



wwPDB EM Validation Summary Report ⓘ

Dec 19, 2023 – 01:20 pm GMT

PDB ID : 8OEU
EMDB ID : EMD-16835
Title : Structure of the mammalian Pol II-SPT6 complex (composite structure, Structure 4)
Authors : Chen, Y.; Kokic, G.; Dienemann, C.; Dybkov, O.; Urlaub, H.; Cramer, P.
Deposited on : 2023-03-13
Resolution : 3.04 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:

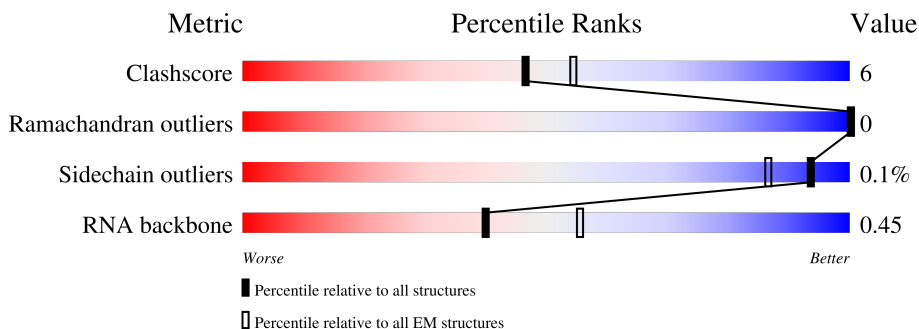
EMDB validation analysis : 0.0.1.dev70
MolProbity : 4.02b-467
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ : 1.9.9
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.36

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 3.04 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.








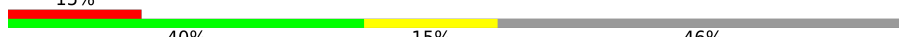



Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826
RNA backbone	4643	859

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 $\leq 5\%$. 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	1970	
2	B	1251	
3	C	275	
4	D	184	
5	E	210	
6	F	127	
7	G	172	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
8	H	150	 89% 10%
9	I	125	 80% 13% 7%
10	J	67	 87% 12%
11	K	117	 86% 12%
12	L	58	 67% 9% 24%
13	N	48	 15% 40% 15% 46%
14	P	46	 7% 22% 7% 72%
15	T	48	 17% 56% 23% 21%
16	S	1729	 37% 10% 53%

2 Entry composition [i](#)

There are 18 unique types of molecules in this entry. The entry contains 39178 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called DNA-directed RNA polymerase II subunit RPB1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	1409	11159	7024	2000	2064	71	0	0

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

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	B	1131	9047	5721	1592	1670	64	0	0

- Molecule 3 is a protein called DNA-directed RNA polymerase II subunit RPB3.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	C	258	2072	1300	356	410	6	0	0

- Molecule 4 is a protein called RNA polymerase II subunit D.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	D	118	967	608	167	188	4	0	0

- Molecule 5 is a protein called DNA-directed RNA polymerase II subunit E.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
5	E	209	1721	1089	300	324	8	0	0

- Molecule 6 is a protein called DNA-directed RNA polymerase II subunit F.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
6	F	78	626	401	106	114	5	0	0

- Molecule 7 is a protein called DNA-directed RNA polymerase II subunit RPB7.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
7	G	171	1347	872	218	249	8	0	0

- Molecule 8 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC3.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
8	H	148	1186	750	194	237	5	0	0

- Molecule 9 is a protein called DNA-directed RNA polymerase II subunit RPB9.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
9	I	116	932	577	165	179	11	0	0

- Molecule 10 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC5.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
10	J	66	524	339	88	91	6	0	0

- Molecule 11 is a protein called DNA-directed RNA polymerase II subunit RPB11-a.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
11	K	115	920	593	152	173	2	0	0

- Molecule 12 is a protein called RNA polymerase II subunit K.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
12	L	44	367	228	69	64	6	0	0

- Molecule 13 is a DNA chain called Non-template DNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
13	N	26	549	255	117	151	26	0	0

- Molecule 14 is a RNA chain called RNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
14	P	13	280	125	54	88	13	0	0

- Molecule 15 is a DNA chain called Template DNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
15	T	38	769	365	130	236	38	0	0

- Molecule 16 is a protein called Transcription elongation factor SPT6.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
16	S	819	6703	4260	1164	1246	33	0	0

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
S	-2	SER	-	expression tag	UNP Q7KZ85
S	-1	ASN	-	expression tag	UNP Q7KZ85
S	0	ALA	-	expression tag	UNP Q7KZ85

- Molecule 17 is MAGNESIUM ION (three-letter code: MG) (formula: Mg) (labeled as "Ligand of Interest" by depositor).

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

- Molecule 18 is ZINC ION (three-letter code: ZN) (formula: Zn) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms		AltConf
18	A	2	Total	Zn	0
			2	2	
18	B	1	Total	Zn	0
			1	1	
18	C	1	Total	Zn	0
			1	1	
18	I	2	Total	Zn	0
			2	2	
18	J	1	Total	Zn	0
			1	1	

Continued on next page...

