



# wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 25, 2024 – 07:14 PM EST

PDB ID : 5N2J  
Title : UDP-Glucose Glycoprotein Glucosyltransferase from Chaetomium thermophilum (closed form)  
Authors : Roversi, P.; Caputo, A.T.; Hill, J.; Alonzi, D.S.; Zitzmann, N.  
Deposited on : 2017-02-07  
Resolution : 4.40 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 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:

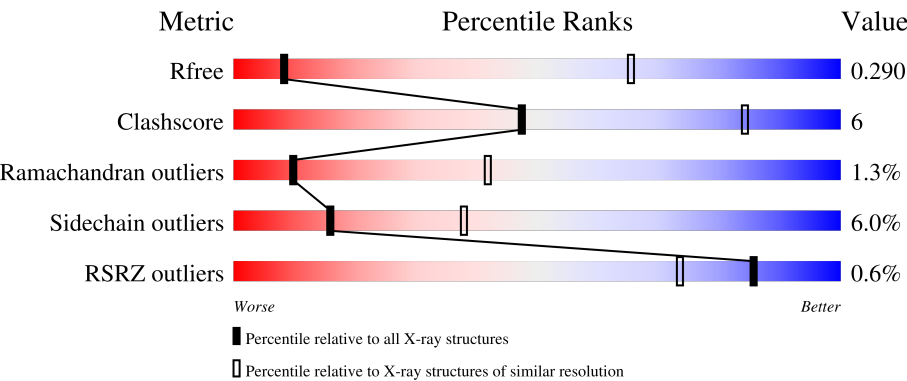
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.21
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.004 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.40

# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:  
*X-RAY DIFFRACTION*

The reported resolution of this entry is 4.40 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.



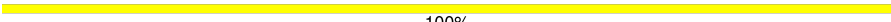
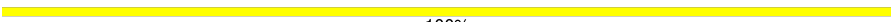
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R <sub>free</sub>	164625	1013 (4.90-3.90)
Clashscore	180529	1066 (4.90-3.90)
Ramachandran outliers	177936	1019 (4.92-3.86)
Sidechain outliers	177891	1003 (4.92-3.86)
RSRZ outliers	164620	1010 (4.90-3.90)

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  $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	1494	<div><div>75%</div><div>16%</div><div>• 8%</div></div>
1	B	1494	<div><div>%</div><div>74%</div><div>17%</div><div>• 8%</div></div>
2	C	3	<div><div>100%</div></div>
2	F	3	<div><div>33%</div><div>67%</div></div>
3	D	5	<div><div>20%</div><div>80%</div></div>
3	G	5	<div><div>20%</div><div>80%</div></div>

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
4	E	2	 100%
4	H	2	 100%

## 2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 44580 atoms, of which 22133 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 UDP-glucose-glycoprotein glucosyltransferase-like protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	1373	Total	C	H	N	O	S	0	0	0
			21929	7057	10895	1880	2065	32			
1	B	1381	Total	C	H	N	O	S	0	0	0
			22060	7098	10964	1889	2077	32			

There are 24 discrepancies between the modelled and reference sequences:

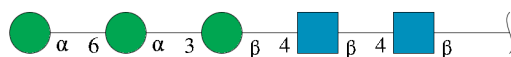
Chain	Residue	Modelled	Actual	Comment	Reference
A	21	GLU	-	expression tag	UNP G0SB58
A	22	THR	-	expression tag	UNP G0SB58
A	23	GLY	-	expression tag	UNP G0SB58
A	1506	GLY	-	expression tag	UNP G0SB58
A	1507	THR	-	expression tag	UNP G0SB58
A	1508	LYS	-	expression tag	UNP G0SB58
A	1509	HIS	-	expression tag	UNP G0SB58
A	1510	HIS	-	expression tag	UNP G0SB58
A	1511	HIS	-	expression tag	UNP G0SB58
A	1512	HIS	-	expression tag	UNP G0SB58
A	1513	HIS	-	expression tag	UNP G0SB58
A	1514	HIS	-	expression tag	UNP G0SB58
B	21	GLU	-	expression tag	UNP G0SB58
B	22	THR	-	expression tag	UNP G0SB58
B	23	GLY	-	expression tag	UNP G0SB58
B	1506	GLY	-	expression tag	UNP G0SB58
B	1507	THR	-	expression tag	UNP G0SB58
B	1508	LYS	-	expression tag	UNP G0SB58
B	1509	HIS	-	expression tag	UNP G0SB58
B	1510	HIS	-	expression tag	UNP G0SB58
B	1511	HIS	-	expression tag	UNP G0SB58
B	1512	HIS	-	expression tag	UNP G0SB58
B	1513	HIS	-	expression tag	UNP G0SB58
B	1514	HIS	-	expression tag	UNP G0SB58

- Molecule 2 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
2	C	3	Total	C	H	N	O	0	0	0
			73	22	34	2	15			
2	F	3	Total	C	H	N	O	0	0	0
			73	22	34	2	15			

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



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	5	Total	C	H	N	O	0	0	0
			113	34	52	2	25			
3	G	5	Total	C	H	N	O	0	0	0
			113	34	52	2	25			

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



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	E	2	Total	C	H	N	O	0	0	0
			53	16	25	2	10			
4	H	2	Total	C	H	N	O	0	0	0
			53	16	25	2	10			

- Molecule 5 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: C<sub>8</sub>H<sub>15</sub>NO<sub>6</sub>).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	A	1	Total	C	H	N	O	0	0
			27	8	13	1	5		
5	A	1	Total	C	H	N	O	0	0
			27	8	13	1	5		
5	B	1	Total	C	H	N	O	0	0
			27	8	13	1	5		
5	B	1	Total	C	H	N	O	0	0
			27	8	13	1	5		

- Molecule 6 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	1	Total	Ca	0	0
			1	1		
6	B	1	Total	Ca	0	0
			1	1		

- Molecule 7 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	1	Total O 1 1	0	0

