



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 7, 2020 – 11:02 AM BST

PDB ID : 5ZSL  
Title : Crystal structure of monkey TLR7 in complex with GGUUGG  
Authors : Zhang, Z.; Ohto, U.; Shimizu, T.  
Deposited on : 2018-04-28  
Resolution : 2.30 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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) : 1.13  
EDS : 2.13.1  
buster-report : 1.1.7 (2018)  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.13.1

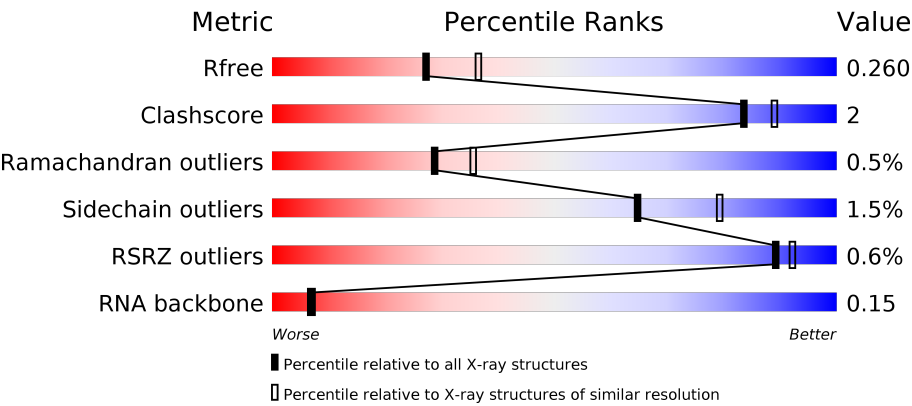
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.30 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.



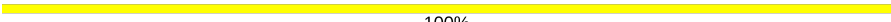
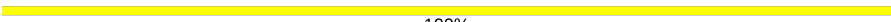
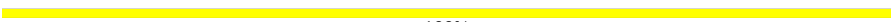
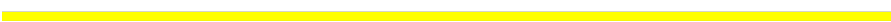
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R <sub>free</sub>	130704	5042 (2.30-2.30)
Clashscore	141614	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)
RSRZ outliers	127900	4938 (2.30-2.30)
RNA backbone	3102	1090 (2.70-1.90)

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  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	823	<div><div>%</div><div><div></div><div>86%</div><div>7%</div><div>6%</div></div></div>
1	B	823	<div><div></div><div>87%</div><div>7%</div><div>6%</div></div>
2	D	6	<div><div></div><div>33%</div><div>33%</div><div>33%</div></div>
2	E	6	<div><div></div><div>50%</div><div>17%</div><div>33%</div></div>

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
3	C	2	 100%
3	F	2	 100%
3	G	2	 100%
3	H	2	 100%

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 13046 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 Toll-like receptor 7.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	B	777	Total	C	N	O	S	0	0	0
			6295	4037	1073	1155	30			
1	A	773	Total	C	N	O	S	0	0	0
			6255	4009	1065	1151	30			

There are 40 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	23	ARG	-	expression tag	UNP B3Y653
B	24	SER	-	expression tag	UNP B3Y653
B	25	PRO	-	expression tag	UNP B3Y653
B	26	TRP	-	expression tag	UNP B3Y653
B	167	GLN	ASN	engineered mutation	UNP B3Y653
B	389	GLN	ASN	engineered mutation	UNP B3Y653
B	440	LEU	SER	see sequence details	UNP B3Y653
B	441	VAL	GLU	see sequence details	UNP B3Y653
B	442	PRO	VAL	see sequence details	UNP B3Y653
B	443	ARG	GLY	see sequence details	UNP B3Y653
B	444	GLY	PHE	see sequence details	UNP B3Y653
B	445	SER	CYS	see sequence details	UNP B3Y653
B	488	GLN	ASN	engineered mutation	UNP B3Y653
B	799	GLN	ASN	engineered mutation	UNP B3Y653
B	840	GLU	-	expression tag	UNP B3Y653
B	841	PHE	-	expression tag	UNP B3Y653
B	842	LEU	-	expression tag	UNP B3Y653
B	843	VAL	-	expression tag	UNP B3Y653
B	844	PRO	-	expression tag	UNP B3Y653
B	845	ARG	-	expression tag	UNP B3Y653
A	23	ARG	-	expression tag	UNP B3Y653
A	24	SER	-	expression tag	UNP B3Y653
A	25	PRO	-	expression tag	UNP B3Y653
A	26	TRP	-	expression tag	UNP B3Y653
A	167	GLN	ASN	engineered mutation	UNP B3Y653

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	389	GLN	ASN	engineered mutation	UNP B3Y653
A	440	LEU	SER	see sequence details	UNP B3Y653
A	441	VAL	GLU	see sequence details	UNP B3Y653
A	442	PRO	VAL	see sequence details	UNP B3Y653
A	443	ARG	GLY	see sequence details	UNP B3Y653
A	444	GLY	PHE	see sequence details	UNP B3Y653
A	445	SER	CYS	see sequence details	UNP B3Y653
A	488	GLN	ASN	engineered mutation	UNP B3Y653
A	799	GLN	ASN	engineered mutation	UNP B3Y653
A	840	GLU	-	expression tag	UNP B3Y653
A	841	PHE	-	expression tag	UNP B3Y653
A	842	LEU	-	expression tag	UNP B3Y653
A	843	VAL	-	expression tag	UNP B3Y653
A	844	PRO	-	expression tag	UNP B3Y653
A	845	ARG	-	expression tag	UNP B3Y653

- Molecule 2 is a RNA chain called RNA (5'-R(P\*UP\*UP\*GP\*A)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	4	Total 67	C 28	N 9	O 26	P 4	0	0	0
2	E	4	Total 67	C 28	N 9	O 26	P 4	0	0	0

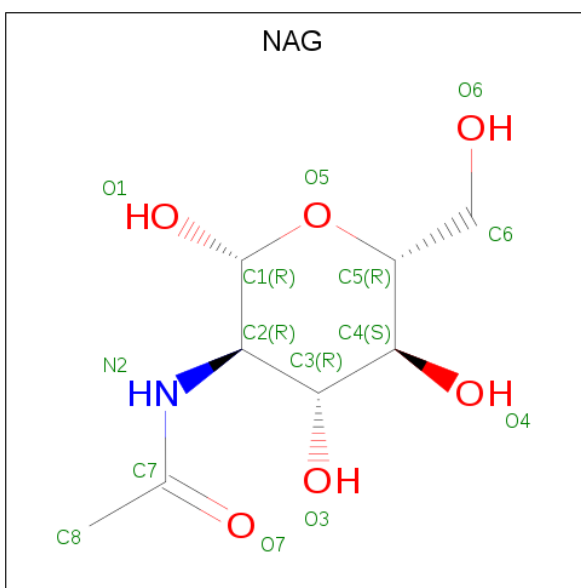
- Molecule 3 is an oligosaccharide called 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	C	2	Total	C	N	O	0	0	0
			28	16	2	10			
3	F	2	Total	C	N	O	0	0	0
			28	16	2	10			
3	G	2	Total	C	N	O	0	0	0
			28	16	2	10			
3	H	2	Total	C	N	O	0	0	0
			28	16	2	10			

- Molecule 4 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula:

C<sub>8</sub>H<sub>15</sub>NO<sub>6</sub>).



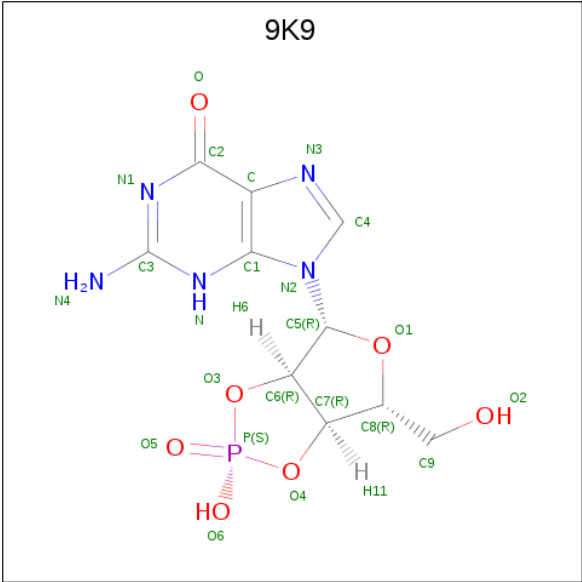
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	B	1	Total	C	N	O	0	0
			14	8	1	5		
4	B	1	Total	C	N	O	0	0
			14	8	1	5		
4	B	1	Total	C	N	O	0	0
			14	8	1	5		
4	B	1	Total	C	N	O	0	0
			14	8	1	5		
4	B	1	Total	C	N	O	0	0
			14	8	1	5		
4	B	1	Total	C	N	O	0	0
			14	8	1	5		
4	A	1	Total	C	N	O	0	0
			14	8	1	5		
4	A	1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

- Molecule 5 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	B	1	Total	O	S	0	0
			5	4	1		
5	B	1	Total	O	S	0	0
			5	4	1		
5	B	1	Total	O	S	0	0
			5	4	1		
5	B	1	Total	O	S	0	0
			5	4	1		
5	A	1	Total	O	S	0	0
			5	4	1		
5	A	1	Total	O	S	0	0
			5	4	1		
5	A	1	Total	O	S	0	0
			5	4	1		

- Molecule 6 is 2-amino-9-[(2S,3aR,4R,6R,6aR)-2-hydroxy-6-(hydroxymethyl)-2-oxotetrahydro-2H-2lambda 5 -furo[3,4-d][1,3,2]dioxaphosphol-4-yl]-3,9-dihydro-6H-purin-6-one (three-letter code: 9K9) (formula: C<sub>10</sub>H<sub>12</sub>N<sub>5</sub>O<sub>7</sub>P).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
6	B	1	Total	C	N	O	P	0	0
			23	10	5	7	1		
6	B	1	Total	C	N	O	P	0	0
			23	10	5	7	1		

- Molecule 7 is water.

