



## wwPDB EM Validation Summary Report ⓘ

Feb 20, 2023 – 04:34 PM JST

PDB ID : 7WSC  
EMDB ID : EMD-32753  
Title : Local structure of BD55-3500 and omicron RBD complex  
Authors : Liu, P.L.  
Deposited on : 2022-01-28  
Resolution : 3.78 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:

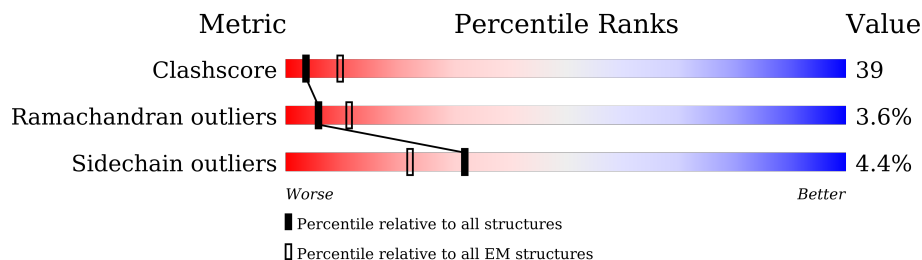
EMDB validation analysis : 0.0.1.dev43  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
MolProbity : 4.02b-467  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
MapQ : 1.9.9  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.32.1

# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:  
*ELECTRON MICROSCOPY*

The reported resolution of this entry is 3.78 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	R	236	
2	H	127	
3	L	108	
4	A	2	

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 3369 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 protein S1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	R	187	1504	969	254	273	8	0	0

There are 28 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
R	314	ALA	-	expression tag	UNP P0DTC2
R	315	ALA	-	expression tag	UNP P0DTC2
R	316	GLY	-	expression tag	UNP P0DTC2
R	317	SER	-	expression tag	UNP P0DTC2
R	318	ALA	-	expression tag	UNP P0DTC2
R	339	ASP	GLY	variant	UNP P0DTC2
R	371	LEU	SER	variant	UNP P0DTC2
R	373	PRO	SER	variant	UNP P0DTC2
R	375	PHE	SER	variant	UNP P0DTC2
R	417	ASN	LYS	variant	UNP P0DTC2
R	440	LYS	ASN	variant	UNP P0DTC2
R	446	SER	GLY	variant	UNP P0DTC2
R	477	ASN	SER	variant	UNP P0DTC2
R	478	LYS	THR	variant	UNP P0DTC2
R	484	ALA	GLU	variant	UNP P0DTC2
R	493	ARG	GLN	variant	UNP P0DTC2
R	496	SER	GLY	variant	UNP P0DTC2
R	498	ARG	GLN	variant	UNP P0DTC2
R	501	TYR	ASN	variant	UNP P0DTC2
R	505	HIS	TYR	variant	UNP P0DTC2
R	542	HIS	-	expression tag	UNP P0DTC2
R	543	HIS	-	expression tag	UNP P0DTC2
R	544	HIS	-	expression tag	UNP P0DTC2
R	545	HIS	-	expression tag	UNP P0DTC2
R	546	HIS	-	expression tag	UNP P0DTC2
R	547	HIS	-	expression tag	UNP P0DTC2
R	548	HIS	-	expression tag	UNP P0DTC2
R	549	HIS	-	expression tag	UNP P0DTC2

- Molecule 2 is a protein called 3500H.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	H	127	1024	658	172	191	3	0	0

- Molecule 3 is a protein called 3500L.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	L	108	813	513	135	161	4	0	0

- Molecule 4 is an oligosaccharide called 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.

