wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 10, 2023 – 12:13 PM EDT

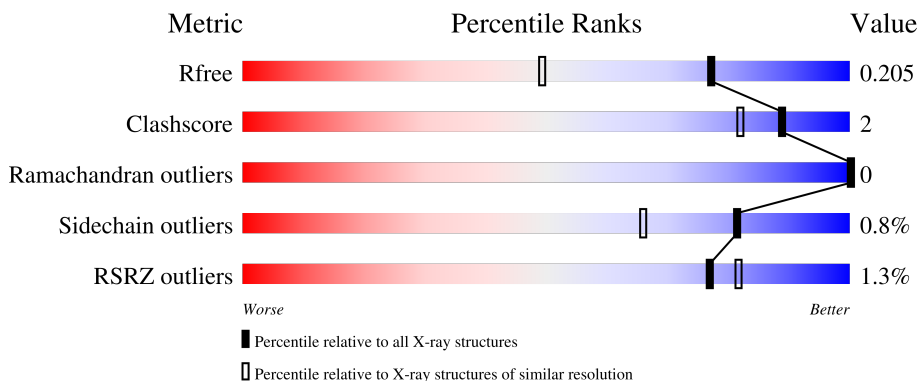|		
PDB ID	:	7KR6
Title	:	Glycoside hydrolase family 16 endo-glucanase from Bacteroides ovatus in complex with G4G3G-2F-DNP
Authors	:	Tamura, K.; Brumer, H.; van Petegem, F.
Deposited on	:	2020-11-18
Resolution	:	1.56 Å(reported)

1 Overall quality at a glance ⓘ

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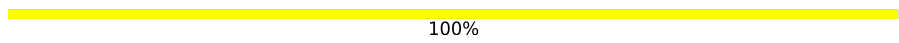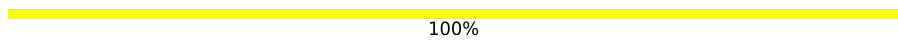*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.56 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	Percentile Ranks	Value
Rfree		0.205
Clashscore		2
Ramachandran outliers		0
Sidechain outliers		0.8%
RSRZ outliers		1.3%

Worse   *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1483 (1.56-1.56)
Clashscore	141614	1529 (1.56-1.56)
Ramachandran outliers	138981	1498 (1.56-1.56)
Sidechain outliers	138945	1495 (1.56-1.56)
RSRZ outliers	127900	1465 (1.56-1.56)

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	AAA	272	% 82% 5% 13%
1	BBB	272	% 81% 5% 13%
2	AaA	3	100%
2	BaB	3	100%

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 7829 atoms, of which 3628 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 Glycoside hydrolase family 16 protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	AAA	236	Total	C	H	N	O	S	113	1	0
			3656	1197	1776	312	364	7			
1	BBB	236	Total	C	H	N	O	S	113	1	0
			3673	1200	1792	314	360	7			

There are 44 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
AAA	0	MET	-	initiating methionine	UNP A0A1Y4PXW9
AAA	1	GLY	-	expression tag	UNP A0A1Y4PXW9
AAA	2	SER	-	expression tag	UNP A0A1Y4PXW9
AAA	3	SER	-	expression tag	UNP A0A1Y4PXW9
AAA	4	HIS	-	expression tag	UNP A0A1Y4PXW9
AAA	5	HIS	-	expression tag	UNP A0A1Y4PXW9
AAA	6	HIS	-	expression tag	UNP A0A1Y4PXW9
AAA	7	HIS	-	expression tag	UNP A0A1Y4PXW9
AAA	8	HIS	-	expression tag	UNP A0A1Y4PXW9
AAA	9	HIS	-	expression tag	UNP A0A1Y4PXW9
AAA	10	SER	-	expression tag	UNP A0A1Y4PXW9
AAA	11	SER	-	expression tag	UNP A0A1Y4PXW9
AAA	12	GLY	-	expression tag	UNP A0A1Y4PXW9
AAA	13	LEU	-	expression tag	UNP A0A1Y4PXW9
AAA	14	VAL	-	expression tag	UNP A0A1Y4PXW9
AAA	15	PRO	-	expression tag	UNP A0A1Y4PXW9
AAA	16	ARG	-	expression tag	UNP A0A1Y4PXW9
AAA	17	GLY	-	expression tag	UNP A0A1Y4PXW9
AAA	18	SER	-	expression tag	UNP A0A1Y4PXW9
AAA	19	HIS	-	expression tag	UNP A0A1Y4PXW9
AAA	20	MET	-	expression tag	UNP A0A1Y4PXW9
AAA	148	ALA	GLU	engineered mutation	UNP A0A1Y4PXW9
BBB	0	MET	-	initiating methionine	UNP A0A1Y4PXW9
BBB	1	GLY	-	expression tag	UNP A0A1Y4PXW9
BBB	2	SER	-	expression tag	UNP A0A1Y4PXW9

