WORLDWIDE wwPDB PROTEIN DATA BANK

Full wwPDB NMR Structure Validation Report (i)

Jun 16, 2024 – 12:58 AM EDT

PDB ID	:	2N7B
BMRB ID	:	25799
Title	:	Solution structure of the human Siglec-8 lectin domain in complex with 6'sulfo sialyl Lewisx
Authors	:	Proepster, J.M.; Yang, F.; Rabbani, S.; Ernst, B.; Allain, F.H.-T.; Schubert, M.
Deposited on	:	2015-09-07

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

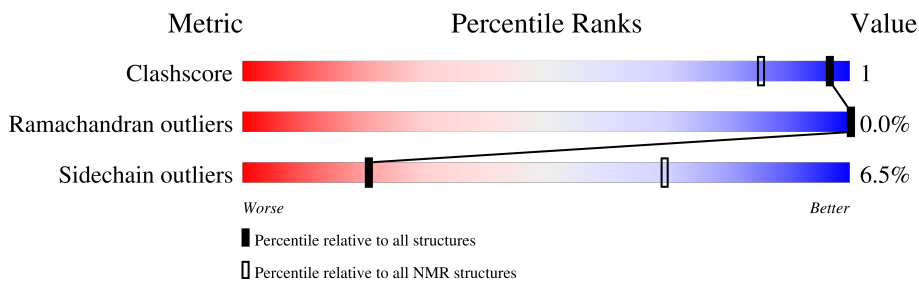|		
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

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 91%.

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	Percentile Ranks	Value
Clashscore		1
Ramachandran outliers		0.0%
Sidechain outliers		6.5%

Worse   *Better*

▌ Percentile relative to all structures
▯ Percentile relative to all NMR structures

Metric	Whole archive (#Entries)	NMR archive (#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 >=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 <=5%

Mol	Chain	Length	Quality of chain
1	A	145	78% • • 17%
2	B	4	75% 25%

2 Ensemble composition and analysis ⓘ

This entry contains 20 models. Model 10 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *lowest energy*.

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:6-A:53, A:61-A:65, A:72-A:138 (120)	0.14	10

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 2 clusters and 1 single-model cluster was found.

Cluster number	Models
1	1, 2, 3, 4, 5, 7, 8, 9, 10, 11, 14, 15, 19, 20
2	6, 12, 13, 16, 17
Single-model clusters	18

3 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 2412 atoms, of which 1172 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Sialic acid-binding Ig-like lectin 8.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	145	2290	738	1114	209	225	4	0

There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	26	SER	CYS	engineered mutation	UNP Q9NYZ4
A	140	GLY	-	expression tag	UNP Q9NYZ4
A	141	SER	-	expression tag	UNP Q9NYZ4
A	142	LEU	-	expression tag	UNP Q9NYZ4
A	143	VAL	-	expression tag	UNP Q9NYZ4
A	144	PRO	-	expression tag	UNP Q9NYZ4
A	145	ARG	-	expression tag	UNP Q9NYZ4

- Molecule 2 is an oligosaccharide called N-acetyl-alpha-neuraminic acid-(2-3)-6-O-sulfo-beta-D-galactopyranose-(1-4)-[alpha-L-fucopyranose-(1-3)]2-acetamido-2-deoxy-beta-D-glucopyranose.
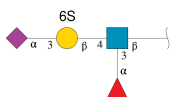
Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
2	B	4	108	31	49	2	25	1	0

- Molecule 3 is 3-aminopropan-1-ol (three-letter code: 0D8) (formula: C_3H_9NO).

0D8

N3

NH$_2$

C7

CA3

OH HO

C8
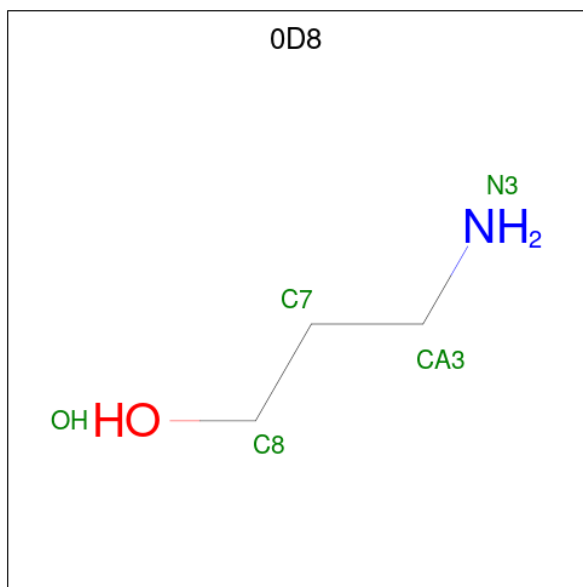
Mol	Chain	Residues	Atoms				
			Total	C	H	N	O
3	A	1	14	3	9	1	1

W O R L D W I D E

PDB

PROTEIN DATA BANK

4 Residue-property plots ⓘ

4.1 Average score per residue in the NMR ensemble

These plots are provided for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic is the same as shown in the summary in section 1 of this report. The second graphic shows the sequence where residues are colour-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 outliers are shown as green connectors. Residues which are classified as ill-defined in the NMR ensemble, are shown in cyan with an underline colour-coded according to the previous scheme. Residues which were present in the experimental sample, but not modelled in the final structure are shown in grey.

