



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 7, 2020 – 09:06 PM BST

PDB ID : 3OGS  
Title : Complex structure of beta-galactosidase from *Trichoderma reesei* with IPTG  
Authors : Maksimainen, M.; Rouvinen, J.  
Deposited on : 2010-08-17  
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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.13.1  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.13.1

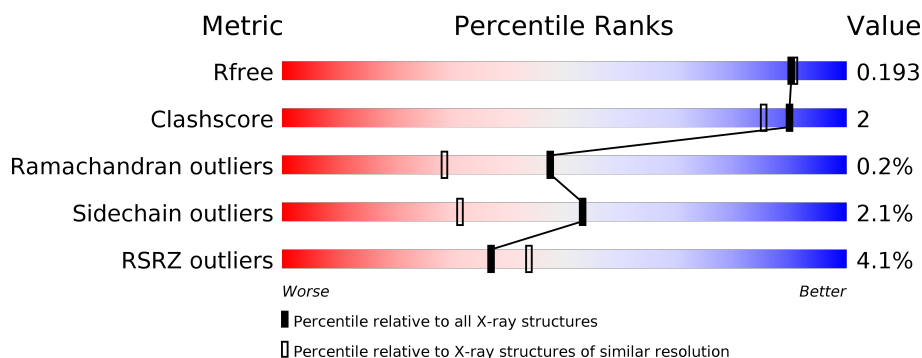
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.75 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))
$R_{free}$	130704	2340 (1.76-1.76)
Clashscore	141614	2466 (1.76-1.76)
Ramachandran outliers	138981	2437 (1.76-1.76)
Sidechain outliers	138945	2437 (1.76-1.76)
RSRZ outliers	127900	2298 (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 on 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	1003	<div> <div>4%</div> <div>87%</div> <div>11%</div> <div>••</div> </div>
2	B	7	<div> <div>14%</div> <div>86%</div> </div>
3	C	9	<div> <div>89%</div> <div>11%</div> </div>
4	D	2	<div> <div>100%</div> </div>

## 2 Entry composition [i](#)

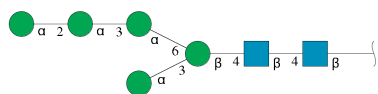
There are 7 unique types of molecules in this entry. The entry contains 8819 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 Beta-galactosidase.

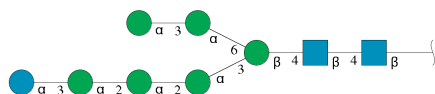
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	986	7634	4906	1289	1431	8	0	3	0

- Molecule 2 is an oligosaccharide called alpha-D-mannopyranose-(1-2)-alpha-D-mannopyranose-(1-3)-alpha-D-mannopyranose-(1-6)-[alpha-D-mannopyranose-(1-3)]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
			Total	C	N	O			
2	B	7	83	46	2	35	0	0	0

- Molecule 3 is an oligosaccharide called alpha-D-glucopyranose-(1-3)-alpha-D-mannopyranose-(1-2)-alpha-D-mannopyranose-(1-2)-alpha-D-mannopyranose-(1-3)-[alpha-D-mannopyranose-(1-3)-alpha-D-mannopyranose-(1-6)]beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.



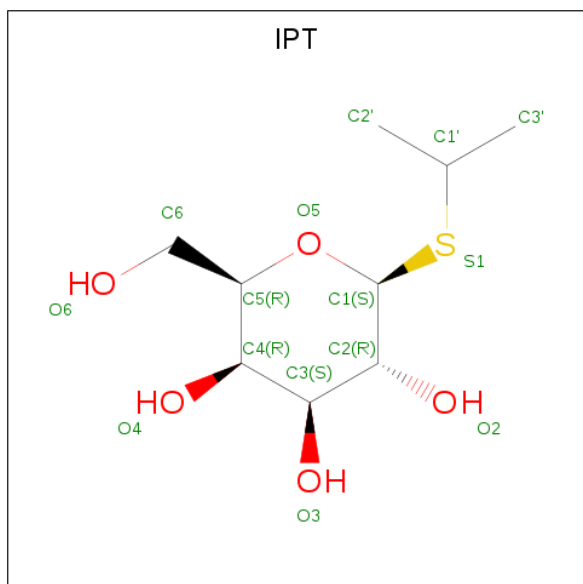
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
3	C	9	105	58	2	45	0	0	0

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



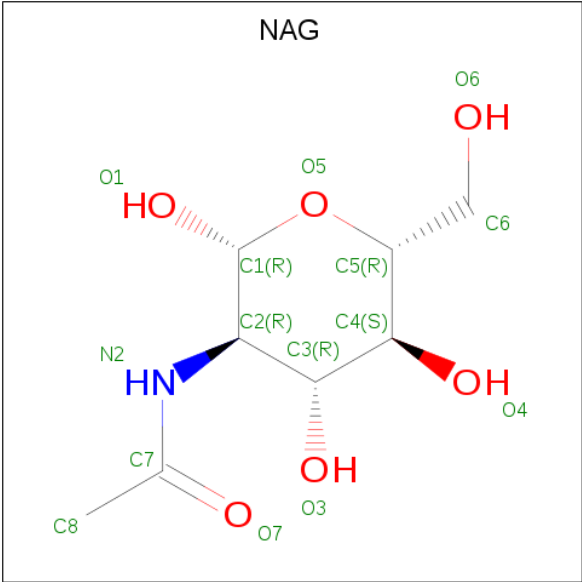
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
4	D	2	Total	C	N	O	0	0	0
			28	16	2	10			

- Molecule 5 is 1-methylethyl 1-thio-beta-D-galactopyranoside (three-letter code: IPT) (formula:  $C_9H_{18}O_5S$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	1	Total	C	O	S	0	0
			15	9	5	1		

- Molecule 6 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



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

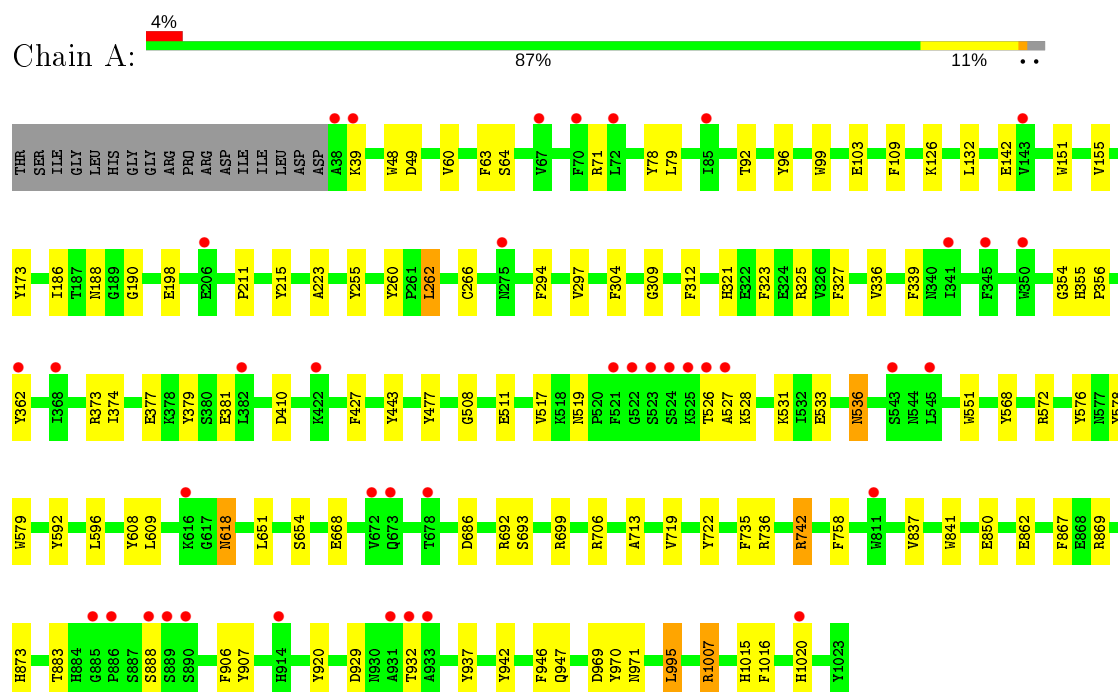
- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	926	Total	O	0	0
			926	926		

