



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 9, 2020 – 05:32 AM BST

PDB ID : 2JEN  
Title : Family 12 xyloglucanase from Bacillus licheniformis in complex with ligand  
Authors : Gloster, T.M.; Ibatullin, F.M.; Macauley, K.; Eklof, J.M.; Roberts, S.; Turkenburg, J.P.; Bjornvad, M.E.; Jorgensen, P.L.; Danielsen, S.; Johansen, K.S.; Borchert, T.V.; Wilson, K.S.; Brumer, H.; Davies, G.J.  
Deposited on : 2007-01-18  
Resolution : 1.40 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)  
Xtrriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
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) : 2.13.1

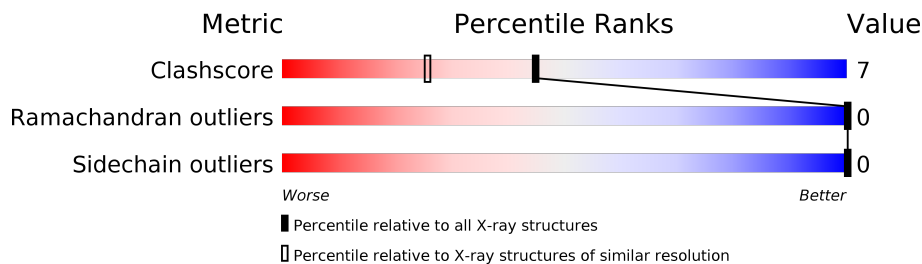
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.40 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))
Clashscore	141614	1812 (1.40-1.40)
Ramachandran outliers	138981	1763 (1.40-1.40)
Sidechain outliers	138945	1762 (1.40-1.40)

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  $\leq 5\%$ .

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	261	
2	B	6	
3	C	4	

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	GLC	B	1[A]	X	-	-	-
2	GLC	B	1[B]	X	-	-	-

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	GLC	B	2	X	-	-	-
3	GLC	C	1	X	-	-	-
3	GLC	C	2	X	-	-	-

## 2 Entry composition [i](#)

There are 7 unique types of molecules in this entry. The entry contains 2483 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.

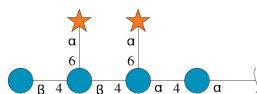
- Molecule 1 is a protein called ENDO-BETA-1,4-GLUCANASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	233	1972	1264	320	386	2	0	29	0

There is a discrepancy between the modelled and reference sequences:

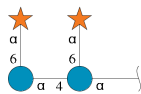
Chain	Residue	Modelled	Actual	Comment	Reference
A	155	ALA	GLU	engineered mutation	UNP Q7X4S4

- Molecule 2 is an oligosaccharide called beta-D-glucopyranose-(1-4)-[alpha-D-xylopyranose-(1-6)]beta-D-glucopyranose-(1-4)-[alpha-D-xylopyranose-(1-6)]alpha-D-glucopyranose-(1-4)-alpha-D-glucopyranose.



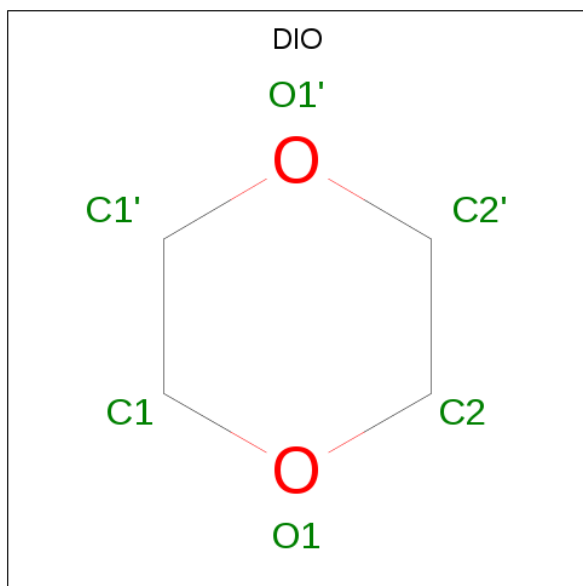
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace
			Total	C	O			
2	B	6	75	40	35	0	1	0

- Molecule 3 is an oligosaccharide called alpha-D-xylopyranose-(1-6)-alpha-D-glucopyranose-(1-4)-[alpha-D-xylopyranose-(1-6)]alpha-D-glucopyranose.



