



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 16, 2021 – 11:22 PM EDT

PDB ID : 1RZK  
Title : HIV-1 YU2 GP120 ENVELOPE GLYCOPROTEIN COMPLEXED WITH CD4 AND INDUCED NEUTRALIZING ANTIBODY 17B  
Authors : Huang, C.C.; Venturi, M.; Majeed, S.; Moore, M.J.; Phogat, S.; Zhang, M.-Y.; Dimitrov, D.S.; Hendrickson, W.A.; Robinson, J.; Sodroski, J.; Wyatt, R.; Choe, H.; Farzan, M.; Kwong, P.D.  
Deposited on : 2003-12-24  
Resolution : 2.90 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

<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	:	2.23.2
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)	:	2.23.2

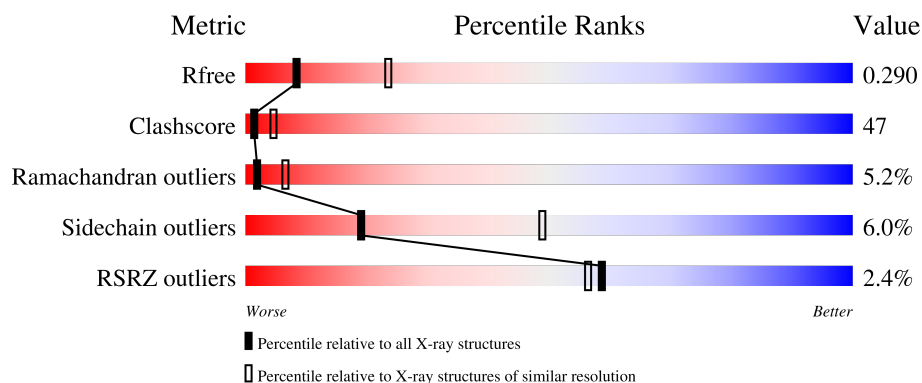
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

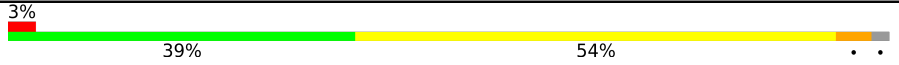
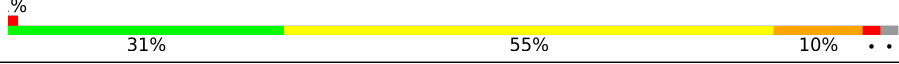
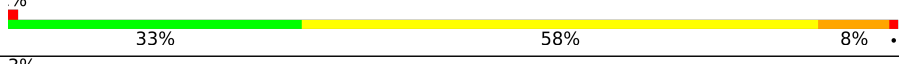

The reported resolution of this entry is 2.90 Å.

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	1957 (2.90-2.90)
Clashscore	141614	2172 (2.90-2.90)
Ramachandran outliers	138981	2115 (2.90-2.90)
Sidechain outliers	138945	2117 (2.90-2.90)
RSRZ outliers	127900	1906 (2.90-2.90)

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 of 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	G	313	
2	C	185	
3	L	214	
4	H	229	

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	G	588	X	-	-	X
5	NAG	G	741	X	-	-	-
5	NAG	G	908	-	-	-	X

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 7706 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 ENVELOPE GLYCOPROTEIN GP120.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	G	306	Total	C	N	O	S	1	0	0
			2385	1494	417	454	20			

- Molecule 2 is a protein called T-CELL SURFACE GLYCOPROTEIN CD4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	181	Total	C	N	O	S	0	0	0
			1412	885	247	276	4			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	184	ASN	SER	engineered mutation	UNP P01730
C	185	THR	ILE	engineered mutation	UNP P01730

- Molecule 3 is a protein called ANTIBODY 17B, LIGHT CHAIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	L	214	Total	C	N	O	S	0	0	0
			1647	1028	282	332	5			

- Molecule 4 is a protein called ANTIBODY 17B, HEAVY CHAIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	H	229	Total	C	N	O	S	0	0	0
			1722	1086	289	342	5			

- Molecule 5 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: C<sub>8</sub>H<sub>15</sub>NO<sub>6</sub>).



