wwPDB X-ray Structure Validation Summary Report (i)

Feb 1, 2016 – 05:57 PM GMT

PDB ID	:	4JUL
Title	:	Crystal structure of H5N1 influenza virus hemagglutinin, clade 2.3.4
Authors	:	DuBois, R.M.; Zaraket, H.; Reddivari, M.; Coop, T.; Heath, R.J.; White, S.W.; Russell, C.J.
Deposited on	:	2013-03-25
Resolution	:	2.79 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:

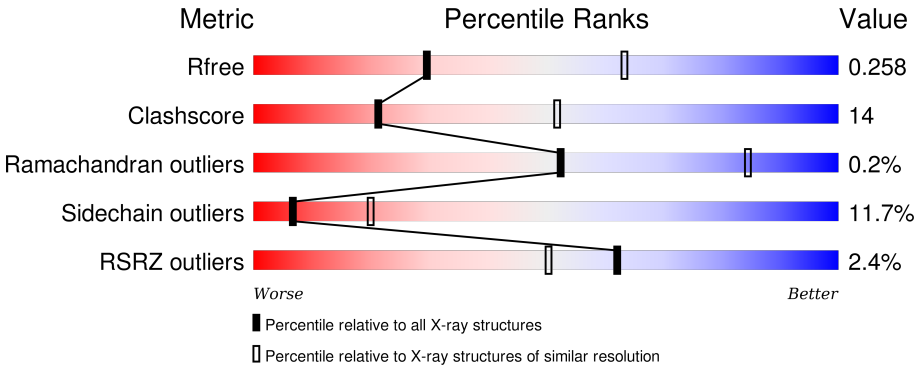
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	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 2.79 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.



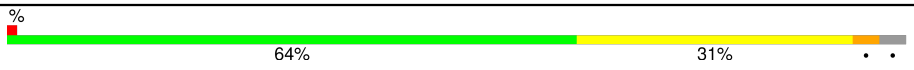
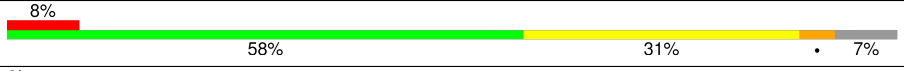
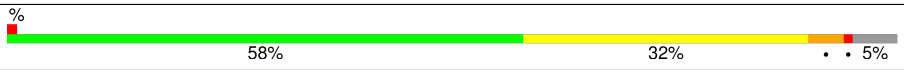
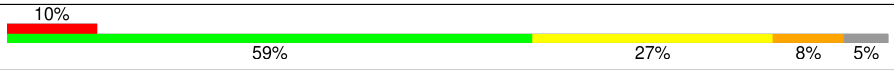
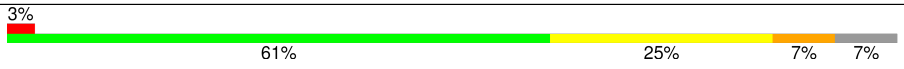
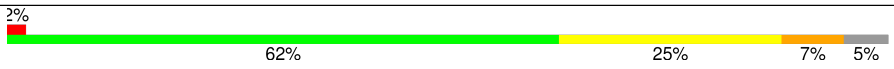
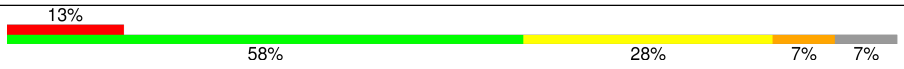
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	91344	2393 (2.80-2.80)
Clashscore	102246	2827 (2.80-2.80)
Ramachandran outliers	100387	2782 (2.80-2.80)
Sidechain outliers	100360	2784 (2.80-2.80)
RSRZ outliers	91569	2404 (2.80-2.80)

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	71% 24%
1	C	329	64% 30%
1	E	329	67% 27%
1	H	329	67% 27%
1	J	329	69% 26%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	L	329	
2	B	182	
2	D	182	
2	F	182	
2	I	182	
2	K	182	
2	M	182	

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	J	401	-	-	X	-
3	NAG	J	402	-	-	-	X

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 23752 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 HA1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	321	Total	C	N	O	S	0	0	0
			2535	1601	436	485	13			
1	C	321	Total	C	N	O	S	0	0	0
			2535	1601	436	485	13			
1	E	322	Total	C	N	O	S	0	0	0
			2542	1606	437	486	13			
1	H	321	Total	C	N	O	S	0	0	0
			2535	1601	436	485	13			
1	J	321	Total	C	N	O	S	0	0	0
			2535	1601	436	485	13			
1	L	319	Total	C	N	O	S	0	0	0
			2523	1595	434	481	13			

