wwPDB X-ray Structure Validation Summary Report (i)

Aug 10, 2020 – 06:34 AM BST

PDB ID	:	1AU1
Title	:	HUMAN INTERFERON-BETA CRYSTAL STRUCTURE
Authors	:	Karpusas, M.; Nolte, M.; Lipscomb, W.
Deposited on	:	1997-09-09
Resolution	:	2.20 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:

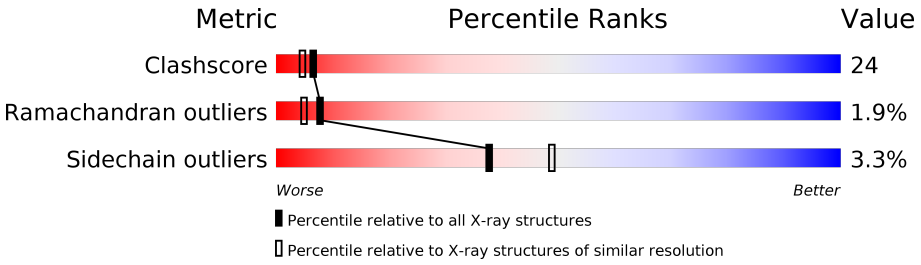
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.13.1

1 Overall quality at a glance

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.20 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	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)

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 on 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	166	73% 22%
1	B	166	70% 26%
2	C	7	86% 14%
3	D	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
3	G6D	D	2	X	-	-	-

2 Entry composition

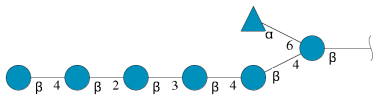
There are 5 unique types of molecules in this entry. The entry contains 3866 atoms, of which 890 are hydrogens and 0 are deuteriums.

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 INTERFERON-BETA.

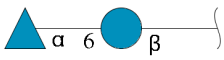
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	166	Total	C	H	N	O	S	0	0	0
			1782	908	369	246	252	7			
1	B	163	Total	C	H	N	O	S	0	0	0
			1728	888	344	243	246	7			

- Molecule 2 is an oligosaccharide called beta-D-glucopyranose-(1-4)-beta-D-glucopyranose-(1-2)-beta-D-glucopyranose-(1-3)-beta-D-glucopyranose-(1-4)-beta-D-glucopyranose-(1-4)-[alpha-D-quinovopyranose-(1-6)]beta-D-glucopyranose.



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	C	7	Total	C	H	O	0	0	0
			87	42	11	34			

- Molecule 3 is an oligosaccharide called alpha-D-quinovopyranose-(1-6)-beta-D-glucopyranose.



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	D	2	Total	C	H	O	0	0	0
			25	12	4	9			

- Molecule 4 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	B	1	Total Zn 1 1	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	29	Total H O 87 58 29	0	0
5	B	52	Total H O 156 104 52	0	0

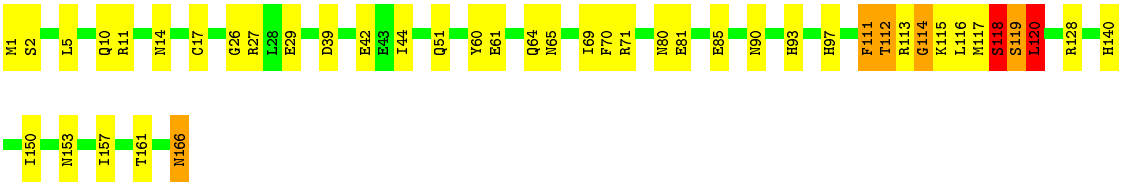
3 Residue-property plots (i)

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.

