

wwPDB X-ray Structure Validation Summary Report (i)

Nov 14, 2023 – 04:27 PM JST

PDB ID : 5ZR2

Title: Crystal Structure of Phosphoserine Phosphatase Mutant (H9A) from Enta-

moeba histolytica in complex with Phosphoserine

Authors: Kumari, P.; Gourinath, S.

Deposited on : 2018-04-23

Resolution : 2.95 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
A user guide is available at

https://www.wwpdb.org/validation/2017/XrayValidationReportHelp with specific help available everywhere you see the (i) 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 (1)) were used in the production of this report:

 $Mol Probity \quad : \quad 4.02b\text{--}467$

Mogul: 1.8.5 (274361), CSD as541be (2020)

Xtriage (Phenix) : 1.13 EDS : 2.36

Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)

 $Refmac \quad : \quad 5.8.0158$

CCP4 : 7.0.044 (Gargrove)

Ideal geometry (proteins) : Engh & Huber (2001) Ideal geometry (DNA, RNA) : Parkinson et al. (1996)

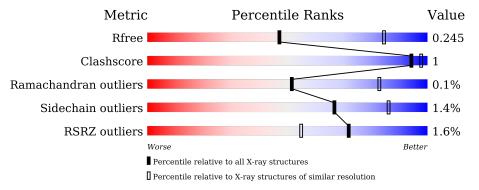
Validation Pipeline (wwPDB-VP) : 2.36

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure: X- $RAY\ DIFFRACTION$

The reported resolution of this entry is 2.95 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	Similar resolution
Metric	$(\# \mathrm{Entries})$	$(\# ext{Entries}, ext{ resolution range}(ext{Å}))$
R_{free}	130704	3104 (3.00-2.92)
Clashscore	141614	3462 (3.00-2.92)
Ramachandran outliers	138981	3340 (3.00-2.92)
Sidechain outliers	138945	3343 (3.00-2.92)
RSRZ outliers	127900	2986 (3.00-2.92)

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 >=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	A	211	88%	5%	7%
1	В	211	90%	•	7%
1	С	211	90%	•	6%
1	D	211	89%	5%	7%



2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 6177 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 Phosphoglycerate mutase family protein, putative.

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace		
1	Λ	197	Total	С	N	О	S	0	0	0
1	A	197	1529	976	256	288	9	U	U	0
1	D	197	Total	С	N	О	S	0	0	0
1	. В	197	1517	970	256	282	9	U	U	
1	С	198	Total	С	N	О	S	0	0	0
1		190	1535	980	259	288	8	0	U	0
1	D	107	Total	С	N	О	S	0	0	0
1		197	1489	952	252	277	8	0	U	U

There are 28 discrepancies between the modelled and reference sequences:

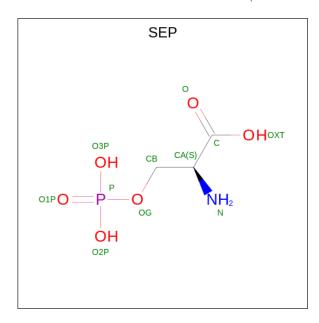
Chain	Residue	Modelled	Actual	Comment	Reference
A	9	ALA	HIS	engineered mutation	UNP N9V397
A	206	HIS	-	expression tag	UNP N9V397
A	207	HIS	-	expression tag	UNP N9V397
A	208	HIS	-	expression tag	UNP N9V397
A	209	HIS	-	expression tag	UNP N9V397
A	210	HIS	-	expression tag	UNP N9V397
A	211	HIS	-	expression tag	UNP N9V397
В	9	ALA	HIS	engineered mutation	UNP N9V397
В	206	HIS	-	expression tag	UNP N9V397
В	207	HIS	-	expression tag	UNP N9V397
В	208	HIS	-	expression tag	UNP N9V397
В	209	HIS	-	expression tag	UNP N9V397
В	210	HIS	-	expression tag	UNP N9V397
В	211	HIS	-	expression tag	UNP N9V397
С	9	ALA	HIS	engineered mutation	UNP N9V397
С	206	HIS	-	expression tag	UNP N9V397
С	207	HIS	-	expression tag	UNP N9V397
С	208	HIS	-	expression tag	UNP N9V397
С	209	HIS	-	expression tag	UNP N9V397
С	210	HIS	-	expression tag	UNP N9V397
С	211	HIS	-	expression tag	UNP N9V397



Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	9	ALA	HIS	engineered mutation	UNP N9V397
D	206	HIS	-	expression tag	UNP N9V397
D	207	HIS	-	expression tag	UNP N9V397
D	208	HIS	-	expression tag	UNP N9V397
D	209	HIS	-	expression tag	UNP N9V397
D	210	HIS	-	expression tag	UNP N9V397
D	211	HIS	-	expression tag	UNP N9V397

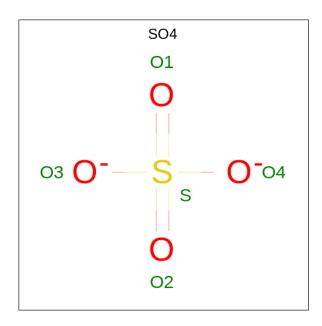
 \bullet Molecule 2 is PHOSPHOSERINE (three-letter code: SEP) (formula: $\mathrm{C_3H_8NO_6P}).$



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf		
2	Λ	1	Total	С	N	О	Р	0	0	
2	A	1	11	3	1	6	1	0	U	
2	D	1	Total	С	N	О	Р	0	0	
2	2 B	1	11	3	1	6	1	0		
2	С	1	Total	С	N	О	Р	0	0	
2		1	11	3	1	6	1	0	U	
2	D	1	Total	С	N	О	Р	0	0	
2	ט	1	11	3	1	6	1	0	U	

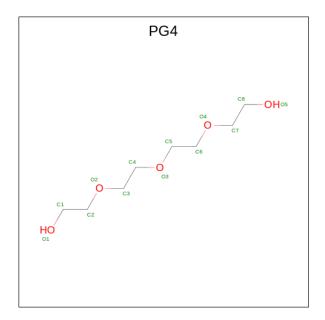
 \bullet Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: $\mathrm{O_4S}).$





Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total O S 5 4 1	0	0
3	С	1	Total O S 5 4 1	0	0

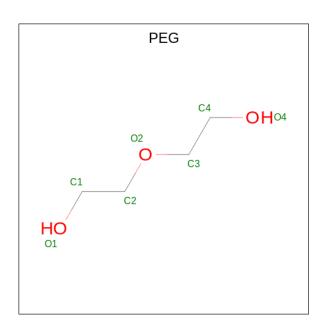
 \bullet Molecule 4 is TETRAETHYLENE GLYCOL (three-letter code: PG4) (formula: $\mathrm{C_8H_{18}O_5}).$



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C O 13 8 5	0	0

 $\bullet \ \ Molecule \ 5 \ is \ DI(HYDROXYETHYL)ETHER \ (three-letter \ code: \ PEG) \ (formula: \ C_4H_{10}O_3). \\$





Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	В	1	Total 7	C 4	O 3	0	0

• Molecule 6 is water.

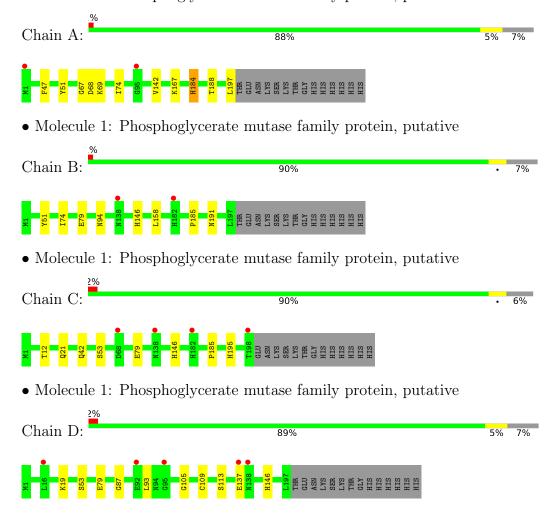
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	10	Total O 10 10	0	0
6	В	6	Total O 6 6	0	0
6	С	6	Total O 6 6	0	0
6	D	11	Total O 11 11	0	0



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: Phosphoglycerate mutase family protein, putative