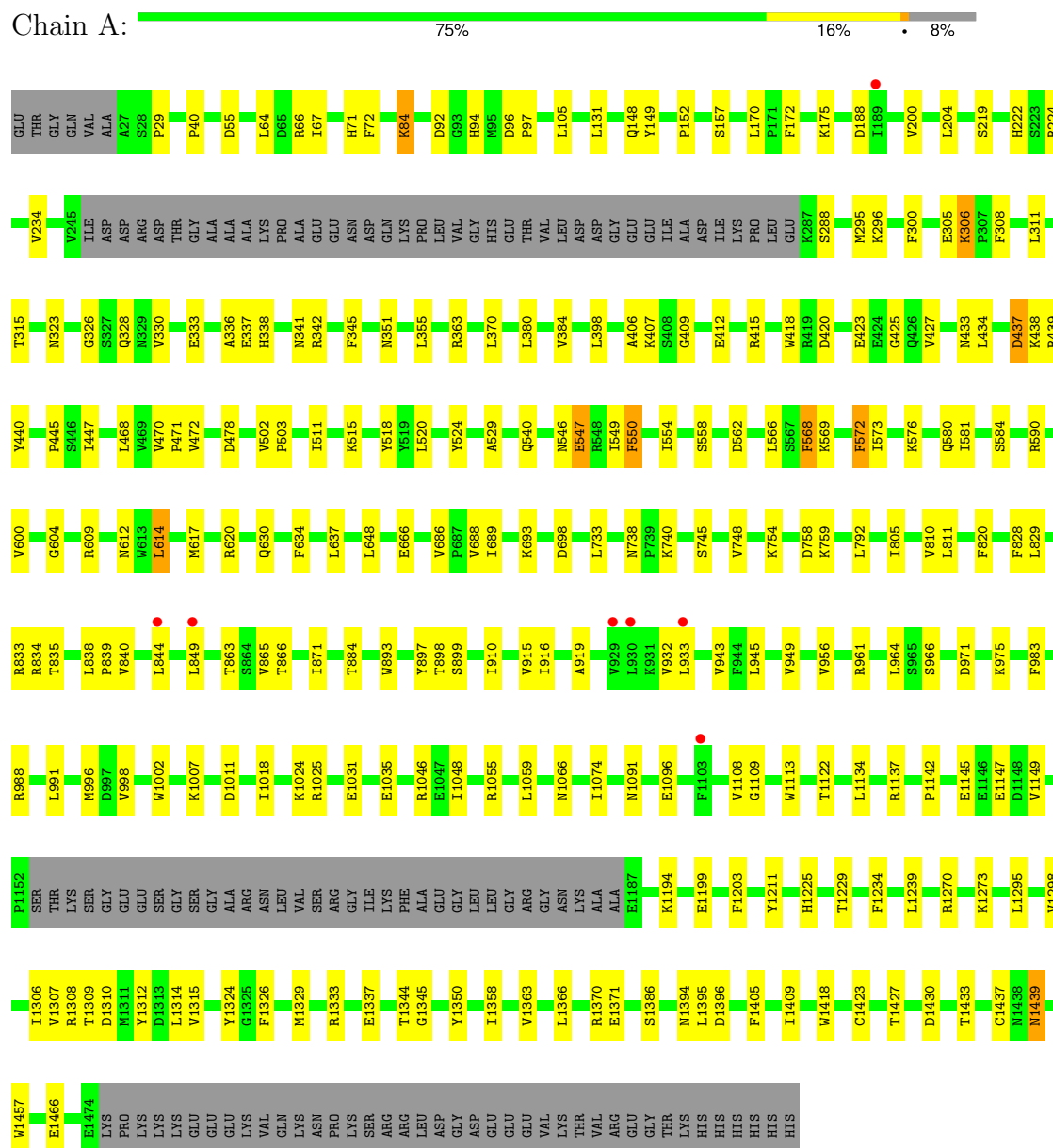
- Molecule 8 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	A	1	Total O 1 1	0	0
8	B	1	Total O 1 1	0	0

### 3 Residue-property plots

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. 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. 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: UDP-glucose-glycoprotein glucosyltransferase-like protein



- Molecule 1: UDP-glucose-glycoprotein glucosyltransferase-like protein





Chain F:  33% 67%

MAG1  
MAG2  
BGL3

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

Chain D:  20% 80%

MAG1  
MAG2  
BGL3  
MAN4  
MAN5

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

Chain G:  20% 80%

MAG1  
MAG2  
BGL3  
MAN4  
MAN5

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

Chain E:  100%

MAG1  
MAG2

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

Chain H:  100%

MAG1  
MAG2

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 43	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	148.60Å 148.60Å 179.97Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	114.59 – 4.40 114.59 – 4.40	Depositor EDS
% Data completeness (in resolution range)	100.0 (114.59-4.40) 100.0 (114.59-4.40)	Depositor EDS
$R_{merge}$	0.78	Depositor
$R_{sym}$	0.81	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.70 (at 4.30Å)	Xtriage
Refinement program	BUSTER 2.10.3	Depositor
R, $R_{free}$	0.201 , 0.259 0.230 , 0.290	Depositor DCC
$R_{free}$ test set	1282 reflections (5.15%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	126.7	Xtriage
Anisotropy	0.358	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 270.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.34$ , $\langle L^2 \rangle = 0.17$	Xtriage
Estimated twinning fraction	0.167 for h,-k,-l	Xtriage
$F_o, F_c$ correlation	0.87	EDS
Total number of atoms	44580	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	243.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.34% 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, CA, MAN, NAG, EDO

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	0.49	0/11298	0.74	2/15323 (0.0%)
1	B	0.55	0/11361	0.77	3/15409 (0.0%)
All	All	0.52	0/22659	0.75	5/30732 (0.0%)

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	305	GLU	C-N-CA	5.80	136.21	121.70
1	B	305	GLU	C-N-CA	5.80	136.21	121.70
1	B	565	GLU	C-N-CA	5.52	135.50	121.70
1	B	306	LYS	N-CA-C	-5.42	96.38	111.00
1	A	306	LYS	N-CA-C	-5.29	96.70	111.00

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	11034	10895	10898	144	0
1	B	11096	10964	10966	105	0
2	C	39	34	34	0	0
2	F	39	34	34	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	D	61	52	52	0	0
3	G	61	52	52	0	0
4	E	28	25	25	0	0
4	H	28	25	25	0	0
5	A	28	26	26	0	0
5	B	28	26	26	0	0
6	A	1	0	0	0	0
6	B	1	0	0	0	0
7	A	1	0	0	0	0
8	A	1	0	0	1	0
8	B	1	0	0	0	0
All	All	22447	22133	22138	247	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.

The worst 5 of 247 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:1225:HIS:ND1	1:A:1308:ARG:HA	1.68	1.07
1:B:243:TYR:O	1:B:285:LEU:HG	1.65	0.96
1:A:1225:HIS:CE1	1:A:1308:ARG:HA	2.04	0.93
1:A:1149:VAL:HG13	1:A:1371:GLU:O	1.71	0.90
1:B:1149:VAL:HG13	1:B:1371:GLU:O	1.73	0.87

There are no symmetry-related clashes.

## 5.3 Torsion angles ⓘ

### 5.3.1 Protein backbone ⓘ

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	1367/1494 (92%)	1297 (95%)	59 (4%)	11 (1%)	16 53

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	1375/1494 (92%)	1265 (92%)	84 (6%)	26 (2%)	6	33
All	All	2742/2988 (92%)	2562 (93%)	143 (5%)	37 (1%)	10	42

5 of 37 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	288	SER
1	A	569	LYS
1	A	1439	ASN
1	B	288	SER
1	B	332	ALA

### 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	1199/1297 (92%)	1150 (96%)	49 (4%)	26	48
1	B	1206/1297 (93%)	1110 (92%)	96 (8%)	10	29
All	All	2405/2594 (93%)	2260 (94%)	145 (6%)	16	38

5 of 145 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	834	ARG
1	B	1439	ASN
1	B	858	SER
1	B	1066	ASN
1	A	1466	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 12 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	400	HIS
1	B	523	ASN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	1058	GLN
1	B	738	ASN
1	A	1058	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 ⓘ

