



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 4, 2026 – 10:57 PM UTC

PDB ID : 4G68 / pdb_00004g68
Title : Biochemical and structural insights into xylan utilization by the thermophilic bacterium *Caldanaerobius polysaccharolyticus*
Authors : Agarwal, V.; Nair, S.K.
Deposited on : 2012-07-18
Resolution : 1.80 Å (reported)

This is a wwPDB X-ray Structure Validation Summary 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 ⓘ 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 ⓘ](#)) were used in the production of this report:

MolProbity	:	4-5-2 with Phenix2.0
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

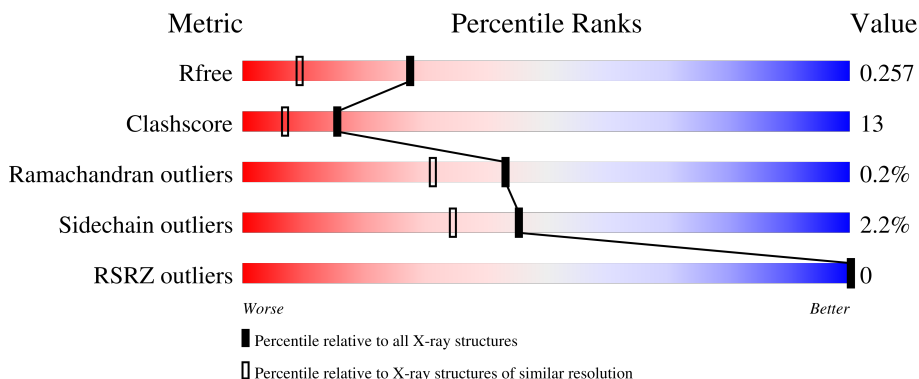
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.80 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}	180053	7662 (1.80-1.80)
Clashscore	190562	8479 (1.80-1.80)
Ramachandran outliers	187476	8391 (1.80-1.80)
Sidechain outliers	187428	8390 (1.80-1.80)
RSRZ outliers	180081	7663 (1.80-1.80)

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 $\leq 5\%$. 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	456	
2	B	432	
3	C	432	
4	D	3	
4	E	3	

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 10210 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 ABC transporter.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	392	Total	C	N	O	S	Se	0	2	0
			3058	1957	494	597	4	6			

- Molecule 2 is a protein called ABC transporter.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	B	392	Total	C	N	O	S	Se	0	0	0
			3050	1951	494	596	4	5			

- Molecule 3 is a protein called ABC transporter.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
3	C	392	Total	C	N	O	S	Se	0	0	0
			3050	1951	494	596	8	1			

- Molecule 4 is an oligosaccharide called beta-D-xylopyranose-(1-4)-beta-D-xylopyranose-(1-4)-alpha-D-xylopyranose.



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace
4	D	3	Total	C	O	0	0	0
			28	15	13			
4	E	3	Total	C	O	0	0	0
			28	15	13			

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	364	Total 364	O 364	0	0
5	B	363	Total 363	O 363	0	0
5	C	269	Total 269	O 269	0	0

i

- Molecule 1: ABC transporter





- Molecule 4: beta-D-xylopyranose-(1-4)-beta-D-xylopyranose-(1-4)-alpha-D-xylopyranose

Chain D: 100%



- Molecule 4: beta-D-xylopyranose-(1-4)-beta-D-xylopyranose-(1-4)-alpha-D-xylopyranose

Chain E: 100%



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	59.33Å 150.86Å 150.88Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	25.00 – 1.80 25.00 – 1.80	Depositor EDS
% Data completeness (in resolution range)	100.0 (25.00-1.80) 99.7 (25.00-1.80)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	6.42 (at 1.81Å)	Xtriage
Refinement program	REFMAC 5.5.0056	Depositor
R, R_{free}	0.230 , 0.261 0.215 , 0.257	Depositor DCC
R_{free} test set	6314 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å ²)	13.8	Xtriage
Anisotropy	0.257	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 20.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.43$, $\langle L^2 \rangle = 0.25$	Xtriage
Estimated twinning fraction	0.478 for -h,l,k	Xtriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	10210	wwPDB-VP
Average B, all atoms (Å ²)	17.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.92% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: XYP, XYS

