



wwPDB EM Validation Summary Report ⓘ

Nov 7, 2022 – 12:26 PM EST

PDB ID : 6DRD
EMDB ID : EMD-7997
Title : RNA Pol II(G)
Authors : Yu, X.; Jishage, M.; Shi, Y.; Ganesan, S.; Sali, A.; Chait, B.T.; Asturias, F.;
Roeder, R.G.
Deposited on : 2018-06-11
Resolution : 3.90 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

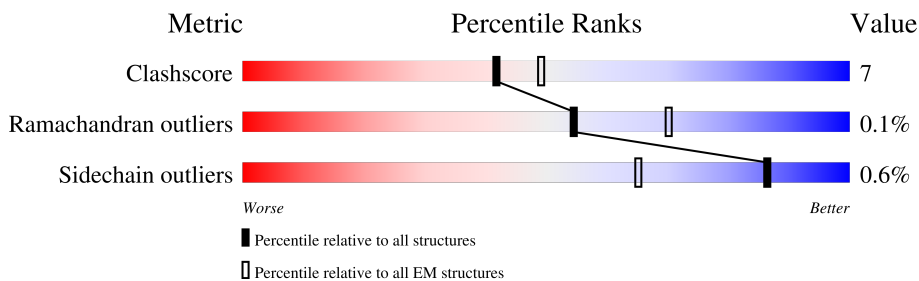
1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 3.90 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



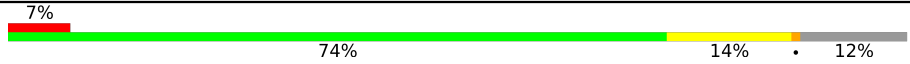

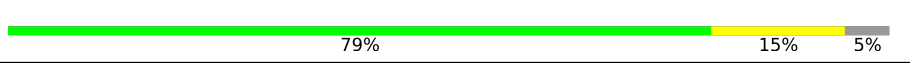


Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826

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	1174	
3	C	275	
4	D	142	
5	E	210	
6	F	127	
7	G	172	
8	H	150	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
9	I	125	
10	J	67	
11	K	117	
12	L	58	
13	M	62	

2 Entry composition [i](#)

There are 14 unique types of molecules in this entry. The entry contains 26708 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	1000	7989	5045	1409	1493	42	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1835	THR	ALA	conflict	UNP P24928

- Molecule 2 is a protein called DNA-directed RNA polymerase II subunit RPB2.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	B	1009	8079	5127	1403	1499	50	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	262	2104	1321	361	416	6	0	0

- Molecule 4 is a protein called DNA-directed RNA polymerase II subunit RPB4.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	D	128	1005	632	172	197	4	0	0

- Molecule 5 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
5	E	185	1529	976	267	279	7	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	44	PHE	SER	conflict	UNP P19388
E	132	GLU	GLN	conflict	UNP P19388
E	157	SER	THR	conflict	UNP P19388
E	186	ARG	LYS	conflict	UNP P19388

- Molecule 6 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC2.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
6	F	68	549	356	95	93	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	167	1307	850	211	238	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	146	1176	744	192	235	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	110	897	554	159	173	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	64	507	328	86	87	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	111	890	576	147	166	1	0	0

- Molecule 12 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC4.

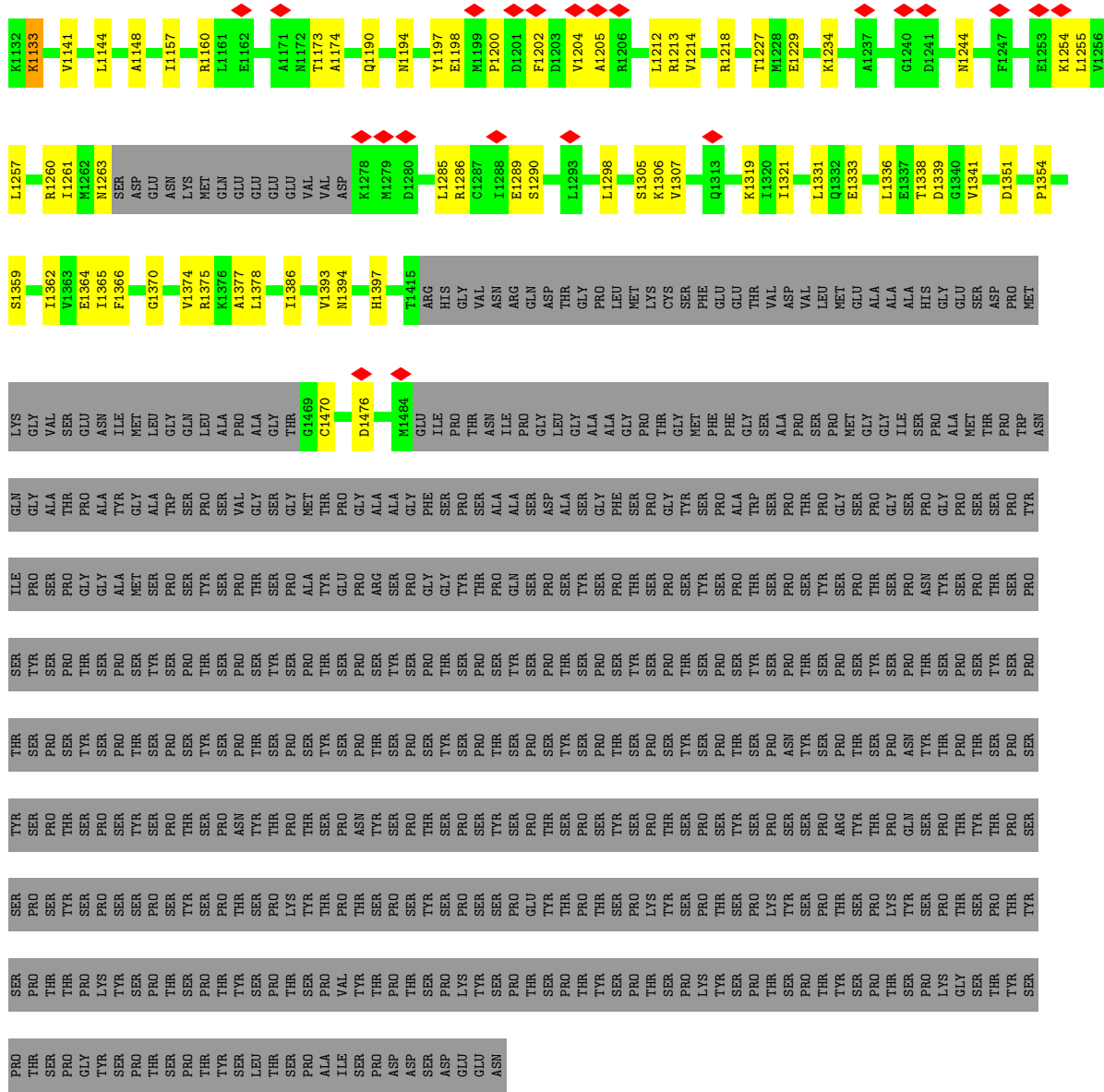
Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
12	L	37	312	195	58	53	6	0	0

- Molecule 13 is a protein called DNA-directed RNA polymerase II subunit GRINL1A.

