Enhancement of protein expression in insect cells by a lobster tropomyosin cDNA leader sequence

Ken-Ichi Sano, Kayo Maeda, Manami Oki, Yuichiro Maéda*

RIKEN Harima Institute at SPring-8, Laboratory for Structural Biochemistry, Mikazuki, Sayo, Hyogo 679-5148, Japan

Received 3 October 2002; revised 25 October 2002; accepted 26 October 2002

First published online 7 November 2002

Edited by Horst Feldmann

Abstract **We describe a *cis* element that dramatically increases the expression levels of exogenous genes in baculovirus-infected insect cells. This 21 bp sequence element is derived from a 5′ untranslated leader sequence of a lobster tropomyosin cDNA (L21). By using a transfer vector carrying L21, the expression levels of tropomyosin and luciferase were 20- and seven-fold higher with L21 than without L21, respectively. L21 has both the Kozak sequence and the A-rich sequence found in the polyhedrin leader sequence. We assume that both sequence elements are essential for the enhancement of protein expression in the baculovirus-based expression system.**

Key words: Baculovirus; Sf9; Recombinant protein expression; Leader sequence; Tropomyosin

1. Introduction

The Sf9-baculovirus expression system has numerous advantages for preparing biologically important eukaryotic proteins [1,2]. In most transfer vectors used for expression in Sf9 cells, the polyhedrin promoter of *Autographa californica* nuclear polyhedrosis virus (AcNPV) is employed. This is because polyhedrin is a highly expressed viral protein, which can compose 50–75% of the total protein in infected cells [1]. Even with the polyhedrin promoter, the expression level of recombinant proteins is usually substantially lower than that of polyhedrin. In order to increase recombinant protein expression levels in Sf9 cells, some attempts have focused on changing the length of the 5′ untranslated region (UTR). Deletions within the original 5′ UTR of the polyhedrin promoter adversely affected the expression level of exogenous protein, which might be due to the reduced amount of mRNA [3,4]. In contrast, an extension of the leader sequence, by the insertion of a foreign gene downstream of the destroyed polyhedrin ATG start codon, often resulted in high levels of protein expression without affecting the mRNA expression level [3,4]. Page and co-workers [5] enhanced the expression of several recombinant proteins by inserting a 24 bp stretch of the 3′ end of the *Escherichia coli trpE* gene immediately upstream of the initiation codon of a foreign gene. In contrast, the insertion of both vertebrate consensus and *trpE* sequences immediately upstream of the translational initiation site had little effect on the expression of fibroblast growth factor [6].

Here we examined the effect of the 21 base leader sequence (L21) on the expression levels of three different foreign proteins. The expression levels of tropomyosin and luciferase were 20- and seven-fold higher with L21 than without L21, respectively. Our results demonstrate that L21 can dramatically increase the expression level, and therefore the yield, of non-fused recombinant proteins in the baculovirus-based expression system.

2. Materials and methods

2.1. Construction of transfer vectors

The construction of the transfer vectors pAcC4-RαTm for the expression of rabbit skeletal muscle α-tropomyosin (RαTm) and pVL1392-*sTm1* for the expression of lobster slow muscle-specific tropomyosin type 1 isoform (*sTm1*) has been previously described [7,8]. The transfer vector for the expression of RαTm with L21 (pVL-L21-RαTm) was constructed by replacement of the *sTm1* gene with the RαTm gene in pVL1392-*sTm1* (Fig. 1A).

The transfer vectors pAcC4-luc, pVL-L21-luc and pVL-L21-α-actin were constructed by PCR. For pAcC4-luc and pVL-L21-luc, the sequences of the PCR primers were: 5′-CTGTTGGTAAAGCCACCATGG-3′ and 5′-GGATCCTTACACGGCGATCTTTCCGC-3′ (restriction sites are underlined). The pGL3-Basic Vector (Promega) was used as the template. The PCR product was inserted into *Nco*I and *Bam*HI sites of pAcC4 and pVL-L21, respectively. For pVL-L21-α-actin, the primer sequences were: 5′-TCTAGA*AACTCCTAAAAAACCGCCACC*ATGTGCGACGAAGACGAGACC-3′ (L21 sequence in italic) and 5′-AAAGGATCCGCTGGAGGTGGAGTGTGTTAGAAG-3′. The human α-actin cDNA was used as the template. The PCR product was cloned into *Xba*I and *Bam*HI sites of pVL1392.

