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

Aug 8, 2020-11:49 PM BST

| PDB ID | : | 1LGC |
| ---: | :--- | :--- |
| Title $:$ | INTERACTION OF A LEGUME LECTIN WITH THE N2 FRAGMENT OF |  |
|  | HUMAN LACTOTRANSFERRIN OR WITH THE ISOLATED BIANTEN- |  |
|  | NARY GLYCOPEPTIDE: ROLE OF THE FUCOSE MOIETY |  |
| Authors $:$ | Bourne, Y.; Cambillau, C. |  |
| Deposited on $:$ | 1994-01-07 |  |
| Resolution $:$ | $2.80 \AA$ (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 following versions of software and data (see references (i)) were used in the production of this report:

```
        MolProbity : 4.02b-467
                            Mogul : 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix) : NOT EXECUTED
                            EDS : NOT EXECUTED
                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) : 2.13.1
```


## 1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:

## X-RAY DIFFRACTION

The reported resolution of this entry is $2.80 \AA$.
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 $(\boldsymbol{\AA})$ ) |
| :---: | :---: | :---: |
| Clashscore | 141614 | $3569(2.80-2.80)$ |
| Ramachandran outliers | 138981 | $3498(2.80-2.80)$ |
| Sidechain outliers | 138945 | $3500(2.80-2.80)$ |

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 on 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 $<=5 \%$

Note EDS was not executed.

| Mol | Chain | Length | Quality of chain |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
| 1 | A | 181 | 57\% | 35\% | 7\% |
| 1 | C | 181 | 61\% | 33\% | 6\% |
| 1 | E | 181 | 48\% | 48\% | . |
| 2 | H | 2 | 50\% | 50\% |  |
| 2 | I | 2 | 50\% | 50\% |  |
| 2 | J | 2 | 50\% | 50\% |  |
| 3 | B | 53 | 55\% | 36\% | 6\% • |
| 3 | D | 53 | 58\% | 28\% | 9\% |

Continued from previous page...

| Mol | Chain | Length | Quality of chain |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 3 | F | 53 |  | $42 \%$ |  |  |  |
| 4 | G | 10 |  | $50 \%$ | $8 \%$ | $11 \%$ |  |
| 5 | K | 9 | $52 \%$ |  | $40 \%$ | $10 \%$ |  |
| 5 | L | 9 |  |  | $44 \%$ |  |  |

## 2 Entry composition (i)

There are 9 unique types of molecules in this entry. The entry contains 6044 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 LEGUME ISOLECTIN II (ALPHA CHAIN).

| Mol | Chain | Residues | Atoms |  |  |  | ZeroOcc | AltConf | Trace |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 1 | A | 181 | Total <br> 1407 | $\begin{gathered} \hline \mathrm{C} \\ 896 \end{gathered}$ | $\begin{gathered} \hline \mathrm{N} \\ 232 \end{gathered}$ | $\begin{gathered} \hline \mathrm{O} \\ 279 \end{gathered}$ | 0 | 0 | 0 |
| 1 | C | 181 | Total 1407 | $\begin{gathered} \mathrm{C} \\ 896 \end{gathered}$ | $\begin{gathered} \mathrm{N} \\ 232 \end{gathered}$ | $\begin{gathered} \mathrm{O} \\ 279 \end{gathered}$ | 0 | 0 | 0 |
| 1 | E | 181 | $\begin{aligned} & \text { Total } \\ & 1397 \end{aligned}$ | $\begin{gathered} \mathrm{C} \\ 889 \end{gathered}$ | $\begin{gathered} \mathrm{N} \\ 231 \end{gathered}$ | $\begin{gathered} \mathrm{O} \\ 277 \end{gathered}$ | 0 | 0 | 0 |

There are 6 discrepancies between the modelled and reference sequences:

| Chain | Residue | Modelled | Actual | Comment | Reference |
| :---: | :---: | :---: | :---: | :---: | :---: |
| A | 16 | PRO | GLN | conflict | UNP P04122 |
| A | 168 | GLY | ALA | conflict | UNP P04122 |
| C | 16 | PRO | GLN | conflict | UNP P04122 |
| C | 168 | GLY | ALA | conflict | UNP P04122 |
| E | 16 | PRO | GLN | conflict | UNP P04122 |
| E | 168 | GLY | ALA | conflict | UNP P04122 |

