Full wwPDB X-ray Structure Validation Report (i)

Feb 1, 2016 – 09:01 PM GMT

PDB ID	:	4UO1
Title	:	Structure of the A_Equine_Richmond_07 H3 haemagglutinin in complex with 3SLN
Authors	:	Vachieri, S.G.; Collins, P.J.; Haire, L.F.; Ogrodowicz, R.W.; Martin, S.R.; Walker, P.A.; Xiong, X.; Gamblin, S.J.; Skehel, J.J.
Deposited on	:	2014-05-31
Resolution	:	3.00 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
A user guide is available at
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

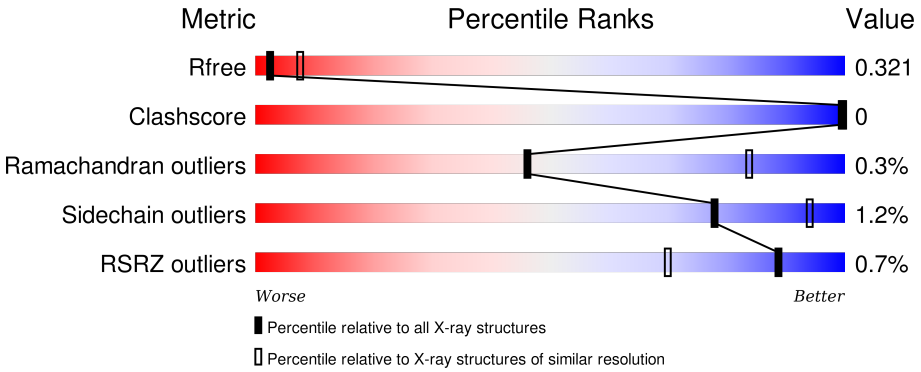
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 3.00 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	1578 (3.00-3.00)
Clashscore	102246	1912 (3.00-3.00)
Ramachandran outliers	100387	1853 (3.00-3.00)
Sidechain outliers	100360	1856 (3.00-3.00)
RSRZ outliers	91569	1592 (3.00-3.00)

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 >=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 <=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	329	96%
1	C	329	95%
1	E	329	95%
2	B	172	99%
2	D	172	99%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	172	98%

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
12	FUC	E	600	X	-	-	-
12	NAG	E	602	X	-	-	-
12	FUC	F	200	X	-	-	-
5	MAN	C	434	X	-	-	-
5	MAN	E	437	X	-	-	-
6	FUC	A	600	X	-	-	-
9	FUC	B	200	X	-	-	X
9	NAG	B	201	-	-	-	X

2 Entry composition

There are 13 unique types of molecules in this entry. The entry contains 12524 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.

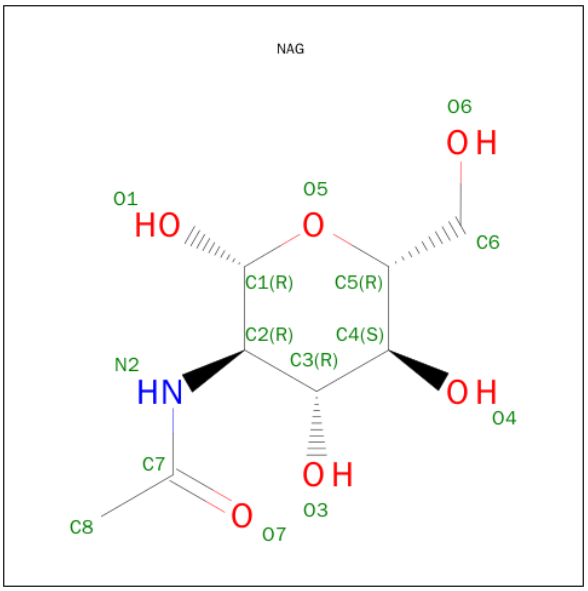
- Molecule 1 is a protein called HEMAGGLUTININ.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	325	Total	C	N	O	S	0	0	0
			2535	1584	449	487	15			
1	C	320	Total	C	N	O	S	0	0	0
			2492	1557	441	479	15			
1	E	325	Total	C	N	O	S	0	0	0
			2526	1578	447	486	15			

- 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
			1399	873	242	278	6			
2	D	171	Total	C	N	O	S	0	0	0
			1384	863	239	276	6			
2	F	172	Total	C	N	O	S	0	0	0
			1403	874	245	278	6			

- 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 14 8 1 5	0	0
3	C	1	Total C N O 14 8 1 5	0	0
3	C	1	Total C N O 14 8 1 5	0	0
3	D	1	Total C N O 14 8 1 5	0	0
3	D	1	Total C N O 14 8 1 5	0	0
3	E	1	Total C N O 14 8 1 5	0	0
3	E	1	Total C N O 14 8 1 5	0	0

