



## wwPDB EM Validation Summary Report ⓘ

Oct 13, 2024 – 04:40 am BST

PDB ID : 8BMW  
EMDB ID : EMD-16126  
Title : SsoCsm  
Authors : Spagnolo, L.; White, M.F.  
Deposited on : 2022-11-11  
Resolution : 3.50 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.dev113  
MolProbity : 4.02b-467  
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)  
MapQ : 1.9.13  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.39

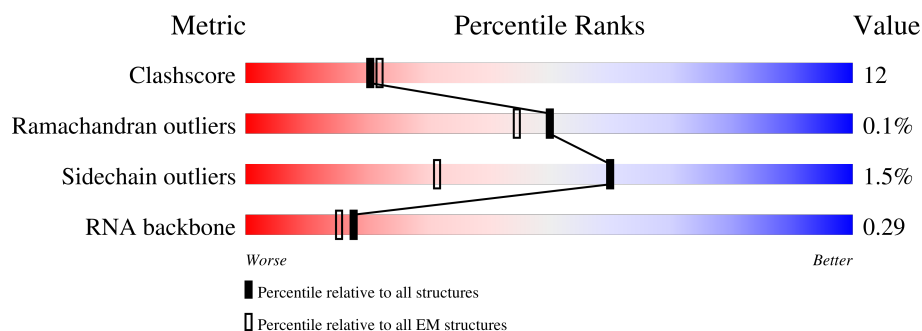
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 3.50 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	210492	15764
Ramachandran outliers	207382	16835
Sidechain outliers	206894	16415
RNA backbone	6643	2191

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	139	
1	B	139	
1	C	139	
1	D	139	
1	E	139	
2	J	272	
3	L	202	

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
4	N	252	
5	R	48	
6	F	248	
6	G	248	
7	H	278	
7	I	278	
8	K	829	
9	M	300	

## 2 Entry composition

There are 9 unique types of molecules in this entry. The entry contains 28658 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 CRISPR-associated small subunit protein (Type III-D).

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	121	Total	C	N	O	S	0	0
			964	617	164	182	1		
1	B	119	Total	C	N	O	S	0	0
			948	606	161	180	1		
1	C	120	Total	C	N	O	S	0	0
			957	612	163	181	1		
1	D	121	Total	C	N	O	S	0	0
			964	617	164	182	1		
1	E	122	Total	C	N	O	S	0	0
			973	622	166	184	1		

- Molecule 2 is a protein called CRISPR-associated Cas7 paralog (Type III-D).

Mol	Chain	Residues	Atoms					AltConf	Trace
2	J	270	Total	C	N	O	S	0	0
			2121	1350	371	390	10		

- Molecule 3 is a protein called CRISPR-associated Cas7 paralog (Type III-D).

Mol	Chain	Residues	Atoms					AltConf	Trace
3	L	201	Total	C	N	O	S	0	0
			1585	1000	268	307	10		

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
L	94	ASP	GLU	conflict	UNP A0A157T2I3

- Molecule 4 is a protein called CRISPR-associated Cas7 paralog (Type III-D).

Mol	Chain	Residues	Atoms					AltConf	Trace
4	N	251	Total	C	N	O	S	0	0
			2029	1283	359	383	4		

- Molecule 5 is a RNA chain called RNA (48-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
5	R	48	Total	C	N	O	P	0	0
			971	438	114	372	47		

- Molecule 6 is a protein called CRISPR-associated Cas7 paralog (Type III-D).

Mol	Chain	Residues	Atoms					AltConf	Trace
6	G	238	Total	C	N	O	S	0	0
			1877	1209	309	348	11		
6	F	239	Total	C	N	O	S	0	0
			1876	1204	309	351	12		

- Molecule 7 is a protein called CRISPR-associated Cas7 paralog (Type III-D).

Mol	Chain	Residues	Atoms					AltConf	Trace
7	H	275	Total	C	N	O	S	0	0
			2168	1381	374	401	12		
7	I	275	Total	C	N	O	S	0	0
			2164	1381	374	401	8		

- Molecule 8 is a protein called CRISPR-associated protein Cas10 (Type III-D).

Mol	Chain	Residues	Atoms					AltConf	Trace
8	K	829	Total	C	N	O	S	0	0
			6678	4277	1101	1274	26		

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
K	421	ASP	ASN	conflict	UNP A0A157T112

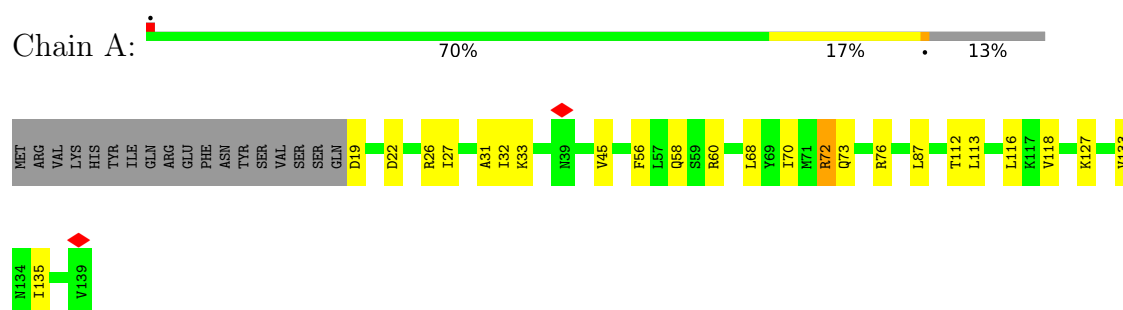
- Molecule 9 is a protein called CRISPR-associated protein Cas5 (Type III-D).

Mol	Chain	Residues	Atoms					AltConf	Trace
9	M	294	Total	C	N	O	S	0	0
			2383	1541	389	444	9		

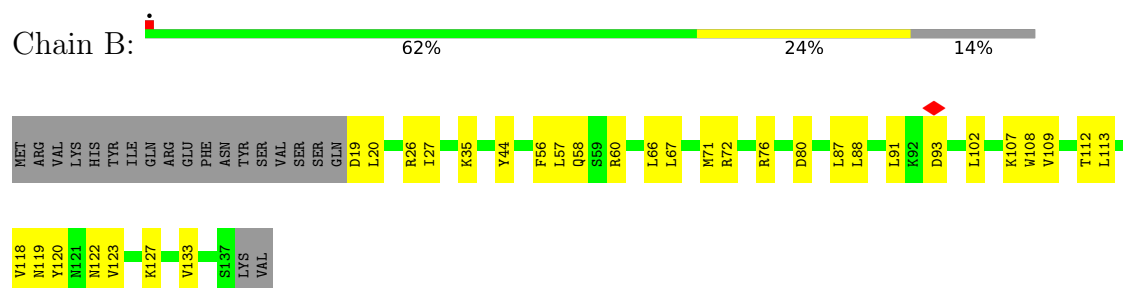
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and atom inclusion in map density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red diamond above a residue indicates a poor fit to the EM map for this residue (all-atom inclusion < 40%). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

