



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 8, 2020 – 09:15 AM BST

PDB ID : 2RFT  
Title : Crystal structure of influenza B virus hemagglutinin in complex with LSTa receptor analog  
Authors : Wang, Q.; Tian, X.; Chen, X.; Ma, J.  
Deposited on : 2007-10-01  
Resolution : 2.80 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](mailto: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 ⓘ symbol.

---

The following versions of software and data (see [references ⓘ](#)) 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

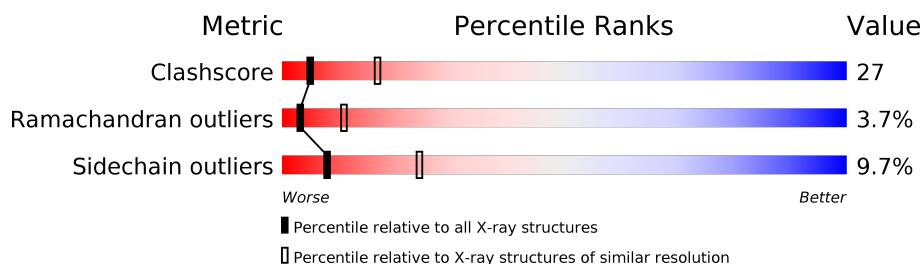
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.80 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	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)

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  $\geq 3$ , 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  $\leq 5\%$ .

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	344	
2	B	176	
3	C	2	
3	D	2	
3	E	2	
3	F	2	
4	G	5	

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	NAG	D	1	-	-	X	-

## 2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 4088 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 Influenza B hemagglutinin (HA).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	342	Total	C	N	O	S	0	0	0
			2583	1626	457	484	16			

- Molecule 2 is a protein called Influenza B hemagglutinin (HA).

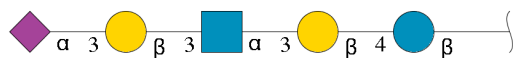
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	169	Total	C	N	O	S	0	0	0
			1281	800	219	256	6			

- 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	D	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	F	2	Total	C	N	O	0	0	0
			28	16	2	10			

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



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

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



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

- Molecule 6 is water.

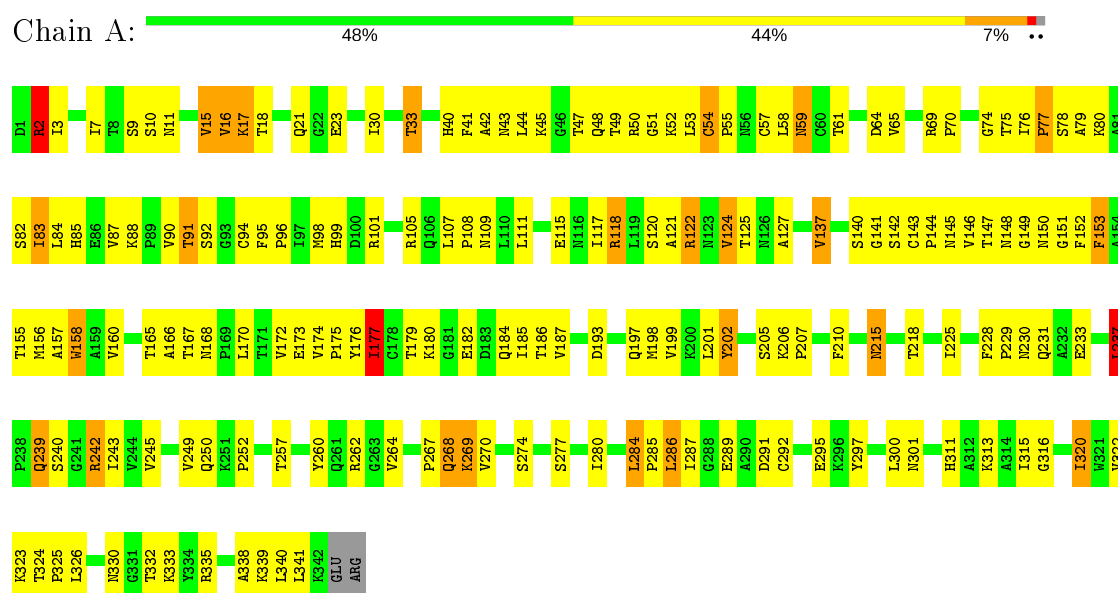
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	2	Total	O	0	0
			2	2		

