wwPDB X-ray Structure Validation Summary Report (i)

Aug 14, 2023 – 06:19 PM EDT

PDB ID	:	1RVT
Title	:	1930 H1 Hemagglutinin in complex with LSTC
Authors	:	Skehel, J.J.; Gamblin, S.J.; Haire, L.F.; Russell, R.J.; Stevens, D.J.; Xiao, B.; Ha, Y.; Vasisht, N.; Steinhauer, D.A.; Daniels, R.S.
Deposited on	:	2003-12-15
Resolution	:	2.50 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:

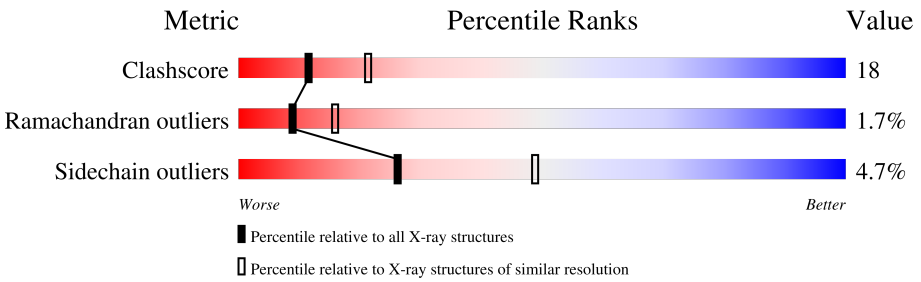
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

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 2.50 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	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)

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	H	328	68% 28% • •
1	J	328	69% 27% • •
1	L	328	68% 28% • •
2	I	160	64% 31% • •
2	K	160	65% 29% • •
2	M	160	64% 31% • •
3	A	2	100%
3	C	2	100%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	B	5	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
5	NDG	H	742	-	-	X	-
5	NDG	J	740	-	-	X	-
5	NDG	L	744	-	-	X	-

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 12143 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.

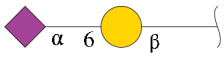
- Molecule 1 is a protein called hemagglutinin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	H	324	2511	1580	436	484	11	0	0	0
1	J	324	2511	1580	436	484	11	0	0	0
1	L	324	2511	1580	436	484	11	0	0	0

- Molecule 2 is a protein called hemagglutinin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	I	160	1281	801	223	252	5	0	0	0
2	K	160	1281	801	223	252	5	0	0	0
2	M	160	1281	801	223	252	5	0	0	0

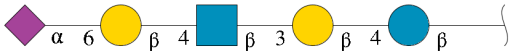
- Molecule 3 is an oligosaccharide called N-acetyl-alpha-neuraminic acid-(2-6)-beta-D-galacto pyranose.



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
3	A	2	31	17	1	13	0	0	0
3	C	2	31	17	1	13	0	0	0

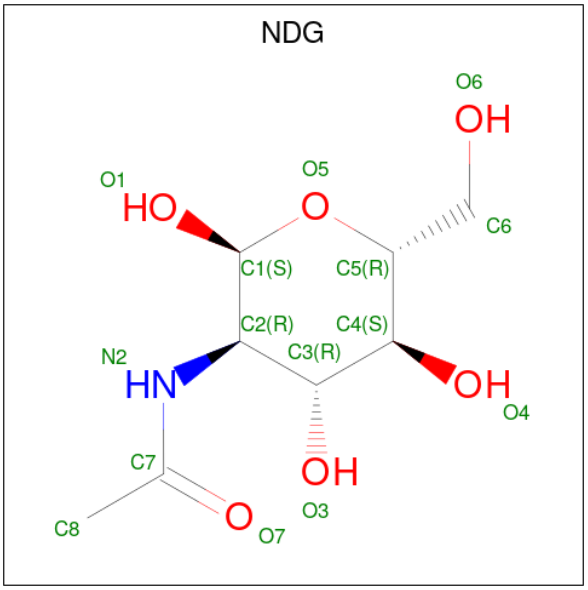
- Molecule 4 is an oligosaccharide called N-acetyl-alpha-neuraminic acid-(2-6)-beta-D-galacto pyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-3)-beta-D-galactopyranose-(1-

4)-beta-D-glucopyranose.



