



wwPDB X-ray Structure Validation Summary Report ⓘ

May 15, 2020 – 11:56 am BST

PDB ID : 1GK9
Title : Crystal structures of penicillin acylase enzyme-substrate complexes: Structural insights into the catalytic mechanism
Authors : McVey, C.E.; Walsh, M.A.; Dodson, G.G.; Wilson, K.S.; Brannigan, J.A.
Deposited on : 2001-08-10
Resolution : 1.30 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

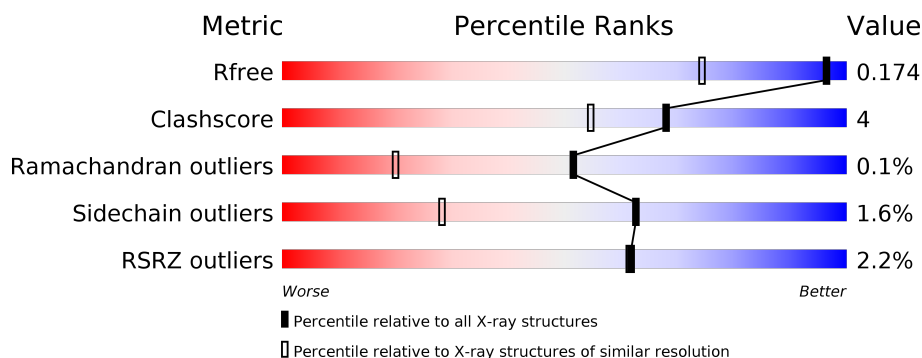
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.30 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	1058 (1.30-1.30)
Clashscore	141614	1101 (1.30-1.30)
Ramachandran outliers	138981	1058 (1.30-1.30)
Sidechain outliers	138945	1058 (1.30-1.30)
RSRZ outliers	127900	1029 (1.30-1.30)

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	260	<div> <div>%</div> <div> <div></div> <div>69%</div> <div>11%</div> <div>20%</div> </div> </div>
2	B	557	<div> <div>3%</div> <div> <div></div> <div>87%</div> <div>12%</div> <div>.</div> </div> </div>

2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 7533 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 PENICILLIN G ACYLASE ALPHA SUBUNIT.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	208	Total	C	N	O	S	5	10	0
			1750	1114	293	334	9			

- Molecule 2 is a protein called PENICILLIN G ACYLASE BETA SUBUNIT.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	557	Total	C	N	O	S	8	40	0
			4751	3016	825	899	11			

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



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			4	2	2		
3	A	1	Total	C	O	0	1
			8	4	4		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total 4	C 2	O 2	0	0
3	B	1	Total 4	C 2	O 2	0	0
3	B	1	Total 4	C 2	O 2	0	0
3	B	1	Total 4	C 2	O 2	0	0
3	B	1	Total 4	C 2	O 2	0	0
3	B	1	Total 4	C 2	O 2	0	0
3	B	1	Total 4	C 2	O 2	0	0
3	B	1	Total 4	C 2	O 2	0	0
3	B	1	Total 4	C 2	O 2	0	0
3	B	1	Total 8	C 4	O 4	0	1
3	B	1	Total 4	C 2	O 2	0	0
3	B	1	Total 4	C 2	O 2	0	0
3	B	1	Total 4	C 2	O 2	0	0
3	B	1	Total 4	C 2	O 2	0	0
3	B	1	Total 4	C 2	O 2	0	0
3	B	1	Total 4	C 2	O 2	0	0
3	B	1	Total 4	C 2	O 2	0	0
3	B	1	Total 4	C 2	O 2	0	0
3	B	1	Total 4	C 2	O 2	0	0
3	B	1	Total 4	C 2	O 2	0	0
3	B	1	Total 4	C 2	O 2	0	0
3	B	1	Total 4	C 2	O 2	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	B	1	Total	C	O	0	1
			8	4	4		

