



wwPDB X-ray Structure Validation Summary Report ⓘ

May 14, 2020 – 04:31 pm BST

PDB ID : 6GZB
Title : Tandem GerMN domains of the sporulation protein GerM from *Bacillus subtilis*
Authors : Trouve, J.; Mohamed, A.; Leisico, F.; Contreras-Martel, C.; Liu, B.; Mas, C.; Rudner, D.Z.; Rodrigues, C.D.A.; Morlot, C.
Deposited on : 2018-07-03
Resolution : 2.10 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

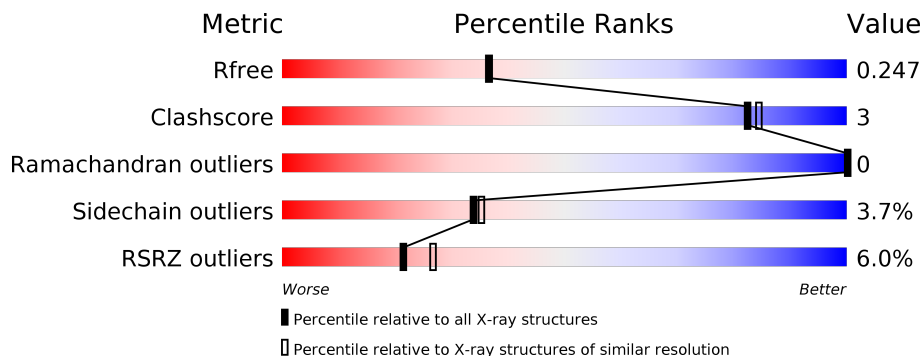
1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.10 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	5197 (2.10-2.10)
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)
RSRZ outliers	127900	5083 (2.10-2.10)

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 ≥ 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	341	<div style="display: flex; align-items: center;"> <div style="width: 100%; height: 15px; background: linear-gradient(to right, red 1%, orange 1%, yellow 6%, green 78%, grey 16%);"></div> <div style="margin-left: 10px;"> <p>%</p> <p>78% 6% 16%</p> </div> </div>
1	B	341	<div style="display: flex; align-items: center;"> <div style="width: 100%; height: 15px; background: linear-gradient(to right, red 1%, orange 1%, yellow 6%, green 77%, grey 17%);"></div> <div style="margin-left: 10px;"> <p>%</p> <p>77% 6% 17%</p> </div> </div>
1	C	341	<div style="display: flex; align-items: center;"> <div style="width: 100%; height: 15px; background: linear-gradient(to right, red 9%, orange 1%, yellow 8%, green 75%, grey 17%);"></div> <div style="margin-left: 10px;"> <p>9%</p> <p>75% 8% 17%</p> </div> </div>
1	D	341	<div style="display: flex; align-items: center;"> <div style="width: 100%; height: 15px; background: linear-gradient(to right, red 8%, orange 1%, yellow 6%, green 77%, grey 17%);"></div> <div style="margin-left: 10px;"> <p>8%</p> <p>77% 6% 17%</p> </div> </div>

2 Entry composition [i](#)

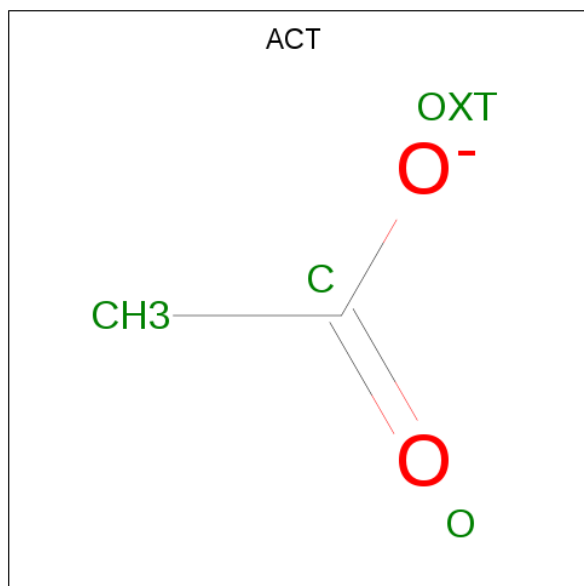
There are 4 unique types of molecules in this entry. The entry contains 9144 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 Spore germination protein GerM.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	287	Total 2218	C 1396	N 365	O 454	S 3	0	0	0
1	B	282	Total 2185	C 1376	N 360	O 446	S 3	0	0	0
1	C	284	Total 2199	C 1386	N 362	O 448	S 3	0	0	0
1	D	282	Total 2174	C 1366	N 359	O 446	S 3	0	0	0

