wwPDB X-ray Structure Validation Summary Report (i)

Oct 5, 2024 – 09:49 PM EDT

PDB ID	:	1JSI
Title	:	CRYSTAL STRUCTURE OF H9 HAEMAGGLUTININ BOUND TO LSTC RECEPTOR ANALOG
Authors	:	Ha, Y.; Stevens, D.J.; Skehel, J.J.; Wiley, D.C.
Deposited on	:	2001-08-17
Resolution	:	2.40 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:

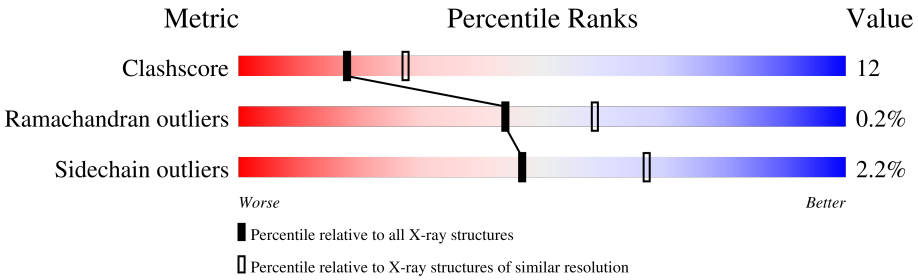
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

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 2.40 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	180529	5218 (2.40-2.40)
Ramachandran outliers	177936	5158 (2.40-2.40)
Sidechain outliers	177891	5159 (2.40-2.40)

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	319	75% 24% .
2	B	176	80% 9% . 9%
3	C	2	50% 50%
3	E	2	50% 50%
3	G	2	100%
4	D	3	67% 33%
5	F	5	20% 60% 20%

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
5	GLC	F	1	X	-	-	-

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 4073 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 HAEMAGGLUTININ (HA1 CHAIN).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	317	Total	C	N	O	S	0	0	0
			2472	1551	435	473	13			

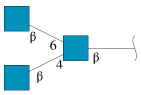
- Molecule 2 is a protein called HAEMAGGLUTININ (HA2 CHAIN).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	160	Total	C	N	O	S	0	0	0
			1251	779	214	250	8			

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

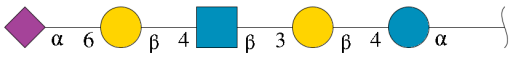
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	C	2	Total	C	N	O	0	0	0
			28	16	2	10			
3	E	2	Total	C	N	O	0	0	0
			28	16	2	10			
3	G	2	Total	C	N	O	0	0	0
			28	16	2	10			

- Molecule 4 is an oligosaccharide called 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-[2-acetamido-2-deoxy-beta-D-glucopyranose-(1-6)]2-acetamido-2-deoxy-beta-D-glucopyranose.



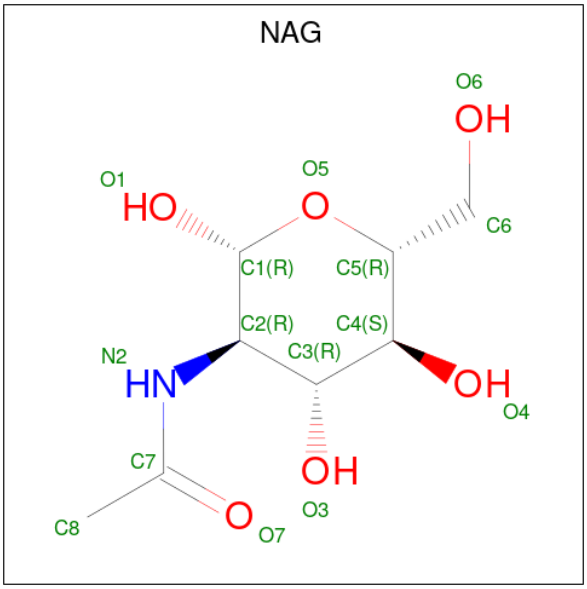
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf	Trace
4	D	3	Total C N O 42 24 3 15	0	0	0

- Molecule 5 is an oligosaccharide called N-acetyl-alpha-neuraminic acid-(2-6)-beta-D-galacto pyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-3)-beta-D-galactopyranose-(1-4)-alpha-D-glucopyranose.



