wwPDB X-ray Structure Validation Summary Report (i)

Nov 20, 2016 – 05:27 AM EST

PDB ID	:	5K9Q
Title	:	Crystal structure of multidonor HV1-18-class broadly neutralizing Influenza A antibody 16.a.26 in complex with A/Hong Kong/1-4-MA21-1/1968 (H3N2) Hemagglutinin
Authors	:	Joyce, M.G.; Thomas, P.V.; Wheatley, A.K.; McDermott, A.B.; Mascola, J.R.; Kwong, P.D.
Deposited on	:	2016-06-01
Resolution	:	2.50 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:

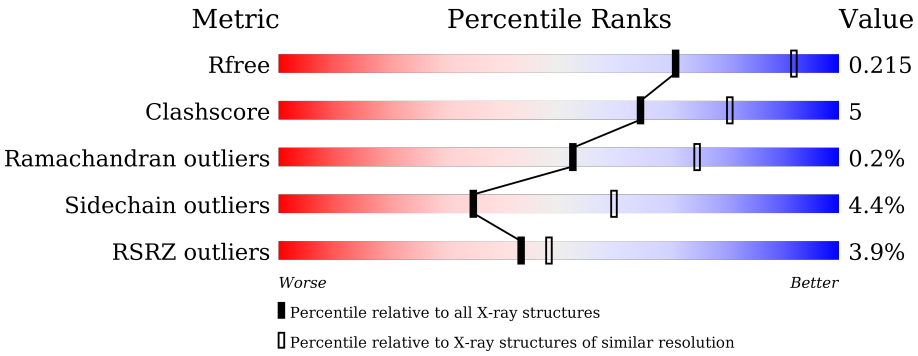
MolProbity	:	4.02b-467
Mogul	:	1.7.1 (RC1), CSD as537be (2016)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20028320
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20028320

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 2.50 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.



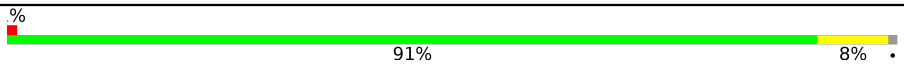
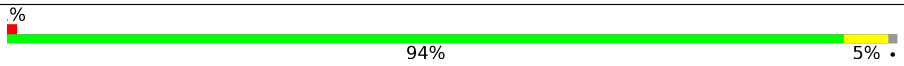
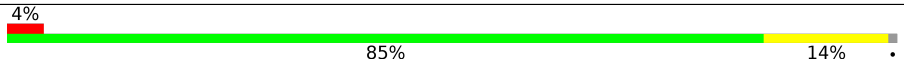
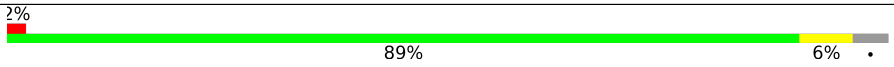
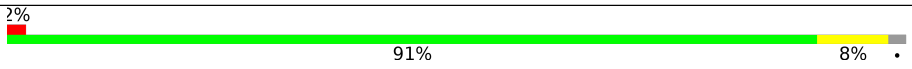
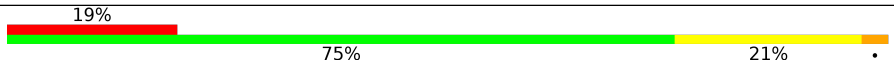
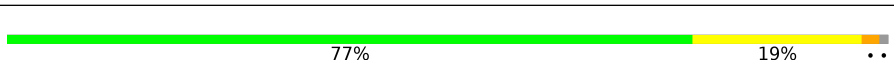
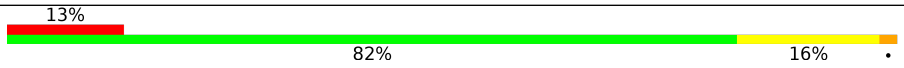
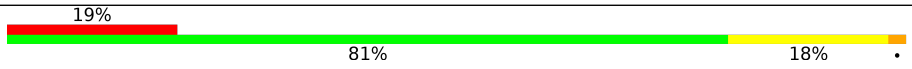
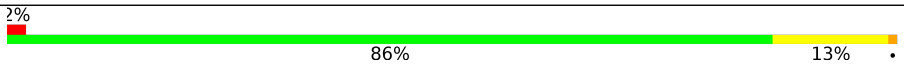
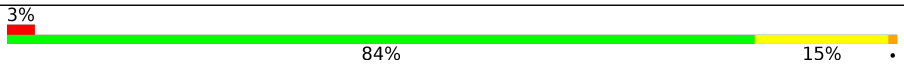
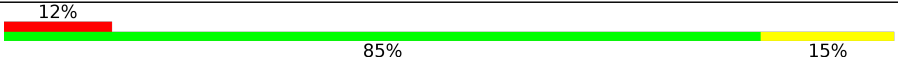
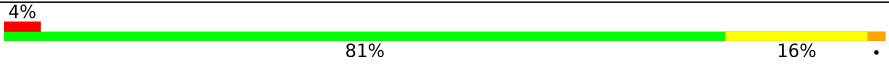
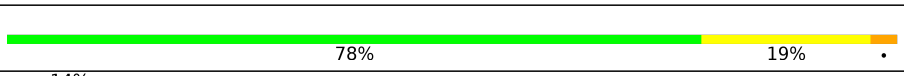
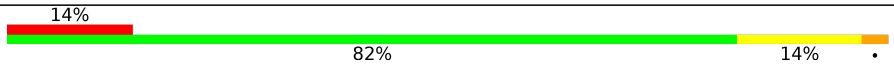
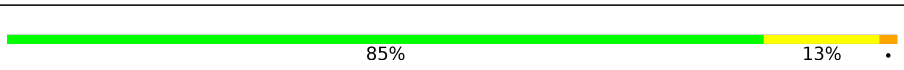
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	91344	3553 (2.50-2.50)
Clashscore	102246	4242 (2.50-2.50)
Ramachandran outliers	100387	4156 (2.50-2.50)
Sidechain outliers	100360	4158 (2.50-2.50)
RSRZ outliers	91569	3562 (2.50-2.50)

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	320	92% 8%
1	C	320	89% 10%
1	E	320	90% 8%
1	M	320	93% 7%
1	O	320	89% 9%
1	Q	320	89% 9%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	B	170	
2	D	170	
2	F	170	
2	N	170	
2	P	170	
2	R	170	
3	G	231	
3	H	231	
3	J	231	
3	S	231	
3	U	231	
3	X	231	
4	I	214	
4	K	214	
4	L	214	
4	T	214	
4	V	214	
4	Y	214	

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
5	NAG	A	402	-	-	-	X
5	NAG	A	403	-	-	-	X
5	NAG	A	404	-	-	-	X
5	NAG	C	403	-	-	-	X
5	NAG	C	404	-	-	-	X
5	NAG	E	403	-	-	-	X
5	NAG	M	405	-	-	-	X

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
5	NAG	Q	401	-	-	-	X
5	NAG	Q	402	-	-	-	X
5	NAG	Q	403	-	-	-	X

