wwPDB X-ray Structure Validation Summary Report (i)

Sep 16, 2021 – 12:01 pm BST

PDB ID	:	7B7N
Title	:	Human herpesvirus-8 gH/gL in complex with EphA2
Authors	:	Pederzoli, R.; Guardado-Calvo, P.; Rey, F.A.; Backovic, M.
Deposited on	:	2020-12-11
Resolution	:	2.69 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 following versions of software and data (see references (i)) were used in the production of this report:

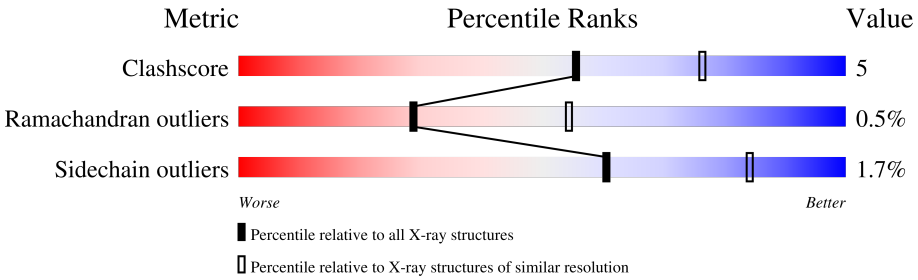
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.1

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 2.69 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	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)

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	E	215	65% 9% 27%
2	H	681	78% 15% 6%
3	L	184	53% 7% 40%
4	A	2	100%
4	B	2	100%

2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 7234 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 Ephrin type-A receptor 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	E	158	Total	C	N	O	S	0	0	0
			1273	820	209	235	9			

There are 35 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	203	ASP	-	expression tag	UNP P29317
E	204	ASP	-	expression tag	UNP P29317
E	205	ASP	-	expression tag	UNP P29317
E	206	ASP	-	expression tag	UNP P29317
E	207	LYS	-	expression tag	UNP P29317
E	208	ALA	-	expression tag	UNP P29317
E	209	GLY	-	expression tag	UNP P29317
E	210	TRP	-	expression tag	UNP P29317
E	211	SER	-	expression tag	UNP P29317
E	212	HIS	-	expression tag	UNP P29317
E	213	PRO	-	expression tag	UNP P29317
E	214	GLN	-	expression tag	UNP P29317
E	215	PHE	-	expression tag	UNP P29317
E	216	GLU	-	expression tag	UNP P29317
E	217	LYS	-	expression tag	UNP P29317
E	218	GLY	-	expression tag	UNP P29317
E	219	GLY	-	expression tag	UNP P29317
E	220	GLY	-	expression tag	UNP P29317
E	221	SER	-	expression tag	UNP P29317
E	222	GLY	-	expression tag	UNP P29317
E	223	GLY	-	expression tag	UNP P29317
E	224	GLY	-	expression tag	UNP P29317
E	225	SER	-	expression tag	UNP P29317
E	226	GLY	-	expression tag	UNP P29317
E	227	GLY	-	expression tag	UNP P29317
E	228	GLY	-	expression tag	UNP P29317
E	229	SER	-	expression tag	UNP P29317

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	230	TRP	-	expression tag	UNP P29317
E	231	SER	-	expression tag	UNP P29317
E	232	HIS	-	expression tag	UNP P29317
E	233	PRO	-	expression tag	UNP P29317
E	234	GLN	-	expression tag	UNP P29317
E	235	PHE	-	expression tag	UNP P29317
E	236	GLU	-	expression tag	UNP P29317
E	237	LYS	-	expression tag	UNP P29317

- Molecule 2 is a protein called Envelope glycoprotein H.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	H	637	Total	C	N	O	S	0	0	0
			4962	3178	836	921	27			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
H	24	ARG	-	cloning artifact	UNP Q98142
H	25	SER	-	cloning artifact	UNP Q98142

- Molecule 3 is a protein called Envelope glycoprotein L.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	L	110	Total	C	N	O	S	0	0	0
			847	537	153	153	4			

