BIOLOGICAL LETT. 2011, 48(1): 47–56
Available online at: http:/www.versita.com/science/lifesciences/bl/ DOI: 10.2478/v10120-011-0005-9

Characterization of the protein gp1 from bacteriophage φIN93

ISAO MATSUSHITA[1] and HIDESHI YANASE[2]

[1] Energy Technology Laboratories, Osaka Gas Co., Ltd., 6-19-9, Torishima Konohana-ku, Osaka 554-0051, Japan
[2] Department of Biotechnology, Tottori University, 4-101 Koyamacho-minami, Tottori 680-8552, Japan
Corresponding author: Isao Matsushita, imatsu@osakagas.co.jp

(*Received on 20 January 2010; Accepted on 6 January 2011*)

Abstract: The protein gp1, encoded by the *ORF1* gene in the genome of the *Thermus* phage φIN93, is similar to the replication protein encoded by *repA* in the *Thermus sp.* plasmid. To confirm that gp1 functions as a replication protein, we constructed the recombinant plasmid pIA1, by inserting *ORF1* and the kanamycin acetyltransferase (*kat*) gene into pBluescriptII SK(+), which enabled replication of pIA1 in *Thermus thermophilus* HB27. By contrast, plasmid pIA1-del, which contained only a part of *ORF1* (the other part deleted using *Kpn*I), was not replicated. These results show that gp1 functions as a replication protein and that pIA1 can be used as a shuttle vector between *Thermus thermophilus* HB27 and *Escherichia coli*.

Keywords: replication protein gene, *Thermus thermophilus* TZ2, *Thermus thermophilus* HB27, bacteriophage, shuttle vector

INTRODUCTION

Some extremely thermophilic bacteriophages are able to infect *Thermus thermophilus* HB8, growing at temperatures ranging from 47°C to 85°C (Oshima & Imahori 1974). Among them, φYS40 (152,372 bp), *Thermus* P23-45 (84,201 bp) and *Thermus* P74-26 (83,319 bp) contain double-stranded DNA, and the functions of many of the genes in the respective genomes have been studied (Naryshkina et al. 2006; Minakhin et al. 2008). It is reported that these phages inherently contain DNA polymerase, but their ability to replicate genomic DNA has not been studied yet.

We isolated an extremely thermophilic phage, φIN93, which was induced from the lysogenic strain *Thermus thermophilus* TZ2 (Matsushita et al. 1995). The nucleotide sequence of the φIN93 genome is available in the DDBJ/EMBL/GenBank databases (accession number AB063393). An analysis of the φIN93 genome showed that gp1, encoded by the *ORF1* gene (1239 bp), has 28.1% similarity to the replication protein (RepA) encoded by *repA* in the *Thermus* sp. plasmid (de Grado et al. 1998; Matsushita et al. 2009). Based on that finding, we proposed that the replicative origin

of *ORF1* is likely situated within an AT-rich region, just as it is in *repA*, and that gp1 is also involved in the replication of φIN93 genomic DNA. To prove this, we inserted *ORF1* into *Escherichia coli* plasmid pBluescript SK+ and examined its replication in *Thermus thermophilus* HB27. A *kat* gene encoding a thermostable protein providing resistance to kanamycin was also inserted into the plasmid as a selection marker.

In this article we show that gp1 is necessary for plasmid replication, and propose that it acts in the same manner as RepA. In addition, although several *Thermus-Escherichia* shuttle vectors, containing a replicative origin derived from a plasmid of *Thermus* sp., have been developed so far (Lasa et al. 1992; Wayne et al. 1997; de Grado et al. 1999), here we describe the first construction of a *Thermus-Escherichia* shuttle vector containing a replication protein gene derived from a *Thermus* bacteriophage.

