An in vitro DNA virus for in vitro protein evolution

Ichiro Tabuchi[a,b], Sayaka Soramoto[a], Naoto Nemoto[b], Yuzuru Husimi[a,*]

[a] *Department of Functional Materials Science, Saitama University, 255 Shimo-okubo, Saitama 338-8570, Japan*
[b] *GenCom Co., 11 Minami-oya, Machida 194-8511, Japan*

Received 27 July 2001; revised 19 October 2001; accepted 19 October 2001

First published online 5 November 2001

Edited by Takashi Gojobori

Abstract In vitro virus is a molecular construct for in vitro protein evolution, which requires some mechanism to link phenotype to genotype. The first in vitro virus was realized by bonding a nascent protein with its coding mRNA via puromycin in in vitro translation. We report a new construct of in vitro DNA virus. The virion was a covalent cDNA–protein fusion, and virion formation did not require any modification of mRNA. Due to intactness of mRNA, this type of in vitro DNA virus will take the next step toward in vitro autonomous evolution, just like in vivo viral evolution in a cellstat.

Key words: cDNA–protein fusion; Evolutionary protein engineering; In vitro selection

1. Introduction

Evolutionary molecular engineering has been demonstrating its effectiveness in designing novel functional biopolymers. The Darwinian selection process for proteins or peptides requires some linking strategy between phenotype and genotype. There are at least three such strategies both in the laboratory and in nature [1]: virus-type, cell-type and external intelligence-type. In virus-type linking strategy, the genotype molecule (DNA or RNA) is bound to the phenotype molecule (protein) directly just as in a simple bacteriophage particle (virion). A phage display [2] and a cellstat [3] were developed for the study of molecular evolution in the laboratory using this strategy. An in vitro version of these methods has been developed [4,5]. We called it in vitro virus [1,4] because it can be regarded as the simplest virus with a single gene and a single coat protein, and its 'life cycle' is turned over in vitro, i.e. its 'host' is a test tube. The main reason is, of course, its genotype–phenotype bonding strategy other than the compartmentalization strategy of a cellular organism [1]. By attaching puromycin to the 3′-end of an mRNA, a covalent linkage between a nascent protein and its encoding mRNA was made on a ribosome during in vitro translation. It dramatically increased the library diversity and the system flexibility in comparison with the in vivo version. In fact, functional proteins have been evolved starting from a random library in this approach [6,7]. The original in vitro virus had some problems. For example, it required a modification to the mRNA template to attach puromycin-linker, and the template could not tolerate harsh selective conditions because the genotype RNA molecule is fragile. Some in vitro selection systems consisting of DNA–peptide fusion have been developed to overcome the latter drawback. There are various types for these DNA–peptide fusions [8–10]. But because these methods still required modifying the evolving template, it was difficult to make an autonomously evolving molecular system.

We developed a new type of in vitro DNA virus, in which the virion formation consisted of the following processes (Fig. 1): (1) hybridization of a DNA-primer with a puromycin-linker to an mRNA, (2) in vitro translation, and (3) reverse transcription, which lead to cDNA–protein fusion. We developed two constructs of the DNA-primer with a puromycin-linker; L-linker-primer and T-linker-primer (Fig. 2A). Thus this in vitro DNA virus is free from necessity of modifications of the evolving template.

2. Materials and methods

2.1. Synthesis of DNA-primer with puromycin-linker

The L-linker-primer and the DNA-primer moiety of the T-linker-primer (5′-$(GCCCCCCG)_{OMe}$(T-NH_2)CCC-3′) were synthesized by the use of a standard DNA synthetic method. The T-linker-primer was synthesized in the first step as two separate parts by a DNA synthesizer and each primary amine on the two parts was linked together via bifunctional *N*-hydroxysuccinimide ester. 5′ protection of the puromycin-linker moiety of the T-linker-primer (5′-NH_2-$(spacer18)_4$-$(CC)_{OMe}$-puromycin-3′-controlled pore glass (CPG)) and base protections of the DNA-primer moiety of T-linker-primer were deprotected by standard protocols. Puromycin-CPG and phosphoramidites were purchased from Glen Research and used according to the protocols recommended by the manufacturer. The puromycin-linker moiety of the T-linker-primer (1 μmol) was activated by FluoroLink Cy3 reactive dye (1 mg) (Amersham Pharmacia) in dimethyl sulfoxide (DMSO) for 30 min at 25°C and washed with DMSO. DNA-primer moiety (1 μmol) was added and incubated in DMSO for 3 h at 25°C and washed with DMSO and TE buffer (10 mM Tris–HCl, 1 mM EDTA, pH 8). Concentrated NH_4OH was added to remove the linker-primers from CPG and deprotection of bases was performed for 8 h at 55°C. The linker-primers were purified with a reverse phase cartridge (Glen Research).