### 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.

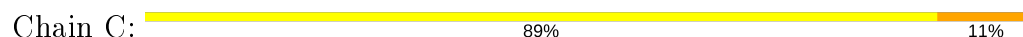
- Molecule 1: Beta-galactosidase



- Molecule 2: alpha-D-mannopyranose-(1-2)-alpha-D-mannopyranose-(1-3)-alpha-D-mannopyranose-(1-6)-[alpha-D-mannopyranose-(1-3)]beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose



- Molecule 3: alpha-D-glucopyranose-(1-3)-alpha-D-mannopyranose-(1-2)-alpha-D-mannopyranose-(1-2)-alpha-D-mannopyranose-(1-3)-[alpha-D-mannopyranose-(1-3)-alpha-D-mannopyranose-(1-6)]beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose



MAG1	MAG2	EM43	MAN4	MAN5	MAN6	GLC7	MAN8	MAN9
------	------	------	------	------	------	------	------	------

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

Chain D:  100%

MAG1	MAG2
------	------

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	69.50 Å 70.30 Å 82.40 Å 108.50° 97.80° 114.40°	Depositor
Resolution (Å)	43.54 – 1.75 43.54 – 1.75	Depositor EDS
% Data completeness (in resolution range)	95.0 (43.54-1.75) 91.2 (43.54-1.75)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.28 (at 1.75 Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, $R_{free}$	0.145 , 0.200 0.159 , 0.193	Depositor DCC
$R_{free}$ test set	5893 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	18.2	Xtriage
Anisotropy	0.590	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 64.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.019 for k,h,-h-k-l	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	8819	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	20.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.62% 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: GLC, BMA, NAG, IPT, MAN

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	1.78	85/7862 (1.1%)	0.99	18/10715 (0.2%)

All (85) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	260	TYR	CD1-CE1	8.49	1.52	1.39
1	A	937	TYR	CD1-CE1	8.14	1.51	1.39
1	A	109	PHE	CE2-CZ	7.80	1.52	1.37
1	A	103	GLU	CB-CG	7.63	1.66	1.52
1	A	362	TYR	CE2-CZ	7.43	1.48	1.38
1	A	920	TYR	CD2-CE2	7.30	1.50	1.39
1	A	377	GLU	CB-CG	7.20	1.65	1.52
1	A	551	TRP	CB-CG	7.18	1.63	1.50
1	A	719	VAL	CB-CG2	6.89	1.67	1.52
1	A	578	TYR	CG-CD1	6.88	1.48	1.39
1	A	906	PHE	CE1-CZ	6.80	1.50	1.37
1	A	942	TYR	CD2-CE2	6.67	1.49	1.39
1	A	443	TYR	CD1-CE1	6.66	1.49	1.39
1	A	151	TRP	CE3-CZ3	6.65	1.49	1.38
1	A	937	TYR	CD2-CE2	6.65	1.49	1.39
1	A	99	TRP	CG-CD1	6.57	1.46	1.36
1	A	142	GLU	CB-CG	6.53	1.64	1.52
1	A	477	TYR	CD1-CE1	6.47	1.49	1.39
1	A	96	TYR	CD1-CE1	6.37	1.49	1.39
1	A	304	PHE	CD1-CE1	6.26	1.51	1.39
1	A	443	TYR	CD2-CE2	6.24	1.48	1.39
1	A	568	TYR	CD1-CE1	6.21	1.48	1.39
1	A	427	PHE	CE1-CZ	6.21	1.49	1.37
1	A	592	TYR	CG-CD2	6.18	1.47	1.39
1	A	722	TYR	CE1-CZ	6.18	1.46	1.38
1	A	266	CYS	CB-SG	6.16	1.92	1.82

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	1016	PHE	CE2-CZ	6.02	1.48	1.37
1	A	294	PHE	CE1-CZ	6.01	1.48	1.37
1	A	758	PHE	CE2-CZ	5.96	1.48	1.37
1	A	379	TYR	CD1-CE1	5.93	1.48	1.39
1	A	850	GLU	CG-CD	5.91	1.60	1.51
1	A	477	TYR	CG-CD1	5.91	1.46	1.39
1	A	309	GLY	N-CA	5.90	1.54	1.46
1	A	568	TYR	CD2-CE2	5.88	1.48	1.39
1	A	354	GLY	N-CA	5.88	1.54	1.46
1	A	533	GLU	CG-CD	5.87	1.60	1.51
1	A	970	TYR	CD1-CE1	5.83	1.48	1.39
1	A	517	VAL	CB-CG1	5.82	1.65	1.52
1	A	155	VAL	CB-CG1	5.82	1.65	1.52
1	A	339	PHE	CD1-CE1	5.77	1.50	1.39
1	A	946	PHE	CE2-CZ	5.71	1.48	1.37
1	A	579	TRP	CZ3-CH2	5.63	1.49	1.40
1	A	907	TYR	CE2-CZ	5.60	1.45	1.38
1	A	592	TYR	CD2-CE2	5.57	1.47	1.39
1	A	198	GLU	CG-CD	5.53	1.60	1.51
1	A	592	TYR	CE1-CZ	5.52	1.45	1.38
1	A	841	TRP	CG-CD1	5.52	1.44	1.36
1	A	381	GLU	CB-CG	5.49	1.62	1.52
1	A	592	TYR	CE2-CZ	5.48	1.45	1.38
1	A	735	PHE	CE2-CZ	5.44	1.47	1.37
1	A	722	TYR	CD1-CE1	5.44	1.47	1.39
1	A	260	TYR	CD2-CE2	5.43	1.47	1.39
1	A	78	TYR	CD1-CE1	5.41	1.47	1.39
1	A	327	PHE	CE2-CZ	5.40	1.47	1.37
1	A	336	VAL	CB-CG1	5.40	1.64	1.52
1	A	841	TRP	CD2-CE2	5.38	1.47	1.41
1	A	578	TYR	CD1-CE1	5.37	1.47	1.39
1	A	576	TYR	CD1-CE1	5.37	1.47	1.39
1	A	869	ARG	CZ-NH1	5.34	1.40	1.33
1	A	215	TYR	CD2-CE2	5.32	1.47	1.39
1	A	970	TYR	CD2-CE2	5.31	1.47	1.39
1	A	255	TYR	CE2-CZ	5.30	1.45	1.38
1	A	48	TRP	CB-CG	5.26	1.59	1.50
1	A	312	PHE	CG-CD1	5.25	1.46	1.38
1	A	294	PHE	CD2-CE2	5.22	1.49	1.39
1	A	508	GLY	N-CA	5.22	1.53	1.46
1	A	841	TRP	CZ3-CH2	5.20	1.48	1.40
1	A	862	GLU	CG-CD	5.20	1.59	1.51

