



wwPDB X-ray Structure Validation Summary Report ⓘ

Dec 15, 2024 – 09:59 AM EST

PDB ID : 5U73
Title : Crystal structure of human Niemann-Pick C1 protein
Authors : Li, X.; Wang, J.; Blobel, G.
Deposited on : 2016-12-11
Resolution : 3.35 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

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

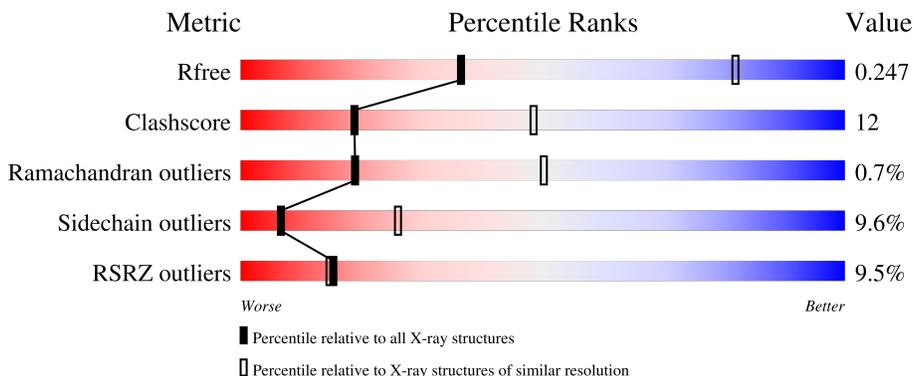
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.35 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}	164625	1325 (3.38-3.30)
Clashscore	180529	1376 (3.38-3.30)
Ramachandran outliers	177936	1376 (3.38-3.30)
Sidechain outliers	177891	1375 (3.38-3.30)
RSRZ outliers	164620	1325 (3.38-3.30)

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 ≥ 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	1278	 7% 47% 21% 5% 30%
2	B	2	 50% 50%
3	C	3	 100%
3	D	3	 100%

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 7262 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 Niemann-Pick C1 protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	898	7044	4581	1136	1281	46	0	0	0

There is a discrepancy between the modelled and reference sequences:

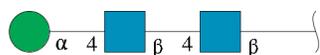
Chain	Residue	Modelled	Actual	Comment	Reference
A	642	ILE	MET	conflict	UNP O15118

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



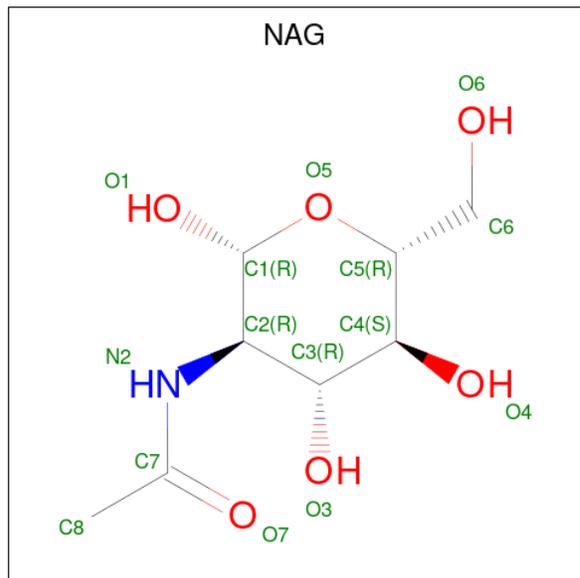
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	B	2	28	16	2	10	0	0	0

