



## wwPDB EM Validation Summary Report ⓘ

Nov 21, 2022 – 06:57 AM EST

PDB ID : 7SCY  
EMDB ID : EMD-25042  
Title : Nuc147 bound to single BRCT  
Authors : Muthurajan, U.M.; Rudolph, J.R.  
Deposited on : 2021-09-29  
Resolution : 4.10 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](mailto: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) : 2.31.3

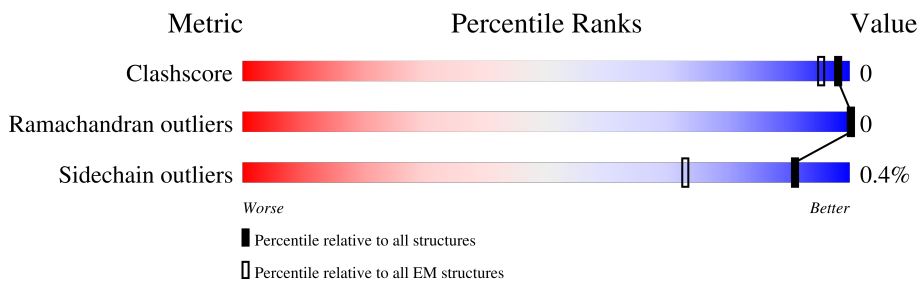
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 4.10 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  $\geq 3$ , 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	I	147	7% 52% 39% .
2	J	147	5% 53% 41% .
3	A	139	63% 5% . 31%
3	E	139	65% 6% 29%
4	B	106	68% . . 27%
4	F	106	6% 73% 9% 18%
5	C	133	75% 6% 19%
5	G	133	74% . 22%

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
6	D	129	 69% 29%
6	H	129	 69% 29%
7	K	131	 5% 71% 27%

## 2 Entry composition [i](#)

There are 7 unique types of molecules in this entry. The entry contains 23206 atoms, of which 10436 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 DNA chain called DNA (147-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace	
			Total	C	H	N	O			P
1	I	147	4682	1435	1651	566	883	147	0	0

- Molecule 2 is a DNA chain called DNA (147-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace	
			Total	C	H	N	O			P
2	J	147	4646	1423	1650	545	881	147	0	0

- Molecule 3 is a protein called Histone H3.1.

Mol	Chain	Residues	Atoms					AltConf	Trace	
			Total	C	H	N	O			S
3	A	96	1616	499	826	151	136	4	0	0
3	E	99	1684	517	865	159	139	4	0	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-3	GLY	-	expression tag	UNP P68431
A	-2	SER	-	expression tag	UNP P68431
A	-1	HIS	-	expression tag	UNP P68431
E	-3	GLY	-	expression tag	UNP P68431
E	-2	SER	-	expression tag	UNP P68431
E	-1	HIS	-	expression tag	UNP P68431

- Molecule 4 is a protein called Histone H4.

Mol	Chain	Residues	Atoms					AltConf	Trace	
			Total	C	H	N	O			S
4	B	77	1271	389	657	119	105	1	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Residues	Atoms					AltConf	Trace	
			Total	C	H	N	O			S
4	F	87	1459	442	756	142	118	1	0	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-3	GLY	-	expression tag	UNP P62805
B	-2	SER	-	expression tag	UNP P62805
B	-1	HIS	-	expression tag	UNP P62805
F	-3	GLY	-	expression tag	UNP P62805
F	-2	SER	-	expression tag	UNP P62805
F	-1	HIS	-	expression tag	UNP P62805

- Molecule 5 is a protein called Histone H2A.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	H	N	O		
5	C	108	1733	526	898	165	144	0	0
5	G	104	1667	508	862	157	140	0	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	-3	GLY	-	expression tag	UNP Q08AJ9
C	-2	SER	-	expression tag	UNP Q08AJ9
C	-1	HIS	-	expression tag	UNP Q08AJ9
G	-3	GLY	-	expression tag	UNP Q08AJ9
G	-2	SER	-	expression tag	UNP Q08AJ9
G	-1	HIS	-	expression tag	UNP Q08AJ9

- Molecule 6 is a protein called Histone H2B type 1-J.

Mol	Chain	Residues	Atoms					AltConf	Trace	
			Total	C	H	N	O			S
6	D	92	1461	453	741	129	136	2	0	0
6	H	91	1450	450	736	128	134	2	0	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	-3	GLY	-	expression tag	UNP P06899
D	-2	SER	-	expression tag	UNP P06899
D	-1	HIS	-	expression tag	UNP P06899
H	-3	GLY	-	expression tag	UNP P06899
H	-2	SER	-	expression tag	UNP P06899
H	-1	HIS	-	expression tag	UNP P06899

- Molecule 7 is a protein called Poly [ADP-ribose] polymerase 1.

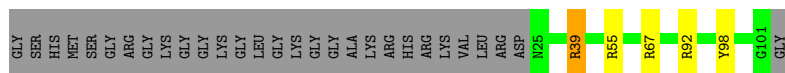
Mol	Chain	Residues	Atoms					AltConf	Trace	
			Total	C	H	N	O			S
7	K	96	1537	468	794	127	143	5	0	0

There are 23 discrepancies between the modelled and reference sequences:

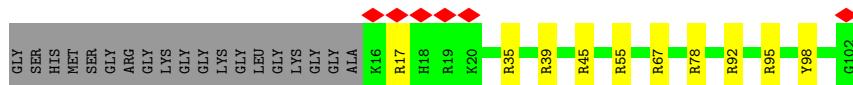
Chain	Residue	Modelled	Actual	Comment	Reference
K	364	MET	-	initiating methionine	UNP P09874
K	365	GLY	-	expression tag	UNP P09874
K	366	SER	-	expression tag	UNP P09874
K	367	SER	-	expression tag	UNP P09874
K	368	HIS	-	expression tag	UNP P09874
K	369	HIS	-	expression tag	UNP P09874
K	370	HIS	-	expression tag	UNP P09874
K	371	HIS	-	expression tag	UNP P09874
K	372	HIS	-	expression tag	UNP P09874
K	373	HIS	-	expression tag	UNP P09874
K	374	SER	-	expression tag	UNP P09874
K	375	SER	-	expression tag	UNP P09874
K	376	GLY	-	expression tag	UNP P09874
K	377	LEU	-	expression tag	UNP P09874
K	378	VAL	-	expression tag	UNP P09874
K	379	PRO	-	expression tag	UNP P09874
K	380	ARG	-	expression tag	UNP P09874
K	381	GLY	-	expression tag	UNP P09874
K	382	SER	-	expression tag	UNP P09874
K	383	HIS	-	expression tag	UNP P09874
K	384	MET	-	expression tag	UNP P09874
K	487	GLY	-	insertion	UNP P09874
K	488	SER	-	insertion	UNP P09874



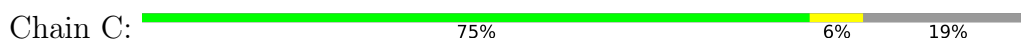
• Molecule 4: Histone H4



• Molecule 4: Histone H4



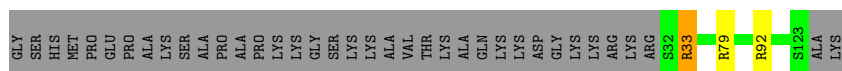
• Molecule 5: Histone H2A



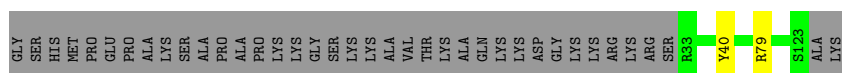
• Molecule 5: Histone H2A



• Molecule 6: Histone H2B type 1-J



• Molecule 6: Histone H2B type 1-J



• Molecule 7: Poly [ADP-ribose] polymerase 1





## 4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	16495	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	NONE	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	60	Depositor
Minimum defocus (nm)	Not provided	
Maximum defocus (nm)	Not provided	
Magnification	Not provided	
Image detector	GATAN K3 (6k x 4k)	Depositor
Maximum map value	0.054	Depositor
Minimum map value	-0.018	Depositor
Average map value	0.000	Depositor
Map value standard deviation	0.002	Depositor
Recommended contour level	0.0095	Depositor
Map size ( $\text{\AA}$ )	272.64, 272.64, 272.64	wwPDB
Map dimensions	256, 256, 256	wwPDB
Map angles ( $^\circ$ )	90.0, 90.0, 90.0	wwPDB
Pixel spacing ( $\text{\AA}$ )	1.065, 1.065, 1.065	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	I	1.68	7/3403 (0.2%)	2.37	287/5255 (5.5%)
2	J	1.69	4/3357 (0.1%)	2.41	292/5174 (5.6%)
3	A	0.71	0/802	1.09	7/1076 (0.7%)
3	E	0.70	0/831	1.08	9/1113 (0.8%)
4	B	0.72	0/621	1.10	5/832 (0.6%)
4	F	0.75	0/711	1.19	11/948 (1.2%)
5	C	0.66	0/845	1.04	7/1139 (0.6%)
5	G	0.66	0/815	1.00	4/1100 (0.4%)
6	D	0.64	0/731	0.93	3/983 (0.3%)
6	H	0.64	0/725	0.90	1/975 (0.1%)
7	K	0.54	0/749	0.92	1/999 (0.1%)
All	All	1.28	11/13590 (0.1%)	1.88	627/19594 (3.2%)

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	I	0	69
2	J	0	69
3	A	0	1
4	B	0	1
6	D	0	1
6	H	0	1
All	All	0	142

The worst 5 of 11 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	I	52	DC	C4-N4	-5.42	1.29	1.33
2	J	11	DG	C2-N2	-5.30	1.29	1.34
1	I	61	DC	C4-N4	-5.29	1.29	1.33

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	J	4	DC	C4-N4	-5.25	1.29	1.33
1	I	-36	DG	C2-N2	-5.17	1.29	1.34

The worst 5 of 627 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	J	-9	DA	N1-C6-N6	-10.89	112.07	118.60
2	J	-69	DA	N1-C6-N6	-10.21	112.48	118.60
2	J	-73	DA	N1-C6-N6	-10.00	112.60	118.60
1	I	-55	DA	N1-C6-N6	-9.90	112.66	118.60
1	I	43	DA	N1-C6-N6	-9.69	112.79	118.60

There are no chirality outliers.

5 of 142 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	I	-66	DG	Sidechain
1	I	-67	DT	Sidechain
1	I	-68	DA	Sidechain
1	I	-69	DG	Sidechain
1	I	-72	DT	Sidechain

## 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	I	3031	1651	1637	3	0
2	J	2996	1650	1639	2	0
3	A	790	826	826	0	0
3	E	819	865	864	0	0
4	B	614	657	656	0	0
4	F	703	756	755	0	0
5	C	835	898	897	0	0
5	G	805	862	861	0	0
6	D	720	741	740	0	0
6	H	714	736	735	0	0
7	K	743	794	792	1	0

Continued on next page...

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	12770	10436	10402	6	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 0.