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 45420 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	320	Total	C	N	O	S	0	0	0
			2468	1544	433	478	13			
1	C	319	Total	C	N	O	S	0	0	0
			2467	1544	433	477	13			
1	E	318	Total	C	N	O	S	0	0	0
			2458	1538	431	476	13			
1	M	318	Total	C	N	O	S	0	0	0
			2458	1538	431	476	13			
1	O	318	Total	C	N	O	S	0	0	0
			2458	1538	431	476	13			
1	Q	318	Total	C	N	O	S	0	0	0
			2459	1540	431	475	13			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	218	GLU	GLY	conflict	UNP Q91MA7
A	327	ALA	GLN	conflict	UNP Q91MA7
C	218	GLU	GLY	conflict	UNP Q91MA7
C	327	ALA	GLN	conflict	UNP Q91MA7
E	218	GLU	GLY	conflict	UNP Q91MA7
E	327	ALA	GLN	conflict	UNP Q91MA7
M	218	GLU	GLY	conflict	UNP Q91MA7
M	327	ALA	GLN	conflict	UNP Q91MA7
O	218	GLU	GLY	conflict	UNP Q91MA7
O	327	ALA	GLN	conflict	UNP Q91MA7
Q	218	GLU	GLY	conflict	UNP Q91MA7
Q	327	ALA	GLN	conflict	UNP Q91MA7

- Molecule 2 is a protein called Hemagglutinin HA2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	168	Total	C	N	O	S	0	0	0
			1366	845	241	274	6			
2	D	169	Total	C	N	O	S	0	0	0
			1377	854	242	275	6			
2	F	168	Total	C	N	O	S	0	0	0
			1371	848	242	275	6			
2	N	169	Total	C	N	O	S	0	0	0
			1368	846	240	276	6			
2	P	163	Total	C	N	O	S	0	0	0
			1340	829	236	269	6			
2	R	167	Total	C	N	O	S	0	0	0
			1362	843	240	273	6			

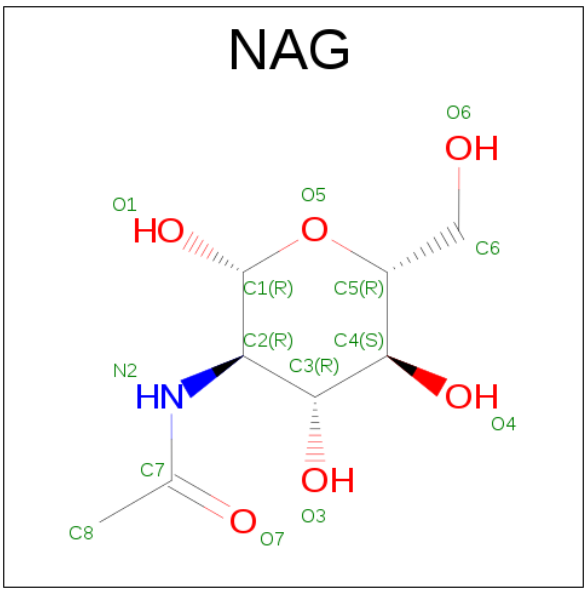
- Molecule 3 is a protein called 16.a.26 Heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	G	231	Total	C	N	O	S	0	0	0
			1728	1089	294	338	7			
3	H	228	Total	C	N	O	S	0	0	0
			1707	1077	290	333	7			
3	J	231	Total	C	N	O	S	0	0	0
			1728	1089	294	338	7			
3	S	231	Total	C	N	O	S	0	0	0
			1728	1089	294	338	7			
3	U	231	Total	C	N	O	S	0	0	0
			1728	1089	294	338	7			
3	X	231	Total	C	N	O	S	0	0	0
			1728	1089	294	338	7			

- Molecule 4 is a protein called 16.a.26 Light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	I	214	Total	C	N	O	S	0	0	0
			1658	1034	283	335	6			
4	K	214	Total	C	N	O	S	0	0	0
			1658	1034	283	335	6			
4	L	214	Total	C	N	O	S	0	0	0
			1658	1034	283	335	6			
4	T	214	Total	C	N	O	S	0	0	0
			1658	1034	283	335	6			
4	V	214	Total	C	N	O	S	0	0	0
			1658	1034	283	335	6			
4	Y	214	Total	C	N	O	S	0	0	0
			1658	1034	283	335	6			

- Molecule 5 is N-ACETYL-D-GLUCOSAMINE (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



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

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	E	1	Total	C	N	O	0	0
			14	8	1	5		
5	E	1	Total	C	N	O	0	0
			14	8	1	5		
5	E	1	Total	C	N	O	0	0
			14	8	1	5		
5	E	1	Total	C	N	O	0	0
			14	8	1	5		
5	F	1	Total	C	N	O	0	0
			14	8	1	5		
5	M	1	Total	C	N	O	0	0
			14	8	1	5		
5	M	1	Total	C	N	O	0	0
			14	8	1	5		
5	M	1	Total	C	N	O	0	0
			14	8	1	5		
5	M	1	Total	C	N	O	0	0
			14	8	1	5		
5	M	1	Total	C	N	O	0	0
			14	8	1	5		
5	N	1	Total	C	N	O	0	0
			14	8	1	5		
5	O	1	Total	C	N	O	0	0
			14	8	1	5		
5	O	1	Total	C	N	O	0	0
			14	8	1	5		
5	O	1	Total	C	N	O	0	0
			14	8	1	5		
5	O	1	Total	C	N	O	0	0
			14	8	1	5		
5	O	1	Total	C	N	O	0	0
			14	8	1	5		
5	P	1	Total	C	N	O	0	0
			14	8	1	5		
5	Q	1	Total	C	N	O	0	0
			14	8	1	5		
5	Q	1	Total	C	N	O	0	0
			14	8	1	5		
5	Q	1	Total	C	N	O	0	0
			14	8	1	5		
5	Q	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
5	Q	1	Total C N O 14 8 1 5	0	0
5	R	1	Total C N O 14 8 1 5	0	0

- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	173	Total O 173 173	0	0
6	B	81	Total O 81 81	0	0
6	C	156	Total O 156 156	0	0
6	D	74	Total O 74 74	0	0
6	E	110	Total O 110 110	0	0
6	F	75	Total O 75 75	0	0
6	G	29	Total O 29 29	0	0
6	H	62	Total O 62 62	0	0
6	I	34	Total O 34 34	0	0
6	J	29	Total O 29 29	0	0
6	K	30	Total O 30 30	0	0
6	L	75	Total O 75 75	0	0
6	M	126	Total O 126 126	0	0
6	N	50	Total O 50 50	0	0
6	O	96	Total O 96 96	0	0
6	P	40	Total O 40 40	0	0
6	Q	128	Total O 128 128	0	0

Continued on next page...

Continued from previous page...

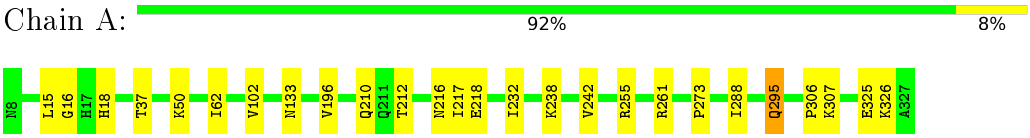
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	R	51	Total O 51 51	0	0
6	S	9	Total O 9 9	0	0
6	T	19	Total O 19 19	0	0
6	U	61	Total O 61 61	0	0
6	V	55	Total O 55 55	0	0
6	X	55	Total O 55 55	0	0
6	Y	51	Total O 51 51	0	0

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.