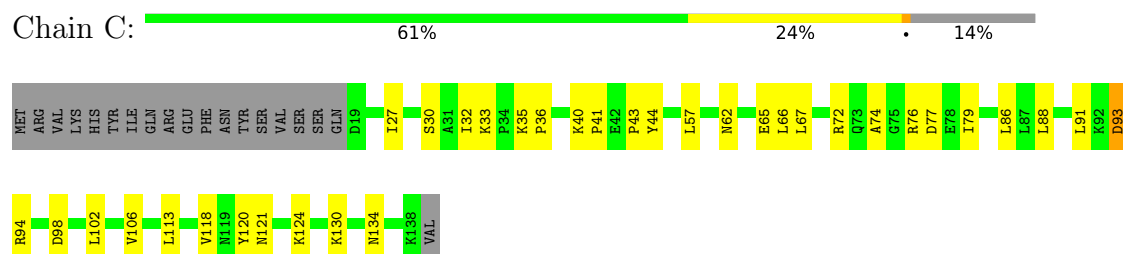
- Molecule 1: CRISPR-associated small subunit protein (Type III-D)



- Molecule 1: CRISPR-associated small subunit protein (Type III-D)



- Molecule 1: CRISPR-associated small subunit protein (Type III-D)



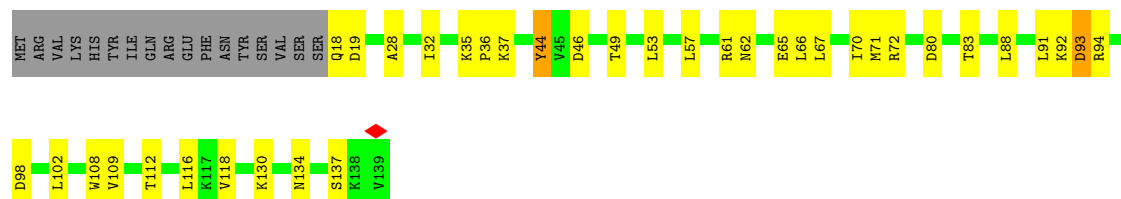
- Molecule 1: CRISPR-associated small subunit protein (Type III-D)





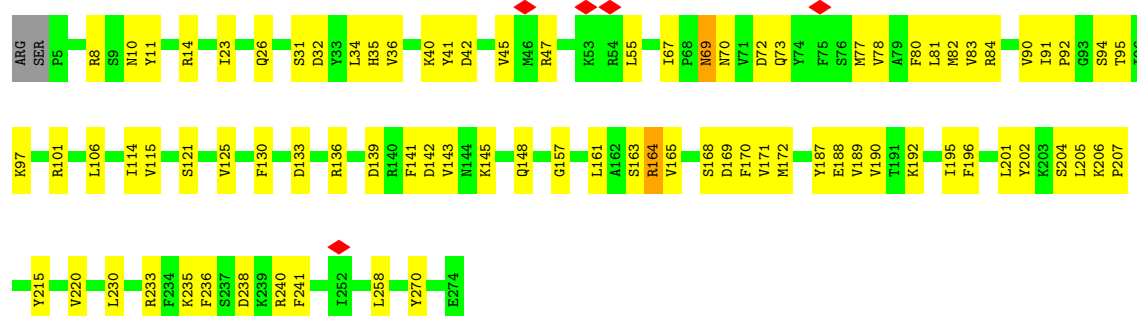
- Molecule 1: CRISPR-associated small subunit protein (Type III-D)

Chain E: 61% 25% 12%



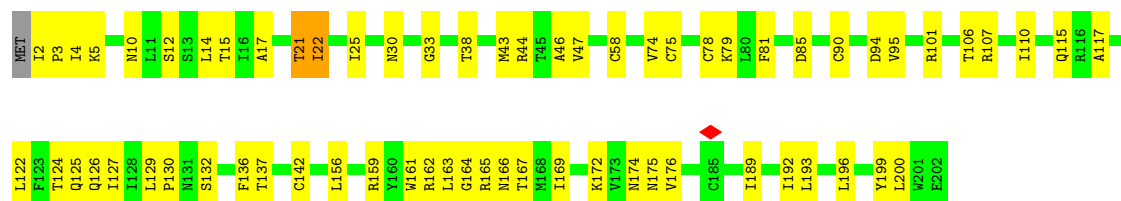
- Molecule 2: CRISPR-associated Cas7 paralog (Type III-D)

Chain J: 68% 30% 2%



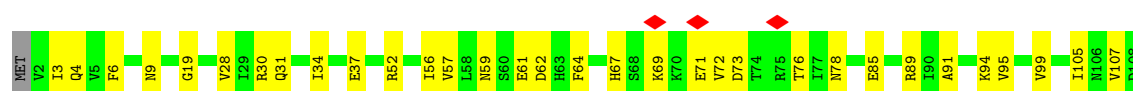
- Molecule 3: CRISPR-associated Cas7 paralog (Type III-D)

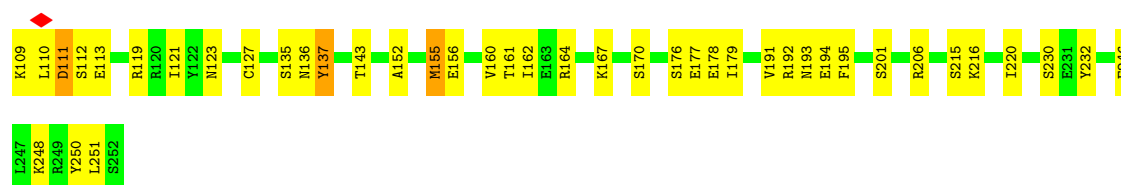
Chain L: 67% 32% 1%



- Molecule 4: CRISPR-associated Cas7 paralog (Type III-D)

Chain N: 70% 28% 2%

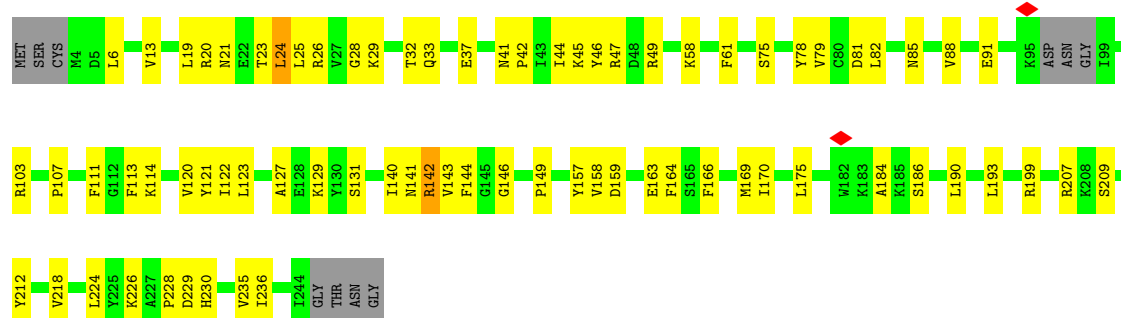




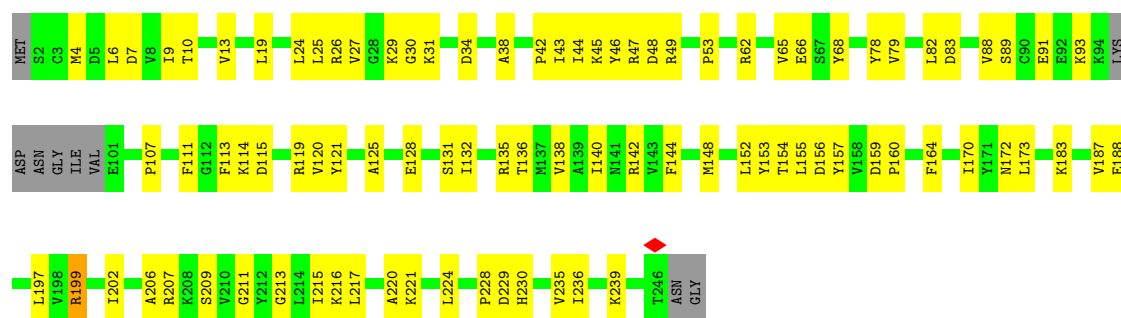
- Molecule 5: RNA (48-MER)