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

- Molecule 5 is 2-acetamido-2-deoxy-alpha-D-glucopyranose (three-letter code: NDG) (formula: $C_8H_{15}NO_6$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
5	H	1	15	8	1	6	0	0
5	J	1	15	8	1	6	0	0
5	L	1	15	8	1	6	0	0

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
6	H	128	128	128	0	0
6	I	36	36	36	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
6	J	151	151	151	0	0
			Total	O		
6	K	44	44	44	0	0
			Total	O		
6	L	191	191	191	0	0
			Total	O		
6	M	42	42	42	0	0

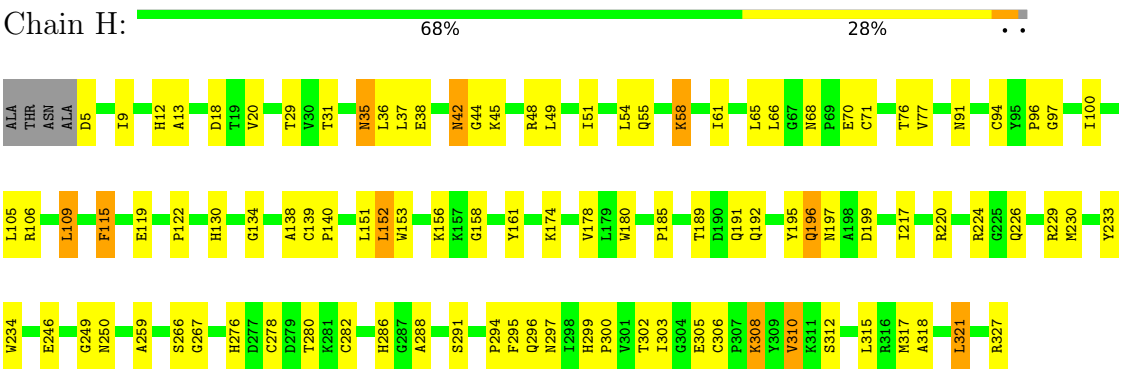
3 Residue-property plots (i)

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.

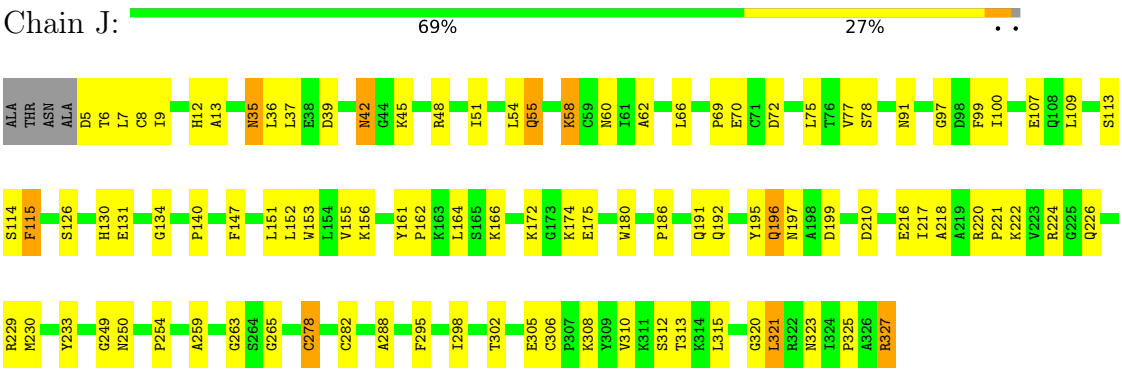
• Molecule 1: hemagglutinin

Chain H:



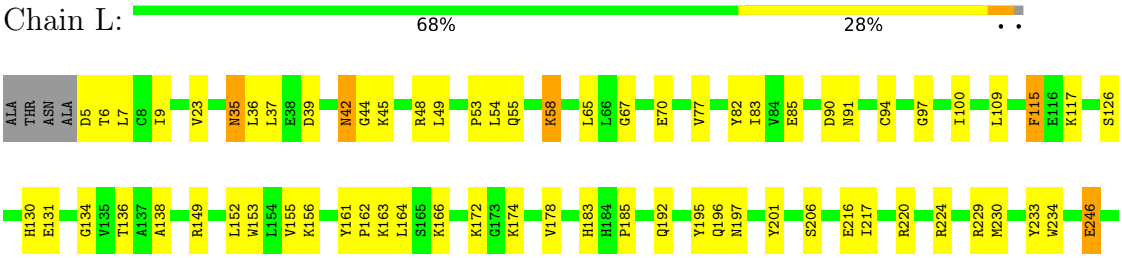
• Molecule 1: hemagglutinin

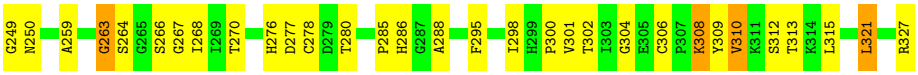
Chain J:



• Molecule 1: hemagglutinin

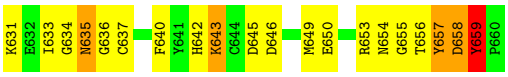
Chain L:





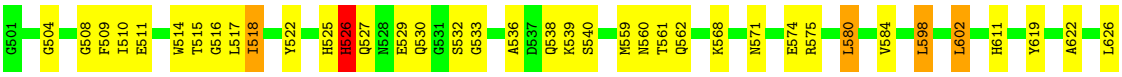
• Molecule 2: hemagglutinin

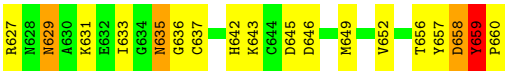
Chain I:



• Molecule 2: hemagglutinin

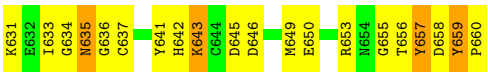
Chain K:





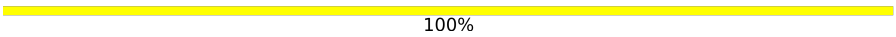
• Molecule 2: hemagglutinin

Chain M:



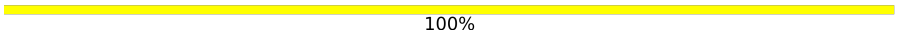
• Molecule 3: N-acetyl-alpha-neuraminic acid-(2-6)-beta-D-galactopyranose

Chain A:



• Molecule 3: N-acetyl-alpha-neuraminic acid-(2-6)-beta-D-galactopyranose

Chain C:



• Molecule 4: 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)-beta-D-glucopyranose

Chain B: 100%

BGC1
GAL2
NAG3
GAL4
SIA5

4 Data and refinement statistics (i)

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

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	204.30Å 83.30Å 177.38Å 90.00° 106.15° 90.00°	Depositor
Resolution (Å)	20.00 – 2.50	Depositor
% Data completeness (in resolution range)	(Not available) (20.00-2.50)	Depositor
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.217 , 0.247	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	12143	wwPDB-VP
Average B, all atoms (Å^2)	44.0	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: SIA, BGC, GAL, NDG, 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	H	0.39	0/2574	0.67	0/3507
1	J	0.36	0/2574	0.65	0/3507
1	L	0.39	0/2574	0.69	0/3507
2	I	0.35	0/1307	0.56	0/1761
2	K	0.34	0/1307	0.56	0/1761
2	M	0.36	0/1307	0.57	0/1761
All	All	0.37	0/11643	0.64	0/15804

There are no bond length outliers.

There are no bond angle outliers.