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

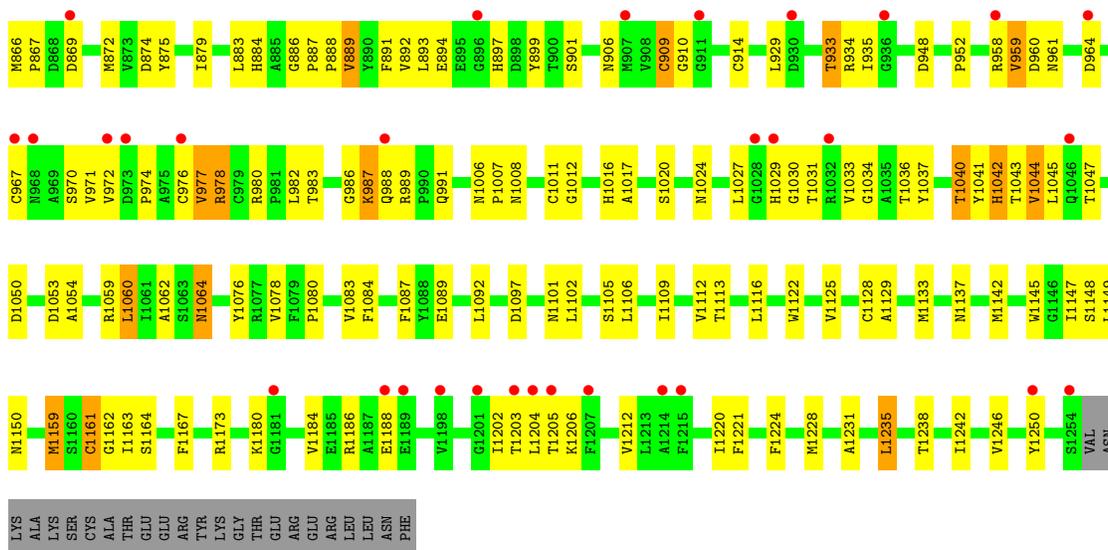


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
3	C	3	39	22	2	15	0	0	0
3	D	3	39	22	2	15	0	0	0

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



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



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

Chain B: 50%

MAG1
MAG2

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

Chain C: 100%

MAG1
MAG2
MAN3

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

Chain D: 100%

MAG1
MAG2
MAN3

4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	181.94Å 223.07Å 63.21Å 90.00° 105.31° 90.00°	Depositor
Resolution (Å)	41.15 – 3.35 41.15 – 3.35	Depositor EDS
% Data completeness (in resolution range)	51.5 (41.15-3.35) 51.4 (41.15-3.35)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.45 (at 3.32Å)	Xtrriage
Refinement program	PHENIX (1.11.1_2575: ???)	Depositor
R, R_{free}	0.233 , 0.258 0.231 , 0.247	Depositor DCC
R_{free} test set	1726 reflections (4.86%)	wwPDB-VP
Wilson B-factor (Å ²)	24.8	Xtrriage
Anisotropy	0.617	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.26 , 33.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.85	EDS
Total number of atoms	7262	wwPDB-VP
Average B, all atoms (Å ²)	51.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.40% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: NAG, MAN

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	0.26	0/7221	0.44	0/9827

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	536	THR	Peptide

5.2 Too-close contacts [i](#)

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	7044	0	6960	168	0
2	B	28	0	25	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	C	39	0	34	0	0
3	D	39	0	34	3	0
4	A	112	0	103	1	0
All	All	7262	0	7156	169	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 12.

The worst 5 of 169 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:1231:ALA:O	1:A:1235:LEU:HB2	1.74	0.86
1:A:404:ARG:NH1	1:A:584:GLU:OE2	2.11	0.82
1:A:906:ASN:OD1	1:A:989:ARG:NH1	2.16	0.78
1:A:914:CYS:SG	1:A:989:ARG:NH2	2.56	0.77
1:A:635:ILE:HD11	1:A:1204:LEU:HD13	1.66	0.76

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	892/1278 (70%)	788 (88%)	98 (11%)	6 (1%)	19 50

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	934	ARG
1	A	529	LEU
1	A	536	THR
1	A	537	PHE

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	987	LYS

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	782/1109 (70%)	707 (90%)	75 (10%)	7 26

5 of 75 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1040	THR
1	A	1163	ILE
1	A	1044	VAL
1	A	1116	LEU
1	A	656	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	447	GLN
1	A	490	ASN
1	A	578	GLN
1	A	775	GLN

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates i

8 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	0.94	1 (7%)	17,19,21	0.87	1 (5%)
2	NAG	B	2	2	14,14,15	0.34	0	17,19,21	0.60	0
3	NAG	C	1	1,3	14,14,15	0.88	1 (7%)	17,19,21	0.74	1 (5%)
3	NAG	C	2	3	14,14,15	1.25	2 (14%)	17,19,21	1.28	1 (5%)
3	MAN	C	3	3	11,11,12	1.69	3 (27%)	15,15,17	1.58	4 (26%)
3	NAG	D	1	1,3	14,14,15	0.96	2 (14%)	17,19,21	0.81	0
3	NAG	D	2	3	14,14,15	1.12	2 (14%)	17,19,21	1.48	2 (11%)
3	MAN	D	3	3	11,11,12	1.67	2 (18%)	15,15,17	1.78	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	-	2/6/23/26	0/1/1/1
2	NAG	B	2	2	-	2/6/23/26	0/1/1/1
3	NAG	C	1	1,3	-	1/6/23/26	0/1/1/1
3	NAG	C	2	3	-	2/6/23/26	0/1/1/1
3	MAN	C	3	3	-	1/2/19/22	0/1/1/1
3	NAG	D	1	1,3	-	2/6/23/26	0/1/1/1
3	NAG	D	2	3	-	1/6/23/26	0/1/1/1
3	MAN	D	3	3	-	0/2/19/22	0/1/1/1