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	867	PHE	CE2-CZ	5.19	1.47	1.37
1	A	356	PRO	N-CA	5.15	1.55	1.47
1	A	867	PHE	CE1-CZ	5.15	1.47	1.37
1	A	173	TYR	CE2-CZ	5.14	1.45	1.38
1	A	297	VAL	CB-CG1	5.13	1.63	1.52
1	A	362	TYR	CD1-CE1	5.13	1.47	1.39
1	A	608	TYR	CD2-CE2	5.13	1.47	1.39
1	A	654	SER	CB-OG	5.13	1.49	1.42
1	A	713	ALA	CA-CB	5.13	1.63	1.52
1	A	568	TYR	CE2-CZ	5.11	1.45	1.38
1	A	132	LEU	C-O	5.11	1.33	1.23
1	A	339	PHE	CE1-CZ	5.10	1.47	1.37
1	A	223	ALA	CA-CB	5.08	1.63	1.52
1	A	381	GLU	CD-OE2	5.04	1.31	1.25
1	A	920	TYR	CG-CD2	5.03	1.45	1.39
1	A	99	TRP	CE3-CZ3	5.03	1.47	1.38
1	A	511	GLU	CD-OE1	5.01	1.31	1.25

All (18) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	869	ARG	NE-CZ-NH1	9.01	124.80	120.30
1	A	869	ARG	NE-CZ-NH2	-8.48	116.06	120.30
1	A	49	ASP	CB-CG-OD1	8.34	125.81	118.30
1	A	572	ARG	NE-CZ-NH1	6.96	123.78	120.30
1	A	373	ARG	NE-CZ-NH1	6.36	123.48	120.30
1	A	995	LEU	CA-CB-CG	6.20	129.56	115.30
1	A	572	ARG	NE-CZ-NH2	-6.12	117.24	120.30
1	A	686	ASP	CB-CG-OD1	5.83	123.55	118.30
1	A	71	ARG	NE-CZ-NH2	-5.78	117.41	120.30
1	A	969	ASP	CB-CG-OD1	5.69	123.42	118.30
1	A	929	ASP	CB-CG-OD1	5.61	123.35	118.30
1	A	995	LEU	CB-CG-CD2	5.28	119.97	111.00
1	A	71	ARG	NE-CZ-NH1	5.26	122.93	120.30
1	A	736	ARG	NE-CZ-NH2	-5.23	117.68	120.30
1	A	325	ARG	NE-CZ-NH2	-5.17	117.71	120.30
1	A	699	ARG	NE-CZ-NH1	-5.16	117.72	120.30
1	A	339	PHE	CB-CG-CD2	-5.06	117.26	120.80
1	A	410	ASP	CB-CG-OD1	5.05	122.85	118.30

There are no chirality outliers.

There are no planarity outliers.

## 5.2 Too-close contacts ⓘ

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	7634	0	7435	26	0
2	B	83	0	70	0	0
3	C	105	0	88	2	0
4	D	28	0	25	0	0
5	A	15	0	18	0	0
6	A	28	0	26	0	0
7	A	926	0	0	15	0
All	All	8819	0	7662	28	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 2.

All (28) 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:1020:HIS:HB2	7:A:1807:HOH:O	1.20	1.34
1:A:188:ASN:HB3	7:A:1881:HOH:O	1.48	1.14
1:A:531:LYS:HB3	7:A:1575:HOH:O	1.62	0.97
1:A:971:ASN:HB3	7:A:1816:HOH:O	1.67	0.95
7:A:1952:HOH:O	3:C:7:GLC:H61	1.79	0.82
1:A:971:ASN:ND2	7:A:1816:HOH:O	2.21	0.67
1:A:536:ASN:OD1	1:A:536:ASN:C	2.33	0.66
1:A:971:ASN:CB	7:A:1816:HOH:O	2.36	0.64
1:A:692:ARG:NH1	7:A:1666:HOH:O	2.30	0.64
1:A:706:ARG:NH1	7:A:1805:HOH:O	2.34	0.60
1:A:262:LEU:HD22	1:A:323:PHE:CE1	2.41	0.56
1:A:262:LEU:HD22	1:A:323:PHE:CD1	2.42	0.54
1:A:321:HIS:HD2	7:A:1189:HOH:O	1.91	0.53
1:A:668:GLU:HB3	7:A:1882:HOH:O	2.12	0.49
1:A:60[B]:VAL:HG11	1:A:596:LEU:CD1	2.43	0.48
1:A:1007:ARG:HD2	7:A:1607:HOH:O	2.12	0.48
1:A:651:LEU:N	1:A:651:LEU:HD23	2.28	0.48
1:A:883:THR:HG21	7:A:1721:HOH:O	2.14	0.48
1:A:651:LEU:H	1:A:651:LEU:HD23	1.80	0.47
1:A:79:LEU:HD22	1:A:126:LYS:HD3	2.00	0.44

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:64:SER:HA	1:A:92:THR:O	2.18	0.43
7:A:1952:HOH:O	3:C:7:GLC:C6	2.53	0.43
1:A:742:ARG:HG3	1:A:837:VAL:HG22	2.01	0.42
1:A:873:HIS:CD2	1:A:873:HIS:H	2.38	0.42
1:A:618:ASN:C	1:A:618:ASN:HD22	2.23	0.42
1:A:883:THR:CG2	7:A:1721:HOH:O	2.67	0.41
1:A:374:ILE:HD11	1:A:609:LEU:HB2	2.03	0.41
1:A:186:ILE:HA	1:A:190:GLY:O	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	987/1003 (98%)	957 (97%)	28 (3%)	2 (0%)	47 29

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	527	ALA
1	A	519	ASN

### 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	808/819 (99%)	791 (98%)	17 (2%)	53 31

All (17) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	39	LYS
1	A	63	PHE
1	A	211	PRO
1	A	262	LEU
1	A	355	HIS
1	A	526	THR
1	A	528	LYS
1	A	536	ASN
1	A	618	ASN
1	A	693	SER
1	A	742	ARG
1	A	888	SER
1	A	932	THR
1	A	947	GLN
1	A	995	LEU
1	A	1007	ARG
1	A	1015	HIS

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

Mol	Chain	Res	Type
1	A	225	ASN
1	A	321	HIS
1	A	412	GLN
1	A	618	ASN
1	A	836	ASN
1	A	873	HIS
1	A	884	HIS
1	A	1015	HIS

