



# Full wwPDB X-ray Structure Validation Report ⓘ

Jan 7, 2024 – 03:12 am GMT

PDB ID : 5MGD  
Title : STRUCTURE OF E298Q-BETA-GALACTOSIDASE FROM AS-  
PERGILLUS NIGER IN COMPLEX WITH 6-Galactosyl-lactose  
Authors : Rico-Diaz, A.; Ramirez-Escudero, M.; Vizoso Vazquez, A.; Cerdan, M.E.; Be-  
cerra, M.; Sanz-Aparicio, J.  
Deposited on : 2016-11-21  
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 : 1.8.4, CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.36  
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) : 2.36

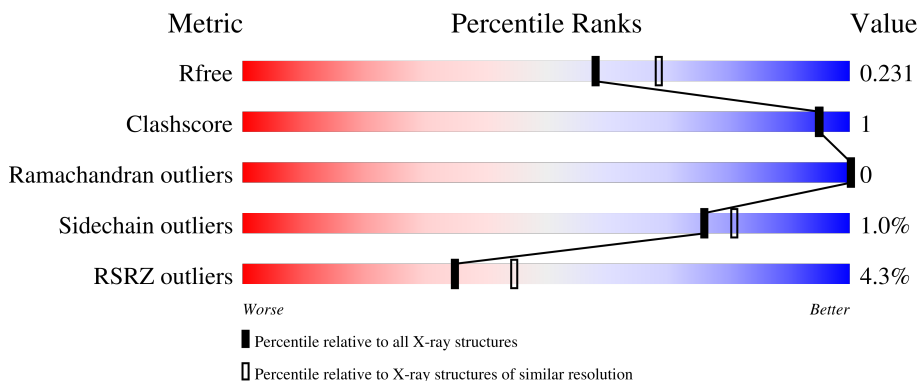
# 1 Overall quality at a glance

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	1013	 4% 92%
2	B	8	 25% 62% 12%
3	C	10	 10% 80% 10%
4	D	6	 17% 67% 17%
5	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:

<b>Mol</b>	<b>Type</b>	<b>Chain</b>	<b>Res</b>	<b>Chirality</b>	<b>Geometry</b>	<b>Clashes</b>	<b>Electron density</b>
5	GLC	E	1	-	-	-	X
5	GAL	E	2	-	-	-	X
6	NAG	A	1156	-	-	-	X

## 2 Entry composition [i](#)

There are 10 unique types of molecules in this entry. The entry contains 8088 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.

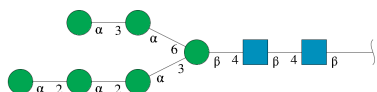
- Molecule 1 is a protein called Probable beta-galactosidase A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	968	7482	4780	1220	1471	11	0	0	0

There are 7 discrepancies between the modelled and reference sequences:

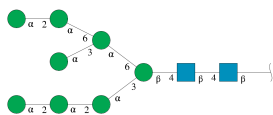
Chain	Residue	Modelled	Actual	Comment	Reference
A	298	GLN	GLU	engineered mutation	UNP A2QAN3
A	1008	HIS	-	expression tag	UNP A2QAN3
A	1009	HIS	-	expression tag	UNP A2QAN3
A	1010	HIS	-	expression tag	UNP A2QAN3
A	1011	HIS	-	expression tag	UNP A2QAN3
A	1012	HIS	-	expression tag	UNP A2QAN3
A	1013	HIS	-	expression tag	UNP A2QAN3

- Molecule 2 is an oligosaccharide called alpha-D-mannopyranose-(1-2)-alpha-D-mannopyranose-(1-2)-alpha-D-mannopyranose-(1-3)-[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.



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	B	8	94	52	2	40	0	0	0

- Molecule 3 is an oligosaccharide called alpha-D-mannopyranose-(1-2)-alpha-D-mannopyranose-(1-2)-alpha-D-mannopyranose-(1-3)-[alpha-D-mannopyranose-(1-2)-alpha-D-mannopyranose-(1-6)-[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.



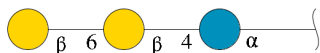
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
3	C	10	116	64	2	50	0	0	0

- Molecule 4 is an oligosaccharide called alpha-D-mannopyranose-(1-2)-alpha-D-mannopyranose-(1-2)-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
			Total	C	N	O			
4	D	6	72	40	2	30	0	0	0

- Molecule 5 is an oligosaccharide called beta-D-galactopyranose-(1-6)-beta-D-galactopyranose-(1-4)-alpha-D-glucopyranose.



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace
			Total	C	O			
5	E	3	34	18	16	0	0	0

- Molecule 6 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
			Total	C	N	O		
6	A	1	Total 14	8	1	5	0	0
6	A	1	Total 14	8	1	5	0	0
6	A	1	Total 14	8	1	5	0	0
6	A	1	Total 14	8	1	5	0	0
6	A	1	Total 14	8	1	5	0	0
6	A	1	Total 14	8	1	5	0	0

- Molecule 7 is DIMETHYL SULFOXIDE (three-letter code: DMS) (formula: C<sub>2</sub>H<sub>6</sub>OS).

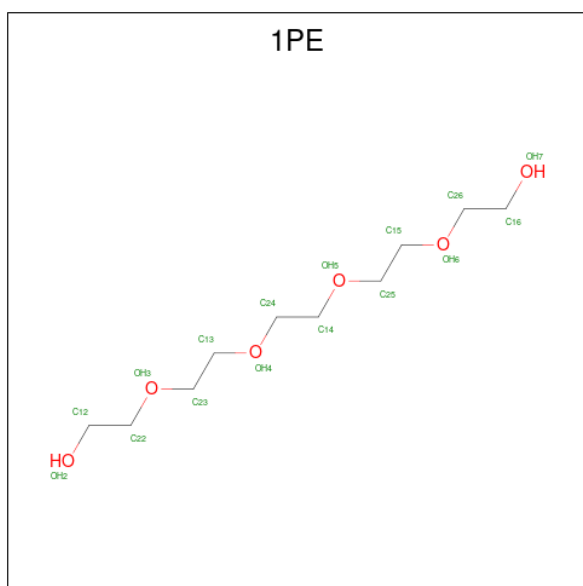


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	O	S		
7	A	1	4	2	1	1	0	0

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

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Cl		
8	A	1	1	1	0	0

- Molecule 9 is PENTAETHYLENE GLYCOL (three-letter code: 1PE) (formula: C<sub>10</sub>H<sub>22</sub>O<sub>6</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
9	A	1	Total	C	O	0	0
			16	10	6		

- Molecule 10 is water.

