



## wwPDB EM Validation Summary Report ⓘ

Nov 19, 2022 – 04:05 pm GMT

PDB ID : 5NJG  
EMDB ID : EMD-3654  
Title : Structure of an ABC transporter: part of the structure that could be built de novo  
Authors : Taylor, N.M.I.; Manolaridis, I.; Jackson, S.M.; Kowal, J.; Stahlberg, H.; Locher, K.P.  
Deposited on : 2017-03-28  
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 : 1.8.4, 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) : 2.31.2

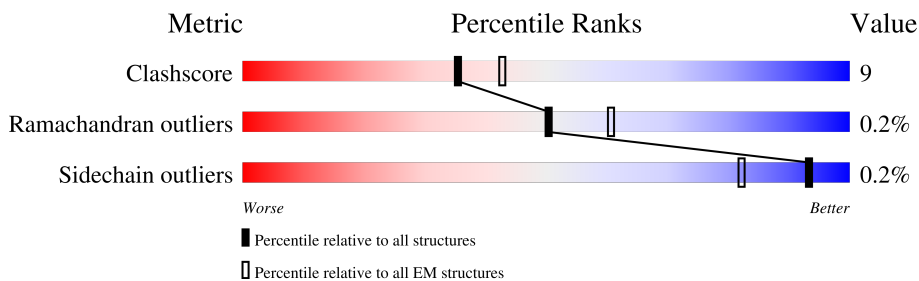
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 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	A	664	
1	B	664	
2	C	221	
2	E	221	
3	D	214	
3	F	214	
4	G	2	
4	H	2	

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 8038 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 ATP-binding cassette sub-family G member 2.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	286	Total	C	N	O	S	0	0
			2249	1492	353	384	20		
1	B	286	Total	C	N	O	S	0	0
			2249	1492	353	384	20		

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-8	ASP	-	expression tag	UNP Q9UNQ0
A	-7	TYR	-	expression tag	UNP Q9UNQ0
A	-6	LYS	-	expression tag	UNP Q9UNQ0
A	-5	ASP	-	expression tag	UNP Q9UNQ0
A	-4	ASP	-	expression tag	UNP Q9UNQ0
A	-3	ASP	-	expression tag	UNP Q9UNQ0
A	-2	ASP	-	expression tag	UNP Q9UNQ0
A	-1	LYS	-	expression tag	UNP Q9UNQ0
A	0	GLY	-	expression tag	UNP Q9UNQ0
A	1	SER	-	expression tag	UNP Q9UNQ0
B	-8	ASP	-	expression tag	UNP Q9UNQ0
B	-7	TYR	-	expression tag	UNP Q9UNQ0
B	-6	LYS	-	expression tag	UNP Q9UNQ0
B	-5	ASP	-	expression tag	UNP Q9UNQ0
B	-4	ASP	-	expression tag	UNP Q9UNQ0
B	-3	ASP	-	expression tag	UNP Q9UNQ0
B	-2	ASP	-	expression tag	UNP Q9UNQ0
B	-1	LYS	-	expression tag	UNP Q9UNQ0
B	0	GLY	-	expression tag	UNP Q9UNQ0
B	1	SER	-	expression tag	UNP Q9UNQ0

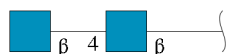
- Molecule 2 is a protein called 5D3-Fab heavy chain.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	C	118	928	592	153	180	3	0	0
2	E	118	928	592	153	180	3	0	0

- Molecule 3 is a protein called 5D3-Fab light chain.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	D	106	814	517	136	159	2	0	0
3	F	106	814	517	136	159	2	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	G	2	28	16	2	10	0	0
4	H	2	28	16	2	10	0	0







## 4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C2	Depositor
Number of particles used	97612	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 ( $e^-/\text{\AA}^2$ )	67	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	0.287	Depositor
Minimum map value	-0.170	Depositor
Average map value	0.001	Depositor
Map value standard deviation	0.008	Depositor
Recommended contour level	0.038	Depositor
Map size (Å)	249.36002, 249.36002, 249.36002	wwPDB
Map dimensions	240, 240, 240	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.039, 1.039, 1.039	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	A	0.64	0/2306	0.72	0/3128
1	B	0.64	0/2306	0.72	0/3128
2	C	0.68	0/953	0.73	1/1297 (0.1%)
2	E	0.68	0/953	0.73	1/1297 (0.1%)
3	D	0.74	1/834 (0.1%)	0.69	0/1133
3	F	0.74	1/834 (0.1%)	0.69	0/1133
All	All	0.67	2/8186 (0.0%)	0.72	2/11116 (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	3
1	B	0	3
2	C	0	1
2	E	0	1
All	All	0	8

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	D	42	ASN	C-N	-6.14	1.20	1.34
3	F	42	ASN	C-N	-6.14	1.20	1.34

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	E	81	LEU	CA-CB-CG	5.87	128.79	115.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
2	C	81	LEU	CA-CB-CG	5.86	128.77	115.30

There are no chirality outliers.

5 of 8 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	601	ASN	Peptide
1	A	602	PRO	Peptide
1	A	603	CYS	Peptide
1	B	601	ASN	Peptide
1	B	602	PRO	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	2249	0	2303	46	0
1	B	2249	0	2303	44	0
2	C	928	0	890	15	0
2	E	928	0	890	17	0
3	D	814	0	793	14	0
3	F	814	0	793	14	0
4	G	28	0	25	0	0
4	H	28	0	25	0	0
All	All	8038	0	8022	138	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 9.

