



wwPDB EM Validation Summary Report ⓘ

Nov 20, 2022 – 03:01 PM JST

PDB ID : 7BWD
EMDB ID : EMD-30232
Title : Structure of Dot1L-H2BK34ub Nucleosome Complex
Authors : Lou, Z.Y.; Liu, L.; Cao, L.; Ai, H.S.; Sun, Z.X.
Deposited on : 2020-04-14
Resolution : 4.32 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
Mogul : 1.8.5 (274361), CSD as541be (2020)
MolProbity : 4.02b-467
buster-report : 1.1.7 (2018)
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

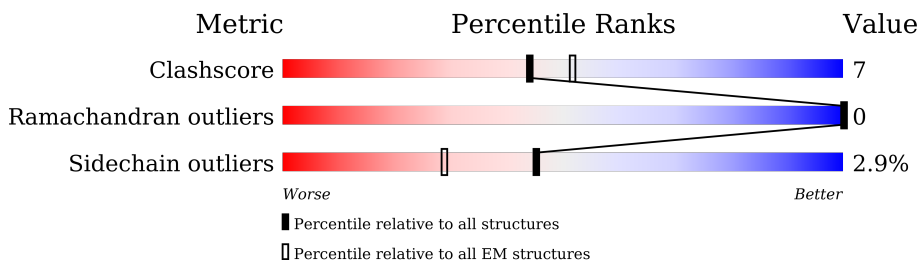
1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 4.32 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	K	416	
2	B	102	
2	F	102	
3	C	129	
3	G	129	
4	D	125	
4	H	125	
5	A	135	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
5	E	135	 61% 10% 28%
6	I	146	 79% 21%
7	J	146	 79% 20%
8	L	76	 97% 88% 12%
8	M	76	 97% 91% 9%

2 Entry composition [i](#)

There are 9 unique types of molecules in this entry. The entry contains 16042 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-lysine N-methyltransferase, H3 lysine-79 specific.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	K	328	2671	1706	455	498	12	0	0

- Molecule 2 is a protein called Histone H4.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	F	87	703	442	142	118	1	0	0
2	B	84	673	424	133	115	1	0	0

- Molecule 3 is a protein called Histone H2A type 1-B/E.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
3	G	109	844	532	167	145	0	0
3	C	112	866	544	170	152	0	0

- Molecule 4 is a protein called Histone H2B type 1-K.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	H	94	735	461	134	138	2	0	0
4	D	94	735	461	134	138	2	0	0

- Molecule 5 is a protein called Histone H3.1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
5	E	97	799	505	153	137	4	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
5	A	99	816	514	158	140	4	0	0

- Molecule 6 is a DNA chain called DNA (146-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
6	I	146	2975	1413	540	876	146	0	0

- Molecule 7 is a DNA chain called DNA (146-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
7	J	146	3011	1425	564	876	146	0	0

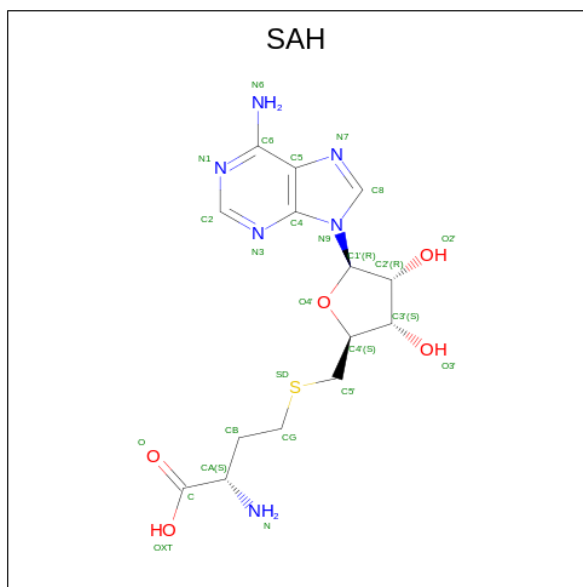
- Molecule 8 is a protein called Ubiquitin.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
8	L	76	594	374	103	116	1	0	0
8	M	76	594	374	103	116	1	0	0

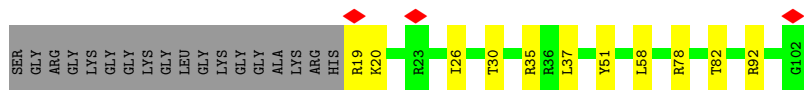
There are 2 discrepancies between the modelled and reference sequences:

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

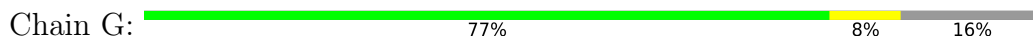
- Molecule 9 is S-ADENOSYL-L-HOMOCYSTEINE (three-letter code: SAH) (formula: $C_{14}H_{20}N_6O_5S$).



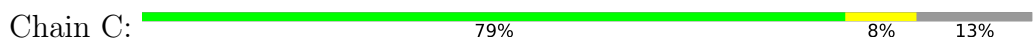
Mol	Chain	Residues	Atoms					AltConf
			Total	C	N	O	S	
9	K	1	26	14	6	5	1	0



