

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

#### Oct 19, 2022 - 04:14 pm BST

PDB ID	:	4UWT
Title	:	Hypocrea jecorina Cel7A E212Q mutant in complex with p-nitrophenyl cel-
		lobioside
Authors	:	Nutt, A.; Momeni, M.H.; Johansson, G.; Stahlberg, J.
Deposited on		
Resolution	:	1.20  Å(reported)

This is a Full wwPDB X-ray Structure Validation 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:

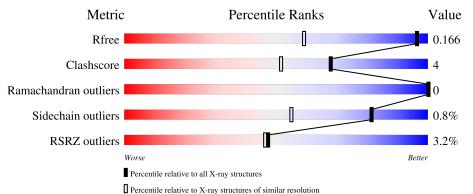
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
$\mathrm{EDS}$	:	2.31.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber $(2001)$
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.31.2

# 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 1.20 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	1223 (1.22-1.18)
Clashscore	141614	1286 (1.22-1.18)
Ramachandran outliers	138981	1240 (1.22-1.18)
Sidechain outliers	138945	1239 (1.22-1.18)
RSRZ outliers	127900	1200 (1.22-1.18)

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	434	3% 94% 6%	•
2	В	2	100%	-
2	С	2	100%	-



#### $4\mathrm{UWT}$

# 2 Entry composition (i)

There are 8 unique types of molecules in this entry. The entry contains 4242 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.

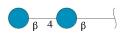
• Molecule 1 is a protein called CELLOBIOHYDROLASE CEL7A.

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace		
1	А	434	Total 3468	C 2130	N 570	O 740	S 28	0	36	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	94	ASP	GLY	cloning artifact	UNP P62694
А	212	GLN	GLU	engineered mutation	UNP P62694

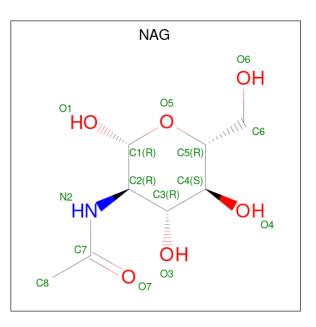
• Molecule 2 is an oligosaccharide called beta-D-glucopyranose-(1-4)-beta-D-glucopyranose.



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf	Trace
2	В	2	Total         C         O           22         12         10	0	0	0
2	С	2	Total         C         O           22         12         10	0	0	0

• Molecule 3 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



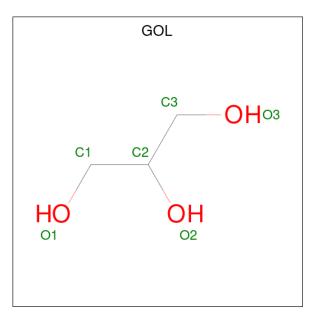


Mol	Chain	Residues	A	tor	ns		ZeroOcc	AltConf
3	А	1	Total 14	C 8	N 1	O 5	0	0

• Molecule 4 is COBALT (II) ION (three-letter code: CO) (formula: Co).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	А	2	Total Co 2 2	0	0

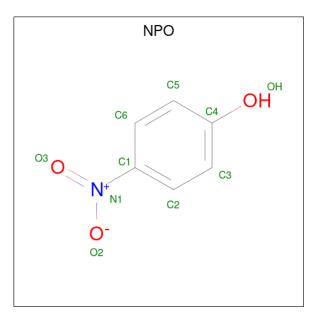
• Molecule 5 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).





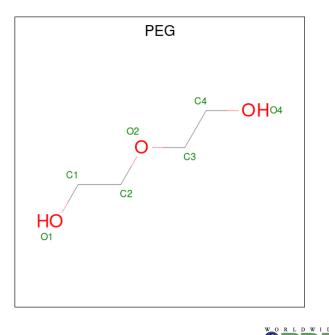
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	А	1	$\begin{array}{ccc} \text{Total} & \text{C} & \text{O} \\ 6 & 3 & 3 \end{array}$	0	0

• Molecule 6 is P-NITROPHENOL (three-letter code: NPO) (formula:  $C_6H_5NO_3$ ).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	А	1	Total C N O 10 6 1 3	0	0
6	А	1	Total         C         N         O           10         6         1         3	0	0

