



## Full wwPDB EM Validation Report ⓘ

Jan 31, 2023 – 07:08 pm GMT

PDB ID : 7Q59  
EMDB ID : EMD-13829  
Title : Cryo-EM structure of Mycobacterium tuberculosis RNA polymerase holoenzyme dimer comprising sigma factor SigB  
Authors : Brodolin, K.  
Deposited on : 2021-11-03  
Resolution : 4.36 Å (reported)  
Based on initial model : 7PP4

This is a Full wwPDB EM Validation 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:

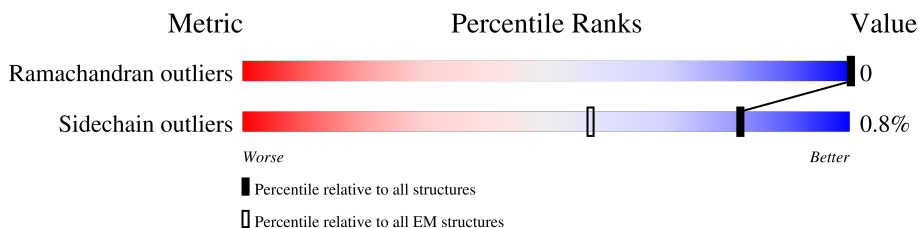
EMDB validation analysis : 0.0.1.dev43  
MolProbity : 4.02b-467  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
MapQ : 1.9.9  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.32.1

# 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 4.36 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)
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	A	347	64% 36%
1	B	347	66% 34%
1	a	347	64% 36%
1	b	347	66% 34%
2	C	1174	95% 5%
2	c	1174	95% 5%
3	D	1319	96%
3	d	1319	95%
4	E	110	75% 25%

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
4	e	110	 <p>75% 25%</p>
5	F	343	 <p>39% 59%</p>
5	f	343	 <p>40% 59%</p>

## 2 Entry composition i

There are 7 unique types of molecules in this entry. The entry contains 47632 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	223	Total	C	N	O	S	0	0
			1700	1068	294	336	2		
1	B	230	Total	C	N	O	S	0	0
			1749	1099	301	347	2		
1	a	223	Total	C	N	O	S	0	0
			1700	1068	294	336	2		
1	b	230	Total	C	N	O	S	0	0
			1749	1099	301	347	2		

- 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	1112	Total	C	N	O	S	0	0
			8640	5408	1515	1678	39		
2	c	1112	Total	C	N	O	S	0	0
			8640	5408	1515	1678	39		

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	5	MET	-	initiating methionine	UNP P9WGY9
C	6	VAL	ILE	conflict	UNP P9WGY9
c	5	MET	-	initiating methionine	UNP P9WGY9
c	6	VAL	ILE	conflict	UNP P9WGY9

- 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	1269	Total	C	N	O	S	0	0
			9951	6231	1805	1874	41		
3	d	1269	Total	C	N	O	S	0	0
			9951	6231	1805	1874	41		

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	1317	HIS	-	expression tag	UNP P9WGY7
D	1318	HIS	-	expression tag	UNP P9WGY7
D	1319	HIS	-	expression tag	UNP P9WGY7
D	1320	HIS	-	expression tag	UNP P9WGY7
D	1321	HIS	-	expression tag	UNP P9WGY7
D	1322	HIS	-	expression tag	UNP P9WGY7
d	1317	HIS	-	expression tag	UNP P9WGY7
d	1318	HIS	-	expression tag	UNP P9WGY7
d	1319	HIS	-	expression tag	UNP P9WGY7
d	1320	HIS	-	expression tag	UNP P9WGY7
d	1321	HIS	-	expression tag	UNP P9WGY7
d	1322	HIS	-	expression tag	UNP P9WGY7

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

Mol	Chain	Residues	Atoms				AltConf	Trace
4	E	83	Total	C	N	O	0	0
			654	417	108	129		
4	e	83	Total	C	N	O	0	0
			654	417	108	129		

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

Mol	Chain	Residues	Atoms				AltConf	Trace	
5	F	142	Total	C	N	O	S	0	0
			1119	705	210	201	3		
5	f	142	Total	C	N	O	S	0	0
			1119	705	210	201	3		

There are 40 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
F	-19	MET	-	initiating methionine	UNP P9WGI5
F	-18	GLY	-	expression tag	UNP P9WGI5
F	-17	SER	-	expression tag	UNP P9WGI5
F	-16	SER	-	expression tag	UNP P9WGI5
F	-15	HIS	-	expression tag	UNP P9WGI5
F	-14	HIS	-	expression tag	UNP P9WGI5
F	-13	HIS	-	expression tag	UNP P9WGI5
F	-12	HIS	-	expression tag	UNP P9WGI5
F	-11	HIS	-	expression tag	UNP P9WGI5

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	-10	HIS	-	expression tag	UNP P9WGI5
F	-9	SER	-	expression tag	UNP P9WGI5
F	-8	SER	-	expression tag	UNP P9WGI5
F	-7	GLY	-	expression tag	UNP P9WGI5
F	-6	LEU	-	expression tag	UNP P9WGI5
F	-5	VAL	-	expression tag	UNP P9WGI5
F	-4	PRO	-	expression tag	UNP P9WGI5
F	-3	ARG	-	expression tag	UNP P9WGI5
F	-2	GLY	-	expression tag	UNP P9WGI5
F	-1	SER	-	expression tag	UNP P9WGI5
F	0	HIS	-	expression tag	UNP P9WGI5
f	-19	MET	-	initiating methionine	UNP P9WGI5
f	-18	GLY	-	expression tag	UNP P9WGI5
f	-17	SER	-	expression tag	UNP P9WGI5
f	-16	SER	-	expression tag	UNP P9WGI5
f	-15	HIS	-	expression tag	UNP P9WGI5
f	-14	HIS	-	expression tag	UNP P9WGI5
f	-13	HIS	-	expression tag	UNP P9WGI5
f	-12	HIS	-	expression tag	UNP P9WGI5
f	-11	HIS	-	expression tag	UNP P9WGI5
f	-10	HIS	-	expression tag	UNP P9WGI5
f	-9	SER	-	expression tag	UNP P9WGI5
f	-8	SER	-	expression tag	UNP P9WGI5
f	-7	GLY	-	expression tag	UNP P9WGI5
f	-6	LEU	-	expression tag	UNP P9WGI5
f	-5	VAL	-	expression tag	UNP P9WGI5
f	-4	PRO	-	expression tag	UNP P9WGI5
f	-3	ARG	-	expression tag	UNP P9WGI5
f	-2	GLY	-	expression tag	UNP P9WGI5
f	-1	SER	-	expression tag	UNP P9WGI5
f	0	HIS	-	expression tag	UNP P9WGI5

