



Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 02:56 AM GMT

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
A user guide is available at
<http://wwpdb.org/validation/2016/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 : 1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix) : **NOT EXECUTED**
EDS : **NOT EXECUTED**
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk26865

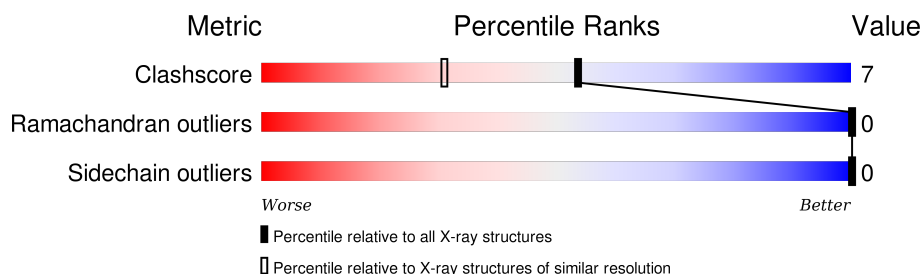
1 Overall quality at a glance

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	102246	1295 (1.40-1.40)
Ramachandran outliers	100387	1259 (1.40-1.40)
Sidechain outliers	100360	1258 (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 ≥ 3 , 2, 1 and 0 types of geometric quality criteria. 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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	261	

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
5	GLC	A	1267[A]	X	-	-	-
5	GLC	A	1267[B]	X	-	-	-
5	GLC	A	1268	X	-	-	-
6	GLC	A	1273	X	-	-	-
6	GLC	A	1274	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.

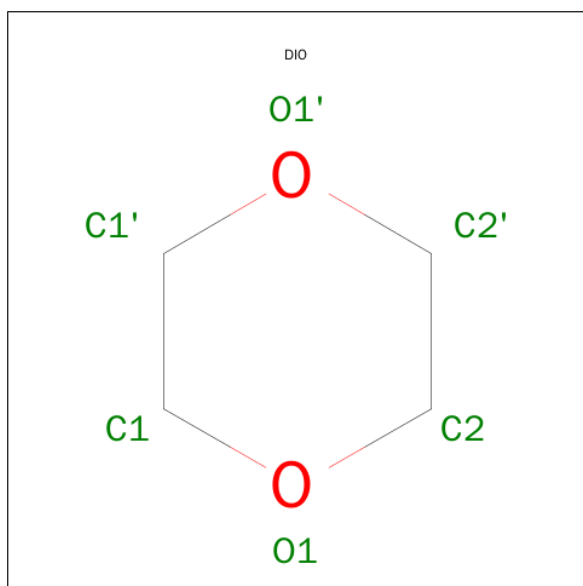
- Molecule 1 is a protein called ENDO-BETA-1,4-GLUCANASE.

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

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 1,4-DIETHYLENE DIOXIDE (three-letter code: DIO) (formula: C₄H₈O₂).



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

- Molecule 3 is GLYCEROL (three-letter code: GOL) (formula: C₃H₈O₃).



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

- Molecule 4 is SULFATE ION (three-letter code: SO4) (formula: O₄S).



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

- Molecule 5 is a polymer of unknown type called SUGAR (6-MER).

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	6	Total	C	O	0	1
			75	40	35		

- Molecule 6 is a polymer of unknown type called SUGAR (4-MER).

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	A	4	Total	C	O	0	0
			41	22	19		

- Molecule 7 is water.


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

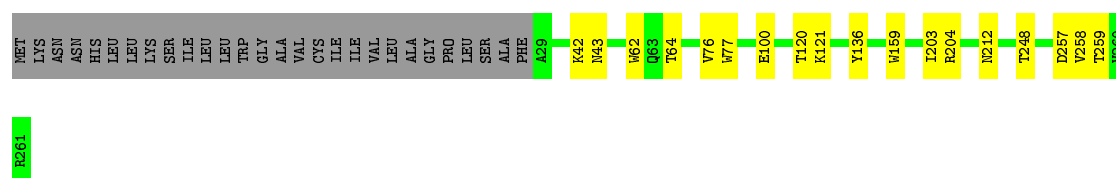
3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

Note EDS was not executed.