There are 38 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
L	19	ARG	-	cloning artifact	UNP Q76RG7
L	20	SER	-	cloning artifact	UNP Q76RG7
L	58	SER	CYS	engineered mutation	UNP Q76RG7
L	168	ASP	-	expression tag	UNP Q76RG7
L	169	ASP	-	expression tag	UNP Q76RG7
L	170	ASP	-	expression tag	UNP Q76RG7
L	171	ASP	-	expression tag	UNP Q76RG7
L	172	LYS	-	expression tag	UNP Q76RG7
L	173	ALA	-	expression tag	UNP Q76RG7
L	174	GLY	-	expression tag	UNP Q76RG7
L	175	TRP	-	expression tag	UNP Q76RG7
L	176	SER	-	expression tag	UNP Q76RG7

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
L	177	HIS	-	expression tag	UNP Q76RG7
L	178	PRO	-	expression tag	UNP Q76RG7
L	179	GLN	-	expression tag	UNP Q76RG7
L	180	PHE	-	expression tag	UNP Q76RG7
L	181	GLU	-	expression tag	UNP Q76RG7
L	182	LYS	-	expression tag	UNP Q76RG7
L	183	GLY	-	expression tag	UNP Q76RG7
L	184	GLY	-	expression tag	UNP Q76RG7
L	185	GLY	-	expression tag	UNP Q76RG7
L	186	SER	-	expression tag	UNP Q76RG7
L	187	GLY	-	expression tag	UNP Q76RG7
L	188	GLY	-	expression tag	UNP Q76RG7
L	189	GLY	-	expression tag	UNP Q76RG7
L	190	SER	-	expression tag	UNP Q76RG7
L	191	GLY	-	expression tag	UNP Q76RG7
L	192	GLY	-	expression tag	UNP Q76RG7
L	193	GLY	-	expression tag	UNP Q76RG7
L	194	SER	-	expression tag	UNP Q76RG7
L	195	TRP	-	expression tag	UNP Q76RG7
L	196	SER	-	expression tag	UNP Q76RG7
L	197	HIS	-	expression tag	UNP Q76RG7
L	198	PRO	-	expression tag	UNP Q76RG7
L	199	GLN	-	expression tag	UNP Q76RG7
L	200	PHE	-	expression tag	UNP Q76RG7
L	201	GLU	-	expression tag	UNP Q76RG7
L	202	LYS	-	expression tag	UNP Q76RG7

- Molecule 4 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
4	A	2	Total	C	N	O	0	0	0
			28	16	2	10			
4	B	2	Total	C	N	O	0	0	0
			28	16	2	10			

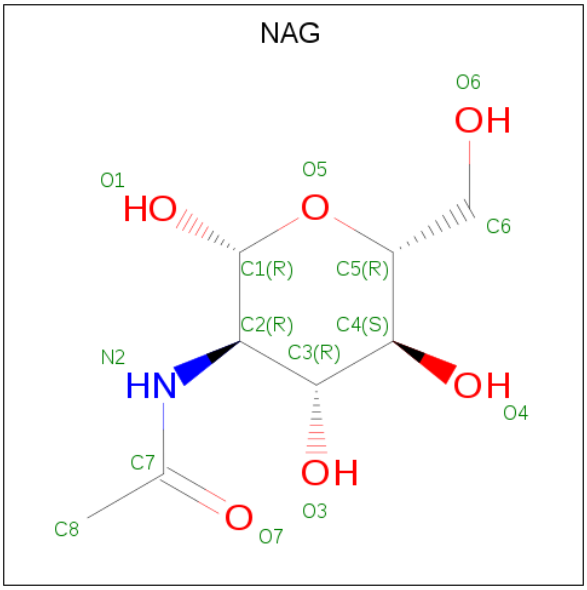
- Molecule 5 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	E	3	Total Cl 3 3	0	0
5	L	1	Total Cl 1 1	0	0

- Molecule 6 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	E	2	Total Na 2 2	0	0
6	H	5	Total Na 5 5	0	0
6	L	3	Total Na 3 3	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
7	H	1	Total C N O 14 8 1 5	0	0
7	L	1	Total C N O 14 8 1 5	0	0

- Molecule 8 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	E	15	Total O 15 15	0	0
8	H	22	Total O 22 22	0	0
8	L	17	Total O 17 17	0	0

