



wwPDB X-ray Structure Validation Summary Report

Oct 5, 2023 – 06:53 AM EDT

PDB ID : 6V49
Title : The crystal structure of hemagglutinin from A/wedge-tailed shearwater/Western Australia/2576/1979 (H15N9)
Authors : Yang, H.; Stevens, J.
Deposited on : 2019-11-27
Resolution : 2.50 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 : **FAILED**
Mogul : 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : **FAILED**
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.35.1

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.

There are no overall percentile quality scores available for this entry.

MolProbity and EDS failed to run properly - the sequence quality summary graphics cannot be shown.

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 |
|-----|------|-------|-----|-----------|----------|---------|------------------|
| 4 | NAG | B | 301 | X | - | - | - |
| 4 | NAG | D | 301 | X | - | - | - |
| 4 | NAG | F | 301 | X | - | - | - |

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 11781 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 Hemagglutinin HA1 chain.

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|-----|-------|----------|-------|------|-----|-----|----|---------|---------|-------|
| | | | Total | C | N | O | S | | | |
| 1 | A | 329 | 2516 | 1566 | 446 | 490 | 14 | 0 | 0 | 0 |
| 1 | C | 329 | 2516 | 1566 | 446 | 490 | 14 | 0 | 0 | 0 |
| 1 | E | 329 | 2516 | 1566 | 446 | 490 | 14 | 0 | 0 | 0 |

There are 9 discrepancies between the modelled and reference sequences:

| Chain | Residue | Modelled | Actual | Comment | Reference |
|-------|---------|----------|--------|----------------|------------|
| A | 0 | GLY | - | expression tag | UNP Q20ND8 |
| A | 132 | SER | THR | conflict | UNP Q20ND8 |
| A | 133 | SER | VAL | conflict | UNP Q20ND8 |
| C | 0 | GLY | - | expression tag | UNP Q20ND8 |
| C | 132 | SER | THR | conflict | UNP Q20ND8 |
| C | 133 | SER | VAL | conflict | UNP Q20ND8 |
| E | 0 | GLY | - | expression tag | UNP Q20ND8 |
| E | 132 | SER | THR | conflict | UNP Q20ND8 |
| E | 133 | SER | VAL | conflict | UNP Q20ND8 |

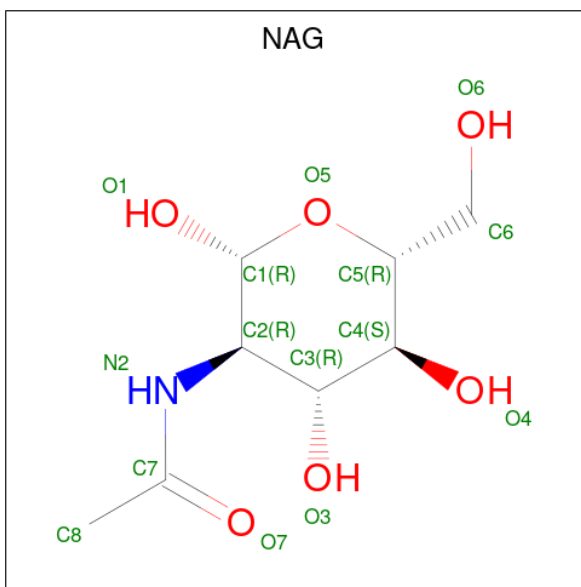
- Molecule 2 is a protein called Hemagglutinin HA2 chain.

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|-----|-------|----------|-------|-----|-----|-----|---|---------|---------|-------|
| | | | Total | C | N | O | S | | | |
| 2 | B | 167 | 1369 | 843 | 245 | 275 | 6 | 0 | 0 | 0 |
| 2 | D | 167 | 1369 | 843 | 245 | 275 | 6 | 0 | 0 | 0 |
| 2 | F | 167 | 1369 | 843 | 245 | 275 | 6 | 0 | 0 | 0 |

- Molecule 3 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 |
|-----|-------|----------|-------|----|---|----|---------|---------|-------|
| 3 | G | 2 | Total | C | N | O | 0 | 0 | 0 |
| | | | 28 | 16 | 2 | 10 | | | |
| 3 | H | 2 | Total | C | N | O | 0 | 0 | 0 |
| | | | 28 | 16 | 2 | 10 | | | |
| 3 | I | 2 | Total | C | N | O | 0 | 0 | 0 |
| | | | 28 | 16 | 2 | 10 | | | |

