



wwPDB EM Validation Summary Report ⓘ

Aug 24, 2026 – 08:05 AM EDT

PDB ID : 9Y7B / pdb_00009y7b
EMDB ID : EMD-72650
Title : Composite map of IF-3a transposase complex with paired transposon ends
Authors : Miller, D.J.; Truong, V.H.; Fatma, S.; Kellogg, E.H.
Deposited on : 2025-09-09
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:

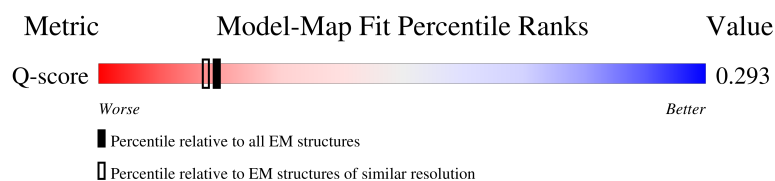
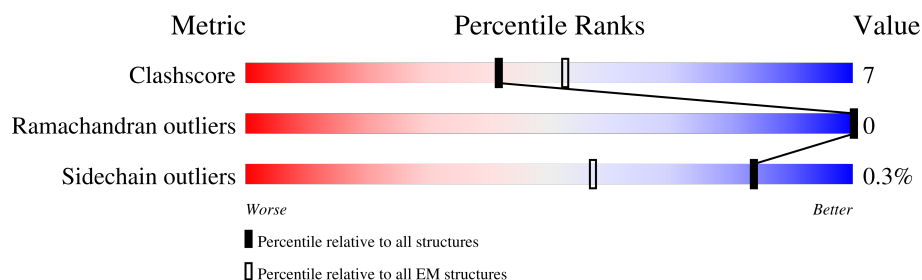
EMDB validation analysis : 0.0.1.dev133
MolProbity : 4-5-2 with Phenix2.0
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)
EM percentile statistics : 202505.v01 (Using data in the EMDB archive up until May 2025)
MapQ : 1.9.13
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.50

1 Overall quality at a glance

The following experimental techniques were used to determine the structure:
ELECTRON MICROSCOPY

The reported resolution of this entry is 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.



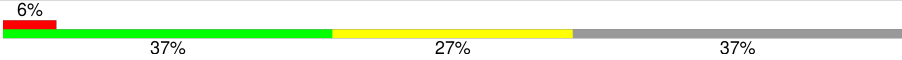
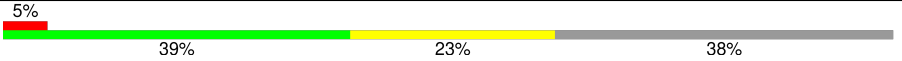
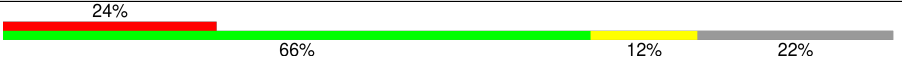
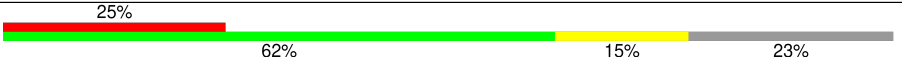
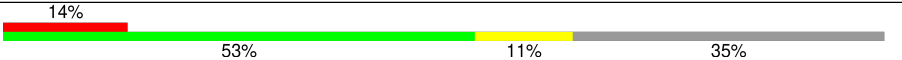
Metric	Whole archive (#Entries)	EM structures (#Entries)	Similar EM resolution (#Entries, resolution range(Å))
Clashscore	229148	23984	-
Ramachandran outliers	224038	23583	-
Sidechain outliers	223484	23102	-
Q-score	-	25397	8855 (3.40 - 4.40)

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	G	121	
2	H	94	
3	K	91	
4	N	192	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
5	E	128	
6	F	133	
7	A	603	
7	B	603	
7	C	603	
7	D	603	

2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 21225 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 Integration host factor subunit alpha.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	G	93	Total	C	N	O	S	0	0
			751	470	140	140	1		

There are 22 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
G	-21	MET	-	initiating methionine	UNP Q14F23
G	-20	GLY	-	expression tag	UNP Q14F23
G	-19	SER	-	expression tag	UNP Q14F23
G	-18	SER	-	expression tag	UNP Q14F23
G	-17	HIS	-	expression tag	UNP Q14F23
G	-16	HIS	-	expression tag	UNP Q14F23
G	-15	HIS	-	expression tag	UNP Q14F23
G	-14	HIS	-	expression tag	UNP Q14F23
G	-13	HIS	-	expression tag	UNP Q14F23
G	-12	HIS	-	expression tag	UNP Q14F23
G	-11	SER	-	expression tag	UNP Q14F23
G	-10	GLN	-	expression tag	UNP Q14F23
G	-9	ASP	-	expression tag	UNP Q14F23
G	-8	PRO	-	expression tag	UNP Q14F23
G	-7	GLU	-	expression tag	UNP Q14F23
G	-6	ASN	-	expression tag	UNP Q14F23
G	-5	LEU	-	expression tag	UNP Q14F23
G	-4	TYR	-	expression tag	UNP Q14F23
G	-3	PHE	-	expression tag	UNP Q14F23
G	-2	GLN	-	expression tag	UNP Q14F23
G	-1	GLY	-	expression tag	UNP Q14F23
G	0	SER	-	expression tag	UNP Q14F23

- Molecule 2 is a protein called Integration host factor subunit beta.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	H	94	Total	C	N	O	S	0	0
			742	461	139	139	3		

- Molecule 3 is a DNA chain called IF3A Right Transposon End.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	K	38	Total	C	N	O	P	0	0
			777	373	146	220	38		