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf	Trace
5	F	5	Total C N O 68 37 2 29	0	0	0

- Molecule 6 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



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

- Molecule 7 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	75	Total O 75 75	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	B	67	Total O 67 67	0	0

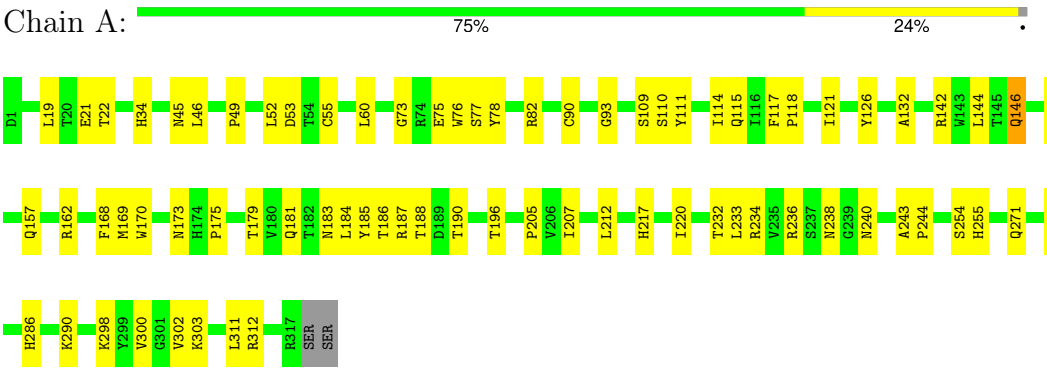
3 Residue-property plots (i)

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.

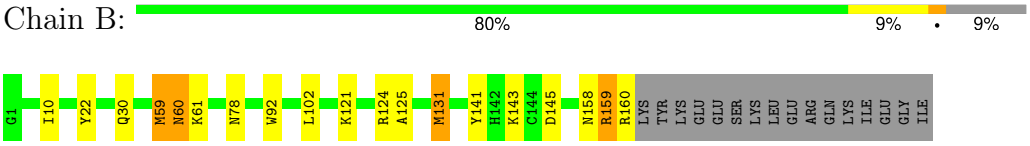
• Molecule 1: HAEMAGGLUTININ (HA1 CHAIN)

Chain A:



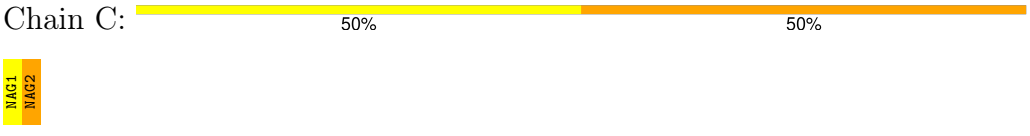
• Molecule 2: HAEMAGGLUTININ (HA2 CHAIN)

Chain B:



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

Chain C:

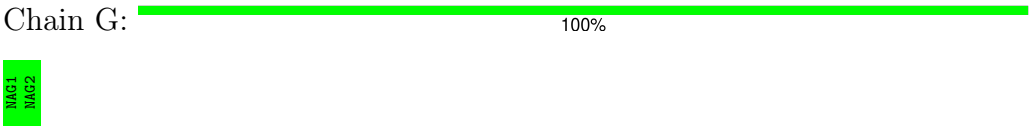


• Molecule 3: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain E:

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

Chain G:

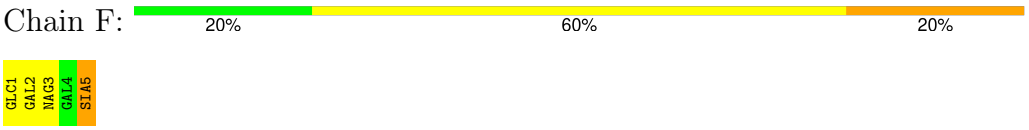


• Molecule 4: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-[2-acetamido-2-deoxy-beta-D-glucopyranose-(1-6)]2-acetamido-2-deoxy-beta-D-glucopyranose

Chain D:

• Molecule 5: N-acetyl-alpha-neuraminic acid-(2-6)-beta-D-galactopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-3)-beta-D-galactopyranose-(1-4)-alpha-D-glucopyranose

Chain F:



4 Data and refinement statistics (i)

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

Property	Value	Source
Space group	P 63	Depositor
Cell constants a, b, c, α, β, γ	108.58Å 108.58Å 149.41Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	40.00 – 2.40	Depositor
% Data completeness (in resolution range)	(Not available) (40.00-2.40)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	CNS	Depositor
R, R_{free}	0.202 , 0.235	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	4073	wwPDB-VP
Average B, all atoms (Å^2)	35.0	wwPDB-VP

