



# Full wwPDB X-ray Structure Validation Report ⓘ

Nov 22, 2023 – 09:50 PM JST

PDB ID : 7XIL  
Title : SARS-CoV-2-Beta-RBD and B38-GWP/P-VK antibody complex  
Authors : Shi, R.; Wang, Y.; Wu, Z.; Han, X.; Yan, J.  
Deposited on : 2022-04-13  
Resolution : 2.91 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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 : 2.36  
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) : 2.36

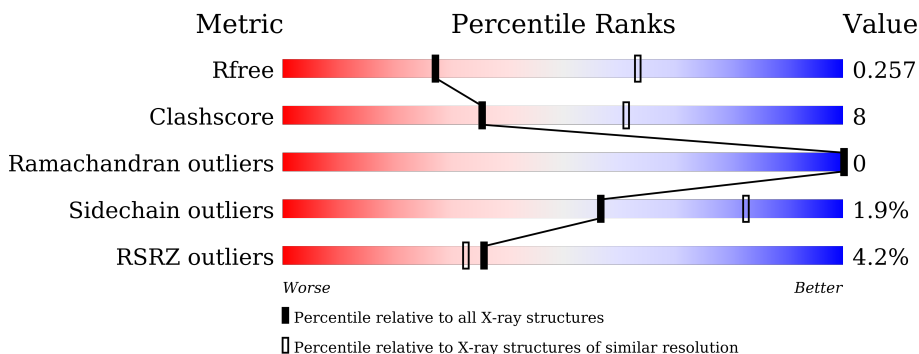
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.91 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	2307 (2.94-2.90)
Clashscore	141614	2531 (2.94-2.90)
Ramachandran outliers	138981	2462 (2.94-2.90)
Sidechain outliers	138945	2464 (2.94-2.90)
RSRZ outliers	127900	2248 (2.94-2.90)

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	A	227	 7% 70% 15% 15%
2	H	222	 % 78% 18% 5%
3	L	219	 3% 78% 21% .

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 4790 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 Spike protein S1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	193	1528	981	253	286	8	0	0	0

There are 11 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	417	ASN	LYS	variant	UNP P0DTC2
A	484	LYS	GLU	variant	UNP P0DTC2
A	501	TYR	ASN	variant	UNP P0DTC2
A	538	HIS	-	expression tag	UNP P0DTC2
A	539	HIS	-	expression tag	UNP P0DTC2
A	540	HIS	-	expression tag	UNP P0DTC2
A	541	HIS	-	expression tag	UNP P0DTC2
A	542	HIS	-	expression tag	UNP P0DTC2
A	543	HIS	-	expression tag	UNP P0DTC2
A	544	HIS	-	expression tag	UNP P0DTC2
A	545	HIS	-	expression tag	UNP P0DTC2

- Molecule 2 is a protein called B38 Fab heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	H	212	1579	998	267	307	7	0	0	0

- Molecule 3 is a protein called B38 Fab light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	L	217	1669	1050	275	337	7	3	1	0

- Molecule 4 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>).

