



Full wwPDB X-ray Structure Validation Report ⓘ

Jan 14, 2026 – 12:34 PM JST

PDB ID : 9UVM / pdb_00009uvm
Title : The SeMet pili protein of Streptococcus Sanguinis
Authors : Takebe, K.; Suzuki, M.; Nakata, M.
Deposited on : 2025-05-10
Resolution : 1.80 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

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ 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 ⓘ](#)) were used in the production of this report:

MolProbity	:	4-5-2 with Phenix2.0
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.47

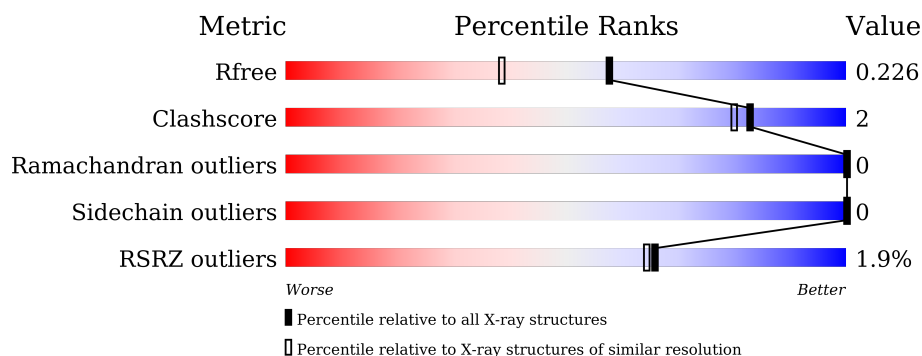
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.80 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}	164625	7108 (1.80-1.80)
Clashscore	180529	8162 (1.80-1.80)
Ramachandran outliers	177936	8077 (1.80-1.80)
Sidechain outliers	177891	8076 (1.80-1.80)
RSRZ outliers	164620	7108 (1.80-1.80)

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 $\leq 5\%$. 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	479	

2 Entry composition [i](#)

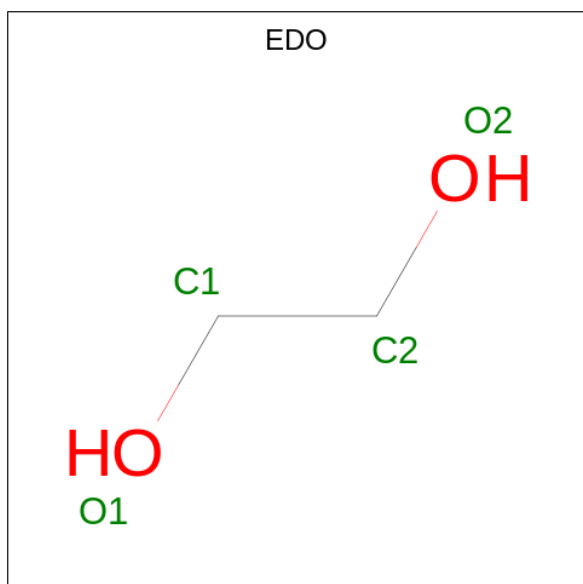
There are 5 unique types of molecules in this entry. The entry contains 3716 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 FimA fimbrial subunit-like protein, putative.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	416	Total	C	N	O	Se	0	12	0
			3242	2010	547	678	7			

- Molecule 2 is 1,2-ETHANEDIOL (CCD ID: EDO) (formula: C₂H₆O₂).