- Molecule 4 is a DNA chain called IF3A Right Transposon End Complement.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	N	40	Total	C	N	O	P	0	0
			823	397	140	246	40		

- Molecule 5 is a DNA chain called IF3A Left Transposon End.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	E	81	Total	C	N	O	P	0	0
			1654	796	299	478	81		

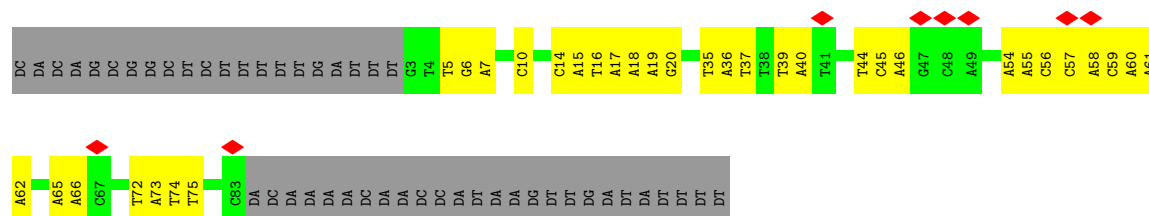
- Molecule 6 is a DNA chain called IF3A Left Transposon End Complement.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	F	82	Total	C	N	O	P	0	0
			1688	813	294	499	82		

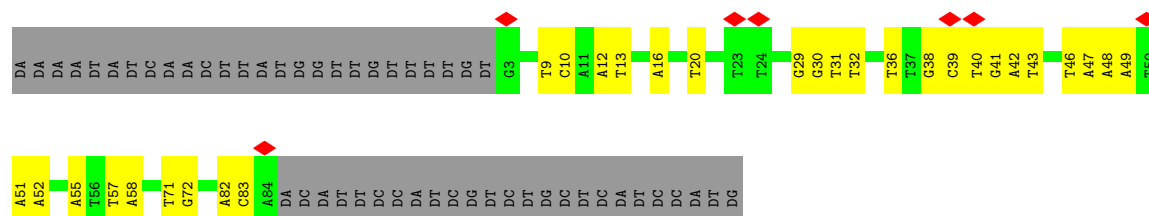
- Molecule 7 is a protein called IF3A Tn6677 TnsB.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	A	475	Total	C	N	O	S	0	0
			3909	2495	673	730	11		
7	B	469	Total	C	N	O	S	0	0
			3861	2470	669	710	12		
7	C	464	Total	C	N	O	S	0	0
			3824	2447	661	704	12		
7	D	389	Total	C	N	O	S	0	0
			3196	2033	545	608	10		

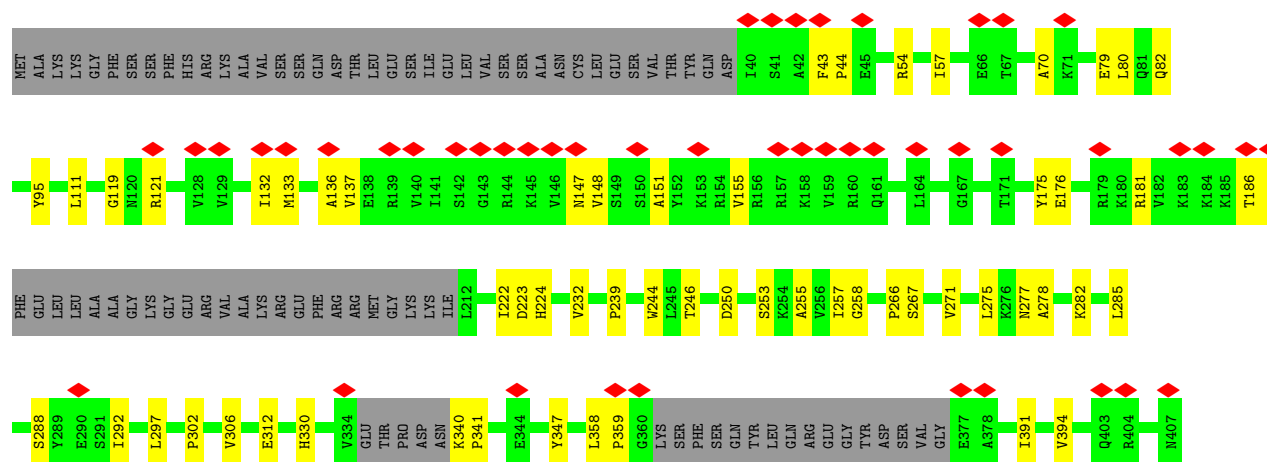
- Molecule 5: IF3A Left Transposon End

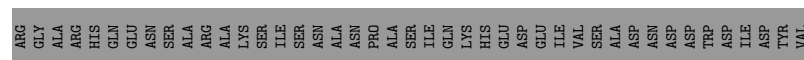


- Molecule 6: IF3A Left Transposon End Complement



- Molecule 7: IF3A Tn6677 TnsB





4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	168668	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TALOS ARCTICA	Depositor
Voltage (kV)	200	Depositor
Electron dose ($e^-/\text{\AA}^2$)	60	Depositor
Minimum defocus (nm)	1000	Depositor
Maximum defocus (nm)	2500	Depositor
Magnification	Not provided	
Image detector	GATAN K3 BIOQUANTUM (6k x 4k)	Depositor
Maximum map value	57.804	Depositor
Minimum map value	-3.147	Depositor
Average map value	4.377	Depositor
Map value standard deviation	1.030	Depositor
Recommended contour level	10	Depositor
Map size (Å)	322.56, 322.56, 322.56	wwPDB
Map dimensions	256, 256, 256	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.26, 1.26, 1.26	Depositor

