



## Full wwPDB EM Validation Report ⓘ

Nov 23, 2022 – 07:51 AM JST

PDB ID : 7V22  
EMDB ID : EMD-31634  
Title : Local CryoEM structure del68-76/del679-688 prefusion-stabilized spike  
Authors : Du, S.; Xiao, J.  
Deposited on : 2021-08-07  
Resolution : 3.38 Å (reported)

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) : 2.31.3

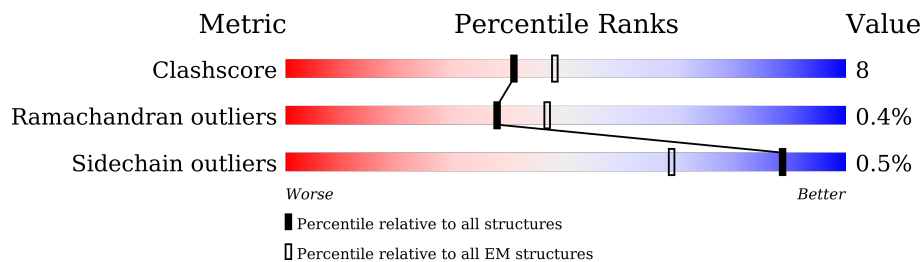
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 3.38 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	A	1269	

## 2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 1958 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 Spike glycoprotein.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	240	1958	1278	321	354	5	0	0

There are 105 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	?	-	ILE	deletion	UNP P0DTC2
A	?	-	HIS	deletion	UNP P0DTC2
A	?	-	VAL	deletion	UNP P0DTC2
A	?	-	SER	deletion	UNP P0DTC2
A	?	-	GLY	deletion	UNP P0DTC2
A	?	-	THR	deletion	UNP P0DTC2
A	?	-	ASN	deletion	UNP P0DTC2
A	?	-	GLY	deletion	UNP P0DTC2
A	?	-	THR	deletion	UNP P0DTC2
A	?	-	ASN	deletion	UNP P0DTC2
A	?	-	SER	deletion	UNP P0DTC2
A	?	-	PRO	deletion	UNP P0DTC2
A	?	-	ARG	deletion	UNP P0DTC2
A	?	-	ARG	deletion	UNP P0DTC2
A	?	-	ALA	deletion	UNP P0DTC2
A	?	-	ARG	deletion	UNP P0DTC2
A	?	-	SER	deletion	UNP P0DTC2
A	?	-	VAL	deletion	UNP P0DTC2
A	?	-	ALA	deletion	UNP P0DTC2
A	807	PRO	PHE	engineered mutation	UNP P0DTC2
A	882	PRO	ALA	engineered mutation	UNP P0DTC2
A	889	PRO	ALA	engineered mutation	UNP P0DTC2
A	932	PRO	ALA	engineered mutation	UNP P0DTC2
A	976	PRO	LYS	engineered mutation	UNP P0DTC2
A	977	PRO	VAL	engineered mutation	UNP P0DTC2
A	1199	GLY	-	expression tag	UNP P0DTC2
A	1200	SER	-	expression tag	UNP P0DTC2
A	1201	GLY	-	expression tag	UNP P0DTC2

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	1202	TYR	-	expression tag	UNP P0DTC2
A	1203	ILE	-	expression tag	UNP P0DTC2
A	1204	PRO	-	expression tag	UNP P0DTC2
A	1205	GLU	-	expression tag	UNP P0DTC2
A	1206	ALA	-	expression tag	UNP P0DTC2
A	1207	PRO	-	expression tag	UNP P0DTC2
A	1208	ARG	-	expression tag	UNP P0DTC2
A	1209	ASP	-	expression tag	UNP P0DTC2
A	1210	GLY	-	expression tag	UNP P0DTC2
A	1211	GLN	-	expression tag	UNP P0DTC2
A	1212	ALA	-	expression tag	UNP P0DTC2
A	1213	TYR	-	expression tag	UNP P0DTC2
A	1214	VAL	-	expression tag	UNP P0DTC2
A	1215	ARG	-	expression tag	UNP P0DTC2
A	1216	LYS	-	expression tag	UNP P0DTC2
A	1217	ASP	-	expression tag	UNP P0DTC2
A	1218	GLY	-	expression tag	UNP P0DTC2
A	1219	GLU	-	expression tag	UNP P0DTC2
A	1220	TRP	-	expression tag	UNP P0DTC2
A	1221	VAL	-	expression tag	UNP P0DTC2
A	1222	LEU	-	expression tag	UNP P0DTC2
A	1223	LEU	-	expression tag	UNP P0DTC2
A	1224	SER	-	expression tag	UNP P0DTC2
A	1225	THR	-	expression tag	UNP P0DTC2
A	1226	PHE	-	expression tag	UNP P0DTC2
A	1227	LEU	-	expression tag	UNP P0DTC2
A	1228	GLY	-	expression tag	UNP P0DTC2
A	1229	ARG	-	expression tag	UNP P0DTC2
A	1230	SER	-	expression tag	UNP P0DTC2
A	1231	LEU	-	expression tag	UNP P0DTC2
A	1232	GLU	-	expression tag	UNP P0DTC2
A	1233	VAL	-	expression tag	UNP P0DTC2
A	1234	LEU	-	expression tag	UNP P0DTC2
A	1235	PHE	-	expression tag	UNP P0DTC2
A	1236	GLN	-	expression tag	UNP P0DTC2
A	1237	GLY	-	expression tag	UNP P0DTC2
A	1238	PRO	-	expression tag	UNP P0DTC2
A	1239	GLY	-	expression tag	UNP P0DTC2
A	1240	HIS	-	expression tag	UNP P0DTC2
A	1241	HIS	-	expression tag	UNP P0DTC2
A	1242	HIS	-	expression tag	UNP P0DTC2
A	1243	HIS	-	expression tag	UNP P0DTC2

