



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 03:40 PM GMT

PDB ID : 4CYW  
Title : Structure of the A\_mallard\_Sweden\_51\_2002 H10 Avian Haemmagglutinin  
in complex with human receptor analog 6-SLN  
Authors : Vachieri, S.G.; Xiong, X.; Collins, P.J.; Walker, P.A.; Martin, S.R.; Haire,  
L.F.; McCauley, J.W.; Gamblin, S.J.; Skehel, J.J.  
Deposited on : 2014-04-15  
Resolution : 2.60 Å(reported)

This is a Full wwPDB X-ray Structure 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/validation/2016/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 : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

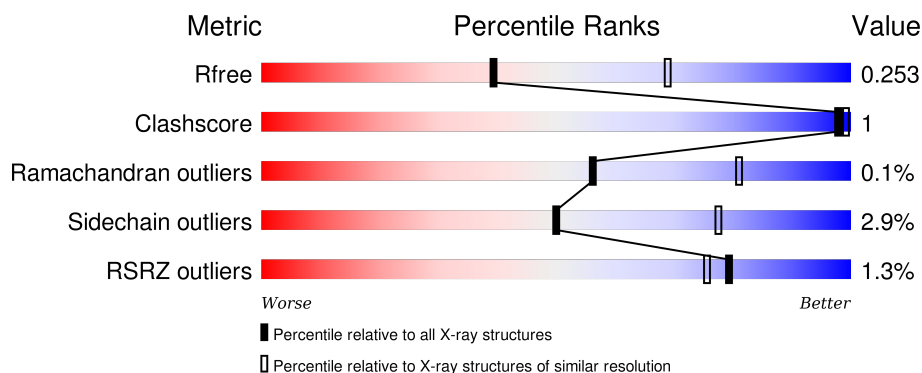
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.60 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.



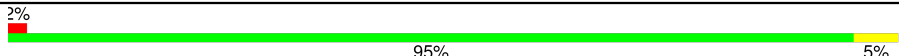
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	91344	2328 (2.60-2.60)
Clashscore	102246	2679 (2.60-2.60)
Ramachandran outliers	100387	2635 (2.60-2.60)
Sidechain outliers	100360	2635 (2.60-2.60)
RSRZ outliers	91569	2334 (2.60-2.60)

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. 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	324	<div> <div>3%</div> <div>93%</div> <div>5%</div> </div>
1	C	324	<div> <div>%</div> <div>95%</div> <div>• •</div> </div>
1	E	324	<div> <div>92%</div> <div>6%</div> <div>• •</div> </div>
2	B	173	<div> <div>%</div> <div>95%</div> <div>• •</div> </div>
2	D	173	<div> <div>96%</div> <div>•</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	F	173	

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
3	NAG	A	420	X	-	-	X
3	NAG	C	420	X	-	-	-
3	NAG	E	401	-	-	-	X
6	MAN	D	207	X	-	-	-

## 2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 12105 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
1	A	318	Total	C	N	O	S	0	0	0
			2411	1490	434	471	16			
1	C	319	Total	C	N	O	S	0	0	0
			2430	1504	437	473	16			
1	E	319	Total	C	N	O	S	0	0	0
			2430	1504	437	473	16			

- Molecule 2 is a protein called HEMAGGLUTININ.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	172	Total	C	N	O	S	0	0	0
			1381	853	238	282	8			
2	D	173	Total	C	N	O	S	0	0	0
			1384	856	239	281	8			
2	F	173	Total	C	N	O	S	0	0	0
			1391	860	240	283	8			

- Molecule 3 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



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

- Molecule 4 is a polymer of unknown type called SUGAR (3-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	3	Total	C	N	O	0	0
			46	25	2	19		
4	C	3	Total	C	N	O	0	0
			46	25	2	19		
4	E	3	Total	C	N	O	0	0
			45	25	2	18		

- Molecule 5 is a polymer of unknown type called SUGAR (5-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	B	5	Total	C	N	O	0	0
			61	34	2	25		

- Molecule 6 is a polymer of unknown type called SUGAR (5-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	D	5	Total	C	N	O	0	0
			61	34	2	25		

