W O R L D W I D E
wwPDB
PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 16, 2024 – 10:05 AM EDT

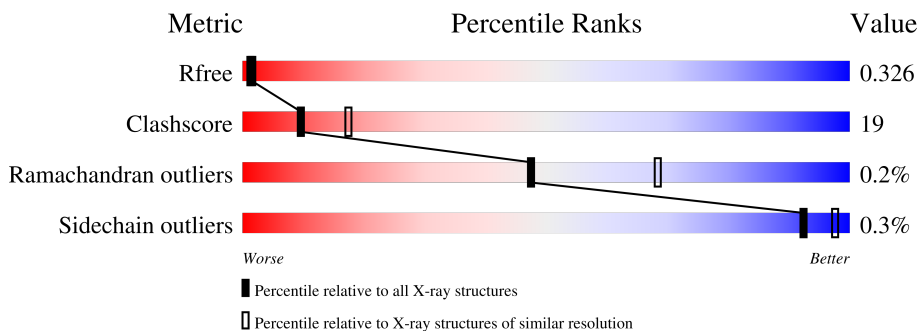|		
PDB ID	:	2GMP
Title	:	Metal-free (apo) P. angolensis seed lectin in complex with GlcNAC-beta(1-2)Man
Authors	:	Garcia-Pino, A.; Buts, L.; Wyns, L.; Loris, R.
Deposited on	:	2006-04-07
Resolution	:	2.50 Å(reported)

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 2.50 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
Rfree		0.326
Clashscore		19
Ramachandran outliers		0.2%
Sidechain outliers		0.3%

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))
R_{free}	130704	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)

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%

Mol	Chain	Length	Quality of chain
1	A	252	59% 31% 10%
1	B	252	62% 30% 8%
2	C	2	100%
2	D	2	100%

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:

W O R L D W I D E
wwPDB
PROTEIN DATA BANK

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	NAG	D	2	-	-	X	-

2 Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 3700 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 lectin.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
1	A	228	Total	C	N	O	0	0	0
			1726	1099	289	338			
1	B	233	Total	C	N	O	0	0	1
			1744	1111	293	340			

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

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	C	2	Total	C	N	O	0	0	0
			26	14	1	11			
2	D	2	Total	C	N	O	0	0	0
			26	14	1	11			

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O_4S).

SO4

O1

O

O3 O⁻ — S — O⁻ O4

S

O

O2

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

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	90	Total 90	O 90	0	0
4	B	83	Total 83	O 83	0	0

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.
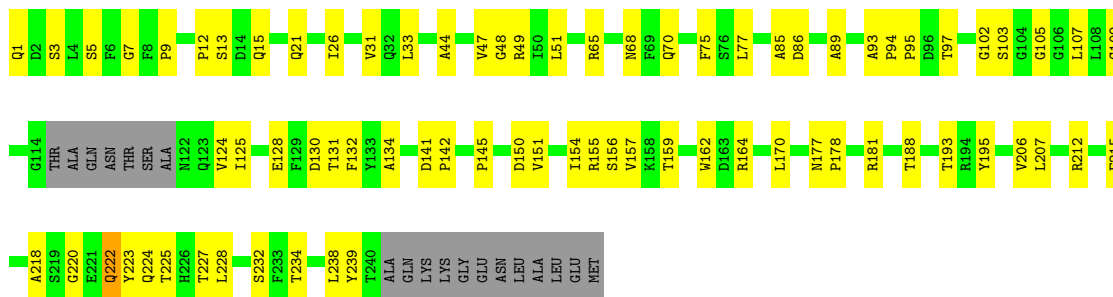
- Molecule 1: lectin

