Vol. 184, No. 1, 1992 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
April 15, 1992 Pages 277-282

# MOLECULAR CLONING OF A SUBTILISIN J GENE FROM *Bacillus stearothermophilus* AND ITS EXPRESSION IN *Bacillus subtilis*+

Jeong Su Jang, Dae Ook Kang, Moon Jin Chun* and Si Myung Byun#

Department of Life Science,
Korea Advanced Institute of Science and Technology,
373-1 Kusong-Dong, Yusung-Gu, Taejon 305-701, Korea

Received January 5, 1992

**SUMMARY**: The structural gene for a subtilisin J from *Bacillus stearothermophilus* NCIMB10278 was cloned in *Bacillus subtilis* using pZ124 as a vector, and its nucleotide sequence was determined. The nucleotide sequence revealed only one large open reading frame, composed of 1,143 base pairs and 381 amino acid residues. A Shine-Dalgarno sequence was found 8 bp upstream from the translation start site (GTG). The deduced amino acid sequence revealed an N-terminal signal peptide and pro-peptide of 106 residues followed by the mature protein comprised of 275 residues. The productivity of subtilisin in the culture broth of the *Bacillus subtilis* was about 46-fold higher than that of the *Bacillus stearothermophilus*. The amino acid sequence of the extracellular alkaline protease subtilisin J is highly homologous to that of subtilisin E and it shows 69% identity with subtilisin Carlsberg, 89% with subtilisin BPN´ and 70% with subtilisin DY. Some properties of the subtilisin J that had been purified from the *Bacillus subtilis* were examined. The subtilisin J has alkaline pH characteristics and a molecular weight of 27,500. It retains about 50% of its activity even after treatment at 60°C for 30 min in the presence of 2 mM calcium chloride.


Many species of the genus *Bacillus* have been reported to produce two major extracellular proteases, a neutral (metallo-) protease and an alkaline serine protease (subtilisin). Both are produced after the exponential growth phase, when the culture enters the stationary phase and begins the process of sporulation (1, 2). Especially, subtilisins have been studied not only with the aim of industrial production but also for the elucidation of the catalytic mechanism in serine proteases (3, 4). The complete amino acid sequences of sutilisins Carlsberg (5), BPN´ (6), Amylosacchariticus (7) and DY (8) have been reported. In recent years, the complete amino acid sequences of thermitase (9), proteinase K (10), aqualysin I (11) and alkaline elastase YaB (12) have also been reported. We are interested in the thermophilic proteases for their contribution to a basic understanding of the molecular basis of thermostability. The ability of some microorganisms to grow at extreme temperatures (13)

+Sequence data from this article have been deposited with the EMBL/GenBank Data Libraries under Accession No. M64743.

# To whom correspondence should be addressed.

Abbreviations: bp, base pairs; kb, kilobase pairs; SD, Shine-Dalgarno sequence; TBAB, tryptose blood agar base.

0006-291X/92 $1.50

implies that their enzymes are stable and active at these temperatures. In general, enzymes from thermophilic organisms are more thermostable than the equivalents isolated from the phylogenetically related mesophilic organisms (14, 15). In this paper, we describe the cloning and expression in *Bacillus subtilis* of *aprJ*, the gene for the alkaline protease from *Bacillus stearothermophilus* NCIMB10278, and compare the amino acid sequence deduced from the nucleotide sequence with those of other subtilisins.