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: GAL, SIA, GLC, 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	0.33	0/2531	0.63	0/3452
2	B	0.36	0/1272	0.61	1/1716 (0.1%)
All	All	0.34	0/3803	0.62	1/5168 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	10	ILE	N-CA-C	-5.30	96.69	111.00

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in 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	2472	0	2406	74	0
2	B	1251	0	1181	13	0
3	C	28	0	25	5	0
3	E	28	0	25	3	0
3	G	28	0	25	0	0
4	D	42	0	37	4	0
5	F	68	0	58	6	0
6	A	14	0	13	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	A	75	0	0	6	0
7	B	67	0	0	1	0
All	All	4073	0	3770	93	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 12.

The worst 5 of 93 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:271:GLN:HE21	1:A:280:ASN:HA	1.45	0.80
1:A:311:LEU:HD23	1:A:311:LEU:H	1.47	0.79
1:A:300:VAL:HG12	1:A:302:VAL:HG22	1.65	0.77
2:B:125:ALA:O	2:B:160:ARG:HA	1.87	0.73
1:A:153:ILE:H	1:A:153:ILE:HD12	1.54	0.73

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	315/319 (99%)	301 (96%)	14 (4%)	0	100 100
2	B	158/176 (90%)	153 (97%)	4 (2%)	1 (1%)	22 33
All	All	473/495 (96%)	454 (96%)	18 (4%)	1 (0%)	44 59

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	159	ARG

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	279/282 (99%)	276 (99%)	3 (1%)	70 84
2	B	132/150 (88%)	126 (96%)	6 (4%)	23 40
All	All	411/432 (95%)	402 (98%)	9 (2%)	47 67

5 of 9 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	121	LYS
2	B	131	MET
2	B	22	TYR
2	B	59	MET
2	B	60	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 14 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	214	ASN
1	A	240	ASN
2	B	78	ASN
2	B	60	ASN
2	B	68	HIS

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)

14 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	1,3	14,14,15	0.52	0	17,19,21	0.63	0
3	NAG	C	2	3	14,14,15	0.71	1 (7%)	17,19,21	0.57	0
4	NAG	D	1	1,4	14,14,15	0.58	0	17,19,21	0.76	1 (5%)
4	NAG	D	2	4	14,14,15	0.53	0	17,19,21	0.80	1 (5%)
4	NAG	D	3	4	14,14,15	0.65	0	17,19,21	0.63	0
3	NAG	E	1	1,3	14,14,15	0.87	1 (7%)	17,19,21	1.27	1 (5%)
3	NAG	E	2	3	14,14,15	0.60	0	17,19,21	0.75	1 (5%)
5	GLC	F	1	5	12,12,12	0.46	0	17,17,17	0.60	0
5	GAL	F	2	5	11,11,12	0.67	0	15,15,17	0.37	0
5	NAG	F	3	5	14,14,15	0.76	0	17,19,21	0.99	2 (11%)
5	GAL	F	4	5	11,11,12	0.57	0	15,15,17	0.45	0
5	SIA	F	5	5	20,20,21	0.62	0	21,28,31	0.96	2 (9%)
3	NAG	G	1	2,3	14,14,15	0.60	0	17,19,21	0.72	0
3	NAG	G	2	3	14,14,15	0.61	0	17,19,21	0.66	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
3	NAG	C	1	1,3	-	4/6/23/26	0/1/1/1
3	NAG	C	2	3	-	5/6/23/26	0/1/1/1
4	NAG	D	1	1,4	-	4/6/23/26	0/1/1/1
4	NAG	D	2	4	-	1/6/23/26	0/1/1/1
4	NAG	D	3	4	-	4/6/23/26	0/1/1/1
3	NAG	E	1	1,3	-	0/6/23/26	0/1/1/1
3	NAG	E	2	3	-	4/6/23/26	0/1/1/1
5	GLC	F	1	5	1/1/5/5	2/2/22/22	0/1/1/1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	GAL	F	2	5	-	2/2/19/22	0/1/1/1
5	NAG	F	3	5	-	2/6/23/26	0/1/1/1
5	GAL	F	4	5	-	0/2/19/22	0/1/1/1
5	SIA	F	5	5	-	0/18/34/38	0/1/1/1
3	NAG	G	1	2,3	-	1/6/23/26	0/1/1/1
3	NAG	G	2	3	-	4/6/23/26	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	E	1	NAG	C1-C2	2.50	1.55	1.52
3	C	2	NAG	C1-C2	2.16	1.55	1.52

