wwPDB X-ray Structure Validation Summary Report (i)

Oct 16, 2023 – 12:57 PM EDT

PDB ID	:	1YF8
Title	:	Crystal structure of Himalayan mistletoe RIP reveals the presence of a natural inhibitor and a new functionally active sugar-binding site
Authors	:	Mishra, V.; Bilgrami, S.; Sharma, R.S.; Kaur, P.; Yadav, S.; Betzel, C.; Babu, C.R.; Singh, T.P.
Deposited on	:	2004-12-31
Resolution	:	2.80 Å(reported)

This is a wwPDB X-ray 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/XrayValidationReportHelp
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 (i)) were used in the production of this report:

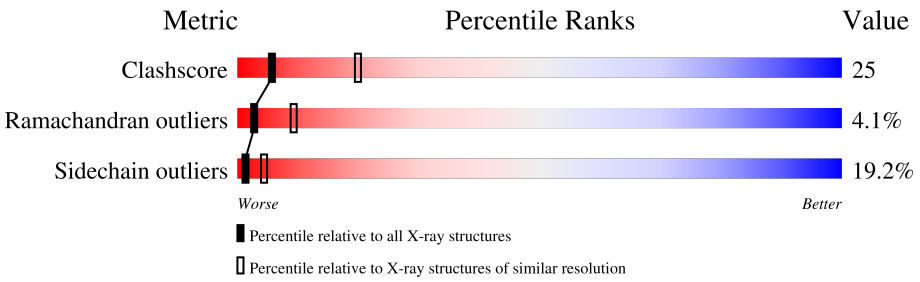
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 2.80 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	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)

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 of 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	240	49% 37% 12% •
2	B	255	56% 36% 7% •
3	C	2	100%
4	D	4	25% 50% 25%
5	E	3	67% 33%
6	F	2	100%
6	G	2	100%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	MAN	D	3	X	-	-	-
5	MAN	E	3	X	-	-	-
6	GAL	G	2	-	-	X	-
7	NAG	A	1003	X	-	-	-

2 Entry composition (i)

There are 10 unique types of molecules in this entry. The entry contains 4125 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 Beta-galactoside-specific lectin 4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	240	1875	1191	319	361	4	0	0	0

- Molecule 2 is a protein called Beta-galactoside-specific lectin 4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	255	1938	1199	343	384	12	0	0	0

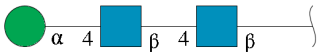
- Molecule 3 is an oligosaccharide called 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
3	C	2	28	16	2	10	0	0	0

- Molecule 4 is an oligosaccharide called beta-D-mannopyranose-(1-3)-alpha-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
4	D	4	50	28	2	20	0	0	0

- Molecule 5 is an oligosaccharide called alpha-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.

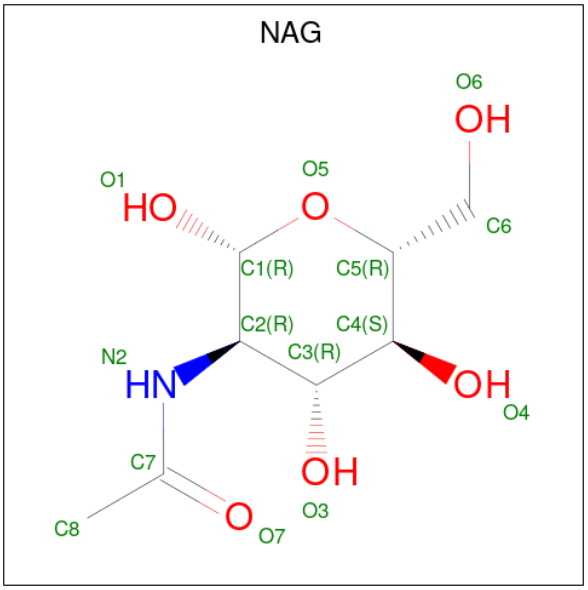


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
5	E	3	39	22	2	15	0	0	0

- Molecule 6 is an oligosaccharide called beta-D-galactopyranose-(1-4)-beta-D-glucopyranose.