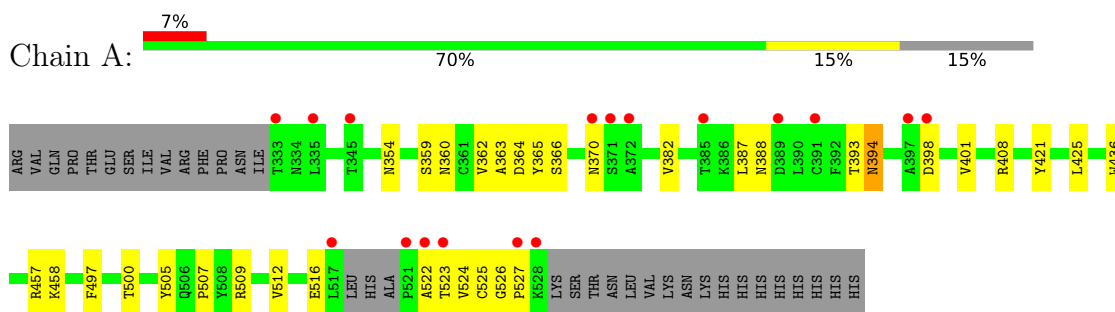


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
4	A	1	14	8	1	5	0	0

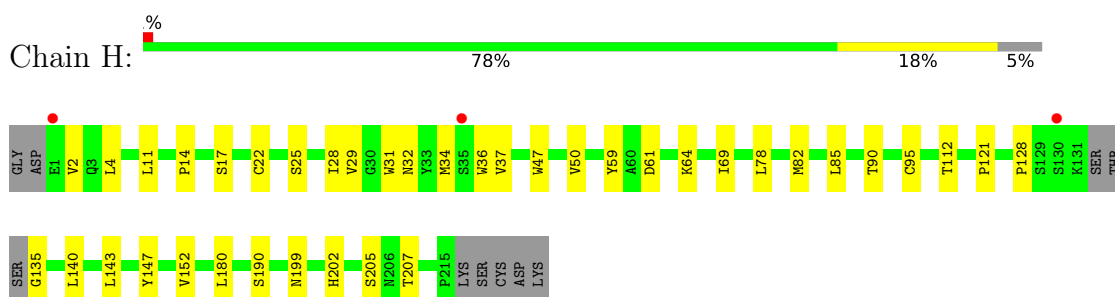
### 3 Residue-property plots [i](#)

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.

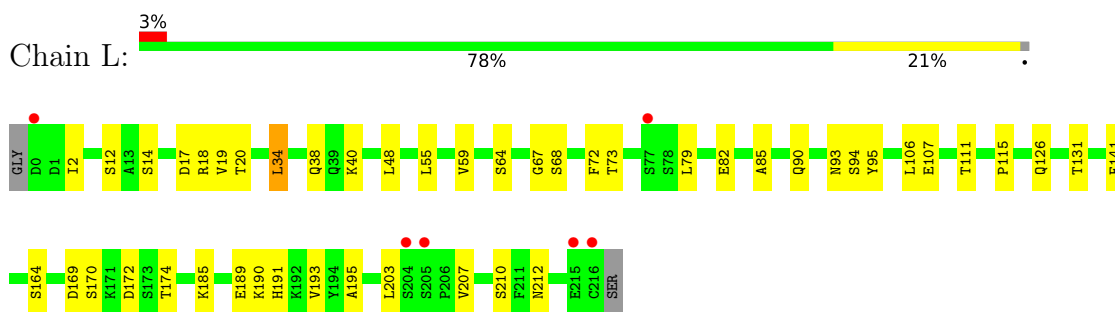
- Molecule 1: Spike protein S1



- Molecule 2: B38 Fab heavy chain



- Molecule 3: B38 Fab light chain



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	194.72Å 84.10Å 56.86Å 90.00° 100.19° 90.00°	Depositor
Resolution (Å)	27.74 – 2.91 27.74 – 2.91	Depositor EDS
% Data completeness (in resolution range)	96.3 (27.74-2.91) 96.3 (27.74-2.91)	Depositor EDS
$R_{merge}$	0.19	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.40 (at 2.90Å)	Xtrriage
Refinement program	PHENIX 1.20.1-4487	Depositor
R, $R_{free}$	0.206 , 0.256 0.207 , 0.257	Depositor DCC
$R_{free}$ test set	1922 reflections (9.99%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	54.5	Xtrriage
Anisotropy	0.393	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 41.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	4790	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	51.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.91% 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	A	0.30	0/1571	0.52	0/2137
2	H	0.35	0/1618	0.54	0/2207
3	L	0.27	0/1710	0.56	1/2324 (0.0%)
All	All	0.31	0/4899	0.54	1/6668 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	L	34	LEU	CB-CG-CD2	-5.29	102.01	111.00

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	A	1528	0	1441	28	0
2	H	1579	0	1538	24	3
3	L	1669	0	1625	28	1
4	A	14	0	13	0	0
All	All	4790	0	4617	76	3

