



wwPDB X-ray Structure Validation Summary Report ⓘ

May 21, 2020 – 06:11 pm BST

PDB ID : 4FAU
Title : Structure of *Oceanobacillus iheyensis* group II intron in the presence of Li⁺,
Mg²⁺ and 5'-exon
Authors : Marcia, M.; Pyle, A.M.
Deposited on : 2012-05-22
Resolution : 2.87 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 ⓘ symbol.

The following versions of software and data (see [references ⓘ](#)) 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	:	2.11
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)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

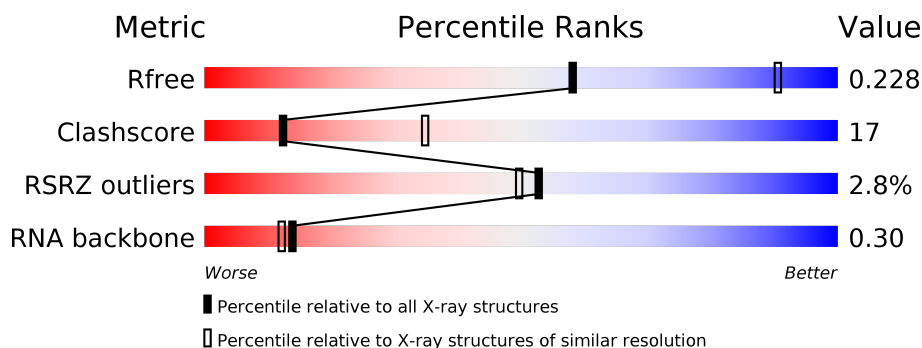
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.87 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}	130704	2691 (2.90-2.86)
Clashscore	141614	2947 (2.90-2.86)
RSRZ outliers	127900	2629 (2.90-2.86)
RNA backbone	3102	1121 (3.16-2.60)

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 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 $\leq 5\%$. 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	398	

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
2	MG	A	401	-	-	-	X
2	MG	A	402	-	-	-	X
2	MG	A	409	-	-	-	X

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	MG	A	414	-	-	-	X
2	MG	A	419	-	-	-	X
2	MG	A	420	-	-	-	X
2	MG	A	424	-	-	-	X
3	SPM	A	425	-	-	-	X
4	EPE	A	427	-	X	-	-

2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 8581 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.

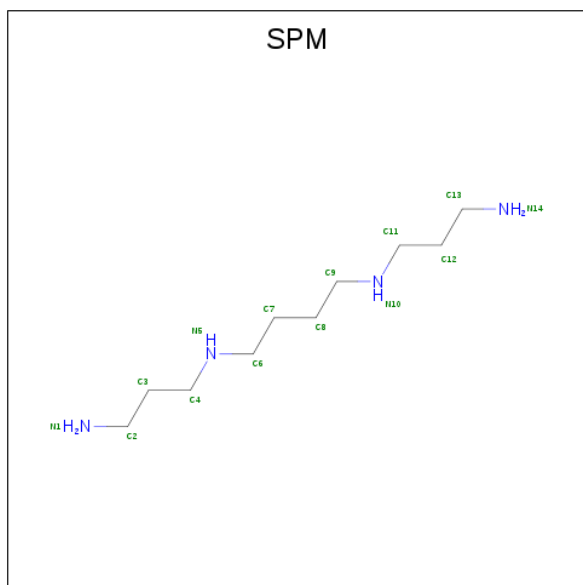
- Molecule 1 is a RNA chain called Group IIC intron.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	393	Total	C	N	O	P	0	2	0
			8466	3777	1568	2726	395			