- Molecule 6 is ZINC ION (three-letter code: ZN) (formula: Zn) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms	AltConf
6	D	2	Total Zn 2 2	0
6	d	2	Total Zn 2 2	0

- Molecule 7 is MAGNESIUM ION (three-letter code: MG) (formula: Mg) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms		AltConf
7	D	1	Total 1	Mg 1	0
7	d	1	Total 1	Mg 1	0









## 4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	115112	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 ( $e^-/\text{\AA}^2$ )	49.6	Depositor
Minimum defocus (nm)	500	Depositor
Maximum defocus (nm)	7000	Depositor
Magnification	Not provided	
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor
Maximum map value	1.279	Depositor
Minimum map value	-0.340	Depositor
Average map value	-0.002	Depositor
Map value standard deviation	0.038	Depositor
Recommended contour level	0.2	Depositor
Map size (Å)	495.0, 495.0, 495.0	wwPDB
Map dimensions	450, 450, 450	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: 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	0.24	0/1726	0.51	0/2348
1	B	0.25	0/1775	0.52	0/2414
1	a	0.24	0/1726	0.51	0/2348
1	b	0.25	0/1775	0.52	0/2414
2	C	0.26	0/8799	0.51	0/11929
2	c	0.26	0/8799	0.51	0/11929
3	D	0.28	0/10118	0.54	0/13675
3	d	0.28	0/10118	0.53	0/13675
4	E	0.24	0/667	0.44	0/908
4	e	0.24	0/667	0.44	0/908
5	F	0.43	0/1134	0.73	0/1525
5	f	0.43	0/1134	0.74	0/1525
All	All	0.27	0/48438	0.53	0/65598

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](#)

Due to software issues we are unable to calculate clashes - this section is therefore empty.

## 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	221/347 (64%)	212 (96%)	9 (4%)	0	100	100
1	B	228/347 (66%)	222 (97%)	6 (3%)	0	100	100
1	a	221/347 (64%)	212 (96%)	9 (4%)	0	100	100
1	b	228/347 (66%)	222 (97%)	6 (3%)	0	100	100
2	C	1110/1174 (94%)	1075 (97%)	35 (3%)	0	100	100
2	c	1110/1174 (94%)	1075 (97%)	35 (3%)	0	100	100
3	D	1265/1319 (96%)	1217 (96%)	48 (4%)	0	100	100
3	d	1265/1319 (96%)	1216 (96%)	49 (4%)	0	100	100
4	E	81/110 (74%)	75 (93%)	6 (7%)	0	100	100
4	e	81/110 (74%)	75 (93%)	6 (7%)	0	100	100
5	F	140/343 (41%)	135 (96%)	5 (4%)	0	100	100
5	f	140/343 (41%)	136 (97%)	4 (3%)	0	100	100
All	All	6090/7280 (84%)	5872 (96%)	218 (4%)	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	192/297 (65%)	191 (100%)	1 (0%)	88	93
1	B	197/297 (66%)	195 (99%)	2 (1%)	76	86

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	a	192/297 (65%)	191 (100%)	1 (0%)	88	93
1	b	197/297 (66%)	195 (99%)	2 (1%)	76	86
2	C	946/995 (95%)	944 (100%)	2 (0%)	93	96
2	c	946/995 (95%)	944 (100%)	2 (0%)	93	96
3	D	1058/1098 (96%)	1049 (99%)	9 (1%)	78	88
3	d	1058/1098 (96%)	1048 (99%)	10 (1%)	78	88
4	E	70/90 (78%)	70 (100%)	0	100	100
4	e	70/90 (78%)	70 (100%)	0	100	100
5	F	111/284 (39%)	103 (93%)	8 (7%)	14	41
5	f	111/284 (39%)	106 (96%)	5 (4%)	27	54
All	All	5148/6122 (84%)	5106 (99%)	42 (1%)	82	89

All (42) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	15	THR
1	B	121	PRO
1	B	148	PRO
2	C	318	LYS
2	C	581	VAL
3	D	60	CYS
3	D	126	GLU
3	D	225	THR
3	D	256	MET
3	D	335	PHE
3	D	537	ASP
3	D	717	LYS
3	D	776	GLU
3	D	963	ARG
5	F	28	ARG
5	F	30	TYR
5	F	62	LEU
5	F	66	LYS
5	F	67	ARG
5	F	73	LYS
5	F	92	GLU
5	F	114	ASP
1	a	15	THR

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	b	121	PRO
1	b	148	PRO
2	c	318	LYS
2	c	581	VAL
3	d	60	CYS
3	d	67	ARG
3	d	225	THR
3	d	256	MET
3	d	259	GLU
3	d	335	PHE
3	d	537	ASP
3	d	717	LYS
3	d	776	GLU
3	d	963	ARG
5	f	30	TYR
5	f	67	ARG
5	f	73	LYS
5	f	92	GLU
5	f	114	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
3	d	369	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

Of 6 ligands modelled in this entry, 6 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.