- Molecule 6: CRISPR-associated Cas7 paralog (Type III-D)



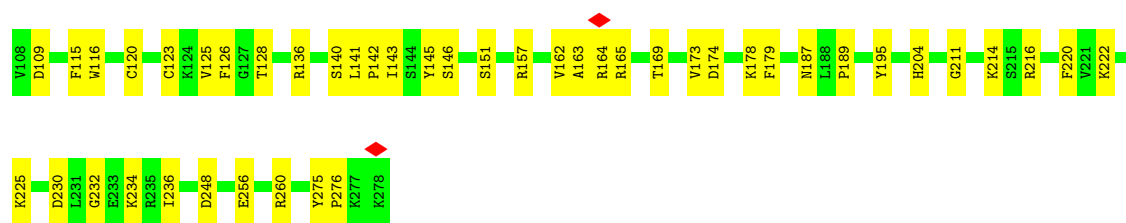
- Molecule 6: CRISPR-associated Cas7 paralog (Type III-D)



- Molecule 7: CRISPR-associated Cas7 paralog (Type III-D)

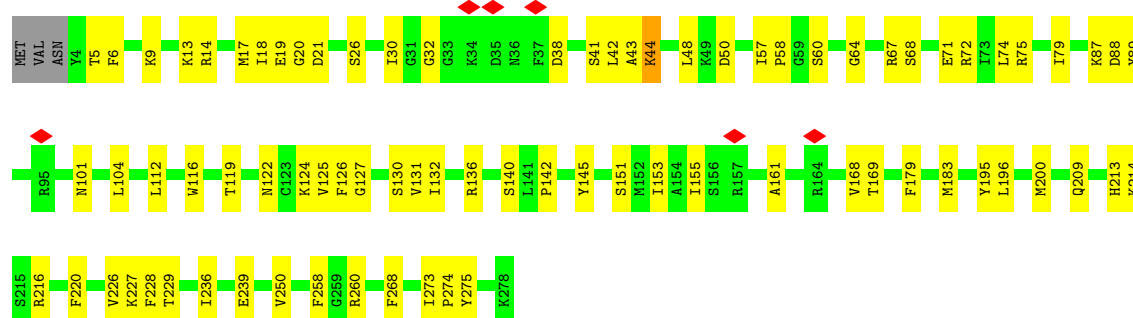






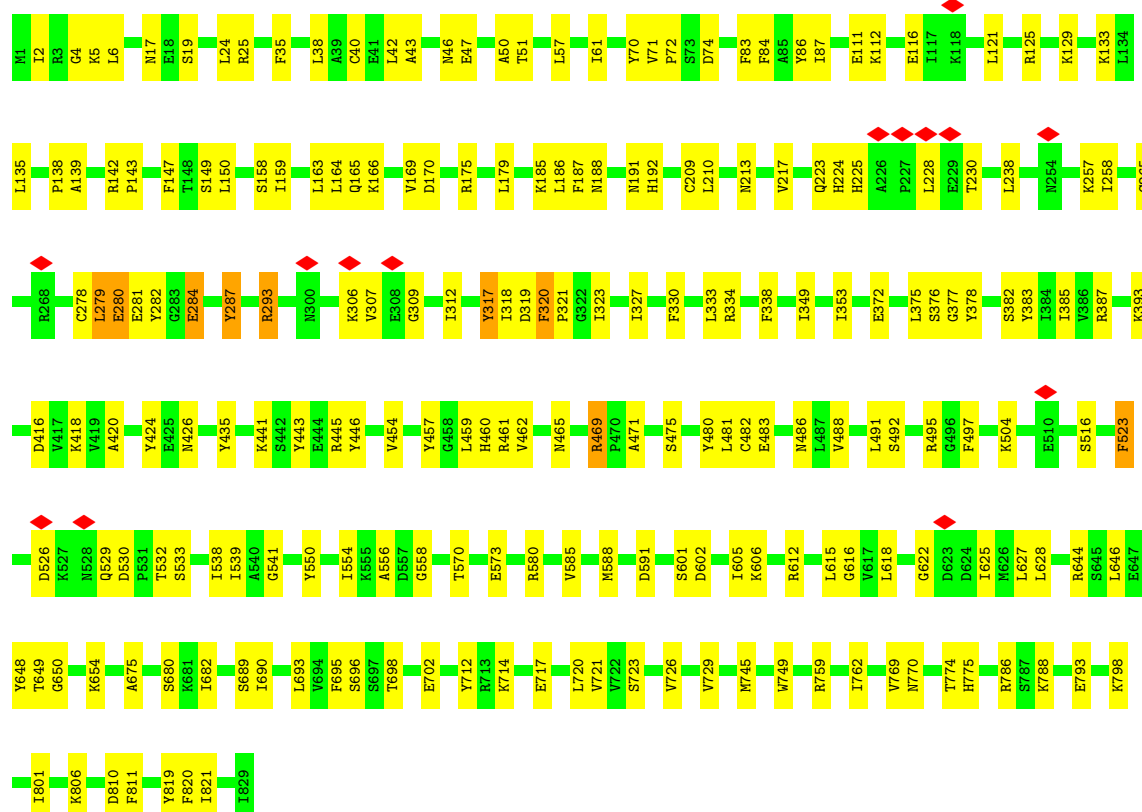
• Molecule 7: CRISPR-associated Cas7 paralog (Type III-D)

Chain I: 70% 28%

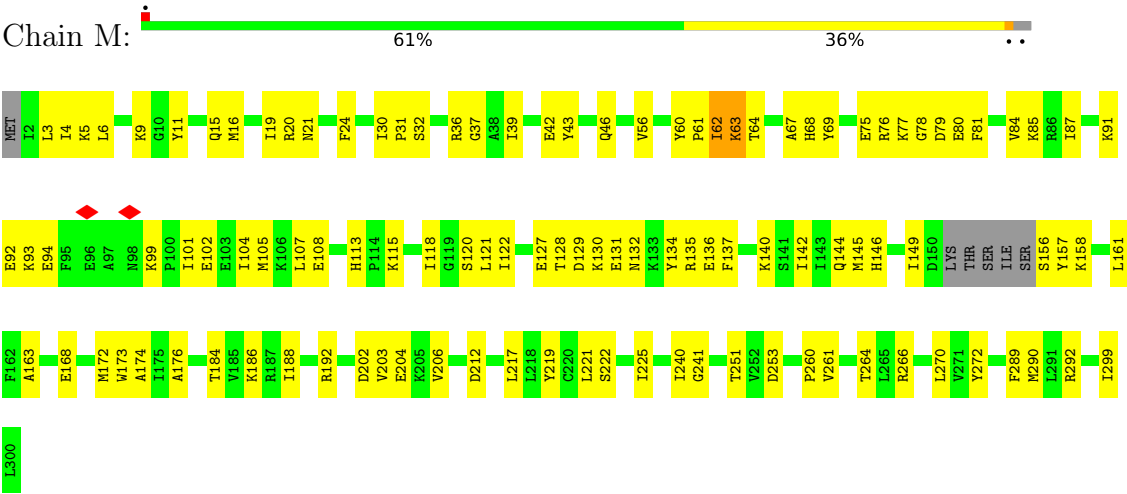


• Molecule 8: CRISPR-associated protein Cas10 (Type III-D)