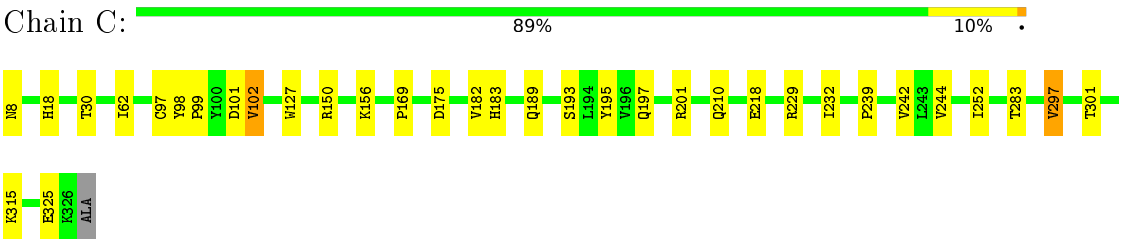
- Molecule 1: Hemagglutinin HA1

Chain A:



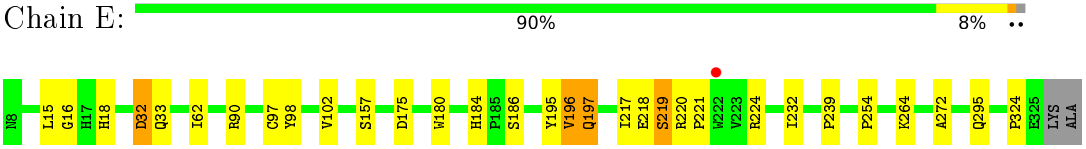
- Molecule 1: Hemagglutinin HA1

Chain C:



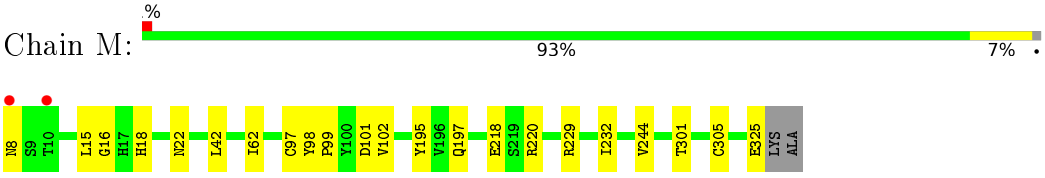
- Molecule 1: Hemagglutinin HA1

Chain E:



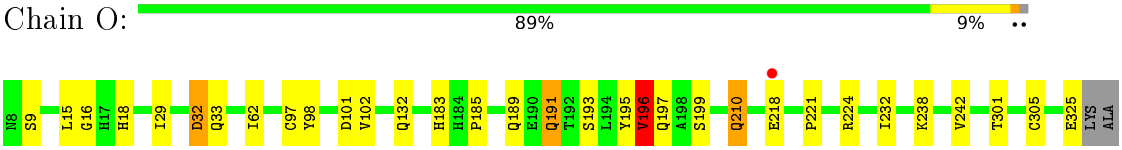
- Molecule 1: Hemagglutinin HA1

Chain M:



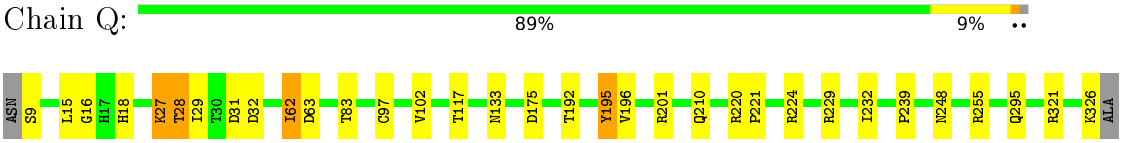
- Molecule 1: Hemagglutinin HA1

Chain O:



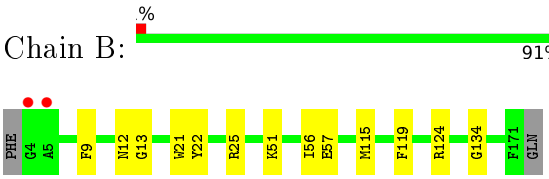
• Molecule 1: Hemagglutinin HA1

Chain Q:



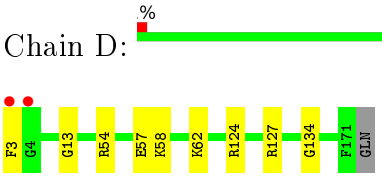
• Molecule 2: Hemagglutinin HA2

Chain B:



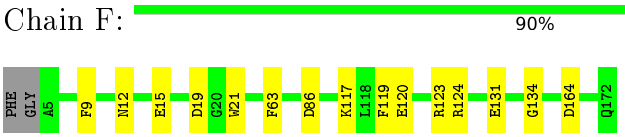
• Molecule 2: Hemagglutinin HA2

Chain D:



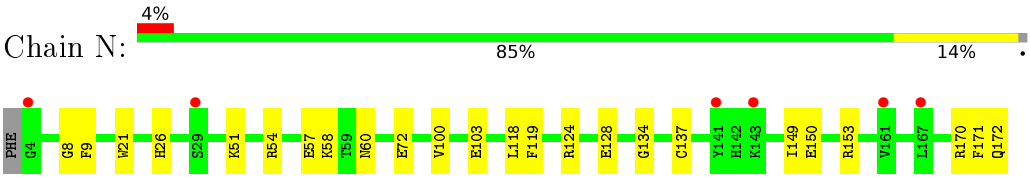
• Molecule 2: Hemagglutinin HA2

Chain F:



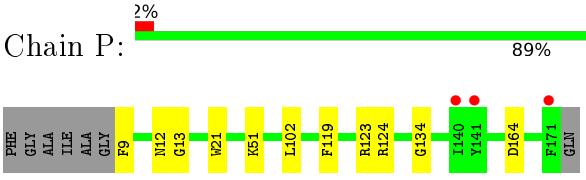
• Molecule 2: Hemagglutinin HA2

Chain N:



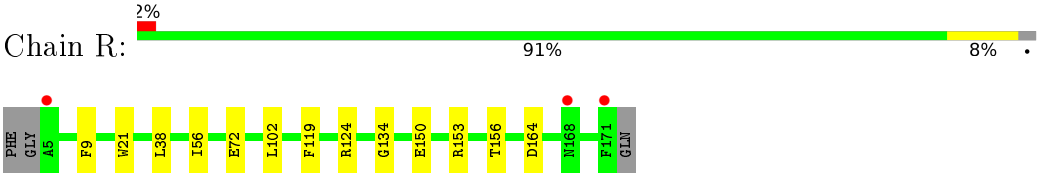
• Molecule 2: Hemagglutinin HA2

Chain P:



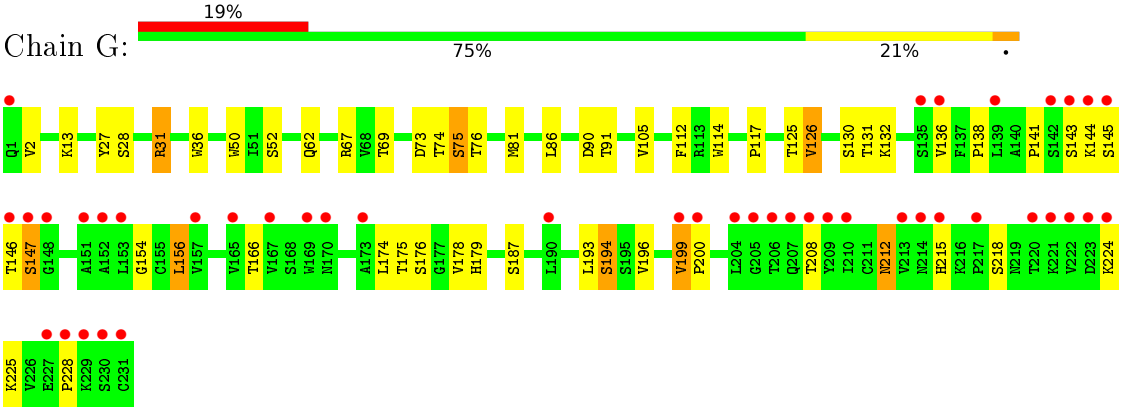
• Molecule 2: Hemagglutinin HA2

Chain R:



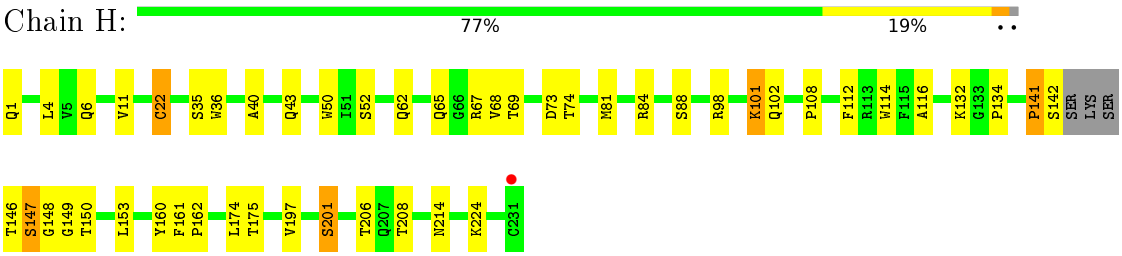
• Molecule 3: 16.a.26 Heavy chain

Chain G:



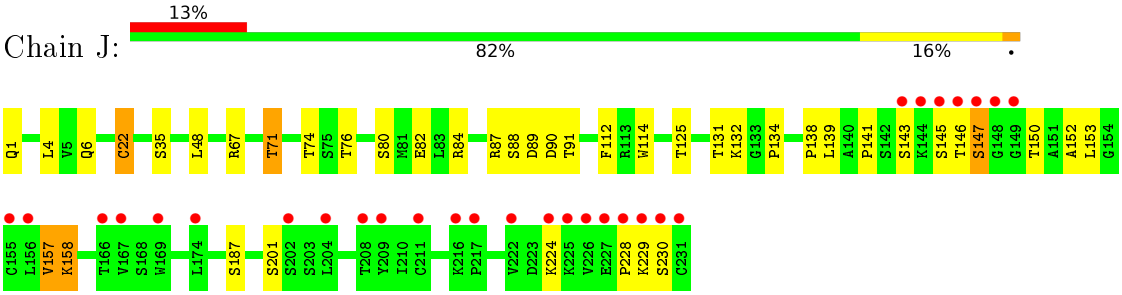
• Molecule 3: 16.a.26 Heavy chain

Chain H:



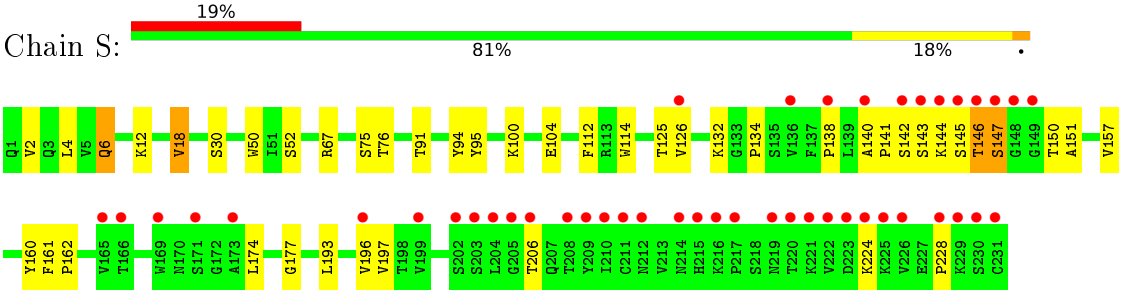
• Molecule 3: 16.a.26 Heavy chain

Chain J:



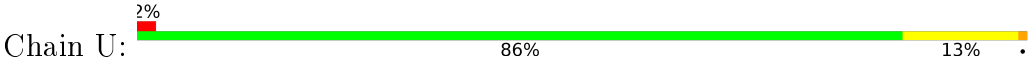
• Molecule 3: 16.a.26 Heavy chain

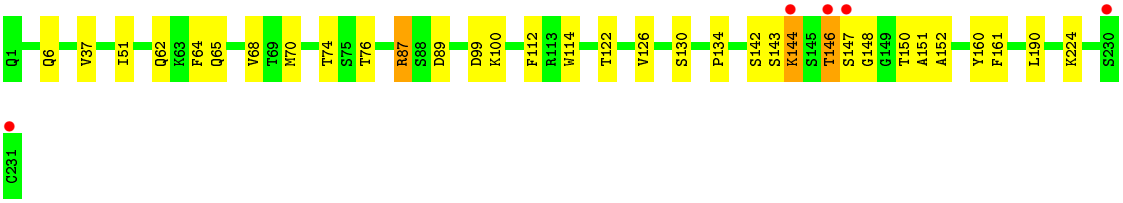
Chain S:



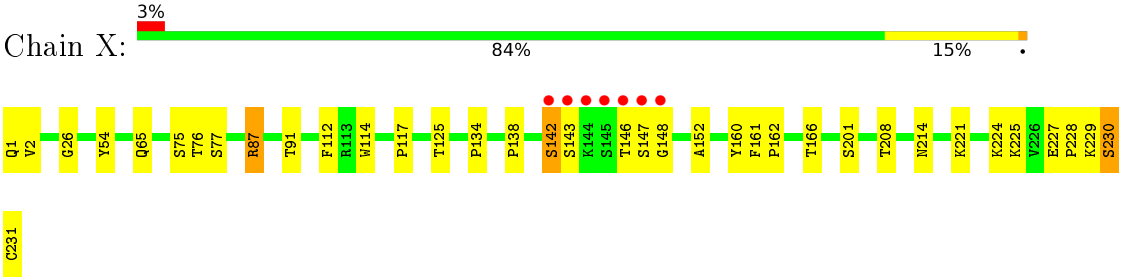
• Molecule 3: 16.a.26 Heavy chain

Chain U:

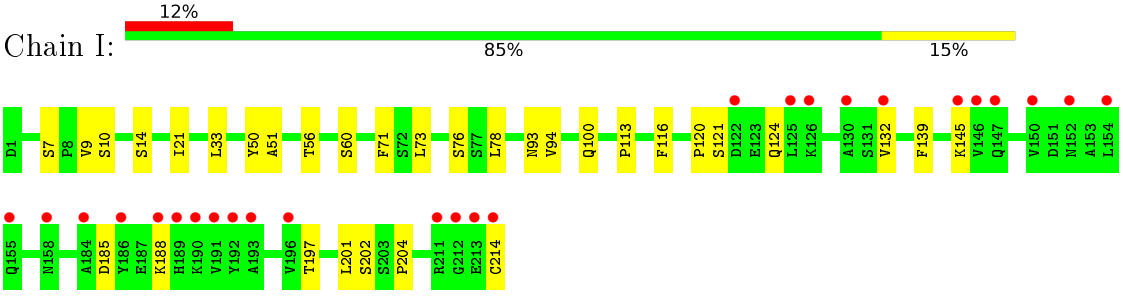




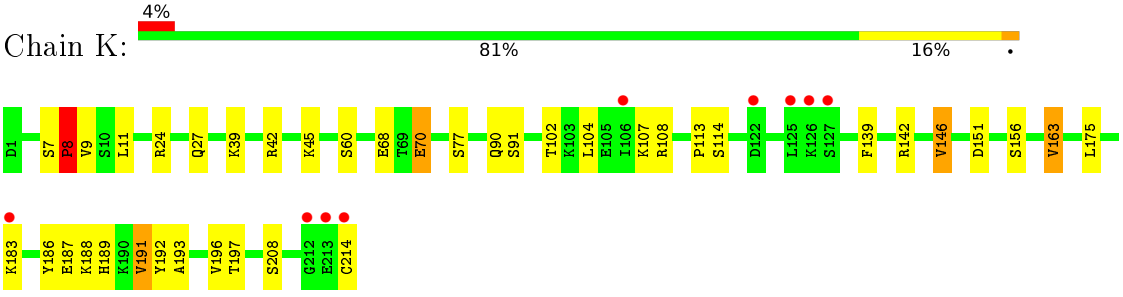
• Molecule 3: 16.a.26 Heavy chain



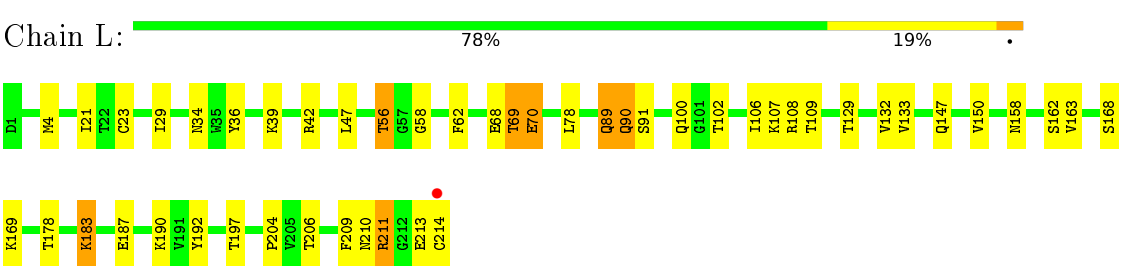
• Molecule 4: 16.a.26 Light chain



• Molecule 4: 16.a.26 Light chain

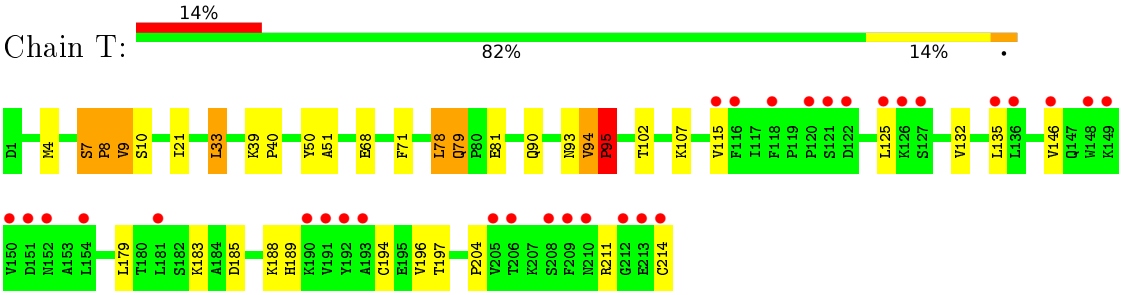


• Molecule 4: 16.a.26 Light chain



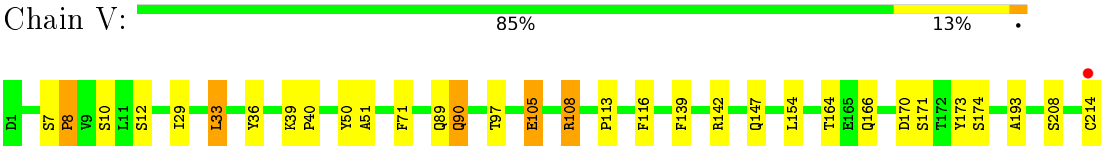
• Molecule 4: 16.a.26 Light chain

Chain T:



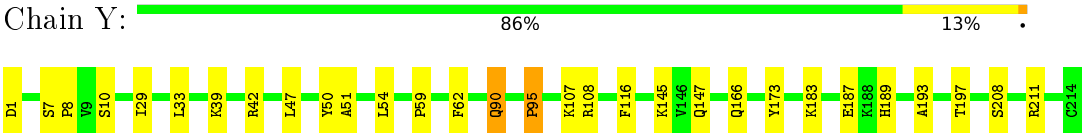
• Molecule 4: 16.a.26 Light chain

Chain V:



• Molecule 4: 16.a.26 Light chain

Chain Y:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	121.81Å 233.91Å 302.45Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.28 – 2.50 49.28 – 2.50	Depositor EDS
% Data completeness (in resolution range)	82.5 (49.28-2.50) 82.5 (49.28-2.50)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.02 (at 2.51Å)	Xtriage
Refinement program	PHENIX (1.10.1_2155: ???)	Depositor
R, R_{free}	0.182 , 0.218 0.180 , 0.215	Depositor DCC
R_{free} test set	12271 reflections (5.02%)	DCC
Wilson B-factor (Å^2)	46.3	Xtriage
Anisotropy	0.015	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.29 , 44.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	45420	wwPDB-VP
Average B, all atoms (Å^2)	61.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.17% 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, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality (i)

5.1 Standard geometry (i)

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	0.33	0/2524	0.49	0/3441
1	C	0.35	0/2523	0.47	0/3438
1	E	0.33	0/2514	0.47	0/3427
1	M	0.32	0/2514	0.48	0/3427
1	O	0.32	0/2514	0.47	0/3427
1	Q	0.32	0/2515	0.50	0/3427
2	B	0.28	0/1389	0.42	0/1867
2	D	0.33	0/1401	0.44	0/1883
2	F	0.35	0/1394	0.45	0/1874
2	N	0.33	0/1391	0.42	0/1870
2	P	0.28	0/1363	0.42	0/1832
2	R	0.28	0/1385	0.43	0/1862
3	G	0.33	0/1771	0.50	0/2411
3	H	0.32	0/1749	0.50	0/2381
3	J	0.29	0/1771	0.48	0/2411
3	S	0.30	0/1771	0.48	0/2411
3	U	0.31	0/1771	0.49	0/2411
3	X	0.34	0/1771	0.52	0/2411
4	I	0.34	0/1693	0.49	0/2294
4	K	0.39	1/1693 (0.1%)	0.54	1/2294 (0.0%)
4	L	0.35	0/1693	0.53	1/2294 (0.0%)
4	T	0.48	3/1693 (0.2%)	0.56	1/2294 (0.0%)
4	V	0.47	2/1693 (0.1%)	0.55	1/2294 (0.0%)
4	Y	0.51	2/1693 (0.1%)	0.56	1/2294 (0.0%)
All	All	0.35	8/44189 (0.0%)	0.49	5/59975 (0.0%)

