



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 8, 2020 – 02:11 PM BST

PDB ID : 4CQW  
Title : H5 (tyTy) Del133/Ile155Thr Mutant Haemagglutinin in Complex with Avian Receptor Analogue 3'SLN  
Authors : Xiong, X.; Xiao, H.; Martin, S.R.; Coombs, P.J.; Liu, J.; Collins, P.J.; Vachieri, S.G.; Walker, P.A.; Lin, Y.P.; McCauley, J.W.; Gamblin, S.J.; Skehel, J.J.  
Deposited on : 2014-02-21  
Resolution : 2.30 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

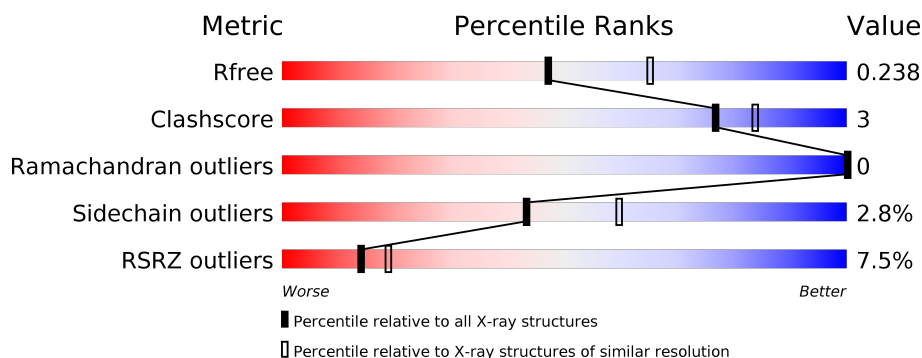
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 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.



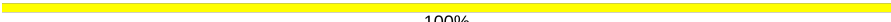

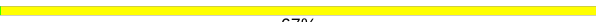
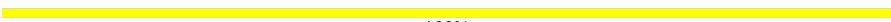
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	5042 (2.30-2.30)
Clashscore	141614	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)
RSRZ outliers	127900	4938 (2.30-2.30)

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	327	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 1%, green 99%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>89%</span> <span>9%</span> </div> </div>
1	C	327	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 1%, green 99%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>91%</span> <span>7%</span> </div> </div>
1	E	327	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 2%, green 98%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>90%</span> <span>8%</span> </div> </div>
2	B	166	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 4%, green 96%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>82%</span> <span>16%</span> </div> </div>
2	D	166	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 25%, green 75%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>84%</span> <span>14%</span> </div> </div>
2	F	166	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 30%, green 70%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>90%</span> <span>7%</span> </div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
3	G	3	 100%
3	H	3	 33%  67%
3	I	3	 100%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	NAG	A	1023	X	-	-	-

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 12709 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 HAEMAGGLUTININ HA1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	324	Total	C	N	O	S	0	0	0
			2572	1618	449	491	14			
1	C	322	Total	C	N	O	S	0	0	0
			2552	1607	443	488	14			
1	E	322	Total	C	N	O	S	0	0	0
			2552	1607	443	488	14			

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	ASP	-	expression tag	UNP Q207Z6
A	0	PRO	-	expression tag	UNP Q207Z6
A	?	-	ALA	deletion	UNP Q207Z6
A	150	THR	ILE	engineered mutation	UNP Q207Z6
A	322	ARG	GLY	conflict	UNP Q207Z6
A	324	THR	ARG	conflict	UNP Q207Z6
C	-1	ASP	-	expression tag	UNP Q207Z6
C	0	PRO	-	expression tag	UNP Q207Z6
C	?	-	ALA	deletion	UNP Q207Z6
C	150	THR	ILE	engineered mutation	UNP Q207Z6
C	322	ARG	GLY	conflict	UNP Q207Z6
C	324	THR	ARG	conflict	UNP Q207Z6
E	-1	ASP	-	expression tag	UNP Q207Z6
E	0	PRO	-	expression tag	UNP Q207Z6
E	?	-	ALA	deletion	UNP Q207Z6
E	150	THR	ILE	engineered mutation	UNP Q207Z6
E	322	ARG	GLY	conflict	UNP Q207Z6
E	324	THR	ARG	conflict	UNP Q207Z6

- Molecule 2 is a protein called HAEMAGGLUTININ HA2.

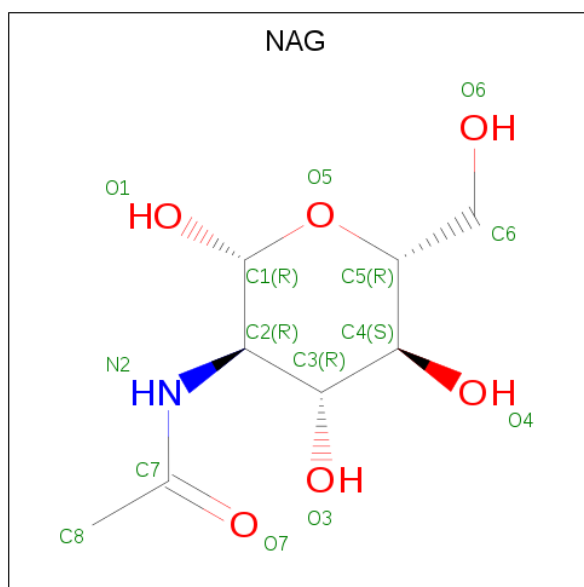
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	163	Total	C	N	O	S	0	0	0
			1320	822	229	261	8			
2	D	163	Total	C	N	O	S	0	0	0
			1320	822	229	261	8			
2	F	163	Total	C	N	O	S	0	0	0
			1320	822	229	261	8			