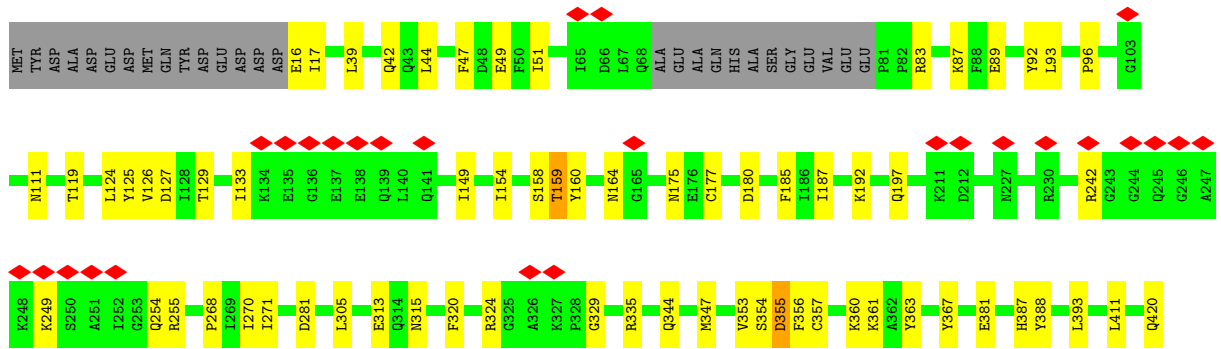
Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
13	M	62	359	219	67	73	0	0

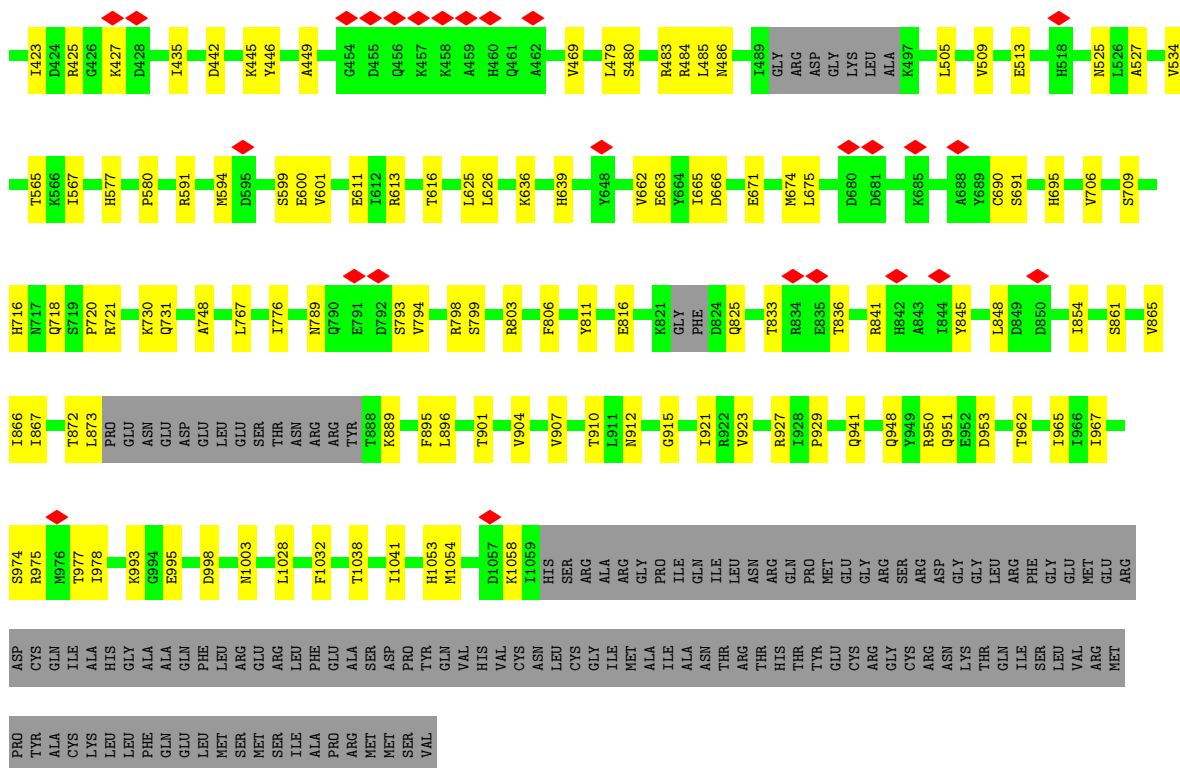
- Molecule 14 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		AltConf
14	C	1	Total	Zn	0
			1	1	
14	I	2	Total	Zn	0
			2	2	
14	J	1	Total	Zn	0
			1	1	
14	L	1	Total	Zn	0
			1	1	

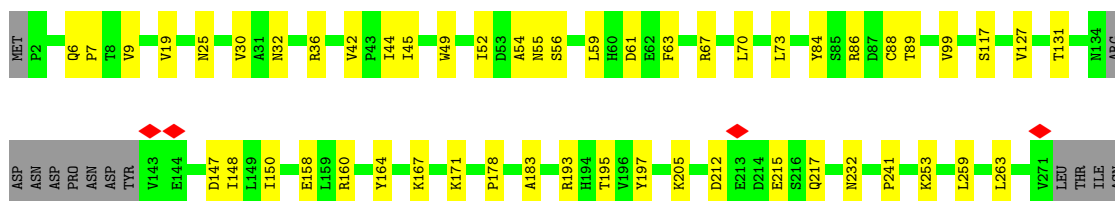
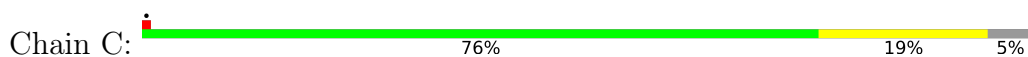