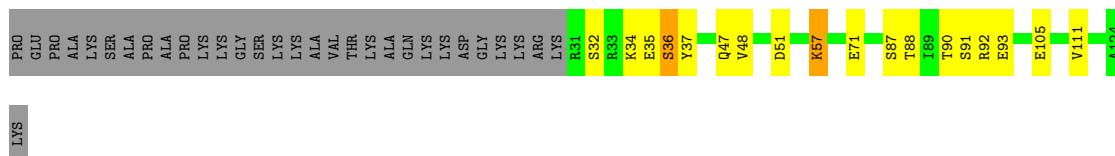
• Molecule 3: Histone H2A type 1-B/E



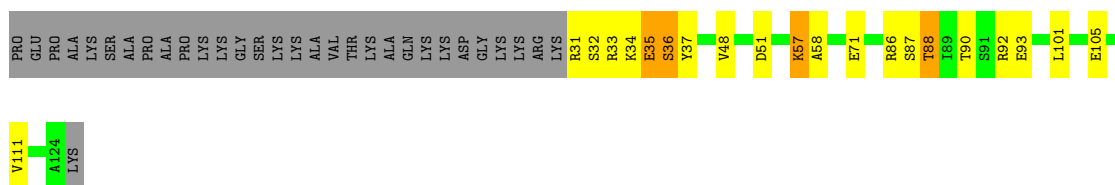
• Molecule 3: Histone H2A type 1-B/E



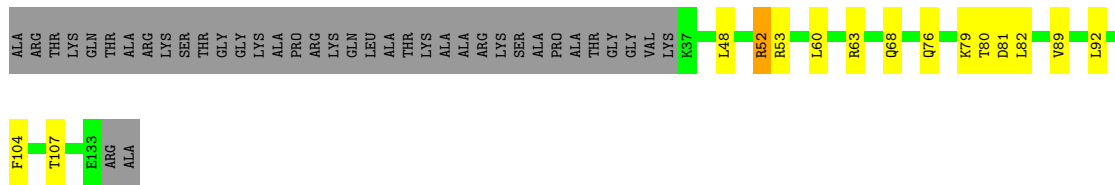
• Molecule 4: Histone H2B type 1-K



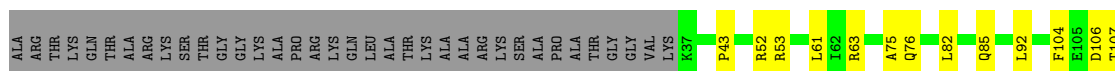
• Molecule 4: Histone H2B type 1-K



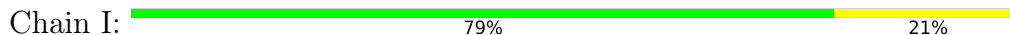
• Molecule 5: Histone H3.1



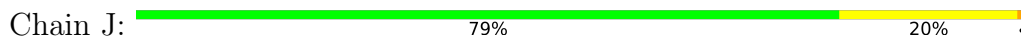
• Molecule 5: Histone H3.1



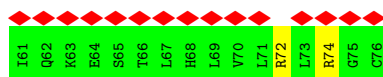
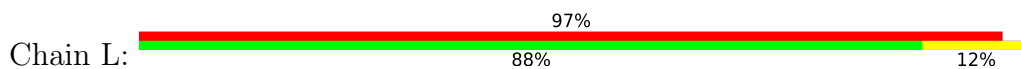
- Molecule 6: DNA (146-MER)



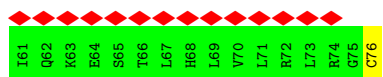
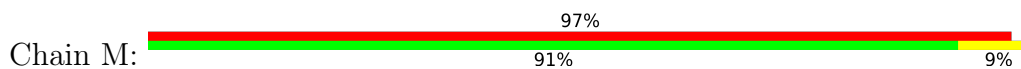
- Molecule 7: DNA (146-MER)



- Molecule 8: Ubiquitin



- Molecule 8: Ubiquitin



4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	22445	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TECNAI F30	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{\AA}^2$)	50	Depositor
Minimum defocus (nm)	Not provided	
Maximum defocus (nm)	Not provided	
Magnification	Not provided	
Image detector	GATAN K2 BASE (4k x 4k)	Depositor
Maximum map value	0.210	Depositor
Minimum map value	-0.126	Depositor
Average map value	0.001	Depositor
Map value standard deviation	0.008	Depositor
Recommended contour level	0.02	Depositor
Map size (Å)	216.00002, 216.00002, 216.00002	wwPDB
Map dimensions	200, 200, 200	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.08, 1.08, 1.08	Depositor

5 Model quality [i](#)

5.1 Standard geometry [i](#)

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

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	K	0.31	0/2741	0.59	1/3718 (0.0%)
2	B	0.53	0/680	0.60	0/908
2	F	0.53	0/711	0.63	0/948
3	C	0.43	0/876	0.55	0/1180
3	G	0.45	0/854	0.59	1/1150 (0.1%)
4	D	0.47	0/746	0.55	0/1003
4	H	0.47	0/746	0.55	0/1003
5	A	0.45	0/828	0.59	0/1109
5	E	0.50	0/811	0.54	0/1088
6	I	0.93	0/3333	0.95	1/5137 (0.0%)
7	J	0.93	0/3381	0.96	1/5221 (0.0%)
8	L	0.27	0/600	0.59	0/809
8	M	0.28	0/600	0.59	0/809
All	All	0.67	0/16907	0.77	4/24083 (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
3	C	0	1

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	K	275	LEU	CA-CB-CG	6.35	129.91	115.30
3	G	58	LEU	CA-CB-CG	5.66	128.33	115.30
6	I	140	DC	P-O3'-C3'	5.25	126.00	119.70

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
7	J	126	DC	O4'-C4'-C3'	-5.04	102.48	104.50

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	C	13	LYS	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	K	2671	0	2615	39	0
2	B	673	0	722	9	0
2	F	703	0	755	9	0
3	C	866	0	928	7	0
3	G	844	0	910	4	0
4	D	735	0	756	19	0
4	H	735	0	756	7	0
5	A	816	0	856	10	0
5	E	799	0	838	11	0
6	I	2975	0	1639	27	0
7	J	3011	0	1639	55	0
8	L	594	0	610	37	0
8	M	594	0	611	15	0
9	K	26	0	13	2	0
All	All	16042	0	13648	178	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 178 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)
7:J:125:DG:C4'	8:L:72:ARG:NH2	1.67	1.54

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
7:J:125:DG:H4'	8:L:72:ARG:CZ	1.38	1.53
7:J:125:DG:H4'	8:L:72:ARG:NH2	1.16	1.42
4:D:31:ARG:N	8:M:76:CYS:H	1.17	1.38
4:D:31:ARG:N	8:M:76:CYS:N	1.73	1.33

