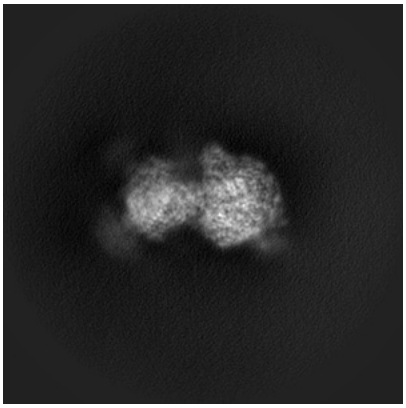


Z

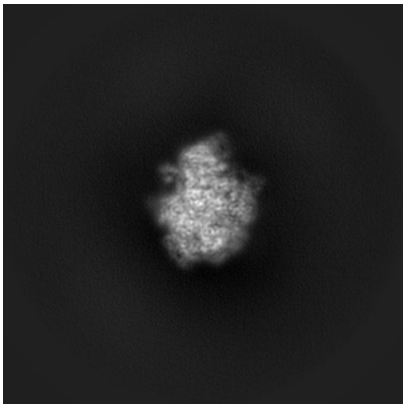
6.1.2 Raw map



X



Y

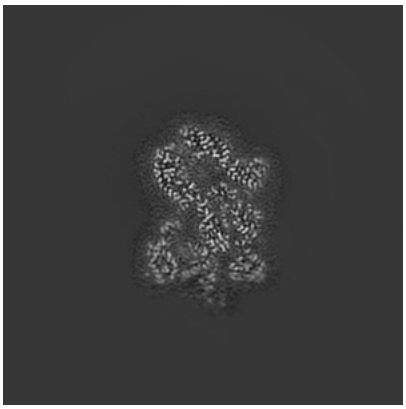


Z

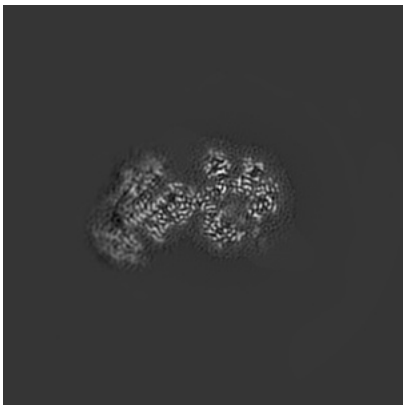
The images above show the map projected in three orthogonal directions.