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	F	3	NAG	C2-N2-C7	-2.46	119.60	122.90
4	D	2	NAG	C2-N2-C7	-2.45	119.62	122.90
5	F	5	SIA	O1B-C1-C2	2.41	118.98	112.71
3	E	1	NAG	C4-C3-C2	-2.30	107.65	111.02
4	D	1	NAG	C2-N2-C7	-2.20	119.95	122.90

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
5	F	1	GLC	C1

5 of 33 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	C	1	NAG	C8-C7-N2-C2
3	C	1	NAG	O7-C7-N2-C2
3	C	2	NAG	C1-C2-N2-C7
3	C	2	NAG	C8-C7-N2-C2
3	C	2	NAG	O7-C7-N2-C2

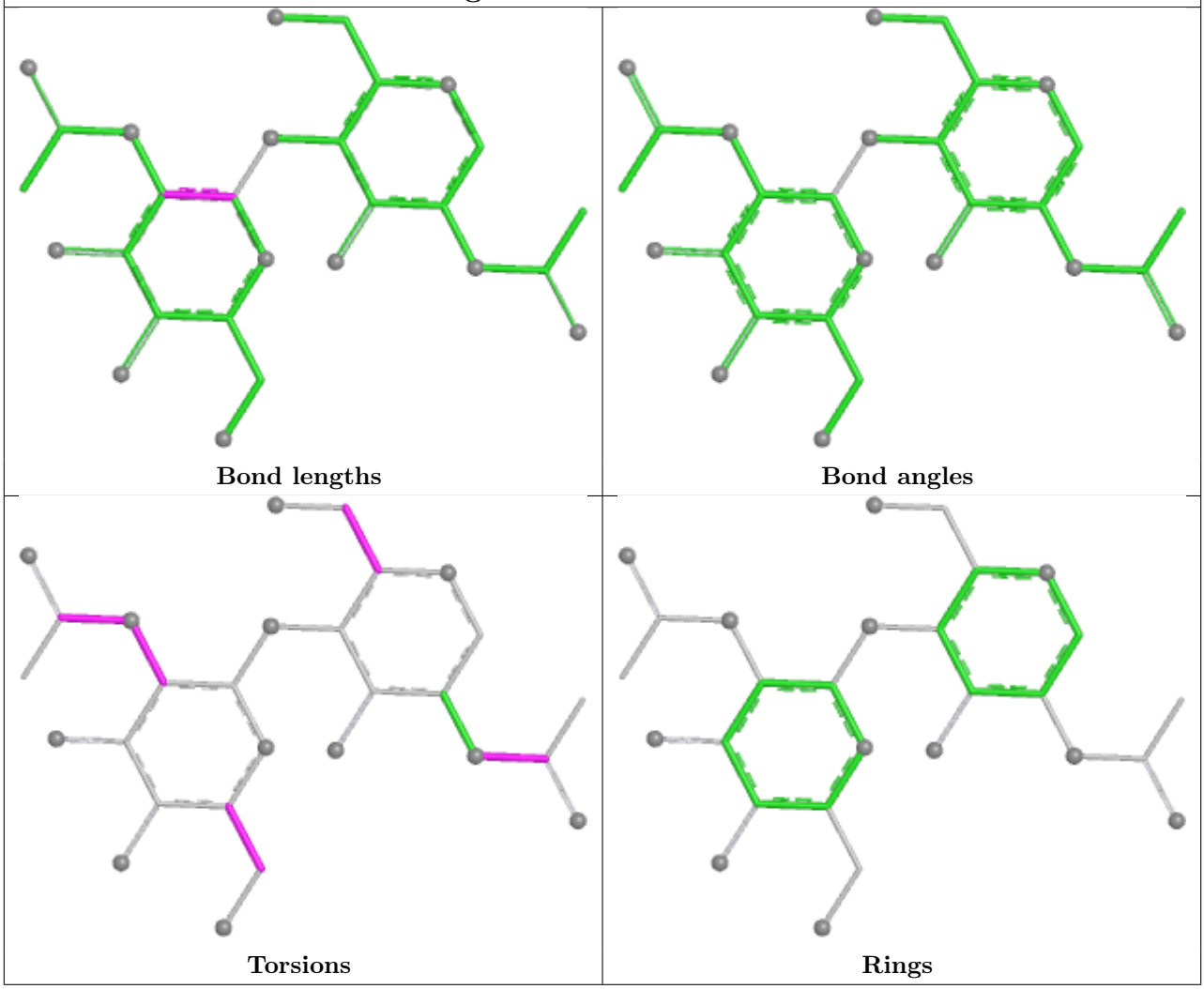
There are no ring outliers.

