



# Full wwPDB X-ray Structure Validation Report ⓘ

Jan 8, 2024 – 12:18 am GMT

PDB ID : 5N0C  
Title : Crystal structure of the tetanus neurotoxin in complex with GM1a  
Authors : Masuyer, G.; Conrad, J.; Stenmark, P.  
Deposited on : 2017-02-02  
Resolution : 2.60 Å(reported)

This is a Full wwPDB X-ray Structure 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/XrayValidationReportHelp>

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:

MolProbity : 4.02b-467  
Mogul : 1.8.4, CSD as541be (2020)  
Xtrriage (Phenix) : 1.13  
EDS : 2.36  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.36

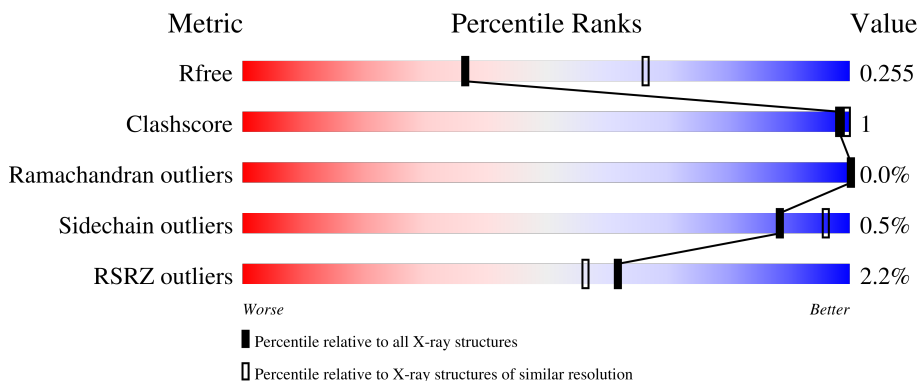
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

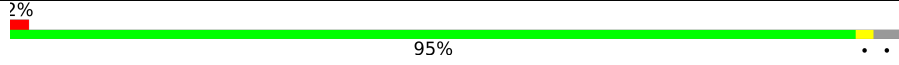
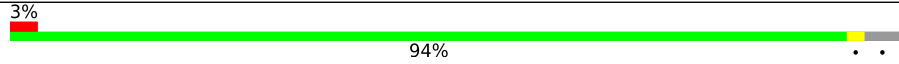


The reported resolution of this entry is 2.60 Å.

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)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	3163 (2.60-2.60)
Clashscore	141614	3518 (2.60-2.60)
Ramachandran outliers	138981	3455 (2.60-2.60)
Sidechain outliers	138945	3455 (2.60-2.60)
RSRZ outliers	127900	3104 (2.60-2.60)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower 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 electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	1335	 2% 95%
1	B	1335	 3% 94%
2	C	5	 20% 80%
2	D	5	 20% 80%

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 21439 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, 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 Tetanus toxin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	1293	10444	6696	1704	2009	35	0	0	0
1	B	1286	10398	6664	1697	2002	35	0	0	0

There are 44 discrepancies between the modelled and reference sequences:

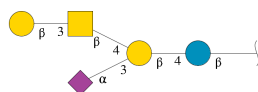
Chain	Residue	Modelled	Actual	Comment	Reference
A	-19	MET	-	initiating methionine	UNP P04958
A	-18	GLY	-	expression tag	UNP P04958
A	-17	SER	-	expression tag	UNP P04958
A	-16	SER	-	expression tag	UNP P04958
A	-15	HIS	-	expression tag	UNP P04958
A	-14	HIS	-	expression tag	UNP P04958
A	-13	HIS	-	expression tag	UNP P04958
A	-12	HIS	-	expression tag	UNP P04958
A	-11	HIS	-	expression tag	UNP P04958
A	-10	HIS	-	expression tag	UNP P04958
A	-9	SER	-	expression tag	UNP P04958
A	-8	SER	-	expression tag	UNP P04958
A	-7	GLY	-	expression tag	UNP P04958
A	-6	LEU	-	expression tag	UNP P04958
A	-5	VAL	-	expression tag	UNP P04958
A	-4	PRO	-	expression tag	UNP P04958
A	-3	ARG	-	expression tag	UNP P04958
A	-2	GLY	-	expression tag	UNP P04958
A	-1	SER	-	expression tag	UNP P04958
A	0	HIS	-	expression tag	UNP P04958
A	372	ALA	ARG	engineered mutation	UNP P04958
A	375	PHE	TYR	engineered mutation	UNP P04958
B	-19	MET	-	initiating methionine	UNP P04958
B	-18	GLY	-	expression tag	UNP P04958
B	-17	SER	-	expression tag	UNP P04958

