千葉大園学報 (Tech. Bull. Fac. Hort. Chiba Univ.)
第47号 45-51 (1993)

Assignment and Similarity of the Chitosanase Gene of *Bacillus circulans* MH-K1.

Akikazu ANDO, Takao KOBAYASHI, Hideaki MOROSAWA, Shoji OKAJIMA, Hirofumi SHINOYAMA, and Takaaki FUJII.
(*Laboratory of Microbial Engineeing*)

ABSTRACT

Molecular structure of chitosanase gene of *Bacillus circulans* MH-K1 was examined. In the 5′ upstream region of the chitosanase open reading frame, there are two kinds of regulatory sequences that might regulate the expression inducible chitosanase gene. Chitosanase protein has a few short segments homologous to lysozyme. One of them just covers the catalytic amino acids domain of lysozyme. One of the inverted repeat sequence of nucleotides in the chitosanase gene suggests that the chitosanase gene evolved by the dimerization of lysozyme. Localization of the amino acids, which correspond to catalytic amino acids of lysozyme, also supports the idea.

INTRODUCTION

The chitosanase of *Bacillus circulans* MH-K 1 was cloned and sequenced (1). Its open reading frame including signal sequence and some regulatory sequences (RNA polymerase binding sites, ribosomal binding site and termination structre) were assigned. Trial of homology assignment of the chitosanase gene showed no overall similarity to any other enzymes.

We further examined the similarity and homology of the chitosanase gene including noncoding region. The chitosanase protein keeps catalytic amino acids sequence of lysozyme. The idea that the chitosanase was evolved from lysozyme by dimerization was supported by the findings of inverted repeat in the center of matured chitosanase protein and two sets of homologous amino acids which correspond to catalytic amino acids of lysozyme.

In the upstream region of the open reading frame, there are two kinds of regulatory sequences that seemed to be responsible for the regulation of the inducible chitosanase gene expression. And other molecular structure of the gene is also examined in this study.

MATERIALS AND METHODS

Sequence similarity analysis was done using the software of Genentyx ver. 7.0 as already reported (1) or done manually.

RESULTS AND DISCUSSION

1. Non coding (5′-upstream and 3′-down stream) region of the chitosanase gene.

Regulatory site. Because MH-K1 chitosanase is an inducible enzyme, there should be some regulatory site. One of the most popular regulatory

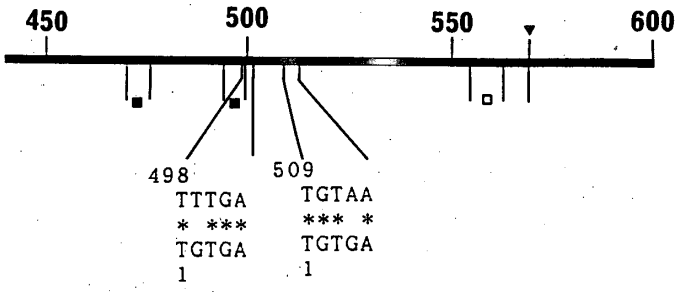

Figure 1 Assignment of regulatory sites of chitosanase gene, which are similar to CAP binding sites of the *lac* promotor.

■ show -35 and -10 sequences of RNA polymerase binding sites. □ shows Shine-Dalgalno sequence for ribosomal binding site. ▼ shows initiation codon for translation. Above numbers show the nucleotide numbers of the gene. Lower nucleotde sequence is the consensus sequence of CAP binding site.

mechanism is the catabolite activator protein (CAP) of *lac* operon (of *Escherichia coli*) (2). There are more than ten sections which possess 80% of the homology to either TGTGA or TCACT, which are the consenseus sequence of CAP binding site. Among them, most plausible ones are those of at 498 to 513 of upstream sequence, which is partially overlap to -10 site of RNA polymerase binding site. (Fig. 1)

Other regulatory mechanism of inducible enzyme is the repressor binding mechanism of carbon catabolite control. *ara* operon of *E. coli* is the typical one (3). Direct repeat sequence of chitinase (of *Streptomyces plicatus*) for the binding of repressor had been identified (4). The direct repeat sequences in the chitosanase gene similar to those of two chitinase genes (*chi*63 and *chi*35) were searched. As shown in Fig. 2, there are some parts in chitosanase gene, which possess homology to those of *chi*63 (TTGTCCAGACCT) and *chi*35 (TGGTCTAGTCCT). These homological sequences of chitosanase are not close to the -10 and -35 sequences (that is transcriptional control site). Rather, they are in the neighbour of SD sequence. That discourages the idea that the repressor binding mechianism is the major regulatory mechanism of the chitosanase gene.

