



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 20, 2020 – 02:31 PM BST

PDB ID : 5ZSM  
Title : Crystal structure of monkey TLR7 in complex with GGUCCC  
Authors : Zhang, Z.; Ohto, U.; Shimizu, T.  
Deposited on : 2018-04-28  
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](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

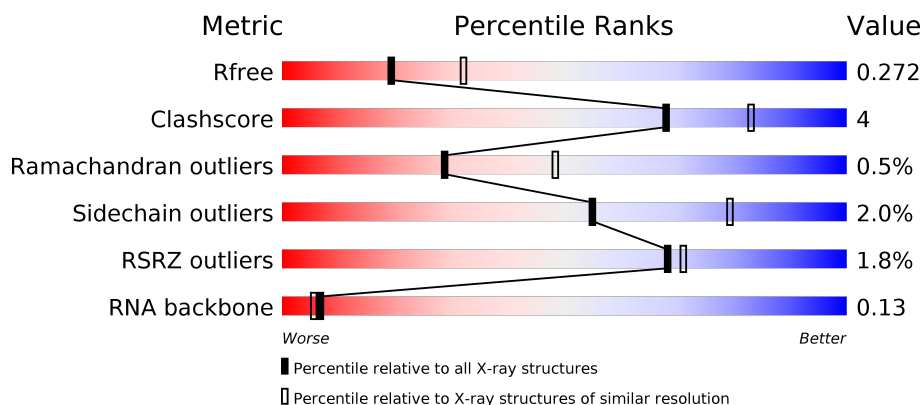
# 1 Overall quality at a glance

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.



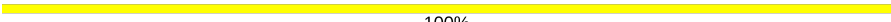
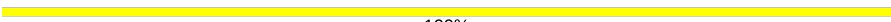
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)
RNA backbone	3102	1008 (2.84-2.16)

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>2%</div> <div>84%</div> <div>10%</div> <div>6%</div> </div>
1	B	823	<div> <div>83%</div> <div>10%</div> <div>6%</div> </div>
2	D	6	<div> <div>50%</div> <div>17%</div> <div>50%</div> <div>17%</div> </div>
2	E	6	<div> <div>83%</div> <div>17%</div> <div>50%</div> <div>17%</div> </div>

*Continued on next page...*

*Continued from previous page...*

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

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 13150 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	776	Total	C	N	O	S	0	2	0
			6308	4043	1078	1157	30			
1	A	774	Total	C	N	O	S	0	2	0
			6291	4030	1077	1154	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(\*GP\*UP\*CP\*CP\*C)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	5	Total	C	N	O	P	0	0	0
			100	46	16	34	4			
2	E	5	Total	C	N	O	P	0	0	0
			100	46	16	34	4			

- 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			

- 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>).



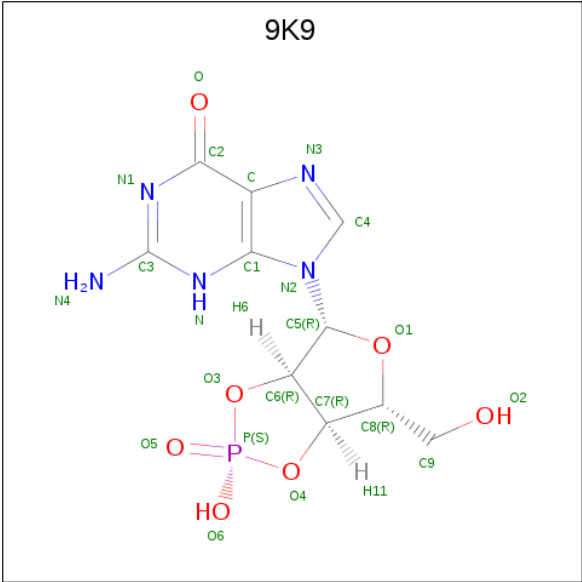
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	B	1	Total	C	N	O	0	0
			14	8	1	5		
4	B	1	Total	C	N	O	0	0
			14	8	1	5		
4	B	1	Total	C	N	O	0	0
			14	8	1	5		
4	B	1	Total	C	N	O	0	0
			14	8	1	5		
4	B	1	Total	C	N	O	0	0
			14	8	1	5		
4	B	1	Total	C	N	O	0	0
			14	8	1	5		
4	B	1	Total	C	N	O	0	0
			14	8	1	5		
4	A	1	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		
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	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		
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	A	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	18	Total	O	0	0
			18	18		
7	A	4	Total	O	0	0
			4	4		



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.

