



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 9, 2020 – 10:08 PM BST

PDB ID : 5XEA  
Title : Structure of Thogoto virus envelope glycoprotein  
Authors : Peng, R.; Shi, Y.; Qi, J.; Gao, G.F.  
Deposited on : 2017-04-03  
Resolution : 2.09 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

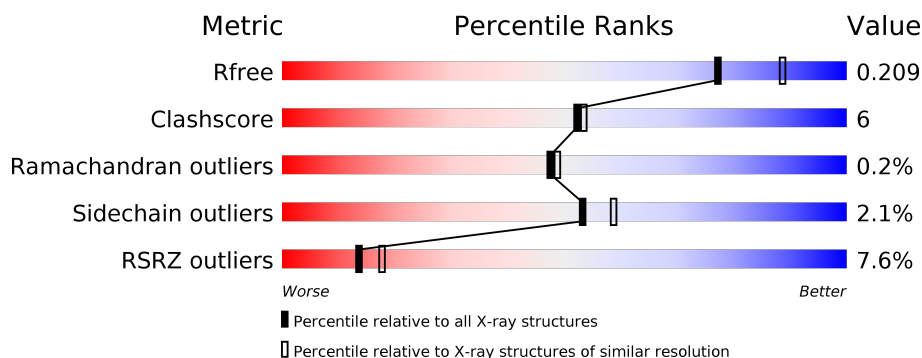
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.09 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	5197 (2.10-2.10)
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)
RSRZ outliers	127900	5083 (2.10-2.10)

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	466	<div> <div>8%</div> <div> <div></div> <div>74%</div> <div>8%</div> <div>17%</div> </div> </div>
1	B	466	<div> <div>6%</div> <div> <div></div> <div>74%</div> <div>11%</div> <div>15%</div> </div> </div>
1	C	466	<div> <div>6%</div> <div> <div></div> <div>74%</div> <div>8%</div> <div>17%</div> </div> </div>
2	D	2	<div> <div></div> <div>100%</div> </div>
2	E	2	<div> <div></div> <div> <div>50%</div> <div>50%</div> </div> </div>
2	F	2	<div> <div></div> <div> <div>50%</div> <div>50%</div> </div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	G	2	<div><div></div>100%</div>
3	H	3	<div><div></div>67%<div></div>33%</div>
3	I	3	<div><div></div>33%<div></div>33%<div></div>33%</div>

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 10621 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 Envelope glycoprotein.

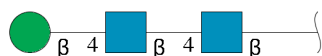
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	385	Total	C	N	O	S	0	4	0
			3108	1983	525	584	16			
1	B	397	Total	C	N	O	S	0	2	0
			3182	2029	536	600	17			
1	C	387	Total	C	N	O	S	0	4	0
			3127	1997	526	587	17			

- Molecule 2 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
2	D	2	Total	C	N	O	0	0	0
			28	16	2	10			
2	E	2	Total	C	N	O	0	0	0
			28	16	2	10			
2	F	2	Total	C	N	O	0	0	0
			28	16	2	10			
2	G	2	Total	C	N	O	0	0	0
			28	16	2	10			

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



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

- Molecule 4 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



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

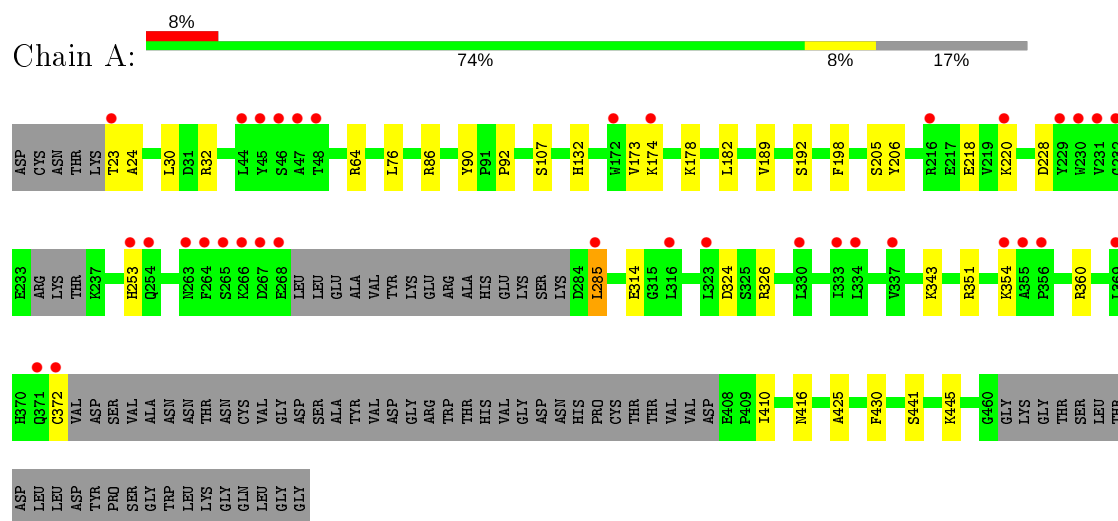
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	300	Total	O	0	0
			300	300		
5	B	318	Total	O	0	0
			318	318		
5	C	354	Total	O	0	0
			354	354		

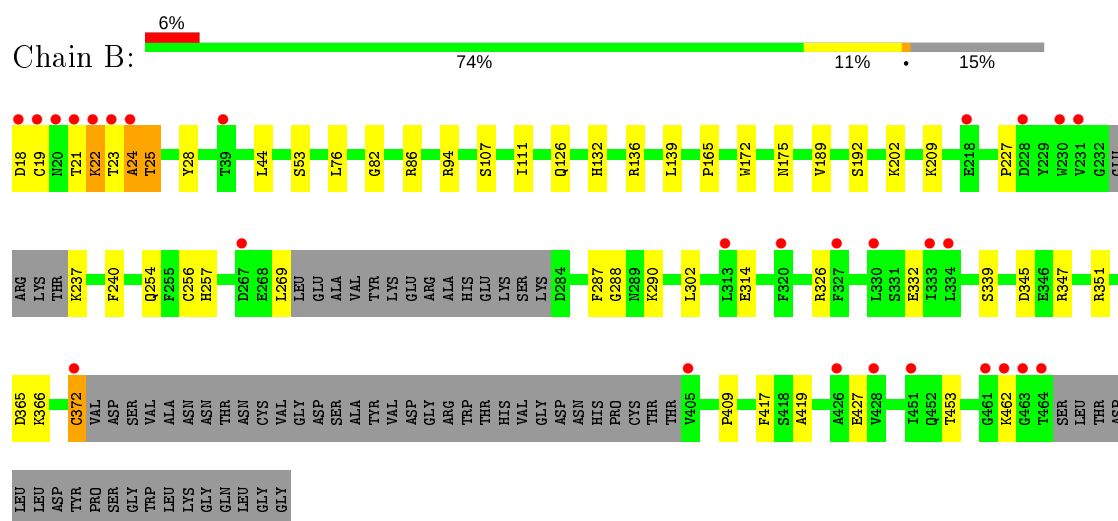
### 3 Residue-property plots [i](#)

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.

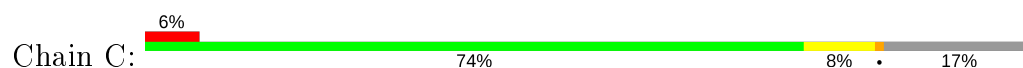
#### • Molecule 1: Envelope glycoprotein

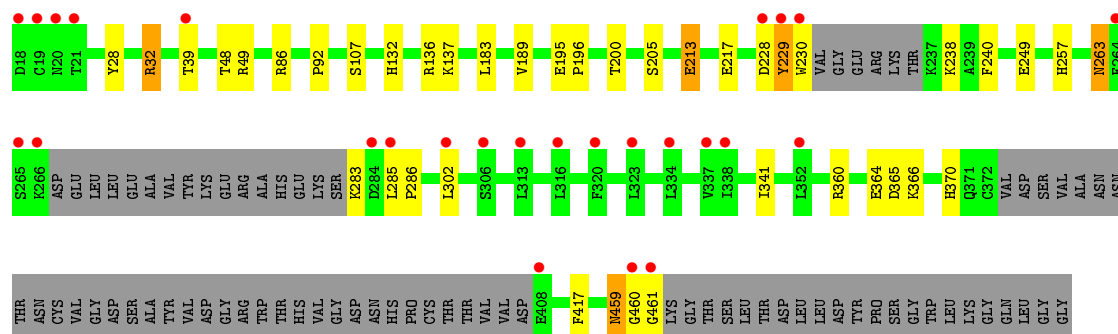


#### • Molecule 1: Envelope glycoprotein



#### • Molecule 1: Envelope glycoprotein





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

Chain D: 100%

HA61  
HA62

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

Chain E: 50%

HA61  
HA62

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

Chain F: 50%

HA61  
HA62

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

Chain G: 100%

HA61  
HA62

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

Chain H: 67%

HA61  
HA62  
BUA3

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

Chain I: 





