



# Full wwPDB X-ray Structure Validation Report ⓘ

Sep 23, 2025 – 01:41 am BST

PDB ID : 9GTW / pdb\_00009gtw  
Title : Crystal structure of human lysosomal acid-alpha-glucosidase, GAA, in complex with iminosugar compound 3g  
Authors : Sulzenbacher, G.; Roig-Zamboni, V.; Moracci, M.; Parenti, G.; Py, S.  
Deposited on : 2024-09-18  
Resolution : 1.75 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 Phenix2.0
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.46

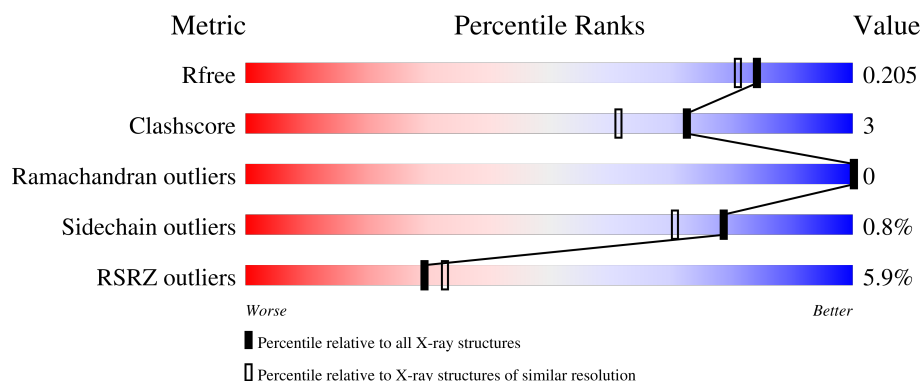
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.75 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	2888 (1.76-1.76)
Clashscore	180529	3097 (1.76-1.76)
Ramachandran outliers	177936	3072 (1.76-1.76)
Sidechain outliers	177891	3072 (1.76-1.76)
RSRZ outliers	164620	2887 (1.76-1.76)

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	<div> <div>5%</div> <div> <div></div> <div>82%</div> <div>6%</div> <div>11%</div> </div> </div>
2	B	3	<div> <div>33%</div> <div>67%</div> </div>
3	C	2	<div> <div>50%</div> <div>50%</div> </div>
4	D	3	<div> <div>33%</div> <div>67%</div> </div>

## 2 Entry composition [i](#)

There are 13 unique types of molecules in this entry. The entry contains 7651 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.

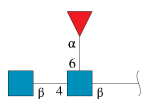
- Molecule 1 is a protein called Lysosomal alpha-glucosidase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	843	Total	C	N	O	S	0	8	0
			6675	4283	1124	1237	31			

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.



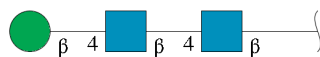
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	B	3	Total	C	N	O	0	0	0
			38	22	2	14			

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



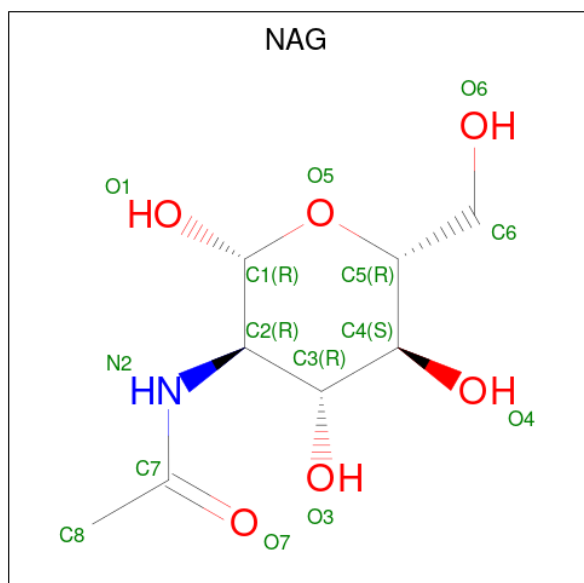
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	C	2	Total	C	N	O	0	0	0
			28	16	2	10			

- Molecule 4 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.



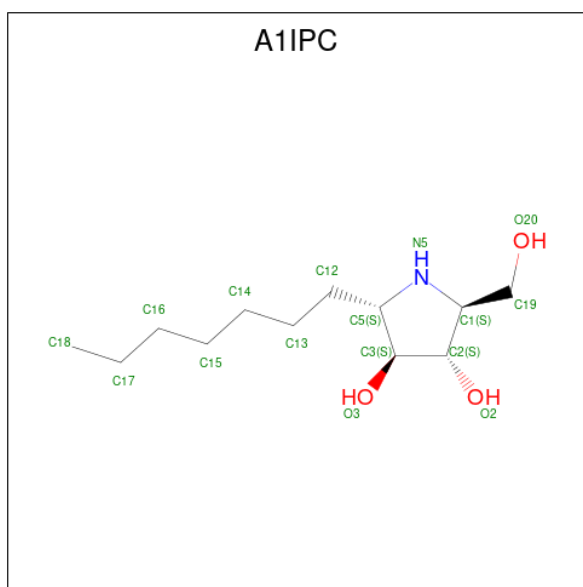
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
4	D	3	Total	C	N	O	0	0	0
			39	22	2	15			

- Molecule 5 is 2-acetamido-2-deoxy-beta-D-glucopyranose (CCD ID: NAG) (formula:  $C_8H_{15}NO_6$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	1	Total	C	N	O	0	0
			14	8	1	5		
5	A	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 6 is (2 {S},3 {S},4 {S},5 {S})-2-heptyl-5-(hydroxymethyl)pyrrolidine-3,4-diol (CCD ID: A1IPC) (formula:  $C_{12}H_{25}NO_3$ ) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	A	1	Total	C	N	O	0	0
			16	12	1	3		

- 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		

*Continued on next page...*

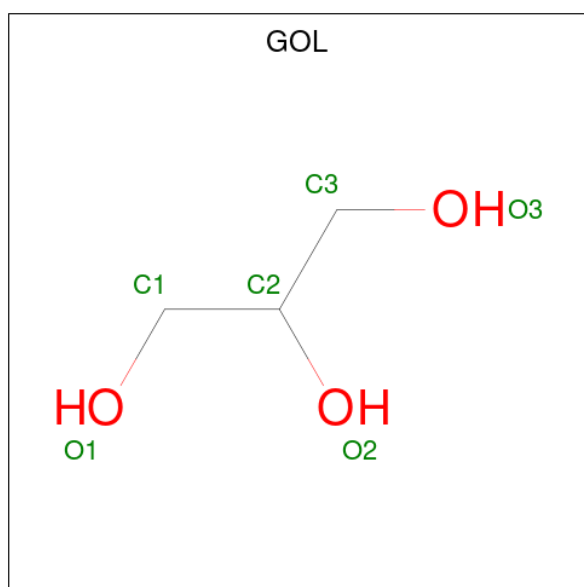
Continued from previous page...