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
BBB	3	SER	-	expression tag	UNP A0A1Y4PXW9
BBB	4	HIS	-	expression tag	UNP A0A1Y4PXW9
BBB	5	HIS	-	expression tag	UNP A0A1Y4PXW9
BBB	6	HIS	-	expression tag	UNP A0A1Y4PXW9
BBB	7	HIS	-	expression tag	UNP A0A1Y4PXW9
BBB	8	HIS	-	expression tag	UNP A0A1Y4PXW9
BBB	9	HIS	-	expression tag	UNP A0A1Y4PXW9
BBB	10	SER	-	expression tag	UNP A0A1Y4PXW9
BBB	11	SER	-	expression tag	UNP A0A1Y4PXW9
BBB	12	GLY	-	expression tag	UNP A0A1Y4PXW9
BBB	13	LEU	-	expression tag	UNP A0A1Y4PXW9
BBB	14	VAL	-	expression tag	UNP A0A1Y4PXW9
BBB	15	PRO	-	expression tag	UNP A0A1Y4PXW9
BBB	16	ARG	-	expression tag	UNP A0A1Y4PXW9
BBB	17	GLY	-	expression tag	UNP A0A1Y4PXW9
BBB	18	SER	-	expression tag	UNP A0A1Y4PXW9
BBB	19	HIS	-	expression tag	UNP A0A1Y4PXW9
BBB	20	MET	-	expression tag	UNP A0A1Y4PXW9
BBB	148	ALA	GLU	engineered mutation	UNP A0A1Y4PXW9

- Molecule 2 is an oligosaccharide called beta-D-glucopyranose-(1-4)-beta-D-glucopyranose-(1-3)-2-deoxy-2-fluoro-alpha-D-glucopyranose.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	AaA	3	Total	C	F	H	O	9	0	0
			63	18	1	30	14			
2	BaB	3	Total	C	F	H	O	8	0	0
			63	18	1	30	14			

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	AAA	200	Total	O	0	0
			200	200		
3	BBB	174	Total	O	0	0
			174	174		

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.
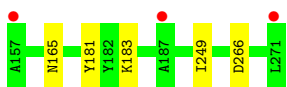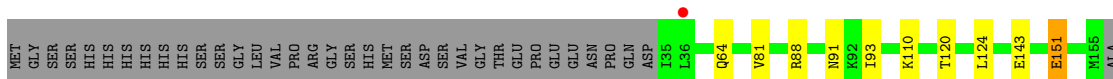
- Molecule 1: Glycoside hydrolase family 16 protein

Chain AAA:
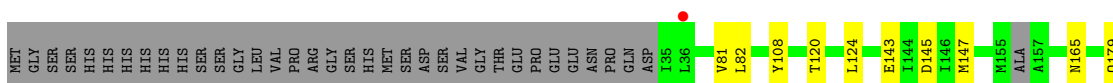
- Molecule 1: Glycoside hydrolase family 16 protein

Chain BBB:

- Molecule 2: beta-D-glucopyranose-(1-4)-beta-D-glucopyranose-(1-3)-2-deoxy-2-fluoro-alpha-D-glucopyranose

Chain AaA:

- Molecule 2: beta-D-glucopyranose-(1-4)-beta-D-glucopyranose-(1-3)-2-deoxy-2-fluoro-alpha-D-glucopyranose

Chain BaB:

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	46.45Å 84.58Å 67.11Å 90.00° 93.98° 90.00°	Depositor
Resolution (Å)	39.40 – 1.56 39.40 – 1.56	Depositor EDS
% Data completeness (in resolution range)	98.6 (39.40-1.56) 98.6 (39.40-1.56)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	2.59 (at 1.56Å)	Xtriage
Refinement program	REFMAC 5.8.0258	Depositor
R, R_{free}	0.172 , 0.199 0.181 , 0.205	Depositor DCC
R_{free} test set	3598 reflections (4.96%)	wwPDB-VP
Wilson B-factor (Å^2)	21.0	Xtriage
Anisotropy	0.383	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.39 , 54.4	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	7829	wwPDB-VP
Average B, all atoms (Å^2)	27.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.33% 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: BGC, G2F

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	AAA	0.74	0/1928	0.92	2/2609 (0.1%)
1	BBB	0.72	0/1929	0.88	0/2608
All	All	0.73	0/3857	0.90	2/5217 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	BBB	0	1

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	AAA	108	TYR	CB-CG-CD1	6.15	124.69	121.00
1	AAA	145	ASP	CB-CG-OD2	-5.26	113.57	118.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	BBB	64	GLN	Peptide

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	AAA	1880	1776	1757	6	0
1	BBB	1881	1792	1773	9	0
2	AaA	33	30	27	0	0
2	BaB	33	30	27	0	0
3	AAA	200	0	0	1	0
3	BBB	174	0	0	2	0
All	All	4201	3628	3584	14	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 2.

The worst 5 of 14 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:BBB:165:ASN:HB3	3:BBB:435:HOH:O	1.97	0.65
1:AAA:165:ASN:HB3	3:AAA:451:HOH:O	2.08	0.54
1:BBB:93:ILE:HD11	1:BBB:249:ILE:CD1	2.39	0.53
1:BBB:120:THR:HA	1:BBB:124:LEU:HB2	1.95	0.48
1:AAA:82:LEU:HD23	1:AAA:265:ILE:HD13	1.97	0.47

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	AAA	233/272 (86%)	230 (99%)	3 (1%)	0	100	100
1	BBB	233/272 (86%)	229 (98%)	4 (2%)	0	100	100
All	All	466/544 (86%)	459 (98%)	7 (2%)	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	AAA	195/230 (85%)	194 (100%)	1 (0%)	88	78
1	BBB	195/230 (85%)	193 (99%)	2 (1%)	76	57
All	All	390/460 (85%)	387 (99%)	3 (1%)	81	66

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	AAA	143	GLU
1	BBB	143	GLU
1	BBB	151	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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	#$|Z| > 2$
2	G2F	AaA	1	2	11,11,12	1.12	0	10,15,17	1.09	1 (10%)
2	BGC	AaA	2	2	11,11,12	1.09	1 (9%)	15,15,17	1.61	3 (20%)
2	BGC	AaA	3	2	11,11,12	0.68	0	15,15,17	1.34	3 (20%)
2	G2F	BaB	1	2	11,11,12	0.73	0	10,15,17	1.75	4 (40%)
2	BGC	BaB	2	2	11,11,12	1.58	3 (27%)	15,15,17	1.74	3 (20%)
2	BGC	BaB	3	2	11,11,12	0.68	0	15,15,17	1.17	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
2	G2F	AaA	1	2	-	0/2/19/22	0/1/1/1
2	BGC	AaA	2	2	-	0/2/19/22	0/1/1/1
2	BGC	AaA	3	2	-	0/2/19/22	0/1/1/1
2	G2F	BaB	1	2	-	0/2/19/22	0/1/1/1
2	BGC	BaB	2	2	-	0/2/19/22	0/1/1/1
2	BGC	BaB	3	2	-	1/2/19/22	0/1/1/1

