Full wwPDB NMR Structure Validation Report (i)

Jun 24, 2024 – 04:10 PM EDT

PDB ID	:	7KS6
BMRB ID	:	30815
Title	:	STRUCTURE OF TETRASACCHARIDE BUILDING BLOCK OF A SULFATED FUCAN FROM LYTECHINUS VARIEGATUS
Authors	:	Kim, S.B.; Thara, R.; Aderibigbe, A.O.; Doerksen, R.J.; Pomin, V.H.
Deposited on	:	2020-11-21

This is a Full wwPDB NMR Structure Validation Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/NMRValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (1)) were used in the production of this report:

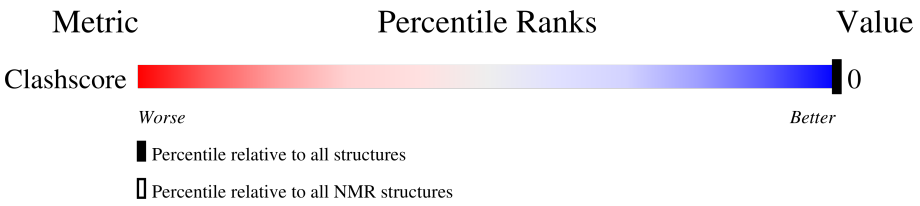
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)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

1 Overall quality at a glance (i)

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	Whole archive (#Entries)	NMR archive (#Entries)
Clashscore	158937	12864

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	B	4	100%

2 Ensemble composition and analysis (i)

This entry contains 10 models. This entry does not contain polypeptide chains, therefore identification of well-defined residues and clustering analysis are not possible. All residues are included in the validation scores.

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 98 atoms, of which 37 are hydrogens and 0 are deuteriums.

- Molecule 1 is an oligosaccharide called 4-O-sulfo-alpha-L-fucopyranose-(1-3)-2,4-di-O-sulfo-alpha-L-fucopyranose-(1-3)-2-O-sulfo-alpha-L-fucopyranose-(1-3)-2-O-sulfo-alpha-L-fucopyranose.

Mol	Chain	Residues	Atoms					Trace
			Total	C	H	O	S	
1	B	4	98	24	37	32	5	0

4 Residue-property plots (i)

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: 4-O-sulfo-alpha-L-fucopyranose-(1-3)-2,4-di-O-sulfo-alpha-L-fucopyranose-(1-3)-2-O-sulfo-alpha-L-fucopyranose-(1-3)-2-O-sulfo-alpha-L-fucopyranose

Chain B: 100%

X6Y1
X6Y2
X2Y3
X344

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: 4-O-sulfo-alpha-L-fucopyranose-(1-3)-2,4-di-O-sulfo-alpha-L-fucopyranose-(1-3)-2-O-sulfo-alpha-L-fucopyranose-(1-3)-2-O-sulfo-alpha-L-fucopyranose

Chain B: 100%

X6Y1
X6Y2
X2Y3
X344

4.2.2 Score per residue for model 2

• Molecule 1: 4-O-sulfo-alpha-L-fucopyranose-(1-3)-2,4-di-O-sulfo-alpha-L-fucopyranose-(1-3)-2-O-sulfo-alpha-L-fucopyranose-(1-3)-2-O-sulfo-alpha-L-fucopyranose

Chain B: 100%

X6Y1
X6Y2
X2Y3
X344

4.2.3 Score per residue for model 3

• Molecule 1: 4-O-sulfo-alpha-L-fucopyranose-(1-3)-2,4-di-O-sulfo-alpha-L-fucopyranose-(1-3)-2-O-sulfo-alpha-L-fucopyranose-(1-3)-2-O-sulfo-alpha-L-fucopyranose

Chain B: 100%

X6Y1
X6Y2
X2Y3
X344

4.2.4 Score per residue for model 4

• Molecule 1: 4-O-sulfo-alpha-L-fucopyranose-(1-3)-2,4-di-O-sulfo-alpha-L-fucopyranose-(1-3)-2-O-sulfo-alpha-L-fucopyranose-(1-3)-2-O-sulfo-alpha-L-fucopyranose

