J. Biochem. 121, 1031-1034 (1997)

Screening of a Mutant Plasmid with High Expression Efficiency of GC-Rich *leuB* Gene of an Extreme Thermophile, *Thermus thermophilus*, in *Escherichia coli*

Toshiharu Suzuki,*,[1] Yuji Tanaka,[†] Masami Ishida,[‡] Morio Ishizuka,[†] Akihiko Yamagishi,*,[§] and Tairo Oshima*,[§]

*Department of Life Science, Tokyo Institute of Technology, Nagatsuta, Yokohama 227; [†]Department of Industrial Chemistry, Chuo University, Bunkyo-ku, Tokyo; [‡]Laboratory of Biochemistry of Marine Resources, Tokyo University of Fisheries, Minato-ku, Tokyo 108; and [§]Department of Molecular Biology, Tokyo University of Pharmacy and Life Science, 1432-1 Horinouchi, Hachioji, Tokyo 192-03

Received for publication, November 25, 1996

A mutant plasmid with elevated expression efficiency of GC-rich *Thermus thermophilus leuB* gene was screened in *Escherichia coli*. A wild-type plasmid pHB2 carrying *T. thermophilus leuB* gene was introduced into *leuB*-deficient *E. coli* C600 cells. During successive cultures of the transformant in leucine-free medium, the original plasmid was spontaneously replaced by a mutant plasmid. The expression efficiency of the *leuB* gene on the mutant plasmid was 4.8-fold higher than that of the wild-type plasmid. Sequencing of the mutant plasmid revealed that the open reading frame (ORF1) in front of the *leuB* gene was shortened from 822 to 306 bp. Several expression vectors were constructed to investigate the effect of the length of ORF1, and the optimal length for the expression of the following *leuB* gene was determined. It was also shown that the stop codon of ORF1 should be overlapped with the initiation codon of *leuB* gene for the highest efficiency.

Key words: 2-cistron expression system, expression efficiency, GC-rich gene, 3-isopropylmalate dehydrogenase, *Thermus thermophilus*.

Thermophilic enzymes are useful for basic research as well as in industrial applications, since they are usually rather stable to chemical and physical treatments. They are also easy to purify. Up to now the genes coding for many heat-stable enzymes have been cloned from an extreme thermophile, *Thermus thermophilus*, but the expression of the thermophile genes in *Escherichia coli* is generally poor (*1-3*). We have studied the expression efficiency of *T. thermophilus* genes using *leuB* gene (coding for 3-isopropylmalate dehydrogenase, an enzyme essential for leucine biosynthesis) as a model (*4*).

The GC content (69%) of the genomic DNA of the thermophile is unusually high (*5*). It has been speculated that secondary structures of the mRNA around the ribosome binding site, owing to the high GC content, inhibit the translation and are responsible for the poor expression of the genes (*1-3, 6*). However, it was found that the thermophile *leuB* gene can be efficiently expressed in *E. coli* if another open reading frame (ORF1) exists in front of the *leuB* gene [for example, a recombinant plasmid pHB2 (*7*)]. This arrangement is often called the 2-cistron expression system (*8, 9*). The expression was drastically decreased when the ORF1 was removed from pHB2 (*4*).

In this study we attempted to improve the expression of the thermophile *leuB* gene in *E. coli* harboring pHB2 by laboratory evolution. A spontaneous mutant that gave a faster growth rate of the *leuB*-deficient *E. coli* in the leucine-free medium was isolated. The mutant plasmid contained a truncated ORF1. This observation led us to investigate the effect of the length of ORF1 in the 2-cistronic structure on the expression efficiency.

MATERIALS AND METHODS

Strains and Media—*Escherichia coli* C600 [F⁻ *leuB6, thr, thi, trp*] or JA221 [F⁻, *hsdR, trpE5, leuB6, lacY, recA1*] was cultured in either LB broth or M9 minimum medium (*10*). The medium was supplemented with ampicillin (10 μg/ml), threonine (50 μg/ml), and/or tryptophan (50 μg/ml) when required.

