



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 27, 2023 – 09:37 AM EDT

PDB ID : 3IGU  
Title : Crystal structure of human alpha-N-acetylgalactosaminidase, covalent intermediate  
Authors : Clark, N.E.; Garman, S.C.  
Deposited on : 2009-07-28  
Resolution : 2.15 Å(reported)

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

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

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

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

---

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

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.35  
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) : 2.35

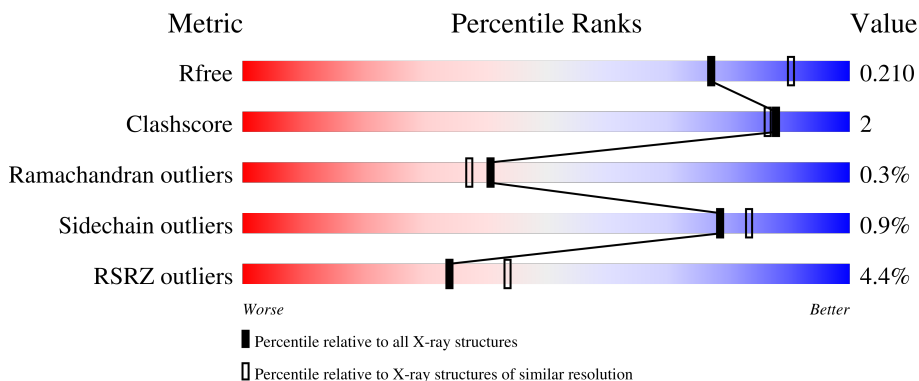
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.15 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	1479 (2.16-2.16)
Clashscore	141614	1585 (2.16-2.16)
Ramachandran outliers	138981	1560 (2.16-2.16)
Sidechain outliers	138945	1559 (2.16-2.16)
RSRZ outliers	127900	1456 (2.16-2.16)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	400	 92% 5%
1	B	400	 7% 90% 7%
2	C	2	 100%
3	D	5	 60% 40%
3	F	5	 40% 60%

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
4	E	3	 33% 67%

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
2	NAG	C	2	-	-	-	X
3	MAN	F	5	-	-	-	X
4	NAG	E	2	-	-	-	X
4	FUC	E	3	-	-	-	X
6	7JZ	B	1001	X	-	-	-
7	NAG	A	885	-	-	-	X
7	NAG	B	885	-	-	-	X

## 2 Entry composition i

There are 8 unique types of molecules in this entry. The entry contains 7197 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 Alpha-N-acetylgalactosaminidase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	387	3103	1981	526	569	27	0	2	0
1	B	387	3089	1973	520	569	27	0	1	0

There are 14 discrepancies between the modelled and reference sequences:

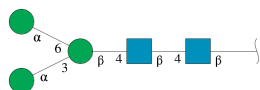
Chain	Residue	Modelled	Actual	Comment	Reference
A	201	GLN	ASN	engineered mutation	UNP P17050
A	412	HIS	-	expression tag	UNP P17050
A	413	HIS	-	expression tag	UNP P17050
A	414	HIS	-	expression tag	UNP P17050
A	415	HIS	-	expression tag	UNP P17050
A	416	HIS	-	expression tag	UNP P17050
A	417	HIS	-	expression tag	UNP P17050
B	201	GLN	ASN	engineered mutation	UNP P17050
B	412	HIS	-	expression tag	UNP P17050
B	413	HIS	-	expression tag	UNP P17050
B	414	HIS	-	expression tag	UNP P17050
B	415	HIS	-	expression tag	UNP P17050
B	416	HIS	-	expression tag	UNP P17050
B	417	HIS	-	expression tag	UNP P17050

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



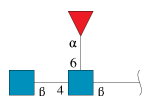
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	C	2	28	16	2	10	0	0	0

- Molecule 3 is an oligosaccharide called 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	D	5	61	34	2	25	0	0	0
3	F	5	61	34	2	25	0	0	0

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



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

- Molecule 5 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



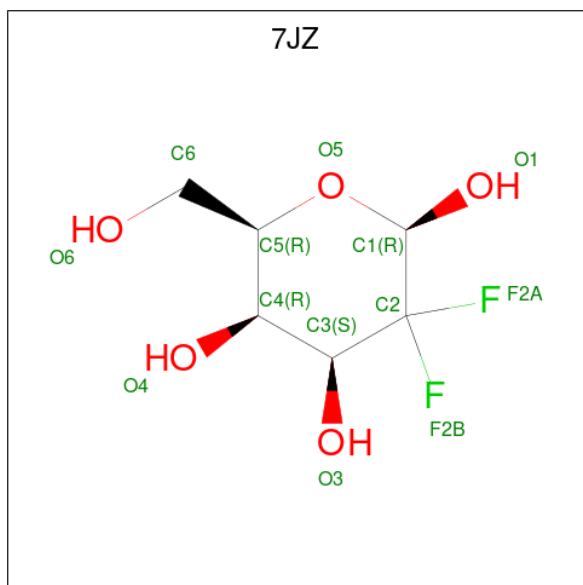
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total C O 6 3 3	0	0
5	A	1	Total C O 6 3 3	0	0
5	A	1	Total C O 6 3 3	0	0
5	A	1	Total C O 6 3 3	0	0
5	A	1	Total C O 6 3 3	0	0
5	A	1	Total C O 6 3 3	0	0
5	A	1	Total C O 6 3 3	0	0
5	A	1	Total C O 6 3 3	0	0
5	A	1	Total C O 6 3 3	0	0
5	A	1	Total C O 6 3 3	0	0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

*Continued on next page...*

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	B	1	Total	C	O	0	0
			6	3	3		

- Molecule 6 is 2-deoxy-2,2-difluoro-beta-D-lyxo-hexopyranose (three-letter code: 7JZ) (formula:  $C_6H_{10}F_2O_5$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	A	1	Total	C	F	O	0	0
			12	6	2	4		
6	B	1	Total	C	F	O	0	0
			12	6	2	4		

- Molecule 7 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



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

- Molecule 8 is water.

