



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 21, 2020 – 09:09 PM BST

PDB ID : 6E57  
Title : Bacteroides ovatus mixed-linkage glucan utilization locus (MLGUL) SGBP-B  
in complex with mixed-linkage heptasaccharide  
Authors : Koropatkin, N.M.; Schnizlein, M.; Bahr, C.M.E.  
Deposited on : 2018-07-19  
Resolution : 2.71 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.13.1  
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) : 2.13.1

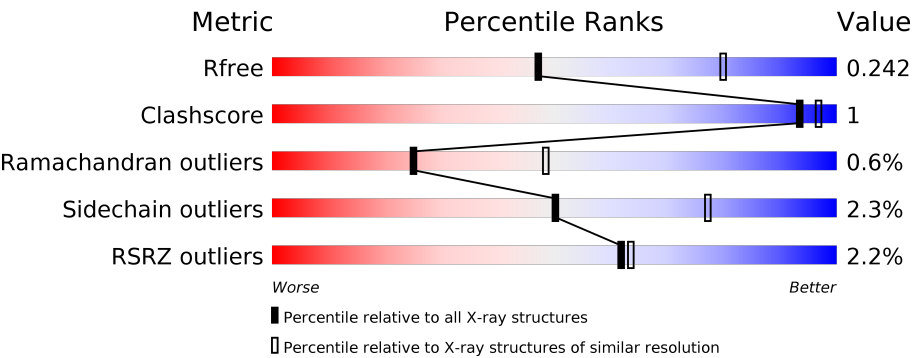
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.71 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.



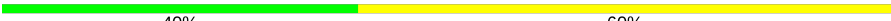
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	3359 (2.74-2.70)
Clashscore	141614	3686 (2.74-2.70)
Ramachandran outliers	138981	3622 (2.74-2.70)
Sidechain outliers	138945	3623 (2.74-2.70)
RSRZ outliers	127900	3276 (2.74-2.70)

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 on 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	420	<div><div></div><div>89%5%5%</div></div>
1	B	420	<div>5%<div></div><div>87%5%8%</div></div>
1	C	420	<div>%<div></div><div>88%•7%</div></div>
1	D	420	<div>2%<div></div><div>88%•8%</div></div>
2	E	5	<div><div></div><div>80%20%</div></div>
2	F	5	<div><div></div><div>40%60%</div></div>

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	G	5	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	BGC	G	1	-	-	-	X

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 12333 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 surface glycan binding protein B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	399	Total	C	N	O	S	0	0	0
			3050	1910	495	637	8			
1	B	385	Total	C	N	O	S	0	0	0
			2941	1842	479	612	8			
1	C	390	Total	C	N	O	S	0	0	0
			2979	1865	486	620	8			
1	D	388	Total	C	N	O	S	0	0	0
			2963	1856	484	615	8			

There are 84 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	initiating methionine	UNP A7LY28
A	2	GLY	-	expression tag	UNP A7LY28
A	3	SER	-	expression tag	UNP A7LY28
A	4	SER	-	expression tag	UNP A7LY28
A	5	HIS	-	expression tag	UNP A7LY28
A	6	HIS	-	expression tag	UNP A7LY28
A	7	HIS	-	expression tag	UNP A7LY28
A	8	HIS	-	expression tag	UNP A7LY28
A	9	HIS	-	expression tag	UNP A7LY28
A	10	HIS	-	expression tag	UNP A7LY28
A	11	SER	-	expression tag	UNP A7LY28
A	12	SER	-	expression tag	UNP A7LY28
A	13	GLY	-	expression tag	UNP A7LY28
A	14	LEU	-	expression tag	UNP A7LY28
A	15	VAL	-	expression tag	UNP A7LY28
A	16	PRO	-	expression tag	UNP A7LY28
A	17	ARG	-	expression tag	UNP A7LY28
A	18	GLY	-	expression tag	UNP A7LY28
A	19	SER	-	expression tag	UNP A7LY28
A	20	HIS	-	expression tag	UNP A7LY28
A	21	MET	-	expression tag	UNP A7LY28

*Continued on next page...*

*Continued from previous page...*

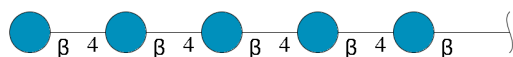
Chain	Residue	Modelled	Actual	Comment	Reference
B	1	MET	-	initiating methionine	UNP A7LY28
B	2	GLY	-	expression tag	UNP A7LY28
B	3	SER	-	expression tag	UNP A7LY28
B	4	SER	-	expression tag	UNP A7LY28
B	5	HIS	-	expression tag	UNP A7LY28
B	6	HIS	-	expression tag	UNP A7LY28
B	7	HIS	-	expression tag	UNP A7LY28
B	8	HIS	-	expression tag	UNP A7LY28
B	9	HIS	-	expression tag	UNP A7LY28
B	10	HIS	-	expression tag	UNP A7LY28
B	11	SER	-	expression tag	UNP A7LY28
B	12	SER	-	expression tag	UNP A7LY28
B	13	GLY	-	expression tag	UNP A7LY28
B	14	LEU	-	expression tag	UNP A7LY28
B	15	VAL	-	expression tag	UNP A7LY28
B	16	PRO	-	expression tag	UNP A7LY28
B	17	ARG	-	expression tag	UNP A7LY28
B	18	GLY	-	expression tag	UNP A7LY28
B	19	SER	-	expression tag	UNP A7LY28
B	20	HIS	-	expression tag	UNP A7LY28
B	21	MET	-	expression tag	UNP A7LY28
C	2	MET	-	initiating methionine	UNP A7LY28
C	3	GLY	-	expression tag	UNP A7LY28
C	4	SER	-	expression tag	UNP A7LY28
C	5	SER	-	expression tag	UNP A7LY28
C	6	HIS	-	expression tag	UNP A7LY28
C	7	HIS	-	expression tag	UNP A7LY28
C	8	HIS	-	expression tag	UNP A7LY28
C	9	HIS	-	expression tag	UNP A7LY28
C	10	HIS	-	expression tag	UNP A7LY28
C	11	HIS	-	expression tag	UNP A7LY28
C	12	SER	-	expression tag	UNP A7LY28
C	13	SER	-	expression tag	UNP A7LY28
C	14	GLY	-	expression tag	UNP A7LY28
C	15	LEU	-	expression tag	UNP A7LY28
C	16	VAL	-	expression tag	UNP A7LY28
C	17	PRO	-	expression tag	UNP A7LY28
C	18	ARG	-	expression tag	UNP A7LY28
C	19	GLY	-	expression tag	UNP A7LY28
C	20	SER	-	expression tag	UNP A7LY28
C	21	HIS	-	expression tag	UNP A7LY28
C	22	MET	-	expression tag	UNP A7LY28