Screening and Sequencing of Mutant Plasmid—*Escherichia coli* C600 was transformed with the plasmid pHB2 (*7*) carrying *T. thermophilus leuB* gene. The transformant was cultivated aerobically in 5-ml of the M9 minimum medium containing ampicillin, threonine, and tryptophan in a 15-ml test tube at 37°C for 20 h. Five microliters of the saturated culture was used to inoculate the next fresh medium.

The sequence of the isolated mutant plasmid was determined by the dideoxy-chain-termination method (*11*).

Construction of 2-Cistron Expression Vectors and Assay of Expression Efficiency of the leuB Gene—Standard DNA manipulations were performed as described by Sambrook *et al.* (*10*). All expression plasmids which have 2-cistron structure were derivatives of pHB2 (see also Figs. 2 and 3). The plasmid pHB2 was digested with *Bgl*II, and the protruding ends were treated with Klenow fragment of DNA polymerase I, and ligated. The resulting plasmid was

[1] To whom correspondence should be addressed. Phone: +81-426-76-7141, Fax: +81-426-76-7145

named pHB2-Bgf. The plasmid pHB2 was digested with *Eco*RV and *Bgl*II, treated with Klenow fragment, and ligated. The resulting plasmid was named pHB2-E5fBg. A 1.1-kbp *Bam*HI fragment of pHB2 including *T. thermophilus leuB* gene was ligated with *Bam*HI-digested plasmid pBR322. The resulting plasmid with the *leuB* gene in the same direction as the *tet* promoter was named pHB2-NB. The plasmid pHB2-NB was partially digested with *Bam*HI, treated with Klenow fragment, and ligated. The plasmid in which only the *Bam*HI site in front of the *leuB* gene was cleaved by the treatment was isolated and named pHB2-Bf. The plasmid pHB2 was digested with *Eco*RV and *Bam*HI, treated with Klenow fragment, and ligated. The resulting plasmid was named pHB2-E5fB. Two fragments were amplified by polymerase chain reaction using pHB2 as the template, and two synthesized oligonucleotide-pairs (5′-GGGAATTCTCAGTTTGACAGCTTATC-3′ 5′-GACGAGGTACCCATTGTTAGATTTC-3′) and (5′-CTCCGGTACCCCAGGAGGACGGAATGA-3′ 5′-AGGCGGTGGAGCGGATTCTAGACGTCGG-3′) as primers. The former set of primers were designed to amplify a 117-bp fragment corresponding to the upper region of the *leuB* gene (see Fig. 3), and the latter set of primers, to amplify a 1.1-kbp fragment containing the whole coding region of the *leuB* gene. As a result of the amplifications, a *Kpn*I site (underlined) was newly introduced into each amplified fragment. The 117-bp fragment amplified with the former primer pair was digested with *Eco*RI (underlined) and *Kpn*I, and the 1.1-kbp fragment amplified with the later pair was digested with *Kpn*I and *Xho*I (located in the *leuB* gene, see Fig. 2). These two digested fragments were introduced into the *Eco*RI-*Xho*I site of pHB2. The resulting plasmid was named pHB2-SP. The sequence of the amplified region was verified by the dideoxy-chain-termination method (*11*). The plasmid pHB2-SP was digested with *Kpn*I, treated with T4 DNA polymerase in the presence of dNTPs, and digested with *Xho*I. The resulting 0.6-kbp *Kpn*I (blunt-ended)-*Xho*I fragment was introduced into the *Eco*RV-*Xho*I site of pHB2. The resultant plasmid was named pHB2-E5fK.

E. coli JA221 cells harboring each plasmid were cultured in LB broth containing ampicillin (50 μg/ml) to overexpress the *leuB* gene. The preparation of the crude extract of the cells and the assay of isopropylmalate dehydrogenase activity in the crude extract followed the methods described previously (*4*).

