Full wwPDB X-ray Structure Validation Report (i)

Aug 8, 2020 – 10:07 AM BST

PDB ID	:	2PIL
Title	:	Crystallographic Structure of Phosphorylated Pilin from Neisseria: Phosphoserine Sites Modify Type IV Pilus Surface Chemistry
Authors	:	Forest, K.T.; Dunham, S.A.; Koomey, M.; Tainer, J.A.
Deposited on	:	1998-03-02
Resolution	:	2.60 Å(reported)

This is a Full wwPDB X-ray 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/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

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

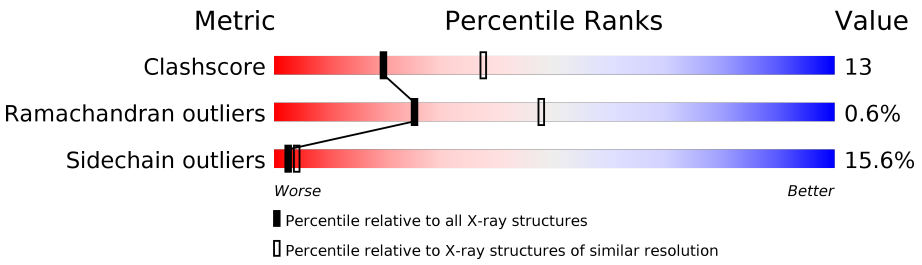
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.13.1

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.60 Å.

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)	Similar resolution (#Entries, resolution range(Å))
Clashscore	141614	3518 (2.60-2.60)
Ramachandran outliers	138981	3455 (2.60-2.60)
Sidechain outliers	138945	3455 (2.60-2.60)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. 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%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	158	53% 36% 11%
2	B	2	50% 50%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 1375 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called TYPE 4 PILIN.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	P	S			
1	A	158	1212	761	206	240	1	4	0	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	68	SEP	SER	modified residue	UNP P02974

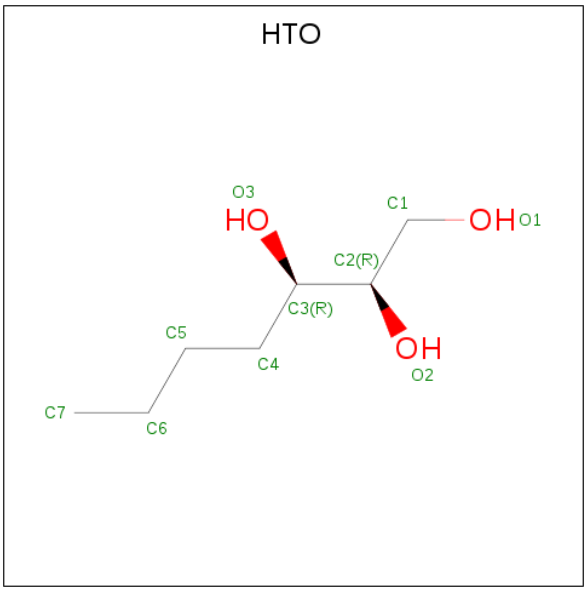
- Molecule 2 is an oligosaccharide called alpha-D-galactopyranose-(1-3)-2-acetamido-2-deoxy-beta-D-glucopyranose.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	B	2	25	14	1	10	0	0	0

- Molecule 3 is PLATINUM (II) ION (three-letter code: PT) (formula: Pt).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Pt		
3	A	1	1	1	0	0

- Molecule 4 is HEPTANE-1,2,3-TRIOL (three-letter code: HTO) (formula: $C_7H_{16}O_3$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
4	A	1	10	7	3	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
5	A	127	127	127	0	0

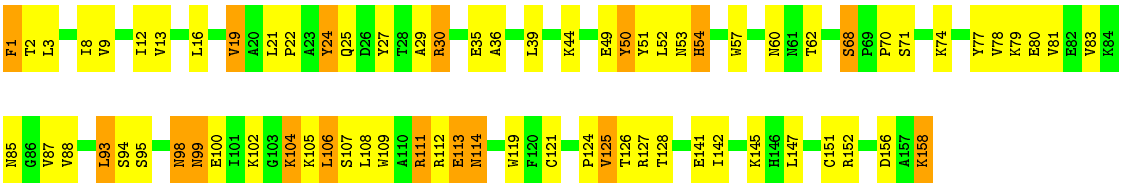
3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-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 outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

Note EDS was not executed.