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	1244	HIS	-	expression tag	UNP P0DTC2
A	1245	HIS	-	expression tag	UNP P0DTC2
A	1246	HIS	-	expression tag	UNP P0DTC2
A	1247	HIS	-	expression tag	UNP P0DTC2
A	1248	SER	-	expression tag	UNP P0DTC2
A	1249	ALA	-	expression tag	UNP P0DTC2
A	1250	TRP	-	expression tag	UNP P0DTC2
A	1251	SER	-	expression tag	UNP P0DTC2
A	1252	HIS	-	expression tag	UNP P0DTC2
A	1253	PRO	-	expression tag	UNP P0DTC2
A	1254	GLN	-	expression tag	UNP P0DTC2
A	1255	PHE	-	expression tag	UNP P0DTC2
A	1256	GLU	-	expression tag	UNP P0DTC2
A	1257	LYS	-	expression tag	UNP P0DTC2
A	1258	GLY	-	expression tag	UNP P0DTC2
A	1259	GLY	-	expression tag	UNP P0DTC2
A	1260	GLY	-	expression tag	UNP P0DTC2
A	1261	SER	-	expression tag	UNP P0DTC2
A	1262	GLY	-	expression tag	UNP P0DTC2
A	1263	GLY	-	expression tag	UNP P0DTC2
A	1264	GLY	-	expression tag	UNP P0DTC2
A	1265	GLY	-	expression tag	UNP P0DTC2
A	1266	SER	-	expression tag	UNP P0DTC2
A	1267	GLY	-	expression tag	UNP P0DTC2
A	1268	GLY	-	expression tag	UNP P0DTC2
A	1269	SER	-	expression tag	UNP P0DTC2
A	1270	ALA	-	expression tag	UNP P0DTC2
A	1271	TRP	-	expression tag	UNP P0DTC2
A	1272	SER	-	expression tag	UNP P0DTC2
A	1273	HIS	-	expression tag	UNP P0DTC2
A	1274	PRO	-	expression tag	UNP P0DTC2
A	1275	GLN	-	expression tag	UNP P0DTC2
A	1276	PHE	-	expression tag	UNP P0DTC2
A	1277	GLU	-	expression tag	UNP P0DTC2
A	1278	LYS	-	expression tag	UNP P0DTC2





## 4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	200144	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 ( $e^-/\text{\AA}^2$ )	50	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	3.800	Depositor
Minimum map value	-2.231	Depositor
Average map value	0.000	Depositor
Map value standard deviation	0.046	Depositor
Recommended contour level	0.888	Depositor
Map size ( $\text{\AA}$ )	432.00003, 432.00003, 432.00003	wwPDB
Map dimensions	400, 400, 400	wwPDB
Map angles ( $^\circ$ )	90.0, 90.0, 90.0	wwPDB
Pixel spacing ( $\text{\AA}$ )	1.08, 1.08, 1.08	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	0.26	0/2013	0.52	0/2740

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	1958	0	1903	29	0
All	All	1958	0	1903	29	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 8.

All (29) 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:A:278:LYS:HD2	1:A:287:ASP:HB3	1.69	0.74
1:A:77:LYS:HD3	1:A:78:ARG:HG3	1.75	0.69
1:A:34:ARG:HG3	1:A:216:LEU:HD23	1.76	0.66
1:A:130:VAL:HG21	1:A:231:ILE:HD12	1.78	0.65
1:A:143:VAL:HG13	1:A:153:MET:HG3	1.78	0.64