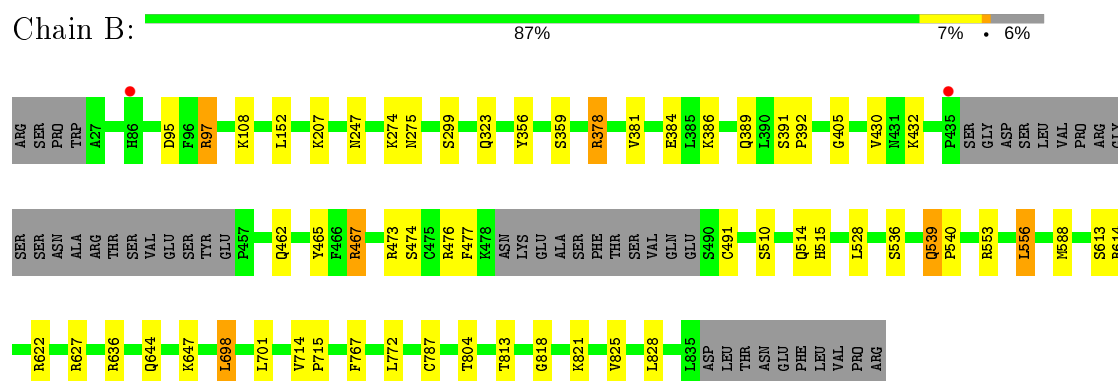
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	B	8	Total	O	0	0
			8	8		
7	A	7	Total	O	0	0
			7	7		



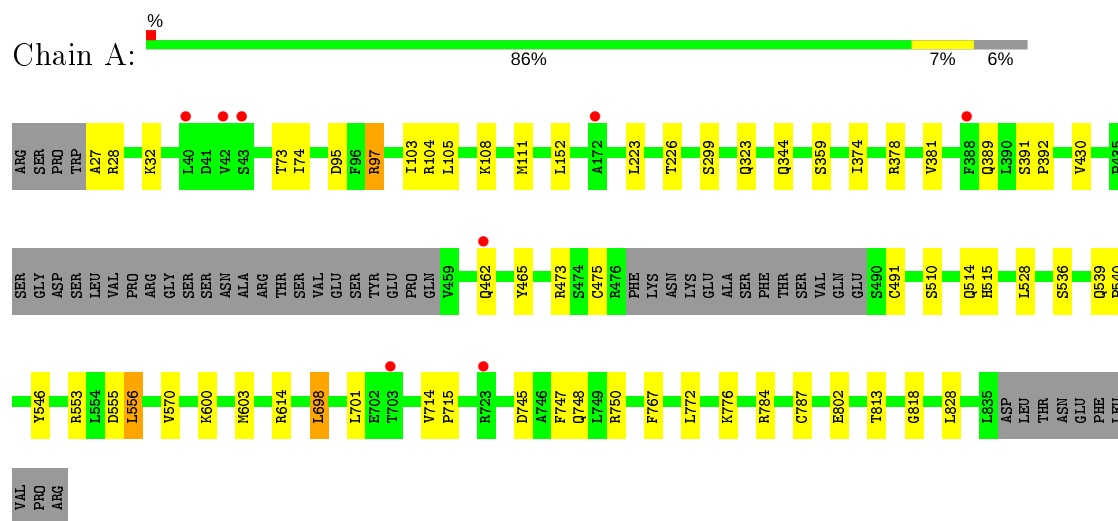
### 3 Residue-property plots

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 and electron density. 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. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). 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.

- Molecule 1: Toll-like receptor 7



- Molecule 1: Toll-like receptor 7

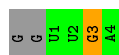


- Molecule 2: RNA (5'-R(P\*UP\*UP\*GP\*A)-3')



- Molecule 2: RNA (5'-R(P\*UP\*UP\*GP\*A)-3')

Chain E:  50% 17% 33%



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

Chain C:  100%



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

Chain F:  100%




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

Chain G:  100%



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

Chain H:  100%



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	99.63Å 139.42Å 150.89Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.36 – 2.30 47.31 – 2.30	Depositor EDS
% Data completeness (in resolution range)	99.9 (47.36-2.30) 99.9 (47.31-2.30)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.46 (at 2.29Å)	Xtriage
Refinement program	REFMAC 5.8.0230	Depositor
R, $R_{free}$	0.217 , 0.258 0.221 , 0.260	Depositor DCC
$R_{free}$ test set	4644 reflections (4.95%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	48.1	Xtriage
Anisotropy	0.045	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 24.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	13046	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	52.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.53% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: 9K9, NAG, SO4

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	0.49	0/6386	0.65	0/8653
1	B	0.50	0/6428	0.67	1/8708 (0.0%)
2	D	0.42	0/73	0.78	0/112
2	E	0.50	0/73	0.75	0/112
All	All	0.49	0/12960	0.66	1/17585 (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	5
1	B	0	6
All	All	0	11

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	627	ARG	NE-CZ-NH2	5.86	123.23	120.30

There are no chirality outliers.

All (11) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	104	ARG	Sidechain
1	A	473	ARG	Sidechain
1	A	750	ARG	Sidechain

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Group
1	A	784	ARG	Sidechain
1	A	97	ARG	Sidechain
1	B	378	ARG	Sidechain
1	B	467	ARG	Sidechain
1	B	473	ARG	Sidechain
1	B	622	ARG	Sidechain
1	B	636	ARG	Sidechain
1	B	97	ARG	Sidechain

## 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	6255	0	6300	35	0
1	B	6295	0	6347	31	0
2	D	67	0	31	1	0
2	E	67	0	31	1	0
3	C	28	0	25	0	0
3	F	28	0	25	0	0
3	G	28	0	25	0	0
3	H	28	0	25	0	0
4	A	56	0	52	1	0
4	B	98	0	91	2	0
5	A	15	0	0	0	0
5	B	20	0	0	0	0
6	B	46	0	0	0	0
7	A	7	0	0	0	0
7	B	8	0	0	0	0
All	All	13046	0	12952	63	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 2.

