wwPDB X-ray Structure Validation Summary Report (i)

Sep 12, 2023 – 09:30 AM EDT

|		
PDB ID	:	4N2Z
Title	:	Crystal Structure of the alpha-L-arabinofuranosidase PaAbf62A from Podospora anserina in complex with cellotriose
Authors	:	Siguier, B.; Dumon, C.; Mourey, L.; Tranier, S.
Deposited on	:	2013-10-06
Resolution	:	1.80 Å(reported)

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.
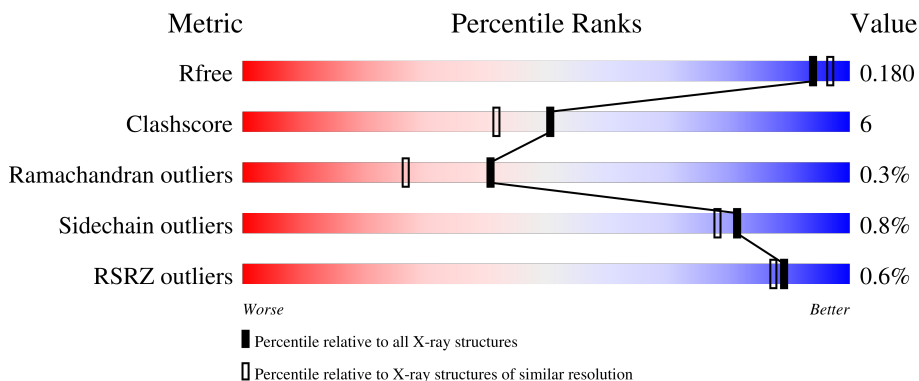
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	5950 (1.80-1.80)
Clashscore	141614	6793 (1.80-1.80)
Ramachandran outliers	138981	6697 (1.80-1.80)
Sidechain outliers	138945	6696 (1.80-1.80)
RSRZ outliers	127900	5850 (1.80-1.80)

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 <=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	360	79% 8% · 12%
2	B	3	33% 67%

2   Entry composition ⓘ

There are 7 unique types of molecules in this entry. The entry contains 3008 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 GH62 arabinofuranosidase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	318	2567	1645	429	481	12	0	11	0

There are 23 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	338	GLY	-	expression tag	UNP E2GHW5
A	339	LEU	-	expression tag	UNP E2GHW5
A	340	GLU	-	expression tag	UNP E2GHW5
A	341	GLN	-	expression tag	UNP E2GHW5
A	342	LYS	-	expression tag	UNP E2GHW5
A	343	LEU	-	expression tag	UNP E2GHW5
A	344	ILE	-	expression tag	UNP E2GHW5
A	345	SER	-	expression tag	UNP E2GHW5
A	346	GLU	-	expression tag	UNP E2GHW5
A	347	GLU	-	expression tag	UNP E2GHW5
A	348	ASP	-	expression tag	UNP E2GHW5
A	349	LEU	-	expression tag	UNP E2GHW5
A	350	ASN	-	expression tag	UNP E2GHW5
A	351	SER	-	expression tag	UNP E2GHW5
A	352	ALA	-	expression tag	UNP E2GHW5
A	353	VAL	-	expression tag	UNP E2GHW5
A	354	ASP	-	expression tag	UNP E2GHW5
A	355	HIS	-	expression tag	UNP E2GHW5
A	356	HIS	-	expression tag	UNP E2GHW5
A	357	HIS	-	expression tag	UNP E2GHW5
A	358	HIS	-	expression tag	UNP E2GHW5
A	359	HIS	-	expression tag	UNP E2GHW5
A	360	HIS	-	expression tag	UNP E2GHW5

- Molecule 2 is an oligosaccharide called beta-D-glucopyranose-(1-4)-beta-D-glucopyranose-(1-4)-alpha-D-glucopyranose.
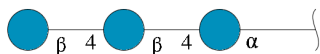
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace
2	B	3	Total 34	C 18	O 16	0	0	0