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	1	MET	-	initiating methionine	UNP A7LY28
D	2	GLY	-	expression tag	UNP A7LY28
D	3	SER	-	expression tag	UNP A7LY28
D	4	SER	-	expression tag	UNP A7LY28
D	5	HIS	-	expression tag	UNP A7LY28
D	6	HIS	-	expression tag	UNP A7LY28
D	7	HIS	-	expression tag	UNP A7LY28
D	8	HIS	-	expression tag	UNP A7LY28
D	9	HIS	-	expression tag	UNP A7LY28
D	10	HIS	-	expression tag	UNP A7LY28
D	11	SER	-	expression tag	UNP A7LY28
D	12	SER	-	expression tag	UNP A7LY28
D	13	GLY	-	expression tag	UNP A7LY28
D	14	LEU	-	expression tag	UNP A7LY28
D	15	VAL	-	expression tag	UNP A7LY28
D	16	PRO	-	expression tag	UNP A7LY28
D	17	ARG	-	expression tag	UNP A7LY28
D	18	GLY	-	expression tag	UNP A7LY28
D	19	SER	-	expression tag	UNP A7LY28
D	20	HIS	-	expression tag	UNP A7LY28
D	21	MET	-	expression tag	UNP A7LY28

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



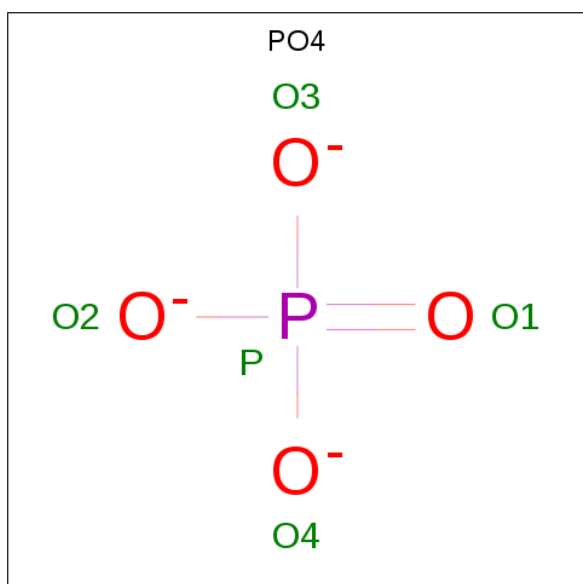
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace
2	E	5	Total	C	O	0	0	0
			56	30	26			
2	F	5	Total	C	O	0	0	0
			56	30	26			
2	G	5	Total	C	O	0	0	0
			56	30	26			

- Molecule 3 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			4	2	2		
3	A	1	Total	C	O	0	0
			4	2	2		
3	B	1	Total	C	O	0	0
			4	2	2		
3	C	1	Total	C	O	0	0
			4	2	2		
3	D	1	Total	C	O	0	0
			4	2	2		

- Molecule 4 is PHOSPHATE ION (three-letter code: PO4) (formula: O<sub>4</sub>P).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total O P 5 4 1	0	0
4	A	1	Total O P 5 4 1	0	0
4	A	1	Total O P 5 4 1	0	0
4	A	1	Total O P 5 4 1	0	0
4	B	1	Total O P 5 4 1	0	0
4	B	1	Total O P 5 4 1	0	0
4	B	1	Total O P 5 4 1	0	0
4	C	1	Total O P 5 4 1	0	0
4	C	1	Total O P 5 4 1	0	0
4	C	1	Total O P 5 4 1	0	0
4	C	1	Total O P 5 4 1	0	0
4	D	1	Total O P 5 4 1	0	0
4	D	1	Total O P 5 4 1	0	0
4	D	1	Total O P 5 4 1	0	0

- Molecule 5 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total Na 1 1	0	0
5	C	1	Total Na 1 1	0	0

- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	48	Total O 48 48	0	0
6	B	19	Total O 19 19	0	0

*Continued on next page...*



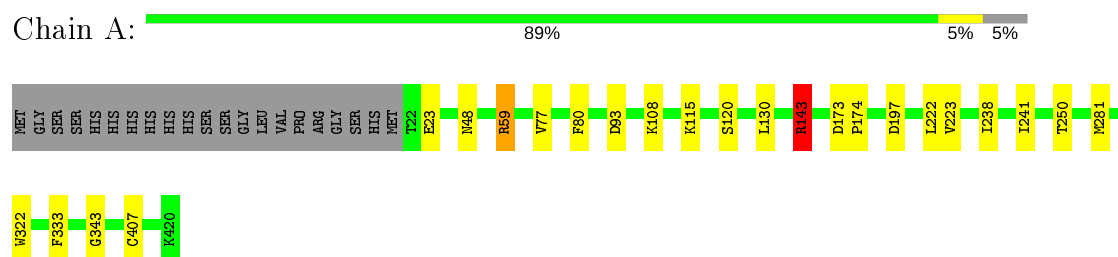
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	C	45	Total 45	O 45	0	0
6	D	28	Total 28	O 28	0	0