Chain K: 74% 24%



● Molecule 9: CRISPR-associated protein Cas5 (Type III-D)



## 4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	192787	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 ( $e^-/\text{\AA}^2$ )	34	Depositor
Minimum defocus (nm)	0	Depositor
Maximum defocus (nm)	0	Depositor
Magnification	Not provided	
Image detector	GATAN K2 BASE (4k x 4k)	Depositor
Maximum map value	0.069	Depositor
Minimum map value	-0.004	Depositor
Average map value	0.000	Depositor
Map value standard deviation	0.002	Depositor
Recommended contour level	0.008	Depositor
Map size (Å)	313.8, 313.8, 313.8	wwPDB
Map dimensions	300, 300, 300	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.046, 1.046, 1.046	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	A	0.24	0/976	0.42	0/1318
1	B	0.24	0/960	0.42	0/1297
1	C	0.25	0/969	0.42	0/1308
1	D	0.24	0/976	0.42	0/1318
1	E	0.25	0/985	0.43	0/1330
2	J	0.29	0/2161	0.51	0/2909
3	L	0.29	0/1613	0.50	0/2188
4	N	0.27	0/2060	0.50	0/2763
5	R	0.27	0/1072	0.78	0/1658
6	F	0.27	0/1907	0.46	0/2568
6	G	0.27	0/1908	0.46	0/2569
7	H	0.32	0/2209	0.51	0/2966
7	I	0.31	0/2205	0.51	0/2962
8	K	0.27	0/6803	0.45	0/9173
9	M	0.30	0/2433	0.51	0/3266
All	All	0.28	0/29237	0.49	0/39593

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	964	0	1019	19	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	B	948	0	997	21	0
1	C	957	0	1010	23	0
1	D	964	0	1019	21	0
1	E	973	0	1027	33	0
2	J	2121	0	2149	66	0
3	L	1585	0	1591	57	0
4	N	2029	0	2059	57	0
5	R	971	0	488	41	0
6	F	1876	0	1921	66	0
6	G	1877	0	1936	50	0
7	H	2168	0	2180	52	0
7	I	2164	0	2172	58	0
8	K	6678	0	6719	135	0
9	M	2383	0	2416	69	0
All	All	28658	0	28703	680	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 12.

The worst 5 of 680 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)
3:L:22:ILE:HD11	8:K:720:LEU:HD22	1.54	0.88
3:L:58:CYS:N	3:L:75:CYS:SG	2.47	0.86
6:F:26:ARG:HG2	6:F:206:ALA:HB2	1.58	0.85
8:K:38:LEU:HD11	8:K:116:GLU:HG3	1.55	0.85
3:L:159:ARG:HD3	3:L:200:LEU:HB2	1.60	0.81

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	119/139 (86%)	113 (95%)	6 (5%)	0	100	100
1	B	117/139 (84%)	113 (97%)	4 (3%)	0	100	100
1	C	118/139 (85%)	115 (98%)	3 (2%)	0	100	100
1	D	119/139 (86%)	116 (98%)	3 (2%)	0	100	100
1	E	120/139 (86%)	115 (96%)	5 (4%)	0	100	100
2	J	268/272 (98%)	250 (93%)	16 (6%)	2 (1%)	19	53
3	L	199/202 (98%)	185 (93%)	14 (7%)	0	100	100
4	N	249/252 (99%)	230 (92%)	19 (8%)	0	100	100
6	F	235/248 (95%)	212 (90%)	23 (10%)	0	100	100
6	G	234/248 (94%)	217 (93%)	17 (7%)	0	100	100
7	H	273/278 (98%)	261 (96%)	12 (4%)	0	100	100
7	I	273/278 (98%)	252 (92%)	21 (8%)	0	100	100
8	K	827/829 (100%)	790 (96%)	37 (4%)	0	100	100
9	M	290/300 (97%)	258 (89%)	30 (10%)	2 (1%)	19	53
All	All	3441/3602 (96%)	3227 (94%)	210 (6%)	4 (0%)	50	79

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	J	45	VAL
9	M	20	ARG
9	M	93	LYS
2	J	69	ASN

### 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	A	110/128 (86%)	108 (98%)	2 (2%)	54	74
1	B	108/128 (84%)	107 (99%)	1 (1%)	75	86
1	C	109/128 (85%)	108 (99%)	1 (1%)	75	86

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	D	110/128 (86%)	109 (99%)	1 (1%)	75	86
1	E	111/128 (87%)	109 (98%)	2 (2%)	54	74
2	J	239/244 (98%)	236 (99%)	3 (1%)	65	81
3	L	185/186 (100%)	183 (99%)	2 (1%)	70	83
4	N	225/226 (100%)	222 (99%)	3 (1%)	65	81
6	F	209/216 (97%)	207 (99%)	2 (1%)	73	84
6	G	209/216 (97%)	206 (99%)	3 (1%)	62	79
7	H	236/239 (99%)	233 (99%)	3 (1%)	65	81
7	I	232/239 (97%)	227 (98%)	5 (2%)	47	70
8	K	745/745 (100%)	734 (98%)	11 (2%)	60	77
9	M	262/268 (98%)	254 (97%)	8 (3%)	35	63
All	All	3090/3219 (96%)	3043 (98%)	47 (2%)	60	77

5 of 47 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
8	K	166	LYS
8	K	320	PHE
8	K	279	LEU
8	K	287	TYR
8	K	469	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	E	18	GLN
1	E	134	ASN
3	L	125	GLN
9	M	15	GLN

### 5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
5	R	47/48 (97%)	15 (31%)	1 (2%)

5 of 15 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
5	R	3	U
5	R	5	A
5	R	6	A
5	R	7	A
5	R	8	G

All (1) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
5	R	4	G

## 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.



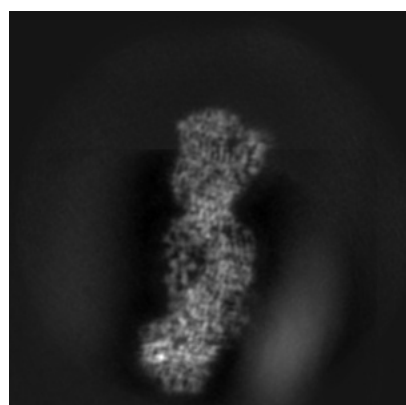
## 6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-16126. These allow visual inspection of the internal detail of the map and identification of artifacts.