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

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

- Molecule 4 is 2-AMINO-2-HYDROXYMETHYL-PROPANE-1,3-DIOL (three-letter code: TRS) (formula: $C_4H_{12}NO_3$).
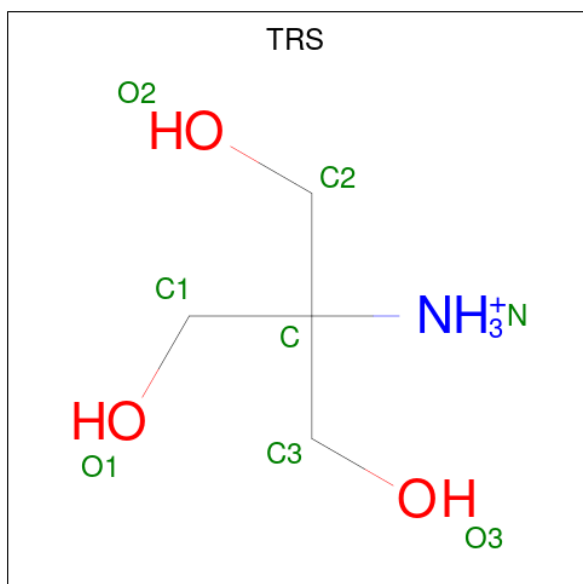
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total 8	C 4	N 1	O 3	0	0

- Molecule 5 is PENTAETHYLENE GLYCOL (three-letter code: 1PE) (formula: $C_{10}H_{22}O_6$).

1PE

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
5	A	1	13	8	5	0	0

- Molecule 6 is 4-(2-HYDROXYETHYL)-1-PIPERAZINE ETHANESULFONIC ACID (three-letter code: EPE) (formula: $C_8H_{18}N_2O_4S$).

EPE
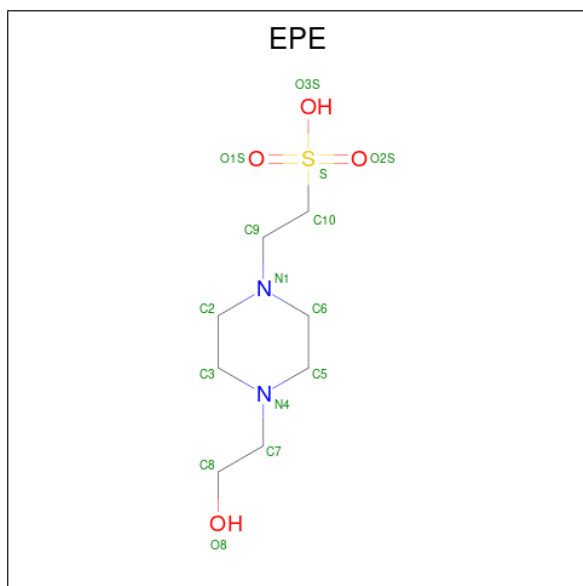
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	S		
6	A	1	15	8	2	4	1	0	0

- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	370	Total 370	O 370	0	0

3 Residue-property plots ⓘ

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: GH62 arabinofuranosidase
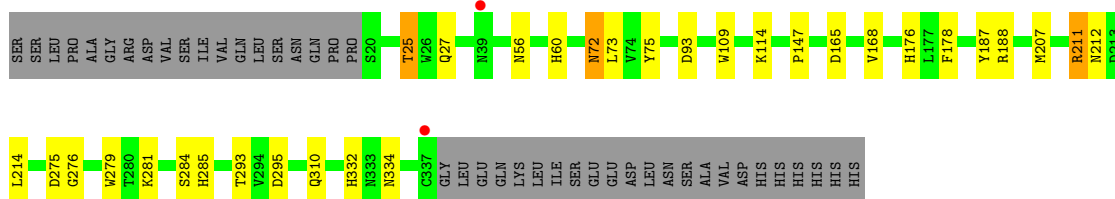
• Molecule 2: beta-D-glucopyranose-(1-4)-beta-D-glucopyranose-(1-4)-alpha-D-glucopyranose

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	102.14Å 66.79Å 60.36Å 90.00° 117.33° 90.00°	Depositor
Resolution (Å)	14.84 – 1.80 14.84 – 1.80	Depositor EDS
% Data completeness (in resolution range)	99.6 (14.84-1.80) 99.8 (14.84-1.80)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.07	Depositor
$< I/\sigma(I) >$ [1]	3.31 (at 1.79Å)	Xtriage
Refinement program	REFMAC 5.7.0032	Depositor
R, R_{free}	0.131 , 0.170 0.146 , 0.180	Depositor DCC
R_{free} test set	1687 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å^2)	11.1	Xtriage
Anisotropy	0.026	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.40 , 53.1	EDS
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.013 for -1/2*h+1/2*k+l,1/2*h-1/2*k+l,1/2*h+1/2*k 0.015 for -1/2*h-1/2*k+l,-1/2*h-1/2*k-l,1/2*h-1/2*k	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	3008	wwPDB-VP
Average B, all atoms (Å^2)	13.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.59% 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 ⓘ