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	-16	SER	-	expression tag	UNP P04958
B	-15	HIS	-	expression tag	UNP P04958
B	-14	HIS	-	expression tag	UNP P04958
B	-13	HIS	-	expression tag	UNP P04958
B	-12	HIS	-	expression tag	UNP P04958
B	-11	HIS	-	expression tag	UNP P04958
B	-10	HIS	-	expression tag	UNP P04958
B	-9	SER	-	expression tag	UNP P04958
B	-8	SER	-	expression tag	UNP P04958
B	-7	GLY	-	expression tag	UNP P04958
B	-6	LEU	-	expression tag	UNP P04958
B	-5	VAL	-	expression tag	UNP P04958
B	-4	PRO	-	expression tag	UNP P04958
B	-3	ARG	-	expression tag	UNP P04958
B	-2	GLY	-	expression tag	UNP P04958
B	-1	SER	-	expression tag	UNP P04958
B	0	HIS	-	expression tag	UNP P04958
B	372	ALA	ARG	engineered mutation	UNP P04958
B	375	PHE	TYR	engineered mutation	UNP P04958

- Molecule 2 is an oligosaccharide called beta-D-galactopyranose-(1-3)-2-acetamido-2-deoxy-beta-D-galactopyranose-(1-4)-[N-acetyl-alpha-neuraminic acid-(2-3)]beta-D-galactopyranose-(1-4)-beta-D-glucopyranose.



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	C	5	Total	C	N	O	0	0	0
			68	37	2	29			
2	D	5	Total	C	N	O	0	0	0
			68	37	2	29			

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total	Zn	0	0
			1	1		
3	B	1	Total	Zn	0	0
			1	1		

- Molecule 4 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: C<sub>4</sub>H<sub>10</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C O 7 4 3	0	0

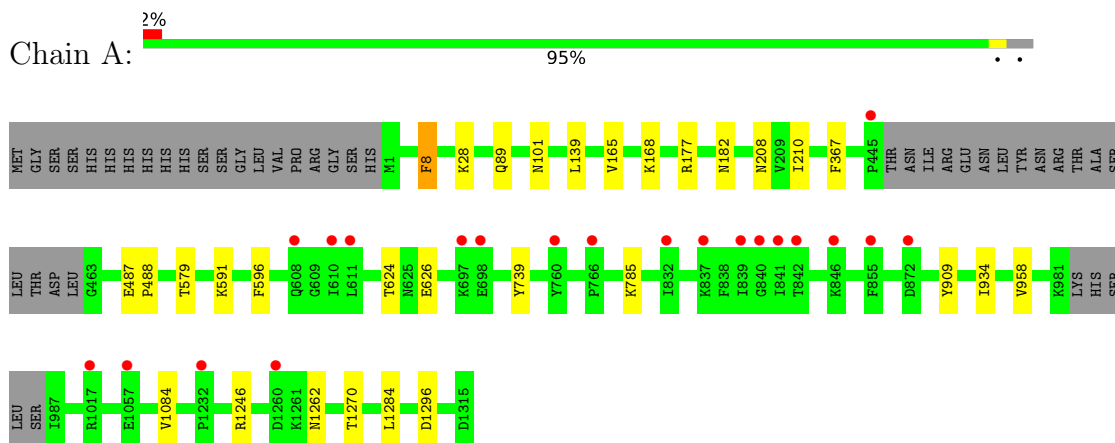
- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	242	Total O 242 242	0	0
5	B	210	Total O 210 210	0	0

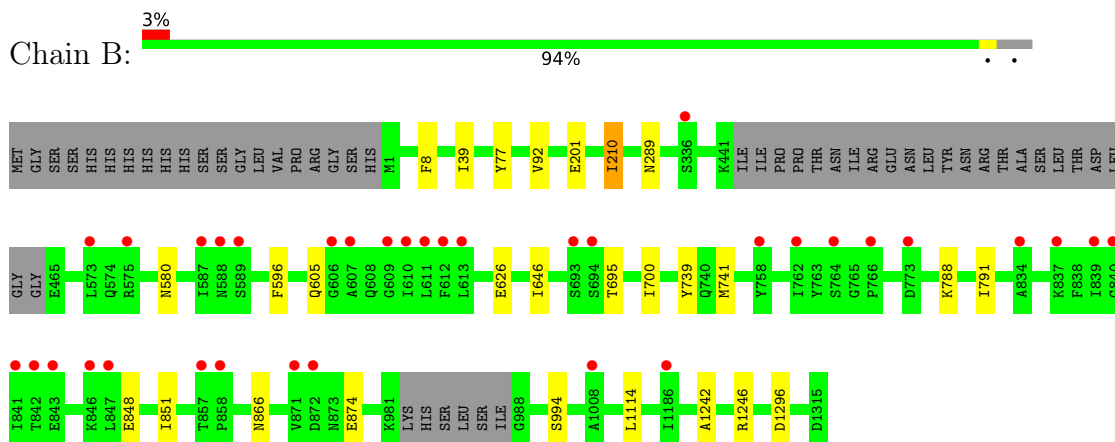
### 3 Residue-property plots

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 electron 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 dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). 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: Tetanus toxin