MATERIALS AND METHODS

Cell growth and phage infection

Thermus thermophilus TZ2 was grown overnight at 70şC in A-2 medium consisting of 0.1% tryptone, 0.1% yeast extract, and Castenholz basal salts (pH 7.0) (Brock & Freeze 1969; Matsushita et al. 1995). To infect *Thermus thermophilus* TZ2 with φIN93, the culture was incubated for 3.5 h at 70şC, with shaking at 220 rpm (until the OD_{610} was around 0.15, measured with a Hitachi U3310 spectrophotometer). φIN93 stock solution was then added to the culture to a multiplicity of infection of around 0.4. The culture was then incubated under the same conditions until complete lysis occurred (2.5 h), after which the cell debris was removed by centrifugation at 3000 *g* for 10 min. The resultant φIN93 lysate was stored at 4şC for use as the phage stock. The titer of the phage was assayed by using the overlay method of Adams with some modifications (Adams 1950; Brock & Freeze 1969).

Isolation of phage DNA

φIN93 virions were prepared as described previously (Yamato et al. 1970). DNase I and RNase A were added to the φIN93 lysate to final concentrations of 1 µg/ml each and incubated for 30 min at 30şC. Thereafter, 2.5 M NaCl solution containing 25% polyethylene glycol 8000 was added to a final concentration of 20% (v/v), and the mixture was incubated for 1 h on ice before being centrifuged at 15 000 *g* for 30 min at 4şC by using an RPR12-2 rotor in Himac CR20B3 centrifuge (Hitachi Co.). The precipitated pellets were suspended in 2 ml of 10-mM ammonium acetate buffer containing 5 mM MgSO4 (pH 6.0).

To extract the φIN93 DNA, an equal volume of phenol-chloroform (1:1, v/v) was added to the purified phage suspension, mixed, and centrifuged at 2800 *g* for 10 min at room temperature by using a T4SS rotor in Himac CT6E centrifuge (Hitachi Co.). The aqueous phase was then extracted, and the DNA was precipitated in ethanol, after which the purified DNA was dissolved in TE buffer (10 mM Tris hydrochloride and 1 mM EDTA, pH 8.0).

Construction of the plasmid vector

A plasmid containing *ORF1* was constructed in 4 steps.

The first step was designed to clone *ORF1* into pDONR221. A DNA fragment (around 1.4 kb) containing the *ORF1* and Shine-Dalgarno sequences was amplified by PCR, using Gene Amp PCR System 9700 (Applied Biosystems Co.) and a PCR kit (Toyobo Co.) with φIN93 genomic DNA as the template and synthetic DNA oligomers (Fig. 1). The amplified DNA fragment and pDONR221 were mixed with BP Clonase™ enzyme mix (Invitrogen Co.) and incubated for 1 h at 25°C, according to the manual, after which the DNA mixtures were transformed into DH5α™ *Escherichia coli* competent cells (Invitrogen Co.). The entry clone was selected on LB plates containing 30 μg/ml kanamycin, and DNA plasmid was extracted using a Plasmid Miniprep Kit (BioRad Co.). The *ORF1*-containing recombinant plasmid was called pDONR221-ORF1.

attB1-EcoRI:

5'GGGGACAAGTTTGTACAAAAAAGCAGGCTGAATTCCCCTCAACAGGAG
GCGTCATGAACTTGAAC3'
*Eco*RI

attB2-XbaI:

5'GGGGACCACTTTGTACAAGAAAGCTGGGTTCTAGACTGCCCTTCAGAA
TGTCCTCCAATGCCTC3'
*Xba*I

Fig. 1. The synthetic DNA oligomers used to amplify *ORF1* by PCR. The nucleotide sequences of attB1 and attB2 are boxed, and the *Eco*RI and *Xba*I recognition sites are underlined

In the second step, the *kat* gene, encoding a thermostable protein providing resistance to kanamycin, was cloned into pBluescriptII SK(+). The *Thermus* plasmid vector pMK18 (Biotools Co.) was digested with *Bam*HI and *Pst*I, after which a DNA fragment (around 1.2 kb, containing thermostable *kat*, its promoter, and the SD sequences) was separated by 1.0% agarose gel electrophoresis and next purified by using a DNA and Gel Band Purification Kit (Amersham Biosciences Co.). The purified DNA fragment was then ligated into pBluescriptII SK (+) digested with *Bam*HI and *Pst*I. The ligated DNA was transformed into DH5α™ *Escherichia coli* competent cells, after which those expressing the recombinant plasmid were selected on LB plates containing 50 μg/ml ampicillin, 40 μg/ml X-gal, and 0.4 mM IPTG. The *kat*-containing recombinant plasmid was called pBluescriptII-Kat.

