



# Full wwPDB X-ray Structure Validation Report ⓘ

Jan 31, 2016 – 06:44 PM GMT

PDB ID : 1CAN  
Title : CRYSTALLOGRAPHIC STUDIES OF THE BINDING OF PROTONATED AND UNPROTONATED INHIBITORS TO CARBONIC ANHYDRASE USING HYDROGEN SULPHIDE AND NITRATE ANIONS  
Authors : Mangani, S.; Hakansson, K.  
Deposited on : 1992-07-02  
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  
Mogul : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
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) : trunk26865

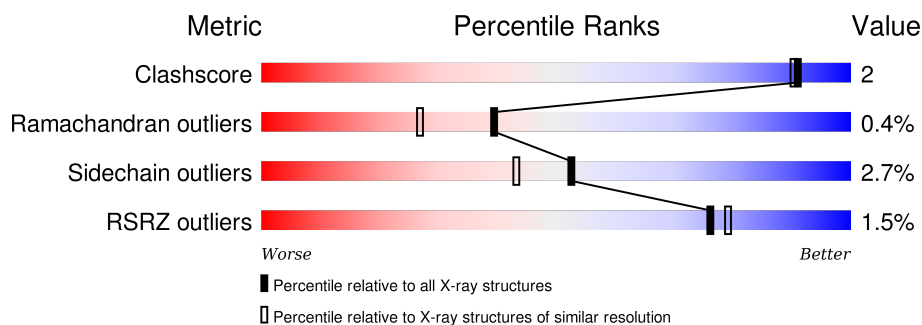
# 1 Overall quality at a glance

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	102246	5398 (1.90-1.90)
Ramachandran outliers	100387	5338 (1.90-1.90)
Sidechain outliers	100360	5339 (1.90-1.90)
RSRZ outliers	91569	4766 (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	A	260	

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 2325 atoms, of which 1 is hydrogen 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.

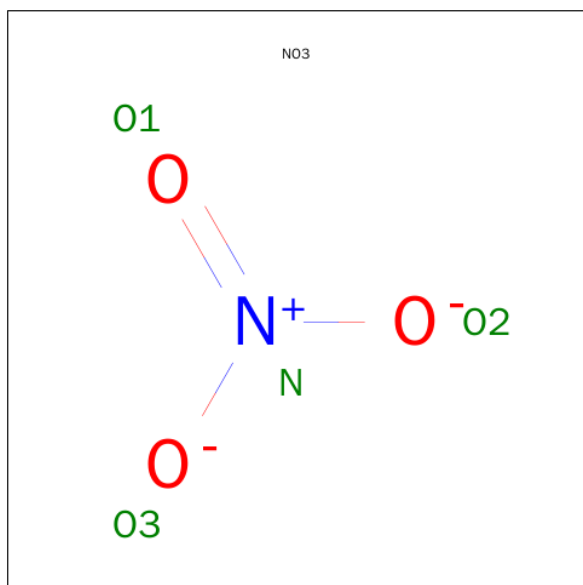
- Molecule 1 is a protein called CARBONIC ANHYDRASE II.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	260	Total	C	H	N	O	S	0	4	1
			2088	1338	1	361	386	2			

- Molecule 2 is MERCURY (II) ION (three-letter code: HG) (formula: Hg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	2	Total	Hg	0	0
			2	2		

- Molecule 3 is NITRATE ION (three-letter code: NO3) (formula: NO<sub>3</sub>).



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

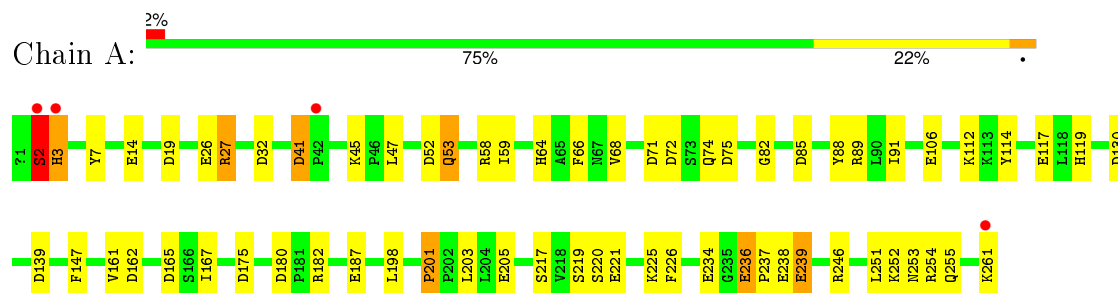
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	231	Total 231	O 231	0	0

### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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: CARBONIC ANHYDRASE II