20 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	1,2	14,14,15	0.34	0	17,19,21	0.93	1 (5%)
2	NAG	C	2	2	14,14,15	0.50	0	17,19,21	1.63	2 (11%)
2	BMA	C	3	2	11,11,12	0.49	0	15,15,17	0.79	1 (6%)
3	NAG	D	1	3,1	14,14,15	0.30	0	17,19,21	2.51	2 (11%)
3	NAG	D	2	3	14,14,15	0.38	0	17,19,21	1.74	2 (11%)
3	BMA	D	3	3	11,11,12	0.32	0	15,15,17	0.51	0
3	MAN	D	4	3	11,11,12	0.46	0	15,15,17	0.86	1 (6%)
3	MAN	D	5	3	11,11,12	0.41	0	15,15,17	1.12	1 (6%)
4	NAG	E	1	4,1	14,14,15	0.39	0	17,19,21	0.89	1 (5%)
4	NAG	E	2	4	14,14,15	0.46	0	17,19,21	1.66	2 (11%)
2	NAG	F	1	1,2	14,14,15	0.33	0	17,19,21	0.86	1 (5%)
2	NAG	F	2	2	14,14,15	0.53	0	17,19,21	1.51	2 (11%)
2	BMA	F	3	2	11,11,12	0.52	0	15,15,17	0.83	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	NAG	G	1	3,1	14,14,15	0.25	0	17,19,21	2.52	2 (11%)
3	NAG	G	2	3	14,14,15	0.37	0	17,19,21	1.87	2 (11%)
3	BMA	G	3	3	11,11,12	0.36	0	15,15,17	0.53	0
3	MAN	G	4	3	11,11,12	0.40	0	15,15,17	0.87	1 (6%)
3	MAN	G	5	3	11,11,12	0.37	0	15,15,17	1.10	1 (6%)
4	NAG	H	1	4,1	14,14,15	0.38	0	17,19,21	0.95	1 (5%)
4	NAG	H	2	4	14,14,15	0.47	0	17,19,21	1.72	2 (11%)

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

There are no bond length outliers.

The worst 5 of 25 bond angle outliers are listed below:



Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
3	G	1	NAG	C1-O5-C5	8.49	123.57	112.19
3	D	1	NAG	C1-O5-C5	8.48	123.56	112.19
3	G	2	NAG	O5-C1-C2	6.26	120.98	111.29
3	D	2	NAG	O5-C1-C2	5.92	120.46	111.29
4	H	2	NAG	C1-O5-C5	5.25	119.22	112.19

There are no chirality outliers.

5 of 26 torsion outliers are listed below:

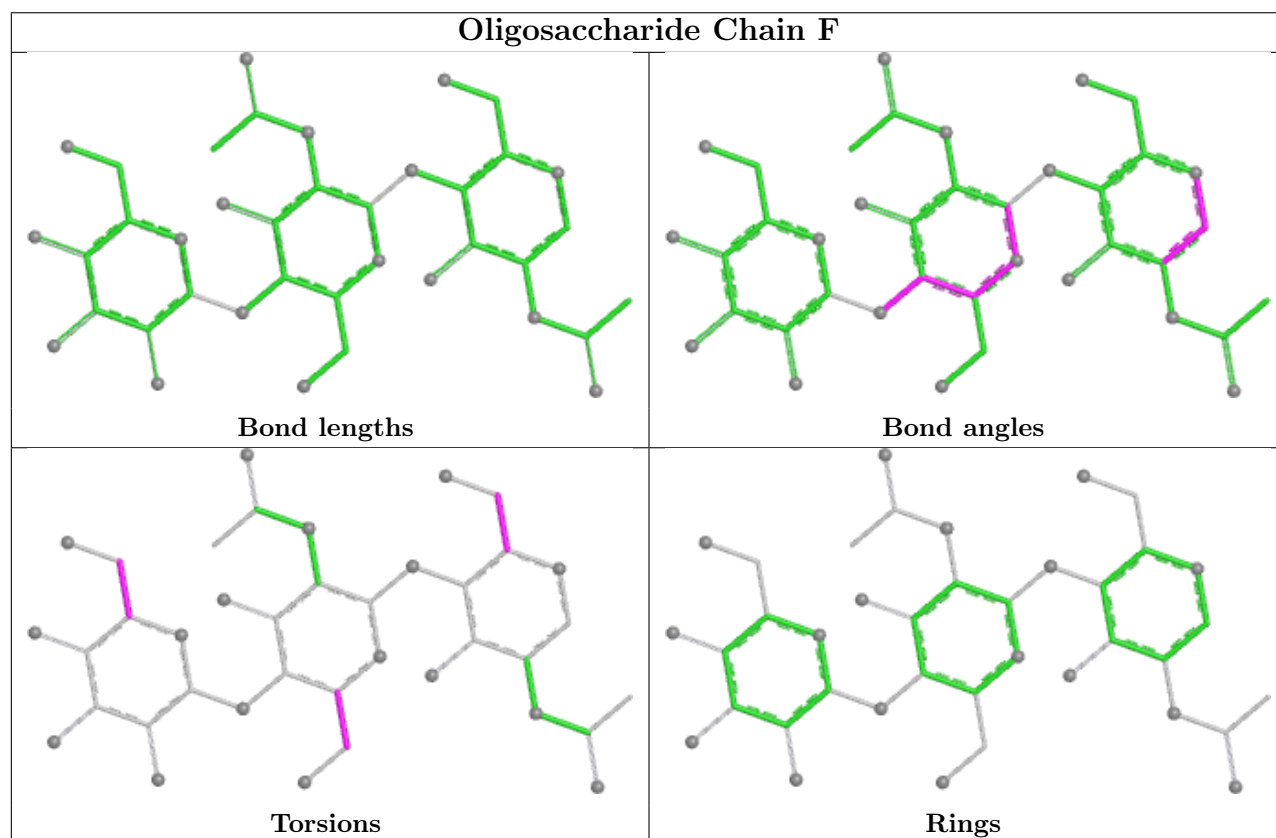
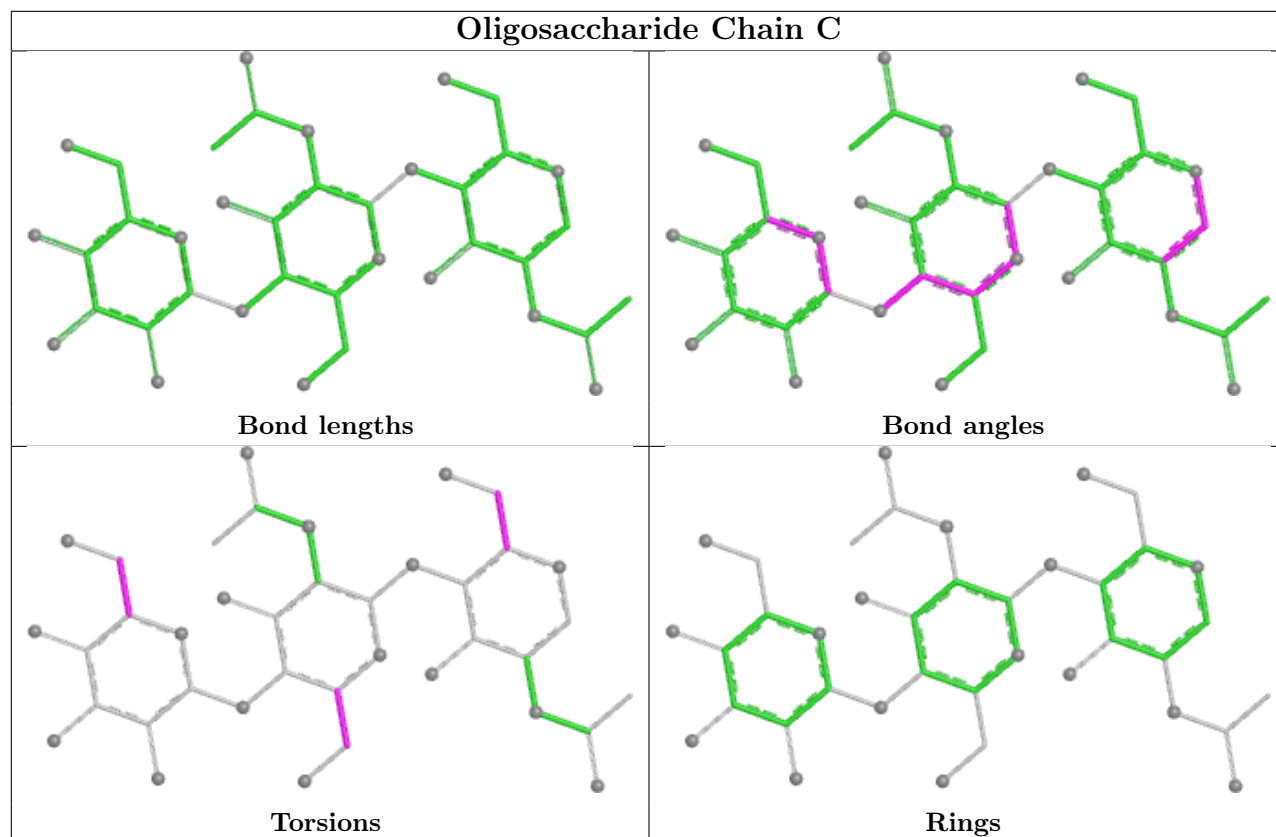
Mol	Chain	Res	Type	Atoms
3	G	4	MAN	O5-C5-C6-O6
3	D	4	MAN	O5-C5-C6-O6
2	F	2	NAG	O5-C5-C6-O6
4	H	1	NAG	C4-C5-C6-O6
2	C	2	NAG	O5-C5-C6-O6