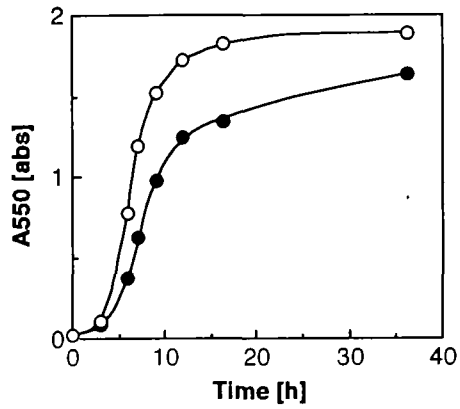

Fig. 1. **The growth curves of *E. coli* JA221 harboring the wild-type (pHB2) and mutant (pA1) plasmids.** Closed circles: *E. coli* JA221 harboring the wild-type plasmid pHB2. Open circles: *E. coli* JA221 harboring the mutant plasmid pA1. Each transformant was cultivated under the same conditions as used in the screening for the mutant plasmid, pA1. Apparent absorbance was measured at 550 nm after appropriate dilution.

Plasmid	Structure	Length of ORF1 (bp)	Enzyme activity (units/ g cells, 10^3)	Relative enzyme activity
pHB2	E H E5 B Bg B X B	822	46.4	1.00
pHB2-Bf	E H E5 (B fill-in) X B	375	152	3.28
pA1	E H B X B	306	223	4.80
pHB2-E5fB	E H (E5 fill-in B) X B	183	318	6.85
pHB2-E5fK	E H B(E5 fill-in KpnI) X B	120	292	6.29
pHB2-Sp	E H K X B	39	213	4.60
pHB2-Bgf	E H E5 B (Bg fill-in) B X B	591	34.3	0.74
pHB2-E5fBg	E H (E5 fill-in Bg) B X B	156	49.2	1.06
pHB2-NB	E H E5 B X B	760	<0.5	N. D.

Fig. 2. **The effect of ORF1 on the expression efficiency of *T. thermophilus leuB* gene.** Filled arrows, ORF1; open arrows, *T. thermophilus leuB* gene; open circles, *tet* promoter; R, 5-bp identical sequences found at both ends of the deleted region; E, *Eco*RI; H, *Hin*dIII; E5, *Eco*RV; B, *Bam*HI; Bg, *Bgl*II; X, *Xho*I. Isopropylmalate dehydrogenase activity of the crude extract was measured, and the relative activity is shown in the figure. N.D., no detectable activity.

RESULTS AND DISCUSSION

Screening and Characterization of Mutant Plasmid—The *leuB*-deficient *E. coli* strain C600 harboring plasmid pHB2 (7) was cultivated in the M9 minimum medium. The saturated culture was used for the next inoculation, and the process was repeated seven times. During the culture, the growth rate increased remarkably. A plasmid prepared from the rapidly growing strain was named pA1. Figure 1 shows the growth curves of the *leuB*-deficient *E. coli* strains (JA221) harboring pHB2 and pA1 in the M9 minimum medium. *E. coli* JA221 harboring pA1 grew faster than the strain harboring pHB2.

Crude extract was prepared from the *E. coli* cells and the isopropylmalate dehydrogenase activity of the extract was measured. *E. coli* JA221 harboring pA1 showed 4.8-fold higher activity of the enzyme than that harboring pHB2 (Fig. 2).

We have previously shown that an open reading frame (ORF1) in front of the *leuB* gene of pHB2 is essential for the efficient expression of the thermophile enzyme (4). The sequencing of pA1 revealed a 561-bp deletion within the ORF1 (Fig. 3). The ORF1 was truncated from 822 to 366 bp without causing a frame shift. Identical 5-bp sequences (5′-CACCG-3′) were found at both ends of the deleted region, and may have played an important role in the spontaneous deletion. No other base replacement was found in the sequenced region (1.5-kbp *Eco*RI-*Bam*HI fragment). The result suggests that the shortening of the ORF1 is responsible for the enhanced expression of the *leuB* gene.