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

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	24	Total	Mg	0	0
			24	24		

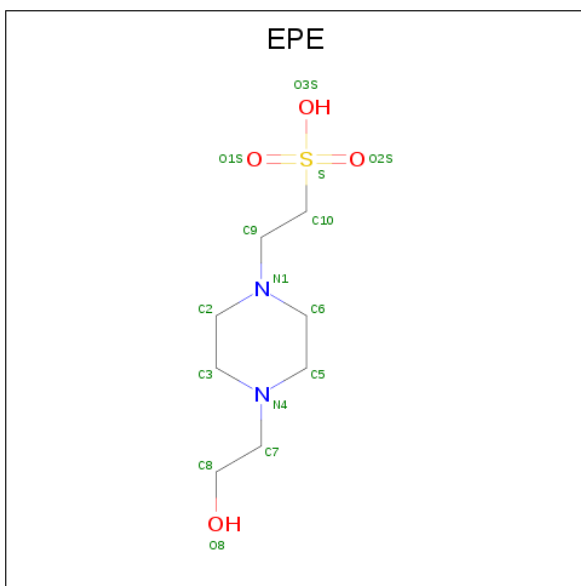
- Molecule 3 is SPERMINE (three-letter code: SPM) (formula: C₁₀H₂₆N₄).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	N	0	0
			14	10	4		

- Molecule 4 is 4-(2-HYDROXYETHYL)-1-PIPERAZINE ETHANESULFONIC ACID

(three-letter code: EPE) (formula: C₈H₁₈N₂O₄S).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	A	1	Total	C	N	O	S	0	0
			12	6	2	3	1		
4	A	1	Total	C	O	S		0	0
			5	1	3	1			
4	A	1	Total	C	O	S		0	0
			5	1	3	1			

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	55	Total	O	0	0
			55	55		

These plots are drawn for all protein, RNA and DNA 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 ($\text{RSRZ} > 2$). 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.

Chain A:

3% 29% 33% 30% 7%

Index	Value	Category
A346	C346	Green
A347	A347	Green
A348	A348	Green
A349	A349	Green
A350	A350	Green
A351	A351	Green
A352	A352	Green
A353	A353	Green
A354	A354	Green
A355	A355	Green
A356	A356	Green
A357	A357	Green
A358	A358	Green
A359	A359	Green
A360	A360	Green
A361	A361	Green
A362	A362	Green
A363	A363	Green
A366	A366	Green
A367	A367	Green
A368	A368	Green
A369	A369	Green
A370	A370	Green
A371	A371	Green
A372	A372	Green
A373	A373	Green
A374	A374	Green
A375	A375	Green
A376	A376	Green
A377	A377	Green
A378	A378	Green
A379	A379	Green
A380	A380	Green
A381	A381	Green
A382	A382	Green
A383	A383	Green
A384	A384	Green
A385	A385	Green
A386	A386	Green
A387	A387	Green
A388	A388	Green
A389	A389	Green
A390	A390	Green
A391	A391	Green
A392	A392	Green
A393	A393	Green
A394	A394	Green
A395	A395	Green
A396	A396	Green
A397	A397	Green
A398	A398	Green
A399	A399	Green
A400	A400	Green
A401	A401	Green
A402	A402	Green
A403	A403	Green
A404	A404	Green
A405	A405	Green
A406	A406	Green
A407	A407	Green
A408	A408	Green
A409	A409	Green
A410	A410	Green
A411	A411	Green
A412	A412	Green
A413	A413	Green
A414	A414	Green
A415	A415	Green
A416	A416	Green
A417	A417	Green
A418	A418	Green
A419	A419	Green
A420	A420	Green
A421	A421	Green
A422	A422	Green
A423	A423	Green
A424	A424	Green
A425	A425	Green
A426	A426	Green
A427	A427	Green
A428	A428	Green
A429	A429	Green
A430	A430	Green
A431	A431	Green
A432	A432	Green
A433	A433	Green
A434	A434	Green
A435	A435	Green
A436	A436	Green
A437	A437	Green
A438	A438	Green
A439	A439	Green
A440	A440	Green
A441	A441	Green
A442	A442	Green
A443	A443	Green
A444	A444	Green
A445	A445	Green
A446	A446	Green
A447	A447	Green
A448	A448	Green
A449	A449	Green
A450	A450	Green
A451	A451	Green
A452	A452	Green
A453	A453	Green
A454	A454	Green
A455	A455	Green
A456	A456	Green
A457	A457	Green
A458	A458	Green
A459	A459	Green
A460	A460	Green
A461	A461	Green
A462	A462	Green
A463	A463	Green
A464	A464	Green
A465	A465	Green
A466	A466	Green
A467	A467	Green
A468	A468	Green
A469	A469	Green
A470	A470	Green
A471	A471	Green
A472	A472	Green
A473	A473	Green
A474	A474	Green
A475	A475	Green
A476	A476	Green
A477	A477	Green
A478	A478	Green
A479	A479	Green
A480	A480	Green
A481	A481	Green
A482	A482	Green
A483	A483	Green
A484	A484	Green
A485	A485	Green
A486	A486	Green
A487	A487	Green
A488	A488	Green
A489	A489	Green
A490	A490	Green
A491	A491	Green
A492	A492	Green
A493	A493	Green
A49		

