



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 13, 2017 – 11:38 am GMT

PDB ID : 1B97  
Title : ANALYSIS OF A MUTATIONAL HOT-SPOT IN THE ECORV RESTRICTION ENDONUCLEASE: A CATALYTIC ROLE FOR A MAIN CHAIN CARBONYL GROUP  
Authors : Thomas, M.P.; Halford, S.E.; Brady, R.L.  
Deposited on : 1999-02-19  
Resolution : 1.90 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](mailto:validation@mail.wwpdb.org)

A user guide is available at

<http://wwpdb.org/validation/2016/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
Xtriage (Phenix)	:	1.9-1692
EDS	:	trunk28620
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	recalc28949

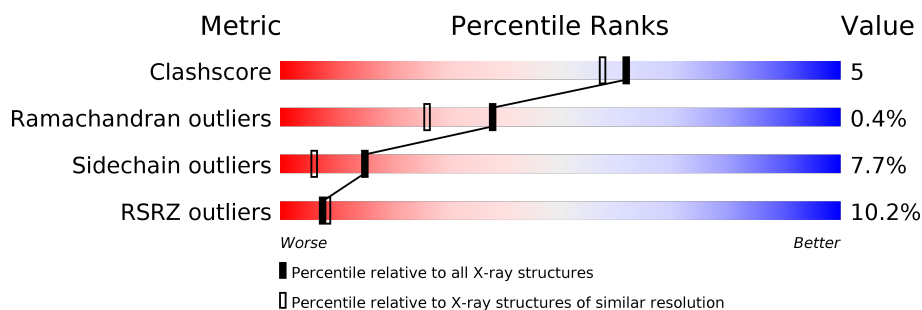
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.90 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))
Clashscore	112137	5731 (1.90-1.90)
Ramachandran outliers	110173	5669 (1.90-1.90)
Sidechain outliers	110143	5670 (1.90-1.90)
RSRZ outliers	101464	5100 (1.90-1.90)

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. 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	C	11	<div> <div>27%</div> <div>45%</div> <div>27%</div> </div>
1	D	11	<div> <div>9%</div> <div>36%</div> <div>45%</div> <div>18%</div> </div>
2	A	244	<div> <div>9%</div> <div>80%</div> <div>19%</div> <div>.</div> </div>
2	B	244	<div> <div>12%</div> <div>82%</div> <div>17%</div> <div>.</div> </div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 4892 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 DNA chain called DNA (5'-D(\*AP\*AP\*AP\*GP\*AP\*TP\*AP\*TP\*CP\*TP\*T)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	C	11	Total	C	N	O	P	0	0	0
			223	109	41	63	10			
1	D	11	Total	C	N	O	P	0	0	0
			223	109	41	63	10			

- Molecule 2 is a protein called RESTRICTION ENDONUCLEASE ECORV.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	A	244	Total	C	N	O	S	0	0	0
			2022	1301	340	380	1			
2	B	244	Total	C	N	O	S	0	5	0
			2066	1328	350	387	1			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	69	LEU	GLN	ENGINEERED	UNP P04390
B	69	LEU	GLN	ENGINEERED	UNP P04390

- Molecule 3 is water.

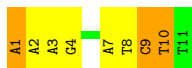
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	159	Total	O	0	0
			159	159		
3	B	156	Total	O	0	0
			156	156		
3	C	21	Total	O	0	0
			21	21		
3	D	22	Total	O	0	0
			22	22		

### 3 Residue-property plots [i](#)

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 ( $RSRZ > 2$ ). 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.

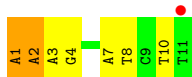
- Molecule 1: DNA (5'-D(\*AP\*AP\*AP\*GP\*AP\*TP\*AP\*TP\*CP\*TP\*T)-3')

Chain C: 




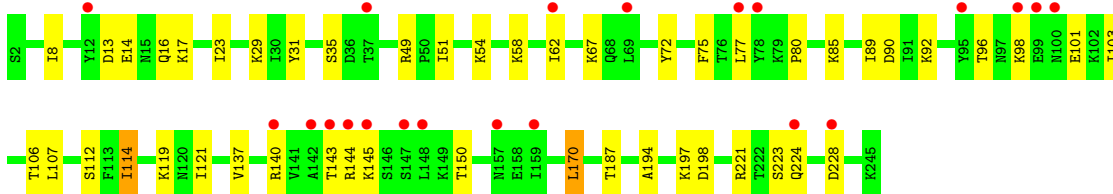
- Molecule 1: DNA (5'-D(\*AP\*AP\*AP\*GP\*AP\*TP\*AP\*TP\*CP\*TP\*T)-3')