● Molecule 2: DNA-directed RNA polymerase II subunit RPB2

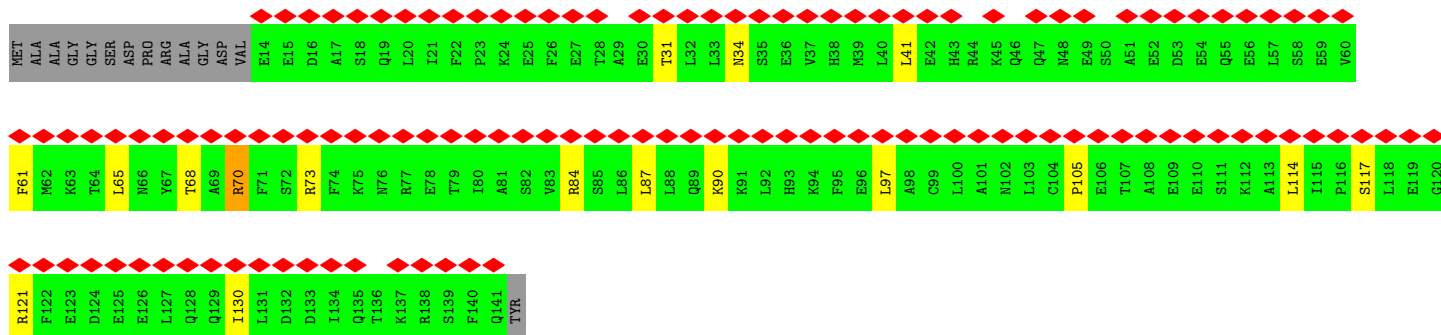
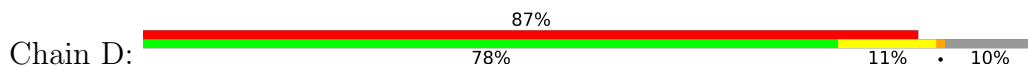




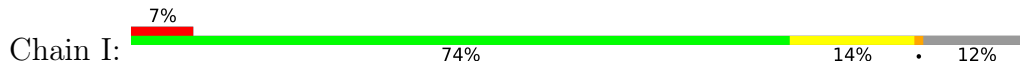
• Molecule 3: DNA-directed RNA polymerase II subunit RPB3



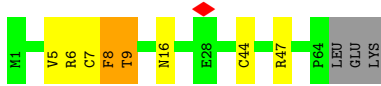
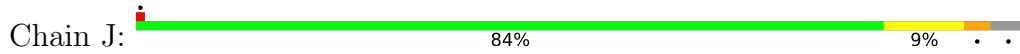
• Molecule 4: DNA-directed RNA polymerase II subunit RPB4



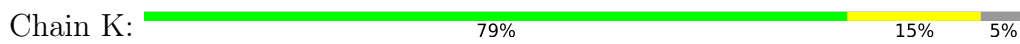
• Molecule 5: DNA-directed RNA polymerases I, II, and III subunit RPABC1



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



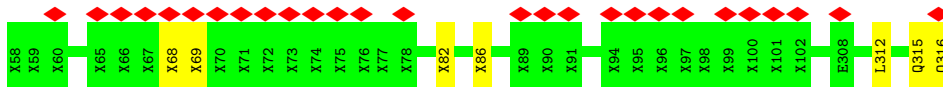
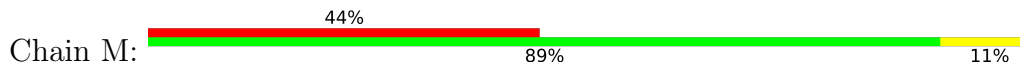
- Molecule 11: DNA-directed RNA polymerase II subunit RPB11-a



- Molecule 12: DNA-directed RNA polymerases I, II, and III subunit RPABC4



- Molecule 13: DNA-directed RNA polymerase II subunit GRINL1A



4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	141619	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	Depositor
Minimum defocus (nm)	Not provided	
Maximum defocus (nm)	Not provided	
Magnification	Not provided	
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor
Maximum map value	0.238	Depositor
Minimum map value	-0.140	Depositor
Average map value	0.001	Depositor
Map value standard deviation	0.007	Depositor
Recommended contour level	0.035	Depositor
Map size (Å)	293.44, 293.44, 293.44	wwPDB
Map dimensions	224, 224, 224	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.3100001, 1.3100001, 1.3100001	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

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	0.46	0/8130	0.62	0/10980
2	B	0.55	0/8240	0.67	0/11128
3	C	0.54	0/2147	0.64	0/2917
4	D	0.26	0/1019	0.51	0/1374
5	E	0.42	0/1552	0.61	0/2091
6	F	0.40	0/558	0.62	0/752
7	G	0.32	0/1336	0.56	0/1811
8	H	0.48	0/1197	0.65	0/1614
9	I	0.41	0/917	0.60	0/1241
10	J	0.59	0/516	0.72	0/696
11	K	0.46	0/909	0.64	0/1232
12	L	0.48	0/315	0.67	0/416
13	M	0.30	0/134	0.81	0/179
All	All	0.48	0/26970	0.63	0/36431

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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	7989	0	8058	125	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	8079	0	8123	121	0
3	C	2104	0	2053	37	0
4	D	1005	0	964	15	0
5	E	1529	0	1565	31	0
6	F	549	0	590	8	0
7	G	1307	0	1305	20	0
8	H	1176	0	1137	28	0
9	I	897	0	831	15	0
10	J	507	0	523	9	0
11	K	890	0	913	14	0
12	L	312	0	311	5	0
13	M	359	0	195	3	0
14	C	1	0	0	0	0
14	I	2	0	0	0	0
14	J	1	0	0	0	0
14	L	1	0	0	0	0
All	All	26708	0	26568	387	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 387 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)
10:J:9:THR:HB	10:J:44:CYS:SG	1.82	1.20
4:D:87:LEU:HB3	4:D:97:LEU:HD11	1.15	1.09
8:H:96:VAL:HG12	8:H:116:VAL:CG2	1.87	1.05
8:H:96:VAL:CG1	8:H:116:VAL:HG22	1.90	1.01
1:A:863:ARG:HH21	1:A:1128:ILE:HG21	1.24	1.01

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	984/1970 (50%)	876 (89%)	106 (11%)	2 (0%)	47	79
2	B	999/1174 (85%)	870 (87%)	128 (13%)	1 (0%)	51	84
3	C	258/275 (94%)	222 (86%)	36 (14%)	0	100	100
4	D	126/142 (89%)	121 (96%)	5 (4%)	0	100	100
5	E	177/210 (84%)	166 (94%)	11 (6%)	0	100	100
6	F	64/127 (50%)	58 (91%)	6 (9%)	0	100	100
7	G	163/172 (95%)	147 (90%)	16 (10%)	0	100	100
8	H	144/150 (96%)	130 (90%)	14 (10%)	0	100	100
9	I	106/125 (85%)	90 (85%)	16 (15%)	0	100	100
10	J	62/67 (92%)	54 (87%)	7 (11%)	1 (2%)	9	44
11	K	109/117 (93%)	103 (94%)	6 (6%)	0	100	100
12	L	33/58 (57%)	26 (79%)	7 (21%)	0	100	100
13	M	15/62 (24%)	12 (80%)	3 (20%)	0	100	100
All	All	3240/4649 (70%)	2875 (89%)	361 (11%)	4 (0%)	54	84