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	106.20Å 66.77Å 120.74Å 90.00° 103.31° 90.00°	Depositor
Resolution (Å)	35.33 – 2.09 35.33 – 2.09	Depositor EDS
% Data completeness (in resolution range)	98.4 (35.33-2.09) 98.5 (35.33-2.09)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.67 (at 2.10Å)	Xtriage
Refinement program	PHENIX (1.10.1_2155)	Depositor
R, $R_{free}$	0.172 , 0.210 0.172 , 0.209	Depositor DCC
$R_{free}$ test set	4794 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	41.5	Xtriage
Anisotropy	0.614	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 56.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	10621	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	54.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.86% 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: BMA, NAG

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 5$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z  >5	RMSZ	# Z  >5
1	A	0.41	0/3203	0.54	0/4346
1	B	0.41	0/3270	0.55	0/4440
1	C	0.41	0/3221	0.56	0/4373
All	All	0.41	0/9694	0.55	0/13159

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	C	0	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	459	ASN	Peptide

### 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	3108	0	2961	29	0
1	B	3182	0	3019	47	0
1	C	3127	0	2977	36	0
2	D	28	0	25	1	0
2	E	28	0	25	2	0
2	F	28	0	25	1	0
2	G	28	0	25	1	0
3	H	39	0	34	4	0
3	I	39	0	34	3	0
4	A	28	0	26	3	0
4	C	14	0	13	0	0
5	A	300	0	0	10	0
5	B	318	0	0	19	0
5	C	354	0	0	10	0
All	All	10621	0	9164	104	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 6.

All (104) 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:32:ARG:NH2	5:A:601:HOH:O	1.81	1.12
1:A:178:LYS:NZ	5:A:602:HOH:O	2.06	0.87
1:B:372:CYS:SG	5:C:823:HOH:O	2.33	0.86
1:A:23:THR:N	5:A:603:HOH:O	2.09	0.84
1:B:136:ARG:HB2	5:B:895:HOH:O	1.78	0.84
1:C:460:GLY:O	5:C:601:HOH:O	1.96	0.83
1:B:365:ASP:OD1	5:B:702:HOH:O	1.98	0.80
1:B:202:LYS:NZ	5:B:701:HOH:O	1.98	0.80
1:B:345:ASP:OD1	5:B:703:HOH:O	2.02	0.78
1:B:419:ALA:O	5:B:704:HOH:O	2.03	0.77
1:B:24:ALA:HB1	1:B:25:THR:HB	1.70	0.74
3:I:1:NAG:H3	3:I:1:NAG:H83	1.69	0.74
1:B:175:ASN:ND2	5:B:706:HOH:O	2.14	0.72
1:C:360:ARG:NH2	5:C:606:HOH:O	2.25	0.70
1:C:213:GLU:OE2	5:C:602:HOH:O	2.10	0.70
1:C:195:GLU:OE2	5:C:603:HOH:O	2.13	0.65
5:C:604:HOH:O	3:H:1:NAG:O6	2.15	0.65
4:A:504:NAG:H81	5:C:851:HOH:O	1.95	0.65
1:A:218:GLU:O	5:A:605:HOH:O	2.15	0.64
1:B:288:GLY:O	5:B:707:HOH:O	2.15	0.63

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:257:HIS:HE1	5:B:881:HOH:O	1.82	0.60
1:B:351[B]:ARG:NH1	1:C:364:GLU:OE2	2.35	0.59
1:C:263:ASN:OD1	1:C:263:ASN:N	2.35	0.59
1:B:139:LEU:O	1:B:453:THR:HG21	2.03	0.58
1:B:209:LYS:NZ	5:B:718:HOH:O	2.36	0.58
1:B:347:ARG:NH1	1:B:351[B]:ARG:HH22	2.03	0.57
1:B:351[B]:ARG:NH2	5:B:716:HOH:O	2.35	0.57
1:B:227:PRO:HA	1:B:237:LYS:N	2.20	0.57
1:C:213:GLU:OE2	5:C:605:HOH:O	2.17	0.57
3:I:1:NAG:H3	3:I:1:NAG:C8	2.35	0.57
1:B:326:ARG:HG2	1:C:417:PHE:CD1	2.41	0.56
1:A:253[A]:HIS:ND1	5:A:616:HOH:O	2.33	0.55
1:A:189:VAL:HG21	2:D:2:NAG:H81	1.88	0.55
1:A:351:ARG:NH2	5:A:627:HOH:O	2.42	0.52
1:C:48:THR:OG1	3:I:1:NAG:H82	2.10	0.52
1:B:269:LEU:O	5:B:708:HOH:O	2.18	0.51
1:A:182:LEU:O	5:A:607:HOH:O	2.19	0.51
1:B:24:ALA:HB3	5:B:736:HOH:O	2.12	0.50
1:B:94:ARG:HD2	5:B:720:HOH:O	2.11	0.50
1:C:238:LYS:NZ	1:C:249:GLU:OE2	2.32	0.49
1:C:196:PRO:HA	5:C:761:HOH:O	2.12	0.49
1:A:285:LEU:HD22	1:B:165:PRO:HG3	1.95	0.49
1:B:351[B]:ARG:NE	5:B:716:HOH:O	2.43	0.49
1:A:173:VAL:HG12	1:A:174:LYS:HD2	1.96	0.48
1:A:430:PHE:H	1:C:283:LYS:NZ	2.11	0.48
1:C:39:THR:OG1	1:C:228:ASP:OD2	2.21	0.48
1:B:25:THR:HG23	1:B:28:TYR:HE2	1.78	0.47
1:C:285[A]:LEU:H	1:C:285[A]:LEU:HD23	1.78	0.47
1:B:189:VAL:HG21	2:E:2:NAG:H81	1.97	0.47
1:C:285[A]:LEU:HG	1:C:286:PRO:HD3	1.96	0.47
1:A:351:ARG:HH22	1:B:339:SER:HB3	1.78	0.46
1:C:48:THR:HB	1:C:263:ASN:OD1	2.16	0.46
4:A:503:NAG:H2	5:A:611:HOH:O	2.14	0.46
1:B:287:PHE:O	1:B:290:LYS:HG3	2.15	0.46
1:A:30:LEU:O	1:A:32:ARG:NH1	2.49	0.46
1:B:366:LYS:NZ	5:B:713:HOH:O	2.31	0.46
1:B:347:ARG:HH11	1:B:351[B]:ARG:HH22	1.63	0.46
1:A:326:ARG:HG2	1:B:417:PHE:CD1	2.51	0.46
1:A:343:LYS:HE2	5:A:613:HOH:O	2.17	0.45
1:A:24:ALA:HB2	1:C:370:HIS:CD2	2.52	0.45
1:A:314:GLU:HG3	1:B:240:PHE:CD1	2.51	0.45

