



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 20, 2020 – 04:50 PM BST

PDB ID : 5Y2M  
Title : Crystal structure of a group 2 HA binding antibody AF4H1K1 Fab in complex with the H4N6 duck isolate (H4-CZ/56) hemagglutinin  
Authors : Xiao, H.; Qi, J.; Gao, F.G.  
Deposited on : 2017-07-26  
Resolution : 3.80 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

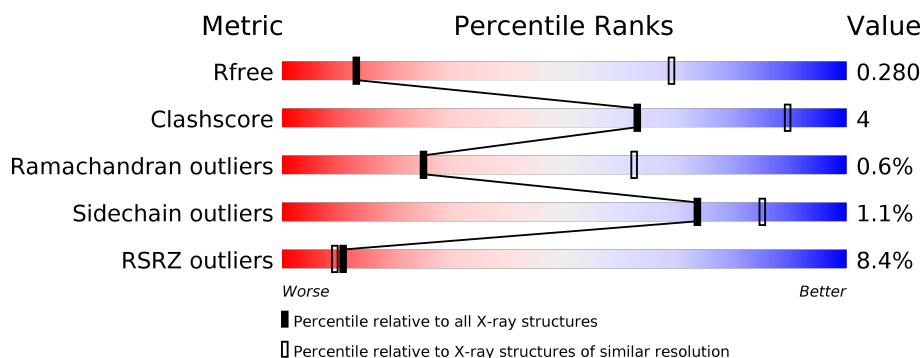
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.80 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	1212 (4.00-3.60)
Clashscore	141614	1288 (4.00-3.60)
Ramachandran outliers	138981	1243 (4.00-3.60)
Sidechain outliers	138945	1237 (4.00-3.60)
RSRZ outliers	127900	1121 (4.00-3.60)

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	327	<div> <div></div> <div>89% 8% ..</div> </div>
1	B	327	<div> <div>2%</div> <div>87% 10% .</div> </div>
1	C	327	<div> <div>2%</div> <div>86% 11% ..</div> </div>
2	D	176	<div> <div>6%</div> <div>88% 10% .</div> </div>
2	E	176	<div> <div>6%</div> <div>90% 8% .</div> </div>
2	F	176	<div> <div>6%</div> <div>91% 7% .</div> </div>

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
3	I	233	
3	K	233	
4	J	220	
4	L	220	

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	601	-	-	-	X

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 18315 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	319	Total	C	N	O	S	0	0	0
			2457	1538	437	470	12			
1	B	319	Total	C	N	O	S	0	0	0
			2457	1538	437	470	12			
1	C	319	Total	C	N	O	S	0	0	0
			2457	1538	437	470	12			

- Molecule 2 is a protein called Hemagglutinin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	172	Total	C	N	O	S	0	0	0
			1404	871	250	279	4			
2	E	172	Total	C	N	O	S	0	0	0
			1404	871	250	279	4			
2	F	172	Total	C	N	O	S	0	0	0
			1404	871	250	279	4			

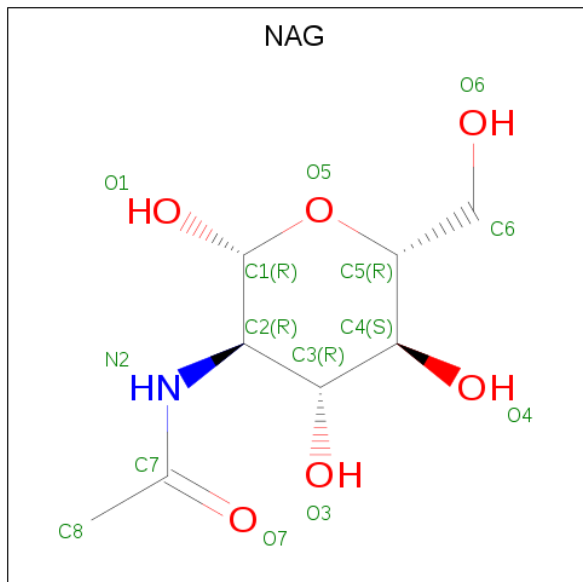
- Molecule 3 is a protein called a group 2 HA binding antibody AF4H1K1 Fab heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	I	224	Total	C	N	O	S	0	0	0
			1694	1077	286	325	6			
3	K	224	Total	C	N	O	S	0	0	0
			1694	1077	286	325	6			

- Molecule 4 is a protein called a group 2 HA binding antibody AF4H1K1 Fab light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	J	212	Total	C	N	O	S	0	0	0
			1630	1018	278	329	5			
4	L	212	Total	C	N	O	S	0	0	0
			1630	1018	278	329	5			

- Molecule 5 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



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

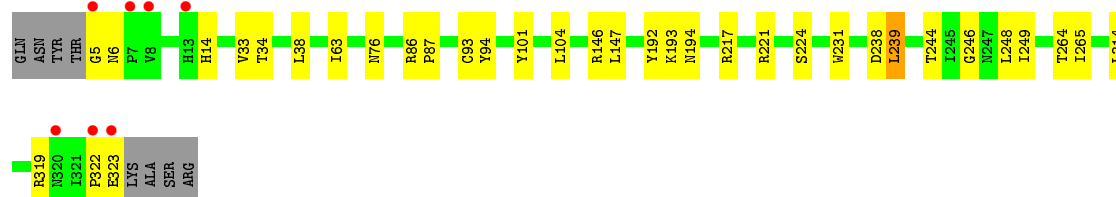
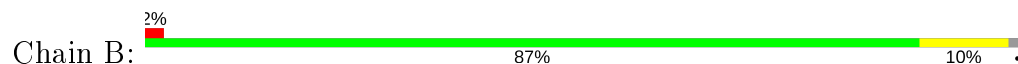
### 3 Residue-property plots [i](#)

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.