## MATERIALS AND METHODS

**Materials:** [$\alpha$-$^{32}$P]dATP (110 TBq/mmole) was purchased from Amersham International. Restriction enzymes, T4 DNA ligase, and *E. coli* DNA polymerase I were obtained from New England Biolabs; *N*-succinyl-L-Ala-L-Ala-L-Pro-L-Phe-*p*-nitroanilide, from Sigma; and calf intestinal alkaline phosphatase, from Boehringer Mannheim.
**Bacterial Strains and Plasmids:** *B. stearothermophilus* NCIMB10278 was used as a DNA donor (KCTC1823, Korean Collection for Type Cultures, Genetic Engineering Research Institute, Korea Institute of Science and Technology, Korea). *B. subtilis* DB104 (16) was a gift from Dr. R. H. Doi. *Escherichia coli* MC1061 and JM109 were used as host strains for cloning. Plasmids used were pZ124 [Km$^r$ (shuttle vector for *E. coli* and *B. subtilis*)] for *B. subtilis* and pUC18 for *E. coli*. Plasmid pZA118 (Km$^r$) containing the subtilisin E structural gene from pKWZ (gift from Dr. R. H. Doi) was constructed in our laboratory.
**Isolation of DNA and Transformation of Plasmid DNA:** Chromosomal DNA of *B. stearothermophilus* was prepared as previously described by Saito and Miura (17). Transformation of *B. subtilis* and *E. coli* was done by the competent cell method (18, 19).
**Cloning of Subtilisin J gene:** Southern hybridization and colony hybridization were carried out with the 1.8-kilobase (kb) *aprE* (*Eco*RI-*Pst*I digested pZA118) probe labeled by nick translation (19).
**Purification of Subtilisin J:** *B. subtilis* DB104(pZS101) was cultivated in 2 liters of the 2X TY medium containing kanamycin (5 $\mu$g/ml) at 37°C for 36 h with vigorous aeration. The culture broth was centrifuged (6,000X*g*, 30 min, 4°C), and three volumes of ethanol (-20°C) was added to the supernatant. After centrifugation (7,000 X *g*, 30 min, 4°C), the precipitate was resuspended in 20 ml of the 10 mM sodium phosphate buffer (pH 6.2) and dialyzed against 4 liters of the same buffer solution overnight at 4°C. The solution after dialysis was desalted by using a Sephadex G-25 column (2.5 X 45 cm). The active fractions were applied to a CM-cellulose column (3.2 X 22 cm), and the enzyme was eluted with the same buffer plus 0.08 M NaCl. The subtilisin peak was identified by the enzyme activity and shown to be >95% pure by 17.5% SDS-polyacrylamide gel electrophoresis. The enzyme was assayed in a solution containing 0.3 mM *N*-succinyl-L-Ala-L-Ala-L-Pro-L-Phe-*p*-nitroanilide and 0.1 M Tris-HCl (pH 8.6) at 25°C. The assays measured the increase in absorbance at 410 nm per min due to hydrolysis and release of the *p*-nitroaniline ($\epsilon_{410}$ = 8,480 $M^{-1}cm^{-1}$) (20).

## RESULTS AND DISCUSSION

Southern blot analysis of *B. stearothermophilus* NCIMB10278 chromosomal DNA digested with *Cla*I and *Eco*RI showed that a 3.1-kb *Cla*I fragment and a 4.8-kb *Eco*RI fragment hybridize with the labeled *aprE* gene probe (Fig. 1). Since the hybridizing *Cla*I fragment (3.1 kb) is sufficiently larger than the *aprE* gene (1.8 kb) which is supposedly homologous to the *aprJ* gene, the *Cla*I fragment was electroeluted from the agarose gel and inserted into the *Acc*I site of pUC18. After *E. coli* JM109 was transformed with this ligation mixture, one out of approximately 1 X $10^3$ colonies showed hybridization to the nick-translated probe. Plasmid purified from this hybridizing colony was designated as pUS111. For expression of this cloned gene in *B. subtilis*, plasmid pUS111 was digested with *Eco*RI and *Bam*HI and then ligated with *Eco*RI-*Bam*HI digested pZ124. The resulting recombinant

