



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 06:29 PM GMT

PDB ID : 4LN6  
Title : The crystal structure of hemagglutinin from a h7n9 influenza virus  
(a/shanghai/2/2013)  
Authors : Yang, H.; Carney, P.J.; Chang, J.C.; Villanueva, J.M.; Stevens, J.  
Deposited on : 2013-07-11  
Resolution : 2.12 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](mailto:validation@mail.wwpdb.org)  
A user guide is available at  
<http://wwpdb.org/validation/2016/XrayValidationReportHelp>  
with specific help available everywhere you see the ⓘ symbol.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

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

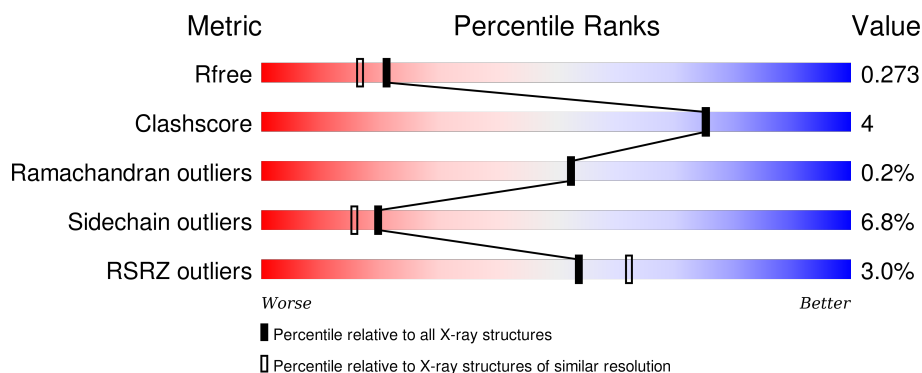
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.12 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	4587 (2.14-2.10)
Clashscore	102246	5132 (2.14-2.10)
Ramachandran outliers	100387	5080 (2.14-2.10)
Sidechain outliers	100360	5081 (2.14-2.10)
RSRZ outliers	91569	4597 (2.14-2.10)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	325	<div> <div>2%</div> <div>86%</div> <div>9%</div> <div>• •</div> </div>
1	C	325	<div> <div>3%</div> <div>85%</div> <div>11%</div> <div>• •</div> </div>
1	E	325	<div> <div>3%</div> <div>85%</div> <div>9%</div> <div>• •</div> </div>
1	G	325	<div> <div>4%</div> <div>85%</div> <div>10%</div> <div>• •</div> </div>
1	I	325	<div> <div>3%</div> <div>86%</div> <div>10%</div> <div>• •</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	K	325	
2	B	181	
2	D	181	
2	F	181	
2	H	181	
2	J	181	
2	L	181	

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
7	NAG	E	401	-	-	-	X

## 2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 23253 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	316	Total	C	N	O	S	1	0	0
			2412	1498	436	463	15			
1	C	316	Total	C	N	O	S	1	0	0
			2412	1498	436	463	15			
1	E	316	Total	C	N	O	S	1	0	0
			2412	1498	436	463	15			
1	G	316	Total	C	N	O	S	1	0	0
			2412	1498	436	463	15			
1	I	316	Total	C	N	O	S	1	0	0
			2412	1498	436	463	15			
1	K	316	Total	C	N	O	S	1	0	0
			2412	1498	436	463	15			

- Molecule 2 is a protein called Hemagglutinin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	167	Total	C	N	O	S	0	0	0
			1360	838	238	277	7			
2	D	167	Total	C	N	O	S	0	0	0
			1360	838	238	277	7			
2	F	167	Total	C	N	O	S	0	0	0
			1360	838	238	277	7			
2	H	169	Total	C	N	O	S	0	0	0
			1375	849	240	279	7			
2	J	169	Total	C	N	O	S	0	0	0
			1375	849	240	279	7			
2	L	169	Total	C	N	O	S	0	0	0
			1375	849	240	279	7			