• Molecule 1: Sialic acid-binding Ig-like lectin 8
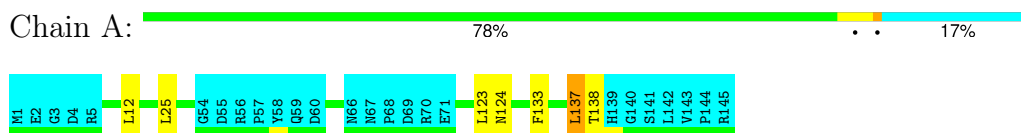
Chain A:

• Molecule 2: N-acetyl-alpha-neuraminic acid-(2-3)-6-O-sulfo-beta-D-galactopyranose-(1-4)-[alpha-L-fucopyranose-(1-3)]2-acetamido-2-deoxy-beta-D-glucopyranose

Chain B:
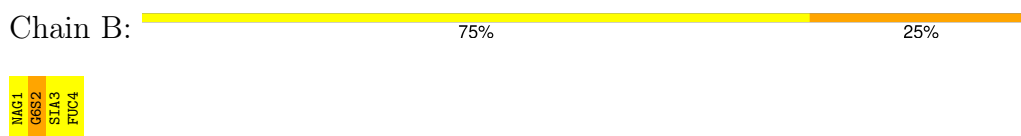
4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

4.2.1 Score per residue for model 1

• Molecule 1: Sialic acid-binding Ig-like lectin 8

Chain A:

• Molecule 2: N-acetyl-alpha-neuraminic acid-(2-3)-6-O-sulfo-beta-D-galactopyranose-(1-4)-[alpha-L-fucopyranose-(1-3)]2-acetamido-2-deoxy-beta-D-glucopyranose
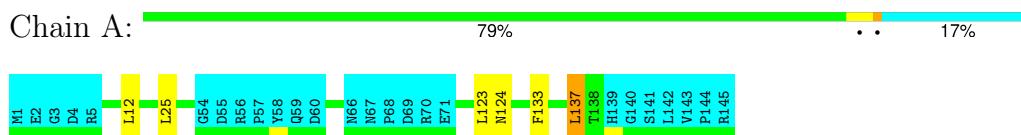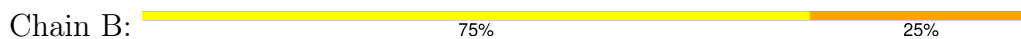
Chain B:

NAG1 G6S2 SIA3 FUC4

4.2.2 Score per residue for model 2

• Molecule 1: Sialic acid-binding Ig-like lectin 8

Chain A:

79%   17%

M1 E2 G3 D4 R5 L12 L25 G54 D55 R56 P57 Y58 Q59 D60 N66 N67 P68 D69 R70 E71 L123 N124 F133 L137 T138 H139 G140 S141 L142 V143 P144 R145

• Molecule 2: N-acetyl-alpha-neuraminic acid-(2-3)-6-O-sulfo-beta-D-galactopyranose-(1-4)-[alpha-L-fucopyranose-(1-3)]2-acetamido-2-deoxy-beta-D-glucopyranose

Chain B:

75%   25%

NAG1 G6S2 SIA3 FUC4

4.2.3 Score per residue for model 3

• Molecule 1: Sialic acid-binding Ig-like lectin 8

Chain A:

78%   17%

M1 E2 G3 D4 R5 L12 L25 G54 D55 R56 P57 Y58 Q59 D60 N66 N67 P68 D69 R70 E71 L123 N124 F133 L137 T138 H139 G140 S141 L142 V143 P144 R145

• Molecule 2: N-acetyl-alpha-neuraminic acid-(2-3)-6-O-sulfo-beta-D-galactopyranose-(1-4)-[alpha-L-fucopyranose-(1-3)]2-acetamido-2-deoxy-beta-D-glucopyranose

Chain B:

75%   25%

NAG1 G6S2 SIA3 FUC4

4.2.4 Score per residue for model 4

• Molecule 1: Sialic acid-binding Ig-like lectin 8

Chain A:

78%   17%

M1 E2 G3 D4 R5 L12 L25 G54 D55 R56 P57 Y58 Q59 D60 N66 N67 P68 D69 R70 E71 L123 N124 F133 L137 T138 H139 G140 S141 L142 V143 P144 R145

• Molecule 2: N-acetyl-alpha-neuraminic acid-(2-3)-6-O-sulfo-beta-D-galactopyranose-(1-4)-[alpha-L-fucopyranose-(1-3)]2-acetamido-2-deoxy-beta-D-glucopyranose

Chain B:
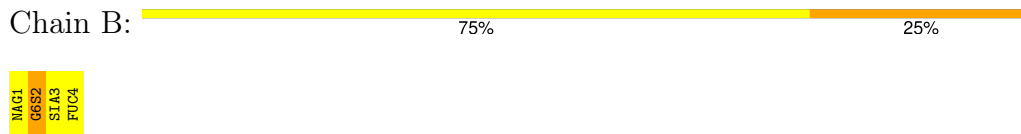
4.2.5 Score per residue for model 5

• Molecule 1: Sialic acid-binding Ig-like lectin 8
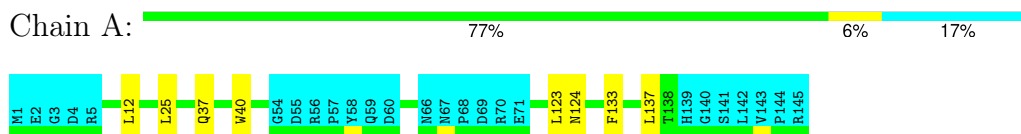
Chain A:

• Molecule 2: N-acetyl-alpha-neuraminic acid-(2-3)-6-O-sulfo-beta-D-galactopyranose-(1-4)-[alpha-L-fucopyranose-(1-3)]2-acetamido-2-deoxy-beta-D-glucopyranose

Chain B:
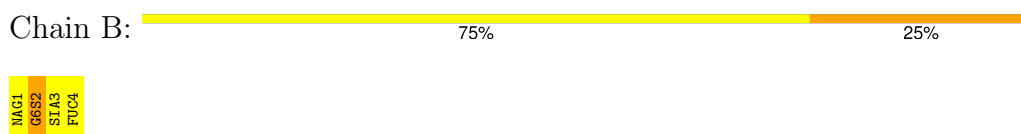
4.2.6 Score per residue for model 6