Repeated sequences in the 5′ upstream regon of chitosanase gene. To find any regular seguence that might have regulatory function, repeated sequence (minimum 75% homology) was searched in the 5′ upstream region of the chitosanase gene (base No 1 to 600). As shown in Fig. 3, three direct repeats (1 in Fig. 3), four inverted repeats (not complementarily) (2 in Fig. 3) and 7 inverted complementarily repeats and 6 palindoromes (3 in Fig. 3) are found. The functional meanings of these repeated sequences are not clear.

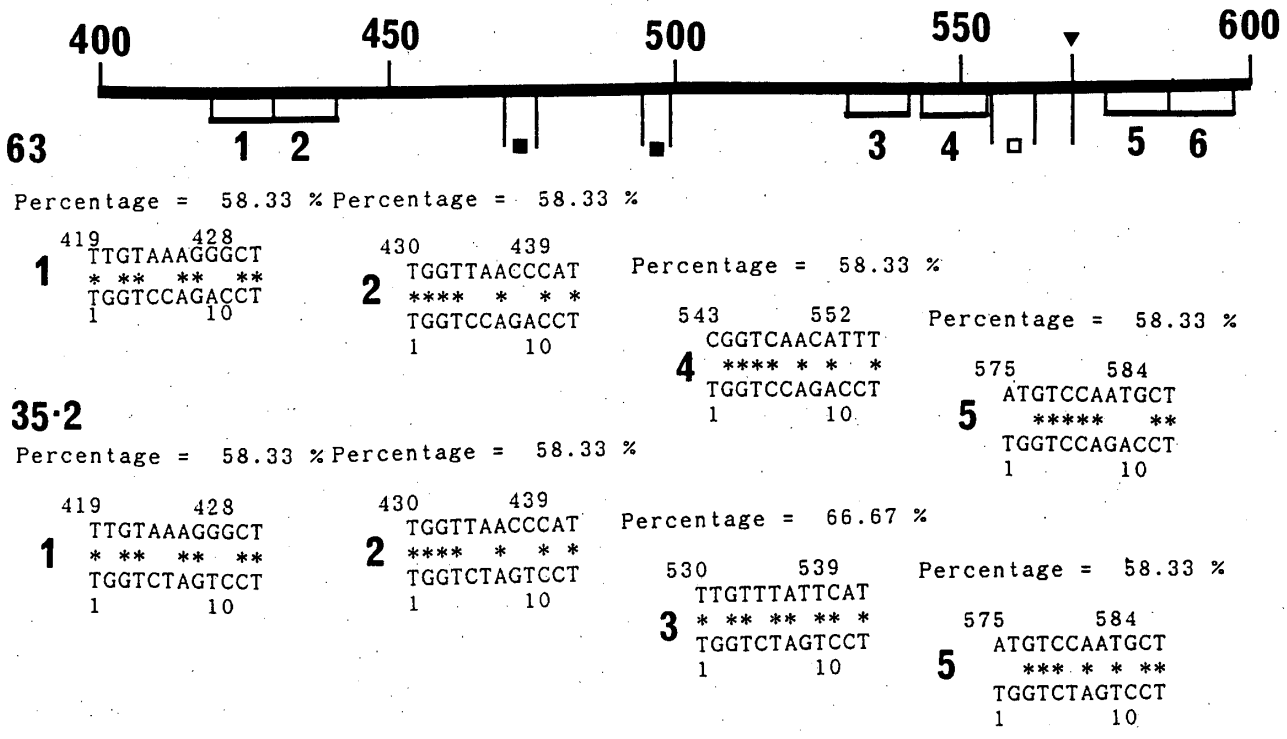

Figure 2 Assignment of direct repeat sequences in the chitosanase gene that might be responsible for cabon catabolite control.