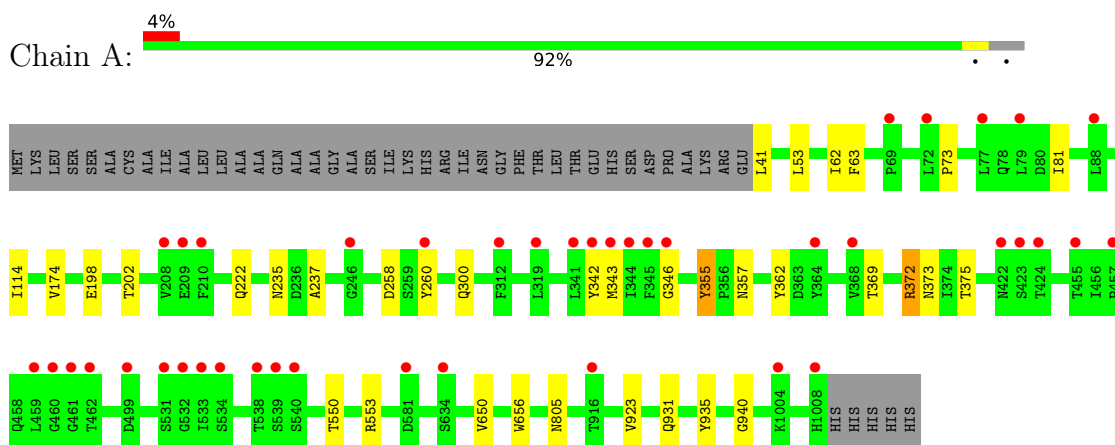
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
10	A	185	Total	O	0	0
			185	185		



### 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.

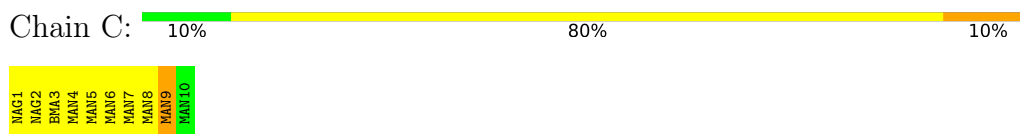
- Molecule 1: Probable beta-galactosidase A



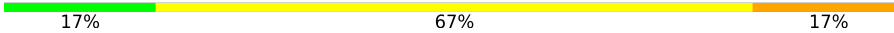
- Molecule 2: alpha-D-mannopyranose-(1-2)-alpha-D-mannopyranose-(1-2)-alpha-D-mannopyranose-(1-3)-[alpha-D-mannopyranose-(1-3)-alpha-D-mannopyranose-(1-6)]beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

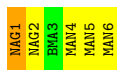


- Molecule 3: alpha-D-mannopyranose-(1-2)-alpha-D-mannopyranose-(1-2)-alpha-D-mannopyranose-(1-3)-[alpha-D-mannopyranose-(1-2)-alpha-D-mannopyranose-(1-6)-[alpha-D-mannopyranose-(1-3)]alpha-D-mannopyranose-(1-6)]beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose



- Molecule 4: alpha-D-mannopyranose-(1-2)-alpha-D-mannopyranose-(1-2)-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:  17% 67% 17%



- Molecule 5: beta-D-galactopyranose-(1-6)-beta-D-galactopyranose-(1-4)-alpha-D-glucopyranose

Chain E:  33% 67%



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	58.28Å 105.87Å 83.72Å 90.00° 99.24° 90.00°	Depositor
Resolution (Å)	82.63 – 2.15 46.11 – 2.15	Depositor EDS
% Data completeness (in resolution range)	99.9 (82.63-2.15) 100.0 (46.11-2.15)	Depositor EDS
$R_{merge}$	0.14	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.06 (at 2.16Å)	Xtrriage
Refinement program	REFMAC 5.8.0155	Depositor
R, $R_{free}$	0.191 , 0.227 0.201 , 0.231	Depositor DCC
$R_{free}$ test set	2768 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	18.9	Xtrriage
Anisotropy	1.137	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 39.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.30$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	8088	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	29.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.71% 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: DMS, NAG, GLC, CL, BMA, MAN, GAL, 1PE

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	1/7685 (0.0%)	0.65	1/10488 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	372	ARG	C-N	-10.24	1.10	1.34

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	372	ARG	O-C-N	-5.11	114.52	122.70

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	7482	0	7179	16	0
2	B	94	0	79	2	0
3	C	116	0	97	1	0
4	D	72	0	61	1	0
5	E	34	0	30	2	0
6	A	84	0	78	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	A	4	0	6	0	0
8	A	1	0	0	0	0
9	A	16	0	22	0	0
10	A	185	0	0	0	0
All	All	8088	0	7552	18	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 (18) 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)
5:E:1:GLC:H62	5:E:2:GAL:C1	2.00	0.92
1:A:375:THR:HG21	2:B:1:NAG:H82	1.85	0.56
1:A:940:GLY:HA2	4:D:1:NAG:H83	1.93	0.51
1:A:202:THR:HG21	1:A:237:ALA:HB3	1.93	0.50
1:A:53:LEU:HD12	1:A:62:ILE:HG13	1.94	0.49
1:A:550:THR:HG21	3:C:9:MAN:H62	1.95	0.48
1:A:174:VAL:HG11	1:A:222:GLN:HB3	1.97	0.46
1:A:346:GLY:HA3	1:A:362:TYR:O	2.15	0.46
1:A:650:VAL:HG22	1:A:656:TRP:CD1	2.50	0.46
1:A:260:TYR:HA	1:A:300:GLN:HB2	1.99	0.45
1:A:369:THR:HB	2:B:1:NAG:H62	1.98	0.45
1:A:73:PRO:HB2	1:A:114:ILE:HB	1.98	0.44
1:A:81:ILE:HD11	1:A:372:ARG:HG2	2.01	0.42
1:A:355:TYR:CZ	1:A:357:ASN:HB2	2.55	0.42
1:A:923:VAL:HB	1:A:935:TYR:HB3	2.02	0.41
5:E:1:GLC:C6	5:E:2:GAL:C1	2.78	0.40
1:A:235:ASN:HB2	1:A:258:ASP:OD1	2.21	0.40
1:A:342:TYR:HA	1:A:343:MET:HA	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	966/1013 (95%)	934 (97%)	32 (3%)	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	808/842 (96%)	800 (99%)	8 (1%)	76	81

All (8) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	41	LEU
1	A	63	PHE
1	A	198	GLU
1	A	355	TYR
1	A	373	ASN
1	A	553	ARG
1	A	805	ASN
1	A	931	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	466	ASN
1	A	911	ASN

