



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 21, 2020 – 07:21 PM BST

PDB ID : 3USU  
Title : Crystal structure of Butea monosperma seed lectin  
Authors : Abhilash, J.; Geethanandan, K.; Bharath, S.R.; Sadasivan, C.; Haridas, M.  
Deposited on : 2011-11-24  
Resolution : 2.46 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](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

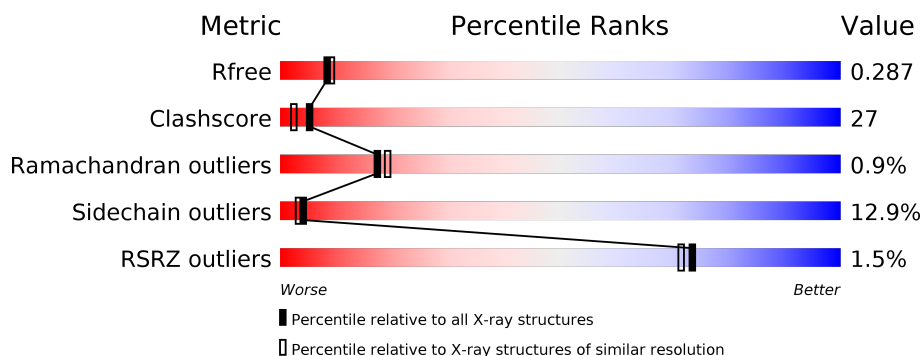
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.46 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	1544 (2.48-2.44)
Clashscore	141614	1613 (2.48-2.44)
Ramachandran outliers	138981	1598 (2.48-2.44)
Sidechain outliers	138945	1598 (2.48-2.44)
RSRZ outliers	127900	1523 (2.48-2.44)

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	256	
1	C	256	
1	E	256	
1	G	256	
2	B	242	
2	D	242	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	F	242	
2	H	242	
3	I	4	
3	K	4	
3	L	4	
4	J	5	
5	M	2	
6	N	2	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
11	XYP	B	306	-	-	X	-
4	BMA	J	3	-	-	X	-
9	ABU	A	280	-	-	X	-
9	ABU	B	276	-	-	X	-
9	ABU	C	282	-	-	X	-
9	ABU	E	288	-	-	X	-
9	ABU	F	290	-	-	X	-

## 2 Entry composition [i](#)

There are 13 unique types of molecules in this entry. The entry contains 15748 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.

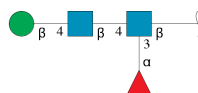
- Molecule 1 is a protein called Lectin Alpha chain.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
1	A	250	Total	C	N	O	0	0	0
			1892	1227	296	369			
1	C	250	Total	C	N	O	0	0	0
			1892	1227	296	369			
1	E	250	Total	C	N	O	0	0	0
			1892	1227	296	369			
1	G	250	Total	C	N	O	0	0	0
			1892	1227	296	369			

- Molecule 2 is a protein called Lectin Beta Chain.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	B	242	Total	C	N	O	0	0	0
			1828	1184	287	357			
2	D	242	Total	C	N	O	0	0	0
			1828	1184	287	357			
2	F	242	Total	C	N	O	0	0	0
			1828	1184	287	357			
2	H	242	Total	C	N	O	0	0	0
			1828	1184	287	357			

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



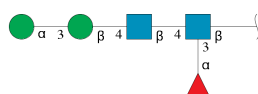
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	I	4	Total	C	N	O	0	0	0
			49	28	2	19			

*Continued on next page...*

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	K	4	Total	C	N	O	0	0	0
			49	28	2	19			
3	L	4	Total	C	N	O	0	0	0
			49	28	2	19			

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



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
4	J	5	Total	C	N	O	0	0	0
			60	34	2	24			

- Molecule 5 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
5	M	2	Total	C	N	O	0	0	0
			28	16	2	10			

- Molecule 6 is an oligosaccharide called alpha-L-fucopyranose-(1-3)-2-acetamido-2-deoxy-beta-D-glucopyranose.



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

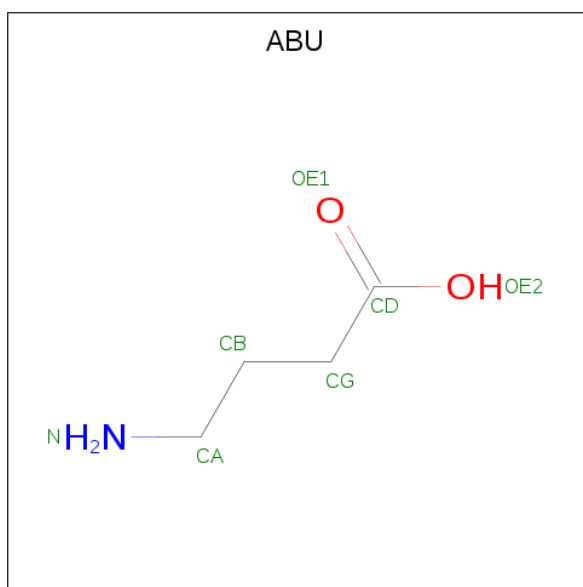
- Molecule 7 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	G	1	Total Ca 1 1	0	0
7	D	1	Total Ca 1 1	0	0
7	E	1	Total Ca 1 1	0	0
7	H	1	Total Ca 1 1	0	0
7	B	1	Total Ca 1 1	0	0
7	C	1	Total Ca 1 1	0	0
7	A	1	Total Ca 1 1	0	0
7	F	1	Total Ca 1 1	0	0

- Molecule 8 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	G	1	Total Mn 1 1	0	0
8	D	1	Total Mn 1 1	0	0
8	E	1	Total Mn 1 1	0	0
8	H	1	Total Mn 1 1	0	0
8	B	1	Total Mn 1 1	0	0
8	C	1	Total Mn 1 1	0	0
8	A	1	Total Mn 1 1	0	0
8	F	1	Total Mn 1 1	0	0

- Molecule 9 is GAMMA-AMINO-BUTANOIC ACID (three-letter code: ABU) (formula: C<sub>4</sub>H<sub>9</sub>NO<sub>2</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
9	A	1	Total	C	N	O	0	0
			7	4	1	2		
9	B	1	Total	C	N	O	0	0
			7	4	1	2		
9	C	1	Total	C	N	O	0	0
			7	4	1	2		
9	D	1	Total	C	N	O	0	0
			7	4	1	2		
9	E	1	Total	C	N	O	0	0
			7	4	1	2		
9	E	1	Total	C	N	O	0	0
			7	4	1	2		
9	F	1	Total	C	N	O	0	0
			7	4	1	2		
9	G	1	Total	C	N	O	0	0
			7	4	1	2		

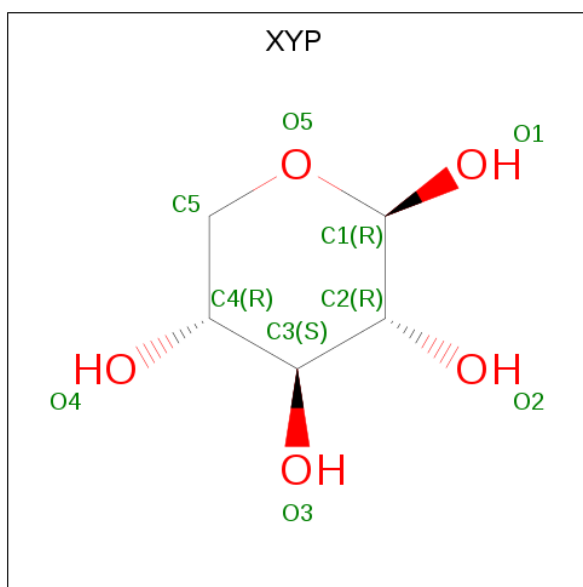
- Molecule 10 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
10	A	1	Total	C	O	0	0
			6	3	3		
10	B	1	Total	C	O	0	0
			6	3	3		
10	C	1	Total	C	O	0	0
			6	3	3		
10	D	1	Total	C	O	0	0
			6	3	3		
10	E	1	Total	C	O	0	0
			6	3	3		
10	E	1	Total	C	O	0	0
			6	3	3		
10	G	1	Total	C	O	0	0
			6	3	3		
10	H	1	Total	C	O	0	0
			6	3	3		

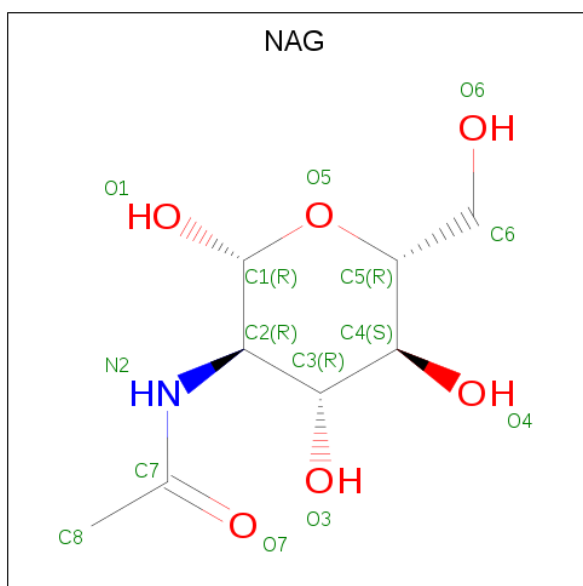
- Molecule 11 is beta-D-xylopyranose (three-letter code: XYP) (formula: C<sub>5</sub>H<sub>10</sub>O<sub>5</sub>).





Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
11	B	1	Total	C	O	0	0
			9	5	4		

- Molecule 12 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).

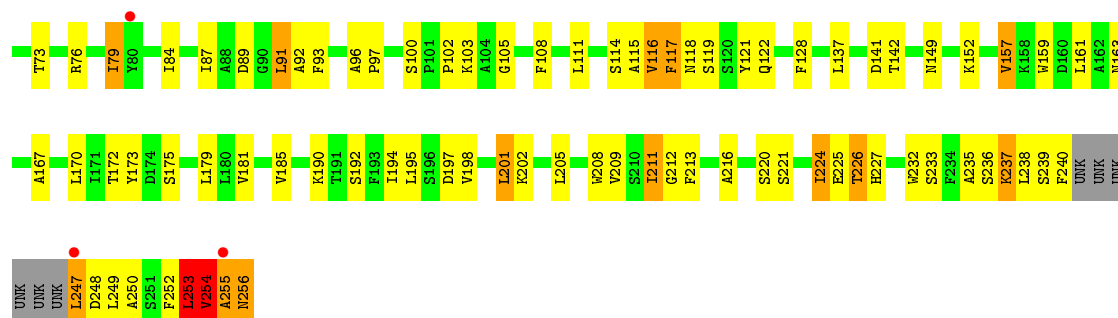


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
12	D	1	Total	C	N	O	0	0
			14	8	1	5		

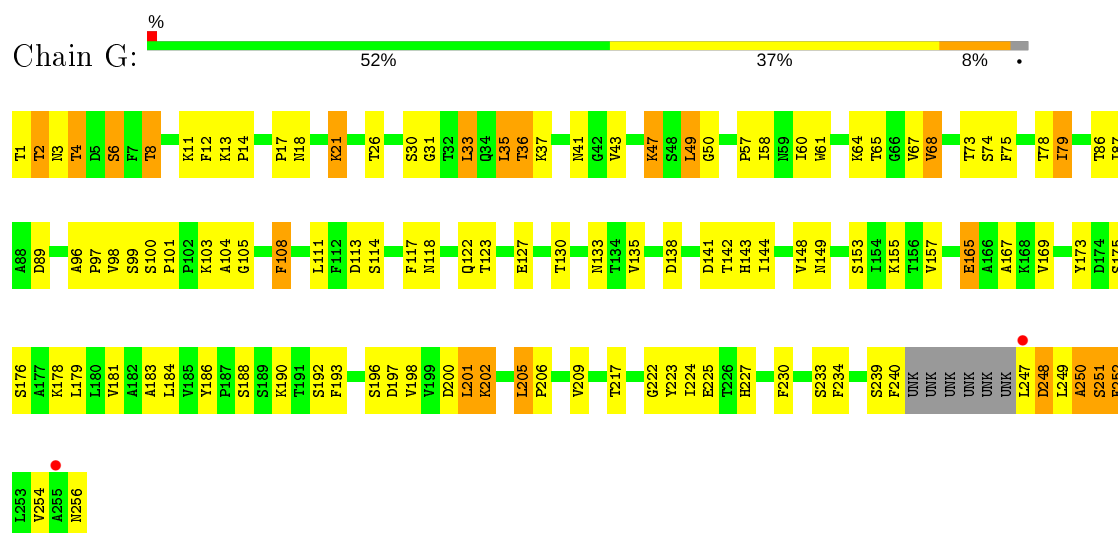
- Molecule 13 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
13	A	65	Total 65	O 65	0	0
13	B	70	Total 70	O 70	0	0
13	C	64	Total 64	O 64	0	0
13	D	47	Total 47	O 47	0	0
13	E	51	Total 51	O 51	0	0
13	F	65	Total 65	O 65	0	0
13	G	60	Total 60	O 60	0	0
13	H	44	Total 44	O 44	0	0

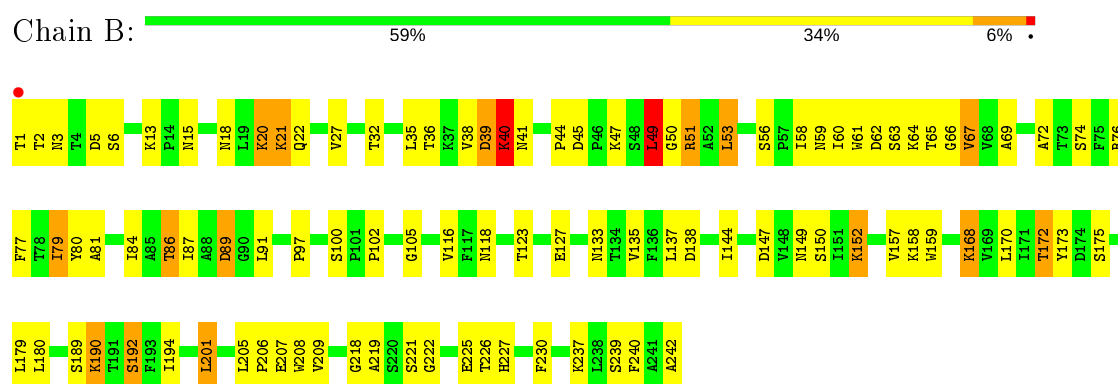




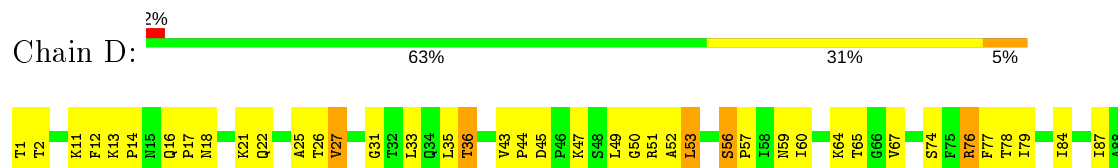
### • Molecule 1: Lectin Alpha chain

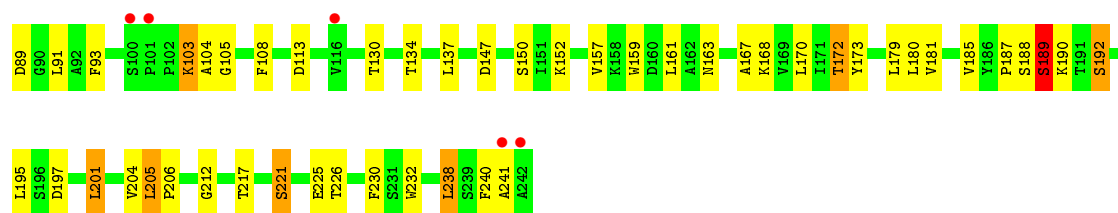


### • Molecule 2: Lectin Beta Chain

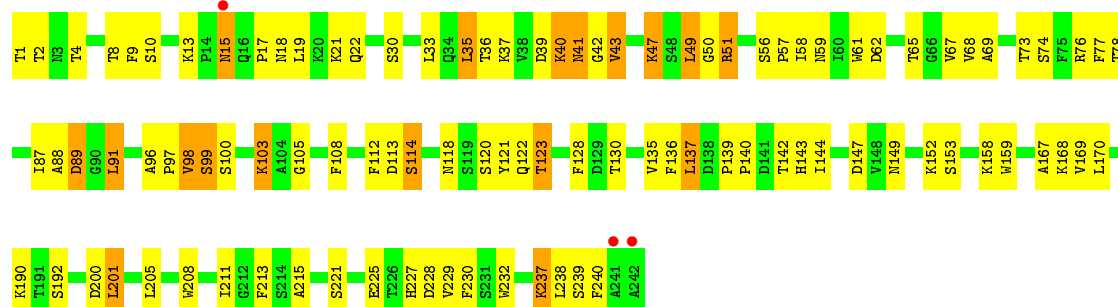


### • Molecule 2: Lectin Beta Chain

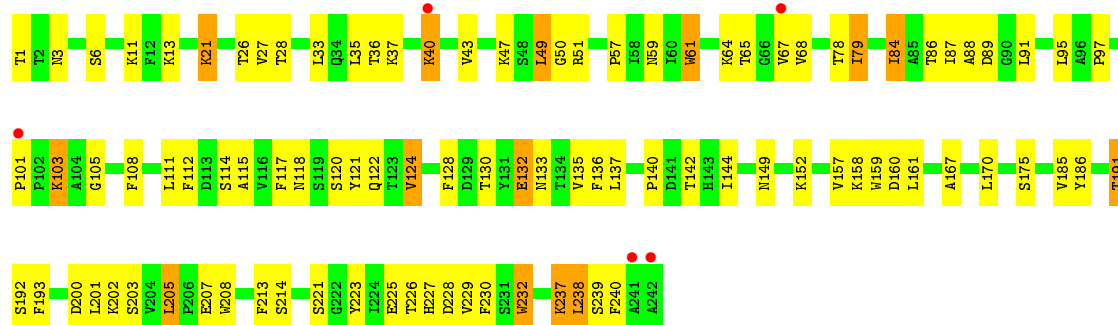




• Molecule 2: Lectin Beta Chain



• Molecule 2: Lectin Beta Chain



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



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



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

Chain L:  50% 50%

MAG1  
MAG2  
BMA3  
FUC4

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

Chain J:  40% 60%

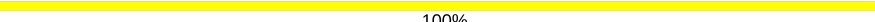
MAG1  
MAG2  
BMA3  
MAA4  
FUC5

- Molecule 5: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain M:  100%

MAG1  
MAG2

- Molecule 6: alpha-L-fucopyranose-(1-3)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain N:  100%