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates ⓘ

18 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	1.28	1 (7%)	17,19,21	1.83	4 (23%)
2	NAG	B	2	2	14,14,15	1.05	0	17,19,21	0.91	0
2	BMA	B	3	2	11,11,12	1.56	2 (18%)	15,15,17	1.16	2 (13%)
2	MAN	B	4	2	11,11,12	1.64	2 (18%)	15,15,17	1.52	2 (13%)
2	MAN	B	5	2	11,11,12	1.28	2 (18%)	15,15,17	1.57	4 (26%)
2	MAN	B	6	2	11,11,12	0.68	0	15,15,17	1.46	4 (26%)
2	MAN	B	7	2	11,11,12	1.27	2 (18%)	15,15,17	2.02	5 (33%)
3	NAG	C	1	1,3	14,14,15	1.78	2 (14%)	17,19,21	1.58	4 (23%)
3	NAG	C	2	3	14,14,15	0.79	0	17,19,21	1.43	3 (17%)
3	BMA	C	3	3	11,11,12	1.09	1 (9%)	15,15,17	1.39	2 (13%)
3	MAN	C	4	3	11,11,12	1.23	1 (9%)	15,15,17	2.37	5 (33%)
3	MAN	C	5	3	11,11,12	0.97	0	15,15,17	1.90	5 (33%)
3	MAN	C	6	3	11,11,12	0.93	1 (9%)	15,15,17	2.27	5 (33%)
3	GLC	C	7	3	11,11,12	0.70	0	15,15,17	2.75	4 (26%)
3	MAN	C	8	3	11,11,12	0.88	0	15,15,17	2.20	2 (13%)
3	MAN	C	9	3	11,11,12	0.63	0	15,15,17	1.10	1 (6%)
4	NAG	D	1	1,4	14,14,15	0.57	0	17,19,21	1.77	4 (23%)
4	NAG	D	2	4	14,14,15	0.64	0	17,19,21	1.22	1 (5%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	NAG	B	1	1,2	-	0/6/23/26	0/1/1/1
2	NAG	B	2	2	-	0/6/23/26	0/1/1/1
2	BMA	B	3	2	-	0/2/19/22	0/1/1/1
2	MAN	B	4	2	-	0/2/19/22	0/1/1/1
2	MAN	B	5	2	-	1/2/19/22	0/1/1/1
2	MAN	B	6	2	-	2/2/19/22	0/1/1/1
2	MAN	B	7	2	-	0/2/19/22	0/1/1/1
3	NAG	C	1	1,3	-	1/6/23/26	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
3	MAN	C	4	3	-	0/2/19/22	0/1/1/1
3	MAN	C	5	3	-	0/2/19/22	0/1/1/1
3	MAN	C	6	3	-	0/2/19/22	0/1/1/1
3	GLC	C	7	3	-	0/2/19/22	0/1/1/1
3	MAN	C	8	3	-	0/2/19/22	0/1/1/1
3	MAN	C	9	3	-	2/2/19/22	0/1/1/1
4	NAG	D	1	1,4	-	0/6/23/26	0/1/1/1
4	NAG	D	2	4	-	2/6/23/26	0/1/1/1

All (14) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	1	NAG	C1-C2	5.02	1.59	1.52
2	B	4	MAN	C2-C3	2.85	1.56	1.52
3	C	1	NAG	C3-C2	2.82	1.58	1.52
2	B	3	BMA	O4-C4	2.73	1.49	1.43
2	B	7	MAN	C2-C3	2.71	1.56	1.52
2	B	5	MAN	O2-C2	2.58	1.48	1.43
2	B	7	MAN	O5-C1	-2.51	1.39	1.43
3	C	4	MAN	C2-C3	2.49	1.56	1.52
2	B	4	MAN	C6-C5	2.43	1.60	1.51
2	B	3	BMA	C2-C3	2.42	1.56	1.52
3	C	6	MAN	C2-C3	2.29	1.55	1.52
2	B	5	MAN	C2-C3	2.13	1.55	1.52
2	B	1	NAG	C2-N2	2.10	1.49	1.46
3	C	3	BMA	O2-C2	2.07	1.47	1.43

All (57) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	7	GLC	C1-O5-C5	7.10	121.81	112.19
3	C	4	MAN	C6-C5-C4	-6.91	96.82	113.00

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	8	MAN	O5-C5-C6	6.57	117.51	107.20
3	C	7	GLC	C1-C2-C3	5.71	116.68	109.67
2	B	7	MAN	C1-O5-C5	5.61	119.80	112.19
4	D	1	NAG	O5-C5-C6	4.98	115.01	107.20
3	C	6	MAN	O3-C3-C2	4.61	118.83	109.99
3	C	6	MAN	O3-C3-C4	-4.24	100.55	110.35
3	C	5	MAN	C1-O5-C5	3.97	117.57	112.19
3	C	6	MAN	O2-C2-C1	-3.82	101.34	109.15
2	B	1	NAG	C4-C3-C2	3.79	116.58	111.02
3	C	4	MAN	C3-C4-C5	3.60	116.66	110.24
3	C	8	MAN	O3-C3-C2	3.42	116.55	109.99
2	B	5	MAN	C1-O5-C5	3.36	116.75	112.19
2	B	1	NAG	O5-C1-C2	-3.36	105.99	111.29
3	C	3	BMA	O3-C3-C2	3.30	116.31	109.99
2	B	4	MAN	C1-O5-C5	3.29	116.66	112.19
2	B	4	MAN	O2-C2-C3	-3.25	103.63	110.14
3	C	7	GLC	O5-C5-C6	3.25	112.29	107.20
3	C	7	GLC	O3-C3-C2	-3.21	103.85	109.99
3	C	5	MAN	O5-C5-C6	3.11	112.09	107.20
2	B	1	NAG	O5-C5-C6	3.06	111.99	107.20
3	C	5	MAN	O2-C2-C1	3.02	115.33	109.15
3	C	2	NAG	C1-O5-C5	3.01	116.27	112.19
2	B	7	MAN	C6-C5-C4	-2.96	106.06	113.00
3	C	1	NAG	C2-N2-C7	2.95	127.11	122.90
3	C	1	NAG	O5-C1-C2	-2.94	106.65	111.29
2	B	3	BMA	O4-C4-C3	-2.85	103.76	110.35
3	C	3	BMA	O3-C3-C4	-2.85	103.77	110.35
2	B	6	MAN	O2-C2-C3	-2.74	104.64	110.14
4	D	1	NAG	O3-C3-C2	-2.66	103.96	109.47
3	C	2	NAG	O7-C7-C8	-2.64	117.16	122.06
3	C	1	NAG	C6-C5-C4	-2.60	106.92	113.00
4	D	1	NAG	C3-C4-C5	-2.58	105.63	110.24
2	B	5	MAN	O6-C6-C5	-2.53	102.61	111.29
3	C	1	NAG	C8-C7-N2	2.52	120.36	116.10
3	C	5	MAN	O2-C2-C3	-2.50	105.13	110.14
2	B	5	MAN	O2-C2-C1	2.43	114.13	109.15
3	C	6	MAN	O2-C2-C3	2.38	114.91	110.14
2	B	5	MAN	O3-C3-C2	2.37	114.54	109.99
2	B	3	BMA	O3-C3-C2	-2.36	105.47	109.99
2	B	6	MAN	O6-C6-C5	-2.34	103.25	111.29
2	B	6	MAN	O5-C1-C2	-2.34	107.16	110.77
3	C	4	MAN	C2-C3-C4	-2.29	106.94	110.89