2.2. Cell culture and manipulations

Spodoptera frugiperda cells, Sf9, were purchased from Invitrogen. The Sf9 cells were cultured at 27°C, in TNM-FH medium containing 10% fetal calf serum (Sigma) for transfection and expression of α-actin, and in modified Sf-900II medium, which contains 90% Sf-900II SFM (Gibco BRL), 9% Grace's insect medium supplemented (Gibco BRL), and 1% fetal calf serum (Sigma), for recombinant virus amplification and recombinant protein expression. In suspension culture, Pluronic F68 (Gibco BRL) was added to final concentrations of 0.1% and 0.2%, for the TNM-FH and modified Sf-900II media, respectively.

The recombinant baculovirus was generated by co-transfection of Sf9 cells with a transfer vector and BaculoGold DNA (Pharmingen). Each recombinant virus was screened by a plaque assay. Sf9 cells and recombinant virus manipulations were performed basically as described by Summers and Smith [2].

*Corresponding author. Fax: (81)-791-58 2836.
E-mail address: ymaeda@spring8.or.jp (Y. Maéda).

Abbreviations: L21, 21 bp 5′ untranslated leader sequence of a lobster tropomyosin cDNA; UTR, untranslated region; AcNPV, *Autographa californica* nuclear polyhedrosis virus; RαTm, rabbit skeletal muscle α-tropomyosin; *sTm1*, lobster slow muscle-specific tropomyosin type 1 isoform

0014-5793/02/$22.00
PII: S0014-5793(02)03659-1

Sf9 cell culture for expression of the large-scale tropomyosins was carried out as described [8,9]. For the expression of luciferase and α-actin, 7.5×10^6 Sf9 cells were seeded in a 10 cm plate with 10 ml of medium, infected with recombinant virus at a multiplicity of infection of 2.0. After 48 h post-infection, the infected cells were dislodged from the plate by pipetting and were collected by centrifugation at $1200\times g$ for 10 min.

2.3. Estimation of expression level

To estimate the expression level, tropomyosin was partially purified from a 3 l culture and was subjected to colorimetry on an SDS–PAGE gel. The tropomyosin from 3 l of Sf9 cell culture was purified as previously described up to the column chromatography step [8–10]. The total protein of this preparation was quantified by the microbiuret method [11], and the relative content of tropomyosin in the total protein was determined by measuring the optical density of a 15% SDS–PAGE gel stained with Coomassie brilliant blue by using the Gel-Doc 2000 image analysis system (Bio-Rad).

The luciferase expression level was measured as follows. The infected and collected cells were lysed with 500 μl of lysis buffer, containing 20 mM Tris–HCl pH 8.0, and 5 mM 2-mercaptoethanol. The cell lysate was centrifuged and the supernatant was diluted 1000-fold with the lysis buffer. The diluted cell lysate (20 μl) was mixed with 100 μl of Luciferase Assay Reagent (Promega), and the luminescence was measured by a single-tube luminometer (BioOrbit), which was programmed to perform an 8 s delay followed by a 10 s read. The readout value was defined as RLU (relative light units).

To detect α-actin expression, collected cells were lysed with 500 μl of 20 mM Tris–HCl pH 8.0, 1 mM EDTA, and 5 mM dithiothreitol. After centrifugation, 10 μl aliquots of supernatant were fractionated on two 12% SDS–PAGE gels. Western blotting was performed with a mini transblot system (Bio-Rad). To detect only the exogenous actin, not the actin species endogenous to the Sf9 cell, the monoclonal anti-α-sarcomeric actin IgM (Sigma) was used.

2.4. Compilation of the S. frugiperda consensus leader sequence

Sequences of *S. frugiperda* genes were retrieved from the GenBank database (accessed 11 June 2001). The 10 base long leader sequences immediately upstream of the initiation codon and the four base long coding sequences including the ATG from 28 genes were aligned. The assignment of the consensus sequence was calculated as follows: the first assignment was made on the basis of Cavener's 75/50 rule (as indicated in capital letters in Table 1) [12], and the second assignment was subjected by the 70/40 rule which was a modification of Cavener's rule (the results are indicated in small letters in Table 1).