6.2 Central slices (i)

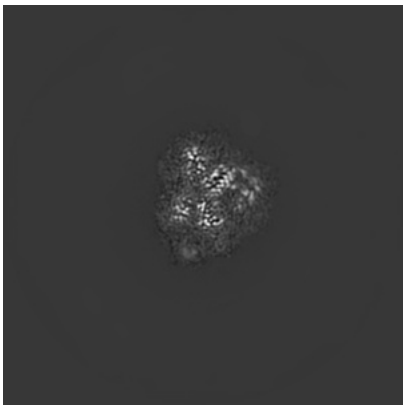
6.2.1 Primary map



X Index: 140

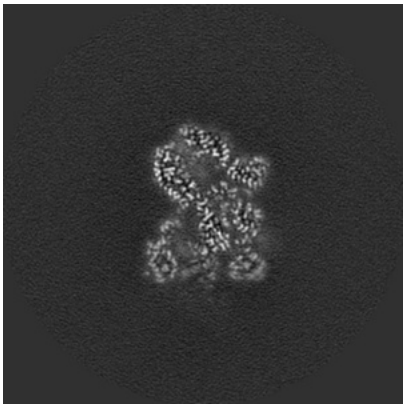


Y Index: 140

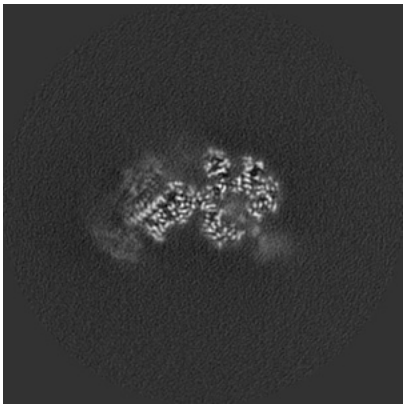


Z Index: 140

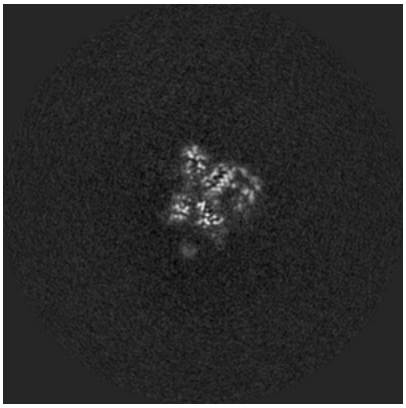
6.2.2 Raw map



X Index: 140



Y Index: 140

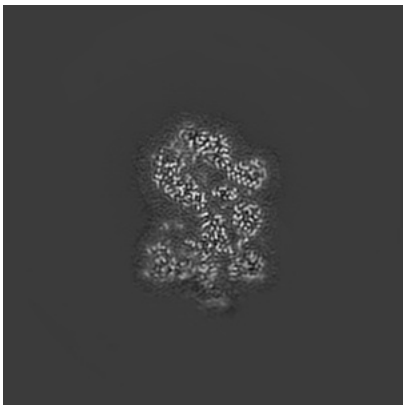


Z Index: 140

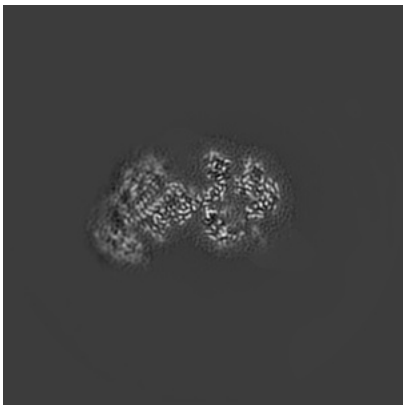
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

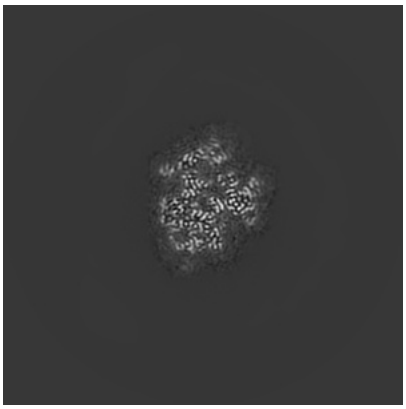
6.3.1 Primary map



X Index: 138

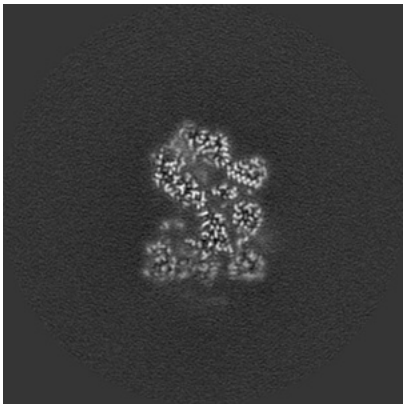


Y Index: 141

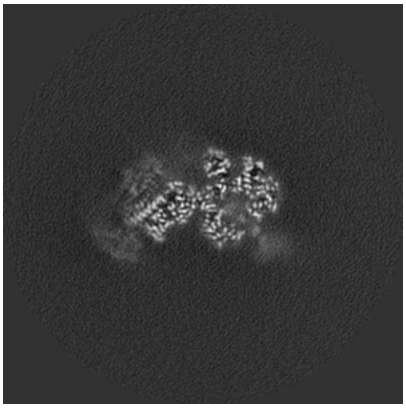


Z Index: 152

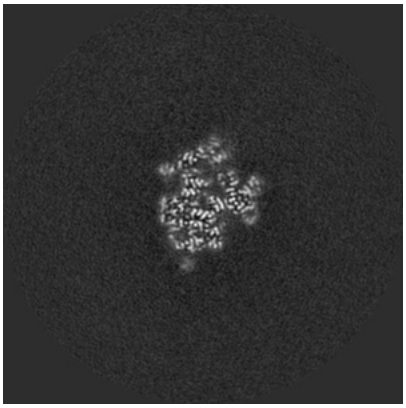
6.3.2 Raw map



X Index: 138



Y Index: 140



Z Index: 152

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 0.13. 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.