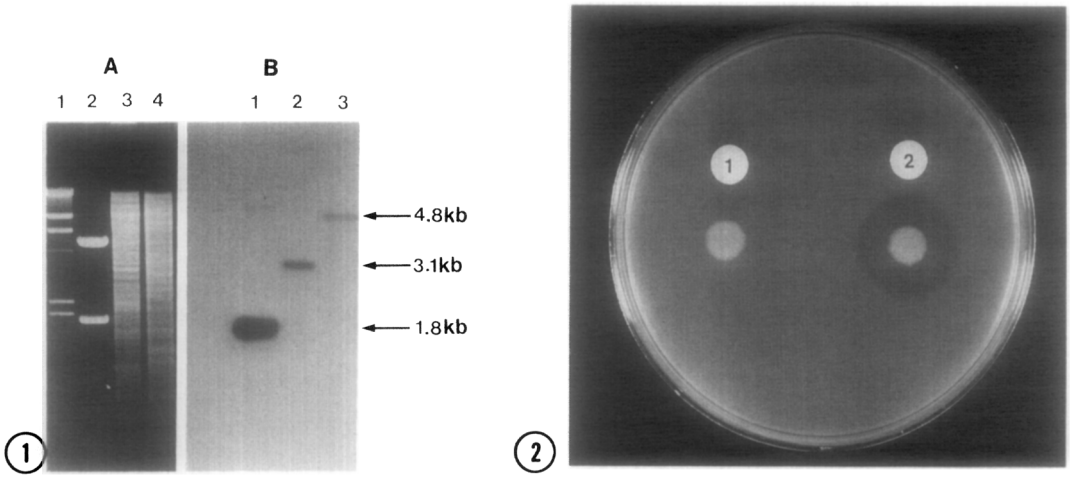


Fig. 1. Electrophoretic separation of the restriction enzyme digests of chromosomal DNA from *B. stearothermophilus* (A) and their southern blot analyses (B) with a probe. Numbers 1-4 in A refer to *Hin*dIII-digested λ DNA, *Eco*RI-*Pst*I digested pZA118, *Cla*I-digested chromosomal DNA, *Eco*RI-digested chromosomal DNA, respectively. Numbers 1-3 in B refer to *Eco*RI-*Pst*I digested pZA118, *Cla*I-digested chromosomal DNA, *Eco*RI-digested chromosomal DNA, respectively.

Fig. 2. Protease assay on a TBAB-1% skim milk plate (Km; final 5 μg/ml). Number 1 refers to *B. subtilis* DB104(pZ124) and 2 refers to *B. subtilis* DB104(pZS101).

plasmid was isolated from *E. coli* MC1061 and designated as pZS101. This plasmid was introduced into *B. subtilis* DB104, a protease-deficient strain, and protease-overproducing colonies were screened by plating on tryptose blood agar base (TBAB) plus 1% skim milk (Fig.2). The culture supernatant of *B. subtilis* DB104(pZS101) was assayed for secreted alkaline protease activity. The protease was completely inhibited by phenylmethylsulfonylfluoride, but not by ethylenediaminetetraacetic acid, indicating that the enzyme is an alkaline protease. In addition, when the 3.1-kb *Bam*HI-*Pst*I fragment from pUS111 was introduced into *B. subtilis* DB104, protease activity showed no detectable change. These results imply that the alkaline protease gene (*aprJ*) from *B. stearothermophilus* NCIMB10278 was cloned in *B. subtilis*.

The sequencing strategy and a restriction map of the 2.5-kb *Eco*RI-*Bam*HI fragment is shown in Fig. 3. Its restriction enzyme sites are almost identical to those of the *aprE* gene. The complete nucleotide sequence of the *aprJ* and its flanking region are shown in Fig. 4. The sequence contains only one large open reading frame of 1,143 nucleotides (381 amino acid residues). A potential Shine-Dalgarno (SD) sequence (GAGAGG) is found 8 bp upstream from a probable translation start codon (GTG). A putative promoter sequence (-35 region: TCTACT; -10 region: TACAAT), which is identical to the known promoter for the subtilisin E gene, *aprE*, is also found 52 bp upstream of the SD sequence (21, 22). The amino acid sequence deduced from the nucleotide sequence is shown in Fig. 4. From the comparison with other subtilisin genes, the sequence of the protein-coding region of subtilisin J was identified to code for a 29 - residue signal peptide and a 77 - residue pro-peptide, which precede the 275 residues of mature subtilisin. A typical transcription terminator sequence is