3. Results and discussion

The present work is based on our experiences with Sf9 cells using the baculovirus system, in which lobster tropomyosin variants, typified by *sTm1*, were expressed at extremely high levels [8,10]. These levels were 20–30-fold higher than the expression level of RαTm [7,9]. The comparison of cDNA sequences indicated that this enhanced expression cannot be accounted for by the distinct codon usage between the lobster and rabbit tropomyosin genes [7,13].

A possible explanation is the difference in the transfer vectors used. The differences here are two-fold (Fig. 1B). First, the translational initiation position used was different; the original polyhedrin initiation codon is destroyed in the transfer vector for lobster tropomyosin (pVL1392-*sTm1*), while in the transfer vector for RαTm (pAcC4-RαTm), the original polyhedrin initiation codon is intact and used for protein expression. Second, pVL1392-*sTm1*, but not pAcC4-RαTm,

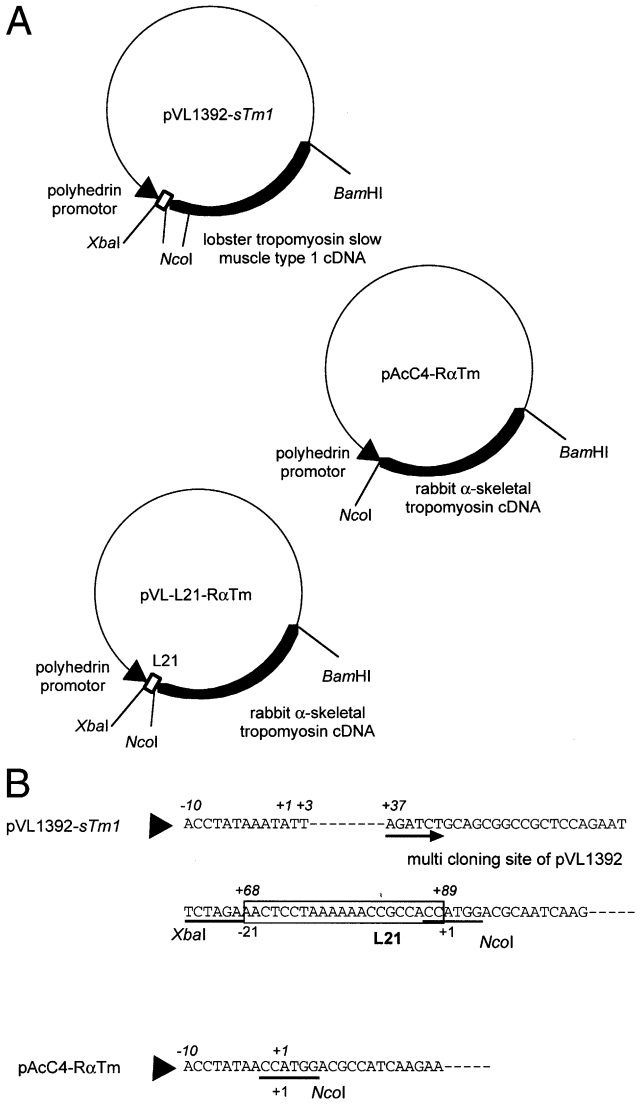

Fig. 1. A: The transfer vector constructs used for the expression of various tropomyosin species. B: The nucleotide sequences of the transfer vectors near the translational start point. Numbers in italics above the sequences are nucleotide numbers counted from the first nucleotide A (numbered as *+1*) within the original polyhedrin translational initiation codon ATG. Numbers below the sequences are nucleotide numbers counted from the first nucleotide A (numbered as +1) within the inserted tropomyosin translational initiation codon ATG. The boxed sequence indicates L21, the leader sequence of lobster tropomyosin cDNA.

Table 1
Frequency of A, C, G, and T around the translational initiation site in *S. frugiperda*

	−10	−9	−8	−7	−6	−5	−4	−3	−2	−1	+1	+2	+3	+4
A	***0.35***	0.30	0.21	0.14	0.21	***0.36***	0.21	**0.68**	**0.43**	0.18	**1**	0	0	**0.29**
C	0.12	0.30	***0.43***	**0.36**	0.21	***0.36***	**0.50**	0.04	**0.36**	**0.50**	0	0	0	0.14
G	***0.38***	0.04	0.14	0.07	0.29	0	0.07	0.29	0.11	**0.29**	0	0	**1**	**0.57**
T	0.15	0.37	0.21	**0.43**	0.29	0.29	0.21	0	0.11	0.04	0	**1**	0	0
Consensus	a/g	N	c	C/T	N	a/c	C	A	A/C	C/G	A	T	G	G/A