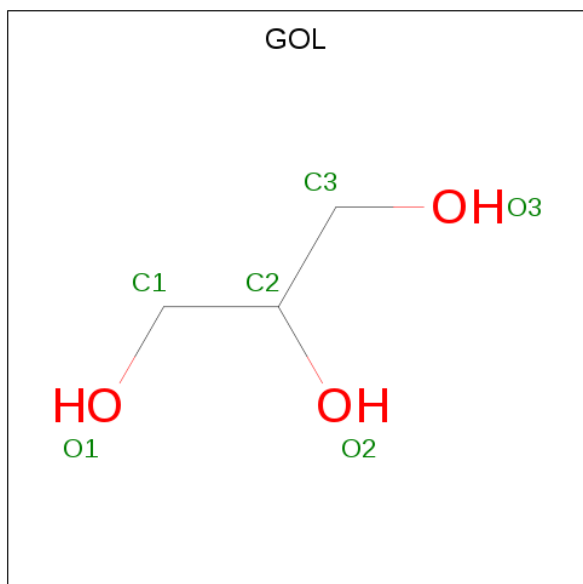
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace
			Total	C	O			
3	C	4	41	22	19	0	0	0

- Molecule 4 is 1,4-DIETHYLENE DIOXIDE (three-letter code: DIO) (formula:  $C_4H_8O_2$ ).



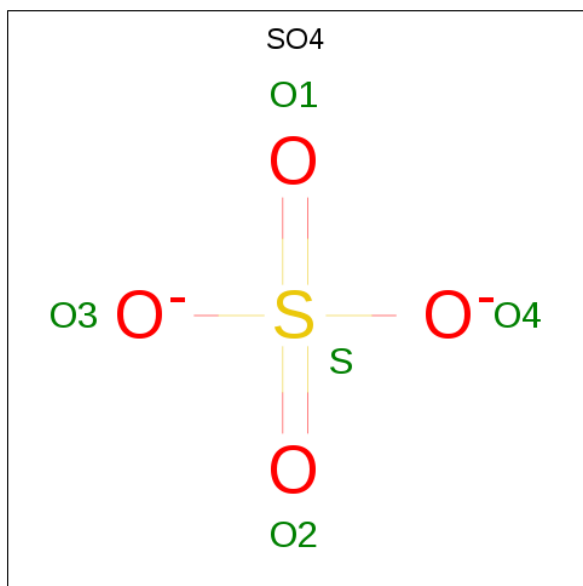
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	1
			12	8	4		
4	A	1	Total	C	O	0	0
			6	4	2		

- Molecule 5 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			6	3	3		

- Molecule 6 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	A	1	Total	O	S	0	0
			5	4	1		
6	A	1	Total	O	S	0	0
			5	4	1		

- Molecule 7 is water.


Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	361	Total	O	0	0
			361	361		

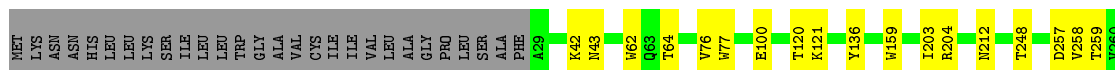
### 3 Residue-property plots

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. 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. 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.

Note EDS was not executed.

- Molecule 1: ENDO-BETA-1,4-GLUCANASE

Chain A: 



 R261

- Molecule 2: beta-D-glucopyranose-(1-4)-[alpha-D-xylopyranose-(1-6)]beta-D-glucopyranose-(1-4)-[alpha-D-xylopyranose-(1-6)]alpha-D-glucopyranose-(1-4)-alpha-D-glucopyranose

Chain B: 

 GLC1  
GLC2  
BGC3  
BGC4  
XYS5  
XYS6

- Molecule 3: alpha-D-xylopyranose-(1-6)-alpha-D-glucopyranose-(1-4)-[alpha-D-xylopyranose-(1-6)]alpha-D-glucopyranose

Chain C: 

 GLC1  
GLC2  
XYS3  
XYS4

## 4 Data and refinement statistics

Xtrriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	H 3 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	114.63Å 114.63Å 110.30Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	74.54 – 1.40	Depositor
% Data completeness (in resolution range)	99.7 (74.54-1.40)	Depositor
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	REFMAC 5.3.0011	Depositor
R, $R_{free}$	0.115 , 0.141	Depositor
Estimated twinning fraction	No twinning to report.	Xtrriage
Total number of atoms	2483	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	17.0	wwPDB-VP