5 Model quality [i](#)

5.1 Standard geometry [i](#)

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	G	0.11	0/761	0.32	0/1016
2	H	0.09	0/755	0.24	0/1010
3	K	0.18	0/873	0.39	0/1343
4	N	0.17	0/921	0.39	0/1421
5	E	0.17	0/1856	0.36	0/2859
6	F	0.19	0/1892	0.39	0/2921
7	A	0.08	0/3992	0.23	0/5398
7	B	0.08	0/3947	0.22	0/5334
7	C	0.08	0/3909	0.23	0/5283
7	D	0.07	0/3261	0.22	0/4405
All	All	0.11	0/22167	0.28	0/30990

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	G	751	0	767	12	0
2	H	742	0	749	16	0
3	K	777	0	429	18	0
4	N	823	0	459	14	0
5	E	1654	0	918	32	0
6	F	1688	0	937	23	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	A	3909	0	3908	43	0
7	B	3861	0	3870	48	0
7	C	3824	0	3829	68	0
7	D	3196	0	3163	48	0
All	All	21225	0	19029	274	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 274 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:C:451:VAL:O	7:C:455:LEU:HB3	1.84	0.76
1:G:59:GLN:HA	1:G:73:ILE:O	1.88	0.73
7:D:273:LEU:HD11	7:D:480:PRO:HG2	1.71	0.72
4:N:31:DT:O2	7:C:118:ARG:NH2	2.24	0.69
7:C:222:ILE:HG12	7:C:248:LEU:HG	1.76	0.68

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	G	91/121 (75%)	90 (99%)	1 (1%)	0	100	100
2	H	92/94 (98%)	92 (100%)	0	0	100	100
7	A	467/603 (77%)	453 (97%)	14 (3%)	0	100	100
7	B	463/603 (77%)	454 (98%)	9 (2%)	0	100	100
7	C	458/603 (76%)	452 (99%)	6 (1%)	0	100	100
7	D	381/603 (63%)	373 (98%)	8 (2%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
All	All	1952/2627 (74%)	1914 (98%)	38 (2%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains ⓘ

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	G	80/106 (76%)	80 (100%)	0	100	100
2	H	78/79 (99%)	78 (100%)	0	100	100
7	A	432/541 (80%)	431 (100%)	1 (0%)	87	87
7	B	424/541 (78%)	424 (100%)	0	100	100
7	C	420/541 (78%)	419 (100%)	1 (0%)	87	87
7	D	353/541 (65%)	350 (99%)	3 (1%)	73	77
All	All	1787/2349 (76%)	1782 (100%)	5 (0%)	84	85

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
7	A	543	GLU
7	C	450	THR
7	D	155	VAL
7	D	226	VAL
7	D	472	HIS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 19 such sidechains are listed below:

Mol	Chain	Res	Type
7	C	494	ASN
7	D	320	GLN
7	D	383	ASN
7	D	277	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
7	B	403	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 oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

There are no ligands in this entry.

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.

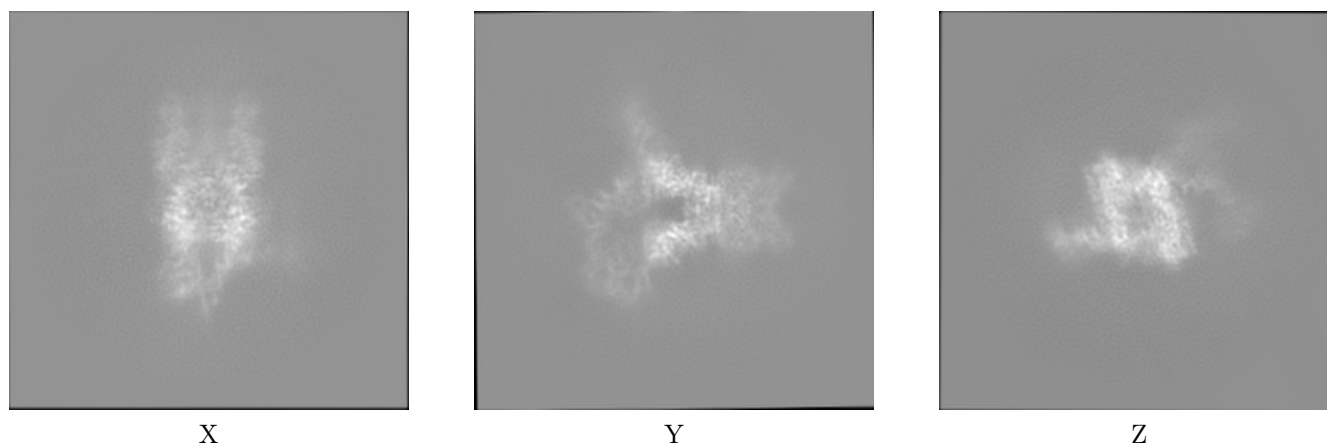
6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-72650. 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](#)

