Cloning and Nucleotide Sequences of the Bacillus stearothermophilus Neutral Protease Gene and Its Transcriptional Activator Gene

YOSHIKI NISHIYA AND TADAYUKI IMANAKA

Biomedical Products Division, Enzyme Factory, Toyoobo Co., Ltd. Tsuruga-shi, Fukui Prefecture 914, and Department of Fermentation Technology, Faculty of Engineering, Osaka University, Yamadaoka, Suita-shi, Osaka 565, Japan

Received 18 January 1990/Accepted 13 June 1990

Both the neutral protease gene (nprS) and its transcriptional activator gene (nprA) from Bacillus stearothermophilus TELNE were cloned in Bacillus subtilis by using pTB53 as a vector plasmid. The presence of the nprA gene enhanced protease synthesis by about fivefold. The nucleotide sequences of nprS and its flanking regions were determined. nprS was composed of 1,653 base pairs and 551 amino acid residues. A Shine-Dalgarno (SD) sequence was found 9 bases upstream from the translation start site (ATG). The deduced amino acid sequence was very similar to that of another thermostable neutral protease gene, nprM (M. Kubo and T. Imanaka, J. Gen. Microbiol. 134:1883–1892, 1988). The amino acid sequence of the extracellular neutral protease NprS was completely identical to that of NprM. By deletion analysis and substitution of the original promoter with a foreign promoter, it was found that the nprA gene existed upstream of nprS. It was also found that a possible target region (palindromic sequence) of the gene product of nprA existed near the promoter sequence of nprS. The nucleotide sequences of nprA and its flanking regions were determined. The DNA sequence revealed only one large open reading frame, composed of 1,218 base pairs (406 amino acids; molecular weight, 49,097). The SD sequence was found 4 bases upstream from the translation start site (GTG). A possible promoter sequence (TTGAAG for the −35 region and AATTIT for the −10 region) was also found about 20 bases upstream of the SD sequence. The nprA gene was separated from nprS by a typical terminator sequence. By constructing an in-frame fusion between the lacZ gene and the 5′ region of the nprA gene, it was demonstrated that the coding region of nprA was indeed translated in vivo. Three palindromic sequences, which are highly homologous with a possible target region by NprA, were also found in the 5′ region of the nprA gene. This suggests that the expression of nprA is autoregulated. From the time course of the production of NprA-LacZ fusion protein, it was indicated that nprA was expressed in late log phase, whereas nprS was expressed in the stationary phase. The NprA protein had consensus regions homologous to the DNA recognition domains of DNA-binding proteins but showed no sequence homology with any other regulatory proteins for protease production. It is inferred that NprA protein binds to the upstream region of nprS promoter and activates transcription of nprS. A new regulatory mechanism by the nprA-nprS genes is discussed.

The extracellular proteases are known to be the most common secretory enzymes of genus Bacillus, and some structural genes have already been cloned. Expression of these structural genes is controlled by the positive or negative regulators. In fact, production of the neutral and alkaline proteases in Bacillus subtilis is controlled by a number of regulatory genes, such as sacU (8), sacQ (30), ptrR (20), hpr (21), and sin (5), which are mapped at various sites on the chromosomes and studied genetically. These mutations or gene amplifications in B. subtilis show pleiotropic effects, such as reduced competency for DNA-mediated transformation, absence of flagellation, altered control of sporulation, and enhancement of synthesis of extracellular enzymes (16). Recently we have cloned such a pleiotropic regulatory gene, degT from Bacillus stearothermophilus, in B. subtilis (26). However, the genetics and biochemistry of the real regulatory mechanisms are not well understood. The regulatory gene is sometimes located in the vicinity of the target gene as has been shown for the Bacillus penicillinate system (9). To analyze the regulation of extracellular protease synthesis from a different angle, we tried to clone the large fragment containing both the neutral protease gene and its regulatory gene from B. stearothermophilus. As a result, an approximately 18-kilobase-pair (kbp) fragment containing these two genes was cloned in B. subtilis. The purpose of this paper is to describe the nucleotide sequences of these genes and to discuss the regulatory mechanism of protease biosynthesis.