• Molecule 1: Sialic acid-binding Ig-like lectin 8

Chain A:

• Molecule 2: N-acetyl-alpha-neuraminic acid-(2-3)-6-O-sulfo-beta-D-galactopyranose-(1-4)-[alpha-L-fucopyranose-(1-3)]2-acetamido-2-deoxy-beta-D-glucopyranose

Chain B:

4.2.7 Score per residue for model 7

• Molecule 1: Sialic acid-binding Ig-like lectin 8

Chain A:

77%   5% ·   17%

M1 E2 G3 D4 R5 L12 L25 Q37 W40 G54 D55 R56 P57 Y58 Q59 D60 N66 N67 P68 D69 R70 E71 L123 N124 F133 L137 T138 H139 G140 S141 L142 V143 P144 R145

• Molecule 2: N-acetyl-alpha-neuraminic acid-(2-3)-6-O-sulfo-beta-D-galactopyranose-(1-4)-[alpha-L-fucopyranose-(1-3)]2-acetamido-2-deoxy-beta-D-glucopyranose

Chain B:

75%   25%

NAG1 G6S2 SIA3 FUC4

4.2.8 Score per residue for model 8

• Molecule 1: Sialic acid-binding Ig-like lectin 8

Chain A:

77%   6% ·   17%

M1 E2 G3 D4 R5 L12 L25 Q37 W40 G54 D55 R56 P57 Y58 Q59 D60 N66 N67 P68 D69 R70 E71 L123 N124 F133 L137 T138 H139 G140 S141 L142 V143 P144 R145
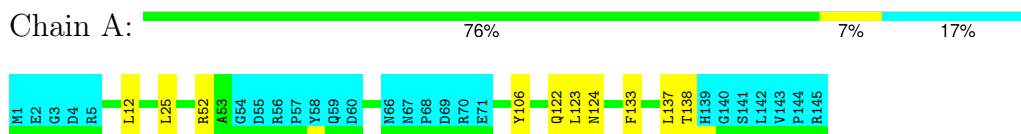
• Molecule 2: N-acetyl-alpha-neuraminic acid-(2-3)-6-O-sulfo-beta-D-galactopyranose-(1-4)-[alpha-L-fucopyranose-(1-3)]2-acetamido-2-deoxy-beta-D-glucopyranose

Chain B:

75%   25%

NAG1 G6S2 SIA3 FUC4

4.2.9 Score per residue for model 9

• Molecule 1: Sialic acid-binding Ig-like lectin 8

Chain A:

78%   · ·   17%

M1 E2 G3 D4 R5 L12 L25 G54 D55 R56 P57 Y58 Q59 D60 N66 N67 P68 D69 R70 E71 L123 N124 F133 L137 T138 H139 G140 S141 L142 V143 P144 R145

• Molecule 2: N-acetyl-alpha-neuraminic acid-(2-3)-6-O-sulfo-beta-D-galactopyranose-(1-4)-[alpha-L-fucopyranose-(1-3)]2-acetamido-2-deoxy-beta-D-glucopyranose
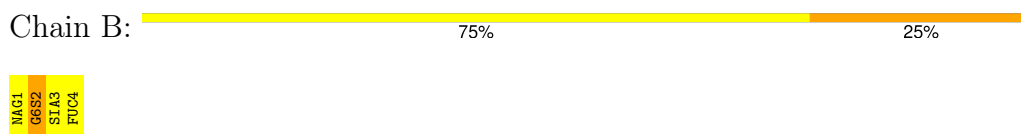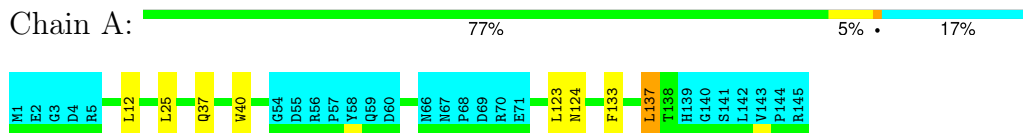
Chain B:

75%   25%

NAG1 G6S2 SIA3 FUC4
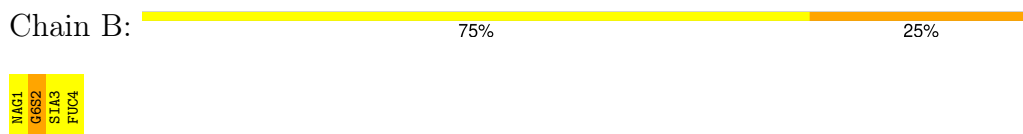
4.2.10   Score per residue for model 10 (medoid)

• Molecule 1: Sialic acid-binding Ig-like lectin 8

Chain A:

77%   5%   •   17%

M1 E2 G3 D4 R5 L12 L25 T41 G54 D55 R56 P57 Y58 Q59 D60 N66 N67 P68 D69 R70 E71 L123 N124 F133 L137 T138 H139 G140 S141 L142 V143 P144 R145

• Molecule 2: N-acetyl-alpha-neuraminic acid-(2-3)-6-O-sulfo-beta-D-galactopyranose-(1-4)-[alph a-L-fucopyranose-(1-3)]2-acetamido-2-deoxy-beta-D-glucopyranose
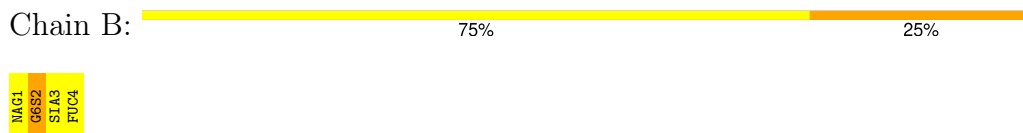
Chain B:

75%   25%

NAG1 G6S2 SIA3 FUC4
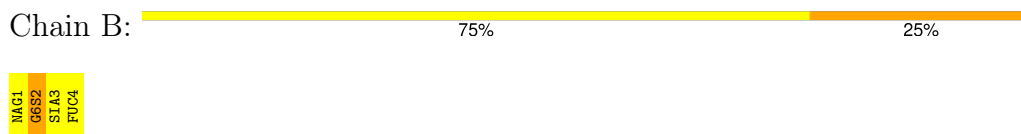
4.2.11   Score per residue for model 11

• Molecule 1: Sialic acid-binding Ig-like lectin 8

Chain A:

77%   6%   17%

M1 E2 G3 D4 R5 L12 L25 Q37 W40 G54 D55 R56 P57 Y58 Q59 D60 N66 N67 P68 D69 R70 E71 L123 N124 F133 L137 T138 H139 G140 S141 L142 V143 P144 R145
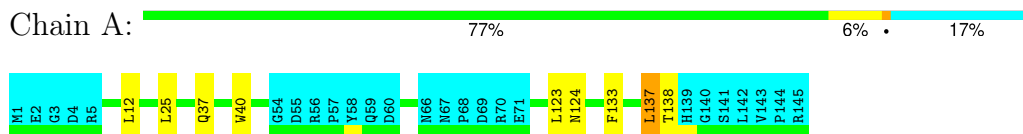
• Molecule 2: N-acetyl-alpha-neuraminic acid-(2-3)-6-O-sulfo-beta-D-galactopyranose-(1-4)-[alph a-L-fucopyranose-(1-3)]2-acetamido-2-deoxy-beta-D-glucopyranose
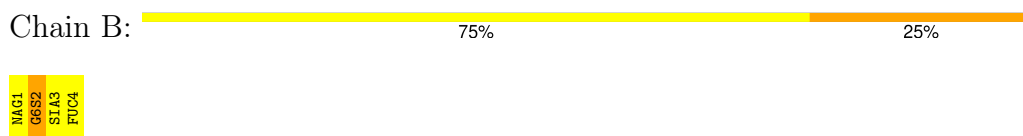
Chain B:

75%   25%

NAG1 G6S2 SIA3 FUC4

4.2.12   Score per residue for model 12

• Molecule 1: Sialic acid-binding Ig-like lectin 8

Chain A:

76%   6%   •   17%

M1 E2 G3 D4 R5 L12 L25 R52 A53 G54 D55 R56 P57 Y58 Q59 D60 N66 N67 P68 D69 R70 E71 Y106 L123 N124 K128 F133 L137 T138 H139 G140 S141 L142 V143 P144 R145
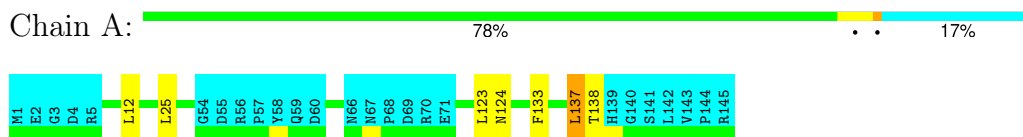
• Molecule 2: N-acetyl-alpha-neuraminic acid-(2-3)-6-O-sulfo-beta-D-galactopyranose-(1-4)-[alph a-L-fucopyranose-(1-3)]2-acetamido-2-deoxy-beta-D-glucopyranose

Chain B:

75%   25%

NAG1 G6S2 SIA3 FUC4

4.2.13 Score per residue for model 13

• Molecule 1: Sialic acid-binding Ig-like lectin 8

Chain A: 77% 6% • 17%

M1 E2 G3 D4 R5 L12 L25 R52 A53 G54 D55 R56 P57 Y58 Q59 D60 N66 N67 P68 D69 R70 E71 Y106 L123 N124 F133 L137 T138 H139 G140 S141 L142 V143 P144 R145

• Molecule 2: N-acetyl-alpha-neuraminic acid-(2-3)-6-O-sulfo-beta-D-galactopyranose-(1-4)-[alpha-L-fucopyranose-(1-3)]2-acetamido-2-deoxy-beta-D-glucopyranose
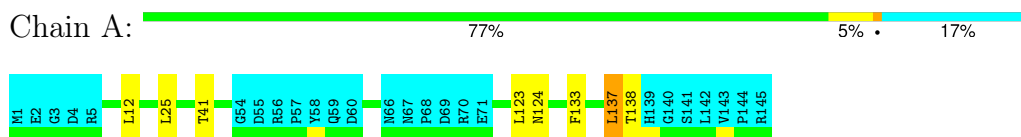
Chain B: 75% 25%

NAG1 G6S2 SIA3 FUC4

4.2.14 Score per residue for model 14

• Molecule 1: Sialic acid-binding Ig-like lectin 8

Chain A: 78% • • 17%

M1 E2 G3 D4 R5 L12 L25 G54 D55 R56 P57 Y58 Q59 D60 N66 N67 P68 D69 R70 E71 L123 N124 F133 L137 T138 H139 G140 S141 L142 V143 P144 R145

• Molecule 2: N-acetyl-alpha-neuraminic acid-(2-3)-6-O-sulfo-beta-D-galactopyranose-(1-4)-[alpha-L-fucopyranose-(1-3)]2-acetamido-2-deoxy-beta-D-glucopyranose