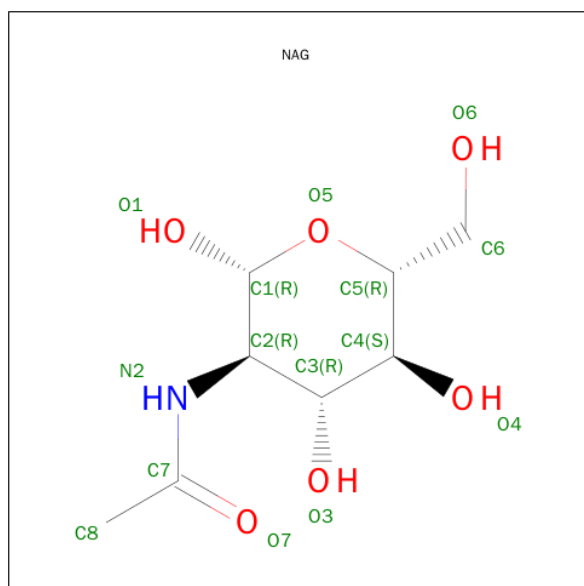
- Molecule 3 is a polymer of unknown type called SUGAR (2-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	2	Total	C	N	O	0	0
			28	16	2	10		
3	G	2	Total	C	N	O	0	0
			28	16	2	10		

- Molecule 4 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	G	1	Total	Ca	0	0
			1	1		
4	K	1	Total	Ca	0	0
			1	1		
4	E	1	Total	Ca	0	0
			1	1		
4	I	1	Total	Ca	0	0
			1	1		
4	C	1	Total	Ca	0	0
			1	1		
4	A	1	Total	Ca	0	0
			1	1		

- Molecule 5 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula: C<sub>8</sub>H<sub>15</sub>NO<sub>6</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	B	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	D	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	H	1	Total	C	N	O	0	0
			14	8	1	5		
5	J	1	Total	C	N	O	0	0
			14	8	1	5		
5	L	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 6 is a polymer of unknown type called SUGAR (3-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	C	3	Total	C	N	O	0	0
			39	22	2	15		
6	E	3	Total	C	N	O	0	0
			39	22	2	15		
6	I	3	Total	C	N	O	0	0
			39	22	2	15		
6	K	3	Total	C	N	O	0	0
			39	22	2	15		

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

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

- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	32	Total	O	0	0
			32	32		
8	B	12	Total	O	0	0
			12	12		
8	C	32	Total	O	0	0
			32	32		
8	D	12	Total	O	0	0
			12	12		
8	E	24	Total	O	0	0
			24	24		

*Continued on next page...*

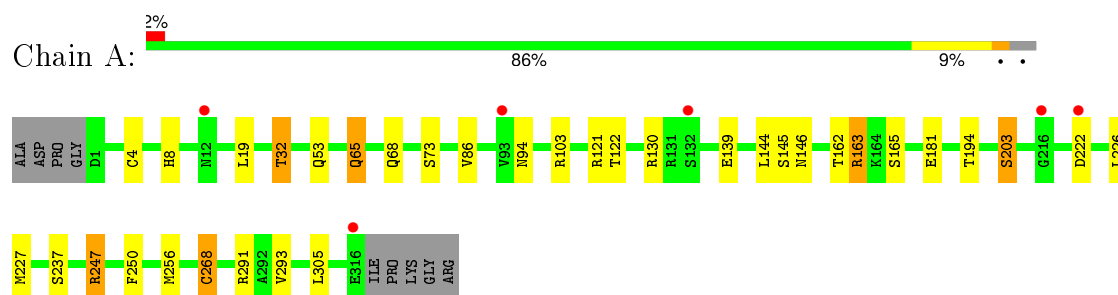
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	F	15	Total 15	O 15	0	0
8	G	24	Total 24	O 24	0	0
8	H	15	Total 15	O 15	0	0
8	I	16	Total 16	O 16	0	0
8	J	14	Total 14	O 14	0	0
8	K	18	Total 18	O 18	0	0
8	L	11	Total 11	O 11	0	0

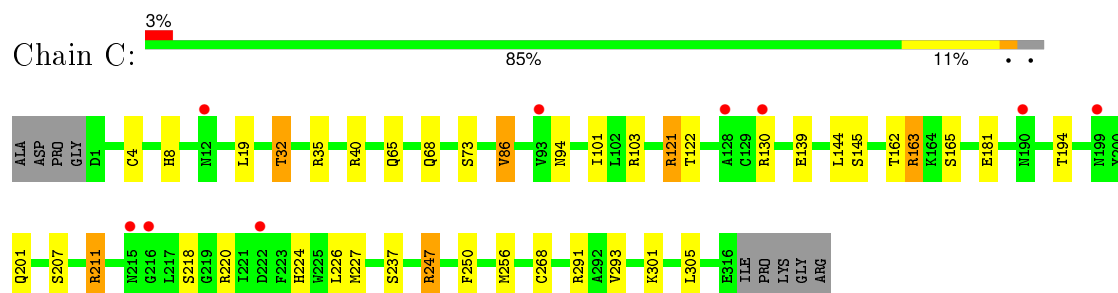
### 3 Residue-property plots [i](#)

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.