There are 36 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	5	GLY	-	EXPRESSION TAG	UNP Q00G25
A	6	SER	-	EXPRESSION TAG	UNP Q00G25
A	7	ALA	-	EXPRESSION TAG	UNP Q00G25
A	8	ASP	-	EXPRESSION TAG	UNP Q00G25
A	9	PRO	-	EXPRESSION TAG	UNP Q00G25
A	10	GLY	-	EXPRESSION TAG	UNP Q00G25
C	5	GLY	-	EXPRESSION TAG	UNP Q00G25
C	6	SER	-	EXPRESSION TAG	UNP Q00G25
C	7	ALA	-	EXPRESSION TAG	UNP Q00G25
C	8	ASP	-	EXPRESSION TAG	UNP Q00G25
C	9	PRO	-	EXPRESSION TAG	UNP Q00G25
C	10	GLY	-	EXPRESSION TAG	UNP Q00G25
E	5	GLY	-	EXPRESSION TAG	UNP Q00G25
E	6	SER	-	EXPRESSION TAG	UNP Q00G25
E	7	ALA	-	EXPRESSION TAG	UNP Q00G25
E	8	ASP	-	EXPRESSION TAG	UNP Q00G25
E	9	PRO	-	EXPRESSION TAG	UNP Q00G25

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	10	GLY	-	EXPRESSION TAG	UNP Q00G25
H	5	GLY	-	EXPRESSION TAG	UNP Q00G25
H	6	SER	-	EXPRESSION TAG	UNP Q00G25
H	7	ALA	-	EXPRESSION TAG	UNP Q00G25
H	8	ASP	-	EXPRESSION TAG	UNP Q00G25
H	9	PRO	-	EXPRESSION TAG	UNP Q00G25
H	10	GLY	-	EXPRESSION TAG	UNP Q00G25
J	5	GLY	-	EXPRESSION TAG	UNP Q00G25
J	6	SER	-	EXPRESSION TAG	UNP Q00G25
J	7	ALA	-	EXPRESSION TAG	UNP Q00G25
J	8	ASP	-	EXPRESSION TAG	UNP Q00G25
J	9	PRO	-	EXPRESSION TAG	UNP Q00G25
J	10	GLY	-	EXPRESSION TAG	UNP Q00G25
L	5	GLY	-	EXPRESSION TAG	UNP Q00G25
L	6	SER	-	EXPRESSION TAG	UNP Q00G25
L	7	ALA	-	EXPRESSION TAG	UNP Q00G25
L	8	ASP	-	EXPRESSION TAG	UNP Q00G25
L	9	PRO	-	EXPRESSION TAG	UNP Q00G25
L	10	GLY	-	EXPRESSION TAG	UNP Q00G25

- Molecule 2 is a protein called Hemagglutinin HA2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	170	Total	C	N	O	S	0	0	0
			1380	859	241	272	8			
2	D	172	Total	C	N	O	S	0	0	0
			1398	869	243	278	8			
2	F	172	Total	C	N	O	S	0	0	0
			1398	869	243	278	8			
2	I	170	Total	C	N	O	S	0	0	0
			1380	859	241	272	8			
2	K	172	Total	C	N	O	S	0	0	0
			1398	869	243	278	8			
2	M	169	Total	C	N	O	S	0	0	0
			1369	853	237	271	8			

There are 36 discrepancies between the modelled and reference sequences:

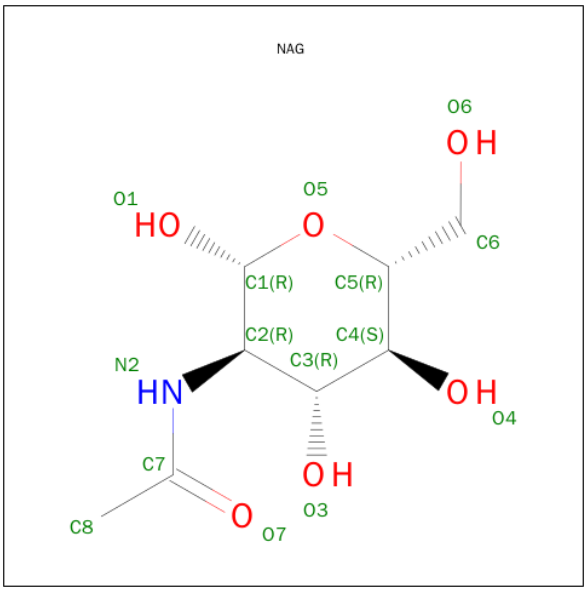
Chain	Residue	Modelled	Actual	Comment	Reference
B	177	ARG	-	EXPRESSION TAG	UNP Q00G25
B	178	SER	-	EXPRESSION TAG	UNP Q00G25
B	179	LEU	-	EXPRESSION TAG	UNP Q00G25

