



wwPDB EM Validation Summary Report ⓘ

May 16, 2022 – 10:43 am BST

PDB ID : 7POV
EMDB ID : EMD-13575
Title : MUC2 Tubules of D1D2D3 domains
Authors : Javitt, G.; Fass, D.
Deposited on : 2021-09-10
Resolution : 3.80 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

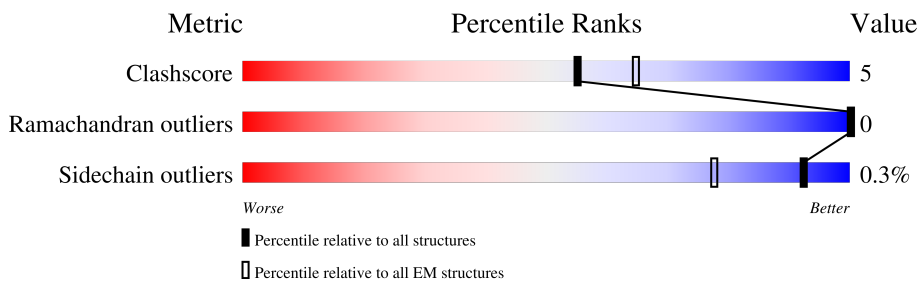
EMDB validation analysis : 0.0.1.dev8
Mogul : 1.8.4, CSD as541be (2020)
MolProbity : 4.02b-467
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.28.1

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 3.80 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	1245	
1	B	1245	
1	C	1245	
1	D	1245	
2	E	2	
2	F	2	

2 Entry composition i

There are 5 unique types of molecules in this entry. The entry contains 33345 atoms, of which 15915 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 Mucin-2.

Mol	Chain	Residues	Atoms						AltConf	Trace
			Total	C	H	N	O	S		
1	A	708	10088	3272	4821	915	1013	67	2	0
1	B	708	10109	3277	4834	918	1013	67	3	0
1	C	440	6460	2084	3098	580	651	47	1	0
1	D	440	6460	2084	3098	580	651	47	1	0

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1260	HIS	-	expression tag	UNP A0A0G2JR65
A	1261	HIS	-	expression tag	UNP A0A0G2JR65
A	1262	HIS	-	expression tag	UNP A0A0G2JR65
A	1263	HIS	-	expression tag	UNP A0A0G2JR65
A	1264	HIS	-	expression tag	UNP A0A0G2JR65
A	1265	HIS	-	expression tag	UNP A0A0G2JR65
B	1260	HIS	-	expression tag	UNP A0A0G2JR65
B	1261	HIS	-	expression tag	UNP A0A0G2JR65
B	1262	HIS	-	expression tag	UNP A0A0G2JR65
B	1263	HIS	-	expression tag	UNP A0A0G2JR65
B	1264	HIS	-	expression tag	UNP A0A0G2JR65
B	1265	HIS	-	expression tag	UNP A0A0G2JR65
C	1260	HIS	-	expression tag	UNP A0A0G2JR65
C	1261	HIS	-	expression tag	UNP A0A0G2JR65
C	1262	HIS	-	expression tag	UNP A0A0G2JR65
C	1263	HIS	-	expression tag	UNP A0A0G2JR65
C	1264	HIS	-	expression tag	UNP A0A0G2JR65
C	1265	HIS	-	expression tag	UNP A0A0G2JR65
D	1260	HIS	-	expression tag	UNP A0A0G2JR65
D	1261	HIS	-	expression tag	UNP A0A0G2JR65
D	1262	HIS	-	expression tag	UNP A0A0G2JR65
D	1263	HIS	-	expression tag	UNP A0A0G2JR65

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	1264	HIS	-	expression tag	UNP A0A0G2JR65
D	1265	HIS	-	expression tag	UNP A0A0G2JR65

- Molecule 2 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	H	N	O		
2	E	2	44	16	16	2	10	0	0
2	F	2	44	16	16	2	10	0	0

- Molecule 3 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		AltConf
3	A	2	Total	Ca	0
			2	2	
3	B	2	Total	Ca	0
			2	2	
3	C	1	Total	Ca	0
			1	1	
3	D	1	Total	Ca	0
			1	1	

- Molecule 4 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: C₈H₁₅NO₆).



Mol	Chain	Residues	Atoms					AltConf
			Total	C	H	N	O	
4	A	1	Total	C	H	N	O	0
			44	16	16	2	10	
4	A	1	Total	C	H	N	O	0
			44	16	16	2	10	
4	B	1	Total	C	H	N	O	0
			44	16	16	2	10	
4	B	1	Total	C	H	N	O	0
			44	16	16	2	10	

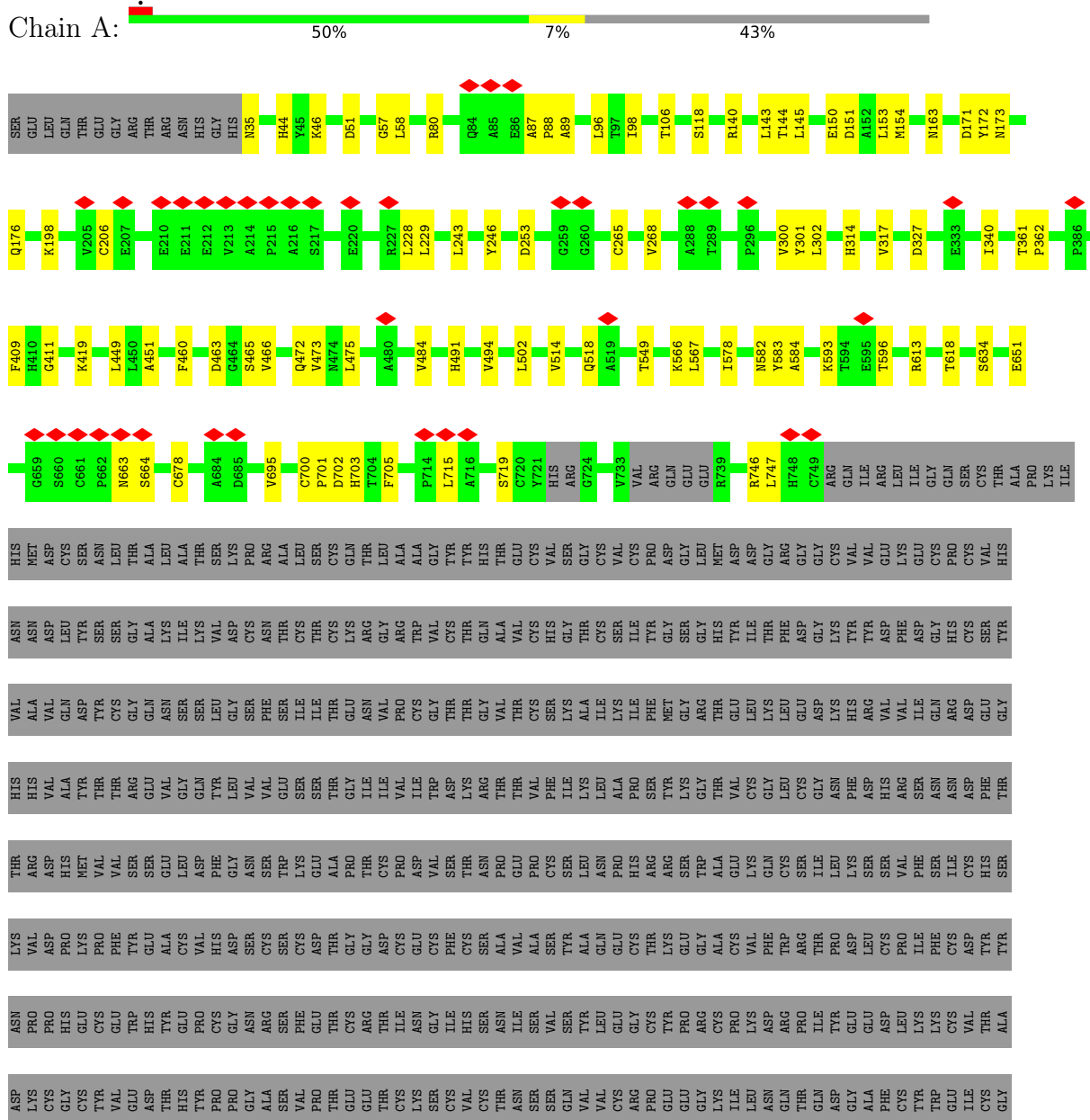
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		AltConf
5	A	11	Total	O	0
			11	11	
5	B	11	Total	O	0
			11	11	
5	C	12	Total	O	0
			12	12	
5	D	12	Total	O	0
			12	12	