The worst 5 of 13 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	3	MAN	C1-C2	4.02	1.61	1.52
3	D	3	MAN	O5-C5	3.70	1.50	1.43

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	2	NAG	C1-C2	3.65	1.57	1.52
3	D	3	MAN	C4-C5	3.30	1.60	1.53
3	D	2	NAG	O5-C1	3.07	1.48	1.43

The worst 5 of 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	D	2	NAG	O4-C4-C5	-5.03	96.93	109.32
3	C	2	NAG	O4-C4-C5	-4.92	97.21	109.32
3	D	3	MAN	C1-C2-C3	-3.96	103.88	109.64
3	C	3	MAN	O2-C2-C1	3.36	116.93	109.22
3	D	3	MAN	C1-O5-C5	3.19	116.46	112.19

There are no chirality outliers.

5 of 11 torsion outliers are listed below:

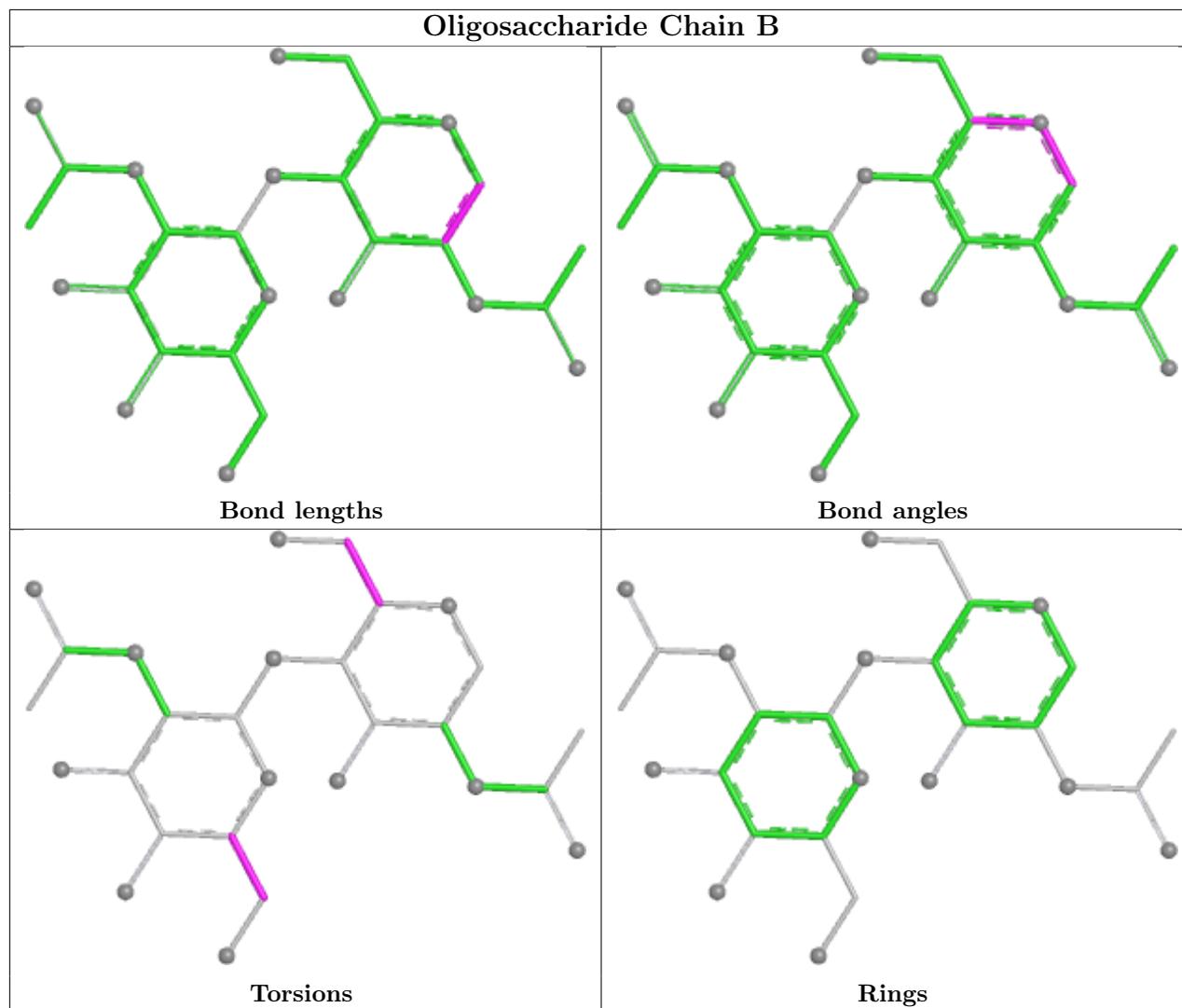
Mol	Chain	Res	Type	Atoms
3	C	2	NAG	O5-C5-C6-O6
2	B	2	NAG	C4-C5-C6-O6
2	B	1	NAG	O5-C5-C6-O6
3	D	1	NAG	O5-C5-C6-O6
2	B	2	NAG	O5-C5-C6-O6