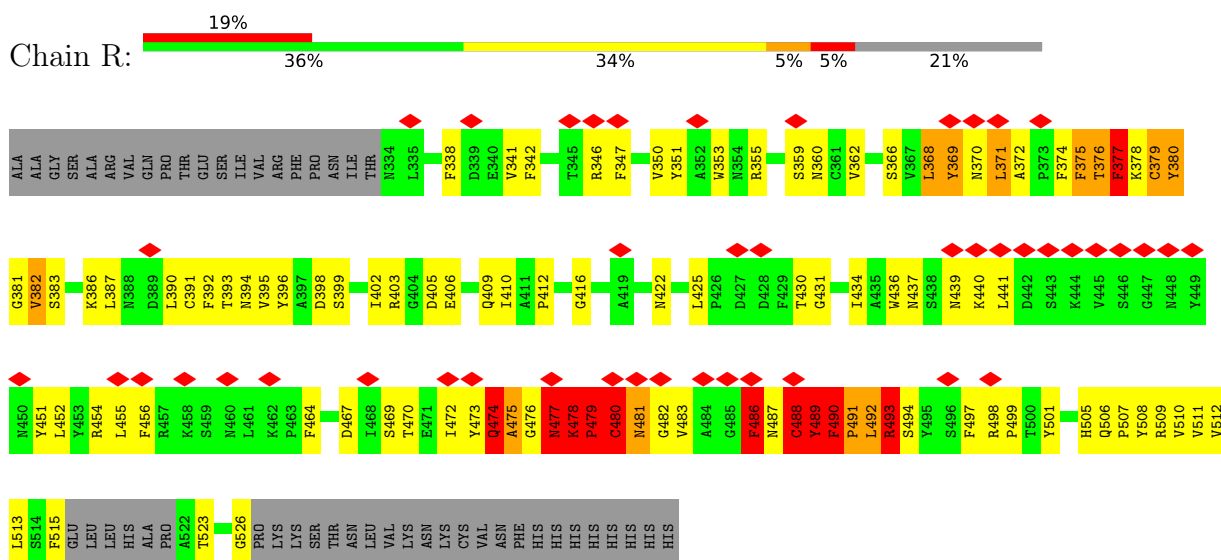


Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
4	A	2	28	16	2	10	0	0

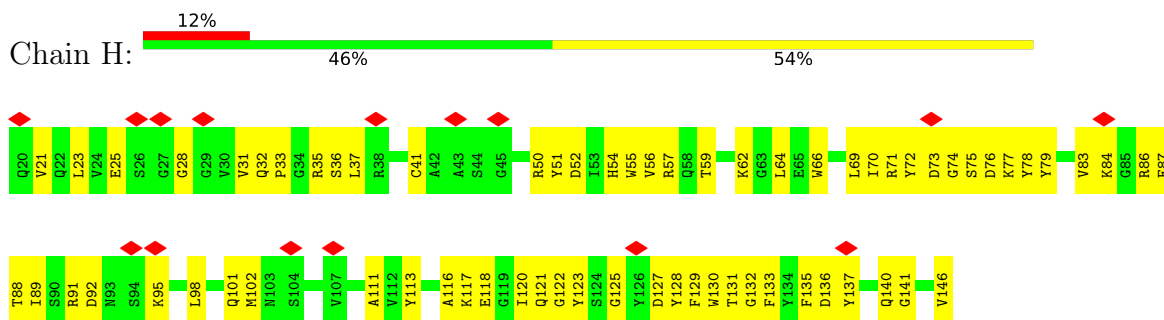
### 3 Residue-property plots i

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.

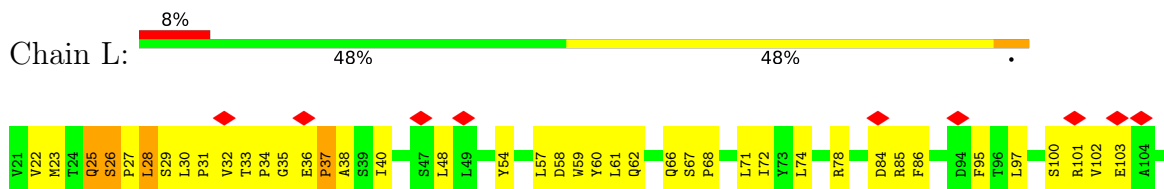
#### • Molecule 1: Spike protein S1



#### • Molecule 2: 3500H



#### • Molecule 3: 3500L





- Molecule 4: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain A: 50% 50%



## 4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	95595	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TITAN	Depositor
Voltage (kV)	300	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	59	Depositor
Minimum defocus (nm)	1200	Depositor
Maximum defocus (nm)	5000	Depositor
Magnification	Not provided	
Image detector	GATAN K3 (6k x 4k)	Depositor
Maximum map value	1.535	Depositor
Minimum map value	-1.104	Depositor
Average map value	-0.000	Depositor
Map value standard deviation	0.019	Depositor
Recommended contour level	0.242	Depositor
Map size ( $\text{\AA}$ )	385.2, 385.2, 385.2	wwPDB
Map dimensions	360, 360, 360	wwPDB
Map angles ( $^\circ$ )	90.0, 90.0, 90.0	wwPDB
Pixel spacing ( $\text{\AA}$ )	1.07, 1.07, 1.07	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: NAG

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	R	0.93	12/1547 (0.8%)	1.00	14/2102 (0.7%)
2	H	0.41	0/1052	0.56	0/1425
3	L	0.44	0/831	0.71	1/1130 (0.1%)
All	All	0.70	12/3430 (0.3%)	0.82	15/4657 (0.3%)

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	R	0	2

The worst 5 of 12 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	R	380	TYR	CB-CG	-11.43	1.34	1.51
1	R	380	TYR	CG-CD2	-8.05	1.28	1.39
1	R	380	TYR	CE1-CZ	-7.30	1.29	1.38
1	R	379	CYS	CB-SG	-6.28	1.71	1.82
1	R	377	PHE	C-O	-5.96	1.12	1.23

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	R	490	PHE	C-N-CD	-11.71	94.85	120.60
1	R	482	GLY	N-CA-C	-11.04	85.51	113.10
1	R	379	CYS	CA-CB-SG	-9.38	97.12	114.00
3	L	37	PRO	CA-N-CD	-9.25	98.55	111.50
1	R	475	ALA	N-CA-C	-7.95	89.55	111.00



There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	R	368	LEU	Mainchain
1	R	477	ASN	Mainchain

## 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	R	1504	0	1433	138	0
2	H	1024	0	970	66	0
3	L	813	0	798	62	0
4	A	28	0	25	2	0
All	All	3369	0	3226	258	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 39.

The worst 5 of 258 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:R:455:LEU:CD2	1:R:493:ARG:HD2	1.40	1.50
1:R:456:PHE:HB3	1:R:473:TYR:CD2	1.66	1.30
1:R:455:LEU:HD22	1:R:493:ARG:CD	1.65	1.26
1:R:456:PHE:HB3	1:R:473:TYR:CE2	1.76	1.19
1:R:473:TYR:O	1:R:488:CYS:O	1.60	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	R	183/236 (78%)	138 (75%)	30 (16%)	15 (8%)	1	12
2	H	125/127 (98%)	109 (87%)	16 (13%)	0	100	100
3	L	106/108 (98%)	91 (86%)	15 (14%)	0	100	100
All	All	414/471 (88%)	338 (82%)	61 (15%)	15 (4%)	6	30