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	6	MAN	C1-C2-C3	2.28	112.47	109.67
4	D	1	NAG	C4-C3-C2	2.27	114.35	111.02
3	C	4	MAN	O5-C1-C2	2.26	114.26	110.77
3	C	5	MAN	C1-C2-C3	-2.26	106.89	109.67
3	C	4	MAN	O4-C4-C3	2.21	115.46	110.35
2	B	7	MAN	O5-C5-C6	-2.20	103.75	107.20
2	B	7	MAN	O3-C3-C2	2.19	114.18	109.99
2	B	6	MAN	C2-C3-C4	-2.19	107.11	110.89
4	D	2	NAG	C1-O5-C5	2.17	115.13	112.19
3	C	2	NAG	O5-C5-C6	2.15	110.57	107.20
3	C	9	MAN	O5-C5-C6	2.14	110.56	107.20
2	B	7	MAN	O5-C1-C2	-2.14	107.47	110.77
2	B	1	NAG	C1-O5-C5	2.13	115.08	112.19

There are no chirality outliers.

All (8) torsion outliers are listed below:

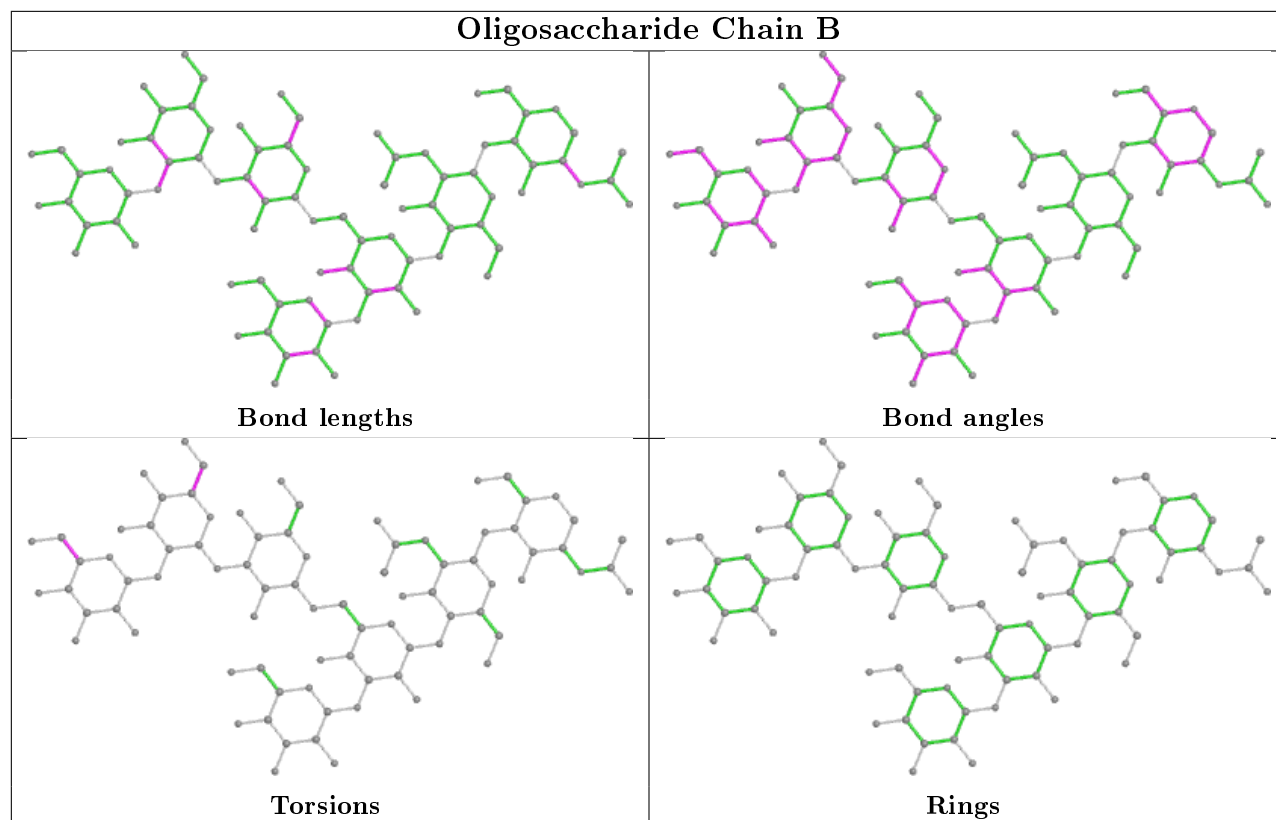
Mol	Chain	Res	Type	Atoms
3	C	9	MAN	C4-C5-C6-O6
3	C	9	MAN	O5-C5-C6-O6
2	B	6	MAN	O5-C5-C6-O6
2	B	6	MAN	C4-C5-C6-O6
4	D	2	NAG	O5-C5-C6-O6
4	D	2	NAG	C4-C5-C6-O6
2	B	5	MAN	C4-C5-C6-O6
3	C	1	NAG	C4-C5-C6-O6