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	180	VAL	-	EXPRESSION TAG	UNP Q00G25
B	181	PRO	-	EXPRESSION TAG	UNP Q00G25
B	182	ARG	-	EXPRESSION TAG	UNP Q00G25
D	177	ARG	-	EXPRESSION TAG	UNP Q00G25
D	178	SER	-	EXPRESSION TAG	UNP Q00G25
D	179	LEU	-	EXPRESSION TAG	UNP Q00G25
D	180	VAL	-	EXPRESSION TAG	UNP Q00G25
D	181	PRO	-	EXPRESSION TAG	UNP Q00G25
D	182	ARG	-	EXPRESSION TAG	UNP Q00G25
F	177	ARG	-	EXPRESSION TAG	UNP Q00G25
F	178	SER	-	EXPRESSION TAG	UNP Q00G25
F	179	LEU	-	EXPRESSION TAG	UNP Q00G25
F	180	VAL	-	EXPRESSION TAG	UNP Q00G25
F	181	PRO	-	EXPRESSION TAG	UNP Q00G25
F	182	ARG	-	EXPRESSION TAG	UNP Q00G25
I	177	ARG	-	EXPRESSION TAG	UNP Q00G25
I	178	SER	-	EXPRESSION TAG	UNP Q00G25
I	179	LEU	-	EXPRESSION TAG	UNP Q00G25
I	180	VAL	-	EXPRESSION TAG	UNP Q00G25
I	181	PRO	-	EXPRESSION TAG	UNP Q00G25
I	182	ARG	-	EXPRESSION TAG	UNP Q00G25
K	177	ARG	-	EXPRESSION TAG	UNP Q00G25
K	178	SER	-	EXPRESSION TAG	UNP Q00G25
K	179	LEU	-	EXPRESSION TAG	UNP Q00G25
K	180	VAL	-	EXPRESSION TAG	UNP Q00G25
K	181	PRO	-	EXPRESSION TAG	UNP Q00G25
K	182	ARG	-	EXPRESSION TAG	UNP Q00G25
M	177	ARG	-	EXPRESSION TAG	UNP Q00G25
M	178	SER	-	EXPRESSION TAG	UNP Q00G25
M	179	LEU	-	EXPRESSION TAG	UNP Q00G25
M	180	VAL	-	EXPRESSION TAG	UNP Q00G25
M	181	PRO	-	EXPRESSION TAG	UNP Q00G25
M	182	ARG	-	EXPRESSION TAG	UNP Q00G25

- 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	E	1	Total C N O 14 8 1 5	0	0
3	H	1	Total C N O 14 8 1 5	0	0
3	J	1	Total C N O 14 8 1 5	0	0
3	L	1	Total C N O 14 8 1 5	0	0
3	J	1	Total C N O 14 8 1 5	0	0
3	L	1	Total C N O 14 8 1 5	0	0
3	L	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	E	2	Total C N O 28 16 2 10	0	0

Continued on next page...

Continued from previous page...

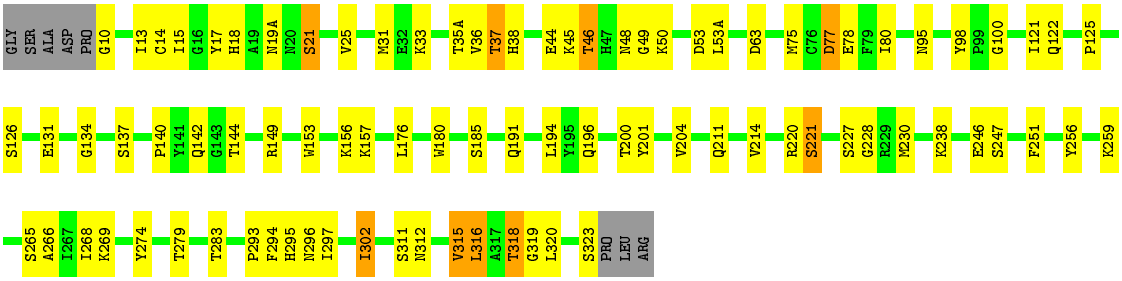
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	H	2	Total	C	N	O	0	0
			28	16	2	10		

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.

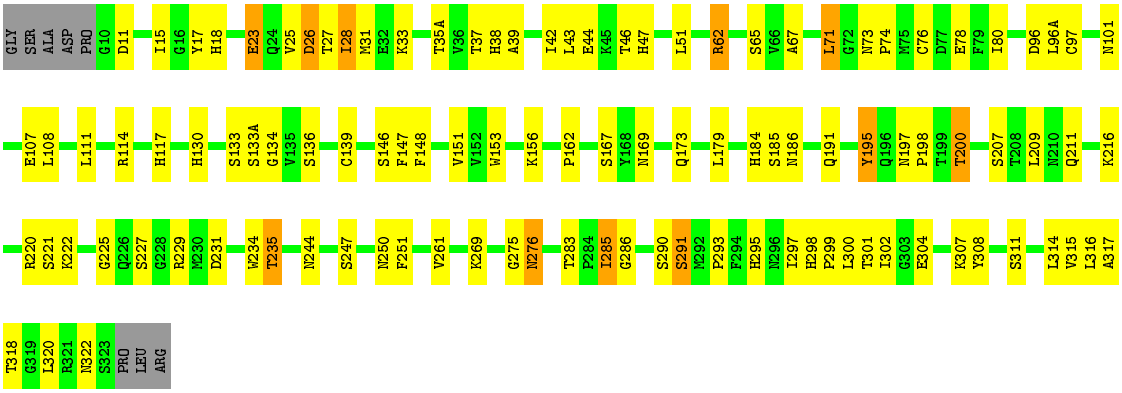
- Molecule 1: Hemagglutinin HA1

Chain A:



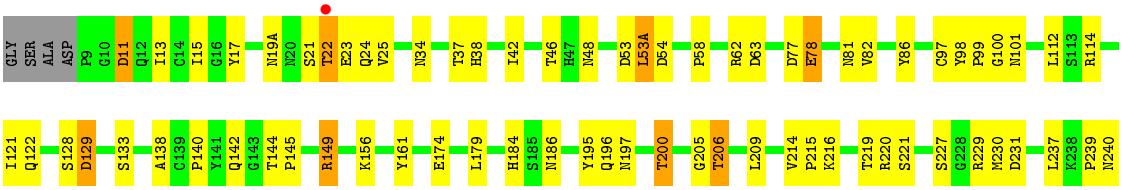
- Molecule 1: Hemagglutinin HA1

Chain C:



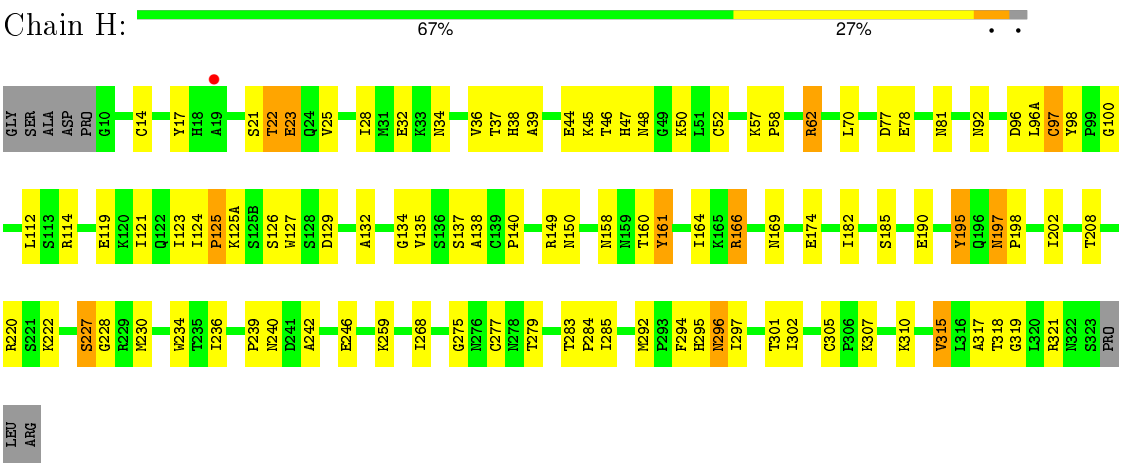
- Molecule 1: Hemagglutinin HA1

Chain E:



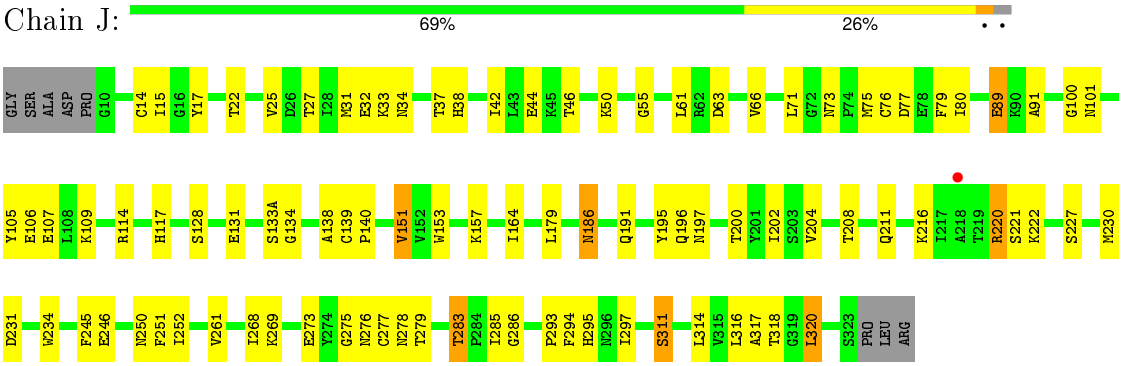
• Molecule 1: Hemagglutinin HA1

Chain H:



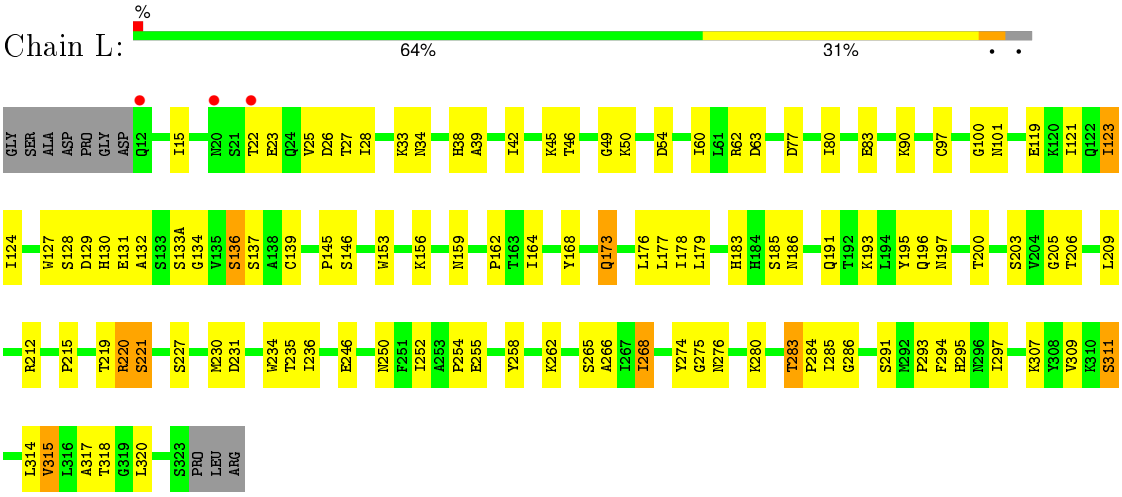
• Molecule 1: Hemagglutinin HA1

Chain J:



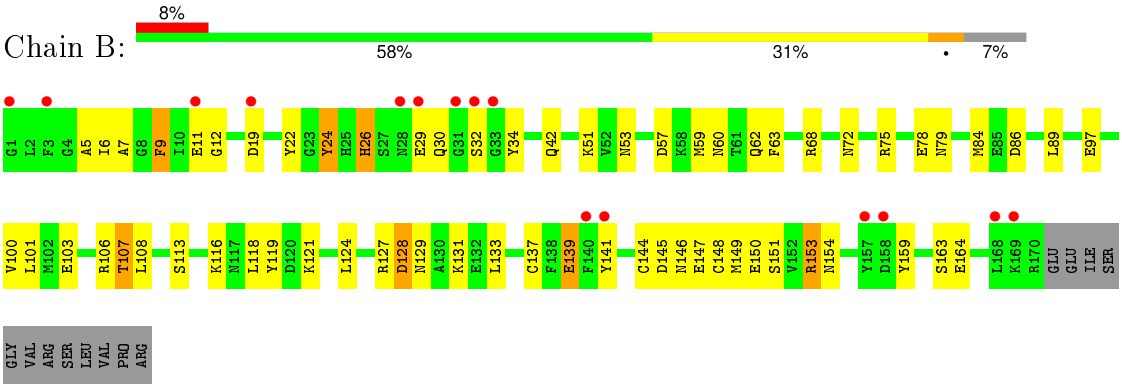
• Molecule 1: Hemagglutinin HA1

Chain L:



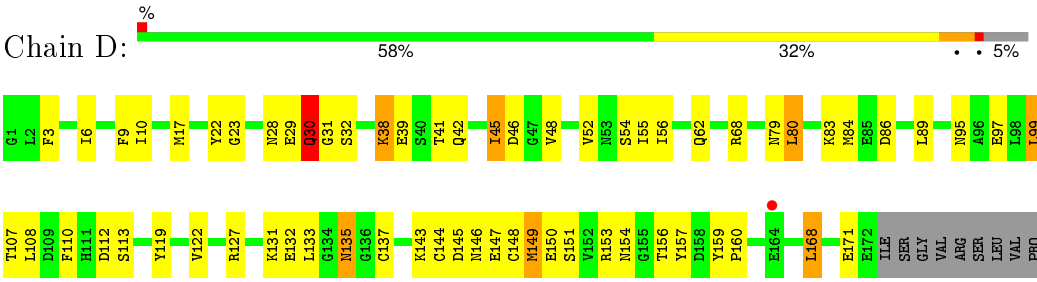
• Molecule 2: Hemagglutinin HA2

Chain B:



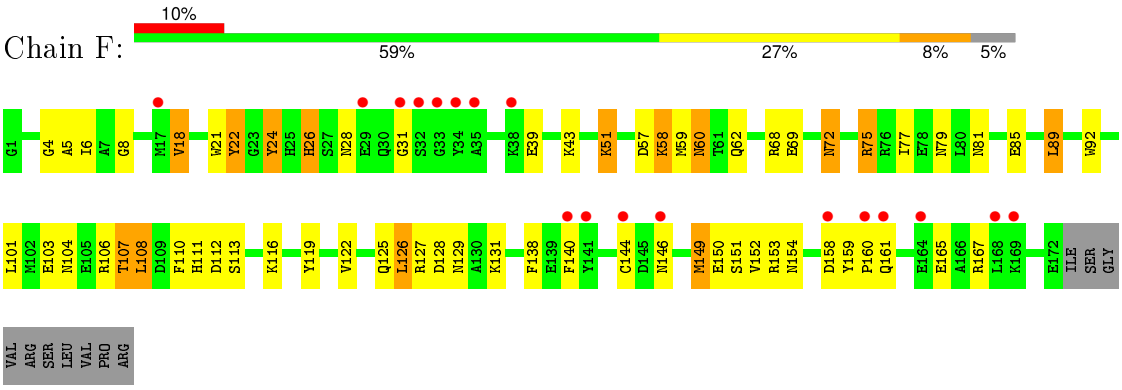
- Molecule 2: Hemagglutinin HA2

Chain D:



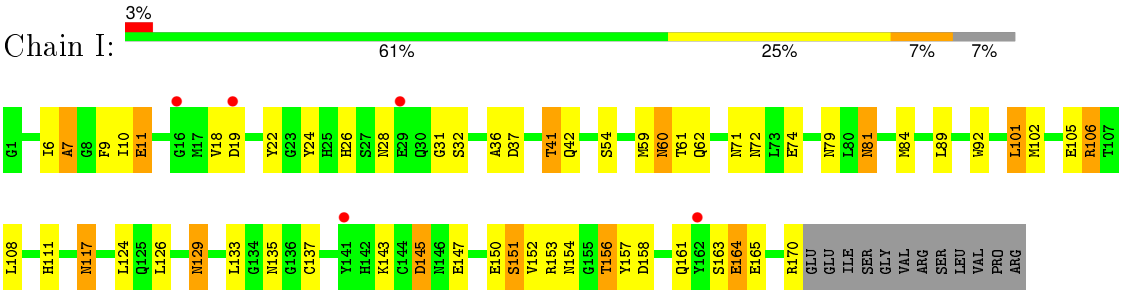
- Molecule 2: Hemagglutinin HA2

Chain F:



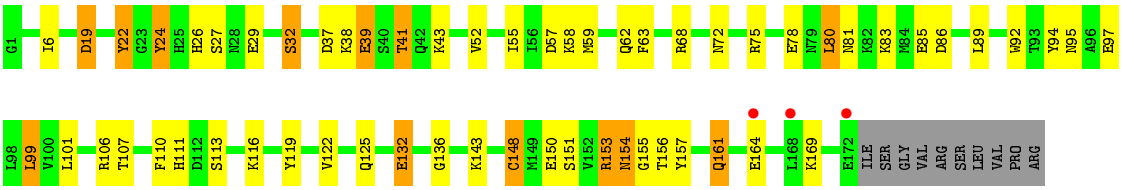
- Molecule 2: Hemagglutinin HA2

Chain I:

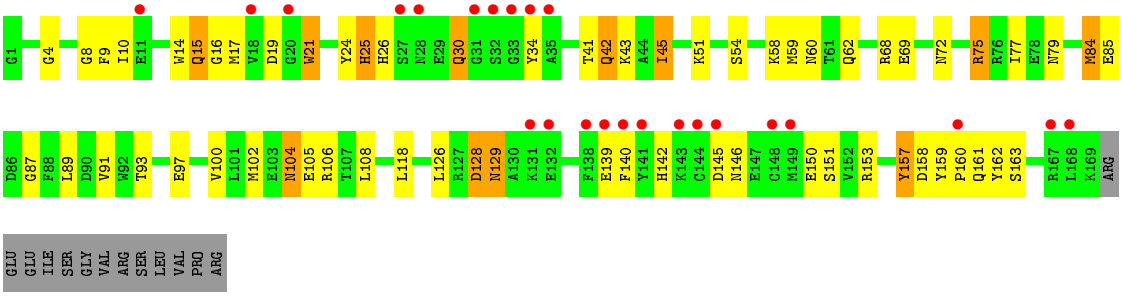


- Molecule 2: Hemagglutinin HA2

Chain K:



- Molecule 2: Hemagglutinin HA2



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	71.58Å 71.56Å 247.05Å 83.94° 85.46° 60.68°	Depositor
Resolution (Å)	48.87 – 2.79 48.87 – 2.79	Depositor EDS
% Data completeness (in resolution range)	81.4 (48.87-2.79) 79.3 (48.87-2.79)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.08 (at 2.81Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.4_1496)	Depositor
R, R_{free}	0.196 , 0.254 0.204 , 0.258	Depositor DCC
R_{free} test set	4355 reflections (5.34%)	DCC
Wilson B-factor (Å^2)	46.5	Xtriage
Anisotropy	0.379	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 40.0	EDS
Estimated twinning fraction	0.018 for -k,-h,-l	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Outliers	0 of 85847 reflections	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	23752	wwPDB-VP
Average B, all atoms (Å^2)	45.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.43% 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: 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	0.52	0/2596	0.68	0/3529
1	C	0.59	0/2596	0.69	0/3529
1	E	0.52	0/2604	0.69	0/3540
1	H	0.53	1/2596 (0.0%)	0.66	0/3529
1	J	0.55	0/2596	0.69	0/3529
1	L	0.51	0/2584	0.69	0/3513
2	B	0.51	0/1407	0.66	1/1891 (0.1%)
2	D	0.51	0/1425	0.68	1/1915 (0.1%)
2	F	0.49	0/1425	0.63	0/1915
2	I	0.52	0/1407	0.71	1/1891 (0.1%)
2	K	0.49	0/1425	0.62	0/1915
2	M	0.48	0/1396	0.65	0/1877
All	All	0.52	1/24057 (0.0%)	0.67	3/32573 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	H	125	PRO	N-CD	-5.58	1.40	1.47

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	D	30	GLN	N-CA-C	-5.55	96.02	111.00
2	B	153	ARG	NE-CZ-NH2	5.17	122.89	120.30
2	I	7	ALA	N-CA-C	-5.03	97.42	111.00

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2535	0	2468	66	0
1	C	2535	0	2469	73	0
1	E	2542	0	2478	81	0
1	H	2535	0	2469	67	0
1	J	2535	0	2469	52	0
1	L	2523	0	2462	74	0
2	B	1380	0	1291	54	0
2	D	1398	0	1303	64	0
2	F	1398	0	1303	57	0
2	I	1380	0	1291	47	0
2	K	1398	0	1303	42	0
2	M	1369	0	1278	63	0
3	A	14	0	13	2	0
3	C	28	0	26	0	0
3	E	14	0	13	5	0
3	H	14	0	13	1	0
3	J	28	0	26	8	0
3	L	42	0	39	4	0
4	A	28	0	25	2	0
4	E	28	0	25	5	0
4	H	28	0	25	2	0
All	All	23752	0	22789	658	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 14.