3 Residue-property plots [i](#)

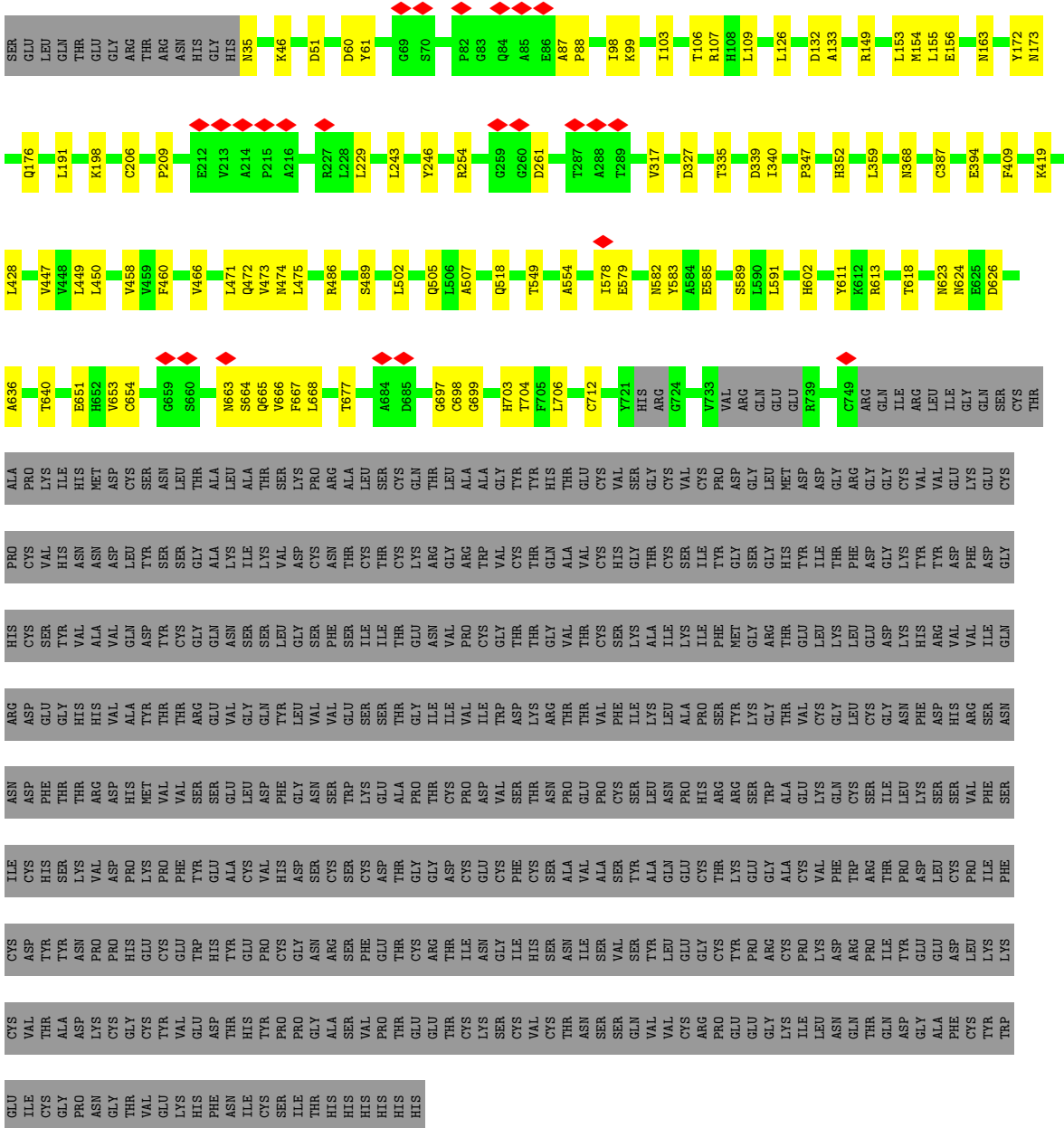
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and atom inclusion 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: Mucin-2

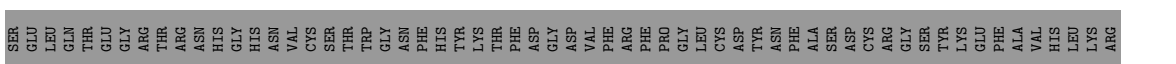


PRO ASN GLY THR VAL GLU LYS HIS PHE ASN ILE CYS SER ILE THR THR HIS HIS HIS HIS HIS HIS

• Molecule 1: Mucin-2



• Molecule 1: Mucin-2



MAG1
MAG2

4 Experimental information

Property	Value	Source
EM reconstruction method	HELICAL	Depositor
Imposed symmetry	HELICAL, twist=83.2°, rise=69.5 Å, axial sym=D2	Depositor
Number of segments used	515422	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 ($e^-/\text{Å}^2$)	48	Depositor
Minimum defocus (nm)	Not provided	
Maximum defocus (nm)	Not provided	
Magnification	Not provided	
Image detector	GATAN K3 BIOQUANTUM (6k x 4k)	Depositor
Maximum map value	0.205	Depositor
Minimum map value	-0.031	Depositor
Average map value	0.000	Depositor
Map value standard deviation	0.013	Depositor
Recommended contour level	0.04	Depositor
Map size (Å)	515.4, 515.4, 515.4	wwPDB
Map dimensions	600, 600, 600	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	0.859, 0.859, 0.859	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, CA

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	0.25	0/5394	0.49	0/7339
1	B	0.25	0/5405	0.49	0/7353
1	C	0.26	0/3448	0.50	0/4682
1	D	0.26	0/3448	0.50	0/4682
All	All	0.25	0/17695	0.49	0/24056

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	5267	4821	4821	60	0
1	B	5275	4834	4834	62	0
1	C	3362	3098	3097	26	0
1	D	3362	3098	3097	26	0
2	E	28	16	25	1	0
2	F	28	16	25	1	0
3	A	2	0	0	0	0
3	B	2	0	0	0	0
3	C	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	D	1	0	0	0	0
4	A	28	16	26	2	0
4	B	28	16	26	2	0
5	A	11	0	0	6	0
5	B	11	0	0	4	0
5	C	12	0	0	8	0
5	D	12	0	0	9	0
All	All	17430	15915	15951	175	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 5.