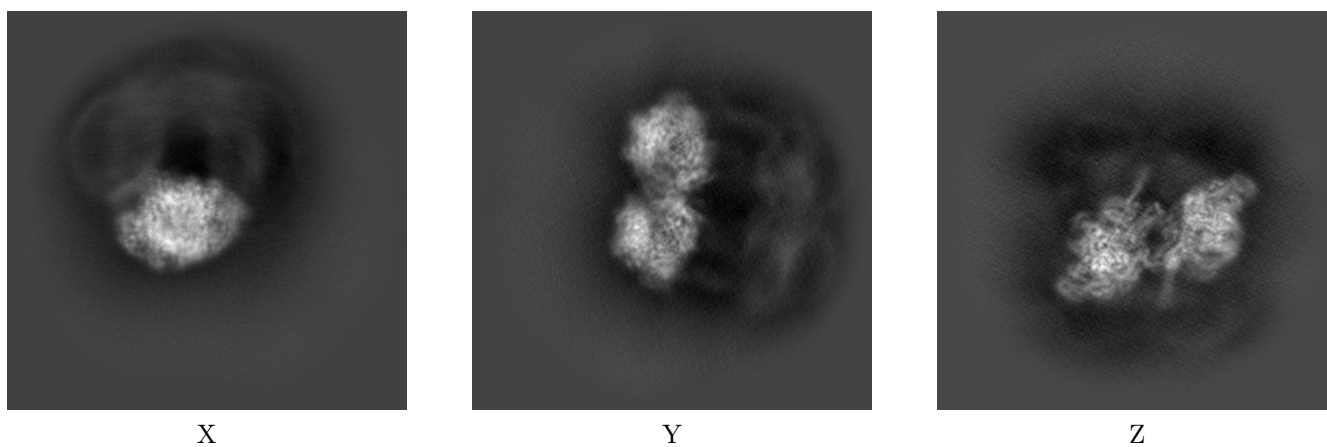
## 6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-13829. 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](#)

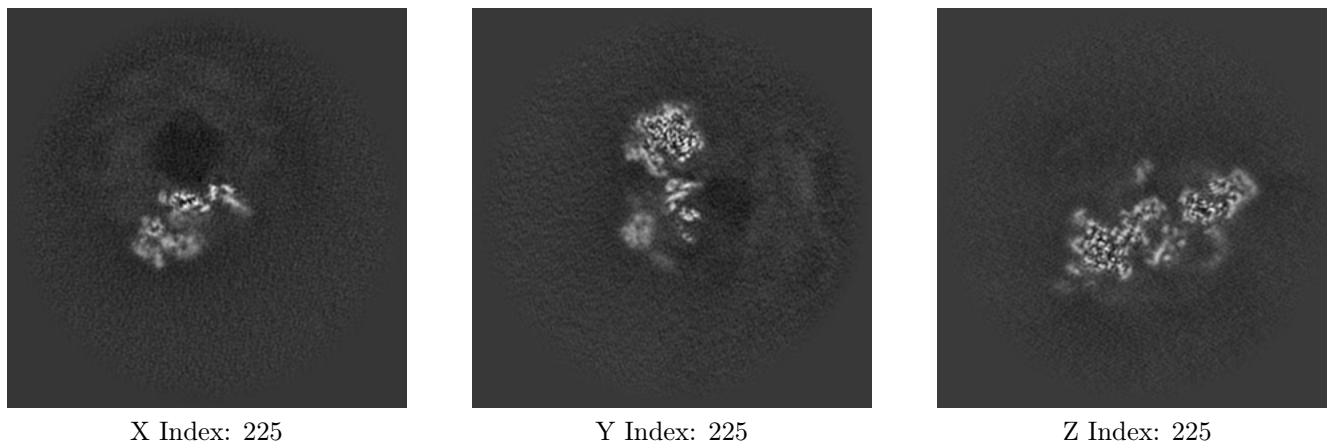
#### 6.1.1 Primary map



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

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

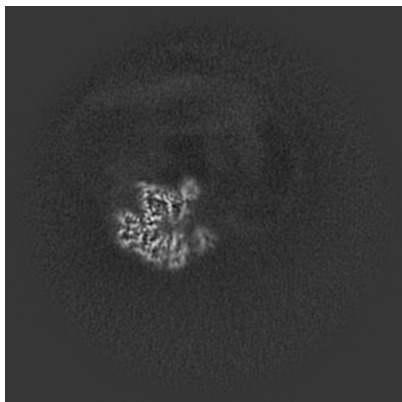
#### 6.2.1 Primary map



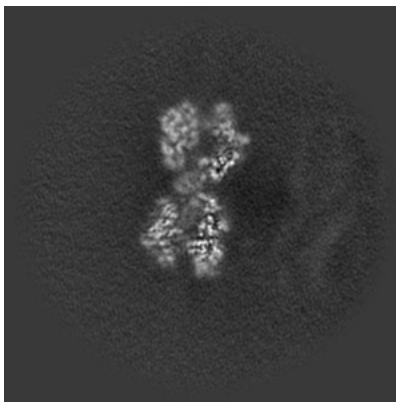
The images above show central slices of the map in three orthogonal directions.