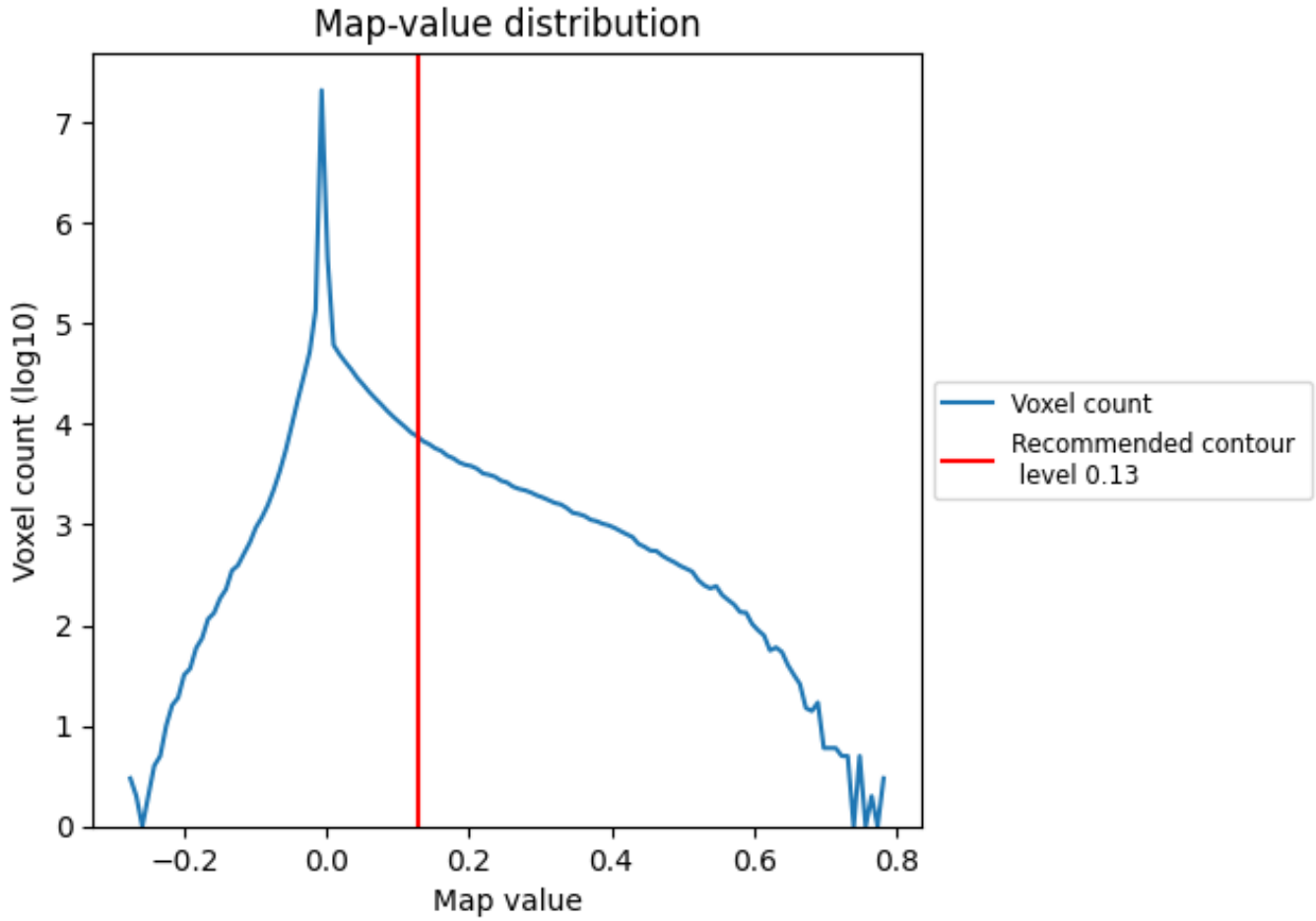
6.5 Mask visualisation (i)