The worst 5 of 175 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:B:35:ASN:N	1:B:172:TYR:HH	1.58	1.00
1:A:484:VAL:HG12	1:A:494:VAL:HG22	1.44	0.97
1:C:1185:ASP:O	1:C:1217:SER:OG	1.85	0.93
1:B:173:ASN:O	4:B:2003:NAG:O3	1.99	0.81
1:B:103:ILE:HD11	1:B:126:LEU:HD11	1.61	0.81

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	704/1245 (56%)	633 (90%)	71 (10%)	0	100	100
1	B	705/1245 (57%)	633 (90%)	72 (10%)	0	100	100
1	C	433/1245 (35%)	395 (91%)	38 (9%)	0	100	100
1	D	433/1245 (35%)	398 (92%)	35 (8%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
All	All	2275/4980 (46%)	2059 (90%)	216 (10%)	0	100	100

There are no Ramachandran outliers to report.

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	561/1077 (52%)	558 (100%)	3 (0%)	88	94
1	B	562/1077 (52%)	558 (99%)	4 (1%)	84	91
1	C	379/1077 (35%)	378 (100%)	1 (0%)	92	96
1	D	379/1077 (35%)	379 (100%)	0	100	100
All	All	1881/4308 (44%)	1873 (100%)	8 (0%)	92	95

5 of 8 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	941	ARG
1	B	419	LYS
1	B	107[B]	ARG
1	B	107[A]	ARG
1	B	149	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	622	GLN
1	C	906	ASN
1	D	906	ASN

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
2	NAG	E	1	2,1	14,14,15	0.21	0	17,19,21	0.39	0
2	NAG	E	2	2	14,14,15	0.19	0	17,19,21	0.45	0
2	NAG	F	1	2,1	14,14,15	0.24	0	17,19,21	0.38	0
2	NAG	F	2	2	14,14,15	0.20	0	17,19,21	0.44	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
2	NAG	E	1	2,1	-	3/6/23/26	0/1/1/1
2	NAG	E	2	2	-	2/6/23/26	0/1/1/1
2	NAG	F	1	2,1	-	4/6/23/26	0/1/1/1
2	NAG	F	2	2	-	2/6/23/26	0/1/1/1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 11 torsion outliers are listed below:

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

Continued on next page...

Continued from previous page...

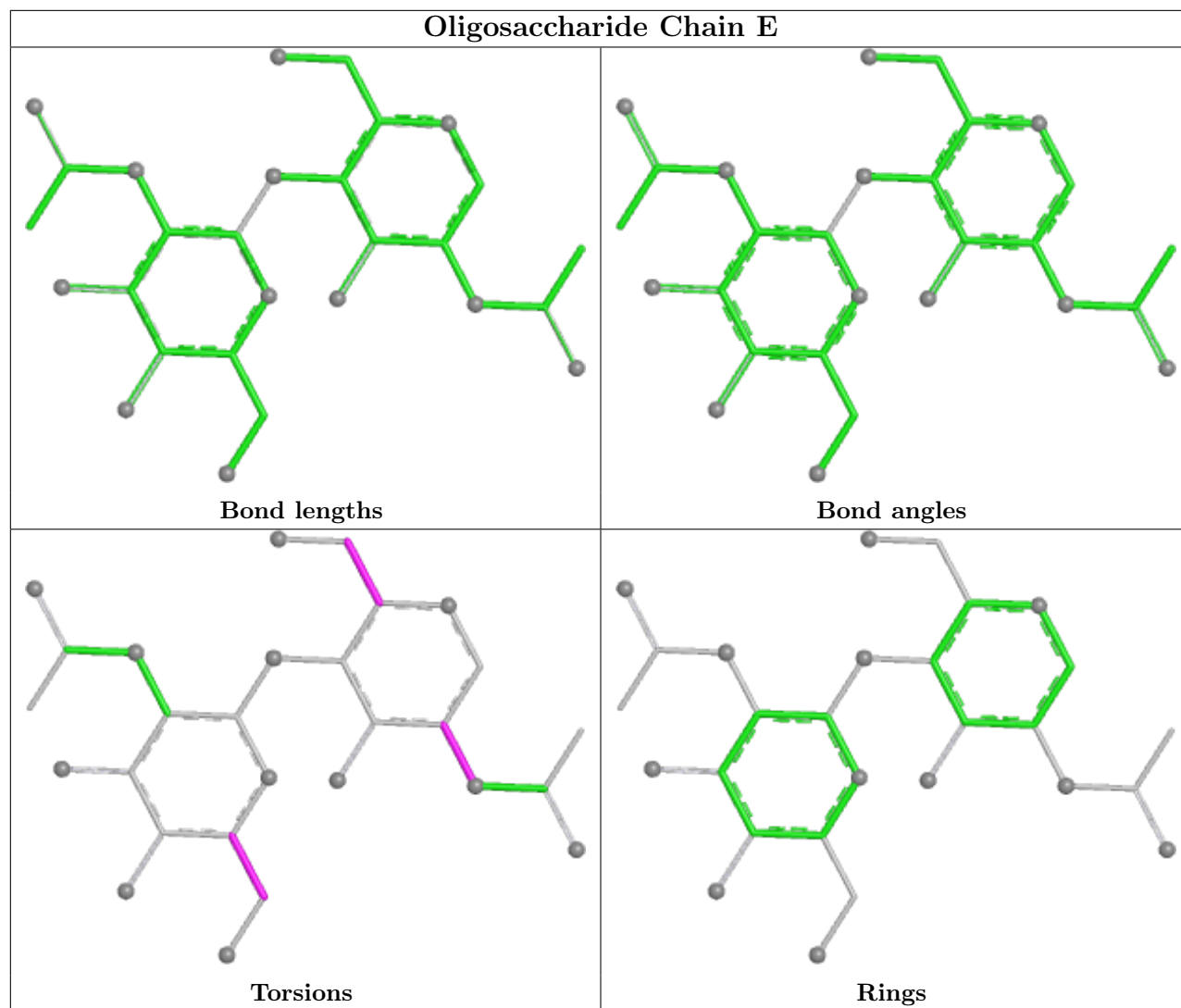
Mol	Chain	Res	Type	Atoms
2	F	1	NAG	C4-C5-C6-O6
2	F	2	NAG	O5-C5-C6-O6
2	F	2	NAG	C4-C5-C6-O6