Chain B: 75% 25%

NAG1 G6S2 SIA3 FUC4

4.2.15 Score per residue for model 15

• Molecule 1: Sialic acid-binding Ig-like lectin 8

Chain A: 77% 6% • 17%

M1 E2 G3 D4 R5 L12 L25 Q37 W40 G54 D55 R56 P57 Y58 Q59 D60 N66 N67 P68 D69 R70 E71 L123 N124 F133 L137 T138 H139 G140 S141 L142 V143 P144 R145
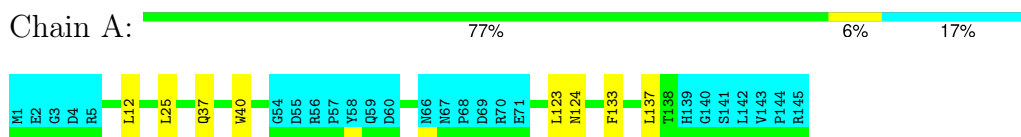
• Molecule 2: N-acetyl-alpha-neuraminic acid-(2-3)-6-O-sulfo-beta-D-galactopyranose-(1-4)-[alpha-L-fucopyranose-(1-3)]2-acetamido-2-deoxy-beta-D-glucopyranose
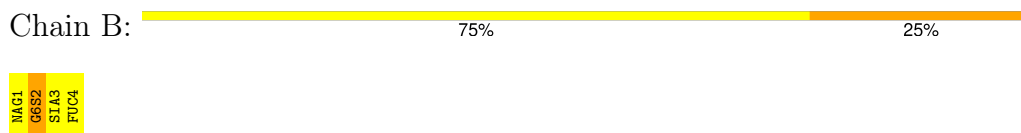
Chain B: 100%

4.2.16 Score per residue for model 16

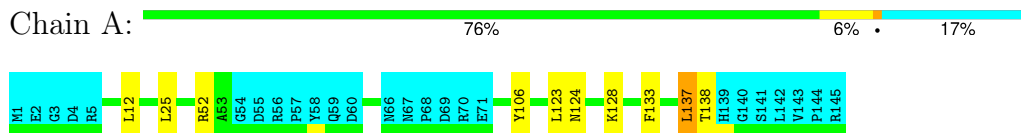• Molecule 1: Sialic acid-binding Ig-like lectin 8

Chain A: 77% 6% · 17%

• Molecule 2: N-acetyl-alpha-neuraminic acid-(2-3)-6-O-sulfo-beta-D-galactopyranose-(1-4)-[alpha-L-fucopyranose-(1-3)]2-acetamido-2-deoxy-beta-D-glucopyranose

Chain B: 75% 25%
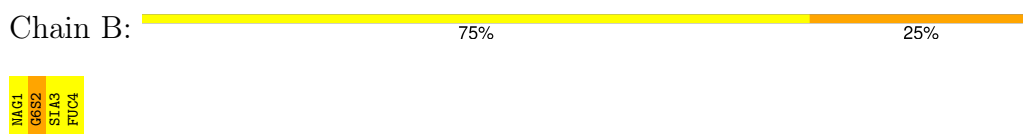
4.2.17 Score per residue for model 17

• Molecule 1: Sialic acid-binding Ig-like lectin 8

Chain A: 77% 6% · 17%

• Molecule 2: N-acetyl-alpha-neuraminic acid-(2-3)-6-O-sulfo-beta-D-galactopyranose-(1-4)-[alpha-L-fucopyranose-(1-3)]2-acetamido-2-deoxy-beta-D-glucopyranose

Chain B: 75% 25%
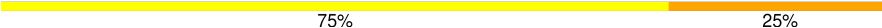
4.2.18 Score per residue for model 18

• Molecule 1: Sialic acid-binding Ig-like lectin 8

Chain A:

78%   17%

M1 E2 G3 D4 R5 L12 L25 G54 D55 R56 P57 Y58 Q59 D60 N66 N67 P68 D69 R70 E71 L123 N124 F133 L137 T138 H139 G140 S141 L142 V143 P144 R145

• Molecule 2: N-acetyl-alpha-neuraminic acid-(2-3)-6-O-sulfo-beta-D-galactopyranose-(1-4)-[alpha-L-fucopyranose-(1-3)]2-acetamido-2-deoxy-beta-D-glucopyranose

Chain B:

75%   25%

NAG1 G6S2 SIA3 FUC4
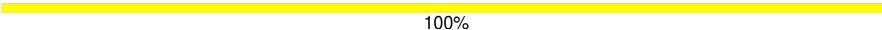
4.2.19 Score per residue for model 19

• Molecule 1: Sialic acid-binding Ig-like lectin 8
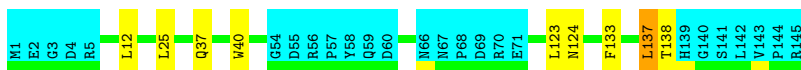
Chain A:

77%   6% •   17%

M1 E2 G3 D4 R5 L12 L25 Q37 W40 G54 D55 R56 P57 Y58 Q59 D60 N66 N67 P68 D69 R70 E71 L123 N124 F133 L137 T138 H139 G140 S141 L142 V143 P144 R145

• Molecule 2: N-acetyl-alpha-neuraminic acid-(2-3)-6-O-sulfo-beta-D-galactopyranose-(1-4)-[alpha-L-fucopyranose-(1-3)]2-acetamido-2-deoxy-beta-D-glucopyranose
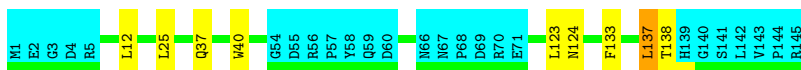
Chain B:

100%

NAG1 G6S2 SIA3 FUC4
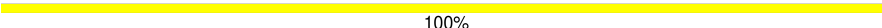
4.2.20 Score per residue for model 20

• Molecule 1: Sialic acid-binding Ig-like lectin 8

Chain A:

78%   17%

M1 E2 G3 D4 R5 L12 L25 G54 D55 R56 P57 Y58 Q59 D60 N66 N67 P68 D69 R70 E71 L123 N124 F133 L137 T138 H139 G140 S141 L142 V143 P144 R145
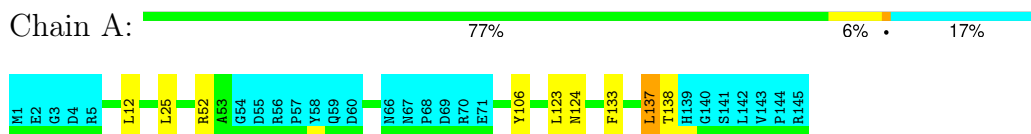
• Molecule 2: N-acetyl-alpha-neuraminic acid-(2-3)-6-O-sulfo-beta-D-galactopyranose-(1-4)-[alpha-L-fucopyranose-(1-3)]2-acetamido-2-deoxy-beta-D-glucopyranose
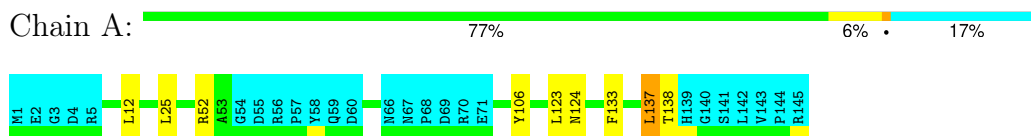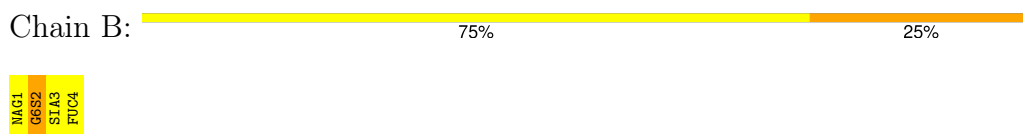
Chain B:

100%

NAG1 G6S2 SIA3 FUC4

5 Refinement protocol and experimental data overview (i)

The models were refined using the following method: *simulated annealing*.

Of the 500 calculated structures, 20 were deposited, based on the following criterion: *structures with the lowest energy*.

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

Software name	Classification	Version
CYANA	structure calculation	3.0
Amber	refinement	12

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	2
Total number of shifts	1819
Number of shifts mapped to atoms	1819
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	91%

6 Model quality ⓘ

6.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: 0D8, G6S, SIA, NAG, FUC

There are no covalent bond-length or bond-angle outliers.

There are no bond-length outliers.

There are no bond-angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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	972	932	932	1±1
2	B	59	49	46	0±0
All	All	20720	19800	19720	29

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.

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

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models Worst	Models Total
1:A:52:ARG:HG2	1:A:106:TYR:CE2	0.45	2.47	12	5
1:A:137:LEU:H	1:A:137:LEU:HD23	0.42	1.74	10	15
1:A:37:GLN:HA	1:A:40:TRP:CZ2	0.41	2.51	7	6
1:A:137:LEU:HD23	1:A:137:LEU:H	0.40	1.76	14	2
1:A:128:LYS:HE3	1:A:128:LYS:HA	0.40	1.93	12	1

6.3 Torsion angles ⓘ

6.3.1 Protein backbone ⓘ

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	120/145 (83%)	117±0 (98±0%)	3±0 (2±0%)	0±0 (0±0%)	100	100
All	All	2400/2900 (83%)	2348 (98%)	51 (2%)	1 (0%)	100	100

All 1 unique Ramachandran outliers are listed below.

Mol	Chain	Res	Type	Models (Total)
1	A	138	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	105/127 (83%)	98±1 (93±1%)	7±1 (7±1%)	21	69
All	All	2100/2540 (83%)	1963 (93%)	137 (7%)	21	69

All 9 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	12	LEU	20
1	A	25	LEU	20
1	A	123	LEU	20
1	A	124	ASN	20
1	A	133	PHE	20
1	A	137	LEU	20
1	A	138	THR	15
1	A	122	GLN	1
1	A	41	THR	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 ⓘ

4 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	Bond lengths		
					Counts	RMSZ	#Z>2
2	NAG	B	1	2,3	14,14,15	1.17±0.03	2±0 (13±2%)
2	G6S	B	2	2	15,15,16	1.02±0.02	1±1 (9±3%)
2	SIA	B	3	2	20,20,21	1.12±0.02	1±0 (5±0%)
2	FUC	B	4	2	10,10,11	1.17±0.03	2±0 (20±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	Bond angles		
					Counts	RMSZ	#Z>2
2	NAG	B	1	2,3	17,19,21	0.90±0.06	0±0 (1±2%)
2	G6S	B	2	2	21,22,24	0.92±0.05	0±0 (1±2%)
2	SIA	B	3	2	21,28,31	1.17±0.04	2±0 (10±1%)
2	FUC	B	4	2	14,14,16	0.63±0.03	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	NAG	B	1	2,3	-	0±0,6,23,26	0±0,1,1,1
2	G6S	B	2	2	-	0±0,6,23,26	0±0,1,1,1
2	SIA	B	3	2	-	0±0,18,34,38	0±0,1,1,1
2	FUC	B	4	2	-	-	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	3	SIA	C2-C1	3.99	1.57	1.52	18	20
2	B	4	FUC	O5-C1	2.66	1.48	1.43	18	20
2	B	2	G6S	O3-C3	2.38	1.48	1.43	17	20
2	B	4	FUC	O5-C5	2.33	1.48	1.43	15	20
2	B	1	NAG	O5-C1	2.31	1.47	1.43	9	17
2	B	1	NAG	O5-C5	2.31	1.47	1.43	4	20
2	B	2	G6S	O5-C5	2.17	1.47	1.43	1	5
2	B	2	G6S	O5-C1	2.03	1.47	1.43	2	2

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(^{o})	Ideal(^{o})	Models Worst	Total
2	B	3	SIA	O6-C2-C1	3.52	114.36	107.72	3	20
2	B	3	SIA	O1A-C1-C2	2.76	116.88	122.85	18	20
2	B	2	G6S	C1-O5-C5	2.64	115.72	112.19	1	6
2	B	1	NAG	O5-C1-C2	2.14	107.98	111.29	18	4
2	B	3	SIA	O6-C2-C3	2.12	107.71	110.56	12	1
2	B	3	SIA	O1B-C1-C2	2.05	118.04	112.71	3	3
2	B	2	G6S	C3-C4-C5	2.04	113.93	110.23	17	1

There are no chirality outliers.

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.

Oligosaccharide Chain B
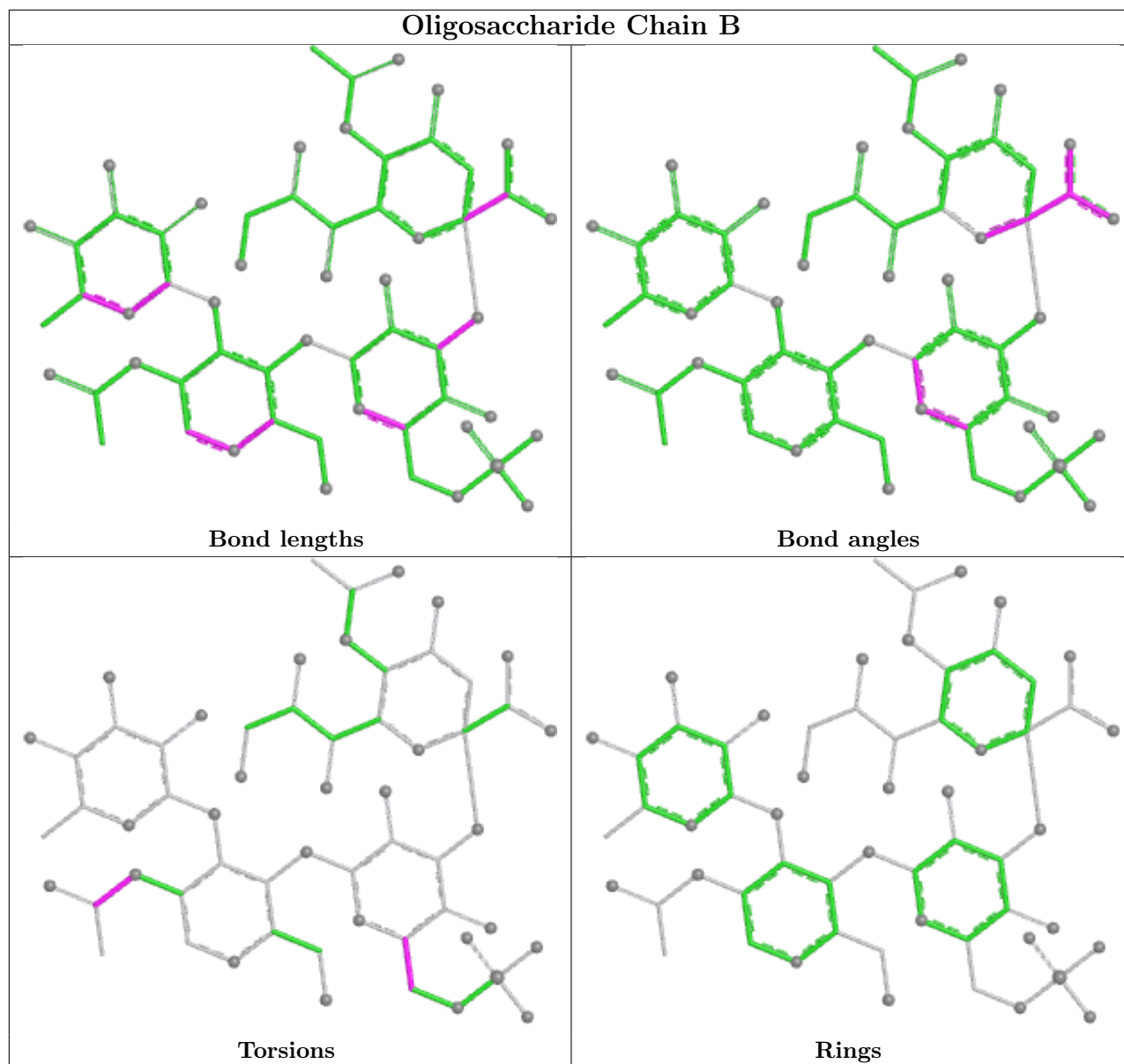
6.6 Ligand geometry ⓘ

1 ligand is 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	Bond lengths		
					Counts	RMSZ	#Z>2

Mol	Type	Chain	Res	Link	Bond lengths		
					Counts	RMSZ	#Z>2
3	0D8	A	201	2	4,4,4	0.28±0.01	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	Bond angles		
					Counts	RMSZ	#Z>2
3	0D8	A	201	2	3,3,3	1.40±0.10	1±0 (31±7%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	0D8	A	201	2	-	0±0,2,2,2	-

There are no bond-length outliers.

All unique angle outliers are listed below.

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models	
								Worst	Total
3	A	201	0D8	CA3-C7-C8	2.76	108.64	112.85	6	19

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

6.7 Other polymers ⓘ

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 (i)

The completeness of assignment taking into account all chemical shift lists is 91% for the well-defined parts and 90% for the entire structure.

7.1 Chemical shift list 1

File name: working_cs.cif

Chemical shift list name: *Siglec8_ bound_ chemical_ shifts*

7.1.1 Bookkeeping (i)

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	1774
Number of shifts mapped to atoms	1774
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	8

7.1.2 Chemical shift referencing (i)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_\alpha$	145	-0.00 ± 0.13	None needed (< 0.5 ppm)
$^{13}C_\beta$	132	0.11 ± 0.14	None needed (< 0.5 ppm)
$^{13}C'$	138	0.36 ± 0.17	None needed (< 0.5 ppm)
^{15}N	130	-1.10 ± 0.45	Should be applied

7.1.3 Completeness of resonance assignments (i)

The following table shows the completeness of the chemical shift assignments for the well-defined regions of the structure. The overall completeness is 91%, i.e. 1499 atoms were assigned a chemical shift out of a possible 1644. 0 out of 21 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	587/601 (98%)	239/245 (98%)	236/240 (98%)	112/116 (97%)
Sidechain	772/858 (90%)	521/555 (94%)	235/264 (89%)	16/39 (41%)

Continued on next page...