The worst 5 of 138 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:F:18:ARG:HA	3:F:75:ILE:O	1.91	0.70
3:D:18:ARG:HA	3:D:75:ILE:O	1.91	0.70
1:A:490:THR:O	1:A:494:TYR:HB2	1.93	0.69

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:490:THR:O	1:B:494:TYR:HB2	1.93	0.69
2:E:28:SER:HB3	2:E:31:SER:HB3	1.83	0.61

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	284/664 (43%)	267 (94%)	16 (6%)	1 (0%)	34	69
1	B	284/664 (43%)	267 (94%)	16 (6%)	1 (0%)	34	69
2	C	116/221 (52%)	105 (90%)	11 (10%)	0	100	100
2	E	116/221 (52%)	105 (90%)	11 (10%)	0	100	100
3	D	104/214 (49%)	96 (92%)	8 (8%)	0	100	100
3	F	104/214 (49%)	96 (92%)	8 (8%)	0	100	100
All	All	1008/2198 (46%)	936 (93%)	70 (7%)	2 (0%)	50	78

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	602	PRO
1	B	602	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	247/568 (44%)	246 (100%)	1 (0%)	91	95
1	B	247/568 (44%)	246 (100%)	1 (0%)	91	95
2	C	102/193 (53%)	102 (100%)	0	100	100
2	E	102/193 (53%)	102 (100%)	0	100	100
3	D	90/189 (48%)	90 (100%)	0	100	100
3	F	90/189 (48%)	90 (100%)	0	100	100
All	All	878/1900 (46%)	876 (100%)	2 (0%)	93	97

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	557	ASN
1	B	557	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	398	GLN
1	B	376	GLN
1	A	601	ASN
1	A	557	ASN
1	B	375	HIS

### 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](#)

4 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	G	1	4,1	14,14,15	1.29	1 (7%)	17,19,21	0.84	0
4	NAG	G	2	4	14,14,15	0.22	0	17,19,21	0.45	0
4	NAG	H	1	4,1	14,14,15	1.05	1 (7%)	17,19,21	0.88	0
4	NAG	H	2	4	14,14,15	0.22	0	17,19,21	0.45	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	G	1	4,1	-	0/6/23/26	0/1/1/1
4	NAG	G	2	4	-	2/6/23/26	0/1/1/1
4	NAG	H	1	4,1	-	1/6/23/26	0/1/1/1
4	NAG	H	2	4	-	2/6/23/26	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	G	1	NAG	O5-C1	-4.71	1.36	1.43
4	H	1	NAG	O5-C1	-3.70	1.37	1.43

There are no bond angle outliers.

There are no chirality outliers.

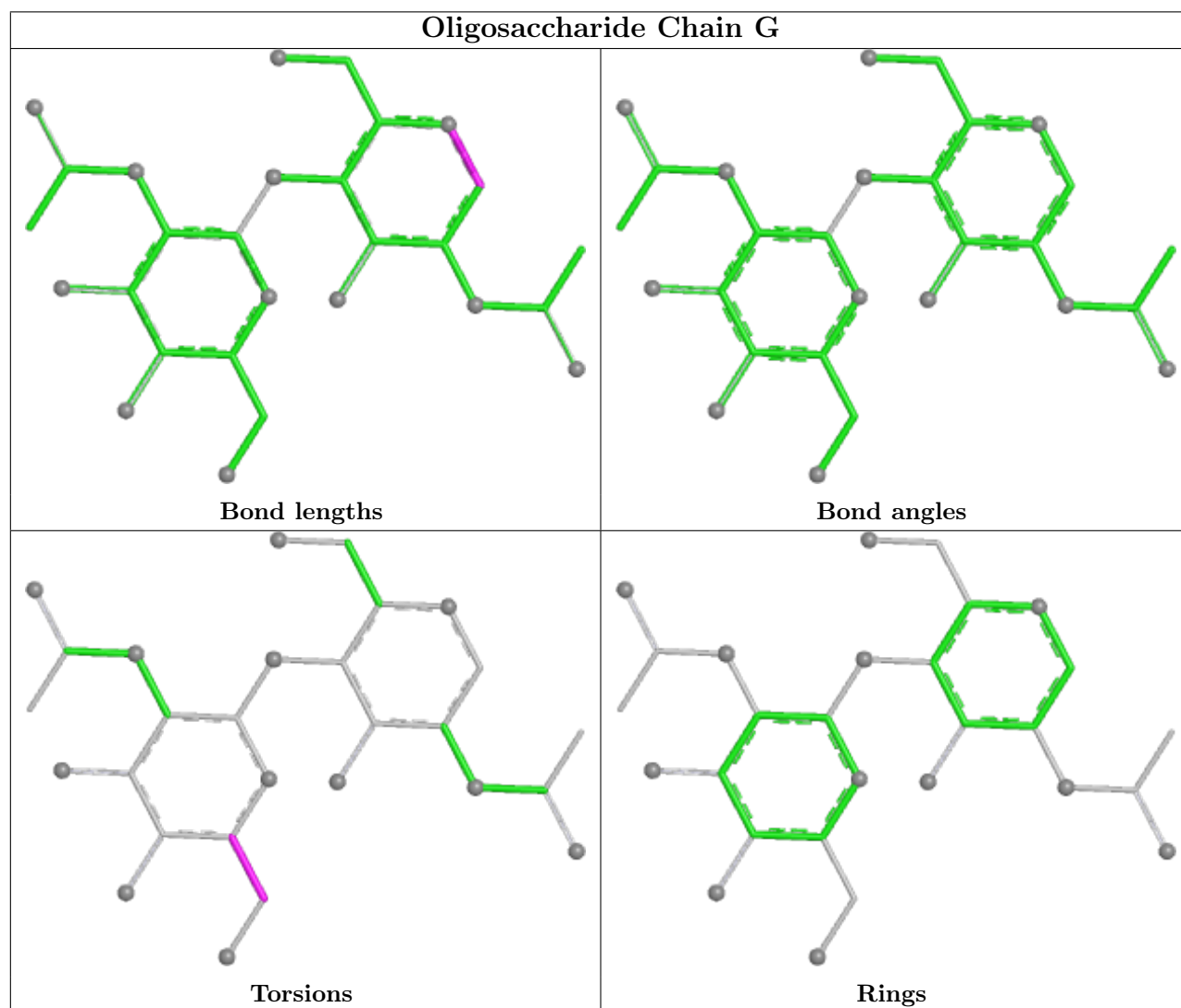
All (5) torsion outliers are listed below:

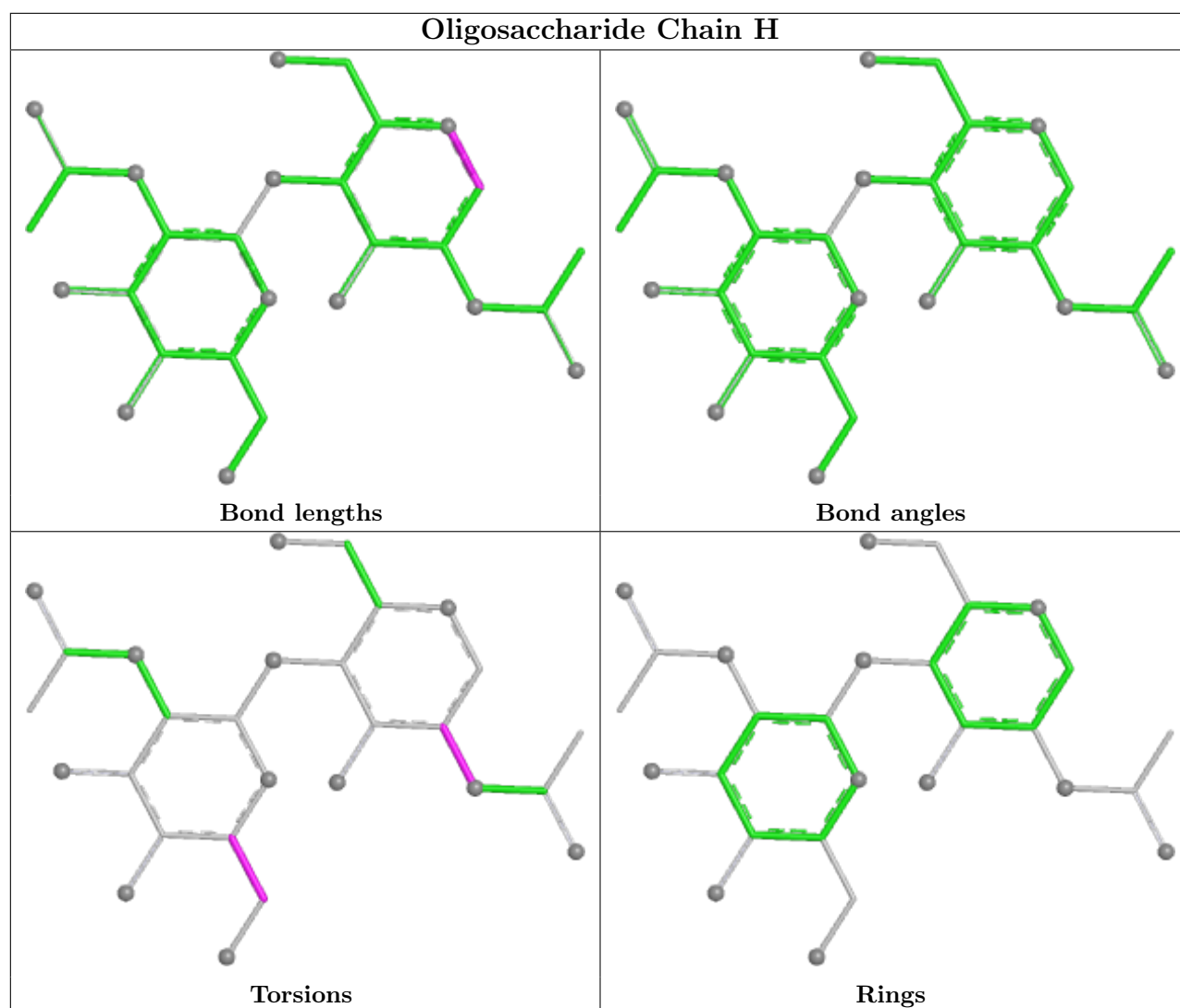
Mol	Chain	Res	Type	Atoms
4	H	2	NAG	O5-C5-C6-O6
4	G	2	NAG	O5-C5-C6-O6
4	H	2	NAG	C4-C5-C6-O6
4	G	2	NAG	C4-C5-C6-O6
4	H	1	NAG	C1-C2-N2-C7

There are no ring outliers.

No monomer is involved in short contacts.

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](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
3	D	1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Number of breaks
3	F	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	D	42:ASN	C	43:ALA	N	1.19
1	F	42:ASN	C	43:ALA	N	1.19



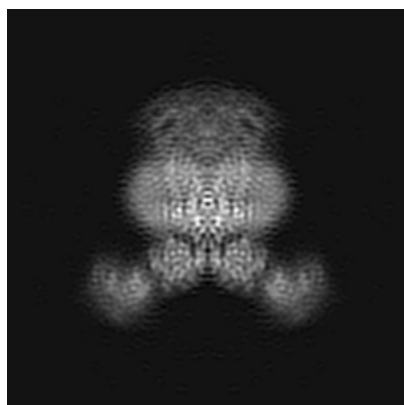
## 6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-3654. These allow visual inspection of the internal detail of the map and identification of artifacts.

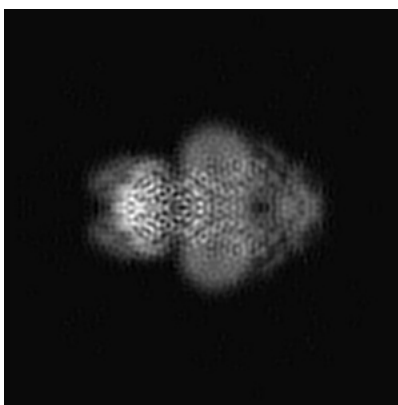
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

