

# Full wwPDB X-ray Structure Validation Report (i)

Dec 17, 2023 – 02:52 pm GMT

PDB ID : 4TTU

Title: N-terminally truncated dextransucrase DSR-E from Leuconostoc mesen-

teroides NRRL B-1299 in complex with isomaltotriose

Authors: Brison, Y.; Remaud-Simeon, M.; Mourey, L.; Tranier, S.

Deposited on : 2014-06-23

Resolution : 2.18 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

https://www.wwpdb.org/validation/2017/XrayValidationReportHelp with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (1)) were used in the production of this report:

 $Mol Probity \quad : \quad 4.02b\text{-}467$ 

Mogul : 1.8.4, CSD as541be (2020)

Xtriage (Phenix) : 1.13 EDS : 2.36

Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)

 $Refmac \quad : \quad 5.8.0158$ 

CCP4 : 7.0.044 (Gargrove)

Ideal geometry (proteins) : Engh & Huber (2001) Ideal geometry (DNA, RNA) : Parkinson et al. (1996)

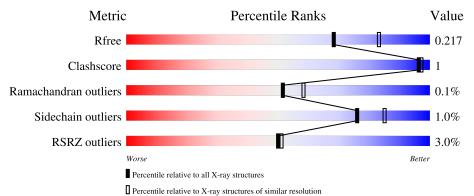
Validation Pipeline (wwPDB-VP) : 2.36

## 1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure: X- $RAY\ DIFFRACTION$ 

The reported resolution of this entry is 2.18 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 $(\# \mathrm{Entries})$	Similar resolution $(\# \text{Entries, resolution range}(\text{\AA}))$
$R_{free}$	130704	6864 (2.20-2.16)
Clashscore	141614	7689 (2.20-2.16)
Ramachandran outliers	138981	7564 (2.20-2.16)
Sidechain outliers	138945	7564 (2.20-2.16)
RSRZ outliers	127900	6738 (2.20-2.16)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=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 <=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	1108	90%	•	6%
2	В	2	100%		



## 2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 8955 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 Dextransucrase.

Mol	Chain	Residues		Atoms			ZeroOcc	AltConf	Trace		
1	Λ	1038	Total	С	I	N	О	S	0	Q	0
1	A	1030	8196	5130	3	1404	1641	18	U	8	U

There are 31 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1758	ALA	-	expression tag	UNP Q8G9Q2
A	2836	LYS	-	expression tag	UNP Q8G9Q2
A	2837	GLY	-	expression tag	UNP Q8G9Q2
A	2838	GLU	-	expression tag	UNP Q8G9Q2
A	2839	LEU	-	expression tag	UNP Q8G9Q2
A	2840	LYS	-	expression tag	UNP Q8G9Q2
A	2841	LEU	-	expression tag	UNP Q8G9Q2
A	2842	GLU	-	expression tag	UNP Q8G9Q2
A	2843	GLY	-	expression tag	UNP Q8G9Q2
A	2844	LYS	-	expression tag	UNP Q8G9Q2
A	2845	PRO	-	expression tag	UNP Q8G9Q2
A	2846	ILE	-	expression tag	UNP Q8G9Q2
A	2847	PRO	-	expression tag	UNP Q8G9Q2
A	2848	ASN	_	expression tag	UNP Q8G9Q2
A	2849	PRO	-	expression tag	UNP Q8G9Q2
A	2850	LEU	-	expression tag	UNP Q8G9Q2
A	2851	LEU	-	expression tag	UNP Q8G9Q2
A	2852	GLY	-	expression tag	UNP Q8G9Q2
A	2853	LEU	_	expression tag	UNP Q8G9Q2
A	2854	ASP	-	expression tag	UNP Q8G9Q2
A	2855	SER	-	expression tag	UNP Q8G9Q2
A	2856	THR	-	expression tag	UNP Q8G9Q2
A	2857	ARG	-	expression tag	UNP Q8G9Q2
A	2858	THR	-	expression tag	UNP Q8G9Q2
A	2859	GLY	-	expression tag	UNP Q8G9Q2
A	2860	HIS	-	expression tag	UNP Q8G9Q2
A	2861	HIS	-	expression tag	UNP Q8G9Q2

Continued on next page...



Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	2862	HIS	-	expression tag	UNP Q8G9Q2
A	2863	HIS	-	expression tag	UNP Q8G9Q2
A	2864	HIS	-	expression tag	UNP Q8G9Q2
A	2865	HIS	-	expression tag	UNP Q8G9Q2

• Molecule 2 is an oligosaccharide called alpha-D-glucopyranose-(1-6)-alpha-D-glucopyranose.



Mol	Chain	Residues	At	oms		ZeroOcc	AltConf	Trace
2	В	2	Total 23	C 12	O 11	0	0	0

• Molecule 3 is CALCIUM ION (three-letter code: CA) (formula: Ca).

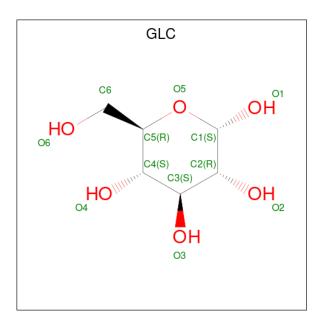
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total Ca 1 1	0	0

• Molecule 4 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total Na 1 1	0	0

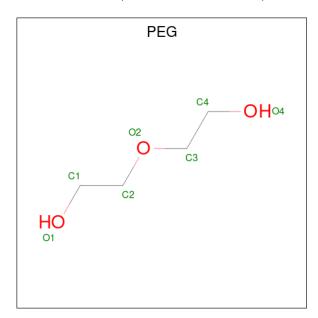
 $\bullet$  Molecule 5 is alpha-D-glucopyranose (three-letter code: GLC) (formula:  $\mathrm{C_6H_{12}O_6}).$ 





Mol	Chain	Residues	Atoms		ZeroOcc	AltConf	
5	A	1	Total 12	C 6	O 6	0	0

 $\bullet \ \ Molecule \ 6 \ is \ DI(HYDROXYETHYL)ETHER \ (three-letter \ code: \ PEG) \ (formula: \ C_4H_{10}O_3).$ 



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

• Molecule 7 is water.



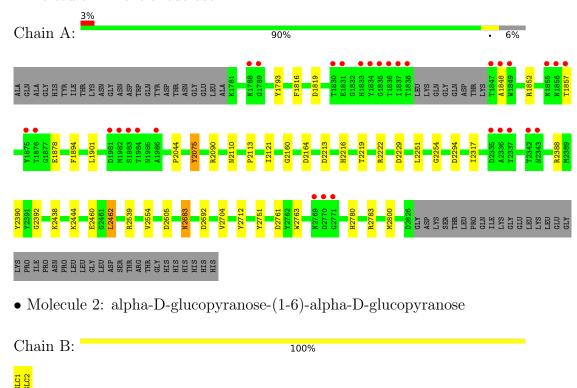
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	715	Total O 715 715	0	0



# 3 Residue-property plots (i)

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: Dextransucrase





## 4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants	68.29Å 100.10Å 184.43Å	Depositor
a, b, c, $\alpha$ , $\beta$ , $\gamma$	90.00° 90.00° 90.00°	Depositor
Resolution (Å)	43.99 - 2.18	Depositor
rtesolution (A)	43.99 - 2.18	EDS
% Data completeness	100.0 (43.99-2.18)	Depositor
(in resolution range)	100.0 (43.99-2.18)	EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	(Not available)	Depositor
$< I/\sigma(I) > 1$	3.05  (at  2.18Å)	Xtriage
Refinement program	REFMAC 5.8.0103	Depositor
P. P.	0.166 , 0.213	Depositor
$R, R_{free}$	0.174 , $0.217$	DCC
$R_{free}$ test set	3388 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	23.8	Xtriage
Anisotropy	0.035	Xtriage
Bulk solvent $k_{sol}(e/Å^3)$ , $B_{sol}(Å^2)$	0.30, 34.9	EDS
L-test for twinning <sup>2</sup>	$ < L > = 0.47, < L^2> = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	8955	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	28.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: The largest off-origin peak in the Patterson function is 3.87% of the height of the origin peak. No significant pseudotranslation is detected.

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



<sup>&</sup>lt;sup>1</sup>Intensities estimated from amplitudes.

## 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: IYR, PEG, GLC, TYI, CA, NA

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).

