



# Full wwPDB X-ray Structure Validation Report ⓘ

Nov 9, 2024 – 11:26 pm GMT

PDB ID : 5NN5  
Title : Crystal structure of human lysosomal acid-alpha-glucosidase, GAA, in complex with 1-deoxynojirimycin  
Authors : Roig-Zamboni, V.; Cobucci-Ponzano, B.; Iacono, R.; Ferrara, M.C.; Germany, S.; Parenti, G.; Bourne, Y.; Moracci, M.  
Deposited on : 2017-04-08  
Resolution : 2.00 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:

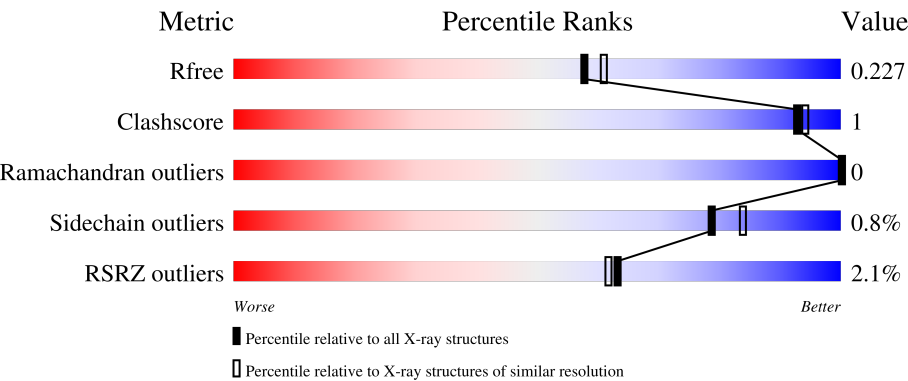
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

# 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 2.00 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.



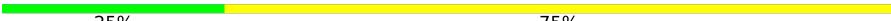
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R <sub>free</sub>	164625	9409 (2.00-2.00)
Clashscore	180529	10737 (2.00-2.00)
Ramachandran outliers	177936	10628 (2.00-2.00)
Sidechain outliers	177891	10627 (2.00-2.00)
RSRZ outliers	164620	9409 (2.00-2.00)

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	872	<div><div>2%</div><div>93%</div><div></div></div>
2	B	3	<div><div>67%</div><div>33%</div><div></div></div>
2	F	3	<div><div>33%</div><div>67%</div><div></div></div>
3	C	3	<div><div>67%</div><div>33%</div><div></div></div>
4	D	2	<div><div>100%</div><div></div><div></div></div>

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
5	E	4	 A horizontal bar chart showing the quality of chain 5. The bar is divided into two segments: a green segment on the left representing 25% and a yellow segment on the right representing 75%. The labels '25%' and '75%' are positioned below their respective segments.

## 2 Entry composition [i](#)

There are 10 unique types of molecules in this entry. The entry contains 7695 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.

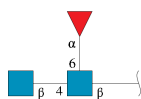
- Molecule 1 is a protein called Lysosomal alpha-glucosidase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	846	6664	4275	1117	1241	31	0	13	0

There are 3 discrepancies between the modelled and reference sequences:

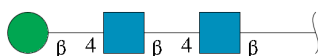
Chain	Residue	Modelled	Actual	Comment	Reference
A	199	ARG	HIS	conflict	UNP P10253
A	223	HIS	ARG	conflict	UNP P10253
A	780	ILE	VAL	conflict	UNP P10253

- Molecule 2 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.



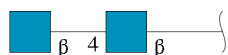
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	B	3	38	22	2	14	0	0	0
2	F	3	38	22	2	14	0	0	0

- 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	C	3	Total	C	N	O	0	0	0
			39	22	2	15			

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



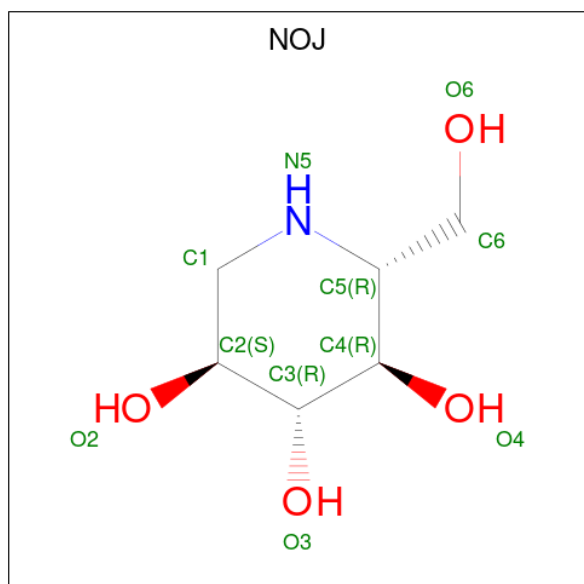
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
4	D	2	Total	C	N	O	0	0	0
			28	16	2	10			

- Molecule 5 is an oligosaccharide called 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.



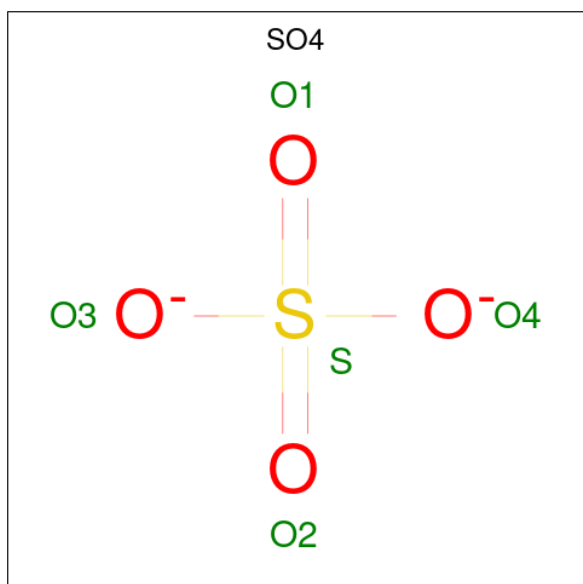
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
5	E	4	Total	C	N	O	0	0	0
			50	28	2	20			

- Molecule 6 is 1-DEOXYNOJIRIMYCIN (three-letter code: NOJ) (formula: C<sub>6</sub>H<sub>13</sub>NO<sub>4</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	A	1	Total	C	N	O	0	0
			11	6	1	4		

- Molecule 7 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	A	1	Total	O	S	0	0
			5	4	1		
7	A	1	Total	O	S	0	0
			5	4	1		
7	A	1	Total	O	S	0	0
			5	4	1		
7	A	1	Total	O	S	0	0
			5	4	1		

- Molecule 8 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	6	Total	Cl	0	0
			6	6		