Solid line shows chitosanase gene. Above numbers show the nucleotide number of the gene. Signs around the solid line are similar to those of Fig. 1. 63 is the fragments of chitosanase gene that show more than 58 % homology to the repressor binding site of chitinase 63 (4). 35-2 is the similar fragments that show homology to 2nd repressor binding site of chitinase 35 (4).

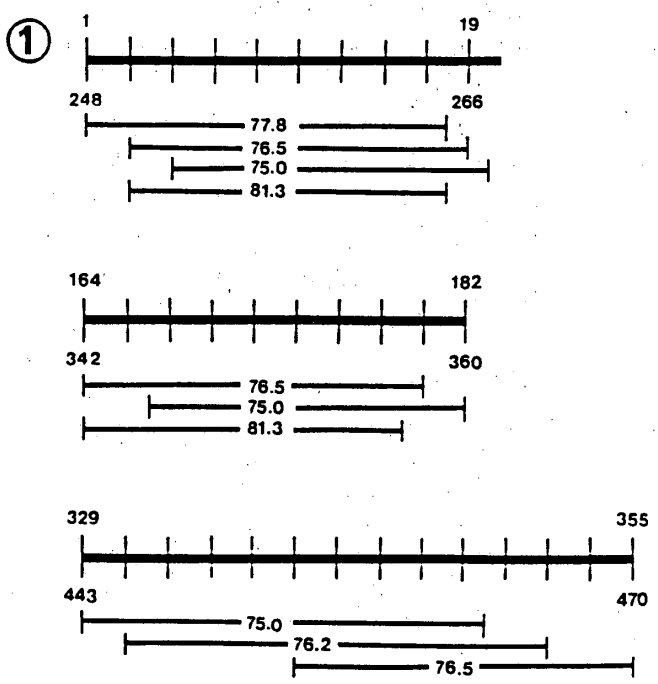

Fig 3 − 1

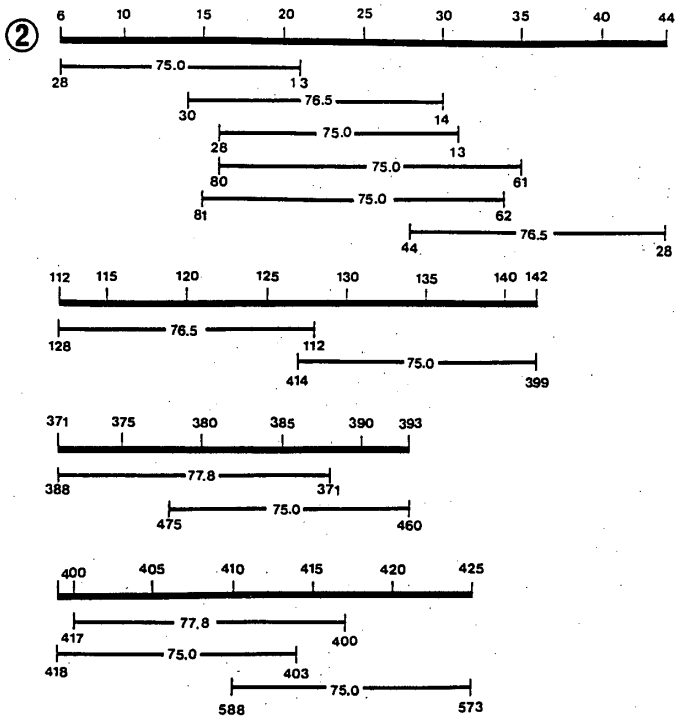

Fig 3 − 2

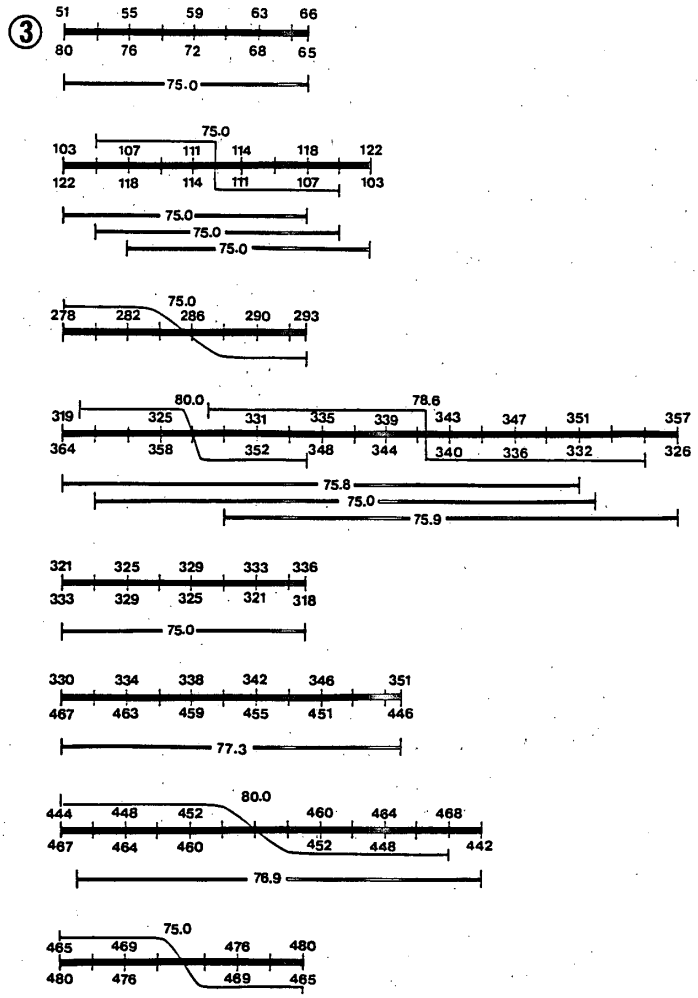

Fig 3 − 3

Figure 3 Repeat sequences in 5′ upstream region of the chitosanase gene.

1, direct repeat; 2, inverted repeat (not complementarily); 3, inverted and complementarily repeat, and palindrome. Numbers above or under the thick lines show the nucleotide numbers of the chitosanase gene. Numbers in the middle of lower thinner lines show percent of homologies of that part. ⌐¬ shows palindrome, and the number above above it shows the percent of its homology.

Open reading frame in 5′ upstream or 3′ down stream region of chitosanase gene. It is popular that inducibie enzyme genes cluster as operon (like *lac* operon of *E. coli*), then open reaging frames

