



# Full wwPDB NMR Structure Validation Report ⓘ

Aug 7, 2020 – 01:59 AM BST

PDB ID : 5UE2  
Title : proMMP-7 with heparin octasaccharide bridging between domains  
Authors : Fulcher, Y.G.; Prior, S.H.; Linhardt, R.J.; Van Doren, S.R.  
Deposited on : 2016-12-29

This is a Full wwPDB NMR Structure Validation 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/NMRValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
RCI : v\_1n\_11\_5\_13\_A (Berjanski et al., 2005)  
PANAV : Wang et al. (2010)  
ShiftChecker : 2.13.1  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.13.1

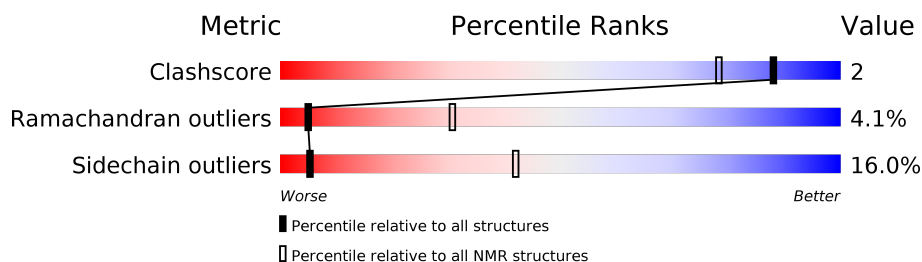
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*SOLUTION NMR*

The overall completeness of chemical shifts assignment is 85%.

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) | NMR archive<br>(#Entries) |
|-----------------------|-----------------------------|---------------------------|
| Clashscore            | 158937                      | 12864                     |
| Ramachandran outliers | 154571                      | 11451                     |
| Sidechain outliers    | 154315                      | 11428                     |

The table below summarises the geometric issues observed across the polymeric chains and their fit to the experimental data. The red, orange, yellow and green segments indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. A cyan segment indicates the fraction of residues that are not part of the well-defined cores, and 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  $\leq 5\%$ .

| Mol | Chain | Length | Quality of chain |
|-----|-------|--------|------------------|
| 1   | A     | 247    |                  |
| 2   | B     | 8      |                  |

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA and RNA chains that are outliers for geometric criteria:

| Mol | Chain | Compound | Res | Total models with violations |          |
|-----|-------|----------|-----|------------------------------|----------|
|     |       |          |     | Chirality                    | Geometry |
| 2   | B     | IDS      | 6   | 1                            | -        |

## 2 Ensemble composition and analysis

This entry contains 16 models. Model 1 is the overall representative, medoid model (most similar to other models).

The following residues are included in the computation of the global validation metrics.

| Well-defined (core) protein residues |   |                   |              |
|--------------------------------------|---|-------------------|--------------|
| Well-defined core                    | Residue range (total)                                     | Backbone RMSD (Å) | Medoid model |
| 1                                    | A:15-A:22, A:34-A:71,<br>A:83-A:216, A:224-A:239<br>(196) | 0.41              | 1            |

Ill-defined regions of proteins are excluded from the global statistics.

Ligands and non-protein polymers are included in the analysis.

The models can be grouped into 3 clusters and 3 single-model clusters were found.

| Cluster number        | Models                       |
|-----------------------|------------------------------|
| 1                     | 1, 2, 4, 5, 7, 8, 12, 14, 15 |
| 2                     | 3, 11                        |
| 3                     | 6, 13                        |
| Single-model clusters | 9; 10; 16                    |

### 3 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 4032 atoms, of which 1945 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Matrilysin.

| Mol | Chain | Residues | Atoms |      |      |     |     |   | Trace |
|-----|-------|----------|-------|------|------|-----|-----|---|-------|
| 1   | A     | 247      | Total | C    | H    | N   | O   | S | 0     |
|     |       |          | 3828  | 1234 | 1885 | 338 | 362 | 9 |       |

There is a discrepancy between the modelled and reference sequences:

| Chain | Residue | Modelled | Actual | Comment             | Reference  |
|-------|---------|----------|--------|---------------------|------------|
| A     | 195     | ALA      | GLU    | engineered mutation | UNP P09237 |

- Molecule 2 is an oligosaccharide called 2-O-sulfo-alpha-L-idopyranuronic acid-(1-4)-2-deoxy-6-O-sulfo-2-(sulfoamino)-alpha-D-glucopyranose-(1-4)-2-O-sulfo-alpha-L-idopyranuronic acid-(1-4)-2-deoxy-6-O-sulfo-2-(sulfoamino)-alpha-D-glucopyranose-(1-4)-2-O-sulfo-alpha-L-idopyranuronic acid-(1-4)-2-deoxy-6-O-sulfo-2-(sulfoamino)-alpha-D-glucopyranose-(1-4)-2-O-sulfo-alpha-L-idopyranuronic acid-(1-4)-2-deoxy-6-O-sulfo-2-(sulfoamino)-alpha-D-glucopyranose.



| Mol | Chain | Residues | Atoms |    |    |   |    |    | Trace |
|-----|-------|----------|-------|----|----|---|----|----|-------|
| 2   | B     | 8        | Total | C  | H  | N | O  | S  | 0     |
|     |       |          | 200   | 48 | 60 | 4 | 76 | 12 |       |

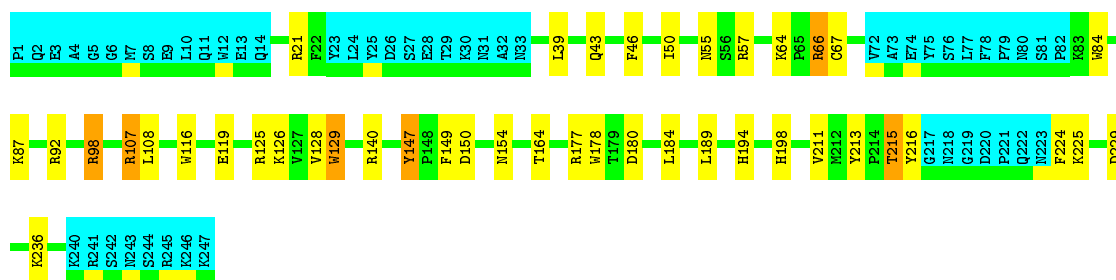
- Molecule 3 is CALCIUM ION (three-letter code: CA) (formula: Ca).

| Mol | Chain | Residues | Atoms |    |
|-----|-------|----------|-------|----|
| 3   | A     | 2        | Total | Ca |
|     |       |          | 2     | 2  |

- Molecule 4 is ZINC ION (three-letter code: ZN) (formula: Zn).

| Mol | Chain | Residues | Atoms |    |
|-----|-------|----------|-------|----|
| 4   | A     | 2        | Total | Zn |
|     |       |          | 2     | 2  |





- Molecule 2: 2-O-sulfo-alpha-L-idopyranuronic acid-(1-4)-2-deoxy-6-O-sulfo-2-(sulfoamino)-alpha-D-glucopyranose-(1-4)-2-O-sulfo-alpha-L-idopyranuronic acid-(1-4)-2-deoxy-6-O-sulfo-2-(sulfoamino)-alpha-D-glucopyranose-(1-4)-2-O-sulfo-alpha-L-idopyranuronic acid-(1-4)-2-deoxy-6-O-sulfo-2-(sulfoamino)-alpha-D-glucopyranose-(1-4)-2-O-sulfo-alpha-L-idopyranuronic acid-(1-4)-2-deoxy-6-O-sulfo-2-(sulfoamino)-alpha-D-glucopyranose

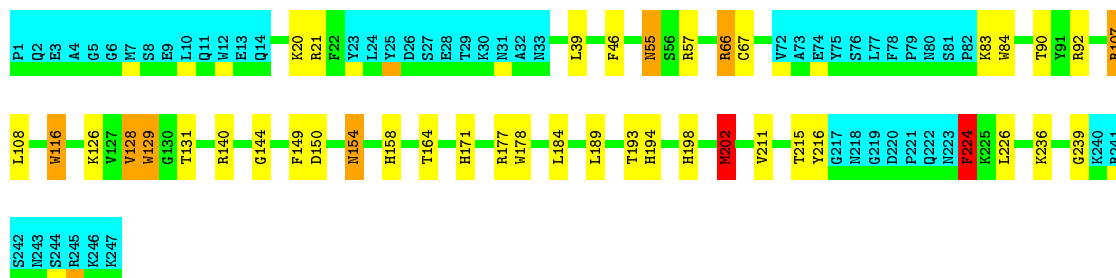
Chain B:  100%



#### 4.2.2 Score per residue for model 2

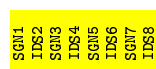
- Molecule 1: Matrilysin

Chain A:  62% 13% 2% 1% 22%



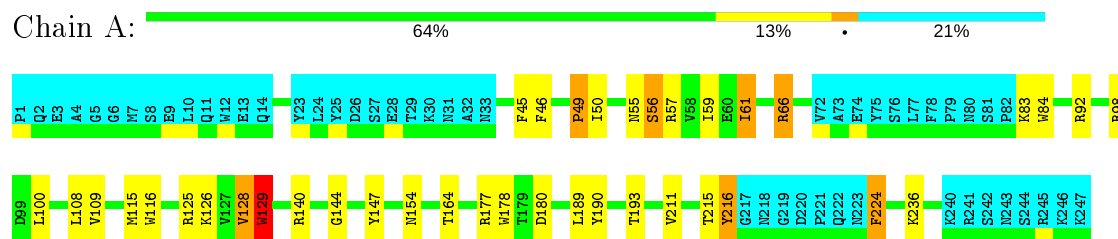
- Molecule 2: 2-O-sulfo-alpha-L-idopyranuronic acid-(1-4)-2-deoxy-6-O-sulfo-2-(sulfoamino)-alpha-D-glucopyranose-(1-4)-2-O-sulfo-alpha-L-idopyranuronic acid-(1-4)-2-deoxy-6-O-sulfo-2-(sulfoamino)-alpha-D-glucopyranose-(1-4)-2-O-sulfo-alpha-L-idopyranuronic acid-(1-4)-2-deoxy-6-O-sulfo-2-(sulfoamino)-alpha-D-glucopyranose-(1-4)-2-O-sulfo-alpha-L-idopyranuronic acid-(1-4)-2-deoxy-6-O-sulfo-2-(sulfoamino)-alpha-D-glucopyranose

