



wwPDB EM Validation Summary Report ⓘ

Nov 14, 2022 – 11:40 PM EST

PDB ID : 7KIN
EMDB ID : EMD-22888
Title : Mycobacterium tuberculosis WT RNAP transcription open promoter complex
with WhiB7 promoter
Authors : Lilic, M.; Darst, S.A.; Campbell, E.A.
Deposited on : 2020-10-23
Resolution : 2.74 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

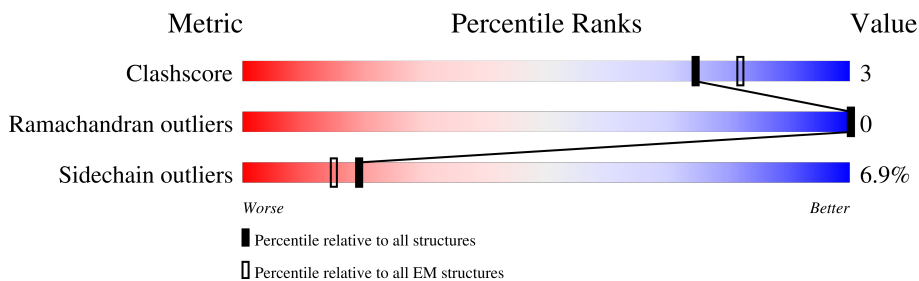
1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 2.74 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.



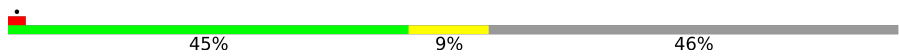

Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826

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	347	56% (green), 9% (yellow), 35% (grey)
1	B	347	59% (green), 8% (yellow), 32% (grey), 6% (red)
2	C	1172	81% (green), 13% (yellow), 5% (grey)
3	D	1318	85% (green), 11% (yellow), 4% (grey), 6% (red)
4	E	110	64% (green), 12% (yellow), 25% (grey), 6% (red)
5	F	528	53% (green), 6% (yellow), 40% (grey), 6% (red)
6	J	111	81% (green), 14% (yellow), 5% (grey), 6% (red)
7	M	162	84% (green), 12% (yellow), 4% (grey), 6% (red)

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
8	O	100	
9	P	100	

2 Entry composition [i](#)

There are 11 unique types of molecules in this entry. The entry contains 29379 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 subunit alpha.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	225	Total	C	N	O	S	0	0
			1716	1080	296	338	2		
1	B	237	Total	C	N	O	S	0	0
			1765	1115	301	346	3		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	C	1111	Total	C	N	O	S	0	0
			8606	5392	1511	1664	39		

- Molecule 3 is a protein called DNA-directed RNA polymerase subunit beta'.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	D	1270	Total	C	N	O	S	0	0
			9914	6208	1802	1862	42		

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	-1	GLY	-	expression tag	UNP A0A045J9E2
D	0	ALA	-	expression tag	UNP A0A045J9E2

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

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
4	E	83	Total	C	N	O	0	0
			649	414	108	127		

- Molecule 5 is a protein called RNA polymerase sigma factor SigA.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	F	317	Total	C	N	O	S	0	0
			2510	1566	455	480	9		

- Molecule 6 is a protein called RNA polymerase-binding protein RbpA.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	J	107	Total	C	N	O	S	0	0
			870	537	164	166	3		

- Molecule 7 is a protein called RNA polymerase-binding transcription factor CarD.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	M	159	Total	C	N	O	S	0	0
			1241	777	224	239	1		

- Molecule 8 is a DNA chain called DNA (54-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
8	O	54	Total	C	N	O	P	0	0
			1096	522	189	331	54		

- Molecule 9 is a DNA chain called DNA (49-MER).

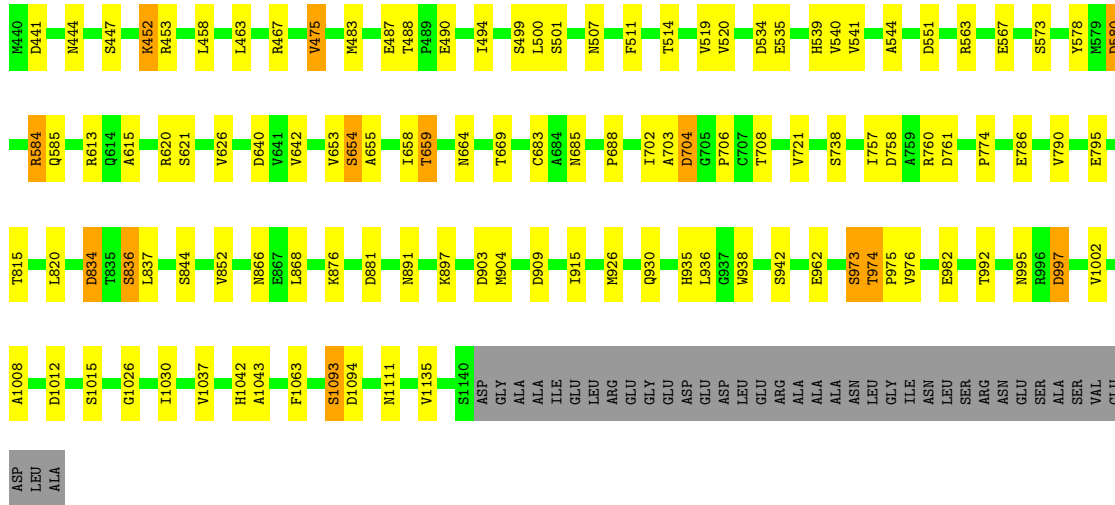
Mol	Chain	Residues	Atoms					AltConf	Trace
9	P	49	Total	C	N	O	P	0	0
			1009	477	195	288	49		

- Molecule 10 is ZINC ION (three-letter code: ZN) (formula: Zn).