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	556	GLU
1	A	582	PRO
2	B	159	THR
10	J	9	THR

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	890/1749 (51%)	885 (99%)	5 (1%)	86	91
2	B	889/1028 (86%)	885 (100%)	4 (0%)	91	94

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
3	C	239/252 (95%)	239 (100%)	0	100	100
4	D	106/126 (84%)	105 (99%)	1 (1%)	78	87
5	E	169/192 (88%)	168 (99%)	1 (1%)	86	91
6	F	59/111 (53%)	59 (100%)	0	100	100
7	G	144/153 (94%)	144 (100%)	0	100	100
8	H	129/131 (98%)	128 (99%)	1 (1%)	81	89
9	I	101/112 (90%)	100 (99%)	1 (1%)	76	86
10	J	53/56 (95%)	51 (96%)	2 (4%)	33	59
11	K	101/106 (95%)	101 (100%)	0	100	100
12	L	34/55 (62%)	33 (97%)	1 (3%)	42	65
13	M	16/16 (100%)	15 (94%)	1 (6%)	18	47
All	All	2930/4087 (72%)	2913 (99%)	17 (1%)	86	91

5 of 17 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
10	J	8	PHE
13	M	312	LEU
2	B	355	ASP
2	B	356	PHE
4	D	70	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 24 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	699	HIS
4	D	93	HIS
2	B	941	GLN
5	E	65	ASN
1	A	1077	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 5 ligands modelled in this entry, 5 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 [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
13	M	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	M	102:UNK	C	300:PRO	N	51.89

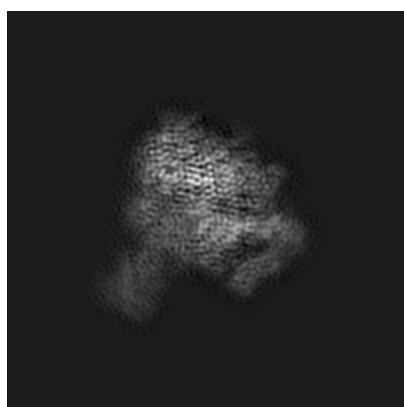
6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-7997. These allow visual inspection of the internal detail of the map and identification of artifacts.

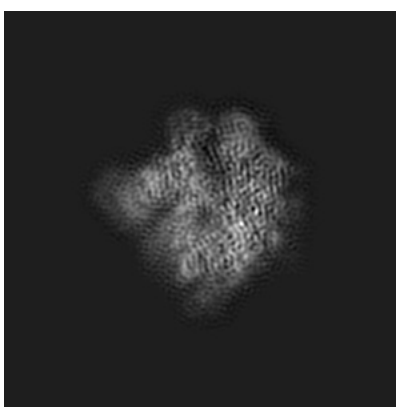
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections [i](#)

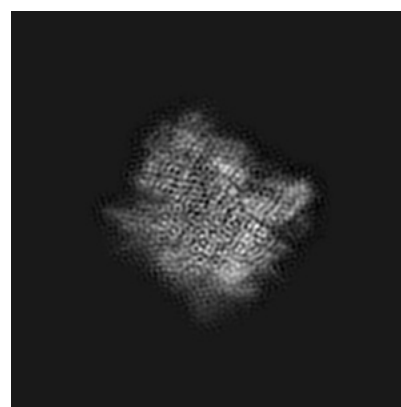
6.1.1 Primary map



X



Y

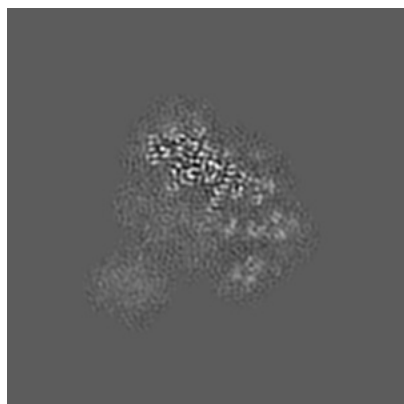


Z

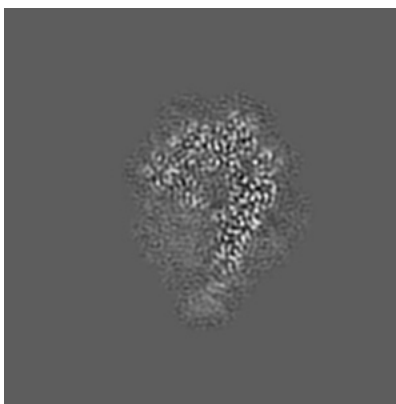
The images above show the map projected in three orthogonal directions.

6.2 Central slices [i](#)

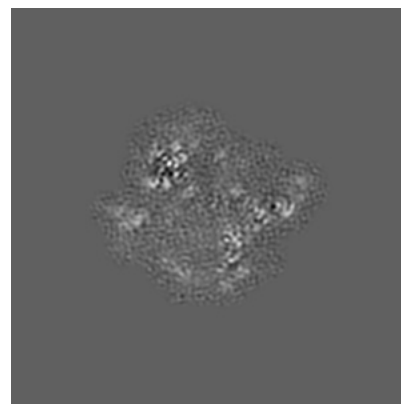
6.2.1 Primary map



X Index: 112



Y Index: 112

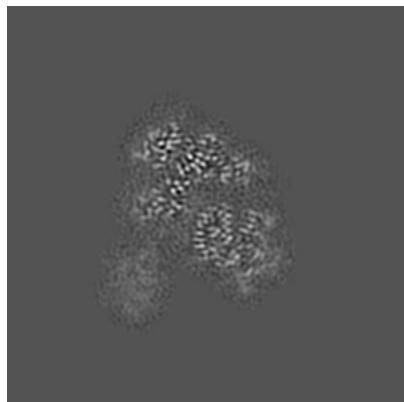


Z Index: 112

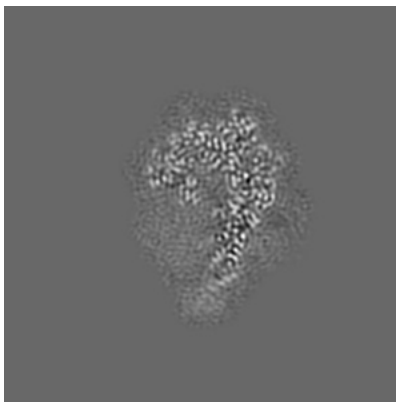
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices [i](#)

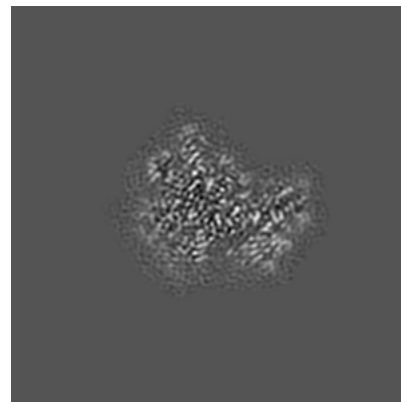
6.3.1 Primary map



X Index: 124



Y Index: 111



Z Index: 131

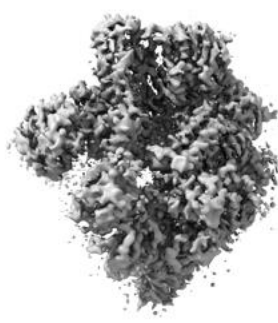
The images above show the largest variance slices of the map in three orthogonal directions.