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

7 Map analysis

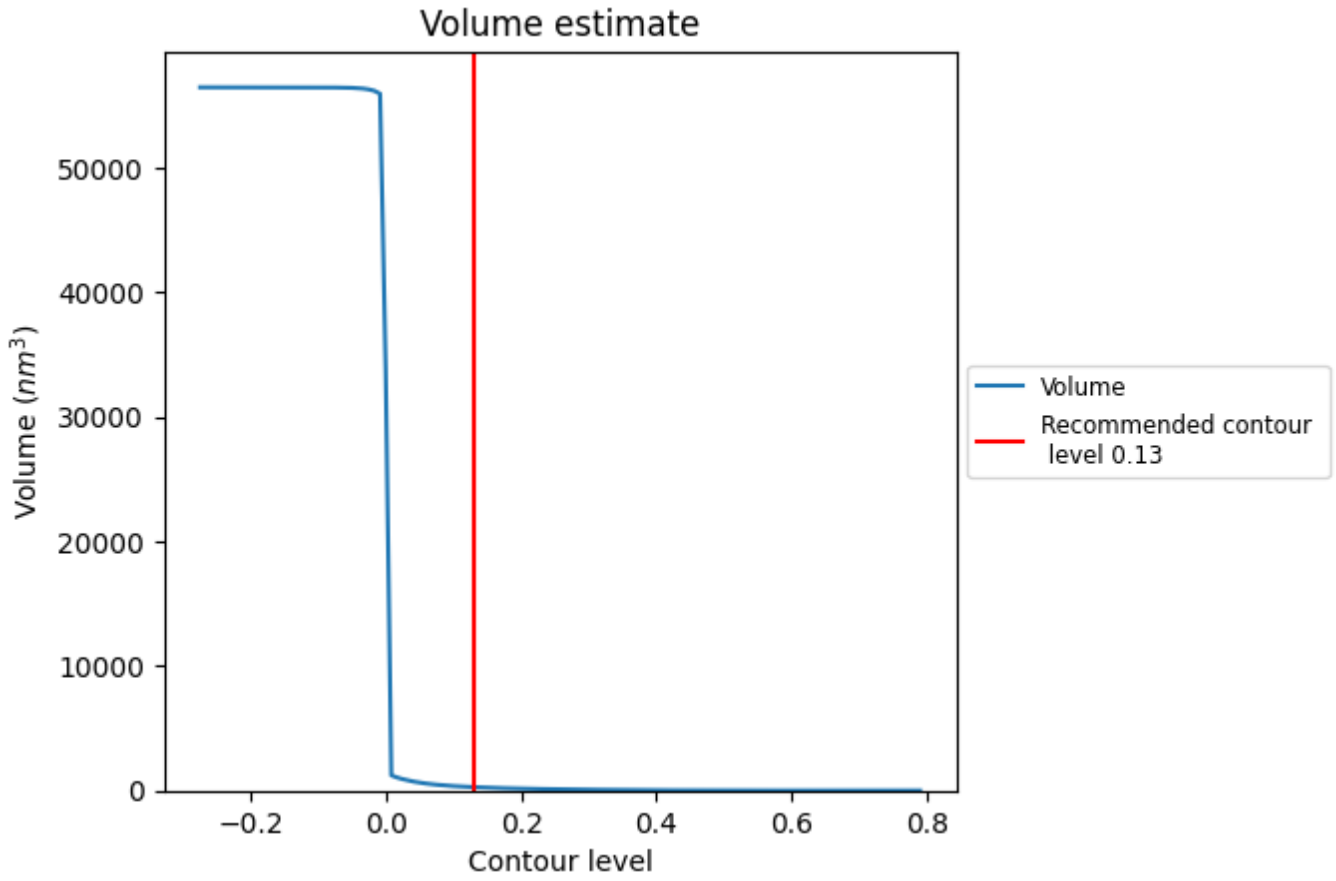
This section contains the results of statistical analysis of the map.

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.