All (63) 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:546:TYR:CD2	1:A:570:VAL:HG21	1.92	1.03
1:A:546:TYR:CD2	1:A:570:VAL:CG2	2.63	0.81
1:A:546:TYR:CE2	1:A:570:VAL:HG21	2.17	0.77
1:A:748:GLN:N	1:A:748:GLN:OE1	2.21	0.73
1:A:73:THR:HG22	1:A:74:ILE:HG13	1.72	0.70
1:A:546:TYR:CG	1:A:570:VAL:CG2	2.82	0.62
1:A:73:THR:HG21	1:A:475:CYS:O	2.00	0.61
1:A:95:ASP:OD1	1:A:97:ARG:HD3	2.03	0.59
1:B:588:MET:CE	1:B:613:SER:HB3	2.33	0.59
1:A:28:ARG:HH21	1:A:32:LYS:HB2	1.68	0.57
1:B:510:SER:OG	1:B:536:SER:O	2.21	0.57
1:A:73:THR:HG23	1:A:97:ARG:HB2	1.87	0.57
1:A:570:VAL:HG12	1:A:600:LYS:HB3	1.87	0.57
1:A:767:PHE:HB3	1:A:772:LEU:HD11	1.86	0.57
1:B:95:ASP:OD1	1:B:97:ARG:HD3	2.05	0.56
1:A:510:SER:OG	1:A:536:SER:O	2.23	0.56
1:B:644:GLN:HG2	1:B:647:LYS:HD2	1.90	0.54
1:A:714:VAL:HB	1:A:715:PRO:HD2	1.89	0.53
1:B:714:VAL:HB	1:B:715:PRO:HD2	1.92	0.52
1:B:152:LEU:HD23	1:B:152:LEU:C	2.30	0.51
1:B:391:SER:N	1:B:392:PRO:CD	2.74	0.50
1:B:767:PHE:HB3	1:B:772:LEU:HD11	1.93	0.50
1:B:274:LYS:HE2	1:B:275:ASN:HD21	1.77	0.50
1:A:745:ASP:HA	1:A:747:PHE:CE1	2.47	0.50
1:B:299:SER:HA	1:B:323:GLN:O	2.13	0.49
1:A:299:SER:HA	1:A:323:GLN:O	2.12	0.49
1:A:787:CYS:HB3	1:A:828:LEU:HD21	1.94	0.49
1:B:588:MET:HE1	1:B:613:SER:HB3	1.94	0.49
1:A:391:SER:N	1:A:392:PRO:CD	2.76	0.48
1:B:528:LEU:HD21	1:A:553:ARG:HD2	1.95	0.48
1:A:152:LEU:HD23	1:A:152:LEU:C	2.33	0.48
1:B:698:LEU:HB3	1:B:701:LEU:HB2	1.94	0.48
1:A:514:GLN:HE21	1:A:515:HIS:CD2	2.32	0.48
1:A:27:ALA:HB1	1:A:802:GLU:OE1	2.14	0.48
1:B:384:GLU:OE2	1:B:386:LYS:HE3	2.14	0.47
1:B:274:LYS:HG2	1:B:275:ASN:ND2	2.29	0.47
1:B:359:SER:HB2	1:B:389:GLN:NE2	2.30	0.47
1:B:467:ARG:NH1	2:E:3:G:H1'	2.30	0.47
1:B:474:SER:OG	1:B:476:ARG:O	2.33	0.47
1:B:787:CYS:HB3	1:B:828:LEU:HD21	1.98	0.46
1:B:514:GLN:HE21	1:B:515:HIS:CD2	2.33	0.46
1:A:359:SER:HB2	1:A:389:GLN:NE2	2.31	0.46

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:546:TYR:CG	1:A:570:VAL:HG23	2.52	0.45
1:B:491:CYS:HB3	4:B:901:NAG:H81	1.97	0.45
1:B:432:LYS:NZ	1:A:555:ASP:OD2	2.50	0.45
1:B:588:MET:HE3	1:B:613:SER:HB3	1.99	0.44
1:A:344:GLN:HG3	1:A:374:ILE:HB	2.00	0.44
1:A:103:ILE:CD1	1:A:111:MET:SD	3.06	0.44
1:A:698:LEU:HB3	1:A:701:LEU:HB2	2.00	0.44
1:B:553:ARG:HD2	1:A:528:LEU:HD21	2.00	0.44
1:B:539:GLN:N	1:B:540:PRO:CD	2.81	0.43
1:A:491:CYS:HB3	4:A:901:NAG:H81	1.99	0.43
1:A:539:GLN:N	1:A:540:PRO:CD	2.82	0.42
1:B:588:MET:HA	1:B:588:MET:HE3	2.01	0.42
1:B:588:MET:HE2	4:B:902:NAG:C1	2.49	0.42
1:B:378:ARG:HA	1:B:405:GLY:O	2.21	0.41
1:A:378:ARG:HD2	1:A:465:TYR:CE2	2.56	0.41
1:B:378:ARG:HD3	1:B:465:TYR:CE2	2.56	0.41
1:A:105:LEU:HB3	2:D:2:U:H4'	2.03	0.41
1:A:103:ILE:HD11	1:A:111:MET:SD	2.61	0.40
1:B:556:LEU:HD23	1:B:556:LEU:HA	1.95	0.40
1:A:556:LEU:HA	1:A:556:LEU:HD23	1.93	0.40
1:B:247:ASN:OD1	1:B:247:ASN:N	2.49	0.40

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	A	767/823 (93%)	721 (94%)	43 (6%)	3 (0%)	34	42
1	B	771/823 (94%)	726 (94%)	41 (5%)	4 (0%)	29	35
All	All	1538/1646 (93%)	1447 (94%)	84 (6%)	7 (0%)	29	35

All (7) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	818	GLY
1	A	818	GLY
1	B	477	PHE
1	B	381	VAL
1	A	381	VAL
1	B	430	VAL
1	A	430	VAL

### 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	725/774 (94%)	715 (99%)	10 (1%)	67	81
1	B	730/774 (94%)	718 (98%)	12 (2%)	62	78
All	All	1455/1548 (94%)	1433 (98%)	22 (2%)	65	79

All (22) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	108	LYS
1	B	207	LYS
1	B	356	TYR
1	B	462	GLN
1	B	539	GLN
1	B	556	LEU
1	B	614	ARG
1	B	698	LEU
1	B	804	THR
1	B	813	THR
1	B	821	LYS
1	B	825	VAL
1	A	108	LYS
1	A	223	LEU
1	A	226	THR
1	A	462	GLN

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
1	A	556	LEU
1	A	603	MET
1	A	614	ARG
1	A	698	LEU
1	A	776	LYS
1	A	813	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	275	ASN
1	B	323	GLN
1	B	732	ASN
1	B	763	GLN
1	A	763	GLN

### 5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	D	2/6 (33%)	1 (50%)	1 (50%)
2	E	2/6 (33%)	0	1 (50%)
All	All	4/12 (33%)	1 (25%)	2 (50%)

All (1) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
2	D	3	G

All (2) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
2	D	3	G
2	E	3	G

## 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 ⓘ

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$
3	NAG	C	1	1,3	14,14,15	0.53	0	17,19,21	1.07	1 (5%)
3	NAG	C	2	3	14,14,15	0.63	0	17,19,21	1.76	3 (17%)
3	NAG	F	1	1,3	14,14,15	0.59	0	17,19,21	1.23	3 (17%)
3	NAG	F	2	3	14,14,15	0.46	0	17,19,21	1.27	1 (5%)
3	NAG	G	1	1,3	14,14,15	0.42	0	17,19,21	1.24	4 (23%)
3	NAG	G	2	3	14,14,15	0.46	0	17,19,21	1.54	1 (5%)
3	NAG	H	1	1,3	14,14,15	0.68	0	17,19,21	1.39	4 (23%)
3	NAG	H	2	3	14,14,15	0.71	0	17,19,21	1.61	4 (23%)

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	-	0/6/23/26	0/1/1/1
3	NAG	C	2	3	-	2/6/23/26	0/1/1/1
3	NAG	F	1	1,3	-	2/6/23/26	0/1/1/1
3	NAG	F	2	3	-	2/6/23/26	0/1/1/1
3	NAG	G	1	1,3	-	0/6/23/26	0/1/1/1
3	NAG	G	2	3	-	0/6/23/26	0/1/1/1
3	NAG	H	1	1,3	-	0/6/23/26	0/1/1/1
3	NAG	H	2	3	-	4/6/23/26	0/1/1/1

There are no bond length outliers.

All (21) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	G	2	NAG	C1-O5-C5	5.56	119.72	112.19
3	C	2	NAG	C1-O5-C5	5.50	119.65	112.19
3	F	2	NAG	C1-O5-C5	4.13	117.79	112.19
3	H	2	NAG	C2-N2-C7	2.88	127.00	122.90
3	C	2	NAG	O3-C3-C4	-2.87	103.71	110.35
3	F	1	NAG	C1-O5-C5	2.72	115.88	112.19
3	C	1	NAG	O5-C1-C2	-2.67	107.07	111.29
3	H	2	NAG	O4-C4-C5	2.58	115.70	109.30
3	H	2	NAG	C3-C4-C5	-2.55	105.69	110.24
3	F	1	NAG	C6-C5-C4	-2.42	107.34	113.00
3	H	1	NAG	O4-C4-C5	2.40	115.25	109.30
3	H	1	NAG	C1-O5-C5	2.34	115.36	112.19
3	G	1	NAG	O5-C1-C2	-2.34	107.59	111.29
3	H	1	NAG	O4-C4-C3	-2.25	105.15	110.35
3	G	1	NAG	C1-O5-C5	2.22	115.20	112.19
3	G	1	NAG	O5-C5-C6	2.21	110.67	107.20
3	H	1	NAG	C8-C7-N2	-2.18	112.41	116.10
3	F	1	NAG	O5-C5-C6	2.17	110.61	107.20
3	H	2	NAG	O5-C5-C4	-2.07	105.78	110.83
3	G	1	NAG	C1-C2-N2	-2.02	107.03	110.49
3	C	2	NAG	C6-C5-C4	-2.00	108.32	113.00

There are no chirality outliers.