### 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: Influenza B hemagglutinin (HA)



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

Chain D:  50% 50%

MAG1  
MAG2

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

Chain E:  100%


MAG1  
MAG2

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

Chain F:  100%

MAG1  
MAG2

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

Chain G:  20% 40% 40%

BGC1  
GAL2  
NDG3  
GAL4  
STAG

## 4 Data and refinement statistics

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

Property	Value	Source
Space group	P 3 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	98.33 Å 98.33 Å 135.99 Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	10.00 – 2.80	Depositor
% Data completeness (in resolution range)	90.9 (10.00-2.80)	Depositor
$R_{merge}$	0.05	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.280 , 0.299	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	4088	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	56.0	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, BGC, NAG, NDG

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	0.41	0/2642	0.64	1/3592 (0.0%)
2	B	0.39	0/1300	0.57	0/1752
All	All	0.40	0/3942	0.62	1/5344 (0.0%)

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	3

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	A	2	ARG	NE-CZ-NH1	5.07	122.83	120.30

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	237	LEU	Peptide
1	A	284	LEU	Peptide
1	A	338	ALA	Peptide

## 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	2583	0	2602	197	0
2	B	1281	0	1251	41	1
3	C	28	0	25	0	0
3	D	28	0	25	9	0
3	E	28	0	25	0	0
3	F	28	0	25	4	0
4	G	68	0	57	7	0
5	A	28	0	26	1	0
5	B	14	0	13	0	0
6	A	2	0	0	0	0
All	All	4088	0	4049	222	1

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 27.

The worst 5 of 222 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:260:TYR:HD1	1:A:264:VAL:HG21	1.11	1.08
1:A:260:TYR:CD1	1:A:264:VAL:HG21	1.92	1.04
1:A:137:VAL:HG21	3:D:1:NAG:H81	1.38	0.98
1:A:172:VAL:HG23	1:A:260:TYR:CE2	1.98	0.98
1:A:18:THR:HG23	2:B:105:GLU:HG2	1.48	0.95

All (1) 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)
2:B:18:ILE:CG1	2:B:18:ILE:CD1[6_766]	2.13	0.07

## 5.3 Torsion angles

### 5.3.1 Protein backbone

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	340/344 (99%)	288 (85%)	38 (11%)	14 (4%)	3	9
2	B	167/176 (95%)	148 (89%)	14 (8%)	5 (3%)	4	15
All	All	507/520 (98%)	436 (86%)	52 (10%)	19 (4%)	3	11

5 of 19 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	122	ARG
1	A	153	PHE
1	A	177	ILE
2	B	8	GLY
2	B	163	SER

### 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	A	287/289 (99%)	260 (91%)	27 (9%)	8	26
2	B	136/141 (96%)	122 (90%)	14 (10%)	7	21
All	All	423/430 (98%)	382 (90%)	41 (10%)	8	24

5 of 41 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	237	LEU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	268	GLN
2	B	116	LEU
1	A	239	GLN
1	A	242	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	150	ASN
2	B	75	ASN
1	A	250	GLN
1	A	126	ASN
1	A	294	HIS

### 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 ⓘ

13 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	0.64	0	17,19,21	1.72	3 (17%)
3	NAG	C	2	3	14,14,15	0.57	0	17,19,21	1.08	2 (11%)
3	NAG	D	1	1,3	14,14,15	0.60	0	17,19,21	1.36	2 (11%)
3	NAG	D	2	3	14,14,15	0.51	0	17,19,21	0.78	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	NAG	E	1	1,3	14,14,15	0.41	0	17,19,21	1.78	2 (11%)
3	NAG	E	2	3	14,14,15	0.69	0	17,19,21	1.18	2 (11%)
3	NAG	F	1	1,3	14,14,15	0.66	0	17,19,21	1.09	2 (11%)
3	NAG	F	2	3	14,14,15	0.92	1 (7%)	17,19,21	1.55	6 (35%)
4	BGC	G	1	4	12,12,12	0.52	0	17,17,17	1.30	2 (11%)
4	GAL	G	2	4	11,11,12	0.91	0	15,15,17	2.37	2 (13%)
4	NDG	G	3	4	14,14,15	0.75	0	17,19,21	1.13	1 (5%)
4	GAL	G	4	4	11,11,12	0.69	0	15,15,17	1.05	0
4	SIA	G	5	4	17,20,21	0.35	0	21,28,31	1.04	2 (9%)

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	-	3/6/23/26	0/1/1/1
3	NAG	C	2	3	-	6/6/23/26	0/1/1/1
3	NAG	D	1	1,3	-	3/6/23/26	0/1/1/1
3	NAG	D	2	3	-	4/6/23/26	0/1/1/1
3	NAG	E	1	1,3	-	6/6/23/26	0/1/1/1
3	NAG	E	2	3	-	5/6/23/26	0/1/1/1
3	NAG	F	1	1,3	-	3/6/23/26	0/1/1/1
3	NAG	F	2	3	-	3/6/23/26	0/1/1/1
4	BGC	G	1	4	-	0/2/22/22	0/1/1/1
4	GAL	G	2	4	-	0/2/19/22	0/1/1/1
4	NDG	G	3	4	-	2/6/23/26	0/1/1/1
4	GAL	G	4	4	-	0/2/19/22	0/1/1/1
4	SIA	G	5	4	-	0/14/34/38	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	F	2	NAG	C1-C2	2.12	1.55	1.52

The worst 5 of 26 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
4	G	2	GAL	C1-O5-C5	7.11	121.83	112.19
3	E	1	NAG	C1-O5-C5	5.65	119.84	112.19
4	G	2	GAL	C1-C2-C3	4.99	115.80	109.67
3	C	1	NAG	O5-C1-C2	-3.98	105.00	111.29
4	G	3	NDG	C1-O5-C5	3.66	117.15	112.19

There are no chirality outliers.

5 of 35 torsion outliers are listed below:

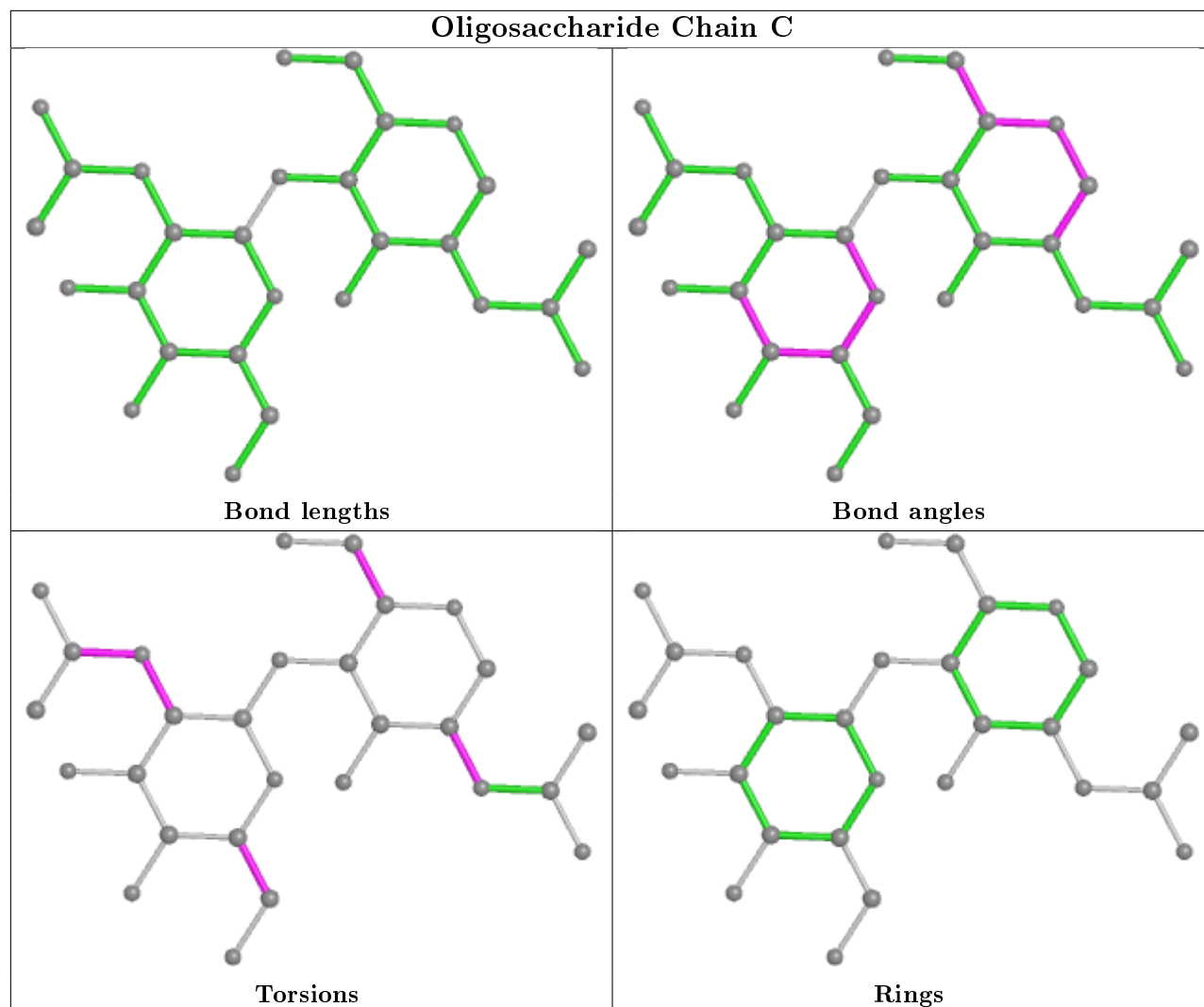
Mol	Chain	Res	Type	Atoms
3	F	1	NAG	C8-C7-N2-C2
3	F	1	NAG	O7-C7-N2-C2
3	D	1	NAG	C3-C2-N2-C7
3	D	1	NAG	C8-C7-N2-C2
3	D	1	NAG	O7-C7-N2-C2

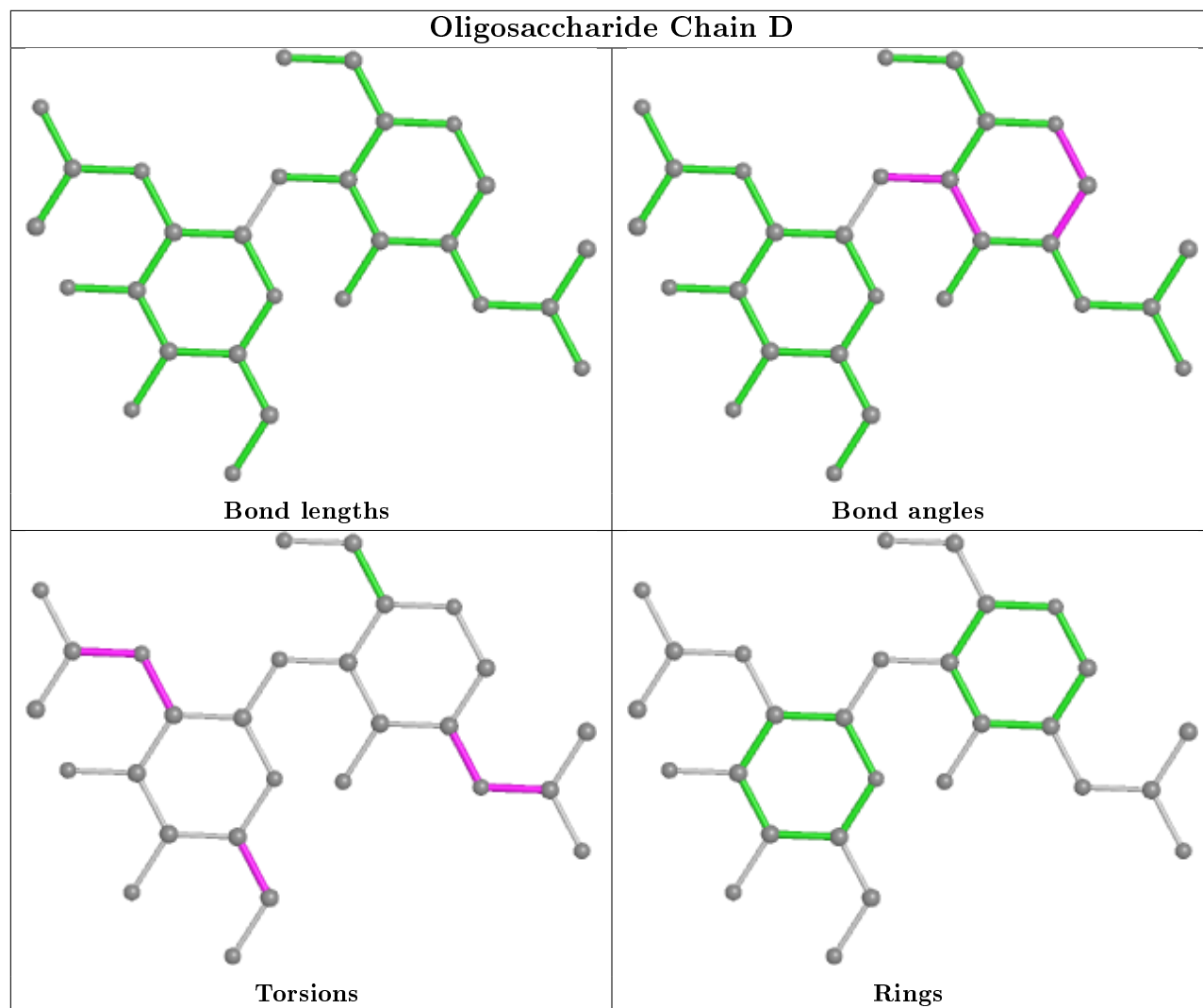
There are no ring outliers.

6 monomers are involved in 20 short contacts:

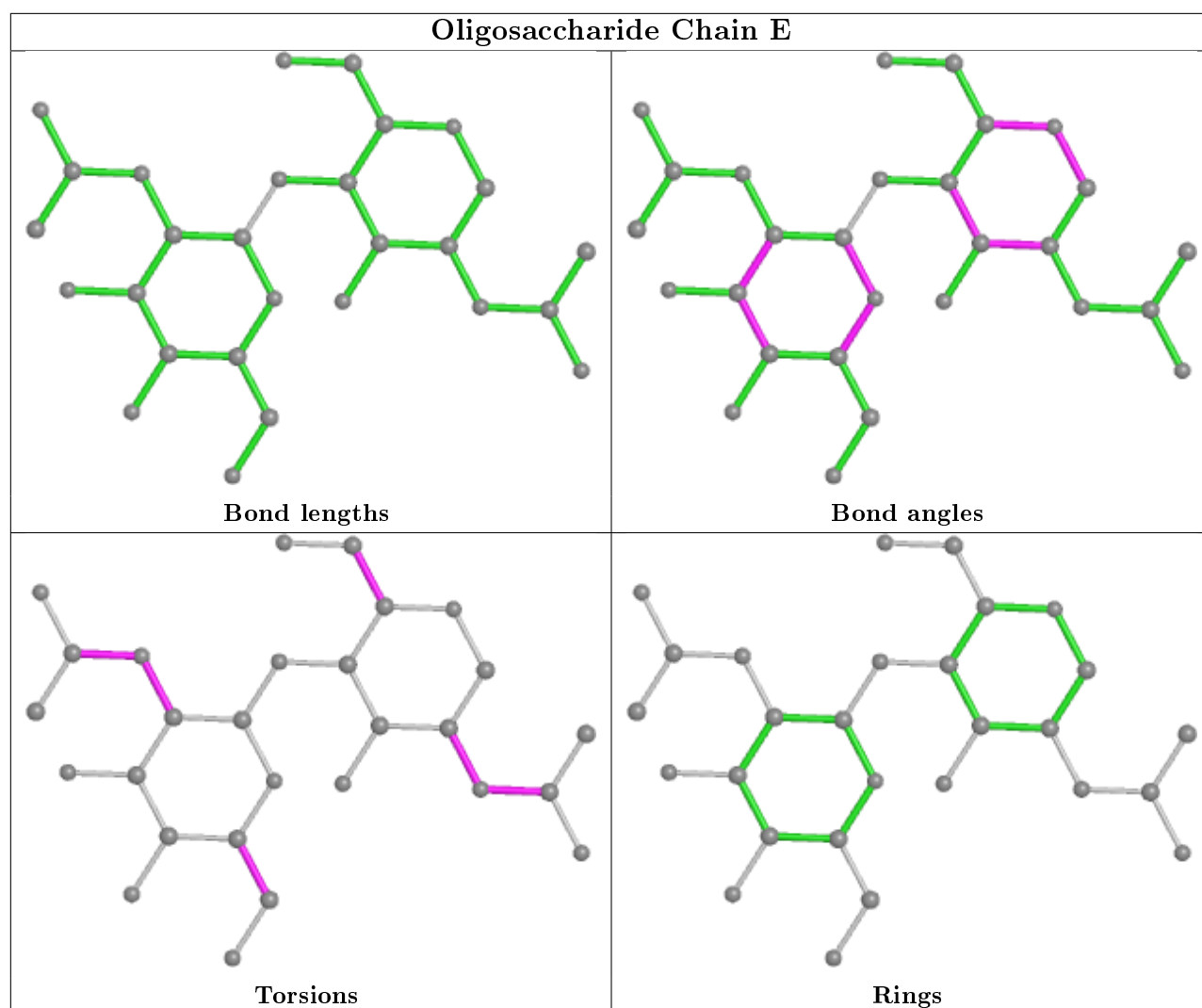
Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	F	1	NAG	4	0
3	D	1	NAG	8	0
4	G	3	NDG	1	0
3	F	2	NAG	2	0
4	G	5	SIA	6	0
3	D	2	NAG	2	0

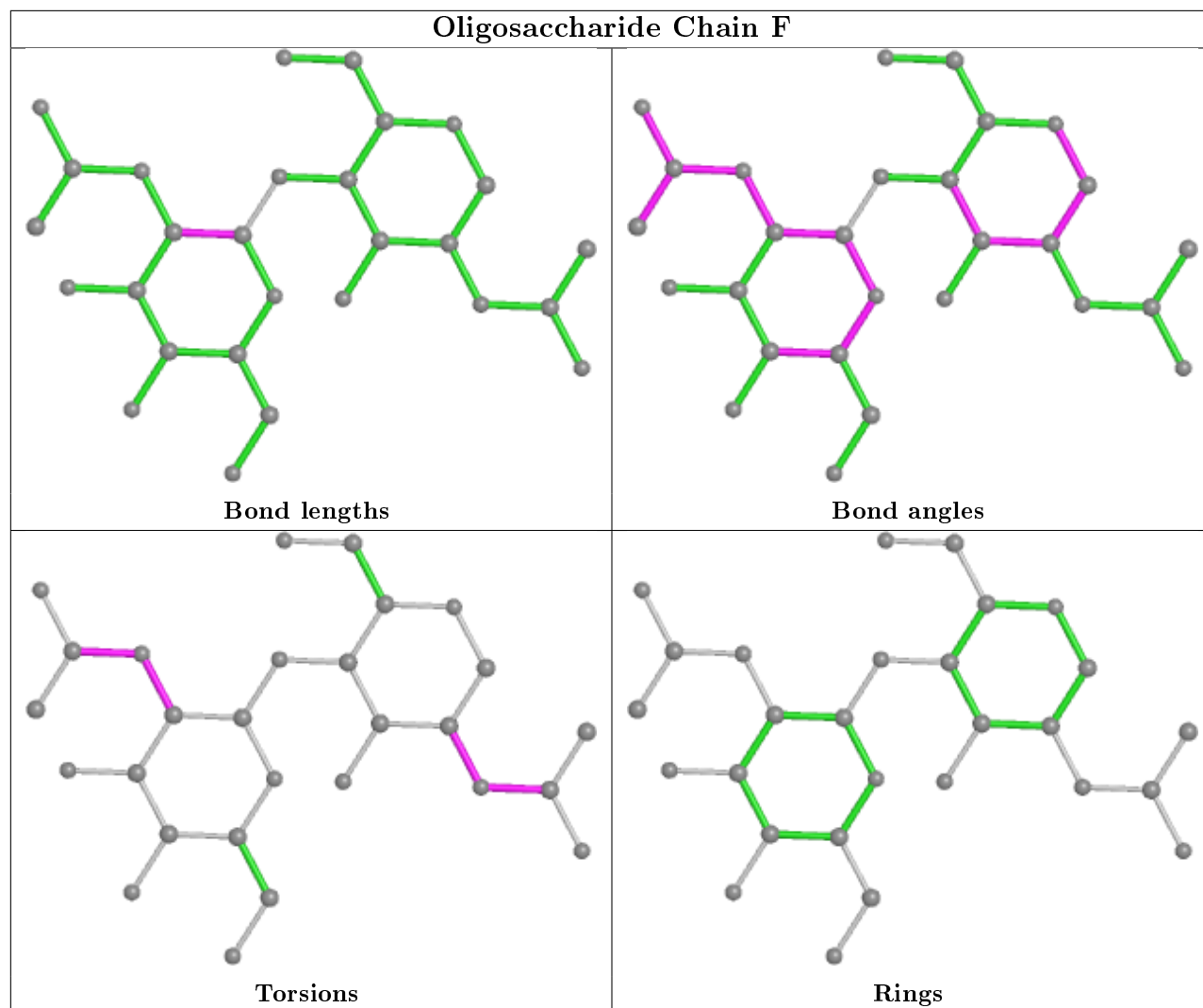
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for oligosaccharide.