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



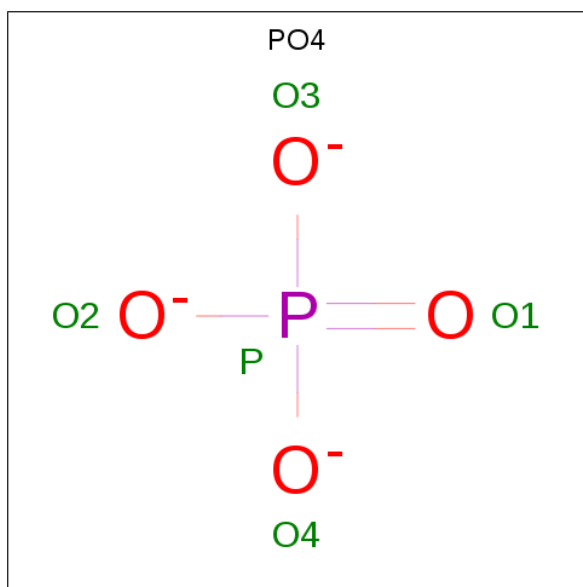
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	G	3	Total	C	N	O	0	0	0
			46	25	2	19			
3	H	3	Total	C	N	O	0	0	0
			46	25	2	19			
3	I	3	Total	C	N	O	0	0	0
			46	25	2	19			

- 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	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	C	1	Total	C	N	O	0	0
			14	8	1	5		
4	C	1	Total	C	N	O	0	0
			14	8	1	5		
4	C	1	Total	C	N	O	0	0
			14	8	1	5		
4	E	1	Total	C	N	O	0	0
			14	8	1	5		
4	E	1	Total	C	N	O	0	0
			14	8	1	5		
4	E	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 5 is PHOSPHATE ION (three-letter code: PO4) (formula: O<sub>4</sub>P).



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

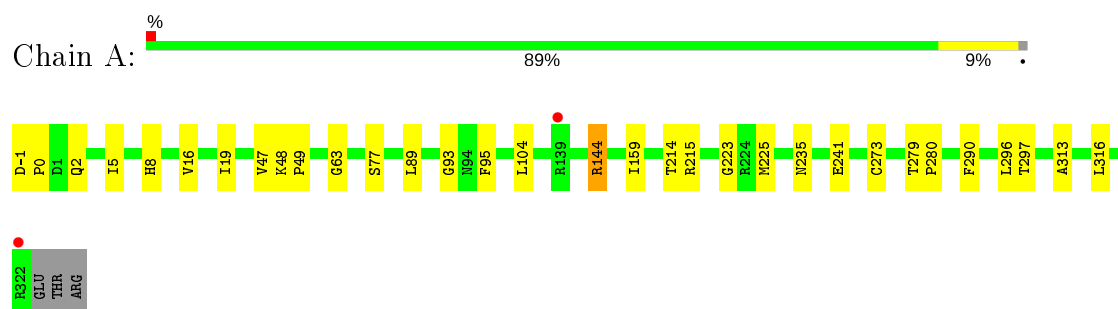
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	197	Total 197	O 197	0	0
6	B	64	Total 64	O 64	0	0
6	C	262	Total 262	O 262	0	0
6	D	30	Total 30	O 30	0	0
6	E	225	Total 225	O 225	0	0
6	F	26	Total 26	O 26	0	0

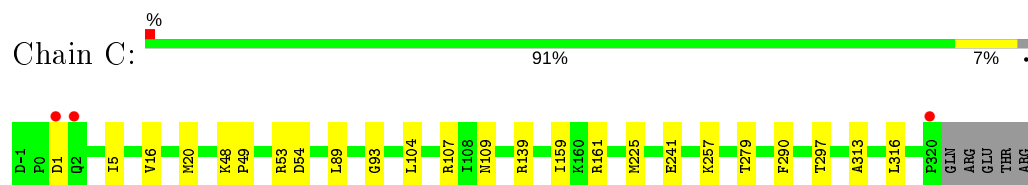
### 3 Residue-property plots [i](#)

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 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.