5.1   Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: CA, GLC, BGC, TRS, 1PE, EPE

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	1.09	2/2685 (0.1%)	1.00	3/3672 (0.1%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	165	ASP	C-O	-5.29	1.13	1.23
1	A	279	TRP	CE3-CZ3	5.27	1.47	1.38

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	211	ARG	NE-CZ-NH1	5.76	123.18	120.30
1	A	295	ASP	CB-CG-OD1	5.58	123.33	118.30
1	A	93	ASP	CB-CG-OD2	-5.44	113.41	118.30

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	2567	0	2417	27	0
2	B	34	0	30	2	0
3	A	1	0	0	0	0

Continued on next page...

W O R L D W I D E
wwPDB
PROTEIN DATA BANK

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	A	8	0	12	0	0
5	A	13	0	17	1	0
6	A	15	0	18	0	0
7	A	370	0	0	7	0
All	All	3008	0	2494	29	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 6.

The worst 5 of 29 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:72:ASN:HD22	1:A:73:LEU:H	1.28	0.82
1:A:176:HIS:HD2	1:A:188:ARG:HH11	1.31	0.78
1:A:60:HIS:HD2	1:A:75:TYR:OH	1.77	0.68
1:A:281[B]:LYS:NZ	7:A:1195:HOH:O	2.30	0.65
1:A:56[B]:ASN:ND2	7:A:1217:HOH:O	2.20	0.64

There are no symmetry-related clashes.

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	327/360 (91%)	314 (96%)	12 (4%)	1 (0%)	41	27

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	285	HIS

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	273/301 (91%)	271 (99%)	2 (1%)	84	81

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	25	THR
1	A	72	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	212	ASN
1	A	240	HIS
1	A	332	HIS
1	A	94	GLN
1	A	156	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 ⓘ

3 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	GLC	B	1	2	12,12,12	1.15	1 (8%)	17,17,17	2.57	9 (52%)
2	BGC	B	2	2	11,11,12	1.27	1 (9%)	15,15,17	2.59	9 (60%)
2	BGC	B	3	2	11,11,12	0.83	0	15,15,17	1.88	5 (33%)

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	B	1	2	-	1/2/22/22	0/1/1/1
2	BGC	B	2	2	-	2/2/19/22	0/1/1/1
2	BGC	B	3	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	B	1	GLC	C1-C2	2.65	1.58	1.52
2	B	2	BGC	O4-C4	-2.49	1.37	1.43

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	1	GLC	O5-C5-C4	5.51	119.69	109.69
2	B	2	BGC	C1-O5-C5	4.27	117.98	112.19
2	B	2	BGC	O4-C4-C5	-3.95	99.48	109.30
2	B	1	GLC	O5-C1-C2	3.92	117.29	110.28
2	B	2	BGC	C6-C5-C4	3.77	121.84	113.00

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	2	BGC	O5-C5-C6-O6
2	B	2	BGC	C4-C5-C6-O6
2	B	1	GLC	C4-C5-C6-O6

