



wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 16, 2021 – 12:12 PM EDT

PDB ID : 6VHO
Title : Glycoside hydrolase family 16 endo-glucanase from Bacteroides ovatus in complex with G4G4G3G-NHCOCH₂Br
Authors : Tamura, K.; Brumer, H.; van Petegem, F.
Deposited on : 2020-01-10
Resolution : 2.15 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.23.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	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)	:	2.23.2

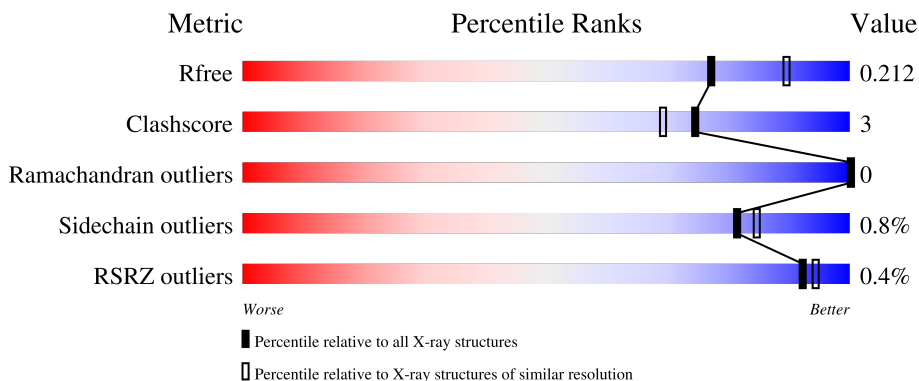
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.15 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	2523 (2.16-2.12)
Clashscore	141614	2653 (2.16-2.12)
Ramachandran outliers	138981	2618 (2.16-2.12)
Sidechain outliers	138945	2617 (2.16-2.12)
RSRZ outliers	127900	2485 (2.16-2.12)

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	AAA	272	<div> <div></div> <div>81%</div> <div>7%</div> <div>13%</div> </div>
1	BBB	272	<div> <div></div> <div>78%</div> <div>9%</div> <div>13%</div> </div>
2	A	3	<div> <div></div> <div>33%</div> <div>67%</div> </div>

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 7688 atoms, of which 3658 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 Glycosyl hydrolase family 16.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	AAA	237	Total	C	H	N	O	S	111	1	0
			3705	1210	1808	315	365	7			
1	BBB	237	Total	C	H	N	O	S	111	1	0
			3712	1211	1815	316	363	7			

There are 42 discrepancies between the modelled and reference sequences:

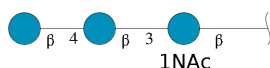
Chain	Residue	Modelled	Actual	Comment	Reference
AAA	0	MET	-	initiating methionine	UNP A7LY25
AAA	1	GLY	-	expression tag	UNP A7LY25
AAA	2	SER	-	expression tag	UNP A7LY25
AAA	3	SER	-	expression tag	UNP A7LY25
AAA	4	HIS	-	expression tag	UNP A7LY25
AAA	5	HIS	-	expression tag	UNP A7LY25
AAA	6	HIS	-	expression tag	UNP A7LY25
AAA	7	HIS	-	expression tag	UNP A7LY25
AAA	8	HIS	-	expression tag	UNP A7LY25
AAA	9	HIS	-	expression tag	UNP A7LY25
AAA	10	SER	-	expression tag	UNP A7LY25
AAA	11	SER	-	expression tag	UNP A7LY25
AAA	12	GLY	-	expression tag	UNP A7LY25
AAA	13	LEU	-	expression tag	UNP A7LY25
AAA	14	VAL	-	expression tag	UNP A7LY25
AAA	15	PRO	-	expression tag	UNP A7LY25
AAA	16	ARG	-	expression tag	UNP A7LY25
AAA	17	GLY	-	expression tag	UNP A7LY25
AAA	18	SER	-	expression tag	UNP A7LY25
AAA	19	HIS	-	expression tag	UNP A7LY25
AAA	20	MET	-	expression tag	UNP A7LY25
BBB	0	MET	-	initiating methionine	UNP A7LY25
BBB	1	GLY	-	expression tag	UNP A7LY25
BBB	2	SER	-	expression tag	UNP A7LY25
BBB	3	SER	-	expression tag	UNP A7LY25