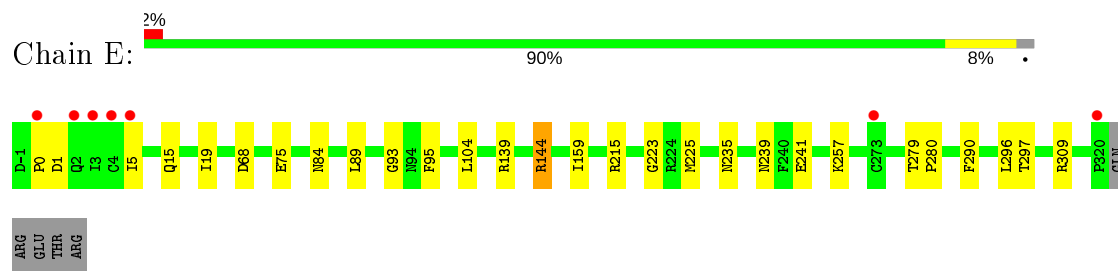
- Molecule 1: HAEMAGGLUTININ HA1



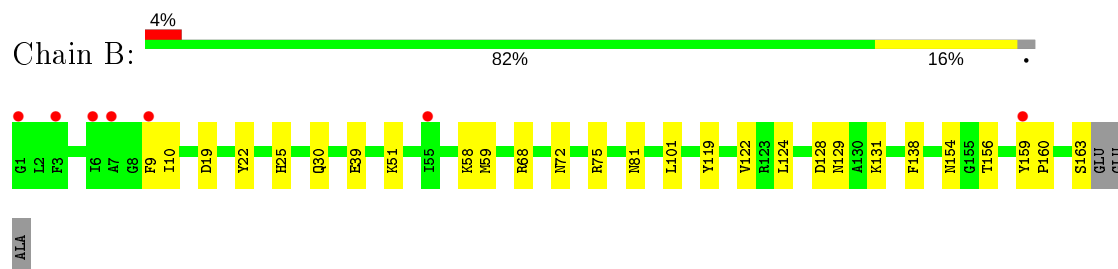
- Molecule 1: HAEMAGGLUTININ HA1



- Molecule 1: HAEMAGGLUTININ HA1

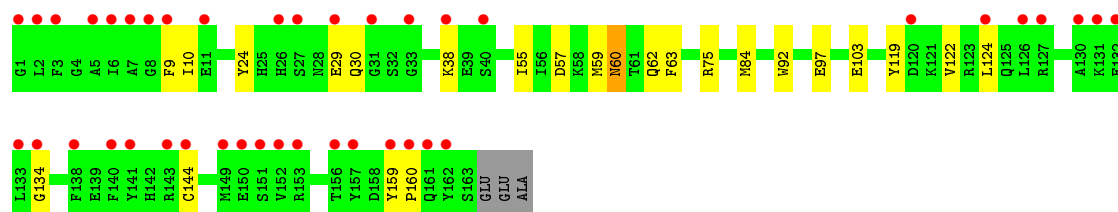
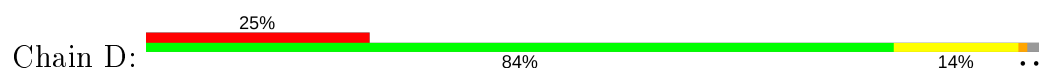


- Molecule 2: HAEMAGGLUTININ HA2

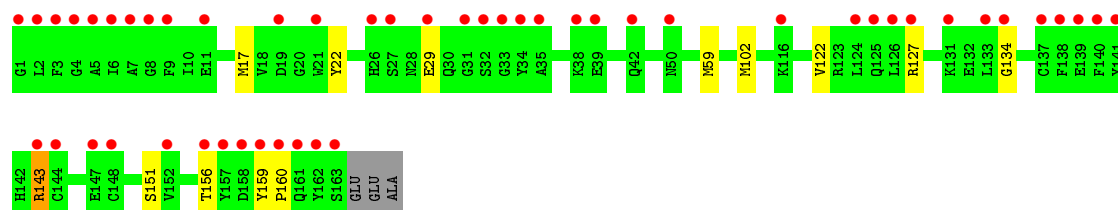
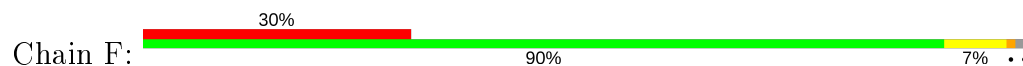


- Molecule 2: HAEMAGGLUTININ HA2





• Molecule 2: HAEMAGGLUTININ HA2



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



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



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



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	87.64Å 117.55Å 101.16Å 90.00° 92.50° 90.00°	Depositor
Resolution (Å)	101.06 – 2.30 87.55 – 2.30	Depositor EDS
% Data completeness (in resolution range)	98.3 (101.06-2.30) 98.3 (87.55-2.30)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.05 (at 2.29Å)	Xtriage
Refinement program	REFMAC 5.8.0046	Depositor
R, $R_{free}$	0.196 , 0.238 0.196 , 0.238	Depositor DCC
$R_{free}$ test set	4479 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	35.4	Xtriage
Anisotropy	0.592	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 41.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.024 for h,-k,-l	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	12709	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 3.52% 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: PO4, GAL, NAG, SIA

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	0.32	0/2634	0.53	0/3580
1	C	0.34	0/2614	0.53	0/3554
1	E	0.33	0/2614	0.54	0/3554
2	B	0.34	0/1347	0.49	0/1813
2	D	0.32	0/1347	0.48	0/1813
2	F	0.32	0/1347	0.47	0/1813
All	All	0.33	0/11903	0.51	0/16127

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	2572	0	2504	18	0
1	C	2552	0	2483	12	0
1	E	2552	0	2483	17	0
2	B	1320	0	1224	17	0
2	D	1320	0	1224	14	0
2	F	1320	0	1224	8	0
3	G	46	0	40	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	H	46	0	40	0	0
3	I	46	0	40	0	0
4	A	42	0	39	2	0
4	C	42	0	39	1	0
4	E	42	0	39	4	0
5	A	5	0	0	0	0
6	A	197	0	0	0	0
6	B	64	0	0	2	0
6	C	262	0	0	0	0
6	D	30	0	0	0	0
6	E	225	0	0	0	0
6	F	26	0	0	0	0
All	All	12709	0	11379	70	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 3.