8 monomers are involved in 18 short contacts:

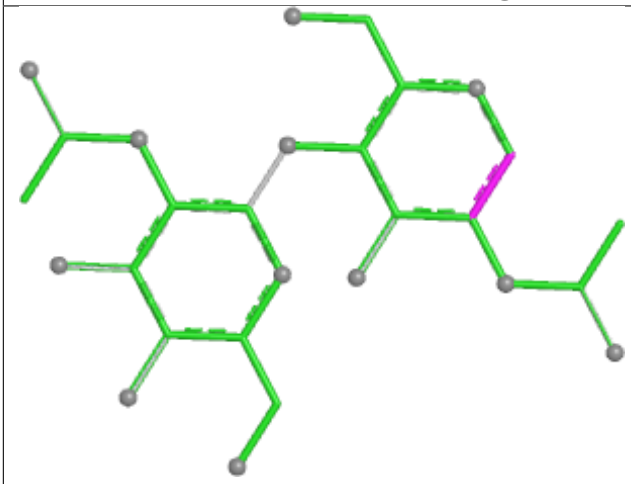
Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	C	1	NAG	5	0
5	F	2	GAL	3	0
5	F	1	GLC	3	0
4	D	2	NAG	3	0
3	C	2	NAG	3	0
4	D	3	NAG	4	0
3	E	1	NAG	3	0
5	F	5	SIA	3	0