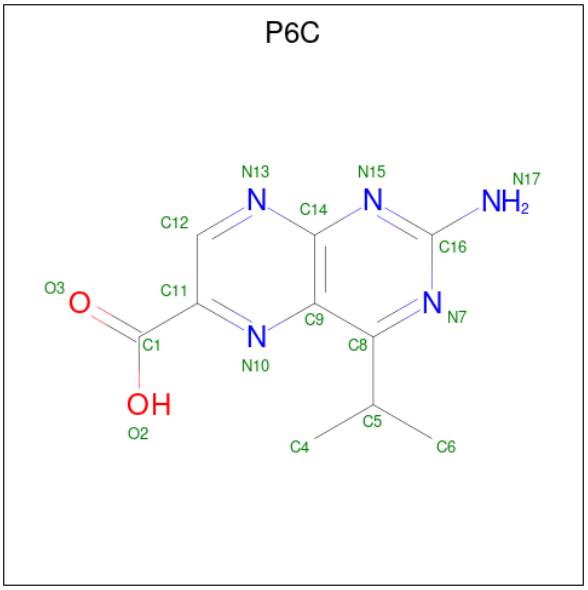
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	O				
6	F	2	23	12	11		0	0	0
6	G	2	23	12	11		0	0	0

- Molecule 7 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



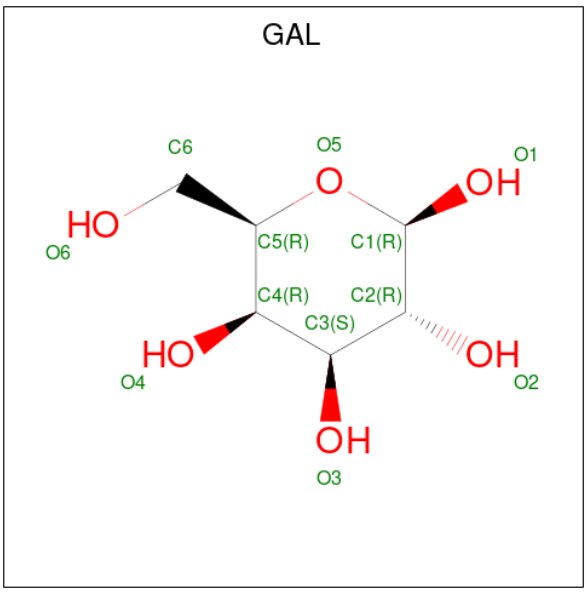
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
7	A	1	14	8	1	5	0	0

- Molecule 8 is 2-AMINO-4-ISOPROPYL-PTERIDINE-6-CARBOXYLIC ACID (three-letter code: P6C) (formula: $C_{10}H_{11}N_5O_2$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
8	A	1	17	10	5	2	0	0

- Molecule 9 is beta-D-galactopyranose (three-letter code: GAL) (formula: $C_6H_{12}O_6$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
9	B	1	11	6	5	0	0

- Molecule 10 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
10	A	55	55	55	0	0
			Total	O		
10	B	52	52	52	0	0

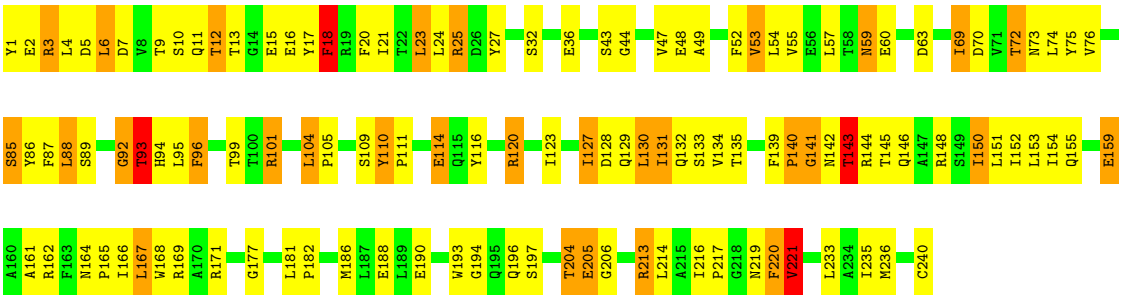
3 Residue-property plots (i)

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.

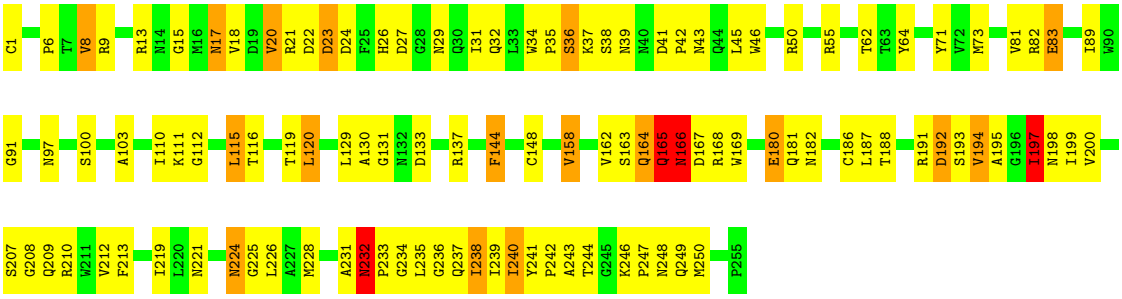
• Molecule 1: Beta-galactoside-specific lectin 4