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	BaB	2	BGC	C2-C3	3.63	1.57	1.52
2	BaB	2	BGC	O5-C5	2.27	1.48	1.43
2	BaB	2	BGC	O4-C4	2.11	1.48	1.43
2	AaA	2	BGC	C2-C3	2.06	1.55	1.52

The worst 5 of 15 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	BaB	2	BGC	C1-C2-C3	-3.93	104.84	109.67
2	AaA	2	BGC	C1-C2-C3	-3.25	105.67	109.67
2	BaB	2	BGC	C6-C5-C4	2.99	120.00	113.00
2	BaB	1	G2F	C1-O5-C5	2.76	115.94	112.19
2	BaB	1	G2F	O5-C5-C6	-2.70	102.97	107.20

There are no chirality outliers.

All (1) torsion outliers are listed below:

| Mol	Chain	Res	Type	Atoms
2	BaB	3	BGC	C4-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.

Oligosaccharide Chain AaA	
Bond lengths	Bond angles
Torsions	Rings

Oligosaccharide Chain BaB

Bond lengths

Bond angles

Torsions

Rings
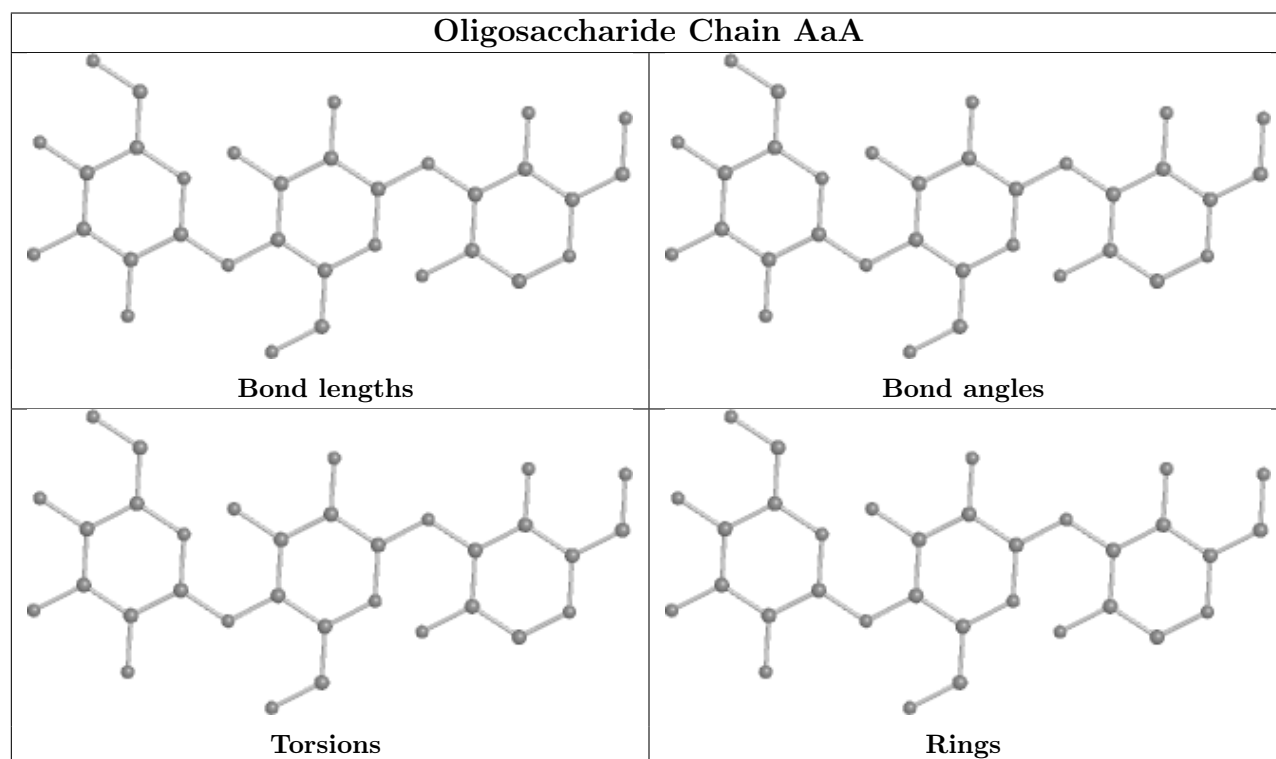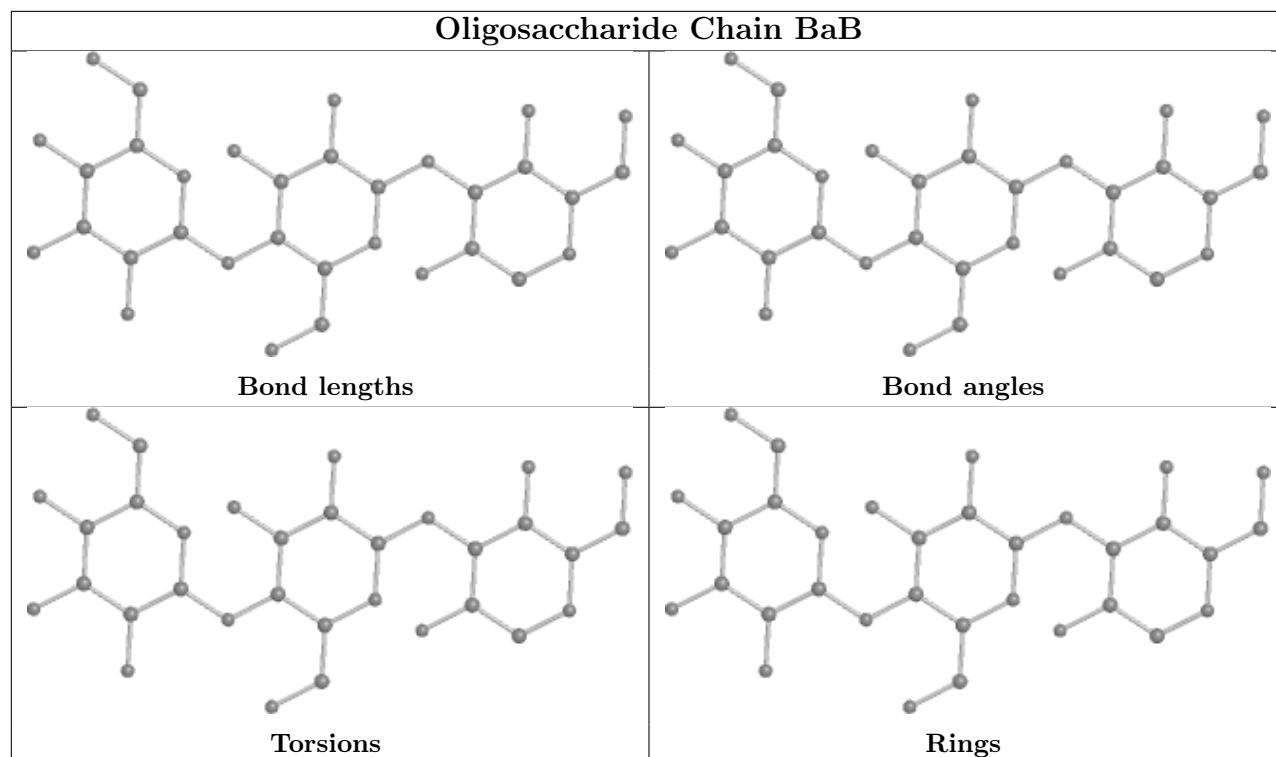
5.6 Ligand geometry (i)

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 ⓘ

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	AAA	236/272 (86%)	-0.27	2 (0%)	86	89	16, 24, 36, 50	0
1	BBB	236/272 (86%)	-0.05	4 (1%)	70	75	17, 27, 43, 73	0
All	All	472/544 (86%)	-0.16	6 (1%)	77	82	16, 25, 41, 73	0

The worst 5 of 6 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	BBB	157	ALA	3.9
1	AAA	36	LEU	3.8
1	BBB	271	LEU	3.4
1	AAA	181[A]	TYR	3.0
1	BBB	187	ALA	2.6

