



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 02:15 PM GMT

PDB ID : 3WNO  
Title : D308A mutant of Bacillus circulans T-3040 cycloisomaltooligosaccharide glucanotransferase complexed with cycloisomaltooctaose  
Authors : Suzuki, N.; Fujimoto, Z.; Kim, Y.M.; Momma, M.; Kishine, N.; Suzuki, R.; Kobayashi, M.; Kimura, A.; Funane, K.  
Deposited on : 2013-12-10  
Resolution : 1.90 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  
<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) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

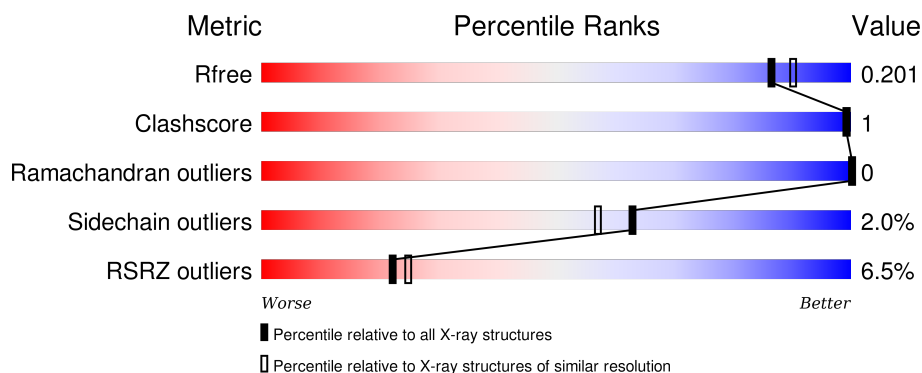
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.90 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))
$R_{free}$	91344	4755 (1.90-1.90)
Clashscore	102246	5398 (1.90-1.90)
Ramachandran outliers	100387	5338 (1.90-1.90)
Sidechain outliers	100360	5339 (1.90-1.90)
RSRZ outliers	91569	4766 (1.90-1.90)

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. 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	710	<div> <div>7%</div> <div>95%</div> <div>..</div> </div>
1	B	710	<div> <div>6%</div> <div>95%</div> <div>..</div> </div>

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
4	SO4	A	804	-	-	-	X
4	SO4	A	807	-	-	-	X
4	SO4	B	814	-	-	-	X
5	EDO	A	810	-	-	-	X
5	EDO	A	811	-	-	-	X
5	EDO	A	812	-	-	-	X
5	EDO	A	813	-	-	-	X
5	EDO	B	818	-	-	-	X
7	GLC	B	808	-	-	-	X

## 2 Entry composition

There are 9 unique types of molecules in this entry. The entry contains 12799 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 Cycloisomaltooligosaccharide glucanotransferase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	700	Total	C	N	O	S	0	5	0
			5532	3475	931	1111	15			
1	B	699	Total	C	N	O	S	0	6	0
			5515	3462	925	1113	15			

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	37	MET	-	EXPRESSION TAG	UNP P94286
A	38	GLY	-	EXPRESSION TAG	UNP P94286
A	278	PHE	SER	SEE REMARK 999	UNP P94286
A	308	ALA	ASP	ENGINEERED MUTATION	UNP P94286
A	739	LEU	-	EXPRESSION TAG	UNP P94286
A	740	GLU	-	EXPRESSION TAG	UNP P94286
A	741	HIS	-	EXPRESSION TAG	UNP P94286
A	742	HIS	-	EXPRESSION TAG	UNP P94286
A	743	HIS	-	EXPRESSION TAG	UNP P94286
A	744	HIS	-	EXPRESSION TAG	UNP P94286
A	745	HIS	-	EXPRESSION TAG	UNP P94286
A	746	HIS	-	EXPRESSION TAG	UNP P94286
B	37	MET	-	EXPRESSION TAG	UNP P94286
B	38	GLY	-	EXPRESSION TAG	UNP P94286
B	278	PHE	SER	SEE REMARK 999	UNP P94286
B	308	ALA	ASP	ENGINEERED MUTATION	UNP P94286
B	739	LEU	-	EXPRESSION TAG	UNP P94286
B	740	GLU	-	EXPRESSION TAG	UNP P94286
B	741	HIS	-	EXPRESSION TAG	UNP P94286
B	742	HIS	-	EXPRESSION TAG	UNP P94286
B	743	HIS	-	EXPRESSION TAG	UNP P94286
B	744	HIS	-	EXPRESSION TAG	UNP P94286
B	745	HIS	-	EXPRESSION TAG	UNP P94286
B	746	HIS	-	EXPRESSION TAG	UNP P94286

