



# Full wwPDB X-ray Structure Validation Report ⓘ

Sep 17, 2017 – 09:15 AM EDT

PDB ID : 4WE7  
Title : Structure and receptor binding preferences of recombinant human A(H3N2) virus hemagglutinins  
Authors : Yang, H.; Carney, P.J.; Chang, J.C.; Guo, Z.; Villanueva, J.M.; Stevens, J.  
Deposited on : unknown  
Resolution : 2.50 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

<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 : 1.7.2 (RC1), CSD as538be (2017)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20029824  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
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-20029824

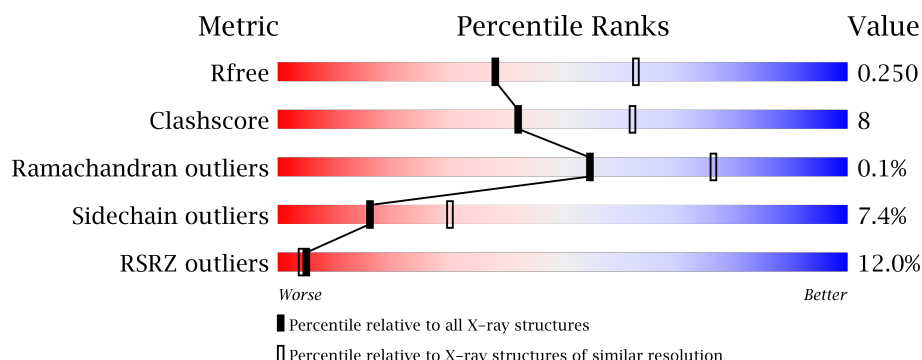
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.50 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))
$R_{free}$	100719	3846 (2.50-2.50)
Clashscore	112137	4554 (2.50-2.50)
Ramachandran outliers	110173	4463 (2.50-2.50)
Sidechain outliers	110143	4465 (2.50-2.50)
RSRZ outliers	101464	3876 (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  $\geq 3$ , 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	284	<div> <div>18%</div> <div> <div></div> <div>73%</div> <div>19%</div> <div>• •</div> </div> </div>
1	B	284	<div> <div>5%</div> <div> <div></div> <div>81%</div> <div>13%</div> <div>• •</div> </div> </div>
1	C	284	<div> <div>17%</div> <div> <div></div> <div>75%</div> <div>19%</div> <div>• •</div> </div> </div>
1	D	284	<div> <div>6%</div> <div> <div></div> <div>82%</div> <div>13%</div> <div>• •</div> </div> </div>

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 crite-

ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	NAG	A	304	X	-	-	-
2	NAG	C	301	X	-	-	-
2	NAG	D	302	X	-	-	-

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 8827 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 chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	272	Total	C	N	O	S	0	0	0
			2128	1333	383	402	10			
1	B	272	Total	C	N	O	S	0	0	0
			2128	1333	383	402	10			
1	C	272	Total	C	N	O	S	0	0	0
			2128	1333	383	402	10			
1	D	272	Total	C	N	O	S	0	0	0
			2128	1333	383	402	10			

There are 68 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-4	ALA	-	expression tag	UNP U3N7S5
A	-3	ASP	-	expression tag	UNP U3N7S5
A	-2	LEU	-	expression tag	UNP U3N7S5
A	-1	GLY	-	expression tag	UNP U3N7S5
A	0	SER	-	expression tag	UNP U3N7S5
A	177	VAL	ILE	conflict	UNP U3N7S5
A	269	SER	-	expression tag	UNP U3N7S5
A	270	LYS	-	expression tag	UNP U3N7S5
A	271	GLN	-	expression tag	UNP U3N7S5
A	272	ASN	-	expression tag	UNP U3N7S5
A	273	SER	-	expression tag	UNP U3N7S5
A	274	GLY	-	expression tag	UNP U3N7S5
A	275	ARG	-	expression tag	UNP U3N7S5
A	276	LEU	-	expression tag	UNP U3N7S5
A	277	VAL	-	expression tag	UNP U3N7S5
A	278	PRO	-	expression tag	UNP U3N7S5
A	279	ARG	-	expression tag	UNP U3N7S5
B	-4	ALA	-	expression tag	UNP U3N7S5
B	-3	ASP	-	expression tag	UNP U3N7S5
B	-2	LEU	-	expression tag	UNP U3N7S5
B	-1	GLY	-	expression tag	UNP U3N7S5