5 of 15 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	R	371	LEU
1	R	375	PHE
1	R	377	PHE
1	R	477	ASN
1	R	479	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	R	163/207 (79%)	150 (92%)	13 (8%)	12	42
2	H	106/106 (100%)	106 (100%)	0	100	100
3	L	92/92 (100%)	89 (97%)	3 (3%)	38	64
All	All	361/405 (89%)	345 (96%)	16 (4%)	32	58

5 of 16 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	L	26	SER
3	L	25	GLN
1	R	486	PHE
1	R	493	ARG
1	R	480	CYS

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

Mol	Chain	Res	Type
1	R	422	ASN
1	R	474	GLN
2	H	121	GLN
3	L	25	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](#)

2 monosaccharides are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
4	NAG	A	1	4,1	14,14,15	0.19	0	17,19,21	0.61	0
4	NAG	A	2	4	14,14,15	0.29	0	17,19,21	0.39	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	NAG	A	1	4,1	-	2/6/23/26	0/1/1/1
4	NAG	A	2	4	-	1/6/23/26	0/1/1/1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (3) torsion outliers are listed below:

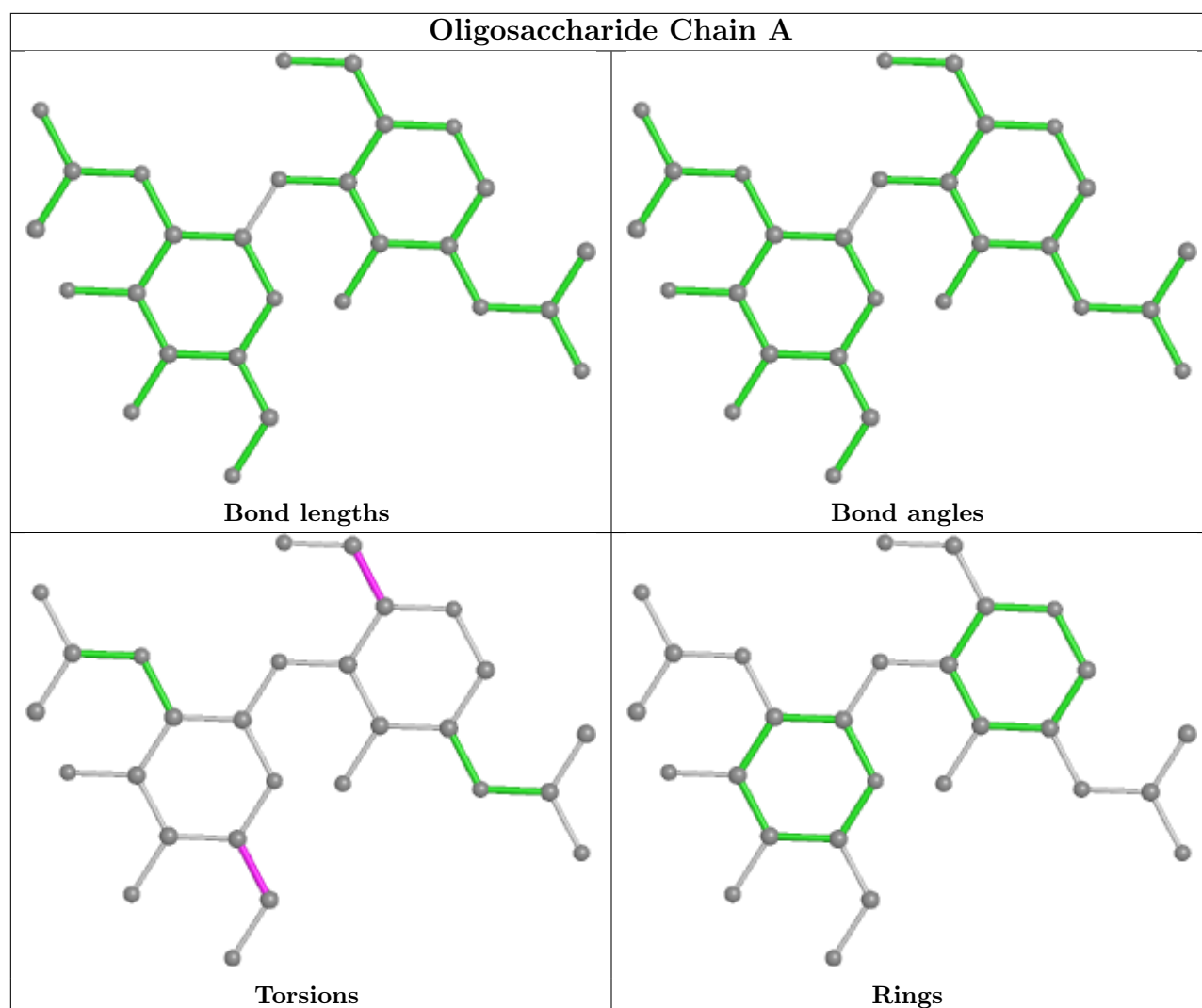
Mol	Chain	Res	Type	Atoms
4	A	1	NAG	C4-C5-C6-O6
4	A	1	NAG	O5-C5-C6-O6
4	A	2	NAG	O5-C5-C6-O6

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	1	NAG	2	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for oligosaccharide.



## 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.

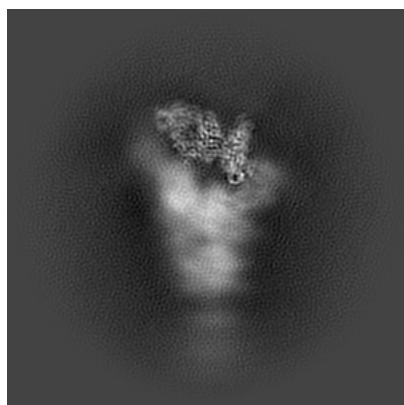
## 6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-32753. These allow visual inspection of the internal detail of the map and identification of artifacts.