The worst 5 of 6 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:I:15:DT:H1'	1:I:16:DA:C8	2.52	0.44
2:J:6:DC:H1'	2:J:7:DC:C6	2.53	0.43
1:I:16:DA:H1'	1:I:17:DA:C8	2.55	0.42
7:K:470:GLN:HA	7:K:473:PHE:CD2	2.55	0.42
1:I:69:DT:H2''	1:I:70:DC:C6	2.56	0.41

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	
3	A	94/139 (68%)	91 (97%)	3 (3%)	0	100	100
3	E	97/139 (70%)	94 (97%)	3 (3%)	0	100	100
4	B	75/106 (71%)	73 (97%)	2 (3%)	0	100	100
4	F	85/106 (80%)	81 (95%)	4 (5%)	0	100	100
5	C	106/133 (80%)	104 (98%)	2 (2%)	0	100	100
5	G	102/133 (77%)	99 (97%)	3 (3%)	0	100	100
6	D	90/129 (70%)	89 (99%)	1 (1%)	0	100	100
6	H	89/129 (69%)	88 (99%)	1 (1%)	0	100	100
7	K	94/131 (72%)	90 (96%)	4 (4%)	0	100	100

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
All	All	832/1145 (73%)	809 (97%)	23 (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	
3	A	84/113 (74%)	83 (99%)	1 (1%)	71	83
3	E	87/113 (77%)	87 (100%)	0	100	100
4	B	63/81 (78%)	63 (100%)	0	100	100
4	F	72/81 (89%)	72 (100%)	0	100	100
5	C	85/102 (83%)	84 (99%)	1 (1%)	71	83
5	G	83/102 (81%)	82 (99%)	1 (1%)	71	83
6	D	79/107 (74%)	79 (100%)	0	100	100
6	H	78/107 (73%)	78 (100%)	0	100	100
7	K	85/113 (75%)	85 (100%)	0	100	100
All	All	716/919 (78%)	713 (100%)	3 (0%)	91	94

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

Mol	Chain	Res	Type
3	A	63	ARG
5	C	79	ILE
5	G	79	ILE

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](#)

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.

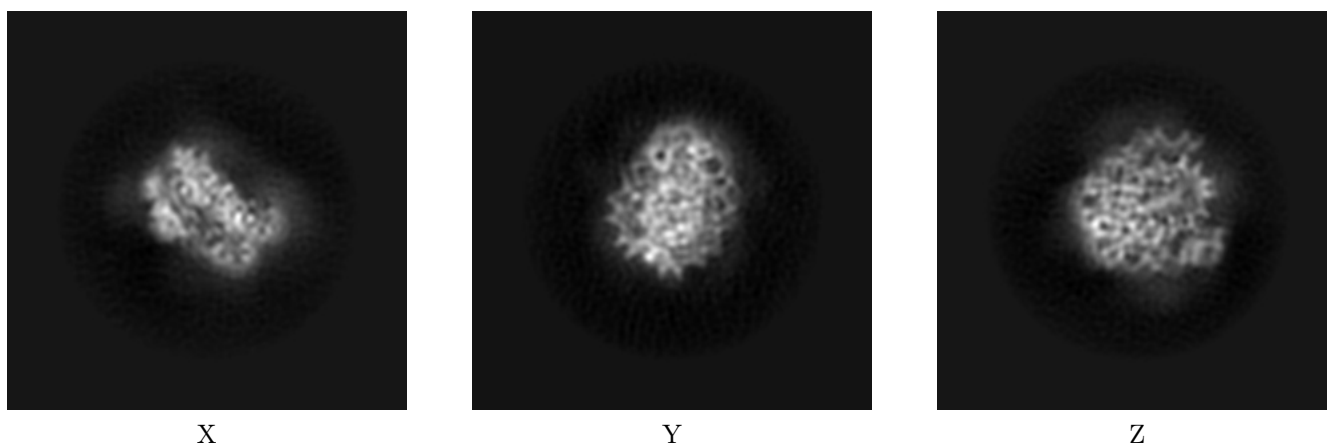
## 6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-25042. 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](#)

