



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 05:34 AM GMT

PDB ID : 2R9K  
Title : Crystal Structure of Misteltoe Lectin I in Complex with Phloretamide  
Authors : Meyer, A.; Rypniewski, W.; Celewicz, L.; Erdmann, V.A.; Voelter, W.; Betzel, C.  
Deposited on : 2007-09-13  
Resolution : 2.70 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  
<http://wwpdb.org/validation/2016/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 : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

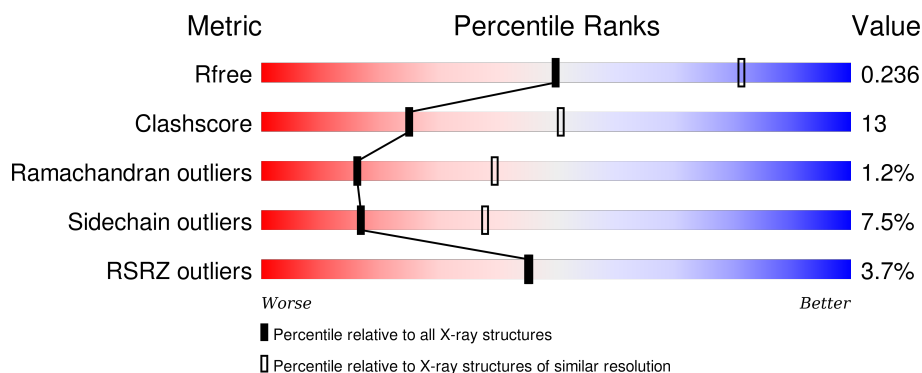
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.70 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))
$R_{free}$	91344	2103 (2.70-2.70)
Clashscore	102246	2422 (2.70-2.70)
Ramachandran outliers	100387	2382 (2.70-2.70)
Sidechain outliers	100360	2382 (2.70-2.70)
RSRZ outliers	91569	2107 (2.70-2.70)

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. 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	254	<div> <div>3%</div> <div>67%</div> <div>25%</div> <div>• • •</div> </div>
2	B	263	<div> <div>4%</div> <div>68%</div> <div>25%</div> <div>6% •</div> </div>

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
4	SO4	B	3	-	-	-	X
5	GOL	A	649	-	-	-	X
5	GOL	B	647	-	-	-	X
5	GOL	B	648	-	-	-	X

## 2 Entry composition

There are 9 unique types of molecules in this entry. The entry contains 4135 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 Beta-galactoside-specific lectin 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	248	Total	C	N	O	S	0	1	0
			1927	1218	330	375	4			

There are 29 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	8	THR	VAL	SEE REMARK 999	UNP P81446
A	9	ASP	THR	SEE REMARK 999	UNP P81446
A	10	GLN	HIS	SEE REMARK 999	UNP P81446
A	15	ALA	GLU	SEE REMARK 999	UNP P81446
A	19	SER	ARG	SEE REMARK 999	UNP P81446
A	23	VAL	LEU	SEE REMARK 999	UNP P81446
A	36	ASN	GLU	SEE REMARK 999	UNP P81446
A	45	VAL	ILE	SEE REMARK 999	UNP P81446
A	49	GLU	ASP	SEE REMARK 999	UNP P81446
A	50	GLY	ALA	SEE REMARK 999	UNP P81446
A	61	ALA	GLU	SEE REMARK 999	UNP P81446
A	65	THR	SER	SEE REMARK 999	UNP P81446
A	81	GLU	GLN	SEE REMARK 999	UNP P81446
A	84	ASN	ASP	SEE REMARK 999	UNP P81446
A	90	SER	ARG	SEE REMARK 999	UNP P81446
A	94	ALA	ARG	SEE REMARK 999	UNP P81446
A	99	GLN	HIS	SEE REMARK 999	UNP P81446
A	100	ASP	LEU	SEE REMARK 999	UNP P81446
A	102	SER	THR	SEE REMARK 999	UNP P81446
A	106	SER	ARG	SEE REMARK 999	UNP P81446
A	109	GLN	LEU	SEE REMARK 999	UNP P81446
A	148	GLN	SER	SEE REMARK 999	UNP P81446
A	150	LYS	ARG	SEE REMARK 999	UNP P81446
A	208	HIS	GLN	SEE REMARK 999	UNP P81446
A	219	ALA	ARG	SEE REMARK 999	UNP P81446
A	223	ALA	PRO	SEE REMARK 999	UNP P81446
A	226	VAL	ASN	SEE REMARK 999	UNP P81446

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	227	ILE	PHE	SEE REMARK 999	UNP P81446
A	233	ILE	VAL	SEE REMARK 999	UNP P81446