MAG1  
FUC2

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	78.45Å 78.91Å 101.85Å 74.30° 76.65° 86.88°	Depositor
Resolution (Å)	69.06 – 2.46 69.06 – 2.46	Depositor EDS
% Data completeness (in resolution range)	94.7 (69.06-2.46) 94.7 (69.06-2.46)	Depositor EDS
$R_{merge}$	0.12	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.99 (at 2.45Å)	Xtriage
Refinement program	REFMAC 5.6.0117	Depositor
R, $R_{free}$	0.205 , 0.287 0.206 , 0.287	Depositor DCC
$R_{free}$ test set	3944 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	26.1	Xtriage
Anisotropy	0.245	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 46.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	0.022 for -k,-h,-l	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	15748	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	23.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.94% 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: XYP, GOL, BMA, NAG, CA, MN, ABU, MAN, FUC

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	0.90	2/1939 (0.1%)	0.96	0/2646
1	C	0.79	3/1939 (0.2%)	0.91	1/2646 (0.0%)
1	E	0.77	2/1939 (0.1%)	0.89	3/2646 (0.1%)
1	G	0.74	1/1939 (0.1%)	0.89	0/2646
2	B	0.85	1/1875 (0.1%)	0.94	3/2560 (0.1%)
2	D	0.75	2/1875 (0.1%)	0.84	2/2560 (0.1%)
2	F	0.83	2/1875 (0.1%)	0.93	3/2560 (0.1%)
2	H	0.72	2/1875 (0.1%)	0.85	0/2560
All	All	0.80	15/15256 (0.1%)	0.90	12/20824 (0.1%)

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
1	C	0	1
1	E	0	1
All	All	0	3

The worst 5 of 15 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	232	TRP	CD2-CE2	5.96	1.48	1.41
2	H	61	TRP	CD2-CE2	5.95	1.48	1.41
1	C	208	TRP	CD2-CE2	5.78	1.48	1.41
1	E	61	TRP	CD2-CE2	5.58	1.48	1.41
2	D	232	TRP	CD2-CE2	5.58	1.48	1.41

The worst 5 of 12 bond angle outliers are listed below:



Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	39	ASP	CB-CG-OD1	7.57	125.11	118.30
1	E	253	LEU	CA-CB-CG	6.04	129.19	115.30
2	B	49	LEU	CA-CB-CG	-5.96	101.60	115.30
1	C	249	LEU	CA-CB-CG	5.89	128.84	115.30
1	E	253	LEU	CB-CG-CD2	-5.72	101.28	111.00

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	98	VAL	Peptide
1	C	247	LEU	Peptide
1	E	22	GLN	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	1892	0	1874	128	0
1	C	1892	0	1875	97	0
1	E	1892	0	1875	112	0
1	G	1892	0	1875	114	0
2	B	1828	0	1809	108	0
2	D	1828	0	1810	80	0
2	F	1828	0	1810	96	0
2	H	1828	0	1810	112	0
3	I	49	0	43	4	0
3	K	49	0	42	0	0
3	L	49	0	43	1	0
4	J	60	0	47	9	0
5	M	28	0	25	0	0
6	N	24	0	21	0	0
7	A	1	0	0	0	0
7	B	1	0	0	0	0
7	C	1	0	0	0	0
7	D	1	0	0	0	0
7	E	1	0	0	0	0
7	F	1	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	G	1	0	0	0	0
7	H	1	0	0	0	0
8	A	1	0	0	0	0
8	B	1	0	0	0	0
8	C	1	0	0	0	0
8	D	1	0	0	0	0
8	E	1	0	0	0	0
8	F	1	0	0	0	0
8	G	1	0	0	0	0
8	H	1	0	0	0	0
9	A	7	0	5	5	0
9	B	7	0	5	6	0
9	C	7	0	5	6	0
9	D	7	0	5	1	0
9	E	14	0	10	11	0
9	F	7	0	5	7	0
9	G	7	0	5	1	0
10	A	6	0	8	2	0
10	B	6	0	8	0	0
10	C	6	0	8	1	0
10	D	6	0	8	0	0
10	E	12	0	16	2	0
10	G	6	0	8	0	0
10	H	6	0	8	3	0
11	B	9	0	0	6	0
12	D	14	0	12	0	0
13	A	65	0	0	16	0
13	B	70	0	0	20	0
13	C	64	0	0	13	0
13	D	47	0	0	11	0
13	E	51	0	0	5	0
13	F	65	0	0	13	0
13	G	60	0	0	17	0
13	H	44	0	0	15	0
All	All	15748	0	15075	806	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 27.

The worst 5 of 806 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)
2:H:238:LEU:HG	13:H:261:HOH:O	1.30	1.28
1:G:101:PRO:HD2	13:G:542:HOH:O	1.34	1.23
11:B:306:XYP:C1	4:J:3:BMA:O2	1.87	1.21
10:H:289:GOL:H2	13:H:411:HOH:O	1.45	1.15
1:A:36:THR:HG21	1:A:227:HIS:HD2	1.04	1.13

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	246/256 (96%)	229 (93%)	12 (5%)	5 (2%)	7	5
1	C	246/256 (96%)	230 (94%)	13 (5%)	3 (1%)	13	12
1	E	246/256 (96%)	228 (93%)	15 (6%)	3 (1%)	13	12
1	G	246/256 (96%)	231 (94%)	12 (5%)	3 (1%)	13	12
2	B	240/242 (99%)	227 (95%)	11 (5%)	2 (1%)	19	22
2	D	240/242 (99%)	224 (93%)	15 (6%)	1 (0%)	34	41
2	F	240/242 (99%)	230 (96%)	9 (4%)	1 (0%)	34	41
2	H	240/242 (99%)	226 (94%)	14 (6%)	0	100	100
All	All	1944/1992 (98%)	1825 (94%)	101 (5%)	18 (1%)	17	19

5 of 18 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	41	ASN
1	A	255	ALA
2	B	40	LYS
1	C	178	LYS
1	E	253	LEU

### 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	210/210 (100%)	175 (83%)	35 (17%)	2	1
1	C	210/210 (100%)	190 (90%)	20 (10%)	8	9
1	E	210/210 (100%)	173 (82%)	37 (18%)	2	1
1	G	210/210 (100%)	179 (85%)	31 (15%)	3	2
2	B	202/202 (100%)	177 (88%)	25 (12%)	4	4
2	D	202/202 (100%)	181 (90%)	21 (10%)	7	7
2	F	202/202 (100%)	178 (88%)	24 (12%)	5	4
2	H	202/202 (100%)	182 (90%)	20 (10%)	8	7
All	All	1648/1648 (100%)	1435 (87%)	213 (13%)	4	3

5 of 213 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	D	192	SER
1	E	117	PHE
2	H	37	LYS
2	D	221	SER
1	E	27	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 33 such sidechains are listed below:

Mol	Chain	Res	Type
2	D	227	HIS
1	E	227	HIS
2	H	59	ASN
1	E	18	ASN
1	E	22	GLN

### 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 ⓘ

21 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
3	NAG	I	1	1,3	14,14,15	0.71	1 (7%)	17,19,21	1.51	3 (17%)
3	NAG	I	2	3	14,14,15	0.59	0	17,19,21	2.18	4 (23%)
3	BMA	I	3	3	11,11,12	1.30	3 (27%)	15,15,17	2.36	5 (33%)
3	FUC	I	4	3	10,10,11	0.73	0	14,14,16	1.52	3 (21%)
4	NAG	J	1	2,4	14,14,15	0.89	1 (7%)	17,19,21	2.35	3 (17%)
4	NAG	J	2	4	14,14,15	1.05	1 (7%)	17,19,21	1.98	4 (23%)
4	BMA	J	3	4	11,11,12	2.32	3 (27%)	15,15,17	4.21	6 (40%)
4	MAN	J	4	4	11,11,12	6.34	8 (72%)	15,15,17	3.98	8 (53%)
4	FUC	J	5	4	10,10,11	0.75	0	14,14,16	1.27	3 (21%)
3	NAG	K	1	1,3	14,14,15	0.58	0	17,19,21	1.35	2 (11%)
3	NAG	K	2	3	14,14,15	0.59	0	17,19,21	1.60	5 (29%)
3	BMA	K	3	3	11,11,12	0.81	0	15,15,17	1.14	1 (6%)
3	FUC	K	4	3	10,10,11	4.32	4 (40%)	14,14,16	4.37	5 (35%)
3	NAG	L	1	1,3	14,14,15	1.18	2 (14%)	17,19,21	1.62	4 (23%)
3	NAG	L	2	3	14,14,15	0.71	0	17,19,21	2.31	7 (41%)
3	BMA	L	3	3	11,11,12	0.81	0	15,15,17	1.45	3 (20%)
3	FUC	L	4	3	10,10,11	0.80	0	14,14,16	1.00	1 (7%)
5	NAG	M	1	1,5	14,14,15	0.71	0	17,19,21	1.49	3 (17%)
5	NAG	M	2	5	14,14,15	0.87	1 (7%)	17,19,21	1.93	4 (23%)
6	NAG	N	1	2,6	14,14,15	1.41	3 (21%)	17,19,21	2.69	6 (35%)
6	FUC	N	2	6	10,10,11	3.78	4 (40%)	14,14,16	4.08	8 (57%)

