



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 11, 2023 – 07:12 PM EDT

PDB ID : 7RBY  
Title : Crystal structure of Nanobody nb112 and SARS-CoV-2 RBD  
Authors : Chen, Y.; Tolbert, W.; Pazgier, M.  
Deposited on : 2021-07-06  
Resolution : 2.82 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 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 : 2.35.1  
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) : 2.35.1

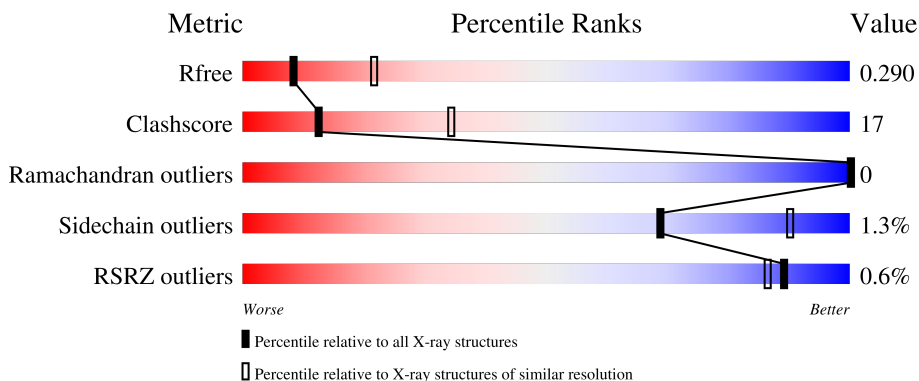
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.82 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	3617 (2.84-2.80)
Clashscore	141614	4060 (2.84-2.80)
Ramachandran outliers	138981	3978 (2.84-2.80)
Sidechain outliers	138945	3980 (2.84-2.80)
RSRZ outliers	127900	3552 (2.84-2.80)

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	210	 69% 24% 7%
1	C	210	 66% 26% 7%
2	B	130	 67% 31% .
2	D	130	 62% 38%

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 5098 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.

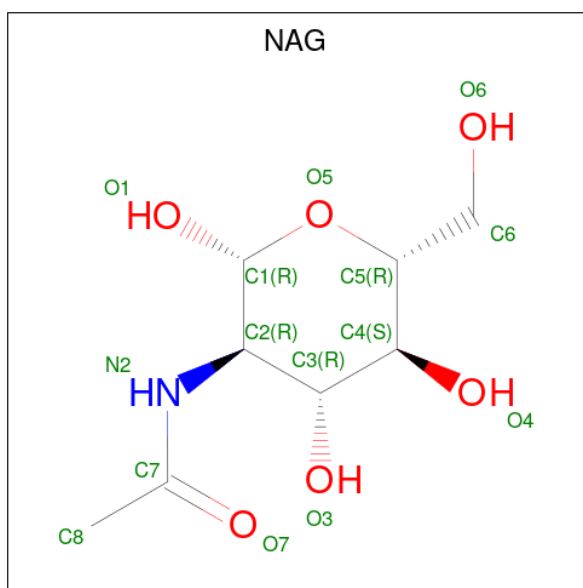
- Molecule 1 is a protein called Spike protein S1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	195	Total 1545	C 991	N 258	O 288	S 8	0	0	0
1	C	195	Total 1545	C 991	N 258	O 288	S 8	0	0	0

- Molecule 2 is a protein called llama-isolated nanobody NIH-CoV nb-112 specific to SARS-CoV-2 RBD.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	130	Total 986	C 616	N 163	O 201	S 6	0	0	0
2	D	130	Total 986	C 616	N 163	O 201	S 6	0	0	0

- Molecule 3 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
3	A	1	Total	C	N	O	0	0
			14	8	1	5		
3	C	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 4 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	C	1	Total	Mg	0	0
			1	1		

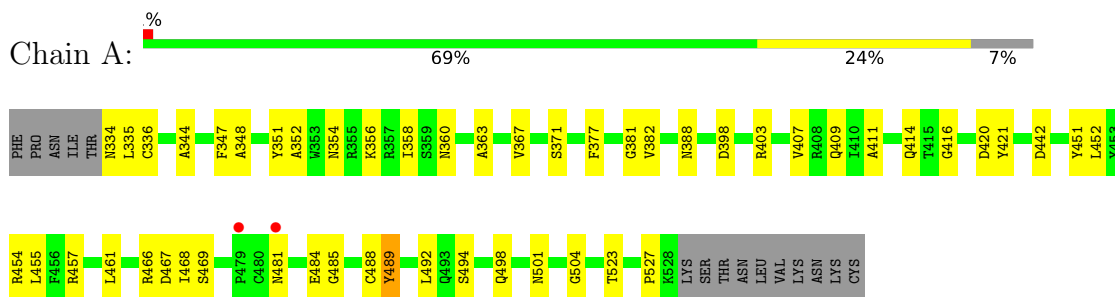
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	1	Total	O	0	0
			1	1		
5	B	1	Total	O	0	0
			1	1		
5	C	3	Total	O	0	0
			3	3		
5	D	2	Total	O	0	0
			2	2		

