wwPDB X-ray Structure Validation Summary Report (i)

Feb 1, 2016 – 08:14 PM GMT

PDB ID	:	4QY1
Title	:	Structure of H10 from human-infecting H10N8 in complex with avian receptor
Authors	:	Wang, M.; Zhang, W.; Qi, J.; Wang, F.; Zhou, J.; Bi, Y.; Wu, Y.; Sun, H.; Liu, J.; Huang, C.; Li, X.; Yan, J.; Shu, Y.; Shi, Y.; Gao, G.F.
Deposited on	:	2014-07-23
Resolution	:	2.59 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
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:

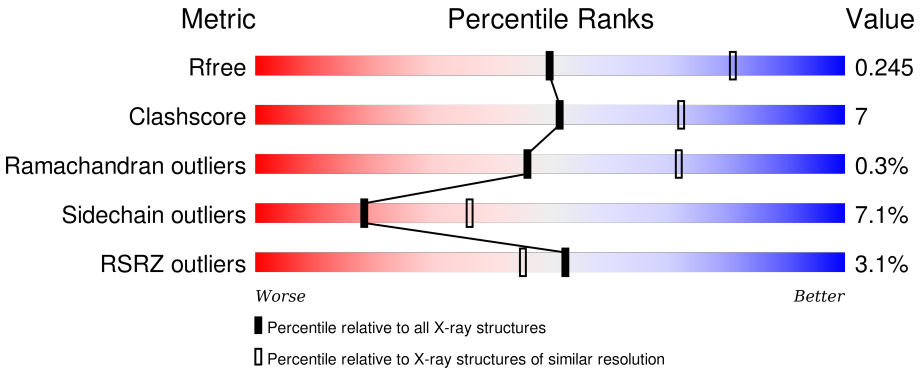
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	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 2.59 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 >=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	318	78% 20% •
1	C	318	2% 81% 17% •
1	E	318	79% 19% •
1	G	318	2% 81% 18% •
1	I	318	2% 82% 15% •

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	K	318	2% 81% 18%
1	M	318	11% 83% 15%
1	O	318	9% 78% 21%
1	Q	318	2% 76% 22%
1	S	318	7% 82% 16%
1	U	318	2% 80% 17%
1	W	318	3% 79% 19%
2	B	174	3% 83% 14%
2	D	174	2% 82% 16%
2	F	174	2% 82% 16%
2	H	174	% 82% 16%
2	J	174	3% 86% 13%
2	L	174	3% 84% 14%
2	N	174	2% 84% 14%
2	P	174	3% 82% 16%
2	R	174	5% 82% 17%
2	T	174	2% 82% 16%
2	V	174	2% 80% 17%
2	X	174	2% 82% 17%

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	U	601	-	-	-	X
3	NAG	W	601	-	-	-	X

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 48138 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
			2437	1506	449	465	17			
1	C	318	Total	C	N	O	S	0	0	0
			2437	1506	449	465	17			
1	E	318	Total	C	N	O	S	0	0	0
			2437	1506	449	465	17			
1	G	318	Total	C	N	O	S	0	0	0
			2437	1506	449	465	17			
1	I	318	Total	C	N	O	S	0	0	0
			2437	1506	449	465	17			
1	K	318	Total	C	N	O	S	0	0	0
			2437	1506	449	465	17			
1	M	318	Total	C	N	O	S	0	0	0
			2437	1506	449	465	17			
1	O	318	Total	C	N	O	S	0	0	0
			2437	1506	449	465	17			
1	Q	318	Total	C	N	O	S	0	0	0
			2437	1506	449	465	17			
1	S	318	Total	C	N	O	S	0	0	0
			2436	1506	449	464	17			
1	U	318	Total	C	N	O	S	0	0	0
			2437	1506	449	465	17			
1	W	318	Total	C	N	O	S	0	0	0
			2437	1506	449	465	17			

- Molecule 2 is a protein called hemagglutinin.

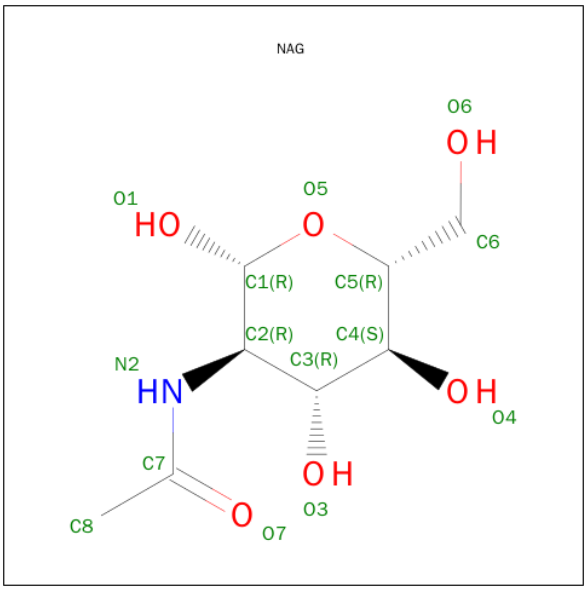
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	174	Total	C	N	O	S	0	0	0
			1402	866	243	285	8			
2	D	174	Total	C	N	O	S	0	0	0
			1402	866	243	285	8			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	F	174	Total	C	N	O	S	0	0	0
			1402	866	243	285	8			
2	H	174	Total	C	N	O	S	0	0	0
			1402	866	243	285	8			
2	J	174	Total	C	N	O	S	0	0	0
			1402	866	243	285	8			
2	L	174	Total	C	N	O	S	0	0	0
			1402	866	243	285	8			
2	N	174	Total	C	N	O	S	0	0	0
			1402	866	243	285	8			
2	P	174	Total	C	N	O	S	0	0	0
			1402	866	243	285	8			
2	R	174	Total	C	N	O	S	0	0	0
			1402	866	243	285	8			
2	T	174	Total	C	N	O	S	0	0	0
			1402	866	243	285	8			
2	V	174	Total	C	N	O	S	0	0	0
			1402	866	243	285	8			
2	X	174	Total	C	N	O	S	0	0	0
			1402	866	243	285	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		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	B	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	D	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		
3	F	1	Total	C	N	O	0	0
			14	8	1	5		
3	G	1	Total	C	N	O	0	0
			14	8	1	5		
3	H	1	Total	C	N	O	0	0
			14	8	1	5		
3	I	1	Total	C	N	O	0	0
			14	8	1	5		
3	J	1	Total	C	N	O	0	0
			14	8	1	5		
3	K	1	Total	C	N	O	0	0
			14	8	1	5		
3	L	1	Total	C	N	O	0	0
			14	8	1	5		
3	M	1	Total	C	N	O	0	0
			14	8	1	5		
3	N	1	Total	C	N	O	0	0
			14	8	1	5		
3	O	1	Total	C	N	O	0	0
			14	8	1	5		
3	P	1	Total	C	N	O	0	0
			14	8	1	5		
3	Q	1	Total	C	N	O	0	0
			14	8	1	5		
3	R	1	Total	C	N	O	0	0
			14	8	1	5		
3	S	1	Total	C	N	O	0	0
			14	8	1	5		
3	T	1	Total	C	N	O	0	0
			14	8	1	5		
3	U	1	Total	C	N	O	0	0
			14	8	1	5		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	V	1	Total C N O 14 8 1 5	0	0
3	W	1	Total C N O 14 8 1 5	0	0
3	X	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	E	2	Total C N O 28 16 2 10	0	0

- Molecule 5 is a polymer of unknown type called SUGAR (3-MER).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	E	3	Total C N O 45 25 2 18	0	0
5	G	3	Total C N O 45 25 2 18	0	0
5	I	3	Total C N O 45 25 2 18	0	0
5	U	3	Total C N O 45 25 2 18	0	0

- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	106	Total O 106 106	0	0
6	B	70	Total O 70 70	0	0
6	C	71	Total O 71 71	0	0
6	D	87	Total O 87 87	0	0
6	E	93	Total O 93 93	0	0
6	F	64	Total O 64 64	0	0
6	G	80	Total O 80 80	0	0

Continued on next page...