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:302:LEU:HD11	1:C:302:LEU:HD11	1.99	0.45
1:B:21:THR:O	1:B:23:THR:N	2.49	0.45
1:B:237:LYS:HB2	1:B:237:LYS:HE3	1.69	0.44
1:B:126:GLN:HG2	5:B:933:HOH:O	2.17	0.44
1:C:189:VAL:HG21	3:H:2:NAG:H81	1.99	0.44
1:C:92:PRO:HG2	1:C:205:SER:HB2	1.98	0.44
1:A:354:LYS:O	1:A:354:LYS:HG3	2.17	0.44
1:B:314:GLU:HG3	1:C:240:PHE:CD1	2.53	0.44
1:B:18:ASP:HA	1:B:19:CYS:HA	1.69	0.44
2:E:1:NAG:H61	2:E:2:NAG:C1	2.48	0.44
1:C:200:THR:O	3:H:3:BMA:O2	2.09	0.44
1:B:332:GLU:HG2	1:C:32:ARG:HG2	1.99	0.43
1:A:441:SER:O	1:A:445:LYS:HG3	2.19	0.43
1:B:172:TRP:CD1	1:B:256:CYS:SG	3.12	0.43
1:B:24:ALA:HB1	1:B:25:THR:CB	2.44	0.43
1:B:347:ARG:CZ	5:B:703:HOH:O	2.67	0.43
1:C:459:ASN:HB3	1:C:461:GLY:N	2.33	0.42
1:B:254:GLN:NE2	5:B:732:HOH:O	2.51	0.42
1:C:257:HIS:HE1	5:C:698:HOH:O	2.01	0.42
1:C:183:LEU:HD13	3:H:1:NAG:H62	2.01	0.42
1:A:92:PRO:HG2	1:A:205:SER:HB2	2.01	0.42
1:C:136:ARG:HG3	1:C:137:LYS:N	2.34	0.42
1:A:220:LYS:N	5:A:620:HOH:O	2.38	0.42
1:B:21:THR:OG1	1:B:22:LYS:N	2.53	0.42
1:A:107:SER:HA	1:A:132:HIS:O	2.20	0.42
1:C:49:ARG:NH2	1:C:217:GLU:OE1	2.50	0.42
1:A:76:LEU:HD23	1:A:76:LEU:HA	1.90	0.41
2:G:1:NAG:H62	2:G:2:NAG:O5	2.20	0.41
1:A:425:ALA:HB3	1:C:229:TYR:OH	2.21	0.41
1:B:462:LYS:NZ	5:B:735:HOH:O	2.53	0.41
1:B:107:SER:HA	1:B:132:HIS:O	2.21	0.41
1:C:341:ILE:HG13	1:C:341:ILE:H	1.73	0.41
1:A:416:ASN:ND2	4:A:504:NAG:H83	2.35	0.41
1:A:64:ARG:HB3	1:A:90:TYR:O	2.21	0.41
1:C:107:SER:HA	1:C:132:HIS:O	2.21	0.41
1:A:410:ILE:HB	1:C:28:TYR:CD1	2.56	0.41
1:B:82:GLY:HA3	1:C:460:GLY:O	2.21	0.41
1:B:409:PRO:HG3	1:C:366:LYS:HE3	2.03	0.40
1:B:76:LEU:HD23	1:B:76:LEU:HA	1.83	0.40
1:A:425:ALA:O	1:C:229:TYR:OH	2.39	0.40
1:A:198:PHE:HA	1:A:206:TYR:O	2.21	0.40

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:F:1:NAG:O4	2:F:2:NAG:H61	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	381/466 (82%)	365 (96%)	16 (4%)	0	100	100
1	B	391/466 (84%)	375 (96%)	14 (4%)	2 (0%)	29	26
1	C	383/466 (82%)	368 (96%)	15 (4%)	0	100	100
All	All	1155/1398 (83%)	1108 (96%)	45 (4%)	2 (0%)	47	49

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	22	LYS
1	B	24	ALA

### 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	338/404 (84%)	331 (98%)	7 (2%)	53	59
1	B	345/404 (85%)	337 (98%)	8 (2%)	50	55

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	C	341/404 (84%)	335 (98%)	6 (2%)	59	65
All	All	1024/1212 (84%)	1003 (98%)	21 (2%)	53	59

All (21) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	86	ARG
1	A	192	SER
1	A	228	ASP
1	A	285	LEU
1	A	324	ASP
1	A	360	ARG
1	A	372	CYS
1	B	25	THR
1	B	44	LEU
1	B	53	SER
1	B	86	ARG
1	B	111	ILE
1	B	192	SER
1	B	372	CYS
1	B	427	GLU
1	C	32	ARG
1	C	86	ARG
1	C	213	GLU
1	C	229	TYR
1	C	230	TRP
1	C	263	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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 ⓘ

14 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	D	1	1,2	14,14,15	0.35	0	17,19,21	0.68	1 (5%)
2	NAG	D	2	2	14,14,15	0.45	0	17,19,21	0.54	0
2	NAG	E	1	1,2	14,14,15	0.52	0	17,19,21	0.78	0
2	NAG	E	2	2	14,14,15	1.59	1 (7%)	17,19,21	1.91	1 (5%)
2	NAG	F	1	1,2	14,14,15	0.28	0	17,19,21	0.61	0
2	NAG	F	2	2	14,14,15	0.72	1 (7%)	17,19,21	1.16	1 (5%)
2	NAG	G	1	1,2	14,14,15	0.48	0	17,19,21	1.05	2 (11%)
2	NAG	G	2	2	14,14,15	0.70	1 (7%)	17,19,21	0.76	0
3	NAG	H	1	1,3	14,14,15	0.56	0	17,19,21	0.56	0
3	NAG	H	2	3	14,14,15	0.53	0	17,19,21	0.75	1 (5%)
3	BMA	H	3	3	11,11,12	0.26	0	15,15,17	0.64	0
3	NAG	I	1	1,3	14,14,15	1.18	1 (7%)	17,19,21	1.99	4 (23%)
3	NAG	I	2	3	14,14,15	0.47	0	17,19,21	0.59	0
3	BMA	I	3	3	11,11,12	0.93	1 (9%)	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	D	1	1,2	-	0/6/23/26	0/1/1/1
2	NAG	D	2	2	-	0/6/23/26	0/1/1/1
2	NAG	E	1	1,2	-	2/6/23/26	0/1/1/1
2	NAG	E	2	2	-	2/6/23/26	0/1/1/1
2	NAG	F	1	1,2	-	0/6/23/26	0/1/1/1
2	NAG	F	2	2	-	1/6/23/26	0/1/1/1
2	NAG	G	1	1,2	-	0/6/23/26	0/1/1/1
2	NAG	G	2	2	-	4/6/23/26	0/1/1/1

*Continued on next page...*



Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	NAG	H	1	1,3	-	0/6/23/26	0/1/1/1
3	NAG	H	2	3	-	1/6/23/26	0/1/1/1
3	BMA	H	3	3	-	0/2/19/22	0/1/1/1
3	NAG	I	1	1,3	-	3/6/23/26	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

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	E	2	NAG	O5-C1	5.84	1.53	1.43
3	I	1	NAG	O5-C1	-4.15	1.37	1.43
3	I	3	BMA	O5-C1	-2.70	1.39	1.43
2	F	2	NAG	O5-C1	2.37	1.47	1.43
2	G	2	NAG	O5-C1	2.07	1.47	1.43

All (10) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	E	2	NAG	C1-O5-C5	7.62	122.52	112.19
3	I	1	NAG	C2-N2-C7	5.44	130.65	122.90
2	F	2	NAG	C1-O5-C5	4.39	118.14	112.19
3	I	1	NAG	C3-C4-C5	3.71	116.86	110.24
3	I	1	NAG	C1-C2-N2	3.12	115.83	110.49
2	G	1	NAG	C1-O5-C5	2.25	115.25	112.19
3	H	2	NAG	O4-C4-C3	-2.25	105.16	110.35
2	D	1	NAG	C1-O5-C5	2.06	114.98	112.19
3	I	1	NAG	O4-C4-C5	-2.04	104.23	109.30
2	G	1	NAG	O4-C4-C3	-2.00	105.72	110.35

There are no chirality outliers.

All (15) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	G	2	NAG	O5-C5-C6-O6
2	E	2	NAG	C4-C5-C6-O6
2	E	2	NAG	O5-C5-C6-O6
2	G	2	NAG	C4-C5-C6-O6
3	I	1	NAG	C8-C7-N2-C2
3	I	1	NAG	O7-C7-N2-C2
3	I	2	NAG	C4-C5-C6-O6

Continued on next page...

