



# Full wwPDB X-ray Structure Validation Report

Feb 28, 2014 – 09:41 AM GMT

PDB ID : 4MC5  
Title : Crystal structure of a subtype H18 hemagglutinin homologue from A/flat-faced bat/Peru/033/2010 (H18N11)  
Authors : Yang, H.; Carney, P.J.; Chang, J.C.; Guo, Z.; Stevens, J.  
Deposited on : 2013-08-21  
Resolution : 2.24 Å(reported)

This is a full wwPDB validation report for a publicly released PDB entry.  
We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)  
A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

---

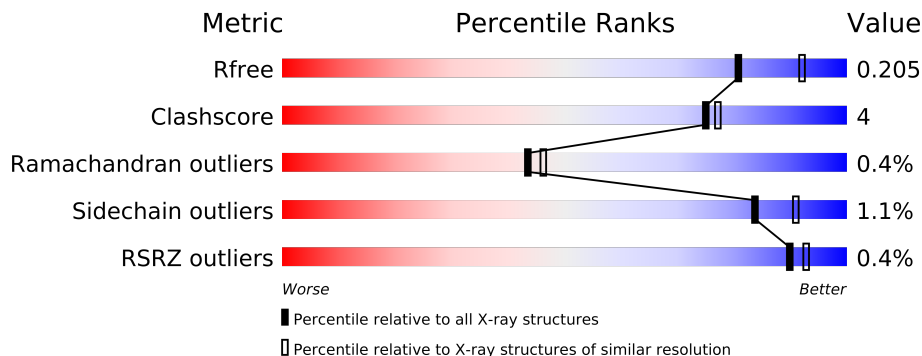
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.15 2013  
Xtriage (Phenix) : dev-1323  
EDS : stable22639  
Percentile statistics : 21963  
Refmac : 5.8.0049  
CCP4 : 6.3.0 (Settle)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : stable22683

# 1 Overall quality at a glance

The reported resolution of this entry is 2.24 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))
$R_{free}$	66092	1112 (2.26-2.22)
Clashscore	79885	1317 (2.26-2.22)
Ramachandran outliers	78287	1282 (2.26-2.22)
Sidechain outliers	78261	1282 (2.26-2.22)
RSRZ outliers	66119	1112 (2.26-2.22)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	510	
1	B	510	
1	C	510	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
5	FUC	A	614	-	X
5	FUC	B	610	-	X
5	FUC	C	608	-	X
6	NAG	B	611	-	X
6	NAG	B	613	-	X

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Geometry	Electron density
6	NAG	C	609	-	X
6	NAG	C	610	-	X

## 2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 13219 atoms, of which 0 are hydrogen and 0 are deuterium.

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
1	A	494	Total	C	N	O	S	0	0	0
			3914	2459	660	777	18			
1	B	494	Total	C	N	O	S	0	0	0
			3914	2459	660	777	18			
1	C	497	Total	C	N	O	S	0	0	0
			3938	2476	664	780	18			

- Molecule 2 is a polymer of unknown type called SUGAR (4-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	4	Total	C	N	O	0	0
			50	28	2	20		

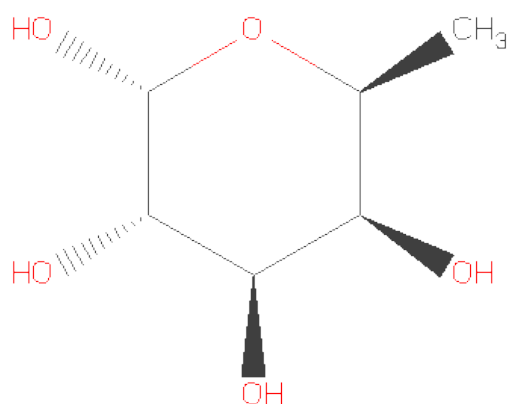
- Molecule 3 is a polymer of unknown type called SUGAR (7-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	7	Total	C	N	O	0	0
			81	46	2	33		
3	B	7	Total	C	N	O	0	0
			81	46	2	33		

- Molecule 4 is a polymer of unknown type called SUGAR (2-MER).

