



# Full wwPDB X-ray Structure Validation Report ⓘ

Sep 16, 2025 – 04:17 pm BST

PDB ID : 9GSV / pdb\_00009gsv  
Title : Crystal structure of human lysosomal acid-alpha-glucosidase, GAA, in complex with iminosugar compound 4c  
Authors : Sulzenbacher, G.; Roig-Zamboni, V.; Moracci, M.; Parenti, G.; Py, S.  
Deposited on : 2024-09-16  
Resolution : 1.80 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](mailto: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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity	:	4-5-2 with Phenix2.0rc1
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	2.0rc1
EDS	:	3.0
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.006 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.45.1

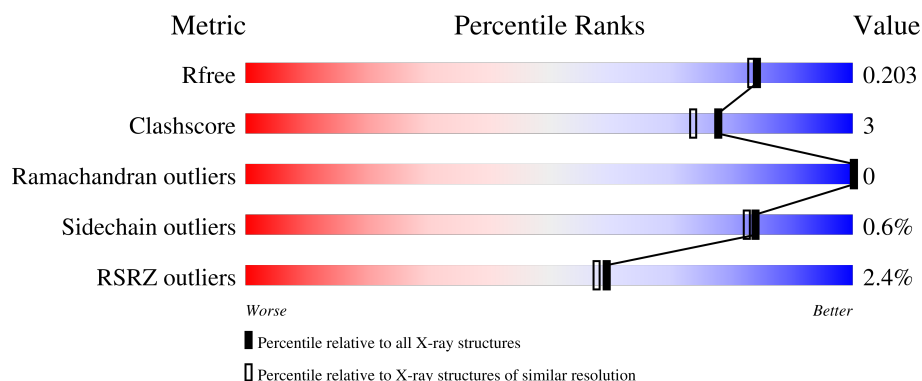
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*


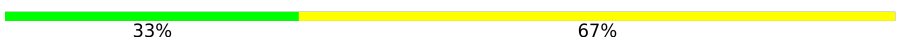
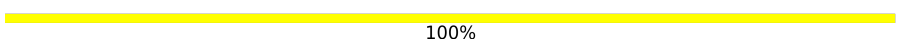
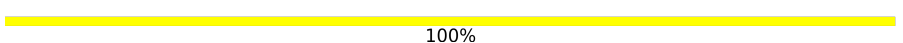
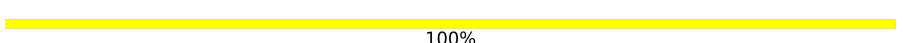
The reported resolution of this entry is 1.80 Å.

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))
$R_{free}$	164625	7108 (1.80-1.80)
Clashscore	180529	8162 (1.80-1.80)
Ramachandran outliers	177936	8077 (1.80-1.80)
Sidechain outliers	177891	8076 (1.80-1.80)
RSRZ outliers	164620	7108 (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  $\geq 3$ , 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	A	952	
2	B	3	
3	C	3	
3	E	3	
4	D	2	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
5	F	2	

## 2 Entry composition [i](#)

There are 12 unique types of molecules in this entry. The entry contains 7419 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.

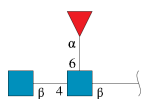
- Molecule 1 is a protein called Lysosomal alpha-glucosidase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	842	Total	C	N	O	S	0	4	0
			6651	4269	1119	1230	33			

There are 3 discrepancies between the modelled and reference sequences:

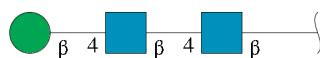
Chain	Residue	Modelled	Actual	Comment	Reference
A	199	ARG	HIS	variant	UNP P10253
A	223	HIS	ARG	variant	UNP P10253
A	780	ILE	VAL	variant	UNP P10253

- Molecule 2 is an oligosaccharide called 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-[alpha-L-fucopyranose-(1-6)]2-acetamido-2-deoxy-beta-D-glucopyranose.



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	B	3	Total	C	N	O	0	0	0
			38	22	2	14			

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



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	C	3	Total	C	N	O	0	0	0
			39	22	2	15			

*Continued on next page...*

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	E	3	Total	C	N	O	0	0	0
			39	22	2	15			

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



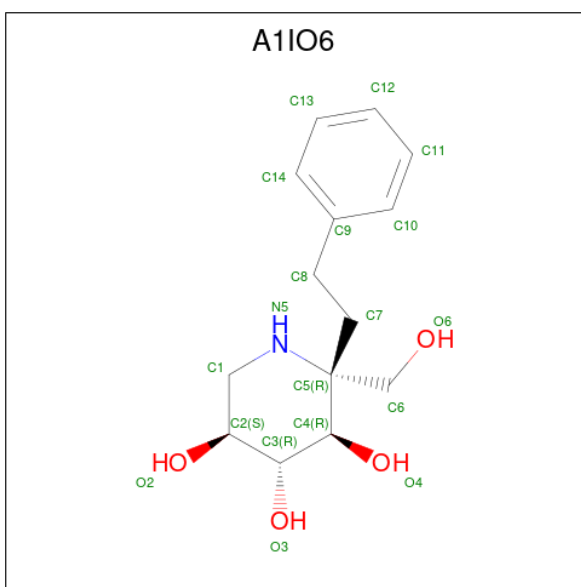
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
4	D	2	Total	C	N	O	0	0	0
			28	16	2	10			

- Molecule 5 is an oligosaccharide called alpha-L-fucopyranose-(1-6)-2-acetamido-2-deoxy-beta-D-glucopyranose.



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
5	F	2	Total	C	N	O	0	0	0
			24	14	1	9			

- Molecule 6 is (2 {R},3 {R},4 {R},5 {S})-2-(hydroxymethyl)-2-(2-phenylethyl)piperidine-3,4,5-triol (CCD ID: A1IO6) (formula: C<sub>14</sub>H<sub>21</sub>NO<sub>4</sub>) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	A	1	Total	C	N	O	0	0
			19	14	1	4		

- Molecule 7 is SULFATE ION (CCD ID: SO4) (formula: O<sub>4</sub>S).

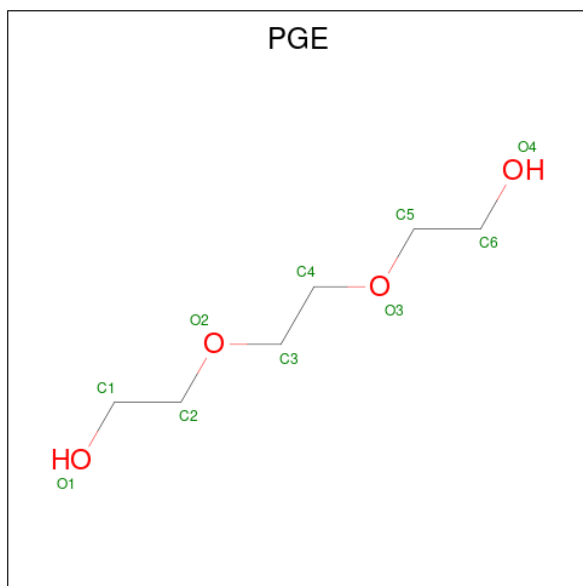


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	A	1	Total	O	S	0	0
			5	4	1		
7	A	1	Total	O	S	0	0
			5	4	1		
7	A	1	Total	O	S	0	0
			5	4	1		

- Molecule 8 is CHLORIDE ION (CCD ID: CL) (formula: Cl).

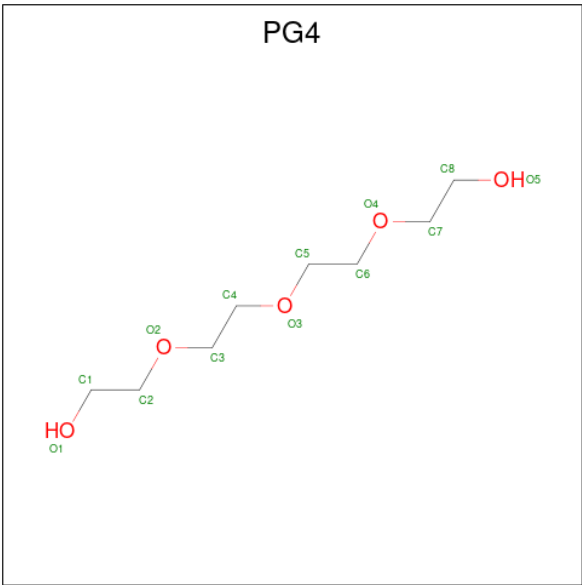
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	A	5	Total Cl 5 5	0	0

- Molecule 9 is TRIETHYLENE GLYCOL (CCD ID: PGE) (formula: C<sub>6</sub>H<sub>14</sub>O<sub>4</sub>).