Chain D: 




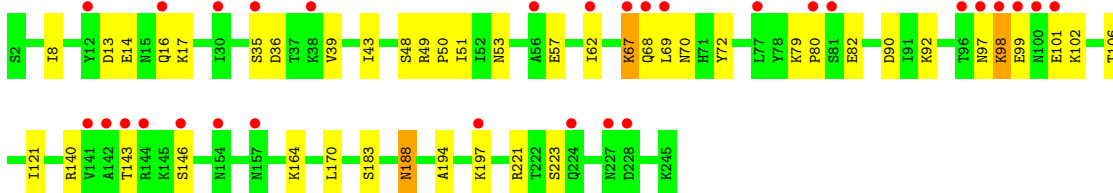
- Molecule 2: RESTRICTION ENDONUCLEASE ECORV

Chain A: 



- Molecule 2: RESTRICTION ENDONUCLEASE ECORV

Chain B: 



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	48.65Å 49.22Å 64.04Å 96.65° 108.91° 107.53°	Depositor
Resolution (Å)	15.00 – 1.90 14.99 – 2.50	Depositor EDS
% Data completeness (in resolution range)	92.7 (15.00-1.90) 87.5 (14.99-2.50)	Depositor EDS
$R_{merge}$	0.05	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	6.38 (at 2.51Å)	Xtriage
Refinement program	REFMAC	Depositor
R, $R_{free}$	0.219 , 0.279 0.191 , (Not available)	Depositor DCC
$R_{free}$ test set	No test flags present.	DCC
Wilson B-factor (Å <sup>2</sup> )	24.5	Xtriage
Anisotropy	0.358	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 62.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	4892	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	28.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.84% of the height of the origin peak. No significant pseudotranslation is detected.*

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

<sup>2</sup>Theoretical 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

### 5.1 Standard geometry

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	C	1.08	0/250	2.15	12/384 (3.1%)
1	D	1.13	0/250	2.07	9/384 (2.3%)
2	A	0.35	0/2073	0.82	2/2803 (0.1%)
2	B	0.35	0/2118	0.83	1/2863 (0.0%)
All	All	0.49	0/4691	1.06	24/6434 (0.4%)

There are no bond length outliers.

All (24) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	8	DT	P-O3'-C3'	14.23	136.78	119.70
1	D	8	DT	P-O3'-C3'	12.97	135.26	119.70
1	D	4	DG	O4'-C1'-N9	9.06	114.34	108.00
1	D	3	DA	P-O3'-C3'	8.82	130.28	119.70
1	D	2	DA	P-O3'-C3'	8.70	130.14	119.70
1	C	2	DA	P-O3'-C3'	7.97	129.26	119.70
1	C	3	DA	P-O3'-C3'	7.71	128.95	119.70
1	D	1	DA	P-O3'-C3'	6.72	127.76	119.70
1	C	10	DT	P-O3'-C3'	6.70	127.73	119.70
1	C	1	DA	O4'-C4'-C3'	-6.56	101.88	104.50
1	D	7	DA	P-O3'-C3'	6.25	127.20	119.70
1	C	7	DA	P-O3'-C3'	6.21	127.16	119.70
1	C	10	DT	O4'-C1'-N1	-6.21	103.66	108.00
1	C	1	DA	P-O3'-C3'	6.20	127.14	119.70
1	C	1	DA	C4'-C3'-C2'	-6.05	97.65	103.10
1	C	4	DG	O4'-C1'-N9	5.96	112.17	108.00
2	B	221	ARG	CD-NE-CZ	5.84	131.78	123.60
1	C	8	DT	C4-C5-C6	5.64	121.38	118.00
2	A	221	ARG	NE-CZ-NH2	-5.59	117.51	120.30
1	D	10	DT	P-O3'-C3'	5.56	126.37	119.70
1	D	7	DA	OP1-P-OP2	-5.52	111.32	119.60
2	A	221	ARG	CD-NE-CZ	5.43	131.20	123.60
1	D	8	DT	N3-C4-O4	5.19	123.01	119.90

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	9	DC	OP1-P-OP2	-5.15	111.87	119.60

There are no chirality outliers.

