



wwPDB EM Validation Summary Report ⓘ

Nov 15, 2022 – 06:08 AM JST

PDB ID : 6KIW
EMDB ID : EMD-0693
Title : Cryo-EM structure of human MLL3-ubNCP complex (4.0 angstrom)
Authors : Huang, J.; Xue, H.; Yao, T.
Deposited on : 2019-07-20
Resolution : 4.00 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.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

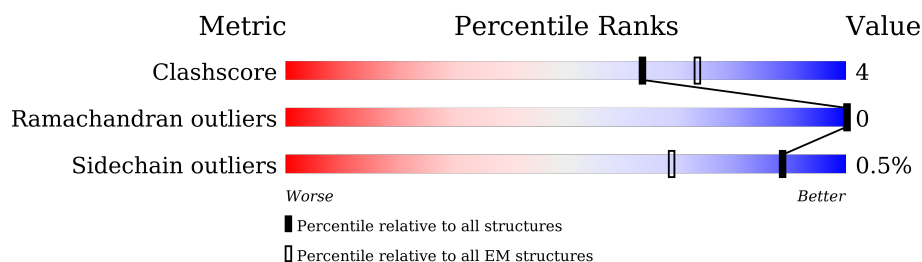
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 4.00 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

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 $\leq 5\%$. 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	135	
1	E	135	
2	B	102	
2	F	102	
3	C	129	
3	G	129	
4	D	122	
4	H	122	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
5	I	144	
6	J	145	
7	K	205	
8	N	538	
9	O	76	
10	R	334	
11	T	534	

2 Entry composition

There are 12 unique types of molecules in this entry. The entry contains 20225 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 Histone H3.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	97	Total	C	N	O	S	0	0
			802	506	155	138	3		
1	E	97	Total	C	N	O	S	0	0
			802	506	155	138	3		

- Molecule 2 is a protein called Histone H4.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	82	Total	C	N	O	S	0	0
			657	416	128	112	1		
2	F	81	Total	C	N	O	S	0	0
			648	410	126	111	1		

- Molecule 3 is a protein called Histone H2A.

Mol	Chain	Residues	Atoms				AltConf	Trace
3	C	107	Total	C	N	O	0	0
			823	519	161	143		
3	G	104	Total	C	N	O	0	0
			800	504	156	140		

- Molecule 4 is a protein called Histone H2B 1.1.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	D	92	Total	C	N	O	S	0	0
			718	451	128	136	3		
4	H	94	Total	C	N	O	S	0	0
			733	460	131	139	3		

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	29	THR	SER	engineered mutation	UNP P02281

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	117	CYS	LYS	engineered mutation	UNP P02281
H	29	THR	SER	engineered mutation	UNP P02281
H	117	CYS	LYS	engineered mutation	UNP P02281

- Molecule 5 is a DNA chain called DNA (144-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
5	I	144	Total	C	N	O	P	0	0
			2935	1393	536	862	144		

- Molecule 6 is a DNA chain called DNA (145-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
6	J	145	Total	C	N	O	P	0	0
			2990	1415	559	871	145		

- Molecule 7 is a protein called Histone-lysine N-methyltransferase 2C.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	K	156	Total	C	N	O	S	0	0
			1259	784	235	228	12		

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
K	4708	SER	CYS	engineered mutation	UNP Q8NEZ4

- Molecule 8 is a protein called Retinoblastoma-binding protein 5.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	N	347	Total	C	N	O	S	0	0
			2724	1715	472	522	15		

- Molecule 9 is a protein called Ubiquitin.

Mol	Chain	Residues	Atoms					AltConf	Trace
9	O	76	Total	C	N	O	S	0	0
			603	379	105	117	2		

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
O	76	CYS	GLY	engineered mutation	UNP P62979

- Molecule 10 is a protein called WD repeat-containing protein 5.

Mol	Chain	Residues	Atoms					AltConf	Trace
10	R	300	Total	C	N	O	S	0	0
			2326	1485	388	444	9		

- Molecule 11 is a protein called Set1/Ash2 histone methyltransferase complex subunit ASH2.

Mol	Chain	Residues	Atoms					AltConf	Trace
11	T	176	Total	C	N	O	S	0	0
			1404	907	235	256	6		

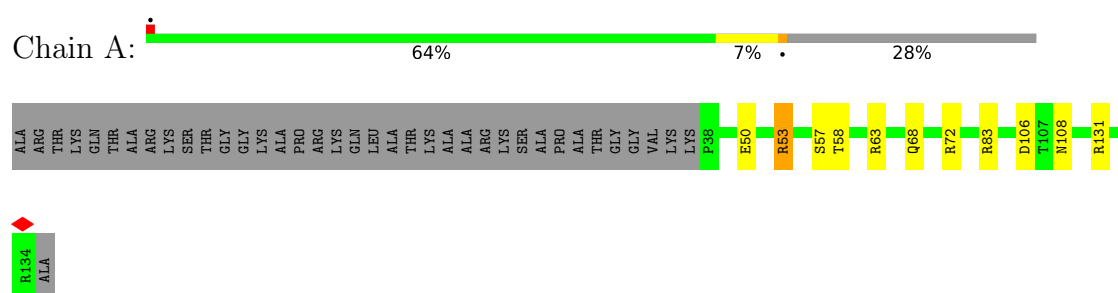
- Molecule 12 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		AltConf
12	K	1	Total	Zn	0
			1	1	

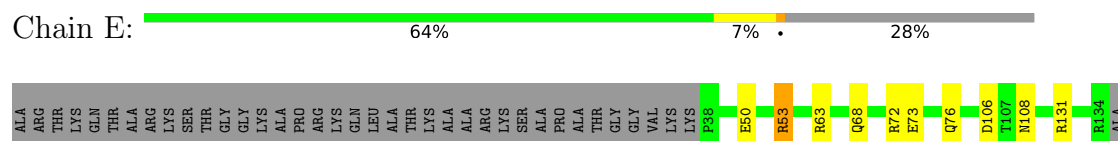
3 Residue-property plots

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.