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	K	326/416 (78%)	299 (92%)	27 (8%)	0	100	100
2	B	82/102 (80%)	79 (96%)	3 (4%)	0	100	100
2	F	85/102 (83%)	75 (88%)	10 (12%)	0	100	100
3	C	110/129 (85%)	101 (92%)	9 (8%)	0	100	100
3	G	107/129 (83%)	102 (95%)	5 (5%)	0	100	100
4	D	92/125 (74%)	85 (92%)	7 (8%)	0	100	100
4	H	92/125 (74%)	85 (92%)	7 (8%)	0	100	100
5	A	97/135 (72%)	93 (96%)	4 (4%)	0	100	100
5	E	95/135 (70%)	91 (96%)	4 (4%)	0	100	100
8	L	74/76 (97%)	73 (99%)	1 (1%)	0	100	100
8	M	74/76 (97%)	73 (99%)	1 (1%)	0	100	100
All	All	1234/1550 (80%)	1156 (94%)	78 (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	
1	K	294/360 (82%)	291 (99%)	3 (1%)	76	86
2	B	69/78 (88%)	69 (100%)	0	100	100
2	F	72/78 (92%)	71 (99%)	1 (1%)	67	81
3	C	89/99 (90%)	89 (100%)	0	100	100
3	G	86/99 (87%)	85 (99%)	1 (1%)	71	84
4	D	80/104 (77%)	69 (86%)	11 (14%)	3	20
4	H	80/104 (77%)	69 (86%)	11 (14%)	3	20
5	A	86/110 (78%)	84 (98%)	2 (2%)	50	70
5	E	85/110 (77%)	83 (98%)	2 (2%)	49	69
8	L	66/69 (96%)	66 (100%)	0	100	100
8	M	66/69 (96%)	66 (100%)	0	100	100
All	All	1073/1280 (84%)	1042 (97%)	31 (3%)	45	64

5 of 31 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	H	105	GLU
4	D	87	SER
5	E	53	ARG
4	D	105	GLU
4	D	48	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 17 such sidechains are listed below:

Mol	Chain	Res	Type
8	L	60	ASN
8	M	60	ASN
4	H	84	ASN
5	E	39	HIS
5	A	68	GLN

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](#)

1 ligand is 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	SAH	K	501	1	24,28,28	1.15	1 (4%)	25,40,40	2.21	10 (40%)

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	SAH	K	501	1	-	1/11/31/31	0/3/3/3

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
9	K	501	SAH	C2-N3	3.57	1.37	1.32

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
9	K	501	SAH	N3-C2-N1	-5.79	119.63	128.68
9	K	501	SAH	C1'-N9-C4	-4.11	119.43	126.64
9	K	501	SAH	C5'-SD-CG	-4.08	90.01	102.27
9	K	501	SAH	OXT-C-O	-3.60	115.92	124.09
9	K	501	SAH	C2'-C3'-C4'	2.96	108.39	102.64

There are no chirality outliers.

All (1) torsion outliers are listed below:

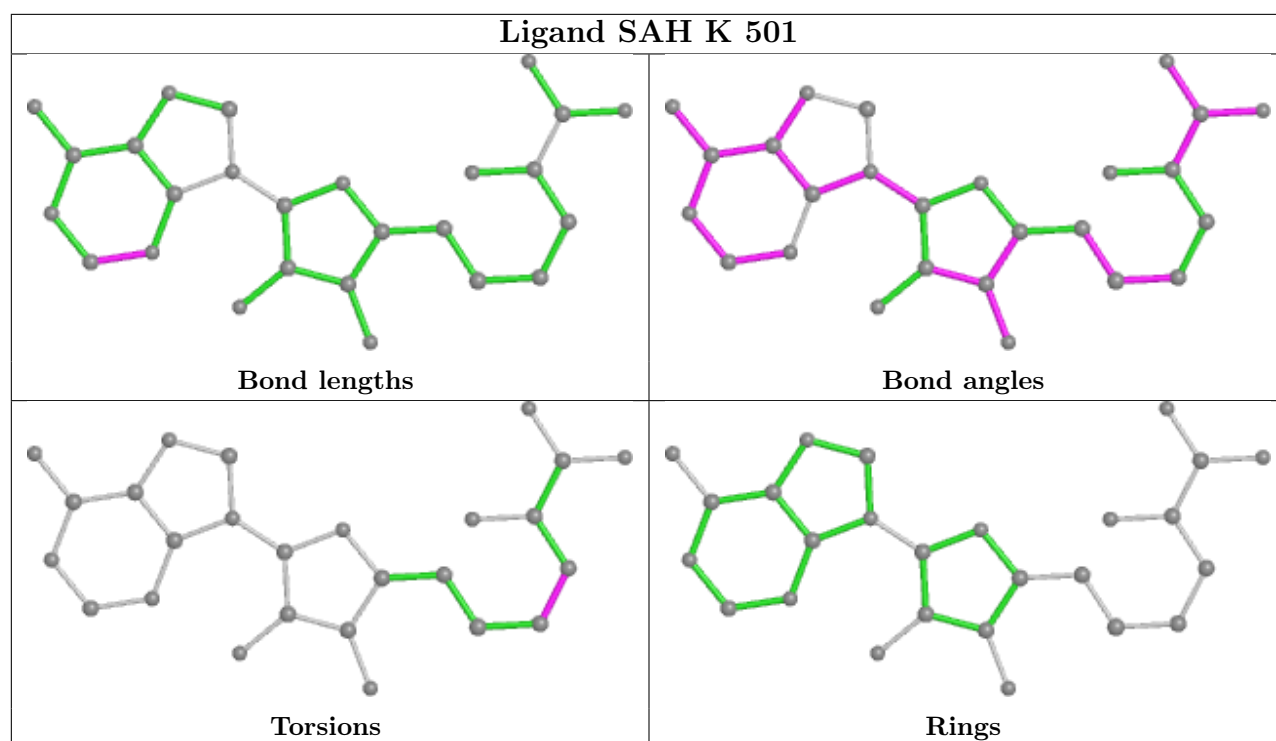
Mol	Chain	Res	Type	Atoms
9	K	501	SAH	CA-CB-CG-SD

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
9	K	501	SAH	2	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.