Chain A:
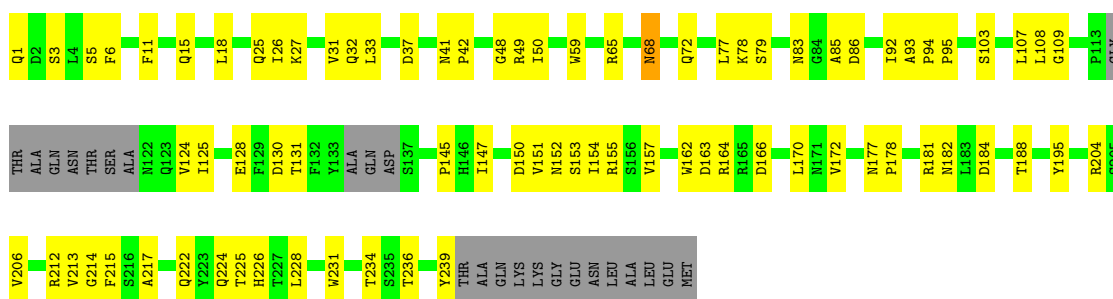
- Molecule 1: lectin

Chain B:

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

Chain C:

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

Chain D:

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	58.11Å 61.00Å 128.52Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 2.50 16.13 – 2.48	Depositor EDS
% Data completeness (in resolution range)	(Not available) (20.00-2.50) 98.5 (16.13-2.48)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.94 (at 2.49Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.198 , 0.255 0.296 , 0.326	Depositor DCC
R_{free} test set	1297 reflections (7.87%)	wwPDB-VP
Wilson B-factor (Å^2)	29.9	Xtriage
Anisotropy	0.435	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.29 , 31.2	EDS
L-test for twinning[2]	$< |L| > = 0.48, < L^2 > = 0.31$	Xtriage
Estimated twinning fraction	0.032 for k,h,-l	Xtriage
F_o,F_c correlation	0.85	EDS
Total number of atoms	3700	wwPDB-VP
Average B, all atoms (Å^2)	28.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.17% 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: PCA, NAG, MAN, SO4

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	0.38	0/1760	0.69	1/2399 (0.0%)
1	B	0.38	0/1780	0.70	0/2428
All	All	0.38	0/3540	0.69	1/4827 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	68	ASN	N-CA-C	-5.03	97.41	111.00

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts (i)

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1726	0	1625	60	0
1	B	1744	0	1628	66	0
2	C	26	0	23	7	0
2	D	26	0	25	9	0
3	A	5	0	0	0	0
4	A	90	0	0	4	0
4	B	83	0	0	0	0
All	All	3700	0	3301	131	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 19.

The worst 5 of 131 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:B:102:GLY:O	2:D:2:NAG:H61	1.53	1.08
1:B:164:ARG:HE	1:B:224:GLN:HE22	1.08	0.95
1:A:164:ARG:HE	1:A:224:GLN:HE22	1.01	0.93
1:A:86:ASP:OD1	2:C:1:MAN:H61	1.70	0.91
1:A:164:ARG:HE	1:A:224:GLN:NE2	1.69	0.90

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	222/252 (88%)	211 (95%)	10 (4%)	1 (0%)	29	48
1	B	229/252 (91%)	215 (94%)	14 (6%)	0	100	100
All	All	451/504 (90%)	426 (94%)	24 (5%)	1 (0%)	47	68

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	108	LEU

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	184/215 (86%)	184 (100%)	0	100	100
1	B	182/215 (85%)	181 (100%)	1 (0%)	88	96
All	All	366/430 (85%)	365 (100%)	1 (0%)	92	97

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	222	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	224	GLN
1	B	182	ASN
1	B	25	GLN
1	B	21	GLN
1	B	41	ASN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

2 non-standard protein/DNA/RNA residues 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	#$|Z| > 2$
1	PCA	A	1	1	7,8,9	2.82	3 (42%)	9,10,12	2.00	3 (33%)
1	PCA	B	1	1	7,8,9	2.79	2 (28%)	9,10,12	1.84	2 (22%)

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	PCA	A	1	1	-	0/0/11/13	0/1/1/1
1	PCA	B	1	1	-	0/0/11/13	0/1/1/1

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	1	PCA	CD-N	6.50	1.50	1.34
1	A	1	PCA	CD-N	6.41	1.50	1.34
1	A	1	PCA	CA-N	3.08	1.50	1.46
1	B	1	PCA	CA-N	2.96	1.49	1.46
1	A	1	PCA	CB-CG	-2.01	1.48	1.53

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	1	PCA	OE-CD-CG	-3.40	120.65	126.72
1	B	1	PCA	OE-CD-CG	-3.13	121.13	126.72
1	A	1	PCA	OE-CD-N	2.87	131.18	124.96
1	B	1	PCA	OE-CD-N	2.61	130.59	124.96
1	A	1	PCA	CB-CA-C	-2.23	109.59	112.66