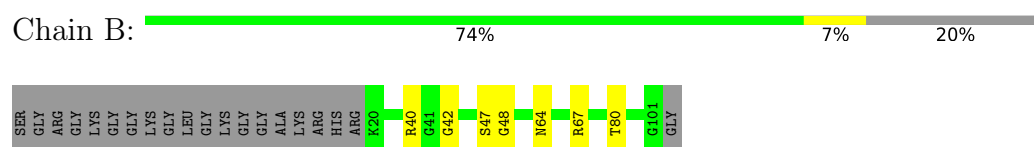
• Molecule 1: Histone H3



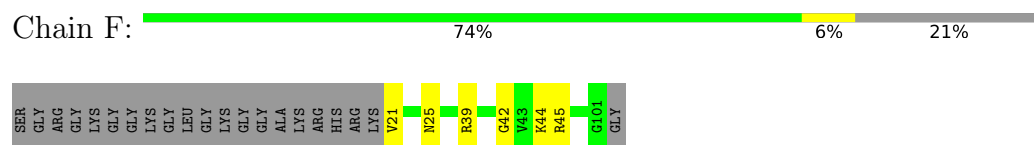
• Molecule 1: Histone H3



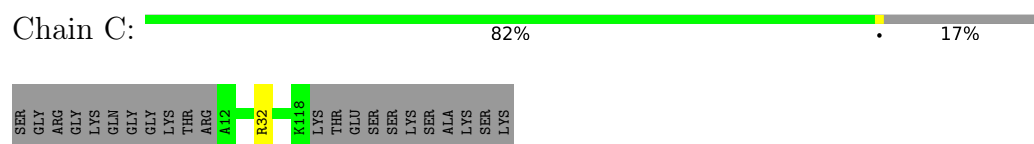
• Molecule 2: Histone H4




• Molecule 2: Histone H4



• Molecule 3: Histone H2A



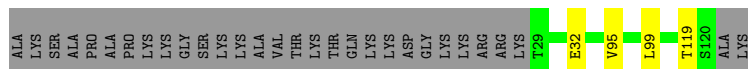
- Molecule 3: Histone H2A

Chain G:  78% 19%



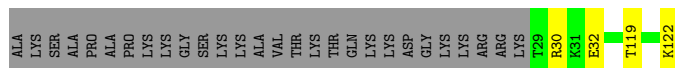
- Molecule 4: Histone H2B 1.1

Chain D:  72% 25%



- Molecule 4: Histone H2B 1.1

Chain H:  74% 23%




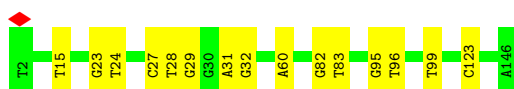
- Molecule 5: DNA (144-MER)

Chain I:  91% 9%



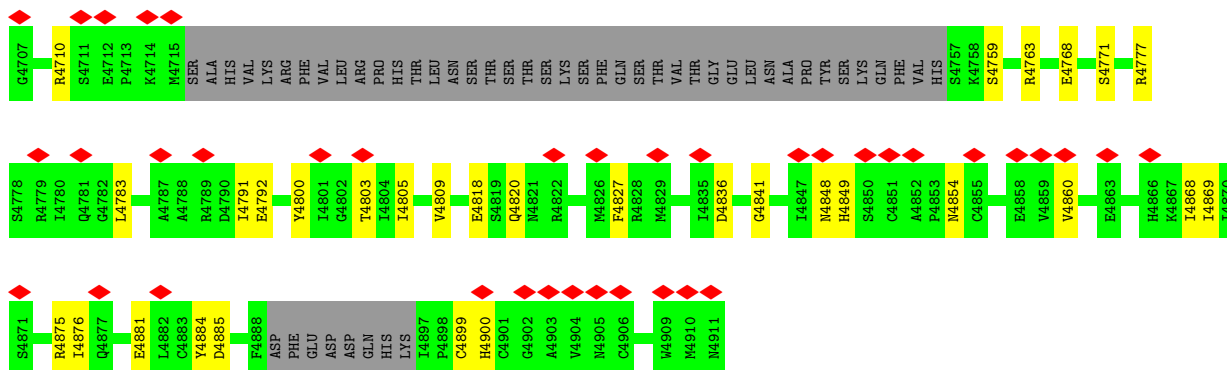
- Molecule 6: DNA (145-MER)

Chain J:  90% 10%



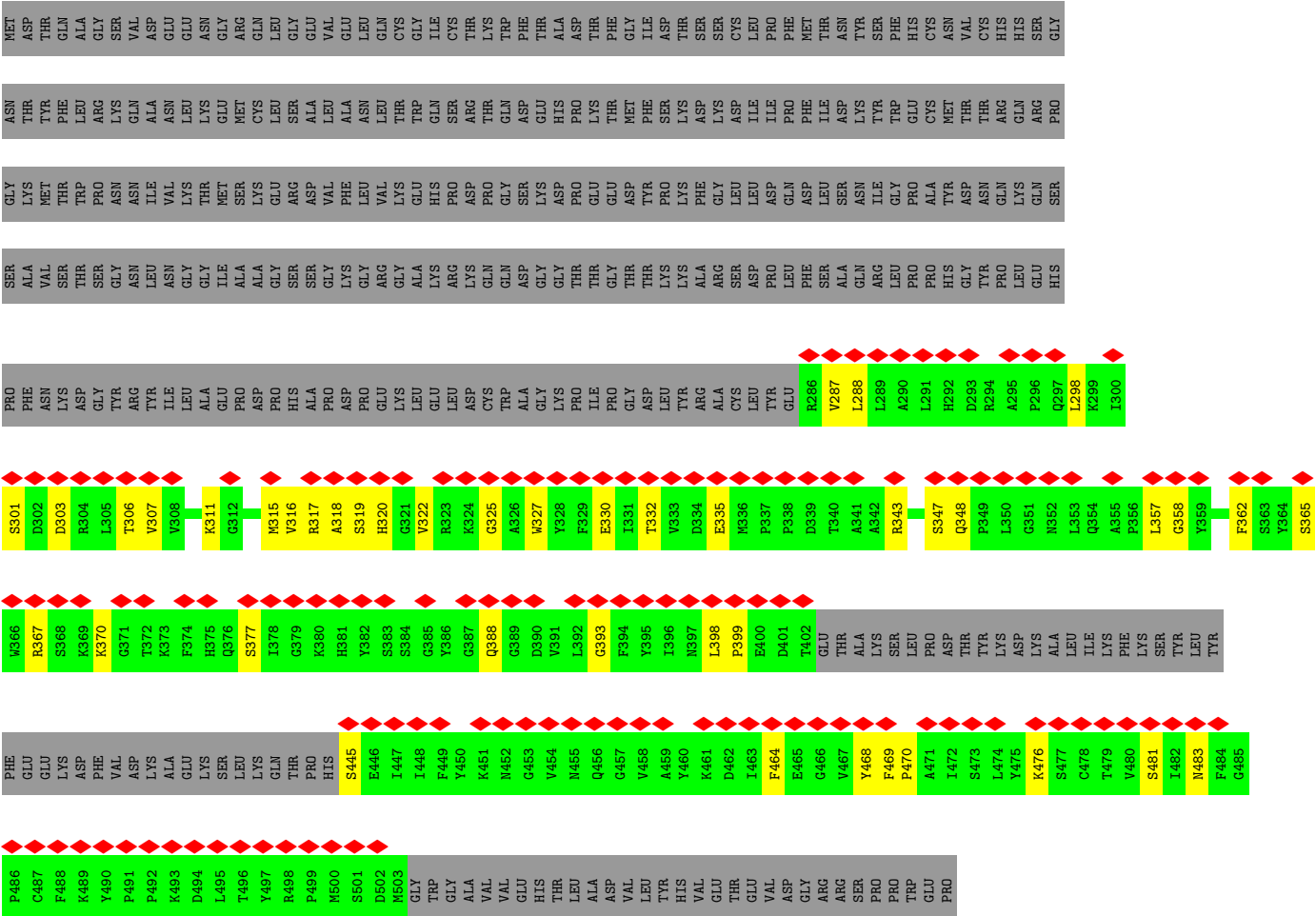
- Molecule 7: Histone-lysine N-methyltransferase 2C