contains the extra DNA stretch of 21 bp, which was brought along accidentally from the lobster tropomyosin cDNA. Altogether, the leader sequence of the putative mRNA of lobster tropomyosin in Sf9 cells must be longer than its counterpart of the rabbit mRNA by 89 bases, i.e. 68 bases from pVL1392 and 21 bases from the lobster *sTm1* cDNA. To test whether the longer leader sequence associated with the transfer vector for *sTm1* also increases the expression level of RαTm, we constructed pVL-L21-RαTm, in which the RαTm gene was inserted downstream of this leader sequence (Fig. 1A), and compared the expression level (Fig. 2A). The expression levels of RαTm are about 20-fold higher with pVL-L21 (120–200 mg/l of infected cells) than with pAcC4 (6–20 mg/l of infected cells). After this observation, we were also able to express RαTm internal deletion mutants [14] and other tropomyosin isoforms (unpublished result) at high levels using the pVL-L21 transfer vector, indicating that the longer leader sequence effectively increased the expression level of tropomyosin isoforms.

To further test if the enhanced expression could be generally applicable for foreign proteins, we expressed firefly, *Photinus pyralis*, luciferase in the presence of L21. We constructed pAcC4-luc and pVL-L21-luc, and compared the expression levels by SDS–PAGE (Fig. 2B) and by measuring luciferin–