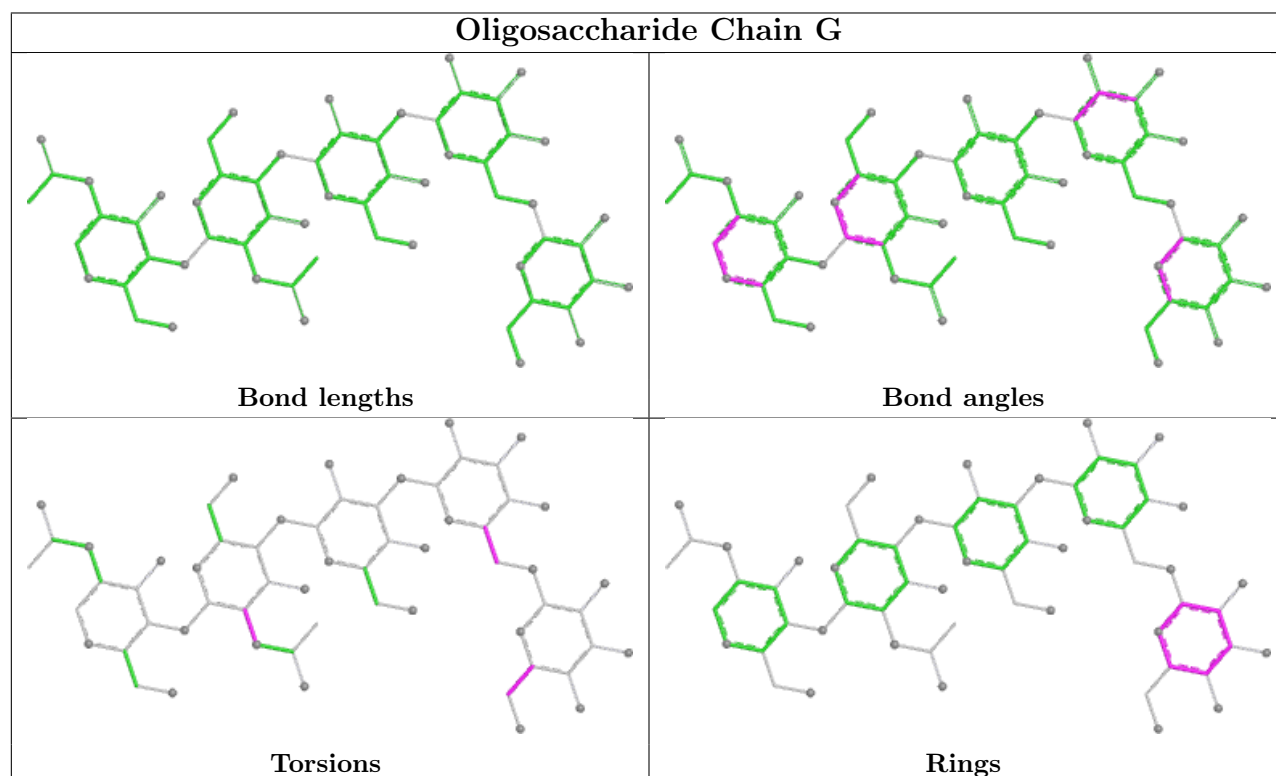
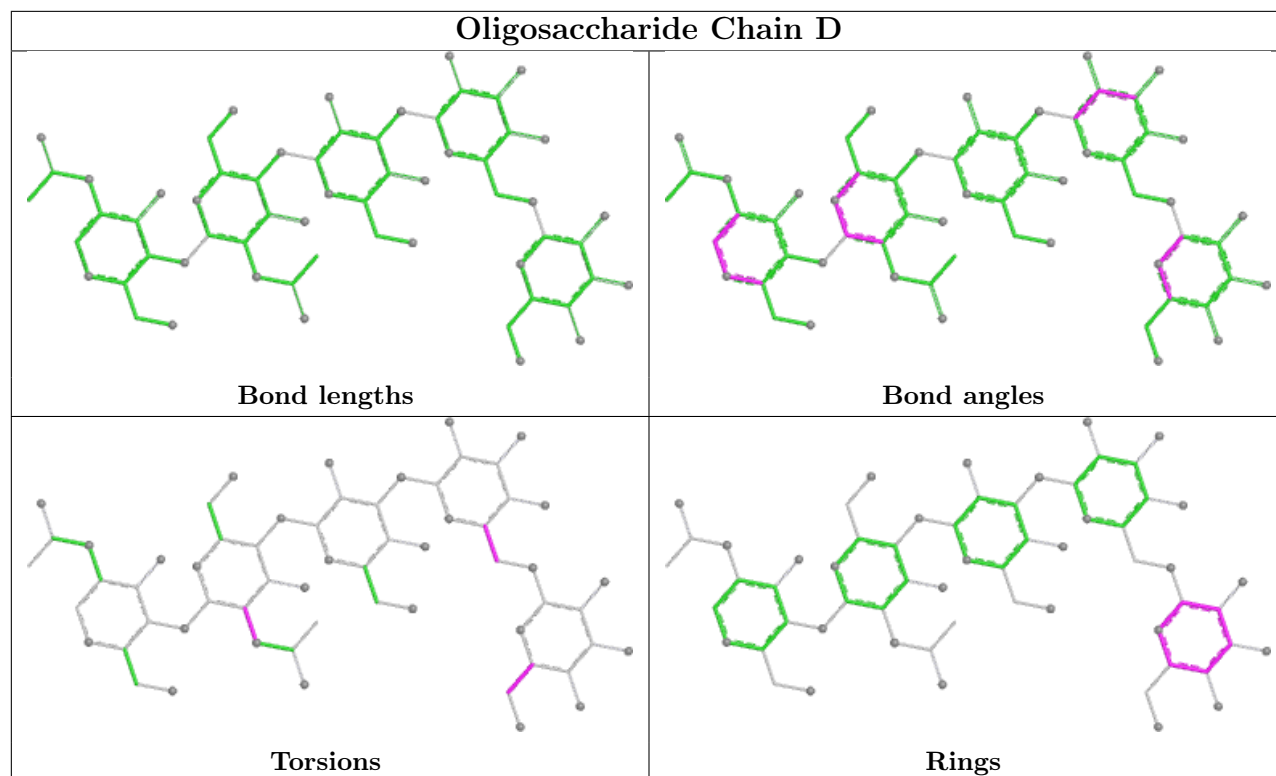
All (2) ring outliers are listed below:

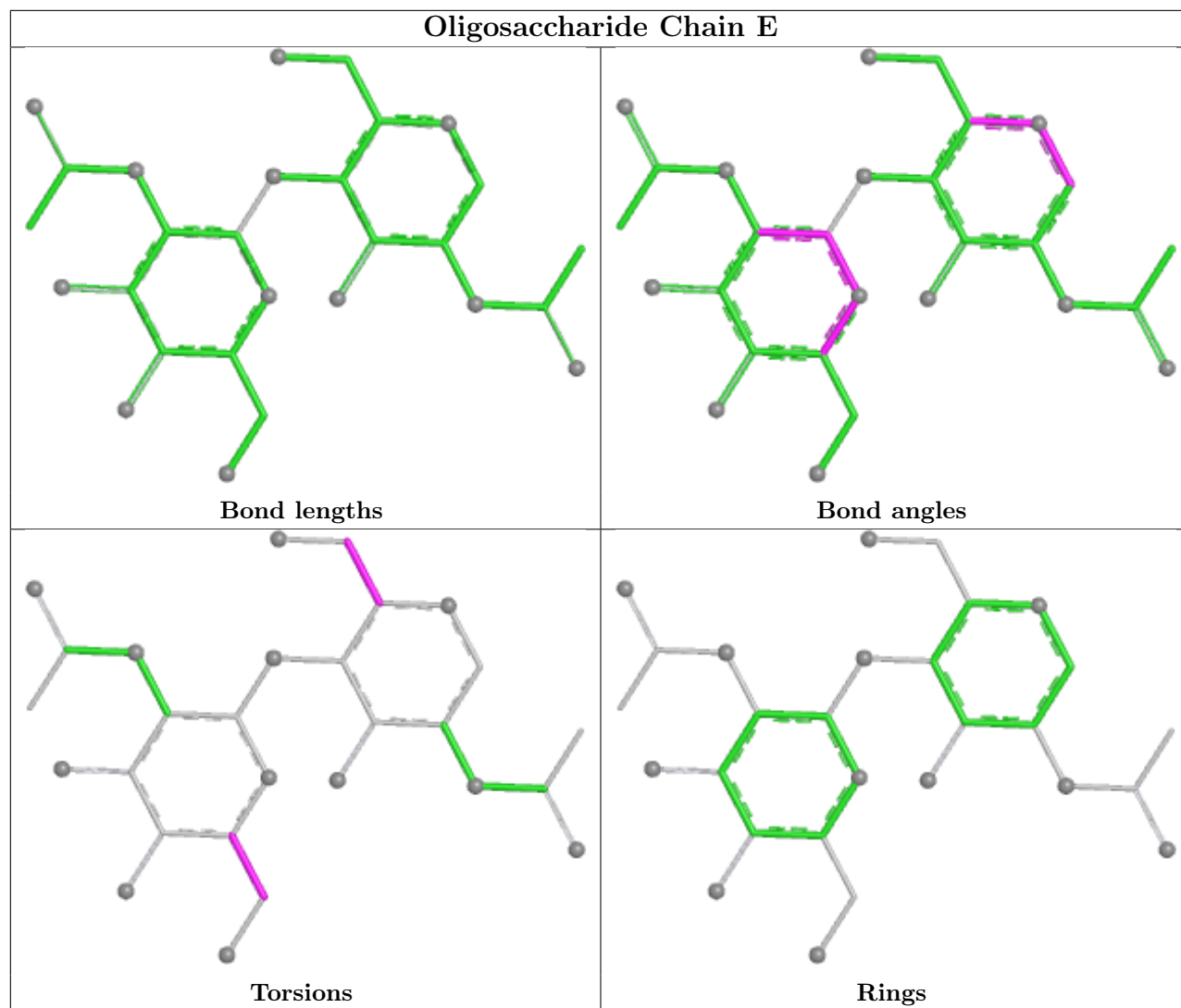
Mol	Chain	Res	Type	Atoms
3	D	5	MAN	C1-C2-C3-C4-C5-O5
3	G	5	MAN	C1-C2-C3-C4-C5-O5