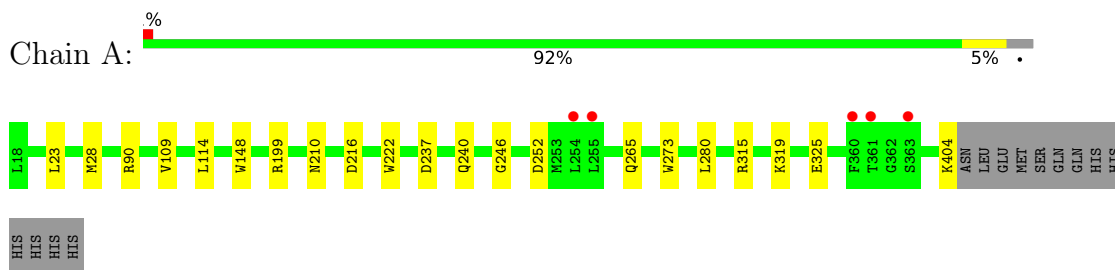
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
8	A	397	397	397	0	0
8	B	278	278	278	0	0



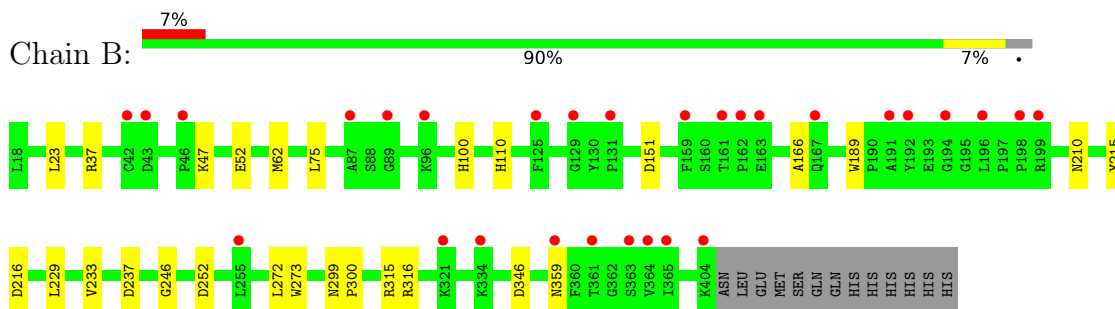
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

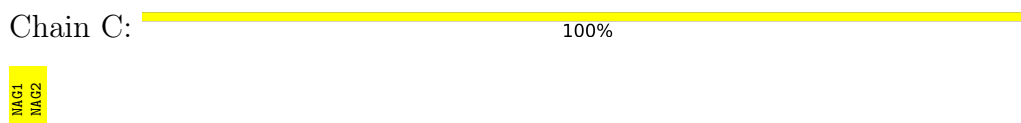
- Molecule 1: Alpha-N-acetylgalactosaminidase



- Molecule 1: Alpha-N-acetylgalactosaminidase



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



- Molecule 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



- Molecule 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

Chain F:  40% 60%



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

Chain E:  33% 67%



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	151.61Å 113.65Å 68.42Å 90.00° 96.14° 90.00°	Depositor
Resolution (Å)	45.98 – 2.15 45.96 – 2.15	Depositor EDS
% Data completeness (in resolution range)	99.6 (45.98-2.15) 99.6 (45.96-2.15)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	0.08	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.55 (at 2.14Å)	Xtrriage
Refinement program	REFMAC 5.5.0070	Depositor
R, $R_{free}$	0.164 , 0.199 0.174 , 0.210	Depositor DCC
$R_{free}$ test set	3150 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	33.7	Xtrriage
Anisotropy	0.069	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 46.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	7197	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	28.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.60% 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: NAG, MAN, FUC, 7JZ, BMA, GOL

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	0.39	0/3188	0.54	0/4335
1	B	0.36	0/3174	0.52	0/4318
All	All	0.37	0/6362	0.53	0/8653

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts [i](#)

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3103	0	3001	10	0
1	B	3089	0	2982	15	0
2	C	28	0	25	0	0
3	D	61	0	52	0	0
3	F	61	0	52	0	0
4	E	38	0	34	0	0
5	A	60	0	80	3	0
5	B	30	0	40	1	0
6	A	12	0	8	0	0
6	B	12	0	8	0	0
7	A	14	0	13	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	B	14	0	13	0	0
8	A	397	0	0	1	0
8	B	278	0	0	0	0
All	All	7197	0	6308	25	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 (25) 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:B:229:LEU:HD11	1:B:272:LEU:HD13	1.55	0.89
1:B:110:HIS:HE1	1:B:151:ASP:OD2	1.72	0.71
1:A:216:ASP:HA	5:A:1200:GOL:H31	1.77	0.65
1:A:90:ARG:NH2	1:A:148:TRP:HE1	2.01	0.57
1:B:52:GLU:OE2	1:B:100:HIS:HD2	1.89	0.55
5:A:1328:GOL:H12	1:B:37:ARG:HD2	1.91	0.53
1:A:237:ASP:OD1	1:A:315[A]:ARG:NH2	2.42	0.52
1:B:23:LEU:HD12	1:B:210:ASN:HB3	1.93	0.50
1:A:240:GLN:HG3	1:A:315[A]:ARG:HG2	1.94	0.49
1:B:346:ASP:HB3	5:B:1328:GOL:H31	1.95	0.49
1:A:319:LYS:HG3	1:A:325:GLU:HG2	1.96	0.48
1:A:109:VAL:HG13	1:A:114:LEU:HB2	1.97	0.47
1:A:222:TRP:CD2	1:A:265:GLN:HG2	2.52	0.45
1:A:199[A]:ARG:HD2	8:A:694:HOH:O	2.17	0.43
1:B:62:MET:SD	1:B:75:LEU:HD23	2.59	0.43
1:B:210:ASN:HA	1:B:246:GLY:O	2.18	0.43
1:B:215:TYR:CG	1:B:216:ASP:N	2.87	0.43
1:A:28:MET:HG2	1:A:280:LEU:HB2	2.00	0.43
1:B:110:HIS:CE1	1:B:151:ASP:OD2	2.62	0.42
1:B:166:ALA:HA	1:B:189:TRP:CZ2	2.54	0.42
1:B:233:VAL:HG11	1:B:316:ARG:CZ	2.49	0.42
1:A:210:ASN:HA	1:A:246:GLY:O	2.21	0.41
5:A:1328:GOL:H12	1:B:37:ARG:HH11	1.86	0.41
1:B:237:ASP:OD1	1:B:315:ARG:NH2	2.51	0.41
1:B:299:ASN:HA	1:B:300:PRO:HD3	1.95	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	387/400 (97%)	374 (97%)	12 (3%)	1 (0%)	41	37
1	B	386/400 (96%)	374 (97%)	11 (3%)	1 (0%)	41	37
All	All	773/800 (97%)	748 (97%)	23 (3%)	2 (0%)	41	37

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	252	ASP
1	B	252	ASP

### 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	335/346 (97%)	332 (99%)	3 (1%)	78	83
1	B	334/346 (96%)	331 (99%)	3 (1%)	78	83
All	All	669/692 (97%)	663 (99%)	6 (1%)	78	83

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

Mol	Chain	Res	Type
1	A	23	LEU
1	A	273	TRP
1	A	404	LYS
1	B	47	LYS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	273	TRP
1	B	359	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (7) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	143	GLN
1	A	235	HIS
1	A	318	HIS
1	B	100	HIS
1	B	110	HIS
1	B	235	HIS
1	B	357	GLN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates [i](#)

