



wwPDB X-ray Structure Validation Summary Report

Feb 27, 2014 – 04:14 PM GMT

PDB ID : 4AW3
Title : Structure of the mixed-function P450 MycG F286V mutant in complex with mycinamicin V in P1 space group
Authors : Li, S.; Tietz, D.R.; Rutaganira, F.U.; Kells, P.M.; Anzai, Y.; Kato, F.; Pochapsky, T.C.; Sherman, D.H.; Podust, L.M.
Deposited on : 2012-05-30
Resolution : 2.05 Å(reported)

This is a wwPDB validation summary report for a publicly released PDB entry.
We welcome your comments at validation@mail.wwpdb.org
A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

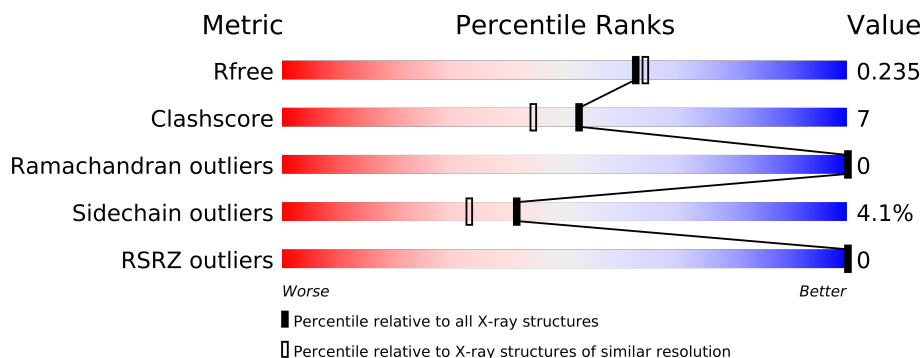
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : **FAILED**
Xtriage (Phenix) : dev-1323
EDS : stable22639
Percentile statistics : 21963
Refmac : 5.8.0049
CCP4 : 6.3.0 (Settle)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : stable22683

1 Overall quality at a glance

The reported resolution of this entry is 2.05 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}	66092	1380 (2.06-2.02)
Clashscore	79885	1577 (2.06-2.02)
Ramachandran outliers	78287	1565 (2.06-2.02)
Sidechain outliers	78261	1565 (2.06-2.02)
RSRZ outliers	66119	1381 (2.06-2.02)

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 ≥ 3 , 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	417	<div><div></div><div></div><div></div><div></div><div></div></div>
1	B	417	<div><div></div><div></div><div></div><div></div><div></div></div>

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
3	MYV	A	500[A]	-	X
3	MYV	A	500[B]	-	X
3	MYV	B	500[A]	-	X
3	MYV	B	500[B]	-	X
4	SO4	A	1398	-	X
4	SO4	A	1399	-	X

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Geometry	Electron density
4	SO4	A	1401	-	X
4	SO4	B	1400	-	X
5	GOL	B	1402	-	X

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 7065 atoms, of which 0 are hydrogen and 0 are deuterium.

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 P-450-LIKE PROTEIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	393	Total	C	N	O	S	0	7	0
			3122	1954	572	584	12			
1	B	393	Total	C	N	O	S	0	2	0
			3075	1930	563	571	11			

There are 42 discrepancies between the modelled and reference sequences:

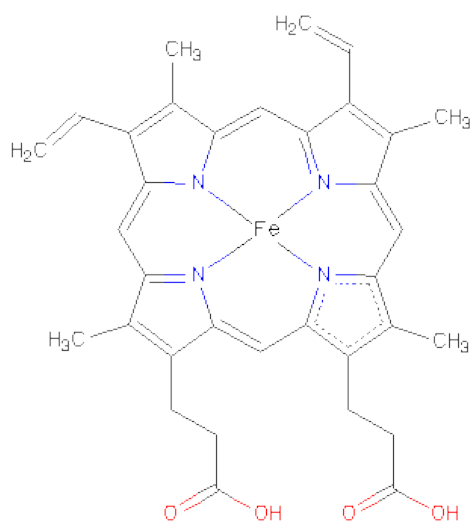
Chain	Residue	Modelled	Actual	Comment	Reference
A	-19	MET	-	EXPRESSION TAG	UNP Q59523
A	-18	GLY	-	EXPRESSION TAG	UNP Q59523
A	-17	SER	-	EXPRESSION TAG	UNP Q59523
A	-16	SER	-	EXPRESSION TAG	UNP Q59523
A	-15	HIS	-	EXPRESSION TAG	UNP Q59523
A	-14	HIS	-	EXPRESSION TAG	UNP Q59523
A	-13	HIS	-	EXPRESSION TAG	UNP Q59523
A	-12	HIS	-	EXPRESSION TAG	UNP Q59523
A	-11	HIS	-	EXPRESSION TAG	UNP Q59523
A	-10	HIS	-	EXPRESSION TAG	UNP Q59523
A	-9	SER	-	EXPRESSION TAG	UNP Q59523
A	-8	SER	-	EXPRESSION TAG	UNP Q59523
A	-7	GLY	-	EXPRESSION TAG	UNP Q59523
A	-6	LEU	-	EXPRESSION TAG	UNP Q59523
A	-5	VAL	-	EXPRESSION TAG	UNP Q59523
A	-4	PRO	-	EXPRESSION TAG	UNP Q59523
A	-3	ARG	-	EXPRESSION TAG	UNP Q59523
A	-2	GLY	-	EXPRESSION TAG	UNP Q59523
A	-1	SER	-	EXPRESSION TAG	UNP Q59523
A	0	HIS	-	EXPRESSION TAG	UNP Q59523
A	286	VAL	PHE	ENGINEERED MUTATION	UNP Q59523
B	-19	MET	-	EXPRESSION TAG	UNP Q59523
B	-18	GLY	-	EXPRESSION TAG	UNP Q59523
B	-17	SER	-	EXPRESSION TAG	UNP Q59523
B	-16	SER	-	EXPRESSION TAG	UNP Q59523

Continued on next page...

Continued from previous page...

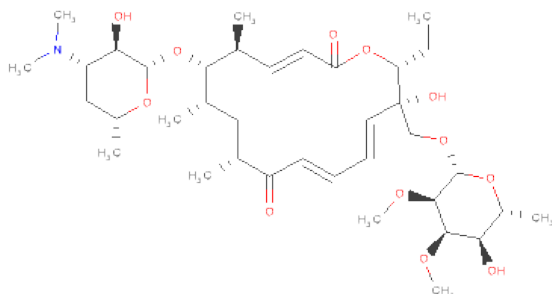
Chain	Residue	Modelled	Actual	Comment	Reference
B	-15	HIS	-	EXPRESSION TAG	UNP Q59523
B	-14	HIS	-	EXPRESSION TAG	UNP Q59523
B	-13	HIS	-	EXPRESSION TAG	UNP Q59523
B	-12	HIS	-	EXPRESSION TAG	UNP Q59523
B	-11	HIS	-	EXPRESSION TAG	UNP Q59523
B	-10	HIS	-	EXPRESSION TAG	UNP Q59523
B	-9	SER	-	EXPRESSION TAG	UNP Q59523
B	-8	SER	-	EXPRESSION TAG	UNP Q59523
B	-7	GLY	-	EXPRESSION TAG	UNP Q59523
B	-6	LEU	-	EXPRESSION TAG	UNP Q59523
B	-5	VAL	-	EXPRESSION TAG	UNP Q59523
B	-4	PRO	-	EXPRESSION TAG	UNP Q59523
B	-3	ARG	-	EXPRESSION TAG	UNP Q59523
B	-2	GLY	-	EXPRESSION TAG	UNP Q59523
B	-1	SER	-	EXPRESSION TAG	UNP Q59523
B	0	HIS	-	EXPRESSION TAG	UNP Q59523
B	286	VAL	PHE	ENGINEERED MUTATION	UNP Q59523

- Molecule 2 is PROTOPORPHYRIN IX CONTAINING FE (three-letter code: HEM) (formula: $C_{34}H_{32}FeN_4O_4$).



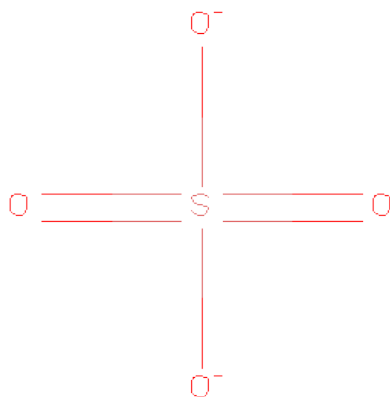
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
2	B	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		

- Molecule 3 is MYCINAMICIN V (three-letter code: MYV) (formula: $C_{37}H_{61}NO_{12}$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	1
			100	74	2	24		
3	B	1	Total	C	N	O	0	1
			100	74	2	24		

- Molecule 4 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



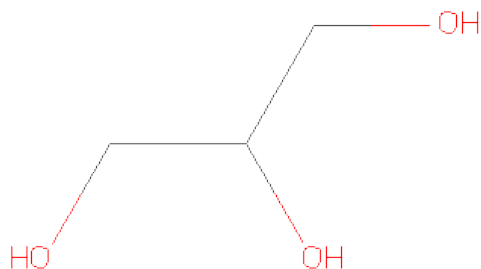
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	O	S	0	0
			5	4	1		

Continued on next page...

