wwPDB NMR Structure Validation Summary Report

Jul 28, 2022 – 08:41 pm BST

PDB ID	:	7QGV
Title	:	Solid-state NMR structure of Teixobactin-Lipid II
Authors	:	Weingarth, M.H.; Shukla, R.
Deposited on	:	2021-12-10

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

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/NMRValidationReportHelp
with specific help available everywhere you see the ⓘ 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 ⓘ) were used in the production of this report:

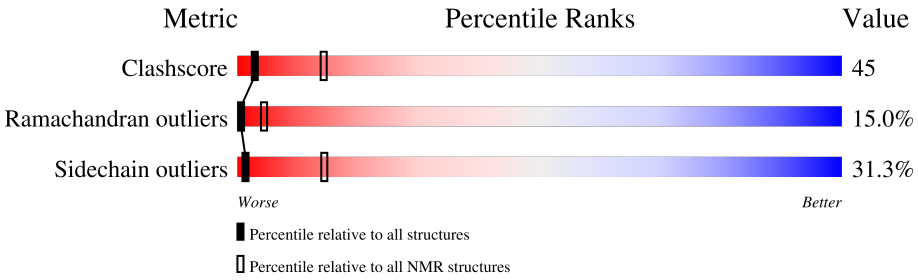
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	2.29
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.29

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
SOLID-STATE NMR

The overall completeness of chemical shifts assignment is 11%.

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
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	11	9% 45% 45%
1	B	11	9% 45% 45%
1	C	11	9% 45% 45%
1	D	11	18% 36% 45%
3	I	2	50% 50%
3	J	2	50% 50%
3	K	2	50% 50%
3	L	2	50% 50%

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
3	I	MUB	1	15	-
3	J	MUB	1	13	-
3	K	MUB	1	16	-
3	L	MUB	1	18	-

2 Ensemble composition and analysis (i)

This entry contains 25 models. Model 3 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *medoid*.

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:2-A:3, A:6-A:7, A:9-A:9, A:11-A:11, B:2-B:3, B:6-B:7, B:9-B:9, B:11-B:11, C:2-C:3, C:6-C:7, C:9-C:9, C:11-C:11, D:2-D:3, D:6-D:7, D:9-D:9, D:11-D:11 (24)	1.39	3

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 4 clusters and 1 single-model cluster was found.

Cluster number	Models
1	1, 2, 3, 4, 5, 6, 9, 11, 13, 15, 16, 17, 19, 25
2	10, 14, 23, 24
3	7, 8, 12, 18
4	21, 22
Single-model clusters	20

3 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 1445 atoms, of which 737 are hydrogens and 0 are deuteriums.

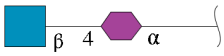
- Molecule 1 is a protein called Teixobactin.

Mol	Chain	Residues	Atoms					Trace
1	A	11	Total	C	H	N	O	0
			184	58	96	15	15	
1	B	11	Total	C	H	N	O	0
			184	58	96	15	15	
1	C	11	Total	C	H	N	O	0
			184	58	96	15	15	
1	D	11	Total	C	H	N	O	0
			184	58	96	15	15	

- Molecule 2 is a protein (with D amino acids) called Lipid II.

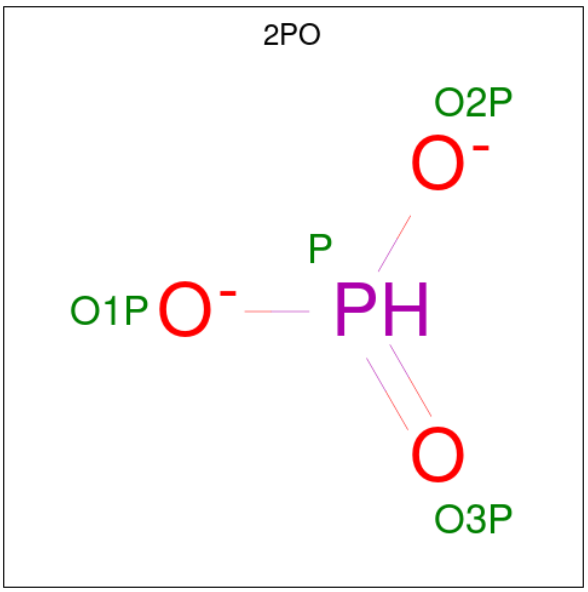
Mol	Chain	Residues	Atoms					Trace
2	E	5	Total	C	H	N	O	0
			66	20	33	6	7	
2	F	5	Total	C	H	N	O	0
			66	20	33	6	7	
2	G	5	Total	C	H	N	O	0
			66	20	33	6	7	
2	H	5	Total	C	H	N	O	0
			67	20	34	6	7	

- Molecule 3 is an oligosaccharide called 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-N-acetyl-alpha-muramic acid.



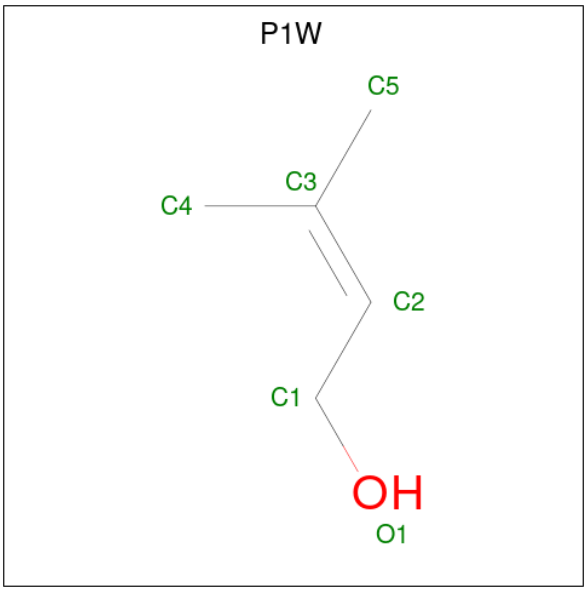
Mol	Chain	Residues	Atoms					Trace
3	I	2	Total	C	H	N	O	0
			63	19	30	2	12	
3	J	2	Total	C	H	N	O	0
			63	19	30	2	12	
3	K	2	Total	C	H	N	O	0
			63	19	30	2	12	
3	L	2	Total	C	H	N	O	0
			63	19	30	2	12	

- Molecule 4 is PHOSPHONATE (three-letter code: 2PO) (formula: HO_3P).