### 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](#)

27 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	1,2	14,14,15	2.23	8 (57%)	17,19,21	1.55	3 (17%)
2	NAG	B	2	2	14,14,15	0.32	0	17,19,21	0.89	0
2	BMA	B	3	2	11,11,12	0.25	0	15,15,17	1.41	2 (13%)
2	MAN	B	4	2	11,11,12	0.27	0	15,15,17	1.26	1 (6%)
2	MAN	B	5	2	11,11,12	0.27	0	15,15,17	1.15	1 (6%)
2	MAN	B	6	2	11,11,12	0.47	0	15,15,17	0.72	0
2	MAN	B	7	2	11,11,12	0.57	0	15,15,17	2.63	3 (20%)
2	MAN	B	8	2	11,11,12	0.33	0	15,15,17	1.58	1 (6%)
3	NAG	C	1	1,3	14,14,15	0.34	0	17,19,21	1.23	3 (17%)
3	MAN	C	10	3	11,11,12	0.39	0	15,15,17	0.71	0
3	NAG	C	2	3	14,14,15	0.46	0	17,19,21	0.87	1 (5%)
3	BMA	C	3	3	11,11,12	0.30	0	15,15,17	1.05	1 (6%)
3	MAN	C	4	3	11,11,12	0.23	0	15,15,17	1.43	1 (6%)
3	MAN	C	5	3	11,11,12	0.27	0	15,15,17	0.84	1 (6%)
3	MAN	C	6	3	11,11,12	0.38	0	15,15,17	1.28	1 (6%)
3	MAN	C	7	3	11,11,12	0.40	0	15,15,17	0.92	1 (6%)
3	MAN	C	8	3	11,11,12	0.40	0	15,15,17	0.81	1 (6%)
3	MAN	C	9	3	11,11,12	0.52	0	15,15,17	1.09	2 (13%)
4	NAG	D	1	4,1	14,14,15	0.38	0	17,19,21	0.98	1 (5%)
4	NAG	D	2	4	14,14,15	0.36	0	17,19,21	1.02	1 (5%)
4	BMA	D	3	4	11,11,12	0.52	0	15,15,17	1.16	0
4	MAN	D	4	4	11,11,12	0.46	0	15,15,17	1.34	3 (20%)
4	MAN	D	5	4	11,11,12	0.35	0	15,15,17	0.97	1 (6%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	MAN	D	6	4	11,11,12	0.32	0	15,15,17	0.91	1 (6%)
5	GLC	E	1	5	12,12,12	1.01	1 (8%)	17,17,17	0.90	1 (5%)
5	GAL	E	2	5	11,11,12	1.60	1 (9%)	15,15,17	2.52	7 (46%)
5	GAL	E	3	5	11,11,12	1.18	2 (18%)	15,15,17	2.12	4 (26%)

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	1,2	-	0/6/23/26	0/1/1/1
2	NAG	B	2	2	-	0/6/23/26	0/1/1/1
2	BMA	B	3	2	-	0/2/19/22	0/1/1/1
2	MAN	B	4	2	-	0/2/19/22	0/1/1/1
2	MAN	B	5	2	-	0/2/19/22	0/1/1/1
2	MAN	B	6	2	-	2/2/19/22	0/1/1/1
2	MAN	B	7	2	-	0/2/19/22	0/1/1/1
2	MAN	B	8	2	-	1/2/19/22	0/1/1/1
3	NAG	C	1	1,3	-	1/6/23/26	0/1/1/1
3	MAN	C	10	3	-	2/2/19/22	0/1/1/1
3	NAG	C	2	3	-	0/6/23/26	0/1/1/1
3	BMA	C	3	3	-	0/2/19/22	0/1/1/1
3	MAN	C	4	3	-	0/2/19/22	0/1/1/1
3	MAN	C	5	3	-	1/2/19/22	0/1/1/1
3	MAN	C	6	3	-	0/2/19/22	0/1/1/1
3	MAN	C	7	3	-	2/2/19/22	0/1/1/1
3	MAN	C	8	3	-	2/2/19/22	0/1/1/1
3	MAN	C	9	3	-	1/2/19/22	0/1/1/1
4	NAG	D	1	4,1	-	0/6/23/26	0/1/1/1
4	NAG	D	2	4	-	0/6/23/26	0/1/1/1
4	BMA	D	3	4	-	0/2/19/22	0/1/1/1
4	MAN	D	4	4	-	0/2/19/22	0/1/1/1
4	MAN	D	5	4	-	2/2/19/22	0/1/1/1
4	MAN	D	6	4	-	0/2/19/22	0/1/1/1
5	GLC	E	1	5	-	0/2/22/22	0/1/1/1
5	GAL	E	2	5	-	0/2/19/22	0/1/1/1
5	GAL	E	3	5	-	1/2/19/22	0/1/1/1

All (12) bond length outliers are listed below:



Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	E	2	GAL	O6-C6	-5.05	1.21	1.42
2	B	1	NAG	C1-C2	4.37	1.58	1.52
2	B	1	NAG	O5-C1	-3.27	1.38	1.43
5	E	1	GLC	O4-C4	-3.12	1.35	1.43
2	B	1	NAG	C3-C2	-2.73	1.46	1.52
2	B	1	NAG	O7-C7	-2.66	1.17	1.23
2	B	1	NAG	O3-C3	-2.54	1.37	1.43
5	E	3	GAL	O5-C1	-2.48	1.39	1.43
2	B	1	NAG	O4-C4	-2.41	1.37	1.43
5	E	3	GAL	O2-C2	-2.29	1.38	1.43
2	B	1	NAG	C2-N2	-2.10	1.42	1.46
2	B	1	NAG	O5-C5	-2.00	1.39	1.43

All (42) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	7	MAN	C1-O5-C5	7.89	122.89	112.19
2	B	8	MAN	C1-O5-C5	5.26	119.32	112.19
5	E	3	GAL	O2-C2-C1	-5.06	98.80	109.15
5	E	2	GAL	O6-C6-C5	4.78	127.69	111.29
3	C	6	MAN	C1-O5-C5	4.65	118.49	112.19
2	B	7	MAN	O5-C1-C2	4.62	117.90	110.77
3	C	4	MAN	C1-O5-C5	4.47	118.25	112.19
2	B	4	MAN	C1-O5-C5	4.15	117.82	112.19
5	E	2	GAL	O5-C1-C2	-4.04	104.53	110.77
2	B	3	BMA	C1-O5-C5	3.98	117.58	112.19
5	E	2	GAL	O5-C5-C6	3.90	113.31	107.20
5	E	2	GAL	O4-C4-C3	-3.74	101.70	110.35
2	B	7	MAN	C1-C2-C3	3.70	114.21	109.67
2	B	5	MAN	C1-O5-C5	3.69	117.19	112.19
5	E	3	GAL	O5-C1-C2	-3.37	105.56	110.77
2	B	1	NAG	O3-C3-C2	-3.31	102.61	109.47
5	E	3	GAL	O2-C2-C3	3.12	116.40	110.14
4	D	4	MAN	C1-O5-C5	3.10	116.39	112.19
5	E	1	GLC	O4-C4-C3	-3.04	103.32	110.35
3	C	9	MAN	O5-C5-C6	2.91	111.77	107.20
5	E	3	GAL	C1-C2-C3	-2.79	106.24	109.67
4	D	5	MAN	C1-O5-C5	2.69	115.83	112.19
2	B	1	NAG	C2-N2-C7	2.64	126.66	122.90
2	B	3	BMA	C6-C5-C4	-2.61	106.90	113.00
2	B	1	NAG	O5-C1-C2	-2.58	107.21	111.29
3	C	5	MAN	C1-O5-C5	2.51	115.60	112.19
3	C	7	MAN	O6-C6-C5	-2.50	102.70	111.29

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	D	6	MAN	C1-O5-C5	2.49	115.56	112.19
5	E	2	GAL	O2-C2-C1	2.46	114.18	109.15
4	D	4	MAN	O2-C2-C3	-2.45	105.23	110.14
3	C	3	BMA	O5-C5-C6	2.44	111.03	107.20
3	C	8	MAN	C1-O5-C5	2.38	115.42	112.19
3	C	1	NAG	O5-C1-C2	-2.29	107.67	111.29
5	E	2	GAL	C2-C3-C4	-2.29	106.93	110.89
3	C	1	NAG	C1-C2-N2	2.11	114.09	110.49
3	C	1	NAG	C1-O5-C5	2.11	115.05	112.19
3	C	9	MAN	C1-C2-C3	2.08	112.22	109.67
4	D	4	MAN	O5-C1-C2	-2.08	107.56	110.77
4	D	1	NAG	O5-C1-C2	-2.07	108.02	111.29
5	E	2	GAL	C3-C4-C5	2.07	113.92	110.24
4	D	2	NAG	C1-O5-C5	2.07	114.99	112.19
3	C	2	NAG	C1-C2-N2	-2.03	107.01	110.49

There are no chirality outliers.

