



wwPDB EM Validation Summary Report ⓘ

Nov 9, 2022 – 01:39 AM JST

PDB ID : 6ACJ
EMDB ID : EMD-9593
Title : Trypsin-cleaved and low pH-treated SARS-CoV spike glycoprotein and ACE2 complex, ACE2-bound conformation 2
Authors : Gui, M.; Song, W.
Deposited on : 2018-07-26
Resolution : 4.20 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

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) : 2.31.2

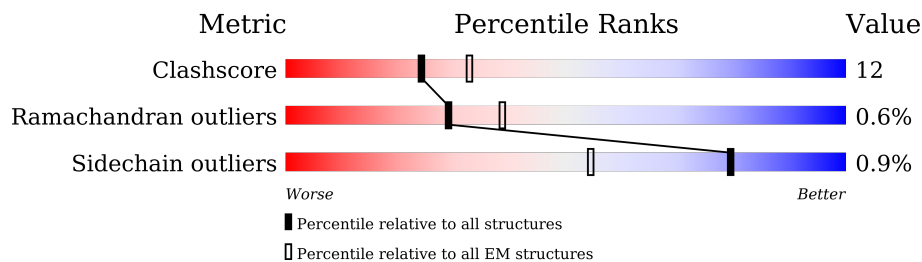
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 4.20 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

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 ≥ 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	1203	
1	B	1203	
1	C	1203	
2	D	603	

2 Entry composition i

There are 2 unique types of molecules in this entry. The entry contains 29715 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	1065	8302	5304	1374	1579	45	0	0
1	B	1065	8302	5304	1374	1579	45	0	0
1	C	1057	8241	5264	1364	1568	45	0	0

There are 21 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1197	SER	-	expression tag	UNP P59594
A	1198	HIS	-	expression tag	UNP P59594
A	1199	PRO	-	expression tag	UNP P59594
A	1200	GLN	-	expression tag	UNP P59594
A	1201	PHE	-	expression tag	UNP P59594
A	1202	GLU	-	expression tag	UNP P59594
A	1203	LYS	-	expression tag	UNP P59594
B	1197	SER	-	expression tag	UNP P59594
B	1198	HIS	-	expression tag	UNP P59594
B	1199	PRO	-	expression tag	UNP P59594
B	1200	GLN	-	expression tag	UNP P59594
B	1201	PHE	-	expression tag	UNP P59594
B	1202	GLU	-	expression tag	UNP P59594
B	1203	LYS	-	expression tag	UNP P59594
C	1197	SER	-	expression tag	UNP P59594
C	1198	HIS	-	expression tag	UNP P59594
C	1199	PRO	-	expression tag	UNP P59594
C	1200	GLN	-	expression tag	UNP P59594
C	1201	PHE	-	expression tag	UNP P59594
C	1202	GLU	-	expression tag	UNP P59594
C	1203	LYS	-	expression tag	UNP P59594

- Molecule 2 is a protein called Angiotensin-converting enzyme 2.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	D	597	4870	3115	806	920	29	0	0

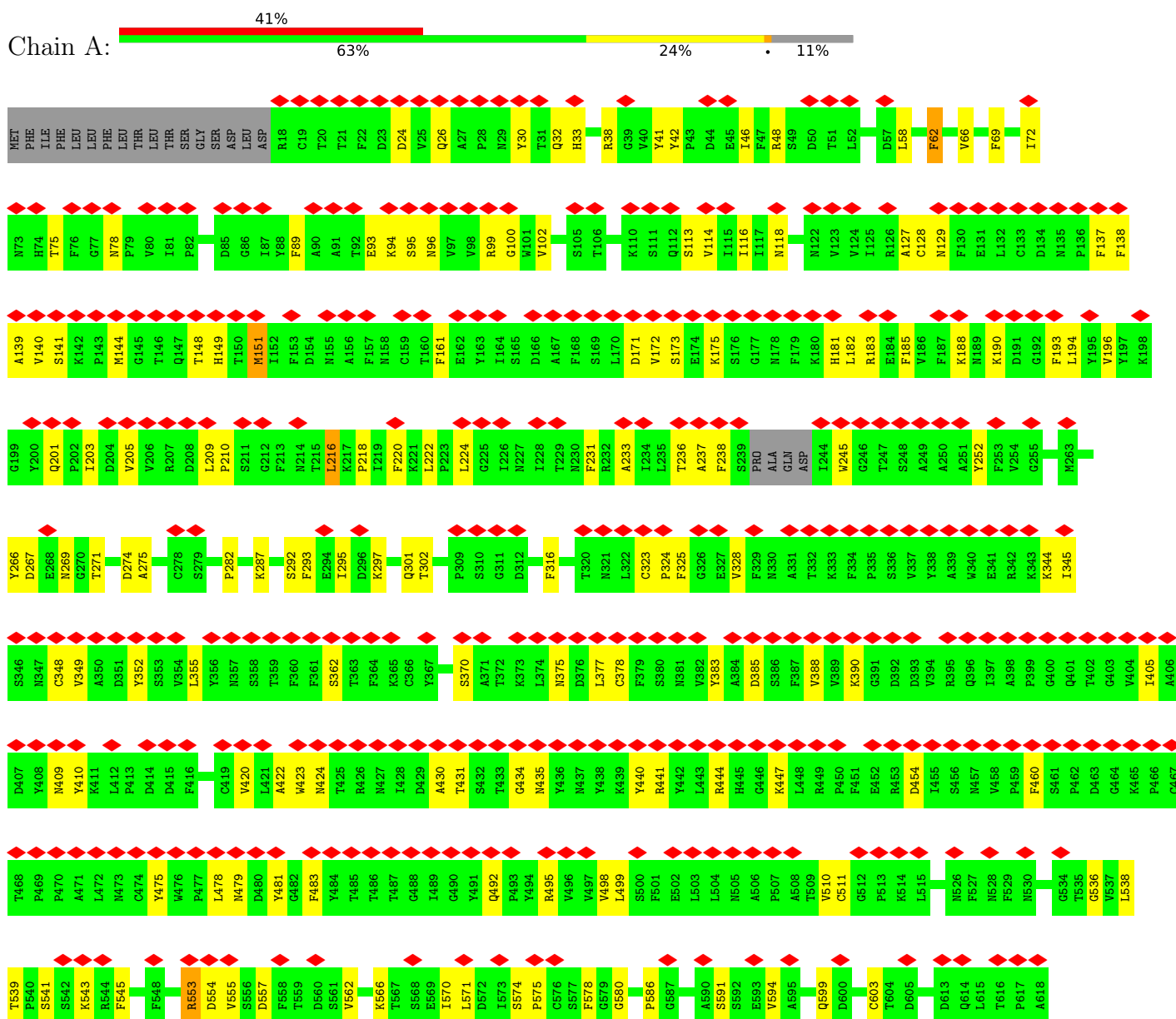
There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	616	HIS	-	expression tag	UNP Q9BYF1
D	617	HIS	-	expression tag	UNP Q9BYF1
D	618	HIS	-	expression tag	UNP Q9BYF1
D	619	HIS	-	expression tag	UNP Q9BYF1
D	620	HIS	-	expression tag	UNP Q9BYF1
D	621	HIS	-	expression tag	UNP Q9BYF1

