wwPDB X-ray Structure Validation Summary Report (i)

Dec 17, 2023 – 09:25 AM EST

PDB ID	:	1EX1
Title	:	BETA-D-GLUCAN EXOHYDROLASE FROM BARLEY
Authors	:	Varghese, J.N.; Hrmova, M.; Fincher, G.B.
Deposited on	:	1998-11-10
Resolution	:	2.20 Å(reported)

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

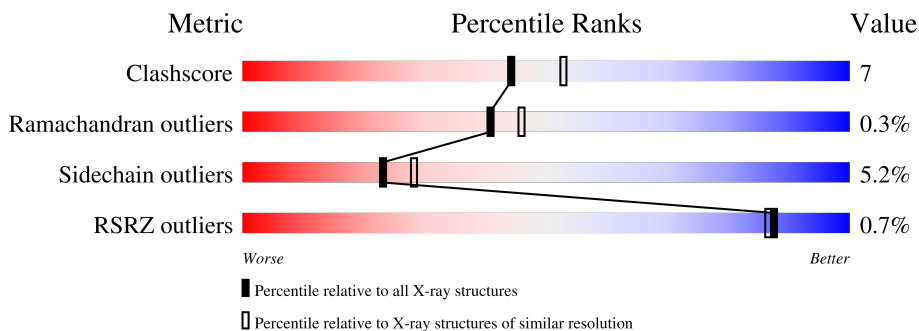|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (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:
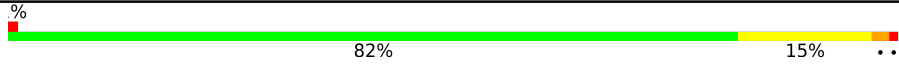*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.20 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	Percentile Ranks	Value
Clashscore		7
Ramachandran outliers		0.3%
Sidechain outliers		5.2%
RSRZ outliers		0.7%

Worse   *Better*

■ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)
RSRZ outliers	127900	4800 (2.20-2.20)

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 <=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	605	.% 82% 15% ··
2	B	6	50% 50%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	MAN	B	3	X	-	-	-
2	FCA	B	6	X	-	-	-

WORLDWIDE wwPDB PROTEIN DATA BANK

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 4885 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 PROTEIN (BETA-D-GLUCAN EXOHYDROLASE ISOEN-ZYME EXO1).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	602	Total	C	N	O	S	0	0	0
			4566	2891	787	862	26			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	320	LYS	ASN	conflict	GB AAD23382

- Molecule 2 is an oligosaccharide called 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-2)-alpha-D-mannopyranose-(1-6)-alpha-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-[alpha-D-fucopyranose-(1-3)]2-acetamido-2-deoxy-beta-D-glucopyranose.
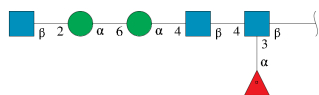
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	B	6	Total	C	N	O	0	0	0
			74	42	3	29			

- Molecule 3 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).

NAG

O6
OH

O1
HO

O5
O

C6

C1(R)   C5(R)

N2   C2(R)   C4(S)
HN   C3(R)

OH
O4

C7

OH
O3

C8   O
O7

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

- Molecule 4 is alpha-D-glucopyranose (three-letter code: GLC) (formula: $C_6H_{12}O_6$).

GLC

C6   O5   O1
HO   OH
O6
C5(R)   C1(S)

C4(S)   C2(R)
C3(S)