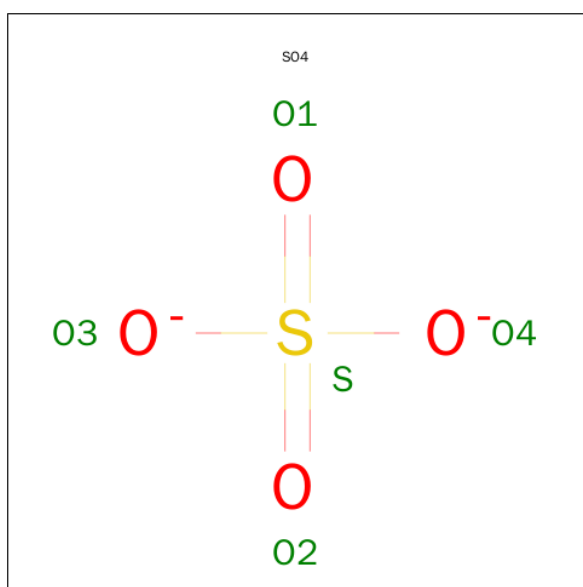
- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	1	Total	Ca	0	0
			1	1		
2	A	1	Total	Ca	0	0
			1	1		

- Molecule 3 is SODIUM ION (three-letter code: NA) (formula: Na).

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

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



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		
4	A	1	Total	O	S	0	0
			5	4	1		
4	A	1	Total	O	S	0	0
			5	4	1		
4	A	1	Total	O	S	0	0
			5	4	1		

*Continued on next page...*

Continued from previous page...

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		
4	B	1	Total	O	S	0	0
			5	4	1		
4	B	1	Total	O	S	0	0
			5	4	1		
4	B	1	Total	O	S	0	0
			5	4	1		
4	B	1	Total	O	S	0	0
			5	4	1		
4	B	1	Total	O	S	0	0
			5	4	1		

- Molecule 5 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula:  $C_2H_6O_2$ ).



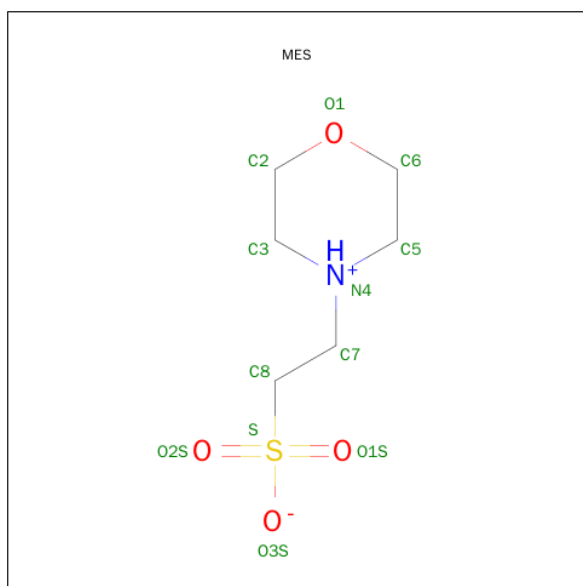
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			4	2	2		
5	A	1	Total	C	O	0	0
			4	2	2		
5	A	1	Total	C	O	0	0
			4	2	2		

Continued on next page...

Continued from previous page...

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

- Molecule 6 is 2-(N-MORPHOLINO)-ETHANESULFONIC ACID (three-letter code: MES) (formula: C<sub>6</sub>H<sub>13</sub>NO<sub>4</sub>S).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
6	A	1	Total	C	N	O	S	0	0
			12	6	1	4	1		
6	B	1	Total	C	N	O	S	0	0
			12	6	1	4	1		

- Molecule 7 is a polymer of unknown type called SUGAR (8-MER).

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	A	8	Total	C	O	0	0
			88	48	40		

Continued on next page...

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	A	8	Total	C	O	0	0
			88	48	40		
7	B	8	Total	C	O	0	0
			88	48	40		
7	B	8	Total	C	O	0	0
			88	48	40		

- Molecule 8 is a polymer of unknown type called SUGAR (4-MER).

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	B	4	Total	C	O	0	0
			45	24	21		

- Molecule 9 is water.