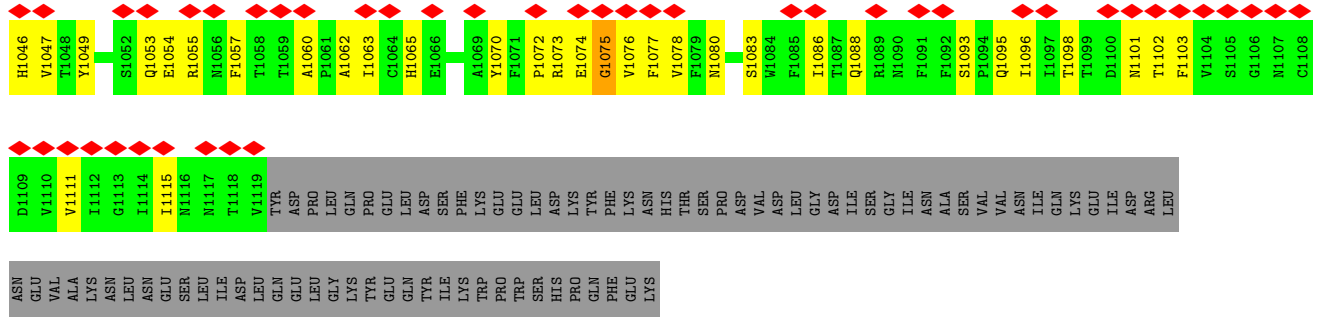
3 Residue-property plots

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.

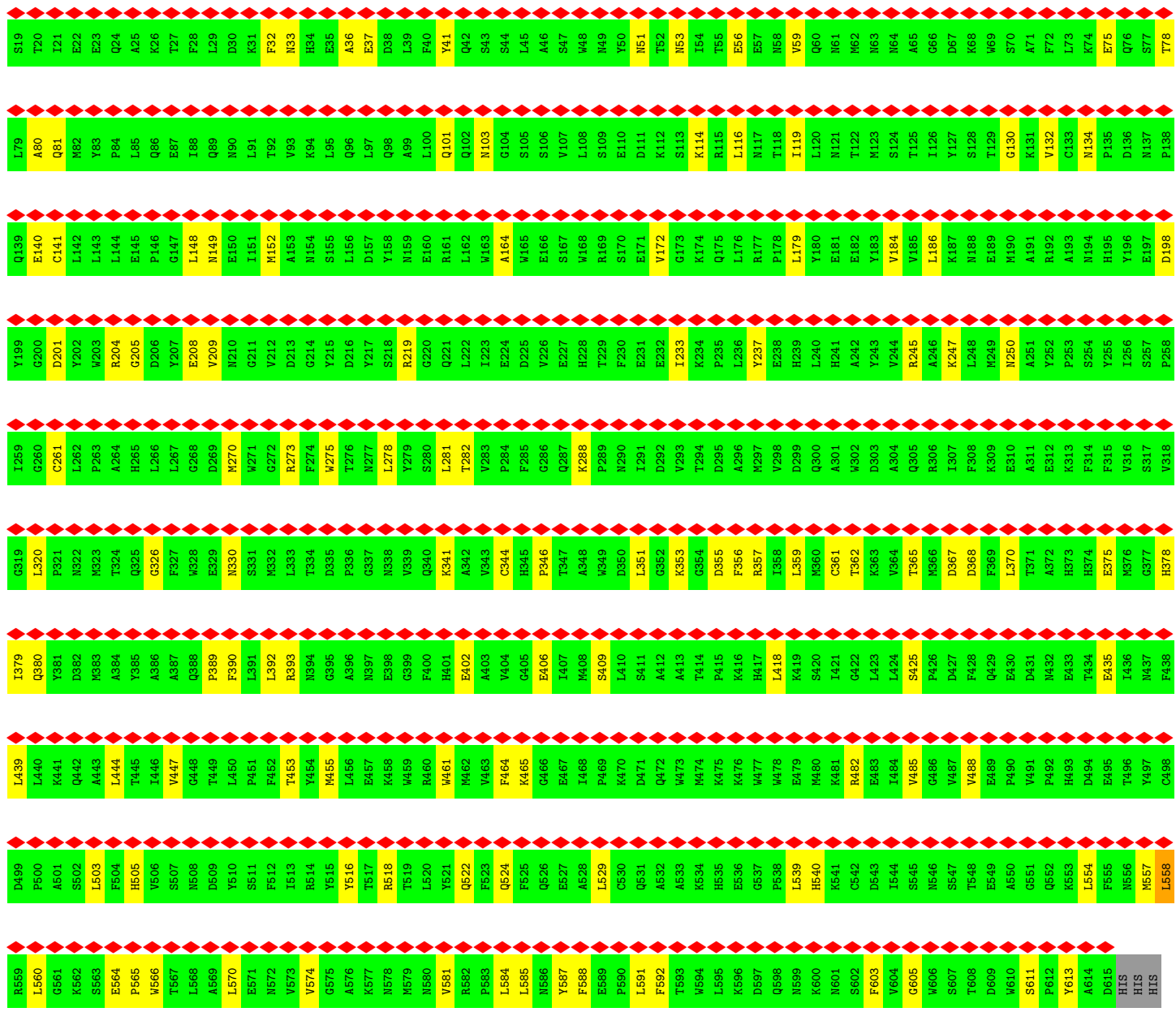
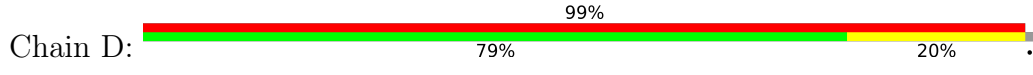
• Molecule 1: Spike glycoprotein



Validation summary table for EMD-9593, 6ACJ. The table consists of 11 columns, each representing a different validation metric. Each row represents a specific residue, identified by its PDB ID (e.g., G68, F69, H70, etc.). The cells in the table are color-coded: green for good, yellow for warning, orange for outlier, and red for failure. Red diamonds above the residue IDs indicate specific validation issues. The table lists residues from G68 at the top to N837 at the bottom, with some residues missing from the sequence.