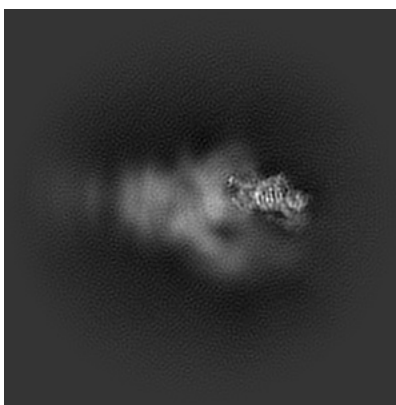
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

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

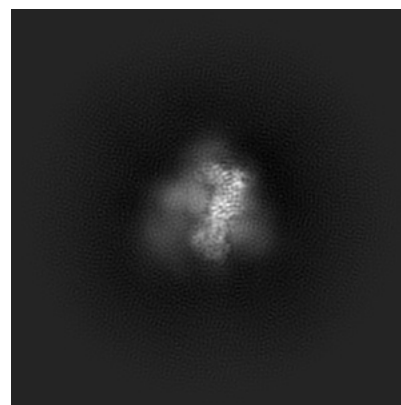
#### 6.1.1 Primary map



X



Y

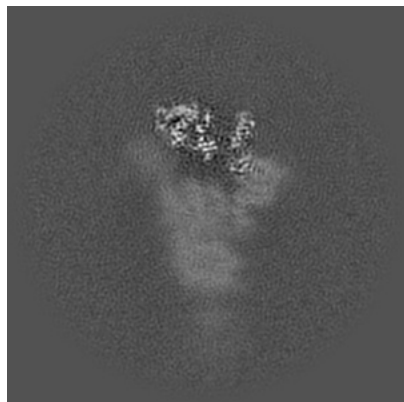


Z

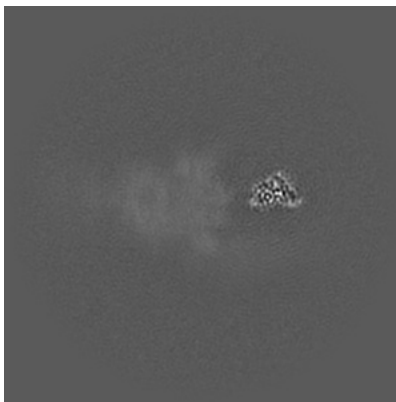
The images above show the map projected in three orthogonal directions.

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

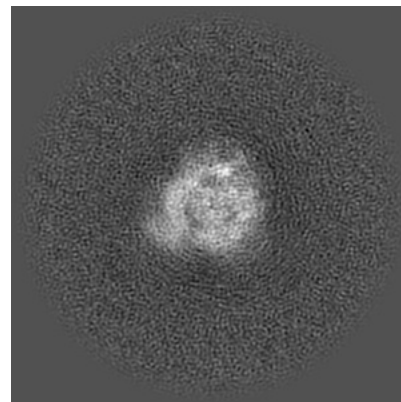
#### 6.2.1 Primary map



X Index: 180



Y Index: 180

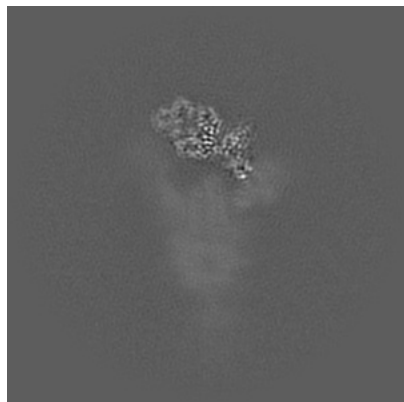


Z Index: 180

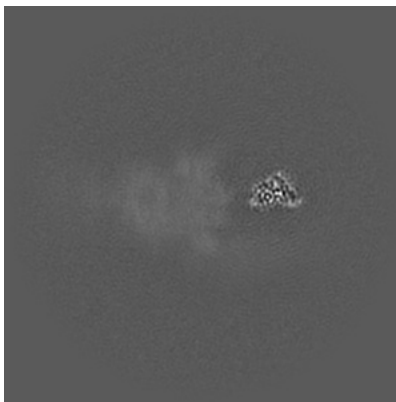
The images above show central slices of the map in three orthogonal directions.

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

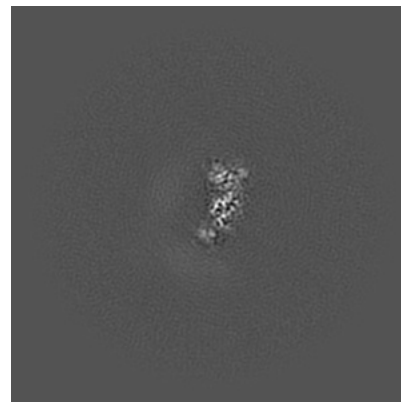
### 6.3.1 Primary map



X Index: 188



Y Index: 180



Z Index: 240

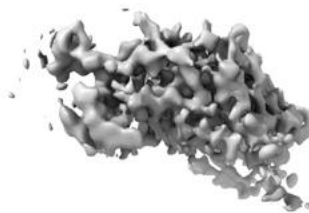
The images above show the largest variance slices of the map in three orthogonal directions.