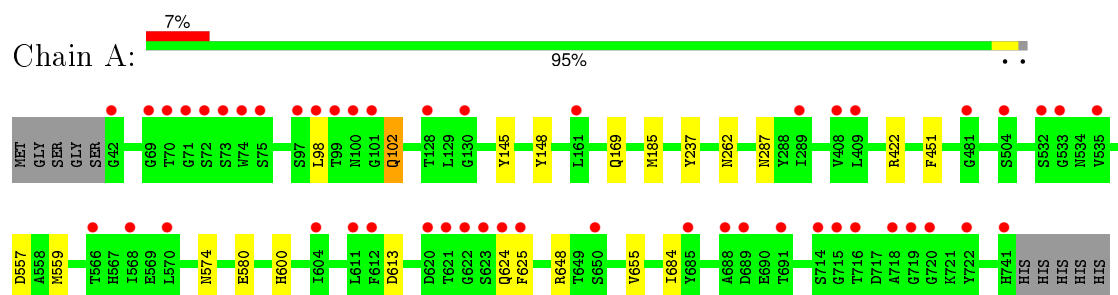
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	A	602	Total	O	0	0
			602	602		
9	B	624	Total	O	0	0
			624	624		



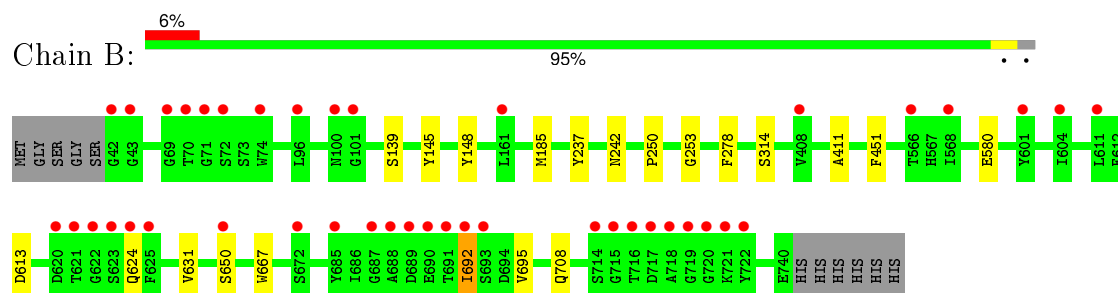
### 3 Residue-property plots [i](#)

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 ( $\text{RSRZ} > 2$ ). 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: Cycloisomaltooligosaccharide glucanotransferase



- Molecule 1: Cycloisomaltooligosaccharide glucanotransferase



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	61.93Å 167.72Å 174.86Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.47 – 1.90 30.45 – 1.90	Depositor EDS
% Data completeness (in resolution range)	99.3 (30.47-1.90) 99.3 (30.45-1.90)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.08	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.24 (at 1.91Å)	Xtriage
Refinement program	REFMAC 5.7.0029	Depositor
R, $R_{free}$	0.170 , 0.195 0.178 , 0.201	Depositor DCC
$R_{free}$ test set	7171 reflections (5.28%)	DCC
Wilson B-factor (Å <sup>2</sup> )	24.7	Xtriage
Anisotropy	0.178	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 46.0	EDS
Estimated twinning fraction	0.016 for -h,l,k	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Outliers	0 of 143084 reflections	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	12799	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	29.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.86% 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, 0.375 respectively for untwinned datasets, and 0.333, 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, CA, GLC, EDO, SO4, MES

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	0.37	0/5694	0.59	0/7754
1	B	0.37	0/5684	0.58	0/7741
All	All	0.37	0/11378	0.59	0/15495

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

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	5532	0	5134	7	0
1	B	5515	0	5115	6	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
4	A	35	0	0	0	0
4	B	30	0	0	0	0
5	A	20	0	30	0	0
5	B	16	0	24	0	0
6	A	12	0	13	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	B	12	0	13	0	0
7	A	176	0	144	1	0
7	B	176	0	144	2	0
8	B	45	0	39	0	0
9	A	602	0	0	1	0
9	B	624	0	0	0	0
All	All	12799	0	10656	13	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 (13) 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:145:TYR:OH	1:B:613:ASP:OD1	1.78	0.99
1:A:145:TYR:OH	1:A:613:ASP:OD1	2.08	0.67
1:A:169:GLN:HE22	1:A:574:ASN:HB2	1.74	0.53
1:A:98:LEU:HD22	1:A:102:GLN:HG3	1.92	0.50
1:A:580:GLU:HB2	7:A:817:GLC:H3	1.95	0.48
1:B:253:GLY:HA2	1:B:278:PHE:HA	1.96	0.47
1:A:422:ARG:HD3	9:A:1205:HOH:O	2.15	0.45
1:A:655:VAL:HG21	1:A:684:ILE:HD13	2.01	0.42
1:B:250:PRO:HG3	7:B:801:GLC:H61	2.01	0.42
1:A:557:ASP:HB3	1:A:600:HIS:CE1	2.54	0.42
1:B:692:ILE:HD11	1:B:695:VAL:HG23	2.01	0.42
1:B:580:GLU:HB2	7:B:823:GLC:H3	2.03	0.41
1:B:411:ALA:HB1	1:B:667:TRP:CH2	2.55	0.41

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	703/710 (99%)	680 (97%)	23 (3%)	0	100	100
1	B	703/710 (99%)	674 (96%)	29 (4%)	0	100	100
All	All	1406/1420 (99%)	1354 (96%)	52 (4%)	0	100	100

There are no Ramachandran outliers to report.