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

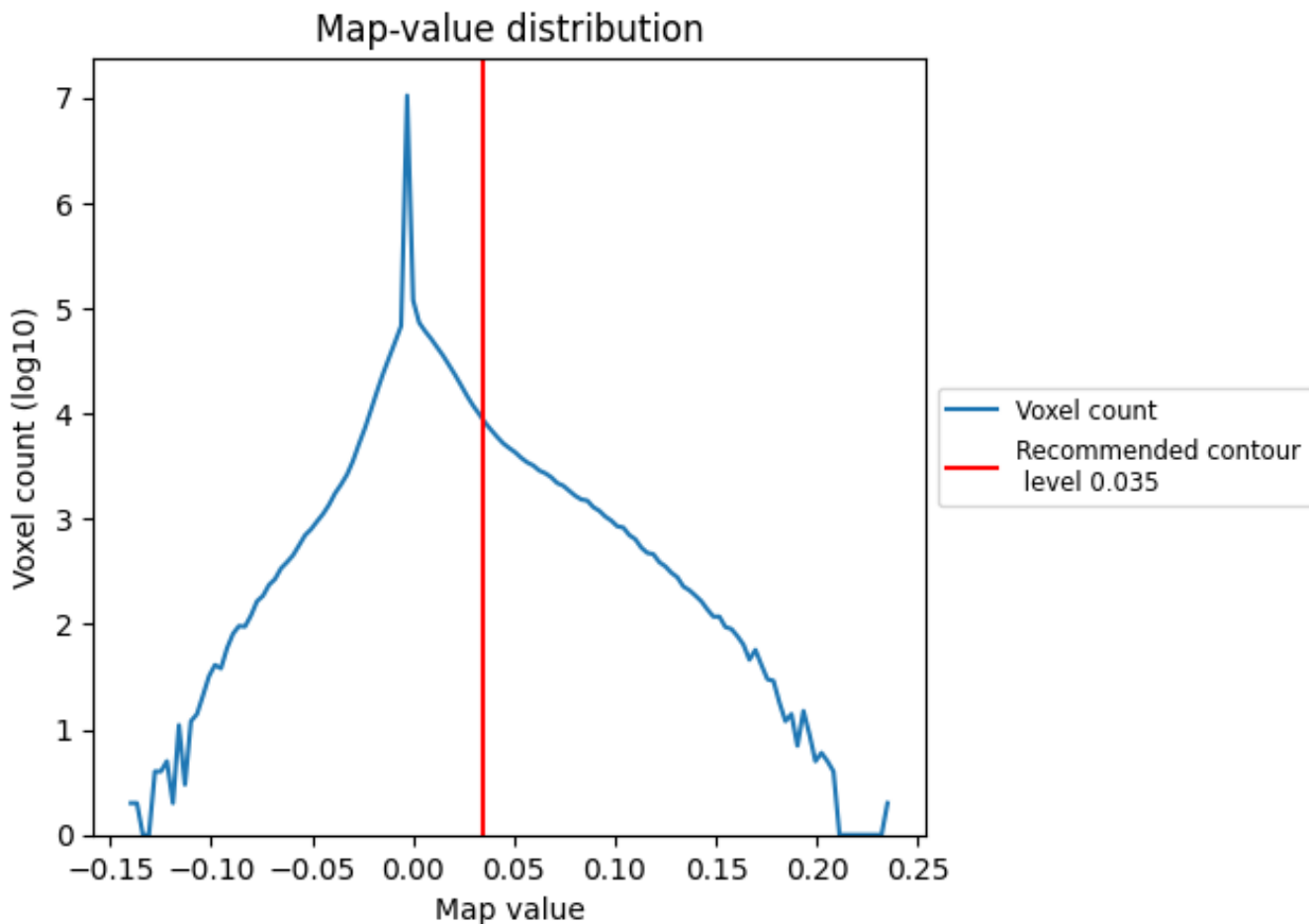
6.5 Mask visualisation

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

7 Map analysis [i](#)

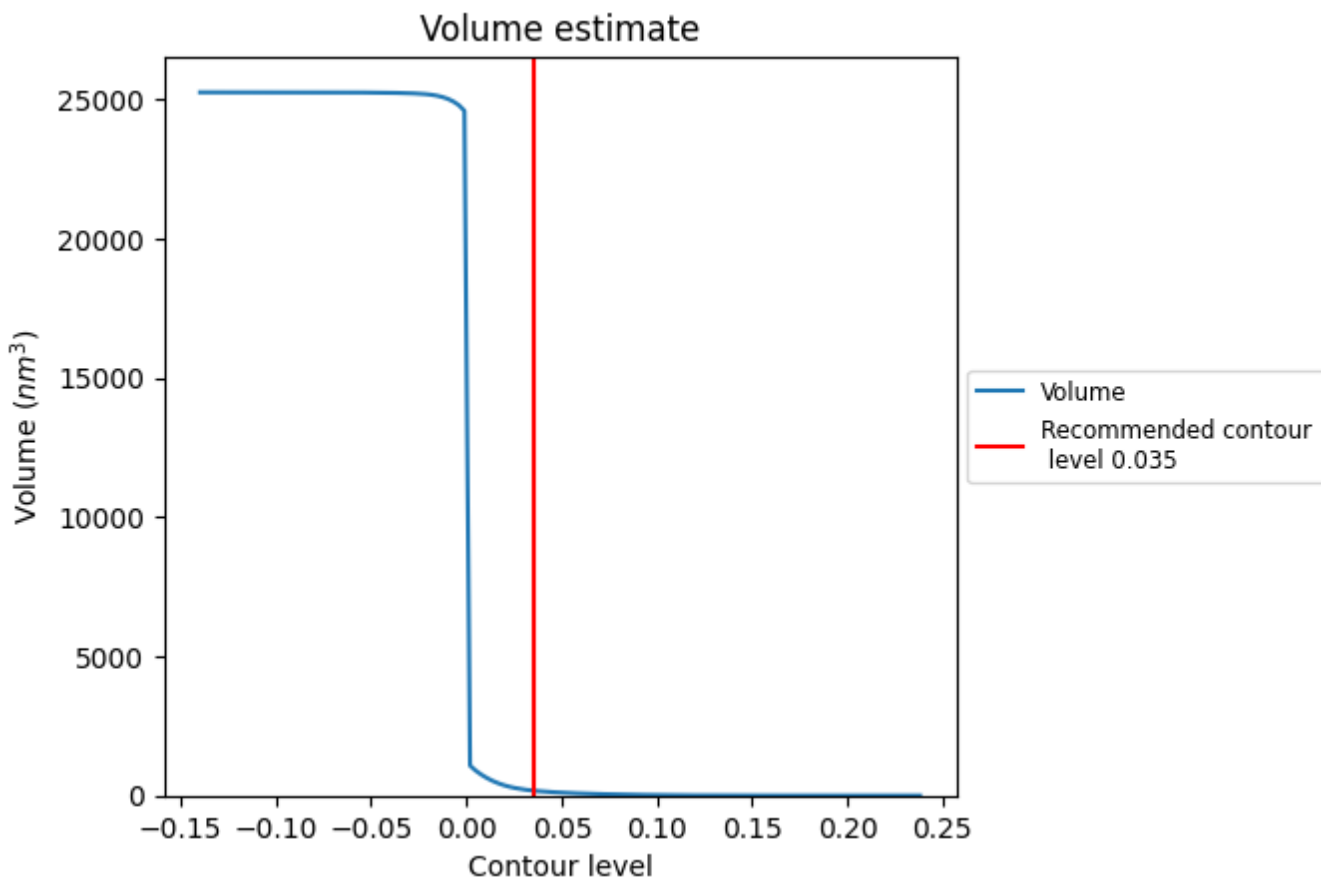
This section contains the results of statistical analysis of the map.