● Molecule 2: Angiotensin-converting enzyme 2



HIS
HIS
HIS

4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	129462	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 K2 SUMMIT (4k x 4k)	Depositor
Maximum map value	33.479	Depositor
Minimum map value	-16.130	Depositor
Average map value	-0.001	Depositor
Map value standard deviation	0.871	Depositor
Recommended contour level	8.0	Depositor
Map size (Å)	380.16, 380.16, 380.16	wwPDB
Map dimensions	288, 288, 288	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.32, 1.32, 1.32	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	0.50	1/8499 (0.0%)	0.70	3/11568 (0.0%)
1	B	0.50	1/8499 (0.0%)	0.70	1/11568 (0.0%)
1	C	0.55	8/8435 (0.1%)	0.77	10/11477 (0.1%)
2	D	0.31	0/5007	0.56	1/6803 (0.0%)
All	All	0.49	10/30440 (0.0%)	0.70	15/41416 (0.0%)

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	A	0	14
1	B	0	19
1	C	0	21
2	D	0	1
All	All	0	55

The worst 5 of 10 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	677	TYR	CE2-CZ	-8.20	1.27	1.38
1	A	725	CYS	CB-SG	-6.94	1.70	1.82
1	C	678	THR	CA-C	-5.93	1.37	1.52
1	C	676	ALA	C-O	-5.67	1.12	1.23
1	C	646	TYR	CE1-CZ	-5.53	1.31	1.38

The worst 5 of 15 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	642	VAL	CB-CA-C	-7.37	97.40	111.40
1	C	647	GLU	C-N-CA	-7.22	103.65	121.70
1	B	557	ASP	CB-CG-OD1	7.20	124.78	118.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	641	HIS	O-C-N	6.86	133.67	122.70
1	C	641	HIS	CA-C-N	-6.62	102.64	117.20

There are no chirality outliers.

5 of 55 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	151	MET	Peptide
1	A	553	ARG	Peptide
1	A	62	PHE	Peptide
1	A	631	THR	Peptide
1	A	633	ALA	Peptide

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	8302	0	8082	200	0
1	B	8302	0	8082	210	0
1	C	8241	0	8011	259	0
2	D	4870	0	4643	73	0
All	All	29715	0	28818	697	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 697 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:C:646:TYR:HB2	1:C:677:TYR:CE2	1.33	1.59
1:C:646:TYR:CB	1:C:677:TYR:CE2	1.90	1.53
1:C:647:GLU:O	1:C:648:CYS:CB	1.77	1.26
1:C:646:TYR:HB2	1:C:677:TYR:CD2	1.78	1.17
1:C:646:TYR:CA	1:C:677:TYR:CE2	2.28	1.16

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	1057/1203 (88%)	853 (81%)	199 (19%)	5 (0%)	29	68
1	B	1057/1203 (88%)	848 (80%)	204 (19%)	5 (0%)	29	68
1	C	1045/1203 (87%)	840 (80%)	193 (18%)	12 (1%)	14	52
2	D	595/603 (99%)	562 (94%)	33 (6%)	0	100	100
All	All	3754/4212 (89%)	3103 (83%)	629 (17%)	22 (1%)	29	64

5 of 22 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	632	GLN
1	C	648	CYS
1	C	676	ALA
1	A	632	GLN
1	A	1072	PRO

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	922/1048 (88%)	917 (100%)	5 (0%)	88	93
1	B	922/1048 (88%)	915 (99%)	7 (1%)	81	89
1	C	914/1048 (87%)	902 (99%)	12 (1%)	69	82
2	D	527/533 (99%)	521 (99%)	6 (1%)	73	84
All	All	3285/3677 (89%)	3255 (99%)	30 (1%)	79	87

5 of 30 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	426	ARG
2	D	288	LYS
1	C	647	GLU
2	D	455	MET
2	D	53	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 61 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	835	GLN
2	D	210	ASN
1	C	135	ASN
2	D	53	ASN
2	D	526	GLN

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

There are no chain breaks in this entry.

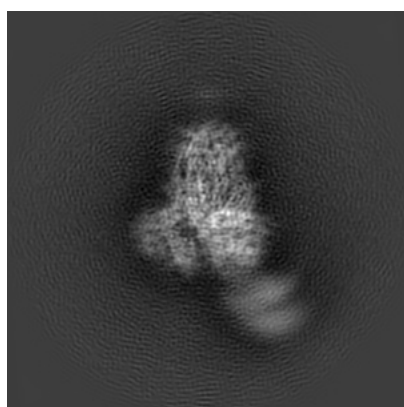
6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-9593. These allow visual inspection of the internal detail of the map and identification of artifacts.

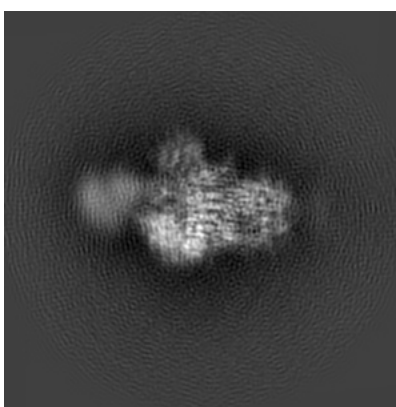
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections [i](#)