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

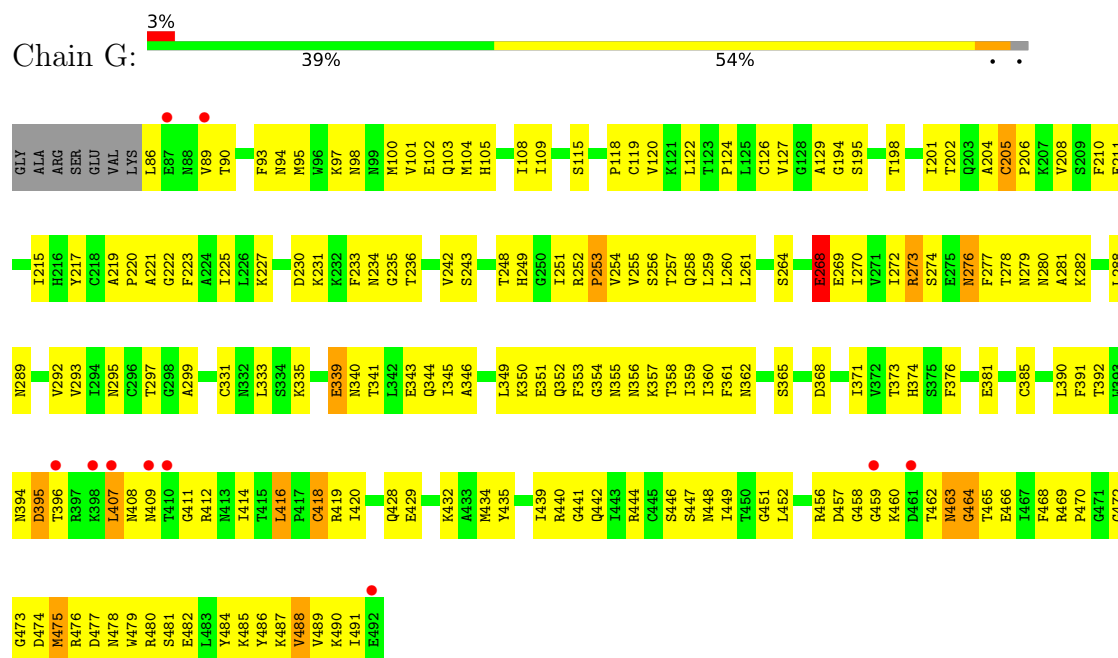
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	G	135	Total 135	O 135	0	0
6	C	57	Total 57	O 57	0	0
6	L	73	Total 73	O 73	0	0
6	H	79	Total 79	O 79	0	0

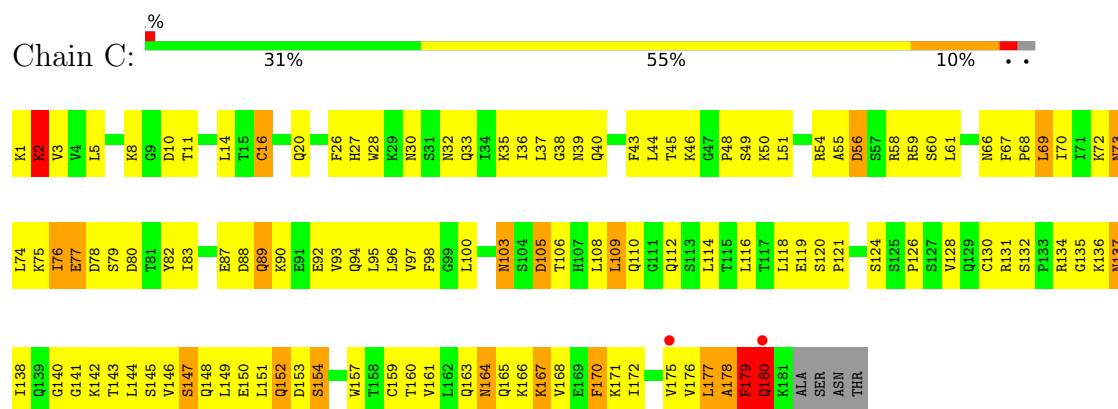
### 3 Residue-property plots

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: ENVELOPE GLYCOPROTEIN GP120