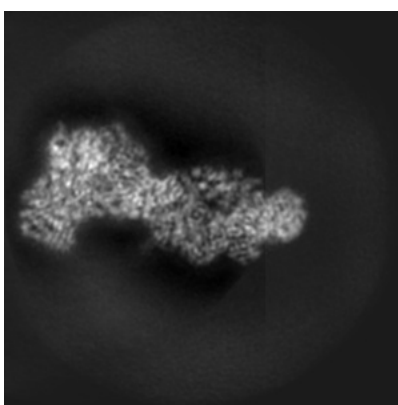
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

### 6.1 Orthogonal projections [i](#)

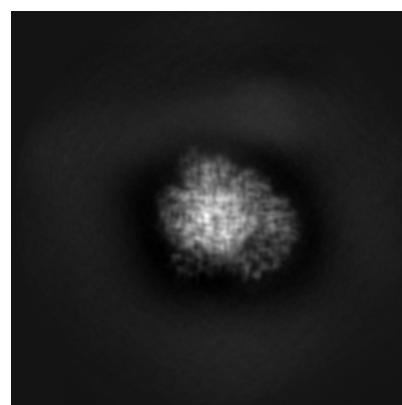
#### 6.1.1 Primary map



X



Y

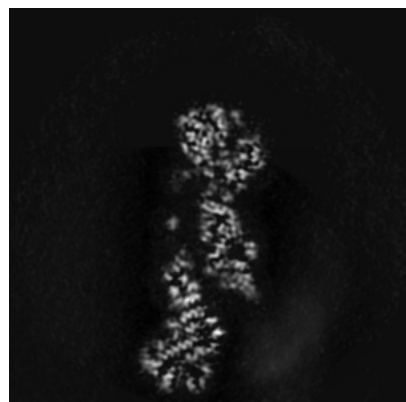


Z

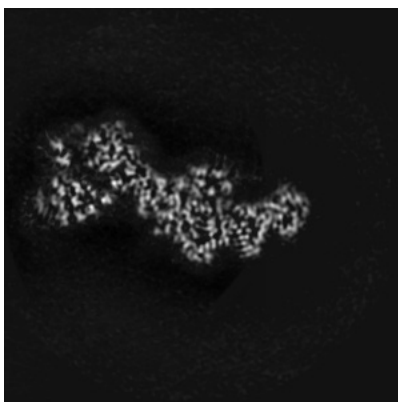
The images above show the map projected in three orthogonal directions.

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

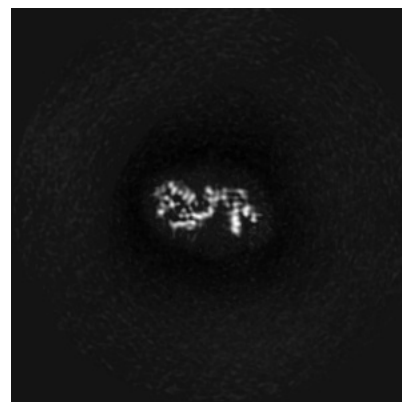
#### 6.2.1 Primary map



X Index: 150



Y Index: 150

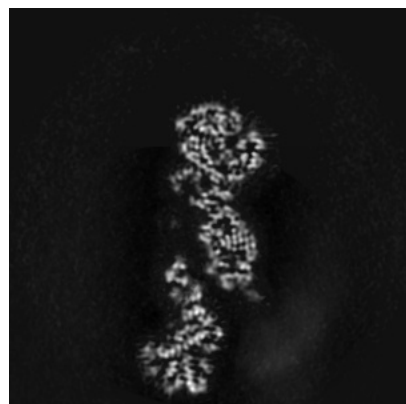


Z Index: 150

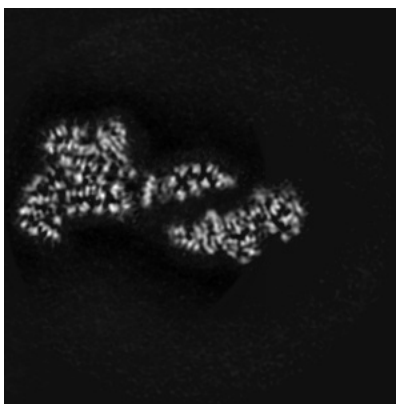
The images above show central slices of the map in three orthogonal directions.

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

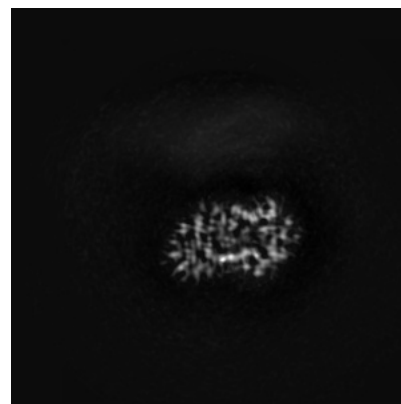
### 6.3.1 Primary map



X Index: 147



Y Index: 137



Z Index: 39

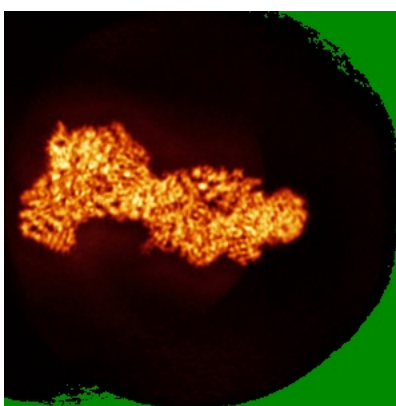
The images above show the largest variance slices of the map in three orthogonal directions.

