WORLDWIDE PDB
PROTEIN DATA BANK

wwPDB NMR Structure Validation Summary Report (i)

Sep 28, 2024 – 06:21 AM EDT

PDB ID : 1GYA
Title : N-GLYCAN AND POLYPEPTIDE NMR SOLUTION STRUCTURES OF THE ADHESION DOMAIN OF HUMAN CD2
Authors : Wyss, D.F.; Choi, J.S.; Wagner, G.
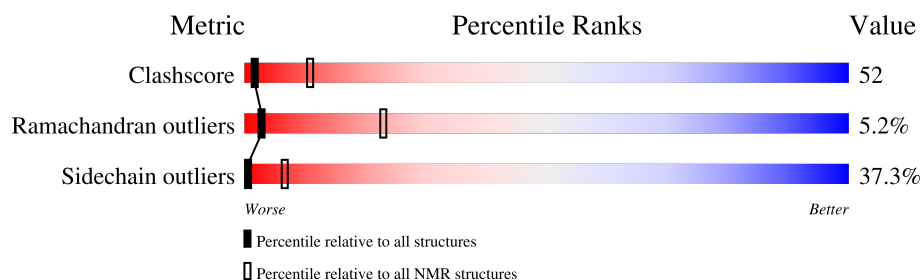Deposited on : 1995-05-26

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 was not calculated.

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		52
Ramachandran outliers		5.2%
Sidechain outliers		37.3%

Worse *Better*

▮ Percentile relative to all structures

▯ Percentile relative to all NMR structures

Metric	Whole archive (#Entries)	NMR archive (#Entries)
Clashscore	210492	14027
Ramachandran outliers	207382	12486
Sidechain outliers	206894	12463

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	105	11%　62%　21%　6%
2	B	9	89%　11%

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	BMA	3	2	-
2	B	MAN	4	1	-
2	B	MAN	8	1	-
2	B	MAN	9	1	-

WORLDWIDE wwPDB PROTEIN DATA BANK

2   Ensemble composition and analysis ⓘ

This entry contains 18 models. Model 11 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:5-A:103 (99)	0.84	11

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 2 single-model clusters were found.

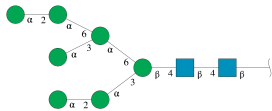| Cluster number	Models
1	6, 8, 9, 12, 15, 16, 17
2	1, 2, 3, 7, 10, 11, 18
3	5, 13
Single-model clusters	4; 14

3 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 1976 atoms, of which 994 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called HUMAN CD2.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	105	1774	560	897	144	172	1	0

- Molecule 2 is an oligosaccharide called alpha-D-mannopyranose-(1-2)-alpha-D-mannopyran ose-(1-6)-[alpha-D-mannopyranose-(1-3)]alpha-D-mannopyranose-(1-6)-[alpha-D-mannopyr anose-(1-2)-alpha-D-mannopyranose-(1-3)]beta-D-mannopyranose-(1-4)-2-acetamido-2-deox y-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.

Mol	Chain	Residues	Atoms					Trace
			Total	C	H	N	O	
2	B	9	202	58	97	2	45	0

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: HUMAN CD2

• Molecule 2: alpha-D-mannopyranose-(1-2)-alpha-D-mannopyranose-(1-6)-[alpha-D-mannopyranose-(1-3)]alpha-D-mannopyranose-(1-6)-[alpha-D-mannopyranose-(1-2)-alpha-D-mannopyranose-(1-3)]beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose
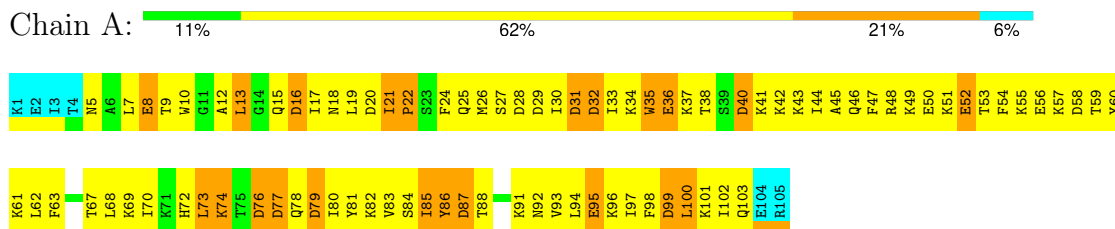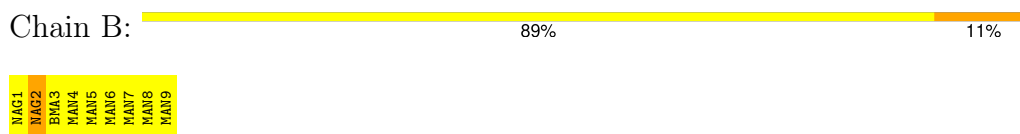