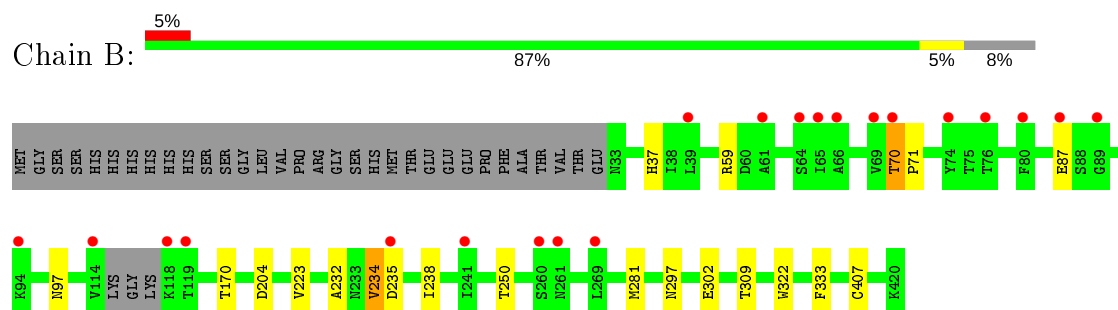
### 3 Residue-property plots

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.

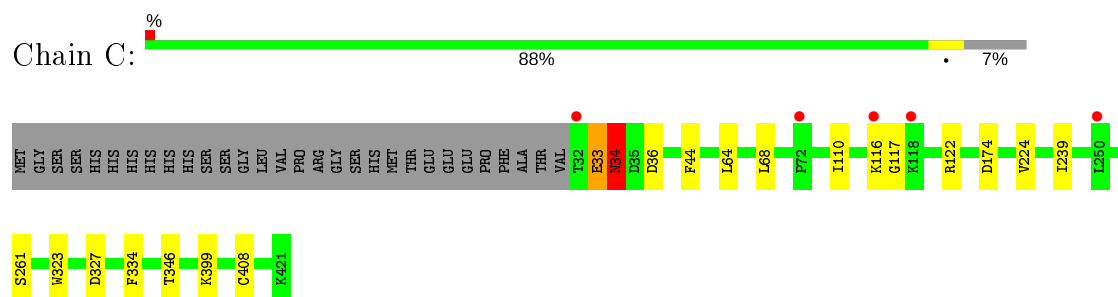
- Molecule 1: surface glycan binding protein B



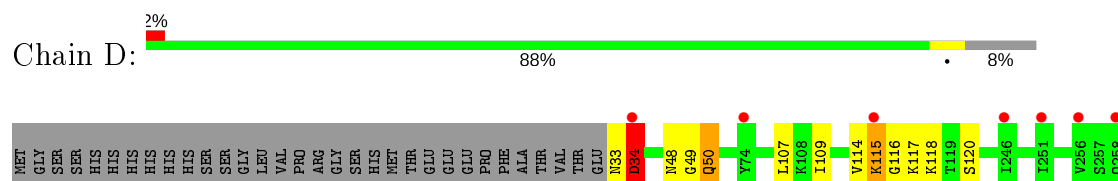
- Molecule 1: surface glycan binding protein B



- Molecule 1: surface glycan binding protein B



- Molecule 1: surface glycan binding protein B





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

Chain E: 80% 20%



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

Chain F: 40% 60%



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

Chain G: 40% 60%



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	156.44Å 243.65Å 76.06Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	78.34 – 2.71 78.22 – 2.71	Depositor EDS
% Data completeness (in resolution range)	97.7 (78.34-2.71) 97.7 (78.22-2.71)	Depositor EDS
$R_{merge}$	0.12	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.36 (at 2.69Å)	Xtriage
Refinement program	REFMAC 5.8.0222	Depositor
R, $R_{free}$	0.194 , 0.238 0.200 , 0.242	Depositor DCC
$R_{free}$ test set	3927 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	49.9	Xtriage
Anisotropy	0.113	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 27.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	12333	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	58.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.76% of the height of the origin peak. No significant pseudotranslation is detected.*

<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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: NA, PO4, BGC, EDO

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	0.30	0/3106	0.52	0/4240
1	B	0.30	0/2994	0.51	0/4087
1	C	0.30	0/3033	0.52	0/4139
1	D	0.30	0/3017	0.53	0/4117
All	All	0.30	0/12150	0.52	0/16583

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	1
1	D	0	1
All	All	0	2

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	143	ARG	Sidechain
1	D	118	LYS	Peptide

## 5.2 Too-close contacts ⓘ

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	3050	0	2952	11	0
1	B	2941	0	2847	7	0
1	C	2979	0	2890	7	0
1	D	2963	0	2877	9	0
2	E	56	0	48	0	0
2	F	56	0	48	0	0
2	G	56	0	48	0	0
3	A	8	0	12	0	0
3	B	4	0	6	0	0
3	C	4	0	6	0	0
3	D	4	0	6	0	0
4	A	20	0	0	0	0
4	B	15	0	0	0	0
4	C	20	0	0	0	0
4	D	15	0	0	0	0
5	A	1	0	0	0	0
5	C	1	0	0	0	0
6	A	48	0	0	0	0
6	B	19	0	0	0	0
6	C	45	0	0	0	0
6	D	28	0	0	0	0
All	All	12333	0	11740	33	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 (33) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:33:GLU:HA	1:C:34:ASN:HB3	1.56	0.86
1:C:33:GLU:HA	1:C:34:ASN:CB	2.19	0.71
1:B:70:THR:HG22	1:B:71:PRO:HD3	1.75	0.67
1:D:33:ASN:HA	1:D:34:ASP:CB	2.34	0.56
1:A:143:ARG:CG	1:A:143:ARG:HH11	2.21	0.53
1:B:333:PHE:HB3	1:B:407:CYS:HB2	1.91	0.52