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

Chain A:  82% 7% 11%



4 Data and refinement statistics

Xtriage (Phenix) and EDS were not executed - this section will therefore be incomplete.

Property	Value	Source
Space group	H 3 2	Depositor
Cell constants a, b, c, α , β , γ	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.	Xtriage
Total number of atoms	2483	wwPDB-VP
Average B, all atoms (Å ²)	17.0	wwPDB-VP

5 Model quality

5.1 Standard geometry

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%)

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
5	A	3	0
6	A	2	0
All	All	5	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

All (5) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
5	A	1267[A]	GLC	C1
5	A	1267[B]	GLC	C1
5	A	1268	GLC	C1
6	A	1273	GLC	C1
6	A	1274	GLC	C1

There are no planarity outliers.

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	1972	0	1882	26	0
2	A	18	0	24	1	0
3	A	6	0	8	0	0
4	A	10	0	0	1	0
5	A	75	0	63	0	0
6	A	41	0	35	0	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	4: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

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
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
2: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 ⓘ

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 ⓘ

11 carbohydrates 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
5	GLC	A	1267[A]	5	12,12,12	0.43	0	17,17,17	1.73	7 (41%)
5	GLC	A	1267[B]	5	12,12,12	0.81	1 (8%)	17,17,17	3.01	8 (47%)
5	GLC	A	1268	5	11,11,12	0.59	0	14,15,17	1.14	1 (7%)
5	BGC	A	1269	5	11,11,12	0.99	1 (9%)	14,15,17	1.33	1 (7%)
5	BGC	A	1270	5	11,11,12	0.44	0	14,15,17	0.66	0
5	XYS	A	1271	5	9,9,10	1.14	1 (11%)	12,12,14	2.17	5 (41%)
5	XYS	A	1272	5	9,9,10	1.28	1 (11%)	12,12,14	1.36	1 (8%)
6	GLC	A	1273	6	12,12,12	0.51	0	17,17,17	0.72	0
6	GLC	A	1274	6	11,11,12	0.59	0	14,15,17	0.97	0
6	XYS	A	1275	6	9,9,10	0.89	1 (11%)	12,12,14	1.18	1 (8%)
6	XYS	A	1276	6	9,9,10	0.58	0	12,12,14	0.97	1 (8%)

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	GLC	A	1267[A]	5	1/1/5/5	0/2/22/22	0/1/1/1
5	GLC	A	1267[B]	5	1/1/5/5	0/2/22/22	0/1/1/1
5	GLC	A	1268	5	1/1/4/5	0/2/19/22	0/1/1/1
5	BGC	A	1269	5	-	0/2/19/22	0/1/1/1
5	BGC	A	1270	5	-	0/2/19/22	0/1/1/1
5	XYS	A	1271	5	-	0/0/14/17	0/1/1/1
5	XYS	A	1272	5	-	0/0/14/17	0/1/1/1
6	GLC	A	1273	6	1/1/5/5	0/2/22/22	0/1/1/1
6	GLC	A	1274	6	1/1/4/5	0/2/19/22	0/1/1/1
6	XYS	A	1275	6	-	0/0/14/17	0/1/1/1
6	XYS	A	1276	6	-	0/0/14/17	0/1/1/1

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	1272	XYS	O5-C1	-3.41	1.36	1.42
5	A	1269	BGC	O5-C1	-2.65	1.39	1.43
6	A	1275	XYS	O5-C1	-2.52	1.38	1.42
5	A	1271	XYS	O5-C5	-2.07	1.38	1.42
5	A	1267[B]	GLC	C1-C2	-2.06	1.48	1.52