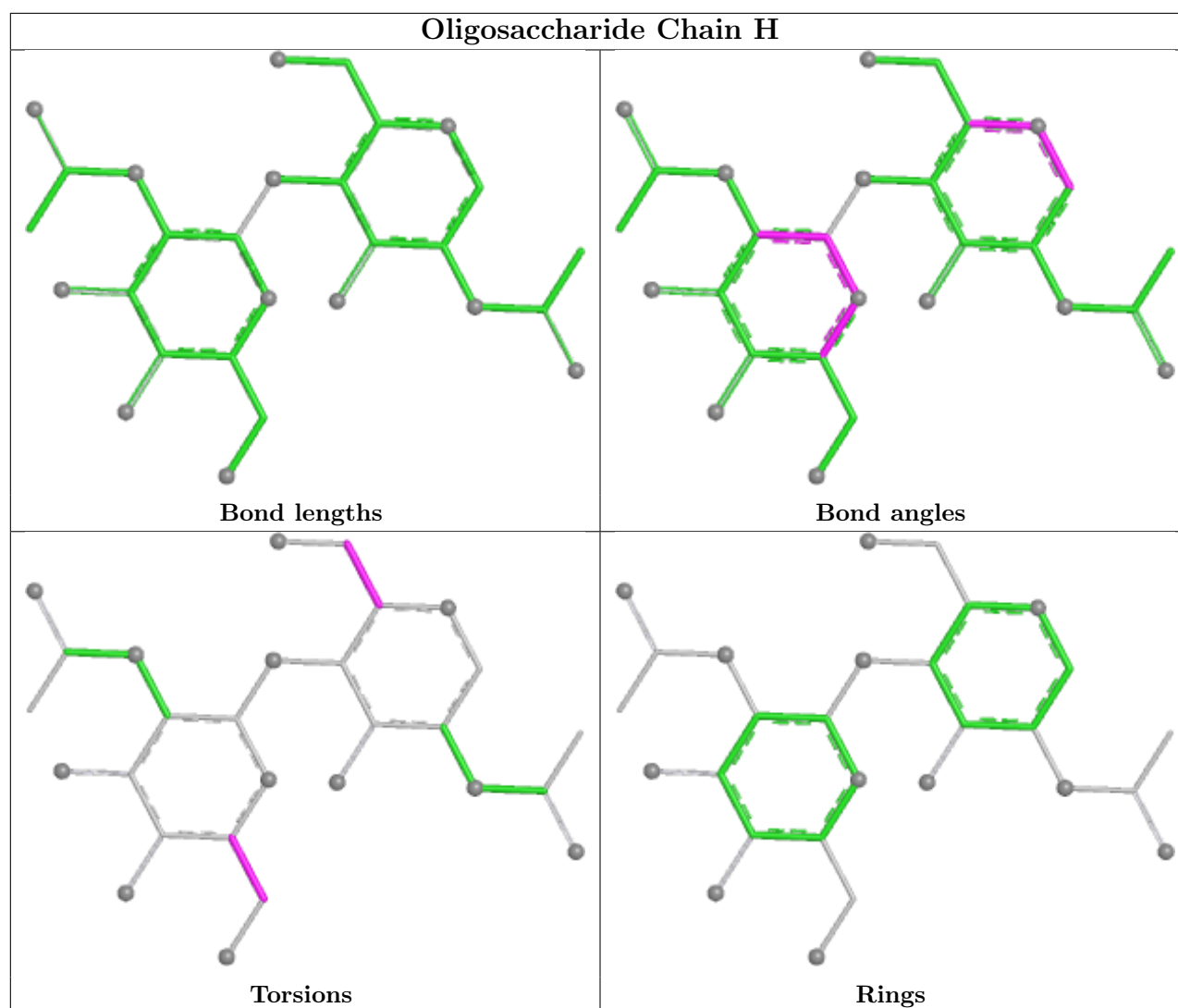
No monomer is involved in short contacts.

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](#)

Of 7 ligands modelled in this entry, 2 are monoatomic and 1 is modelled with single atom - leaving 4 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	$\# Z  > 2$
5	NAG	A	1601	1	14,14,15	0.41	0	17,19,21	0.74	0
5	NAG	B	1612	1	14,14,15	0.48	0	17,19,21	2.18	4 (23%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
5	NAG	A	1612	1	14,14,15	0.44	0	17,19,21	2.18	4 (23%)
5	NAG	B	1601	1	14,14,15	0.33	0	17,19,21	0.59	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
5	NAG	A	1601	1	-	0/6/23/26	0/1/1/1
5	NAG	B	1612	1	-	2/6/23/26	0/1/1/1
5	NAG	A	1612	1	-	2/6/23/26	0/1/1/1
5	NAG	B	1601	1	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	1612	NAG	C1-O5-C5	6.97	121.53	112.19
5	B	1612	NAG	C1-O5-C5	6.79	121.29	112.19
5	B	1612	NAG	C1-C2-N2	3.94	116.64	110.43
5	A	1612	NAG	C1-C2-N2	3.66	116.19	110.43
5	A	1612	NAG	C2-N2-C7	2.42	126.14	122.90

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	1612	NAG	C3-C2-N2-C7
5	B	1612	NAG	C3-C2-N2-C7
5	A	1612	NAG	C1-C2-N2-C7
5	B	1612	NAG	C1-C2-N2-C7

There are no ring outliers.

No monomer is involved in short contacts.

## 5.7 Other polymers ⓘ

There are no such residues in this entry.

## 5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

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

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

**Warning:** The R factor obtained from EDS is 0.2623, which does not match the depositor's R factor of 0.201. Please interpret the results in this section carefully.

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	1373/1494 (91%)	-0.02	7 (0%) 87 76	169, 249, 284, 297	0
1	B	1381/1494 (92%)	0.01	9 (0%) 84 71	160, 243, 281, 296	0
All	All	2754/2988 (92%)	-0.00	16 (0%) 85 73	160, 246, 282, 297	0

The worst 5 of 16 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	1311	MET	3.5
1	A	189	ILE	3.2
1	A	933	LEU	3.2
1	B	460	GLN	3.1
1	B	1138	LEU	2.9

### 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	MAN	D	5	11/12	-0.07	0.18	298,299,300,300	0
3	MAN	D	4	11/12	0.22	0.15	297,299,300,300	0