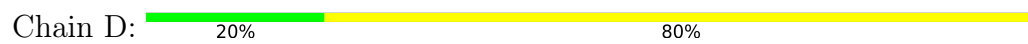
- Molecule 1: Tetanus toxin



- Molecule 2: beta-D-galactopyranose-(1-3)-2-acetamido-2-deoxy-beta-D-galactopyranose-(1-4)-[N-acetyl-alpha-neuraminic acid-(2-3)]beta-D-galactopyranose-(1-4)-beta-D-glucopyranose



- Molecule 2: beta-D-galactopyranose-(1-3)-2-acetamido-2-deoxy-beta-D-galactopyranose-(1-4)-[N-acetyl-alpha-neuraminic acid-(2-3)]beta-D-galactopyranose-(1-4)-beta-D-glucopyranose



8GG1
GALL2
8GK3
GALL4
8IA5

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	92.23Å 137.45Å 152.40Å 90.00° 90.12° 90.00°	Depositor
Resolution (Å)	68.50 – 2.60 68.48 – 2.60	Depositor EDS
% Data completeness (in resolution range)	99.8 (68.50-2.60) 99.8 (68.48-2.60)	Depositor EDS
$R_{merge}$	0.13	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.98 (at 2.62Å)	Xtrriage
Refinement program	REFMAC 5.8.0158	Depositor
R, $R_{free}$	0.221 , 0.253 0.224 , 0.255	Depositor DCC
$R_{free}$ test set	5742 reflections (4.93%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	36.6	Xtrriage
Anisotropy	0.098	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 41.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.29$	Xtrriage
Estimated twinning fraction	0.031 for h,-k,-l	Xtrriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	21439	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	49.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 34.92 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 6.3269e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.



## 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: ZN, GAL, BGC, NGA, PEG, SIA

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	0.40	0/10667	0.56	0/14454
1	B	0.40	0/10619	0.55	0/14387
All	All	0.40	0/21286	0.55	0/28841

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	10444	0	10375	14	0
1	B	10398	0	10322	11	0
2	C	68	0	58	0	0
2	D	68	0	58	0	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
4	A	7	0	10	0	0
5	A	242	0	0	0	0
5	B	210	0	0	0	0
All	All	21439	0	20823	25	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 1.

All (25) 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:210:ILE:HD11	1:B:739:TYR:CE2	2.35	0.62
1:A:579:THR:HG21	1:A:591:LYS:HE2	1.91	0.51
1:B:1246:ARG:NH2	1:B:1296:ASP:O	2.43	0.51
1:A:1270:THR:HG22	1:A:1284:LEU:HD12	1.92	0.51
1:A:210:ILE:HD11	1:A:739:TYR:CE2	2.47	0.49
1:A:958:VAL:HG12	1:A:1084:VAL:HG12	1.95	0.49
1:B:788:LYS:HA	1:B:791:ILE:HD12	1.97	0.47
1:B:39:ILE:CD1	1:B:92:VAL:HG22	2.45	0.46
1:A:1246:ARG:NH2	1:A:1296:ASP:O	2.49	0.46
1:A:177:ARG:NH1	1:A:182:ASN:OD1	2.50	0.45
1:B:646:ILE:HG22	1:B:741:MET:SD	2.57	0.45
1:A:909:TYR:CE1	1:A:934:ILE:HD12	2.52	0.45
1:A:579:THR:HG21	1:A:591:LYS:CE	2.46	0.44
1:A:624:THR:HG23	1:A:785:LYS:HD2	2.00	0.42
1:B:596:PHE:CD1	1:B:626:GLU:HB3	2.55	0.42
1:B:848:GLU:HA	1:B:851:ILE:HG22	2.01	0.41
1:B:77:TYR:OH	1:B:201:GLU:OE1	2.28	0.41
1:A:8:PHE:O	1:A:89:GLN:NE2	2.46	0.41
1:A:487:GLU:N	1:A:488:PRO:CD	2.84	0.41
1:B:210:ILE:HD11	1:B:739:TYR:CZ	2.55	0.41
1:A:101:ASN:HB3	1:A:367:PHE:CZ	2.56	0.41
1:A:596:PHE:CD1	1:A:626:GLU:HB3	2.56	0.41
1:B:695:THR:HG21	1:B:700:ILE:HD12	2.02	0.41
1:B:1114:LEU:CD1	1:B:1242:ALA:HB1	2.52	0.40
1:A:165:VAL:HG11	1:A:168:LYS:HG3	2.02	0.40

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 X-ray entries followed by that with respect to entries of similar resolution.

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	1287/1335 (96%)	1242 (96%)	45 (4%)	0	100	100
1	B	1280/1335 (96%)	1232 (96%)	47 (4%)	1 (0%)	51	75
All	All	2567/2670 (96%)	2474 (96%)	92 (4%)	1 (0%)	100	100

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	605	GLN

### 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 X-ray entries followed by that with respect to entries of similar resolution.