- Molecule 2 is a protein called DIPEPTIDE.

| Mol | Chain | Residues | Atoms |  |  | ZeroOcc | AltConf | Trace |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 2 | H | 1 | $\begin{array}{c}\text { Total } \\ 9\end{array}$ | $\begin{array}{c}\text { C }\end{array}$ | $\begin{array}{c}\text { N }\end{array}$ | O |  |  |
| 4 | 3 |  |  |  |  |  |  |  |$)$

- Molecule 3 is a protein called LEGUME ISOLECTIN II (BETA CHAIN).

| Mol | Chain | Residues | Atoms |  |  |  | ZeroOcc | AltConf | Trace |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 3 | B | 51 | Total 401 | C | N 61 |  | 0 | 0 | 1 |

Continued on next page...

Continued from previous page...

| Mol | Chain | Residues | Atoms |  |  |  | ZeroOcc | AltConf | Trace |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 3 | D | 48 | $\begin{array}{c}\text { Total } \\ 380\end{array}$ | $\begin{array}{c}\text { C }\end{array}$ | N | O | 58 | 73 | 0 |
| 0 | 1 |  |  |  |  |  |  |  |  |
| 3 | F | 47 | $\begin{array}{c}\text { Total } \\ 371\end{array}$ | $\begin{array}{c}\text { C }\end{array}$ | N | O | 244 | 57 | 70 |$) 0$| 1 |
| :---: |

- Molecule 4 is an oligosaccharide called alpha-L-fucopyranose-(1-3)-[beta-D-galactopyranose-(1-4)]2-acetamido-2-deoxy-beta-D-glucopyranose-(1-2)-alpha-D-mannopyranose-(1-3)-[2-ace tamido-2-deoxy-beta-D-glucopyranose-(1-2)-alpha-D-mannopyranose-(1-6)]beta-D-mannopy ranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-[alpha-L-fucopyranose-(1-6)]2 -acetamido-2-deoxy-beta-D-glucopyranose.


| Mol | Chain | Residues | Atoms |  |  | ZeroOcc | AltConf | Trace |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 4 | G | 10 | $\begin{array}{c}\text { Total } \\ 120\end{array}$ | $\begin{array}{c}\mathrm{C}\end{array}$ | N | O | 4 | 48 |$) 0$| 0 |
| :---: |

- Molecule 5 is an oligosaccharide called N -acetyl-alpha-neuraminic acid-(2-6)-beta-D-galacto pyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-2)-alpha-D-mannopyranose-(1 -3)-[alpha-D-mannopyranose-(1-6)]beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-[alpha-L-fucopyranose-(1-6)]2-acetamido-2-deoxy-beta-D-glucopyran ose.


| Mol | Chain | Residues | Atoms | ZeroOcc | AltConf | Trace |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 5 | K | 9 | Total C N O <br> 116 65 4 47 | 0 | 0 | 0 |
| 5 | L | 9 | $\begin{array}{cccc}\text { Total } & \mathrm{C} & \mathrm{N} & \mathrm{O} \\ 116 & 65 & 4 & 47\end{array}$ | 0 | 0 | 0 |

- Molecule 6 is CALCIUM ION (three-letter code: CA) (formula: Ca).

| Mol | Chain | Residues | Atoms | ZeroOcc | AltConf |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 6 | A | 1 | Total  <br> 1 Ca <br> 1  | 0 | 0 |
| 6 | C | 1 | Total Ca <br> 1 1 | 0 | 0 |

Continued from previous page...

| Mol | Chain | Residues | Atoms | ZeroOcc | AltConf |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 6 | E | 1 | Total <br> 1 | Ca <br> 1 | 0 | 0 |

- Molecule 7 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

| Mol | Chain | Residues | Atoms | ZeroOcc | AltConf |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 7 | A | 1 | Total <br> 1 | Mn <br> 1 | 0 | 0 |
| 7 | C | 1 | Total <br> 1 | Mn <br> 1 | 0 | 0 |
| 7 | E | 1 | Total <br> 1 | Mn <br> 1 | 0 | 0 |