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.

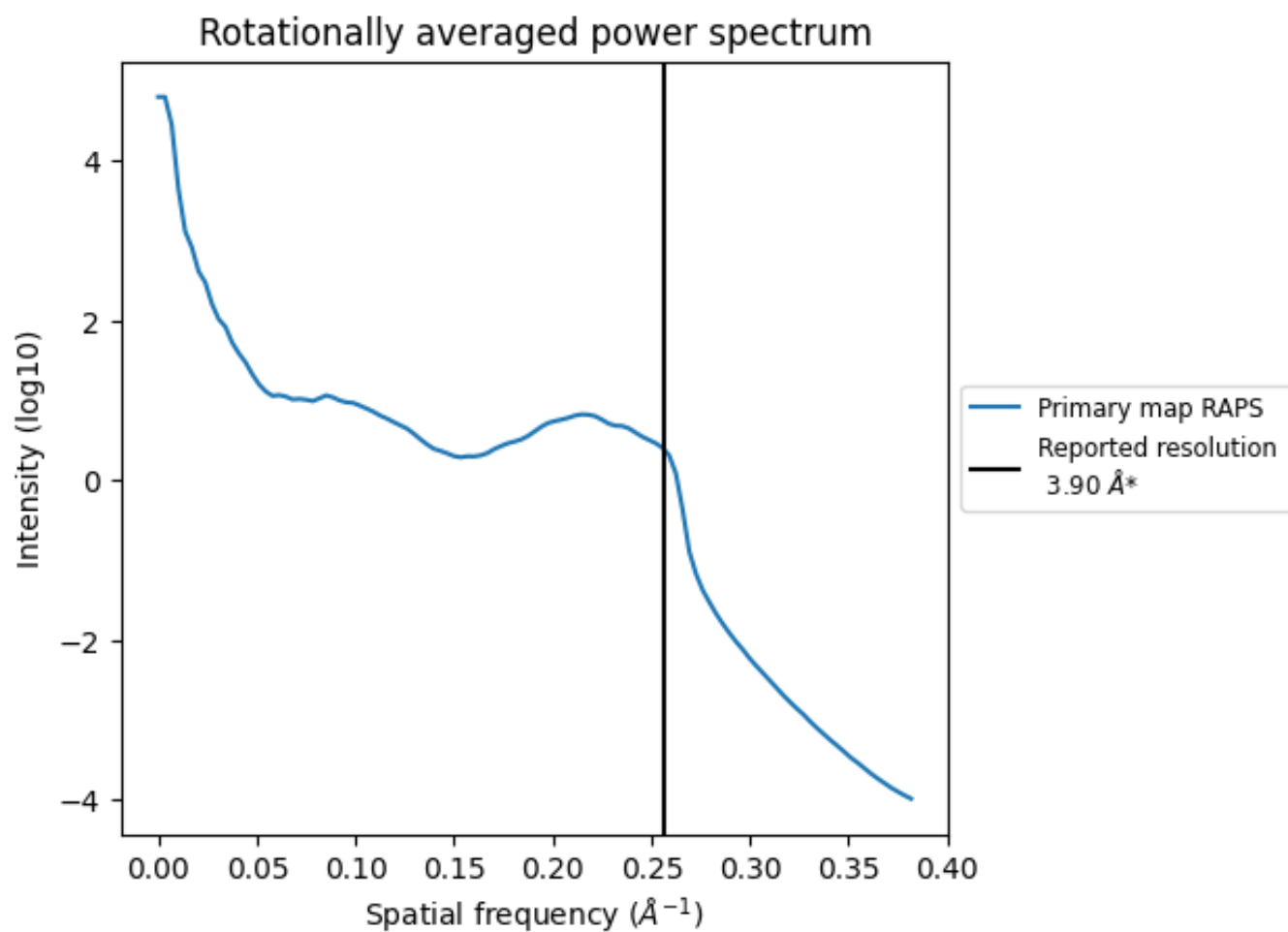
7.2 Volume estimate [i](#)



The volume at the recommended contour level is 178 nm³; this corresponds to an approximate mass of 161 kDa.

The volume estimate graph shows how the enclosed volume varies with the contour level. The recommended contour level is shown as a vertical line and the intersection between the line and the curve gives the volume of the enclosed surface at the given level.

