



# wwPDB EM Validation Summary Report ⓘ

Feb 27, 2024 – 10:22 AM EST

PDB ID : 6VBW  
EMDB ID : EMD-21146  
Title : Cryo-EM structure of Cascade-TniQ-dsDNA ternary complex  
Authors : Jia, N.; Patel, D.J.  
Deposited on : 2019-12-19  
Resolution : 3.20 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:

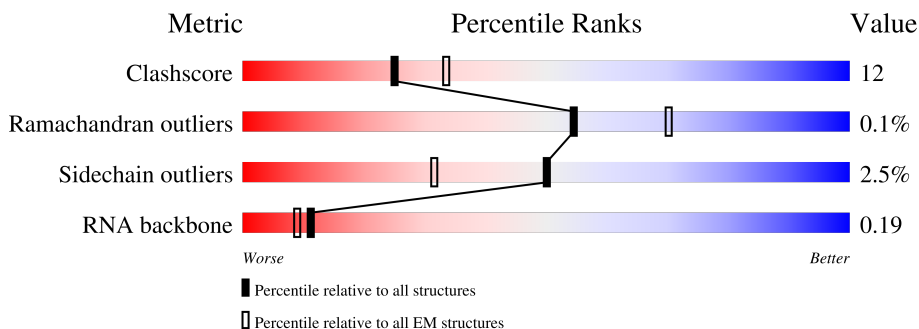
EMDB validation analysis : 0.0.1.dev70  
MolProbity : 4.02b-467  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
MapQ : 1.9.13  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.36

# 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 3.20 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
RNA backbone	4643	859

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	M	22	
2	L	100	
3	K	61	
4	A	640	
5	B	352	
5	C	352	
5	D	352	

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
5	E	352	 78% 18% ..
5	F	352	 78% 18% ..
5	G	352	 66% 19% . 13%
6	H	199	 52% 15% . 32%
7	I	394	 37% 77% 18% 5%
7	J	394	 58% 79% 17% ..

## 2 Entry composition [i](#)

There are 8 unique types of molecules in this entry. The entry contains 29087 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 DNA chain called DNA (5'-D(\*GP\*CP\*AP\*GP\*TP\*CP\*AP\*TP\*CP\*AP\*CP\*CP\*AP\*AP\*TP\*TP\*TP\*AP\*TP\*TP\*TP\*A)-3').

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
1	M	7	140	67	26	40	7	0	0

- Molecule 2 is a DNA chain called DNA (100-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
2	L	39	807	382	161	225	39	0	0

- Molecule 3 is a RNA chain called RNA (61-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
3	K	61	1275	569	219	427	60	0	0

- Molecule 4 is a protein called Cas8.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	A	450	3538	2244	615	661	18	0	0

- Molecule 5 is a protein called Cas7.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
5	B	338	2704	1720	464	505	15	0	0
5	C	339	2713	1726	466	506	15	0	0
5	D	341	2729	1736	469	509	15	0	0
5	F	341	2729	1736	469	509	15	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms					AltConf	Trace
5	E	340	Total	C	N	O	S	0	0
			2720	1730	467	508	15		
5	G	307	Total	C	N	O	S	0	0
			2490	1591	428	457	14		

- Molecule 6 is a protein called Cas6.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	H	135	Total	C	N	O	S	0	0
			1113	712	203	192	6		

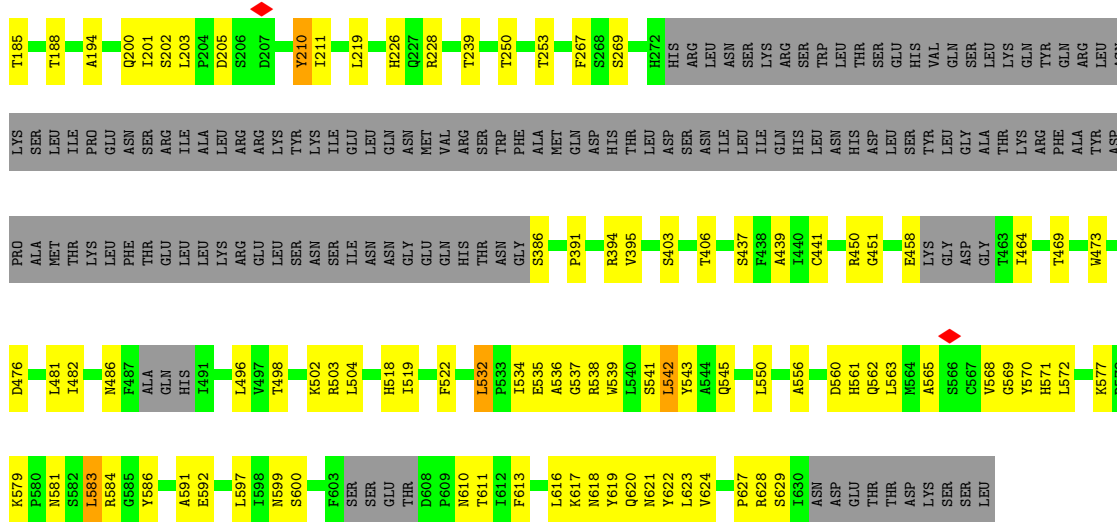
- Molecule 7 is a protein called TniQ.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	I	374	Total	C	N	O	S	0	0
			3050	1959	522	550	19		
7	J	378	Total	C	N	O	S	0	0
			3075	1975	526	555	19		

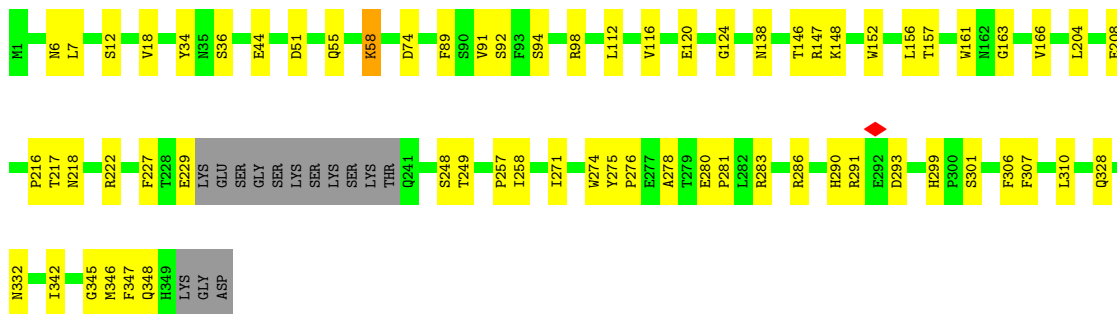
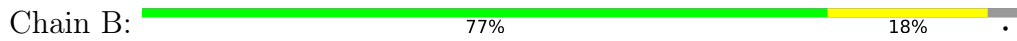
- Molecule 8 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		AltConf
8	I	2	Total	Zn	0
			2	2	
8	J	2	Total	Zn	0
			2	2	