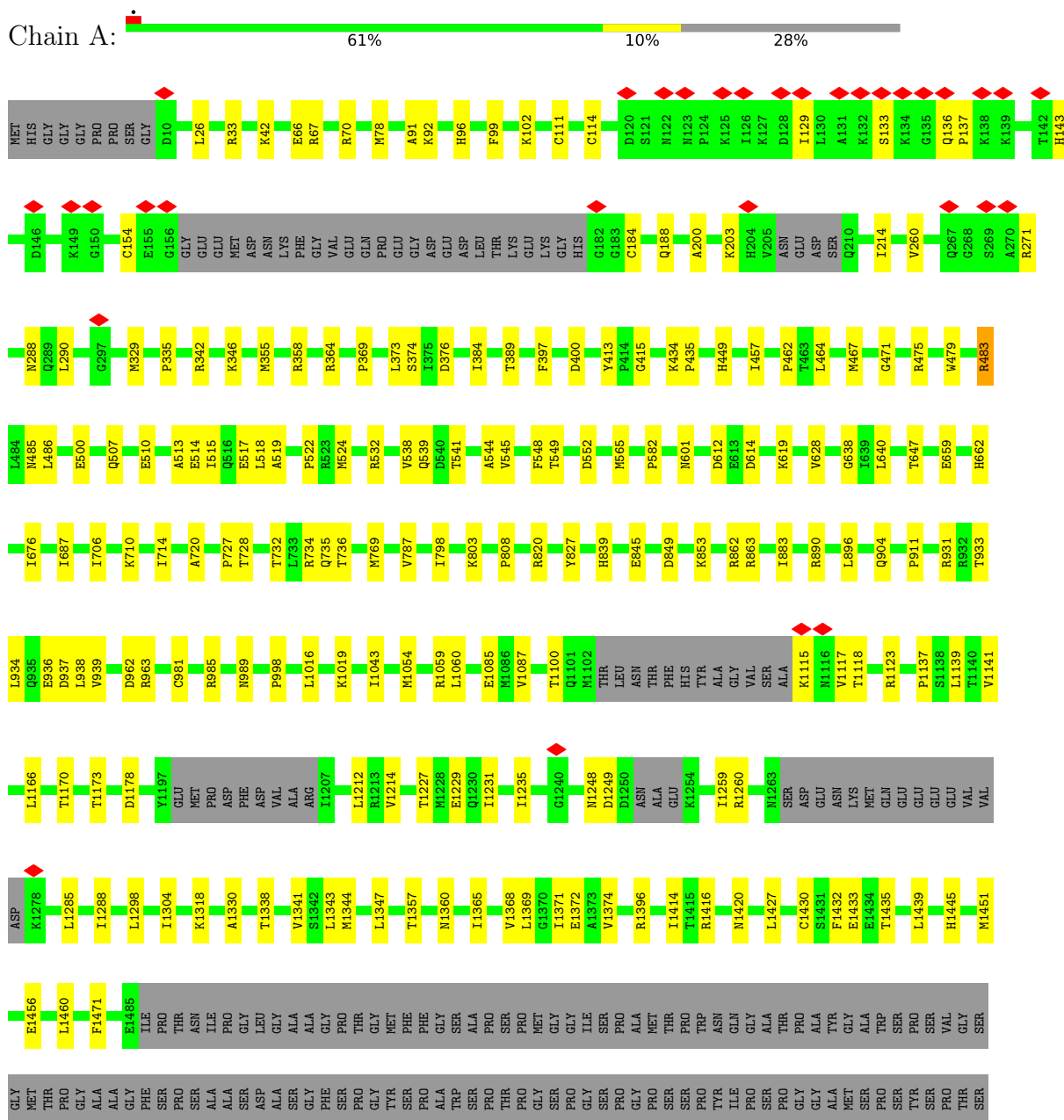
Continued from previous page...

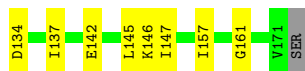
Mol	Chain	Residues	Atoms		AltConf
			Total	Zn	
18	L	1	1	1	0

3 Residue-property plots

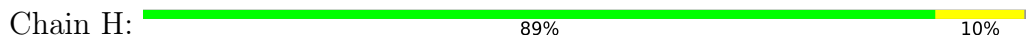
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and atom inclusion 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: DNA-directed RNA polymerase II subunit RPB1

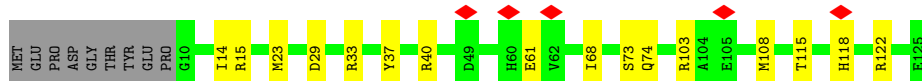
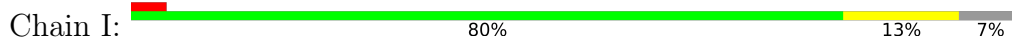




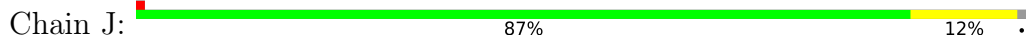
- Molecule 8: DNA-directed RNA polymerases I, II, and III subunit RPABC3



- Molecule 9: DNA-directed RNA polymerase II subunit RPB9



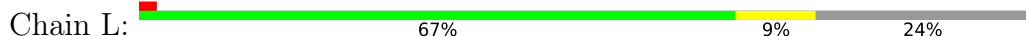
- Molecule 10: DNA-directed RNA polymerases I, II, and III subunit RPABC5



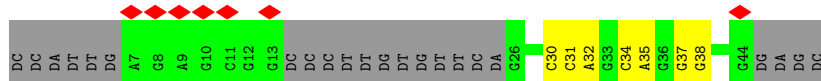
- Molecule 11: DNA-directed RNA polymerase II subunit RPB11-a



- Molecule 12: RNA polymerase II subunit K



- Molecule 13: Non-template DNA



4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	174029	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 ($e^-/\text{\AA}^2$)	40.09	Depositor
Minimum defocus (nm)	350	Depositor
Maximum defocus (nm)	7500	Depositor
Magnification	Not provided	
Image detector	GATAN K3 BIOQUANTUM (6k x 4k)	Depositor
Maximum map value	59.731	Depositor
Minimum map value	-28.119	Depositor
Average map value	0.011	Depositor
Map value standard deviation	1.105	Depositor
Recommended contour level	3.5	Depositor
Map size (Å)	461.99997, 461.99997, 461.99997	wwPDB
Map dimensions	440, 440, 440	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.05, 1.05, 1.05	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: 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	A	0.32	0/11360	0.52	2/15328 (0.0%)
2	B	0.35	0/9227	0.53	2/12454 (0.0%)
3	C	0.36	0/2115	0.49	0/2873
4	D	0.36	0/979	0.71	1/1312 (0.1%)
5	E	0.32	0/1752	0.52	0/2366
6	F	0.35	0/636	0.55	0/859
7	G	0.46	0/1378	0.67	1/1870 (0.1%)
8	H	0.33	0/1207	0.52	0/1628
9	I	0.32	0/954	0.54	0/1293
10	J	0.36	0/533	0.51	0/719
11	K	0.33	0/939	0.54	1/1271 (0.1%)
12	L	0.35	0/372	0.57	0/493
13	N	0.50	0/619	0.74	0/954
14	P	0.33	0/313	0.79	0/486
15	T	0.57	0/857	0.95	0/1319
16	S	0.36	0/6826	0.57	0/9194
All	All	0.35	0/40067	0.56	7/54419 (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
1	A	0	2
7	G	0	1
16	S	0	1
All	All	0	4

There are no bond length outliers.