- 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 28 16 2 10	0	0
4	A	2	Total C N O 28 16 2 10	0	0
4	C	2	Total C N O 28 16 2 10	0	0
4	E	2	Total C N O 28 16 2 10	0	0

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

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

- Molecule 6 is a polymer of unknown type called SUGAR (4-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	A	4	Total	C	N	O	0	0
			49	28	2	19		

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

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
7	A	3	Total	C	N	O	0	0
			39	22	2	15		

- Molecule 8 is a polymer of unknown type called SUGAR (2-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
8	A	2	Total	C	N	O	0	0
			31	17	1	13		

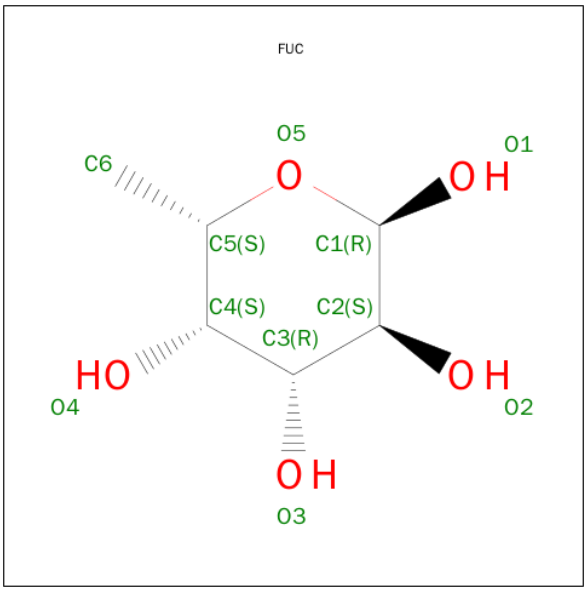
- Molecule 9 is a polymer of unknown type called SUGAR (2-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
9	B	2	Total	C	N	O	0	0
			24	14	1	9		

- Molecule 10 is a polymer of unknown type called SUGAR (3-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
10	C	3	Total	C	N	O	0	0
			46	25	2	19		
10	E	3	Total	C	N	O	0	0
			46	25	2	19		

- Molecule 11 is SUGAR (ALPHA-L-FUCOSE) (three-letter code: FUC) (formula: $C_6H_{12}O_5$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
11	D	1	Total C O 10 6 4	0	0

- Molecule 12 is a polymer of unknown type called SUGAR (3-MER).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
12	E	3	Total C N O 38 22 2 14	0	0
12	F	3	Total C N O 38 22 2 14	0	0

- Molecule 13 is water.

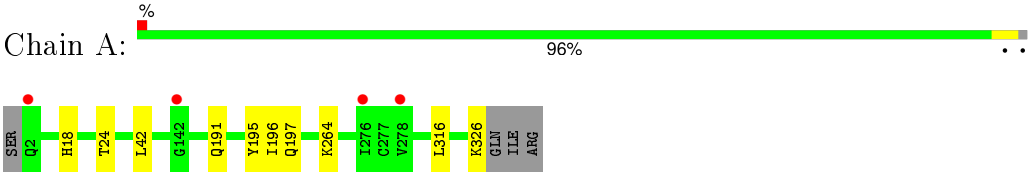
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
13	A	20	Total O 20 20	0	0
13	B	5	Total O 5 5	0	0
13	C	17	Total O 17 17	0	0
13	D	9	Total O 9 9	0	0
13	E	12	Total O 12 12	0	0
13	F	8	Total O 8 8	0	0

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.

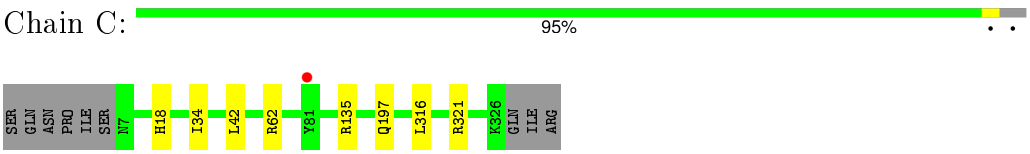
- Molecule 1: HEMAGGLUTININ

Chain A:



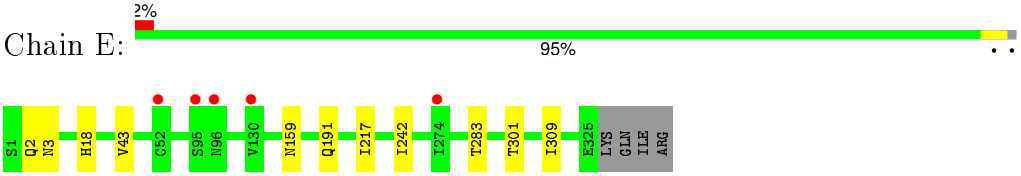
- Molecule 1: HEMAGGLUTININ

Chain C:



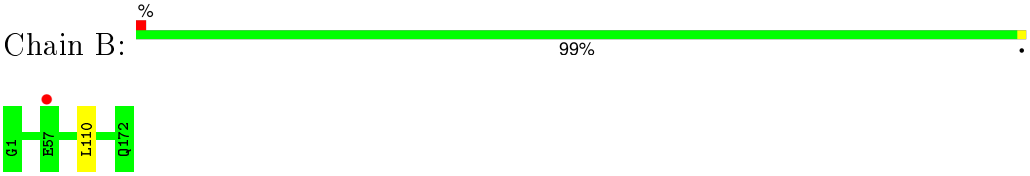
- Molecule 1: HEMAGGLUTININ

Chain E:



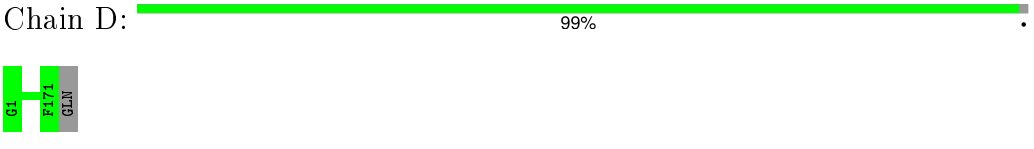
- Molecule 2: HEMAGGLUTININ

Chain B:



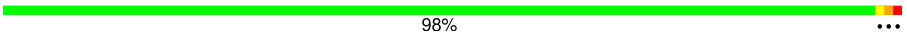
- Molecule 2: HEMAGGLUTININ

Chain D:



- Molecule 2: HEMAGGLUTININ

Chain F:




G1
R58
T59
D158
I161
Q172

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	78.81Å 129.94Å 194.12Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	107.98 – 3.00 46.67 – 3.00	Depositor EDS
% Data completeness (in resolution range)	95.8 (107.98-3.00) 95.9 (46.67-3.00)	Depositor EDS
R_{merge}	0.13	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.10 (at 3.01Å)	Xtriage
Refinement program	REFMAC 5.8.0069	Depositor
R, R_{free}	0.275 , 0.327 0.272 , 0.321	Depositor DCC
R_{free} test set	1958 reflections (5.28%)	DCC
Wilson B-factor (Å^2)	59.3	Xtriage
Anisotropy	0.951	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.29 , 23.8	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.44$, $< L^2 > = 0.27$	Xtriage
Outliers	4 of 39035 reflections (0.010%)	Xtriage
F_o,F_c correlation	0.88	EDS
Total number of atoms	12524	wwPDB-VP
Average B, all atoms (Å^2)	83.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.52% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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: BMA, NAG, SIA, GAL, FUC, 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	0.25	0/2589	0.44	0/3514
1	C	0.25	0/2543	0.43	0/3449
1	E	0.26	0/2580	0.44	0/3504
2	B	0.27	0/1424	0.42	0/1915
2	D	0.26	0/1409	0.42	0/1897
2	F	0.26	0/1428	0.42	0/1920
All	All	0.26	0/11973	0.43	0/16199

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
5	C	1	0
5	E	1	0
6	A	1	0
9	B	1	0
12	E	2	0
12	F	1	0
All	All	7	0

There are no bond length outliers.

There are no bond angle outliers.

All (7) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
6	A	600	FUC	C1
9	B	200	FUC	C1
5	C	434	MAN	C1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atom
5	E	437	MAN	C1
12	E	600	FUC	C1
12	E	602	NAG	C1
12	F	200	FUC	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	2535	0	2484	2	0
1	C	2492	0	2435	2	0
1	E	2526	0	2469	3	0
2	B	1399	0	1317	0	0
2	D	1384	0	1294	0	0
2	F	1403	0	1324	2	0
3	A	14	0	13	0	0
3	C	28	0	26	0	0
3	D	28	0	26	0	0
3	E	28	0	26	0	0
4	A	56	0	50	0	0
4	C	28	0	25	0	0
4	E	28	0	25	0	0
5	A	61	0	52	0	0
5	C	61	0	52	0	0
5	E	61	0	52	0	0
6	A	49	0	43	0	0
7	A	39	0	34	0	0
8	A	31	0	26	0	0
9	B	24	0	22	0	0
10	C	46	0	40	0	0
10	E	46	0	40	0	0
11	D	10	0	10	0	0
12	E	38	0	34	0	0
12	F	38	0	34	1	0
13	A	20	0	0	0	0
13	B	5	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
13	C	17	0	0	0	0
13	D	9	0	0	0	0
13	E	12	0	0	0	0
13	F	8	0	0	0	0
All	All	12524	0	11953	10	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 0.