Chain A: 49% 37% 12% •



• Molecule 2: Beta-galactoside-specific lectin 4

Chain B: 56% 36% 7% •



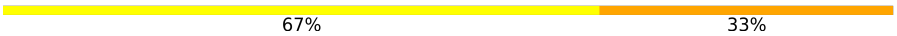
• Molecule 3: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain C: 100%

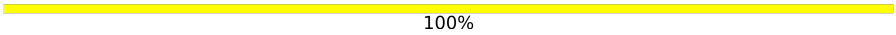
• Molecule 4: beta-D-mannopyranose-(1-3)-alpha-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain D:

• Molecule 5: alpha-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain E: 

• Molecule 6: beta-D-galactopyranose-(1-4)-beta-D-glucopyranose

Chain F: 

• Molecule 6: beta-D-galactopyranose-(1-4)-beta-D-glucopyranose

Chain G:

4 Data and refinement statistics (i)

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

Property	Value	Source
Space group	P 65 2 2	Depositor
Cell constants a, b, c, α, β, γ	109.42Å 109.42Å 309.80Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	20.00 – 2.80	Depositor
% Data completeness (in resolution range)	100.0 (20.00-2.80)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	0.06	Depositor
Refinement program	REFMAC 5.1.19	Depositor
R, R_{free}	0.234 , 0.287	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	4125	wwPDB-VP
Average B, all atoms (Å^2)	86.0	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: BMA, MAN, GAL, BGC, P6C, NAG

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	0.92	3/1913 (0.2%)	1.06	11/2602 (0.4%)
2	B	0.92	0/1978	1.04	5/2697 (0.2%)
All	All	0.92	3/3891 (0.1%)	1.05	16/5299 (0.3%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	89	SER	CB-OG	7.99	1.52	1.42
1	A	36	GLU	CG-CD	5.64	1.60	1.51
1	A	168	TRP	CB-CG	5.36	1.59	1.50

The worst 5 of 16 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	96	PHE	N-CA-CB	8.72	126.29	110.60
1	A	6	LEU	CA-CB-CG	8.39	134.59	115.30
1	A	88	LEU	CA-CB-CG	7.57	132.70	115.30
1	A	4	LEU	CA-CB-CG	6.77	130.87	115.30
2	B	236	GLY	N-CA-C	6.66	129.75	113.10

There are no chirality outliers.

There are no planarity outliers.

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	1875	0	1849	103	0
2	B	1938	0	1851	93	0
3	C	28	0	25	2	0
4	D	50	0	43	5	0
5	E	39	0	34	1	0
6	F	23	0	21	0	0
6	G	23	0	21	8	0
7	A	14	0	13	0	0
8	A	17	0	9	2	0
9	B	11	0	10	5	0
10	A	55	0	0	4	0
10	B	52	0	0	3	0
All	All	4125	0	3876	201	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 25.

The worst 5 of 201 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:18:PHE:HB2	1:A:167:LEU:HD21	1.27	1.16
3:C:1:NAG:H62	3:C:2:NAG:H2	1.17	1.10
4:D:2:NAG:H83	4:D:2:NAG:H3	1.39	1.02
1:A:139:PHE:HB3	1:A:140:PRO:HD2	1.45	0.96
1:A:101:ARG:HG2	1:A:101:ARG:HH11	1.32	0.94

There are no symmetry-related clashes.

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	238/240 (99%)	200 (84%)	27 (11%)	11 (5%)	2 7

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
2	B	253/255 (99%)	223 (88%)	21 (8%)	9 (4%)	3 11
All	All	491/495 (99%)	423 (86%)	48 (10%)	20 (4%)	3 9

5 of 20 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	110	TYR
1	A	140	PRO
1	A	143	THR
1	A	204	THR
1	A	217	PRO

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	205/205 (100%)	158 (77%)	47 (23%)	1 2
2	B	211/211 (100%)	178 (84%)	33 (16%)	2 8
All	All	416/416 (100%)	336 (81%)	80 (19%)	1 4

5 of 80 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	115	LEU
2	B	212	VAL
2	B	120	LEU
2	B	182	ASN
2	B	235	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 17 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	221	ASN
2	B	232	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	B	97	ASN
2	B	145	ASN
2	B	165	GLN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates (i)