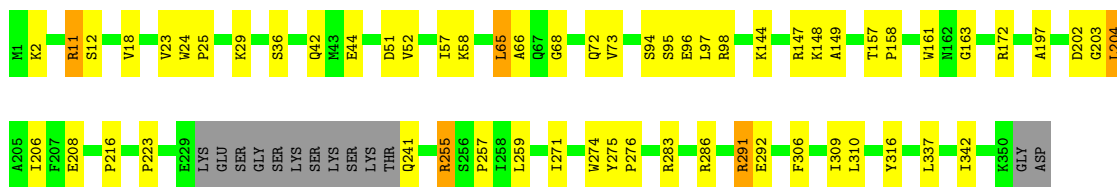
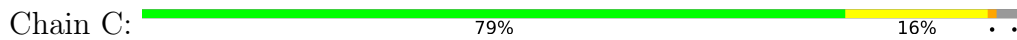




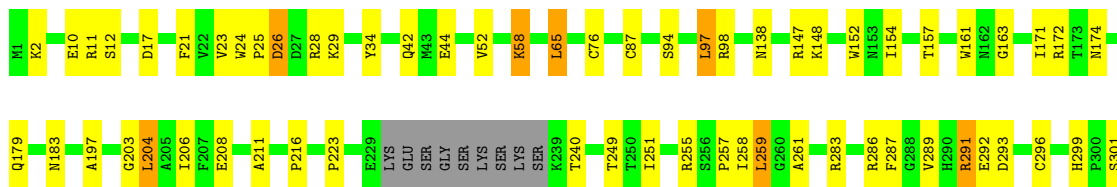
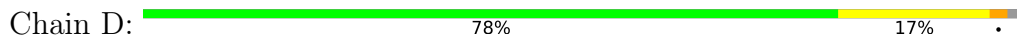
• Molecule 5: Cas7



• Molecule 5: Cas7

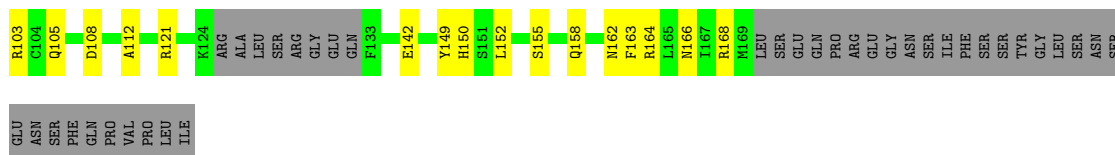


• Molecule 5: Cas7

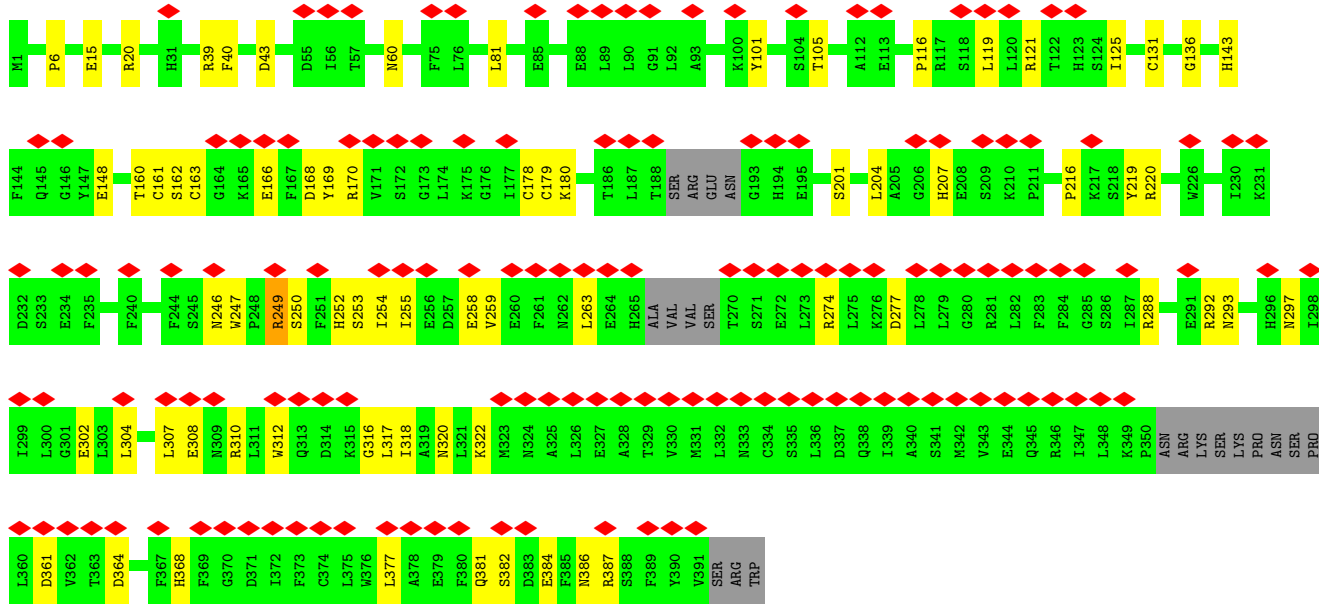
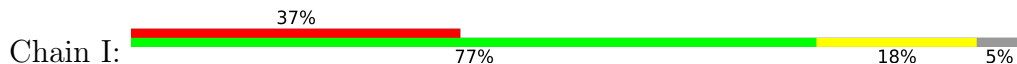




