



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 09:22 AM GMT

PDB ID : 3I5X  
Title : Structure of Mss116p bound to ssRNA and AMP-PNP  
Authors : Del Campo, M.; Lambowitz, A.M.  
Deposited on : 2009-07-06  
Resolution : 1.90 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

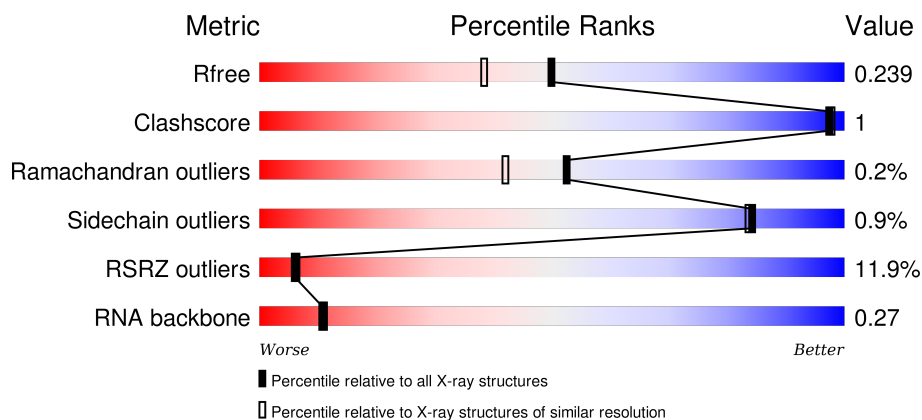
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.90 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	4755 (1.90-1.90)
Clashscore	102246	5398 (1.90-1.90)
Ramachandran outliers	100387	5338 (1.90-1.90)
Sidechain outliers	100360	5339 (1.90-1.90)
RSRZ outliers	91569	4766 (1.90-1.90)
RNA backbone	2183	1028 (2.70-1.10)

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	563	<div> <div>10%</div> <div> <div>87%</div> <div>10%</div> </div> </div>
2	B	10	<div> <div>40%</div> <div>60%</div> <div>30%</div> <div>10%</div> </div>

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 4529 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 ATP-dependent RNA helicase MSS116.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	509	4018	2573	670	759	16	0	5	0

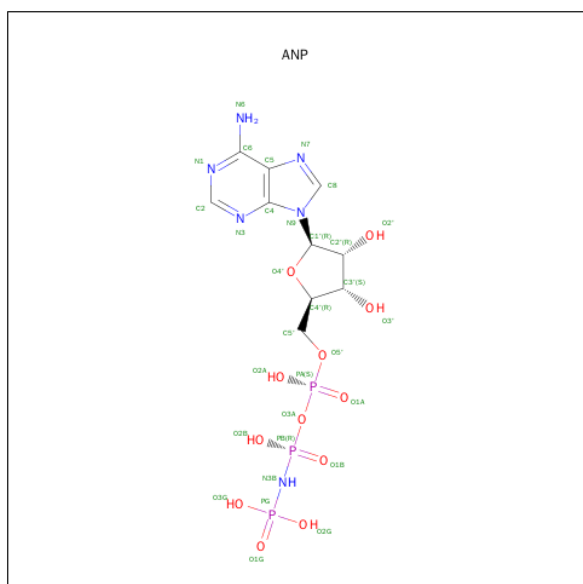
There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	35	GLY	-	EXPRESSION TAG	UNP P15424
A	36	SER	-	EXPRESSION TAG	UNP P15424

- Molecule 2 is a RNA chain called 5'-R(\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*U)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	B	10	190	86	19	76	9	0	0	0

- Molecule 3 is PHOSPHOAMINOPHOSPHONIC ACID-ADENYLATE ESTER (three-letter code: ANP) (formula: C<sub>10</sub>H<sub>17</sub>N<sub>6</sub>O<sub>12</sub>P<sub>3</sub>).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	P	0	0
			31	10	6	12	3		

- Molecule 4 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	1	Total	Mg	0	0
			1	1		