*Continued on next page...*

*Continued from previous page...*

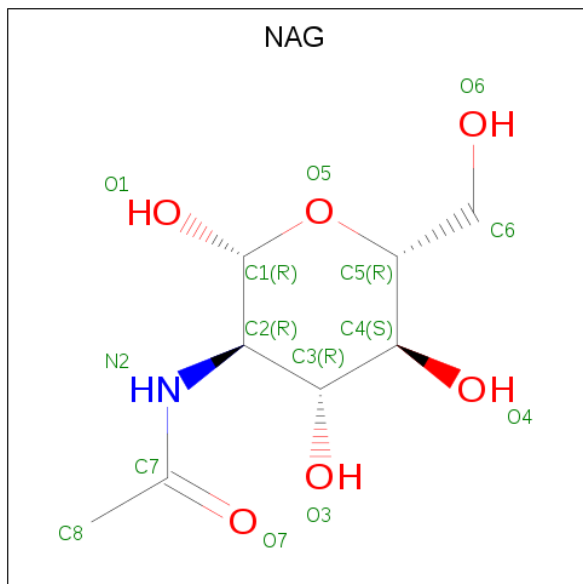
Chain	Residue	Modelled	Actual	Comment	Reference
B	0	SER	-	expression tag	UNP U3N7S5
B	177	VAL	ILE	conflict	UNP U3N7S5
B	269	SER	-	expression tag	UNP U3N7S5
B	270	LYS	-	expression tag	UNP U3N7S5
B	271	GLN	-	expression tag	UNP U3N7S5
B	272	ASN	-	expression tag	UNP U3N7S5
B	273	SER	-	expression tag	UNP U3N7S5
B	274	GLY	-	expression tag	UNP U3N7S5
B	275	ARG	-	expression tag	UNP U3N7S5
B	276	LEU	-	expression tag	UNP U3N7S5
B	277	VAL	-	expression tag	UNP U3N7S5
B	278	PRO	-	expression tag	UNP U3N7S5
B	279	ARG	-	expression tag	UNP U3N7S5
C	-4	ALA	-	expression tag	UNP U3N7S5
C	-3	ASP	-	expression tag	UNP U3N7S5
C	-2	LEU	-	expression tag	UNP U3N7S5
C	-1	GLY	-	expression tag	UNP U3N7S5
C	0	SER	-	expression tag	UNP U3N7S5
C	177	VAL	ILE	conflict	UNP U3N7S5
C	269	SER	-	expression tag	UNP U3N7S5
C	270	LYS	-	expression tag	UNP U3N7S5
C	271	GLN	-	expression tag	UNP U3N7S5
C	272	ASN	-	expression tag	UNP U3N7S5
C	273	SER	-	expression tag	UNP U3N7S5
C	274	GLY	-	expression tag	UNP U3N7S5
C	275	ARG	-	expression tag	UNP U3N7S5
C	276	LEU	-	expression tag	UNP U3N7S5
C	277	VAL	-	expression tag	UNP U3N7S5
C	278	PRO	-	expression tag	UNP U3N7S5
C	279	ARG	-	expression tag	UNP U3N7S5
D	-4	ALA	-	expression tag	UNP U3N7S5
D	-3	ASP	-	expression tag	UNP U3N7S5
D	-2	LEU	-	expression tag	UNP U3N7S5
D	-1	GLY	-	expression tag	UNP U3N7S5
D	0	SER	-	expression tag	UNP U3N7S5
D	177	VAL	ILE	conflict	UNP U3N7S5
D	269	SER	-	expression tag	UNP U3N7S5
D	270	LYS	-	expression tag	UNP U3N7S5
D	271	GLN	-	expression tag	UNP U3N7S5
D	272	ASN	-	expression tag	UNP U3N7S5
D	273	SER	-	expression tag	UNP U3N7S5
D	274	GLY	-	expression tag	UNP U3N7S5

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	275	ARG	-	expression tag	UNP U3N7S5
D	276	LEU	-	expression tag	UNP U3N7S5
D	277	VAL	-	expression tag	UNP U3N7S5
D	278	PRO	-	expression tag	UNP U3N7S5
D	279	ARG	-	expression tag	UNP U3N7S5

- Molecule 2 is N-ACETYL-D-GLUCOSAMINE (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



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

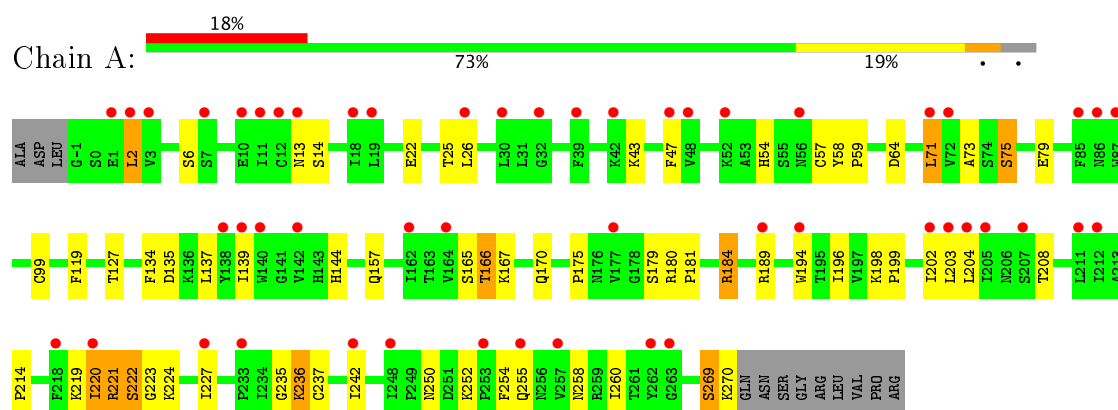
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	19	Total	O	0	0
			19	19		
3	B	38	Total	O	0	0
			38	38		
3	C	17	Total	O	0	0
			17	17		
3	D	45	Total	O	0	0
			45	45		