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:223:VAL:HG13	1:A:238:ILE:HG23	1.94	0.50
1:A:333:PHE:HB3	1:A:407:CYS:HB2	1.94	0.49
1:D:33:ASN:HA	1:D:34:ASP:HB2	1.93	0.49
1:C:33:GLU:CA	1:C:34:ASN:CB	2.90	0.48
1:C:334:PHE:HB3	1:C:408:CYS:HB2	1.95	0.48
1:A:343:GLY:HA3	1:B:302:GLU:HB2	1.95	0.47
1:A:222:LEU:HB3	1:A:241:ILE:HG22	1.95	0.47
1:D:33:ASN:CA	1:D:34:ASP:HB2	2.45	0.47
1:B:232:ALA:HB2	1:B:309:THR:HG22	1.95	0.47
1:B:223:VAL:HG13	1:B:238:ILE:HG23	1.96	0.46
1:A:143:ARG:HH11	1:A:143:ARG:HG2	1.82	0.45
1:A:59:ARG:NH1	1:A:130:LEU:HD12	2.33	0.44
1:A:80:PHE:HB2	1:A:108:LYS:HB2	1.99	0.44
1:C:68:LEU:HD21	1:C:110:ILE:HG21	1.99	0.43
1:D:373:THR:HG21	1:D:398:LYS:HE3	2.01	0.43
1:B:250:THR:O	1:B:281:MET:HA	2.20	0.42
1:A:77:VAL:HG22	1:A:93:ASP:O	2.19	0.42
1:D:115:LYS:HA	1:D:116:GLY:HA2	1.68	0.42
1:A:173:ASP:N	1:A:174:PRO:CD	2.83	0.42
1:B:234:VAL:HG22	1:B:235:ASP:N	2.35	0.41
1:D:114:VAL:O	1:D:115:LYS:C	2.59	0.41
1:D:109:ILE:O	1:D:120:SER:HA	2.19	0.41
1:D:333:PHE:HB3	1:D:407:CYS:HB2	2.02	0.41
1:D:49:GLY:H	1:D:50:GLN:HA	1.85	0.41
1:C:224:VAL:HG13	1:C:239:ILE:HG23	2.03	0.41
1:A:250:THR:O	1:A:281:MET:HA	2.21	0.41
1:C:44:PHE:HD2	1:C:122:ARG:HH11	1.69	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles ⓘ

### 5.3.1 Protein backbone ⓘ

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	397/420 (94%)	380 (96%)	17 (4%)	0	100	100
1	B	381/420 (91%)	361 (95%)	19 (5%)	1 (0%)	41	65
1	C	388/420 (92%)	369 (95%)	14 (4%)	5 (1%)	12	28
1	D	386/420 (92%)	360 (93%)	23 (6%)	3 (1%)	19	41
All	All	1552/1680 (92%)	1470 (95%)	73 (5%)	9 (1%)	25	48

All (9) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	34	ASN
1	C	116	LYS
1	C	261	SER
1	D	115	LYS
1	D	117	LYS
1	C	36	ASP
1	D	34	ASP
1	B	234	VAL
1	C	117	GLY

### 5.3.2 Protein sidechains ⓘ

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	340/358 (95%)	332 (98%)	8 (2%)	49	76
1	B	328/358 (92%)	319 (97%)	9 (3%)	44	72
1	C	332/358 (93%)	324 (98%)	8 (2%)	49	76
1	D	330/358 (92%)	325 (98%)	5 (2%)	65	85
All	All	1330/1432 (93%)	1300 (98%)	30 (2%)	50	77

All (30) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	23	GLU
1	A	48	ASN

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
1	A	59	ARG
1	A	115	LYS
1	A	120	SER
1	A	143	ARG
1	A	197	ASP
1	A	322	TRP
1	B	37	HIS
1	B	59	ARG
1	B	70	THR
1	B	87	GLU
1	B	97	ASN
1	B	170	THR
1	B	204	ASP
1	B	297	ASN
1	B	322	TRP
1	C	33	GLU
1	C	34	ASN
1	C	64	LEU
1	C	174	ASP
1	C	323	TRP
1	C	327	ASP
1	C	346	THR
1	C	399	LYS
1	D	34	ASP
1	D	48	ASN
1	D	50	GLN
1	D	107	LEU
1	D	322	TRP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (7) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	48	ASN
1	B	62	ASN
1	B	97	ASN
1	B	261	ASN
1	D	33	ASN
1	D	233	ASN
1	D	297	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

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

## 5.5 Carbohydrates ⓘ

15 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	E	1	2	12,12,12	0.44	0	17,17,17	0.79	0
2	BGC	E	2	2	11,11,12	0.20	0	15,15,17	1.00	1 (6%)
2	BGC	E	3	2	11,11,12	0.27	0	15,15,17	0.66	0
2	BGC	E	4	2	11,11,12	0.30	0	15,15,17	0.63	0
2	BGC	E	5	2	11,11,12	0.26	0	15,15,17	0.70	0
2	BGC	F	1	2	12,12,12	0.57	0	17,17,17	1.02	2 (11%)
2	BGC	F	2	2	11,11,12	0.37	0	15,15,17	0.47	0
2	BGC	F	3	2	11,11,12	0.42	0	15,15,17	0.72	0
2	BGC	F	4	2	11,11,12	0.33	0	15,15,17	0.94	1 (6%)
2	BGC	F	5	2	11,11,12	0.44	0	15,15,17	0.87	1 (6%)
2	BGC	G	1	2	12,12,12	0.61	0	17,17,17	0.90	1 (5%)
2	BGC	G	2	2	11,11,12	0.48	0	15,15,17	0.77	0
2	BGC	G	3	2	11,11,12	0.43	0	15,15,17	0.79	0
2	BGC	G	4	2	11,11,12	0.48	0	15,15,17	1.07	1 (6%)
2	BGC	G	5	2	11,11,12	0.43	0	15,15,17	1.11	2 (13%)