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	NAG	I	1	1,3	-	0/6/23/26	0/1/1/1
3	NAG	I	2	3	-	2/6/23/26	0/1/1/1
3	BMA	I	3	3	-	0/2/19/22	0/1/1/1
3	FUC	I	4	3	-	-	0/1/1/1
4	NAG	J	1	2,4	-	2/6/23/26	0/1/1/1
4	NAG	J	2	4	-	0/6/23/26	0/1/1/1
4	BMA	J	3	4	-	1/2/19/22	0/1/1/1
4	MAN	J	4	4	-	2/2/19/22	0/1/1/1
4	FUC	J	5	4	-	-	0/1/1/1
3	NAG	K	1	1,3	-	0/6/23/26	0/1/1/1
3	NAG	K	2	3	-	0/6/23/26	0/1/1/1
3	BMA	K	3	3	-	1/2/19/22	0/1/1/1
3	FUC	K	4	3	-	-	0/1/1/1
3	NAG	L	1	1,3	-	1/6/23/26	0/1/1/1
3	NAG	L	2	3	-	2/6/23/26	0/1/1/1
3	BMA	L	3	3	-	2/2/19/22	0/1/1/1
3	FUC	L	4	3	-	-	0/1/1/1
5	NAG	M	1	1,5	-	0/6/23/26	0/1/1/1
5	NAG	M	2	5	-	0/6/23/26	0/1/1/1
6	NAG	N	1	2,6	-	1/6/23/26	0/1/1/1
6	FUC	N	2	6	-	-	0/1/1/1

The worst 5 of 31 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	K	4	FUC	C2-C3	-12.58	1.34	1.52
4	J	4	MAN	O5-C5	-12.03	1.19	1.43
4	J	4	MAN	C2-C3	-9.27	1.38	1.52
4	J	4	MAN	O2-C2	8.96	1.62	1.43
6	N	2	FUC	C4-C5	-7.19	1.36	1.52

The worst 5 of 88 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	K	4	FUC	C1-C2-C3	12.71	125.29	109.67
4	J	3	BMA	O4-C4-C5	11.56	138.01	109.30
6	N	2	FUC	C6-C5-C4	10.45	132.39	113.07

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	J	3	BMA	O3-C3-C2	9.83	128.82	109.99
4	J	4	MAN	C1-C2-C3	8.28	119.84	109.67

There are no chirality outliers.

5 of 14 torsion outliers are listed below:

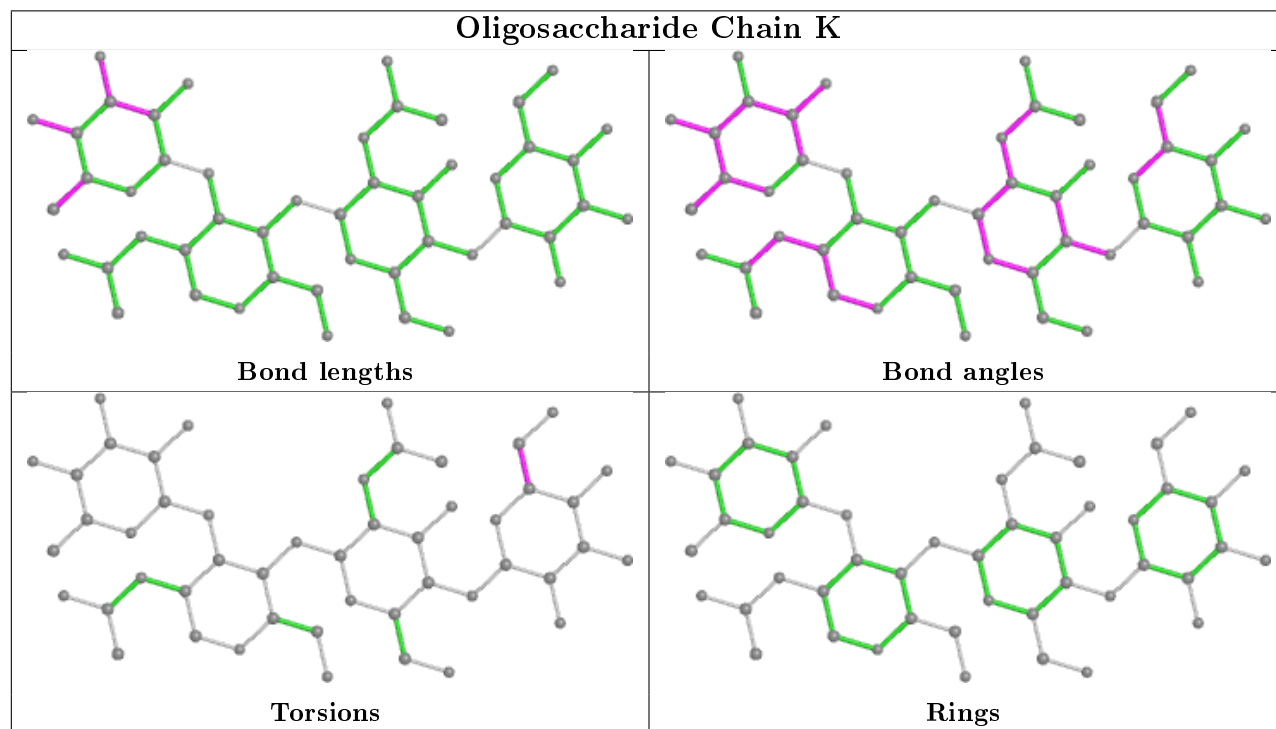
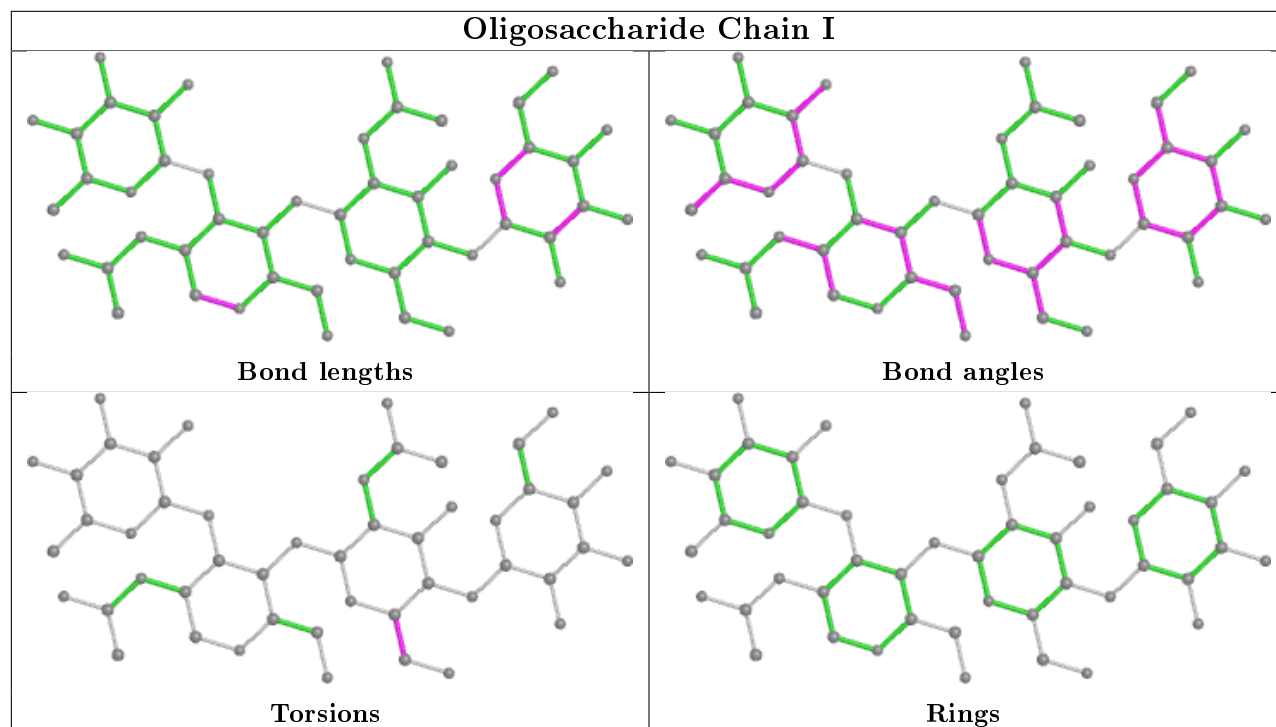
Mol	Chain	Res	Type	Atoms
3	I	2	NAG	O5-C5-C6-O6
3	I	2	NAG	C4-C5-C6-O6
3	L	3	BMA	C4-C5-C6-O6
4	J	1	NAG	C4-C5-C6-O6
4	J	4	MAN	C4-C5-C6-O6

There are no ring outliers.