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	H	2511	0	2437	92	0
1	J	2511	0	2437	93	0
1	L	2511	0	2437	85	0
2	I	1281	0	1201	56	0
2	K	1281	0	1201	54	0
2	M	1281	0	1201	52	0
3	A	31	0	25	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	C	31	0	25	0	0
4	B	68	0	57	0	0
5	H	15	0	12	11	0
5	J	15	0	12	7	0
5	L	15	0	12	8	0
6	H	128	0	0	8	0
6	I	36	0	0	7	0
6	J	151	0	0	6	0
6	K	44	0	0	3	0
6	L	191	0	0	8	0
6	M	42	0	0	4	0
All	All	12143	0	11057	397	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 18.

The worst 5 of 397 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:L:91:ASN:HD21	5:L:744:NDG:C1	1.55	1.19
1:J:58:LYS:H	1:J:58:LYS:HD3	1.06	1.11
1:L:58:LYS:H	1:L:58:LYS:HD3	1.16	1.09
1:H:58:LYS:H	1:H:58:LYS:HD3	1.17	1.04
1:J:224:ARG:NH2	5:J:740:NDG:H2	1.77	0.98

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	H	322/328 (98%)	299 (93%)	22 (7%)	1 (0%)	41 61

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	J	322/328 (98%)	304 (94%)	17 (5%)	1 (0%)	41	61
1	L	322/328 (98%)	301 (94%)	20 (6%)	1 (0%)	41	61
2	I	158/160 (99%)	134 (85%)	16 (10%)	8 (5%)	2	2
2	K	158/160 (99%)	139 (88%)	12 (8%)	7 (4%)	2	3
2	M	158/160 (99%)	139 (88%)	13 (8%)	6 (4%)	3	4
All	All	1440/1464 (98%)	1316 (91%)	100 (7%)	24 (2%)	9	16

5 of 24 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	I	635	ASN
2	K	657	TYR
2	K	658	ASP
2	K	659	TYR
2	M	635	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	H	279/281 (99%)	268 (96%)	11 (4%)	32	57
1	J	279/281 (99%)	266 (95%)	13 (5%)	26	49
1	L	279/281 (99%)	265 (95%)	14 (5%)	24	46
2	I	134/134 (100%)	128 (96%)	6 (4%)	27	51
2	K	134/134 (100%)	126 (94%)	8 (6%)	19	37
2	M	134/134 (100%)	128 (96%)	6 (4%)	27	51
All	All	1239/1245 (100%)	1181 (95%)	58 (5%)	26	49

5 of 58 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	J	321	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	M	598	LEU
2	K	629	ASN
2	M	580	LEU
1	L	313	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 34 such sidechains are listed below:

Mol	Chain	Res	Type
1	L	191	GLN
1	L	196	GLN
2	M	625	GLN
1	J	42	ASN
1	J	35	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)

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

5.5 Carbohydrates (i)

9 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	GAL	A	1	3	11,11,12	3.06	6 (54%)	15,15,17	1.55	3 (20%)
3	SIA	A	2	3	20,20,21	4.51	7 (35%)	24,28,31	3.29	9 (37%)
4	BGC	B	1	4	12,12,12	2.59	6 (50%)	17,17,17	0.61	0
4	GAL	B	2	4	11,11,12	2.31	3 (27%)	15,15,17	1.13	1 (6%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
4	NAG	B	3	4	14,14,15	1.94	6 (42%)	17,19,21	0.90	0
4	GAL	B	4	4	11,11,12	2.74	6 (54%)	15,15,17	1.45	3 (20%)
4	SIA	B	5	4	20,20,21	4.30	6 (30%)	24,28,31	3.34	10 (41%)
3	GAL	C	1	3	11,11,12	2.86	6 (54%)	15,15,17	1.15	1 (6%)
3	SIA	C	2	3	20,20,21	4.56	7 (35%)	24,28,31	3.25	8 (33%)

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	GAL	A	1	3	-	0/2/19/22	0/1/1/1
3	SIA	A	2	3	-	3/18/34/38	0/1/1/1
4	BGC	B	1	4	-	0/2/22/22	0/1/1/1
4	GAL	B	2	4	-	0/2/19/22	0/1/1/1
4	NAG	B	3	4	-	0/6/23/26	0/1/1/1
4	GAL	B	4	4	-	0/2/19/22	0/1/1/1
4	SIA	B	5	4	-	3/18/34/38	0/1/1/1
3	GAL	C	1	3	-	0/2/19/22	0/1/1/1
3	SIA	C	2	3	-	4/18/34/38	0/1/1/1