2.2. Construction of in vitro virus genome

Three constructs of in vitro virus genome, 'His'(169 bp), 'Flag'(127 bp), and 'FLH'(177 bp)(Fig. 2B), were made. The plasmid pETHisKGFP was constructed by inserting green fluorescent protein (GFP) gene of phGFPS65T (Clontech) into pETHisK vector [11]. DNA genomes were prepared by PCR with different primer sets in two steps using the plasmid as an initial template.

Capped mRNAs were prepared using the DNA genomes and T7

*Corresponding author. Fax: (81)-48-858 3531.
E-mail address: husimi@fms.saitama-u.ac.jp (Y. Husimi).

Abbreviations: DMSO, dimethyl sulfoxide; CPG, controlled pore glass; NTA, nitrilotriacetic acid; GFP, green fluorescent protein; OMe, 2′-*O*-methyl ribonucleotide; WGE, wheat-germ extract

0014-5793/01/$20.00
PII: S0014-5793(01)03075-7

RNA polymerase transcription kit (RiboMAX-T7; Promega). mRNAs (1 μM) were hybridized with each linker (3 μM) in an annealing condition of 15 min/from 94 to 25°C in 10 mM Tris–HCl, 1 mM $MgCl_2$, 50 mM KCl (pH 8) and purified by a spin column (S-200HR; Amersham Pharmacia).

2.3. Formation of in vitro DNA virion

The mRNAs were translated in wheat-germ extract (WGE; Promega) for 15 min at 25°C in standard condition and for 30 min at 10°C in a high salt condition (finally, 50 mM $MgCl_2$, 500 mM KCl). Translated products were desalted with the spin column. Reverse transcription was performed using *C. therm.* Polymerase (Roche) for 30 min at 65°C.

Virion formation was analyzed for 'FLH' genome. Virion formation procedure for this purpose was the same as described above except that the virion was labeled with a fluorescent nucleotide (FluoroLink Cy5-dCTP; Amersham Pharmacia) and a fluorescent amino acid (FluoreTect Green in vitro translation labeling system; Promega), and WGE was purchased from Toyobo. The products were analyzed using 6 M urea 10% SDS–Tricin PAGE [12] and visualized using a fluorescence imager (Molecular Imager FX; Bio-Rad). Virion formation analysis using PCR was performed through the same procedure described below as the selective enrichment analysis using Ni–nitrilotriacetic acid (Ni–NTA) agarose gel, except that thiopropyl Sepharose purification was omitted and PCR was performed normally (25 cycles).

2.4. Selective enrichment analysis

Affinity columns of Ni–NTA agarose (QIAGEN; 200 μl volume (V_c)) and anti-Flag M2 antibody agarose (Sigma) were used. In vitro DNA virions were purified with thiopropyl Sepharose (Sigma) as described in [13]. Purified virion was diluted in a TBS buffer (50 mM Tris–HCl, 150 mM NaCl, 0.1% Tween 20, pH 8.0) and incubated in each affinity column for 30 min at 25°C. Ni–NTA agarose column was washed with 25 V_c of a buffer (50 mM Tris–HCl, 1 M NaCl, 0.1% Tween 20, 10 mM imidazole, pH 8.0) and eluted with 2.5 V_c of a buffer (50 mM Tris–HCl, 0.1% Tween 20, 500 mM imidazole, pH 8.0). Anti-Flag agarose was washed with 10 V_c of TBS and eluted with 2.5 V_c of a buffer (0.1 M glycine HCl, 0.1% Tween 20, pH 3.5). A competitive PCR was performed (40 cycles) using KOD' polymerase (Toyobo). PCR products were analyzed by electrophoresis on 3% agarose gel (MetaPhor agarose; FMC) using triethanolamine/Tricin running buffer [14].