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:93:ALA:O	1:A:266:TYR:N	2.35	0.58
1:A:121:ASN:O	1:A:121:ASN:ND2	2.33	0.56
1:A:105:ILE:HD11	1:A:239:GLN:HB3	1.91	0.51
1:A:126:VAL:HB	1:A:174:PRO:HA	1.94	0.50
1:A:127:VAL:HG12	1:A:171:VAL:HG12	1.94	0.49
1:A:144:TYR:HB2	1:A:154:GLU:HB3	1.93	0.49
1:A:187:LYS:NZ	1:A:212:LEU:O	2.46	0.48
1:A:220:PHE:HE1	1:A:285:ILE:HG22	1.79	0.47
1:A:231:ILE:HG22	1:A:233:ILE:HG23	1.97	0.47
1:A:31:SER:HA	1:A:216:LEU:HD22	1.98	0.45
1:A:131:CYS:HA	1:A:166:CYS:HB3	1.98	0.45
1:A:39:PRO:HG3	1:A:55:PHE:HZ	1.82	0.45
1:A:276:LEU:HB3	1:A:289:VAL:HG22	1.98	0.45
1:A:91:TYR:N	1:A:268:GLY:O	2.38	0.44
1:A:115:GLN:HB3	1:A:132:GLU:HG3	1.98	0.44
1:A:156:GLU:OE1	1:A:158:ARG:NH2	2.51	0.43
1:A:103:GLY:HA3	1:A:119:ILE:O	2.18	0.43
1:A:109:THR:HA	1:A:237:ARG:HH21	1.83	0.43
1:A:273:ARG:HH11	1:A:273:ARG:HG3	1.84	0.43
1:A:25:PRO:HA	1:A:26:PRO:HD3	1.97	0.41
1:A:93:ALA:HB3	1:A:266:TYR:HB2	2.01	0.41
1:A:102:ARG:HG3	1:A:141:LEU:HD22	2.02	0.41
1:A:34:ARG:NH1	1:A:219:GLY:O	2.54	0.41
1:A:102:ARG:HA	1:A:102:ARG:HD3	1.78	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
1	A	232/1269 (18%)	203 (88%)	28 (12%)	1 (0%)	34 68

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	79	PHE

### 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	217/1102 (20%)	216 (100%)	1 (0%)	<a href="#">88</a> <a href="#">94</a>

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	121	ASN

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.

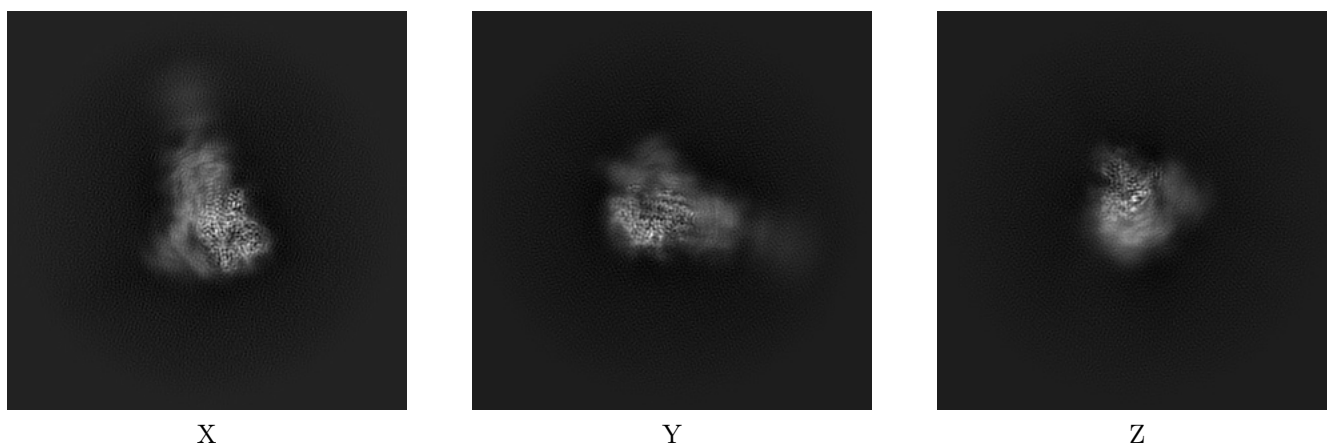
## 6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-31634. 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](#)

