



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 12, 2024 – 09:23 AM EDT

PDB ID : 2I8A  
Title : Salmonella typhimurium liganded by phosphate ion at 1.64Å resolution  
Authors : Timofeev, V.I.; Gabdoulkhakov, A.G.; Mikhailov, A.M.  
Deposited on : 2006-09-01  
Resolution : 1.64 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](mailto: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	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	2.36.2
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)	:	2.36.2

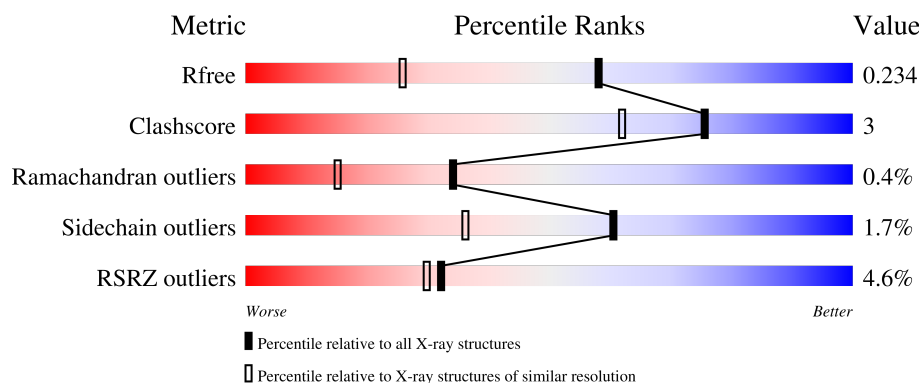
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.64 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	3122 (1.66-1.62)
Clashscore	141614	3268 (1.66-1.62)
Ramachandran outliers	138981	3215 (1.66-1.62)
Sidechain outliers	138945	3215 (1.66-1.62)
RSRZ outliers	127900	3079 (1.66-1.62)

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  $\geq 3$ , 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	252	<div> <div>7%</div> <div> <div></div> <div>88%</div> <div>12%</div> </div> </div>
1	B	252	<div> <div>4%</div> <div> <div></div> <div>85%</div> <div>10%</div> <div>5%</div> </div> </div>
1	C	252	<div> <div>4%</div> <div> <div></div> <div>88%</div> <div>7%</div> <div>5%</div> </div> </div>
1	D	252	<div> <div>3%</div> <div> <div></div> <div>91%</div> <div>5%</div> </div> </div>
1	E	252	<div> <div>6%</div> <div> <div></div> <div>90%</div> <div>8%</div> <div>••</div> </div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	F	252	 A horizontal bar chart showing the quality of chain F. The bar is divided into three segments: a small red segment at the beginning labeled '3%', a large green segment in the middle labeled '90%', and a small yellow segment at the end labeled '7%'. A small grey dot is visible at the far right end of the bar.

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 12217 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 Uridine phosphorylase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	251	Total	C	N	O	S	0	18	0
			1988	1253	342	379	14			
1	B	239	Total	C	N	O	S	0	17	0
			1897	1201	334	350	12			
1	C	239	Total	C	N	O	S	0	12	0
			1864	1173	328	352	11			
1	D	244	Total	C	N	O	S	0	12	0
			1905	1200	332	362	11			
1	E	248	Total	C	N	O	S	0	13	0
			1938	1219	342	365	12			
1	F	245	Total	C	N	O	S	0	12	0
			1901	1201	330	358	12			

- Molecule 2 is PHOSPHATE ION (three-letter code: PO4) (formula: O<sub>4</sub>P).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total O P 5 4 1	0	0
2	B	1	Total O P 5 4 1	0	0
2	C	1	Total O P 5 4 1	0	0
2	D	1	Total O P 5 4 1	0	0
2	E	1	Total O P 5 4 1	0	0
2	F	1	Total O P 5 4 1	0	0

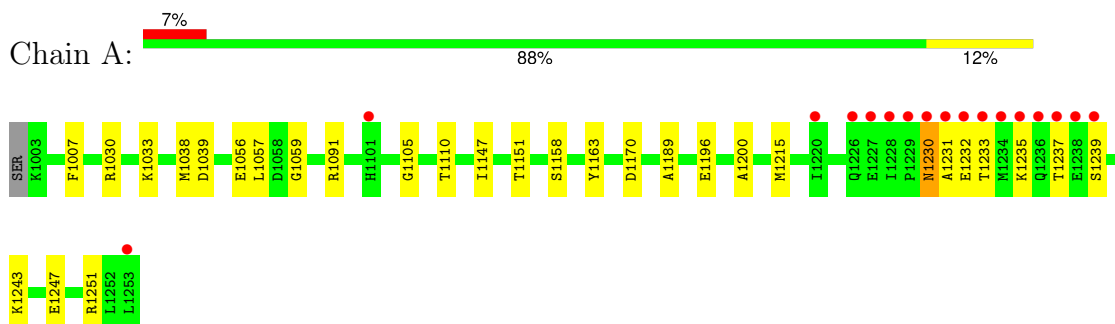
- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	134	Total O 136 136	0	2
3	B	107	Total O 108 108	0	1
3	C	107	Total O 107 107	0	0
3	D	112	Total O 112 112	0	0
3	E	116	Total O 116 116	0	0
3	F	112	Total O 115 115	0	3

