



# Full wwPDB X-ray Structure Validation Report ⓘ

May 17, 2020 – 11:35 am BST

PDB ID : 5GR6  
Title : Crystal structure of branching enzyme Y500A/D501A double mutant from  
Cyanotheca sp. ATCC 51142  
Authors : Suzuki, R.; Suzuki, E.  
Deposited on : 2016-08-08  
Resolution : 2.00 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 : 2.11  
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) : 2.11

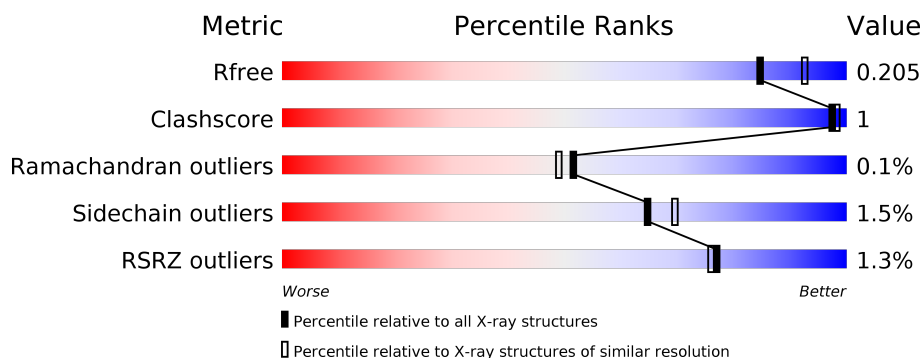
# 1 Overall quality at a glance

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))
$R_{free}$	130704	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (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 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	793	<div> <div></div> <div>88%</div> <div>7%</div> <div>5%</div> </div>

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 7043 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 1,4-alpha-glucan branching enzyme GlgB.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	755	Total	C	N	O	S	0	0	0
			6263	4050	1042	1147	24			

There are 22 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-19	MET	-	initiating methionine	UNP B1WPM8
A	-18	GLY	-	expression tag	UNP B1WPM8
A	-17	SER	-	expression tag	UNP B1WPM8
A	-16	SER	-	expression tag	UNP B1WPM8
A	-15	HIS	-	expression tag	UNP B1WPM8
A	-14	HIS	-	expression tag	UNP B1WPM8
A	-13	HIS	-	expression tag	UNP B1WPM8
A	-12	HIS	-	expression tag	UNP B1WPM8
A	-11	HIS	-	expression tag	UNP B1WPM8
A	-10	HIS	-	expression tag	UNP B1WPM8
A	-9	SER	-	expression tag	UNP B1WPM8
A	-8	SER	-	expression tag	UNP B1WPM8
A	-7	GLY	-	expression tag	UNP B1WPM8
A	-6	LEU	-	expression tag	UNP B1WPM8
A	-5	VAL	-	expression tag	UNP B1WPM8
A	-4	PRO	-	expression tag	UNP B1WPM8
A	-3	ARG	-	expression tag	UNP B1WPM8
A	-2	GLY	-	expression tag	UNP B1WPM8
A	-1	SER	-	expression tag	UNP B1WPM8
A	0	HIS	-	expression tag	UNP B1WPM8
A	500	ALA	TYR	engineered mutation	UNP B1WPM8
A	501	ALA	ASP	engineered mutation	UNP B1WPM8

- Molecule 2 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			6	3	3		
2	A	1	Total	C	O	0	0
			6	3	3		
2	A	1	Total	C	O	0	0
			6	3	3		
2	A	1	Total	C	O	0	0
			6	3	3		
2	A	1	Total	C	O	0	0
			6	3	3		

- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	4	Total	Mg	0	0
			4	4		

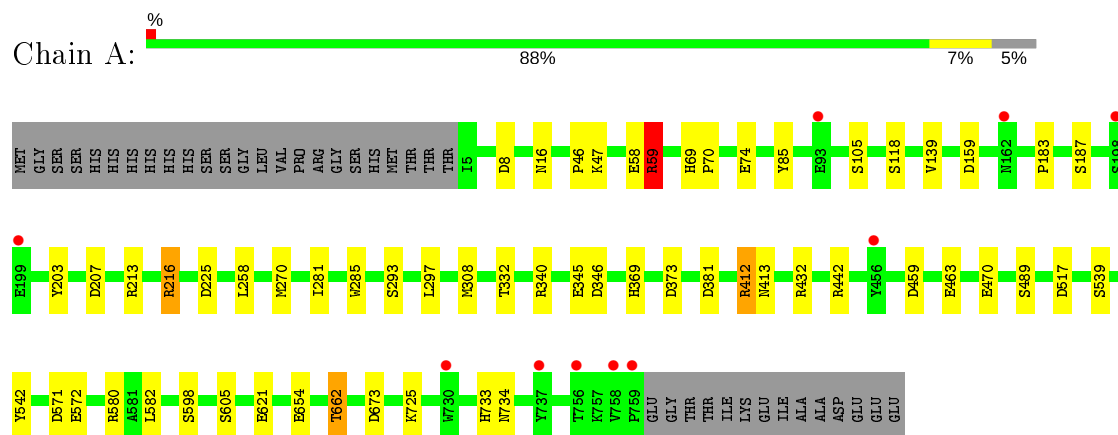
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	740	Total	O	0	0
			740	740		