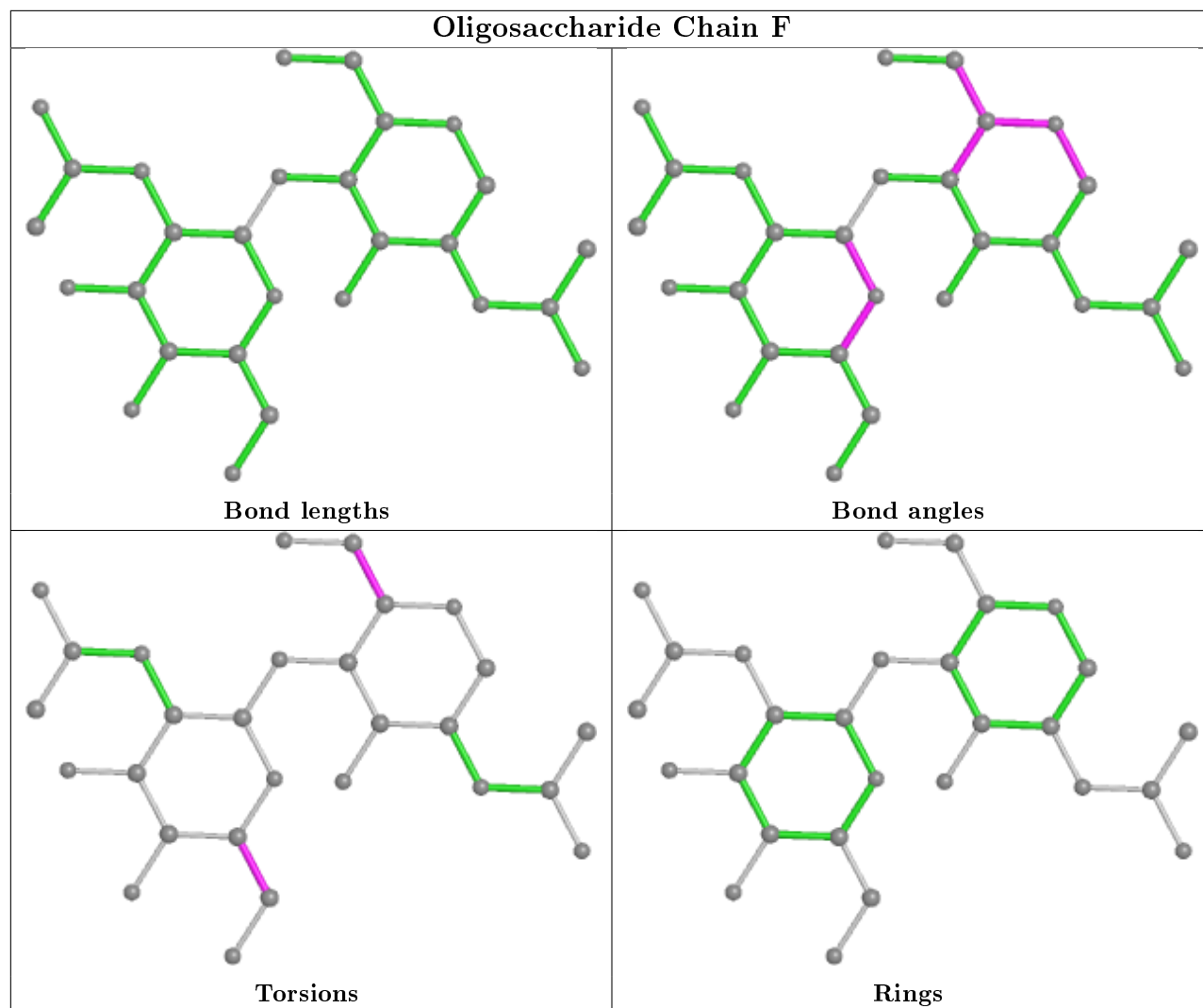
All (10) torsion outliers are listed below:

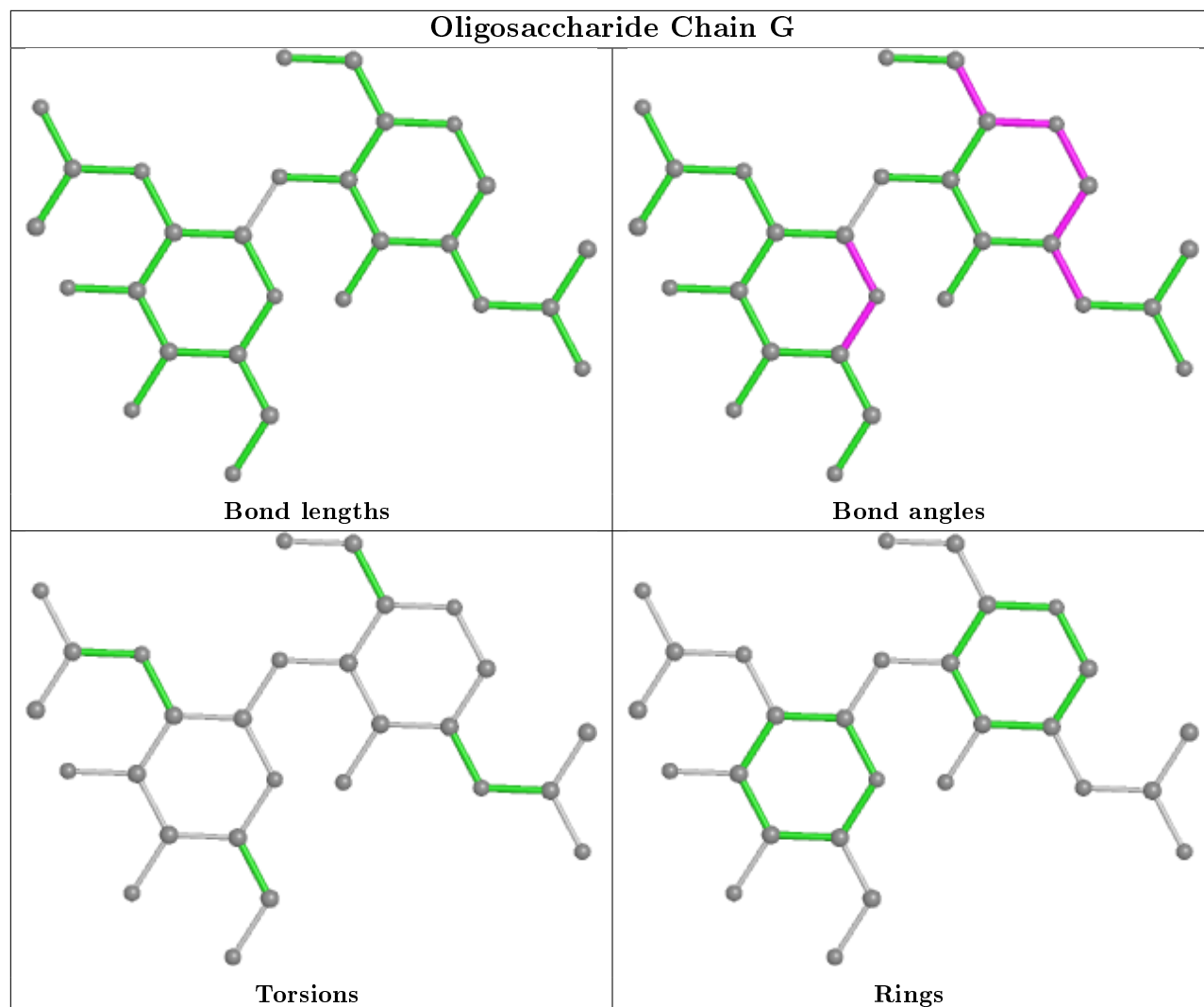
Mol	Chain	Res	Type	Atoms
3	H	2	NAG	C4-C5-C6-O6
3	H	2	NAG	O5-C5-C6-O6
3	F	2	NAG	O5-C5-C6-O6
3	C	2	NAG	O5-C5-C6-O6
3	C	2	NAG	C4-C5-C6-O6
3	F	2	NAG	C4-C5-C6-O6
3	F	1	NAG	C4-C5-C6-O6
3	F	1	NAG	O5-C5-C6-O6
3	H	2	NAG	C1-C2-N2-C7
3	H	2	NAG	C3-C2-N2-C7