The worst 5 of 53 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	2	SIA	C2-C1	15.17	1.65	1.52
3	A	2	SIA	C2-C1	14.37	1.65	1.52
4	B	5	SIA	C2-C1	13.49	1.64	1.52
3	C	2	SIA	O4-C4	-7.84	1.26	1.43
4	B	5	SIA	O4-C4	-7.45	1.27	1.43

The worst 5 of 35 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	2	SIA	C6-O6-C2	12.33	137.70	111.34
4	B	5	SIA	C6-O6-C2	12.26	137.56	111.34
3	C	2	SIA	C6-O6-C2	11.95	136.89	111.34
4	B	5	SIA	O6-C2-C3	-4.70	103.99	110.46
3	C	2	SIA	O6-C2-C3	-4.63	104.09	110.46

There are no chirality outliers.

5 of 10 torsion outliers are listed below:

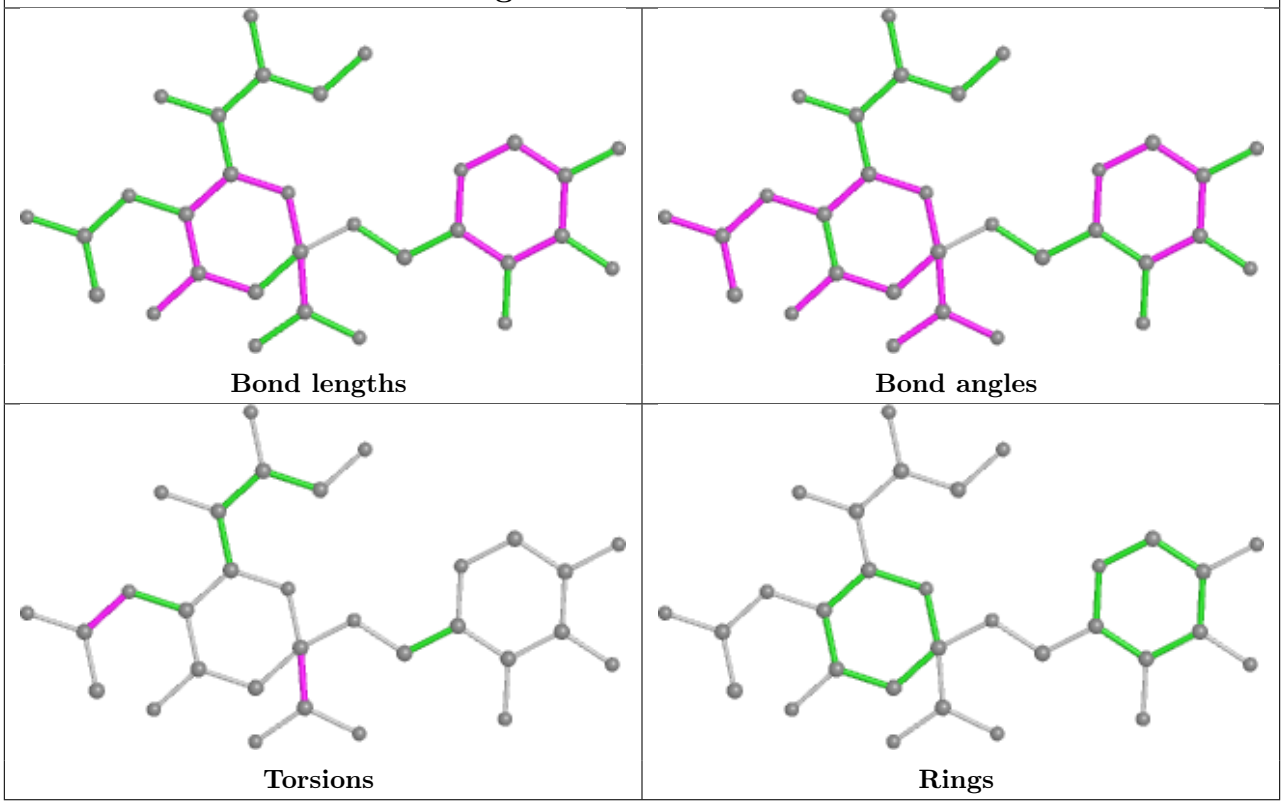
Mol	Chain	Res	Type	Atoms
3	A	2	SIA	C11-C10-N5-C5
3	A	2	SIA	O10-C10-N5-C5
3	C	2	SIA	C11-C10-N5-C5
3	C	2	SIA	O10-C10-N5-C5
4	B	5	SIA	C11-C10-N5-C5