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

These plots are drawn for all protein, RNA and DNA 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: 1,4-alpha-glucan branching enzyme GlgB



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	133.60 Å   133.60 Å   184.23 Å 90.00°   90.00°   90.00°	Depositor
Resolution (Å)	46.06 – 2.00 42.82 – 2.00	Depositor EDS
% Data completeness (in resolution range)	98.5 (46.06-2.00) 98.5 (42.82-2.00)	Depositor EDS
$R_{merge}$	0.04	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	9.73 (at 2.00 Å)	Xtriage
Refinement program	REFMAC 5.8.0049	Depositor
R, $R_{free}$	0.162   ,   0.194 0.174   ,   0.205	Depositor DCC
$R_{free}$ test set	5560 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	25.2	Xtriage
Anisotropy	0.013	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.39 , 49.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.35$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	7043	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	28.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.75% 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: GOL, MG

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	1.25	18/6482 (0.3%)	1.09	29/8821 (0.3%)

All (18) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	74	GLU	CD-OE2	9.06	1.35	1.25
1	A	285	TRP	CB-CG	-8.84	1.34	1.50
1	A	654	GLU	CD-OE2	-8.36	1.16	1.25
1	A	58	GLU	CD-OE2	8.26	1.34	1.25
1	A	74	GLU	CG-CD	7.29	1.62	1.51
1	A	605	SER	CB-OG	-7.03	1.33	1.42
1	A	85	TYR	CB-CG	-6.57	1.41	1.51
1	A	58	GLU	CG-CD	5.99	1.60	1.51
1	A	621	GLU	CD-OE2	-5.93	1.19	1.25
1	A	187	SER	CB-OG	-5.84	1.34	1.42
1	A	662	THR	CB-CG2	-5.80	1.33	1.52
1	A	572	GLU	CD-OE2	-5.67	1.19	1.25
1	A	74	GLU	CB-CG	-5.62	1.41	1.52
1	A	412	ARG	CZ-NH2	5.48	1.40	1.33
1	A	74	GLU	CD-OE1	-5.27	1.19	1.25
1	A	203	TYR	CE1-CZ	5.27	1.45	1.38
1	A	489	SER	CB-OG	5.15	1.49	1.42
1	A	159	ASP	CB-CG	5.01	1.62	1.51

All (29) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	412	ARG	NE-CZ-NH1	-9.87	115.36	120.30
1	A	459	ASP	CB-CG-OD1	9.37	126.73	118.30
1	A	74	GLU	CG-CD-OE1	-7.81	102.68	118.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	459	ASP	CB-CG-OD2	-7.74	111.33	118.30
1	A	571	ASP	CB-CG-OD1	7.38	124.94	118.30
1	A	74	GLU	CG-CD-OE2	7.28	132.85	118.30
1	A	381	ASP	CB-CG-OD1	6.80	124.42	118.30
1	A	8	ASP	CB-CG-OD1	6.71	124.34	118.30
1	A	225	ASP	CB-CG-OD2	-6.54	112.41	118.30
1	A	673	ASP	CB-CG-OD2	6.18	123.86	118.30
1	A	654	GLU	OE1-CD-OE2	-6.16	115.91	123.30
1	A	216	ARG	NE-CZ-NH1	6.08	123.34	120.30
1	A	463	GLU	OE1-CD-OE2	-5.84	116.29	123.30
1	A	225	ASP	CB-CG-OD1	5.78	123.50	118.30
1	A	442	ARG	NE-CZ-NH2	-5.70	117.45	120.30
1	A	308	MET	CG-SD-CE	5.59	109.14	100.20
1	A	582	LEU	CB-CG-CD1	5.56	120.45	111.00
1	A	118	SER	CA-CB-OG	-5.49	96.37	111.20
1	A	442	ARG	NE-CZ-NH1	5.45	123.03	120.30
1	A	345	GLU	OE1-CD-OE2	-5.42	116.80	123.30
1	A	580	ARG	NE-CZ-NH1	5.41	123.01	120.30
1	A	207	ASP	CB-CG-OD2	-5.37	113.47	118.30
1	A	346	ASP	CB-CG-OD1	-5.34	113.49	118.30
1	A	412	ARG	NE-CZ-NH2	5.28	122.94	120.30
1	A	59	ARG	NE-CZ-NH1	5.20	122.90	120.30
1	A	58	GLU	CG-CD-OE2	5.17	128.65	118.30
1	A	432	ARG	NE-CZ-NH2	-5.14	117.73	120.30
1	A	340	ARG	NE-CZ-NH1	5.12	122.86	120.30
1	A	517	ASP	CB-CG-OD2	-5.10	113.71	118.30

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	6263	0	5854	10	0
2	A	36	0	48	0	0
3	A	4	0	0	0	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	A	740	0	0	1	0
All	All	7043	0	5902	10	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 1.