Chain B:  100%

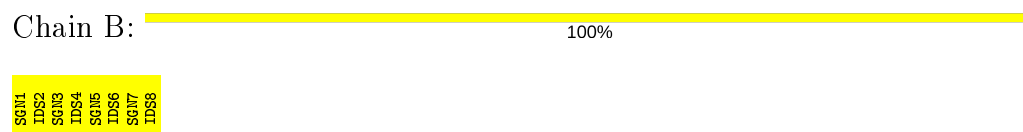


### 4.2.3 Score per residue for model 3

- Molecule 1: Matrilysin

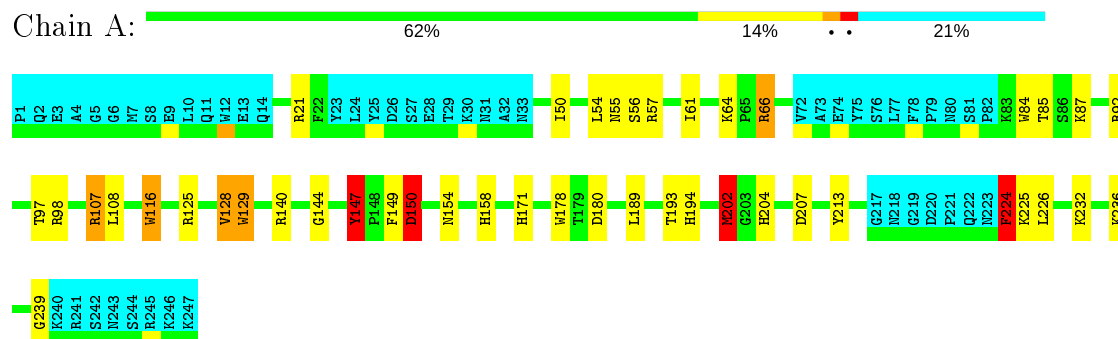


- Molecule 2: 2-O-sulfo-alpha-L-idopyranuronic acid-(1-4)-2-deoxy-6-O-sulfo-2-(sulfoamino)-alpha-D-glucopyranose-(1-4)-2-O-sulfo-alpha-L-idopyranuronic acid-(1-4)-2-deoxy-6-O-sulfo-2-(sulfoamino)-alpha-D-glucopyranose-(1-4)-2-O-sulfo-alpha-L-idopyranuronic acid-(1-4)-2-deoxy-6-O-sulfo-2-(sulfoamino)-alpha-D-glucopyranose-(1-4)-2-O-sulfo-alpha-L-idopyranuronic acid-(1-4)-2-deoxy-6-O-sulfo-2-(sulfoamino)-alpha-D-glucopyranose

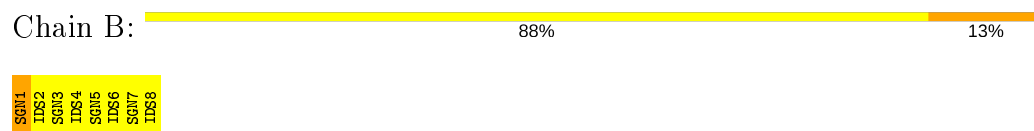


### 4.2.4 Score per residue for model 4

- Molecule 1: Matrilysin

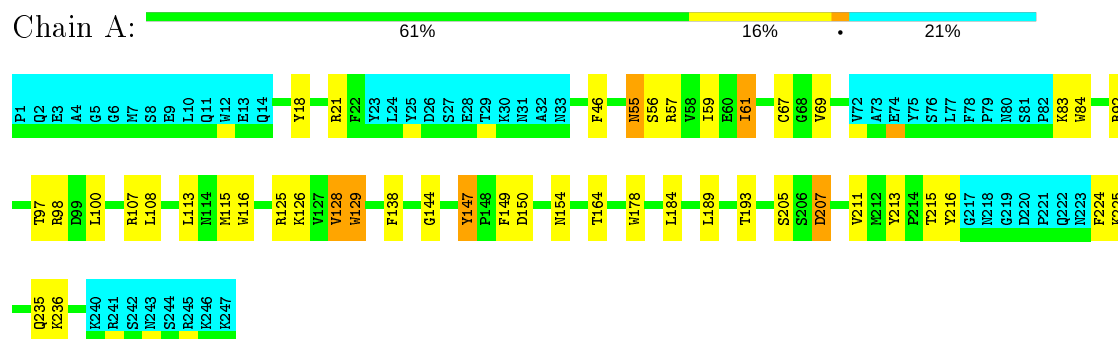


- Molecule 2: 2-O-sulfo-alpha-L-idopyranuronic acid-(1-4)-2-deoxy-6-O-sulfo-2-(sulfoamino)-alpha-D-glucopyranose-(1-4)-2-O-sulfo-alpha-L-idopyranuronic acid-(1-4)-2-deoxy-6-O-sulfo-2-(sulfoamino)-alpha-D-glucopyranose-(1-4)-2-O-sulfo-alpha-L-idopyranuronic acid-(1-4)-2-deoxy-6-O-sulfo-2-(sulfoamino)-alpha-D-glucopyranose-(1-4)-2-O-sulfo-alpha-L-idopyranuronic acid-(1-4)-2-deoxy-6-O-sulfo-2-(sulfoamino)-alpha-D-glucopyranose

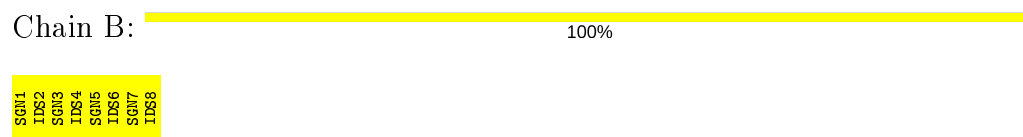


### 4.2.5 Score per residue for model 5

- Molecule 1: Matrilysin

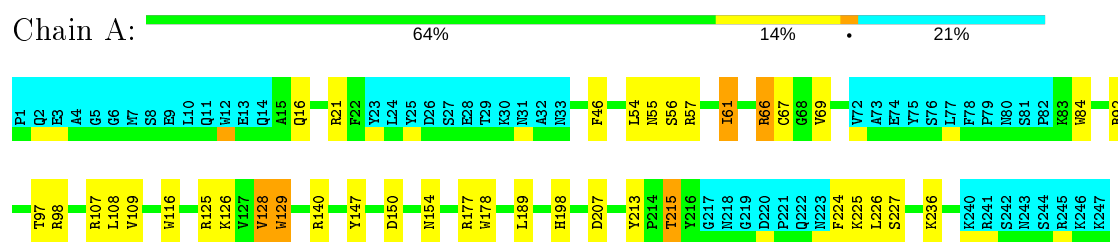


- Molecule 2: 2-O-sulfo-alpha-L-idopyranuronic acid-(1-4)-2-deoxy-6-O-sulfo-2-(sulfoamino)-alpha-D-glucopyranose-(1-4)-2-O-sulfo-alpha-L-idopyranuronic acid-(1-4)-2-deoxy-6-O-sulfo-2-(sulfoamino)-alpha-D-glucopyranose-(1-4)-2-O-sulfo-alpha-L-idopyranuronic acid-(1-4)-2-deoxy-6-O-sulfo-2-(sulfoamino)-alpha-D-glucopyranose-(1-4)-2-O-sulfo-alpha-L-idopyranuronic acid-(1-4)-2-deoxy-6-O-sulfo-2-(sulfoamino)-alpha-D-glucopyranose

