FEBS Letters 434 (1998) 8–12 FEBS 20701

The distance between the 3′-pyrimidine-rich tract and the AUG codon modulates internal initiation of translation of hepatitis A virus RNA

Jane Silveira Carneiro[b], M. Bucci[a], M. Equestre[c], P. Pagnotti[a], Alessandra Pierangeli[a], R. Pérez Bercoff[a,*]

[a] *Department of Cellular and Developmental Biology, University of Rome 'La Sapienza', V.le di Porta Tiburtina 28, 00185 Rome, Italy*
[b] *IBAMA, Brasilia, Brazil*
[c] *Istituto Superiore di Sanità, Rome, Italy*

Received 25 May 1998; revised version received 22 July 1998

Abstract Protein synthesis directed by hepatitis A virus (HAV) RNA is mediated by a mechanism involving the recognition of internal sequences. Two in-frame AUG codons initiate the long open reading frame (positions 734–736 and 740–742). The extracistronic region extending between the uncapped 5′-end and the ORF contains two pyrimidine-rich tracts (PRTs): one 12 nucleotides in length in the close vicinity of the initiator AUG, and a longer one between bases 94 and 140. In order to study the relative contribution of these elements to the process of internal initiation of translation, cDNA representations of the 5′-terminal extra-cistronic region of HAV RNA were inserted in the intergenic region of the bi-cistronic plasmid pSV-GH/CAT, between the genes encoding the human growth hormone (GH) and the bacterial enzyme chloramphenicol acetyltransferase (CAT), and following transfection of COS-1 cells, the transient expression of both genes was quantified. The importance of the 3′-PRT appeared to be strongly influenced by the length of the 'spacer' sequence extending between this structure and the translation initiation site: placed 45 nucleotides upstream from the initiator codon of a reporter gene, its integrity was stringently required for initiation to occur. Bringing the length of the 'spacer' back to its actual size in HAV RNA (i.e. 11 or 17 nt) reduced considerably the overall rate of internal initiation of translation, and the relative contribution to this process of the 3′-PRT became marginal. Concomitantly, the importance of the functional domains previously identified in the 5′-PRT fluctuated: while integrity of domain 100–106 was always stringently required for initiation to occur, the activity of domain 113–118 paralleled that of the 3′-PRT, and the opposite applied to domain 121–126, whose contribution became relevant only after switching off the 3′-PRT. Systematic mutations introduced in the 'spacer' sequences suggest that the length of this region may be responsible for the down regulation of translation of HAV RNA and, possibly, for its lengthy replication cycle.

Key words: Hepatitis A virus RNA; Pyrimidine-rich tract; Translation; Bi-cistronic mRNA; Spacer

1. Introduction

The genomic RNA of picornaviruses is exceptional among eukaryotic mRNAs in that it lacks a 5′-terminal cap structure, and previous work from this laboratory provided the first experimental evidence supporting the contention that translation of picornavirus RNA involves a most efficient mechanism of internal (rather than 5′-terminal) initiation that can by-pass the need for either a cap or a 5′-end [1–3]. Several lines of evidence indicate that this mechanism may not be confined to the RNA of picornaviruses, and that other eukaryotic mRNAs may use a similar mechanism of translation, an observation suggesting that the above findings may have broader biological implications than so far suspected ([4] and references therein).

The causative agent of human hepatitis A is also a picornavirus of the hepatovirus group which, in contrast to other members of this family, replicates very slowly (several days instead of a few hours) in all tissue culture systems so far studied. There is no definitive explanation for this unusual behavior. The genome of hepatitis A virus (HAV) is a single-stranded RNA molecule, 7478 nucleotides in length, of positive polarity, whose uncapped 5′-end is blocked by a small protein (VPg) [5,6]. The general organization of HAV genomic RNA does not depart substantially from that of other members of the Picornaviridae family: a single open reading frame (ORF) extends from base 734 through position 7410, followed by a short extracistronic region (68 nt), and a 3′-terminal poly(A) tract. Two in-frame AUG codons initiate the ORF at positions 734–736 and 740–742 (nucleotide numbering according to the sequence published by Cohen et al. [7], GenBank accession number M14707). While the former may be used in vitro, there is some evidence suggesting that the second is preferentially selected in vivo [8].

An unusually long extra-cistronic region (5′-ECR) containing nine (apparently unused) AUG codons precedes the uninterrupted coding sequences. Two main features distinguish the 5′-ECR of HAV RNA (Fig. 1): the presence of two (instead of just one) pyrimidine-rich tracts (PRTs), on the one hand, and the exceedingly short distance ('spacer' sequence) extending between the 3′-PRT and the translation initiation sites, on the other.

This structural organization, with a polypyrimidine sequence canonically located in the vicinity of the main (ORF), and a second, longer one in an abnormal location (nt 94–140), far removed from the translation initiation site, looked rather unusual, because the PRT sequence and the distance between this element and the closest downstream AUG triplet are among the more constant features throughout the Picornaviridae family [9]. Moreover, in HAV RNA the length of the 5′-PRT may vary in different isolates although its basic structure does not: it consists of a conserved hexanucleotide of the form (U)UUUCCC, reiterated 3–5 times. Mutational analysis of this tract allowed us to identify four functional domains differently involved in securing the internal initiation of translation of a reporter gene in bi-cistronic mRNAs [10]. Experiments designed to assess the role of the polypyrimidine sequences in translation driven by

*Corresponding author. Fax: (39) (6) 446 2306.

PII: S0014-5793(98)00939-9

the 5′-ECR of HAV RNA led us to postulate that the 5′-PRT is indeed part of the *cis*-acting elements required to initiate translation, and that the 3′ one might play a merely vicarious role due, perhaps, to the too short distance intervening between this element and the downstream AUG codon [4]. This view was indirectly supported by the observation that the reduction of the spacer between the PRT and a downstream (silent) AUG triplet to less than 14 nt abrogated translation driven by poliovirus 5′-ECR [11].

In contrast, when the 'spacer' of encephalomiocarditis virus (EMCV) RNA was shortened from 25 to 11–14 nt by mutations and/or deletions, only a 30–35% decrease in the rate of in vitro translation was observed [12]. Interestingly, the general organization of EMCV RNA is reminiscent of that of the HAV genome in that besides a PRT in the vicinity of the translation initiation site, it contains a second, longer polypyrimidine sequence, namely the poly(C) tract, 5′-distal to the IRES.

Taken together, these observations suggested that the PRTs and the 'spacer' of HAV RNA may be mutually balanced, and that the notoriously low rate of translation of HAV RNA may be the result of an unfavorable arrangement of these elements.

In an attempt to characterize the nature of the *cis*-acting elements that secure the internal initiation of translation of HAV RNA, and in order to assess the relative contribution of each PRT in this process, a series of bi-cistronic plasmids were constructed containing cDNA representations of the 5′-ECR of HAV RNA (or its mutants) inserted between the genes encoding the human growth hormone (GH) and the bacterial chloramphenicol acetyltransferase (CAT) under the transcriptional control of the late promoter of SV-40. In these constructs we engineered 'spacers' between the 3′-PRT of the HAV genome and the AUG that initiates the reporter gene CAT which reproduced either the situation existing in HAV RNA (i.e. 11 or 17 nt), or created a novel environment with the 3′-PRT 45 nt upstream of the initiator codon, and we report here on the different weight that each element seems to acquire in this process as a function of the activity of the other PRT.

2. Material and methods

2.1. Cells and bacteria

COS-1 cells were routinely maintained for up to 6 passages after thawing as monolayers in Dulbecco's modified minimum essential medium (DMEM), supplemented with 2% of the synthetic serum substitute Ultroser G (Sepracor, France) and antibiotics. Cultures of *Escherichia coli* TG-1 were transformed by treatment with calcium chloride and selected for ampicillin resistance according to standard procedures [13].

2.2. Oligonucleotide-directed mutagenesis

Mutated cDNA sequences were generated in polymerase chain reactions (PCR) driven by pairs of perfectly complementary oligonucleotides containing the desired mutations. Basically, this was done as already described [10,14]. Briefly, pairs of linear cDNA fragments were generated in separate reactions using one of the mutated oligonucleotides and an external primer. Following purification, they were allowed to hybridize their overlapping sequences, extended and the fragments so generated were further amplified in PCRs driven in this case by the external primers. The mutated linear cDNA fragments were gel purified, digested with appropriate enzymes as indicated in the text, and introduced by 'cassette exchange' into the enzyme-digested, phosphatase-treated and gel purified plasmid vector. The correct insertion of the properly mutated fragments were ascertained by direct DNA sequencing.

2.3. Construction of bi-cistronic plasmids

The construction of bi-cistronic plasmids containing cDNA representations of either the native or mutated 5′-ECR of HAV RNA inserted between the genes encoding human GH and bacterial CAT has been described in detail [4,10].

2.4. Transfection

$4–6\times10^5$ COS-1 cells in 10 ml DMEM, supplemented with 2% synthetic Ultroser and antibiotics, were seeded in 10 cm diameter plastic dishes (Nunc clone). Two days later, media were removed, and cultures transfected with 10 μg of the bi-cistronic plasmid DNA according to the DEAE-dextran/chloroquine procedure as described [14]. Following incubation at 37°C, the presence of human GH in the supernatant and of CAT in the cellular lysates was assayed as described [10].

2.5. Computer assisted sequence analysis

This was performed using the GCG Software package, University of Wisconsin [15].

3. Results and discussion

3.1. The length of the 'spacer' and the activity of the 3′-PRT

We first investigated the effect that changes in the length of the sequences extending between the 3′-PRT and the AUG translation initiation codon of a reporter gene had on both the overall level of its expression and the relative contribution to this process of the 3′-PRT.

To this end, cDNA representations of the 5′-ECR of HAV RNA were introduced in the *Eco*RV site of the bi-cistronic plasmid pSV-GH/CAT [9], between the genes encoding the human GH and the bacterial CAT. Note that the two in-frame initiator AUG codons of HAV RNA (positions 734–736 and 740–742) have been removed in all constructions to secure that the translation of the CAT gene is initiated at its legitimate site. This prevents the synthesis of a fusion protein devoid of CAT activity as already shown [10].

In the first series of constructs, a foreign string of 45 nt that was part of the intergenic region of the bi-cistronic plasmid extended between the 3′-PRT and the initiator codon of the

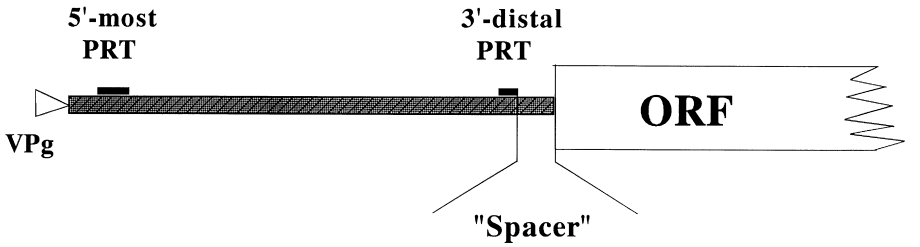

Fig. 1. The 5′-terminal extra-cistronic region of HAV RNA. The PRTs (thick bars at the top of the main line), and the 'spacer' sequence between the 3′-PRT and the translation initiation sites are out of scale.

reporter gene CAT. Following transfection of COS-1 cells, the transient expression of GH and CAT was determined as described [10]. Briefly, the levels of GH present in the supernatants of the cultures 72 h post-transfection were assessed in triplicate 100 μl samples, and served as an indicator of the efficiency of transfection, which was identical for all constructions.

Under these conditions, the efficiency of internal initiation of translation driven by the 5′-ECR of HAV RNA was substantially similar to that of the 5′-ECR of poliovirus RNA [10]. More relevant to our point, the 3′-PRT proved to be a functional structure and to participate actively in directing internal initiation of translation: Mutations were engineered in different domains of the 12 nt long 3′-PRT as depicted in Fig. 2A, and cDNA copies of HAV RNA carrying these substitutions were inserted in the bi-cistronic constructs. Following transfection of COS-1 cells, the expression of the reporter gene was quantified.

Removal of the 'core' 12 nt (Pyr-Del in Fig. 2) resulted in an almost 90% decrease in the expression of the CAT gene, and mutations introduced in the 3′-terminal half of the PRT

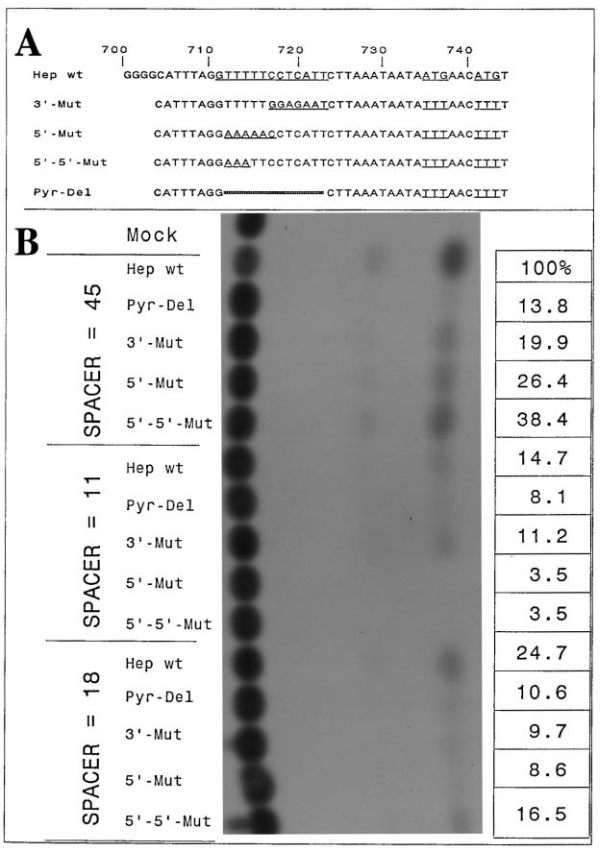

Fig. 2. Effect of mutations of the 3′-PRT on internal initiation of translation. A: Mutations (underlined) introduced in the 3′- PRT. In the wild type sequence (upper line), the 3′-PRT and the two in-frame AUG codons are indicated. Note that the latter have been removed in the mutants to secure translation from the AUG codon of the reporter CAT gene. B: CAT activity in lysates of COS-1 cells transfected with the bi-cistronic plasmids described in the text. The ratio of acetylated to non-acetylated chloramphenicol obtained after incubation of the reactions with the lysate of cells transfected with the construct carrying the wt 5′-ECR of HAV RNA and a 'spacer' of 45 nt is considered 100% activity.

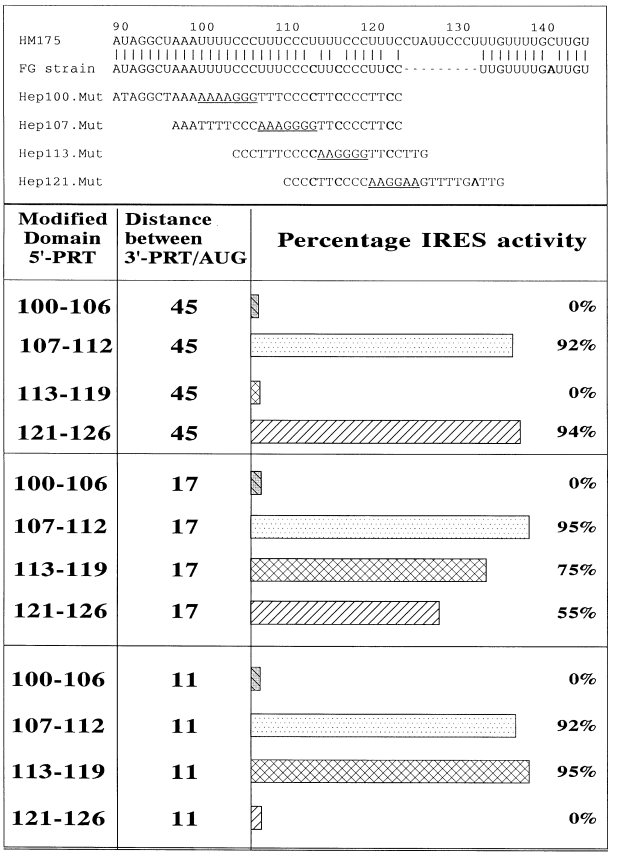

Fig. 3. Effect of the length of the 'spacer' on the activity of functional domains of the 5′-PRT. Upper panel: Mutations introduced in the 5′-PRT [10]. HM175 is the HAV isolate reported by Cohen et al. [7] (GenBank accession number M14707); FG ('fast-growing') is the strain used in these studies [16]. Lower panel: Activity of individual domains of the 5′-PRT as a function of the length of the spacer. CAT activity in lysates of COS-1 cells was quantified as in [10]. Histogram bars indicate the relative IRES activity.

caused a severe (5-fold) reduction of translation. Changes in its 5′-distal domain appeared to be better tolerated. The primary sequence of the 'spacer', on the other hand, seems to be totally irrelevant (Fig. 2).

In HAV RNA, however, the 3′-PRT is located 11 and 17 nt upstream of the first and second in-frame AUG codons of the main ORF, respectively. In order to study the role of the PRTs under 'physiological' conditions, i.e. in constructs that reproduce such an environment, *Kpn*I restriction sites were introduced in the 'spacer' sequence of the above mutants so that upon digestion of the PCR-generated fragments, ligation and PCR re-amplification of the products, the distance between the 3′-PRT and the AUG codon of the reporter CAT gene was reduced to 11 or 17 nt. Under these conditions a totally different picture emerged: the overall efficiency of translation was reduced by 85% and 75%, respectively compared to the levels of expression observed when the 3′-PRT was 45 nt from the reporter gene (Fig. 2), and mutations introduced in any domain of the 3′-PRT had little effect on the already low translation ability of these messengers. This appeared to reproduce a condition reminiscent of that observed with HAV RNA, which seems to be a very poor mRNA in all cell-free systems so far tested [8].

The general conclusion from these experiments was that the

shorter the 'spacer' between the 3′-PRT and the initiator AUG codon, the lower the efficiency of the whole 5′-ECR in translation, and the less significant (and more dispensable with) the role played by the 3′-PRT. These findings raised also the question of why the evolutionary pressure has not selected for HAV genomes with longer 'spacers' between the 3′-PRT and the initiator codon. While for the time being there is no obvious answer, it might be proper to recall that following serial plaque transfer of foot-and-mouth disease virus, mutants have been isolated that carried a progressively longer A-rich sequence immediately upstream the initiator AUG. In this case, the longer the A string, the poorer the replication ability of the mutants [17].

3.2. Functional organization of the 5′-PRT

The relative contribution of the 3′-PRT to the process of translation appeared to be directly proportional to the length of the 'spacer' intervening between this structure and the initiation codon. The question, therefore, arose as to whether the shortening of the 'spacer' and the concomitant switching off of the activity of the 3′-PRT could modulate or modify the activity of the different functional domains that we had previously identified in the 5′-PRT.

To address this issue, the bi-cistronic constructs containing a cDNA representation of HAV 5′-ECR with a native, unmodified 3′-PRT at either 45, 17 or 11 nt from the target AUG were digested with *Avr*II (which cuts at positions 82 and 241 in the sequence of HAV 5′-ECR), and mutations in the four well defined domains of the 5′-PRT were introduced by 'cassette exchange'. Plasmids were propagated in competent bacteria, and in every case all the mutations introduced were verified by direct DNA sequencing.

COS-1 cells were then transfected with these constructs, and the transient expression of the CAT gene was measured as described. The picture emerging from these experiments is clearly depicted in Fig. 3, and is summarized in its right-hand panel: both PRTs present in the 5′-terminal extra-cistronic region of HAV RNA are part of the *cis*-acting elements required to secure internal initiation of translation, but they seem to have a very different weight in directing this process.

When the 3′-PRT was 45 nt upstream of the target AUG, domains 100–106 and 113–119 of the 5′-PRT appeared to be stringently required for translation, and changes introduced in these short stretches abolished the ability of HAV 5′-ECR to drive internal initiation of translation. Domains 107–112 and 121–126, on the other hand, appeared to play a minor role, and substitutions introduced in these sequences did not significantly change the levels of translation [10].

By bringing the initiation AUG codon closer to the 3′-PRT, the contribution of the latter structure to the efficiency of the process of translation diminished considerably, and a concomitant redistribution of the roles of each domain within the 5′-PRT could be observed: while the sequences encompassing nt 100–106 seem to be always indispensable, and under no condition tolerate any change, translation seems to be totally indifferent to the sequences spanning nt 107–112, which could be modified without any effect on the expression of the reporter CAT gene.

The substitutions introduced in domain 121–126 (which had no effect when the 3′-PRT was active, i.e. at 45 nt from the target) abrogated the translation of the downstream gene as soon as the 3′-PRT was switched off, as if the activation of the 3′-PRT provided a function normally covered by domain 121–126 in the 5′-PRT.

The opposite seems to apply to domain 113–119, whose activation moved in parallel to that of the 3′-PRT: stringently required when the 3′-PRT was active (spacer: 45 nt), its role became less prominent with a dimmer 3′-PRT (spacer: 17 nt), and was totally dispensed with when the 3′-PRT switched off, 11 nt upstream of the target AUG.

Three main features emerge from these findings. (a) The notoriously poor translation ability of HAV RNA is, in the first place, due to a large extent to the too short 'spacer' sequence separating the 3′-PRT and the translation initiation codon AUG. Increasing the distance between these elements considerably improves the efficiency of internal initiation of translation.

(b) Under the structural conditions of HAV RNA, the 3′-PRT seems to play a vicarious role, helping perhaps in the correct positioning of the ribosome at initiation of translation. The preferential use of the second AUG codon, 17 nt downstream of the 3′-PRT, may therefore be a consequence of its location at a better (or less penalized) distance from the 3′-PRT.

(c) The 5′-most poly-pyrimidine tract, an element of variable length, where the basic module (U)UUUCCC is repeated 4–5 times, emerges from these studies as a functional mosaic of activated and inactive domains, whose function can only in part be provided by the 3′-PRT.

Although for the time being the nature of the function(s) of the PRTs remains to a large extent a matter of speculation, the available evidence tends to support the notion that the non-contiguous PRTs and the 'spacer' sequence are part of a larger structure, possibly brought together by secondary and tertiary folding, and required to direct internal initiation of translation. Some of these elements have also been identified in the 5′-ECR of eukaryotic mRNAs (notably those encoding ribosomal proteins, among others), and the possibility should be entertained that they may be translated using a similar mechanism, at least during defined stages of development. The efficiency of the recognition of these structure by cellular factor(s) may constitute the limiting step that determines the overall efficiency of translation.

Acknowledgements: The authors are grateful to Dr. Ned Mantei (ETH, Zürich, Switzerland), to Dr. Jean Lucas-Lenard (Storrs, CT, USA) for most useful comments and discussions and to the anonymous reviewer of this manuscript for helpful suggestions. J.S.C., on leave of absence from IBAMA (Brazil), was a Fellow of the ICGEB (Trieste, Italy). This work was partly supported by grants to R.P.B. from the Italian National Research Council (CNR), the Institut Pasteur/Fondazione Cenci-Bolognetti (Rome), and the Targeted Project 'Viral Hepatitis' of the Italian National Institute of Health (ISS).

References

[1] Pérez Bercoff, R. (1982) in: Protein Biosynthesis in Eukaryotes (Pérez Bercoff, R., Ed.), pp. 242–252, Plenum Press, New York.
[2] Pérez Bercoff, R. and Kaempfer, R. (1982) J. Virol. 41, 30–41.
[3] Degener, A.M., Pagnotti, P., Facchini, J. and Pérez Bercoff, R. (1983) J. Virol. 45, 889–894.
[4] Degener, A.M., Silveira Carneiro, J., Cassetti, C., Pierangeli, A., Pagnotti, P., Bucci, M. and Pérez Bercoff, R. (1995) Virus Res. 37, 291–303.
[5] Wang, C. and Siddiqui, A. (1995) in: Cap-independent Translation (Sarnow, P., Ed.), pp. 99–116, Springer Verlag, Berlin.
[6] Iizuka, N., Chen, C., Yang, Q., Johannes, G. and Sarnow, P.

(1995) in: Cap-independent Translation (Sarnow, P., Ed.), pp. 155–177, Springer Verlag, Berlin.
[7] Cohen, J.J., Ticehurst, J.R., Purcell, R.H., Buckler-White, A. and Baroudy, B.M. (1987) J. Virol. 61, 50–59.
[8] Tesar, M., Harmon, S.A., Summers, D.F. and Ehrenfeld, E. (1992) Virology 186, 609–618.
[9] Nicholson, R., Pelletier, J., Le, S.Y. and Sonenberg, N. (1991) J. Virol. 65, 5886–5894.
[10] Silveira Carneiro, J., Equestre, M., Pagnotti, P., Gradi, A., Sonenberg, N. and Pérez Bercoff, R. (1995) J. Gen. Virol. 76, 1189–1196.
[11] Pilipenko, E.V., Gmyl, A.P., Maslova, S.v., Svitkin, Y.V., Sinyakov, A.N. and Agol, V.I. (1992) Cell 68, 119–131.
[12] Kaminski, A., Belsham, J.G. and Jackson, R.J. (1994) EMBO J. 13, 1673–1681.
[13] Gorman, C. (1985) in: DNA Cloning: A Practical Approach (Glover, D.M., Ed.), pp. 143–190, IRL Press, Oxford.
[14] Pierangeli, A., Bucci, M., Pagnotti, P., Degener, A.M. and Pérez Bercoff, R. (1995) FEBS Lett. 374, 327–332.
[15] Devereux, J., Haeberli, P. and Smithies, O. (1984) Nucleic Acids Res. 12, 387–395.
[16] Venuti, A., Di Russo, C., Del Grosso, N., Patti, A.M., Rugeri, F., De Stasio, P.R., Martiniello, M.G., Pagnotti, P., Degener, A.M., Midulla, M., Panà, A. and Pérez Bercoff, R. (1985) J. Virol. 56, 579–588.
[17] Escaramis, C., Dávila, M., Charpentier, N., Bracho, A., Moya, A. and Domingo, E. (1996) J. Mol. Biol. 264, 255–267.