Mal	Chain	Boı	nd lengths	Во	ond angles
MIOI	$\begin{array}{c c} \text{Col} & \text{Chain} & \text{RMSZ} \end{array}$		# Z  > 5	RMSZ	# Z  > 5
1	A	0.85	1/8378 (0.0%)	0.84	9/11387 (0.1%)

#### All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(A)	$\operatorname{Ideal}( ext{\AA})$
1	A	2783	ARG	CZ-NH2	5.66	1.40	1.33

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	$\mathbf{Z}$	$\mathbf{Observed}(^o)$	$\operatorname{Ideal}({}^{o})$
1	A	2222	ARG	NE-CZ-NH1	6.81	123.70	120.30
1	A	2692	ASP	CB-CG-OD1	6.33	123.99	118.30
1	A	2692	ASP	CB-CG-OD2	-6.02	112.88	118.30
1	A	2164	ASP	CB-CG-OD1	5.90	123.61	118.30
1	A	2704	VAL	CB-CA-C	-5.81	100.37	111.40
1	A	1852	ARG	NE-CZ-NH1	5.64	123.12	120.30
1	A	2605	ASP	CB-CG-OD2	-5.46	113.39	118.30
1	A	2800	MET	CG-SD-CE	5.20	108.52	100.20
1	A	2213	ASP	CB-CG-OD2	-5.07	113.74	118.30

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	8196	0	7713	16	0
2	В	23	0	21	0	0
3	A	1	0	0	0	0
4	A	1	0	0	0	0
5	A	12	0	12	0	0
6	A	7	0	10	0	0
7	A	715	0	0	2	0
All	All	8955	0	7756	16	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 (16) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	$\begin{array}{c} {\rm Interatomic} \\ {\rm distance} \ ({\rm \AA}) \end{array}$	Clash overlap (Å)
1:A:2075:TYI:I1	7:A:3686:HOH:O	2.62	0.86
1:A:2444:LYS:HB2	1:A:2460:GLU:OE1	2.06	0.55
1:A:2462:LEU:HB3	1:A:2554:VAL:HG12	1.88	0.55
1:A:2761:ASP:OD1	1:A:2780[A]:HIS:ND1	2.40	0.53
1:A:2254:GLY:HA3	1:A:2763:TRP:O	2.09	0.51
1:A:2683:ASN:ND2	7:A:3015:HOH:O	2.45	0.50
1:A:2044:PRO:HG3	1:A:2121:ILE:HD11	1.94	0.50
1:A:2390:TYR:CE2	1:A:2392:GLY:HA3	2.48	0.48
1:A:1894:PHE:CZ	1:A:1901:LEU:HD12	2.49	0.47
1:A:2229:ASP:HB3	1:A:2751:TYR:CD1	2.51	0.46
1:A:2317:ILE:HG22	1:A:2388:ARG:HB3	1.99	0.43
1:A:1793:TYR:HB2	1:A:1816:PHE:CZ	2.55	0.41
1:A:2160:GLY:HA3	1:A:2216:HIS:CD2	2.54	0.41
1:A:1848:ALA:HB1	1:A:1878:GLU:HG3	2.03	0.41
1:A:2251:LEU:HD12	1:A:2294:ASP:HB3	2.02	0.41
1:A:2044:PRO:CG	1:A:2121:ILE:HD11	2.50	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	1040/1108 (94%)	1005 (97%)	34 (3%)	1 (0%)	51	58

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1819	ASP

#### 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	868/923 (94%)	858 (99%)	10 (1%)	71 81

All (10) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1857	THR
1	A	2090[A]	ARG
1	A	2090[B]	ARG
1	A	2110	ASN
1	A	2113	PRO
1	A	2219	THR
1	A	2438	LYS
1	A	2462	LEU
1	A	2539	ARG
1	A	2683	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	1927	GLN
1	A	2683	ASN
1	A	2728	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)

2 non-standard protein/DNA/RNA residues 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	Dec	Dec	Dec	Dog	Dec	Link	Bo	Bond lengths			Bond angles		
WIOI	Type	Chain	nes	Lilik	Counts	RMSZ	# Z  > 2	Counts	RMSZ	$\mid \# Z  > 2$						
1	IYR	A	2712	1	12,13,14	1.38	1 (8%)	14,17,19	2.27	5 (35%)						
1	TYI	A	2075	1	13,14,15	1.71	2 (15%)	16,19,21	0.98	1 (6%)						

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
1	IYR	A	2712	1	-	0/5/6/8	0/1/1/1
1	TYI	A	2075	1	=	0/5/6/8	0/1/1/1

#### All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	$\mathbf{Z}$	$\operatorname{Observed}(\text{\AA})$	Ideal(A)
1	A	2075	TYI	CZ-CE2	4.05	1.49	1.40
1	A	2075	TYI	CZ-CE1	3.96	1.49	1.40
1	A	2712	IYR	CF-CE	3.52	1.46	1.39

#### All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	$\mathbf{Z}$	$Observed(^o)$	$Ideal(^{o})$
1	A	2712	IYR	CF-CE-IE	-5.48	114.37	119.81
1	A	2712	IYR	CD-CE-IE	3.70	125.43	118.61
1	A	2712	IYR	CB-CC-CD	2.91	125.43	120.44
1	A	2712	IYR	CB-CC-CH	-2.68	115.59	120.91

Continued on next page...



