



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 21, 2020 – 04:04 PM BST

PDB ID : 2H0W  
Title : Post-cleavage state of the Thermoanaerobacter tengcongensis glmS ribozyme  
Authors : Klein, D.J.; Ferre-D'Amare, A.R.  
Deposited on : 2006-05-15  
Resolution : 2.40 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

<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 : 2.13.1  
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) : 2.13.1

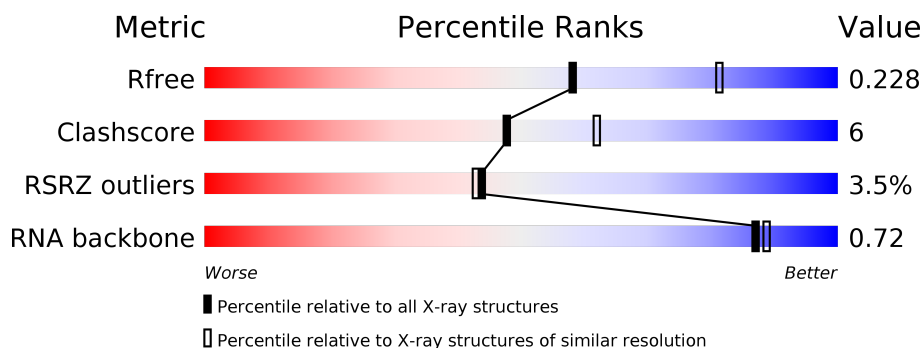
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.40 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	3907 (2.40-2.40)
Clashscore	141614	4398 (2.40-2.40)
RSRZ outliers	127900	3811 (2.40-2.40)
RNA backbone	3102	1174 (2.80-2.00)

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 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	26	<div> <div>8%</div> <div>42%</div> <div>27%</div> <div>• •</div> <div>23%</div> </div>
2	B	125	<div> <div>2%</div> <div>66%</div> <div>22%</div> <div>6%</div> <div>• •</div> </div>

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	MES	B	601	-	X	-	-

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 3231 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 RNA chain called glmS ribozyme product oligonucleotide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	20	Total	C	N	O	P	0	0	0
			414	186	74	135	19			

- Molecule 2 is a RNA chain called glmS ribozyme RNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	123	Total	C	N	O	P	0	0	0
			2641	1174	485	858	124			

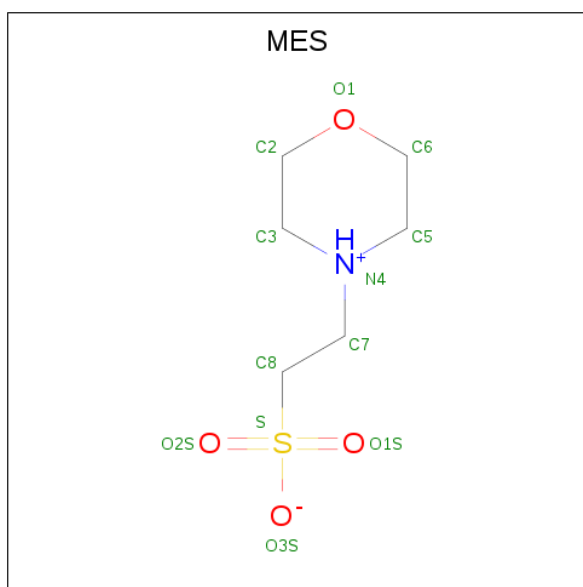
There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	21	C	-	INSERTION	GB 20517198
B	22	C	-	INSERTION	GB 20517198
B	145	A23	A	MODIFIED RESIDUE	GB 20517198

- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	4	Total	Mg	0	0
			4	4		

- Molecule 4 is 2-(N-MORPHOLINO)-ETHANESULFONIC ACID (three-letter code: MES) (formula: C<sub>6</sub>H<sub>13</sub>NO<sub>4</sub>S).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	B	1	Total	C	O	S	0	0
			5	1	3	1		
4	B	1	Total	C	N	O	S	0
			12	6	1	4	1	

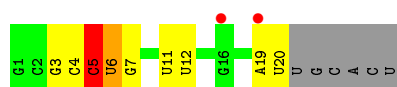
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	16	Total	O	0	0
			16	16		
5	B	139	Total	O	0	0
			139	139		

