WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Sep 7, 2023 – 01:18 AM EDT

PDB ID	:	4GP2
Title	:	Crystal structure of ISOPRENOID SYNTHASE A3MSH1 (TARGET EFI-501992) from pyrobaculum calidifontis complexed with DMAPP and Magnesium
Authors	:	Patskovsky, Y.; Toro, R.; Bhosle, R.; Hillerich, B.; Seidel, R.D.; Washington, E.; Scott Glenn, A.; Chowdhury, S.; Evans, B.; Hammonds, J.; Zencheck, W.D.; Imker, H.J.; Poulter, C.D.; Gerlt, J.A.; Almo, S.C.; Enzyme Function Initiative (EFI)
Deposited on	:	2012-08-20
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*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) 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 (i)) were used in the production of this report:

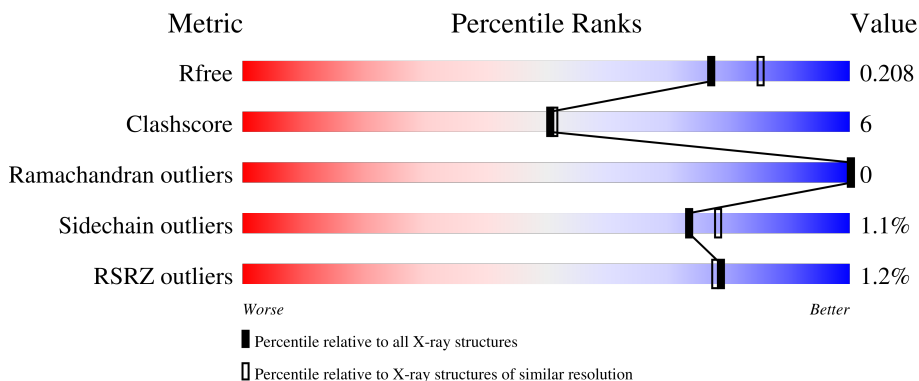|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)

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	Percentile Ranks	Value
Rfree		0.208
Clashscore		6
Ramachandran outliers		0
Sidechain outliers		1.1%
RSRZ outliers		1.2%

Worse   *Better*

▌ Percentile relative to all X-ray structures

◻ Percentile relative to X-ray structures of similar resolution

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 of the lower bar indicate the fraction of residues that contain outliers for >=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 <=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	358	2% — 84% — 11% · ·
1	B	358	% — 86% — 6% 8%

Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

WORLDWIDE
ᴡᴡPDB
PROTEIN DATA BANK

2 Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 5607 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 Polyprenyl synthetase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	342	Total 2705	C 1723	N 482	O 497	S 3	0	4	0
1	B	331	Total 2585	C 1644	N 458	O 481	S 2	0	1	0

There are 48 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	expression tag	UNP A3MSH1
A	2	VAL	-	expression tag	UNP A3MSH1
A	337	ALA	-	expression tag	UNP A3MSH1
A	338	GLU	-	expression tag	UNP A3MSH1
A	339	ASN	-	expression tag	UNP A3MSH1
A	340	LEU	-	expression tag	UNP A3MSH1
A	341	TYR	-	expression tag	UNP A3MSH1
A	342	PHE	-	expression tag	UNP A3MSH1
A	343	GLN	-	expression tag	UNP A3MSH1
A	344	SER	-	expression tag	UNP A3MSH1
A	345	HIS	-	expression tag	UNP A3MSH1
A	346	HIS	-	expression tag	UNP A3MSH1
A	347	HIS	-	expression tag	UNP A3MSH1
A	348	HIS	-	expression tag	UNP A3MSH1
A	349	HIS	-	expression tag	UNP A3MSH1
A	350	HIS	-	expression tag	UNP A3MSH1
A	351	TRP	-	expression tag	UNP A3MSH1
A	352	SER	-	expression tag	UNP A3MSH1
A	353	HIS	-	expression tag	UNP A3MSH1
A	354	PRO	-	expression tag	UNP A3MSH1
A	355	GLN	-	expression tag	UNP A3MSH1
A	356	PHE	-	expression tag	UNP A3MSH1
A	357	GLU	-	expression tag	UNP A3MSH1
A	358	LYS	-	expression tag	UNP A3MSH1
B	1	MET	-	expression tag	UNP A3MSH1

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	2	VAL	-	expression tag	UNP A3MSH1
B	337	ALA	-	expression tag	UNP A3MSH1
B	338	GLU	-	expression tag	UNP A3MSH1
B	339	ASN	-	expression tag	UNP A3MSH1
B	340	LEU	-	expression tag	UNP A3MSH1
B	341	TYR	-	expression tag	UNP A3MSH1
B	342	PHE	-	expression tag	UNP A3MSH1
B	343	GLN	-	expression tag	UNP A3MSH1
B	344	SER	-	expression tag	UNP A3MSH1
B	345	HIS	-	expression tag	UNP A3MSH1
B	346	HIS	-	expression tag	UNP A3MSH1
B	347	HIS	-	expression tag	UNP A3MSH1
B	348	HIS	-	expression tag	UNP A3MSH1
B	349	HIS	-	expression tag	UNP A3MSH1
B	350	HIS	-	expression tag	UNP A3MSH1
B	351	TRP	-	expression tag	UNP A3MSH1
B	352	SER	-	expression tag	UNP A3MSH1
B	353	HIS	-	expression tag	UNP A3MSH1
B	354	PRO	-	expression tag	UNP A3MSH1
B	355	GLN	-	expression tag	UNP A3MSH1
B	356	PHE	-	expression tag	UNP A3MSH1
B	357	GLU	-	expression tag	UNP A3MSH1
B	358	LYS	-	expression tag	UNP A3MSH1

- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	3	Total 3	Mg 3	0	0
2	B	3	Total 3	Mg 3	0	0

- Molecule 3 is DIMETHYLALLYL DIPHOSPHATE (three-letter code: DMA) (formula: $C_5H_{12}O_7P_2$).

DMA
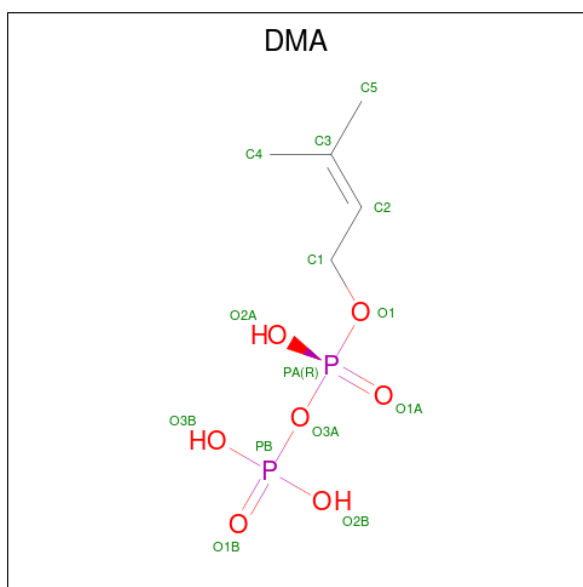
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total 14	C 5	O 7	P 2	0	0
3	A	1	Total 14	C 5	O 7	P 2	0	0
3	B	1	Total 14	C 5	O 7	P 2	0	0
3	B	1	Total 14	C 5	O 7	P 2	0	0

- Molecule 4 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	138	Total 138	O 138	0	0
4	B	117	Total 117	O 117	0	0

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: Polyprenyl synthetase

Chain A:

• Molecule 1: Polyprenyl synthetase

Chain B:
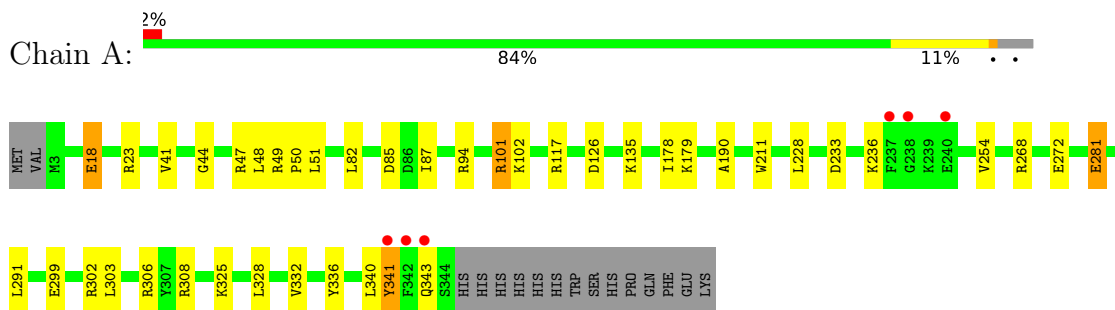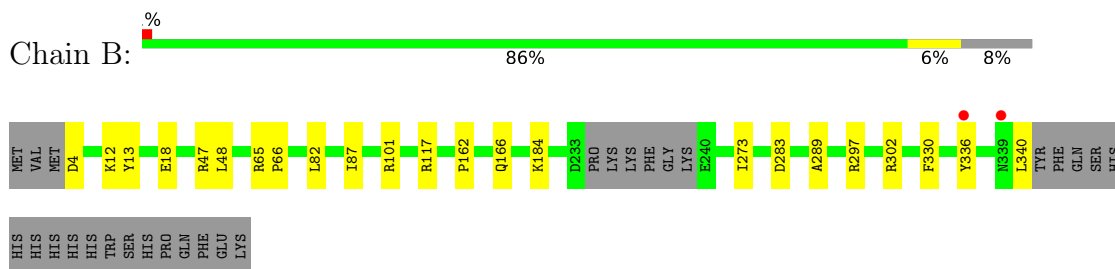
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	53.99Å 93.56Å 146.89Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.81 – 2.00	Depositor
	46.76 – 2.00	EDS
% Data completeness	99.9 (46.81-2.00)	Depositor
(in resolution range)	99.9 (46.76-2.00)	EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.07	Depositor
$< I/\sigma(I) >$ [1]	3.17 (at 2.00Å)	Xtriage
Refinement program	REFMAC 5.7.0029	Depositor
R, R_{free}	0.171 , 0.207	Depositor
	0.173 , 0.208	DCC
R_{free} test set	1589 reflections (3.11%)	wwPDB-VP
Wilson B-factor (Å^2)	37.3	Xtriage
Anisotropy	0.110	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 43.8	EDS
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	5607	wwPDB-VP
Average B, all atoms (Å^2)	44.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.00% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< |L| >$, $< L^2 >$ 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: MG, DMA

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	0.49	0/2761	0.61	0/3731
1	B	0.49	0/2628	0.61	0/3557
All	All	0.49	0/5389	0.61	0/7288

There are no bond length outliers.