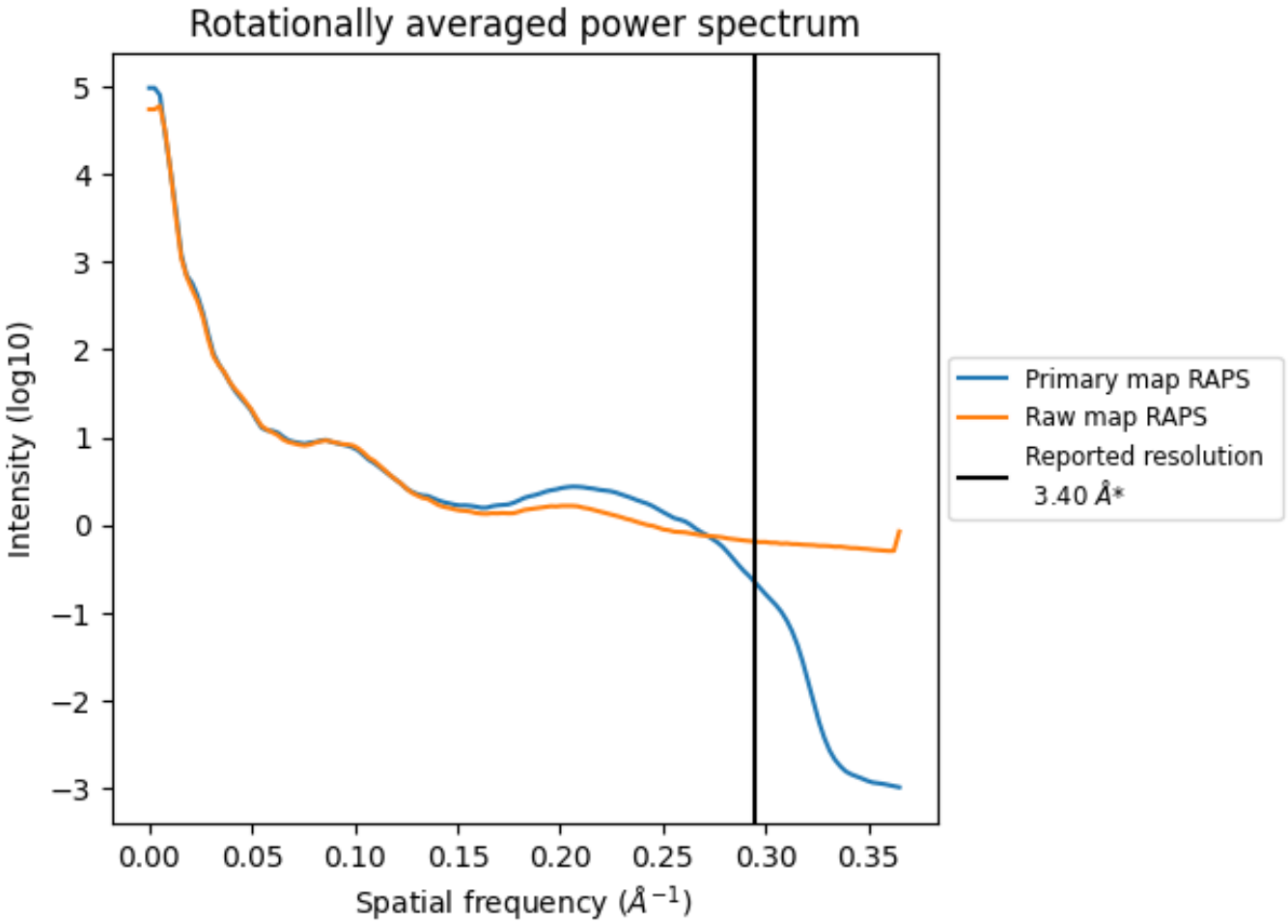
7.2 Volume estimate (i)



The volume at the recommended contour level is 275 nm^3; this corresponds to an approximate mass of 248 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)