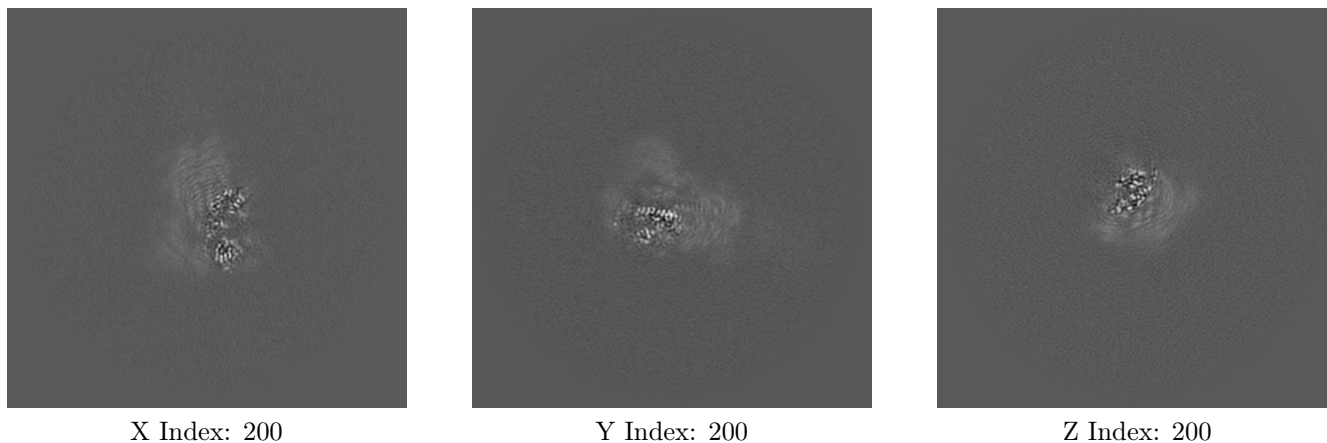
#### 6.1.1 Primary map



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

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

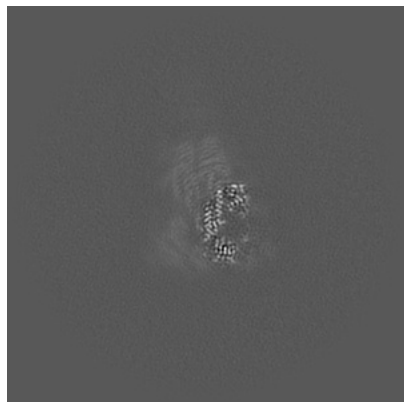
#### 6.2.1 Primary map



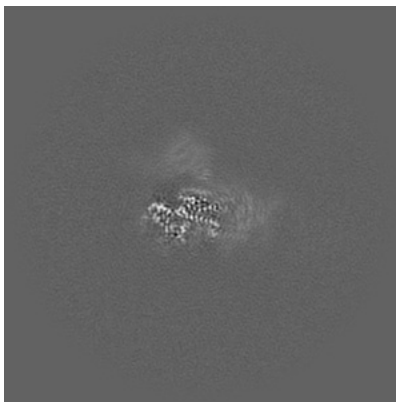
The images above show central slices of the map in three orthogonal directions.

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

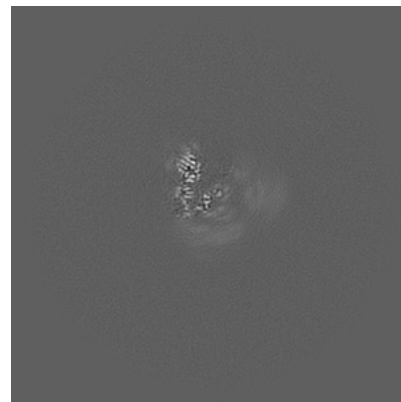
### 6.3.1 Primary map



X Index: 196



Y Index: 209

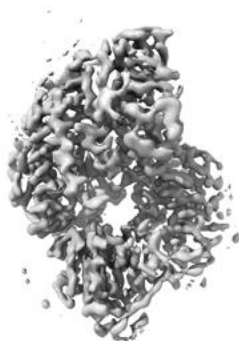


Z Index: 179

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

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

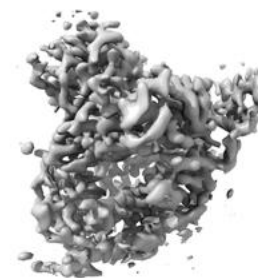
### 6.4.1 Primary map



X



Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.888. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