Continued from previous page...

Mol	Chain	Res	Type	Atoms	$\mathbf{Z}$	$\mathbf{Observed}(^o)$	$\operatorname{Ideal}({}^{o})$
1	A	2712	IYR	CC-CB-CA	-2.47	109.09	114.10
1	A	2075	TYI	CZ-CE2-I2	2.21	122.83	119.42

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	2075	TYI	1	0

### 5.5 Carbohydrates (i)

2 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	Trino	Chain	Dag	Link	Во	ond leng	$ ag{ths}$	Bond angles		
IVIOI	Type	Chain	Res	Lilik	Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	GLC	В	1	2	12,12,12	0.47	0	17,17,17	1.25	1 (5%)
2	GLC	В	2	2	11,11,12	0.98	0	15,15,17	1.44	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	GLC	В	1	2	-	0/2/22/22	0/1/1/1
2	GLC	В	2	2	-	2/2/19/22	0/1/1/1

There are no bond length outliers.

All (3) bond angle outliers are listed below:



Mol	Chain	Res	Type	Atoms	Z	$Observed(^o)$	$\operatorname{Ideal}({}^{o})$
2	В	2	GLC	O3-C3-C2	-2.73	104.77	109.99
2	В	2	GLC	C3-C4-C5	2.56	114.81	110.24
2	В	1	GLC	C4-C3-C2	-2.53	106.41	110.82

There are no chirality outliers.

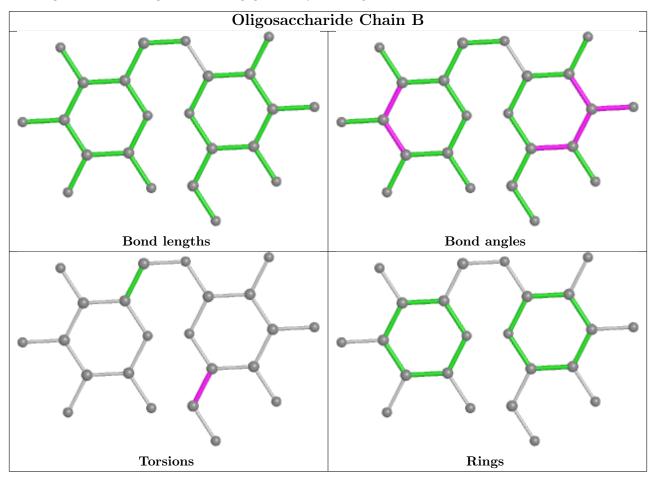
All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	В	2	GLC	C4-C5-C6-O6
2	В	2	GLC	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)

Of 4 ligands modelled in this entry, 2 are monoatomic - leaving 2 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	Bo	ond leng	$ ag{ths}$	Bond angles		
MIOI	туре	Chain	nes	Lilik	Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
5	GLC	A	2905	-	12,12,12	1.37	1 (8%)	17,17,17	1.35	2 (11%)
6	PEG	A	2906	-	6,6,6	0.69	0	5,5,5	0.58	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	GLC	A	2905	-	-	2/2/22/22	0/1/1/1
6	PEG	A	2906	-	-	1/4/4/4	-

#### All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$\operatorname{Observed}(\text{\AA})$	$\operatorname{Ideal}( ext{\AA})$
5	A	2905	GLC	C1-C2	2.39	1.58	1.52

#### All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$\operatorname{Observed}(^{o})$	$\operatorname{Ideal}({}^{o})$
5	A	2905	GLC	C1-C2-C3	2.89	116.32	110.31
5	A	2905	GLC	O5-C5-C4	2.09	113.49	109.69

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	2905	GLC	C4-C5-C6-O6
5	A	2905	GLC	O5-C5-C6-O6
6	A	2906	PEG	O2-C3-C4-O4



There are no ring outliers.