## 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: GOL, XYS, GLC, DIO, BGC, SO4

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	0.71	0/2118	0.75	1/2890 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	A	204	ARG	NE-CZ-NH1	5.99	123.30	120.30

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	1972	0	1882	26	0
2	B	75	0	63	0	0
3	C	41	0	35	0	0
4	A	18	0	24	1	0
5	A	6	0	8	0	0
6	A	10	0	0	1	0
7	A	361	0	0	10	0
All	All	2483	0	2012	27	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 7.

All (27) 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:A:64[B]:THR:CG2	1:A:76[B]:VAL:HG22	1.77	1.15
1:A:42[A]:LYS:NZ	7:A:2022:HOH:O	1.79	1.13
1:A:100[A]:GLU:OE2	7:A:2125:HOH:O	1.75	1.04
1:A:42[A]:LYS:CE	7:A:2019:HOH:O	2.06	1.02
1:A:64[B]:THR:HG22	1:A:76[B]:VAL:HG22	1.46	0.94
1:A:120[B]:THR:HG23	1:A:258:VAL:HG22	1.49	0.93
1:A:64[B]:THR:HG22	1:A:76[B]:VAL:CG2	2.03	0.88
1:A:42[A]:LYS:HE2	7:A:2019:HOH:O	1.67	0.87
1:A:212[A]:ASN:OD1	7:A:2287:HOH:O	1.94	0.86
1:A:64[B]:THR:CG2	1:A:76[B]:VAL:CG2	2.54	0.85
1:A:64[B]:THR:HG23	1:A:76[B]:VAL:HG22	1.61	0.81
1:A:42[A]:LYS:HG3	7:A:2018:HOH:O	1.91	0.69
1:A:62:TRP:HZ3	1:A:76[B]:VAL:HG23	1.57	0.68
1:A:121:LYS:HG2	1:A:212[B]:ASN:HD22	1.60	0.67
1:A:120[B]:THR:HG21	1:A:136:TYR:CE1	2.36	0.61
1:A:257[B]:ASP:OD2	7:A:2338:HOH:O	2.17	0.59
1:A:43[B]:ASN:ND2	6:A:1266:SO4:O3	2.40	0.54
1:A:64[B]:THR:HG21	7:A:2069:HOH:O	2.15	0.47
1:A:120[B]:THR:HG21	1:A:136:TYR:HE1	1.78	0.47
1:A:42[A]:LYS:NZ	7:A:2019:HOH:O	2.37	0.47
1:A:121:LYS:HG2	1:A:212[B]:ASN:ND2	2.29	0.46
1:A:120[A]:THR:HG22	1:A:258:VAL:HG22	1.97	0.46
1:A:257[B]:ASP:OD2	1:A:259[B]:THR:HG23	2.16	0.46
1:A:159:TRP:CD2	1:A:203[B]:ILE:HD12	2.52	0.45
4:A:1262[A]:DIO:C2	7:A:2343:HOH:O	2.67	0.42
1:A:77:TRP:O	1:A:248[A]:THR:HG23	2.19	0.42
1:A:120[A]:THR:HG22	1:A:258:VAL:HG13	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	261/261 (100%)	253 (97%)	8 (3%)	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 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	224/218 (103%)	224 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	227	GLN
1	A	252	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](#)