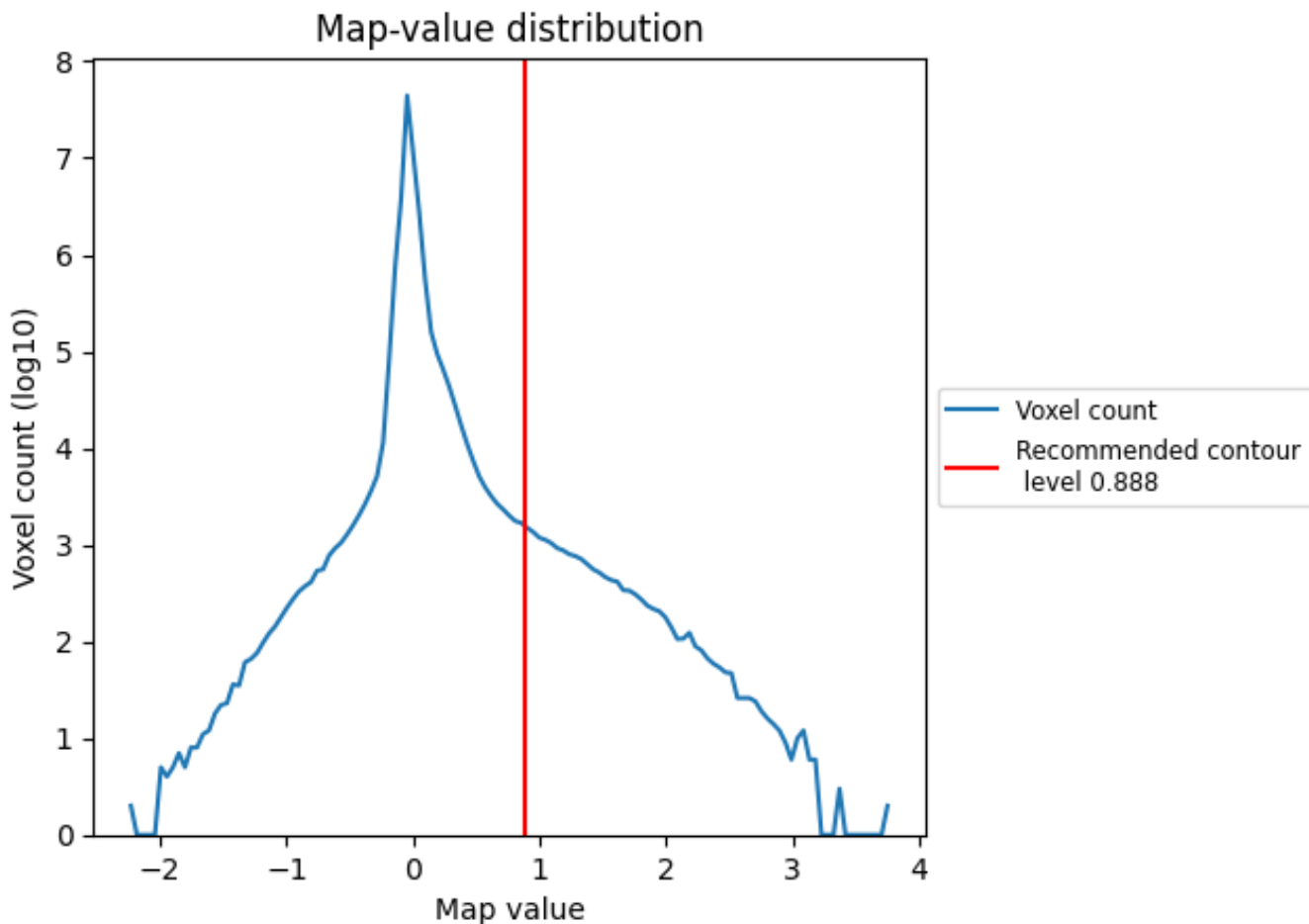
## 6.5 Mask visualisation

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.

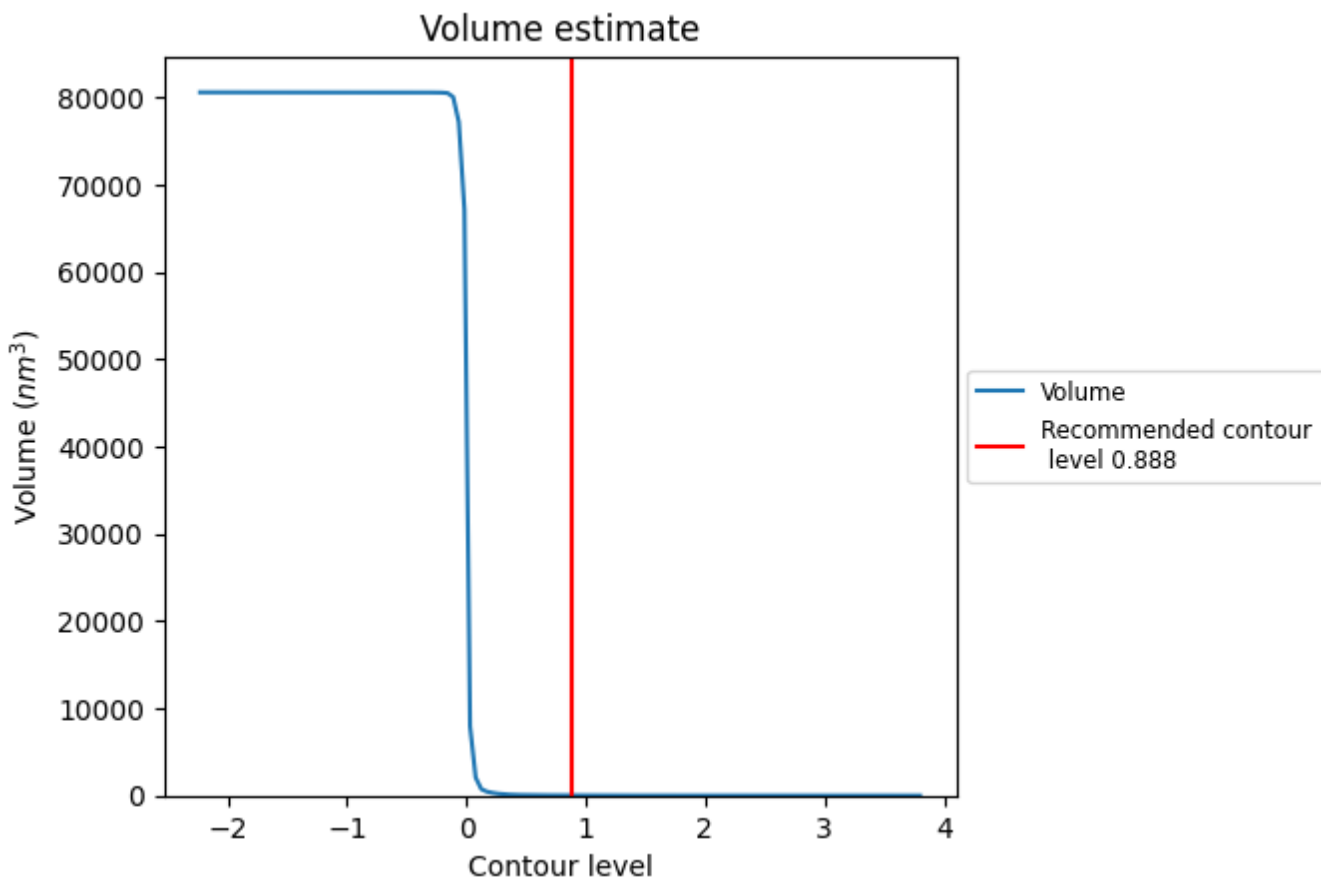
### 7.1 Map-value distribution [i](#)