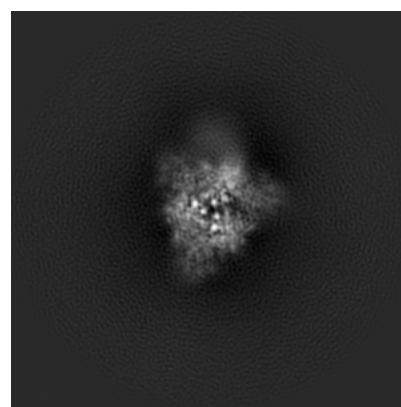
6.1.1 Primary map



X



Y

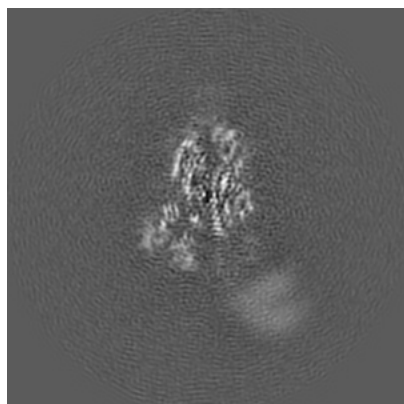


Z

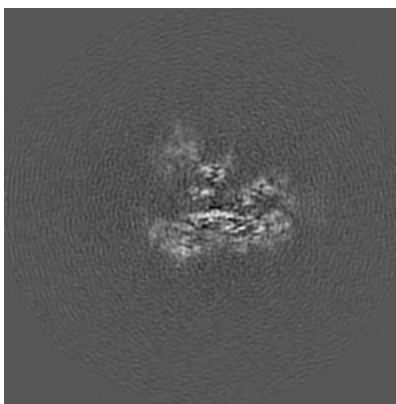
The images above show the map projected in three orthogonal directions.

6.2 Central slices [i](#)

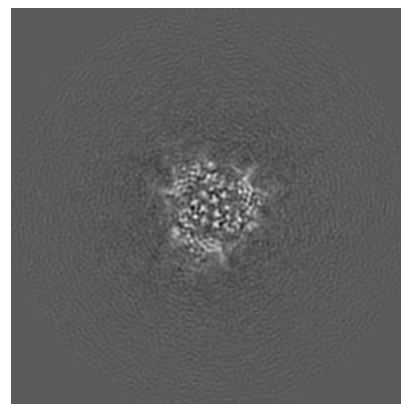
6.2.1 Primary map



X Index: 144



Y Index: 144

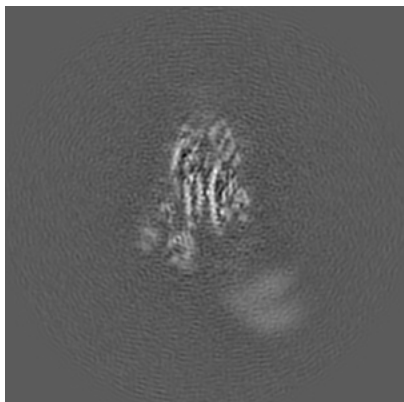


Z Index: 144

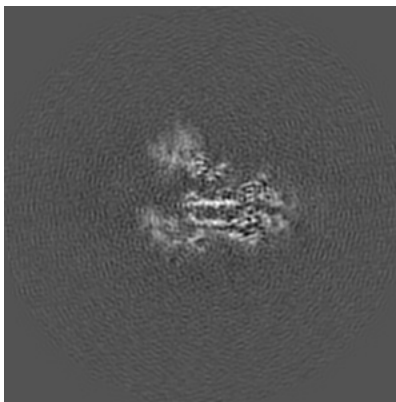
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices [i](#)

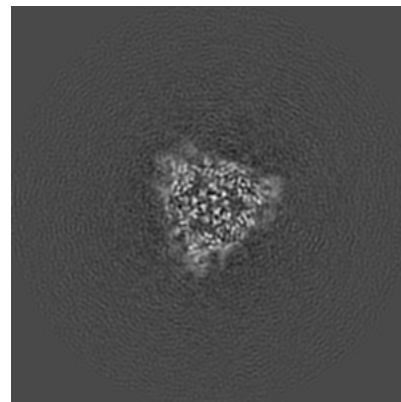
6.3.1 Primary map



X Index: 146



Y Index: 149



Z Index: 141

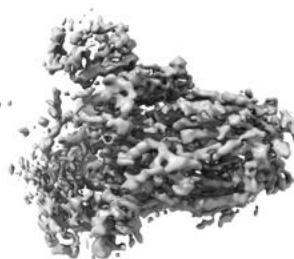
The images above show the largest variance slices of the map in three orthogonal directions.

6.4 Orthogonal surface views [i](#)

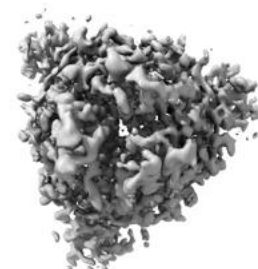
6.4.1 Primary map



X



Y



Z

The images above show the 3D surface view of the map at the recommended contour level 8.0. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

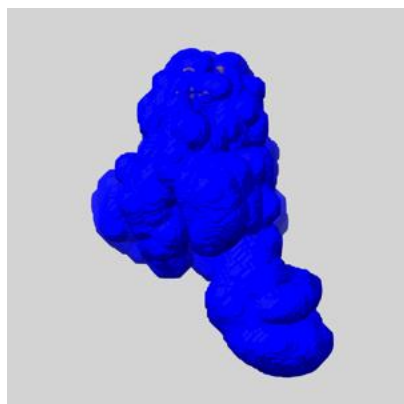
6.5 Mask visualisation [i](#)

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

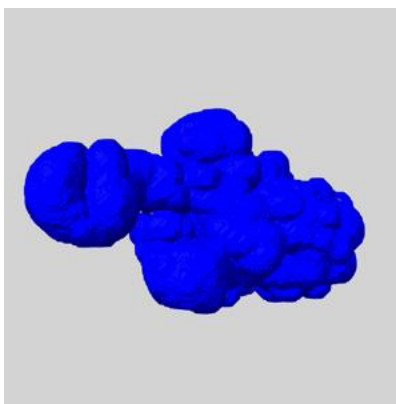
A mask typically either:

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

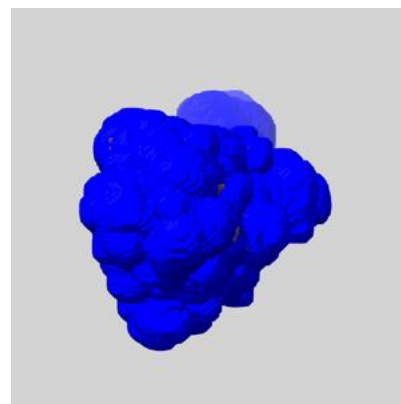
6.5.1 emd_9593_msk_1.map [i](#)