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.

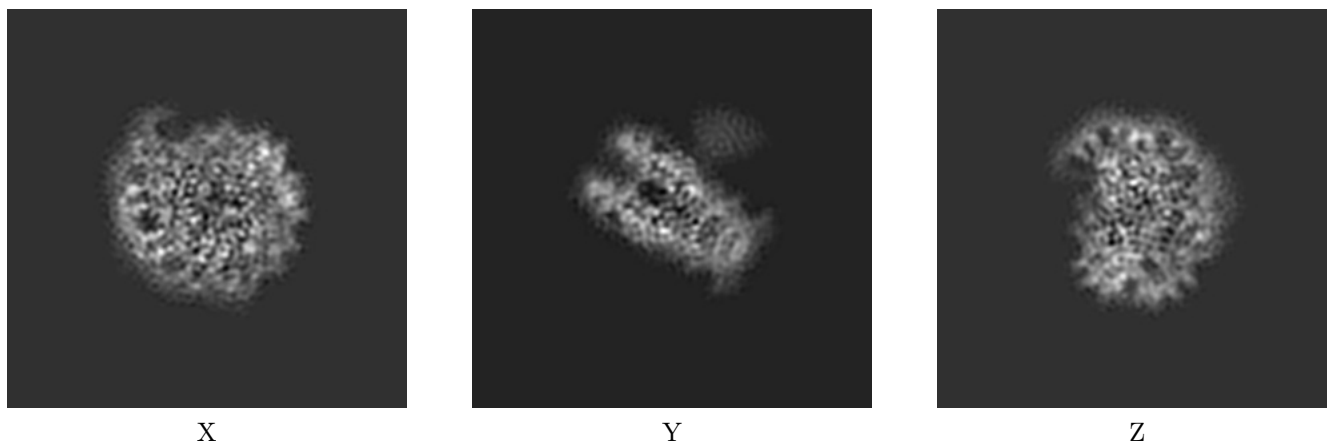
6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-30232. 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](#)

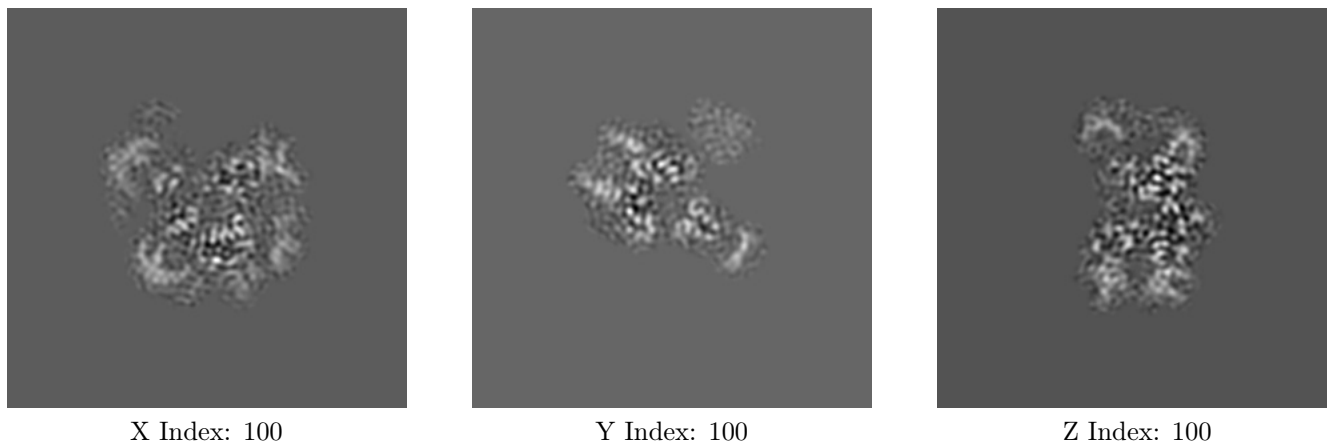
6.1.1 Primary map



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

6.2 Central slices [i](#)

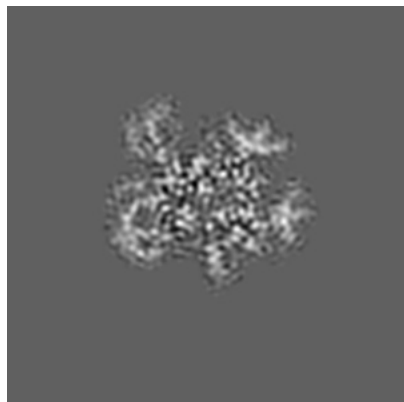
6.2.1 Primary map



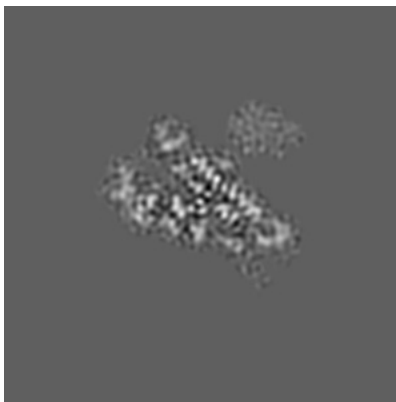
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices [i](#)