All (15) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	D	5	MAN	C4-C5-C6-O6
4	D	5	MAN	O5-C5-C6-O6
2	B	6	MAN	O5-C5-C6-O6
2	B	6	MAN	C4-C5-C6-O6
3	C	9	MAN	O5-C5-C6-O6
3	C	7	MAN	C4-C5-C6-O6
5	E	3	GAL	O5-C5-C6-O6
2	B	8	MAN	O5-C5-C6-O6
3	C	10	MAN	C4-C5-C6-O6
3	C	7	MAN	O5-C5-C6-O6
3	C	1	NAG	C4-C5-C6-O6
3	C	5	MAN	C4-C5-C6-O6
3	C	8	MAN	C4-C5-C6-O6
3	C	8	MAN	O5-C5-C6-O6
3	C	10	MAN	O5-C5-C6-O6

There are no ring outliers.

5 monomers are involved in 6 short contacts:

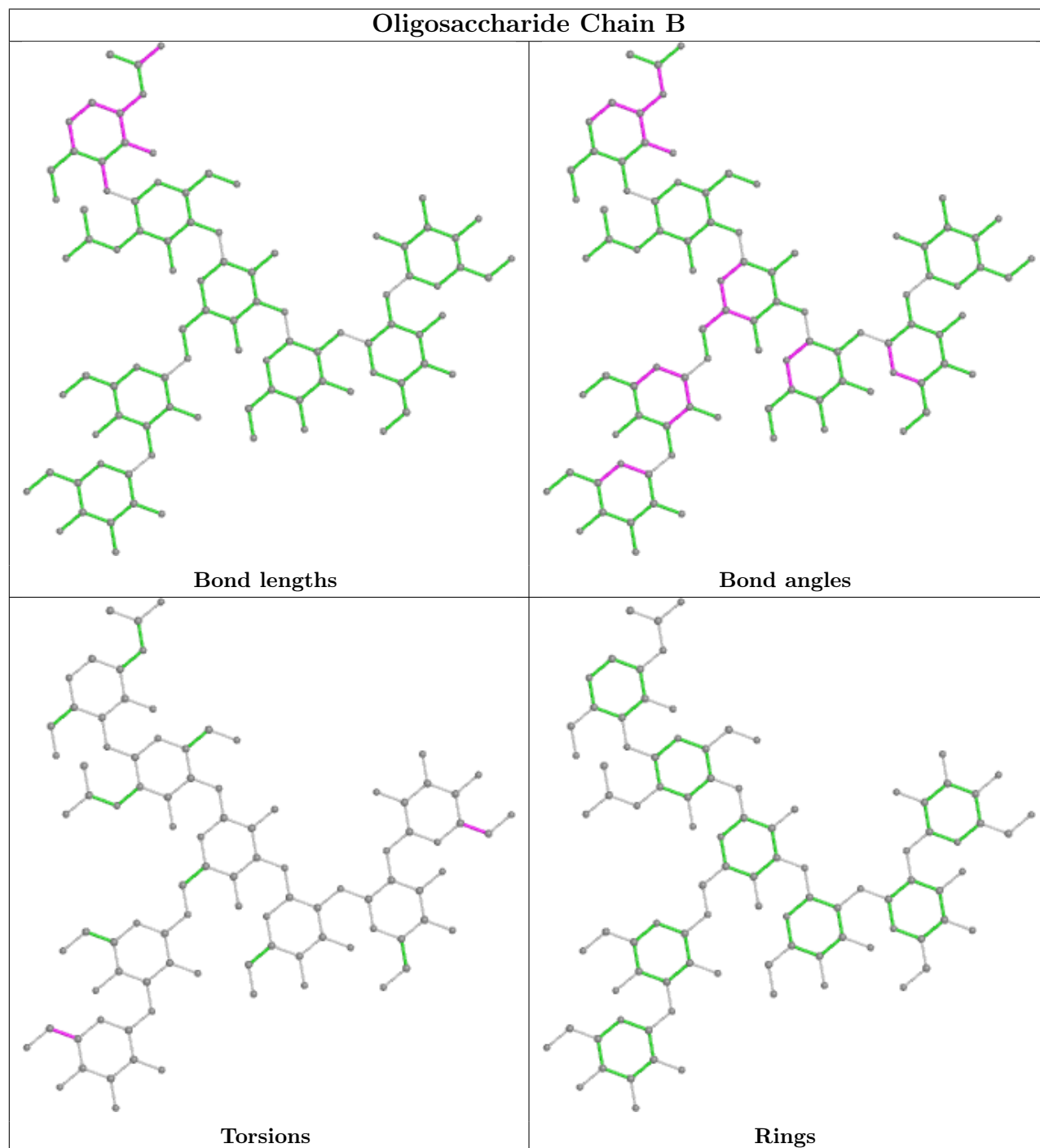
Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	C	9	MAN	1	0