There are no planarity outliers.

## 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	C	223	0	127	5	0
1	D	223	0	127	3	0
2	A	2022	0	1990	19	0
2	B	2066	0	2034	17	0
3	A	159	0	0	2	0
3	B	156	0	0	2	0
3	C	21	0	0	0	0
3	D	22	0	0	0	0
All	All	4892	0	4278	44	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 5.

All (44) 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:C:1:DA:H2'	1:C:1:DA:O5'	1.65	0.96
1:C:1:DA:O5'	1:C:1:DA:N3	2.09	0.84
1:C:1:DA:C2'	1:C:1:DA:O5'	2.36	0.73
2:A:14:GLU:HG2	2:A:51:ILE:HD11	1.79	0.64
2:A:29:LYS:HE2	2:A:150:THR:HG21	1.81	0.63
2:B:49:ARG:HB2	2:B:50:PRO:HD3	1.80	0.63
2:A:13:ASP:HA	2:A:16:GLN:HE21	1.66	0.60
2:A:92:LYS:HE3	2:A:107:LEU:HA	1.84	0.59
2:B:14:GLU:HG2	2:B:51:ILE:HD11	1.85	0.58
1:C:9:DC:H2'	1:C:10:DT:H72	1.88	0.56
2:B:70[B]:ASN:HB3	3:B:370:HOH:O	2.05	0.55

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:8:ILE:HG12	2:B:170:LEU:HB3	1.90	0.54
2:B:48:SER:HA	2:B:51:ILE:HD12	1.90	0.54
2:B:39:VAL:O	2:B:43:ILE:HG12	2.10	0.52
2:B:70[A]:ASN:HB3	3:B:370:HOH:O	2.10	0.52
2:A:114:ILE:HG23	3:A:326:HOH:O	2.10	0.51
1:D:1:DA:H2'	1:D:2:DA:C8	2.46	0.50
2:A:23:ILE:CG1	2:A:31:TYR:HB2	2.41	0.50
2:A:103:ILE:HD13	2:A:137:VAL:HG21	1.94	0.50
2:B:53:ASN:O	2:B:57:GLU:HG3	2.12	0.49
2:B:69[B]:LEU:HD12	2:B:69[B]:LEU:H	1.78	0.49
2:B:102:LYS:HD3	2:B:194:ALA:HA	1.94	0.48
2:B:98:LYS:HB2	2:B:101:GLU:HG3	1.96	0.48
2:A:112:SER:HA	2:A:119:LYS:HD3	1.96	0.47
2:A:49:ARG:HG2	2:A:75:PHE:CZ	2.51	0.45
2:A:80:PRO:HD2	3:A:366:HOH:O	2.17	0.45
2:A:49:ARG:HG2	2:A:75:PHE:HZ	1.81	0.45
2:A:8:ILE:HD13	2:A:170:LEU:HB2	1.99	0.44
2:A:92:LYS:HD3	2:A:106:THR:O	2.16	0.44
2:A:194:ALA:HB1	2:A:198:ASP:OD2	2.16	0.44
2:A:72:TYR:O	2:A:121:ILE:HA	2.18	0.43
2:A:98:LYS:O	2:A:101:GLU:HG3	2.19	0.43
2:B:97:ASN:HB2	2:B:101:GLU:OE1	2.18	0.43
2:A:35:SER:HB2	2:A:140:ARG:HH21	1.83	0.42
2:B:183:SER:OG	2:B:188:ASN:HB2	2.19	0.42
1:C:9:DC:H2'	1:C:10:DT:C7	2.49	0.42
2:A:13:ASP:HA	2:A:16:GLN:NE2	2.33	0.41
2:B:72:TYR:O	2:B:121:ILE:HA	2.21	0.41
2:B:62:ILE:HD13	2:B:80:PRO:HD3	2.03	0.41
1:D:1:DA:H1'	1:D:2:DA:H5'	2.03	0.41
2:B:92:LYS:HE3	2:B:106:THR:O	2.20	0.40
1:D:1:DA:O5'	1:D:1:DA:H2'	2.21	0.40
2:B:67[B]:LYS:HG3	2:B:68[B]:GLN:OE1	2.21	0.40
2:A:77:LEU:HD13	2:A:89:ILE:HD12	2.03	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	
2	A	242/244 (99%)	231 (96%)	10 (4%)	1 (0%)	38	26
2	B	247/244 (101%)	241 (98%)	5 (2%)	1 (0%)	38	26
All	All	489/488 (100%)	472 (96%)	15 (3%)	2 (0%)	38	26