X



Y

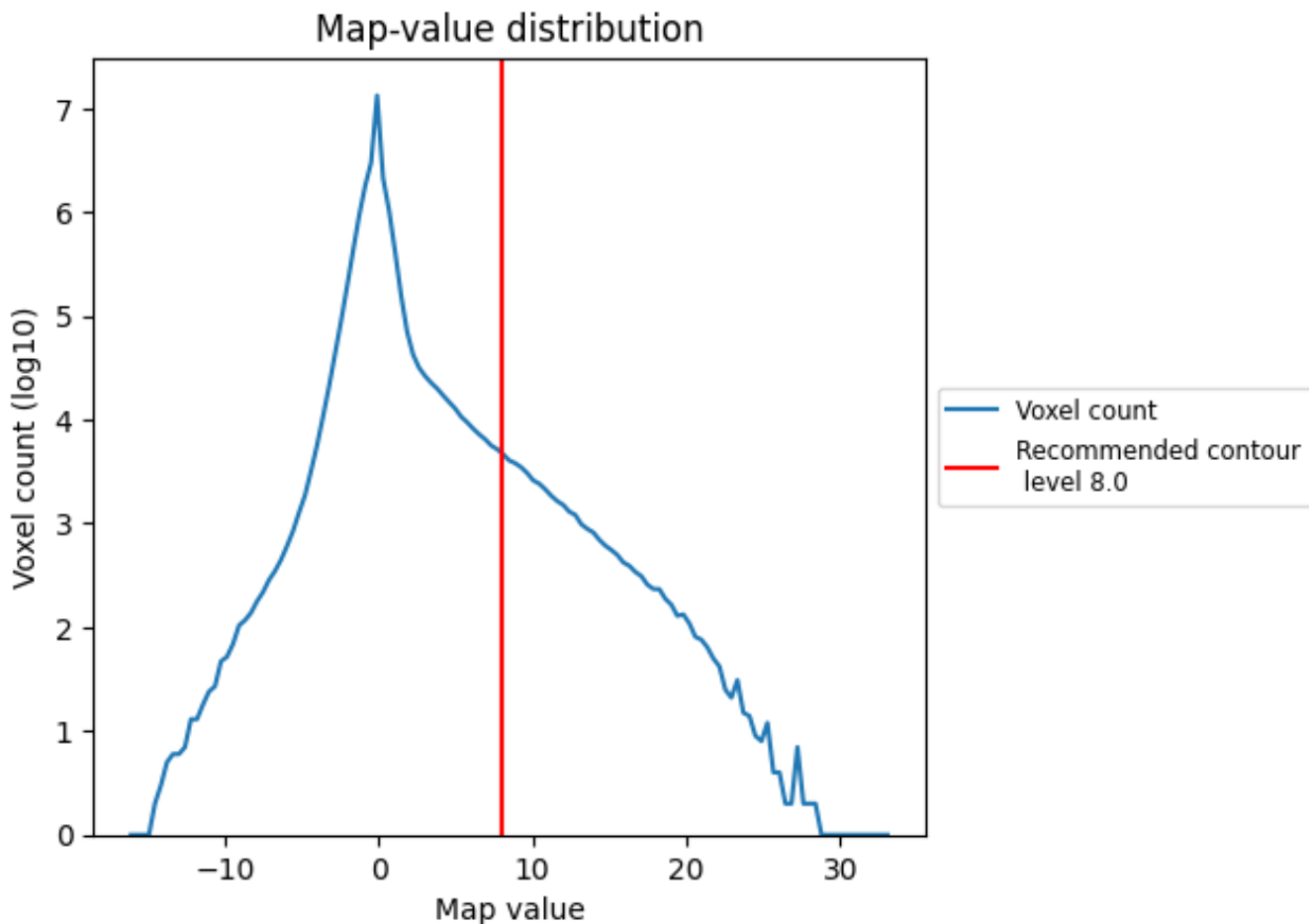


Z

7 Map analysis [i](#)

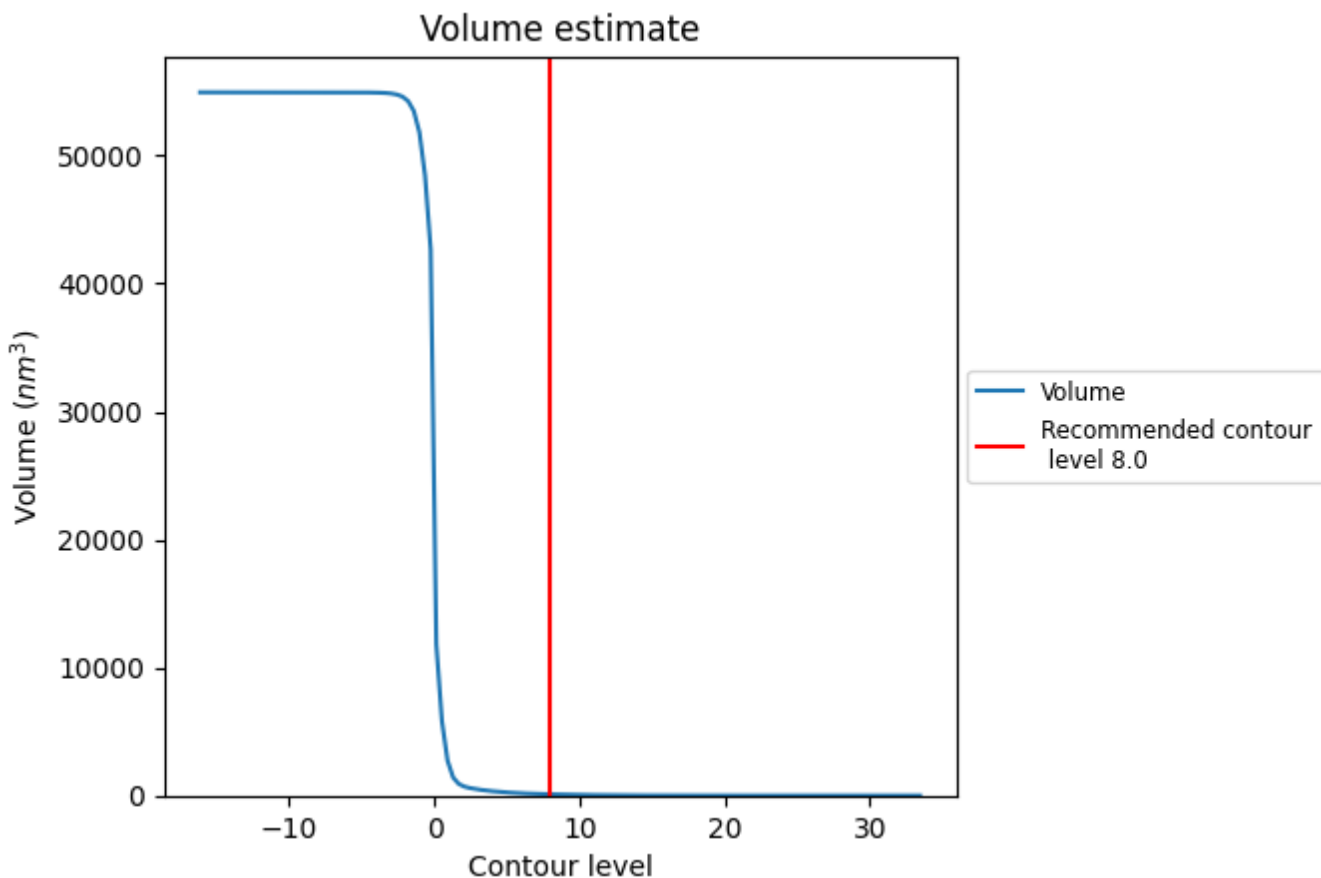
This section contains the results of statistical analysis of the map.