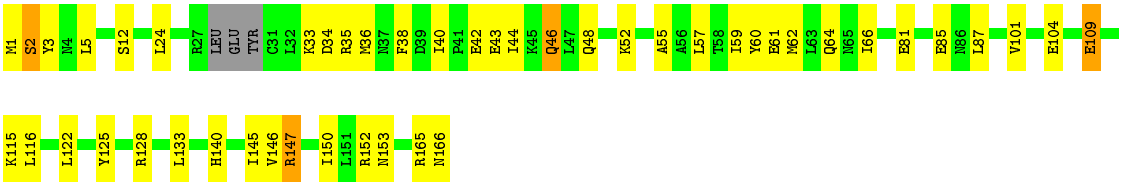
• Molecule 1: INTERFERON-BETA

Chain A:



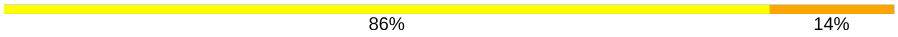
• Molecule 1: INTERFERON-BETA

Chain B:



• Molecule 2: beta-D-glucopyranose-(1-4)-beta-D-glucopyranose-(1-2)-beta-D-glucopyranose-(1-3)-beta-D-glucopyranose-(1-4)-beta-D-glucopyranose-(1-4)-[alpha-D-quinovopyranose-(1-6)]beta-D-glucopyranose

Chain C:



• Molecule 3: alpha-D-quinovopyranose-(1-6)-beta-D-glucopyranose

Chain D:

4 Data and refinement statistics (i)

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

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	55.30Å 65.90Å 121.50Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 2.20	Depositor
% Data completeness (in resolution range)	83.1 (30.00-2.20)	Depositor
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	X-PLOR 3.81	Depositor
R, R_{free}	0.223 , 0.283	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	3866	wwPDB-VP
Average B, all atoms (Å^2)	31.0	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: ZN, BGC, G6D

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	0.66	0/1443	0.82	2/1940 (0.1%)
1	B	0.65	1/1412 (0.1%)	0.76	1/1896 (0.1%)
All	All	0.66	1/2855 (0.0%)	0.79	3/3836 (0.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	2	SER	CA-CB	5.39	1.61	1.52

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	26	GLY	N-CA-C	-5.55	99.22	113.10
1	A	128	ARG	NE-CZ-NH2	-5.20	117.70	120.30
1	B	116	LEU	CA-CB-CG	5.15	127.14	115.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	111	PHE	Sidechain

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	1413	369	1405	91	11
1	B	1384	344	1382	79	11
2	C	76	11	55	1	0
3	D	21	4	10	0	0
4	B	1	0	0	0	0
5	A	29	58	0	6	0
5	B	52	104	0	1	0
All	All	2976	890	2852	136	11

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 24.

The worst 5 of 136 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:116:LEU:CD2	1:B:5:LEU:HD23	1.19	1.60
1:A:115:LYS:HG3	1:A:118:SER:CB	1.27	1.60
1:A:116:LEU:CD1	1:B:2:SER:HB2	1.35	1.51
1:A:115:LYS:CG	1:A:118:SER:HB2	1.39	1.50
1:A:113:ARG:CZ	1:B:104:GLU:OE1	1.72	1.36

The worst 5 of 11 symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:71:ARG:NH2	1:B:104:GLU:CG[3_645]	1.25	0.95
1:A:71:ARG:HH12	1:B:104:GLU:CA[3_645]	0.73	0.87
1:A:71:ARG:NH1	1:B:104:GLU:CA[3_645]	1.46	0.74
1:A:71:ARG:HH22	1:B:104:GLU:CG[3_645]	0.90	0.70
1:A:71:ARG:HH22	1:B:104:GLU:CB[3_645]	1.30	0.30

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	164/166 (99%)	151 (92%)	7 (4%)	6 (4%)	3 1
1	B	159/166 (96%)	154 (97%)	5 (3%)	0	100 100
All	All	323/332 (97%)	305 (94%)	12 (4%)	6 (2%)	8 5

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	112	THR
1	A	114	GLY
1	A	118	SER
1	A	119	SER
1	A	120	LEU

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	154/154 (100%)	150 (97%)	4 (3%)	46 58
1	B	151/154 (98%)	145 (96%)	6 (4%)	31 40
All	All	305/308 (99%)	295 (97%)	10 (3%)	38 49

5 of 10 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	12	SER

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	46	GLN
1	B	109	GLU
1	A	166	ASN
1	B	87	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	94	GLN
1	B	121	HIS
1	B	48	GLN
1	A	51	GLN
1	B	46	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)