13 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	$\#\|Z\| > 2$
3	NAG	C	1	2,3	14,14,15	0.90	1 (7%)	17,19,21	3.21	9 (52%)
3	NAG	C	2	3	14,14,15	0.74	0	17,19,21	1.77	4 (23%)
4	NAG	D	1	2,4	14,14,15	0.77	0	17,19,21	1.56	3 (17%)
4	NAG	D	2	4	14,14,15	0.58	0	17,19,21	2.25	7 (41%)
4	MAN	D	3	4	11,11,12	0.60	0	15,15,17	2.38	4 (26%)
4	BMA	D	4	4	11,11,12	0.56	0	15,15,17	1.69	3 (20%)
5	NAG	E	1	2,5	14,14,15	0.55	0	17,19,21	1.96	5 (29%)
5	NAG	E	2	5	14,14,15	0.72	0	17,19,21	1.78	3 (17%)
5	MAN	E	3	5	11,11,12	0.47	0	15,15,17	1.66	3 (20%)
6	BGC	F	1	6	12,12,12	0.57	0	17,17,17	2.08	5 (29%)
6	GAL	F	2	6	11,11,12	0.56	0	15,15,17	2.19	6 (40%)
6	BGC	G	1	6	12,12,12	0.92	1 (8%)	17,17,17	1.98	8 (47%)
6	GAL	G	2	6	11,11,12	0.65	0	15,15,17	2.11	6 (40%)

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	NAG	C	1	2,3	-	2/6/23/26	0/1/1/1
3	NAG	C	2	3	-	6/6/23/26	0/1/1/1
4	NAG	D	1	2,4	-	0/6/23/26	0/1/1/1
4	NAG	D	2	4	-	5/6/23/26	0/1/1/1
4	MAN	D	3	4	2/2/4/5	2/2/19/22	0/1/1/1
4	BMA	D	4	4	-	2/2/19/22	0/1/1/1
5	NAG	E	1	2,5	-	2/6/23/26	0/1/1/1
5	NAG	E	2	5	-	4/6/23/26	0/1/1/1
5	MAN	E	3	5	1/1/4/5	1/2/19/22	1/1/1/1
6	BGC	F	1	6	-	2/2/22/22	0/1/1/1
6	GAL	F	2	6	-	1/2/19/22	0/1/1/1
6	BGC	G	1	6	-	1/2/22/22	0/1/1/1
6	GAL	G	2	6	-	2/2/19/22	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	G	1	BGC	O1-C1	2.14	1.46	1.39
3	C	1	NAG	C2-N2	2.02	1.49	1.46

The worst 5 of 66 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	C	1	NAG	C1-C2-N2	7.05	122.53	110.49
3	C	1	NAG	C1-O5-C5	6.72	121.30	112.19
4	D	3	MAN	O3-C3-C2	5.43	120.39	109.99
6	F	2	GAL	C1-O5-C5	5.09	119.09	112.19
4	D	4	BMA	O5-C5-C6	4.76	114.66	107.20

All (3) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
4	D	3	MAN	C1
4	D	3	MAN	C3
5	E	3	MAN	C1

5 of 30 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	C	2	NAG	C3-C2-N2-C7
4	D	2	NAG	C3-C2-N2-C7
6	G	2	GAL	O5-C5-C6-O6
3	C	2	NAG	O5-C5-C6-O6
6	F	1	BGC	O5-C5-C6-O6

All (1) ring outliers are listed below:

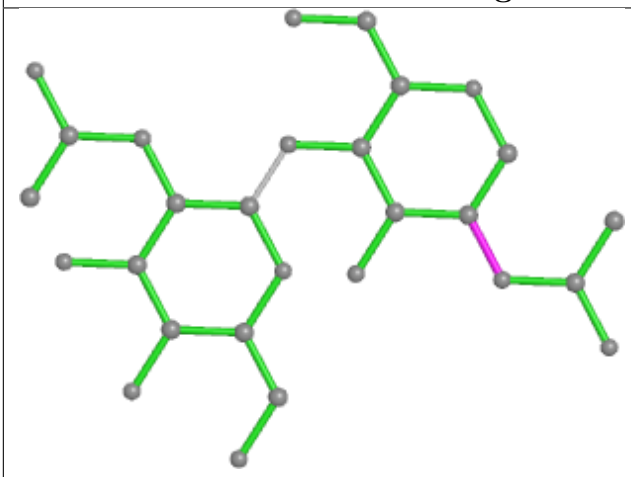
Mol	Chain	Res	Type	Atoms
5	E	3	MAN	C1-C2-C3-C4-C5-O5

8 monomers are involved in 16 short contacts:

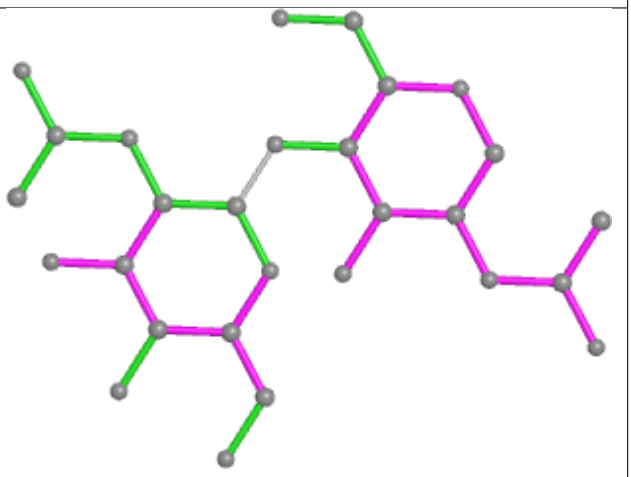
Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	E	1	NAG	1	0
4	D	2	NAG	4	0
4	D	4	BMA	1	0
6	G	1	BGC	3	0
3	C	1	NAG	2	0
4	D	3	MAN	1	0
3	C	2	NAG	2	0
6	G	2	GAL	6	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for oligosaccharide.