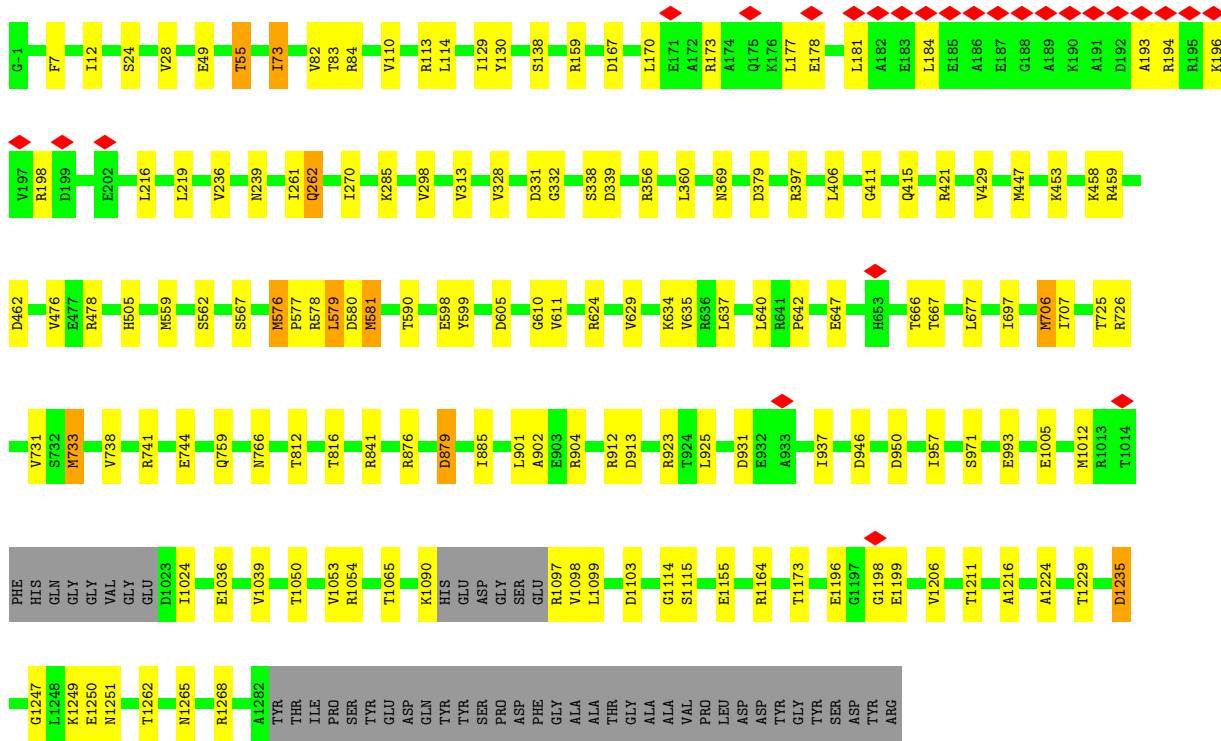
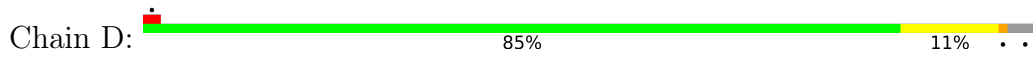
Mol	Chain	Residues	Atoms		AltConf
10	D	2	Total	Zn	0
			2	2	

- Molecule 11 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

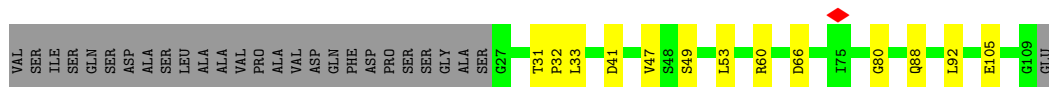
Mol	Chain	Residues	Atoms		AltConf
11	D	1	Total	Mg	0
			1	1	



• Molecule 3: DNA-directed RNA polymerase subunit beta'



• Molecule 4: DNA-directed RNA polymerase subunit omega



• Molecule 5: RNA polymerase sigma factor SigA

4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	296097	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 ($e^-/\text{\AA}^2$)	64	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	4.851	Depositor
Minimum map value	-3.227	Depositor
Average map value	0.004	Depositor
Map value standard deviation	0.119	Depositor
Recommended contour level	0.3	Depositor
Map size (Å)	330.0, 330.0, 330.0	wwPDB
Map dimensions	300, 300, 300	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.1, 1.1, 1.1	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: MG, 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	0.73	0/1742	0.64	1/2370 (0.0%)
1	B	0.60	0/1792	0.61	0/2442
2	C	0.77	0/8765	0.64	2/11885 (0.0%)
3	D	0.74	0/10078	0.63	3/13624 (0.0%)
4	E	0.71	0/662	0.66	0/901
5	F	0.49	0/2541	0.56	1/3427 (0.0%)
6	J	0.43	0/886	0.53	0/1196
7	M	0.37	0/1257	0.55	0/1700
8	O	0.93	1/1223 (0.1%)	1.09	0/1881
9	P	0.86	0/1133	0.99	0/1744
All	All	0.72	1/30079 (0.0%)	0.67	7/41170 (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
2	C	0	3
3	D	0	1
All	All	0	4