4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants	95.11Å 111.74Å 178.25Å	Depositor
a, b, c, α , β , γ	90.00° 90.00° 90.00°	Depositor
Resolution (Å)	89.28 - 2.95	Depositor
rtesolution (A)	26.66 - 2.95	EDS
% Data completeness	97.5 (89.28-2.95)	Depositor
(in resolution range)	97.7 (26.66-2.95)	EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) > 1$	2.51 (at 2.94Å)	Xtriage
Refinement program	REFMAC 5.8.0135	Depositor
D D.	0.207 , 0.250	Depositor
R, R_{free}	0.208 , 0.245	DCC
R_{free} test set	944 reflections (4.74%)	wwPDB-VP
Wilson B-factor (Å ²)	46.3	Xtriage
Anisotropy	0.474	Xtriage
Bulk solvent $k_{sol}(e/Å^3)$, $B_{sol}(Å^2)$	0.33 , 35.0	EDS
L-test for twinning ²	$ < L > = 0.47, < L^2> = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	6177	wwPDB-VP
Average B, all atoms $(Å^2)$	43.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: The largest off-origin peak in the Patterson function is 11.61% of the height of the origin peak. No significant pseudotranslation is detected.

²Theoretical 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.



¹Intensities estimated from amplitudes.

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: SEP, PG4, SO4, PEG

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		
IVIOI		RMSZ	# Z > 5	RMSZ	# Z > 5	
1	A	0.40	0/1562	0.64	0/2115	
1	В	0.38	0/1549	0.59	0/2096	
1	С	0.38	0/1567	0.60	0/2121	
1	D	0.39	0/1521	0.58	0/2063	
All	All	0.39	0/6199	0.60	0/8395	

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	1529	0	1508	6	0
1	В	1517	0	1504	4	0
1	С	1535	0	1523	4	0
1	D	1489	0	1442	4	0
2	A	11	0	5	0	0
2	В	11	0	5	0	0
2	С	11	0	5	0	0
2	D	11	0	5	0	0
3	A	5	0	0	0	0



Continued	trom	mmoninonic	maaa
COHABABACA		DIEUIUU	DUIUE
0 0 1000100000			

Mol	Chain	Non-H	H(model)	$\mathbf{H}(\mathbf{added})$	Clashes	Symm-Clashes
3	С	5	0	0	0	0
4	A	13	0	18	0	0
5	В	7	0	10	0	0
6	A	10	0	0	0	0
6	В	6	0	0	0	0
6	С	6	0	0	0	0
6	D	11	0	0	0	0
All	All	6177	0	6025	16	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 1.

The worst 5 of 16 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	$egin{aligned} & ext{Interatomic} \ & ext{distance} \ & ext{(Å)} \end{aligned}$	$\begin{array}{c} \text{Clash} \\ \text{overlap } (\text{\AA}) \end{array}$
1:C:42:GLN:HG3	1:C:195:HIS:HB3	1.87	0.57
1:C:12:THR:HG21	1:C:21:GLN:NE2	2.24	0.53
1:A:47:PHE:HB3	1:A:142:VAL:HG23	1.91	0.53
1:A:184:HIS:ND1	1:A:184:HIS:N	2.59	0.51
1:B:79:GLU:HB2	1:B:146:HIS:CG	2.48	0.49

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	Perce	entiles
1	A	195/211 (92%)	185 (95%)	9 (5%)	1 (0%)	29	64
1	В	195/211 (92%)	185 (95%)	10 (5%)	0	100	100
1	С	196/211 (93%)	188 (96%)	8 (4%)	0	100	100
1	D	195/211 (92%)	185 (95%)	10 (5%)	0	100	100



Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
All	All	781/844 (92%)	743 (95%)	37 (5%)	1 (0%)	51 83	

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	67	GLY

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	166/183 (91%)	163 (98%)	3 (2%)	59	82	
1	В	163/183 (89%)	161 (99%)	2 (1%)	71	88	
1	С	166/183 (91%)	165 (99%)	1 (1%)	86	94	
1	D	154/183 (84%)	151 (98%)	3 (2%)	57	81	
All	All	$649/732 \ (89\%)$	640 (99%)	9 (1%)	67	86	

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

Mol	Chain	Res	Type
1	D	113	SER
1	D	137	GLU
1	В	94	ASN
1	В	185	PRO
1	С	185	PRO

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	146	HIS
1	С	172	ASN
1	С	21	GLN
1	С	15	ASN



Continued from previous page...