Chain K:  19% 61% 15% 24%





● Molecule 11: Set1/Ash2 histone methyltransferase complex subunit ASH2



4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	81945	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING ONLY	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{\AA}^2$)	40.0	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	0.166	Depositor
Minimum map value	-0.114	Depositor
Average map value	0.000	Depositor
Map value standard deviation	0.006	Depositor
Recommended contour level	0.015	Depositor
Map size (\AA)	294.30002, 294.30002, 294.30002	wwPDB
Map dimensions	270, 270, 270	wwPDB
Map angles ($^\circ$)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (\AA)	1.09, 1.09, 1.09	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

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	0.36	0/814	0.59	0/1092
1	E	0.36	0/814	0.59	0/1092
2	B	0.45	0/664	0.63	0/889
2	F	0.45	0/655	0.59	0/878
3	C	0.37	0/833	0.54	0/1124
3	G	0.35	0/810	0.55	0/1095
4	D	0.41	0/729	0.55	0/983
4	H	0.37	0/744	0.55	0/1001
5	I	0.72	0/3289	1.00	1/5069 (0.0%)
6	J	0.71	0/3357	0.98	0/5184
7	K	0.28	0/1281	0.53	0/1715
8	N	0.30	0/2781	0.56	0/3777
9	O	0.26	0/609	0.57	0/819
10	R	0.27	0/2382	0.54	1/3231 (0.0%)
11	T	0.28	0/1445	0.55	0/1954
All	All	0.48	0/21207	0.74	2/29903 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
8	N	0	1

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	I	4	DG	O4'-C4'-C3'	-6.81	101.78	104.50

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
10	R	309	CYS	C-N-CA	5.75	136.08	121.70

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
8	N	327	GLU	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	A	802	0	841	8	0
1	E	802	0	841	7	0
2	B	657	0	706	6	0
2	F	648	0	693	5	0
3	C	823	0	882	1	0
3	G	800	0	851	2	0
4	D	718	0	734	3	0
4	H	733	0	752	3	0
5	I	2935	0	1615	9	0
6	J	2990	0	1628	10	0
7	K	1259	0	1252	22	0
8	N	2724	0	2677	35	0
9	O	603	0	631	11	0
10	R	2326	0	2309	48	0
11	T	1404	0	1369	24	0
12	K	1	0	0	0	0
All	All	20225	0	17781	163	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 163 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:E:63:ARG:HH21	6:J:60:DA:H5''	1.56	0.71
9:O:44:ILE:HB	9:O:68:HIS:HB2	1.78	0.65
1:E:76:GLN:HE22	2:F:21:VAL:HG23	1.62	0.64
8:N:27:ALA:H	8:N:316:GLY:HA2	1.61	0.63
11:T:315:MET:SD	11:T:317:ARG:NH1	2.73	0.62

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all EM entries.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	95/135 (70%)	89 (94%)	6 (6%)	0	100	100
1	E	95/135 (70%)	89 (94%)	6 (6%)	0	100	100
2	B	80/102 (78%)	73 (91%)	7 (9%)	0	100	100
2	F	79/102 (78%)	77 (98%)	2 (2%)	0	100	100
3	C	105/129 (81%)	99 (94%)	6 (6%)	0	100	100
3	G	102/129 (79%)	96 (94%)	6 (6%)	0	100	100
4	D	90/122 (74%)	88 (98%)	2 (2%)	0	100	100
4	H	92/122 (75%)	89 (97%)	3 (3%)	0	100	100
7	K	150/205 (73%)	132 (88%)	18 (12%)	0	100	100
8	N	343/538 (64%)	294 (86%)	49 (14%)	0	100	100
9	O	74/76 (97%)	68 (92%)	6 (8%)	0	100	100
10	R	298/334 (89%)	263 (88%)	35 (12%)	0	100	100
11	T	172/534 (32%)	157 (91%)	15 (9%)	0	100	100
All	All	1775/2663 (67%)	1614 (91%)	161 (9%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all EM entries.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	85/110 (77%)	84 (99%)	1 (1%)	71	84
1	E	85/110 (77%)	84 (99%)	1 (1%)	71	84
2	B	68/78 (87%)	68 (100%)	0	100	100
2	F	67/78 (86%)	67 (100%)	0	100	100
3	C	84/101 (83%)	84 (100%)	0	100	100
3	G	82/101 (81%)	81 (99%)	1 (1%)	71	84
4	D	79/102 (78%)	79 (100%)	0	100	100
4	H	80/102 (78%)	80 (100%)	0	100	100
7	K	136/182 (75%)	136 (100%)	0	100	100
8	N	304/462 (66%)	299 (98%)	5 (2%)	62	79
9	O	69/69 (100%)	69 (100%)	0	100	100
10	R	262/291 (90%)	262 (100%)	0	100	100
11	T	150/460 (33%)	150 (100%)	0	100	100
All	All	1551/2246 (69%)	1543 (100%)	8 (0%)	89	93

5 of 8 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
8	N	348	ARG
8	N	324	ASN
8	N	250	ASN
8	N	34	ARG
8	N	294	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 15 such sidechains are listed below:

Mol	Chain	Res	Type
7	K	4807	ASN
10	R	136	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
8	N	246	GLN
10	R	295	GLN
8	N	328	ASN

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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 1 ligands modelled in this entry, 1 is 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.

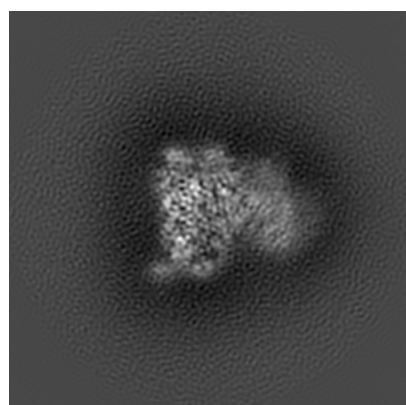
6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-0693. These allow visual inspection of the internal detail of the map and identification of artifacts.