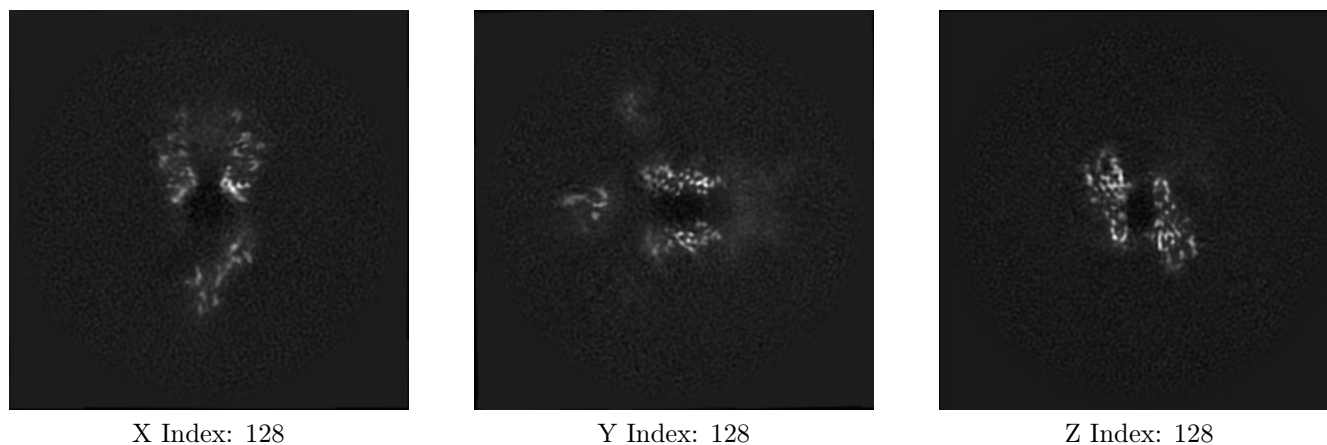
6.1.1 Primary map



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

6.2 Central slices [i](#)

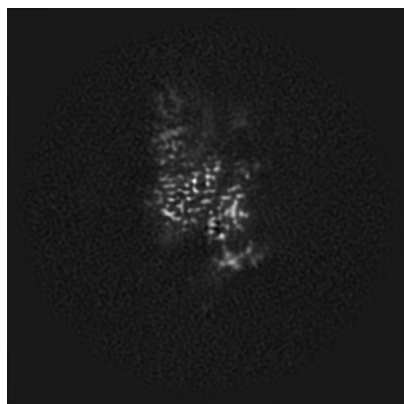
6.2.1 Primary map



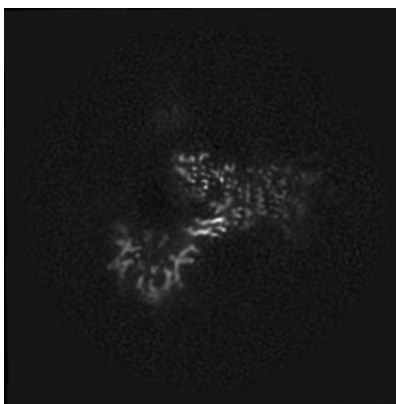
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices [i](#)

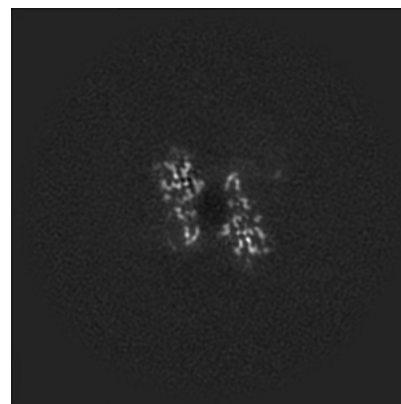
6.3.1 Primary map



X Index: 144



Y Index: 110

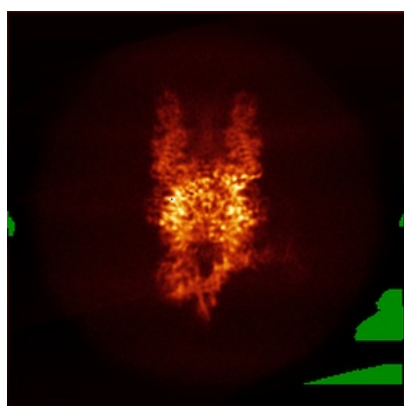


Z Index: 126

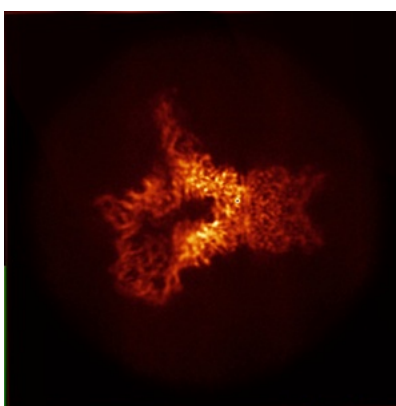
The images above show the largest variance slices of the map in three orthogonal directions.