MATERIALS AND METHODS

Bacterial strains, plasmids, bacteriophages and culture conditions. The Bacillus strains used were B. stearothermophilus TELNE (Npr+ [neutral protease production]) (FERM-P-9439, Fermentation Research Institute, Agency of Industrial Science and Technology, Ministry of International Trade and Industry, Japan), B. subtilis MT-2 (trpC2 leuC7 hsmM hsrM Npr−) (4), and B. subtilis MT113 (arg-15 trpC2 hsmM Npr+ Apr+ [alkaline protease production]) (11). Escherichia coli JM109 was used as a host strain for phages M13mp18 and M13mp19. Plasmids used were pTB53 (Km− Te−) (10) and pHY300PLK (Te−) (14) for B. subtilis and pUC19 (Ap+) for E. coli. Bacteria were grown in LB broth (10 g of Polypeptone, 5 g of yeast extract, and 5 g of NaCl per liter [pH 7.2]), on L agar (L broth plus 1.5% agar), or on LB agar (L agar plus 1% casein). The antibiotics used were

*Corresponding author.
tetracycline (25 μg/ml) for *B. subtilis* and ampicillin (50 μg/ml) for *E. coli*.

**Transformation.** Transformation of *B. subtilis* and *E. coli* was done by the competent-cell method. Competent cells of *B. subtilis* were prepared as described previously (11). Those of *E. coli* were prepared by the standard method (17).

**Liquid culture of recombinant strains.** Recombinant cells were grown at 37°C in 100 ml of L broth containing an antibiotic on a rotary shaker (180 rpm). Bacterial growth was measured as the optical density at 660 nm.

**Isolation of plasmid and chromosomal DNA.** Plasmid DNA was prepared by either rapid alkaline extraction or CsCl-ethidium bromide equilibrium density gradient centrifugation as described previously (12). Chromosomal DNA was prepared as described earlier (13).

**Manipulation and analysis of DNA.** Cleavage of DNA with restriction enzymes, ligation of DNA with T4 DNA ligase, and exonuclease treatment of DNA were done as recommended by the manufacturer (Toyobo Co., Osaka, Japan). Agarose or polyacrylamide gel electrophoresis was done under standard conditions (17). Recovery of DNA from either low-melting-point agarose or polyacrylamide gel was done by the standard methods (17). DNA was sequenced by the dideoxy method (23) with an M13 cloning kit and Sequenase TM (Toyobo Co.).

**Purification and analysis of extracellular protease.** Purification of neutral protease was carried out as described previously (25). The amino-terminal amino acid sequence was determined by Edman degradation as described elsewhere (25).

**Enzyme assay.** Protease-producing colonies were detected on LC agar plates by halo formation around the colonies. Protease activity was assayed as casein hydrolytic activity (4), while β-galactosidase activity was assayed as o-nitrophenyl-β-D-galactopyranoside hydrolytic activity (28). One unit of protease was defined as the quantity required to increase the A_{280} by the equivalent of 1 μg of tyrosine per min at 35°C and pH 7.5, while 1 U of β-galactosidase causes the liberation of 1 μmol of o-nitrophenol per min at 37°C and pH 7.3. Culture supernatant was used for neutral protease assay. *B. subtilis* cells were prepared for the β-galactosidase assay as described by Nagami and Tanaka (20).

**Analysis of amino acid sequence homology.** Amino acid sequence homology was analyzed with an NEC PC-9801 computer (Nippon Electric Co., Tokyo, Japan) and the GENIAS system (Mitsui Information Co., Tokyo, Japan).

**RESULTS**

**Cloning of neutral protease gene.** Chromosomal DNA (5 μg) from *B. stearothermophilus* TELNE was partially digested with *Sau3AI* and ligated with a *BamHI* digest of pTB53 (1 μg). The ligation mixture was used to transform *B. subtilis* MT-2 (neutral protease-deficient mutant). From about 5,000 Km<sup>Tc</sup> transformants, 1 halo-forming colony was obtained on an LC agar plate. The plasmid prepared from this transformant was capable of transforming *B. subtilis* MT-2 competent cells to Km<sup>Tc</sup> and protease<sup>Tc</sup>. The recombinant plasmid was designated pSP53. pSP53 contained an approximately 18-kbp insert of *B. stearothermophilus* TELNE DNA (Fig. 1). The recombinant plasmid pSP53 was very stable because the copy number is fairly low (10). The protease activity produced by strain MT-2(pSP53) was inhibited by EDTA but not by phenylmethylsulfonyl fluoride, and the optimum pH was between 7 and 8 (data not shown). These results indicate that the enzyme encoded on pSP53 is a neutral protease.