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
BBB	4	HIS	-	expression tag	UNP A7LY25
BBB	5	HIS	-	expression tag	UNP A7LY25
BBB	6	HIS	-	expression tag	UNP A7LY25
BBB	7	HIS	-	expression tag	UNP A7LY25
BBB	8	HIS	-	expression tag	UNP A7LY25
BBB	9	HIS	-	expression tag	UNP A7LY25
BBB	10	SER	-	expression tag	UNP A7LY25
BBB	11	SER	-	expression tag	UNP A7LY25
BBB	12	GLY	-	expression tag	UNP A7LY25
BBB	13	LEU	-	expression tag	UNP A7LY25
BBB	14	VAL	-	expression tag	UNP A7LY25
BBB	15	PRO	-	expression tag	UNP A7LY25
BBB	16	ARG	-	expression tag	UNP A7LY25
BBB	17	GLY	-	expression tag	UNP A7LY25
BBB	18	SER	-	expression tag	UNP A7LY25
BBB	19	HIS	-	expression tag	UNP A7LY25
BBB	20	MET	-	expression tag	UNP A7LY25

- Molecule 2 is an oligosaccharide called beta-D-glucopyranose-(1-4)-beta-D-glucopyranose-(1-3)-N-acetyl-beta-D-glucopyranosylamine.



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	A	3	Total	C	H	N	O	7	0	0
			72	20	35	1	16			


- Molecule 3 is water.

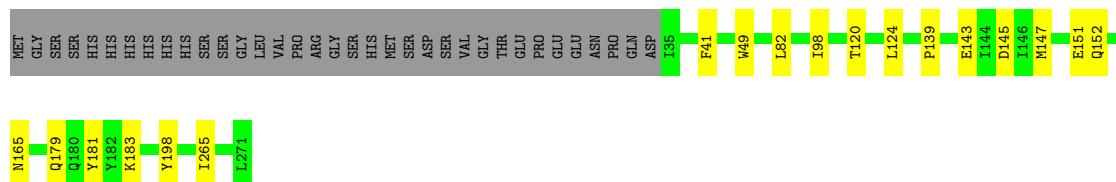
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	AAA	106	Total	O	0	0
			106	106		
3	BBB	93	Total	O	0	0
			93	93		

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.

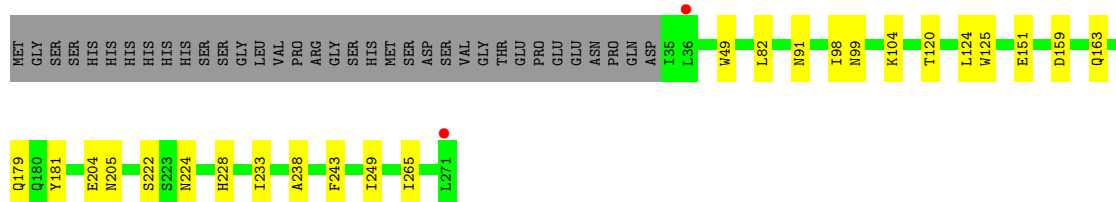
- Molecule 1: Glycosyl hydrolase family 16

Chain AAA: 




- Molecule 1: Glycosyl hydrolase family 16

Chain BBB: 



- Molecule 2: beta-D-glucopyranose-(1-4)-beta-D-glucopyranose-(1-3)-N-acetyl-beta-D-glucopyranosylamine

Chain A: 



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	168.58Å 61.15Å 49.41Å 90.00° 93.61° 90.00°	Depositor
Resolution (Å)	84.12 – 2.15 84.12 – 2.15	Depositor EDS
% Data completeness (in resolution range)	99.5 (84.12-2.15) 99.5 (84.12-2.15)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.93 (at 2.14Å)	Xtriage
Refinement program	REFMAC 5.8.0258	Depositor
R, R_{free}	0.186 , 0.207 0.192 , 0.212	Depositor DCC
R_{free} test set	1311 reflections (4.77%)	wwPDB-VP
Wilson B-factor (Å ²)	27.4	Xtriage
Anisotropy	0.389	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.41 , 51.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.51$, $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	7688	wwPDB-VP
Average B, all atoms (Å ²)	31.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.95% 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: NBG, 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	AAA	0.64	0/1946	0.78	0/2632
1	BBB	0.64	0/1946	0.77	0/2631
All	All	0.64	0/3892	0.77	0/5263

There are no bond length outliers.

There are no bond angle outliers.

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	AAA	1897	1808	1792	9	0
1	BBB	1897	1815	1799	16	0
2	A	37	35	31	0	0
3	AAA	106	0	0	0	0
3	BBB	93	0	0	1	0
All	All	4030	3658	3622	24	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 3.