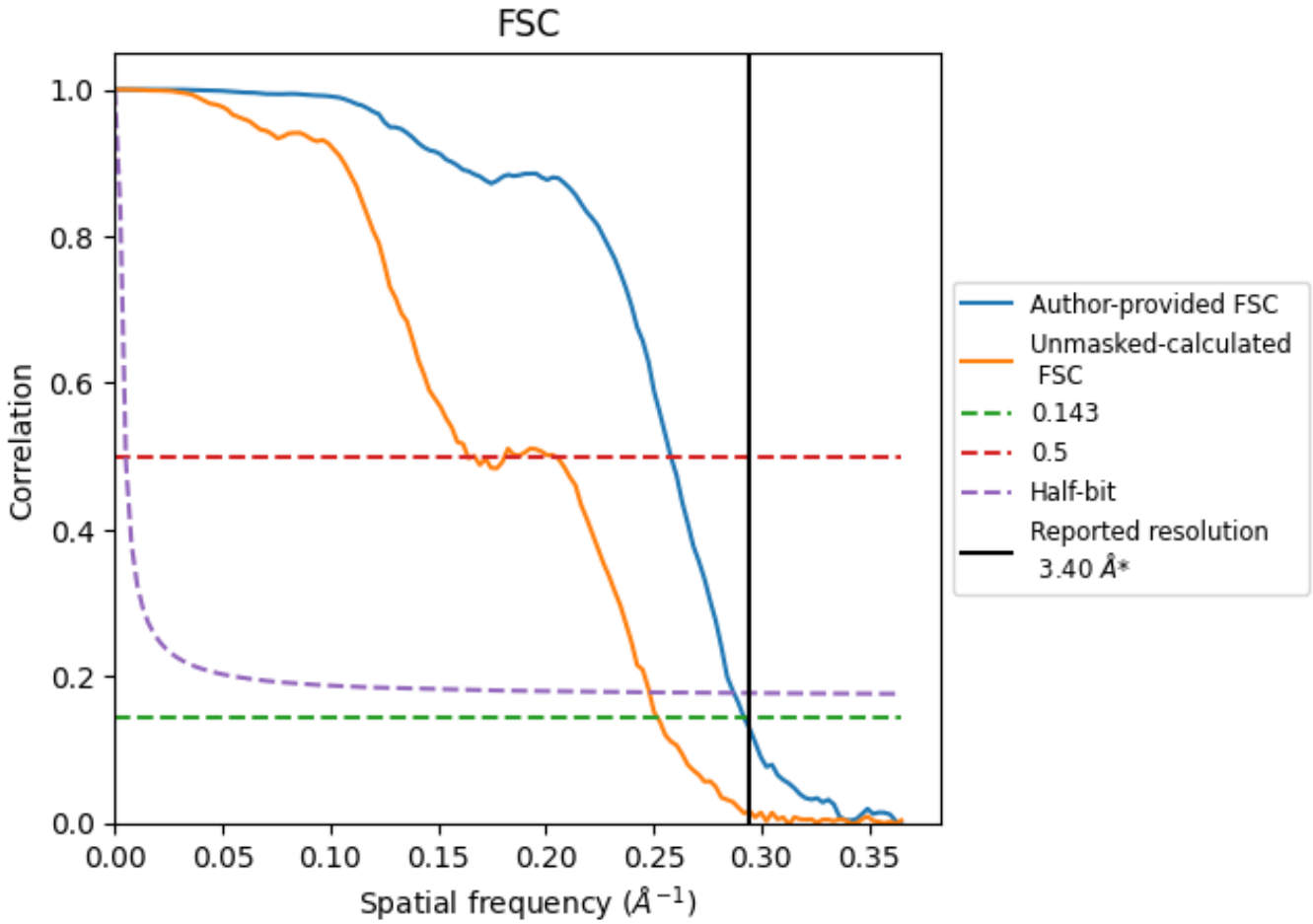


*Reported resolution corresponds to spatial frequency of 0.294 Å^{-1}

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 0.294 Å^{-1}

8.2 Resolution estimates (i)

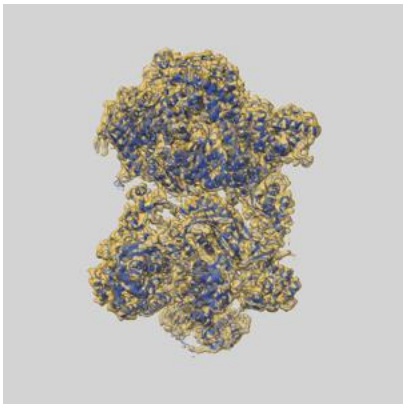
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	3.40	-	-
Author-provided FSC curve	3.42	3.88	3.48
Unmasked-calculated*	3.97	6.11	4.03

*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 3.97 differs from the reported value 3.4 by more than 10 %

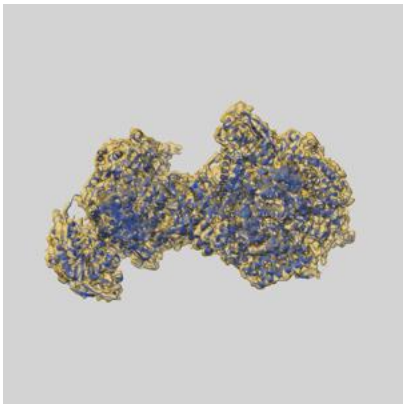
9 Map-model fit

This section contains information regarding the fit between EMDB map EMD-10659 and PDB model 6XZD. Per-residue inclusion information can be found in section 3 on page 4.