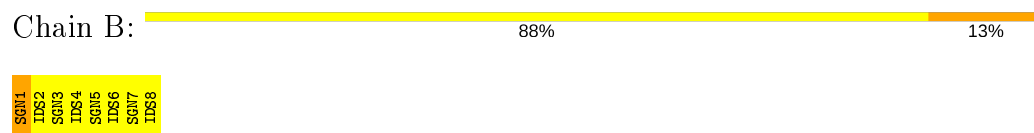


### 4.2.6 Score per residue for model 6

- Molecule 1: Matrilysin



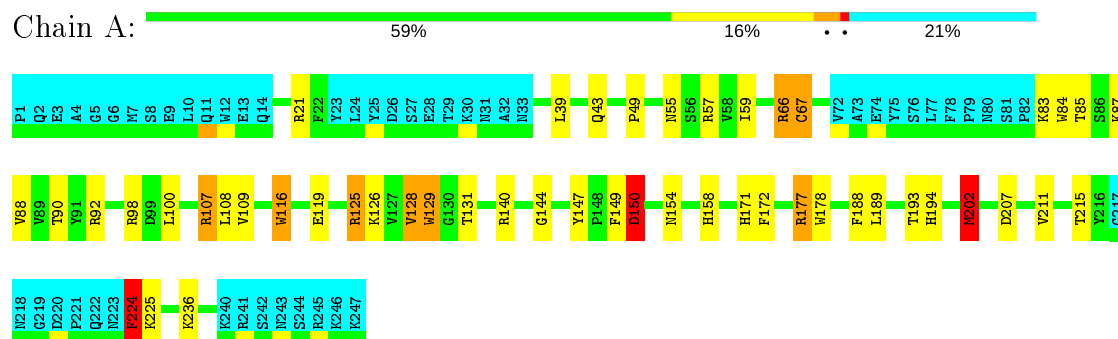
- Molecule 2: 2-O-sulfo-alpha-L-idopyranuronic acid-(1-4)-2-deoxy-6-O-sulfo-2-(sulfoamino)-alpha-D-glucopyranose-(1-4)-2-O-sulfo-alpha-L-idopyranuronic acid-(1-4)-2-deoxy-6-O-sulfo-2-(sulfoamino)-alpha-D-glucopyranose-(1-4)-2-O-sulfo-alpha-L-idopyranuronic acid-(1-4)-2-deoxy-6-O-sulfo-2-(sulfoamino)-alpha-D-glucopyranose-(1-4)-2-O-sulfo-alpha-L-idopyranuronic acid-(1-4)-2-deoxy-6-O-sulfo-2-(sulfoamino)-alpha-D-glucopyranose



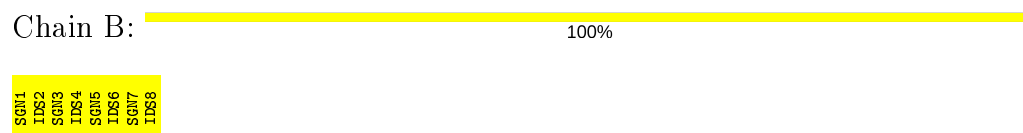


### 4.2.7 Score per residue for model 7

#### • Molecule 1: Matrilysin

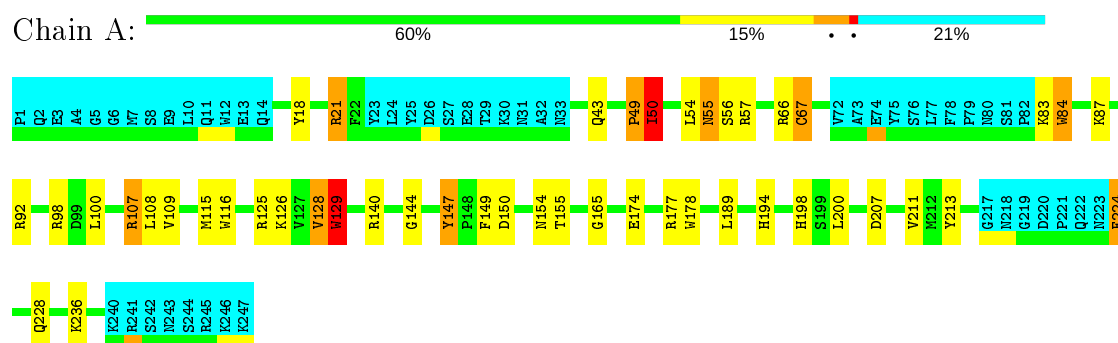


- Molecule 2: 2-O-sulfo-alpha-L-idopyranuronic acid-(1-4)-2-deoxy-6-O-sulfo-2-(sulfoamino)-alpha-D-glucopyranose-(1-4)-2-O-sulfo-alpha-L-idopyranuronic acid-(1-4)-2-deoxy-6-O-sulfo-2-(sulfoamino)-alpha-D-glucopyranose-(1-4)-2-O-sulfo-alpha-L-idopyranuronic acid-(1-4)-2-deoxy-6-O-sulfo-2-(sulfoamino)-alpha-D-glucopyranose-(1-4)-2-O-sulfo-alpha-L-idopyranuronic acid-(1-4)-2-deoxy-6-O-sulfo-2-(sulfoamino)-alpha-D-glucopyranose



### 4.2.8 Score per residue for model 8

#### • Molecule 1: Matrilysin



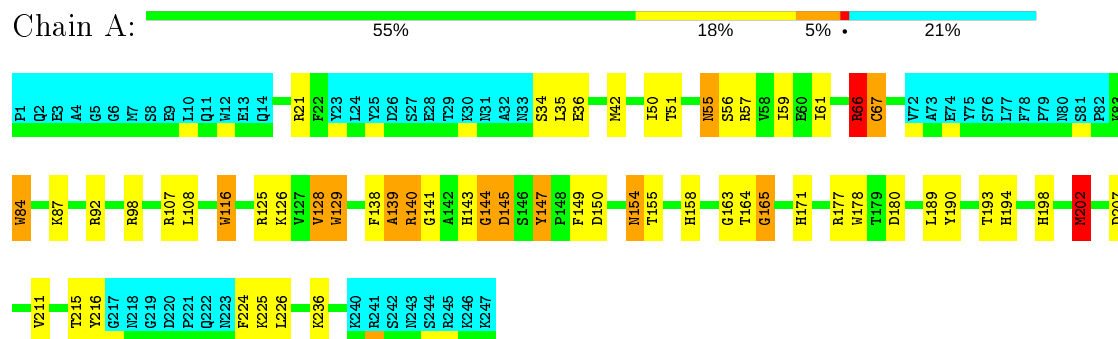
- Molecule 2: 2-O-sulfo-alpha-L-idopyranuronic acid-(1-4)-2-deoxy-6-O-sulfo-2-(sulfoamino)-alpha-D-glucopyranose-(1-4)-2-O-sulfo-alpha-L-idopyranuronic acid-(1-4)-2-deoxy-6-O-sulfo-2-(sulfoamino)-alpha-D-glucopyranose-(1-4)-2-O-sulfo-alpha-L-idopyranuronic acid-(1-4)-2-deoxy-6-O-sulfo-2-(sulfoamino)-alpha-D-glucopyranose-(1-4)-2-O-sulfo-alpha-L-idopyranuronic acid-(1-4)-2-deoxy-6-O-sulfo-2-(sulfoamino)-alpha-D-glucopyranose



SGN1  
IDS2  
SGN3  
IDS4  
SGN5  
IDS6  
SGN7  
IDS8

#### 4.2.9 Score per residue for model 9

- Molecule 1: Matrilysin



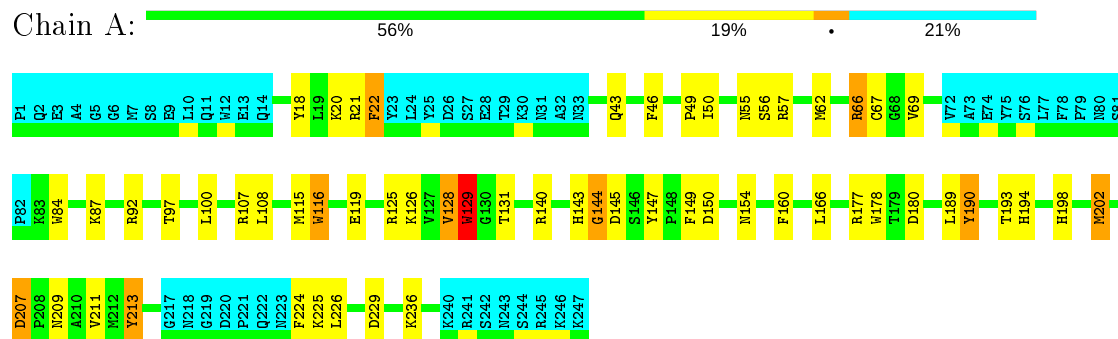
- Molecule 2: 2-O-sulfo-alpha-L-idopyranuronic acid-(1-4)-2-deoxy-6-O-sulfo-2-(sulfoamino)-alpha-D-glucopyranose-(1-4)-2-O-sulfo-alpha-L-idopyranuronic acid-(1-4)-2-deoxy-6-O-sulfo-2-(sulfoamino)-alpha-D-glucopyranose-(1-4)-2-O-sulfo-alpha-L-idopyranuronic acid-(1-4)-2-deoxy-6-O-sulfo-2-(sulfoamino)-alpha-D-glucopyranose-(1-4)-2-O-sulfo-alpha-L-idopyranuronic acid-(1-4)-2-deoxy-6-O-sulfo-2-(sulfoamino)-alpha-D-glucopyranose



SGN1  
IDS2  
SGN3  
IDS4  
SGN5  
IDS6  
SGN7  
IDS8


#### 4.2.10 Score per residue for model 10

- Molecule 1: Matrilysin



- Molecule 2: 2-O-sulfo-alpha-L-idopyranuronic acid-(1-4)-2-deoxy-6-O-sulfo-2-(sulfoamino)-alpha-D-glucopyranose-(1-4)-2-O-sulfo-alpha-L-idopyranuronic acid-(1-4)-2-deoxy-6-O-sulfo-2-(sulfoamino)-alpha-D-glucopyranose-(1-4)-2-O-sulfo-alpha-L-idopyranuronic acid-(1-4)-2-deoxy-6-O-sulfo-2-(sulfoamino)-alpha-D-glucopyranose

o-2-(sulfoamino)-alpha-D-glucopyranose-(1-4)-2-O-sulfo-alpha-L-idopyranuronic acid-(1-4)-2-deoxy-6-O-sulfo-2-(sulfoamino)-alpha-D-glucopyranose

