*Polish J. Chem.*, **75**, 1449–1456 (2001)

# On the Peptides Related to Nucleotide Sequences of Archeal tRNAs

**by I.Z. Siemion[1*], I. Strug[1], E. Wyszko[2] and J. Barciszewski[2]**

[1]*Faculty of Chemistry, Wrocław University, 50-383 Wrocław, Poland*
[2]*Institute of Bioorganic Chemistry of the Polish Academy of Sciences, 61-704 Poznań, Poland*

*(Received May 14th, 2001)*

It is shown that the nucleotide sequences of archeal tRNAs correspond to peptide sequences, some of which resemble very much the contemporary regulatory peptides. In particular, the nucleotide sequence of archeal A0980 $tRNA^{Ala}$ corresponds to a putative polypeptide, which is homologous to thymopoietin, a peptide hormone produced in thymus and endowed with immunomodulatory activity. The 23 amino acid long synthetic peptide corresponding to A0980 interacts strongly with TAR RNA fragment of HIV mRNAs as well as binds specifically to $tRNA^{Phe}$, whereas $tRNA^{Leu}$ and $tRNA^{Met}$ are deprived of such ability.

**Key words**: archeal tRNAs, RNA world, peptides related to tRNAs, thymopoietin homology

The discovery of catalytic activities of RNA molecules [1,2] prompted many scientists to speculate about the primary of "RNA world" in the genesis of life [3,4,5,6]. The idea on RNA world as an early stage in evolution of life is now widely accepted [7]. Looking for the relics of RNA world in the contemporary forms of life, Jeffares *et al*. [8] concluded that tRNAs molecules may possess such a character, because they fulfil the criteria to be such the relics, *i.e*. they are ubiquitous, highly conserved and central to the metabolism of all life. It should be also noted that in 1981 Cedergren *et al*. [9] suggested that tRNAs are the molecular equivalents of living fossils. In this context we decided to examine, if the polynucleotide sequences of archeal tRNAs could be translated into the polypeptide sequences. Of course, the nowadays role of archeal tRNAs has nothing common with polypeptide coding. However, such a role cannot be excluded for the precursors of archeal tRNAs, existing in "RNA world". Then, the eventual ability of archeal tRNAs to code the uninterrupted polypeptide sequences can be considered as the relict of this "RNA world".

*Corresponding author; Tel: (048) (71 ) 32 04 206; Fax: (048) (71 ) 32 82 348 or (048) (71 ) 32 04 215; E-mail:siemion@wchuwr.chem.uni.wroc.pl

## EXPERIMENTAL

*A0980 peptide synthesis*: The peptide was synthesized by solid-phase Fmoc-strategy, using a Milligen/Biosearch 9500 automatic peptide synthesizer. The synthesis was performed using Wang and TGT resins (both supplied by Novabiochem). In the case of procedure with Wang resin obtained material did not contain A0980 peptide. The successful synthesis can be performed with TGT resin only. A mixture of trifluoroacetic acid/$H_2O$/phenol/ thioanisole/ethanedithiol (82.5:5:5:5:2.5) was used as the deprotecting and cleaving reagent on the final step of the synthesis. The deprotection time should not exceed 10 min. The peptide is acid labile and it is destroyed when the action of deprotecting mixture exceeds 10 min. The peptide solution was filtered into cold diethyl ether and precipitating peptide was filtered, washed with ether, and dried *in vacuum*. The obtained material was prepurified by extraction with water and further lyophilized. The crude peptide was additionally purified by HPLC (RP-C4), then transformed into acetate form by lyophilization from 10% acetic acid. After the purification the HPLC picture was not changed sufficiently; the purity of the sample was 89% (Fig. 1). The purity and homogeneity of the sample was checked by HPLC (C4 Vaydack) and mass spectrometry (Finnigan, MAT TSQ 700 spectrometer equipped with a Finnigan electrospray ionization source). In the mass spectra beside the signals of the +7, +4, and +3 main peptide ions of the low intensity, the signals originated from the fragments: RIPAG, VRIPAG, RSRVRIPAG, GPVVQRK, GPVVQRKD, GPVVQRKDAR, GPVVQRKDARL, and GPVVQRKDARLARG were visible. HPLC retention time of the A0980 peptide was 45 min, but the signal obtained was not sharp.