11 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	GLC	B	1[A]	2	12,12,12	0.46	0	17,17,17	1.72	7 (41%)
2	GLC	B	1[B]	2	12,12,12	0.66	0	17,17,17	2.62	8 (47%)
2	GLC	B	2	2	11,11,12	0.60	0	15,15,17	1.13	1 (6%)
2	BGC	B	3	2	11,11,12	1.04	1 (9%)	15,15,17	1.29	1 (6%)
2	BGC	B	4	2	11,11,12	0.46	0	15,15,17	0.79	1 (6%)
2	XYS	B	5	2	9,9,10	1.14	1 (11%)	10,12,14	2.10	4 (40%)
2	XYS	B	6	2	9,9,10	1.30	1 (11%)	10,12,14	1.45	1 (10%)
3	GLC	C	1	3	12,12,12	0.49	0	17,17,17	0.69	0
3	GLC	C	2	3	11,11,12	0.61	0	15,15,17	0.93	0
3	XYS	C	3	3	9,9,10	0.90	1 (11%)	10,12,14	1.29	1 (10%)
3	XYS	C	4	3	9,9,10	0.58	0	10,12,14	0.87	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	GLC	B	1[A]	2	1/1/5/5	0/2/22/22	0/1/1/1
2	GLC	B	1[B]	2	1/1/5/5	0/2/22/22	0/1/1/1
2	GLC	B	2	2	1/1/4/5	2/2/19/22	0/1/1/1
2	BGC	B	3	2	-	0/2/19/22	0/1/1/1
2	BGC	B	4	2	-	2/2/19/22	0/1/1/1
2	XYS	B	5	2	-	-	0/1/1/1
2	XYS	B	6	2	-	-	0/1/1/1
3	GLC	C	1	3	1/1/5/5	0/2/22/22	0/1/1/1
3	GLC	C	2	3	1/1/4/5	0/2/19/22	0/1/1/1
3	XYS	C	3	3	-	-	0/1/1/1
3	XYS	C	4	3	-	-	0/1/1/1

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	6	XYS	O5-C1	-3.44	1.36	1.42

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	3	BGC	O5-C1	-2.77	1.39	1.43
3	C	3	XYS	O5-C1	-2.54	1.38	1.42
2	B	5	XYS	O5-C5	-2.09	1.38	1.42

All (24) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	1[B]	GLC	O5-C1-C2	5.36	119.86	110.28
2	B	1[B]	GLC	C1-C2-C3	-4.71	100.55	110.31
2	B	1[B]	GLC	O1-C1-O5	4.52	123.95	110.38
2	B	5	XYS	C5-C4-C3	3.99	114.57	109.67
2	B	6	XYS	C5-O5-C1	3.58	117.03	111.52
2	B	3	BGC	C1-O5-C5	3.56	117.01	112.19
2	B	5	XYS	C5-O5-C1	3.54	116.96	111.52
2	B	1[B]	GLC	C1-O5-C5	-3.39	107.27	113.66
2	B	1[B]	GLC	O4-C4-C5	-3.24	101.24	109.30
2	B	1[A]	GLC	O3-C3-C4	2.92	117.09	110.35
2	B	1[A]	GLC	O2-C2-C1	2.81	115.67	109.16
2	B	1[A]	GLC	C3-C4-C5	-2.72	105.39	110.24
2	B	5	XYS	O3-C3-C2	2.54	114.87	109.99
2	B	1[B]	GLC	C3-C4-C5	2.39	114.51	110.24
3	C	3	XYS	C5-O5-C1	2.39	115.20	111.52
2	B	1[B]	GLC	O5-C5-C6	2.35	112.27	106.44
2	B	1[A]	GLC	O2-C2-C3	-2.30	105.02	110.35
2	B	1[A]	GLC	O3-C3-C2	-2.28	105.07	110.35
2	B	4	BGC	O5-C5-C6	2.26	110.74	107.20
2	B	1[B]	GLC	O1-C1-C2	2.25	115.38	109.03
2	B	1[A]	GLC	O5-C1-C2	-2.22	106.32	110.28
2	B	2	GLC	O3-C3-C4	2.12	115.25	110.35
2	B	1[A]	GLC	O5-C5-C6	2.12	111.70	106.44
2	B	5	XYS	C4-C3-C2	-2.08	108.45	110.92

All (5) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
3	C	2	GLC	C1
2	B	2	GLC	C1
2	B	1[B]	GLC	C1
2	B	1[A]	GLC	C1
3	C	1	GLC	C1