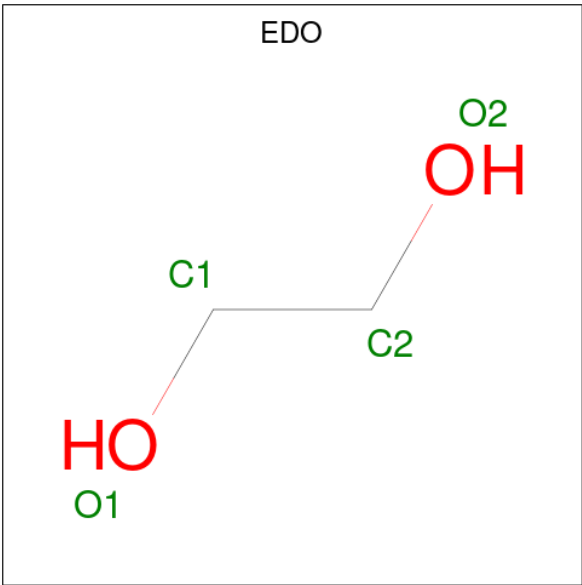
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
9	A	1	Total C O 10 6 4	0	0
9	A	1	Total C O 10 6 4	0	0

- Molecule 10 is TETRAETHYLENE GLYCOL (CCD ID: PG4) (formula: C<sub>8</sub>H<sub>18</sub>O<sub>5</sub>).



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

- Molecule 11 is 1,2-ETHANEDIOL (CCD ID: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
11	A	1	Total	C	O	0	0
			4	2	2		
11	A	1	Total	C	O	0	0
			4	2	2		

Continued on next page...



*Continued from previous page...*

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
11	A	1	Total C O 4 2 2	0	0
11	A	1	Total C O 4 2 2	0	0
11	A	1	Total C O 4 2 2	0	0
11	A	1	Total C O 4 2 2	0	0
11	A	1	Total C O 4 2 2	0	0
11	A	1	Total C O 4 2 2	0	0
11	A	1	Total C O 4 2 2	0	0
11	A	1	Total C O 4 2 2	0	0

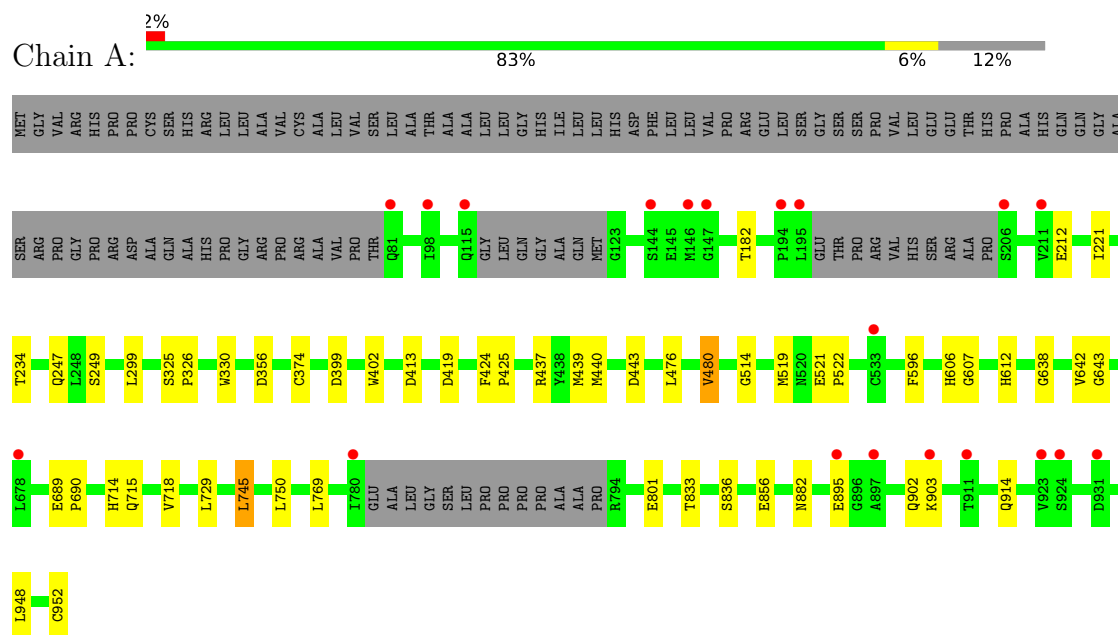
- Molecule 12 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
12	A	475	Total O 475 475	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: Lysosomal alpha-glucosidase



- Molecule 2: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-[alpha-L-fucopyranose-(1-6)]2-acetamido-2-deoxy-beta-D-glucopyranose



- Molecule 3: beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose




- Molecule 3: beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain E:  100%

MAG1  
MAG2  
BGA3

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

Chain D:  100%

MAG1  
MAG2

- Molecule 5: alpha-L-fucopyranose-(1-6)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain F:  50% 50%

MAG1  
FUC2

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	96.37Å 102.30Å 128.16Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.23 – 1.80 48.23 – 1.80	Depositor EDS
% Data completeness (in resolution range)	99.8 (48.23-1.80) 99.9 (48.23-1.80)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.66 (at 1.79Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, $R_{free}$	0.160 , 0.192 0.171 , 0.203	Depositor DCC
$R_{free}$ test set	5855 reflections (3.79%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	29.1	Xtriage
Anisotropy	0.200	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 37.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	7419	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	33.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.49% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<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.

## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: PGE, SO4, A1IO6, PG4, BMA, CSO, NAG, FUC, CL, EDO

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	1.00	2/6853 (0.0%)	1.13	2/9360 (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	A	0	1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	606	HIS	CE1-NE2	5.75	1.38	1.32
1	A	643	GLY	C-O	5.48	1.29	1.23

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	413	ASP	CA-CB-CG	5.30	117.90	112.60
1	A	356	ASP	CA-CB-CG	5.23	117.83	112.60

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	745	LEU	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	A	6651	0	6436	35	0
2	B	38	0	34	0	0
3	C	39	0	34	0	0
3	E	39	0	34	0	0
4	D	28	0	25	0	0
5	F	24	0	22	0	0
6	A	19	0	0	0	0
7	A	15	0	0	0	0
8	A	5	0	0	0	0
9	A	20	0	28	2	0
10	A	26	0	36	1	0
11	A	40	0	60	4	0
12	A	475	0	0	4	1
All	All	7419	0	6709	37	1

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.

All (37) 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:902:GLN:HB3	11:A:1016:EDO:H12	1.54	0.86
11:A:1017:EDO:H22	12:A:1102:HOH:O	1.89	0.72
1:A:902:GLN:CB	11:A:1016:EDO:H12	2.23	0.68
1:A:882:ASN:ND2	12:A:1101:HOH:O	2.17	0.57
1:A:521:GLU:N	1:A:522:PRO:HA	2.20	0.56
1:A:234:THR:HA	1:A:249:SER:O	2.06	0.56
1:A:902:GLN:HG3	1:A:903:LYS:HG3	1.89	0.55
1:A:948:LEU:C	1:A:948:LEU:HD23	2.33	0.53
1:A:325:SER:N	1:A:326:PRO:HA	2.25	0.51
1:A:439[A]:MET:SD	1:A:514:GLY:HA3	2.50	0.51
1:A:182:THR:HG21	9:A:1010:PGE:C1	2.41	0.50
1:A:769:LEU:HD12	1:A:801:GLU:HA	1.92	0.50
1:A:437:ARG:CZ	11:A:1022:EDO:H12	2.42	0.50
1:A:612:HIS:O	1:A:642:VAL:HA	2.12	0.50