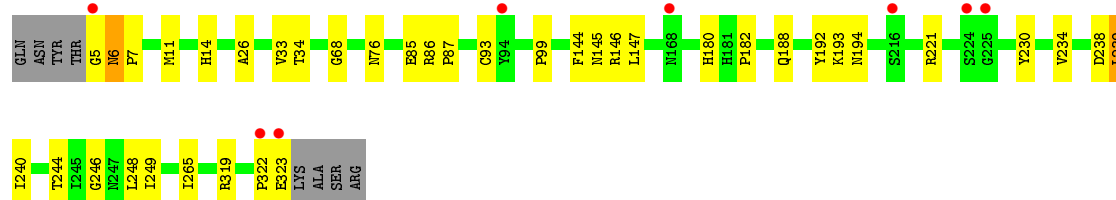
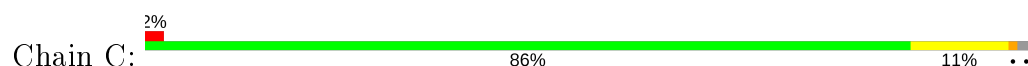
#### • Molecule 1: Hemagglutinin



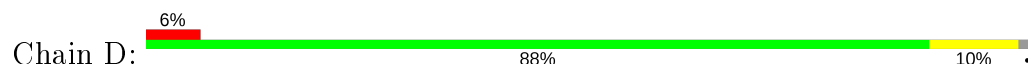
#### • Molecule 1: Hemagglutinin



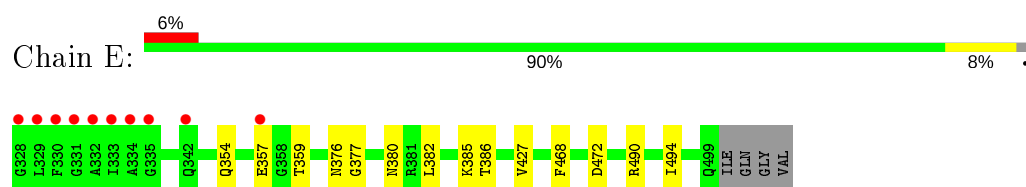
#### • Molecule 1: Hemagglutinin



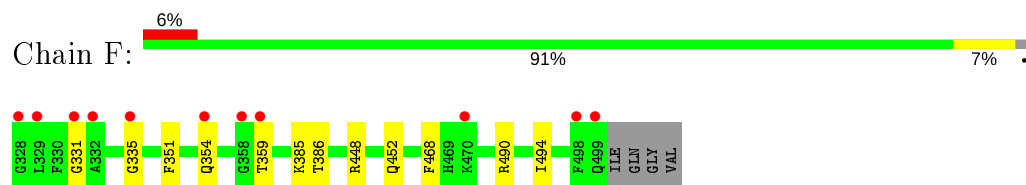
#### • Molecule 2: Hemagglutinin



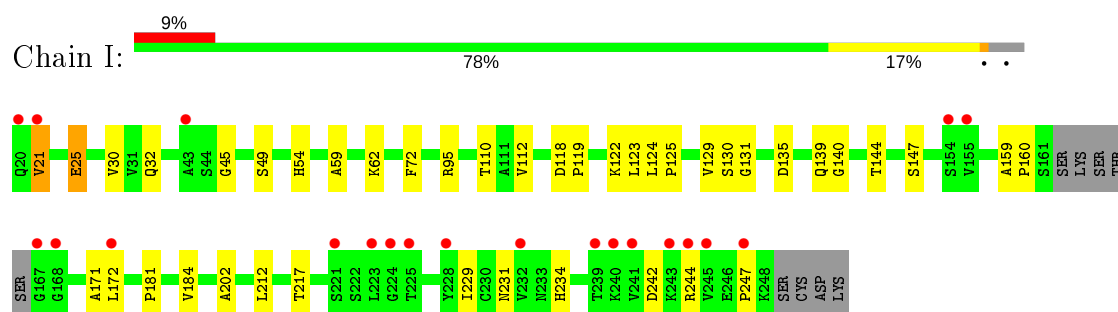
#### • Molecule 2: Hemagglutinin



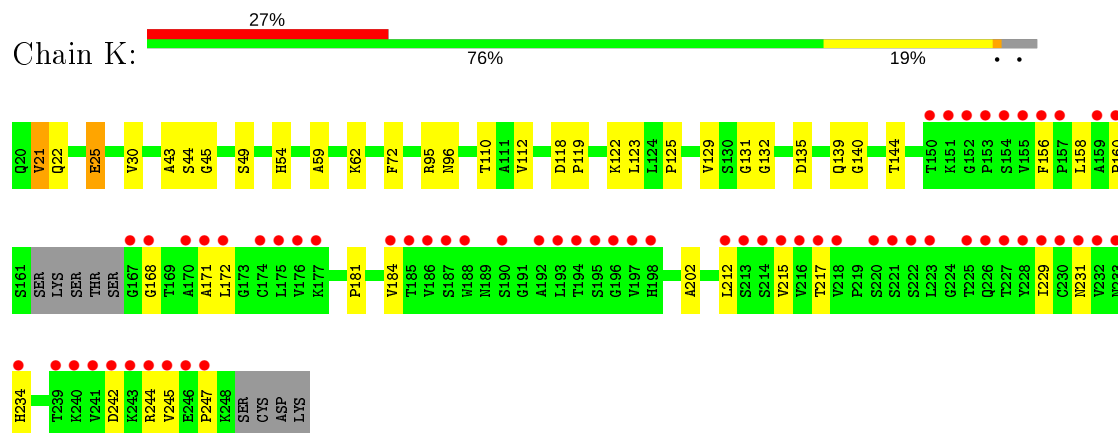
- Molecule 2: Hemagglutinin



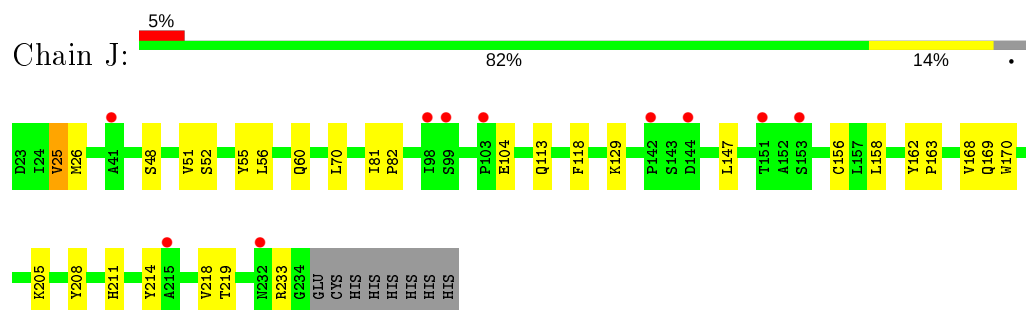
- Molecule 3: a group 2 HA binding antibody AF4H1K1 Fab heavy chain



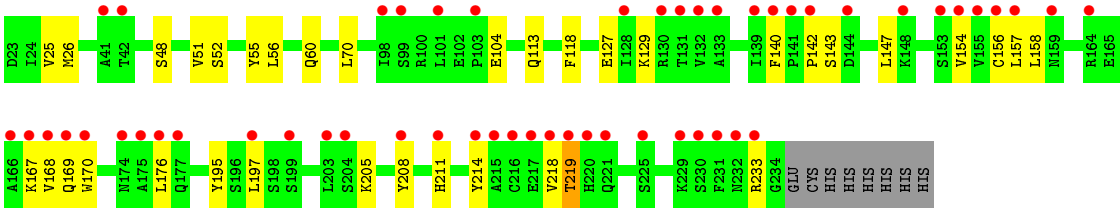
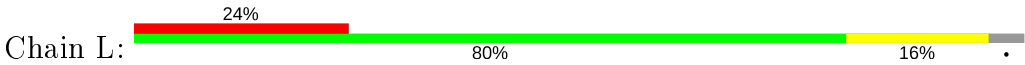
- Molecule 3: a group 2 HA binding antibody AF4H1K1 Fab heavy chain



- Molecule 4: a group 2 HA binding antibody AF4H1K1 Fab light chain



- Molecule 4: a group 2 HA binding antibody AF4H1K1 Fab light chain





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	114.17Å 140.94Å 138.10Å 90.00° 95.31° 90.00°	Depositor
Resolution (Å)	49.21 – 3.80 49.21 – 3.79	Depositor EDS
% Data completeness (in resolution range)	99.3 (49.21-3.80) 99.5 (49.21-3.79)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.27 (at 3.77Å)	Xtriage
Refinement program	PHENIX 1.9_1692	Depositor
