



# wwPDB X-ray Structure Validation Summary Report ⓘ

Dec 17, 2023 – 04:05 pm GMT

PDB ID : 4CPY  
Title : Structure of the Neuraminidase from the B/Lyon/CHU/15.216/2011 virus in complex with Oseltamivir  
Authors : Vachieri, S.G.; Collins, P.J.; Escuret, V.; Casalegno, J.S.; Cattle, N.; Ferraris, O.; Sabatier, M.; Frobert, E.; Caro, V.; Skehel, J.J.; Gamblin, S.J.; Valla, F.; Valette, M.; Ottmann, M.; McCauley, J.W.; Daniels, R.S.; Lina, B.  
Deposited on : 2014-02-09  
Resolution : 1.80 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](mailto: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](#) ①) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.8.4, CSD as541be (2020)  
Xtrriage (Phenix) : 1.13  
EDS : 2.36  
buster-report : 1.1.7 (2018)  
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

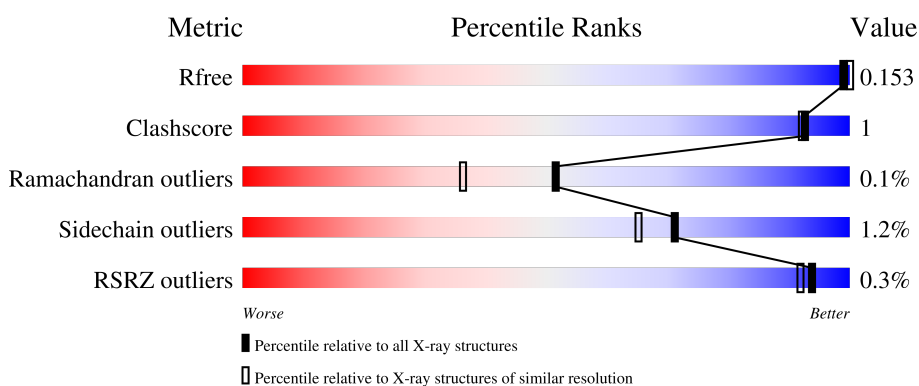
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.80 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<br>(#Entries) | Similar resolution<br>(#Entries, resolution range(Å)) |
|-----------------------|-----------------------------|---|
| $R_{free}$            | 130704                      | 5950 (1.80-1.80)                                      |
| Clashscore            | 141614                      | 6793 (1.80-1.80)                                      |
| Ramachandran outliers | 138981                      | 6697 (1.80-1.80)                                      |
| Sidechain outliers    | 138945                      | 6696 (1.80-1.80)                                      |
| RSRZ outliers         | 127900                      | 5850 (1.80-1.80)                                      |

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  $\geq 3$ , 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     | 466    |                  |
| 1   | B     | 466    |                  |
| 2   | C     | 2      |                  |
| 2   | D     | 2      |                  |

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     | 1469 | X         | -        | -       | -                |

## 2 Entry composition [i](#)

There are 7 unique types of molecules in this entry. The entry contains 6958 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 NEURAMINIDASE.

| Mol | Chain | Residues | Atoms         |           |          |          |         | ZeroOcc | AltConf | Trace |
|-----|-------|----------|---------------|-----------|----------|----------|---------|---------|---------|-------|
|     |       |          | Total         | C         | N        | O        | S       |         |         |       |
| 1   | A     | 390      | Total<br>3037 | C<br>1904 | N<br>521 | O<br>583 | S<br>29 | 0       | 5       | 0     |
| 1   | B     | 390      | Total<br>3024 | C<br>1896 | N<br>520 | O<br>579 | S<br>29 | 0       | 2       | 0     |

- 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   | C     | 2        | Total<br>28 | C<br>16 | N<br>2 | O<br>10 | 0       | 0       | 0     |
| 2   | D     | 2        | Total<br>28 | C<br>16 | N<br>2 | O<br>10 | 0       | 0       | 0     |

- Molecule 3 is CALCIUM ION (three-letter code: CA) (formula: Ca).

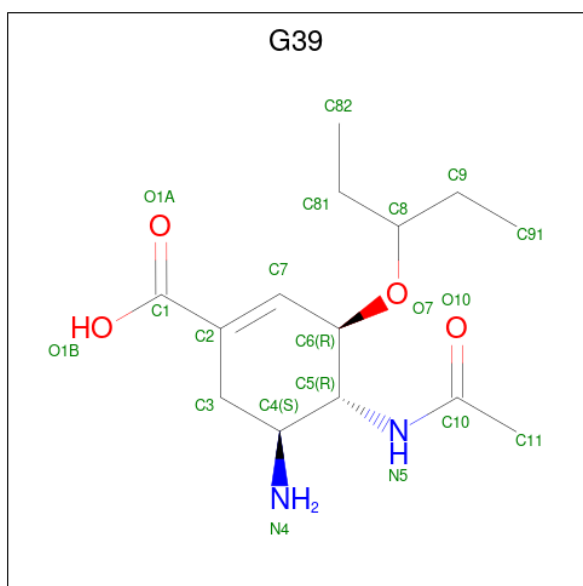
| Mol | Chain | Residues | Atoms              | ZeroOcc | AltConf |
|-----|-------|----------|--------------------|---------|---------|
| 3   | A     | 1        | Total<br>Ca<br>1 1 | 0       | 0       |
| 3   | B     | 1        | Total<br>Ca<br>1 1 | 0       | 0       |

- Molecule 4 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: C<sub>8</sub>H<sub>15</sub>NO<sub>6</sub>).



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

- Molecule 5 is (3R,4R,5S)-4-(acetylamino)-5-amino-3-(pentan-3-yloxy)cyclohex-1-ene-1-carboxylic acid (three-letter code: G39) (formula:  $C_{14}H_{24}N_2O_4$ ).



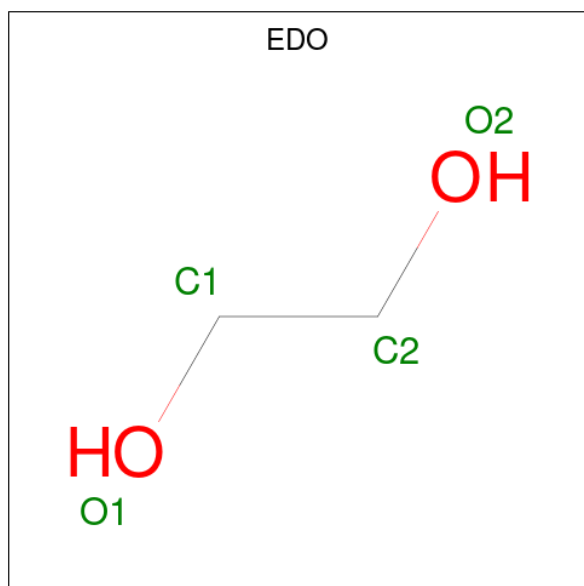
| Mol | Chain | Residues | Atoms |    |   |   | ZeroOcc | AltConf |
|-----|-------|----------|-------|----|---|---|---------|---------|
|     |       |          | Total | C  | N | O |         |         |
| 5   | A     | 1        | 20    | 14 | 2 | 4 | 0       | 0       |

*Continued on next page...*

Continued from previous page...