### 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	593/596 (100%)	582 (98%)	11 (2%)	65	59
1	B	593/596 (100%)	578 (98%)	15 (2%)	55	47
All	All	1186/1192 (100%)	1160 (98%)	26 (2%)	63	53

All (26) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	102	GLN
1	A	148	TYR
1	A	185	MET
1	A	237	TYR
1	A	262	ASN
1	A	287	ASN
1	A	451	PHE
1	A	559	MET
1	A	624	GLN
1	A	625	PHE
1	A	648	ARG
1	B	139[A]	SER
1	B	139[B]	SER
1	B	148	TYR
1	B	185	MET
1	B	237	TYR
1	B	242[A]	ASN
1	B	242[B]	ASN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	314[A]	SER
1	B	314[B]	SER
1	B	451	PHE
1	B	624	GLN
1	B	631	VAL
1	B	650	SER
1	B	692	ILE
1	B	708	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (12) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	262	ASN
1	A	315	ASN
1	A	320	ASN
1	A	389	ASN
1	A	479	ASN
1	A	521	GLN
1	A	619	ASN
1	A	624	GLN
1	B	320	ASN
1	B	496	GLN
1	B	624	GLN
1	B	666	GLN

### 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](#)

36 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$
7	GLC	A	816	7	11,11,12	0.55	0	14,15,17	0.65	0
7	GLC	A	817	7	11,11,12	0.60	0	14,15,17	0.53	0
7	GLC	A	818	7	11,11,12	0.60	0	14,15,17	0.75	0
7	GLC	A	819	7	11,11,12	0.47	0	14,15,17	0.96	1 (7%)
7	GLC	A	820	7	11,11,12	0.64	0	14,15,17	0.57	0
7	GLC	A	821	7	11,11,12	0.54	0	14,15,17	0.92	1 (7%)
7	GLC	A	822	7	11,11,12	0.53	0	14,15,17	0.83	0
7	GLC	A	823	7	11,11,12	0.57	0	14,15,17	0.79	0
7	GLC	A	824	7	11,11,12	0.55	0	14,15,17	0.78	1 (7%)
7	GLC	A	825	7	11,11,12	0.54	0	14,15,17	0.89	1 (7%)
7	GLC	A	826	7	11,11,12	0.57	0	14,15,17	0.98	1 (7%)
7	GLC	A	827	7	11,11,12	0.62	0	14,15,17	0.96	1 (7%)
7	GLC	A	828	7	11,11,12	0.54	0	14,15,17	0.90	1 (7%)
7	GLC	A	829	7	11,11,12	0.52	0	14,15,17	0.92	1 (7%)
7	GLC	A	830	7	11,11,12	0.54	0	14,15,17	0.84	1 (7%)
7	GLC	A	831	7	11,11,12	0.55	0	14,15,17	0.74	0
7	GLC	B	801	7	11,11,12	0.58	0	14,15,17	0.91	0
7	GLC	B	802	7	11,11,12	0.58	0	14,15,17	1.26	1 (7%)
7	GLC	B	803	7	11,11,12	0.58	0	14,15,17	0.99	0
7	GLC	B	804	7	11,11,12	0.63	0	14,15,17	2.15	3 (21%)
7	GLC	B	805	7	11,11,12	0.53	0	14,15,17	1.25	1 (7%)
7	GLC	B	806	7	11,11,12	0.50	0	14,15,17	0.95	0
7	GLC	B	807	7	11,11,12	0.51	0	14,15,17	2.43	2 (14%)
7	GLC	B	808	7	11,11,12	0.66	0	14,15,17	1.32	2 (14%)
7	GLC	B	822	7	11,11,12	0.62	0	14,15,17	0.79	0
7	GLC	B	823	7	11,11,12	0.69	0	14,15,17	0.57	0
7	GLC	B	824	7	11,11,12	0.50	0	14,15,17	0.67	0
7	GLC	B	825	7	11,11,12	0.42	0	14,15,17	0.76	0
7	GLC	B	826	7	11,11,12	0.54	0	14,15,17	0.79	0
7	GLC	B	827	7	11,11,12	0.55	0	14,15,17	0.92	0
7	GLC	B	828	7	11,11,12	0.60	0	14,15,17	0.78	0