15 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	C	1	2,1	14,14,15	0.51	0	17,19,21	1.09	1 (5%)
2	NAG	C	2	2	14,14,15	0.56	0	17,19,21	1.06	2 (11%)
3	NAG	D	1	3,1	14,14,15	0.56	0	17,19,21	0.92	0
3	NAG	D	2	3	14,14,15	0.56	0	17,19,21	0.66	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	BMA	D	3	3	11,11,12	0.52	0	15,15,17	1.09	1 (6%)
3	MAN	D	4	3	11,11,12	0.68	0	15,15,17	1.06	1 (6%)
3	MAN	D	5	3	11,11,12	0.65	0	15,15,17	0.95	0
4	NAG	E	1	4,1	14,14,15	0.62	0	17,19,21	0.87	1 (5%)
4	NAG	E	2	4	14,14,15	0.49	0	17,19,21	0.82	1 (5%)
4	FUC	E	3	4	10,10,11	0.62	0	14,14,16	0.57	0
3	NAG	F	1	3,1	14,14,15	0.60	0	17,19,21	0.71	0
3	NAG	F	2	3	14,14,15	0.58	0	17,19,21	0.82	1 (5%)
3	BMA	F	3	3	11,11,12	0.46	0	15,15,17	0.70	0
3	MAN	F	4	3	11,11,12	0.62	0	15,15,17	0.91	1 (6%)
3	MAN	F	5	3	11,11,12	0.60	0	15,15,17	1.13	2 (13%)

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	C	1	2,1	-	0/6/23/26	0/1/1/1
2	NAG	C	2	2	-	2/6/23/26	0/1/1/1
3	NAG	D	1	3,1	-	0/6/23/26	0/1/1/1
3	NAG	D	2	3	-	0/6/23/26	0/1/1/1
3	BMA	D	3	3	-	2/2/19/22	0/1/1/1
3	MAN	D	4	3	-	2/2/19/22	0/1/1/1
3	MAN	D	5	3	-	0/2/19/22	0/1/1/1
4	NAG	E	1	4,1	-	4/6/23/26	0/1/1/1
4	NAG	E	2	4	-	2/6/23/26	0/1/1/1
4	FUC	E	3	4	-	-	0/1/1/1
3	NAG	F	1	3,1	-	0/6/23/26	0/1/1/1
3	NAG	F	2	3	-	0/6/23/26	0/1/1/1
3	BMA	F	3	3	-	2/2/19/22	0/1/1/1
3	MAN	F	4	3	-	2/2/19/22	0/1/1/1
3	MAN	F	5	3	-	2/2/19/22	0/1/1/1

There are no bond length outliers.

All (11) bond angle outliers are listed below:



Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	D	4	MAN	C1-C2-C3	3.08	113.46	109.67
3	F	5	MAN	C1-O5-C5	2.92	116.15	112.19
2	C	2	NAG	C4-C3-C2	2.61	114.84	111.02
4	E	1	NAG	C4-C3-C2	2.46	114.62	111.02
4	E	2	NAG	C1-O5-C5	2.40	115.44	112.19
3	D	3	BMA	O5-C5-C6	2.37	110.92	107.20
3	F	4	MAN	C1-C2-C3	2.34	112.54	109.67
2	C	2	NAG	O5-C5-C6	2.27	110.76	107.20
3	F	2	NAG	C4-C3-C2	2.19	114.23	111.02
3	F	5	MAN	C1-C2-C3	2.08	112.23	109.67
2	C	1	NAG	O5-C5-C6	2.00	110.35	107.20

There are no chirality outliers.

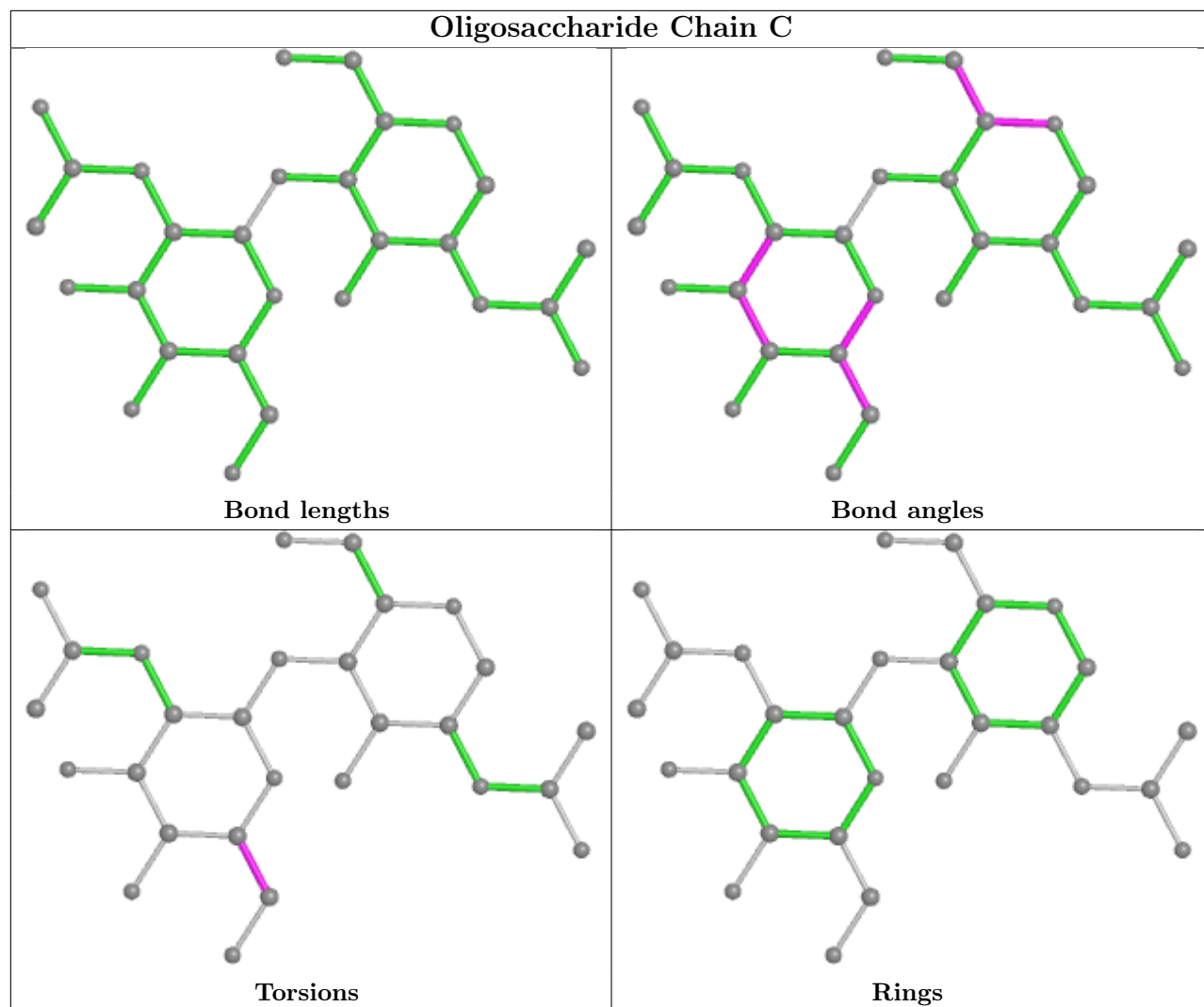
All (18) torsion outliers are listed below:

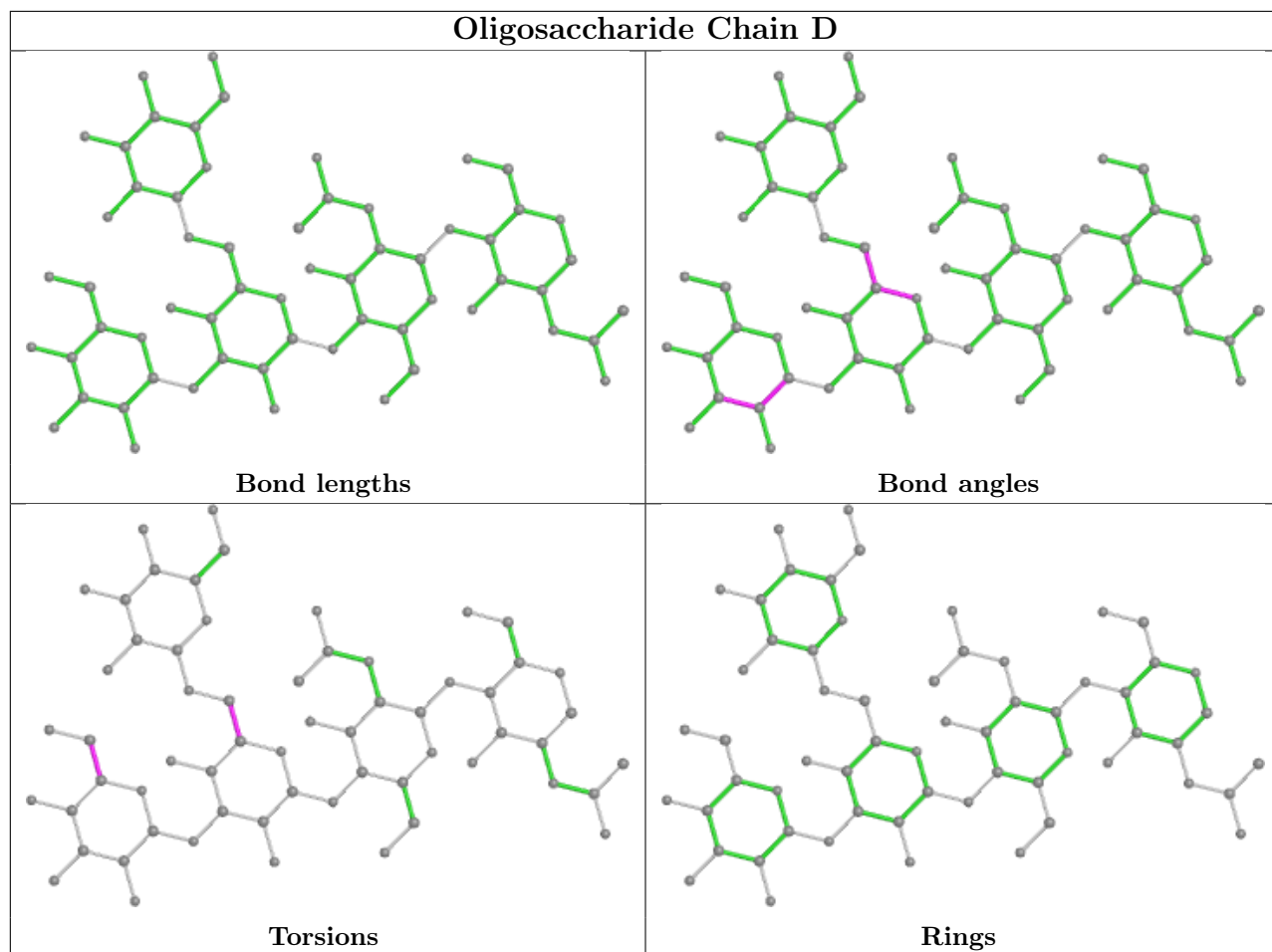
Mol	Chain	Res	Type	Atoms
4	E	1	NAG	C8-C7-N2-C2
4	E	1	NAG	O7-C7-N2-C2
4	E	2	NAG	C8-C7-N2-C2
4	E	2	NAG	O7-C7-N2-C2
3	D	3	BMA	O5-C5-C6-O6
3	D	4	MAN	O5-C5-C6-O6
3	F	4	MAN	O5-C5-C6-O6
3	F	5	MAN	O5-C5-C6-O6
3	D	4	MAN	C4-C5-C6-O6
3	D	3	BMA	C4-C5-C6-O6
3	F	4	MAN	C4-C5-C6-O6
3	F	5	MAN	C4-C5-C6-O6
3	F	3	BMA	O5-C5-C6-O6
3	F	3	BMA	C4-C5-C6-O6
2	C	2	NAG	C4-C5-C6-O6
2	C	2	NAG	O5-C5-C6-O6
4	E	1	NAG	C4-C5-C6-O6
4	E	1	NAG	O5-C5-C6-O6