HO   OH
O4   O2

OH
O3
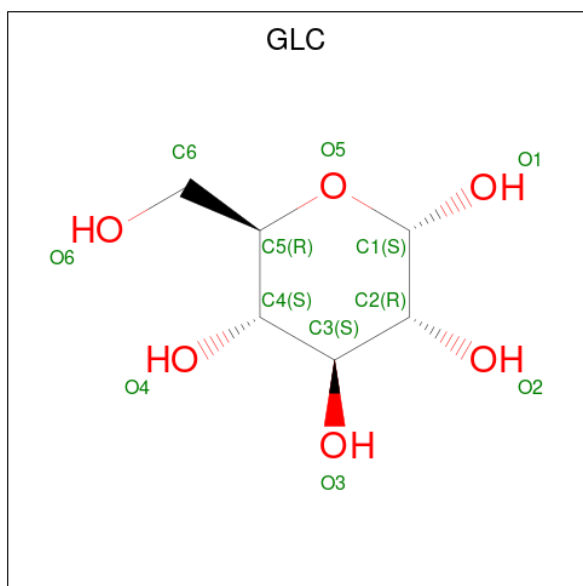
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
4	A	1	11	6	5	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	220	Total 220	O 220	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 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: PROTEIN (BETA-D-GLUCAN EXOHYDROLASE ISOENZYME EXO1)
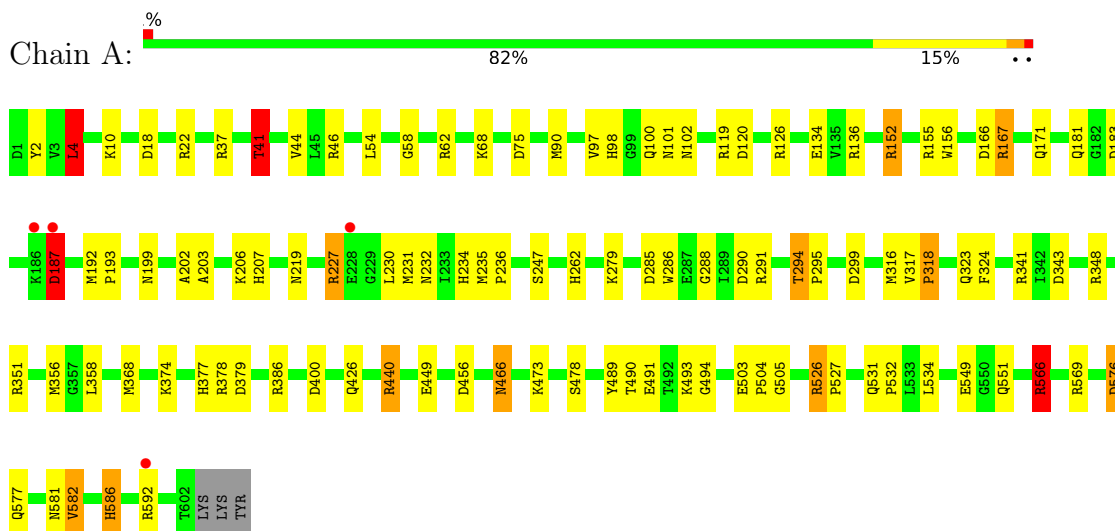
• Molecule 2: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-2)-alpha-D-mannopyranose-(1-6)-alpha-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-[alpha-D-fucopyranose-(1-3)]2-acetamido-2-deoxy-beta-D-glucopyranose

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, α, β, γ	102.09Å 102.09Å 184.50Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 2.20 12.51 – 2.20	Depositor EDS
% Data completeness (in resolution range)	83.3 (20.00-2.20) 83.4 (12.51-2.20)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.44 (at 2.19Å)	Xtriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.170 , 0.202 0.146 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å^2)	24.8	Xtriage
Anisotropy	0.105	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.38 , 60.1	EDS
L-test for twinning[2]	$< |L| > = 0.50, < L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	4885	wwPDB-VP
Average B, all atoms (Å^2)	26.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.41% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $< |L| >$, $< L^2 >$ 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: GLC, FCA, 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	0.69	0/4663	1.47	53/6334 (0.8%)

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.

The worst 5 of 53 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	166	ASP	CB-CG-OD1	13.13	130.11	118.30
1	A	152	ARG	NE-CZ-NH2	-12.92	113.84	120.30
1	A	62	ARG	NE-CZ-NH2	-12.44	114.08	120.30
1	A	386	ARG	NE-CZ-NH2	-12.33	114.13	120.30
1	A	566	ARG	NE-CZ-NH2	-11.42	114.59	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	187	ASP	Peptide

5.2 Too-close contacts ⓘ

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	4566	0	4551	60	0
2	B	74	0	64	1	0
3	A	14	0	13	0	0
4	A	11	0	10	2	0
5	A	220	0	0	5	0
All	All	4885	0	4638	61	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 7.