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.

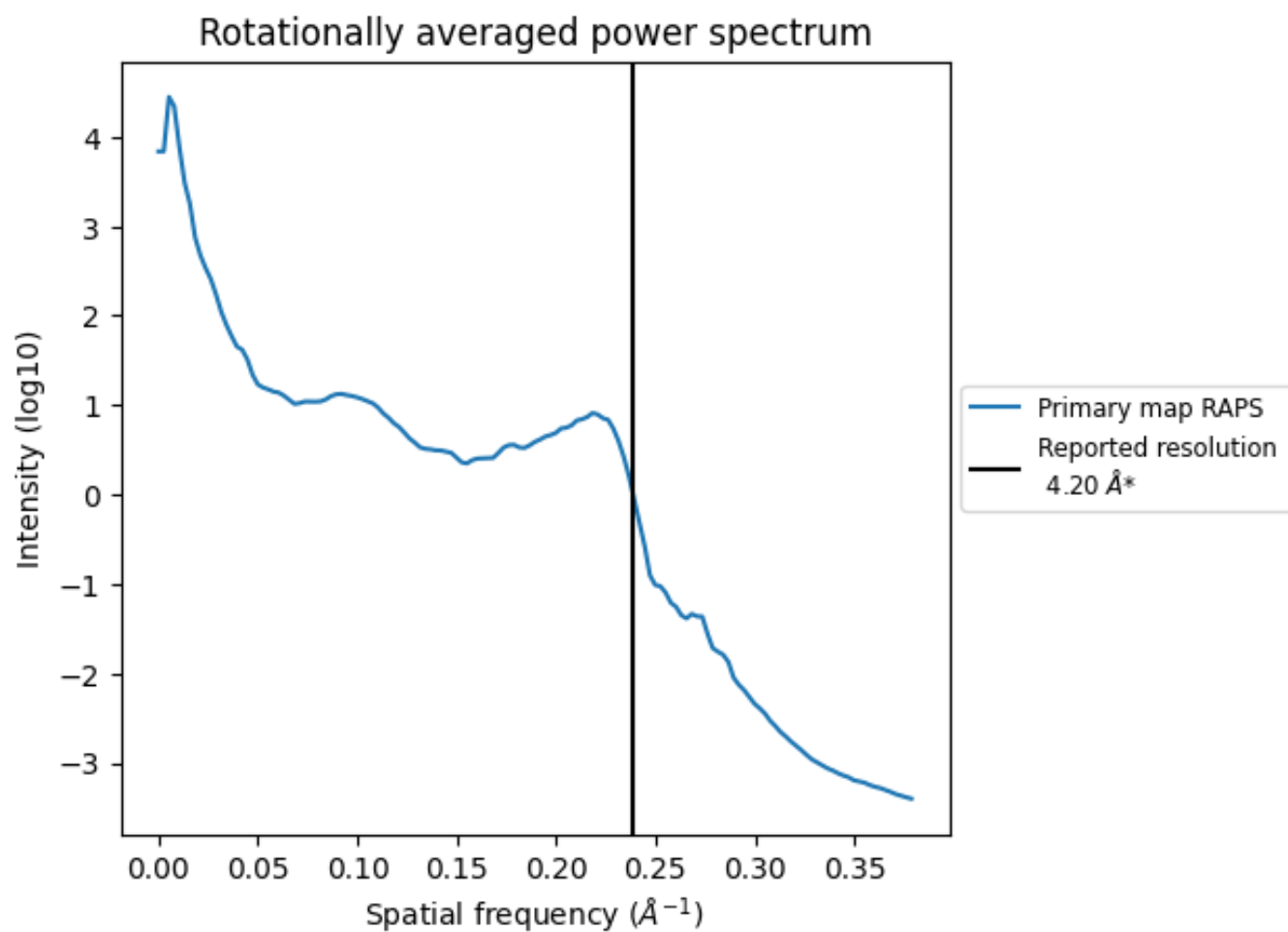
7.2 Volume estimate [\(i\)](#)



The volume at the recommended contour level is 100 nm^3 ; this corresponds to an approximate mass of 90 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](#)