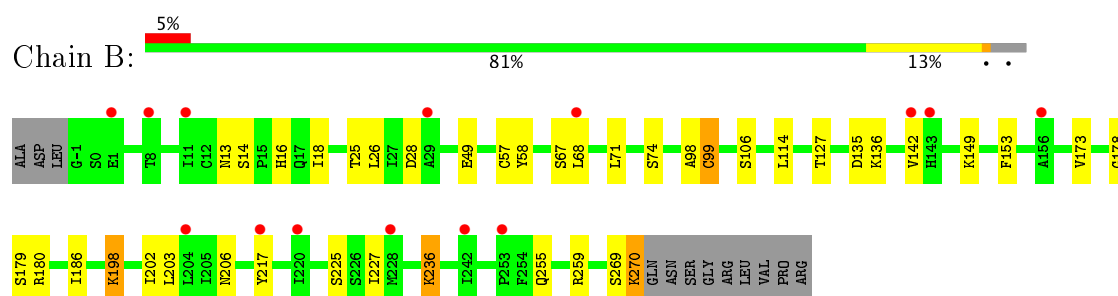
### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA 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.

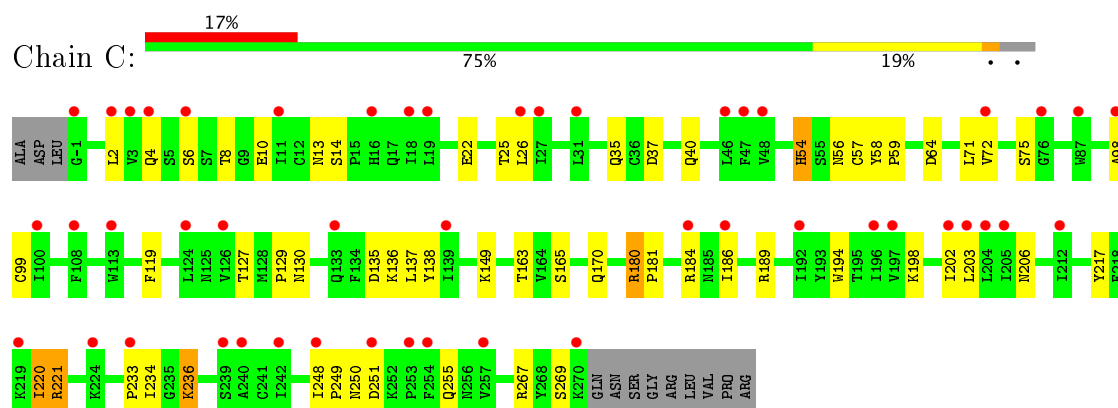
- Molecule 1: Hemagglutinin HA1 chain



- Molecule 1: Hemagglutinin HA1 chain

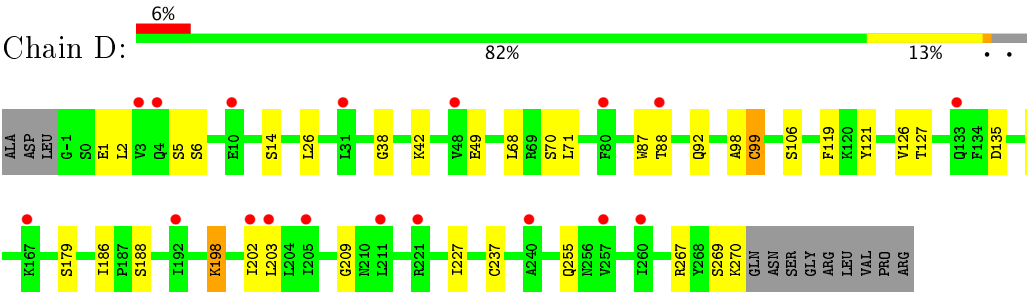


- Molecule 1: Hemagglutinin HA1 chain



- Molecule 1: Hemagglutinin HA1 chain





GLOBAL-STATISTICS INFOmissingINFO

## 4 Model quality [i](#)

### 4.1 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	0.41	0/2180	0.59	0/2962
1	B	0.46	0/2180	0.63	0/2962
1	C	0.40	0/2180	0.58	0/2962
1	D	0.46	0/2180	0.63	0/2962
All	All	0.43	0/8720	0.61	0/11848

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 4.2 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	2128	0	2069	56	2
1	B	2128	0	2071	18	0
1	C	2128	0	2068	46	2
1	D	2128	0	2071	16	0
2	A	56	0	52	1	0
2	B	42	0	39	0	0
2	C	56	0	52	2	0
2	D	42	0	39	2	0
3	A	19	0	0	3	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	38	0	0	3	0
3	C	17	0	0	2	0
3	D	45	0	0	3	0
All	All	8827	0	8461	135	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 8.