In the third step, a destination vector was constructed from pBluescriptII-Kat by digesting the plasmid with *Eco*RV and ligating it with the Reading Frame A (RFA) cassette containing the chloramphenicol resistance and *ccd*B genes (Invitrogen Co.). The resultant DNA was transformed into the DB3.1™ *Escherichia coli* competent cells (Invitrogen Co.), and those expressing the recombinant plasmid were selected

on LB plates containing 30 μg/ml chloramphenicol, as instructed by the manual. The RFA-containing destination vector was called pBlue-Kat-RFA. The direction of the inserted RFA was determined by sequencing, using M13 primer.

In the fourth step, *ORF1* was cloned into the destination vector. For this purpose, pDONR221-ORF1 and pBlue-Kat-RFA were mixed with LR Clonase™ enzyme mix (Invitrogen Co.), incubated for 1 h at 25°C and transformed into DH5α™ *Escherichia coli* competent cells, according to the manual. Cells expressing the recombinant plasmid were selected on LB plates containing 50 μg/ml ampicillin. The recombinant plasmid containing *ORF1* and *kat* was called pIA1.

The insertion of *ORF1* and *kat* into pIA1 was confirmed by sequencing, using an ALF DNA sequencer (Amersham Pharmacia Co.) with a T7 primer: GTAATACGACTCACTATAGGGC, a Kat (f) primer: GCCAGTTCGTAATGTCTGG, and a Kat (r) primer: CCCGAGACGGGCGAGTTGCC.

Transformation of the plasmid vector into Thermus *sp*.

The following transformation protocol was based on the methods of de Grado et al. (1999) and Koyama et al. (1986). *Thermus thermophilus* HB27 was used as the host in the transformation experiment because its restriction system is not strict. *Thermus thermophilus* HB27 cells were incubated overnight in a rich medium containing 0.8% trypticase, 0.4% yeast extract, 0.3% NaCl, and Castenholz basal salts (pH 7.0) at 70°C, with shaking at 220 rpm. The overnight culture was diluted 1:100 in the fresh rich medium and incubated for about 4 h (to an OD_{610} of 0.8) at 70°C, with shaking. Thereafter, 1-ml aliquots of the culture were removed and mixed with 10 μl of the required amount of pIA1, after which this mixture was incubated for 2 h at 60şC, with shaking. The growing culture was then spread on 3% agar plates with the aforementioned rich medium containing kanamycin (30 μg/ml), and incubated for 72 h at 60°C. Finally, the plasmid DNA from *Thermus* sp. was extracted by using a Plasmid Miniprep Kit (BioRad Co.).

Primer extension experiments

Total RNA was prepared from pIA1-transformed *Thermus thermophilus* HB27 cells. The cells were harvested after incubation for 7 h at 70°C, with shaking at 220 rpm in the rich medium containing kanamycin (30 μg/ml), and RNA was extracted using an Aurum Total RNA Mini Kit (BioRad Co.). Primer extension analysis was performed using a synthetic biotinylated primer (gp1 primer: 5'-GACGCCTCCTGTTGAGGGG-3') with a Primer Extension System-AMV Reverse Transcriptase Kit (Promega Co.). Following reverse transcription, the synthesized DNA was analyzed by electrophoresis on an 8% SDS-polyacrylamide gel containing 8 M urea. As a standard, a sequence ladder was prepared with pIA1 DNA, using a Sequencing High Non-Radio Isotopic DNA Sequencing Kit (Toyobo Co.). An Imaging High Non-Radio Isotopic Detection Kit (Toyobo Co.) was next used to detect the synthesized DNA.

Construction of an ORF1-*deleted plasmid*

Plasmid pIA1 was digested with *Kpn*I at 2 restriction sites located within and just downstream of *ORF1*, respectively. The 2 resultant DNA fragments were separated by 1% agarose gel electrophoresis, after which the fragment (around 4.7 kb)

containing only part of *ORF1* was purified, using a DNA and Gel Band Purification Kit (Amersham Biosciences Co.). After self-ligation, the fragment was transformed into DH5α^{TM} *E. coli* competent cells on LB plates containing 50 μg/ml ampicillin. The pIA1-del vector, containing the partially deleted *ORF1*, was confirmed in a *Kpn*I digestion experiment.

