wwPDB X-ray Structure Validation Summary Report (i)

Jul 31, 2023 – 02:26 AM EDT

PDB ID	:	1GC1
Title	:	HIV-1 GP120 CORE COMPLEXED WITH CD4 AND A NEUTRALIZING HUMAN ANTIBODY
Authors	:	Kwong, P.D.; Wyatt, R.; Robinson, J.; Sweet, R.W.; Sodroski, J.; Hendrickson, W.A.
Deposited on	:	1998-06-15
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:

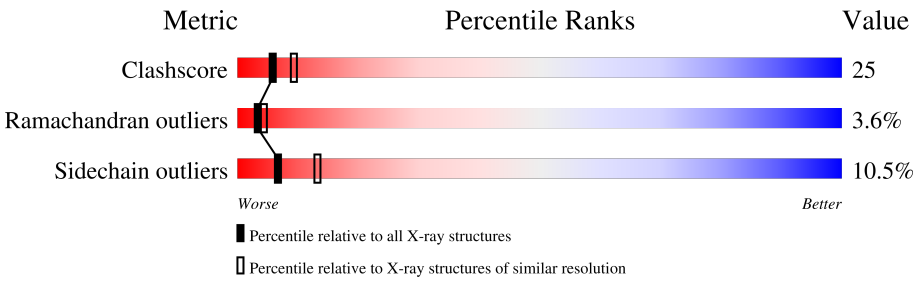
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.34

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	G	321	44% 42% 6% • 7%
2	C	185	41% 51% 6% •
3	L	213	53% 38% 8%
4	H	229	62% 30% 8%
5	A	2	50% 50%
5	B	2	50% 50%
5	D	2	100%
5	E	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
5	NAG	A	1	X	-	-	-

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 7877 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 ENVELOPE PROTEIN GP120.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	G	297	2300	1441	401	438	20	0	0	0

There are 11 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
G	79	GLY	-	expression tag	UNP P04578
G	80	ALA	-	expression tag	UNP P04578
G	81	ARG	-	expression tag	UNP P04578
G	82	SER	-	expression tag	UNP P04578
G	128	GLY	-	linker	UNP P04578
G	129	ALA	-	linker	UNP P04578
G	194	GLY	-	linker	UNP P04578
G	298	GLY	-	linker	UNP P04578
G	299	ALA	-	linker	UNP P04578
G	329	GLY	-	linker	UNP P04578
G	403	LYS	GLU	conflict	UNP P04578

- Molecule 2 is a protein called CD4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	C	181	1412	885	247	276	4	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	184	ASN	SER	engineered mutation	UNP P01730
C	185	THR	ILE	engineered mutation	UNP P01730

- Molecule 3 is a protein called ANTIBODY 17B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	L	213	1645	1028	281	332	4	0	0	0

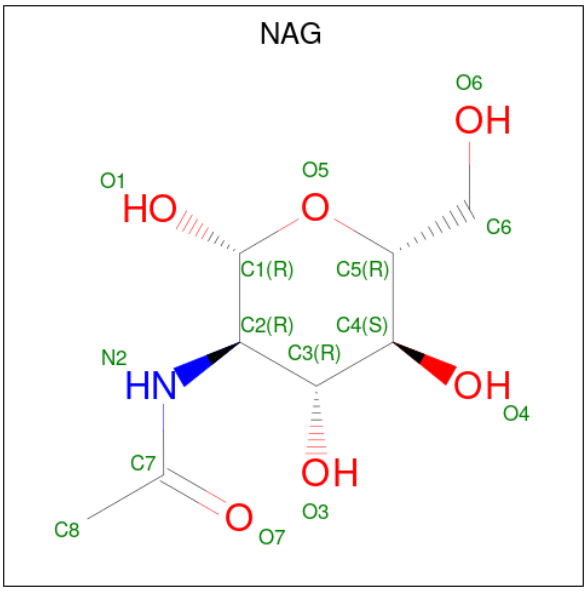
- Molecule 4 is a protein called ANTIBODY 17B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	H	229	1723	1088	291	339	5	0	0	0