All (135) 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:A:75:SER:HA	1:A:221:ARG:CG	1.72	1.18
1:A:75:SER:CA	1:A:221:ARG:CG	2.24	1.14
1:A:220:ILE:N	1:A:220:ILE:HD12	1.56	1.13
1:A:75:SER:HA	1:A:221:ARG:HG3	1.16	1.12
1:C:220:ILE:N	1:C:220:ILE:HD12	1.54	1.12
1:C:220:ILE:CD1	1:C:220:ILE:H	1.63	1.12
1:C:221:ARG:HH21	1:C:221:ARG:HG2	0.96	1.08
1:A:220:ILE:H	1:A:220:ILE:HD12	0.92	1.08
1:A:220:ILE:CD1	1:A:220:ILE:H	1.67	1.07
1:A:75:SER:CA	1:A:221:ARG:HG2	1.85	1.03
1:A:75:SER:CB	1:A:221:ARG:HG2	1.93	0.97
1:C:220:ILE:HD12	1:C:220:ILE:H	0.78	0.94
1:C:221:ARG:CG	1:C:221:ARG:HH21	1.78	0.94
1:A:75:SER:CA	1:A:221:ARG:HG3	1.89	0.94
1:C:4:GLN:HE22	1:C:249:PRO:HD2	1.33	0.91
1:A:26:LEU:HD22	1:A:227:ILE:HD11	1.53	0.88
1:C:54:HIS:HE1	2:C:301:NAG:H61	1.38	0.86
2:D:303:NAG:O7	3:D:438:HOH:O	1.95	0.84
1:C:221:ARG:NH2	1:C:221:ARG:HG2	1.64	0.84
1:C:4:GLN:NE2	1:C:249:PRO:HD2	1.93	0.82
1:A:221:ARG:O	1:A:222:SER:CB	2.30	0.80
1:D:135:ASP:OD2	3:D:401:HOH:O	2.00	0.79
1:A:184:ARG:HH11	1:A:184:ARG:HG3	1.49	0.76
1:A:260:ILE:O	3:A:418:HOH:O	2.04	0.76
1:C:75:SER:O	1:C:221:ARG:CD	2.37	0.73
1:A:135:ASP:O	1:A:220:ILE:HD11	1.90	0.71
2:D:303:NAG:H2	3:D:438:HOH:O	1.89	0.71
1:A:13:ASN:HD21	1:A:236:LYS:HA	1.56	0.71
1:A:180:ARG:HG3	1:A:181:PRO:HD2	1.71	0.71
1:C:75:SER:O	1:C:221:ARG:NE	2.24	0.70

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:221:ARG:O	1:A:222:SER:HB2	1.92	0.69
1:C:75:SER:O	1:C:221:ARG:HD2	1.93	0.69
1:B:127:THR:HB	1:B:202:ILE:HD11	1.75	0.68
1:C:54:HIS:CE1	2:C:301:NAG:H61	2.27	0.67
1:C:4:GLN:NE2	1:C:248:ILE:HB	2.10	0.67
1:A:73:ALA:HB2	1:A:227:ILE:HD12	1.77	0.66
1:A:242:ILE:O	3:A:416:HOH:O	2.13	0.65
1:B:16:HIS:HB2	3:B:431:HOH:O	1.97	0.65
1:B:114:LEU:O	3:B:406:HOH:O	2.15	0.64
1:A:13:ASN:HD21	1:A:236:LYS:HG2	1.64	0.63
1:C:4:GLN:HE22	1:C:248:ILE:HB	1.63	0.62
1:A:166:THR:OG1	1:A:167:LYS:N	2.33	0.61
1:A:47:PHE:HB3	1:A:227:ILE:HG12	1.83	0.60
1:A:184:ARG:HG3	1:A:184:ARG:NH1	2.17	0.60
1:C:137:LEU:HB2	1:C:220:ILE:HG13	1.83	0.60
1:A:254:PHE:HD1	1:A:269:SER:HB2	1.67	0.58
1:B:178:GLY:O	1:B:180:ARG:NH1	2.36	0.57
1:D:1:GLU:OE1	1:D:1:GLU:N	2.26	0.56
1:A:54:HIS:HE1	2:A:301:NAG:O5	1.89	0.56
1:C:251:ASP:OD1	3:C:401:HOH:O	2.18	0.56
1:A:222:SER:H	1:A:223:GLY:HA3	1.72	0.55
1:D:49:GLU:HG3	1:D:227:ILE:HD11	1.88	0.55
1:D:87:TRP:HB2	1:D:92:GLN:NE2	2.21	0.55
1:C:10:GLU:HG2	1:C:233:PRO:HG2	1.87	0.55
1:A:13:ASN:ND2	1:A:236:LYS:HA	2.20	0.54
1:A:135:ASP:O	1:A:220:ILE:CD1	2.56	0.54
1:D:127:THR:HB	1:D:202:ILE:HD11	1.90	0.54
1:B:25:THR:HG23	1:B:28:ASP:H	1.74	0.53
1:C:130:ASN:HB2	1:C:136:LYS:HE2	1.91	0.52
1:A:26:LEU:HD22	1:A:227:ILE:CD1	2.32	0.52
1:B:49:GLU:HG3	1:B:227:ILE:HD11	1.93	0.51
1:C:4:GLN:HE22	1:C:249:PRO:CD	2.15	0.51
1:A:222:SER:N	1:A:223:GLY:HA3	2.25	0.51
1:A:75:SER:HB2	1:A:221:ARG:HG2	1.88	0.50
1:B:153:PHE:O	3:B:409:HOH:O	2.20	0.50
1:A:64:ASP:HB3	1:A:194:TRP:CH2	2.46	0.50
1:A:13:ASN:ND2	1:A:235:GLY:O	2.45	0.50
1:B:13:ASN:HA	1:B:18:ILE:HD12	1.94	0.50
1:A:258:ASN:HB3	3:A:418:HOH:O	2.11	0.50
1:C:181:PRO:O	1:C:189:ARG:NH2	2.29	0.50
1:A:144:HIS:CE1	1:A:175:PRO:HA	2.47	0.49