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
8	O	48	DT	N1-C6	-5.85	1.34	1.38

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	475	VAL	C-N-CA	6.87	138.88	121.70
3	D	579	LEU	CA-CB-CG	5.93	128.93	115.30
2	C	308	LEU	CA-CB-CG	5.69	128.38	115.30
1	A	128	LEU	CA-CB-CG	5.52	127.99	115.30
3	D	578	ARG	C-N-CA	5.42	135.25	121.70

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	C	1093	SER	Peptide
2	C	774	PRO	Peptide
2	C	962	GLU	Peptide
3	D	579	LEU	Mainchain

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	1716	0	1756	13	0
1	B	1765	0	1794	11	0
2	C	8606	0	8544	67	0
3	D	9914	0	9986	62	0
4	E	649	0	645	5	0
5	F	2510	0	2537	23	0
6	J	870	0	848	9	0
7	M	1241	0	1259	10	0
8	O	1096	0	611	6	0
9	P	1009	0	549	5	0
10	D	2	0	0	0	0
11	D	1	0	0	0	0
All	All	29379	0	28529	186	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 3.

The worst 5 of 186 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:B:183:VAL:HA	1:B:188:ASP:H	1.51	0.73
1:A:40:ARG:HH12	2:C:903:ASP:HB3	1.60	0.66
5:F:306:LEU:HD11	5:F:348:THR:HG23	1.77	0.65
3:D:1265:ASN:OD1	3:D:1268:ARG:NH2	2.31	0.64
3:D:181:LEU:HA	3:D:198:ARG:HH21	1.64	0.63

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	223/347 (64%)	210 (94%)	13 (6%)	0	100	100
1	B	235/347 (68%)	220 (94%)	15 (6%)	0	100	100
2	C	1109/1172 (95%)	1036 (93%)	73 (7%)	0	100	100
3	D	1264/1318 (96%)	1198 (95%)	66 (5%)	0	100	100
4	E	81/110 (74%)	76 (94%)	5 (6%)	0	100	100
5	F	315/528 (60%)	305 (97%)	10 (3%)	0	100	100
6	J	105/111 (95%)	92 (88%)	13 (12%)	0	100	100
7	M	157/162 (97%)	145 (92%)	12 (8%)	0	100	100
All	All	3489/4095 (85%)	3282 (94%)	207 (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	A	194/297 (65%)	181 (93%)	13 (7%)	16	29
1	B	195/297 (66%)	177 (91%)	18 (9%)	9	17
2	C	937/993 (94%)	860 (92%)	77 (8%)	11	21
3	D	1048/1095 (96%)	989 (94%)	59 (6%)	21	36
4	E	69/90 (77%)	63 (91%)	6 (9%)	10	19
5	F	264/427 (62%)	254 (96%)	10 (4%)	33	54
6	J	92/97 (95%)	85 (92%)	7 (8%)	13	23
7	M	129/131 (98%)	118 (92%)	11 (8%)	10	20
All	All	2928/3427 (85%)	2727 (93%)	201 (7%)	19	27

5 of 201 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	D	219	LEU
3	D	759	GLN
7	M	133	GLN
3	D	298	VAL
3	D	581	MET

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

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 3 ligands modelled in this entry, 3 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](#)

There are no chain breaks in this entry.

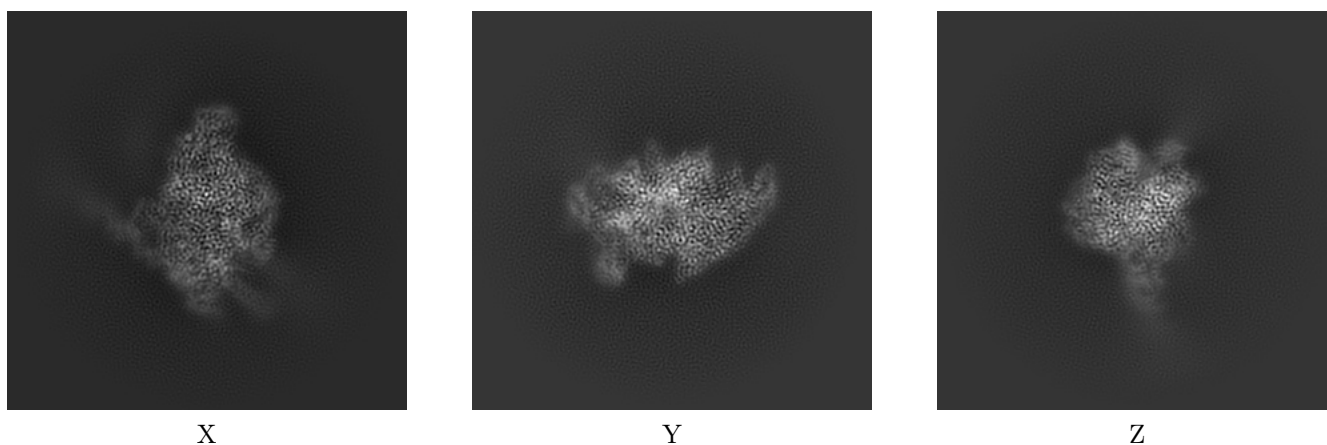
6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-22888. 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](#)