The worst 5 of 70 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:E:19:ILE:HD11	2:F:102:MET:HA	1.57	0.85
4:E:1023:NAG:H83	4:E:1023:NAG:H3	1.66	0.77
1:E:290:PHE:HZ	2:F:59:MET:HG3	1.57	0.69
1:C:290:PHE:HZ	2:D:59:MET:HG3	1.60	0.66
1:E:0:PRO:HA	2:F:143:ARG:HH22	1.60	0.66

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	322/327 (98%)	312 (97%)	10 (3%)	0	100	100
1	C	320/327 (98%)	308 (96%)	12 (4%)	0	100	100
1	E	320/327 (98%)	311 (97%)	9 (3%)	0	100	100
2	B	161/166 (97%)	156 (97%)	5 (3%)	0	100	100
2	D	161/166 (97%)	155 (96%)	6 (4%)	0	100	100
2	F	161/166 (97%)	158 (98%)	3 (2%)	0	100	100
All	All	1445/1479 (98%)	1400 (97%)	45 (3%)	0	100	100

There are no Ramachandran outliers to report.

### 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	291/294 (99%)	286 (98%)	5 (2%)	60	76
1	C	289/294 (98%)	282 (98%)	7 (2%)	49	66
1	E	289/294 (98%)	283 (98%)	6 (2%)	53	70
2	B	139/141 (99%)	134 (96%)	5 (4%)	35	49
2	D	139/141 (99%)	132 (95%)	7 (5%)	24	34
2	F	139/141 (99%)	133 (96%)	6 (4%)	29	40
All	All	1286/1305 (98%)	1250 (97%)	36 (3%)	43	60

5 of 36 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	316	LEU
2	D	38	LYS
2	F	143	ARG
2	D	29	GLU
2	D	60	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	110	HIS
2	F	30	GLN
2	D	60	ASN
1	A	235	ASN
1	E	235	ASN

### 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	$\# Z  > 2$
3	NAG	G	1	3	15,15,15	0.47	0	21,21,21	1.29	2 (9%)
3	GAL	G	2	3	11,11,12	0.33	0	15,15,17	1.12	1 (6%)
3	SIA	G	3	3	17,20,21	0.39	0	21,28,31	1.12	2 (9%)
3	NAG	H	1	3	15,15,15	0.45	0	21,21,21	2.02	4 (19%)
3	GAL	H	2	3	11,11,12	0.30	0	15,15,17	0.90	0
3	SIA	H	3	3	17,20,21	0.42	0	21,28,31	1.02	2 (9%)
3	NAG	I	1	3	15,15,15	0.47	0	21,21,21	1.23	2 (9%)
3	GAL	I	2	3	11,11,12	0.35	0	15,15,17	1.19	3 (20%)
3	SIA	I	3	3	17,20,21	0.50	0	21,28,31	1.10	3 (14%)

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	G	1	3	-	0/6/26/26	0/1/1/1
3	GAL	G	2	3	-	0/2/19/22	0/1/1/1
3	SIA	G	3	3	-	0/14/34/38	0/1/1/1
3	NAG	H	1	3	-	0/6/26/26	0/1/1/1
3	GAL	H	2	3	-	0/2/19/22	0/1/1/1
3	SIA	H	3	3	-	0/14/34/38	0/1/1/1
3	NAG	I	1	3	-	0/6/26/26	0/1/1/1
3	GAL	I	2	3	-	0/2/19/22	0/1/1/1
3	SIA	I	3	3	-	0/14/34/38	0/1/1/1

There are no bond length outliers.

The worst 5 of 19 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	H	1	NAG	C1-C2-N2	-6.05	103.72	110.73
3	G	1	NAG	C1-C2-N2	-4.79	105.18	110.73
3	H	1	NAG	C3-C2-N2	-4.39	102.33	110.62
3	I	1	NAG	C1-C2-N2	-3.57	106.59	110.73
3	I	3	SIA	C4-C5-N5	-3.30	103.85	110.38

There are no chirality outliers.