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	1178/1216 (97%)	1173 (100%)	5 (0%)	91	97
1	B	1173/1216 (96%)	1166 (99%)	7 (1%)	86	95
All	All	2351/2432 (97%)	2339 (100%)	12 (0%)	88	96

All (12) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	8	PHE
1	A	28	LYS
1	A	139	LEU
1	A	208	ASN
1	A	1262	ASN
1	B	8	PHE
1	B	210	ILE
1	B	289	ASN
1	B	580	ASN
1	B	866	ASN
1	B	874	GLU
1	B	994	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (17) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	113	ASN
1	A	208	ASN
1	A	608	GLN
1	A	770	GLN
1	A	780	ASN
1	A	903	ASN
1	A	972	ASN
1	A	998	ASN
1	A	1078	ASN
1	B	278	GLN
1	B	348	GLN
1	B	421	ASN
1	B	580	ASN
1	B	588	ASN
1	B	1068	ASN
1	B	1219	ASN
1	B	1293	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](#)

10 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	BGC	C	1	2	12,12,12	0.78	0	17,17,17	1.39	4 (23%)
2	GAL	C	2	2	11,11,12	0.53	0	15,15,17	1.36	3 (20%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	NGA	C	3	2	14,14,15	0.40	0	17,19,21	0.91	0
2	GAL	C	4	2	11,11,12	0.32	0	15,15,17	1.13	1 (6%)
2	SIA	C	5	2	20,20,21	0.69	0	24,28,31	1.51	4 (16%)
2	BGC	D	1	2	12,12,12	0.75	0	17,17,17	1.38	4 (23%)
2	GAL	D	2	2	11,11,12	0.49	0	15,15,17	1.20	3 (20%)
2	NGA	D	3	2	14,14,15	0.32	0	17,19,21	1.01	0
2	GAL	D	4	2	11,11,12	0.33	0	15,15,17	0.81	1 (6%)
2	SIA	D	5	2	20,20,21	0.66	0	24,28,31	1.35	3 (12%)

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	BGC	C	1	2	-	0/2/22/22	0/1/1/1
2	GAL	C	2	2	-	2/2/19/22	0/1/1/1
2	NGA	C	3	2	-	1/6/23/26	0/1/1/1
2	GAL	C	4	2	-	2/2/19/22	0/1/1/1
2	SIA	C	5	2	-	4/18/34/38	0/1/1/1
2	BGC	D	1	2	-	0/2/22/22	0/1/1/1
2	GAL	D	2	2	-	2/2/19/22	0/1/1/1
2	NGA	D	3	2	-	2/6/23/26	0/1/1/1
2	GAL	D	4	2	-	2/2/19/22	0/1/1/1
2	SIA	D	5	2	-	3/18/34/38	0/1/1/1

There are no bond length outliers.

All (23) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	5	SIA	C6-O6-C2	3.81	119.49	111.34
2	C	4	GAL	O5-C5-C6	3.68	112.97	107.20
2	C	1	BGC	C1-O5-C5	3.49	120.24	113.66
2	C	5	SIA	C11-C10-N5	3.43	121.90	116.10
2	D	1	BGC	C1-O5-C5	2.99	119.30	113.66
2	D	5	SIA	C6-O6-C2	2.96	117.67	111.34
2	C	2	GAL	O5-C5-C6	2.86	111.69	107.20
2	C	2	GAL	C1-O5-C5	2.80	115.98	112.19
2	D	4	GAL	O5-C5-C6	2.65	111.36	107.20

Continued on next page...

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	5	SIA	C11-C10-N5	2.59	120.49	116.10
2	D	5	SIA	O1B-C1-C2	2.46	120.05	113.03
2	D	2	GAL	C1-O5-C5	2.39	115.43	112.19
2	D	1	BGC	O4-C4-C3	2.38	115.85	110.35
2	D	2	GAL	C1-C2-C3	2.36	112.57	109.67
2	D	1	BGC	C1-C2-C3	2.35	115.18	110.31
2	C	5	SIA	C5-N5-C10	2.33	128.84	123.18
2	D	1	BGC	O5-C1-C2	2.33	114.44	110.28
2	C	5	SIA	O10-C10-C11	-2.26	117.86	122.06
2	C	2	GAL	C1-C2-C3	2.20	112.37	109.67
2	C	1	BGC	O5-C5-C4	2.18	113.65	109.69
2	C	1	BGC	C1-C2-C3	2.10	114.68	110.31
2	D	2	GAL	O3-C3-C2	-2.05	106.06	109.99
2	C	1	BGC	O4-C4-C5	2.03	114.33	109.30

There are no chirality outliers.