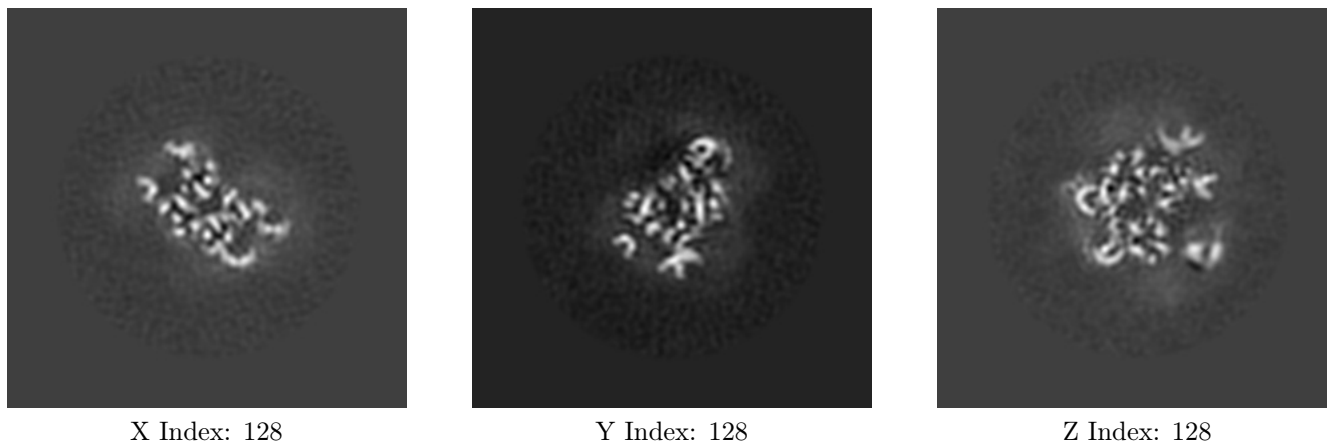
#### 6.1.1 Primary map



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

### 6.2 Central slices [i](#)

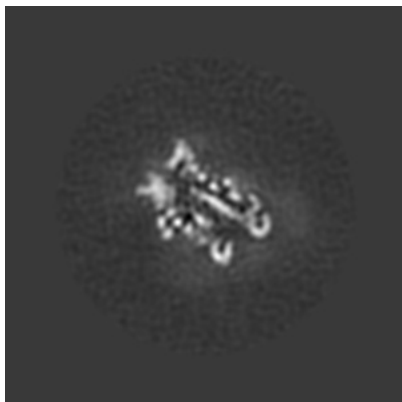
#### 6.2.1 Primary map



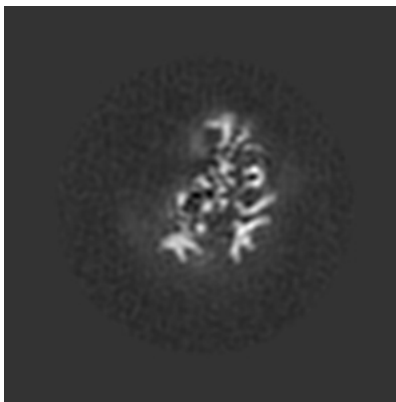
The images above show central slices of the map in three orthogonal directions.

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

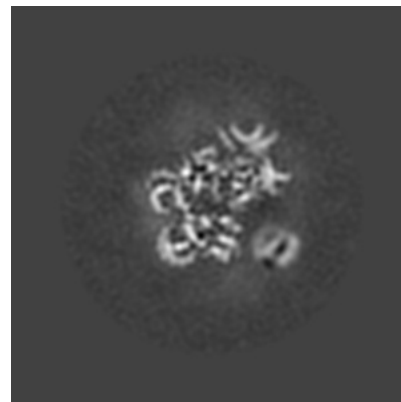
### 6.3.1 Primary map



X Index: 118



Y Index: 113

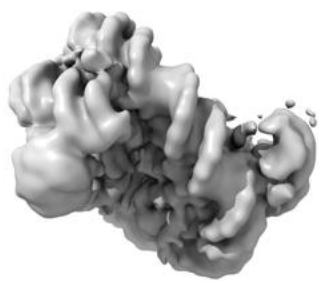


Z Index: 125

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

## 6.4 Orthogonal surface views [i](#)

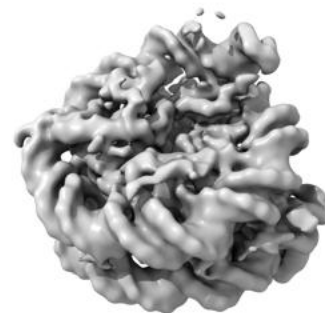
### 6.4.1 Primary map



X



Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.0095. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.