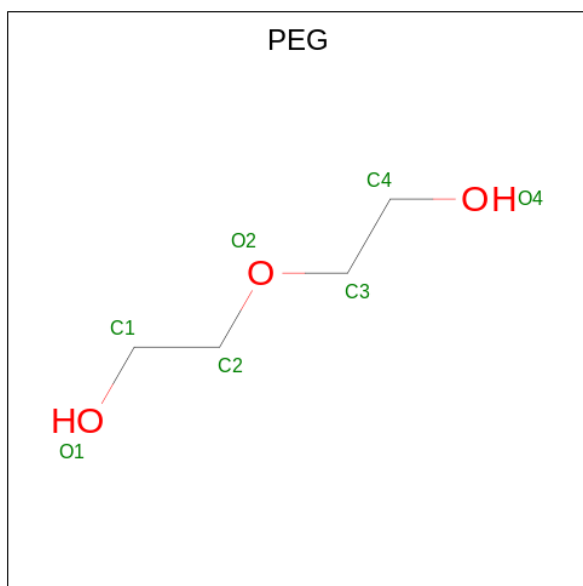
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			4	2	2		
2	A	1	Total	C	O	0	0
			4	2	2		
2	A	1	Total	C	O	0	0
			4	2	2		
2	A	1	Total	C	O	0	0
			4	2	2		
2	A	1	Total	C	O	0	0
			4	2	2		

Continued on next page...

Continued from previous page...

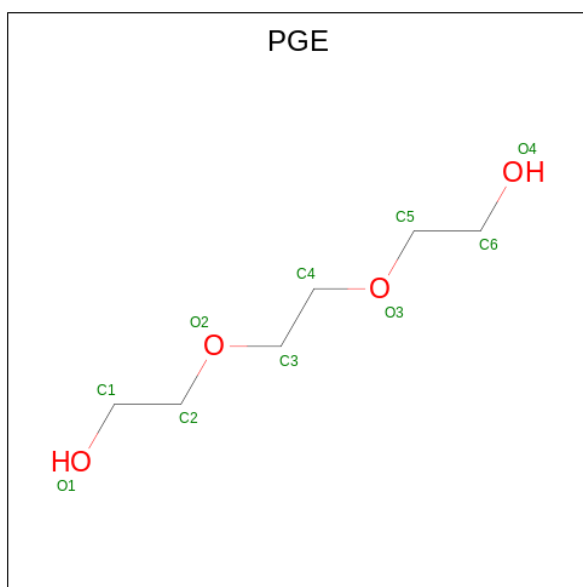
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			4	2	2		
2	A	1	Total	C	O	0	0
			4	2	2		
2	A	1	Total	C	O	0	0
			4	2	2		
2	A	1	Total	C	O	0	0
			4	2	2		
2	A	1	Total	C	O	0	0
			4	2	2		
2	A	1	Total	C	O	0	0
			4	2	2		

- Molecule 3 is DI(HYDROXYETHYL)ETHER (CCD ID: PEG) (formula: $C_4H_{10}O_3$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			7	4	3		
3	A	1	Total	C	O	0	0
			7	4	3		

- Molecule 4 is TRIETHYLENE GLYCOL (CCD ID: PGE) (formula: $C_6H_{14}O_4$).

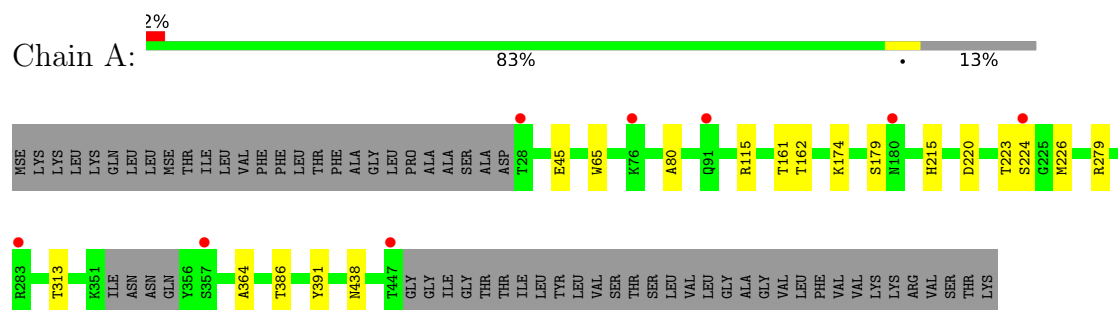


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			10	6	4		
4	A	1	Total	C	O	0	0
			10	6	4		