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:64:ASP:HB3	1:C:194:TRP:CZ3	2.47	0.49
1:A:71:LEU:HD11	1:A:196:ILE:HD11	1.93	0.49
1:C:13:ASN:OD1	1:C:236:LYS:HD2	2.12	0.49
1:A:137:LEU:HB2	1:A:220:ILE:HG13	1.94	0.49
1:A:75:SER:HB2	1:A:221:ARG:H	1.76	0.49
1:A:47:PHE:HB3	1:A:227:ILE:CG1	2.43	0.49
1:B:135:ASP:OD1	1:B:198:LYS:HE2	2.13	0.49
1:D:87:TRP:HB2	1:D:92:GLN:HE21	1.77	0.49
1:A:221:ARG:O	1:A:222:SER:OG	2.30	0.48
1:C:64:ASP:HB3	1:C:194:TRP:CH2	2.48	0.48
1:A:165:SER:HB3	1:A:170:GLN:HG3	1.96	0.48
1:A:135:ASP:C	1:A:220:ILE:HD11	2.33	0.48
1:C:13:ASN:OD1	1:C:236:LYS:HA	2.14	0.48
1:C:8:THR:HG23	3:C:411:HOH:O	2.13	0.48
1:C:119:PHE:CE2	1:D:149:LYS:HD2	2.48	0.48
1:C:35:GLN:OE1	1:C:54:HIS:NE2	2.46	0.47
1:D:38:GLY:O	1:D:42:LYS:HE3	2.14	0.47
1:C:163:THR:HB	1:C:206:ASN:HB2	1.97	0.47
1:A:139:ILE:O	1:A:214:PRO:HB3	2.15	0.47
1:A:57:CYS:SG	1:A:58:TYR:N	2.83	0.47
1:C:59:PRO:HB2	1:C:189:ARG:HD3	1.97	0.46
1:A:59:PRO:HB2	1:A:189:ARG:HD3	1.97	0.46
1:B:98:ALA:HB2	1:B:186:ILE:HG13	1.96	0.46
1:A:75:SER:OG	1:A:221:ARG:HG2	2.15	0.45
1:A:254:PHE:CD1	1:A:269:SER:HB2	2.49	0.45
1:A:64:ASP:HB3	1:A:194:TRP:CZ3	2.51	0.45
1:B:269:SER:OG	1:B:270:LYS:N	2.50	0.45
1:D:98:ALA:HB2	1:D:186:ILE:HG13	1.99	0.44
1:C:165:SER:HB3	1:C:170:GLN:HG3	2.00	0.44
1:C:129:PRO:HA	1:C:202:ILE:HG13	1.99	0.44
1:C:136:LYS:HE3	1:C:138:TYR:HE2	1.83	0.44
1:C:136:LYS:HD2	1:C:217:TYR:CD2	2.53	0.44
1:C:35:GLN:NE2	1:C:56:ASN:HB2	2.32	0.43
1:C:135:ASP:C	1:C:220:ILE:HD11	2.39	0.43
1:B:142:VAL:HG21	1:B:173:VAL:HG11	2.00	0.43
1:A:119:PHE:CE2	1:B:149:LYS:HD2	2.53	0.43
1:C:57:CYS:SG	1:C:58:TYR:N	2.88	0.43
1:C:37:ASP:O	1:C:40:GLN:HG3	2.18	0.43
1:D:99:CYS:O	1:D:106:SER:N	2.49	0.43
1:A:134:PHE:CZ	1:A:219:LYS:HE3	2.53	0.43
1:C:75:SER:CA	1:C:221:ARG:HD2	2.49	0.43

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:57:CYS:SG	1:B:58:TYR:N	2.90	0.43
1:C:136:LYS:HE3	1:C:138:TYR:CE2	2.54	0.43
1:C:180:ARG:HB3	1:C:181:PRO:HD2	2.01	0.42
1:A:2:LEU:O	1:A:252:LYS:HB3	2.19	0.42
1:C:220:ILE:CD1	1:C:220:ILE:N	2.30	0.42
1:D:142:VAL:HG22	1:D:162:ILE:HD12	2.00	0.42
1:A:181:PRO:O	1:A:189:ARG:NH2	2.33	0.42
1:D:121:TYR:CZ	1:D:209:GLY:HA2	2.54	0.42
1:D:135:ASP:OD1	1:D:198:LYS:HE2	2.20	0.42
1:A:157:GLN:NE2	1:A:208:THR:O	2.49	0.42
1:A:127:THR:HB	1:A:202:ILE:HD11	2.01	0.42
1:C:98:ALA:HB2	1:C:186:ILE:HG13	2.02	0.42
1:B:236:LYS:HD2	1:B:236:LYS:N	2.34	0.41
1:B:136:LYS:HD2	1:B:217:TYR:CD2	2.55	0.41
1:A:135:ASP:OD1	1:A:199:PRO:HD3	2.21	0.41
1:D:87:TRP:CZ3	1:D:126:VAL:HG11	2.56	0.41
1:C:149:LYS:HD2	1:D:119:PHE:CE2	2.56	0.41
1:D:270:LYS:HG3	1:D:270:LYS:HZ3	1.70	0.41
1:B:74:SER:HA	1:B:225:SER:O	2.20	0.41
1:B:99:CYS:O	1:B:106:SER:N	2.53	0.41
1:C:127:THR:HB	1:C:202:ILE:HD11	2.03	0.41
1:A:43:LYS:HZ1	1:A:79:GLU:CD	2.25	0.40
1:C:26:LEU:HD21	1:C:72:VAL:HG12	2.03	0.40