*TAR RNA – AO980 peptide complex formation*: TAR RNA sample was obtained according to the procedure described by Wyszko *et al*. [10]. The sample was labeled at 5′-end with [$^{32}$P] ATP and 10 U T4 polynucleotide kinase (Promega) [ 11 ]. The formation of TAR RNA – AO980 peptide complex was examined by electrophoresis on 0.7% agarose gel. $1\times10^{-9}$ M TAR RNA was incubated with 0.26, 0.4, $0.53\times10^{-9}$ M peptide solution. The experiments were carried out in buffer I (50 mM Tris-HCl, pH 7.5, 70 mM NaCl,1 mM EDTA, 0.1% Nonidet) and buffer II (25 mM Tris-HCI, pH 7.5, 100 mM NaCl, 1 mM $MgCl_2$, 0.1% Triton X-100) for 30 min at 25°C. The results are shown in Fig. 2.

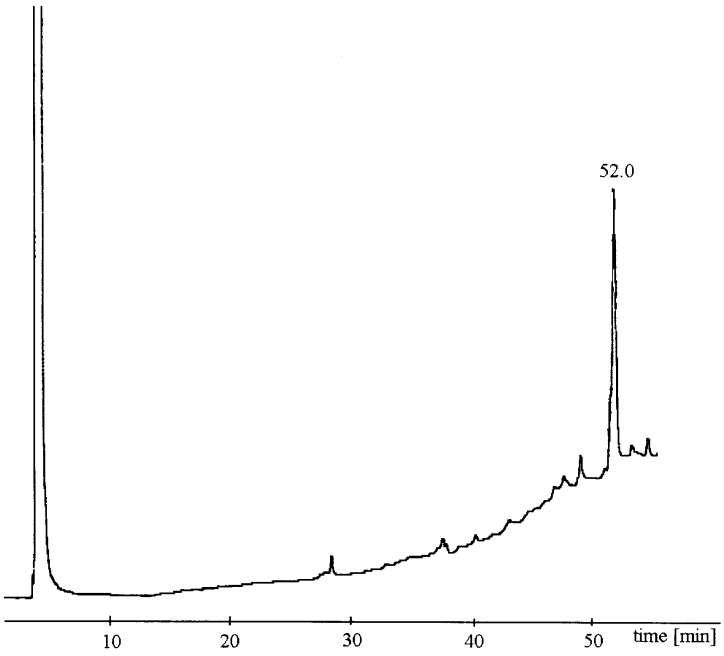


**Figure 1.** HPLC chromatogram of the crude A0980 peptide.

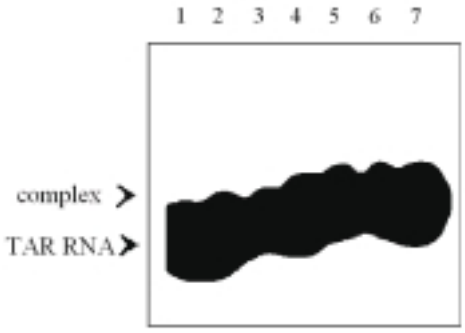


**Figure 2.** Complex formation between $1\times10^{-9}$ M (10000 cpm) TAR RNA and 0.26, 0.4 and $0.53\times10^{-9}$ M peptide, carried out in buffer I or buffer II for 30 min at 25°C. Lane 1 – control RNA in water; lanes 2–4 – complex formation in buffer I; lanes 5–7 – complex formation in buffer II. Analysis was done on 0.7% agarose gel electrophoresis.

*Isolation and labeling of tRNAs*: The $tRNA^{Phe}$ from yeast (Boehringer), $tRNA^{Leu}$ and $tRNA^{Met}$ from *E. coli* (Sigma) were purified on 15% polyacrylamide gel with 7 M urea in 0.05 Tris/borate buffer. The samples of tRNAs were 3′-end labeled with [$^{32}$P] pCp and RNA ligase [11]. [$^{32}$P]-labeled tRNAs were purified by electrophoresis on 10% polyacrylamide gel (PAGE) with 7 M urea, eluted from the gel and renatured [12].

*tRNAs binding reactions and electrophoretic mobility gel shift assays*: The tRNA-peptide complex formation assay was performed at 25°C for 30 min in pH 7.5 buffer (25 mM Tris-HCl, 100 mM NaCl, 1 mM $MgCl_2$, 0.1% Triton X-100).$1\times10^{-9}$ M (10000 cpm) $tRNA^{Phe}$, $tRNA^{Leu}$, and $tRNA^{Met}$ were incubated with $0.4\times10^{-9}$ M peptide in total volume of 10 µl. Analysis of the complexation was carried out on 0.7% agarose gel in 0.05 M Tris/borate buffer. The association constants $K_a$ for binding of RNAs with peptide was estimated by an autoradiogram densitometric analysis [13]. *L. luteus* $tRNA^{Phe}$ was used in the last experiments.