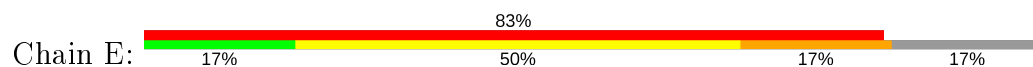
- Chain B:
- 
- 83% 10% 6%
- | Amino Acid | Frequency (%) |
|------------|---------------|
| ALA        | 0.5           |
| SER        | 0.5           |
| PHE        | 0.5           |
| THR        | 0.5           |
| SER        | 0.5           |
| VAL        | 0.5           |
| GLN        | 0.5           |
| GLU        | 0.5           |
| S490       | 0.5           |
| M503       | 0.5           |
| S510       | 0.5           |
| Q514       | 0.5           |
| H515       | 0.5           |
| S525       | 0.5           |
| S536       | 0.5           |
| F538       | 0.5           |
| Q539       | 0.5           |
| P540       | 0.5           |
| S550       | 0.5           |
| R553       | 0.5           |
| L556       | 0.5           |
| S559       | 0.5           |
| ASN        | 0.5           |
| L565       | 0.5           |
| S577       | 0.5           |
| Q581       | 0.5           |
| M588       | 0.5           |
| S613       | 0.5           |
| R622       | 0.5           |
| R636       | 0.5           |
| R641       | 0.5           |
| W692       | 0.5           |
| L698       | 0.5           |
| L701       | 0.5           |
| S707       | 0.5           |
| V714       | 0.5           |
| P715       | 0.5           |
| R723       | 0.5           |
| K731       | 0.5           |
| S755       | 0.5           |
| K788       | 0.5           |
| L779       | 0.5           |
| R784       | 0.5           |
| T804       | 0.5           |
| I805       | 0.5           |
| P806       | 0.5           |
| Y807       | 0.5           |
| T810       | 0.5           |
| D811       | 0.5           |
| H812       | 0.5           |
| T813       | 0.5           |
| P817       | 0.5           |
| G818       | 0.5           |
| A819       | 0.5           |
| V825       | 0.5           |
| L835       | 0.5           |
| ASP        | 0.5           |
| LEU        | 0.5           |
| THR        | 0.5           |
| ASN        | 0.5           |
| GLU        | 0.5           |
| PHE        | 0.5           |
| LEU        | 0.5           |
| VAL        | 0.5           |
| PRO        | 0.5           |
| ARG        | 0.5           |
| R357       | 0.5           |
| S363       | 0.5           |
| K370       | 0.5           |
| R378       | 0.5           |
| V381       | 0.5           |
| F382       | 0.5           |
| S391       | 0.5           |
| F392       | 0.5           |
| G405       | 0.5           |
| V430       | 0.5           |
| M431       | 0.5           |
| K432       | 0.5           |
| I433       | 0.5           |
| S434       | 0.5           |
| SER        | 0.5           |
| GLY        | 0.5           |
| ASP        | 0.5           |
| SER        | 0.5           |
| LEU        | 0.5           |
| VAL        | 0.5           |
| PRO        | 0.5           |
| ARG        | 0.5           |
| GLY        | 0.5           |
| SER        | 0.5           |
| SER        | 0.5           |
| ASN        | 0.5           |
| ALA        | 0.5           |
| ARG        | 0.5           |
| THR        | 0.5           |
| SER        | 0.5           |
| VAL        | 0.5           |
| GLU        | 0.5           |
| SER        | 0.5           |
| TTR        | 0.5           |
| GLU        | 0.5           |
| PRO        | 0.5           |
| Q458       | 0.5           |
| L463       | 0.5           |
| R467       | 0.5           |
| A472       | 0.5           |
| R473       | 0.5           |
| S474       | 0.5           |
| C475       | 0.5           |
| R476       | 0.5           |
| F477       | 0.5           |
| K478       | 0.5           |
| ASN        | 0.5           |
| LYS        | 0.5           |
| GLU        | 0.5           |
| V856       | 0.5           |
| ARG        | 0.5           |
| SER        | 0.5           |
| PRO        | 0.5           |
| TRP        | 0.5           |
| A27        | 0.5           |
| R28        | 0.5           |
| W29        | 0.5           |
| L40        | 0.5           |
| T73        | 0.5           |
| I74        | 0.5           |
| V89        | 0.5           |
| H90        | 0.5           |
| D95        | 0.5           |
| F96        | 0.5           |
| R97        | 0.5           |
| V101       | 0.5           |
| L105       | 0.5           |
| G106       | 0.5           |
| S107       | 0.5           |
| K108       | 0.5           |
| S109       | 0.5           |
| L152       | 0.5           |
| E156       | 0.5           |
| L168       | 0.5           |
| G180       | 0.5           |
| L206       | 0.5           |
| K207       | 0.5           |
| V208       | 0.5           |
| D213       | 0.5           |
| P224       | 0.5           |
| S225       | 0.5           |
| T226       | 0.5           |
| E229       | 0.5           |
| M234       | 0.5           |
| S277       | 0.5           |
| S299       | 0.5           |
| S322       | 0.5           |
| Q323       | 0.5           |
| S348       | 0.5           |
| V356       | 0.5           |

- Chain A: 

- Chain D: 



- Molecule 2: RNA (5'-R(\*GP\*UP\*CP\*CP\*C)-3')



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



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



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	99.66Å 139.17Å 150.79Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.36 – 2.50 47.31 – 2.50	Depositor EDS
% Data completeness (in resolution range)	99.9 (47.36-2.50) 99.9 (47.31-2.50)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.61 (at 2.51Å)	Xtriage
Refinement program	REFMAC 5.8.0230	Depositor
R, $R_{free}$	0.228 , 0.273 0.233 , 0.272	Depositor DCC
$R_{free}$ test set	3596 reflections (4.92%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	49.2	Xtriage
Anisotropy	0.047	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 28.2	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	0.93	EDS
Total number of atoms	13150	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	50.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.70% 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 [i](#)