There are no ring outliers.

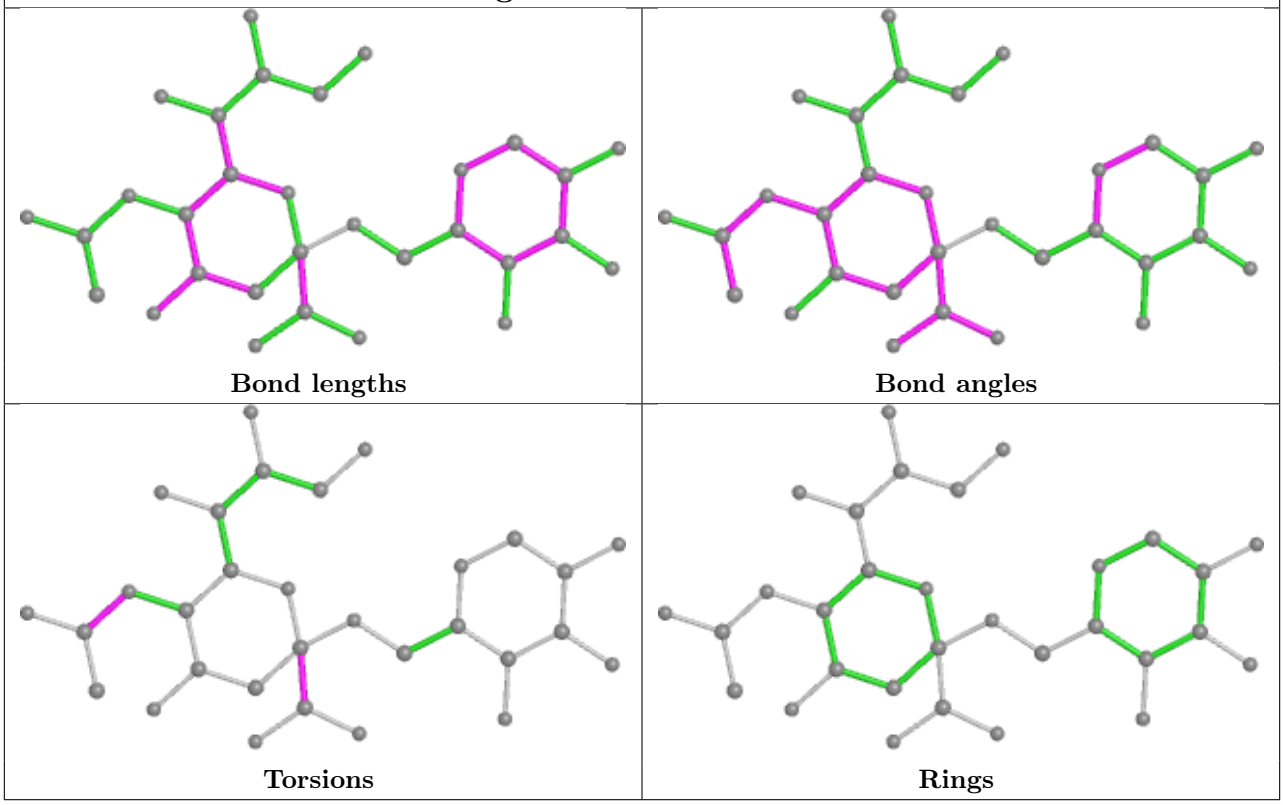
No monomer is involved in short contacts.