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

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

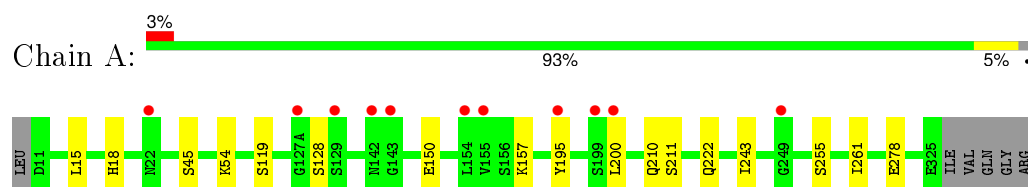
- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	48	Total	O	0	0
			48	48		
8	B	47	Total	O	0	0
			47	47		
8	C	52	Total	O	0	0
			52	52		
8	D	42	Total	O	0	0
			42	42		
8	E	73	Total	O	0	0
			73	73		
8	F	23	Total	O	0	0
			23	23		

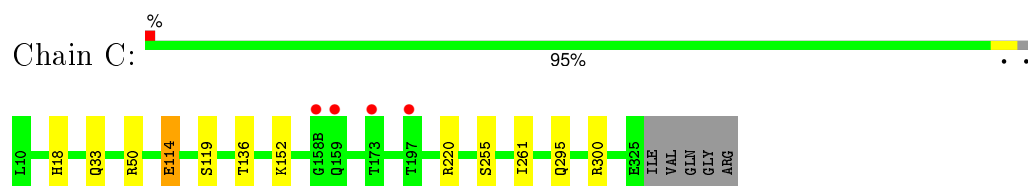
### 3 Residue-property plots [i](#)

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.

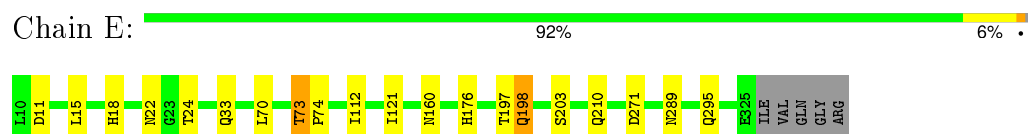
- Molecule 1: HEMAGGLUTININ



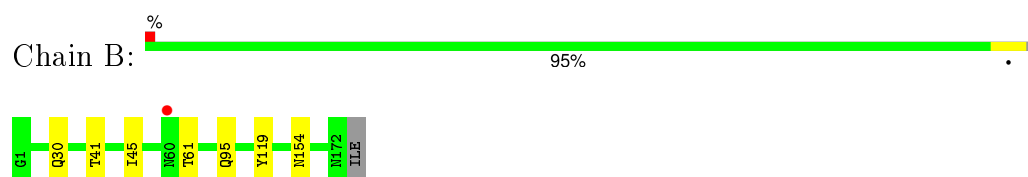
- Molecule 1: HEMAGGLUTININ



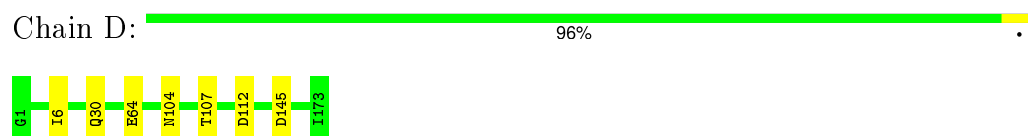
- Molecule 1: HEMAGGLUTININ



- Molecule 2: HEMAGGLUTININ

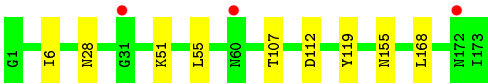


- Molecule 2: HEMAGGLUTININ



- Molecule 2: HEMAGGLUTININ







## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	68.97Å 215.45Å 79.36Å 90.00° 104.91° 90.00°	Depositor
Resolution (Å)	107.72 – 2.60 48.85 – 2.60	Depositor EDS
% Data completeness (in resolution range)	99.1 (107.72-2.60) 99.1 (48.85-2.60)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.32 (at 2.61Å)	Xtriage
Refinement program	REFMAC 5.8.0049	Depositor
R, $R_{free}$	0.208 , 0.254 0.211 , 0.253	Depositor DCC
$R_{free}$ test set	3428 reflections (5.32%)	DCC
Wilson B-factor (Å <sup>2</sup> )	48.4	Xtriage
Anisotropy	0.689	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 33.9	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	0 of 67867 reflections	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	12105	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	62.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.24% 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, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