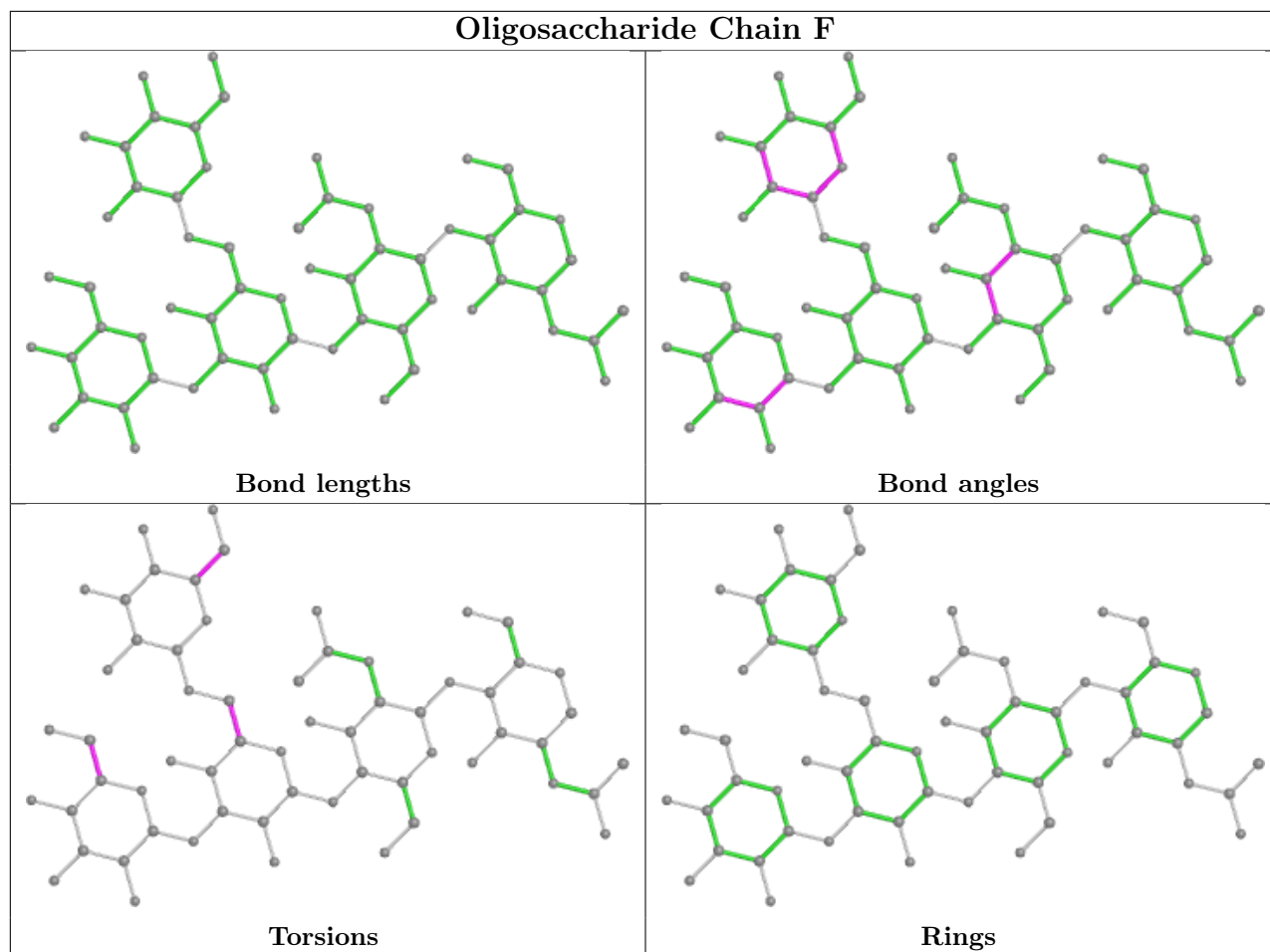
There are no ring outliers.

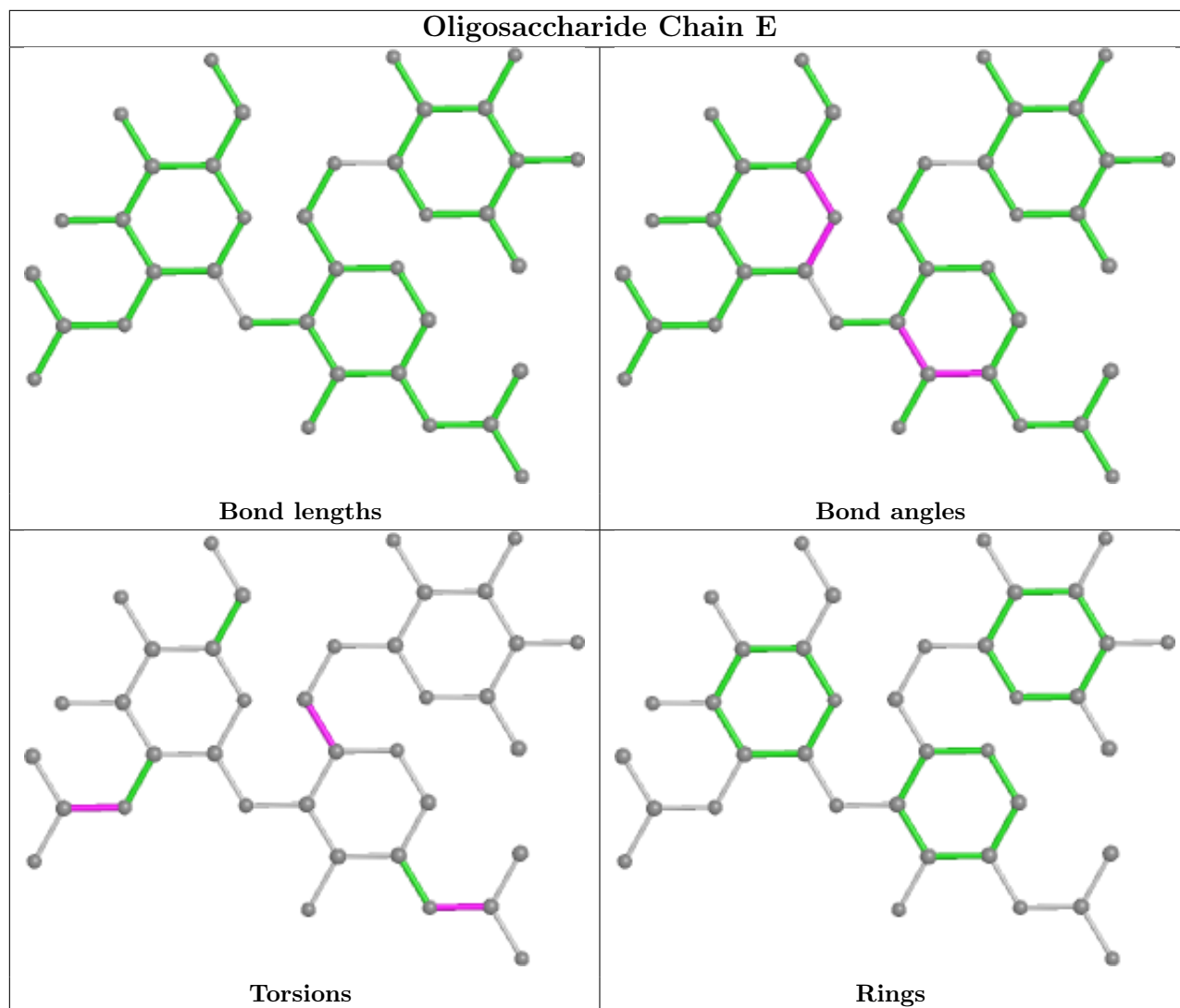
No monomer is involved in short contacts.