- Molecule 1: TYPE 4 PILIN

Chain A:



- Molecule 2: alpha-D-galactopyranose-(1-3)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain B:

4 Data and refinement statistics (i)

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	C 2 2 2	Depositor
Cell constants a, b, c, α, β, γ	127.58Å 121.08Å 26.86Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 2.60	Depositor
% Data completeness (in resolution range)	95.0 (20.00-2.60)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	0.08	Depositor
Refinement program	X-PLOR 3.8	Depositor
R, R_{free}	0.187 , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	1375	wwPDB-VP
Average B, all atoms (Å^2)	42.0	wwPDB-VP

5 Model quality (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: MEA, NAG, PT, SEP, HTO, GLA

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 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	0.91	0/1210	1.77	36/1638 (2.2%)

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

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1

There are no bond length outliers.

All (36) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	106	LEU	CA-CB-CG	9.41	136.94	115.30
1	A	119	TRP	CD1-CG-CD2	8.67	113.23	106.30
1	A	2	THR	N-CA-C	8.62	134.28	111.00
1	A	109	TRP	CD1-CG-CD2	8.33	112.97	106.30
1	A	152	ARG	NE-CZ-NH2	-8.27	116.17	120.30
1	A	27	TYR	CB-CG-CD1	-8.04	116.18	121.00
1	A	119	TRP	CE2-CD2-CG	-7.67	101.16	107.30
1	A	127	ARG	NE-CZ-NH1	7.57	124.08	120.30
1	A	109	TRP	CE2-CD2-CG	-7.53	101.28	107.30
1	A	57	TRP	CD1-CG-CD2	7.25	112.10	106.30
1	A	77	TYR	CB-CG-CD2	-7.19	116.69	121.00
1	A	57	TRP	CE2-CD2-CG	-7.00	101.70	107.30
1	A	100	GLU	CA-CB-CG	6.90	128.58	113.40
1	A	119	TRP	CG-CD2-CE3	6.60	139.84	133.90
1	A	113	GLU	CA-C-N	-6.49	102.92	117.20
1	A	19	VAL	CA-CB-CG2	-6.29	101.46	110.90

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	30	ARG	NE-CZ-NH2	-5.82	117.39	120.30
1	A	2	THR	CB-CA-C	-5.66	96.31	111.60
1	A	24	TYR	CB-CG-CD1	-5.53	117.68	121.00
1	A	19	VAL	CA-CB-CG1	5.52	119.18	110.90
1	A	51	TYR	CB-CG-CD1	-5.44	117.74	121.00
1	A	119	TRP	CG-CD1-NE1	-5.43	104.67	110.10
1	A	104	LYS	CA-CB-CG	5.43	125.34	113.40
1	A	125	VAL	N-CA-CB	-5.42	99.58	111.50
1	A	152	ARG	NE-CZ-NH1	5.39	123.00	120.30
1	A	156	ASP	CB-CG-OD2	-5.37	113.46	118.30
1	A	156	ASP	CB-CG-OD1	5.33	123.09	118.30
1	A	57	TRP	CG-CD2-CE3	5.30	138.67	133.90
1	A	109	TRP	CG-CD1-NE1	-5.29	104.81	110.10
1	A	9	VAL	CA-CB-CG2	-5.27	103.00	110.90
1	A	119	TRP	CB-CG-CD1	-5.24	120.19	127.00
1	A	57	TRP	CB-CG-CD1	-5.18	120.26	127.00
1	A	87	VAL	CG1-CB-CG2	-5.15	102.66	110.90
1	A	16	LEU	CA-CB-CG	5.07	126.97	115.30
1	A	128	THR	CA-CB-CG2	5.03	119.45	112.40
1	A	127	ARG	NE-CZ-NH2	-5.03	117.79	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	50	TYR	Sidechain