## 5 Model quality

### 5.1 Standard geometry

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

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	0.28	0/2460	0.50	0/3337
1	C	0.28	0/2479	0.50	0/3361
1	E	0.30	0/2479	0.53	0/3361
2	B	0.31	0/1406	0.50	0/1899
2	D	0.31	0/1409	0.49	0/1904
2	F	0.30	0/1416	0.49	0/1912
All	All	0.30	0/11649	0.50	0/15774

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
6	D	1	0

There are no bond length outliers.

There are no bond angle outliers.

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
6	D	207	MAN	C1

There are no planarity outliers.

### 5.2 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen

atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2411	0	2330	6	0
1	C	2430	0	2367	3	0
1	E	2430	0	2367	7	0
2	B	1381	0	1268	2	0
2	D	1384	0	1273	5	0
2	F	1391	0	1286	4	0
3	A	28	0	26	0	0
3	C	28	0	26	3	0
3	E	28	0	26	0	0
4	A	46	0	40	1	0
4	C	46	0	40	0	0
4	E	45	0	38	0	0
5	B	61	0	52	1	0
6	D	61	0	52	0	0
7	F	50	0	43	0	0
8	A	48	0	0	0	0
8	B	47	0	0	0	0
8	C	52	0	0	0	0
8	D	42	0	0	0	0
8	E	73	0	0	0	0
8	F	23	0	0	0	0
All	All	12105	0	11234	26	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 1.

All (26) 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)
3:C:401:NAG:H83	3:C:401:NAG:H3	1.87	0.57
1:A:128:SER:O	1:A:157:LYS:NZ	2.40	0.55
2:D:104:ASN:HA	2:D:107:THR:HG22	1.89	0.54
1:E:121:ILE:HD13	1:E:176:HIS:CE1	2.42	0.54
1:E:70:LEU:HD11	1:E:112:ILE:HD11	1.92	0.52
5:B:201:NAG:H81	1:C:114:GLU:OE2	2.09	0.52
1:C:114:GLU:OE2	2:D:64:GLU:OE1	2.29	0.50
3:C:401:NAG:H3	3:C:401:NAG:C8	2.44	0.48
1:A:119:SER:OG	1:A:261:ILE:HD11	2.15	0.47
1:A:54:LYS:HB3	1:A:278:GLU:HB3	1.97	0.45

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:15:LEU:HD13	2:B:119:TYR:HA	1.98	0.45
1:A:243:ILE:HG23	1:A:243:ILE:O	2.17	0.44
2:D:30:GLN:HE22	2:D:145:ASP:HA	1.82	0.44
1:E:73:THR:HG22	1:E:74:PRO:HD2	1.99	0.44
1:E:15:LEU:HD13	2:F:119:TYR:HA	1.99	0.44
2:D:6:ILE:HD12	2:D:112:ASP:HA	1.98	0.44
1:C:119:SER:OG	1:C:261:ILE:HD11	2.19	0.43
1:A:222:GLN:NE2	4:A:702:GAL:O3	2.51	0.43
2:F:6:ILE:HD12	2:F:112:ASP:HA	2.02	0.42
1:E:160:ASN:CB	1:E:198:GLN:HE21	2.32	0.42
2:F:51:LYS:NZ	2:F:107:THR:HG23	2.34	0.42
2:B:41:THR:HG22	2:B:45:ILE:HD12	2.02	0.42
1:E:11:ASP:HB3	2:F:28:ASN:HA	2.02	0.42
1:E:160:ASN:HB3	1:E:198:GLN:HE21	1.85	0.41
3:C:401:NAG:H83	3:C:401:NAG:C3	2.51	0.41
2:D:104:ASN:HA	2:D:107:THR:CG2	2.51	0.41

There are no symmetry-related clashes.