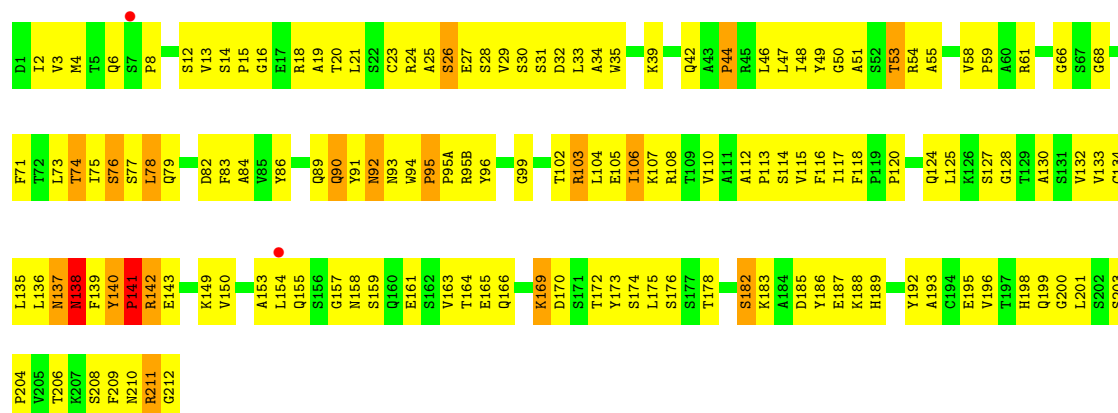


#### • Molecule 2: T-CELL SURFACE GLYCOPROTEIN CD4

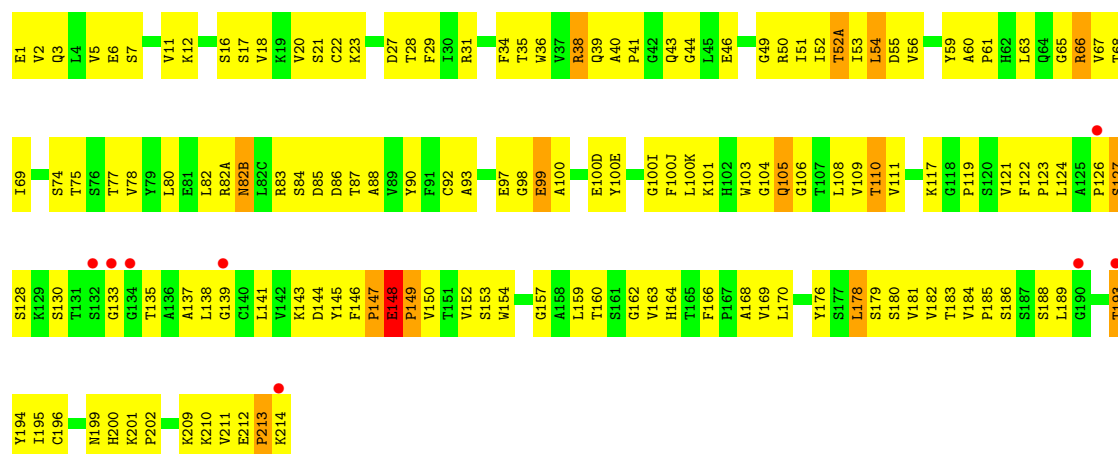


#### • Molecule 3: ANTIBODY 17B, LIGHT CHAIN





● Molecule 4: ANTIBODY 17B, HEAVY CHAIN





## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	174.98Å 81.71Å 74.48Å 90.00° 90.37° 90.00°	Depositor
Resolution (Å)	20.00 – 2.90 17.90 – 2.80	Depositor EDS
% Data completeness (in resolution range)	98.6 (20.00-2.90) 98.3 (17.90-2.80)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	0.09	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.89 (at 2.78Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, $R_{free}$	0.205 , 0.297 0.215 , 0.290	Depositor DCC
$R_{free}$ test set	1293 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	56.1	Xtriage
Anisotropy	0.310	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.24 , 54.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	0.037 for -h,-k,l	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	7706	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	66.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.11% 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 [i](#)

### 5.1 Standard geometry [i](#)

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

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 5$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\# Z  > 5$	RMSZ	$\# Z  > 5$
1	G	0.45	0/2432	0.66	0/3296
2	C	0.40	0/1432	0.72	2/1930 (0.1%)
3	L	0.43	0/1684	0.86	3/2288 (0.1%)
4	H	0.42	0/1762	0.64	0/2399
All	All	0.43	0/7310	0.72	5/9913 (0.1%)