- Molecule 8 is (4S)-2-METHYL-2,4-PENTANEDIOL (three-letter code: MPD) (formula: $\mathrm{C}_{6} \mathrm{H}_{14} \mathrm{O}_{2}$ ).


| Mol | Chain | Residues | Atoms |  | ZeroOcc | AltConf |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 8 | E | 1 | Total   <br> 8 6 C | 2 | 0 | 0 |

- Molecule 9 is water.

| Mol | Chain | Residues | Atoms | ZeroOcc | AltConf |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 9 | A | 74 | Total  <br> 74 O <br> 7 | 0 | 0 |

Continued on next page...

Continued from previous page...

| Mol | Chain | Residues | Atoms | ZeroOcc | AltConf |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 9 | H | 15 | Total <br> 15 | O <br> 15 | 0 | 0 |
| 9 | B | 22 | Total <br> 22 | O <br> 22 | 0 | 0 |
| 9 | C | 74 | Total <br> 74 | O <br> 7 | 0 | 0 |
| 9 | I | 8 | Total <br> 8 | O <br> 8 | 0 | 0 |
| 9 | D | 21 | Total <br> 21 | O <br> 21 | 0 | 0 |
| 9 | E | 48 | Total <br> 48 | O <br> 48 | 0 | 0 |
| 9 | J | 6 | Total <br> 6 | O <br> 6 | 0 | 0 |
| 9 | F | 11 | Total <br> 11 | O <br> 11 | 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. 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.

Note EDS was not executed.

- Molecule 1: LEGUME ISOLECTIN II (ALPHA CHAIN)

- Molecule 1: LEGUME ISOLECTIN II (ALPHA CHAIN)

- Molecule 1: LEGUME ISOLECTIN II (ALPHA CHAIN)

- Molecule 2: DIPEPTIDE

Chain H:

## 管昜

－Molecule 2：DIPEPTIDE
Chain I：

－Molecule 2：DIPEPTIDE

Chain J：
50\％
50\％
戠总
－Molecule 3：LEGUME ISOLECTIN II（BETA CHAIN）

－Molecule 3：LEGUME ISOLECTIN II（BETA CHAIN）

－Molecule 3：LEGUME ISOLECTIN II（BETA CHAIN）


## 象品

－Molecule 4：alpha－L－fucopyranose－（1－3）－［beta－D－galactopyranose－（1－4）］2－acetamido－2－deoxy－bet a－D－glucopyranose－（1－2）－alpha－D－mannopyranose－（1－3）－［2－acetamido－2－deoxy－beta－D－glucopyra nose－（1－2）－alpha－D－mannopyranose－（1－6）］beta－D－mannopyranose－（1－4）－2－acetamido－2－deoxy－bet a－D－glucopyranose－（1－4）－［alpha－L－fucopyranose－（1－6）］2－acetamido－2－deoxy－beta－D－glucopyranos e

| Chain G： | $50 \%$ | $40 \%$ | $10 \%$ |
| :--- | :--- | :--- | :--- |

## 

－Molecule 5：N－acetyl－alpha－neuraminic acid－（2－6）－beta－D－galactopyranose－（1－4）－2－acetamido－2－ deoxy－beta－D－glucopyranose－（1－2）－alpha－D－mannopyranose－（1－3）－［alpha－D－mannopyranose－（1－6）］
beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-[alpha-L-fucopyra nose-(1-6)]2-acetamido-2-deoxy-beta-D-glucopyranose

Chain K: 22\% 44\% 33\%

## 

- Molecule 5: N-acetyl-alpha-neuraminic acid-(2-6)-beta-D-galactopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-2)-alpha-D-mannopyranose-(1-3)-[alpha-D-mannopyranose-(1-6)] beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-[alpha-L-fucopyra nose-(1-6)]2-acetamido-2-deoxy-beta-D-glucopyranose