*RNase footprint assay*: In a footprint reaction of tRNA-peptide complexes, the following amounts of RNases (Pharmacia) were used: T1 – ($2\times10^{-4}$U), V1 – ($6\times10^{-2}$ U), and S1 – (3U). For localization of the peptide binding site, 4 µg of tRNA and 40000 cpm labeled $tRNA^{Phe}$ were digested with 0.04 U of T1 RNase in 20 mM sodium acetate buffer (pH 4.5, 7 M urea, 1 mM EDTA and 0.05% xylene cyanol) and analyzed on 20% polyacrylamide gel with 7 M urea in 0.09 M Tris/borate buffer [14].

*Lead ion induced RNA hydrolysis*: Labeled $tRNA^{Phe}$ was supplemented with 4 µg of non-specific RNA. The complex formation was performed as described above for 30 min, 1 µl of 20 mM $Pb(CH_3COO)_2$ was added for RNA hydrolysis. Reaction (15 min) was stopped by mixing with 10 µl of 8 M urea/dyes 20 mM EDTA and loaded on 20% polyacrylamide gel.

## RESULTS AND DISCUSSION

We analyzed the nucleotide sequences of 60 archeal tRNAs collected in EMBL Data Bank. The unexpected result was that 41.7% of the sequences, read from their 5′-terminus, gave open reading frames of 24–25 amino acid residues. The 40% of the sequences were interrupted by stop codon in one position, and only 18.3% were interrupted in such a way in two positions. Ten of 60 (16.7%) of archeal tRNAs analyzed were of $tRNA^{Ala}$ type. The tRNAs of other amino acids were distinctly less represented in the whole group. Therefore, we concentrated our attention on tRNAs be-

longing to $tRNA^{Ala}$ type. In this subgroup two tRNAs (designated as A0580 and A0660) code for the same polynucleotide sequence, and one (A0380) can be translated to the shortened peptide consisted of 16 amino acids. Therefore, our experimental group consisted of eight polypeptide sequences, derived from $tRNA^{Ala}$ molecules:

**Table 1.** The amino acid sequences of peptides derived from translation of $tRNA^{Ala}$. The sequences are arranged-in the order of diminishing homology to the AO980 peptide.

| Amino acid sequences | Accession number |
|---|---|
| GPVV$^{x}$RKDARLARGRSRVRIPAGP | AO980 |
| GLVAQRESAAFARPRPRVQIPPSP | AO340 |
| GPIAQW$^{x}$SASFARRMPWVRIPVGP | AO385 |
| GPIAQW$^{x}$SASFARRMPWVGIPVGP | AO420 |
| GLVAQLVERRLCKAEALGPNPSKS | AO670 |
| GPVAQTGRAPPLQGGGPGFKSRWV | AO580 |
| GPIAQPGRAPPLQGGGPGFKSRWVQ | AO680 |
| GPVAQPGRASALQAEGPGFESPPVH | AO940 |

From these sequences five are uninterrupted (ORF) ones, one (AO980) is interrupted in position 5, and two (AO385 and AO420) are interrupted in position 7. The position 5 is occupied by Gln in all other sequences. As we can see, a one-step mutation of the transitional type (CAA – UAA or CAG – UAG mutation) leads from Gln codon into the Stop codon. Therefore, we can assume that ORF A0980 peptide precursor had the Gln residue in position 5. It should be also noted that the same situation appears for AO385 and AO420 peptides. The position 7 in other peptides of this group is occupied by Glu, Lys, Gly or Val. Excluding Val, in all other cases one-step mutation of transversional type transforms the correspondent codons into the Stop codons. The uninterrupted amino acid sequence GPVVQRKDARLARGRSRVRIPAGP (ORF), which corresponds to AO980 $tRNA^{Ala}$, shows a peculiar and unexpected homology to C-terminal part of thymopoietin, the peptide hormone produced in thymus and used in human and animal organisms for the regulation of the immune system. Thymopoietins were firstly isolated by G. Goldstein's group from bovine thymi [15]. According to the results of Harris *et al*. [16], human thymopoietin is of the following sequence:

PEFLEDPSVLTKDKLKSELVANNVTLPAGEQRKDVYVQLYLQHLTARN

The comparison of this sequence with the sequence of AO980 peptide:

–PAGEQRKDVYVQLYLQHLTARNRP–

GPVVQRKDARLARGRSRVRIPAGP

shows (when not only identical but also similar residues: Ile – Leu; Leu – Val are taken into account) a 25% homology. The thymopoietin homology diminishes in the series of sequences shown in Table l. However, all the sequences shown demonstrate the common origin from the same precursor RNA sequence.