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	E	1	2	-	2/2/22/22	0/1/1/1
2	BGC	E	2	2	-	0/2/19/22	0/1/1/1
2	BGC	E	3	2	-	0/2/19/22	0/1/1/1
2	BGC	E	4	2	-	0/2/19/22	0/1/1/1
2	BGC	E	5	2	-	0/2/19/22	0/1/1/1
2	BGC	F	1	2	-	2/2/22/22	0/1/1/1
2	BGC	F	2	2	-	1/2/19/22	0/1/1/1
2	BGC	F	3	2	-	2/2/19/22	0/1/1/1
2	BGC	F	4	2	-	2/2/19/22	0/1/1/1
2	BGC	F	5	2	-	2/2/19/22	0/1/1/1
2	BGC	G	1	2	-	2/2/22/22	0/1/1/1
2	BGC	G	2	2	-	1/2/19/22	0/1/1/1
2	BGC	G	3	2	-	0/2/19/22	0/1/1/1
2	BGC	G	4	2	-	2/2/19/22	0/1/1/1
2	BGC	G	5	2	-	1/2/19/22	0/1/1/1

There are no bond length outliers.

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	E	2	BGC	O5-C5-C6	2.89	111.73	107.20
2	G	5	BGC	C1-C2-C3	2.81	113.12	109.67
2	F	1	BGC	C3-C4-C5	2.62	114.91	110.24
2	G	4	BGC	C1-O5-C5	2.45	115.51	112.19
2	G	5	BGC	O5-C5-C6	2.27	110.77	107.20
2	F	4	BGC	C1-O5-C5	2.21	115.19	112.19
2	F	1	BGC	O5-C5-C4	2.12	113.55	109.69
2	G	1	BGC	C3-C4-C5	2.03	113.86	110.24
2	F	5	BGC	C1-C2-C3	2.01	112.14	109.67

There are no chirality outliers.

All (17) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	F	5	BGC	O5-C5-C6-O6
2	F	4	BGC	C4-C5-C6-O6
2	F	3	BGC	C4-C5-C6-O6
2	F	4	BGC	O5-C5-C6-O6
2	E	1	BGC	C4-C5-C6-O6
2	G	1	BGC	O5-C5-C6-O6