Chain L: $67 \% \quad 22 \% \quad 11 \%$


## 4 Data and refinement statistics (i

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

| Property | Value | Source |
| :---: | :---: | :---: |
| Space group | P 43 21 2 | Depositor |
| Cell constants <br> $\mathrm{a}, \mathrm{b}, \mathrm{c}, \alpha, \beta, \gamma$ | $117.00 \AA 120.10 \AA$ <br> $90.00^{\circ}$ <br> Resolution $(\AA)$ <br> $90.00^{\circ}$$\quad 90.00^{\circ}$ | Depositor |
| \% Data completeness <br> (in resolution range) | (Not available) $(6.00-00-2.80)$ | Depositor |
| $\mathrm{R}_{\text {merge }}$ | (Not available) | Depositor |
| $\mathrm{R}_{\text {sym }}$ | (Not available) | Depositor |
| Refinement program | X-PLOR | Depositor |
| $\mathrm{R}, \mathrm{R}_{\text {free }}$ | $0.185 \quad, \quad$ (Not available) | Depositor |
| Estimated twinning fraction | No twinning to report. | Xtriage |
| Total number of atoms | 6044 | wwPDB-VP |
| Average B, all atoms $\left(\AA^{2}\right)$ | 41.0 | wwPDB-VP |

## 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: MPD, BMA, NAG, CA, MN, SIA, GAL, FUC, MAN

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z|>5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

| Mol | Chain | Bond lengths |  | Bond angles |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | RMSZ | $\#\|Z\|>5$ | RMSZ | $\#\|Z\|>5$ |
| 1 | A | 0.44 | $0 / 1442$ | 0.76 | $1 / 1967(0.1 \%)$ |
| 1 | C | 0.44 | $0 / 1442$ | 0.74 | $0 / 1967$ |
| 1 | E | 0.44 | $0 / 1431$ | 0.74 | $0 / 1952$ |
| 2 | H | 0.93 | $0 / 8$ | 0.72 | $0 / 8$ |
| 2 | I | 0.74 | $0 / 8$ | 0.66 | $0 / 8$ |
| 2 | J | 0.65 | $0 / 17$ | 0.56 | $0 / 20$ |
| 3 | B | 0.47 | $0 / 413$ | 0.66 | $0 / 566$ |
| 3 | D | 0.50 | $0 / 392$ | 0.70 | $0 / 537$ |
| 3 | F | 0.50 | $0 / 383$ | 0.65 | $0 / 525$ |
| All | All | 0.45 | $0 / 5536$ | 0.73 | $1 / 7550(0.0 \%)$ |

There are no bond length outliers.
All (1) bond angle outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed $\left({ }^{\circ}\right)$ | Ideal $\left({ }^{\circ}\right)$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 1 | A | 154 | LEU | CA-CB-CG | 5.12 | 127.08 | 115.30 |

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 | 1407 | 0 | 1357 | 61 | 0 |

Continued on next page...

Continued from previous page...

| Mol | Chain | Non-H | H(model) | H(added) | Clashes | Symm-Clashes |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 1 | C | 1407 | 0 | 1357 | 51 | 0 |
| 1 | E | 1397 | 0 | 1346 | 83 | 0 |
| 2 | H | 9 | 0 | 4 | 1 | 0 |
| 2 | I | 9 | 0 | 4 | 0 | 0 |
| 2 | J | 18 | 0 | 11 | 2 | 0 |
| 3 | B | 401 | 0 | 377 | 23 | 0 |
| 3 | D | 380 | 0 | 353 | 18 | 0 |
| 3 | F | 371 | 0 | 344 | 41 | 0 |
| 4 | G | 120 | 0 | 103 | 5 | 0 |
| 5 | K | 116 | 0 | 98 | 5 | 0 |
| 5 | L | 116 | 0 | 98 | 3 | 0 |
| 6 | A | 1 | 0 | 0 | 0 | 0 |
| 6 | C | 1 | 0 | 0 | 0 | 0 |
| 6 | E | 1 | 0 | 0 | 0 | 0 |
| 7 | A | 1 | 0 | 0 | 0 | 0 |
| 7 | C | 1 | 0 | 0 | 0 | 0 |
| 7 | E | 1 | 0 | 0 | 0 | 0 |
| 8 | E | 8 | 0 | 14 | 1 | 0 |
| 9 | A | 74 | 0 | 0 | 2 | 0 |
| 9 | B | 22 | 0 | 0 | 1 | 0 |
| 9 | C | 74 | 0 | 0 | 2 | 0 |
| 9 | D | 21 | 0 | 0 | 0 | 0 |
| 9 | E | 48 | 0 | 0 | 4 | 0 |
| 9 | F | 11 | 0 | 0 | 1 | 0 |
| 9 | H | 15 | 0 | 0 | 1 | 0 |
| 9 | I | 8 | 0 | 0 | 0 | 0 |
| 9 | J | 6 | 0 | 0 | 0 | 0 |
| All | All | 6044 | 0 | 5466 | 228 | 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 20 .