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

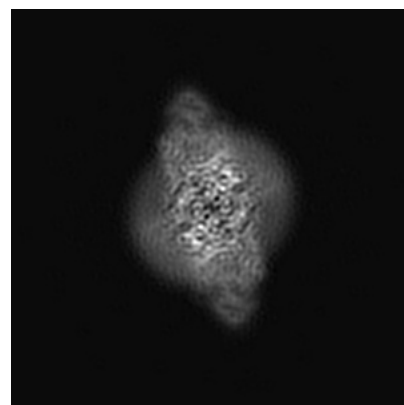
#### 6.1.1 Primary map



X



Y

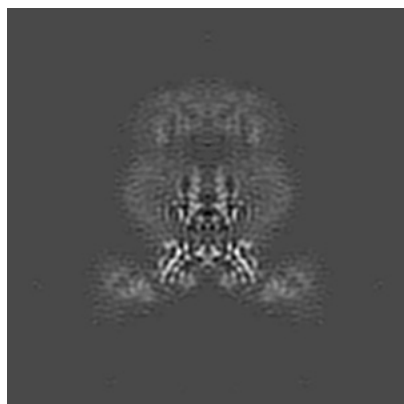


Z

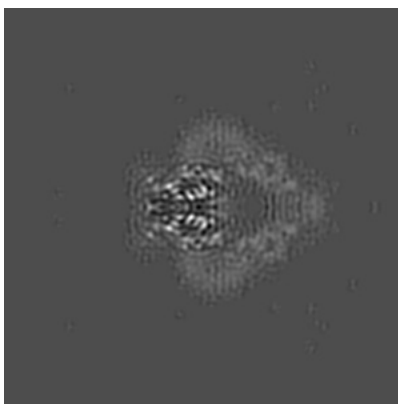
The images above show the map projected in three orthogonal directions.

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

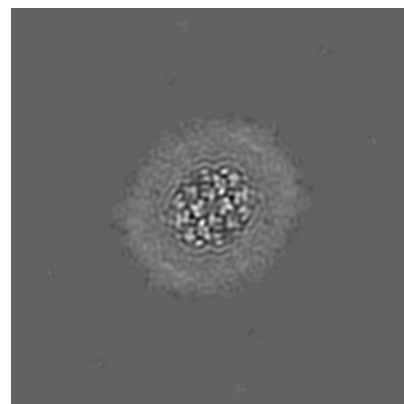
#### 6.2.1 Primary map



X Index: 120



Y Index: 120

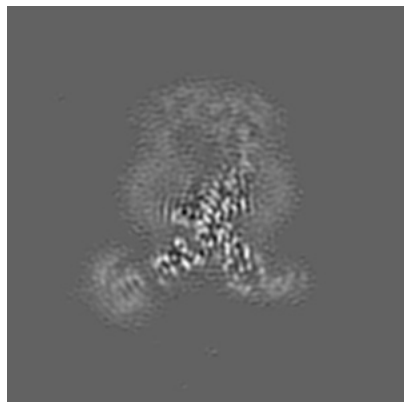


Z Index: 120

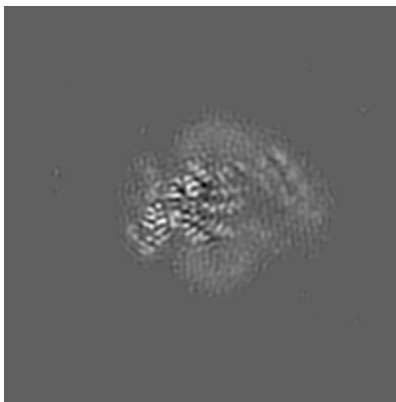
The images above show central slices of the map in three orthogonal directions.

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

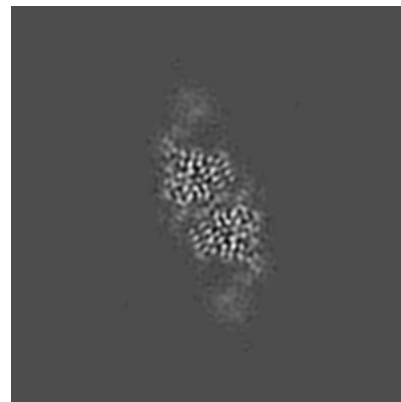
### 6.3.1 Primary map



X Index: 127



Y Index: 126

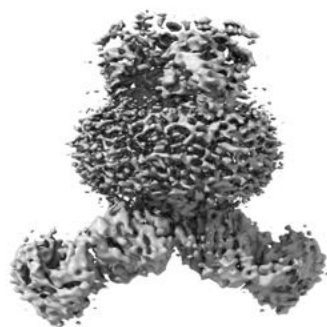


Z Index: 88

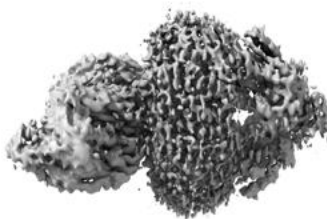
The images above show the largest variance slices of the map in three orthogonal directions.

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

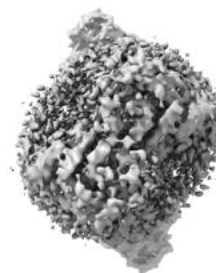
### 6.4.1 Primary map



X



Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.038. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