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

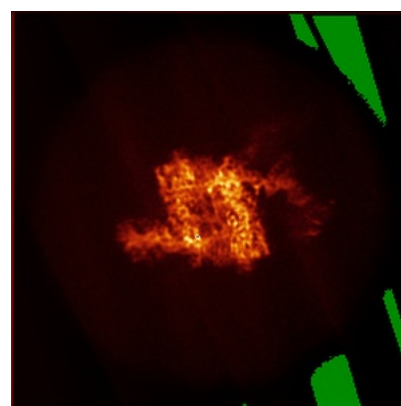
6.4.1 Primary map



X



Y

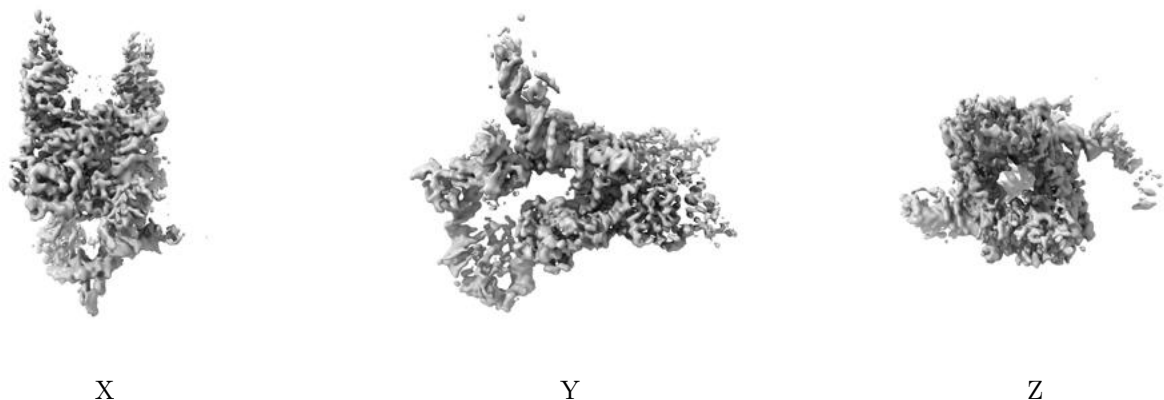


Z

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

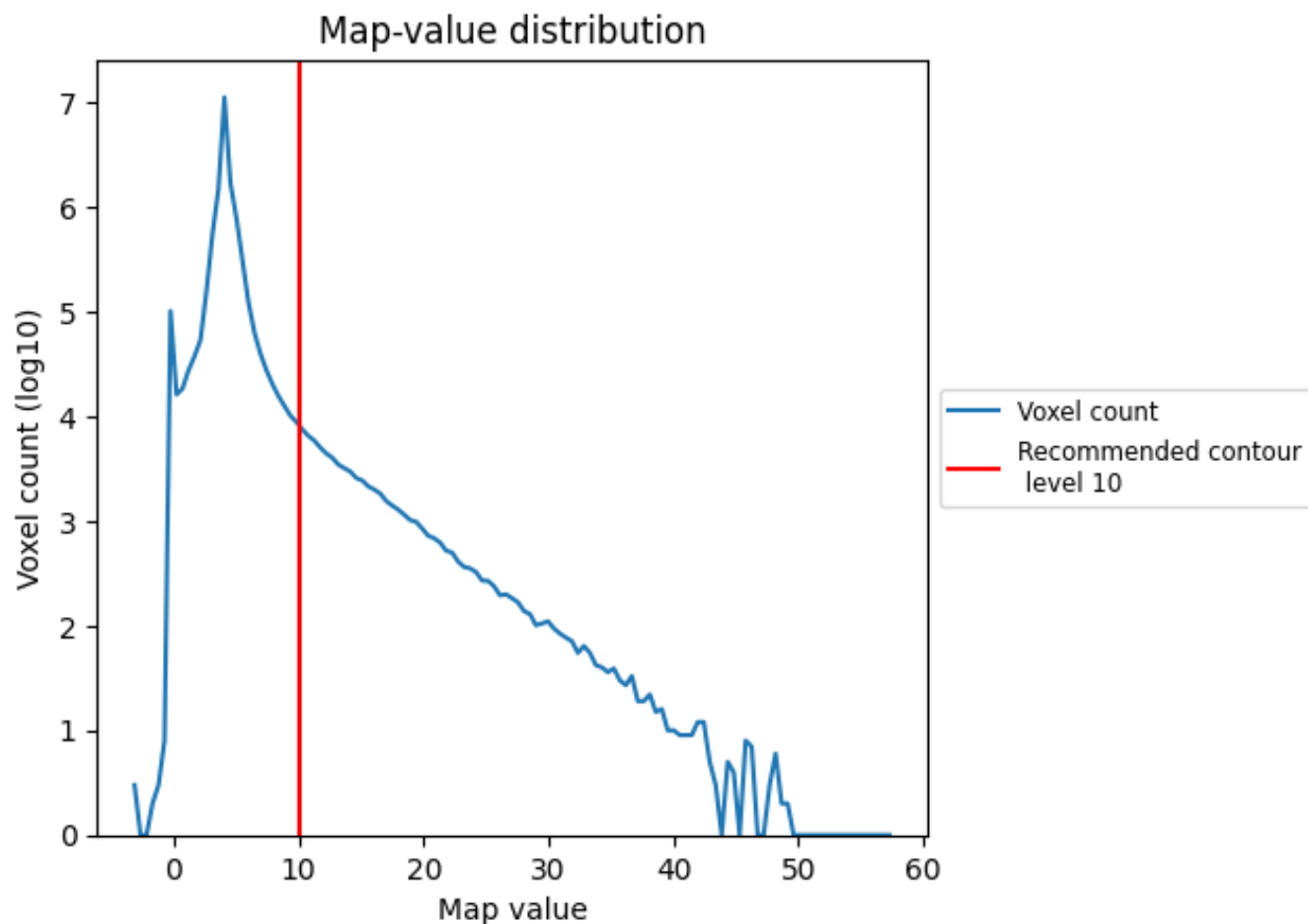
6.6 Mask visualisation [i](#)