The worst 5 of 228 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

| Atom-1 | Atom-2 | Interatomic <br> distance $(\AA)$ | Clash <br> overlap $(\AA)$ |
| :---: | :---: | :---: | :---: |
| 1:C:160:ALA:HB1 | 1:C:177:LEU:HD11 | 1.38 | 1.04 |
| 1:E:54:ASP:HB2 | 3:F:46:LEU:HD21 | 1.62 | 0.82 |
| 1:C:14:ASP:HA | 9:C:306:HOH:O | 1.80 | 0.79 |
| 3:F:16:VAL:HG22 | 3:F:17:PRO:HD2 | 1.66 | 0.77 |
| 1:A:33:LEU:O | 1:A:42:GLY:HA3 | 1.86 | 0.76 |

There are no symmetry-related clashes.

### 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 | $179 / 181(99 \%)$ | $164(92 \%)$ | $14(8 \%)$ | $1(1 \%)$ | 25 | 56 |
| 1 | C | $179 / 181(99 \%)$ | $145(81 \%)$ | $29(16 \%)$ | $5(3 \%)$ | 5 | 17 |
| 1 | E | $179 / 181(99 \%)$ | $147(82 \%)$ | $25(14 \%)$ | $7(4 \%)$ | 3 | 10 |
| 3 | B | $49 / 53(92 \%)$ | $44(90 \%)$ | $5(10 \%)$ | 0 | 100 | 100 |
| 3 | D | $46 / 53(87 \%)$ | $41(89 \%)$ | $5(11 \%)$ | 0 | 100 | 100 |
| 3 | F | $45 / 53(85 \%)$ | $39(87 \%)$ | $6(13 \%)$ | 0 | 100 | 100 |
| All | All | $677 / 702(96 \%)$ | $580(86 \%)$ | $84(12 \%)$ | $13(2 \%)$ | 8 | 26 |

5 of 13 Ramachandran outliers are listed below:

| Mol | Chain | Res | Type |
| :---: | :---: | :---: | :---: |
| 1 | C | 39 | ASN |
| 1 | C | 134 | ASP |
| 1 | E | 36 | ALA |
| 1 | E | 39 | ASN |
| 1 | A | 101 | 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 | $156 / 156(100 \%)$ | $137(88 \%)$ | $19(12 \%)$ | 5 | 15 |

Continued from previous page...

| Mol | Chain | Analysed | Rotameric | Outliers |  | Percentiles |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 1 | C | $156 / 156(100 \%)$ | $138(88 \%)$ | $18(12 \%)$ | 5 | 17 |  |
| 1 | E | $154 / 156(99 \%)$ | $140(91 \%)$ | $14(9 \%)$ | 9 | 27 |  |
| 2 | H | $1 / 2(50 \%)$ | 0 | $1(100 \%)$ | 0 | 0 |  |
| 2 | I | $1 / 2(50 \%)$ | $1(100 \%)$ | 0 | 100 | 100 |  |
| 2 | J | $2 / 2(100 \%)$ | 0 | $2(100 \%)$ | 0 | 0 |  |
| 3 | B | $44 / 47(94 \%)$ | $39(89 \%)$ | $5(11 \%)$ | 5 | 18 |  |
| 3 | D | $40 / 47(85 \%)$ | $37(92 \%)$ | $3(8 \%)$ | 13 | 37 |  |
| 3 | F | $39 / 47(83 \%)$ | $34(87 \%)$ | $5(13 \%)$ | 4 | 13 |  |
| All | All | $593 / 615(96 \%)$ | $526(89 \%)$ | $67(11 \%)$ | 6 | 18 |  |