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		

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

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	3	Total	Cl	0	0
			3	3		

- Molecule 9 is GLYCEROL (CCD ID: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



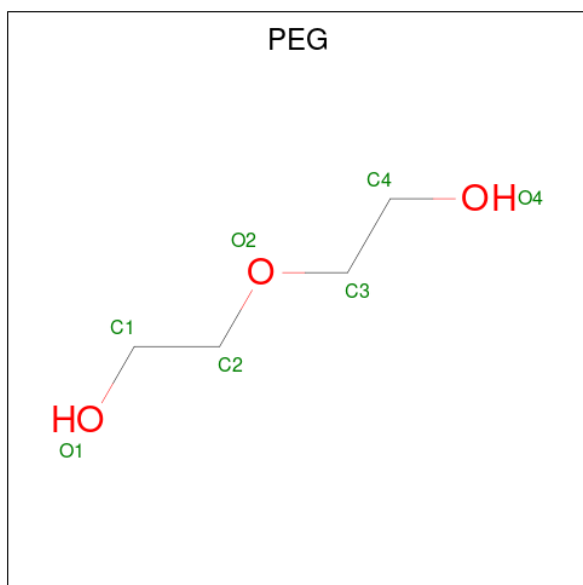
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
9	A	1	Total	C	O	0	0
			6	3	3		
9	A	1	Total	C	O	0	0
			6	3	3		

- Molecule 10 is TRIETHYLENE GLYCOL (CCD ID: PGE) (formula: C<sub>6</sub>H<sub>14</sub>O<sub>4</sub>).



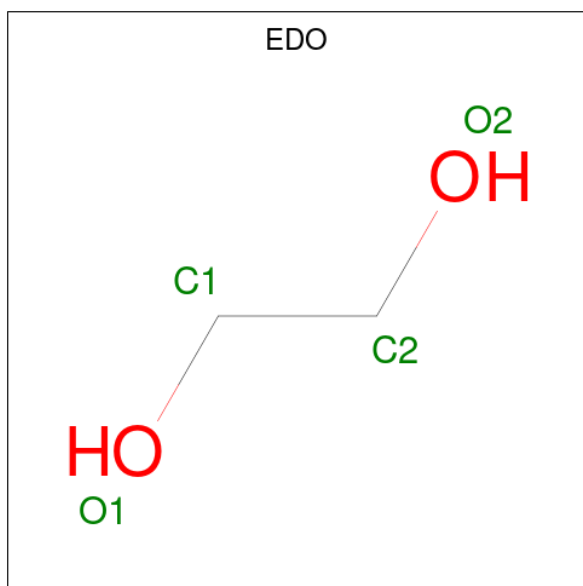
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
10	A	1	Total	C	O	0	0
			10	6	4		
10	A	1	Total	C	O	0	0
			10	6	4		
10	A	1	Total	C	O	0	0
			10	6	4		
10	A	1	Total	C	O	0	0
			10	6	4		

- Molecule 11 is DI(HYDROXYETHYL)ETHER (CCD ID: PEG) (formula:  $C_4H_{10}O_3$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
11	A	1	Total	C	O	0	0
			7	4	3		

- Molecule 12 is 1,2-ETHANEDIOL (CCD ID: EDO) (formula:  $C_2H_6O_2$ ).



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

*Continued on next page...*



*Continued from previous page...*

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

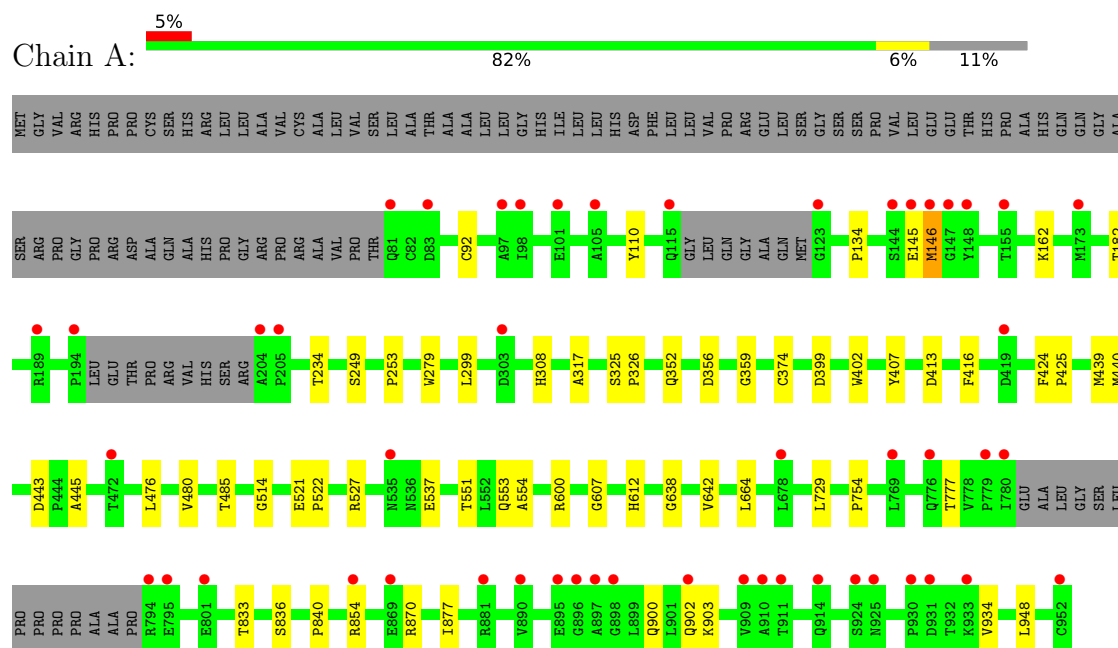
- Molecule 13 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
13	A	692	Total	O	0	0
			692	692		

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



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

Chain D:  33% 67%

MAG1  
MAG2  
BOL3

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	96.95Å 101.94Å 128.66Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	77.55 – 1.75 77.55 – 1.75	Depositor EDS
% Data completeness (in resolution range)	99.8 (77.55-1.75) 99.9 (77.55-1.75)	Depositor EDS
$R_{merge}$	0.18	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.36 (at 1.75Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, $R_{free}$	0.164 , 0.192 0.178 , 0.205	Depositor DCC
$R_{free}$ test set	6432 reflections (4.16%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	20.9	Xtriage
Anisotropy	0.110	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 40.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	0.011 for k,h,-l	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	7651	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	34.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.70% 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 [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: PEG, SO4, GOL, PGE, A1IPC, CL, FUC, BMA, EDO, NAG, CSO

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	0.93	0/6890	1.09	0/9414

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	A	6675	0	6462	41	0
2	B	38	0	34	0	0
3	C	28	0	25	0	0
4	D	39	0	34	0	0
5	A	28	0	26	0	0
6	A	16	0	0	0	0
7	A	25	0	0	0	0
8	A	3	0	0	0	0
9	A	12	0	16	0	0
10	A	40	0	56	2	0
11	A	7	0	10	1	0
12	A	48	0	72	6	0
13	A	692	0	0	3	0
All	All	7651	0	6735	42	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.