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 ⓘ

4 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	MAN	C	1	2	12,12,12	1.93	2 (16%)	17,17,17	1.96	2 (11%)
2	NAG	C	2	2	14,14,15	0.77	1 (7%)	17,19,21	0.86	1 (5%)
2	MAN	D	1	2	12,12,12	0.44	0	17,17,17	1.99	2 (11%)
2	NAG	D	2	2	14,14,15	0.75	1 (7%)	17,19,21	0.85	1 (5%)

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	MAN	C	1	2	-	2/2/22/22	0/1/1/1
2	NAG	C	2	2	-	1/6/23/26	0/1/1/1
2	MAN	D	1	2	-	0/2/22/22	0/1/1/1
2	NAG	D	2	2	-	1/6/23/26	0/1/1/1

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	1	MAN	O6-C6	-6.19	1.16	1.42
2	C	2	NAG	C1-C2	2.39	1.55	1.52
2	D	2	NAG	C1-C2	2.31	1.55	1.52
2	C	1	MAN	C6-C5	-2.03	1.45	1.51

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	D	1	MAN	O2-C2-C1	7.51	126.59	109.25
2	C	1	MAN	C6-C5-C4	7.32	130.98	113.02
2	C	2	NAG	C1-O5-C5	2.42	115.44	112.19
2	D	2	NAG	C1-O5-C5	2.37	115.37	112.19
2	D	1	MAN	C3-C4-C5	2.08	114.00	110.23

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	C	1	MAN	O5-C5-C6-O6
2	C	1	MAN	C4-C5-C6-O6
2	C	2	NAG	O5-C5-C6-O6

Continued on next page...

W O R L D W I D E
wwPDB
PROTEIN DATA BANK

Continued from previous page...

| Mol	Chain	Res	Type	Atoms
2	D	2	NAG	O5-C5-C6-O6

There are no ring outliers.

4 monomers are involved in 16 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	D	2	NAG	8	0
2	C	1	MAN	5	0
2	C	2	NAG	6	0
2	D	1	MAN	4	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 C

