

# wwPDB X-ray Structure Validation Summary Report (i)

#### Oct 11, 2023 - 04:09 AM EDT

PDB ID	:	7LAQ
Title	:	Crystal structure of Campylobacter jejuni Cj0843c lytic transglycosylase in
		complex with N,N'-diacetylchitobiose
Authors	:	van den Akker, F.; Kumar, V.
Deposited on	:	2021-01-06
Resolution	:	2.58  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 types of validation reports are described at http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (1)) were used in the production of this report:

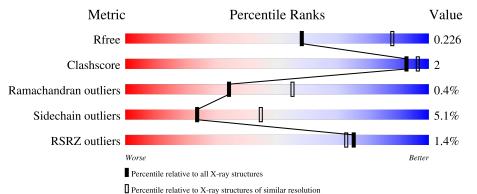
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
$\mathrm{EDS}$	:	2.35.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber $(2001)$
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

# 1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:  $X\text{-}RAY \, DIFFRACTION$ 

The reported resolution of this entry is 2.58 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	$\begin{array}{c} \textbf{Whole archive} \\ (\#\textbf{Entries}) \end{array}$	${f Similar\ resolution}\ (\#{ m Entries,\ resolution\ range}({ m \AA}))$		
$R_{free}$	130704	3676 (2.60-2.56)		
Clashscore	141614	4049 (2.60-2.56)		
Ramachandran outliers	138981	3979 (2.60-2.56)		
Sidechain outliers	138945	3979 (2.60-2.56)		
RSRZ outliers	127900	3614 (2.60-2.56)		

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	А	544	% 87% 8% • 5%
2	В	2	100%
2	D	2	100%



# 2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 4544 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 Lytic transglycosylase domain-containing protein.

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace		
1	А	516	Total 4276	C 2770	N 704	0 787	S 15	0	0	0

Chain	Residue	Modelled	Actual	Comment	Reference
А	-2	MET	-	initiating methionine	UNP A0A3I4HTM2
А	-1	GLY	-	expression tag	UNP A0A3I4HTM2
А	0	SER	-	expression tag	UNP A0A3I4HTM2
А	1	SER	-	expression tag	UNP A0A3I4HTM2
А	2	HIS	-	expression tag	UNP A0A3I4HTM2
А	3	HIS	-	expression tag	UNP A0A3I4HTM2
А	4	HIS	-	expression tag	UNP A0A3I4HTM2
А	5	HIS	-	expression tag	UNP A0A3I4HTM2
А	6	HIS	-	expression tag	UNP A0A3I4HTM2
А	7	HIS	-	expression tag	UNP A0A3I4HTM2
А	8	SER	-	expression tag	UNP A0A3I4HTM2
А	9	SER	-	expression tag	UNP A0A3I4HTM2
А	10	GLY	-	expression tag	UNP A0A3I4HTM2
А	11	LEU	-	expression tag	UNP A0A3I4HTM2
А	12	VAL	-	expression tag	UNP A0A3I4HTM2
А	13	PRO	-	expression tag	UNP A0A3I4HTM2
А	14	ARG	-	expression tag	UNP A0A3I4HTM2
А	15	GLY	-	expression tag	UNP A0A3I4HTM2
А	16	SER	-	expression tag	UNP A0A3I4HTM2
А	17	HIS	-	expression tag	UNP A0A3I4HTM2
А	18	MET	-	expression tag	UNP A0A3I4HTM2

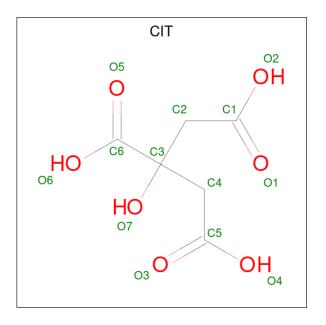
There are 21 discrepancies between the modelled and reference sequences:

• Molecule 2 is an oligosaccharide called 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-a cetamido-2-deoxy-beta-D-glucopyranose.





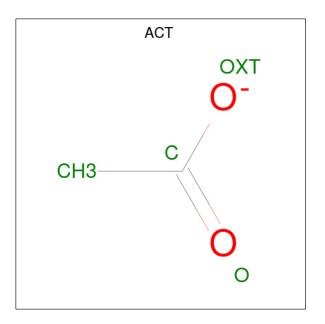
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf	Trace
2	В	2	Total         C         N         O           29         16         2         11	0	0	0
2	D	2	Total         C         N         O           29         16         2         11	0	0	0



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	А	1	Total         C         O           13         6         7	0	0

• Molecule 4 is ACETATE ION (three-letter code: ACT) (formula:  $C_2H_3O_2$ ).





Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	А	1	Total 4	$\begin{array}{c} \mathrm{C} \\ \mathrm{2} \end{array}$	O 2	0	0

• Molecule 5 is water.