All (42) 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:182:THR:HG21	11:A:1017:PEG:H21	1.60	0.82
1:A:840:PRO:HB3	1:A:902:GLN:HG3	1.72	0.72
1:A:900:GLN:HE22	10:A:1016:PGE:H42	1.56	0.70
1:A:134:PRO:CG	12:A:1029:EDO:H12	2.24	0.67
1:A:134:PRO:HG2	12:A:1029:EDO:H12	1.75	0.67
1:A:900:GLN:HE22	10:A:1016:PGE:C4	2.12	0.63
1:A:551[A]:THR:OG1	1:A:553[A]:GLN:OE1	2.17	0.63
12:A:1023:EDO:H22	13:A:1199:HOH:O	2.00	0.61
1:A:551[A]:THR:HG23	1:A:554:ALA:H	1.71	0.55
1:A:854[B]:ARG:NH2	1:A:877:ILE:HD13	2.25	0.52
1:A:521:GLU:N	1:A:522:PRO:HA	2.25	0.51
1:A:402:TRP:HA	1:A:439:MET:O	2.12	0.50
1:A:325:SER:N	1:A:326:PRO:HA	2.27	0.49
1:A:374:CYS:HB2	1:A:402:TRP:HB2	1.94	0.49
1:A:527[B]:ARG:NH2	1:A:553[B]:GLN:OE1	2.46	0.49
1:A:607:GLY:HA3	1:A:638:GLY:O	2.14	0.48
1:A:234:THR:HA	1:A:249:SER:O	2.14	0.47
1:A:279:TRP:CE3	1:A:317:ALA:HB2	2.50	0.47
1:A:424:PHE:HB3	1:A:425:PRO:HD3	1.97	0.47
1:A:145:GLU:O	1:A:146:MET:HB2	2.14	0.47
1:A:352:GLN:NE2	1:A:356[A]:ASP:OD2	2.49	0.46
1:A:664:LEU:HD13	1:A:754:PRO:HG3	1.97	0.46
1:A:162:LYS:NZ	1:A:537:GLU:OE1	2.49	0.45
1:A:407:TYR:O	1:A:416:PHE:HA	2.18	0.44
1:A:443:ASP:HB3	13:A:1642:HOH:O	2.16	0.44
1:A:612:HIS:O	1:A:642:VAL:HA	2.16	0.44
1:A:299:LEU:C	1:A:299:LEU:HD23	2.43	0.44
1:A:439:MET:SD	1:A:514:GLY:HA3	2.58	0.43
1:A:445:ALA:HB1	1:A:485:THR:HB	2.00	0.43
1:A:948:LEU:C	1:A:948:LEU:HD23	2.43	0.43
1:A:359:GLY:HA2	12:A:1019:EDO:H22	2.01	0.43
1:A:134:PRO:CD	12:A:1029:EDO:H12	2.48	0.42
1:A:833:THR:HA	1:A:836:SER:OG	2.20	0.41
1:A:600:ARG:O	1:A:612:HIS:CE1	2.74	0.41
1:A:146:MET:HE3	1:A:146:MET:HB3	1.88	0.41
1:A:777:THR:HB	1:A:934:VAL:HG21	2.03	0.41
1:A:870:ARG:HD2	13:A:1143:HOH:O	2.21	0.41

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:413:ASP:OD2	1:A:443:ASP:O	2.39	0.41
1:A:253:PRO:HG3	1:A:308:HIS:HE1	1.86	0.40
1:A:476:LEU:C	1:A:476:LEU:HD23	2.46	0.40
1:A:92:CYS:O	1:A:92:CYS:SG	2.79	0.40
1:A:110:TYR:O	12:A:1025:EDO:H11	2.21	0.40

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	A	842/952 (88%)	818 (97%)	24 (3%)	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	A	727/805 (90%)	721 (99%)	6 (1%)	79	71

All (6) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	146	MET

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	399	ASP
1	A	440	MET
1	A	480	VAL
1	A	729	LEU
1	A	903	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (8) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	177	ASN
1	A	188	ASN
1	A	352	GLN
1	A	401	GLN
1	A	540	ASN
1	A	584	HIS
1	A	692	GLN
1	A	900	GLN

### 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	0.94	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 [i](#)

8 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	1,2	14,14,15	0.61	0	17,19,21	1.05	1 (5%)
2	NAG	B	2	2	14,14,15	0.38	0	17,19,21	1.28	2 (11%)
2	FUC	B	3	2	10,10,11	0.32	0	14,14,16	0.78	0
3	NAG	C	1	1,3	14,14,15	0.65	1 (7%)	17,19,21	1.35	3 (17%)
3	NAG	C	2	3	14,14,15	0.36	0	17,19,21	0.84	0
4	NAG	D	1	1,4	14,14,15	0.63	0	17,19,21	1.07	2 (11%)
4	NAG	D	2	4	14,14,15	0.50	0	17,19,21	1.22	1 (5%)
4	BMA	D	3	4	11,11,12	0.28	0	15,15,17	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
2	NAG	B	1	1,2	-	0/6/23/26	0/1/1/1
2	NAG	B	2	2	-	2/6/23/26	0/1/1/1
2	FUC	B	3	2	-	-	0/1/1/1
3	NAG	C	1	1,3	-	0/6/23/26	0/1/1/1
3	NAG	C	2	3	-	0/6/23/26	0/1/1/1
4	NAG	D	1	1,4	-	0/6/23/26	0/1/1/1
4	NAG	D	2	4	-	0/6/23/26	0/1/1/1
4	BMA	D	3	4	-	0/2/19/22	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	1	NAG	C1-C2	2.01	1.55	1.52

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	1	NAG	O5-C1-C2	-2.93	106.66	111.29
3	C	1	NAG	O5-C1-C2	-2.88	106.75	111.29
2	B	2	NAG	C1-O5-C5	-2.86	108.31	112.19
4	D	2	NAG	O7-C7-N2	-2.64	117.10	121.95
4	D	1	NAG	O5-C1-C2	-2.39	107.52	111.29
3	C	1	NAG	O5-C5-C6	2.29	110.79	107.20
3	C	1	NAG	C1-O5-C5	2.26	115.25	112.19
4	D	1	NAG	C1-O5-C5	2.16	115.11	112.19
2	B	2	NAG	C4-C3-C2	2.10	114.09	111.02

There are no chirality outliers.