(initiation codon; AAA, ATG, GTG: termination codon; TAA, TAG, TGA) was searched in 5′ upstream or 3′ downstream region of the chitosanase gene. Only complementary sequence of 5′ upstream region showed open reading frame structure (1 of Fig. 4). For 3′ downstream region, reverse, complementarily, and reverse and complementarily sequences showed open reading frame (2 of Fig. 4). The length of open reading frames shown in Fig 4, is too small to code a protein of reasonable size. For further analysis, we should get further continuing sequences.

2. Open reading frame of the chitosanase gene

Amino acid sequence homology. When we surveyed the sequence homology of *B. circulans* MH-K 1 chitosanase with other enzymes including chitinase, no meaningful sequence was obtained (in the previous study, ref. 1). Then we searched the homology of shorter length.

As shown in Fig. 5, some homological sequences of MH-K 1 chitosanase to lysozymes of various origines were found. Among these sequences, most important one is that of No 98-115, which corresponds to No 35-52 of lysozyme of domestic pigeon (and to No 53-72 of human lysozyme). These 35 Glu and 52 Asp are catalytic amino acids of lysozyme. Catalytic amino acid contained part of other lysozyme (36 Glu-52 Asp, of human langur) also showed homology to 185-197 of chitosanase. The latter similarity is lower than that of the former. Signal sequence of chitosanase also showed some similarity to that of chiken