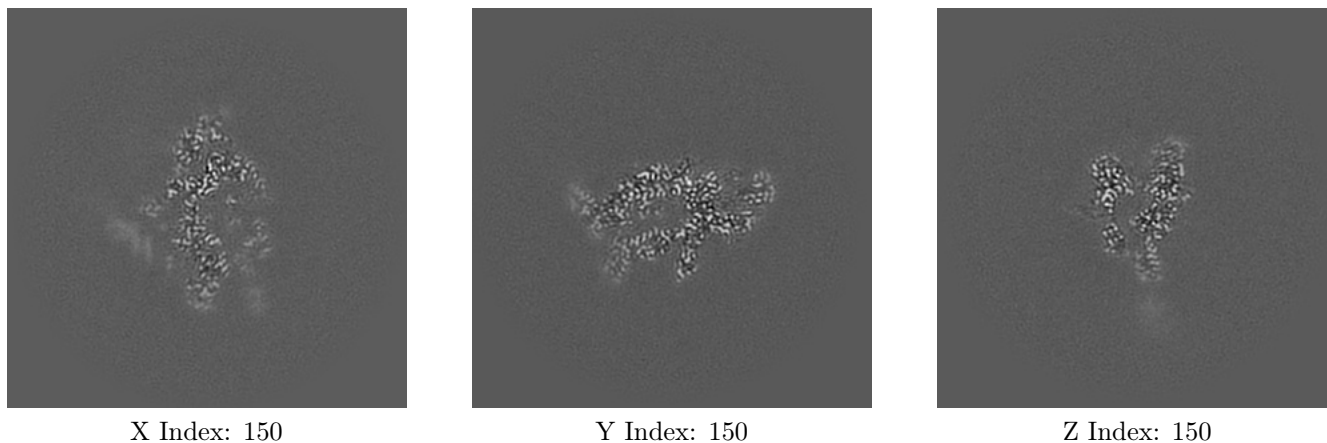
6.1.1 Primary map



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

6.2 Central slices [i](#)

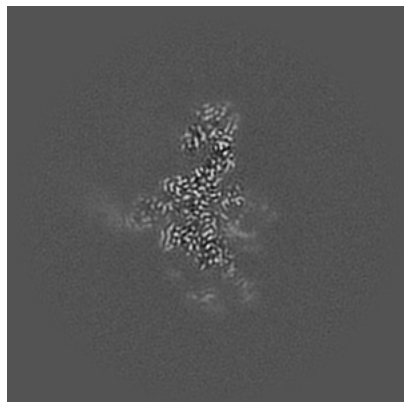
6.2.1 Primary map



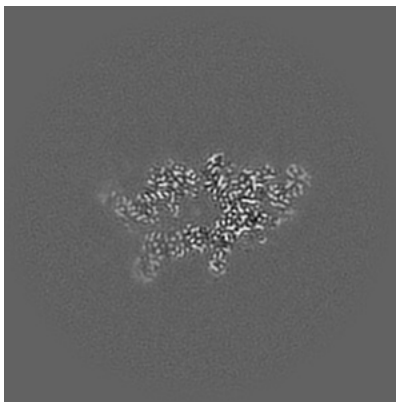
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices [i](#)

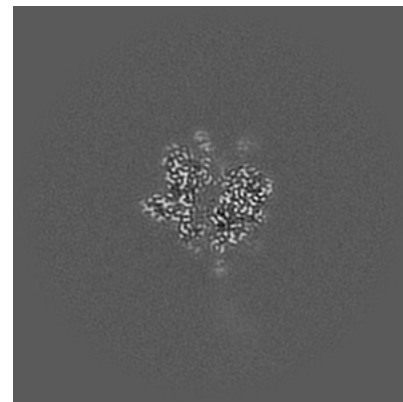
6.3.1 Primary map



X Index: 160



Y Index: 154



Z Index: 158

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

6.4 Orthogonal surface views [i](#)

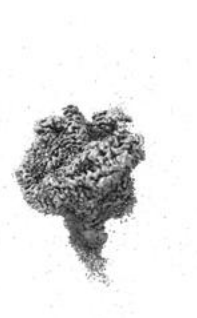
6.4.1 Primary map



X



Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.3. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