- Molecule 5 is an oligosaccharide called alpha-L-fucopyranose-(1-6)-2-acetamido-2-deoxy-beta-D-glucopyranose.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
5	A	2	24	14	1	9	0	0	0
5	B	2	24	14	1	9	0	0	0
5	D	2	24	14	1	9	0	0	0
5	E	2	24	14	1	9	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
			Total	C	N	O		
6	G	1	14	8	1	5	0	0
6	G	1	14	8	1	5	0	0
6	G	1	14	8	1	5	0	0
6	G	1	14	8	1	5	0	0
6	G	1	14	8	1	5	0	0
6	G	1	14	8	1	5	0	0
6	G	1	14	8	1	5	0	0

- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
7	G	208	208	208	1	0
7	C	116	116	116	0	0
7	L	136	136	136	0	0
7	H	143	143	143	0	0

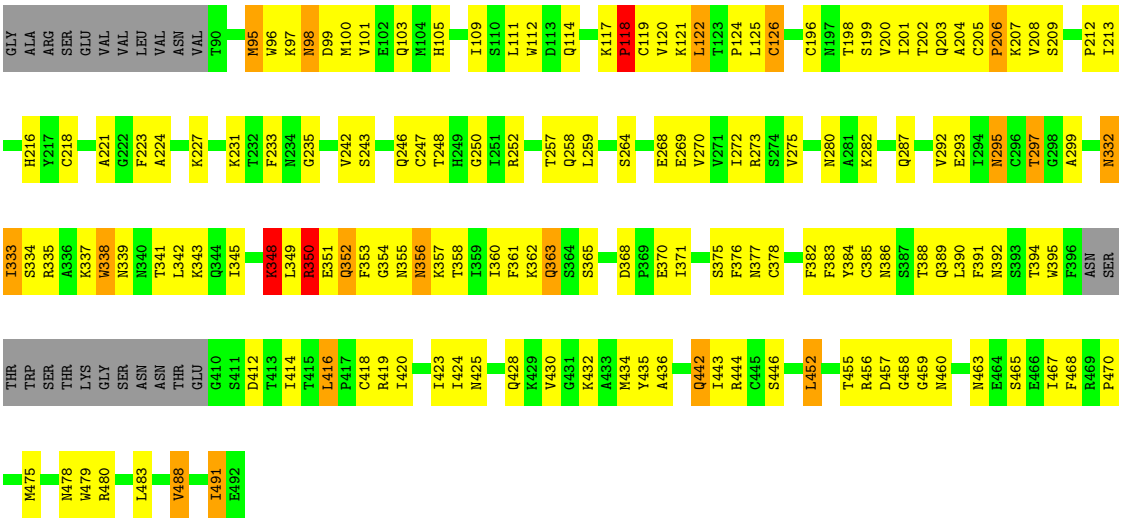
3 Residue-property plots (i)

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.