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









## 5.6 Ligand geometry [i](#)

19 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	$\# Z  > 2$
5	GOL	A	1328	-	5,5,5	0.38	0	5,5,5	0.20	0
5	GOL	B	1024	-	5,5,5	0.37	0	5,5,5	0.33	0
5	GOL	B	1200	-	5,5,5	0.38	0	5,5,5	0.24	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
5	GOL	B	1108	-	5,5,5	0.35	0	5,5,5	0.27	0
5	GOL	A	1108	-	5,5,5	0.37	0	5,5,5	0.25	0
5	GOL	A	1165	-	5,5,5	0.37	0	5,5,5	0.36	0
5	GOL	A	1210	-	5,5,5	0.34	0	5,5,5	0.32	0
6	7JZ	A	1001	1	11,12,13	1.15	1 (9%)	11,18,20	1.97	3 (27%)
5	GOL	B	1113	-	5,5,5	0.39	0	5,5,5	0.33	0
5	GOL	A	1201	-	5,5,5	0.36	0	5,5,5	0.34	0
5	GOL	A	1113	-	5,5,5	0.38	0	5,5,5	0.30	0
5	GOL	A	1200	-	5,5,5	0.39	0	5,5,5	0.36	0
5	GOL	A	1030	-	5,5,5	0.38	0	5,5,5	0.29	0
5	GOL	B	1328	-	5,5,5	0.36	0	5,5,5	0.34	0
5	GOL	A	1024	-	5,5,5	0.39	0	5,5,5	0.41	0
6	7JZ	B	1001	1	11,12,13	0.93	0	11,18,20	1.39	2 (18%)
5	GOL	A	1039	-	5,5,5	0.37	0	5,5,5	0.30	0
7	NAG	A	885	1	14,14,15	0.47	0	17,19,21	0.89	1 (5%)
7	NAG	B	885	1	14,14,15	0.55	0	17,19,21	0.93	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
5	GOL	A	1328	-	-	0/4/4/4	-
5	GOL	B	1024	-	-	4/4/4/4	-
5	GOL	B	1200	-	-	2/4/4/4	-
5	GOL	B	1108	-	-	3/4/4/4	-
5	GOL	A	1108	-	-	2/4/4/4	-
5	GOL	A	1165	-	-	2/4/4/4	-
5	GOL	A	1210	-	-	2/4/4/4	-
6	7JZ	A	1001	1	-	0/2/22/26	0/1/1/1
5	GOL	B	1113	-	-	3/4/4/4	-
5	GOL	A	1201	-	-	0/4/4/4	-
5	GOL	A	1113	-	-	2/4/4/4	-
5	GOL	A	1200	-	-	4/4/4/4	-
5	GOL	A	1030	-	-	2/4/4/4	-
5	GOL	B	1328	-	-	0/4/4/4	-
5	GOL	A	1024	-	-	4/4/4/4	-
6	7JZ	B	1001	1	1/1/4/5	0/2/22/26	0/1/1/1

Continued on next page...

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	GOL	A	1039	-	-	4/4/4/4	-
7	NAG	A	885	1	-	4/6/23/26	0/1/1/1
7	NAG	B	885	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)
6	A	1001	7JZ	C1-C2	2.31	1.56	1.52

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	A	1001	7JZ	F2B-C2-C3	-3.46	105.32	109.15
6	A	1001	7JZ	O5-C5-C6	3.15	112.14	107.20
6	A	1001	7JZ	F2B-C2-C1	3.13	114.11	109.67
6	B	1001	7JZ	F2B-C2-C3	-2.73	106.13	109.15
7	A	885	NAG	C1-O5-C5	2.70	115.85	112.19
6	B	1001	7JZ	O5-C5-C6	2.44	111.02	107.20
7	B	885	NAG	C4-C3-C2	2.25	114.32	111.02

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
6	B	1001	7JZ	C1

All (40) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	1024	GOL	C1-C2-C3-O3
5	A	1039	GOL	O1-C1-C2-C3
5	A	1165	GOL	O1-C1-C2-O2
5	A	1165	GOL	O1-C1-C2-C3
5	A	1200	GOL	O1-C1-C2-O2
5	A	1200	GOL	O1-C1-C2-C3
5	A	1030	GOL	C1-C2-C3-O3
5	A	1210	GOL	O1-C1-C2-O2
5	A	1210	GOL	O1-C1-C2-C3
5	B	1024	GOL	C1-C2-C3-O3
5	B	1113	GOL	O1-C1-C2-C3
5	B	1200	GOL	O1-C1-C2-O2
7	A	885	NAG	C8-C7-N2-C2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms
7	A	885	NAG	O7-C7-N2-C2
5	A	1024	GOL	O1-C1-C2-O2
5	A	1200	GOL	O2-C2-C3-O3
5	B	1108	GOL	O1-C1-C2-O2
7	A	885	NAG	O5-C5-C6-O6
5	A	1024	GOL	O1-C1-C2-C3
5	A	1039	GOL	C1-C2-C3-O3
5	A	1108	GOL	C1-C2-C3-O3
5	A	1113	GOL	C1-C2-C3-O3
5	A	1200	GOL	C1-C2-C3-O3
5	B	1024	GOL	O1-C1-C2-C3
5	B	1108	GOL	O1-C1-C2-C3
5	B	1108	GOL	C1-C2-C3-O3
5	B	1113	GOL	C1-C2-C3-O3
5	B	1200	GOL	O1-C1-C2-C3
7	B	885	NAG	C4-C5-C6-O6
5	A	1024	GOL	O2-C2-C3-O3
5	A	1039	GOL	O1-C1-C2-O2
5	B	1024	GOL	O1-C1-C2-O2
5	B	1024	GOL	O2-C2-C3-O3
7	B	885	NAG	O5-C5-C6-O6
5	A	1113	GOL	O2-C2-C3-O3
5	B	1113	GOL	O1-C1-C2-O2
5	A	1108	GOL	O2-C2-C3-O3
5	A	1030	GOL	O2-C2-C3-O3
5	A	1039	GOL	O2-C2-C3-O3
7	A	885	NAG	C4-C5-C6-O6

There are no ring outliers.