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.
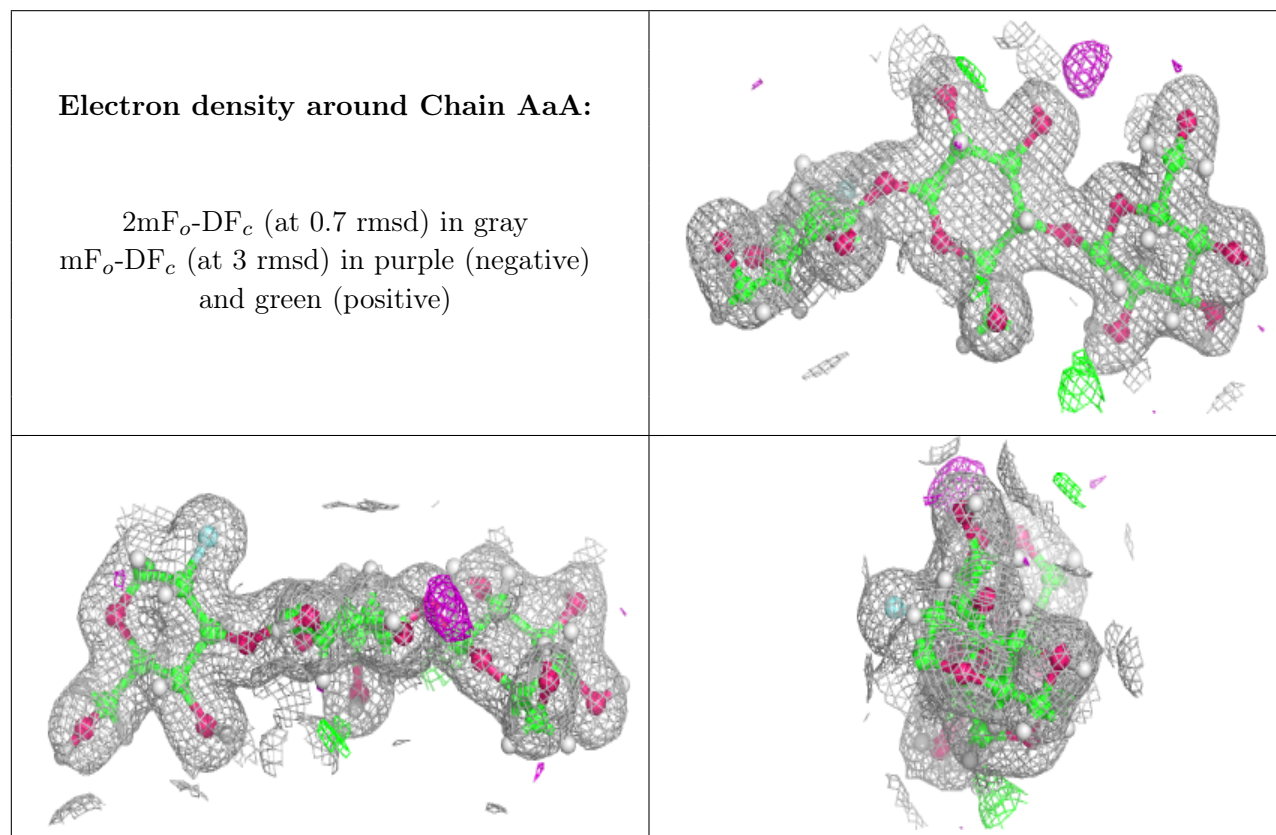
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	BGC	AaA	3	11/12	0.95	0.06	0,31,33,36	4
2	BGC	BaB	2	11/12	0.96	0.06	0,21,25,25	3
2	BGC	BaB	3	11/12	0.96	0.08	0,35,39,41	3
2	G2F	BaB	1	11/12	0.98	0.07	18,19,21,23	2
2	BGC	AaA	2	11/12	0.98	0.08	0,19,21,22	3

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors($\mathring{A}^2$)	Q<0.9
2	G2F	AaA	1	11/12	0.98	0.10	16,18,19,20	2

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 AaA:

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

Electron density around Chain BaB:

$2mF_o$-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.