R, $R_{free}$	0.235 , 0.277 0.240 , 0.280	Depositor DCC
$R_{free}$ test set	2168 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	75.8	Xtriage
Anisotropy	0.611	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 52.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.86	EDS
Total number of atoms	18315	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	98.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.14% 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: 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	0/2511	0.45	0/3418
1	B	0.25	0/2511	0.45	0/3418
1	C	0.24	0/2511	0.44	0/3418
2	D	0.28	0/1428	0.43	0/1922
2	E	0.27	0/1428	0.43	0/1922
2	F	0.28	0/1428	0.43	0/1922
3	I	0.24	0/1739	0.43	0/2372
3	K	0.24	0/1739	0.42	0/2372
4	J	0.23	0/1665	0.43	0/2258
4	L	0.24	0/1665	0.43	0/2258
All	All	0.25	0/18625	0.43	0/25280

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	2457	0	2402	16	0
1	B	2457	0	2402	21	0
1	C	2457	0	2402	22	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	D	1404	0	1320	13	0
2	E	1404	0	1320	11	0
2	F	1404	0	1320	8	0
3	I	1694	0	1644	24	0
3	K	1694	0	1644	27	0
4	J	1630	0	1581	15	0
4	L	1630	0	1581	19	0
5	A	28	0	26	0	0
5	B	28	0	26	0	0
5	C	28	0	26	0	0
All	All	18315	0	17694	155	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 4.