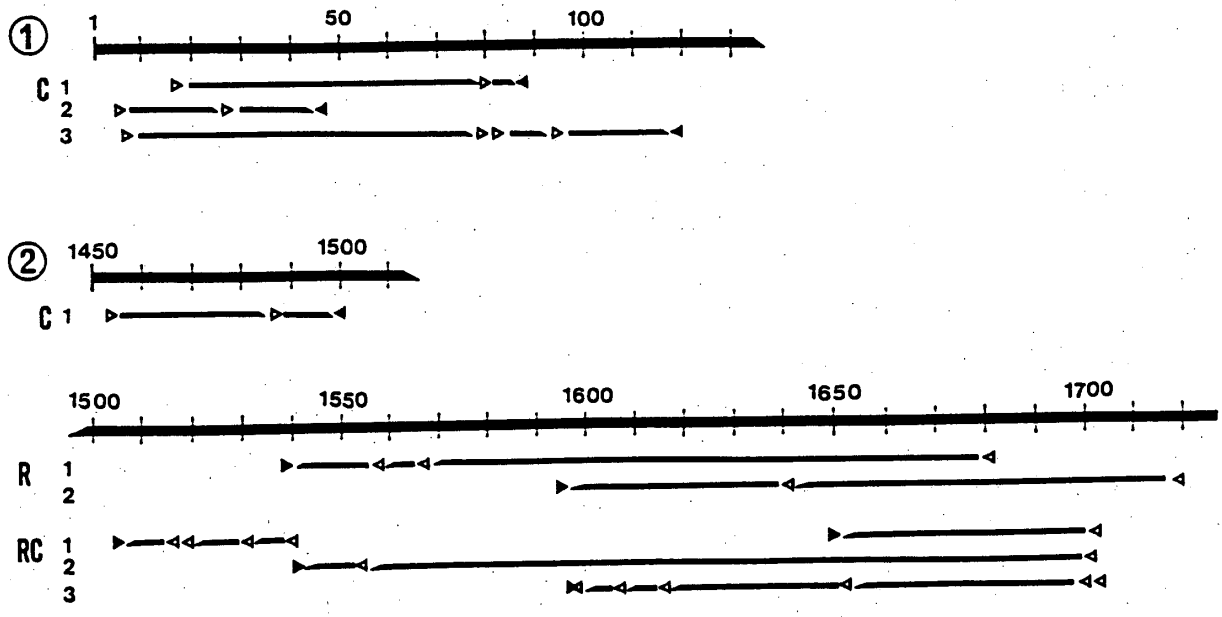

Figure 4 Survey of open reading frame in 5′ upstream side (1) or 3′ downstream side (2) of the chitosanase gene.

For the survey of open reading frame, initiation codon (AAA, ATG or GTG; ▷) and termination codon (TAA, TAG or TGA; ◀) were searched in three reading frames. Numbers above the thick line show nucleotide numbers. Thinner line between ▷ and ◀ shows open reading frame. Open reading frames are found in complementarily (C), reverse (R), or reverse and complementarily (RC), sequences. C1 means reading frame No 1 of the complementarily sequence.