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	89.80 Å 95.83 Å 227.30 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.88 – 2.87 48.88 – 2.87	Depositor EDS
% Data completeness (in resolution range)	98.1 (48.88-2.87) 98.1 (48.88-2.87)	Depositor EDS
R_{merge}	0.04	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.65 (at 2.86 Å)	Xtriage
Refinement program	REFMAC 5.6.0117	Depositor
R, R_{free}	0.196 , 0.232 0.194 , 0.228	Depositor DCC
R_{free} test set	2262 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å ²)	85.3	Xtriage
Anisotropy	0.374	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.22 , 50.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	8581	wwPDB-VP
Average B, all atoms (Å ²)	102.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.06% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 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, EPE, SPM

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	1.97	4/9495 (0.0%)	1.97	251/14809 (1.7%)

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	289[A]	C	C1'-N1	130.81	3.44	1.48
1	A	289[B]	C	C1'-N1	130.81	3.44	1.48
1	A	69	C	O3'-P	5.45	1.67	1.61
1	A	168	U	O3'-P	-5.11	1.55	1.61

The worst 5 of 251 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	289[A]	C	O4'-C1'-N1	-97.86	29.91	108.20
1	A	289[B]	C	O4'-C1'-N1	-97.86	29.91	108.20
1	A	289[A]	C	C2-N1-C1'	47.04	170.55	118.80
1	A	289[B]	C	C2-N1-C1'	47.04	170.55	118.80
1	A	289[A]	C	C6-N1-C1'	-46.65	64.82	120.80

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	8466	0	4250	209	0
2	A	24	0	0	0	0
3	A	14	0	26	0	0
4	A	22	0	15	5	0
5	A	55	0	0	8	0
All	All	8581	0	4291	209	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 17.

The worst 5 of 209 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:288[A]:G:N3	1:A:288[A]:G:H5''	1.33	1.41
1:A:66:U:H5'	5:A:535:HOH:O	1.23	1.33
1:A:288[A]:G:C5'	1:A:288[A]:G:N3	1.98	1.25
1:A:39:C:H2'	1:A:40:C:H5''	1.13	1.10
1:A:39:C:C2'	1:A:40:C:H5''	1.83	1.09

There are no symmetry-related clashes.

5.3 Torsion angles [i](#)

5.3.1 Protein backbone [i](#)

There are no protein molecules in this entry.

5.3.2 Protein sidechains [i](#)

There are no protein molecules in this entry.

5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	A	388/398 (97%)	112 (28%)	20 (5%)

5 of 112 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	A	5	G
1	A	7	C
1	A	12	A
1	A	13	U
1	A	16	G

5 of 20 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
1	A	167	A
1	A	168	U
1	A	320	G
1	A	165	C
1	A	166	A

5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

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 28 ligands modelled in this entry, 24 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
4	EPE	A	428	-	4,4,15	1.57	1 (25%)	5,6,20	1.16	1 (20%)
3	SPM	A	425	-	13,13,13	0.59	0	12,12,12	1.24	2 (16%)
4	EPE	A	426	-	12,12,15	2.09	2 (16%)	14,16,20	5.51	6 (42%)
4	EPE	A	427	-	4,4,15	1.45	1 (25%)	5,6,20	5.34	4 (80%)