All (10) 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:283:THR:HG22	1:E:301:THR:HG22	1.88	0.55
2:F:158:ASP:O	2:F:161:ILE:HG22	2.12	0.49
1:C:34:ILE:HD11	1:C:321:ARG:HD2	1.95	0.48
1:C:42:LEU:HD11	1:C:316:LEU:HB2	1.95	0.48
1:A:42:LEU:HD11	1:A:316:LEU:HB2	1.94	0.48
1:A:195:TYR:O	1:A:197:GLN:N	2.48	0.47
2:F:58:ARG:O	2:F:59:THR:HG23	2.15	0.47
12:F:201:NAG:O4	12:F:202:NAG:O7	2.33	0.47
1:E:191:GLN:HG2	1:E:217:ILE:HD11	1.99	0.45
1:E:43:VAL:HG22	1:E:309:ILE:HD11	2.02	0.42

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	323/329 (98%)	303 (94%)	19 (6%)	1 (0%)	46 84
1	C	317/329 (96%)	302 (95%)	14 (4%)	1 (0%)	46 84

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	E	323/329 (98%)	307 (95%)	15 (5%)	1 (0%)	46 84
2	B	170/172 (99%)	162 (95%)	8 (5%)	0	100 100
2	D	169/172 (98%)	158 (94%)	11 (6%)	0	100 100
2	F	170/172 (99%)	157 (92%)	12 (7%)	1 (1%)	30 72
All	All	1472/1503 (98%)	1389 (94%)	79 (5%)	4 (0%)	46 84

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	62	ARG
2	F	58	ARG
1	E	3	ASN
1	A	196	ILE

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	288/292 (99%)	283 (98%)	5 (2%)	68 91
1	C	281/292 (96%)	278 (99%)	3 (1%)	80 94
1	E	286/292 (98%)	282 (99%)	4 (1%)	74 93
2	B	145/146 (99%)	144 (99%)	1 (1%)	88 96
2	D	143/146 (98%)	143 (100%)	0	100 100
2	F	146/146 (100%)	144 (99%)	2 (1%)	74 93
All	All	1289/1314 (98%)	1274 (99%)	15 (1%)	78 94

All (15) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	18	HIS
1	A	24	THR
1	A	191	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	264	LYS
1	A	326	LYS
2	B	110	LEU
1	C	18	HIS
1	C	135	ARG
1	C	197	GLN
1	E	2	GLN
1	E	18	HIS
1	E	159	ASN
1	E	242	ILE
2	F	58	ARG
2	F	59	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	197	GLN