### 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: glmS ribozyme product oligonucleotide



- Molecule 2: glmS ribozyme RNA



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	181.09 Å 40.58 Å 71.80 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 2.40 28.13 – 2.40	Depositor EDS
% Data completeness (in resolution range)	(Not available) (30.00-2.40) 97.2 (28.13-2.40)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.92 (at 2.39 Å)	Xtriage
Refinement program	CNS	Depositor
R, $R_{free}$	0.219 , 0.240 0.204 , 0.228	Depositor DCC
$R_{free}$ test set	2014 reflections (9.42%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	35.7	Xtriage
Anisotropy	0.552	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.27 , 30.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	3231	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	34.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.38% 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

Bond lengths and bond angles in the following residue types are not validated in this section: MG, A23, MES

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	0.37	0/462	0.83	2/718 (0.3%)
2	B	0.36	0/2928	0.86	16/4568 (0.4%)
All	All	0.37	0/3390	0.86	18/5286 (0.3%)

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	1	0
2	B	0	2
All	All	1	2

There are no bond length outliers.

All (18) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	5	C	C2'-C3'-O3'	8.95	129.19	109.50
2	B	48	U	N1-C1'-C2'	8.70	125.31	114.00
2	B	57	C	C2'-C3'-O3'	8.29	127.73	109.50
2	B	103	U	C2'-C3'-O3'	7.45	125.89	109.50
2	B	130	G	N9-C1'-C2'	7.13	123.28	114.00
2	B	138	G	N9-C1'-C2'	6.18	122.03	114.00
2	B	41	G	N9-C1'-C2'	5.95	121.73	114.00
2	B	55	G	N9-C1'-C2'	5.94	121.73	114.00
2	B	94	G	O4'-C1'-N9	5.88	112.90	108.20
2	B	58	A	N9-C1'-C2'	5.86	121.62	114.00
2	B	48	U	O4'-C1'-N1	5.76	112.80	108.20
2	B	102	G	C2'-C3'-O3'	5.50	122.50	113.70

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	94	G	N9-C1'-C2'	5.43	121.06	114.00
1	A	5	C	C4'-C3'-C2'	5.23	107.83	102.60
2	B	59	U	N1-C1'-C2'	5.17	120.72	114.00
2	B	130	G	O4'-C1'-N9	5.10	112.28	108.20
2	B	57	C	C4'-C3'-C2'	5.09	107.69	102.60
2	B	91	U	N1-C1'-C2'	5.06	120.58	114.00

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	5	C	C3'

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	114	G	Sidechain
2	B	55	G	Sidechain

## 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	414	0	214	6	0
2	B	2641	0	1324	22	0
3	B	4	0	0	0	0
4	B	17	0	13	1	0
5	A	16	0	0	0	0
5	B	139	0	0	2	0
All	All	3231	0	1551	28	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 6.

All (28) 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:5:C:H2'	5:B:353:HOH:O	1.64	0.97
2:B:144:A:H2'	2:B:145:A23:H5'	1.57	0.86
2:B:144:A:C2'	2:B:145:A23:H5'	2.25	0.67
2:B:55:G:H5'	5:B:263:HOH:O	1.95	0.65
2:B:101:G:O2'	2:B:102:G:H5'	1.98	0.64
2:B:48:U:O2'	2:B:49:U:H5''	2.00	0.61
2:B:27:U:H2'	2:B:28:U:C6	2.45	0.52
2:B:144:A:C6	2:B:145:A23:N1	2.79	0.51
2:B:55:G:H4'	2:B:56:A:OP2	2.09	0.50
2:B:47:U:O2'	2:B:48:U:H5'	2.12	0.49
2:B:103:U:O2'	2:B:104:A:OP1	2.22	0.49
2:B:23:G:H2'	2:B:24:G:H8	1.77	0.48
1:A:11:U:H2'	1:A:12:U:C6	2.49	0.48
2:B:100:U:H2'	2:B:101:G:O4'	2.14	0.48
2:B:34:G:H4'	2:B:59:U:OP2	2.15	0.47
1:A:19:A:H2'	1:A:20:U:C6	2.50	0.46
2:B:105:A:H2'	2:B:106:A:O4'	2.16	0.46
1:A:20:U:H6	1:A:20:U:O5'	1.97	0.45
2:B:94:G:O2'	2:B:95:A:OP2	2.27	0.45
2:B:33:U:H2'	2:B:34:G:O4'	2.17	0.45
1:A:6:U:O2'	1:A:7:G:H5'	2.18	0.44
2:B:54:A:H1'	2:B:105:A:N1	2.32	0.44
2:B:144:A:N6	2:B:145:A23:N1	2.65	0.44
2:B:46:G:O2'	2:B:47:U:H5'	2.18	0.44
2:B:65:G:N2	4:B:602:MES:O3S	2.48	0.43
2:B:144:A:H2'	2:B:145:A23:C5'	2.37	0.41
2:B:125:A:H2'	2:B:126:U:O4'	2.20	0.41
1:A:3:G:O2'	1:A:4:C:H5'	2.21	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