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 (76) 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:4:LEU:HD22	2:H:22:CYS:SG	2.13	0.89
2:H:4:LEU:HD22	2:H:95:CYS:SG	2.17	0.85
3:L:79:LEU:HD21	3:L:106:LEU:HD21	1.58	0.83
1:A:360:ASN:OD1	1:A:523:THR:HG21	1.77	0.82
2:H:4:LEU:CD2	2:H:22:CYS:SG	2.72	0.78
2:H:2:VAL:HA	2:H:25:SER:O	1.83	0.78
1:A:388:ASN:HB2	1:A:527:PRO:HD2	1.68	0.76
1:A:382:VAL:HG21	1:A:387:LEU:HD13	1.68	0.74
1:A:360:ASN:OD1	1:A:523:THR:CG2	2.37	0.72
2:H:90:THR:HG23	2:H:112:THR:HA	1.73	0.71
1:A:360:ASN:HA	1:A:523:THR:HG22	1.74	0.69
3:L:34:LEU:HA	3:L:90:GLN:O	1.96	0.65
2:H:28:ILE:O	2:H:32:ASN:ND2	2.33	0.61
2:H:2:VAL:O	2:H:2:VAL:HG13	2.01	0.60
3:L:34:LEU:HD21	3:L:72:PHE:CG	2.36	0.60
2:H:128:PRO:HG3	2:H:140:LEU:HB3	1.84	0.60
3:L:115:PRO:HB3	3:L:141:PHE:HB3	1.84	0.60
3:L:169:ASP:OD2	3:L:170:SER:N	2.37	0.58
3:L:40:LYS:HD2	3:L:85:ALA:HB2	1.88	0.56
3:L:185:LYS:O	3:L:189:GLU:HG3	2.05	0.55
3:L:190:LYS:HE3	3:L:191:HIS:NE2	2.21	0.55
2:H:37:VAL:HG22	2:H:47:TRP:HA	1.88	0.55
1:A:393:THR:HA	1:A:522:ALA:HA	1.89	0.55
1:A:364:ASP:OD2	1:A:365:TYR:N	2.40	0.54
2:H:36:TRP:HD1	2:H:69:ILE:HD12	1.72	0.54
1:A:425:LEU:HD21	1:A:512:VAL:HG11	1.91	0.53
1:A:354:ASN:O	1:A:398:ASP:HA	2.08	0.53
1:A:366:SER:O	1:A:370:ASN:HB2	2.09	0.53
2:H:29:VAL:HG13	2:H:34:MET:HG3	1.91	0.53
1:A:359:SER:OG	1:A:394:ASN:HB3	2.08	0.52
1:A:497:PHE:CG	1:A:507:PRO:HG3	2.44	0.52
3:L:195:ALA:HB2	3:L:210:SER:HB3	1.90	0.52
1:A:505:TYR:OH	3:L:2:ILE:HD11	2.09	0.51
3:L:34:LEU:HD21	3:L:72:PHE:CD1	2.45	0.51
3:L:38:GLN:HB2	3:L:48:LEU:HD11	1.93	0.50
1:A:362:VAL:HG23	1:A:525:CYS:O	2.12	0.50
2:H:2:VAL:CA	2:H:25:SER:O	2.58	0.49
1:A:401:VAL:HG22	1:A:509:ARG:HG2	1.95	0.48

*Continued on next page...*



Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:L:203:LEU:HD13	3:L:207:VAL:HG12	1.95	0.48
1:A:363:ALA:O	1:A:527:PRO:HD3	2.14	0.47
1:A:362:VAL:HA	1:A:525:CYS:O	2.14	0.47
3:L:93:ASN:O	3:L:94:SER:OG	2.32	0.47
1:A:362:VAL:CG2	1:A:526:GLY:C	2.82	0.47
2:H:59:TYR:O	2:H:64:LYS:NZ	2.47	0.47
1:A:365:TYR:CD2	1:A:387:LEU:HG	2.51	0.46
2:H:121:PRO:HB3	2:H:147:TYR:HB3	1.97	0.46
1:A:365:TYR:HB2	1:A:388:ASN:HB3	1.96	0.46
1:A:388:ASN:OD1	1:A:388:ASN:N	2.46	0.46
1:A:359:SER:HA	1:A:524:VAL:HG13	1.97	0.46
2:H:61:ASP:HA	2:H:64:LYS:HE2	1.98	0.45
3:L:93:ASN:C	3:L:95:TYR:H	2.18	0.45
3:L:172:ASP:OD1	3:L:174:THR:OG1	2.27	0.45
1:A:408:ARG:NH2	3:L:95:TYR:CZ	2.85	0.45
2:H:47:TRP:HZ2	2:H:50:VAL:HG12	1.82	0.45
2:H:152:VAL:HG12	2:H:202:HIS:CD2	2.52	0.45
2:H:17:SER:HA	2:H:82:MET:O	2.17	0.45
3:L:34:LEU:HD21	3:L:72:PHE:CD2	2.52	0.44
3:L:12:SER:OG	3:L:107:GLU:OE2	2.35	0.44
3:L:193:VAL:HG22	3:L:212:ASN:OD1	2.18	0.43
1:A:364:ASP:C	1:A:366:SER:H	2.20	0.43
3:L:55:LEU:CD2	3:L:59:VAL:HB	2.48	0.43
2:H:69:ILE:HD11	2:H:78:LEU:HD11	1.99	0.43
3:L:20:THR:HG23	3:L:73:THR:HG23	2.00	0.43
1:A:421:TYR:CD1	1:A:457:ARG:HB3	2.53	0.43
3:L:126:GLN:HG2	3:L:131:THR:O	2.18	0.42
2:H:143:LEU:HD12	2:H:180:LEU:O	2.19	0.42
3:L:67:GLY:HA3	3:L:72:PHE:HA	2.02	0.42
3:L:14:SER:O	3:L:17:ASP:HB2	2.19	0.42
2:H:202:HIS:ND1	2:H:205:SER:OG	2.46	0.41
3:L:19:VAL:HG21	3:L:79:LEU:HD22	2.01	0.41
3:L:40:LYS:NZ	3:L:82:GLU:O	2.31	0.41
1:A:408:ARG:NH2	3:L:95:TYR:OH	2.53	0.41
1:A:458:LYS:HD3	2:H:31:TRP:CD1	2.57	0.40
2:H:14:PRO:HA	2:H:85:LEU:O	2.21	0.40
2:H:147:TYR:CE1	2:H:152:VAL:HG13	2.57	0.40
1:A:436:TRP:CE2	1:A:509:ARG:HB2	2.57	0.40