4.2   Residue scores for the representative (medoid) model from the NMR ensemble

The representative model is number 11. Colouring as in section 4.1 above.

• Molecule 1: HUMAN CD2
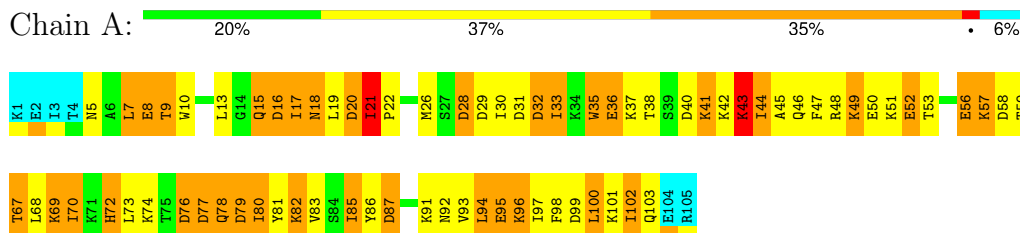
• Molecule 2: alpha-D-mannopyranose-(1-2)-alpha-D-mannopyranose-(1-6)-[alpha-D-mannopyran ose-(1-3)]alpha-D-mannopyranose-(1-6)-[alpha-D-mannopyranose-(1-2)-alpha-D-mannopyranose-(1-3)]beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2 -deoxy-beta-D-glucopyranose

Chain B:

89%   11%

NAG1 NAG2 BMA3 MAN4 MAN5 MAN6 MAN7 MAN8 MAN9

5 Refinement protocol and experimental data overview ⓘ

The models were refined using the following method: *distance geometry.*

Of the 120 calculated structures, 18 were deposited, based on the following criterion: *PRO-STAT/STRUCT_CHECK.*

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

Software name	Classification	Version
DGII	refinement	
DGII	structure solution	

No chemical shift data was provided.

6 Model quality ⓘ

6.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: BMA, NAG, 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 (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	1.40±0.00	19±0/837 (2.3± 0.0%)	1.38±0.01	30±1/1120 (2.7± 0.0%)
All	All	1.40	342/15066 (2.3%)	1.38	536/20160 (2.7%)

5 of 38 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
1	A	8	GLU	CD-OE2	10.12	1.36	1.25	7	11
1	A	56	GLU	CD-OE2	10.07	1.36	1.25	18	6
1	A	36	GLU	CD-OE1	10.06	1.36	1.25	2	10
1	A	50	GLU	CD-OE1	10.05	1.36	1.25	10	9
1	A	36	GLU	CD-OE2	10.05	1.36	1.25	6	8

5 of 30 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	48	ARG	NE-CZ-NH1	8.26	124.43	120.30	12	18
1	A	99	ASP	CB-CG-OD1	7.60	125.14	118.30	12	18
1	A	28	ASP	CB-CG-OD1	7.16	124.75	118.30	8	17
1	A	32	ASP	CB-CG-OD1	7.16	124.74	118.30	9	18
1	A	99	ASP	CB-CG-OD2	-7.14	111.88	118.30	12	18

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	823	839	838	96±15
2	B	105	97	88	1±1
All	All	16704	16848	16668	1735

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 52.