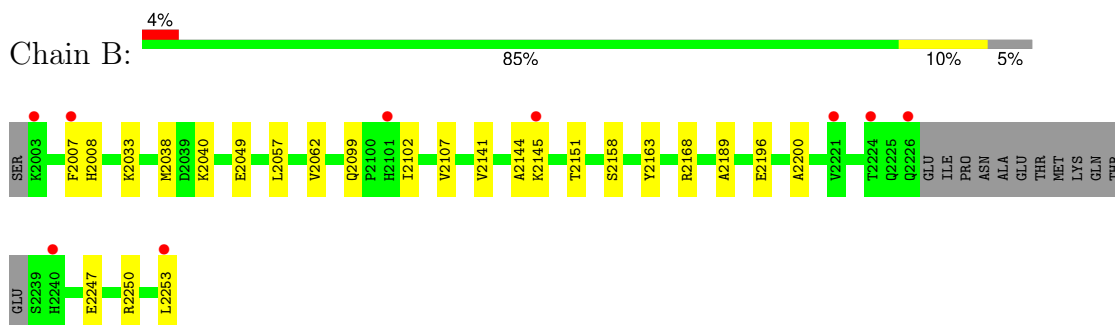
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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.

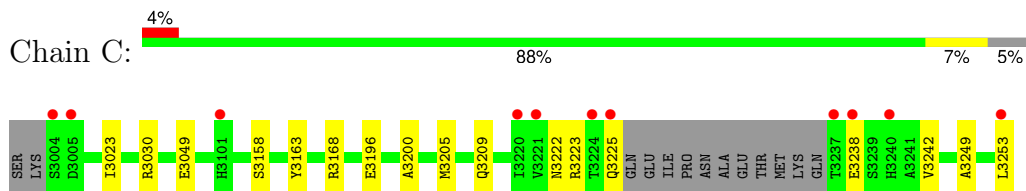
#### • Molecule 1: Uridine phosphorylase



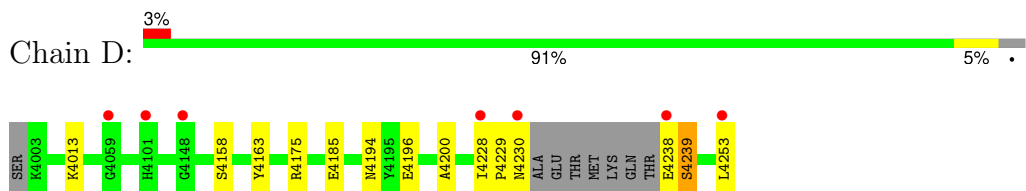
#### • Molecule 1: Uridine phosphorylase



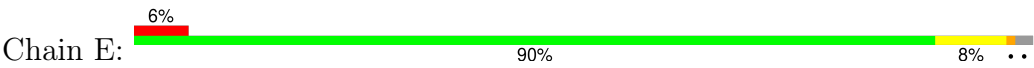
#### • Molecule 1: Uridine phosphorylase



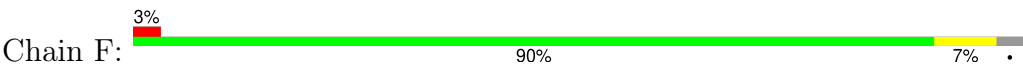
#### • Molecule 1: Uridine phosphorylase



● Molecule 1: Uridine phosphorylase



● Molecule 1: Uridine phosphorylase



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	89.60Å 125.03Å 134.95Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	27.00 – 1.64 26.95 – 1.64	Depositor EDS
% Data completeness (in resolution range)	95.4 (27.00-1.64) 95.4 (26.95-1.64)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.81 (at 1.64Å)	Xtriage
Refinement program	REFMAC 5.2.0003	Depositor
R, $R_{free}$	0.197 , 0.237 0.196 , 0.234	Depositor DCC
$R_{free}$ test set	8837 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	12.2	Xtriage
Anisotropy	0.430	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 45.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.28$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	12217	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	15.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.70% 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 [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: PO4

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	0.47	0/2073	0.61	0/2804
1	B	0.42	0/1979	0.54	0/2671
1	C	0.45	0/1929	0.59	0/2609
1	D	0.45	0/1972	0.60	0/2668
1	E	0.46	0/2004	0.61	0/2708
1	F	0.47	0/1967	0.59	0/2660
All	All	0.45	0/11924	0.59	0/16120