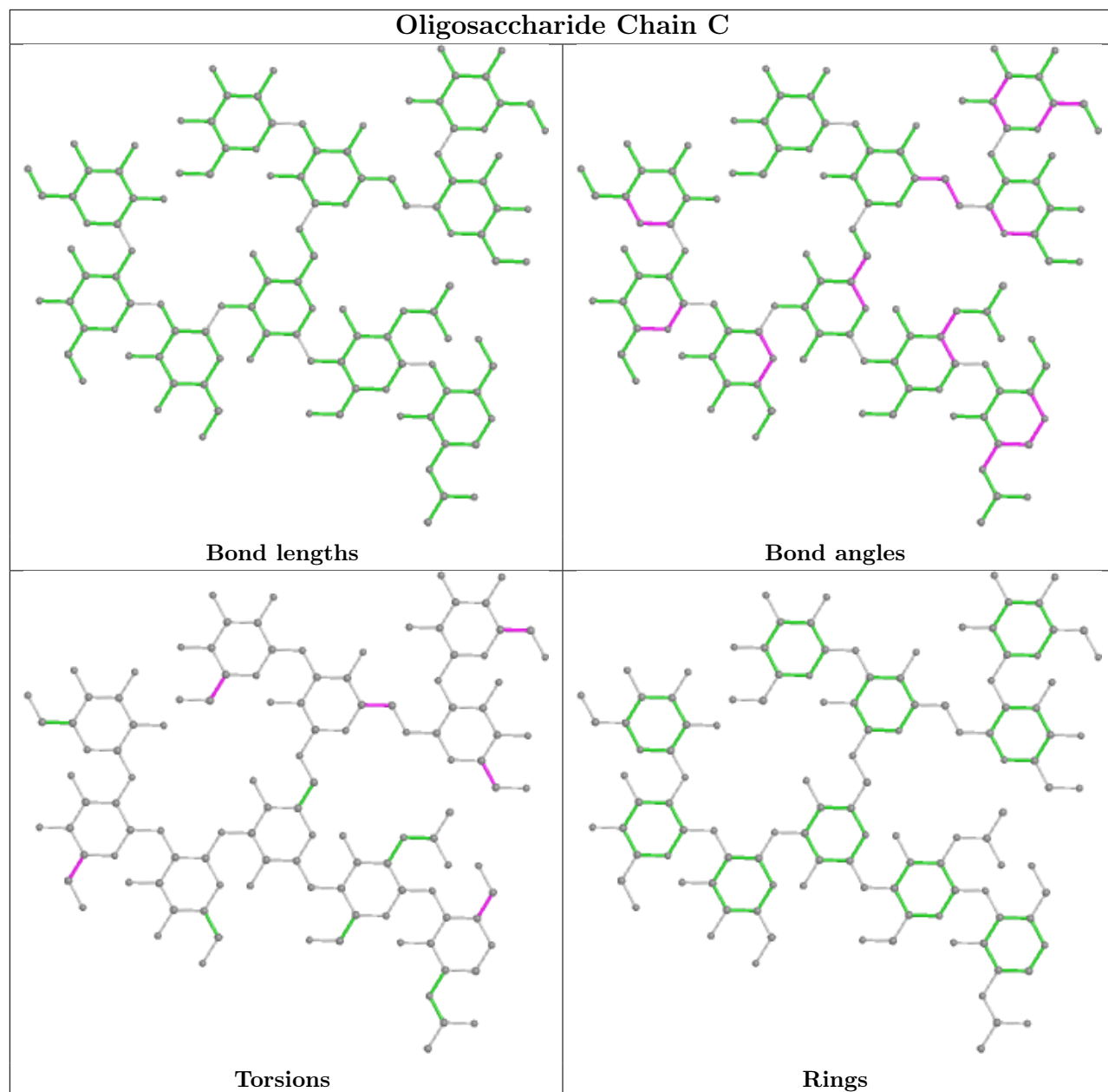
Continued on next page...

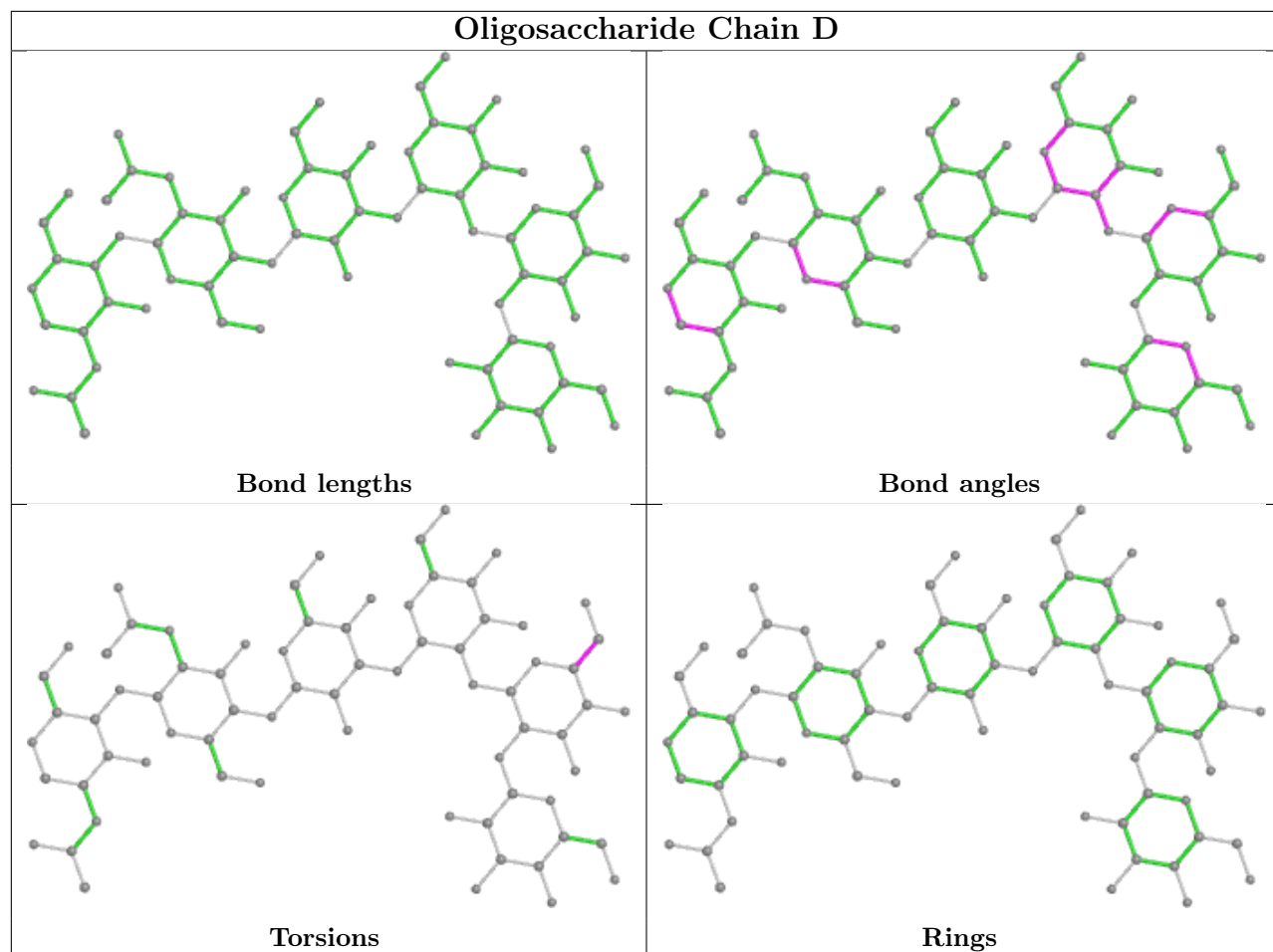
Continued from previous page...