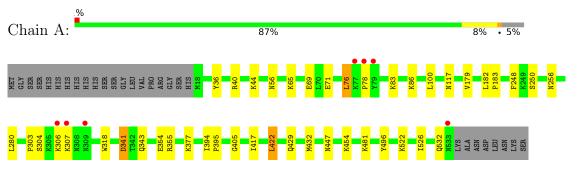
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	А	193	Total O 193 193	0	0



# 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: Lytic transglycosylase domain-containing protein



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

Chain B:

100%

NAG 1 NAG 2

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

Chain D:

100%

NAG1 NAG2



# 4 Data and refinement statistics (i)

Property	Value	Source
Space group	I 2 3	Depositor
Cell constants	176.99Å $176.99$ Å $176.99$ Å	Depositor
a, b, c, $\alpha$ , $\beta$ , $\gamma$	$90.00^{\circ}$ $90.00^{\circ}$ $90.00^{\circ}$	Depositor
Resolution (Å)	29.50 - 2.58	Depositor
Resolution (A)	29.50 - 2.58	EDS
% Data completeness	99.5(29.50-2.58)	Depositor
(in resolution range)	$99.6\ (29.50-2.58)$	EDS
R <sub>merge</sub>	0.17	Depositor
R <sub>sym</sub>	(Not available)	Depositor
$< I/\sigma(I) > 1$	$1.34 (at 2.57 \text{\AA})$	Xtriage
Refinement program	REFMAC 5.8.0258	Depositor
$R, R_{free}$	0.176 , $0.229$	Depositor
$\mathbf{n}, \mathbf{n}_{free}$	0.180 , $0.226$	DCC
$R_{free}$ test set	1574 reflections $(5.44%)$	wwPDB-VP
Wilson B-factor $(Å^2)$	39.9	Xtriage
Anisotropy	0.000	Xtriage
Bulk solvent $k_{sol}(e/Å^3)$ , $B_{sol}(Å^2)$	0.30 , $40.4$	EDS
L-test for twinning <sup>2</sup>	$<  L  > = 0.51, < L^2 > = 0.35$	Xtriage
Estimated twinning fraction	0.002 for -l,-k,-h	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	4544	wwPDB-VP
Average B, all atoms $(Å^2)$	45.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: The largest off-origin peak in the Patterson function is 1.98% of the height of the origin peak. No significant pseudotranslation is detected.

<sup>&</sup>lt;sup>2</sup>Theoretical 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.



<sup>&</sup>lt;sup>1</sup>Intensities estimated from amplitudes.

# 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: CIT, ACT, NAG

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		
IVIOI	Chain	RMSZ	# Z  > 5	RMSZ	# Z  > 5	
1	А	0.68	0/4377	0.82	1/5905~(0.0%)	

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$Observed(^{o})$	$Ideal(^{o})$
1	А	496	TYR	CB-CA-C	6.96	124.33	110.40

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	А	4276	0	4280	13	0
2	В	29	0	27	0	0
2	D	29	0	27	0	0
3	А	13	0	5	0	0
4	А	4	0	3	0	0
5	А	193	0	0	1	0
All	All	4544	0	4342	13	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.



Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:481:LYS:NZ	5:A:701:HOH:O	2.38	0.56
1:A:303:PRO:HB3	1:A:341:ASP:OD1	2.11	0.50
1:A:76:LEU:O	1:A:76:LEU:HD23	2.14	0.48
1:A:417:ILE:O	1:A:422:LEU:HB2	2.14	0.47
1:A:429:GLN:O	1:A:432:MET:HG2	2.16	0.46

The worst 5 of 13 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

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	А	514/544~(94%)	500 (97%)	12 (2%)	2~(0%)	34 55	

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	А	341	ASP
1	А	78	PRO

#### 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 side chain conformation was analysed, and the total number of residues.

Continued on next page...



Conti	nued	from	previous	page	

	$\mathbf{Mol}$	Chain	Analysed	Rotameric	Outliers	Percentiles
Г	Mal	Chain	Analysed	Dotomoria	Outliana	Demoentiles
	IVIOI	Chain	Analysed	Rotameric	Outliers	Percentiles
	1	А	472/496~(95%)	448~(95%)	24~(5%)	24 44

5 of 24 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	А	343	GLN
1	А	377	LYS
1	А	355	ARG
1	А	422	LEU
1	А	117	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	А	193	GLN