All (155) 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:B:244:THR:HG22	1:B:246:GLY:H	1.37	0.89
1:A:244:THR:HG22	1:A:246:GLY:H	1.40	0.86
1:C:244:THR:HG22	1:C:246:GLY:H	1.40	0.86
3:I:229:ILE:HG12	3:I:244:ARG:HG3	1.69	0.74
1:B:5:GLY:HA2	2:E:468:PHE:HA	1.68	0.74
1:C:5:GLY:HA2	2:F:468:PHE:HA	1.71	0.72
1:A:5:GLY:HA2	2:D:468:PHE:HA	1.72	0.70
4:L:211:HIS:O	4:L:233:ARG:NH2	2.26	0.69
3:K:229:ILE:HG12	3:K:244:ARG:HG3	1.75	0.68
4:J:211:HIS:O	4:J:233:ARG:NH2	2.27	0.66
4:L:60:GLN:HB2	4:L:70:LEU:HD11	1.78	0.66
1:C:322:PRO:O	1:C:323:GLU:HB2	1.95	0.66
1:B:322:PRO:O	1:B:323:GLU:HB2	1.95	0.66
3:K:231:ASN:ND2	3:K:242:ASP:OD1	2.28	0.66
1:A:322:PRO:O	1:A:323:GLU:HB2	1.95	0.65
3:K:202:ALA:HB2	3:K:212:LEU:HD23	1.80	0.64
4:J:60:GLN:HB2	4:J:70:LEU:HD11	1.80	0.63
3:K:160:PRO:HG3	3:K:172:LEU:HB3	1.80	0.63
3:I:160:PRO:HG3	3:I:172:LEU:HB3	1.79	0.63
4:J:25:VAL:HG22	4:J:48:SER:HB2	1.83	0.61
2:E:376:ASN:OD1	3:K:123:LEU:HB3	2.00	0.61
2:E:354:GLN:HG3	2:E:359:THR:HG22	1.82	0.61
3:I:125:PRO:HD2	3:I:129:VAL:HG21	1.84	0.60

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:J:26:MET:HE3	4:J:113:GLN:HB3	1.84	0.60
1:A:93:CYS:O	1:A:221:ARG:NH1	2.35	0.60
3:K:30:VAL:HG21	3:K:181:PRO:HG3	1.82	0.59
3:I:231:ASN:ND2	3:I:242:ASP:OD1	2.33	0.59
3:I:30:VAL:HG21	3:I:181:PRO:HG3	1.84	0.58
3:I:202:ALA:HB2	3:I:212:LEU:HD23	1.85	0.58
4:L:25:VAL:HG22	4:L:48:SER:HB2	1.86	0.58
2:D:354:GLN:HG3	2:D:359:THR:HG22	1.86	0.58
1:B:93:CYS:O	1:B:221:ARG:NH1	2.36	0.57
4:L:25:VAL:H	4:L:48:SER:HB3	1.70	0.57
1:C:33:VAL:HG12	1:C:34:THR:HG23	1.86	0.57
3:I:119:PRO:HD3	3:I:135:ASP:HB2	1.87	0.56
1:C:76:ASN:OD1	1:C:146:ARG:NH2	2.37	0.56
3:K:119:PRO:HD3	3:K:135:ASP:HB2	1.86	0.56
3:K:125:PRO:HD2	3:K:129:VAL:HG21	1.87	0.56
2:D:377:GLY:HA3	1:C:26:ALA:O	2.07	0.55
3:I:49:SER:O	3:I:72:PHE:HB2	2.07	0.55
2:F:354:GLN:HG3	2:F:359:THR:HG22	1.89	0.55
3:K:172:LEU:HB2	3:K:245:VAL:HG11	1.87	0.54
3:K:25:GLU:OE1	3:K:140:GLY:N	2.31	0.54
1:A:192:TYR:O	1:A:194:ASN:N	2.40	0.54
2:E:385:LYS:HD2	2:E:386:THR:N	2.23	0.54
3:K:158:LEU:HB3	4:L:140:PHE:CG	2.43	0.54
3:K:171:ALA:HB2	3:K:217:THR:HG22	1.89	0.54
1:A:26:ALA:O	2:E:377:GLY:HA3	2.08	0.53
3:I:171:ALA:HB2	3:I:217:THR:HG22	1.90	0.53
1:B:319:ARG:HG2	1:B:319:ARG:HH11	1.73	0.53
2:F:385:LYS:HD2	2:F:386:THR:N	2.24	0.53
1:B:33:VAL:HG12	1:B:34:THR:HG23	1.90	0.53
1:A:147:LEU:HD12	1:A:249:ILE:HG22	1.91	0.52
3:I:25:GLU:OE1	3:I:139:GLN:N	2.43	0.52
2:F:385:LYS:HD2	2:F:386:THR:H	1.73	0.52
1:A:33:VAL:HG12	1:A:34:THR:HG23	1.91	0.52
1:B:63:ILE:HG13	1:B:101:TYR:CZ	2.43	0.52
4:J:169:GLN:HG2	4:J:176:LEU:HD11	1.92	0.52
2:E:385:LYS:HD2	2:E:386:THR:H	1.75	0.52
4:J:25:VAL:H	4:J:48:SER:HB3	1.76	0.51
4:L:26:MET:HE3	4:L:113:GLN:HB3	1.91	0.51
4:L:169:GLN:HG2	4:L:176:LEU:HD11	1.93	0.51
2:D:385:LYS:HD2	2:D:386:THR:N	2.26	0.51
1:C:93:CYS:O	1:C:221:ARG:NH1	2.43	0.51