*Continued on next page...*

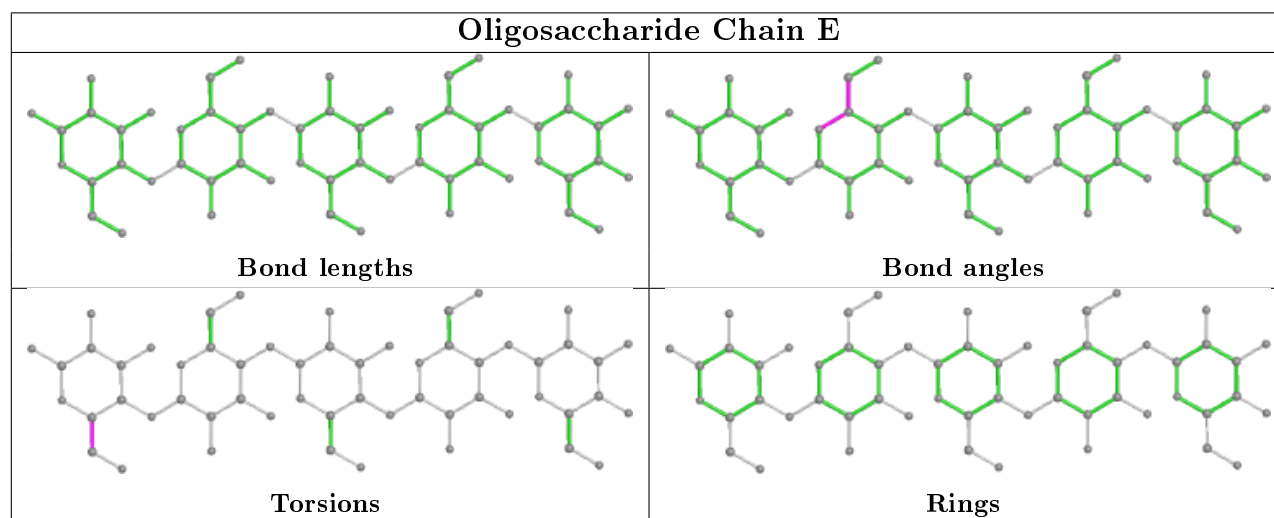
*Continued from previous page...*

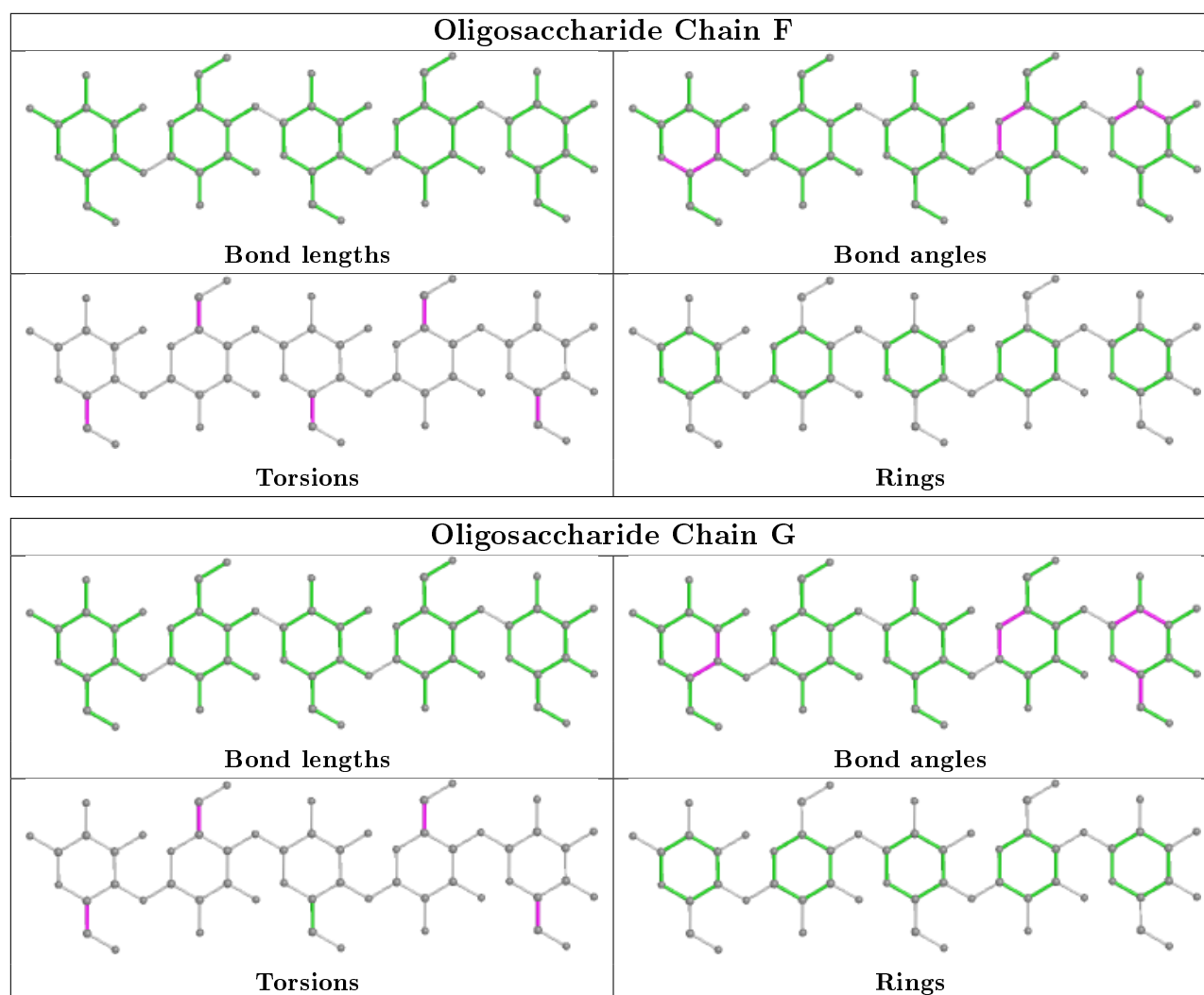
Mol	Chain	Res	Type	Atoms
2	E	1	BGC	O5-C5-C6-O6
2	F	5	BGC	C4-C5-C6-O6
2	F	1	BGC	C4-C5-C6-O6
2	F	3	BGC	O5-C5-C6-O6
2	G	4	BGC	O5-C5-C6-O6
2	G	1	BGC	C4-C5-C6-O6
2	F	2	BGC	O5-C5-C6-O6
2	F	1	BGC	O5-C5-C6-O6
2	G	2	BGC	O5-C5-C6-O6
2	G	5	BGC	O5-C5-C6-O6
2	G	4	BGC	C4-C5-C6-O6

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 21 ligands modelled in this entry, 2 are monoatomic - leaving 19 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	PO4	A	510	-	4,4,4	0.80	0	6,6,6	0.51	0
3	EDO	B	504	-	3,3,3	0.47	0	2,2,2	0.19	0
4	PO4	D	509	-	4,4,4	0.83	0	6,6,6	0.48	0
4	PO4	D	507	-	4,4,4	0.83	0	6,6,6	0.44	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	PO4	B	503	-	4,4,4	0.77	0	6,6,6	0.46	0
4	PO4	C	507	-	4,4,4	0.80	0	6,6,6	0.41	0
4	PO4	B	501	-	4,4,4	0.93	0	6,6,6	0.45	0
4	PO4	A	509	-	4,4,4	0.89	0	6,6,6	0.45	0
4	PO4	A	511	-	4,4,4	0.84	0	6,6,6	0.42	0
4	PO4	C	509	-	4,4,4	0.86	0	6,6,6	0.44	0
4	PO4	C	506	-	4,4,4	0.93	0	6,6,6	0.35	0
3	EDO	C	511	-	3,3,3	0.50	0	2,2,2	0.22	0
3	EDO	D	506	-	3,3,3	0.50	0	2,2,2	0.27	0
3	EDO	A	506	-	3,3,3	0.48	0	2,2,2	0.30	0
4	PO4	A	508	-	4,4,4	0.85	0	6,6,6	0.45	0
4	PO4	D	508	-	4,4,4	0.88	0	6,6,6	0.42	0
3	EDO	A	507	-	3,3,3	0.53	0	2,2,2	0.22	0
4	PO4	C	508	-	4,4,4	0.91	0	6,6,6	0.42	0
4	PO4	B	502	-	4,4,4	0.86	0	6,6,6	0.41	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
3	EDO	A	506	-	-	0/1/1/1	-
3	EDO	A	507	-	-	1/1/1/1	-
3	EDO	B	504	-	-	0/1/1/1	-
3	EDO	C	511	-	-	0/1/1/1	-
3	EDO	D	506	-	-	1/1/1/1	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	D	506	EDO	O1-C1-C2-O2
3	A	507	EDO	O1-C1-C2-O2

There are no ring outliers.

No monomer is involved in short contacts.