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	1988	0	2045	23	0
1	B	1897	0	1964	14	0
1	C	1864	0	1905	9	0
1	D	1905	0	1936	7	0
1	E	1938	0	1982	16	0
1	F	1901	0	1952	8	0
2	A	5	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	5	0	0	0	0
2	C	5	0	0	0	0
2	D	5	0	0	0	0
2	E	5	0	0	0	0
2	F	5	0	0	0	0
3	A	136	0	0	3	0
3	B	108	0	0	2	0
3	C	107	0	0	2	0
3	D	112	0	0	3	0
3	E	116	0	0	1	0
3	F	115	0	0	0	0
All	All	12217	0	11784	73	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 73 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:1039[B]:ASP:HB3	1:A:1056[B]:GLU:CG	1.72	1.18
1:A:1039[B]:ASP:CB	1:A:1056[B]:GLU:HG3	1.81	1.09
1:D:4185[B]:GLU:HG3	3:D:57:HOH:O	1.69	0.90
1:B:2144:ALA:HB1	1:B:2151:THR:HG21	1.59	0.84
1:E:5149:ALA:HB2	1:E:5240[B]:HIS:HE1	1.45	0.82

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	267/252 (106%)	261 (98%)	5 (2%)	1 (0%)	34 15

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	252/252 (100%)	248 (98%)	3 (1%)	1 (0%)	34	15
1	C	247/252 (98%)	241 (98%)	5 (2%)	1 (0%)	34	15
1	D	252/252 (100%)	248 (98%)	3 (1%)	1 (0%)	34	15
1	E	257/252 (102%)	254 (99%)	2 (1%)	1 (0%)	34	15
1	F	253/252 (100%)	250 (99%)	2 (1%)	1 (0%)	34	15
All	All	1528/1512 (101%)	1502 (98%)	20 (1%)	6 (0%)	34	15

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1163	TYR
1	C	3163	TYR
1	F	6163	TYR
1	B	2163	TYR
1	D	4163	TYR

### 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	
1	A	218/201 (108%)	211 (97%)	7 (3%)	39	12
1	B	206/201 (102%)	203 (98%)	3 (2%)	65	42
1	C	201/201 (100%)	198 (98%)	3 (2%)	65	42
1	D	206/201 (102%)	202 (98%)	4 (2%)	57	32
1	E	209/201 (104%)	205 (98%)	4 (2%)	57	32
1	F	206/201 (102%)	203 (98%)	3 (2%)	65	42
All	All	1246/1206 (103%)	1222 (98%)	24 (2%)	60	32

5 of 24 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	4196	GLU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	E	5003	LYS
1	D	4253	LEU
1	E	5092[A]	ILE
1	A	1230	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	4240	HIS
1	E	5225	GLN
1	F	6225	GLN
1	F	6101	HIS
1	D	4188	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 monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

6 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	PO4	B	7002	-	4,4,4	0.93	0	6,6,6	0.49	0
2	PO4	D	7004	-	4,4,4	0.86	0	6,6,6	0.71	0
2	PO4	E	7005	-	4,4,4	0.96	0	6,6,6	0.41	0
2	PO4	C	7003	-	4,4,4	0.95	0	6,6,6	0.53	0
2	PO4	F	7006	-	4,4,4	1.01	0	6,6,6	0.64	0
2	PO4	A	7001	-	4,4,4	0.91	0	6,6,6	0.54	0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

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, 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(Å²)	Q<0.9
1	A	251/252 (99%)	0.03	17 (6%) 17 15	5, 11, 35, 56	12 (4%)
1	B	239/252 (94%)	0.11	9 (3%) 40 38	7, 16, 31, 38	1 (0%)
1	C	239/252 (94%)	0.07	11 (4%) 32 30	7, 14, 32, 40	0
1	D	244/252 (96%)	-0.05	7 (2%) 51 50	5, 13, 25, 33	0
1	E	248/252 (98%)	0.07	15 (6%) 21 19	5, 11, 29, 37	3 (1%)
1	F	245/252 (97%)	-0.08	8 (3%) 46 44	5, 12, 26, 31	1 (0%)
All	All	1466/1512 (96%)	0.02	67 (4%) 32 30	5, 13, 29, 56	17 (1%)

The worst 5 of 67 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1228[A]	ILE	10.0
1	E	5231	ALA	7.6
1	E	5230	ASN	7.1
1	E	5233	THR	6.5
1	F	6002	SER	6.3

### 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 monosaccharides 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, 95<sup>th</sup> 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
2	PO4	B	7002	5/5	0.86	0.19	29,29,30,30	5
2	PO4	E	7005	5/5	0.90	0.17	34,34,35,35	5
2	PO4	C	7003	5/5	0.98	0.06	27,28,28,29	0
2	PO4	D	7004	5/5	0.99	0.06	16,17,17,18	0
2	PO4	A	7001	5/5	0.99	0.04	18,19,19,19	0
2	PO4	F	7006	5/5	0.99	0.07	14,16,16,17	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.