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

7 Map analysis [i](#)

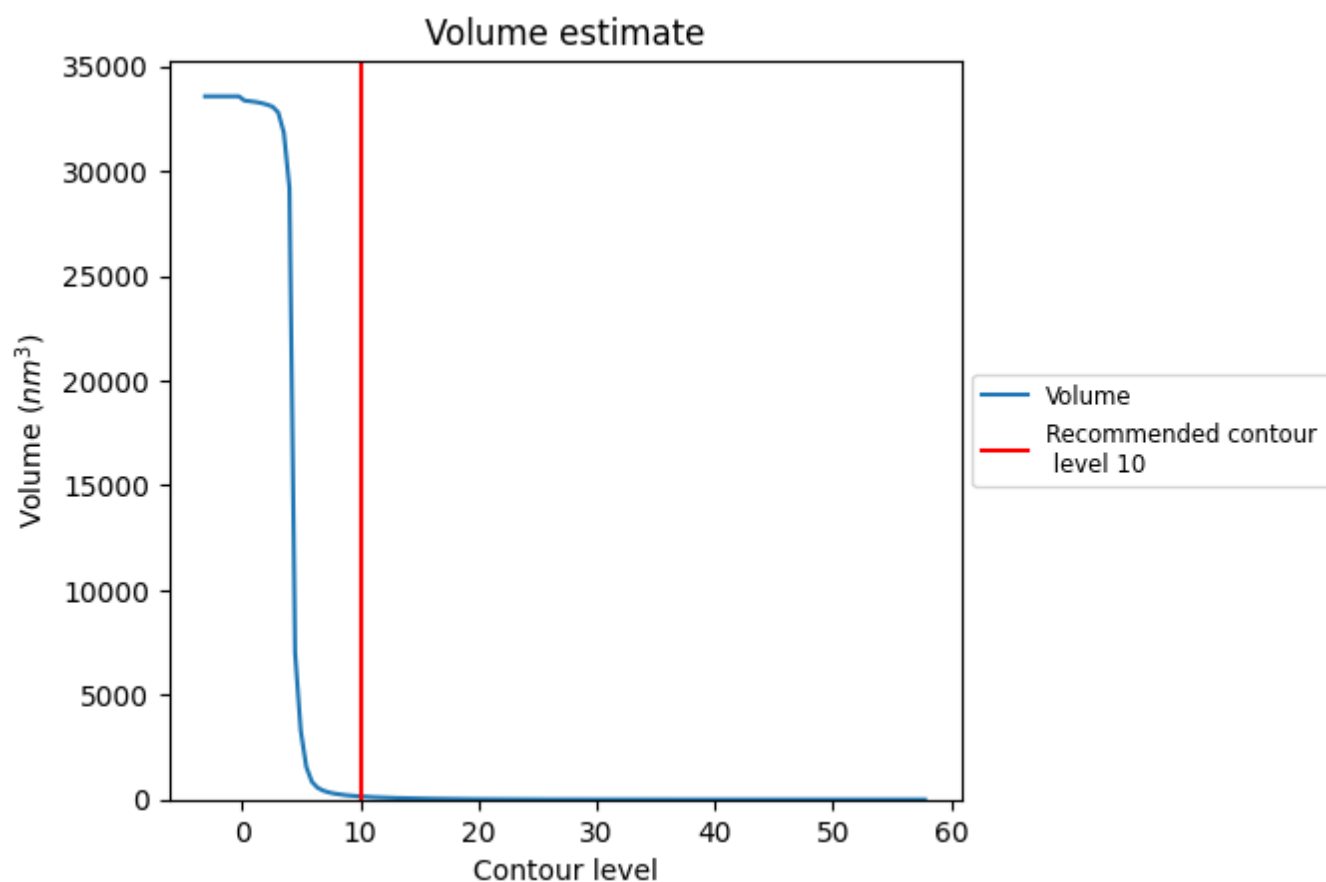
This section contains the results of statistical analysis of the map.

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.

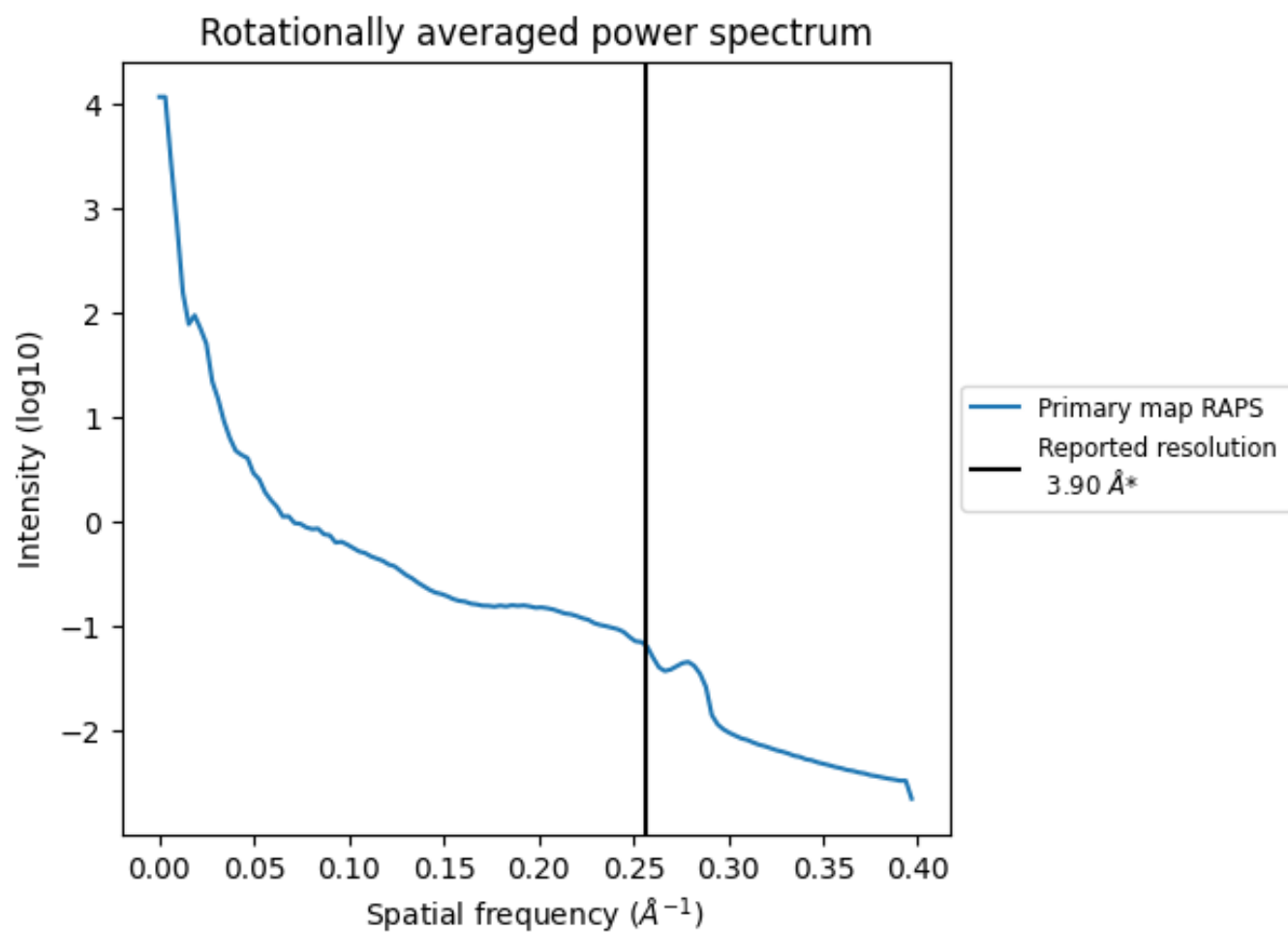
7.2 Volume estimate [i](#)