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	L	140	TYR	C-N-CD	-21.37	73.58	120.60
3	L	140	TYR	C-N-CA	13.72	179.64	122.00
2	C	179	PHE	N-CA-C	-9.53	85.27	111.00
2	C	180	GLN	N-CA-C	8.14	132.98	111.00
3	L	141	PRO	N-CA-C	-5.44	97.97	112.10

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts [i](#)

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	G	2385	0	2327	220	0
2	C	1412	0	1444	161	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	L	1647	0	1593	179	0
4	H	1722	0	1691	164	0
5	G	196	0	182	22	0
6	C	57	0	0	9	0
6	G	135	0	0	13	0
6	H	79	0	0	10	0
6	L	73	0	0	11	0
All	All	7706	0	7237	692	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 47.

The worst 5 of 692 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:C:176:VAL:C	2:C:177:LEU:HD12	1.62	1.17
2:C:108:LEU:HD21	2:C:112:GLN:HB3	1.28	1.12
4:H:148:GLU:HG3	4:H:149:PRO:HA	1.34	1.07
2:C:178:ALA:O	2:C:179:PHE:HD1	1.38	1.06
2:C:179:PHE:O	2:C:180:GLN:HB2	1.62	0.99

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	G	304/313 (97%)	242 (80%)	48 (16%)	14 (5%)	<a href="#">2</a> <a href="#">9</a>
2	C	179/185 (97%)	127 (71%)	38 (21%)	14 (8%)	<a href="#">1</a> <a href="#">2</a>
3	L	212/214 (99%)	171 (81%)	28 (13%)	13 (6%)	<a href="#">1</a> <a href="#">4</a>
4	H	227/229 (99%)	184 (81%)	36 (16%)	7 (3%)	<a href="#">4</a> <a href="#">16</a>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	922/941 (98%)	724 (78%)	150 (16%)	48 (5%)	<b>2</b> <b>6</b>

5 of 48 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	G	268	GLU
1	G	409	ASN
1	G	475	MET
2	C	109	LEU
2	C	165	GLN

### 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	G	271/276 (98%)	257 (95%)	14 (5%)	<b>23</b> 55
2	C	164/167 (98%)	149 (91%)	15 (9%)	<b>9</b> 28
3	L	184/184 (100%)	174 (95%)	10 (5%)	<b>22</b> 54
4	H	193/193 (100%)	183 (95%)	10 (5%)	<b>23</b> 55
All	All	812/820 (99%)	763 (94%)	49 (6%)	<b>19</b> 49

5 of 49 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	C	177	LEU
3	L	137	ASN
3	L	53	THR
3	L	92	ASN
3	L	141	PRO

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 20 such sidechains are listed below:

Mol	Chain	Res	Type
3	L	147	GLN
4	H	82(B)	ASN
4	H	200	HIS
4	H	199	ASN
1	G	355	ASN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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