• Molecule 7 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula:  $C_4H_{10}O_3$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	А	1	Total 7	$\begin{array}{c} \mathrm{C} \\ 4 \end{array}$	O 3	0	0

• Molecule 8 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	А	646	Total         O           681         681	0	34



# 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.

- Chain A:

  94%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6%
  6
- Molecule 1: CELLOBIOHYDROLASE CEL7A



# 4 Data and refinement statistics (i)

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants	83.30Å 81.78Å 110.53Å	Depositor
a, b, c, $\alpha$ , $\beta$ , $\gamma$	90.00° 90.00° 90.00°	-
Resolution (Å)	29.19 - 1.20	Depositor
	29.18 - 1.20	EDS
% Data completeness	$100.0 \ (29.19-1.20)$	Depositor
(in resolution range)	$100.0\ (29.18-1.20)$	EDS
R <sub>merge</sub>	0.09	Depositor
R <sub>sym</sub>	(Not available)	Depositor
$< I/\sigma(I) > 1$	$3.62 (at 1.20 \text{\AA})$	Xtriage
Refinement program	REFMAC 5.8.0073	Depositor
D D	0.145 , $0.159$	Depositor
$R, R_{free}$	0.154 , $0.166$	DCC
$R_{free}$ test set	5905 reflections $(5.03\%)$	wwPDB-VP
Wilson B-factor $(Å^2)$	6.5	Xtriage
Anisotropy	0.016	Xtriage
Bulk solvent $k_{sol}(e/Å^3), B_{sol}(Å^2)$	(Not available), (Not available)	EDS
L-test for twinning <sup>2</sup>	$< L >=0.49, < L^2>=0.32$	Xtriage
Estimated twinning fraction	0.014 for -k,-h,-l	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	4242	wwPDB-VP
Average B, all atoms $(Å^2)$	8.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: The largest off-origin peak in the Patterson function is 7.44% 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: PCA, CO, NPO, PEG, BGC, GOL, 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
	Unain	RMSZ	# Z  > 5	RMSZ	# Z  > 5
1	А	0.34	0/3580	0.61	0/4875

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 (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	3468	0	3163	25	0
2	В	22	0	19	0	0
2	С	22	0	19	0	0
3	А	14	0	13	0	0
4	А	2	0	0	0	0
5	А	6	0	8	0	0
6	А	20	0	8	0	0
7	А	7	0	10	1	0
8	А	681	0	0	10	2
All	All	4242	0	3240	26	2

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 4.



		Interatomic	Clash
Atom-1	Atom-2	distance $(\text{\AA})$	overlap (Å)
1:A:353[B]:LYS:HZ3	1:A:353[B]:LYS:HB3	1.33	0.92
1:A:315[B]:ASN:OD1	8:A:1512:HOH:O	1.88	0.89
1:A:353[B]:LYS:NZ	1:A:353[B]:LYS:CB	2.37	0.86
1:A:315[B]:ASN:ND2	8:A:1512:HOH:O	2.16	0.79
1:A:315[B]:ASN:CG	8:A:1512:HOH:O	2.22	0.69
1:A:353[B]:LYS:HB2	1:A:353[B]:LYS:HZ2	1.59	0.68
1:A:353[B]:LYS:HB3	1:A:353[B]:LYS:NZ	2.02	0.63
1:A:353[B]:LYS:NZ	1:A:353[B]:LYS:HB2	2.13	0.60
1:A:253:GLY:N	8:A:1416[B]:HOH:O	2.27	0.57
1:A:315[B]:ASN:ND2	8:A:1500:HOH:O	2.37	0.57
1:A:26[B]:THR:HG21	8:A:1066[B]:HOH:O	2.05	0.57
1:A:140:LEU:CD1	1:A:407[B]:VAL:HG11	2.36	0.55
1:A:278:SER:HA	8:A:1450:HOH:O	2.11	0.51
1:A:203[B]:ILE:HD11	1:A:247:TYR:HD1	1.75	0.50
1:A:202:GLY:O	1:A:203[B]:ILE:HD13	2.12	0.50
1:A:203[B]:ILE:HD12	1:A:247:TYR:HB3	1.94	0.49
1:A:26[B]:THR:HG23	1:A:28:GLN:NE2	2.28	0.48
1:A:318:LEU:HD11	1:A:337[B]:GLU:HG2	1.95	0.47
1:A:45[A]:ASN:OD1	1:A:45[A]:ASN:N	2.50	0.44
1:A:140:LEU:HD11	1:A:407[B]:VAL:HG11	1.98	0.44
1:A:249[B]:ASP:OD1	1:A:250:ASN:N	2.53	0.42
1:A:227:PRO:HD2	1:A:261:CYS:O	2.20	0.41
1:A:203[B]:ILE:HD11	1:A:247:TYR:CD1	2.55	0.40
7:A:445:PEG:H22	8:A:1761:HOH:O	2.20	0.40
1:A:372:ALA:HB1	8:A:1776:HOH:O	2.21	0.40
1:A:286[A]:LYS:HE3	8:A:1455[A]:HOH:O	2.22	0.40