The volume at the recommended contour level is 150 nm³; this corresponds to an approximate mass of 136 kDa.

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

7.3 Rotationally averaged power spectrum ⓘ



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

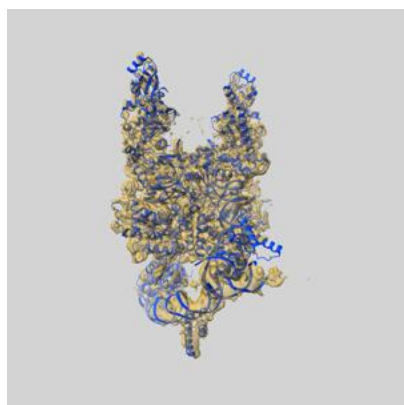
8 Fourier-Shell correlation

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

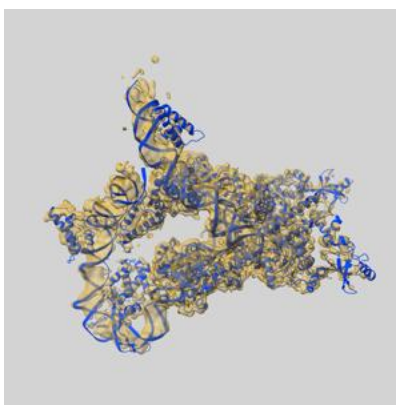
9 Map-model fit [i](#)

This section contains information regarding the fit between EMDB map EMD-72650 and PDB model 9Y7B. Per-residue inclusion information can be found in section [3](#) on page [6](#).

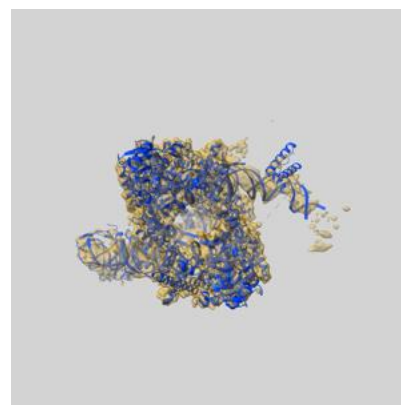
9.1 Map-model overlay [i](#)