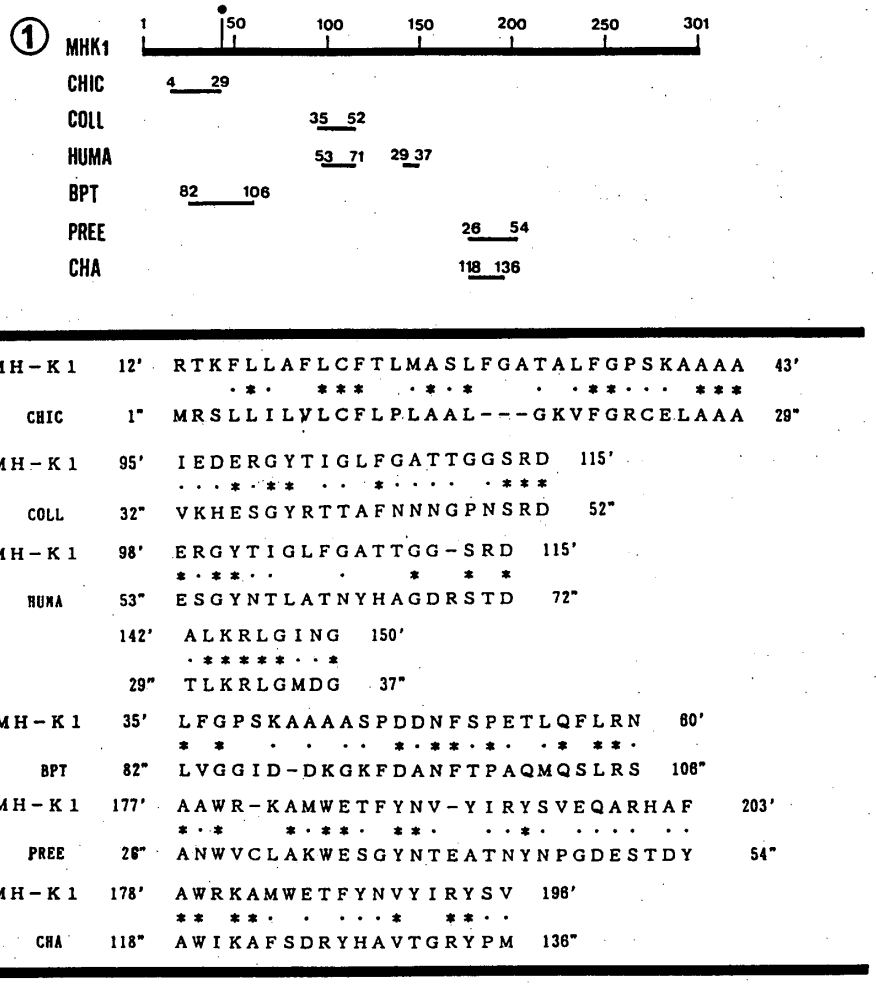

Figure 5 Amino acid sequence homology of the chitosanase to lysozymes of various origin.

1 shows an overview of homologies, and 2 shows corresponding amino acid sequences. Numbers in 1 show their amino acid number. ↑ shows N terminal amino acid of chitosanase protein. MH-K1 is chitosanase protein of *B. circulans* MH-K1. The origin of lysozyme are as follows : chic, *Gallas gallus* (chicken) lysozyme C precursor; coll, *Columba livia* (domestic pigeon) lysozyme C; huma, human lysozyme C precursor; bpt, *Presbytis entellus* (human langur) lysozyme C; cha, *Chalaropsis* (imperfect fungi) lysozyme CH.

lysozyme, because of its hydrophobic nature.

Repeated nucleotide sequence in the frame. Molecular weight of lysozymye is about 15k, which is just half of the MH-K1 chitosanase protein. Dimerization of the prototype gene is one of the popular means for the evolution of new enzyme. We searched repeatrd nucleotide sequence within the open reading frame. As shown in Fig. 6, there are one direct repeat, two inverted repeats (not complementarily), one inverted complementarily repeat. These repeated fragments do not contain nucleotide No 859-912 (corresponded to amino acid No 98-115). The central (or turning) point of the inverted repeat (1047-1072 to 1120-1095) is between 1083 and 1084. That point is also the central and half-way point of matured protein (between amino acid No 172 and 173, these numhers include signal sequence). The amino acids, which have simiar distance from that point and correspond to 98 Glu (17 amino acids from 115 Asp) and 115 Asp (57 amino acids from the center), are 226 Asp (53 amino acids from the center) and 241 Glu (15 amino acido from 226 Asp). These repeated sequences suggest the idea that the chitosanase is evolved by the inverted dimerization of the prototype lysozyme, although the chitosanase did not show lysozyme activity using glycol chitin as a substrate.

The hypothesis is enough attractive to be evaluated through the examination of the active sites by site-directed mutagenesis.

ACKNOWLEDGMENT

We thank D. Takeshi Watanabe of Niigata University for his kind help on searching sequence homology.