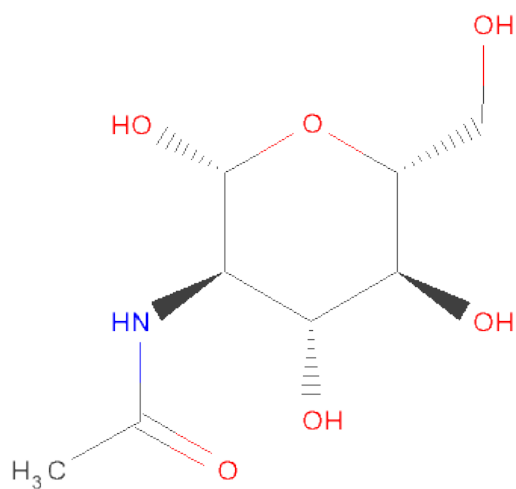
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	2	Total	C	N	O	0	0
			28	16	2	10		
4	B	2	Total	C	N	O	0	0
			28	16	2	10		
4	C	2	Total	C	N	O	0	0
			28	16	2	10		

- Molecule 5 is SUGAR (ALPHA-L-FUCOSE) (three-letter code: FUC) (formula: C<sub>6</sub>H<sub>12</sub>O<sub>5</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			10	6	4		
5	B	1	Total	C	O	0	0
			10	6	4		
5	C	1	Total	C	O	0	0
			10	6	4		

- Molecule 6 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



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

- Molecule 7 is a polymer of unknown type called SUGAR (4-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
7	C	4	Total	C	N	O	0	0
			48	28	2	18		

- Molecule 8 is water.

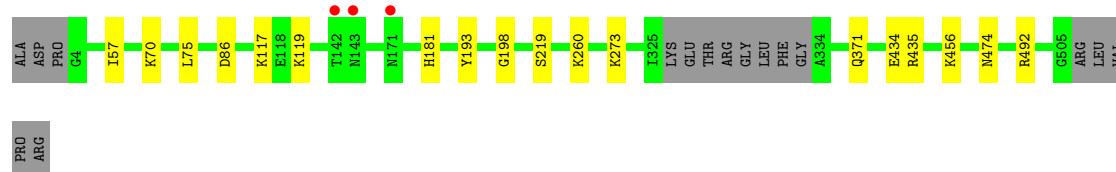
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	292	Total	O	0	0
			292	292		
8	B	343	Total	O	0	0
			343	343		
8	C	332	Total	O	0	0
			332	332		

### 3 Residue-property plots i

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

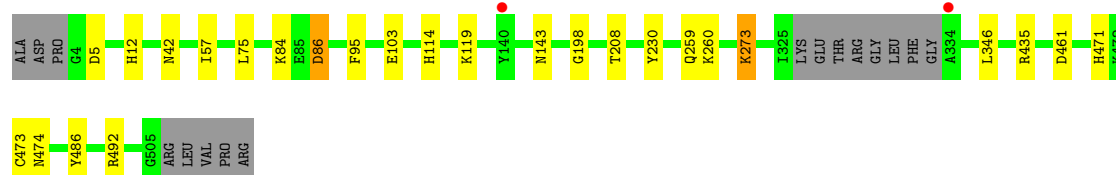
#### • Molecule 1: Hemagglutinin

Chain A:



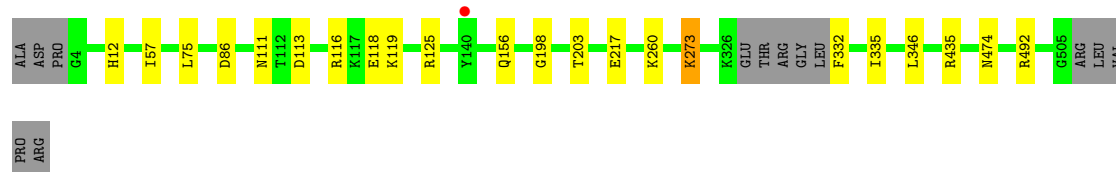
#### • Molecule 1: Hemagglutinin

Chain B:



#### • Molecule 1: Hemagglutinin

Chain C:



## 4 Data and refinement statistics

Property	Value	Source
Space group	I 4 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	239.09Å 239.09Å 161.23Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.02 – 2.24 48.02 – 2.24	Depositor EDS
% Data completeness (in resolution range)	99.8 (48.02-2.24) 95.7 (48.02-2.24)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.79 (at 2.24Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.1069)	Depositor
R, $R_{free}$	0.175 , 0.204 0.176 , 0.205	Depositor DCC
$R_{free}$ test set	5316 reflections (4.99%)	DCC
Wilson B-factor (Å <sup>2</sup> )	32.2	Xtriage
Anisotropy	0.166	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 39.7	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 111081 reflections	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	13219	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	40.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.47% of the height of the origin peak. No significant pseudotranslation is detected.*

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

## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: MAN, FUL, BMA, NAG, FUC

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	0.30	0/3998	0.46	0/5425
1	B	0.31	0/3998	0.47	0/5425
1	C	0.30	0/4023	0.46	0/5457
All	All	0.30	0/12019	0.46	0/16307

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3914	0	0	13	0
1	B	3914	0	0	18	1
1	C	3938	0	0	15	0
2	A	50	0	0	0	0
3	A	81	0	0	1	1
3	B	81	0	0	0	1
4	A	28	0	0	0	0
4	B	28	0	0	0	0
4	C	28	0	0	3	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	A	10	0	0	0	0
5	B	10	0	0	0	0
5	C	10	0	0	0	0
6	A	28	0	0	0	0
6	B	42	0	0	1	0
6	C	42	0	0	0	0
7	C	48	0	0	0	0
8	A	292	0	0	10	0
8	B	343	0	0	12	1
8	C	332	0	0	14	0
All	All	13219	0	0	49	2

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 4.

All (49) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:C:125:ARG:NH2	8:C:980:HOH:O	2.01	0.91
4:C:607:NAG:O4	8:C:997:HOH:O	1.91	0.87
1:C:113:ASP:OD2	8:C:927:HOH:O	1.96	0.82
1:C:474:ASN:OD1	8:C:811:HOH:O	1.99	0.80
1:A:474:ASN:OD1	8:A:812:HOH:O	2.01	0.78
1:B:119:LYS:NZ	8:B:811:HOH:O	2.15	0.77
1:A:117:LYS:NZ	8:A:807:HOH:O	2.18	0.76
1:C:217:GLU:OE1	8:C:897:HOH:O	2.04	0.76
1:C:111:ASN:ND2	8:C:903:HOH:O	2.17	0.75
1:C:119:LYS:NZ	8:C:736:HOH:O	2.19	0.75
1:A:371:GLN:OE1	8:A:899:HOH:O	2.04	0.75
1:B:208:THR:OG1	8:B:770:HOH:O	2.06	0.73
1:A:119:LYS:NZ	8:A:712:HOH:O	2.22	0.73
3:A:607:BMA:O2	8:A:896:HOH:O	2.07	0.72
4:C:606:NAG:O3	8:C:981:HOH:O	2.07	0.71
1:A:219:SER:OG	8:A:856:HOH:O	2.09	0.69
1:C:203:THR:OG1	8:C:784:HOH:O	2.10	0.68
1:B:260:LYS:NZ	8:B:1024:HOH:O	2.26	0.67
1:B:84:LYS:NZ	8:B:718:HOH:O	2.28	0.67
1:B:86:ASP:OD2	8:B:785:HOH:O	2.15	0.65
1:C:273:LYS:NZ	8:C:970:HOH:O	2.31	0.64
1:A:260:LYS:NZ	8:A:804:HOH:O	2.32	0.61
4:C:607:NAG:O3	8:C:850:HOH:O	2.18	0.56
1:B:103:GLU:OE2	8:B:842:HOH:O	2.18	0.55
1:A:181:HIS:ND1	1:A:193:TYR:OH	2.40	0.54

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:70:LYS:NZ	8:A:945:HOH:O	2.41	0.53
1:A:435:ARG:NH2	8:A:834:HOH:O	2.41	0.52
1:B:435:ARG:NE	8:B:739:HOH:O	2.42	0.52
1:C:12:HIS:ND1	1:C:346:LEU:O	2.44	0.49
1:C:116:ARG:NH1	1:C:118:GLU:OE1	2.47	0.48
1:B:5:ASP:OD2	1:B:473:CYS:N	2.48	0.47
6:B:611:NAG:O4	8:B:851:HOH:O	2.21	0.47
1:B:474:ASN:N	1:B:474:ASN:OD1	2.48	0.46
1:A:435:ARG:NE	8:A:834:HOH:O	2.48	0.46
1:C:156:GLN:NE2	8:C:860:HOH:O	2.47	0.46
1:C:260:LYS:NZ	8:C:909:HOH:O	2.48	0.46
1:A:456:LYS:NZ	1:B:461:ASP:O	2.49	0.46
1:B:42:ASN:N	8:B:750:HOH:O	2.48	0.45
1:A:474:ASN:OD1	1:A:474:ASN:N	2.50	0.45
1:B:435:ARG:NH2	8:B:739:HOH:O	2.50	0.43
1:B:471:HIS:NE2	1:B:486:TYR:OH	2.51	0.43
1:C:113:ASP:O	8:C:941:HOH:O	2.21	0.42
1:C:474:ASN:N	1:C:474:ASN:OD1	2.52	0.42
1:B:12:HIS:ND1	1:B:346:LEU:O	2.53	0.42
1:B:95:PHE:O	1:B:230:TYR:OH	2.38	0.41
1:B:5:ASP:N	8:B:954:HOH:O	2.54	0.41
1:B:114:HIS:ND1	1:B:259:GLN:OE1	2.54	0.40
1:B:143:ASN:N	8:B:865:HOH:O	2.53	0.40
1:A:434:GLU:OE2	1:C:435:ARG:NH2	2.53	0.40

All (2) 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	Distance(Å)	Clash(Å)
3:A:608:MAN:O4	3:B:601:NAG:O7[3_755]	2.16	0.04
1:B:273:LYS:NZ	8:B:969:HOH:O[6_575]	2.18	0.02

## 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	490/510 (96%)	480 (98%)	8 (2%)	2 (0%)	43	46
1	B	490/510 (96%)	479 (98%)	9 (2%)	2 (0%)	43	46
1	C	493/510 (97%)	482 (98%)	9 (2%)	2 (0%)	43	46
All	All	1473/1530 (96%)	1441 (98%)	26 (2%)	6 (0%)	43	46

All (6) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	57	ILE
1	A	198	GLY
1	B	57	ILE
1	C	57	ILE
1	C	198	GLY
1	B	198	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	437/450 (97%)	433 (99%)	4 (1%)	87	93
1	B	437/450 (97%)	433 (99%)	4 (1%)	87	93
1	C	439/450 (98%)	433 (99%)	6 (1%)	78	88
All	All	1313/1350 (97%)	1299 (99%)	14 (1%)	84	91

All (14) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	75	LEU
1	A	86	ASP
1	A	273	LYS
1	A	492	ARG
1	B	75	LEU
1	B	86	ASP
1	B	273	LYS
1	B	492	ARG
1	C	75	LEU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	C	86	ASP
1	C	273	LYS
1	C	332	PHE
1	C	335	ILE
1	C	492	ARG

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 chains 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 ⓘ

28 carbohydrates 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	A	601	1,2	12,14,15	0.60	0	15,19,21	1.17	2 (13%)
2	NAG	A	602	2	12,14,15	0.65	0	15,19,21	1.08	1 (6%)
2	BMA	A	603	2	10,11,12	0.82	1 (10%)	11,15,17	1.12	2 (18%)
2	MAN	A	604	2	10,11,12	0.70	0	11,15,17	0.68	0
3	NAG	A	605	1,3	12,14,15	0.64	0	15,19,21	1.31	1 (6%)
3	NAG	A	606	3	12,14,15	0.62	0	15,19,21	1.05	1 (6%)
3	BMA	A	607	3	10,11,12	0.78	0	11,15,17	1.41	1 (9%)
3	MAN	A	608	3	10,11,12	0.84	1 (10%)	11,15,17	1.13	1 (9%)
3	FUC	A	609	3	9,10,11	0.64	0	10,14,16	0.64	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	FUL	A	610	3	9,10,11	0.96	1 (11%)	10,14,16	1.00	1 (10%)
3	MAN	A	611	3	10,11,12	0.70	0	11,15,17	1.23	1 (9%)
4	NAG	A	612	1,4	12,14,15	0.59	0	15,19,21	1.05	2 (13%)
4	NAG	A	613	4	12,14,15	0.54	0	15,19,21	1.88	2 (13%)
3	NAG	B	601	1,3	12,14,15	0.68	0	15,19,21	1.65	3 (20%)
3	NAG	B	602	3	12,14,15	0.65	0	15,19,21	1.15	2 (13%)
3	BMA	B	603	3	10,11,12	0.79	0	11,15,17	1.38	1 (9%)
3	MAN	B	604	3	10,11,12	0.71	0	11,15,17	1.33	2 (18%)
3	FUC	B	605	3	9,10,11	0.64	0	10,14,16	0.50	0
3	FUL	B	606	3	9,10,11	0.84	0	10,14,16	1.06	1 (10%)
3	MAN	B	607	3	10,11,12	0.63	0	11,15,17	2.00	4 (36%)
4	NAG	B	608	1,4	12,14,15	0.64	0	15,19,21	1.05	0
4	NAG	B	609	4	12,14,15	0.51	0	15,19,21	2.01	3 (20%)
7	NAG	C	602	1,7	12,14,15	0.53	0	15,19,21	1.61	2 (13%)
7	NAG	C	603	7	12,14,15	0.65	0	15,19,21	1.20	2 (13%)
7	FUC	C	604	7	9,10,11	1.07	1 (11%)	10,14,16	1.34	1 (10%)
7	FUL	C	605	7	9,10,11	0.78	0	10,14,16	0.85	0
4	NAG	C	606	1,4	12,14,15	0.59	0	15,19,21	0.79	0
4	NAG	C	607	4	12,14,15	0.66	0	15,19,21	0.83	0

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	A	601	1,2	-	0/6/23/26	0/1/1/1
2	NAG	A	602	2	-	0/6/23/26	0/1/1/1
2	BMA	A	603	2	-	0/2/19/22	0/1/1/1
2	MAN	A	604	2	-	0/2/19/22	0/1/1/1
3	NAG	A	605	1,3	-	0/6/23/26	0/1/1/1
3	NAG	A	606	3	-	0/6/23/26	0/1/1/1
3	BMA	A	607	3	-	0/2/19/22	0/1/1/1
3	MAN	A	608	3	-	0/2/19/22	0/1/1/1
3	FUC	A	609	3	-	0/0/17/20	0/1/1/1
3	FUL	A	610	3	-	0/0/17/20	0/1/1/1
3	MAN	A	611	3	-	0/2/19/22	0/1/1/1
4	NAG	A	612	1,4	-	0/6/23/26	0/1/1/1
4	NAG	A	613	4	-	0/6/23/26	0/1/1/1

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	NAG	B	601	1,3	-	0/6/23/26	0/1/1/1
3	NAG	B	602	3	-	0/6/23/26	0/1/1/1
3	BMA	B	603	3	-	0/2/19/22	0/1/1/1
3	MAN	B	604	3	-	0/2/19/22	0/1/1/1
3	FUC	B	605	3	-	0/0/17/20	0/1/1/1
3	FUL	B	606	3	-	0/0/17/20	0/1/1/1
3	MAN	B	607	3	-	0/2/19/22	0/1/1/1
4	NAG	B	608	1,4	-	0/6/23/26	0/1/1/1
4	NAG	B	609	4	-	0/6/23/26	0/1/1/1
7	NAG	C	602	1,7	-	0/6/23/26	0/1/1/1
7	NAG	C	603	7	-	0/6/23/26	0/1/1/1
7	FUC	C	604	7	-	0/0/17/20	0/1/1/1
7	FUL	C	605	7	-	0/0/17/20	0/1/1/1
4	NAG	C	606	1,4	-	0/6/23/26	0/1/1/1
4	NAG	C	607	4	-	0/6/23/26	0/1/1/1

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	610	FUL	C4-C5	2.24	1.57	1.52
7	C	604	FUC	C4-C5	2.22	1.57	1.52
2	A	603	BMA	O5-C5	-2.14	1.41	1.45
3	A	608	MAN	O5-C5	-2.11	1.41	1.45

All (36) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	B	609	NAG	O5-C5-C4	6.28	118.62	110.65
4	A	613	NAG	O5-C5-C4	5.76	117.97	110.65
7	C	602	NAG	O5-C5-C6	4.94	112.16	106.98
3	B	601	NAG	O5-C5-C6	4.45	111.65	106.98
3	B	607	MAN	O5-C5-C4	4.34	116.16	110.65
3	A	605	NAG	O5-C5-C6	3.98	111.16	106.98
4	A	613	NAG	C3-C4-C5	3.45	116.36	110.20
7	C	604	FUC	O5-C5-C4	3.36	115.29	110.22
3	A	611	MAN	O5-C5-C6	3.35	110.50	106.98
3	B	607	MAN	O5-C5-C6	3.03	110.16	106.98
3	B	604	MAN	O5-C5-C4	2.99	114.45	110.65
3	A	606	NAG	O5-C5-C6	2.98	110.11	106.98
3	B	604	MAN	C4-C3-C2	-2.92	106.58	110.50
3	B	601	NAG	C2-N2-C7	-2.90	118.22	123.09
3	A	607	BMA	C6-C5-C4	-2.89	106.01	113.00

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	602	NAG	O5-C5-C6	2.88	110.01	106.98
3	B	607	MAN	C3-C4-C5	2.87	115.33	110.20
3	B	603	BMA	O5-C5-C6	2.83	109.95	106.98
2	A	602	NAG	O5-C5-C6	2.78	109.89	106.98
2	A	603	BMA	O5-C5-C6	2.67	109.78	106.98
3	B	606	FUL	O5-C5-C4	2.65	114.23	110.22
3	A	608	MAN	O5-C5-C6	2.60	109.71	106.98
7	C	603	NAG	C3-C4-C5	2.46	114.60	110.20
4	B	609	NAG	C2-N2-C7	-2.46	118.96	123.09
3	B	607	MAN	C4-C3-C2	2.36	113.67	110.50
2	A	603	BMA	O5-C5-C4	-2.36	107.66	110.65
4	B	609	NAG	C4-C3-C2	-2.35	105.56	111.32
2	A	601	NAG	O4-C4-C3	-2.33	105.14	110.35
7	C	603	NAG	O5-C5-C6	2.30	109.39	106.98
4	A	612	NAG	C2-N2-C7	-2.29	119.25	123.09
7	C	602	NAG	O3-C3-C2	2.28	113.87	109.09
2	A	601	NAG	O5-C5-C4	2.21	113.46	110.65
3	A	610	FUL	O5-C5-C4	2.20	113.54	110.22
4	A	612	NAG	O5-C5-C6	2.15	109.24	106.98
3	B	601	NAG	O5-C5-C4	-2.10	107.98	110.65
3	B	602	NAG	C2-N2-C7	-2.05	119.64	123.09

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 5.6 Ligand geometry ⓘ

11 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
5	FUC	A	614	-	9,10,11	0.97	1 (11%)	10,14,16	1.41	2 (20%)
6	NAG	A	615	1	12,14,15	0.65	0	15,19,21	0.68	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
6	NAG	A	616	1	12,14,15	0.70	1 (8%)	15,19,21	1.10	1 (6%)
5	FUC	B	610	-	9,10,11	0.93	1 (11%)	10,14,16	1.37	2 (20%)
6	NAG	B	611	1	12,14,15	0.66	0	15,19,21	1.01	1 (6%)
6	NAG	B	612	1	12,14,15	0.63	0	15,19,21	0.85	0
6	NAG	B	613	1	12,14,15	0.61	0	15,19,21	0.88	1 (6%)
6	NAG	C	601	1	12,14,15	0.47	0	15,19,21	1.30	3 (20%)
5	FUC	C	608	-	9,10,11	0.78	1 (11%)	10,14,16	1.24	1 (10%)
6	NAG	C	609	1	12,14,15	0.68	1 (8%)	15,19,21	0.61	0
6	NAG	C	610	1	12,14,15	0.71	1 (8%)	15,19,21	1.31	3 (20%)

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
5	FUC	A	614	-	-	0/0/17/20	0/1/1/1
6	NAG	A	615	1	-	0/6/23/26	0/1/1/1
6	NAG	A	616	1	1/1/5/7	0/6/23/26	0/1/1/1
5	FUC	B	610	-	-	0/0/17/20	0/1/1/1
6	NAG	B	611	1	-	0/6/23/26	0/1/1/1
6	NAG	B	612	1	1/1/5/7	0/6/23/26	0/1/1/1
6	NAG	B	613	1	-	0/6/23/26	0/1/1/1
6	NAG	C	601	1	-	0/6/23/26	0/1/1/1
5	FUC	C	608	-	-	0/0/17/20	0/1/1/1
6	NAG	C	609	1	-	0/6/23/26	0/1/1/1
6	NAG	C	610	1	1/1/5/7	0/6/23/26	0/1/1/1

All (6) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	B	610	FUC	O5-C5	2.47	1.50	1.45
5	A	614	FUC	O5-C5	2.36	1.49	1.45
6	C	610	NAG	O5-C5	-2.12	1.41	1.45
6	C	609	NAG	O5-C5	-2.09	1.41	1.45
5	C	608	FUC	O5-C5	2.07	1.49	1.45
6	A	616	NAG	O5-C5	-2.04	1.41	1.45

All (14) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	614	FUC	O5-C5-C4	3.42	115.39	110.22
6	C	610	NAG	O5-C5-C6	3.26	110.41	106.98
6	A	616	NAG	O5-C5-C6	3.24	110.38	106.98
5	C	608	FUC	O5-C5-C4	3.06	114.85	110.22
5	B	610	FUC	O5-C5-C4	2.95	114.68	110.22
6	B	611	NAG	O5-C5-C6	2.87	109.99	106.98
6	C	601	NAG	O5-C5-C4	2.84	114.26	110.65
6	B	613	NAG	O5-C5-C6	2.78	109.90	106.98
6	C	610	NAG	O5-C5-C4	-2.61	107.34	110.65
5	B	610	FUC	C4-C3-C2	-2.36	107.34	110.50
6	C	601	NAG	C4-C3-C2	-2.31	105.66	111.32
6	C	601	NAG	O5-C5-C6	2.24	109.33	106.98
6	C	610	NAG	C3-C2-N2	-2.16	108.48	111.76
5	A	614	FUC	C4-C3-C2	-2.02	107.79	110.50

All (3) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
6	C	610	NAG	C1
6	B	612	NAG	C1
6	A	616	NAG	C1

There are no torsion outliers.

There are no ring outliers.