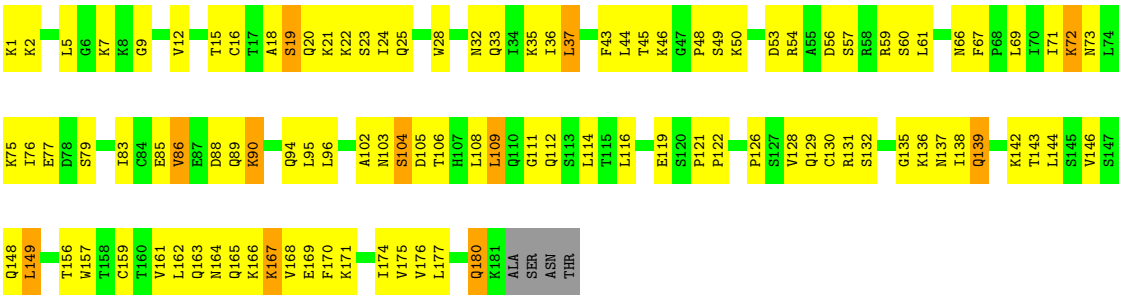
• Molecule 1: ENVELOPE PROTEIN GP120

Chain G: 44% 42% 6% • 7%



• Molecule 2: CD4

Chain C: 41% 51% 6% •



• Molecule 3: ANTIBODY 17B

Chain L: 53% 38% 8%

E1 L2 E3 L4 T5 Q6 S7 P8 A9 T10 L11 S14 P15 G16 E17 R18 A19 T20 L21 R24 A25 S30 S31 D32 L33 A34 W35 Y36 K39 P40 R45 L46 L47 A55 P59 A60 R61 S65 T72 S76 S77 S80 E81 D82 F83 A84 V85 Y86

Q89 Q90 Y91 N92 N93 W94 P95 P96 R97 Y98 T104 R105 L106 E107 I108 K109 P115 F118 I119 P122 S123 D124 L127 K128 S133 L137 L138 N139 N140 F141 Y142 P143 Q149 W150 K151 V152 D153 N154 A155 L156 S161 Q162 E163 K171 S176 L177 S178 T182

L183 S184 K185 A186 D187 Y188 E189 K190 H191 K192 V193 Y194 A195 H200 Q201 G202 L203 S204 S205 P206 V207 T208 K209 S210 F211 N212 R213

- Molecule 4: ANTIBODY 17B

Chain H: 62% 30% 8%

Q1 L5 V18 K19 K23 D27 T28 F29 I30 R31 Y32 T35 R38 P41 E46 W47 M48 G49 I52 T53 I54 L55 D56 Y60 H63 L64 V68 T69 I70 K74 T78 V79 Y80 L81 E82 L83 R84 N85 L86 R87 T91 Y94 V99

Y100 E101 G102 E103 A104 D105 Y109 N112 Q120 G121 T125 V126 T127 S128 A129 P134 P138 L139 A140 P141 S142 S143 T146 A151 A152 L153 G154 Y160 F161 P162 Q163 P164 V165 T166 V167 S168 W169 L174 L185 Y191 V196 V197 T198 V199 L204 G205

T206 Y209 I210 C211 N212 H215 K216 P217 K224 K225 V226 K229

- Molecule 5: alpha-L-fucopyranose-(1-6)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain A: 50% 50%

NAG1 FUC2

- Molecule 5: alpha-L-fucopyranose-(1-6)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain B: 50% 50%

NAG1 FUC2

- Molecule 5: alpha-L-fucopyranose-(1-6)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain D: 100%

NAG1 FUC2

- Molecule 5: alpha-L-fucopyranose-(1-6)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain E: 100%

NAG1 FUC2

4 Data and refinement statistics (i)

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

Property	Value	Source
Space group	P 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	71.64Å 88.13Å 196.70Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	5.00 – 2.50	Depositor
% Data completeness (in resolution range)	74.9 (5.00-2.50)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	0.09	Depositor
Refinement program	X-PLOR 3.8	Depositor
R, R_{free}	0.210 , 0.302	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	7877	wwPDB-VP
Average B, all atoms (Å^2)	21.0	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: FUC, 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	G	0.39	0/2346	0.75	0/3177
2	C	0.38	0/1432	0.71	0/1930
3	L	0.40	0/1682	0.74	0/2287
4	H	0.41	0/1763	0.71	0/2402
All	All	0.40	0/7223	0.73	0/9796