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:A:224:LYS:O	1:C:267:ARG:NH2[2_546]	2.05	0.15
1:A:250:ASN:O	1:C:250:ASN:ND2[2_546]	2.18	0.02

## 4.3 Torsion angles [i](#)

### 4.3.1 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	270/284 (95%)	255 (94%)	14 (5%)	1 (0%)	38	59
1	B	270/284 (95%)	258 (96%)	12 (4%)	0	100	100
1	C	270/284 (95%)	259 (96%)	11 (4%)	0	100	100
1	D	270/284 (95%)	256 (95%)	14 (5%)	0	100	100
All	All	1080/1136 (95%)	1028 (95%)	51 (5%)	1 (0%)	55	76

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	222	SER

#### 4.3.2 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	239/249 (96%)	218 (91%)	21 (9%)	12	22
1	B	239/249 (96%)	225 (94%)	14 (6%)	23	42
1	C	239/249 (96%)	221 (92%)	18 (8%)	16	29
1	D	239/249 (96%)	221 (92%)	18 (8%)	16	29
All	All	956/996 (96%)	885 (93%)	71 (7%)	16	30

All (71) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	2	LEU
1	A	6	SER
1	A	14	SER
1	A	22	GLU
1	A	25	THR
1	A	71	LEU
1	A	75	SER
1	A	99	CYS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	166	THR
1	A	179	SER
1	A	184	ARG
1	A	198	LYS
1	A	203	LEU
1	A	204	LEU
1	A	220	ILE
1	A	221	ARG
1	A	236	LYS
1	A	237	CYS
1	A	255	GLN
1	A	269	SER
1	A	270	LYS
1	B	14	SER
1	B	26	LEU
1	B	67	SER
1	B	68	LEU
1	B	71	LEU
1	B	99	CYS
1	B	179	SER
1	B	198	LYS
1	B	203	LEU
1	B	206	ASN
1	B	236	LYS
1	B	255	GLN
1	B	259	ARG
1	B	270	LYS
1	C	2	LEU
1	C	6	SER
1	C	14	SER
1	C	22	GLU
1	C	25	THR
1	C	54	HIS
1	C	71	LEU
1	C	99	CYS
1	C	180	ARG
1	C	184	ARG
1	C	198	LYS
1	C	203	LEU
1	C	220	ILE
1	C	221	ARG
1	C	234	ILE

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
1	C	236	LYS
1	C	255	GLN
1	C	269	SER
1	D	2	LEU
1	D	5	SER
1	D	6	SER
1	D	14	SER
1	D	26	LEU
1	D	68	LEU
1	D	70	SER
1	D	71	LEU
1	D	88	THR
1	D	99	CYS
1	D	179	SER
1	D	188	SER
1	D	198	LYS
1	D	203	LEU
1	D	237	CYS
1	D	255	GLN
1	D	267	ARG
1	D	269	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (6) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	13	ASN
1	A	54	HIS
1	B	4	GLN
1	B	92	GLN
1	C	4	GLN
1	D	92	GLN

#### 4.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

## 4.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 4.6 Ligand geometry [i](#)

14 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$
2	NAG	A	301	1	14,14,15	0.43	0	15,19,21	1.15	2 (13%)
2	NAG	A	302	1	14,14,15	0.43	0	15,19,21	1.42	2 (13%)
2	NAG	A	303	1	14,14,15	0.47	0	15,19,21	0.80	1 (6%)
2	NAG	A	304	1	14,14,15	0.38	0	15,19,21	2.18	2 (13%)
2	NAG	B	301	1	14,14,15	0.34	0	15,19,21	1.51	2 (13%)
2	NAG	B	302	1	14,14,15	0.51	0	15,19,21	1.21	1 (6%)
2	NAG	B	303	1	14,14,15	0.50	0	15,19,21	1.34	2 (13%)
2	NAG	C	301	1	14,14,15	0.73	0	15,19,21	1.82	2 (13%)
2	NAG	C	302	1	14,14,15	0.41	0	15,19,21	0.93	1 (6%)
2	NAG	C	303	1	14,14,15	0.47	0	15,19,21	0.81	1 (6%)
2	NAG	C	304	1	14,14,15	1.13	2 (14%)	15,19,21	2.17	4 (26%)
2	NAG	D	301	1	14,14,15	0.29	0	15,19,21	1.73	3 (20%)
2	NAG	D	302	1	14,14,15	0.57	0	15,19,21	1.52	3 (20%)
2	NAG	D	303	1	14,14,15	0.46	0	15,19,21	1.09	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
2	NAG	A	301	1	-	0/6/23/26	0/1/1/1
2	NAG	A	302	1	-	0/6/23/26	0/1/1/1
2	NAG	A	303	1	-	0/6/23/26	0/1/1/1