The worst 5 of 61 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:41:THR:HG22	1:A:44:VAL:H	1.35	0.88
1:A:181:GLN:HE21	1:A:203:ALA:H	1.20	0.88
1:A:286:TRP:HB2	1:A:316:MET:HE1	1.56	0.87
1:A:316:MET:SD	5:A:869:HOH:O	2.34	0.84
1:A:286:TRP:HE3	1:A:316:MET:CE	1.90	0.83

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	600/605 (99%)	574 (96%)	24 (4%)	2 (0%)	41	46

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	187	ASP
1	A	505	GLY

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	485/488 (99%)	460 (95%)	25 (5%)	23	28

5 of 25 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	358	LEU
1	A	473	LYS
1	A	586	HIS
1	A	466	ASN
1	A	478	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	262	HIS
1	A	377	HIS
1	A	586	HIS
1	A	581	ASN
1	A	199	ASN

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 ⓘ

6 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	1.31	1 (7%)	17,19,21	2.20	8 (47%)
2	NAG	B	2	2	14,14,15	1.19	1 (7%)	17,19,21	1.33	2 (11%)
2	MAN	B	3	2	11,11,12	0.86	0	15,15,17	1.41	3 (20%)
2	MAN	B	4	2	11,11,12	1.21	1 (9%)	15,15,17	1.76	5 (33%)
2	NAG	B	5	2	14,14,15	1.30	1 (7%)	17,19,21	1.80	4 (23%)
2	FCA	B	6	2	10,10,11	1.13	1 (10%)	14,14,16	1.73	4 (28%)

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	-	2/6/23/26	0/1/1/1
2	NAG	B	2	2	-	0/6/23/26	0/1/1/1
2	MAN	B	3	2	1/1/4/5	2/2/19/22	0/1/1/1
2	MAN	B	4	2	-	0/2/19/22	0/1/1/1
2	NAG	B	5	2	-	1/6/23/26	0/1/1/1
2	FCA	B	6	2	5/5/4/5	-	0/1/1/1

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	5	NAG	O7-C7	-3.96	1.14	1.23

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	1	NAG	O7-C7	-3.73	1.14	1.23
2	B	2	NAG	O7-C7	-3.33	1.15	1.23
2	B	4	MAN	C1-C2	2.57	1.58	1.52
2	B	6	FCA	O5-C1	-2.13	1.40	1.43

The worst 5 of 26 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	1	NAG	C1-O5-C5	4.19	117.87	112.19
2	B	5	NAG	C1-O5-C5	4.00	117.61	112.19
2	B	1	NAG	O3-C3-C2	3.99	117.72	109.47
2	B	1	NAG	O5-C5-C6	-3.56	101.63	107.20
2	B	6	FCA	C1-O5-C5	3.38	120.43	112.78

5 of 6 chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	B	3	MAN	C1
2	B	6	FCA	C2
2	B	6	FCA	C3
2	B	6	FCA	C4
2	B	6	FCA	C1

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	1	NAG	O5-C5-C6-O6
2	B	1	NAG	C4-C5-C6-O6
2	B	3	MAN	O5-C5-C6-O6
2	B	3	MAN	C4-C5-C6-O6
2	B	5	NAG	C8-C7-N2-C2

There are no ring outliers.

2 monomers are involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	2	NAG	1	0
2	B	4	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.