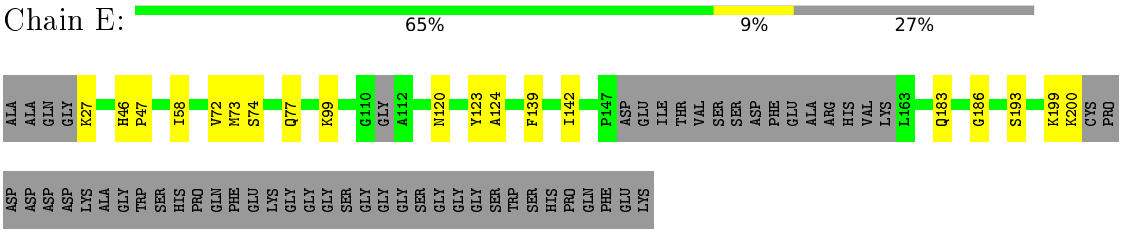
3 Residue-property plots (i)

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.

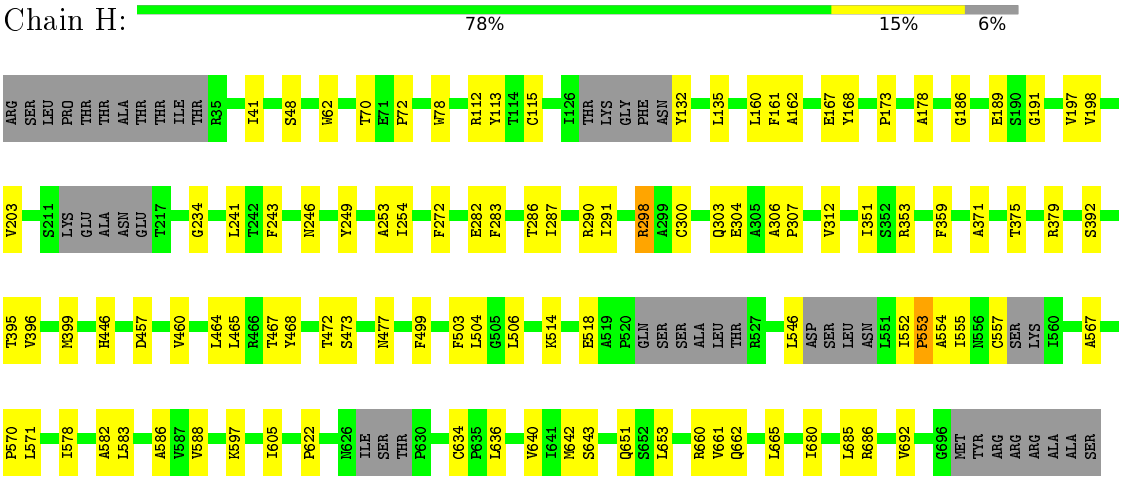
- Molecule 1: Ephrin type-A receptor 2

Chain E:



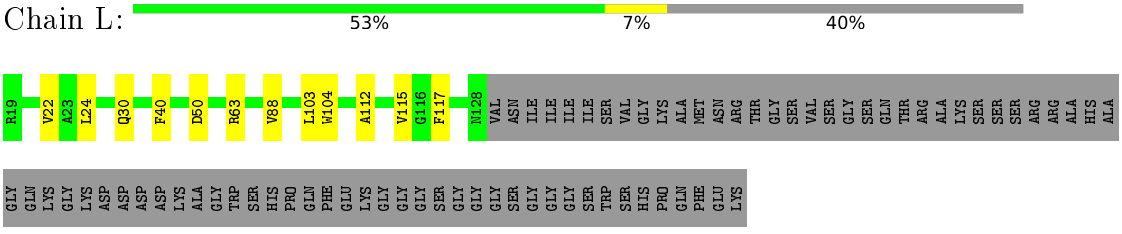
- Molecule 2: Envelope glycoprotein H

Chain H:



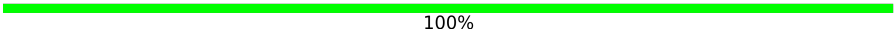
- Molecule 3: Envelope glycoprotein L

Chain L:



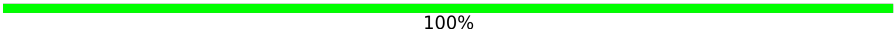
- Molecule 4: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain A:



• Molecule 4: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-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 21	Depositor
Cell constants a, b, c, α, β, γ	72.94Å 129.00Å 267.69Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.44 – 2.69	Depositor
% Data completeness (in resolution range)	99.6 (46.44-2.69)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	PHENIX 1.17.1_3660, BUSTER	Depositor
R, R_{free}	0.223 , 0.265	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	7234	wwPDB-VP
Average B, all atoms (Å^2)	76.0	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: CL, NAG, NA

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	E	0.26	0/1303	0.47	0/1762
2	H	0.25	0/5067	0.44	0/6891
3	L	0.26	0/866	0.42	0/1180
All	All	0.25	0/7236	0.44	0/9833

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

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	E	1273	0	1224	11	0
2	H	4962	0	4913	57	0
3	L	847	0	835	6	0
4	A	28	0	25	0	0
4	B	28	0	25	0	0
5	E	3	0	0	0	0
5	L	1	0	0	0	0
6	E	2	0	0	0	0
6	H	5	0	0	0	0
6	L	3	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	H	14	0	13	0	0
7	L	14	0	13	0	0
8	E	15	0	0	0	0
8	H	22	0	0	0	0
8	L	17	0	0	0	0
All	All	7234	0	7048	72	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 5.

The worst 5 of 72 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)
2:H:472:THR:HG22	2:H:506:LEU:H	1.51	0.75
1:E:99:LYS:HB2	1:E:193:SER:HB3	1.72	0.71
2:H:477:ASN:ND2	2:H:651:GLN:OE1	2.30	0.64
2:H:198:VAL:HG22	2:H:203:VAL:HG22	1.82	0.62
2:H:642:MET:HE2	2:H:685:LEU:HD11	1.81	0.61

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	E	152/215 (71%)	143 (94%)	9 (6%)	0	100 100
2	H	623/681 (92%)	594 (95%)	25 (4%)	4 (1%)	25 50
3	L	108/184 (59%)	106 (98%)	2 (2%)	0	100 100
All	All	883/1080 (82%)	843 (96%)	36 (4%)	4 (0%)	29 54

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	H	553	PRO
2	H	234	GLY
2	H	555	ILE
2	H	570	PRO

5.3.2 Protein sidechains

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	E	135/177 (76%)	135 (100%)	0	100	100
2	H	543/591 (92%)	533 (98%)	10 (2%)	59	83
3	L	92/146 (63%)	89 (97%)	3 (3%)	38	67
All	All	770/914 (84%)	757 (98%)	13 (2%)	60	84

5 of 13 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	H	557	CYS
2	H	660	ARG
3	L	117	PHE
3	L	50	ASP
3	L	104	TRP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates (i)

4 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	$\#\|Z\|>2$
4	NAG	A	1	4,2	14,14,15	0.40	0	17,19,21	0.48	0
4	NAG	A	2	4	14,14,15	0.27	0	17,19,21	0.38	0
4	NAG	B	1	4,2	14,14,15	0.29	0	17,19,21	0.44	0
4	NAG	B	2	4	14,14,15	0.23	0	17,19,21	0.51	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
4	NAG	A	1	4,2	-	1/6/23/26	0/1/1/1
4	NAG	A	2	4	-	0/6/23/26	0/1/1/1
4	NAG	B	1	4,2	-	0/6/23/26	0/1/1/1
4	NAG	B	2	4	-	2/6/23/26	0/1/1/1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (3) torsion outliers are listed below:

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

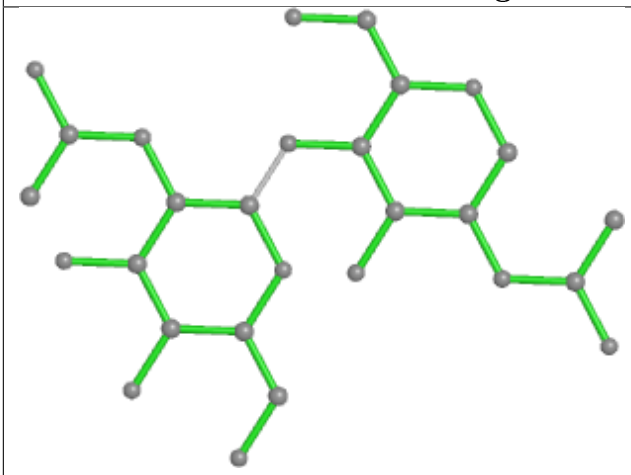
There are no ring outliers.