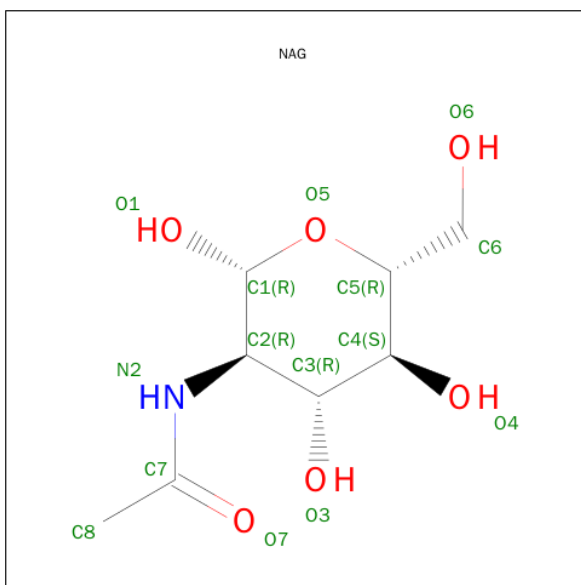
- Molecule 2 is a protein called Beta-galactoside-specific lectin 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	262	Total 1995	C 1239	N 352	O 392	S 12	0	0	0

There are 21 discrepancies between the modelled and reference sequences:

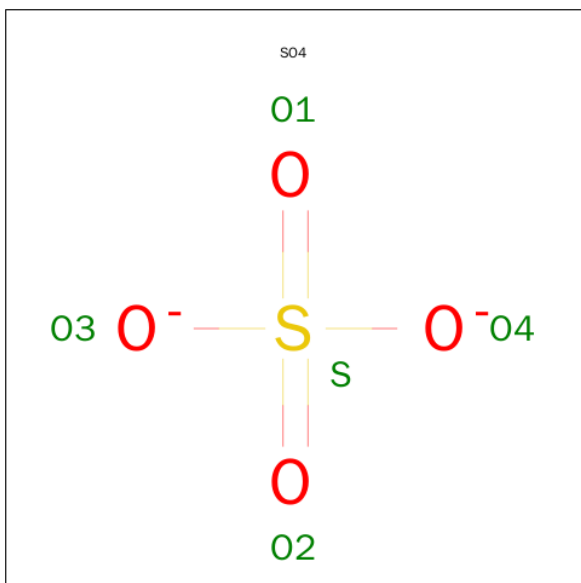
Chain	Residue	Modelled	Actual	Comment	Reference
B	249	ALA	ASP	SEE REMARK 999	UNP P81446
B	253	THR	SER	SEE REMARK 999	UNP P81446
B	258	ILE	THR	SEE REMARK 999	UNP P81446
B	268	THR	CYS	SEE REMARK 999	UNP P81446
B	301	LYS	ARG	SEE REMARK 999	UNP P81446
B	337	ILE	LEU	SEE REMARK 999	UNP P81446
B	339	GLN	GLU	SEE REMARK 999	UNP P81446
B	390	THR	VAL	SEE REMARK 999	UNP P81446
B	403	ALA	ASN	SEE REMARK 999	UNP P81446
B	408	TYR	TRP	SEE REMARK 999	UNP P81446
B	413	THR	VAL	SEE REMARK 999	UNP P81446
B	414	ALA	ILE	SEE REMARK 999	UNP P81446
B	415	GLY	SER	SEE REMARK 999	UNP P81446
B	417	GLU	GLN	SEE REMARK 999	UNP P81446
B	434	LEU	ASN	SEE REMARK 999	UNP P81446
B	436	SER	ASP	SEE REMARK 999	UNP P81446
B	441	ASN	CYS	SEE REMARK 999	UNP P81446
B	446	ILE	VAL	SEE REMARK 999	UNP P81446
B	489	SER	LYS	SEE REMARK 999	UNP P81446
B	491	GLN	ARG	SEE REMARK 999	UNP P81446
B	501	ASN	LYS	SEE REMARK 999	UNP P81446

- Molecule 3 is SUGAR (2,6-ANHYDRO-1-DEOXY-D-GALACTITOL) (three-letter code: NAG) (formula: C<sub>8</sub>H<sub>15</sub>NO<sub>6</sub>).