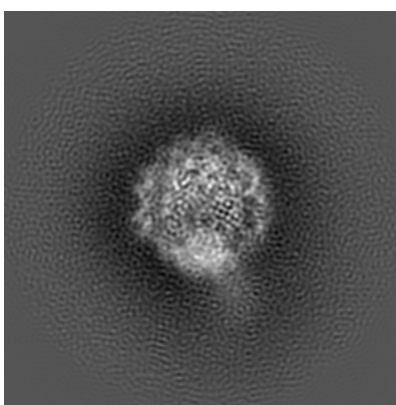
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections [i](#)

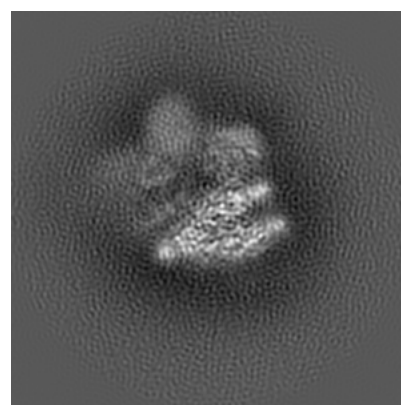
6.1.1 Primary map



X



Y

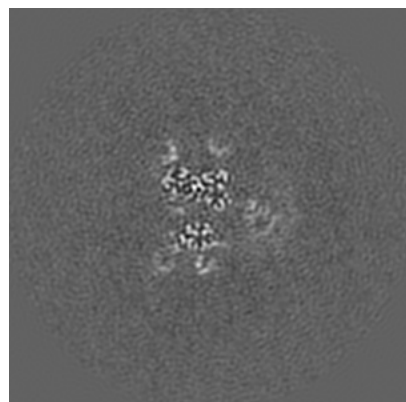


Z

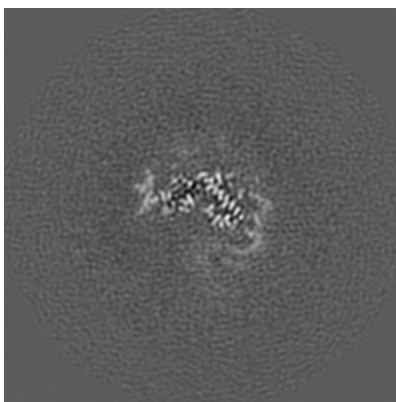
The images above show the map projected in three orthogonal directions.

6.2 Central slices [i](#)

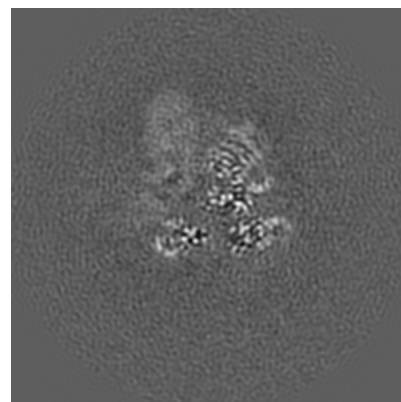
6.2.1 Primary map



X Index: 135



Y Index: 135

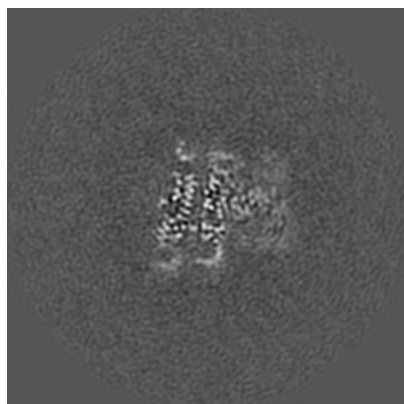


Z Index: 135

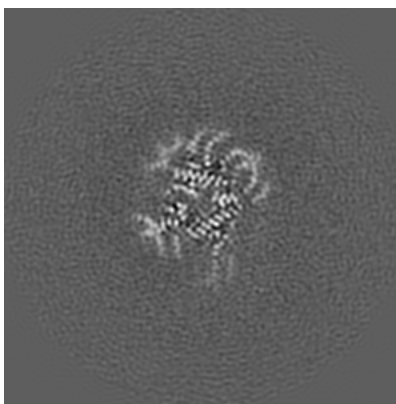
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices [i](#)

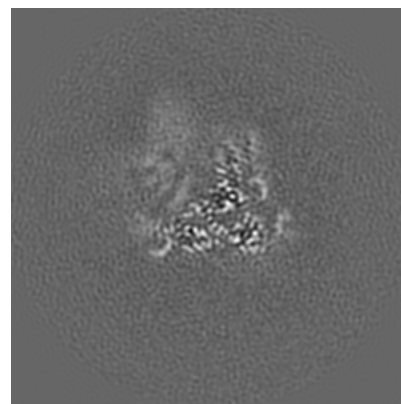
6.3.1 Primary map



X Index: 153



Y Index: 120

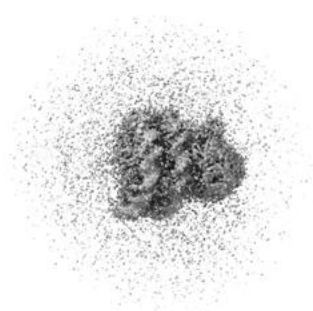


Z Index: 142

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

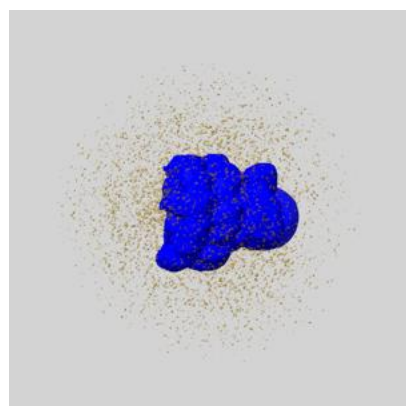
6.5 Mask visualisation [i](#)

This section shows the 3D surface view of the primary map at 50% transparency overlaid with the specified mask at 0% transparency

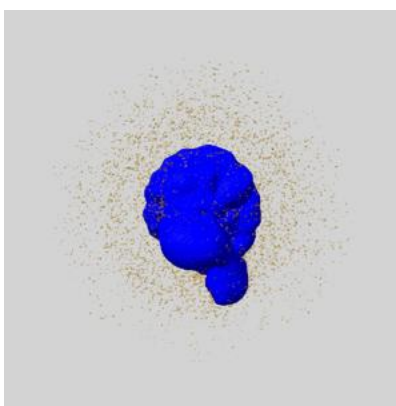
A mask typically either:

- Encompasses the whole structure
- Separates out a domain, a functional unit, a monomer or an area of interest from a larger structure