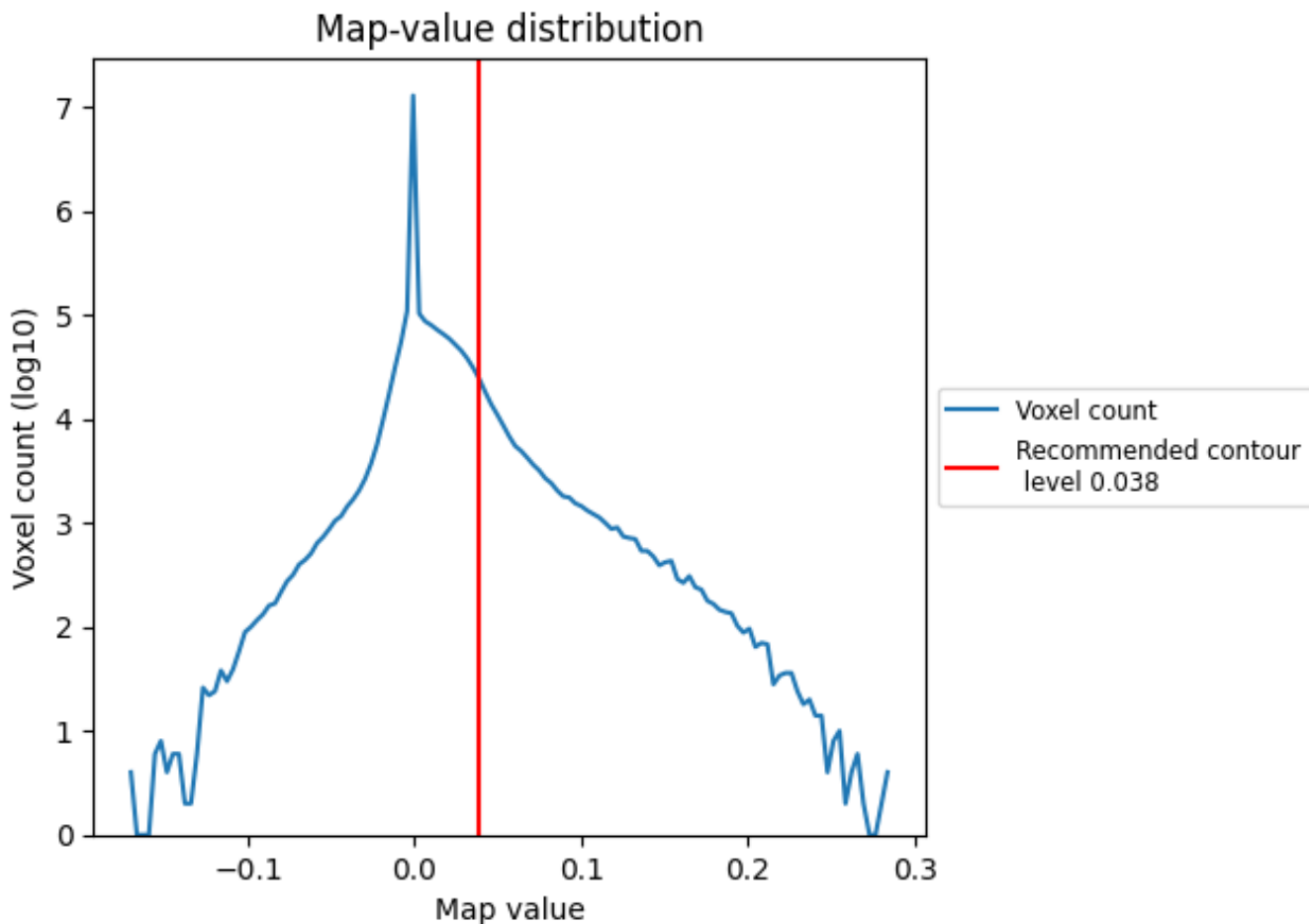
## 6.5 Mask visualisation

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

## 7 Map analysis [i](#)

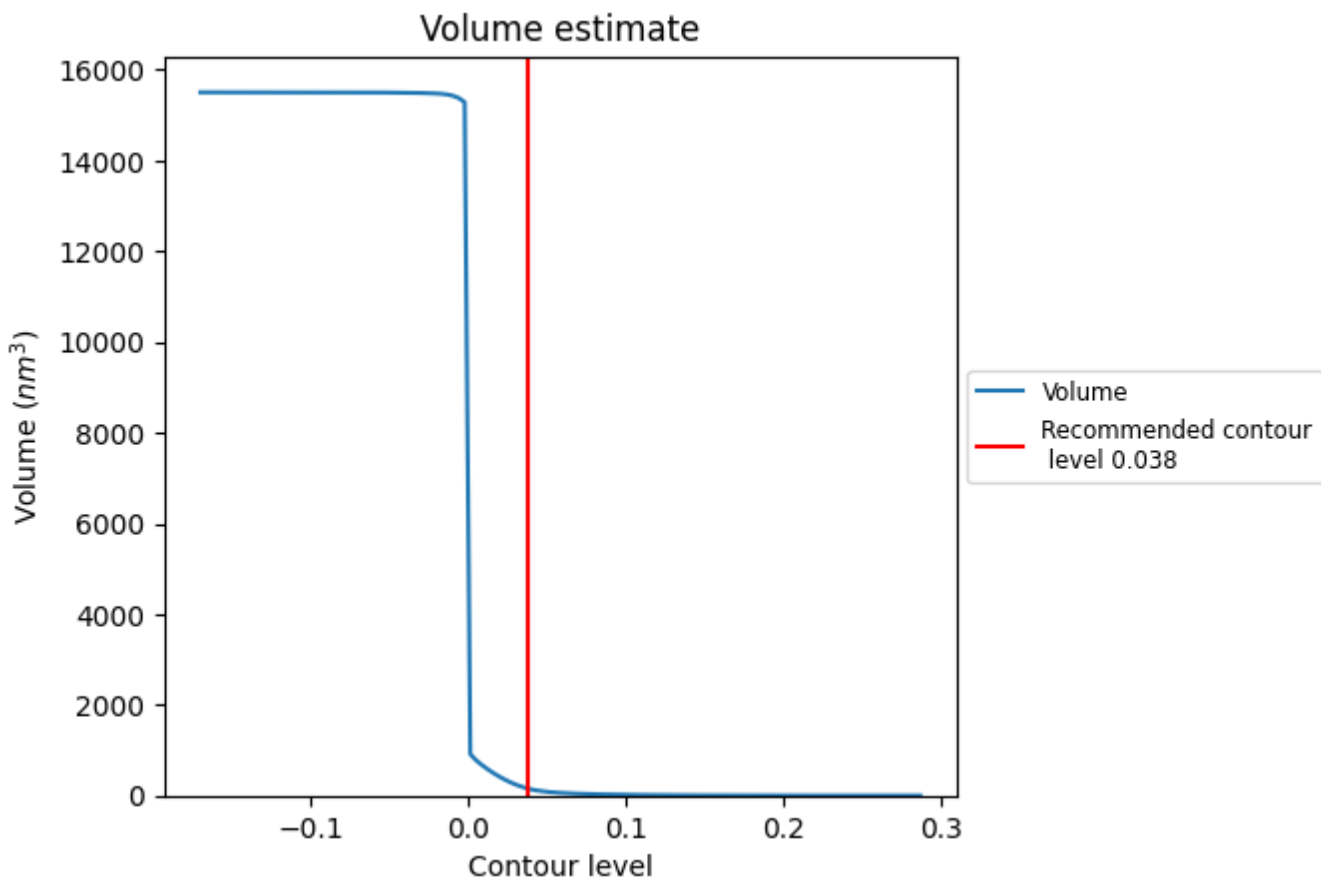
This section contains the results of statistical analysis of the map.