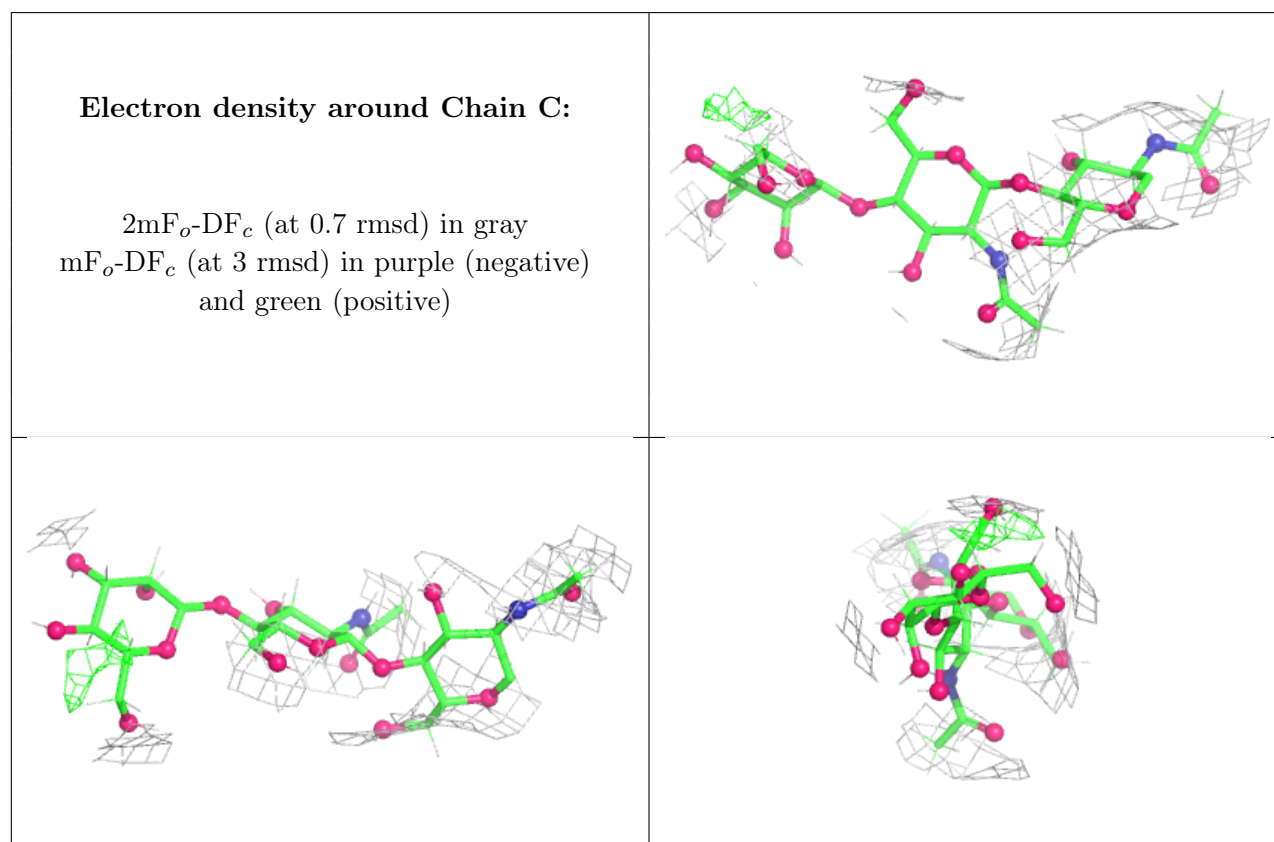
*Continued on next page...*



*Continued from previous page...*

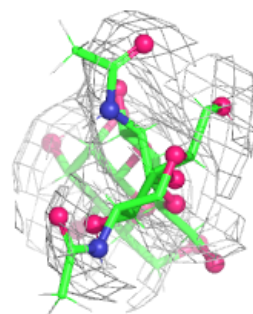
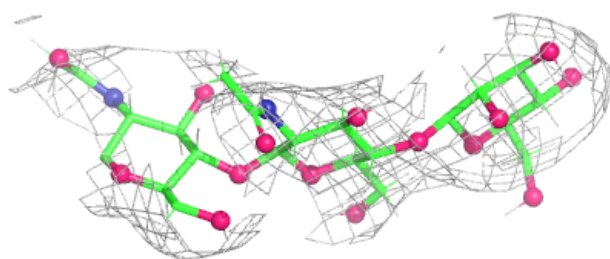
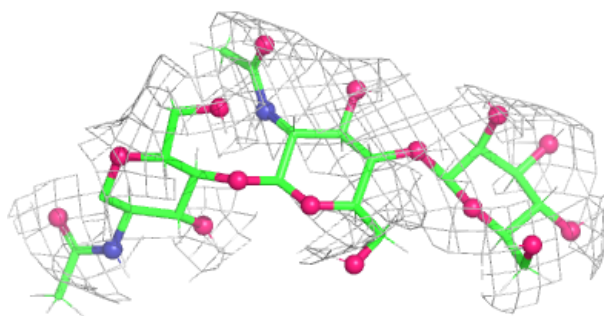
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
4	NAG	H	2	14/15	0.26	0.18	278,290,293,293	0
3	MAN	G	4	11/12	0.49	0.12	292,296,299,299	0
2	BMA	F	3	11/12	0.53	0.10	292,295,297,298	0
2	BMA	C	3	11/12	0.54	0.12	289,295,298,299	0
3	MAN	G	5	11/12	0.56	0.10	270,288,293,298	0
4	NAG	E	2	14/15	0.60	0.11	294,295,295,295	0
3	BMA	D	3	11/12	0.62	0.13	296,297,299,299	0
3	BMA	G	3	11/12	0.66	0.12	274,281,296,296	0
4	NAG	H	1	14/15	0.68	0.12	291,292,293,294	0
2	NAG	F	2	14/15	0.68	0.12	290,294,296,296	0
4	NAG	E	1	14/15	0.75	0.10	288,293,295,295	0
2	NAG	C	2	14/15	0.78	0.11	286,294,297,299	0
3	NAG	D	2	14/15	0.79	0.11	275,288,291,292	0
3	NAG	D	1	14/15	0.81	0.09	263,268,273,281	0
2	NAG	C	1	14/15	0.83	0.14	284,291,292,292	0
2	NAG	F	1	14/15	0.83	0.09	276,285,289,291	0
3	NAG	G	2	14/15	0.87	0.11	268,272,277,283	0
3	NAG	G	1	14/15	0.96	0.07	256,258,260,264	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.

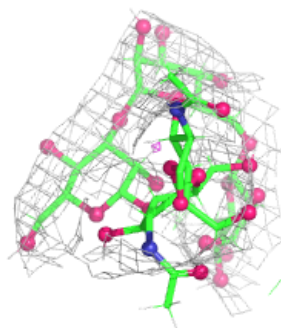
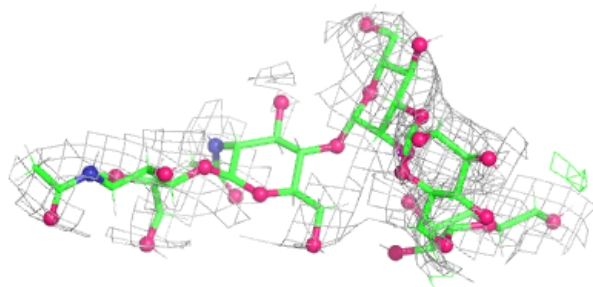
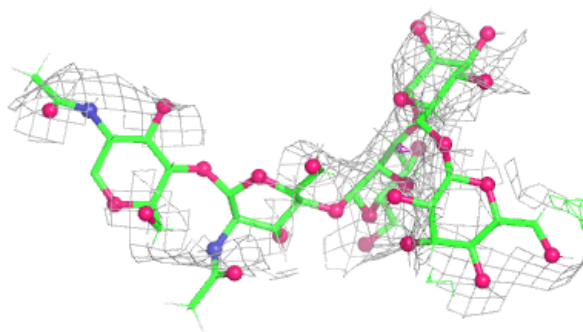