- 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 |
|-----|-------|----------|-------|---|---|---|---------|---------|
| 4 | B | 1 | Total | C | N | O | 0 | 0 |
| | | | 14 | 8 | 1 | 5 | | |
| 4 | D | 1 | Total | C | N | O | 0 | 0 |
| | | | 14 | 8 | 1 | 5 | | |
| 4 | F | 1 | Total | C | N | O | 0 | 0 |
| | | | 14 | 8 | 1 | 5 | | |

MolProbity and EDS failed to run properly - this section is therefore empty.

3 Data and refinement statistics i

EDS failed to run properly - this section is therefore incomplete.

| Property | Value | Source |
|--|---|-----------|
| Space group | C 1 2 1 | Depositor |
| Cell constants a, b, c, α , β , γ | 108.67Å 101.38Å 163.27Å 90.00° 90.74° 90.00° | Depositor |
| Resolution (Å) | 44.09 – 2.50 | Depositor |
| % Data completeness (in resolution range) | 99.4 (44.09-2.50) | Depositor |
| R_{merge} | (Not available) | Depositor |
| R_{sym} | 0.07 | Depositor |
| $\langle I/\sigma(I) \rangle$ ¹ | 1.68 (at 2.51Å) | Xtrriage |
| Refinement program | REFMAC 5.8.0238 | Depositor |
| R, R_{free} | 0.230 , 0.265 | Depositor |
| Wilson B-factor (Å ²) | 54.3 | Xtrriage |
| Anisotropy | 0.316 | Xtrriage |
| L-test for twinning ² | $\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$ | Xtrriage |
| Estimated twinning fraction | 0.012 for -h,-k,l | Xtrriage |
| Total number of atoms | 11781 | wwPDB-VP |
| Average B, all atoms (Å ²) | 70.0 | wwPDB-VP |

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.81% 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.

4 Model quality [i](#)

4.1 Standard geometry [i](#)

MolProbity failed to run properly - this section is therefore empty.

4.2 Too-close contacts [i](#)

MolProbity failed to run properly - this section is therefore empty.

4.3 Torsion angles [i](#)

4.3.1 Protein backbone [i](#)

MolProbity failed to run properly - this section is therefore empty.

4.3.2 Protein sidechains [i](#)

MolProbity failed to run properly - this section is therefore empty.

4.3.3 RNA [i](#)

MolProbity failed to run properly - this section is therefore empty.

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

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

4.5 Carbohydrates [i](#)

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 |
| 3 | NAG | G | 1 | 3,1 | 14,14,15 | 0.55 | 0 | 17,19,21 | 2.44 | 9 (52%) |
| 3 | NAG | G | 2 | 3 | 14,14,15 | 1.56 | 3 (21%) | 17,19,21 | 2.57 | 7 (41%) |
| 3 | NAG | H | 1 | 3,1 | 14,14,15 | 0.90 | 1 (7%) | 17,19,21 | 2.78 | 8 (47%) |
| 3 | NAG | H | 2 | 3 | 14,14,15 | 1.23 | 1 (7%) | 17,19,21 | 2.36 | 9 (52%) |
| 3 | NAG | I | 1 | 3,1 | 14,14,15 | 0.67 | 0 | 17,19,21 | 1.94 | 4 (23%) |
| 3 | NAG | I | 2 | 3 | 14,14,15 | 0.65 | 0 | 17,19,21 | 1.70 | 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 |
|-----|------|-------|-----|------|---------|-----------|---------|
| 3 | NAG | G | 1 | 3,1 | - | 0/6/23/26 | 0/1/1/1 |
| 3 | NAG | G | 2 | 3 | - | 4/6/23/26 | 0/1/1/1 |
| 3 | NAG | H | 1 | 3,1 | - | 1/6/23/26 | 0/1/1/1 |
| 3 | NAG | H | 2 | 3 | - | 0/6/23/26 | 0/1/1/1 |
| 3 | NAG | I | 1 | 3,1 | - | 2/6/23/26 | 0/1/1/1 |
| 3 | NAG | I | 2 | 3 | - | 2/6/23/26 | 0/1/1/1 |

All (5) bond length outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed(Å) | Ideal(Å) |
|-----|-------|-----|------|-------|------|-------------|----------|
| 3 | G | 2 | NAG | C1-C2 | 3.80 | 1.58 | 1.52 |
| 3 | G | 2 | NAG | O5-C1 | 2.51 | 1.47 | 1.43 |
| 3 | H | 1 | NAG | O7-C7 | 2.30 | 1.28 | 1.23 |
| 3 | G | 2 | NAG | C2-N2 | 2.24 | 1.50 | 1.46 |
| 3 | H | 2 | NAG | O4-C4 | 2.02 | 1.47 | 1.43 |

The worst 5 of 41 bond angle outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed(°) | Ideal(°) |
|-----|-------|-----|------|----------|-------|-------------|----------|
| 3 | H | 1 | NAG | C1-O5-C5 | -6.13 | 103.89 | 112.19 |
| 3 | G | 1 | NAG | C1-O5-C5 | -6.07 | 103.97 | 112.19 |
| 3 | G | 2 | NAG | C1-O5-C5 | 5.66 | 119.86 | 112.19 |
| 3 | G | 2 | NAG | O5-C1-C2 | 5.60 | 120.13 | 111.29 |
| 3 | H | 1 | NAG | C2-N2-C7 | -4.81 | 116.05 | 122.90 |

There are no chirality outliers.

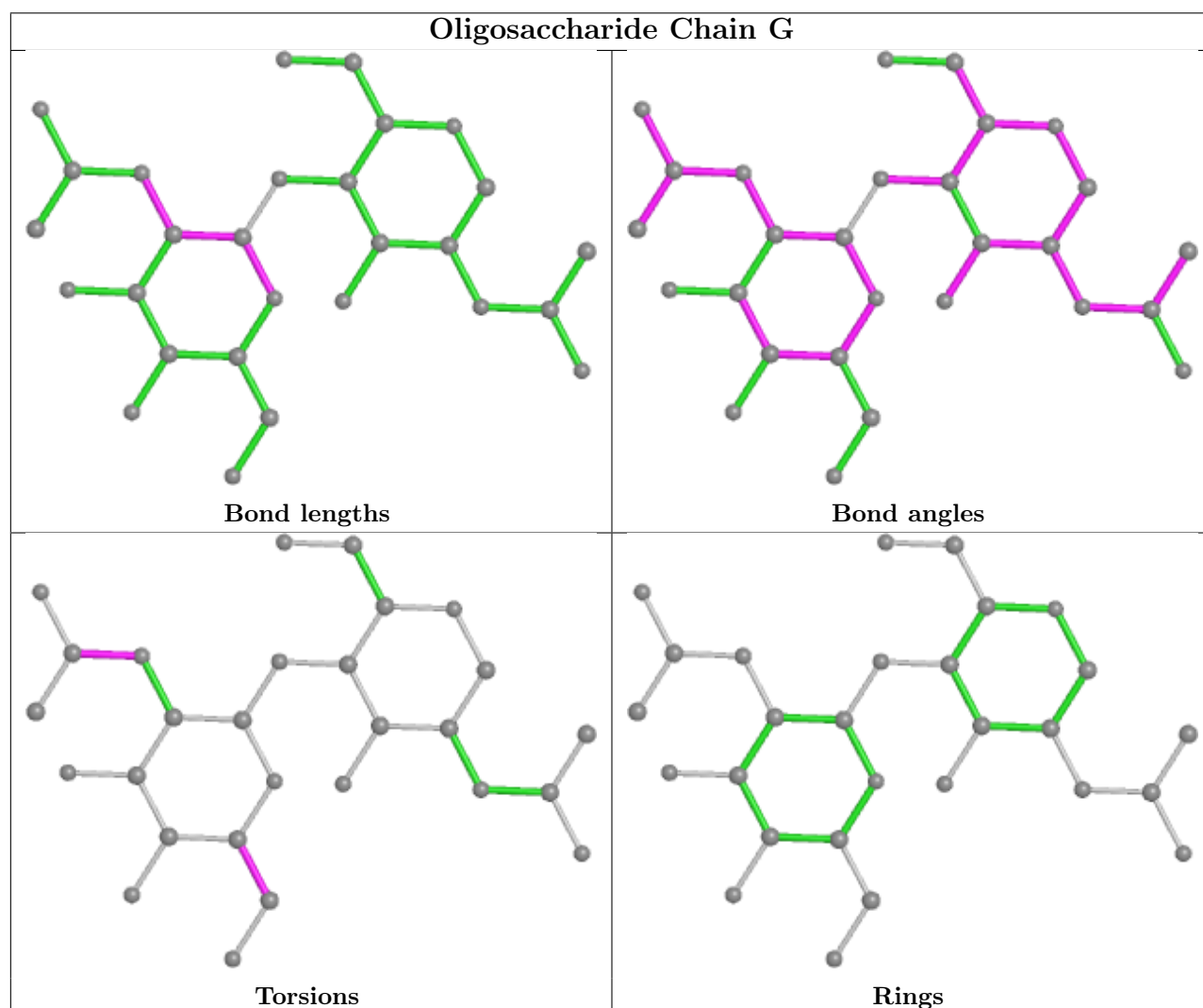
5 of 9 torsion outliers are listed below:

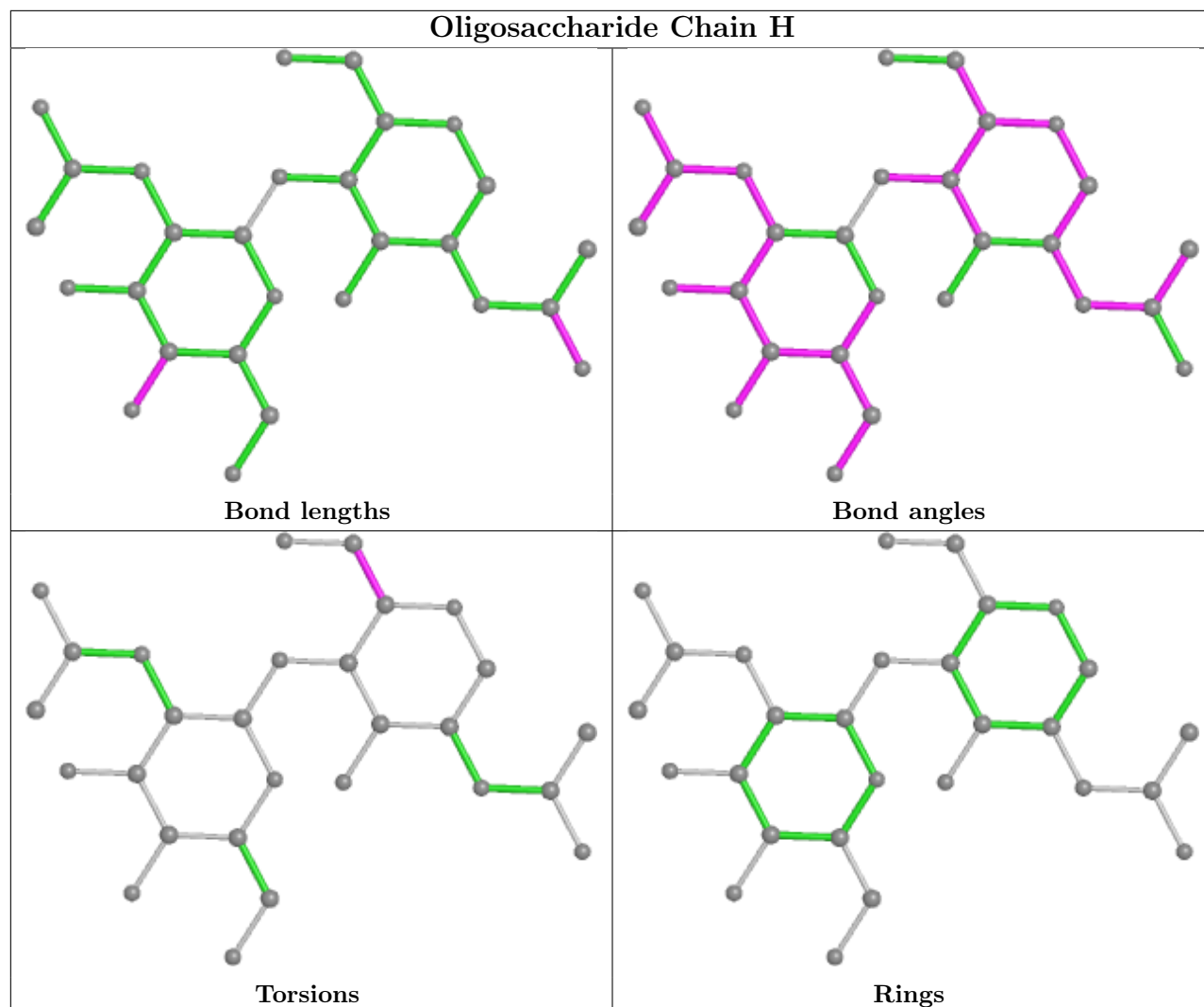
| Mol | Chain | Res | Type | Atoms |
|-----|-------|-----|------|-------------|
| 3 | G | 2 | NAG | C4-C5-C6-O6 |
| 3 | G | 2 | NAG | O5-C5-C6-O6 |
| 3 | I | 2 | NAG | C4-C5-C6-O6 |
| 3 | G | 2 | NAG | C8-C7-N2-C2 |
| 3 | I | 1 | NAG | O5-C5-C6-O6 |

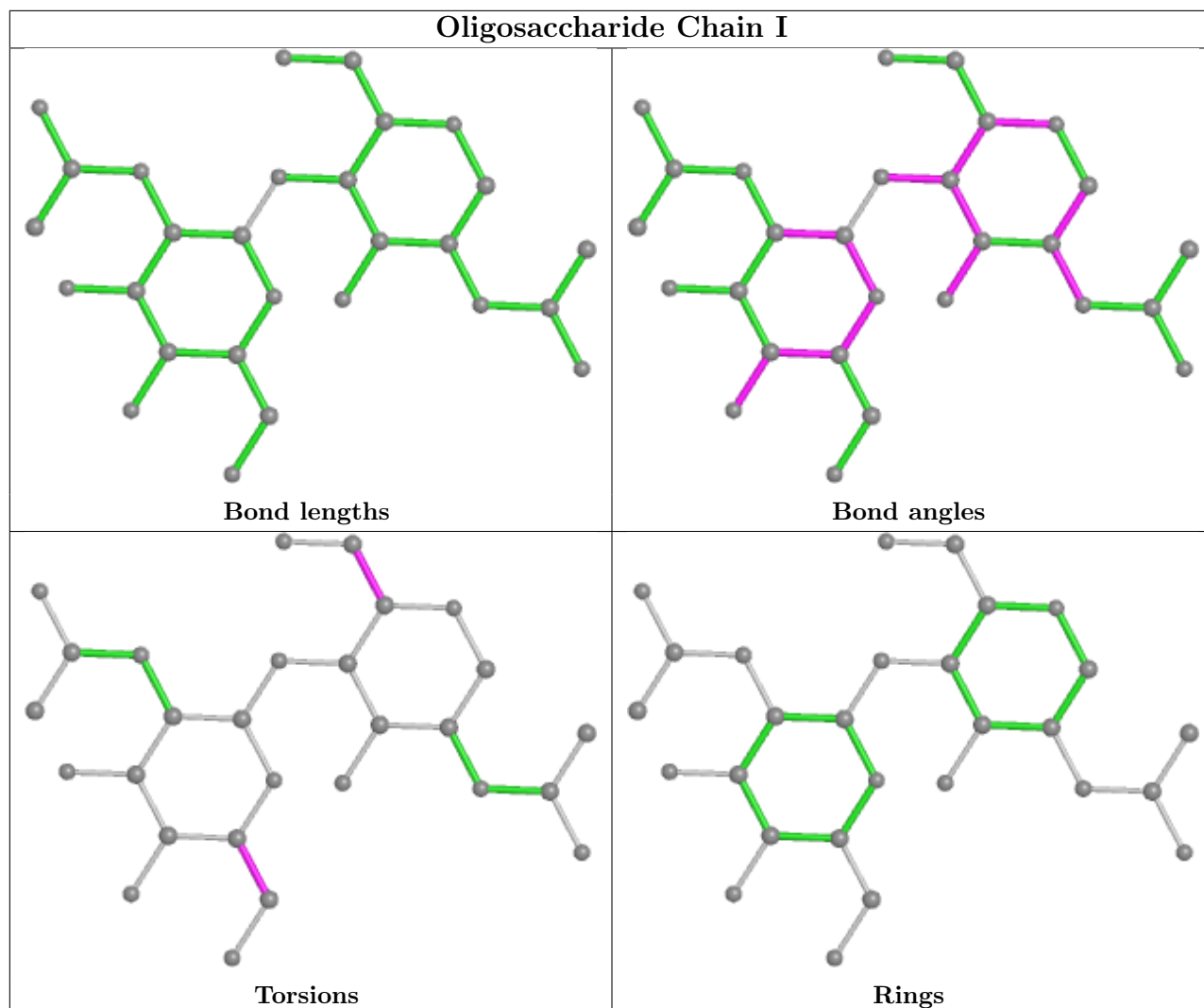
There are no ring outliers.

No monomer is involved in short contacts.

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







4.6 Ligand geometry [i](#)

3 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 | F | 301 | 2 | 14,14,15 | 0.98 | 1 (7%) | 17,19,21 | 1.90 | 6 (35%) |
| 4 | NAG | B | 301 | 2 | 14,14,15 | 1.32 | 1 (7%) | 17,19,21 | 2.51 | 10 (58%) |
| 4 | NAG | D | 301 | 2 | 14,14,15 | 1.25 | 2 (14%) | 17,19,21 | 1.83 | 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 |
|-----|------|-------|-----|------|---------|-----------|---------|
| 4 | NAG | F | 301 | 2 | 1/1/5/7 | 3/6/23/26 | 0/1/1/1 |
| 4 | NAG | B | 301 | 2 | 1/1/5/7 | 2/6/23/26 | 0/1/1/1 |
| 4 | NAG | D | 301 | 2 | 1/1/5/7 | 2/6/23/26 | 0/1/1/1 |

All (4) bond length outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed(Å) | Ideal(Å) |
|-----|-------|-----|------|-------|------|-------------|----------|
| 4 | B | 301 | NAG | C1-C2 | 4.21 | 1.58 | 1.52 |
| 4 | D | 301 | NAG | C1-C2 | 3.28 | 1.57 | 1.52 |
| 4 | F | 301 | NAG | C1-C2 | 2.55 | 1.56 | 1.52 |
| 4 | D | 301 | NAG | C3-C2 | 2.31 | 1.57 | 1.52 |

The worst 5 of 20 bond angle outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed(°) | Ideal(°) |
|-----|-------|-----|------|----------|------|-------------|----------|
| 4 | D | 301 | NAG | C4-C3-C2 | 4.41 | 117.48 | 111.02 |
| 4 | B | 301 | NAG | C1-O5-C5 | 4.25 | 117.96 | 112.19 |
| 4 | B | 301 | NAG | C1-C2-N2 | 4.18 | 117.64 | 110.49 |
| 4 | F | 301 | NAG | C4-C3-C2 | 3.79 | 116.58 | 111.02 |
| 4 | B | 301 | NAG | O5-C1-C2 | 3.72 | 117.17 | 111.29 |

All (3) chirality outliers are listed below:

| Mol | Chain | Res | Type | Atom |
|-----|-------|-----|------|------|
| 4 | B | 301 | NAG | C1 |
| 4 | D | 301 | NAG | C1 |
| 4 | F | 301 | NAG | C1 |

5 of 7 torsion outliers are listed below:

| Mol | Chain | Res | Type | Atoms |
|-----|-------|-----|------|-------------|
| 4 | B | 301 | NAG | O5-C5-C6-O6 |
| 4 | D | 301 | NAG | C8-C7-N2-C2 |
| 4 | B | 301 | NAG | C4-C5-C6-O6 |
| 4 | D | 301 | NAG | O7-C7-N2-C2 |
| 4 | F | 301 | NAG | C8-C7-N2-C2 |

There are no ring outliers.

No monomer is involved in short contacts.

4.7 Other polymers

There are no such residues in this entry.

4.8 Polymer linkage issues

There are no chain breaks in this entry.

5 Fit of model and data

5.1 Protein, DNA and RNA chains

EDS failed to run properly - this section is therefore empty.

5.2 Non-standard residues in protein, DNA, RNA chains

EDS failed to run properly - this section is therefore empty.

5.3 Carbohydrates

EDS failed to run properly - this section is therefore empty.

5.4 Ligands

EDS failed to run properly - this section is therefore empty.

5.5 Other polymers

EDS failed to run properly - this section is therefore empty.