# wwPDB X-ray Structure Validation Summary Report (i)

Oct 5, 2023 – 03:53 AM EDT

| | | |
|---|---|---|
| PDB ID | : | 6V80 |
| Title | : | Crystal structure of human CD1d presenting alpha-Galactosylceramide in complex with NKT12 TCR and VHH nanobody 1D12 |
| Authors | : | Shahine, A.; Rossjohn, J. |
| Deposited on | : | 2019-12-10 |
| Resolution | : | 3.53 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 (i) 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 (1)) were used in the production of this report:

| | | |
|---|---|---|
| MolProbity | : | FAILED |
| Mogul | : | 1.8.5 (274361), CSD as541be (2020) |
| Xtriage (Phenix) | : | 1.13 |
| EDS | : | FAILED |
| buster-report | : | 1.1.7 (2018) |
| 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 3.53 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.

# 2 Entry composition (i)

There are 8 unique types of molecules in this entry. The entry contains 14208 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 Antigen-presenting glycoprotein CD1d.

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | A | 274 | Total | C | N | O | S | 0 | 0 | 0 |
| | | | 2130 | 1362 | 372 | 389 | 7 | | | |
| 1 | F | 270 | Total | C | N | O | S | 0 | 0 | 0 |
| | | | 2075 | 1327 | 366 | 376 | 6 | | | |

There are 2 discrepancies between the modelled and reference sequences:

| Chain | Residue | Modelled | Actual | Comment | Reference |
|---|---|---|---|---|---|
| A | 4 | MET | - | initiating methionine | UNP P15813 |
| F | 4 | MET | - | initiating methionine | UNP P15813 |

- Molecule 2 is a protein called Beta-2-microglobulin.

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|---|
| 2 | B | 98 | Total | C | N | O | S | 0 | 0 | 0 |
| | | | 743 | 477 | 116 | 148 | 2 | | | |
| 2 | G | 94 | Total | C | N | O | S | 0 | 0 | 0 |
| | | | 745 | 473 | 126 | 144 | 2 | | | |

There are 2 discrepancies between the modelled and reference sequences:

| Chain | Residue | Modelled | Actual | Comment | Reference |
|---|---|---|---|---|---|
| B | 0 | MET | - | initiating methionine | UNP P61769 |
| G | 0 | MET | - | initiating methionine | UNP P61769 |

- Molecule 3 is a protein called T cell receptor alpha variable 10, nkt tcr alpha chain fusion.

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|---|
| 3 | C | 198 | Total | C | N | O | S | 0 | 0 | 0 |
| | | | 1514 | 937 | 256 | 312 | 9 | | | |
| 3 | H | 181 | Total | C | N | O | S | 0 | 0 | 0 |
| | | | 1376 | 844 | 237 | 287 | 8 | | | |

There are 4 discrepancies between the modelled and reference sequences:

| Chain | Residue | Modelled | Actual | Comment | Reference |
|---|---|---|---|---|---|
| C | 1 | HIS | - | expression tag | UNP A0A0B4J240 |
| C | 2 | MET | - | expression tag | UNP A0A0B4J240 |
| H | 1 | HIS | - | expression tag | UNP A0A0B4J240 |
| H | 2 | MET | - | expression tag | UNP A0A0B4J240 |

- Molecule 4 is a protein called nkt tcr beta chain.

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|---|
| 4 | D | 240 | Total | C | N | O | S | 0 | 0 | 0 |
| | | | 1848 | 1166 | 317 | 357 | 8 | | | |
| 4 | I | 237 | Total | C | N | O | S | 0 | 0 | 0 |
| | | | 1888 | 1190 | 327 | 363 | 8 | | | |

