



# wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 17, 2024 – 02:09 PM JST

PDB ID : 8WEM  
Title : Bacteroides fragilis toxin 1  
Authors : Guo, Y.; Wen, Y.  
Deposited on : 2023-09-18  
Resolution : 1.80 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  
Xtriage (Phenix) : 1.13  
EDS : 3.0  
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)  
CCP4 : 9.0.002 (Gargrove)  
Density-Fitness : 1.0.11  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.38.2

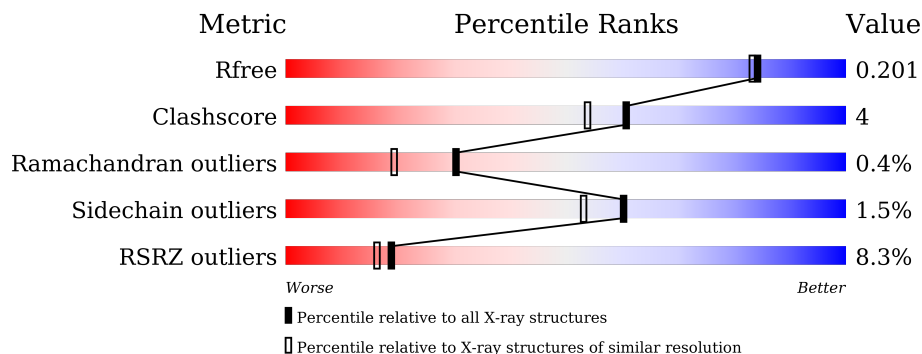
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

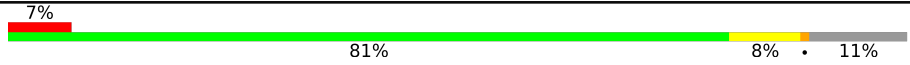

The reported resolution of this entry is 1.80 Å.

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  $\geq 3$ , 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	397	 7% 81% 8% • 11%
1	C	397	 7% 79% 8% • 13%

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 6106 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 Fragilysin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	C	347	2751	1744	447	545	15	0	5	0
1	A	353	2801	1774	456	556	15	0	6	0

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn) (labeled as "Ligand of Interest" by depositor).

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

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	C	287	Total	O	0	0
			287	287		
3	A	265	Total	O	0	0
			265	265		



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	82.67Å 82.67Å 280.07Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.37 – 1.80 46.37 – 1.80	Depositor EDS
% Data completeness (in resolution range)	99.9 (46.37-1.80) 99.9 (46.37-1.80)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.04 (at 1.81Å)	Xtrriage
Refinement program	PHENIX 1.19_4092	Depositor
R, $R_{free}$	0.176 , 0.202 0.175 , 0.201	Depositor DCC
$R_{free}$ test set	88841 reflections (2.20%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	33.0	Xtrriage
Anisotropy	0.462	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 46.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	6106	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	43.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.83% 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:  
ZN

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	0.41	0/2877	0.58	0/3892
1	C	0.43	0/2825	0.57	0/3820
All	All	0.42	0/5702	0.58	0/7712

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	2801	0	2684	23	0
1	C	2751	0	2634	28	0
2	A	1	0	0	0	0
2	C	1	0	0	0	0
3	A	265	0	0	2	0
3	C	287	0	0	5	0
All	All	6106	0	5318	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 4.

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:A:379:GLU:OE2	3:A:501:HOH:O	1.92	0.87
1:C:57:ASP:OD1	1:C:79:LYS:NZ	2.13	0.79
1:C:258[B]:VAL:HG11	1:C:389:LEU:HD22	1.64	0.78
1:A:258[B]:VAL:HG11	1:A:389:LEU:HD22	1.65	0.77
1:C:195:TYR:OH	3:C:501:HOH:O	1.98	0.74

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	355/397 (89%)	347 (98%)	6 (2%)	2 (1%)	22	11
1	C	346/397 (87%)	342 (99%)	3 (1%)	1 (0%)	37	25
All	All	701/794 (88%)	689 (98%)	9 (1%)	3 (0%)	30	19

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	147	ARG
1	C	149	ILE
1	A	146	THR

### 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	305/340 (90%)	299 (98%)	6 (2%)	50	40
1	C	299/340 (88%)	296 (99%)	3 (1%)	73	68
All	All	604/680 (89%)	595 (98%)	9 (2%)	60	53

5 of 9 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	289	SER
1	A	341	MET
1	A	161	SER
1	A	174	LYS
1	A	197	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	C	290	ASN
1	A	299	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 oligosaccharides in this entry.