5.2 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 the 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 within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1212	0	1208	32	1
2	B	25	0	22	0	0
3	A	1	0	0	0	0
4	A	10	0	16	2	0
5	A	127	0	0	1	1
All	All	1375	0	1246	32	1

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 13.

All (32) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:1:MEA:HE2	4:A:162:HTO:H41	1.53	0.91
1:A:62:THR:OG1	1:A:68:SEP:HB2	1.90	0.71
1:A:99:ASN:HA	1:A:102:LYS:HE3	1.81	0.62
1:A:8:ILE:O	1:A:12:ILE:HG23	2.01	0.61
1:A:111:ARG:HG2	1:A:158:LYS:HD3	1.86	0.57
1:A:60:ASN:HA	1:A:83:VAL:O	2.08	0.53
1:A:112:ARG:NH1	1:A:114:ASN:HA	2.24	0.53
1:A:112:ARG:HD3	1:A:114:ASN:OD1	2.09	0.53
1:A:107:SER:OG	1:A:124:PRO:HA	2.10	0.52
1:A:94:SER:HA	1:A:102:LYS:HB2	1.92	0.52
1:A:36:ALA:HA	1:A:78:VAL:HG21	1.91	0.51
1:A:70:PRO:HB2	1:A:80:GLU:HB2	1.93	0.51
1:A:39:LEU:HB3	1:A:81:VAL:HG21	1.92	0.50
1:A:53:ASN:HB2	1:A:54:HIS:ND1	2.29	0.48
1:A:104:LYS:HB2	1:A:142:ILE:HD13	1.95	0.48
1:A:35:GLU:OE1	1:A:78:VAL:HG23	2.14	0.47
1:A:113:GLU:HG3	1:A:113:GLU:O	2.15	0.47
1:A:1:MEA:CE2	4:A:162:HTO:H41	2.36	0.47
1:A:112:ARG:HH11	1:A:114:ASN:HA	1.78	0.46
1:A:24:TYR:HE2	5:A:397:HOH:O	1.99	0.46
1:A:25:GLN:OE1	1:A:25:GLN:HA	2.16	0.45
1:A:121:CYS:HB3	1:A:151:CYS:HA	1.98	0.44
1:A:79:LYS:HB2	1:A:93:LEU:HD13	1.99	0.43
1:A:21:LEU:HB3	1:A:22:PRO:HD2	2.00	0.43
1:A:12:ILE:HG13	1:A:13:VAL:N	2.33	0.43
1:A:50:TYR:O	1:A:54:HIS:ND1	2.51	0.43
1:A:29:ALA:HB2	1:A:98:ASN:ND2	2.34	0.43
1:A:124:PRO:HD2	1:A:141:GLU:HB3	2.00	0.43
1:A:3:LEU:HD12	1:A:3:LEU:HA	1.90	0.42
1:A:30:ARG:HE	1:A:30:ARG:HB3	1.73	0.41
1:A:44:LYS:HD3	1:A:108:LEU:HD23	2.01	0.41
1:A:158:LYS:HZ3	1:A:158:LYS:HB3	1.86	0.41

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:68:SEP:O1P	5:A:608:HOH:O[1_554]	2.00	0.20

5.3 Torsion angles (i)

5.3.1 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 X-ray entries followed by that with respect to entries of similar resolution.

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	155/158 (98%)	146 (94%)	8 (5%)	1 (1%)	25 47

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	114	ASN

5.3.2 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 X-ray entries followed by that with respect to entries of similar resolution.