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

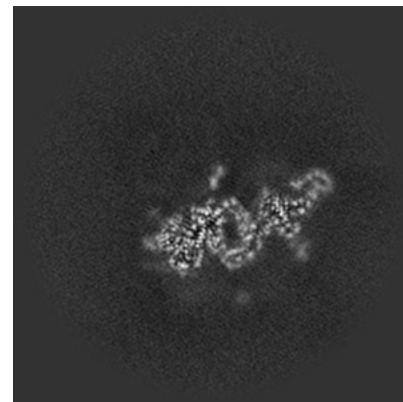
### 6.3.1 Primary map



X Index: 184



Y Index: 186



Z Index: 233

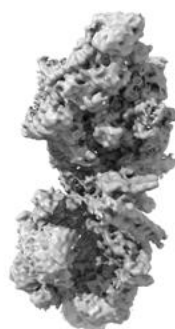
The images above show the largest variance slices of the map in three orthogonal directions.

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

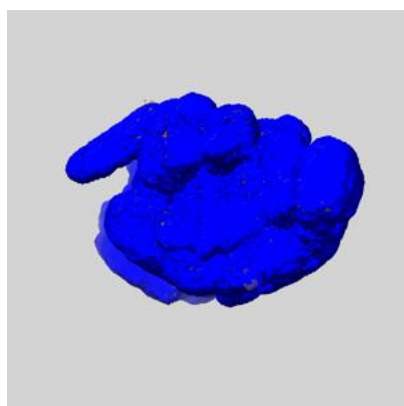
## 6.5 Mask visualisation [i](#)

This section shows the 3D surface view of the primary map at 50% transparency overlaid with the specified mask at 0% transparency

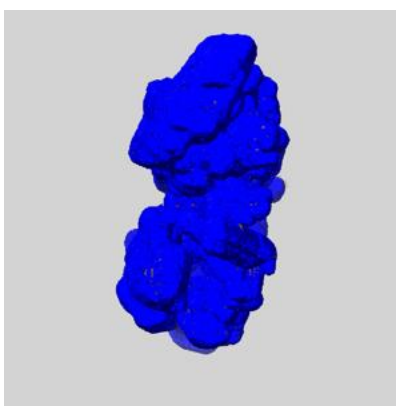
A mask typically either:

- Encompasses the whole structure
- Separates out a domain, a functional unit, a monomer or an area of interest from a larger structure

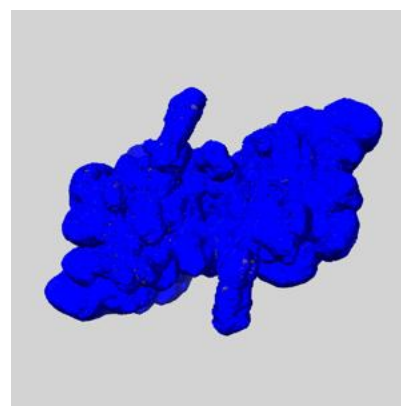
### 6.5.1 emd\_13829\_msk\_1.map [i](#)