Mol	Chain	Residues	Atoms		
4	A	1	Total	O	P
			4	3	1
4	A	1	Total	O	P
			4	3	1
4	B	1	Total	O	P
			4	3	1
4	B	1	Total	O	P
			4	3	1
4	C	1	Total	O	P
			4	3	1
4	C	1	Total	O	P
			4	3	1
4	D	1	Total	O	P
			4	3	1
4	D	1	Total	O	P
			4	3	1

- Molecule 5 is 3-methylbut-2-en-1-ol (three-letter code: P1W) (formula: $C_5H_{10}O$).



Mol	Chain	Residues	Atoms		
5	A	1	Total	C	H
			13	5	8
5	A	1	Total	C	H
			13	5	8
5	A	1	Total	C	H
			14	5	9
5	B	1	Total	C	H
			14	5	9
5	B	1	Total	C	H
			13	5	8
5	B	1	Total	C	H
			13	5	8
5	C	1	Total	C	H
			14	5	9
5	C	1	Total	C	H
			14	5	9
5	C	1	Total	C	H
			13	5	8
5	C	1	Total	C	H
			13	5	8
5	D	1	Total	C	H
			13	5	8
5	D	1	Total	C	H
			13	5	8

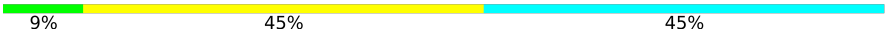
4 Residue-property plots

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.

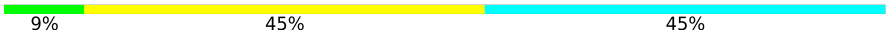
• Molecule 1: Teixobactin

Chain A:



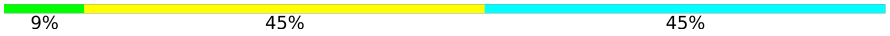
• Molecule 1: Teixobactin

Chain B:



• Molecule 1: Teixobactin

Chain C:



• Molecule 1: Teixobactin

Chain D:

• Molecule 3: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-N-acetyl-alpha-muramic acid

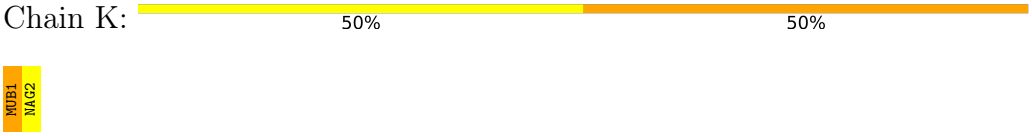
Chain I:

• Molecule 3: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-N-acetyl-alpha-muramic acid

Chain J:

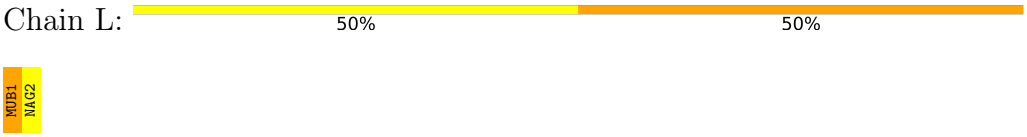
• Molecule 3: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-N-acetyl-alpha-muramic acid

Chain K:



• Molecule 3: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-N-acetyl-alpha-muramic acid

Chain L:

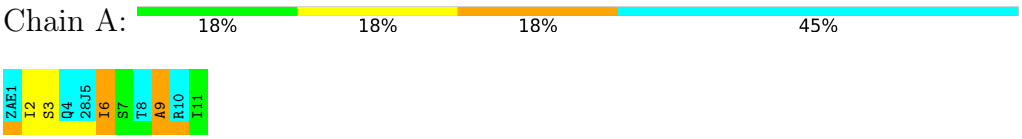


4.2 Residue scores for the representative (medoid) model from the NMR ensemble

The representative model is number 3. Colouring as in section 4.1 above.

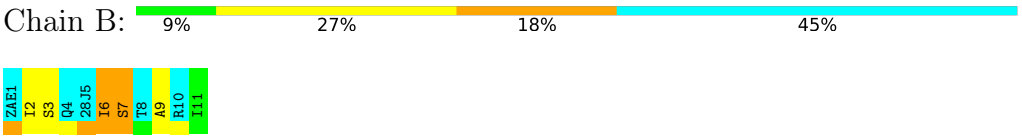
• Molecule 1: Teixobactin

Chain A:



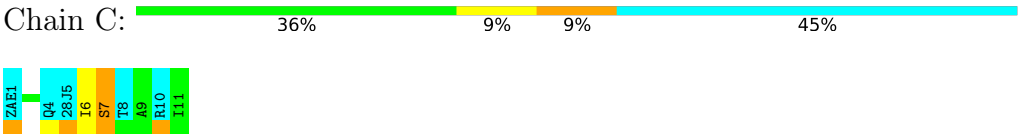
• Molecule 1: Teixobactin

Chain B:



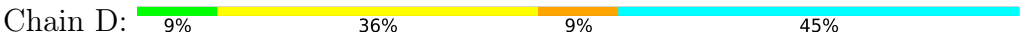
• Molecule 1: Teixobactin

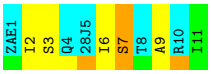
Chain C:



• Molecule 1: Teixobactin

Chain D:





• Molecule 3: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-N-acetyl-alpha-muramic acid

Chain I:

• Molecule 3: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-N-acetyl-alpha-muramic acid

Chain J:

• Molecule 3: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-N-acetyl-alpha-muramic acid

Chain K:

• Molecule 3: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-N-acetyl-alpha-muramic acid

Chain L:

5 Refinement protocol and experimental data overview

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

Of the 25 calculated structures, 25 were deposited, based on the following criterion: *all calculated structures submitted*.

