



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 13, 2024 – 05:24 AM EDT

PDB ID : 2EIG  
Title : Lotus tetragonolobus seed lectin (Isoform)  
Authors : Moreno, F.B.M.B.; Vicoti, M.M.; Abrego, J.R.B.; de Oliveira, T.M.; Bezerra, G.A.; Cavada, B.S.; Filgueira de Azevedo Jr., W.  
Deposited on : 2007-03-13  
Resolution : 2.00 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	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

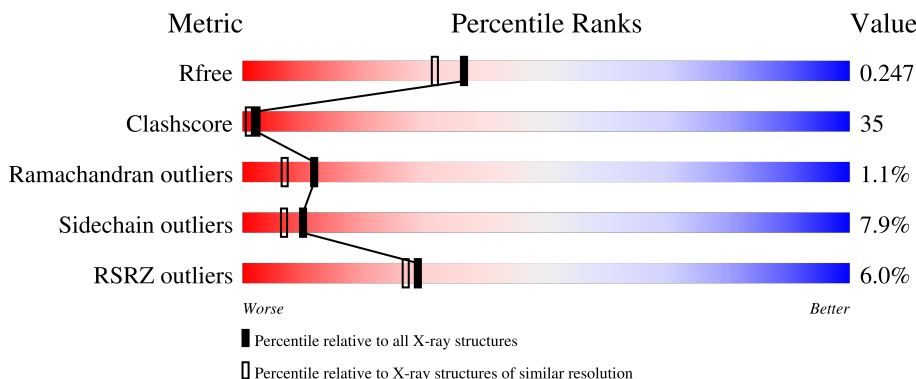
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.00 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))
$R_{free}$	164625	9409 (2.00-2.00)
Clashscore	180529	10737 (2.00-2.00)
Ramachandran outliers	177936	10628 (2.00-2.00)
Sidechain outliers	177891	10627 (2.00-2.00)
RSRZ outliers	164620	9409 (2.00-2.00)

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	234	<div> <div>5%</div> <div>61%</div> <div>28%</div> <div>9%</div> <div>..</div> </div>
1	B	234	<div> <div>4%</div> <div>64%</div> <div>27%</div> <div>6%</div> <div>..</div> </div>
1	C	234	<div> <div>7%</div> <div>60%</div> <div>32%</div> <div>..</div> <div>..</div> </div>
1	D	234	<div> <div>8%</div> <div>50%</div> <div>38%</div> <div>6%</div> <div>..</div> </div>

## 2 Entry composition [i](#)

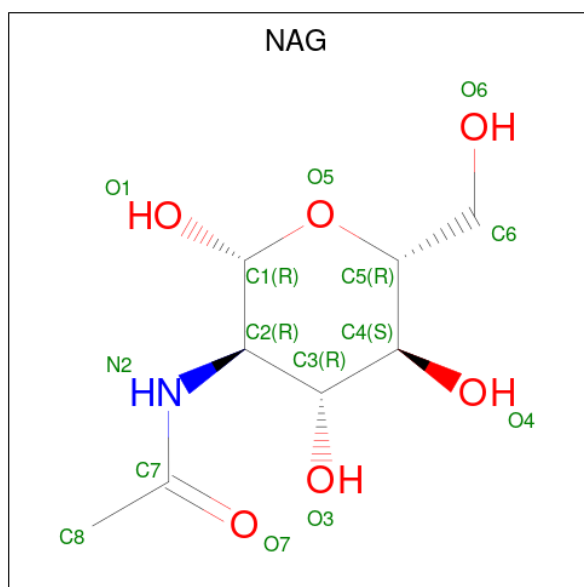
There are 5 unique types of molecules in this entry. The entry contains 7876 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 lectin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	230	Total	C	N	O	S	0	0	0
			1807	1151	301	352	3			
1	B	230	Total	C	N	O	S	0	0	0
			1807	1151	301	352	3			
1	C	230	Total	C	N	O	S	0	0	0
			1807	1151	301	352	3			
1	D	230	Total	C	N	O	S	0	0	0
			1807	1151	301	352	3			

- Molecule 2 is 2-acetamido-2-deoxy-beta-D-glucopyranose (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	B	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		