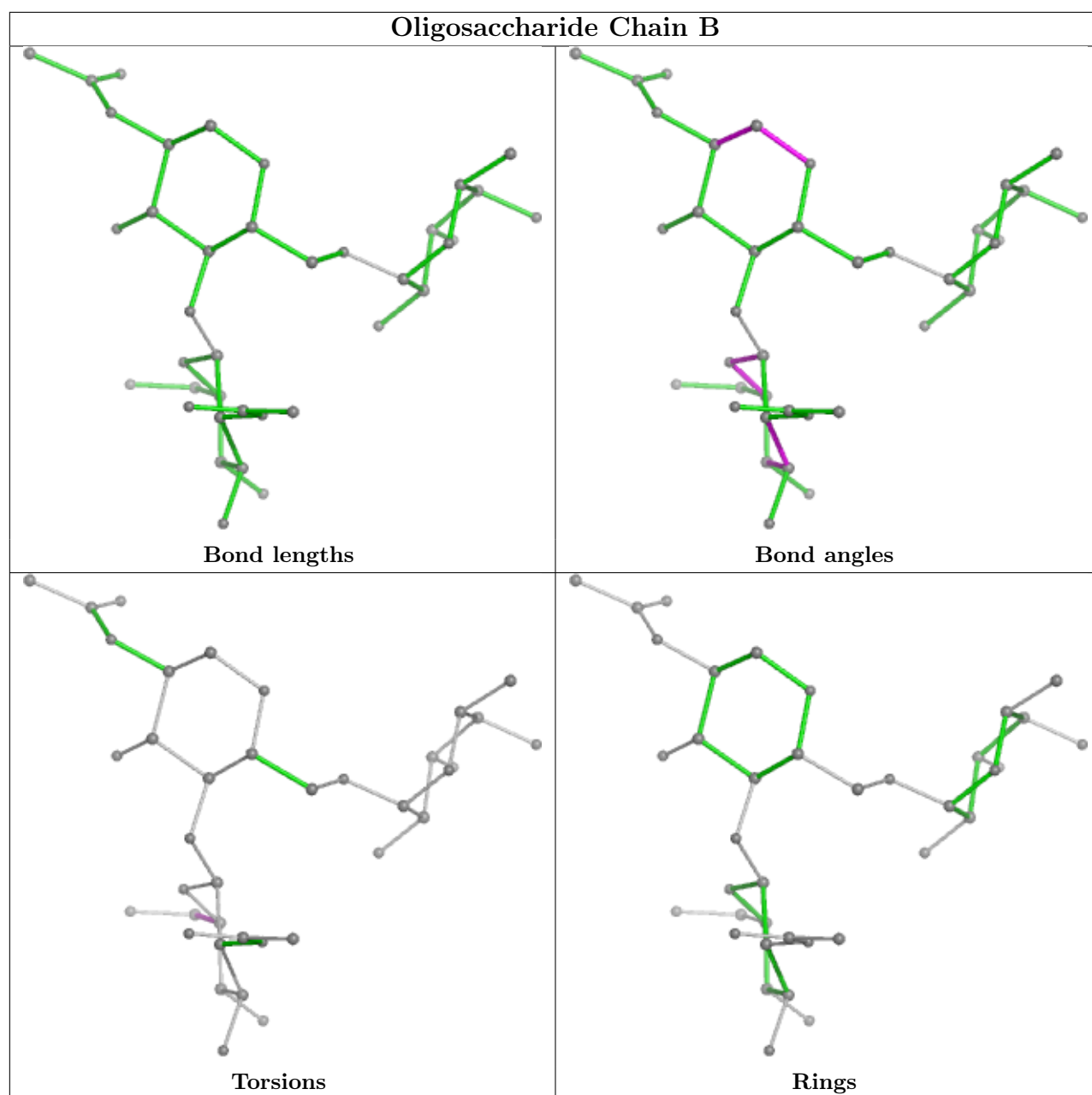
All (2) torsion outliers are listed below:

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

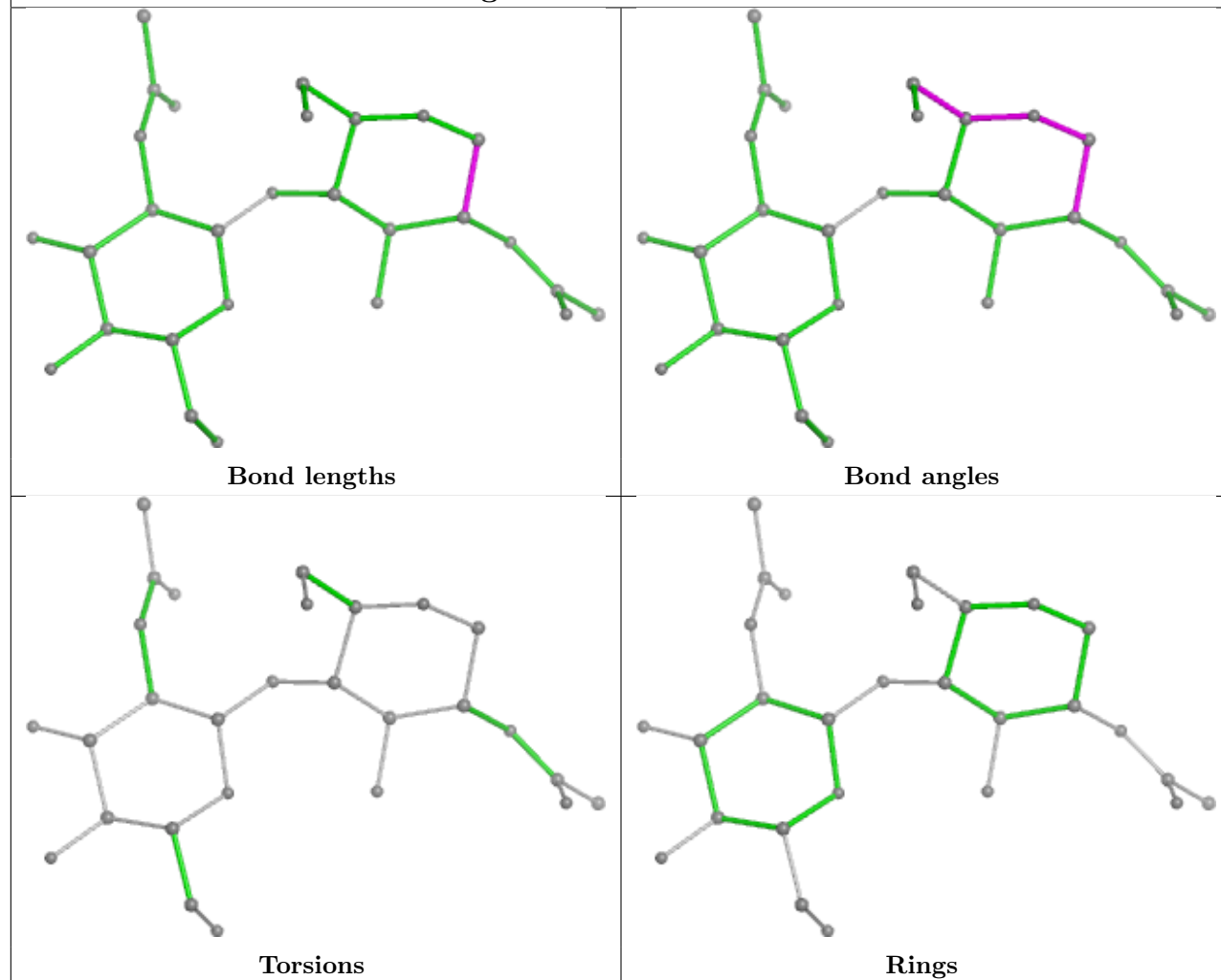
There are no ring outliers.