### 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

### 5.6 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	$\# Z  > 2$
5	NAG	G	734	1	14,14,15	0.63	0	17,19,21	0.52	0
5	NAG	G	789	1	14,14,15	0.65	0	17,19,21	0.88	0
5	NAG	G	894	1	14,14,15	0.66	0	17,19,21	0.74	1 (5%)
5	NAG	G	697	1	14,14,15	0.65	0	17,19,21	0.64	0
5	NAG	G	588	1	14,14,15	0.63	0	17,19,21	0.79	0
5	NAG	G	776	1	14,14,15	0.59	0	17,19,21	0.78	0
5	NAG	G	886	1	14,14,15	0.66	0	17,19,21	1.04	2 (11%)
5	NAG	G	948	1	14,14,15	0.89	1 (7%)	17,19,21	0.95	1 (5%)
5	NAG	G	856	1	14,14,15	0.77	0	17,19,21	0.71	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
5	NAG	G	741	1	14,14,15	0.56	0	17,19,21	0.60	0
5	NAG	G	963	1	14,14,15	0.75	0	17,19,21	0.66	0
5	NAG	G	795	1	14,14,15	0.52	0	17,19,21	0.74	0
5	NAG	G	762	1	14,14,15	0.62	0	17,19,21	0.73	1 (5%)
5	NAG	G	908	1	14,14,15	0.62	0	17,19,21	0.60	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	NAG	G	734	1	-	5/6/23/26	0/1/1/1
5	NAG	G	789	1	-	4/6/23/26	0/1/1/1
5	NAG	G	894	1	-	4/6/23/26	0/1/1/1
5	NAG	G	697	1	-	5/6/23/26	0/1/1/1
5	NAG	G	588	1	1/1/5/7	4/6/23/26	0/1/1/1
5	NAG	G	776	1	-	2/6/23/26	0/1/1/1
5	NAG	G	886	1	-	4/6/23/26	0/1/1/1
5	NAG	G	948	1	-	4/6/23/26	0/1/1/1
5	NAG	G	856	1	-	4/6/23/26	0/1/1/1
5	NAG	G	741	1	1/1/5/7	4/6/23/26	0/1/1/1
5	NAG	G	963	1	-	5/6/23/26	0/1/1/1
5	NAG	G	795	1	-	2/6/23/26	0/1/1/1
5	NAG	G	762	1	-	6/6/23/26	0/1/1/1
5	NAG	G	908	1	-	2/6/23/26	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	G	948	NAG	C1-C2	2.55	1.56	1.52

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	G	886	NAG	C2-N2-C7	-2.69	119.08	122.90
5	G	762	NAG	C2-N2-C7	-2.17	119.81	122.90
5	G	894	NAG	C2-N2-C7	-2.13	119.86	122.90

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	G	886	NAG	C4-C3-C2	-2.07	107.98	111.02
5	G	948	NAG	C2-N2-C7	-2.02	120.02	122.90

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
5	G	588	NAG	C1
5	G	741	NAG	C1

5 of 55 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	G	588	NAG	C8-C7-N2-C2
5	G	588	NAG	O7-C7-N2-C2
5	G	697	NAG	C1-C2-N2-C7
5	G	697	NAG	C8-C7-N2-C2
5	G	697	NAG	O7-C7-N2-C2

There are no ring outliers.

9 monomers are involved in 22 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	G	789	NAG	5	0
5	G	894	NAG	3	0
5	G	776	NAG	2	0
5	G	948	NAG	3	0
5	G	856	NAG	1	0
5	G	741	NAG	1	0
5	G	963	NAG	2	0
5	G	762	NAG	1	0
5	G	908	NAG	4	0

## 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	G	306/313 (97%)	-0.34	10 (3%) 46 41	22, 52, 102, 125	1 (0%)
2	C	181/185 (97%)	-0.05	2 (1%) 80 80	22, 80, 113, 123	0
3	L	214/214 (100%)	-0.24	2 (0%) 84 84	36, 73, 102, 111	0
4	H	229/229 (100%)	-0.19	8 (3%) 44 38	24, 57, 123, 134	0
All	All	930/941 (98%)	-0.22	22 (2%) 59 56	22, 64, 114, 134	1 (0%)

The worst 5 of 22 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	H	132	SER	5.9
1	G	398	LYS	3.9
1	G	410	THR	3.6
4	H	133	GLY	3.5
1	G	461	ASP	3.4

### 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	G	908	14/15	0.47	0.42	121,123,124,124	0
5	NAG	G	894	14/15	0.76	0.30	99,101,103,103	0
5	NAG	G	588	14/15	0.80	0.40	132,133,134,134	0
5	NAG	G	948	14/15	0.82	0.39	70,73,75,76	0
5	NAG	G	963	14/15	0.83	0.47	122,124,125,125	0
5	NAG	G	741	14/15	0.84	0.35	98,101,102,102	0
5	NAG	G	697	14/15	0.85	0.33	95,98,99,100	0
5	NAG	G	856	14/15	0.87	0.33	97,100,101,102	0
5	NAG	G	776	14/15	0.90	0.19	66,69,72,73	0
5	NAG	G	886	14/15	0.90	0.15	46,52,65,66	0
5	NAG	G	734	14/15	0.91	0.31	98,99,99,99	0
5	NAG	G	789	14/15	0.92	0.18	61,65,71,73	0
5	NAG	G	762	14/15	0.96	0.12	50,56,58,59	0
5	NAG	G	795	14/15	0.96	0.13	31,34,38,44	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.