7	GLC	B	829	7	11,11,12	0.60	0	14,15,17	0.87	0
8	GLC	B	830	8	11,11,12	0.48	0	14,15,17	0.70	0
8	GLC	B	831	8	11,11,12	0.57	0	14,15,17	0.89	0
8	GLC	B	832	8	11,11,12	0.49	0	14,15,17	1.23	1 (7%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
8	GLC	B	833	8	12,12,12	0.50	0	17,17,17	0.86	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
7	GLC	A	816	7	-	0/2/19/22	0/1/1/1
7	GLC	A	817	7	-	0/2/19/22	0/1/1/1
7	GLC	A	818	7	-	0/2/19/22	0/1/1/1
7	GLC	A	819	7	-	0/2/19/22	0/1/1/1
7	GLC	A	820	7	-	0/2/19/22	0/1/1/1
7	GLC	A	821	7	-	0/2/19/22	0/1/1/1
7	GLC	A	822	7	-	0/2/19/22	0/1/1/1
7	GLC	A	823	7	-	0/2/19/22	0/1/1/1
7	GLC	A	824	7	-	0/2/19/22	0/1/1/1
7	GLC	A	825	7	-	0/2/19/22	0/1/1/1
7	GLC	A	826	7	-	0/2/19/22	0/1/1/1
7	GLC	A	827	7	-	0/2/19/22	0/1/1/1
7	GLC	A	828	7	-	0/2/19/22	0/1/1/1
7	GLC	A	829	7	-	0/2/19/22	0/1/1/1
7	GLC	A	830	7	-	0/2/19/22	0/1/1/1
7	GLC	A	831	7	-	0/2/19/22	0/1/1/1
7	GLC	B	801	7	-	0/2/19/22	0/1/1/1
7	GLC	B	802	7	-	0/2/19/22	0/1/1/1
7	GLC	B	803	7	-	0/2/19/22	0/1/1/1
7	GLC	B	804	7	-	0/2/19/22	0/1/1/1
7	GLC	B	805	7	-	0/2/19/22	0/1/1/1
7	GLC	B	806	7	-	0/2/19/22	0/1/1/1
7	GLC	B	807	7	-	0/2/19/22	0/1/1/1
7	GLC	B	808	7	-	0/2/19/22	0/1/1/1
7	GLC	B	822	7	-	0/2/19/22	0/1/1/1
7	GLC	B	823	7	-	0/2/19/22	0/1/1/1
7	GLC	B	824	7	-	0/2/19/22	0/1/1/1
7	GLC	B	825	7	-	0/2/19/22	0/1/1/1
7	GLC	B	826	7	-	0/2/19/22	0/1/1/1
7	GLC	B	827	7	-	0/2/19/22	0/1/1/1
7	GLC	B	828	7	-	0/2/19/22	0/1/1/1
7	GLC	B	829	7	-	0/2/19/22	0/1/1/1
8	GLC	B	830	8	-	0/2/19/22	0/1/1/1
8	GLC	B	831	8	-	0/2/19/22	0/1/1/1
8	GLC	B	832	8	-	0/2/19/22	0/1/1/1

*Continued on next page...*



*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
8	GLC	B	833	8	-	0/2/22/22	0/1/1/1

There are no bond length outliers.

All (19) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
8	B	832	GLC	O5-C1-C2	-2.26	107.20	110.86
7	A	827	GLC	O6-C6-C5	-2.23	103.96	111.33
7	A	830	GLC	C1-O5-C5	2.03	114.82	112.25
7	A	819	GLC	C1-O5-C5	2.04	114.83	112.25
7	A	824	GLC	C1-O5-C5	2.22	115.06	112.25
7	B	804	GLC	O5-C1-C2	2.33	114.63	110.86
7	B	808	GLC	C3-C4-C5	2.35	114.30	110.20
7	A	828	GLC	C1-O5-C5	2.36	115.24	112.25
7	A	825	GLC	C1-O5-C5	2.39	115.28	112.25
7	A	829	GLC	C1-O5-C5	2.59	115.53	112.25
7	A	821	GLC	C1-O5-C5	2.61	115.56	112.25
7	A	826	GLC	C1-O5-C5	2.95	115.99	112.25
7	B	808	GLC	C1-O5-C5	3.09	116.17	112.25
7	B	802	GLC	C1-C2-C3	3.69	113.90	109.54
7	B	805	GLC	C1-O5-C5	4.11	117.46	112.25
7	B	807	GLC	C1-C2-C3	4.39	114.74	109.54
7	B	804	GLC	C1-O5-C5	5.06	118.67	112.25
7	B	804	GLC	C1-C2-C3	5.38	115.90	109.54
7	B	807	GLC	C1-O5-C5	7.50	121.77	112.25