Continued from previous page...

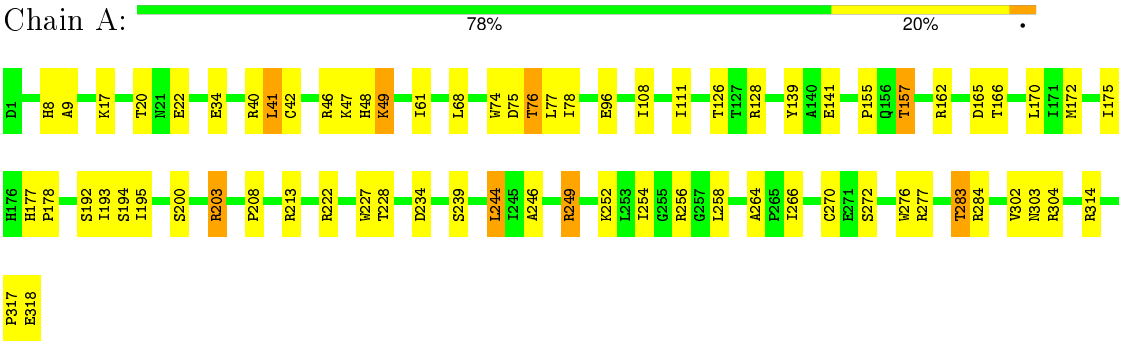
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	H	53	Total O 53 53	0	0
6	I	102	Total O 102 102	0	0
6	J	43	Total O 43 43	0	0
6	K	93	Total O 93 93	0	0
6	L	53	Total O 53 53	0	0
6	M	27	Total O 27 27	0	0
6	N	46	Total O 46 46	0	0
6	O	40	Total O 40 40	0	0
6	P	48	Total O 48 48	0	0
6	Q	84	Total O 84 84	0	0
6	R	42	Total O 42 42	0	0
6	S	64	Total O 64 64	0	0
6	T	49	Total O 49 49	0	0
6	U	76	Total O 76 76	0	0
6	V	44	Total O 44 44	0	0
6	W	50	Total O 50 50	0	0
6	X	28	Total O 28 28	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.

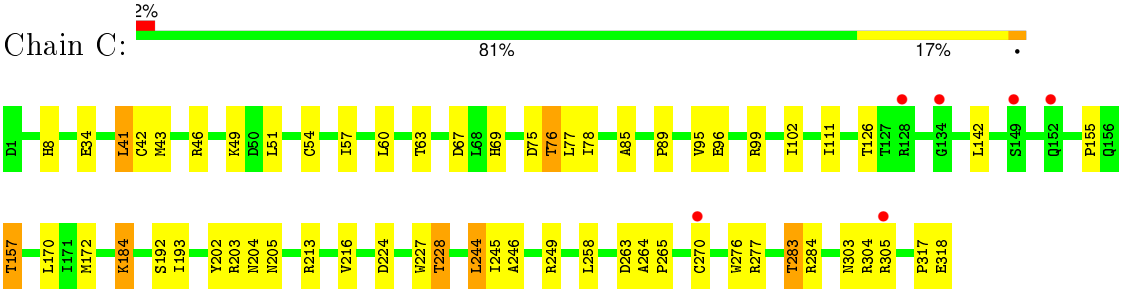
- Molecule 1: hemagglutinin

Chain A:



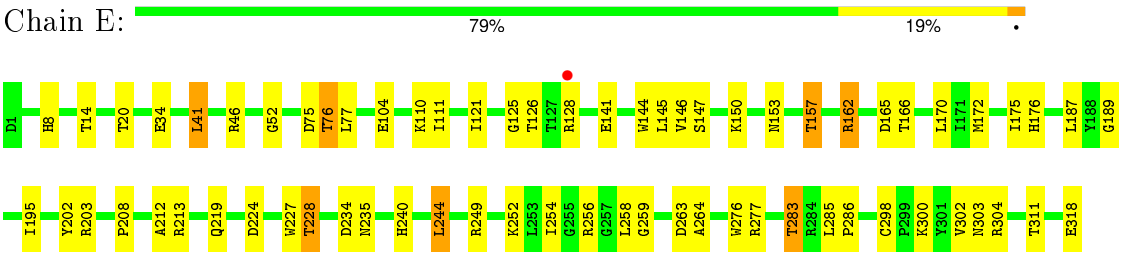
- Molecule 1: hemagglutinin

Chain C:



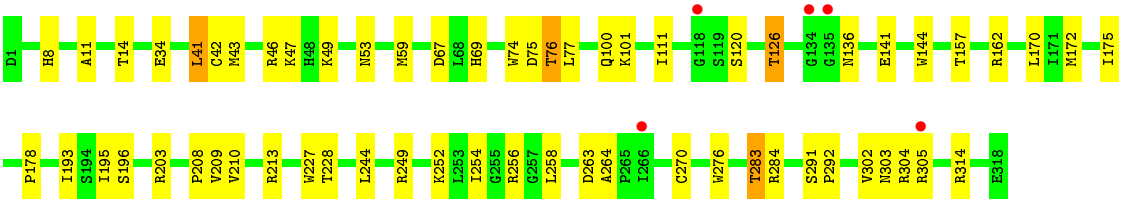
- Molecule 1: hemagglutinin

Chain E:



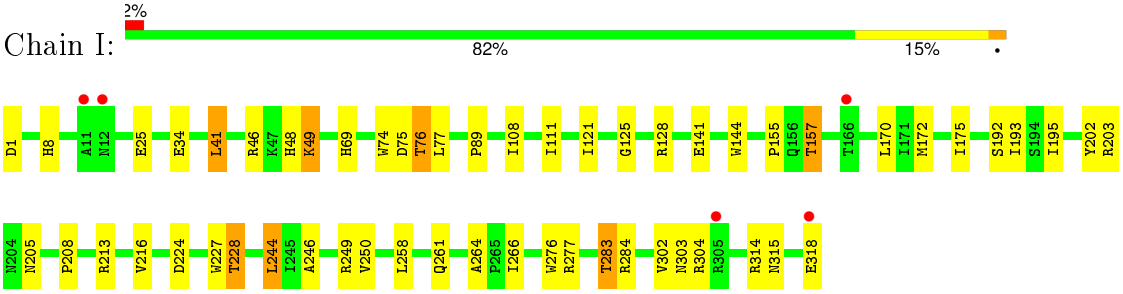
- Molecule 1: hemagglutinin

Chain G:



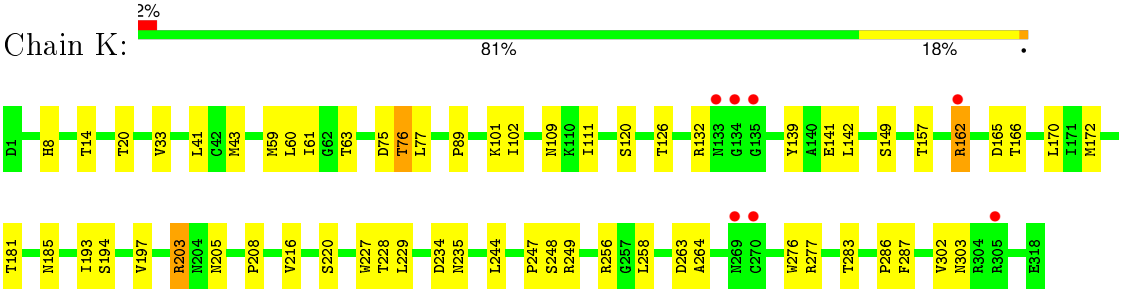
• Molecule 1: hemagglutinin

Chain I:



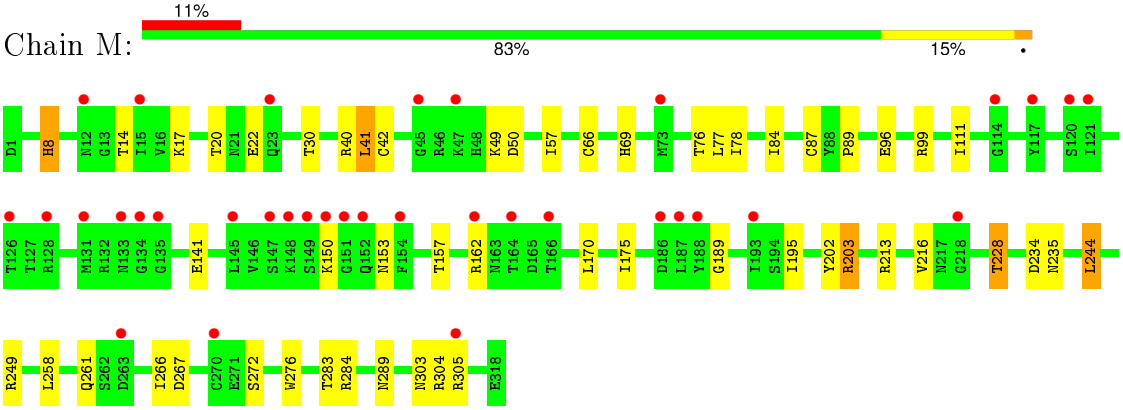
• Molecule 1: hemagglutinin

Chain K:



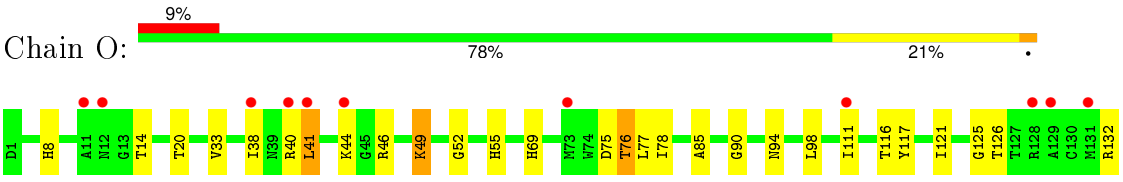
• Molecule 1: hemagglutinin

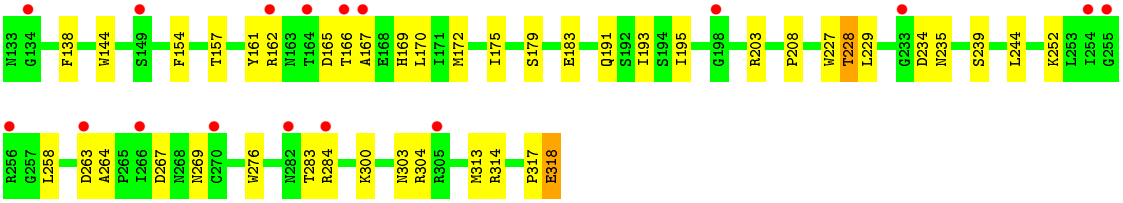
Chain M:



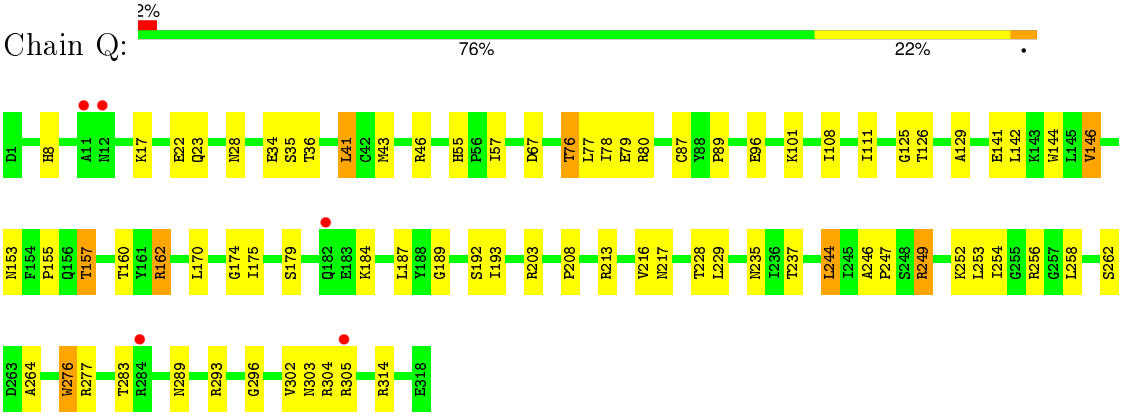
• Molecule 1: hemagglutinin

Chain O:

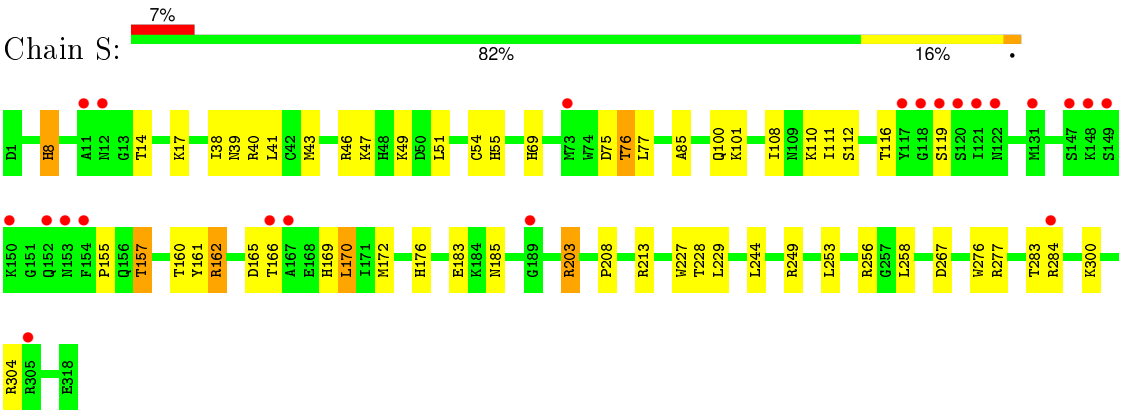




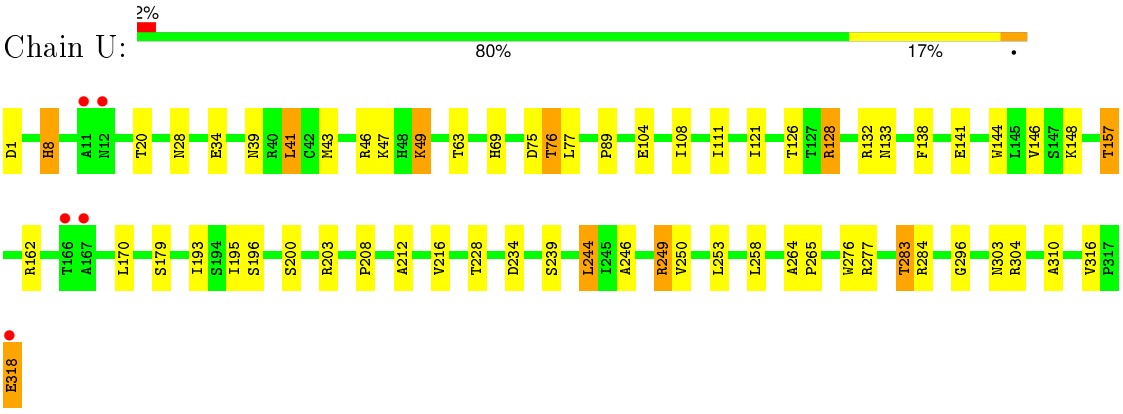
• Molecule 1: hemagglutinin



• Molecule 1: hemagglutinin

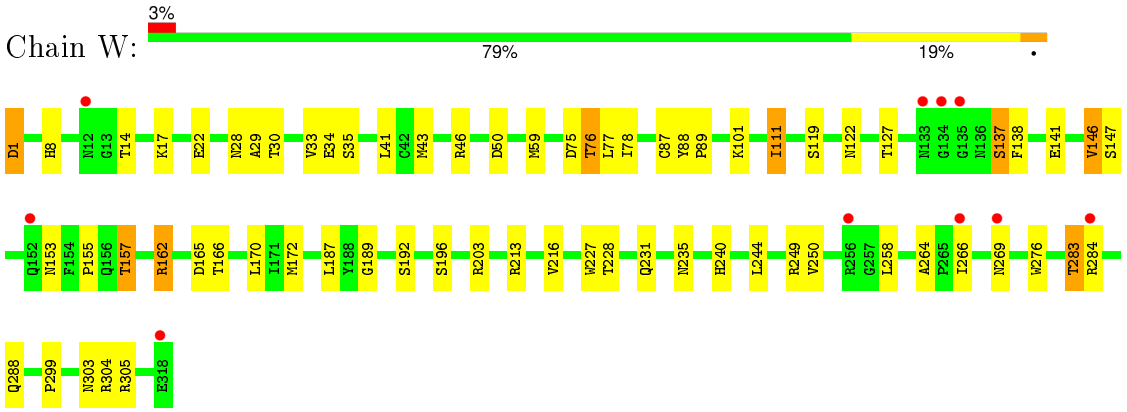


• Molecule 1: hemagglutinin



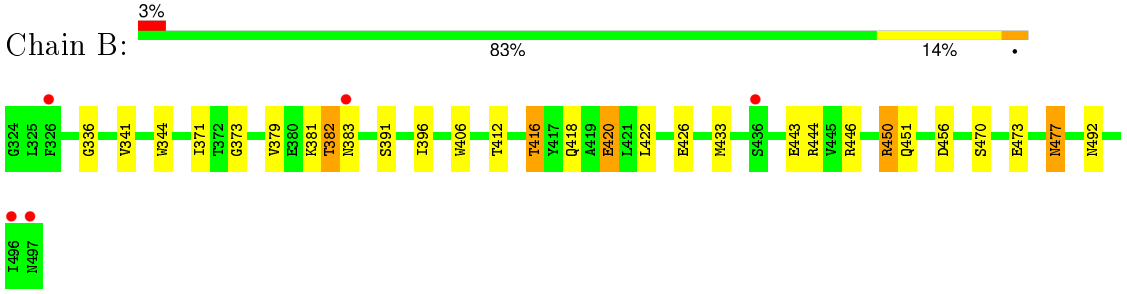
• Molecule 1: hemagglutinin

Chain W:



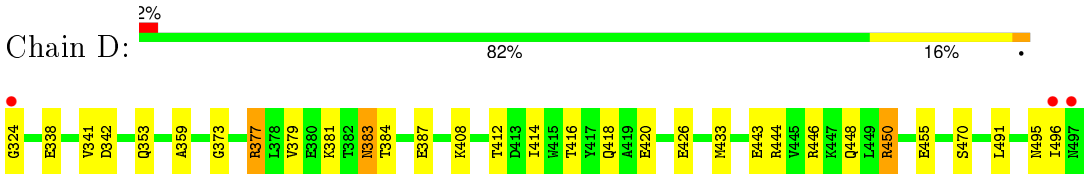
- Molecule 2: hemagglutinin

Chain B:



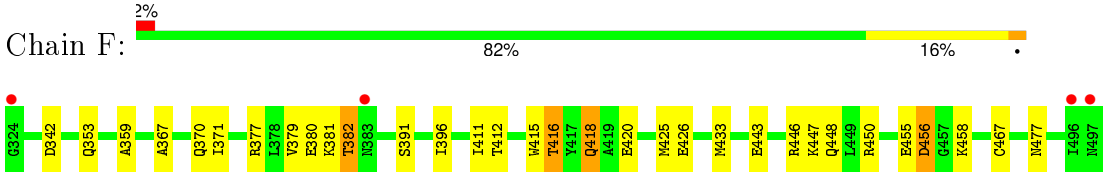
- Molecule 2: hemagglutinin

Chain D:



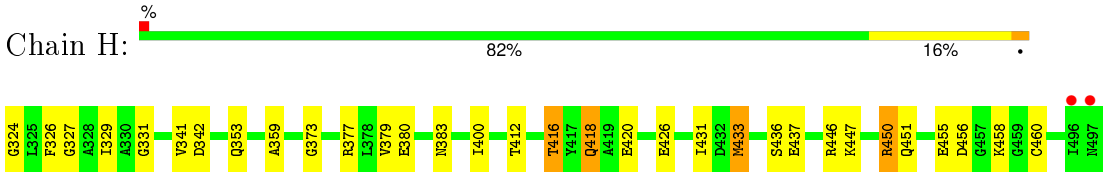
- Molecule 2: hemagglutinin

Chain F:



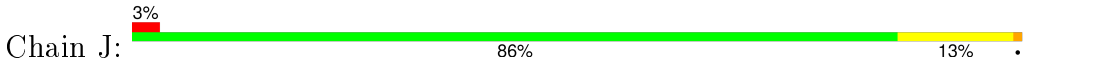
- Molecule 2: hemagglutinin

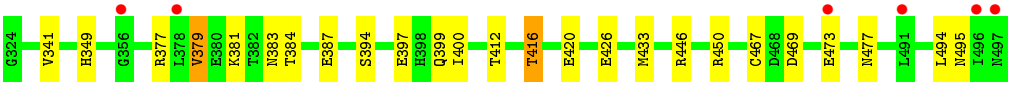
Chain H:



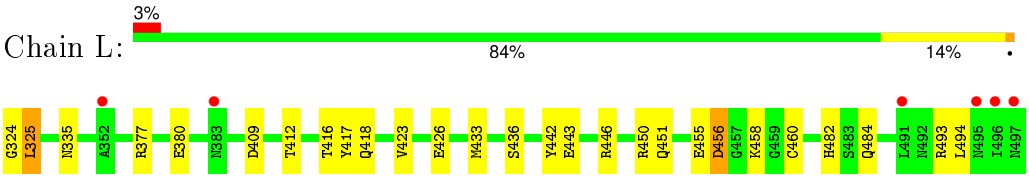
- Molecule 2: hemagglutinin

Chain J:

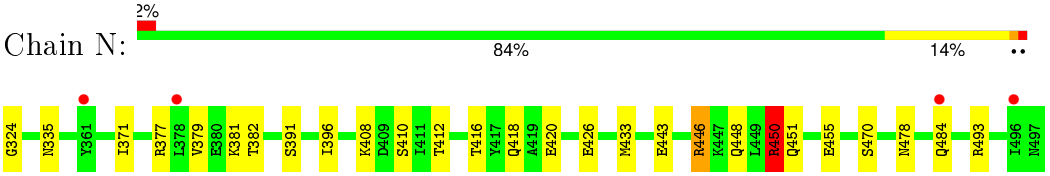




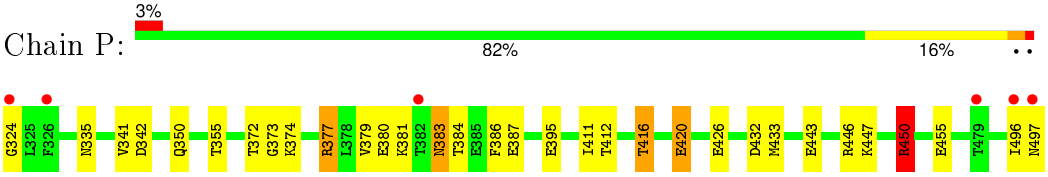
• Molecule 2: hemagglutinin



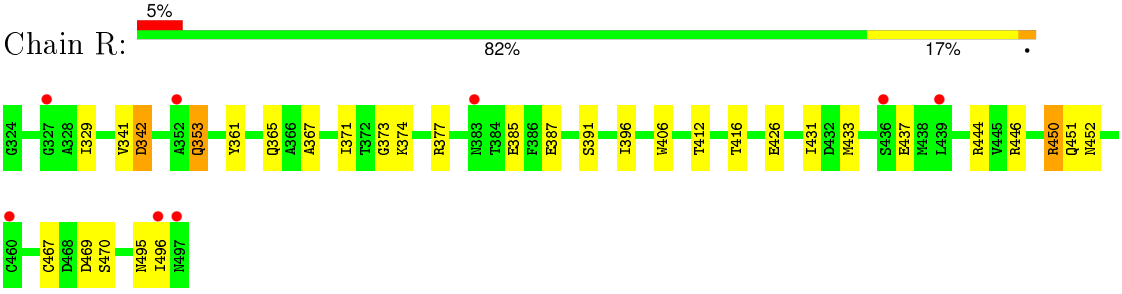
• Molecule 2: hemagglutinin



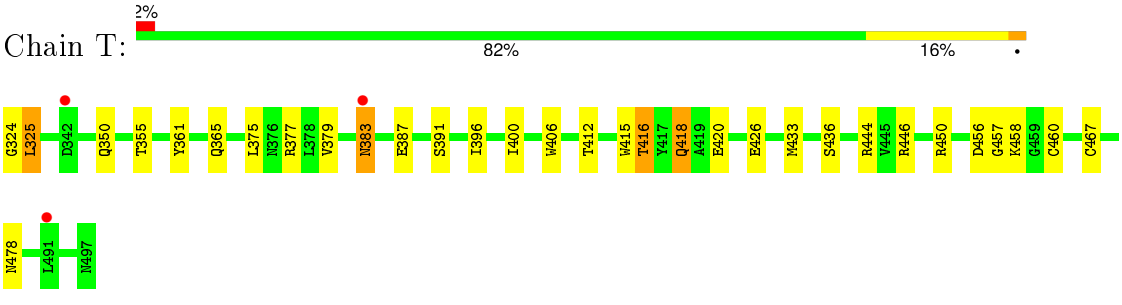
• Molecule 2: hemagglutinin



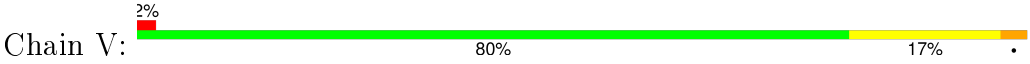
• Molecule 2: hemagglutinin

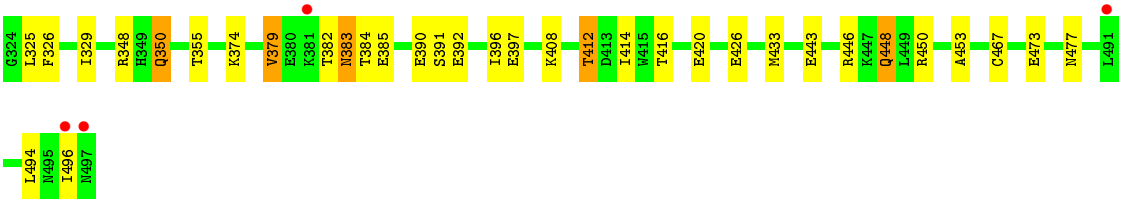


• Molecule 2: hemagglutinin

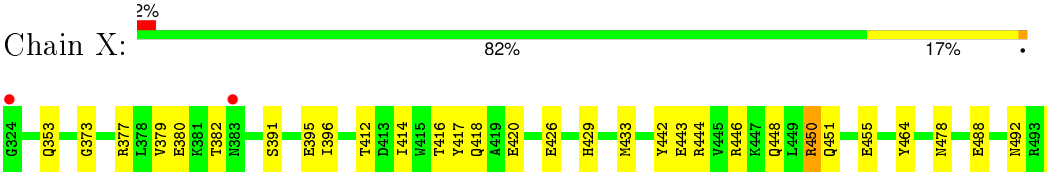


• Molecule 2: hemagglutinin





- Molecule 2: hemagglutinin



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	93.26Å 111.44Å 222.35Å 94.96° 101.17° 91.90°	Depositor
Resolution (Å)	37.07 – 2.59 37.07 – 2.59	Depositor EDS
% Data completeness (in resolution range)	97.2 (37.07-2.59) 86.9 (37.07-2.59)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.70 (at 2.61Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.2_1309)	Depositor
R, R_{free}	0.197 , 0.246 0.194 , 0.245	Depositor DCC
R_{free} test set	12085 reflections (5.02%)	DCC
Wilson B-factor (Å^2)	38.7	Xtriage
Anisotropy	0.572	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 42.6	EDS
Estimated twinning fraction	0.004 for h,-k,-h-l	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Outliers	0 of 262928 reflections	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	48138	wwPDB-VP
Average B, all atoms (Å^2)	38.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.62% 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: SIA, GAL, NAG

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	0.33	0/2486	0.50	0/3368
1	C	0.30	0/2486	0.48	0/3368
1	E	0.31	0/2486	0.49	0/3368
1	G	0.31	0/2486	0.48	0/3368
1	I	0.31	0/2486	0.49	0/3368
1	K	0.31	0/2486	0.49	0/3368
1	M	0.27	0/2486	0.45	0/3368
1	O	0.28	0/2486	0.46	0/3368
1	Q	0.31	0/2486	0.47	0/3368
1	S	0.29	0/2485	0.49	0/3366
1	U	0.30	0/2486	0.48	0/3368
1	W	0.29	0/2486	0.47	0/3368
2	B	0.33	0/1427	0.46	0/1926
2	D	0.34	0/1427	0.50	0/1926
2	F	0.34	0/1427	0.48	0/1926
2	H	0.33	0/1427	0.46	0/1926
2	J	0.31	0/1427	0.47	0/1926
2	L	0.32	0/1427	0.46	0/1926
2	N	0.31	0/1427	0.45	0/1926
2	P	0.33	0/1427	0.47	0/1926
2	R	0.31	0/1427	0.45	0/1926
2	T	0.32	0/1427	0.45	0/1926
2	V	0.31	0/1427	0.48	0/1926
2	X	0.30	0/1427	0.45	0/1926
All	All	0.31	0/46955	0.47	0/63526