There are 16 discrepancies between the modelled and reference sequences:

| Chain | Residue | Modelled | Actual | Comment | Reference |
|---|---|---|---|---|---|
| D | 92 | THR | SER | linker | UNP K7N5M4 |
| D | 93 | SER | GLU | linker | UNP K7N5M4 |
| D | 94 | ARG | ASN | linker | UNP K7N5M4 |
| D | 95 | ARG | SER | linker | UNP K7N5M4 |
| D | 97 | SER | THR | linker | UNP K7N5M4 |
| D | ? | - | GLY | deletion | UNP K7N5M4 |
| D | ? | - | ARG | deletion | UNP K7N5M4 |
| D | ? | - | ILE | deletion | UNP K7N5M4 |
| I | 92 | THR | SER | linker | UNP K7N5M4 |
| I | 93 | SER | GLU | linker | UNP K7N5M4 |
| I | 94 | ARG | ASN | linker | UNP K7N5M4 |
| I | 95 | ARG | SER | linker | UNP K7N5M4 |
| I | 97 | SER | THR | linker | UNP K7N5M4 |
| I | ? | - | GLY | deletion | UNP K7N5M4 |
| I | ? | - | ARG | deletion | UNP K7N5M4 |
| I | ? | - | ILE | deletion | UNP K7N5M4 |

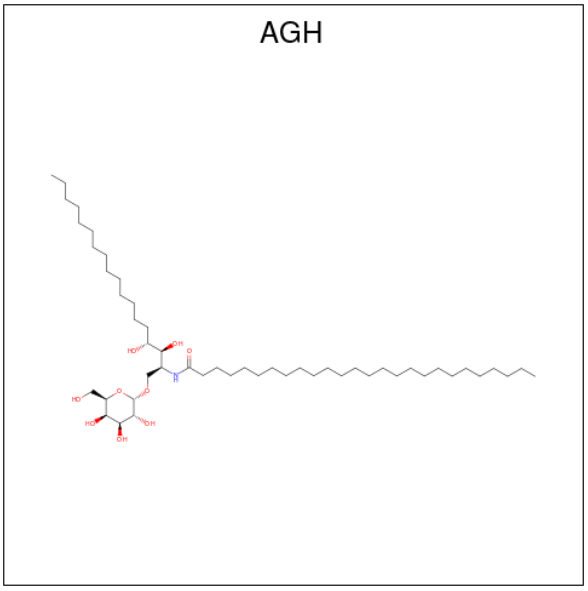
- Molecule 5 is a protein called Nanobody VHH ID12.

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|---|
| 5 | E | 117 | Total | C | N | O | S | 0 | 0 | 0 |
| | | | 851 | 529 | 144 | 174 | 4 | | | |
| 5 | J | 118 | Total | C | N | O | S | 0 | 1 | 0 |
| | | | 834 | 522 | 140 | 168 | 4 | | | |

- Molecule 6 is an oligosaccharide called 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.

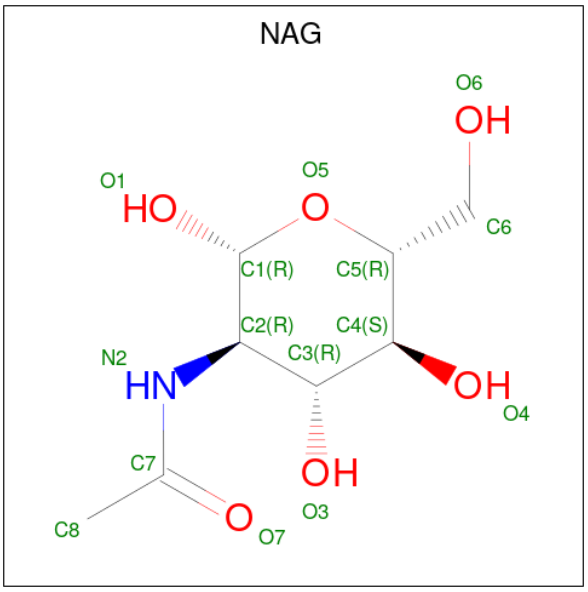
| Mol | Chain | Residues | Atoms | | | | ZeroOcc | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 6 | K | 2 | Total | C | N | O | 0 | 0 | 0 |
| | | | 28 | 16 | 2 | 10 | | | |
| 6 | L | 2 | Total | C | N | O | 0 | 0 | 0 |
| | | | 28 | 16 | 2 | 10 | | | |