*Continued from previous page...*

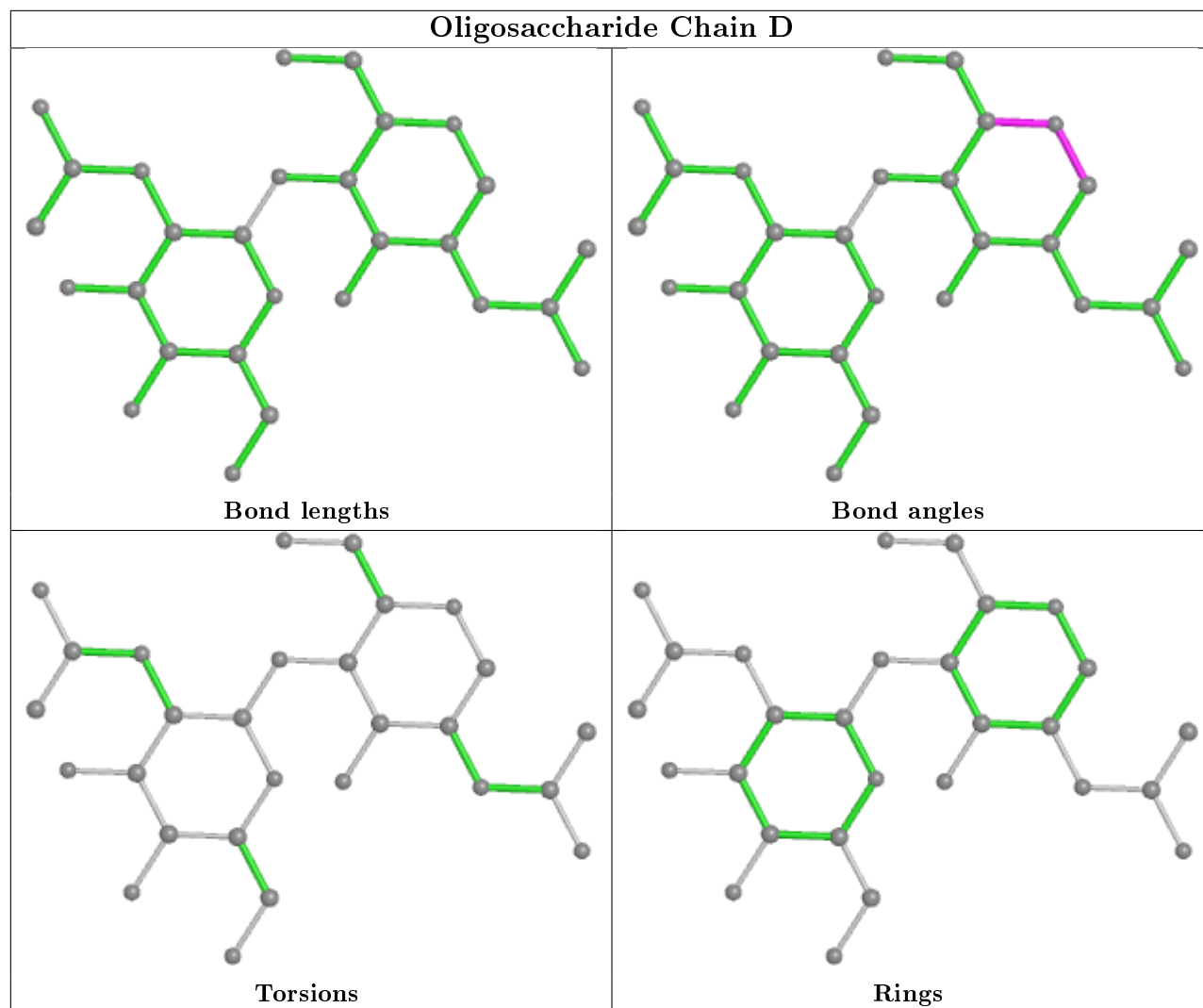
Mol	Chain	Res	Type	Atoms
2	E	1	NAG	O5-C5-C6-O6
2	E	1	NAG	C4-C5-C6-O6
3	I	1	NAG	C3-C2-N2-C7
3	I	2	NAG	O5-C5-C6-O6
2	G	2	NAG	C1-C2-N2-C7
3	H	2	NAG	C4-C5-C6-O6
2	F	2	NAG	C4-C5-C6-O6
2	G	2	NAG	C3-C2-N2-C7

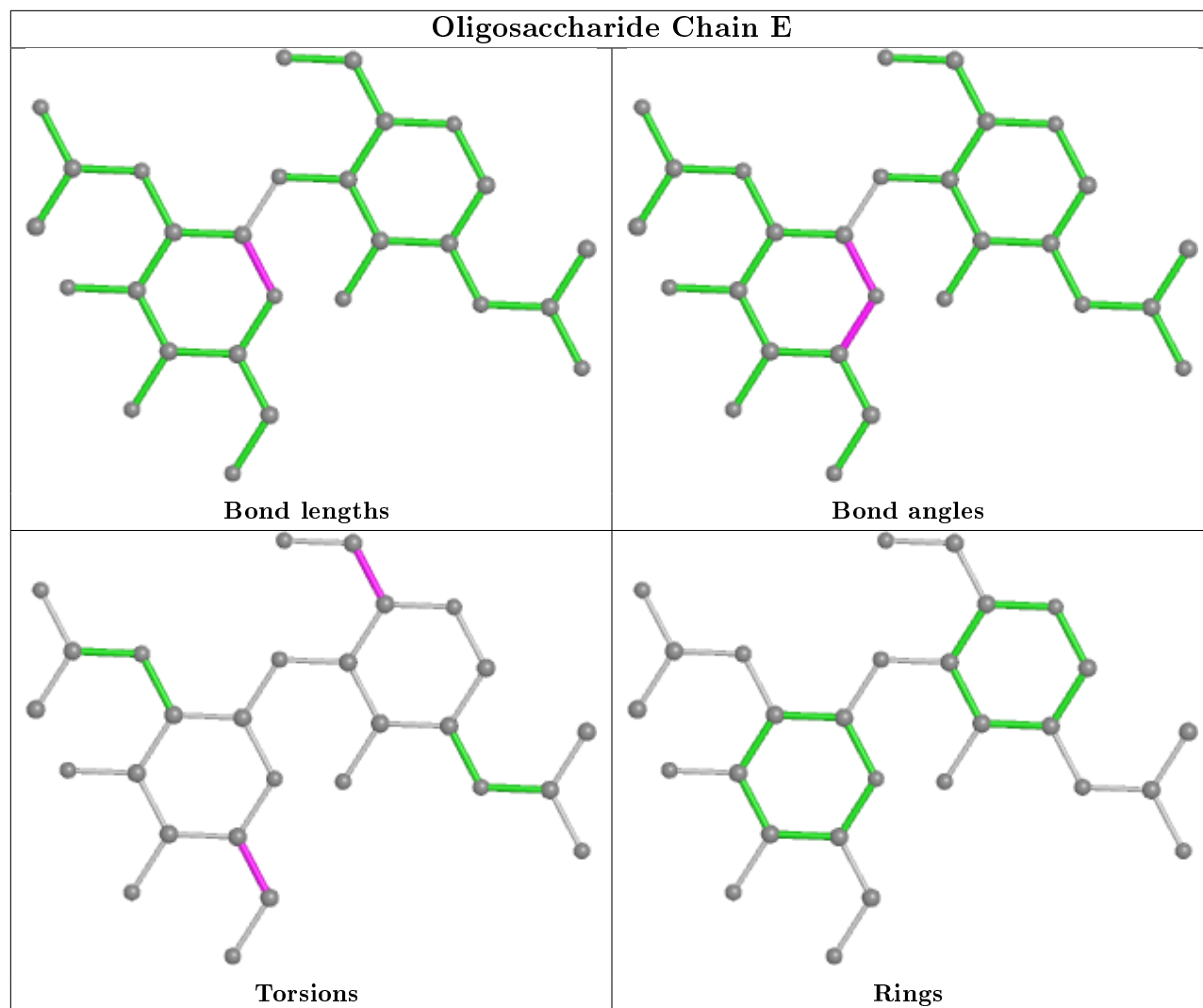
There are no ring outliers.