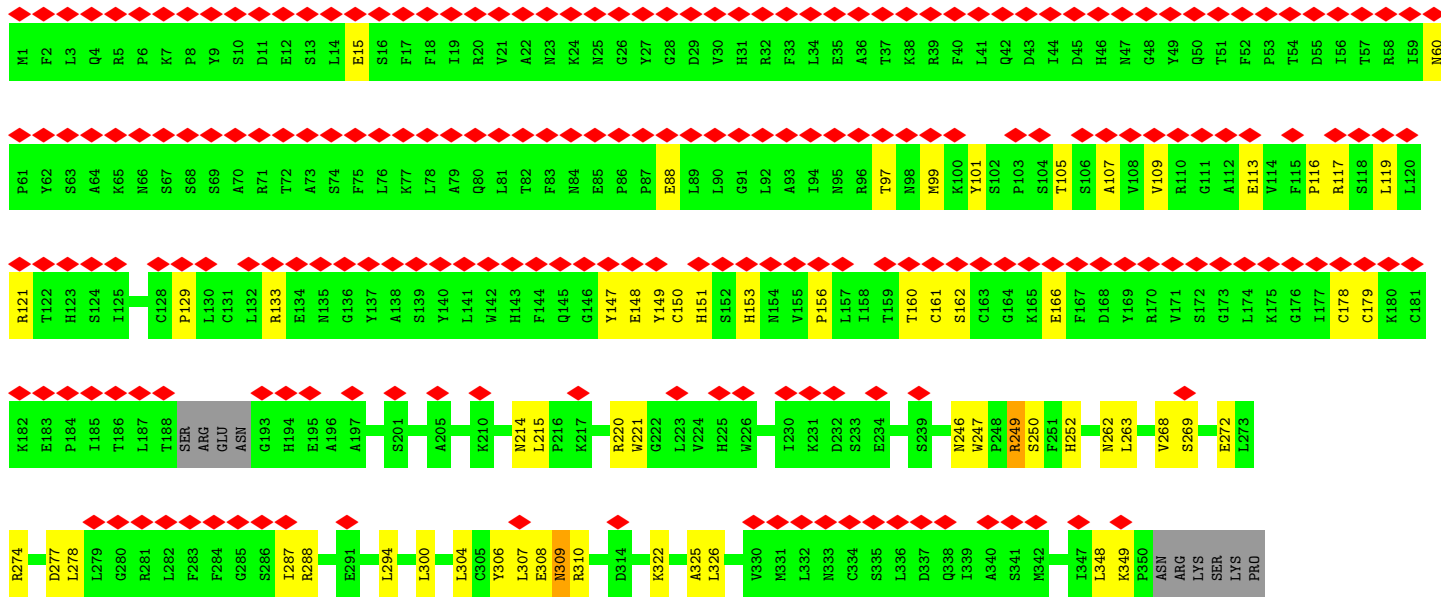


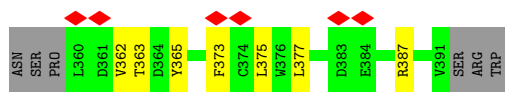


• Molecule 7: ThiQ



• Molecule 7: ThiQ





## 4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	55900	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 ( $e^-/\text{\AA}^2$ )	2.16	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	0.179	Depositor
Minimum map value	-0.102	Depositor
Average map value	0.000	Depositor
Map value standard deviation	0.004	Depositor
Recommended contour level	0.01	Depositor
Map size ( $\text{\AA}$ )	306.432, 306.432, 306.432	wwPDB
Map dimensions	288, 288, 288	wwPDB
Map angles ( $^\circ$ )	90.0, 90.0, 90.0	wwPDB
Pixel spacing ( $\text{\AA}$ )	1.064, 1.064, 1.064	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

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	M	0.74	0/156	0.84	0/237
2	L	1.27	0/909	0.91	0/1401
3	K	0.95	0/1423	1.05	0/2216
4	A	0.45	0/3611	0.65	0/4894
5	B	0.51	0/2775	0.65	0/3771
5	C	0.61	0/2784	0.70	0/3782
5	D	0.66	0/2800	0.73	0/3803
5	E	0.63	0/2791	0.72	0/3792
5	F	0.61	0/2800	0.71	0/3803
5	G	0.54	0/2557	0.68	0/3470
6	H	0.34	0/1135	0.57	0/1519
7	I	0.30	0/3134	0.51	0/4243
7	J	0.30	0/3160	0.48	0/4281
All	All	0.58	0/30035	0.69	0/41212

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	M	140	0	79	1	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	L	807	0	436	15	0
3	K	1275	0	642	60	0
4	A	3538	0	3528	158	0
5	B	2704	0	2632	60	0
5	C	2713	0	2645	48	0
5	D	2729	0	2665	64	0
5	E	2720	0	2652	56	0
5	F	2729	0	2665	46	0
5	G	2490	0	2433	111	0
6	H	1113	0	1113	52	0
7	I	3050	0	2960	53	0
7	J	3075	0	2992	58	0
8	I	2	0	0	0	0
8	J	2	0	0	0	0
All	All	29087	0	27442	653	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 653 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)
4:A:560:ASP:CB	4:A:563:LEU:HD13	1.46	1.44
4:A:611:THR:HG22	4:A:629:SER:CB	1.52	1.39
4:A:619:TYR:HD2	4:A:622:TYR:CE2	1.46	1.30
6:H:12:PRO:CG	6:H:15:CYS:HB3	1.65	1.26
4:A:619:TYR:CD2	4:A:622:TYR:CE2	2.28	1.21

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	
4	A	432/640 (68%)	380 (88%)	50 (12%)	2 (0%)	29	67
5	B	334/352 (95%)	301 (90%)	33 (10%)	0	100	100
5	C	335/352 (95%)	304 (91%)	31 (9%)	0	100	100
5	D	337/352 (96%)	315 (94%)	22 (6%)	0	100	100
5	E	336/352 (96%)	305 (91%)	31 (9%)	0	100	100
5	F	337/352 (96%)	303 (90%)	34 (10%)	0	100	100
5	G	301/352 (86%)	274 (91%)	27 (9%)	0	100	100
6	H	125/199 (63%)	113 (90%)	12 (10%)	0	100	100
7	I	366/394 (93%)	344 (94%)	22 (6%)	0	100	100
7	J	372/394 (94%)	354 (95%)	17 (5%)	1 (0%)	41	74
All	All	3275/3739 (88%)	2993 (91%)	279 (8%)	3 (0%)	54	83

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	A	536	ALA
4	A	620	GLN
7	J	307	LEU

### 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	
4	A	400/573 (70%)	391 (98%)	9 (2%)	50	78
5	B	296/308 (96%)	291 (98%)	5 (2%)	60	83
5	C	297/308 (96%)	286 (96%)	11 (4%)	34	68
5	D	299/308 (97%)	285 (95%)	14 (5%)	26	62
5	E	298/308 (97%)	291 (98%)	7 (2%)	50	78
5	F	299/308 (97%)	289 (97%)	10 (3%)	38	71
5	G	275/308 (89%)	262 (95%)	13 (5%)	26	62
6	H	121/181 (67%)	118 (98%)	3 (2%)	47	77

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
7	I	337/356 (95%)	336 (100%)	1 (0%)	92	96
7	J	340/356 (96%)	338 (99%)	2 (1%)	86	94
All	All	2962/3314 (89%)	2887 (98%)	75 (2%)	50	77

5 of 75 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
5	G	11	ARG
6	H	102	ARG
5	G	95	SER
5	G	283	ARG
5	C	337	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 34 such sidechains are listed below:

Mol	Chain	Res	Type
7	I	194	HIS
7	I	320	ASN
7	J	237	HIS
5	C	77	HIS
5	C	72	GLN

### 5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
3	K	59/61 (96%)	35 (59%)	12 (20%)

5 of 35 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
3	K	3	G
3	K	6	A
3	K	9	U
3	K	10	U
3	K	14	G

5 of 12 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
3	K	37	U
3	K	41	G
3	K	46	C
3	K	42	U
3	K	19	C

#### 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

#### 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

#### 5.6 Ligand geometry [i](#)

Of 4 ligands modelled in this entry, 4 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

#### 5.7 Other polymers [i](#)

There are no such residues in this entry.