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	Y	95	PRO	N-CD	5.64	1.55	1.47
4	T	95	PRO	N-CD	5.57	1.55	1.47

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	K	8	PRO	N-CD	5.55	1.55	1.47
4	Y	8	PRO	N-CD	5.45	1.55	1.47
4	V	8	PRO	N-CD	5.31	1.55	1.47

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	K	39	LYS	C-N-CD	5.95	140.89	128.40
4	Y	39	LYS	C-N-CD	5.92	140.82	128.40
4	T	39	LYS	C-N-CD	5.81	140.60	128.40
4	L	58	GLY	C-N-CD	5.67	140.31	128.40
4	V	39	LYS	C-N-CD	5.58	140.13	128.40

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	2468	0	2406	19	0
1	C	2467	0	2412	21	0
1	E	2458	0	2399	27	0
1	M	2458	0	2399	18	0
1	O	2458	0	2399	29	0
1	Q	2459	0	2406	35	0
2	B	1366	0	1282	15	0
2	D	1377	0	1291	8	0
2	F	1371	0	1287	15	0
2	N	1368	0	1278	17	0
2	P	1340	0	1255	13	0
2	R	1362	0	1279	16	0
3	G	1728	0	1705	36	0
3	H	1707	0	1681	23	1
3	J	1728	0	1705	20	0
3	S	1728	0	1705	43	0
3	U	1728	0	1705	24	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	X	1728	0	1705	31	2
4	I	1658	0	1609	14	0
4	K	1658	0	1609	12	1
4	L	1658	0	1609	26	0
4	T	1658	0	1611	25	0
4	V	1658	0	1609	19	0
4	Y	1658	0	1609	11	0
5	A	70	0	65	0	0
5	B	14	0	13	0	0
5	C	70	0	65	2	0
5	D	14	0	13	0	0
5	E	70	0	65	0	0
5	F	14	0	13	0	0
5	M	70	0	65	1	0
5	N	14	0	13	0	0
5	O	70	0	65	0	0
5	P	14	0	13	0	0
5	Q	70	0	65	0	0
5	R	14	0	13	0	0
6	A	173	0	0	2	0
6	B	81	0	0	1	0
6	C	156	0	0	3	0
6	D	74	0	0	0	0
6	E	110	0	0	3	0
6	F	75	0	0	2	0
6	G	29	0	0	1	0
6	H	62	0	0	2	0
6	I	34	0	0	0	0
6	J	29	0	0	0	0
6	K	30	0	0	1	0
6	L	75	0	0	3	0
6	M	126	0	0	2	0
6	N	50	0	0	1	0
6	O	96	0	0	2	0
6	P	40	0	0	0	0
6	Q	128	0	0	3	0
6	R	51	0	0	2	0
6	S	9	0	0	0	0
6	T	19	0	0	0	0
6	U	61	0	0	2	0
6	V	55	0	0	1	0
6	X	55	0	0	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	Y	51	0	0	1	0
All	All	45420	0	42423	438	2

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 5.

The worst 5 of 438 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:S:146:THR:HG21	3:S:151:ALA:CA	1.48	1.43
3:S:146:THR:CG2	3:S:151:ALA:HA	1.52	1.38
3:S:146:THR:HG21	3:S:151:ALA:CB	1.54	1.36
3:J:147:SER:CB	3:J:150:THR:O	1.77	1.32
3:J:147:SER:HB2	3:J:150:THR:O	1.27	1.28

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:K:156:SER:OG	3:X:231:CYS:SG[1_455]	1.94	0.26
3:H:1:GLN:N	3:X:26:GLY:O[1_455]	2.15	0.05

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	318/320 (99%)	306 (96%)	11 (4%)	1 (0%)	46 68
1	C	317/320 (99%)	308 (97%)	8 (2%)	1 (0%)	46 68
1	E	316/320 (99%)	306 (97%)	9 (3%)	1 (0%)	46 68
1	M	316/320 (99%)	306 (97%)	9 (3%)	1 (0%)	46 68

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	O	316/320 (99%)	308 (98%)	6 (2%)	2 (1%)	30 50
1	Q	316/320 (99%)	307 (97%)	8 (2%)	1 (0%)	46 68
2	B	166/170 (98%)	160 (96%)	6 (4%)	0	100 100
2	D	167/170 (98%)	159 (95%)	8 (5%)	0	100 100
2	F	166/170 (98%)	160 (96%)	6 (4%)	0	100 100
2	N	167/170 (98%)	160 (96%)	7 (4%)	0	100 100
2	P	161/170 (95%)	155 (96%)	6 (4%)	0	100 100
2	R	165/170 (97%)	158 (96%)	7 (4%)	0	100 100
3	G	229/231 (99%)	218 (95%)	11 (5%)	0	100 100
3	H	224/231 (97%)	217 (97%)	6 (3%)	1 (0%)	39 61
3	J	229/231 (99%)	222 (97%)	7 (3%)	0	100 100
3	S	229/231 (99%)	218 (95%)	10 (4%)	1 (0%)	39 61
3	U	229/231 (99%)	219 (96%)	10 (4%)	0	100 100
3	X	229/231 (99%)	219 (96%)	10 (4%)	0	100 100
4	I	212/214 (99%)	206 (97%)	6 (3%)	0	100 100
4	K	212/214 (99%)	207 (98%)	5 (2%)	0	100 100
4	L	212/214 (99%)	208 (98%)	3 (1%)	1 (0%)	34 55
4	T	212/214 (99%)	204 (96%)	7 (3%)	1 (0%)	34 55
4	V	212/214 (99%)	207 (98%)	5 (2%)	0	100 100
4	Y	212/214 (99%)	208 (98%)	4 (2%)	0	100 100
All	All	5532/5610 (99%)	5346 (97%)	175 (3%)	11 (0%)	52 75