Oligosaccharide Chain B

Bond lengths

Bond angles

Torsions

Rings
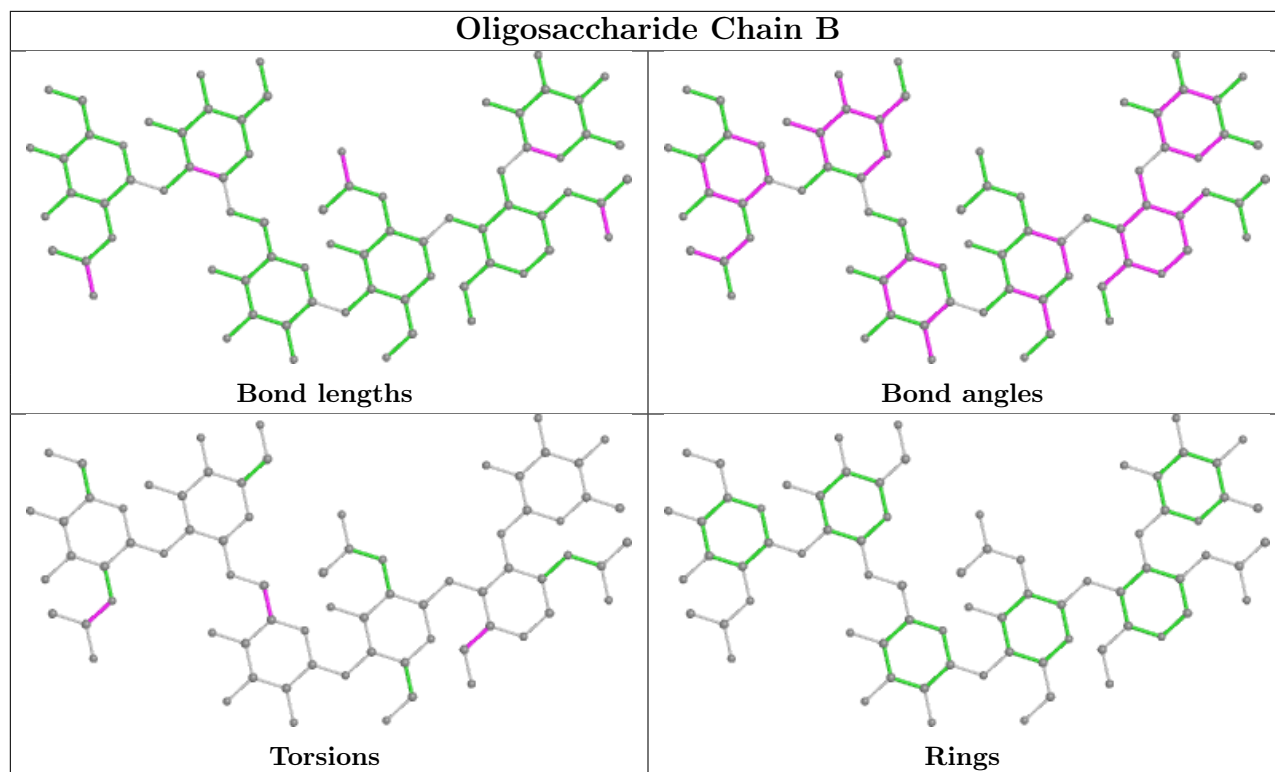
5.6   Ligand geometry ⓘ

2 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	GLC	A	617	-	11,11,12	1.86	4 (36%)	15,15,17	2.90	8 (53%)
3	NAG	A	610	1	14,14,15	1.24	1 (7%)	17,19,21	1.70	3 (17%)

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	GLC	A	617	-	-	2/2/19/22	0/1/1/1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	NAG	A	610	1	-	0/6/23/26	0/1/1/1

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	610	NAG	O7-C7	-3.63	1.15	1.23
4	A	617	GLC	O6-C6	3.33	1.56	1.42
4	A	617	GLC	O4-C4	3.10	1.50	1.43
4	A	617	GLC	O5-C1	3.06	1.48	1.43
4	A	617	GLC	C1-C2	2.05	1.56	1.52

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	617	GLC	C1-O5-C5	5.97	120.28	112.19
4	A	617	GLC	C1-C2-C3	4.48	115.17	109.67
3	A	610	NAG	C4-C3-C2	-4.07	105.06	111.02
4	A	617	GLC	O3-C3-C4	-3.93	101.27	110.35
3	A	610	NAG	C1-O5-C5	3.28	116.64	112.19

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	617	GLC	O5-C5-C6-O6
4	A	617	GLC	C4-C5-C6-O6

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	617	GLC	2	0

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)

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	602/605 (99%)	-0.97	4 (0%)	87	86	13, 23, 39, 63	0

All (4) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	187	ASP	3.9
1	A	186	LYS	3.2
1	A	228	GLU	3.0
1	A	592	ARG	2.4