- Molecule 9 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
9	A	1	Total	C	O	0	0
			4	2	2		
9	A	1	Total	C	O	0	0
			4	2	2		
9	A	1	Total	C	O	0	0
			4	2	2		
9	A	1	Total	C	O	0	0
			4	2	2		
9	A	1	Total	C	O	0	0
			4	2	2		
9	A	1	Total	C	O	0	0
			4	2	2		

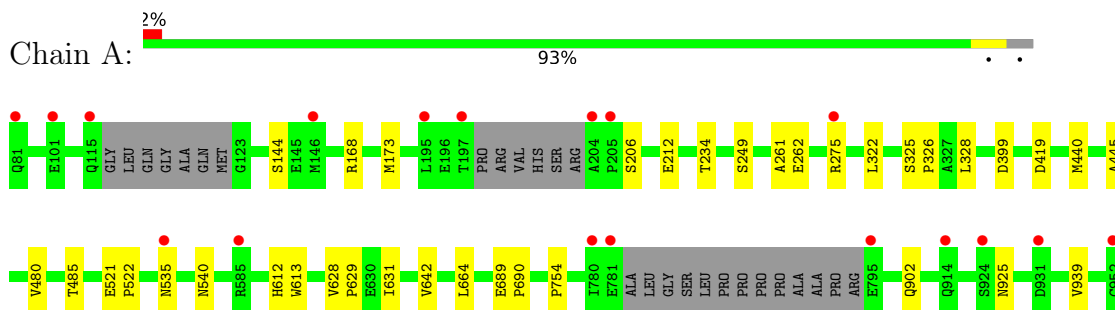
- Molecule 10 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
10	A	773	Total	O	0	0
			773	773		

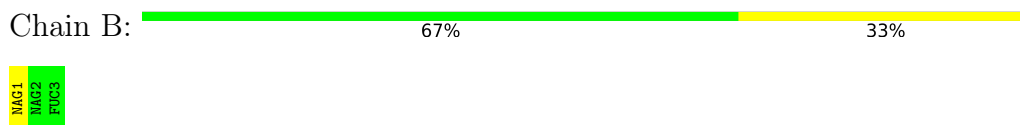
### 3 Residue-property plots

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 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: Lysosomal alpha-glucosidase



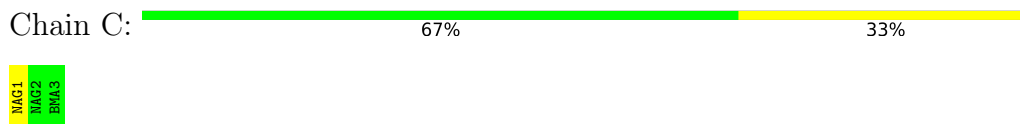
- Molecule 2: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-[alpha-L-fucopyranose-(1-6)]2-acetamido-2-deoxy-beta-D-glucopyranose



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



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



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



NAG1  
NAG2

- Molecule 5: 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 E:  25% 75%

NAG1  
NAG2  
BMA3  
MAN4

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	97.49Å 102.97Å 129.03Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	53.86 – 2.00 53.86 – 2.00	Depositor EDS
% Data completeness (in resolution range)	94.3 (53.86-2.00) 94.3 (53.86-2.00)	Depositor EDS
$R_{merge}$	0.17	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.98 (at 2.00Å)	Xtriage
Refinement program	REFMAC 5.7.0029	Depositor
R, $R_{free}$	0.183 , 0.218 0.194 , 0.227	Depositor DCC
$R_{free}$ test set	4191 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	15.7	Xtriage
Anisotropy	0.144	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 44.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	7695	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	19.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.06% 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: SO4, CL, CSO, NAG, FUC, MAN, EDO, BMA, NOJ

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	0.48	0/6861	0.66	0/9376

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	6664	0	6427	19	0
2	B	38	0	34	0	0
2	F	38	0	34	0	0
3	C	39	0	34	0	0
4	D	28	0	25	0	0
5	E	50	0	43	0	0
6	A	11	0	13	0	0
7	A	20	0	0	0	0
8	A	6	0	0	0	0
9	A	28	0	42	2	0
10	A	773	0	0	5	0
All	All	7695	0	6652	19	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 1.

All (19) 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:925:ASN:HA	10:A:1638:HOH:O	1.91	0.70
1:A:206:SER:O	9:A:1030:EDO:H22	2.02	0.59
1:A:275:ARG:NH2	10:A:1101:HOH:O	2.11	0.59
1:A:168[A]:ARG:NH1	10:A:1105:HOH:O	2.40	0.53
1:A:445:ALA:HB1	1:A:485:THR:HB	1.92	0.51
1:A:325:SER:N	1:A:326:PRO:HA	2.27	0.50
1:A:521:GLU:N	1:A:522:PRO:HA	2.27	0.49
1:A:419:ASP:HB2	10:A:1159:HOH:O	2.13	0.47
1:A:212[A]:GLU:HG3	10:A:1263:HOH:O	2.17	0.45
1:A:612:HIS:O	1:A:642:VAL:HA	2.16	0.45
1:A:261:ALA:HB1	1:A:262:GLU:HA	2.00	0.44
1:A:234:THR:HA	1:A:249:SER:O	2.18	0.43
1:A:664:LEU:HD13	1:A:754:PRO:HG3	2.00	0.43
1:A:689:GLU:HB3	1:A:690:PRO:HD3	2.01	0.43
1:A:322:LEU:CD2	1:A:328:LEU:HD12	2.48	0.42
1:A:902:GLN:HA	1:A:939:VAL:O	2.19	0.42
1:A:535:ASN:OD1	1:A:540:ASN:ND2	2.53	0.42
1:A:173:MET:O	9:A:1027:EDO:H22	2.19	0.42
1:A:628:VAL:HB	1:A:629:PRO:HD3	2.01	0.41

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	839/872 (96%)	821 (98%)	18 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 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	724/741 (98%)	718 (99%)	6 (1%)	79 84

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

Mol	Chain	Res	Type
1	A	144[A]	SER
1	A	399	ASP
1	A	440	MET
1	A	480	VAL
1	A	613	TRP
1	A	631	ILE

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

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

1 non-standard protein/DNA/RNA residue is 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	$\# Z  > 2$
1	CSO	A	938	1	3,6,7	0.76	0	0,6,8	-	-

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
1	CSO	A	938	1	-	0/1/5/7	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.5 Carbohydrates