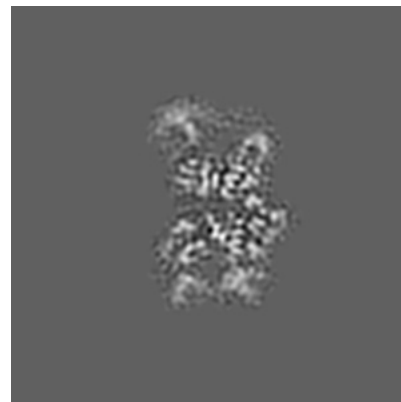
6.3.1 Primary map



X Index: 89



Y Index: 112

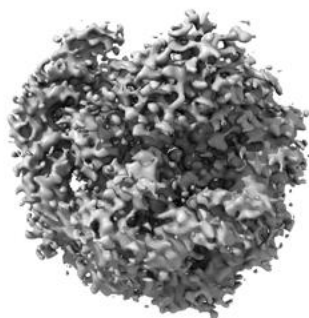


Z Index: 98

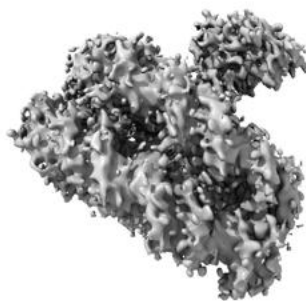
The images above show the largest variance slices of the map in three orthogonal directions.

6.4 Orthogonal surface views [i](#)

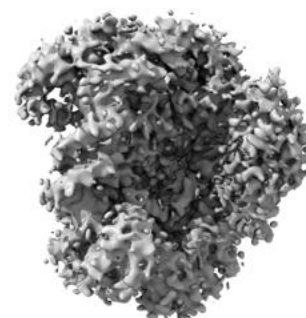
6.4.1 Primary map



X



Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.02. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