W O R L D W I D E
wwPDB
PROTEIN DATA BANK

Continued from previous page...

	Total	^{1}H	^{13}C	^{15}N
Aromatic	140/185 (76%)	82/90 (91%)	54/89 (61%)	4/6 (67%)
Overall	1499/1644 (91%)	842/890 (95%)	525/593 (89%)	132/161 (82%)

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 90%, i.e. 1772 atoms were assigned a chemical shift out of a possible 1977. 0 out of 23 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	698/723 (97%)	285/295 (97%)	283/290 (98%)	130/138 (94%)
Sidechain	924/1053 (88%)	622/676 (92%)	282/323 (87%)	20/54 (37%)
Aromatic	150/201 (75%)	88/98 (90%)	58/96 (60%)	4/7 (57%)
Overall	1772/1977 (90%)	995/1069 (93%)	623/709 (88%)	154/199 (77%)

### 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.

List Id	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	A	79	ARG	HH12	10.28	5.04 − 8.65	9.5
1	A	79	ARG	HH22	9.58	5.04 − 8.54	8.0
1	A	117	TRP	HD1	4.59	5.46 − 8.81	-7.6
1	A	117	TRP	HZ2	5.44	5.71 − 8.86	-5.9
1	A	95	ILE	HG13	-1.05	-0.82 − 3.23	-5.5
1	A	108	PHE	HB2	1.02	1.20 − 4.80	-5.5
1	A	112	ARG	HD3	1.69	1.81 − 4.39	-5.5
1	A	51	PHE	HB2	1.10	1.20 − 4.80	-5.2

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.If well-defined core and ill-defined regions are not identified then it is shown as gray bars.

Random coil index (RCI) for chain A:
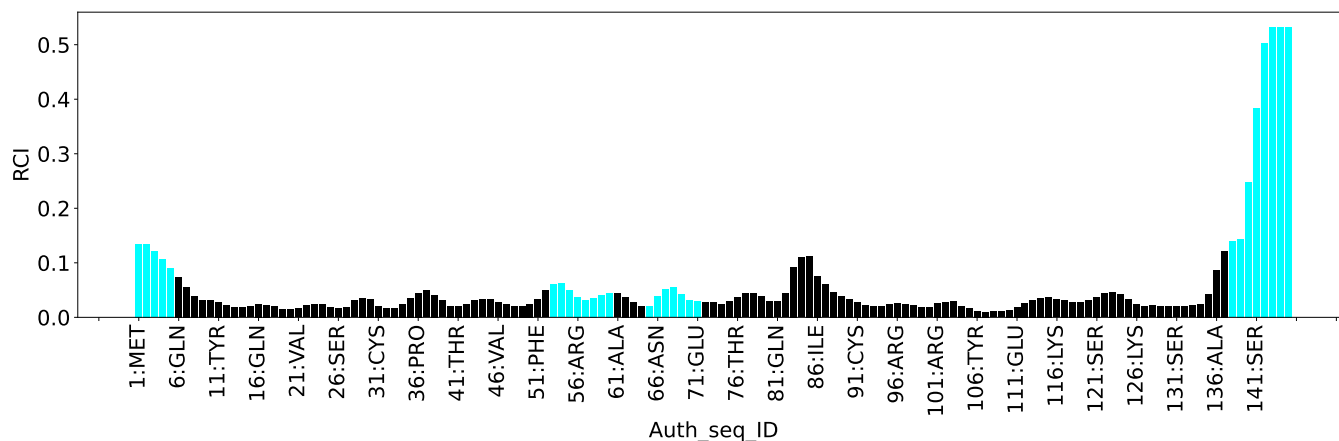
7.2 Chemical shift list 2

File name: working_cs.cif

Chemical shift list name: *sugar_bound_chemical_shifts*

7.2.1 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	45
Number of shifts mapped to atoms	45
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	0

7.2.2 Chemical shift referencing ⓘ

No chemical shift referencing corrections were calculated (not enough data).

7.2.3 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 0%, i.e. 0 atoms were assigned a chemical shift out of a possible 1644. 0 out of 21 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	0/601 (0%)	0/245 (0%)	0/240 (0%)	0/116 (0%)

Continued on next page...

Continued from previous page...

	Total	**¹H**	**¹³C**	**¹⁵N**
Sidechain	0/858 (0%)	0/555 (0%)	0/264 (0%)	0/39 (0%)
Aromatic	0/185 (0%)	0/90 (0%)	0/89 (0%)	0/6 (0%)
Overall	0/1644 (0%)	0/890 (0%)	0/593 (0%)	0/161 (0%)

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 0%, i.e. 0 atoms were assigned a chemical shift out of a possible 1977. 0 out of 23 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	**¹H**	**¹³C**	**¹⁵N**
Backbone	0/723 (0%)	0/295 (0%)	0/290 (0%)	0/138 (0%)
Sidechain	0/1053 (0%)	0/676 (0%)	0/323 (0%)	0/54 (0%)
Aromatic	0/201 (0%)	0/98 (0%)	0/96 (0%)	0/7 (0%)
Overall	0/1977 (0%)	0/1069 (0%)	0/709 (0%)	0/199 (0%)

7.2.4　Statistically unusual chemical shifts ⓘ

There are no statistically unusual chemical shifts.

7.2.5　Random Coil Index (RCI) plots ⓘ

No *random coil index* (RCI) plot could be generated from the current chemical shift list. RCI is only applicable to proteins