#### 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.



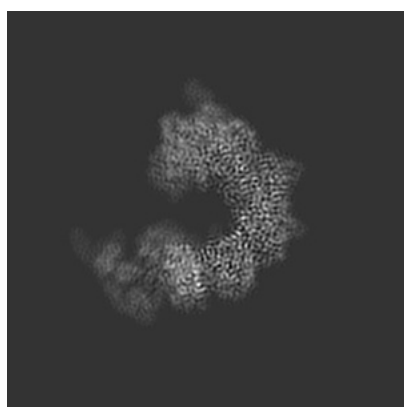
## 6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-21146. These allow visual inspection of the internal detail of the map and identification of artifacts.

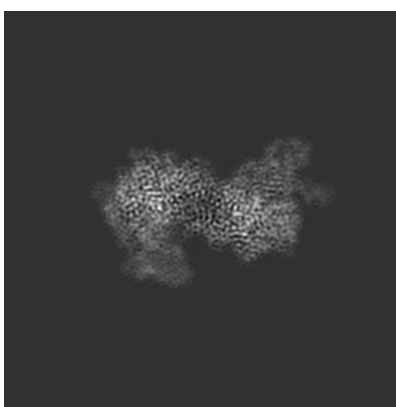
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

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

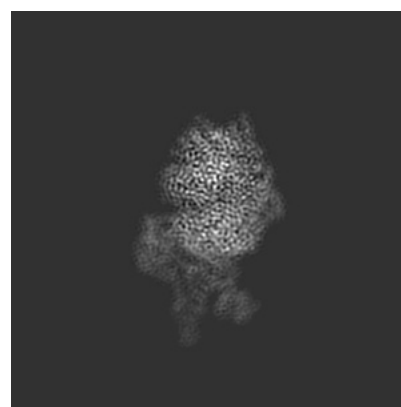
#### 6.1.1 Primary map



X



Y

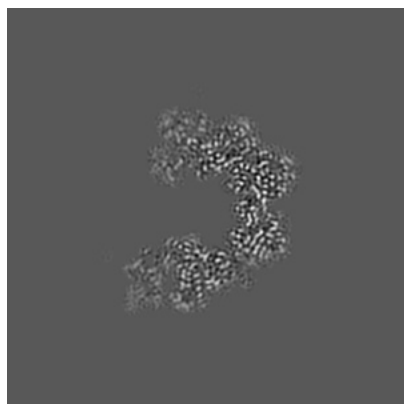


Z

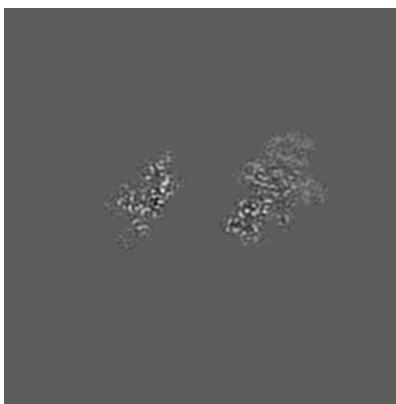
The images above show the map projected in three orthogonal directions.

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

#### 6.2.1 Primary map



X Index: 144



Y Index: 144



Z Index: 144

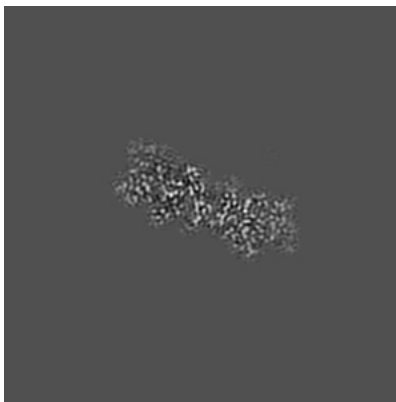
The images above show central slices of the map in three orthogonal directions.

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

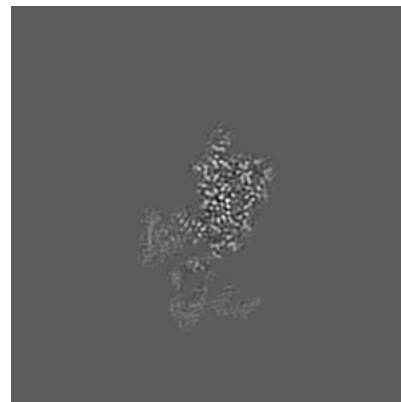
### 6.3.1 Primary map



X Index: 146



Y Index: 169

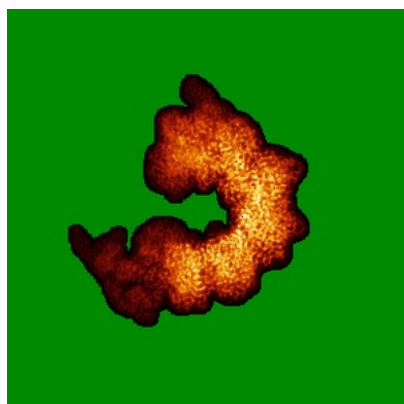


Z Index: 107

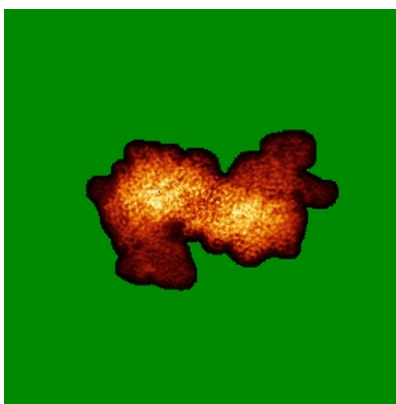
The images above show the largest variance slices of the map in three orthogonal directions.

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

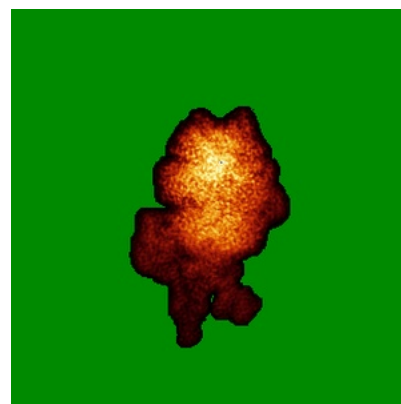
### 6.4.1 Primary map



X



Y

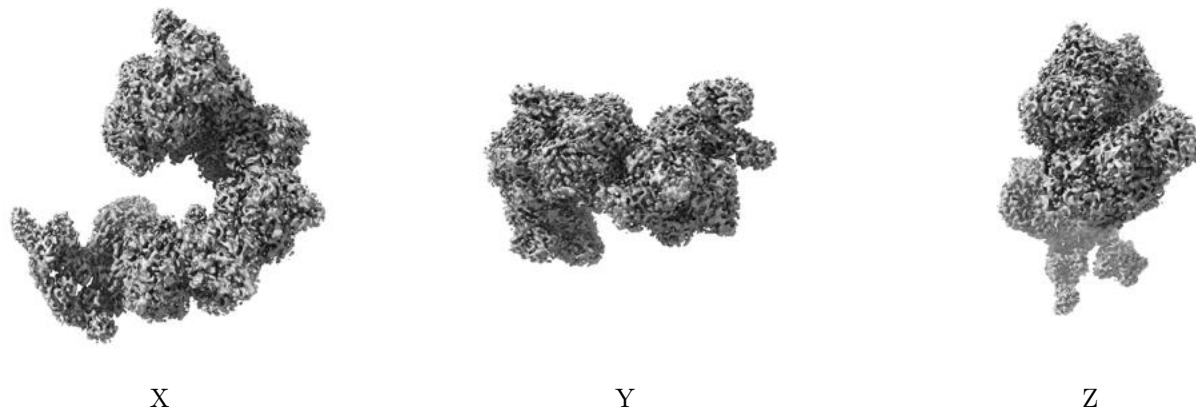


Z

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

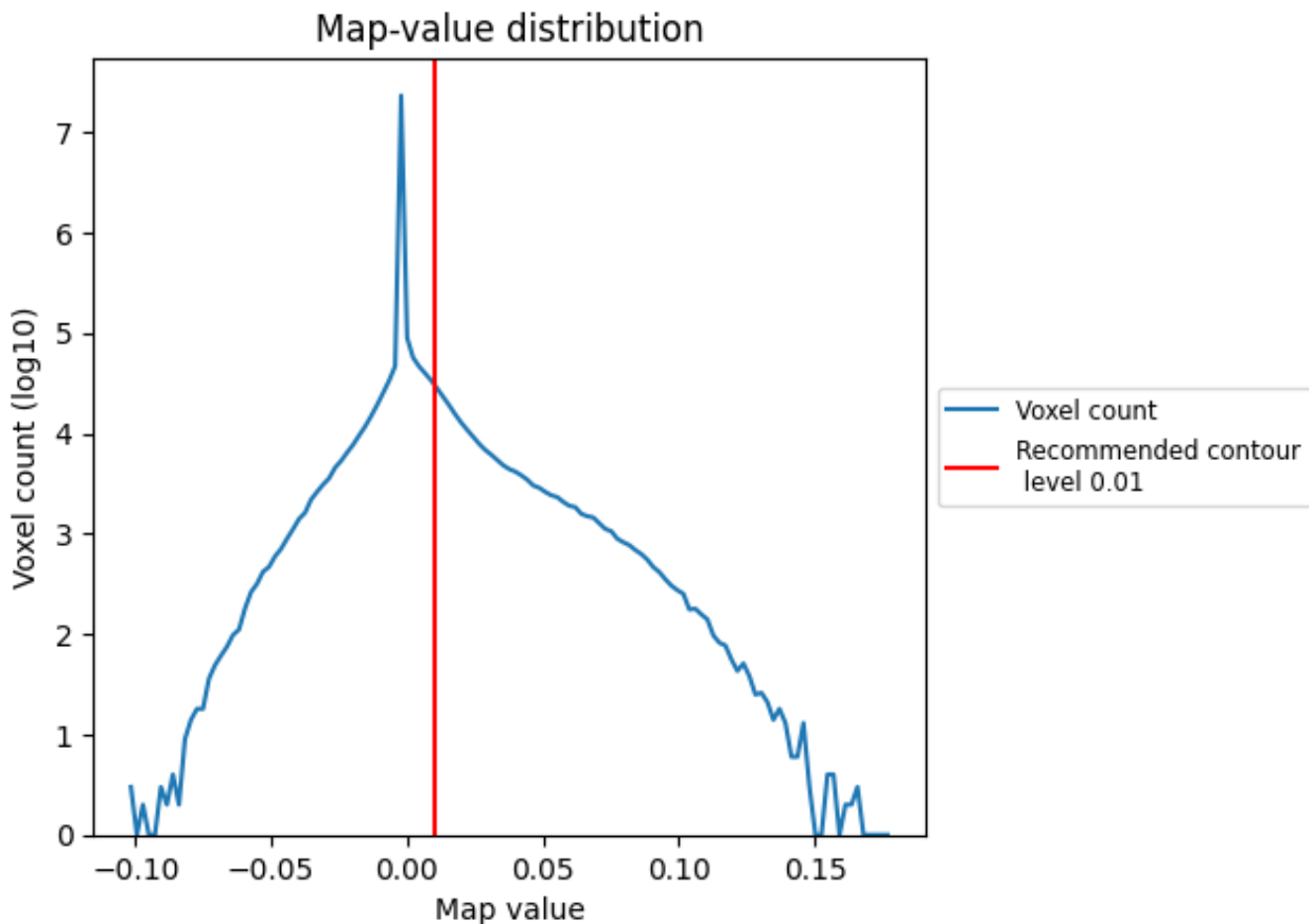
## 6.6 Mask visualisation [i](#)

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

## 7 Map analysis [i](#)

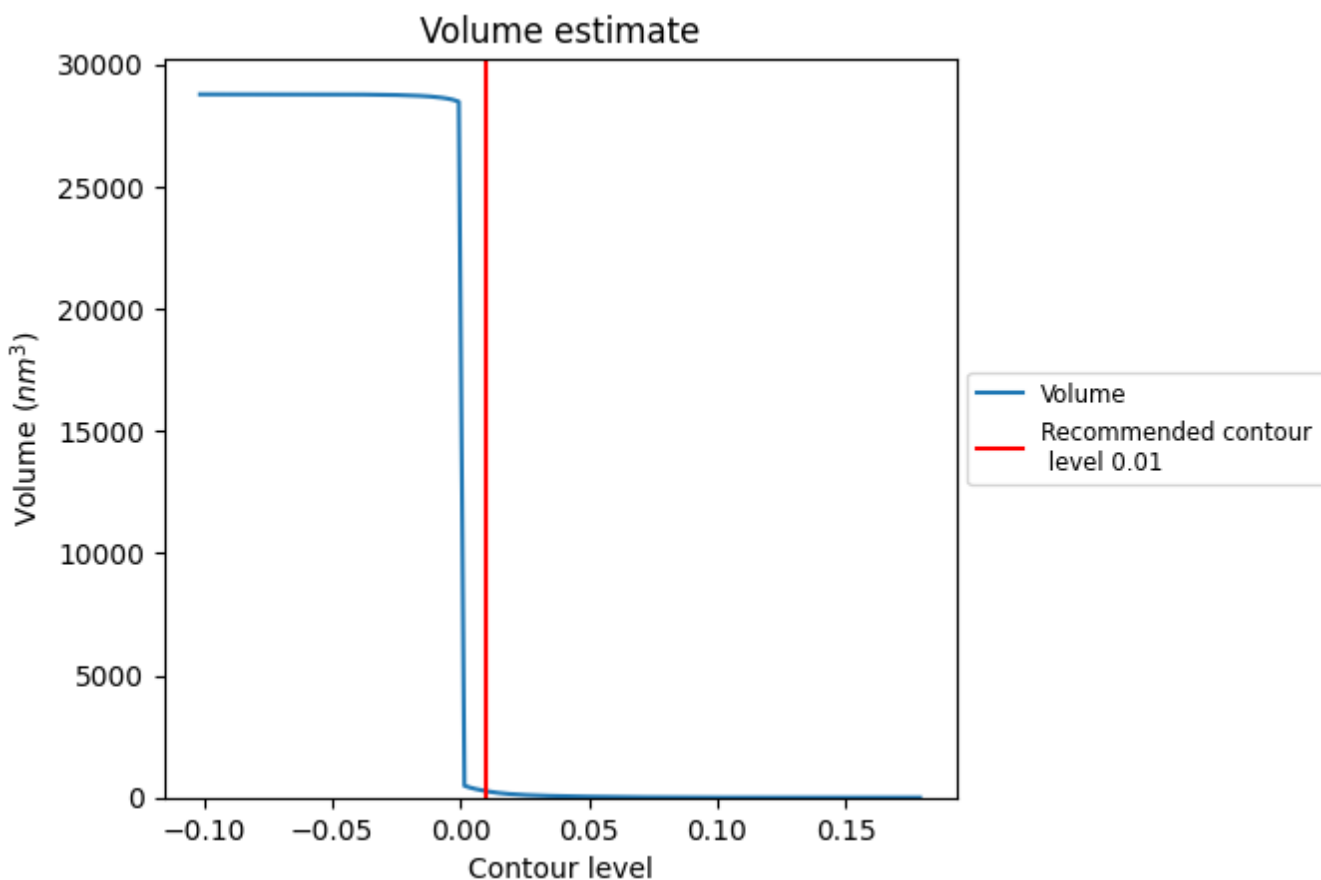
This section contains the results of statistical analysis of the map.

### 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.

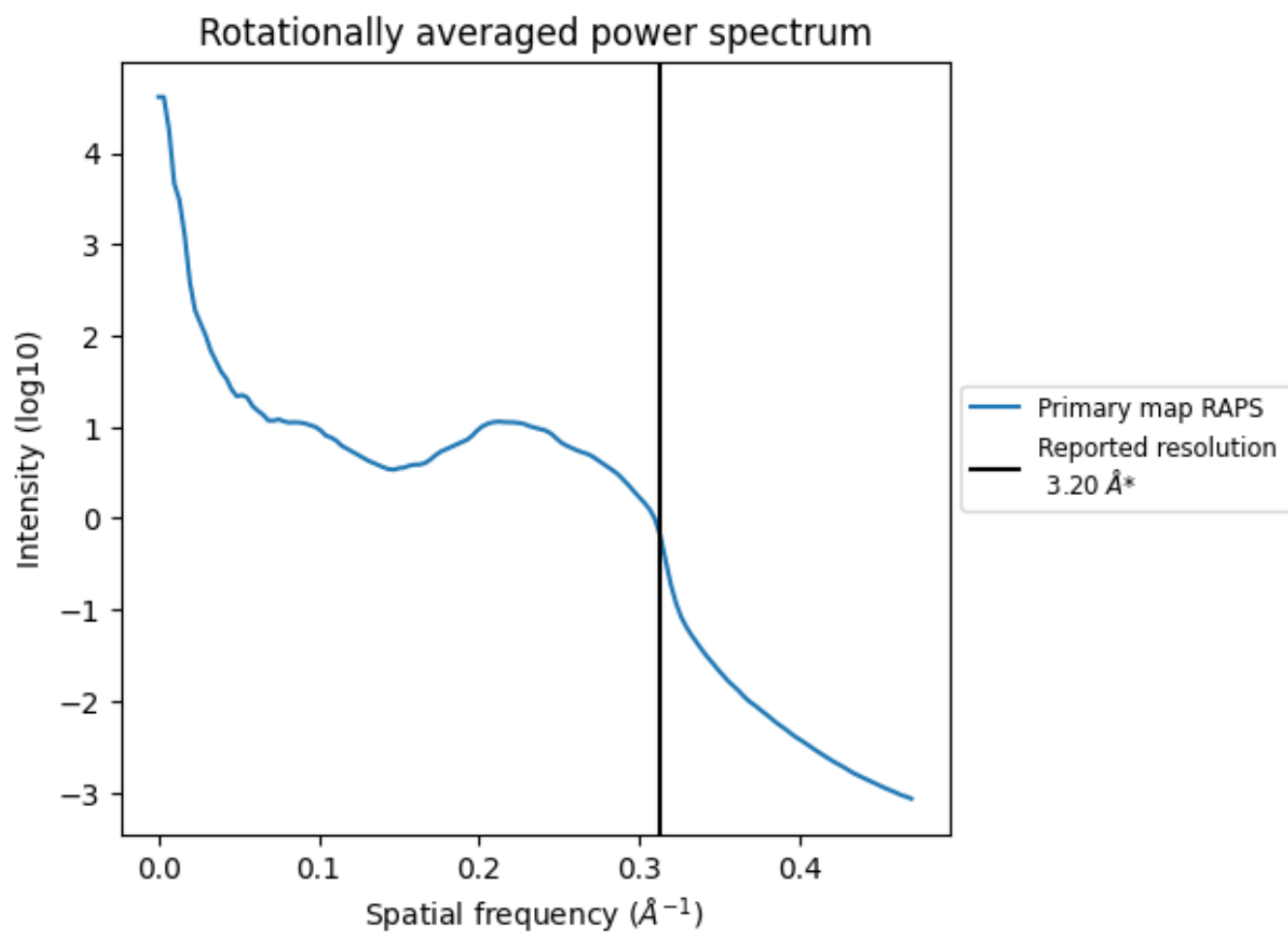
## 7.2 Volume estimate [i](#)



The volume at the recommended contour level is 261 nm<sup>3</sup>; this corresponds to an approximate mass of 236 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](#)

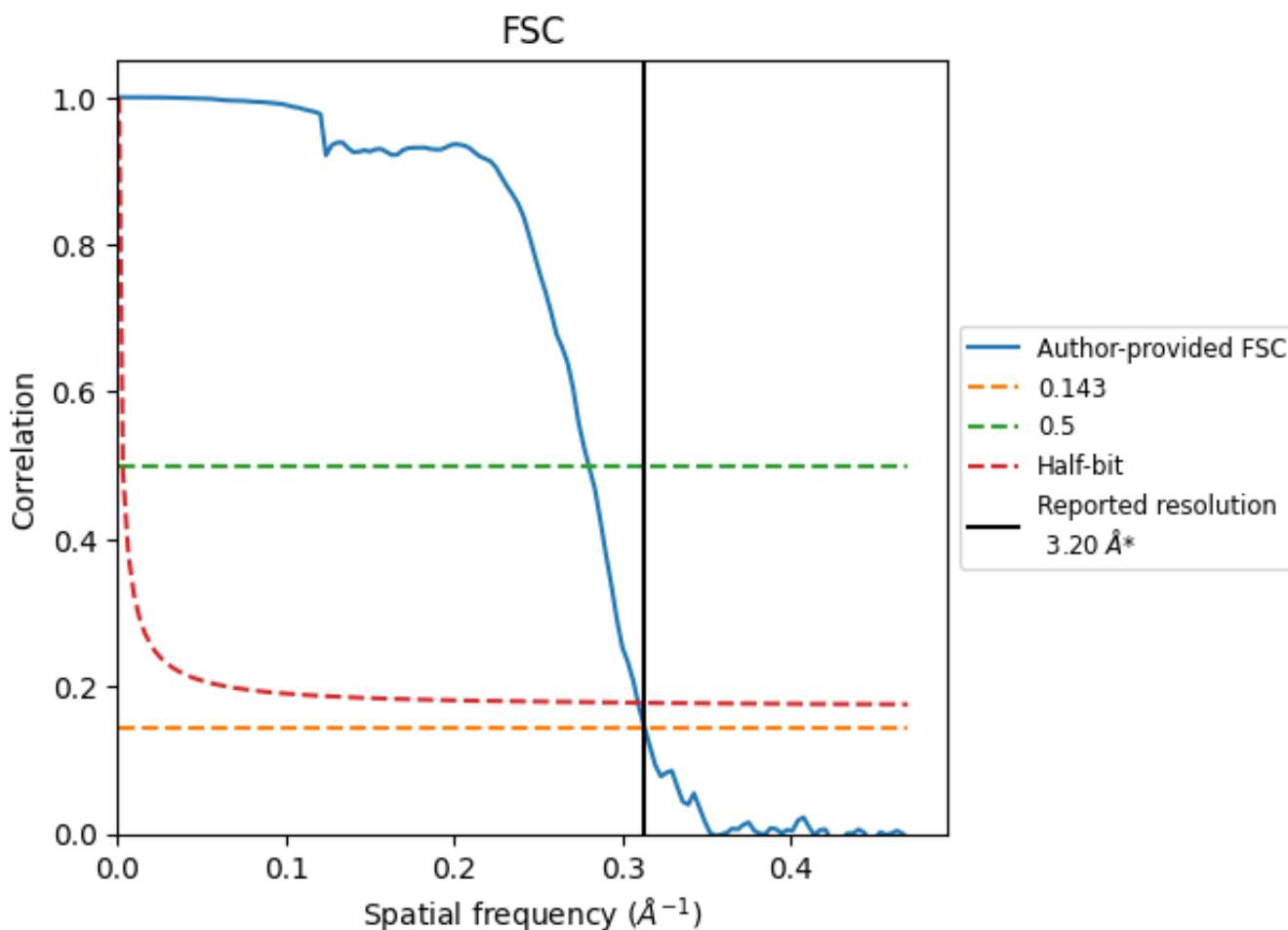


\*Reported resolution corresponds to spatial frequency of 0.312 Å<sup>-1</sup>

## 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 0.312 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	3.20	-	-
Author-provided FSC curve	3.19	3.57	3.23
Unmasked-calculated*	-	-	-

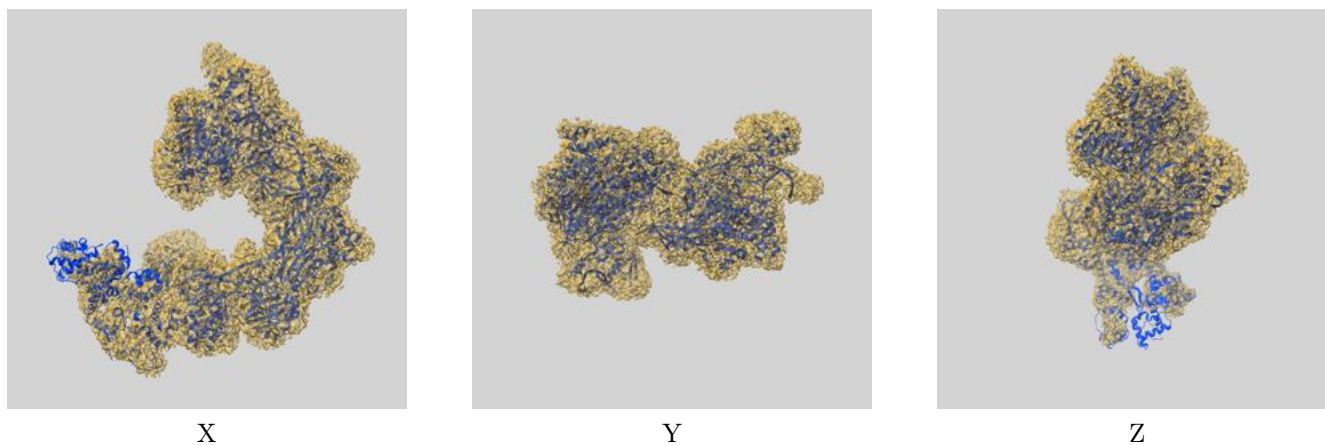
\*Resolution estimate based on FSC curve calculated by comparison of deposited half-maps.



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

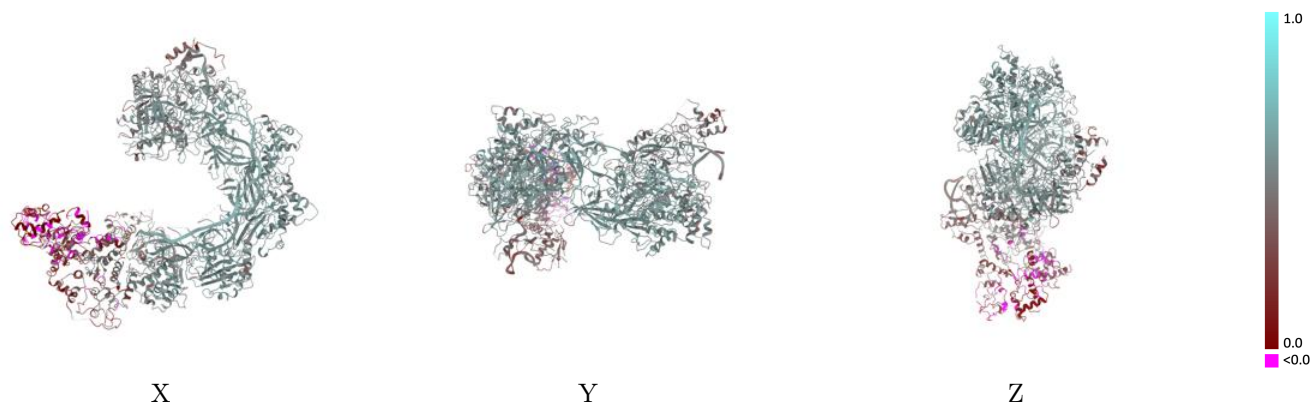
This section contains information regarding the fit between EMDB map EMD-21146 and PDB model 6VBW. Per-residue inclusion information can be found in section 3 on page 6.

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



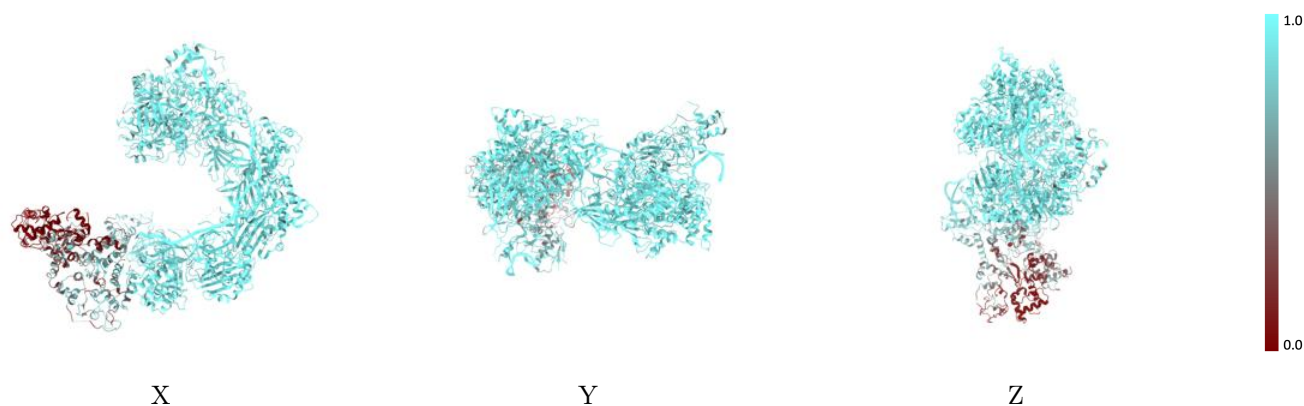
The images above show the 3D surface view of the map at the recommended contour level 0.01 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

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



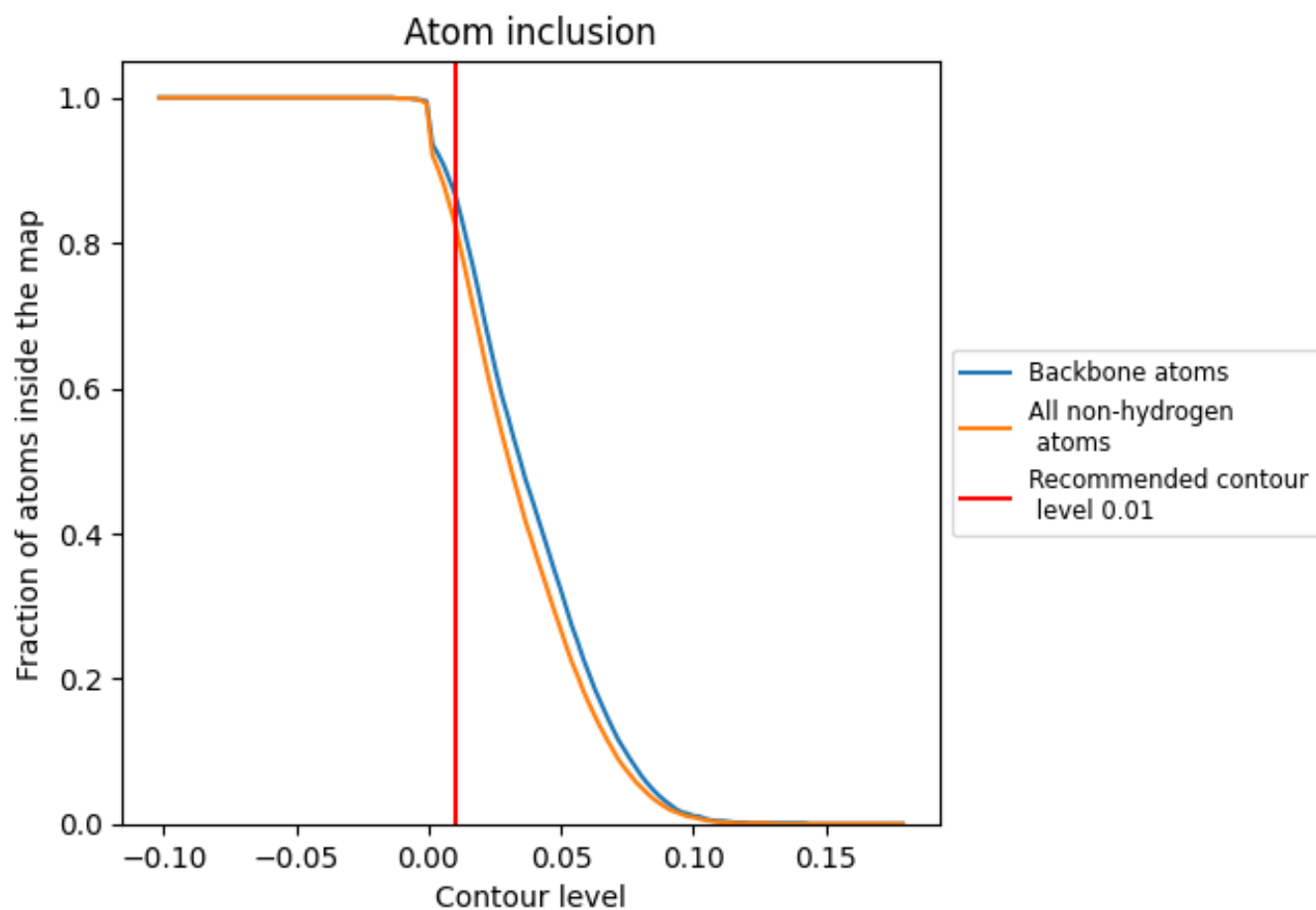
The images above show the model with each residue coloured according to its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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 (0.01).





























## 9.4 Atom inclusion [i](#)



At the recommended contour level, 87% of all backbone atoms, 83% 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 (0.01) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	 0.8260	 0.4660
A	 0.8870	 0.4870
B	 0.9320	 0.5260
C	 0.9550	 0.5590
D	 0.9620	 0.5690
E	 0.9570	 0.5680
F	 0.9540	 0.5600
G	 0.9380	 0.5360
H	 0.8000	 0.3640
I	 0.5020	 0.2630
J	 0.3410	 0.1730
K	 0.9470	 0.5150
L	 0.9860	 0.5890
M	 0.9710	 0.4940