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

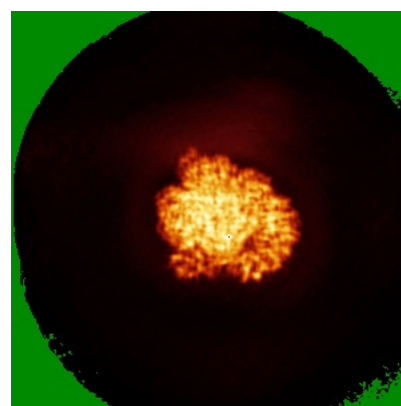
### 6.4.1 Primary map



X



Y

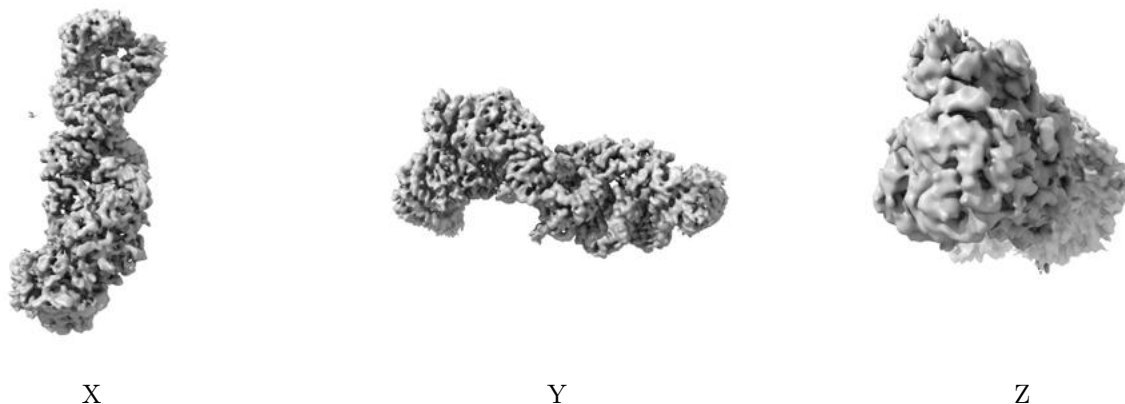


Z

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

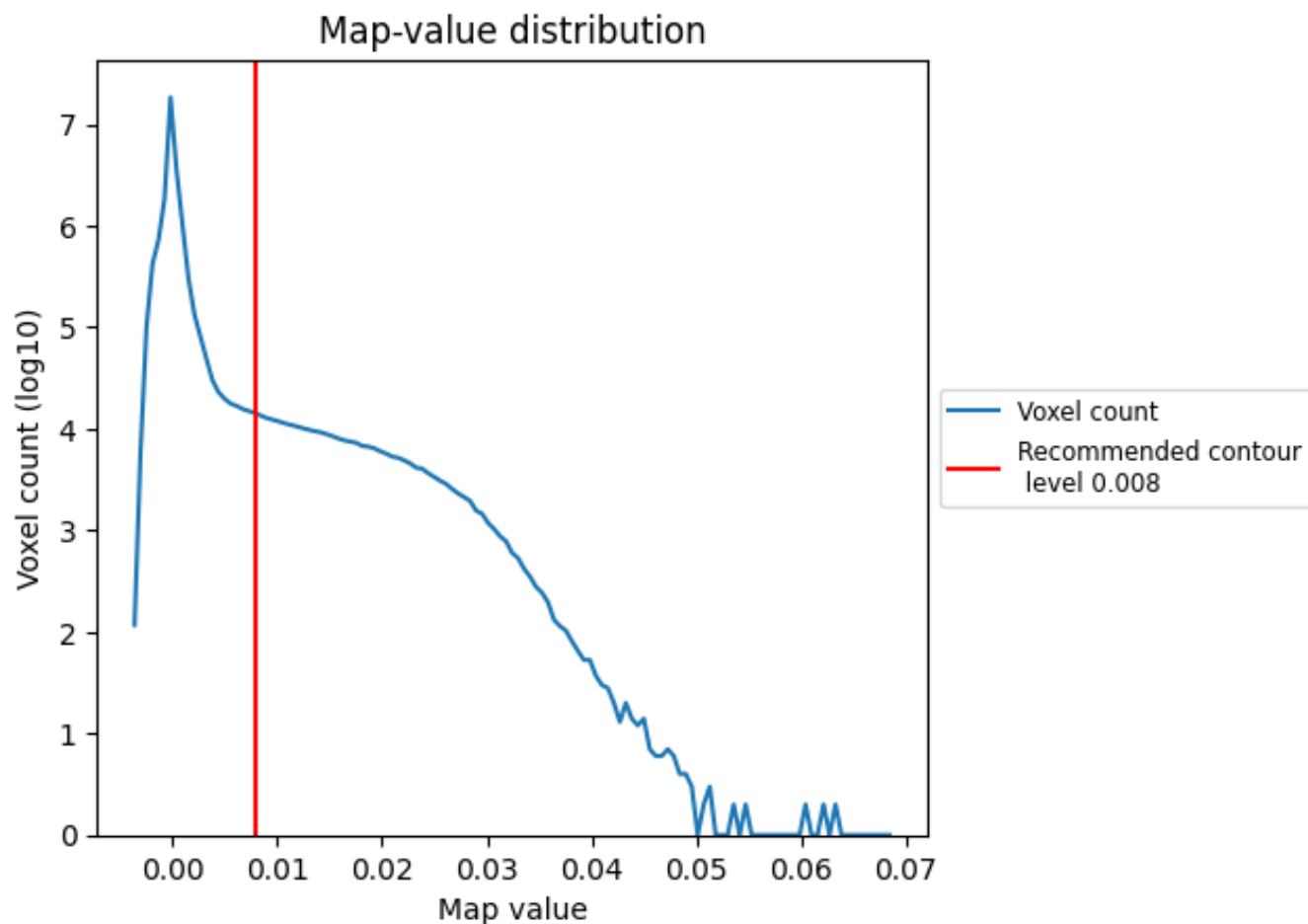
## 6.6 Mask visualisation [i](#)

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

## 7 Map analysis [i](#)

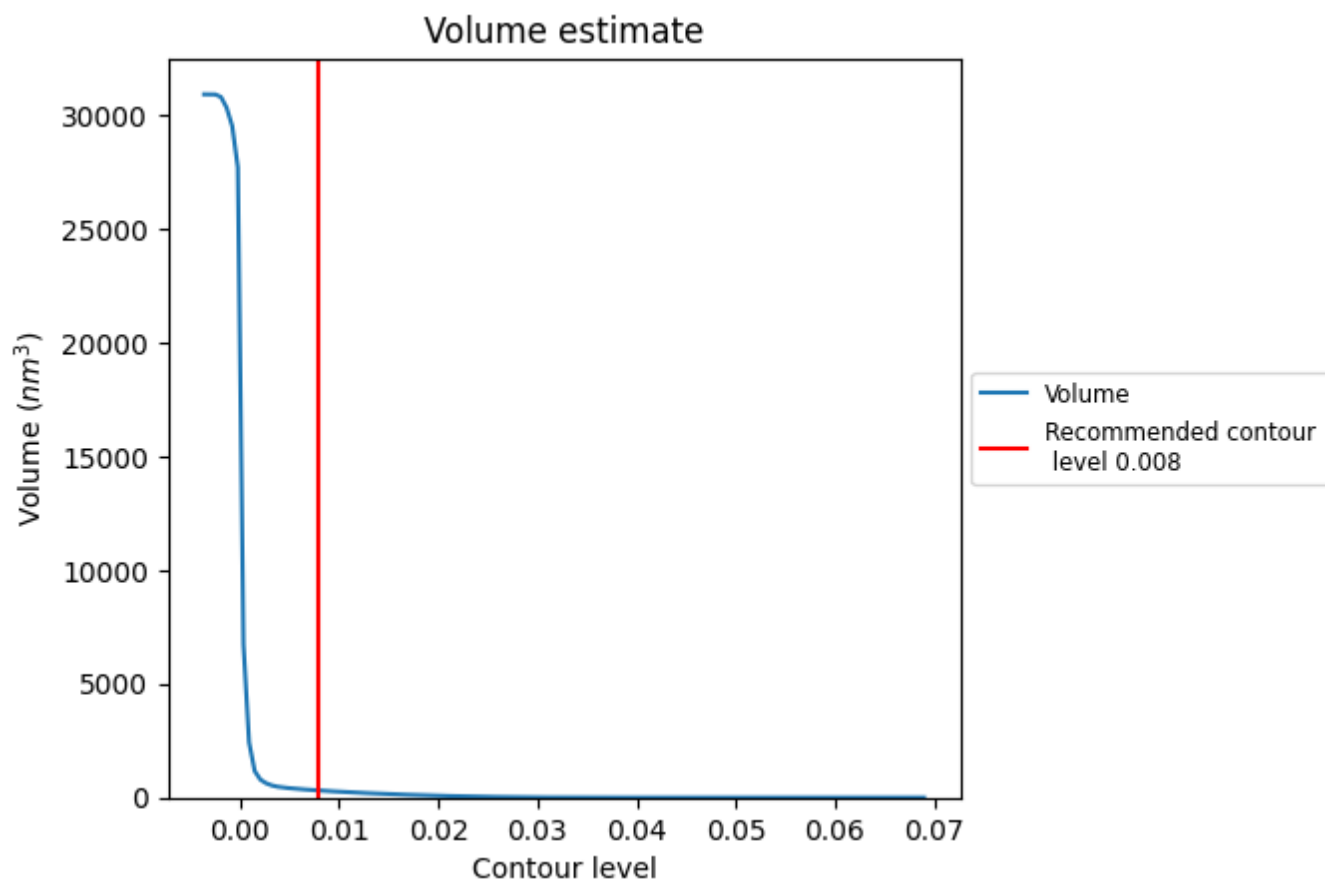
This section contains the results of statistical analysis of the map.