- Molecule 7 is N-{(1S,2R,3S)-1-[(ALPHA-D-GALACTOPYRANOSYLOXY)METHYL]-2,3-DIHYDROXYHEPTADECYL}HEXACOSANAMIDE (three-letter code: AGH) (formula: $C_{50}H_{99}NO_9$).




| Mol | Chain | Residues | Atoms | | | | ZeroOcc | AltConf |
|---|---|---|---|---|---|---|---|---|
| 7 | A | 1 | Total | C | N | O | 7 | 0 |
| | | | 60 | 50 | 1 | 9 | | |
| 7 | F | 1 | Total | C | N | O | 16 | 0 |
| | | | 60 | 50 | 1 | 9 | | |

- Molecule 8 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).




| Mol | Chain | Residues | Atoms | ZeroOcc | AltConf |
|---|---|---|---|---|---|
| 8 | F | 1 | Total C N O<br>14 8 1 5 | 0 | 0 |
| 8 | F | 1 | Total C N O<br>14 8 1 5 | 0 | 0 |

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<br>a, b, c, $\alpha$, $\beta$, $\gamma$ | 210.67Å 165.25Å 84.50Å<br>90.00° 90.00° 90.00° | Depositor |
| Resolution (Å) | 82.63 – 3.53 | Depositor |
| % Data completeness<br>(in resolution range) | 100.0 (82.63-3.53) | Depositor |
| $R_{merge}$ | (Not available) | Depositor |
| $R_{sym}$ | (Not available) | Depositor |
| $< I/\sigma(I) >$ [1] | 1.76 (at 3.49Å) | Xtriage |
| Refinement program | PHENIX 1.14_3260 | Depositor |
| R, $R_{free}$ | 0.195 , 0.254 | Depositor |
| Wilson B-factor (Å$^2$) | 104.5 | Xtriage |
| Anisotropy | 0.327 | Xtriage |
| L-test for twinning[2] | $< \|L\| > = 0.36$, $< L^2 > = 0.19$ | Xtriage |
| Estimated twinning fraction | 0.297 for -h,-k,l | Xtriage |
| Reported twinning fraction | 0.390 for -h,-k,l | Depositor |
| Outliers | 0 of 35558 reflections | Xtriage |
| Total number of atoms | 14208 | wwPDB-VP |
| Average B, all atoms (Å$^2$) | 120.0 | wwPDB-VP |

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.70% 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.

# 4 Model quality

## 4.1 Standard geometry

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

## 4.2 Too-close contacts

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

## 4.3 Torsion angles

### 4.3.1 Protein backbone

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

### 4.3.2 Protein sidechains

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

### 4.3.3 RNA

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

## 4.4 Non-standard residues in protein, DNA, RNA chains

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

## 4.5 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 | $\#\|Z\|>2$ |
| 6 | NAG | K | 1 | 6,1 | 14,14,15 | 0.47 | 0 | 17,19,21 | 0.69 | 0 |
| 6 | NAG | K | 2 | 6 | 14,14,15 | 0.25 | 0 | 17,19,21 | 0.28 | 0 |
| 6 | NAG | L | 1 | 6,1 | 14,14,15 | 1.48 | 4 (28%) | 17,19,21 | 3.88 | 8 (47%) |
| 6 | NAG | L | 2 | 6 | 14,14,15 | 0.69 | 0 | 17,19,21 | 1.99 | 6 (35%) |

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 | K | 1 | 6,1 | - | 2/6/23/26 | 0/1/1/1 |
| 6 | NAG | K | 2 | 6 | - | 1/6/23/26 | 0/1/1/1 |
| 6 | NAG | L | 1 | 6,1 | - | 3/6/23/26 | 0/1/1/1 |
| 6 | NAG | L | 2 | 6 | - | 4/6/23/26 | 0/1/1/1 |

All (4) bond length outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed(Å) | Ideal(Å) |
|---|---|---|---|---|---|---|---|
| 6 | L | 1 | NAG | O5-C1 | -2.64 | 1.39 | 1.43 |
| 6 | L | 1 | NAG | C2-N2 | -2.35 | 1.42 | 1.46 |
| 6 | L | 1 | NAG | C4-C5 | -2.04 | 1.48 | 1.53 |
| 6 | L | 1 | NAG | C1-C2 | -2.02 | 1.49 | 1.52 |

The worst 5 of 14 bond angle outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed($^o$) | Ideal($^o$) |
|---|---|---|---|---|---|---|---|
| 6 | L | 1 | NAG | C2-N2-C7 | -11.12 | 107.06 | 122.90 |
| 6 | L | 1 | NAG | C6-C5-C4 | -6.43 | 97.95 | 113.00 |
| 6 | L | 1 | NAG | O3-C3-C4 | -5.10 | 98.56 | 110.35 |
| 6 | L | 2 | NAG | O5-C5-C6 | 4.86 | 114.83 | 107.20 |
| 6 | L | 1 | NAG | O4-C4-C3 | 4.19 | 120.05 | 110.35 |

There are no chirality outliers.

5 of 10 torsion outliers are listed below:

| Mol | Chain | Res | Type | Atoms |
|---|---|---|---|---|
| 6 | L | 1 | NAG | C3-C2-N2-C7 |
| 6 | L | 1 | NAG | C8-C7-N2-C2 |
| 6 | L | 1 | NAG | O7-C7-N2-C2 |

*Continued on next page...*

*Continued from previous page...*

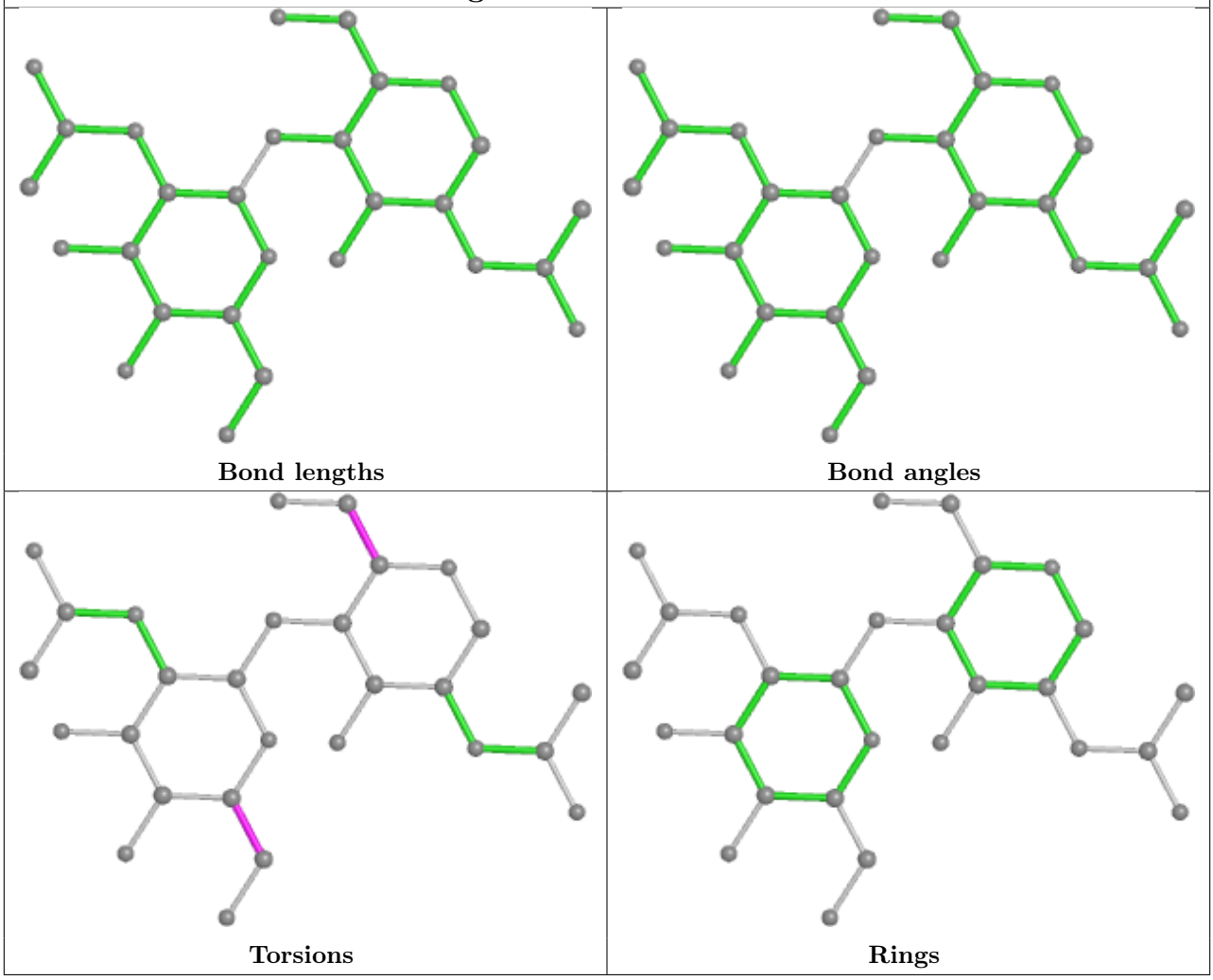
| Mol | Chain | Res | Type | Atoms |
|---|---|---|---|---|
| 6 | L | 2 | NAG | C8-C7-N2-C2 |
| 6 | K | 1 | NAG | O5-C5-C6-O6 |

There are no ring outliers.

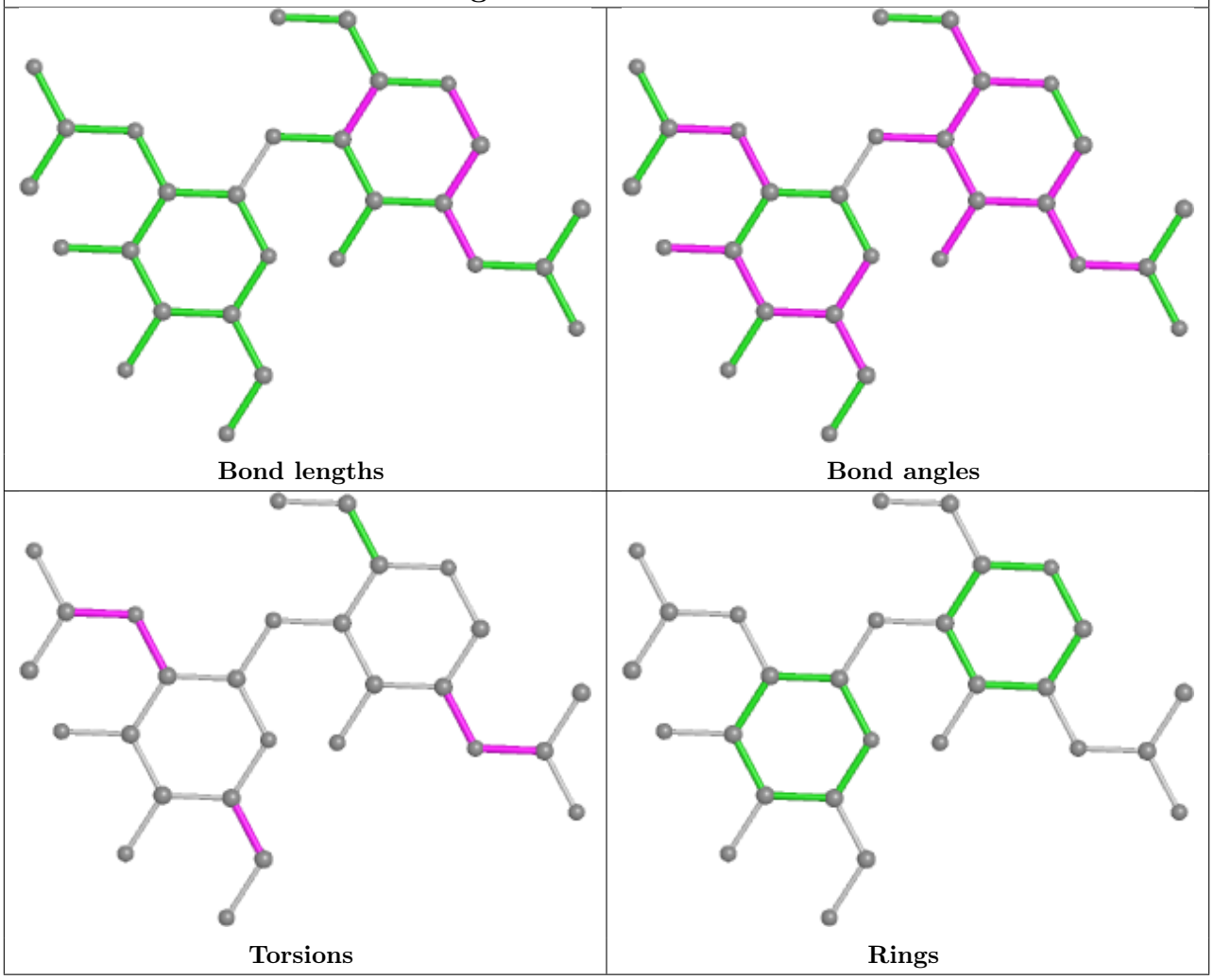
No monomer is involved in short contacts.

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

**Oligosaccharide Chain K**



**Oligosaccharide Chain L**




## 4.6 Ligand geometry (i)

4 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 | $\#\|Z\| > 2$ |
| 7 | AGH | A | 401 | - | 60,60,60 | 1.40 | 5 (8%) | 65,69,69 | 1.02 | 4 (6%) |
| 8 | NAG | F | 404 | 1 | 14,14,15 | 0.76 | 0 | 17,19,21 | 1.55 | 4 (23%) |
| 7 | AGH | F | 401 | - | 60,60,60 | 1.38 | 5 (8%) | 65,69,69 | 1.15 | 4 (6%) |

| Mol | Type | Chain | Res | Link | Bond lengths | | | Bond angles | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Counts | RMSZ | $\#\|Z\|>2$ | Counts | RMSZ | $\#\|Z\|>2$ |
| 8 | NAG | F | 405 | 1 | 14,14,15 | 1.10 | 2 (14%) | 17,19,21 | 3.02 | 6 (35%) |

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 |
|---|---|---|---|---|---|---|---|
| 7 | AGH | A | 401 | - | - | 15/58/78/78 | 0/1/1/1 |
| 8 | NAG | F | 404 | 1 | - | 4/6/23/26 | 0/1/1/1 |
| 7 | AGH | F | 401 | - | - | 22/58/78/78 | 0/1/1/1 |
| 8 | NAG | F | 405 | 1 | - | 5/6/23/26 | 0/1/1/1 |

The worst 5 of 12 bond length outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed(Å) | Ideal(Å) |
|---|---|---|---|---|---|---|---|
| 7 | A | 401 | AGH | CAA-N2 | 7.02 | 1.49 | 1.34 |
| 7 | F | 401 | AGH | CAA-N2 | 6.96 | 1.48 | 1.34 |
| 7 | F | 401 | AGH | O1A-C1 | -3.08 | 1.38 | 1.43 |
| 7 | A | 401 | AGH | O1A-C1 | -2.98 | 1.38 | 1.43 |
| 7 | A | 401 | AGH | CAB-CAA | 2.78 | 1.56 | 1.51 |

The worst 5 of 18 bond angle outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed($^o$) | Ideal($^o$) |
|---|---|---|---|---|---|---|---|
| 8 | F | 405 | NAG | C1-O5-C5 | 7.61 | 122.50 | 112.19 |
| 8 | F | 405 | NAG | O5-C1-C2 | -6.09 | 101.66 | 111.29 |
| 8 | F | 405 | NAG | C1-C2-N2 | -4.26 | 103.20 | 110.49 |
| 8 | F | 405 | NAG | C2-N2-C7 | -3.98 | 117.23 | 122.90 |
| 7 | F | 401 | AGH | C6-C5-C4 | -3.67 | 108.15 | 114.18 |

There are no chirality outliers.

5 of 46 torsion outliers are listed below:

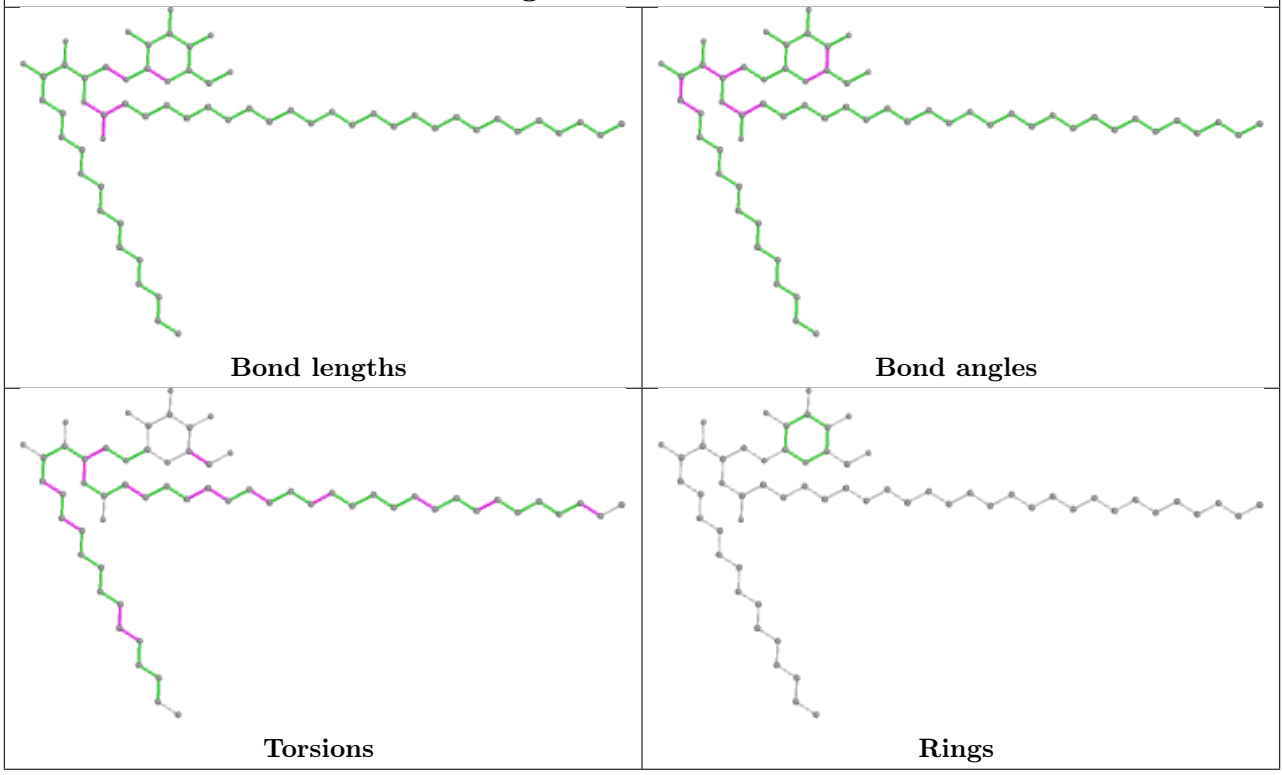
| Mol | Chain | Res | Type | Atoms |
|---|---|---|---|---|
| 7 | A | 401 | AGH | O1A-C1-C2-C3 |
| 7 | F | 401 | AGH | N2-C2-C3-O3 |
| 8 | F | 404 | NAG | C8-C7-N2-C2 |
| 8 | F | 404 | NAG | O7-C7-N2-C2 |
| 8 | F | 405 | NAG | C3-C2-N2-C7 |

There are no ring outliers.

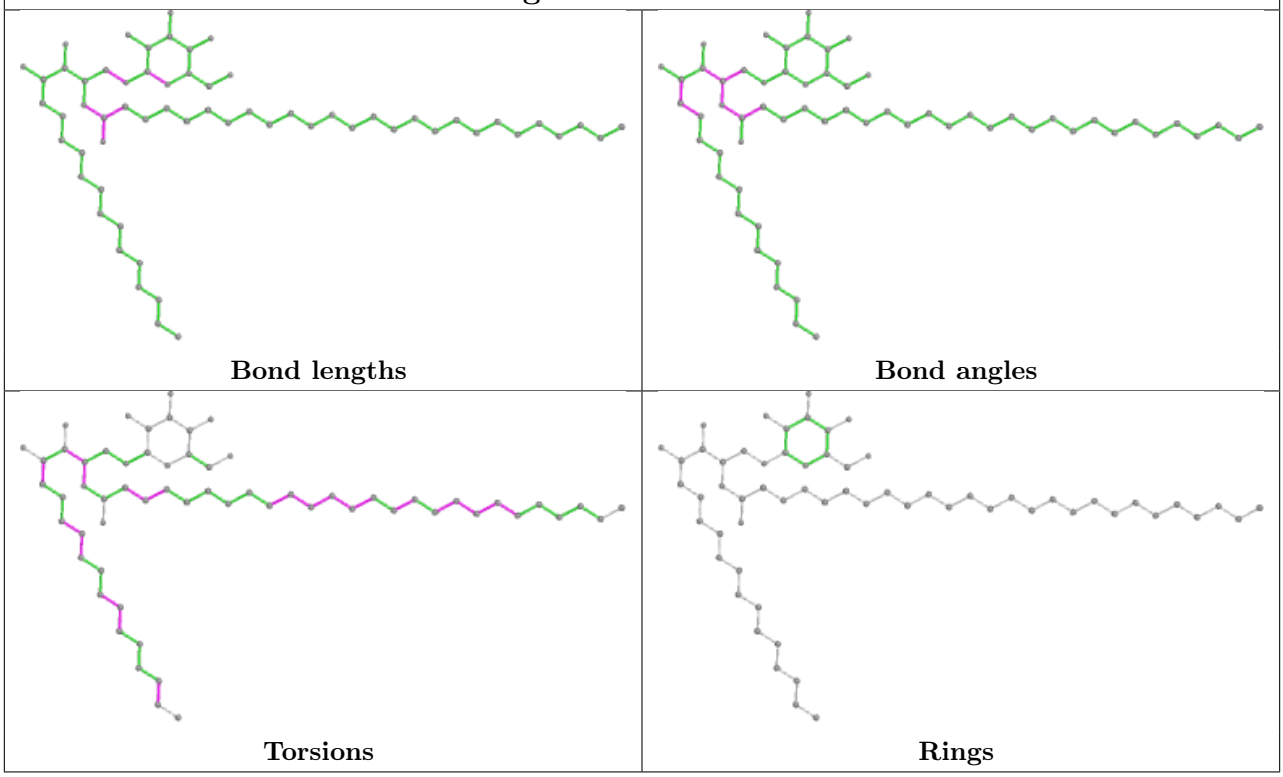
No monomer is involved in short contacts.

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

**Ligand AGH A 401**



**Ligand AGH F 401**




## 4.7 Other polymers (i)

There are no such residues in this entry.

## 4.8 Polymer linkage issues (i)

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.