



# Full wwPDB X-ray Structure Validation Report i

Feb 27, 2014 – 04:16 AM GMT

PDB ID : 2XT6  
Title : CRYSTAL STRUCTURE OF MYCOBACTERIUM SMEGMATIS ALPHA-KETOGLUTARATE DECARBOXYLASE HOMODIMER (ORTHORHOMBIC FORM)  
Authors : Wagner, T.; Bellinzoni, M.; Wehenkel, A.M.; O'Hare, H.M.; Alzari, P.M.  
Deposited on : 2010-10-05  
Resolution : 2.74 Å(reported)

This is a full wwPDB 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/ValidationPDFNotes.html>

---

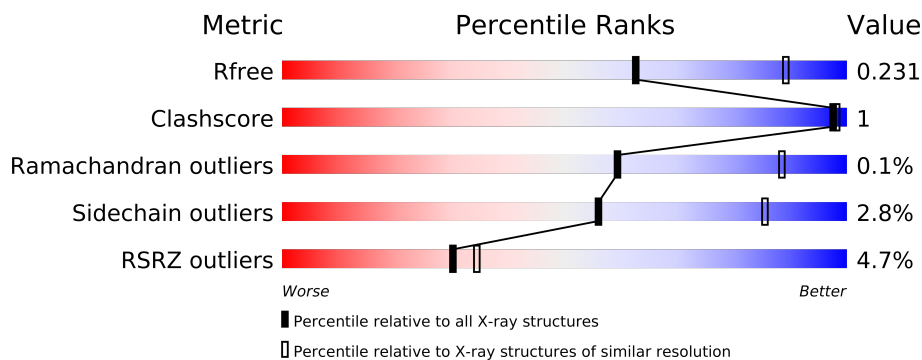
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
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 2.74 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	2164 (2.78-2.70)
Clashscore	79885	2639 (2.78-2.70)
Ramachandran outliers	78287	2594 (2.78-2.70)
Sidechain outliers	78261	2595 (2.78-2.70)
RSRZ outliers	66119	2166 (2.78-2.70)

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. 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	1113	
1	B	1113	

## 2 Entry composition i

There are 5 unique types of molecules in this entry. The entry contains 15902 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.

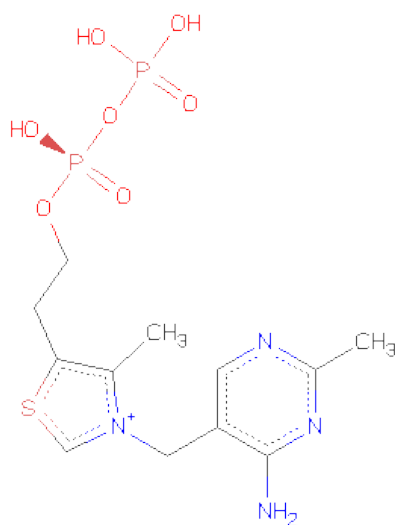
- Molecule 1 is a protein called 2-OXOGLUTARATE DECARBOXYLASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1055	Total	C	N	O	S	0	0	0
			8153	5136	1447	1540	30			
1	B	989	Total	C	N	O	S	0	0	0
			7574	4776	1342	1429	27			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	115	GLY	-	EXPRESSION TAG	UNP A0R2B1
B	115	GLY	-	EXPRESSION TAG	UNP A0R2B1

- Molecule 2 is THIAMINE DIPHOSPHATE (three-letter code: TPP) (formula: C<sub>12</sub>H<sub>19</sub>N<sub>4</sub>O<sub>7</sub>P<sub>2</sub>S).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	P	0	0
			26	12	4	7	2		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	B	1	Total	C	N	O	P	S	
			26	12	4	7	2	1	
								0	0

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

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	1	Total	Mg		
			1	1	0	0
3	A	1	Total	Mg		
			1	1	0	0