**Construction of deletion derivatives from pSP53 and protease production by the plasmid carrier.** To determine the location of the neutral protease gene, *nprS*, we constructed some deletion derivatives of pSP53 as follows. pSP53 was partially digested with *SalI* and 2.5-kbp or longer fragments were isolated by agarose gel electrophoresis. The *Sau3AI* fragments were digested with *SalI* and the *SalI*- *Sau3AI* fragments were ligated with the *BamHI*-PstI fragment (Km<sup>Tc</sup> of pTB53. The *SalI*-PstI fragments thus obtained were connected to the PstI linker and were self-ligated. The ligation mixture was used to transform *B. subtilis* MT-2, and Tc<sup>Tc</sup> Npr<sup>T</sup> transforms were selected. A recombinant plasmid, designated pSP135, was isolated from one of the transformants. Since the PstI site can be ligated with EcoT22I, rearrangement of the DNA fragment was easy. To obtain a deletion plasmid, pSP135 was digested with EcoT22I and PstI and ligated, and it then was used to transform strain MT-2. Three kinds of plasmids (pSP164, pSP16, and pSP146) were isolated from transformants. When a 3.3-kbp fragment was inserted into the EcoT22I site of pSP164 and pSP146, the recombinant plasmids were designated pSP12 and pSP11, respectively. pSP13 lacks the 3.5 kbp PstI-EcoT22I fragment, and pETB53 contains only the 2.0-kbp EcoT22I fragment. Deletion derivatives of pSP135 are summarized in Fig. 1. Protease production by *B. subtilis* MT-2 carrying each plasmid is also shown in Fig. 1. By deletion analysis, it was found that the neutral protease gene, designated *nprS*, was located around a *BamHI* site of pSP135 and that there may be a positive regulatory gene.

**Nucleotide sequence of neutral protease gene.** The nucleotide sequences of the *nprS* gene and its flanking regions were determined (Fig. 2). Only one large open reading frame was found at the *BamHI* site region. *nprS* was composed of 1,653 bp and 551 amino acid residues. A Shine-Dalgarno (SD) sequence was found 9 bases upstream from the translation start site (ATG at position 1570). A possible promoter sequence (TTCCTTC for the −35 region and AATTTT for the −10 region), which was very similar to the known promoter sequence (TTTCCT for the −35 region and TATTTT for the −10 region) of another protease gene, *nprT* (25), was also found 28 bp upstream of the SD sequence.

The deduced amino acid sequence was very similar to those of other thermostable neutral proteases (15, 25, 27). In particular, the amino acid sequence of the extracellular neutral protease NprS was identical to that of another protease, NprM (15), though the amino acid sequence of the pre-pro region of NprS was 96% homologous with that of NprM.

**Nucleotide sequence of positive regulatory gene for protease production.** Since the positive regulatory gene for protease production was found on the 2.0-kbp EcoT22I fragment, the nucleotide sequence was determined (Fig. 2). Only one large open reading frame was found upstream of the *nprS* gene. A potential SD sequence was found 4 bp upstream from a probable translation start site (GTG) at position 1. The open reading frame was composed of 1,218 bp (406 amino acid residues; molecular weight, 49,097). A typical terminator sequence was also found 186 bp downstream from the open reading frame. To analyze the enhancing activity of enzyme production, several deletion derivatives were constructed from pSP164 (Fig. 3). A possible promoter sequence (TGAGAAG for the −35 region and AATTTT for the −10 region) was also found 20 bases upstream of the SD sequence. The *BamHI*-EcoT22I fragment could enhance the production of neutral protease (Fig. 3). In contrast, all the other fragments cleaved with one of the
B. STEAROTHERMOPHILUS nprS AND nprA

FIG. 1. Construction of deletion derivatives from plasmid pSP53 containing neutral protease gene and its activator gene. Symbols: □ and ■, chromosomal DNA of B. stearothermophilus TELNE; ——, vector DNA; Δ, insertion DNA. BamHI, PstI, SalI, XbaI, and EcoT22I cleavage sites are indicated by B, P, S, X, and ET, respectively. The two EcoT22I sites are numbered 1 and 2.