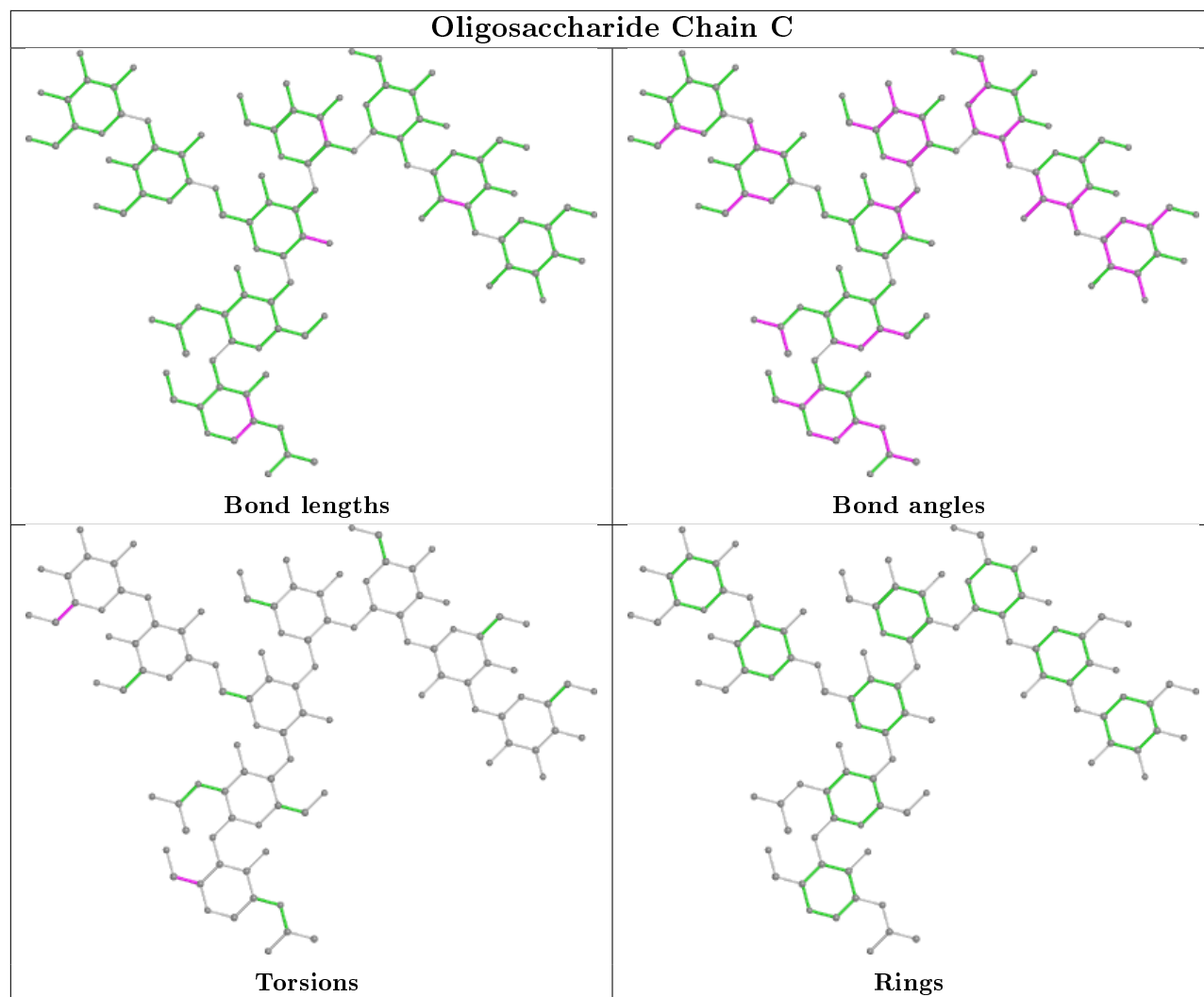
There are no ring outliers.

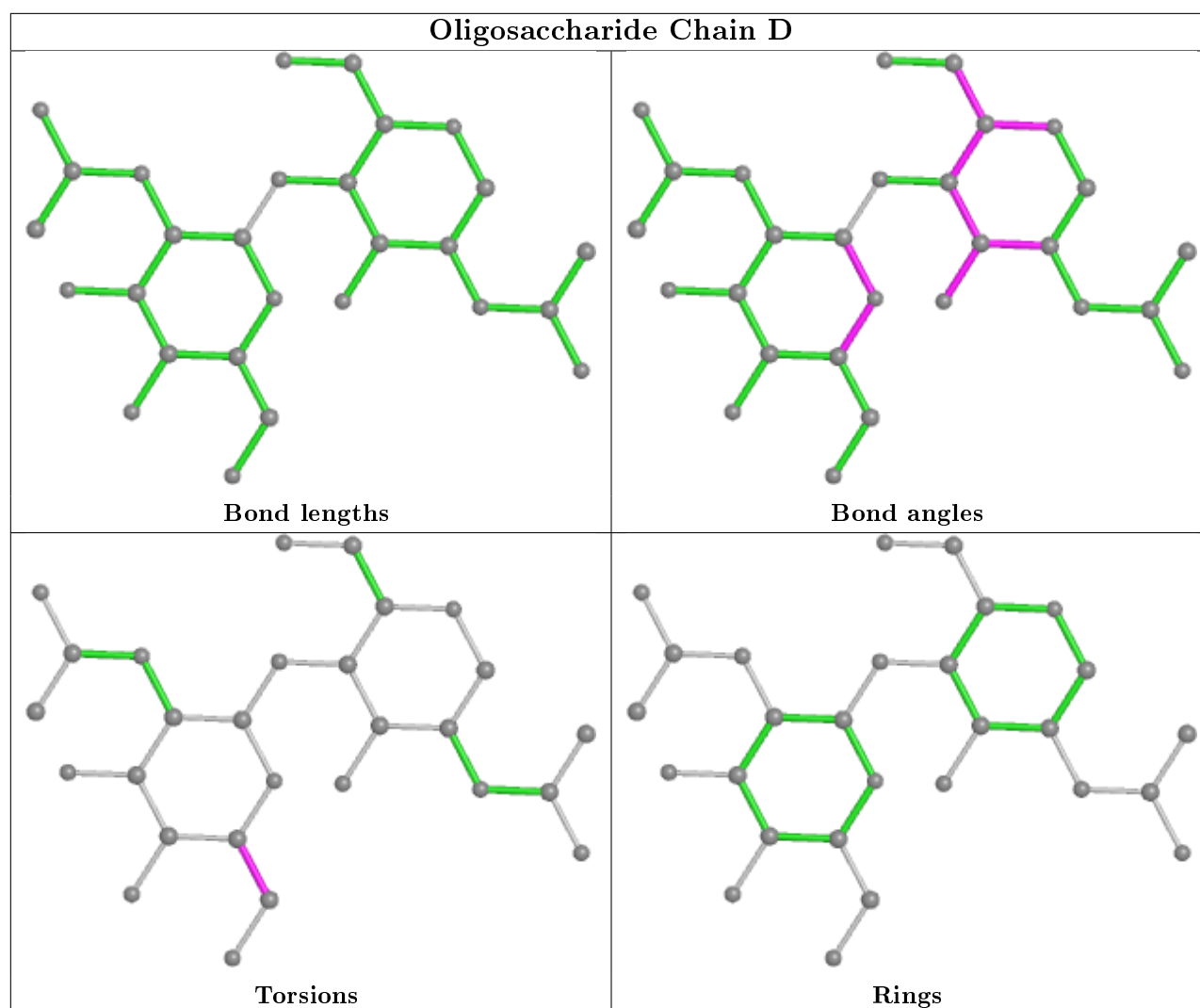
1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	C	7	GLC	2	0

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 ⓘ

3 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
6	NAG	A	1042	1	14,14,15	0.54	0	17,19,21	1.75	5 (29%)
5	IPT	A	1024	-	14,15,15	1.33	1 (7%)	18,21,21	1.78	6 (33%)
6	NAG	A	1041	1	14,14,15	0.58	0	17,19,21	2.34	4 (23%)

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
6	NAG	A	1042	1	-	0/6/23/26	0/1/1/1
5	IPT	A	1024	-	-	3/6/26/26	0/1/1/1
6	NAG	A	1041	1	-	1/6/23/26	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	1024	IPT	C4-C5	3.12	1.59	1.53

