wwPDB X-ray Structure Validation Summary Report (i)

Mar 24, 2022 – 03:32 pm GMT

PDB ID	:	6R5T
Title	:	The crystal structure of Glycoside Hydrolase BglX inactive mutant D286N from P. aeruginosa in complex with lactose
Authors	:	Batuecas, M.T.; Hermoso, J.A.
Deposited on	:	2019-03-25
Resolution	:	1.60 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 (i) symbol.

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

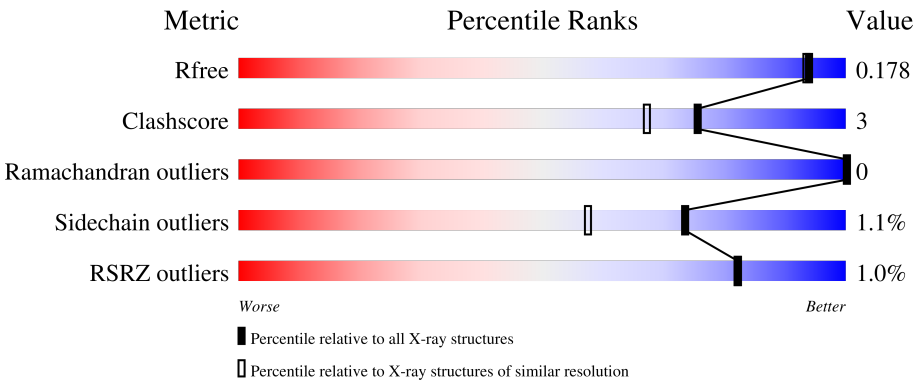
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.27
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.27

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 1.60 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}	130704	3398 (1.60-1.60)
Clashscore	141614	3665 (1.60-1.60)
Ramachandran outliers	138981	3564 (1.60-1.60)
Sidechain outliers	138945	3563 (1.60-1.60)
RSRZ outliers	127900	3321 (1.60-1.60)

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	733	2% 92% 8%
1	B	733	94% 5%
2	C	2	100%
2	D	2	100%

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 12876 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.

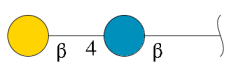
- Molecule 1 is a protein called Periplasmic beta-glucosidase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	B	733	5687	3579	1017	1071	20	0	12	0
1	A	733	5711	3601	1018	1071	21	0	18	0

There are 2 discrepancies between the modelled and reference sequences:

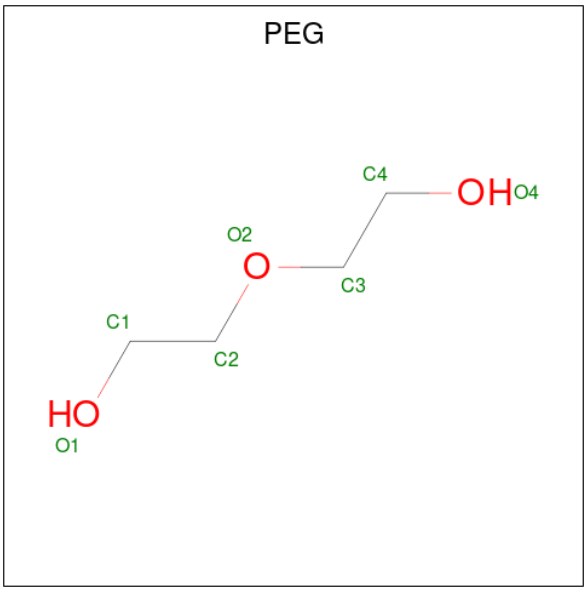
Chain	Residue	Modelled	Actual	Comment	Reference
B	286	ASN	ASP	engineered mutation	UNP Q9I311
A	286	ASN	ASP	engineered mutation	UNP Q9I311

- Molecule 2 is an oligosaccharide called beta-D-galactopyranose-(1-4)-beta-D-glucopyranose.



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace
			Total	C	O			
2	C	2	23	12	11	0	0	0
2	D	2	23	12	11	0	0	0

- Molecule 3 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: $C_4H_{10}O_3$).