5 of 864 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models Worst	Total
1:A:83:VAL:HG22	1:A:85:ILE:HD11	1.14	1.16	17	9
1:A:73:LEU:HD21	1:A:100:LEU:HD11	1.10	1.19	8	3
1:A:38:THR:HG21	1:A:80:ILE:HG22	1.09	1.23	15	8
1:A:33:ILE:HG23	1:A:85:ILE:HG23	1.06	1.26	6	14
1:A:19:LEU:HD12	1:A:68:LEU:HD22	1.06	1.19	3	2

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	99/105 (94%)	79±2 (80±2%)	15±2 (15±2%)	5±1 (5±1%)	3	23
All	All	1782/1890 (94%)	1424 (80%)	265 (15%)	93 (5%)	3	23

5 of 17 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	72	HIS	16
1	A	22	PRO	14

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	27	SER	12
1	A	52	GLU	12
1	A	43	LYS	7

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	92/98 (94%)	58±4 (63±5%)	34±4 (37±5%)	0	7
All	All	1656/1764 (94%)	1038 (63%)	618 (37%)	0	7

5 of 82 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	16	ASP	18
1	A	74	LYS	17
1	A	9	THR	16
1	A	78	GLN	16
1	A	31	ASP	15

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 ⓘ

9 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	1,2	14,14,15	1.39±0.01	1±0 (7±0%)
2	NAG	B	2	2	14,14,15	1.36±0.01	1±0 (7±0%)
2	BMA	B	3	2	11,11,12	1.34±0.01	1±0 (9±0%)
2	MAN	B	4	2	11,11,12	1.31±0.01	1±0 (9±0%)
2	MAN	B	5	2	11,11,12	1.28±0.02	1±0 (9±0%)
2	MAN	B	6	2	11,11,12	1.23±0.01	1±0 (7±3%)
2	MAN	B	7	2	11,11,12	1.23±0.02	1±0 (7±3%)
2	MAN	B	8	2	11,11,12	1.30±0.01	1±0 (9±0%)
2	MAN	B	9	2	11,11,12	1.21±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
2	NAG	B	1	1,2	17,19,21	1.06±0.06	1±0 (5±0%)
2	NAG	B	2	2	17,19,21	1.16±0.16	1±0 (6±2%)
2	BMA	B	3	2	15,15,17	0.87±0.06	1±0 (6±0%)
2	MAN	B	4	2	15,15,17	1.02±0.05	1±0 (6±0%)
2	MAN	B	5	2	15,15,17	1.07±0.04	1±0 (6±0%)
2	MAN	B	6	2	15,15,17	1.14±0.03	1±0 (6±0%)
2	MAN	B	7	2	15,15,17	1.10±0.05	1±0 (6±0%)
2	MAN	B	8	2	15,15,17	1.09±0.04	1±0 (6±0%)
2	MAN	B	9	2	15,15,17	1.15±0.02	1±0 (6±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	1,2	-	0±0,6,23,26	0±0,1,1,1
2	NAG	B	2	2	-	0±0,6,23,26	0±0,1,1,1
2	BMA	B	3	2	-	0±0,2,19,22	0±0,1,1,1
2	MAN	B	4	2	-	0±0,2,19,22	0±0,1,1,1
2	MAN	B	5	2	-	0±0,2,19,22	0±0,1,1,1
2	MAN	B	6	2	-	0±0,2,19,22	0±0,1,1,1
2	MAN	B	7	2	-	0±0,2,19,22	0±0,1,1,1
2	MAN	B	8	2	-	0±0,2,19,22	0±0,1,1,1
2	MAN	B	9	2	-	0±0,2,19,22	0±0,1,1,1

5 of 8 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	1	NAG	C1-C2	2.73	1.56	1.52	2	18
2	B	2	NAG	C1-C2	2.69	1.56	1.52	6	18
2	B	3	BMA	C2-C3	2.42	1.56	1.52	14	18
2	B	5	MAN	C2-C3	2.41	1.56	1.52	8	18
2	B	4	MAN	C2-C3	2.29	1.56	1.52	8	18

5 of 10 unique angle outliers are listed below. They are sorted according to the Z-score of the worst occurrence in the ensemble.