RESULTS AND DISCUSSION

We previously compared the amino acid sequences of gp1, encoded by *ORF1* in the genome of *Thermus* phage φIN93, with those of RepA, encoded by *repA* in the *Thermus* sp. plasmid. Thanks to this, we identified a highly conserved sequence, R*PWRDLE*D*R*GRTA. Downstream of the conserved sequence was an AT-rich region (116 bp, AT content: 54.3%) like that found in *repA*, which we had predicted to be the region of replicative origin (Matsushita et al. 2009).

To clarify the function of gp1, we constructed a novel recombinant plasmid, pIA1 (5540 bp), by inserting *ORF1* and *kat* into pBluescriptII SK(+) (Fig. 2). We

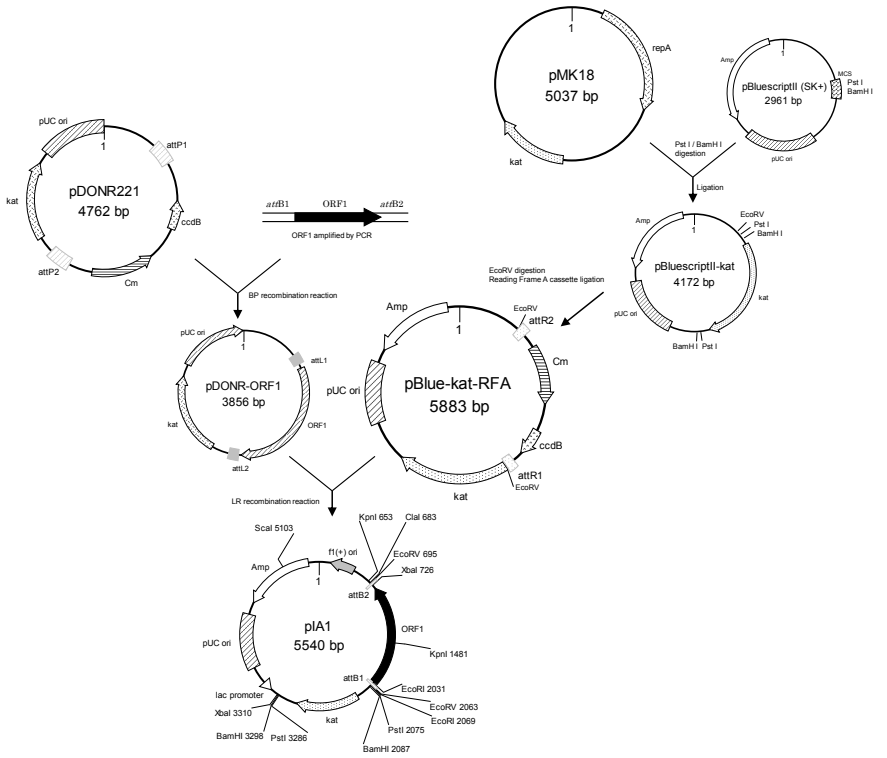


Fig. 2. The four-step construction of pIA1, containing *ORF1* from the *Thermus* bacteriophage φIN93. pIA1 was constructed by insertion of *ORF1* from φIN93 and *kat* from pMK18 into pBluescriptII SK (+). Amp and Cm indicate the genes coding for resistance to ampicillin and chloramphenicol, respectively

took this approach because the *Thermus* plasmid pMK18, which was developed by DE GRADO et al. (DE GRADO et al. 1999), contains no appropriate restriction sites through which to substitute *ORF1* from φIN93 for *repA* from pMK18. pIA1 containing ColE1 *ori* and amplified in *Escherichia coli* DH5α™ was called pIA1^{ec}. In the transformation experiment, *Thermus thermophilus* HB27 was transformed with pIA1^{ec}, which was then replicated and amplified (Table 1). The transformation efficiency of pIA1^{ec} was 2.4×10^3 cfu/μg, and the pIA1 amplified in *T. thermophilus* HB27 was called pIA1th.