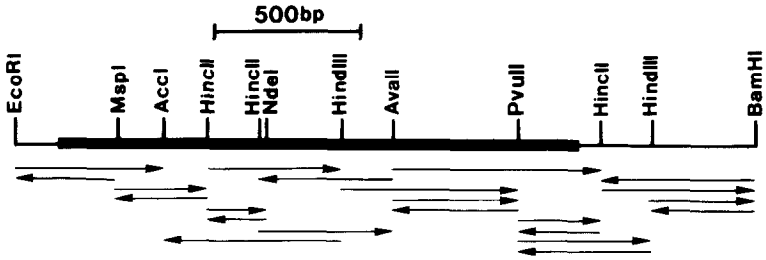


Fig. 3. Restriction map and sequencing strategy for the subtilisin J gene. The arrows indicate the starting point and the direction of individual sequencing. The black region indicates the subtilisin J coding gene (*aprJ*).

```
GAATTCTCCATTTTCTTCTGCTATCAAAATAACAGACTCGTGATTTTCCAAACGAGCTTTCAAAAAAGCCTCTGCC -533
CCTTGCAAATCGGATGCCTGTCTATAAAATTCCCGATATTGGTTAAACAACGGCGCAATGGCGGCCGCATCTGATG -457
TCTTTGCTTGGCGAATGTTCATCTTATTTCTTCCTCCCTCTCAATAATTTTTTCATTCTATCCCTTTTCTGTAAAG -381
TTTATTTTTCAGAATACTTTTATCATCATGCTTTGAAAAAATATCACGATAATATCCATTGTTCTCACGGAAGCAC -305
ACGCAGGTCATTTGAACGAATTTTTTCGACAGAAATTTGCCGGGACTCAGGAGCATTTAACCTAAAAAAGCATGAC -229
ATTTCAGCATAATGAACATTTACTCATGTCTATTTTCGTTCTTTTCTGTATGAAAATAGTTATTTCGAGTCTCTAC -153
GGAAATAGCGAGAGATGATATACCTAAATAGAGATAAAATCATCTCAAAAAAATGGGTCTACTAAAATATTATTCC  -77
                                                          -35
ATCTATTACAATAAATTCACAGAATAGTCTTTTAAGTAAGTCTACTCTGAATTTTTTTAAAAGGAGAGGGTAAAGA   -1
      -10                                                  SD
GTGAGAAGCAAAAAATTGTGGATCAGCTTGTTGTTTGCGTTAACGTTAATCTTTACGATGGCGTTCAGCAACATG    75
 M  R  S  K  K  L  W  I  S  L  L  F  A  L  T  L  I  F  T  M  A  F  S  N  M   25
TCTGTGCAGGCTGCCGGAAAAAGCAGTACAGAAAAGAAATACATTGTCGGATTTAAACAGACAATGAGTGCCATG   150
 S  V  Q  A  A  G  K  S  S  T  E  K  K  Y  I  V  G  F  K  Q  T  M  S  A  M   50
AGTTCCGCCAAGAAAAAGGATGTTATTTCTGAAAAAGGCGGAAAGGTTCAAAAGCAATTTAAGTATGTTAACGCG   225
 S  S  A  K  K  K  D  V  I  S  E  K  G  G  K  V  Q  K  Q  F  K  Y  V  N  A   75
GCCGCAGCAACATTGCATGAAAAAGCTGTAAAAGAATTGAAAAAAGATCCGAGCGTTGCATATGTGGAAGAAGAT   300
 A  A  A  T  L  D  E  K  A  V  K  E  L  K  K  D  P  S  V  A  Y  V  E  E  D  100
CATATTGCACATGAATATGCGCAATCTGTTCCTTATGGCATTTCTCAAATTAAAGCGCCGGCTCTTCACTCTCAA   375
 H  I  A  H  E  Y  A  Q  S  V  P  Y  G  I  S  Q  I  K  A  P  A  L  H  S  Q  125
GGCTACACAGGCTCTAACGTAAAAGTAGCTGTTATCGACAGCGGAATTGACTCTTCTCATCCTGACTTAAACGTC   450
 G  Y  T  G  S  N  V  K  V  A  V  I  D  S  G  I  D  S  S  H  P  D  L  N  V  150
AGAGGCGGAGCAAGCTTCGTACCTTCTGAAACAAACCCATACCAGGACGGCAGTTCTCACGGTACGCATGTAGCC   525
 R  G  G  A  S  F  V  P  S  E  T  N  P  Y  Q  D  G  S  S  H  G  T  H  V  A  175
GGTACGATTGCCGCTCTTAATAACTCAATCGGTGTTCTGGGCGTCAGCCCAAGCGCATCATTATATGCAGTAAAA   600