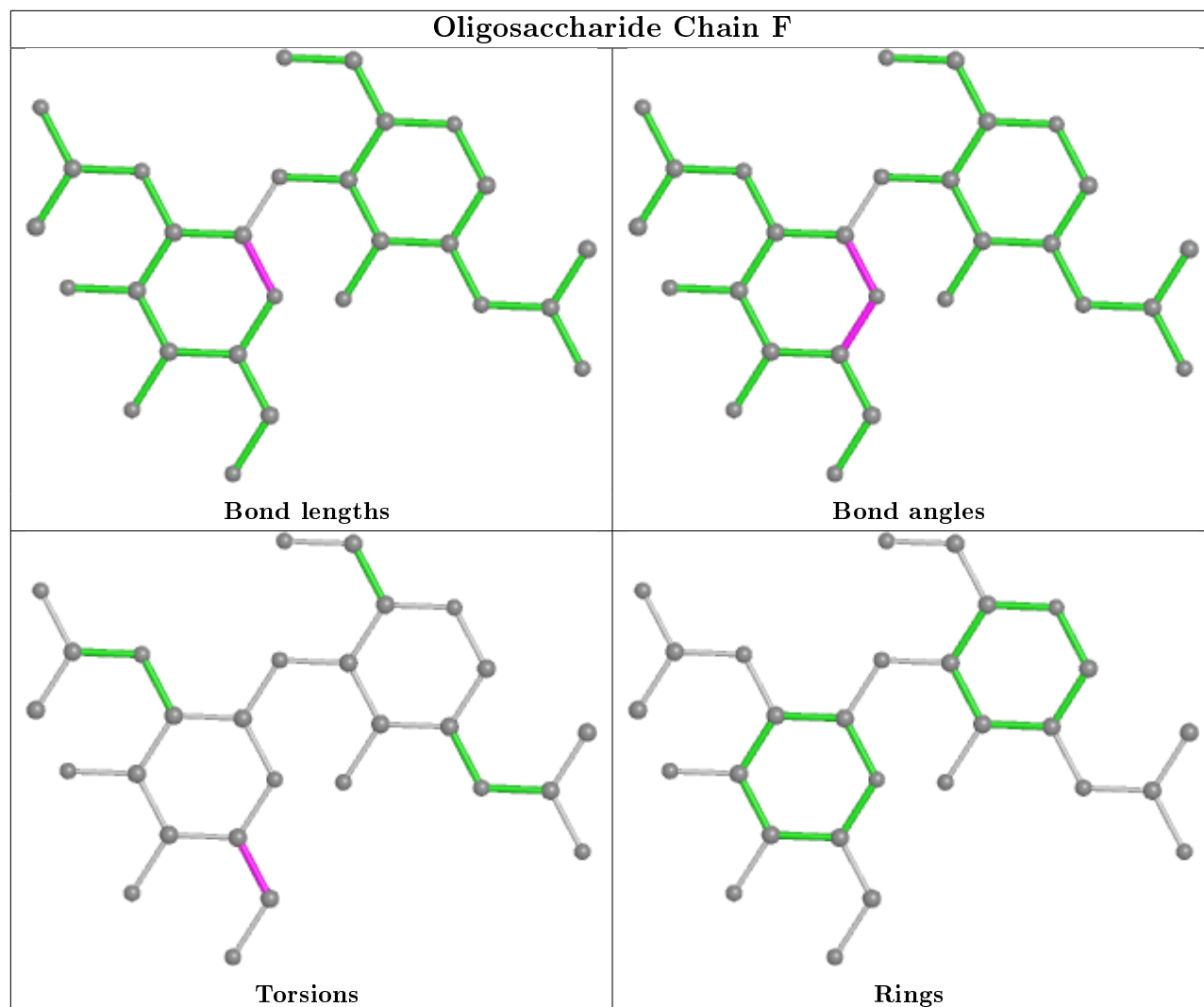
11 monomers are involved in 12 short contacts:

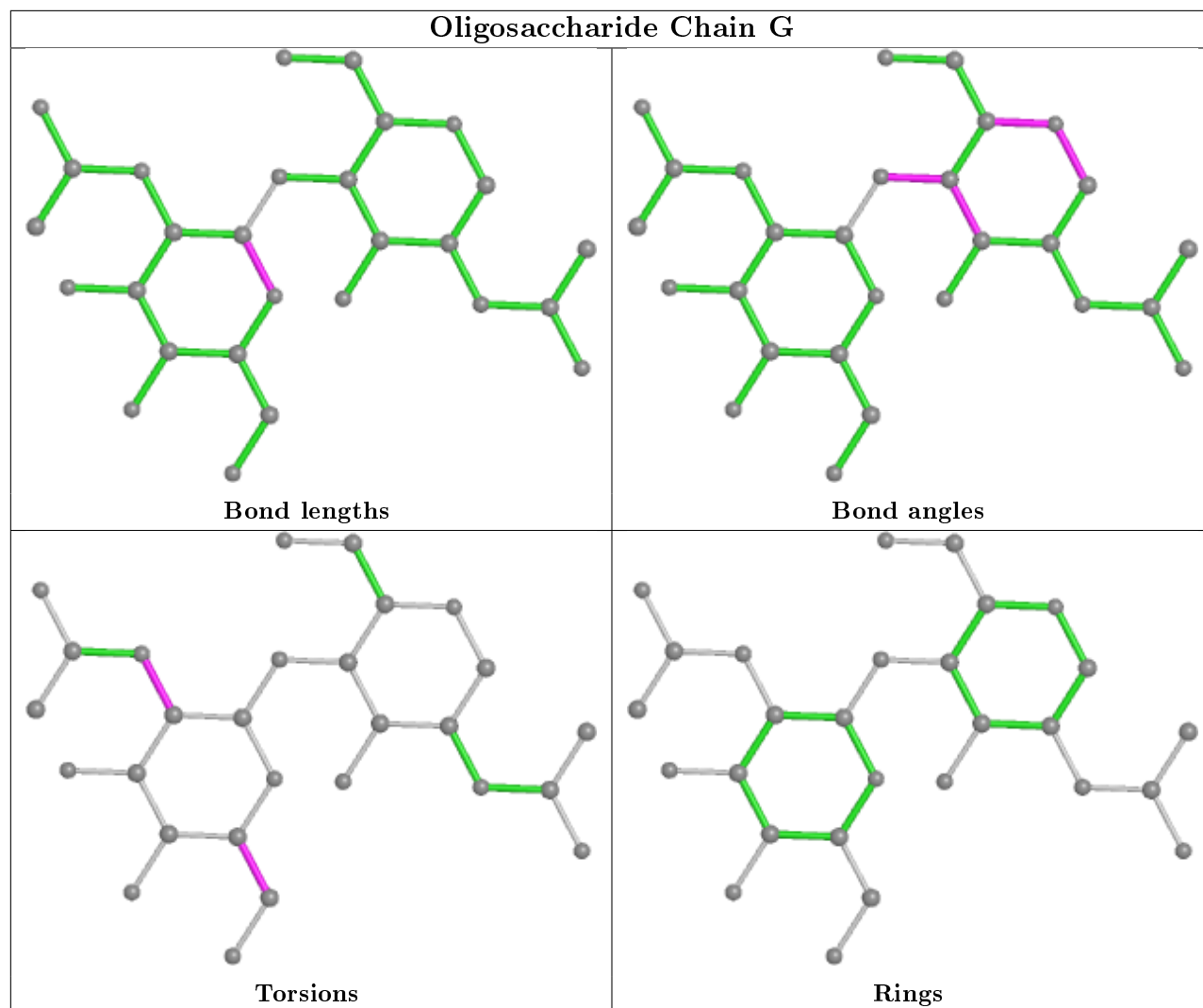
Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	G	2	NAG	1	0
2	E	1	NAG	1	0
2	G	1	NAG	1	0
2	E	2	NAG	2	0
2	F	2	NAG	1	0
2	D	2	NAG	1	0
3	I	1	NAG	3	0
2	F	1	NAG	1	0
3	H	1	NAG	2	0
3	H	2	NAG	1	0
3	H	3	BMA	1	0