#### 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	Bo	ond leng	$\mathbf{ths}$	В	ond ang	les
IVI01	туре	Ullalli	nes		Counts	RMSZ	# Z >2	Counts	RMSZ	# Z >2
Mol	Type	Chain	n Res Link		Bo	ond leng		B	ond ang	
	туре	Ullalli	Ites		Counts	RMSZ	# Z >2	Counts	RMSZ	# Z  > 2
2	NAG	В	1	2	$15,\!15,\!15$	0.27	0	21,21,21	0.73	0
2	NAG	В	2	2	14,14,15	0.69	0	17,19,21	0.71	0
2	NAG	D	1	2	$15,\!15,\!15$	0.44	0	21,21,21	0.86	1 (4%)
2	NAG	D	2	2	14,14,15	0.47	0	17,19,21	0.79	1 (5%)

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	NAG	В	1	2	-	0/6/26/26	0/1/1/1
2	NAG	В	2	2	-	0/6/23/26	0/1/1/1
2	NAG	D	1	2	-	2/6/26/26	0/1/1/1
2	NAG	D	2	2	-	1/6/23/26	0/1/1/1

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$\mathbf{Observed}(^{o})$	$Ideal(^{o})$
2	D	1	NAG	C1-C2-N2	2.70	113.86	110.73
2	D	2	NAG	C1-C2-N2	2.02	113.94	110.49

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	D	1	NAG	C4-C5-C6-O6
2	D	1	NAG	O5-C5-C6-O6
2	D	2	NAG	C4-C5-C6-O6

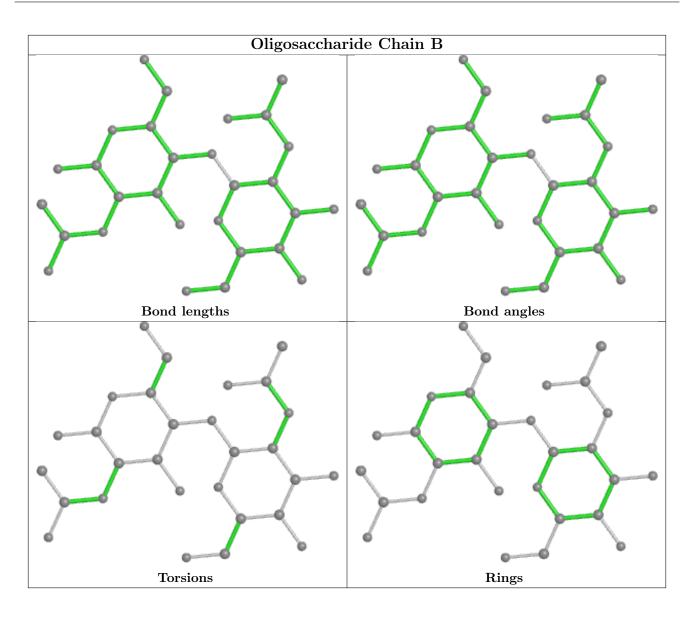
There are no ring outliers.

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.

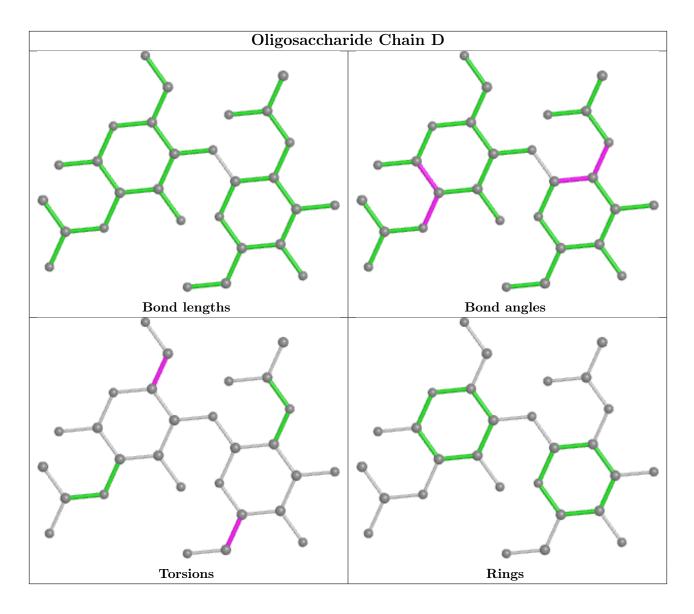












### 5.6 Ligand geometry (i)

2 ligands 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	Chain	Chain	Dec	Link	Bo	ond leng	ths	B	ond ang	les
	туре	Chain	$\operatorname{Res}$	LIIIK	Counts	RMSZ	# Z >2	Counts	RMSZ	# Z  > 2		
3	CIT	А	601	-	12,12,12	1.25	1 (8%)	$17,\!17,\!17$	1.47	4 (23%)		
4	ACT	А	602	-	3,3,3	0.97	0	3,3,3	0.81	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	CIT	А	601	-	-	8/16/16/16	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	А	601	CIT	C3-C6	2.65	1.56	1.53

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$\mathbf{Observed}(^{o})$	$Ideal(^{o})$
3	А	601	CIT	O5-C6-C3	-3.39	117.45	122.25
3	А	601	CIT	O6-C6-C3	2.35	117.13	113.05
3	А	601	CIT	C3-C4-C5	2.26	119.28	113.81
3	А	601	CIT	O3-C5-C4	-2.20	116.52	122.94

There are no chirality outliers.