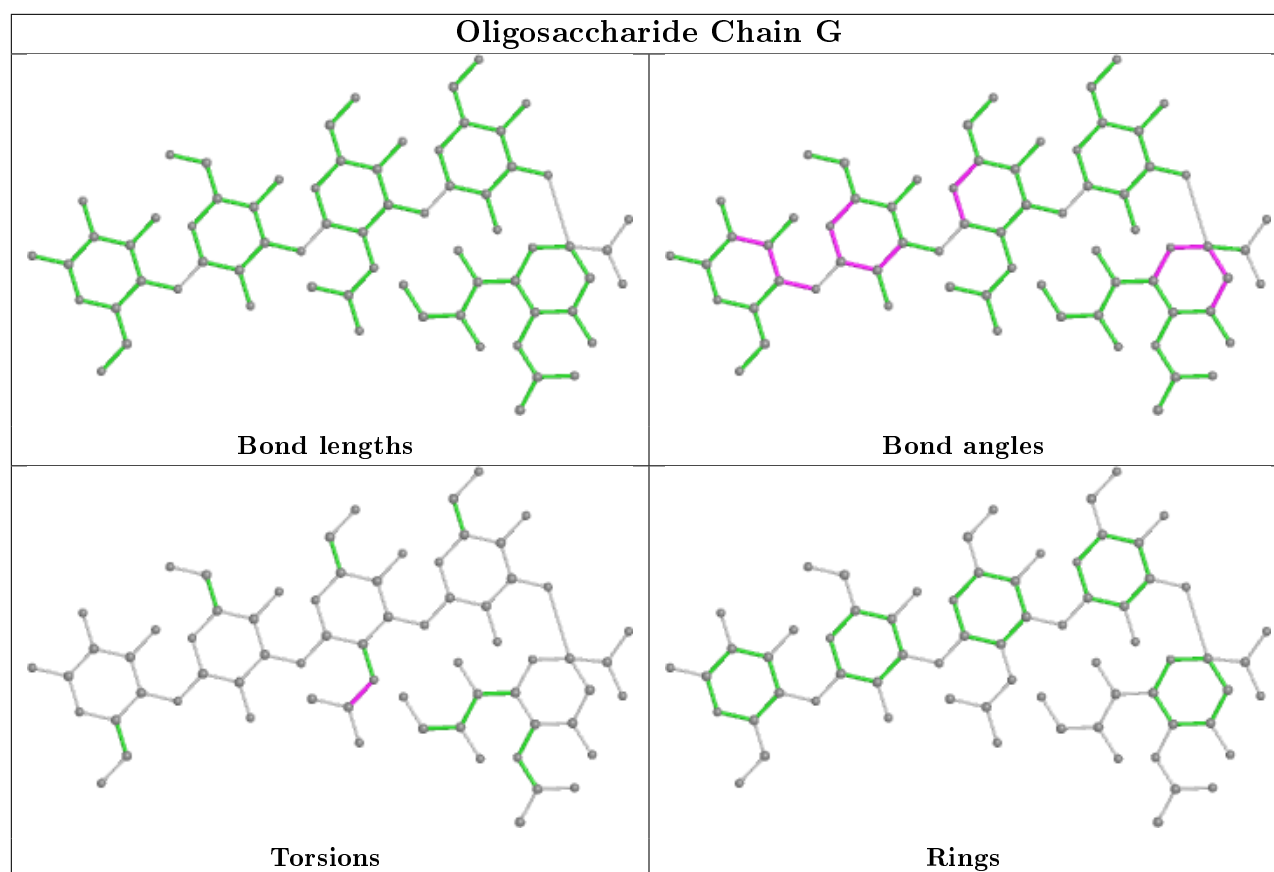












## 5.6 Ligand geometry [i](#)

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	NAG	B	170	2	14,14,15	0.50	0	17,19,21	0.92	1 (5%)
5	NAG	A	347	1	14,14,15	0.62	0	17,19,21	1.45	2 (11%)
5	NAG	A	348	1	14,14,15	0.66	0	17,19,21	1.20	2 (11%)

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	B	170	2	-	2/6/23/26	0/1/1/1
5	NAG	A	347	1	-	6/6/23/26	0/1/1/1
5	NAG	A	348	1	-	4/6/23/26	0/1/1/1

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	347	NAG	C1-O5-C5	4.56	118.37	112.19
5	A	348	NAG	O5-C5-C4	-2.69	104.29	110.83
5	A	347	NAG	O5-C1-C2	2.34	114.98	111.29
5	B	170	NAG	O5-C1-C2	-2.25	107.73	111.29
5	A	348	NAG	O5-C5-C6	2.21	110.67	107.20

There are no chirality outliers.

5 of 12 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	347	NAG	C8-C7-N2-C2
5	A	347	NAG	O7-C7-N2-C2
5	A	347	NAG	O5-C5-C6-O6
5	A	347	NAG	C4-C5-C6-O6
5	A	348	NAG	O5-C5-C6-O6

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	348	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

### 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.