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	1.33	7/3125 (0.2%)	1.15	4/4224 (0.1%)
2	B	1.32	7/3112 (0.2%)	1.13	2/4209 (0.0%)
3	C	1.14	5/3116 (0.2%)	1.16	6/4221 (0.1%)
All	All	1.27	19/9353 (0.2%)	1.15	12/12654 (0.1%)

The worst 5 of 19 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	364	MET	SD-CE	-8.28	1.58	1.79
3	C	187	TRP	C-O	-5.86	1.19	1.24
1	A	150	PHE	N-CA	5.83	1.53	1.46
2	B	383	SER	CA-C	-5.74	1.45	1.52
1	A	195	MSE	CA-C	-5.43	1.46	1.52

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	195	MET	CG-SD-CE	-8.67	81.82	100.90
3	C	365	ASN	N-CA-C	7.36	119.09	111.14
1	A	195	MSE	CG-SE-CE	-6.65	84.30	98.92
1	A	184	LYS	N-CA-C	5.93	118.50	111.33
3	C	380	ILE	CB-CA-C	-5.86	103.39	111.65

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	3058	0	2995	36	1
2	B	3050	0	2981	37	1
3	C	3050	0	2981	173	2
4	D	28	0	9	0	0
4	E	28	0	9	0	0
5	A	364	0	0	15	0
5	B	363	0	0	14	0
5	C	269	0	0	126	0
All	All	10210	0	8975	244	2

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 13.

The worst 5 of 244 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)
2:B:218:ASN:HB3	5:B:1245:HOH:O	1.21	1.29
3:C:69:LYS:HD2	5:C:701:HOH:O	1.25	1.27
3:C:363:ILE:HG22	5:C:764:HOH:O	1.13	1.27
3:C:198:LEU:HG	5:C:765:HOH:O	1.28	1.27
3:C:279:VAL:HG21	5:C:584:HOH:O	1.26	1.26

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:162:SER:O	3:C:154:ASN:ND2[1_455]	2.02	0.18
1:A:347:LYS:NZ	3:C:295:LYS:NZ[1_455]	2.13	0.07

5.3 Torsion angles

5.3.1 Protein backbone

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	392/456 (86%)	386 (98%)	5 (1%)	1 (0%)	36	25
2	B	390/432 (90%)	382 (98%)	8 (2%)	0	100	100
3	C	390/432 (90%)	379 (97%)	10 (3%)	1 (0%)	36	25
All	All	1172/1320 (89%)	1147 (98%)	23 (2%)	2 (0%)	43	31

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	C	154	ASN
1	A	65	ASN

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	321/375 (86%)	316 (98%)	5 (2%)	55	47
2	B	319/352 (91%)	313 (98%)	6 (2%)	50	41
3	C	319/356 (90%)	309 (97%)	10 (3%)	35	23
All	All	959/1083 (89%)	938 (98%)	21 (2%)	45	34

5 of 21 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	C	138	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	C	244	LEU
3	C	372	LYS
3	C	280	LYS
3	C	228	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 26 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	350	ASN
3	C	48	GLN
3	C	353	GLN
2	B	362	GLN
3	C	57	GLN

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 ⓘ

6 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	$\# Z > 2$
4	XYS	D	1	4	10,10,10	2.01	3 (30%)	14,14,14	2.63	8 (57%)
4	XYP	D	2	4	9,9,10	0.97	0	10,12,14	2.09	4 (40%)
4	XYP	D	3	4	9,9,10	1.00	0	10,12,14	1.31	1 (10%)
4	XYS	E	1	4	10,10,10	1.93	3 (30%)	14,14,14	2.45	7 (50%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
4	XYP	E	2	4	9,9,10	1.19	1 (11%)	10,12,14	2.06	2 (20%)
4	XYP	E	3	4	9,9,10	0.92	0	10,12,14	1.03	1 (10%)

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	XYS	D	1	4	-	-	0/1/1/1
4	XYP	D	2	4	-	-	0/1/1/1
4	XYP	D	3	4	-	-	0/1/1/1
4	XYS	E	1	4	-	-	0/1/1/1
4	XYP	E	2	4	-	-	0/1/1/1
4	XYP	E	3	4	-	-	0/1/1/1

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	D	1	XYS	O5-C5	-4.87	1.35	1.43
4	E	1	XYS	O5-C5	-4.30	1.36	1.43
4	E	1	XYS	O5-C1	-2.69	1.38	1.43
4	E	2	XYP	C2-C3	2.46	1.56	1.52
4	E	1	XYS	O2-C2	-2.32	1.37	1.43