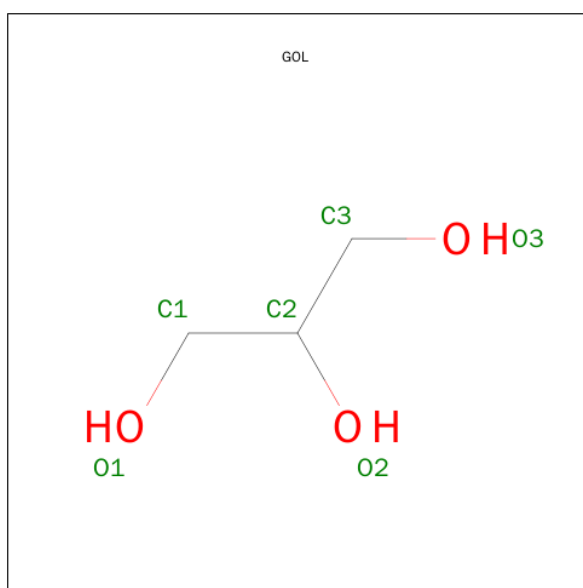
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			14	8	1	5		
3	B	1	Total	C	N	O	0	0
			14	8	1	5		
3	B	1	Total	C	N	O	0	0
			14	8	1	5		
3	B	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 4 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total O S 5 4 1	0	0
4	A	1	Total O S 5 4 1	0	0
4	A	1	Total O S 5 4 1	0	0
4	B	1	Total O S 5 4 1	0	0

- Molecule 5 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).

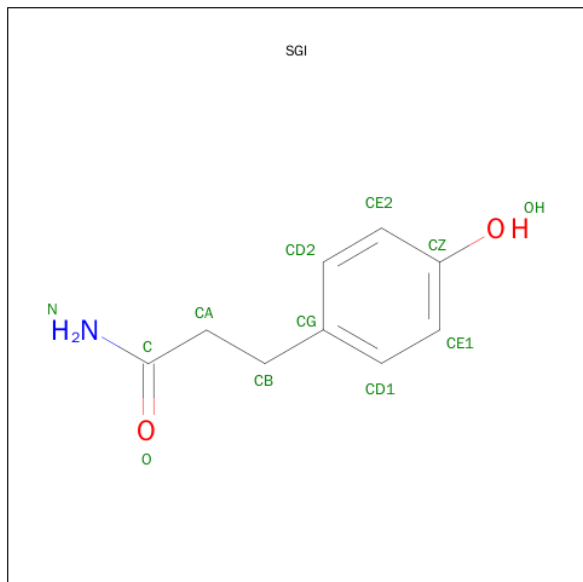


Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total C O 6 3 3	0	0
5	A	1	Total C O 6 3 3	0	0
5	B	1	Total C O 6 3 3	0	0
5	B	1	Total C O 6 3 3	0	0

- Molecule 6 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	1	Total Cl 1 1	0	0

- Molecule 7 is 3-(4-HYDROXYPHENYL)PROPANAMIDE (three-letter code: SGI) (formula: C<sub>9</sub>H<sub>11</sub>NO<sub>2</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
7	A	1	Total	C	N	O	0	0
			12	9	1	2		

- Molecule 8 is a polymer of unknown type called SUGAR (N-ACETYL-D-GLUCOSAMINE).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
8	B	2	Total	C	N	O	0	0
			28	16	2	10		

- Molecule 9 is water.

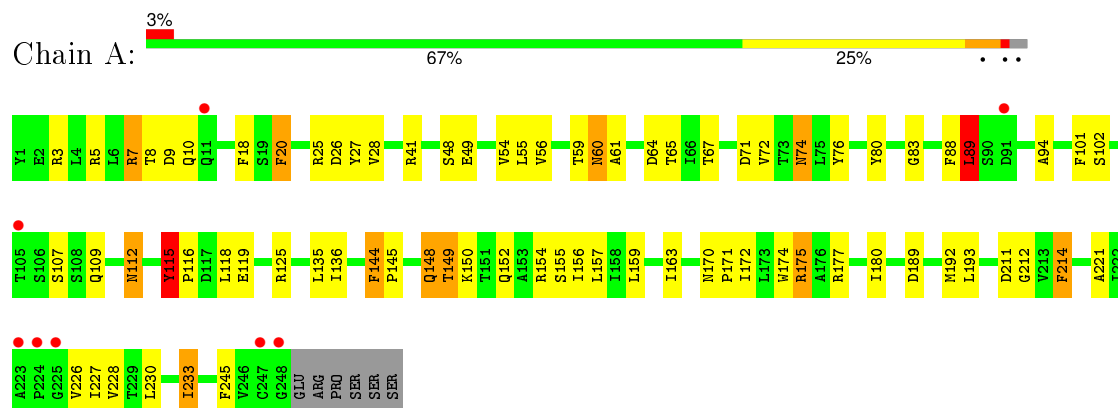
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	A	27	Total	O	0	0
			27	27		
9	B	45	Total	O	0	0
			45	45		