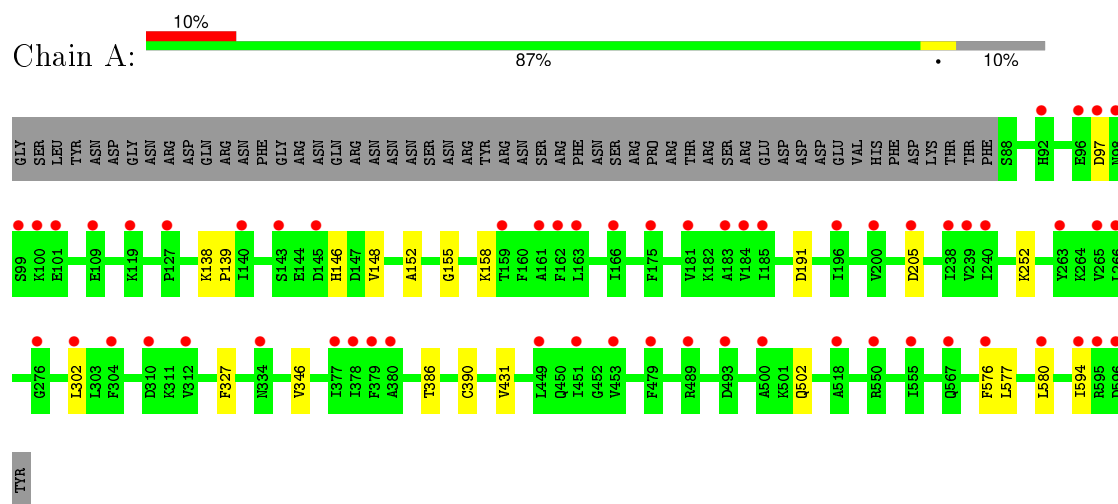
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	279	Total	O	0	0
			279	279		
5	B	10	Total	O	0	0
			10	10		

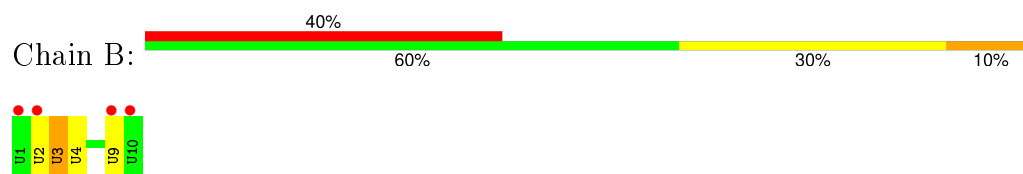
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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: ATP-dependent RNA helicase MSS116



- Molecule 2: 5'-R(\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*U)-3'



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	88.54Å 126.52Å 55.52Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	34.00 – 1.90 29.79 – 1.94	Depositor EDS
% Data completeness (in resolution range)	96.3 (34.00-1.90) 96.9 (29.79-1.94)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	5.52 (at 1.95Å)	Xtriage
Refinement program	REFMAC 5.5.0070	Depositor
R, $R_{free}$	0.191 , 0.225 0.206 , 0.239	Depositor DCC
$R_{free}$ test set	2276 reflections (5.30%)	DCC
Wilson B-factor (Å <sup>2</sup> )	27.2	Xtriage
Anisotropy	0.327	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 41.5	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	1 of 45214 reflections (0.002%)	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	4529	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	16.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.23% 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 [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: MG, ANP

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	0.40	0/4099	0.53	0/5529
2	B	0.84	0/208	1.19	1/320 (0.3%)
All	All	0.43	0/4307	0.59	1/5849 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	3	U	C3'-C2'-C1'	5.04	105.53	101.50

There are no chirality outliers.

There are no planarity outliers.

### 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	4018	0	4081	10	0
2	B	190	0	99	0	0
3	A	31	0	13	1	0
4	A	1	0	0	0	0
5	A	279	0	0	1	0
5	B	10	0	0	0	0
All	All	4529	0	4193	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:252:LYS:NZ	5:A:1155:HOH:O	2.29	0.65
1:A:155:GLY:H	3:A:1000:ANP:HNB1	1.47	0.60
1:A:577:LEU:HD22	1:A:594:ILE:HD11	1.86	0.58
1:A:390[A]:CYS:SG	1:A:431:VAL:HB	2.45	0.56
1:A:576:PHE:CE2	1:A:580:LEU:HD11	2.41	0.56
1:A:152:ALA:HB3	1:A:158:LYS:HD3	1.96	0.47
1:A:138:LYS:HB3	1:A:139:PRO:HD3	1.98	0.46
1:A:148:VAL:HB	1:A:302:LEU:HD22	2.00	0.43
1:A:346:VAL:HG12	1:A:502:GLN:HG2	2.00	0.43
1:A:386:THR:HG23	1:A:431:VAL:HG12	2.01	0.43

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	512/563 (91%)	506 (99%)	5 (1%)	1 (0%)	52 42

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	97	ASP

### 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	442/509 (87%)	438 (99%)	4 (1%)	84	83

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	146	HIS
1	A	191	ASP
1	A	205	ASP
1	A	327	PHE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	B	9/10 (90%)	4 (44%)	1 (11%)