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	G	2300	0	2242	145	0
2	C	1412	0	1444	79	0
3	L	1645	0	1591	77	0
4	H	1723	0	1699	64	0
5	A	24	0	22	0	0
5	B	24	0	22	2	0
5	D	24	0	22	4	0
5	E	24	0	22	4	0
6	G	98	0	91	9	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	C	116	0	0	10	0
7	G	208	0	0	30	0
7	H	143	0	0	4	0
7	L	136	0	0	10	0
All	All	7877	0	7155	363	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 363 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)
4:H:138:PRO:HD3	4:H:224:LYS:HE2	1.45	0.96
2:C:111:GLY:HA2	2:C:148:GLN:HG3	1.49	0.94
1:G:124:PRO:HD2	1:G:198:THR:HA	1.52	0.90
3:L:151:LYS:HE2	3:L:154:ASN:HA	1.59	0.85
1:G:363:GLN:HE21	1:G:470:PRO:HG3	1.43	0.83

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	G	293/321 (91%)	242 (83%)	40 (14%)	11 (4%)	3 4
2	C	179/185 (97%)	150 (84%)	24 (13%)	5 (3%)	5 7
3	L	211/213 (99%)	185 (88%)	19 (9%)	7 (3%)	4 5
4	H	227/229 (99%)	201 (88%)	16 (7%)	10 (4%)	2 3
All	All	910/948 (96%)	778 (86%)	99 (11%)	33 (4%)	3 4

5 of 33 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	G	99	ASP
1	G	118	PRO
1	G	258	GLN
1	G	350	ARG
1	G	352	GLN

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	G	262/283 (93%)	240 (92%)	22 (8%)	11 21
2	C	164/167 (98%)	144 (88%)	20 (12%)	5 9
3	L	184/184 (100%)	161 (88%)	23 (12%)	4 8
4	H	193/193 (100%)	174 (90%)	19 (10%)	8 15
All	All	803/827 (97%)	719 (90%)	84 (10%)	7 13

5 of 84 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	L	153	ASP
4	H	53	THR
3	L	161	SER
4	H	5	LEU
4	H	81	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 22 such sidechains are listed below:

Mol	Chain	Res	Type
3	L	89	GLN
3	L	191	HIS
3	L	93	ASN
4	H	85	ASN
1	G	363	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)

8 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	$\#\|Z\| > 2$
5	NAG	A	1	5,1	14,14,15	0.64	0	17,19,21	1.22	1 (5%)
5	FUC	A	2	5	10,10,11	0.42	0	14,14,16	0.47	0
5	NAG	B	1	5,1	14,14,15	0.52	0	17,19,21	1.07	1 (5%)
5	FUC	B	2	5	10,10,11	0.60	0	14,14,16	0.62	0
5	NAG	D	1	5,1	14,14,15	0.50	0	17,19,21	0.58	0
5	FUC	D	2	5	10,10,11	0.57	0	14,14,16	0.84	1 (7%)
5	NAG	E	1	5,1	14,14,15	0.59	0	17,19,21	0.86	0
5	FUC	E	2	5	10,10,11	0.59	0	14,14,16	0.69	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	NAG	A	1	5,1	1/1/5/7	1/6/23/26	0/1/1/1
5	FUC	A	2	5	-	-	0/1/1/1
5	NAG	B	1	5,1	-	2/6/23/26	0/1/1/1
5	FUC	B	2	5	-	-	0/1/1/1
5	NAG	D	1	5,1	-	2/6/23/26	0/1/1/1
5	FUC	D	2	5	-	-	0/1/1/1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	NAG	E	1	5,1	-	1/6/23/26	0/1/1/1
5	FUC	E	2	5	-	-	0/1/1/1

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	A	1	NAG	C1-O5-C5	2.68	115.82	112.19
5	B	1	NAG	C4-C3-C2	-2.58	107.24	111.02
5	D	2	FUC	C1-C2-C3	2.25	112.43	109.67