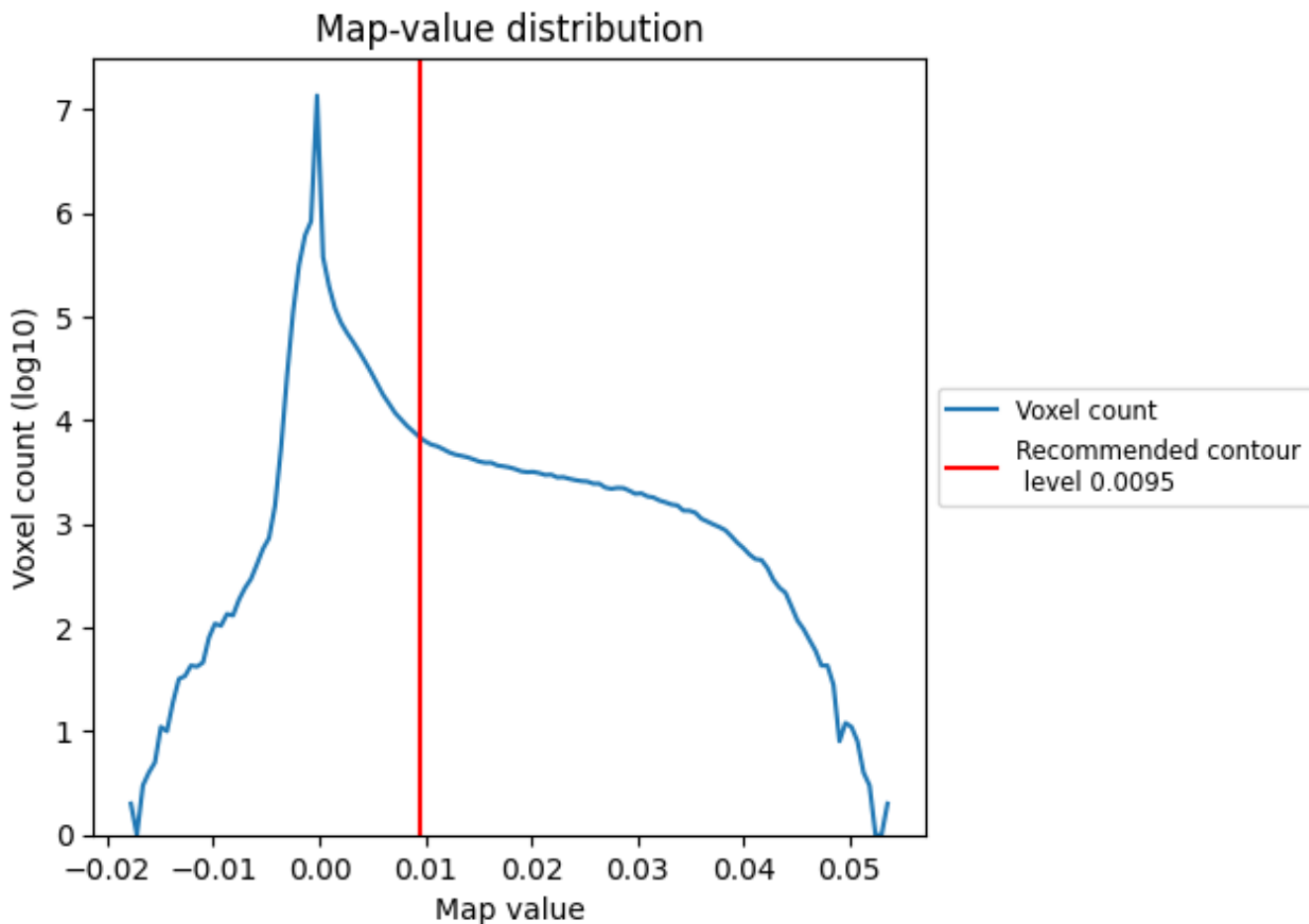
## 6.5 Mask visualisation

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

## 7 Map analysis [i](#)

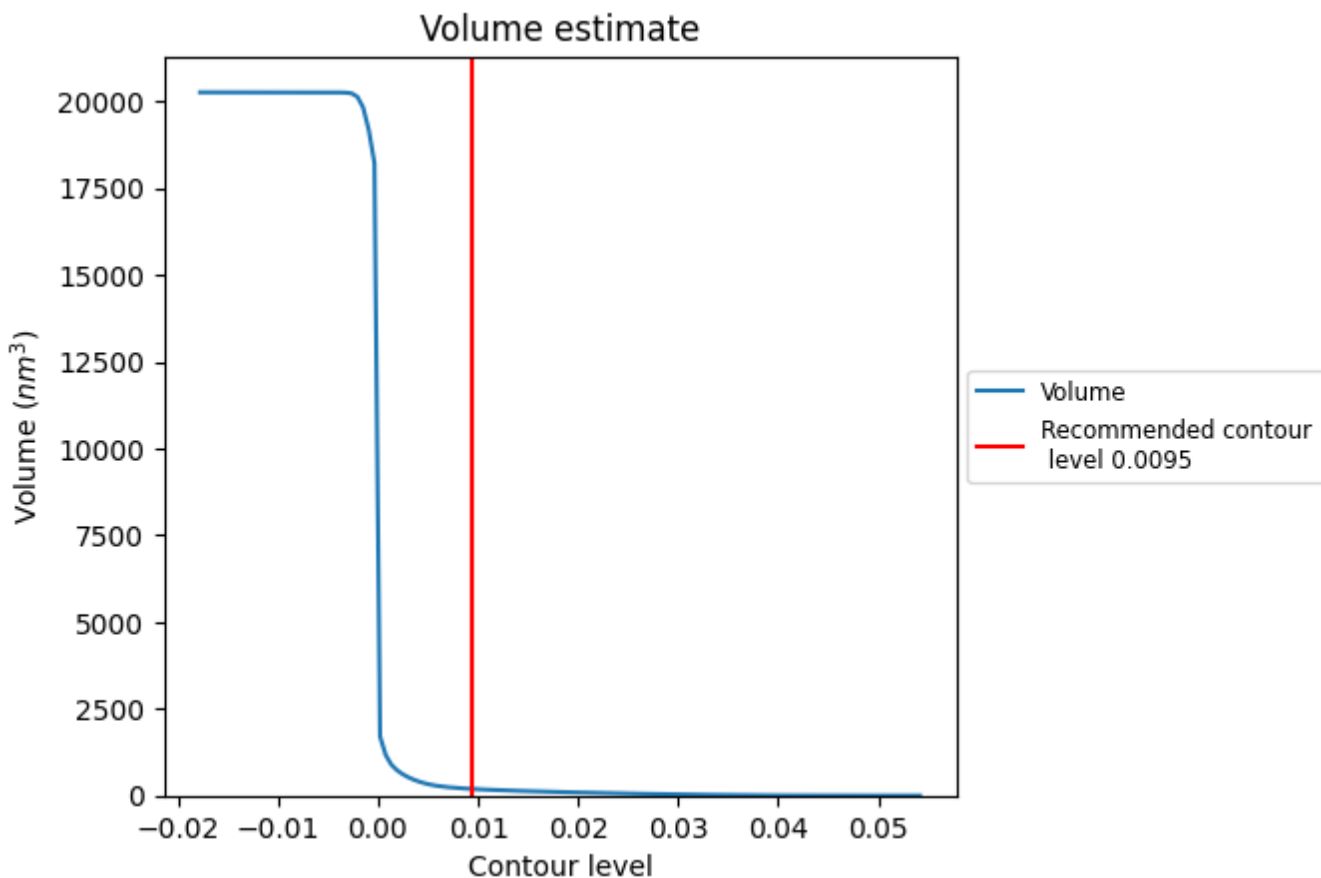
This section contains the results of statistical analysis of the map.