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

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	2437	0	2389	41	0
1	C	2437	0	2389	31	0
1	E	2437	0	2387	40	0
1	G	2437	0	2389	36	0
1	I	2437	0	2389	34	0
1	K	2437	0	2389	27	0
1	M	2437	0	2389	26	0
1	O	2437	0	2389	39	0
1	Q	2437	0	2389	38	0
1	S	2436	0	2389	33	0
1	U	2437	0	2389	36	0
1	W	2437	0	2389	37	0
2	B	1402	0	1298	21	0
2	D	1402	0	1298	27	0
2	F	1402	0	1298	20	0
2	H	1402	0	1298	29	0
2	J	1402	0	1298	20	0
2	L	1402	0	1298	21	0
2	N	1402	0	1298	19	0
2	P	1402	0	1298	22	0
2	R	1402	0	1298	26	0
2	T	1402	0	1298	23	0
2	V	1402	0	1298	26	0
2	X	1402	0	1298	22	0
3	A	14	0	13	0	0
3	B	14	0	13	0	0
3	C	14	0	13	0	0
3	D	14	0	13	1	0
3	E	28	0	26	1	0
3	F	14	0	13	0	0
3	G	14	0	13	0	0
3	H	14	0	13	0	0
3	I	14	0	13	0	0
3	J	14	0	13	1	0
3	K	14	0	13	1	0
3	L	14	0	13	0	0
3	M	14	0	13	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	N	14	0	13	1	0
3	O	14	0	13	1	0
3	P	14	0	13	0	0
3	Q	14	0	13	0	0
3	R	14	0	13	0	0
3	S	14	0	13	0	0
3	T	14	0	13	1	0
3	U	14	0	13	1	0
3	V	14	0	13	0	0
3	W	14	0	13	0	0
3	X	14	0	13	0	0
4	E	28	0	25	1	0
5	E	45	0	38	4	0
5	G	45	0	38	4	0
5	I	45	0	38	3	0
5	U	45	0	38	4	0
6	A	106	0	0	8	0
6	B	70	0	0	2	0
6	C	71	0	0	5	0
6	D	87	0	0	9	0
6	E	93	0	0	11	0
6	F	64	0	0	3	0
6	G	80	0	0	11	0
6	H	53	0	0	6	0
6	I	102	0	0	8	0
6	J	43	0	0	8	0
6	K	93	0	0	3	0
6	L	53	0	0	3	0
6	M	27	0	0	5	0
6	N	46	0	0	4	0
6	O	40	0	0	5	0
6	P	48	0	0	6	0
6	Q	84	0	0	5	0
6	R	42	0	0	8	0
6	S	64	0	0	7	0
6	T	49	0	0	3	0
6	U	76	0	0	7	0
6	V	44	0	0	9	0
6	W	50	0	0	11	0
6	X	28	0	0	4	0
All	All	48138	0	44744	619	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 619 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)
2:L:409:ASP:OD2	6:L:738:HOH:O	1.81	0.98
1:A:96:GLU:OE1	6:A:712:HOH:O	1.86	0.92
2:R:444:ARG:O	6:R:723:HOH:O	1.89	0.89
2:V:390:GLU:OE2	6:V:709:HOH:O	1.90	0.89
1:G:11:ALA:O	6:G:712:HOH:O	1.91	0.88

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	316/318 (99%)	309 (98%)	7 (2%)	0	100	100
1	C	316/318 (99%)	306 (97%)	10 (3%)	0	100	100
1	E	316/318 (99%)	306 (97%)	10 (3%)	0	100	100
1	G	316/318 (99%)	310 (98%)	6 (2%)	0	100	100
1	I	316/318 (99%)	307 (97%)	9 (3%)	0	100	100
1	K	316/318 (99%)	306 (97%)	10 (3%)	0	100	100
1	M	316/318 (99%)	302 (96%)	13 (4%)	1 (0%)	46	72
1	O	316/318 (99%)	303 (96%)	13 (4%)	0	100	100
1	Q	316/318 (99%)	304 (96%)	12 (4%)	0	100	100
1	S	316/318 (99%)	305 (96%)	10 (3%)	1 (0%)	46	72
1	U	316/318 (99%)	307 (97%)	8 (2%)	1 (0%)	46	72
1	W	316/318 (99%)	305 (96%)	11 (4%)	0	100	100
2	B	172/174 (99%)	167 (97%)	4 (2%)	1 (1%)	30	56

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	D	172/174 (99%)	166 (96%)	4 (2%)	2 (1%)	16	33
2	F	172/174 (99%)	167 (97%)	4 (2%)	1 (1%)	30	56
2	H	172/174 (99%)	168 (98%)	3 (2%)	1 (1%)	30	56
2	J	172/174 (99%)	164 (95%)	7 (4%)	1 (1%)	30	56
2	L	172/174 (99%)	169 (98%)	2 (1%)	1 (1%)	30	56
2	N	172/174 (99%)	167 (97%)	4 (2%)	1 (1%)	30	56
2	P	172/174 (99%)	163 (95%)	8 (5%)	1 (1%)	30	56
2	R	172/174 (99%)	166 (96%)	5 (3%)	1 (1%)	30	56
2	T	172/174 (99%)	166 (96%)	5 (3%)	1 (1%)	30	56
2	V	172/174 (99%)	163 (95%)	8 (5%)	1 (1%)	30	56
2	X	172/174 (99%)	166 (96%)	5 (3%)	1 (1%)	30	56
All	All	5856/5904 (99%)	5662 (97%)	178 (3%)	16 (0%)	46	72

5 of 16 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	D	450	ARG
2	F	450	ARG
2	H	450	ARG
2	J	450	ARG
2	L	450	ARG

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	269/269 (100%)	248 (92%)	21 (8%)	16	30
1	C	269/269 (100%)	249 (93%)	20 (7%)	17	34
1	E	269/269 (100%)	247 (92%)	22 (8%)	14	27
1	G	269/269 (100%)	248 (92%)	21 (8%)	16	30
1	I	269/269 (100%)	250 (93%)	19 (7%)	18	36

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	K	269/269 (100%)	246 (91%)	23 (9%)	13 25
1	M	269/269 (100%)	250 (93%)	19 (7%)	18 36
1	O	269/269 (100%)	249 (93%)	20 (7%)	17 34
1	Q	269/269 (100%)	246 (91%)	23 (9%)	13 25
1	S	269/269 (100%)	250 (93%)	19 (7%)	18 36
1	U	269/269 (100%)	241 (90%)	28 (10%)	9 16
1	W	269/269 (100%)	245 (91%)	24 (9%)	12 23
2	B	148/148 (100%)	139 (94%)	9 (6%)	23 46
2	D	148/148 (100%)	138 (93%)	10 (7%)	20 39
2	F	148/148 (100%)	137 (93%)	11 (7%)	17 34
2	H	148/148 (100%)	144 (97%)	4 (3%)	52 79
2	J	148/148 (100%)	140 (95%)	8 (5%)	27 52
2	L	148/148 (100%)	140 (95%)	8 (5%)	27 52
2	N	148/148 (100%)	141 (95%)	7 (5%)	32 59
2	P	148/148 (100%)	137 (93%)	11 (7%)	17 34
2	R	148/148 (100%)	141 (95%)	7 (5%)	32 59
2	T	148/148 (100%)	141 (95%)	7 (5%)	32 59
2	V	148/148 (100%)	139 (94%)	9 (6%)	23 46
2	X	148/148 (100%)	143 (97%)	5 (3%)	44 72
All	All	5004/5004 (100%)	4649 (93%)	355 (7%)	18 36