Continued from previous page...

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

- Molecule 5 is GLYCEROL (three-letter code: GOL) (formula: C₃H₈O₃).



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

- Molecule 6 is water.

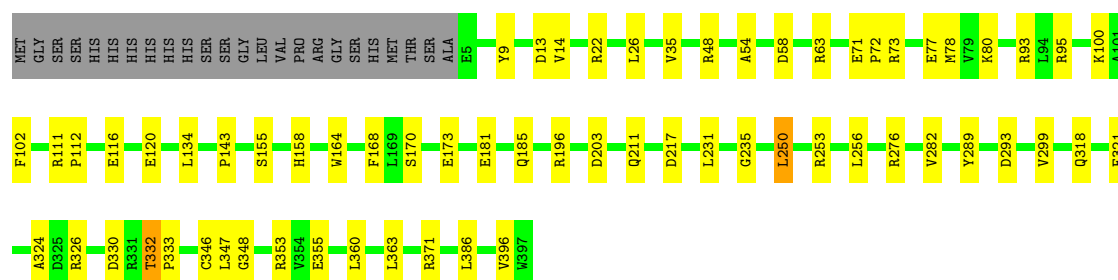
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	273	Total 273	O 273	0	0
6	B	245	Total 245	O 245	0	0

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 ($RSRZ > 2$). 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.

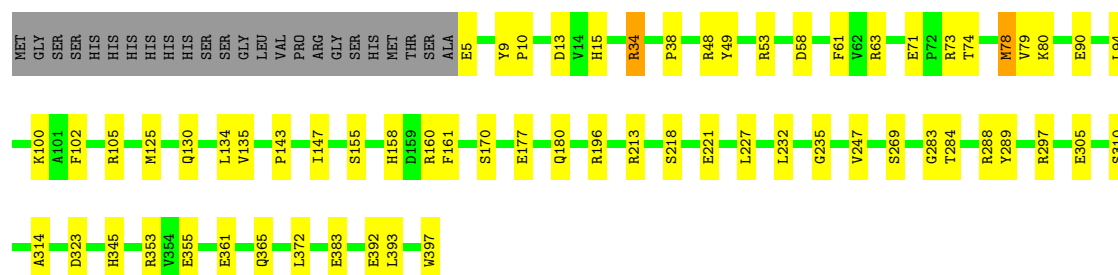
• Molecule 1: P-450-LIKE PROTEIN

Chain A:



• Molecule 1: P-450-LIKE PROTEIN

Chain B:



4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	55.45Å 56.11Å 76.42Å 90.43° 97.22° 102.11°	Depositor
Resolution (Å)	75.77 – 2.05 75.77 – 2.05	Depositor EDS
% Data completeness (in resolution range)	96.4 (75.77-2.05) 96.4 (75.77-2.05)	Depositor EDS
R_{merge}	0.17	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.81 (at 2.05Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, R_{free}	0.166 , 0.235 0.168 , 0.235	Depositor DCC
R_{free} test set	2752 reflections (5.37%)	DCC
Wilson B-factor (Å ²)	21.8	Xtriage
Anisotropy	0.333	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.38 , 30.2	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 54004 reflections	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	7065	wwPDB-VP
Average B, all atoms (Å ²)	22.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.59% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

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: GOL, SO4, HEM, MYV

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	1.09	1/3186 (0.0%)	1.03	13/4336 (0.3%)
1	B	1.05	2/3139 (0.1%)	0.94	3/4274 (0.1%)
All	All	1.07	3/6325 (0.0%)	0.99	16/8610 (0.2%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	90	GLU	CG-CD	6.06	1.61	1.51
1	A	95	ARG	CB-CG	-5.04	1.39	1.52
1	B	221	GLU	CG-CD	5.03	1.59	1.51

The worst 5 of 16 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	63	ARG	NE-CZ-NH2	-7.69	116.46	120.30
1	A	253	ARG	NE-CZ-NH2	-7.20	116.70	120.30
1	B	63	ARG	NE-CZ-NH1	7.02	123.81	120.30
1	A	95	ARG	NE-CZ-NH2	-6.50	117.05	120.30
1	A	276	ARG	NE-CZ-NH2	-6.42	117.09	120.30

There are no chirality outliers.