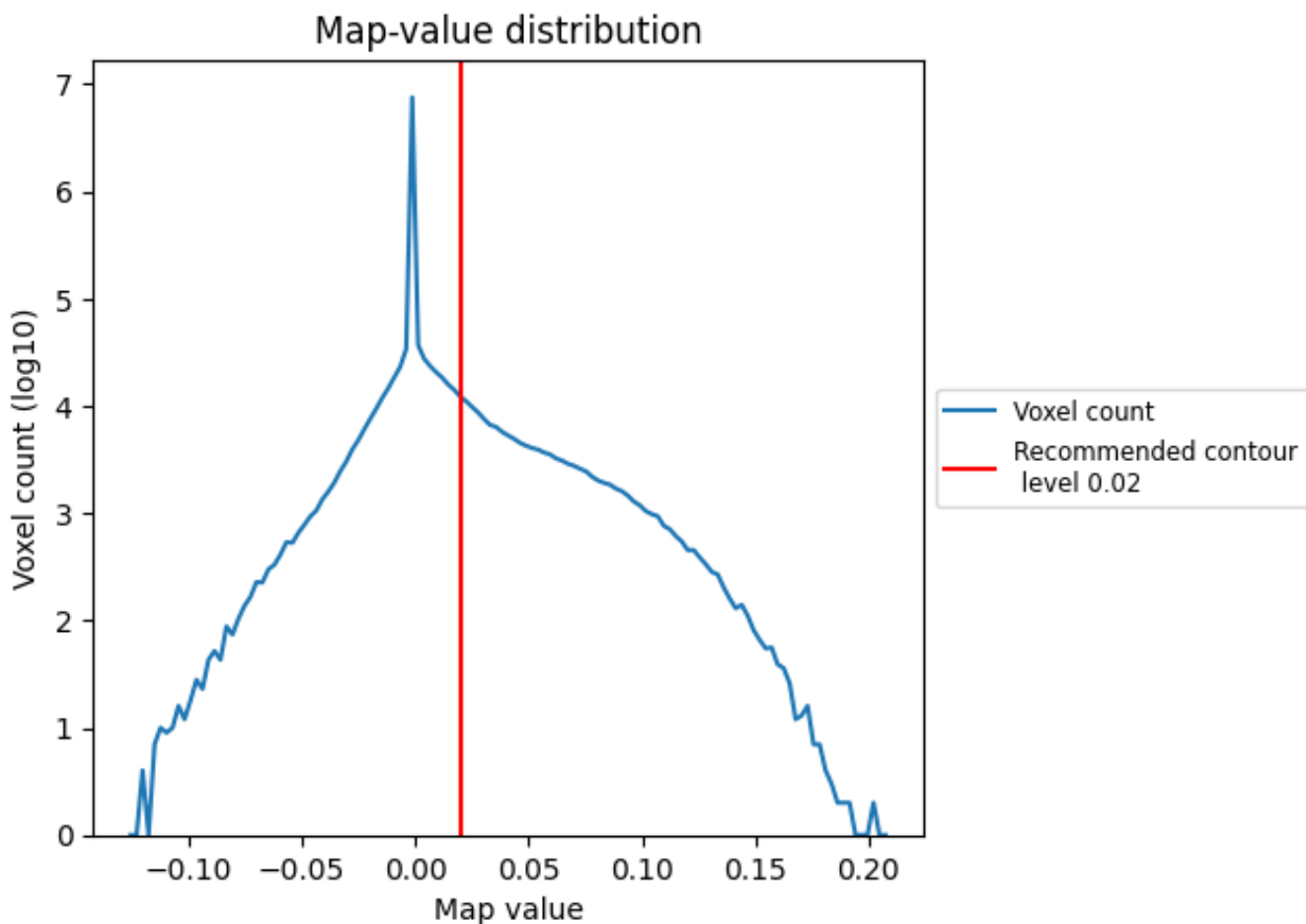
6.5 Mask visualisation

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

7 Map analysis [i](#)

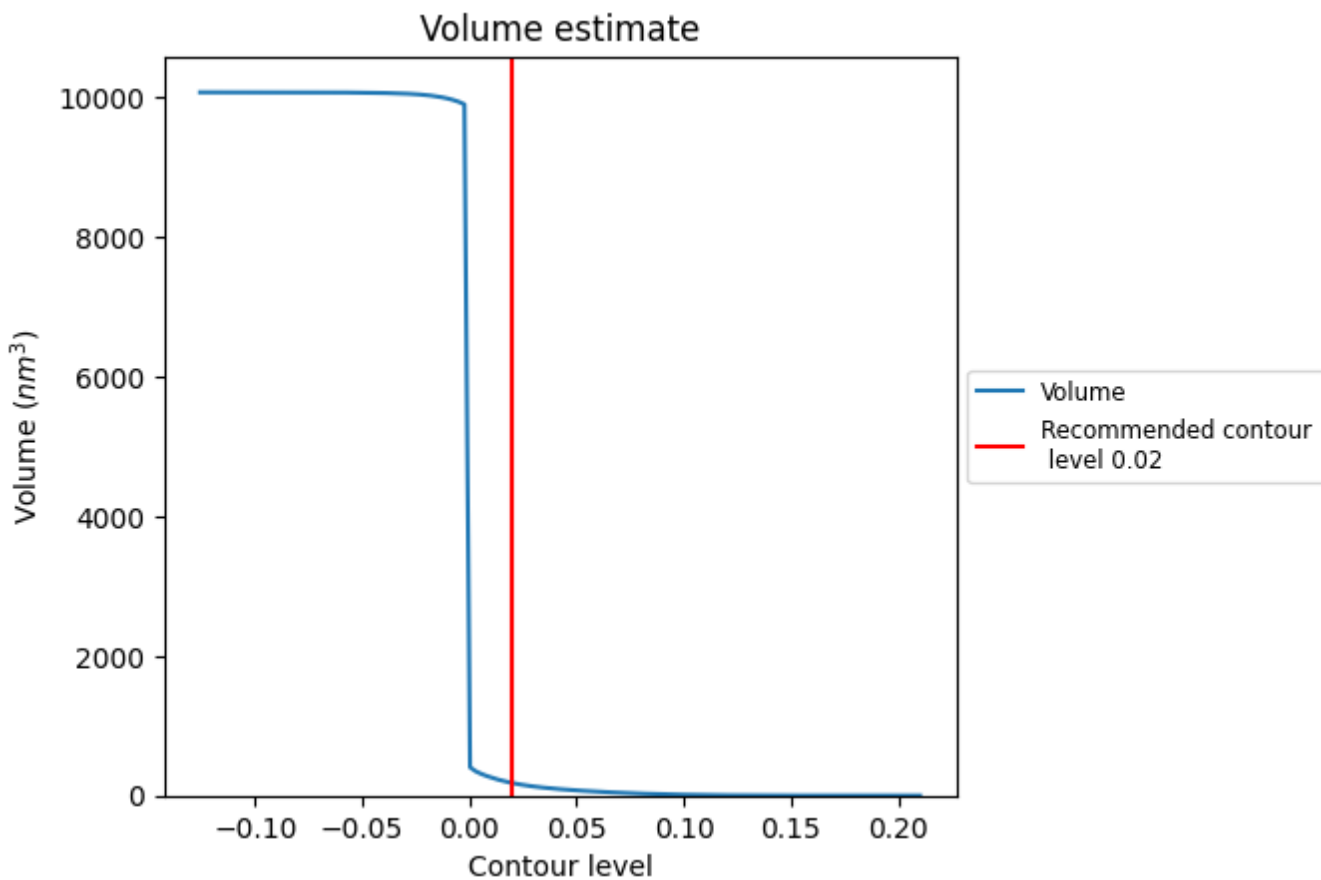
This section contains the results of statistical analysis of the map.

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.

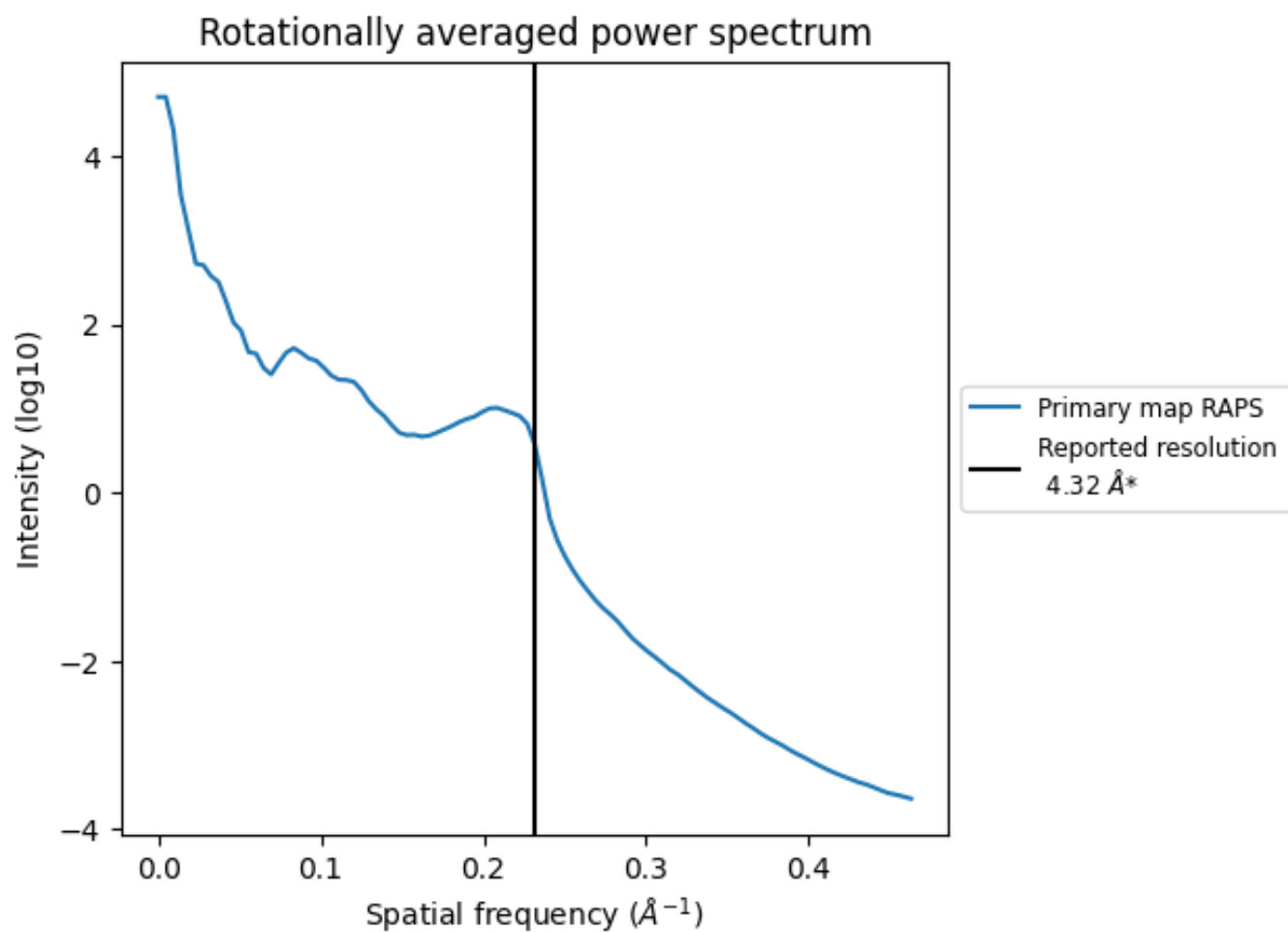
7.2 Volume estimate [i](#)