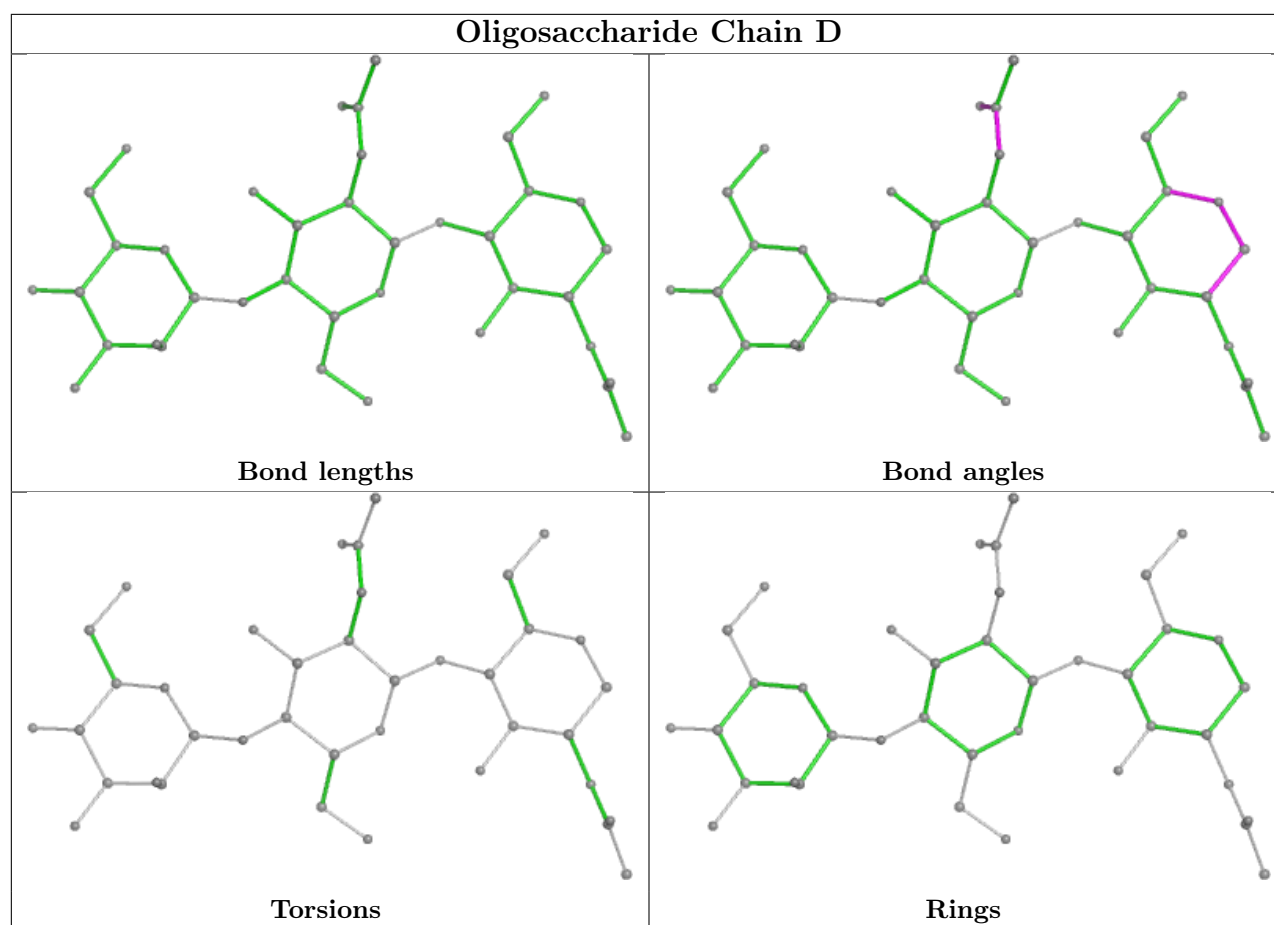
No monomer is involved in short contacts.

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



## Oligosaccharide Chain C





## 5.6 Ligand geometry [i](#)

Of 30 ligands modelled in this entry, 3 are monoatomic - leaving 27 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	1005	-	4,4,4	0.29	0	6,6,6	0.12	0
12	EDO	A	1022	-	3,3,3	0.07	0	2,2,2	0.23	0
10	PGE	A	1015	-	9,9,9	0.16	0	8,8,8	0.13	0
11	PEG	A	1017	-	6,6,6	0.22	0	5,5,5	0.19	0
12	EDO	A	1029	-	3,3,3	0.15	0	2,2,2	0.25	0
7	SO4	A	1007	-	4,4,4	0.38	0	6,6,6	0.09	0
6	A1IPC	A	1003	-	16,16,16	0.81	0	15,20,20	1.07	1 (6%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
9	GOL	A	1013	-	5,5,5	0.13	0	5,5,5	0.42	0
12	EDO	A	1019	-	3,3,3	0.03	0	2,2,2	0.24	0
12	EDO	A	1030	-	3,3,3	0.18	0	2,2,2	0.28	0
12	EDO	A	1020	-	3,3,3	0.17	0	2,2,2	0.38	0
12	EDO	A	1028	-	3,3,3	0.08	0	2,2,2	0.21	0
5	NAG	A	1001	1	14,14,15	0.60	0	17,19,21	1.08	1 (5%)
12	EDO	A	1027	-	3,3,3	0.12	0	2,2,2	0.27	0
9	GOL	A	1012	-	5,5,5	0.09	0	5,5,5	0.17	0
10	PGE	A	1014	-	9,9,9	0.20	0	8,8,8	0.18	0
7	SO4	A	1008	-	4,4,4	0.24	0	6,6,6	0.35	0
5	NAG	A	1002	1	14,14,15	0.44	0	17,19,21	1.16	1 (5%)
12	EDO	A	1026	-	3,3,3	0.08	0	2,2,2	0.16	0
10	PGE	A	1018	-	9,9,9	0.27	0	8,8,8	0.20	0
12	EDO	A	1024	-	3,3,3	0.14	0	2,2,2	0.03	0
10	PGE	A	1016	-	9,9,9	0.15	0	8,8,8	0.14	0
7	SO4	A	1006	-	4,4,4	0.37	0	6,6,6	0.27	0
12	EDO	A	1021	-	3,3,3	0.15	0	2,2,2	0.28	0
12	EDO	A	1025	-	3,3,3	0.13	0	2,2,2	0.17	0
12	EDO	A	1023	-	3,3,3	0.30	0	2,2,2	0.22	0
7	SO4	A	1004	-	4,4,4	0.34	0	6,6,6	0.21	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
12	EDO	A	1022	-	-	0/1/1/1	-
10	PGE	A	1015	-	-	0/7/7/7	-
11	PEG	A	1017	-	-	0/4/4/4	-
12	EDO	A	1029	-	-	1/1/1/1	-
6	A1IPC	A	1003	-	-	1/9/25/25	0/1/1/1
9	GOL	A	1013	-	-	2/4/4/4	-
12	EDO	A	1019	-	-	0/1/1/1	-
12	EDO	A	1030	-	-	0/1/1/1	-
12	EDO	A	1020	-	-	0/1/1/1	-
12	EDO	A	1028	-	-	0/1/1/1	-
5	NAG	A	1001	1	-	0/6/23/26	0/1/1/1
12	EDO	A	1027	-	-	1/1/1/1	-
9	GOL	A	1012	-	-	0/4/4/4	-
10	PGE	A	1014	-	-	1/7/7/7	-
5	NAG	A	1002	1	-	0/6/23/26	0/1/1/1

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
12	EDO	A	1026	-	-	1/1/1/1	-
10	PGE	A	1018	-	-	3/7/7/7	-
12	EDO	A	1024	-	-	0/1/1/1	-
10	PGE	A	1016	-	-	2/7/7/7	-
12	EDO	A	1021	-	-	0/1/1/1	-
12	EDO	A	1025	-	-	0/1/1/1	-
12	EDO	A	1023	-	-	0/1/1/1	-

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	1001	NAG	O5-C1-C2	-2.59	107.20	111.29
6	A	1003	A1IPC	C3-C2-C1	2.53	106.55	102.55
5	A	1002	NAG	C1-O5-C5	2.39	115.43	112.19

There are no chirality outliers.