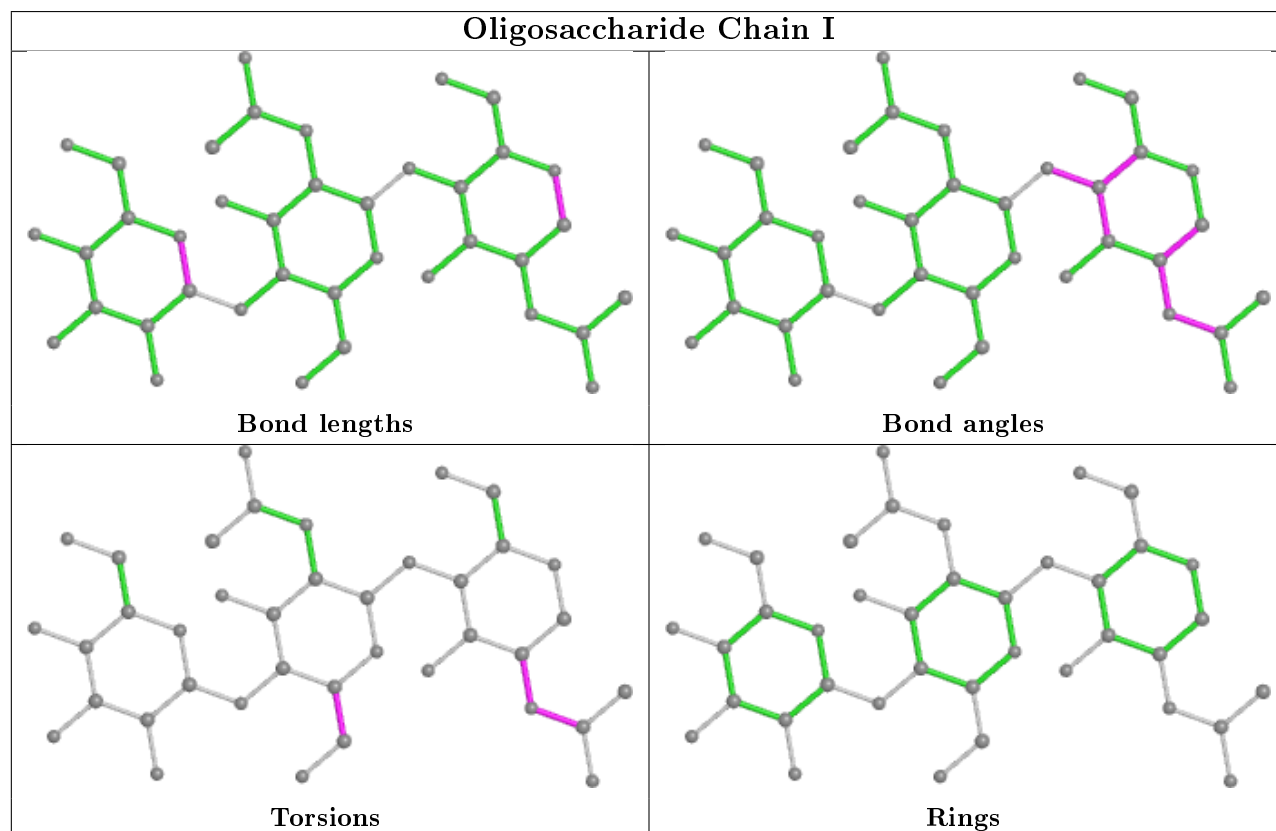
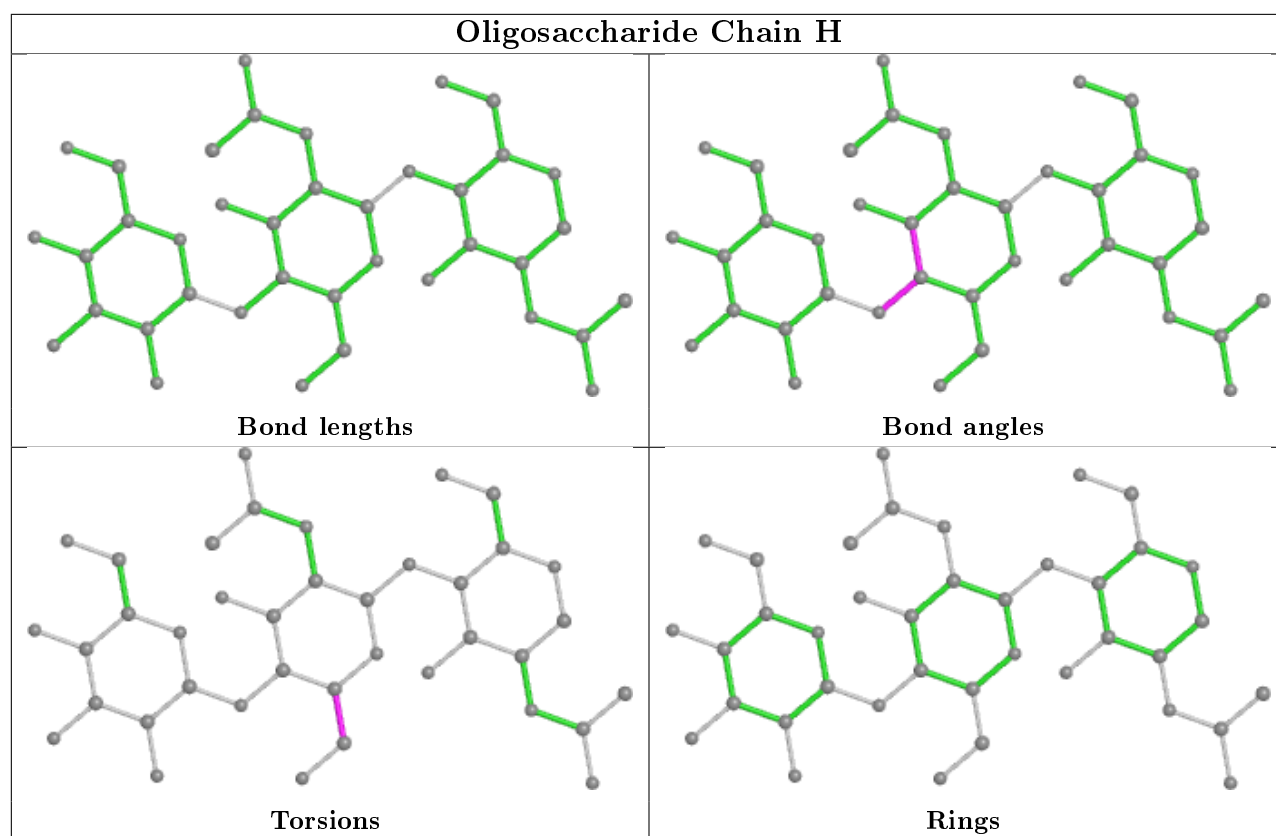
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for oligosaccharide.











## 5.6 Ligand geometry

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
4	NAG	A	503	1	14,14,15	0.54	0	17,19,21	0.58	0
4	NAG	A	504	1	14,14,15	0.97	1 (7%)	17,19,21	1.40	1 (5%)
4	NAG	C	507	1	14,14,15	0.42	0	17,19,21	0.39	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
4	NAG	A	503	1	-	0/6/23/26	0/1/1/1
4	NAG	A	504	1	-	4/6/23/26	0/1/1/1
4	NAG	C	507	1	-	2/6/23/26	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	504	NAG	O5-C1	3.50	1.49	1.43

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	504	NAG	C1-O5-C5	5.47	119.61	112.19

There are no chirality outliers.

All (6) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	C	507	NAG	O5-C5-C6-O6

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms
4	C	507	NAG	C4-C5-C6-O6
4	A	504	NAG	C8-C7-N2-C2
4	A	504	NAG	O7-C7-N2-C2
4	A	504	NAG	C4-C5-C6-O6
4	A	504	NAG	O5-C5-C6-O6

There are no ring outliers.

2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	503	NAG	1	0
4	A	504	NAG	2	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

### 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	385/466 (82%)	0.22	35 (9%) 9 12	29, 48, 88, 132	0
1	B	397/466 (85%)	0.23	28 (7%) 16 20	32, 49, 91, 155	0
1	C	387/466 (83%)	0.13	26 (6%) 17 22	28, 46, 84, 140	0
All	All	1169/1398 (83%)	0.19	89 (7%) 13 18	28, 48, 89, 155	0