 G  T  I  A  A  L  N  N  S  I  G  V  L  G  V  S  P  S  A  S  L  Y  A  V  K  200
GTGCTTGATTCAACAGGAAGCGGCCAATATAGCTGGATTATTAACGGCATTGAATGGGCCATTTCCAACAATATG   675
 V  L  D  S  T  G  S  G  Q  Y  S  W  I  I  N  G  I  E  W  A  I  S  N  N  M  225
GATGTTATTAACATGAGCCTTGGCGGACCTTCTGGTTCTACAGCGCTGAAAACAGTCGTTGATAAAGCCGTTTCC   750
 D  V  I  N  M  S  L  G  G  P  S  G  S  T  A  L  K  T  V  V  D  K  A  V  S  250
AGCGGTATCGTCGTTGCTGCCGCTGCCGGAAACGAAGGTTCGTCCGGAAGCTCAAGCACAGTCGGCTACCCTGCA   825
 S  G  I  V  V  A  A  A  A  G  N  E  G  S  S  G  S  S  S  T  V  G  Y  P  A  275
AAATATCCTTCTACTATTGCGGTAGGTGCGGTAAACAGCAGCAACCAAAGAGCTTCATTCTCAAGCGCAGGTTCT   900
 K  Y  P  S  T  I  A  V  G  A  V  N  S  S  N  Q  R  A  S  F  S  S  A  G  S  300
GAGCTTGATGTGATGGCTCCTGGCGTATCCATCCAAAGCACACTTCCTGGAGGCACTTACGGTGCTTACAACGGC   975
 E  L  D  V  M  A  P  G  V  S  I  Q  S  T  L  P  G  G  T  Y  G  A  Y  N  G  325
ACGTCCATGGCGACTCCTCACGTTGCCGGAGCAGCAGCGCTAATTCTTTCTAAGCACCCGACTTGGACAAACGCG  1050
 T  S  M  A  T  P  H  V  A  G  A  A  A  L  I  L  S  K  H  P  T  W  T  N  A  350
CAAGTCCGTGATCGTTTAGAAAGCACTGCAACATATCTTGGAAACTCTTTCTACTATGGAAAAGGGTTAATCAAC  1125
 Q  V  R  D  R  L  E  S  T  A  T  Y  L  G  N  S  F  Y  Y  G  K  G  L  I  N  375
GTACAAGCAGCTGCACAATAATAGTAAAAAGAAGCAGGTTCCTCCATACCTGCTTCTTTTTATTTGTCAGCATCC  1200
 V  Q  A  A  A  Q  *  *  *                                                    381
TGATGTTCCGGCGCATTCTCTTCTTTCTCCGTATGTTGAATCCGTTCCATGATCGAAGGATGGCTGCCTCTGAAA  1275
ATCTTCACAAGCACCGGAGGATCAACCTGGCTCAGCCCCGTCACGGCCAAATCCTGAAACGTTTTAACAGCGGCT  1350
TCTCTGTTCTCTGTCAACTCGATCCCATACTGGTCAGCCTTATTCTCCTGATAACGCGAGACAGCATTAGAAAAA  1425
GGCGTCACCGCAAAGCTCAAAACAGAAAACAAAAGCAATAACAGCGGAAGTGCCGCAAGATCATGCCGCCCTTCT  1500
AAATGAAACATGCTGCGGGTTAGGCGAACCGTCCGCTTGTAAAGCTTATCAATGACATAAAATCCGGCGAGCGAC  1575
ACGAGCAAATAACCAGCCAGACCGATGTAAACGTGCTTCATGACATAATGGCCCATTTCGTGGCCCATAATAAAC  1650
AGGATTTCGTGAATCGTCAAGTTTGTTCAGCGTCGTATCCCACAATACAATCCGTTTATTGGCCCCAATTCCTGTA  1725
ACATAGGCATTCAGCGCATTTGTTTTTTCTGACATGTTCACTTCATATACATGGTCAGCCGGAATATTGGCTTCA  1800
TCTGCCAGCTCTAAAATTTTGCTTTCAAGCTCTTTGTTTTTCAGCGGATAAAAATCATTGTATAAAGGATCGACT  1875
CTAGAGGATCC                                                                   1886
```