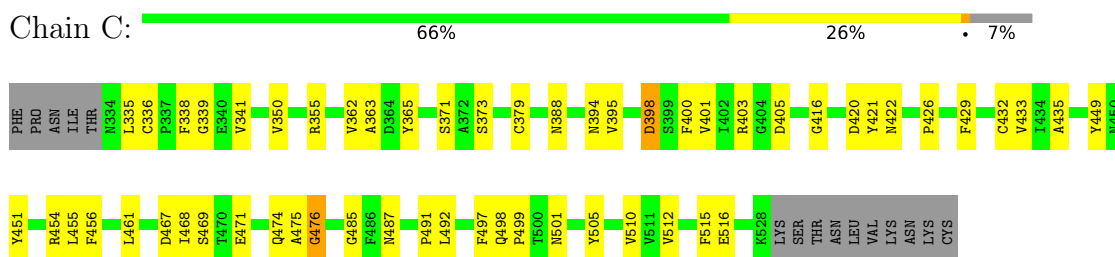
### 3 Residue-property plots

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.

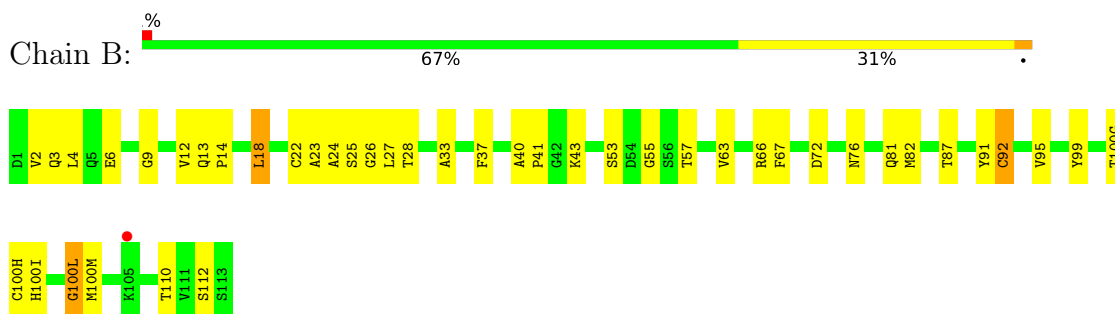
- Molecule 1: Spike protein S1



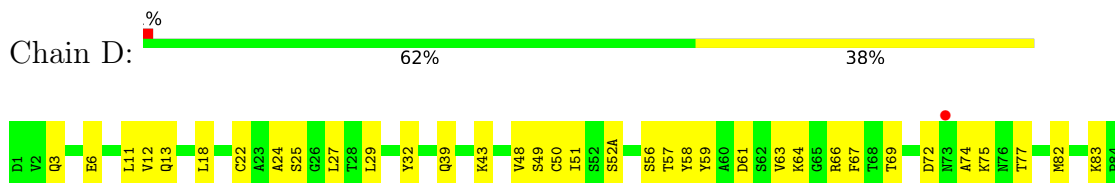
- Molecule 1: Spike protein S1



- Molecule 2: Ilama-isolated nanobody NIH-CoV nb-112 specific to SARS-CoV-2 RBD



- Molecule 2: Ilama-isolated nanobody NIH-CoV nb-112 specific to SARS-CoV-2 RBD



E85	D86	T87	Y90	Y95	F96	T100C	H100I	P100J	G100K	G104	T107	G108	V109	T110	V111	S112	S113
-----	-----	-----	-----	-----	-----	-------	-------	-------	-------	------	------	------	------	------	------	------	------

## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	32.70Å 59.00Å 216.88Å 90.00° 91.34° 90.00°	Depositor
Resolution (Å)	56.93 – 2.82 72.27 – 2.82	Depositor EDS
% Data completeness (in resolution range)	79.7 (56.93-2.82) 72.4 (72.27-2.82)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.15	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.86 (at 2.82Å)	Xtrriage
Refinement program	PHENIX 1.19.2_4158	Depositor
R, $R_{free}$	0.201 , 0.287 0.239 , 0.290	Depositor DCC
$R_{free}$ test set	1602 reflections (9.90%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	20.3	Xtrriage
Anisotropy	0.631	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , -3.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtrriage
Estimated twinning fraction	0.120 for h,-k,-l	Xtrriage
Reported twinning fraction	0.090 for h,-k,-l	Depositor
Outliers	1 of 16180 reflections (0.006%)	Xtrriage
$F_o, F_c$ correlation	0.84	EDS
Total number of atoms	5098	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	15.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.87% 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: MG, 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	1.08	0/1589	1.04	6/2162 (0.3%)
1	C	0.97	0/1589	1.02	5/2162 (0.2%)
2	B	1.12	0/1010	1.18	8/1372 (0.6%)
2	D	0.73	1/1010 (0.1%)	0.84	0/1372
All	All	1.00	1/5198 (0.0%)	1.03	19/7068 (0.3%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	22	CYS	CB-SG	-5.54	1.72	1.81