### 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.

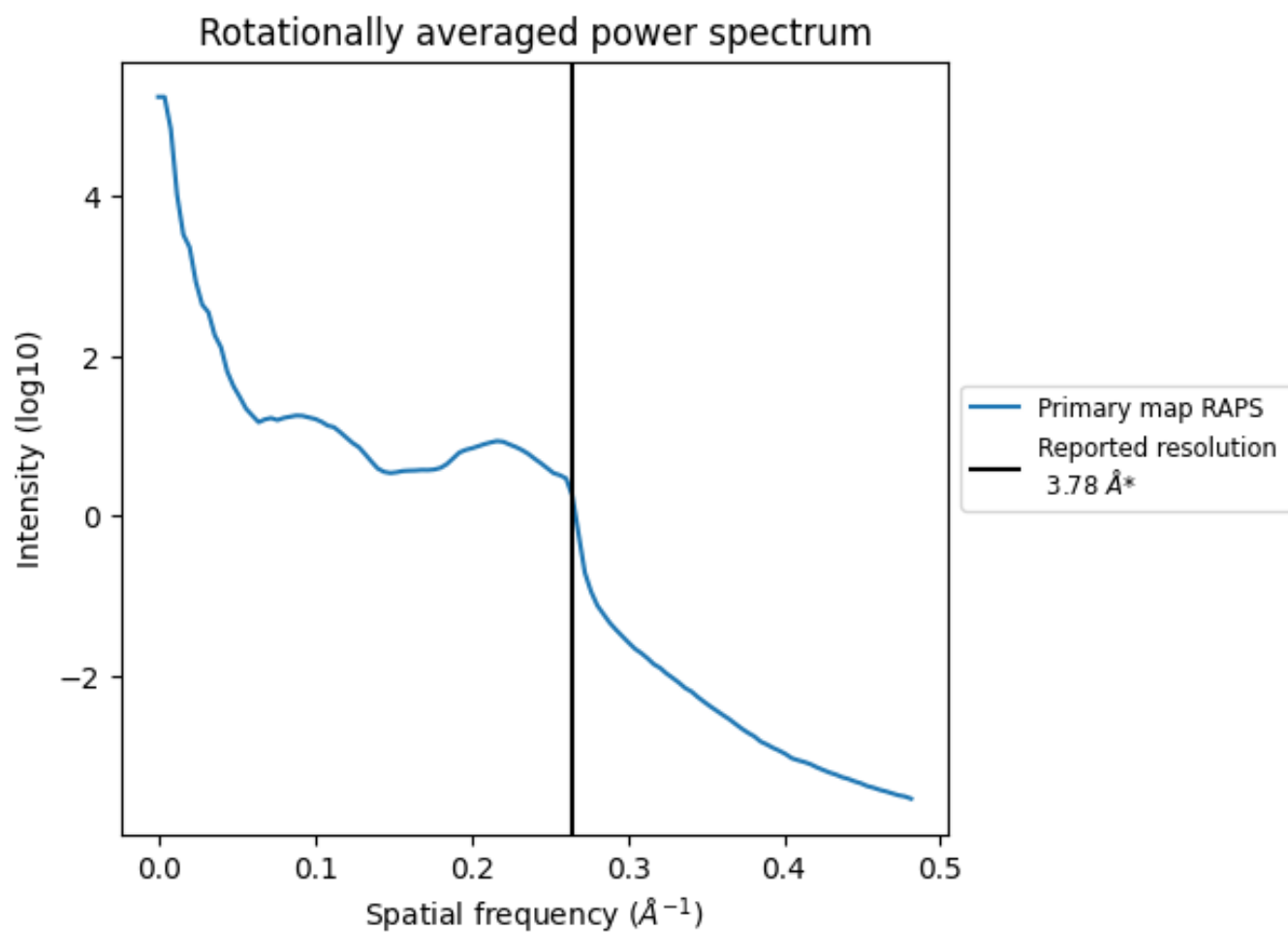
## 7.2 Volume estimate [i](#)



The volume at the recommended contour level is 155 nm<sup>3</sup>; this corresponds to an approximate mass of 140 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](#)