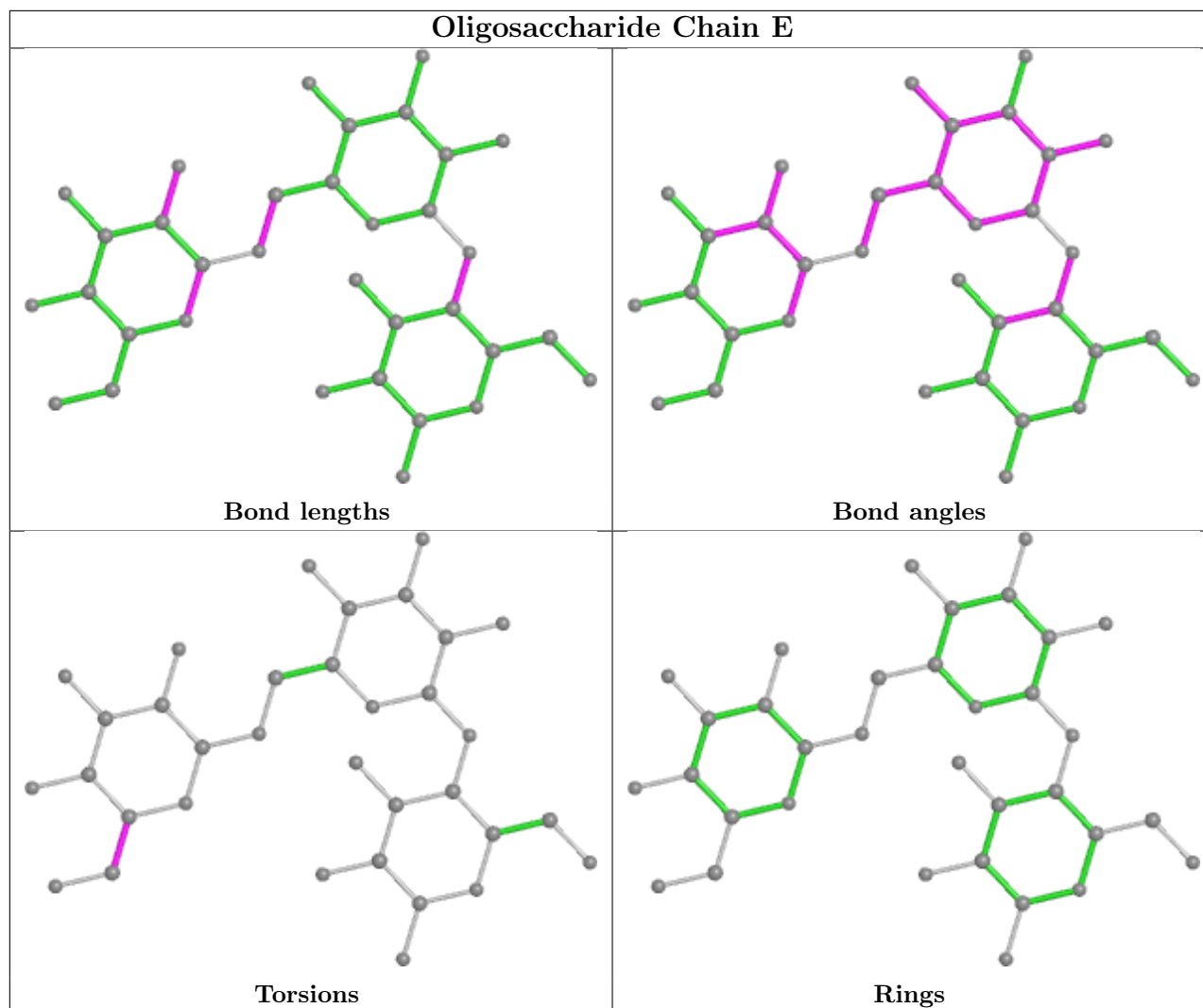
Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	1	NAG	2	0
4	D	1	NAG	1	0
5	E	2	GAL	2	0
5	E	1	GLC	2	0

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









## 5.6 Ligand geometry [i](#)

Of 9 ligands modelled in this entry, 1 is monoatomic - leaving 8 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	$\# Z  > 2$
7	DMS	A	3001	-	3,3,3	0.50	0	3,3,3	0.63	0
6	NAG	A	1777	1	14,14,15	0.36	0	17,19,21	0.90	0
6	NAG	A	1739	1	14,14,15	0.48	0	17,19,21	1.97	4 (23%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
6	NAG	A	1402	1	14,14,15	0.65	0	17,19,21	1.42	1 (5%)
6	NAG	A	1478	1	14,14,15	0.45	0	17,19,21	1.01	1 (5%)
9	1PE	A	3003	-	15,15,15	0.51	0	14,14,14	0.55	0
6	NAG	A	1760	1	14,14,15	0.34	0	17,19,21	1.27	1 (5%)
6	NAG	A	1156	1	14,14,15	0.68	0	17,19,21	1.58	4 (23%)

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
6	NAG	A	1777	1	-	0/6/23/26	0/1/1/1
6	NAG	A	1739	1	-	3/6/23/26	0/1/1/1
6	NAG	A	1402	1	-	0/6/23/26	0/1/1/1
6	NAG	A	1478	1	-	0/6/23/26	0/1/1/1
9	1PE	A	3003	-	-	9/13/13/13	-
6	NAG	A	1760	1	-	2/6/23/26	0/1/1/1
6	NAG	A	1156	1	-	1/6/23/26	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(°)
6	A	1739	NAG	C2-N2-C7	5.38	130.56	122.90
6	A	1402	NAG	C1-O5-C5	4.71	118.58	112.19
6	A	1760	NAG	C1-O5-C5	4.33	118.06	112.19
6	A	1156	NAG	C4-C3-C2	3.92	116.76	111.02
6	A	1739	NAG	C8-C7-N2	3.57	122.15	116.10
6	A	1739	NAG	C1-C2-N2	2.65	115.02	110.49
6	A	1739	NAG	O5-C1-C2	-2.47	107.39	111.29
6	A	1156	NAG	O5-C1-C2	2.45	115.16	111.29
6	A	1478	NAG	C1-O5-C5	2.30	115.31	112.19
6	A	1156	NAG	O5-C5-C4	-2.24	105.37	110.83
6	A	1156	NAG	C2-N2-C7	2.01	125.76	122.90

There are no chirality outliers.

All (15) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	A	1739	NAG	C8-C7-N2-C2
6	A	1739	NAG	O7-C7-N2-C2
6	A	1760	NAG	O5-C5-C6-O6
9	A	3003	1PE	C24-C14-OH5-C25
9	A	3003	1PE	C15-C25-OH5-C14
6	A	1760	NAG	C4-C5-C6-O6
6	A	1156	NAG	O5-C5-C6-O6
9	A	3003	1PE	OH7-C16-C26-OH6
9	A	3003	1PE	C16-C26-OH6-C15
9	A	3003	1PE	C14-C24-OH4-C13
9	A	3003	1PE	C23-C13-OH4-C24
9	A	3003	1PE	OH2-C12-C22-OH3
6	A	1739	NAG	C3-C2-N2-C7
9	A	3003	1PE	C12-C22-OH3-C23
9	A	3003	1PE	OH6-C15-C25-OH5

There are no ring outliers.

No monomer is involved in short contacts.