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:D:385:LYS:HD2	2:D:386:THR:H	1.76	0.50
3:K:25:GLU:OE1	3:K:139:GLN:N	2.43	0.50
3:K:49:SER:O	3:K:72:PHE:HB2	2.10	0.50
1:B:192:TYR:O	1:B:194:ASN:N	2.44	0.50
1:C:319:ARG:HG2	1:C:319:ARG:HH11	1.75	0.50
3:K:21:VAL:HA	3:K:45:GLY:HA3	1.92	0.50
3:I:184:VAL:HG12	3:I:234:HIS:HB2	1.94	0.50
4:L:142:PRO:HD3	4:L:154:VAL:HG22	1.93	0.50
3:K:215:VAL:HG21	4:L:157:LEU:HD13	1.93	0.50
1:A:238:ASP:OD1	1:A:239:LEU:N	2.40	0.50
1:A:319:ARG:HH11	1:A:319:ARG:HG2	1.76	0.50
1:A:76:ASN:OD1	1:A:146:ARG:NH2	2.45	0.49
1:C:192:TYR:O	1:C:194:ASN:N	2.45	0.49
2:D:369:GLN:NE2	3:I:130:SER:OG	2.45	0.49
3:K:43:ALA:O	3:K:96:ASN:ND2	2.45	0.49
2:E:380:ASN:OD1	3:K:122:LYS:HG3	2.12	0.49
4:L:147:LEU:O	4:L:205:LYS:HD2	2.13	0.49
3:K:22:GLN:N	3:K:44:SER:O	2.45	0.48
3:K:184:VAL:HG12	3:K:234:HIS:HB2	1.94	0.48
1:B:314:LEU:HD22	2:E:382:LEU:HD23	1.94	0.48
3:I:32:GLN:HG2	3:I:147:SER:HA	1.94	0.48
1:C:147:LEU:HD12	1:C:249:ILE:HG22	1.97	0.47
4:J:156:CYS:HB2	4:J:170:TRP:CH2	2.49	0.47
4:J:147:LEU:O	4:J:205:LYS:HD2	2.15	0.47
1:A:63:ILE:HG13	1:A:101:TYR:CZ	2.50	0.47
4:L:127:GLU:OE2	4:L:195:TYR:OH	2.28	0.47
1:C:238:ASP:OD1	1:C:239:LEU:N	2.44	0.46
1:C:182:PRO:HG2	1:C:188:GLN:OE1	2.15	0.46
1:B:76:ASN:OD1	1:B:146:ARG:NH2	2.49	0.46
2:D:331:GLY:O	2:D:335:GLY:HA3	2.15	0.46
2:E:357:GLU:OE2	2:E:472:ASP:HB2	2.16	0.45
4:L:208:TYR:O	4:L:214:TYR:OH	2.32	0.45
1:C:11:MET:HE1	2:F:351:PHE:CE2	2.52	0.45
1:C:244:THR:HG21	1:C:248:LEU:HB2	1.98	0.45
3:K:54:HIS:CD2	3:K:118:ASP:HB2	2.52	0.45
4:L:168:VAL:HG12	4:L:218:VAL:HG22	1.98	0.45
1:A:244:THR:HG21	1:A:248:LEU:HB2	1.98	0.44
1:B:147:LEU:HD12	1:B:249:ILE:HG22	1.99	0.44
1:C:85:GLU:HB2	1:C:265:ILE:HD11	1.99	0.44
1:C:68:GLY:HA3	1:C:145:ASN:OD1	2.16	0.44
2:D:376:ASN:OD1	3:I:123:LEU:HB3	2.17	0.44