Mol	Chain	Res	Type
1	С	146	HIS

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)

8 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	Thurs Chair Day Lin		Link	Bond lengths			Bond angles		
MIOI	Type	Chain	Res	Lilik	Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z >2
2	SEP	A	301	-	9,10,10	0.90	1 (11%)	12,14,14	2.06	3 (25%)
4	PG4	A	303	-	12,12,12	0.64	0	11,11,11	0.67	0
2	SEP	С	301	-	9,10,10	0.91	1 (11%)	12,14,14	1.90	2 (16%)
2	SEP	D	301	-	9,10,10	0.88	1 (11%)	12,14,14	2.02	4 (33%)
3	SO4	С	302	-	4,4,4	0.34	0	6,6,6	0.07	0
3	SO4	A	302	-	4,4,4	0.34	0	6,6,6	0.08	0
5	PEG	В	302	-	6,6,6	0.57	0	5,5,5	0.58	0
2	SEP	В	301	-	9,10,10	0.87	1 (11%)	12,14,14	2.23	3 (25%)

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.



, ,	moone	no	outliers	α f	that	kind	woro	idontific	$^{\circ}$
_	means	110	outhers	OI	unat	MIIIU	were	raemim	zu.

Mol	\mathbf{Type}	Chain	Res	Link	Chirals	Torsions	Rings
2	SEP	A	301	-	-	3/10/10/10	-
4	PG4	A	303	-	-	5/10/10/10	-
2	SEP	С	301	-	-	3/10/10/10	-
2	SEP	D	301	-	-	3/10/10/10	-
5	PEG	В	302	-	-	3/4/4/4	-
2	SEP	В	301	-	-	1/10/10/10	-

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	\mathbf{Z}	Observed(A)	Ideal(Å)
2	A	301	SEP	OXT-C	-2.16	1.23	1.30
2	С	301	SEP	OXT-C	-2.14	1.23	1.30
2	D	301	SEP	OXT-C	-2.11	1.23	1.30
2	В	301	SEP	OXT-C	-2.09	1.23	1.30

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$Observed(^o)$	$\operatorname{Ideal}({}^{o})$
2	В	301	SEP	OG-CB-CA	6.50	113.72	108.06
2	A	301	SEP	OG-CB-CA	5.64	112.98	108.06
2	D	301	SEP	OG-CB-CA	5.12	112.53	108.06
2	С	301	SEP	OG-CB-CA	5.08	112.49	108.06
2	A	301	SEP	OXT-C-O	-2.62	118.15	124.09

There are no chirality outliers.

5 of 18 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	301	SEP	CB-OG-P-O1P
2	С	301	SEP	N-CA-CB-OG
2	D	301	SEP	N-CA-CB-OG
4	A	303	PG4	O2-C3-C4-O3
4	A	303	PG4	O1-C1-C2-O2

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^{th} 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>	$\#\mathrm{RSRZ}{>}2$	$OWAB(A^2)$	Q < 0.9
1	A	197/211 (93%)	-0.06	2 (1%) 82 68	26, 37, 56, 71	0
1	В	197/211 (93%)	-0.05	2 (1%) 82 68	29, 41, 55, 65	0
1	С	198/211 (93%)	0.00	4 (2%) 65 48	28, 41, 58, 74	0
1	D	197/211 (93%)	0.13	5 (2%) 57 40	32, 48, 61, 75	0
All	All	789/844 (93%)	0.00	13 (1%) 72 55	26, 41, 59, 75	0

The worst 5 of 13 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	С	198	THR	4.0
1	С	182	HIS	3.7
1	В	138	ASN	3.2
1	D	138	ASN	2.9
1	A	95	GLY	2.6

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^{th} 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	$\mathbf{B} ext{-}\mathbf{factors}(\mathring{\mathbf{A}}^2)$	Q<0.9
3	SO4	A	302	5/5	0.72	0.38	109,110,114,116	0
5	PEG	В	302	7/7	0.75	0.37	62,63,64,64	0
3	SO4	С	302	5/5	0.83	0.27	95,95,98,99	0
4	PG4	A	303	13/13	0.86	0.24	39,39,41,42	0
2	SEP	A	301	11/11	0.91	0.25	61,63,64,65	0
2	SEP	В	301	11/11	0.91	0.21	66,67,70,71	0
2	SEP	D	301	11/11	0.92	0.19	66,69,71,71	0
2	SEP	С	301	11/11	0.94	0.18	50,54,57,60	0

6.5 Other polymers (i)

There are no such residues in this entry.