It has been shown recently that the thymopoietin-like segments appear in the huge number of various nucleic acid binding domains of proteins [17]. In particular the thymopoietin homology appears in archeal histone-like proteins, what suggests a very ancient origin of thymopoietin-like sequences. Of course, the appearance of thymopoietin-like sequence among the sequences related to archeal tRNAs may be considered as the natural curiosity only. However, such a curiosity seems to be worth of attention. We showed that AO980 peptide can also interact with RNAs in a specific manner. To examine this phenomenon we performed the synthesis of 23 amino acid long peptide, denoted in the text as an AO980 peptide. To make the synthesis easier we eliminated the proline residue from its C-terminus.

The synthesis of AO980 peptide was performed by solid phase method, using the Fmoc procedure. Application of the Boc-procedure appeared to be impossible, because of instability of the peptide in the acid media. The splitting of the peptide from the resin must also be performed with caution. We used for this purpose the trifluoroacetic acid-water-thioanisole-ethanedithiol-phenol mixture. When the exposition time of the peptide to the action of the mixture exceeded 10 min, its degrada-

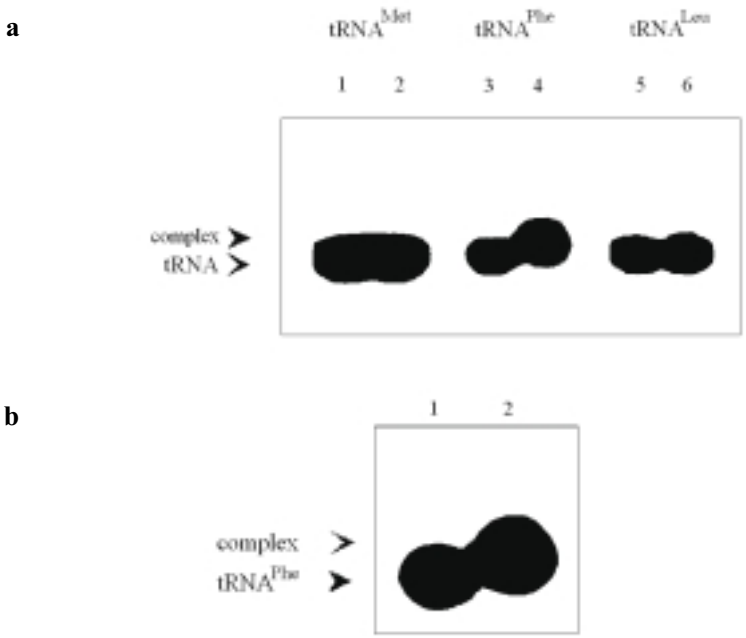


**Figure 3.** **a**: Autoradiogram of peptide binding to tRNAs. Complex formation between $1\times10^{-9}$ M (10000 cpm) $tRNA^{Met}$, $tRNA^{Phe}$, $tRNA^{Leu}$ and $0.4\times10^{-9}$ M peptide was carried out in buffer II for 30 min at 25°C. Lanes 1, 3, 5 – control RNA in buffer; lanes 2, 4, 6 – comlexes. Analysis was done on 0.7% agarose gel electrophoresis. **b**: Complex formation between $1\times10^{-9}$ M (10000 cpm) $tRNA^{Phe}$ and $0.53\times10^{-9}$ M peptide carried out in bufer II. Lane 1 – control RNA in buffer; lane 2 – complex. Conditions as above.

tion to the smaller fragments has been observed. The peptide is also unstable in the relatively mild conditions of MS experiments (ESI technique). The analysis of the signals observed in MS spectra enabled us to conclude that -Arg-Gly-Arg-fragment, situated near the center of the peptide chain is a principal region of peptide fragmentation. It is visible that its purity was satisfactory (Fig. 1). The crude product was additionally purified by HPLC to the purity of 89%.