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

Software name	Classification	Version
HADDOCK	structure calculation	2.4
CNS	refinement	

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	77
Number of shifts mapped to atoms	76
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	1
Assignment completeness (well-defined parts)	11%

Note: This is a solid-state NMR structure, where hydrogen atoms are typically not assigned a chemical shift value, which may lead to lower completeness of assignment measure.

6 Model quality

6.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: P1W, DGN, DAL, NAG, 28J, DTH, DGL, MUB, 2PO, ZAE, EI4

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	0.94±0.19	0±0/37 (0.0± 0.0%)	1.23±0.21	0±0/44 (0.2± 0.6%)
1	B	1.08±0.22	0±0/37 (0.0± 0.0%)	1.42±0.22	0±0/44 (0.2± 0.6%)
1	C	1.12±0.19	0±0/37 (0.0± 0.0%)	1.39±0.21	0±0/44 (0.0± 0.0%)
1	D	0.86±0.17	0±0/37 (0.1± 0.5%)	1.18±0.13	0±0/44 (0.0± 0.0%)
2	E	1.78±0.18	0±0/12 (0.0± 0.0%)	2.69±0.21	1±0/12 (9.3± 2.7%)
2	F	1.82±0.20	0±0/12 (0.0± 0.0%)	2.71±0.23	1±0/12 (8.7± 1.6%)
2	G	1.85±0.22	0±0/12 (0.7± 2.3%)	2.62±0.28	1±0/12 (8.7± 1.6%)
2	H	1.83±0.20	0±0/12 (0.0± 0.0%)	2.65±0.27	1±0/12 (8.3± 0.0%)
All	All	1.27	3/4900 (0.1%)	1.71	109/5600 (1.9%)

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	G	3	LYS	CD-CE	5.29	1.64	1.51	9	2
1	D	3	SER	CA-CB	-5.09	1.45	1.52	12	1

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	E	3	LYS	CD-CE-NZ	9.09	132.61	111.70	12	25
2	G	3	LYS	CD-CE-NZ	8.93	132.24	111.70	12	25
2	F	3	LYS	CD-CE-NZ	8.77	131.88	111.70	12	25
2	H	3	LYS	CD-CE-NZ	8.70	131.70	111.70	12	25
1	A	3	SER	N-CA-CB	6.33	120.00	110.50	11	2

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
1	A	41	48	48	4±3
1	B	41	48	48	4±3
1	C	41	48	48	4±3
1	D	41	48	48	4±2
2	E	33	33	30	2±2
2	F	33	33	31	2±2
2	G	33	33	30	2±2
2	H	33	34	29	2±3
3	I	33	30	29	5±2
3	J	33	30	29	6±2
3	K	33	30	29	7±2
3	L	33	30	29	6±2
4	A	8	0	0	2±1
4	B	8	0	0	2±2
4	C	8	0	0	2±1
4	D	8	0	0	2±2
5	B	15	25	0	1±2
5	C	20	34	0	1±1
5	D	10	16	0	0±1
5	A	15	25	0	0±1
All	All	13000	13625	10653	1058

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 45.

5 of 422 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
2:H:4:DAL:CA	2:H:4:DAL:CB	1.65	1.75	24	3
2:F:4:DAL:CA	2:F:4:DAL:CB	1.63	1.74	24	3
2:H:5:DAL:CA	2:H:5:DAL:C	1.62	1.74	25	1
2:G:4:DAL:CA	2:G:4:DAL:CB	1.61	1.79	23	4
2:E:4:DAL:CA	2:E:4:DAL:CB	1.60	1.74	24	3

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	5/11 (45%)	3±1 (67±12%)	1±1 (18±12%)	1±1 (15±13%)	0 4
1	B	5/11 (45%)	3±1 (61±11%)	1±1 (24±13%)	1±1 (15±12%)	0 4
1	C	5/11 (45%)	3±1 (63±11%)	1±1 (25±13%)	1±1 (12±11%)	1 7
1	D	5/11 (45%)	3±1 (67±14%)	1±1 (15±13%)	1±1 (18±10%)	0 3
2	E	0	-	-	-	-
2	F	0	-	-	-	-
2	G	0	-	-	-	-
2	H	0	-	-	-	-
All	All	500/1600 (31%)	323 (65%)	102 (20%)	75 (15%)	0 4

5 of 9 unique Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
1	B	7	SER	14
1	D	7	SER	14
1	A	7	SER	12
1	C	7	SER	11
1	D	9	ALA	8

6.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all 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	5/5 (100%)	4±1 (82±19%)	1±1 (18±19%)	4 39
1	B	5/5 (100%)	4±1 (71±23%)	1±1 (29±23%)	2 18
1	C	5/5 (100%)	4±1 (78±14%)	1±1 (22±14%)	3 29

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	D	5/5 (100%)	4±1 (82±19%)	1±1 (18±19%)	4 37
2	E	1/1 (100%)	0±0 (32±47%)	1±0 (68±47%)	0 0
2	F	1/1 (100%)	0±0 (20±40%)	1±0 (80±40%)	0 0
2	G	1/1 (100%)	0±0 (20±40%)	1±0 (80±40%)	0 0
2	H	1/1 (100%)	0±0 (12±32%)	1±0 (88±32%)	0 0
All	All	600/600 (100%)	412 (69%)	188 (31%)	1 14

5 of 24 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)
2	H	3	LYS	22
2	G	3	LYS	20
2	F	3	LYS	20
2	E	3	LYS	17
1	C	2	ILE	14

6.3.3 RNA (i)

There are no RNA molecules in this entry.