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	B	1	2,1	14,14,15	0.49	0	17,19,21	1.06	1 (5%)
2	NAG	B	2	2	14,14,15	0.44	0	17,19,21	0.81	0
2	FUC	B	3	2	10,10,11	0.56	0	14,14,16	0.54	0
3	NAG	C	1	3,1	14,14,15	0.61	0	17,19,21	1.01	1 (5%)
3	NAG	C	2	3	14,14,15	0.60	0	17,19,21	0.79	0
3	BMA	C	3	3	11,11,12	0.34	0	15,15,17	0.86	0
4	NAG	D	1	4,1	14,14,15	0.54	0	17,19,21	0.91	0
4	NAG	D	2	4	14,14,15	0.62	0	17,19,21	0.84	0
5	NAG	E	1	5,1	14,14,15	0.77	1 (7%)	17,19,21	0.91	0
5	NAG	E	2	5	14,14,15	0.51	0	17,19,21	1.06	1 (5%)
5	BMA	E	3	5	11,11,12	0.35	0	15,15,17	1.14	1 (6%)
5	MAN	E	4	5	11,11,12	0.55	0	15,15,17	0.75	0
2	NAG	F	1	2,1	14,14,15	0.58	0	17,19,21	0.95	1 (5%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	NAG	F	2	2	14,14,15	0.49	0	17,19,21	0.88	1 (5%)
2	FUC	F	3	2	10,10,11	0.57	0	14,14,16	0.74	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	B	1	2,1	-	1/6/23/26	0/1/1/1
2	NAG	B	2	2	-	2/6/23/26	0/1/1/1
2	FUC	B	3	2	-	-	0/1/1/1
3	NAG	C	1	3,1	-	0/6/23/26	0/1/1/1
3	NAG	C	2	3	-	2/6/23/26	0/1/1/1
3	BMA	C	3	3	-	0/2/19/22	0/1/1/1
4	NAG	D	1	4,1	-	2/6/23/26	0/1/1/1
4	NAG	D	2	4	-	2/6/23/26	0/1/1/1
5	NAG	E	1	5,1	-	0/6/23/26	0/1/1/1
5	NAG	E	2	5	-	0/6/23/26	0/1/1/1
5	BMA	E	3	5	-	2/2/19/22	0/1/1/1
5	MAN	E	4	5	-	2/2/19/22	0/1/1/1
2	NAG	F	1	2,1	-	0/6/23/26	0/1/1/1
2	NAG	F	2	2	-	2/6/23/26	0/1/1/1
2	FUC	F	3	2	-	-	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	E	1	NAG	O5-C1	-2.36	1.39	1.43

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	E	3	BMA	O5-C5-C6	3.51	112.71	107.20
2	B	1	NAG	O5-C1-C2	-2.79	106.88	111.29
5	E	2	NAG	C1-O5-C5	2.67	115.81	112.19
2	F	2	NAG	O5-C5-C6	2.47	111.07	107.20
3	C	1	NAG	O4-C4-C3	-2.38	104.84	110.35
2	F	1	NAG	O5-C1-C2	-2.36	107.57	111.29

There are no chirality outliers.

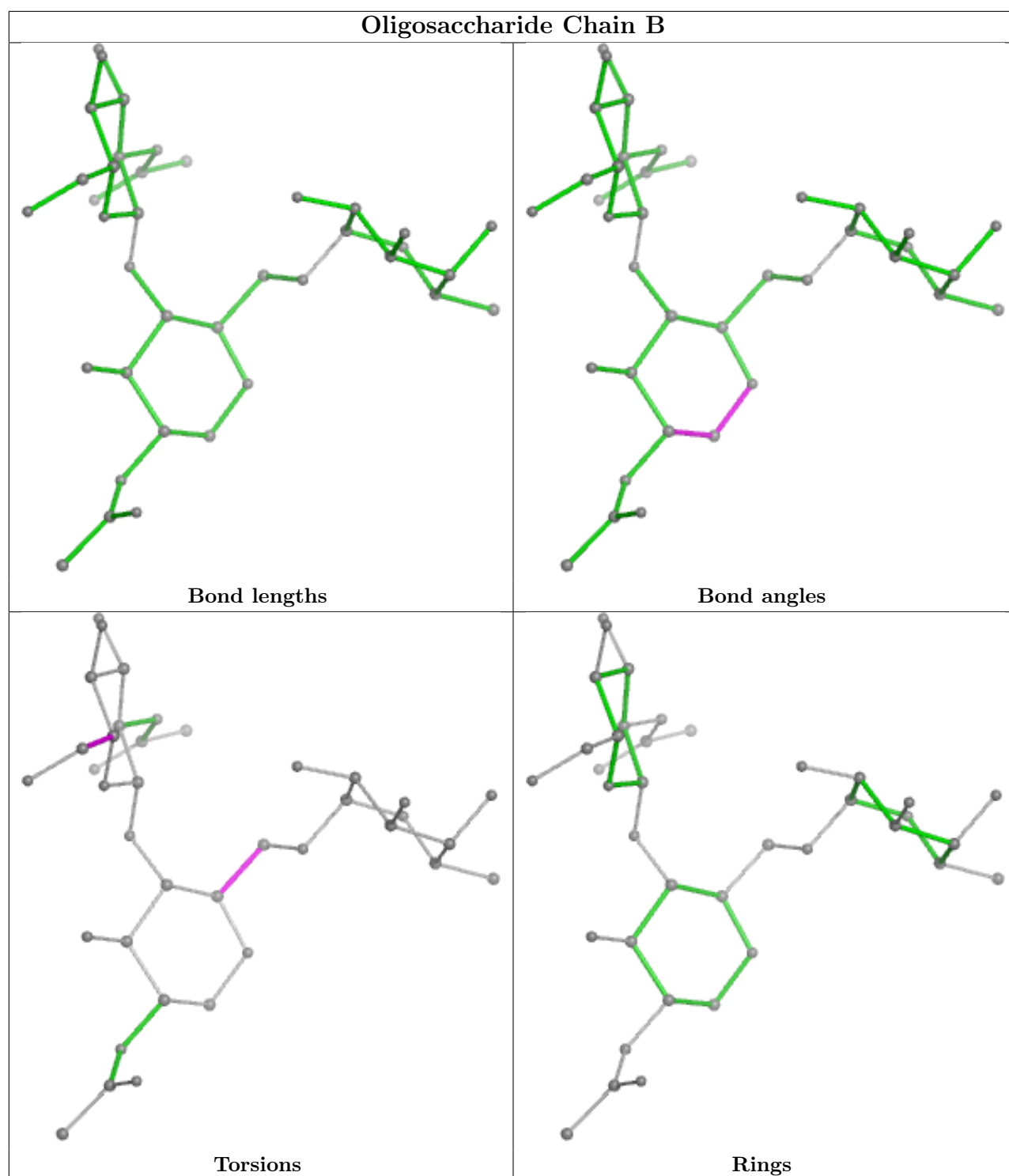
All (15) torsion outliers are listed below:

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

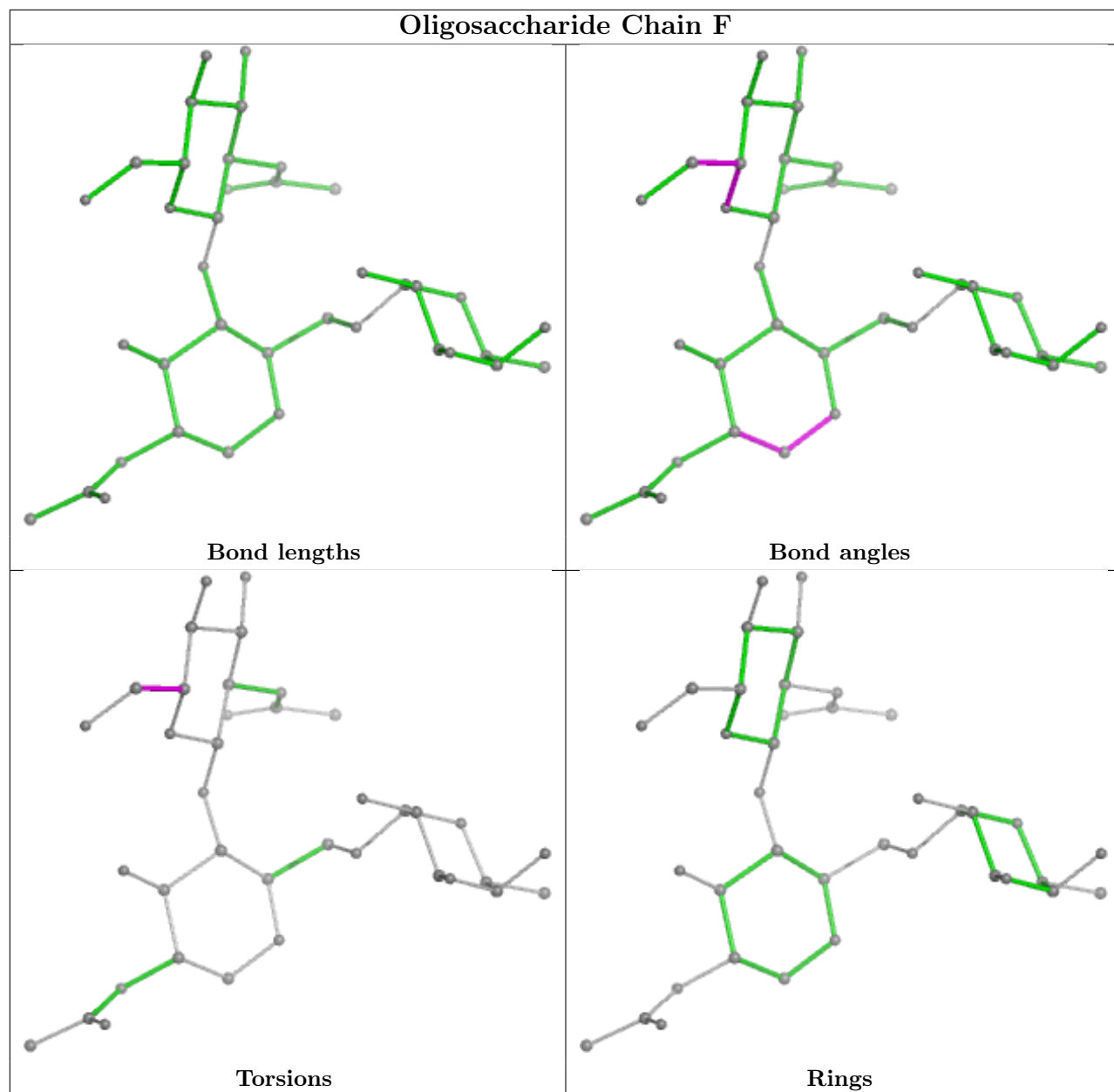
There are no ring outliers.

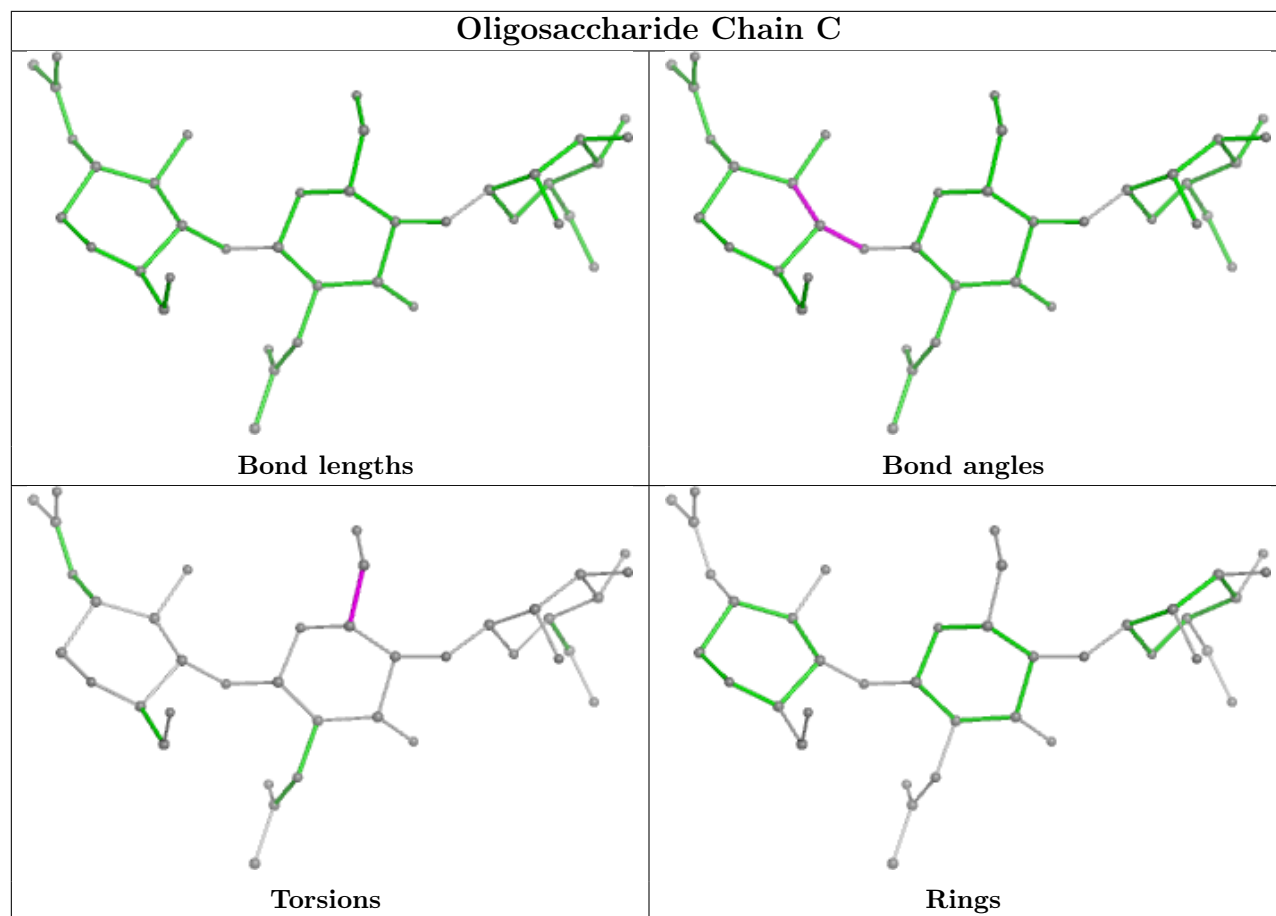
No monomer is involved in short contacts.