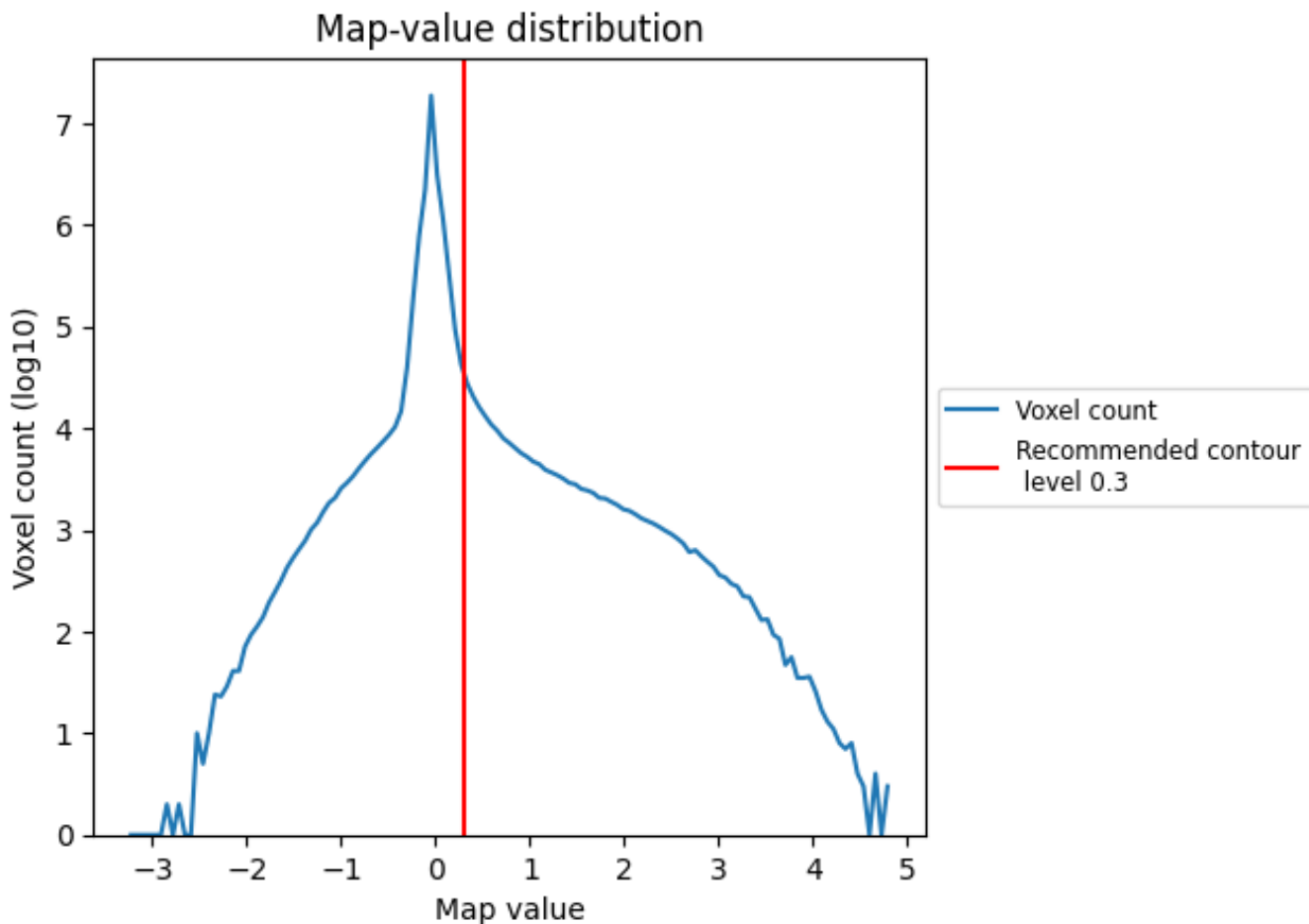
6.5 Mask visualisation

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

7 Map analysis [i](#)

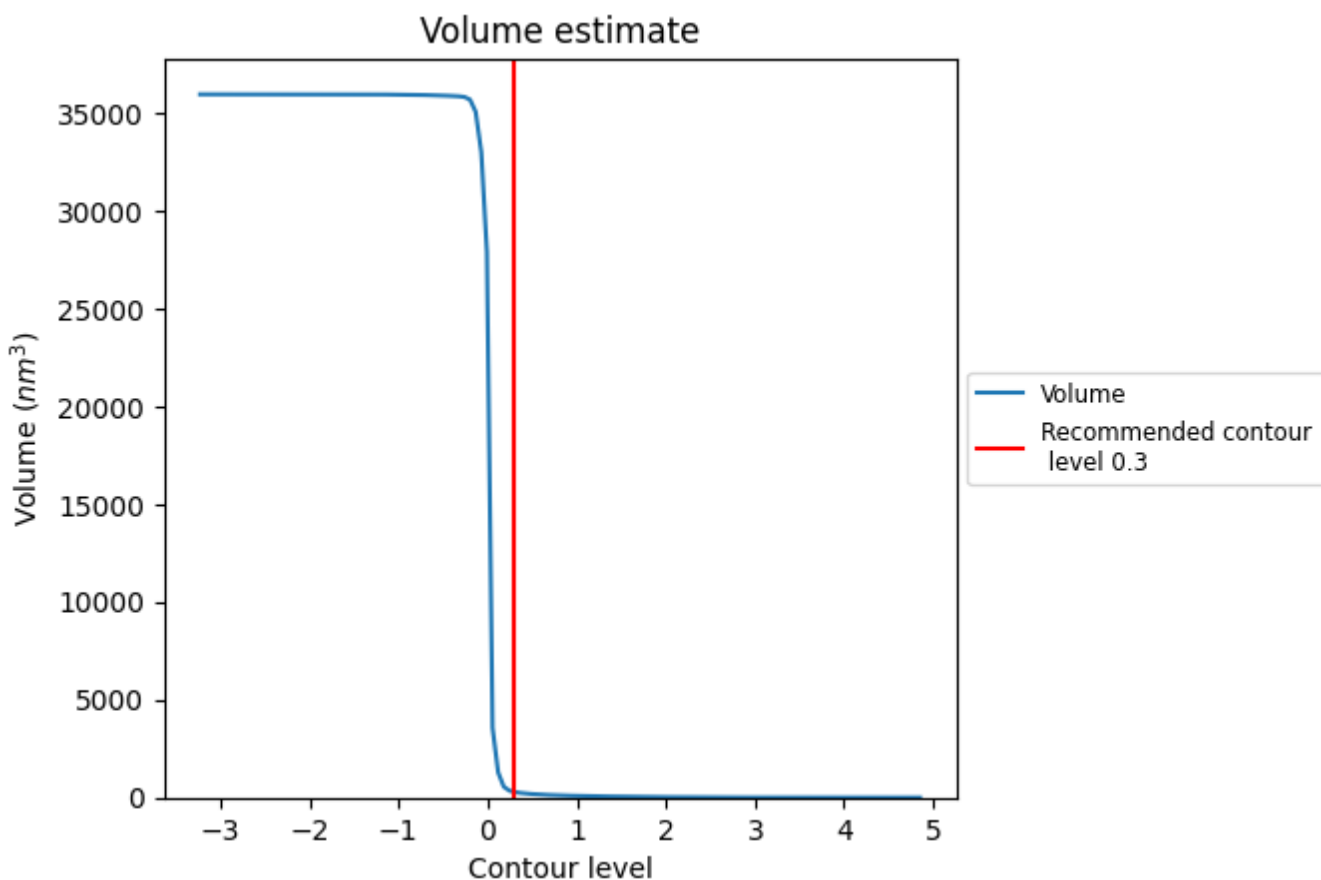
This section contains the results of statistical analysis of the map.