All (18) torsion outliers are listed below:

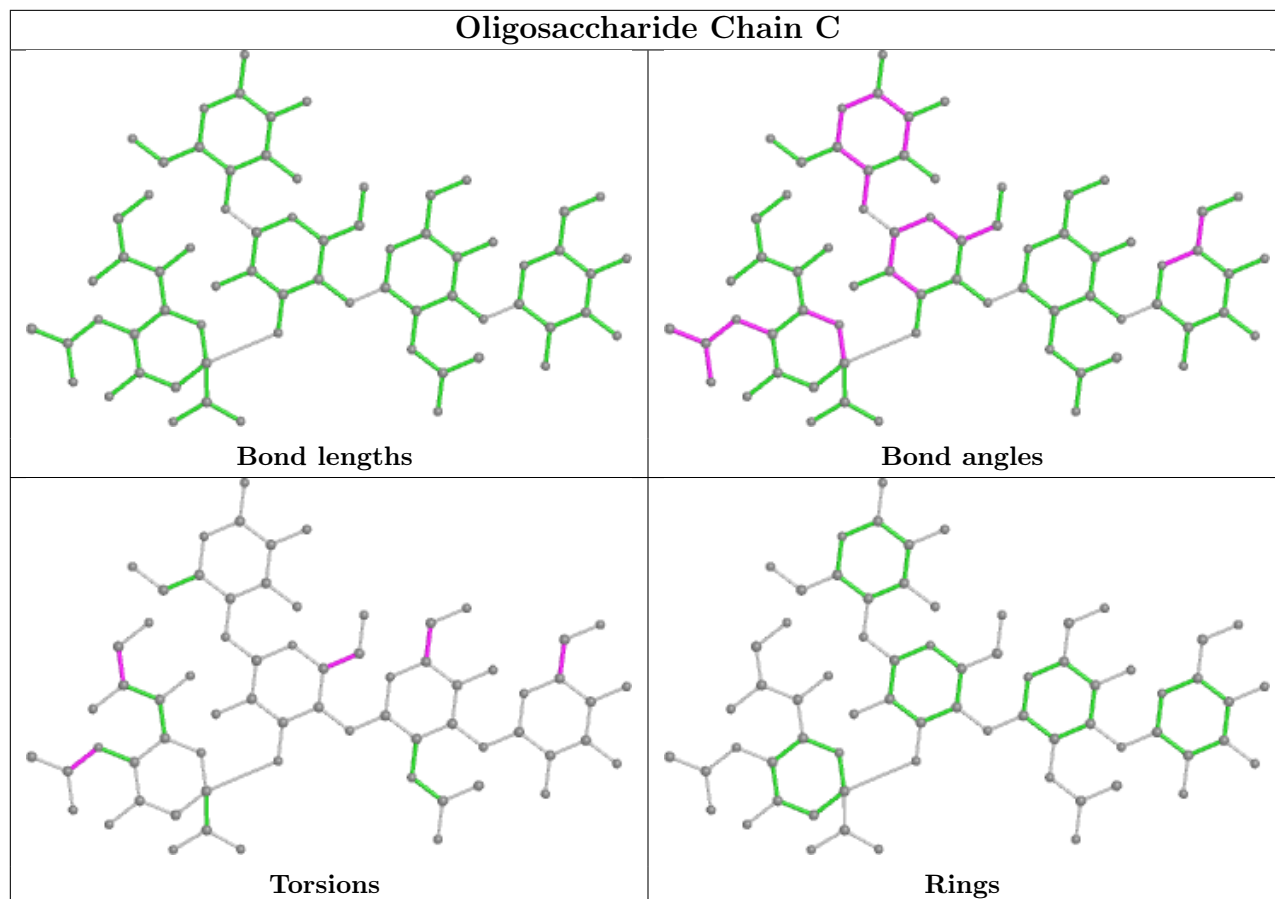
Mol	Chain	Res	Type	Atoms
2	C	5	SIA	O8-C8-C9-O9
2	C	4	GAL	O5-C5-C6-O6
2	C	2	GAL	O5-C5-C6-O6
2	C	5	SIA	C7-C8-C9-O9
2	C	4	GAL	C4-C5-C6-O6
2	D	3	NGA	O5-C5-C6-O6
2	D	2	GAL	C4-C5-C6-O6
2	C	2	GAL	C4-C5-C6-O6
2	C	5	SIA	C11-C10-N5-C5
2	C	5	SIA	O10-C10-N5-C5
2	D	5	SIA	C11-C10-N5-C5
2	D	5	SIA	O10-C10-N5-C5
2	D	2	GAL	O5-C5-C6-O6
2	D	4	GAL	C4-C5-C6-O6
2	D	4	GAL	O5-C5-C6-O6
2	D	3	NGA	C4-C5-C6-O6
2	C	3	NGA	O5-C5-C6-O6
2	D	5	SIA	C6-C7-C8-O8