Table 1. Transformation efficiencies [a] obtained by *T. thermophilus* HB27 with pIA1 isolated from *E.coli* and *T. thermophilus* HB27

	Thermus thermophilus HB27
pIA1^{ec}	2.4×10^3
pIA1th	2.2×10^6

[a] Given in transformants per microgram of pIA1

In addition, to confirm the function of *ORF1*, we constructed plasmid pIA1-del, in which a 707-bp sequence, containing the conserved sequence and the AT-rich region, was deleted from *ORF1* in pIA1^{ec} using *Kpn*I. The plasmid pIA1-del was not replicated in *T. thermophilus* HB27 (Fig. 3), making it clear that *ORF1* is crucial for replication, and that gp1 functions as a replication protein in the same manner as RepA. Thus φIN93 appears to have the same replication system as the *Thermus* sp. plasmid, one that differs from the rolling circle system of *Escherichia coli* phage λ (which contains terminal proteins and *cos* sites).

In the constructed plasmid pIA1th, we initially thought the *lac* promoter derived from pBluescriptII served as the promoter for *ORF1*. However, primer extension analysis showed that the transcription initiation site (+1) is at guanine 2094, i.e. 82

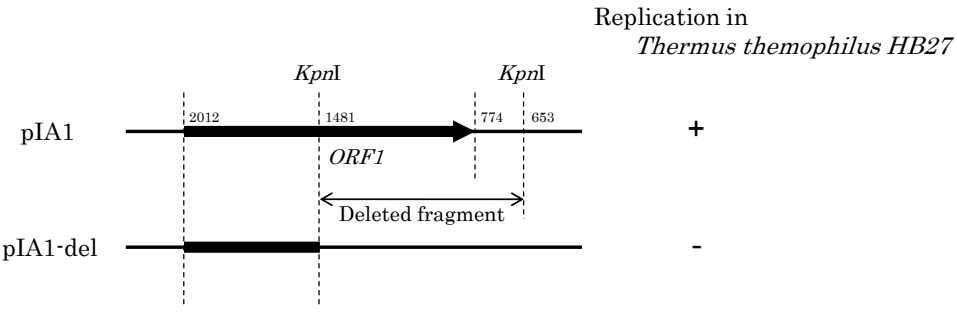
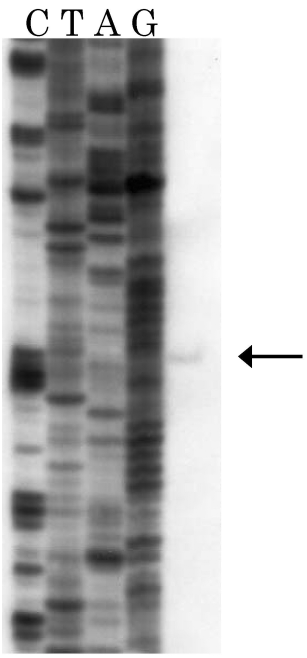


Fig. 3. Replication experiment carried out using intact pIA1 or pIA1-del. The latter contains a partially deleted *ORF1* constructed through self-ligation of *Kpn*I-digested pIA1. *ORF1* is indicated by the thick arrow. *Kpn*I-deleted *ORF1* is indicated by the thick bar

A



B

```
2300  GGTCCATTCATATGCCTCACACCTCCTTAAGGGTCGTGGGCGGGAACCCGAGACGGGCGA  2241
      CCAGGTAAGTATACGGAGTGTGGAGGAATTCCCAGCACCCGCCCTTGGGCTCTGCCCGCT
      kat gene

2240  GTTGCCGCGTTTCCTCTCCGCCCAGGTCCGCCCGGTGCGGGGAAAACCCCCCAAAAGGAG  2181
      CAACGGCGCAAAGGAGAGGCGGGTCCAGGCGGGCCACGCCCCTTTTGGGGGGTTTTCCTC

                                                  promoter of ORF1
                                                  -35 box
2180  CCCTTTTTCCCCGCATCCGGCGCTATCGTAAAAACCTCACGCGCCCTTGTCAAGCGGTCG  2121
      GGGAAAAAGGGGCGTAGGCCGCGATAGCATTTTTGGAGTGCGCGGGAACAGTTCGCCAGC
                          -10 box                 -35 box
                                promoter of kat gene
                    -10 box     +1
2120  GGCCTTAAGGTTTCTGTTATACTCCCGGGGATCCCCCGGGCTGCAGGAATTCGATATCAC  2061
      CCGGAATTCCAAAGACAATATGAGGGCCCCTAGGGGGCCCGACGTCCTTAAGCTATAGTG

                                                  ORF1
2060  AAGTTTGTACAAAAAAGCAGGCTGGAATTCCCCTCAACAGGAGGCGTCATGAACTTGAAC  2001
      TTCAAACATGTTTTTTCGTCCGACCTTAAGGGGAGTTGTCCTCCGCAGTACTTGAACTTG
                                gp1 primer
```