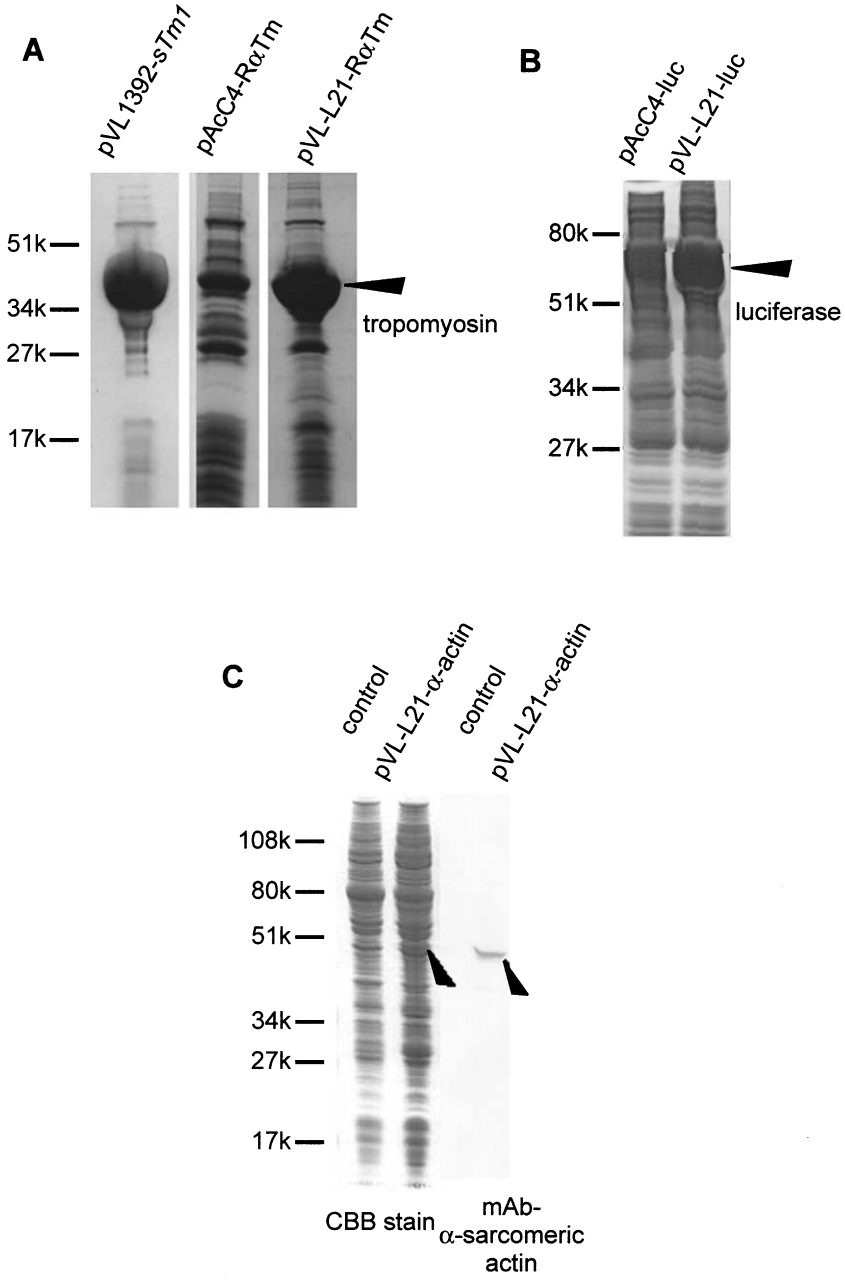

Fig. 2. Analysis of the expression levels of various tropomyosin species (A), luciferase (B) and α-actin (C) by SDS–PAGE. In A, on each lane, partially purified protein preparations were loaded in equivalent amounts on the basis of culture volume, while in B and C, Sf9 cells in the same culture volume were loaded. The gels were stained by Coomassie brilliant blue. Numbers on the left indicate molecular weight markers. The arrowhead indicates each expressed protein. The two lanes on the right in C were Western blotted with an anti-α-sarcomeric actin antibody (Sigma).

luciferase luminescence. The expression levels of firefly luciferase are about seven-fold higher with pVL-L21 ($3.8 \pm 1.2 \times 10^7$ RLU/ml of infected cells) than with pAcC4 ($5.4 \pm 2.1 \times 10^6$ RLU/ml of infected cells). Consistent with our hypothesis the long leader sequence including L21 enhanced the luciferase expression level. For the enhanced expression level, the sequence of L21, rather than the length of the inserted sequence, must be crucial. This is because the insertion of 21 bases from a *Bacillus subtilis* sequence, instead of L21, resulted in an even lower luciferase expression level than that with pAcC4-luc (data not shown).

Comparison of the L21 sequence with known leader sequences (Fig. 3) indicated that L21 completely agrees with the Kozak sequence in positions −8 to −1, which is distinct from the invertebrate and AcNPV consensus sequences. In mammalian cells, the sequence at positions −5 to −3, and especially −3 to −1, is reportedly the most important for the enhancement of translational efficiency [15,16]. In the baculovirus-Sf9 cell, although the importance of the leader sequence immediately upstream of the initiation codon was proposed [3–5], the Kozak sequence did not increase the expression level of fibroblast growth factor [6]. On the other hand, several reports have indicated that sequences further upstream influence protein expression. An elevated level of translation was observed with mRNA at positions −9 to −7 [16], and moreover with mRNA having tandem repeats of that sequence motif in positions −12 to −7 [17–19]. L21 also contains the A-rich sequence in positions −15 to −9 (TAAAAAA), which is almost identical to a part of the A-rich sequence of the original polyhedrin leader sequence in positions −17 to −10 (TAAAAAAA) [20]. It is well known that the A-rich sequence increased the expression level in Sf9 cells [3]. Therefore, we assume that, in the case of baculovirus-infected insect cells, not only the Kozak sequence but also the A-rich sequence upstream of the Kozak sequence is essential for enhanced recombinant protein expression.

Encouraged by the ability of L21 to increase the expression level of foreign proteins, we employed L21 for the expression of a protein useful in our research, which had not been successfully expressed for unknown reasons. We were able to express human skeletal muscle α-actin successfully by the use of L21 (Fig. 2C). On Coomassie brilliant blue-stained SDS–PAGE gels, the soluble fraction of the cell lysate shows an extra band that migrates with the same mobility as actin and reacts with the anti-α-sarcomeric actin monoclonal antibody. This result indicates that L21 improves the expression level of recombinant proteins so that otherwise undetectable expressed proteins could be detected.

Our results demonstrate that inclusion of an L21 cassette in the 5′ UTR can be a potent tool for enhancing protein expression. Further studies are required to elucidate the molecular mechanism of its action.