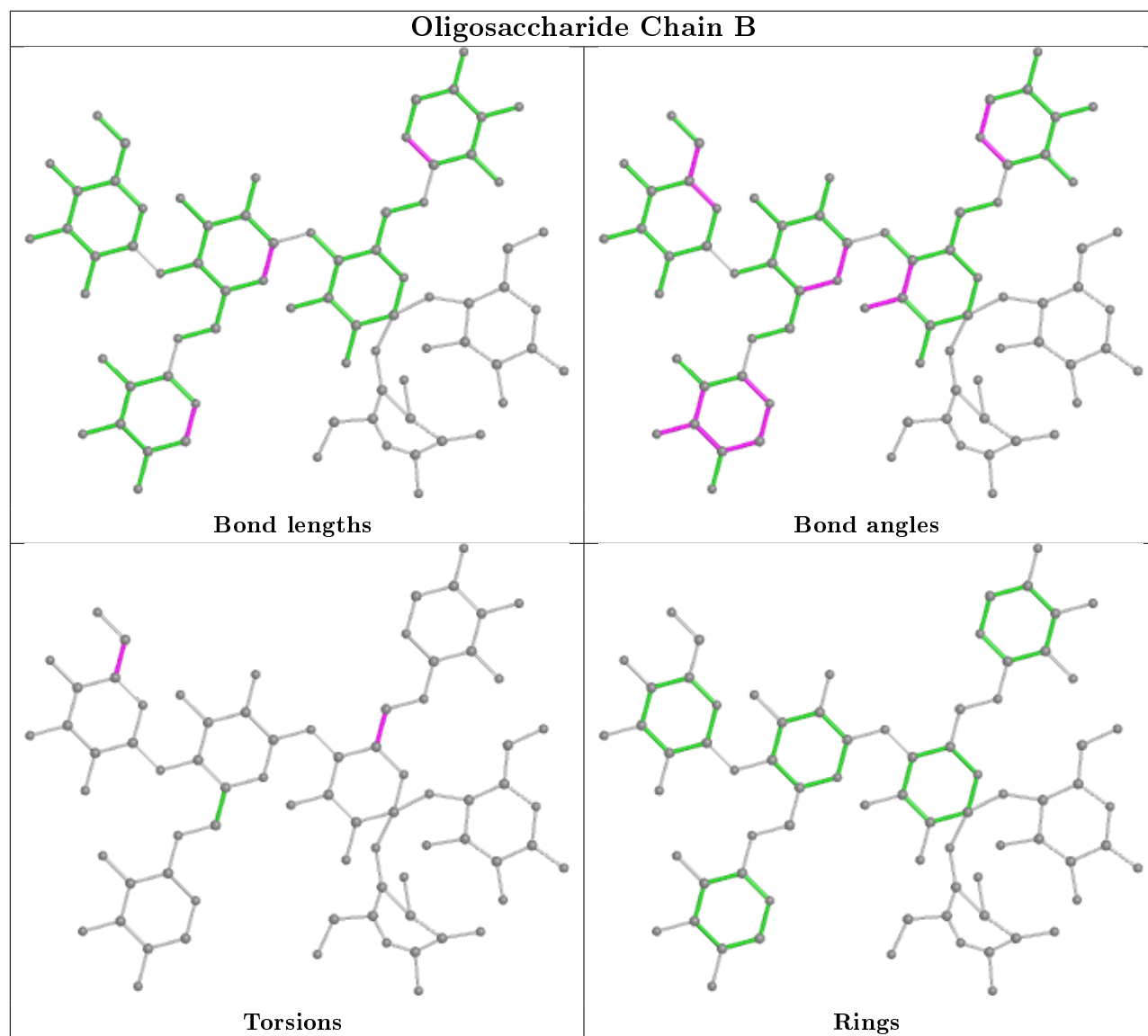
All (4) torsion outliers are listed below:

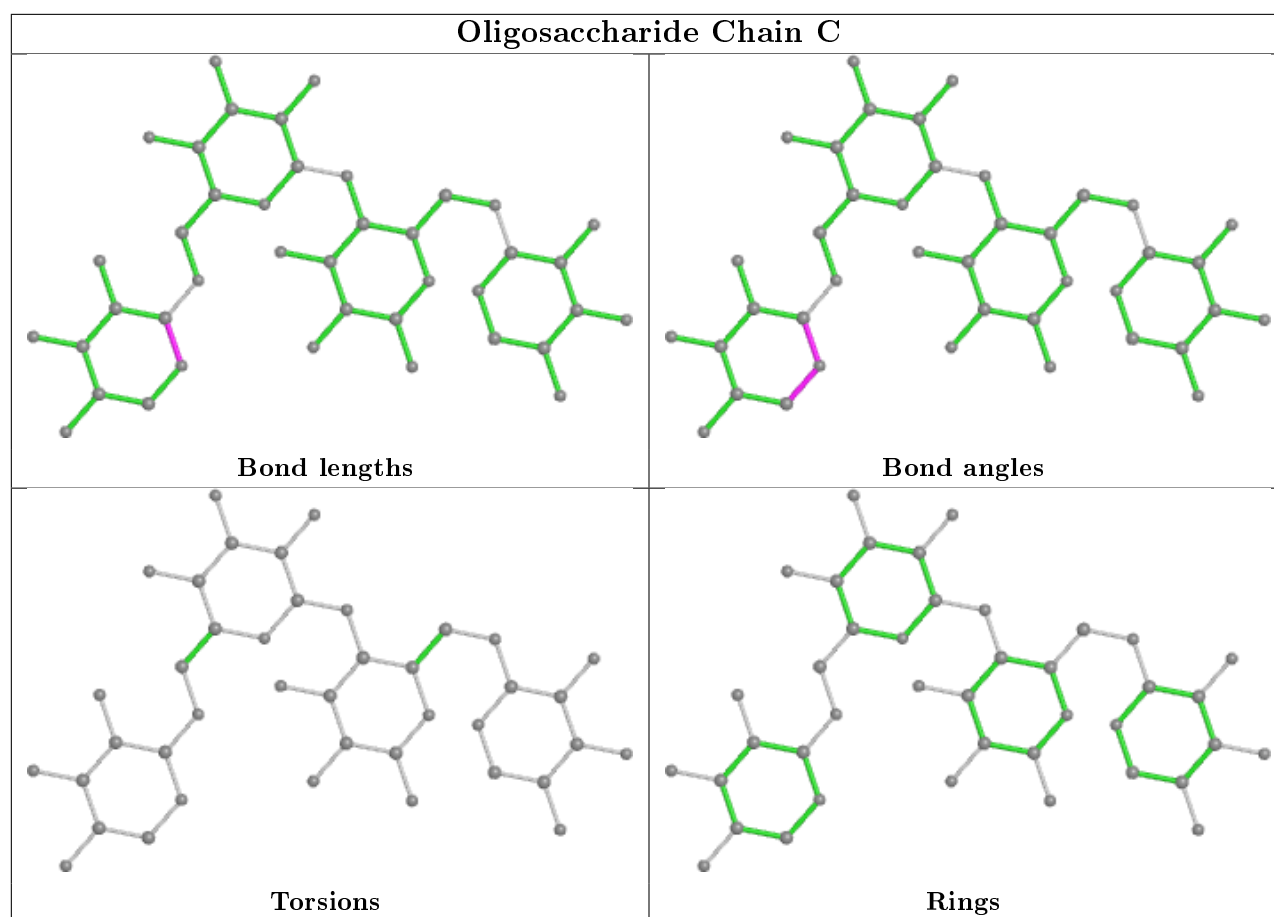
Mol	Chain	Res	Type	Atoms
2	B	2	GLC	O5-C5-C6-O6
2	B	2	GLC	C4-C5-C6-O6
2	B	4	BGC	C4-C5-C6-O6
2	B	4	BGC	O5-C5-C6-O6

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](#)

6 ligands 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	DIO	A	1263	-	6,6,6	0.51	0	6,6,6	1.22	1 (16%)
4	DIO	A	1262[B]	-	6,6,6	0.54	0	6,6,6	1.30	0
5	GOL	A	1264	-	5,5,5	0.41	0	5,5,5	0.48	0
6	SO4	A	1266	-	4,4,4	0.31	0	6,6,6	0.51	0
6	SO4	A	1265	-	4,4,4	0.14	0	6,6,6	1.50	1 (16%)
4	DIO	A	1262[A]	-	6,6,6	0.50	0	6,6,6	1.14	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
5	GOL	A	1264	-	-	0/4/4/4	-
4	DIO	A	1263	-	-	-	0/1/1/1
4	DIO	A	1262[A]	-	-	-	0/1/1/1
4	DIO	A	1262[B]	-	-	-	0/1/1/1

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	1263	DIO	C2'-O1'-C1'	2.39	117.86	109.89
6	A	1265	SO4	O3-S-O1	-2.33	97.14	109.31

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	A	1266	SO4	1	0
4	A	1262[A]	DIO	1	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 Fit of model and data

### 6.1 Protein, DNA and RNA chains

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

### 6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

### 6.4 Ligands

EDS was not executed - this section is therefore empty.

### 6.5 Other polymers

EDS was not executed - this section is therefore empty.