Chain B:  88% 13%

SGN1  
IDS2  
SGN3  
IDS4  
SGN5  
IDS6  
SGN7  
IDS8

#### 4.2.11 Score per residue for model 11

- Molecule 1: Matrilysin

Chain A:  58% 19% 21%

P1 Q2 E3 A4 G5 G6 W7 S8 E9 L10 Q11 M12 E13 Q14 L19 R20 F22 Y23 L24 D26 Y25 S27 E28 T29 K30 N31 A32 N33 F46 P49 N55 S56 R57 I61 R66 V69 V72 A73 E74 Y75 S76 L77 F78 P79 N80 S81 P82 K83 W84 K87  
R98 R107 L108 V109 M115 W116 R125 K126 V127 V128 W129 R140 H143 G144 Y147 P148 F149 D150 M154 T155 H158 A161 P162 G163 T164 H171 F172 D173 R177 W178 T179 D180 L184 F188 L189 H194 H204 D207 V211 T215 Y216  
G217 N218 G219 D220 P221 G222 N223 F224 K236 L237 K240 R241 N242 S243 R244 K246 K247

- Molecule 2: 2-O-sulfo-alpha-L-idopyranuronic acid-(1-4)-2-deoxy-6-O-sulfo-2-(sulfoamino)-alpha-D-glucopyranose-(1-4)-2-O-sulfo-alpha-L-idopyranuronic acid-(1-4)-2-deoxy-6-O-sulfo-2-(sulfoamino)-alpha-D-glucopyranose-(1-4)-2-O-sulfo-alpha-L-idopyranuronic acid-(1-4)-2-deoxy-6-O-sulfo-2-(sulfoamino)-alpha-D-glucopyranose-(1-4)-2-O-sulfo-alpha-L-idopyranuronic acid-(1-4)-2-deoxy-6-O-sulfo-2-(sulfoamino)-alpha-D-glucopyranose

Chain B:  100%

SGN1  
IDS2  
SGN3  
IDS4  
SGN5  
IDS6  
SGN7  
IDS8

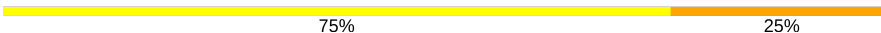
#### 4.2.12 Score per residue for model 12

- Molecule 1: Matrilysin

Chain A:  62% 15% 21%

P1 Q2 E3 A4 G5 G6 W7 S8 E9 L10 Q11 M12 E13 Q14 R21 F22 Y23 L24 D26 Y25 S27 E28 T29 K30 N31 A32 N33 F46 P49 N55 S56 R57 I61 R66 V69 V72 A73 E74 Y75 S76 L77 F78 P79 N80 S81  
R98 R107 L108 V109 M115 W116 R125 K126 V127 V128 W129 R140 H143 G144 Y147 P148 F149 D150 M154 T155 H158 A161 P162 G163 T164 H171 F172 D173 R177 W178 T179 D180 L184 F188 L189 H194 H204 D207 V211 T215 Y216  
K236 K240 R241 N242 S243 R244 K246 K247

- Molecule 2: 2-O-sulfo-alpha-L-idopyranuronic acid-(1-4)-2-deoxy-6-O-sulfo-2-(sulfoamino)-alpha-D-glucopyranose-(1-4)-2-O-sulfo-alpha-L-idopyranuronic acid-(1-4)-2-deoxy-6-O-sulfo-2-(sulfoamino)-alpha-D-glucopyranose-(1-4)-2-O-sulfo-alpha-L-idopyranuronic acid-(1-4)-2-deoxy-6-O-sulfo-2-(sulfoamino)-alpha-D-glucopyranose-(1-4)-2-O-sulfo-alpha-L-idopyranuronic acid-(1-4)-2-deoxy-6-O-sulfo-2-(sulfoamino)-alpha-D-glucopyranose

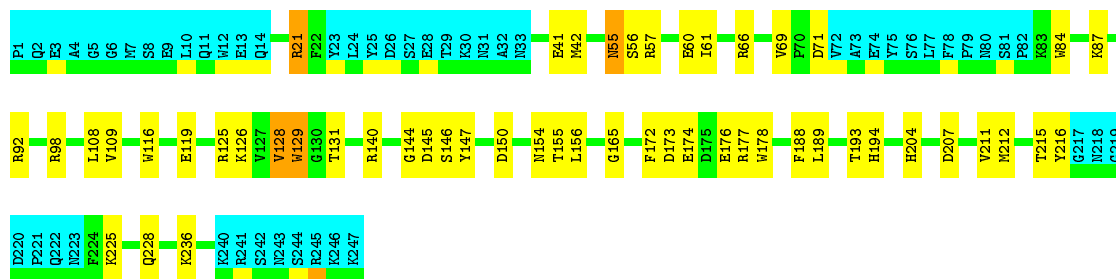
Chain B:  75% 25%

SGN1  
IDS2  
SGN3  
IDS4  
SGN5  
IDS6  
SGN7  
IDS8

#### 4.2.13 Score per residue for model 13

- Molecule 1: Matrilysin

Chain A:  58% 20% 21%



- Molecule 2: 2-O-sulfo-alpha-L-idopyranuronic acid-(1-4)-2-deoxy-6-O-sulfo-2-(sulfoamino)-alpha-D-glucopyranose-(1-4)-2-O-sulfo-alpha-L-idopyranuronic acid-(1-4)-2-deoxy-6-O-sulfo-2-(sulfoamino)-alpha-D-glucopyranose-(1-4)-2-O-sulfo-alpha-L-idopyranuronic acid-(1-4)-2-deoxy-6-O-sulfo-2-(sulfoamino)-alpha-D-glucopyranose-(1-4)-2-O-sulfo-alpha-L-idopyranuronic acid-(1-4)-2-deoxy-6-O-sulfo-2-(sulfoamino)-alpha-D-glucopyranose

Chain B:  100%

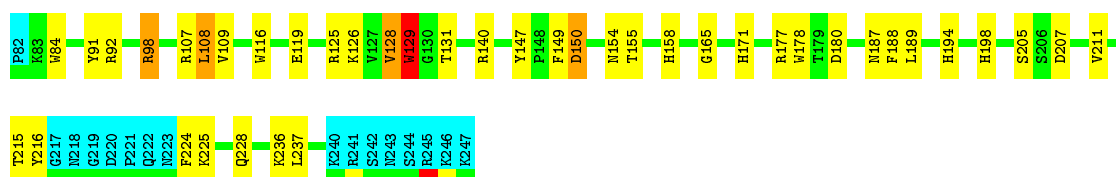
SGN1  
IDS2  
SGN3  
IDS4  
SGN5  
IDS6  
SGN7  
IDS8

#### 4.2.14 Score per residue for model 14

- Molecule 1: Matrilysin

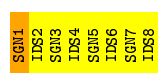
Chain A:  57% 19% 21%





- Molecule 2: 2-O-sulfo-alpha-L-idopyranuronic acid-(1-4)-2-deoxy-6-O-sulfo-2-(sulfoamino)-alpha-D-glucopyranose-(1-4)-2-O-sulfo-alpha-L-idopyranuronic acid-(1-4)-2-deoxy-6-O-sulfo-2-(sulfoamino)-alpha-D-glucopyranose-(1-4)-2-O-sulfo-alpha-L-idopyranuronic acid-(1-4)-2-deoxy-6-O-sulfo-2-(sulfoamino)-alpha-D-glucopyranose-(1-4)-2-O-sulfo-alpha-L-idopyranuronic acid-(1-4)-2-deoxy-6-O-sulfo-2-(sulfoamino)-alpha-D-glucopyranose