There are no bond angle outliers.

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	2705	0	2762	45	0
1	B	2585	0	2622	16	0
2	A	3	0	0	0	0
2	B	3	0	0	0	0
3	A	28	0	18	5	0
3	B	28	0	18	3	0
4	A	138	0	0	9	0
4	B	117	0	0	2	0
All	All	5607	0	5420	61	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 6.

The worst 5 of 61 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:B:184:LYS:HE2	3:B:404:DMA:H2	1.35	1.05
1:B:184:LYS:CE	3:B:404:DMA:H2	2.07	0.85
1:A:18:GLU:HG2	1:A:47[A]:ARG:HH12	1.41	0.83
1:A:47[B]:ARG:HH11	1:A:47[B]:ARG:CG	1.92	0.82
1:A:233:ASP:OD2	1:A:236:LYS:HD2	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	344/358 (96%)	341 (99%)	3 (1%)	0	100	100
1	B	328/358 (92%)	321 (98%)	7 (2%)	0	100	100
All	All	672/716 (94%)	662 (98%)	10 (2%)	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	271/285 (95%)	266 (98%)	5 (2%)	59	63

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	B	258/285 (90%)	257 (100%)	1 (0%)	91	93
All	All	529/570 (93%)	523 (99%)	6 (1%)	73	78

5 of 6 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	306	ARG
1	A	341	TYR
1	B	4	ASP
1	A	101	ARG
1	A	18	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	11	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 monosaccharides in this entry.

5.6   Ligand geometry ⓘ

Of 10 ligands modelled in this entry, 6 are monoatomic - leaving 4 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	DMA	A	404	-	11,13,13	0.62	0	15,19,19	1.52	3 (20%)
3	DMA	A	403	2	11,13,13	0.63	0	15,19,19	1.58	2 (13%)
3	DMA	B	405	-	11,13,13	0.65	0	15,19,19	1.30	2 (13%)
3	DMA	B	404	2	11,13,13	0.52	0	15,19,19	1.35	1 (6%)

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	DMA	A	404	-	-	1/13/13/13	-
3	DMA	A	403	2	-	1/13/13/13	-
3	DMA	B	405	-	-	0/13/13/13	-
3	DMA	B	404	2	-	1/13/13/13	-

There are no bond length outliers.

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	404	DMA	C1-C2-C3	-4.08	120.15	126.75
3	A	403	DMA	C5-C3-C4	3.24	121.76	114.60
3	A	403	DMA	C1-C2-C3	-2.86	122.12	126.75
3	B	405	DMA	C1-C2-C3	-2.82	122.19	126.75
3	B	404	DMA	C5-C3-C4	2.79	120.77	114.60

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	B	404	DMA	PB-O3A-PA-O2A
3	A	403	DMA	C1-O1-PA-O3A
3	A	404	DMA	C1-O1-PA-O1A

There are no ring outliers.

3 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	404	DMA	2	0
3	A	403	DMA	4	0
3	B	404	DMA	3	0

5.7 Other polymers ⓘ

There are no such residues in this entry.

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	342/358 (95%)	-0.08	6 (1%)	68	66	27, 38, 80, 144	0
1	B	331/358 (92%)	-0.33	2 (0%)	89	88	26, 40, 78, 131	0
All	All	673/716 (93%)	-0.20	8 (1%)	79	78	26, 39, 79, 144	0

The worst 5 of 8 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	341	TYR	6.2
1	A	237	PHE	3.1
1	A	238	GLY	3.0
1	A	342	PHE	2.8
1	B	336	TYR	2.8

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 monosaccharides 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. 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	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	MG	B	403	1/1	0.95	0.03	48,48,48,48	0
2	MG	B	401	1/1	0.96	0.05	40,40,40,40	0
3	DMA	A	403	14/14	0.96	0.09	41,64,82,83	0
3	DMA	B	404	14/14	0.96	0.09	37,53,69,72	0
2	MG	A	402	1/1	0.97	0.04	45,45,45,45	0
3	DMA	B	405	14/14	0.97	0.10	40,48,110,114	0
2	MG	A	401	1/1	0.98	0.02	42,42,42,42	0
3	DMA	A	404	14/14	0.98	0.11	45,56,91,101	0
2	MG	B	402	1/1	0.98	0.04	40,40,40,40	0
2	MG	A	405	1/1	0.98	0.03	45,45,45,45	0

6.5　Other polymers ⓘ

There are no such residues in this entry.