| Mol | Chain | Residues | Atoms |    |   | ZeroOcc | AltConf |   |
|-----|-------|----------|-------|----|---|---------|---------|---|
|     |       |          | Total | C  | N |         |         | O |
| 5   | B     | 1        | 20    | 14 | 2 | 4       | 0       | 0 |

- Molecule 6 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



| Mol | Chain | Residues | Atoms |   |   | ZeroOcc | AltConf |
|-----|-------|----------|-------|---|---|---------|---------|
|     |       |          | Total | C | O |         |         |
| 6   | A     | 1        | 4     | 2 | 2 | 0       | 0       |
| 6   | A     | 1        | 4     | 2 | 2 | 0       | 0       |
| 6   | A     | 1        | 4     | 2 | 2 | 0       | 0       |
| 6   | A     | 1        | 4     | 2 | 2 | 0       | 0       |
| 6   | A     | 1        | 4     | 2 | 2 | 0       | 0       |
| 6   | A     | 1        | 4     | 2 | 2 | 0       | 0       |
| 6   | A     | 1        | 4     | 2 | 2 | 0       | 0       |
| 6   | A     | 1        | 4     | 2 | 2 | 0       | 0       |
| 6   | A     | 1        | 4     | 2 | 2 | 0       | 0       |
| 6   | A     | 1        | 4     | 2 | 2 | 0       | 0       |

Continued on next page...

*Continued from previous page...*

| Mol | Chain | Residues | Atoms              | ZeroOcc | AltConf |
|-----|-------|----------|--------------------|---------|---------|
| 6   | A     | 1        | Total C O<br>4 2 2 | 0       | 0       |
| 6   | B     | 1        | Total C O<br>4 2 2 | 0       | 0       |
| 6   | B     | 1        | Total C O<br>4 2 2 | 0       | 0       |
| 6   | B     | 1        | Total C O<br>4 2 2 | 0       | 0       |
| 6   | B     | 1        | Total C O<br>4 2 2 | 0       | 0       |
| 6   | B     | 1        | Total C O<br>4 2 2 | 0       | 0       |
| 6   | B     | 1        | Total C O<br>4 2 2 | 0       | 0       |

- Molecule 7 is water.


| Mol | Chain | Residues | Atoms              | ZeroOcc | AltConf |
|-----|-------|----------|--------------------|---------|---------|
| 7   | A     | 381      | Total O<br>381 381 | 0       | 0       |
| 7   | B     | 322      | Total O<br>322 322 | 0       | 0       |



### 3 Residue-property plots [i](#)

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: NEURAMINIDASE

Chain A:  80% 16%

MET LEU PRO SER THR ILE GLN THR LEU LEU THR LEU PHE LEU THR LEU PRO GLY VAL LEU LEU SER LEU TYR VAL SER ALA LEU LEU TYR LEU TYR SER ASP ILE LEU LEU LYS PHE SER PRO THR GLU ILE THR ALA PRO THR MET SER LEU ASP CYS ALA ASN ALA SER ASN VAL

GLN ALA VAL ASN ARG ARG ALA THR LYS LEU VAL VAL PHE LEU LEU LEU PRO GLY E76 C86 T105 H133 Y134 M150 L200 T212 M238 G246 K271 H272 T273 E274 Y295 R314 W387 W455 L465

- Molecule 1: NEURAMINIDASE

Chain B:  81% 16%

MET LEU PRO SER THR ILE GLN THR LEU LEU THR LEU PHE LEU THR LEU PRO GLY VAL LEU LEU SER LEU TYR TYR LEU LEU TYR SER ASP ILE LEU LEU LYS PHE SER PRO THR GLU ILE THR ALA PRO THR MET SER LEU ASP CYS ALA ASN ALA SER ASN VAL

GLN ALA VAL ASN ARG ARG ALA THR LYS LEU VAL VAL PHE LEU LEU LEU PRO GLY E76 P77 Y134 L220 Q224 M288 V302 T313 R314 W387 S401 W455 L465

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

Chain C:  50% 50%

MAG1  
MAG2

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

Chain D:  100%

MAG1  
MAG2

## 4 Data and refinement statistics i

| Property  | Value   | Source           |
|---|---|------------------|
| Space group   | P 32 2 1  | Depositor        |
| Cell constants<br>a, b, c, $\alpha$ , $\beta$ , $\gamma$                | 160.09Å 160.09Å 89.80Å<br>90.00° 90.00° 120.00°             | Depositor        |
| Resolution (Å)  | 45.26 – 1.80<br>45.26 – 1.80                                | Depositor<br>EDS |
| % Data completeness<br>(in resolution range)                            | 99.5 (45.26-1.80)<br>99.5 (45.26-1.80)                      | Depositor<br>EDS |
| $R_{merge}$   | 0.10  | Depositor        |
| $R_{sym}$   | (Not available)   | Depositor        |
| $\langle I/\sigma(I) \rangle$ <sup>1</sup>                              | 2.02 (at 1.79Å)   | Xtriage          |
| Refinement program  | REFMAC 5.8.0046   | Depositor        |
| R, $R_{free}$   | 0.172 , 0.179<br>0.145 , 0.153                              | Depositor<br>DCC |
| $R_{free}$ test set   | 6106 reflections (5.02%)                                    | wwPDB-VP         |
| Wilson B-factor (Å <sup>2</sup> )                                       | 14.4  | Xtriage          |
| Anisotropy  | 0.197   | Xtriage          |
| Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> ) | 0.32 , 22.3   | EDS              |
| L-test for twinning <sup>2</sup>  | $\langle  L  \rangle = 0.39$ , $\langle L^2 \rangle = 0.21$ | Xtriage          |
| Estimated twinning fraction   | 0.316 for -h,-k,l   | Xtriage          |
| Reported twinning fraction  | 0.673 for H, K, L<br>0.327 for -h,-k,l                      | Depositor        |
| Outliers  | 2 of 121682 reflections (0.002%)                            | Xtriage          |
| $F_o, F_c$ correlation  | 0.96  | EDS              |
| Total number of atoms   | 6958  | wwPDB-VP         |
| Average B, all atoms (Å <sup>2</sup> )                                  | 16.0  | wwPDB-VP         |

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.34% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: CA, EDO, G39, NAG

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     | 0.42         | 0/3120  | 0.66        | 0/4217  |
| 1   | B     | 0.43         | 0/3104  | 0.67        | 0/4194  |
| All | All   | 0.43         | 0/6224  | 0.67        | 0/8411  |

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 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     | 3037  | 0        | 2911     | 10      | 0            |
| 1   | B     | 3024  | 0        | 2903     | 4       | 0            |
| 2   | C     | 28    | 0        | 25       | 0       | 0            |
| 2   | D     | 28    | 0        | 25       | 1       | 0            |
| 3   | A     | 1     | 0        | 0        | 0       | 0            |
| 3   | B     | 1     | 0        | 0        | 0       | 0            |
| 4   | A     | 14    | 0        | 13       | 0       | 0            |
| 4   | B     | 14    | 0        | 13       | 0       | 0            |
| 5   | A     | 20    | 0        | 23       | 1       | 0            |
| 5   | B     | 20    | 0        | 23       | 2       | 0            |
| 6   | A     | 44    | 0        | 66       | 0       | 0            |

*Continued on next page...*

Continued from previous page...

| Mol | Chain | Non-H | H(model) | H(added) | Clashes | Symm-Clashes |
|-----|-------|-------|----------|----------|---------|--------------|
| 6   | B     | 24    | 0        | 36       | 0       | 0            |
| 7   | A     | 381   | 0        | 0        | 2       | 11           |
| 7   | B     | 322   | 0        | 0        | 0       | 12           |
| All | All   | 6958  | 0        | 6038     | 18      | 12           |

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 1.