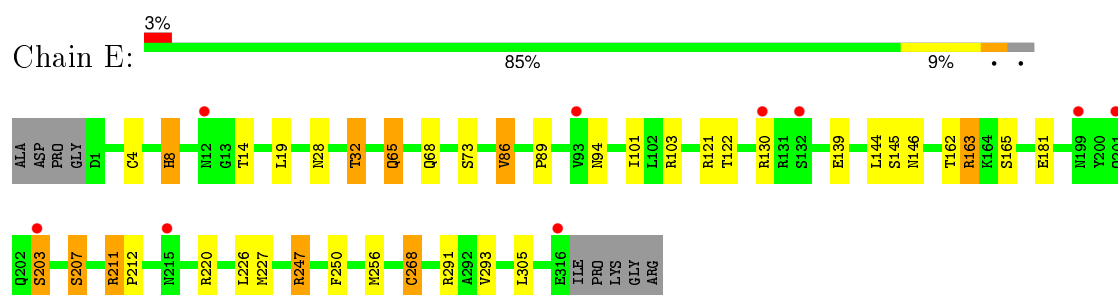
#### • Molecule 1: Hemagglutinin



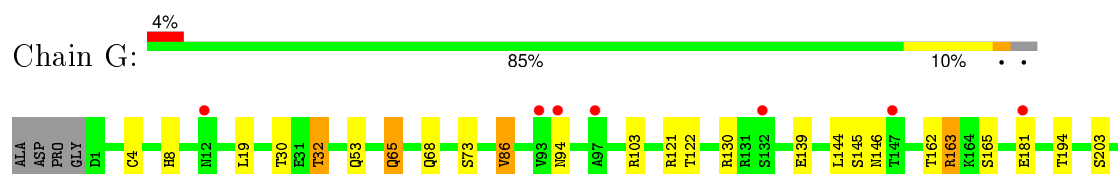
#### • Molecule 1: Hemagglutinin

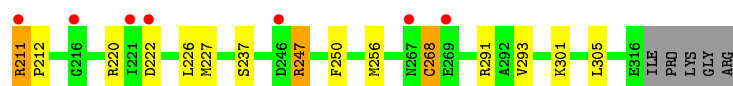


#### • Molecule 1: Hemagglutinin

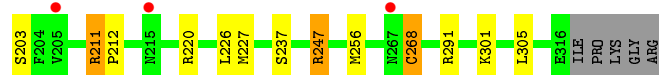
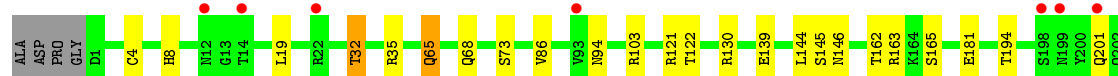
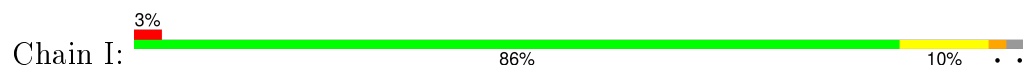


#### • Molecule 1: Hemagglutinin

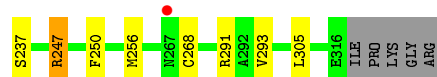
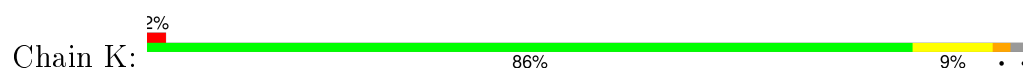




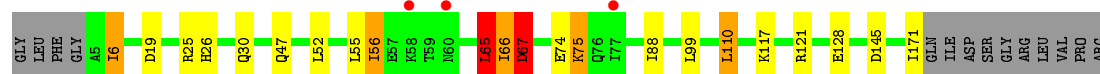
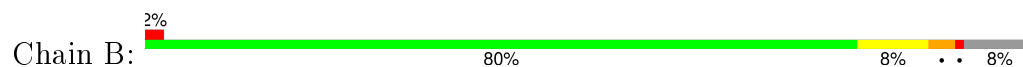
• Molecule 1: Hemagglutinin