There are no planarity outliers.

5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the

chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3122	0	3083	37	0
1	B	3075	0	3053	44	0
2	A	43	0	30	4	0
2	B	43	0	30	1	0
3	A	100	0	122	7	0
3	B	100	0	122	10	0
4	A	20	0	0	0	0
4	B	20	0	0	3	0
5	A	12	0	16	1	0
5	B	12	0	16	3	0
6	A	273	0	0	5	0
6	B	245	0	0	10	0
All	All	7065	0	6472	93	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 7.

The worst 5 of 93 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:B:48:ARG:NH2	6:B:2050:HOH:O	1.98	0.96
3:B:500[B]:MYV:H73C	3:B:500[B]:MYV:O4	1.78	0.83
3:B:500[B]:MYV:H72C	3:B:500[B]:MYV:HB	1.62	0.82
1:A:250:LEU:HD12	1:A:256:LEU:HG	1.61	0.81
1:B:5:GLU:HG2	6:B:2002:HOH:O	1.82	0.79

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone ⓘ

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	398/417 (95%)	386 (97%)	12 (3%)	0	100	100
1	B	393/417 (94%)	384 (98%)	9 (2%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
All	All	791/834 (95%)	770 (97%)	21 (3%)	0	100	100

There are no Ramachandran outliers to report.

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	326/346 (94%)	310 (95%)	16 (5%)	35	25
1	B	321/346 (93%)	308 (96%)	13 (4%)	42	33
All	All	647/692 (94%)	618 (96%)	29 (4%)	41	28

5 of 29 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	330[A]	ASP
1	B	34	ARG
1	B	297	ARG
1	A	330[B]	ASP
1	B	78	MET

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	320	GLN
1	B	318	GLN
1	B	130	GLN
1	A	211	GLN
1	B	15	HIS

5.3.3 RNA ⓘ

There are no RNA chains in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

Mogul failed to run properly - this section will therefore be empty.

5.5 Carbohydrates ⓘ

Mogul failed to run properly - this section will therefore be empty.

5.6 Ligand geometry ⓘ

Mogul failed to run properly - this section will therefore be empty.

5.7 Other polymers ⓘ

Mogul failed to run properly - this section will therefore be empty.

5.8 Polymer linkage issues

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, 95th 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 ²)	Q<0.9
1	A	393/417 (94%)	-0.51	0 100 100	9, 19, 34, 48	0
1	B	393/417 (94%)	-0.47	0 100 100	10, 21, 37, 51	0
All	All	786/834 (94%)	-0.49	0 100 100	9, 20, 36, 51	0

There are no RSRZ outliers to report.

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

6.4 Ligands ⓘ

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, 95th 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	RSR	LLDF	B-factors(Å ²)	Q<0.9
4	SO4	A	1399	5/5	0.18	10.56	50,53,54,55	0
4	SO4	A	1398	5/5	0.20	8.98	58,59,60,60	0
5	GOL	B	1402	6/6	0.13	6.70	36,42,42,44	0
3	MYV	A	500[A]	50/50	0.15	3.64	17,23,34,35	50
3	MYV	B	500[B]	50/50	0.15	3.55	19,24,30,30	50

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors(\AA^2)	Q<0.9
3	MYV	B	500[A]	50/50	0.15	3.33	5,17,29,31	50
4	SO4	B	1400	5/5	0.17	3.10	55,55,56,57	0
3	MYV	A	500[B]	50/50	0.15	2.80	6,25,36,37	50
4	SO4	A	1401	5/5	0.12	2.19	61,62,64,64	0
5	GOL	A	1402	6/6	0.11	1.85	45,45,47,48	0
4	SO4	A	1400	5/5	0.12	1.32	52,55,57,57	0
4	SO4	B	1398	5/5	0.10	0.91	25,25,28,31	0
5	GOL	B	1403	6/6	0.09	0.62	22,26,28,29	0
5	GOL	A	1403	6/6	0.09	0.44	20,23,28,29	0
2	HEM	A	450	43/43	0.09	0.29	8,14,16,19	0
4	SO4	B	1399	5/5	0.11	0.22	60,60,61,63	0
4	SO4	B	1401	5/5	0.10	0.20	40,45,46,46	0
2	HEM	B	450	43/43	0.09	0.19	8,13,17,23	0

6.5 Other polymers ⓘ

There are no such residues in this entry.