The worst 5 of 24 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:179:GLN:OE1	1:BBB:181[A]:TYR:HE1	1.87	0.58
1:AAA:147:MET:HG3	1:AAA:198:TYR:CE2	2.39	0.57
1:BBB:179:GLN:OE1	1:BBB:181[A]:TYR:CE1	2.58	0.57
1:AAA:120:THR:HA	1:AAA:124:LEU:HB2	1.89	0.55
1:BBB:205:ASN:HD22	1:BBB:222:SER:HB3	1.75	0.51

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	AAA	236/272 (87%)	234 (99%)	2 (1%)	0	100	100
1	BBB	236/272 (87%)	232 (98%)	4 (2%)	0	100	100
All	All	472/544 (87%)	466 (99%)	6 (1%)	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	AAA	198/231 (86%)	196 (99%)	2 (1%)	76	79
1	BBB	198/231 (86%)	197 (100%)	1 (0%)	88	91
All	All	396/462 (86%)	393 (99%)	3 (1%)	81	85

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

Mol	Chain	Res	Type
1	AAA	139	PRO
1	AAA	152	GLN
1	BBB	104	LYS

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 ⓘ

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	NBG	A	1	2	15,15,15	0.94	2 (13%)	21,21,21	2.99	7 (33%)
2	BGC	A	2	2	11,11,12	1.06	1 (9%)	15,15,17	1.46	3 (20%)
2	BGC	A	3	2	11,11,12	0.53	0	15,15,17	0.93	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
2	NBG	A	1	2	-	2/6/26/26	0/1/1/1
2	BGC	A	2	2	-	0/2/19/22	0/1/1/1
2	BGC	A	3	2	-	0/2/19/22	0/1/1/1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	2	BGC	O4-C4	-3.30	1.35	1.43
2	A	1	NBG	C1-N1	2.23	1.46	1.43
2	A	1	NBG	C2-C1	2.03	1.55	1.52

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	1	NBG	O5-C1-C2	7.69	117.55	109.83
2	A	1	NBG	C5-O5-C1	6.88	121.85	112.52
2	A	1	NBG	C2-C1-N1	-5.21	105.20	111.30
2	A	1	NBG	C3-C2-C1	3.87	115.57	109.94
2	A	2	BGC	C1-O5-C5	3.32	116.69	112.19

There are no chirality outliers.

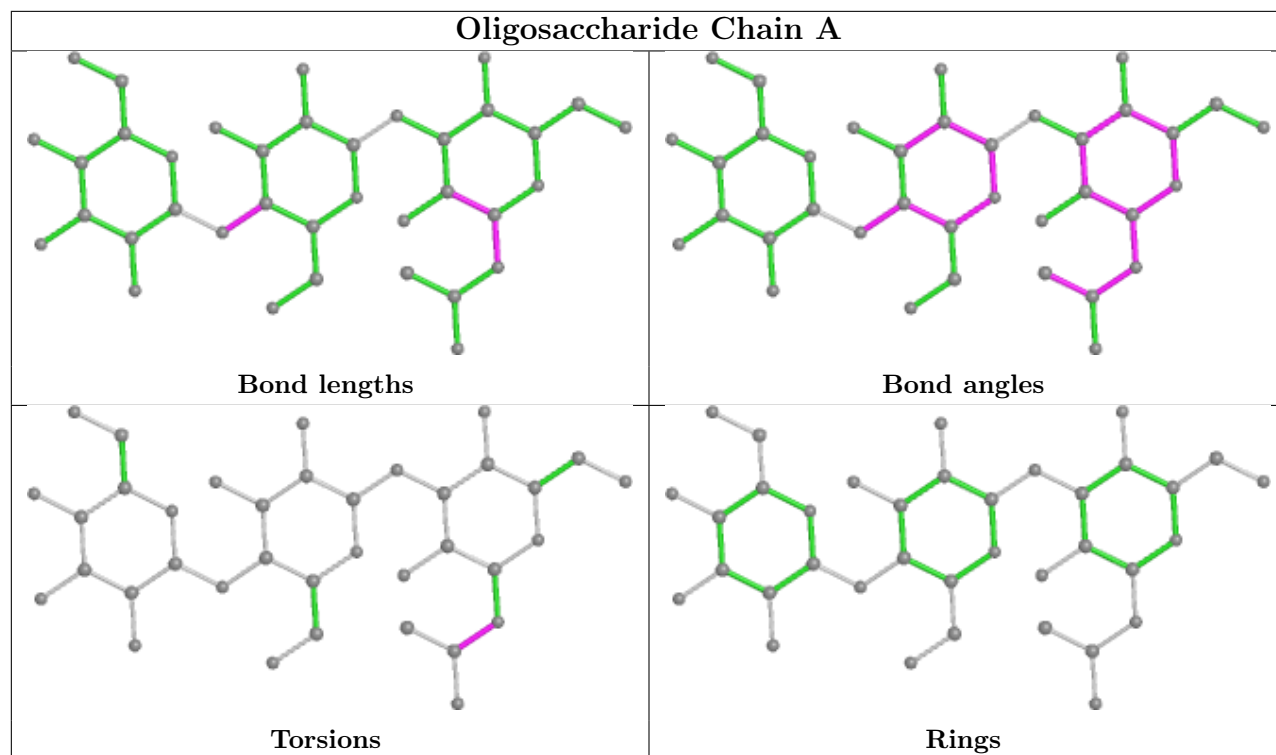
All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	1	NBG	C8-C7-N1-C1
2	A	1	NBG	O7-C7-N1-C1