All (10) 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:213:ARG:HH12	1:A:332:THR:HG1	1.53	0.54
1:A:412:ARG:HD2	1:A:470:GLU:OE1	2.15	0.47
1:A:69:HIS:CG	1:A:70:PRO:HD2	2.50	0.45
1:A:258:LEU:HD22	1:A:297:LEU:HD13	1.99	0.45
1:A:59:ARG:HG3	4:A:1333:HOH:O	2.17	0.44
1:A:270:MET:CG	1:A:293:SER:HB3	2.48	0.44
1:A:733:HIS:O	1:A:734:ASN:HB2	2.19	0.42
1:A:139:VAL:HG11	1:A:183:PRO:HB3	2.01	0.42
1:A:539:SER:HA	1:A:542:TYR:CE2	2.55	0.41
1:A:46:PRO:O	1:A:47:LYS:HB2	2.21	0.41

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	753/793 (95%)	733 (97%)	19 (2%)	1 (0%)	51 49

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	598	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	666/698 (95%)	656 (98%)	10 (2%)	65 69

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

Mol	Chain	Res	Type
1	A	16	ASN
1	A	59	ARG
1	A	105	SER
1	A	216	ARG
1	A	281	ILE
1	A	369	HIS
1	A	373	ASP
1	A	413	ASN
1	A	662	THR
1	A	725	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	16	ASN
1	A	65	ASN
1	A	634	ASN
1	A	638	GLN
1	A	688	GLN

### 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 carbohydrates in this entry.

## 5.6 Ligand geometry

Of 10 ligands modelled in this entry, 4 are monoatomic - leaving 6 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	GOL	A	805	-	5,5,5	0.33	0	5,5,5	1.02	0
2	GOL	A	806	-	5,5,5	0.96	0	5,5,5	1.70	2 (40%)
2	GOL	A	801	-	5,5,5	1.07	1 (20%)	5,5,5	1.35	1 (20%)
2	GOL	A	802	-	5,5,5	0.75	0	5,5,5	0.81	0
2	GOL	A	803	-	5,5,5	0.44	0	5,5,5	0.77	0
2	GOL	A	804	-	5,5,5	0.46	0	5,5,5	0.84	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	GOL	A	805	-	-	2/4/4/4	-
2	GOL	A	806	-	-	3/4/4/4	-
2	GOL	A	801	-	-	0/4/4/4	-
2	GOL	A	802	-	-	0/4/4/4	-
2	GOL	A	803	-	-	4/4/4/4	-
2	GOL	A	804	-	-	0/4/4/4	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	801	GOL	O2-C2	2.10	1.49	1.43

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	806	GOL	O3-C3-C2	2.54	122.40	110.20
2	A	801	GOL	C3-C2-C1	-2.43	102.25	111.70
2	A	806	GOL	O1-C1-C2	2.21	120.81	110.20

There are no chirality outliers.

All (9) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	803	GOL	O1-C1-C2-C3
2	A	803	GOL	C1-C2-C3-O3
2	A	806	GOL	O2-C2-C3-O3
2	A	803	GOL	O2-C2-C3-O3
2	A	806	GOL	O1-C1-C2-O2
2	A	805	GOL	O2-C2-C3-O3
2	A	803	GOL	O1-C1-C2-O2
2	A	805	GOL	C1-C2-C3-O3
2	A	806	GOL	C1-C2-C3-O3

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 [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	755/793 (95%)	-0.32	10 (1%) 77 76	16, 25, 46, 82	0

All (10) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	758	VAL	5.7
1	A	759	PRO	4.2
1	A	456	TYR	4.1
1	A	737	TYR	3.1
1	A	756	THR	2.9
1	A	730	TRP	2.8
1	A	93	GLU	2.4
1	A	199	GLU	2.4
1	A	198	SER	2.4
1	A	162	ASN	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 carbohydrates 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	GOL	A	804	6/6	0.88	0.21	42,58,59,62	0
2	GOL	A	801	6/6	0.91	0.20	32,34,42,47	0
3	MG	A	810	1/1	0.91	0.10	58,58,58,58	0
2	GOL	A	803	6/6	0.92	0.33	51,59,60,62	0
2	GOL	A	802	6/6	0.93	0.14	40,42,44,44	0
3	MG	A	807	1/1	0.96	0.09	60,60,60,60	0
2	GOL	A	805	6/6	0.96	0.10	29,39,46,54	0
3	MG	A	809	1/1	0.96	0.23	40,40,40,40	0
2	GOL	A	806	6/6	0.97	0.17	28,42,43,44	0
3	MG	A	808	1/1	0.97	0.04	25,25,25,25	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.