5 of 11 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	T	9	VAL
1	O	196	VAL
1	A	62	ILE
1	C	62	ILE
1	E	62	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	281/282 (100%)	275 (98%)	6 (2%)	61 85
1	C	282/282 (100%)	275 (98%)	7 (2%)	55 82
1	E	281/282 (100%)	275 (98%)	6 (2%)	61 85
1	M	281/282 (100%)	280 (100%)	1 (0%)	93 98
1	O	281/282 (100%)	269 (96%)	12 (4%)	35 61
1	Q	281/282 (100%)	273 (97%)	8 (3%)	51 78
2	B	144/146 (99%)	143 (99%)	1 (1%)	88 97
2	D	145/146 (99%)	145 (100%)	0	100 100
2	F	145/146 (99%)	144 (99%)	1 (1%)	88 97
2	N	144/146 (99%)	142 (99%)	2 (1%)	74 91
2	P	143/146 (98%)	141 (99%)	2 (1%)	74 91
2	R	144/146 (99%)	141 (98%)	3 (2%)	61 85
3	G	194/194 (100%)	172 (89%)	22 (11%)	7 13
3	H	191/194 (98%)	175 (92%)	16 (8%)	14 25
3	J	194/194 (100%)	175 (90%)	19 (10%)	10 19
3	S	194/194 (100%)	183 (94%)	11 (6%)	25 46
3	U	194/194 (100%)	188 (97%)	6 (3%)	47 75
3	X	194/194 (100%)	187 (96%)	7 (4%)	42 69
4	I	189/189 (100%)	174 (92%)	15 (8%)	15 28
4	K	189/189 (100%)	169 (89%)	20 (11%)	8 16
4	L	189/189 (100%)	169 (89%)	20 (11%)	8 16
4	T	189/189 (100%)	179 (95%)	10 (5%)	28 50
4	V	189/189 (100%)	181 (96%)	8 (4%)	36 62
4	Y	189/189 (100%)	180 (95%)	9 (5%)	31 55
All	All	4847/4866 (100%)	4635 (96%)	212 (4%)	35 60

5 of 212 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	K	9	VAL
4	L	90	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
4	V	214	CYS
4	K	45	LYS
4	K	142	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 23 such sidechains are listed below:

Mol	Chain	Res	Type
4	I	93	ASN
1	M	18	HIS
3	X	65	GLN
4	L	89	GLN
1	M	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

There are no carbohydrates in this entry.

5.6 Ligand geometry

36 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
5	NAG	A	401	1	14,14,15	0.19	0	15,19,21	0.64	1 (6%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
5	NAG	A	402	1	14,14,15	0.32	0	15,19,21	0.33	0
5	NAG	A	403	1	14,14,15	0.21	0	15,19,21	0.38	0
5	NAG	A	404	1	14,14,15	0.45	0	15,19,21	0.66	1 (6%)
5	NAG	A	405	1	14,14,15	0.22	0	15,19,21	0.21	0
5	NAG	B	201	2	14,14,15	0.26	0	15,19,21	0.26	0
5	NAG	C	401	1	14,14,15	0.40	0	15,19,21	0.46	0
5	NAG	C	402	1	14,14,15	0.40	0	15,19,21	0.44	0
5	NAG	C	403	1	14,14,15	1.21	1 (7%)	15,19,21	1.24	2 (13%)
5	NAG	C	404	1	14,14,15	0.24	0	15,19,21	0.27	0
5	NAG	C	405	1	14,14,15	0.20	0	15,19,21	0.23	0
5	NAG	D	201	2	14,14,15	0.18	0	15,19,21	0.29	0
5	NAG	E	401	1	14,14,15	0.37	0	15,19,21	0.55	0
5	NAG	E	402	1	14,14,15	0.29	0	15,19,21	0.28	0
5	NAG	E	403	1	14,14,15	0.17	0	15,19,21	0.34	0
5	NAG	E	404	1	14,14,15	0.32	0	15,19,21	0.27	0
5	NAG	E	405	1	14,14,15	0.22	0	15,19,21	0.37	0
5	NAG	F	201	2	14,14,15	0.18	0	15,19,21	0.40	0
5	NAG	M	401	1	14,14,15	0.47	0	15,19,21	0.48	0
5	NAG	M	402	1	14,14,15	0.42	0	15,19,21	0.48	0
5	NAG	M	403	1	14,14,15	0.95	1 (7%)	15,19,21	1.24	1 (6%)
5	NAG	M	404	1	14,14,15	0.26	0	15,19,21	0.25	0
5	NAG	M	405	1	14,14,15	0.22	0	15,19,21	0.35	0
5	NAG	N	201	2	14,14,15	0.18	0	15,19,21	0.34	0
5	NAG	O	401	1	14,14,15	0.24	0	15,19,21	0.34	0
5	NAG	O	402	1	14,14,15	0.38	0	15,19,21	0.40	0
5	NAG	O	403	1	14,14,15	0.19	0	15,19,21	0.46	0
5	NAG	O	404	1	14,14,15	0.25	0	15,19,21	0.27	0
5	NAG	O	405	1	14,14,15	0.27	0	15,19,21	0.27	0
5	NAG	P	201	2	14,14,15	0.21	0	15,19,21	0.37	0
5	NAG	Q	401	1	14,14,15	0.31	0	15,19,21	0.22	0
5	NAG	Q	402	1	14,14,15	0.33	0	15,19,21	0.43	0
5	NAG	Q	403	1	14,14,15	0.15	0	15,19,21	0.32	0
5	NAG	Q	404	1	14,14,15	0.28	0	15,19,21	0.32	0
5	NAG	Q	405	1	14,14,15	0.19	0	15,19,21	0.28	0
5	NAG	R	201	2	14,14,15	0.15	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
5	NAG	A	401	1	-	0/6/23/26	0/1/1/1
5	NAG	A	402	1	-	0/6/23/26	0/1/1/1
5	NAG	A	403	1	-	0/6/23/26	0/1/1/1
5	NAG	A	404	1	-	0/6/23/26	0/1/1/1
5	NAG	A	405	1	-	0/6/23/26	0/1/1/1
5	NAG	B	201	2	-	0/6/23/26	0/1/1/1
5	NAG	C	401	1	-	0/6/23/26	0/1/1/1
5	NAG	C	402	1	-	0/6/23/26	0/1/1/1
5	NAG	C	403	1	-	0/6/23/26	0/1/1/1
5	NAG	C	404	1	-	0/6/23/26	0/1/1/1
5	NAG	C	405	1	-	0/6/23/26	0/1/1/1
5	NAG	D	201	2	-	0/6/23/26	0/1/1/1
5	NAG	E	401	1	-	0/6/23/26	0/1/1/1
5	NAG	E	402	1	-	0/6/23/26	0/1/1/1
5	NAG	E	403	1	-	0/6/23/26	0/1/1/1
5	NAG	E	404	1	-	0/6/23/26	0/1/1/1
5	NAG	E	405	1	-	0/6/23/26	0/1/1/1
5	NAG	F	201	2	-	0/6/23/26	0/1/1/1
5	NAG	M	401	1	-	0/6/23/26	0/1/1/1
5	NAG	M	402	1	-	0/6/23/26	0/1/1/1
5	NAG	M	403	1	-	0/6/23/26	0/1/1/1
5	NAG	M	404	1	-	0/6/23/26	0/1/1/1
5	NAG	M	405	1	-	0/6/23/26	0/1/1/1
5	NAG	N	201	2	-	0/6/23/26	0/1/1/1
5	NAG	O	401	1	-	0/6/23/26	0/1/1/1
5	NAG	O	402	1	-	0/6/23/26	0/1/1/1
5	NAG	O	403	1	-	0/6/23/26	0/1/1/1
5	NAG	O	404	1	-	0/6/23/26	0/1/1/1
5	NAG	O	405	1	-	0/6/23/26	0/1/1/1
5	NAG	P	201	2	-	0/6/23/26	0/1/1/1
5	NAG	Q	401	1	-	0/6/23/26	0/1/1/1
5	NAG	Q	402	1	-	0/6/23/26	0/1/1/1
5	NAG	Q	403	1	-	0/6/23/26	0/1/1/1
5	NAG	Q	404	1	-	0/6/23/26	0/1/1/1
5	NAG	Q	405	1	-	0/6/23/26	0/1/1/1
5	NAG	R	201	2	-	0/6/23/26	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	M	403	NAG	O5-C1	3.37	1.49	1.43
5	C	403	NAG	C1-C2	4.05	1.58	1.52