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	392	Total	O	0	0
			392	392		

- Molecule 1: FimA fimbrial subunit-like protein, putative



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	37.30Å 58.87Å 214.02Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	39.60 – 1.80 39.60 – 1.80	Depositor EDS
% Data completeness (in resolution range)	99.9 (39.60-1.80) 100.0 (39.60-1.80)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.89 (at 1.79Å)	Xtriage
Refinement program	PHENIX (1.20.1_4487: ???)	Depositor
R, R_{free}	0.208 , 0.226 0.208 , 0.226	Depositor DCC
R_{free} test set	2248 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å ²)	20.2	Xtriage
Anisotropy	1.194	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 30.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	3716	wwPDB-VP
Average B, all atoms (Å ²)	28.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.15% 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, PGE, PEG

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	0.17	0/3286	0.38	0/4451

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 [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	3242	0	3169	14	0
2	A	48	0	72	5	0
3	A	14	0	20	1	0
4	A	20	0	28	2	0
5	A	392	0	0	2	0
All	All	3716	0	3289	15	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 2.

All (15) 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:224[A]:SER:H	4:A:516:PGE:H62	1.49	0.78
1:A:224[B]:SER:H	4:A:516:PGE:H62	1.49	0.76
1:A:313:THR:HB	2:A:504:EDO:H22	1.70	0.74
1:A:223:THR:HG21	1:A:226[B]:MSE:HE3	1.82	0.61
1:A:391:TYR:CD1	1:A:438:ASN:HB3	2.47	0.49
1:A:174:LYS:HE3	1:A:220:ASP:OD2	2.12	0.49
1:A:45:GLU:OE2	1:A:115:ARG:NH1	2.49	0.45
2:A:504:EDO:H21	5:A:684:HOH:O	2.16	0.45
1:A:161:THR:HA	2:A:509:EDO:H22	1.99	0.44
1:A:179[B]:SER:O	3:A:513:PEG:H21	2.17	0.44
1:A:162:THR:H	2:A:509:EDO:H22	1.85	0.42
1:A:215:HIS:HE1	5:A:949:HOH:O	2.02	0.41
1:A:364:ALA:HB2	1:A:386:THR:HG21	2.03	0.41
1:A:65:TRP:CD1	1:A:80:ALA:HB2	2.56	0.40
1:A:279:ARG:HB2	2:A:510:EDO:H11	2.02	0.40

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	424/479 (88%)	414 (98%)	10 (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	363/397 (91%)	363 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	34	GLN
1	A	58	ASN
1	A	215	HIS
1	A	255	ASN
1	A	309	ASN

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 oligosaccharides 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	EDO	A	506	-	3,3,3	0.48	0	2,2,2	0.39	0
3	PEG	A	513	-	6,6,6	0.10	0	5,5,5	0.11	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
2	EDO	A	507	-	3,3,3	0.45	0	2,2,2	0.36	0
2	EDO	A	508	-	3,3,3	0.47	0	2,2,2	0.36	0
2	EDO	A	509	-	3,3,3	0.44	0	2,2,2	0.36	0
2	EDO	A	510	-	3,3,3	0.48	0	2,2,2	0.18	0
4	PGE	A	516	-	9,9,9	0.36	0	8,8,8	0.28	0
2	EDO	A	501	-	3,3,3	0.44	0	2,2,2	0.42	0
3	PEG	A	514	-	6,6,6	0.15	0	5,5,5	0.10	0
2	EDO	A	503	-	3,3,3	0.45	0	2,2,2	0.40	0
2	EDO	A	511	-	3,3,3	0.47	0	2,2,2	0.36	0
4	PGE	A	515	-	9,9,9	0.32	0	8,8,8	0.28	0
2	EDO	A	502	-	3,3,3	0.45	0	2,2,2	0.35	0
2	EDO	A	512	-	3,3,3	0.47	0	2,2,2	0.38	0
2	EDO	A	504	-	3,3,3	0.51	0	2,2,2	0.19	0
2	EDO	A	505	-	3,3,3	0.46	0	2,2,2	0.27	0