### 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.

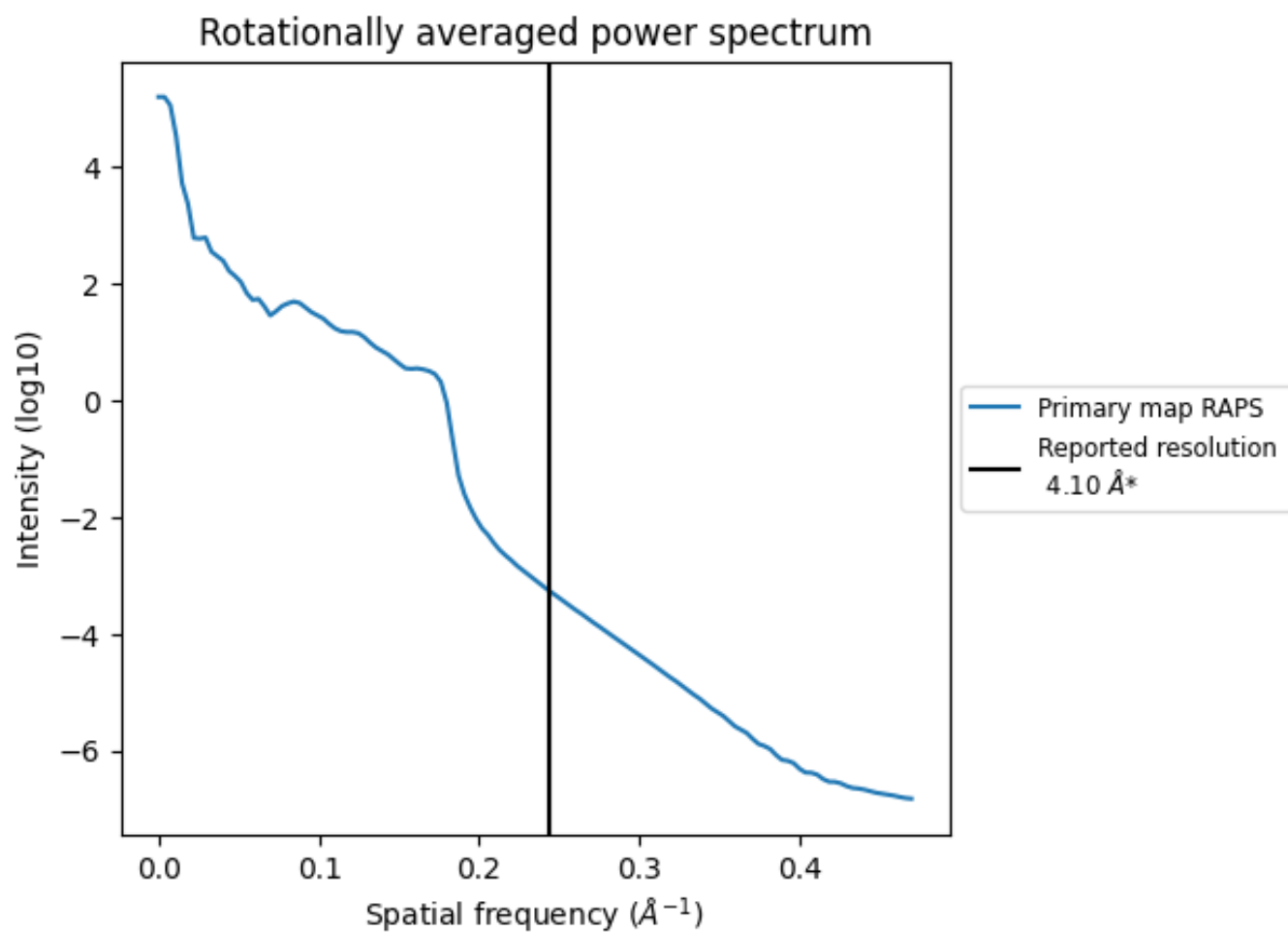
## 7.2 Volume estimate [i](#)



The volume at the recommended contour level is 189  $\text{nm}^3$ ; this corresponds to an approximate mass of 171 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