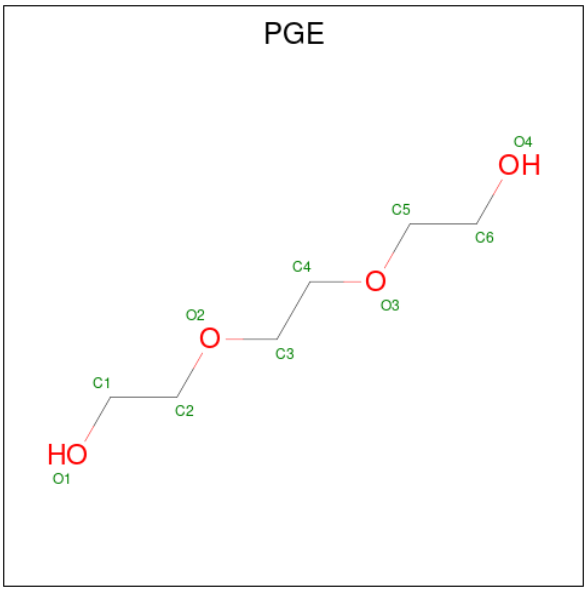


Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	1	Total C O 7 4 3	0	0
3	A	1	Total C O 7 4 3	0	0
3	A	1	Total C O 7 4 3	0	0
3	A	1	Total C O 7 4 3	0	0

- Molecule 4 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	B	1	Total Mg 1 1	0	0
4	A	1	Total Mg 1 1	0	0

- Molecule 5 is TRIETHYLENE GLYCOL (three-letter code: PGE) (formula: $C_6H_{14}O_4$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
5	A	1	10	6	4	0	0

- Molecule 6 is water.

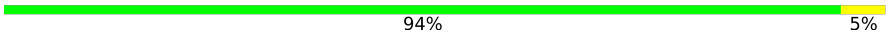
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
6	B	718	718	718	0	0
			Total	O		
6	A	674	674	674	0	0

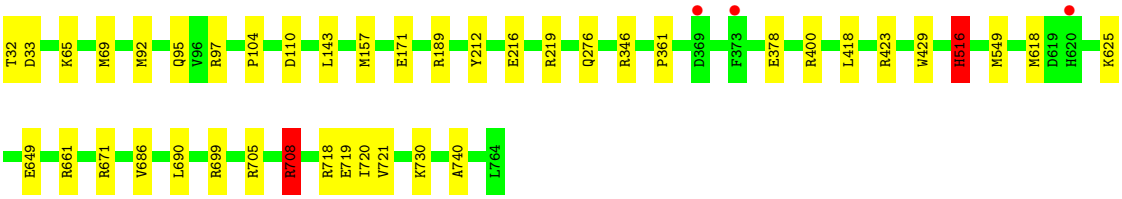
3 Residue-property plots (i)

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: Periplasmic beta-glucosidase

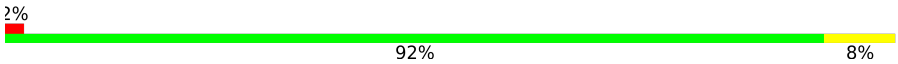
Chain B:

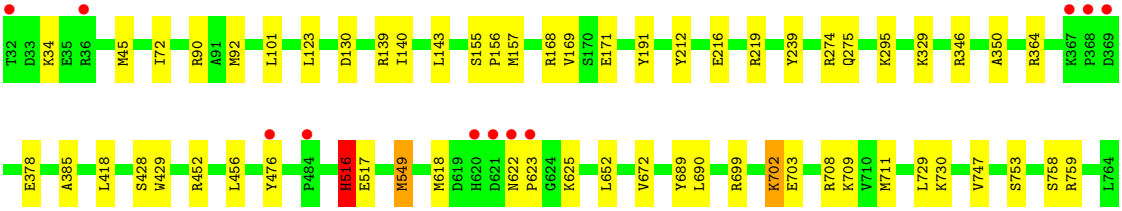




• Molecule 1: Periplasmic beta-glucosidase

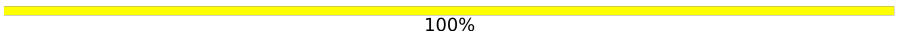
Chain A:





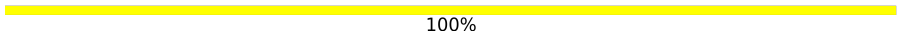
• Molecule 2: beta-D-galactopyranose-(1-4)-beta-D-glucopyranose

Chain C:



• Molecule 2: beta-D-galactopyranose-(1-4)-beta-D-glucopyranose

Chain D:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	70.98Å 73.59Å 81.12Å 65.55° 73.42° 69.68°	Depositor
Resolution (Å)	45.17 – 1.60 45.17 – 1.60	Depositor EDS
% Data completeness (in resolution range)	96.1 (45.17-1.60) 96.1 (45.17-1.60)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.38 (at 1.60Å)	Xtriage
Refinement program	REFMAC 5.8.0238	Depositor
R, R_{free}	0.139 , 0.168 0.153 , 0.178	Depositor DCC
R_{free} test set	8830 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å^2)	17.9	Xtriage
Anisotropy	0.042	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	12876	wwPDB-VP
Average B, all atoms (Å^2)	23.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.54% of the height of the origin peak. No significant pseudotranslation is detected.*

[1]Intensities estimated from amplitudes.

[2]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.

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: PEG, GAL, MG, PGE, BGC

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	0.77	3/5876 (0.1%)	0.94	17/7963 (0.2%)
1	B	0.78	2/5837 (0.0%)	0.92	6/7911 (0.1%)
All	All	0.77	5/11713 (0.0%)	0.93	23/15874 (0.1%)

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	378	GLU	CD-OE2	-7.96	1.16	1.25
1	A	216	GLU	CD-OE1	-5.75	1.19	1.25
1	B	216	GLU	CD-OE1	-5.64	1.19	1.25
1	A	378	GLU	CD-OE1	-5.23	1.19	1.25
1	A	171	GLU	CD-OE1	5.05	1.31	1.25

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	699	ARG	NE-CZ-NH2	-9.66	115.47	120.30
1	A	274	ARG	NE-CZ-NH2	-9.64	115.48	120.30
1	A	219	ARG	NE-CZ-NH1	8.22	124.41	120.30
1	A	549	MET	CG-SD-CE	-7.66	87.95	100.20
1	A	168	ARG	NE-CZ-NH2	-6.83	116.89	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	5711	0	5761	41	0
1	B	5687	0	5710	29	0
2	C	23	0	21	0	0
2	D	23	0	21	0	0
3	A	21	0	30	2	0
3	B	7	0	10	1	0
4	A	1	0	0	0	0
4	B	1	0	0	0	0
5	A	10	0	14	1	0
6	A	674	0	0	16	9
6	B	718	0	0	18	11
All	All	12876	0	11567	72	11

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 3.

The worst 5 of 72 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)
3:A:803:PEG:H11	6:A:991:HOH:O	1.36	1.21
1:B:516:HIS:HD2	6:B:1393:HOH:O	1.34	1.08
1:A:295:LYS:CE	6:A:947:HOH:O	2.02	1.06
1:A:295:LYS:HE2	6:A:947:HOH:O	1.59	0.99
1:A:130:ASP:OD2	6:A:901:HOH:O	1.90	0.88

The worst 5 of 11 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)
6:B:1551:HOH:O	6:A:1506:HOH:O[1_556]	1.26	0.94
6:B:1400:HOH:O	6:A:1135:HOH:O[1_545]	1.36	0.84
6:B:1529:HOH:O	6:A:1392:HOH:O[1_546]	1.42	0.78
6:B:1532:HOH:O	6:A:1495:HOH:O[1_645]	1.55	0.65
6:B:1477:HOH:O	6:A:1437:HOH:O[1_546]	1.62	0.58

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	749/733 (102%)	723 (96%)	26 (4%)	0	100	100
1	B	744/733 (102%)	723 (97%)	21 (3%)	0	100	100
All	All	1493/1466 (102%)	1446 (97%)	47 (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	604/586 (103%)	598 (99%)	6 (1%)	76	61
1	B	599/586 (102%)	592 (99%)	7 (1%)	71	54
All	All	1203/1172 (103%)	1190 (99%)	13 (1%)	73	57

5 of 13 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	157	MET
1	A	346	ARG
1	A	702	LYS
1	A	516	HIS
1	A	622	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	B	384	GLN
1	B	516	HIS
1	A	516	HIS

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)

4 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	BGC	C	1	2	12,12,12	0.78	0	17,17,17	1.27	3 (17%)
2	GAL	C	2	2	11,11,12	1.01	1 (9%)	15,15,17	1.31	2 (13%)
2	BGC	D	1	2	12,12,12	1.01	1 (8%)	17,17,17	1.26	2 (11%)
2	GAL	D	2	2	11,11,12	0.66	0	15,15,17	1.50	3 (20%)

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	BGC	C	1	2	-	0/2/22/22	0/1/1/1
2	GAL	C	2	2	-	0/2/19/22	0/1/1/1
2	BGC	D	1	2	-	0/2/22/22	0/1/1/1
2	GAL	D	2	2	-	0/2/19/22	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	1	BGC	C3-C2	2.24	1.58	1.52
2	C	2	GAL	O5-C5	2.02	1.47	1.43

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	D	2	GAL	C1-O5-C5	-3.29	107.73	112.19
2	D	1	BGC	C4-C3-C2	-3.19	105.25	110.82
2	D	1	BGC	C1-O5-C5	-2.90	108.20	113.66
2	C	1	BGC	C4-C3-C2	-2.87	105.82	110.82
2	C	2	GAL	C1-O5-C5	-2.86	108.31	112.19

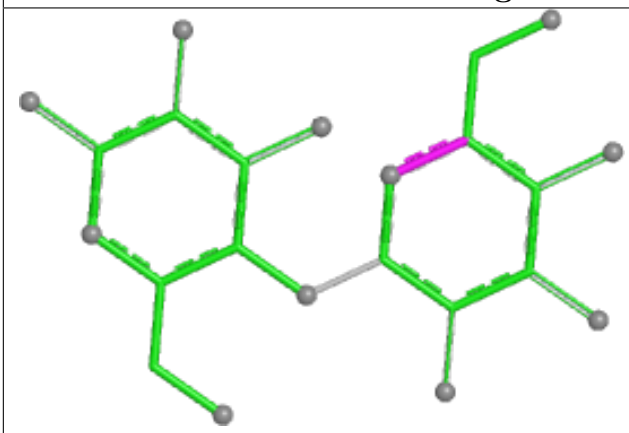
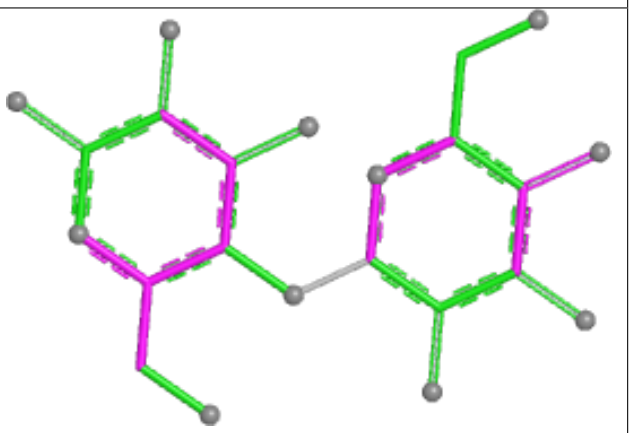
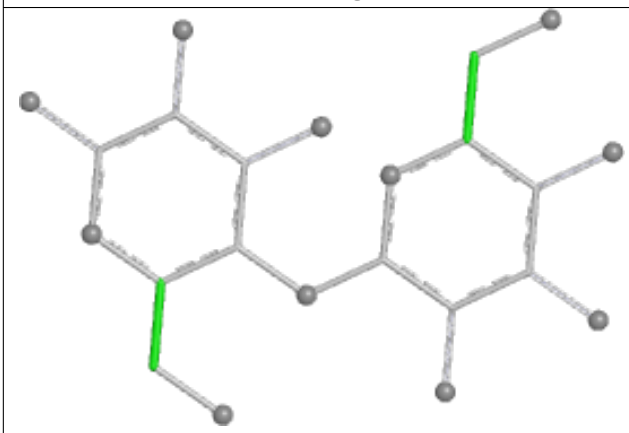
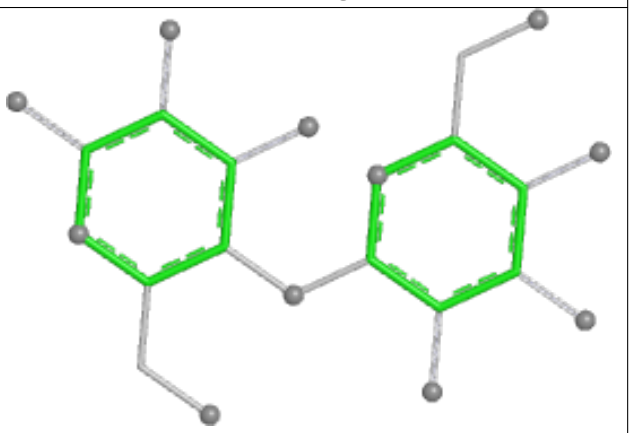
There are no chirality outliers.