All (15) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	A	1041	NAG	C1-O5-C5	7.32	122.11	112.19
5	A	1024	IPT	C1-O5-C5	4.24	120.39	112.58
6	A	1041	NAG	O5-C1-C2	3.74	117.19	111.29
6	A	1042	NAG	C1-O5-C5	3.69	117.19	112.19
6	A	1042	NAG	O3-C3-C2	-3.24	102.75	109.47
6	A	1041	NAG	C1-C2-N2	-3.16	105.09	110.49
6	A	1042	NAG	O7-C7-C8	-2.97	116.55	122.06
5	A	1024	IPT	O2-C2-C3	-2.51	104.54	110.35
5	A	1024	IPT	C2-C1-S1	-2.49	107.46	111.30
6	A	1042	NAG	C8-C7-N2	2.37	120.11	116.10
6	A	1041	NAG	O5-C5-C4	2.17	116.10	110.83
5	A	1024	IPT	O4-C4-C3	-2.15	105.37	110.35
6	A	1042	NAG	O6-C6-C5	-2.10	104.10	111.29
5	A	1024	IPT	O5-C5-C4	-2.07	105.93	109.69
5	A	1024	IPT	C1-S1-C1'	-2.04	94.01	100.26

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	1024	IPT	C2'-C1'-S1-C1
5	A	1024	IPT	C3'-C1'-S1-C1
6	A	1041	NAG	O5-C5-C6-O6
5	A	1024	IPT	O5-C5-C6-O6

There are no ring outliers.

No monomer is involved in short contacts.

## 5.7 Other polymers

There are no such residues in this entry.

## 5.8 Polymer linkage issues

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	986/1003 (98%)	0.07	40 (4%)	37 44	10, 17, 33, 64	0

All (40) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	933	ALA	8.5
1	A	526	THR	7.9
1	A	932	THR	6.6
1	A	525	LYS	6.5
1	A	931	ALA	5.5
1	A	523	SER	4.9
1	A	543	SER	4.0
1	A	522	GLY	3.8
1	A	811	TRP	3.7
1	A	524	SER	3.7
1	A	886	PRO	3.6
1	A	888	SER	3.5
1	A	889	SER	3.1
1	A	678	THR	3.1
1	A	885	GLY	2.9
1	A	368	ILE	2.9
1	A	527	ALA	2.9
1	A	38	ALA	2.8
1	A	616	LYS	2.8
1	A	39	LYS	2.7
1	A	673	GLN	2.7
1	A	275	ASN	2.6
1	A	672	VAL	2.6
1	A	545	LEU	2.5
1	A	345	PHE	2.5
1	A	72	LEU	2.5
1	A	341	ILE	2.4

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	206	GLU	2.4
1	A	382	LEU	2.3
1	A	890	SER	2.3
1	A	350	TRP	2.3
1	A	85	ILE	2.3
1	A	521	PHE	2.2
1	A	362	TYR	2.2
1	A	67	VAL	2.2
1	A	143	VAL	2.2
1	A	70	PHE	2.1
1	A	422	LYS	2.1
1	A	1020	HIS	2.1
1	A	914	HIS	2.0

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

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

## 6.3 Carbohydrates [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<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.

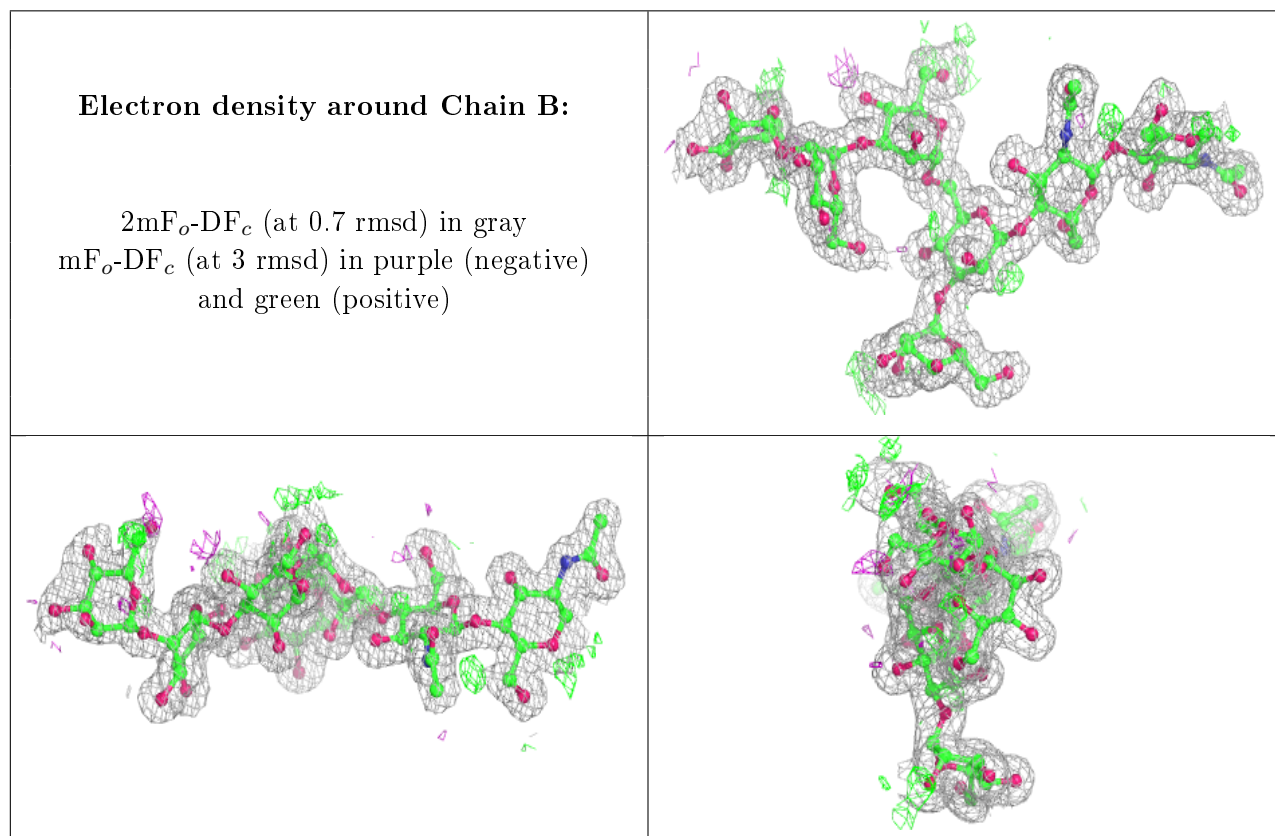
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
3	MAN	C	9	11/12	0.68	0.38	55,62,64,65	0
4	NAG	D	2	14/15	0.76	0.33	46,56,63,63	0
3	GLC	C	7	11/12	0.80	0.16	41,47,48,49	0
2	MAN	B	6	11/12	0.88	0.29	30,35,41,47	0
3	MAN	C	8	11/12	0.89	0.22	34,40,45,48	0
3	MAN	C	6	11/12	0.92	0.13	24,30,35,38	0
3	NAG	C	1	14/15	0.93	0.11	13,18,24,30	0
4	NAG	D	1	14/15	0.93	0.15	32,37,44,46	0
3	MAN	C	4	11/12	0.93	0.13	22,25,31,34	0
2	MAN	B	7	11/12	0.94	0.09	24,26,28,30	0
3	MAN	C	5	11/12	0.94	0.14	17,22,26,27	0
2	MAN	B	4	11/12	0.95	0.09	13,15,18,25	0
3	NAG	C	2	14/15	0.96	0.13	16,23,37,39	0
2	BMA	B	3	11/12	0.96	0.07	13,16,19,19	0
3	BMA	C	3	11/12	0.97	0.11	19,21,25,29	0