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	$\#\|Z\|>2$
2	BGC	C	1	1,2	11,11,12	1.29	2 (18%)	15,15,17	2.23	7 (46%)
2	BGC	C	2	2	11,11,12	1.70	2 (18%)	15,15,17	3.08	7 (46%)
2	BGC	C	3	2	11,11,12	2.34	4 (36%)	15,15,17	2.98	8 (53%)
2	BGC	C	4	2	11,11,12	2.15	4 (36%)	15,15,17	3.60	8 (53%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
2	BGC	C	5	2	11,11,12	1.58	4 (36%)	15,15,17	3.51	5 (33%)
2	BGC	C	6	2	11,11,12	2.36	3 (27%)	15,15,17	2.30	6 (40%)
2	G6D	C	7	2	10,10,11	1.29	1 (10%)	14,14,16	2.31	3 (21%)
3	BGC	D	1	1,3	11,11,12	1.35	1 (9%)	15,15,17	3.45	7 (46%)
3	G6D	D	2	3	10,10,11	1.68	3 (30%)	14,14,16	2.25	4 (28%)

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
2	BGC	C	1	1,2	-	2/2/19/22	0/1/1/1
2	BGC	C	2	2	-	2/2/19/22	0/1/1/1
2	BGC	C	3	2	-	2/2/19/22	0/1/1/1
2	BGC	C	4	2	-	0/2/19/22	0/1/1/1
2	BGC	C	5	2	-	1/2/19/22	0/1/1/1
2	BGC	C	6	2	-	1/2/19/22	0/1/1/1
2	G6D	C	7	2	-	-	0/1/1/1
3	BGC	D	1	1,3	-	2/2/19/22	0/1/1/1
3	G6D	D	2	3	1/1/4/5	-	0/1/1/1

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	6	BGC	C2-C3	4.99	1.59	1.52
2	C	2	BGC	C4-C5	4.26	1.62	1.53
2	C	4	BGC	C2-C3	4.24	1.58	1.52
2	C	3	BGC	C2-C3	4.09	1.58	1.52
2	C	3	BGC	C4-C5	4.04	1.61	1.53

The worst 5 of 55 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	C	5	BGC	C1-O5-C5	10.47	126.38	112.19
2	C	2	BGC	C1-O5-C5	7.85	122.83	112.19
3	D	1	BGC	C1-O5-C5	7.66	122.57	112.19
3	D	1	BGC	O5-C5-C6	-7.59	95.30	107.20
2	C	4	BGC	O5-C1-C2	7.56	122.44	110.77

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
3	D	2	G6D	C1

5 of 10 torsion outliers are listed below:

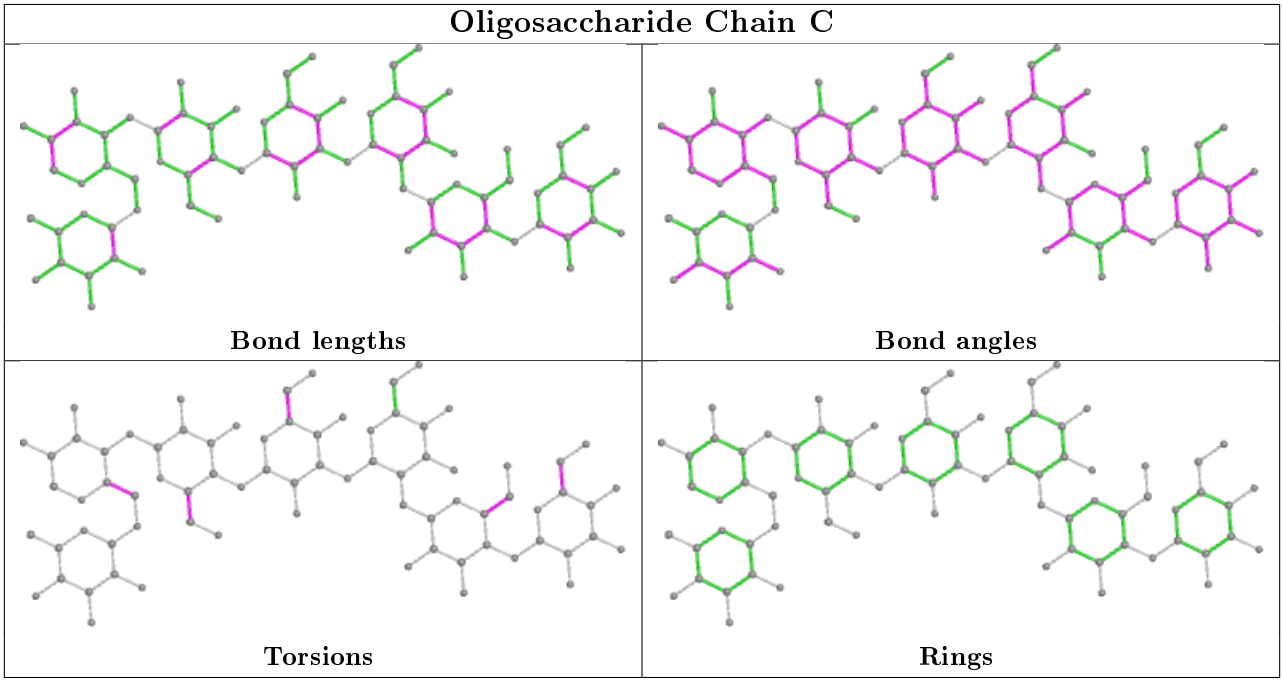
Mol	Chain	Res	Type	Atoms
3	D	1	BGC	O5-C5-C6-O6
2	C	2	BGC	C4-C5-C6-O6
2	C	1	BGC	C4-C5-C6-O6
2	C	2	BGC	O5-C5-C6-O6
2	C	3	BGC	O5-C5-C6-O6