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
5	A	1	NAG	C1

5 of 6 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	1	NAG	C3-C2-N2-C7
5	B	1	NAG	C3-C2-N2-C7
5	D	1	NAG	O5-C5-C6-O6
5	D	1	NAG	C4-C5-C6-O6
5	E	1	NAG	O5-C5-C6-O6

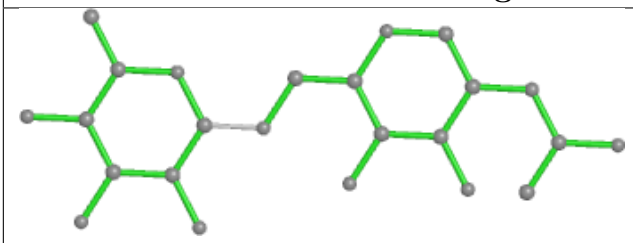
There are no ring outliers.

5 monomers are involved in 10 short contacts:

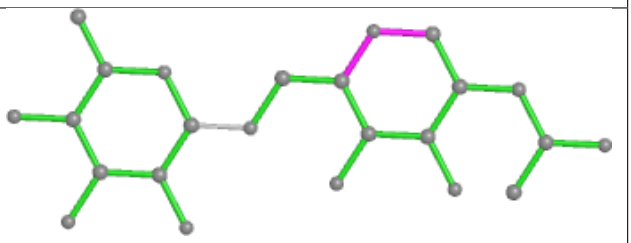
Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	D	1	NAG	4	0
5	E	1	NAG	4	0
5	E	2	FUC	2	0
5	B	1	NAG	2	0
5	B	2	FUC	1	0

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

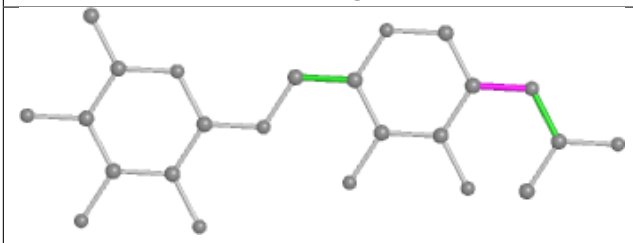
Oligosaccharide Chain A



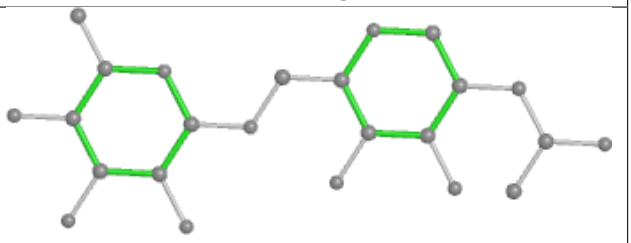
Bond lengths



Bond angles

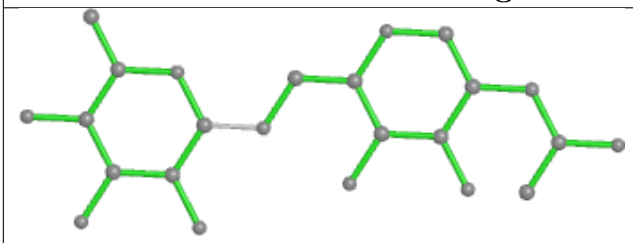


Torsions

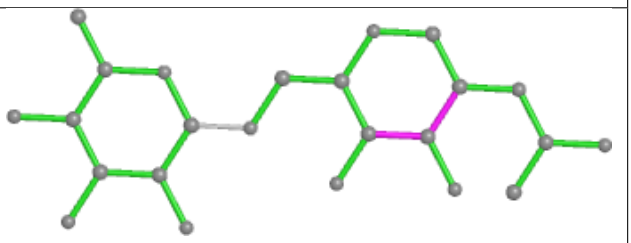


Rings

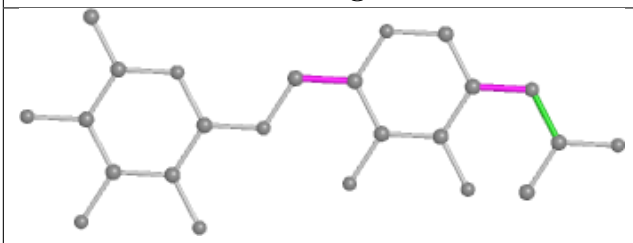
Oligosaccharide Chain B



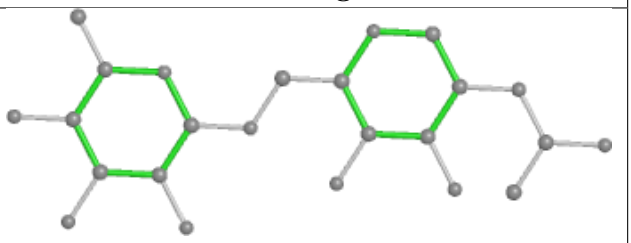
Bond lengths



Bond angles

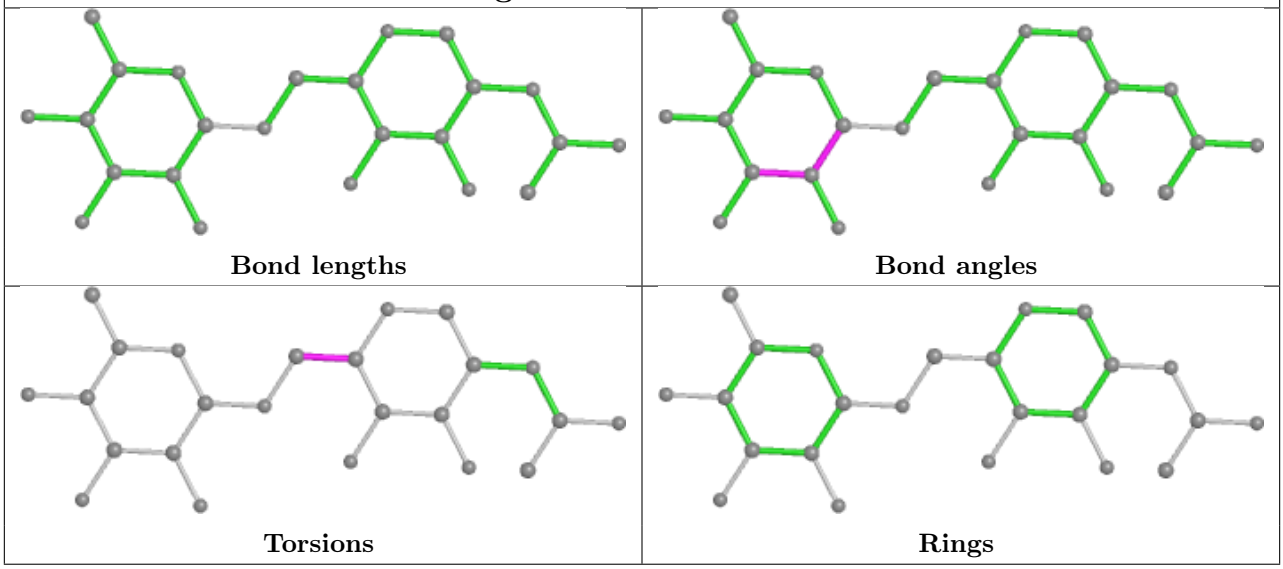


Torsions

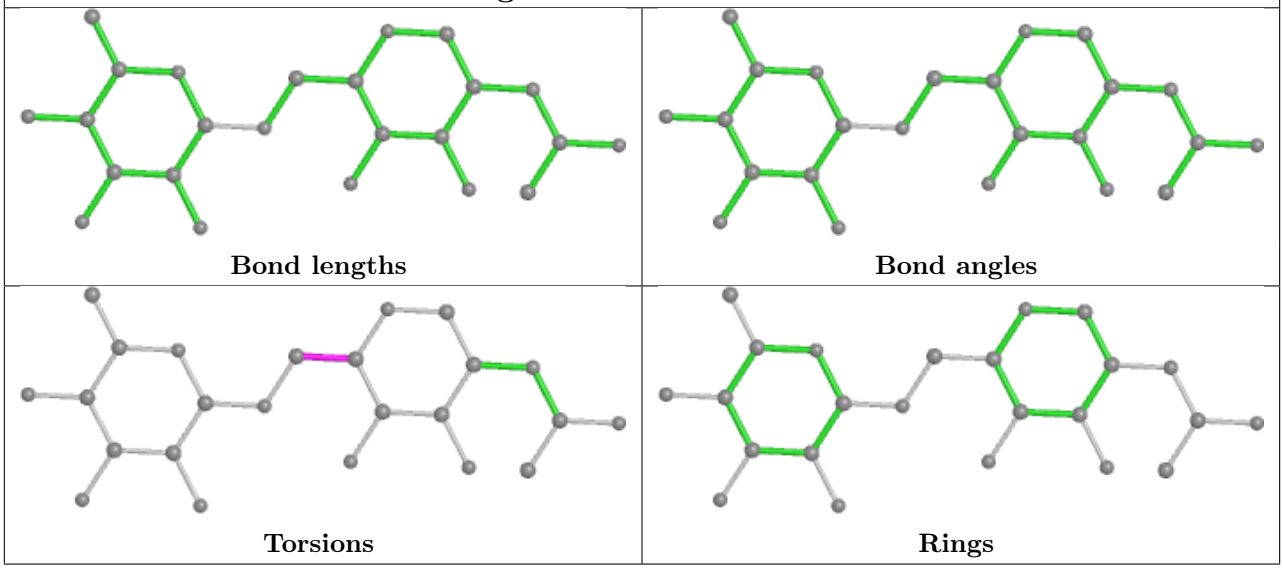


Rings

Oligosaccharide Chain D



Oligosaccharide Chain E



5.6 Ligand geometry (i)

7 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	$\#\|Z\| > 2$
6	NAG	G	762	1	14,14,15	0.44	0	17,19,21	0.74	0
6	NAG	G	839	1	14,14,15	0.49	0	17,19,21	0.83	0
6	NAG	G	886	1	14,14,15	0.52	0	17,19,21	0.85	1 (5%)
6	NAG	G	789	1	14,14,15	0.54	0	17,19,21	0.54	0
6	NAG	G	832	1	14,14,15	0.68	0	17,19,21	0.69	0
6	NAG	G	776	1	14,14,15	0.43	0	17,19,21	0.46	0
6	NAG	G	697	1	14,14,15	0.55	0	17,19,21	0.67	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
6	NAG	G	762	1	-	0/6/23/26	0/1/1/1
6	NAG	G	839	1	-	1/6/23/26	0/1/1/1
6	NAG	G	886	1	-	2/6/23/26	0/1/1/1
6	NAG	G	789	1	-	0/6/23/26	0/1/1/1
6	NAG	G	832	1	-	3/6/23/26	0/1/1/1
6	NAG	G	776	1	-	0/6/23/26	0/1/1/1
6	NAG	G	697	1	-	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})
6	G	886	NAG	C4-C3-C2	-2.33	107.61	111.02

There are no chirality outliers.

5 of 6 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	G	886	NAG	O5-C5-C6-O6
6	G	832	NAG	O5-C5-C6-O6
6	G	832	NAG	C4-C5-C6-O6
6	G	886	NAG	C4-C5-C6-O6
6	G	839	NAG	C3-C2-N2-C7

There are no ring outliers.

4 monomers are involved in 9 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	G	839	NAG	3	0
6	G	789	NAG	3	0
6	G	832	NAG	2	0
6	G	697	NAG	1	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.