- Molecule 2 is ACETATE ION (three-letter code: ACT) (formula: C₂H₃O₂).



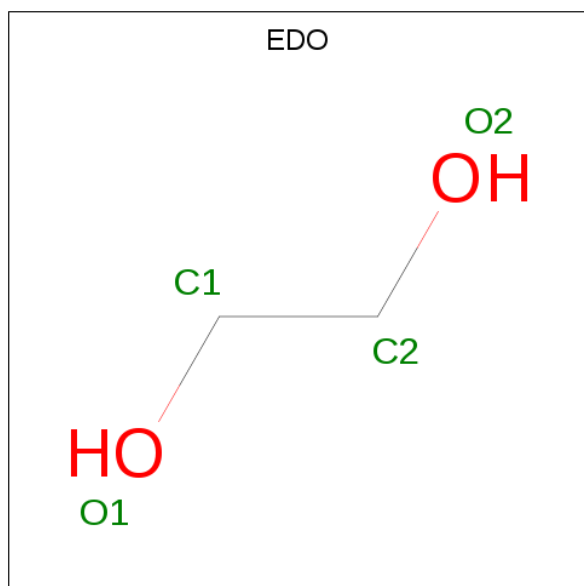
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
2	A	1	Total 4	C 2	O 2	0	0
2	A	1	Total 4	C 2	O 2	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total C O 4 2 2	0	0
2	B	1	Total C O 4 2 2	0	0
2	B	1	Total C O 4 2 2	0	0
2	B	1	Total C O 4 2 2	0	0
2	C	1	Total C O 4 2 2	0	0
2	D	1	Total C O 4 2 2	0	0
2	D	1	Total C O 4 2 2	0	0
2	D	1	Total C O 4 2 2	0	0

- Molecule 3 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C₂H₆O₂).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C O 4 2 2	0	0
3	A	1	Total C O 4 2 2	0	0
3	B	1	Total C O 4 2 2	0	0

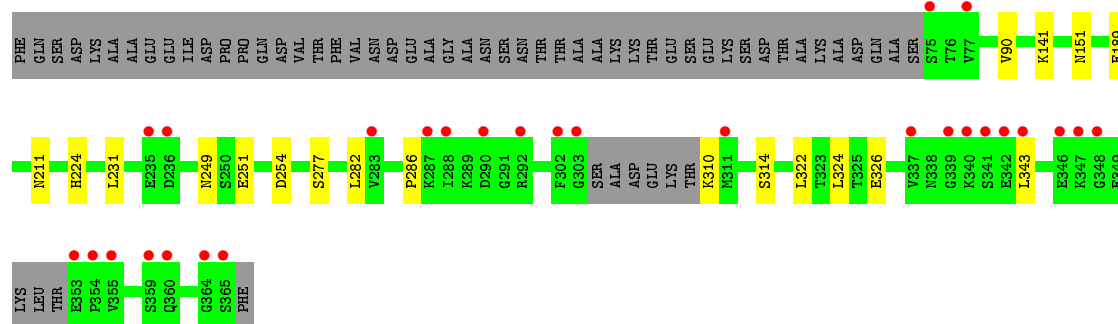
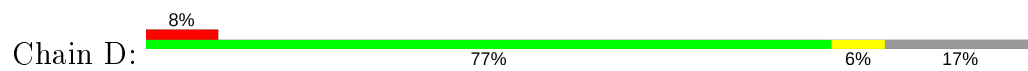
Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	B	1	Total	C	O	0	0
			4	2	2		
3	C	1	Total	C	O	0	0
			4	2	2		
3	D	1	Total	C	O	0	0
			4	2	2		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	124	Total	O	0	0
			124	124		
4	B	61	Total	O	0	0
			61	61		
4	C	33	Total	O	0	0
			33	33		
4	D	86	Total	O	0	0
			86	86		