All (12) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
9	A	1013	GOL	O1-C1-C2-O2
9	A	1013	GOL	O1-C1-C2-C3
10	A	1018	PGE	O2-C3-C4-O3
10	A	1018	PGE	O1-C1-C2-O2
6	A	1003	A1IPC	C14-C15-C16-C17
10	A	1014	PGE	O2-C3-C4-O3
10	A	1016	PGE	O2-C3-C4-O3
10	A	1018	PGE	O3-C5-C6-O4
12	A	1027	EDO	O1-C1-C2-O2
12	A	1029	EDO	O1-C1-C2-O2
12	A	1026	EDO	O1-C1-C2-O2
10	A	1016	PGE	C1-C2-O2-C3

There are no ring outliers.

6 monomers are involved in 9 short contacts:

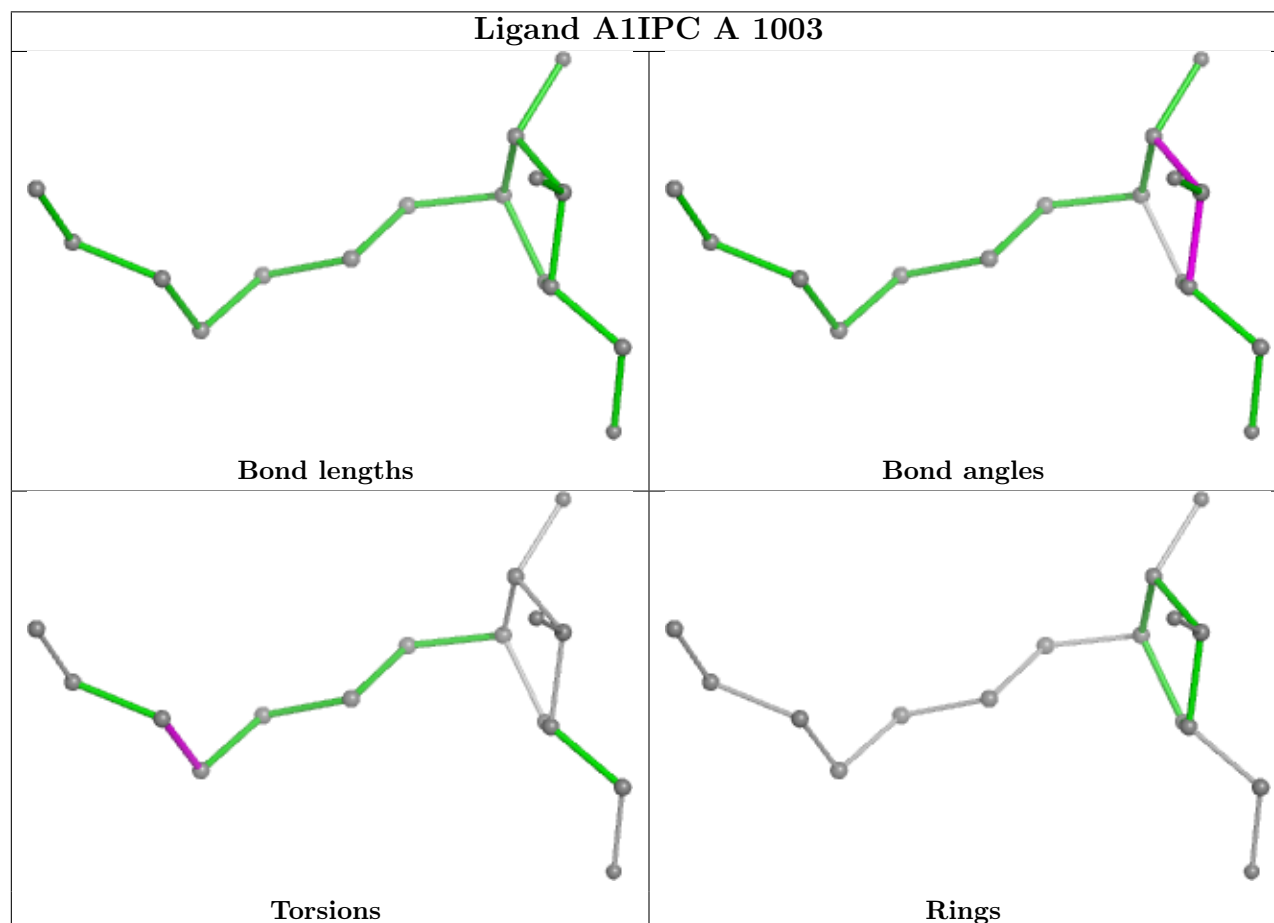
Mol	Chain	Res	Type	Clashes	Symm-Clashes
11	A	1017	PEG	1	0
12	A	1029	EDO	3	0
12	A	1019	EDO	1	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Clashes	Symm-Clashes
10	A	1016	PGE	2	0
12	A	1025	EDO	1	0
12	A	1023	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

### 6.1 Protein, DNA and RNA chains

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 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	842/952 (88%)	0.36	50 (5%)	29 33	17, 30, 47, 89	8 (0%)

All (50) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	205	PRO	5.7
1	A	204	ALA	5.6
1	A	81	GLN	5.0
1	A	780	ILE	4.9
1	A	146	MET	4.7
1	A	145	GLU	4.6
1	A	897	ALA	4.5
1	A	98	ILE	3.8
1	A	896	GLY	3.4
1	A	144	SER	3.4
1	A	194	PRO	3.3
1	A	779	PRO	3.3
1	A	189	ARG	3.2
1	A	911	THR	3.2
1	A	952	CYS	3.1
1	A	769	LEU	3.0
1	A	795	GLU	3.0
1	A	794	ARG	3.0
1	A	83	ASP	2.9
1	A	933	LYS	2.9
1	A	931	ASP	2.8
1	A	105	ALA	2.8
1	A	854[A]	ARG	2.7
1	A	147	GLY	2.7
1	A	115	GLN	2.6
1	A	678	LEU	2.6
1	A	925	ASN	2.6

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	898	GLY	2.6
1	A	419	ASP	2.5
1	A	910	ALA	2.5
1	A	173	MET	2.5
1	A	924	SER	2.4
1	A	902	GLN	2.4
1	A	303	ASP	2.4
1	A	930	PRO	2.3
1	A	155	THR	2.3
1	A	881	ARG	2.3
1	A	890	VAL	2.3
1	A	535	ASN	2.3
1	A	101	GLU	2.3
1	A	472	THR	2.2
1	A	776	GLN	2.2
1	A	895	GLU	2.2
1	A	97	ALA	2.2
1	A	123	GLY	2.1
1	A	909	VAL	2.1
1	A	914	GLN	2.1
1	A	148	TYR	2.1
1	A	801	GLU	2.1
1	A	869	GLU	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<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.