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:I:21:VAL:HA	3:I:45:GLY:HA3	1.99	0.44
2:D:448:ARG:O	2:D:452:GLN:HG3	2.18	0.44
1:A:11:MET:HE1	2:D:351:PHE:CE2	2.52	0.43
3:K:59:ALA:HB3	3:K:62:LYS:HB2	1.99	0.43
2:D:380:ASN:OD1	3:I:122:LYS:HG3	2.18	0.43
3:I:131:GLY:HA3	4:J:118:PHE:CE1	2.52	0.43
1:B:38:LEU:HD11	1:B:314:LEU:HB2	2.00	0.43
3:I:110:THR:HG23	3:I:144:THR:HA	2.00	0.43
1:A:6:ASN:H	1:A:7:PRO:CD	2.32	0.43
1:B:238:ASP:OD1	1:B:239:LEU:N	2.42	0.43
1:B:93:CYS:SG	1:B:94:TYR:N	2.89	0.43
1:B:244:THR:HG21	1:B:248:LEU:HB2	2.01	0.43
3:I:124:LEU:HD12	3:I:129:VAL:HB	2.00	0.42
4:L:158:LEU:HB2	4:L:197:LEU:HB3	2.00	0.42
1:B:104:LEU:HB2	1:B:231:TRP:CZ3	2.53	0.42
3:K:202:ALA:HA	3:K:212:LEU:HB3	2.00	0.42
1:C:234:VAL:HG21	1:C:240:ILE:HB	2.01	0.42
3:I:54:HIS:CD2	3:I:118:ASP:HB2	2.55	0.42
4:J:208:TYR:O	4:J:214:TYR:OH	2.35	0.42
3:K:131:GLY:HA3	4:L:118:PHE:CE1	2.54	0.42
3:K:156:PHE:HB3	4:L:143:SER:OG	2.20	0.42
1:B:63:ILE:HG13	1:B:101:TYR:CE2	2.55	0.42
1:C:6:ASN:H	1:C:7:PRO:CD	2.33	0.42
3:I:202:ALA:HA	3:I:212:LEU:HB3	2.02	0.42
4:L:167:LYS:HB3	4:L:219:THR:HG23	2.01	0.42
3:I:25:GLU:OE1	3:I:140:GLY:N	2.33	0.42
1:C:180:HIS:ND1	1:C:192:TYR:OH	2.39	0.42
3:I:59:ALA:HB3	3:I:62:LYS:HB2	2.01	0.42
2:D:378:LYS:NZ	2:D:430:GLU:O	2.53	0.42
2:F:331:GLY:O	2:F:335:GLY:HA3	2.20	0.41
1:B:264:THR:OG1	1:B:265:ILE:N	2.52	0.41
4:L:156:CYS:HB2	4:L:170:TRP:CH2	2.56	0.41
1:B:86:ARG:HA	1:B:87:PRO:HD3	1.95	0.41
2:E:490:ARG:O	2:E:494:ILE:HG12	2.21	0.41
4:J:81:ILE:HA	4:J:82:PRO:HD2	1.84	0.41
1:A:314:LEU:HD22	2:D:382:LEU:HD23	2.02	0.41
4:J:168:VAL:HG12	4:J:218:VAL:HG22	2.02	0.41
2:F:448:ARG:O	2:F:452:GLN:HG3	2.20	0.41
4:J:147:LEU:H	4:J:147:LEU:HD12	1.83	0.41
1:C:86:ARG:HA	1:C:87:PRO:HD3	1.94	0.41
1:B:38:LEU:HD12	2:E:427:VAL:HG22	2.02	0.41

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:K:132:GLY:H	4:L:118:PHE:HE1	1.68	0.41
1:C:144:PHE:HB2	1:C:147:LEU:HB2	2.03	0.41
3:I:159:ALA:HA	3:I:160:PRO:HD3	1.82	0.41
4:J:158:LEU:HB2	4:J:197:LEU:HB3	2.03	0.41
1:B:217:ARG:HB2	1:B:224:SER:O	2.21	0.40
4:J:162:TYR:CG	4:J:163:PRO:HA	2.56	0.40
1:C:99:PRO:HG2	1:C:230:TYR:CE2	2.57	0.40
2:F:490:ARG:O	2:F:494:ILE:HG12	2.21	0.40
3:K:110:THR:HG23	3:K:144:THR:HA	2.02	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	317/327 (97%)	304 (96%)	11 (4%)	2 (1%)	25	62
1	B	317/327 (97%)	304 (96%)	11 (4%)	2 (1%)	25	62
1	C	317/327 (97%)	304 (96%)	11 (4%)	2 (1%)	25	62
2	D	170/176 (97%)	161 (95%)	9 (5%)	0	100	100
2	E	170/176 (97%)	161 (95%)	9 (5%)	0	100	100
2	F	170/176 (97%)	160 (94%)	10 (6%)	0	100	100
3	I	220/233 (94%)	205 (93%)	14 (6%)	1 (0%)	29	66
3	K	220/233 (94%)	204 (93%)	14 (6%)	2 (1%)	17	54
4	J	210/220 (96%)	194 (92%)	13 (6%)	3 (1%)	11	46
4	L	210/220 (96%)	194 (92%)	13 (6%)	3 (1%)	11	46
All	All	2321/2415 (96%)	2191 (94%)	115 (5%)	15 (1%)	25	62

All (15) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	J	51	VAL
4	L	51	VAL
3	K	168	GLY
1	A	6	ASN
1	A	193	LYS
1	B	6	ASN
1	B	193	LYS
4	J	55	TYR
1	C	6	ASN
1	C	193	LYS
3	I	247	PRO
3	K	247	PRO
4	L	52	SER
4	L	55	TYR
4	J	52	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	274/282 (97%)	272 (99%)	2 (1%)	84	91
1	B	274/282 (97%)	272 (99%)	2 (1%)	84	91
1	C	274/282 (97%)	272 (99%)	2 (1%)	84	91
2	D	147/150 (98%)	147 (100%)	0	100	100
2	E	147/150 (98%)	147 (100%)	0	100	100
2	F	147/150 (98%)	147 (100%)	0	100	100
3	I	187/196 (95%)	183 (98%)	4 (2%)	53	74
3	K	187/196 (95%)	183 (98%)	4 (2%)	53	74
4	J	184/192 (96%)	179 (97%)	5 (3%)	44	69
4	L	184/192 (96%)	180 (98%)	4 (2%)	52	72
All	All	2005/2072 (97%)	1982 (99%)	23 (1%)	73	85

All (23) residues with a non-rotameric sidechain are listed below:



Mol	Chain	Res	Type
1	A	14	HIS
1	A	239	LEU
1	B	14	HIS
1	B	239	LEU
1	C	14	HIS
1	C	239	LEU
3	I	21	VAL
3	I	25	GLU
3	I	95	ARG
3	I	112	VAL
4	J	25	VAL
4	J	56	LEU
4	J	104	GLU
4	J	129	LYS
4	J	219	THR
3	K	21	VAL
3	K	25	GLU
3	K	95	ARG
3	K	112	VAL
4	L	56	LEU
4	L	104	GLU
4	L	129	LYS
4	L	219	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	74	HIS
2	D	369	GLN
2	E	452	GLN
1	C	74	HIS

### 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 monosaccharides in this entry.

## 5.6 Ligand geometry

6 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	C	602	1	14,14,15	0.22	0	17,19,21	0.43	0
5	NAG	B	601	1	14,14,15	0.35	0	17,19,21	0.32	0
5	NAG	A	601	1	14,14,15	0.28	0	17,19,21	0.36	0
5	NAG	C	601	1	14,14,15	0.40	0	17,19,21	0.31	0
5	NAG	B	602	1	14,14,15	0.29	0	17,19,21	0.48	0
5	NAG	A	602	1	14,14,15	0.29	0	17,19,21	0.69	1 (5%)

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	C	602	1	-	2/6/23/26	0/1/1/1
5	NAG	B	601	1	-	2/6/23/26	0/1/1/1
5	NAG	A	601	1	-	1/6/23/26	0/1/1/1
5	NAG	C	601	1	-	2/6/23/26	0/1/1/1
5	NAG	B	602	1	-	2/6/23/26	0/1/1/1
5	NAG	A	602	1	-	2/6/23/26	0/1/1/1

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	602	NAG	C1-O5-C5	2.01	114.91	112.19

There are no chirality outliers.

All (11) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	C	602	NAG	C8-C7-N2-C2
5	C	602	NAG	O7-C7-N2-C2
5	B	602	NAG	C8-C7-N2-C2
5	B	602	NAG	O7-C7-N2-C2
5	A	602	NAG	C8-C7-N2-C2
5	A	602	NAG	O7-C7-N2-C2
5	C	601	NAG	O5-C5-C6-O6
5	B	601	NAG	O5-C5-C6-O6
5	C	601	NAG	C4-C5-C6-O6
5	A	601	NAG	O5-C5-C6-O6
5	B	601	NAG	C4-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	319/327 (97%)	0.06	3 (0%) 84 79	33, 68, 102, 181	0
1	B	319/327 (97%)	0.14	7 (2%) 62 54	36, 71, 106, 205	0
1	C	319/327 (97%)	0.16	8 (2%) 57 49	41, 78, 110, 170	0
2	D	172/176 (97%)	0.40	10 (5%) 23 18	40, 90, 171, 309	0
2	E	172/176 (97%)	0.39	10 (5%) 23 18	45, 102, 173, 346	0
2	F	172/176 (97%)	0.35	11 (6%) 19 15	34, 110, 175, 299	0
3	I	224/233 (96%)	0.58	21 (9%) 8 7	41, 95, 171, 209	0
3	K	224/233 (96%)	1.36	62 (27%) 0 0	55, 101, 254, 347	0
4	J	212/220 (96%)	0.59	11 (5%) 27 24	55, 94, 148, 187	0
4	L	212/220 (96%)	1.42	53 (25%) 0 0	72, 143, 239, 296	0
All	All	2345/2415 (97%)	0.50	196 (8%) 11 9	33, 87, 191, 347	0

All (196) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	K	217	THR	11.0
3	K	175	LEU	9.7
3	K	193	LEU	9.4
3	K	215	VAL	9.3
3	K	214	SER	8.5
3	K	167	GLY	8.4
1	B	322	PRO	7.8
3	K	228	TYR	7.7
4	L	141	PRO	7.3
4	L	230	SER	6.6
2	E	332	ALA	6.5
3	K	198	HIS	6.5
2	E	334	ALA	6.3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
3	K	154	SER	6.2
3	K	216	VAL	6.1
3	K	176	VAL	6.1
2	E	331	GLY	5.9
2	E	335	GLY	5.9
3	K	233	ASN	5.8
3	K	232	VAL	5.6
4	L	156	CYS	5.6
3	K	174	CYS	5.4
3	K	218	VAL	5.4
1	B	5	GLY	5.2
3	K	239	THR	5.2
3	K	159	ALA	5.1
4	L	232	ASN	5.1
3	K	155	VAL	5.1
4	L	167	LYS	5.1
4	L	142	PRO	5.1
4	L	157	LEU	5.0
4	L	153	SER	5.0
3	K	185	THR	5.0
3	K	168	GLY	5.0
3	K	172	LEU	4.9
2	D	335	GLY	4.9
3	K	197	VAL	4.9
3	K	241	VAL	4.8
3	I	224	GLY	4.8
3	K	186	VAL	4.7
4	L	216	CYS	4.6
3	I	167	GLY	4.4
3	I	228	TYR	4.4
3	I	223	LEU	4.4
3	K	221	SER	4.4
3	K	195	SER	4.3
2	F	329	LEU	4.3
4	L	219	THR	4.3
4	L	215	ALA	4.2
3	K	227	THR	4.2
2	D	331	GLY	4.2
3	K	153	PRO	4.1
2	E	330	PHE	4.1
2	D	334	ALA	4.1
3	K	196	GLY	4.0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
3	K	222	SER	4.0
4	L	144	ASP	4.0
4	L	148	LYS	3.9
3	K	184	VAL	3.9
2	E	333	ILE	3.9
4	L	221	GLN	3.8
3	I	225	THR	3.8
4	L	169	GLN	3.8
4	L	154	VAL	3.7
3	I	221	SER	3.7
2	D	328	GLY	3.6
3	I	20	GLN	3.6
1	C	5	GLY	3.5
3	K	151	LYS	3.5
3	K	188	TRP	3.5
2	E	342	GLN	3.5
3	K	194	THR	3.5
4	L	139	ILE	3.5
4	L	211	HIS	3.5
1	A	5	GLY	3.4
4	L	174	ASN	3.4
4	L	133	ALA	3.4
3	K	160	PRO	3.4
3	K	231	ASN	3.3
4	L	220	HIS	3.3
3	I	241	VAL	3.3
4	L	218	VAL	3.3
4	L	229	LYS	3.3
3	K	242	ASP	3.3
3	K	244	ARG	3.2
3	K	230	CYS	3.2
4	L	233	ARG	3.2
3	K	156	PHE	3.2
3	K	226	GLN	3.2
3	I	243	LYS	3.1
4	J	99	SER	3.1
4	L	231	PHE	3.1
1	A	7	PRO	3.1
2	D	329	LEU	3.0
2	D	332	ALA	3.0
4	J	142	PRO	3.0
4	L	41	ALA	3.0