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:374:CYS:HB2	1:A:402:TRP:HB2	1.94	0.49
1:A:419:ASP:HB2	12:A:1345:HOH:O	2.13	0.48
1:A:689:GLU:N	1:A:690:PRO:CD	2.77	0.48
1:A:715:GLN:NE2	1:A:856:GLU:OE1	2.47	0.48
1:A:607:GLY:HA3	1:A:638:GLY:O	2.15	0.46
1:A:212:GLU:HG3	1:A:221:ILE:HB	1.98	0.46
1:A:476:LEU:C	1:A:476:LEU:HD23	2.42	0.44
1:A:182:THR:HG21	9:A:1010:PGE:H12	2.00	0.44
1:A:895:GLU:OE1	1:A:895:GLU:HA	2.16	0.43
1:A:439[B]:MET:SD	1:A:596:PHE:HD2	2.42	0.43
1:A:424:PHE:HB3	1:A:425:PRO:HD3	2.00	0.43
1:A:247:GLN:HA	1:A:330:TRP:O	2.19	0.43
1:A:714:HIS:CE1	1:A:718:VAL:HG11	2.54	0.43
1:A:439[B]:MET:HE2	1:A:439[B]:MET:HB3	1.92	0.43
1:A:833:THR:HA	1:A:836:SER:OG	2.19	0.43
1:A:439[B]:MET:SD	1:A:596:PHE:CD2	3.13	0.42
1:A:443:ASP:HB3	12:A:1499:HOH:O	2.19	0.41
1:A:914:GLN:N	1:A:952:CYS:OXT	2.53	0.41
1:A:480:VAL:CG1	1:A:519[B]:MET:HE1	2.50	0.41
1:A:299:LEU:C	1:A:299:LEU:HD23	2.46	0.40
1:A:745:LEU:HA	1:A:750:LEU:O	2.21	0.40
10:A:1013:PG4:H31	10:A:1013:PG4:H12	1.67	0.40

All (1) 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)
12:A:1311:HOH:O	12:A:1538:HOH:O[4_546]	2.18	0.02

## 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	A	837/952 (88%)	815 (97%)	22 (3%)	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	A	723/805 (90%)	719 (99%)	4 (1%)	84	82

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	399	ASP
1	A	440	MET
1	A	480	VAL
1	A	729	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (6) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	177	ASN
1	A	223	HIS
1	A	316	ASN
1	A	352	GLN
1	A	401	GLN
1	A	570	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.



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

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	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
1	CSO	A	938	1	3,6,7	0.66	0	0,6,8	-	-

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	CSO	A	938	1	-	0/1/5/7	-

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.

## 5.5 Carbohydrates ⓘ

13 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	NAG	B	1	2,1	14,14,15	0.62	0	17,19,21	1.03	1 (5%)
2	NAG	B	2	2	14,14,15	0.42	0	17,19,21	0.76	0
2	FUC	B	3	2	10,10,11	0.61	0	14,14,16	1.00	1 (7%)
3	NAG	C	1	3,1	14,14,15	0.91	1 (7%)	17,19,21	1.19	1 (5%)
3	NAG	C	2	3	14,14,15	0.59	0	17,19,21	1.12	1 (5%)
3	BMA	C	3	3	11,11,12	0.36	0	15,15,17	1.34	2 (13%)
4	NAG	D	1	4,1	14,14,15	0.37	0	17,19,21	1.37	2 (11%)
4	NAG	D	2	4	14,14,15	0.56	0	17,19,21	1.38	3 (17%)
3	NAG	E	1	3,1	14,14,15	0.87	1 (7%)	17,19,21	1.11	0
3	NAG	E	2	3	14,14,15	0.44	0	17,19,21	1.12	1 (5%)
3	BMA	E	3	3	11,11,12	0.48	0	15,15,17	1.50	2 (13%)
5	NAG	F	1	5,1	14,14,15	0.59	0	17,19,21	1.50	4 (23%)
5	FUC	F	2	5	10,10,11	0.52	0	14,14,16	1.04	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	NAG	B	1	2,1	-	0/6/23/26	0/1/1/1
2	NAG	B	2	2	-	0/6/23/26	0/1/1/1
2	FUC	B	3	2	-	-	0/1/1/1
3	NAG	C	1	3,1	-	0/6/23/26	0/1/1/1
3	NAG	C	2	3	-	0/6/23/26	0/1/1/1
3	BMA	C	3	3	-	0/2/19/22	0/1/1/1
4	NAG	D	1	4,1	-	0/6/23/26	0/1/1/1
4	NAG	D	2	4	-	0/6/23/26	0/1/1/1
3	NAG	E	1	3,1	-	0/6/23/26	0/1/1/1
3	NAG	E	2	3	-	0/6/23/26	0/1/1/1
3	BMA	E	3	3	-	0/2/19/22	0/1/1/1
5	NAG	F	1	5,1	-	1/6/23/26	0/1/1/1
5	FUC	F	2	5	-	-	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	1	NAG	C1-C2	2.42	1.56	1.52
3	E	1	NAG	O5-C1	-2.06	1.40	1.43

All (18) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	E	3	BMA	C1-O5-C5	3.30	116.66	112.19
3	C	3	BMA	O5-C1-C2	3.12	115.59	110.77
4	D	2	NAG	O5-C1-C2	3.08	116.16	111.29
4	D	1	NAG	O5-C1-C2	-3.08	106.42	111.29
5	F	1	NAG	O5-C5-C6	2.84	111.66	107.20
3	C	3	BMA	C1-C2-C3	2.74	113.04	109.67
2	B	1	NAG	O5-C5-C6	2.55	111.20	107.20
4	D	2	NAG	O5-C5-C4	-2.54	104.65	110.83
5	F	1	NAG	C8-C7-N2	-2.46	111.93	116.10
5	F	1	NAG	C1-C2-N2	2.44	114.66	110.49
3	C	1	NAG	C2-N2-C7	-2.44	119.43	122.90
4	D	2	NAG	C3-C4-C5	-2.38	105.99	110.24
5	F	1	NAG	O7-C7-N2	2.24	126.07	121.95
3	E	3	BMA	C3-C4-C5	2.24	114.23	110.24
4	D	1	NAG	C1-C2-N2	2.23	114.29	110.49
3	C	2	NAG	O5-C5-C4	-2.14	105.62	110.83
2	B	3	FUC	C1-O5-C5	2.05	117.42	112.78
3	E	2	NAG	O7-C7-N2	-2.03	118.21	121.95

There are no chirality outliers.

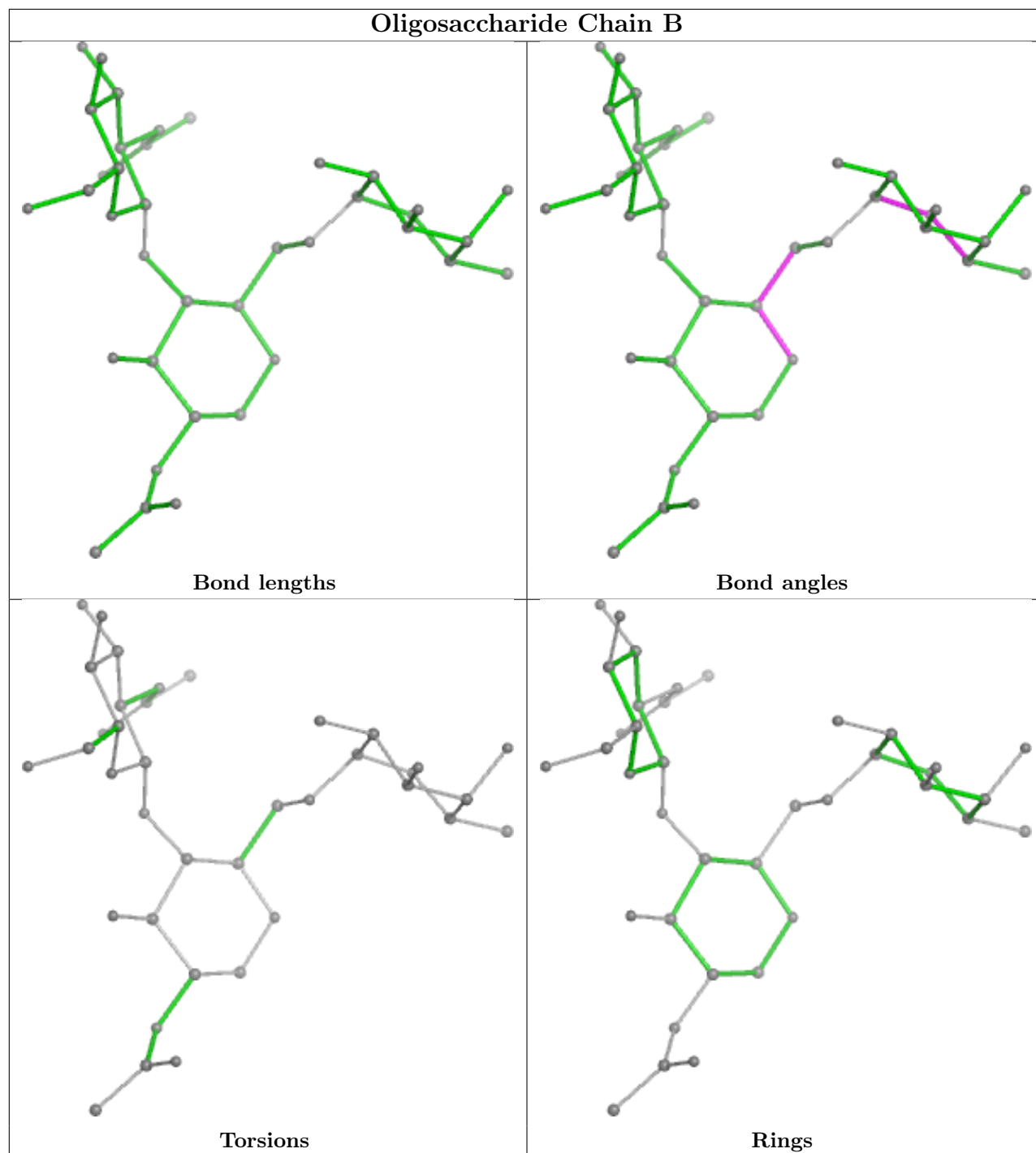
All (1) torsion outliers are listed below:

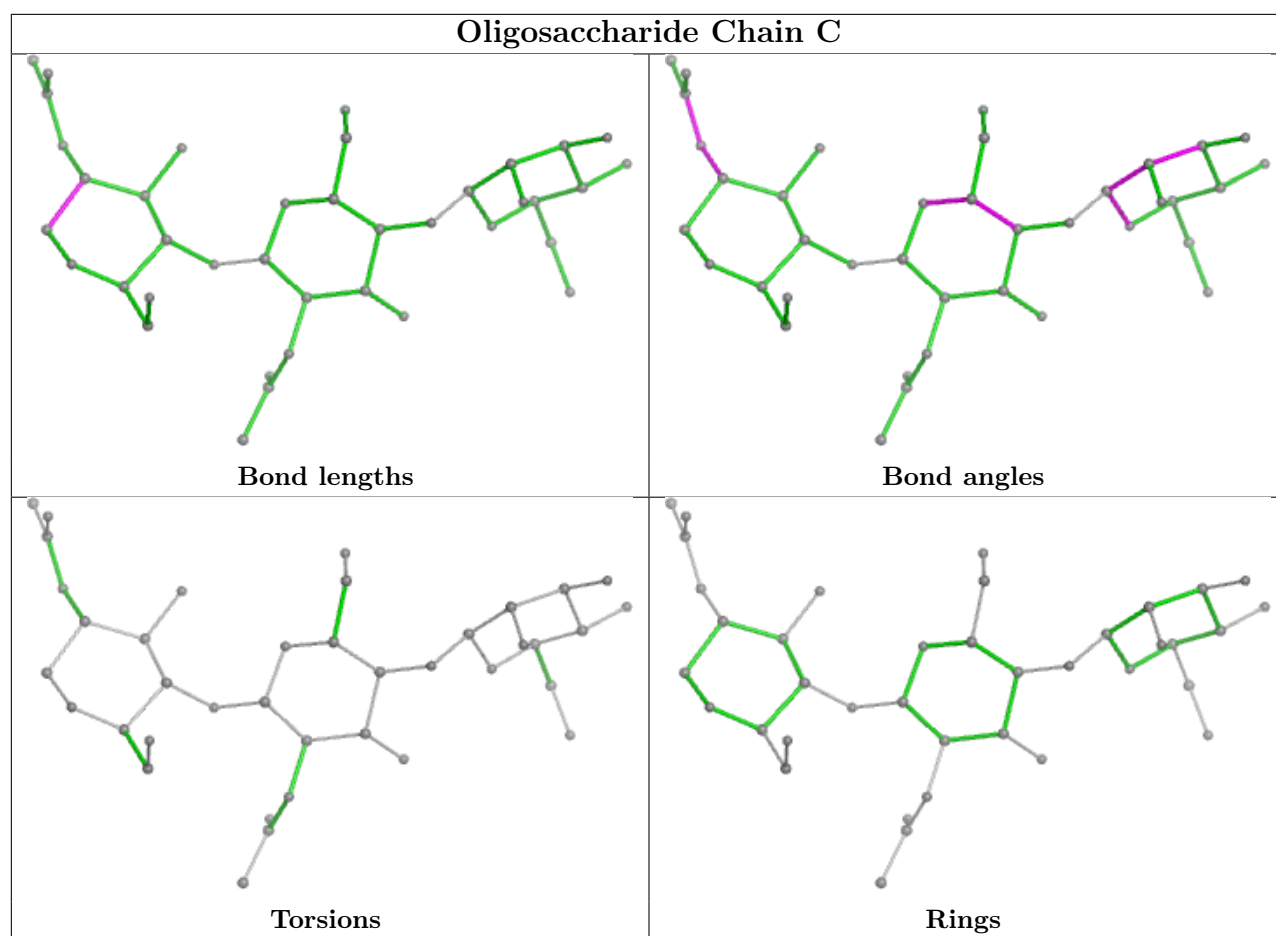
Mol	Chain	Res	Type	Atoms
5	F	1	NAG	C4-C5-C6-O6

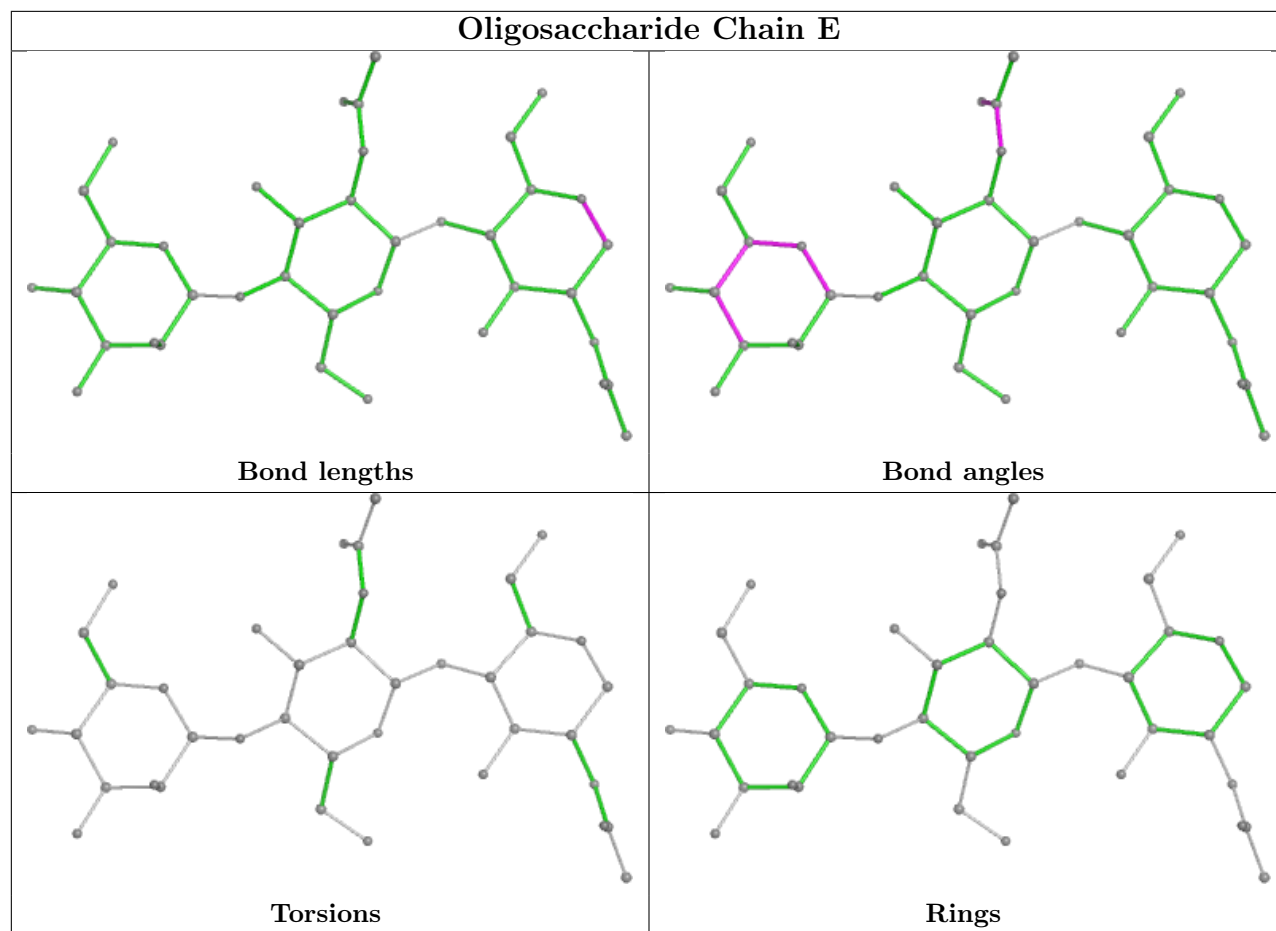
There are no ring outliers.