All (3) 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)
2:H:207:THR:CG2	3:L:18:ARG:NH1[1_554]	1.92	0.28
2:H:135:GLY:N	2:H:135:GLY:N[2_555]	2.03	0.17
2:H:135:GLY:N	2:H:135:GLY:O[2_555]	2.04	0.16

### 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	189/227 (83%)	179 (95%)	10 (5%)	0	100	100
2	H	208/222 (94%)	200 (96%)	8 (4%)	0	100	100
3	L	216/219 (99%)	211 (98%)	5 (2%)	0	100	100
All	All	613/668 (92%)	590 (96%)	23 (4%)	0	100	100

There are no Ramachandran outliers to report.

#### 5.3.2 Protein sidechains [i](#)

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	166/200 (83%)	163 (98%)	3 (2%)	59	83
2	H	174/184 (95%)	171 (98%)	3 (2%)	60	84
3	L	191/191 (100%)	187 (98%)	4 (2%)	53	80
All	All	531/575 (92%)	521 (98%)	10 (2%)	57	83

All (10) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	394	ASN
1	A	500	THR
1	A	516	GLU
2	H	11	LEU
2	H	190	SER
2	H	199	ASN
3	L	64	SER
3	L	68	SER
3	L	111	THR
3	L	164	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	474	GLN
2	H	194	GLN

### 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](#)

1 ligand 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
4	NAG	A	601	1	14,14,15	0.37	0	17,19,21	0.65	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
4	NAG	A	601	1	-	1/6/23/26	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(°)
4	A	601	NAG	C1-O5-C5	2.36	115.38	112.19

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	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	193/227 (85%)	0.24	17 (8%) 10 7	35, 51, 96, 103	0
2	H	212/222 (95%)	-0.07	3 (1%) 75 76	34, 45, 62, 82	0
3	L	217/219 (99%)	-0.03	6 (2%) 53 50	32, 48, 64, 93	0
All	All	622/668 (93%)	0.04	26 (4%) 36 33	32, 48, 83, 103	0

All (26) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	L	216	CYS	6.6
1	A	333	THR	6.2
1	A	523	THR	3.6
1	A	528	LYS	3.6
1	A	521	PRO	3.6
3	L	0	ASP	3.2
1	A	522	ALA	2.9
2	H	1	GLU	2.9
1	A	370	ASN	2.9
1	A	385	THR	2.7
1	A	335	LEU	2.7
1	A	372	ALA	2.7
1	A	527	PRO	2.4
1	A	345	THR	2.4
1	A	389	ASP	2.3
3	L	204	SER	2.3
2	H	130	SER	2.3
1	A	398	ASP	2.3
1	A	517	LEU	2.3
3	L	77	SER	2.2
1	A	371	SER	2.2
1	A	391	CYS	2.2
3	L	205	SER	2.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
3	L	215	GLU	2.1
2	H	35	SER	2.1
1	A	397	ALA	2.1

## 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(Å <sup>2</sup> )	Q<0.9
4	NAG	A	601	14/15	0.81	0.24	59,74,77,80	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.