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	483	ARG	CG-CD-NE	7.76	128.09	111.80
1	A	883	ILE	CG1-CB-CG2	-6.33	97.47	111.40
11	K	53	ASP	CB-CG-OD1	5.66	123.40	118.30
2	B	141	PRO	CA-N-CD	-5.16	104.28	111.50
4	D	130	LEU	CB-CG-CD1	5.11	119.69	111.00

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	538	VAL	Peptide
1	A	539	GLN	Peptide
7	G	124	ASN	Peptide
16	S	1156	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	11159	0	11304	130	0
2	B	9047	0	9080	126	0
3	C	2072	0	2019	36	0
4	D	967	0	973	19	0
5	E	1721	0	1737	23	0
6	F	626	0	657	5	0
7	G	1347	0	1347	29	0
8	H	1186	0	1147	10	0
9	I	932	0	856	14	0
10	J	524	0	540	7	0
11	K	920	0	942	9	0
12	L	367	0	367	5	0
13	N	549	0	289	5	0
14	P	280	0	142	1	0
15	T	769	0	429	10	0
16	S	6703	0	6683	120	0
17	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
18	A	2	0	0	0	0
18	B	1	0	0	0	0
18	C	1	0	0	0	0
18	I	2	0	0	0	0
18	J	1	0	0	0	0
18	L	1	0	0	0	0
All	All	39178	0	38512	490	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 6.

The worst 5 of 490 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)
16:S:836:LEU:HD12	16:S:867:ILE:CD1	1.85	1.05
16:S:836:LEU:CD1	16:S:867:ILE:CD1	2.52	0.87
5:E:115:LYS:NZ	5:E:129:GLN:OE1	2.07	0.86
16:S:836:LEU:HD12	16:S:867:ILE:HD13	1.56	0.85
16:S:836:LEU:CD1	16:S:867:ILE:HD13	2.11	0.78

There are no symmetry-related clashes.

5.3 Torsion angles [i](#)

5.3.1 Protein backbone [i](#)

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	1395/1970 (71%)	1362 (98%)	33 (2%)	0	100	100
2	B	1123/1251 (90%)	1090 (97%)	33 (3%)	0	100	100
3	C	254/275 (92%)	252 (99%)	2 (1%)	0	100	100
4	D	114/184 (62%)	109 (96%)	5 (4%)	0	100	100
5	E	207/210 (99%)	202 (98%)	5 (2%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
6	F	76/127 (60%)	75 (99%)	1 (1%)	0	100	100
7	G	169/172 (98%)	148 (88%)	21 (12%)	0	100	100
8	H	146/150 (97%)	144 (99%)	2 (1%)	0	100	100
9	I	114/125 (91%)	111 (97%)	3 (3%)	0	100	100
10	J	64/67 (96%)	63 (98%)	1 (2%)	0	100	100
11	K	113/117 (97%)	110 (97%)	3 (3%)	0	100	100
12	L	42/58 (72%)	41 (98%)	1 (2%)	0	100	100
16	S	795/1729 (46%)	756 (95%)	39 (5%)	0	100	100
All	All	4612/6435 (72%)	4463 (97%)	149 (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	
1	A	1239/1749 (71%)	1239 (100%)	0	100	100
2	B	991/1084 (91%)	989 (100%)	2 (0%)	93	98
3	C	235/252 (93%)	235 (100%)	0	100	100
4	D	109/160 (68%)	109 (100%)	0	100	100
5	E	191/192 (100%)	191 (100%)	0	100	100
6	F	68/111 (61%)	68 (100%)	0	100	100
7	G	151/153 (99%)	151 (100%)	0	100	100
8	H	129/131 (98%)	129 (100%)	0	100	100
9	I	102/112 (91%)	102 (100%)	0	100	100
10	J	55/56 (98%)	55 (100%)	0	100	100
11	K	104/106 (98%)	104 (100%)	0	100	100
12	L	40/55 (73%)	40 (100%)	0	100	100
16	S	726/1524 (48%)	725 (100%)	1 (0%)	93	98

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	4140/5685 (73%)	4137 (100%)	3 (0%)	93 98

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	762	LYS
2	B	764	VAL
16	S	869	HIS

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

Mol	Chain	Res	Type
1	A	1445	HIS
2	B	538	GLN
2	B	902	GLN
2	B	990	GLN
16	S	608	GLN

5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
14	P	12/46 (26%)	2 (16%)	0

All (2) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
14	P	36	G
14	P	37	G

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

Of 9 ligands modelled in this entry, 9 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

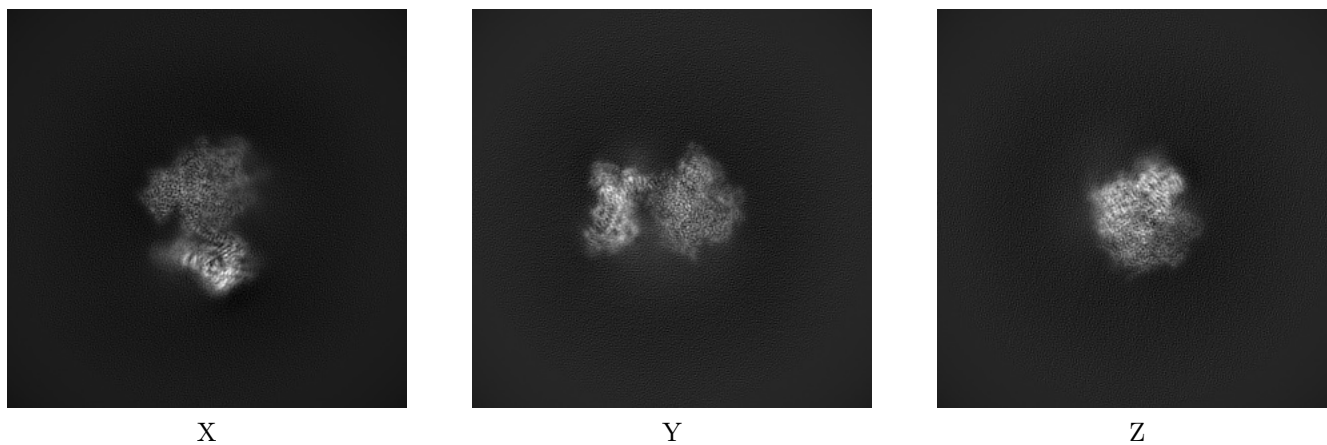
6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-16835. 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](#)