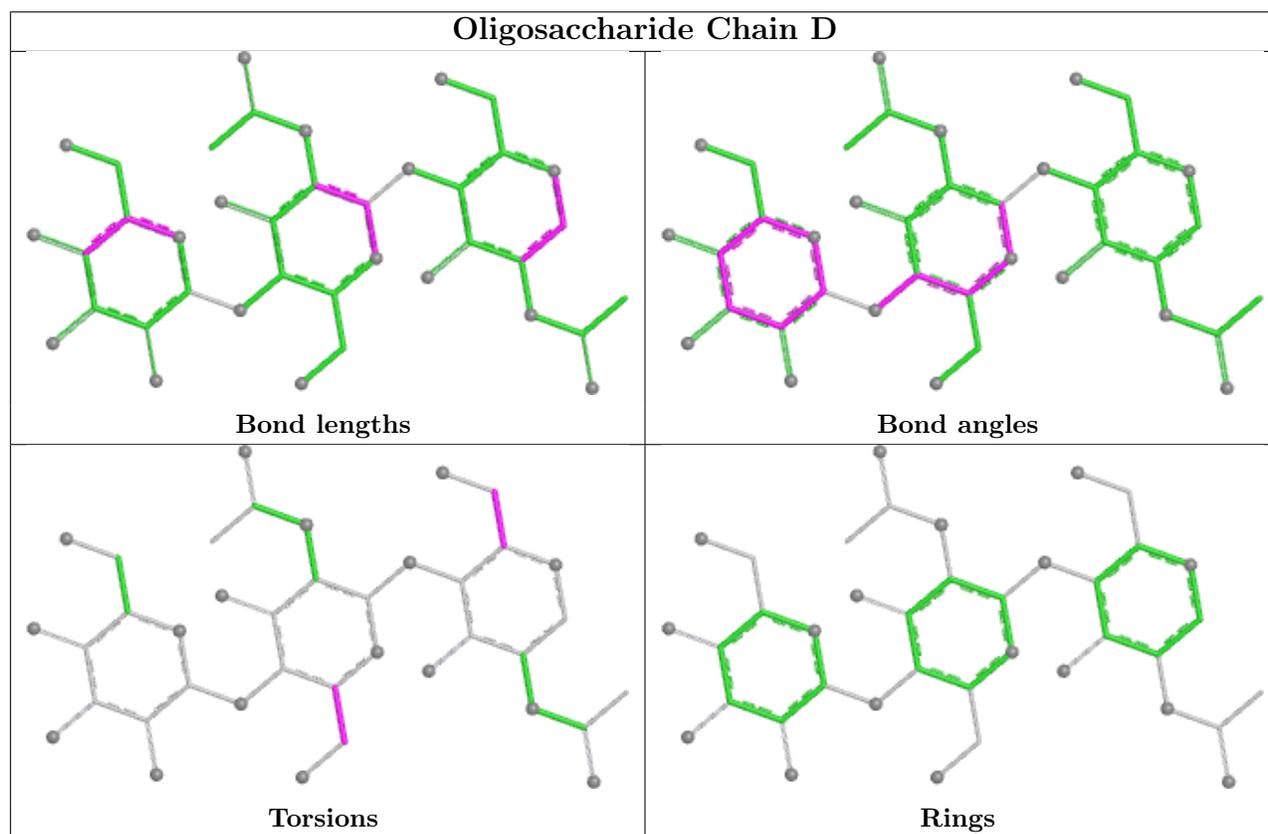
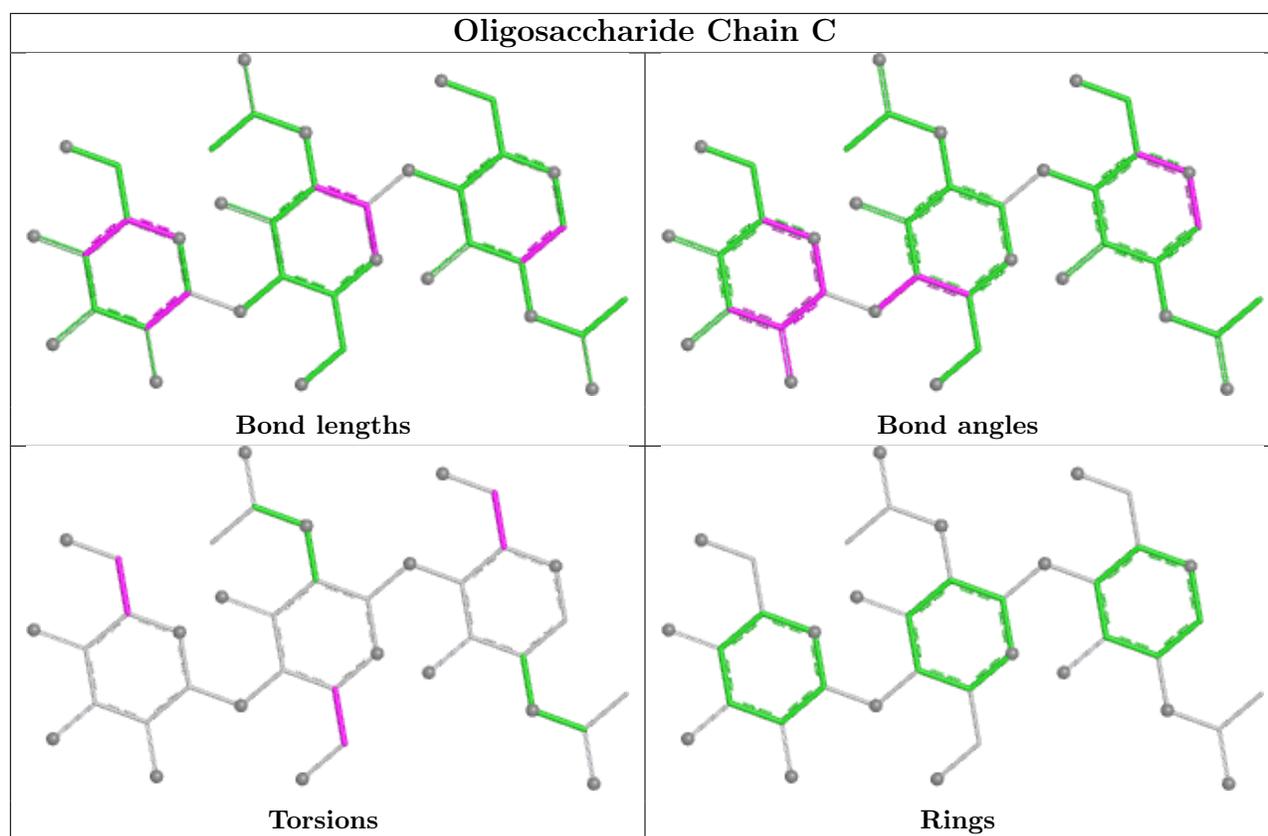
There are no ring outliers.