No monomer is involved in short contacts.

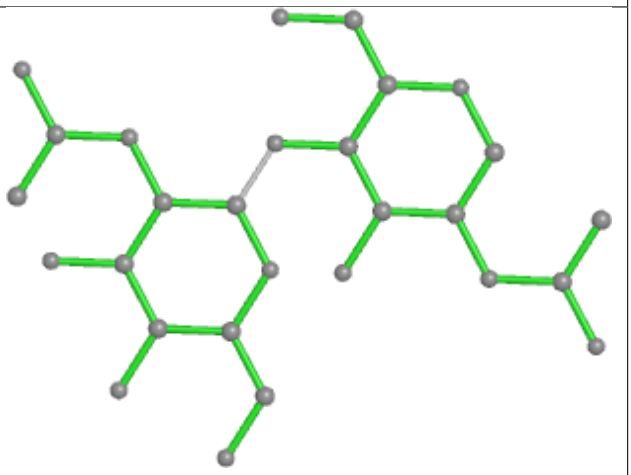
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths,

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

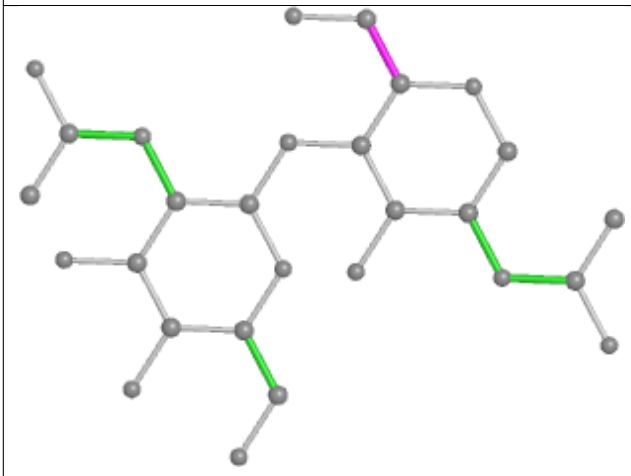
Oligosaccharide Chain A



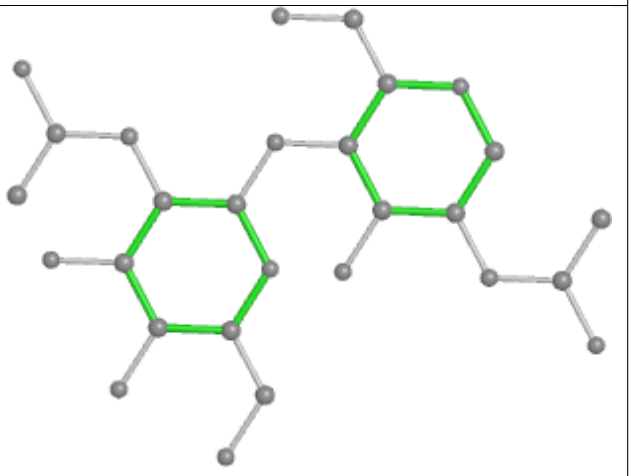
Bond lengths



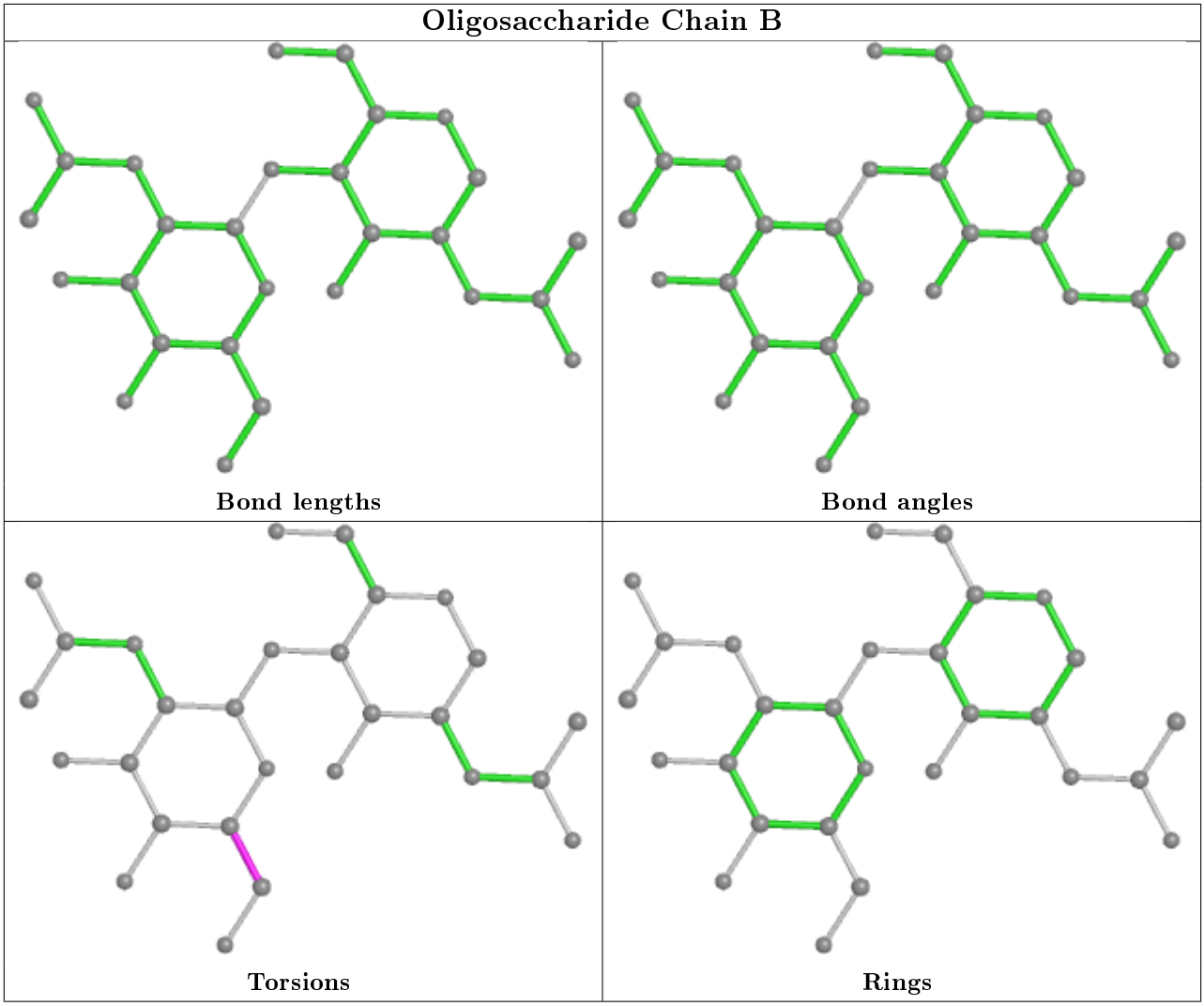
Bond angles



Torsions



Rings



5.6 Ligand geometry

Of 16 ligands modelled in this entry, 14 are monoatomic - leaving 2 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	$\#\|Z\| > 2$
7	NAG	L	301	3	14,14,15	0.47	0	17,19,21	0.65	1 (5%)
7	NAG	H	801	2	14,14,15	0.21	0	17,19,21	0.40	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
7	NAG	L	301	3	-	2/6/23/26	0/1/1/1
7	NAG	H	801	2	-	0/6/23/26	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})
7	L	301	NAG	C1-O5-C5	2.29	115.30	112.19

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
7	L	301	NAG	C4-C5-C6-O6
7	L	301	NAG	O5-C5-C6-O6

There are no ring outliers.

No monomer is involved in short contacts.

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

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.