X



Y

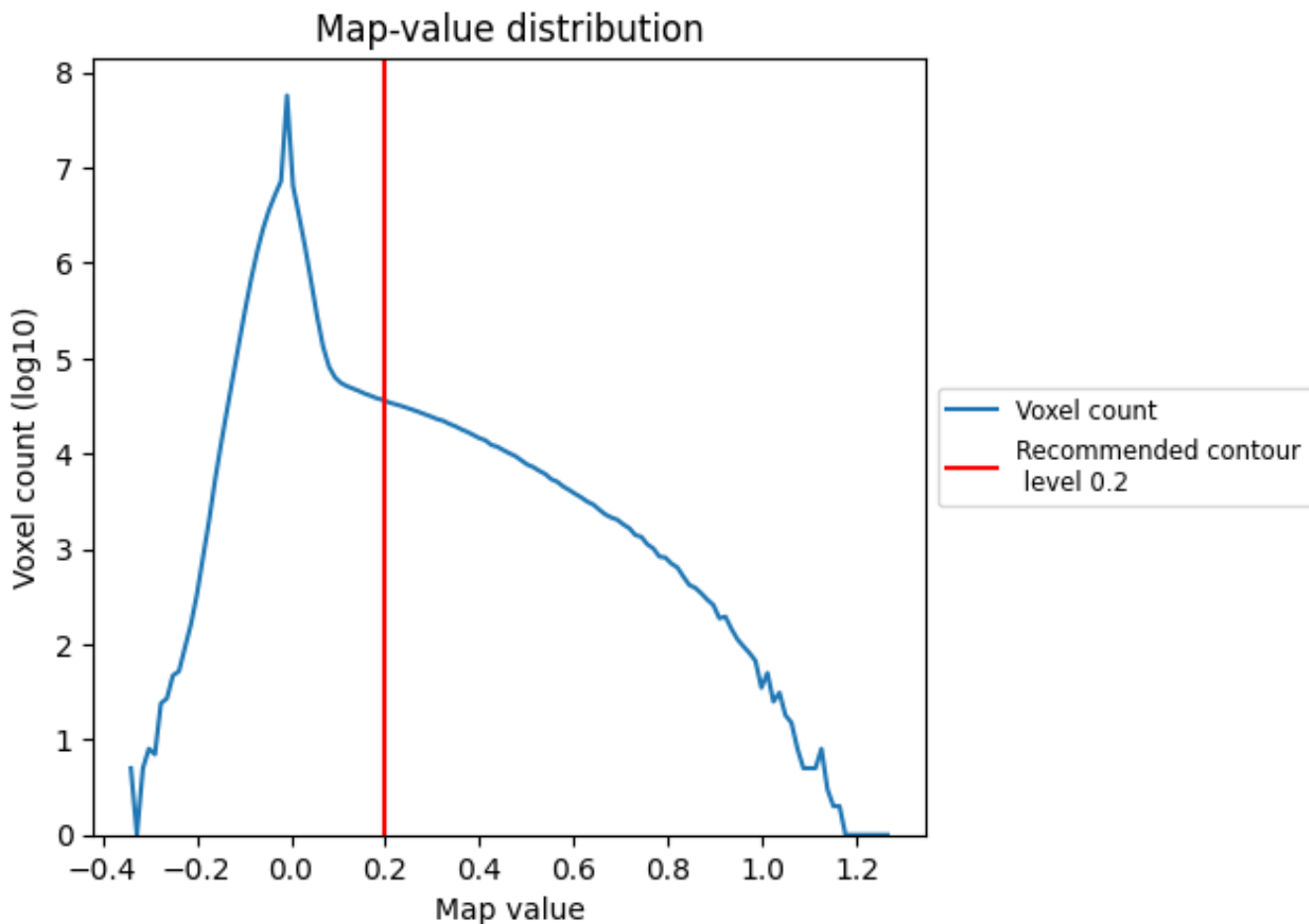


Z

## 7 Map analysis [i](#)

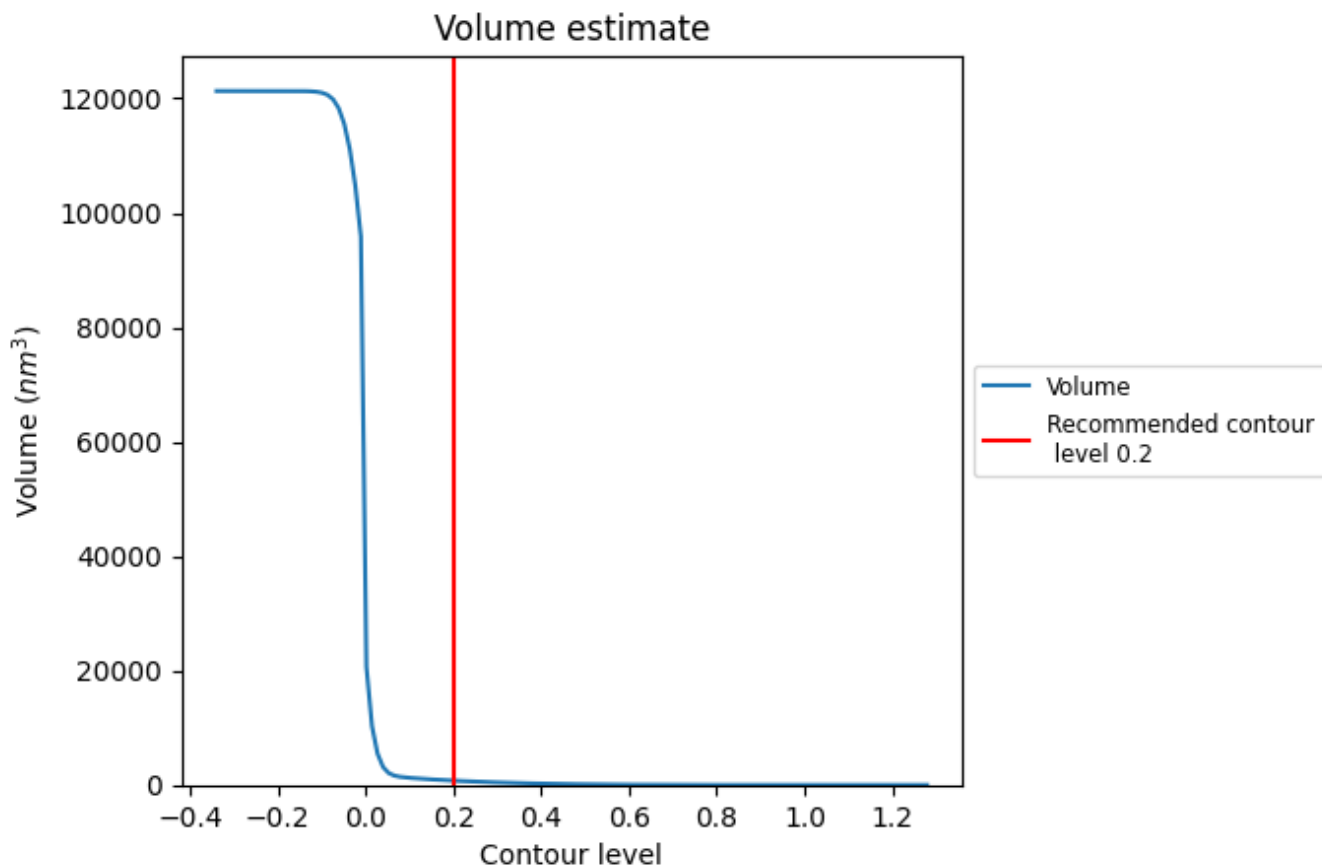
This section contains the results of statistical analysis of the map.

### 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.

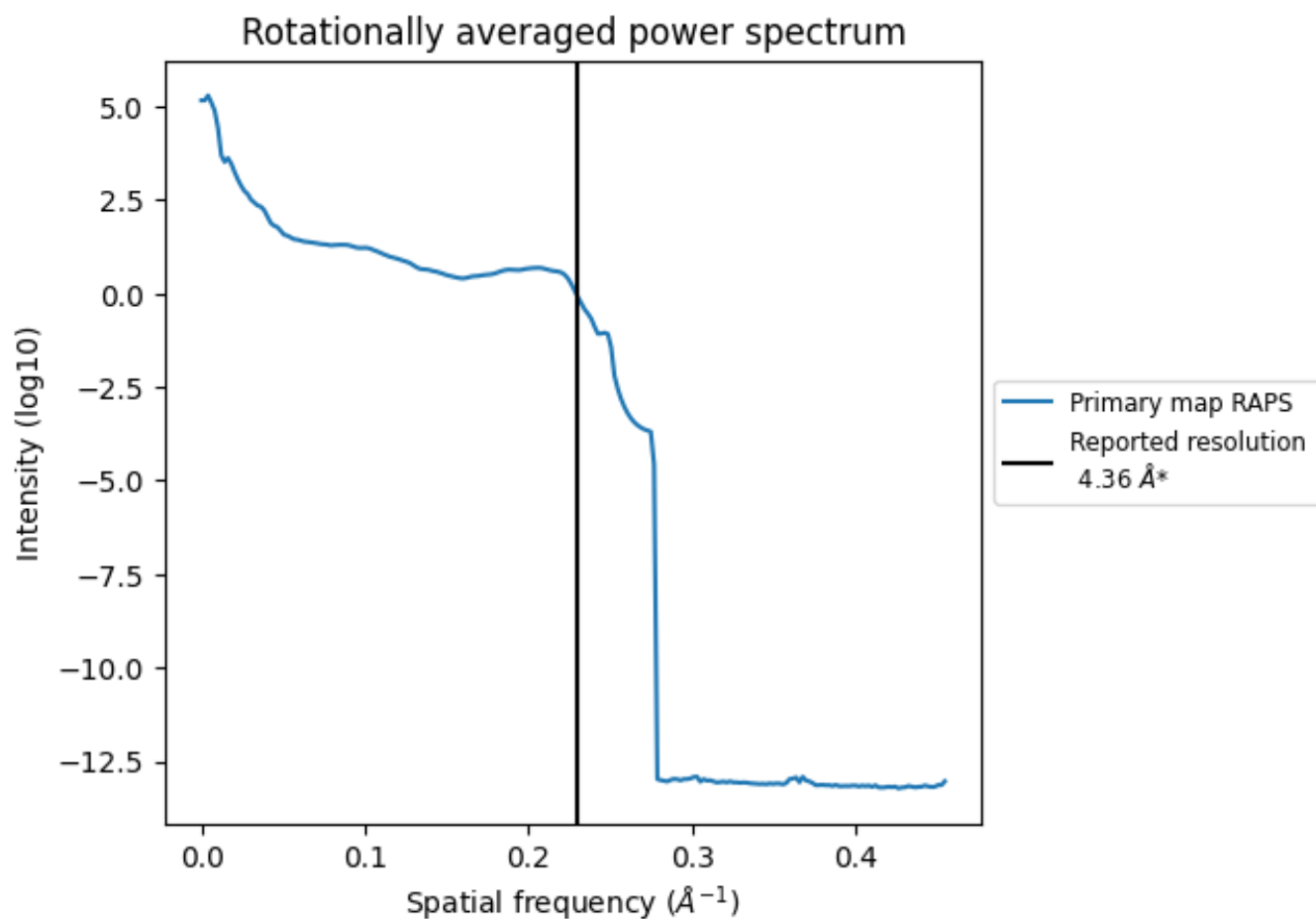
## 7.2 Volume estimate [i](#)



The volume at the recommended contour level is 759  $\text{nm}^3$ ; this corresponds to an approximate mass of 685 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](#)