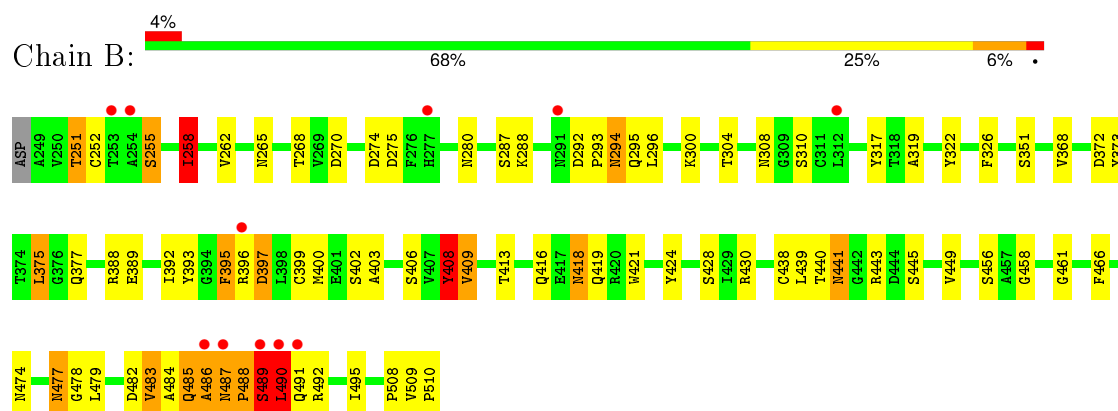
### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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 ( $\text{RSRZ} > 2$ ). 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: Beta-galactoside-specific lectin 1



#### • Molecule 2: Beta-galactoside-specific lectin 1



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 65 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	107.06Å 107.06Å 312.38Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	19.74 – 2.70 19.93 – 2.70	Depositor EDS
% Data completeness (in resolution range)	100.0 (19.74-2.70) 99.9 (19.93-2.70)	Depositor EDS
$R_{merge}$	0.60	Depositor
$R_{sym}$	0.47	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.55 (at 2.71Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.226 , 0.267 0.225 , 0.236	Depositor DCC
$R_{free}$ test set	1518 reflections (5.33%)	DCC
Wilson B-factor (Å <sup>2</sup> )	52.8	Xtriage
Anisotropy	0.051	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 33.9	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 30114 reflections	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	4135	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	46.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.31% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: SGI, GOL, SO4, NAG, CL

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	0.50	0/1965	0.65	0/2678
2	B	0.52	0/2034	0.68	0/2774
All	All	0.51	0/3999	0.66	0/5452

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	33
2	B	0	37
All	All	0	70

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (70) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	107	SER	Mainchain
1	A	115	TYR	Sidechain
1	A	116	PRO	Mainchain
1	A	118	LEU	Mainchain
1	A	135	LEU	Mainchain
1	A	144	PHE	Sidechain
1	A	154	ARG	Sidechain
1	A	156	ILE	Mainchain
1	A	171	PRO	Mainchain

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Group
1	A	174	TRP	Mainchain
1	A	175[A]	ARG	Sidechain
1	A	177	ARG	Mainchain
1	A	18	PHE	Sidechain
1	A	180	ILE	Mainchain
1	A	20	PHE	Sidechain
1	A	211	ASP	Mainchain
1	A	214	PHE	Sidechain
1	A	221	ALA	Mainchain
1	A	227	ILE	Mainchain
1	A	245	PHE	Sidechain
1	A	26	ASP	Mainchain
1	A	27	TYR	Sidechain
1	A	3	ARG	Mainchain
1	A	5	ARG	Sidechain
1	A	54	VAL	Mainchain
1	A	61	ALA	Mainchain
1	A	65	THR	Mainchain
1	A	7	ARG	Mainchain
1	A	83	GLY	Mainchain
1	A	88	PHE	Sidechain
1	A	89	LEU	Mainchain
1	A	94	ALA	Mainchain
2	B	251	THR	Mainchain
2	B	252	CYS	Mainchain
2	B	255	SER	Mainchain
2	B	258	ILE	Mainchain
2	B	270	ASP	Sidechain
2	B	274	ASP	Sidechain
2	B	287	SER	Mainchain
2	B	300	LYS	Mainchain
2	B	304	THR	Mainchain
2	B	317	TYR	Sidechain
2	B	319	ALA	Mainchain
2	B	322	TYR	Sidechain
2	B	351	SER	Mainchain
2	B	368	VAL	Mainchain
2	B	388	ARG	Sidechain
2	B	392	ILE	Mainchain
2	B	393	TYR	Sidechain
2	B	396	ARG	Sidechain
2	B	397	ASP	Sidechain