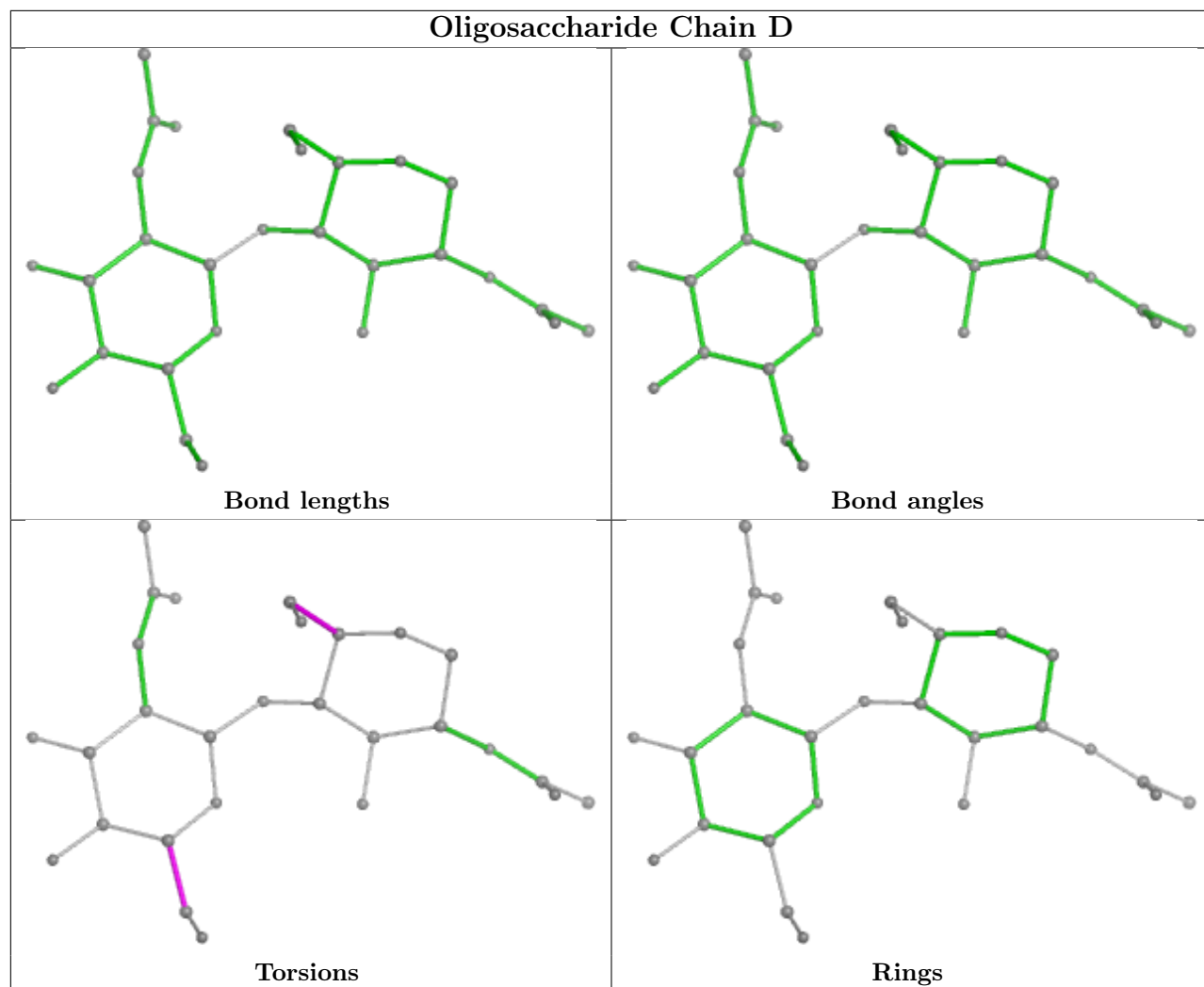
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for oligosaccharide.