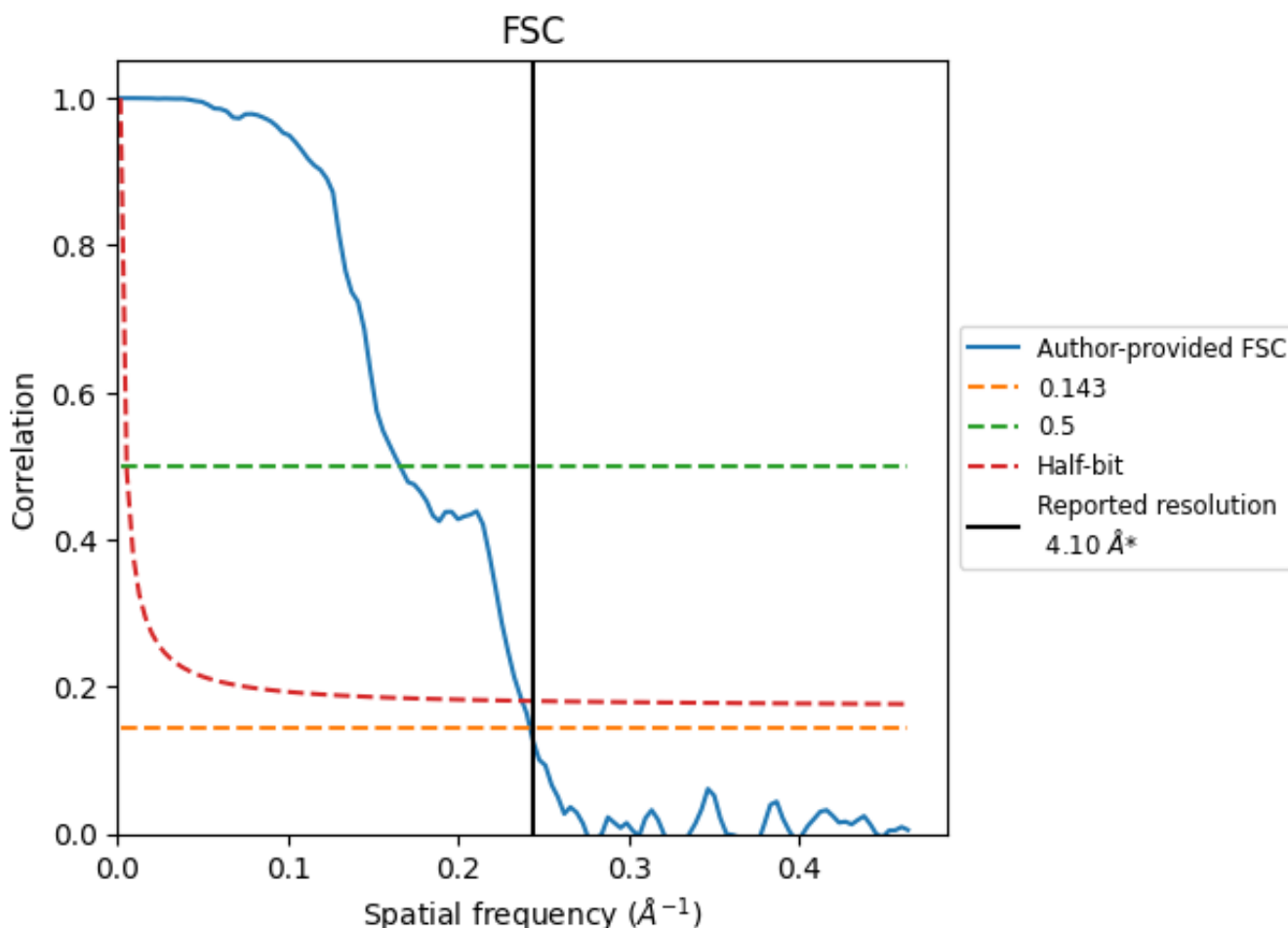


\*Reported resolution corresponds to spatial frequency of  $0.244 \text{ \AA}^{-1}$

## 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 0.244 Å<sup>-1</sup>

## 8.2 Resolution estimates [i](#)

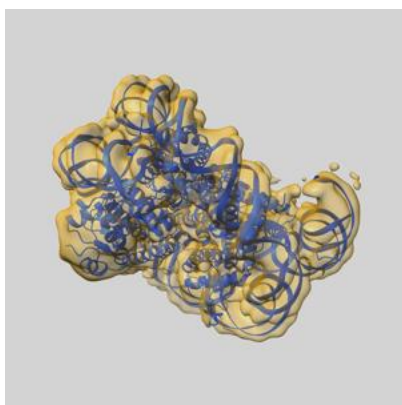
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	4.10	-	-
Author-provided FSC curve	4.13	6.03	4.21
Unmasked-calculated*	-	-	-

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

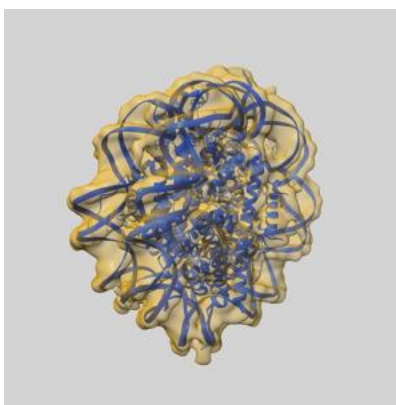
## 9 Map-model fit [i](#)

This section contains information regarding the fit between EMDB map EMD-25042 and PDB model 7SCY. Per-residue inclusion information can be found in section 3 on page 7.

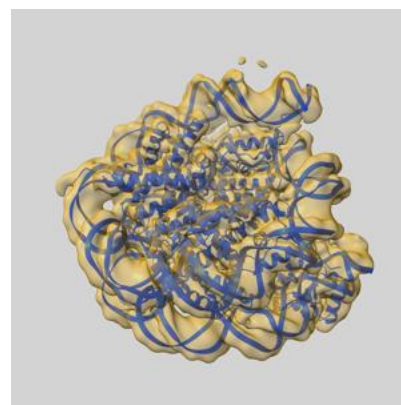
### 9.1 Map-model overlay [i](#)