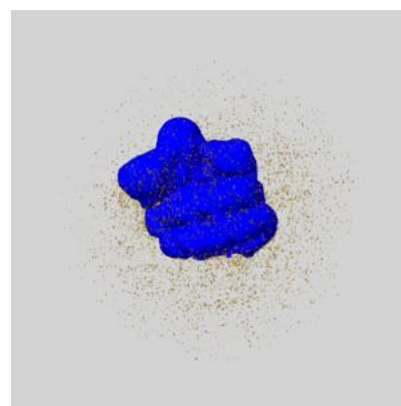
6.5.1 emd_0693_msk_1.map [i](#)



X



Y

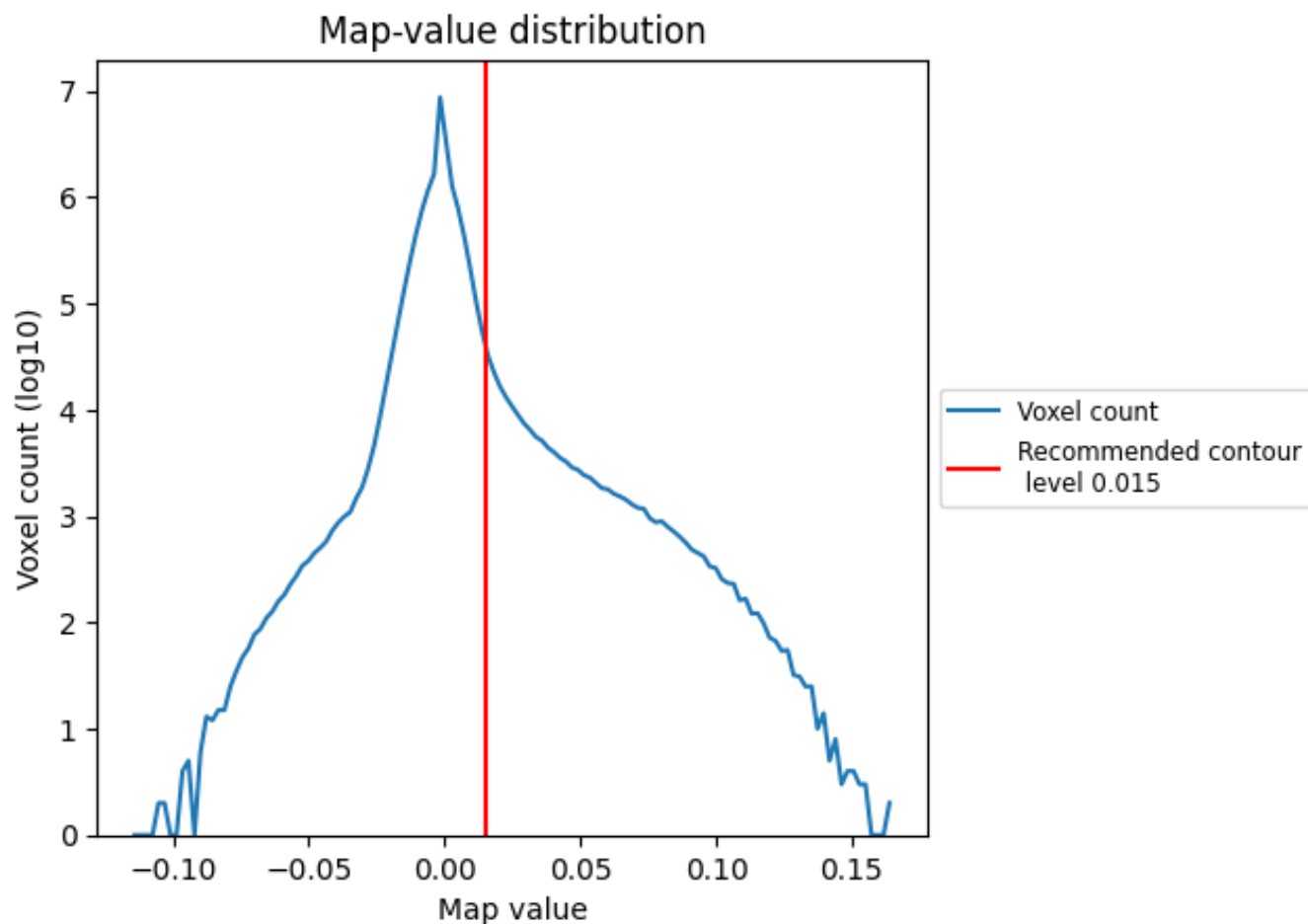


Z

7 Map analysis [i](#)

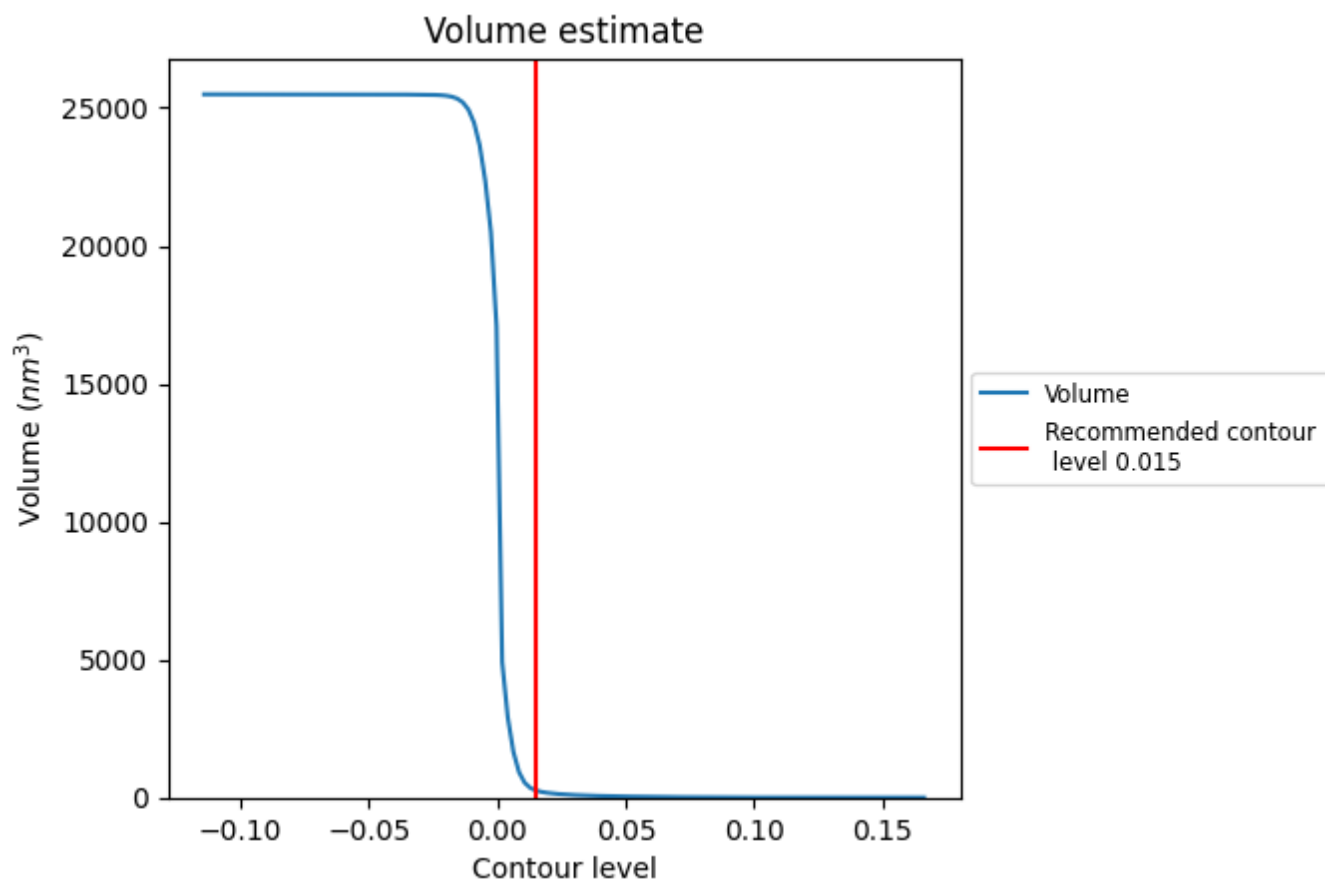
This section contains the results of statistical analysis of the map.