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	36	ASP
2	A	187	THR

### 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	
2	A	220/220 (100%)	203 (92%)	17 (8%)	15	6
2	B	225/220 (102%)	207 (92%)	18 (8%)	14	5
All	All	445/440 (101%)	410 (92%)	35 (8%)	15	6

All (35) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	A	17	LYS
2	A	54	LYS
2	A	58	LYS
2	A	62	ILE

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
2	A	67	LYS
2	A	85	LYS
2	A	90	ASP
2	A	96	THR
2	A	114	ILE
2	A	143	THR
2	A	144	ARG
2	A	145	LYS
2	A	170	LEU
2	A	197	LYS
2	A	223	SER
2	A	224	GLN
2	A	228	ASP
2	B	13	ASP
2	B	16	GLN
2	B	17	LYS
2	B	35	SER
2	B	67[A]	LYS
2	B	67[B]	LYS
2	B	79	LYS
2	B	82	GLU
2	B	90	ASP
2	B	98	LYS
2	B	99	GLU
2	B	140	ARG
2	B	143	THR
2	B	146	SER
2	B	164	LYS
2	B	188	ASN
2	B	197	LYS
2	B	223	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
2	A	16	GLN
2	A	224	GLN
2	B	16	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](#)

There are no ligands in this entry.

### 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 ⓘ

### 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	C	11/11 (100%)	0.11	0 100 100	19, 22, 47, 54	0
1	D	11/11 (100%)	0.16	1 (9%) 10 11	18, 22, 40, 56	0
2	A	244/244 (100%)	0.63	21 (8%) 11 12	14, 26, 45, 58	0
2	B	244/244 (100%)	0.70	30 (12%) 5 5	14, 27, 46, 64	0
All	All	510/510 (100%)	0.64	52 (10%) 7 8	14, 26, 47, 64	0

All (52) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	A	144	ARG	5.4
2	B	99	GLU	4.8
2	B	144	ARG	4.5
2	B	100	ASN	4.3
2	B	67[A]	LYS	4.2
2	A	228	ASP	4.2
2	A	98	LYS	4.2
2	A	37	THR	4.1
2	B	97	ASN	3.7
1	D	11	DT	3.4
2	A	145	LYS	3.4
2	B	146	SER	3.4
2	B	224	GLN	3.3
2	B	12	TYR	3.3
2	B	35	SER	3.2
2	A	143	THR	3.1
2	B	143	THR	3.1
2	B	69[A]	LEU	3.0
2	B	98	LYS	3.0
2	B	142	ALA	3.0
2	A	100	ASN	3.0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	B	157	ASN	2.9
2	A	95	TYR	2.9
2	A	99	GLU	2.9
2	B	81	SER	2.9
2	A	69	LEU	2.7
2	A	142	ALA	2.6
2	A	12	TYR	2.6
2	B	141	VAL	2.6
2	B	228	ASP	2.6
2	A	140	ARG	2.5
2	B	227	ASN	2.5
2	B	77	LEU	2.5
2	A	147	SER	2.5
2	B	96	THR	2.5
2	A	62	ILE	2.4
2	A	148	LEU	2.4
2	B	197	LYS	2.4
2	B	68[A]	GLN	2.4
2	A	78	TYR	2.3
2	B	62	ILE	2.3
2	B	80	PRO	2.3
2	A	224	GLN	2.3
2	B	16	GLN	2.3
2	B	38	LYS	2.2
2	B	154	ASN	2.2
2	A	159	ILE	2.2
2	B	56	ALA	2.2
2	A	157	ASN	2.1
2	B	30	ILE	2.1
2	A	77	LEU	2.1
2	B	101	GLU	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 carbohydrates in this entry.

## 6.4 Ligands [i](#)

There are no ligands in this entry.

## 6.5 Other polymers [i](#)

There are no such residues in this entry.