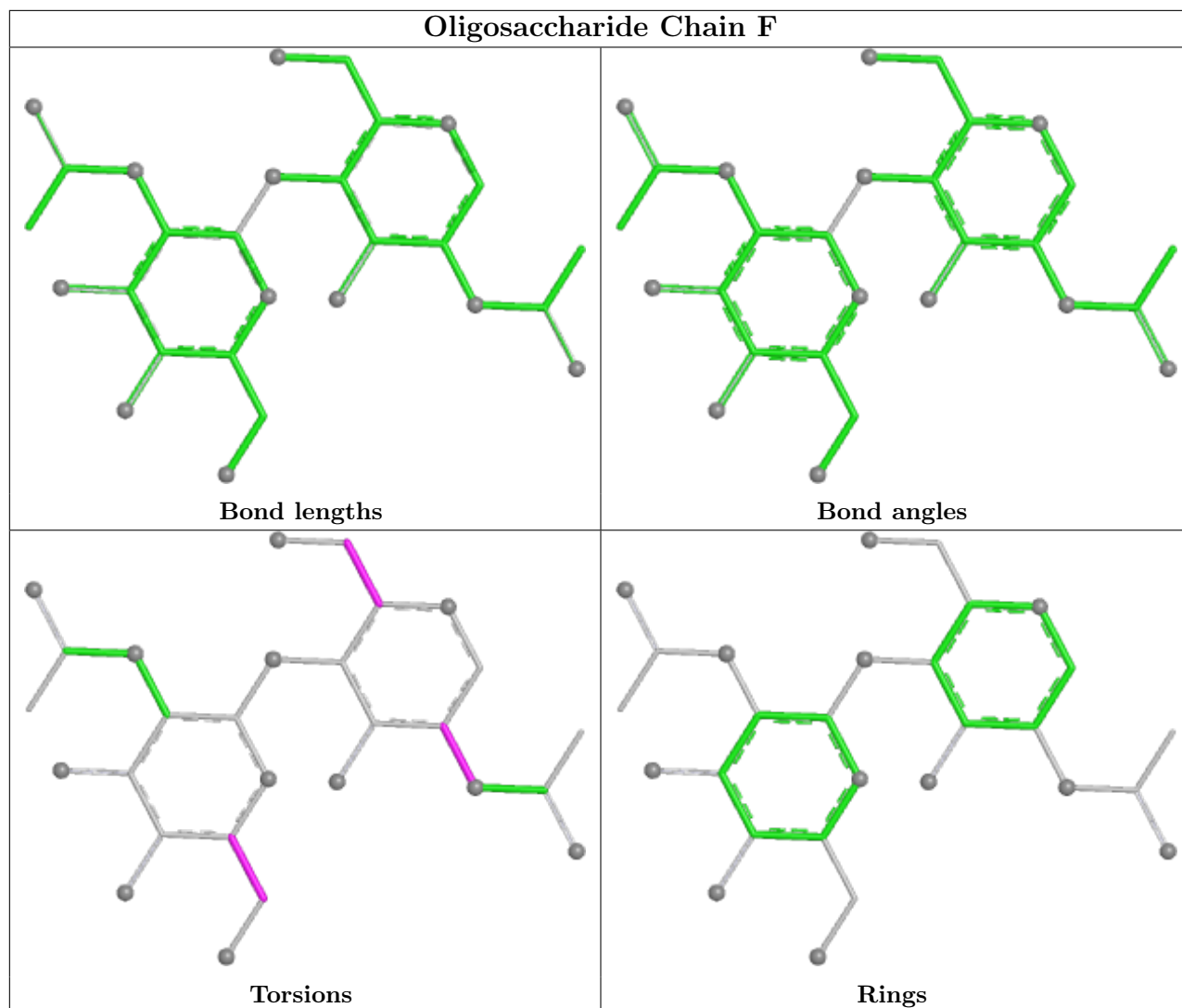
There are no ring outliers.

4 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	F	2	NAG	1	0
2	F	1	NAG	1	0
2	E	1	NAG	1	0
2	E	2	NAG	1	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](#)

Of 10 ligands modelled in this entry, 6 are monoatomic - leaving 4 for Mogul analysis.

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	B	2004	1	14,14,15	0.42	0	17,19,21	0.36	0
4	NAG	B	2003	1	14,14,15	0.68	1 (7%)	17,19,21	0.88	1 (5%)
4	NAG	A	2003	1	14,14,15	0.73	1 (7%)	17,19,21	0.95	1 (5%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
4	NAG	A	2004	1	14,14,15	0.40	0	17,19,21	0.36	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	B	2004	1	-	4/6/23/26	0/1/1/1
4	NAG	B	2003	1	-	2/6/23/26	0/1/1/1
4	NAG	A	2003	1	-	2/6/23/26	0/1/1/1
4	NAG	A	2004	1	-	4/6/23/26	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	2003	NAG	O5-C1	2.42	1.47	1.43
4	B	2003	NAG	O5-C1	2.20	1.47	1.43

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	2003	NAG	C1-O5-C5	3.67	117.17	112.19
4	B	2003	NAG	C1-O5-C5	3.34	116.72	112.19

There are no chirality outliers.

5 of 12 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	2003	NAG	O5-C5-C6-O6
4	A	2004	NAG	O5-C5-C6-O6
4	B	2004	NAG	O5-C5-C6-O6
4	A	2003	NAG	C4-C5-C6-O6
4	B	2004	NAG	C4-C5-C6-O6

There are no ring outliers.

2 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	B	2003	NAG	2	0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	2003	NAG	2	0

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.

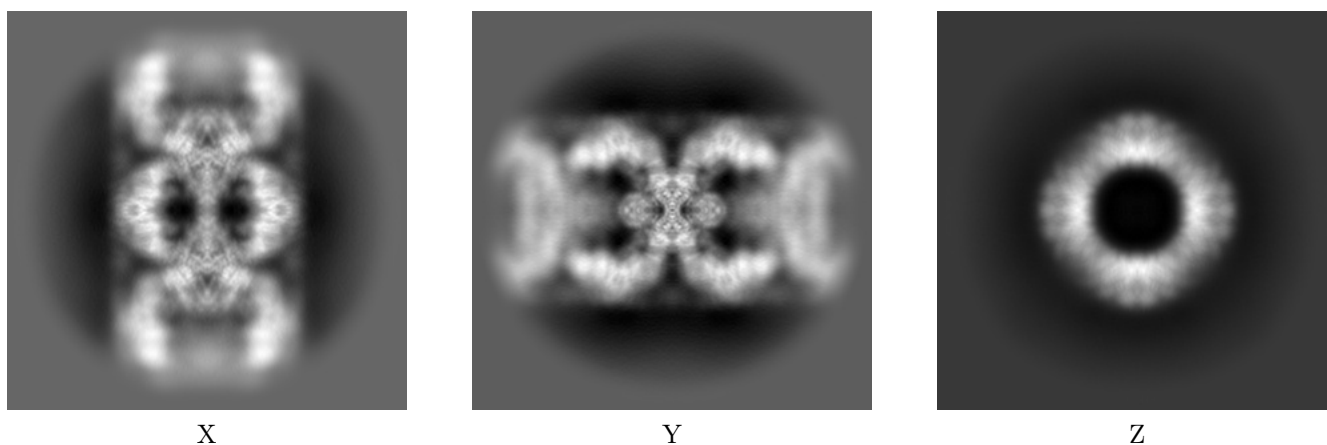
6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-13575. 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](#)