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	AAA	237/272 (87%)	-0.10	0 100 100	22, 28, 43, 50	0
1	BBB	237/272 (87%)	-0.07	2 (0%) 86 89	20, 30, 45, 62	0
All	All	474/544 (87%)	-0.08	2 (0%) 92 94	20, 30, 44, 62	0

All (2) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	BBB	36	LEU	2.3
1	BBB	271	LEU	2.1

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

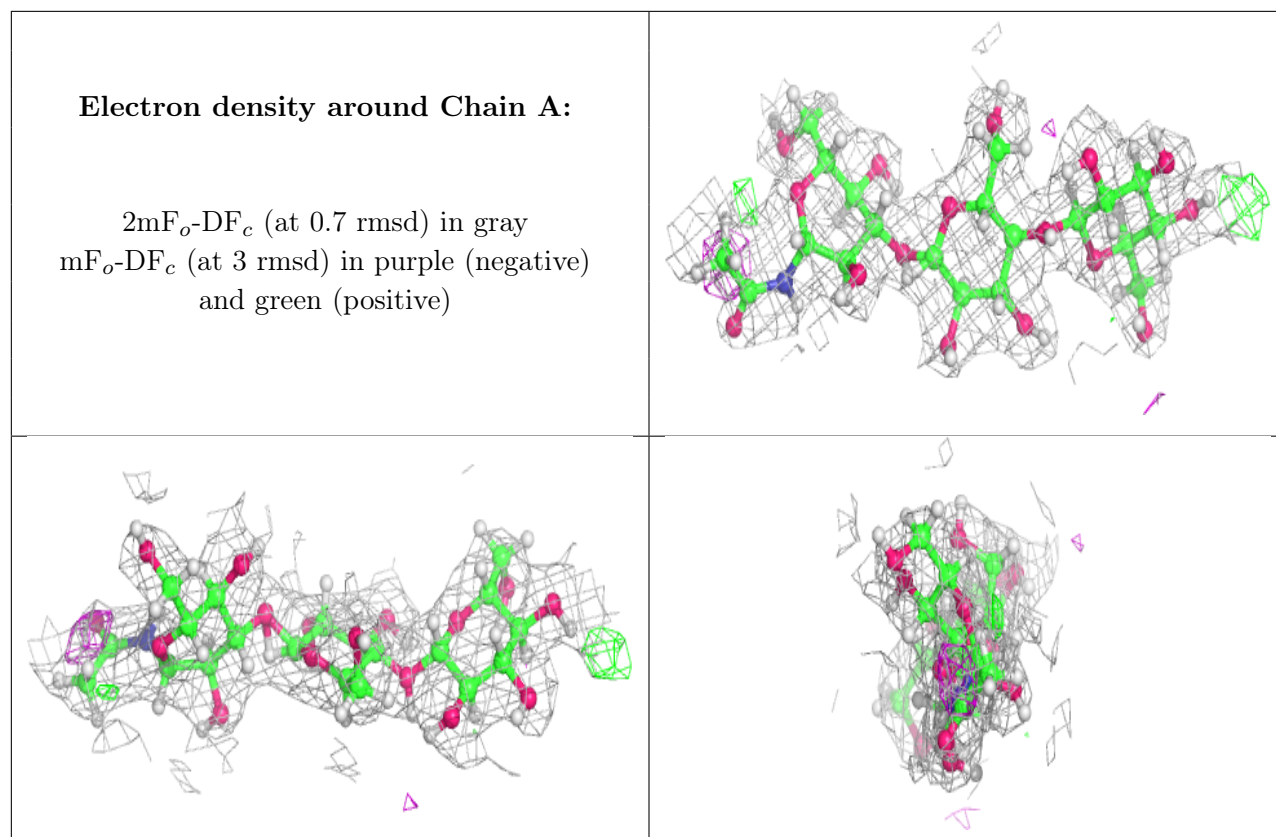
There are no non-standard protein/DNA/RNA residues in this entry.

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	NBG	A	1	15/15	0.91	0.16	37,41,55,61	3
2	BGC	A	3	11/12	0.91	0.11	0,38,45,45	4
2	BGC	A	2	11/12	0.97	0.12	0,30,34,38	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](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.