### 5.6 Ligand geometry [i](#)

Of 2 ligands modelled in this entry, 2 are monoatomic - leaving 0 for Mogul analysis.

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(Å <sup>2</sup> )	Q<0.9
1	A	353/397 (88%)	0.37	29 (8%) 19 16	17, 40, 78, 101	6 (1%)
1	C	347/397 (87%)	0.27	29 (8%) 18 16	18, 38, 73, 97	5 (1%)
All	All	700/794 (88%)	0.32	58 (8%) 19 16	17, 39, 76, 101	11 (1%)

The worst 5 of 58 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	199	PRO	4.8
1	C	189	CYS	4.6
1	C	188	GLU	4.6
1	C	58	PHE	4.5
1	C	167	GLY	4.4

### 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](#)

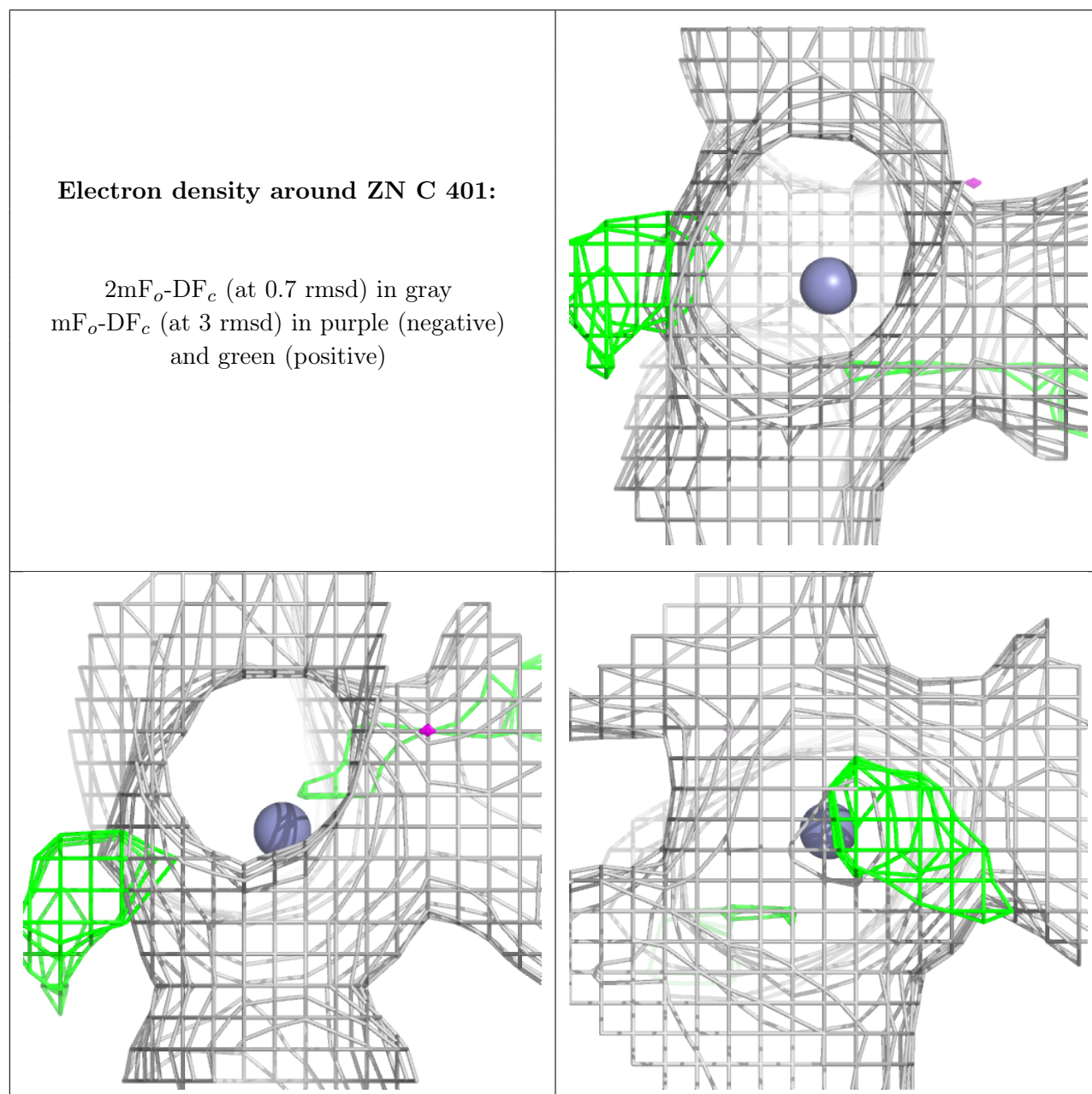
There are no monosaccharides in this entry.