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	2	BGC	2	0
2	B	3	BGC	2	0

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 B

Bond lengths

Bond angles

Torsions

Rings
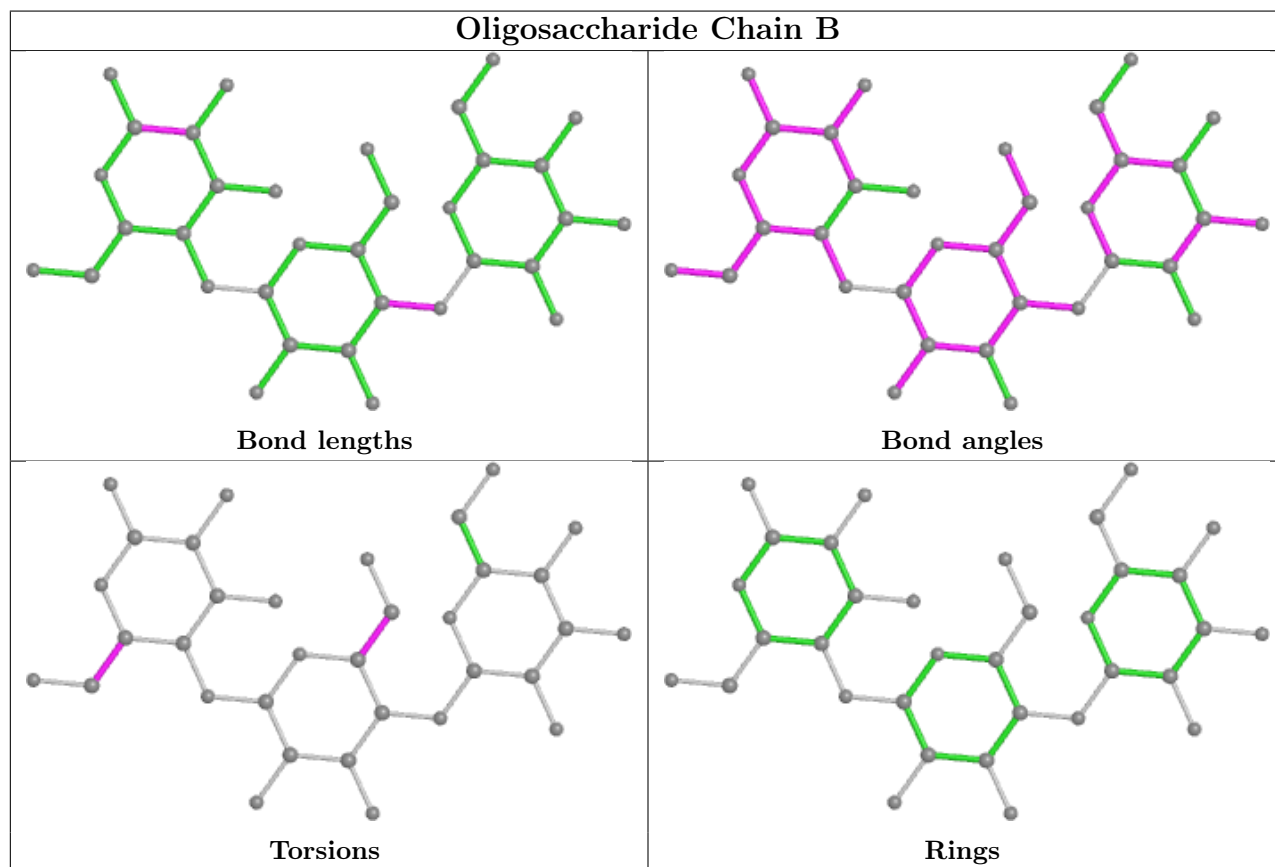
5.6 Ligand geometry ⓘ

Of 4 ligands modelled in this entry, 1 is monoatomic - leaving 3 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
6	EPE	A	404	-	15,15,15	1.89	2 (13%)	18,20,20	3.45	9 (50%)
4	TRS	A	402	-	7,7,7	0.83	0	9,9,9	0.71	0
5	1PE	A	403	-	12,12,15	0.55	0	11,11,14	0.76	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
6	EPE	A	404	-	-	2/9/19/19	0/1/1/1
4	TRS	A	402	-	-	0/9/9/9	-
5	1PE	A	403	-	-	4/10/10/13	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	A	404	EPE	C10-S	-5.75	1.69	1.77
6	A	404	EPE	O2S-S	3.56	1.55	1.45

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
6	A	404	EPE	C5-N4-C3	7.71	126.18	108.83
6	A	404	EPE	C6-N1-C2	7.66	126.07	108.83
6	A	404	EPE	O1S-S-C10	4.10	111.85	106.92
6	A	404	EPE	C9-N1-C2	4.06	121.61	111.23
6	A	404	EPE	C3-C2-N1	3.82	118.47	110.64

There are no chirality outliers.