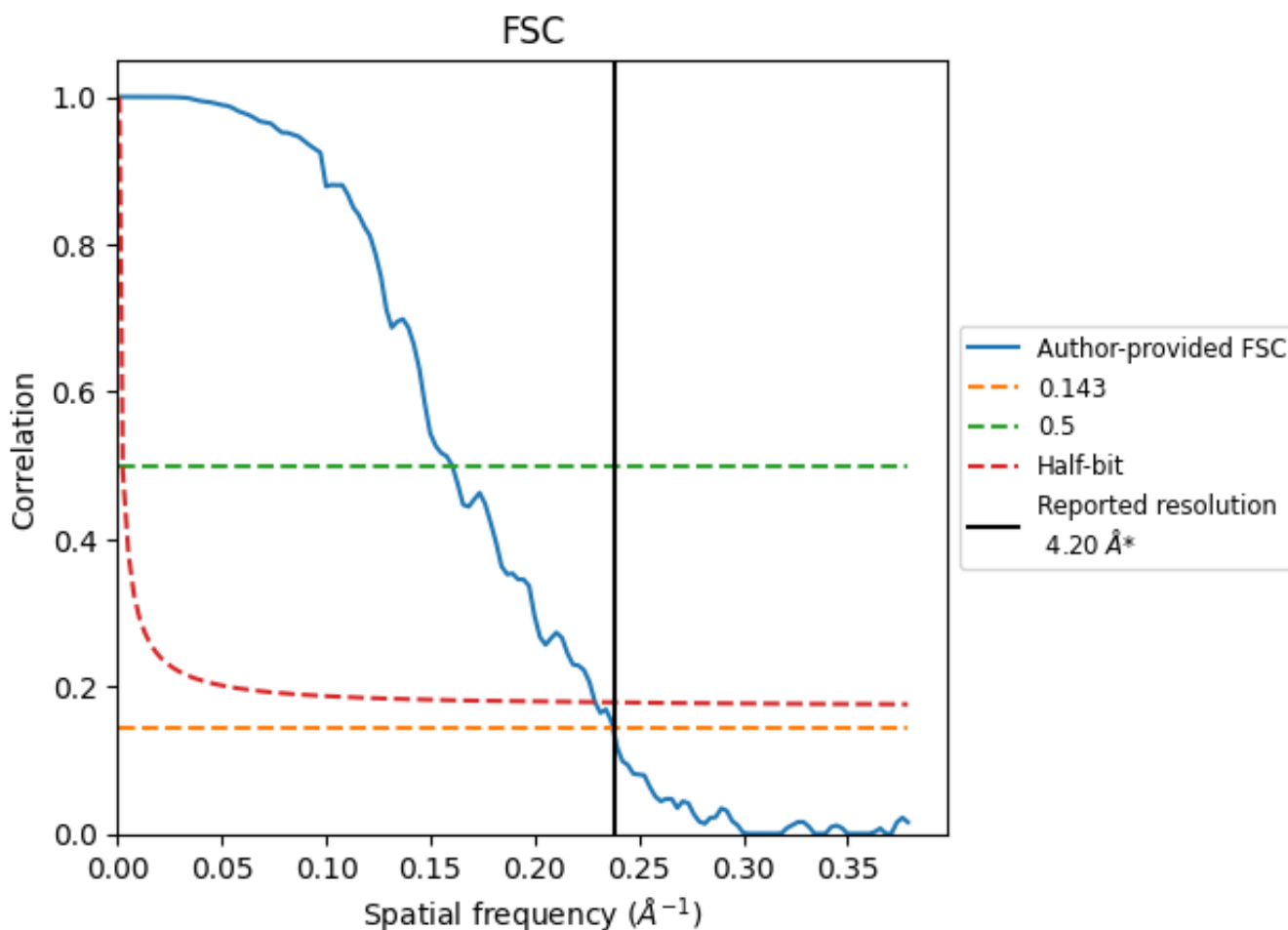


*Reported resolution corresponds to spatial frequency of 0.238\AA^{-1}

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 0.238 Å⁻¹

8.2 Resolution estimates [i](#)

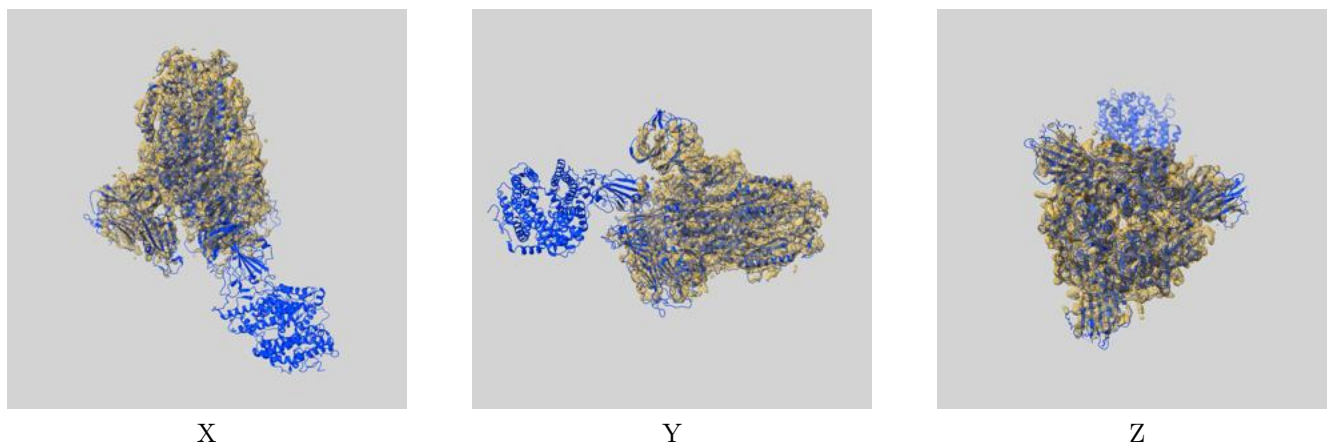
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	4.20	-	-
Author-provided FSC curve	4.21	6.24	4.37
Unmasked-calculated*	-	-	-

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

9 Map-model fit [i](#)

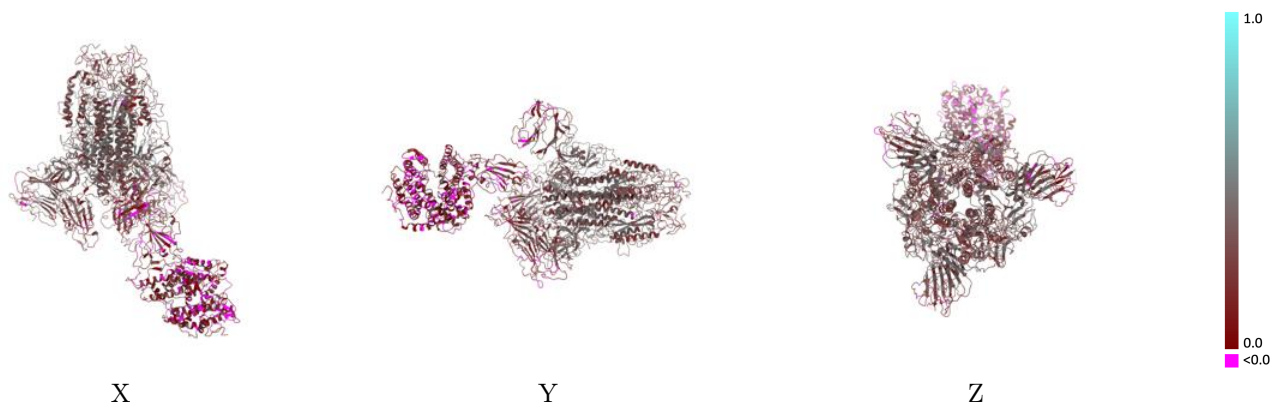
This section contains information regarding the fit between EMDB map EMD-9593 and PDB model 6ACJ. Per-residue inclusion information can be found in section 3 on page 5.