- Molecule 3 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total	Mn	0	0
			1	1		
3	B	1	Total	Mn	0	0
			1	1		
3	C	1	Total	Mn	0	0
			1	1		
3	D	1	Total	Mn	0	0
			1	1		

- Molecule 4 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	1	Total	Ca	0	0
			1	1		
4	B	1	Total	Ca	0	0
			1	1		
4	C	1	Total	Ca	0	0
			1	1		
4	D	1	Total	Ca	0	0
			1	1		

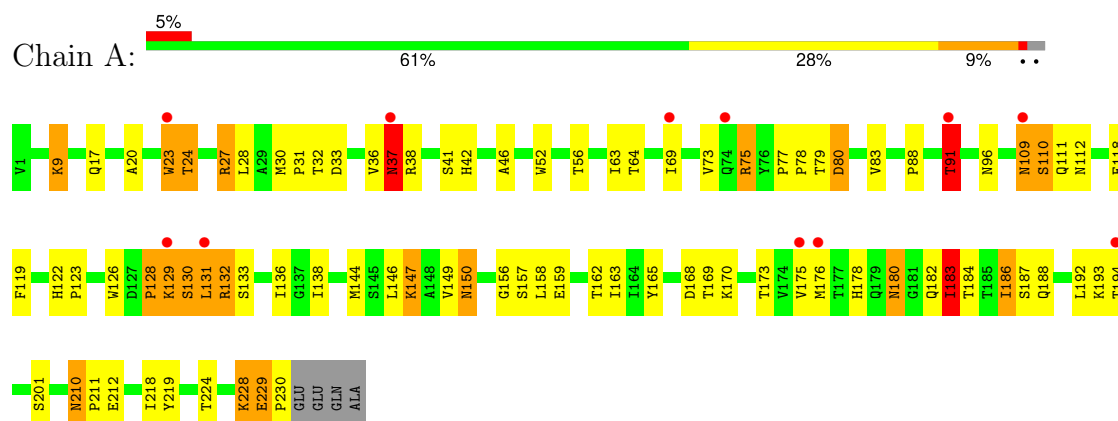
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	149	Total	O	0	0
			149	149		
5	B	181	Total	O	0	0
			181	181		
5	C	139	Total	O	0	0
			139	139		
5	D	129	Total	O	0	0
			129	129		

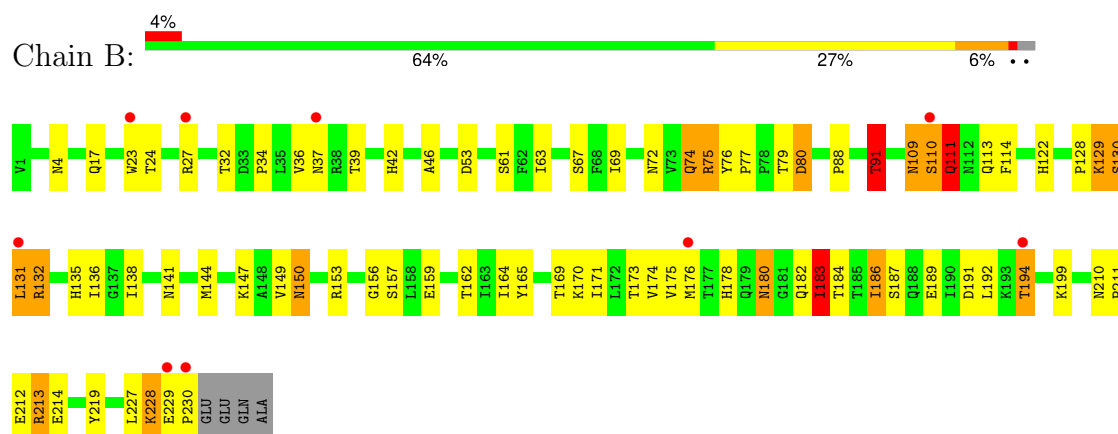
### 3 Residue-property plots [i](#)

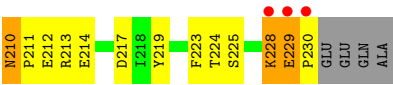
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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: lectin