The worst 5 of 18 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:224:GLN:HE21 | 1:B:238:MET:H    | 1.27                     | 0.82              |
| 1:A:246:GLY:O    | 1:A:272:HIS:HD2  | 1.76                     | 0.68              |
| 1:A:86:CYS:HB2   | 7:A:2016:HOH:O   | 1.98                     | 0.62              |
| 1:A:272:HIS:HE1  | 1:A:274:GLU:OE2  | 1.82                     | 0.61              |
| 1:A:133:HIS:HD2  | 1:A:150:ASN:HD21 | 1.55                     | 0.55              |

The worst 5 of 12 symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

| Atom-1         | Atom-2                | Interatomic distance (Å) | Clash overlap (Å) |
|----------------|-----------------------|--------------------------|-------------------|
| 7:A:2123:HOH:O | 7:B:2103:HOH:O[6_555] | 0.95                     | 1.25              |
| 7:A:2080:HOH:O | 7:B:2037:HOH:O[6_555] | 1.03                     | 1.17              |
| 7:A:2105:HOH:O | 7:B:2027:HOH:O[6_555] | 1.03                     | 1.17              |
| 7:A:2149:HOH:O | 7:B:2028:HOH:O[6_555] | 1.07                     | 1.13              |
| 7:A:2126:HOH:O | 7:B:2047:HOH:O[6_555] | 1.12                     | 1.08              |

## 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     | 393/466 (84%) | 383 (98%) | 10 (2%) | 0        | 100         | 100 |
| 1   | B     | 390/466 (84%) | 378 (97%) | 11 (3%) | 1 (0%)   | 41          | 27  |
| All | All   | 783/932 (84%) | 761 (97%) | 21 (3%) | 1 (0%)   | 51          | 36  |

All (1) Ramachandran outliers are listed below:

| Mol | Chain | Res | Type |
|-----|-------|-----|------|
| 1   | B     | 220 | LEU  |

### 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     | 327/390 (84%) | 324 (99%) | 3 (1%)   | 78          | 75 |
| 1   | B     | 325/390 (83%) | 320 (98%) | 5 (2%)   | 65          | 56 |
| All | All   | 652/780 (84%) | 644 (99%) | 8 (1%)   | 71          | 65 |

5 of 8 residues with a non-rotameric sidechain are listed below:

| Mol | Chain | Res | Type |
|-----|-------|-----|------|
| 1   | B     | 455 | TRP  |
| 1   | B     | 401 | SER  |
| 1   | B     | 134 | TYR  |
| 1   | B     | 77  | PRO  |
| 1   | B     | 238 | MET  |

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     | 133 | HIS  |
| 1   | A     | 272 | HIS  |
| 1   | B     | 214 | HIS  |
| 1   | B     | 224 | 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](#)

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   | NAG  | C     | 1   | 2,1  | 14,14,15     | 0.60 | 0        | 17,19,21    | 0.98 | 2 (11%)  |
| 2   | NAG  | C     | 2   | 2    | 14,14,15     | 0.46 | 0        | 17,19,21    | 1.03 | 0        |
| 2   | NAG  | D     | 1   | 2,1  | 14,14,15     | 0.45 | 0        | 17,19,21    | 1.23 | 3 (17%)  |
| 2   | NAG  | D     | 2   | 2    | 14,14,15     | 0.48 | 0        | 17,19,21    | 1.45 | 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. '2' means no outliers of that kind were identified.

| Mol | Type | Chain | Res | Link | Chirals | Torsions  | Rings   |
|-----|------|-------|-----|------|---------|-----------|---------|
| 2   | NAG  | C     | 1   | 2,1  | -       | 0/6/23/26 | 0/1/1/1 |
| 2   | NAG  | C     | 2   | 2    | -       | 0/6/23/26 | 0/1/1/1 |
| 2   | NAG  | D     | 1   | 2,1  | -       | 2/6/23/26 | 0/1/1/1 |
| 2   | NAG  | D     | 2   | 2    | -       | 2/6/23/26 | 0/1/1/1 |

There are no bond length outliers.

The worst 5 of 9 bond angle outliers are listed below:

| Mol | Chain | Res | Type | Atoms    | Z     | Observed(°) | Ideal(°) |
|-----|-------|-----|------|----------|-------|-------------|----------|
| 2   | D     | 2   | NAG  | C8-C7-N2 | 3.45  | 121.94      | 116.10   |
| 2   | D     | 1   | NAG  | C8-C7-N2 | 2.58  | 120.46      | 116.10   |
| 2   | D     | 1   | NAG  | C2-N2-C7 | 2.55  | 126.53      | 122.90   |
| 2   | D     | 2   | NAG  | C1-C2-N2 | 2.50  | 114.77      | 110.49   |
| 2   | D     | 2   | NAG  | O7-C7-C8 | -2.27 | 117.84      | 122.06   |

There are no chirality outliers.

All (4) torsion outliers are listed below:

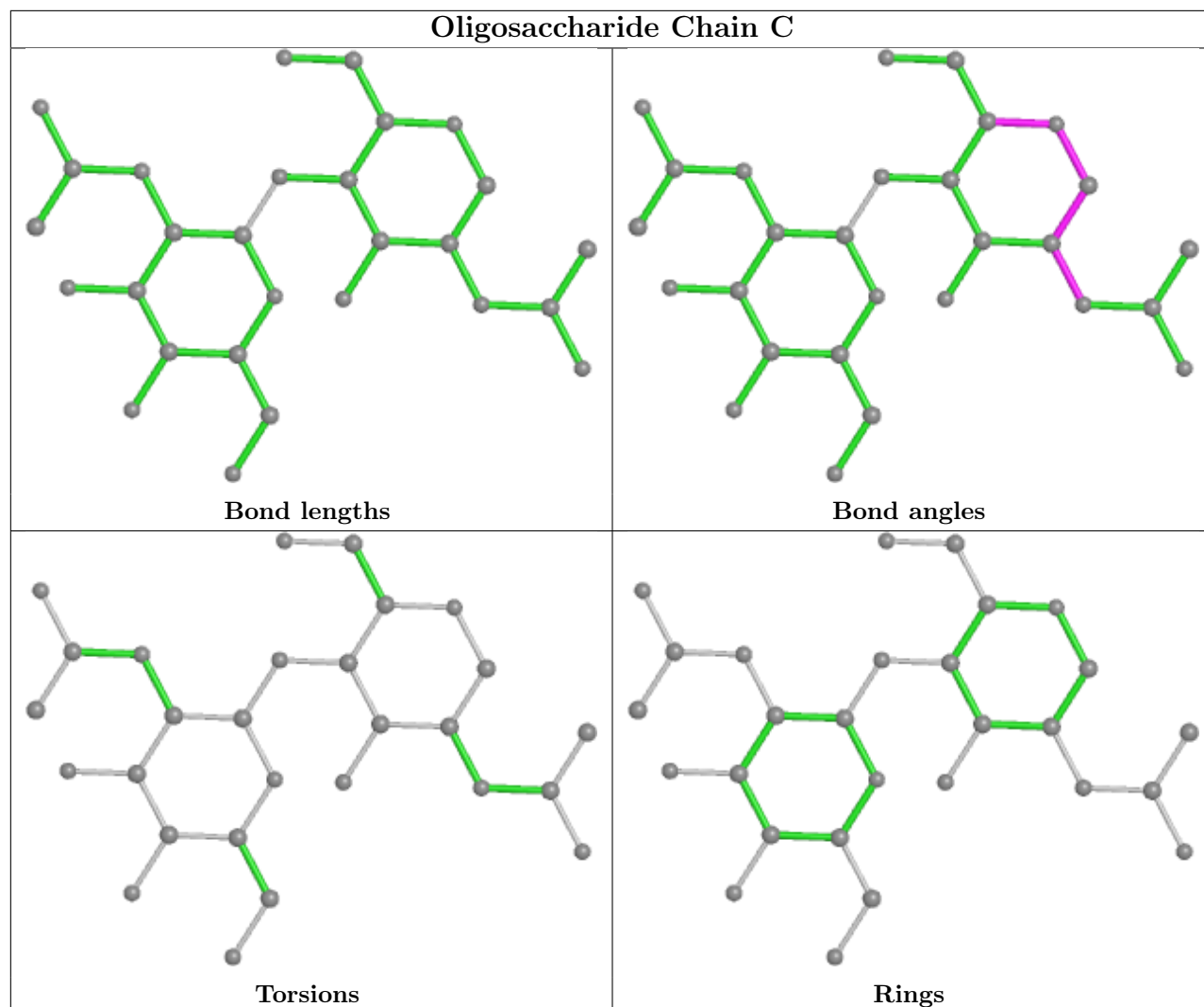
| Mol | Chain | Res | Type | Atoms       |
|-----|-------|-----|------|-------------|
| 2   | D     | 1   | NAG  | C8-C7-N2-C2 |
| 2   | D     | 1   | NAG  | O7-C7-N2-C2 |
| 2   | D     | 2   | NAG  | C8-C7-N2-C2 |
| 2   | D     | 2   | NAG  | O7-C7-N2-C2 |