No monomer is involved in short contacts.

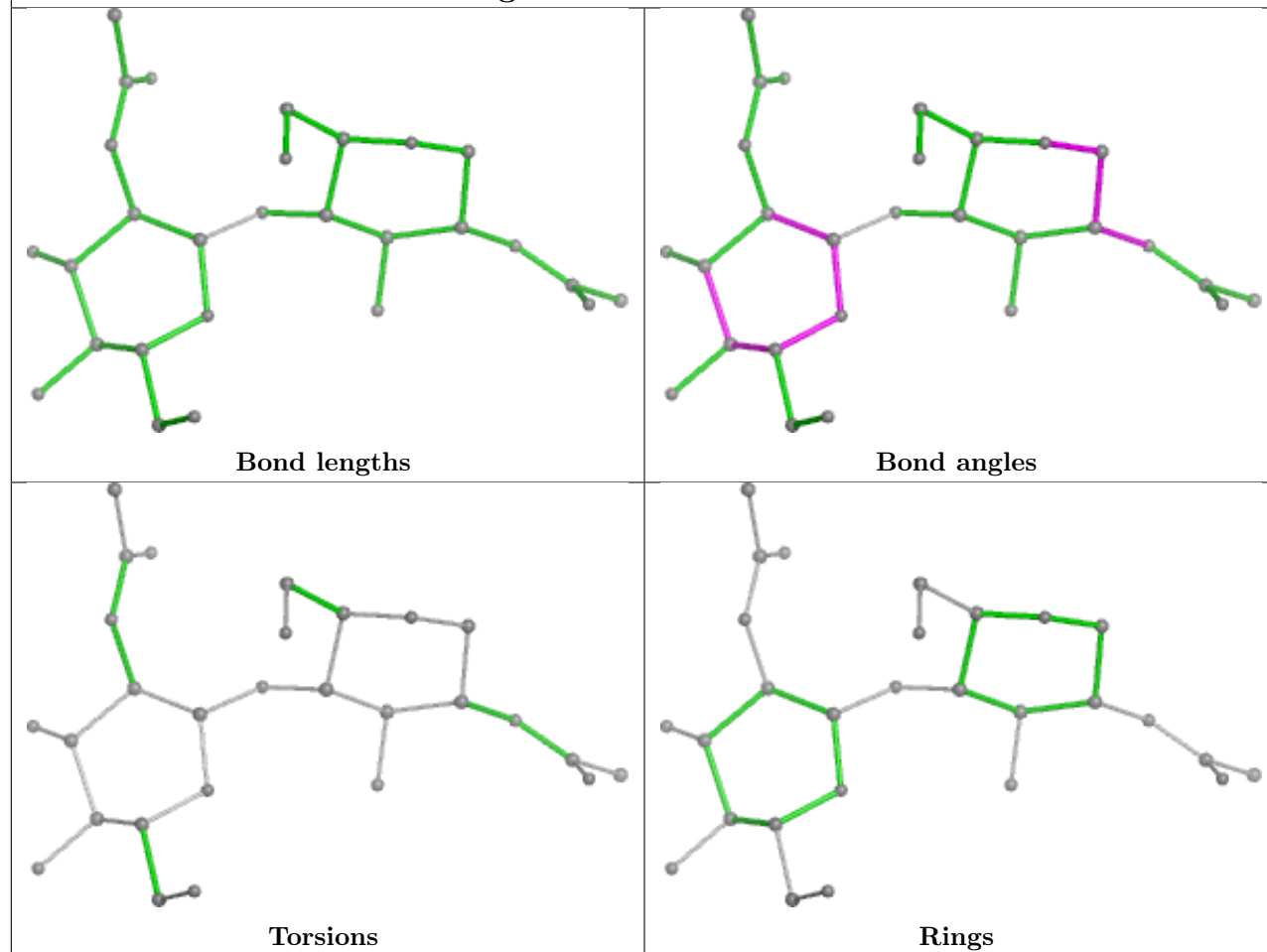
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for oligosaccharide.



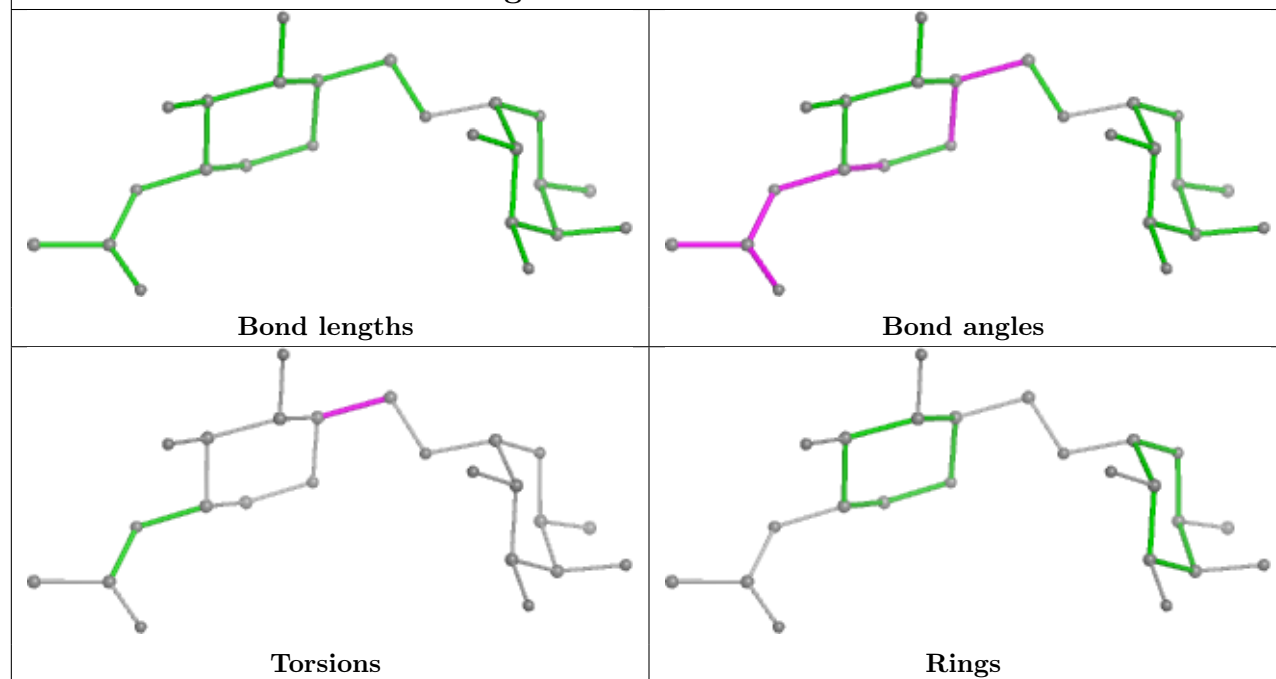




## Oligosaccharide Chain D



## Oligosaccharide Chain F



## 5.6 Ligand geometry

Of 23 ligands modelled in this entry, 5 are monoatomic - leaving 18 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
7	SO4	A	1002	-	4,4,4	0.36	0	6,6,6	0.22	0
11	EDO	A	1016	-	3,3,3	0.36	0	2,2,2	0.03	0
7	SO4	A	1003	-	4,4,4	0.30	0	6,6,6	0.14	0
9	PGE	A	1010	-	9,9,9	0.29	0	8,8,8	0.21	0
9	PGE	A	1011	-	9,9,9	0.24	0	8,8,8	0.15	0
11	EDO	A	1020	-	3,3,3	0.71	0	2,2,2	0.36	0
10	PG4	A	1013	-	12,12,12	0.20	0	11,11,11	0.13	0
7	SO4	A	1004	-	4,4,4	0.38	0	6,6,6	0.38	0
11	EDO	A	1021	-	3,3,3	0.24	0	2,2,2	0.22	0
11	EDO	A	1018	-	3,3,3	0.76	0	2,2,2	0.52	0
11	EDO	A	1015	-	3,3,3	0.46	0	2,2,2	1.17	0
10	PG4	A	1012	-	12,12,12	0.36	0	11,11,11	0.33	0
11	EDO	A	1023	-	3,3,3	0.17	0	2,2,2	0.19	0
6	A1IO6	A	1001	-	16,20,20	0.74	0	22,28,28	1.07	2 (9%)
11	EDO	A	1014	-	3,3,3	0.24	0	2,2,2	0.20	0
11	EDO	A	1017	-	3,3,3	0.57	0	2,2,2	0.21	0
11	EDO	A	1019	-	3,3,3	0.20	0	2,2,2	0.26	0
11	EDO	A	1022	-	3,3,3	0.40	0	2,2,2	0.74	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
11	EDO	A	1016	-	-	1/1/1/1	-
9	PGE	A	1010	-	-	3/7/7/7	-
9	PGE	A	1011	-	-	4/7/7/7	-
11	EDO	A	1020	-	-	0/1/1/1	-
10	PG4	A	1013	-	-	1/10/10/10	-
11	EDO	A	1021	-	-	1/1/1/1	-