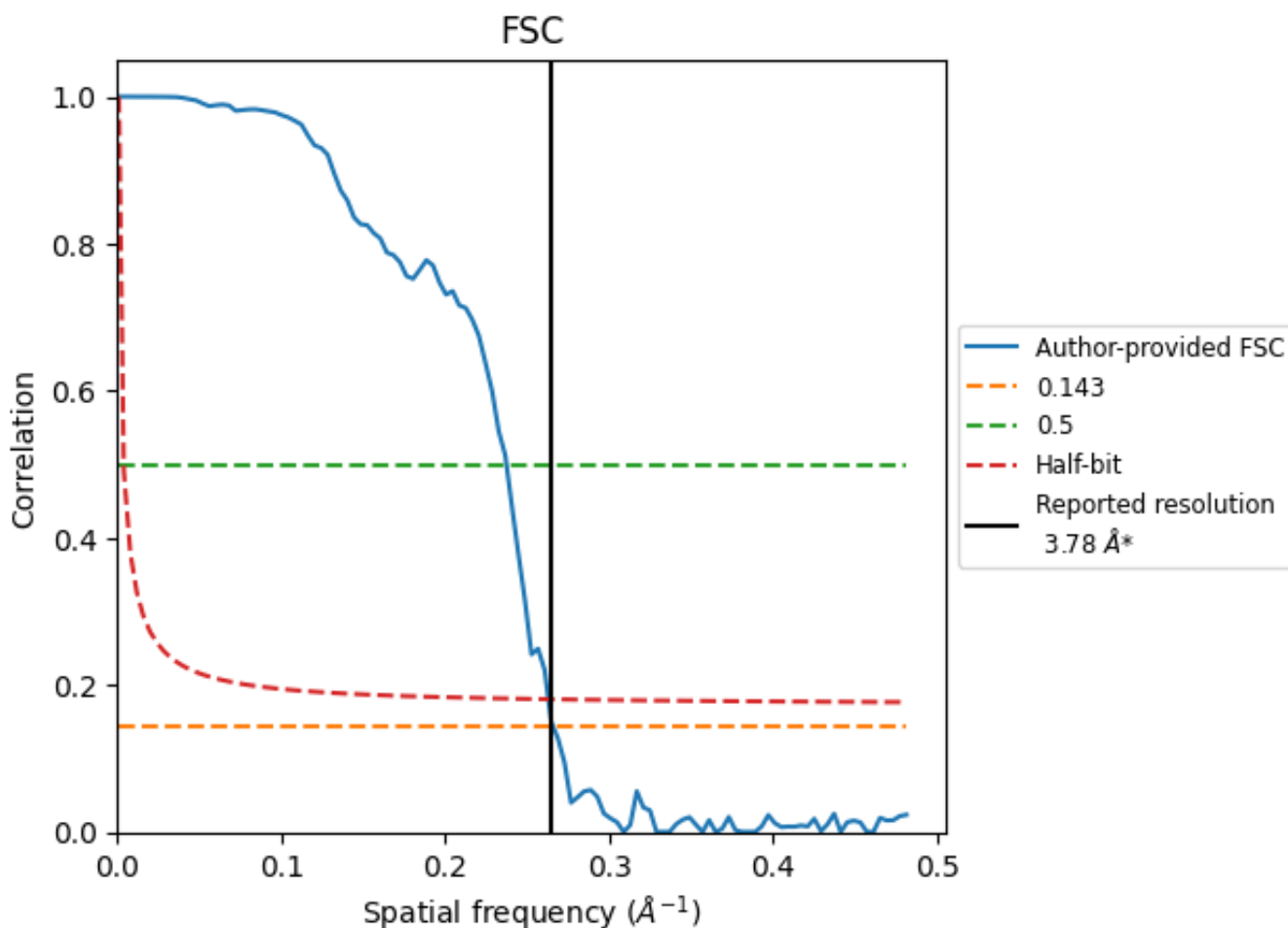


\*Reported resolution corresponds to spatial frequency of  $0.265 \text{\AA}^{-1}$

## 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 0.265 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	3.78	-	-
Author-provided FSC curve	3.76	4.21	3.80
Unmasked-calculated*	-	-	-

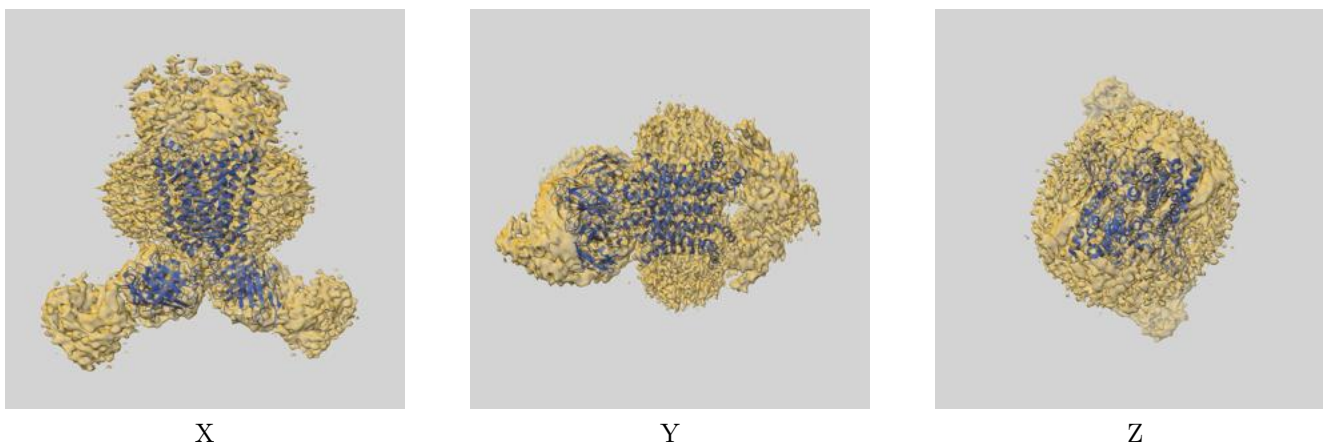
\*Resolution estimate based on FSC curve calculated by comparison of deposited half-maps.



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

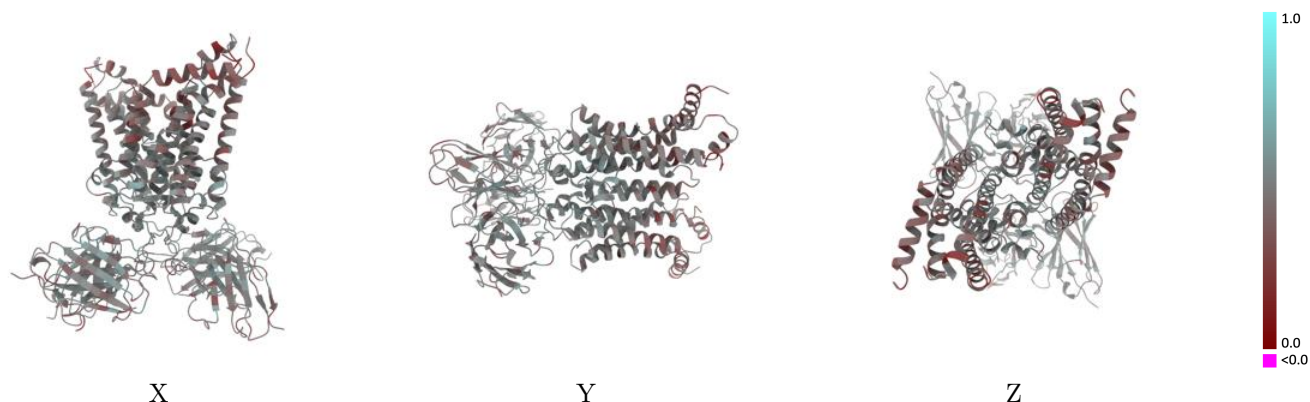
This section contains information regarding the fit between EMDB map EMD-3654 and PDB model 5NJG. Per-residue inclusion information can be found in section 3 on page 5.