Fig. 4. Promoter analysis of *ORF1*. (A) The arrow indicates the estimated 5' end of the mRNA. The DNA sequence ladders shown were synthesized using the same gp-1 primer. (B) The putative -10 and -35 promoter regions are boxed, and the transcription initiation site is indicated in bold (+1). The nucleotide sequences of the ORFs are shown in bold and italics. Shine-Dalgarno sequences are underlined

nucleotides upstream from the translational start codon (ATG). Seven nucleotides farther upstream is a putative –10 box (TGTTAT), and another 22 nucleotides still farther upstream is a putative –35 box (TTGTCA). These putative promoter sequences were found in the DNA fragment containing the *kat* gene from pMK18 digested with *Bam*HI and *Pst* I. If indeed the sequence of the –35 box is TTGTCA, then the TGTCA sequences overlap the –35 box TTGACA in the promoter upstream of *kat* (Fig. 4A–B).

We repeated the transformation experiments summarized above using pIA1th isolated from *Thermus thermophilus* HB27. The transformation efficiency obtained with pIA1th was 2.2×10^{6} cfu/μg in *Thermus thermophilus* HB27, which was 1000-fold higher than the 2.4×10^{3} cfu/μg obtained with pIA1^{ec} (Table 1). Moreover, there was a linear relationship between transformation efficiency and DNA concentration over a 1000-fold range with both pIA1th and pIA1^{ec} (Fig. 5).

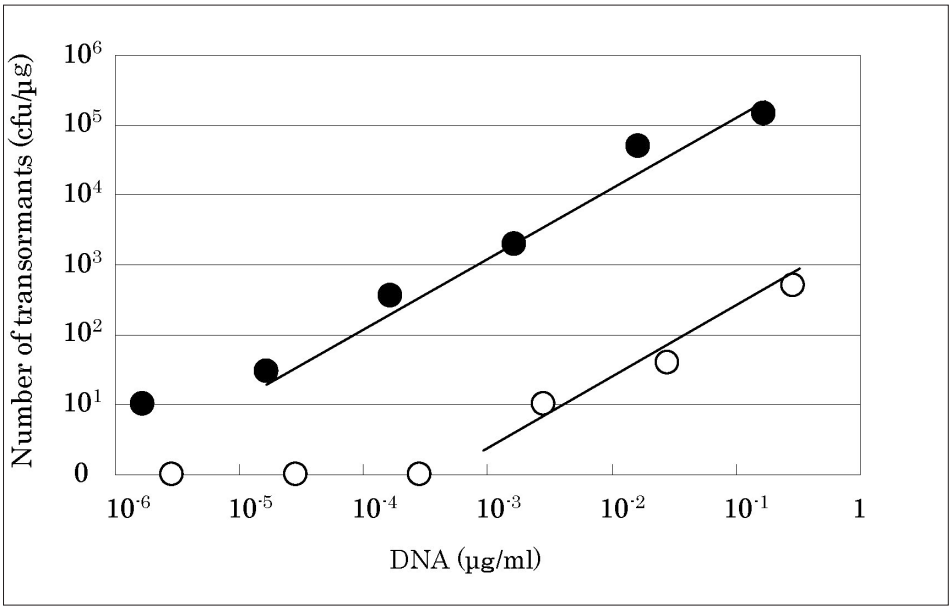


Fig. 5. Dependence of transformation on DNA concentration. The numbers of transformants obtained with pIA1th are indicated by solid circles, while those obtained with pIA1^{ec} are indicated by hollow circles

Several *Thermus-Escherichia* shuttle vectors have been developed so far (Lasa et al. 1992; Wayne & Xu 1997; de Grado et al. 1999). Among them, pMK18 is thought to have the highest transformation efficiency in *Thermus thermophilus* HB27. We found that the transformation efficiencies of pIA1th and pIA1^{ec} were both lower than the previously reported transformation efficiencies of pMK18th ($10^{8\sim9}$ cfu/

μg) and pMK18^{ec} (10^5 cfu/μg) in *Thermus thermophilus* HB27 (de Grado et al. 1999). However, since the transformation experiments with pMK18 were not carried out under the same conditions that we used with pIA1, we do not think this is a useful comparison.