3 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	D	2	NAG	2	0
3	D	1	NAG	1	0
3	D	3	MAN	1	0

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





5.6 Ligand geometry

8 ligands are modelled in this entry.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
4	NAG	A	1307	1	14,14,15	0.81	1 (7%)	17,19,21	1.15	1 (5%)
4	NAG	A	1302	1	14,14,15	0.27	0	17,19,21	0.48	0
4	NAG	A	1301	1	14,14,15	0.25	0	17,19,21	0.41	0
4	NAG	A	1308	1	14,14,15	0.99	1 (7%)	17,19,21	0.88	1 (5%)
4	NAG	A	1305	1	14,14,15	0.52	0	17,19,21	0.77	1 (5%)
4	NAG	A	1306	1	14,14,15	0.40	0	17,19,21	0.41	0
4	NAG	A	1303	1	14,14,15	0.25	0	17,19,21	0.56	0
4	NAG	A	1304	1	14,14,15	0.22	0	17,19,21	0.35	0

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	NAG	A	1307	1	-	3/6/23/26	0/1/1/1
4	NAG	A	1302	1	-	2/6/23/26	0/1/1/1
4	NAG	A	1301	1	-	1/6/23/26	0/1/1/1
4	NAG	A	1308	1	-	2/6/23/26	0/1/1/1
4	NAG	A	1305	1	-	0/6/23/26	0/1/1/1
4	NAG	A	1306	1	-	2/6/23/26	0/1/1/1
4	NAG	A	1303	1	-	2/6/23/26	0/1/1/1
4	NAG	A	1304	1	-	2/6/23/26	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	1308	NAG	O5-C1	-3.29	1.38	1.43
4	A	1307	NAG	C1-C2	2.85	1.56	1.52

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	1307	NAG	C1-O5-C5	4.15	117.75	112.19
4	A	1305	NAG	C1-O5-C5	2.68	115.78	112.19
4	A	1308	NAG	C3-C4-C5	2.22	114.27	110.23

There are no chirality outliers.