**Electron density around Chain F:**

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

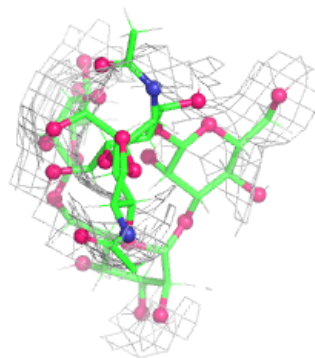
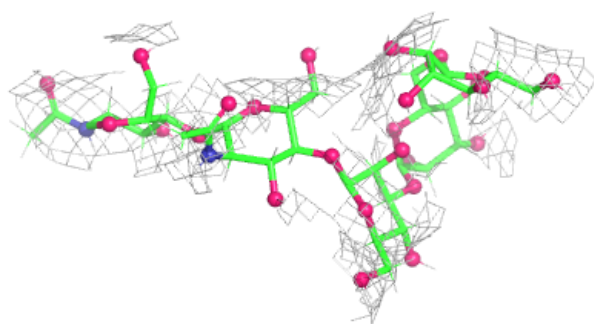
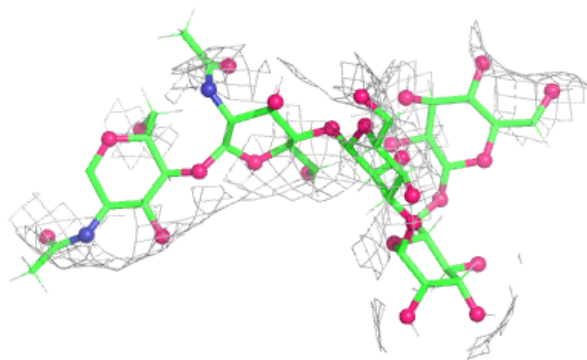
**Electron density around Chain D:**

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



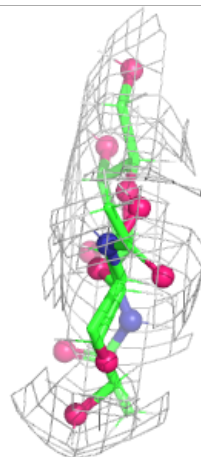
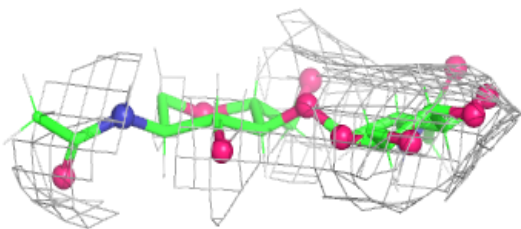
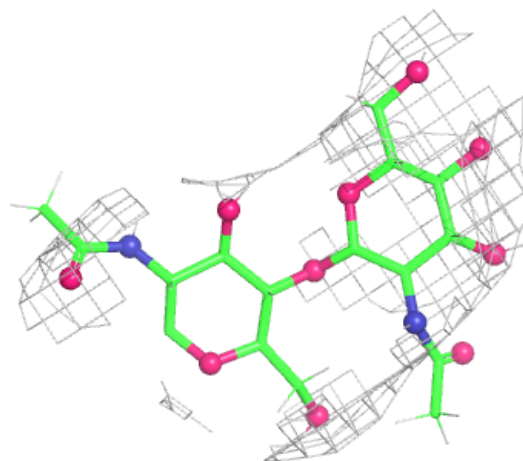
**Electron density around Chain G:**

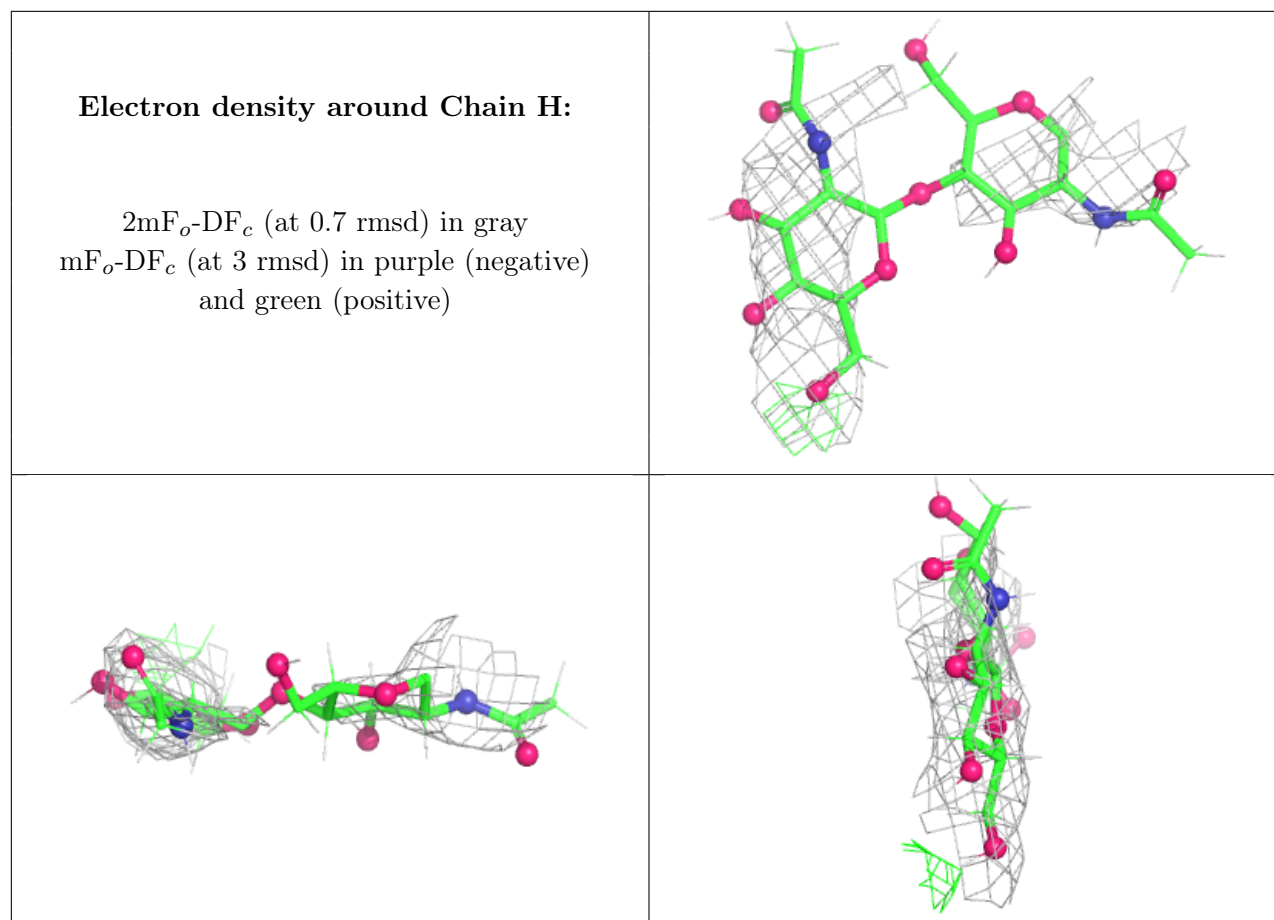
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



**Electron density around Chain E:**

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





## 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q<0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
6	CA	A	1613	1/1	-0.10	0.26	277,277,277,277	1
5	NAG	A	1601	14/15	0.49	0.17	279,286,289,290	0
5	NAG	B	1612	14/15	0.50	0.14	293,296,297,298	0
5	NAG	A	1612	14/15	0.63	0.12	296,300,300,300	0
5	NAG	B	1601	14/15	0.76	0.11	276,278,279,281	0
7	EDO	A	1614	1/4	0.82	0.19	110,110,110,110	0
6	CA	B	1613	1/1	0.86	0.16	257,257,257,257	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.