The worst 5 of 658 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:H:34:ASN:HD21	3:H:401:NAG:C1	1.04	1.59
1:E:34:ASN:HD21	3:E:401:NAG:C1	0.95	1.58
1:J:34:ASN:HD21	3:J:401:NAG:C1	1.14	1.57
1:E:53(A):LEU:HD23	1:E:53(A):LEU:O	1.44	1.15
2:M:158:ASP:O	2:M:161:GLN:HB3	1.51	1.10

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	319/329 (97%)	305 (96%)	13 (4%)	1 (0%)	46 79
1	C	319/329 (97%)	305 (96%)	14 (4%)	0	100 100
1	E	320/329 (97%)	296 (92%)	22 (7%)	2 (1%)	30 65
1	H	319/329 (97%)	299 (94%)	20 (6%)	0	100 100
1	J	319/329 (97%)	298 (93%)	21 (7%)	0	100 100
1	L	317/329 (96%)	288 (91%)	29 (9%)	0	100 100
2	B	168/182 (92%)	159 (95%)	9 (5%)	0	100 100
2	D	170/182 (93%)	161 (95%)	9 (5%)	0	100 100
2	F	170/182 (93%)	154 (91%)	15 (9%)	1 (1%)	30 65
2	I	168/182 (92%)	154 (92%)	14 (8%)	0	100 100
2	K	170/182 (93%)	157 (92%)	11 (6%)	2 (1%)	16 47
2	M	167/182 (92%)	155 (93%)	12 (7%)	0	100 100
All	All	2926/3066 (95%)	2731 (93%)	189 (6%)	6 (0%)	52 84

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	K	161	GLN
1	E	140	PRO
1	A	77	ASP
1	E	78	GLU
2	K	39	GLU

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	286/292 (98%)	265 (93%)	21 (7%)	17 44
1	C	286/292 (98%)	258 (90%)	28 (10%)	10 28
1	E	287/292 (98%)	256 (89%)	31 (11%)	8 23
1	H	286/292 (98%)	262 (92%)	24 (8%)	14 37
1	J	286/292 (98%)	259 (91%)	27 (9%)	11 31
1	L	285/292 (98%)	259 (91%)	26 (9%)	12 33
2	B	145/156 (93%)	121 (83%)	24 (17%)	3 8
2	D	147/156 (94%)	126 (86%)	21 (14%)	4 12
2	F	147/156 (94%)	124 (84%)	23 (16%)	3 9
2	I	145/156 (93%)	117 (81%)	28 (19%)	2 5
2	K	147/156 (94%)	122 (83%)	25 (17%)	2 7
2	M	144/156 (92%)	119 (83%)	25 (17%)	2 7
All	All	2591/2688 (96%)	2288 (88%)	303 (12%)	7 20

5 of 303 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	F	75	ARG
1	H	302	ILE
2	M	21	TRP
2	F	107	THR
1	H	81	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 101 such sidechains are listed below:

Mol	Chain	Res	Type
1	E	226	GLN
2	F	154	ASN
1	L	295	HIS
2	F	15	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	F	60	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

6 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	402	1,4	14,14,15	0.85	0	15,19,21	1.61	1 (6%)
4	NAG	A	403	4	14,14,15	0.71	0	15,19,21	1.90	4 (26%)
4	NAG	E	402	1,4	14,14,15	0.76	0	15,19,21	2.98	5 (33%)
4	NAG	E	403	4	14,14,15	0.57	0	15,19,21	0.97	1 (6%)
4	NAG	H	402	1,4	14,14,15	1.03	1 (7%)	15,19,21	2.75	4 (26%)
4	NAG	H	403	4	14,14,15	0.42	0	15,19,21	1.93	4 (26%)

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	402	1,4	-	0/6/23/26	0/1/1/1
4	NAG	A	403	4	-	0/6/23/26	0/1/1/1
4	NAG	E	402	1,4	-	0/6/23/26	0/1/1/1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	NAG	E	403	4	-	0/6/23/26	0/1/1/1
4	NAG	H	402	1,4	-	0/6/23/26	0/1/1/1
4	NAG	H	403	4	-	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)
4	H	402	NAG	O4-C4	3.07	1.50	1.43

The worst 5 of 19 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	H	402	NAG	O4-C4-C5	-8.68	86.23	109.24
4	E	402	NAG	C2-N2-C7	-8.47	112.16	123.04
4	A	402	NAG	C2-N2-C7	-4.61	117.12	123.04
4	H	402	NAG	C4-C3-C2	-3.25	106.18	111.23
4	E	402	NAG	O4-C4-C5	-3.13	100.93	109.24

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5 monomers are involved in 9 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	402	NAG	2	0
4	A	403	NAG	1	0
4	E	402	NAG	3	0
4	E	403	NAG	2	0
4	H	402	NAG	2	0