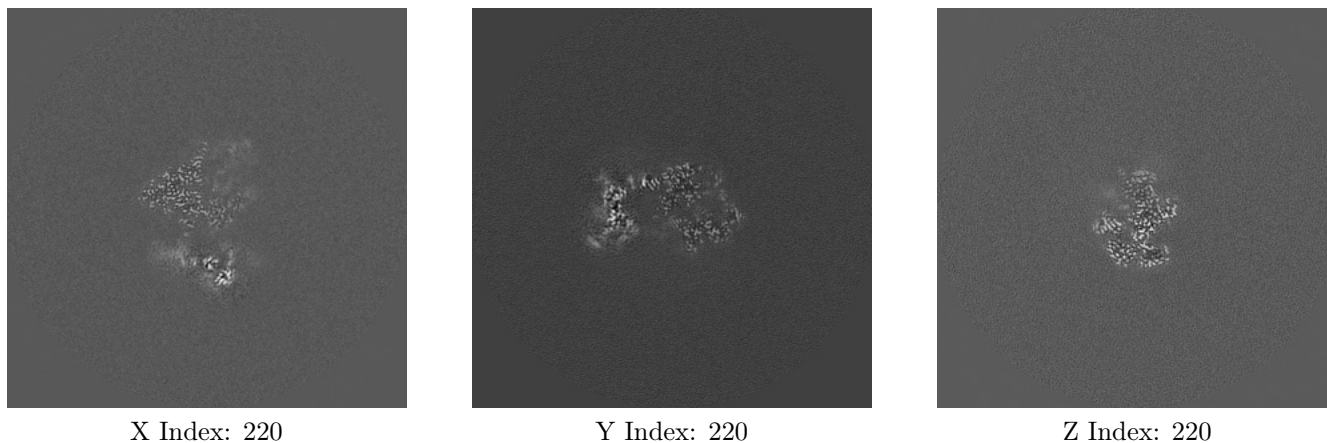
6.1.1 Primary map



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

6.2 Central slices [i](#)

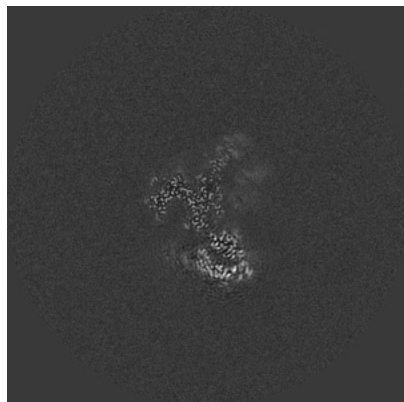
6.2.1 Primary map



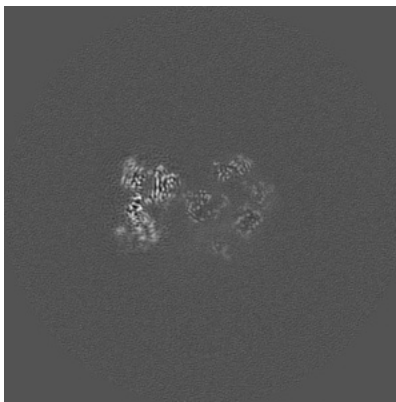
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices [\(i\)](#)

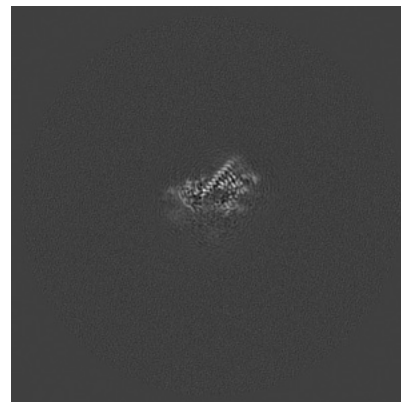
6.3.1 Primary map



X Index: 240



Y Index: 241

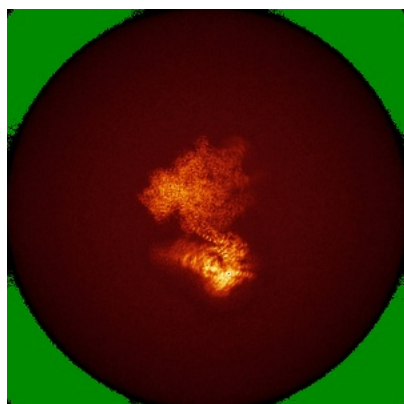


Z Index: 148

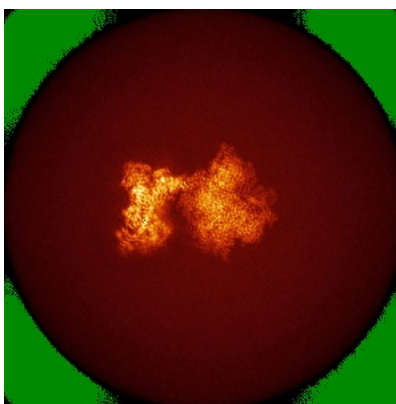
The images above show the largest variance slices of the map in three orthogonal directions.