- Molecule 4 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	1	Total	Ca		
			1	1	0	0
4	A	1	Total	Ca		
			1	1	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	56	Total	O		
			56	56	0	0
5	B	63	Total	O		
			63	63	0	0



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	151.99Å 247.72Å 79.98Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	40.12 – 2.74 72.65 – 2.74	Depositor EDS
% Data completeness (in resolution range)	99.1 (40.12-2.74) 99.1 (72.65-2.74)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.14 (at 2.73Å)	Xtriage
Refinement program	BUSTER-TNT 2.9.3	Depositor
R, $R_{free}$	0.185 , 0.224 0.193 , 0.231	Depositor DCC
$R_{free}$ test set	4027 reflections (5.32%)	DCC
Wilson B-factor (Å <sup>2</sup> )	57.9	Xtriage
Anisotropy	0.620	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 50.5	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 79783 reflections	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	15902	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	70.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.45% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

## 5 Model quality

### 5.1 Standard geometry

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

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	0.51	0/8319	0.67	0/11288
1	B	0.51	0/7721	0.67	0/10472
All	All	0.51	0/16040	0.67	0/21760

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 5.2 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 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	8153	0	0	9	0
1	B	7574	0	0	3	0
2	A	26	0	0	0	0
2	B	26	0	0	0	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
4	A	1	0	0	0	0
4	B	1	0	0	0	0
5	A	56	0	0	0	0
5	B	63	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	15902	0	0	12	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 1.

All (12) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:1095:GLU:OE1	1:A:1106:ARG:NH1	2.30	0.64
1:A:523:ASP:OD2	1:A:722:ARG:NH2	2.39	0.55
1:B:703:GLY:O	1:B:733:LYS:NZ	2.46	0.49
1:A:703:GLY:O	1:A:733:LYS:NZ	2.47	0.47
1:A:607:GLU:OE1	1:A:643:HIS:ND1	2.49	0.45
1:A:446:HIS:ND1	1:A:713:ASP:OD2	2.51	0.44
1:B:1111:ARG:NH1	1:B:1225:ALA:O	2.52	0.43
1:B:744:ARG:NH1	1:B:767:LYS:O	2.52	0.43
1:A:284:GLU:O	1:A:305:LEU:N	2.53	0.41
1:A:193:HIS:ND1	1:A:220:GLU:OE2	2.54	0.41
1:A:744:ARG:NH1	1:A:767:LYS:O	2.55	0.40
1:A:150:THR:O	1:A:781:ARG:NH1	2.55	0.40

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	1045/1113 (94%)	1010 (97%)	34 (3%)	1 (0%)	59 88
1	B	973/1113 (87%)	938 (96%)	33 (3%)	2 (0%)	56 85
All	All	2018/2226 (91%)	1948 (96%)	67 (3%)	3 (0%)	59 88

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	399	SER

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	682	PHE
1	B	682	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	838/926 (90%)	813 (97%)	25 (3%)	53	84
1	B	767/926 (83%)	747 (97%)	20 (3%)	59	87
All	All	1605/1852 (87%)	1560 (97%)	45 (3%)	56	85

All (45) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	158	THR
1	A	161	LEU
1	A	198	LEU
1	A	233	ASP
1	A	253	ILE
1	A	293	SER
1	A	306	ILE
1	A	314	HIS
1	A	316	ILE
1	A	341	GLU
1	A	367	ASN
1	A	394	ASN
1	A	467	VAL
1	A	471	HIS
1	A	682	PHE
1	A	740	LEU
1	A	868	ARG
1	A	950	LEU
1	A	951	SER
1	A	953	PHE
1	A	976	GLN
1	A	1112	LEU
1	A	1114	LEU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	1188	LEU
1	A	1220	GLU
1	B	158	THR
1	B	228	ILE
1	B	234	ILE
1	B	341	GLU
1	B	367	ASN
1	B	402	ASP
1	B	427	VAL
1	B	432	LEU
1	B	467	VAL
1	B	506	PHE
1	B	682	PHE
1	B	868	ARG
1	B	950	LEU
1	B	951	SER
1	B	953	PHE
1	B	976	GLN
1	B	1112	LEU
1	B	1114	LEU
1	B	1187	ILE
1	B	1220	GLU