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	A	817	GLC	1	0
7	B	801	GLC	1	0
7	B	823	GLC	1	0

## 5.6 Ligand geometry

Of 28 ligands modelled in this entry, 4 are monoatomic - leaving 24 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	SO4	A	803	-	4,4,4	0.16	0	6,6,6	0.17	0
4	SO4	A	804	-	4,4,4	0.48	0	6,6,6	0.30	0
4	SO4	A	805	-	4,4,4	0.32	0	6,6,6	0.16	0
4	SO4	A	806	-	4,4,4	0.45	0	6,6,6	0.32	0
4	SO4	A	807	-	4,4,4	0.38	0	6,6,6	0.15	0
4	SO4	A	808	-	4,4,4	0.42	0	6,6,6	0.06	0
4	SO4	A	809	-	4,4,4	0.37	0	6,6,6	0.09	0
5	EDO	A	810	-	3,3,3	0.47	0	2,2,2	0.27	0
5	EDO	A	811	-	3,3,3	0.39	0	2,2,2	0.68	0
5	EDO	A	812	-	3,3,3	0.41	0	2,2,2	0.63	0
5	EDO	A	813	-	3,3,3	0.45	0	2,2,2	0.43	0
5	EDO	A	814	-	3,3,3	0.43	0	2,2,2	0.50	0
6	MES	A	815	-	11,12,12	0.53	0	14,16,16	1.62	1 (7%)
4	SO4	B	811	-	4,4,4	0.21	0	6,6,6	0.24	0
4	SO4	B	812	-	4,4,4	0.34	0	6,6,6	0.08	0
4	SO4	B	813	-	4,4,4	0.43	0	6,6,6	0.40	0
4	SO4	B	814	-	4,4,4	0.48	0	6,6,6	0.10	0
4	SO4	B	815	-	4,4,4	0.38	0	6,6,6	0.11	0
4	SO4	B	816	-	4,4,4	0.29	0	6,6,6	0.14	0
5	EDO	B	817	-	3,3,3	0.39	0	2,2,2	0.59	0
5	EDO	B	818	-	3,3,3	0.42	0	2,2,2	0.49	0
5	EDO	B	819	-	3,3,3	0.40	0	2,2,2	0.41	0
5	EDO	B	820	-	3,3,3	0.43	0	2,2,2	0.30	0
6	MES	B	821	-	11,12,12	0.50	0	14,16,16	1.42	2 (14%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	SO4	A	803	-	-	0/0/0/0	0/0/0/0
4	SO4	A	804	-	-	0/0/0/0	0/0/0/0
4	SO4	A	805	-	-	0/0/0/0	0/0/0/0
4	SO4	A	806	-	-	0/0/0/0	0/0/0/0

Continued on next page...

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	SO4	A	807	-	-	0/0/0/0	0/0/0/0
4	SO4	A	808	-	-	0/0/0/0	0/0/0/0
4	SO4	A	809	-	-	0/0/0/0	0/0/0/0
5	EDO	A	810	-	-	0/1/1/1	0/0/0/0
5	EDO	A	811	-	-	0/1/1/1	0/0/0/0
5	EDO	A	812	-	-	0/1/1/1	0/0/0/0
5	EDO	A	813	-	-	0/1/1/1	0/0/0/0
5	EDO	A	814	-	-	0/1/1/1	0/0/0/0
6	MES	A	815	-	-	0/6/14/14	0/1/1/1
4	SO4	B	811	-	-	0/0/0/0	0/0/0/0
4	SO4	B	812	-	-	0/0/0/0	0/0/0/0
4	SO4	B	813	-	-	0/0/0/0	0/0/0/0
4	SO4	B	814	-	-	0/0/0/0	0/0/0/0
4	SO4	B	815	-	-	0/0/0/0	0/0/0/0
4	SO4	B	816	-	-	0/0/0/0	0/0/0/0
5	EDO	B	817	-	-	0/1/1/1	0/0/0/0
5	EDO	B	818	-	-	0/1/1/1	0/0/0/0
5	EDO	B	819	-	-	0/1/1/1	0/0/0/0
5	EDO	B	820	-	-	0/1/1/1	0/0/0/0
6	MES	B	821	-	-	0/6/14/14	0/1/1/1

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	B	821	MES	O1-C6-C5	-2.07	107.10	111.84
6	B	821	MES	O2S-S-C8	4.05	110.36	106.91
6	A	815	MES	O2S-S-C8	4.68	110.90	106.91