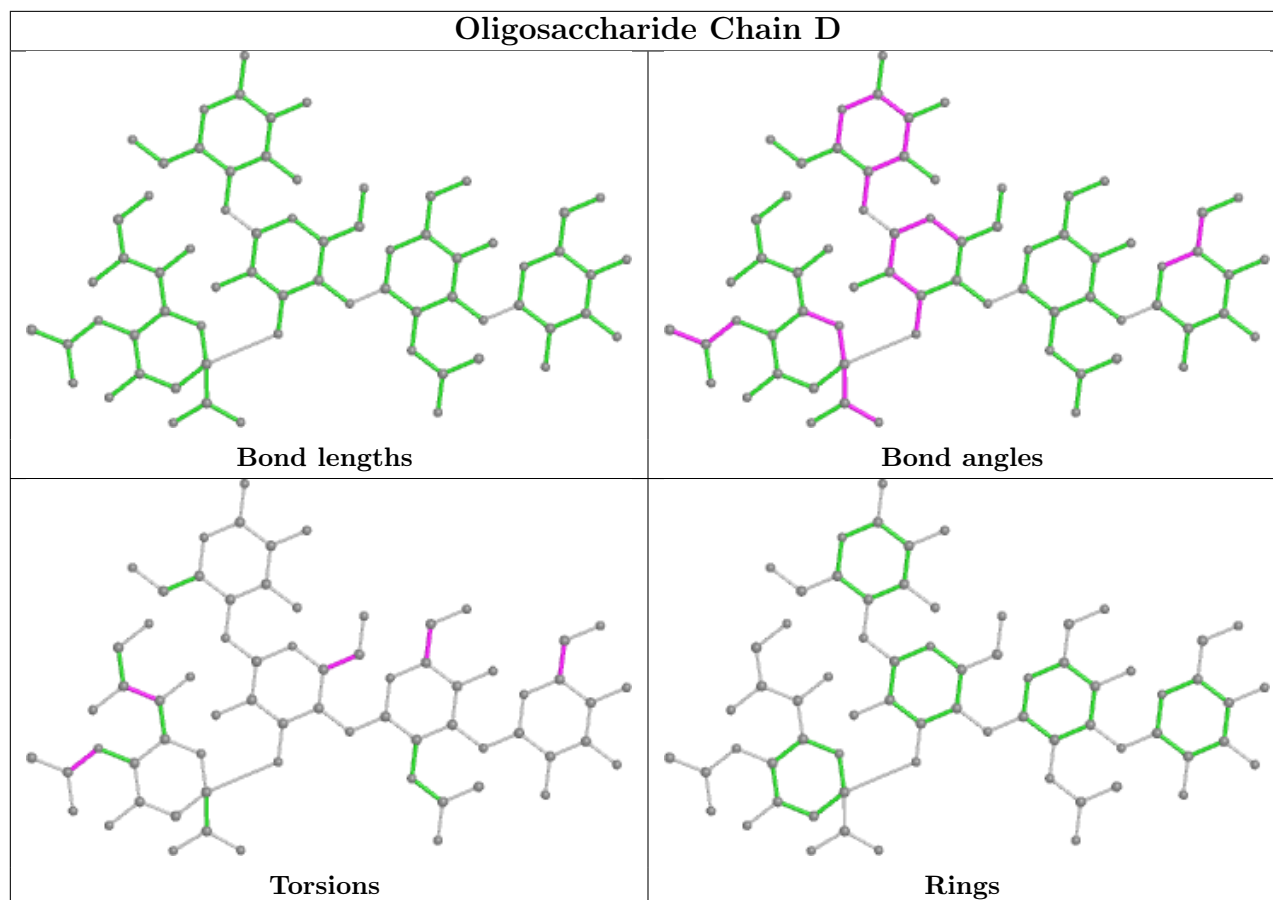
There are no ring outliers.

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](#)

Of 3 ligands modelled in this entry, 2 are monoatomic - leaving 1 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	PEG	A	1407	-	6,6,6	0.44	0	5,5,5	0.33	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. '2' means no outliers of that kind were identified.



Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	PEG	A	1407	-	-	0/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 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 Fit of model and data

### 6.1 Protein, DNA and RNA chains

In the following table, the column labelled ‘#RSRZ > 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q < 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	1293/1335 (96%)	-0.23	21 (1%) 72 68	18, 41, 90, 138	0
1	B	1286/1335 (96%)	-0.10	35 (2%) 54 48	16, 48, 103, 159	0
All	All	2579/2670 (96%)	-0.17	56 (2%) 62 56	16, 44, 96, 159	0

All (56) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	841	ILE	9.1
1	A	842	THR	8.2
1	B	841	ILE	6.9
1	A	839	ILE	5.8
1	B	610	ILE	5.8
1	B	609	GLY	5.6
1	B	840	GLY	5.4
1	A	611	LEU	5.4
1	B	842	THR	4.7
1	B	837	LYS	4.5
1	B	613	LEU	4.4
1	B	764	SER	4.3
1	B	858	PRO	4.3
1	B	606	GLY	3.8
1	B	846	LYS	3.5
1	B	607	ALA	3.4
1	B	693	SER	3.3
1	B	758	TYR	3.2
1	B	336	SER	3.2
1	A	445	PRO	3.2
1	A	608	GLN	3.1
1	A	837	LYS	3.1
1	B	587	ILE	3.0
1	A	840	GLY	2.9