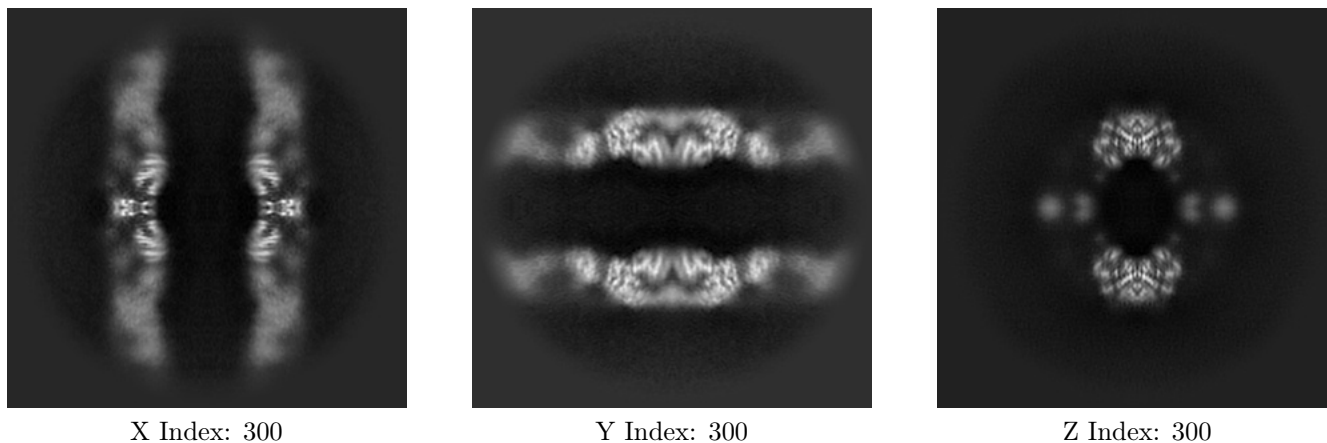
6.1.1 Primary map



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

6.2 Central slices [i](#)

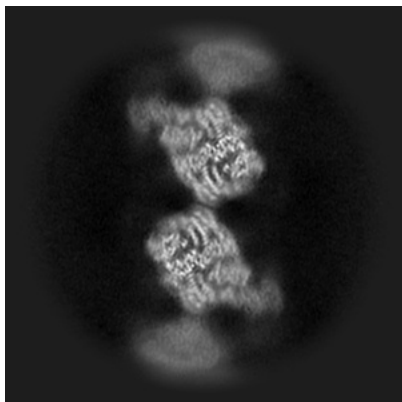
6.2.1 Primary map



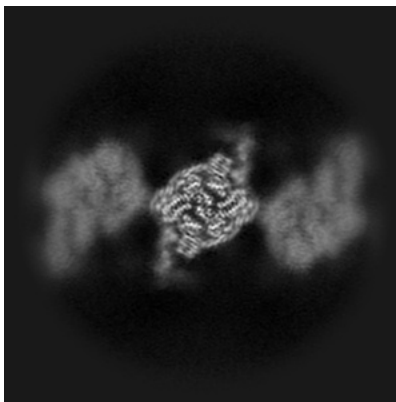
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices [i](#)