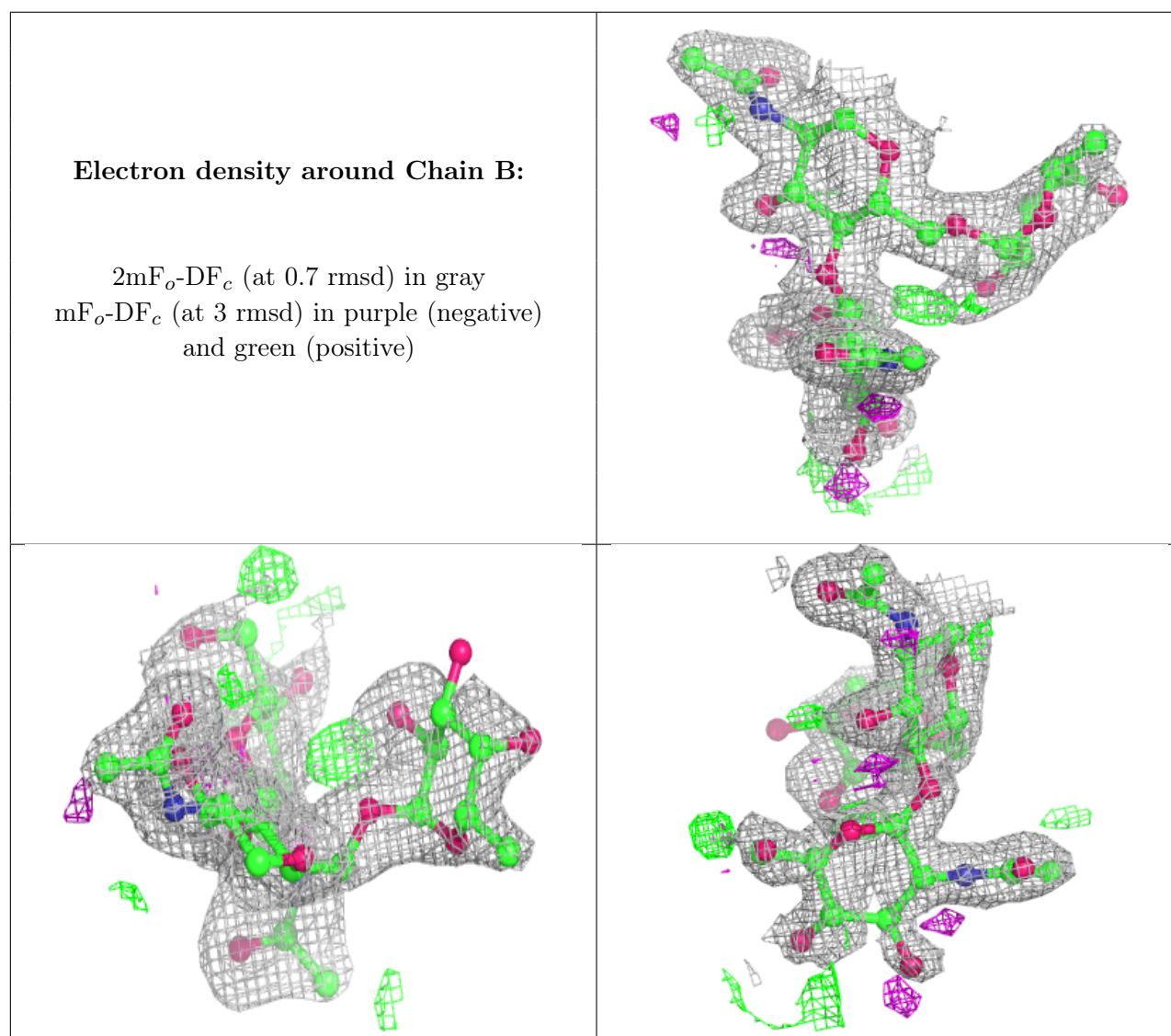
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
1	CSO	A	938	7/8	0.95	0.09	31,33,48,53	0

## 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<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.

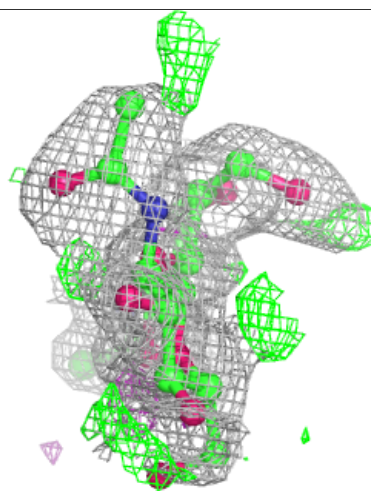
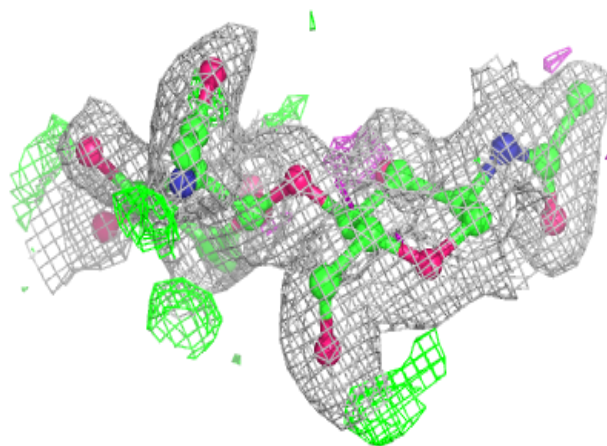
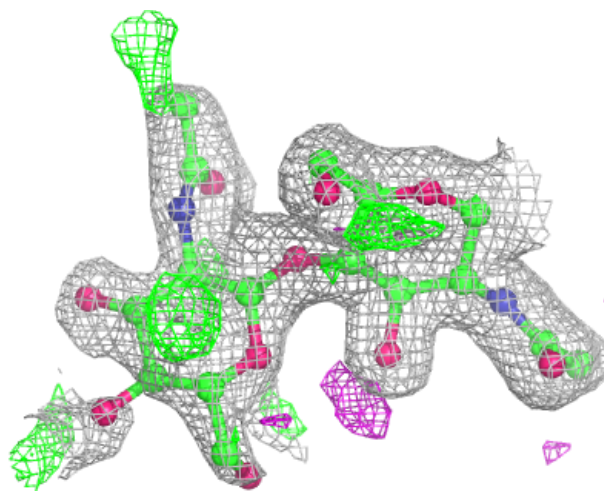
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
4	BMA	D	3	11/12	0.64	0.21	57,74,83,94	0
3	NAG	C	2	14/15	0.65	0.21	58,63,70,77	0
2	NAG	B	2	14/15	0.67	0.20	51,64,78,80	0
2	FUC	B	3	10/11	0.73	0.18	58,71,77,89	0
4	NAG	D	2	14/15	0.91	0.13	35,40,49,56	0
3	NAG	C	1	14/15	0.94	0.10	30,38,44,46	0
2	NAG	B	1	14/15	0.94	0.11	38,47,52,57	0
4	NAG	D	1	14/15	0.97	0.07	27,33,41,42	0

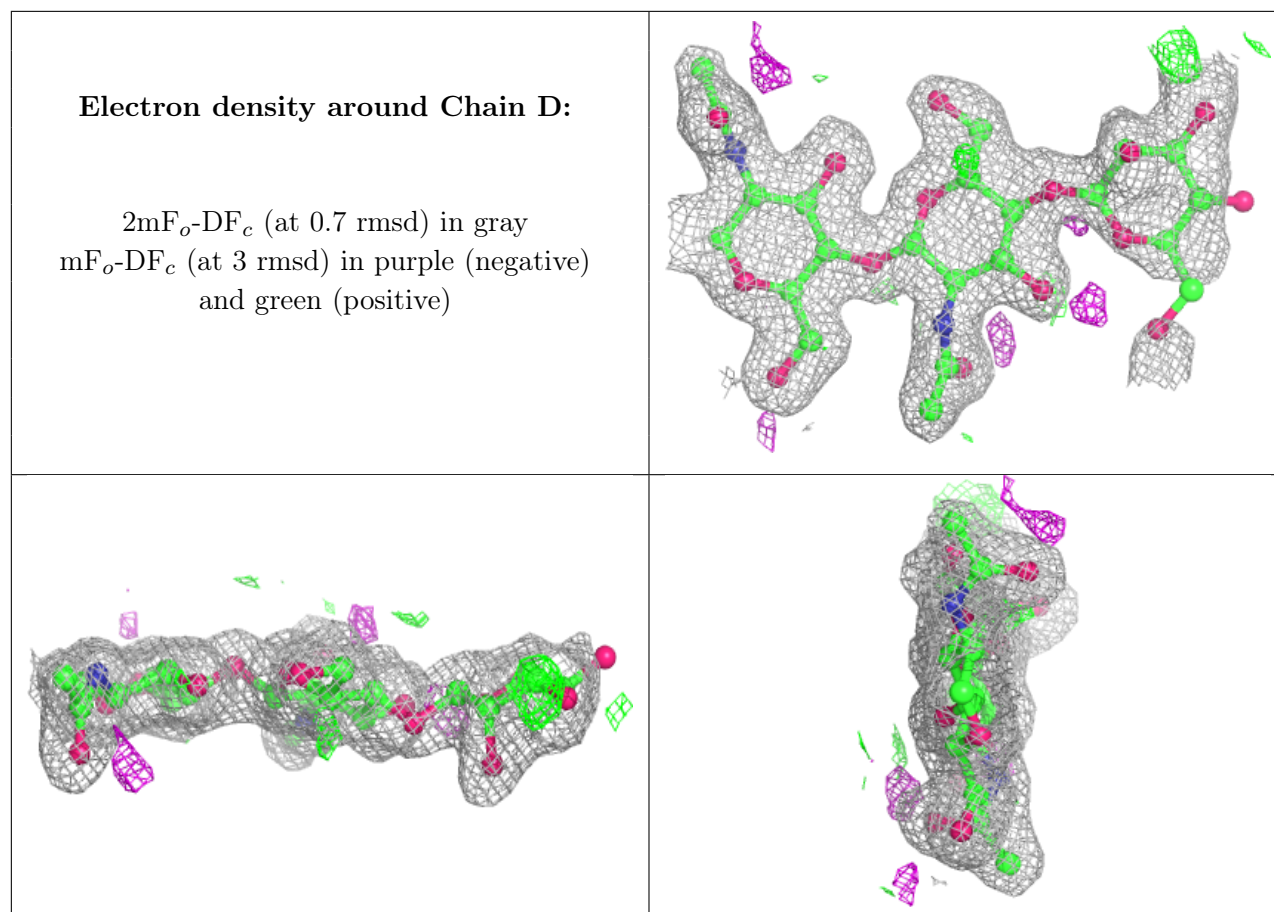
The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.



**Electron density around Chain C:**

$2mF_o-DF_c$  (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
7	SO4	A	1007	5/5	0.69	0.17	70,75,84,89	0
12	EDO	A	1028	4/4	0.72	0.31	70,74,75,78	0
10	PGE	A	1014	10/10	0.74	0.29	63,69,77,83	0
12	EDO	A	1026	4/4	0.75	0.24	68,69,69,72	0
10	PGE	A	1015	10/10	0.76	0.29	71,77,86,107	0
5	NAG	A	1002	14/15	0.77	0.21	39,58,77,79	0
11	PEG	A	1017	7/7	0.77	0.26	48,64,75,78	0
10	PGE	A	1018	10/10	0.79	0.24	48,57,69,70	0
12	EDO	A	1030	4/4	0.80	0.19	51,51,55,60	0
7	SO4	A	1005	5/5	0.81	0.13	56,59,76,78	0
12	EDO	A	1025	4/4	0.83	0.22	45,45,66,69	0
7	SO4	A	1006	5/5	0.83	0.17	51,54,73,82	0

*Continued on next page...*

*Continued from previous page...*

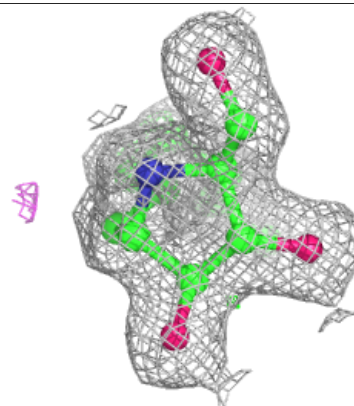
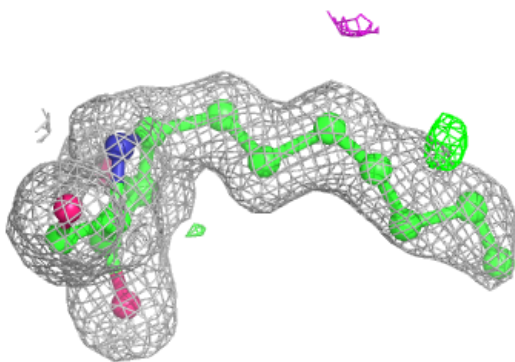
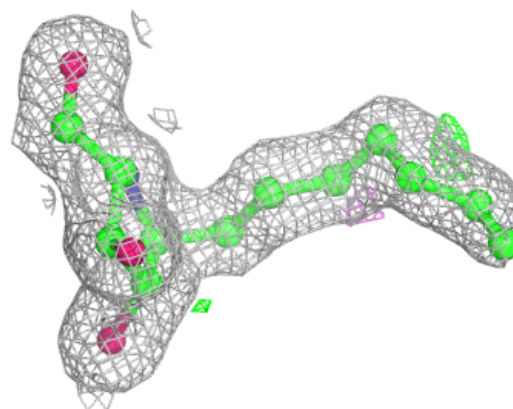
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
10	PGE	A	1016	10/10	0.84	0.23	45,61,85,98	0
12	EDO	A	1029	4/4	0.84	0.17	44,48,49,54	0
12	EDO	A	1024	4/4	0.84	0.26	60,65,65,67	0
12	EDO	A	1019	4/4	0.85	0.21	63,68,68,70	0
12	EDO	A	1021	4/4	0.86	0.21	49,57,64,66	0
12	EDO	A	1020	4/4	0.87	0.20	54,58,62,64	0
8	CL	A	1010	1/1	0.87	0.21	68,68,68,68	0
12	EDO	A	1023	4/4	0.88	0.18	34,51,51,57	0
9	GOL	A	1013	6/6	0.89	0.18	37,47,51,52	0
12	EDO	A	1022	4/4	0.90	0.18	51,61,64,71	0
7	SO4	A	1008	5/5	0.92	0.12	31,40,46,49	0
5	NAG	A	1001	14/15	0.92	0.12	33,41,55,55	0
12	EDO	A	1027	4/4	0.92	0.15	49,61,63,63	0
8	CL	A	1009	1/1	0.93	0.21	74,74,74,74	0
8	CL	A	1011	1/1	0.93	0.24	77,77,77,77	0
9	GOL	A	1012	6/6	0.94	0.12	38,40,48,53	0
7	SO4	A	1004	5/5	0.97	0.07	38,39,40,40	0
6	A1IPC	A	1003	16/16	0.98	0.09	22,26,54,60	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 A1IPC A 1003:**

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



## 6.5 Other polymers ⓘ

There are no such residues in this entry.