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

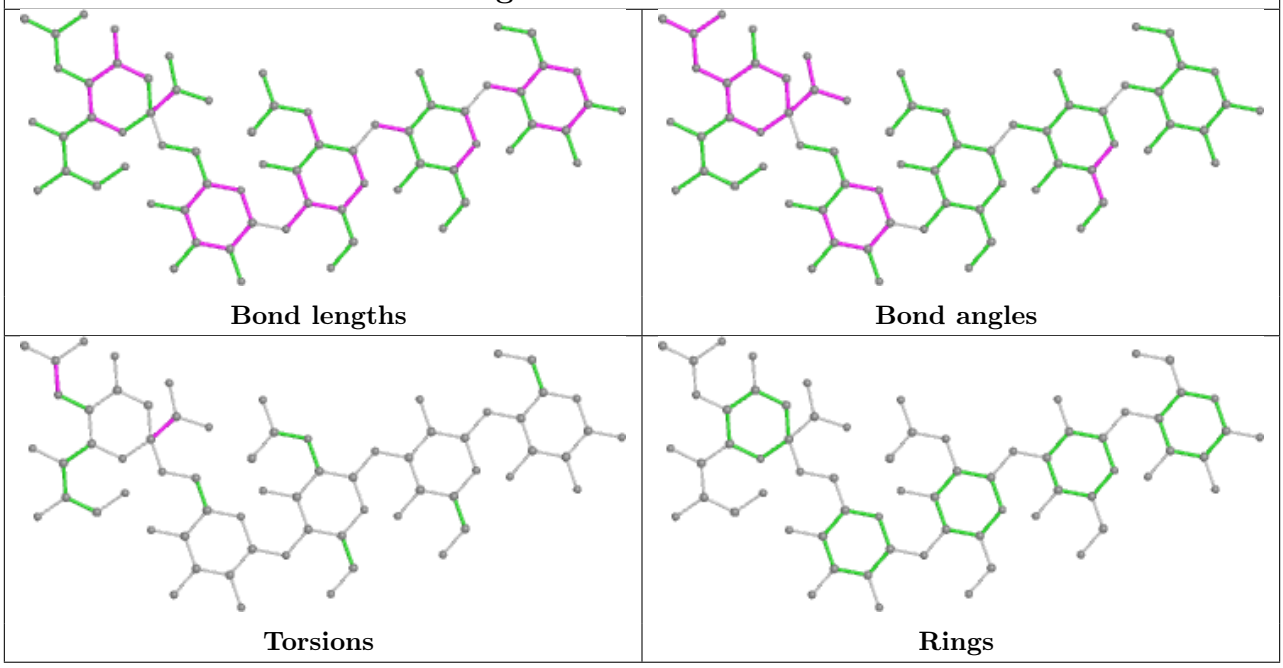
Oligosaccharide Chain A



Oligosaccharide Chain C



Oligosaccharide Chain B



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	$\#\|Z\| > 2$
5	NDG	J	740	-	15,15,15	0.44	0	21,21,21	0.71	0
5	NDG	L	744	-	15,15,15	0.49	0	21,21,21	0.79	0
5	NDG	H	742	-	15,15,15	0.43	0	21,21,21	0.80	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	NDG	J	740	-	-	6/6/26/26	0/1/1/1
5	NDG	L	744	-	-	2/6/26/26	0/1/1/1
5	NDG	H	742	-	-	4/6/26/26	0/1/1/1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 12 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	H	742	NDG	C3-C2-N2-C7
5	J	740	NDG	C3-C2-N2-C7
5	L	744	NDG	C3-C2-N2-C7
5	H	742	NDG	C8-C7-N2-C2
5	J	740	NDG	C8-C7-N2-C2

There are no ring outliers.

3 monomers are involved in 26 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	J	740	NDG	7	0
5	L	744	NDG	8	0
5	H	742	NDG	11	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.