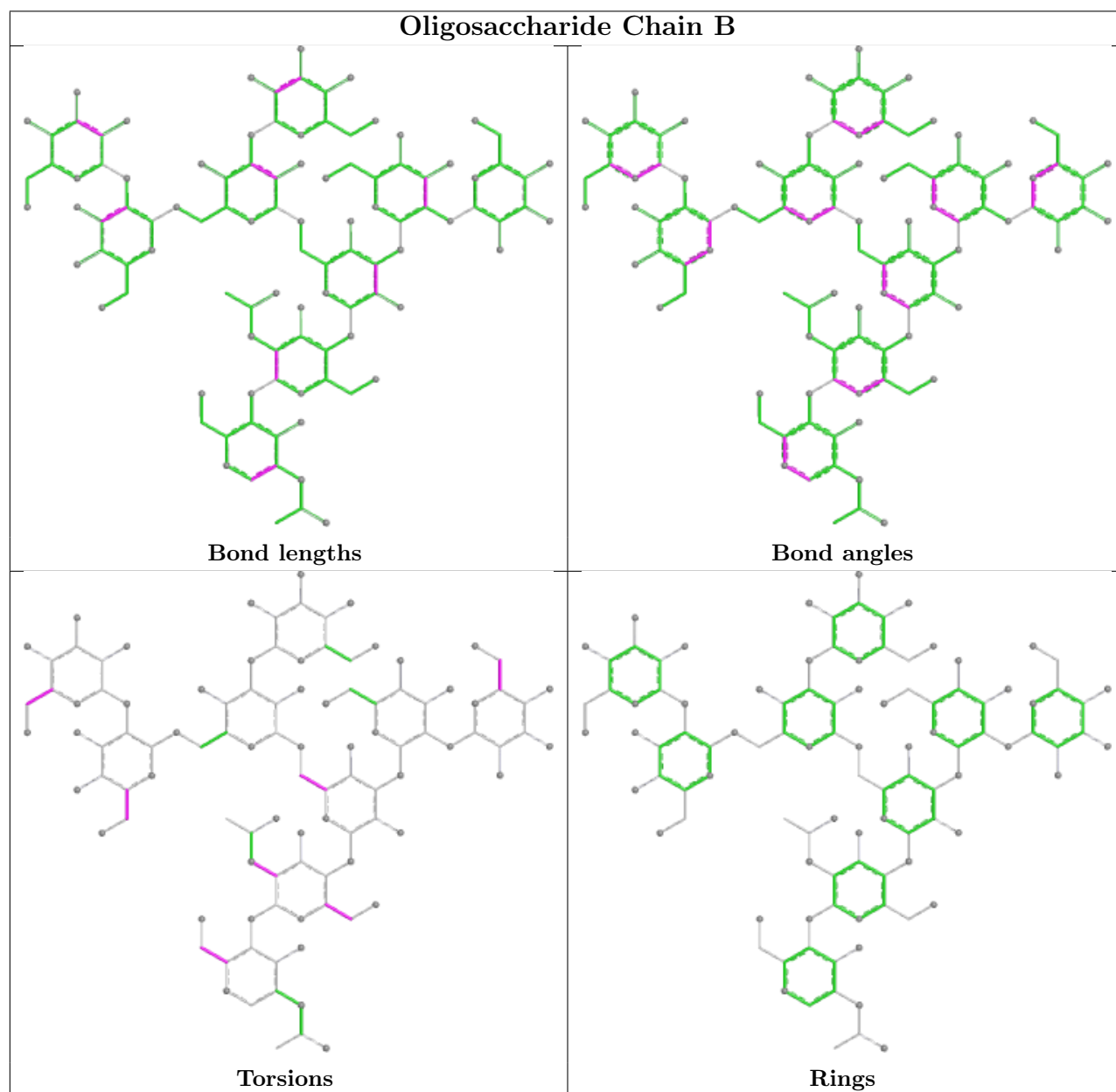| Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models Worst	Total
2	B	6	MAN	C1-O5-C5	4.35	118.01	112.19	3	18
2	B	2	NAG	C1-O5-C5	4.28	117.93	112.19	16	18
2	B	8	MAN	C1-O5-C5	4.24	117.86	112.19	18	18
2	B	9	MAN	C1-O5-C5	4.22	117.84	112.19	3	18
2	B	7	MAN	C1-O5-C5	4.21	117.82	112.19	16	18

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

Mol	Chain	Res	Type	Atoms	Models (Total)
2	B	3	BMA	C1	2
2	B	4	MAN	C1	1
2	B	8	MAN	C1	1
2	B	9	MAN	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.

Oligosaccharide Chain B

Bond lengths | Bond angles

Torsions | Rings

6.6 Ligand geometry ⓘ

There are no ligands in this entry.

6.7 Other polymers ⓘ

There are no such molecules in this entry.

6.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

7   Chemical shift validation (i)

No chemical shift data were provided