6.4 Orthogonal standard-deviation projections (False-color) [\(i\)](#)

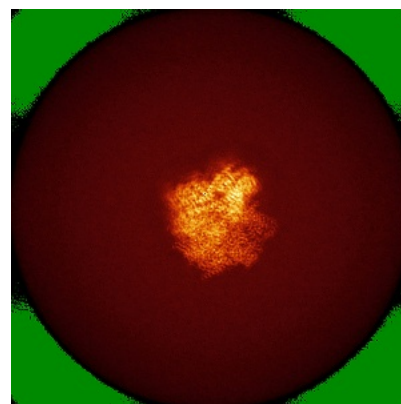
6.4.1 Primary map



X



Y

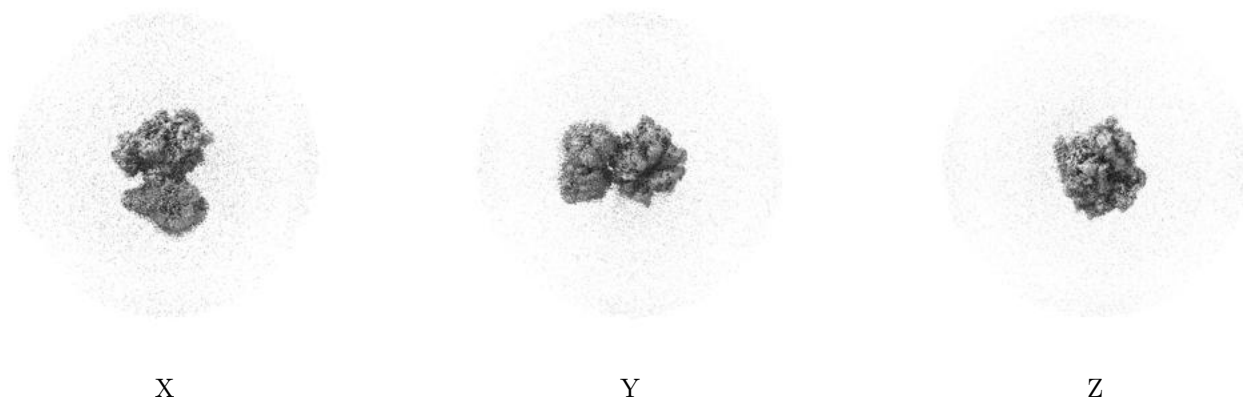


Z

The images above show the map standard deviation projections with false color in three orthogonal directions. Minimum values are shown in green, max in blue, and dark to light orange shades represent small to large values respectively.

6.5 Orthogonal surface views [i](#)

6.5.1 Primary map



The images above show the 3D surface view of the map at the recommended contour level 3.5. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

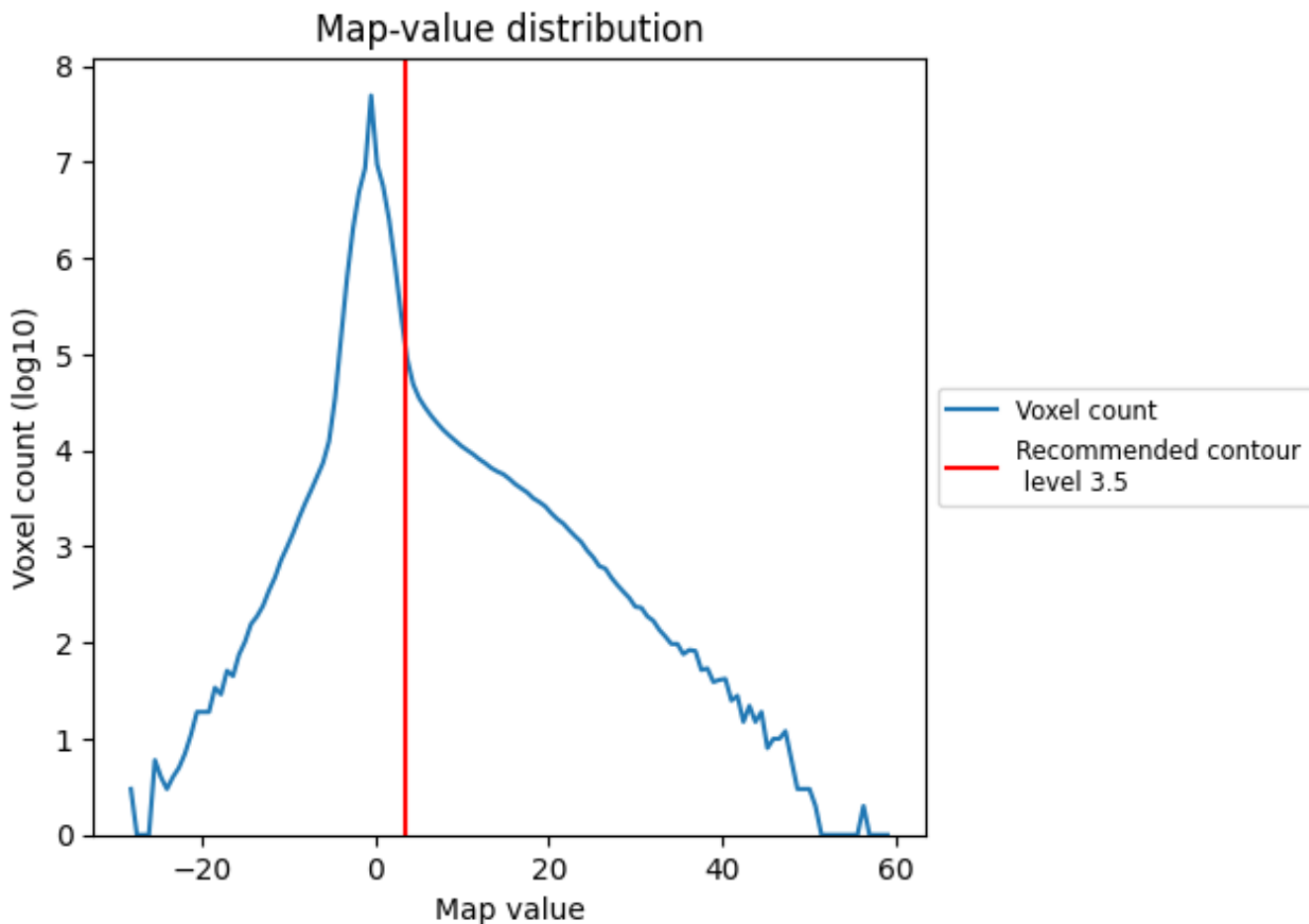
6.6 Mask visualisation [i](#)