The worst 5 of 23 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	D	1	XYS	C5-O5-C1	6.29	125.54	112.46
4	E	1	XYS	C5-O5-C1	5.63	124.17	112.46
4	E	2	XYP	C4-C3-C2	-4.77	105.26	110.92
4	D	2	XYP	C4-C3-C2	-4.40	105.70	110.92
4	D	1	XYS	O5-C5-C4	3.44	118.99	110.79

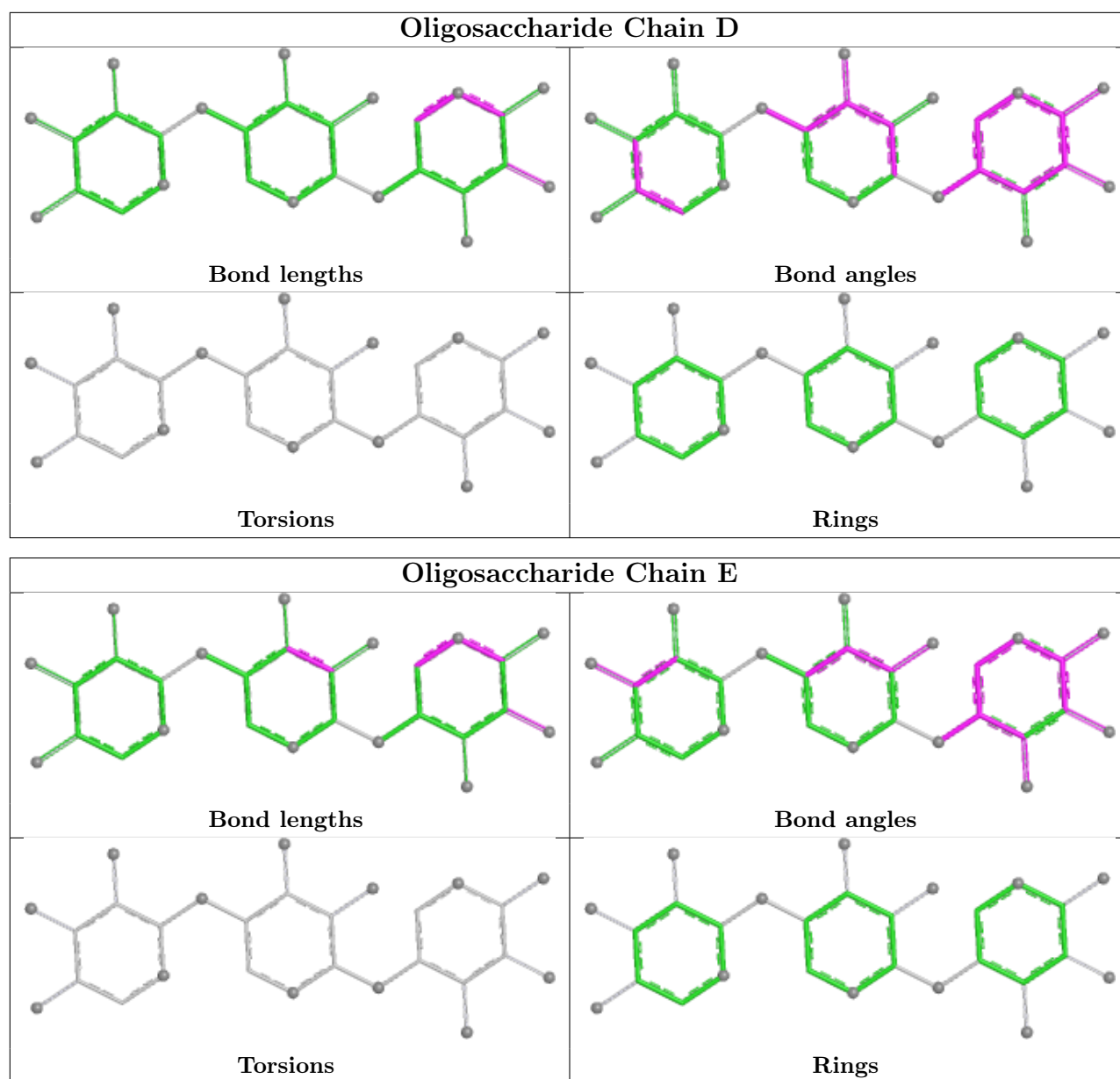
There are no chirality outliers.