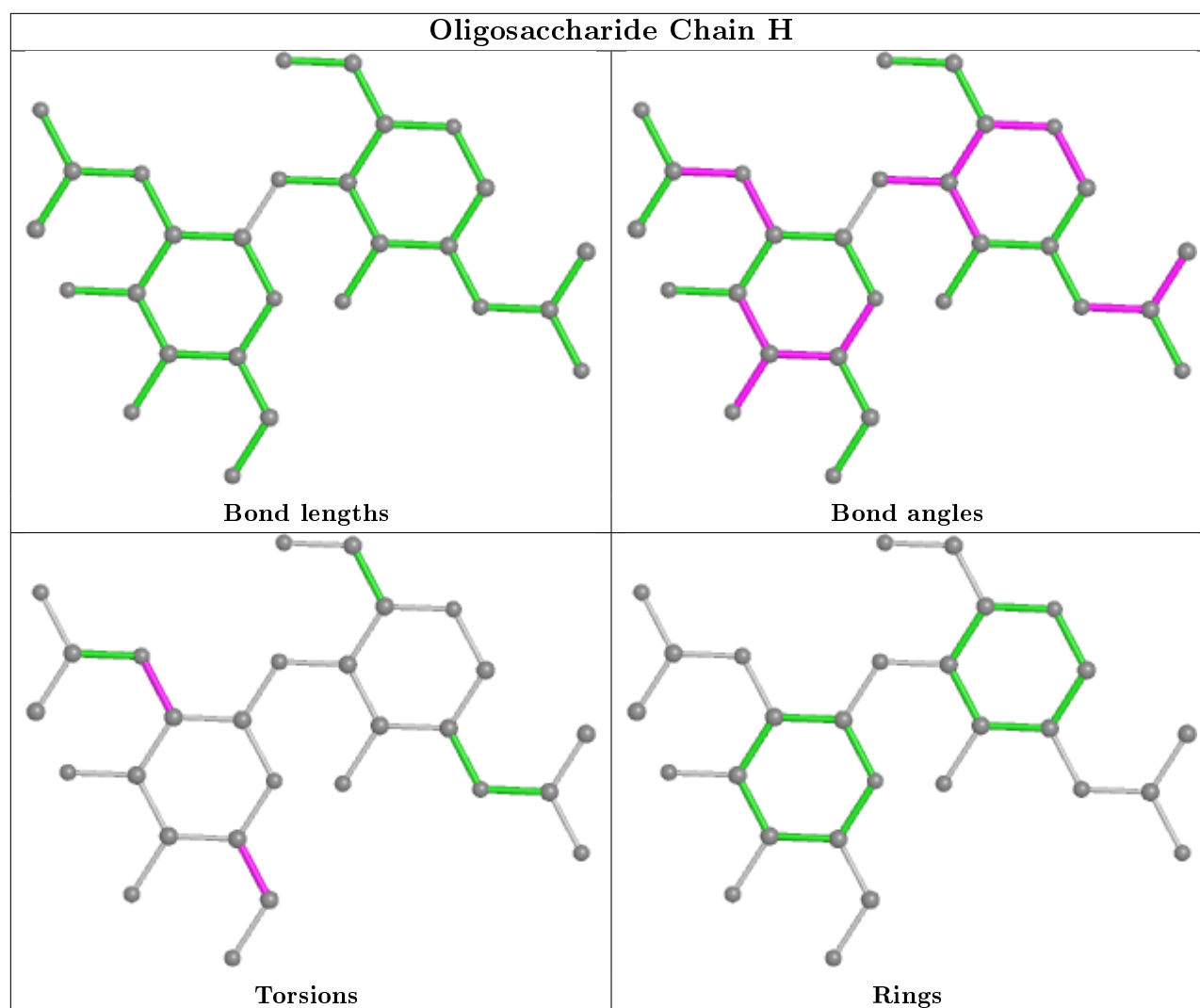
There are no ring outliers.

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.







## 5.6 Ligand geometry [i](#)

20 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	$\# Z  > 2$
4	NAG	A	901	1	14,14,15	0.48	0	17,19,21	1.32	1 (5%)
6	9K9	B	917	-	19,26,26	1.45	4 (21%)	21,41,41	2.88	10 (47%)
5	SO4	B	913	-	4,4,4	0.48	0	6,6,6	0.22	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
5	SO4	B	914	-	4,4,4	0.76	0	6,6,6	0.85	0
5	SO4	A	911	-	4,4,4	0.63	0	6,6,6	0.44	0
4	NAG	B	909	1	14,14,15	0.76	0	17,19,21	1.44	2 (11%)
5	SO4	B	912	-	4,4,4	0.45	0	6,6,6	0.74	0
5	SO4	B	915	-	4,4,4	0.56	0	6,6,6	0.58	0
4	NAG	B	901	1	14,14,15	0.71	0	17,19,21	1.47	4 (23%)
4	NAG	A	907	1	14,14,15	0.52	0	17,19,21	2.32	8 (47%)
4	NAG	B	902	1	14,14,15	0.53	0	17,19,21	1.57	3 (17%)
4	NAG	B	905	1	14,14,15	0.89	0	17,19,21	1.93	3 (17%)
5	SO4	A	909	-	4,4,4	0.43	0	6,6,6	0.78	0
4	NAG	A	902	1	14,14,15	0.52	0	17,19,21	1.43	4 (23%)
5	SO4	A	910	-	4,4,4	0.58	0	6,6,6	0.62	0
4	NAG	A	908	1	14,14,15	0.50	0	17,19,21	1.07	1 (5%)
4	NAG	B	906	1	14,14,15	0.43	0	17,19,21	1.04	1 (5%)
6	9K9	B	916	-	19,26,26	1.59	4 (21%)	21,41,41	2.38	7 (33%)
4	NAG	B	908	1	14,14,15	0.57	0	17,19,21	1.33	2 (11%)
4	NAG	B	907	1	14,14,15	0.65	0	17,19,21	1.33	1 (5%)

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	901	1	-	0/6/23/26	0/1/1/1
6	9K9	B	917	-	-	0/2/32/32	0/4/4/4
4	NAG	B	909	1	-	0/6/23/26	0/1/1/1
4	NAG	B	901	1	-	1/6/23/26	0/1/1/1
4	NAG	A	907	1	-	0/6/23/26	0/1/1/1
4	NAG	B	905	1	-	0/6/23/26	0/1/1/1
4	NAG	B	902	1	-	0/6/23/26	0/1/1/1
4	NAG	A	902	1	-	2/6/23/26	0/1/1/1
6	9K9	B	916	-	-	0/2/32/32	0/4/4/4
4	NAG	A	908	1	-	0/6/23/26	0/1/1/1
4	NAG	B	906	1	-	0/6/23/26	0/1/1/1
4	NAG	B	908	1	-	1/6/23/26	0/1/1/1
4	NAG	B	907	1	-	2/6/23/26	0/1/1/1

All (8) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	B	916	9K9	C2-C	4.07	1.48	1.41
6	B	917	9K9	C2-C	3.71	1.47	1.41
6	B	916	9K9	O1-C5	3.25	1.45	1.41
6	B	916	9K9	C-C1	2.73	1.48	1.40
6	B	917	9K9	C-C1	2.34	1.47	1.40
6	B	916	9K9	O3-C6	-2.27	1.40	1.45
6	B	917	9K9	O1-C5	2.24	1.44	1.41
6	B	917	9K9	O4-C7	-2.20	1.40	1.45

All (47) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	B	905	NAG	C1-O5-C5	6.30	120.73	112.19
6	B	917	9K9	C2-N1-C3	5.68	124.95	115.93
6	B	916	9K9	C2-C-C1	-5.55	115.50	120.80
6	B	917	9K9	C2-C-C1	-5.36	115.68	120.80
4	A	907	NAG	O5-C5-C6	4.89	114.87	107.20
6	B	917	9K9	N-C3-N1	-4.76	120.88	127.22
6	B	917	9K9	C-C2-N1	-4.70	117.01	123.43
6	B	917	9K9	C3-N-C1	4.53	120.53	115.36
4	A	907	NAG	O5-C5-C4	-4.41	100.09	110.83
6	B	916	9K9	C2-N1-C3	4.16	122.55	115.93
4	B	907	NAG	C1-O5-C5	4.07	117.71	112.19
6	B	916	9K9	C3-N-C1	3.90	119.81	115.36
6	B	916	9K9	O4-P-O5	-3.85	105.59	115.76
4	B	902	NAG	C2-N2-C7	-3.67	117.68	122.90
4	A	907	NAG	O5-C1-C2	3.64	117.04	111.29
4	B	909	NAG	C1-O5-C5	3.54	116.99	112.19
4	A	902	NAG	O5-C5-C6	3.41	112.55	107.20
6	B	917	9K9	O6-P-O5	3.37	120.78	109.89
4	B	902	NAG	C4-C3-C2	-3.33	106.14	111.02
4	A	907	NAG	C1-O5-C5	-3.22	107.83	112.19
6	B	916	9K9	N-C3-N1	-3.20	122.95	127.22
6	B	916	9K9	C-C2-N1	-3.20	119.05	123.43
4	A	908	NAG	O5-C1-C2	-3.18	106.27	111.29
4	B	901	NAG	C2-N2-C7	3.02	127.20	122.90
4	A	901	NAG	C1-C2-N2	-2.95	105.44	110.49
4	B	908	NAG	O5-C1-C2	-2.81	106.85	111.29
6	B	917	9K9	O4-P-O5	-2.76	108.48	115.76
4	B	908	NAG	C1-C2-N2	-2.75	105.79	110.49
6	B	917	9K9	C5-N2-C1	-2.47	122.30	126.64
4	B	905	NAG	O5-C1-C2	2.46	115.17	111.29
4	B	901	NAG	C3-C4-C5	-2.36	106.02	110.24