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

### 5.3.3 RNA ⓘ

There are no RNA chains 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 6 ligands modelled in this entry, 4 are monoatomic - leaving 2 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	TPP	A	2001	3	27,27,27	1.99	5 (18%)	40,40,40	1.84	10 (25%)
2	TPP	B	2001	3	27,27,27	2.12	5 (18%)	40,40,40	1.83	11 (27%)

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	TPP	A	2001	3	-	0/17/17/17	0/2/2/2
2	TPP	B	2001	3	-	0/17/17/17	0/2/2/2

All (10) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	2001	TPP	C4-N3	-6.64	1.32	1.39
2	A	2001	TPP	C4-N3	-5.69	1.33	1.39
2	B	2001	TPP	PB-O3A	-5.56	1.50	1.60
2	A	2001	TPP	PB-O3A	-4.89	1.51	1.60
2	A	2001	TPP	C5-S1	3.77	1.76	1.73
2	A	2001	TPP	PB-O1B	3.13	1.61	1.51
2	B	2001	TPP	PB-O1B	2.91	1.61	1.51
2	B	2001	TPP	C5-S1	2.66	1.75	1.73
2	B	2001	TPP	PB-O3B	2.44	1.63	1.54
2	A	2001	TPP	PB-O3B	2.03	1.62	1.54

All (21) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	2001	TPP	CM2-C2'-N1'	5.10	123.10	117.02

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	2001	TPP	CM2-C2'-N1'	4.71	122.63	117.02
2	A	2001	TPP	C5-C4-N3	4.59	116.86	107.53
2	A	2001	TPP	CM4-C4-C5	-4.54	117.76	129.10
2	B	2001	TPP	CM4-C4-C5	-4.34	118.26	129.10
2	B	2001	TPP	C5-C4-N3	4.23	116.12	107.53
2	A	2001	TPP	C6'-N1'-C2'	3.22	121.34	115.68
2	B	2001	TPP	C2-S1-C5	-3.08	89.56	91.63
2	A	2001	TPP	C4-C5-S1	-2.98	107.49	109.93
2	B	2001	TPP	C6'-N1'-C2'	2.90	120.77	115.68
2	A	2001	TPP	C2-S1-C5	-2.77	89.77	91.63
2	A	2001	TPP	N1'-C2'-N3'	-2.70	120.78	125.65
2	A	2001	TPP	C6-C5-C4	2.70	129.40	127.44
2	B	2001	TPP	N1'-C2'-N3'	-2.54	121.07	125.65
2	B	2001	TPP	C4-C5-S1	-2.45	107.93	109.93
2	B	2001	TPP	C6-C5-C4	2.44	129.22	127.44
2	B	2001	TPP	CM4-C4-N3	2.37	125.62	122.53
2	B	2001	TPP	C5'-C6'-N1'	-2.34	119.63	123.86
2	A	2001	TPP	C5'-C6'-N1'	-2.13	120.01	123.86
2	A	2001	TPP	CM4-C4-N3	2.13	125.31	122.53
2	B	2001	TPP	C2'-N3'-C4'	2.01	121.98	118.18

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 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, 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	1055/1113 (94%)	0.17	35 (3%)	44 51	39, 65, 102, 140	0
1	B	989/1113 (88%)	0.31	61 (6%)	20 23	41, 65, 140, 173	0
All	All	2044/2226 (91%)	0.24	96 (4%)	30 34	39, 65, 115, 173	0