Bond lengths

Bond angles

Torsions

Rings

Oligosaccharide Chain D
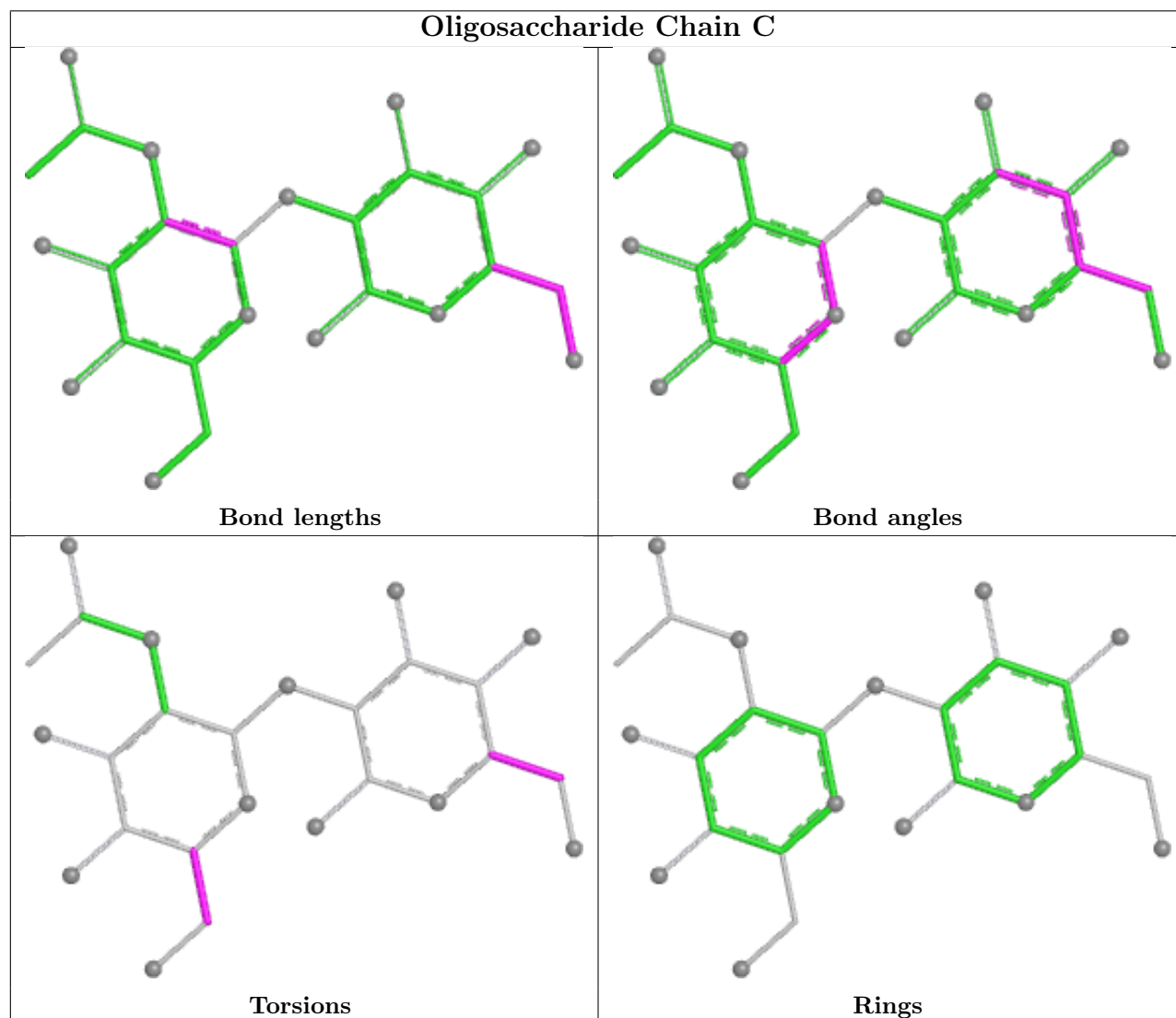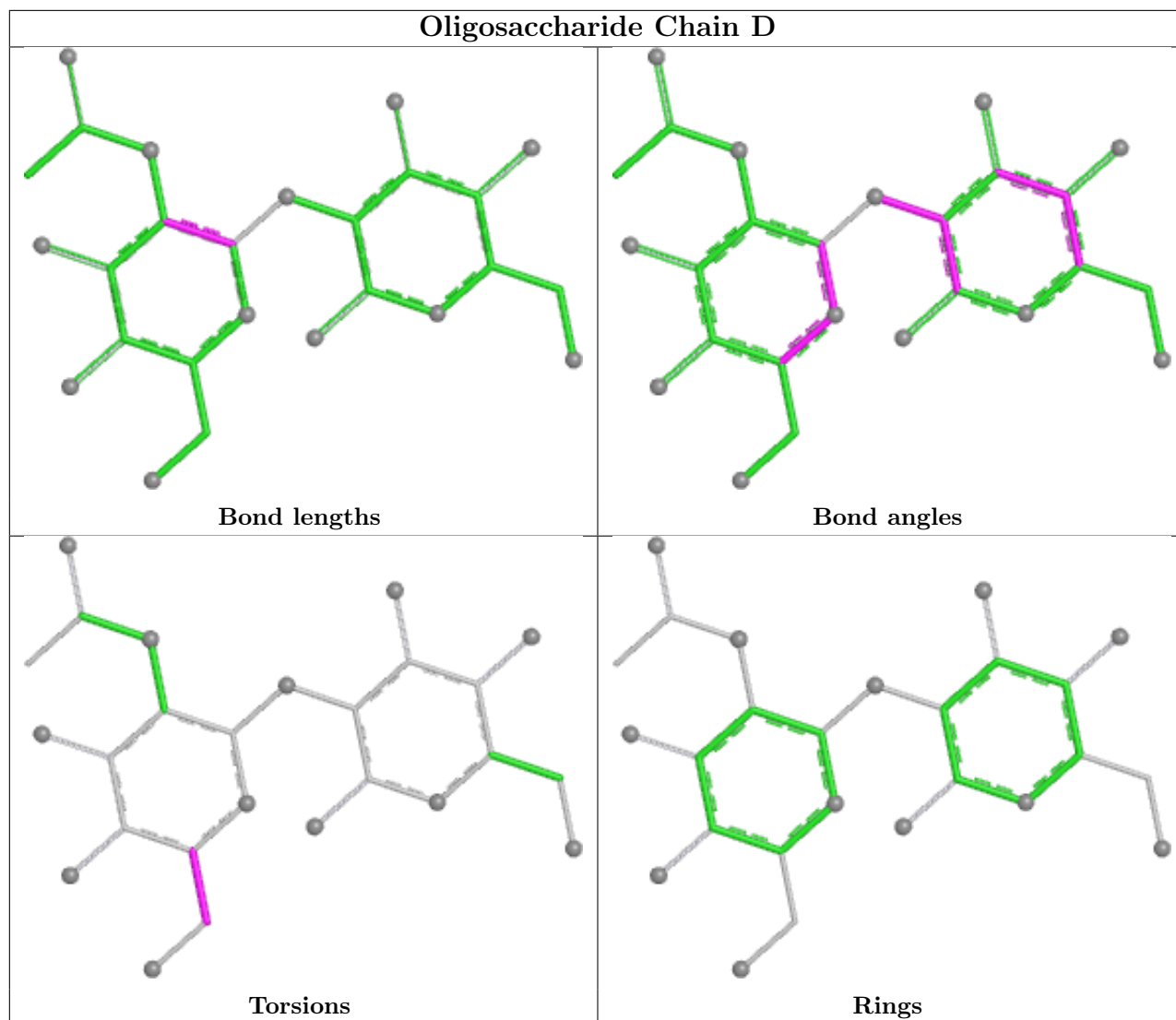
5.6 Ligand geometry (i)