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

46 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	NAG	A	421	1,4	14,14,15	0.70	0	15,19,21	1.02	1 (6%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
4	NAG	A	422	4	14,14,15	0.48	0	15,19,21	0.58	0
5	NAG	A	431	1,5	14,14,15	0.49	0	15,19,21	0.78	0
5	NAG	A	432	5	14,14,15	0.47	0	15,19,21	0.61	0
5	BMA	A	433	5	11,11,12	0.50	0	14,15,17	1.71	2 (14%)
5	MAN	A	434	5	11,11,12	0.57	0	14,15,17	1.03	1 (7%)
5	MAN	A	437	5	11,11,12	0.55	0	14,15,17	2.53	3 (21%)
4	NAG	A	441	1,4	14,14,15	0.38	0	15,19,21	1.44	3 (20%)
4	NAG	A	442	4	14,14,15	0.52	0	15,19,21	1.22	3 (20%)
6	FUC	A	600	6	10,10,11	0.57	0	14,14,16	0.71	0
6	NAG	A	601	1,6	14,14,15	0.51	0	15,19,21	0.85	0
6	NAG	A	602	6	14,14,15	0.46	0	15,19,21	0.76	0
6	BMA	A	603	6	11,11,12	0.39	0	14,15,17	1.89	2 (14%)
7	NAG	A	631	1,7	14,14,15	0.51	0	15,19,21	0.60	0
7	NAG	A	632	7	14,14,15	0.63	0	15,19,21	1.37	1 (6%)
7	BMA	A	633	7	11,11,12	0.28	0	14,15,17	0.66	0
8	SIA	A	701	8	16,20,21	0.23	0	18,28,31	0.88	2 (11%)
8	GAL	A	702	8	11,11,12	0.59	0	14,15,17	0.73	0
9	FUC	B	200	9	10,10,11	0.60	0	14,14,16	1.08	1 (7%)
9	NAG	B	201	9,2	14,14,15	0.53	0	15,19,21	0.86	0
4	NAG	C	411	1,4	14,14,15	0.47	0	15,19,21	0.86	0
4	NAG	C	412	4	14,14,15	0.49	0	15,19,21	0.66	0
5	NAG	C	431	1,5	14,14,15	0.47	0	15,19,21	0.61	0
5	NAG	C	432	5	14,14,15	0.49	0	15,19,21	0.58	0
5	BMA	C	433	5	11,11,12	0.53	0	14,15,17	1.48	3 (21%)
5	MAN	C	434	5	11,11,12	0.51	0	14,15,17	2.51	3 (21%)
5	MAN	C	437	5	11,11,12	0.51	0	14,15,17	1.15	1 (7%)
10	SIA	C	701	10	16,20,21	0.32	0	18,28,31	0.91	2 (11%)
10	GAL	C	702	10	11,11,12	0.63	0	14,15,17	1.05	1 (7%)
10	NAG	C	703	10	15,15,15	0.48	0	17,21,21	1.01	1 (5%)
5	NAG	E	431	1,5	14,14,15	0.49	0	15,19,21	0.81	0
5	NAG	E	432	5	14,14,15	0.51	0	15,19,21	0.59	0
5	BMA	E	433	5	11,11,12	0.44	0	14,15,17	1.14	2 (14%)
5	MAN	E	434	5	11,11,12	0.60	0	14,15,17	1.49	2 (14%)
5	MAN	E	437	5	11,11,12	0.48	0	14,15,17	1.91	2 (14%)
12	FUC	E	600	12	10,10,11	0.54	0	14,14,16	1.42	2 (14%)
12	NAG	E	601	1,12	14,14,15	0.52	0	15,19,21	0.76	0
12	NAG	E	602	12	14,14,15	0.72	0	15,19,21	1.07	1 (6%)
4	NAG	E	631	1,4	14,14,15	0.57	0	15,19,21	0.78	1 (6%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
4	NAG	E	632	4	14,14,15	0.62	0	15,19,21	1.07	1 (6%)
10	SIA	E	701	10	16,20,21	0.32	0	18,28,31	1.33	2 (11%)
10	GAL	E	702	10	11,11,12	0.57	0	14,15,17	0.87	1 (7%)
10	NAG	E	703	10	15,15,15	0.42	0	17,21,21	0.77	0
12	FUC	F	200	12	10,10,11	0.60	0	14,14,16	1.11	1 (7%)
12	NAG	F	201	12,2	14,14,15	0.46	0	15,19,21	0.99	0
12	NAG	F	202	12	14,14,15	0.55	0	15,19,21	1.41	1 (6%)