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.

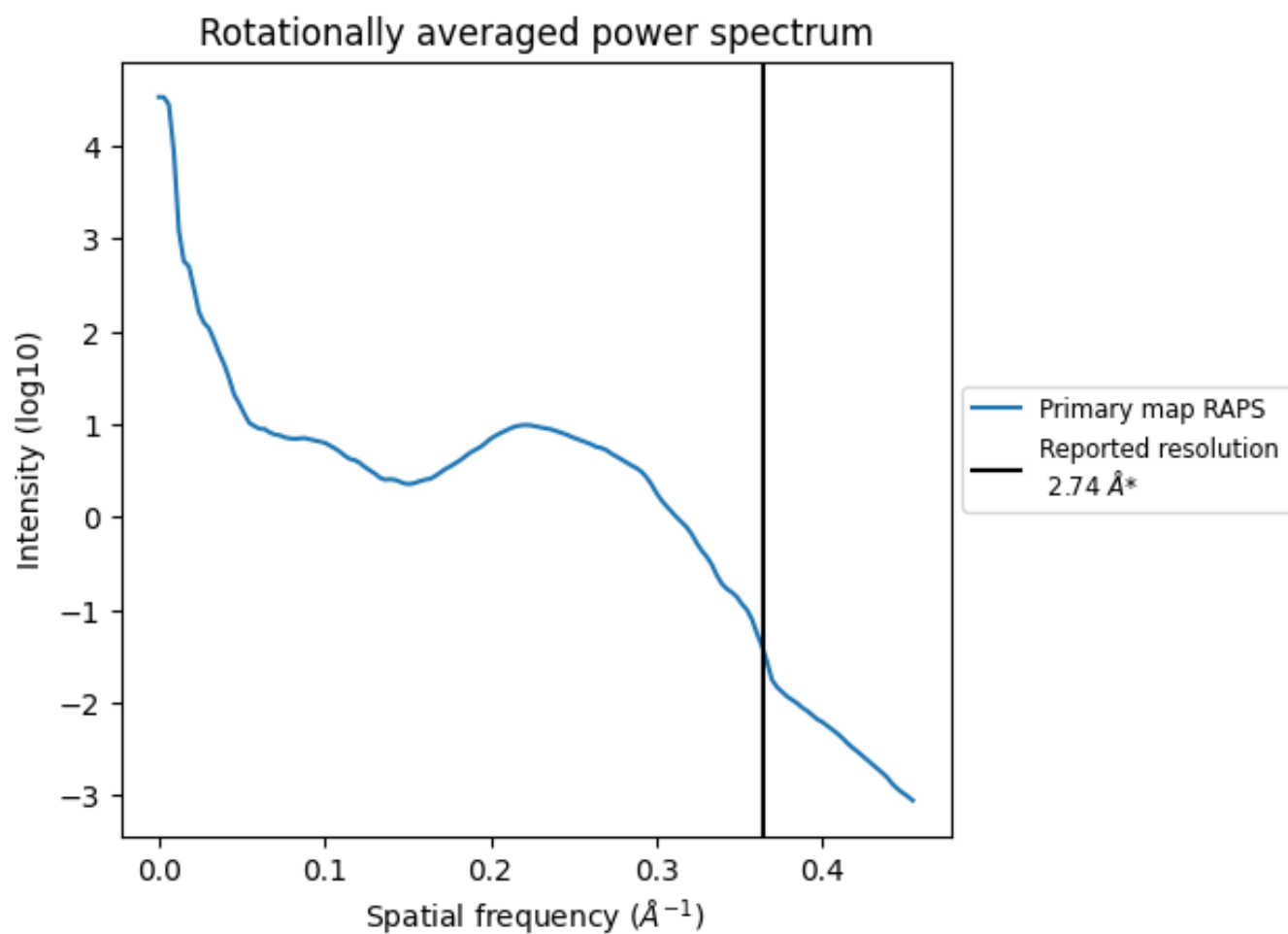
7.2 Volume estimate [i](#)