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.



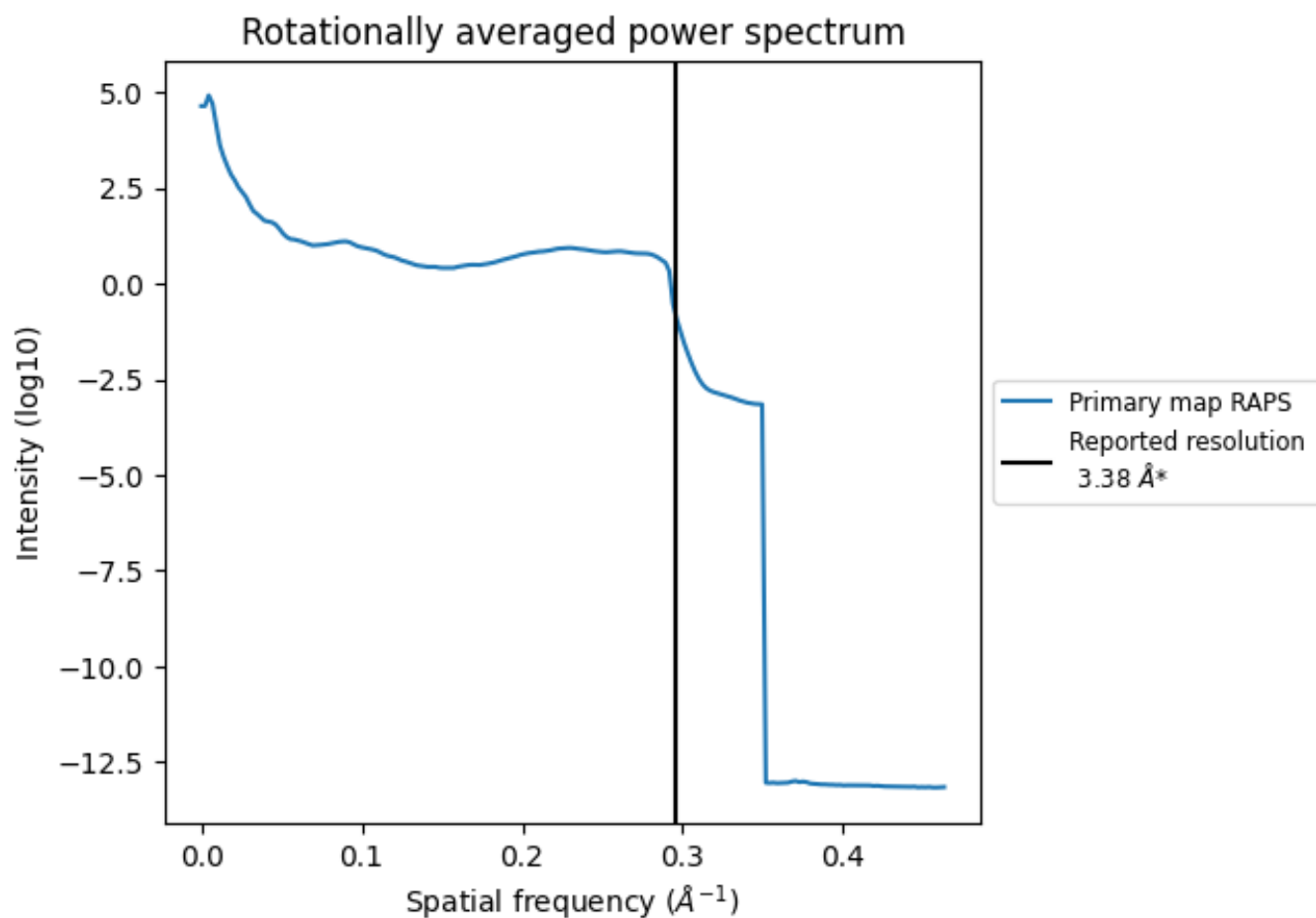
## 7.2 Volume estimate [i](#)



The volume at the recommended contour level is 21 nm<sup>3</sup>; this corresponds to an approximate mass of 19 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 0.296 Å<sup>-1</sup>