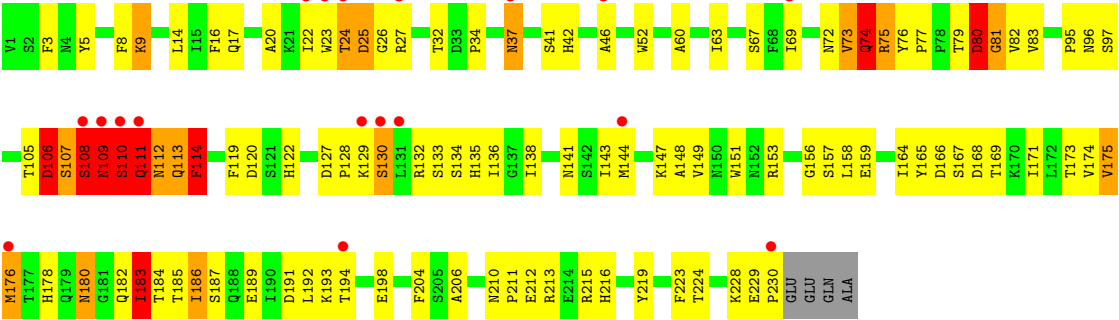


#### • Molecule 1: lectin





● Molecule 1: lectin



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	68.89Å 65.83Å 102.54Å 90.00° 92.00° 90.00°	Depositor
Resolution (Å)	34.46 – 2.00 34.46 – 2.00	Depositor EDS
% Data completeness (in resolution range)	99.4 (34.46-2.00) 99.4 (34.46-2.00)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	0.10	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.36 (at 2.00Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, $R_{free}$	0.188 , 0.251 0.187 , 0.247	Depositor DCC
$R_{free}$ test set	3128 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	23.0	Xtriage
Anisotropy	0.028	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 41.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.016 for k,h,-l 0.016 for -k,-h,-l 0.029 for h,-k,-l	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	7876	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	22.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.22% 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: MN, NAG, CA

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	0.94	1/1857 (0.1%)	0.99	6/2540 (0.2%)
1	B	1.06	2/1857 (0.1%)	1.04	8/2540 (0.3%)
1	C	0.93	1/1857 (0.1%)	1.00	7/2540 (0.3%)
1	D	0.87	0/1857	1.05	9/2540 (0.4%)
All	All	0.95	4/7428 (0.1%)	1.02	30/10160 (0.3%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1
1	B	1	2
1	D	0	9
All	All	1	12

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	91	THR	CA-CB	8.52	1.75	1.53
1	B	132	ARG	CD-NE	-7.63	1.33	1.46
1	C	91	THR	CA-CB	6.60	1.70	1.53
1	B	91	THR	CB-CG2	6.39	1.73	1.52

The worst 5 of 30 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	130	SER	N-CA-C	-8.41	88.30	111.00
1	C	176	MET	CB-CG-SD	-8.02	88.34	112.40

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	130	SER	CB-CA-C	6.76	122.95	110.10
1	C	183	ILE	CB-CA-C	-6.67	98.25	111.60
1	D	111	GLN	C-N-CA	6.66	138.35	121.70

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	B	110	SER	CA

5 of 12 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	73	VAL	Peptide
1	B	109	ASN	Peptide
1	B	110	SER	Peptide
1	D	73	VAL	Peptide
1	D	80	ASP	Peptide

## 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	1807	0	1748	139	0
1	B	1807	0	1749	135	0
1	C	1807	0	1749	134	0
1	D	1807	0	1749	163	0
2	A	14	0	13	0	0
2	B	14	0	13	2	0
2	C	14	0	13	1	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
3	D	1	0	0	0	0
4	A	1	0	0	0	0
4	B	1	0	0	0	0
4	C	1	0	0	0	0
4	D	1	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	A	149	0	0	13	0
5	B	181	0	0	19	0
5	C	139	0	0	4	0
5	D	129	0	0	15	0
All	All	7876	0	7034	502	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 35.