X



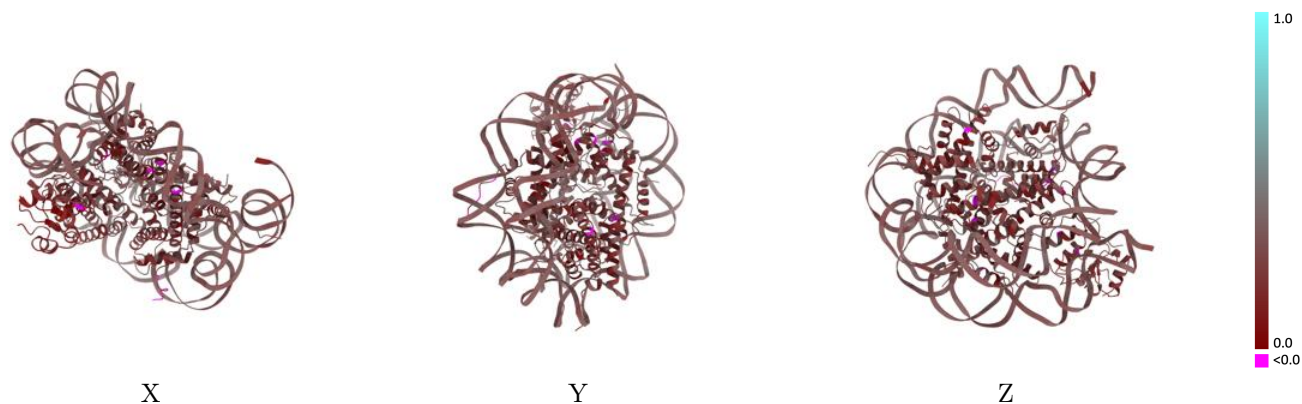
Y



Z

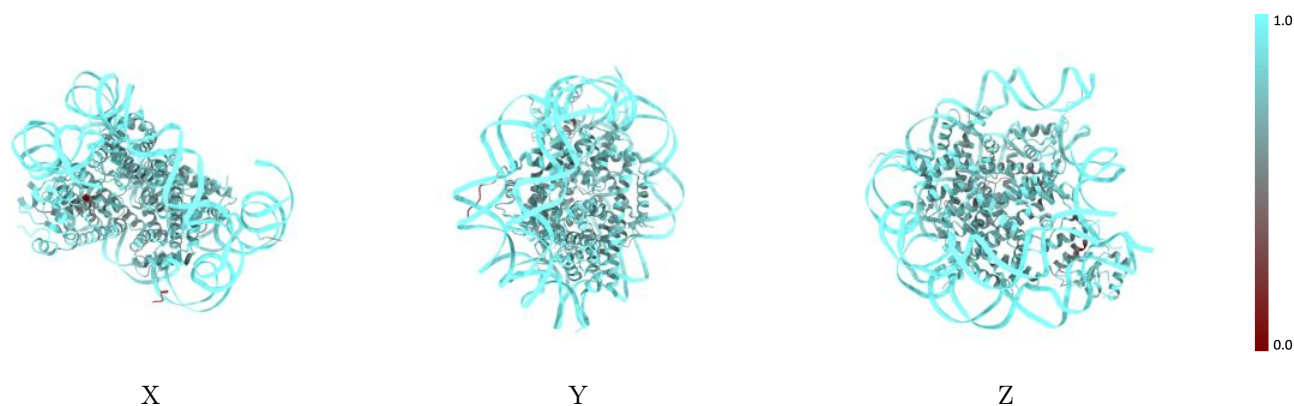
The images above show the 3D surface view of the map at the recommended contour level 0.0095 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

## 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.

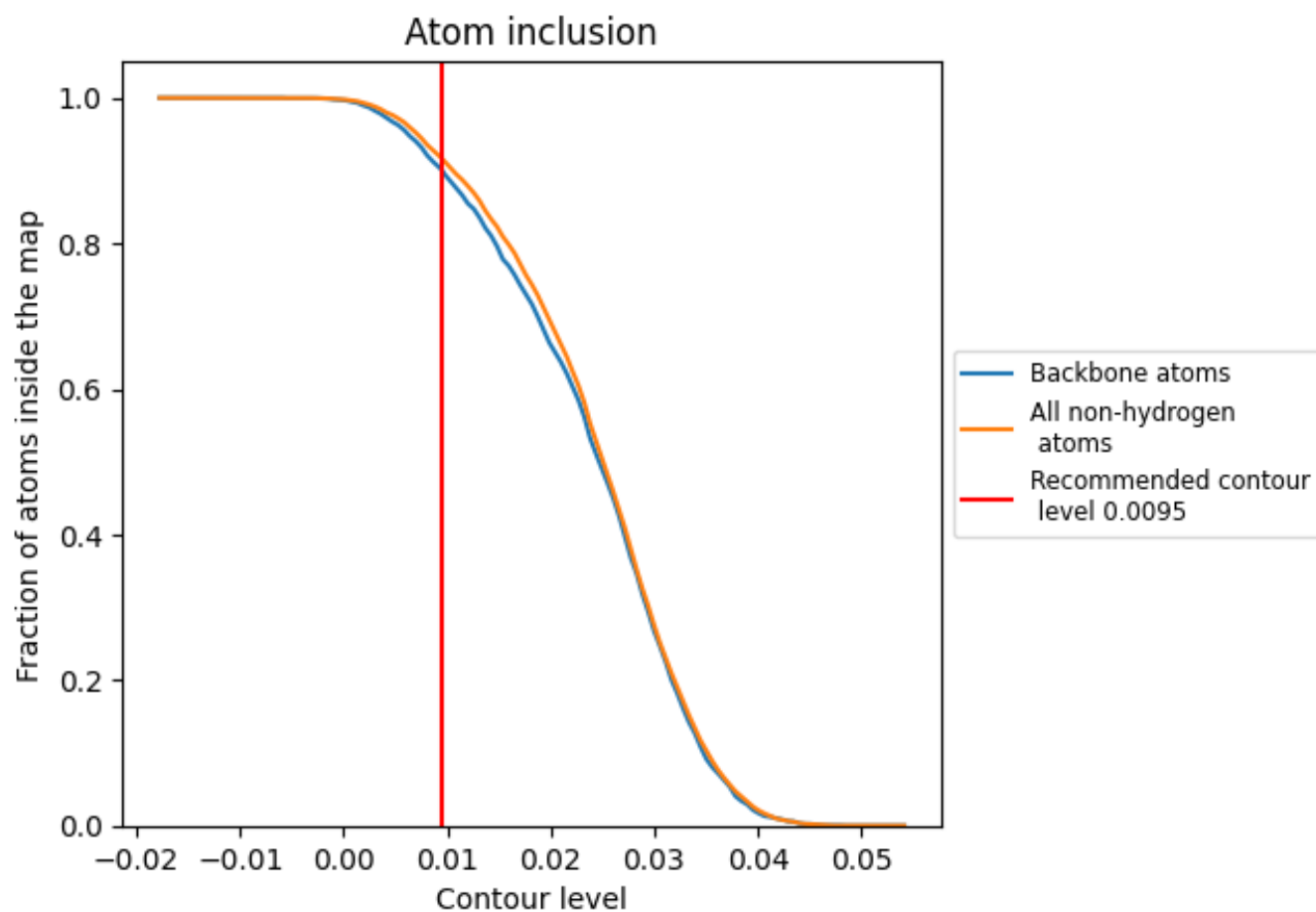
## 9.3 Atom inclusion mapped to coordinate model [i](#)



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 (0.0095).



























## 9.4 Atom inclusion [i](#)



At the recommended contour level, 90% 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 (0.0095) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	 0.9168	 0.2730
A	 0.8344	 0.2320
B	 0.8576	 0.2360
C	 0.8564	 0.2490
D	 0.8478	 0.2320
E	 0.8198	 0.2270
F	 0.7875	 0.2160
G	 0.8513	 0.2480
H	 0.8364	 0.2310
I	 0.9987	 0.3230
J	 0.9983	 0.3250
K	 0.7954	 0.1690