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	128/128 (100%)	108 (84%)	20 (16%)	2 4

All (20) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	19	VAL
1	A	49	GLU
1	A	52	LEU
1	A	54	HIS
1	A	71	SER

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	74	LYS
1	A	85	ASN
1	A	88	VAL
1	A	93	LEU
1	A	95	SER
1	A	98	ASN
1	A	99	ASN
1	A	105	LYS
1	A	106	LEU
1	A	111	ARG
1	A	125	VAL
1	A	126	THR
1	A	145	LYS
1	A	147	LEU
1	A	158	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	43	GLN
1	A	53	ASN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

2 non-standard protein/DNA/RNA residues are modelled in this entry.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
1	SEP	A	68	1	8,9,10	2.27	1 (12%)	8,12,14	4.40	2 (25%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
1	MEA	A	1	1	10,11,13	1.51	2 (20%)	10,13,16	1.63	2 (20%)

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	SEP	A	68	1	-	5/5/8/10	-
1	MEA	A	1	1	-	1/5/6/10	0/1/1/1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	68	SEP	P-O1P	5.73	1.69	1.50
1	A	1	MEA	CB-CG	2.53	1.57	1.51
1	A	1	MEA	CD1-CG	2.08	1.43	1.38

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	68	SEP	OG-CB-CA	11.89	119.72	108.14
1	A	1	MEA	CG-CB-CA	3.29	120.76	114.10
1	A	1	MEA	CB-CA-C	3.16	117.39	111.47
1	A	68	SEP	OG-P-O1P	2.58	113.70	106.47

There are no chirality outliers.

All (6) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	68	SEP	N-CA-CB-OG
1	A	68	SEP	CB-OG-P-O1P
1	A	68	SEP	CB-OG-P-O2P
1	A	68	SEP	CB-OG-P-O3P
1	A	1	MEA	C-CA-CB-CG
1	A	68	SEP	CA-CB-OG-P

There are no ring outliers.

2 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	68	SEP	1	1
1	A	1	MEA	2	0

5.5 Carbohydrates (i)

2 monosaccharides are modelled in this entry.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
2	NAG	B	1	1,2	14,14,15	0.56	0	17,19,21	1.20	2 (11%)
2	GLA	B	2	2	11,11,12	0.40	0	15,15,17	0.64	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	-	2/6/23/26	0/1/1/1
2	GLA	B	2	2	-	1/2/19/22	0/1/1/1

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	1	NAG	C4-C3-C2	-3.09	106.50	111.02
2	B	1	NAG	C2-N2-C7	-3.03	118.58	122.90

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	1	NAG	O5-C5-C6-O6

Continued on next page...

Continued from previous page...