5 of 355 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	K	244	LEU
2	N	433	MET
1	W	30	THR
1	K	276	TRP
1	M	57	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 12 such sidechains are listed below:

Mol	Chain	Res	Type
2	J	399	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	S	8	HIS
1	U	303	ASN
1	I	226	HIS
2	T	418	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

14 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	E	603	1,4	14,14,15	0.22	0	15,19,21	0.35	0
4	NAG	E	604	4	14,14,15	0.81	0	15,19,21	0.70	0
5	SIA	E	605	5	16,20,21	0.23	0	18,28,31	0.70	1 (5%)
5	GAL	E	606	5	11,11,12	0.74	0	14,15,17	1.08	0
5	NAG	E	607	5	14,14,15	0.29	0	15,19,21	0.32	0
5	SIA	G	602	5	16,20,21	0.23	0	18,28,31	0.70	1 (5%)
5	GAL	G	603	5	11,11,12	0.28	0	14,15,17	0.79	1 (7%)
5	NAG	G	604	5	14,14,15	0.25	0	15,19,21	0.21	0
5	SIA	I	602	5	16,20,21	0.23	0	18,28,31	0.70	1 (5%)
5	GAL	I	603	5	11,11,12	0.86	0	14,15,17	1.40	2 (14%)
5	NAG	I	604	5	14,14,15	0.36	0	15,19,21	0.40	0
5	SIA	U	602	5	16,20,21	1.95	6 (37%)	18,28,31	2.75	10 (55%)
5	GAL	U	603	5	11,11,12	0.73	0	14,15,17	1.21	1 (7%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
5	NAG	U	604	5	14,14,15	0.23	0	15,19,21	0.45	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
4	NAG	E	603	1,4	-	0/6/23/26	0/1/1/1
4	NAG	E	604	4	-	0/6/23/26	0/1/1/1
5	SIA	E	605	5	-	0/14/34/38	0/1/1/1
5	GAL	E	606	5	-	0/2/19/22	0/1/1/1
5	NAG	E	607	5	-	0/6/23/26	0/1/1/1
5	SIA	G	602	5	-	0/14/34/38	0/1/1/1
5	GAL	G	603	5	-	0/2/19/22	0/1/1/1
5	NAG	G	604	5	-	0/6/23/26	0/1/1/1
5	SIA	I	602	5	-	0/14/34/38	0/1/1/1
5	GAL	I	603	5	-	0/2/19/22	0/1/1/1
5	NAG	I	604	5	-	0/6/23/26	0/1/1/1
5	SIA	U	602	5	-	0/14/34/38	0/1/1/1
5	GAL	U	603	5	-	0/2/19/22	0/1/1/1
5	NAG	U	604	5	-	0/6/23/26	0/1/1/1

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	U	602	SIA	C8-C7	-3.21	1.46	1.53
5	U	602	SIA	O10-C10	-2.94	1.16	1.23
5	U	602	SIA	C7-C6	-2.69	1.49	1.52
5	U	602	SIA	C11-C10	-2.22	1.46	1.50
5	U	602	SIA	C3-C2	2.24	1.56	1.52

The worst 5 of 17 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	U	602	SIA	O6-C2-C3	-6.20	97.95	109.86
5	U	602	SIA	O9-C9-C8	-5.36	99.45	111.10
5	U	602	SIA	O10-C10-C11	-4.06	114.62	122.06
5	U	602	SIA	O8-C8-C9	-2.40	103.62	109.22
5	G	602	SIA	C7-C6-C5	-2.23	110.95	114.32

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5 monomers are involved in 16 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	E	603	NAG	1	0
5	E	605	SIA	4	0
5	G	602	SIA	4	0
5	I	602	SIA	3	0
5	U	602	SIA	4	0

5.6 Ligand geometry (i)

25 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	601	1	14,14,15	0.27	0	15,19,21	0.47	0
3	NAG	B	601	2	14,14,15	0.34	0	15,19,21	0.45	0
3	NAG	C	601	1	14,14,15	0.19	0	15,19,21	0.33	0
3	NAG	D	601	2	14,14,15	0.29	0	15,19,21	0.39	0
3	NAG	E	601	1	14,14,15	0.37	0	15,19,21	0.45	0
3	NAG	E	602	1	14,14,15	0.87	1 (7%)	15,19,21	0.96	0
3	NAG	F	601	2	14,14,15	0.29	0	15,19,21	0.59	0
3	NAG	G	601	1	14,14,15	0.17	0	15,19,21	0.41	0
3	NAG	H	601	2	14,14,15	0.42	0	15,19,21	0.58	0
3	NAG	I	601	1	14,14,15	0.22	0	15,19,21	0.40	0
3	NAG	J	601	2	14,14,15	0.68	0	15,19,21	0.75	0
3	NAG	K	601	1	14,14,15	0.34	0	15,19,21	0.50	0
3	NAG	L	601	2	14,14,15	0.37	0	15,19,21	0.38	0
3	NAG	M	601	1	14,14,15	0.34	0	15,19,21	0.53	0
3	NAG	N	601	2	14,14,15	0.46	0	15,19,21	0.37	0
3	NAG	O	601	1	14,14,15	0.19	0	15,19,21	0.28	0
3	NAG	P	601	2	14,14,15	0.35	0	15,19,21	0.42	0
3	NAG	Q	601	1	14,14,15	0.27	0	15,19,21	0.35	0
3	NAG	R	601	2	14,14,15	0.19	0	15,19,21	0.62	1 (6%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
3	NAG	S	601	1	14,14,15	0.27	0	15,19,21	0.50	0
3	NAG	T	601	2	14,14,15	0.56	0	15,19,21	0.30	0
3	NAG	U	601	1	14,14,15	0.41	0	15,19,21	0.36	0
3	NAG	V	601	2	14,14,15	0.38	0	15,19,21	0.48	0
3	NAG	W	601	1	14,14,15	0.27	0	15,19,21	0.53	0
3	NAG	X	601	2	14,14,15	0.51	0	15,19,21	0.50	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
3	NAG	A	601	1	-	0/6/23/26	0/1/1/1
3	NAG	B	601	2	-	0/6/23/26	0/1/1/1
3	NAG	C	601	1	-	0/6/23/26	0/1/1/1
3	NAG	D	601	2	-	0/6/23/26	0/1/1/1
3	NAG	E	601	1	-	0/6/23/26	0/1/1/1
3	NAG	E	602	1	-	0/6/23/26	0/1/1/1
3	NAG	F	601	2	-	0/6/23/26	0/1/1/1
3	NAG	G	601	1	-	0/6/23/26	0/1/1/1
3	NAG	H	601	2	-	0/6/23/26	0/1/1/1
3	NAG	I	601	1	-	0/6/23/26	0/1/1/1
3	NAG	J	601	2	-	0/6/23/26	0/1/1/1
3	NAG	K	601	1	-	0/6/23/26	0/1/1/1
3	NAG	L	601	2	-	0/6/23/26	0/1/1/1
3	NAG	M	601	1	-	0/6/23/26	0/1/1/1
3	NAG	N	601	2	-	0/6/23/26	0/1/1/1
3	NAG	O	601	1	-	0/6/23/26	0/1/1/1
3	NAG	P	601	2	-	0/6/23/26	0/1/1/1
3	NAG	Q	601	1	-	0/6/23/26	0/1/1/1
3	NAG	R	601	2	-	0/6/23/26	0/1/1/1
3	NAG	S	601	1	-	0/6/23/26	0/1/1/1
3	NAG	T	601	2	-	0/6/23/26	0/1/1/1
3	NAG	U	601	1	-	0/6/23/26	0/1/1/1
3	NAG	V	601	2	-	0/6/23/26	0/1/1/1
3	NAG	W	601	1	-	0/6/23/26	0/1/1/1
3	NAG	X	601	2	-	0/6/23/26	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	E	602	NAG	C2-N2	2.17	1.50	1.46