## 5.7 Other polymers

There are no such residues in this entry.

## 5.8 Polymer linkage issues

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	399/420 (95%)	0.19	0 100 100	26, 49, 82, 100	0
1	B	385/420 (91%)	0.47	21 (5%) 25 24	30, 65, 105, 123	0
1	C	390/420 (92%)	0.30	5 (1%) 77 78	32, 58, 86, 108	0
1	D	388/420 (92%)	0.32	8 (2%) 63 65	33, 55, 93, 103	0
All	All	1562/1680 (92%)	0.32	34 (2%) 62 63	26, 56, 93, 123	0

All (34) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	74	TYR	5.0
1	D	115	LYS	4.8
1	B	89	GLY	4.1
1	B	65	ILE	4.1
1	D	74	TYR	4.0
1	B	94	LYS	3.9
1	B	69	VAL	3.9
1	D	251	ILE	3.8
1	B	70	THR	3.1
1	D	256	VAL	3.1
1	B	66	ALA	3.0
1	C	118	LYS	3.0
1	B	39	LEU	2.9
1	B	260	SER	2.7
1	C	116	LYS	2.7
1	B	261	ASN	2.6
1	B	235	ASP	2.6
1	B	61	ALA	2.4
1	B	76	THR	2.4
1	B	114	VAL	2.4
1	B	118	LYS	2.3

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	D	258	GLN	2.2
1	D	246	ILE	2.2
1	B	64	SER	2.1
1	B	241	ILE	2.1
1	C	72	PRO	2.1
1	B	269	LEU	2.1
1	C	32	THR	2.1
1	B	80	PHE	2.1
1	B	87	GLU	2.0
1	B	119	THR	2.0
1	D	34	ASP	2.0
1	D	326	ASP	2.0
1	C	250	LEU	2.0

## 6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

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

## 6.3 Carbohydrates ⓘ

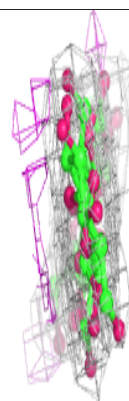
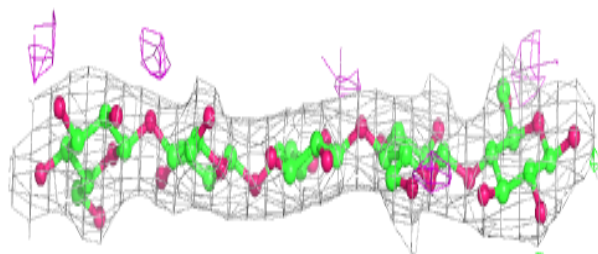
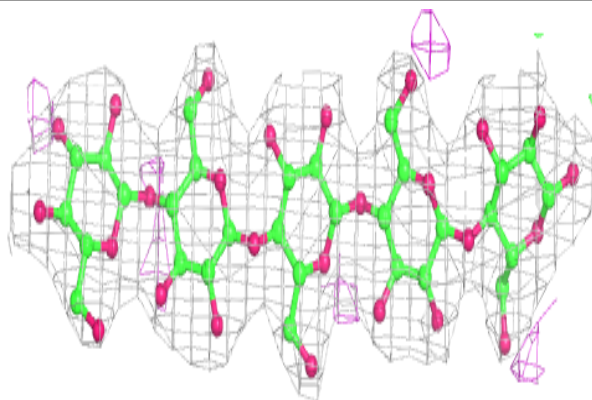
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 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	G	1	12/12	0.66	0.69	99,109,116,121	0
2	BGC	F	1	12/12	0.77	0.34	87,107,112,117	0
2	BGC	G	2	11/12	0.81	0.38	90,93,97,97	0
2	BGC	G	3	11/12	0.84	0.24	80,84,88,91	0
2	BGC	F	2	11/12	0.86	0.23	85,92,98,98	0
2	BGC	F	5	11/12	0.91	0.19	76,79,84,85	0
2	BGC	G	5	11/12	0.93	0.27	80,82,86,87	0
2	BGC	F	3	11/12	0.94	0.17	82,85,90,92	0
2	BGC	F	4	11/12	0.95	0.19	75,79,81,83	0
2	BGC	G	4	11/12	0.96	0.21	73,76,80,85	0
2	BGC	E	1	12/12	0.97	0.19	41,45,53,54	0
2	BGC	E	5	11/12	0.97	0.24	42,45,49,52	0
2	BGC	E	3	11/12	0.98	0.19	38,39,40,41	0
2	BGC	E	4	11/12	0.99	0.17	37,39,41,41	0
2	BGC	E	2	11/12	0.99	0.17	37,40,42,42	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.

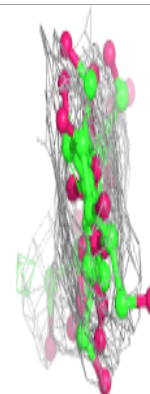
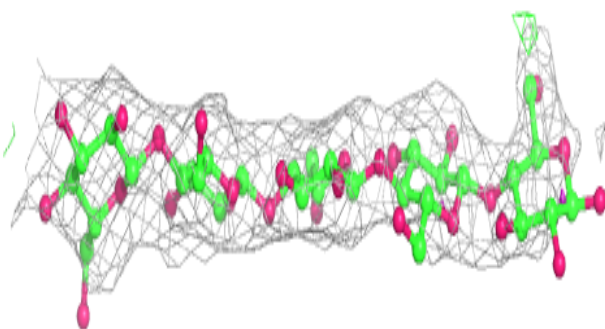
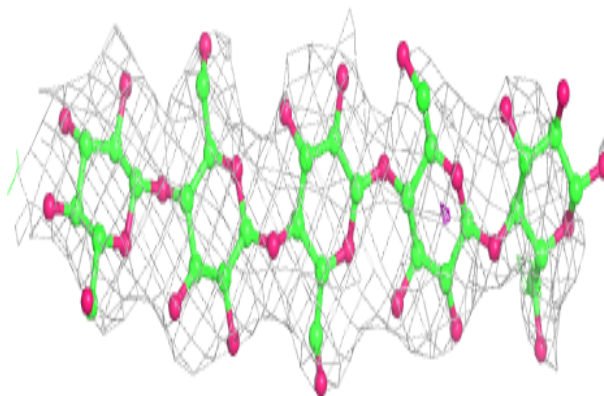
**Electron density around Chain E:**