3 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	1328	GOL	2	0
5	A	1200	GOL	1	0
5	B	1328	GOL	1	0

## 5.7 Other polymers [i](#)

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	387/400 (96%)	-0.05	5 (1%) 77 82	14, 22, 37, 56	0
1	B	387/400 (96%)	0.36	29 (7%) 14 19	15, 25, 40, 60	0
All	All	774/800 (96%)	0.16	34 (4%) 34 43	14, 23, 40, 60	0

All (34) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	363	SER	5.9
1	B	404	LYS	4.9
1	B	194	GLY	4.3
1	B	198	PRO	4.3
1	A	363	SER	4.2
1	B	365	ILE	4.1
1	B	87	ALA	4.0
1	B	361	THR	3.7
1	B	125	PHE	3.5
1	A	361	THR	3.4
1	B	196	LEU	3.3
1	B	364	VAL	3.2
1	B	96	LYS	3.0
1	B	159	PHE	2.8
1	B	162	PRO	2.8
1	B	163	GLU	2.6
1	B	334	LYS	2.6
1	B	199	ARG	2.6
1	A	255	LEU	2.5
1	B	321	LYS	2.5
1	B	43	ASP	2.5
1	B	161	THR	2.4
1	B	167	GLN	2.4
1	B	131	PRO	2.4

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	129	GLY	2.2
1	B	89	GLY	2.2
1	A	360	PHE	2.2
1	A	254	LEU	2.1
1	B	192	TYR	2.1
1	B	46	PRO	2.1
1	B	359	ASN	2.1
1	B	255	LEU	2.1
1	B	42	CYS	2.1
1	B	191	ALA	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](#)

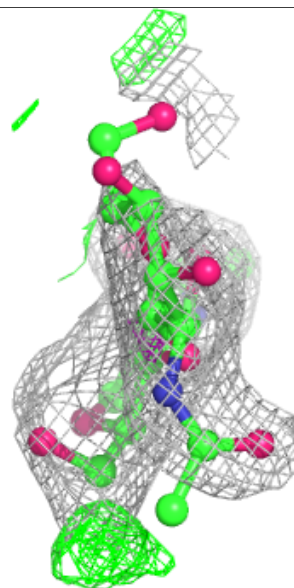
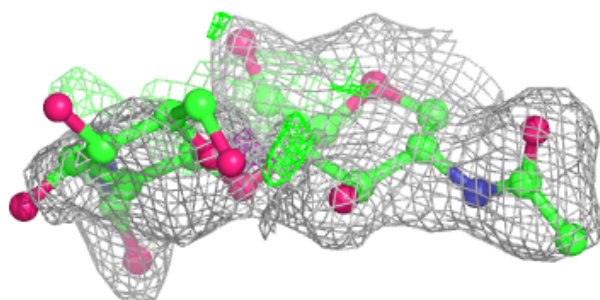
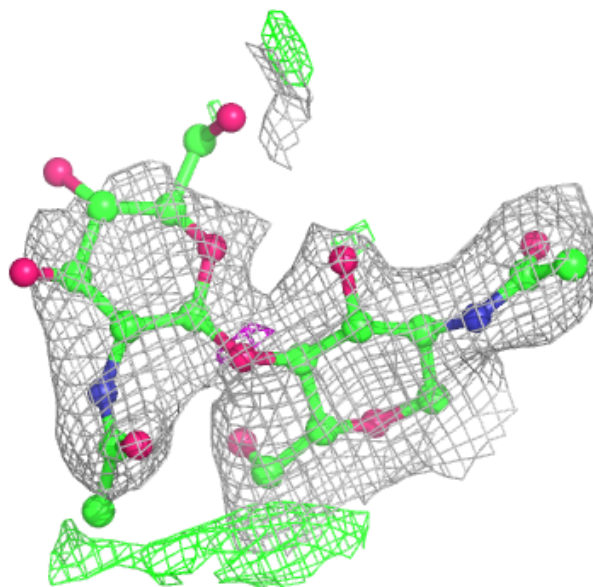
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q<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
4	FUC	E	3	10/11	0.50	0.57	111,112,114,114	0
4	NAG	E	2	14/15	0.56	0.52	104,110,112,112	0
3	MAN	F	5	11/12	0.67	0.65	111,112,113,113	0
3	MAN	D	5	11/12	0.76	0.37	80,82,84,86	0
4	NAG	E	1	14/15	0.76	0.30	83,92,104,107	0
3	BMA	F	3	11/12	0.77	0.23	88,96,102,107	0
2	NAG	C	2	14/15	0.80	0.42	91,94,98,98	0
3	MAN	D	4	11/12	0.80	0.22	76,82,85,87	0
2	NAG	C	1	14/15	0.82	0.23	59,67,76,84	0
3	MAN	F	4	11/12	0.82	0.22	102,104,106,106	0
3	BMA	D	3	11/12	0.87	0.13	46,59,67,76	0
3	NAG	F	2	14/15	0.91	0.18	65,71,73,80	0
3	NAG	D	2	14/15	0.92	0.15	42,45,48,51	0
3	NAG	F	1	14/15	0.93	0.15	46,54,61,61	0
3	NAG	D	1	14/15	0.96	0.11	30,37,42,47	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.

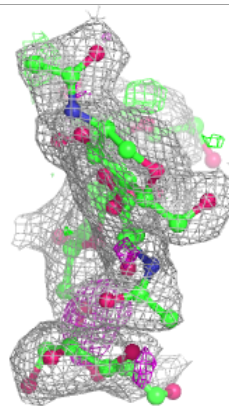
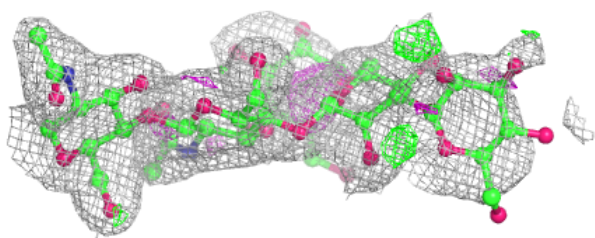
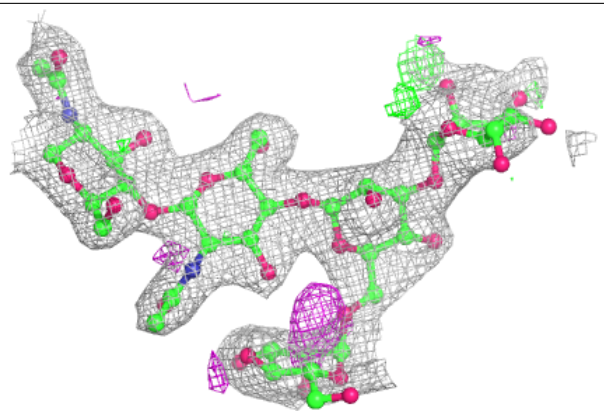
**Electron density around Chain C:**