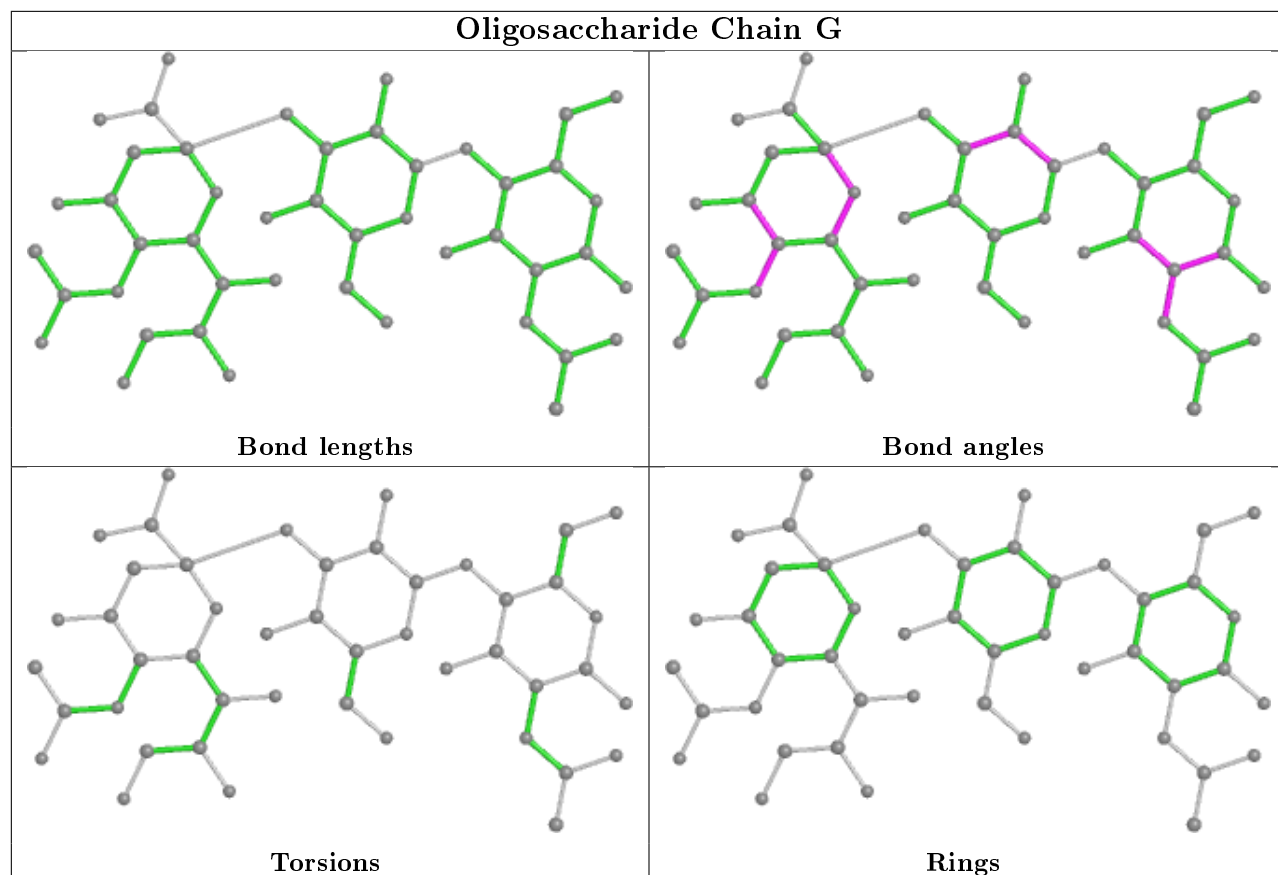
There are no torsion outliers.

There are no ring outliers.

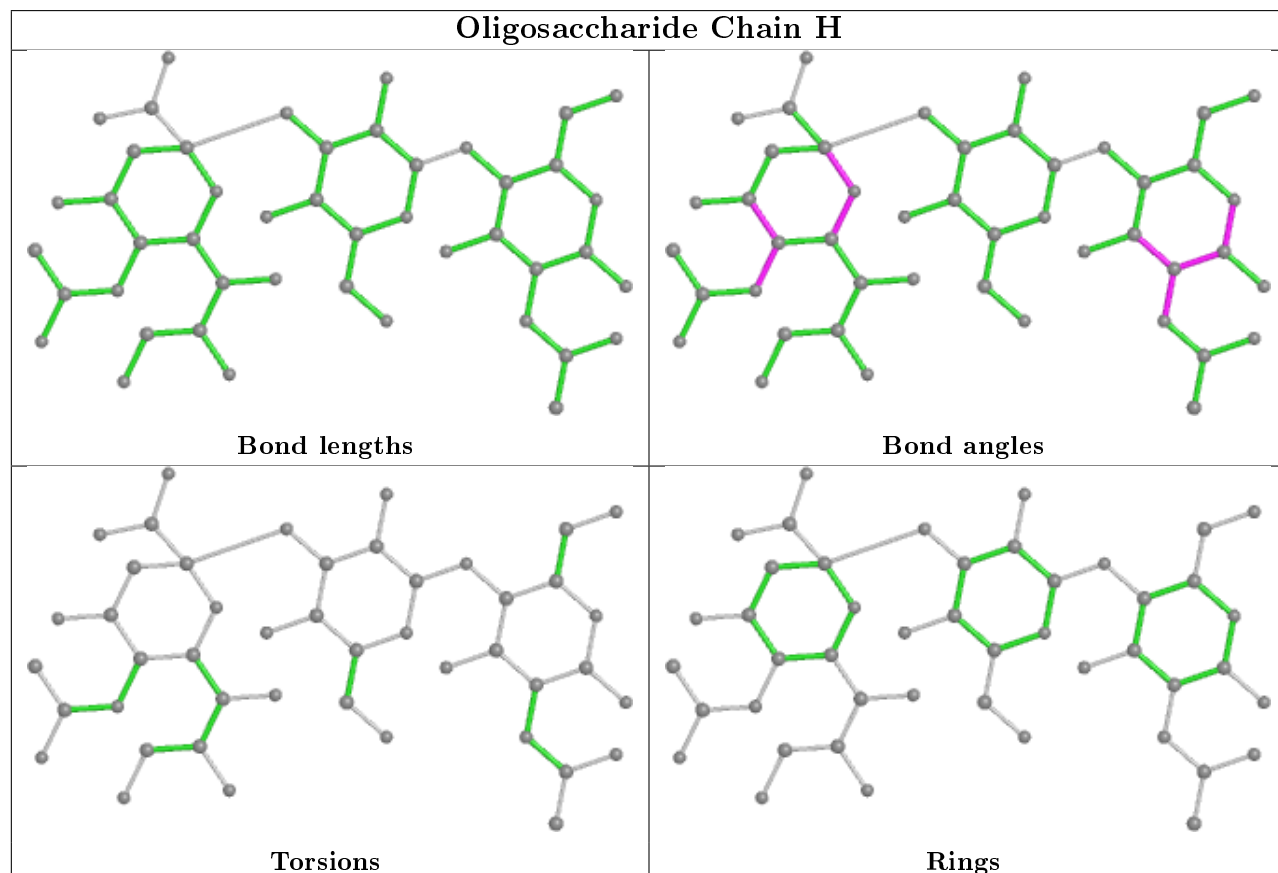
No monomer is involved in short contacts.

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

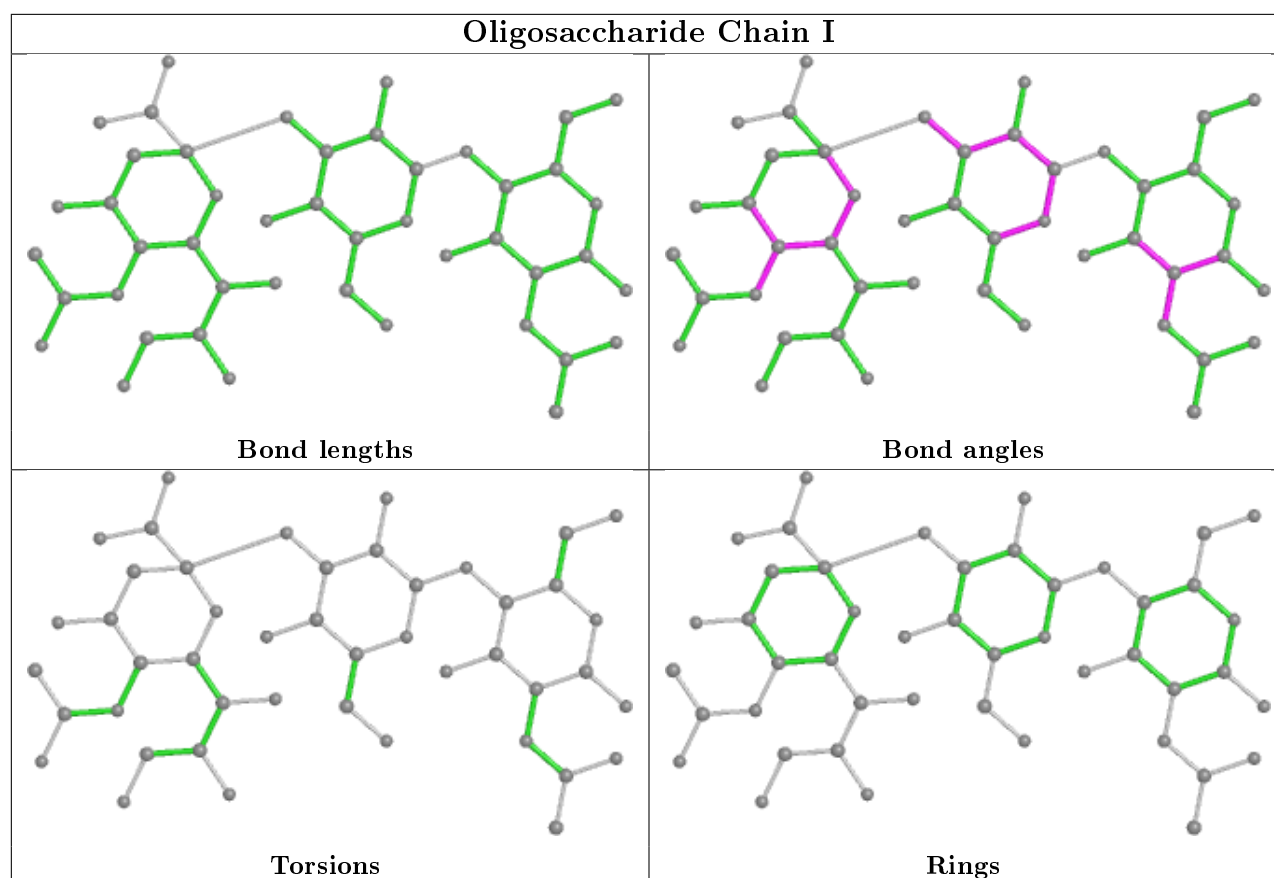
## Oligosaccharide Chain G



## Oligosaccharide Chain H







## 5.6 Ligand geometry ⓘ

10 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	E	1164	1	14,14,15	0.42	0	17,19,21	1.03	1 (5%)
4	NAG	C	1011	1	14,14,15	0.47	0	17,19,21	1.11	2 (11%)
4	NAG	C	1164	1	14,14,15	0.48	0	17,19,21	1.27	3 (17%)
4	NAG	A	1164	1	14,14,15	0.46	0	17,19,21	1.06	1 (5%)
5	PO4	A	1325	-	4,4,4	0.89	0	6,6,6	0.53	0
4	NAG	A	1011	1	14,14,15	0.45	0	17,19,21	1.25	1 (5%)
4	NAG	A	1023	1	14,14,15	0.70	1 (7%)	17,19,21	2.49	6 (35%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	NAG	C	1023	1	14,14,15	0.66	0	17,19,21	2.21	5 (29%)
4	NAG	E	1023	1	14,14,15	0.66	0	17,19,21	2.00	5 (29%)
4	NAG	E	1011	1	14,14,15	0.50	0	17,19,21	1.17	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	E	1164	1	-	2/6/23/26	0/1/1/1
4	NAG	C	1011	1	-	2/6/23/26	0/1/1/1
4	NAG	C	1164	1	-	2/6/23/26	0/1/1/1
4	NAG	A	1164	1	-	0/6/23/26	0/1/1/1
4	NAG	A	1011	1	-	3/6/23/26	0/1/1/1
4	NAG	A	1023	1	1/1/5/7	5/6/23/26	0/1/1/1
4	NAG	C	1023	1	-	3/6/23/26	0/1/1/1
4	NAG	E	1023	1	-	5/6/23/26	0/1/1/1
4	NAG	E	1011	1	-	0/6/23/26	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	1023	NAG	C1-C2	2.03	1.55	1.52

The worst 5 of 25 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	C	1023	NAG	C2-N2-C7	6.62	132.33	122.90
4	A	1023	NAG	C2-N2-C7	6.51	132.17	122.90
4	E	1023	NAG	C2-N2-C7	5.57	130.83	122.90
4	A	1023	NAG	C1-O5-C5	4.04	117.67	112.19
4	A	1011	NAG	C1-O5-C5	3.82	117.37	112.19

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
4	A	1023	NAG	C1

5 of 22 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	1023	NAG	C3-C2-N2-C7
4	C	1023	NAG	C3-C2-N2-C7
4	E	1023	NAG	O5-C5-C6-O6
4	C	1011	NAG	O5-C5-C6-O6
4	A	1011	NAG	C8-C7-N2-C2

There are no ring outliers.

5 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	E	1164	NAG	2	0
4	A	1164	NAG	1	0
4	A	1023	NAG	1	0
4	C	1023	NAG	1	0
4	E	1023	NAG	2	0

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 6 Fit of model and data [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	324/327 (99%)	-0.10	2 (0%) 89 92	27, 42, 57, 84	0
1	C	322/327 (98%)	0.02	3 (0%) 84 88	24, 35, 60, 114	0
1	E	322/327 (98%)	0.06	7 (2%) 62 69	25, 37, 71, 122	0
2	B	163/166 (98%)	0.26	7 (4%) 35 42	23, 51, 80, 99	0
2	D	163/166 (98%)	1.31	41 (25%) 0 0	25, 74, 147, 156	0
2	F	163/166 (98%)	1.59	50 (30%) 0 0	23, 80, 145, 171	0
All	All	1457/1479 (98%)	0.35	110 (7%) 14 19	23, 41, 126, 171	0

The worst 5 of 110 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	F	33	GLY	9.4
2	F	160	PRO	9.0
2	F	29	GLU	7.0
2	F	5	ALA	6.9
2	D	159	TYR	6.8

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

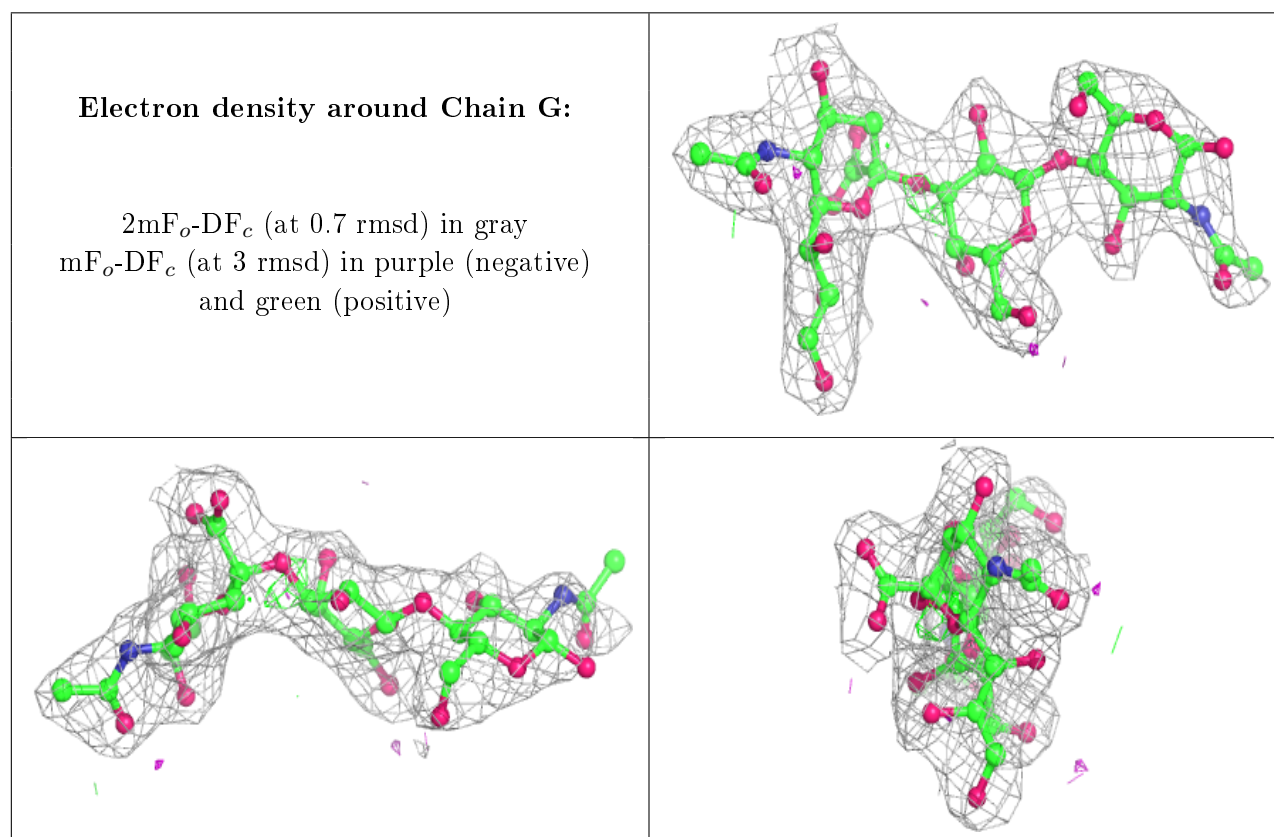
There are no non-standard protein/DNA/RNA residues in this entry.

### 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.

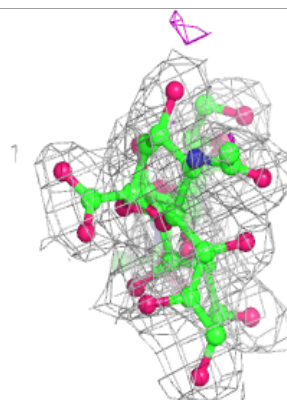
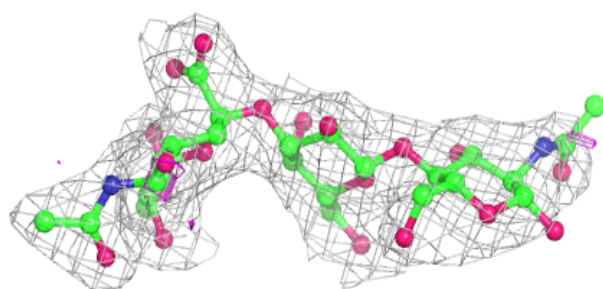
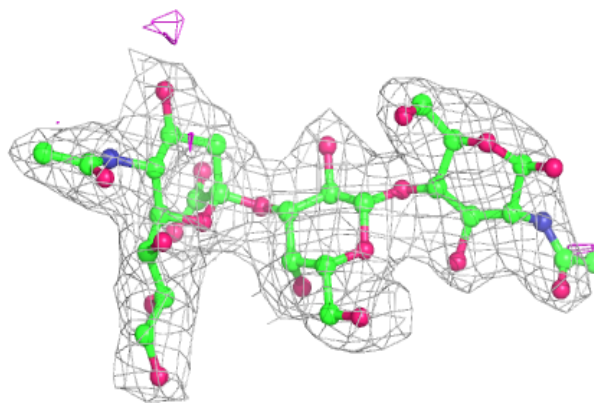
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
3	NAG	I	1	15/15	0.87	0.22	52,60,66,67	0
3	NAG	G	1	15/15	0.91	0.22	59,68,76,77	0
3	NAG	H	1	15/15	0.92	0.14	48,58,67,67	0
3	GAL	I	2	11/12	0.93	0.10	38,43,45,47	0
3	GAL	G	2	11/12	0.94	0.11	44,47,51,53	0
3	SIA	G	3	20/21	0.95	0.12	34,39,43,44	0
3	GAL	H	2	11/12	0.96	0.10	35,37,41,42	0
3	SIA	H	3	20/21	0.96	0.12	27,31,33,35	0
3	SIA	I	3	20/21	0.98	0.14	29,31,36,36	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.

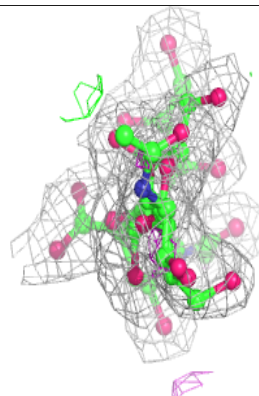
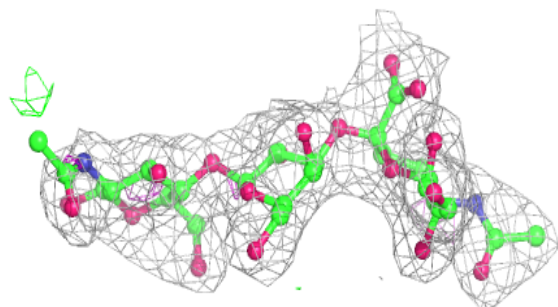
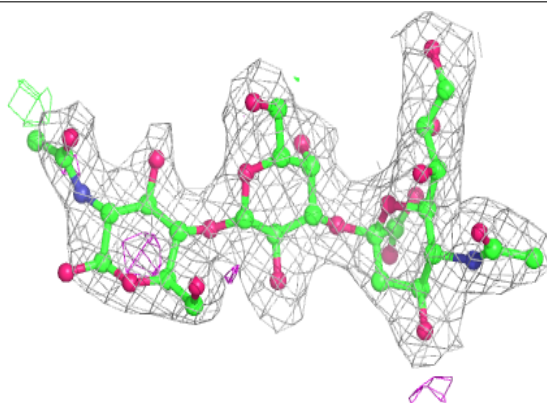


**Electron density around Chain H:**

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

**Electron density around Chain I:**

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



## 6.4 Ligands [i](#)

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	PO4	A	1325	5/5	0.85	0.30	79,82,86,87	0
4	NAG	A	1023	14/15	0.86	0.19	50,61,65,68	0
4	NAG	C	1011	14/15	0.87	0.25	47,54,58,59	0
4	NAG	C	1023	14/15	0.87	0.27	65,78,87,90	0
4	NAG	E	1023	14/15	0.88	0.24	60,70,76,76	0
4	NAG	E	1011	14/15	0.90	0.27	55,64,68,70	0
4	NAG	E	1164	14/15	0.92	0.10	33,38,41,44	0
4	NAG	A	1164	14/15	0.94	0.12	35,40,46,47	0
4	NAG	C	1164	14/15	0.94	0.13	33,39,45,51	0
4	NAG	A	1011	14/15	0.97	0.12	33,40,45,46	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.