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	R	601	NAG	C1-O5-C5	2.03	114.83	112.25

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

8 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	D	601	NAG	1	0
3	E	601	NAG	1	0
3	J	601	NAG	1	0
3	K	601	NAG	1	0
3	N	601	NAG	1	0
3	O	601	NAG	1	0
3	T	601	NAG	1	0
3	U	601	NAG	1	0

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	318/318 (100%)	-0.14	0 100 100	10, 25, 42, 69	0
1	C	318/318 (100%)	0.17	6 (1%) 70 64	10, 36, 64, 95	0
1	E	318/318 (100%)	0.01	1 (0%) 94 93	11, 30, 48, 70	0
1	G	318/318 (100%)	0.10	5 (1%) 74 69	15, 34, 54, 80	0
1	I	318/318 (100%)	0.03	5 (1%) 74 69	11, 29, 52, 101	0
1	K	318/318 (100%)	-0.03	7 (2%) 65 59	13, 29, 52, 84	0
1	M	318/318 (100%)	0.70	35 (11%) 7 4	24, 56, 87, 123	0
1	O	318/318 (100%)	0.54	28 (8%) 12 8	22, 52, 82, 112	0
1	Q	318/318 (100%)	-0.04	5 (1%) 74 69	12, 33, 54, 92	0
1	S	318/318 (100%)	0.38	22 (6%) 20 14	17, 43, 70, 107	0
1	U	318/318 (100%)	0.00	5 (1%) 74 69	16, 31, 53, 83	0
1	W	318/318 (100%)	0.28	10 (3%) 52 45	24, 44, 71, 94	0
2	B	174/174 (100%)	0.29	5 (2%) 55 48	12, 26, 54, 108	0
2	D	174/174 (100%)	0.31	3 (1%) 73 68	11, 25, 52, 130	0
2	F	174/174 (100%)	0.23	4 (2%) 64 57	10, 27, 56, 101	0
2	H	174/174 (100%)	0.24	2 (1%) 82 79	12, 29, 55, 125	0
2	J	174/174 (100%)	0.46	6 (3%) 49 41	12, 37, 74, 132	0
2	L	174/174 (100%)	0.37	6 (3%) 49 41	14, 35, 67, 108	0
2	N	174/174 (100%)	0.38	4 (2%) 64 57	23, 38, 71, 109	0
2	P	174/174 (100%)	0.43	6 (3%) 49 41	18, 37, 66, 125	0
2	R	174/174 (100%)	0.36	8 (4%) 36 29	16, 33, 55, 110	0
2	T	174/174 (100%)	0.25	3 (1%) 73 68	13, 37, 61, 119	0
2	V	174/174 (100%)	0.35	4 (2%) 64 57	14, 39, 71, 127	0
2	X	174/174 (100%)	0.41	4 (2%) 64 57	19, 45, 69, 119	0

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å^2)	Q<0.9
All	All	5904/5904 (100%)	0.23	184 (3%) 52 45	10, 35, 69, 132	0

The worst 5 of 184 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	497	ASN	13.5
2	P	497	ASN	13.2
2	J	497	ASN	11.5
2	V	496	ILE	11.1
2	X	497	ASN	10.6

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. 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
5	SIA	G	602	20/21	0.87	0.24	1.72	44,57,71,72	0
5	SIA	I	602	20/21	0.89	0.19	1.10	26,35,46,47	0
5	GAL	E	606	11/12	0.86	0.24	0.44	90,98,116,120	0
5	SIA	E	605	20/21	0.87	0.19	0.33	34,45,58,60	0
5	SIA	U	602	20/21	0.94	0.15	-0.02	24,34,46,47	0
5	GAL	U	603	11/12	0.78	0.27	-	58,69,81,85	0
5	NAG	U	604	14/15	0.76	0.58	-	99,115,125,128	0
5	NAG	G	604	14/15	0.71	0.79	-	125,140,147,147	0
5	GAL	G	603	11/12	0.74	0.24	-	72,90,112,114	0
5	NAG	E	607	14/15	0.71	0.55	-	100,109,111,112	0
5	NAG	I	604	14/15	0.77	0.49	-	97,103,110,110	0
4	NAG	E	603	14/15	0.94	0.22	-	46,54,64,68	0
4	NAG	E	604	14/15	0.89	0.26	-	48,69,77,80	0
5	GAL	I	603	11/12	0.78	0.24	-	66,76,99,100	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^{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
3	NAG	U	601	14/15	0.84	0.31	3.17	80,98,105,113	0
3	NAG	W	601	14/15	0.73	0.30	2.02	79,87,91,92	0
3	NAG	K	601	14/15	0.88	0.26	1.21	46,71,77,79	0
3	NAG	Q	601	14/15	0.87	0.20	1.18	76,87,98,99	0
3	NAG	A	601	14/15	0.92	0.18	0.99	34,57,68,74	0
3	NAG	V	601	14/15	0.90	0.18	0.93	28,38,40,44	0
3	NAG	O	601	14/15	0.72	0.30	0.91	61,88,97,101	0
3	NAG	M	601	14/15	0.87	0.28	0.74	66,83,89,92	0
3	NAG	G	601	14/15	0.86	0.17	0.67	43,64,72,74	0
3	NAG	I	601	14/15	0.92	0.17	0.56	52,62,65,65	0
3	NAG	S	601	14/15	0.78	0.20	0.40	54,75,88,92	0
3	NAG	R	601	14/15	0.86	0.15	0.33	32,39,49,50	0
3	NAG	C	601	14/15	0.87	0.19	0.19	46,67,77,77	0
3	NAG	L	601	14/15	0.85	0.17	-0.08	23,32,38,38	0
3	NAG	X	601	14/15	0.87	0.16	-0.09	37,46,54,54	0
3	NAG	F	601	14/15	0.89	0.14	-0.10	25,30,33,36	0
3	NAG	P	601	14/15	0.93	0.13	-0.26	34,38,44,44	0
3	NAG	J	601	14/15	0.95	0.13	-0.55	18,28,36,38	0
3	NAG	E	601	14/15	0.91	0.14	-0.72	48,62,69,70	0
3	NAG	N	601	14/15	0.91	0.13	-1.01	40,50,51,54	0
3	NAG	D	601	14/15	0.95	0.11	-1.22	20,28,36,37	0
3	NAG	T	601	14/15	0.93	0.12	-1.28	23,34,41,43	0
3	NAG	H	601	14/15	0.91	0.13	-1.94	22,30,33,33	0
3	NAG	B	601	14/15	0.93	0.10	-2.34	29,38,43,50	0
3	NAG	E	602	14/15	0.88	0.27	-	54,58,73,77	0

6.5 Other polymers

There are no such residues in this entry.