5 of 67 residues with a non-rotameric sidechain are listed below:

| Mol | Chain | Res | Type |
| :---: | :---: | :---: | :---: |
| 1 | C | 45 | LEU |
| 1 | C | 118 | VAL |
| 2 | J | 513 | GLN |
| 1 | C | 66 | SER |
| 1 | C | 75 | ASN |

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 16 such sidechains are listed below:

| Mol | Chain | Res | Type |
| :---: | :---: | :---: | :---: |
| 1 | C | 59 | ASN |
| 1 | C | 75 | ASN |
| 1 | E | 78 | ASN |
| 1 | C | 51 | HIS |
| 1 | E | 181 | ASN |

### 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)

28 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 | $\#\|Z\|>2$ |
| 4 | NAG | G | 1 | 2,4 | 14,14,15 | 0.58 | 0 | 17,19,21 | 0.60 | 0 |
| 4 | FUC | G | 10 | 4 | 10,10,11 | 0.47 | 0 | 14,14,16 | 0.56 | 0 |
| 4 | NAG | G | 2 | 4 | 14,14,15 | 0.42 | 0 | 17,19,21 | 0.77 | 0 |
| 4 | BMA | G | 3 | 4 | 11,11,12 | 0.45 | 0 | 15,15,17 | 1.11 | 2 (13\%) |
| 4 | MAN | G | 4 | 4 | 11,11,12 | 0.62 | 0 | 15,15,17 | 0.79 | 0 |
| 4 | NAG | G | 5 | 4 | 14,14,15 | 1.03 | 0 | 17,19,21 | 0.91 | 1 (5\%) |
| 4 | FUC | G | 6 | 4 | 10,10,11 | 0.68 | 0 | 14,14,16 | 1.00 | 0 |
| 4 | GAL | G | 7 | 4 | 11,11,12 | 0.51 | 0 | 15,15,17 | 0.80 | 0 |
| 4 | MAN | G | 8 | 4 | 11,11,12 | 0.68 | 0 | 15,15,17 | 0.70 | 0 |
| 4 | NAG | G | 9 | 4 | 14,14,15 | 0.56 | 0 | 17,19,21 | 0.56 | 0 |
| 5 | NAG | K | 1 | 2,5 | 14,14,15 | 0.58 | 0 | 17,19,21 | 0.76 | 1 (5\%) |
| 5 | NAG | K | 2 | 5 | 14,14,15 | 0.53 | 0 | 17,19,21 | 1.20 | 2 (11\%) |
| 5 | BMA | K | 3 | 5 | 11,11,12 | 0.69 | 0 | 15,15,17 | 0.60 | 0 |
| 5 | MAN | K | 4 | 5 | 11,11,12 | 0.39 | 0 | 15,15,17 | 0.82 | 0 |
| 5 | NAG | K | 5 | 5 | 14,14,15 | 0.56 | 0 | 17,19,21 | 0.56 | 0 |
| 5 | GAL | K | 6 | 5 | 11,11,12 | 0.44 | 0 | 15,15,17 | 0.55 | 0 |
| 5 | SIA | K | 7 | 5 | 17,20,21 | 0.71 | 0 | 21,28,31 | 0.77 | 0 |
| 5 | MAN | K | 8 | 5 | 11,11,12 | 0.59 | 0 | 15,15,17 | 1.06 | 1 (6\%) |
| 5 | FUC | K | 9 | 5 | 10,10,11 | 0.36 | 0 | 14,14,16 | 0.57 | 0 |
| 5 | NAG | L | 1 | 2,5 | 14,14,15 | 0.56 | 0 | 17,19,21 | 0.88 | 0 |
| 5 | NAG | L | 2 | 5 | 14,14,15 | 0.64 | 0 | 17,19,21 | 1.04 | 3 (17\%) |
| 5 | BMA | L | 3 | 5 | 11,11,12 | 0.61 | 0 | 15,15,17 | 0.51 | 0 |
| 5 | MAN | L | 4 | 5 | 11,11,12 | 0.56 | 0 | 15,15,17 | 0.79 | 0 |