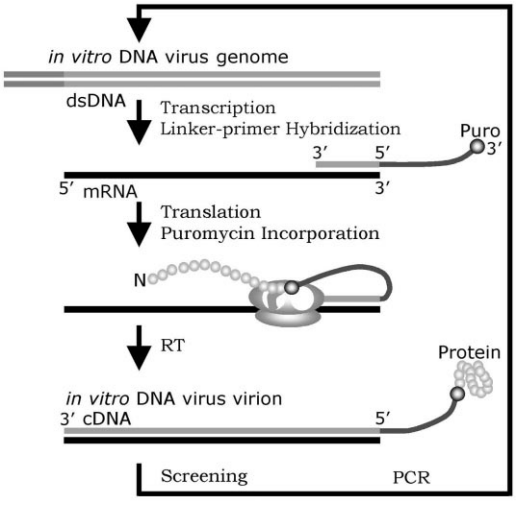

Fig. 1. 'Life cycle' of 'in vitro virus'. In vitro virus 'genome'(dsDNA) is transcribed to mRNA. The mRNA is hybridized to a DNA-primer with a puromycin-linker, and translated. On a ribosome, puromycin is incorporated to the C-terminus of the nascent protein. Reverse transcription starting from the DNA-primer gives a stable cDNA–protein fusion (in vitro virus 'virion'). The DNA moiety of screened virions judging from the 'fitness' of the protein moiety is amplified.

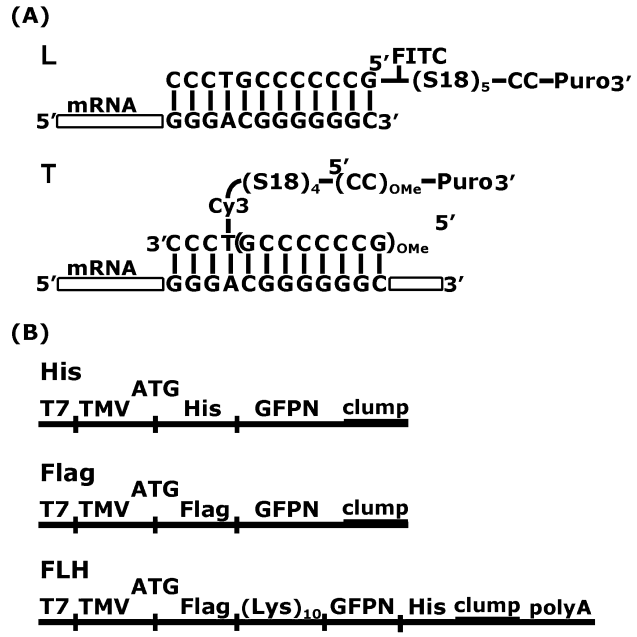

Fig. 2. A: Two constructs of a DNA-primer with a puromycin-linker; L-linker-primer and T-linker-primer. S18: spacer18, a polyethylene glycol-like spacer. Puro: puromycin. OMe: 2′-O-Me-RNA. B: Three constructs of in vitro virus genomes; 'His', 'Flag', and 'FLH'. T7: T7 promoter. TMV: translational enhancer of TMV 5′UTR [24]. ATG: Initiation codon with good context [25]. His: His-tag [26]. Flag: Flag epitope [27]. HMK: HMK site [11]. GFPN: N-terminus sequence in GFP gene with a mutation G4C. Clump: GC-rich sequence for hybridization to the DNA-primer, a GC clump.

2.5. Puromycin incorporation analysis

Fluorescein-dC-puromycin was synthesized using a standard DNA synthetic method and purified in the reverse phase HPLC. DNA-linker-primer/mRNA hybridization was performed through the same procedure as described in Section 2.2. The translation reaction was performed in WGE (Promega) according to the supplier's recommendation with 10 μM fluorescein-dC-puromycin. The translation products were analyzed using 15% SDS–PAGE.

3. Results and discussion

3.1. Incorporation of puromycin at a double-stranded site

It is known that low concentration puromycin is incorporated into the C-terminus of the growing peptide at the mRNA terminus [15] or at RNA–DNA junction [5]. We examined whether a hybridized double-stranded region on mRNA could make the elongation reaction pause or not, and puromycin could or could not be incorporated in this condition. Full length mRNA for GFP was hybridized with the anti-sense DNA-oligomer at about a third from the N-terminus site of the coding region, and was translated with fluorescence-labeled puromycin. The fluorescence-labeled protein of a third length was actually made (Fig. 3). This result indicates that mRNA/DNA-primer hybridization region makes peptide elongation reaction pause, and puromycin incorporation at the position follows. This process is applicable to the puromycin-linker attachment process in the making of in vitro virus. It has an advantage of rapidity and easiness without any chemical or enzymatic modification of mRNA. This method can attach the puromycin-linker at any position of the mRNA target even if it contains the stop codon, where-

as the original ligation method can attach only at the terminus.