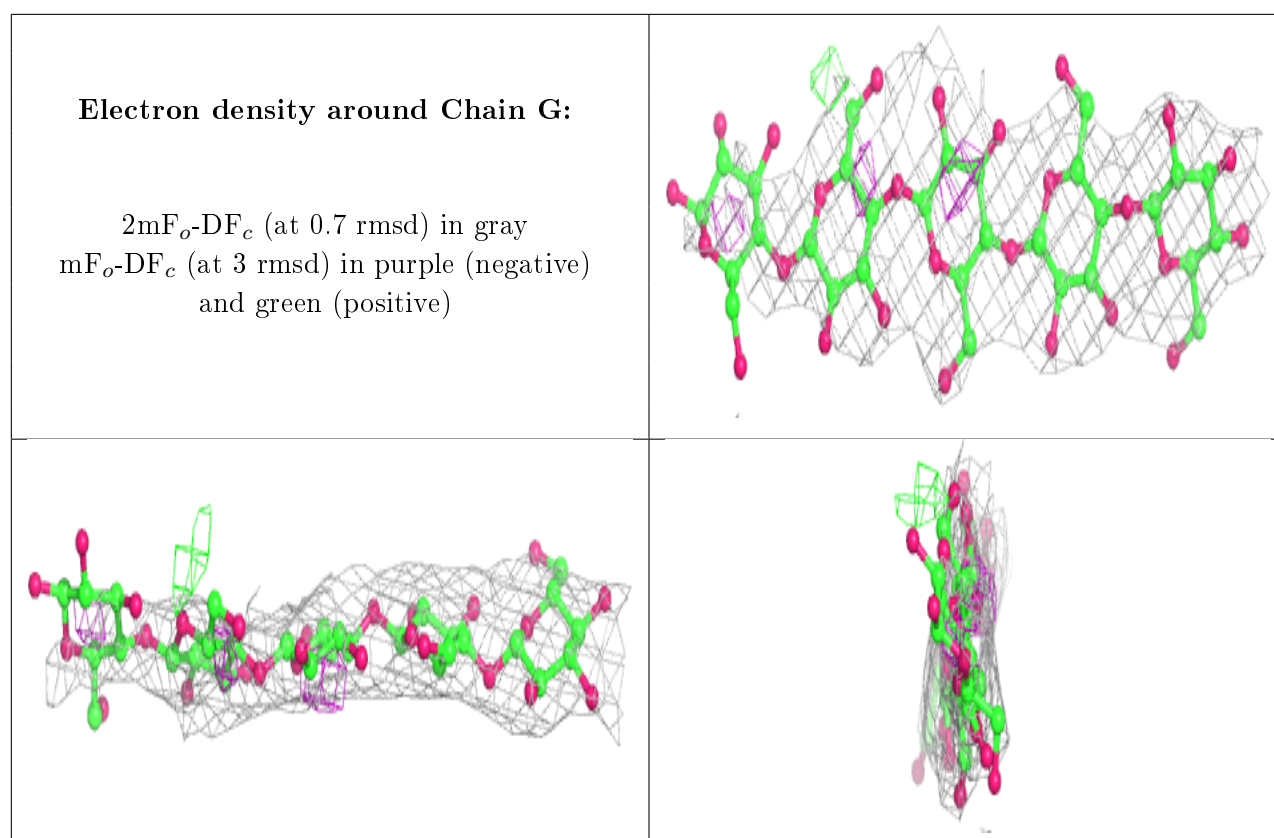
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



**Electron density around Chain F:**

$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)





## 6.4 Ligands ⓘ

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
3	EDO	B	504	4/4	0.79	0.20	61,63,65,65	0
3	EDO	C	511	4/4	0.80	0.14	73,75,77,77	0
4	PO4	A	510	5/5	0.88	0.17	93,100,105,107	0
3	EDO	A	506	4/4	0.88	0.15	75,77,78,79	0
3	EDO	A	507	4/4	0.88	0.25	55,61,62,65	0
4	PO4	D	509	5/5	0.90	0.23	87,97,105,108	0
4	PO4	C	509	5/5	0.93	0.17	86,88,91,94	0
4	PO4	B	502	5/5	0.93	0.21	88,89,95,98	0
3	EDO	D	506	4/4	0.94	0.17	59,60,60,62	0
4	PO4	C	507	5/5	0.94	0.18	76,76,89,90	0
4	PO4	C	508	5/5	0.95	0.18	79,81,86,87	0
5	NA	A	512	1/1	0.95	0.13	59,59,59,59	0
4	PO4	A	511	5/5	0.96	0.21	76,82,86,91	0
4	PO4	D	508	5/5	0.96	0.29	85,93,97,101	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
5	NA	C	510	1/1	0.96	0.12	54,54,54,54	0
4	PO4	A	508	5/5	0.97	0.17	57,61,63,65	0
4	PO4	A	509	5/5	0.97	0.14	59,69,76,78	0
4	PO4	B	501	5/5	0.97	0.17	68,72,75,77	0
4	PO4	D	507	5/5	0.98	0.12	50,51,54,54	0
4	PO4	B	503	5/5	0.98	0.16	52,58,60,62	0
4	PO4	C	506	5/5	0.98	0.14	43,48,51,51	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.