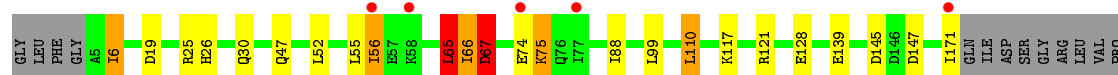
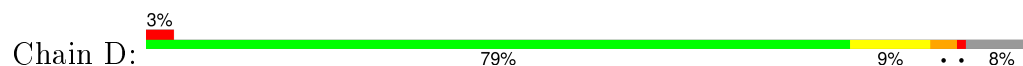
• Molecule 1: Hemagglutinin



• Molecule 2: Hemagglutinin

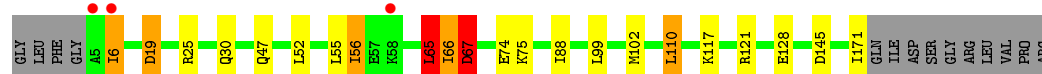
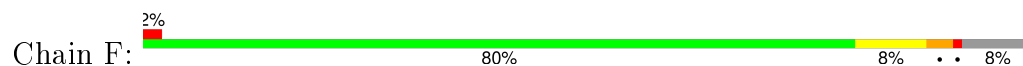


• Molecule 2: Hemagglutinin

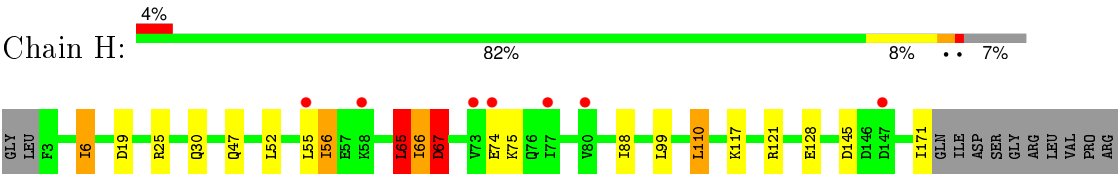


ARG

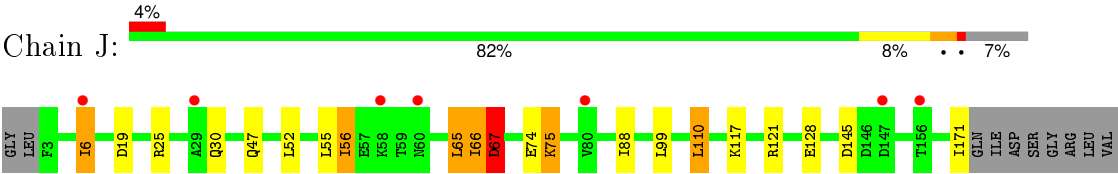
• Molecule 2: Hemagglutinin



• Molecule 2: Hemagglutinin

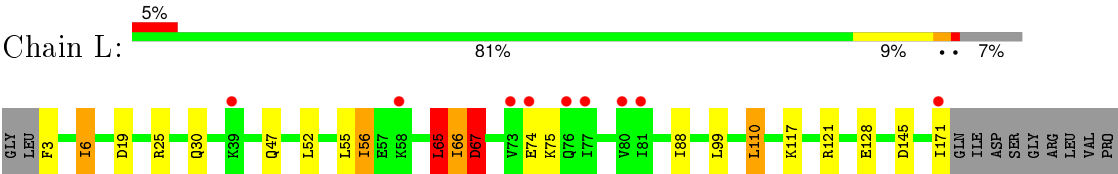


• Molecule 2: Hemagglutinin



PRO  
ARG

• Molecule 2: Hemagglutinin



ARG

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	154.02Å 152.83Å 155.30Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.07 – 2.12 49.07 – 2.12	Depositor EDS
% Data completeness (in resolution range)	95.3 (49.07-2.12) 95.9 (49.07-2.12)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.24 (at 2.12Å)	Xtriage
Refinement program	REFMAC 5.7.0029	Depositor
R, $R_{free}$	0.238 , 0.259 0.251 , 0.273	Depositor DCC
$R_{free}$ test set	10141 reflections (5.35%)	DCC
Wilson B-factor (Å <sup>2</sup> )	33.1	Xtriage
Anisotropy	0.316	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 33.1	EDS
Estimated twinning fraction	0.852 for H, K, L 0.063 for L, H, K 0.085 for -K, -L, H 0.000 for l,-k,h 0.000 for -h,-l,-k 0.000 for k,h,-l 0.096 for l,h,k 0.096 for k,l,h	Xtriage
Reported twinning fraction	0.852 for H, K, L 0.063 for L, H, K 0.085 for -K, -L, H	Depositor
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.54$ , $\langle L^2 \rangle = 0.39$	Xtriage
Outliers	0 of 201333 reflections	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	23253	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	38.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.02% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup> Intensities estimated from amplitudes.

<sup>2</sup> Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

## 5 Model quality ⓘ

### 5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: CA, BMA, NAG, FUC

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	1.02	2/2458 (0.1%)	0.98	6/3322 (0.2%)
1	C	1.18	2/2458 (0.1%)	0.99	9/3322 (0.3%)
1	E	1.08	3/2458 (0.1%)	1.00	9/3322 (0.3%)
1	G	0.98	2/2458 (0.1%)	0.98	9/3322 (0.3%)
1	I	1.02	2/2458 (0.1%)	0.96	8/3322 (0.2%)
1	K	0.88	1/2458 (0.0%)	1.00	7/3322 (0.2%)
2	B	0.99	4/1383 (0.3%)	0.99	6/1864 (0.3%)
2	D	0.98	4/1383 (0.3%)	1.02	6/1864 (0.3%)
2	F	0.96	3/1383 (0.2%)	1.00	8/1864 (0.4%)
2	H	0.92	4/1399 (0.3%)	0.99	6/1885 (0.3%)
2	J	0.90	3/1399 (0.2%)	0.99	7/1885 (0.4%)
2	L	0.96	4/1399 (0.3%)	1.01	7/1885 (0.4%)
All	All	1.00	34/23094 (0.1%)	0.99	88/31179 (0.3%)

The worst 5 of 34 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	268	CYS	CB-SG	-35.85	1.21	1.82
1	I	268	CYS	CB-SG	27.56	2.29	1.82
1	E	268	CYS	CB-SG	23.03	2.21	1.82
1	G	268	CYS	CB-SG	-22.18	1.44	1.82
1	A	268	CYS	CB-SG	-22.00	1.44	1.82

The worst 5 of 88 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	K	268	CYS	CA-CB-SG	16.51	143.71	114.00
2	D	67	ASP	CB-CG-OD1	-14.75	105.03	118.30
2	L	67	ASP	CB-CG-OD1	-14.46	105.28	118.30
2	H	67	ASP	CB-CG-OD1	-14.29	105.44	118.30
1	E	247	ARG	NE-CZ-NH1	13.75	127.17	120.30

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	2412	0	2375	19	0
1	C	2412	0	2375	28	0
1	E	2412	0	2374	21	0
1	G	2412	0	2375	18	1
1	I	2412	0	2374	13	1
1	K	2412	0	2375	15	0
2	B	1360	0	1262	15	0
2	D	1360	0	1262	17	0
2	F	1360	0	1262	12	0
2	H	1375	0	1274	11	0
2	J	1375	0	1274	15	0
2	L	1375	0	1274	16	0
3	A	28	0	25	0	0
3	G	28	0	25	1	0
4	A	1	0	0	0	0
4	C	1	0	0	0	0
4	E	1	0	0	0	0
4	G	1	0	0	0	0
4	I	1	0	0	0	0
4	K	1	0	0	0	0
5	B	14	0	13	0	0
5	D	14	0	13	0	0
5	F	14	0	13	0	0
5	H	14	0	13	0	0
5	J	14	0	13	1	0
5	L	14	0	13	0	1
6	C	39	0	34	0	0
6	E	39	0	34	0	0
6	I	39	0	34	1	0
6	K	39	0	34	1	0
7	E	49	0	43	3	0
8	A	32	0	0	1	0
8	B	12	0	0	1	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
8	C	32	0	0	5	0
8	D	12	0	0	4	0
8	E	24	0	0	1	0
8	F	15	0	0	0	0
8	G	24	0	0	3	1
8	H	15	0	0	0	0
8	I	16	0	0	0	0
8	J	14	0	0	0	0
8	K	18	0	0	0	0
8	L	11	0	0	3	0
All	All	23253	0	22163	175	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 4.

The worst 5 of 175 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:C:121:ARG:NH1	1:C:145:SER:O	1.77	1.18
1:C:121:ARG:HG3	1:C:121:ARG:HH11	1.15	1.12
1:A:53:GLN:OE1	8:A:501:HOH:O	1.89	0.89
1:A:222:ASP:OD1	1:C:201:GLN:NE2	2.07	0.87
1:G:53:GLN:OE1	8:G:501:HOH:O	1.96	0.83

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)
1:G:222:ASP:OD2	1:I:201:GLN:NE2[3_545]	1.66	0.54
5:L:500:NAG:O7	8:G:508:HOH:O[3_555]	2.07	0.13

## 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	314/325 (97%)	307 (98%)	7 (2%)	0	100	100
1	C	314/325 (97%)	306 (98%)	8 (2%)	0	100	100
1	E	314/325 (97%)	305 (97%)	9 (3%)	0	100	100
1	G	314/325 (97%)	305 (97%)	9 (3%)	0	100	100
1	I	314/325 (97%)	305 (97%)	9 (3%)	0	100	100
1	K	314/325 (97%)	305 (97%)	9 (3%)	0	100	100
2	B	165/181 (91%)	160 (97%)	4 (2%)	1 (1%)	30	24
2	D	165/181 (91%)	161 (98%)	3 (2%)	1 (1%)	30	24
2	F	165/181 (91%)	160 (97%)	4 (2%)	1 (1%)	30	24
2	H	167/181 (92%)	163 (98%)	3 (2%)	1 (1%)	30	24
2	J	167/181 (92%)	163 (98%)	3 (2%)	1 (1%)	30	24
2	L	167/181 (92%)	162 (97%)	4 (2%)	1 (1%)	30	24
All	All	2880/3036 (95%)	2802 (97%)	72 (2%)	6 (0%)	52	52

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	H	56	ILE
2	J	56	ILE
2	L	56	ILE
2	B	56	ILE
2	D	56	ILE

### 5.3.2 Protein sidechains ⓘ

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	265/271 (98%)	249 (94%)	16 (6%)	24	20
1	C	265/271 (98%)	249 (94%)	16 (6%)	24	20
1	E	265/271 (98%)	247 (93%)	18 (7%)	20	16

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	G	265/271 (98%)	248 (94%)	17 (6%)	22	17
1	I	265/271 (98%)	248 (94%)	17 (6%)	22	17
1	K	265/271 (98%)	249 (94%)	16 (6%)	24	20
2	B	144/155 (93%)	133 (92%)	11 (8%)	16	12
2	D	144/155 (93%)	132 (92%)	12 (8%)	14	9
2	F	144/155 (93%)	133 (92%)	11 (8%)	16	12
2	H	145/155 (94%)	134 (92%)	11 (8%)	16	12
2	J	145/155 (94%)	134 (92%)	11 (8%)	16	12
2	L	145/155 (94%)	134 (92%)	11 (8%)	16	12
All	All	2457/2556 (96%)	2290 (93%)	167 (7%)	20	16

5 of 167 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	F	52	LEU
1	G	130	ARG
1	K	227	MET
2	F	66	ILE
1	G	8	HIS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 18 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	224	HIS
2	D	26	HIS
2	H	30	GLN
1	C	154	GLN
1	C	201	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 ⓘ

20 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$
3	NAG	A	401	1,3	14,14,15	0.78	0	15,19,21	1.88	4 (26%)
3	NAG	A	402	3	14,14,15	0.47	0	15,19,21	1.47	2 (13%)
6	NAG	C	401	1,6	14,14,15	0.82	0	15,19,21	2.36	3 (20%)
6	NAG	C	402	6	14,14,15	0.63	0	15,19,21	1.44	4 (26%)
6	BMA	C	403	6	11,11,12	0.85	0	14,15,17	1.82	3 (21%)
7	NAG	E	401	1,7	14,14,15	2.10	3 (21%)	15,19,21	3.90	7 (46%)
7	NAG	E	402	7	14,14,15	0.97	1 (7%)	15,19,21	1.57	4 (26%)
7	BMA	E	403	7	11,11,12	0.77	0	14,15,17	2.08	3 (21%)
7	FUC	E	404	7	10,10,11	0.87	0	14,14,16	1.96	3 (21%)
6	NAG	E	405	1,6	14,14,15	1.16	1 (7%)	15,19,21	0.87	0
6	NAG	E	406	6	14,14,15	0.65	0	15,19,21	1.24	1 (6%)
6	BMA	E	407	6	11,11,12	0.78	0	14,15,17	2.85	4 (28%)
3	NAG	G	401	1,3	14,14,15	0.59	0	15,19,21	1.47	4 (26%)
3	NAG	G	402	3	14,14,15	0.59	0	15,19,21	1.47	2 (13%)
6	NAG	I	401	1,6	14,14,15	0.64	0	15,19,21	1.06	1 (6%)
6	NAG	I	402	6	14,14,15	0.66	0	15,19,21	1.09	0
6	BMA	I	403	6	11,11,12	0.67	0	14,15,17	2.35	2 (14%)
6	NAG	K	401	1,6	14,14,15	0.94	0	15,19,21	1.70	5 (33%)
6	NAG	K	402	6	14,14,15	0.68	0	15,19,21	1.31	1 (6%)
6	BMA	K	403	6	11,11,12	0.64	0	14,15,17	1.00	1 (7%)

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	1,3	-	0/6/23/26	0/1/1/1
3	NAG	A	402	3	-	0/6/23/26	0/1/1/1
6	NAG	C	401	1,6	-	0/6/23/26	0/1/1/1
6	NAG	C	402	6	-	0/6/23/26	0/1/1/1
6	BMA	C	403	6	-	0/2/19/22	0/1/1/1
7	NAG	E	401	1,7	-	0/6/23/26	0/1/1/1
7	NAG	E	402	7	-	0/6/23/26	0/1/1/1
7	BMA	E	403	7	-	0/2/19/22	0/1/1/1
7	FUC	E	404	7	-	0/0/17/20	0/1/1/1
6	NAG	E	405	1,6	-	0/6/23/26	0/1/1/1
6	NAG	E	406	6	-	0/6/23/26	0/1/1/1
6	BMA	E	407	6	-	0/2/19/22	0/1/1/1
3	NAG	G	401	1,3	-	0/6/23/26	0/1/1/1
3	NAG	G	402	3	-	0/6/23/26	0/1/1/1
6	NAG	I	401	1,6	-	0/6/23/26	0/1/1/1
6	NAG	I	402	6	-	0/6/23/26	0/1/1/1
6	BMA	I	403	6	-	0/2/19/22	0/1/1/1
6	NAG	K	401	1,6	-	0/6/23/26	0/1/1/1
6	NAG	K	402	6	-	0/6/23/26	0/1/1/1
6	BMA	K	403	6	-	0/2/19/22	0/1/1/1

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	E	401	NAG	C3-C2	2.29	1.57	1.52
7	E	401	NAG	O5-C1	2.29	1.47	1.43
7	E	402	NAG	C1-C2	2.84	1.56	1.52
6	E	405	NAG	C1-C2	2.88	1.56	1.52
7	E	401	NAG	C1-C2	6.01	1.60	1.52

The worst 5 of 54 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	401	NAG	C2-N2-C7	-3.93	117.99	123.04
6	K	401	NAG	O7-C7-C8	-3.44	115.74	122.06
6	K	402	NAG	O7-C7-C8	-3.16	116.27	122.06
7	E	402	NAG	C3-C2-N2	-3.12	103.09	110.56
7	E	401	NAG	O7-C7-C8	-3.09	116.40	122.06

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

6 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	E	401	NAG	2	0
7	E	404	FUC	1	0
3	G	401	NAG	1	0
6	I	402	NAG	1	0
6	I	403	BMA	1	0
6	K	401	NAG	1	0

## 5.6 Ligand geometry

Of 12 ligands modelled in this entry, 6 are monoatomic - leaving 6 for Mogul analysis.

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	B	500	2	14,14,15	0.59	0	15,19,21	2.40	5 (33%)
5	NAG	D	500	2	14,14,15	0.93	1 (7%)	15,19,21	1.45	1 (6%)
5	NAG	F	500	2	14,14,15	0.86	0	15,19,21	2.14	6 (40%)
5	NAG	H	500	2	14,14,15	1.05	1 (7%)	15,19,21	2.53	5 (33%)
5	NAG	J	500	2	14,14,15	0.87	1 (7%)	15,19,21	2.05	1 (6%)
5	NAG	L	500	2	14,14,15	0.74	0	15,19,21	2.77	5 (33%)

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	B	500	2	-	0/6/23/26	0/1/1/1
5	NAG	D	500	2	-	0/6/23/26	0/1/1/1
5	NAG	F	500	2	-	0/6/23/26	0/1/1/1
5	NAG	H	500	2	-	0/6/23/26	0/1/1/1
5	NAG	J	500	2	-	0/6/23/26	0/1/1/1
5	NAG	L	500	2	-	0/6/23/26	0/1/1/1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	H	500	NAG	O5-C1	2.14	1.47	1.43
5	D	500	NAG	O5-C1	2.38	1.47	1.43
5	J	500	NAG	C1-C2	2.61	1.56	1.52

The worst 5 of 23 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	B	500	NAG	O6-C6-C5	-3.54	99.63	111.33
5	F	500	NAG	C8-C7-N2	-3.45	109.50	116.11
5	D	500	NAG	O6-C6-C5	-3.45	99.95	111.33
5	L	500	NAG	C3-C4-C5	-3.10	104.79	110.20
5	L	500	NAG	O7-C7-C8	-3.00	116.55	122.06

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	J	500	NAG	1	0
5	L	500	NAG	0	1

## 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 [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	316/325 (97%)	0.16	6 (1%) 70 75	23, 33, 55, 76	1 (0%)
1	C	316/325 (97%)	0.22	9 (2%) 56 64	24, 34, 53, 74	1 (0%)
1	E	316/325 (97%)	0.22	9 (2%) 56 64	24, 33, 52, 81	1 (0%)
1	G	316/325 (97%)	0.26	14 (4%) 38 47	25, 37, 57, 81	1 (0%)
1	I	316/325 (97%)	0.23	10 (3%) 51 60	25, 35, 54, 79	1 (0%)
1	K	316/325 (97%)	0.20	5 (1%) 74 79	26, 36, 56, 81	1 (0%)
2	B	167/181 (92%)	0.28	3 (1%) 71 76	21, 35, 58, 102	0
2	D	167/181 (92%)	0.32	5 (2%) 54 62	22, 36, 54, 103	0
2	F	167/181 (92%)	0.32	3 (1%) 71 76	22, 36, 53, 90	0
2	H	169/181 (93%)	0.40	7 (4%) 41 49	25, 39, 60, 98	0
2	J	169/181 (93%)	0.43	7 (4%) 41 49	23, 38, 57, 100	0
2	L	169/181 (93%)	0.39	9 (5%) 30 38	24, 38, 56, 94	0
All	All	2904/3036 (95%)	0.26	87 (2%) 54 62	21, 36, 56, 103	6 (0%)

The worst 5 of 87 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	12	ASN	6.3
2	D	58	LYS	5.7
1	E	12	ASN	5.5
1	E	132	SER	4.1
2	F	58	LYS	3.9

### 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

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
7	NAG	E	401	14/15	0.53	0.54	2.64	69,82,89,90	0
6	BMA	C	403	11/12	0.42	0.30	-	84,95,97,100	0
7	NAG	E	402	14/15	0.63	0.46	-	70,91,94,98	0
3	NAG	A	401	14/15	0.63	0.28	-	50,67,76,77	0
3	NAG	G	402	14/15	0.75	0.40	-	62,85,91,92	0
6	BMA	I	403	11/12	0.46	0.39	-	87,99,102,103	0
6	NAG	I	402	14/15	0.75	0.27	-	58,71,88,96	0
6	NAG	K	401	14/15	0.81	0.18	-	44,49,52,58	0
6	NAG	K	402	14/15	0.78	0.33	-	56,70,84,97	0
6	BMA	K	403	11/12	0.35	0.36	-	92,100,109,112	0
6	NAG	I	401	14/15	0.86	0.17	-	49,60,64,65	0
7	BMA	E	403	11/12	0.18	0.62	-	103,109,112,115	0
7	FUC	E	404	10/11	0.60	0.57	-	76,86,91,92	0
6	NAG	E	405	14/15	0.69	0.29	-	45,57,62,71	0
6	NAG	E	406	14/15	0.67	0.29	-	63,76,85,90	0
6	NAG	C	402	14/15	0.76	0.35	-	69,79,84,90	0
6	NAG	C	401	14/15	0.72	0.23	-	41,66,73,73	0
6	BMA	E	407	11/12	0.27	0.40	-	76,87,96,99	0
3	NAG	G	401	14/15	0.81	0.23	-	46,61,68,74	0
3	NAG	A	402	14/15	0.80	0.29	-	61,76,84,89	0

### 6.4 Ligands

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
5	NAG	D	500	14/15	0.76	0.19	1.45	44,51,56,59	0
5	NAG	J	500	14/15	0.70	0.23	1.13	44,54,59,62	0
5	NAG	F	500	14/15	0.72	0.17	0.44	43,49,54,57	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
5	NAG	L	500	14/15	0.80	0.15	-0.22	45,50,55,55	0
5	NAG	B	500	14/15	0.82	0.14	-0.40	39,46,51,55	0
5	NAG	H	500	14/15	0.88	0.13	-1.03	41,46,53,56	0
4	CA	K	404	1/1	0.91	0.11	-	44,44,44,44	0
4	CA	C	404	1/1	0.87	0.07	-	46,46,46,46	0
4	CA	E	408	1/1	0.94	0.07	-	41,41,41,41	0
4	CA	I	404	1/1	0.95	0.04	-	43,43,43,43	0
4	CA	A	403	1/1	0.88	0.08	-	42,42,42,42	0
4	CA	G	403	1/1	0.92	0.06	-	35,35,35,35	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.