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

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

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	A	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	372:ARG	C	373:ASN	N	1.10



## 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	968/1013 (95%)	0.38	42 (4%) 35 45	16, 27, 43, 68	0

All (42) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	459	LEU	4.6
1	A	539	SER	4.1
1	A	209	GLU	3.9
1	A	531	SER	3.7
1	A	457	PRO	3.6
1	A	344	ILE	3.5
1	A	1004	LYS	3.4
1	A	246	GLY	3.4
1	A	533	ILE	3.3
1	A	345	PHE	3.3
1	A	499	ASP	3.2
1	A	341	LEU	3.2
1	A	534	SER	3.1
1	A	208	VAL	3.0
1	A	538	THR	2.9
1	A	916	THR	2.8
1	A	532	GLY	2.8
1	A	364	TYR	2.8
1	A	461	GLY	2.7
1	A	319	LEU	2.7
1	A	368	VAL	2.7
1	A	1008	HIS	2.7
1	A	462	THR	2.7
1	A	581	ASP	2.5
1	A	422	ASN	2.5
1	A	210	PHE	2.4
1	A	540	SER	2.4

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	343	MET	2.3
1	A	72	LEU	2.3
1	A	460	GLY	2.3
1	A	312	PHE	2.3
1	A	423	SER	2.2
1	A	455	THR	2.2
1	A	634	SER	2.2
1	A	88	LEU	2.2
1	A	342	TYR	2.2
1	A	424	THR	2.2
1	A	77	LEU	2.1
1	A	260	TYR	2.1
1	A	69	PRO	2.1
1	A	79	LEU	2.0
1	A	346	GLY	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.

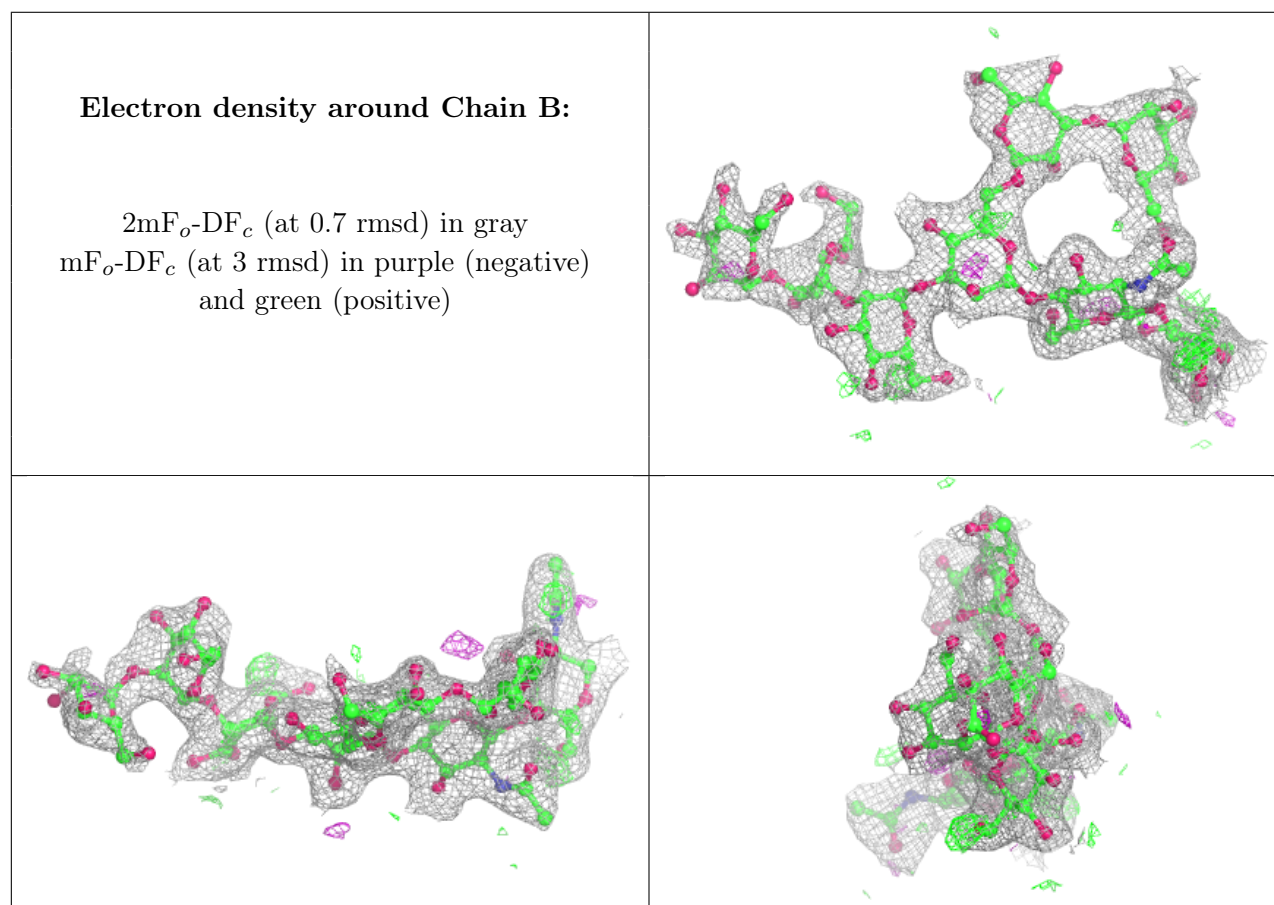
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
5	GLC	E	1	12/12	0.58	0.51	20,20,20,20	0
2	MAN	B	6	11/12	0.74	0.35	65,67,68,69	0
5	GAL	E	2	11/12	0.75	0.43	20,20,20,20	0
4	MAN	D	6	11/12	0.81	0.41	72,74,76,77	0
3	MAN	C	5	11/12	0.81	0.15	51,53,56,58	0
3	MAN	C	9	11/12	0.81	0.32	47,49,51,53	0
4	MAN	D	5	11/12	0.82	0.20	56,57,60,67	0
2	MAN	B	7	11/12	0.82	0.23	56,62,64,65	0
4	MAN	D	4	11/12	0.83	0.15	43,46,51,54	0
3	MAN	C	6	11/12	0.86	0.34	62,67,69,70	0
3	NAG	C	1	14/15	0.87	0.17	27,29,34,34	0
2	MAN	B	8	11/12	0.88	0.18	59,61,62,62	0
2	NAG	B	1	14/15	0.88	0.18	23,24,26,27	0

*Continued on next page...*

Continued from previous page...