*Continued on next page...*



*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
11	EDO	A	1018	-	-	1/1/1/1	-
11	EDO	A	1015	-	-	0/1/1/1	-
10	PG4	A	1012	-	-	2/10/10/10	-
11	EDO	A	1023	-	-	0/1/1/1	-
6	A1IO6	A	1001	-	-	2/9/29/29	0/2/2/2
11	EDO	A	1014	-	-	0/1/1/1	-
11	EDO	A	1017	-	-	1/1/1/1	-
11	EDO	A	1019	-	-	1/1/1/1	-
11	EDO	A	1022	-	-	0/1/1/1	-

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
6	A	1001	A1IO6	C2-C3-C4	-2.48	107.70	110.48
6	A	1001	A1IO6	C8-C7-C5	-2.48	112.40	115.31

There are no chirality outliers.

All (17) torsion outliers are listed below:

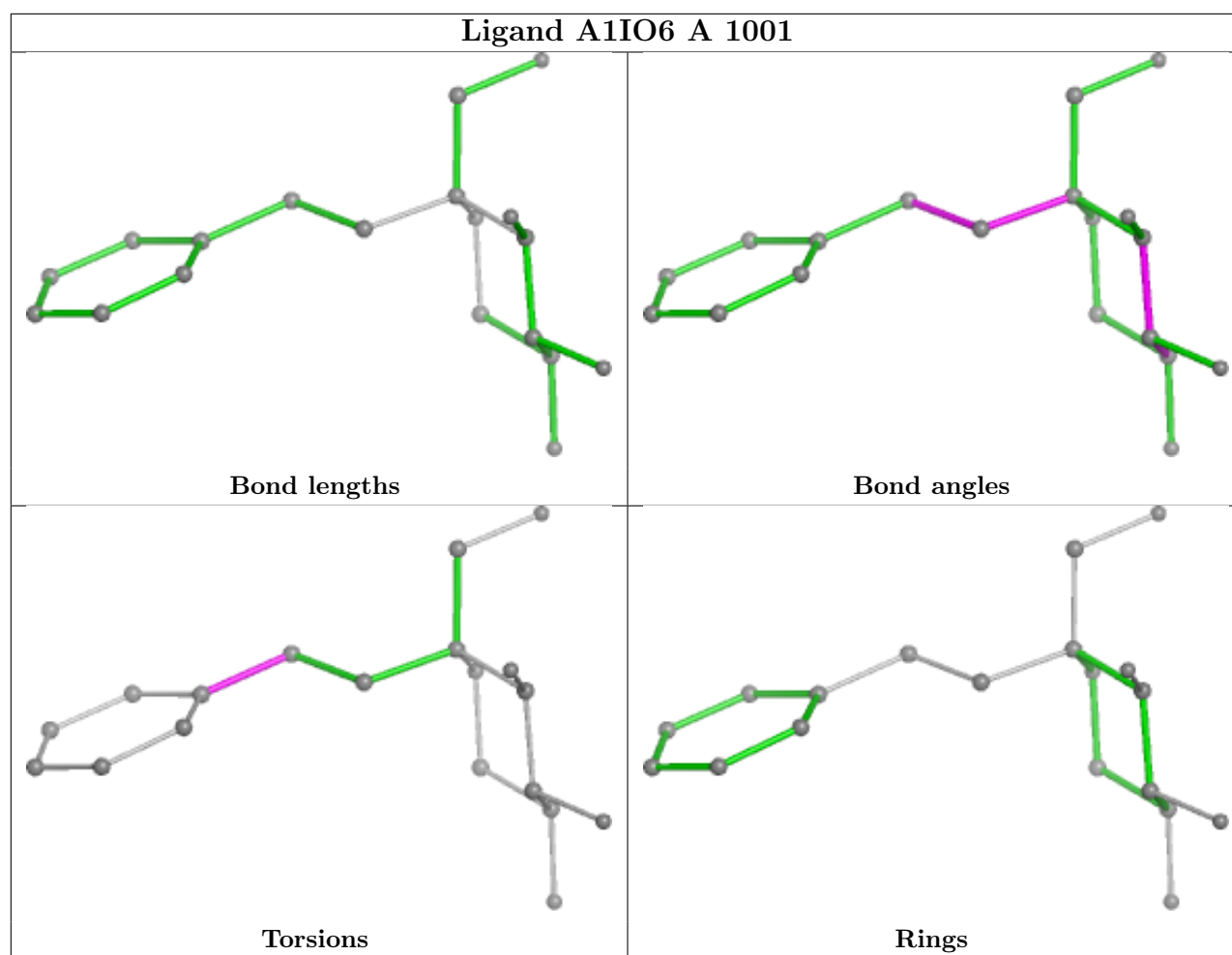
Mol	Chain	Res	Type	Atoms
9	A	1010	PGE	O1-C1-C2-O2
10	A	1013	PG4	C1-C2-O2-C3
11	A	1019	EDO	O1-C1-C2-O2
10	A	1012	PG4	O1-C1-C2-O2
9	A	1011	PGE	C1-C2-O2-C3
11	A	1017	EDO	O1-C1-C2-O2
11	A	1018	EDO	O1-C1-C2-O2
9	A	1011	PGE	C4-C3-O2-C2
9	A	1010	PGE	O2-C3-C4-O3
9	A	1011	PGE	C3-C4-O3-C5
6	A	1001	A1IO6	C7-C8-C9-C14
9	A	1010	PGE	C1-C2-O2-C3
6	A	1001	A1IO6	C7-C8-C9-C10
11	A	1016	EDO	O1-C1-C2-O2
11	A	1021	EDO	O1-C1-C2-O2
10	A	1012	PG4	O4-C7-C8-O5
9	A	1011	PGE	O2-C3-C4-O3

There are no ring outliers.

5 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
11	A	1016	EDO	2	0
9	A	1010	PGE	2	0
10	A	1013	PG4	1	0
11	A	1017	EDO	1	0
11	A	1022	EDO	1	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.



## 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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	841/952 (88%)	-0.10	20 (2%) 59 58	16, 29, 50, 76	4 (0%)

All (20) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	195	LEU	5.0
1	A	678	LEU	3.6
1	A	924	SER	3.6
1	A	780	ILE	3.6
1	A	895	GLU	3.1
1	A	115	GLN	3.1
1	A	98	ILE	3.0
1	A	144	SER	3.0
1	A	81	GLN	2.8
1	A	897	ALA	2.7
1	A	911	THR	2.6
1	A	146	MET	2.6
1	A	206	SER	2.6
1	A	923	VAL	2.5
1	A	931	ASP	2.4
1	A	147	GLY	2.4
1	A	194	PRO	2.3
1	A	211	VAL	2.2
1	A	903	LYS	2.1
1	A	533	CYS	2.1

### 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<sup>th</sup> 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( $\text{\AA}^2$ )	Q<0.9
1	CSO	A	938	7/8	0.93	0.10	32,37,49,66	0