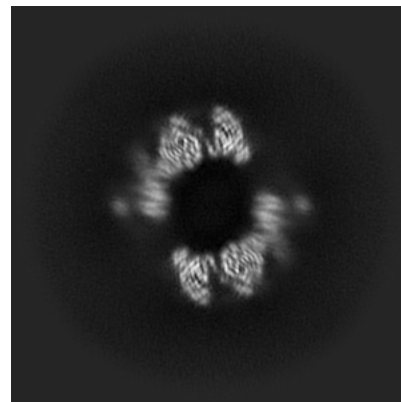
6.3.1 Primary map



X Index: 211



Y Index: 216



Z Index: 322

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

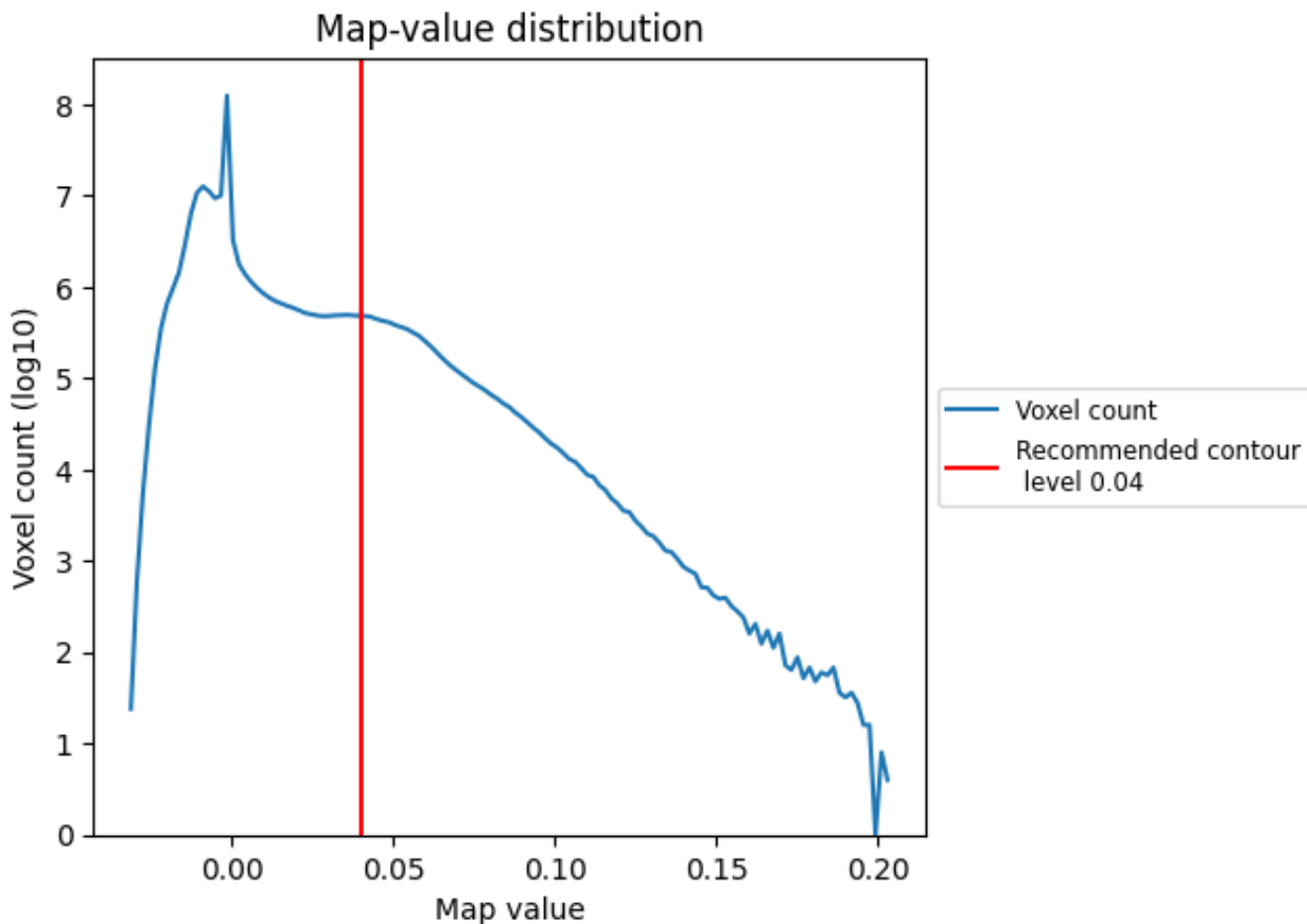
6.5 Mask visualisation

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

7 Map analysis [i](#)

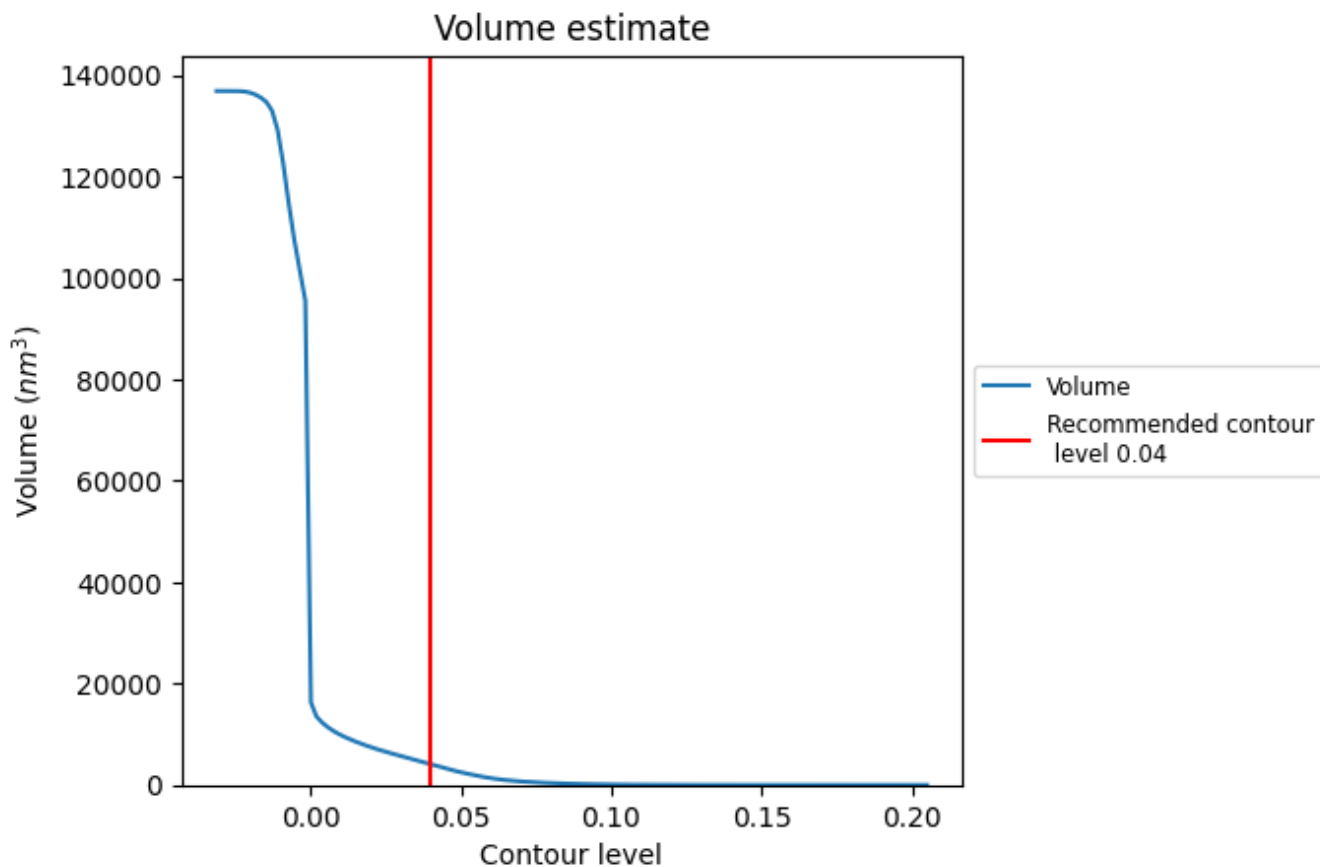
This section contains the results of statistical analysis of the map.

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.

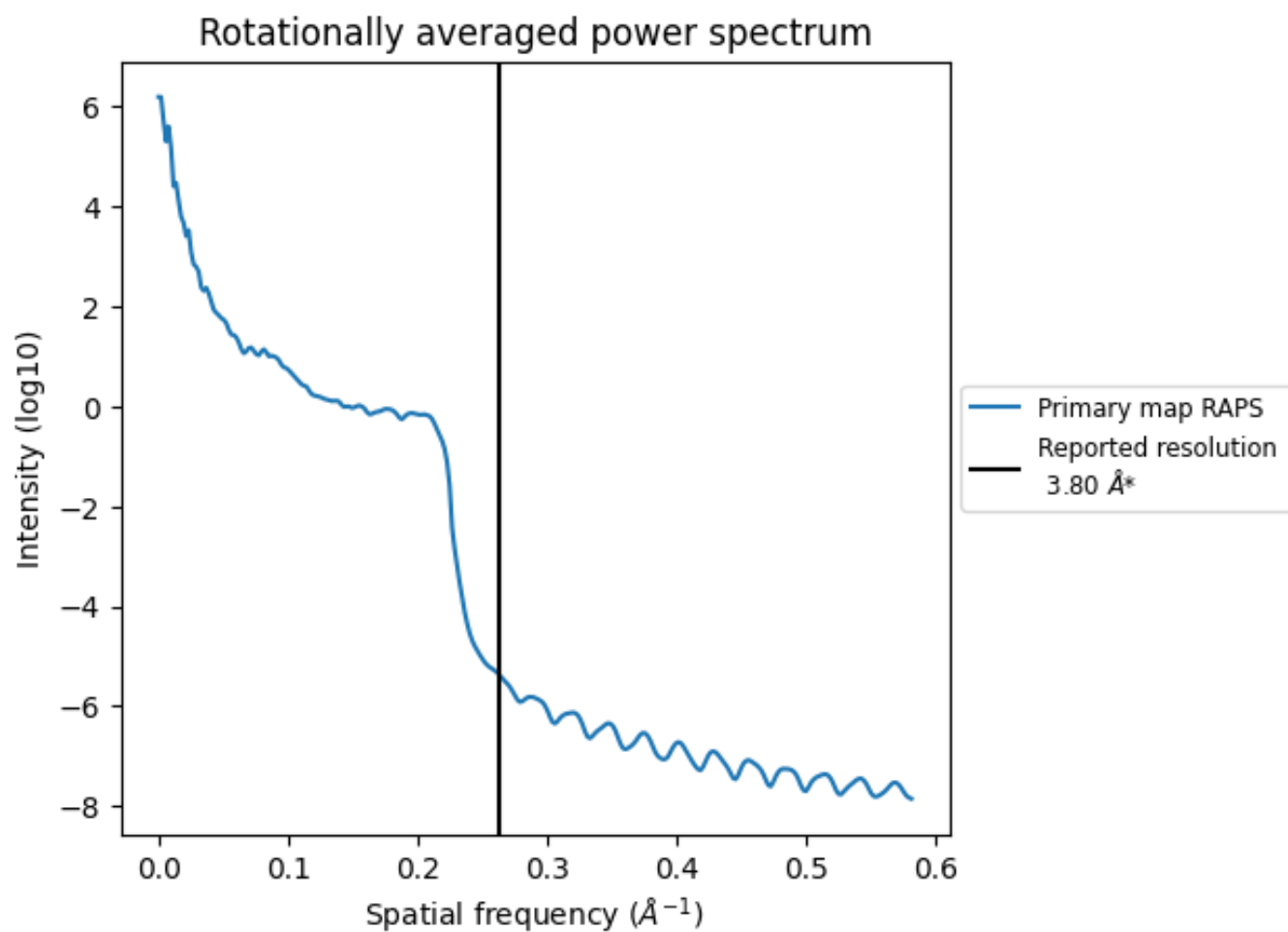
7.2 Volume estimate [\(i\)](#)