7.1 Map-value distribution [i](#)



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.

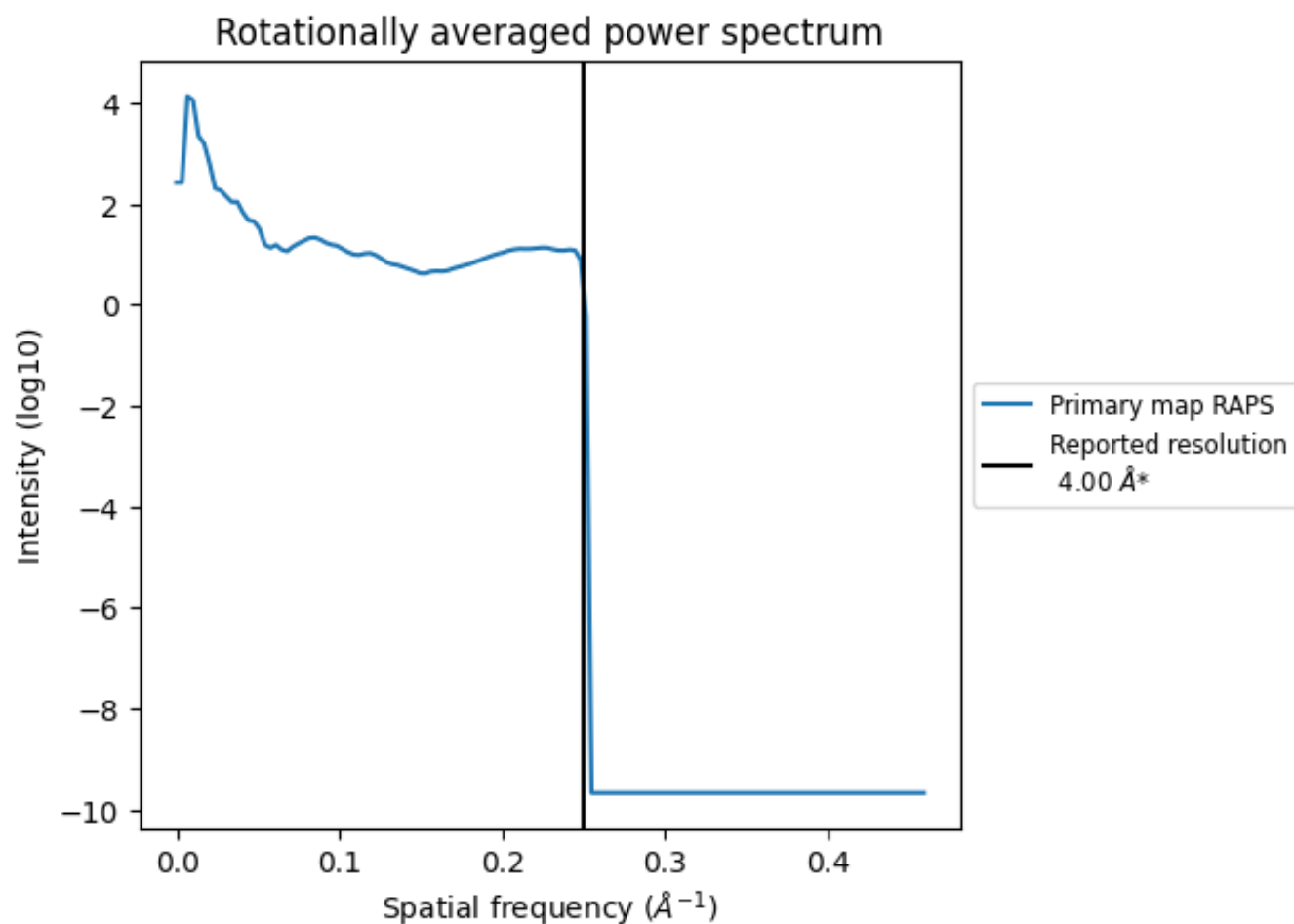
7.2 Volume estimate [i](#)



The volume at the recommended contour level is 258 nm³; this corresponds to an approximate mass of 233 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 ⓘ