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	42.70 Å 41.70 Å 73.00 Å 90.00° 104.60° 90.00°	Depositor
Resolution (Å)	(Not available) – 1.90 29.35 – 1.84	Depositor EDS
% Data completeness (in resolution range)	(Not available) ((Not available)-1.90) 91.1 (29.35-1.84)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	0.96 (at 1.84 Å)	Xtriage
Refinement program	TNT	Depositor
R, $R_{free}$	0.141 , (Not available) 0.197 , (Not available)	Depositor DCC
$R_{free}$ test set	No test flags present.	DCC
Wilson B-factor (Å <sup>2</sup> )	12.0	Xtriage
Anisotropy	0.234	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 77.5	EDS
Estimated twinning fraction	0.033 for h,-k,-h-l	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	1 of 19792 reflections (0.005%)	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	2325	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	16.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.91% 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, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: HG, ACE, NO3

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	1.20	8/2167 (0.4%)	1.84	63/2939 (2.1%)

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	2

All (8) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	239	GLU	CD-OE1	8.84	1.35	1.25
1	A	236	GLU	CD-OE2	8.45	1.34	1.25
1	A	187	GLU	CD-OE1	8.19	1.34	1.25
1	A	234	GLU	CD-OE1	8.01	1.34	1.25
1	A	26	GLU	CD-OE2	7.09	1.33	1.25
1	A	14[A]	GLU	CD-OE1	6.23	1.32	1.25
1	A	14[B]	GLU	CD-OE1	6.23	1.32	1.25
1	A	117	GLU	CD-OE2	5.14	1.31	1.25

All (63) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	175	ASP	CB-CG-OD1	12.67	129.70	118.30
1	A	182	ARG	NE-CZ-NH1	12.57	126.59	120.30
1	A	89	ARG	NE-CZ-NH2	-11.94	114.33	120.30
1	A	52	ASP	CB-CG-OD2	-11.92	107.57	118.30
1	A	66	PHE	CB-CG-CD2	-10.21	113.65	120.80

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	52	ASP	CB-CG-OD1	9.90	127.21	118.30
1	A	72	ASP	CB-CG-OD2	9.87	127.18	118.30
1	A	162	ASP	CB-CG-OD2	-9.75	109.52	118.30
1	A	19	ASP	CB-CG-OD2	-9.49	109.76	118.30
1	A	88	TYR	CZ-CE2-CD2	-9.31	111.42	119.80
1	A	2	SER	C-N-CA	8.59	143.16	121.70
1	A	58	ARG	NE-CZ-NH2	8.51	124.55	120.30
1	A	19	ASP	CB-CG-OD1	8.09	125.58	118.30
1	A	88	TYR	CG-CD2-CE2	7.98	127.68	121.30
1	A	182	ARG	NE-CZ-NH2	-7.91	116.35	120.30
1	A	221	GLU	CG-CD-OE1	7.72	133.74	118.30
1	A	27	ARG	NE-CZ-NH1	7.66	124.13	120.30
1	A	165	ASP	CB-CG-OD2	-7.61	111.45	118.30
1	A	246	ARG	NE-CZ-NH2	-7.60	116.50	120.30
1	A	3	HIS	N-CA-CB	7.52	124.14	110.60
1	A	75	ASP	CB-CG-OD1	7.24	124.82	118.30
1	A	162	ASP	CB-CG-OD1	7.12	124.71	118.30
1	A	27	ARG	NE-CZ-NH2	-6.96	116.82	120.30
1	A	58	ARG	N-CA-CB	-6.80	98.36	110.60
1	A	71	ASP	CB-CG-OD1	6.80	124.42	118.30
1	A	139	ASP	CB-CG-OD2	6.79	124.41	118.30
1	A	130	ASP	CB-CG-OD1	6.76	124.39	118.30
1	A	254	ARG	NE-CZ-NH2	-6.50	117.05	120.30
1	A	147	PHE	CB-CG-CD2	-6.45	116.28	120.80
1	A	41	ASP	CB-CG-OD1	-6.44	112.50	118.30
1	A	226	PHE	CB-CG-CD2	-6.27	116.41	120.80
1	A	3	HIS	N-CA-C	-5.94	94.96	111.00
1	A	175	ASP	CB-CG-OD2	-5.92	112.97	118.30
1	A	7	TYR	CG-CD2-CE2	-5.81	116.65	121.30
1	A	53	GLN	CG-CD-NE2	-5.72	102.96	116.70
1	A	180	ASP	CB-CG-OD1	-5.63	113.23	118.30
1	A	89	ARG	NE-CZ-NH1	5.60	123.10	120.30
1	A	74	GLN	CB-CA-C	-5.55	99.30	110.40
1	A	47	LEU	CB-CG-CD2	-5.54	101.58	111.00
1	A	85	ASP	CB-CG-OD2	5.54	123.29	118.30
1	A	64[A]	HIS	N-CA-CB	5.53	120.55	110.60
1	A	64[B]	HIS	N-CA-CB	5.53	120.55	110.60
1	A	75	ASP	CB-CG-OD2	-5.45	113.40	118.30
1	A	32	ASP	CB-CG-OD2	-5.44	113.41	118.30
1	A	27	ARG	CB-CA-C	5.40	121.19	110.40
1	A	7	TYR	CB-CG-CD2	-5.38	117.77	121.00
1	A	261	LYS	CA-C-O	-5.28	109.01	120.10