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Group
2	B	399	CYS	Mainchain
2	B	400	MET	Mainchain
2	B	408	TYR	Sidechain
2	B	409	VAL	Mainchain
2	B	438	CYS	Mainchain
2	B	441	ASN	Mainchain
2	B	443	ARG	Sidechain
2	B	445	SER	Mainchain
2	B	449	VAL	Mainchain
2	B	456	SER	Mainchain
2	B	482	ASP	Sidechain
2	B	483	VAL	Mainchain
2	B	485	GLN	Mainchain
2	B	486	ALA	Mainchain,Peptide
2	B	489	SER	Mainchain
2	B	490	LEU	Peptide
2	B	508	PRO	Mainchain

## 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	1927	0	1896	35	0
2	B	1995	0	1934	65	0
3	A	14	0	13	1	0
3	B	42	0	39	3	0
4	A	15	0	0	0	0
4	B	5	0	0	0	0
5	A	12	0	16	2	0
5	B	12	0	16	3	0
6	A	1	0	0	0	0
7	A	12	0	11	0	0
8	B	28	0	22	3	0
9	A	27	0	0	1	0
9	B	45	0	0	0	0
All	All	4135	0	3947	102	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 13.

All (102) 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:B:489:SER:H	2:B:490:LEU:HA	1.29	0.98
1:A:148:GLN:HG2	1:A:150:LYS:HB3	1.46	0.96
2:B:408:TYR:HD2	2:B:491:GLN:O	1.48	0.95
2:B:489:SER:N	2:B:490:LEU:HA	1.80	0.94
2:B:488:PRO:HA	2:B:489:SER:HB2	1.50	0.93
1:A:148:GLN:HG3	1:A:150:LYS:H	1.35	0.91
2:B:288:LYS:H	2:B:295:GLN:HE22	1.14	0.90
2:B:265:ASN:HD22	2:B:430:ARG:HH22	1.22	0.87
1:A:115:TYR:O	1:A:119:GLU:OE2	1.92	0.86
1:A:148:GLN:CG	1:A:150:LYS:HB3	2.07	0.83
2:B:489:SER:N	2:B:490:LEU:CA	2.42	0.82
2:B:477:ASN:HD22	2:B:479:LEU:H	1.29	0.79
1:A:148:GLN:HG3	1:A:150:LYS:N	1.97	0.79
2:B:474:ASN:HD22	2:B:477:ASN:H	1.28	0.78
2:B:477:ASN:ND2	2:B:479:LEU:H	1.85	0.74
1:A:9:ASP:HA	1:A:136:ILE:HD11	1.69	0.72
2:B:262:VAL:HG22	2:B:268:THR:HG22	1.72	0.72
2:B:466:PHE:H	5:B:647:GOL:H11	1.54	0.71
2:B:490:LEU:H	2:B:491:GLN:CG	2.03	0.70
2:B:416:GLN:HE21	2:B:418:ASN:HD21	1.39	0.70
2:B:258:ILE:CD1	2:B:296:LEU:HB3	2.22	0.69
2:B:484:ALA:HA	2:B:485:GLN:HG3	1.73	0.69
2:B:441:ASN:HD22	2:B:495:ILE:HD13	1.56	0.69
1:A:25:ARG:HH12	1:A:170:ASN:HD21	1.46	0.63
2:B:408:TYR:CD2	2:B:491:GLN:O	2.40	0.62
2:B:477:ASN:C	2:B:477:ASN:HD22	2.02	0.62
2:B:484:ALA:CA	2:B:485:GLN:HG3	2.30	0.62
2:B:474:ASN:ND2	2:B:477:ASN:H	1.97	0.61
2:B:490:LEU:H	2:B:491:GLN:HG3	1.65	0.61
2:B:409:VAL:HG23	2:B:488:PRO:HG3	1.82	0.61
2:B:372:ASP:H	2:B:377:GLN:HE22	1.48	0.60
2:B:490:LEU:H	2:B:491:GLN:HG2	1.65	0.60
1:A:212:GLY:HA2	1:A:233:ILE:HG23	1.84	0.59
2:B:258:ILE:HD11	2:B:296:LEU:HB3	1.84	0.59
8:B:602:NAG:H83	8:B:602:NAG:H3	1.83	0.59
2:B:488:PRO:HA	2:B:489:SER:CB	2.29	0.58
2:B:490:LEU:N	2:B:491:GLN:HG2	2.19	0.57
1:A:71:ASP:HB3	1:A:74:ASN:HD21	1.70	0.57