5 of 8 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	А	601	CIT	C2-C3-C4-C5
3	А	601	CIT	C6-C3-C4-C5
3	А	601	CIT	O7-C3-C4-C5
3	А	601	CIT	C2-C3-C6-O6
3	А	601	CIT	O7-C3-C6-O6

There are no ring outliers.

No monomer is involved in short contacts.

#### 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,  $95^{th}$  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	$\langle RSRZ \rangle$	#RSRZ>2	$OWAB(Å^2)$	Q<0.9
1	А	516/544~(94%)	-0.32	7 (1%) 75 73	27, 40, 73, 136	0

The worst 5 of 7 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	А	533	ASN	3.8
1	А	78	PRO	3.7
1	А	79	TYR	3.1
1	А	309	ASN	3.0
1	А	306	LYS	2.8

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

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

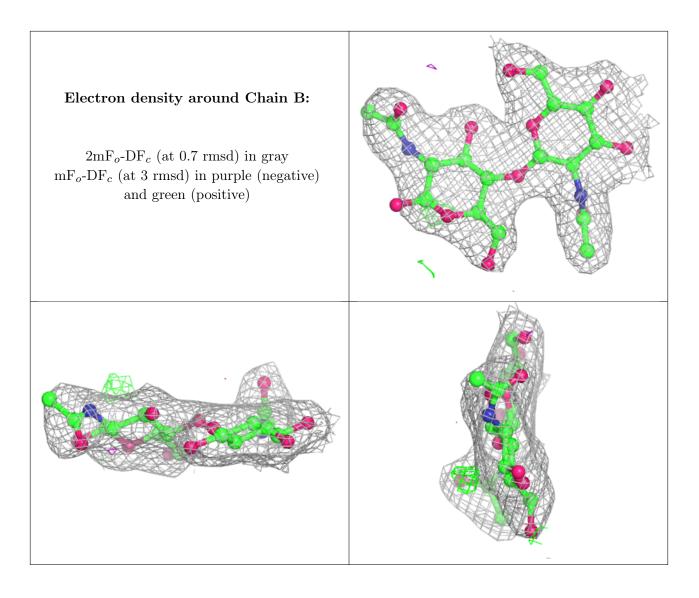
### 6.3 Carbohydrates (i)

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,  $95^{th}$  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.

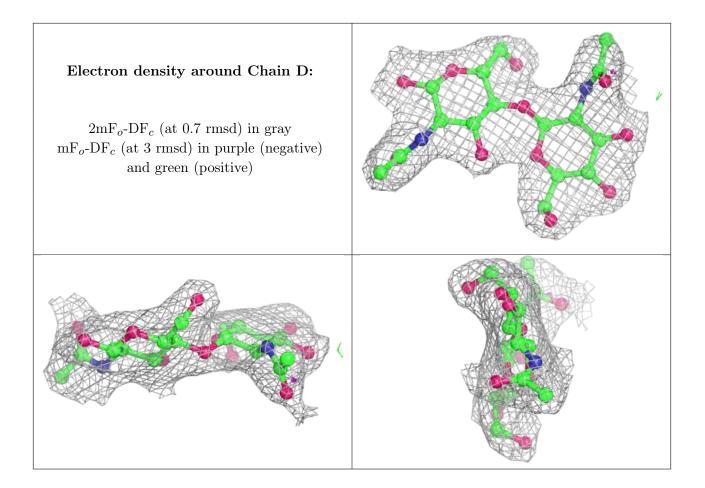
Mol	Type	Chain	Res	Atoms	RSCC	RSR	$\mathbf{B} ext{-factors}(\mathbf{A}^2)$	Q<0.9
2	NAG	D	1	15/15	0.94	0.12	33,39,47,49	0
2	NAG	В	2	14/15	0.96	0.12	34,38,41,42	0
2	NAG	В	1	15/15	0.97	0.11	37,48,54,69	0
2	NAG	D	2	14/15	0.97	0.20	45,51,56,58	0

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)

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,  $95^{th}$  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	$\mathbf{B} ext{-factors}(\mathrm{\AA}^2)$	Q < 0.9
3	CIT	А	601	13/13	0.86	0.20	$57,\!67,\!79,\!80$	0
4	ACT	А	602	4/4	0.90	0.40	56,58,69,70	0

### 6.5 Other polymers (i)

There are no such residues in this entry.