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
3	K	240	LYS	3.0
3	K	212	LEU	3.0
2	F	335	GLY	3.0
4	L	203	LEU	3.0
4	J	151	THR	2.9
4	L	177	GLN	2.9
3	K	192	ALA	2.9
3	K	213	SER	2.9
3	K	225	THR	2.9
4	L	197	LEU	2.9
1	B	323	GLU	2.9
3	I	245	VAL	2.9
4	L	168	VAL	2.8
3	K	152	GLY	2.8
4	L	130	ARG	2.8
4	L	199	SER	2.8
3	I	155	VAL	2.8
3	K	171	ALA	2.8
4	L	131	THR	2.8
1	C	216	SER	2.8
2	F	328	GLY	2.8
3	K	223	LEU	2.8
1	C	322	PRO	2.7
3	I	43	ALA	2.7
2	F	470	LYS	2.7
2	F	359	THR	2.7
4	J	153	SER	2.6
4	L	166	ALA	2.6
3	I	244	ARG	2.6
3	K	220	SER	2.6
4	L	155	VAL	2.6
3	I	168	GLY	2.6
4	L	132	VAL	2.6
4	L	159	ASN	2.5
4	L	214	TYR	2.5
4	J	41	ALA	2.5
1	C	168	ASN	2.5
3	I	154	SER	2.5
3	I	232	VAL	2.5
3	K	245	VAL	2.5
2	F	331	GLY	2.5
1	A	323	GLU	2.5

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	D	474	ASN	2.5
2	E	357	GLU	2.5
4	L	225	SER	2.5
4	L	42	THR	2.4
2	F	499	GLN	2.4
3	K	177	LYS	2.4
1	C	224	SER	2.4
3	K	157	PRO	2.4
4	L	98	ILE	2.4
4	L	217	GLU	2.4
2	F	354	GLN	2.4
3	I	247	PRO	2.4
3	K	247	PRO	2.4
2	E	329	LEU	2.4
4	J	215	ALA	2.4
4	J	202	THR	2.4
4	L	101	LEU	2.4
2	D	467	ILE	2.3
4	L	208	TYR	2.3
2	F	332	ALA	2.3
4	L	99	SER	2.3
3	K	229	ILE	2.3
3	K	150	THR	2.3
4	J	232	ASN	2.3
2	D	469	HIS	2.3
3	K	243	LYS	2.2
1	B	8	VAL	2.2
2	D	471	CYS	2.2
3	K	190	SER	2.2
1	B	7	PRO	2.2
4	L	103	PRO	2.2
3	I	240	LYS	2.2
4	L	128	ILE	2.2
3	K	234	HIS	2.2
3	I	239	THR	2.2
3	I	21	VAL	2.2
4	L	175	ALA	2.1
1	C	94	TYR	2.1
4	L	170	TRP	2.1
4	J	144	ASP	2.1
3	K	246	GLU	2.1
4	L	176	LEU	2.1

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
4	J	98	ILE	2.1
3	I	172	LEU	2.1
3	K	170	ALA	2.1
2	F	358	GLY	2.1
4	J	103	PRO	2.1
2	E	328	GLY	2.1
4	L	204	SER	2.0
1	B	320	ASN	2.0
2	F	498	PHE	2.0
3	K	187	SER	2.0
1	C	323	GLU	2.0
4	L	140	PHE	2.0
1	C	225	GLY	2.0
4	L	164	ARG	2.0
1	B	13	HIS	2.0

## 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 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
5	NAG	A	601	14/15	0.65	0.41	102,123,129,129	0
5	NAG	B	602	14/15	0.78	0.33	67,77,82,93	0
5	NAG	C	601	14/15	0.79	0.37	104,123,130,132	0
5	NAG	A	602	14/15	0.79	0.29	65,78,151,153	0
5	NAG	C	602	14/15	0.82	0.37	63,69,78,86	0
5	NAG	B	601	14/15	0.83	0.34	114,125,131,132	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.