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:265:ASN:ND2	2:B:430:ARG:HH22	1.98	0.57
1:A:170:ASN:ND2	5:A:649:GOL:H32	2.20	0.57
3:B:603:NAG:O4	3:B:604:NAG:H2	2.04	0.57
2:B:466:PHE:H	5:B:647:GOL:C1	2.17	0.56
2:B:488:PRO:HG2	2:B:488:PRO:O	2.06	0.55
2:B:492:ARG:HG3	2:B:492:ARG:HH11	1.71	0.55
2:B:490:LEU:CB	2:B:491:GLN:HG2	2.36	0.55
2:B:466:PHE:N	5:B:647:GOL:H11	2.19	0.55
2:B:490:LEU:HB3	2:B:491:GLN:HG2	1.90	0.54
1:A:112:ASN:OD1	3:A:500:NAG:C1	2.56	0.54
2:B:373:TYR:HE1	8:B:602:NAG:H81	1.73	0.53
1:A:25:ARG:HH22	1:A:170:ASN:ND2	2.07	0.53
1:A:74:ASN:ND2	1:A:76:TYR:H	2.07	0.52
2:B:492:ARG:CG	2:B:492:ARG:HH11	2.23	0.52
1:A:59:THR:HA	1:A:64:ASP:O	2.10	0.52
2:B:413:THR:H	2:B:419:GLN:HE22	1.59	0.51
2:B:288:LYS:NZ	2:B:294:ASN:HD21	2.08	0.51
2:B:490:LEU:CA	2:B:491:GLN:HG2	2.41	0.50
2:B:490:LEU:N	2:B:491:GLN:CG	2.73	0.50
1:A:56:VAL:HG11	1:A:163:ILE:HD13	1.94	0.50
1:A:149:THR:HA	1:A:152:GLN:HE21	1.78	0.49
2:B:488:PRO:CA	2:B:489:SER:HB2	2.33	0.48
1:A:71:ASP:HB3	1:A:74:ASN:ND2	2.27	0.48
2:B:308:ASN:HD21	3:B:600:NAG:C1	2.26	0.48
1:A:148:GLN:HE21	1:A:149:THR:H	1.63	0.47
1:A:175[A]:ARG:NH2	1:A:189:ASP:OD1	2.47	0.47
1:A:125:ARG:HG3	1:A:193:LEU:CD2	2.45	0.47
2:B:373:TYR:CE1	8:B:602:NAG:H81	2.49	0.47
2:B:477:ASN:HD22	2:B:478:GLY:N	2.13	0.46
1:A:148:GLN:HG3	1:A:150:LYS:HB3	1.94	0.46
2:B:486:ALA:O	2:B:487:ASN:ND2	2.49	0.46
2:B:490:LEU:C	2:B:491:GLN:HG2	2.37	0.46
2:B:509:VAL:HA	2:B:510:PRO:HD2	1.82	0.46
1:A:67:THR:O	1:A:80:TYR:HA	2.16	0.46
1:A:20:PHE:CZ	1:A:56:VAL:HG13	2.51	0.45
1:A:8:THR:O	1:A:60:ASN:HA	2.16	0.45
2:B:275:ASP:HB3	2:B:280:ASN:ND2	2.31	0.45
1:A:25:ARG:HH12	1:A:170:ASN:ND2	2.14	0.45
1:A:125:ARG:HG3	1:A:193:LEU:HD22	1.98	0.45
2:B:440:THR:HA	2:B:461:GLY:O	2.16	0.45
2:B:294:ASN:HD22	2:B:294:ASN:H	1.65	0.45

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:25:ARG:NH1	1:A:170:ASN:HD21	2.13	0.44
2:B:418:ASN:C	2:B:418:ASN:HD22	2.21	0.44
1:A:89:LEU:HD22	1:A:109:GLN:HE22	1.83	0.44
1:A:144:PHE:HA	1:A:145:PRO:HD3	1.85	0.44
2:B:488:PRO:CA	2:B:489:SER:CB	2.95	0.43
1:A:148:GLN:NE2	9:A:658:HOH:O	2.51	0.43
2:B:308:ASN:HD21	3:B:600:NAG:C2	2.31	0.43
2:B:402:SER:HB3	2:B:421:TRP:CH2	2.54	0.43
2:B:416:GLN:HE21	2:B:418:ASN:ND2	2.11	0.42
2:B:395:PHE:HE2	2:B:487:ASN:HA	1.83	0.42
2:B:424:TYR:HB2	2:B:428:SER:OG	2.19	0.42
2:B:477:ASN:C	2:B:477:ASN:ND2	2.72	0.42
2:B:484:ALA:CB	2:B:485:GLN:HG3	2.50	0.42
1:A:175[A]:ARG:NH1	5:A:650:GOL:O1	2.43	0.42
1:A:155:SER:O	1:A:159:LEU:HG	2.19	0.42
2:B:403:ALA:HB3	2:B:408:TYR:HE1	1.84	0.42
2:B:292:ASP:HA	2:B:293:PRO:HD3	1.84	0.41
1:A:28:VAL:HG12	1:A:72:VAL:HG22	2.02	0.41
2:B:375:LEU:HD22	2:B:458:GLY:HA2	2.03	0.41
1:A:214:PHE:HE1	1:A:233:ILE:HG22	1.85	0.41
2:B:488:PRO:CG	2:B:488:PRO:O	2.60	0.41
1:A:172:ILE:HG12	1:A:192:MET:HG3	2.04	0.40
2:B:310:SER:HB2	2:B:326:PHE:O	2.21	0.40