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	SPM	A	425	-	-	9/11/11/11	-
4	EPE	A	426	-	-	3/6/14/19	0/1/1/1

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	426	EPE	O3S-S	5.06	1.65	1.47
4	A	426	EPE	C10-S	-4.43	1.71	1.77
4	A	428	EPE	O3S-S	2.72	1.65	1.48
4	A	427	EPE	O3S-S	2.37	1.63	1.48

The worst 5 of 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	426	EPE	O3S-S-O2S	-10.85	84.77	111.27
4	A	426	EPE	O3S-S-O1S	-10.51	85.59	111.27
4	A	426	EPE	O1S-S-C10	9.59	118.46	106.92
4	A	427	EPE	O1S-S-C10	9.37	127.06	107.54
4	A	426	EPE	O3S-S-C10	-7.66	93.37	105.77

There are no chirality outliers.

5 of 12 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	425	SPM	C12-C11-N10-C9
3	A	425	SPM	C2-C3-C4-N5
3	A	425	SPM	C7-C8-C9-N10
3	A	425	SPM	N10-C11-C12-C13
3	A	425	SPM	C3-C4-N5-C6

There are no ring outliers.

2 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	426	EPE	1	0
4	A	427	EPE	4	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, 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	393/398 (98%)	0.69	11 (2%) 53 50	43, 86, 194, 317	1 (0%)

The worst 5 of 11 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	311	G	3.3
1	A	161	G	3.2
1	A	287	A	3.2
1	A	207	A	3.0
1	A	83	G	2.9

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](#)

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. 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
3	SPM	A	425	14/14	0.58	0.55	37,45,58,59	14
2	MG	A	402	1/1	0.60	0.84	106,106,106,106	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
2	MG	A	409	1/1	0.63	0.71	105,105,105,105	0
2	MG	A	420	1/1	0.68	1.07	105,105,105,105	0
2	MG	A	405	1/1	0.76	0.23	61,61,61,61	0
2	MG	A	401	1/1	0.78	0.85	71,71,71,71	0
2	MG	A	414	1/1	0.78	0.80	72,72,72,72	0
2	MG	A	419	1/1	0.79	0.71	90,90,90,90	0
4	EPE	A	426	12/15	0.79	0.34	74,80,107,116	12
2	MG	A	424	1/1	0.80	0.43	73,73,73,73	0
2	MG	A	415	1/1	0.82	0.61	81,81,81,81	0
2	MG	A	407	1/1	0.85	0.27	80,80,80,80	0
2	MG	A	418	1/1	0.86	0.31	50,50,50,50	0
2	MG	A	417	1/1	0.87	0.44	69,69,69,69	0
2	MG	A	423	1/1	0.87	0.17	112,112,112,112	0
2	MG	A	406	1/1	0.87	0.42	110,110,110,110	0
2	MG	A	421	1/1	0.89	0.28	73,73,73,73	0
2	MG	A	412	1/1	0.90	0.17	68,68,68,68	0
2	MG	A	413	1/1	0.92	0.72	68,68,68,68	0
2	MG	A	422	1/1	0.92	0.58	76,76,76,76	0
2	MG	A	403	1/1	0.92	0.11	91,91,91,91	0
2	MG	A	404	1/1	0.92	0.25	92,92,92,92	0
4	EPE	A	428	5/15	0.92	0.37	56,69,74,84	5
2	MG	A	411	1/1	0.94	0.16	36,36,36,36	0
2	MG	A	410	1/1	0.95	0.47	69,69,69,69	0
4	EPE	A	427	5/15	0.95	0.30	44,54,65,68	5
2	MG	A	416	1/1	0.97	0.39	79,79,79,79	0
2	MG	A	408	1/1	0.98	0.29	54,54,54,54	0

6.5 Other polymers ⓘ

There are no such residues in this entry.