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

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

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	254	Total	O	0	5
			254	254		
5	B	669	Total	O	0	10
			669	669		

4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	51.30 Å 131.60 Å 63.90 Å 90.00° 105.90° 90.00°	Depositor
Resolution (Å)	20.00 – 1.30 19.99 – 1.30	Depositor EDS
% Data completeness (in resolution range)	95.5 (20.00-1.30) 95.3 (19.99-1.30)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.20 (at 1.30 Å)	Xtriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.148 , 0.169 0.154 , 0.174	Depositor DCC
R_{free} test set	5756 reflections (3.03%)	wwPDB-VP
Wilson B-factor (Å ²)	10.8	Xtriage
Anisotropy	0.118	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.37 , 54.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	7533	wwPDB-VP
Average B, all atoms (Å ²)	15.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.58% 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: CA, EDO

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	0.99	1/1794 (0.1%)	1.51	24/2432 (1.0%)
2	B	0.91	1/4895 (0.0%)	1.58	54/6669 (0.8%)
All	All	0.93	2/6689 (0.0%)	1.56	78/9101 (0.9%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1
2	B	0	6
All	All	0	7

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	60	LYS	CD-CE	-14.38	1.15	1.51
2	B	482	GLU	CD-OE2	6.04	1.32	1.25

The worst 5 of 78 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	455[A]	ARG	NE-CZ-NH2	-24.69	107.95	120.30
2	B	455[B]	ARG	NE-CZ-NH2	-24.69	107.95	120.30
2	B	495[A]	ARG	NE-CZ-NH2	17.09	128.85	120.30
2	B	495[B]	ARG	NE-CZ-NH2	17.09	128.85	120.30
2	B	269	ARG	NE-CZ-NH2	-16.06	112.27	120.30

There are no chirality outliers.

5 of 7 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	5[B]	SER	Mainchain
2	B	109[B]	LYS	Mainchain
2	B	168[A]	GLN	Sidechain
2	B	168[B]	GLN	Mainchain
2	B	367[B]	PHE	Mainchain

5.2 Too-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 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	1750	0	1679	16	0
2	B	4751	0	4546	35	0
3	A	16	0	24	3	0
3	B	92	0	136	7	0
4	B	1	0	0	0	0
5	A	254	0	0	4	0
5	B	669	0	0	6	0
All	All	7533	0	6385	49	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 4.

The worst 5 of 49 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:149[A]:SER:HB2	5:B:2302[A]:HOH:O	1.48	1.14
1:A:199:LEU:HD11	2:B:225[B]:MET:HE1	1.29	1.14
2:B:168[A]:GLN:NE2	5:B:2215[A]:HOH:O	1.86	1.05
2:B:233[A]:GLN:NE2	5:B:2287[A]:HOH:O	1.76	0.98
2:B:488[B]:PHE:CE1	2:B:497:VAL:HG22	2.05	0.92

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	216/260 (83%)	212 (98%)	4 (2%)	0	100	100
2	B	595/557 (107%)	582 (98%)	12 (2%)	1 (0%)	47	19
All	All	811/817 (99%)	794 (98%)	16 (2%)	1 (0%)	51	20

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	251	SER

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	189/211 (90%)	184 (97%)	5 (3%)	46	10
2	B	500/460 (109%)	494 (99%)	6 (1%)	71	40
All	All	689/671 (103%)	678 (98%)	11 (2%)	62	28

5 of 11 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	183	ASN
2	B	20	ASN
2	B	173	GLN
1	A	181[B]	LEU
2	B	154	TRP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 12 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	231	ASN
2	B	245	GLN
2	B	440	ASN
2	B	187	ASN
2	B	401	GLN

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](#)