The volume at the recommended contour level is 180 nm³; this corresponds to an approximate mass of 163 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 0.231\AA^{-1}

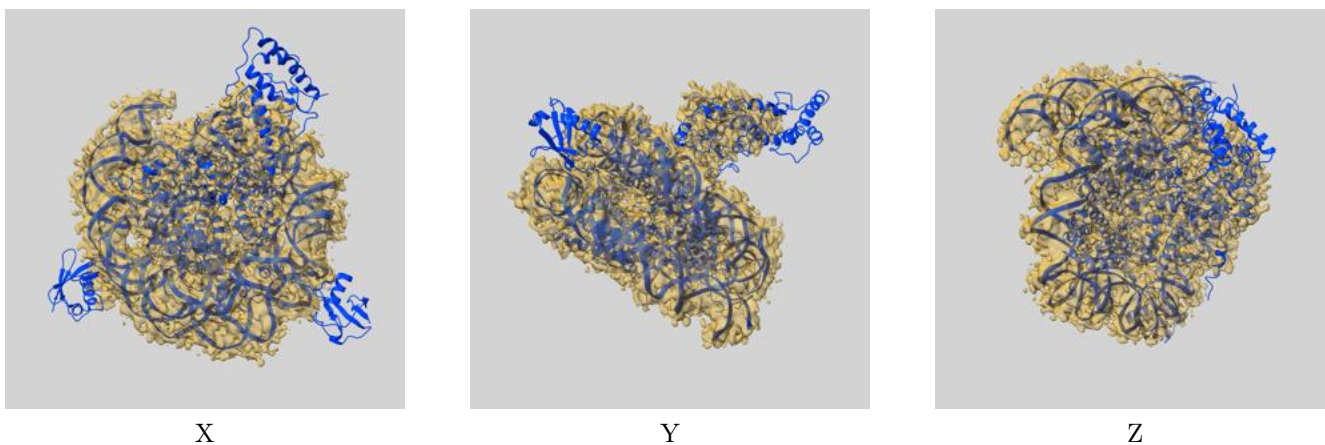
8 Fourier-Shell correlation

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

9 Map-model fit [i](#)

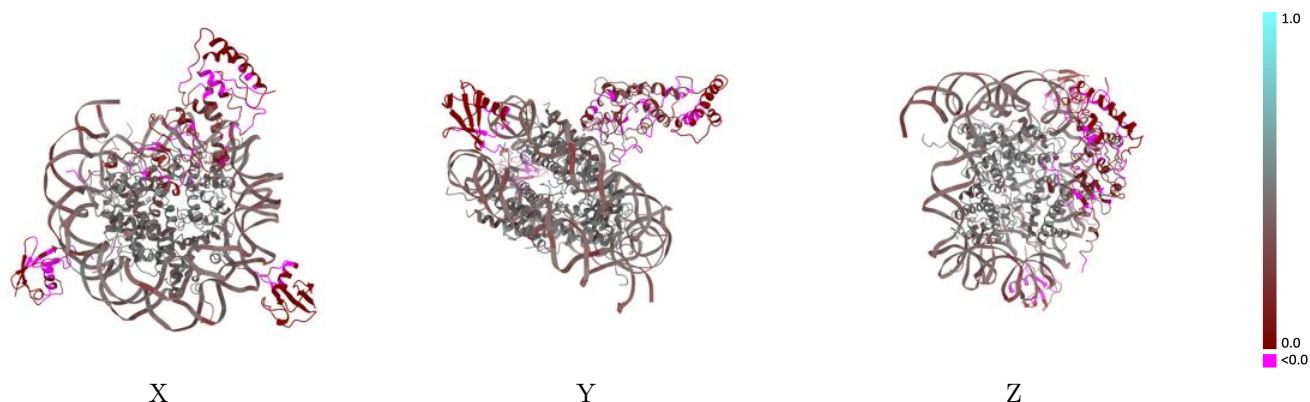
This section contains information regarding the fit between EMDB map EMD-30232 and PDB model 7BWD. Per-residue inclusion information can be found in section 3 on page 7.