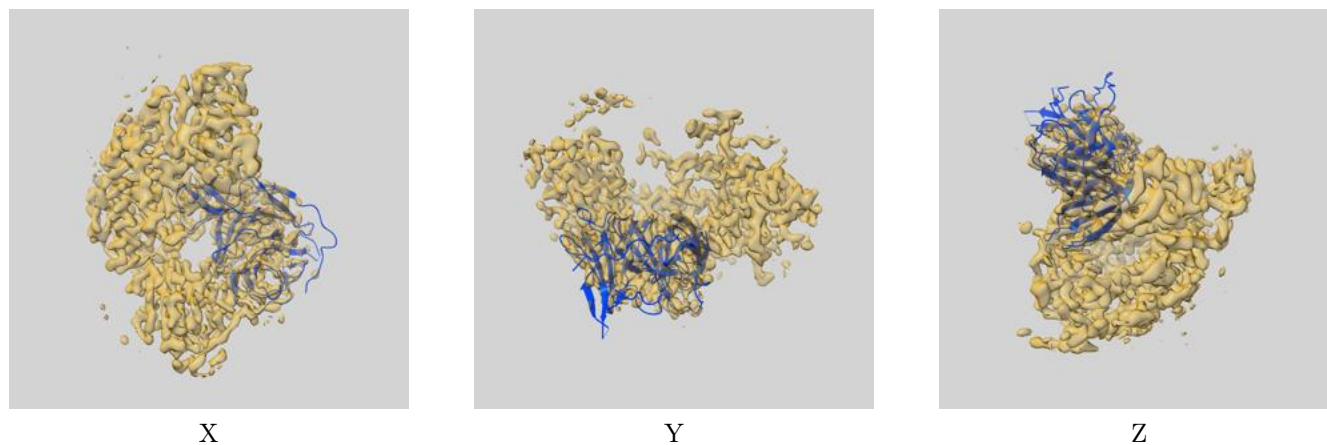
## 8 Fourier-Shell correlation

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

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

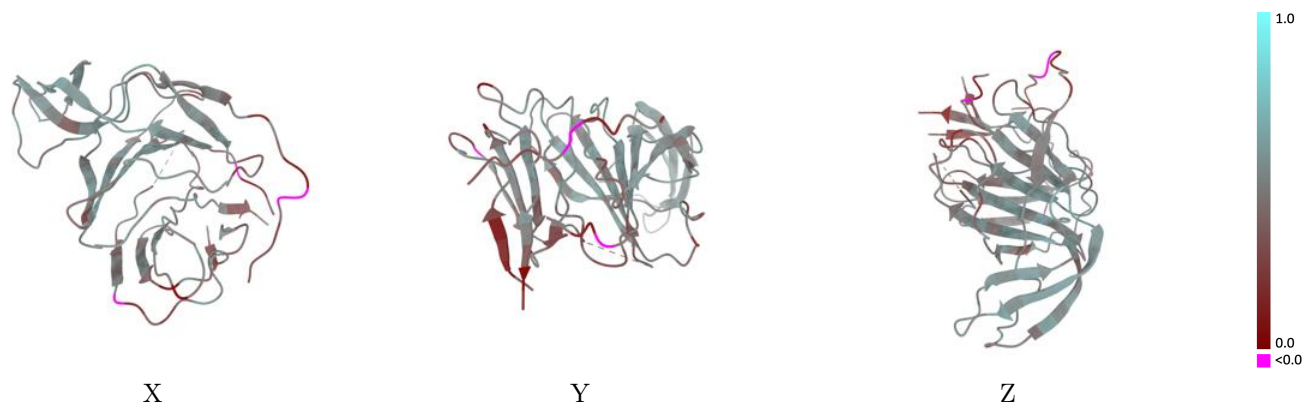
This section contains information regarding the fit between EMDB map EMD-31634 and PDB model 7V22. Per-residue inclusion information can be found in section 3 on page 6.

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



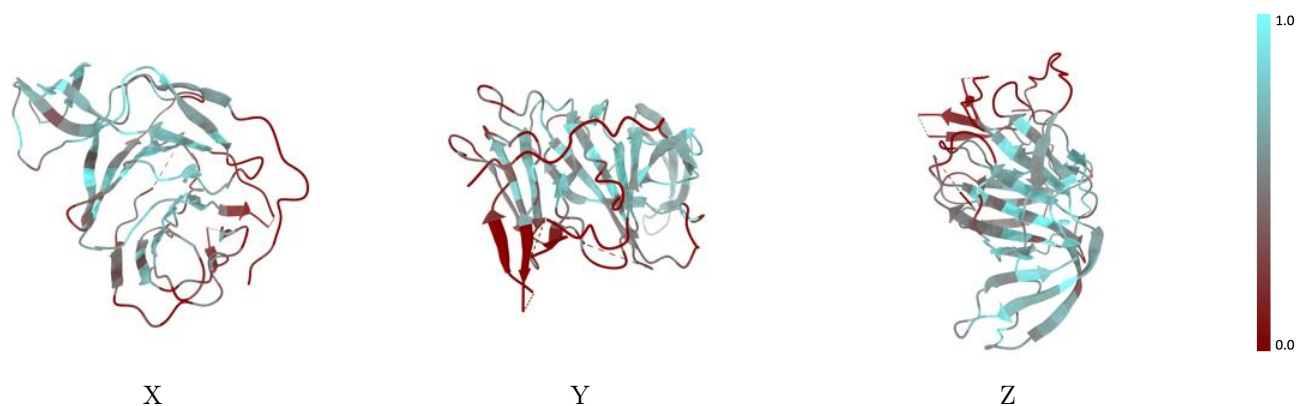
The images above show the 3D surface view of the map at the recommended contour level 0.888 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