The latter has another drawback associated with enzymatic ligation by the use of a splint DNA and undesirable mRNA terminal heterogeneity originated from T7 run-off transcription [16]. Recently, a linker attachment using photo-cross-linking was reported [17]. Photo-cross-linking reaction, however, needs a long (15 min) UV irradiation. Whereas, mRNA/DNA-primer hybridization seemed to be stable enough in this experiment, because a full length protein was not observed in lane D of Fig. 3. Therefore, there is no need to make a covalent bonding between mRNA and the puromycin-linker for the translation process.

3.2. Virion formation of in vitro DNA virus

As a model system to examine the virion formation, we used the genome 'His', which encoded GFP fragment with a His-tag (Fig. 2B). The virion was made and screened through a Ni–NTA-immobilized affinity column. Bound virion molecules were eluted and their moiety of DNA were amplified and analyzed (Fig. 4). Two constructs with each type of the linker-primers and a control without puromycin-linker were examined.

During a single turnover in the 'viral life cycle', both puromycin-linker-primers amplified the viral genomes (and also the virions) well. The T-linker-primer showed about three-fold higher amplification than the L-linker-primer. As the L-linker-primer has a hybrid double-stranded region of low flexibility, puromycin might not be able to approach a ribosome A-site efficiently. The T-linker-primer has a flexible spacer placed near the 3′ terminus of the DNA-primer. At the termination of translation, the ribosome might be located just before the 3′ end of the DNA-primer. So, puromycin is closer to the A-site, and the local concentration of puromycin might be higher. A spacer placed near the 3′ end of the DNA-primer did not affect the reverse transcription productivity (data not shown). The T-linker-primer has oligonucleotide moiety of 2′-O-Me-RNA, which bound to RNA with much higher thermostability than corresponding DNA [18]. Puromycin $(CC)_{OMe}$ might interact stably to the ribosome A-site. 2′-O-Me-RNA/mRNA is not cleaved by RNaseH activity, so mRNA degradation might be lower. The oligonucleotide moiety of the L-linker-primer was synthesized in 5′ to 3′ orientation using 5′-cyanoethyl phosphoramidites, but the L-linker-primer was not with 2′-O-Me construction, because 5′-cyanoethyl 2′-O-Me phosphoramidites were not commercially available.

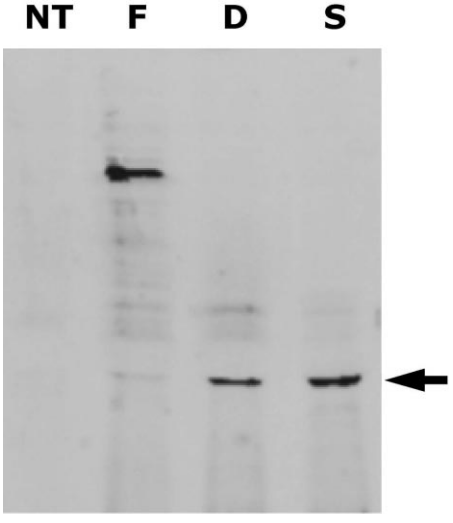

Fig. 3. Gel electrophoretic pattern showing incorporation of puromycin. Various mRNAs were translated with 10 mM fluorescein-dC-puromycin. Lane NT: control, without template. Lane F: Full length GFP (261 aa) mRNA. Lane D: Full length GFP mRNA hybridized with the DNA-oligomer at the region beginning at the 92nd aa site. Lane S: GFP N-terminus fragment (91 aa) mRNA. The products were analyzed using 15% SDS–PAGE and visualized using a fluorescence imager.

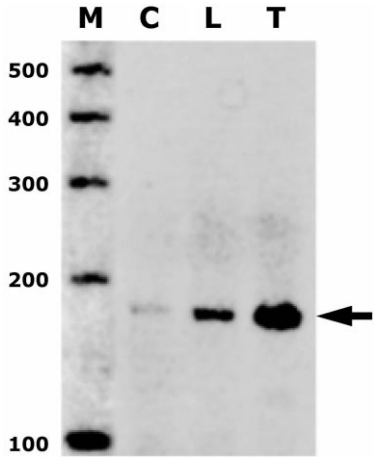

Fig. 4. Virion formation of in vitro DNA virus. Genome after single turnover of viral life cycle with each of two linkers was electrophoresed. Lane M: Fluorescent 100 base DNA ladder. Lane C: Control, with the DNA-primer without the puromycin-linker. Lane L: With L-linker-primer. Lane T: With T-linker-primer. The products were analyzed by electrophoresis on 3% agarose gel.

The electrophoretic pattern of virions (Fig. 5) showed a successful virion formation. The cDNA and the peptide were labeled with fluorescent dyes, Cy5 [19] and BODYPY-FL [20], respectively. Detected fluorescent bands were in correct mobility. A lower mobility band of Cy5 emission and a band of BODYPY-FL emission had the same mobility. It was concluded that these bands corresponded to in vitro DNA virus virion. The yield of cDNA–peptide linkage was about 10%.

3.3. Selective enrichment analysis

As a model system for demonstrating selection processes, we used a mixture of 'His' and 'Flag' genomes (Fig. 2B) that

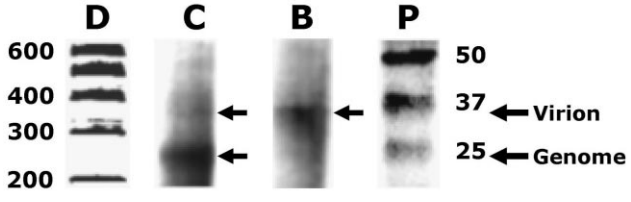

Fig. 5. Virion formation of in vitro DNA virus as revealed by double fluorescence label. Lane D: Fluorescent DNA ladder. Lane C: Virion in which cDNA moiety was labeled with fluorescent Cy5. Lane B: Virion in which protein moiety was labeled with fluorescent BODYPY-FL. Lane P: Protein molecular weight markers (precision prestained protein standard; Bio-Rad). The imager with 635/670 nm laser/filter set or 488/515 nm was used in lanes C and P, or, in lanes B and D, respectively. The products were analyzed using 10% SDS–Tricin PAGE with 6 M urea.

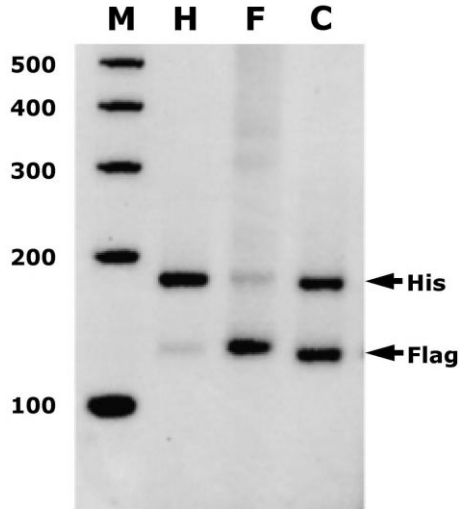

Fig. 6. Selective enrichment analysis. Genome after a single turnover in the viral life cycle with two screening conditions was electrophoresed. Lane M: Fluorescent 100 base DNA ladder. Lane C: Control, without column elution. Lane H: Trapped in Ni–NTA agarose column. Lane F: Trapped in anti-Flag agarose column. The products were analyzed by the same procedure as described in the legend of Fig. 4.

contained a coding sequence of His-tag and Flag epitopes, respectively. Starting from a 1:1 mixture of 'His' genome and 'Flag' genome we made virions and screened them through a Ni–NTA-immobilized affinity column (for His-tag) and through an anti-FLAG-immobilized affinity column. Bound virions were eluted and the DNA moiety was amplified and analyzed by gel electrophoresis. 'Flag' genome was slightly shorter than 'His' genome (Fig. 2B). PCR was performed in a competitive mode and a fluorescence labeled primer was used. Therefore the initial ratio of screened virions could be analyzed using the imager (Fig. 6). A 'His' band was enriched using a Ni–NTA agarose column and a 'Flag' band was enriched using anti-Flag agarose column. The control sample without screening was amplified, showing maintenance of the original ratio of two genomes. The enrichment of about two orders of magnitude was achieved per cycle. This value will be large enough to select target protein from a library of large diversity.

3.4. Conclusion

We demonstrated virion formation and selective enrichment in a model system of in vitro DNA virus that consists of cDNA–protein fusion without any mRNA modifications.

In vitro DNA virus virion consists of a cDNA/mRNA hybrid double helix and a protein bound covalently to the cDNA. This construct must be more stable than mRNA-peptide fusion, and tolerate harsh and stringent selection conditions, such as temperature stress-guided selection [21]. If higher stability will be required, the virion should be converted to a double-stranded DNA.

This method will be able to be applied to the in vitro selection not only for an artificial library but also for a natural mRNA library just as in the case of a linker attachment method using photo-cross-linking [22].

As it is free from necessity of mRNA modifications, in vitro virus will take the next step toward autonomous evolution, such as in vivo viral evolution in a cellstat. Combining the in vitro DNA virus with an isothermal amplification method for RNA/DNA such as 3SR (self-sustained sequence replication) [23], a continuous 'culture' of the in vitro virus will be realized. A flow reactor for this viral life cycle reaction will be a 'natural selection'-type evolution reactor [3] for a protein evolution.

Acknowledgements: We thank Dr. T. Sasaki for his helpful discussions. This work was supported in part by JSPS Grant 'Research for the Future' 96100101.

References

[1] Nemoto, N. and Husimi, Y. (1995) J. Theor. Biol. 176, 67–77.
[2] Smith, G.P. (1985) Science 228, 1315–1317.
[3] Husimi, Y., Nishigaki, K., Kinoshita, Y. and Tanaka, T. (1982) Rev. Sci. Instrum. 53, 517–522.
[4] Nemoto, N., Miyamoto-Sato, E., Husimi, Y. and Yanagawa, H. (1997) FEBS Lett. 414, 405–408.
[5] Roberts, R.W. and Szostak, J.W. (1997) Proc. Natl. Acad. Sci. USA 94, 12297–12300.
[6] Wilson, D.S., Keefe, A.D. and Szostak, J.W. (2001) Proc. Natl. Acad. Sci. USA 98, 3750–3755.
[7] Keefe, A.D. and Szostak, J.W. (2001) Nature 410, 715–718.
[8] Brenner, S. and Lerner, R.A. (1992) Proc. Natl. Acad. Sci. USA 89, 5381–5383.
[9] Doi, N. and Yanagawa, H. (1999) FEBS Lett. 457, 227–230.
[10] Kurz, K., Gu, K., Al-Gawara, A. and Lohse, P. (2001) Chem Biochem. 2, 666–672.
[11] Chen, B.P. and Hai, T. (1994) Gene 139, 73–75.
[12] Schagger, H. and Jagow, G.V. (1987) Anal. Biochem. 166, 368–379.
[13] Barrick, J.E., Takahashi, T.T., Balakin, A. and Roberts, R.W. (2001) Methods 23, 287–293.
[14] Lue, Q., Li, X. and Sommer, S.S. (1999) Anal. Biochem. 270, 112–122.
[15] Nemoto, N., Miyamoto-Sato, E. and Yanagawa, H. (1999) FEBS Lett. 462, 43–46.
[16] Liu, R., Barrick, J.E., Szostak, J.W. and Roberts, R.W. (2000) Methods Enzymol. 318, 268–293.
[17] Kurz, M., Gu, K. and Lohse, P.A. (2000) Nucleic Acids Res. 28, e83.
[18] Majlessi, M., Nelson, N.C. and Becker, M.M. (1998) Nucleic Acids Res. 26, 2224–2229.
[19] Yu, H., Chao, J., Patek, D., Mujumdar, R., Mujumdar, S. and Waggoner, A.S. (1994) Nucleic Acids Res. 22, 3226–3232.
[20] Gite, S., Mamaev, S., Olejnik, J. and Rothschild, K. (2000) Anal Biochem. 279, 218–225.
[21] Jung, S., Honegger, A. and Plückthun, A. (1999) J. Mol. Biol. 294, 163–180.
[22] Hammond, P.W., Alpin, J., Rise, C.E., Wright, M. and Kreider, B.L. (2001) J. Biol. Chem. 276, 20898–20906.
[23] Guatelli, J.C., Whitfield, K.M., Kwoh, D.Y., Barringer, K.J., Richman, D.D. and Gingeras, T.R. (1990) Proc. Natl. Acad. Sci. USA 87, 1874–1878.
[24] Gallie, D.R. and Watts, J.W. (1992) Nucleic Acids Res. 20, 4631–4638.
[25] Kozak, M. (1999) Gene 234, 187–208.
[26] Hochuli, E., Bannwarth, W., Döbeli, H., Gentz, R. and Stüber, D. (1988) Biotechnology 6, 1321–1325.
[27] Hopp, T.P., Prickett, K.S., Price, V.L., Libby, R.T., March, C.J., Cerretti, D.P., Urdal, D.L. and Conlon, P.J. (1988) Biotechnology 6, 1205–1210.