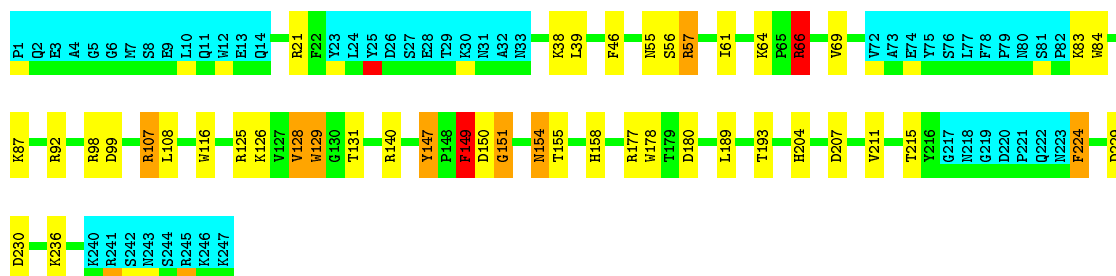
Chain B: 88% 13%



#### 4.2.15 Score per residue for model 15

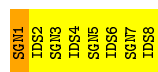
- Molecule 1: Matrilysin

Chain A: 61% 15% 21%



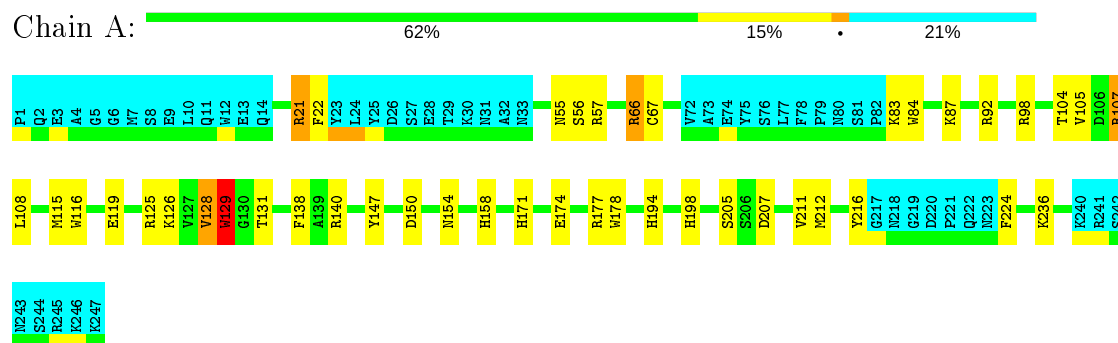
- Molecule 2: 2-O-sulfo-alpha-L-idopyranuronic acid-(1-4)-2-deoxy-6-O-sulfo-2-(sulfoamino)-alpha-D-glucopyranose-(1-4)-2-O-sulfo-alpha-L-idopyranuronic acid-(1-4)-2-deoxy-6-O-sulfo-2-(sulfoamino)-alpha-D-glucopyranose-(1-4)-2-O-sulfo-alpha-L-idopyranuronic acid-(1-4)-2-deoxy-6-O-sulfo-2-(sulfoamino)-alpha-D-glucopyranose-(1-4)-2-O-sulfo-alpha-L-idopyranuronic acid-(1-4)-2-deoxy-6-O-sulfo-2-(sulfoamino)-alpha-D-glucopyranose

Chain B: 88% 13%

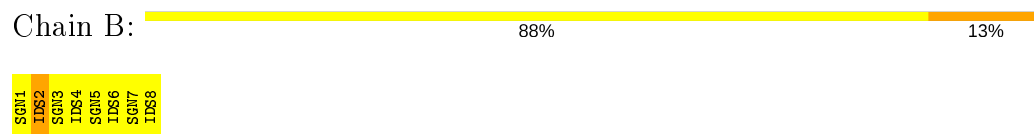


#### 4.2.16 Score per residue for model 16

- Molecule 1: Matrilysin



- Molecule 2: 2-O-sulfo-alpha-L-idopyranuronic acid-(1-4)-2-deoxy-6-O-sulfo-2-(sulfoamino)-alpha-D-glucopyranose-(1-4)-2-O-sulfo-alpha-L-idopyranuronic acid-(1-4)-2-deoxy-6-O-sulfo-2-(sulfoamino)-alpha-D-glucopyranose-(1-4)-2-O-sulfo-alpha-L-idopyranuronic acid-(1-4)-2-deoxy-6-O-sulfo-2-(sulfoamino)-alpha-D-glucopyranose-(1-4)-2-O-sulfo-alpha-L-idopyranuronic acid-(1-4)-2-deoxy-6-O-sulfo-2-(sulfoamino)-alpha-D-glucopyranose



## 5 Refinement protocol and experimental data overview

The models were refined using the following method: *torsion angle dynamics*.

Of the 120 calculated structures, 16 were deposited, based on the following criterion: *structures with the least restraint violations*.

The following table shows the software used for structure solution, optimisation and refinement.

| Software name | Classification        | Version |
|---------------|-----------------------|---------|
| CYANA         | refinement            | 2.1     |
| SYBYL-X       | refinement            | 2.1.1   |
| CYANA         | structure calculation | 2.1     |

The following table shows chemical shift validation statistics as aggregates over all chemical shift files. Detailed validation can be found in section 7 of this report.

|  |                |
|--|----------------|
| Chemical shift file(s)                       | working_cs.cif |
| Number of chemical shift lists               | 1              |
| Total number of shifts                       | 2746           |
| Number of shifts mapped to atoms             | 2746           |
| Number of unparsed shifts                    | 0              |
| Number of shifts with mapping errors         | 0              |
| Number of shifts with mapping warnings       | 0              |
| Assignment completeness (well-defined parts) | 85%            |

No validations of the models with respect to experimental NMR restraints is performed at this time.

## 6 Model quality i

### 6.1 Standard geometry i

Bond lengths and bond angles in the following residue types are not validated in this section: ZN, IDS, CA, SGN

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 (average) 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     | 0.99±0.01    | 0±0/1582 ( 0.0± 0.0%) | 1.17±0.02   | 13±1/2141 ( 0.6± 0.1%) |
| All | All   | 0.99         | 0/25312 ( 0.0%)       | 1.17        | 204/34256 ( 0.6%)      |

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

| Mol | Chain | Chirality | Planarity |
|-----|-------|-----------|-----------|
| 1   | A     | 0.0±0.0   | 1.2±1.2   |
| All | All   | 0         | 19        |

There are no bond-length outliers.

All unique angle outliers are listed below. They are sorted according to the Z-score of the worst occurrence in the ensemble.

| Mol | Chain | Res | Type | Atoms      | Z     | Observed(°) | Ideal(°) | Models |       |
|-----|-------|-----|------|------------|-------|-------------|----------|--------|-------|
|     |       |     |      |            |       |             |          | Worst  | Total |
| 1   | A     | 107 | ARG  | NE-CZ-NH1  | 8.23  | 124.41      | 120.30   | 6      | 13    |
| 1   | A     | 140 | ARG  | NE-CZ-NH1  | 7.97  | 124.28      | 120.30   | 14     | 15    |
| 1   | A     | 21  | ARG  | NE-CZ-NH1  | 7.69  | 124.14      | 120.30   | 14     | 14    |
| 1   | A     | 125 | ARG  | NE-CZ-NH1  | 7.60  | 124.10      | 120.30   | 9      | 15    |
| 1   | A     | 66  | ARG  | NE-CZ-NH1  | 7.22  | 123.91      | 120.30   | 10     | 10    |
| 1   | A     | 177 | ARG  | NE-CZ-NH1  | 7.07  | 123.83      | 120.30   | 14     | 14    |
| 1   | A     | 57  | ARG  | NE-CZ-NH1  | 6.95  | 123.78      | 120.30   | 8      | 16    |
| 1   | A     | 84  | TRP  | CE2-CD2-CG | -6.90 | 101.78      | 107.30   | 14     | 16    |
| 1   | A     | 98  | ARG  | NE-CZ-NH1  | 6.83  | 123.71      | 120.30   | 4      | 12    |
| 1   | A     | 178 | TRP  | CE2-CD2-CG | -6.78 | 101.88      | 107.30   | 16     | 16    |
| 1   | A     | 92  | ARG  | NE-CZ-NH1  | 6.73  | 123.67      | 120.30   | 12     | 15    |
| 1   | A     | 116 | TRP  | CE2-CD2-CG | -6.38 | 102.20      | 107.30   | 12     | 16    |

*Continued on next page...*



Continued from previous page...

| Mol | Chain | Res | Type | Atoms      | Z     | Observed(°) | Ideal(°) | Models |       |
|-----|-------|-----|------|------------|-------|-------------|----------|--------|-------|
|     |       |     |      |            |       |             |          | Worst  | Total |
| 1   | A     | 129 | TRP  | CE2-CD2-CG | -6.22 | 102.33      | 107.30   | 4      | 16    |
| 1   | A     | 202 | MET  | CG-SD-CE   | 6.00  | 109.81      | 100.20   | 4      | 4     |
| 1   | A     | 204 | HIS  | CG-CD2-NE2 | -5.69 | 98.38       | 109.20   | 15     | 1     |
| 1   | A     | 190 | TYR  | CB-CG-CD2  | -5.45 | 117.73      | 121.00   | 10     | 1     |
| 1   | A     | 22  | PHE  | CB-CG-CD1  | -5.41 | 117.01      | 120.80   | 10     | 2     |
| 1   | A     | 18  | TYR  | CB-CG-CD2  | -5.33 | 117.80      | 121.00   | 8      | 1     |
| 1   | A     | 198 | HIS  | CG-CD2-NE2 | -5.29 | 99.15       | 109.20   | 9      | 2     |
| 1   | A     | 150 | ASP  | N-CA-CB    | 5.26  | 120.08      | 110.60   | 4      | 2     |
| 1   | A     | 160 | PHE  | CB-CG-CD2  | 5.07  | 124.35      | 120.80   | 12     | 1     |
| 1   | A     | 22  | PHE  | CB-CG-CD2  | 5.03  | 124.32      | 120.80   | 10     | 1     |
| 1   | A     | 160 | PHE  | CB-CG-CD1  | -5.01 | 117.29      | 120.80   | 12     | 1     |

There are no chirality outliers.

All unique planar outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

| Mol | Chain | Res | Type | Group     | Models (Total) |
|-----|-------|-----|------|-----------|----------------|
| 1   | A     | 216 | TYR  | Sidechain | 5              |
| 1   | A     | 46  | PHE  | Sidechain | 3              |
| 1   | A     | 224 | PHE  | Sidechain | 2              |
| 1   | A     | 147 | TYR  | Sidechain | 2              |
| 1   | A     | 149 | PHE  | Sidechain | 1              |
| 1   | A     | 190 | TYR  | Sidechain | 1              |
| 1   | A     | 160 | PHE  | Sidechain | 1              |
| 1   | A     | 213 | TYR  | Sidechain | 1              |
| 1   | A     | 91  | TYR  | Sidechain | 1              |
| 1   | A     | 18  | TYR  | Sidechain | 1              |
| 1   | A     | 138 | PHE  | Sidechain | 1              |

## 6.2 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 each 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 averaged over the ensemble.

| Mol | Chain | Non-H | H(model) | H(added) | Clashes |
|-----|-------|-------|----------|----------|---------|
| 1   | A     | 1538  | 1512     | 1511     | 6±2     |
| 2   | B     | 140   | 60       | 45       | 0±0     |
| All | All   | 26912 | 25152    | 24897    | 96      |

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 2.