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	90.36Å 103.94Å 162.49Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.46 – 2.10 46.46 – 2.10	Depositor EDS
% Data completeness (in resolution range)	99.8 (46.46-2.10) 99.8 (46.46-2.10)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.06	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.29 (at 2.10Å)	Xtrriage
Refinement program	REFMAC 5.8.0189	Depositor
R, R_{free}	0.204 , 0.236 0.215 , 0.247	Depositor DCC
R_{free} test set	3583 reflections (3.99%)	wwPDB-VP
Wilson B-factor (Å ²)	39.2	Xtrriage
Anisotropy	0.026	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 50.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	9144	wwPDB-VP
Average B, all atoms (Å ²)	50.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.86% 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: EDO, ACT

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	0.43	0/2250	0.50	0/3049
1	B	0.60	0/2216	0.53	0/3004
1	C	0.62	0/2231	0.56	0/3023
1	D	0.55	1/2204 (0.0%)	0.52	0/2985
All	All	0.56	1/8901 (0.0%)	0.53	0/12061

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	189	GLU	C-N	-5.83	1.20	1.34

There are no bond angle outliers.

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	2218	0	2245	7	0
1	B	2185	0	2218	14	0
1	C	2199	0	2228	26	0
1	D	2174	0	2195	7	0
2	A	12	0	9	0	0
2	B	12	0	9	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	C	4	0	3	0	0
2	D	12	0	9	0	0
3	A	8	0	12	0	0
3	B	8	0	12	0	0
3	C	4	0	6	0	0
3	D	4	0	6	0	0
4	A	124	0	0	0	0
4	B	61	0	0	0	0
4	C	33	0	0	0	0
4	D	86	0	0	1	0
All	All	9144	0	8952	45	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 3.

The worst 5 of 45 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:246:ARG:HH12	1:C:269:VAL:CG1	1.72	1.02
1:B:246:ARG:NH1	1:C:269:VAL:CG1	2.32	0.92
1:B:246:ARG:HH12	1:C:269:VAL:HG12	1.34	0.92
1:B:246:ARG:NH1	1:C:269:VAL:HG11	1.85	0.91
1:B:90:VAL:H	1:B:211:ASN:HD21	1.21	0.88

There are no symmetry-related clashes.

5.3 Torsion angles [i](#)

5.3.1 Protein backbone [i](#)

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	283/341 (83%)	279 (99%)	4 (1%)	0	100	100
1	B	278/341 (82%)	274 (99%)	4 (1%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	280/341 (82%)	271 (97%)	9 (3%)	0	100	100
1	D	276/341 (81%)	267 (97%)	9 (3%)	0	100	100
All	All	1117/1364 (82%)	1091 (98%)	26 (2%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains [i](#)

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	258/300 (86%)	246 (95%)	12 (5%)	26	25
1	B	254/300 (85%)	245 (96%)	9 (4%)	36	38
1	C	255/300 (85%)	245 (96%)	10 (4%)	32	33
1	D	252/300 (84%)	245 (97%)	7 (3%)	43	47
All	All	1019/1200 (85%)	981 (96%)	38 (4%)	34	35

5 of 38 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	269	VAL
1	C	79	ARG
1	D	314	SER
1	B	350	LYS
1	C	85	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 18 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	87	ASN
1	C	151	ASN
1	D	87	ASN
1	B	211	ASN
1	B	224	HIS

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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](#)

16 ligands are modelled in this entry.