1 ligand 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	$\#|Z| > 2$
3	SO4	A	255	-	4,4,4	0.41	0	6,6,6	0.08	0

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.

5.7 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 ⓘ

6.1 Protein, DNA and RNA chains ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

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

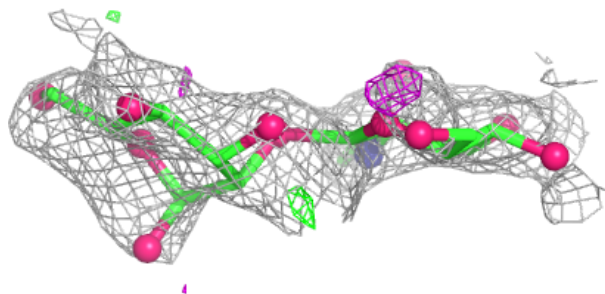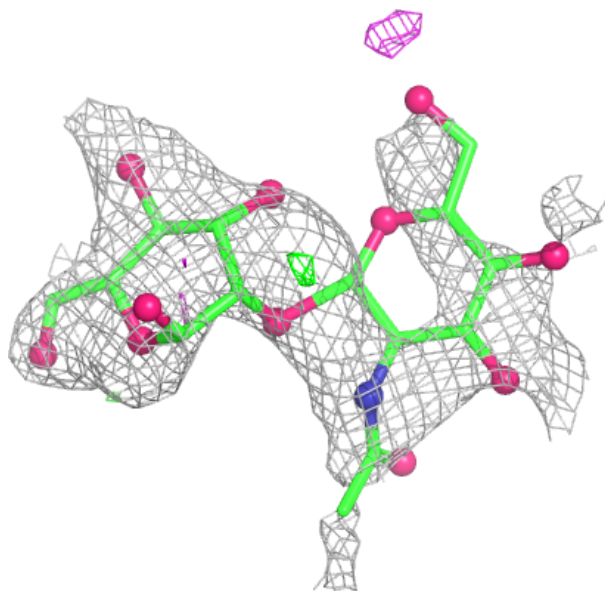Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

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:

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
$mF_o\text{-}DF_c$ (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)
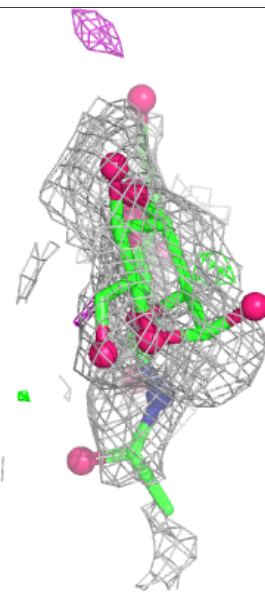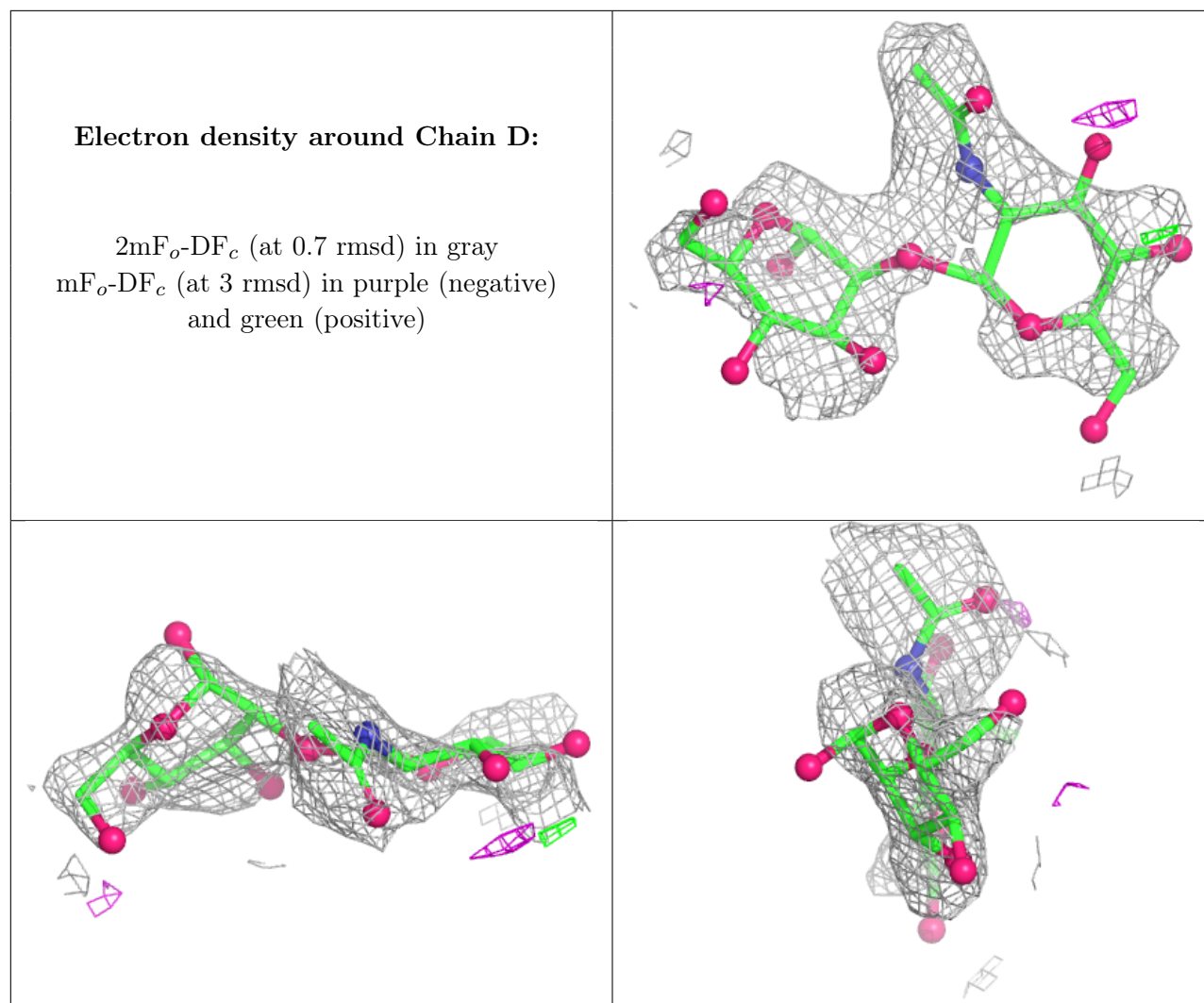
6.4 Ligands ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.