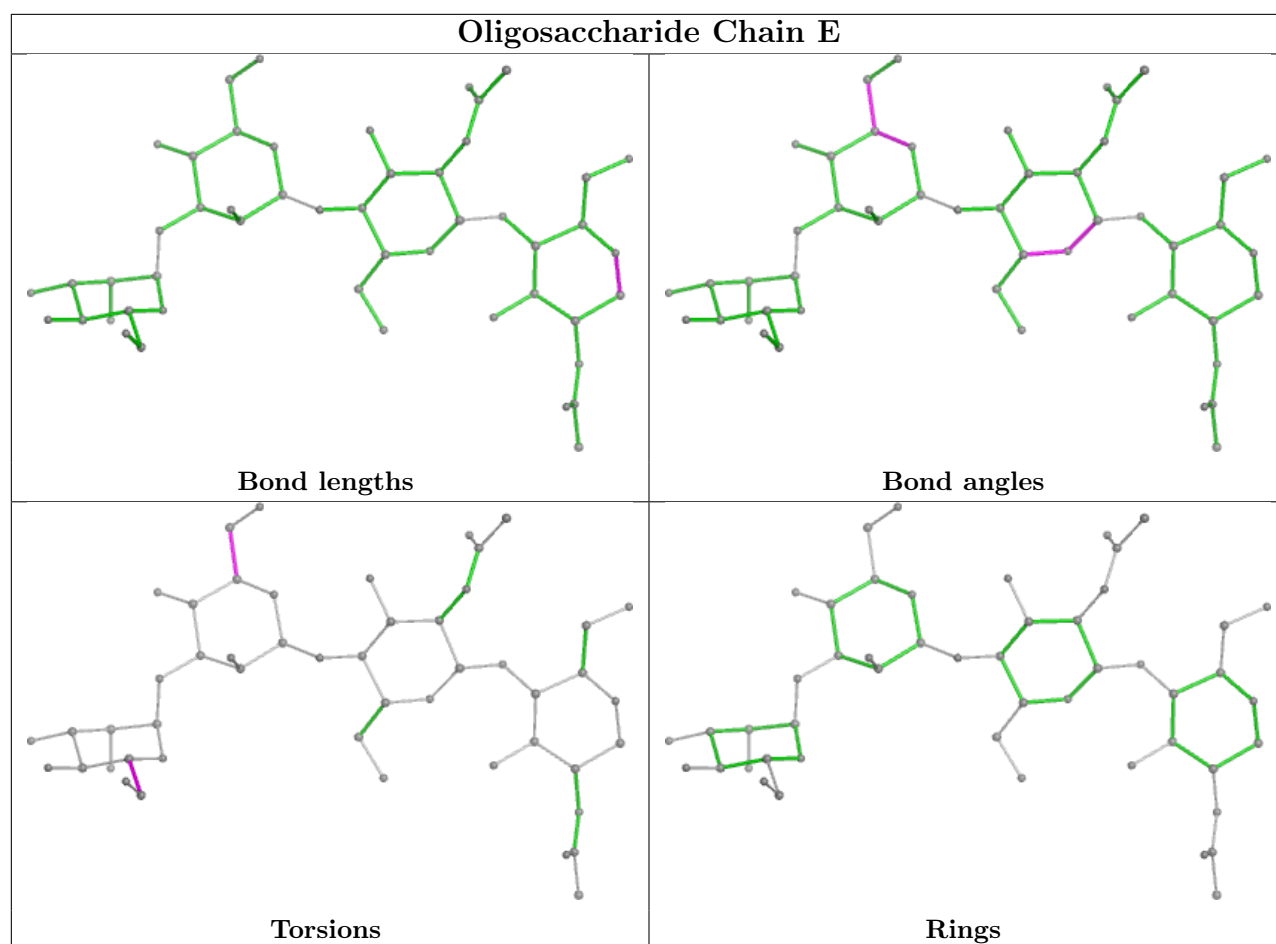


## Oligosaccharide Chain F









## 5.6 Ligand geometry [i](#)

Of 18 ligands modelled in this entry, 6 are monoatomic - leaving 12 for Mogul analysis.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
9	EDO	A	1032	-	3,3,3	0.49	0	2,2,2	0.24	0
9	EDO	A	1029	-	3,3,3	0.27	0	2,2,2	0.65	0
7	SO4	A	1017	-	4,4,4	0.31	0	6,6,6	0.40	0
9	EDO	A	1027	-	3,3,3	0.68	0	2,2,2	0.63	0
9	EDO	A	1033	-	3,3,3	0.50	0	2,2,2	0.43	0
9	EDO	A	1031	-	3,3,3	0.56	0	2,2,2	0.18	0
9	EDO	A	1028	-	3,3,3	0.37	0	2,2,2	0.60	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
7	SO4	A	1020	-	4,4,4	0.52	0	6,6,6	0.26	0
7	SO4	A	1019	-	4,4,4	0.40	0	6,6,6	0.25	0
7	SO4	A	1018	-	4,4,4	0.47	0	6,6,6	0.42	0
9	EDO	A	1030	-	3,3,3	0.44	0	2,2,2	0.26	0
6	NOJ	A	1016	-	11,11,11	0.89	0	13,15,15	1.50	1 (7%)

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
9	EDO	A	1032	-	-	0/1/1/1	-
9	EDO	A	1029	-	-	1/1/1/1	-
6	NOJ	A	1016	-	-	0/2/19/19	0/1/1/1
9	EDO	A	1027	-	-	1/1/1/1	-
9	EDO	A	1033	-	-	1/1/1/1	-
9	EDO	A	1031	-	-	1/1/1/1	-
9	EDO	A	1028	-	-	0/1/1/1	-
9	EDO	A	1030	-	-	0/1/1/1	-

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	A	1016	NOJ	C1-N5-C5	5.15	120.75	109.61

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
9	A	1033	EDO	O1-C1-C2-O2
9	A	1029	EDO	O1-C1-C2-O2
9	A	1027	EDO	O1-C1-C2-O2
9	A	1031	EDO	O1-C1-C2-O2

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
9	A	1027	EDO	1	0
9	A	1030	EDO	1	0

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

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

There are no chain breaks in this entry.

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

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

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	845/872 (96%)	-0.20	18 (2%) 63 62	9, 15, 32, 64	2 (0%)

All (18) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	780	ILE	4.8
1	A	197	THR	4.6
1	A	195	LEU	3.5
1	A	146	MET	3.5
1	A	781	GLU	3.5
1	A	924	SER	3.3
1	A	795	GLU	3.3
1	A	205	PRO	3.2
1	A	81	GLN	2.8
1	A	204	ALA	2.6
1	A	914	GLN	2.5
1	A	115	GLN	2.4
1	A	535	ASN	2.4
1	A	101	GLU	2.3
1	A	931	ASP	2.3
1	A	585	ARG	2.1
1	A	952	CYS	2.0
1	A	275	ARG	2.0

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

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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
1	CSO	A	938	7/8	0.96	0.07	19,21,23,28	0