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	247/254 (97%)	234 (95%)	11 (4%)	2 (1%)	24	51
2	B	260/263 (99%)	241 (93%)	15 (6%)	4 (2%)	13	32

Continued on next page...



*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
All	All	507/517 (98%)	475 (94%)	26 (5%)	6 (1%)	16	39

All (6) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	487	ASN
2	B	488	PRO
1	A	101	PHE
2	B	489	SER
1	A	115	TYR
2	B	395	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	210/215 (98%)	192 (91%)	18 (9%)	13	29
2	B	218/219 (100%)	204 (94%)	14 (6%)	22	47
All	All	428/434 (99%)	396 (92%)	32 (8%)	17	38

All (32) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	7	ARG
1	A	10	GLN
1	A	41	ARG
1	A	48	SER
1	A	49	GLU
1	A	55	LEU
1	A	60	ASN
1	A	74	ASN
1	A	89	LEU
1	A	102	SER
1	A	112	ASN
1	A	148	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	149	THR
1	A	157	LEU
1	A	226	VAL
1	A	228	VAL
1	A	230	LEU
1	A	233	ILE
2	B	251	THR
2	B	255	SER
2	B	258	ILE
2	B	294	ASN
2	B	375	LEU
2	B	389	GLU
2	B	397	ASP
2	B	406	SER
2	B	408	TYR
2	B	418	ASN
2	B	439	LEU
2	B	477	ASN
2	B	483	VAL
2	B	490	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (24) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	10	GLN
1	A	60	ASN
1	A	74	ASN
1	A	85	GLN
1	A	109	GLN
1	A	112	ASN
1	A	148	GLN
1	A	152	GLN
1	A	161	GLN
1	A	170	ASN
1	A	215	ASN
1	A	216	ASN
2	B	265	ASN
2	B	294	ASN
2	B	295	GLN
2	B	308	ASN
2	B	348	ASN
2	B	377	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
2	B	418	ASN
2	B	419	GLN
2	B	462	GLN
2	B	474	ASN
2	B	477	ASN
2	B	487	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 ⓘ

2 carbohydrates 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$
8	NAG	B	602	8,2	14,14,15	0.63	0	15,19,21	2.15	5 (33%)
8	NAG	B	605	8	14,14,15	0.75	0	15,19,21	1.26	2 (13%)

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
8	NAG	B	602	8,2	-	0/6/23/26	0/1/1/1
8	NAG	B	605	8	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
8	B	602	NAG	C3-C4-C5	-2.19	106.38	110.20
8	B	605	NAG	O7-C7-C8	-2.15	118.11	122.06
8	B	605	NAG	C1-O5-C5	2.28	115.14	112.25
8	B	602	NAG	C8-C7-N2	2.78	121.42	116.11
8	B	602	NAG	O4-C4-C5	2.87	116.84	109.24
8	B	602	NAG	C1-O5-C5	4.04	117.38	112.25
8	B	602	NAG	C2-N2-C7	4.94	129.39	123.04

