Effect of genomic and subgenomic leader sequences of potato leafroll virus on gene expression

Marek Juszczuk, Elżbieta Pączkowska, Ewa Sadowy, Włodzimierz Zagórski, Danuta M. Hulanicka*

Instytut Biochemii i Biofizyki, PAN, Ul, Pawińskiego 5A, 02-106 Warsaw, Poland

Received 12 July 2000; revised 2 October 2000; accepted 2 October 2000

Edited by Takashi Gojobori

Abstract The effect of the genomic and subgenomic leader sequence of potato leafroll polerovirus on the efficiency of translation of the downstream located genes has been studied. The results obtained in vitro and in vivo indicate that neither leader sequence functions as translational enhancer, a generally important feature of leader sequences. Deletion analyses demonstrated that both leader sequences not only decrease translation of the downstream located genes but also alter the ratio of the synthesized proteins. A correlation between the in vitro and in vivo results can be established in the case of the subgenomic leader sequence.

Key words: Leader sequence; Potato leafroll polerovirus; Translation

1. Introduction

During the last decade, particular attention has been directed towards structural elements influencing gene expression. Leader sequences (LSs) were found to have multiple effects on the translation of the downstream located genes and on mRNA stability [1–5]. The length of plant viral LSs is generally between 12 and 250 nucleotides (nt). In caulimoviruses and picornaviruses, this length can reach 600–1200 nt. 'Typical' plant viral LSs appear to lack extensive secondary and/or tertiary structure, or any particular nucleotide sequence. Generally, the yield of proteins synthesized from viral RNAs bearing their translational enhancers is several fold to a few dozen fold higher than from RNAs with unrelated random LSs [6]. However, in spite of considerable work, the mechanism of LS action is poorly understood.

Potato leafroll polerovirus (PLRV) employs a variety of mechanisms to express its monopartite single-stranded RNA genome of positive polarity [7], among which the production of a subgenomic (sg) RNA that encodes the 3′ co-terminal open reading frames (ORFs). The genome of PLRV contains six ORFs organized in two gene clusters. The first three ORFs (ORF0, ORF1 and ORF2) are expressed from the genomic (g) RNA, the three others (ORF3, ORF4 and ORF5) from a sgRNA. The possible role of the gLS and sgLS on translation has not been examined in detail. Preliminary results [8] obtained in vitro suggested that the presence of the sgLS not only decreases translation of downstream located genes, but also affects the ratio of the proteins synthesized. This interesting observation prompted us to further investigate the role of gLS and sgLS on translation.

In this study in vitro transcription and translation, and transient expression in transfected protoplasts have been used to evaluate the role of the PLRV gLS and sgLS on the expression of the viral genes and on the reporter gene for β-glucuronidase (GUS).

2. Materials and methods

2.1. PLRV

The Polish isolate of PLRV-P was described previously [9]. Its nucleotide sequence is available in EMBL under accession number X74789.

2.2. Construction of plasmids

Plasmids for in vitro studies were obtained as follows. Plasmid pFC [10] carrying the full-length cDNA genome of PLRV was used as template in PCR reactions. The two pairs of primers (Table 1) PH1+PBAM and PORF0+PBAM were used to amplify the appropriate cDNAs. The amplified fragments were cleaved with *Cla*I and *Bam*HI and ligated into pUC118 cleaved with *Acc*I and *Bam*HI, yielding pGLS1B and pGLS0B, respectively. Transcript (t) tGLS1B contains the full-length gLS of 69 nt, whereas tGLS0B contains only the 12 nt 5′ proximal to ORF0 (Fig. 1).

A set of plasmids carrying different lengths of the sgLS was constructed (Fig. 1). Plasmid pVAP3 which contains the PLRV *Eco*RI/*Eco*RI cDNA fragment, nt 3160–5176 [9], served as template to amplify the cDNA by PCR. The three pairs of primers P1+P4, P2+P4, and P3+P4 (Table 1) were used to amplify the appropriate cDNAs. The amplified fragments were cleaved with *Kpn*I and *Sph*I and ligated into similarly cleaved pUC118. The plasmids obtained are denoted pWT, pHWT and pAWT, respectively. The three transcripts are truncated at their 3′ end (position 4457) and upon translation should yield C-terminally truncated ORF5 proteins (Fig. 1). Transcript tWT contains the full-length sgLS of 212 nt; transcript tHWT is truncated from the 5′ end and has a 112 nt long LS; tAWT is further truncated and has only a 14 nt long LS.

Plasmids for in vivo studies were obtained as follows. Plasmid pFF19G [11] carrying the GUS gene under the control of the cauliflower mosaic virus (CaMV) 35S promoter was used to clone the cDNA of the gLS or sgLS for in vivo studies (Fig. 2A). pFC carrying the full-length PLRV genome served as template to amplify the cDNA of the gLS and the sgLS by PCR. The following primers (Table 1) were used for the PCR reactions: PK1+P35UP for the gLS (69 nt), and P35SG+PALS for the sgLS (212 nt). It could be anticipated that the presence of the AUG start codon of either ORF0 or ORF3 in the amplified fragments and its introduction upstream of the start codon of the GUS gene might affect translation and therefore affect the results. Consequently, the initiator AUG codons of ORF0 and ORF3 were removed by cleavage of the amplified LS fragments with *Sph*I (which recognizes GCATGC; initiation codon underlined) followed by treatment with mung bean nuclease and cleavage with *Kpn*I. pFF19G was cleaved with *Xba*I, filled-in with the Klenow fragment of DNA polymerase I, and subsequently cleaved with *Kpn*I. The

*Corresponding author. Fax: (48)-39-12 16 23.
E-mail: dankahul@ibb.waw.pl

PII: S0014-5793(00)02122-0

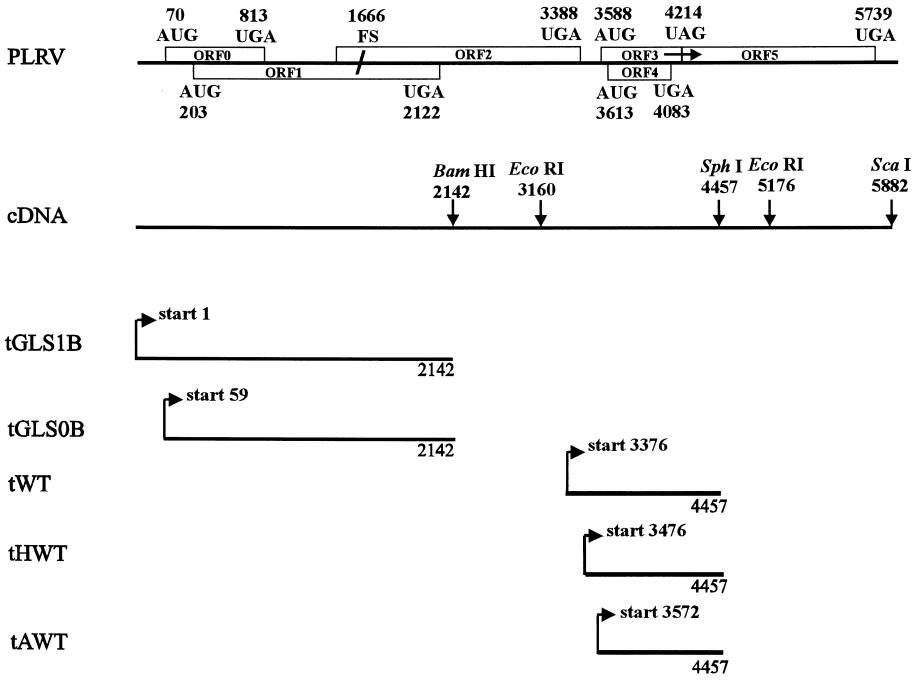

Fig. 1. Organization of the PLRV genome. Boxes represent ORFs; positions of start and stop codons for each ORF, of frameshift site (*/*) and readthrough (→) are indicated. The line below indicates the cDNA of PLRV with restriction sites used for construction of plasmids or linearization of templates. *Bam*HI in nt position 2142 was introduced with primer PBAM by nucleotide substitution to linearize pGLS1B and pGLS0B at the end of ORF1. Diagrammatic representation of transcripts used to study the effect of gLS and sgLS are presented below. Broken arrows, initiation of transcription.

cDNAs of both gLS and sgLS were inserted into cleaved pFF19G, yielding pGUSgLS and pGUSsgLS, respectively.

2.3. In vitro transcription, translation and analysis of products

Transcription and translation were performed as described previously [12]. The translation products were analyzed by 0.1% SDS–12.5% polyacrylamide gels [13]. After electrophoresis, the gels were dried and loaded onto a phosphorimager camera or exposed to an X-ray (β-max) film (Amersham). Quantification was performed using Image Quant.

2.4. Protoplast isolation, polyethylene glycol-mediated transformation and assay of GUS activity and determination of protein concentration

For the isolation and transfection of cowpea protoplasts, the protocols of Hibi et al. [14] and van Beek et al. [15] were adopted. Assay of GUS activity was as described by Jefferson et al. [16] and protein concentration as described by Bradford [17].

2.5. Software

The following computer programs were used to analyze the nucleotide and amino acids sequences, and the autoradiograms: IMAGE QUANT from Molecular Dynamics and GCG-Genetic Computer Group, University of Wisconsin, Madison, WI.

3. Results and discussion

3.1. Effect of gLS on the expression of ORF0 and ORF1 in vitro

To investigate the influence of the gLS on the expression of ORF0 and ORF1, pGLS1B and pGLS0B were linearized by *Bam*HI and used for the synthesis of uncapped transcripts. tGLS1B contains the full-length 69 nt long gLS preceding ORF0 (Fig. 1). tGLS0B starts at nt 59 and contains only the 12 nt of the gLS preceding ORF0; this is reported to be the minimal indispensable length of a LS that allows ribosome entry and translation initiation [18]. Both RNAs are termi-

Table 1
Primers used for amplification of viral cDNA fragments

Primers	Coordinate numbers	Restriction site	Sequence
PH1	1–20 (F)	*Cla*I	GGG*ATCGAT*TAATACGACTCACTATAG**ACAAAAGAATACCAGGGGAAA**
PORF0	59–76 (F)	*Cla*I	GGG*ATCGAT*TAATACGACTCACTATAG**CGTTATAGCATATGATTG**
PBAM	2130–2149 (R)	*Bam*HI	**GTT***GGATCC***ACCTGGCAGCC**
P1	3376–3395 (F)	*Kpn*I	GGG*GTACC*TAATACGACTCACTATAG**ACAAAAGAACACTGAAGGAG**
P2	3476–3498 (F)	*Kpn*I	GGG*GTACC*TAATACGACTCACTATAG**TATAAATTCTTAGCGGGATTTG**
P3	3572–3593 (F)	*Kpn*I	GGG*GTACC*TAATACGACTCACTATAG**GTGCGATCAATTGTTAATGAG**
P4	4441–4457 (R)	*Sph*I	**GACGGCTT***GCATGC***TCG**
PK1	1–20 (F)	*Kpn*I	GGG*GTACC***ACAAAAGAATACCAGGGGAA**
P35UP	62–79 (R)	*Sph*I	GGG*CATGC***GCTATAACGTTTGCAAT**
P35SG	3376–3395 (R)	*Kpn*I	GGG*GTACC***ACAAAAGAACACTGAAGG**
PALS	3572–3586 (R)	*Sph*I	GGG*GCATGC***ACAATTGATCGCACG**

Sequences for the creation of restriction sites are in italics, PLRV sequences are in bold and the T7 promoter is underlined. Other letters are sequences added for better cleavage by restriction enzymes. F: forward primer; R: reverse primer.

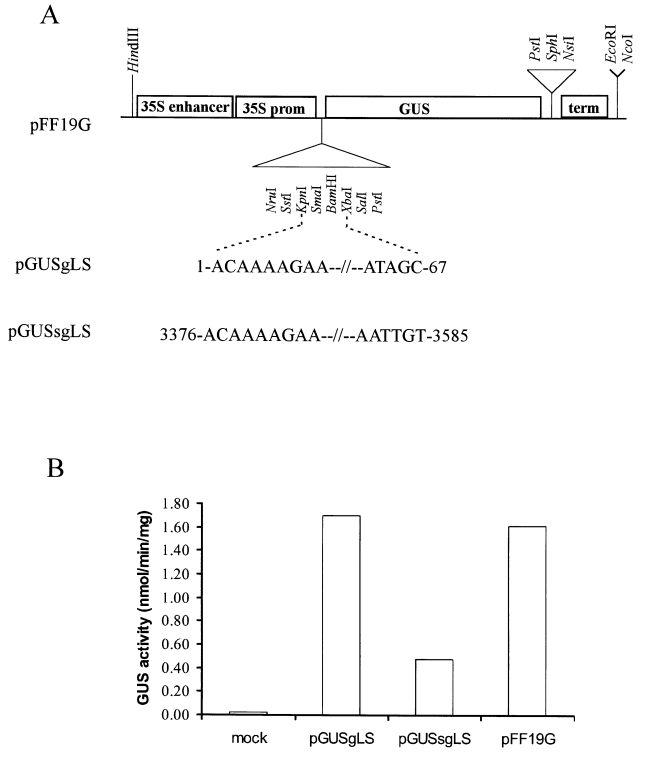

Fig. 2. Effect of the LSs on GUS expression. (A) Diagrammatic representation of pFF19G with restriction sites; CaMV 35S enhancer, promoter (prom), and terminator (term), and GUS gene are in boxes. The LSs were inserted into pFF19G upstream of the GUS gene. The sequence at the beginning and end of the gLS and sgLS are depicted. (B) GUS expression in mock-transfected protoplast and in protoplasts electroporated with pGUSgLS, pGUSsgLS or pFF19G. GUS activity is expressed in nmol of methylumbelliferone produced per min per mg of cellular proteins. The data represent averages of three separate experiments.

nated at nt 2142 downstream of the UGA (nt 2122–2124) of ORF1.

The translation products using either transcript are presented in Fig. 3A. The 28 kDa (ORF0) and 70 kDa (ORF1) proteins synthesized by either tGLS1B or tGLS0B (lanes 3 and 4) present a similar protein pattern, and a pattern similar to the one produced by translation of the full-length PLRV RNA obtained from transcription of pFC (lane 2). Densitometric quantification of the bands seen on the gel and calculation of these values relative to the number of methionine residues contained in each protein (four methionine residues in ORF0 and 13 in ORF1) revealed an inhibitory effect of the LS on the expression of ORF0 and ORF1 (Fig. 3B), since removing most of the gLS resulted in a 30% higher level of translation. Moreover, the lack of gLS also changed the ratio ORF0/ORF1 which was 4/1 for tGLS1B and 2.7/1 in the case of tGLS0B.

3.2. *Effect of sgLS on the expression of ORF3 and ORF4 in vitro*

The 212 nt long LS of the sgRNA begins at position 3376 of the PLRV genome [19]. To determine its role on translation, pWT, pHWT and pAWT were linearized with *Hin*dIII present in the multiple cloning site (MCS) of pUC118 and transcribed. tWT contains the full-length sgLS whereas in transcripts tHWT and tAWT, the sgLS is 112 and 16 nt long, respectively. All three transcripts are 3′ terminally truncated versions of the sgRNA (Fig. 1). Translation of the uncapped transcripts and analysis of their products were performed as in the experiments with the gLS.

Fig. 4 presents the translation products of tWT, tHWT and tAWT. Densitometric measurements indicate a strong dependence of the level of expression of the coat protein (CP; ORF3; 23kDa) and of the 17 kDa protein (ORF4) on the length of the sgLS. Significantly, lack of the first 100 nt of the sgLS led to a fivefold increase in translation efficiency. Removal of most of the sgLS does not further increase translation efficiency. It is interesting, that as in the case of gLS the lack of sgLS changes the ratio CP/17 kDa of the proteins synthesized (three and five methionine residues in the CP and 17 kDa, respectively). In the case of tWT this ratio is 1/1, with tHWT it changes to 1.5/1, and it is 1/1.5 in the case of tAWT. The above results were obtained with 3′ terminally truncated transcripts of sgRNA. Since the 3′ untranslated region (UTR) might affect translation, experiments with the same 5′ UTR truncated transcripts but containing the full-length 3′ UTR were performed [8]. Similar results were obtained (not shown), indicating that the sgLS does not interact with the 3′ UTR to modify translation in vitro.

3.3. *Influence of LSs on the expression of the GUS gene in vivo*

Fusions of LSs to a reporter gene are often used to study their effect on translation efficiency. To investigate the influence of the gLS and sgLS on the expression of the reporter

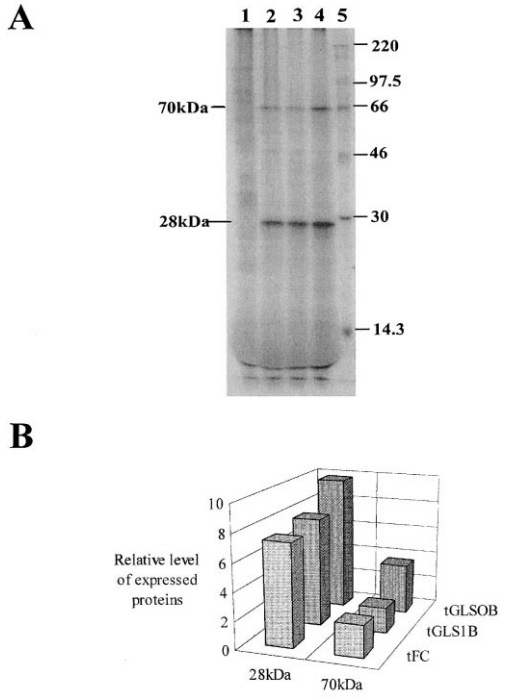

Fig. 3. Effect of gLS on the expression of ORF0 (28 kDa) and ORF1 (70 kDa) in vitro. (A) Analysis of translation products. Identical amounts of transcripts (200 ng) were translated and identical aliquots of the translation mixture (2 μl) were loaded onto each lane of the gel. Lane 1, translational products of TMV RNA. Lanes 2–4, translation products of tFC, tGLS1B and tGLS0B, respectively. Lane 5, labeled protein markers (220 to 14.3 kDa). (B) Relative translation level of tFC, tGLS1B and tGLS0B. The dried gel was loaded into a phosphorimager, scanned and the translation products were quantified.

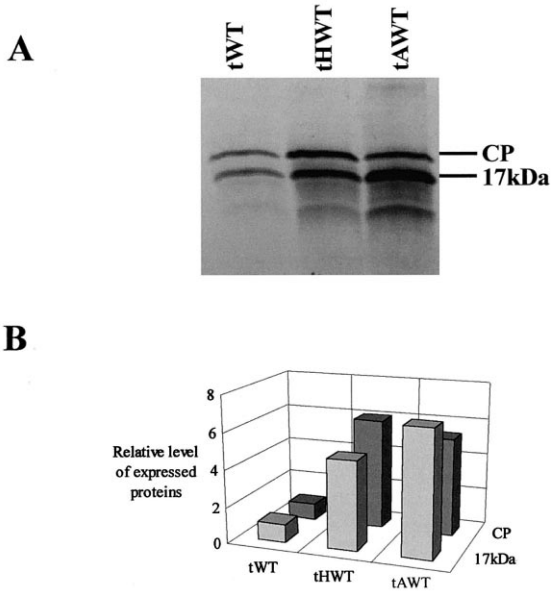

Fig. 4. Effect of the sgLS on the expression ORF3 and ORF4. (A) Translation products obtained with tWT (lane 1), tHWT (lane 2) and tAWT (lane 3). (B) Relative level of expression of the CP and 17 kDa protein as derived from A.

gene GUS, pGUSgLS and pGUSsgLS were constructed. They contain sequences corresponding to the full-length gLS and sgLS of the PLRV genome respectively, flanked by the 35S promoter and terminator. The plasmid DNAs were transfected into cowpea protoplasts, samples collected after 24 h, and the total protein in each protoplast sample as well as GUS activity were measured. The assay of GUS activity in mock- transfected protoplast samples was used as negative control. The level of expression of GUS in protoplasts provided with plasmid pFF19G served as the baseline control. The relative level of GUS activity is presented in Fig. 2B.

The effects of the gLS and sgLS on the expression of the GUS gene are significantly different. A comparison of the GUS activity in protoplasts transfected with pGUSgLS or pFF19G shows the same high level of activity, whereas in pGUSsgLS the activity is reduced nearly threefold.

A correlation can be established between the in vitro and in vivo results in the case of the sgLS. Indeed, the presence of the relatively long sgLS decreases the level of translation of ORF3 and ORF4 in vitro, and also decreases the level of GUS nearly fourfold in vivo with respect to the influence of the relatively short gLS. Experiments using deleted versions of sgLS will further clarify the effect of the length of the sgLS on the level of GUS in vivo, and similar experiments with gLS should determine whether in this case also a correlation can be established between in vitro and in vivo experiments.

Our results indicate that the LSs of PLRV are not translational enhancers. Since decrease of translation efficiency caused by LSs is a rare phenomenon (only lack of enhancement of translation, but not decrease of translation has been reported [6,20]) and because a correlation exists between the results of in vitro and in vivo systems for the sgLS, this might suggest that LSs of PLRV serve other functions than those of enhancing gene expression.

Probably these LSs function as attachment sites for the viral replicase and take part in initiation of replication. Shortening of either LS causes increased translation efficiency of the downstream located genes. This can be explained by the shorter distance that ribosomes must migrate to reach the initiator AUG. However, it has been shown that if the LS is too short, initiation of translation is not efficient [21], and that subsequent systematic lengthening of the LS causes an increase in translation efficiency as has been shown for the CAT reporter gene using an in vitro system [22].

We propose that the main function of both LSs is to maintain synthesis of the proteins in the proper ratio, 4/1 in the case of ORF0/ORF1 and 1/1 for CP/17 kDa. The efficiency of translation could be less important than the ratio. Nevertheless, we can not exclude the possibility that in plants expression of viral proteins may be different from what is observed in vitro.

Acknowledgements: We thank A.-L. Haenni for a critical reading of manuscript. We thank the editorial board of *Acta Polonica Biochemica* for allowing us to reproduce results presented in Fig. 3 in this paper. This work was supported by Grant 6P04B01615 of the State Committee for Scientific Research.

References

[1] Sleat, D.E. and Wilson, T.M.A. (1992) in: Genetic Engineering with Plant Viruses (Davies I.W. and Wilson T.M.A. Eds.), pp 55–80, CRC, Boca Raton, FL.
[2] Mushegian, A.R. and Shepherd, R.J. (1995) Microbiol. Rev. 59, 548–578.
[3] Klaff, P., Riesner, D. and Steger, G. (1996) Plant Mol. Biol. 32, 89–106.
[4] Johnston, J.C. and Rochon, D.M. (1996) Virology 221, 232–239.
[5] Turner, R.L., Glynn, M., Taylor, S.C., Cheung, M.K., Spurr, C., Twell, D. and Foster, G.D. (1999) Arch. Virol. 144, 1451–1461.
[6] Gallie, D.R., Sleat, D.E., Watts, J.W., Turner, P.C. and Wilson, T.M.A. (1987) Nucleic Acids Res. 15, 8693–8711.
[7] Mayo, M.A. and Ziegler-Graff, V. (1996) Adv. Virus Res. 46, 416–460.
[8] Juszczuk, M., Zagorski-Ostoja, W. and Hulanicka, D.M. (1997) Acta Biochim. Pol. 44, 69–78.
[9] Palucha, A., Sadowy, E., Kujawa, A.B., Juszczuk, M., Zagorski-Ostoja, W. and Hulanicka, D.M. (1994) Acta Biochim. Pol. 41, 405–414.
[10] Sadowy, E., Pluta, K., Gronenborn, B. and Hulanicka, D.M. (1998) Acta Biochim. Pol. 45, 611–619.
[11] Timmermans, M.C., Maliga, P., Vieira, J. and Messing, J. (1990) J. Biotechnol. 14, 333–344.
[12] Kujawa, A.B., Drugeon, G., Hulanicka, D. and Haenni, A.-L. (1993) Nucleic Acids Res. 21, 2165–2171.
[13] Laemmli, U.K. (1970) Nature 227, 680–685.
[14] Hibi, T., Rezelman, G. and van Kammen, A. (1975) Virology 64, 308–318.
[15] van Beek, N.A.M., Derksen, A.C.G. and Dijkstra, J. (1985) J. Gen. Virol. 66, 551–557.
[16] Jefferson, R.A., Kavanagh, T.A. and Bevan, M.W. (1987) EMBO J. 6, 3901–3907.
[17] Bradford, M.M. (1976) Anal. Biochem. 72, 248–254.
[18] Kozak, M. (1988) Mol. Cell. Biol. 8, 2737–2744.
[19] Miller, J.S. and Mayo, M.A. (1991) J. Gen. Virol. 72, 2633–2638.
[20] Jobling, S.A., Cuthbert, C.M., Rogers, S.G., Fraley, R.T. and Gehrke, L. (1988) Nucleic Acids Res. 16, 4483–4498.
[21] Kozak, M. (1991) Gene Expr. 1, 111–116.
[22] Kozak, M. (1991) Gene Expr. 1, 117–125.