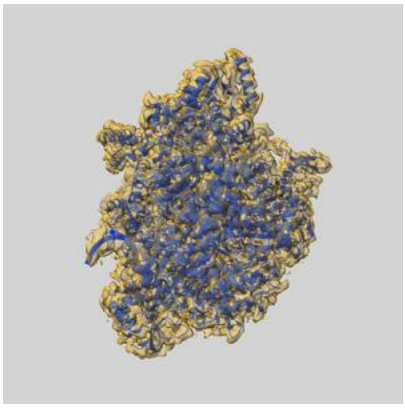
9.1 Map-model overlay



X



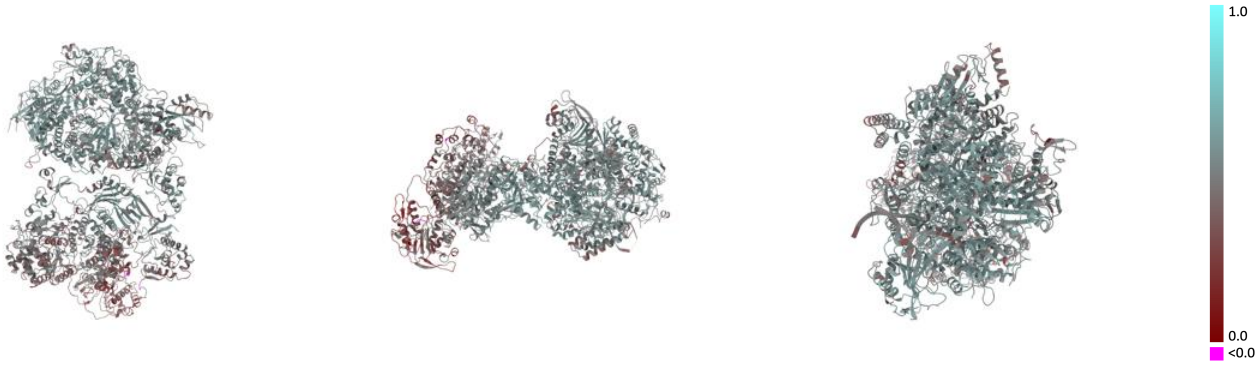
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.13 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.

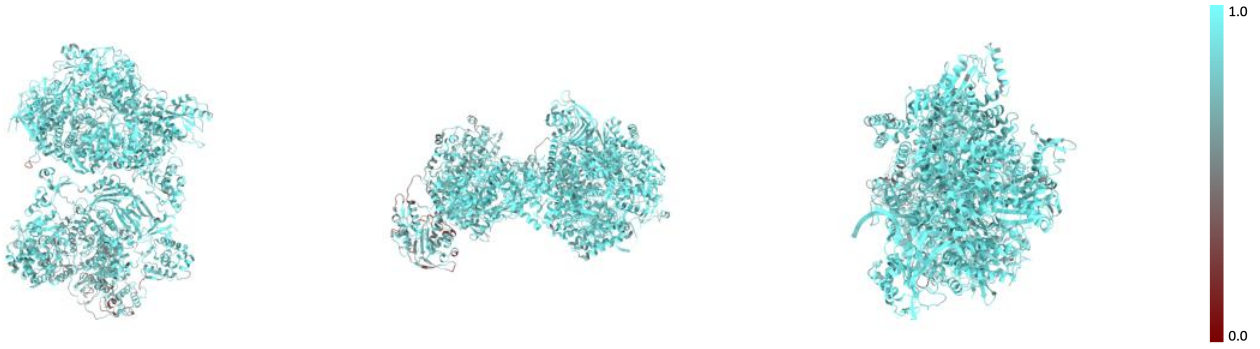
9.2 Q-score mapped to coordinate model (i)



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.