### 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.

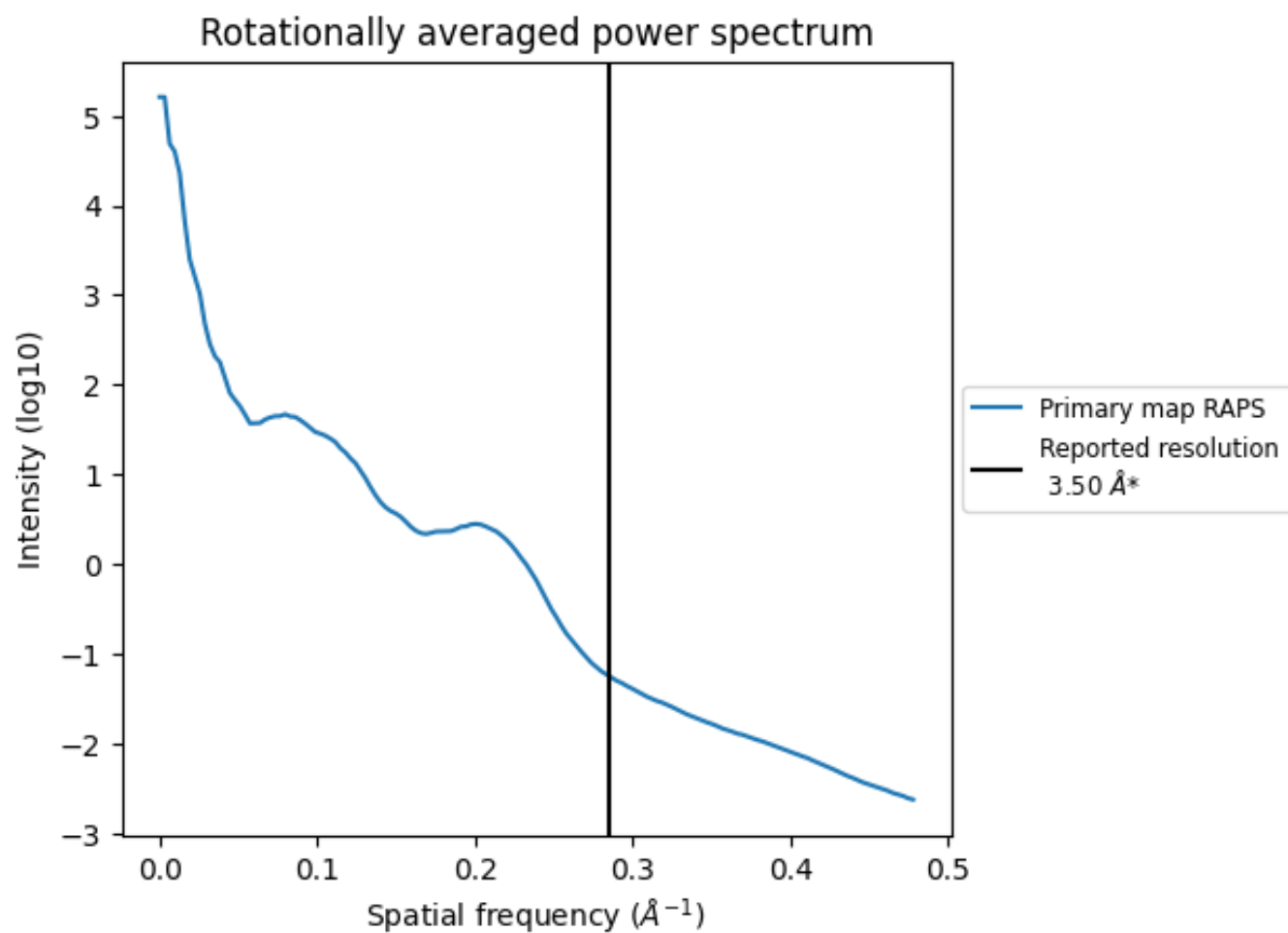
## 7.2 Volume estimate [i](#)



The volume at the recommended contour level is 311 nm<sup>3</sup>; this corresponds to an approximate mass of 281 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 0.286 Å<sup>-1</sup>

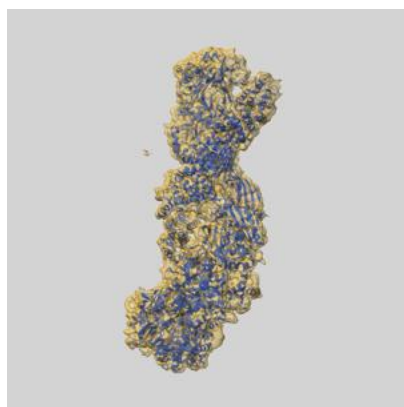
## 8 Fourier-Shell correlation

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

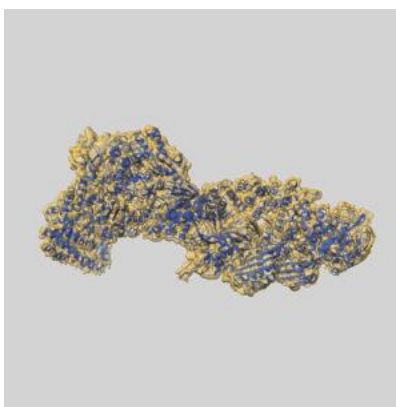
## 9 Map-model fit [i](#)

This section contains information regarding the fit between EMDB map EMD-16126 and PDB model 8BMW. Per-residue inclusion information can be found in section 3 on page 6.

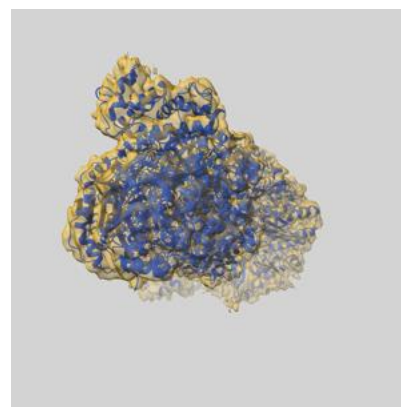
### 9.1 Map-model overlay [i](#)



X



Y

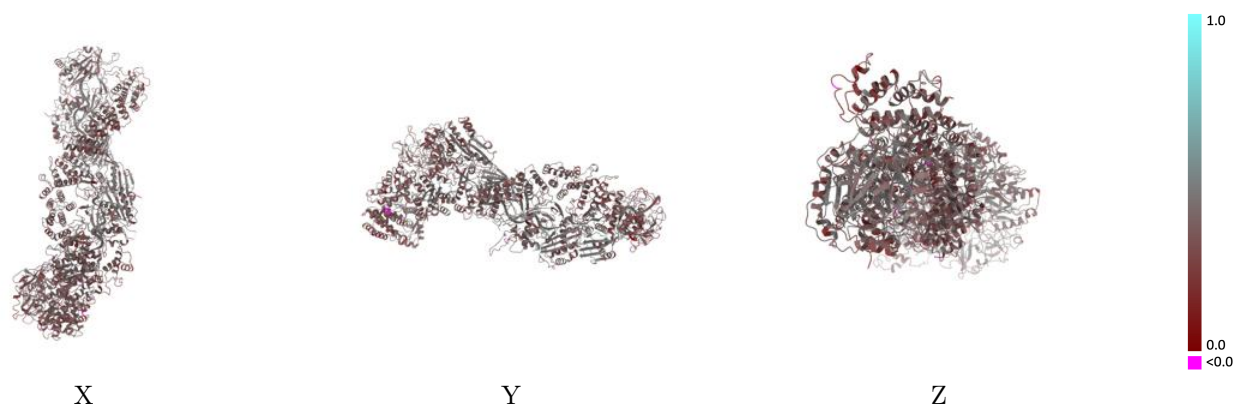


Z

The images above show the 3D surface view of the map at the recommended contour level 0.008 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.