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	A	421	1,4	-	0/6/23/26	0/1/1/1
4	NAG	A	422	4	-	0/6/23/26	0/1/1/1
5	NAG	A	431	1,5	-	0/6/23/26	0/1/1/1
5	NAG	A	432	5	-	0/6/23/26	0/1/1/1
5	BMA	A	433	5	-	0/2/19/22	0/1/1/1
5	MAN	A	434	5	-	0/2/19/22	0/1/1/1
5	MAN	A	437	5	-	0/2/19/22	0/1/1/1
4	NAG	A	441	1,4	-	0/6/23/26	0/1/1/1
4	NAG	A	442	4	-	0/6/23/26	0/1/1/1
6	FUC	A	600	6	1/1/4/5	0/0/17/20	0/1/1/1
6	NAG	A	601	1,6	-	0/6/23/26	0/1/1/1
6	NAG	A	602	6	-	0/6/23/26	0/1/1/1
6	BMA	A	603	6	-	0/2/19/22	0/1/1/1
7	NAG	A	631	1,7	-	0/6/23/26	0/1/1/1
7	NAG	A	632	7	-	0/6/23/26	0/1/1/1
7	BMA	A	633	7	-	0/2/19/22	0/1/1/1
8	SIA	A	701	8	-	0/14/34/38	0/1/1/1
8	GAL	A	702	8	-	0/2/19/22	0/1/1/1
9	FUC	B	200	9	1/1/4/5	0/0/17/20	0/1/1/1
9	NAG	B	201	9,2	-	0/6/23/26	0/1/1/1
4	NAG	C	411	1,4	-	0/6/23/26	0/1/1/1
4	NAG	C	412	4	-	0/6/23/26	0/1/1/1
5	NAG	C	431	1,5	-	0/6/23/26	0/1/1/1
5	NAG	C	432	5	-	0/6/23/26	0/1/1/1
5	BMA	C	433	5	-	0/2/19/22	0/1/1/1
5	MAN	C	434	5	1/1/4/5	0/2/19/22	0/1/1/1
5	MAN	C	437	5	-	0/2/19/22	1/1/1/1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
10	SIA	C	701	10	-	0/14/34/38	0/1/1/1
10	GAL	C	702	10	-	0/2/19/22	0/1/1/1
10	NAG	C	703	10	-	0/6/26/26	0/1/1/1
5	NAG	E	431	1,5	-	0/6/23/26	0/1/1/1
5	NAG	E	432	5	-	0/6/23/26	0/1/1/1
5	BMA	E	433	5	-	0/2/19/22	0/1/1/1
5	MAN	E	434	5	-	0/2/19/22	0/1/1/1
5	MAN	E	437	5	1/1/4/5	0/2/19/22	0/1/1/1
12	FUC	E	600	12	1/1/4/5	0/0/17/20	0/1/1/1
12	NAG	E	601	1,12	-	0/6/23/26	0/1/1/1
12	NAG	E	602	12	1/1/5/7	0/6/23/26	0/1/1/1
4	NAG	E	631	1,4	-	0/6/23/26	0/1/1/1
4	NAG	E	632	4	-	0/6/23/26	0/1/1/1
10	SIA	E	701	10	-	0/14/34/38	0/1/1/1
10	GAL	E	702	10	-	0/2/19/22	0/1/1/1
10	NAG	E	703	10	-	0/6/26/26	0/1/1/1
12	FUC	F	200	12	1/1/4/5	0/0/17/20	0/1/1/1
12	NAG	F	201	12,2	-	0/6/23/26	0/1/1/1
12	NAG	F	202	12	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

All (46) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
10	E	701	SIA	C7-C6-C5	-3.28	109.36	114.32
4	A	441	NAG	C4-C3-C2	-2.81	106.86	111.23
10	C	701	SIA	C7-C6-C5	-2.31	110.83	114.32
12	F	200	FUC	O5-C1-C2	-2.06	107.51	110.86
8	A	701	SIA	C7-C6-C5	-2.05	111.22	114.32
5	E	433	BMA	C1-C2-C3	2.09	112.01	109.54
4	A	441	NAG	C8-C7-N2	2.12	120.16	116.11
8	A	701	SIA	O6-C6-C5	2.12	111.96	108.48
4	A	442	NAG	C2-N2-C7	2.18	125.84	123.04
4	E	631	NAG	C4-C3-C2	2.19	114.63	111.23
4	A	442	NAG	C3-C2-N2	2.22	115.88	110.56
5	C	433	BMA	C2-C3-C4	2.35	115.04	111.04
4	A	442	NAG	C1-O5-C5	2.38	115.27	112.25
5	E	433	BMA	C3-C4-C5	2.41	114.40	110.20
5	C	433	BMA	C3-C4-C5	2.46	114.48	110.20
10	E	702	GAL	C1-C2-C3	2.46	112.45	109.54
10	C	701	SIA	O6-C6-C5	2.52	112.61	108.48
5	C	434	MAN	C1-C2-C3	2.55	112.56	109.54

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	A	433	BMA	C2-C3-C4	2.60	115.45	111.04
12	E	600	FUC	C1-C2-C3	2.67	112.70	109.54
5	E	434	MAN	C1-C2-C3	2.77	112.82	109.54
10	C	703	NAG	C4-C3-C2	2.79	114.30	110.43
12	E	602	NAG	C4-C3-C2	2.82	115.61	111.23
5	A	434	MAN	C1-O5-C5	2.85	115.87	112.25
9	B	200	FUC	C1-O5-C5	2.87	116.81	112.38
4	E	632	NAG	C4-C3-C2	2.90	115.75	111.23
4	A	441	NAG	C1-O5-C5	3.19	116.30	112.25
4	A	421	NAG	C4-C3-C2	3.24	116.26	111.23
5	A	437	MAN	O5-C1-C2	3.26	116.15	110.86
10	C	702	GAL	C1-C2-C3	3.37	113.52	109.54
12	E	600	FUC	C1-O5-C5	3.68	118.06	112.38
5	C	433	BMA	C1-C2-C3	3.75	113.98	109.54
12	F	202	NAG	C2-N2-C7	3.85	127.99	123.04
5	C	437	MAN	C1-O5-C5	3.97	117.29	112.25
6	A	603	BMA	C3-C4-C5	4.02	117.21	110.20
10	E	701	SIA	O6-C6-C5	4.13	115.25	108.48
7	A	632	NAG	C4-C3-C2	4.18	117.73	111.23
5	E	437	MAN	C1-O5-C5	4.24	117.63	112.25
5	A	437	MAN	C1-C2-C3	4.44	114.79	109.54
5	E	434	MAN	C1-O5-C5	4.56	118.03	112.25
5	A	433	BMA	C1-C2-C3	4.96	115.41	109.54
5	C	434	MAN	C1-O5-C5	5.19	118.83	112.25
6	A	603	BMA	C1-O5-C5	5.24	118.89	112.25
5	E	437	MAN	O5-C1-C2	5.31	119.47	110.86
5	C	434	MAN	O5-C1-C2	7.25	122.62	110.86
5	A	437	MAN	C1-O5-C5	7.46	121.72	112.25