No monomer is involved in short contacts.

# 5.7 Other polymers (i)

There are no such residues in this entry.

## 5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.



## 6 Fit of model and data (i)

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

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,  $95^{th}$  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>	$\# \mathrm{RSRZ}{>}2$	$OWAB(Å^2)$	Q<0.9
1	A	1036/1108 (93%)	-0.37	31 (2%) 50 51	12, 23, 55, 107	0

All (31) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	2769	ASN	5.3
1	A	1984	THR	4.7
1	A	1982	MET	4.4
1	A	1983	SER	4.4
1	A	1837	ILE	4.4
1	A	2771	GLY	3.7
1	A	1849	TRP	3.6
1	A	1835	GLY	3.5
1	A	1789	GLY	3.3
1	A	1788	ASN	3.2
1	A	1847	THR	3.1
1	A	1848	ALA	3.1
1	A	1838	THR	3.0
1	A	1834	TYR	3.0
1	A	1981	ASP	3.0
1	A	1986	ALA	2.9
1	A	1875	TYR	2.9
1	A	2336	ALA	2.9
1	A	2335	ASP	2.7
1	A	1836	THR	2.7
1	A	1857	THR	2.7
1	A	2770	ASP	2.6
1	A	1830	THR	2.6
1	A	1855	ASN	2.6
1	A	1876	THR	2.5
1	A	1856	ASN	2.4
1	A	2342	THR	2.4

Continued on next page...



Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	2337	THR	2.2
1	A	1833	HIS	2.2
1	A	2343	ASN	2.1
1	A	1831	GLU	2.0

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

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median,  $95^{th}$  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	${f B-factors}({f \AA}^2)$	Q<0.9
1	IYR	A	2712	13/14	0.95	0.08	18,27,35,49	1
1	TYI	A	2075	14/15	0.96	0.13	28,47,70,74	2

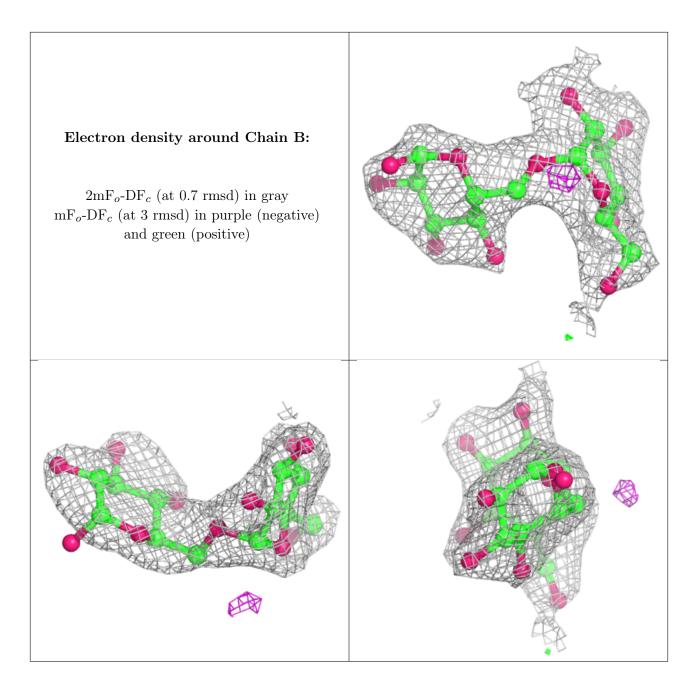
## 6.3 Carbohydrates (i)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median,  $95^{th}$  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	$\mathbf{B} ext{-}\mathbf{factors}(\mathbf{\mathring{A}}^2)$	Q<0.9
2	GLC	В	2	11/12	0.85	0.16	56,61,65,65	0
2	GLC	В	1	12/12	0.91	0.20	60,64,70,76	0

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





## 6.4 Ligands (i)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median,  $95^{th}$  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	$\mathbf{B} ext{-}\mathbf{factors}(\mathbf{\mathring{A}}^2)$	Q<0.9
5	GLC	A	2905	12/12	0.77	0.33	41,64,69,72	0
6	PEG	A	2906	7/7	0.85	0.12	43,46,54,54	0
4	NA	A	2902	1/1	0.98	0.09	33,33,33,33	0
3	CA	A	2901	1/1	1.00	0.13	15,15,15,15	0



## 6.5 Other polymers (i)

There are no such residues in this entry.