### 5.1 Standard geometry [i](#)

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	0.56	0/6421	0.75	0/8697
1	B	0.60	0/6440	0.78	0/8723
2	D	0.60	0/110	0.76	0/169
2	E	0.63	0/110	0.76	0/169
All	All	0.58	0/13081	0.76	0/17758

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	5
All	All	0	10

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 10 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	467	ARG	Sidechain
1	B	473	ARG	Sidechain
1	B	641	ARG	Sidechain
1	B	723	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	6291	0	6346	49	0
1	B	6308	0	6358	51	0
2	D	100	0	56	4	0
2	E	100	0	56	3	0
3	C	28	0	25	0	0
3	F	28	0	25	0	0
4	A	70	0	65	0	0
4	B	112	0	104	1	0
5	A	30	0	0	2	0
5	B	15	0	0	1	0
6	A	23	0	0	1	0
6	B	23	0	0	0	0
7	A	4	0	0	0	0
7	B	18	0	0	0	0
All	All	13150	0	13035	99	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 4.

The worst 5 of 99 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:B:476[A]:ARG:HD2	1:B:476[A]:ARG:H	1.30	0.93
1:B:472:ALA:HB3	2:E:3:C:O4'	1.83	0.79
1:A:433:ILE:H	1:A:503:ASN:HD22	1.35	0.74
1:B:431:ASN:HB2	1:B:503:ASN:HD21	1.55	0.71
1:B:433:ILE:H	1:B:503:ASN:HD22	1.37	0.70

There are no symmetry-related clashes.

## 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	769/823 (93%)	708 (92%)	58 (8%)	3 (0%)	34	54
1	B	772/823 (94%)	720 (93%)	47 (6%)	5 (1%)	25	43
All	All	1541/1646 (94%)	1428 (93%)	105 (7%)	8 (0%)	29	48

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	818	GLY
1	A	818	GLY
1	B	381	VAL
1	B	477	PHE
1	A	381	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	729/774 (94%)	715 (98%)	14 (2%)	57	80
1	B	731/774 (94%)	715 (98%)	16 (2%)	52	77
All	All	1460/1548 (94%)	1430 (98%)	30 (2%)	55	78

5 of 30 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	804	THR

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	108	LYS
1	A	810	THR
1	B	825	VAL
1	A	168	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 26 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	800	HIS
1	A	90	HIS
1	A	734	GLN
1	A	66	ASN
1	A	76	HIS

### 5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	D	4/6 (66%)	2 (50%)	0
2	E	5/6 (83%)	2 (40%)	1 (20%)
All	All	9/12 (75%)	4 (44%)	1 (11%)

All (4) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
2	D	3	C
2	D	4	C
2	E	3	C
2	E	4	C

All (1) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
2	E	1	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 ⓘ

4 monosaccharides are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# $ Z  > 2$	Counts	RMSZ	# $ Z  > 2$
3	NAG	C	1	1,3	14,14,15	0.70	0	17,19,21	1.63	4 (23%)
3	NAG	C	2	3	14,14,15	0.71	0	17,19,21	2.40	5 (29%)
3	NAG	F	1	1,3	14,14,15	0.56	0	17,19,21	1.44	3 (17%)
3	NAG	F	2	3	14,14,15	0.49	0	17,19,21	2.75	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
3	NAG	C	1	1,3	-	2/6/23/26	0/1/1/1
3	NAG	C	2	3	-	0/6/23/26	0/1/1/1
3	NAG	F	1	1,3	-	0/6/23/26	0/1/1/1
3	NAG	F	2	3	-	1/6/23/26	0/1/1/1

There are no bond length outliers.

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	F	2	NAG	C1-O5-C5	10.07	125.84	112.19
3	C	2	NAG	C1-O5-C5	7.78	122.73	112.19
3	C	1	NAG	O5-C1-C2	-4.17	104.70	111.29
3	C	2	NAG	C6-C5-C4	-2.79	106.46	113.00
3	F	1	NAG	C1-C2-N2	-2.64	105.97	110.49

There are no chirality outliers.

All (3) torsion outliers are listed below:

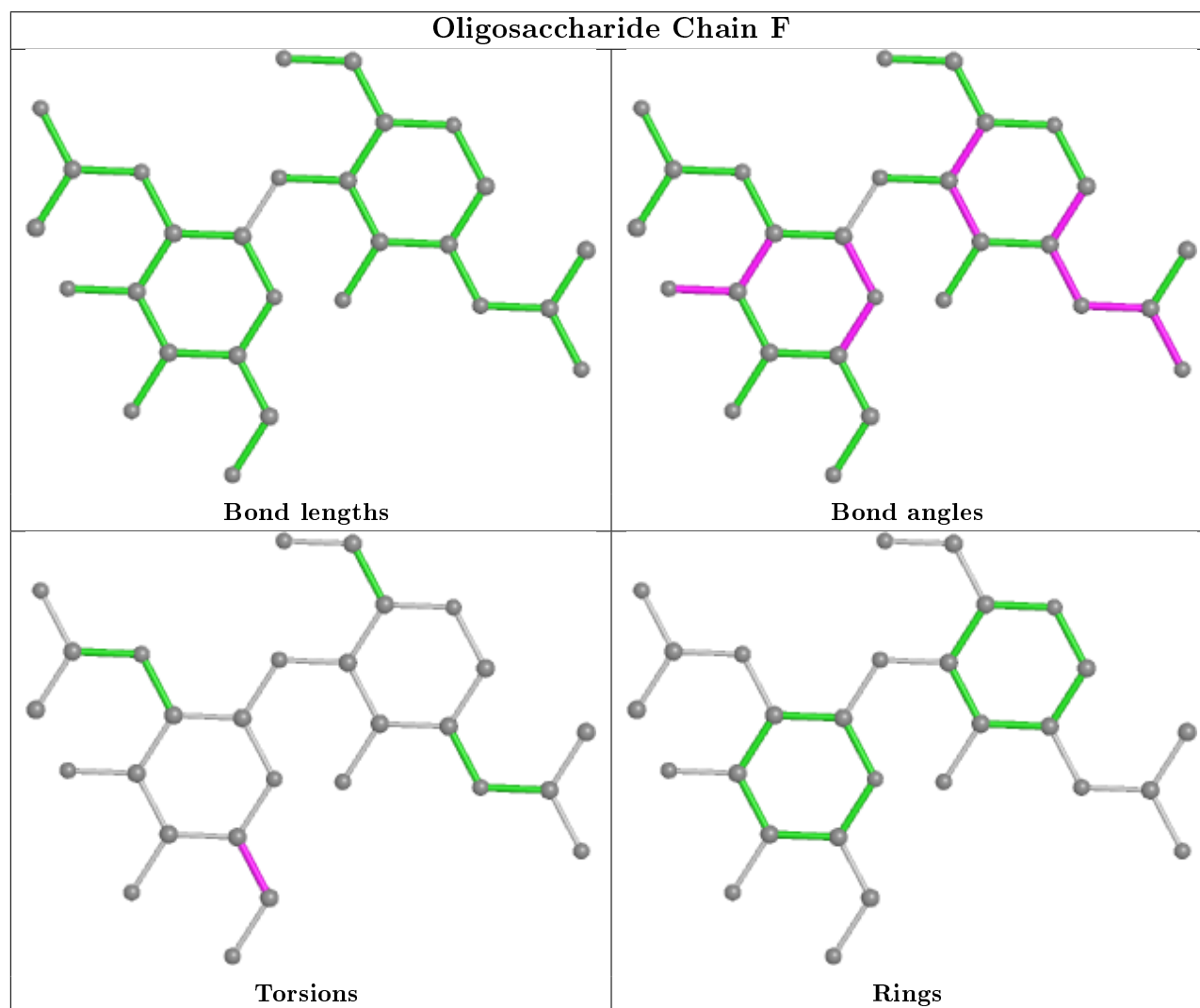


Mol	Chain	Res	Type	Atoms
3	C	1	NAG	O5-C5-C6-O6
3	C	1	NAG	C4-C5-C6-O6
3	F	2	NAG	C4-C5-C6-O6

There are no ring outliers.

No monomer is involved in short contacts.

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



## 5.6 Ligand geometry [i](#)

24 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	B	905	1	14,14,15	0.49	0	17,19,21	1.71	5 (29%)
4	NAG	A	902	1	14,14,15	0.95	1 (7%)	17,19,21	2.10	7 (41%)
5	SO4	B	912	-	4,4,4	0.49	0	6,6,6	0.92	0
4	NAG	B	906	1	14,14,15	0.65	0	17,19,21	2.06	5 (29%)
4	NAG	A	905	1	14,14,15	1.13	2 (14%)	17,19,21	2.10	8 (47%)
6	9K9	A	914	-	19,26,26	1.56	4 (21%)	21,41,41	2.71	11 (52%)
5	SO4	A	909	-	4,4,4	0.30	0	6,6,6	1.21	1 (16%)
5	SO4	A	910	-	4,4,4	0.71	0	6,6,6	0.55	0
4	NAG	A	901	1	14,14,15	0.43	0	17,19,21	1.54	4 (23%)
5	SO4	B	911	-	4,4,4	0.44	0	6,6,6	1.17	0
5	SO4	A	911	-	4,4,4	0.76	0	6,6,6	0.86	0
4	NAG	A	907	1	14,14,15	0.62	0	17,19,21	2.06	3 (17%)
4	NAG	B	901	1	14,14,15	0.62	0	17,19,21	2.04	4 (23%)
5	SO4	A	912	-	4,4,4	0.62	0	6,6,6	0.52	0
5	SO4	A	908	-	4,4,4	0.47	0	6,6,6	1.00	0
4	NAG	B	908	1	14,14,15	0.90	0	17,19,21	2.91	9 (52%)
4	NAG	B	907	1	14,14,15	0.56	0	17,19,21	1.76	4 (23%)
4	NAG	B	909	1	14,14,15	1.03	2 (14%)	17,19,21	2.00	5 (29%)
4	NAG	A	906	1	14,14,15	0.63	0	17,19,21	1.17	2 (11%)
4	NAG	B	902	1	14,14,15	0.86	0	17,19,21	2.11	8 (47%)
4	NAG	B	910	1	14,14,15	0.94	0	17,19,21	2.37	6 (35%)
5	SO4	A	913	-	4,4,4	0.47	0	6,6,6	0.37	0
5	SO4	B	913	-	4,4,4	0.53	0	6,6,6	0.34	0
6	9K9	B	914	-	19,26,26	1.40	2 (10%)	21,41,41	3.14	10 (47%)