9.1 Map-model overlay [i](#)



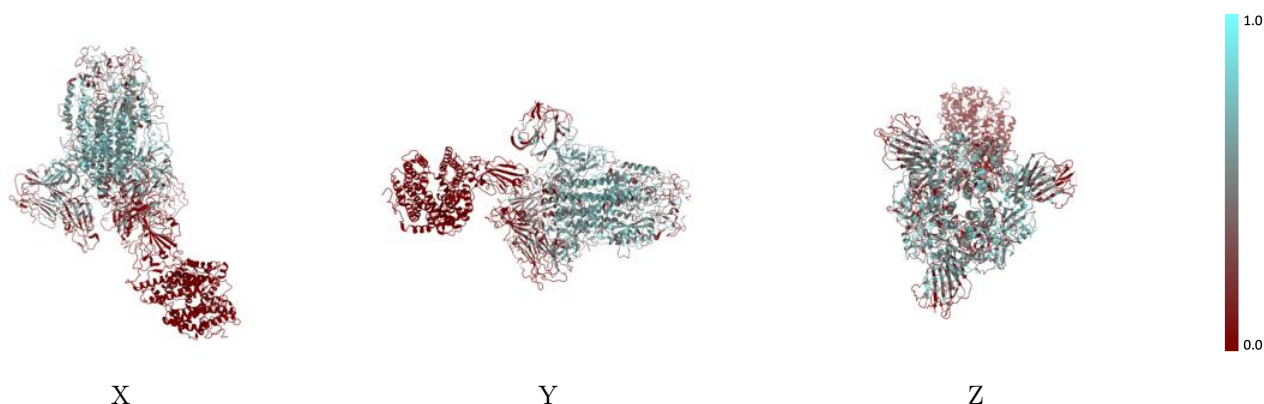
The images above show the 3D surface view of the map at the recommended contour level 8.0 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

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



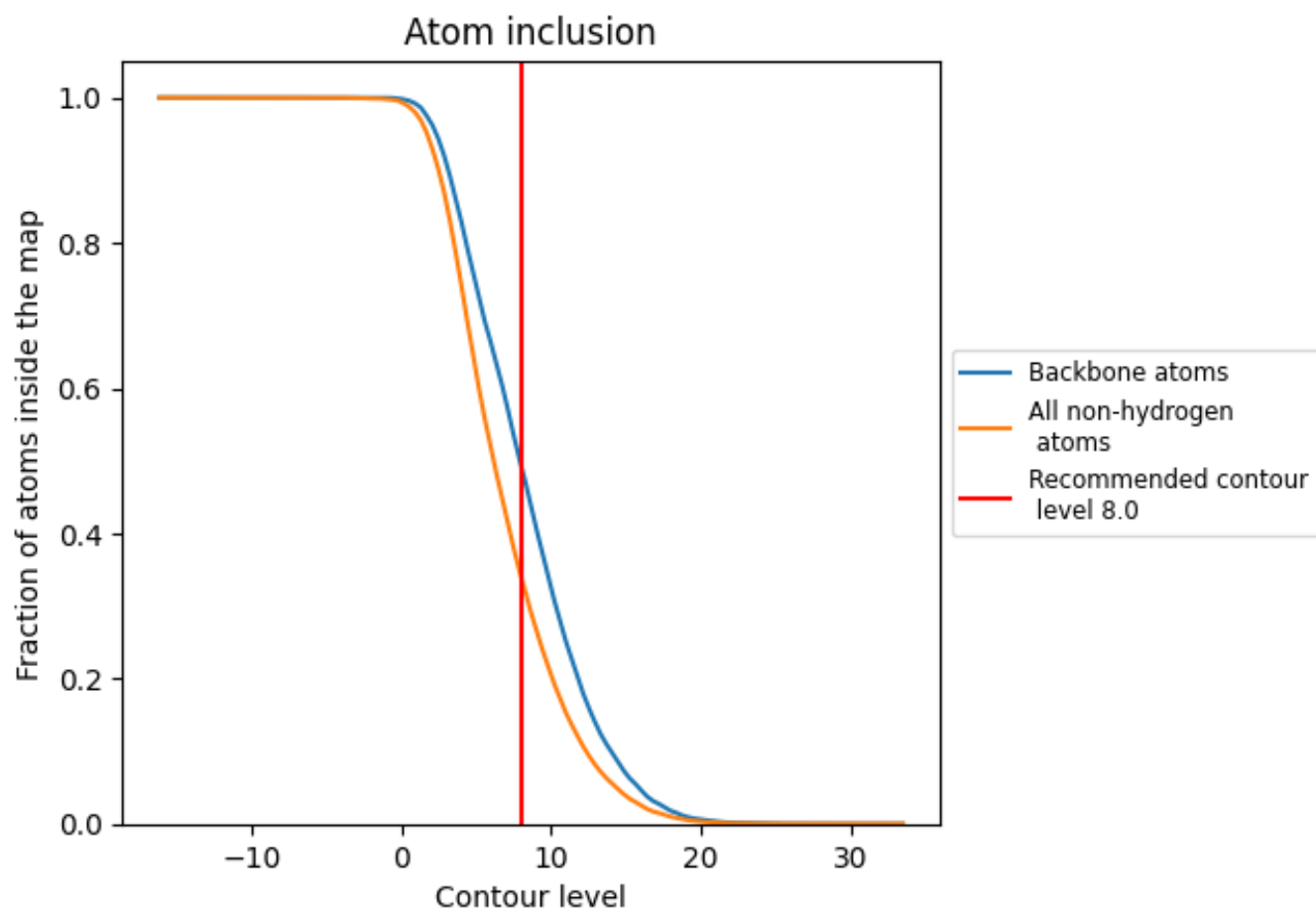
The images above show the model with each residue coloured according to its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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 (8.0).











9.4 Atom inclusion [i](#)



At the recommended contour level, 49% of all backbone atoms, 34% 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 (8.0) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	 0.3399	 0.2770
A	 0.4188	 0.3200
B	 0.4253	 0.3080
C	 0.3748	 0.2960
D	 0.0000	 0.1150