9.1 Map-model overlay [i](#)



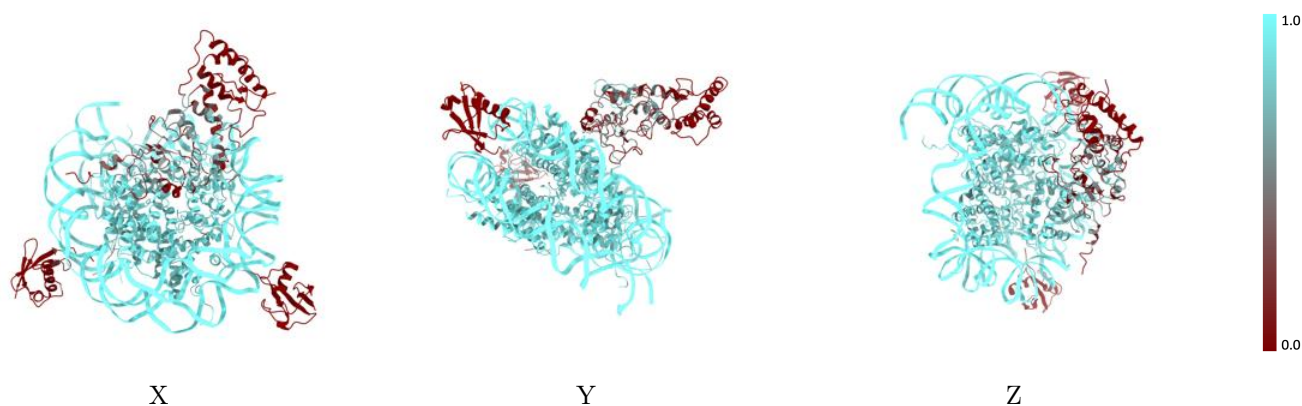
The images above show the 3D surface view of the map at the recommended contour level 0.02 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

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



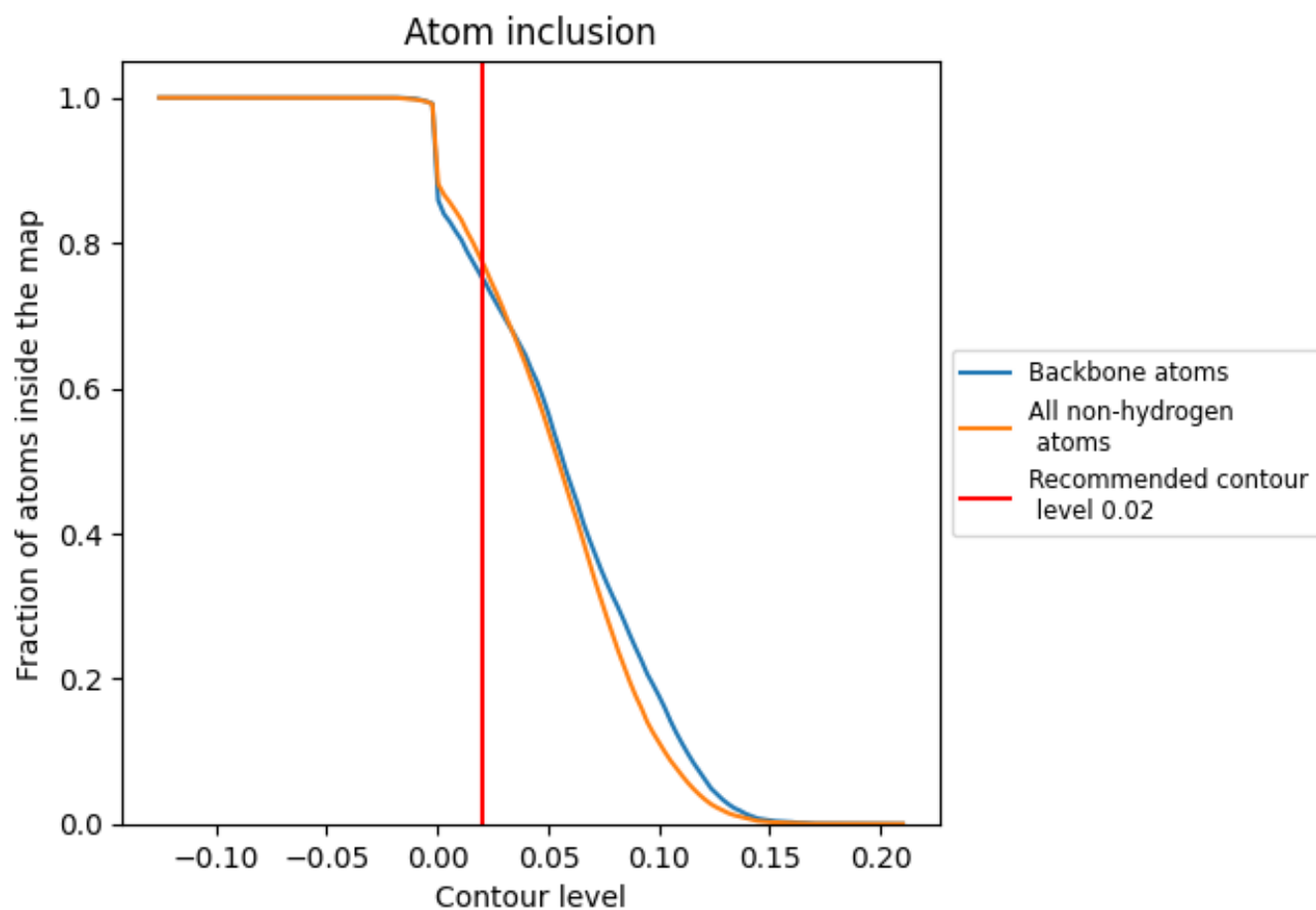
The images above show the model with each residue coloured according to its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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 (0.02).





























9.4 Atom inclusion [i](#)



At the recommended contour level, 75% of all backbone atoms, 77% 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 (0.02) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	 0.7749	 0.3320
A	 0.9376	 0.4540
B	 0.9225	 0.4620
C	 0.9476	 0.4560
D	 0.9260	 0.4400
E	 0.9338	 0.4570
F	 0.9034	 0.4530
G	 0.9364	 0.4570
H	 0.9204	 0.4230
I	 0.9876	 0.3730
J	 0.9917	 0.3650
K	 0.2704	 0.1320
L	 0.0376	 -0.0140
M	 0.0188	 -0.0160