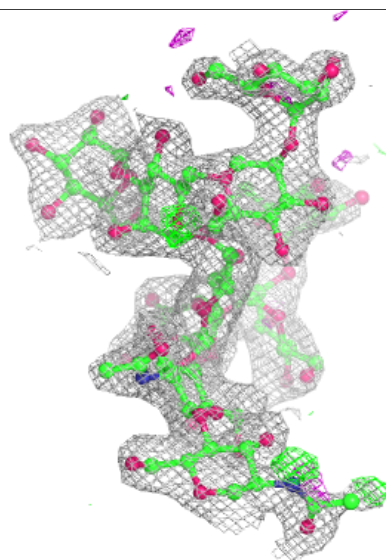
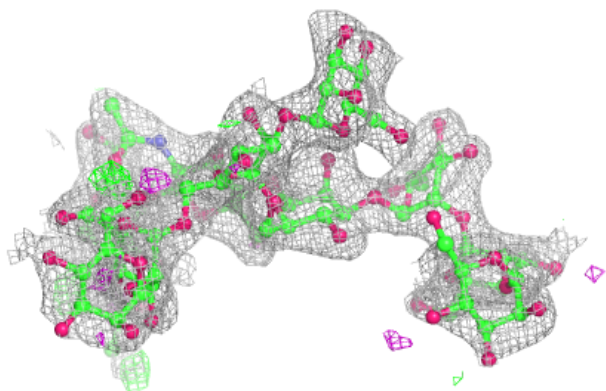
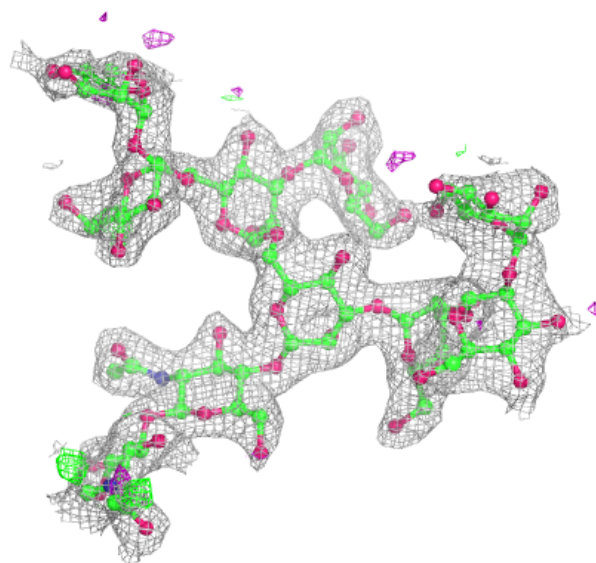
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
5	GAL	E	3	11/12	0.88	0.20	20,20,20,20	0
4	BMA	D	3	11/12	0.89	0.22	43,45,47,49	0
2	BMA	B	3	11/12	0.90	0.12	28,34,41,48	0
2	MAN	B	5	11/12	0.90	0.20	50,54,57,59	0
2	MAN	B	4	11/12	0.91	0.12	35,36,39,43	0
3	MAN	C	10	11/12	0.91	0.17	39,41,42,42	0
3	MAN	C	8	11/12	0.91	0.18	40,42,44,44	0
3	MAN	C	4	11/12	0.93	0.12	39,42,42,47	0
3	NAG	C	2	14/15	0.94	0.12	24,27,29,31	0
4	NAG	D	2	14/15	0.94	0.14	30,32,38,38	0
3	MAN	C	7	11/12	0.94	0.14	30,33,36,38	0
4	NAG	D	1	14/15	0.95	0.10	22,26,28,29	0
3	BMA	C	3	11/12	0.95	0.09	27,30,31,35	0
2	NAG	B	2	14/15	0.95	0.13	23,26,27,29	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.



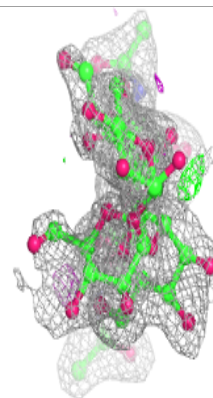
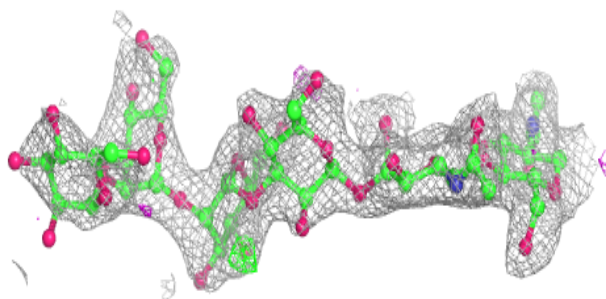
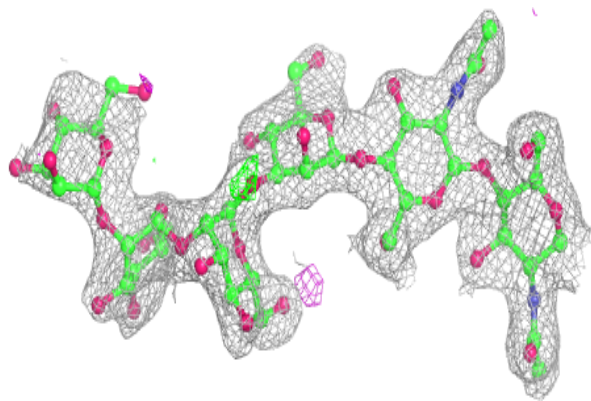
**Electron density around Chain C:**

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

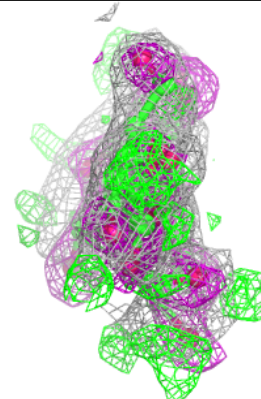
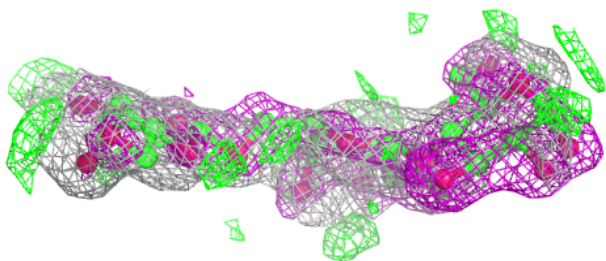
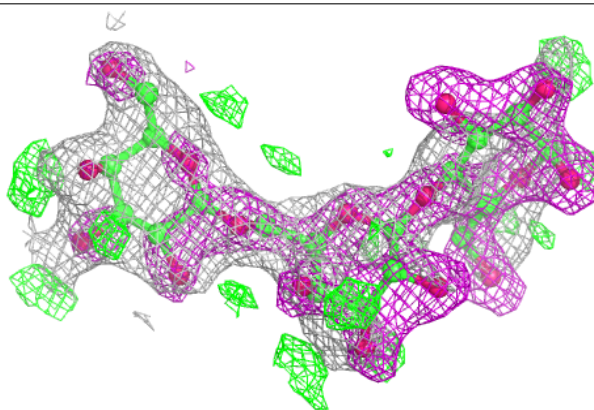


**Electron density around Chain D:**

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

**Electron density around Chain 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	NAG	A	1402	14/15	0.57	0.39	72,76,88,89	0
9	1PE	A	3003	16/16	0.60	0.34	66,75,77,77	0
6	NAG	A	1156	14/15	0.67	0.43	69,74,79,79	0
7	DMS	A	3001	4/4	0.79	0.23	77,77,78,80	0
6	NAG	A	1739	14/15	0.80	0.31	60,65,67,69	0
6	NAG	A	1478	14/15	0.84	0.30	50,55,58,58	0
6	NAG	A	1777	14/15	0.91	0.21	31,34,38,39	0
6	NAG	A	1760	14/15	0.92	0.12	33,35,40,42	0
8	CL	A	3002	1/1	0.99	0.07	26,26,26,26	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.