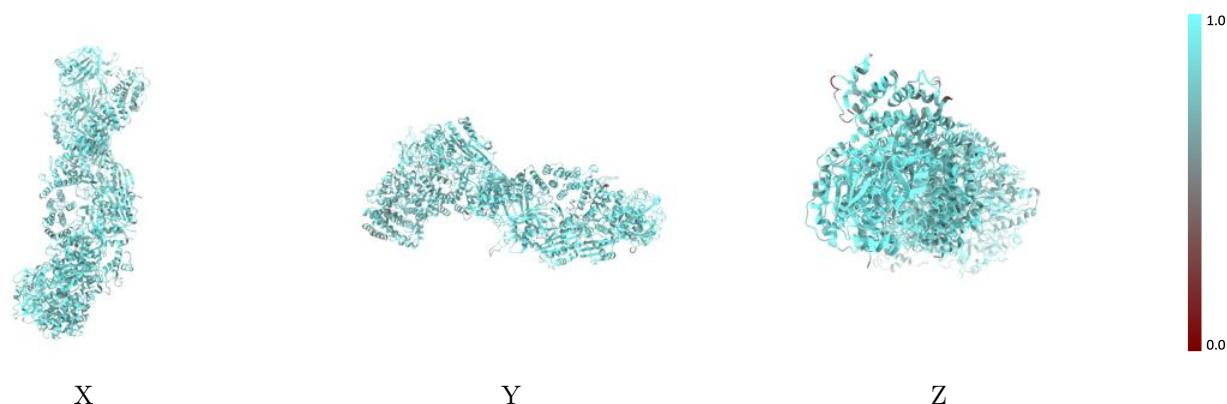


## 9.2 Q-score mapped to coordinate model [i](#)



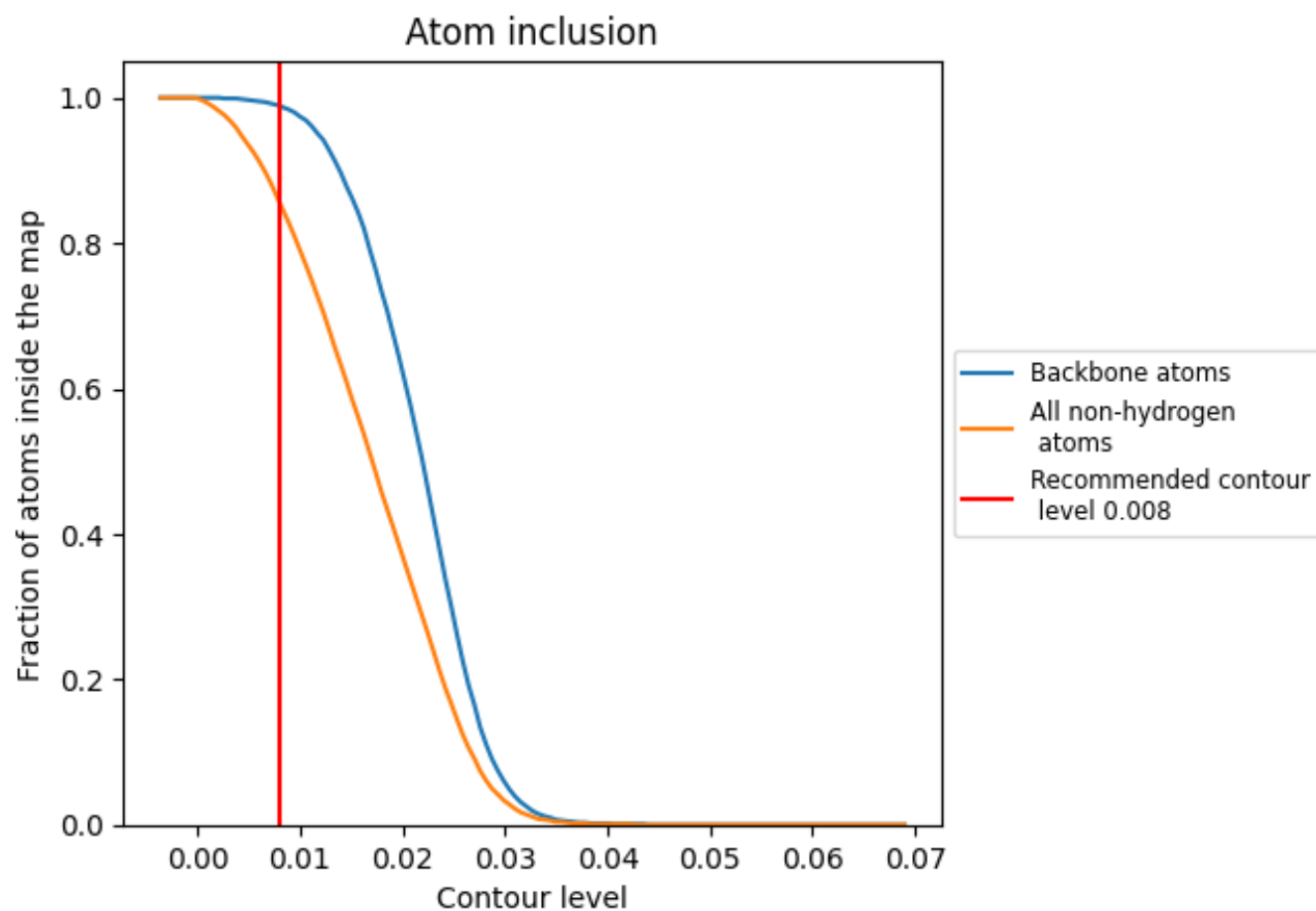
The images above show the model with each residue coloured according to its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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 (0.008).

## 9.4 Atom inclusion [i](#)



At the recommended contour level, 99% of all backbone atoms, 86% 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 (0.008) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	<div></div> 0.8560	<div></div> 0.3700
A	<div></div> 0.8250	<div></div> 0.3500
B	<div></div> 0.8510	<div></div> 0.3690
C	<div></div> 0.8630	<div></div> 0.3660
D	<div></div> 0.8490	<div></div> 0.3670
E	<div></div> 0.8340	<div></div> 0.3290
F	<div></div> 0.8660	<div></div> 0.3900
G	<div></div> 0.8570	<div></div> 0.3980
H	<div></div> 0.8770	<div></div> 0.4020
I	<div></div> 0.8660	<div></div> 0.3840
J	<div></div> 0.8750	<div></div> 0.3650
K	<div></div> 0.8210	<div></div> 0.3430
L	<div></div> 0.8580	<div></div> 0.3660
M	<div></div> 0.8810	<div></div> 0.3650
N	<div></div> 0.8370	<div></div> 0.3780
R	<div></div> 0.9890	<div></div> 0.4290

1.0

0.0

<0.0