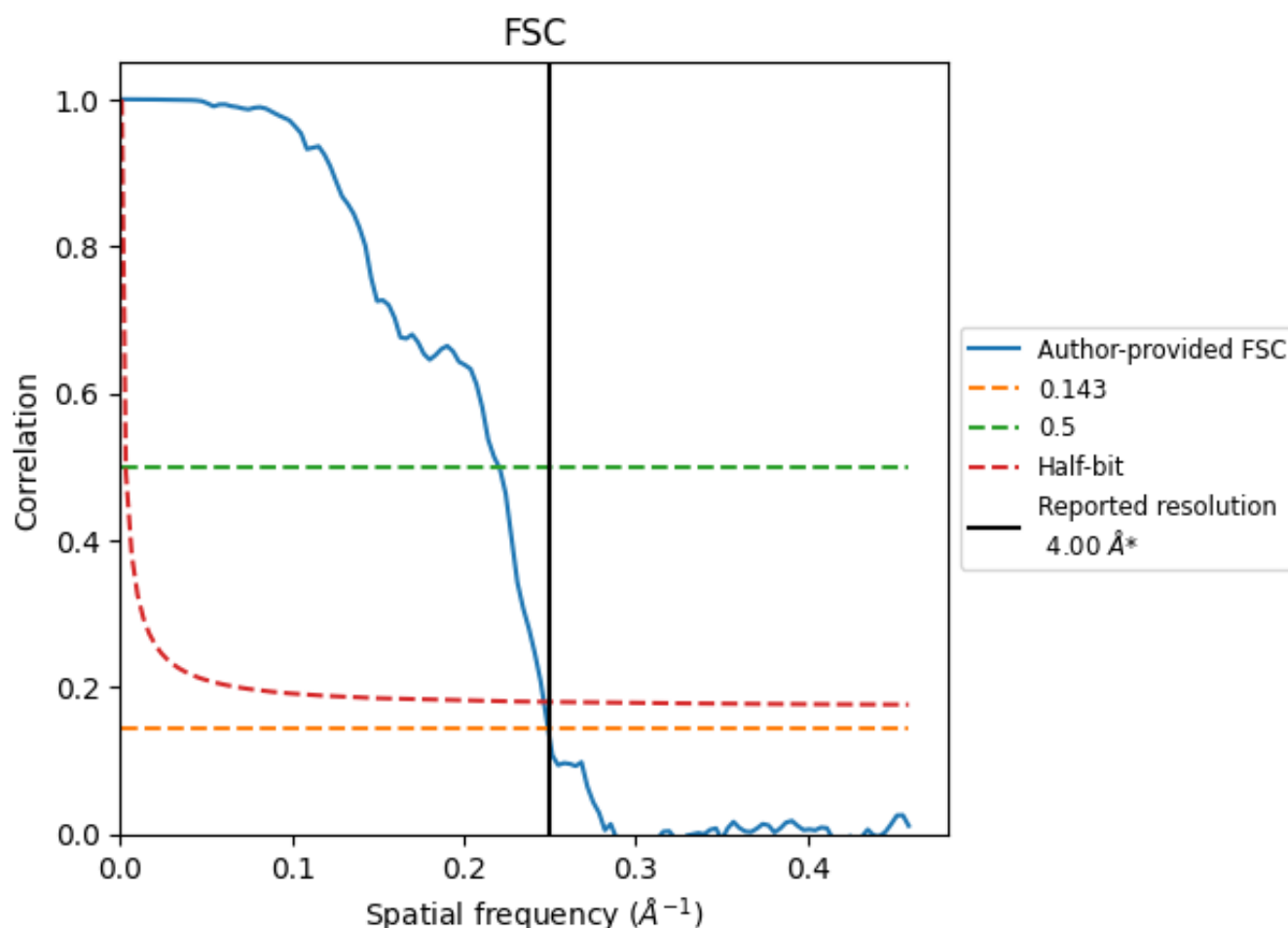


*Reported resolution corresponds to spatial frequency of 0.250 Å⁻¹

8 Fourier-Shell correlation [i](#)

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 [i](#)



*Reported resolution corresponds to spatial frequency of 0.250 Å⁻¹

8.2 Resolution estimates [i](#)

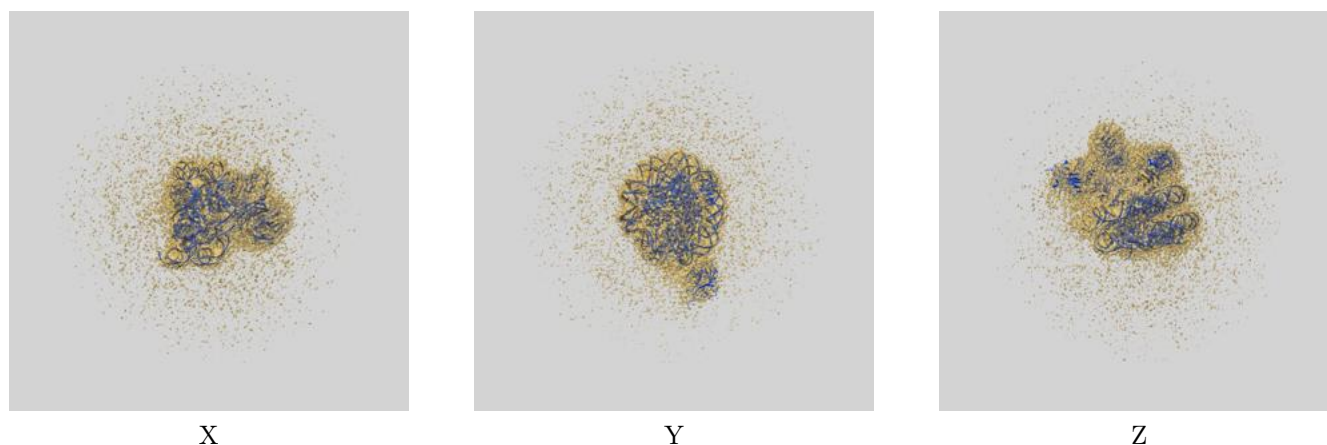
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	4.00	-	-
Author-provided FSC curve	4.02	4.54	4.06
Unmasked-calculated*	-	-	-

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

9 Map-model fit [i](#)

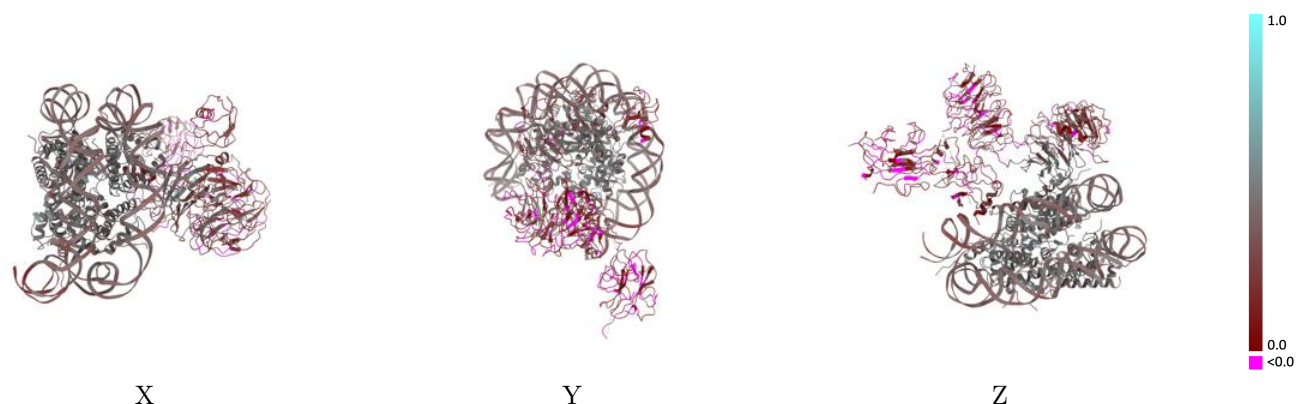
This section contains information regarding the fit between EMDB map EMD-0693 and PDB model 6KIW. Per-residue inclusion information can be found in section [3](#) on page [7](#).