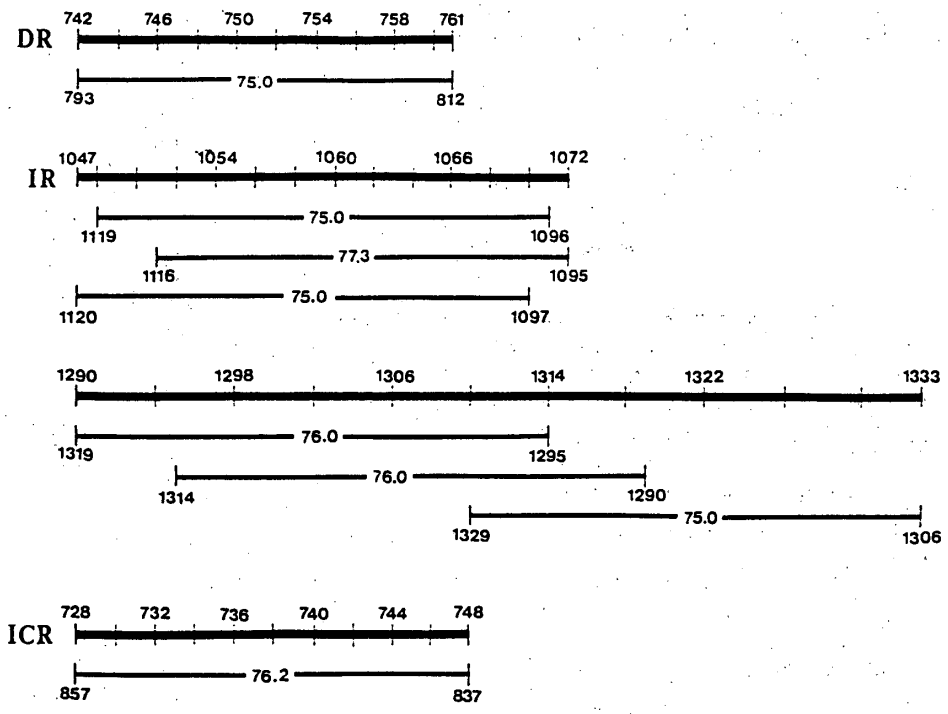

Figure 6 Repeated nucleotide sequences in the open reading frame of chitosanase gene.

Thick lines show chitosanase gene. Thinner lines show repeated part. Numbers above or under lines are nucleotide numbers. Numbers in the middle of the thinner lines are percent of homology. DR, direct repeat; IR, inverted repeat; ICR, inverted and complementarily repeat.

REFERENCES

(1) Ando, A., K. Noguchi, M. Yanagi, H. Shinoyama, Y. Kagawa, H. Hirata, M. Yabuki, and T. Fujii (1992): Primary structure of chitosanase produced by *Bacillus circulans* MH-K1, *J. Gen. Appl. Microbiol.*, 38, 135-144.

(2) Gartenberg, M. R. and D. M. Crothers (1988): DNA sequence determinants of CAP-induced bending and protein binding affinity, *Nature*, 333, 824-829.

(3) Brunelle, A. and R. Schleif (1989): Determining residue-base interactions between araC protein and *ara*I DNA, *J. Mol. Biol.*, 209, 607-622.

(4) Delic, I., P. Robbins, and J. Westpheling (1992): Direct repeat sequences are implicated in the regulation of two *Streptomyces* chitinase promoters that are subject to carbon carbon catabolite control, *Proc. Natl. Acad. USA*, 89, 1885-1889.

Bacillus circulans MH－K1のキトサナーゼ遺伝子の構造について

安藤昭一、小林賢生、諸沢英明、岡嶋正治，篠山浩文、藤井貴明
（微生物工学研究室）

摘　要

Bacillus circulans MH－K1のキトサナーゼ遺伝子構造の解明を目的として、本遺伝子のオープンリーディングフレーム及びその前後のノンコーディング部分について検討した.

キトサナーゼの構造遺伝子の5′上流側にこの誘導酵素の発現調節に関与していると推定される2種類の制御部位が認められた.

キトサナーゼ蛋白はリゾチームとホモロジーのある幾つかの短い部分配列を有していた．そのうちの一つはリゾチームの触媒アミノ酸の部分と重なっていた．核酸塩基配列中の繰り返し配列検索の結果、及びリゾチームの触媒アミノ酸類似のアミノ酸配列検索によると、進化論的は本キトサナーゼはリゾチーム二個融合して出来たとも考えられた.