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	610	ILE	2.9
1	B	1186	ILE	2.9
1	B	843	GLU	2.9
1	B	762	ILE	2.9
1	B	839	ILE	2.8
1	A	1017	ARG	2.7
1	B	766	PRO	2.6
1	B	1008	ALA	2.6
1	B	575	ARG	2.6
1	A	766	PRO	2.6
1	A	846	LYS	2.5
1	A	760	TYR	2.5
1	B	573	LEU	2.5
1	B	611	LEU	2.5
1	B	589	SER	2.5
1	B	847	LEU	2.4
1	A	1232	PRO	2.4
1	B	871	VAL	2.4
1	A	698	GLU	2.3
1	A	697	LYS	2.3
1	B	872	ASP	2.2
1	B	857	THR	2.2
1	A	872	ASP	2.1
1	A	1057	GLU	2.1
1	A	855	PHE	2.1
1	B	588	ASN	2.1
1	B	834	ALA	2.1
1	A	1260	ASP	2.1
1	B	773	ASP	2.1
1	B	612	PHE	2.1
1	A	832	ILE	2.0
1	B	694	SER	2.0

## 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

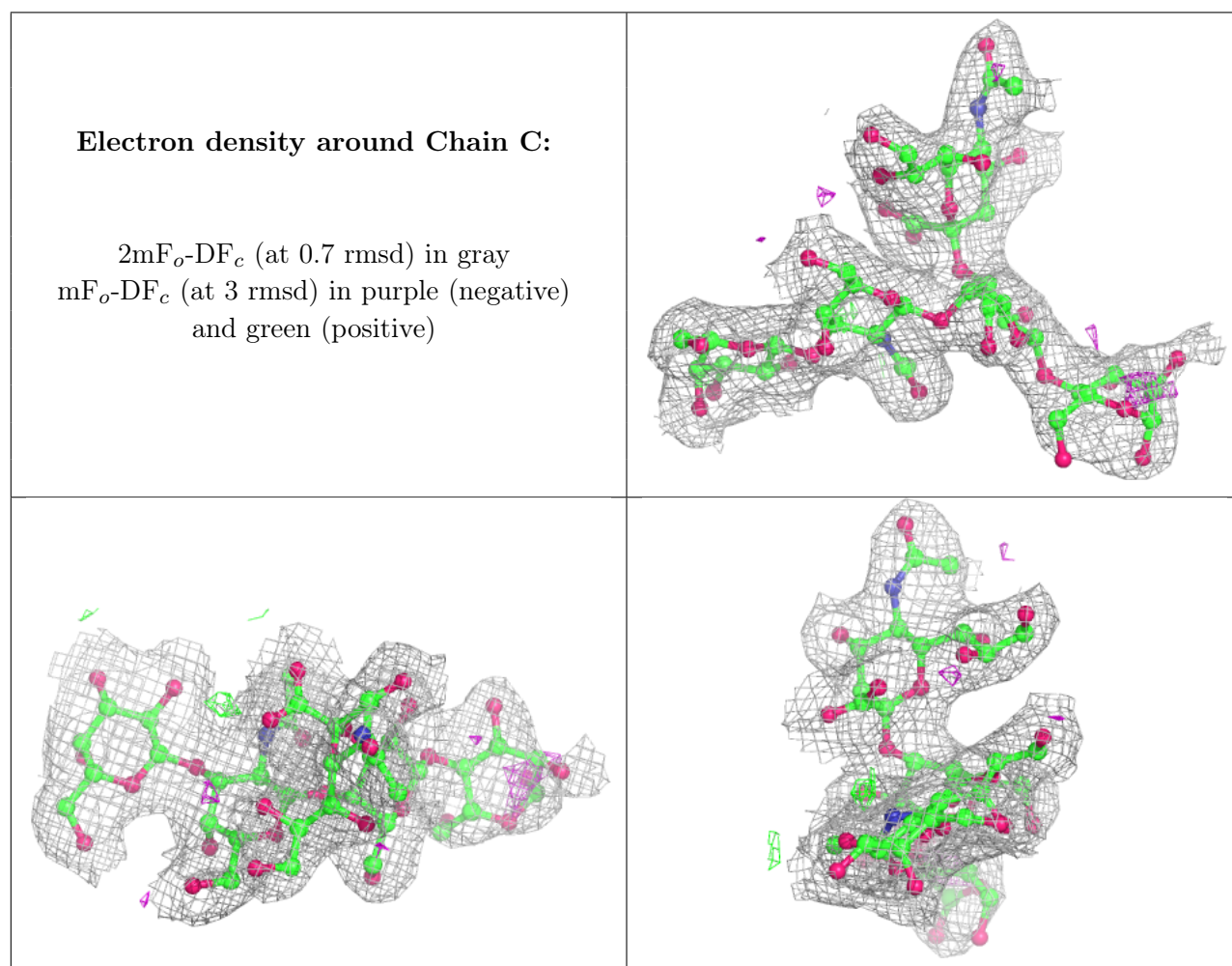
## 6.3 Carbohydrates [i](#)