All (26) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

All (2) 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)
8:A:1759:HOH:O	8:A:1759:HOH:O[3_657]	1.31	0.89
8:A:1557:HOH:O	8:A:1557:HOH:O[2_665]	1.61	0.59



### 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	468/434~(108%)	462 (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	А	392/356~(110%)	389~(99%)	3 (1%)	81 55	

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

Mol	Chain	Res	Type
1	А	74	ASP
1	А	223	GLU
1	А	288	LEU

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

#### 5.3.3 RNA (i)

There are no RNA molecules in this entry.



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

1 non-standard protein/DNA/RNA residue is 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	Dog	Link	B	ond leng	gths	В	ond ang	gles
	Type	Ullalli	nes	LINK	Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z >2
1	PCA	А	1	1	7,8,9	0.61	0	$9,\!10,\!12$	1.22	1 (11%)

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
1	PCA	А	1	1	-	0/0/11/13	0/1/1/1

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	1	PCA	CB-CA-C	-2.90	108.72	112.70

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

### 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



Mol	Turne	Chain	Res	Link	Bo	ond leng	$\mathbf{ths}$	Bond angles		
	Type	Chain	nes	LIIIK	Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z >2
2	BGC	В	1	6,2	11,11,12	0.30	0	$15,\!15,\!17$	0.87	0
2	BGC	В	2	2	11,11,12	0.34	0	$15,\!15,\!17$	0.65	0
2	BGC	С	1	6,2	11,11,12	0.28	0	$15,\!15,\!17$	0.74	0
2	BGC	С	2	2	11,11,12	0.39	0	$15,\!15,\!17$	0.56	0

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).

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	В	1	6,2	-	0/2/19/22	0/1/1/1
2	BGC	В	2	2	-	0/2/19/22	0/1/1/1
2	BGC	С	1	6,2	-	0/2/19/22	0/1/1/1
2	BGC	С	2	2	-	0/2/19/22	0/1/1/1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

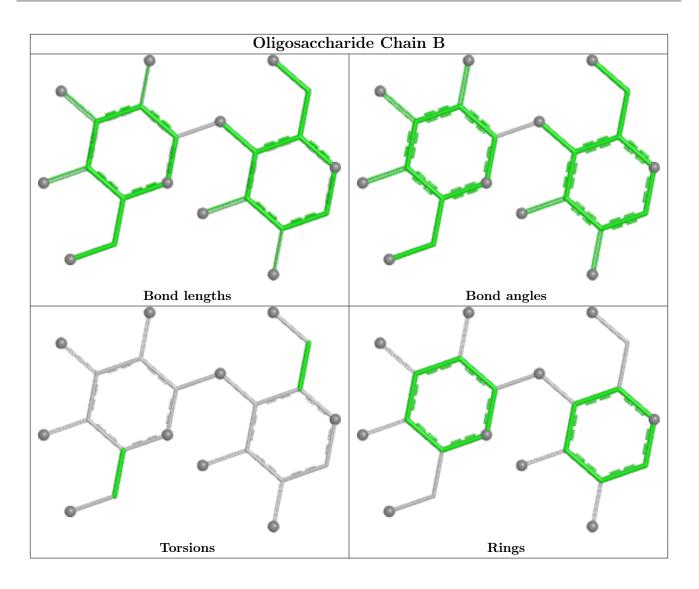
There are no torsion outliers.