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
4	NAG	B	907	1	-	0/6/23/26	0/1/1/1

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	NAG	A	902	1	-	0/6/23/26	0/1/1/1
4	NAG	A	906	1	-	0/6/23/26	0/1/1/1
4	NAG	B	902	1	-	2/6/23/26	0/1/1/1
4	NAG	B	910	1	-	0/6/23/26	0/1/1/1
4	NAG	B	909	1	-	2/6/23/26	0/1/1/1
4	NAG	A	905	1	-	0/6/23/26	0/1/1/1
4	NAG	B	905	1	-	2/6/23/26	0/1/1/1
4	NAG	A	907	1	-	0/6/23/26	0/1/1/1
4	NAG	B	901	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	-	2/6/23/26	0/1/1/1
6	9K9	A	914	-	-	0/2/32/32	0/4/4/4
6	9K9	B	914	-	-	0/2/32/32	0/4/4/4

The worst 5 of 11 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	A	914	9K9	C2-C	4.40	1.48	1.41
6	B	914	9K9	C2-C	4.19	1.48	1.41
6	A	914	9K9	C-C1	2.66	1.48	1.40
6	A	914	9K9	O3-C6	-2.57	1.39	1.45
4	A	902	NAG	C1-C2	-2.54	1.48	1.52

The worst 5 of 92 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	B	914	9K9	C3-N-C1	7.21	123.59	115.36
4	B	910	NAG	C1-O5-C5	6.85	121.48	112.19
6	B	914	9K9	C2-C-C1	-6.73	114.37	120.80
4	A	907	NAG	O5-C5-C6	6.57	117.50	107.20
6	B	914	9K9	N-C3-N1	-5.64	119.70	127.22

There are no chirality outliers.

5 of 8 torsion outliers are listed below:

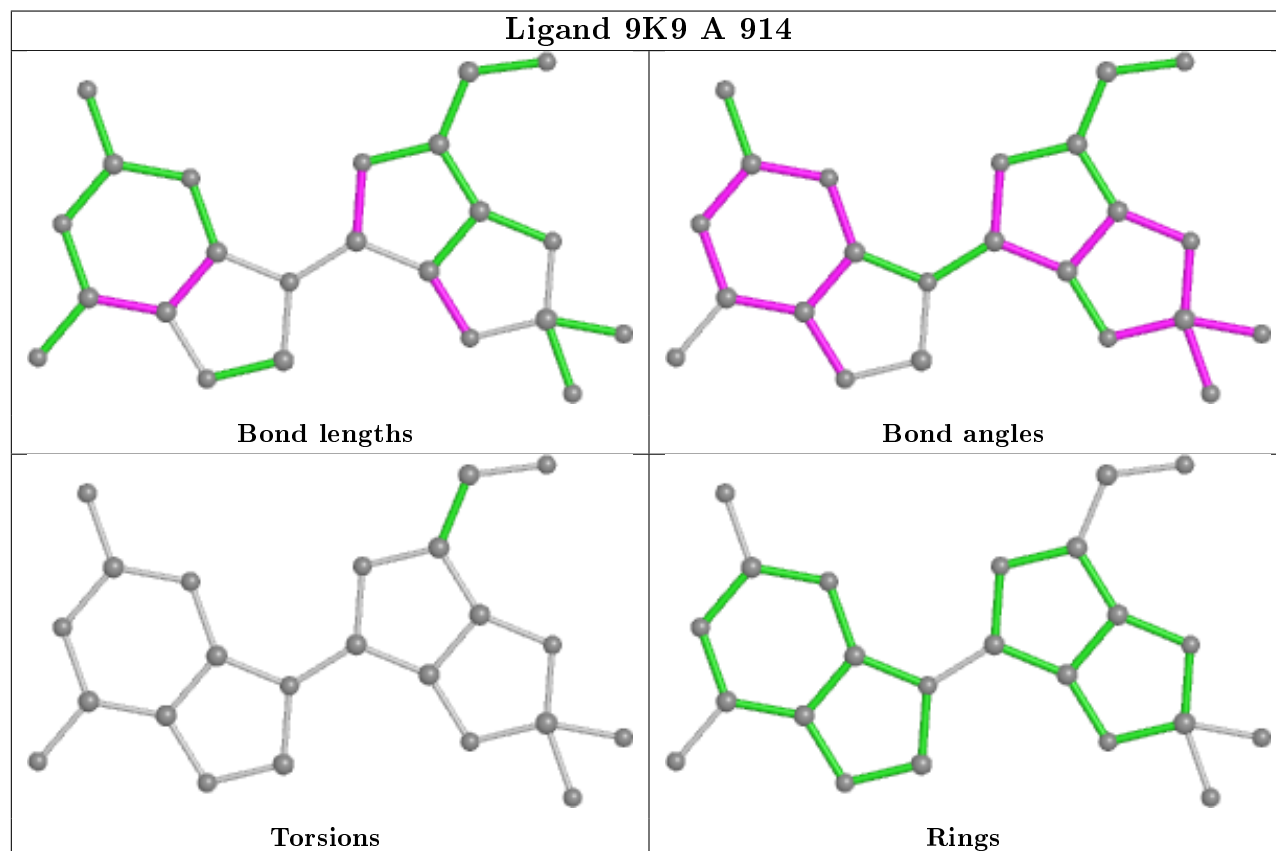
Mol	Chain	Res	Type	Atoms
4	B	908	NAG	O5-C5-C6-O6
4	B	908	NAG	C4-C5-C6-O6
4	B	902	NAG	C4-C5-C6-O6
4	B	909	NAG	C4-C5-C6-O6
4	B	905	NAG	C4-C5-C6-O6