### 6.3 Carbohydrates

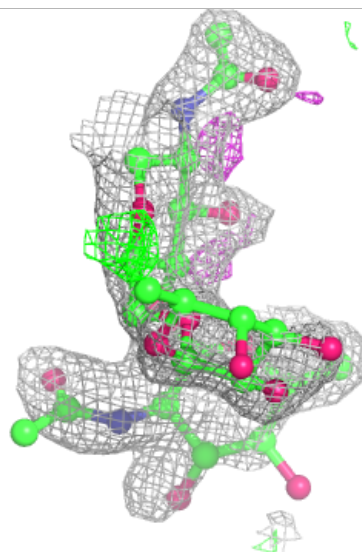
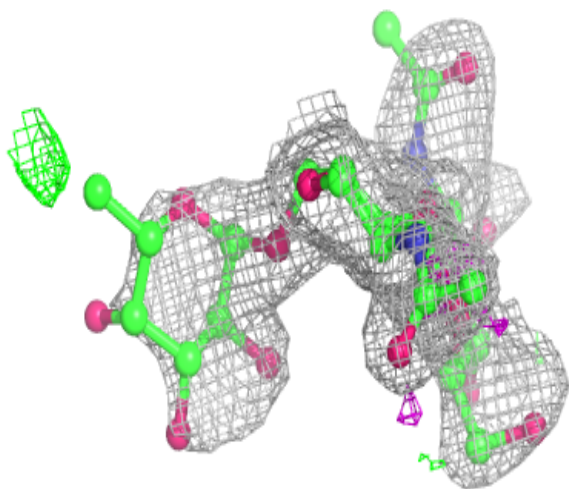
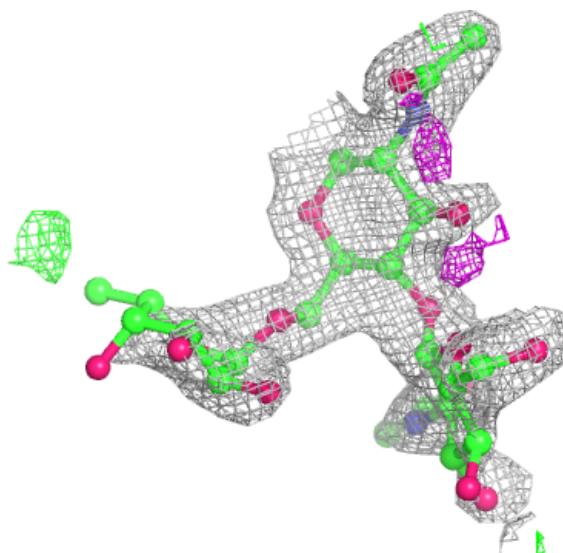
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
3	BMA	C	3	11/12	0.44	0.15	83,102,122,123	0
5	FUC	F	2	10/11	0.67	0.19	80,100,114,120	0
4	NAG	D	2	14/15	0.70	0.15	67,88,99,112	0
2	NAG	B	2	14/15	0.70	0.14	59,87,117,119	0
3	BMA	E	3	11/12	0.74	0.14	54,68,77,80	0
2	FUC	B	3	10/11	0.75	0.16	76,93,115,118	0
3	NAG	C	2	14/15	0.79	0.15	50,66,93,116	0
5	NAG	F	1	14/15	0.88	0.13	37,57,81,88	0
2	NAG	B	1	14/15	0.89	0.12	38,42,68,77	0
3	NAG	E	2	14/15	0.91	0.10	31,35,44,53	0
4	NAG	D	1	14/15	0.93	0.09	30,37,50,65	0
3	NAG	C	1	14/15	0.93	0.09	31,37,42,47	0
3	NAG	E	1	14/15	0.97	0.06	24,28,40,40	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.

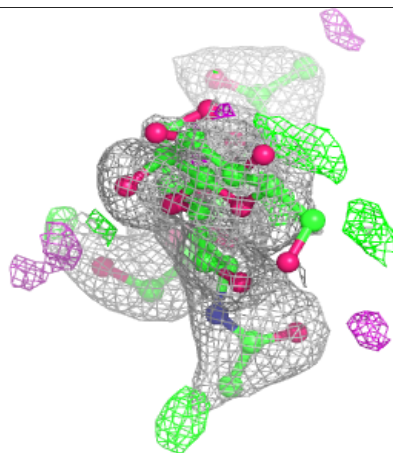
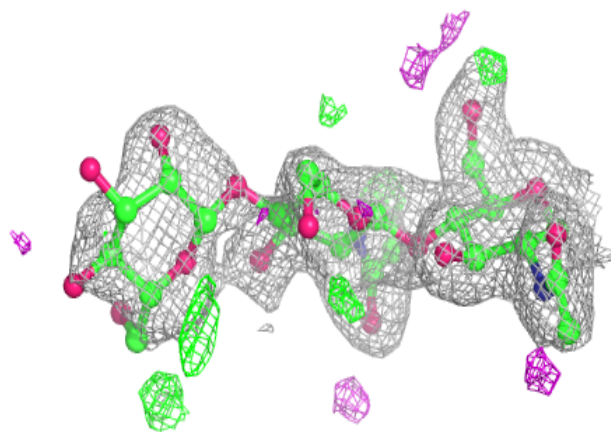
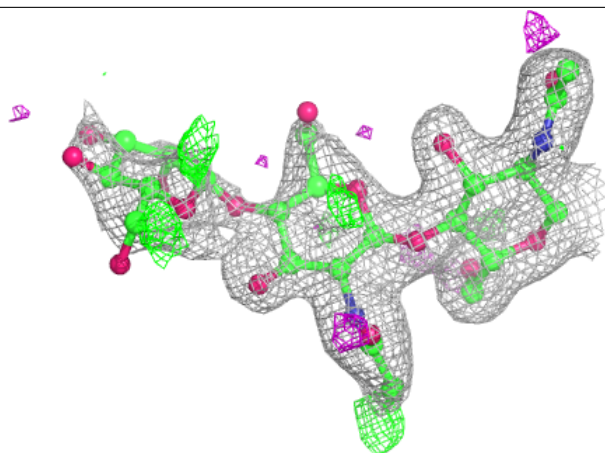
**Electron density around Chain B:**

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



**Electron density around Chain C:**

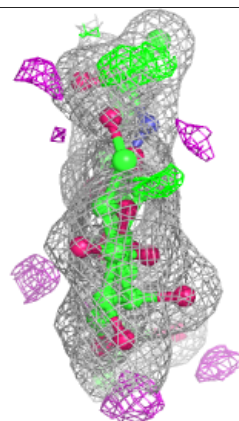
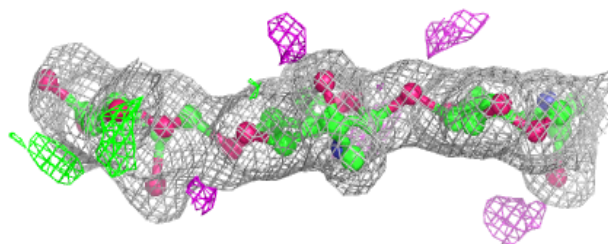
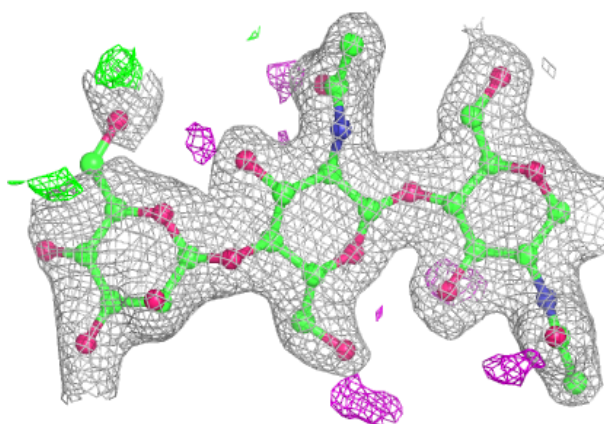
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



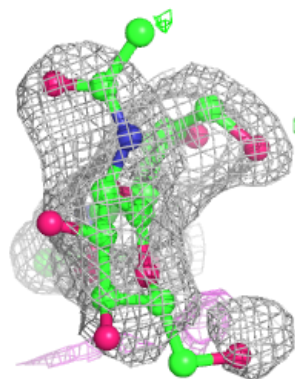
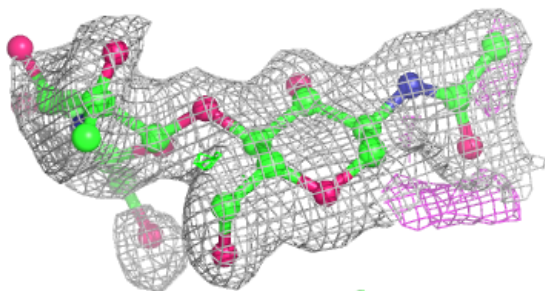
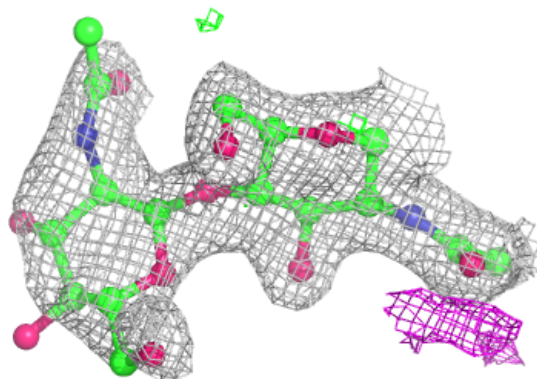