| 5 | NAG | L | 5 | 5 | 14,14,15 | 0.45 | 0 | 17,19,21 | 0.54 | 0 |
| 5 | GAL | L | 6 | 5 | 11,11,12 | 0.38 | 0 | 15,15,17 | 0.77 | 0 |
| 5 | SIA | L | 7 | 5 | 17,20,21 | 0.67 | 0 | 21,28,31 | 0.76 | 1 (4\%) |
| 5 | MAN | L | 8 | 5 | 11,11,12 | 0.54 | 0 | 15,15,17 | 0.53 | 0 |
| 5 | FUC | L | 9 | 5 | 10,10,11 | 0.30 | 0 | 14,14,16 | 0.39 | 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 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 4 | NAG | G | 1 | 2,4 | - | $0 / 6 / 23 / 26$ | $0 / 1 / 1 / 1$ |
| 4 | FUC | G | 10 | 4 | - | - | $0 / 1 / 1 / 1$ |
| 4 | NAG | G | 2 | 4 | - | $2 / 6 / 23 / 26$ | $0 / 1 / 1 / 1$ |
| 4 | BMA | G | 3 | 4 | - | $0 / 2 / 19 / 22$ | $0 / 1 / 1 / 1$ |
| 4 | MAN | G | 4 | 4 | - | $1 / 2 / 19 / 22$ | $0 / 1 / 1 / 1$ |
| 4 | NAG | G | 5 | 4 | - | $0 / 6 / 23 / 26$ | $0 / 1 / 1 / 1$ |
| 4 | FUC | G | 6 | 4 | - | - | $0 / 1 / 1 / 1$ |
| 4 | GAL | G | 7 | 4 | - | $2 / 2 / 19 / 22$ | $0 / 1 / 1 / 1$ |
| 4 | MAN | G | 8 | 4 | - | $0 / 2 / 19 / 22$ | $0 / 1 / 1 / 1$ |
| 4 | NAG | G | 9 | 4 | - | $3 / 6 / 23 / 26$ | $0 / 1 / 1 / 1$ |
| 5 | NAG | K | 1 | 2,5 | - | $0 / 6 / 23 / 26$ | $0 / 1 / 1 / 1$ |
| 5 | NAG | K | 2 | 5 | - | $1 / 6 / 23 / 26$ | $0 / 1 / 1 / 1$ |
| 5 | BMA | K | 3 | 5 | - | $2 / 2 / 19 / 22$ | $0 / 1 / 1 / 1$ |
| 5 | MAN | K | 4 | 5 | - | $2 / 2 / 19 / 22$ | $0 / 1 / 1 / 1$ |
| 5 | NAG | K | 5 | 5 | - | $0 / 6 / 23 / 26$ | $0 / 1 / 1 / 1$ |
| 5 | GAL | K | 6 | 5 | - | $1 / 2 / 19 / 22$ | $0 / 1 / 1 / 1$ |
| 5 | SIA | K | 7 | 5 | - | $2 / 14 / 34 / 38$ | $0 / 1 / 1 / 1$ |
| 5 | MAN | K | 8 | 5 | - | $0 / 2 / 19 / 22$ | $0 / 1 / 1 / 1$ |
| 5 | FUC | K | 9 | 5 | - | - | $0 / 1 / 1 / 1$ |
| 5 | NAG | L | 1 | 2,5 | - | $0 / 6 / 23 / 26$ | $0 / 1 / 1 / 1$ |
| 5 | NAG | L | 2 | 5 | - | $1 / 6 / 23 / 26$ | $0 / 1 / 1 / 1$ |
| 5 | BMA | L | 3 | 5 | - | $2 / 2 / 19 / 22$ | $0 / 1 / 1 / 1$ |
| 5 | MAN | L | 4 | 5 | - | $1 / 2 / 19 / 22$ | $0 / 1 / 1 / 1$ |
| 5 | NAG | L | 5 | 5 | - | $1 / 6 / 23 / 26$ | $0 / 1 / 1 / 1$ |
| 5 | GAL | L | 6 | 5 | - | $2 / 2 / 19 / 22$ | $0 / 1 / 1 / 1$ |
| 5 | SIA | L | 7 | 5 | - | $5 / 14 / 34 / 38$ | $0 / 1 / 1 / 1$ |
| 5 | MAN | L | 8 | 5 | - | $0 / 2 / 19 / 22$ | $0 / 1 / 1 / 1$ |
| 5 | FUC | L | 9 | 5 | - | - | $0 / 1 / 1 / 1$ |

There are no bond length outliers.
The worst 5 of 11 bond angle outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed $\left({ }^{\circ}\right)$ | Ideal $\left({ }^{\circ}\right)$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 5 | K | 2 | NAG | C4-C3-C2 | -3.69 | 105.61 | 111.02 |
| 4 | G | 3 | BMA | C6-C5-C4 | -3.07 | 105.81 | 113.00 |
| 5 | K | 8 | MAN | C1-O5-C5 | 2.95 | 116.19 | 112.19 |
| 5 | L | 2 | NAG | C4-C3-C2 | -2.40 | 107.50 | 111.02 |

Continued on next page...

Continued from previous page...

| Mol | Chain | Res | Type | Atoms | Z | Observed $\left({ }^{\circ}\right)$ | Ideal $\left({ }^{\circ}\right)$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 4 | G | 5 | NAG | C6-C5-C4 | 2.37 | 118.55 | 113.00 |

There are no chirality outliers.
5 of 28 torsion outliers are listed below:

| Mol | Chain | Res | Type | Atoms |
| :---: | :---: | :---: | :---: | :---: |
| 5 | L | 7 | SIA | C5-C6-C7-O7 |
| 5 | L | 7 | SIA | O6-C6-C7-O7 |
| 5 | K | 7 | SIA | C7-C8-C9-O9 |
| 5 | K | 7 | SIA | O8-C8-C9-O9 |
| 5 | L | 2 | NAG | C1-C2-N2-C7 |

There are no ring outliers.
13 monomers are involved in 13 short contacts:

| Mol | Chain | Res | Type | Clashes | Symm-Clashes |
| :---: | :---: | :---: | :---: | :---: | :---: |
| 5 | K | 1 | NAG | 1 | 0 |
| 5 | L | 1 | NAG | 2 | 0 |
| 5 | K | 4 | MAN | 1 | 0 |
| 5 | K | 2 | NAG | 1 | 0 |
| 4 | G | 6 | FUC | 3 | 0 |
| 5 | K | 3 | BMA | 2 | 0 |
| 5 | L | 2 | NAG | 1 | 0 |
| 4 | G | 5 | NAG | 3 | 0 |
| 4 | G | 8 | MAN | 1 | 0 |
| 5 | K | 5 | NAG | 1 | 0 |
| 5 | K | 8 | MAN | 2 | 0 |
| 5 | K | 7 | SIA | 1 | 0 |
| 4 | G | 7 | GAL | 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.
Onanancharide Chain


### 5.6 Ligand geometry (i)

Of 7 ligands modelled in this entry, 6 are monoatomic - leaving 1 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 | $\#\|Z\|>2$ |
| 8 | MPD | E | 303 | - | $7,7,7$ | 0.52 | 0 | $9,10,10$ | 0.38 | 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 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 8 | MPD | E | 303 | - | - | $1 / 5 / 5 / 5$ | - |

There are no bond length outliers.
There are no bond angle outliers.
There are no chirality outliers.
All (1) torsion outliers are listed below:

| Mol | Chain | Res | Type | Atoms |
| :---: | :---: | :---: | :---: | :---: |
| 8 | E | 303 | MPD | C2-C3-C4-C5 |

There are no ring outliers.
1 monomer is involved in 1 short contact:

| Mol | Chain | Res | Type | Clashes | Symm-Clashes |
| :---: | :---: | :---: | :---: | :---: | :---: |
| 8 | E | 303 | MPD | 1 | 0 |

### 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.

PROTEIN DATA BANK

## 6 Fit of model and data (i)

### 6.1 Protein, DNA and RNA chains (i)

EDS was not executed - this section is therefore empty.
6.2 Non-standard residues in protein, DNA, RNA chains (i)

EDS was not executed - this section is therefore empty.

### 6.3 Carbohydrates (i)

EDS was not executed - this section is therefore empty.

### 6.4 Ligands (i)

EDS was not executed - this section is therefore empty.

### 6.5 Other polymers (i)

EDS was not executed - this section is therefore empty.