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	114	TYR	CB-CG-CD1	-5.24	117.86	121.00
1	A	91	ILE	CA-CB-CG2	5.22	121.33	110.90
1	A	198	LEU	CB-CG-CD2	-5.19	102.18	111.00
1	A	220	SER	N-CA-CB	-5.17	102.74	110.50
1	A	217	SER	CB-CA-C	-5.17	100.28	110.10
1	A	238	GLU	CG-CD-OE2	5.15	128.59	118.30
1	A	205	GLU	OE1-CD-OE2	-5.14	117.13	123.30
1	A	220	SER	CA-CB-OG	-5.12	97.37	111.20
1	A	221	GLU	OE1-CD-OE2	-5.08	117.20	123.30
1	A	68	VAL	CG1-CB-CG2	5.08	119.03	110.90
1	A	252	LYS	CA-CB-CG	-5.08	102.22	113.40
1	A	147	PHE	CD1-CE1-CZ	-5.07	114.01	120.10
1	A	219	SER	CA-C-O	5.07	130.75	120.10
1	A	236	GLU	CG-CD-OE1	5.06	128.43	118.30
1	A	47	LEU	CB-CA-C	-5.06	100.59	110.20
1	A	251	LEU	CB-CG-CD2	-5.03	102.45	111.00

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	2	SER	Mainchain,Peptide

## 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	2087	1	2026	9	1
2	A	2	0	0	0	0
3	A	4	0	0	0	0
4	A	231	0	0	1	2
All	All	2324	1	2026	9	2

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 (9) 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:112:LYS:NZ	4:A:423:HOH:O	2.33	0.61
1:A:161:VAL:CG1	1:A:225:LYS:HD2	2.39	0.53
1:A:236:GLU:HB3	1:A:237:PRO:CD	2.39	0.53
1:A:201:PRO:HA	1:A:203:LEU:HG	1.94	0.48
1:A:236:GLU:HB3	1:A:237:PRO:HD2	1.95	0.47
1:A:41:ASP:C	1:A:41:ASP:OD1	2.53	0.47
1:A:45:LYS:O	1:A:82:GLY:HA2	2.18	0.44
1:A:106:GLU:OE1	1:A:119:HIS:HE1	2.03	0.41
1:A:59:ILE:HG12	1:A:167:ILE:HD13	2.03	0.41

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:A:317:HOH:O	4:A:496:HOH:O[1_565]	2.08	0.12
1:A:53:GLN:OE1	4:A:508:HOH:O[2_455]	2.16	0.04

## 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
1	A	261/260 (100%)	249 (95%)	11 (4%)	1 (0%)	39 27

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	3	HIS

### 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	228/224 (102%)	222 (97%)	6 (3%)	54 45

All (6) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	2	SER
1	A	27	ARG
1	A	201	PRO
1	A	239	GLU
1	A	253	ASN
1	A	255	GLN

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

Mol	Chain	Res	Type
1	A	67	ASN
1	A	137	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 3 ligands modelled in this entry, 2 are monoatomic - leaving 1 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	NO3	A	511	2	3,3,3	0.19	0	3,3,3	0.22	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	NO3	A	511	2	-	0/0/0/0	0/0/0/0

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(Å <sup>2</sup> )	Q<0.9
1	A	259/260 (99%)	-0.35	4 (1%) 76 79	4, 13, 30, 43	3 (1%)

All (4) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	2	SER	5.7
1	A	3	HIS	3.3
1	A	261	LYS	2.2
1	A	42	PRO	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](#)

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
3	NO3	A	511	4/4	0.89	0.17	1.59	18,18,28,28	4

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
2	HG	A	510	1/1	1.00	0.06	-0.92	12,12,12,12	1
2	HG	A	262	1/1	1.00	0.40	-	7,7,7,7	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.