There are no torsion outliers.

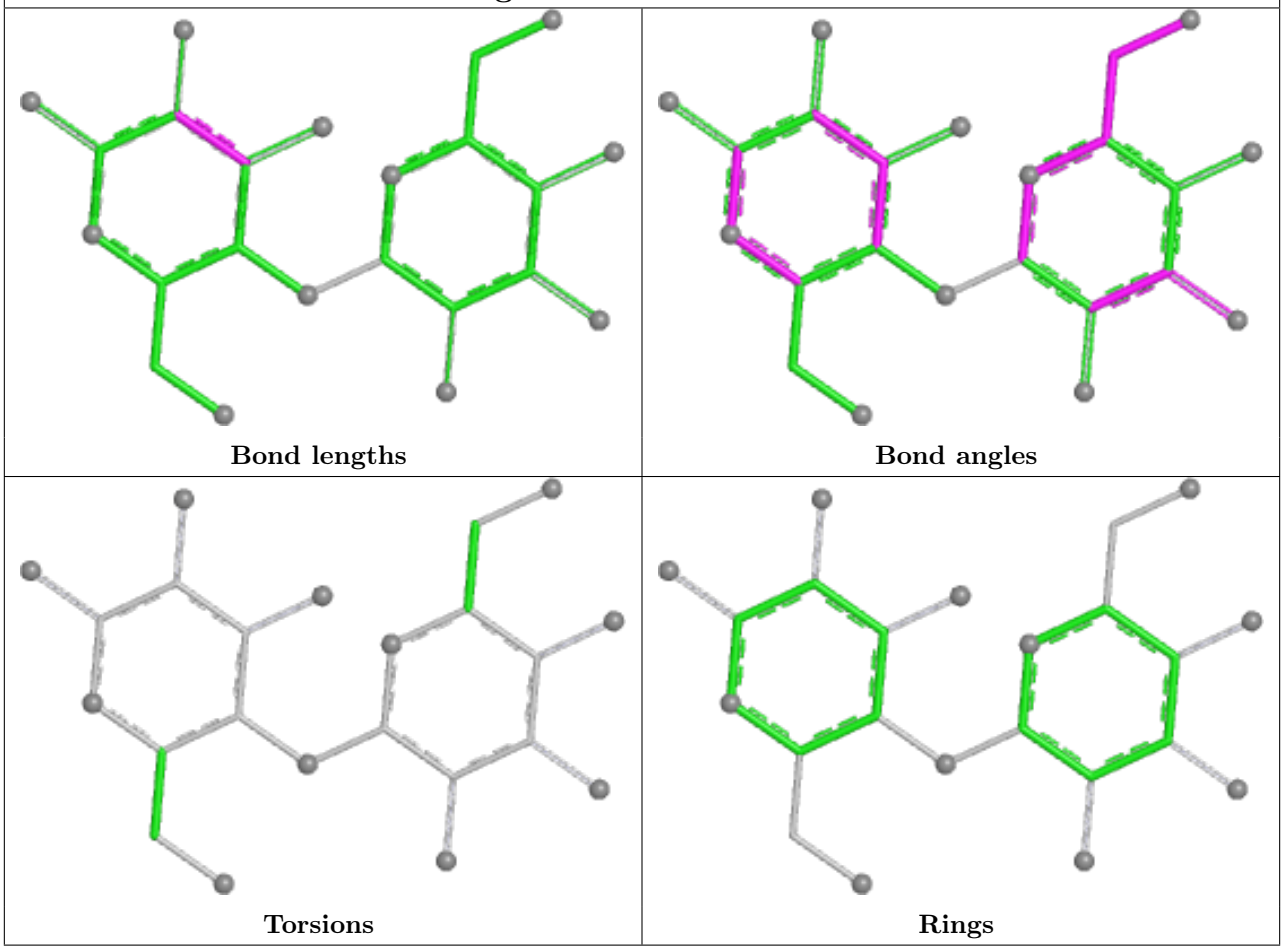
There are no ring outliers.