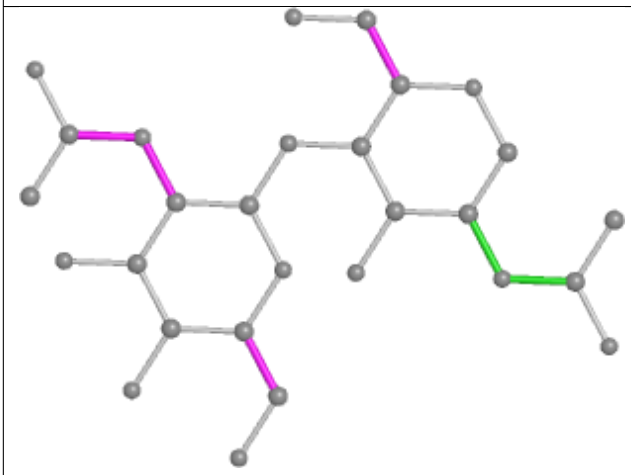
Oligosaccharide Chain C



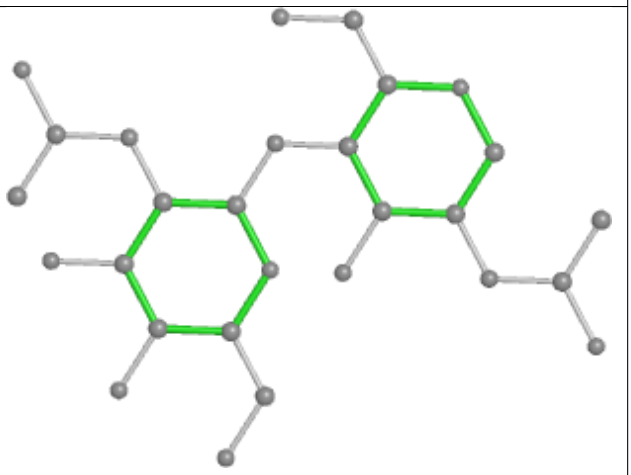
Bond lengths



Bond angles

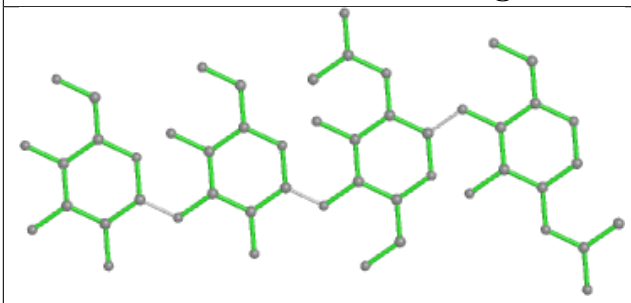


Torsions

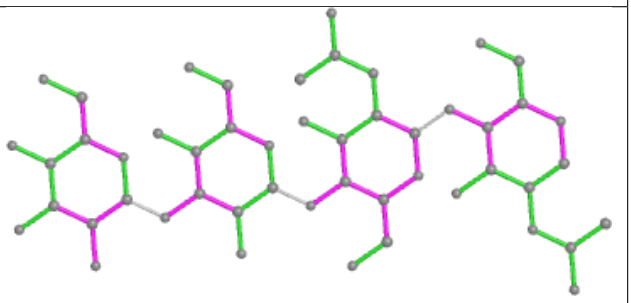


Rings

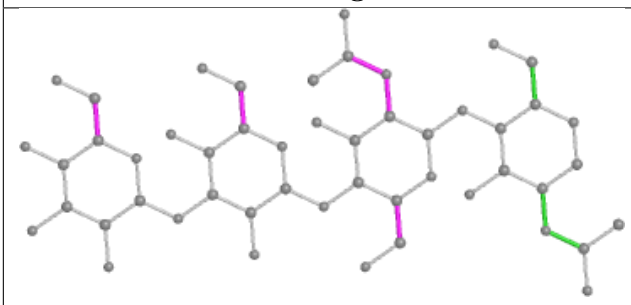
Oligosaccharide Chain D



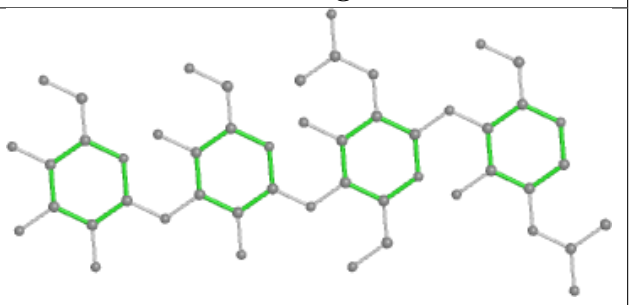
Bond lengths



Bond angles

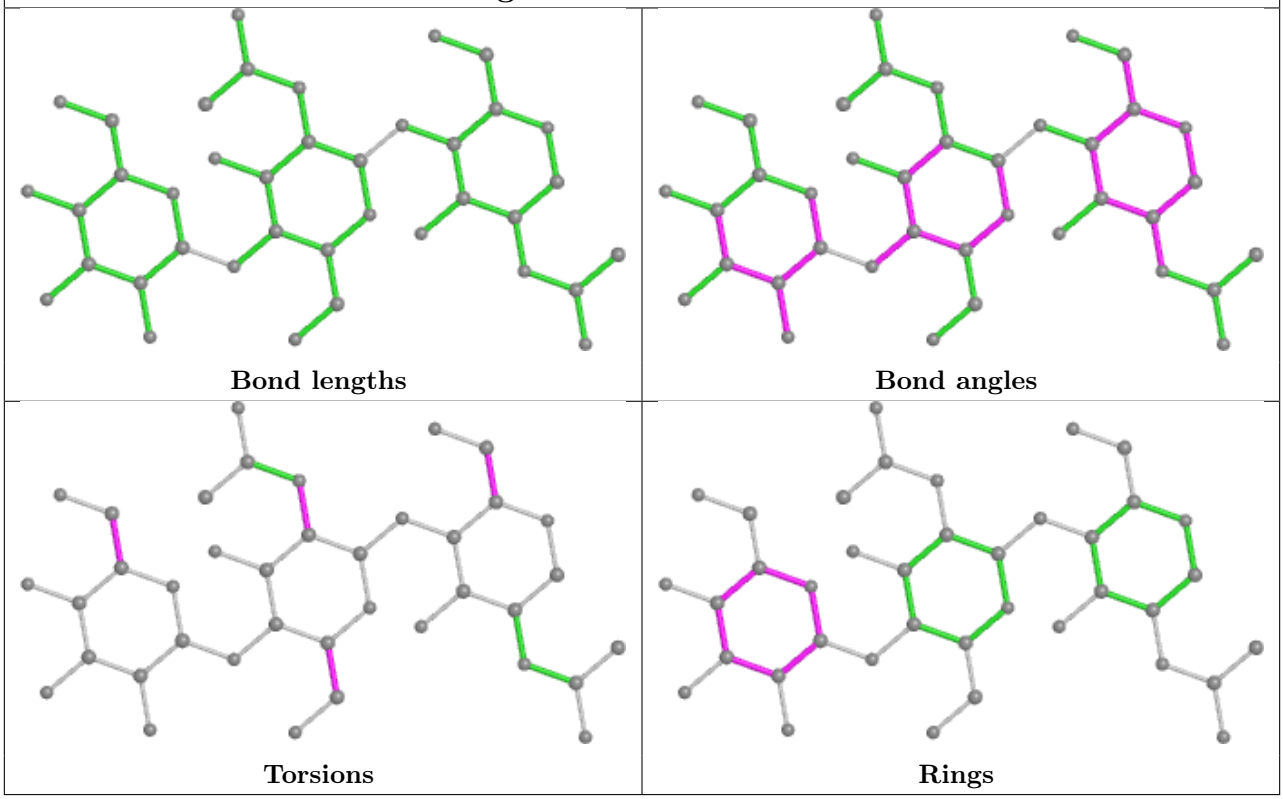


Torsions

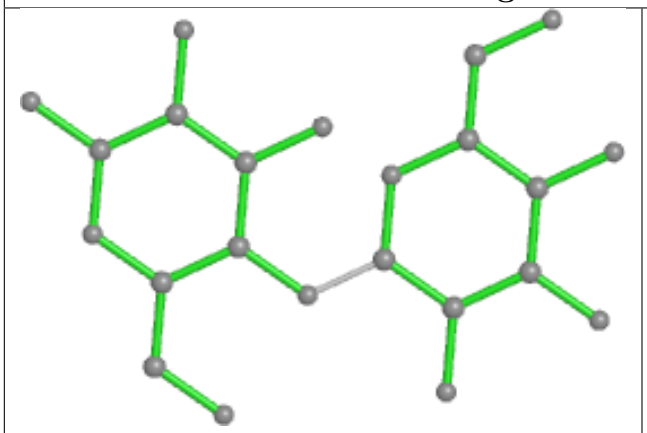


Rings

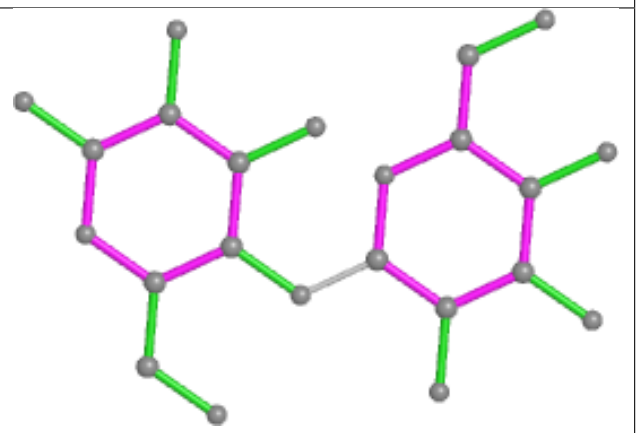
Oligosaccharide Chain E



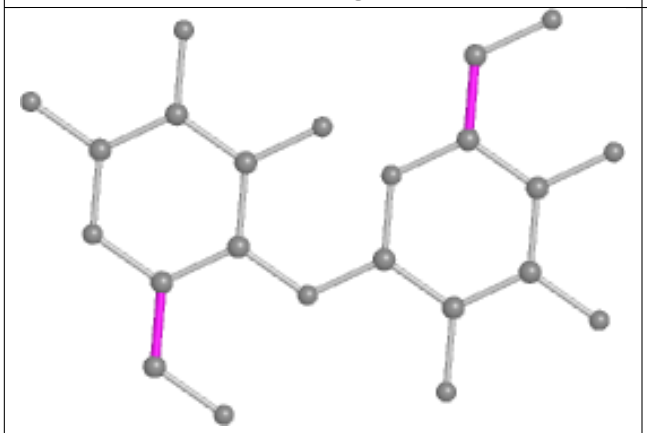
Oligosaccharide Chain F



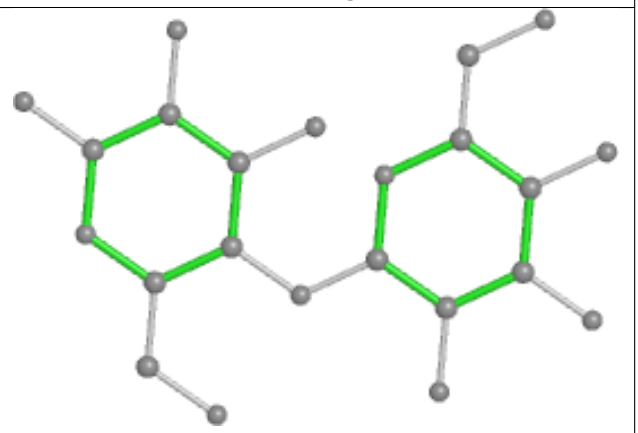
Bond lengths



Bond angles