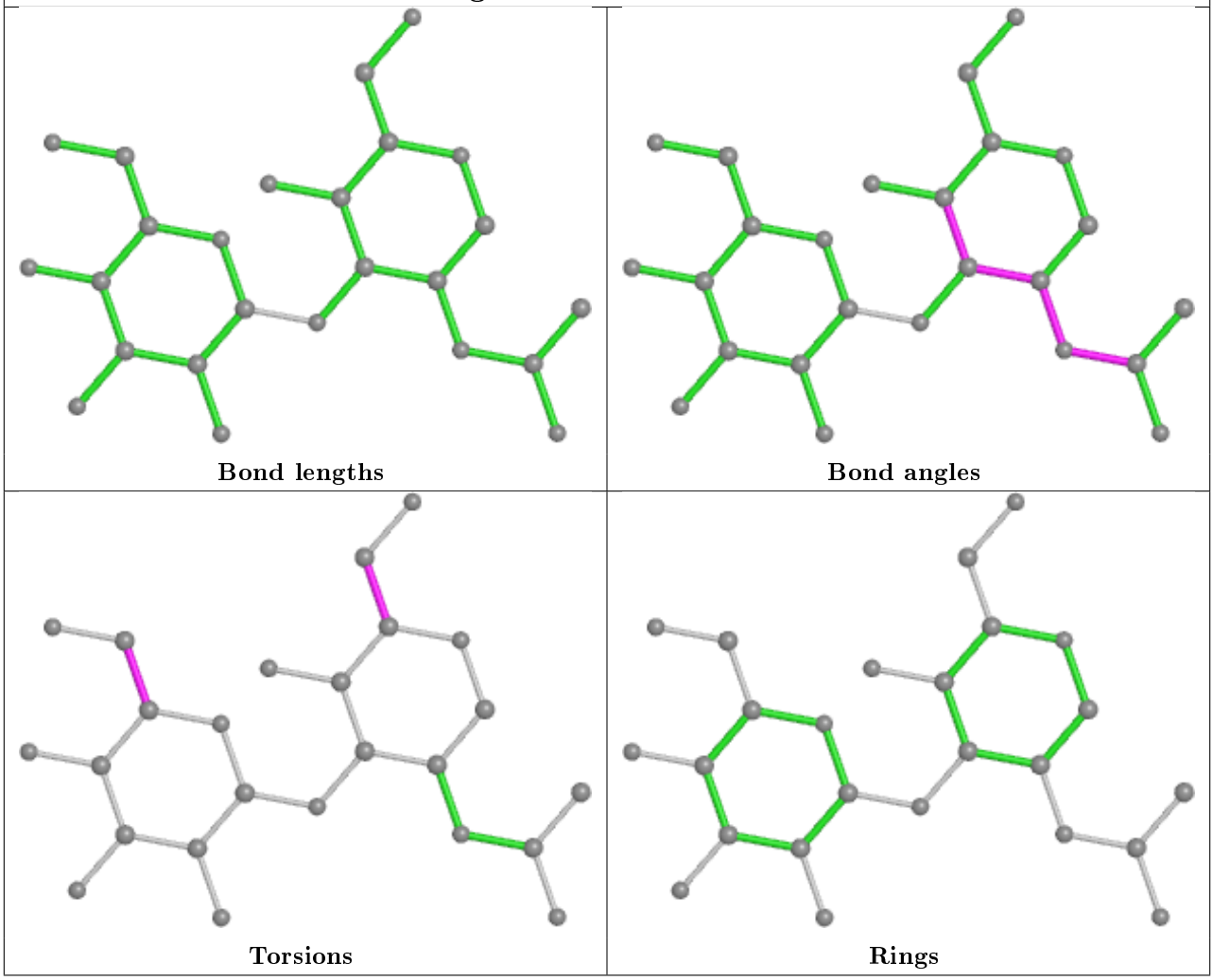
Mol	Chain	Res	Type	Atoms
2	B	2	GLA	O5-C5-C6-O6
2	B	1	NAG	C4-C5-C6-O6

There are no ring outliers.

No monomer is involved in short contacts.

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



5.6 Ligand geometry (i)

Of 2 ligands modelled in this entry, 1 is monoatomic - leaving 1 for Mogul analysis.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
4	HTO	A	162	-	9,9,9	1.00	1 (11%)	10,10,10	1.00	1 (10%)

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
4	HTO	A	162	-	-	2/10/10/10	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	162	HTO	C3-C2	2.49	1.59	1.52

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	A	162	HTO	O3-C3-C2	-2.03	105.56	109.72

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	162	HTO	O1-C1-C2-O2
4	A	162	HTO	O1-C1-C2-C3

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	162	HTO	2	0

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data (i)

6.1 Protein, DNA and RNA chains (i)

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates (i)

EDS was not executed - this section is therefore empty.

6.4 Ligands (i)

EDS was not executed - this section is therefore empty.

6.5 Other polymers (i)

EDS was not executed - this section is therefore empty.