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

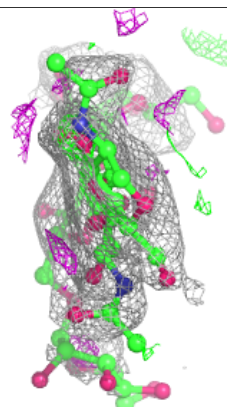
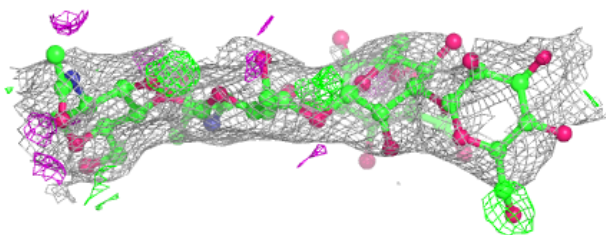
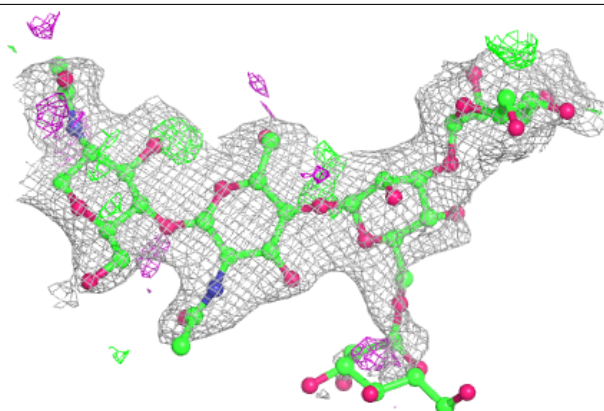


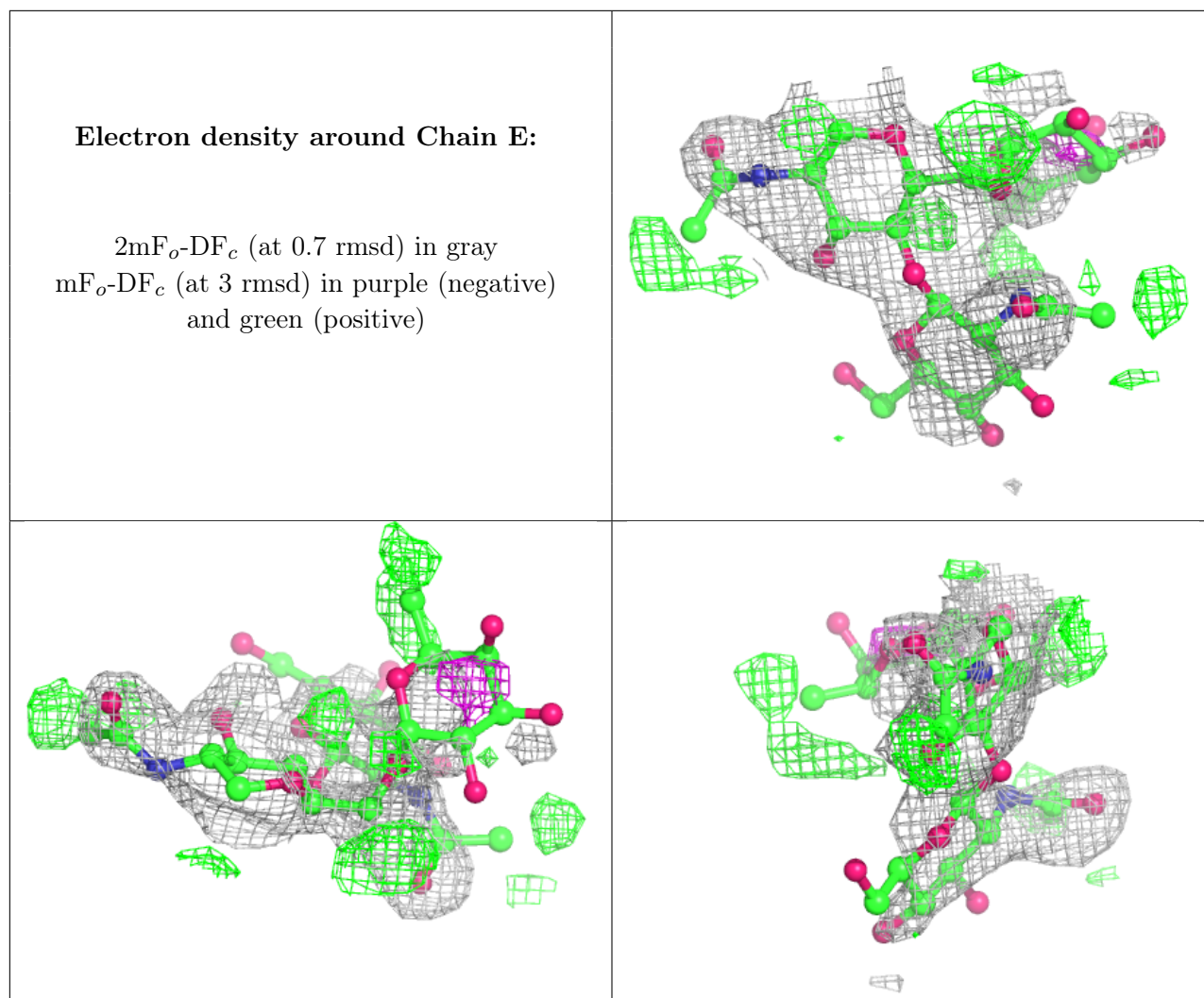
**Electron density around Chain D:**

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

**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)





## 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
5	GOL	B	1200	6/6	0.51	0.28	79,81,84,84	0
7	NAG	B	885	14/15	0.65	0.47	78,84,89,90	0
7	NAG	A	885	14/15	0.67	0.50	81,88,93,94	0
5	GOL	A	1108	6/6	0.68	0.29	81,81,83,85	0
5	GOL	B	1024	6/6	0.75	0.21	68,69,70,71	0
6	7JZ	B	1001	12/13	0.76	0.19	39,46,53,57	0
5	GOL	B	1108	6/6	0.78	0.21	89,89,90,90	0
5	GOL	B	1113	6/6	0.80	0.31	82,82,84,85	0
6	7JZ	A	1001	12/13	0.84	0.17	22,28,40,56	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
5	GOL	A	1200	6/6	0.85	0.23	49,58,60,63	0
5	GOL	A	1210	6/6	0.85	0.25	80,81,83,85	0
5	GOL	A	1039	6/6	0.86	0.23	79,79,81,81	0
5	GOL	A	1201	6/6	0.86	0.18	45,59,62,64	0
5	GOL	A	1030	6/6	0.86	0.18	77,80,80,81	0
5	GOL	B	1328	6/6	0.87	0.18	49,54,58,65	0
5	GOL	A	1328	6/6	0.88	0.16	75,76,76,77	0
5	GOL	A	1113	6/6	0.89	0.23	67,71,73,75	0
5	GOL	A	1024	6/6	0.90	0.12	55,57,59,60	0
5	GOL	A	1165	6/6	0.91	0.16	40,52,54,56	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.