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.

Oligosaccharide Chain C	
 Bond lengths	 Bond angles
 Torsions	 Rings

Oligosaccharide Chain D



5.6 Ligand geometry (i)

Of 7 ligands modelled in this entry, 2 are monoatomic - leaving 5 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	$\#\|Z\| > 2$
3	PEG	A	803	-	6,6,6	0.38	0	5,5,5	0.49	0
3	PEG	A	802	-	6,6,6	0.30	0	5,5,5	0.30	0
5	PGE	A	805	-	9,9,9	0.40	0	8,8,8	0.34	0
3	PEG	A	804	-	6,6,6	0.19	0	5,5,5	0.28	0
3	PEG	B	802	-	6,6,6	0.61	0	5,5,5	1.00	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
3	PEG	A	803	-	-	2/4/4/4	-
3	PEG	A	802	-	-	1/4/4/4	-
5	PGE	A	805	-	-	5/7/7/7	-
3	PEG	A	804	-	-	2/4/4/4	-
3	PEG	B	802	-	-	2/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 12 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	803	PEG	C4-C3-O2-C2
3	B	802	PEG	O1-C1-C2-O2
5	A	805	PGE	C4-C3-O2-C2
5	A	805	PGE	O2-C3-C4-O3
3	A	804	PEG	C4-C3-O2-C2

There are no ring outliers.

3 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	803	PEG	2	0
5	A	805	PGE	1	0
3	B	802	PEG	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 (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	733/733 (100%)	-0.23	11 (1%)	73	73	11, 19, 41, 82	0
1	B	733/733 (100%)	-0.39	3 (0%)	92	92	11, 18, 34, 59	0
All	All	1466/1466 (100%)	-0.31	14 (0%)	82	82	11, 19, 37, 82	0

The worst 5 of 14 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	621	ASP	6.1
1	A	368	PRO	5.6
1	A	620	HIS	3.9
1	A	369	ASP	3.9
1	A	622	ASN	3.6

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)

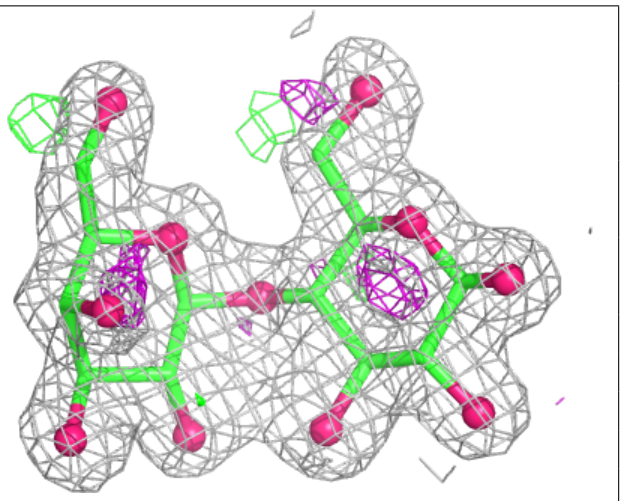
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists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