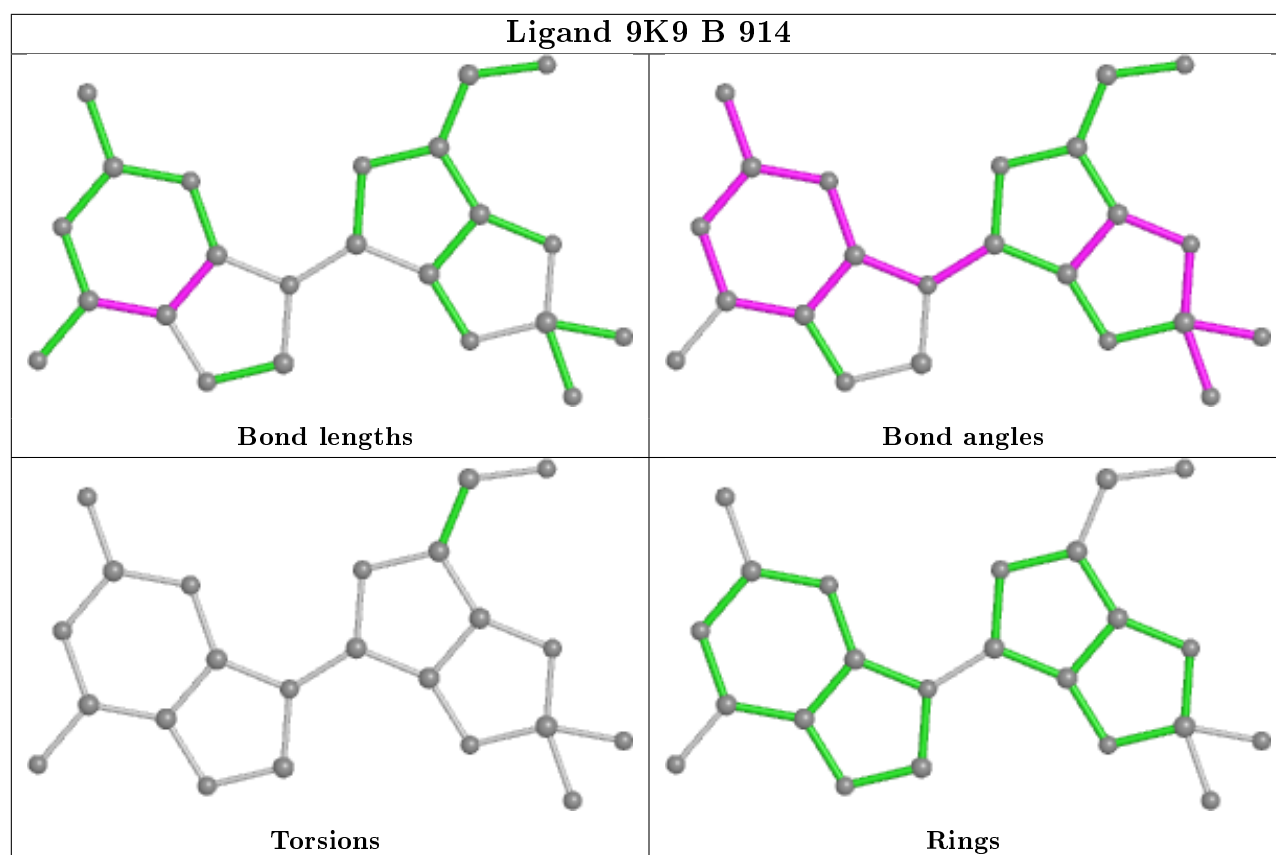
There are no ring outliers.

5 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	A	914	9K9	1	0
5	A	910	SO4	1	0
5	B	911	SO4	1	0
5	A	912	SO4	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	774/823 (94%)	0.24	16 (2%) 63 66	31, 51, 81, 113	0
1	B	776/823 (94%)	0.06	4 (0%) 91 91	25, 42, 68, 104	0
2	D	5/6 (83%)	2.35	3 (60%) 0 0	83, 93, 109, 115	0
2	E	5/6 (83%)	3.11	5 (100%) 0 0	84, 106, 110, 125	0
All	All	1560/1658 (94%)	0.17	28 (1%) 68 71	25, 46, 79, 125	0

The worst 5 of 28 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	3	C	4.6
2	E	1	G	3.9
1	A	42	VAL	3.8
1	A	40	LEU	3.5
2	E	5	C	3.4

### 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.

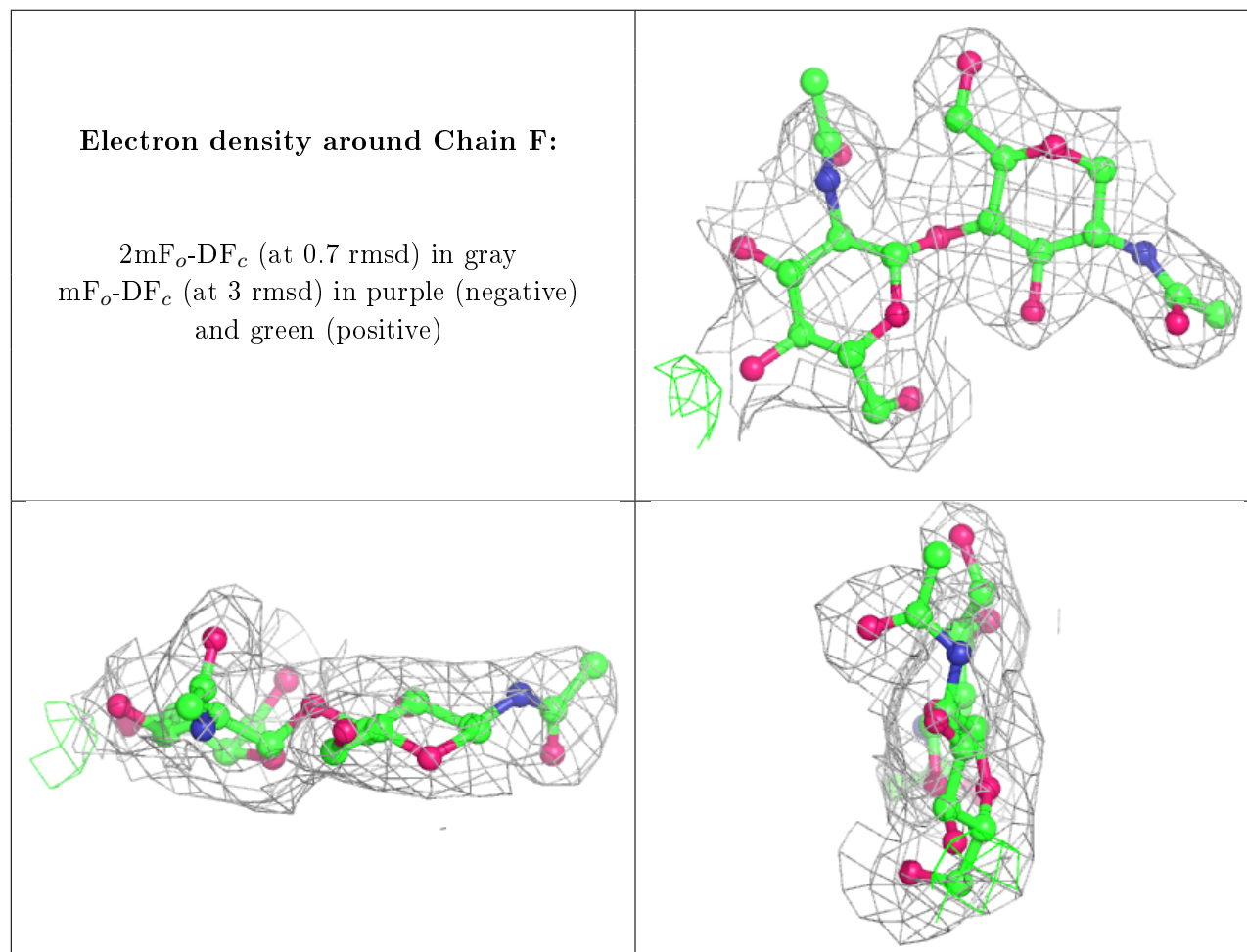
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
3	NAG	F	2	14/15	0.84	0.13	63,80,94,97	0
3	NAG	C	2	14/15	0.89	0.17	57,74,83,88	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
3	NAG	C	1	14/15	0.94	0.14	40,47,53,62	0
3	NAG	F	1	14/15	0.96	0.13	45,49,59,68	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.



## 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( $\text{\AA}^2$ )	Q<0.9
5	SO4	A	911	5/5	0.80	0.30	71,86,103,111	0
5	SO4	A	912	5/5	0.80	0.36	83,95,103,109	0
4	NAG	B	908	14/15	0.82	0.22	76,84,100,105	0

*Continued on next page...*

*Continued from previous page...*

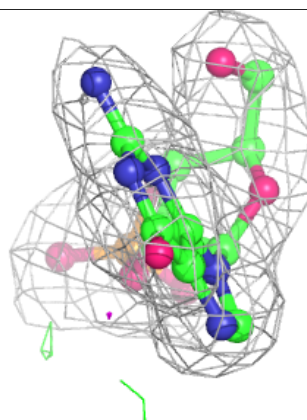
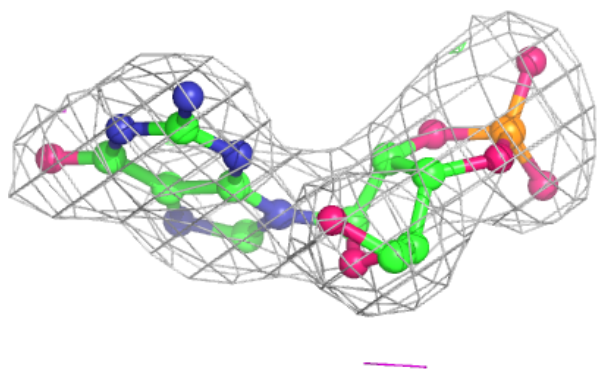
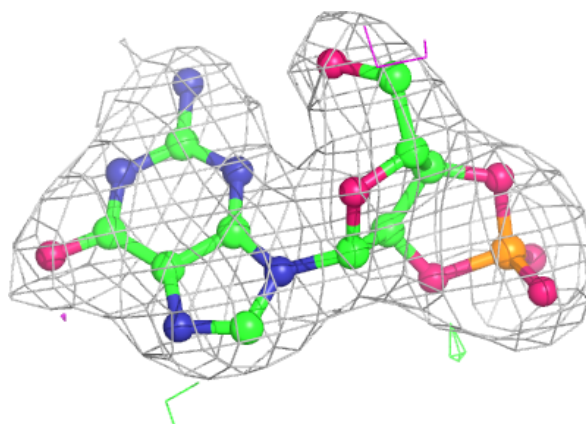
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
4	NAG	A	907	14/15	0.84	0.20	66,88,100,111	0
4	NAG	B	910	14/15	0.85	0.25	61,69,73,74	0
4	NAG	A	905	14/15	0.86	0.15	54,62,67,72	0
4	NAG	A	906	14/15	0.89	0.17	56,72,79,81	0
4	NAG	B	909	14/15	0.90	0.16	54,58,67,68	0
5	SO4	A	913	5/5	0.90	0.27	82,87,102,104	0
5	SO4	B	913	5/5	0.90	0.28	78,83,87,108	0
4	NAG	B	907	14/15	0.92	0.15	60,64,75,79	0
4	NAG	B	905	14/15	0.93	0.20	50,55,66,69	0
4	NAG	A	901	14/15	0.93	0.14	47,54,63,67	0
5	SO4	B	912	5/5	0.93	0.20	68,71,82,93	0
4	NAG	B	906	14/15	0.93	0.11	49,59,65,67	0
5	SO4	A	910	5/5	0.94	0.23	62,66,75,78	0
5	SO4	B	911	5/5	0.95	0.15	50,52,61,88	0
4	NAG	B	902	14/15	0.95	0.15	33,36,43,53	0
4	NAG	A	902	14/15	0.95	0.14	39,44,49,51	0
5	SO4	A	909	5/5	0.95	0.12	58,62,74,80	0
4	NAG	B	901	14/15	0.95	0.12	42,47,59,60	0
6	9K9	A	914	23/23	0.97	0.16	33,39,44,48	0
5	SO4	A	908	5/5	0.97	0.17	47,49,56,64	0
6	9K9	B	914	23/23	0.97	0.16	29,37,42,48	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.

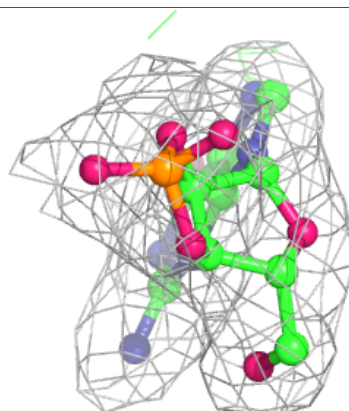
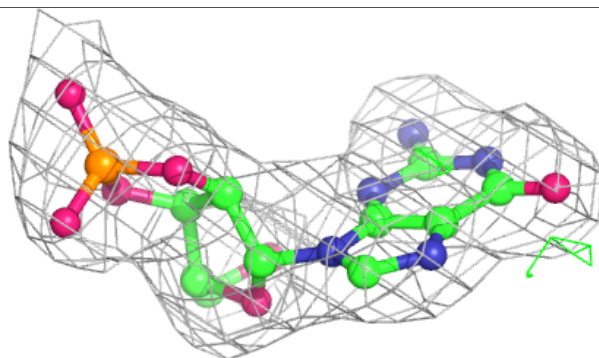
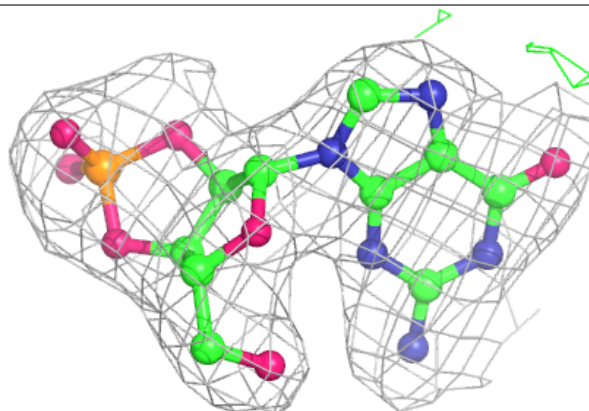


**Electron density around 9K9 A 914:**

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

**Electron density around 9K9 B 914:**

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



## 6.5 Other polymers [i](#)

There are no such residues in this entry.