## 5.3 Torsion angles ⓘ

### 5.3.1 Protein backbone ⓘ

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	316/324 (98%)	306 (97%)	10 (3%)	0	100	100
1	C	317/324 (98%)	304 (96%)	13 (4%)	0	100	100
1	E	317/324 (98%)	310 (98%)	6 (2%)	1 (0%)	46	72
2	B	170/173 (98%)	165 (97%)	5 (3%)	0	100	100
2	D	171/173 (99%)	165 (96%)	6 (4%)	0	100	100
2	F	171/173 (99%)	165 (96%)	6 (4%)	0	100	100
All	All	1462/1491 (98%)	1415 (97%)	46 (3%)	1 (0%)	56	81

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	E	198	GLN

### 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	266/275 (97%)	258 (97%)	8 (3%)	48	76
1	C	270/275 (98%)	260 (96%)	10 (4%)	41	69
1	E	270/275 (98%)	259 (96%)	11 (4%)	37	66
2	B	144/147 (98%)	140 (97%)	4 (3%)	51	78
2	D	144/147 (98%)	144 (100%)	0	100	100
2	F	146/147 (99%)	143 (98%)	3 (2%)	61	85
All	All	1240/1266 (98%)	1204 (97%)	36 (3%)	50	77

All (36) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	18	HIS
1	A	45	SER
1	A	150	GLU
1	A	195	TYR
1	A	200	LEU
1	A	210	GLN
1	A	211	SER
1	A	255	SER
2	B	30	GLN
2	B	61	THR
2	B	95	GLN
2	B	154	ASN
1	C	18	HIS
1	C	33	GLN
1	C	50	ARG
1	C	114	GLU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	C	136	THR
1	C	152	LYS
1	C	220	ARG
1	C	255	SER
1	C	295	GLN
1	C	300	ARG
1	E	18	HIS
1	E	22	ASN
1	E	24	THR
1	E	33	GLN
1	E	73	THR
1	E	197	THR
1	E	203	SER
1	E	210	GLN
1	E	271	ASP
1	E	289	ASN
1	E	295	GLN
2	F	55	LEU
2	F	155	ASN
2	F	168	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (11) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	63	ASN
1	A	104	ASN
1	A	166	ASN
1	A	222	GLN
2	B	172	ASN
1	C	22	ASN
2	D	30	GLN
2	D	161	GLN
1	E	276	ASN
2	F	30	GLN
2	F	155	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 ⓘ

23 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	$\# Z  > 2$
4	SIA	A	701	4	16,20,21	0.34	0	18,28,31	1.69	3 (16%)
4	GAL	A	702	4	11,11,12	0.63	0	14,15,17	0.70	0
4	NAG	A	703	4	15,15,15	0.46	0	17,21,21	1.20	2 (11%)
5	NAG	B	201	2,5	14,14,15	0.48	0	15,19,21	1.53	3 (20%)
5	NAG	B	202	5	14,14,15	0.53	0	15,19,21	1.01	2 (13%)
5	BMA	B	203	5	11,11,12	0.38	0	14,15,17	0.69	0
5	MAN	B	204	5	11,11,12	0.59	0	14,15,17	1.11	2 (14%)
5	MAN	B	207	5	11,11,12	0.53	0	14,15,17	1.22	2 (14%)
4	SIA	C	701	4	16,20,21	0.34	0	18,28,31	1.22	1 (5%)
4	GAL	C	702	4	11,11,12	0.59	0	14,15,17	1.03	1 (7%)
4	NAG	C	703	4	15,15,15	0.48	0	17,21,21	1.06	1 (5%)
6	NAG	D	201	2,6	14,14,15	0.57	0	15,19,21	0.92	1 (6%)
6	NAG	D	202	6	14,14,15	0.41	0	15,19,21	1.83	2 (13%)
6	BMA	D	203	6	11,11,12	0.45	0	14,15,17	0.95	0
6	BMA	D	204	6	11,11,12	0.93	1 (9%)	14,15,17	2.23	3 (21%)
6	MAN	D	207	6	11,11,12	0.67	0	14,15,17	1.75	3 (21%)
4	SIA	E	701	4	16,20,21	0.25	0	18,28,31	0.68	0
4	GAL	E	702	4	11,11,12	0.63	0	14,15,17	1.48	2 (14%)
4	NAG	E	703	4	14,14,15	0.64	0	15,19,21	1.61	2 (13%)
7	NAG	F	201	2,7	14,14,15	0.60	0	15,19,21	1.11	1 (6%)
7	NAG	F	202	7	14,14,15	0.38	0	15,19,21	1.56	2 (13%)
7	BMA	F	203	7	11,11,12	0.62	0	14,15,17	2.32	6 (42%)
7	MAN	F	204	7	11,11,12	0.54	0	14,15,17	1.23	2 (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
4	SIA	A	701	4	-	0/14/34/38	0/1/1/1
4	GAL	A	702	4	-	0/2/19/22	0/1/1/1
4	NAG	A	703	4	-	0/6/26/26	0/1/1/1
5	NAG	B	201	2,5	-	0/6/23/26	0/1/1/1
5	NAG	B	202	5	-	0/6/23/26	0/1/1/1
5	BMA	B	203	5	-	0/2/19/22	0/1/1/1
5	MAN	B	204	5	-	0/2/19/22	0/1/1/1
5	MAN	B	207	5	-	0/2/19/22	0/1/1/1
4	SIA	C	701	4	-	0/14/34/38	0/1/1/1
4	GAL	C	702	4	-	0/2/19/22	0/1/1/1
4	NAG	C	703	4	-	0/6/26/26	0/1/1/1
6	NAG	D	201	2,6	-	0/6/23/26	0/1/1/1
6	NAG	D	202	6	-	0/6/23/26	0/1/1/1
6	BMA	D	203	6	-	0/2/19/22	0/1/1/1
6	BMA	D	204	6	-	0/2/19/22	0/1/1/1
6	MAN	D	207	6	1/1/4/5	0/2/19/22	0/1/1/1
4	SIA	E	701	4	-	0/14/34/38	0/1/1/1
4	GAL	E	702	4	-	0/2/19/22	0/1/1/1
4	NAG	E	703	4	-	0/6/23/26	0/1/1/1
7	NAG	F	201	2,7	-	0/6/23/26	0/1/1/1
7	NAG	F	202	7	-	0/6/23/26	0/1/1/1
7	BMA	F	203	7	-	0/2/19/22	0/1/1/1
7	MAN	F	204	7	-	0/2/19/22	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	D	204	BMA	C2-C3	2.28	1.55	1.52

All (41) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
7	F	202	NAG	C4-C3-C2	-2.61	107.17	111.23
5	B	201	NAG	C4-C3-C2	-2.37	107.55	111.23
5	B	202	NAG	C3-C4-C5	2.02	113.73	110.20
4	C	702	GAL	C3-C4-C5	2.06	113.78	110.20
7	F	204	MAN	C1-C2-C3	2.07	111.99	109.54
6	D	202	NAG	C3-C4-C5	2.09	113.84	110.20

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	701	SIA	C5-N5-C10	2.12	128.56	123.10
5	B	204	MAN	C1-O5-C5	2.15	114.98	112.25
7	F	203	BMA	C3-C4-C5	2.23	114.08	110.20
5	B	202	NAG	C1-O5-C5	2.27	115.12	112.25
6	D	201	NAG	C1-O5-C5	2.27	115.13	112.25
5	B	207	MAN	C1-C2-C3	2.28	112.24	109.54
4	A	703	NAG	C3-C4-C5	2.30	114.22	110.20
7	F	203	BMA	O5-C1-C2	2.35	114.67	110.86
5	B	201	NAG	C8-C7-N2	2.39	120.69	116.11
4	C	703	NAG	C3-C4-C5	2.46	114.48	110.20
7	F	201	NAG	C1-O5-C5	2.50	115.42	112.25
6	D	204	BMA	C1-O5-C5	2.52	115.45	112.25
7	F	203	BMA	O5-C5-C6	2.54	112.85	107.35
4	A	701	SIA	C7-C6-C5	2.61	118.27	114.32
4	A	703	NAG	C4-C3-C2	2.72	114.20	110.43
4	E	702	GAL	C1-C2-C3	2.85	112.91	109.54
7	F	203	BMA	C2-C3-C4	2.90	115.97	111.04
5	B	204	MAN	C1-C2-C3	2.96	113.04	109.54
6	D	204	BMA	C2-C3-C4	3.08	116.28	111.04
6	D	207	MAN	C1-C2-C3	3.16	113.27	109.54
6	D	207	MAN	O5-C1-C2	3.45	116.46	110.86
4	C	701	SIA	C4-C5-N5	3.54	118.11	110.41
4	E	703	NAG	C3-C4-C5	3.62	116.51	110.20
7	F	204	MAN	C1-O5-C5	3.64	116.87	112.25
5	B	207	MAN	C1-O5-C5	3.68	116.91	112.25
4	E	703	NAG	C4-C3-C2	3.94	117.35	111.23
6	D	207	MAN	C1-O5-C5	3.98	117.31	112.25
4	E	702	GAL	C1-O5-C5	4.16	117.52	112.25
7	F	203	BMA	C1-C2-C3	4.16	114.47	109.54
5	B	201	NAG	C1-O5-C5	4.37	117.80	112.25
7	F	202	NAG	C1-O5-C5	4.56	118.03	112.25
4	A	701	SIA	C6-C5-N5	5.19	120.12	111.07
7	F	203	BMA	C1-O5-C5	5.46	119.18	112.25
6	D	202	NAG	C1-O5-C5	6.28	120.22	112.25
6	D	204	BMA	C1-C2-C3	7.00	117.83	109.54

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
6	D	207	MAN	C1

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	702	GAL	1	0
5	B	201	NAG	1	0

## 5.6 Ligand geometry

6 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	$\# Z  > 2$
3	NAG	A	401	1	14,14,15	0.57	0	15,19,21	1.50	2 (13%)
3	NAG	A	420	1	14,14,15	0.46	0	15,19,21	1.60	2 (13%)
3	NAG	C	401	1	14,14,15	0.41	0	15,19,21	1.97	5 (33%)
3	NAG	C	420	1	14,14,15	0.61	0	15,19,21	0.59	0
3	NAG	E	401	1	14,14,15	0.41	0	15,19,21	1.50	4 (26%)
3	NAG	E	420	1	14,14,15	0.35	0	15,19,21	1.56	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
3	NAG	A	401	1	-	0/6/23/26	0/1/1/1
3	NAG	A	420	1	1/1/5/7	0/6/23/26	0/1/1/1
3	NAG	C	401	1	-	0/6/23/26	0/1/1/1
3	NAG	C	420	1	1/1/5/7	0/6/23/26	0/1/1/1
3	NAG	E	401	1	-	0/6/23/26	0/1/1/1
3	NAG	E	420	1	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

All (16) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	401	NAG	C4-C3-C2	-3.05	106.49	111.23
3	C	401	NAG	O7-C7-C8	-2.24	117.95	122.06
3	E	401	NAG	C4-C3-C2	-2.17	107.86	111.23
3	E	420	NAG	C4-C3-C2	-2.09	107.98	111.23
3	C	401	NAG	C3-C2-N2	2.08	115.53	110.56
3	E	420	NAG	C8-C7-N2	2.16	120.25	116.11
3	E	401	NAG	C3-C2-N2	2.25	115.96	110.56
3	E	401	NAG	C2-N2-C7	2.53	126.29	123.04
3	E	401	NAG	C1-O5-C5	2.66	115.63	112.25
3	A	420	NAG	C3-C4-C5	2.72	114.95	110.20
3	C	401	NAG	C8-C7-N2	3.06	121.95	116.11
3	A	401	NAG	C3-C2-N2	3.31	118.48	110.56
3	C	401	NAG	C2-N2-C7	3.66	127.75	123.04
3	E	420	NAG	C1-O5-C5	3.71	116.96	112.25
3	C	401	NAG	C1-O5-C5	4.41	117.84	112.25
3	A	420	NAG	C1-O5-C5	4.57	118.05	112.25

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
3	A	420	NAG	C1
3	C	420	NAG	C1

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	C	401	NAG	3	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 ⓘ

### 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	318/324 (98%)	0.19	11 (3%) 48 40	45, 69, 110, 120	0
1	C	319/324 (98%)	-0.08	4 (1%) 79 75	38, 62, 87, 99	0
1	E	319/324 (98%)	-0.24	0 100 100	36, 52, 83, 94	0
2	B	172/173 (99%)	-0.07	1 (0%) 90 88	38, 55, 74, 94	0
2	D	173/173 (100%)	-0.09	0 100 100	35, 51, 70, 81	0
2	F	173/173 (100%)	0.01	3 (1%) 73 68	36, 56, 80, 90	0
All	All	1474/1491 (98%)	-0.05	19 (1%) 79 75	35, 58, 94, 120	0

All (19) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	200	LEU	5.0
1	A	249	GLY	4.7
1	C	158(B)	GLY	4.7
1	A	127(A)	GLY	4.1
1	A	195	TYR	3.9
1	A	155	VAL	3.5
1	C	197	THR	3.2
1	A	143	GLY	3.1
2	F	60	ASN	2.9
2	B	60	ASN	2.8
1	C	159	GLN	2.6
1	A	142	ASN	2.5
2	F	172	ASN	2.5
1	A	199	SER	2.2
1	C	173	THR	2.2
1	A	154	LEU	2.1
1	A	22	ASN	2.1
1	A	129	SER	2.1
2	F	31	GLY	2.0

## 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	RSCC	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
4	SIA	C	701	20/21	0.93	0.15	0.07	67,70,73,74	0
4	SIA	E	701	20/21	0.96	0.15	-0.46	68,72,79,81	0
4	SIA	A	701	20/21	0.91	0.21	-0.64	101,105,109,110	0
7	NAG	F	201	14/15	0.94	0.12	-1.25	51,54,59,64	0
6	NAG	D	201	14/15	0.93	0.13	-1.56	53,56,62,68	0
5	NAG	B	201	14/15	0.89	0.15	-2.98	57,62,66,73	0
5	MAN	B	204	11/12	0.56	0.41	-	114,116,122,126	0
6	NAG	D	202	14/15	0.92	0.20	-	76,83,87,89	0
4	GAL	E	702	11/12	0.83	0.28	-	94,105,109,109	0
4	GAL	C	702	11/12	0.86	0.29	-	86,96,102,102	0
5	NAG	B	202	14/15	0.91	0.20	-	79,83,92,97	0
5	MAN	B	207	11/12	0.64	0.41	-	108,112,116,117	0
4	GAL	A	702	11/12	0.78	0.33	-	118,128,132,133	0
6	MAN	D	207	11/12	0.79	0.28	-	97,101,104,104	0
7	MAN	F	204	11/12	0.43	0.51	-	111,120,125,127	0
7	NAG	F	202	14/15	0.87	0.17	-	70,74,82,90	0
4	NAG	E	703	14/15	0.79	0.38	-	110,115,122,122	0
4	NAG	C	703	15/15	0.71	0.35	-	104,107,110,110	0
7	BMA	F	203	11/12	0.62	0.27	-	98,101,107,114	0
5	BMA	B	203	11/12	0.82	0.25	-	102,109,115,115	0
4	NAG	A	703	15/15	0.76	0.36	-	132,137,143,145	0
6	BMA	D	203	11/12	0.79	0.30	-	95,101,105,109	0
6	BMA	D	204	11/12	0.45	0.36	-	103,105,110,110	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	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
3	NAG	A	420	14/15	0.72	0.36	13.46	90,97,98,100	0
3	NAG	E	401	14/15	0.78	0.24	3.79	79,85,88,90	0
3	NAG	A	401	14/15	0.88	0.20	1.24	77,79,82,83	0
3	NAG	C	420	14/15	0.82	0.28	-	77,83,89,90	0
3	NAG	C	401	14/15	0.80	0.36	-	97,102,106,108	0
3	NAG	E	420	14/15	0.87	0.29	-	77,83,86,88	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.