All (7) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
5	C	434	MAN	C1
9	B	200	FUC	C1
12	E	600	FUC	C1
5	E	437	MAN	C1
6	A	600	FUC	C1
12	E	602	NAG	C1
12	F	200	FUC	C1

There are no torsion outliers.

All (1) ring outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	C	437	MAN	C1-C2-C3-C4-C5-O5

2 monomers are involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
12	F	201	NAG	1	0
12	F	202	NAG	1	0

5.6 Ligand geometry (i)

8 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	411	1	14,14,15	0.52	0	15,19,21	0.76	0
3	NAG	C	631	-	14,14,15	0.59	0	15,19,21	1.39	1 (6%)
3	NAG	C	632	-	14,14,15	0.57	0	15,19,21	0.81	0
11	FUC	D	200	-	10,10,11	0.62	0	14,14,16	0.84	0
3	NAG	D	201	2	14,14,15	0.61	0	15,19,21	1.15	2 (13%)
3	NAG	D	202	-	14,14,15	0.51	0	15,19,21	0.64	0
3	NAG	E	441	1	14,14,15	0.50	0	15,19,21	0.99	0
3	NAG	E	621	1	14,14,15	0.65	0	15,19,21	1.29	2 (13%)

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	411	1	-	0/6/23/26	0/1/1/1
3	NAG	C	631	-	-	0/6/23/26	0/1/1/1
3	NAG	C	632	-	-	0/6/23/26	0/1/1/1
11	FUC	D	200	-	-	0/0/17/20	0/1/1/1
3	NAG	D	201	2	-	0/6/23/26	0/1/1/1
3	NAG	D	202	-	-	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	E	441	1	-	0/6/23/26	0/1/1/1
3	NAG	E	621	1	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	E	621	NAG	C3-C4-C5	2.03	113.74	110.20
3	D	201	NAG	C3-C4-C5	2.22	114.06	110.20
3	D	201	NAG	C4-C3-C2	3.02	115.93	111.23
3	E	621	NAG	C4-C3-C2	3.82	117.17	111.23
3	C	631	NAG	C3-C4-C5	4.25	117.61	110.20

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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^{th} 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^2)	Q<0.9
1	A	325/329 (98%)	0.00	4 (1%) 81 55	60, 80, 107, 123	0
1	C	320/329 (97%)	-0.10	1 (0%) 94 84	63, 78, 95, 106	0
1	E	325/329 (98%)	0.21	5 (1%) 76 49	67, 97, 118, 146	0
2	B	172/172 (100%)	-0.07	1 (0%) 90 73	55, 74, 90, 107	0
2	D	171/172 (99%)	-0.08	0 100 100	55, 74, 88, 93	0
2	F	172/172 (100%)	-0.06	0 100 100	58, 73, 91, 99	0
All	All	1485/1503 (98%)	-0.00	11 (0%) 89 70	55, 79, 110, 146	0

All (11) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	278	VAL	3.5
1	A	2	GLN	3.0
1	A	142	GLY	2.6
1	E	96	ASN	2.5
2	B	57	GLU	2.5
1	E	130	VAL	2.5
1	E	274	ILE	2.3
1	E	95	SER	2.3
1	A	276	ILE	2.3
1	E	52	CYS	2.1
1	C	81	TYR	2.1