The volume at the recommended contour level is 286 nm³; this corresponds to an approximate mass of 258 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 0.365 Å⁻¹

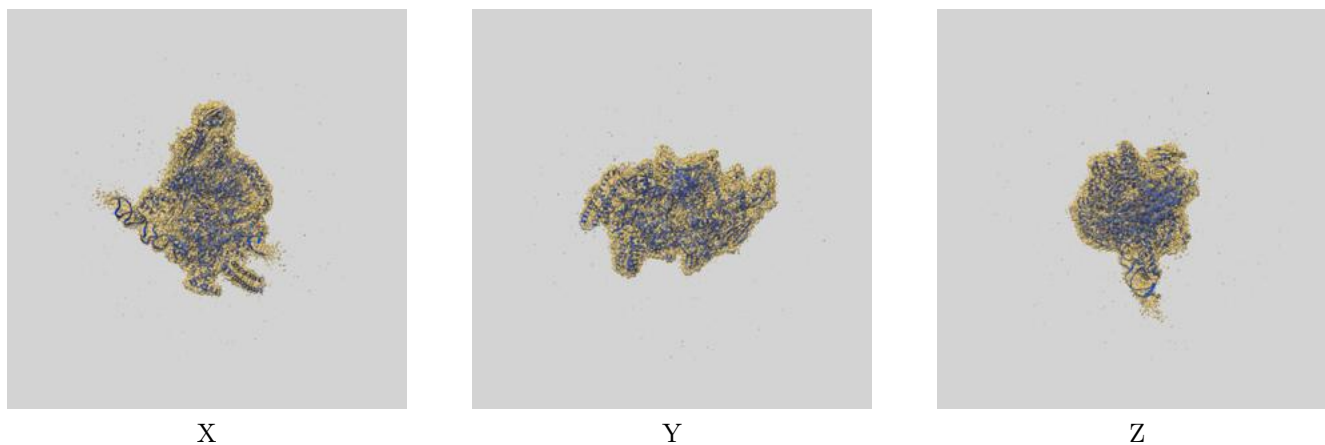
8 Fourier-Shell correlation

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

9 Map-model fit [i](#)

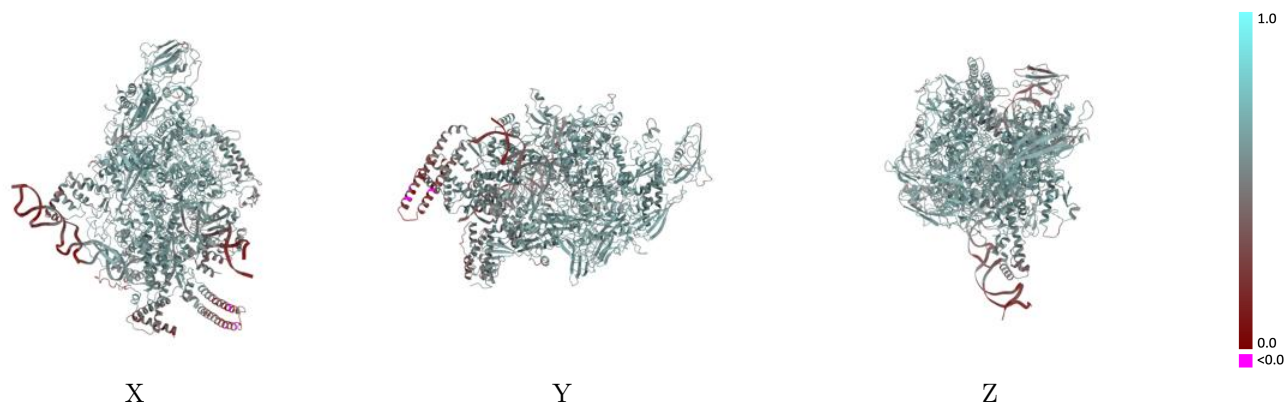
This section contains information regarding the fit between EMDB map EMD-22888 and PDB model 7KIN. Per-residue inclusion information can be found in section [3](#) on page [6](#).

9.1 Map-model overlay [i](#)



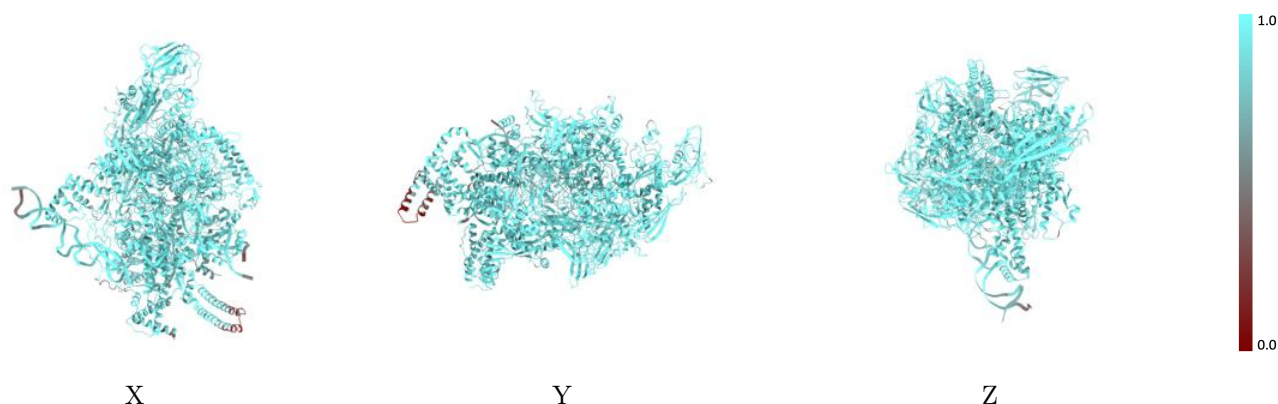
The images above show the 3D surface view of the map at the recommended contour level 0.3 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