7.3 Rotationally averaged power spectrum [i](#)

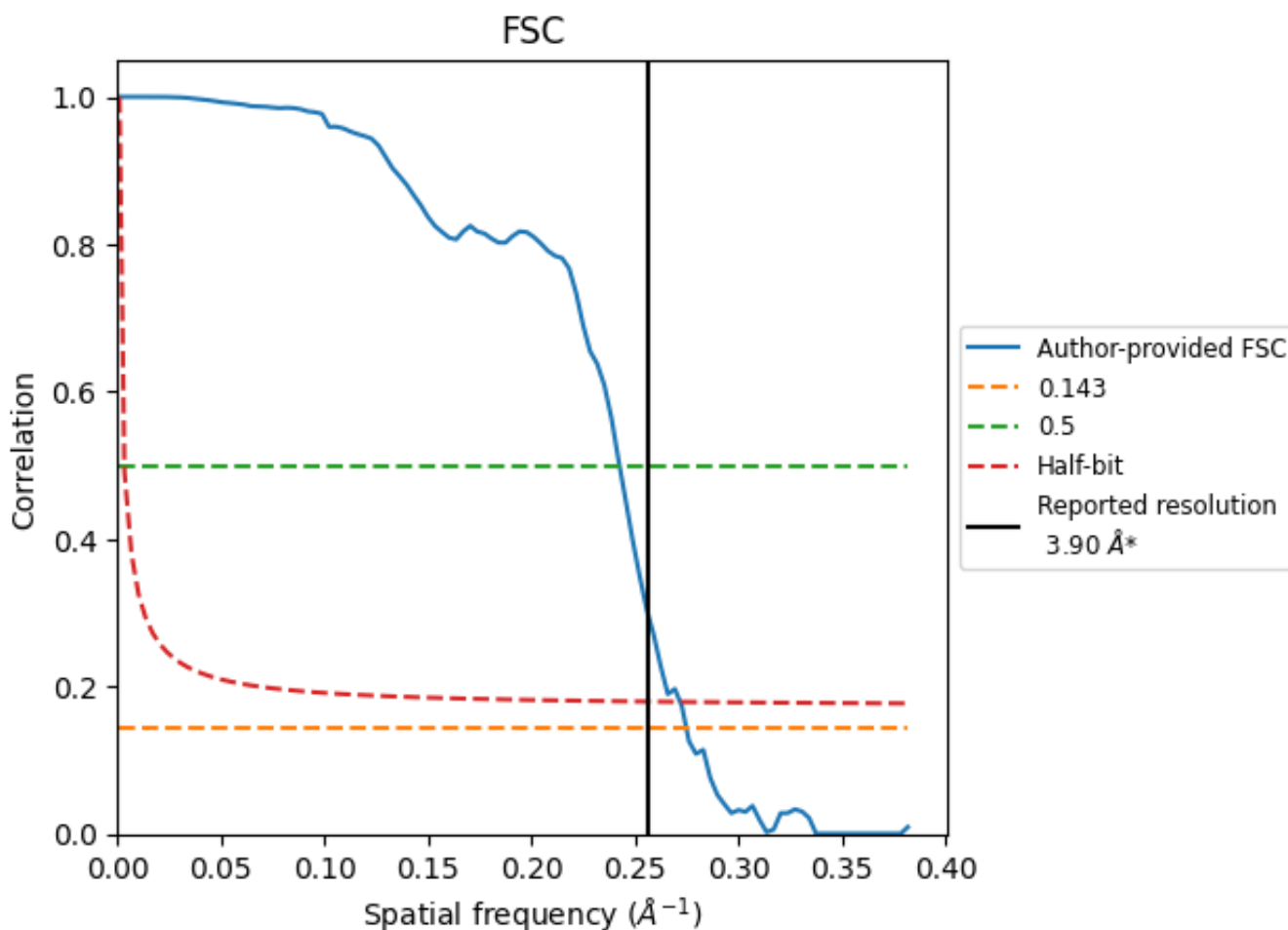


*Reported resolution corresponds to spatial frequency of 0.256 Å⁻¹

8 Fourier-Shell correlation [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 0.256 Å⁻¹

8.2 Resolution estimates [i](#)

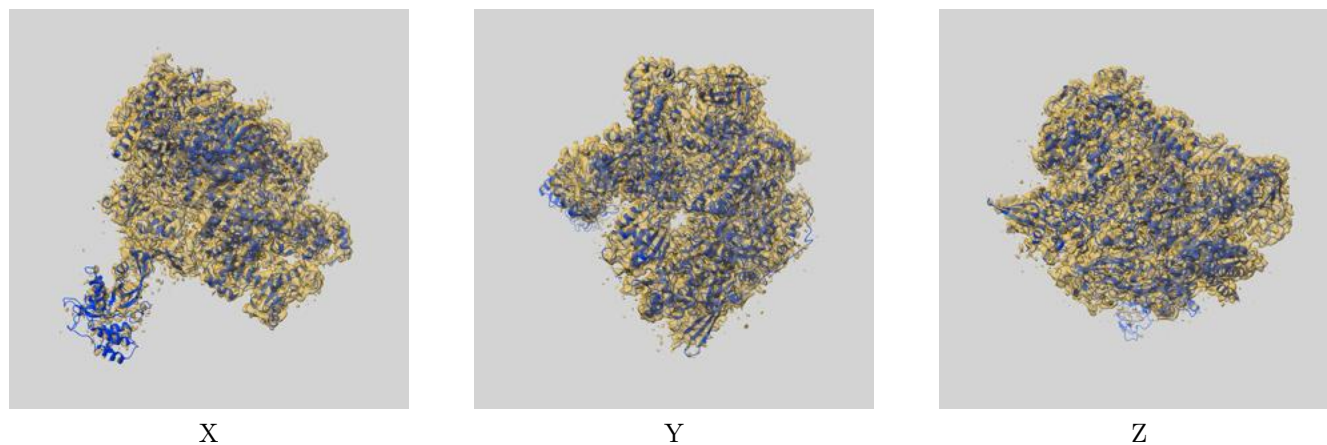
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	3.90	-	-
Author-provided FSC curve	3.64	4.13	3.68
Unmasked-calculated*	-	-	-

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

9 Map-model fit [i](#)

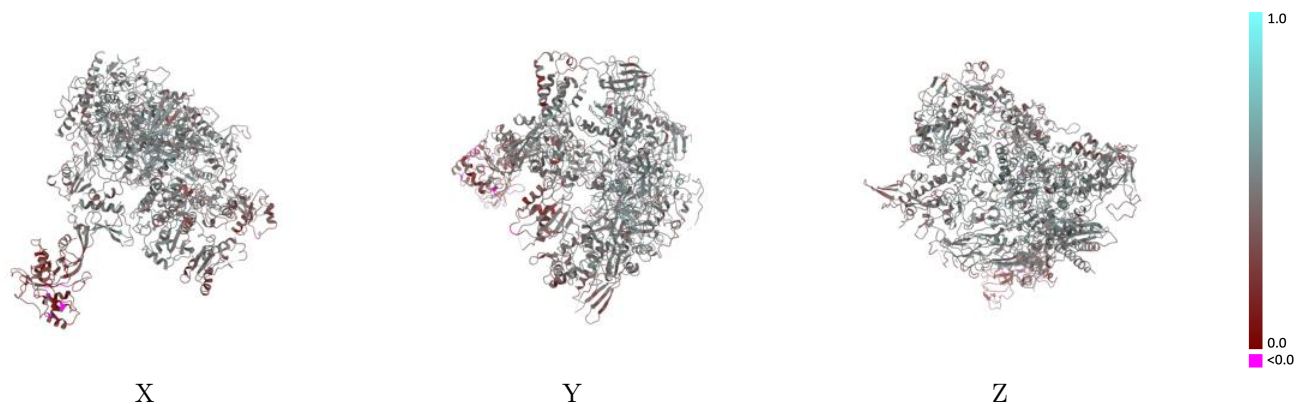
This section contains information regarding the fit between EMDB map EMD-7997 and PDB model 6DRD. Per-residue inclusion information can be found in section [3](#) on page [7](#).