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

32 non-standard protein/DNA/RNA residues 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	28J	A	5	1	6,7,8	0.86±0.14	0±0 (2±5%)
1	28J	B	5	1	6,7,8	1.04±0.18	0±1 (8±11%)
1	EI4	A	10	1	8,11,12	2.04±0.21	3±1 (41±8%)
1	28J	D	5	1	6,7,8	0.92±0.15	0±0 (2±5%)

Mol	Type	Chain	Res	Link	Bond lengths		
					Counts	RMSZ	#Z>2
1	EI4	D	10	1	8,11,12	1.87±0.16	3±1 (39±12%)
1	EI4	C	10	1	8,11,12	1.98±0.16	3±1 (37±9%)
1	ZAE	B	1	1	11,12,13	0.65±0.17	0±0 (1±3%)
1	ZAE	C	1	1	11,12,13	0.67±0.20	0±0 (1±4%)
1	EI4	B	10	1	8,11,12	2.05±0.16	3±1 (43±8%)
1	ZAE	D	1	1	11,12,13	0.59±0.10	0±0 (0±1%)
1	ZAE	A	1	1	11,12,13	0.54±0.12	0±0 (0±0%)
1	28J	C	5	1	6,7,8	0.94±0.17	0±0 (3±6%)

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	28J	A	5	1	5,8,10	1.39±0.22	1±1 (18±11%)
1	28J	B	5	1	5,8,10	1.68±0.37	1±0 (22±8%)
1	EI4	A	10	1	4,14,16	1.00±0.27	0±0 (3±8%)
1	28J	D	5	1	5,8,10	1.34±0.18	1±1 (15±11%)
1	EI4	D	10	1	4,14,16	0.91±0.16	0±0 (2±6%)
1	EI4	C	10	1	4,14,16	0.94±0.22	0±0 (2±9%)
1	ZAE	B	1	1	13,14,16	1.01±0.45	1±1 (6±7%)
1	ZAE	C	1	1	13,14,16	0.95±0.27	1±1 (6±7%)
1	EI4	B	10	1	4,14,16	0.88±0.14	0±0 (0±0%)
1	ZAE	D	1	1	13,14,16	0.72±0.16	0±0 (1±2%)
1	ZAE	A	1	1	13,14,16	0.69±0.14	0±0 (1±2%)
1	28J	C	5	1	5,8,10	1.95±0.50	1±0 (21±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
1	EI4	A	10	1	-	0±0,5,15,17	0±0,1,1,1
1	EI4	B	10	1	-	0±0,5,15,17	0±0,1,1,1
1	ZAE	B	1	1	-	0±0,5,8,10	0±0,1,1,1
1	EI4	C	10	1	-	0±0,5,15,17	0±0,1,1,1
1	ZAE	D	1	1	-	0±0,5,8,10	0±0,1,1,1
1	ZAE	C	1	1	-	0±0,5,8,10	0±0,1,1,1
1	28J	B	5	1	-	0±0,7,8,10	-
1	EI4	D	10	1	-	0±0,5,15,17	0±0,1,1,1
1	28J	C	5	1	-	0±0,7,8,10	-
1	28J	D	5	1	-	0±0,7,8,10	-
1	28J	A	5	1	-	0±0,7,8,10	-
1	ZAE	A	1	1	-	0±0,5,8,10	0±0,1,1,1

5 of 35 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	10	EI4	CG-ND1	5.19	1.52	1.46	22	25
1	A	10	EI4	CG-ND1	4.96	1.52	1.46	13	24
1	C	10	EI4	CG-ND1	4.79	1.52	1.46	1	22
1	D	10	EI4	CG-ND1	4.59	1.52	1.46	5	23
1	C	10	EI4	CE1-NE2	4.13	1.41	1.35	11	21

5 of 33 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	1	ZAE	CG-CB-CA	7.74	124.71	113.63	18	7
1	C	5	28J	CB-CA-C	6.50	102.89	112.83	18	23
1	B	5	28J	CB-CA-C	5.24	104.81	112.83	17	22
1	C	1	ZAE	CG-CB-CA	4.59	120.19	113.63	2	10
1	B	1	ZAE	CB-CA-N	4.38	103.86	110.65	7	4

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

6.5 Carbohydrates (i)

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	Bond lengths		
					Counts	RMSZ	#Z>2
3	MUB	I	1	3,4,2	18,19,20	1.00±0.13	1±0 (6±2%)
3	NAG	I	2	3	14,14,15	0.62±0.05	0±0 (0±1%)
3	MUB	J	1	3,4,2	18,19,20	0.97±0.07	1±0 (6±2%)
3	NAG	J	2	3	14,14,15	0.63±0.06	0±0 (0±1%)
3	MUB	K	1	3,4,2	18,19,20	0.97±0.10	1±0 (5±1%)
3	NAG	K	2	3	14,14,15	0.63±0.05	0±0 (0±1%)
3	MUB	L	1	3,4,2	18,19,20	0.96±0.09	1±0 (6±1%)
3	NAG	L	2	3	14,14,15	0.63±0.05	0±0 (0±1%)

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	MUB	I	1	3,4,2	21,26,28	1.36±0.08	1±1 (6±2%)
3	NAG	I	2	3	17,19,21	0.61±0.05	0±0 (0±0%)
3	MUB	J	1	3,4,2	21,26,28	1.39±0.13	2±1 (7±3%)
3	NAG	J	2	3	17,19,21	0.58±0.08	0±0 (0±0%)
3	MUB	K	1	3,4,2	21,26,28	1.38±0.10	1±1 (6±2%)
3	NAG	K	2	3	17,19,21	0.58±0.07	0±0 (0±0%)
3	MUB	L	1	3,4,2	21,26,28	1.33±0.07	1±1 (6±2%)
3	NAG	L	2	3	17,19,21	0.60±0.08	0±0 (0±1%)

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	MUB	I	1	3,4,2	1±0,1,8,9	0±0,10,32,34	0±0,1,1,1
3	NAG	I	2	3	-	0±0,6,23,26	0±0,1,1,1
3	MUB	J	1	3,4,2	1±0,1,8,9	0±0,10,32,34	0±0,1,1,1
3	NAG	J	2	3	-	0±0,6,23,26	0±0,1,1,1
3	MUB	K	1	3,4,2	1±0,1,8,9	0±0,10,32,34	0±0,1,1,1
3	NAG	K	2	3	-	0±0,6,23,26	0±0,1,1,1
3	MUB	L	1	3,4,2	1±0,1,8,9	0±0,10,32,34	0±0,1,1,1
3	NAG	L	2	3	-	0±0,6,23,26	0±0,1,1,1