The worst 5 of 502 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:91:THR:CB	1:A:91:THR:CA	1.75	1.64
1:B:4:ASN:HD21	2:B:3001:NAG:C1	1.30	1.41
1:C:173:THR:HB	5:C:2131:HOH:O	1.31	1.27
1:B:91:THR:HB	5:B:3167:HOH:O	1.39	1.19
1:B:169:THR:HG23	1:C:147:LYS:HZ2	1.09	1.18

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	228/234 (97%)	211 (92%)	14 (6%)	3 (1%)	10	5
1	B	228/234 (97%)	212 (93%)	16 (7%)	0	100	100
1	C	228/234 (97%)	215 (94%)	12 (5%)	1 (0%)	30	27
1	D	228/234 (97%)	210 (92%)	12 (5%)	6 (3%)	4	1
All	All	912/936 (97%)	848 (93%)	54 (6%)	10 (1%)	12	7

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	37	ASN
1	A	110	SER
1	C	37	ASN
1	D	112	ASN
1	D	114	PHE

### 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	206/209 (99%)	186 (90%)	20 (10%)	6	4
1	B	206/209 (99%)	193 (94%)	13 (6%)	15	12
1	C	206/209 (99%)	189 (92%)	17 (8%)	9	6
1	D	206/209 (99%)	191 (93%)	15 (7%)	11	8
All	All	824/836 (99%)	759 (92%)	65 (8%)	10	7

5 of 65 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	110	SER
1	D	133	SER
1	B	111	GLN
1	B	91	THR
1	D	175	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 32 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	141	ASN
1	D	178	HIS
1	B	74	GLN
1	B	72	ASN
1	D	180	ASN

### 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 oligosaccharides in this entry.

## 5.6 Ligand geometry ⓘ

Of 11 ligands modelled in this entry, 8 are monoatomic - leaving 3 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	$\# Z  > 2$
2	NAG	C	2001	-	14,14,15	0.59	0	17,19,21	1.15	1 (5%)
2	NAG	A	1001	1	14,14,15	0.77	0	17,19,21	1.49	3 (17%)
2	NAG	B	3001	-	14,14,15	0.52	0	17,19,21	0.85	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	C	2001	-	-	4/6/23/26	0/1/1/1
2	NAG	A	1001	1	-	2/6/23/26	0/1/1/1
2	NAG	B	3001	-	-	1/6/23/26	0/1/1/1

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	1001	NAG	C4-C3-C2	3.95	116.80	111.02
2	A	1001	NAG	C1-O5-C5	2.66	115.76	112.19
2	A	1001	NAG	C1-C2-N2	-2.20	106.96	110.43
2	C	2001	NAG	O5-C5-C6	2.15	111.84	107.66

There are no chirality outliers.

5 of 7 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	C	2001	NAG	C8-C7-N2-C2
2	C	2001	NAG	O7-C7-N2-C2
2	C	2001	NAG	C4-C5-C6-O6
2	C	2001	NAG	O5-C5-C6-O6
2	A	1001	NAG	O5-C5-C6-O6

There are no ring outliers.

2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	C	2001	NAG	1	0
2	B	3001	NAG	2	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	230/234 (98%)	-0.19	11 (4%) 36 35	12, 20, 34, 44	0
1	B	230/234 (98%)	-0.32	9 (3%) 44 42	10, 16, 30, 49	0
1	C	230/234 (98%)	-0.13	17 (7%) 22 20	13, 20, 35, 51	0
1	D	230/234 (98%)	0.18	18 (7%) 20 19	15, 24, 40, 54	0
All	All	920/936 (98%)	-0.11	55 (5%) 29 27	10, 20, 36, 54	0

The worst 5 of 55 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	23	TRP	5.7
1	D	110	SER	5.5
1	A	23	TRP	4.9
1	D	129	LYS	4.5
1	A	131	LEU	4.3

### 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
2	NAG	C	2001	14/15	0.64	0.18	58,64,66,67	0
2	NAG	B	3001	14/15	0.77	0.16	51,55,57,58	0
2	NAG	A	1001	14/15	0.84	0.11	44,52,54,55	0
3	MN	A	1101	1/1	0.99	0.02	18,18,18,18	0
3	MN	B	1201	1/1	0.99	0.02	17,17,17,17	0
4	CA	A	1102	1/1	0.99	0.02	19,19,19,19	0
4	CA	B	1202	1/1	0.99	0.02	14,14,14,14	0
4	CA	C	1302	1/1	0.99	0.01	18,18,18,18	0
4	CA	D	1402	1/1	0.99	0.02	21,21,21,21	0
3	MN	C	1301	1/1	1.00	0.02	20,20,20,20	0
3	MN	D	1401	1/1	1.00	0.01	20,20,20,20	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.