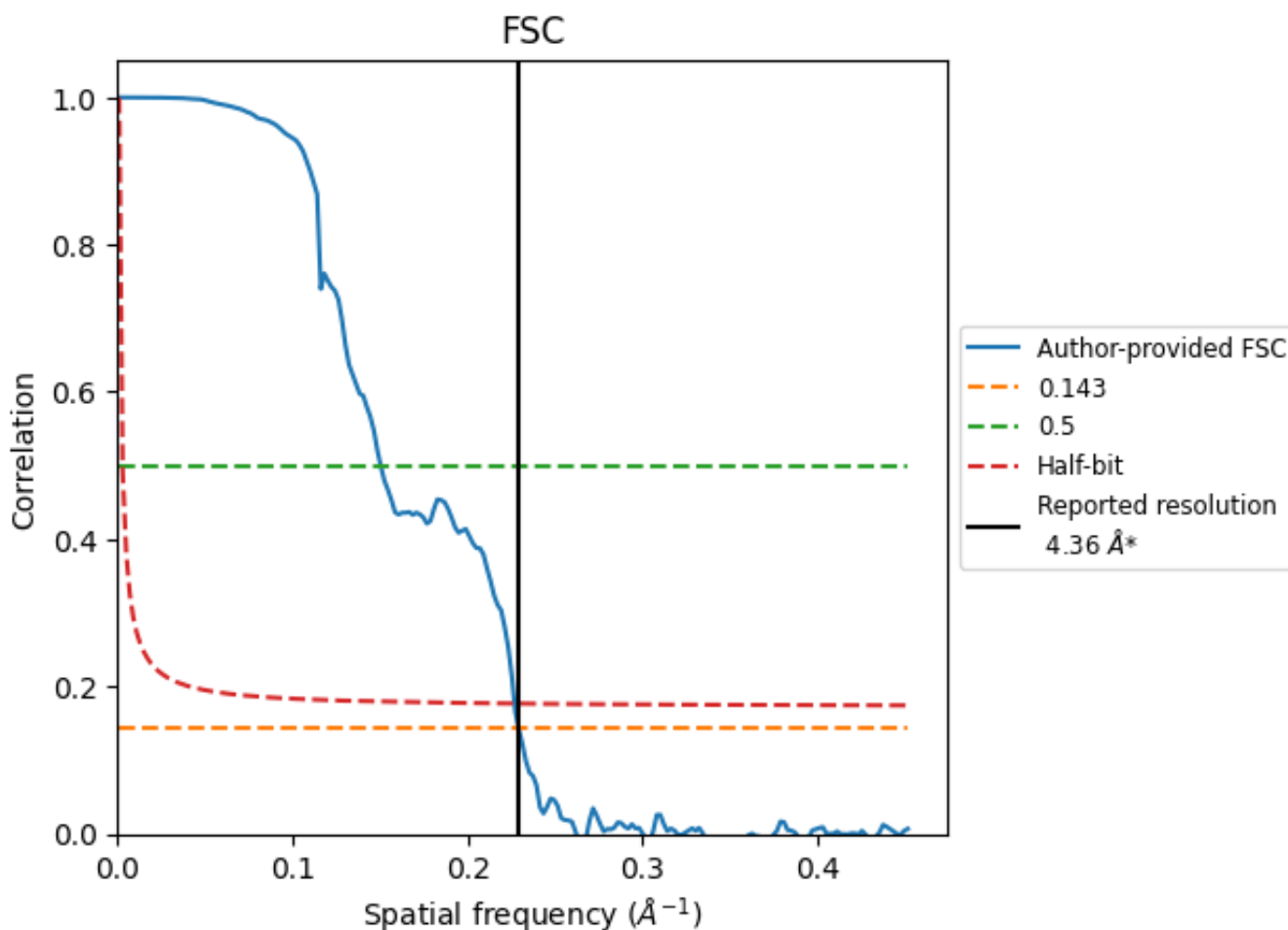


\*Reported resolution corresponds to spatial frequency of 0.229 Å<sup>-1</sup>

## 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 0.229 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	4.36	-	-
Author-provided FSC curve	4.36	6.65	4.41
Unmasked-calculated*	-	-	-

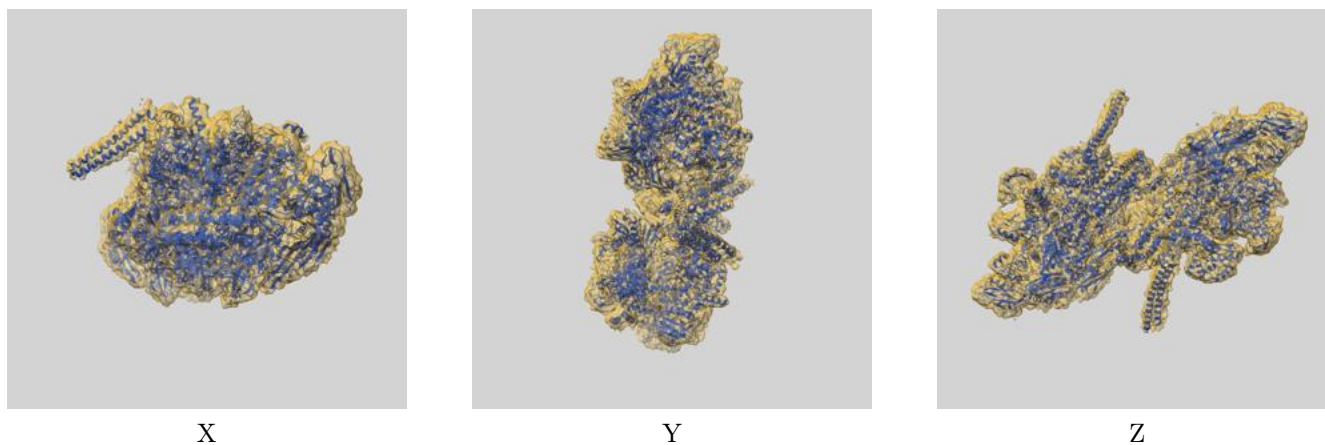
\*Resolution estimate based on FSC curve calculated by comparison of deposited half-maps.



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

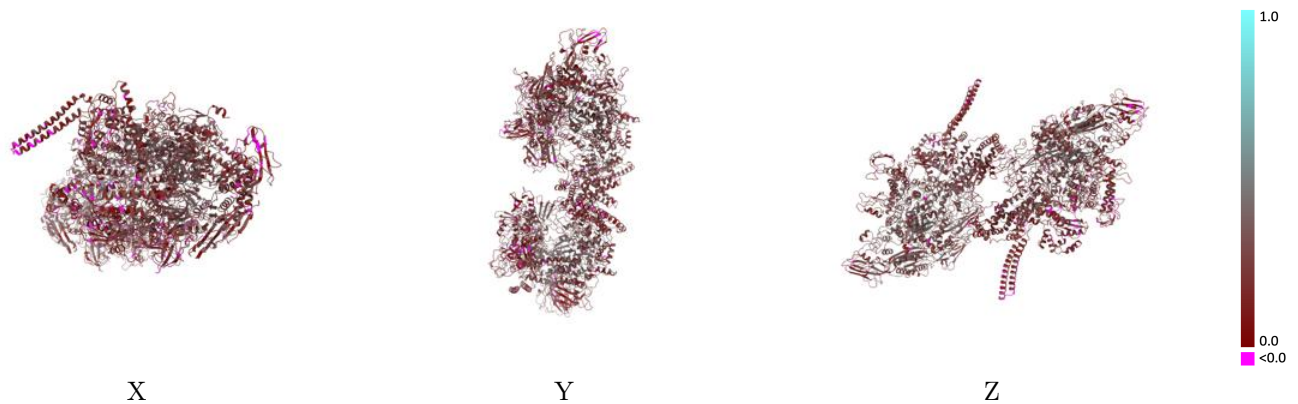
This section contains information regarding the fit between EMDB map EMD-13829 and PDB model 7Q59. Per-residue inclusion information can be found in section 3 on page 8.