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	NAG	A	304	1	1/1/5/7	0/6/23/26	0/1/1/1
2	NAG	B	301	1	-	0/6/23/26	0/1/1/1
2	NAG	B	302	1	-	0/6/23/26	0/1/1/1
2	NAG	B	303	1	-	0/6/23/26	0/1/1/1
2	NAG	C	301	1	1/1/5/7	0/6/23/26	0/1/1/1
2	NAG	C	302	1	-	0/6/23/26	0/1/1/1
2	NAG	C	303	1	-	0/6/23/26	0/1/1/1
2	NAG	C	304	1	-	0/6/23/26	0/1/1/1
2	NAG	D	301	1	-	0/6/23/26	0/1/1/1
2	NAG	D	302	1	1/1/5/7	0/6/23/26	0/1/1/1
2	NAG	D	303	1	-	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)
2	C	304	NAG	O5-C5	-2.31	1.38	1.43
2	C	304	NAG	O5-C1	-2.14	1.40	1.43

All (26) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	301	NAG	O5-C1-C2	-5.30	104.09	111.47
2	C	304	NAG	C2-N2-C7	-4.27	116.72	122.94
2	C	304	NAG	O7-C7-C8	-3.55	115.59	122.06
2	D	301	NAG	C6-C5-C4	-3.19	105.54	113.00
2	C	301	NAG	C6-C5-C4	-3.09	105.77	113.00
2	B	303	NAG	O5-C1-C2	-3.08	107.19	111.47
2	D	302	NAG	O5-C1-C2	-2.85	107.50	111.47
2	A	301	NAG	O5-C1-C2	-2.61	107.85	111.47
2	B	301	NAG	C6-C5-C4	-2.35	107.51	113.00
2	C	303	NAG	C1-O5-C5	2.03	114.96	112.17
2	D	301	NAG	C3-C4-C5	2.08	113.88	110.22
2	B	303	NAG	C1-O5-C5	2.24	115.26	112.17
2	A	303	NAG	C1-O5-C5	2.25	115.27	112.17
2	A	302	NAG	O5-C1-C2	2.29	114.67	111.47
2	C	302	NAG	C1-O5-C5	2.41	115.49	112.17
2	D	302	NAG	C3-C4-C5	2.43	114.50	110.22
2	C	304	NAG	C1-C2-N2	2.70	115.10	110.49
2	A	301	NAG	C1-O5-C5	2.91	116.17	112.17
2	D	302	NAG	C4-C3-C2	3.51	116.16	111.02
2	B	302	NAG	C1-O5-C5	3.79	117.39	112.17
2	A	302	NAG	C1-O5-C5	3.85	117.47	112.17

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	304	NAG	C1-O5-C5	3.85	117.47	112.17
2	B	301	NAG	C1-O5-C5	3.96	117.62	112.17
2	D	301	NAG	C1-O5-C5	4.58	118.48	112.17
2	A	304	NAG	O5-C1-C2	4.93	118.33	111.47
2	A	304	NAG	C1-O5-C5	6.32	120.88	112.17

All (3) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	A	304	NAG	C1
2	D	302	NAG	C1
2	C	301	NAG	C1

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	301	NAG	1	0
2	C	301	NAG	2	0
2	D	303	NAG	2	0

## 4.7 Other polymers [i](#)

There are no such residues in this entry.

## 4.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 5 Fit of model and data ⓘ

### 5.1 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	272/284 (95%)	1.02	51 (18%) 1 1	24, 62, 92, 109	0
1	B	272/284 (95%)	0.64	14 (5%) 29 30	13, 35, 67, 100	0
1	C	272/284 (95%)	1.06	48 (17%) 2 1	25, 62, 95, 114	0
1	D	272/284 (95%)	0.71	18 (6%) 19 19	14, 34, 66, 103	0
All	All	1088/1136 (95%)	0.86	131 (12%) 5 4	13, 47, 88, 114	0

All (131) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	2	LEU	7.5
1	C	254	PHE	6.7
1	C	6	SER	6.4
1	C	202	ILE	5.4
1	A	203	LEU	5.4
1	C	203	LEU	5.1
1	C	-1	GLY	5.0
1	C	72	VAL	4.8
1	A	18	ILE	4.8
1	A	11	ILE	4.7
1	C	48	VAL	4.7
1	A	3	VAL	4.7
1	C	205	ILE	4.6
1	C	98	ALA	4.3
1	C	11	ILE	4.3
1	C	46	LEU	4.2
1	C	240	ALA	4.2
1	A	10	GLU	4.1
1	C	239	SER	3.8
1	A	227	ILE	3.7
1	A	52	LYS	3.6