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 (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. 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^{th} 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^2)	Q<0.9
9	FUC	B	200	10/11	0.87	0.37	3.34	101,103,107,110	0
9	NAG	B	201	14/15	0.85	0.25	2.34	96,101,104,107	0
5	NAG	C	431	14/15	0.89	0.25	1.03	80,82,86,87	0
5	NAG	A	432	14/15	0.82	0.29	0.42	115,119,126,131	0
4	NAG	C	411	14/15	0.63	0.27	0.37	101,106,110,114	0
6	NAG	A	601	14/15	0.89	0.20	0.14	97,101,107,109	0
10	SIA	E	701	20/21	0.87	0.25	0.12	119,125,131,131	0
6	NAG	A	602	14/15	0.90	0.22	-0.17	112,115,120,121	0
12	NAG	F	201	14/15	0.89	0.17	-0.25	95,101,104,107	0
8	SIA	A	701	20/21	0.88	0.19	-0.28	72,77,83,85	0
10	SIA	C	701	20/21	0.90	0.19	-0.76	83,86,93,95	0
5	NAG	E	432	14/15	0.91	0.20	-1.05	103,109,114,117	0
10	GAL	C	702	11/12	0.91	0.12	-	95,96,100,101	0
6	BMA	A	603	11/12	0.84	0.20	-	124,130,135,135	0
4	NAG	A	442	14/15	0.64	0.28	-	91,97,101,103	0
5	BMA	C	433	11/12	0.88	0.17	-	101,105,110,111	0
4	NAG	C	412	14/15	0.65	0.37	-	117,120,126,128	0
5	BMA	E	433	11/12	0.76	0.14	-	119,122,124,128	0
5	NAG	C	432	14/15	0.93	0.30	-	89,91,96,98	0
10	NAG	E	703	15/15	0.74	0.30	-	138,143,150,151	0
4	NAG	E	632	14/15	0.84	0.20	-	99,101,104,107	0
5	BMA	A	433	11/12	0.75	0.24	-	134,139,147,148	0
10	GAL	E	702	11/12	0.88	0.25	-	127,129,135,135	0
12	NAG	E	601	14/15	0.79	0.27	-	107,111,115,118	0
4	NAG	A	422	14/15	0.51	0.55	-	127,130,136,138	0
5	MAN	A	434	11/12	0.73	0.37	-	144,153,163,167	0
12	FUC	E	600	10/11	0.77	0.27	-	114,117,120,122	0
4	NAG	A	421	14/15	0.83	0.32	-	114,118,124,125	0
7	NAG	A	631	14/15	0.89	0.29	-	90,93,98,99	0
12	NAG	F	202	14/15	0.73	0.29	-	100,111,118,121	0
4	NAG	E	631	14/15	0.88	0.14	-	89,94,98,98	0
7	BMA	A	633	11/12	0.75	0.30	-	105,111,115,119	0
5	MAN	C	437	11/12	0.76	0.19	-	108,110,112,115	0
5	MAN	E	437	11/12	0.76	0.24	-	126,130,132,133	0
7	NAG	A	632	14/15	0.74	0.42	-	98,105,109,109	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å^2)	Q<0.9
6	FUC	A	600	10/11	0.89	0.31	-	109,111,114,114	0
5	NAG	A	431	14/15	0.91	0.18	-	103,106,111,113	0
5	MAN	E	434	11/12	0.79	0.25	-	128,132,134,136	0
12	NAG	E	602	14/15	0.56	0.36	-	121,123,130,130	0
8	GAL	A	702	11/12	0.90	0.14	-	81,82,85,85	0
4	NAG	A	441	14/15	0.89	0.20	-	85,89,92,93	0
5	NAG	E	431	14/15	0.94	0.17	-	100,103,105,107	0
12	FUC	F	200	10/11	0.82	0.29	-	99,102,106,108	0
5	MAN	C	434	11/12	0.81	0.21	-	104,111,118,120	0
10	NAG	C	703	15/15	0.80	0.35	-	104,110,115,117	0
5	MAN	A	437	11/12	0.60	0.27	-	143,146,158,159	0

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. 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, 95th 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^2)	Q<0.9
3	NAG	E	441	14/15	0.87	0.26	1.46	91,95,97,98	0
3	NAG	C	631	14/15	0.82	0.23	0.78	81,85,87,87	0
11	FUC	D	200	10/11	0.73	0.28	-	117,122,124,127	0
3	NAG	A	411	14/15	0.87	0.32	-	118,122,124,128	0
3	NAG	C	632	14/15	0.83	0.24	-	96,102,103,104	0
3	NAG	D	202	14/15	0.63	0.56	-	116,124,129,130	0
3	NAG	D	201	14/15	0.83	0.30	-	93,98,100,102	0
3	NAG	E	621	14/15	0.65	0.45	-	98,101,108,109	0

6.5 Other polymers (i)

There are no such residues in this entry.