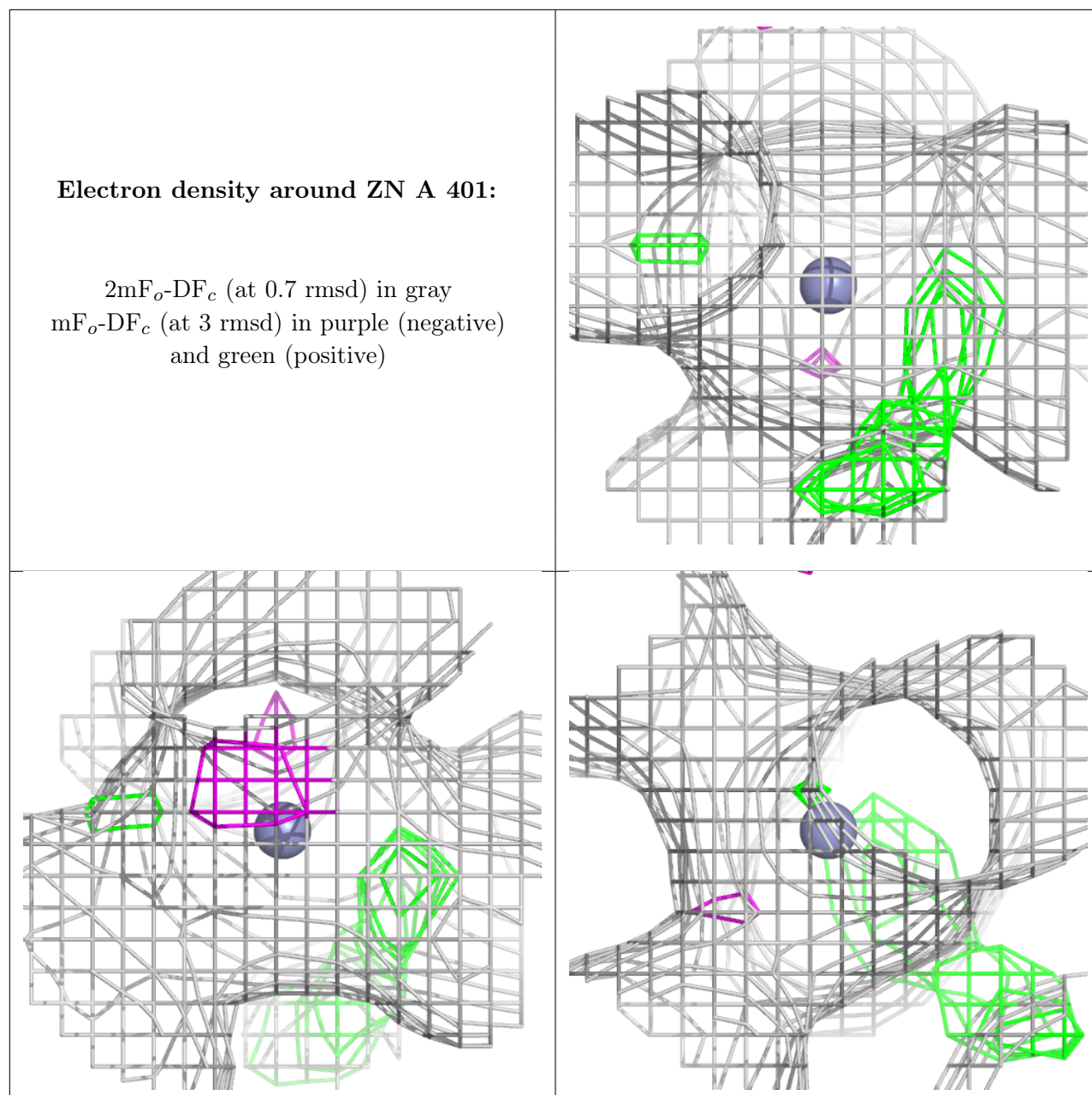
### 6.4 Ligands [i](#)

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	ZN	C	401	1/1	0.99	0.02	33,33,33,33	0
2	ZN	A	401	1/1	0.99	0.02	34,34,34,34	0

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.





## 6.5 Other polymers [i](#)

There are no such residues in this entry.