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	ACT	D	402	-	1,3,3	1.06	0	0,3,3	0.00	-
3	EDO	B	405	-	3,3,3	0.44	0	2,2,2	0.35	0
2	ACT	B	402	-	1,3,3	1.36	0	0,3,3	0.00	-
2	ACT	A	402	-	1,3,3	1.31	0	0,3,3	0.00	-
2	ACT	D	401	-	1,3,3	1.42	0	0,3,3	0.00	-
2	ACT	D	403	-	1,3,3	0.82	0	0,3,3	0.00	-
2	ACT	B	403	-	1,3,3	0.85	0	0,3,3	0.00	-
3	EDO	C	402	-	3,3,3	0.27	0	2,2,2	0.47	0
2	ACT	A	403	-	1,3,3	0.25	0	0,3,3	0.00	-
3	EDO	D	404	-	3,3,3	0.50	0	2,2,2	0.28	0
3	EDO	A	405	-	3,3,3	0.48	0	2,2,2	0.25	0
3	EDO	A	404	-	3,3,3	0.49	0	2,2,2	0.18	0
3	EDO	B	404	-	3,3,3	0.48	0	2,2,2	0.27	0
2	ACT	A	401	-	1,3,3	1.34	0	0,3,3	0.00	-
2	ACT	B	401	-	1,3,3	0.61	0	0,3,3	0.00	-
2	ACT	C	401	-	1,3,3	0.09	0	0,3,3	0.00	-

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	EDO	B	405	-	-	1/1/1/1	-
3	EDO	C	402	-	-	0/1/1/1	-
3	EDO	A	405	-	-	0/1/1/1	-
3	EDO	A	404	-	-	1/1/1/1	-
3	EDO	B	404	-	-	1/1/1/1	-
3	EDO	D	404	-	-	0/1/1/1	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	404	EDO	O1-C1-C2-O2
3	B	404	EDO	O1-C1-C2-O2
3	B	405	EDO	O1-C1-C2-O2

There are no ring outliers.

No monomer is involved in short contacts.

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	287/341 (84%)	-0.07	4 (1%) 75 78	24, 37, 62, 91	0
1	B	282/341 (82%)	0.08	4 (1%) 75 78	27, 43, 77, 121	0
1	C	284/341 (83%)	0.63	32 (11%) 5 6	31, 62, 93, 112	0
1	D	282/341 (82%)	0.38	28 (9%) 7 9	28, 46, 90, 128	0
All	All	1135/1364 (83%)	0.25	68 (5%) 21 27	24, 46, 87, 128	0

The worst 5 of 68 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	309	THR	8.5
1	D	353	GLU	6.4
1	C	310	LYS	6.3
1	D	342	GLU	6.3
1	C	366	PHE	5.5

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(\AA^2)	Q<0.9
3	EDO	A	405	4/4	0.78	0.14	47,53,55,59	0
3	EDO	D	404	4/4	0.80	0.16	54,55,60,62	0
2	ACT	B	403	4/4	0.83	0.12	57,64,77,81	0
3	EDO	B	404	4/4	0.86	0.11	57,60,64,64	0
3	EDO	C	402	4/4	0.88	0.10	47,53,55,59	0
2	ACT	C	401	4/4	0.88	0.11	68,70,73,74	0
2	ACT	A	401	4/4	0.89	0.09	52,56,59,62	0
2	ACT	D	401	4/4	0.92	0.14	69,70,70,74	0
2	ACT	D	403	4/4	0.92	0.12	52,56,59,62	0
2	ACT	A	402	4/4	0.92	0.16	45,47,56,57	0
3	EDO	B	405	4/4	0.93	0.07	57,58,64,66	0
2	ACT	B	402	4/4	0.94	0.14	36,51,53,59	0
3	EDO	A	404	4/4	0.94	0.10	46,47,48,55	0
2	ACT	A	403	4/4	0.95	0.12	56,56,59,71	0
2	ACT	B	401	4/4	0.96	0.16	41,59,64,66	0
2	ACT	D	402	4/4	0.96	0.12	49,54,71,71	0

6.5 Other polymers [i](#)

There are no such residues in this entry.