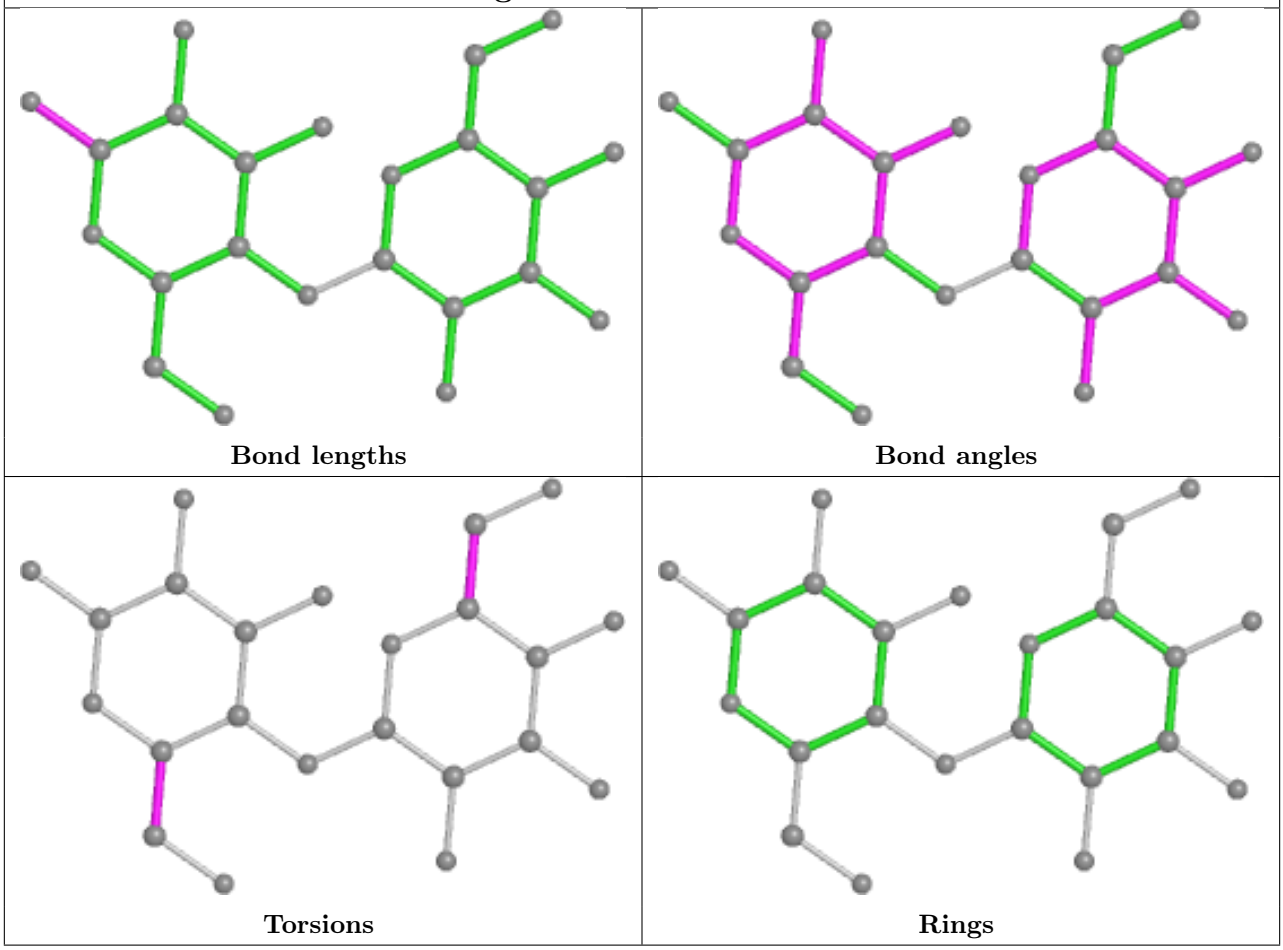


Torsions



Rings

Oligosaccharide Chain G



5.6 Ligand geometry (i)

3 ligands 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	$\#\|Z\| > 2$
7	NAG	A	1003	1	14,14,15	0.58	0	17,19,21	1.76	4 (23%)
9	GAL	B	1013	-	11,11,12	0.56	0	15,15,17	2.81	8 (53%)
8	P6C	A	1001	-	18,18,18	2.88	4 (22%)	23,26,26	4.16	12 (52%)

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
7	NAG	A	1003	1	1/1/5/7	4/6/23/26	0/1/1/1
9	GAL	B	1013	-	-	2/2/19/22	0/1/1/1
8	P6C	A	1001	-	-	2/8/8/8	0/2/2/2

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
8	A	1001	P6C	C8-C5	8.41	1.63	1.51
8	A	1001	P6C	C9-C14	5.58	1.50	1.40
8	A	1001	P6C	C8-N7	4.55	1.36	1.31
8	A	1001	P6C	C8-C9	2.74	1.47	1.43

The worst 5 of 24 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
8	A	1001	P6C	C4-C5-C8	11.13	129.74	110.96
8	A	1001	P6C	N13-C14-N15	8.00	124.96	115.82
8	A	1001	P6C	C12-N13-C14	6.63	123.35	116.69
8	A	1001	P6C	C9-C8-N7	-6.39	116.46	122.81
8	A	1001	P6C	C16-N7-C8	6.20	121.30	116.26

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
7	A	1003	NAG	C1

5 of 8 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
7	A	1003	NAG	C8-C7-N2-C2
7	A	1003	NAG	O7-C7-N2-C2
8	A	1001	P6C	C4-C5-C8-N7
7	A	1003	NAG	C4-C5-C6-O6
7	A	1003	NAG	O5-C5-C6-O6

There are no ring outliers.

2 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
9	B	1013	GAL	5	0
8	A	1001	P6C	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.