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

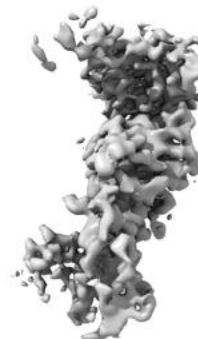
### 6.4.1 Primary map



X



Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.242. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

## 6.5 Mask visualisation

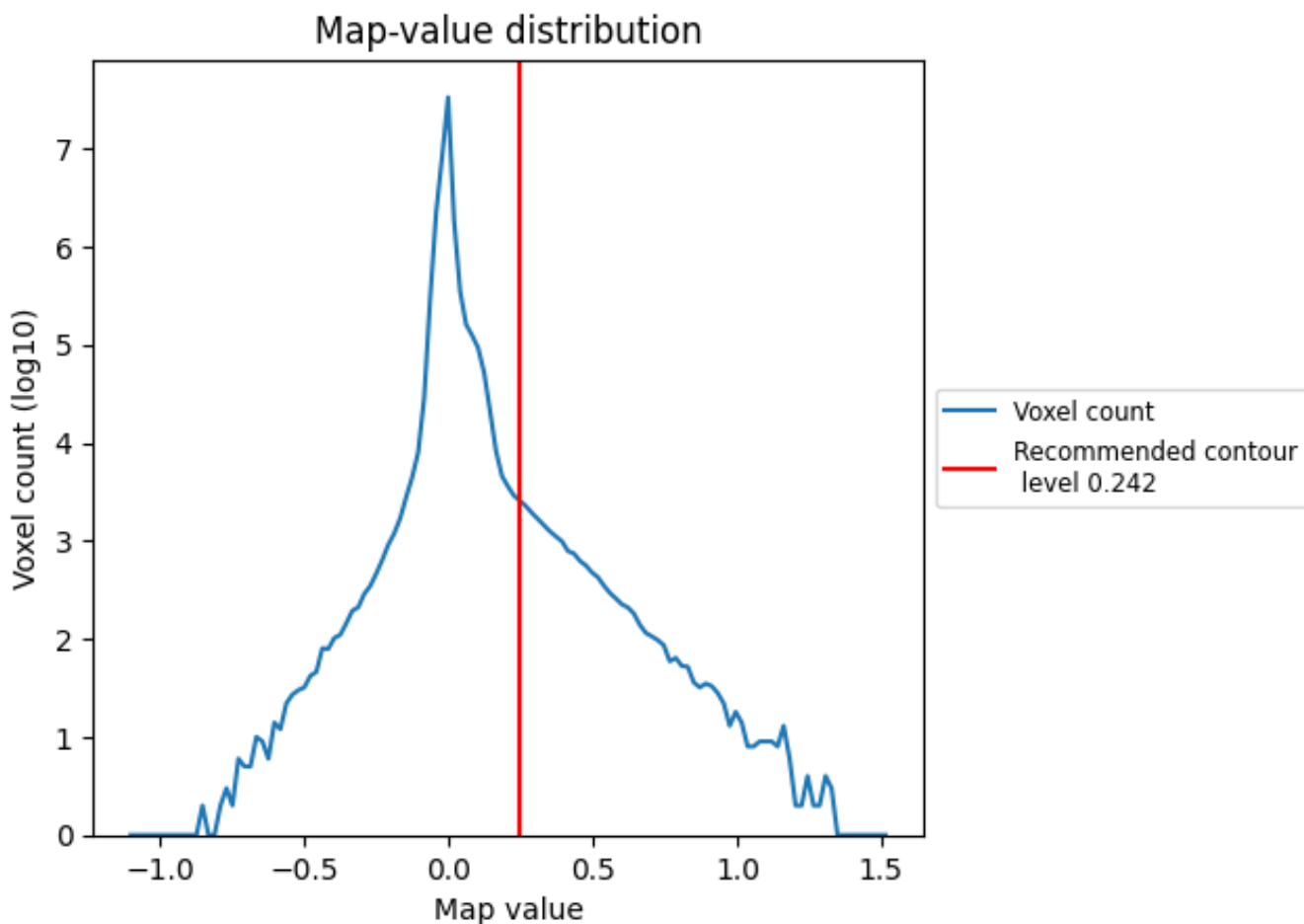
This section was not generated. No masks/segmentation were deposited.



## 7 Map analysis [i](#)

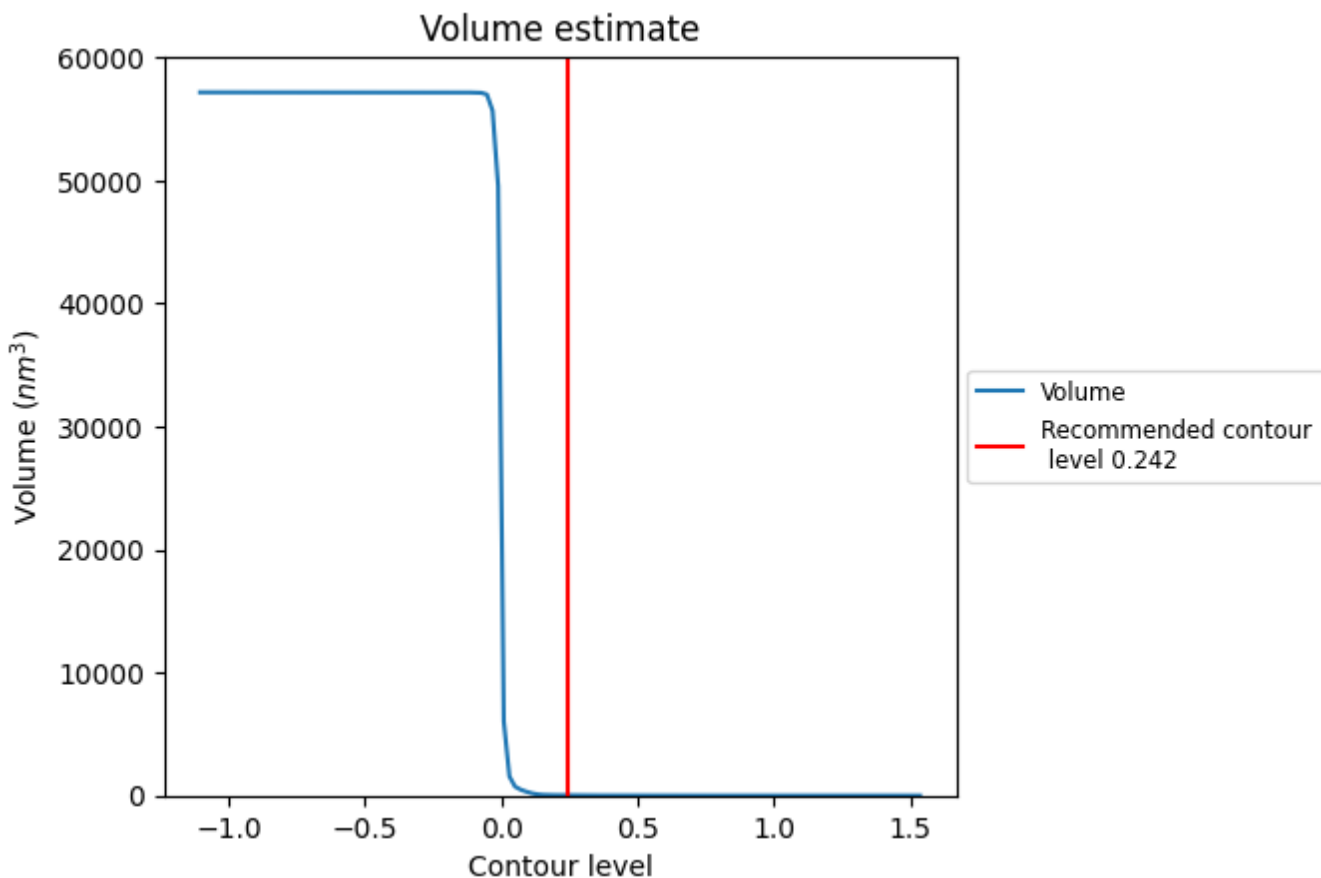
This section contains the results of statistical analysis of the map.

### 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.

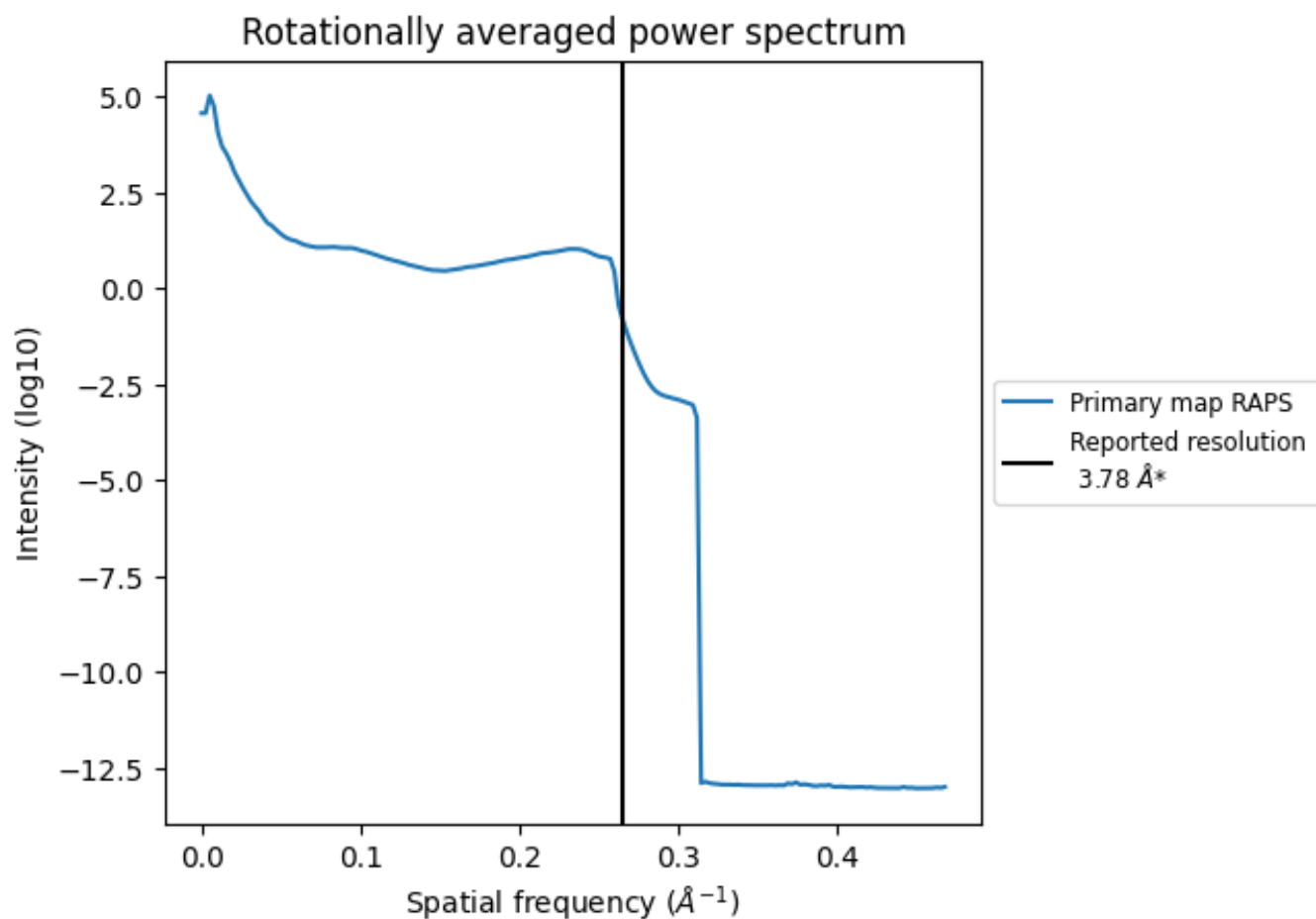
## 7.2 Volume estimate [i](#)



The volume at the recommended contour level is 25 nm<sup>3</sup>; this corresponds to an approximate mass of 22 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 0.265 Å<sup>-1</sup>

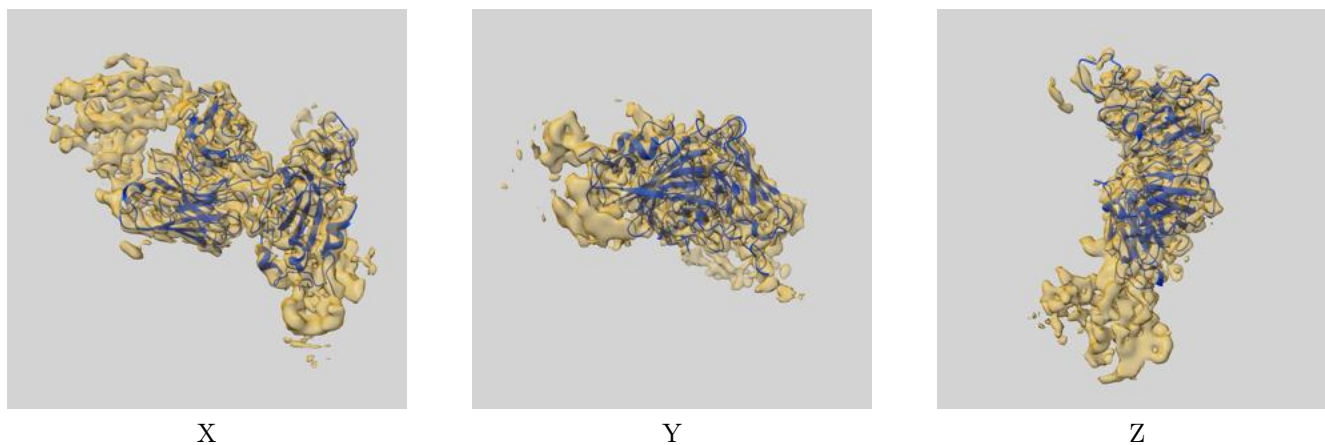
## 8 Fourier-Shell correlation

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

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

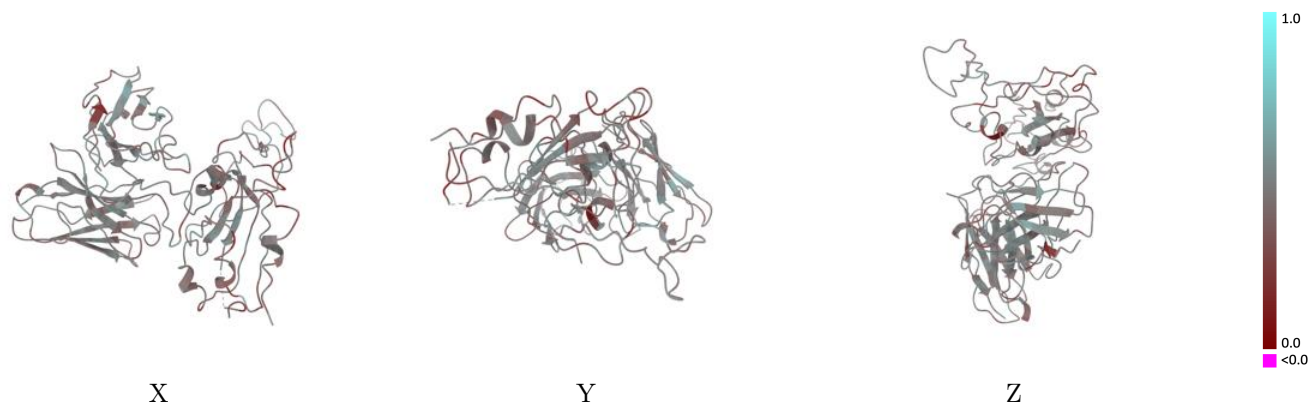
This section contains information regarding the fit between EMDB map EMD-32753 and PDB model 7WSC. Per-residue inclusion information can be found in section [3](#) on page [5](#).