All (89) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	19	CYS	10.2
1	B	464	THR	8.9
1	A	230	TRP	8.9
1	C	18	ASP	8.7
1	B	231	VAL	7.4
1	C	230	TRP	7.3
1	B	230	TRP	6.4
1	B	24	ALA	6.3
1	B	405	VAL	6.1
1	B	18	ASP	5.8
1	B	23	THR	5.8
1	B	19	CYS	5.3
1	A	231	VAL	5.3
1	A	267	ASP	5.2
1	C	460	GLY	4.7
1	C	284	ASP	4.5
1	A	253[A]	HIS	4.4
1	A	264	PHE	4.3
1	B	39	THR	4.2
1	C	266	LYS	4.2
1	A	285	LEU	4.0
1	C	21	THR	4.0
1	A	356	PRO	4.0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	20	ASN	3.8
1	B	461	GLY	3.5
1	C	228	ASP	3.5
1	B	20	ASN	3.4
1	A	232	GLY	3.3
1	A	268	GLU	3.2
1	C	316	LEU	3.2
1	C	338	ILE	3.2
1	A	46	SER	3.2
1	A	354	LYS	3.2
1	A	44	LEU	3.1
1	A	330	LEU	3.1
1	A	47	ALA	3.0
1	C	334	LEU	3.0
1	C	229	TYR	3.0
1	A	45	TYR	3.0
1	A	216	ARG	3.0
1	A	220	LYS	3.0
1	A	372	CYS	2.9
1	A	48	THR	2.9
1	C	285[A]	LEU	2.9
1	C	39	THR	2.8
1	A	371	GLN	2.8
1	B	463	GLY	2.8
1	A	334	LEU	2.8
1	A	266	LYS	2.7
1	A	355	ALA	2.6
1	B	327	PHE	2.6
1	B	462	LYS	2.6
1	B	267	ASP	2.6
1	A	337	VAL	2.5
1	A	174	LYS	2.5
1	C	313	LEU	2.5
1	B	21	THR	2.5
1	A	323	LEU	2.5
1	B	313	LEU	2.4
1	A	172	TRP	2.4
1	B	330	LEU	2.4
1	C	323	LEU	2.4
1	B	372	CYS	2.4
1	B	320	PHE	2.3
1	C	461	GLY	2.3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	228	ASP	2.3
1	C	337	VAL	2.3
1	B	451	ILE	2.3
1	C	408	GLU	2.3
1	C	306	SER	2.3
1	C	264	PHE	2.2
1	B	218	GLU	2.2
1	A	265	SER	2.2
1	B	22	LYS	2.2
1	B	333	ILE	2.2
1	C	352	LEU	2.2
1	A	316	LEU	2.2
1	A	254	GLN	2.2
1	B	426	ALA	2.2
1	C	265	SER	2.2
1	A	263	ASN	2.1
1	A	23	THR	2.1
1	C	320	PHE	2.1
1	A	333	ILE	2.1
1	A	369	LEU	2.1
1	C	302	LEU	2.1
1	B	428	VAL	2.0
1	A	229	TYR	2.0
1	B	334	LEU	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.

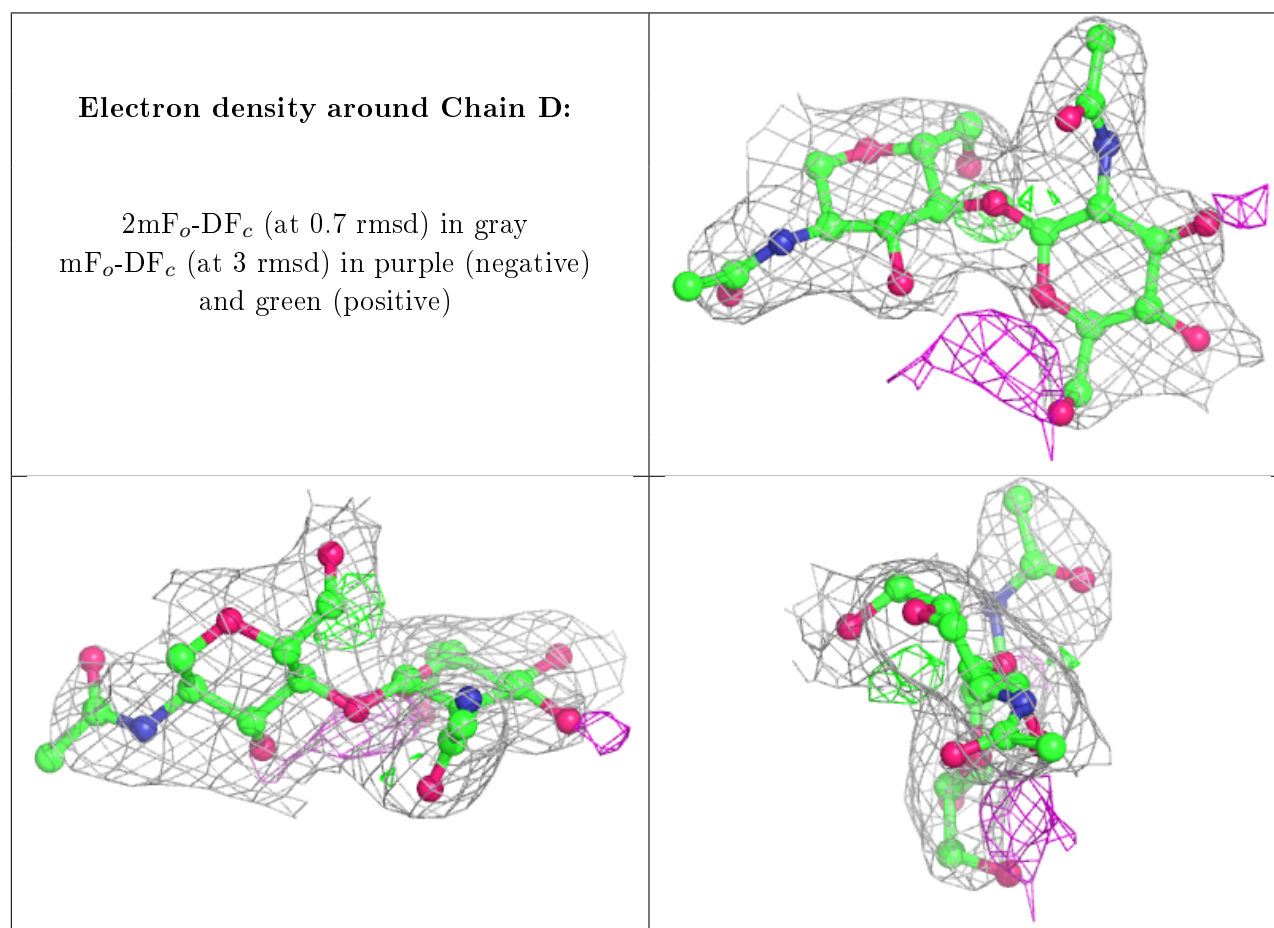
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
2	NAG	F	2	14/15	0.68	0.33	111,116,120,121	0
3	BMA	H	3	11/12	0.70	0.36	75,87,94,100	0
3	BMA	I	3	11/12	0.73	0.21	93,97,106,106	0
2	NAG	E	2	14/15	0.73	0.26	51,80,94,97	0

*Continued on next page...*

*Continued from previous page...*

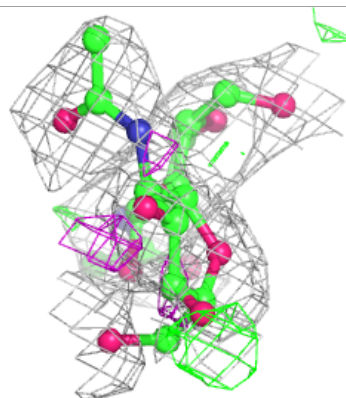
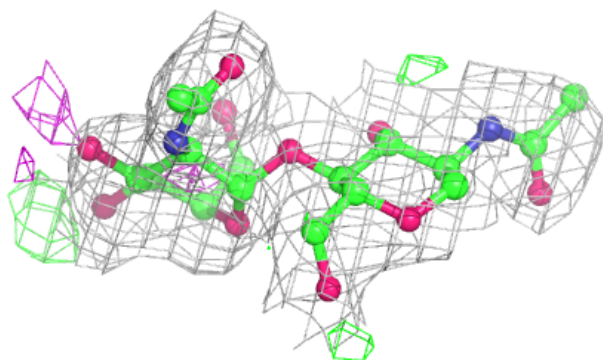
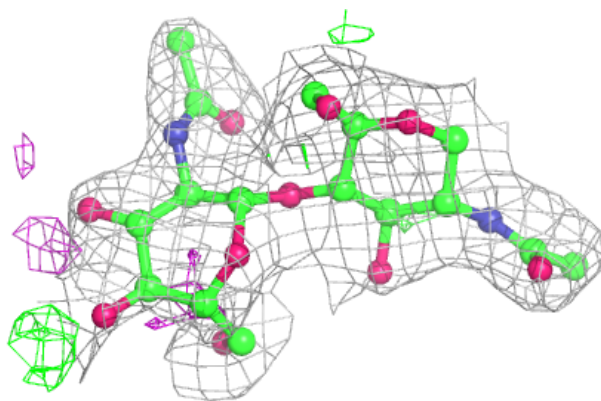
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
3	NAG	I	1	14/15	0.78	0.30	87,101,105,110	0
3	NAG	H	2	14/15	0.81	0.21	56,79,87,89	0
3	NAG	I	2	14/15	0.83	0.29	95,107,112,115	0
2	NAG	F	1	14/15	0.86	0.18	80,95,106,112	0
2	NAG	G	2	14/15	0.86	0.34	105,114,120,121	0
2	NAG	D	2	14/15	0.88	0.21	54,67,78,83	0
2	NAG	G	1	14/15	0.88	0.16	62,83,90,98	0
2	NAG	E	1	14/15	0.93	0.10	44,69,77,84	0
2	NAG	D	1	14/15	0.94	0.11	45,70,79,81	0
3	NAG	H	1	14/15	0.96	0.10	45,72,78,83	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.

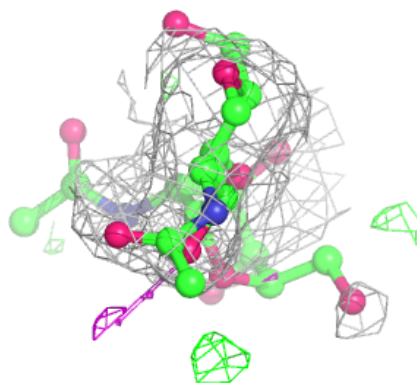
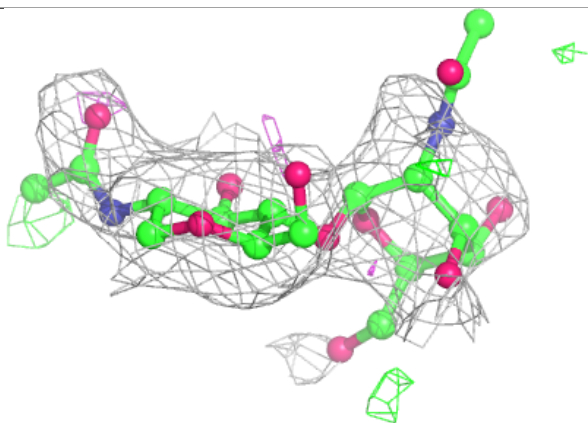
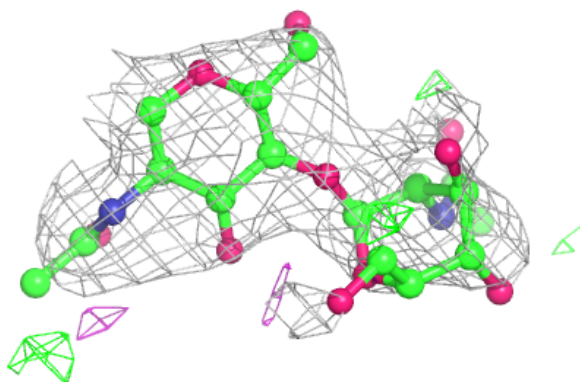


**Electron density around Chain E:**

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

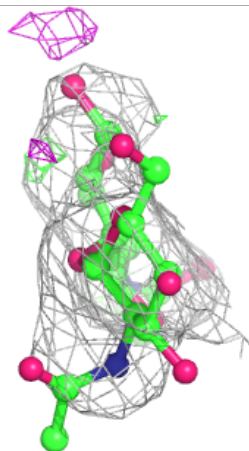
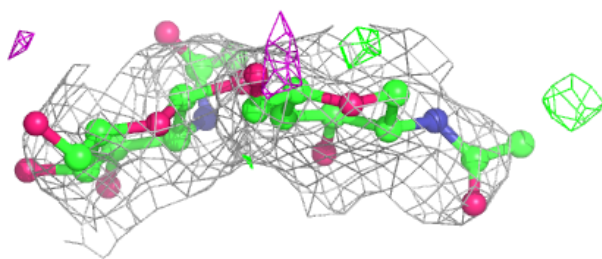
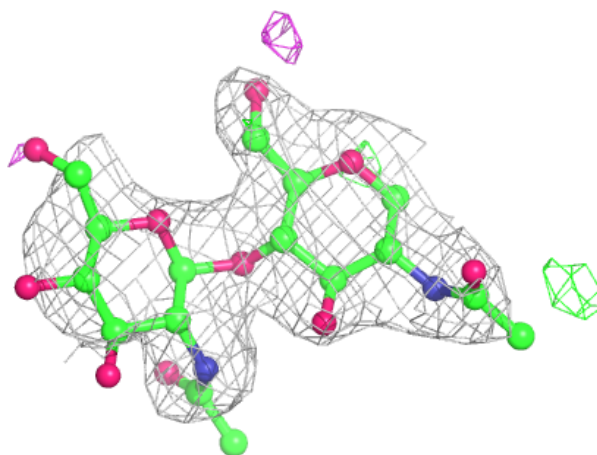
**Electron density around Chain F:**

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



**Electron density around Chain G:**

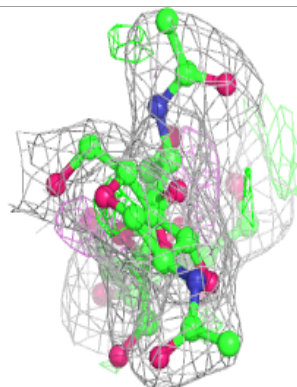
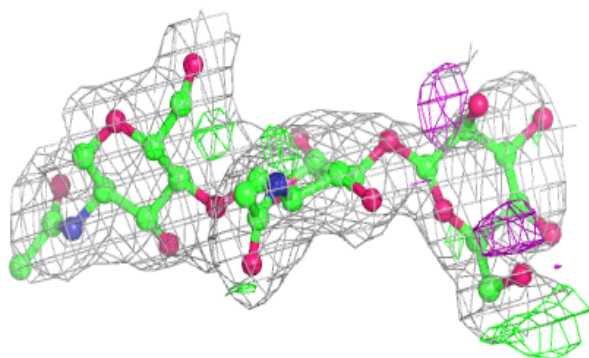
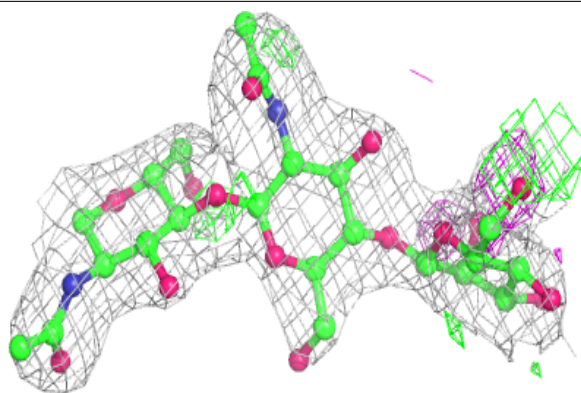
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



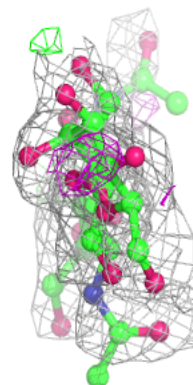
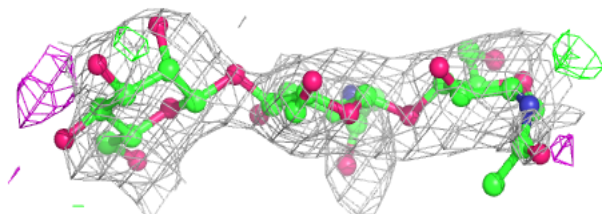
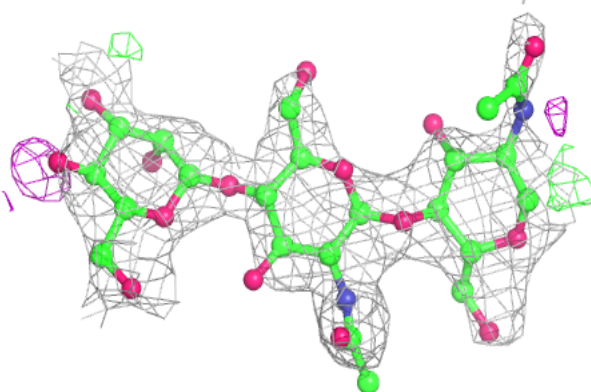


**Electron density around Chain H:**

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

**Electron density around Chain I:**

$2mF_o-DF_c$  (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
4	NAG	C	507	14/15	0.77	0.21	88,95,102,102	0
4	NAG	A	504	14/15	0.81	0.37	91,103,109,113	0
4	NAG	A	503	14/15	0.84	0.47	108,119,128,128	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.