5 of 14 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	1307	NAG	O5-C5-C6-O6
4	A	1303	NAG	O5-C5-C6-O6
4	A	1304	NAG	O5-C5-C6-O6
4	A	1302	NAG	O5-C5-C6-O6
4	A	1303	NAG	C4-C5-C6-O6

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	1306	NAG	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, 95th 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 ²)	Q<0.9
1	A	898/1278 (70%)	0.72	85 (9%) 15 15	8, 47, 93, 144	0

The worst 5 of 85 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	529	LEU	5.9
1	A	1250	TYR	5.9
1	A	818	PHE	5.1
1	A	571	TYR	5.1
1	A	1215	PHE	4.8

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, 95th 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.

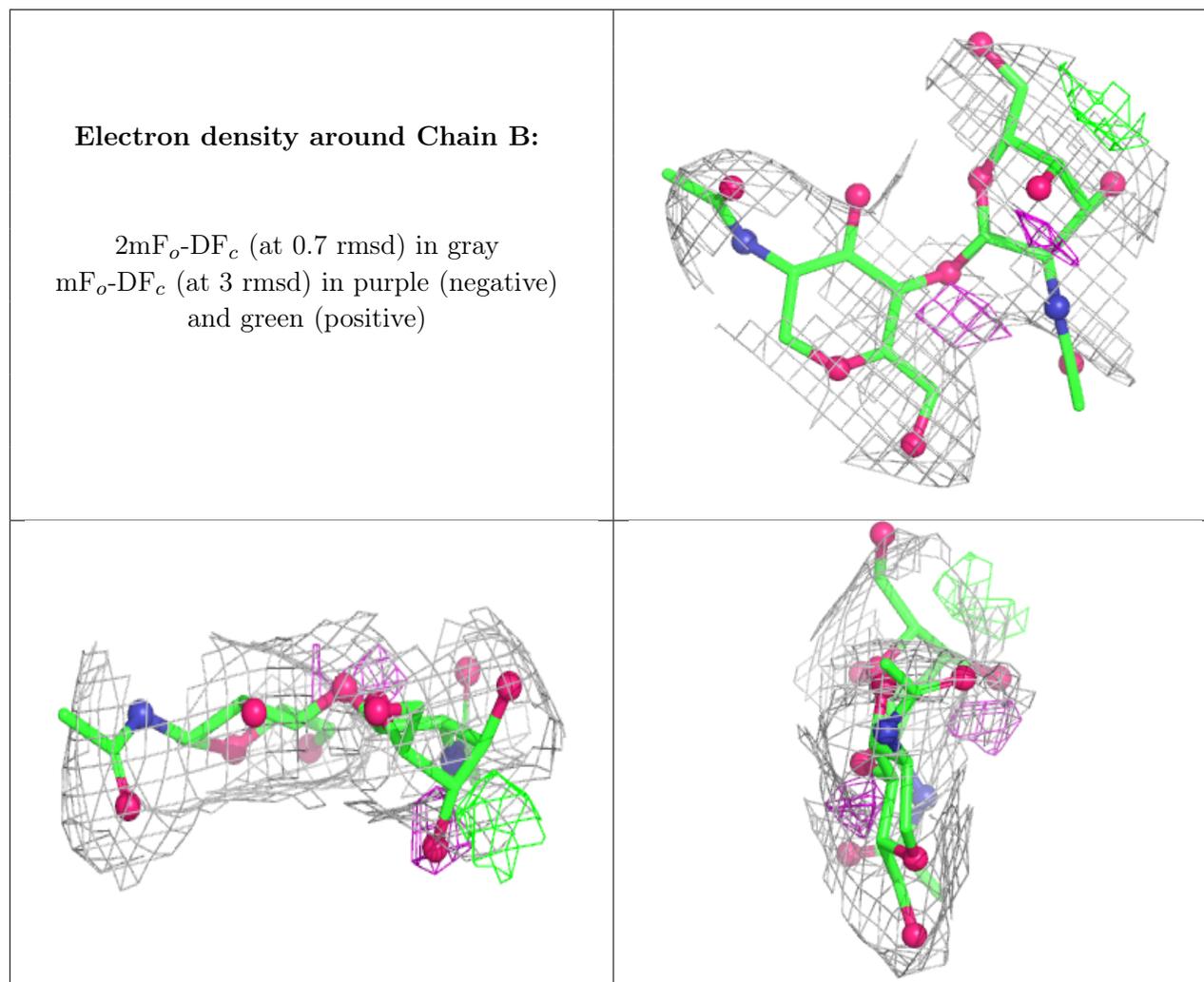
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
3	MAN	D	3	11/12	0.46	0.29	108,118,127,127	0
3	MAN	C	3	11/12	0.56	0.17	48,63,73,77	0
2	NAG	B	2	14/15	0.74	0.19	77,86,91,100	0
3	NAG	D	2	14/15	0.77	0.18	107,113,118,121	0
3	NAG	C	2	14/15	0.84	0.13	35,58,65,68	0
3	NAG	C	1	14/15	0.84	0.13	34,48,57,60	0
2	NAG	B	1	14/15	0.88	0.12	28,50,62,72	0