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
8	B	602	NAG	3	0

## 5.6 Ligand geometry

Of 14 ligands modelled in this entry, 1 is monoatomic - leaving 13 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	$\# Z  > 2$
4	SO4	A	255	-	4,4,4	0.17	0	6,6,6	0.18	0
4	SO4	A	256	-	4,4,4	0.18	0	6,6,6	0.25	0
4	SO4	A	257	-	4,4,4	0.20	0	6,6,6	0.11	0
3	NAG	A	500	-	14,14,15	0.44	0	15,19,21	1.07	1 (6%)
7	SGI	A	600	-	12,12,12	0.51	0	15,15,15	0.62	0
5	GOL	A	649	-	5,5,5	0.30	0	5,5,5	0.30	0
5	GOL	A	650	-	5,5,5	0.40	0	5,5,5	0.31	0
4	SO4	B	3	-	4,4,4	0.25	0	6,6,6	0.23	0
3	NAG	B	600	-	14,14,15	0.60	0	15,19,21	1.36	2 (13%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	NAG	B	603	2	14,14,15	0.51	0	15,19,21	1.54	2 (13%)
3	NAG	B	604	-	14,14,15	0.45	0	15,19,21	1.64	2 (13%)
5	GOL	B	647	-	5,5,5	0.27	0	5,5,5	0.55	0
5	GOL	B	648	-	5,5,5	0.27	0	5,5,5	1.08	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
4	SO4	A	255	-	-	0/0/0/0	0/0/0/0
4	SO4	A	256	-	-	0/0/0/0	0/0/0/0
4	SO4	A	257	-	-	0/0/0/0	0/0/0/0
3	NAG	A	500	-	-	0/6/23/26	0/1/1/1
7	SGI	A	600	-	-	0/5/5/5	0/1/1/1
5	GOL	A	649	-	-	0/4/4/4	0/0/0/0
5	GOL	A	650	-	-	0/4/4/4	0/0/0/0
4	SO4	B	3	-	-	0/0/0/0	0/0/0/0
3	NAG	B	600	-	-	0/6/23/26	0/1/1/1
3	NAG	B	603	2	-	0/6/23/26	0/1/1/1
3	NAG	B	604	-	-	0/6/23/26	0/1/1/1
5	GOL	B	647	-	-	0/4/4/4	0/0/0/0
5	GOL	B	648	-	-	0/4/4/4	0/0/0/0

There are no bond length outliers.

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	603	NAG	C6-C5-C4	-2.32	107.28	113.02
3	A	500	NAG	C4-C3-C2	-2.14	107.90	111.23
3	B	600	NAG	C3-C4-C5	2.92	115.29	110.20
3	B	600	NAG	C1-O5-C5	3.47	116.65	112.25
3	B	604	NAG	C3-C4-C5	3.84	116.90	110.20
3	B	604	NAG	C1-O5-C5	4.28	117.68	112.25
3	B	603	NAG	C1-O5-C5	4.74	118.26	112.25

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

7 monomers are involved in 9 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	500	NAG	1	0
5	A	649	GOL	1	0
5	A	650	GOL	1	0
3	B	600	NAG	2	0
3	B	603	NAG	1	0
3	B	604	NAG	1	0
5	B	647	GOL	3	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	248/254 (97%)	0.02	8 (3%) 51 51	35, 48, 68, 72	1 (0%)
2	B	262/263 (99%)	0.03	11 (4%) 40 39	32, 41, 57, 72	0
All	All	510/517 (98%)	0.02	19 (3%) 45 45	32, 44, 65, 72	1 (0%)

All (19) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	486	ALA	8.1
2	B	487	ASN	6.8
2	B	491	GLN	5.0
1	A	224	PRO	5.0
1	A	223	ALA	4.8
2	B	253	THR	4.6
1	A	248	GLY	3.9
1	A	225	GLY	3.9
2	B	490	LEU	3.6
2	B	277	HIS	3.3
2	B	489	SER	2.7
1	A	105	THR	2.7
1	A	91	ASP	2.5
2	B	254	ALA	2.4
1	A	247	CYS	2.3
2	B	396	ARG	2.3
2	B	312	LEU	2.2
1	A	11	GLN	2.2
2	B	291	ASN	2.2

### 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](#)

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
8	NAG	B	602	14/15	0.87	0.20	-	56,60,64,68	0
8	NAG	B	605	14/15	0.80	0.27	-	71,74,78,78	0

### 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
4	SO4	B	3	5/5	0.82	0.30	6.76	92,93,93,94	0
5	GOL	B	648	6/6	0.87	0.23	5.75	49,50,51,51	0
5	GOL	B	647	6/6	0.83	0.28	4.72	67,71,71,72	0
5	GOL	A	649	6/6	0.86	0.27	2.99	59,62,62,63	0
5	GOL	A	650	6/6	0.91	0.24	1.54	50,53,53,56	0
7	SGI	A	600	12/12	0.93	0.20	1.31	41,42,43,43	0
4	SO4	A	255	5/5	0.90	0.26	1.23	101,101,101,101	0
3	NAG	B	603	14/15	0.96	0.14	-0.58	40,43,47,48	0
4	SO4	A	256	5/5	0.97	0.09	-1.38	54,55,55,55	0
3	NAG	A	500	14/15	0.70	0.27	-	93,94,95,95	0
4	SO4	A	257	5/5	0.98	0.28	-	86,86,86,87	0
3	NAG	B	604	14/15	0.72	0.32	-	81,88,89,89	0
6	CL	A	258	1/1	0.92	0.19	-	77,77,77,77	0
3	NAG	B	600	14/15	0.62	0.27	-	80,83,86,86	0

### 6.5 Other polymers [i](#)

There are no such residues in this entry.