**Electron density around Chain E:**

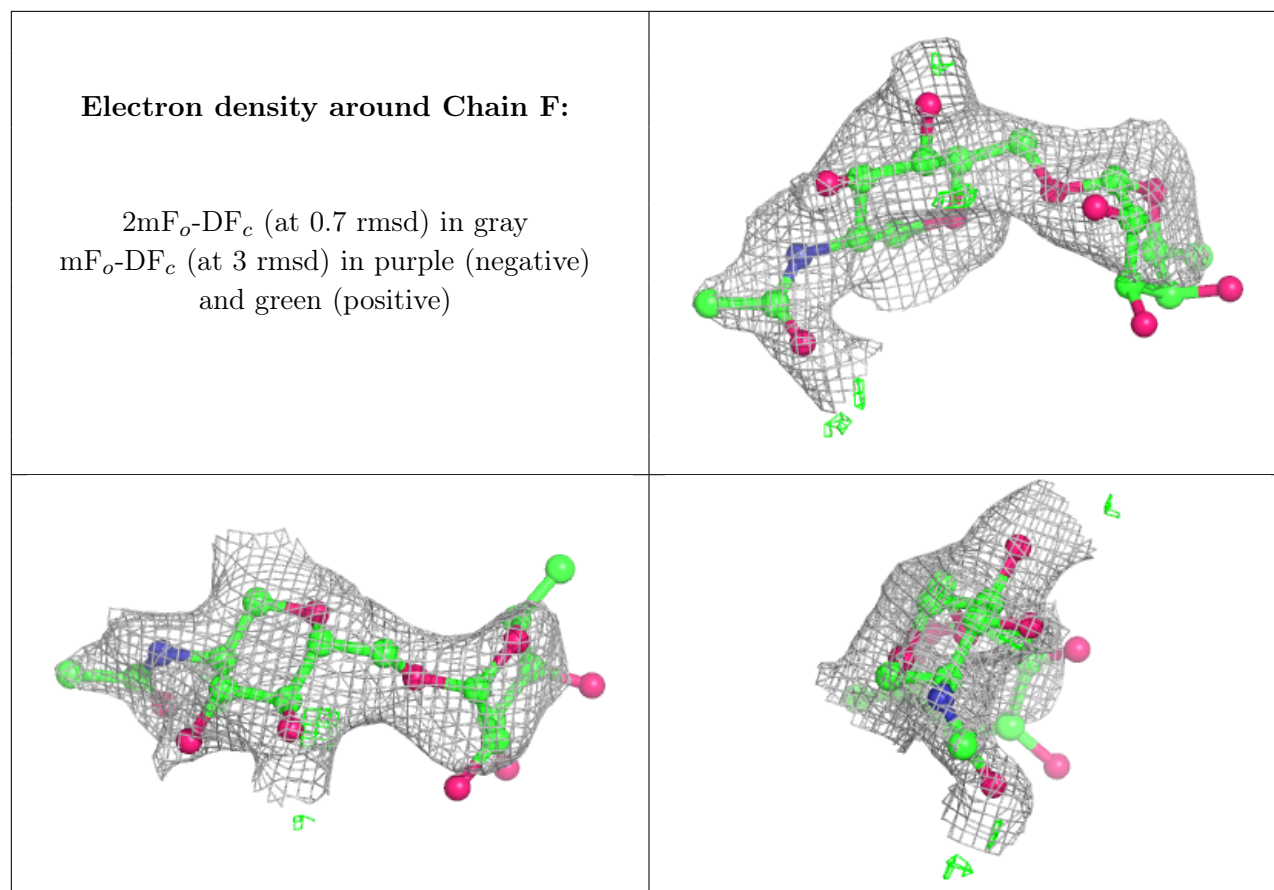
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)

**Electron density around Chain D:**

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







## 6.4 Ligands [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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
11	EDO	A	1015	4/4	0.70	0.21	36,40,44,46	0
9	PGE	A	1011	10/10	0.72	0.24	62,69,76,81	0
11	EDO	A	1019	4/4	0.73	0.19	43,48,52,54	0
11	EDO	A	1018	4/4	0.78	0.17	32,34,34,41	0
11	EDO	A	1020	4/4	0.79	0.18	44,47,48,56	0
11	EDO	A	1017	4/4	0.81	0.18	39,43,48,49	0
8	CL	A	1007	1/1	0.81	0.16	62,62,62,62	0
10	PG4	A	1012	13/13	0.84	0.21	54,69,80,82	0
11	EDO	A	1016	4/4	0.84	0.15	39,40,43,56	0
9	PGE	A	1010	10/10	0.86	0.19	52,66,72,73	0
11	EDO	A	1021	4/4	0.86	0.14	45,46,48,51	0
11	EDO	A	1014	4/4	0.87	0.15	34,38,46,52	0
7	SO4	A	1003	5/5	0.88	0.08	55,60,82,84	0

*Continued on next page...*

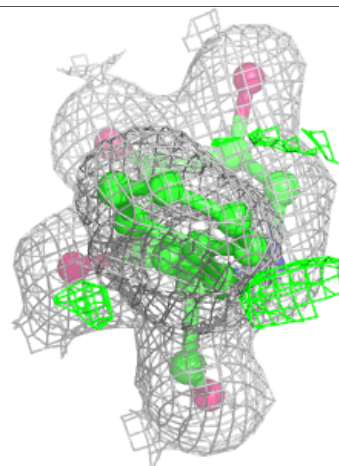
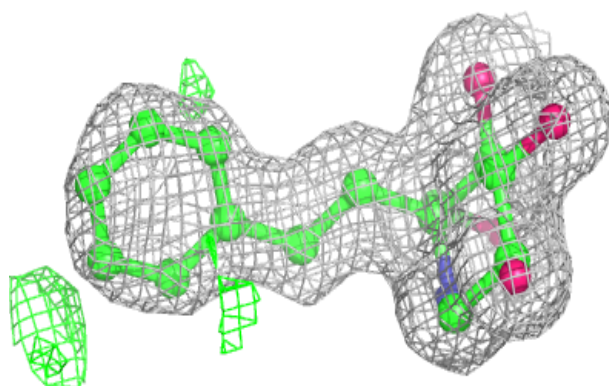
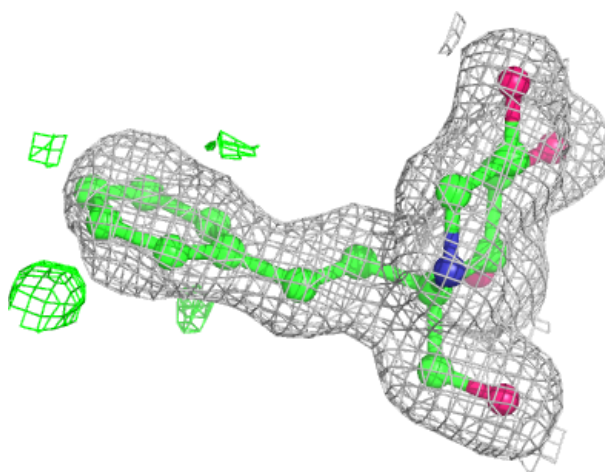
Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
10	PG4	A	1013	13/13	0.88	0.15	51,63,83,85	0
8	CL	A	1005	1/1	0.89	0.18	65,65,65,65	0
11	EDO	A	1023	4/4	0.89	0.14	42,42,42,46	0
8	CL	A	1008	1/1	0.91	0.13	57,57,57,57	0
11	EDO	A	1022	4/4	0.92	0.11	40,41,48,48	0
7	SO4	A	1002	5/5	0.93	0.09	44,48,52,52	0
8	CL	A	1009	1/1	0.93	0.12	61,61,61,61	0
7	SO4	A	1004	5/5	0.96	0.08	25,33,39,39	5
8	CL	A	1006	1/1	0.97	0.09	45,45,45,45	0
6	A1IO6	A	1001	19/19	0.97	0.08	19,22,51,66	0

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

**Electron density around A1IO6 A 1001:**

2mF<sub>o</sub>-DF<sub>c</sub> (at 0.7 rmsd) in gray  
mF<sub>o</sub>-DF<sub>c</sub> (at 3 rmsd) in purple (negative)  
and green (positive)



## 6.5 Other polymers [i](#)

There are no such residues in this entry.