5 of 14 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
3	I	1	MUB	C1-C2	4.12	1.48	1.52	4	1
3	K	1	MUB	C11-C9	3.81	1.59	1.51	13	24
3	I	1	MUB	C11-C9	3.53	1.58	1.51	4	24
3	J	1	MUB	C11-C9	3.52	1.58	1.51	13	24
3	L	1	MUB	C11-C9	3.42	1.58	1.51	13	25

5 of 24 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
3	J	1	MUB	O3-C9-C11	5.98	123.67	107.48	5	25
3	I	1	MUB	O3-C9-C11	5.20	121.54	107.48	11	25
3	K	1	MUB	O3-C9-C11	5.04	121.11	107.48	25	25
3	L	1	MUB	O3-C9-C11	4.94	120.86	107.48	10	25
3	J	1	MUB	C1-C2-N2	3.48	114.76	110.73	7	7

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)
3	L	1	MUB	C9	18
3	K	1	MUB	C9	16
3	I	1	MUB	C9	15
3	J	1	MUB	C9	13

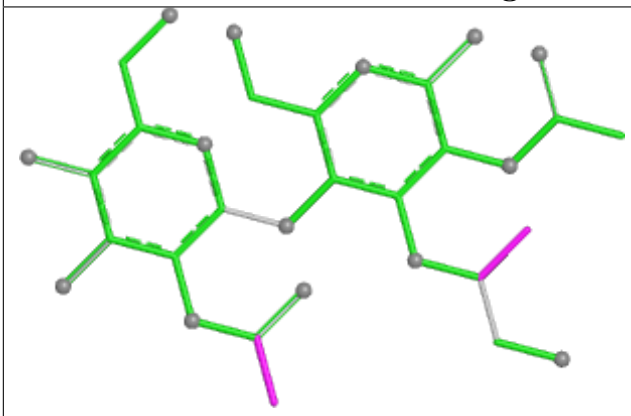
There are no torsion outliers.

There are no ring outliers.

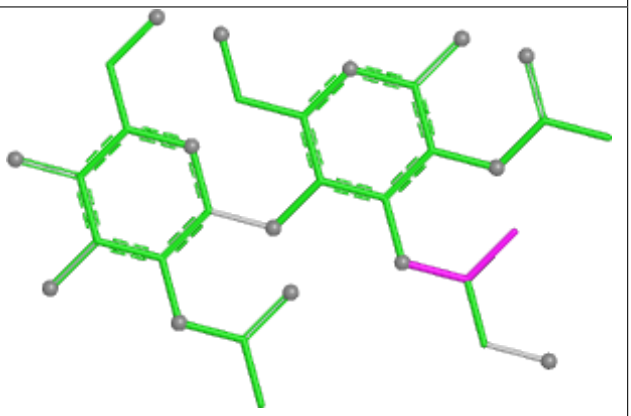
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths,

bond angles, torsion angles, and ring geometry for oligosaccharide.