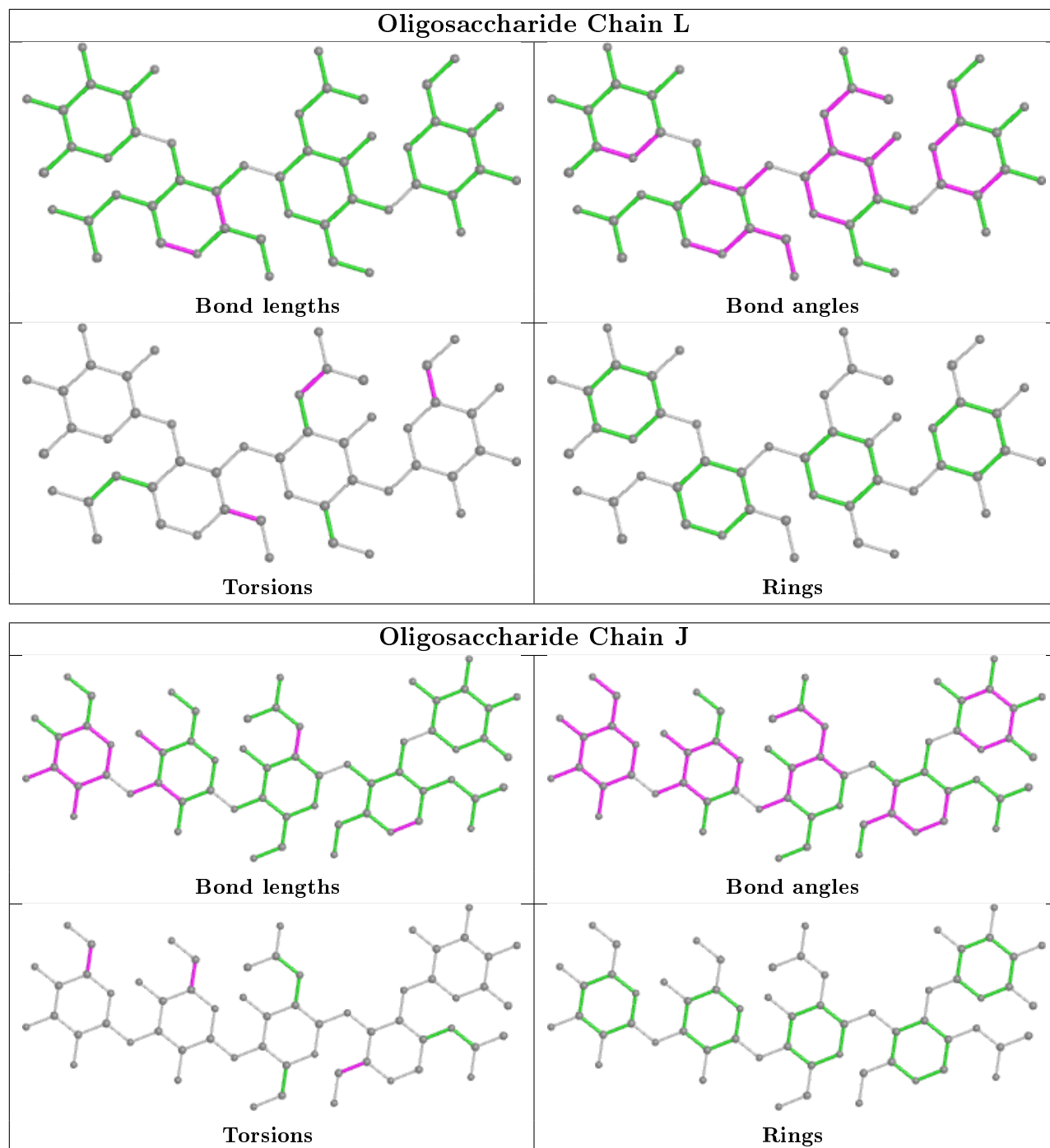
9 monomers are involved in 14 short contacts:

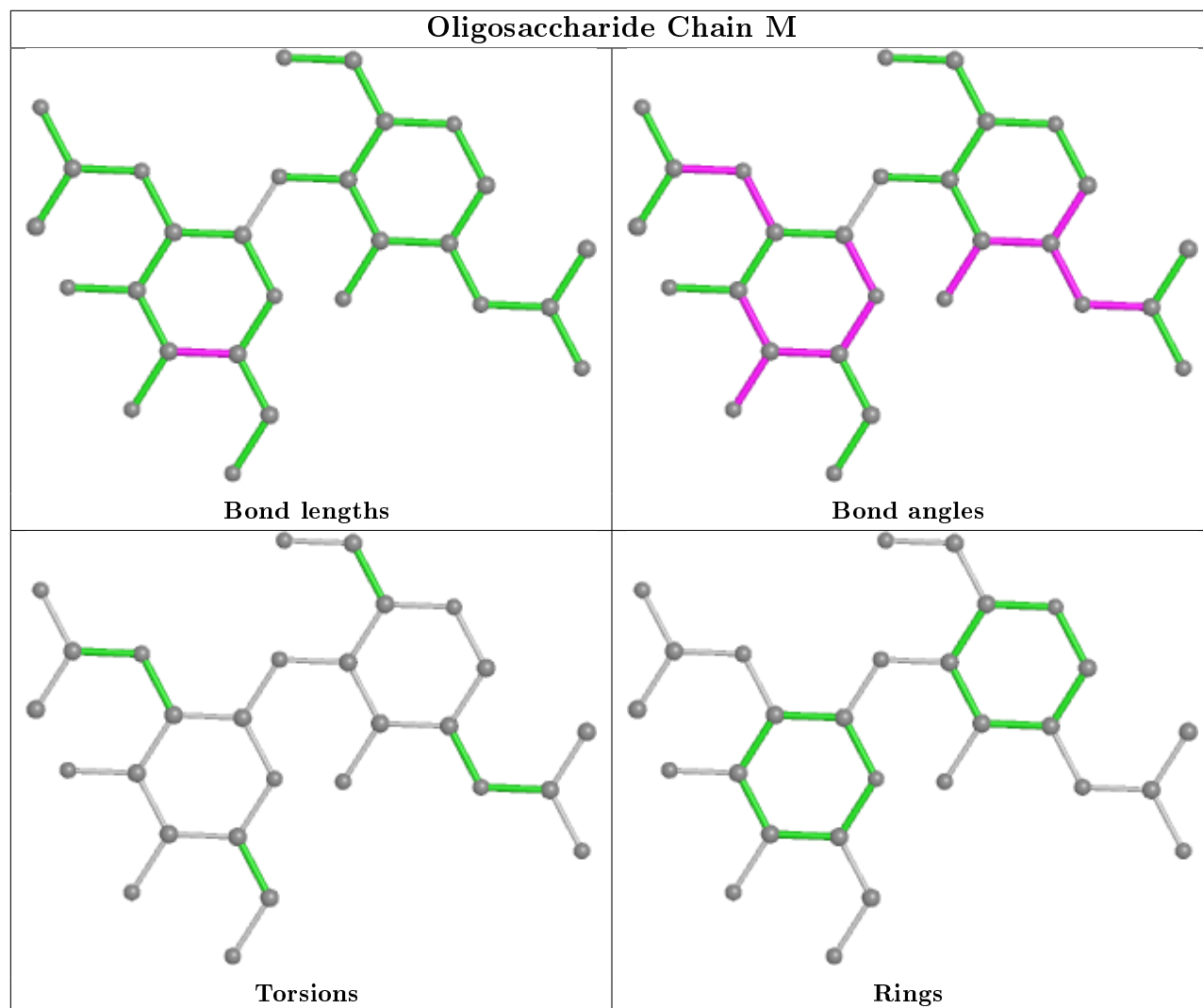
Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	L	4	FUC	1	0
4	J	1	NAG	1	0
3	I	1	NAG	2	0
4	J	3	BMA	8	0
3	I	2	NAG	2	0
3	I	3	BMA	1	0
3	I	4	FUC	1	0
4	J	4	MAN	2	0
3	L	2	NAG	1	0

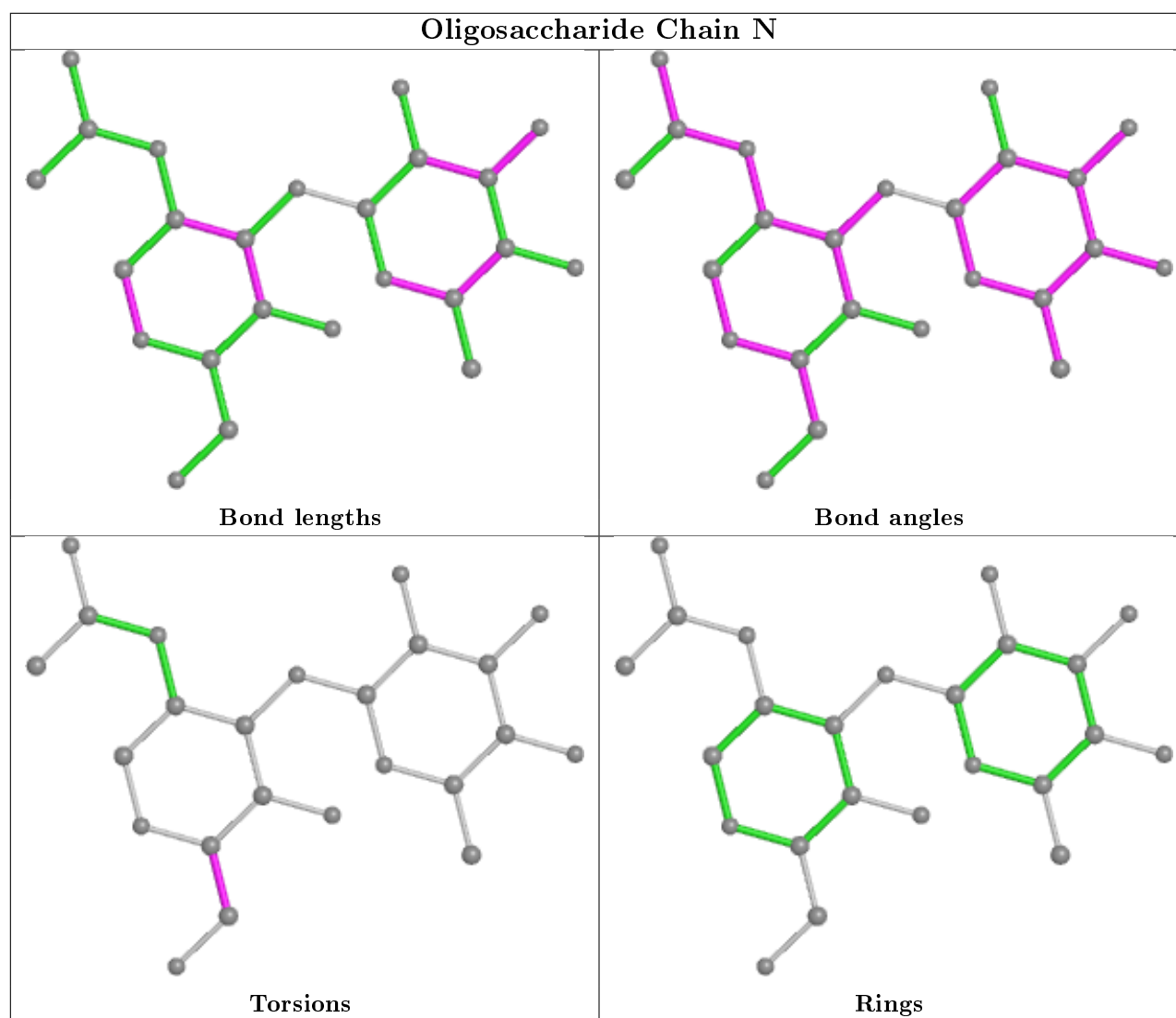
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for oligosaccharide.











## 5.6 Ligand geometry [i](#)

Of 34 ligands modelled in this entry, 16 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	$\# Z  > 2$
9	ABU	G	286	-	3,6,6	0.35	0	2,6,6	0.36	0
10	GOL	B	281	-	5,5,5	0.29	0	5,5,5	0.92	0
9	ABU	D	278	-	3,6,6	0.22	0	2,6,6	0.83	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
10	GOL	C	283	-	5,5,5	0.35	0	5,5,5	0.70	0
9	ABU	A	280	-	3,6,6	0.29	0	2,6,6	0.56	0
10	GOL	G	287	-	5,5,5	0.31	0	5,5,5	0.32	0
10	GOL	E	279	-	5,5,5	0.27	0	5,5,5	1.14	1 (20%)
9	ABU	B	276	-	3,6,6	0.39	0	2,6,6	0.13	0
10	GOL	E	291	-	5,5,5	0.42	0	5,5,5	0.64	0
9	ABU	E	284	-	3,6,6	0.59	0	2,6,6	0.53	0
9	ABU	F	290	-	3,6,6	0.39	0	2,6,6	1.72	1 (50%)
10	GOL	A	277	-	5,5,5	0.28	0	5,5,5	0.60	0
10	GOL	D	285	-	5,5,5	0.32	0	5,5,5	0.65	0
11	XYP	B	306	-	9,9,10	2.45	3 (33%)	10,12,14	6.50	5 (50%)
9	ABU	E	288	-	3,6,6	0.14	0	2,6,6	0.58	0
10	GOL	H	289	-	5,5,5	0.31	0	5,5,5	0.43	0
9	ABU	C	282	-	3,6,6	0.25	0	2,6,6	0.08	0
12	NAG	D	301	2	14,14,15	1.81	3 (21%)	17,19,21	1.64	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
9	ABU	G	286	-	-	2/2/4/4	-
10	GOL	B	281	-	-	0/4/4/4	-
9	ABU	D	278	-	-	2/2/4/4	-
10	GOL	C	283	-	-	0/4/4/4	-
9	ABU	A	280	-	-	1/2/4/4	-
10	GOL	G	287	-	-	2/4/4/4	-
10	GOL	E	279	-	-	2/4/4/4	-
9	ABU	B	276	-	-	1/2/4/4	-
10	GOL	E	291	-	-	4/4/4/4	-
9	ABU	E	284	-	-	0/2/4/4	-
9	ABU	F	290	-	-	1/2/4/4	-
10	GOL	A	277	-	-	2/4/4/4	-
10	GOL	D	285	-	-	4/4/4/4	-
11	XYP	B	306	-	-	-	0/1/1/1
9	ABU	E	288	-	-	1/2/4/4	-
10	GOL	H	289	-	-	2/4/4/4	-
9	ABU	C	282	-	-	1/2/4/4	-
12	NAG	D	301	2	-	2/6/23/26	0/1/1/1

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
11	B	306	XYP	O5-C1	5.75	1.54	1.42
12	D	301	NAG	C1-C2	-4.27	1.46	1.52
12	D	301	NAG	C3-C2	-4.03	1.43	1.52
11	B	306	XYP	C5-C4	2.91	1.58	1.52
11	B	306	XYP	C4-C3	2.80	1.56	1.52

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
11	B	306	XYP	C1-C2-C3	16.27	129.67	109.67
11	B	306	XYP	C5-O5-C1	9.43	126.02	111.52
11	B	306	XYP	O4-C4-C3	-6.03	98.05	110.14
11	B	306	XYP	O4-C4-C5	-4.76	99.42	109.15
12	D	301	NAG	C1-O5-C5	4.14	117.80	112.19

There are no chirality outliers.

5 of 27 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
10	H	289	GOL	O1-C1-C2-C3
10	A	277	GOL	C1-C2-C3-O3
9	E	288	ABU	CA-CB-CG-CD
10	G	287	GOL	O1-C1-C2-C3
10	D	285	GOL	O1-C1-C2-C3

There are no ring outliers.

13 monomers are involved in 51 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
9	G	286	ABU	1	0
9	D	278	ABU	1	0
10	C	283	GOL	1	0
9	A	280	ABU	5	0
10	E	279	GOL	2	0
9	B	276	ABU	6	0
9	E	284	ABU	3	0
9	F	290	ABU	7	0
10	A	277	GOL	2	0
11	B	306	XYP	6	0
9	E	288	ABU	8	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Clashes	Symm-Clashes
10	H	289	GOL	3	0
9	C	282	ABU	6	0

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 6 Fit of model and data [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<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	250/256 (97%)	-0.02	1 (0%) 92 93	12, 20, 30, 48	1 (0%)
1	C	250/256 (97%)	-0.03	4 (1%) 72 69	12, 21, 39, 63	0
1	E	250/256 (97%)	0.22	8 (3%) 47 44	14, 24, 40, 59	0
1	G	250/256 (97%)	-0.05	2 (0%) 86 86	14, 22, 35, 52	0
2	B	242/242 (100%)	-0.05	1 (0%) 92 93	11, 18, 26, 35	2 (0%)
2	D	242/242 (100%)	0.11	5 (2%) 63 60	15, 26, 41, 58	0
2	F	242/242 (100%)	-0.15	3 (1%) 79 77	14, 19, 28, 50	1 (0%)
2	H	242/242 (100%)	0.03	5 (2%) 63 60	17, 26, 39, 55	0
All	All	1968/1992 (98%)	0.01	29 (1%) 73 71	11, 22, 38, 63	4 (0%)

The worst 5 of 29 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	G	255	ALA	5.8
2	H	241	ALA	4.6
1	E	247	LEU	4.3
2	F	241	ALA	4.1
1	E	255	ALA	3.9

### 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.

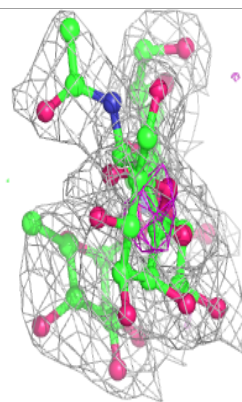
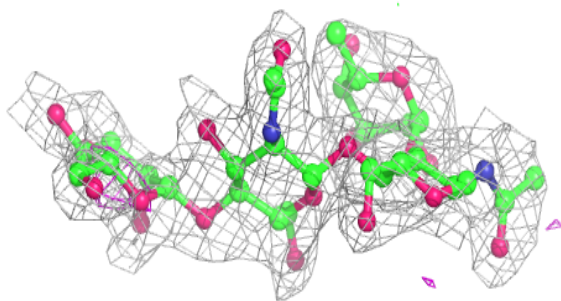
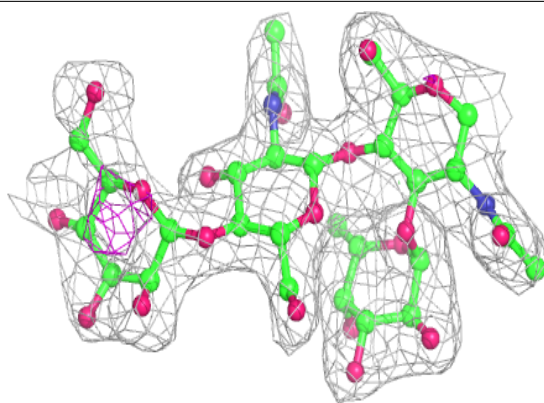
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
3	BMA	I	3	11/12	0.58	0.26	39,42,44,46	0
5	NAG	M	2	14/15	0.69	0.33	60,70,79,81	0
3	BMA	L	3	11/12	0.72	0.19	42,46,48,49	0
6	FUC	N	2	10/11	0.76	0.27	60,63,68,70	0
3	BMA	K	3	11/12	0.79	0.15	42,46,48,48	0
4	BMA	J	3	11/12	0.81	0.17	34,38,39,42	0
4	MAN	J	4	11/12	0.83	0.23	39,41,44,45	0
5	NAG	M	1	14/15	0.89	0.21	44,52,59,60	0
6	NAG	N	1	14/15	0.89	0.14	40,46,48,55	0
3	NAG	K	2	14/15	0.90	0.17	36,39,42,42	0
3	NAG	K	1	14/15	0.90	0.14	28,32,33,35	0
4	NAG	J	1	14/15	0.93	0.13	20,22,25,25	0
3	NAG	L	1	14/15	0.93	0.15	24,26,28,30	0
3	FUC	L	4	10/11	0.94	0.13	28,30,32,32	0
4	FUC	J	5	10/11	0.94	0.15	25,27,29,30	0
3	NAG	I	2	14/15	0.94	0.12	32,33,36,37	0
3	NAG	L	2	14/15	0.94	0.10	32,34,38,40	0
3	FUC	I	4	10/11	0.95	0.18	33,35,37,37	0
3	NAG	I	1	14/15	0.95	0.14	26,29,31,32	0
3	FUC	K	4	10/11	0.95	0.09	32,34,35,36	0
4	NAG	J	2	14/15	0.97	0.13	26,28,31,32	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.

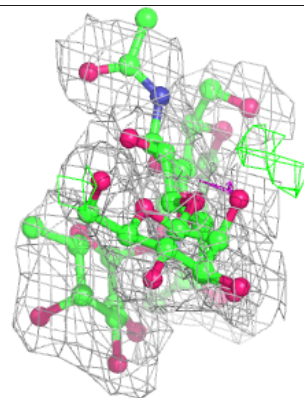
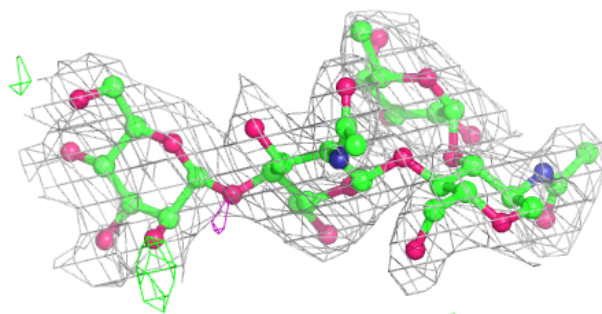
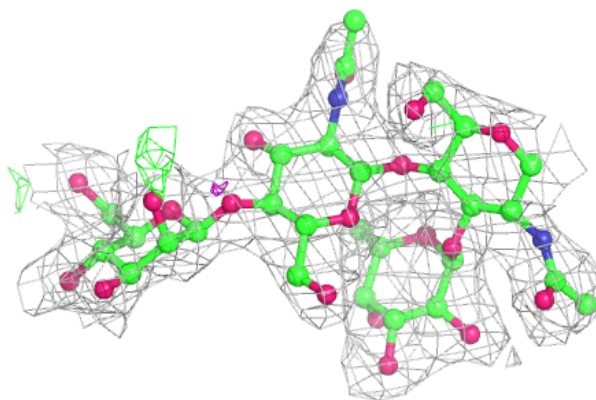


**Electron density around Chain I:**

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

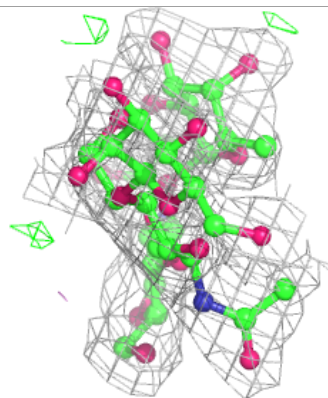
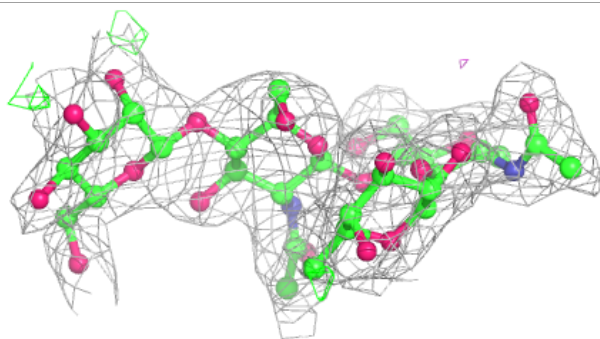
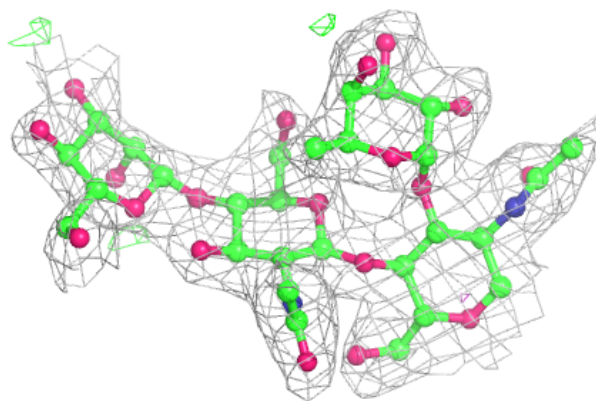
**Electron density around Chain K:**

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

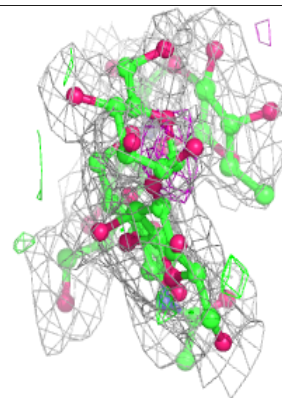
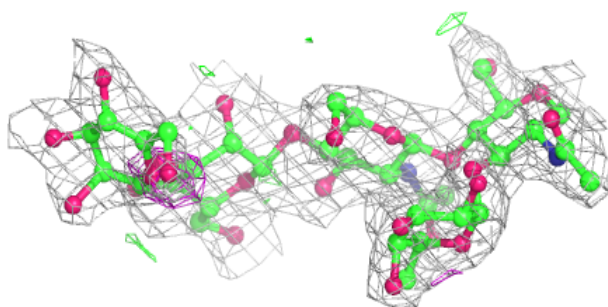
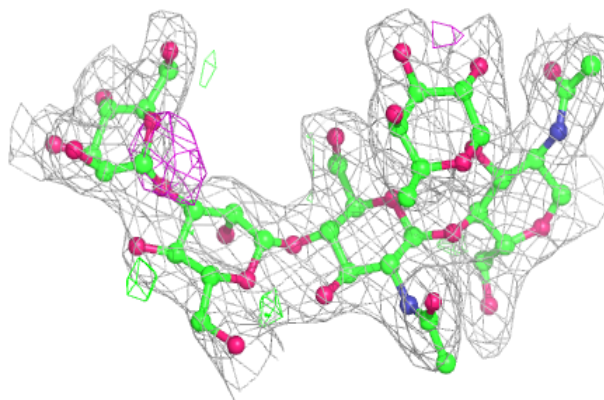