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



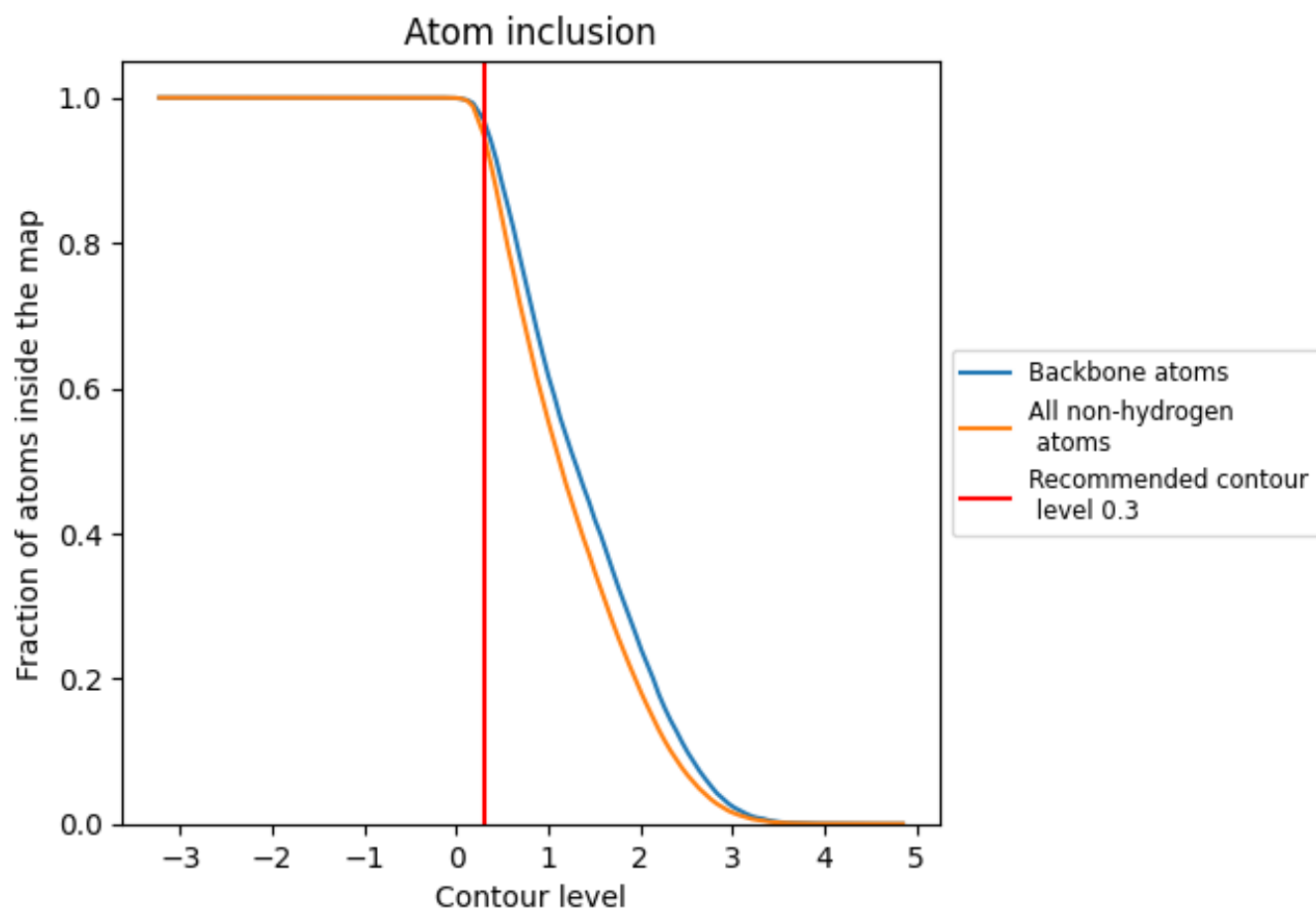
The images above show the model with each residue coloured according to its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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 (0.3).























9.4 Atom inclusion [i](#)



At the recommended contour level, 97% of all backbone atoms, 95% 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 (0.3) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	 0.9477	 0.5560
A	 0.9768	 0.6000
B	 0.9516	 0.5650
C	 0.9692	 0.5930
D	 0.9513	 0.5740
E	 0.9512	 0.5870
F	 0.9450	 0.5330
J	 0.8840	 0.5010
M	 0.9026	 0.4750
O	 0.8650	 0.3730
P	 0.8781	 0.3620