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 ⓘ

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	700/710 (98%)	0.32	49 (7%)	19 21	16, 26, 53, 100	0
1	B	699/710 (98%)	0.18	42 (6%)	25 28	16, 25, 51, 92	0
All	All	1399/1420 (98%)	0.25	91 (6%)	22 25	16, 25, 52, 100	0

All (91) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	621	THR	12.1
1	A	623	SER	10.7
1	B	718	ALA	8.4
1	B	621	THR	7.6
1	A	718	ALA	7.4
1	B	688	ALA	7.3
1	B	622	GLY	6.7
1	A	624	GLN	6.3
1	B	689	ASP	6.0
1	B	623	SER	5.8
1	A	622	GLY	5.8
1	B	719	GLY	5.7
1	B	42	GLY	5.6
1	B	716	THR	5.2
1	A	99	THR	4.8
1	A	101	GLY	4.7
1	A	70	THR	4.6
1	A	72	SER	4.5
1	A	716	THR	4.3
1	A	688	ALA	4.2
1	A	42	GLY	4.1
1	B	72	SER	3.9
1	A	741	HIS	3.8
1	B	71	GLY	3.8

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	692	ILE	3.8
1	A	481	GLY	3.8
1	A	532	SER	3.7
1	A	719	GLY	3.7
1	A	73	SER	3.7
1	B	720	GLY	3.6
1	B	722	TYR	3.6
1	A	504	SER	3.6
1	B	717	ASP	3.6
1	B	624	GLN	3.6
1	B	70	THR	3.6
1	A	604	ILE	3.5
1	B	101	GLY	3.5
1	A	722	TYR	3.5
1	A	535	VAL	3.5
1	B	685	TYR	3.4
1	B	721	LYS	3.4
1	B	96	LEU	3.2
1	A	408	VAL	3.2
1	A	100	ASN	3.2
1	B	625	PHE	3.1
1	A	685	TYR	3.1
1	B	620	ASP	3.1
1	A	720	GLY	3.0
1	A	71	GLY	3.0
1	A	689	ASP	3.0
1	A	128	THR	2.9
1	A	74	TRP	2.9
1	A	75	SER	2.9
1	A	130	GLY	2.9
1	B	650	SER	2.9
1	B	43	GLY	2.9
1	B	568	ILE	2.8
1	B	687	GLY	2.8
1	B	69	GLY	2.8
1	A	97	SER	2.8
1	B	691	THR	2.7
1	A	533	GLY	2.7
1	A	568	ILE	2.7
1	A	691	THR	2.7
1	B	611	LEU	2.7
1	A	409	LEU	2.7