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



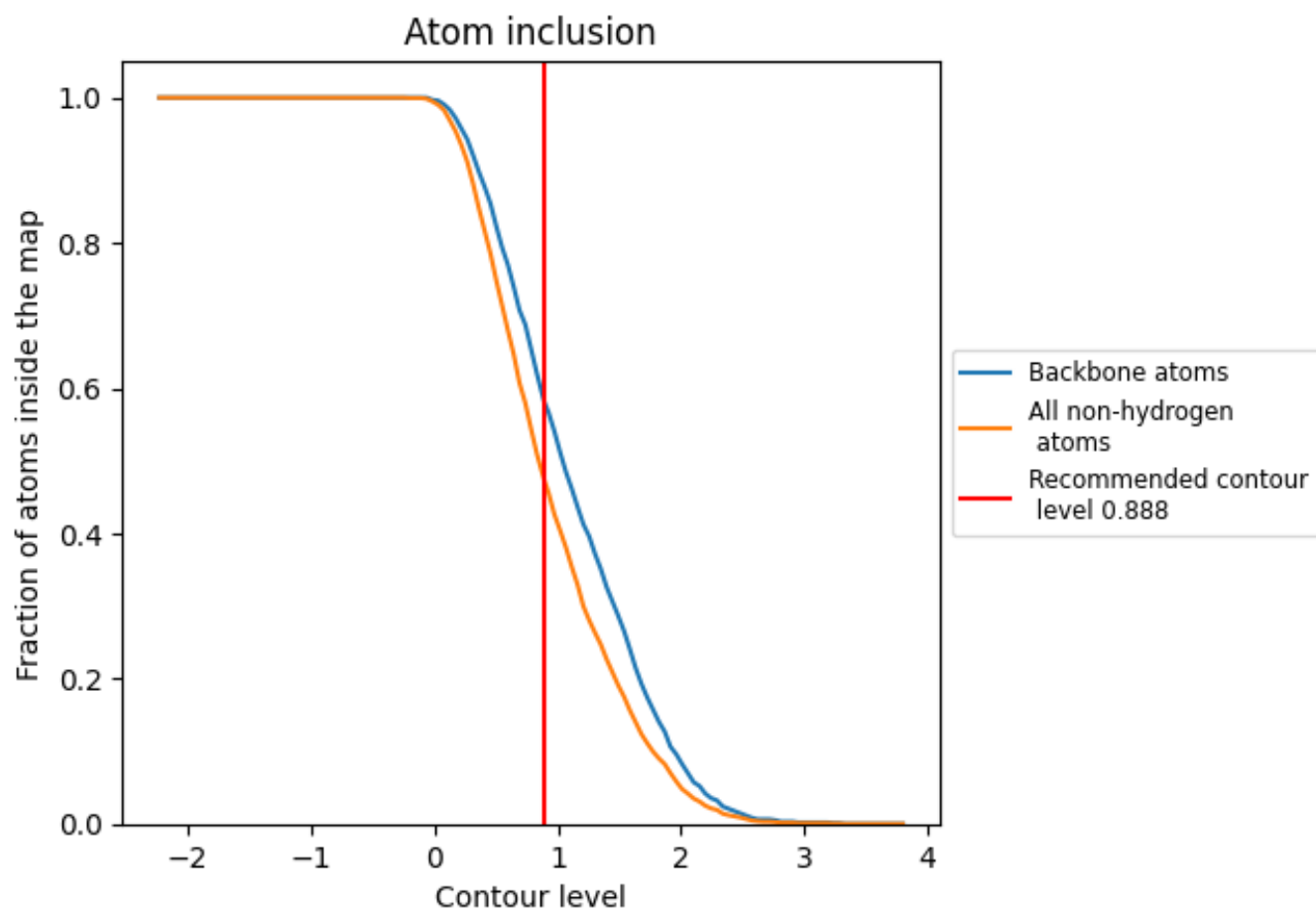
The images above show the model with each residue coloured according to its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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 (0.888).

## 9.4 Atom inclusion [i](#)



At the recommended contour level, 58% of all backbone atoms, 47% of all non-hydrogen atoms, are inside the map.

## 9.5 Map-model fit summary [i](#)

The table lists the average atom inclusion at the recommended contour level (0.888) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	■ 0.4747	■ 0.4520
A	■ 0.4747	■ 0.4520