The volume at the recommended contour level is 40570 nm³; this corresponds to an approximate mass of 3665 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](#)

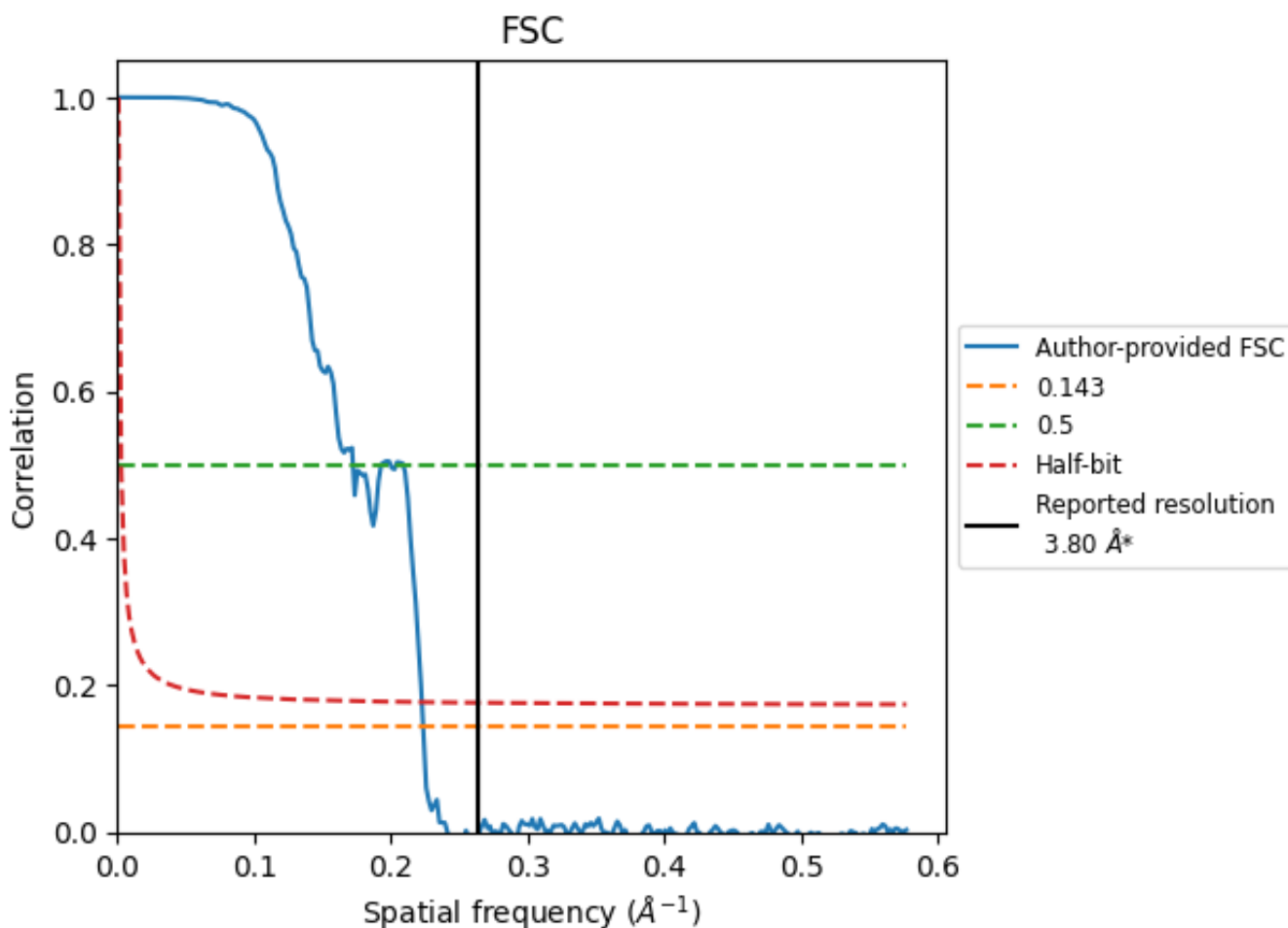


*Reported resolution corresponds to spatial frequency of 0.263 Å⁻¹

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 0.263 Å⁻¹

8.2 Resolution estimates [i](#)

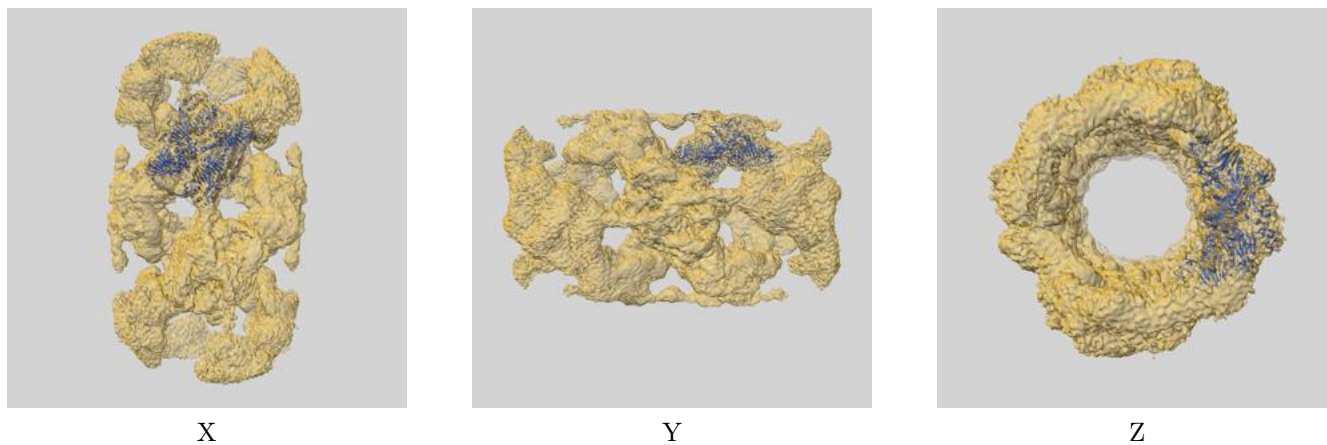
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	3.80	-	-
Author-provided FSC curve	4.47	5.80	4.49
Unmasked-calculated*	-	-	-

