Full wwPDB X-ray Structure Validation Report ⓘ

Aug 16, 2020 – 07:17 PM BST

PDB ID	:	6PSK
Title	:	Crystal structure of the complex between periplasmic domains of antiholin RI and holin T from T4 phage, in P6522
Authors	:	Kuznetsov, V.B.; Krieger, I.V.; Sacchettini, J.C.
Deposited on	:	2019-07-12
Resolution	:	2.20 Å(reported)

The following versions of software and data (see references ⓘ) were used in the production of this report:

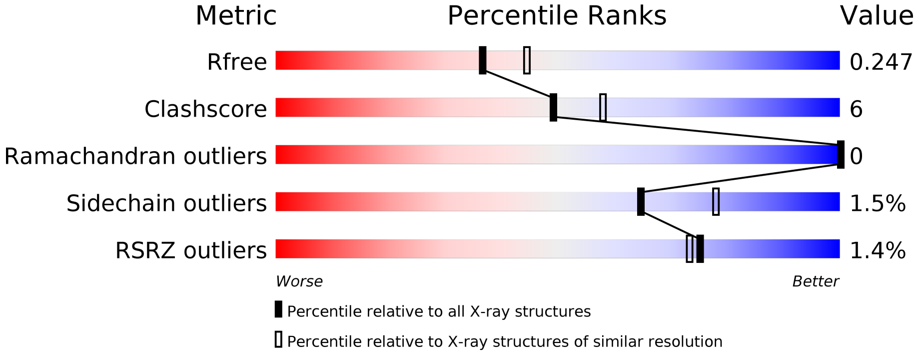
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.13.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (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

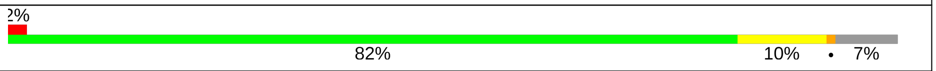
The reported resolution of this entry is 2.20 Å.

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	4898 (2.20-2.20)
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)
RSRZ outliers	127900	4800 (2.20-2.20)

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 >=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 <=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	R	74	
2	T	155	

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
3	SO4	R	101	-	-	X	-

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 1891 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 Antiholin.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	R	72	596	374	98	116	2	6	0	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
R	24	MSE	-	initiating methionine	UNP I7AU04

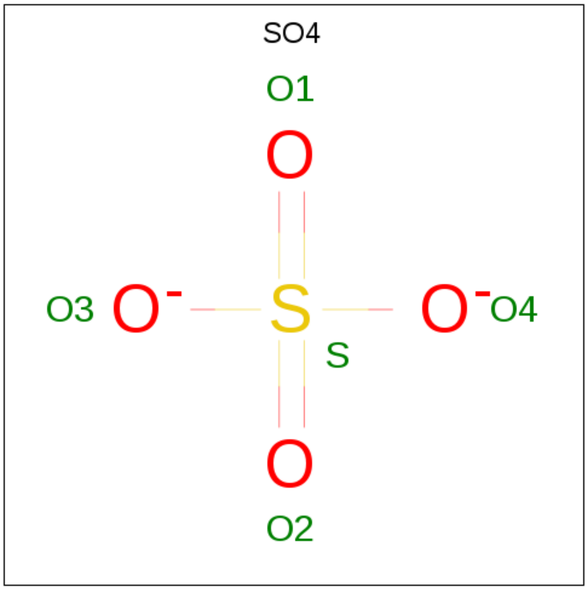
- Molecule 2 is a protein called Holin.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
2	T	144	1176	757	195	219	2	3	0	0	0

There are 13 discrepancies between the modelled and reference sequences:

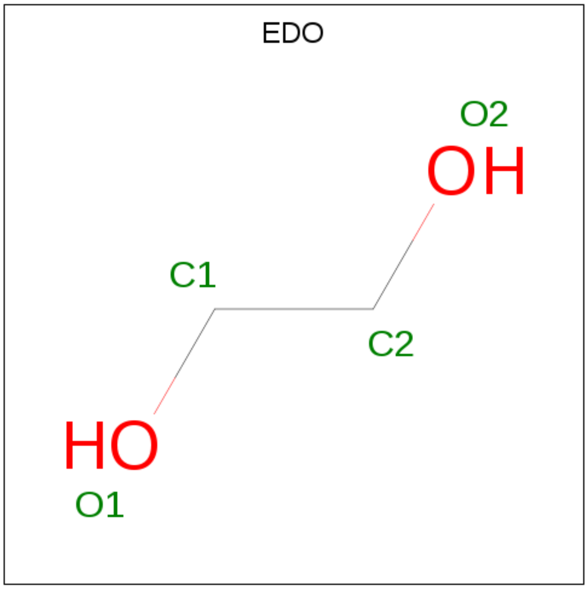
Chain	Residue	Modelled	Actual	Comment	Reference
T	219	LYS	-	expression tag	UNP A0A2U8QQK7
T	220	LEU	-	expression tag	UNP A0A2U8QQK7
T	221	ALA	-	expression tag	UNP A0A2U8QQK7
T	222	ALA	-	expression tag	UNP A0A2U8QQK7
T	223	ALA	-	expression tag	UNP A0A2U8QQK7
T	224	LEU	-	expression tag	UNP A0A2U8QQK7
T	225	GLU	-	expression tag	UNP A0A2U8QQK7
T	226	HIS	-	expression tag	UNP A0A2U8QQK7
T	227	HIS	-	expression tag	UNP A0A2U8QQK7
T	228	HIS	-	expression tag	UNP A0A2U8QQK7
T	229	HIS	-	expression tag	UNP A0A2U8QQK7
T	230	HIS	-	expression tag	UNP A0A2U8QQK7
T	231	HIS	-	expression tag	UNP A0A2U8QQK7

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	R	1	Total O S 5 4 1	0	0
3	T	1	Total O S 5 4 1	0	0
3	T	1	Total O S 5 4 1	0	0

- Molecule 4 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).



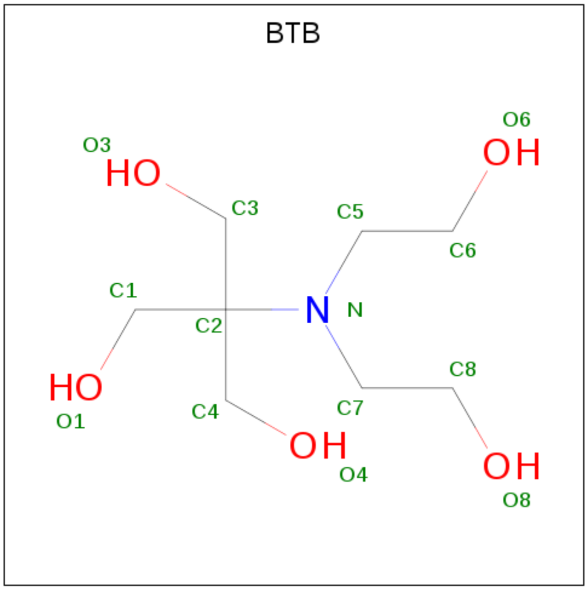
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	R	1	Total C O 4 2 2	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	R	1	Total C O 4 2 2	0	0
4	T	1	Total C O 4 2 2	0	0
4	T	1	Total C O 4 2 2	0	0

- Molecule 5 is 2-[BIS-(2-HYDROXY-ETHYL)-AMINO]-2-HYDROXYMETHYL-PROPANE-1,3-DIOL (three-letter code: BTB) (formula: $C_8H_{19}NO_5$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	T	1	Total C N O 14 8 1 5	0	0

- Molecule 6 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	T	1	Total Cl 1 1	0	0

- Molecule 7 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	R	23	Total O 23 23	0	0

Continued on next page...

Continued from previous page...

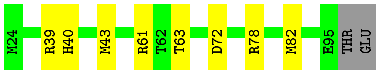
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
7	T	50	50	50	0	0

3 Residue-property plots ⓘ

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: Antiholin

Chain R:



- Molecule 2: Holin

Chain T:

4 Data and refinement statistics ⓘ

Property	**Value**	**Source**
Space group	P 65 2 2	Depositor
Cell constants a, b, c, α, β, γ	120.67Å 120.67Å 85.51Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	44.59 – 2.20 44.59 – 2.20	Depositor EDS
% Data completeness (in resolution range)	99.9 (44.59-2.20) 99.9 (44.59-2.20)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.16	Depositor
$< I/\sigma(I) >$ [1]	2.07 (at 2.20Å)	Xtriage
Refinement program	PHENIX 1.11.1_2575	Depositor
R, R_{free}	0.213 , 0.247 0.213 , 0.247	Depositor DCC
R_{free} test set	918 reflections (4.80%)	wwPDB-VP
Wilson B-factor (Å^2)	42.2	Xtriage
Anisotropy	0.067	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 43.3	EDS
L-test for twinning[2]	$< \vert L \vert > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	1891	wwPDB-VP
Average B, all atoms (Å^2)	44.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.26% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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: SO4, CL, EDO, BTB

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	R	0.40	0/604	0.49	0/799
2	T	0.44	0/1205	0.54	0/1628
All	All	0.43	0/1809	0.53	0/2427

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 ⓘ

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	R	596	0	555	8	0
2	T	1176	0	1141	12	0
3	R	5	0	0	3	0
3	T	10	0	0	0	0
4	R	8	0	12	1	0
4	T	8	0	12	0	0
5	T	14	0	19	1	0
6	T	1	0	0	1	0
7	R	23	0	0	2	0
7	T	50	0	0	0	0
All	All	1891	0	1739	21	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 (21) 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)
2:T:212:ARG:NH2	6:T:306:CL:CL	2.40	0.91
2:T:139:MSE:HG2	2:T:158:ALA:HB2	1.75	0.68
1:R:72:ASP:H	4:R:102:EDO:H12	1.59	0.66
2:T:190:MSE:HE2	2:T:203:LEU:HD22	1.88	0.55
2:T:86:LEU:HD11	2:T:98:SER:HB3	1.89	0.53
2:T:77:ARG:HD3	2:T:220:LEU:HD12	1.90	0.53
7:R:219:HOH:O	2:T:129:SER:HA	2.11	0.50
1:R:39:ARG:NH1	3:R:101:SO4:S	2.85	0.49
1:R:63:THR:HG21	1:R:82:MSE:HG2	1.93	0.49
5:T:303:BTB:H51	5:T:303:BTB:H12	1.35	0.48
1:R:61:ARG:NH2	7:R:201:HOH:O	2.21	0.47
2:T:93:SER:HB3	2:T:190:MSE:HE3	1.95	0.47
2:T:77:ARG:HD2	2:T:77:ARG:HA	1.63	0.46
1:R:40:HIS:NE2	3:R:101:SO4:O3	2.42	0.46
1:R:39:ARG:NH1	3:R:101:SO4:O3	2.49	0.45
1:R:78:ARG:O	1:R:82:MSE:HB2	2.16	0.45
2:T:77:ARG:CD	2:T:220:LEU:HD12	2.50	0.41
1:R:43:MSE:HG3	1:R:43:MSE:O	2.20	0.41
2:T:96:ASP:OD1	2:T:194:ARG:NH1	2.53	0.41
2:T:87:GLN:HG3	2:T:118:GLU:OE2	2.21	0.41
2:T:218:LYS:HD2	2:T:220:LEU:H	1.86	0.41

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	
1	R	70/74 (95%)	65 (93%)	5 (7%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	T	142/155 (92%)	138 (97%)	4 (3%)	0	100	100
All	All	212/229 (93%)	203 (96%)	9 (4%)	0	100	100

There are no Ramachandran outliers to report.

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	R	67/63 (106%)	67 (100%)	0	100	100
2	T	128/133 (96%)	125 (98%)	3 (2%)	50	63
All	All	195/196 (100%)	192 (98%)	3 (2%)	65	78

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	T	77	ARG
2	T	147	ASN
2	T	218	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 9 ligands modelled in this entry, 1 is monoatomic - leaving 8 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	EDO	R	102	-	3,3,3	0.44	0	2,2,2	0.49	0
4	EDO	R	103	-	3,3,3	0.35	0	2,2,2	0.61	0
5	BTB	T	303	-	13,13,13	0.50	0	7,16,16	1.03	0
3	SO4	T	301	-	4,4,4	0.17	0	6,6,6	0.18	0
4	EDO	T	304	-	3,3,3	0.38	0	2,2,2	0.58	0
3	SO4	R	101	-	4,4,4	0.19	0	6,6,6	0.12	0
3	SO4	T	302	-	4,4,4	0.31	0	6,6,6	0.28	0
4	EDO	T	305	-	3,3,3	0.49	0	2,2,2	0.24	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
4	EDO	T	304	-	-	1/1/1/1	-
4	EDO	R	103	-	-	1/1/1/1	-
4	EDO	R	102	-	-	0/1/1/1	-
4	EDO	T	305	-	-	0/1/1/1	-
5	BTB	T	303	-	-	4/21/21/21	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (6) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	T	303	BTB	C4-C2-C3-O3
5	T	303	BTB	C8-C7-N-C5
5	T	303	BTB	N-C5-C6-O6
4	R	103	EDO	O1-C1-C2-O2
4	T	304	EDO	O1-C1-C2-O2
5	T	303	BTB	N-C7-C8-O8

There are no ring outliers.

3 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	R	102	EDO	1	0
5	T	303	BTB	1	0
3	R	101	SO4	3	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, 95th 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($Å^2$)	Q<0.9
1	R	66/74 (89%)	-0.11	0	100	100	35, 45, 60, 73	0
2	T	141/155 (90%)	-0.33	3 (2%)	63	61	30, 39, 62, 80	1 (0%)
All	All	207/229 (90%)	-0.26	3 (1%)	75	73	30, 41, 62, 80	1 (0%)

All (3) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	T	220	LEU	2.8
2	T	129	SER	2.6
2	T	127	GLU	2.4

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

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

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, 95th 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($Å^2$)	Q<0.9
5	BTB	T	303	14/14	0.65	0.29	43,61,69,70	0
3	SO4	T	302	5/5	0.82	0.32	55,61,72,80	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors($Å^2$)	Q<0.9
4	EDO	R	102	4/4	0.83	0.13	58,58,61,64	0
4	EDO	T	305	4/4	0.84	0.20	43,50,51,56	0
4	EDO	T	304	4/4	0.89	0.21	57,61,70,70	0
4	EDO	R	103	4/4	0.92	0.14	57,59,63,66	0
3	SO4	R	101	5/5	0.94	0.14	70,73,81,82	0
3	SO4	T	301	5/5	0.96	0.13	65,67,73,76	0
6	CL	T	306	1/1	0.99	0.18	53,53,53,53	1

6.5 Other polymers ⓘ

There are no such residues in this entry.