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

7 Map analysis [i](#)

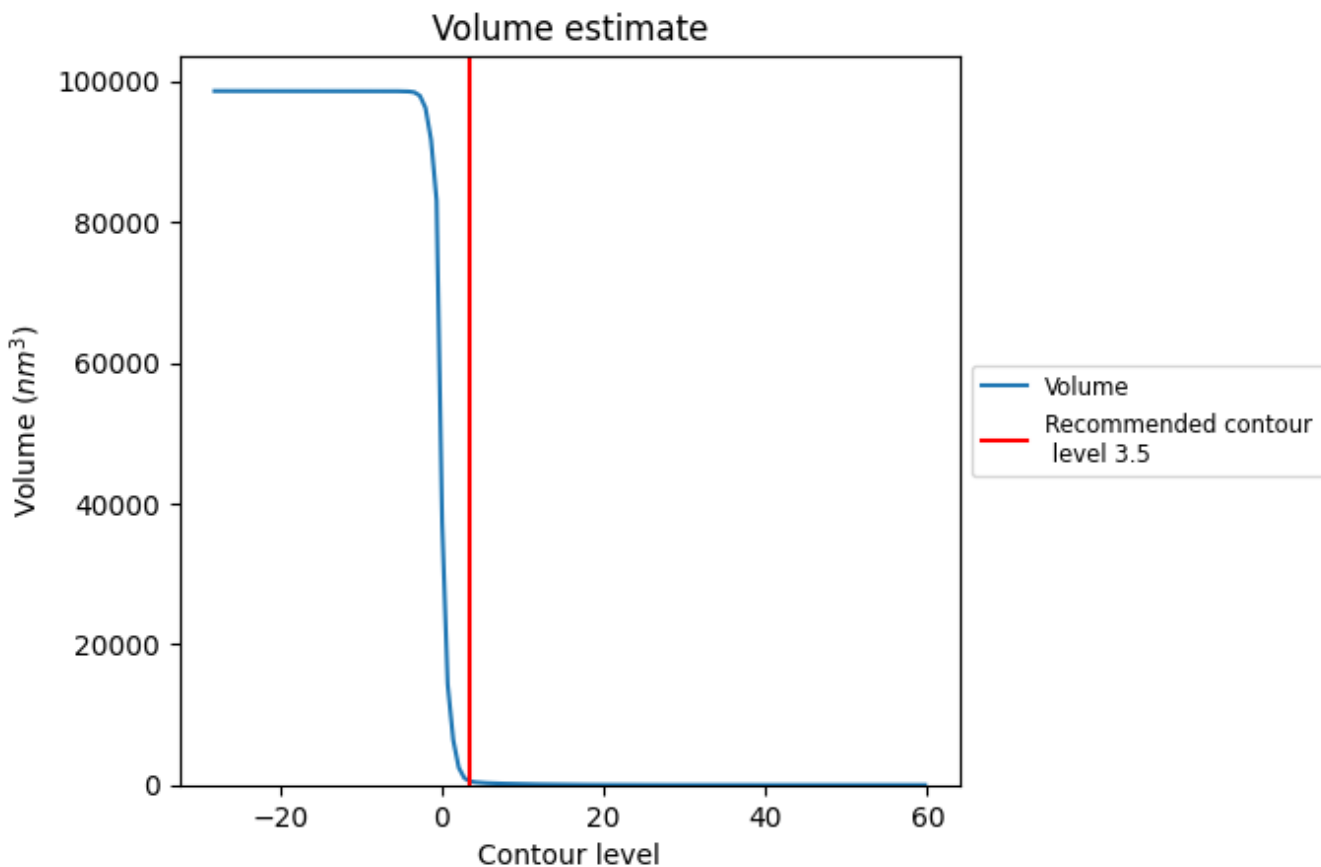
This section contains the results of statistical analysis of the map.

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.

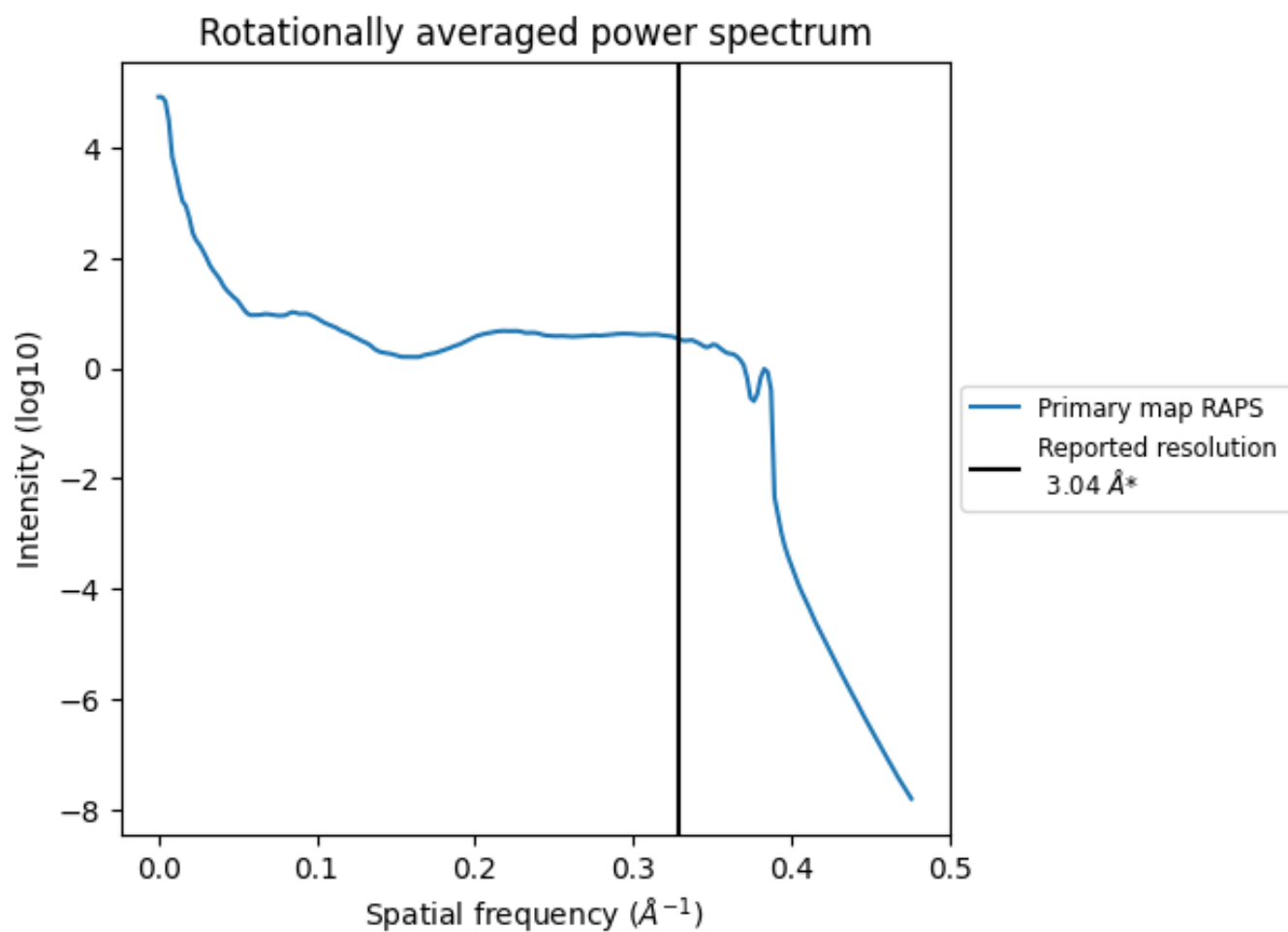
7.2 Volume estimate [i](#)