Expression of nprA in vivo. To analyze the in vivo expression of nprA, the 5'-terminal region of nprA was fused in frame with the lacZ gene as follows. We subcloned the 2.0-kbp EcoT22I fragment in pHY300PLK (14). The recombinant plasmid was cleaved with Cfr101 (Fig. 2 and 3) and was polymerized with Klenow fragment. To remove the 3'-terminal portion of nprA, the fragment was cleaved at the PstI site on the vector plasmid. The large fragment containing the 5'-region of nprA was purified by agarose gel electrophoresis and was ligated with the β-galactosidase gene fusion vector pMC1871 (7,460 bp; Pharmacia Co.) (24), which had been digested with SmaI and PstI. By gene manipulation, the plasmid pETLAC2 carrying an in-frame fusion between the 5'-terminal region of nprA and the lacZ gene cartridge of E. coli was constructed (Fig. 4). Since B. subtilis(pETLAC2) produced the NprA-LacZ fusion protein (data not shown), it was demonstrated that the nprA gene was really translated in vivo.

Role of nprA in enhanced production of protease. To analyze the function of nprA or its gene product, a deletion

restriction enzymes, i.e., HpaI, HincII, AatI, or Cfr101, on the open reading frame or the SD sequence could not enhance the production of neutral protease (Fig. 3). The production of neutral protease was not enhanced by deleting only 2 bp of the PvuI site of the open reading frame (Fig. 3). These results show that this open reading frame is the regulatory gene for enzyme production, and the gene was designated nprA.

Expression of nprA in vivo. To analyze the in vivo expression of nprA, the 5'-terminal region of nprA was fused in frame with the lacZ gene as follows. We subcloned the 2.0-kbp EcoT22I fragment in pHY300PLK (14). The recombinant plasmid was cleaved with Cfr101 (Fig. 2 and 3) and was polymerized with Klenow fragment. To remove the 3'-terminal portion of nprA, the fragment was cleaved at the PstI site on the vector plasmid. The large fragment containing the 5'-region of nprA was purified by agarose gel electrophoresis and was ligated with the β-galactosidase gene fusion vector pMC1871 (7,460 bp; Pharmacia Co.) (24), which had been digested with SmaI and PstI. By gene manipulation, the plasmid pETLAC2 carrying an in-frame fusion between the 5'-terminal region of nprA and the lacZ gene cartridge of E. coli was constructed (Fig. 4). Since B. subtilis(pETLAC2) produced the NprA-LacZ fusion protein (data not shown), it was demonstrated that the nprA gene was really translated in vivo.

Role of nprA in enhanced production of protease. To analyze the function of nprA or its gene product, a deletion
FIG. 2. Nucleotide and amino acid sequences of neutral protease gene, nprS, and its activator gene, nprA. The nucleotide sequence is numbered from base 1 of the initiation codon, GTG. The amino acids of the nprA gene product are numbered from the first amino acid residue. A probable SD sequence and the putative promoter (-35 and -10 regions) are shown by solid lines below the nucleotide sequence. The 98-bp EcoT22I-SspI fragment containing the palindromic sequence and the nprS promoter is surrounded by a box. Palindromic sequences are indicated by broken lines. Probable transcription terminators are indicated by arrows (→ and ←). Asterisks indicate a stop codon. The amino-terminal amino acid sequence of the extracellular protease is indicated by solid lines below the amino acid.
plasmid was constructed from pSP16 by eliminating the EcoT22I-SspI fragment upstream of nprS (Fig. 2). The plasmid pSPF10 (Fig. 5) carries both the structural gene, nprS, and its SD sequence but lacks the putative promoter sequence. *B. subtilis* (pSP16) produced 100 U of neutral protease per ml, whereas pSPF10 carrier did not (Table 1). When the 2.0-kbp EcoT22I fragment carrying the nprA gene was added to pSPF10, the plasmid (pSPF10ET) carrier could not produce neutral protease either (Table 1). This result shows that the promoter sequence of nprS exists on the 98-bp EcoT22I-SspI fragment as mentioned above.