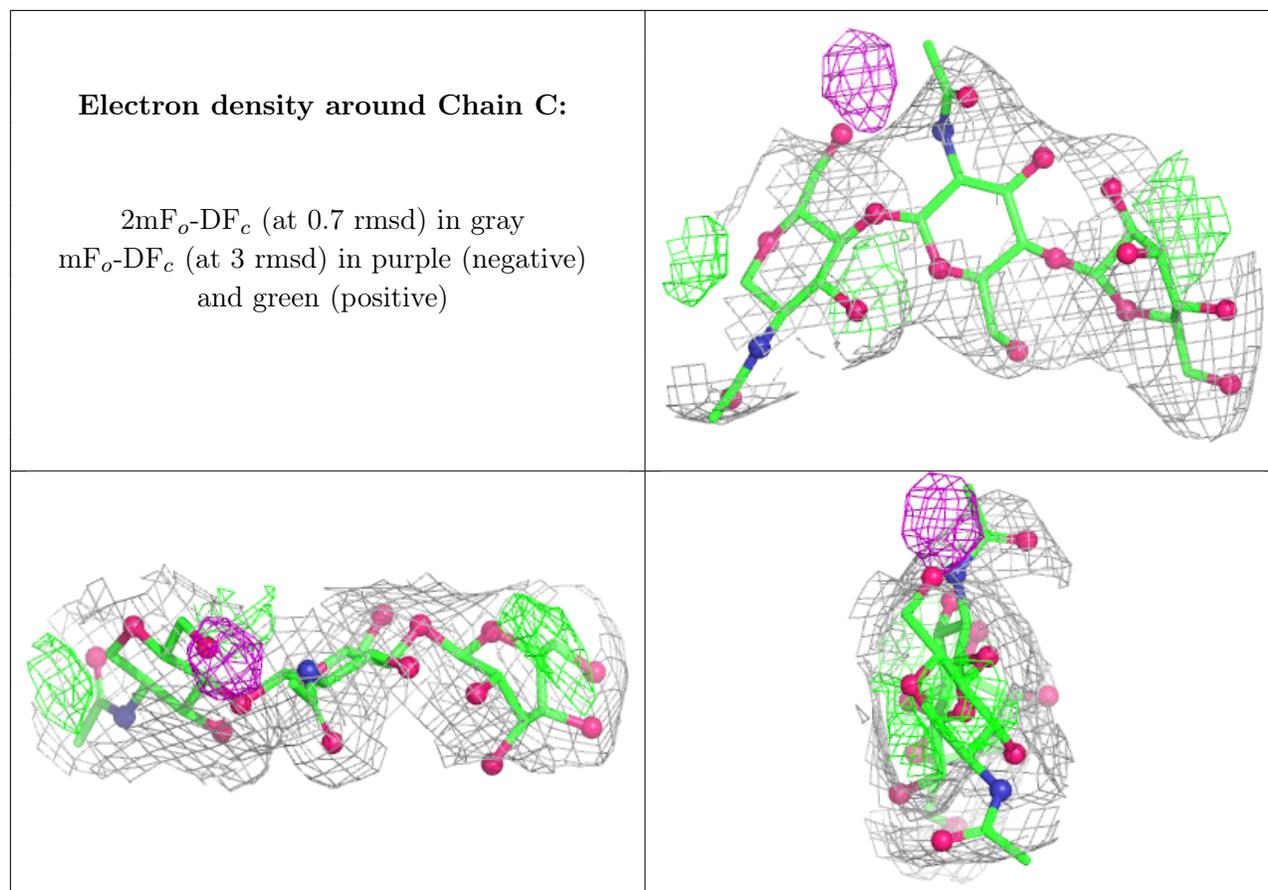
Continued on next page...