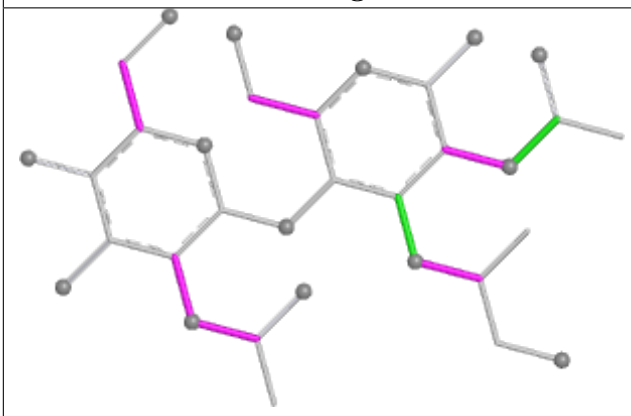
Oligosaccharide Chain I



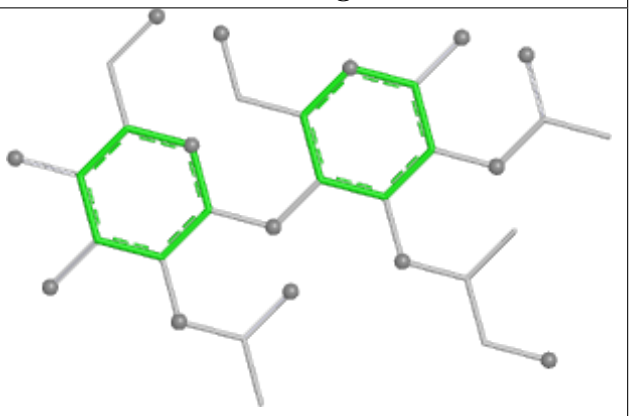
Bond lengths



Bond angles

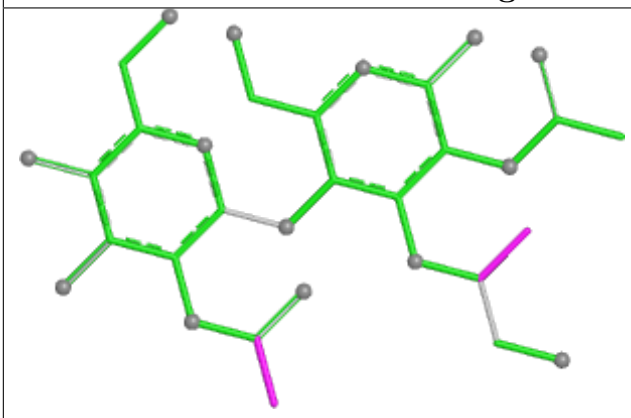


Torsions

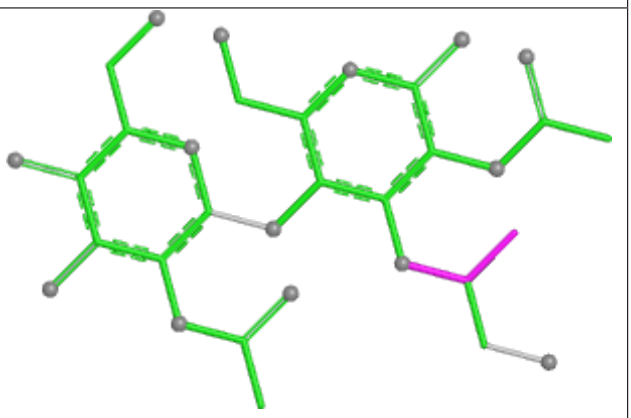


Rings

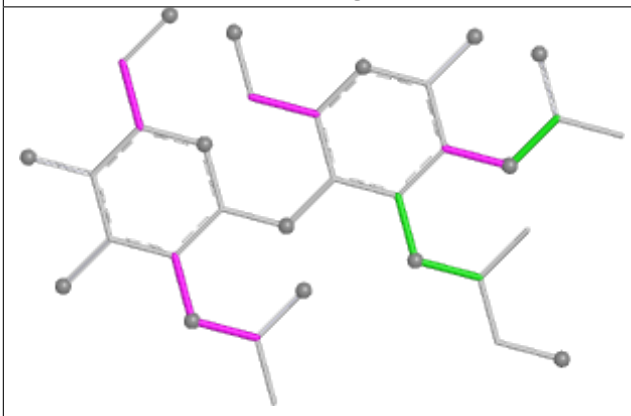
Oligosaccharide Chain J



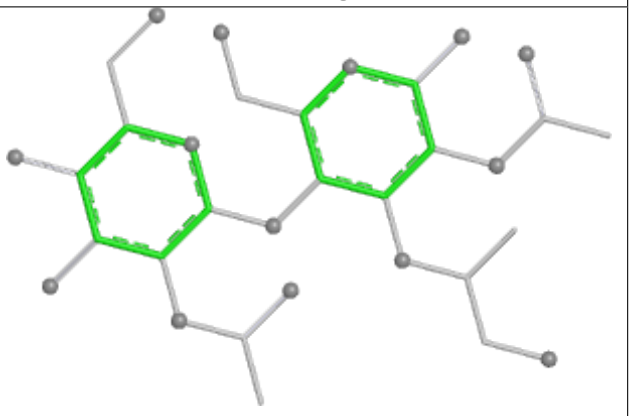
Bond lengths



Bond angles

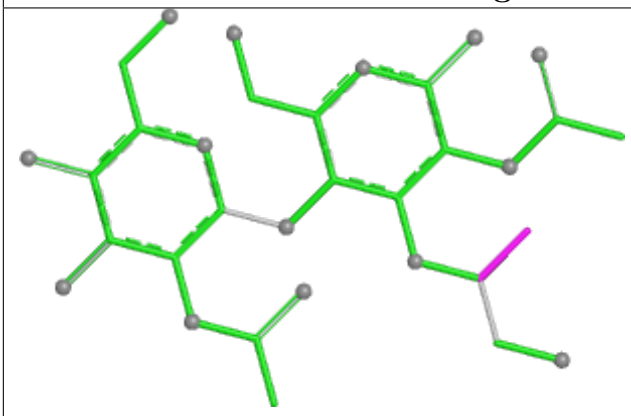


Torsions

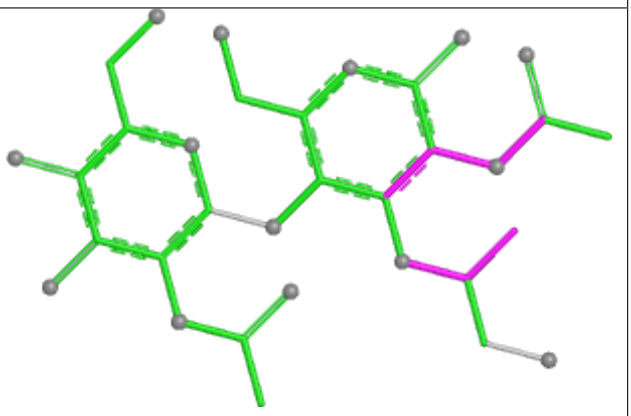


Rings

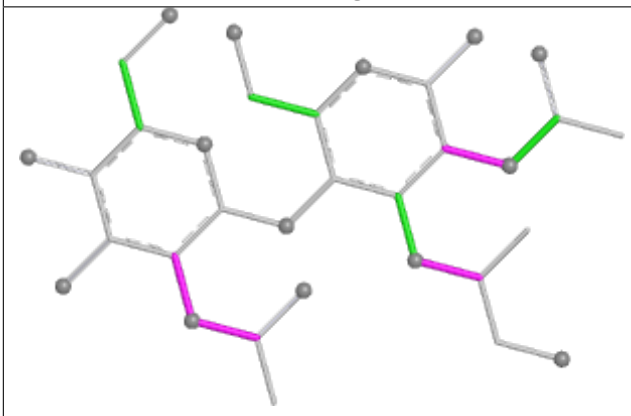
Oligosaccharide Chain K



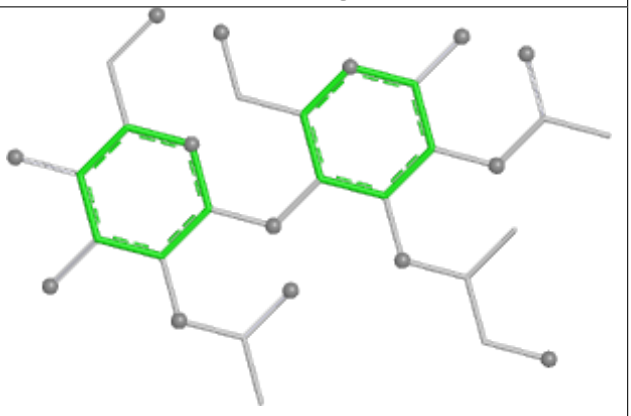
Bond lengths



Bond angles

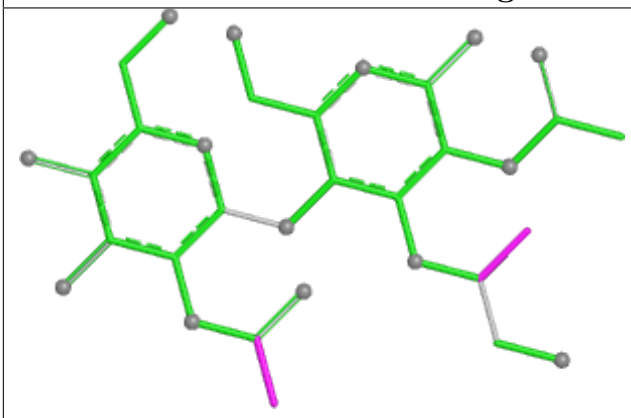


Torsions

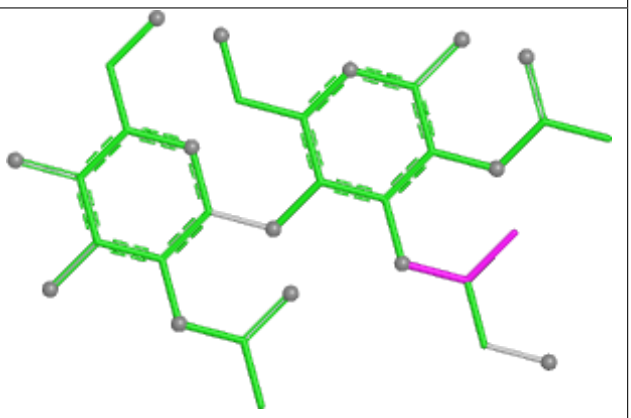


Rings

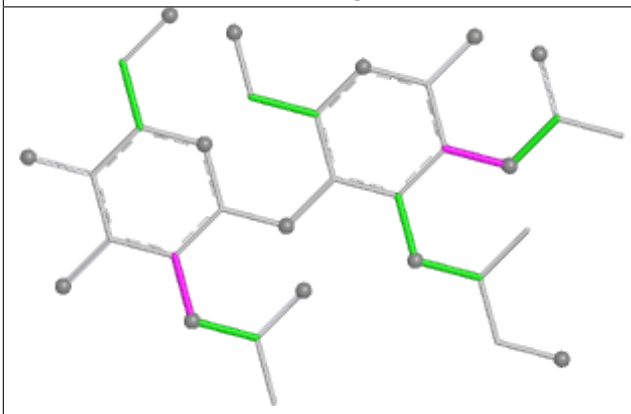
Oligosaccharide Chain L



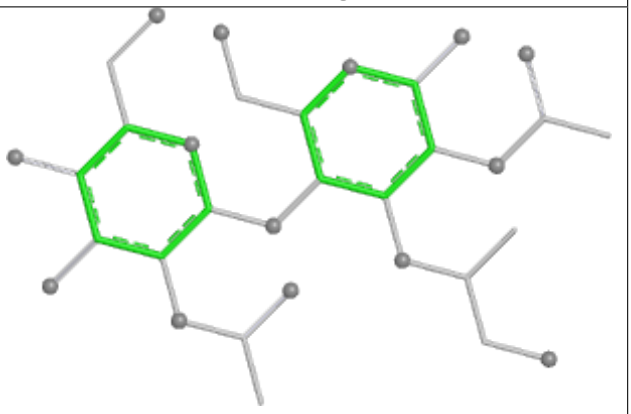
Bond lengths



Bond angles



Torsions



Rings

6.6 Ligand geometry (i)