The EcoT22I-PstI fragment of pHY300PLK containing foreign promoter P1 (promoter of the replication protein, Rep-α1, of pHY300PLK) (14) was inserted into the PstI sites of pSPF10 and pSP16 by using the same cohesive ends of EcoT22I and PstI. The recombinant plasmids were designated pSPF10-α1 and pSP16-α1, respectively (Fig. 5). The 2.0-kbp EcoT22I fragment carrying the nprA gene was added to these two plasmids, and the recombinant plasmids were designated pSPF10-α1ET and pSP16-α1ET, respectively. Protease production by these plasmid carriers was examined (Table 1). *B. subtilis* MT-2 (pSPF10-α1ET) produced nearly the same amount of protease as did the pSPF10-α1 carrier. In other words, the nprA gene product could not enhance transcription from foreign promoter P1. Since pSP16-α1 contains two tandem promoters (P1 promoter and the original nprS promoter), enzyme production by the plasmid carrier was about equal to the sum (280 U/ml) of production by the pSP16 carrier (100 U/ml) and the pSPF10-α1 carrier (150 U/ml). A similar additive effect on enzyme production

FIG. 3. Localization of nprA gene in 2.0-kbp EcoT22I fragment. Symbols: □, EcoT22I fragment and its deletion derivatives; △AT, 2-bp deletion at the PvuI site; +, production of neutral protease enhanced; −, production of neutral protease not enhanced.

FIG. 4. Structure of plasmid pETLAC2. Symbols: ■, N-terminal region of nprA; □, lacZ cartridge of pMC1871; ——, pH300PLK DNA. PstI, Cfr10I, and Smal cleavage sites are indicated by P, C, and S, respectively. Reaction with Klenow fragment is indicated by →K. N-terminal and C-terminal regions of the fusion protein gene are indicated by NH₂ and COOH, respectively.
by pSP16-a1ET carrier was also found, e.g., 560 U/ml by the pSPT2 carrier plus 140 U/ml by the pSP10-a1ET carrier. These results indicate that the npra gene product can specifically enhance transcription from the nprs promoter. It is most likely that the 98-bp EcoT22I-SspI fragment contains both the promoter sequence (~35 region and ~10 region) and the target region for NprA. We expected that the npra gene product was the transcriptional activator of the nprs promoter, as has been shown for the other transcriptional activators (22). Actually, the 28-bp palindromic sequence, which was similar to the targets of general transcriptional regulators (LacI [3], cyclic AMP receptor protein [1], DeoR [29], etc.), was found 5 bp upstream of the nprs promoter, and the sequence might be used as a regulatory region.

**Production phases of NprA and NprS.** To obtain information on the activation mechanism for enzyme production, the production phases of NprA and NprS were investigated by cultivating *B. subtilis* MT-2(pSP16) (NprA⁺ NprS⁺), MT-2(pSPT2) (NprA⁺ NprS⁺), and MT-2(pETLAC2) (NprA⁺ NprS⁺). 

**DISCUSSION**

Production of the extracellular proteases from *B. subtilis* is regulated at the transcription step by many genes, such as *sacQ* (30) and *prtR* (20). These genes are located away from the locus of proteases, and they also regulate the production of other extracellular enzymes. In contrast, a neutral protease-specific regulatory gene, *npra*, was on the same locus as the neutral protease gene, *nprs*. The amino acid sequence of NprA exhibited no homology with sequences of the pleiotropic regulatory gene products. Actually, when the *npa* gene was introduced into *B. subtilis* MI113 (Npr⁺ Apr⁺), the production of neutral and alkaline proteases was not changed (data not shown). Therefore, it seems that *npa* differs in property and function from the known genes for regulating the production of proteases. The structure of *npa-nprs* genes looks like those of general operons, such as the *lac* operon (2).

Some information on the mechanism of the *npa-nprs*
in consequence, the production of NprA is repressed; i.e., there is an autoregulation system. Penl, the negative regulator of the penicillinase gene from Bacillus licheniformis, is a famous example of a regulatory protein being autoregulated (9).

From the information given above, we infer the following mechanism of regulation for extracellular neutral protease production. At first, NprA is produced in the late log phase, but the expression of nprS is autoregulated, because NprA binds to the specific palindromic sequences of nprS. Since NprA also binds to the regulatory sequence of nprS, the promoter of nprS is activated. In the stationary phase, the nprS gene is highly expressed by using the activated promoter. To further investigate the biochemical basis of the mechanism, purification of NprA protein is now in progress.

ACKNOWLEDGMENTS

We are grateful to S. Tarama and Y. Sogabe for helpful discussion.

LITERATURE CITED


13. Imanaka, T., T. Tanaka, H. Tsunekawa, and S. Aiba. 1981. Cloning of the genes for penicillinase, penP and penl, of Bacillus licheniformis in some vector plasmids and their expression in Escherichia coli, Bacillus subtilis, and Bacillus licheni-