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	C	403	NAG	O5-C5-C4	-2.84	105.44	110.13
5	A	401	NAG	C1-O5-C5	2.02	115.12	112.14
5	A	404	NAG	C1-O5-C5	2.08	115.20	112.14
5	C	403	NAG	C1-O5-C5	2.65	116.03	112.14
5	M	403	NAG	C1-O5-C5	4.51	118.77	112.14

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	C	403	NAG	1	0
5	C	405	NAG	1	0
5	M	401	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	320/320 (100%)	-0.43	0 100 100	27, 38, 64, 120	0
1	C	319/320 (99%)	-0.42	0 100 100	27, 43, 68, 109	1 (0%)
1	E	318/320 (99%)	-0.35	1 (0%) 94 95	33, 50, 79, 106	0
1	M	318/320 (99%)	-0.38	2 (0%) 90 91	30, 43, 72, 124	2 (0%)
1	O	318/320 (99%)	-0.26	1 (0%) 94 95	36, 53, 87, 134	1 (0%)
1	Q	318/320 (99%)	-0.40	0 100 100	31, 44, 66, 102	2 (0%)
2	B	168/170 (98%)	-0.12	2 (1%) 81 83	28, 45, 70, 161	2 (1%)
2	D	169/170 (99%)	-0.14	2 (1%) 81 83	28, 47, 72, 167	0
2	F	168/170 (98%)	-0.22	0 100 100	31, 46, 65, 131	2 (1%)
2	N	169/170 (99%)	0.03	6 (3%) 46 51	32, 69, 105, 139	2 (1%)
2	P	163/170 (95%)	-0.02	3 (1%) 71 75	31, 64, 101, 148	2 (1%)
2	R	167/170 (98%)	-0.17	3 (1%) 71 75	30, 55, 98, 122	0
3	G	231/231 (100%)	0.57	44 (19%) 2 1	35, 78, 164, 266	0
3	H	228/231 (98%)	-0.36	1 (0%) 93 93	28, 52, 86, 142	0
3	J	231/231 (100%)	0.44	29 (12%) 5 5	35, 77, 166, 268	1 (0%)
3	S	231/231 (100%)	0.87	45 (19%) 1 1	51, 94, 161, 280	0
3	U	231/231 (100%)	-0.20	5 (2%) 65 69	33, 54, 88, 174	0
3	X	231/231 (100%)	-0.22	7 (3%) 54 59	35, 57, 100, 201	0
4	I	214/214 (100%)	0.25	26 (12%) 6 5	34, 74, 161, 221	1 (0%)
4	K	214/214 (100%)	0.07	9 (4%) 40 45	33, 79, 114, 187	0
4	L	214/214 (100%)	-0.25	1 (0%) 91 92	25, 52, 105, 163	0
4	T	214/214 (100%)	0.50	31 (14%) 3 3	24, 83, 169, 223	2 (0%)
4	V	214/214 (100%)	-0.24	1 (0%) 91 92	24, 51, 79, 153	0
4	Y	214/214 (100%)	-0.27	0 100 100	24, 56, 91, 118	1 (0%)

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å^2)	Q<0.9
All	All	5582/5610 (99%)	-0.10	219 (3%) 43 48	24, 53, 125, 280	19 (0%)

The worst 5 of 219 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	J	148	GLY	11.0
4	T	214	CYS	11.0
3	G	204	LEU	10.6
3	S	146	THR	9.5
3	J	231	CYS	8.9

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)

There are no carbohydrates in this entry.

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
5	NAG	A	403	14/15	0.86	0.28	27.21	78,94,104,106	0
5	NAG	Q	403	14/15	0.87	0.25	13.05	86,95,107,113	0
5	NAG	C	403	14/15	0.73	0.27	9.41	105,118,126,127	0
5	NAG	Q	401	14/15	0.74	0.58	9.23	138,157,160,161	0
5	NAG	A	404	14/15	0.74	0.24	3.95	93,112,127,130	0
5	NAG	Q	402	14/15	0.87	0.18	3.63	81,93,105,110	0
5	NAG	A	402	14/15	0.89	0.23	3.62	76,87,92,97	0
5	NAG	E	403	14/15	0.95	0.18	2.75	57,72,79,82	0
5	NAG	C	404	14/15	0.88	0.33	2.41	104,126,132,137	0
5	NAG	M	405	14/15	0.93	0.17	2.13	59,72,86,86	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å^2)	Q<0.9
5	NAG	C	402	14/15	0.92	0.17	1.44	65,77,83,93	0
5	NAG	E	402	14/15	0.92	0.18	1.28	72,82,91,92	0
5	NAG	M	404	14/15	0.83	0.20	1.25	93,104,108,110	0
5	NAG	O	403	14/15	0.96	0.17	1.20	53,62,70,72	0
5	NAG	O	402	14/15	0.86	0.16	1.12	80,99,106,109	0
5	NAG	O	405	14/15	0.94	0.13	0.81	59,75,87,88	0
5	NAG	M	402	14/15	0.92	0.13	-0.02	77,88,91,92	0
5	NAG	E	405	14/15	0.95	0.10	-1.05	49,62,70,76	0
5	NAG	M	401	14/15	0.75	0.59	-	142,162,166,166	0
5	NAG	B	201	14/15	0.94	0.12	-	52,62,73,85	0
5	NAG	D	201	14/15	0.94	0.15	-	52,67,75,78	0
5	NAG	A	405	14/15	0.92	0.13	-	56,67,76,83	0
5	NAG	Q	404	14/15	0.84	0.24	-	102,123,131,135	0
5	NAG	C	401	14/15	0.43	0.71	-	144,160,163,163	0
5	NAG	M	403	14/15	0.83	0.23	-	95,104,111,111	0
5	NAG	P	201	14/15	0.88	0.22	-	79,93,96,100	0
5	NAG	R	201	14/15	0.91	0.18	-	87,101,112,118	0
5	NAG	O	401	14/15	0.59	0.63	-	149,162,172,172	0
5	NAG	Q	405	14/15	0.94	0.19	-	62,76,87,97	0
5	NAG	N	201	14/15	0.92	0.17	-	83,94,99,105	0
5	NAG	O	404	14/15	0.85	0.18	-	93,109,116,121	0
5	NAG	E	404	14/15	0.90	0.17	-	89,105,123,129	0
5	NAG	C	405	14/15	0.93	0.17	-	64,69,85,86	0
5	NAG	A	401	14/15	0.61	0.62	-	143,153,161,163	0
5	NAG	E	401	14/15	0.53	0.45	-	144,157,161,163	0
5	NAG	F	201	14/15	0.97	0.07	-	48,66,72,75	0

6.5 Other polymers (i)

There are no such residues in this entry.