*Resolution estimate based on FSC curve calculated by comparison of deposited half-maps. The value from author-provided FSC intersecting FSC 0.143 CUT-OFF 4.47 differs from the reported value 3.8 by more than 10 %

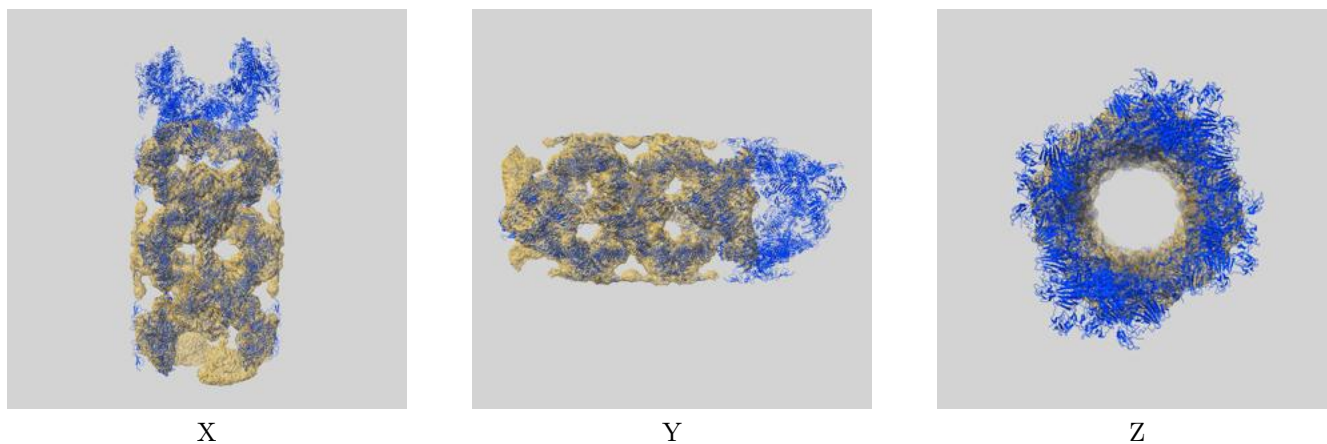
9 Map-model fit [i](#)

This section contains information regarding the fit between EMDB map EMD-13575 and PDB model 7POV. Per-residue inclusion information can be found in section 3 on page 6.

9.0.1 Map-model overlay [i](#)

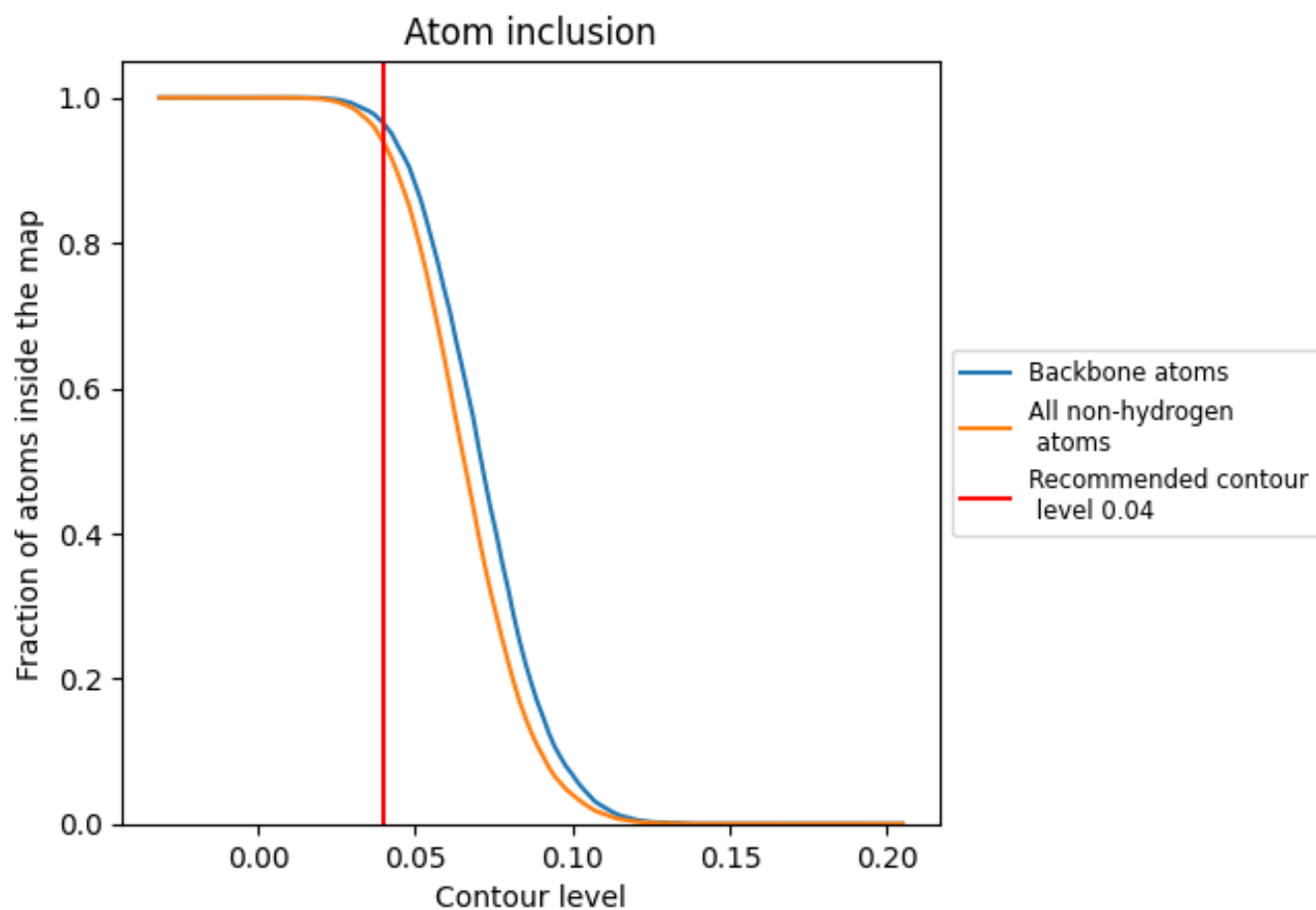


9.0.2 Map-model assembly overlay [i](#)



The images above show the 3D surface view of the map at the recommended contour level 0.04 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.

9.1 Atom inclusion [i](#)



At the recommended contour level, 96% of all backbone atoms, 94% of all non-hydrogen atoms, are inside the map.