The volume at the recommended contour level is 516 nm^3 ; this corresponds to an approximate mass of 466 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 0.329\AA^{-1}

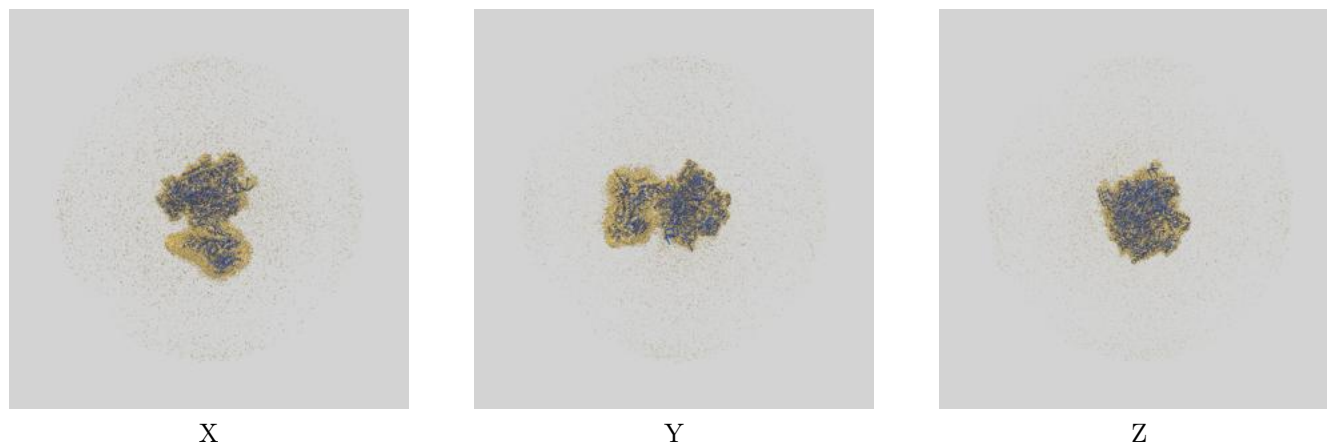
8 Fourier-Shell correlation

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

9 Map-model fit [i](#)

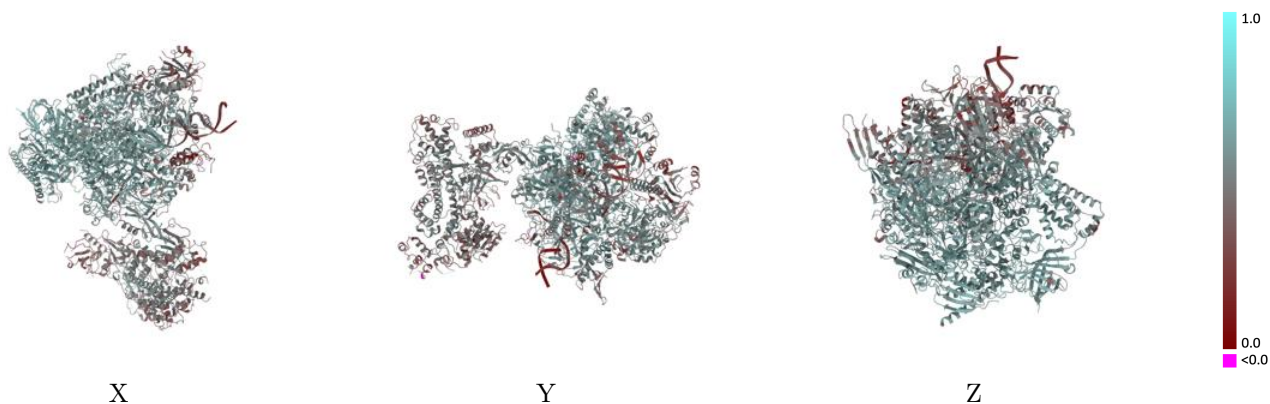
This section contains information regarding the fit between EMDB map EMD-16835 and PDB model 8OEU. Per-residue inclusion information can be found in section 3 on page 8.