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

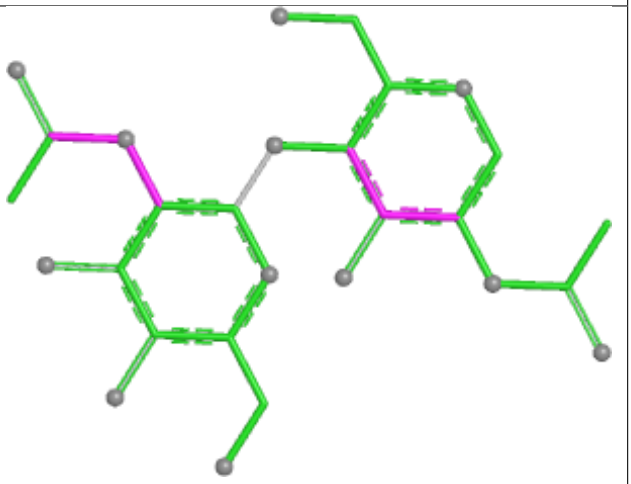
Oligosaccharide Chain C



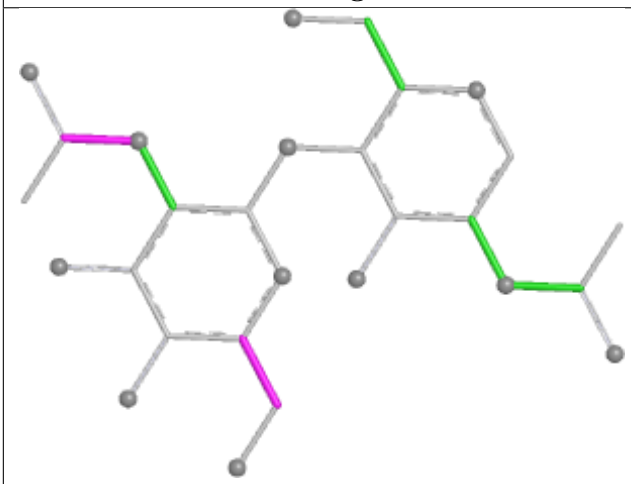
Oligosaccharide Chain E



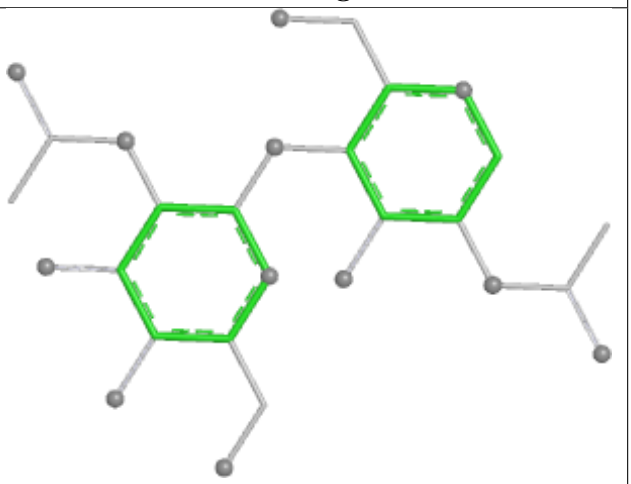
Bond lengths



Bond angles

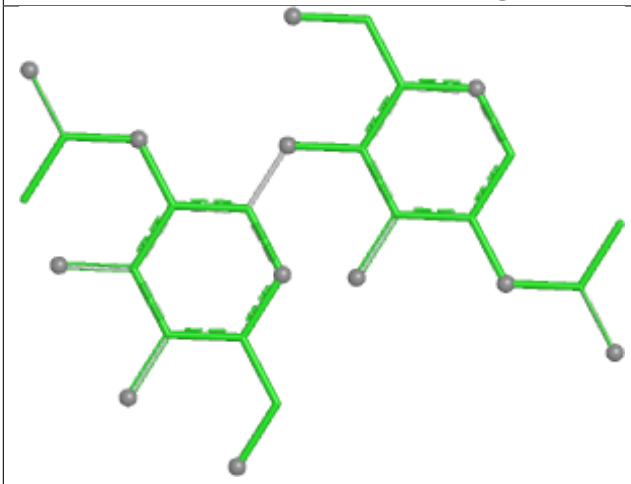


Torsions

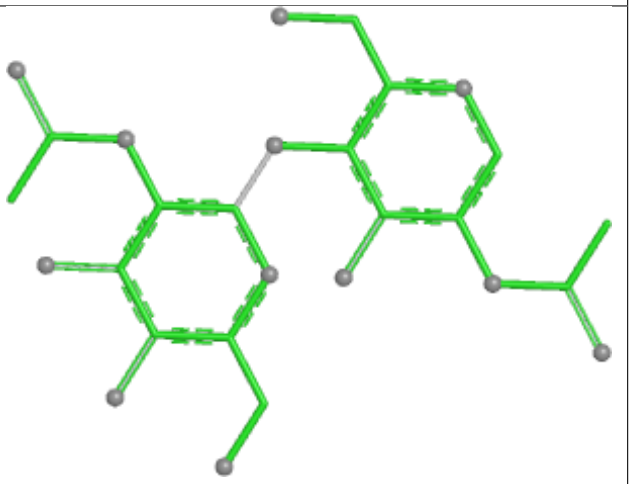


Rings

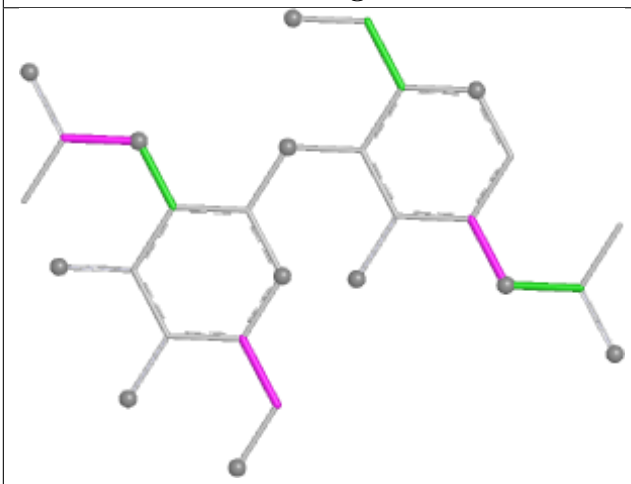
Oligosaccharide Chain G



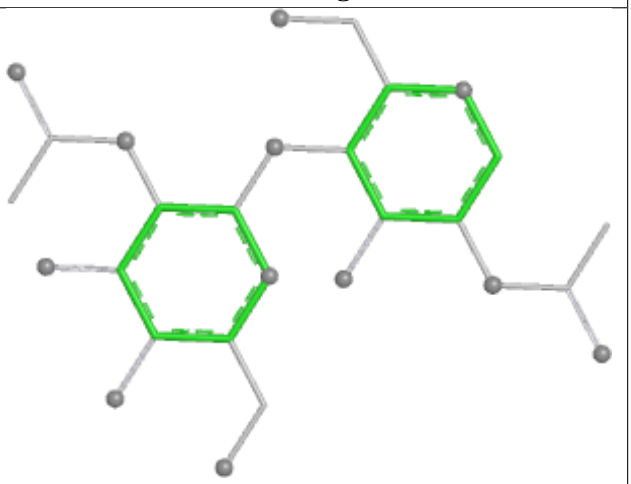
Bond lengths



Bond angles

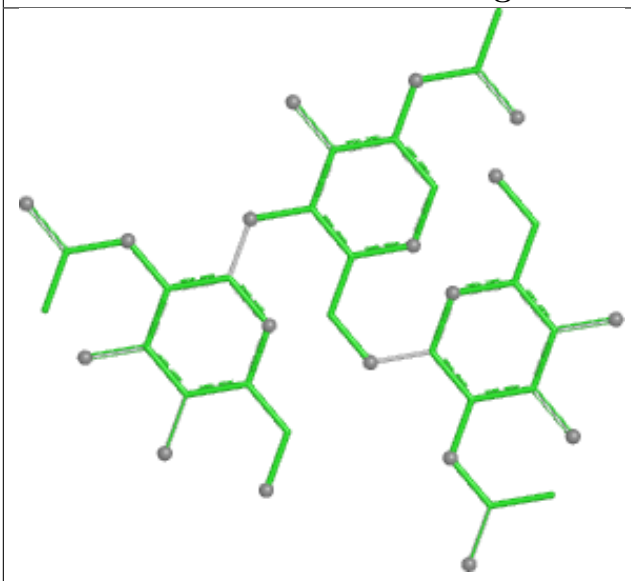


Torsions