Fig. 4. Nucleotide sequence of *aprJ*. The amino acid sequence is shown below the nucleotide sequence. The first amino acid of translation (Met) is numbered 1. A probable SD sequence and the putative promoter (-35 and -10 regions) are shown by solid lines below the nucleotide sequence. The probable transcription terminator is indicated by inverted arrows. Asterisks indicate stop codons.

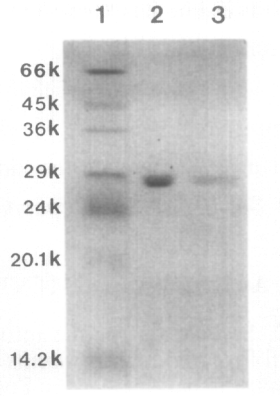


Fig. 5. SDS-polyacrylamide gel electrophoresis of purified subtilisin E (lane 2) and subtilisin J (lane 3). Molecular mass standards (in kilodaltons) are in lane 1.

also found immediately downstream from the open reading frame. As shown in Fig. 4, the deduced amino acid sequence is very similar to that of the subtilisin E coding gene *aprE* (used as a hybridization probe in this experiment). The amino acid sequence of the subtilisin J was compared with the published sequences of the other subtilisins by sequence alignment. The sequence of the extracellular mature form of this subtilisin J is identical to that of the subtilisin E from *B. subtilis*, except for two amino acid substitutions (Thr-130 to Ser-130 and Thr-162 to Ser-162). Also, there is one amino acid substitution in the signal peptide and there are two amino acid substitutions in the pro-peptide (see Fig. 4). However, the amino acid sequence of subtilisin J is completely identical to that of subtilisin Amylosacchariticus. In comparison with other subtilisins, subtilisin J shows 69% identity with subtilisin Carlsberg, 89% with subtilisin BPN´ and 70% with subtilisin DY. The amino acid sequence alignment showed that the probable active-site residues of subtilisin J, Asp-32, His-64 and Ser-221, are exactly identical to those of other subtilisins from *Bacilli*.

The proteolytic activities found in the culture broths were compared after aerobic cultivation of *B. subtilis* DB104(pZS101) and *B. stearothermophilus* NCIMB10278 for 24 h in TBAB. The culture supernatant of the *B. subtilis* showed the activity 46-fold higher than that of the *B. stearothermophilus*. The productivity of subtilisin J in the culture broth of the *B. subtilis*, however, was about 80% compared to that of subtilisin E. This difference in the enzyme productivity might be explained as follows: Since the signal sequence and the following pro-region of the subtilisin J are not the same as those of the subtilisin E, the secretion efficiencies of the two enzymes might be different.

Subtilisin J was purified from the culture supernatant of the cloned *B. subtilis*. The molecular weight of the enzyme was determined to be 27,500 (Fig. 5). Subtilisin J was most active toward *N*-succinyl-L-Ala-L-Ala-L-Pro-L-Phe-*p*-nitroanilide at pH 9.0. $Ca^{2+}$ ions stabilized the enzyme against heat treatment. When the enzyme was heated at 60°C for 30 min, the remaining activity of subtilisin J was about 50%. Even after the treatment at 60°C for 1 h, 30% of the activity of subtilisin J still remained. The enzymic properties of subtilisin

J were similar to those of subtilisin E, but the thermostability of protease subtilisin J is higher than that of subtilisin Amylosacchariticus (23).