Effect of 2-Cistronic Structure on Expression of the leuB Gene—Several expression vectors were constructed to investigate the effect of the length of ORF1 on the expression efficiency of the following *leuB* gene, and the results are listed in Fig. 2. In general, the expression efficiency was higher when ORF1 was shorter (for instance, pHB2, -Bf, pA1, and pHB2-E5fB). However, the efficiency decreased when ORF1 was shorter than 120 bp (pHB2-E5fK and

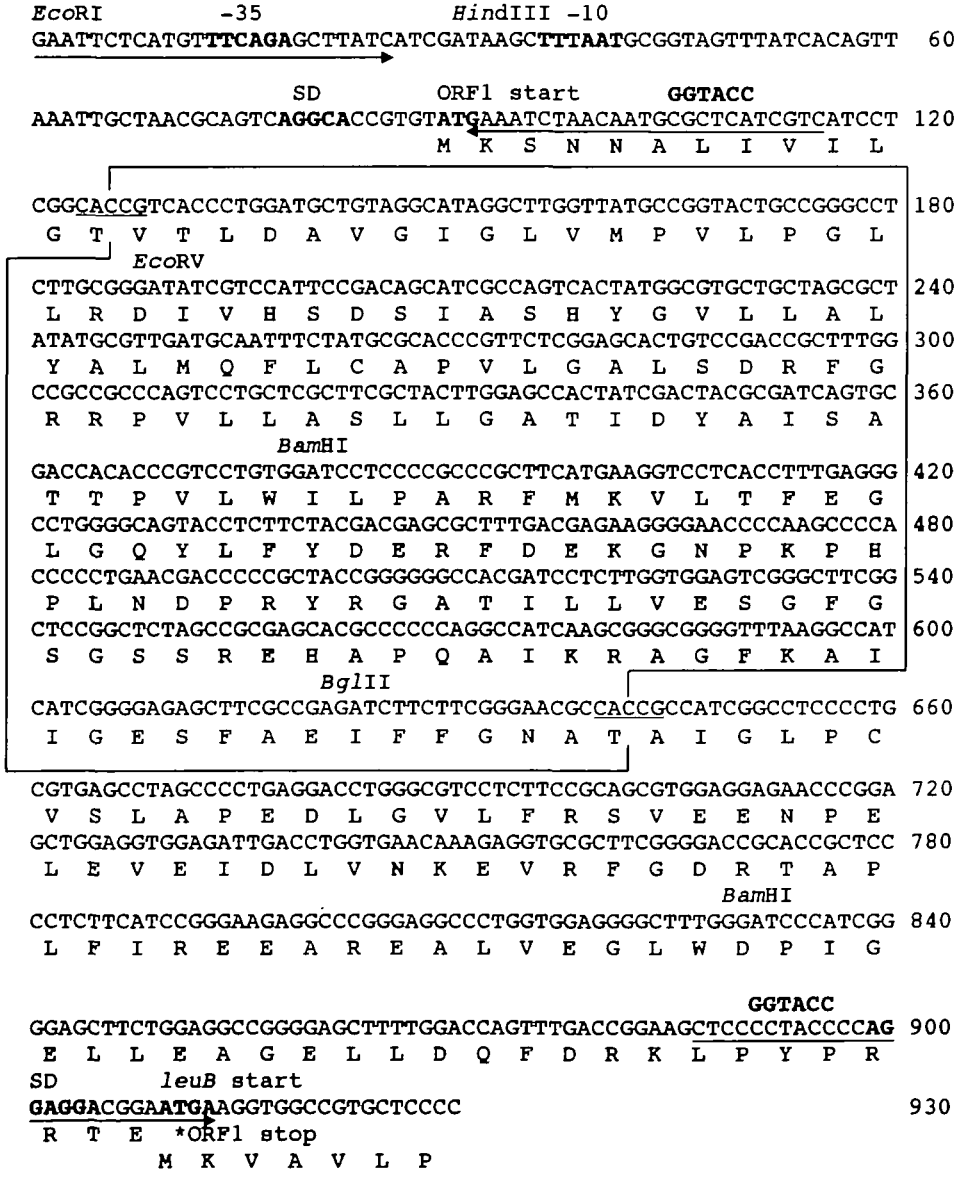

Fig. 3. **The sequence of the plasmid pHB2 around the ORF1.** The boxed region indicates bases which were lost in the mutant plasmid pA1. Underlined bases indicate a 5-bp identical sequence, 5′-CACCG-3′, found at both ends of the deleted sequence. Transcription signals (*tet* promoter) appear at −10 and −35. The sequences complementary to those of the oligonucleotide primers for polymerase chain reaction amplification are indicated with arrows (see "MATERIALS AND METHODS"). Two *Kpn*I recognition sequences (GGTACC) introduced by the amplifications are indicated with bold letters above the nucleotide sequence.