There are no ring outliers.

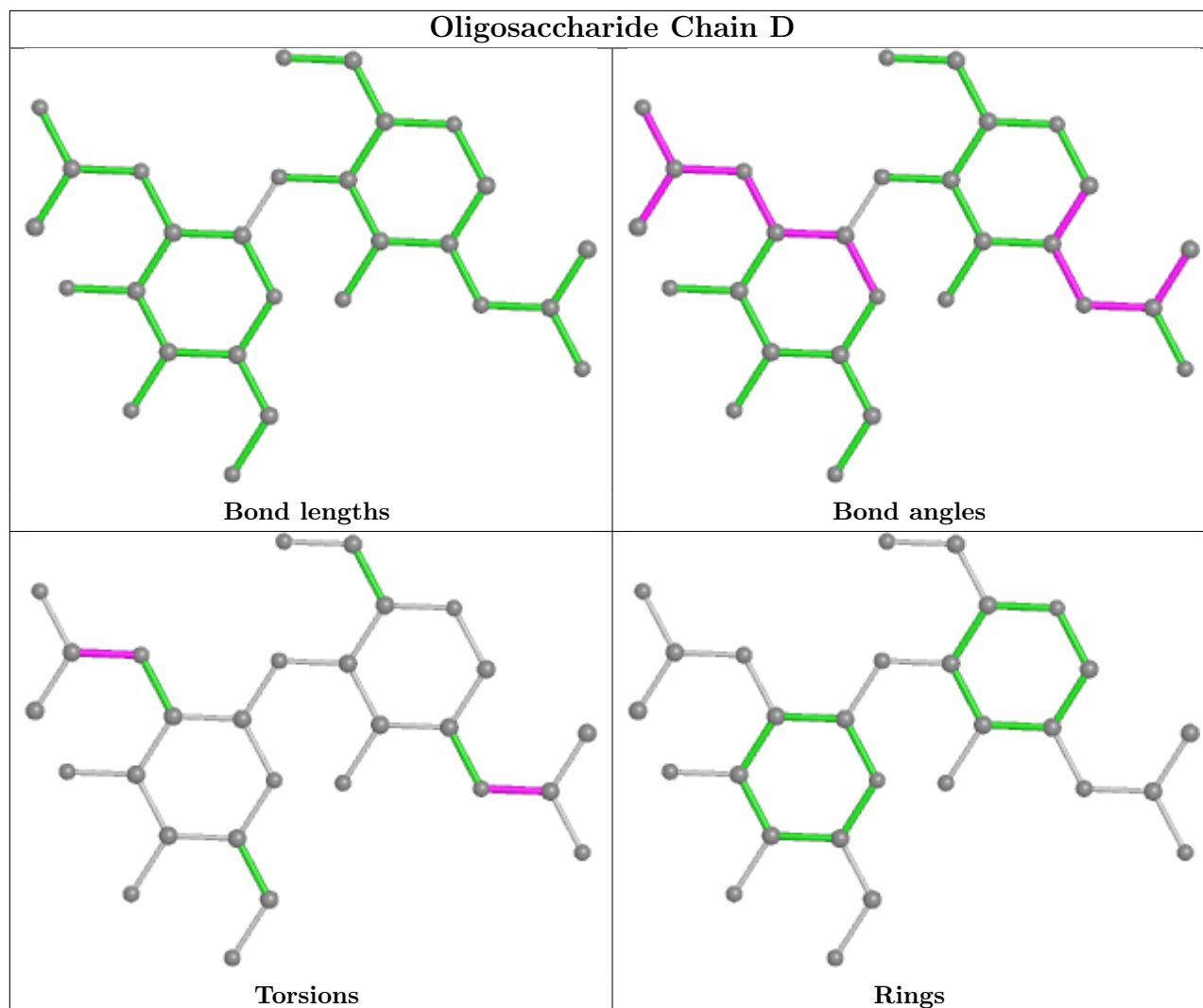
2 monomers are involved in 1 short contact:

| Mol | Chain | Res | Type | Clashes | Symm-Clashes |
|-----|-------|-----|------|---------|--------------|
| 2   | D     | 1   | NAG  | 1       | 0            |
| 2   | D     | 2   | NAG  | 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 [i](#)

Of 23 ligands modelled in this entry, 2 are monoatomic - leaving 21 for Mogul analysis.