All (25) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	1267[B]	GLC	C1-C2-C3	-6.64	100.55	110.43
5	A	1267[B]	GLC	C1-O5-C5	-3.35	107.27	113.47
5	A	1267[B]	GLC	O4-C4-C5	-3.02	101.24	109.24
5	A	1271	XYS	C4-C3-C2	-2.82	108.45	111.24
5	A	1267[A]	GLC	C3-C4-C5	-2.75	105.39	110.20
5	A	1267[A]	GLC	O2-C2-C3	-2.36	105.02	110.34
5	A	1267[A]	GLC	O3-C3-C2	-2.34	105.07	110.34
6	A	1276	XYS	C4-C3-C2	-2.28	108.98	111.24
5	A	1267[A]	GLC	O5-C1-C2	-2.18	106.32	109.80
5	A	1267[A]	GLC	O5-C5-C6	2.12	111.70	106.36
5	A	1268	GLC	O3-C3-C4	2.18	115.25	110.34
6	A	1275	XYS	C5-O5-C1	2.25	115.20	111.57
5	A	1267[B]	GLC	O1-C1-C2	2.30	115.38	109.21
5	A	1267[B]	GLC	O5-C5-C6	2.34	112.27	106.36
5	A	1271	XYS	O5-C1-C2	2.42	114.83	110.31
5	A	1267[B]	GLC	C3-C4-C5	2.47	114.51	110.20

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	1267[A]	GLC	O2-C2-C1	2.65	115.67	109.82
5	A	1271	XYS	O3-C3-C2	2.70	114.87	110.00
5	A	1267[A]	GLC	O3-C3-C4	3.00	117.09	110.34
5	A	1271	XYS	C5-O5-C1	3.34	116.96	111.57
5	A	1272	XYS	C5-O5-C1	3.38	117.03	111.57
5	A	1269	BGC	C1-O5-C5	3.75	117.01	112.25
5	A	1271	XYS	C5-C4-C3	4.25	114.57	109.54
5	A	1267[B]	GLC	O1-C1-O5	5.01	123.95	110.25
5	A	1267[B]	GLC	O5-C1-C2	6.31	119.86	109.80

All (5) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
5	A	1267[A]	GLC	C1
6	A	1274	GLC	C1
5	A	1268	GLC	C1
6	A	1273	GLC	C1
5	A	1267[B]	GLC	C1

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.6 Ligand geometry ⓘ

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	$\# Z > 2$
2	DIO	A	1262[A]	-	6,6,6	0.49	0	6,6,6	1.10	0
2	DIO	A	1262[B]	-	6,6,6	0.53	0	6,6,6	1.27	0
2	DIO	A	1263	-	6,6,6	0.50	0	6,6,6	1.19	1 (16%)
3	GOL	A	1264	-	5,5,5	0.40	0	5,5,5	0.43	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
4	SO4	A	1265	-	4,4,4	0.36	0	6,6,6	1.18	0
4	SO4	A	1266	-	4,4,4	0.50	0	6,6,6	0.31	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	DIO	A	1262[A]	-	-	0/0/6/6	0/1/1/1
2	DIO	A	1262[B]	-	-	0/0/6/6	0/1/1/1
2	DIO	A	1263	-	-	0/0/6/6	0/1/1/1
3	GOL	A	1264	-	-	0/4/4/4	0/0/0/0
4	SO4	A	1265	-	-	0/0/0/0	0/0/0/0
4	SO4	A	1266	-	-	0/0/0/0	0/0/0/0

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	1263	DIO	C2'-O1'-C1'	2.37	117.86	109.89

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
2	A	1262[A]	DIO	1	0
4	A	1266	SO4	1	0

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 [i](#)

6.1 Protein, DNA and RNA chains [i](#)

EDS was not executed - this section will therefore be empty.

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

EDS was not executed - this section will therefore be empty.

6.3 Carbohydrates [i](#)

EDS was not executed - this section will therefore be empty.

6.4 Ligands [i](#)

EDS was not executed - this section will therefore be empty.

6.5 Other polymers [i](#)

EDS was not executed - this section will therefore be empty.