X



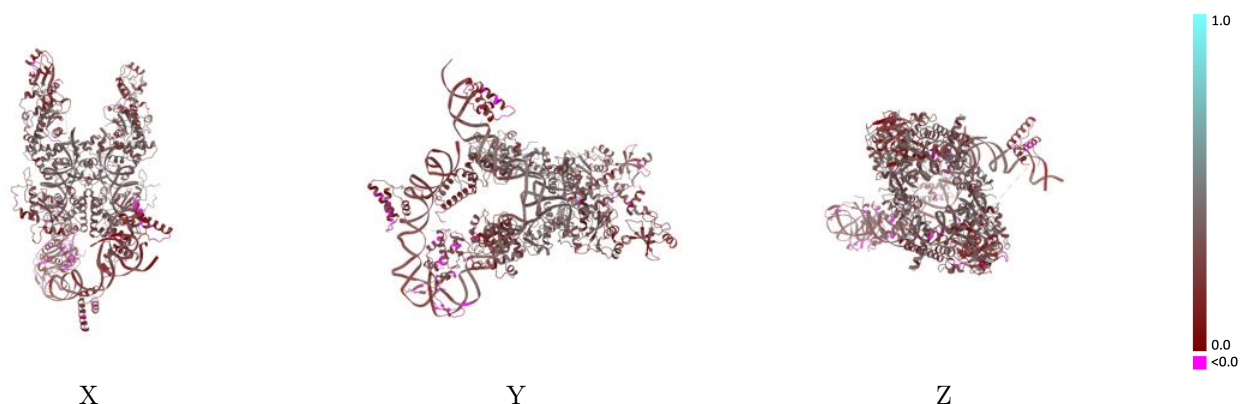
Y



Z

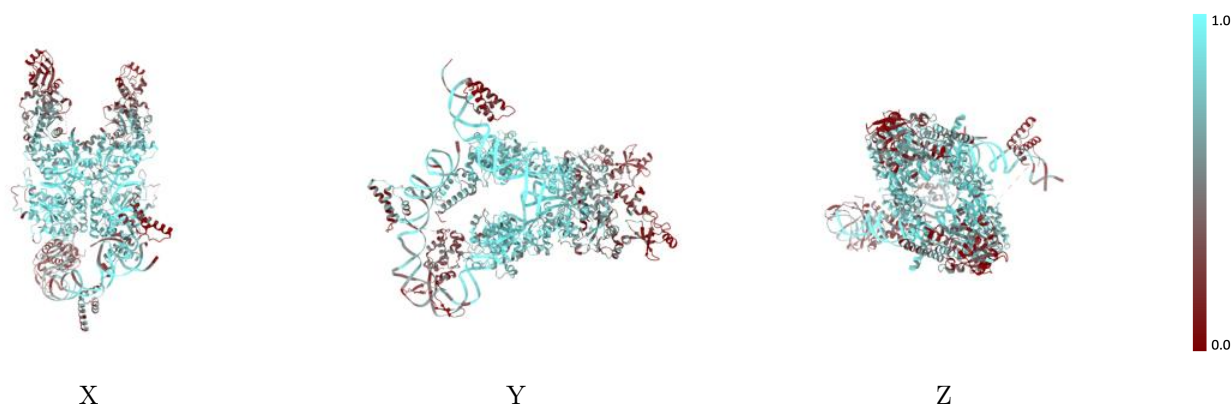
The images above show the 3D surface view of the map at the recommended contour level 10.0 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

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



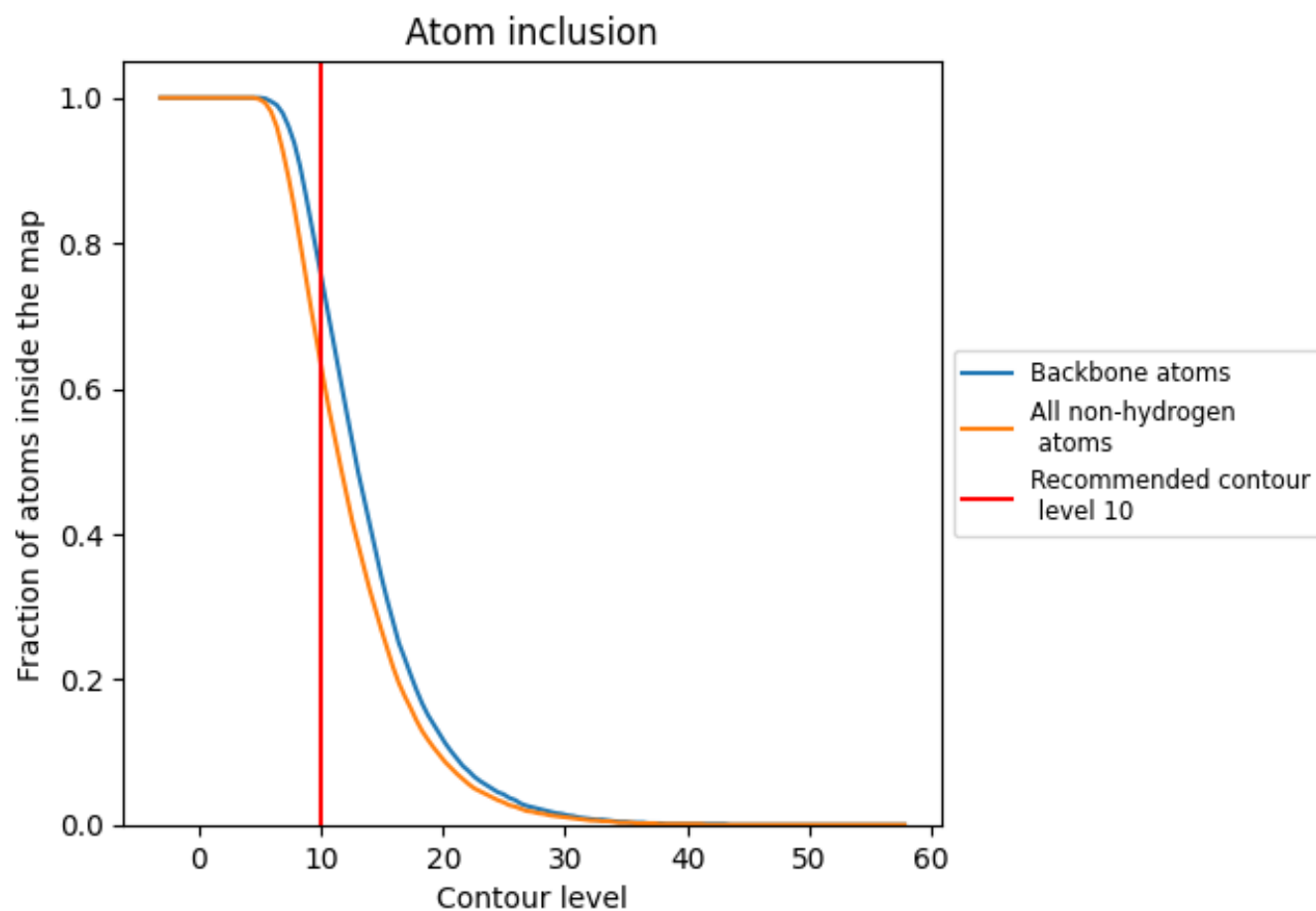
The images above show the model with each residue coloured according to its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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 (10).

9.4 Atom inclusion [i](#)



At the recommended contour level, 76% of all backbone atoms, 63% 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 (10) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.6320	0.2930
A	0.6790	0.2910
B	0.5560	0.3120
C	0.5610	0.3170
D	0.6550	0.3020
E	0.7820	0.2840
F	0.7790	0.2810
G	0.2530	0.1440
H	0.2420	0.1400
K	0.8820	0.3360
N	0.8350	0.3330

1.0
0.0
<0.0