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	715	GLY	2.7
1	B	100	ASN	2.6
1	B	715	GLY	2.6
1	B	408	VAL	2.5
1	B	714	SER	2.5
1	A	620	ASP	2.5
1	A	650	SER	2.5
1	A	566	THR	2.5
1	A	69	GLY	2.5
1	A	625	PHE	2.4
1	A	98	LEU	2.4
1	B	601	TYR	2.4
1	B	693	SER	2.3
1	A	611	LEU	2.3
1	A	714	SER	2.3
1	B	74	TRP	2.3
1	A	570	LEU	2.2
1	B	161	LEU	2.2
1	A	289	ILE	2.0
1	A	612	PHE	2.0
1	B	566	THR	2.0
1	B	672	SER	2.0
1	A	161	LEU	2.0
1	B	690	GLU	2.0
1	B	604	ILE	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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
7	GLC	B	808	11/12	0.76	0.24	15.10	49,53,54,57	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
7	GLC	B	801	11/12	0.84	0.17	1.10	35,40,45,50	0
7	GLC	B	823	11/12	0.98	0.12	0.47	15,16,17,18	0
8	GLC	B	832	11/12	0.95	0.10	0.19	26,30,33,35	0
7	GLC	B	829	11/12	0.98	0.08	-0.51	18,18,20,20	0
7	GLC	A	827	11/12	0.87	0.12	-0.60	45,48,50,54	0
7	GLC	B	822	11/12	0.98	0.08	-0.79	15,16,17,17	0
7	GLC	A	816	11/12	0.98	0.08	-0.89	15,16,16,17	0
7	GLC	A	818	11/12	0.97	0.10	-0.95	19,20,21,22	0
7	GLC	B	824	11/12	0.97	0.09	-1.04	17,18,20,20	0
7	GLC	B	825	11/12	0.97	0.07	-1.29	20,21,23,24	0
7	GLC	A	819	11/12	0.97	0.07	-1.47	22,23,24,25	0
7	GLC	B	827	11/12	0.97	0.07	-1.56	19,20,20,21	0
7	GLC	B	828	11/12	0.98	0.06	-1.64	19,20,22,24	0
7	GLC	A	817	11/12	0.98	0.09	-1.74	16,18,19,19	0
7	GLC	A	823	11/12	0.99	0.06	-1.99	17,18,19,19	0
7	GLC	A	822	11/12	0.98	0.06	-2.02	20,20,22,24	0
7	GLC	A	825	11/12	0.83	0.23	-	60,62,65,67	0
7	GLC	B	826	11/12	0.98	0.07	-	20,21,22,23	0
7	GLC	B	805	11/12	0.85	0.27	-	55,58,59,59	0
7	GLC	B	804	11/12	0.75	0.24	-	51,54,56,57	0
8	GLC	B	831	11/12	0.94	0.18	-	32,33,35,38	0
7	GLC	A	829	11/12	0.87	0.24	-	62,69,70,70	0
7	GLC	A	824	11/12	0.67	0.38	-	67,70,71,71	0
7	GLC	A	831	11/12	0.82	0.25	-	58,64,67,70	0
7	GLC	B	802	11/12	0.83	0.28	-	43,47,49,51	0
7	GLC	A	826	11/12	0.90	0.23	-	55,57,60,60	0
7	GLC	B	806	11/12	0.78	0.35	-	48,51,55,56	0
7	GLC	B	807	11/12	0.80	0.30	-	50,51,53,53	0
7	GLC	B	803	11/12	0.76	0.24	-	45,49,50,50	0
8	GLC	B	833	12/12	0.75	0.27	-	40,65,69,69	0
7	GLC	A	820	11/12	0.97	0.07	-	23,23,24,26	0
7	GLC	A	821	11/12	0.98	0.06	-	21,22,24,24	0
7	GLC	A	828	11/12	0.88	0.28	-	53,58,60,60	0
8	GLC	B	830	11/12	0.96	0.21	-	33,36,39,43	0
7	GLC	A	830	11/12	0.72	0.39	-	69,74,75,75	0

## 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
5	EDO	A	813	4/4	0.91	0.18	12.05	36,37,37,38	0
4	SO4	B	814	5/5	0.74	0.24	8.58	59,64,66,70	0
5	EDO	A	811	4/4	0.83	0.23	7.12	51,52,53,56	0
5	EDO	A	812	4/4	0.95	0.23	5.18	34,35,35,36	0
4	SO4	A	804	5/5	0.97	0.19	3.61	41,43,44,44	0
4	SO4	A	807	5/5	0.88	0.18	3.25	69,69,70,73	0
5	EDO	A	810	4/4	0.84	0.16	2.91	26,31,33,36	0
5	EDO	B	818	4/4	0.90	0.15	2.48	33,33,33,36	0
6	MES	A	815	12/12	0.96	0.12	0.03	28,29,31,31	0
6	MES	B	821	12/12	0.96	0.11	-0.34	26,26,28,30	0
3	NA	B	810	1/1	0.96	0.10	-0.55	28,28,28,28	0
3	NA	A	802	1/1	0.99	0.07	-1.22	24,24,24,24	0
2	CA	A	801	1/1	0.97	0.06	-2.41	30,30,30,30	0
2	CA	B	809	1/1	1.00	0.04	-2.82	26,26,26,26	0
4	SO4	B	811	5/5	0.98	0.12	-	40,41,42,45	0
5	EDO	A	814	4/4	0.90	0.17	-	52,53,53,54	0
5	EDO	B	817	4/4	0.88	0.14	-	41,44,44,45	0
4	SO4	A	808	5/5	0.91	0.26	-	63,66,68,68	0
4	SO4	A	809	5/5	0.91	0.34	-	83,84,85,87	0
4	SO4	B	812	5/5	0.90	0.34	-	76,76,77,78	0
4	SO4	A	805	5/5	0.91	0.40	-	80,81,83,84	0
4	SO4	B	815	5/5	0.85	0.23	-	76,77,79,79	0
4	SO4	B	816	5/5	0.97	0.26	-	52,53,55,55	0
4	SO4	A	806	5/5	0.90	0.19	-	46,51,54,55	0
5	EDO	B	819	4/4	0.94	0.14	-	32,34,34,37	0
4	SO4	B	813	5/5	0.87	0.29	-	53,57,62,64	0
5	EDO	B	820	4/4	0.93	0.19	-	41,43,44,44	0
4	SO4	A	803	5/5	0.98	0.13	-	41,41,43,45	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.