There are no protein molecules in this entry.

### 5.3.2 Protein sidechains [i](#)

There are no protein molecules in this entry.

### 5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	A	19/26 (73%)	1 (5%)	1 (5%)
2	B	122/125 (97%)	12 (9%)	4 (3%)
All	All	141/151 (93%)	13 (9%)	5 (3%)

All (13) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	A	6	U
2	B	37	G
2	B	49	U
2	B	55	G
2	B	58	A
2	B	59	U
2	B	67	U
2	B	68	G
2	B	80	U
2	B	95	A
2	B	103	U
2	B	104	A
2	B	139	G

All (5) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
1	A	5	C
2	B	55	G
2	B	57	C
2	B	67	U
2	B	102	G

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

1 non-standard protein/DNA/RNA residue is 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	A23	B	145	2	19,28,29	0.70	0	19,43,46	1.58	4 (21%)

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
2	A23	B	145	2	-	0/3/35/36	0/4/4/4

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	145	A23	O3'-PC-O1C	-3.35	106.92	115.76
2	B	145	A23	C1'-N9-C4	-2.51	122.23	126.64
2	B	145	A23	O2C-PC-O1C	2.50	117.97	109.89
2	B	145	A23	C5-C6-N6	2.19	123.68	120.35

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	145	A23	5	0

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 6 ligands modelled in this entry, 4 are monoatomic - leaving 2 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	$\# Z  > 2$
4	MES	B	602	-	12,12,12	1.03	1 (8%)	14,16,16	1.20	2 (14%)
4	MES	B	601	-	4,4,12	1.05	0	5,6,16	4.76	4 (80%)

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
4	MES	B	602	-	-	0/6/14/14	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	B	602	MES	C8-S	2.65	1.81	1.77

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	B	601	MES	O1S-S-C8	7.89	123.98	107.54
4	B	601	MES	O2S-S-C8	4.35	116.60	107.54
4	B	601	MES	O3S-S-O1S	-4.07	83.97	109.48
4	B	601	MES	O3S-S-O2S	-3.85	85.40	109.48
4	B	602	MES	O2S-S-C8	2.45	109.86	106.92
4	B	602	MES	O1S-S-C8	2.09	109.43	106.92

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	B	602	MES	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

### 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	A	20/26 (76%)	0.16	2 (10%) <b>7</b> <b>6</b>	19, 36, 88, 91	0
2	B	122/125 (97%)	-0.44	3 (2%) 57 55	18, 30, 51, 87	0
All	All	142/151 (94%)	-0.35	5 (3%) 44 43	18, 30, 60, 91	0

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	23	G	5.0
2	B	103	U	3.3
1	A	19	A	3.3
1	A	16	G	2.0
2	B	24	G	2.0

### 6.2 Non-standard residues in protein, DNA, RNA chains

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(Å <sup>2</sup> )	Q<0.9
2	A23	B	145	25/26	0.70	0.30	67,74,76,78	0

### 6.3 Carbohydrates

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(Å <sup>2</sup> )	Q<0.9
3	MG	B	502	1/1	0.84	0.27	51,51,51,51	0
4	MES	B	602	12/12	0.93	0.26	41,45,50,50	0
4	MES	B	601	5/12	0.94	0.23	63,63,64,65	0
3	MG	B	501	1/1	0.95	0.12	25,25,25,25	0
3	MG	B	503	1/1	0.96	0.14	24,24,24,24	0
3	MG	B	504	1/1	0.96	0.19	21,21,21,21	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.