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	184	ARG	3.6
1	C	251	ASP	3.5
1	C	47	PHE	3.4
1	C	186	ILE	3.4
1	A	162	ILE	3.4
1	A	87	TRP	3.4
1	B	204	LEU	3.3
1	A	72	VAL	3.3
1	A	220	ILE	3.3
1	B	8	THR	3.3
1	B	11	ILE	3.3
1	A	47	PHE	3.2
1	A	255	GLN	3.2
1	D	4	GLN	3.2
1	D	203	LEU	3.2
1	C	3	VAL	3.2
1	C	19	LEU	3.2
1	A	202	ILE	3.1
1	C	248	ILE	3.1
1	C	219	LYS	3.1
1	A	42	LYS	3.1
1	A	26	LEU	3.0
1	C	197	VAL	2.9
1	A	1	GLU	2.9
1	A	262	TYR	2.8
1	A	71	LEU	2.8
1	B	29	ALA	2.8
1	C	18	ILE	2.8
1	D	202	ILE	2.7
1	A	257	VAL	2.7
1	B	1	GLU	2.7
1	C	139	ILE	2.7
1	A	56	ASN	2.7
1	A	30	LEU	2.7
1	D	31	LEU	2.7
1	C	87	TRP	2.7
1	A	139	ILE	2.7
1	C	27	ILE	2.7
1	C	108	PHE	2.7
1	B	142	VAL	2.6
1	A	194	TRP	2.6
1	B	217	TYR	2.6

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	212	ILE	2.6
1	D	80	PHE	2.6
1	A	39	PHE	2.6
1	B	220	ILE	2.6
1	C	100	ILE	2.6
1	A	138	TYR	2.6
1	A	263	GLY	2.6
1	A	7	SER	2.6
1	C	204	LEU	2.5
1	D	133	GLN	2.5
1	C	76	GLY	2.5
1	A	211	LEU	2.5
1	D	260	ILE	2.5
1	C	113	TRP	2.4
1	C	26	LEU	2.4
1	A	164	VAL	2.4
1	A	212	ILE	2.4
1	A	204	LEU	2.4
1	D	10	GLU	2.4
1	B	156	ALA	2.4
1	D	240	ALA	2.4
1	C	2	LEU	2.4
1	C	270	LYS	2.4
1	A	233	PRO	2.4
1	B	253	PRO	2.4
1	C	253	PRO	2.4
1	C	224	LYS	2.3
1	C	126	VAL	2.3
1	B	68	LEU	2.3
1	A	218	PHE	2.3
1	A	19	LEU	2.3
1	A	142	VAL	2.3
1	B	242	ILE	2.3
1	D	3	VAL	2.3
1	A	248	ILE	2.3
1	C	192	ILE	2.3
1	A	253	PRO	2.2
1	D	88	THR	2.2
1	D	257	VAL	2.2
1	C	242	ILE	2.2
1	C	124	LEU	2.2
1	A	13	ASN	2.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	D	205	ILE	2.2
1	A	189	ARG	2.2
1	B	228	MET	2.2
1	A	242	ILE	2.2
1	D	167	LYS	2.2
1	A	140	TRP	2.2
1	D	192	ILE	2.2
1	A	85	PHE	2.2
1	C	233	PRO	2.2
1	A	48	VAL	2.2
1	C	133	GLN	2.2
1	C	257	VAL	2.2
1	C	31	LEU	2.2
1	A	32	GLY	2.1
1	D	211	LEU	2.1
1	A	207	SER	2.1
1	A	177	VAL	2.1
1	B	143	HIS	2.1
1	C	4	GLN	2.1
1	A	86	ASN	2.1
1	D	48	VAL	2.1
1	A	205	ILE	2.1
1	C	16	HIS	2.0
1	A	12	CYS	2.0
1	D	221	ARG	2.0
1	C	196	ILE	2.0

## 5.2 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95<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( $\text{\AA}^2$ )	Q<0.9
2	NAG	A	302	14/15	0.48	0.21	-	93,104,109,110	0
2	NAG	C	302	14/15	0.61	0.18	-	116,126,129,130	0
2	NAG	D	303	14/15	0.78	0.16	-	52,67,77,80	0
2	NAG	C	303	14/15	0.90	0.14	-	63,74,80,82	0
2	NAG	B	303	14/15	0.77	0.20	-	35,59,78,90	0
2	NAG	A	303	14/15	0.82	0.18	-	56,67,73,73	0
2	NAG	C	304	14/15	0.85	0.17	-	90,97,104,104	0
2	NAG	B	302	14/15	0.78	0.20	-	62,83,90,91	0
2	NAG	A	304	14/15	0.80	0.23	-	91,108,110,111	0
2	NAG	D	302	14/15	0.71	0.23	-	83,93,103,103	0
2	NAG	A	301	14/15	0.78	0.14	-	86,92,97,103	0
2	NAG	D	301	14/15	0.78	0.15	-	46,60,68,70	0
2	NAG	B	301	14/15	0.83	0.15	-	44,59,69,73	0
2	NAG	C	301	14/15	0.48	0.16	-	75,101,105,106	0

## 5.5 Other polymers [i](#)

There are no such residues in this entry.