5.6 Ligand geometry (i)

10 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	0.48	0	15,19,21	1.58	2 (13%)
3	NAG	C	401	1	14,14,15	0.54	0	15,19,21	1.73	4 (26%)
3	NAG	C	402	1	14,14,15	0.75	1 (7%)	15,19,21	1.78	5 (33%)
3	NAG	E	401	1	14,14,15	0.58	0	15,19,21	1.71	4 (26%)
3	NAG	H	401	1	14,14,15	0.49	0	15,19,21	1.57	2 (13%)
3	NAG	J	401	1	14,14,15	0.57	0	15,19,21	1.28	1 (6%)
3	NAG	J	402	1	14,14,15	0.54	0	15,19,21	0.69	0
3	NAG	L	401	3	14,14,15	0.56	0	15,19,21	0.97	1 (6%)
3	NAG	L	402	1	14,14,15	0.63	0	15,19,21	1.54	3 (20%)
3	NAG	L	403	1,3	14,14,15	0.56	0	15,19,21	1.77	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	C	401	1	-	0/6/23/26	0/1/1/1
3	NAG	C	402	1	-	0/6/23/26	0/1/1/1
3	NAG	E	401	1	-	0/6/23/26	0/1/1/1
3	NAG	H	401	1	-	0/6/23/26	0/1/1/1
3	NAG	J	401	1	-	0/6/23/26	0/1/1/1
3	NAG	J	402	1	-	0/6/23/26	0/1/1/1
3	NAG	L	401	3	-	0/6/23/26	0/1/1/1
3	NAG	L	402	1	-	0/6/23/26	0/1/1/1
3	NAG	L	403	1,3	-	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	C	402	NAG	O5-C1	-2.27	1.39	1.43

The worst 5 of 25 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	C	401	NAG	C2-N2-C7	-4.43	117.35	123.04
3	L	401	NAG	C4-C3-C2	-2.78	106.91	111.23
3	C	402	NAG	O3-C3-C4	-2.71	104.23	110.34
3	L	403	NAG	C4-C3-C2	-2.65	107.11	111.23

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	L	403	NAG	O3-C3-C2	-2.55	104.06	109.11

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

8 monomers are involved in 20 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	401	NAG	2	0
3	E	401	NAG	5	0
3	H	401	NAG	1	0
3	J	401	NAG	7	0
3	J	402	NAG	1	0
3	L	401	NAG	2	0
3	L	402	NAG	2	0
3	L	403	NAG	2	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^{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	321/329 (97%)	-0.34	0 100 100	19, 36, 54, 81	0
1	C	321/329 (97%)	-0.23	0 100 100	20, 39, 55, 71	0
1	E	322/329 (97%)	-0.27	1 (0%) 94 92	19, 36, 57, 93	0
1	H	321/329 (97%)	-0.18	1 (0%) 94 92	20, 40, 59, 80	0
1	J	321/329 (97%)	-0.21	1 (0%) 94 92	20, 38, 55, 65	0
1	L	319/329 (96%)	-0.19	3 (0%) 85 79	20, 40, 65, 95	0
2	B	170/182 (93%)	0.49	15 (8%) 12 6	17, 58, 92, 104	0
2	D	172/182 (94%)	0.12	1 (0%) 90 86	23, 55, 72, 84	0
2	F	172/182 (94%)	0.51	18 (10%) 8 4	18, 55, 95, 102	0
2	I	170/182 (93%)	0.27	5 (2%) 55 43	18, 56, 84, 91	0
2	K	172/182 (94%)	0.06	3 (1%) 73 63	19, 50, 75, 99	0
2	M	169/182 (92%)	0.61	24 (14%) 4 2	17, 63, 106, 122	0
All	All	2950/3066 (96%)	-0.04	72 (2%) 62 50	17, 41, 81, 122	0

The worst 5 of 72 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	F	168	LEU	6.2
2	M	141	TYR	4.7
2	F	35	ALA	4.5
1	L	20	ASN	4.5
2	B	31	GLY	4.4

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
4	NAG	E	402	14/15	0.92	0.19	1.14	49,55,61,71	0
4	NAG	H	402	14/15	0.93	0.17	0.61	54,64,79,80	0
4	NAG	A	402	14/15	0.91	0.13	-0.75	33,47,56,61	0
4	NAG	H	403	14/15	0.84	0.35	-	81,92,100,104	0
4	NAG	A	403	14/15	0.92	0.25	-	62,75,81,93	0
4	NAG	E	403	14/15	0.78	0.38	-	91,97,107,108	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, 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	J	402	14/15	0.85	0.20	2.20	39,55,68,68	0
3	NAG	C	402	14/15	0.92	0.18	1.02	52,59,64,73	0
3	NAG	L	403	14/15	0.89	0.17	0.92	57,65,83,84	0
3	NAG	A	401	14/15	0.84	0.26	-	89,95,105,119	0
3	NAG	E	401	14/15	0.82	0.23	-	71,84,90,92	0
3	NAG	L	402	14/15	0.80	0.35	-	78,88,95,97	0
3	NAG	H	401	14/15	0.75	0.30	-	91,99,112,113	0
3	NAG	L	401	14/15	0.69	0.35	-	85,93,108,114	0
3	NAG	J	401	14/15	0.73	0.42	-	94,106,111,113	0
3	NAG	C	401	14/15	0.82	0.36	-	82,88,98,98	0

6.5 Other polymers (i)

There are no such residues in this entry.