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



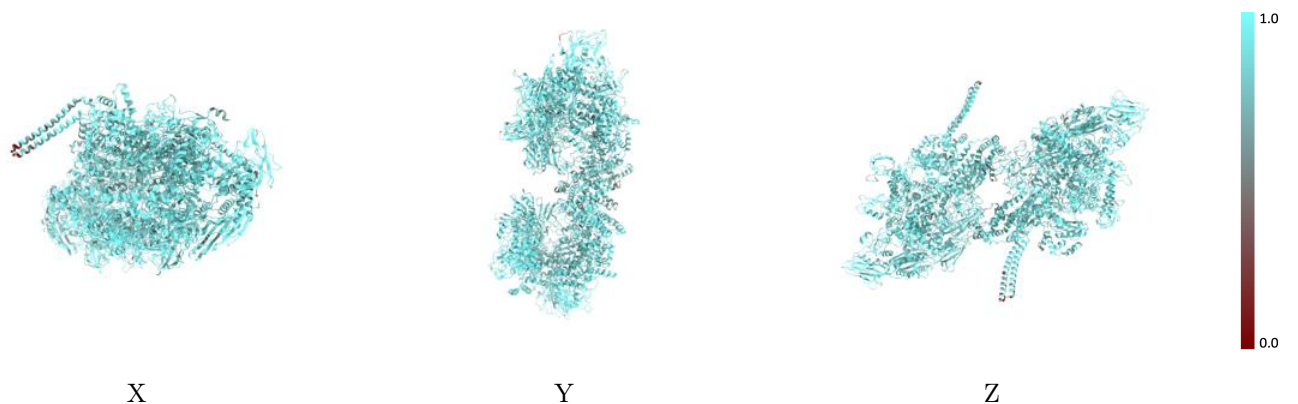
The images above show the 3D surface view of the map at the recommended contour level 0.2 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

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



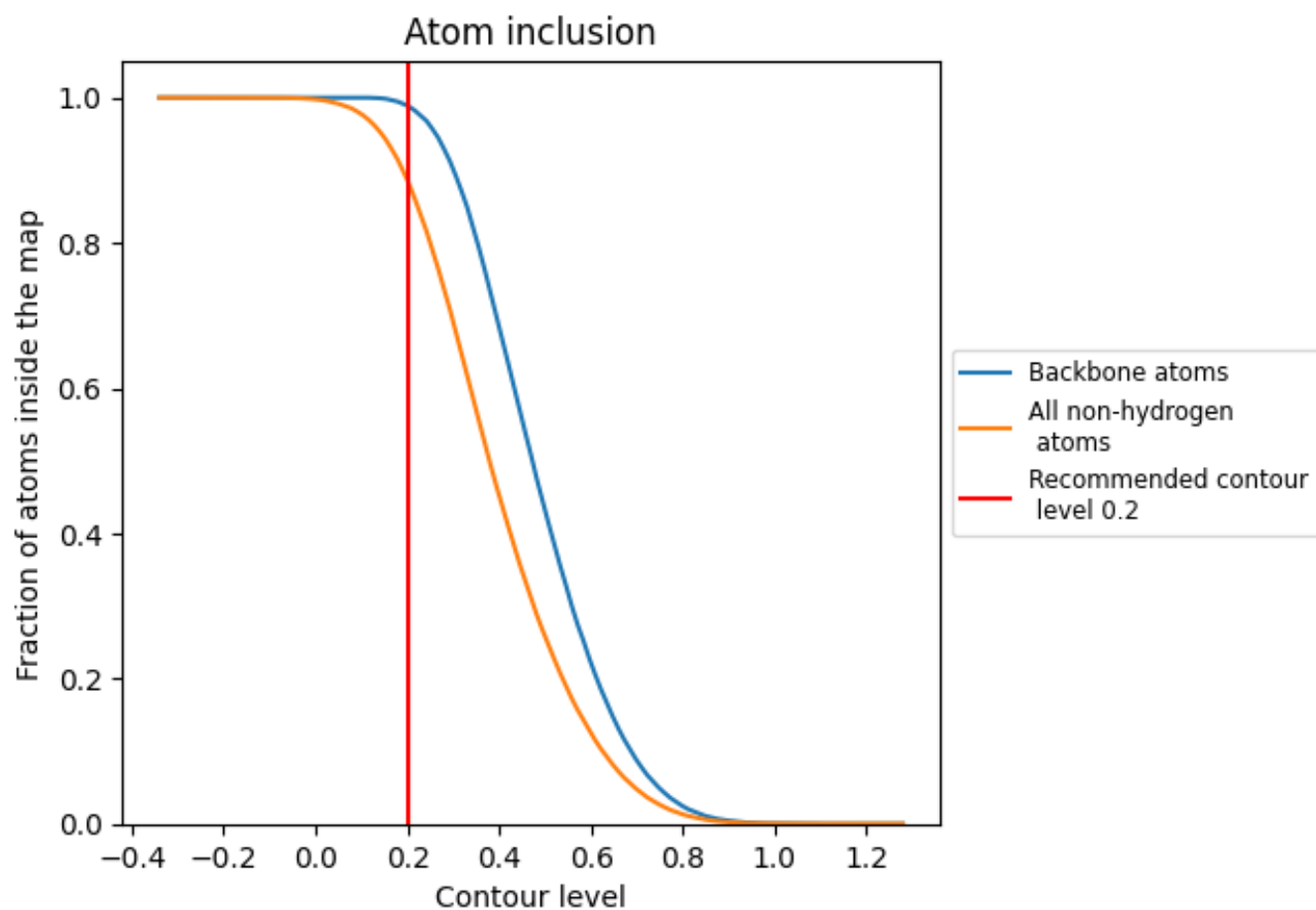
The images above show the model with each residue coloured according to its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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 (0.2).



























## 9.4 Atom inclusion [i](#)



At the recommended contour level, 99% of all backbone atoms, 89% 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 (0.2) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	 0.8867	 0.2570
A	 0.8962	 0.2770
B	 0.9068	 0.2670
C	 0.9039	 0.2590
D	 0.8889	 0.2760
E	 0.9016	 0.3000
F	 0.8438	 0.2480
a	 0.8866	 0.2430
b	 0.8957	 0.2190
c	 0.8773	 0.2370
d	 0.8801	 0.2580
e	 0.8984	 0.2870
f	 0.8299	 0.2040