### 6.3 Carbohydrates

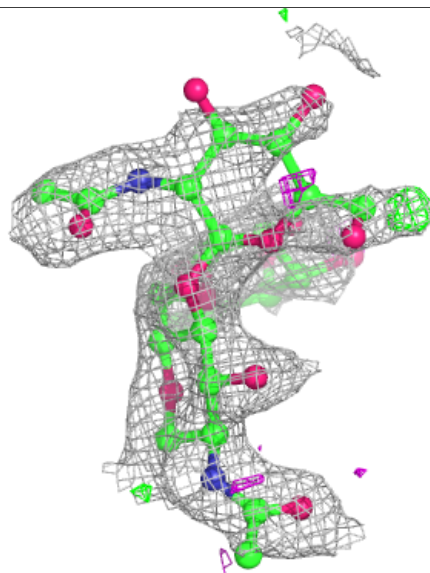
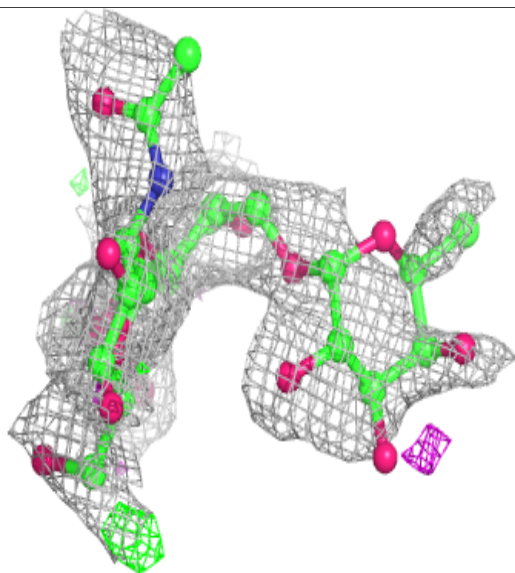
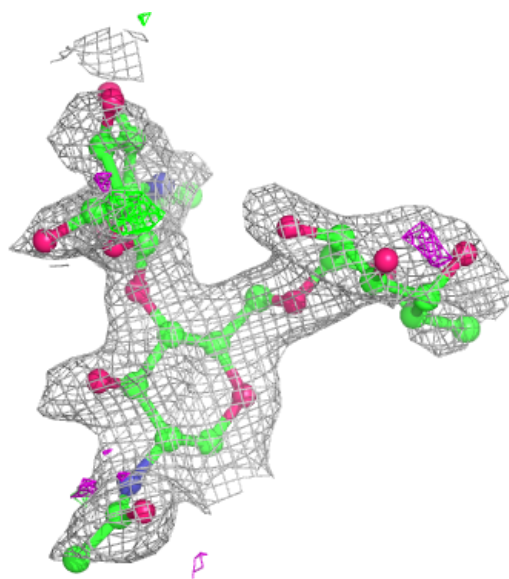
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
5	MAN	E	4	11/12	0.41	0.21	66,80,88,91	0
3	BMA	C	3	11/12	0.45	0.20	70,84,90,91	0
4	NAG	D	2	14/15	0.53	0.26	53,64,73,77	0
2	FUC	F	3	10/11	0.59	0.25	65,77,79,79	0
2	FUC	B	3	10/11	0.61	0.23	55,63,72,72	0
2	NAG	F	2	14/15	0.62	0.19	67,79,90,94	0
2	NAG	B	2	14/15	0.67	0.18	43,61,75,76	0
5	BMA	E	3	11/12	0.72	0.17	42,57,65,72	0
3	NAG	C	2	14/15	0.72	0.17	47,56,77,84	0
2	NAG	F	1	14/15	0.75	0.17	34,46,72,77	0
4	NAG	D	1	14/15	0.89	0.12	17,27,41,49	0
2	NAG	B	1	14/15	0.89	0.12	25,33,44,51	0
3	NAG	C	1	14/15	0.92	0.09	19,26,28,35	0
5	NAG	E	2	14/15	0.92	0.09	18,24,29,35	0
5	NAG	E	1	14/15	0.96	0.07	14,19,21,23	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.

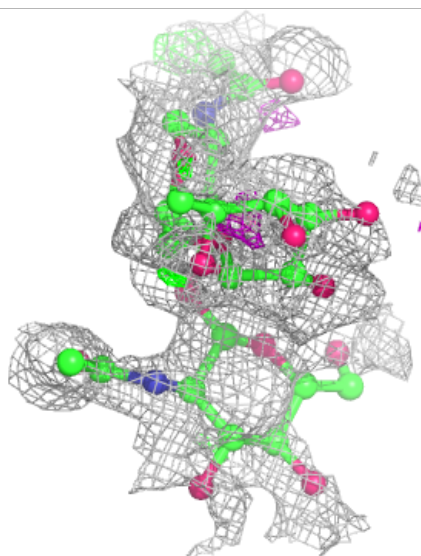
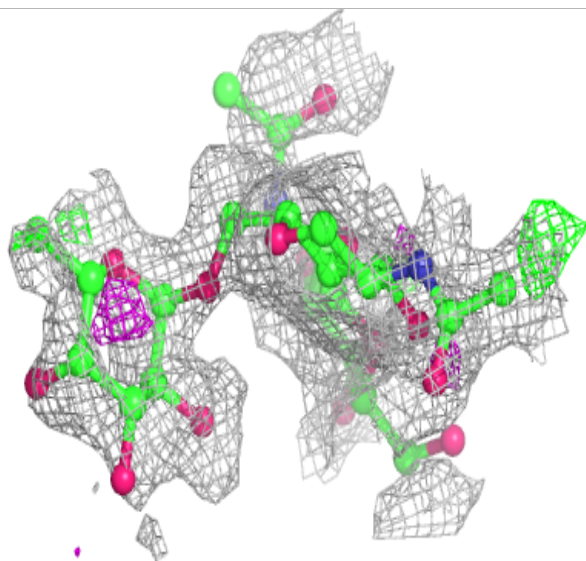
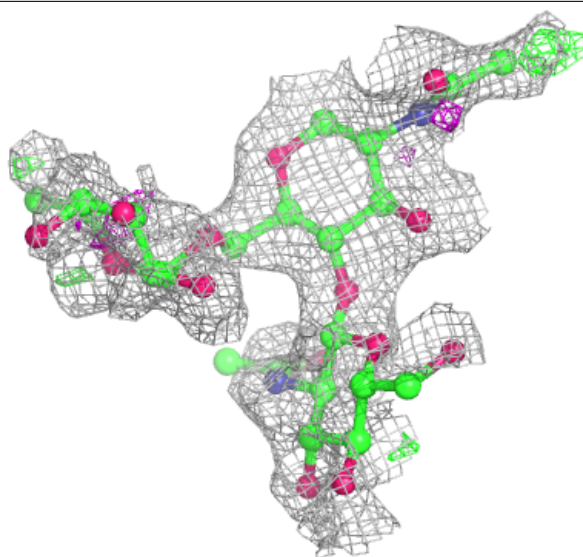
**Electron density around Chain B:**

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



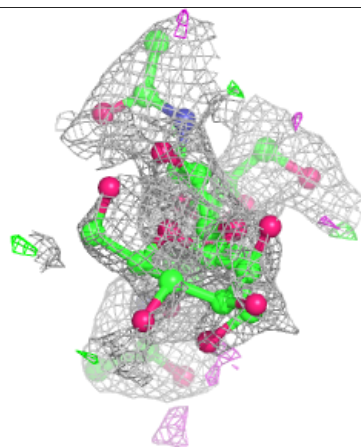
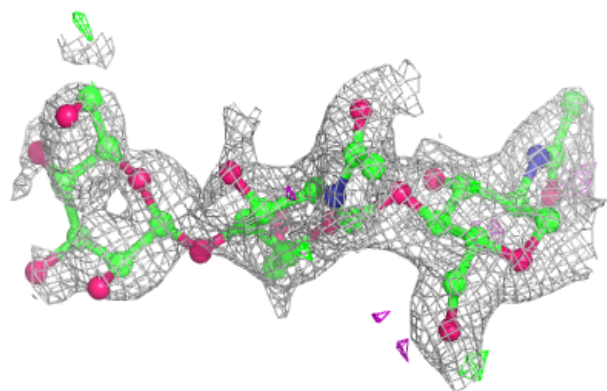
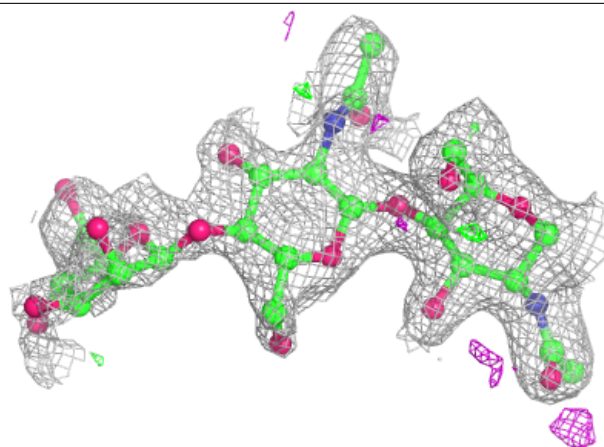
**Electron density around Chain F:**