9.1 Map-model overlay [i](#)



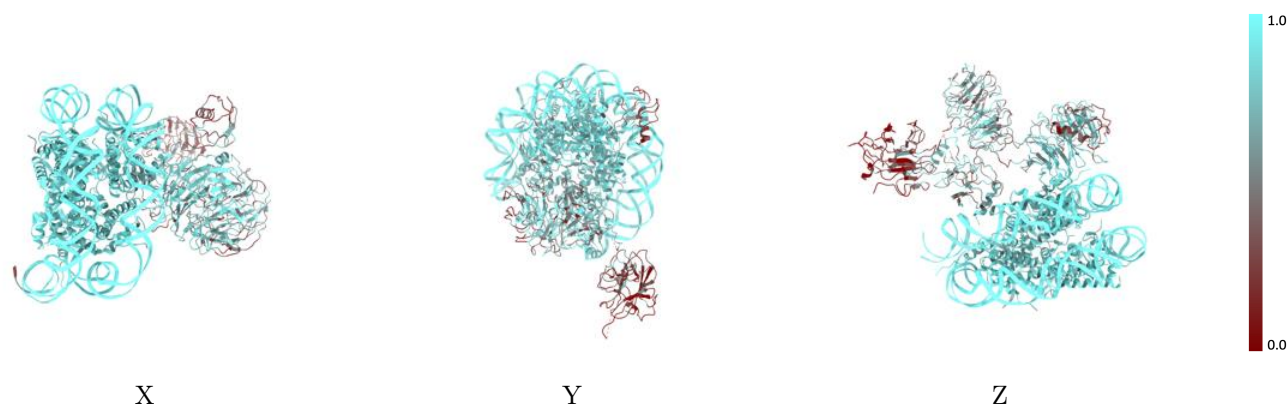
The images above show the 3D surface view of the map at the recommended contour level 0.015 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

9.2 Q-score mapped to coordinate model [i](#)



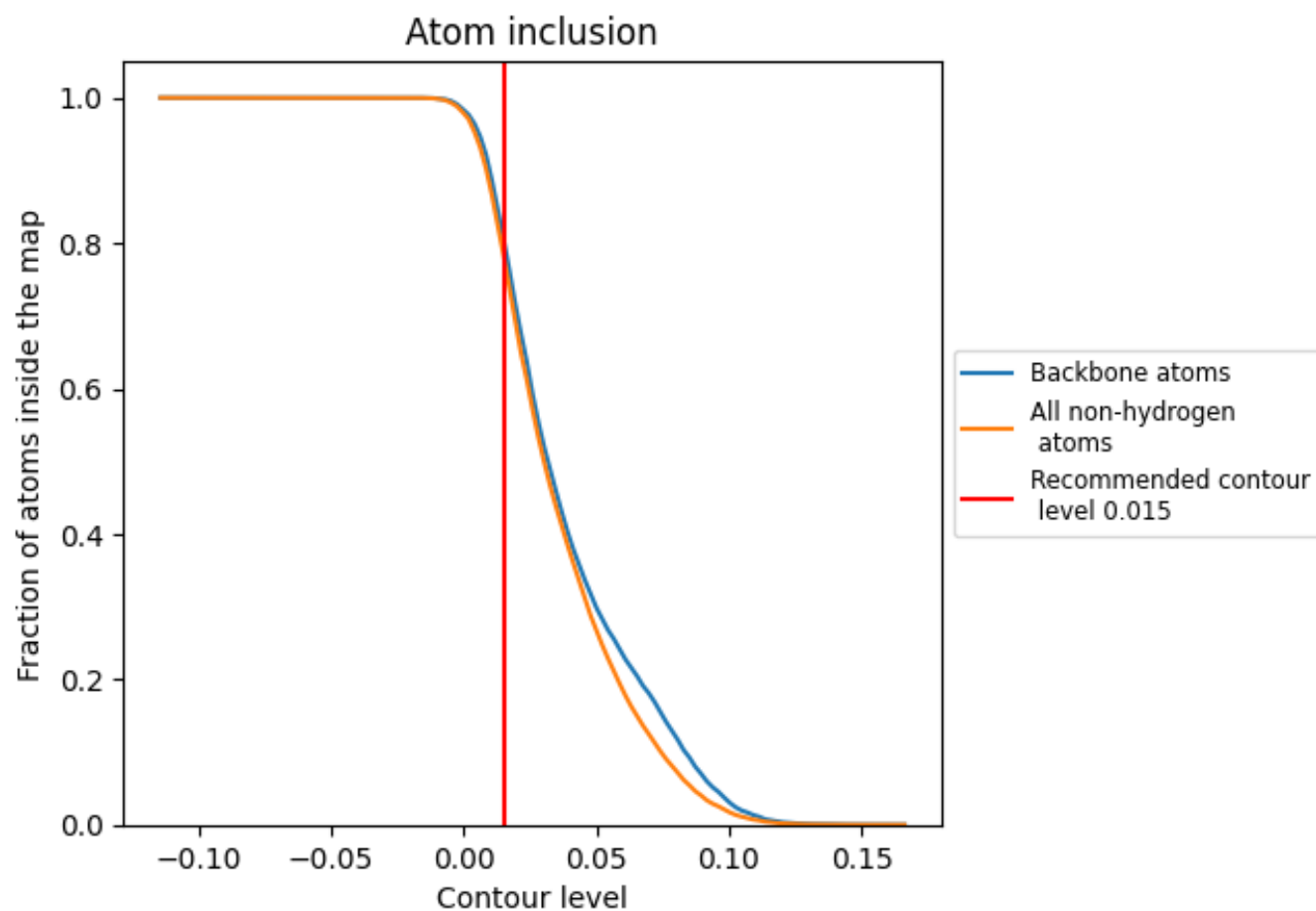
The images above show the model with each residue coloured according to its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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](#)



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 (0.015).

































9.4 Atom inclusion [i](#)



At the recommended contour level, 81% of all backbone atoms, 78% 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 (0.015) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	 0.7843	 0.3300
A	 0.8755	 0.4420
B	 0.9113	 0.4610
C	 0.9098	 0.4730
D	 0.8973	 0.4540
E	 0.8885	 0.4510
F	 0.9212	 0.4630
G	 0.9161	 0.4670
H	 0.8966	 0.4470
I	 0.9693	 0.3730
J	 0.9612	 0.3740
K	 0.5995	 0.1750
N	 0.7597	 0.3190
O	 0.3704	 0.1930
R	 0.5837	 0.1560
T	 0.2369	 0.1070