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	B	916	9K9	O4-C7-C6	2.29	109.27	105.08
6	B	917	9K9	O4-C7-C6	2.28	109.26	105.08
4	A	902	NAG	C1-O5-C5	2.28	115.28	112.19
4	B	909	NAG	O3-C3-C2	-2.26	104.80	109.47
4	B	902	NAG	O5-C1-C2	-2.22	107.79	111.29
4	A	907	NAG	C2-N2-C7	2.21	126.05	122.90
4	A	907	NAG	C3-C4-C5	-2.21	106.30	110.24
4	B	901	NAG	O7-C7-C8	-2.19	117.99	122.06
4	A	907	NAG	C6-C5-C4	2.18	118.11	113.00
4	A	902	NAG	C6-C5-C4	-2.16	107.95	113.00
4	B	905	NAG	C4-C3-C2	-2.13	107.90	111.02
4	A	907	NAG	C1-C2-N2	-2.11	106.88	110.49
6	B	917	9K9	C1-C-N3	-2.10	107.21	109.40
4	B	906	NAG	C1-C2-N2	2.10	114.08	110.49
4	A	902	NAG	C1-C2-N2	2.09	114.06	110.49
4	B	901	NAG	O4-C4-C5	2.00	114.26	109.30

There are no chirality outliers.

All (6) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	B	907	NAG	O5-C5-C6-O6
4	B	907	NAG	C4-C5-C6-O6
4	A	902	NAG	O5-C5-C6-O6
4	B	908	NAG	O5-C5-C6-O6
4	A	902	NAG	C4-C5-C6-O6
4	B	901	NAG	C4-C5-C6-O6

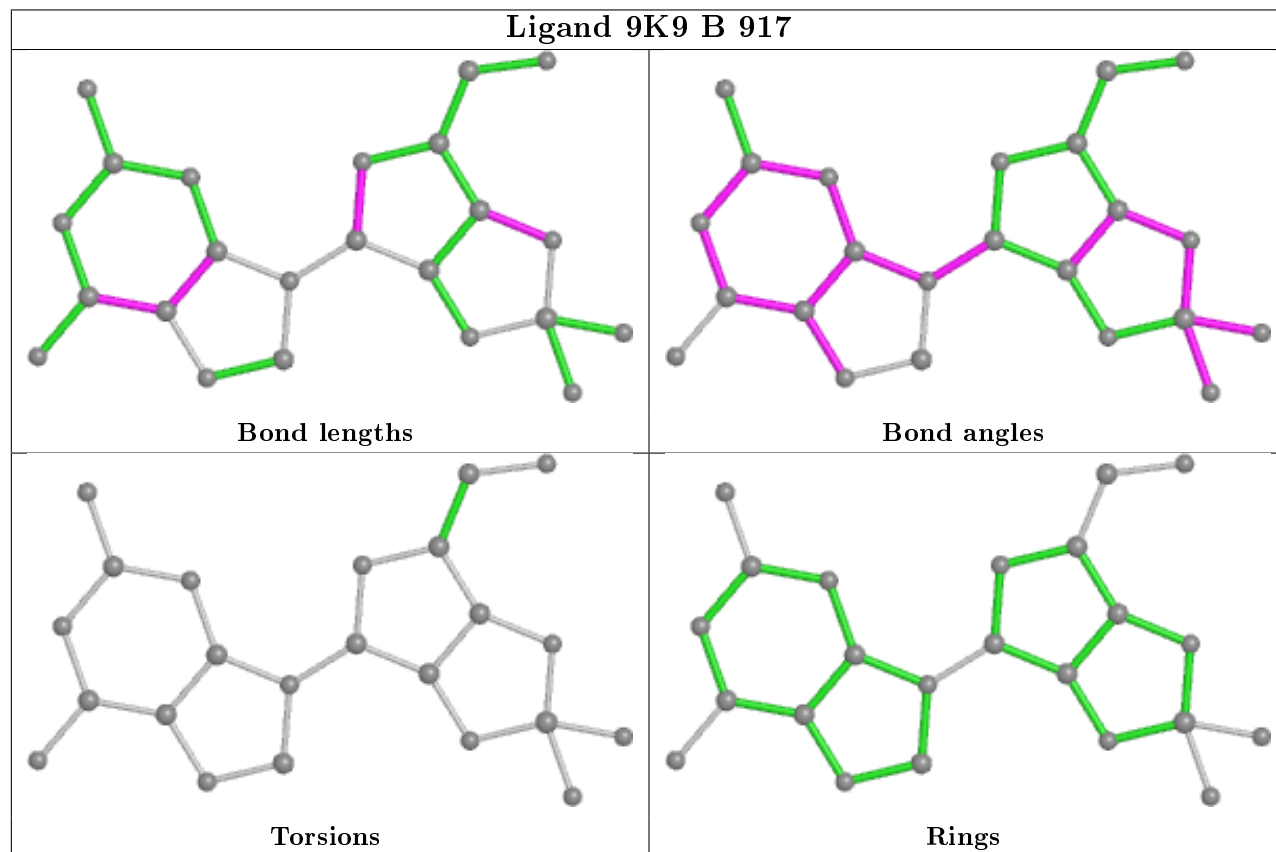
There are no ring outliers.

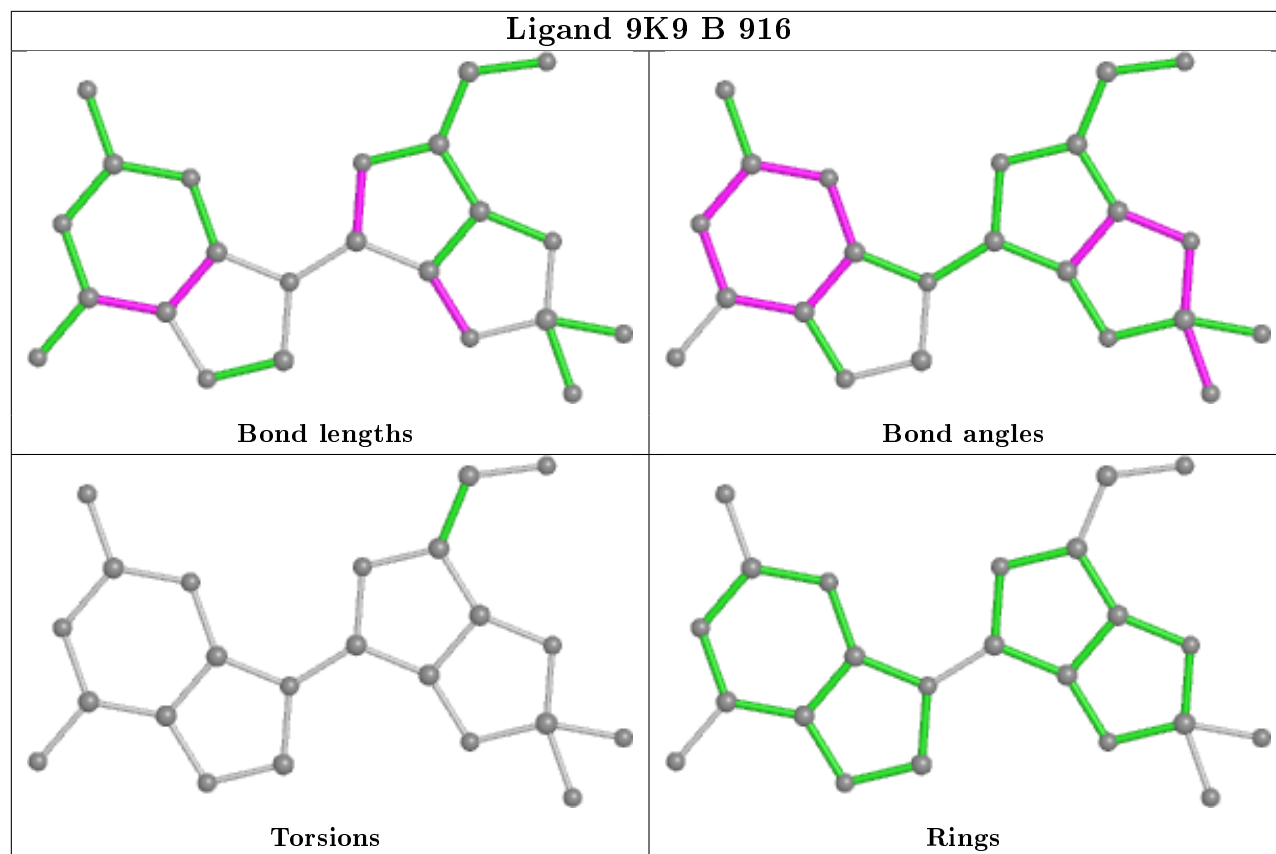
3 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	901	NAG	1	0
4	B	901	NAG	1	0
4	B	902	NAG	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 all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be

highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.





## 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](#)

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	773/823 (93%)	-0.01	8 (1%) 82 86	33, 54, 82, 120	0
1	B	777/823 (94%)	-0.13	2 (0%) 94 96	29, 44, 70, 111	0
2	D	4/6 (66%)	0.30	0 100 100	57, 69, 74, 117	0
2	E	4/6 (66%)	0.45	0 100 100	56, 59, 75, 108	0
All	All	1558/1658 (93%)	-0.07	10 (0%) 89 92	29, 49, 78, 120	0

All (10) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	40	LEU	3.5
1	A	42	VAL	3.5
1	A	43	SER	3.4
1	A	723	ARG	3.1
1	A	172	ALA	2.8
1	A	462	GLN	2.8
1	A	388	PHE	2.4
1	B	435	PRO	2.3
1	B	86	HIS	2.1
1	A	703	THR	2.1

### 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

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

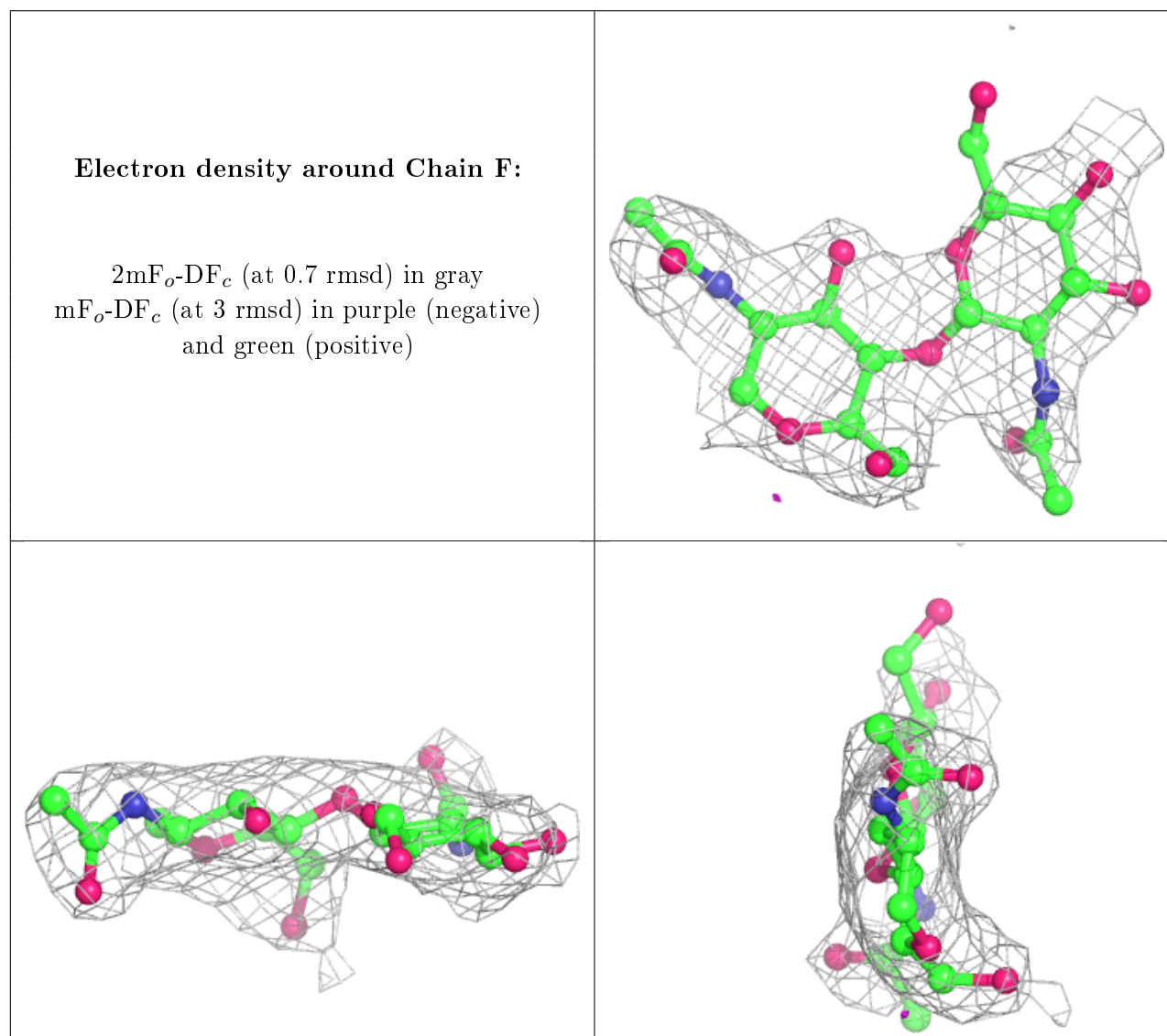
### 6.3 Carbohydrates [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum,

median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q<0.9' lists the number of atoms with occupancy less than 0.9.

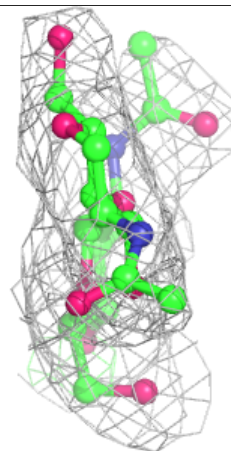
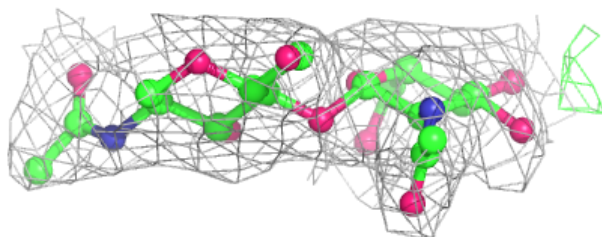
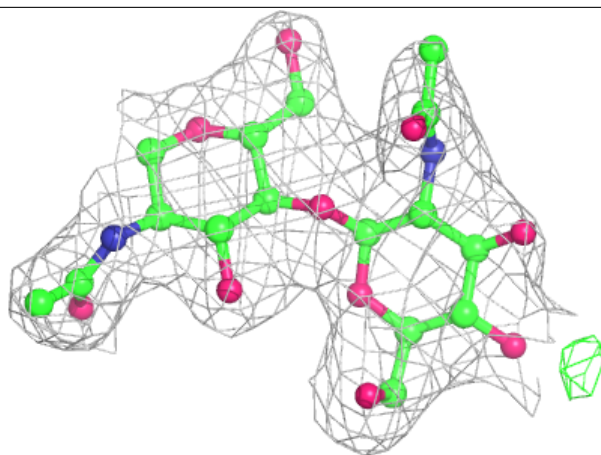
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
3	NAG	H	2	14/15	0.84	0.17	101,107,114,123	0
3	NAG	F	2	14/15	0.90	0.24	79,84,98,100	0
3	NAG	G	2	14/15	0.90	0.09	76,85,90,91	0
3	NAG	H	1	14/15	0.91	0.11	65,74,85,91	0
3	NAG	C	2	14/15	0.92	0.13	61,67,74,76	0
3	NAG	C	1	14/15	0.96	0.11	44,46,53,56	0
3	NAG	F	1	14/15	0.96	0.13	59,64,70,79	0
3	NAG	G	1	14/15	0.97	0.11	42,50,57,66	0

