



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 9, 2020 – 01:26 AM BST

PDB ID : 6A0Z  
Title : Crystal structure of broadly neutralizing antibody 13D4 bound to H5N1 influenza hemagglutinin, HA head region  
Authors : Li, S.; Li, T.  
Deposited on : 2018-06-06  
Resolution : 2.33 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<https://www.wwpdb.org/validation/2017/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 : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.13.1  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.13.1

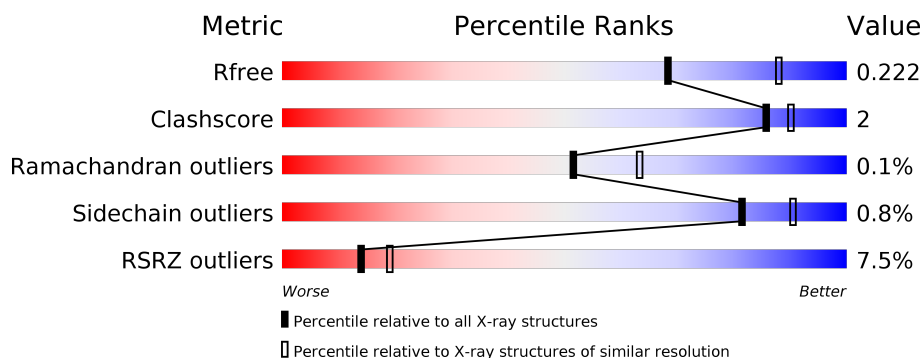
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.33 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))
$R_{free}$	130704	5974 (2.34-2.30)
Clashscore	141614	6604 (2.34-2.30)
Ramachandran outliers	138981	6523 (2.34-2.30)
Sidechain outliers	138945	6523 (2.34-2.30)
RSRZ outliers	127900	5855 (2.34-2.30)

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 respectively. 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	551	<div> <div>9%</div> <div>46%</div> <div>51%</div> </div>
2	H	224	<div> <div>2%</div> <div>95%</div> <div>5%</div> </div>
3	L	214	<div> <div>94%</div> <div>5%</div> </div>

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 5804 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,Envelope glycoprotein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	270	Total	C	N	O	S	0	0	0
			2152	1369	367	403	13			

There are 19 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	10	MET	-	expression tag	UNP Q45ZR9
A	508	SER	-	linker	UNP Q45ZR9
A	509	GLY	-	linker	UNP Q45ZR9
A	510	ARG	-	linker	UNP Q45ZR9
A	511	LEU	-	linker	UNP Q45ZR9
A	512	VAL	-	linker	UNP Q45ZR9
A	513	PRO	-	linker	UNP Q45ZR9
A	514	ARG	-	linker	UNP Q45ZR9
A	515	GLY	-	linker	UNP Q45ZR9
A	516	SER	-	linker	UNP Q45ZR9
A	517	PRO	-	linker	UNP Q45ZR9
A	518	GLY	-	linker	UNP Q45ZR9
A	519	SER	-	linker	UNP Q45ZR9
A	548	HIS	-	expression tag	UNP M1E1E4
A	549	HIS	-	expression tag	UNP M1E1E4
A	550	HIS	-	expression tag	UNP M1E1E4
A	551	HIS	-	expression tag	UNP M1E1E4
A	552	HIS	-	expression tag	UNP M1E1E4
A	553	HIS	-	expression tag	UNP M1E1E4

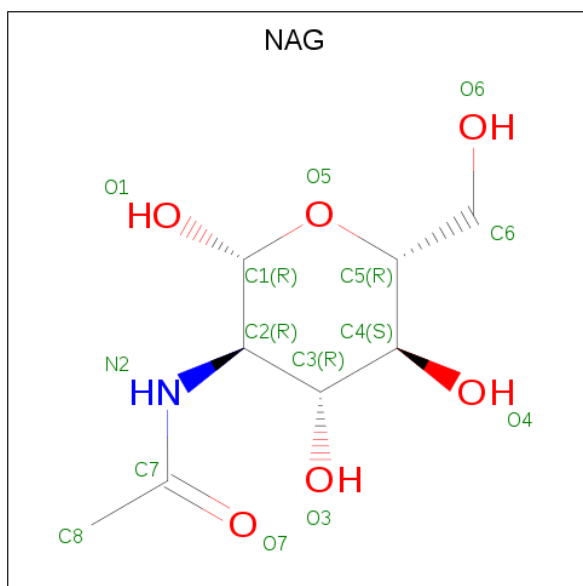
- Molecule 2 is a protein called Antibody 13D4, Fab Heavy Chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	H	224	Total	C	N	O	S	0	0	0
			1681	1065	276	333	7			

- Molecule 3 is a protein called Antibody 13D4, Fab Light Chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	L	212	Total	C	N	O	S	0	0	0
			1644	1023	281	333	7			

- Molecule 4 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



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

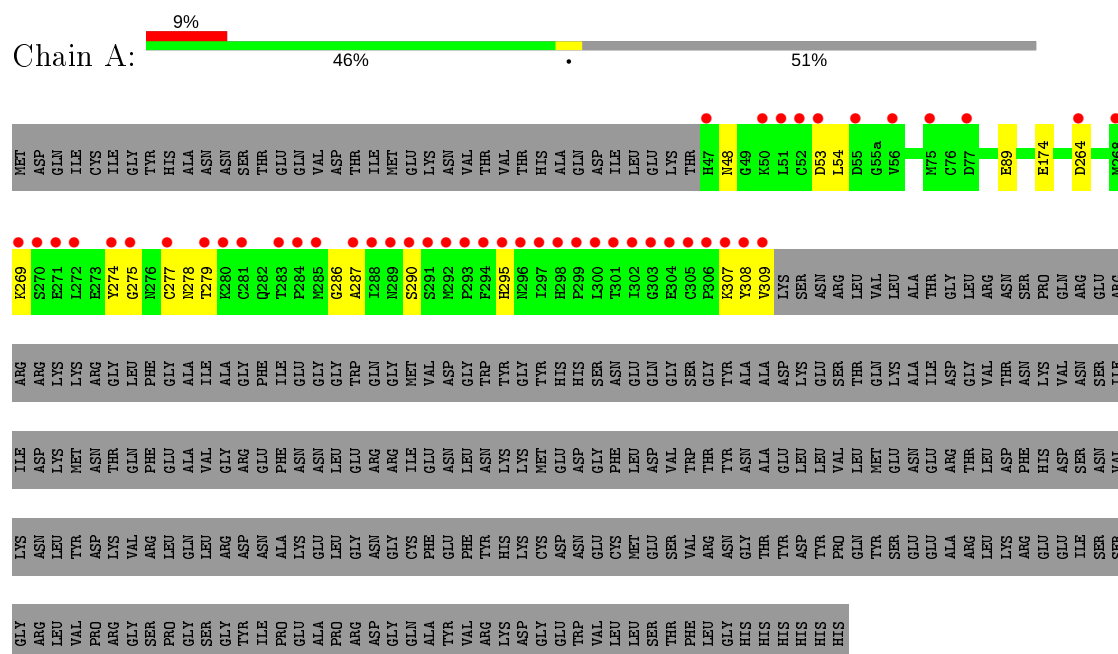
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	78	Total	O	0	0
			78	78		
5	H	115	Total	O	0	0
			115	115		
5	L	106	Total	O	0	0
			106	106		

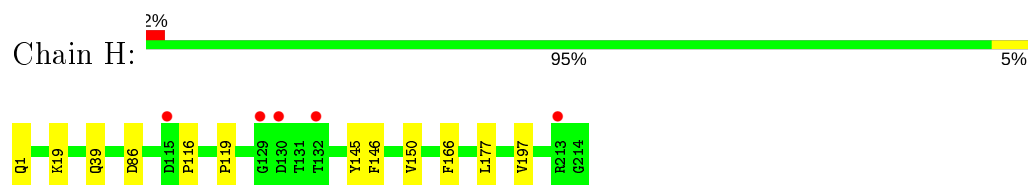
### 3 Residue-property plots

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes 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.

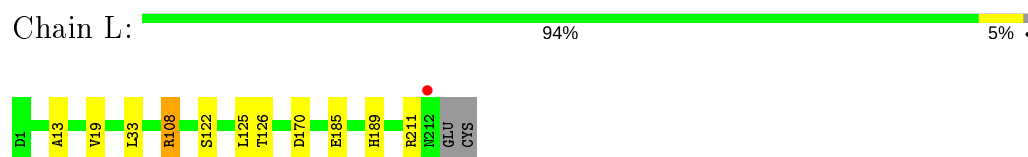
#### • Molecule 1: Hemagglutinin, Envelope glycoprotein



#### • Molecule 2: Antibody 13D4, Fab Heavy Chain



#### • Molecule 3: Antibody 13D4, Fab Light Chain



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	72.25Å 72.67Å 76.70Å 90.00° 93.37° 90.00°	Depositor
Resolution (Å)	34.67 – 2.33 34.67 – 2.33	Depositor EDS
% Data completeness (in resolution range)	99.7 (34.67-2.33) 99.7 (34.67-2.33)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.09	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.83 (at 2.34Å)	Xtriage
Refinement program	PHENIX 1.10.1 _2155	Depositor
R, $R_{free}$	0.172 , 0.218 0.177 , 0.222	Depositor DCC
$R_{free}$ test set	1591 reflections (4.68%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	35.4	Xtriage
Anisotropy	0.372	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 42.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	5804	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	48.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.49% 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, 0.333 respectively for untwinned datasets, and 0.375, 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: PCA, 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	0.25	0/2210	0.46	0/3003
2	H	0.25	0/1720	0.48	0/2354
3	L	0.26	0/1681	0.46	0/2281
All	All	0.25	0/5611	0.47	0/7638

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2152	0	2093	9	0
2	H	1681	0	1623	7	0
3	L	1644	0	1568	7	0
4	A	28	0	26	0	0
5	A	78	0	0	1	0
5	H	115	0	0	3	0
5	L	106	0	0	4	0
All	All	5804	0	5310	23	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 2.

All (23) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:H:19:LYS:NZ	5:H:302:HOH:O	2.17	0.77
3:L:122:SER:O	5:L:301:HOH:O	2.02	0.77
2:H:86:ASP:OD1	5:H:301:HOH:O	2.06	0.74
3:L:185:GLU:OE1	5:L:302:HOH:O	2.06	0.72
3:L:189:HIS:O	3:L:211:ARG:NH1	2.22	0.71
1:A:53:ASP:OD1	1:A:54:LEU:N	2.35	0.60
3:L:126:THR:N	5:L:301:HOH:O	2.20	0.59
1:A:48:ASN:HB2	1:A:286:GLY:HA3	1.88	0.55
1:A:264:ASP:O	5:A:701:HOH:O	2.19	0.53
2:H:39:GLN:OE1	5:H:303:HOH:O	2.19	0.52
3:L:125:LEU:N	5:L:301:HOH:O	2.45	0.50
3:L:13:ALA:HB2	3:L:19:VAL:HG11	1.95	0.48
2:H:166:PHE:O	2:H:177:LEU:CD2	2.64	0.45
1:A:279:THR:HG21	1:A:287:ALA:HB1	1.99	0.44
1:A:307:LYS:CG	1:A:308:TYR:H	2.31	0.44
1:A:278:ASN:OD1	1:A:279:THR:N	2.53	0.42
1:A:89:GLU:O	1:A:269:LYS:HA	2.20	0.42
1:A:174:GLU:N	1:A:174:GLU:OE1	2.53	0.41
2:H:116:PRO:HA	2:H:146:PHE:O	2.21	0.40
2:H:150:VAL:CG2	2:H:197:VAL:HG13	2.51	0.40
1:A:274:TYR:HA	1:A:275:GLY:HA3	1.87	0.40
3:L:108:ARG:HD2	3:L:170:ASP:O	2.20	0.40
2:H:119:PRO:HB3	2:H:145:TYR:HB3	2.03	0.40

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	268/551 (49%)	246 (92%)	21 (8%)	1 (0%)	34 41

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	H	222/224 (99%)	217 (98%)	5 (2%)	0	100	100
3	L	210/214 (98%)	204 (97%)	6 (3%)	0	100	100
All	All	700/989 (71%)	667 (95%)	32 (5%)	1 (0%)	51	63

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	290	SER

### 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	243/485 (50%)	240 (99%)	3 (1%)	71	83
2	H	186/186 (100%)	186 (100%)	0	100	100
3	L	186/188 (99%)	184 (99%)	2 (1%)	73	85
All	All	615/859 (72%)	610 (99%)	5 (1%)	81	90

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	277	CYS
1	A	295	HIS
1	A	309	VAL
3	L	33	LEU
3	L	108	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

1 non-standard protein/DNA/RNA residue is 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	$\# Z  > 2$
2	PCA	H	1	2	7,8,9	2.23	2 (28%)	9,10,12	2.27	5 (55%)

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
2	PCA	H	1	2	-	0/0/11/13	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	H	1	PCA	CD-N	4.70	1.47	1.34
2	H	1	PCA	CA-N	3.41	1.50	1.46

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	H	1	PCA	CB-CA-C	-3.53	107.85	112.70
2	H	1	PCA	OE-CD-CG	-3.10	121.36	126.76
2	H	1	PCA	CA-N-CD	-2.93	103.53	113.58
2	H	1	PCA	CG-CD-N	2.46	114.75	108.39
2	H	1	PCA	CB-CA-N	2.42	110.25	103.30

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.5 Carbohydrates

There are no monosaccharides in this entry.

## 5.6 Ligand geometry

2 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	$\# Z  > 2$
4	NAG	A	602	1	14,14,15	1.46	2 (14%)	17,19,21	1.29	2 (11%)
4	NAG	A	601	1	14,14,15	1.48	2 (14%)	17,19,21	1.26	2 (11%)

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	602	1	-	1/6/23/26	0/1/1/1
4	NAG	A	601	1	-	2/6/23/26	0/1/1/1

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	601	NAG	O5-C1	3.65	1.49	1.43
4	A	602	NAG	O5-C1	3.65	1.49	1.43
4	A	602	NAG	C7-N2	2.61	1.43	1.34
4	A	601	NAG	C7-N2	2.57	1.43	1.34

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	601	NAG	C1-O5-C5	2.87	116.08	112.19
4	A	602	NAG	C1-O5-C5	2.77	115.94	112.19

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	601	NAG	C8-C7-N2	2.23	119.87	116.10
4	A	602	NAG	C8-C7-N2	2.17	119.78	116.10

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	601	NAG	C4-C5-C6-O6
4	A	601	NAG	O5-C5-C6-O6
4	A	602	NAG	O5-C5-C6-O6

There are no ring outliers.

No monomer is involved in short contacts.

## 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<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	270/551 (49%)	0.63	47 (17%) 1 1	22, 48, 136, 173	0
2	H	223/224 (99%)	-0.12	5 (2%) 62 69	23, 37, 67, 161	0
3	L	212/214 (99%)	-0.23	1 (0%) 91 94	22, 38, 66, 98	0
All	All	705/989 (71%)	0.14	53 (7%) 14 19	22, 40, 112, 173	0

All (53) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	294	PHE	12.2
1	A	305	CYS	10.3
1	A	297	ILE	9.5
1	A	304	GLU	7.8
1	A	306	PRO	6.8
1	A	289	ASN	6.8
2	H	130	ASP	6.3
1	A	290	SER	6.2
1	A	285	MET	5.9
1	A	308	TYR	5.9
1	A	295	HIS	5.8
1	A	292	MET	5.8
1	A	47	HIS	5.7
1	A	296	ASN	5.3
1	A	291	SER	5.3
2	H	129	GLY	5.3
1	A	288	ILE	4.8
2	H	213	ARG	4.5
1	A	277	CYS	4.4
1	A	281	CYS	4.3
1	A	309	VAL	4.0
1	A	300	LEU	3.9
1	A	293	PRO	3.7

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	302	ILE	3.7
1	A	56	VAL	3.6
2	H	132	THR	3.6
1	A	272	LEU	3.5
1	A	307	LYS	3.5
1	A	284	PRO	3.4
3	L	212	ASN	3.4
1	A	287	ALA	3.3
1	A	77	ASP	2.8
1	A	274	TYR	2.7
1	A	268	MET	2.6
1	A	269	LYS	2.6
1	A	283	THR	2.6
1	A	303	GLY	2.5
1	A	280	LYS	2.5
1	A	50	LYS	2.4
1	A	298	HIS	2.4
1	A	299	PRO	2.4
1	A	270	SER	2.4
1	A	55	ASP	2.3
1	A	275	GLY	2.3
1	A	301	THR	2.3
1	A	279	THR	2.2
1	A	271	GLU	2.2
2	H	115	ASP	2.2
1	A	51	LEU	2.1
1	A	53	ASP	2.1
1	A	52	CYS	2.1
1	A	75	MET	2.1
1	A	264	ASP	2.0

## 6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

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. 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	B-factors(Å <sup>2</sup> )	Q<0.9
2	PCA	H	1	8/9	0.96	0.12	42,45,50,52	0

### 6.3 Carbohydrates [i](#)

There are no monosaccharides 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. 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	B-factors( $\text{\AA}^2$ )	Q<0.9
4	NAG	A	601	14/15	0.91	0.15	48,59,69,70	0
4	NAG	A	602	14/15	0.92	0.30	58,67,86,91	0

### 6.5 Other polymers [i](#)

There are no such residues in this entry.