We examined a complexation of AO980 peptide with different RNA molecules like TAR RNAand $tRNA^{Phe}$, $tRNA^{Leu}$, and $tRNA^{Met}$. All of them are rich in guanosine moieties in their single stranded fragments. Because the peptide AO980 contains six Arg residues its interactions with tRNA molecules could be mostly due to arginine - guanosine recognition [18]. An additional argument for using of $tRNA^{Met}$ consisted in its known role in polypeptide synthesis initiation. TAR RNA was chosen as the control molecule because it contains G rich loop.

From the investigated tRNAs only one – $tRNA^{Phe}$ – forms a complex with the AO980 peptide (see Fig. 3a and 3b). To check it we carried out the agarose gel shift assay. The agarose gel electrophoresis was performed in two buffers pH 7.5: buffer I (50 mM Tris-HCl, 70 mM NaCl, 1 mM EDTA, 0.1 % Nonidet) and buffer II (25 mM Tris-HCl, 100 mM NaCl, 0.1% Triton X-100,1 mM $MgCl_2$), with increasing amounts of the peptide (0.26, 0.40, and 0.52 nM). For the experiments shown in Fig. 3 buffer II was used. The complexation of AO980 peptide with $tRNA^{Phe}$ was additionally examined (Fig. 3b) for increased amounts of the peptide.

The fact that AO980 peptide interacts with $tRNA^{Phe}$ only suggests the specific character of peptide – nucleic acid recognition. However, in the case of non-specific interactions with the basic peptide, no big differences in the complexation should be observed for all three specific tRNAs. There are also the evidences that the peptides, rich in Arg and /or Lys residues derived from aminoacyl-tRNA synthetase (AARS) do not interact non-specifically with tRNA molecules [19]. Some eukaryotic AARS have different extensions at N or C ends but their influence on specificity is rather limited.

Using the method described by Wyszko and Barciszewska [13] we determined the association constant of AO980 peptide – $tRNA^{Phe}$ complex. The $K_a$ value was found to be equal approximately $1.84\times10^9$ $M^{-1}$. It arguments for the quite strong stability of the complex.

The observation of the specificity of $tRNA^{Phe}$ – AO980 peptide interaction is, however, weakened by the finding that TAR RNA forms also the complex with the peptide (see Fig. 2). The affinity of TAR RNA to AO980 was of the same range as that of $tRNA^{Phe}$. TAR RNA is a double stranded RNA hairpin fragment, located at the 5′-end of all HIV mRNAs. It binds specifically viral Tat-protein (a small protein which plays a vital role in the initiation of transcription process [20]). Our experiments show that the AO980 peptide resembles in this respect Tat-protein.

The AO980 peptide shows some similarity to the RKKRRQRRC sequence of Tat, which is the binding site of the protein to TAR RNA. The peptide with the latter sequence was used recently to obtain the inhibitors of the HIV-1 replication [21].

For determination of the binding site of $tRNA^{Phe}$, a limited specific hydrolysis of $tRNA^{Phe}$ was used. We applied RNase T1 (guanosine specific), S1 (single stranded RNA specific), V1 (double stranded RNA specific) and $Pb^{+2}$-induced limited hydrolysis of $tRNA^{Phe}$. The results of these experiments are shown in Fig. 4. The complexation of the peptide does not influence the hydrolysis accomplished by guanosine specific T1 RNase, but inhibits to some extent $Pb^{+2}$ -induced hydrolysis, as well as the hydrolysis with double stranded specific V1 RNase. At the same time the hydrolysis with S1 RNase seems to be stimulated by the peptide complexation. Thus, we can conclude that the peptide interacts probably with double stranded fragments of $tRNA^{Phe}$. In particular, acceptor stem of the $tRNA^{Phe}$ molecule seems to be especially well adjusted for the interaction with the peptide. If this is true, the influence of the AO980 peptide on aminoacylation reaction of $tRNA^{Phe}$ could be expected.

Thus, nucleotide sequences of archeal tRNAs could really correspond to the peptide sequences and some of them resemble the nowadays peptide sequences. As it is shown in this work these peptides may interact (presumably specifically) with RNA molecules. This result could have some interesting connections to the problem of life evolution, and, in particular, to the problem of RNA world – RNA-protein world transition [8]. However, we have not enough evidences now to put the problem in such a broader perspective.