## 5.7 Other polymers ⓘ

There are no such residues in this entry.

## 5.8 Polymer linkage issues

There are no chain breaks in this entry.

## 6 Fit of model and data ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

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	494/510 (96%)	-0.39	3 (0%) 86 89	23, 39, 59, 110	0
1	B	494/510 (96%)	-0.55	2 (0%) 90 93	21, 34, 56, 86	0
1	C	497/510 (97%)	-0.58	1 (0%) 93 96	22, 36, 56, 82	0
All	All	1485/1530 (97%)	-0.51	6 (0%) 90 93	21, 37, 57, 110	0

All (6) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	143	ASN	7.4
1	A	142	THR	6.3
1	B	334	ALA	3.4
1	C	140	TYR	2.5
1	B	140	TYR	2.5
1	A	171	ASN	2.5

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

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

### 6.3 Carbohydrates ⓘ

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
4	NAG	B	609	14/15	0.21	14.30	66,73,86,86	0
3	MAN	B	604	11/12	0.40	13.56	103,109,114,114	0
4	NAG	A	613	14/15	0.19	6.52	72,74,80,81	0
3	MAN	B	607	11/12	0.23	6.38	84,90,96,96	0
4	NAG	C	607	14/15	0.29	6.10	73,84,90,90	0
3	MAN	A	608	11/12	0.26	5.69	82,88,94,94	0
4	NAG	A	612	14/15	0.23	4.80	61,69,75,81	0
3	MAN	A	611	11/12	0.29	4.36	67,76,86,87	0
4	NAG	C	606	14/15	0.25	3.90	55,64,76,79	0
3	BMA	B	603	11/12	0.19	3.39	74,82,89,92	0
3	BMA	A	607	11/12	0.18	2.71	65,69,75,75	0
3	FUL	B	606	10/11	0.16	2.71	63,73,84,85	0
3	FUL	A	610	10/11	0.22	2.37	58,68,71,72	0
4	NAG	B	608	14/15	0.15	2.30	57,65,76,82	0
3	NAG	A	606	14/15	0.16	1.35	53,58,63,65	0
7	NAG	C	602	14/15	0.21	0.86	80,87,100,108	0
3	NAG	B	602	14/15	0.10	0.24	52,59,63,63	0
3	NAG	A	605	14/15	0.13	-0.07	42,52,62,69	0
3	NAG	B	601	14/15	0.09	-1.05	37,47,57,66	0
2	BMA	A	603	11/12	0.24	-	66,72,77,80	0
2	NAG	A	601	14/15	0.26	-	73,79,85,85	0
2	NAG	A	602	14/15	0.26	-	78,80,82,83	0
7	FUC	C	604	10/11	0.39	-	102,104,105,106	0
2	MAN	A	604	11/12	0.40	-	82,87,91,91	0
7	FUL	C	605	10/11	0.37	-	93,106,135,135	0
7	NAG	C	603	14/15	0.29	-	104,114,120,120	0
3	FUC	A	609	10/11	0.10	-	52,58,63,65	0
3	FUC	B	605	10/11	0.10	-	52,55,59,60	0

## 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
5	FUC	B	610	10/11	0.40	20.87	81,86,87,87	0
6	NAG	B	613	14/15	0.39	14.91	85,90,93,95	0
6	NAG	C	610	14/15	0.22	11.47	81,88,89,90	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
5	FUC	A	614	10/11	0.27	11.03	75,83,85,86	0
5	FUC	C	608	10/11	0.34	9.16	86,89,92,94	0
6	NAG	B	611	14/15	0.19	2.45	63,68,73,75	0
6	NAG	C	609	14/15	0.17	2.41	58,68,71,73	0
6	NAG	A	616	14/15	0.18	1.79	82,89,91,92	0
6	NAG	B	612	14/15	0.20	1.33	83,91,92,93	0
6	NAG	A	615	14/15	0.14	0.99	64,73,75,76	0
6	NAG	C	601	14/15	0.23	-	95,103,107,108	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.