All (4) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
2	B	2	U
2	B	3	U
2	B	4	U
2	B	9	U

All (1) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
2	B	3	U

## 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 [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 2 ligands modelled in this entry, 1 is monoatomic - leaving 1 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	ANP	A	1000	4	27,33,33	1.78	7 (25%)	30,52,52	2.24	8 (26%)

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	ANP	A	1000	4	-	0/12/38/38	0/3/3/3

All (7) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	1000	ANP	PG-O3G	-2.20	1.50	1.56
3	A	1000	ANP	PB-O2B	-2.04	1.51	1.56
3	A	1000	ANP	C5-C4	2.66	1.46	1.40
3	A	1000	ANP	PG-N3B	2.68	1.70	1.63
3	A	1000	ANP	PB-O1B	3.11	1.49	1.46
3	A	1000	ANP	PB-N3B	3.55	1.72	1.63
3	A	1000	ANP	PG-O1G	5.06	1.51	1.46

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	1000	ANP	N3-C2-N1	-8.27	122.56	128.89
3	A	1000	ANP	C1'-N9-C4	-4.04	120.84	126.94

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	1000	ANP	O3A-PB-N3B	-2.75	98.87	106.44
3	A	1000	ANP	O1G-PG-N3B	-2.11	108.66	111.90
3	A	1000	ANP	C4-C5-N7	-2.10	107.55	109.48
3	A	1000	ANP	O3G-PG-O2G	2.27	114.31	107.58
3	A	1000	ANP	C2-N1-C6	2.37	123.00	118.77
3	A	1000	ANP	O2B-PB-O1B	4.92	120.26	110.00

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	1000	ANP	1	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 ⓘ

### 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, 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	509/563 (90%)	0.70	58 (11%) 7 7	2, 16, 23, 28	0
2	B	10/10 (100%)	1.54	4 (40%) 0 0	3, 16, 37, 42	0
All	All	519/573 (90%)	0.71	62 (11%) 6 6	2, 16, 24, 42	0

All (62) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	98	ASN	8.2
1	A	185	ILE	5.5
1	A	596	ASP	5.2
2	B	9	U	4.9
2	B	1	U	4.7
1	A	594	ILE	4.6
1	A	310	ASP	4.3
1	A	97	ASP	4.3
1	A	184	VAL	4.3
1	A	99	SER	4.1
1	A	166	ILE	4.0
1	A	163	LEU	3.8
1	A	183	ALA	3.7
1	A	239	VAL	3.6
1	A	162	PHE	3.6
1	A	451	ILE	3.6
1	A	127	PRO	3.4
1	A	265	VAL	3.4
1	A	161	ALA	3.4
1	A	378	ILE	3.4
1	A	567	GLN	3.3
1	A	334	ASN	3.2
1	A	240	ILE	3.1
2	B	10	U	3.0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	238	ILE	2.9
1	A	96	GLU	2.8
1	A	175	PHE	2.8
1	A	140	ILE	2.7
1	A	449	LEU	2.7
2	B	2	U	2.6
1	A	489	ARG	2.5
1	A	379	PHE	2.5
1	A	580	LEU	2.5
1	A	276	GLY	2.5
1	A	312	VAL	2.5
1	A	302	LEU	2.5
1	A	479	PHE	2.4
1	A	377	ILE	2.4
1	A	518	ALA	2.4
1	A	493	ASP	2.4
1	A	145	ASP	2.4
1	A	101	GLU	2.4
1	A	304	PHE	2.4
1	A	453	VAL	2.4
1	A	109	GLU	2.3
1	A	159	THR	2.3
1	A	266	LEU	2.2
1	A	100	LYS	2.2
1	A	380	ALA	2.2
1	A	92	HIS	2.2
1	A	200	VAL	2.2
1	A	196	ILE	2.2
1	A	595	ARG	2.1
1	A	119	LYS	2.1
1	A	576	PHE	2.1
1	A	263	TYR	2.1
1	A	550	ARG	2.1
1	A	143	SER	2.1
1	A	555	ILE	2.1
1	A	500	ALA	2.1
1	A	205	ASP	2.1
1	A	181	VAL	2.0

## 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 [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. 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( $\text{\AA}^2$ )	Q<0.9
3	ANP	A	1000	31/31	0.97	0.11	-0.55	2,8,10,12	0
4	MG	A	900	1/1	0.94	0.14	-	20,20,20,20	0

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

There are no such residues in this entry.