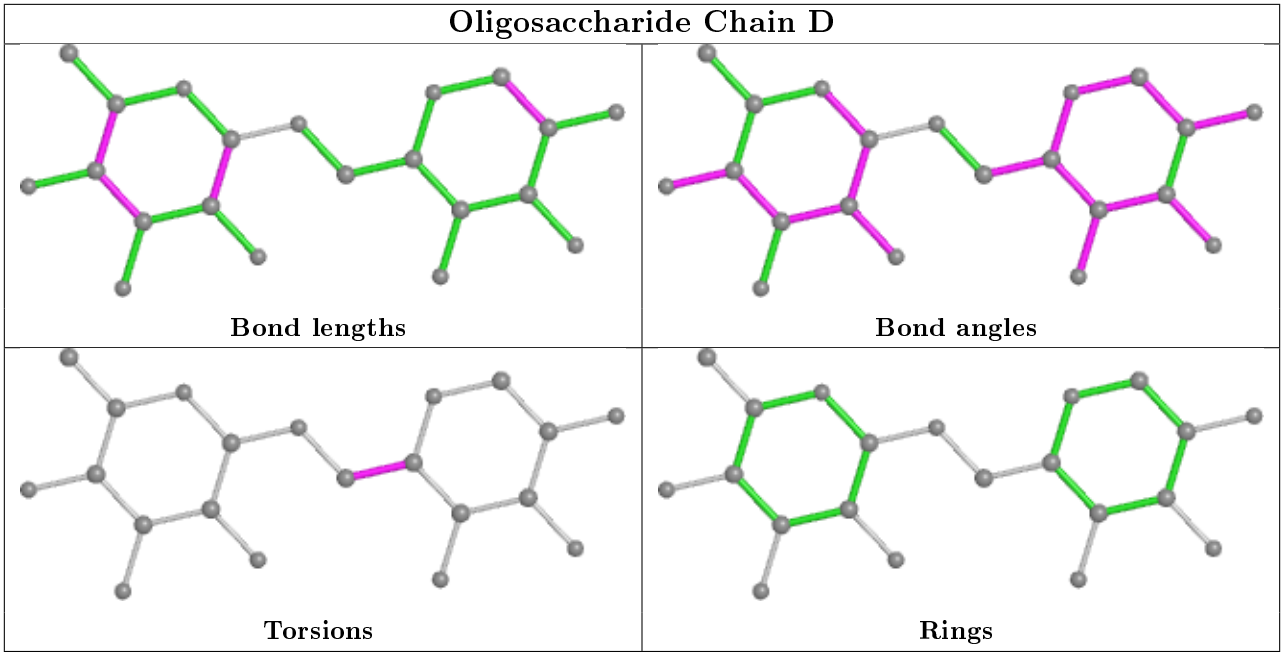
There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	C	1	BGC	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.





5.6 Ligand geometry

Of 1 ligands modelled in this entry, 1 is monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

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.