-SP). The results indicate that there is an optimum length of ORF1 for efficient expression. The optimum length may depend on the sequence of ORF1. The effects of the base sequence on the optimum length will be the subject of future studies.

The position of the stop codon of ORF1 is another factor affecting the expression. Most of the plasmids tested in Fig. 2 have a 4-bp overlap between ORF1 and the *leuB* gene (pHB2, -Bf, pA1, pHB2-E5fB, -K5fK, and -SP). Expression of the *leuB* gene was decreased when the stop codon of the ORF1 was placed 231-bp upstream of the *leuB* gene (pHB2-Bgf and -E5fBg). No expression of the gene was detected when the stop codon of ORF1 was placed 510-bp downstream of the start codon of the *leuB* gene (pHB2-NB). These results indicate that the stop codon of ORF1 should be close to the start codon of the second gene for efficient expression, as reported (*6, 8*).

This tactic, that is, the addition of a short leader ORF in front of or overlapped with the target ORF, seems to be effective for the efficient expression of GC-rich genes in *E. coli*. Though the possibility that transcription was improved by the mutation can not be ruled out, the experimental data suggest that the expression was increased by raising the translation efficiency of the GC-rich gene from the thermophile.

REFERENCES

1. Nureki, O., Muramatsu, T., Suzuki, K., Kohda, D., Matsuzawa, H., Ohta, T., Miyazawa, T., and Yokoyama, S. (1991) Methionyl-tRNA synthetase gene from an extreme thermophile, *Thermus thermophilus* HB8. *J. Biol. Chem.* **266**, 3268-3277
2. Nureki, O., Suzuki, K., Hara-Yokoyame, M., Kohno, T., Matsuzawa, H., Ohta, T., Shimizu, T., Morikawa, K., Miyazawa, T., and Yokoyama, S. (1992) Glutamyl-tRNA synthetase from *Thermus thermophilus* HB8: molecular cloning of the *gltX* gene and crystallization of the overproduced protein. *Eur. J. Biochem.* **204**, 465-472
3. Kanaya, S. and Itaya, M. (1992) Expression, purification, and characterization of a recombinant ribonuclease H from *Thermus thermophilus* HB8. *J. Biol. Chem.* **267**, 10184-10192
4. Ishida, M. and Oshima, T. (1994) Overexpression of genes of an extreme thermophile, *Thermus thermophilus*, in *Escherichia coli* cell. *J. Bacteriol.* **176**, 2767-2770
5. Oshima, T. and Imahori, K. (1974) Description of *Thermus thermophilus* (Yoshida and Oshima) comb. nov., a nonsporulation thermophilic bacterium from a Japanese thermal spa. *Init. J. Syst. Bacteriol.* **24**, 102-112
6. Nishiyama, M., Horinouchi, S., and Beppu, T. (1991) Characterization of an operon encoding succinyl-CoA synthetase and malate dehydrogenase from *Thermus flavus* AT-62 and its expression in *Escherichia coli*. *Mol. Gen. Genet.* **226**, 1-9
7. Tanaka, T., Kawano, N., and Oshima, T. (1981) Cloning of 3-isopropylmalate dehydrogenase gene of an extreme thermophilic and partial purification of the gene product. *J. Biochem.* **89**, 677-682
8. Schoner, B.E., Belagaje, R.M., and Scöner, R.G. (1986) Translation of a synthetic two-cistron mRNA in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **83**, 8506-8510
9. Makoff, A.J. and Smallwood, A.E. (1990) The use of two-cistron constructions in improving the expression of a heterologous gene in *E. coli*. *Nucleic Acids Res.* **18**, 1711-1718
10. Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
11. Sanger, F., Nicklen, S., and Coulson, A.R. (1977) DNA sequencing with chain-termination inhibitors. *Proc. Natl. Acad. Sci. USA* **74**, 5463-5467