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



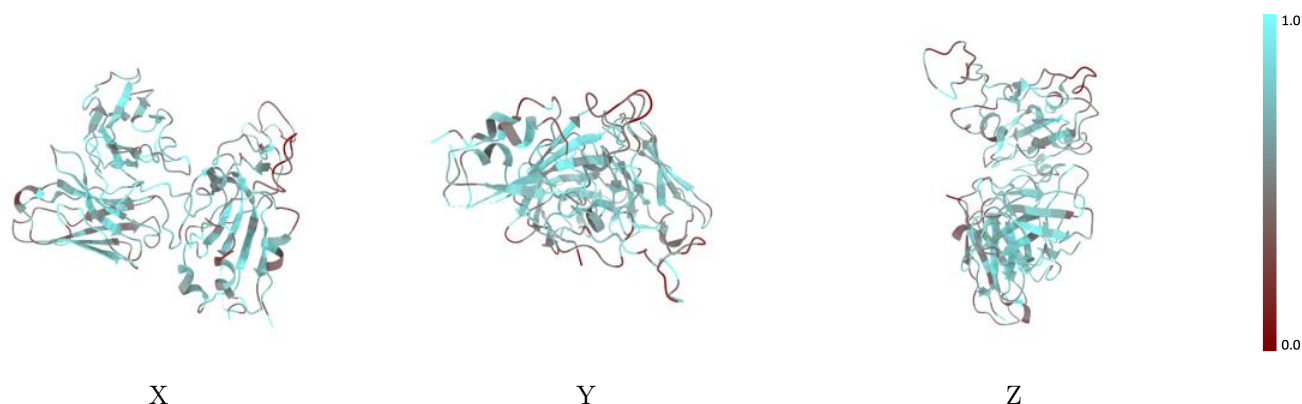
The images above show the 3D surface view of the map at the recommended contour level 0.242 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

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



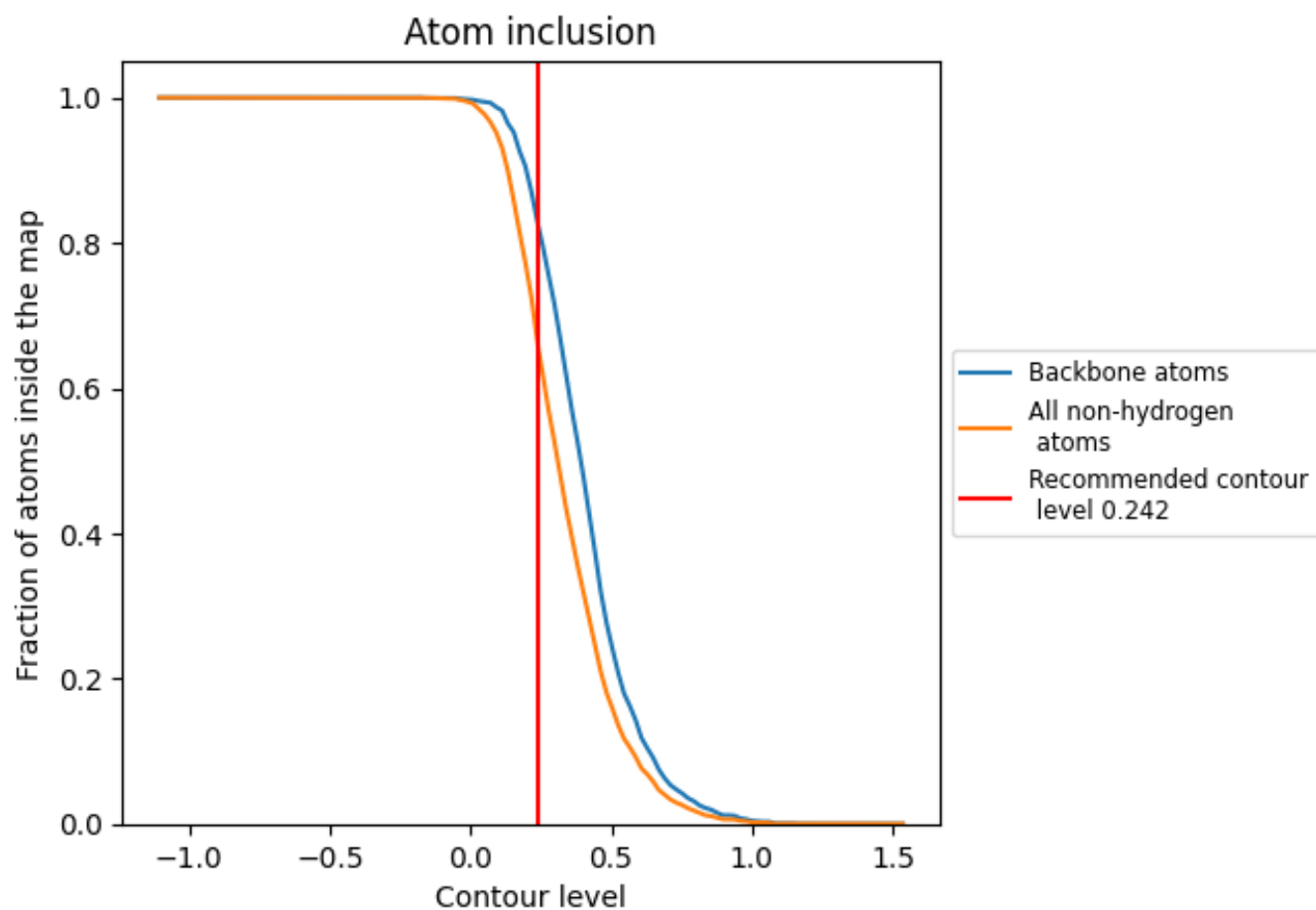
The images above show the model with each residue coloured according to its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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 (0.242).











## 9.4 Atom inclusion [i](#)



At the recommended contour level, 82% of all backbone atoms, 65% 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 (0.242) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	 0.6534	 0.4440
A	 0.4643	 0.4350
H	 0.6764	 0.4510
L	 0.6955	 0.4640
R	 0.6184	 0.4280