5 of 6 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	A	404	EPE	C10-C9-N1-C2
6	A	404	EPE	N4-C7-C8-O8
5	A	403	1PE	OH2-C12-C22-OH3
5	A	403	1PE	C15-C25-OH5-C14
5	A	403	1PE	C13-C23-OH3-C22

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	403	1PE	1	0

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

6 Fit of model and data ⓘ

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, 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	318/360 (88%)	-0.63	2 (0%)	89	87	6, 10, 24, 33	0

All (2) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	337	CYS	3.0
1	A	39	ASN	2.7

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

There are no non-standard protein/DNA/RNA residues in this entry.
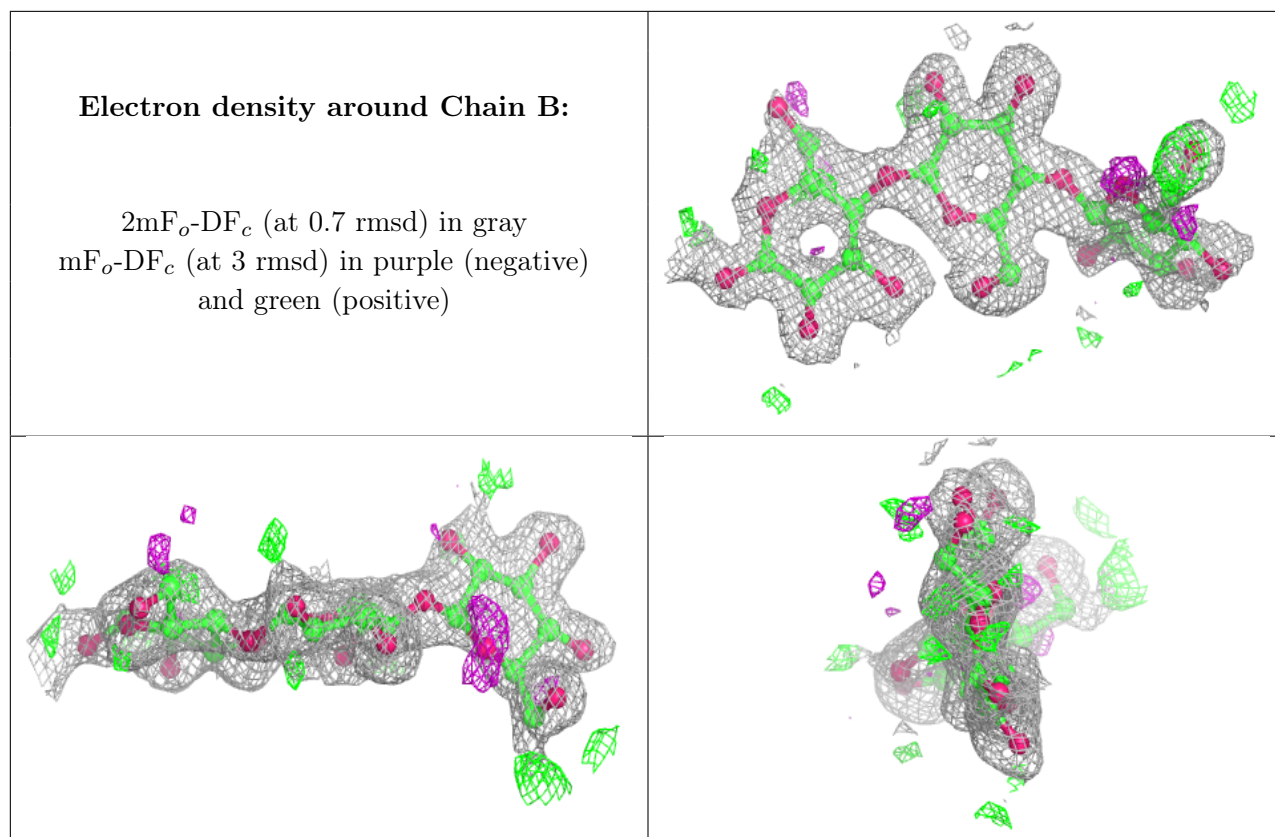
6.3 Carbohydrates ⓘ

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	GLC	B	1	12/12	0.79	0.21	35,42,47,47	0
2	BGC	B	3	11/12	0.86	0.30	24,38,40,42	0
2	BGC	B	2	11/12	0.93	0.11	11,25,31,35	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 B:

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

6.4 Ligands ⓘ

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
5	1PE	A	403	13/16	0.89	0.17	25,28,35,37	13
6	EPE	A	404	15/15	0.92	0.17	25,29,45,46	0
4	TRS	A	402	8/8	0.97	0.09	8,9,9,10	0
3	CA	A	401	1/1	1.00	0.02	11,11,11,11	0

6.5 Other polymers ⓘ

There are no such residues in this entry.