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



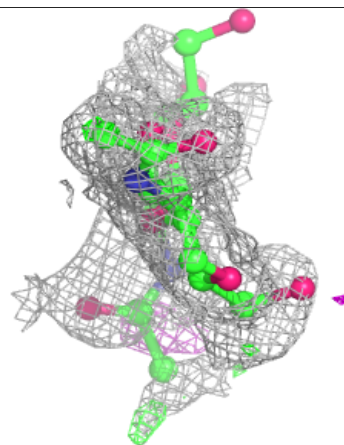
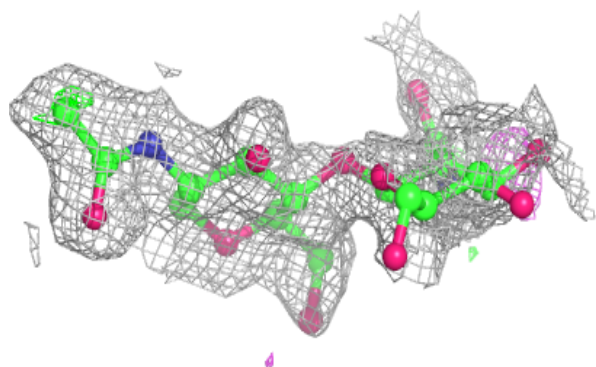
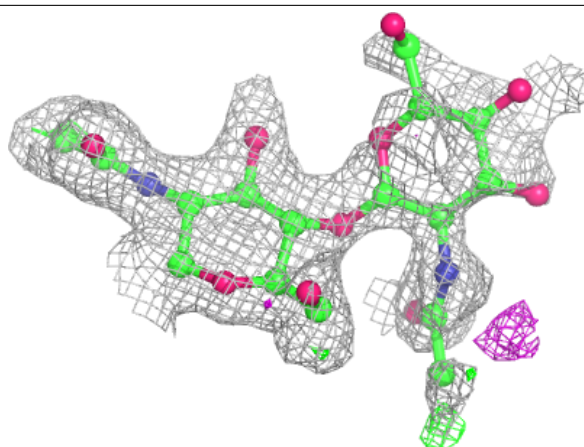
**Electron density around Chain C:**

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

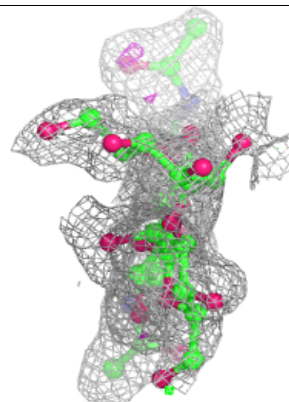
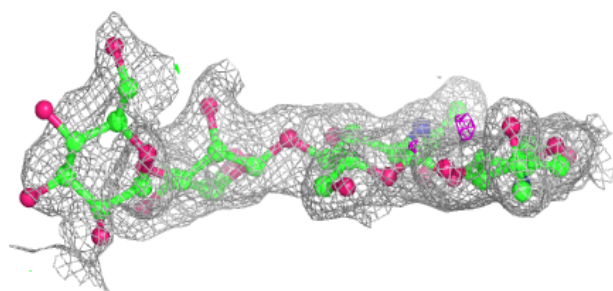
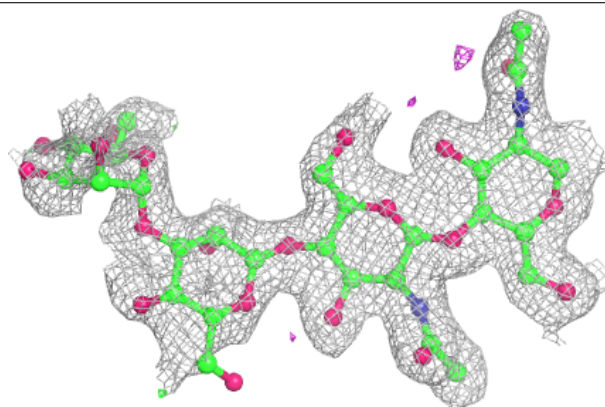


**Electron density around Chain D:**

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

**Electron density around Chain 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( $\text{\AA}^2$ )	Q<0.9
7	SO4	A	1020	5/5	0.61	0.16	46,55,60,71	0
9	EDO	A	1033	4/4	0.65	0.20	43,43,47,53	0
9	EDO	A	1031	4/4	0.76	0.16	42,43,43,45	0
7	SO4	A	1019	5/5	0.77	0.15	53,62,73,73	0
9	EDO	A	1032	4/4	0.81	0.16	37,37,39,40	0
9	EDO	A	1029	4/4	0.81	0.16	26,30,30,31	0
8	CL	A	1026	1/1	0.83	0.18	59,59,59,59	0
9	EDO	A	1028	4/4	0.84	0.18	41,42,46,47	0
9	EDO	A	1030	4/4	0.86	0.13	37,38,39,41	0
9	EDO	A	1027	4/4	0.87	0.20	30,33,34,39	0
8	CL	A	1021	1/1	0.89	0.13	50,50,50,50	0
8	CL	A	1025	1/1	0.90	0.10	50,50,50,50	0
8	CL	A	1022	1/1	0.92	0.10	42,42,42,42	0
7	SO4	A	1018	5/5	0.94	0.15	24,26,33,38	0
6	NOJ	A	1016	11/11	0.96	0.06	11,12,13,13	0
7	SO4	A	1017	5/5	0.97	0.11	29,30,35,38	0
8	CL	A	1024	1/1	0.98	0.05	28,28,28,28	0
8	CL	A	1023	1/1	0.98	0.07	23,23,23,23	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.