There are no torsion outliers.

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](#)

There are no ligands in this entry.

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, 95th 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 ²)	Q<0.9
1	A	386/456 (84%)	-1.69	0 100 100	4, 9, 19, 36	2 (0%)
2	B	387/432 (89%)	-1.69	0 100 100	4, 9, 18, 35	0
3	C	391/432 (90%)	-1.37	0 100 100	15, 28, 41, 54	0
All	All	1164/1320 (88%)	-1.58	0 100 100	4, 13, 36, 54	2 (0%)

There are no RSRZ outliers to report.

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](#)

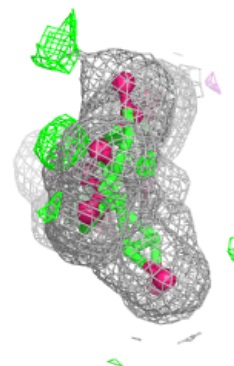
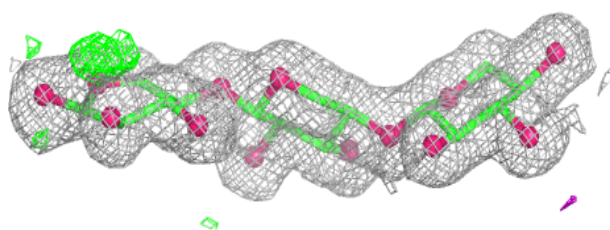
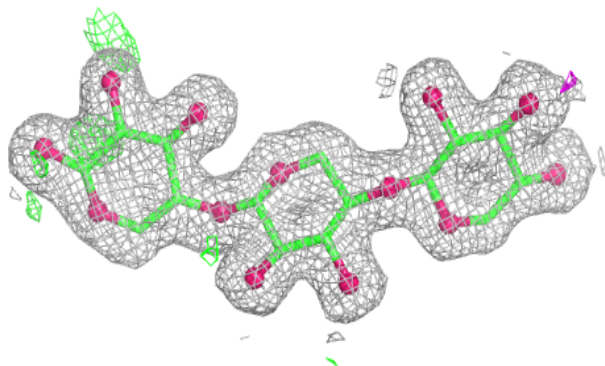
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q< 0.9’ lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
4	XYS	D	1	10/10	0.99	0.02	8,10,15,17	0
4	XYP	D	2	9/10	1.00	0.01	5,5,6,6	0
4	XYP	D	3	9/10	1.00	0.01	4,4,4,4	0
4	XYS	E	1	10/10	1.00	0.02	9,11,14,19	0
4	XYP	E	2	9/10	1.00	0.01	4,5,5,5	0
4	XYP	E	3	9/10	1.00	0.01	3,3,4,4	0

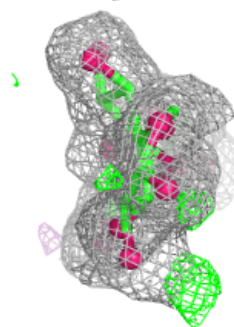
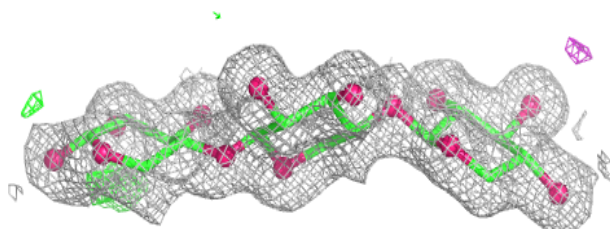
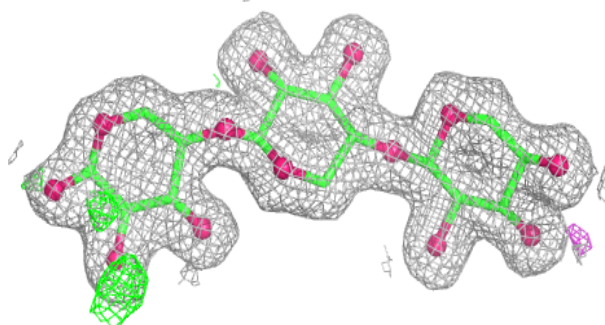
The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around Chain D:

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

**Electron density around Chain E:**

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



6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.