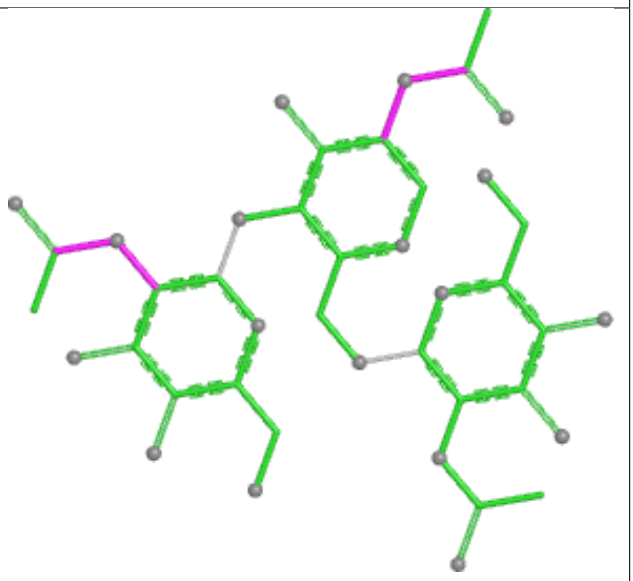


Rings

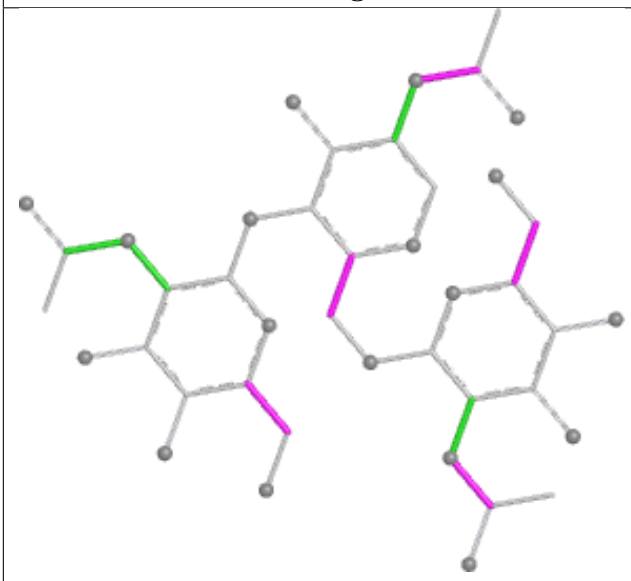
Oligosaccharide Chain D



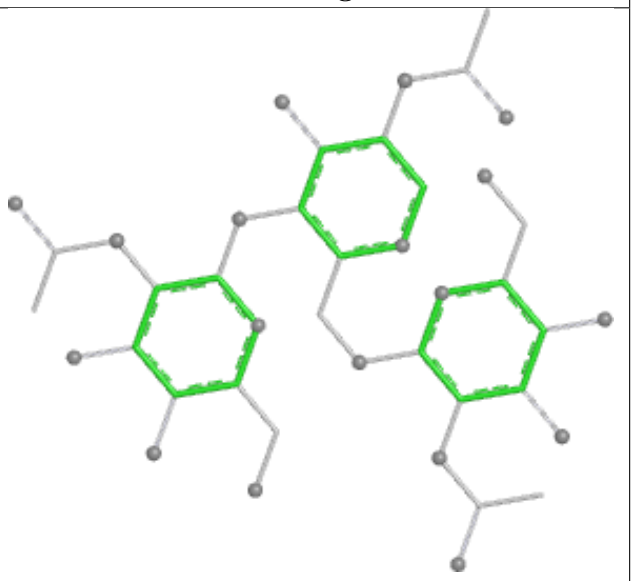
Bond lengths



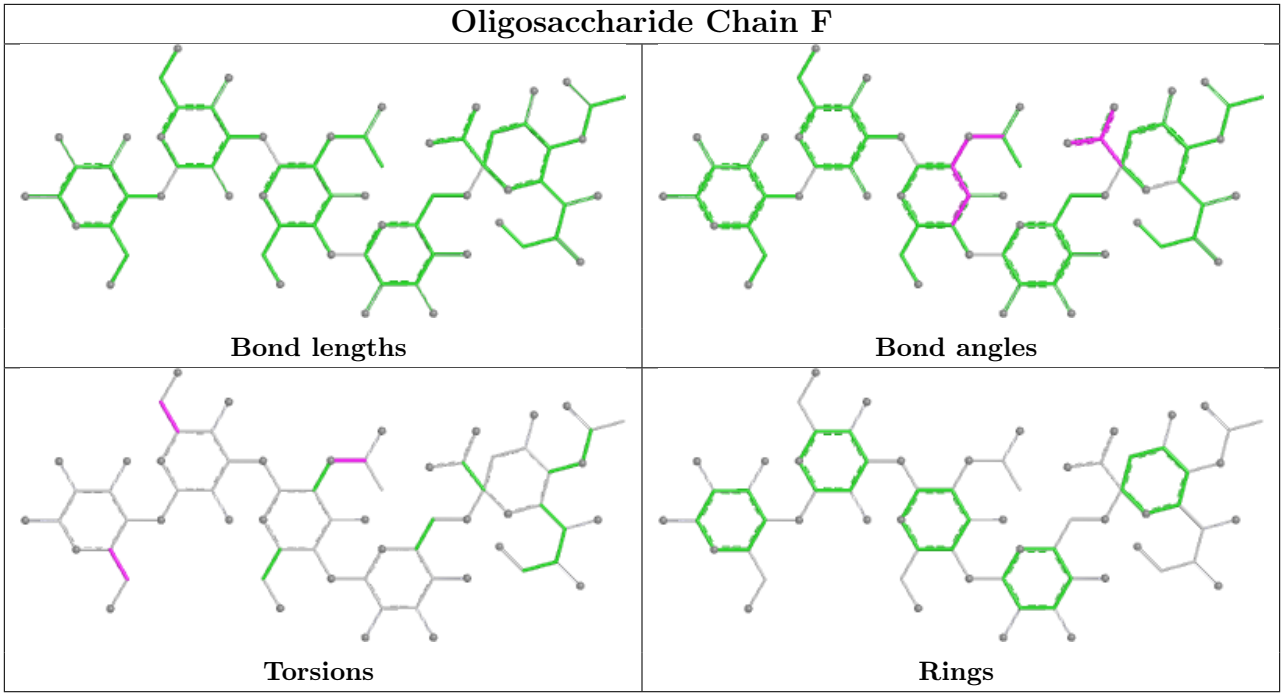
Bond angles



Torsions



Rings



5.6 Ligand geometry (i)

1 ligand is 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	$\#\|Z\| > 2$
6	NAG	A	320	1	14,14,15	0.44	0	17,19,21	0.73	1 (5%)

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
6	NAG	A	320	1	-	4/6/23/26	0/1/1/1

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
6	A	320	NAG	C2-N2-C7	-2.37	119.72	122.90

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	A	320	NAG	O5-C5-C6-O6
6	A	320	NAG	C4-C5-C6-O6
6	A	320	NAG	C8-C7-N2-C2
6	A	320	NAG	O7-C7-N2-C2

There are no ring outliers.

No monomer is involved in short contacts.

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

6.1 Protein, DNA and RNA chains

EDS was not executed - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

6.4 Ligands

EDS was not executed - this section is therefore empty.

6.5 Other polymers

EDS was not executed - this section is therefore empty.