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   | EDO  | A     | 1476 | -    | 3,3,3        | 0.45 | 0           | 2,2,2       | 0.37 | 0           |
| 6   | EDO  | A     | 1479 | -    | 3,3,3        | 0.39 | 0           | 2,2,2       | 0.43 | 0           |
| 6   | EDO  | A     | 1481 | -    | 3,3,3        | 0.55 | 0           | 2,2,2       | 0.20 | 0           |
| 6   | EDO  | B     | 1476 | -    | 3,3,3        | 0.50 | 0           | 2,2,2       | 0.19 | 0           |

| Mol | Type | Chain | Res  | Link | Bond lengths |      |          | Bond angles |      |          |
|-----|------|-------|------|------|--------------|------|----------|-------------|------|----------|
|     |      |       |      |      | Counts       | RMSZ | # Z  > 2 | Counts      | RMSZ | # Z  > 2 |
| 5   | G39  | B     | 1470 | -    | 20,20,20     | 0.73 | 0        | 19,27,27    | 0.85 | 0        |
| 6   | EDO  | A     | 1477 | -    | 3,3,3        | 0.49 | 0        | 2,2,2       | 0.14 | 0        |
| 6   | EDO  | B     | 1472 | -    | 3,3,3        | 0.47 | 0        | 2,2,2       | 0.09 | 0        |
| 6   | EDO  | A     | 1475 | -    | 3,3,3        | 0.48 | 0        | 2,2,2       | 0.28 | 0        |
| 6   | EDO  | A     | 1478 | -    | 3,3,3        | 0.36 | 0        | 2,2,2       | 0.45 | 0        |
| 6   | EDO  | A     | 1471 | -    | 3,3,3        | 0.50 | 0        | 2,2,2       | 0.13 | 0        |
| 6   | EDO  | B     | 1473 | -    | 3,3,3        | 0.45 | 0        | 2,2,2       | 0.56 | 0        |
| 6   | EDO  | A     | 1480 | -    | 3,3,3        | 0.41 | 0        | 2,2,2       | 0.39 | 0        |
| 4   | NAG  | A     | 1469 | 1    | 14,14,15     | 0.48 | 0        | 17,19,21    | 0.79 | 0        |
| 6   | EDO  | A     | 1472 | -    | 3,3,3        | 0.33 | 0        | 2,2,2       | 0.49 | 0        |
| 4   | NAG  | B     | 1469 | 1    | 14,14,15     | 0.74 | 0        | 17,19,21    | 1.57 | 2 (11%)  |
| 6   | EDO  | A     | 1474 | -    | 3,3,3        | 0.46 | 0        | 2,2,2       | 0.29 | 0        |
| 5   | G39  | A     | 1470 | -    | 20,20,20     | 0.69 | 0        | 19,27,27    | 1.03 | 0        |
| 6   | EDO  | B     | 1474 | -    | 3,3,3        | 0.65 | 0        | 2,2,2       | 0.19 | 0        |
| 6   | EDO  | A     | 1473 | -    | 3,3,3        | 0.39 | 0        | 2,2,2       | 0.41 | 0        |
| 6   | EDO  | B     | 1471 | -    | 3,3,3        | 0.46 | 0        | 2,2,2       | 0.40 | 0        |
| 6   | EDO  | B     | 1475 | -    | 3,3,3        | 0.56 | 0        | 2,2,2       | 0.22 | 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   |
|-----|------|-------|------|------|---------|------------|---------|
| 6   | EDO  | A     | 1476 | -    | -       | 1/1/1/1    | -       |
| 6   | EDO  | A     | 1479 | -    | -       | 1/1/1/1    | -       |
| 6   | EDO  | A     | 1481 | -    | -       | 1/1/1/1    | -       |
| 6   | EDO  | B     | 1476 | -    | -       | 0/1/1/1    | -       |
| 5   | G39  | B     | 1470 | -    | -       | 2/16/32/32 | 0/1/1/1 |
| 6   | EDO  | A     | 1477 | -    | -       | 0/1/1/1    | -       |
| 6   | EDO  | B     | 1472 | -    | -       | 0/1/1/1    | -       |
| 6   | EDO  | A     | 1475 | -    | -       | 0/1/1/1    | -       |
| 6   | EDO  | A     | 1478 | -    | -       | 0/1/1/1    | -       |
| 6   | EDO  | A     | 1471 | -    | -       | 0/1/1/1    | -       |
| 6   | EDO  | B     | 1473 | -    | -       | 1/1/1/1    | -       |
| 6   | EDO  | A     | 1480 | -    | -       | 1/1/1/1    | -       |
| 4   | NAG  | A     | 1469 | 1    | -       | 0/6/23/26  | 0/1/1/1 |
| 6   | EDO  | A     | 1472 | -    | -       | 0/1/1/1    | -       |
| 4   | NAG  | B     | 1469 | 1    | 1/1/5/7 | 0/6/23/26  | 0/1/1/1 |
| 6   | EDO  | A     | 1474 | -    | -       | 0/1/1/1    | -       |
| 5   | G39  | A     | 1470 | -    | -       | 3/16/32/32 | 0/1/1/1 |
| 6   | EDO  | B     | 1474 | -    | -       | 0/1/1/1    | -       |

*Continued on next page...*

*Continued from previous page...*

| Mol | Type | Chain | Res  | Link | Chirals | Torsions | Rings |
|-----|------|-------|------|------|---------|----------|-------|
| 6   | EDO  | A     | 1473 | -    | -       | 0/1/1/1  | -     |
| 6   | EDO  | B     | 1471 | -    | -       | 1/1/1/1  | -     |
| 6   | EDO  | B     | 1475 | -    | -       | 1/1/1/1  | -     |

There are no bond length outliers.

All (2) bond angle outliers are listed below:

| Mol | Chain | Res  | Type | Atoms    | Z     | Observed(°) | Ideal(°) |
|-----|-------|------|------|----------|-------|-------------|----------|
| 4   | B     | 1469 | NAG  | O5-C1-C2 | -4.96 | 103.45      | 111.29   |
| 4   | B     | 1469 | NAG  | C1-C2-N2 | 2.30  | 114.41      | 110.49   |

All (1) chirality outliers are listed below:

| Mol | Chain | Res  | Type | Atom |
|-----|-------|------|------|------|
| 4   | B     | 1469 | NAG  | C1   |

5 of 12 torsion outliers are listed below:

| Mol | Chain | Res  | Type | Atoms         |
|-----|-------|------|------|---------------|
| 5   | A     | 1470 | G39  | O7-C8-C81-C82 |
| 5   | B     | 1470 | G39  | O7-C8-C81-C82 |
| 6   | A     | 1481 | EDO  | O1-C1-C2-O2   |
| 6   | B     | 1471 | EDO  | O1-C1-C2-O2   |
| 5   | A     | 1470 | G39  | C9-C8-C81-C82 |

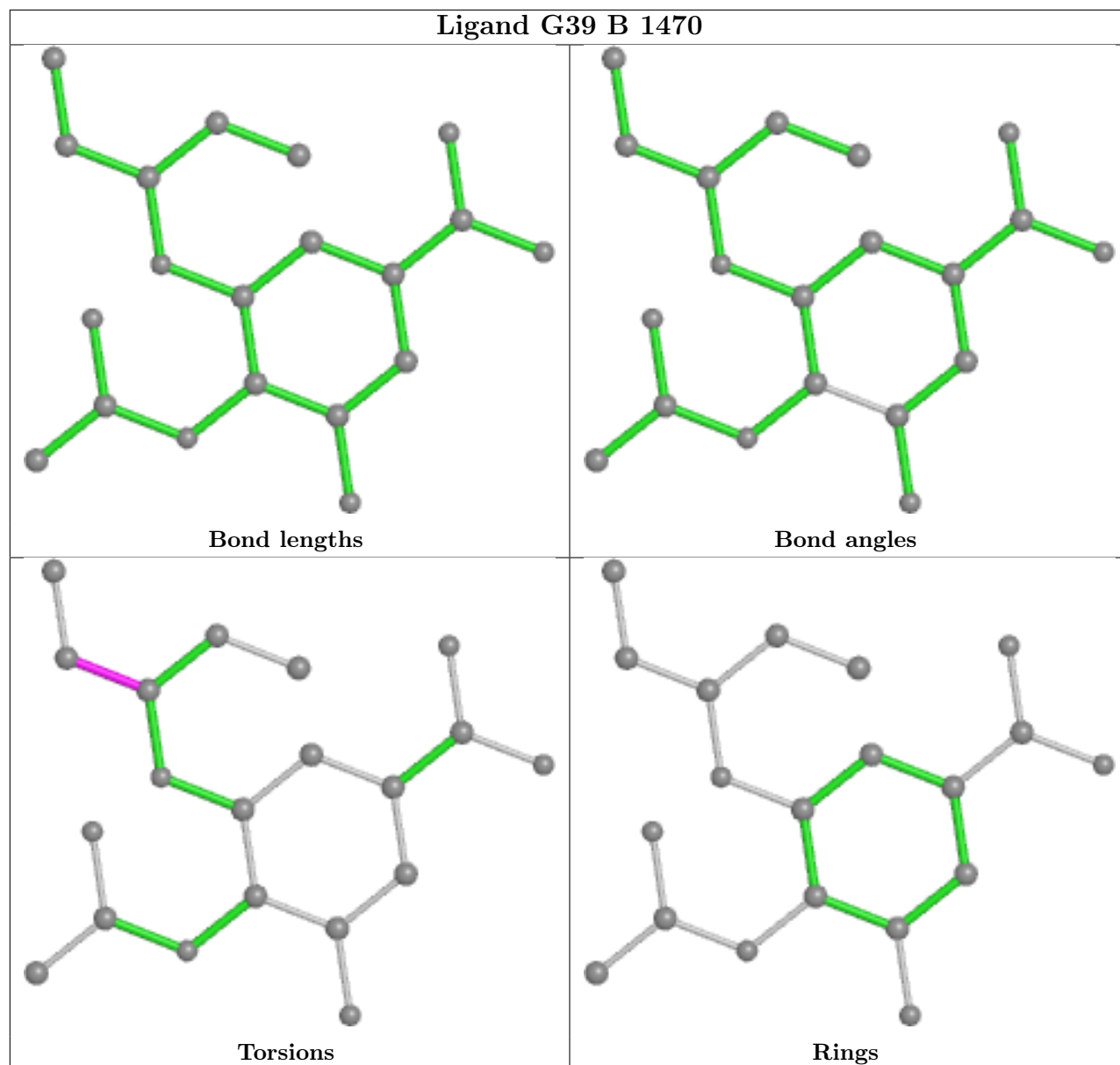
There are no ring outliers.

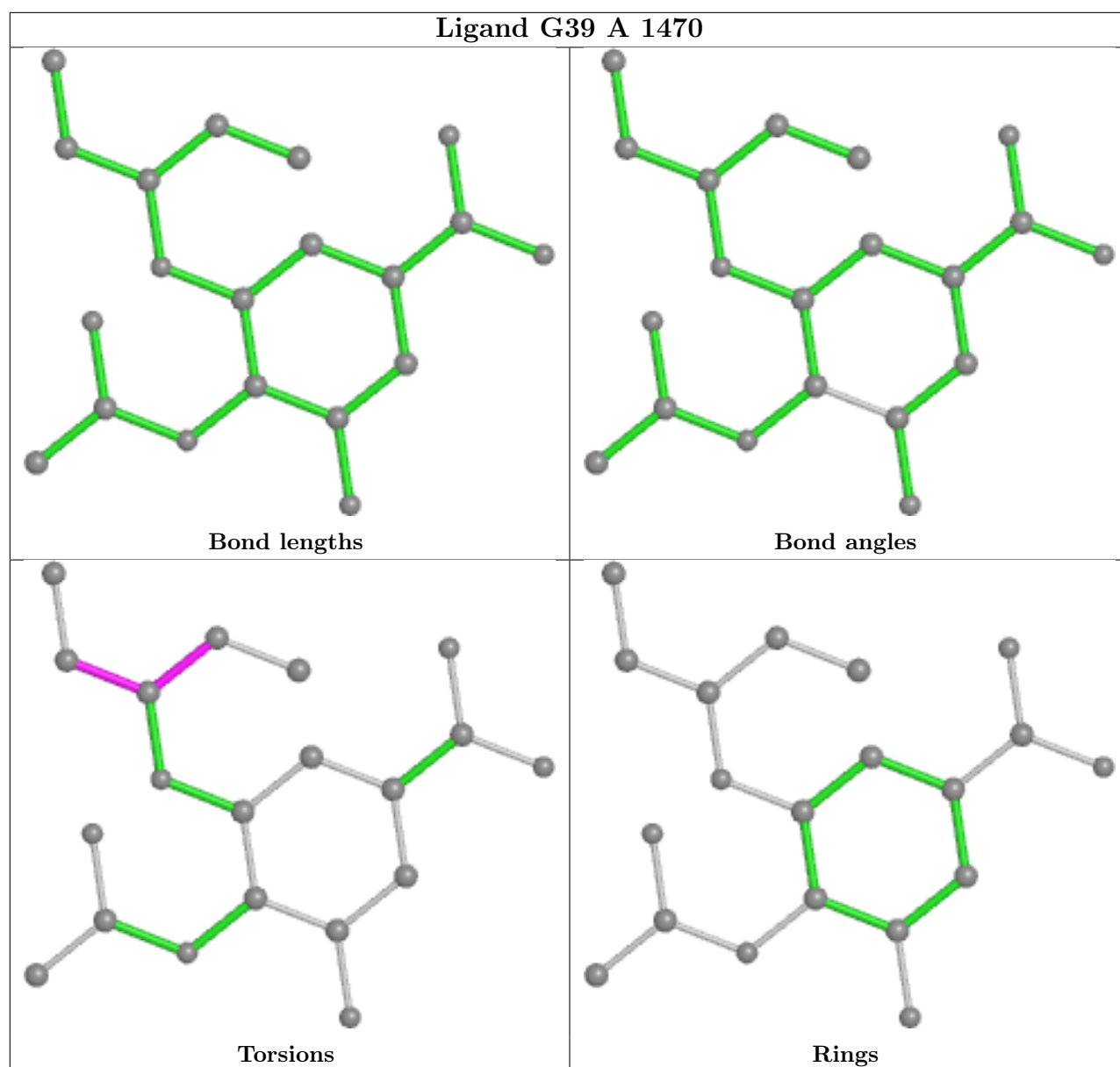
2 monomers are involved in 3 short contacts:

| Mol | Chain | Res  | Type | Clashes | Symm-Clashes |
|-----|-------|------|------|---------|--------------|
| 5   | B     | 1470 | G39  | 2       | 0            |
| 5   | A     | 1470 | G39  | 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 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 than 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.





## 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, 95<sup>th</sup> 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 <sup>2</sup> ) | Q<0.9 |
|-----|-------|---------------|--------|--------------|-----------------------|-------|
| 1   | A     | 390/466 (83%) | -0.82  | 1 (0%) 94 92 | 10, 15, 22, 54        | 0     |
| 1   | B     | 390/466 (83%) | -0.82  | 1 (0%) 94 92 | 10, 15, 22, 52        | 0     |
| All | All   | 780/932 (83%) | -0.82  | 2 (0%) 94 92 | 10, 15, 22, 54        | 0     |

All (2) RSRZ outliers are listed below:

| Mol | Chain | Res | Type | RSRZ |
|-----|-------|-----|------|------|
| 1   | A     | 76  | GLU  | 3.7  |
| 1   | B     | 76  | GLU  | 3.6  |

### 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](#)

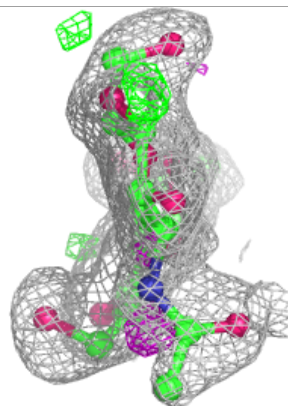
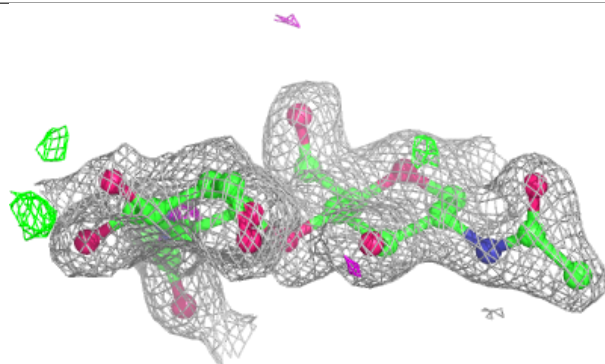
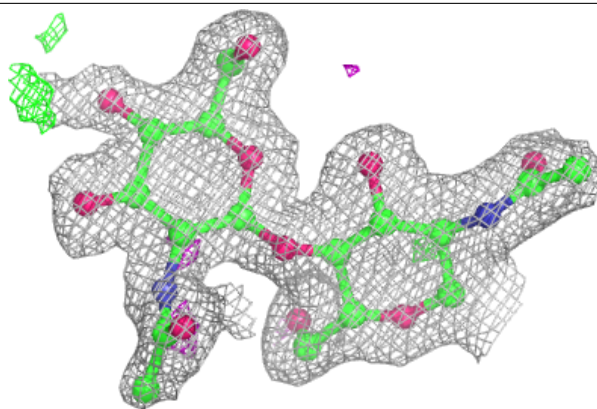
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q< 0.9’ lists the number of atoms with occupancy less than 0.9.

| Mol | Type | Chain | Res | Atoms | RSCC | RSR  | B-factors(Å <sup>2</sup> ) | Q<0.9 |
|-----|------|-------|-----|-------|------|------|----------------------------|-------|
| 2   | NAG  | D     | 2   | 14/15 | 0.84 | 0.23 | 35,39,40,41                | 0     |
| 2   | NAG  | C     | 2   | 14/15 | 0.85 | 0.21 | 37,39,41,43                | 0     |
| 2   | NAG  | D     | 1   | 14/15 | 0.92 | 0.10 | 28,30,31,33                | 0     |
| 2   | NAG  | C     | 1   | 14/15 | 0.95 | 0.14 | 27,29,31,34                | 0     |

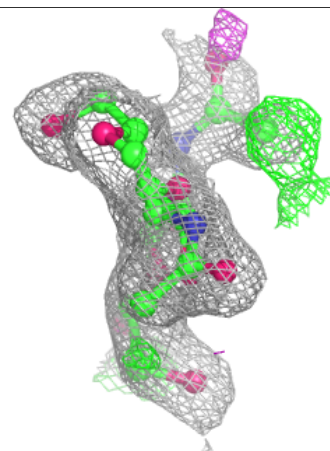
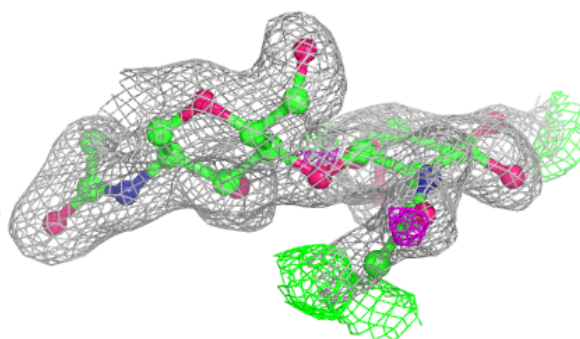
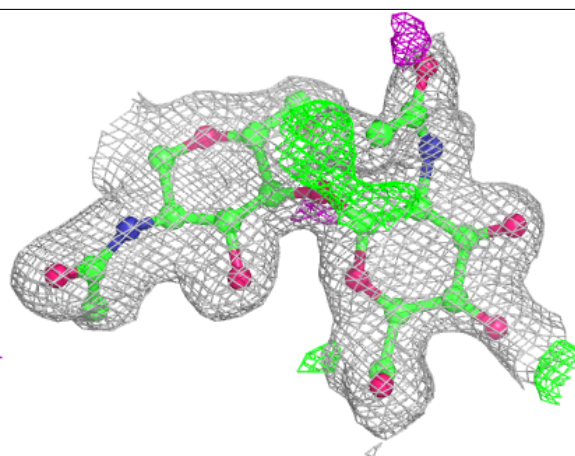
The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.

**Electron density around Chain C:**

$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (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 [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, 95<sup>th</sup> 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.