20 ligands 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
5	P1W	A	103	4,5	4,4,5	0.80±0.30	0±0 (2±6%)
5	P1W	B	105	5	4,4,5	0.77±0.15	0±0 (0±0%)
5	P1W	C	103	5	4,4,5	0.73±0.09	0±0 (0±0%)
4	2PO	B	102	3,4	0,3,3	0.00±0.00	-
4	2PO	A	101	3,4	0,3,3	0.00±0.00	-
5	P1W	B	104	4,5	4,4,5	0.85±0.28	0±0 (4±9%)

Mol	Type	Chain	Res	Link	Bond lengths		
					Counts	RMSZ	#Z>2
5	P1W	A	104	5	4,4,5	0.76±0.16	0±0 (1±4%)
5	P1W	D	103	4,5	4,4,5	0.90±0.26	0±0 (2±6%)
4	2PO	C	105	4,5	0,3,3	0.00±0.00	-
4	2PO	D	102	4,5	0,3,3	0.00±0.00	-
4	2PO	B	103	4,5	0,3,3	0.00±0.00	-
5	P1W	A	105	5	4,4,5	0.87±0.12	0±0 (0±0%)
4	2PO	C	104	3,4	0,3,3	0.00±0.00	-
5	P1W	D	104	5	4,4,5	0.75±0.11	0±0 (0±0%)
4	2PO	D	101	3,4	0,3,3	0.00±0.00	-
5	P1W	C	102	5	4,4,5	0.88±0.12	0±0 (0±0%)
5	P1W	B	101	5	4,4,5	0.85±0.16	0±0 (2±6%)
5	P1W	C	106	4,5	4,4,5	0.76±0.24	0±0 (1±4%)
4	2PO	A	102	4,5	0,3,3	0.00±0.00	-
5	P1W	C	101	5	4,4,5	0.94±0.18	0±0 (4±9%)