All unique clashes are listed below, sorted by their clash magnitude.

| Atom-1           | Atom-2           | Clash(Å) | Distance(Å) | Models |       |
|------------------|------------------|----------|-------------|--------|-------|
|                  |                  |          |             | Worst  | Total |
| 1:A:46:PHE:CD1   | 1:A:61:ILE:HD12  | 0.62     | 2.29        | 14     | 3     |
| 1:A:158:HIS:NE2  | 1:A:171:HIS:ND1  | 0.58     | 2.46        | 16     | 8     |
| 1:A:46:PHE:CD2   | 1:A:61:ILE:HD12  | 0.57     | 2.34        | 11     | 2     |
| 1:A:46:PHE:CE1   | 1:A:61:ILE:HD12  | 0.56     | 2.36        | 15     | 2     |
| 1:A:194:HIS:CD2  | 1:A:198:HIS:NE2  | 0.55     | 2.73        | 14     | 5     |
| 1:A:116:TRP:CZ3  | 1:A:202:MET:SD   | 0.55     | 3.00        | 2      | 5     |
| 1:A:115:MET:SD   | 1:A:224:PHE:CZ   | 0.55     | 3.00        | 3      | 3     |
| 1:A:224:PHE:CD1  | 1:A:224:PHE:C    | 0.54     | 2.81        | 7      | 1     |
| 1:A:224:PHE:CD2  | 1:A:224:PHE:C    | 0.54     | 2.80        | 2      | 2     |
| 1:A:67:CYS:SG    | 1:A:194:HIS:CD2  | 0.53     | 3.02        | 10     | 3     |
| 1:A:194:HIS:CD2  | 1:A:198:HIS:HE2  | 0.51     | 2.22        | 14     | 3     |
| 1:A:22:PHE:CD1   | 1:A:66:ARG:CZ    | 0.50     | 2.94        | 10     | 1     |
| 2:B:1:SGN:O1S    | 2:B:1:SGN:C1     | 0.49     | 2.61        | 4      | 2     |
| 1:A:108:LEU:HD23 | 1:A:188:PHE:HB3  | 0.49     | 1.85        | 14     | 1     |
| 1:A:172:PHE:CE1  | 1:A:188:PHE:CZ   | 0.48     | 3.01        | 13     | 2     |
| 1:A:49:PRO:O     | 1:A:50:ILE:C     | 0.48     | 2.52        | 8      | 3     |
| 1:A:172:PHE:CD1  | 1:A:188:PHE:CE1  | 0.48     | 3.02        | 13     | 1     |
| 1:A:143:HIS:CG   | 1:A:144:GLY:N    | 0.48     | 2.81        | 9      | 4     |
| 1:A:194:HIS:CE1  | 1:A:204:HIS:CE1  | 0.47     | 3.02        | 13     | 3     |
| 1:A:108:LEU:HD23 | 1:A:188:PHE:CB   | 0.47     | 2.40        | 14     | 1     |
| 1:A:116:TRP:CE3  | 1:A:226:LEU:HD11 | 0.47     | 2.44        | 4      | 1     |
| 1:A:224:PHE:C    | 1:A:224:PHE:CD2  | 0.46     | 2.88        | 4      | 1     |
| 1:A:207:ASP:O    | 1:A:213:TYR:CD2  | 0.45     | 2.69        | 8      | 2     |
| 1:A:149:PHE:CE2  | 1:A:158:HIS:CE1  | 0.45     | 3.05        | 15     | 1     |
| 2:B:1:SGN:O2S    | 2:B:1:SGN:C1     | 0.44     | 2.65        | 14     | 1     |
| 1:A:172:PHE:CD1  | 1:A:188:PHE:CZ   | 0.44     | 3.05        | 13     | 1     |
| 1:A:128:VAL:C    | 1:A:129:TRP:CG   | 0.44     | 2.91        | 10     | 6     |
| 1:A:67:CYS:HG    | 1:A:198:HIS:CD2  | 0.43     | 2.30        | 6      | 1     |
| 1:A:190:TYR:CE1  | 1:A:216:TYR:CZ   | 0.43     | 3.05        | 3      | 1     |
| 1:A:46:PHE:CG    | 1:A:61:ILE:HD12  | 0.43     | 2.49        | 11     | 1     |
| 1:A:54:LEU:O     | 1:A:55:ASN:CG    | 0.43     | 2.57        | 8      | 1     |
| 1:A:161:ALA:O    | 1:A:163:GLY:N    | 0.43     | 2.52        | 11     | 1     |
| 1:A:119:GLU:OE2  | 1:A:226:LEU:HD13 | 0.43     | 2.13        | 10     | 1     |
| 1:A:150:ASP:O    | 1:A:151:GLY:C    | 0.42     | 2.57        | 15     | 1     |
| 1:A:207:ASP:O    | 1:A:213:TYR:CD1  | 0.42     | 2.72        | 4      | 2     |
| 1:A:163:GLY:O    | 1:A:165:GLY:N    | 0.42     | 2.52        | 9      | 1     |
| 1:A:34:SER:O     | 1:A:35:LEU:C     | 0.42     | 2.57        | 9      | 1     |
| 1:A:46:PHE:CE2   | 1:A:61:ILE:HD12  | 0.42     | 2.50        | 5      | 1     |

*Continued on next page...*

Continued from previous page...

| Atom-1          | Atom-2          | Clash(Å) | Distance(Å) | Models |       |
|-----------------|-----------------|----------|-------------|--------|-------|
|                 |                 |          |             | Worst  | Total |
| 2:B:1:SGN:O3S   | 2:B:1:SGN:O3    | 0.42     | 2.38        | 10     | 1     |
| 2:B:1:SGN:C1    | 2:B:1:SGN:O3S   | 0.42     | 2.68        | 15     | 1     |
| 1:A:172:PHE:CE2 | 1:A:188:PHE:CZ  | 0.42     | 3.08        | 7      | 1     |
| 1:A:213:TYR:CE2 | 1:A:215:THR:HB  | 0.41     | 2.49        | 1      | 1     |
| 1:A:138:PHE:O   | 1:A:139:ALA:CB  | 0.41     | 2.68        | 9      | 1     |
| 1:A:22:PHE:CE1  | 1:A:66:ARG:CZ   | 0.41     | 3.03        | 10     | 1     |
| 1:A:143:HIS:CE1 | 1:A:171:HIS:CE1 | 0.41     | 3.08        | 9      | 1     |
| 1:A:104:THR:O   | 1:A:105:VAL:C   | 0.41     | 2.58        | 16     | 1     |
| 1:A:21:ARG:NH2  | 2:B:2:IDS:O3S   | 0.41     | 2.54        | 16     | 1     |
| 1:A:67:CYS:SG   | 1:A:194:HIS:CE1 | 0.41     | 3.14        | 8      | 1     |
| 1:A:38:LYS:O    | 1:A:39:LEU:C    | 0.41     | 2.58        | 15     | 2     |
| 1:A:115:MET:SD  | 1:A:224:PHE:CE1 | 0.41     | 3.14        | 8      | 1     |
| 2:B:3:SGN:O3    | 2:B:4:IDS:O5    | 0.40     | 2.40        | 12     | 1     |
| 1:A:213:TYR:CE1 | 1:A:215:THR:HB  | 0.40     | 2.51        | 6      | 1     |
| 1:A:84:TRP:CZ2  | 1:A:200:LEU:CD2 | 0.40     | 3.04        | 8      | 1     |

## 6.3 Torsion angles [i](#)

### 6.3.1 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 PDB entries followed by that with respect to all NMR entries. 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     | 196/247 (79%)   | 170±3 (87±2%) | 18±3 (9±1%) | 8±2 (4±1%) | 5           | 31 |
| All | All   | 3136/3952 (79%) | 2716 (87%)    | 291 (9%)    | 129 (4%)   | 5           | 31 |

All 23 unique Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

| Mol | Chain | Res | Type | Models (Total) |
|-----|-------|-----|------|----------------|
| 1   | A     | 154 | ASN  | 16             |
| 1   | A     | 55  | ASN  | 16             |
| 1   | A     | 147 | TYR  | 15             |
| 1   | A     | 144 | GLY  | 11             |
| 1   | A     | 66  | ARG  | 10             |
| 1   | A     | 56  | SER  | 10             |
| 1   | A     | 128 | VAL  | 9              |

Continued on next page...

*Continued from previous page...*

| Mol | Chain | Res | Type | Models (Total) |
|-----|-------|-----|------|----------------|
| 1   | A     | 150 | ASP  | 8              |
| 1   | A     | 49  | PRO  | 6              |
| 1   | A     | 155 | THR  | 4              |
| 1   | A     | 165 | GLY  | 4              |
| 1   | A     | 50  | ILE  | 3              |
| 1   | A     | 224 | PHE  | 3              |
| 1   | A     | 129 | TRP  | 3              |
| 1   | A     | 173 | ASP  | 2              |
| 1   | A     | 239 | GLY  | 2              |
| 1   | A     | 151 | GLY  | 1              |
| 1   | A     | 84  | TRP  | 1              |
| 1   | A     | 141 | GLY  | 1              |
| 1   | A     | 139 | ALA  | 1              |
| 1   | A     | 162 | PRO  | 1              |
| 1   | A     | 145 | ASP  | 1              |
| 1   | A     | 164 | THR  | 1              |

### 6.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all NMR entries. 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     | 161/205 (79%)   | 135±5 (84±3%) | 26±5 (16±3%) | 5           | 42 |
| All | All   | 2576/3280 (79%) | 2165 (84%)    | 411 (16%)    | 5           | 42 |

All 86 unique residues with a non-rotameric sidechain are listed below. They are sorted by the frequency of occurrence in the ensemble.

| Mol | Chain | Res | Type | Models (Total) |
|-----|-------|-----|------|----------------|
| 1   | A     | 236 | LYS  | 16             |
| 1   | A     | 128 | VAL  | 16             |
| 1   | A     | 108 | LEU  | 16             |
| 1   | A     | 129 | TRP  | 16             |
| 1   | A     | 224 | PHE  | 15             |
| 1   | A     | 126 | LYS  | 15             |
| 1   | A     | 189 | LEU  | 15             |
| 1   | A     | 211 | VAL  | 14             |
| 1   | A     | 149 | PHE  | 12             |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Res | Type | Models (Total) |
|-----|-------|-----|------|----------------|
| 1   | A     | 215 | THR  | 12             |
| 1   | A     | 87  | LYS  | 10             |
| 1   | A     | 207 | ASP  | 10             |
| 1   | A     | 150 | ASP  | 10             |
| 1   | A     | 225 | LYS  | 9              |
| 1   | A     | 193 | THR  | 9              |
| 1   | A     | 67  | CYS  | 9              |
| 1   | A     | 180 | ASP  | 8              |
| 1   | A     | 109 | VAL  | 8              |
| 1   | A     | 61  | ILE  | 8              |
| 1   | A     | 131 | THR  | 8              |
| 1   | A     | 66  | ARG  | 7              |
| 1   | A     | 83  | LYS  | 7              |
| 1   | A     | 107 | ARG  | 7              |
| 1   | A     | 50  | ILE  | 6              |
| 1   | A     | 69  | VAL  | 6              |
| 1   | A     | 147 | TYR  | 6              |
| 1   | A     | 202 | MET  | 5              |
| 1   | A     | 119 | GLU  | 5              |
| 1   | A     | 55  | ASN  | 5              |
| 1   | A     | 39  | LEU  | 5              |
| 1   | A     | 164 | THR  | 5              |
| 1   | A     | 43  | GLN  | 5              |
| 1   | A     | 97  | THR  | 5              |
| 1   | A     | 100 | LEU  | 5              |
| 1   | A     | 64  | LYS  | 4              |
| 1   | A     | 21  | ARG  | 4              |
| 1   | A     | 184 | LEU  | 4              |
| 1   | A     | 98  | ARG  | 4              |
| 1   | A     | 59  | ILE  | 4              |
| 1   | A     | 154 | ASN  | 3              |
| 1   | A     | 228 | GLN  | 3              |
| 1   | A     | 226 | LEU  | 3              |
| 1   | A     | 115 | MET  | 3              |
| 1   | A     | 205 | SER  | 3              |
| 1   | A     | 216 | TYR  | 3              |
| 1   | A     | 145 | ASP  | 3              |
| 1   | A     | 90  | THR  | 3              |
| 1   | A     | 174 | GLU  | 3              |
| 1   | A     | 42  | MET  | 3              |
| 1   | A     | 229 | ASP  | 3              |
| 1   | A     | 62  | MET  | 2              |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Res | Type | Models (Total) |
|-----|-------|-----|------|----------------|
| 1   | A     | 212 | MET  | 2              |
| 1   | A     | 155 | THR  | 2              |
| 1   | A     | 36  | GLU  | 2              |
| 1   | A     | 20  | LYS  | 2              |
| 1   | A     | 237 | LEU  | 2              |
| 1   | A     | 85  | THR  | 2              |
| 1   | A     | 138 | PHE  | 1              |
| 1   | A     | 41  | GLU  | 1              |
| 1   | A     | 146 | SER  | 1              |
| 1   | A     | 176 | GLU  | 1              |
| 1   | A     | 190 | TYR  | 1              |
| 1   | A     | 140 | ARG  | 1              |
| 1   | A     | 88  | VAL  | 1              |
| 1   | A     | 60  | GLU  | 1              |
| 1   | A     | 232 | LYS  | 1              |
| 1   | A     | 156 | LEU  | 1              |
| 1   | A     | 22  | PHE  | 1              |
| 1   | A     | 18  | TYR  | 1              |
| 1   | A     | 19  | LEU  | 1              |
| 1   | A     | 125 | ARG  | 1              |
| 1   | A     | 235 | GLN  | 1              |
| 1   | A     | 99  | ASP  | 1              |
| 1   | A     | 57  | ARG  | 1              |
| 1   | A     | 113 | LEU  | 1              |
| 1   | A     | 177 | ARG  | 1              |
| 1   | A     | 71  | ASP  | 1              |
| 1   | A     | 166 | LEU  | 1              |
| 1   | A     | 187 | ASN  | 1              |
| 1   | A     | 227 | SER  | 1              |
| 1   | A     | 16  | GLN  | 1              |
| 1   | A     | 53  | MET  | 1              |
| 1   | A     | 209 | ASN  | 1              |
| 1   | A     | 51  | THR  | 1              |
| 1   | A     | 35  | LEU  | 1              |
| 1   | A     | 230 | ASP  | 1              |

### 6.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 6.4 Non-standard residues in protein, DNA, RNA chains ⓘ

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

## 6.5 Carbohydrates ⓘ

8 monosaccharides are modelled in this entry.

In the following table, the Counts columns list the number of bonds for which Mogul statistics could be retrieved, the number of bonds that are observed in the model and the number of bonds 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 is the number of standard deviations the observed value is removed from the expected value. A bond length with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the average root-mean-square of all Z scores of the bond lengths.

| Mol | Type | Chain | Res | Link | Counts   | Bond lengths |            |
|-----|------|-------|-----|------|----------|--------------|------------|
|     |      |       |     |      |          | RMSZ         | #Z>2       |
| 2   | SGN  | B     | 1   | 2    | 18,19,20 | 3.66±0.00    | 1±0 (5±0%) |
| 2   | IDS  | B     | 2   | 2    | 13,16,17 | 1.25±0.00    | 0±0 (0±0%) |
| 2   | SGN  | B     | 3   | 2    | 18,19,20 | 3.70±0.00    | 1±0 (5±0%) |
| 2   | IDS  | B     | 4   | 2    | 13,16,17 | 1.26±0.00    | 0±0 (0±0%) |
| 2   | SGN  | B     | 5   | 2    | 18,19,20 | 3.71±0.00    | 1±0 (5±0%) |
| 2   | IDS  | B     | 6   | 2    | 13,16,17 | 1.25±0.00    | 0±0 (0±0%) |
| 2   | SGN  | B     | 7   | 2    | 18,19,20 | 3.68±0.00    | 1±0 (5±0%) |
| 2   | IDS  | B     | 8   | 2    | 13,16,17 | 1.21±0.00    | 0±0 (0±0%) |

In the following table, the Counts columns list the number of angles for which Mogul statistics could be retrieved, the number of angles that are observed in the model and the number of 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 angle is the number of standard deviations the observed value is removed from the expected value. A bond angle with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the average root-mean-square of all Z scores of the bond angles.

| Mol | Type | Chain | Res | Link | Counts   | Bond angles |            |
|-----|------|-------|-----|------|----------|-------------|------------|
|     |      |       |     |      |          | RMSZ        | #Z>2       |
| 2   | SGN  | B     | 1   | 2    | 22,29,31 | 1.55±0.00   | 0±0 (0±0%) |
| 2   | IDS  | B     | 2   | 2    | 15,24,26 | 1.14±0.00   | 0±0 (0±0%) |
| 2   | SGN  | B     | 3   | 2    | 22,29,31 | 1.56±0.00   | 0±0 (0±0%) |
| 2   | IDS  | B     | 4   | 2    | 15,24,26 | 0.99±0.00   | 0±0 (0±0%) |
| 2   | SGN  | B     | 5   | 2    | 22,29,31 | 1.71±0.00   | 0±0 (0±0%) |
| 2   | IDS  | B     | 6   | 2    | 15,24,26 | 1.12±0.00   | 0±0 (0±0%) |
| 2   | SGN  | B     | 7   | 2    | 22,29,31 | 1.43±0.00   | 0±0 (0±0%) |
| 2   | IDS  | B     | 8   | 2    | 15,24,26 | 1.06±0.00   | 0±0 (0±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     |
|-----|------|-------|-----|------|---------|--------------|-----------|
| 2   | SGN  | B     | 1   | 2    | -       | 0±0,11,28,31 | 0±0,1,1,1 |
| 2   | IDS  | B     | 2   | 2    | -       | 0±0,5,26,29  | 0±0,1,1,1 |
| 2   | SGN  | B     | 3   | 2    | -       | 0±0,11,28,31 | 0±0,1,1,1 |
| 2   | IDS  | B     | 4   | 2    | -       | 0±0,5,26,29  | 0±0,1,1,1 |
| 2   | SGN  | B     | 5   | 2    | -       | 0±0,11,28,31 | 0±0,1,1,1 |
| 2   | IDS  | B     | 6   | 2    | -       | 0±0,5,26,29  | 0±0,1,1,1 |
| 2   | SGN  | B     | 7   | 2    | -       | 0±0,11,28,31 | 0±0,1,1,1 |
| 2   | IDS  | B     | 8   | 2    | -       | 0±0,5,26,29  | 0±0,1,1,1 |

All unique bond outliers are listed below. They are sorted according to the Z-score of the worst occurrence in the ensemble.

| Mol | Chain | Res | Type | Atoms | Z     | Observed(Å) | Ideal(Å) | Models |       |
|-----|-------|-----|------|-------|-------|-------------|----------|--------|-------|
|     |       |     |      |       |       |             |          | Worst  | Total |
| 2   | B     | 5   | SGN  | S1-N2 | 14.49 | 1.79        | 1.59     | 10     | 16    |
| 2   | B     | 3   | SGN  | S1-N2 | 14.43 | 1.79        | 1.59     | 12     | 16    |
| 2   | B     | 1   | SGN  | S1-N2 | 14.39 | 1.78        | 1.59     | 14     | 16    |
| 2   | B     | 7   | SGN  | S1-N2 | 14.33 | 1.78        | 1.59     | 5      | 16    |

There are no bond-angle outliers.

All unique chiral outliers are listed below.

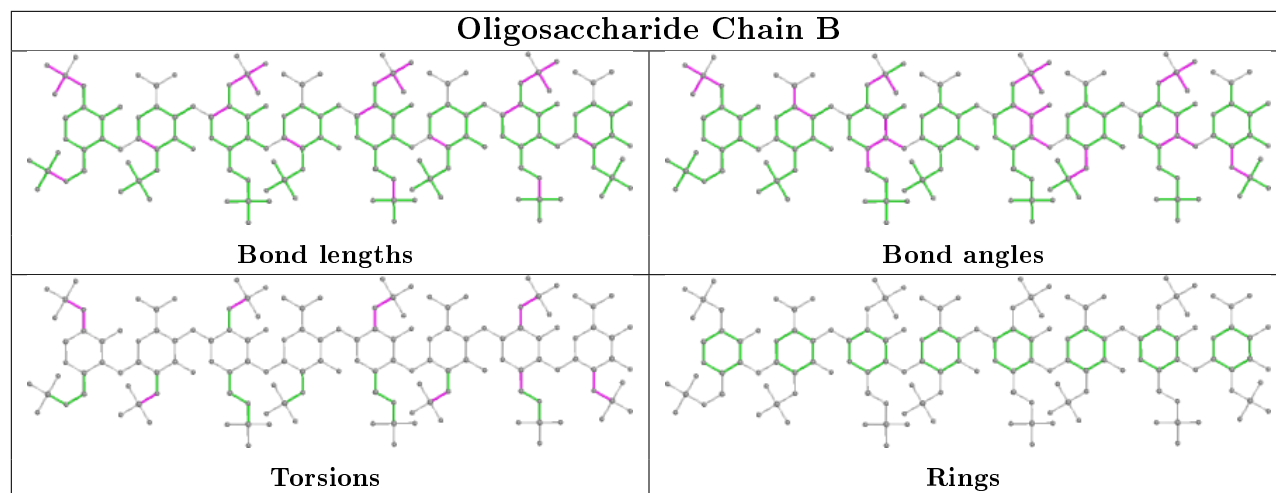
| Mol | Chain | Res | Type | Atoms | Models (Total) |
|-----|-------|-----|------|-------|----------------|
| 2   | B     | 6   | IDS  | C1    | 1              |

There are no torsion outliers.

There are no ring outliers.

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





## 6.6 Ligand geometry [i](#)

Of 4 ligands modelled in this entry, 4 are monoatomic - leaving 0 for Mogul analysis.

## 6.7 Other polymers [i](#)

There are no such molecules in this entry.

## 6.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 7 Chemical shift validation

The completeness of assignment taking into account all chemical shift lists is 85% for the well-defined parts and 82% for the entire structure.

### 7.1 Chemical shift list 1

File name: working\_cs.cif

Chemical shift list name: *assigned\_chem\_shift\_list\_1*

#### 7.1.1 Bookkeeping

The following table shows the results of parsing the chemical shift list and reports the number of nuclei with statistically unusual chemical shifts.

|   |      |
|---|------|
| Total number of shifts                  | 2746 |
| Number of shifts mapped to atoms        | 2746 |
| Number of unparsed shifts               | 0    |
| Number of shifts with mapping errors    | 0    |
| Number of shifts with mapping warnings  | 0    |
| Number of shift outliers (ShiftChecker) | 13   |

#### 7.1.2 Chemical shift referencing

The following table shows the suggested chemical shift referencing corrections.

| Nucleus                | # values | Correction $\pm$ precision, ppm | Suggested action           |
|------------------------|----------|---------------------------------|----------------------------|
| $^{13}\text{C}_\alpha$ | 236      | $-0.28 \pm 0.18$                | None needed ( $< 0.5$ ppm) |
| $^{13}\text{C}_\beta$  | 212      | $0.28 \pm 0.08$                 | None needed ( $< 0.5$ ppm) |
| $^{13}\text{C}'$       | 232      | $0.04 \pm 0.14$                 | None needed ( $< 0.5$ ppm) |
| $^{15}\text{N}$        | 220      | $-0.11 \pm 0.42$                | None needed ( $< 0.5$ ppm) |

#### 7.1.3 Completeness of resonance assignments

The following table shows the completeness of the chemical shift assignments for the well-defined regions of the structure. The overall completeness is 85%, i.e. 2041 atoms were assigned a chemical shift out of a possible 2391. 24 out of 27 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

|           | Total          | $^1\text{H}$  | $^{13}\text{C}$ | $^{15}\text{N}$ |
|-----------|----------------|---------------|-----------------|-----------------|
| Backbone  | 931/960 (97%)  | 367/382 (96%) | 383/392 (98%)   | 181/186 (97%)   |
| Sidechain | 942/1173 (80%) | 574/692 (83%) | 358/431 (83%)   | 10/50 (20%)     |

*Continued on next page...*

*Continued from previous page...*

|          | <b>Total</b>    | <b><sup>1</sup>H</b> | <b><sup>13</sup>C</b> | <b><sup>15</sup>N</b> |
|----------|-----------------|----------------------|-----------------------|-----------------------|
| Aromatic | 168/258 (65%)   | 87/138 (63%)         | 77/108 (71%)          | 4/12 (33%)            |
| Overall  | 2041/2391 (85%) | 1028/1212 (85%)      | 818/931 (88%)         | 195/248 (79%)         |

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 82%, i.e. 2474 atoms were assigned a chemical shift out of a possible 3016. 27 out of 31 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

|           | <b>Total</b>    | <b><sup>1</sup>H</b> | <b><sup>13</sup>C</b> | <b><sup>15</sup>N</b> |
|-----------|-----------------|----------------------|-----------------------|-----------------------|
| Backbone  | 1134/1207 (94%) | 446/480 (93%)        | 468/494 (95%)         | 220/233 (94%)         |
| Sidechain | 1136/1506 (75%) | 693/892 (78%)        | 431/544 (79%)         | 12/70 (17%)           |
| Aromatic  | 204/303 (67%)   | 105/161 (65%)        | 94/129 (73%)          | 5/13 (38%)            |
| Overall   | 2474/3016 (82%) | 1244/1533 (81%)      | 993/1167 (85%)        | 237/316 (75%)         |

#### 7.1.4 Statistically unusual chemical shifts ⓘ

The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

| Mol | Chain | Res | Type | Atom | Shift, ppm | Expected range, ppm | Z-score |
|-----|-------|-----|------|------|------------|---------------------|---------|
| 1   | A     | 107 | ARG  | NE   | 116.85     | 92.63 – 76.73       | 20.2    |
| 1   | A     | 57  | ARG  | NE   | 115.24     | 92.63 – 76.73       | 19.2    |
| 1   | A     | 92  | ARG  | NE   | 114.04     | 92.63 – 76.73       | 18.5    |
| 1   | A     | 66  | ARG  | NE   | 112.31     | 92.63 – 76.73       | 17.4    |
| 1   | A     | 247 | LYS  | CE   | 36.10      | 46.00 – 37.80       | -7.1    |
| 1   | A     | 162 | PRO  | HB2  | -0.32      | 3.82 – 0.32         | -6.8    |
| 1   | A     | 60  | GLU  | HB2  | 0.60       | 3.08 – 0.98         | -6.8    |
| 1   | A     | 60  | GLU  | CG   | 43.68      | 42.24 – 29.94       | 6.2     |
| 1   | A     | 60  | GLU  | HB3  | 0.64       | 3.10 – 0.90         | -6.2    |
| 1   | A     | 65  | PRO  | HA   | 2.44       | 6.05 – 2.75         | -6.0    |
| 1   | A     | 3   | GLU  | CG   | 43.11      | 42.24 – 29.94       | 5.7     |
| 1   | A     | 62  | MET  | HG2  | 0.38       | 4.23 – 0.63         | -5.7    |
| 1   | A     | 242 | SER  | N    | 135.65     | 134.24 – 98.34      | 5.4     |

#### 7.1.5 Random Coil Index (RCI) plots ⓘ

The image below reports *random coil index* values for the protein chains in the structure. The height of each bar gives a probability of a given residue to be disordered, as predicted from the available chemical shifts and the amino acid sequence. A value above 0.2 is an indication of significant predicted disorder. The colour of the bar shows whether the residue is in the well-defined core (black) or in the ill-defined residue ranges (cyan), as described in section 2 on ensemble composition.

Random coil index (RCI) for chain A:

