



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 25, 2024 – 09:58 PM EDT

PDB ID : 6TWH  
Title : Crystal structure of the haemagglutinin mutant (Gln226Leu, Gly228Ser) from an H10N7 seal influenza virus isolated in Germany  
Authors : Zhang, J.; Xiong, X.; Purkiss, A.; Walker, P.; Gamblin, S.; Skehel, J.J.  
Deposited on : 2020-01-13  
Resolution : 2.68 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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 : 2.37.1  
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) : 2.37.1

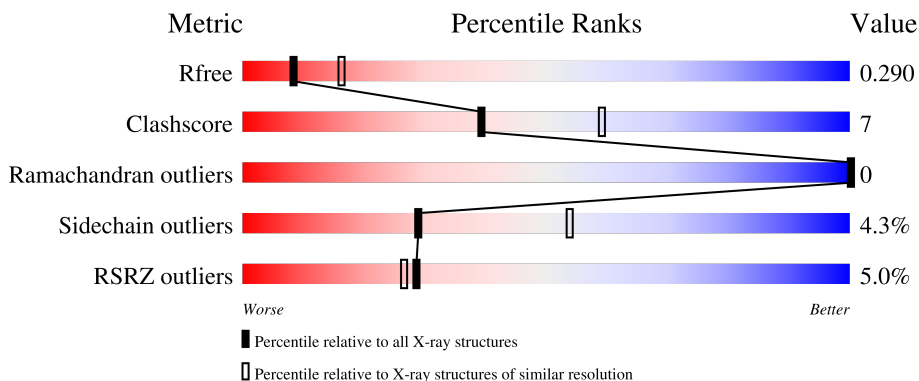
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

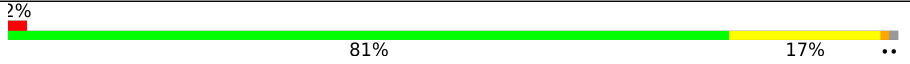
The reported resolution of this entry is 2.68 Å.

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	3863 (2.70-2.66)
Clashscore	141614	4210 (2.70-2.66)
Ramachandran outliers	138981	4141 (2.70-2.66)
Sidechain outliers	138945	4141 (2.70-2.66)
RSRZ outliers	127900	3780 (2.70-2.66)

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 of 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	325	
1	C	325	
1	E	325	
1	L	325	
1	N	325	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	P	325	<p>2% 83% 16%</p>
2	B	177	<p>6% 86% 11%</p>
2	D	177	<p>5% 84% 12%</p>
2	F	177	<p>5% 80% 16%</p>
2	M	177	<p>5% 85% 12%</p>
2	O	177	<p>6% 83% 14%</p>
2	Q	177	<p>6% 84% 13%</p>
3	G	2	<p>100%</p>
3	H	2	<p>100%</p>

## 2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 23138 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 Hemagglutinin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	316	2414	1496	436	466	16	0	0	0
1	C	315	2402	1488	434	464	16	0	0	0
1	E	322	2462	1529	445	472	16	0	0	0
1	L	316	2412	1497	434	465	16	0	0	0
1	N	320	2424	1502	438	468	16	0	0	0
1	P	322	2462	1529	445	472	16	0	0	0

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	ASP	-	expression tag	UNP A0A0A7HR51
A	0	PRO	-	expression tag	UNP A0A0A7HR51
A	221	SER	GLY	engineered mutation	UNP A0A0A7HR51
C	-1	ASP	-	expression tag	UNP A0A0A7HR51
C	0	PRO	-	expression tag	UNP A0A0A7HR51
C	221	SER	GLY	engineered mutation	UNP A0A0A7HR51
E	-1	ASP	-	expression tag	UNP A0A0A7HR51
E	0	PRO	-	expression tag	UNP A0A0A7HR51
E	221	SER	GLY	engineered mutation	UNP A0A0A7HR51
L	-1	ASP	-	expression tag	UNP A0A0A7HR51
L	0	PRO	-	expression tag	UNP A0A0A7HR51
L	221	SER	GLY	engineered mutation	UNP A0A0A7HR51
N	-1	ASP	-	expression tag	UNP A0A0A7HR51
N	0	PRO	-	expression tag	UNP A0A0A7HR51
N	221	SER	GLY	engineered mutation	UNP A0A0A7HR51
P	-1	ASP	-	expression tag	UNP A0A0A7HR51
P	0	PRO	-	expression tag	UNP A0A0A7HR51

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
P	221	SER	GLY	engineered mutation	UNP A0A0A7HR51

- Molecule 2 is a protein called Hemagglutinin HA2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	172	Total	C	N	O	S	0	0	0
			1386	857	241	280	8			
2	D	172	Total	C	N	O	S	0	0	0
			1380	854	238	280	8			
2	F	172	Total	C	N	O	S	0	0	0
			1386	857	241	280	8			
2	M	172	Total	C	N	O	S	0	0	0
			1386	857	241	280	8			
2	O	172	Total	C	N	O	S	0	0	0
			1386	857	241	280	8			
2	Q	172	Total	C	N	O	S	0	0	0
			1386	857	241	280	8			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	177	LYS	-	expression tag	UNP A0A0A7HR51
D	177	LYS	-	expression tag	UNP A0A0A7HR51
F	177	LYS	-	expression tag	UNP A0A0A7HR51
M	177	LYS	-	expression tag	UNP A0A0A7HR51
O	177	LYS	-	expression tag	UNP A0A0A7HR51
Q	177	LYS	-	expression tag	UNP A0A0A7HR51

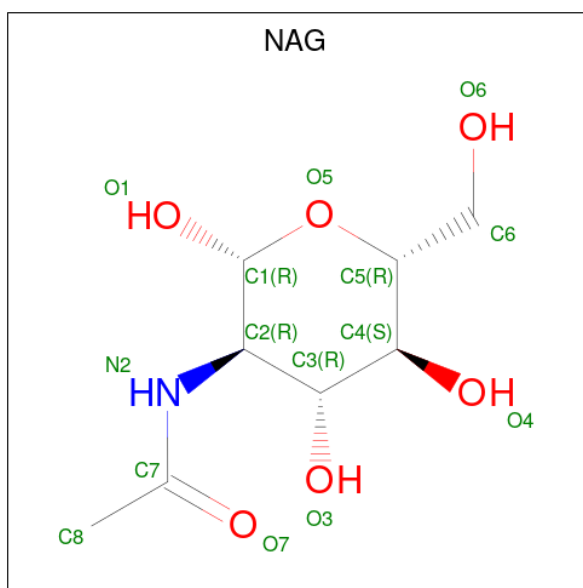
- Molecule 3 is an oligosaccharide called 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	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	C	1	Total	C	N	O	0	0
			14	8	1	5		
4	F	1	Total	C	N	O	0	0
			14	8	1	5		
4	F	1	Total	C	N	O	0	0
			14	8	1	5		
4	M	1	Total	C	N	O	0	0
			14	8	1	5		
4	N	1	Total	C	N	O	0	0
			14	8	1	5		
4	O	1	Total	C	N	O	0	0
			14	8	1	5		
4	P	1	Total	C	N	O	0	0
			14	8	1	5		
4	Q	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 5 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	E	1	Total	Ca	0	0
			1	1		
5	N	1	Total	Ca	0	0
			1	1		
5	P	1	Total	Ca	0	0
			1	1		

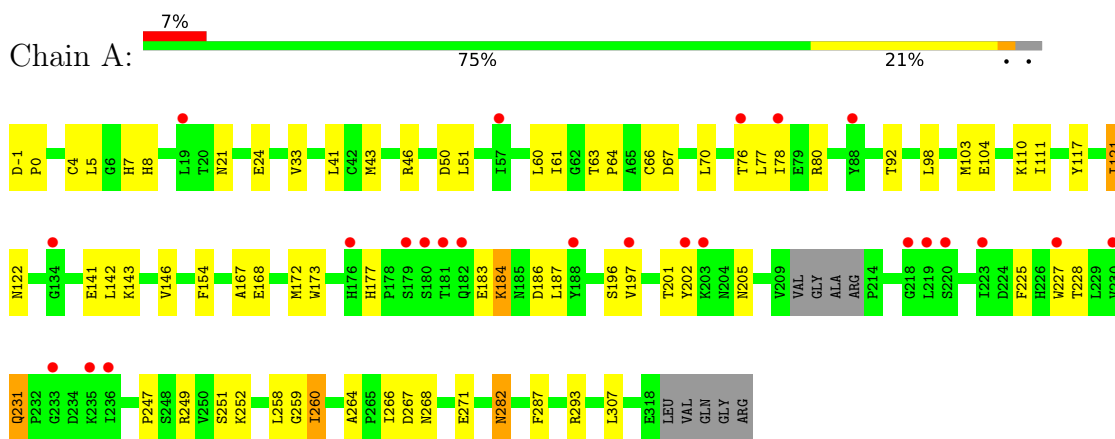
- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	8	Total O 8 8	0	0
6	B	6	Total O 6 6	0	0
6	C	2	Total O 2 2	0	0
6	D	7	Total O 7 7	0	0
6	E	8	Total O 8 8	0	0
6	F	7	Total O 7 7	0	0
6	L	9	Total O 9 9	0	0
6	M	3	Total O 3 3	0	0
6	N	2	Total O 2 2	0	0
6	O	7	Total O 7 7	0	0
6	P	17	Total O 17 17	0	0
6	Q	5	Total O 5 5	0	0

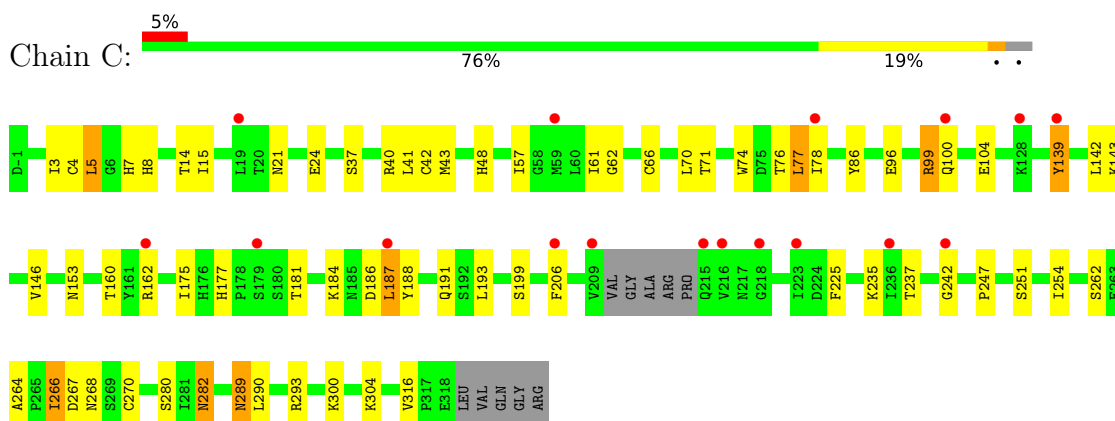
### 3 Residue-property plots [i](#)

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.

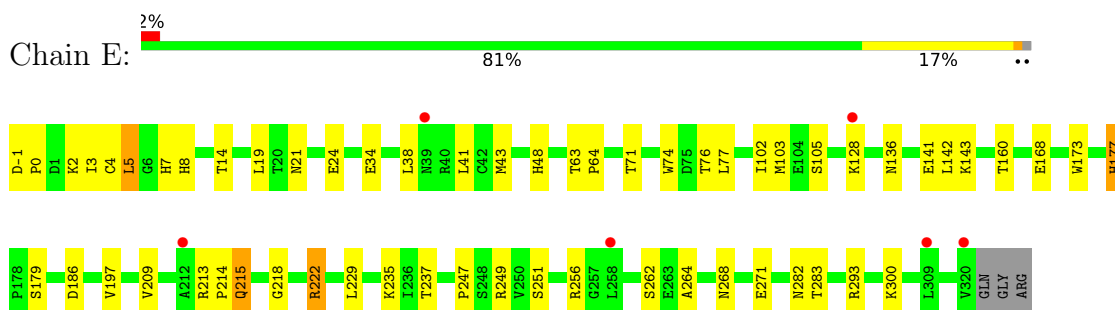
- Molecule 1: Hemagglutinin



- Molecule 1: Hemagglutinin

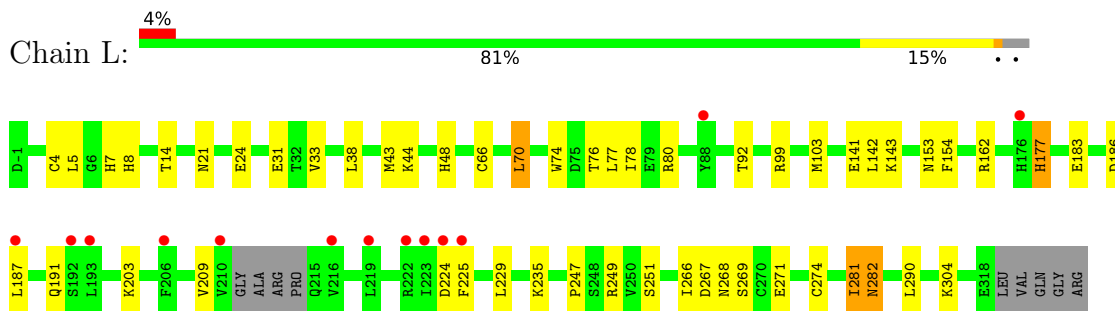


- Molecule 1: Hemagglutinin

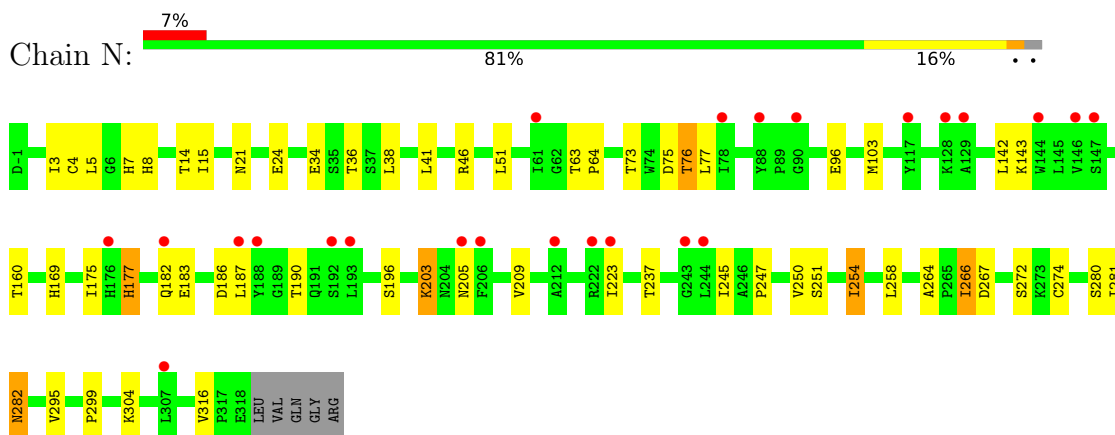




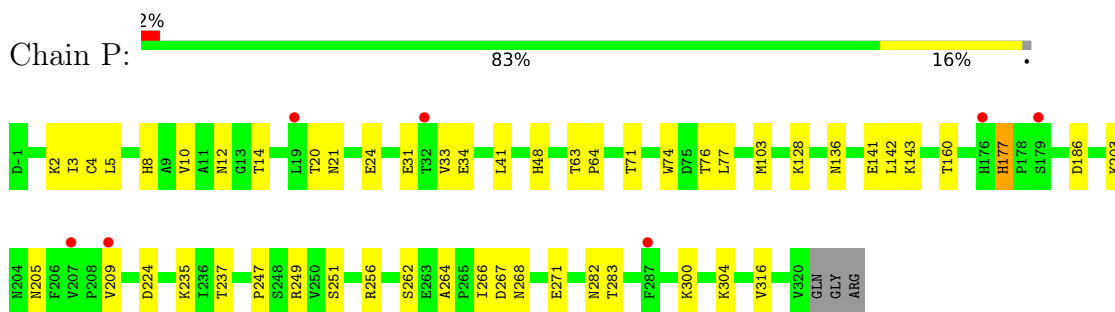
• Molecule 1: Hemagglutinin



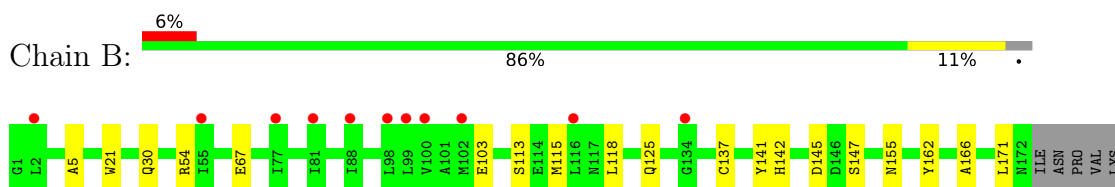
• Molecule 1: Hemagglutinin



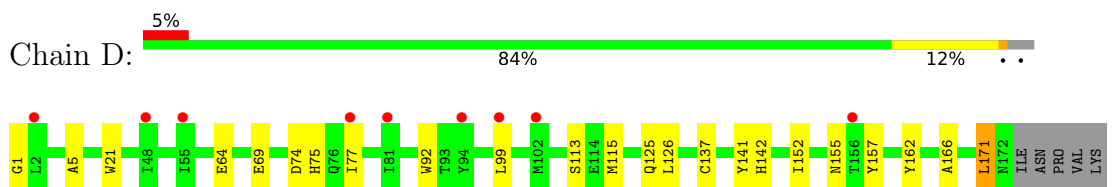
• Molecule 1: Hemagglutinin



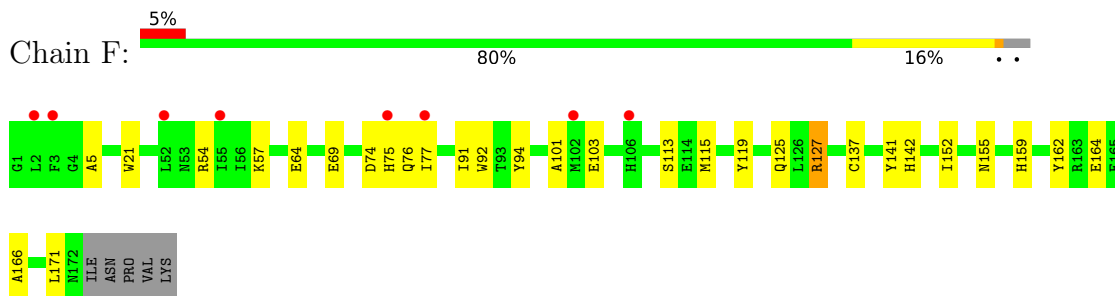
• Molecule 2: Hemagglutinin HA2



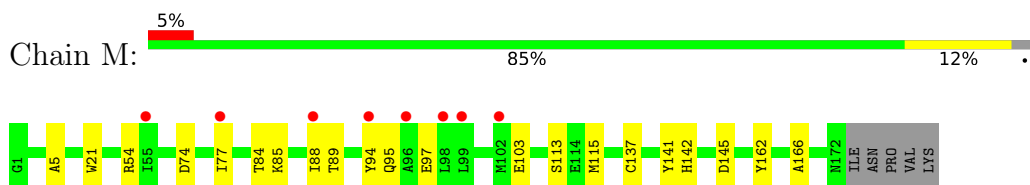
• Molecule 2: Hemagglutinin HA2



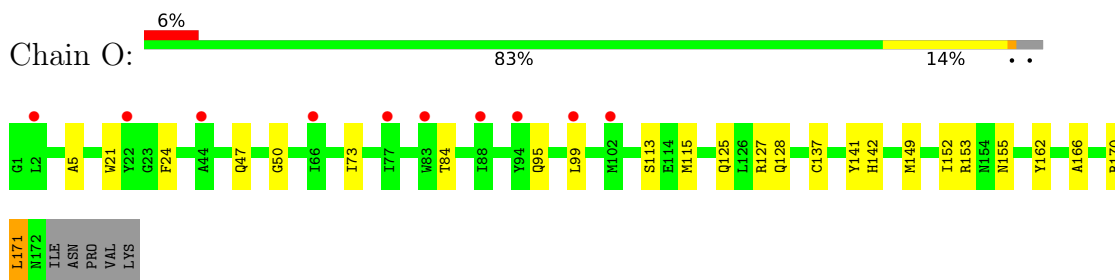
- Molecule 2: Hemagglutinin HA2



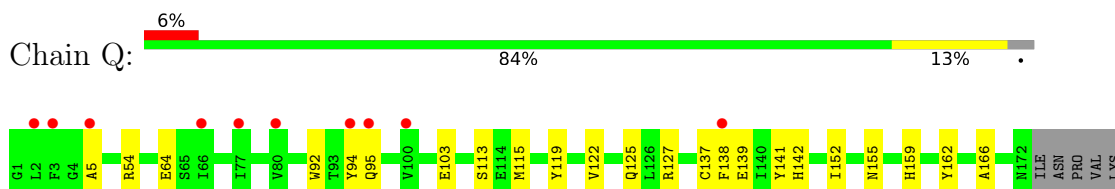
- Molecule 2: Hemagglutinin HA2



- Molecule 2: Hemagglutinin HA2



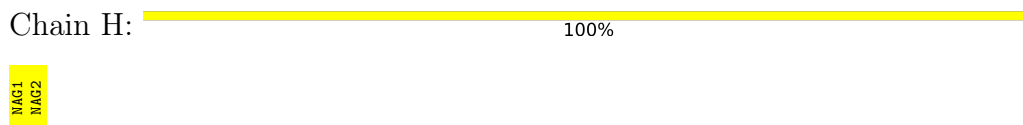
- Molecule 2: Hemagglutinin HA2



- 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 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	69.31Å 213.61Å 156.80Å 90.00° 102.29° 90.00°	Depositor
Resolution (Å)	72.11 – 2.68 72.11 – 2.68	Depositor EDS
% Data completeness (in resolution range)	99.9 (72.11-2.68) 99.9 (72.11-2.68)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.77 (at 2.69Å)	Xtrriage
Refinement program	REFMAC 5.8.0158	Depositor
R, $R_{free}$	0.269 , 0.302 0.260 , 0.290	Depositor DCC
$R_{free}$ test set	6188 reflections (4.97%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	54.0	Xtrriage
Anisotropy	0.714	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 42.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	0.000 for h,-k,-h-l	Xtrriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	23138	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	91.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 30.45 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 1.3055e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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: NAG, CA

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	0.70	0/2462	0.85	0/3334
1	C	0.69	0/2449	0.79	0/3317
1	E	0.66	0/2512	0.79	0/3404
1	L	0.66	0/2460	0.79	0/3333
1	N	0.68	0/2473	0.79	0/3353
1	P	0.66	0/2512	0.79	0/3404
2	B	0.64	0/1411	0.74	0/1903
2	D	0.65	0/1405	0.74	0/1896
2	F	0.66	0/1411	0.75	0/1903
2	M	0.65	0/1411	0.75	0/1903
2	O	0.65	0/1411	0.76	0/1903
2	Q	0.66	0/1411	0.77	0/1903
All	All	0.67	0/23328	0.78	0/31556

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	2414	0	2373	60	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	2402	0	2355	58	0
1	E	2462	0	2430	33	0
1	L	2412	0	2364	38	0
1	N	2424	0	2370	49	0
1	P	2462	0	2429	35	0
2	B	1386	0	1291	12	0
2	D	1380	0	1280	26	0
2	F	1386	0	1290	23	0
2	M	1386	0	1291	15	0
2	O	1386	0	1291	29	0
2	Q	1386	0	1291	20	0
3	G	28	0	25	0	0
3	H	28	0	25	0	0
4	C	14	0	13	0	0
4	F	28	0	26	0	0
4	M	14	0	13	0	0
4	N	14	0	13	0	0
4	O	14	0	13	0	0
4	P	14	0	13	0	0
4	Q	14	0	13	0	0
5	E	1	0	0	0	0
5	N	1	0	0	0	0
5	P	1	0	0	0	0
6	A	8	0	0	0	0
6	B	6	0	0	0	0
6	C	2	0	0	0	0
6	D	7	0	0	2	0
6	E	8	0	0	0	0
6	F	7	0	0	0	0
6	L	9	0	0	0	0
6	M	3	0	0	0	0
6	N	2	0	0	0	0
6	O	7	0	0	0	0
6	P	17	0	0	1	0
6	Q	5	0	0	0	0
All	All	23138	0	22209	337	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 7.

The worst 5 of 337 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:146:VAL:HG12	1:C:187:LEU:HD11	1.36	1.07
1:A:121:ILE:HD11	1:A:154:PHE:CZ	1.91	1.05
2:O:149:MET:O	2:O:152:ILE:HG13	1.62	0.99
1:C:146:VAL:CG1	1:C:187:LEU:HD11	1.96	0.96
1:C:289:ASN:C	1:C:289:ASN:HD22	1.68	0.94

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	312/325 (96%)	309 (99%)	3 (1%)	0	100	100
1	C	311/325 (96%)	308 (99%)	3 (1%)	0	100	100
1	E	320/325 (98%)	316 (99%)	4 (1%)	0	100	100
1	L	312/325 (96%)	308 (99%)	4 (1%)	0	100	100
1	N	318/325 (98%)	314 (99%)	4 (1%)	0	100	100
1	P	320/325 (98%)	316 (99%)	4 (1%)	0	100	100
2	B	170/177 (96%)	161 (95%)	9 (5%)	0	100	100
2	D	170/177 (96%)	162 (95%)	8 (5%)	0	100	100
2	F	170/177 (96%)	160 (94%)	10 (6%)	0	100	100
2	M	170/177 (96%)	161 (95%)	9 (5%)	0	100	100
2	O	170/177 (96%)	161 (95%)	9 (5%)	0	100	100
2	Q	170/177 (96%)	162 (95%)	8 (5%)	0	100	100
All	All	2913/3012 (97%)	2838 (97%)	75 (3%)	0	100	100

There are no Ramachandran outliers to report.

### 5.3.2 Protein sidechains [i](#)

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	269/276 (98%)	256 (95%)	13 (5%)	25	49
1	C	267/276 (97%)	244 (91%)	23 (9%)	10	22
1	E	274/276 (99%)	260 (95%)	14 (5%)	24	46
1	L	268/276 (97%)	253 (94%)	15 (6%)	21	42
1	N	267/276 (97%)	251 (94%)	16 (6%)	19	39
1	P	274/276 (99%)	261 (95%)	13 (5%)	26	50
2	B	146/151 (97%)	144 (99%)	2 (1%)	67	85
2	D	145/151 (96%)	143 (99%)	2 (1%)	67	85
2	F	146/151 (97%)	142 (97%)	4 (3%)	44	71
2	M	146/151 (97%)	144 (99%)	2 (1%)	67	85
2	O	146/151 (97%)	144 (99%)	2 (1%)	67	85
2	Q	146/151 (97%)	145 (99%)	1 (1%)	84	93
All	All	2494/2562 (97%)	2387 (96%)	107 (4%)	29	54

5 of 107 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	F	164	GLU
1	L	281	ILE
1	P	143	LYS
1	L	14	THR
1	L	177	HIS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 30 such sidechains are listed below:

Mol	Chain	Res	Type
1	E	240	HIS
2	Q	12	ASN
1	L	215	GLN
2	Q	161	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	P	48	HIS

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates [i](#)

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	G	1	5,3,2	14,14,15	0.59	0	17,19,21	1.26	1 (5%)
3	NAG	G	2	3	14,14,15	0.46	0	17,19,21	1.49	2 (11%)
3	NAG	H	1	3,2	14,14,15	0.43	0	17,19,21	0.95	1 (5%)
3	NAG	H	2	3	14,14,15	0.53	0	17,19,21	1.41	3 (17%)

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	5,3,2	-	2/6/23/26	0/1/1/1
3	NAG	G	2	3	-	2/6/23/26	0/1/1/1
3	NAG	H	1	3,2	-	2/6/23/26	0/1/1/1
3	NAG	H	2	3	-	0/6/23/26	0/1/1/1



There are no bond length outliers.

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	G	2	NAG	C1-O5-C5	3.61	117.08	112.19
3	H	2	NAG	O5-C5-C6	3.60	112.84	107.20
3	G	2	NAG	O5-C1-C2	-3.38	105.95	111.29
3	G	1	NAG	C1-O5-C5	2.94	116.17	112.19
3	H	2	NAG	O5-C5-C4	-2.76	104.10	110.83

There are no chirality outliers.

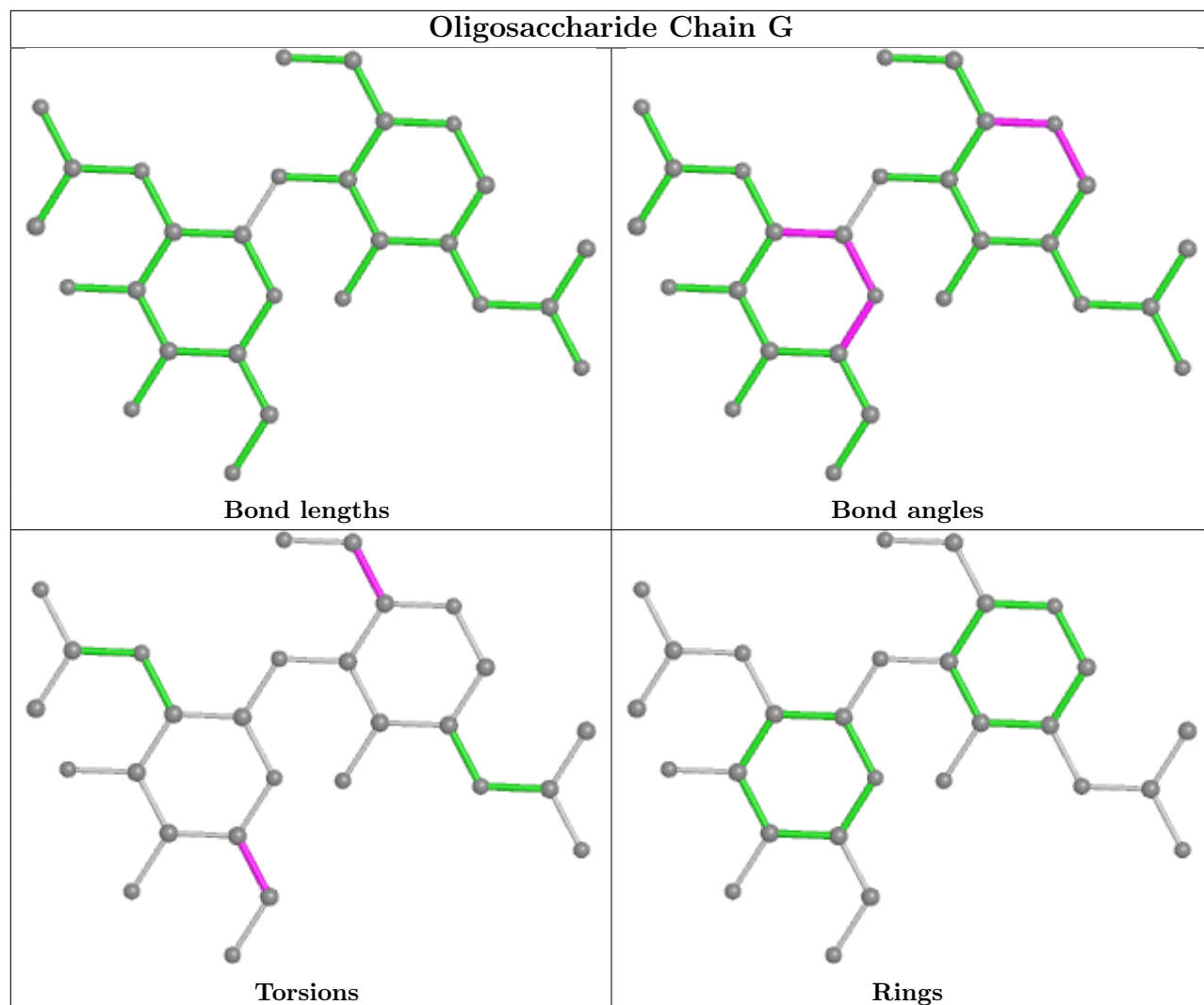
5 of 6 torsion outliers are listed below:

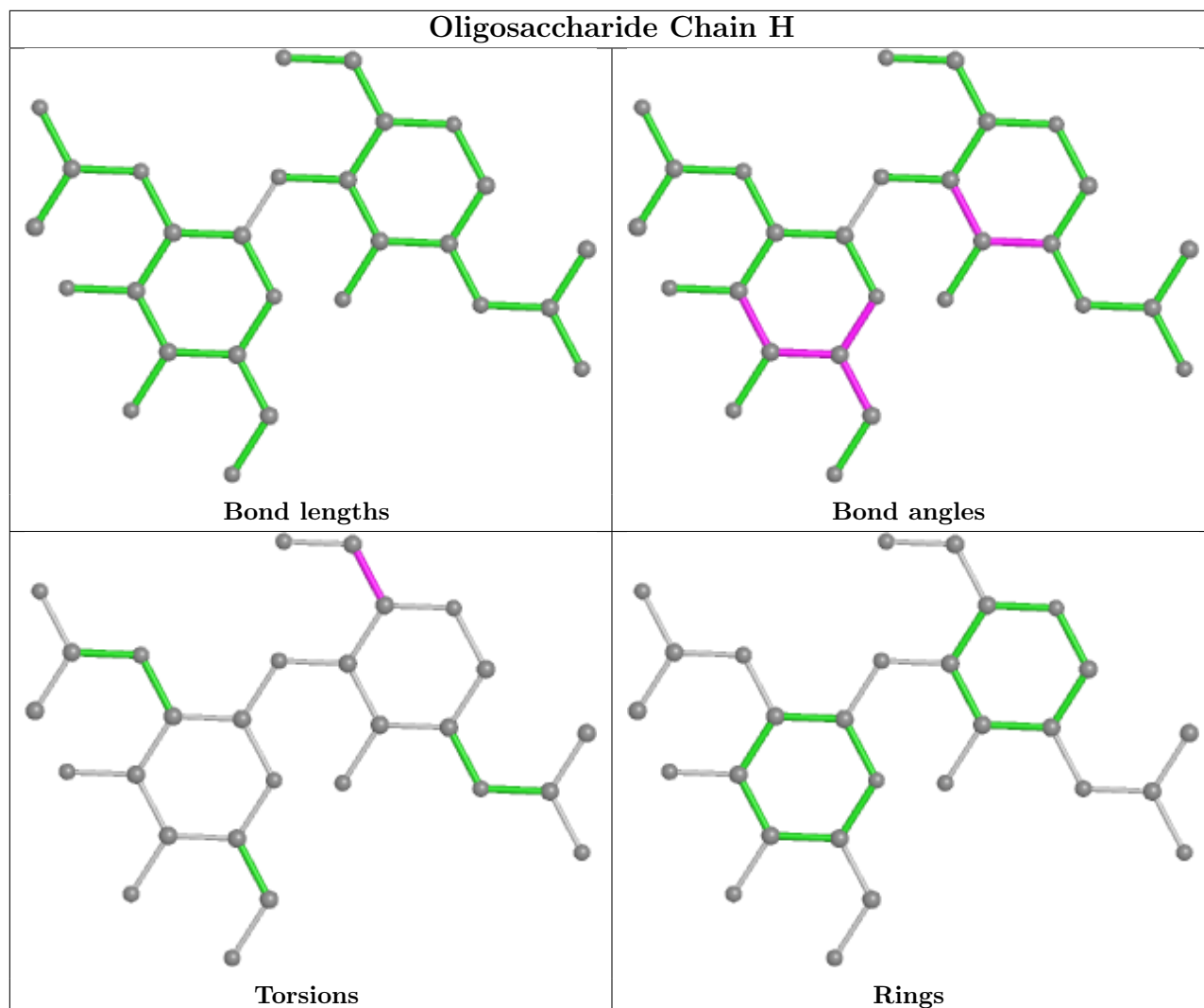
Mol	Chain	Res	Type	Atoms
3	G	2	NAG	C4-C5-C6-O6
3	G	2	NAG	O5-C5-C6-O6
3	H	1	NAG	O5-C5-C6-O6
3	H	1	NAG	C4-C5-C6-O6
3	G	1	NAG	C4-C5-C6-O6

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](#)

Of 11 ligands modelled in this entry, 3 are monoatomic - leaving 8 for Mogul analysis.

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	F	201	2	14,14,15	0.56	0	17,19,21	1.36	3 (17%)
4	NAG	Q	201	5,2	14,14,15	0.46	0	17,19,21	1.31	2 (11%)
4	NAG	F	202	2	14,14,15	0.55	0	17,19,21	1.04	1 (5%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	NAG	C	401	1	14,14,15	0.64	0	17,19,21	0.91	1 (5%)
4	NAG	M	201	5,2	14,14,15	0.95	0	17,19,21	1.70	1 (5%)
4	NAG	N	401	1	14,14,15	0.71	0	17,19,21	1.23	1 (5%)
4	NAG	P	401	1	14,14,15	1.04	1 (7%)	17,19,21	1.16	1 (5%)
4	NAG	O	201	2	14,14,15	0.62	0	17,19,21	1.44	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
4	NAG	F	201	2	-	0/6/23/26	0/1/1/1
4	NAG	Q	201	5,2	-	0/6/23/26	0/1/1/1
4	NAG	F	202	2	-	1/6/23/26	0/1/1/1
4	NAG	C	401	1	-	0/6/23/26	0/1/1/1
4	NAG	M	201	5,2	-	1/6/23/26	0/1/1/1
4	NAG	N	401	1	-	3/6/23/26	0/1/1/1
4	NAG	P	401	1	-	2/6/23/26	0/1/1/1
4	NAG	O	201	2	-	2/6/23/26	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	P	401	NAG	C1-C2	2.94	1.56	1.52

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	M	201	NAG	C1-O5-C5	5.49	119.63	112.19
4	O	201	NAG	C1-O5-C5	4.15	117.81	112.19
4	N	401	NAG	C2-N2-C7	3.26	127.55	122.90
4	Q	201	NAG	C1-O5-C5	3.07	116.35	112.19
4	F	202	NAG	C1-O5-C5	2.83	116.03	112.19

There are no chirality outliers.

5 of 9 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	N	401	NAG	C3-C2-N2-C7

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms
4	N	401	NAG	C4-C5-C6-O6
4	N	401	NAG	O5-C5-C6-O6
4	O	201	NAG	C4-C5-C6-O6
4	F	202	NAG	O5-C5-C6-O6

There are no ring outliers.

No monomer is involved in short contacts.

## 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	316/325 (97%)	0.54	24 (7%) 13 11	54, 105, 163, 213	0
1	C	315/325 (96%)	0.47	17 (5%) 25 24	64, 102, 153, 198	0
1	E	322/325 (99%)	0.32	6 (1%) 66 67	55, 79, 113, 152	0
1	L	316/325 (97%)	0.41	13 (4%) 37 35	52, 93, 142, 175	0
1	N	320/325 (98%)	0.57	24 (7%) 14 12	63, 112, 174, 241	0
1	P	322/325 (99%)	0.33	7 (2%) 62 61	54, 75, 110, 156	0
2	B	172/177 (97%)	0.51	11 (6%) 19 17	48, 75, 98, 117	0
2	D	172/177 (97%)	0.54	9 (5%) 27 25	60, 84, 109, 119	0
2	F	172/177 (97%)	0.45	8 (4%) 31 29	53, 78, 107, 132	0
2	M	172/177 (97%)	0.48	8 (4%) 31 29	48, 73, 97, 121	0
2	O	172/177 (97%)	0.54	10 (5%) 23 21	63, 86, 114, 137	0
2	Q	172/177 (97%)	0.48	10 (5%) 23 21	50, 75, 104, 114	0
All	All	2943/3012 (97%)	0.46	147 (4%) 28 26	48, 87, 146, 241	0

The worst 5 of 147 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	N	146	VAL	7.3
1	N	182	GLN	7.2
1	N	188	TYR	4.8
1	N	129	ALA	4.6
1	A	134	GLY	4.5

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

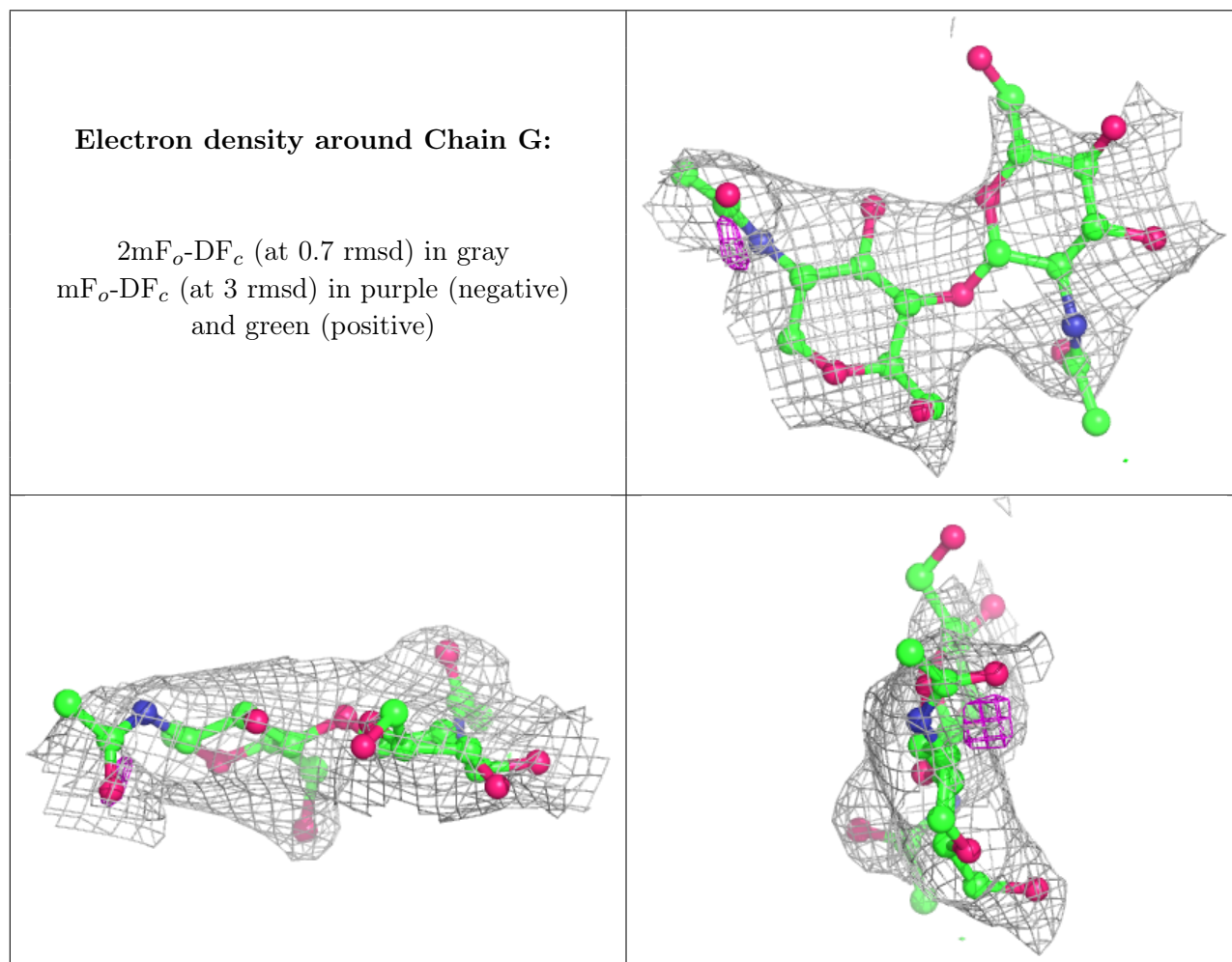
There are no non-standard protein/DNA/RNA residues in this entry.

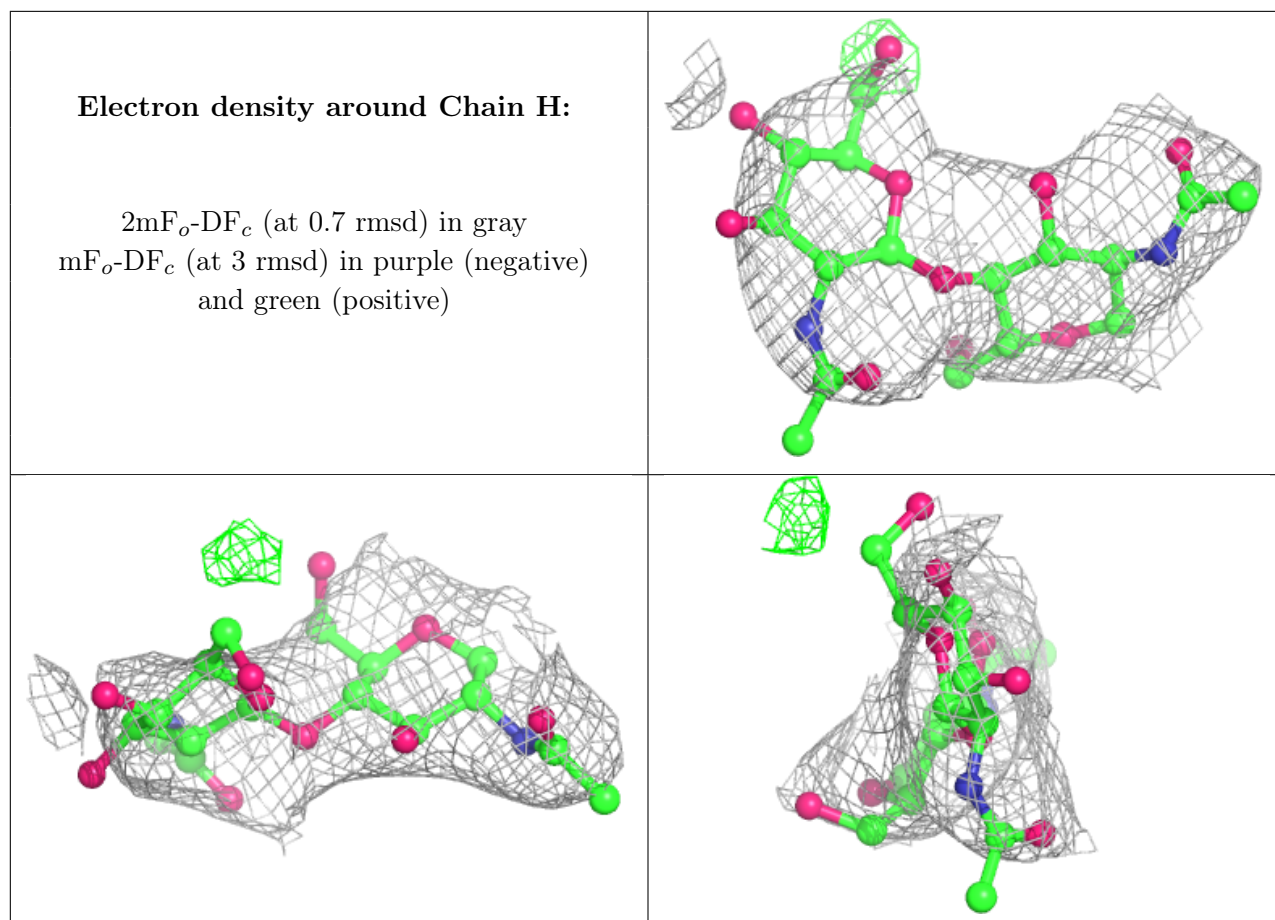
### 6.3 Carbohydrates [i](#)

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	H	2	14/15	0.83	0.24	120,145,154,155	0
3	NAG	G	2	14/15	0.84	0.29	98,125,136,137	0
3	NAG	H	1	14/15	0.88	0.13	107,118,126,130	0
3	NAG	G	1	14/15	0.95	0.15	84,90,98,102	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 [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
4	NAG	O	201	14/15	0.72	0.17	110,127,130,132	0
4	NAG	C	401	14/15	0.78	0.18	113,139,147,149	0
4	NAG	P	401	14/15	0.80	0.17	88,117,136,141	0
4	NAG	N	401	14/15	0.81	0.13	122,140,152,163	0
4	NAG	Q	201	14/15	0.82	0.17	103,108,113,114	0
4	NAG	F	201	14/15	0.86	0.19	97,106,114,119	0
4	NAG	F	202	14/15	0.87	0.13	119,145,153,154	0
4	NAG	M	201	14/15	0.91	0.14	70,78,84,86	0
5	CA	P	402	1/1	0.96	0.15	67,67,67,67	0
5	CA	N	402	1/1	0.97	0.04	106,106,106,106	0
5	CA	E	401	1/1	0.97	0.12	69,69,69,69	0



## 6.5 Other polymers [i](#)

There are no such residues in this entry.