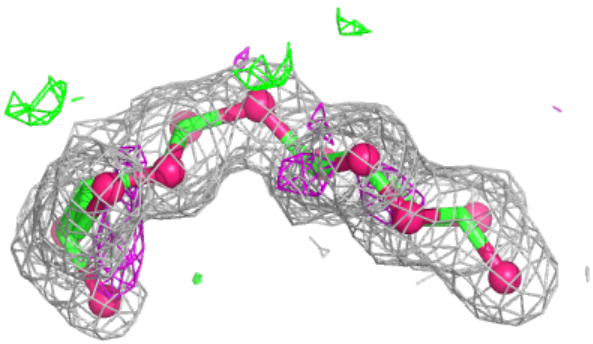
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	GAL	C	2	11/12	0.91	0.13	27,31,34,34	0
2	BGC	D	1	12/12	0.94	0.10	21,23,24,26	0
2	BGC	C	1	12/12	0.95	0.09	21,24,27,27	0
2	GAL	D	2	11/12	0.96	0.09	24,26,30,31	0

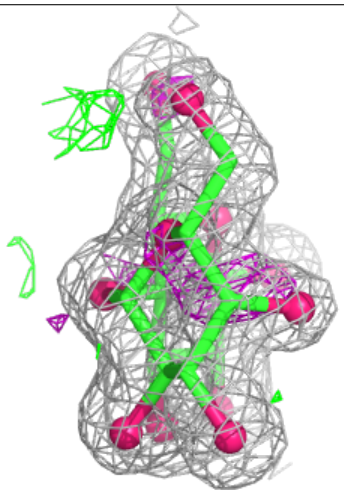
The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around Chain C:

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

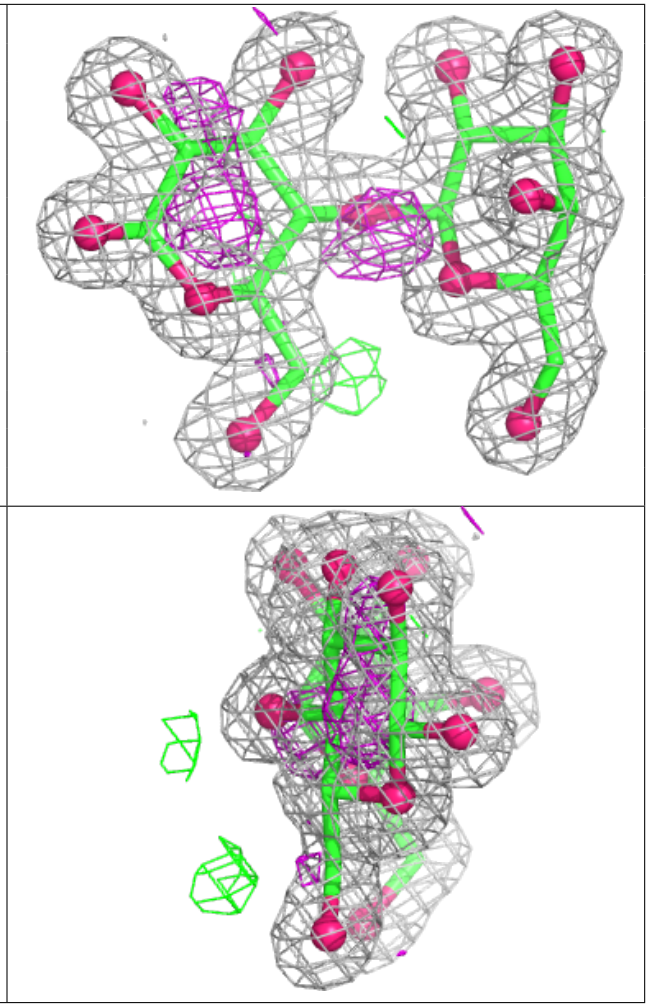






Electron density around Chain D:

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



6.4 Ligands (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
3	PEG	B	802	7/7	0.80	0.17	34,38,44,46	0
3	PEG	A	802	7/7	0.83	0.14	37,40,41,48	0
3	PEG	A	803	7/7	0.84	0.16	40,41,48,53	0
3	PEG	A	804	7/7	0.90	0.11	41,45,48,49	0
5	PGE	A	805	10/10	0.91	0.17	41,47,50,54	0
4	MG	A	806	1/1	0.99	0.05	18,18,18,18	0
4	MG	B	803	1/1	0.99	0.07	17,17,17,17	0

6.5 Other polymers (i)

There are no such residues in this entry.