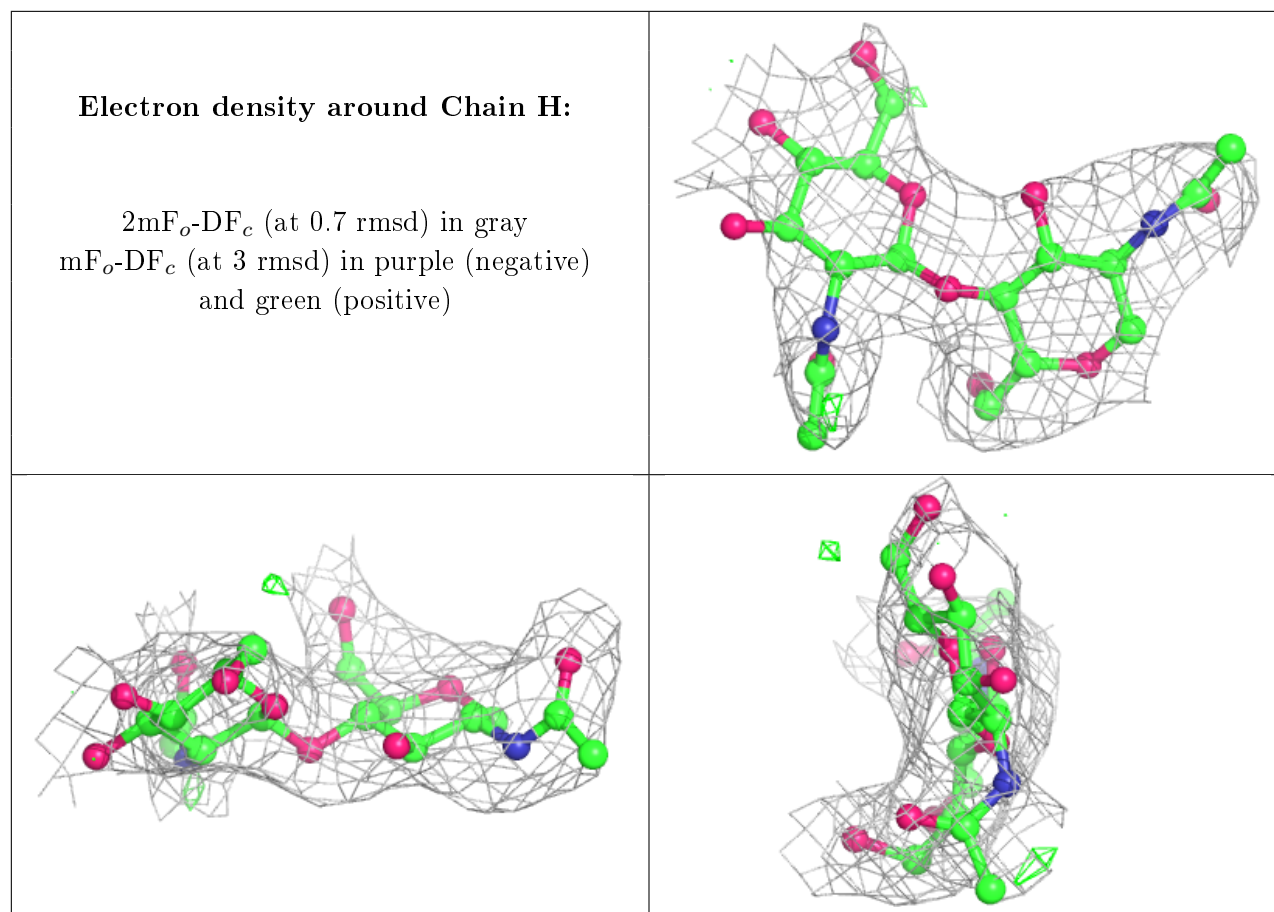
The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.



**Electron density around Chain G:**

$2mF_o - DF_c$  (at 0.7 rmsd) in gray  
 $mF_o - DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)





## 6.4 Ligands ⓘ

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q < 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
5	SO4	B	915	5/5	0.80	0.37	75,87,100,107	0
5	SO4	B	914	5/5	0.82	0.25	78,81,100,101	0
4	NAG	A	907	14/15	0.82	0.20	78,94,104,105	0
4	NAG	B	908	14/15	0.89	0.22	84,96,107,111	0
5	SO4	A	911	5/5	0.90	0.17	75,78,99,102	0
4	NAG	B	907	14/15	0.91	0.13	57,69,80,81	0
5	SO4	A	910	5/5	0.92	0.17	64,78,82,88	0
5	SO4	B	913	5/5	0.92	0.16	86,90,99,99	0
4	NAG	B	909	14/15	0.92	0.20	59,67,70,73	0
4	NAG	A	908	14/15	0.93	0.15	59,70,75,75	0
4	NAG	B	906	14/15	0.93	0.12	56,69,71,73	0
4	NAG	B	905	14/15	0.94	0.10	49,55,65,66	0

*Continued on next page...*

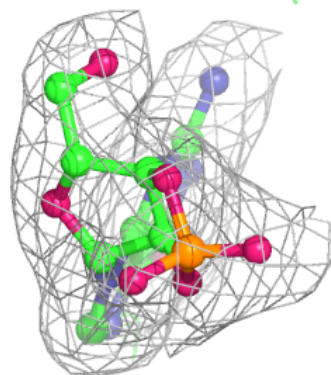
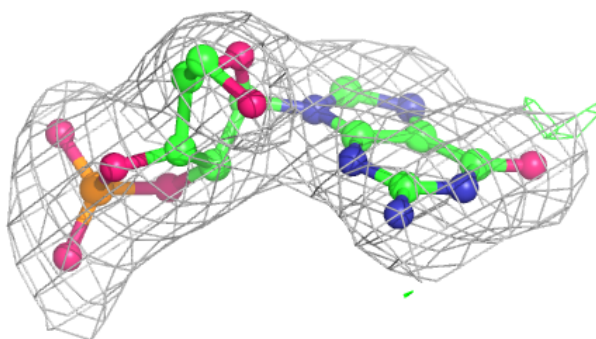
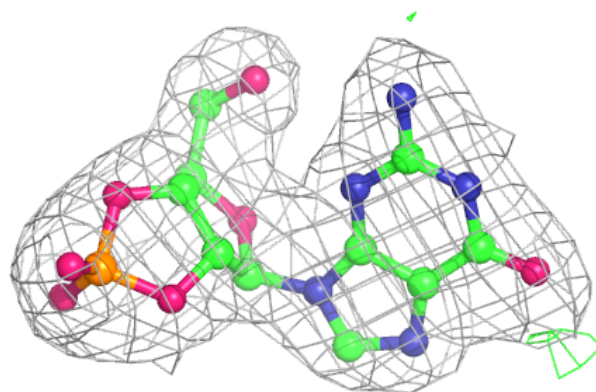
*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
4	NAG	A	902	14/15	0.95	0.11	51,58,64,66	0
5	SO4	B	912	5/5	0.95	0.13	53,57,72,83	0
4	NAG	B	902	14/15	0.95	0.12	53,59,64,70	0
4	NAG	A	901	14/15	0.97	0.11	42,50,56,56	0
5	SO4	A	909	5/5	0.97	0.10	54,69,78,81	0
4	NAG	B	901	14/15	0.97	0.11	36,39,44,45	0
6	9K9	B	917	23/23	0.98	0.14	32,41,49,53	0
6	9K9	B	916	23/23	0.98	0.14	36,41,50,53	0

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

**Electron density around 9K9 B 917:**

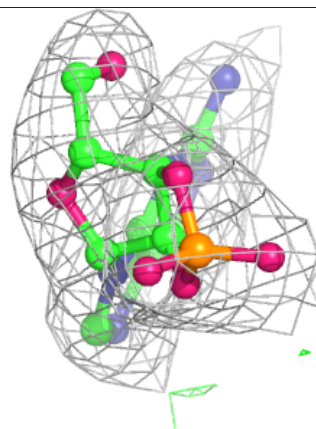
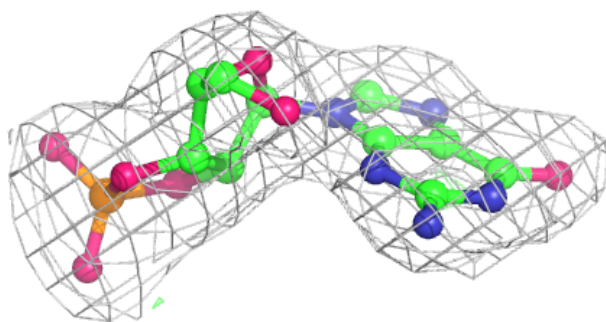
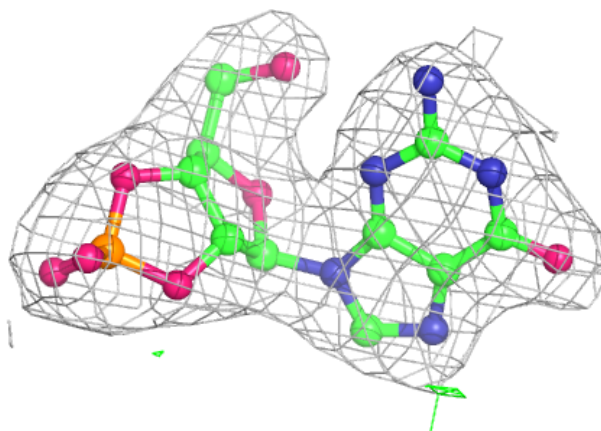
2mF<sub>o</sub>-DF<sub>c</sub> (at 0.7 rmsd) in gray  
mF<sub>o</sub>-DF<sub>c</sub> (at 3 rmsd) in purple (negative)  
and green (positive)





**Electron density around 9K9 B 916:**

$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



## 6.5 Other polymers ⓘ

There are no such residues in this entry.