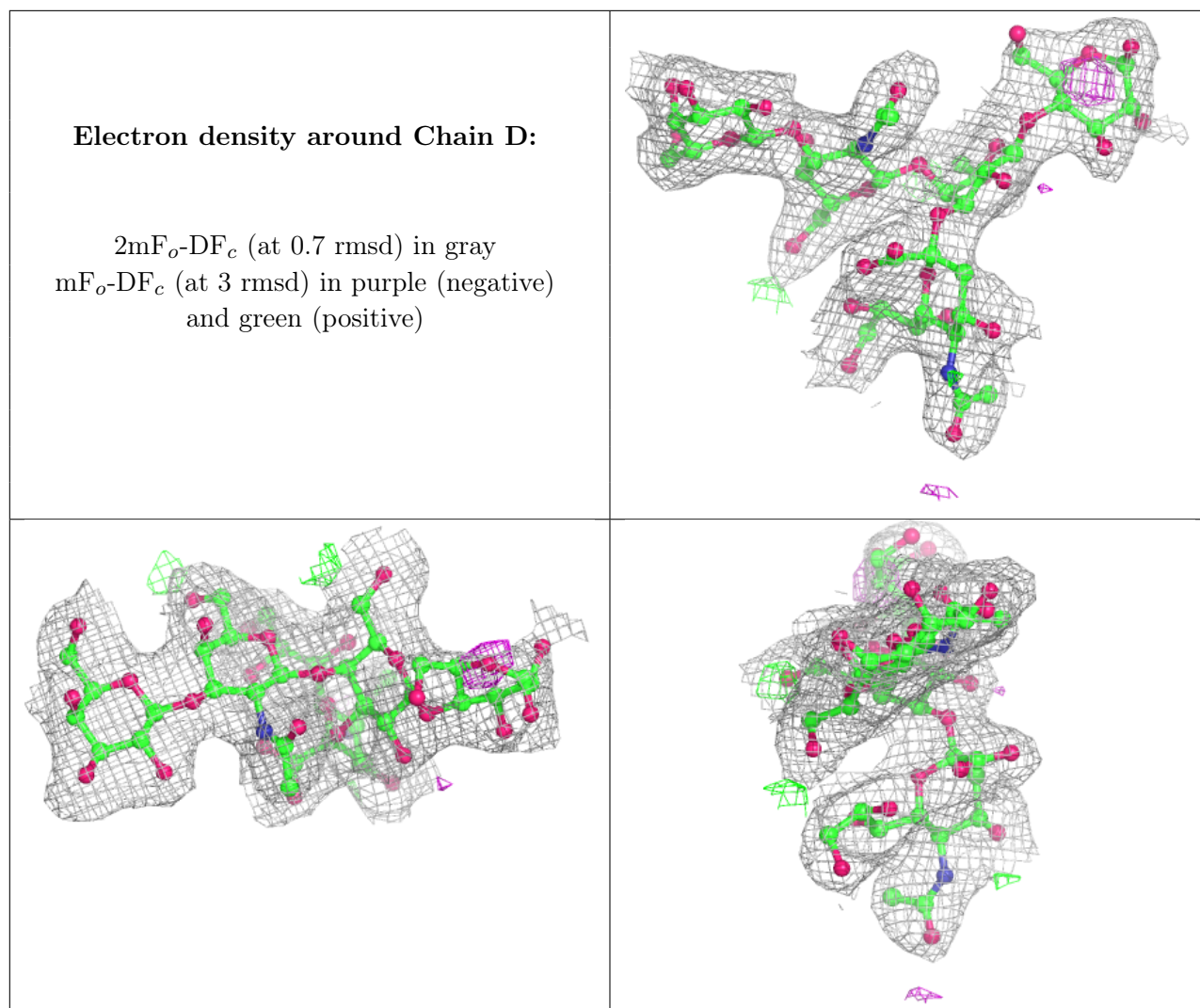
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum,

median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q < 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
2	BGC	C	1	12/12	0.85	0.25	57,68,70,70	0
2	BGC	D	1	12/12	0.86	0.27	60,68,70,72	0
2	GAL	D	2	11/12	0.93	0.10	46,49,52,54	0
2	SIA	D	5	20/21	0.94	0.13	40,43,46,47	0
2	NGA	D	3	14/15	0.95	0.11	40,41,42,43	0
2	SIA	C	5	20/21	0.95	0.16	41,43,44,44	0
2	GAL	C	2	11/12	0.96	0.12	40,44,47,49	0
2	GAL	C	4	11/12	0.97	0.11	32,32,32,33	0
2	GAL	D	4	11/12	0.97	0.08	37,38,38,39	0
2	NGA	C	3	14/15	0.97	0.12	32,35,36,37	0

The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.





## 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
4	PEG	A	1407	7/7	0.84	0.17	63,67,71,72	0
3	ZN	B	1406	1/1	0.99	0.10	52,52,52,52	0
3	ZN	A	1406	1/1	0.99	0.14	42,42,42,42	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.