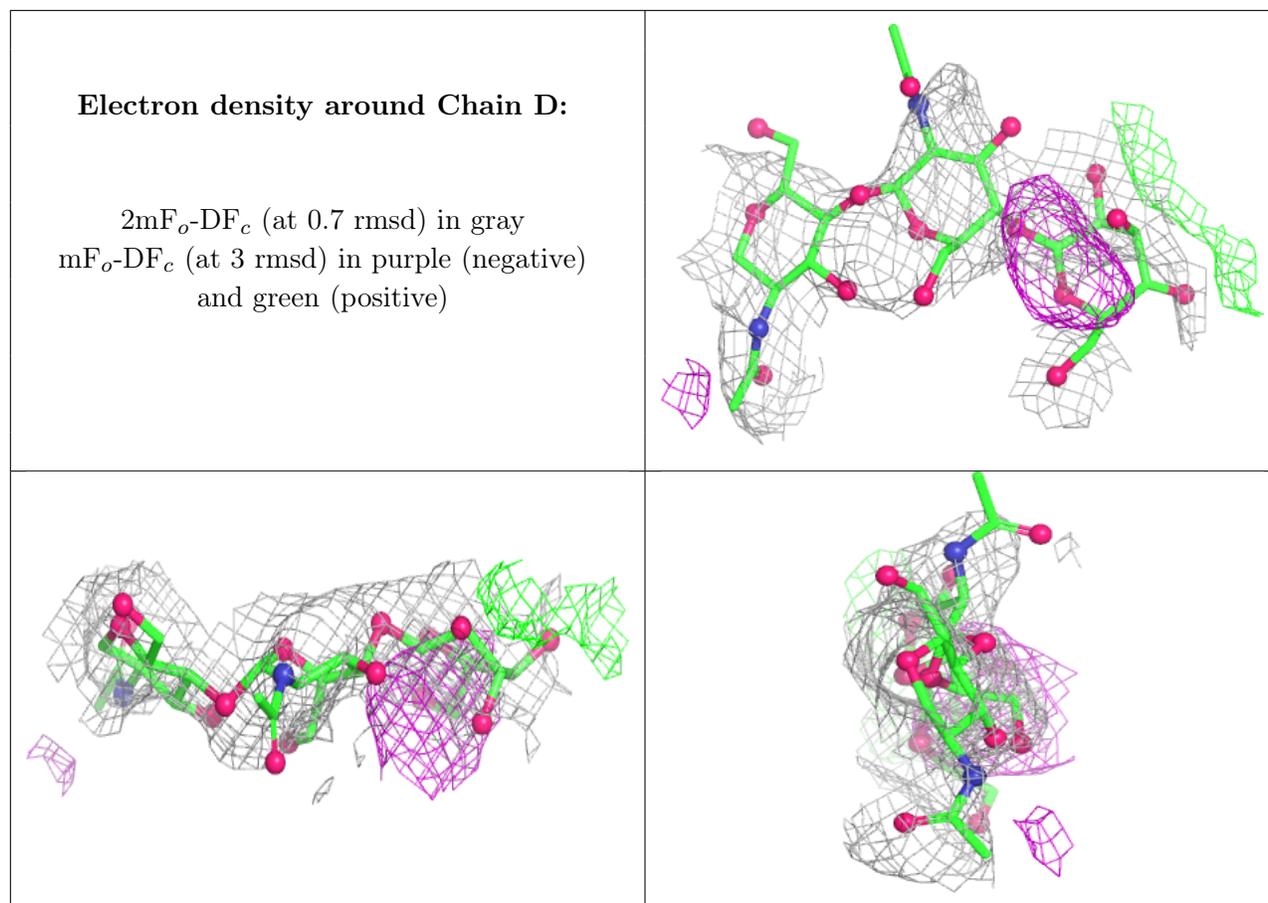
Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
3	NAG	D	1	14/15	0.91	0.10	37,78,88,99	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.







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, 95th 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 ²)	Q<0.9
4	NAG	A	1302	14/15	0.14	0.24	114,124,126,128	0
4	NAG	A	1301	14/15	0.61	0.18	72,83,89,90	0
4	NAG	A	1308	14/15	0.72	0.20	63,82,93,94	0
4	NAG	A	1306	14/15	0.79	0.15	70,76,93,96	0
4	NAG	A	1307	14/15	0.84	0.20	48,74,105,108	0
4	NAG	A	1305	14/15	0.84	0.15	65,91,94,94	0
4	NAG	A	1304	14/15	0.85	0.11	40,53,72,74	0
4	NAG	A	1303	14/15	0.88	0.11	62,70,87,91	0

6.5 Other polymers [i](#)

There are no such residues in this entry.