9.1 Map-model overlay [i](#)



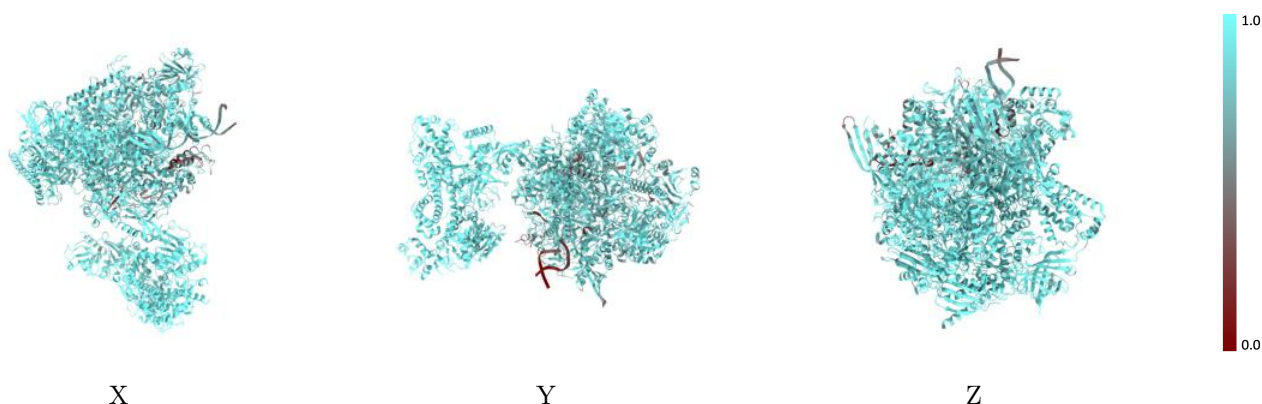
The images above show the 3D surface view of the map at the recommended contour level 3.5 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

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



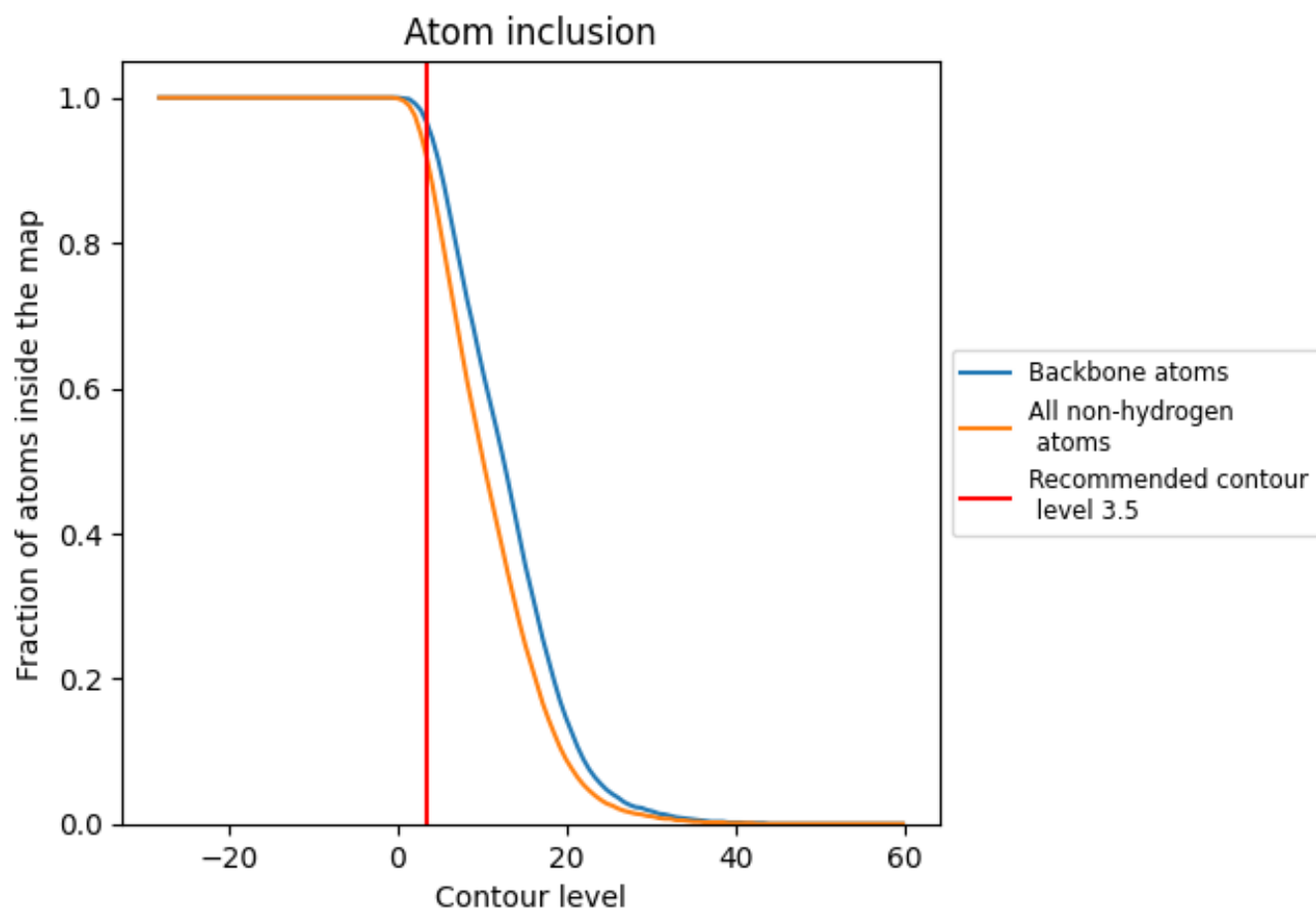
The images above show the model with each residue coloured according to its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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 (3.5).



































9.4 Atom inclusion [i](#)



At the recommended contour level, 96% of all backbone atoms, 92% 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 (3.5) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	 0.9160	 0.5260
A	 0.8980	 0.5530
B	 0.9080	 0.5640
C	 0.9510	 0.6010
D	 0.9620	 0.4280
E	 0.9120	 0.5260
F	 0.9390	 0.5780
G	 0.9610	 0.4980
H	 0.9430	 0.5860
I	 0.8530	 0.4750
J	 0.9570	 0.6030
K	 0.9550	 0.6010
L	 0.8890	 0.5220
N	 0.6140	 0.2890
P	 0.7890	 0.5060
S	 0.9800	 0.4350
T	 0.7060	 0.4170