From the information given above, we propose that the alkaline protease from *Bacillus stearothermophilus* is a subtilisin. Subtilisin J is the first protein cloned from *Bacillus stearothermophilus*. For further investigations on the kinetic properties of the enzyme, the purification of subtilisin J from *Bacillus stearothermophilus* is now in progress.

## ACKNOWLEDGMENTS

We are grateful to Miss C. S. Pearson for the information on *B. stearothermophilus* NCIMB10278 and Dr. R. H. Doi for providing *B. subtilis* DB104 and pKWZ. This work was carried out with the finantial support in part by Genetic Engineering Research Funds of Korea Ministry of Education.

## REFERENCES

1. Hoch, J.A. (1976) Adv. Genet. 18, 69-98.
2. Piggot, P.J., and Coote, J.G. (1976) Bacteriol. Rev. 40, 908-962.
3. Mitsui, T., Satow, Y., Watanabe, Y., Hirono, S., and Iitaka, Y. (1979) Nature. 227, 447-452.
4. Warshel, A., Naray-Szabo, G., Sussman, F., and Hwang, J.K. (1989) Biochemistry. 28, 3629-3637.
5. Jacobs, M., Eliasson, M., Uhlen, M., and Flock, J.I (1985) Nucleic Acids Res.13, 8913-8926.
6. Wells, J.A., Ferarri, E., Henner, D.J., Estell, D.A., and Chen, E.Y. (1983) Nucleic Acids Res. 11, 7911-7925.
7. Yoshimoto, T., Oyama, H., Honda, T., Tone, H., Takeshita, T., Kamiyama, T., and Tsuru, D. (1988) J. Biochem. 103, 1060-1065.
8. Nedkov, P., Oberthur, W., and Braunitzer, G. (1985) Biol.Chem. Hoppe-Seyler 366, 421-430.
9. Meloun, B., Bandys, M., Kostka, V., Hausdorf, G., Frömmel, C., and Höhne, W.E. (1985) FEBS Lett. 183, 195-200.
10. Jany, K.D., Lederer, G., and Mayer, B. (1986) FEBS Lett. 199, 139-144.
11. Kwon, S.T., Terada, I., Matsuzawa, H., and Ohta, T. (1988) Eur. J. Biochem. 173, 491-497.
12. Kaneko, R., Koyama, N., Tsai, Y.C., Juang, R.Y., Yoda, K., and Yamasaki, M. (1989) J. Bacteriol. 171, 5232-5236.
13. Brock, T.D. (1978) Thermophilic Microorganisms and Life at High Temperatures (Springer, New York).
14. Zuber, H. (1979) in Strategies of Microbial Life in Extreme Environments, ed. Shilo, M. (Verlag Chem, New York), pp. 393-415.
15. Ohta, Y. (1967) J. Biol. Chem. 242, 509-515.
16. Kawamura, F., and Doi, R.H. (1984) J. Bacteriol. 160, 442-444.
17. Saito, H., and Miura, K. (1963) Biochim. Biophys. Acta. 72, 619-629.
18. Imanaka, T., Fujii, M., and Aiba, S. (1981) J. Bacteriol. 146, 1091-1097.
19. Maniatis, T., Fritsch, E.F., and Sambrook, J. (1982) Molecular Cloning: a Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
20. DelMar, E.G., Largman, C., Brodrick, J.W., and Goekas, M.C. (1979) Anal. Biochem. 99, 316-320.
21. Ferrari, E., Henner, D.J., Perego, M., and Hoch, J.A. (1988) J. Bacteriol. 170, 289-295.
22. Park, S.S., Wong, S.L., Wang, L.F., and Doi, R.H. (1989) J. Bacteriol. 171, 2657-2665.
23. Tsuru, D., Kira, H., Yamamoto, T., and Fukumoto, J (1966) Agr. Biol. Chem. 30, 1261-1268.