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
5	P1W	A	103	4,5	4,4,5	0.59±0.28	0±0 (0±0%)
5	P1W	B	105	5	4,4,5	0.56±0.14	0±0 (0±0%)
5	P1W	C	103	5	4,4,5	0.59±0.13	0±0 (0±0%)
4	2PO	B	102	3,4	0,3,3	0.00±0.00	-
4	2PO	A	101	3,4	0,3,3	0.00±0.00	-
5	P1W	B	104	4,5	4,4,5	0.61±0.25	0±0 (0±0%)
5	P1W	A	104	5	4,4,5	0.57±0.11	0±0 (0±0%)
5	P1W	D	103	4,5	4,4,5	0.52±0.34	0±0 (2±9%)
4	2PO	C	105	4,5	0,3,3	0.00±0.00	-
4	2PO	D	102	4,5	0,3,3	0.00±0.00	-
4	2PO	B	103	4,5	0,3,3	0.00±0.00	-
5	P1W	A	105	5	4,4,5	0.59±0.17	0±0 (1±4%)
4	2PO	C	104	3,4	0,3,3	0.00±0.00	-
5	P1W	D	104	5	4,4,5	0.54±0.13	0±0 (0±0%)
4	2PO	D	101	3,4	0,3,3	0.00±0.00	-
5	P1W	C	102	5	4,4,5	0.57±0.15	0±0 (0±0%)
5	P1W	B	101	5	4,4,5	0.65±0.15	0±0 (0±0%)

Mol	Type	Chain	Res	Link	Bond angles		
					Counts	RMSZ	#Z>2
5	P1W	C	106	4,5	4,4,5	0.60±0.25	0±0 (0±0%)
4	2PO	A	102	4,5	0,3,3	0.00±0.00	-
5	P1W	C	101	5	4,4,5	0.58±0.18	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
5	P1W	A	104	5	-	0±0,2,2,3	-
5	P1W	A	105	5	-	0±0,2,2,3	-
5	P1W	D	103	4,5	-	0±0,2,2,3	-
5	P1W	B	105	5	-	0±0,2,2,3	-
5	P1W	A	103	4,5	-	0±0,2,2,3	-
5	P1W	D	104	5	-	0±0,2,2,3	-
5	P1W	C	106	4,5	-	0±0,2,2,3	-
5	P1W	C	101	5	-	0±0,2,2,3	-
5	P1W	B	104	4,5	-	0±0,2,2,3	-
5	P1W	B	101	5	-	0±0,2,2,3	-
5	P1W	C	103	5	-	0±0,2,2,3	-
5	P1W	C	102	5	-	0±0,2,2,3	-

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
5	D	103	P1W	C2-C3	2.65	1.22	1.33	5	1
5	A	103	P1W	C5-C3	2.56	1.57	1.50	4	2
5	B	104	P1W	C5-C3	2.45	1.57	1.50	13	4
5	D	103	P1W	C4-C3	2.35	1.56	1.50	2	1
5	C	101	P1W	C4-C3	2.23	1.56	1.50	1	4

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
5	D	103	P1W	C1-C2-C3	2.87	118.20	126.31	5	1
5	D	103	P1W	C5-C3-C4	2.15	119.36	114.60	5	1
5	A	105	P1W	C1-C2-C3	2.01	120.62	126.31	15	1

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

The following chains have linkage breaks:

Mol	Chain	Number of breaks
2	23-F	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
23	F	2:DGL	C	3:LYS	N	2.99

7 Chemical shift validation

The completeness of assignment taking into account all chemical shift lists is 11% for the well-defined parts and 11% 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

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	77
Number of shifts mapped to atoms	76
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	1
Number of shift outliers (ShiftChecker)	0

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

- No matching atoms found in structure. The only occurence is reported below.

Chain	Res	Type	Atom	Shift Data		
				Value	Uncertainty	Ambiguity
A	5	28J	H	8.5	-1.0	1

7.1.2 Chemical shift referencing

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

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 11%, i.e. 39 atoms were assigned a chemical shift out of a possible 360. 0 out of 0 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

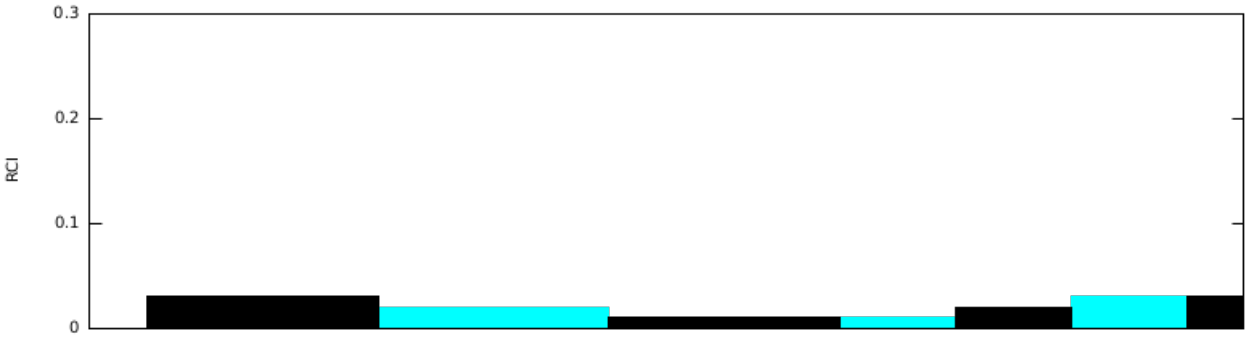
	Total	^{1}H	^{13}C	^{15}N
Backbone	24/160 (15%)	6/64 (9%)	12/64 (19%)	6/32 (19%)
Sidechain	15/200 (8%)	0/116 (0%)	15/80 (19%)	0/4 (0%)
Aromatic	0/0 (—%)	0/0 (—%)	0/0 (—%)	0/0 (—%)
Overall	39/360 (11%)	6/180 (3%)	27/144 (19%)	6/36 (17%)

Note: This is a solid-state NMR structure, where hydrogen atoms are typically not assigned a chemical shift value, which may lead to lower completeness of assignment measure.

7.1.4 Statistically unusual chemical shifts (i)

There are no statistically unusual chemical shifts.

7.1.5 Random Coil Index (RCI) plots (i)

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: