



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 22, 2020 – 12:50 PM BST

PDB ID : 6TYQ  
Title : Salmonella Typhi PltB Homopentamer with Neu-5NAc-9OAc-alpha-2-6-Gal-beta-1-4-GlcNAc Glycans  
Authors : Nguyen, T.; Milano, S.K.; Yang, Y.A.; Song, J.  
Deposited on : 2019-08-09  
Resolution : 1.88 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  
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

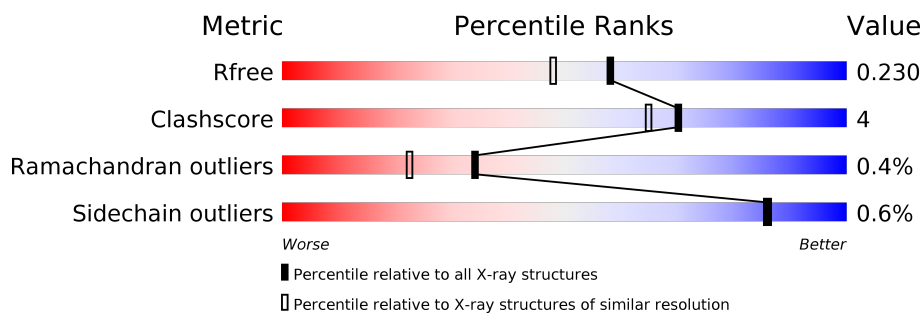
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.88 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.



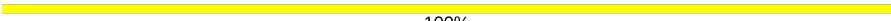
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	9470 (1.90-1.86)
Clashscore	141614	10282 (1.90-1.86)
Ramachandran outliers	138981	10152 (1.90-1.86)
Sidechain outliers	138945	10152 (1.90-1.86)

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  $\leq 5\%$ .

Mol	Chain	Length	Quality of chain
1	A	114	<div> <div style="width: 92%;"></div> <div style="width: 8%;"></div> <div>92%</div> <div>8%</div> </div>
1	B	114	<div> <div style="width: 93%;"></div> <div style="width: 6%;"></div> <div>93%</div> <div>6%</div> </div>
1	C	114	<div> <div style="width: 91%;"></div> <div style="width: 8%;"></div> <div>91%</div> <div>8%</div> </div>
1	D	114	<div> <div style="width: 94%;"></div> <div style="width: 6%;"></div> <div>94%</div> <div>6%</div> </div>
1	E	114	<div> <div style="width: 96%;"></div> <div style="width: 4%;"></div> <div>96%</div> <div>4%</div> </div>
2	F	3	<div> <div style="width: 100%;"></div> <div>100%</div> </div>
2	G	3	<div> <div style="width: 100%;"></div> <div>100%</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	H	3	 100%

## 2 Entry composition [i](#)

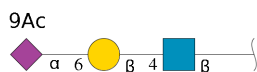
There are 3 unique types of molecules in this entry. The entry contains 4878 atoms, of which 129 are hydrogens and 0 are deuteriums.

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 Pertussis-like toxin subunit B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	114	Total	C	N	O	S	0	0	0
			885	566	145	170	4			
1	B	114	Total	C	N	O	S	0	0	0
			885	566	145	170	4			
1	C	114	Total	C	N	O	S	0	0	0
			885	566	145	170	4			
1	D	114	Total	C	N	O	S	0	0	0
			885	566	145	170	4			
1	E	114	Total	C	N	O	S	0	0	0
			885	566	145	170	4			

- Molecule 2 is an oligosaccharide called 9-O-acetyl-5-acetamido-3,5-dideoxy-D-glycero- $\alpha$ -D-galacto-non-2-ulopyranosonic acid-(2-6)- $\beta$ -D-galactopyranose-(1-4)-2-acetamido-2-deoxy- $\beta$ -D-glucopyranose.



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	F	3	Total	C	H	N	O	0	0	0
			92	27	43	2	20			
2	G	3	Total	C	H	N	O	0	0	0
			92	27	43	2	20			
2	H	3	Total	C	H	N	O	0	0	0
			92	27	43	2	20			

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	39	Total	O	0	0
			39	39		

*Continued on next page...*

*Continued from previous page...*

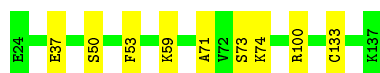
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	33	Total 33	O 33	0	0
3	C	35	Total 35	O 35	0	0
3	D	32	Total 32	O 32	0	0
3	E	38	Total 38	O 38	0	0

### 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.

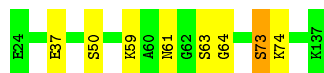
- Molecule 1: Pertussis-like toxin subunit B

Chain A:  92% 8%



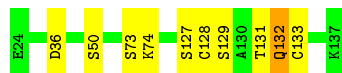
- Molecule 1: Pertussis-like toxin subunit B

Chain B:  93% 6%



- Molecule 1: Pertussis-like toxin subunit B

Chain C:  91% 8%



- Molecule 1: Pertussis-like toxin subunit B

Chain D:  94% 6%



- Molecule 1: Pertussis-like toxin subunit B

Chain E:  96%



- Molecule 2: 9-O-acetyl-5-acetamido-3,5-dideoxy-D-glycero-alpha-D-galacto-non-2-ulopyranosonic acid-(2-6)-beta-D-galactopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain F:  100%

NAG1  
GAL2  
5163

- Molecule 2: 9-O-acetyl-5-acetamido-3,5-dideoxy-D-glycero-alpha-D-galacto-non-2-ulopyranosonic acid-(2-6)-beta-D-galactopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain G:

100%

NAG1  
GAL2  
5163

- Molecule 2: 9-O-acetyl-5-acetamido-3,5-dideoxy-D-glycero-alpha-D-galacto-non-2-ulopyranosonic acid-(2-6)-beta-D-galactopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain H:

100%

NAG1  
GAL2  
5163

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	69.75Å 98.92Å 99.63Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	99.63 – 1.88 57.14 – 1.88	Depositor EDS
% Data completeness (in resolution range)	98.5 (99.63-1.88) 98.5 (57.14-1.88)	Depositor EDS
$R_{merge}$	0.13	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.12 (at 1.88Å)	Xtriage
Refinement program	PHENIX (1.16_3549: ???)	Depositor
R, $R_{free}$	(Not available) , (Not available) 0.199 , 0.230	Depositor DCC
$R_{free}$ test set	2018 reflections (3.64%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	22.2	Xtriage
Anisotropy	0.280	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.39 , 49.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.008 for -h,l,k	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	4878	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	23.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.06% 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: GAL, NAG, 5N6

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	0.38	0/909	0.56	0/1240
1	B	0.41	0/909	0.59	0/1240
1	C	0.42	0/909	0.60	0/1240
1	D	0.40	0/909	0.55	0/1240
1	E	0.38	0/909	0.56	0/1240
All	All	0.40	0/4545	0.57	0/6200

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts [i](#)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	885	0	846	4	0
1	B	885	0	846	6	1
1	C	885	0	846	15	1
1	D	885	0	846	4	0
1	E	885	0	846	3	0
2	F	49	43	23	0	0
2	G	49	43	23	0	0
2	H	49	43	23	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	A	39	0	0	0	0
3	B	33	0	0	0	0
3	C	35	0	0	0	0
3	D	32	0	0	0	0
3	E	38	0	0	0	0
All	All	4749	129	4299	32	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 4.

The worst 5 of 32 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:C:129:SER:H	1:C:132:GLN:HB3	1.05	1.20
1:B:61:ASN:HD21	1:B:63:SER:HB3	1.15	1.03
1:C:129:SER:N	1:C:132:GLN:HB3	1.82	0.94
1:B:61:ASN:ND2	1:B:63:SER:HB3	1.88	0.89
1:B:50:SER:OG	1:B:73:SER:O	1.91	0.89

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)
1:B:61:ASN:ND2	1:C:131:THR:OG1[4_455]	2.06	0.14

## 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	112/114 (98%)	109 (97%)	3 (3%)	0	100	100
1	B	112/114 (98%)	106 (95%)	5 (4%)	1 (1%)	17	7

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	112/114 (98%)	105 (94%)	6 (5%)	1 (1%)	17	7
1	D	112/114 (98%)	108 (96%)	4 (4%)	0	100	100
1	E	112/114 (98%)	109 (97%)	3 (3%)	0	100	100
All	All	560/570 (98%)	537 (96%)	21 (4%)	2 (0%)	34	22

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	132	GLN
1	B	64	GLY

### 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	98/98 (100%)	97 (99%)	1 (1%)	76	73
1	B	98/98 (100%)	97 (99%)	1 (1%)	76	73
1	C	98/98 (100%)	98 (100%)	0	100	100
1	D	98/98 (100%)	98 (100%)	0	100	100
1	E	98/98 (100%)	97 (99%)	1 (1%)	76	73
All	All	490/490 (100%)	487 (99%)	3 (1%)	86	86

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	73	SER
1	B	73	SER
1	E	137	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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 ⓘ

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	NAG	F	1	2	15,15,15	1.65	4 (26%)	21,21,21	1.26	4 (19%)
2	GAL	F	2	2	11,11,12	1.89	2 (18%)	15,15,17	1.27	1 (6%)
2	5N6	F	3	2	20,23,24	2.09	7 (35%)	25,32,35	1.59	4 (16%)
2	NAG	G	1	2	15,15,15	1.60	3 (20%)	21,21,21	1.86	6 (28%)
2	GAL	G	2	2	11,11,12	1.92	2 (18%)	15,15,17	2.81	7 (46%)
2	5N6	G	3	2	20,23,24	2.10	6 (30%)	25,32,35	1.48	7 (28%)
2	NAG	H	1	2	15,15,15	1.78	6 (40%)	21,21,21	1.21	2 (9%)
2	GAL	H	2	2	11,11,12	1.84	2 (18%)	15,15,17	1.55	2 (13%)
2	5N6	H	3	2	20,23,24	2.17	7 (35%)	25,32,35	1.57	3 (12%)

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	NAG	F	1	2	-	2/6/26/26	0/1/1/1
2	GAL	F	2	2	-	0/2/19/22	0/1/1/1
2	5N6	F	3	2	-	0/17/37/41	0/1/1/1
2	NAG	G	1	2	-	1/6/26/26	0/1/1/1

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	GAL	G	2	2	-	1/2/19/22	0/1/1/1
2	5N6	G	3	2	-	4/17/37/41	0/1/1/1
2	NAG	H	1	2	-	2/6/26/26	0/1/1/1
2	GAL	H	2	2	-	0/2/19/22	0/1/1/1
2	5N6	H	3	2	-	3/17/37/41	0/1/1/1

The worst 5 of 39 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	H	3	5N6	O6-C6	5.08	1.51	1.44
2	F	3	5N6	O6-C6	5.02	1.51	1.44
2	G	3	5N6	O6-C6	4.74	1.51	1.44
2	G	3	5N6	C10-N5	4.02	1.48	1.34
2	F	2	GAL	O5-C5	3.94	1.51	1.43

The worst 5 of 36 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	G	2	GAL	C1-C2-C3	6.88	118.12	109.67
2	G	1	NAG	O5-C5-C4	4.92	118.63	109.69
2	H	2	GAL	C1-C2-C3	4.26	114.90	109.67
2	G	2	GAL	O5-C5-C6	4.09	113.62	107.20
2	G	2	GAL	C6-C5-C4	-4.00	103.64	113.00

There are no chirality outliers.

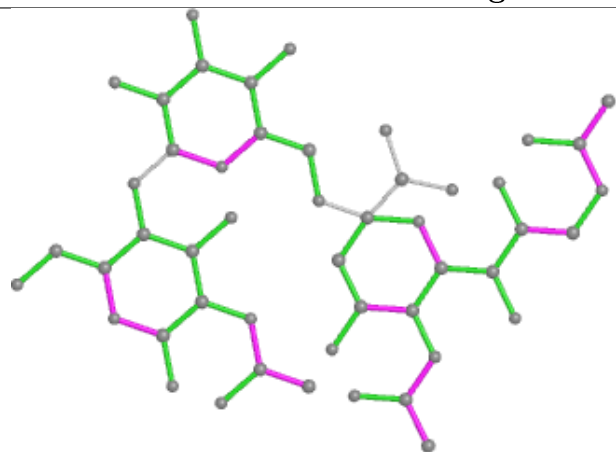
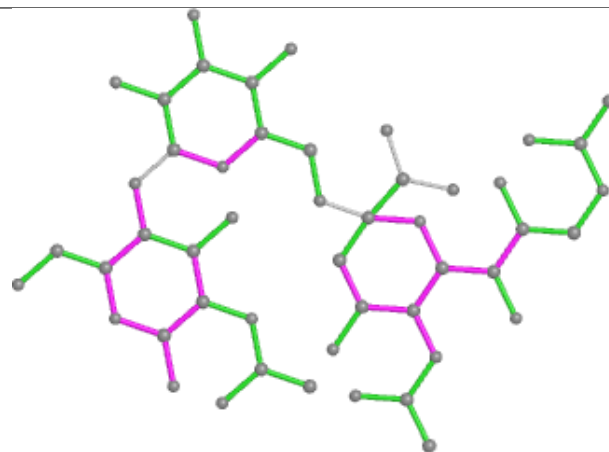
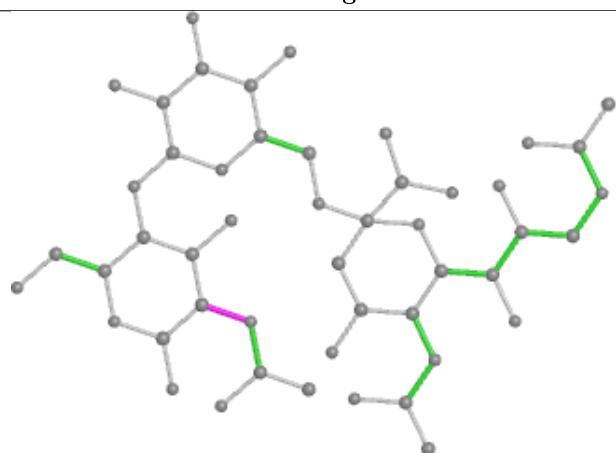
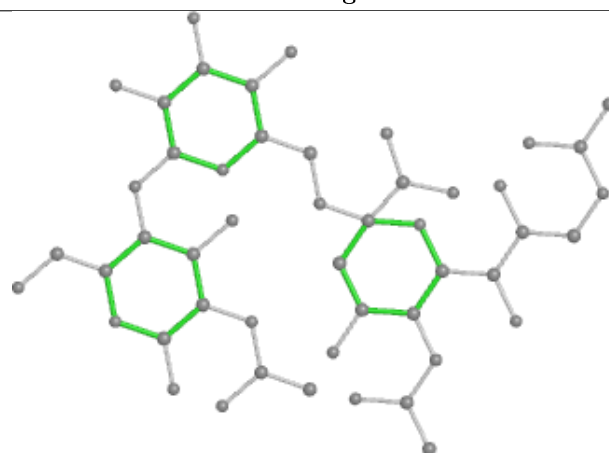
5 of 13 torsion outliers are listed below:

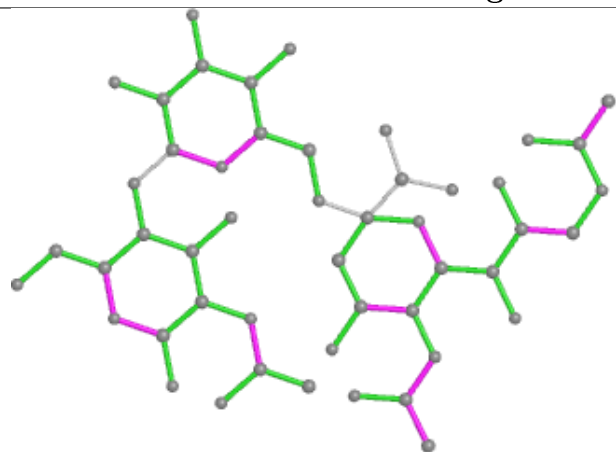
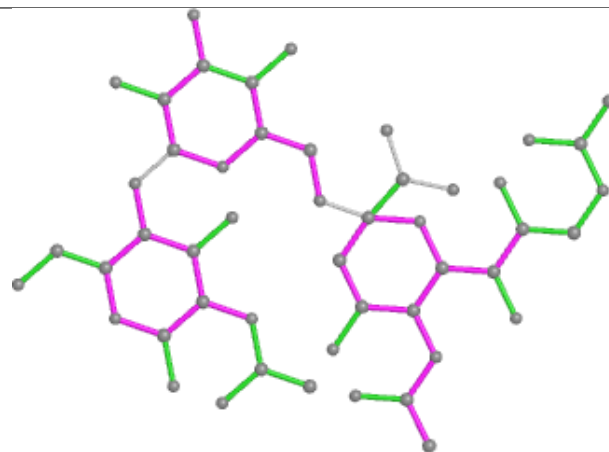
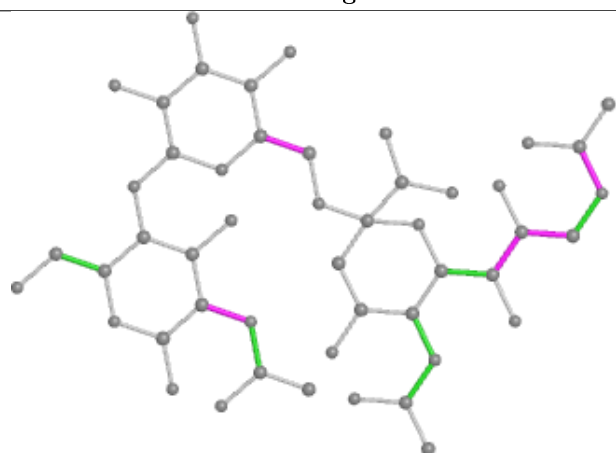
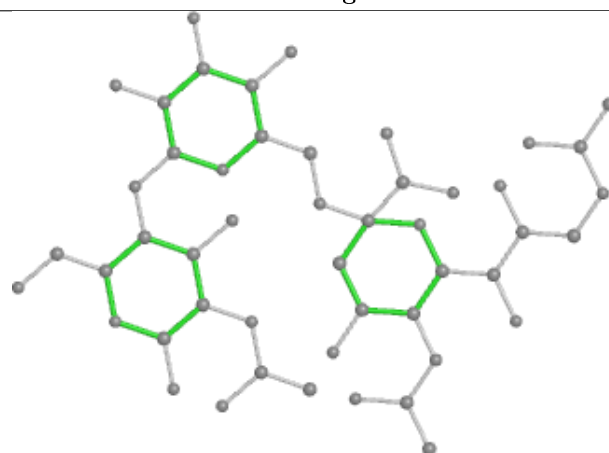
Mol	Chain	Res	Type	Atoms
2	G	1	NAG	C3-C2-N2-C7
2	G	3	5N6	CAF-CAG-O9-C9
2	G	3	5N6	OBJ-CAG-O9-C9
2	F	1	NAG	C3-C2-N2-C7
2	G	2	GAL	O5-C5-C6-O6

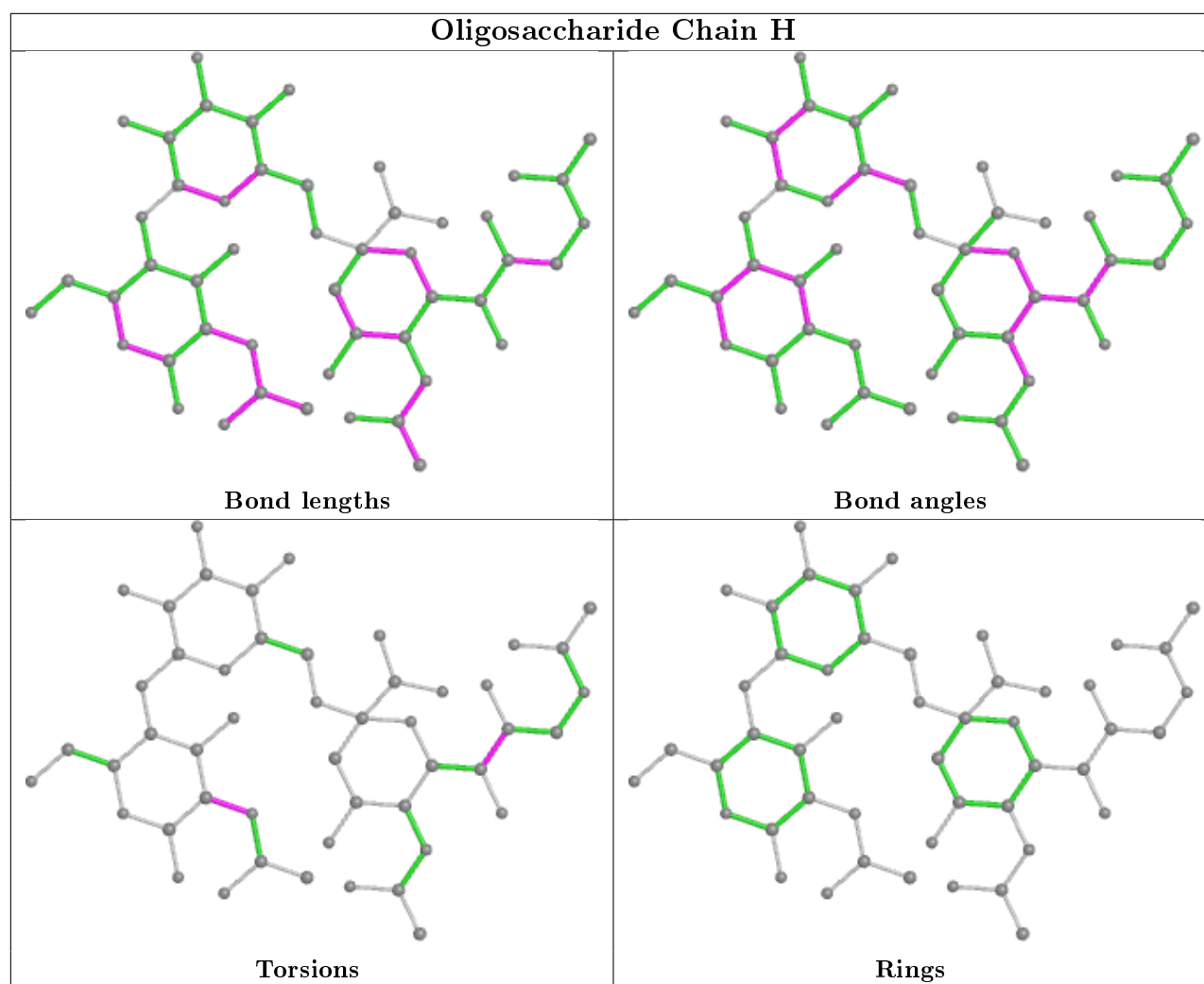
There are no ring outliers.

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.

**Oligosaccharide Chain F****Bond lengths****Bond angles****Torsions****Rings**

**Oligosaccharide Chain G****Bond lengths****Bond angles****Torsions****Rings**



## 5.6 Ligand geometry [i](#)

There are no ligands in this entry.

## 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 ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

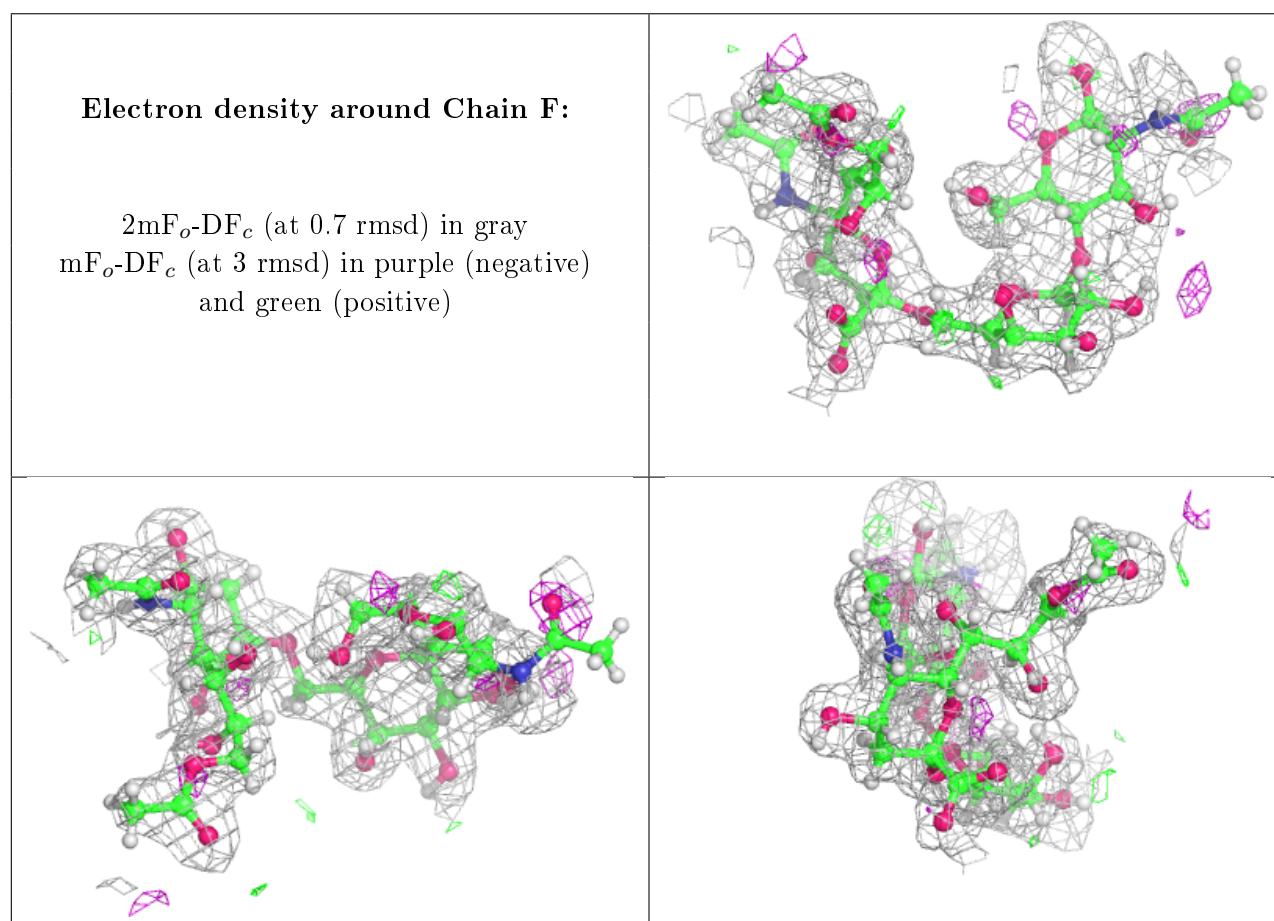
### 6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.3 Carbohydrates ⓘ

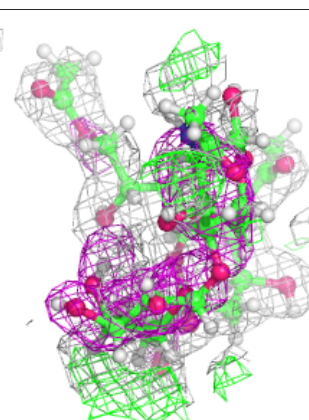
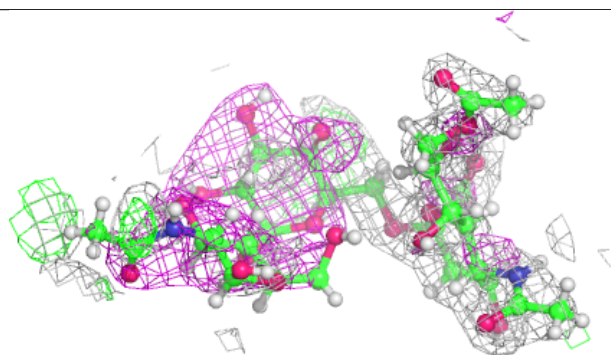
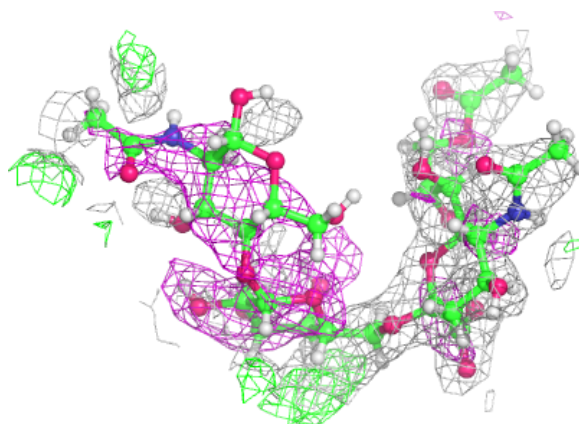
Unable to reproduce the depositors R factor - this section is therefore empty.

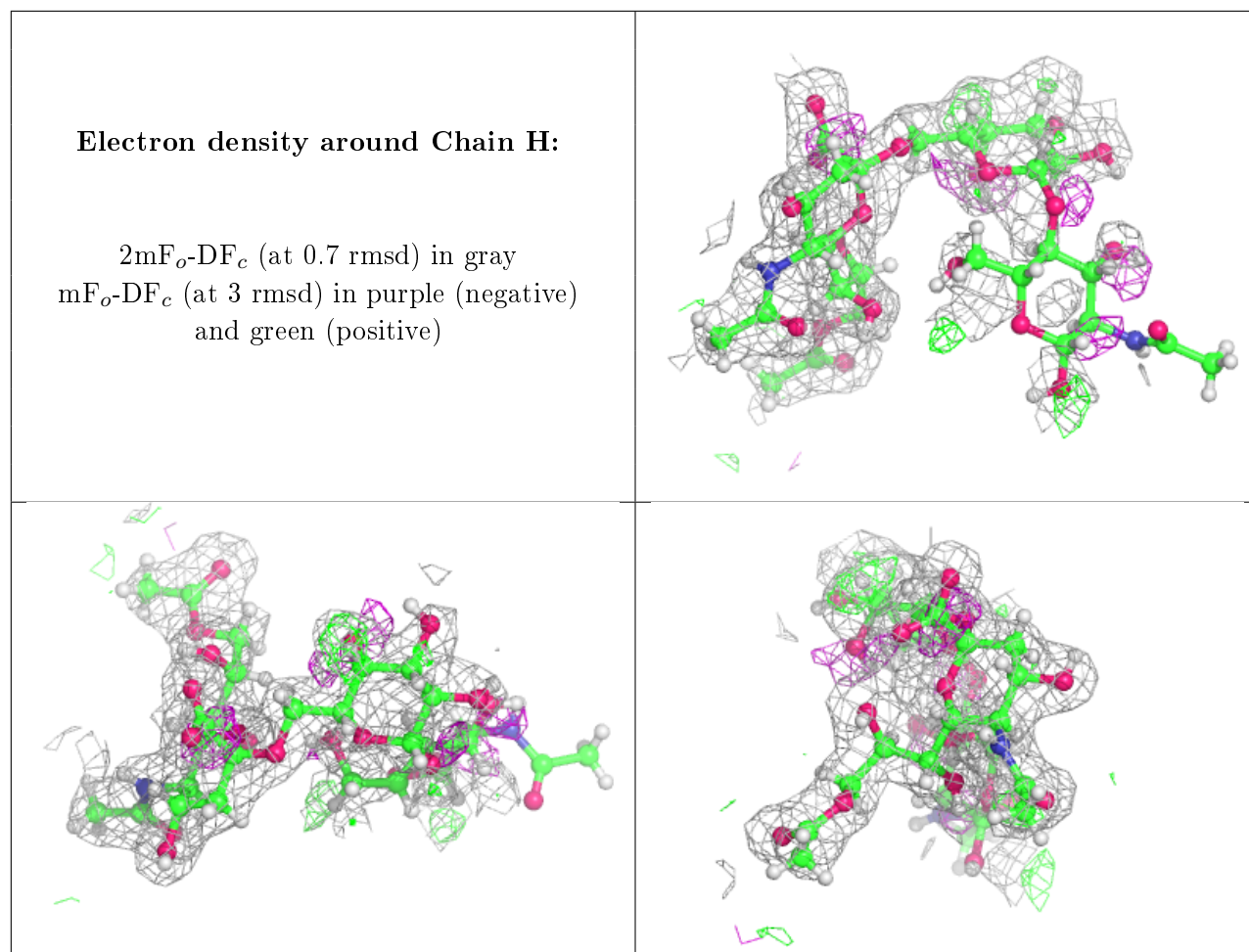
The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.



**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 [i](#)

Unable to reproduce the depositors R factor - this section is therefore empty.

## 6.5 Other polymers [i](#)

Unable to reproduce the depositors R factor - this section is therefore empty.