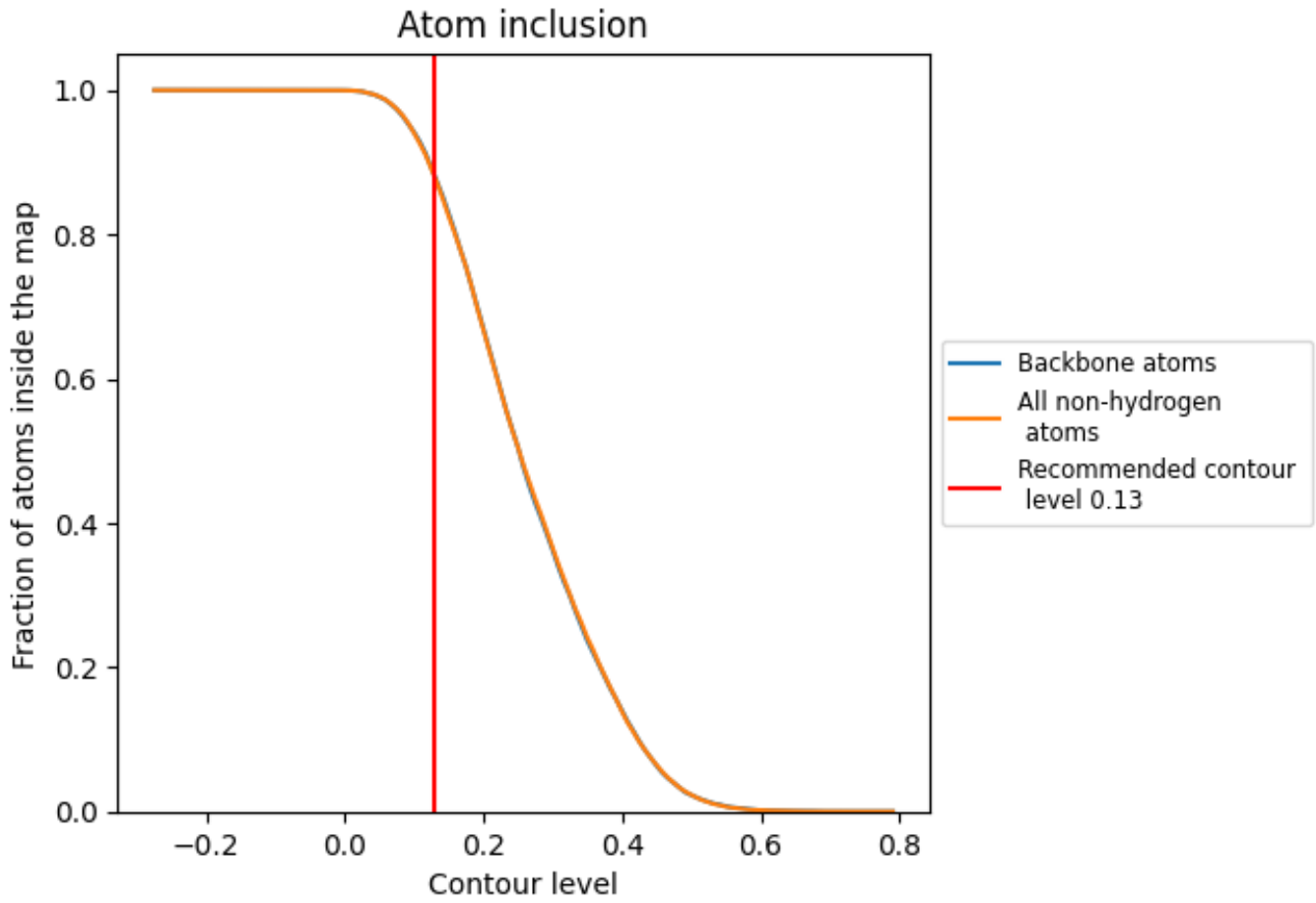
9.3 Atom inclusion mapped to coordinate model (i)



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 (0.13).

9.4 Atom inclusion (i)



At the recommended contour level, 88% of all backbone atoms, 88% 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 (0.13) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.8775	0.4910
AP1	0.9189	0.5320
BP1	0.9323	0.5410
CP1	0.9132	0.5250
DP1	0.9026	0.5110
EP1	0.8495	0.4490
FP1	0.7290	0.3660
IN1	0.9091	0.4880