*Continued on next page...*

*Continued from previous page...*

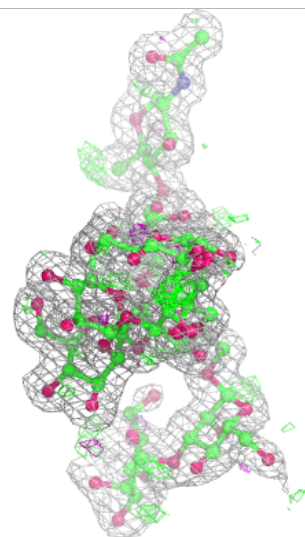
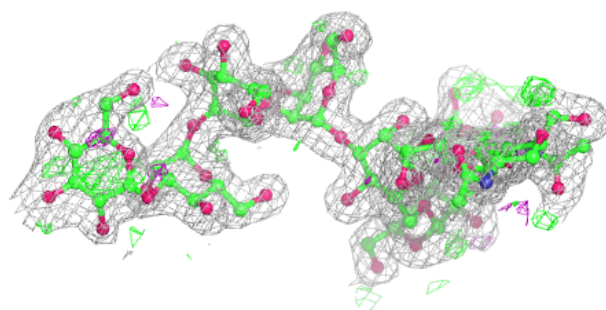
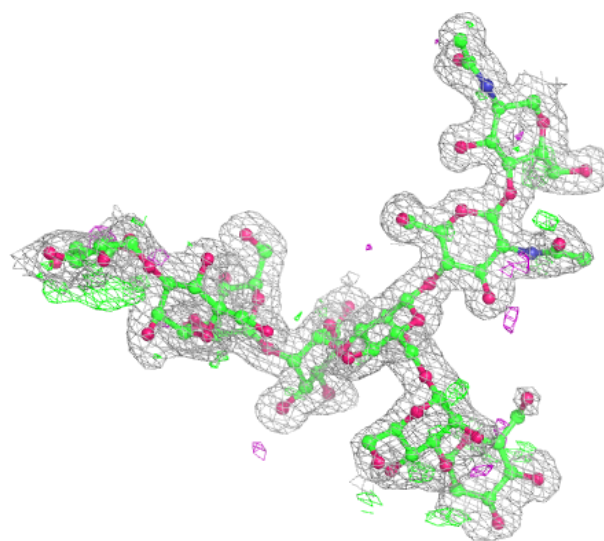
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
2	NAG	B	1	14/15	0.97	0.08	13,15,25,27	0
2	NAG	B	2	14/15	0.97	0.07	13,15,21,22	0
2	MAN	B	5	11/12	0.97	0.07	15,16,28,35	0

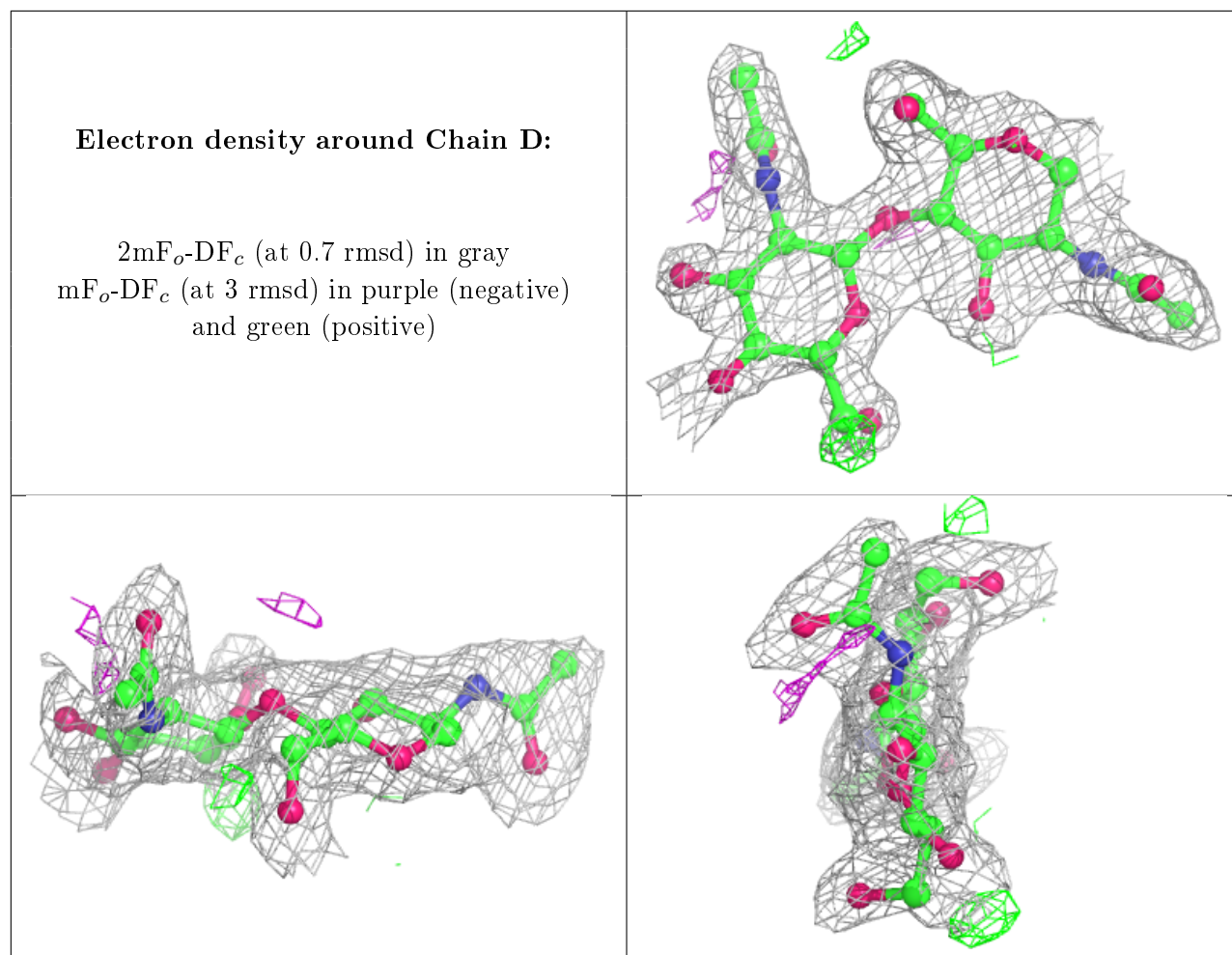
The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.



**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
6	NAG	A	1041	14/15	0.77	0.22	39,48,56,59	0
6	NAG	A	1042	14/15	0.92	0.25	33,39,48,52	0
5	IPT	A	1024	15/15	0.97	0.11	9,12,26,27	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.