Chain B: 100%

X6Y1
X6Y2
X2Y3
X344

4.2.5 Score per residue for model 5

• Molecule 1: 4-O-sulfo-alpha-L-fucopyranose-(1-3)-2,4-di-O-sulfo-alpha-L-fucopyranose-(1-3)-2-O-sulfo-alpha-L-fucopyranose-(1-3)-2-O-sulfo-alpha-L-fucopyranose

Chain B: 100%

X6Y1
X6Y2
X2Y3
X344

4.2.6 Score per residue for model 6

• Molecule 1: 4-O-sulfo-alpha-L-fucopyranose-(1-3)-2,4-di-O-sulfo-alpha-L-fucopyranose-(1-3)-2-O-sulfo-alpha-L-fucopyranose-(1-3)-2-O-sulfo-alpha-L-fucopyranose

Chain B: 100%

X6Y1
X6Y2
X2Y3
X344

4.2.7 Score per residue for model 7

• Molecule 1: 4-O-sulfo-alpha-L-fucopyranose-(1-3)-2,4-di-O-sulfo-alpha-L-fucopyranose-(1-3)-2-O-sulfo-alpha-L-fucopyranose-(1-3)-2-O-sulfo-alpha-L-fucopyranose

Chain B: 100%

X6Y1
X6Y2
X2Y3
X344

4.2.8 Score per residue for model 8

• Molecule 1: 4-O-sulfo-alpha-L-fucopyranose-(1-3)-2,4-di-O-sulfo-alpha-L-fucopyranose-(1-3)-2-O-sulfo-alpha-L-fucopyranose-(1-3)-2-O-sulfo-alpha-L-fucopyranose

Chain B: 100%

X6Y1
X6Y2
X2Y3
X344

4.2.9 Score per residue for model 9

• Molecule 1: 4-O-sulfo-alpha-L-fucopyranose-(1-3)-2,4-di-O-sulfo-alpha-L-fucopyranose-(1-3)-2-O-sulfo-alpha-L-fucopyranose-(1-3)-2-O-sulfo-alpha-L-fucopyranose

Chain B: 100%

X6Y1
X6Y2
X2Y3
X344

4.2.10 Score per residue for model 10

• Molecule 1: 4-O-sulfo-alpha-L-fucopyranose-(1-3)-2,4-di-O-sulfo-alpha-L-fucopyranose-(1-3)-2-O-sulfo-alpha-L-fucopyranose-(1-3)-2-O-sulfo-alpha-L-fucopyranose

Chain B: 100%

X6Y1
X6Y2
X2Y3
X344

5 Refinement protocol and experimental data overview (i)

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

Of the 10 calculated structures, 10 were deposited, based on the following criterion: *back calculated data agree with experimental NOESY spectrum*.

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

Software name	**Classification**	**Version**
MacroModel	structure calculation	

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	48
Number of shifts mapped to atoms	0
Number of unparsed shifts	0
Number of shifts with mapping errors	48
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	0%

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: X6Y, X34, X2Y

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 (i)

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
All	All	610	370	0	-

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is -.

There are no clashes.

6.3 Torsion angles (i)

6.3.1 Protein backbone (i)

There are no protein molecules in this entry.

6.3.2 Protein sidechains (i)

There are no protein molecules in this entry.

6.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

6.5 Carbohydrates (i)

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
1	X6Y	B	1	1	15,15,15	1.23±0.02	2±0 (13±0%)
1	X6Y	B	2	1	14,14,15	1.46±0.02	3±0 (21±0%)
1	X2Y	B	3	1	18,18,19	1.47±0.01	4±0 (21±1%)
1	X34	B	4	1	14,14,15	1.32±0.03	1±0 (9±3%)

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
1	X6Y	B	1	1	20,23,23	1.26±0.02	1±0 (5±0%)
1	X6Y	B	2	1	16,21,23	1.16±0.07	0±0 (1±2%)
1	X2Y	B	3	1	20,28,30	1.37±0.13	1±0 (7±2%)
1	X34	B	4	1	18,21,23	1.03±0.04	0±0 (2±2%)

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
1	X6Y	B	1	1	-	0±0,5,25,25	0±0,1,1,1
1	X6Y	B	2	1	-	0±0,5,22,25	0±0,1,1,1
1	X2Y	B	3	1	-	0±0,10,27,30	0±0,1,1,1
1	X34	B	4	1	-	0±0,5,22,25	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
1	B	4	X34	O4-S1	3.20	1.66	1.57	5	10
1	B	3	X2Y	O2-S1	3.17	1.66	1.57	8	10
1	B	3	X2Y	O4-S2	3.11	1.66	1.57	10	10
1	B	2	X6Y	O2-S1	3.07	1.66	1.57	9	10
1	B	1	X6Y	O2-S1	3.02	1.66	1.57	8	10
1	B	2	X6Y	C1-C2	2.99	1.56	1.51	8	10
1	B	2	X6Y	O2-C2	2.69	1.43	1.47	2	10
1	B	3	X2Y	O2-C2	2.57	1.43	1.47	2	10
1	B	1	X6Y	C1-C2	2.53	1.54	1.52	6	10
1	B	3	X2Y	C1-C2	2.38	1.55	1.51	2	9
1	B	4	X34	C2-C3	2.16	1.55	1.52	10	3

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
1	B	3	X2Y	O2-C2-C3	3.99	112.53	106.95	1	10
1	B	3	X2Y	O3-C3-C2	3.88	118.88	109.32	7	3
1	B	1	X6Y	O2-C2-C1	3.75	112.63	107.58	7	10
1	B	2	X6Y	O2-C2-C3	3.32	111.58	106.95	6	3
1	B	4	X34	O2-C2-C1	2.43	114.79	109.22	6	1
1	B	4	X34	O2-C2-C3	2.10	114.51	110.15	4	4
1	B	3	X2Y	C3-C4-C5	2.03	107.32	110.30	6	1

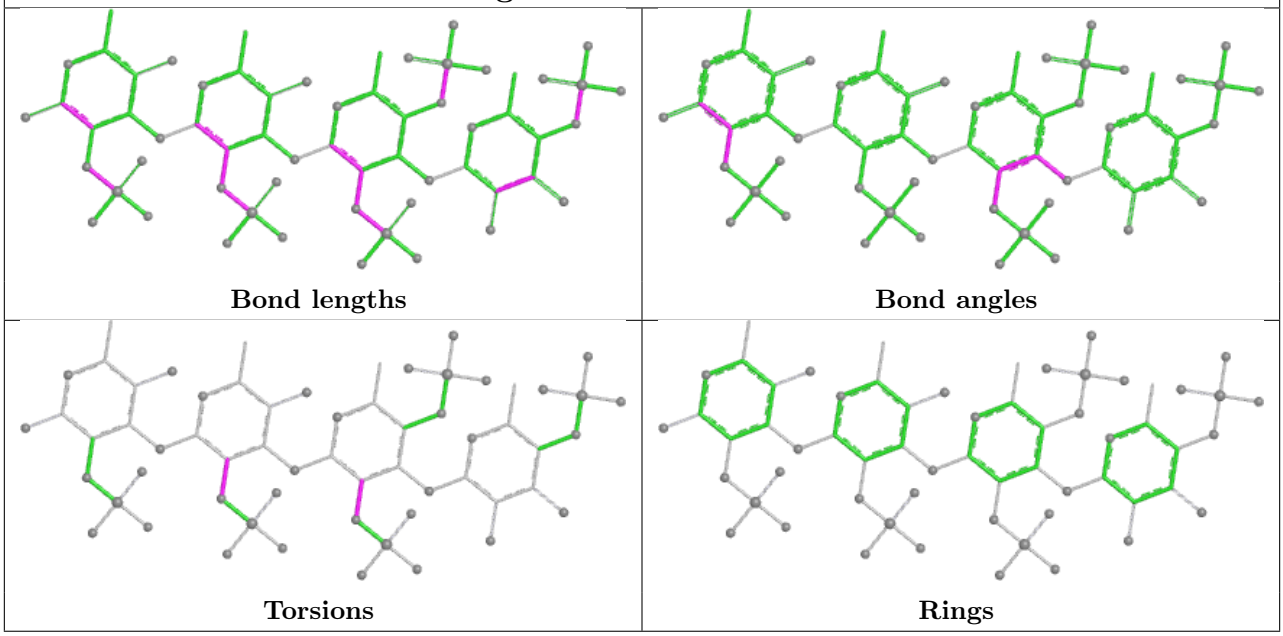
There are no chirality outliers.

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 (i)

There are no ligands in this entry.

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 (i)

The completeness of assignment taking into account all chemical shift lists is 0% for the well-defined parts and 0% for the entire structure.

7.1 Chemical shift list 1

File name: working_cs.cif

Chemical shift list name: *starch_output*

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	48
Number of shifts mapped to atoms	0
Number of unparsed shifts	0
Number of shifts with mapping errors	48
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	0

The following assigned chemical shifts were not mapped to the molecules present in the coordinate file.

- No matching atom found in the structure. All 48 occurrences are reported below.

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	?	1	X6Y	H1	5.51	.	.
1	?	1	X6Y	H2	4.52	.	.
1	?	1	X6Y	H3	4.27	.	.
1	?	1	X6Y	H4	4.11	.	.
1	?	1	X6Y	H5	4.24	.	.
1	?	1	X6Y	H6	1.18	.	.
1	?	1	X6Y	C1	90.3	.	.
1	?	1	X6Y	C2	73.1	.	.
1	?	1	X6Y	C3	73.3	.	.
1	?	1	X6Y	C4	69.0	.	.
1	?	1	X6Y	C5	66.3	.	.
1	?	1	X6Y	C6	16.5	.	.
1	?	2	X6Y	H1	5.4	.	.
1	?	2	X6Y	H2	4.54	.	.

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	?	2	X6Y	H3	4.25	.	.
1	?	2	X6Y	H4	4.85	.	.
1	?	2	X6Y	H5	4.46	.	.
1	?	2	X6Y	H6	1.25	.	.
1	?	2	X6Y	C1	94.1	.	.
1	?	2	X6Y	C2	73.1	.	.
1	?	2	X6Y	C3	74.1	.	.
1	?	2	X6Y	C4	79.6	.	.
1	?	2	X6Y	C5	66.3	.	.
1	?	2	X6Y	C6	15.2	.	.
1	?	3	X2Y	H1	5.36	.	.
1	?	3	X2Y	H2	4.57	.	.
1	?	3	X2Y	H3	4.13	.	.
1	?	3	X2Y	H4	4.11	.	.
1	?	3	X2Y	H5	4.11	.	.
1	?	3	X2Y	H6	1.18	.	.
1	?	3	X2Y	C1	94.6	.	.
1	?	3	X2Y	C2	73.1	.	.
1	?	3	X2Y	C3	73.3	.	.
1	?	3	X2Y	C4	69.0	.	.
1	?	3	X2Y	C5	65.6	.	.
1	?	3	X2Y	C6	16.5	.	.
1	?	4	X34	H1	5.14	.	.
1	?	4	X34	H2	3.75	.	.
1	?	4	X34	H3	4.04	.	.
1	?	4	X34	H4	4.63	.	.
1	?	4	X34	H5	4.55	.	.
1	?	4	X34	H6	1.29	.	.
1	?	4	X34	C1	98.1	.	.
1	?	4	X34	C2	68.6	.	.
1	?	4	X34	C3	73.3	.	.
1	?	4	X34	C4	80.8	.	.
1	?	4	X34	C5	66.3	.	.
1	?	4	X34	C6	15.8	.	.

7.1.2 Chemical shift referencing (i)

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

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

	Total	^{1}H	^{13}C	^{15}N

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

	Total	^{1}H	^{13}C	^{15}N

7.1.4 Statistically unusual chemical shifts (i)

There are no statistically unusual chemical shifts.

7.1.5 Random Coil Index (RCI) plots (i)

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