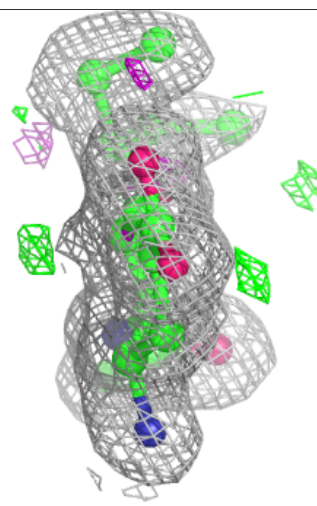
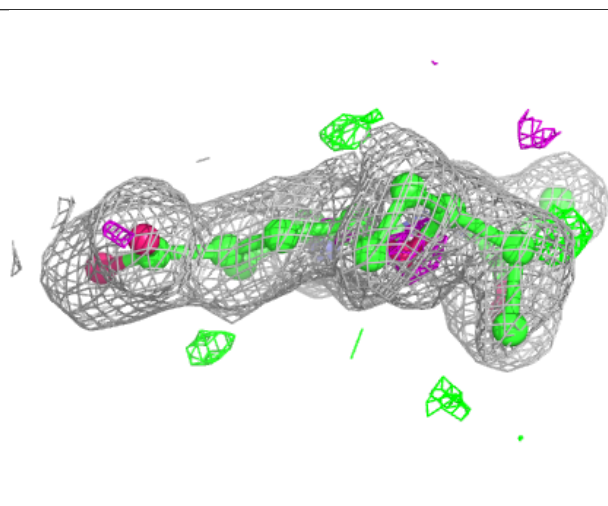
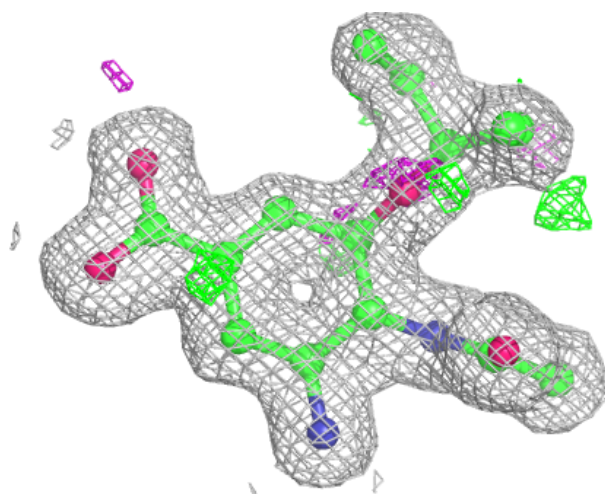
| Mol | Type | Chain | Res  | Atoms | RSCC | RSR  | B-factors(Å <sup>2</sup> ) | Q<0.9 |
|-----|------|-------|------|-------|------|------|----------------------------|-------|
| 4   | NAG  | B     | 1469 | 14/15 | 0.77 | 0.29 | 40,48,51,51                | 0     |
| 6   | EDO  | B     | 1473 | 4/4   | 0.78 | 0.22 | 36,37,37,37                | 0     |
| 6   | EDO  | A     | 1475 | 4/4   | 0.79 | 0.23 | 35,36,36,37                | 0     |
| 4   | NAG  | A     | 1469 | 14/15 | 0.79 | 0.20 | 31,37,39,39                | 0     |
| 6   | EDO  | A     | 1479 | 4/4   | 0.86 | 0.21 | 32,32,33,33                | 0     |
| 6   | EDO  | A     | 1481 | 4/4   | 0.88 | 0.14 | 29,30,30,32                | 0     |
| 6   | EDO  | B     | 1474 | 4/4   | 0.90 | 0.15 | 21,25,26,27                | 0     |
| 6   | EDO  | B     | 1476 | 4/4   | 0.90 | 0.17 | 25,28,29,30                | 0     |
| 6   | EDO  | A     | 1476 | 4/4   | 0.91 | 0.17 | 40,40,40,41                | 0     |
| 6   | EDO  | A     | 1471 | 4/4   | 0.92 | 0.10 | 25,25,26,26                | 0     |
| 6   | EDO  | A     | 1478 | 4/4   | 0.92 | 0.12 | 36,36,36,37                | 0     |
| 6   | EDO  | B     | 1471 | 4/4   | 0.92 | 0.15 | 25,27,27,30                | 0     |
| 6   | EDO  | B     | 1475 | 4/4   | 0.93 | 0.15 | 27,27,27,28                | 0     |
| 6   | EDO  | A     | 1480 | 4/4   | 0.93 | 0.16 | 21,26,26,29                | 0     |
| 5   | G39  | A     | 1470 | 20/20 | 0.96 | 0.09 | 14,16,22,23                | 0     |
| 5   | G39  | B     | 1470 | 20/20 | 0.97 | 0.09 | 14,18,23,24                | 0     |
| 6   | EDO  | A     | 1473 | 4/4   | 0.97 | 0.06 | 15,16,16,17                | 0     |
| 6   | EDO  | A     | 1477 | 4/4   | 0.97 | 0.09 | 20,21,22,22                | 0     |
| 6   | EDO  | A     | 1474 | 4/4   | 0.98 | 0.07 | 14,15,16,16                | 0     |
| 6   | EDO  | A     | 1472 | 4/4   | 0.99 | 0.05 | 14,15,15,15                | 0     |
| 6   | EDO  | B     | 1472 | 4/4   | 0.99 | 0.05 | 16,16,16,16                | 0     |
| 3   | CA   | B     | 1466 | 1/1   | 1.00 | 0.05 | 16,16,16,16                | 0     |
| 3   | CA   | A     | 1466 | 1/1   | 1.00 | 0.06 | 13,13,13,13                | 0     |

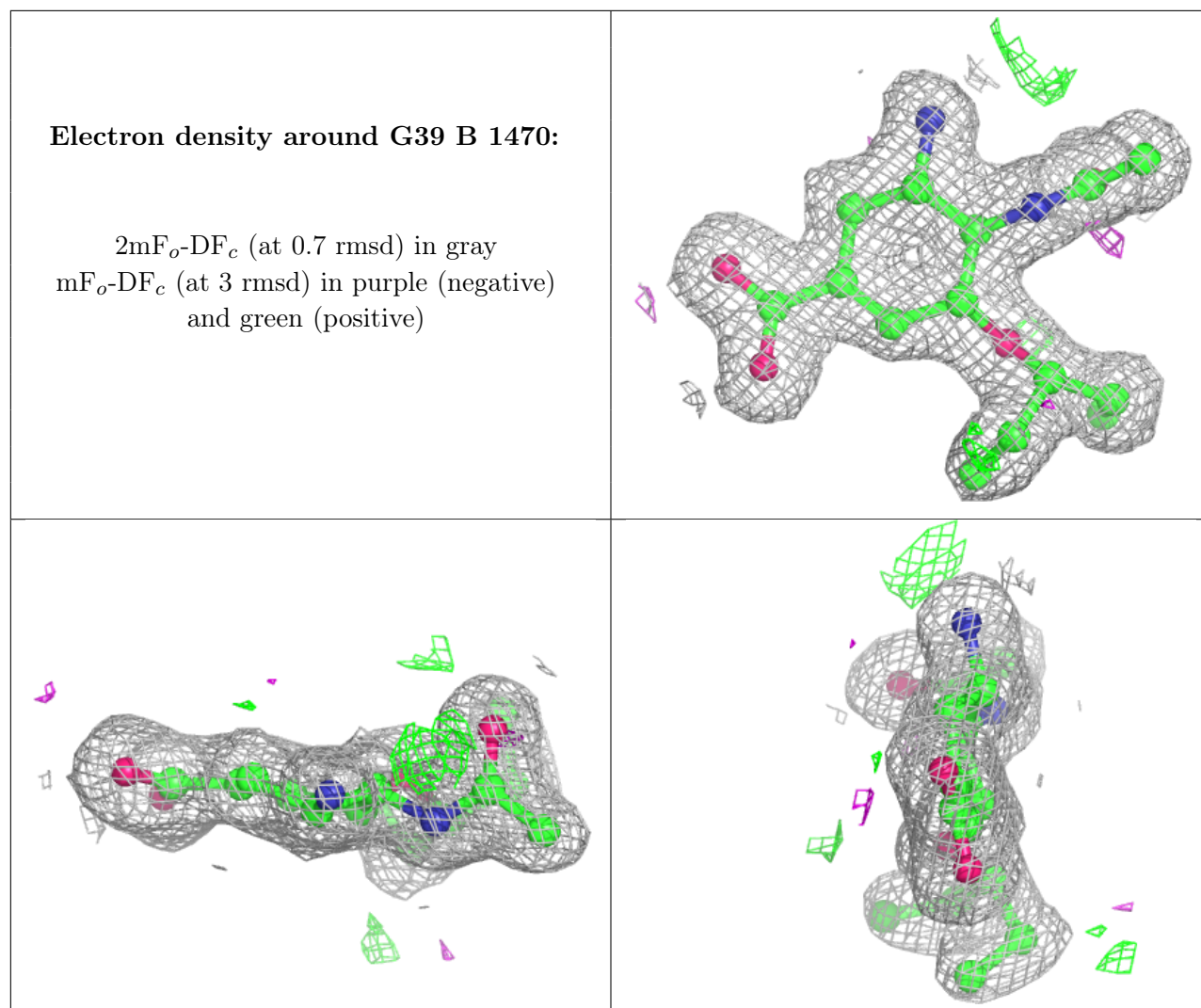
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.



**Electron density around G39 A 1470:**

$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)





## 6.5 Other polymers [i](#)

There are no such residues in this entry.