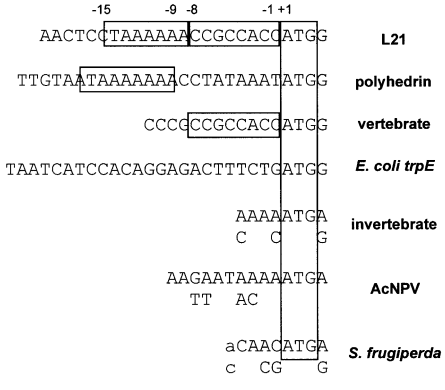

Fig. 3. Sequence comparison among leader sequences. The leader sequences of polyhedrin and *E. coli trpE* are from Van Iddekinge et al. (modified to represent the correct sequence) [20], and Peakman et al. [5], respectively. The consensus sequences among the vertebrate, invertebrate, and AcNPV leader sequences are from Kozak [21], Cavener and Ray [12], and Ranjan and Hasnain [22], respectively. The consensus sequence among the *S. frugiperda* leader sequences is from the present work (see Table 1). The open boxed indicate the sequences which are shared with L21, the leader sequence of lobster tropomyosin cDNA.

Acknowledgements: This work was partly supported by the Special Coordination Funds for Promoting Science and Technology from the Ministry of Education, Culture, Sports, Science and Technology, the Japanese Government to Y.M., President's Special Research Grant, RIKEN, and Special Postdoctoral Researchers Program, RIKEN (K.-I.S.). We thank Dr. J.L. Mandel (CNRS, Strasbourg, France) for providing human α-actin cDNA, Dr. S. Ichiyama (Gakusyuin University) for technical advice regarding the measurement of luciferase activity, and Dr. Y. Takaoka (Iwate Medical University) for helpful discussions.

References

[1] Luckow, V.A. (1990) in: Recombinant DNA Technology and Applications (Ho, C., Prokop, A. and Bajpai, R., Eds.), McGraw-Hill, New York, pp. 1–25, New York.
[2] Summers, M.D. and Smith, G.E. (1987) Texas Agric. Exp. Stn. Bull. 1555.
[3] Luckow, V.A. and Summers, M.D. (1988) Virology 167, 56–71.
[4] Luckow, V.A. and Summers, M.D. (1989) Virology 170, 31–39.
[5] Peakman, T.C., Charles, I.G., Sydenham, M.A., Gewert, D.R., Page, M.J. and Makoff, A.J. (1992) Nucleic Acids Res. 20, 6111–6112.
[6] Hills, D. and Crane-Robinson, C. (1995) Biochim. Biophys. Acta 1260, 14–20.
[7] Kluwe, L., Maeda, K., Miegel, A., Fujita-Becker, S., Maéda, Y., Talbo, G., Houthaeve, T. and Kellner, R. (1995) J. Muscle Res. Cell Motil. 16, 103–110.
[8] Sano, K.I., Maeda, K., Taniguchi, H. and Maéda, Y. (2000) Eur. J. Biochem. 267, 4870–4877.
[9] Sano, K.I., Maeda, K., Oda, T. and Maéda, Y. (2000) J. Biochem. (Tokyo) 127, 1095–1102.
[10] Miegel, A., Sano, K.I., Yamamoto, K., Maeda, K., Maéda, Y., Taniguchi, H., Yao, M. and Wakatsuki, S. (1996) FEBS Lett. 394, 201–205.
[11] Itzhaki, R.F. and Gill, D.M. (1964) Anal. Biochem. 9, 401–410.
[12] Cavener, D.R. and Ray, S.C. (1991) Nucleic Acids Res. 19, 3185–3192.
[13] Mykles, D.L., Cotton, J.L., Taniguchi, H., Sano, K.I. and Maéda, Y. (1998) J. Muscle Res. Cell Motil. 19, 105–115.
[14] Hai, H., Sano, K.I., Maeda, K., Maéda, Y. and Miki, M. (2002) J. Biochem. (Tokyo) 131, 407–418.
[15] Kozak, M. (1986) Cell 44, 283–292.
[16] Kozak, M. (1987) J. Mol. Biol. 196, 947–950.
[17] Lawrence, C.B. and Jackson, K.J. (1982) J. Mol. Biol. 162, 317–334.
[18] van Ormondt, H. and Galibert, F. (1984) Curr. Top. Microbiol. Immunol. 110, 73–142.
[19] Berkner, K.L. and Sharp, P.A. (1984) Nucleic Acids Res. 12, 1925–1941.
[20] Van Iddekinge, B.J.L.H., Smith, G.E. and Summers, M.D. (1983) Virology 131, 561–565.
[21] Kozak, M. (1987) Nucleic Acids Res. 15, 8125–8148.
[22] Ranjan, A. and Hasnain, S.E. (1995) Virus Genes 9, 149–153.