Finally, pIA1 is the first *Thermus-Escherichia* shuttle vector containing a replication protein derived from a *Thermus* bacteriophage. Increasing the transformation efficiency of pIA1 and applying it to an expression vector should make it possible to take advantage of the promoters and regulatory mechanisms of bacteriophage φIN93.

Acknowledgments: We thank A. Nakata, M. Moriya, and A. Ueno for technical assistance and discussion. We are also grateful to Professor Emeritus H. Kawasaki (Osaka Prefecture University) for valuable discussion. This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

REFERENCES

Adams M. H. 1950. Methods of study of bacterial viruses. Methods Med. Res. 2: 1–73.

Brock T. D., Freeze H. 1969. *Thermus aquaticus* gen. nor. and sp. nor., a nonsporulating extreme thermophile. J. Bacteriol. 98: 289–297.

de Grado M., Castán, P., Berenguer J. 1999. A high-transformation-efficiency cloning vector for *Thermus thermophilus*. Plasmid 42: 241–245.

de Grado M., Lasa I., Berenguer J. 1998. Characterization of a plasmid replicative origin from an extreme thermophile. FEMS Microbiol. Lett. 165: 51–57.

Koyama Y., Hoshino T., Tomizuka N., Furukawa K. 1986. Genetic transformation of the extreme thermophile *Thermus thermophilus* and of other *Thermus* spp. J. Bacteriol. 166: 338–340.

Lasa I., de Grado M., de Pedro M. A., Berenguer J. 1992. Development of *Thermus-Escherichia* shuttle vectors and their use for the expression of the *Clostridium thermocellum celA* gene in *Thermus thermophilus*. J. Bacteriol. 20: 6424–6431.

Matsushita I., Yamashita N., Yokota A. 1995. Isolation and characterization of bacteriophage induced from a new isolate of *Thermus aquaticus*. Microbiol. Cult. Coll. 11: 133–138.

Matsushita I., Yanase E. 2009. The genomic structure of *Thermus* bacteriophage φIN93. J. Biochem. 146: 775–785.

Minakhin L., Berdygulova Z., Severinov K., Goel M., Florens L., Washburn M., Mushegian A., Ramanculov E., Glazko G., Karamychev V. N., Slesarev A. I., Kozyavkin S. A., Khromov I., Severinov K., Ackermann H., Mushegian A., Severionv K. 2008. Genomic comparison and proteomic characterization of *Thermus thermophilus* bacteriophages P23-45 and P74-26: Siphoviruses with triplex-forming sequences and the longest known tails. J. Mol. Biol. 378: 468–480.

Naryshkina T., Liu J., Florens L., Swanson S. K., Pavlov A. R., Pavlova N. V., Inman R., Kozyavkin S. A., Washburn M., Mushegian A., Severinov K. 2006. *Thermus thermophilus* bacteriophage φYS40 genome and proteomic characterization of virions. J. Mol. Biol. 364: 667–677.

Oshima T., Imahori K. 1974. Description of *Thermus thermophilus* (Yoshida and Oshima) comb. nov., a nonsporulating thermophilic bacterium from a Japanese thermal spa. Int. J. Syst. Bacteriol. 24: 102–112.

Wayne J., Xu S. 1997. Identification of a thermophilic plasmid origin and its cloning within a new *Thermus*- *E. coli* shuttle vector. Gene 185: 321–328.

Yamato K. R., Alberts B. M., Benzinger R., Lawwhorne L., Treiber G. 1970. Rapid bacteriophage sedimentation in the presence of polyethylene glycol and its application to large scale virus purification. Virology 40: 734–744.