There are no ring outliers.

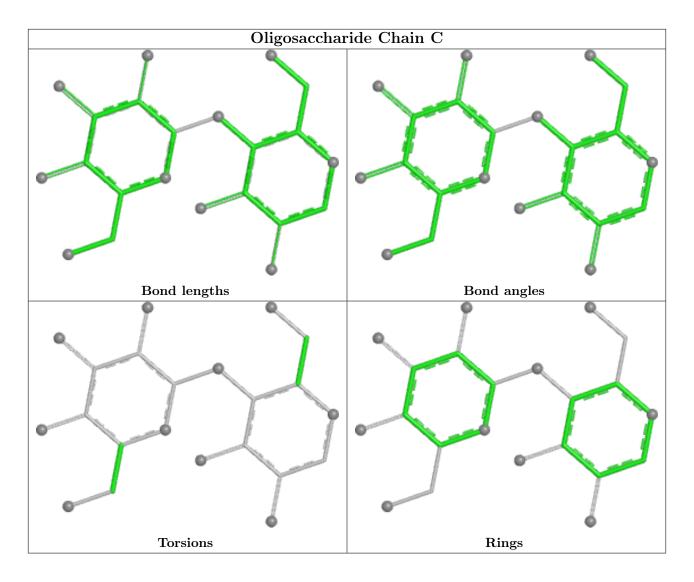
No monomer is involved in short contacts.

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for oligosaccharide.









## 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).

Mal	Mol Type		Res	Link	ink Bond lengths				Bond angles			
	туре	Chain	nes		Counts	RMSZ	# Z >2	Counts	RMSZ	# Z >2		
3	NAG	А	435	1	14,14,15	0.42	0	17,19,21	0.94	1 (5%)		
5	GOL	А	438	-	$5,\!5,\!5$	0.31	0	$5,\!5,\!5$	0.18	0		
6	NPO	А	442	2	9,10,10	<mark>3.53</mark>	1 (11%)	11,13,13	0.51	0		
7	PEG	А	445	-	$6,\!6,\!6$	0.48	0	$5,\!5,\!5$	0.39	0		



Mol	Tuno	Chain	Dog	Link	Bond lengths			Bond angles		
WIOI	туре	Ullaili	nes		Counts	RMSZ	# Z >2	Counts	RMSZ	# Z >2
6	NPO	А	439	2	9,10,10	<mark>3.53</mark>	1 (11%)	11,13,13	0.60	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	NAG	А	435	1	-	0/6/23/26	0/1/1/1
5	GOL	А	438	-	-	0/4/4/4	-
6	NPO	А	442	2	-	2/2/4/4	0/1/1/1
7	PEG	А	445	-	-	2/4/4/4	-
6	NPO	А	439	2	-	0/2/4/4	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	А	439	NPO	O3-N1	10.35	1.40	1.22
6	А	442	NPO	O3-N1	10.33	1.40	1.22

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$Observed(^{o})$	$Ideal(^{o})$
3	А	435	NAG	C1-O5-C5	2.48	115.56	112.19

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	А	442	NPO	C6-C1-N1-O3
7	А	445	PEG	O1-C1-C2-O2
6	А	442	NPO	C2-C1-N1-O3
7	А	445	PEG	C1-C2-O2-C3

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	А	445	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,  $95^{th}$  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	<RSRZ $>$	#RS	$\mathbf{RZ}>$	$\cdot 2$	$OWAB(Å^2)$	Q<0.9
1	А	433/434 (99%)	0.19	14 (3%)	47	46	3,6,12,23	2(0%)