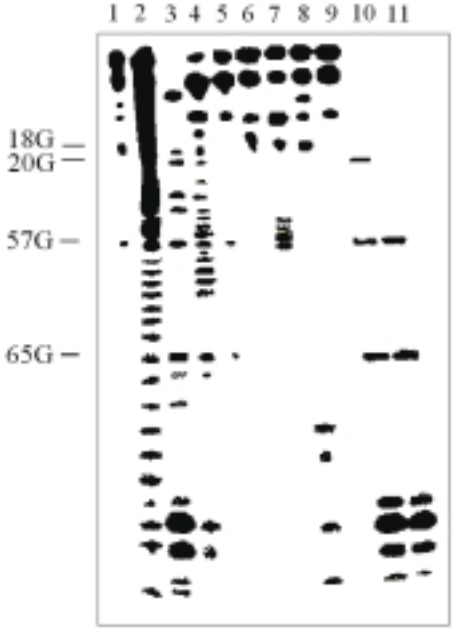


**Figure 4.** RNase and $Pb^{+2}$ footprinting assay. Autoradiogram of 20% polyacrylamide gel with 7 M urea showing hydrolysis products of $[^{32}P]$ $tRNA^{Phe}$ – peptide complex obtained by $Pb^{+2}$ (lanes 4, 5), RNase T1 (lanes 10, 11 ), RNase V1 (lanes 8, 9) and RNase S1 (lanes 6, 7). Lane 1 – control $tRNA^{Phe}$ incubated in the reaction buffer, 2 – ladder, 3 – limited hydrolysis of $tRNA^{Phe}$ with RNase T1; 5, 7, 9,11 – hydrolysis of complex $tRNA^{Phe}$ and peptide; 4, 6, 8,10 – hydrolysis of free $tRNA^{Phe}$. The numbers on the left side correspond to nucleotides in primary sequence of $tRNA^{Phe}$.

## REFERENCES


1. Westheimer F.H., *Nature*, **319**, 534 (1986).
2. Zaug, A.J. and Cech T.R., *Science*, **231**, 470 (1986).
3. Gilbert W., *Nature*, **319**, 618 (1986).
4. Pace N.R. and Marsh T.L., *Origins Life*, **16**, 97 (1985).
5. Darnell J.E. and Doolittle W.F., *Proc. Natl. Acad. Sci. USA*, **83**,1271 (1986).
6. Cedergren R. and Grosjean H., *BioSystems*, **20**, 175 (1987).
7. The RNA World, Second Edition (1999). Cold Spring Harbor, Cold Spring Harbor NY.
8. Jeffares D.C., Poole A.M. and Penny D., *J. Mol. Evol*., **46**, 18 (1998).
9. Cedergren R., Larue B., Sankoff D. and Grosjean H., *Crit. Rev. Biochem*., **11**, 35 (1981).
10. Wyszko E., Fuerste J.P., Barciszewska M., Szymanski M., Admiak R., Erdmann V.A. and Barciszewski J., *Japanese Biochem. Soc*., **126**, 326 (1999).
11. Barciszewska M., Dirhiemer G. and Keith G., *Biochem. Biophys. Res. Commun*., **114**, 1161 (1993).
12. Pieler T. and Erdmann V.A., *Proc. Natl. Acad. Sci. USA*, **79**, 4599 (1982).
13. Wyszko E. and Barciszewska M., *Eur. J. Biochem*., **249**, 107 (1997).
14. Digweed M., Pieler T. and Erdmann V.A., *Advanced Methods in Protein Microsequence Analysis*, Wittmann-Liebold B. *et al*. (eds). Springer-Verlag, Heidelberg (1986).
15. Audhya T.K., Schlesinger D.H. and Goldstein G., *Biochem*., **20**, 6195 (1981).
16. Harris C.A., Andryuk P.J., Cline S., Chan H.K., Naturajan A., Siekierka J.J. and Goldstein G., *Proc. Natl. Acad. Sci. USA*, **91**, 6283 (1994).
17. Siemion I.Z., Strug I. and Szewczuk Z., *Current Topics Pept. Protein Res*., **2**, 1 (1997).
18. Giel-Pietraszuk M., Barciszewska M.Z., Mucha P., Rekowski P., Kupryszewski G. and Barciszewski J., *Acta Biochim. Polon*., **44**, 591 (1997).
19. Siatecka M., Rozek M., Barciszewski J. and Mirande M., *Eur. J. Biochem*., **256**, 80 (1998).
20. Naryashkin N.A., Gait M.J. and Ivanovskaya M.G., *Biochem. (Moscow)*, **63**, 489 (1998).
21. Wang Jihang., Stein S., Leibowitz M.J. and Rabson A.B., *C.A*., **129**, 326081 (1998).