The worst 5 of 19 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	92	CYS	CA-CB-SG	7.99	128.38	114.00
1	C	476	GLY	N-CA-C	7.66	132.25	113.10
2	B	18	LEU	CA-CB-CG	6.95	131.29	115.30
1	A	371	SER	N-CA-C	6.80	129.36	111.00
2	B	100(L)	GLY	N-CA-C	6.76	130.00	113.10

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	A	1545	0	1465	32	0
1	C	1545	0	1465	39	0
2	B	986	0	922	33	0
2	D	986	0	922	65	0
3	A	14	0	13	1	0
3	C	14	0	13	0	0
4	C	1	0	0	0	0
5	A	1	0	0	0	0
5	B	1	0	0	0	0
5	C	3	0	0	0	0
5	D	2	0	0	0	0
All	All	5098	0	4800	166	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 17.

The worst 5 of 166 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:D:87:THR:CG2	2:D:111:VAL:H	1.44	1.28
2:D:67:PHE:CD1	2:D:82:MET:HG2	1.83	1.11
2:D:67:PHE:CE1	2:D:82:MET:HG2	1.88	1.09
2:D:87:THR:CG2	2:D:111:VAL:N	2.23	1.00
2:D:63:VAL:HG22	2:D:67:PHE:CD2	1.98	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	A	193/210 (92%)	187 (97%)	6 (3%)	0	100	100
1	C	193/210 (92%)	184 (95%)	9 (5%)	0	100	100

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	B	128/130 (98%)	124 (97%)	4 (3%)	0	100	100
2	D	128/130 (98%)	122 (95%)	6 (5%)	0	100	100
All	All	642/680 (94%)	617 (96%)	25 (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	168/183 (92%)	166 (99%)	2 (1%)	71	91
1	C	168/183 (92%)	167 (99%)	1 (1%)	86	95
2	B	105/105 (100%)	103 (98%)	2 (2%)	57	84
2	D	105/105 (100%)	103 (98%)	2 (2%)	57	84
All	All	546/576 (95%)	539 (99%)	7 (1%)	69	90

5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	99	TYR
1	C	456	PHE
2	D	107	THR
2	D	100(C)	THR
2	B	22	CYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
2	B	100(I)	HIS
1	C	474	GLN
2	D	100(I)	HIS

### 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](#)

Of 3 ligands modelled in this entry, 1 is monoatomic - leaving 2 for Mogul analysis.

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
3	NAG	C	1001	1	14,14,15	1.02	1 (7%)	17,19,21	0.51	0
3	NAG	A	1001	1	14,14,15	0.80	1 (7%)	17,19,21	0.52	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
3	NAG	C	1001	1	-	4/6/23/26	0/1/1/1
3	NAG	A	1001	1	-	4/6/23/26	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	1001	NAG	C1-C2	3.14	1.57	1.52

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	1001	NAG	C1-C2	2.34	1.55	1.52

There are no bond angle outliers.

There are no chirality outliers.

5 of 8 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	C	1001	NAG	O5-C5-C6-O6
3	A	1001	NAG	O5-C5-C6-O6
3	A	1001	NAG	C8-C7-N2-C2
3	A	1001	NAG	O7-C7-N2-C2
3	C	1001	NAG	C8-C7-N2-C2

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	1001	NAG	1	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	A	195/210 (92%)	-0.20	2 (1%) 82 77	5, 12, 35, 51	0
1	C	195/210 (92%)	-0.32	0 100 100	5, 10, 27, 51	0
2	B	130/130 (100%)	-0.10	1 (0%) 86 82	8, 16, 34, 46	0
2	D	130/130 (100%)	-0.21	1 (0%) 86 82	6, 14, 34, 51	0
All	All	650/680 (95%)	-0.22	4 (0%) 89 86	5, 13, 34, 51	0

All (4) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	481	ASN	3.4
1	A	479	PRO	2.2
2	B	105	LYS	2.1
2	D	73	ASN	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( $\text{\AA}^2$ )	Q<0.9
3	NAG	A	1001	14/15	0.81	0.29	41,50,63,65	0
3	NAG	C	1001	14/15	0.87	0.18	33,41,54,57	0
4	MG	C	1002	1/1	0.96	0.13	14,14,14,14	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.