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	EDO	A	506	-	-	0/1/1/1	-
3	PEG	A	513	-	-	3/4/4/4	-
2	EDO	A	507	-	-	0/1/1/1	-
2	EDO	A	508	-	-	0/1/1/1	-
2	EDO	A	509	-	-	1/1/1/1	-
2	EDO	A	510	-	-	0/1/1/1	-
4	PGE	A	516	-	-	3/7/7/7	-
2	EDO	A	501	-	-	0/1/1/1	-
3	PEG	A	514	-	-	2/4/4/4	-
2	EDO	A	503	-	-	0/1/1/1	-
2	EDO	A	511	-	-	0/1/1/1	-
4	PGE	A	515	-	-	6/7/7/7	-
2	EDO	A	502	-	-	0/1/1/1	-
2	EDO	A	512	-	-	0/1/1/1	-
2	EDO	A	504	-	-	1/1/1/1	-
2	EDO	A	505	-	-	1/1/1/1	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (17) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	515	PGE	O2-C3-C4-O3
3	A	513	PEG	O2-C3-C4-O4
4	A	515	PGE	O1-C1-C2-O2
3	A	514	PEG	O1-C1-C2-O2
4	A	516	PGE	O1-C1-C2-O2
4	A	516	PGE	O2-C3-C4-O3
4	A	515	PGE	C4-C3-O2-C2
3	A	513	PEG	O1-C1-C2-O2
2	A	504	EDO	O1-C1-C2-O2
4	A	516	PGE	C1-C2-O2-C3
2	A	509	EDO	O1-C1-C2-O2
4	A	515	PGE	C1-C2-O2-C3
4	A	515	PGE	C3-C4-O3-C5
3	A	514	PEG	C1-C2-O2-C3
4	A	515	PGE	O3-C5-C6-O4
3	A	513	PEG	C1-C2-O2-C3
2	A	505	EDO	O1-C1-C2-O2

There are no ring outliers.

5 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	513	PEG	1	0
2	A	509	EDO	2	0
2	A	510	EDO	1	0
4	A	516	PGE	2	0
2	A	504	EDO	2	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	411/479 (85%)	0.06	8 (1%) 66 64	11, 27, 37, 47	10 (2%)

All (8) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	180[A]	ASN	3.3
1	A	28	THR	3.0
1	A	447	THR	3.0
1	A	357	SER	2.4
1	A	224[A]	SER	2.4
1	A	91	GLN	2.2
1	A	283[A]	ARG	2.1
1	A	76	LYS	2.1

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 oligosaccharides 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($\text{\AA}^2$)	Q<0.9
4	PGE	A	516	10/10	0.51	0.20	30,31,33,34	0
3	PEG	A	513	7/7	0.72	0.16	32,33,35,36	0
2	EDO	A	509	4/4	0.73	0.17	33,34,34,35	0
3	PEG	A	514	7/7	0.75	0.14	25,29,34,35	0
2	EDO	A	504	4/4	0.75	0.23	29,31,33,35	0
4	PGE	A	515	10/10	0.77	0.15	28,33,35,35	0
2	EDO	A	508	4/4	0.81	0.18	39,39,40,40	0
2	EDO	A	512	4/4	0.81	0.19	33,34,35,37	0
2	EDO	A	510	4/4	0.83	0.15	28,29,29,29	0
2	EDO	A	503	4/4	0.83	0.22	36,38,38,38	0
2	EDO	A	505	4/4	0.85	0.15	24,29,30,33	0
2	EDO	A	501	4/4	0.87	0.12	33,35,36,37	0
2	EDO	A	502	4/4	0.89	0.16	30,30,30,30	0
2	EDO	A	507	4/4	0.89	0.12	32,34,36,37	0
2	EDO	A	506	4/4	0.90	0.14	25,25,26,27	0
2	EDO	A	511	4/4	0.93	0.09	26,28,28,28	0

6.5 Other polymers [i](#)

There are no such residues in this entry.