Of 28 ligands modelled in this entry, 1 is monoatomic - leaving 27 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	EDO	B	1566[A]	-	3,3,3	0.43	0	2,2,2	0.71	0
3	EDO	B	1559	-	3,3,3	0.53	0	2,2,2	0.25	0
3	EDO	B	1558	-	3,3,3	0.68	0	2,2,2	1.16	0
3	EDO	B	1568	-	3,3,3	0.57	0	2,2,2	0.02	0
3	EDO	B	1570	-	3,3,3	0.54	0	2,2,2	0.60	0
3	EDO	B	1563	-	3,3,3	1.04	0	2,2,2	1.04	0
3	EDO	B	1571	-	3,3,3	0.58	0	2,2,2	0.36	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
3	EDO	B	1567	-	3,3,3	0.41	0	2,2,2	0.39	0
3	EDO	B	1564	-	3,3,3	0.42	0	2,2,2	1.04	0
3	EDO	B	1576	-	3,3,3	0.44	0	2,2,2	0.17	0
3	EDO	B	1573	-	3,3,3	0.55	0	2,2,2	0.72	0
3	EDO	B	1561	-	3,3,3	0.99	0	2,2,2	0.61	0
3	EDO	A	1211[A]	-	3,3,3	0.52	0	2,2,2	0.25	0
3	EDO	A	1211[B]	-	3,3,3	0.49	0	2,2,2	0.13	0
3	EDO	A	1212	-	3,3,3	0.57	0	2,2,2	0.57	0
3	EDO	B	1569	-	3,3,3	0.53	0	2,2,2	0.21	0
3	EDO	B	1578[B]	-	3,3,3	0.56	0	2,2,2	0.28	0
3	EDO	B	1578[A]	-	3,3,3	0.67	0	2,2,2	0.54	0
3	EDO	B	1575	-	3,3,3	1.11	0	2,2,2	1.15	0
3	EDO	B	1562	-	3,3,3	0.59	0	2,2,2	0.74	0
3	EDO	B	1574	-	3,3,3	0.45	0	2,2,2	0.70	0
3	EDO	B	1577	-	3,3,3	0.56	0	2,2,2	0.53	0
3	EDO	A	1210	-	3,3,3	0.59	0	2,2,2	0.24	0
3	EDO	B	1566[B]	-	3,3,3	0.36	0	2,2,2	0.54	0
3	EDO	B	1565	-	3,3,3	0.28	0	2,2,2	0.13	0
3	EDO	B	1560	-	3,3,3	0.80	0	2,2,2	0.22	0
3	EDO	B	1572	-	3,3,3	0.44	0	2,2,2	0.47	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
3	EDO	B	1566[A]	-	-	0/1/1/1	-
3	EDO	B	1559	-	-	0/1/1/1	-
3	EDO	B	1558	-	-	0/1/1/1	-
3	EDO	B	1568	-	-	0/1/1/1	-
3	EDO	B	1570	-	-	0/1/1/1	-
3	EDO	B	1563	-	-	0/1/1/1	-
3	EDO	B	1571	-	-	0/1/1/1	-
3	EDO	B	1567	-	-	1/1/1/1	-
3	EDO	B	1564	-	-	0/1/1/1	-
3	EDO	B	1576	-	-	0/1/1/1	-
3	EDO	B	1573	-	-	0/1/1/1	-
3	EDO	B	1561	-	-	0/1/1/1	-
3	EDO	A	1211[A]	-	-	0/1/1/1	-
3	EDO	A	1211[B]	-	-	1/1/1/1	-
3	EDO	A	1212	-	-	0/1/1/1	-
3	EDO	B	1569	-	-	0/1/1/1	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	EDO	B	1578[B]	-	-	1/1/1/1	-
3	EDO	B	1578[A]	-	-	0/1/1/1	-
3	EDO	B	1575	-	-	0/1/1/1	-
3	EDO	B	1562	-	-	0/1/1/1	-
3	EDO	B	1574	-	-	0/1/1/1	-
3	EDO	B	1577	-	-	0/1/1/1	-
3	EDO	A	1210	-	-	1/1/1/1	-
3	EDO	B	1566[B]	-	-	1/1/1/1	-
3	EDO	B	1565	-	-	0/1/1/1	-
3	EDO	B	1560	-	-	0/1/1/1	-
3	EDO	B	1572	-	-	0/1/1/1	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	B	1566[B]	EDO	O1-C1-C2-O2
3	B	1578[B]	EDO	O1-C1-C2-O2
3	A	1210	EDO	O1-C1-C2-O2
3	B	1567	EDO	O1-C1-C2-O2
3	A	1211[B]	EDO	O1-C1-C2-O2

There are no ring outliers.

7 monomers are involved in 10 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	1559	EDO	1	0
3	B	1563	EDO	1	0
3	A	1211[B]	EDO	1	0
3	B	1578[B]	EDO	3	0
3	B	1575	EDO	1	0
3	A	1210	EDO	2	0
3	B	1566[B]	EDO	1	0

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 [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	208/260 (80%)	-0.22	2 (0%) 82 83	8, 12, 20, 68	2 (0%)
2	B	557/557 (100%)	-0.15	15 (2%) 54 52	6, 12, 26, 62	4 (0%)
All	All	765/817 (93%)	-0.17	17 (2%) 62 61	6, 12, 24, 68	6 (0%)

The worst 5 of 17 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	2	GLN	10.0
2	B	132	THR	7.3
2	B	131	GLN	4.7
2	B	133	THR	4.1
2	B	380[A]	GLN	3.9

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(\AA^2)	Q<0.9
3	EDO	A	1210	4/4	0.81	0.12	20,26,30,33	0
3	EDO	B	1567	4/4	0.84	0.11	26,29,31,35	0
3	EDO	A	1212	4/4	0.88	0.14	22,25,27,28	0
3	EDO	B	1577	4/4	0.92	0.25	18,22,24,25	0
3	EDO	B	1575	4/4	0.93	0.13	15,20,21,23	0
3	EDO	B	1565	4/4	0.93	0.07	18,19,24,25	0
3	EDO	B	1573	4/4	0.94	0.19	19,21,21,22	0
3	EDO	B	1570	4/4	0.95	0.12	17,23,23,23	0
3	EDO	B	1563	4/4	0.95	0.09	12,12,16,16	0
3	EDO	B	1558	4/4	0.95	0.09	12,14,15,18	0
3	EDO	B	1578[B]	4/4	0.96	0.19	20,21,25,29	4
3	EDO	B	1578[A]	4/4	0.96	0.19	13,19,21,21	4
3	EDO	A	1211[A]	4/4	0.96	0.14	14,17,19,22	4
3	EDO	B	1574	4/4	0.96	0.09	10,15,19,21	0
3	EDO	A	1211[B]	4/4	0.96	0.14	12,16,18,20	4
3	EDO	B	1576	4/4	0.96	0.19	16,21,28,29	0
3	EDO	B	1569	4/4	0.96	0.08	13,15,17,18	0
3	EDO	B	1564	4/4	0.97	0.07	15,16,17,18	0
3	EDO	B	1562	4/4	0.97	0.06	11,11,13,16	0
3	EDO	B	1561	4/4	0.97	0.08	16,18,20,23	0
3	EDO	B	1572	4/4	0.97	0.16	10,22,24,29	0
3	EDO	B	1559	4/4	0.98	0.06	9,10,11,13	0
3	EDO	B	1566[B]	4/4	0.98	0.07	13,13,16,17	4
3	EDO	B	1566[A]	4/4	0.98	0.07	8,10,10,10	4
3	EDO	B	1571	4/4	0.98	0.08	11,12,13,13	0
3	EDO	B	1560	4/4	0.99	0.05	8,8,9,9	0
3	EDO	B	1568	4/4	0.99	0.05	13,13,14,14	0
4	CA	B	1579	1/1	1.00	0.02	7,7,7,7	0

6.5 Other polymers

There are no such residues in this entry.