**Electron density around Chain L:**

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

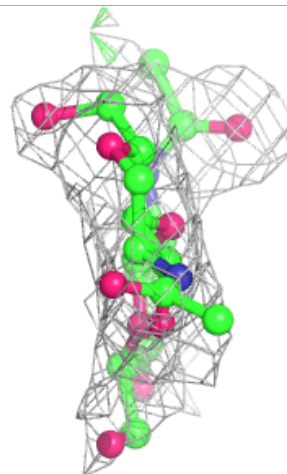
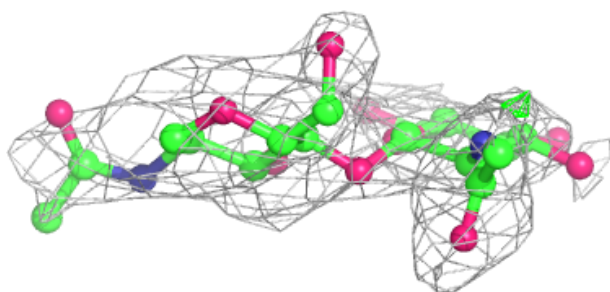
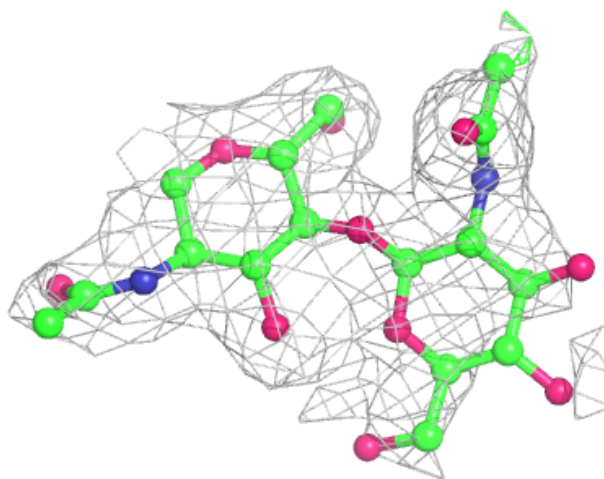
**Electron density around Chain J:**

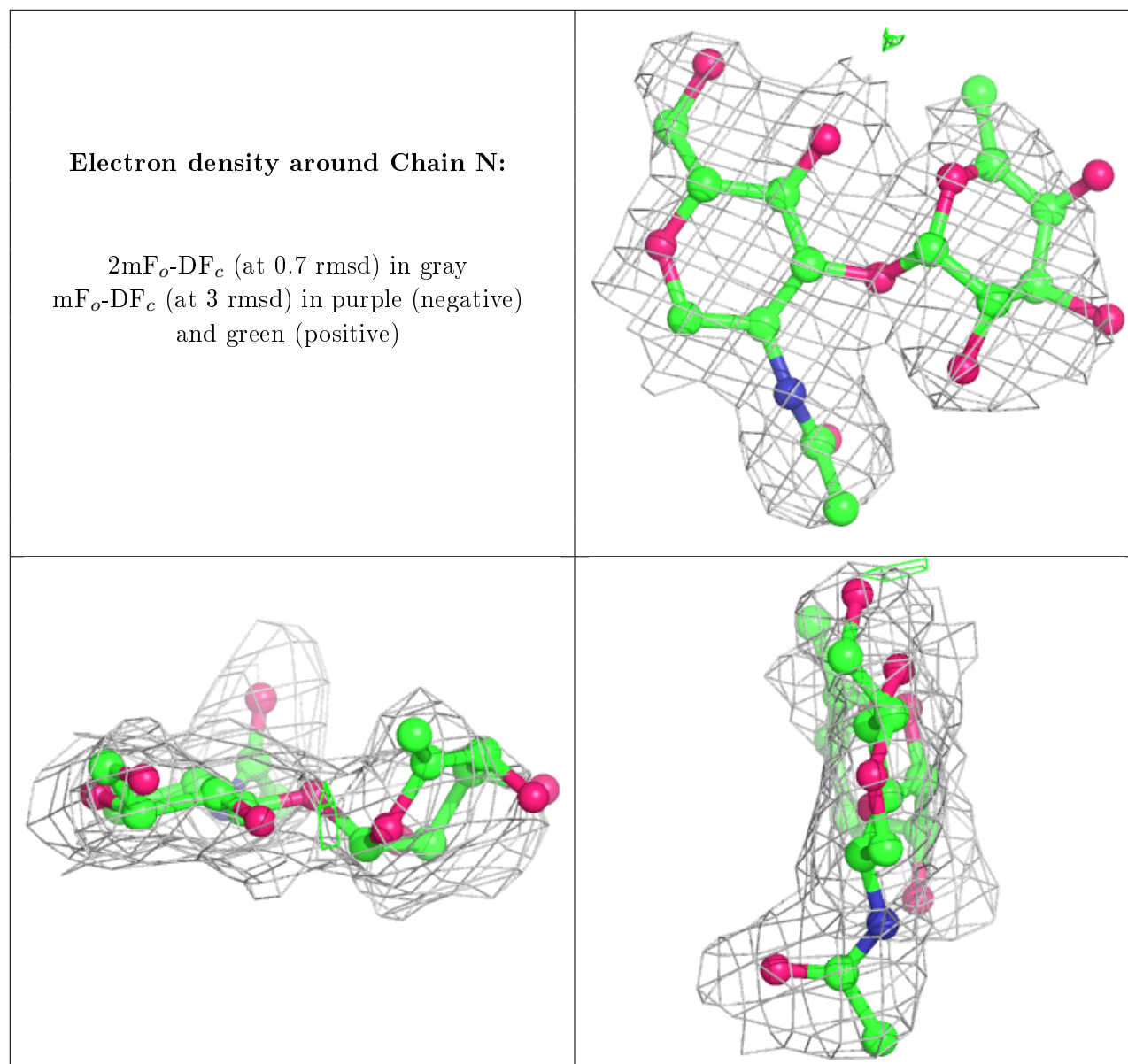
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



**Electron density around Chain M:**

$2mF_o-DF_c$  (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
12	NAG	D	301	14/15	0.77	0.23	50,56,59,60	0
11	XYP	B	306	9/10	0.85	0.17	41,41,43,43	0
9	ABU	E	284	7/7	0.88	0.26	25,28,33,33	0
10	GOL	D	285	6/6	0.89	0.20	32,33,33,34	0
10	GOL	A	277	6/6	0.90	0.18	30,33,34,36	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	GOL	H	289	6/6	0.90	0.15	29,31,31,34	0
9	ABU	D	278	7/7	0.91	0.29	29,30,33,35	0
9	ABU	B	276	7/7	0.92	0.31	27,30,34,35	0
9	ABU	E	288	7/7	0.92	0.38	32,36,38,38	0
10	GOL	E	291	6/6	0.93	0.13	29,29,31,33	0
10	GOL	E	279	6/6	0.93	0.19	31,32,33,35	0
9	ABU	F	290	7/7	0.93	0.28	14,16,20,22	0
10	GOL	C	283	6/6	0.93	0.17	25,25,26,26	0
9	ABU	C	282	7/7	0.94	0.29	29,30,33,34	0
9	ABU	A	280	7/7	0.94	0.20	22,24,27,28	0
10	GOL	G	287	6/6	0.95	0.14	29,29,30,31	0
9	ABU	G	286	7/7	0.95	0.27	31,33,35,37	0
10	GOL	B	281	6/6	0.95	0.13	20,21,22,23	0
7	CA	F	274	1/1	0.96	0.07	23,23,23,23	0
7	CA	H	272	1/1	0.97	0.11	30,30,30,30	0
7	CA	D	268	1/1	0.97	0.08	31,31,31,31	0
8	MN	C	267	1/1	0.98	0.09	20,20,20,20	0
8	MN	E	263	1/1	0.98	0.06	25,25,25,25	0
8	MN	H	273	1/1	0.98	0.09	23,23,23,23	0
7	CA	G	270	1/1	0.99	0.05	17,17,17,17	0
8	MN	B	261	1/1	0.99	0.07	13,13,13,13	0
7	CA	C	266	1/1	0.99	0.10	19,19,19,19	0
7	CA	E	262	1/1	0.99	0.06	20,20,20,20	0
8	MN	F	275	1/1	0.99	0.07	21,21,21,21	0
7	CA	A	264	1/1	0.99	0.09	19,19,19,19	0
8	MN	D	269	1/1	0.99	0.04	38,38,38,38	0
8	MN	A	265	1/1	1.00	0.10	13,13,13,13	0
7	CA	B	260	1/1	1.00	0.10	15,15,15,15	0
8	MN	G	271	1/1	1.00	0.06	19,19,19,19	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.