All (14) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	А	254	GLY	3.9
1	А	388	SER	3.2
1	А	389	THR	3.2
1	А	45[A]	ASN	3.0
1	А	371	TYR	2.7
1	А	390	PRO	2.6
1	А	249[A]	ASP	2.5
1	А	250	ASN	2.4
1	А	247	TYR	2.3
1	А	100	ALA	2.2
1	А	387	SER	2.2
1	А	158[A]	TYR	2.1
1	А	65[A]	GLU	2.0
1	А	384	ASN	2.0

## 6.2 Non-standard residues in protein, DNA, RNA chains (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	$B-factors(A^2)$	Q<0.9
1	PCA	А	1	8/9	0.98	0.06	$7,\!8,\!8,\!8$	0

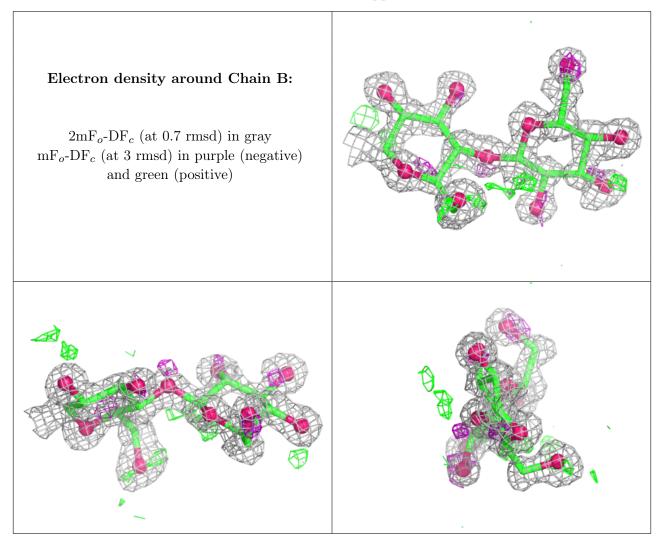


### 6.3 Carbohydrates (i)

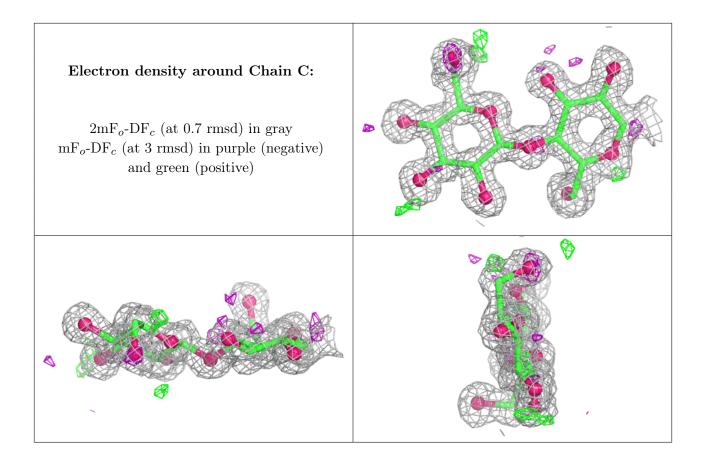
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median,  $95^{th}$  percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9'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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	$B-factors(Å^2)$	Q<0.9
2	BGC	В	2	11/12	0.89	0.21	$11,\!12,\!12,\!12$	11
2	BGC	В	1	11/12	0.90	0.20	9,11,11,12	11
2	BGC	С	1	11/12	0.96	0.08	8,9,10,11	0
2	BGC	С	2	11/12	0.97	0.08	$7,\!7,\!8,\!9$	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	$B-factors(Å^2)$	Q<0.9
7	PEG	А	445	7/7	0.62	0.23	$30,\!31,\!32,\!33$	0
6	NPO	А	442	10/10	0.81	0.22	$13,\!17,\!19,\!19$	2
6	NPO	А	439	10/10	0.82	0.23	13,14,18,18	10
4	CO	А	437	1/1	0.85	0.11	$27,\!27,\!27,\!27$	0
5	GOL	А	438	6/6	0.91	0.19	$14,\!15,\!16,\!16$	0
3	NAG	А	435	14/15	0.93	0.12	6, 8, 12, 13	0
4	CO	А	436	1/1	1.00	0.04	2,2,2,2	1

### 6.5 Other polymers (i)

There are no such residues in this entry.