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

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

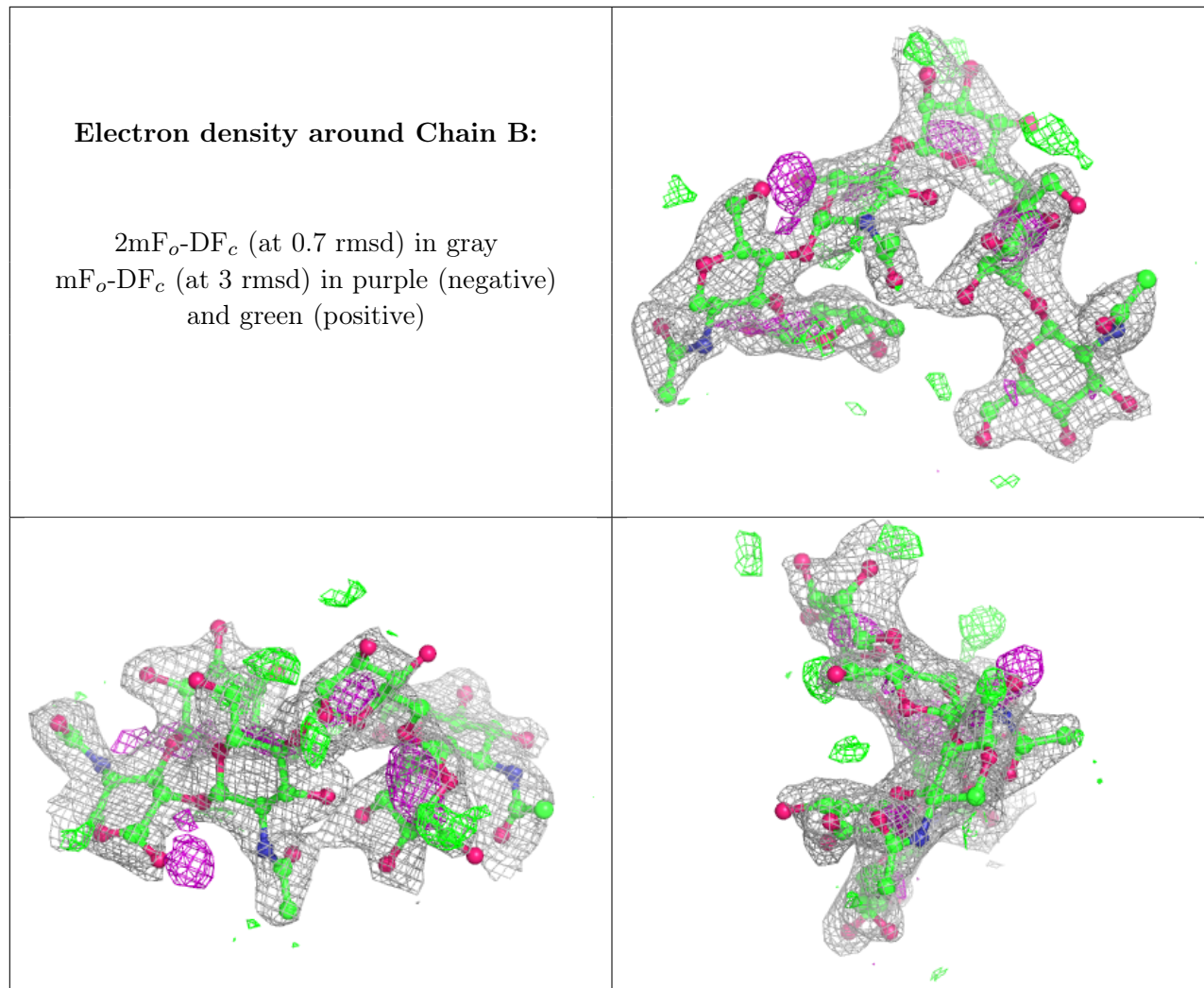## 6.3 Carbohydrates (i)

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
2	MAN	B	3	11/12	0.75	0.36	67,69,71,71	0
2	MAN	B	4	11/12	0.75	0.27	63,66,67,67	0
2	NAG	B	2	14/15	0.81	0.23	58,62,67,67	0
2	NAG	B	5	14/15	0.85	0.24	59,63,64,65	0
2	FCA	B	6	10/11	0.88	0.31	55,55,57,60	0
2	NAG	B	1	14/15	0.93	0.24	51,54,55,59	0

The following is a graphical depiction of the model fit to experimental electron density for oligosac-

charide. 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)

6.4 Ligands ⓘ

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	NAG	A	610	14/15	0.88	0.23	47,49,53,55	0
4	GLC	A	617	11/12	0.91	0.14	2,11,18,19	0

6.5 Other polymers ⓘ

There are no such residues in this entry.