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



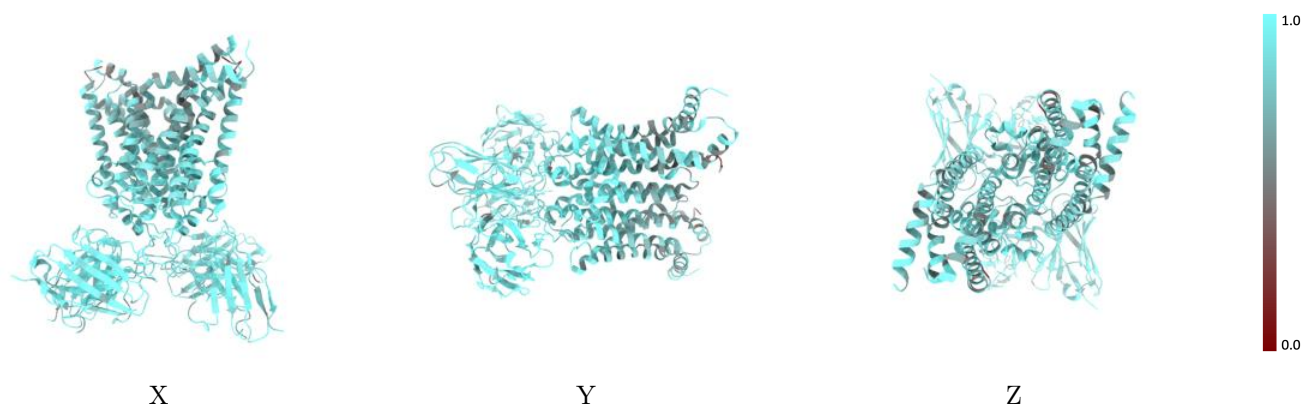
The images above show the 3D surface view of the map at the recommended contour level 0.038 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

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



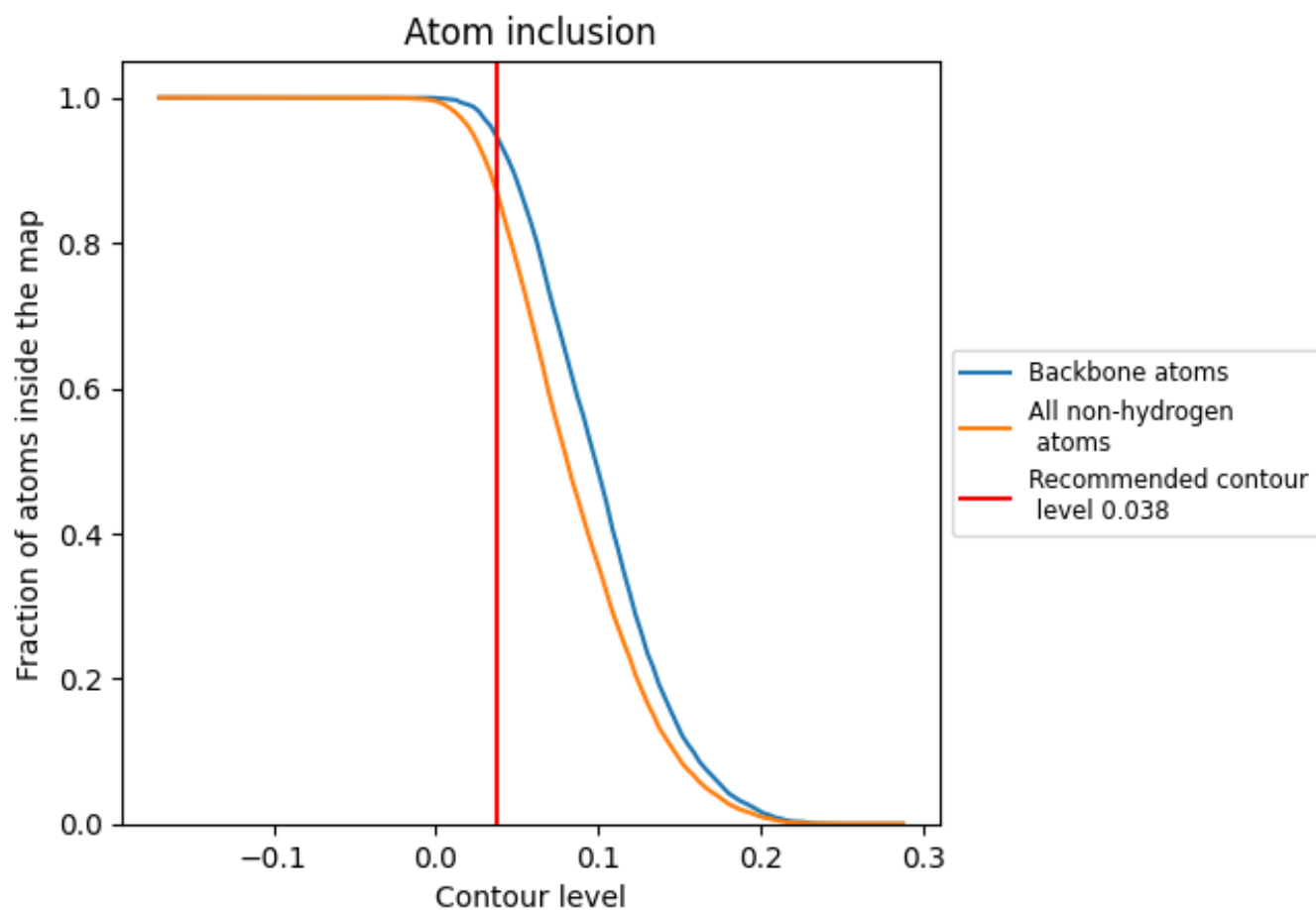
The images above show the model with each residue coloured according to its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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 (0.038).



















## 9.4 Atom inclusion [i](#)



At the recommended contour level, 94% of all backbone atoms, 87% 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 (0.038) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	 0.8662	 0.4550
A	 0.8382	 0.4450
B	 0.8405	 0.4440
C	 0.9063	 0.4680
D	 0.9006	 0.4670
E	 0.9140	 0.4710
F	 0.8893	 0.4690
G	 0.7500	 0.4050
H	 0.7500	 0.4390