9.1 Map-model overlay [i](#)



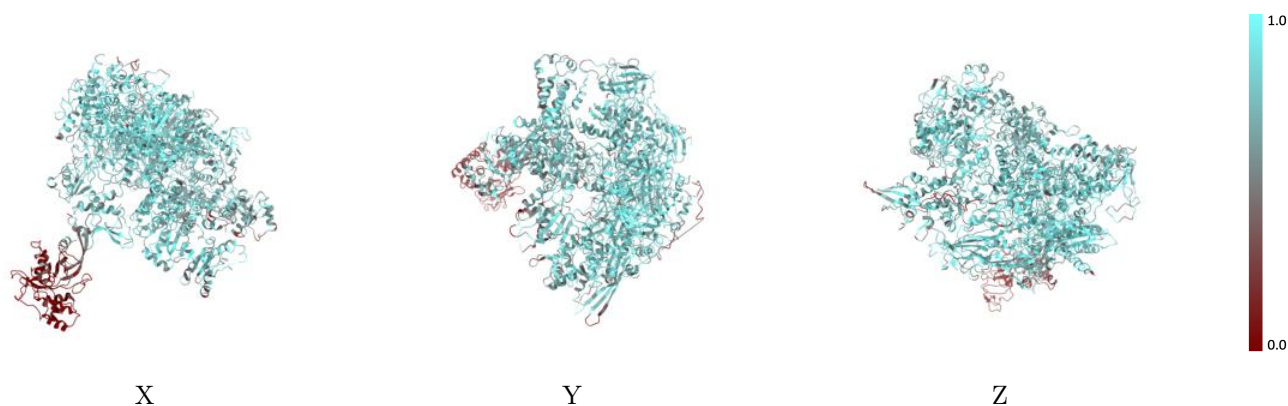
The images above show the 3D surface view of the map at the recommended contour level 0.035 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

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



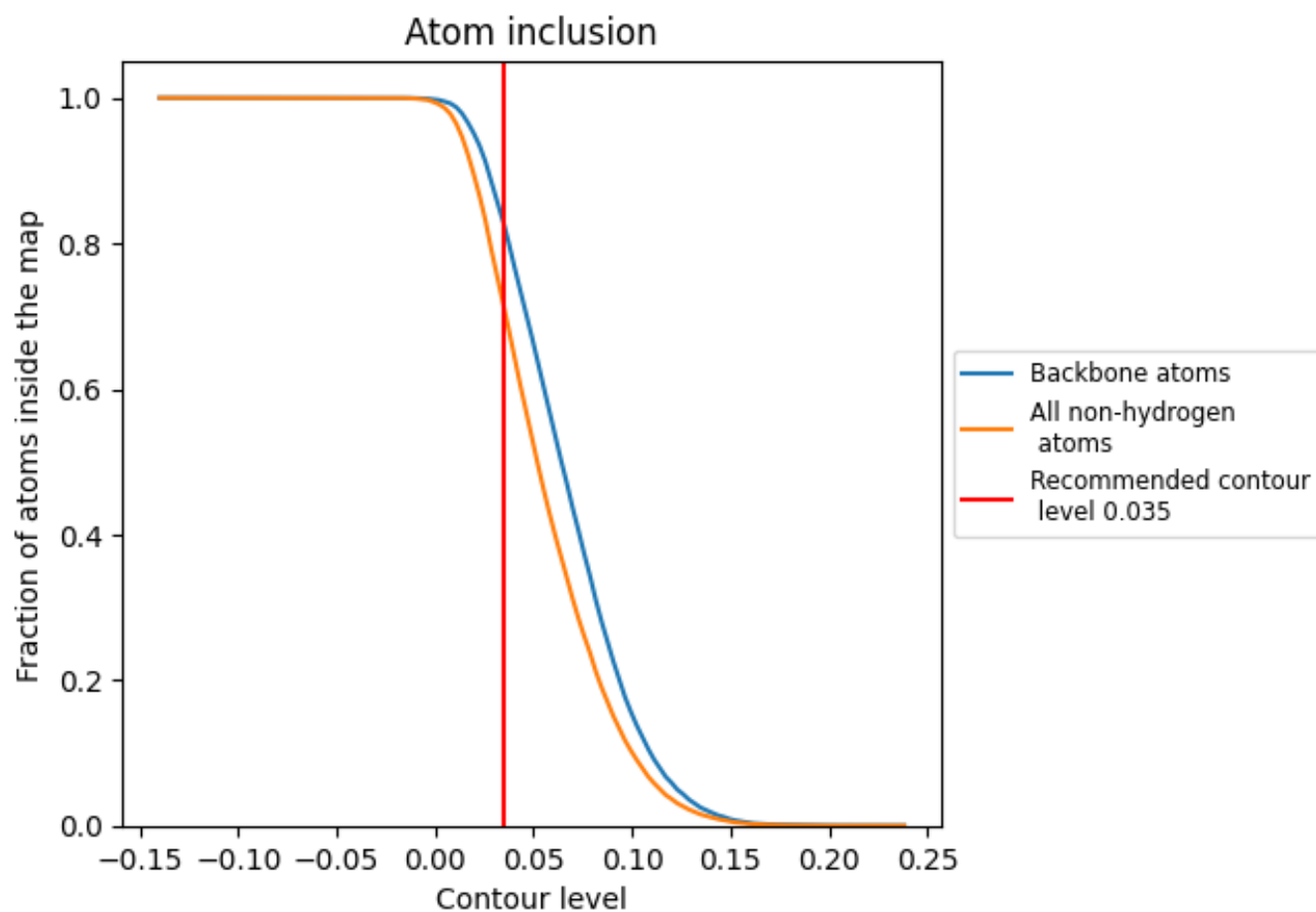
The images above show the model with each residue coloured according to its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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 (0.035).





























9.4 Atom inclusion [i](#)



At the recommended contour level, 83% of all backbone atoms, 71% 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 (0.035) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	 0.7131	 0.4460
A	 0.7542	 0.4590
B	 0.7756	 0.4730
C	 0.8214	 0.4850
D	 0.0960	 0.2450
E	 0.7393	 0.4250
F	 0.7604	 0.4620
G	 0.2120	 0.2750
H	 0.7511	 0.4530
I	 0.7489	 0.4410
J	 0.8626	 0.5000
K	 0.7925	 0.4710
L	 0.8000	 0.4300
M	 0.4624	 0.4140