All (96) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	214	ALA	9.1
1	B	302	ILE	8.4
1	B	303	GLY	7.9
1	B	282	ALA	7.6
1	B	136	MET	7.3
1	B	325	PHE	7.2
1	B	281	GLY	6.9
1	B	213	VAL	6.7
1	B	329	ILE	6.6
1	B	238	ALA	6.5
1	B	309	THR	5.3
1	A	212	VAL	5.2
1	B	308	LEU	5.2
1	B	306	ILE	5.0
1	A	561	PHE	4.7
1	B	133	ALA	4.6
1	B	333	LEU	4.5
1	A	122	VAL	4.3
1	B	132	PRO	4.3
1	B	297	ILE	4.2
1	B	233	ASP	4.2
1	B	310	SER	4.2
1	B	304	LYS	4.1
1	A	198	LEU	4.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	247	ASP	4.1
1	B	164	ALA	4.1
1	B	280	ALA	4.1
1	B	130	ALA	4.0
1	B	231	TYR	3.9
1	B	255	LEU	3.9
1	B	234	ILE	3.9
1	B	307	THR	3.7
1	B	810	LEU	3.7
1	B	237	ARG	3.6
1	B	300	LEU	3.6
1	B	212	VAL	3.4
1	B	327	ARG	3.4
1	A	240	ASP	3.3
1	B	215	ALA	3.3
1	A	289	PHE	3.3
1	B	199	ALA	3.2
1	A	590	PHE	3.2
1	A	199	ALA	3.1
1	B	503	GLN	3.1
1	A	476	VAL	3.0
1	B	305	LEU	3.0
1	A	308	LEU	3.0
1	B	301	GLY	2.9
1	B	160	LEU	2.9
1	A	121	GLU	2.9
1	B	134	LYS	2.9
1	B	227	PHE	2.8
1	B	1123	LEU	2.8
1	A	200	ILE	2.8
1	B	1135	VAL	2.8
1	A	213	VAL	2.7
1	B	1112	LEU	2.6
1	A	258	PRO	2.6
1	B	131	ILE	2.6
1	B	279	GLY	2.6
1	A	128	VAL	2.6
1	B	283	MET	2.6
1	A	560	GLU	2.6
1	B	248	PHE	2.5
1	B	157	PHE	2.5
1	A	631	SER	2.4

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	554	TYR	2.4
1	A	291	GLY	2.4
1	B	239	ARG	2.4
1	A	326	LEU	2.4
1	B	326	LEU	2.4
1	B	330	HIS	2.3
1	A	189	ILE	2.3
1	B	235	VAL	2.3
1	A	634	ARG	2.3
1	A	231	TYR	2.3
1	B	355	TRP	2.3
1	A	810	LEU	2.3
1	A	179	PHE	2.2
1	A	174	ASN	2.2
1	A	784	ILE	2.2
1	B	196	LEU	2.2
1	A	123	PRO	2.1
1	B	502	GLY	2.1
1	A	130	ALA	2.1
1	A	589	MET	2.1
1	B	253	ILE	2.1
1	A	483	LEU	2.1
1	A	783	ASP	2.1
1	B	321	GLU	2.1
1	A	774	TYR	2.1
1	A	235	VAL	2.1
1	B	159	HIS	2.0
1	B	320	ALA	2.0
1	B	419	PHE	2.0
1	A	499	LYS	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 ⓘ

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, 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	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
4	CA	B	2003	1/1	0.16	0.54	67,67,67,67	0
3	MG	B	2002	1/1	0.23	0.51	32,32,32,32	0
2	TPP	A	2001	26/26	0.16	-0.13	35,51,66,74	0
2	TPP	B	2001	26/26	0.18	-0.25	23,41,70,73	0
3	MG	A	2002	1/1	0.13	-0.63	41,41,41,41	0
4	CA	A	2003	1/1	0.10	-1.61	72,72,72,72	0

## 6.5 Other polymers

There are no such residues in this entry.