

Multiple domain structure in a chitinase gene (*chiC*) of *Streptomyces lividans*

TAKESHI FUJII* and KIYOTAKA MIYASHITA

Laboratory of Soil General Microbiology, National Institute of Agro-Environmental Sciences, 3-1-1 Kan-nondai Tsukuba, Ibaraki 305, Japan

(Received 3 August 1992; revised 5 November 1992; accepted 24 November 1992)

One of the chitinases of *Streptomyces lividans*, chitinase C, was encoded by a 2 kb *smaI*–*XhoI* restriction fragment contained in the recombinant plasmid pEMJ7. DNA sequence analysis of this region revealed the presence of two open reading frames (ORF1 and ORF2) which had opposite orientations. Northern analysis showed that only the mRNA complementary to ORF1 was transcribed, and that this transcription was induced by chitin and repressed by glucose. ORF1 showed a codon distribution typical of *Streptomyces*. A sequence identical to that of the N-terminus of mature secreted chitinase C was found from amino acid residue 31 in the deduced amino acid sequence of ORF1 (619 amino acids), implying that ORF1 encodes a pre-protein of chitinase C. The pre-protein of chitinase C consisted of four discrete domains. The 30 amino acid N-terminal sequence, domain 1, was characteristic of a signal peptide. Domain 2 consisted of 105 N-terminal amino acids of mature chitinase C, and was similar to cellulose-binding domains of several cellulases. Domain 3 (94 amino acids) showed homology with type III homology units of fibronectin. Domain 4, a C-terminal 390 amino acid sequence, is probably the catalytic domain of the chitinase, since it exhibited identity with several other chitinolytic enzymes.

Introduction

Chitin (poly- β -1,4-*N*-acetyl-glucosamine) exists abundantly in nature as a major structural component of the cell walls of fungi, arthropod exoskeletons, and crustacean shells. Many organisms, from bacteria to higher plants, are reported to produce chitinases (EC 3.2.1.14) which hydrolyse chitin to chitodextrins. These organisms usually possess several chitinase genes, whose expression is induced by extracellular chitin or its derivatives (Berger & Reynolds, 1958; Monreal & Reese, 1969; Watanabe *et al.*, 1990a; Legrand *et al.*, 1987; Herget *et al.*, 1990). They secrete multiple chitinases to use chitin as a carbon and/or nitrogen source (in bacteria), or to defend against pathogen attack (in plants), suggesting that these various chitinases work together to hydrolyse chitin. The amino acid sequences of several plant and bacterial chitinases have been reported (Jones *et al.*, 1986; Broglie *et al.*, 1986; Kamei *et al.*, 1989; Harpster & Dunsmuir, 1989;

Laflamme & Roxby, 1989; Metraux *et al.*, 1989; Parsons *et al.*, 1989; Payne *et al.*, 1990; Watanabe *et al.*, 1990b; Samac *et al.*, 1990; Zhu & Lamb, 1991; Watanabe *et al.*, 1992; Kuranda & Robbins, 1991), but these sequences are not always homologous. This suggests that chitinase genes have spread through various organisms in the process of evolution and have formed several gene families.

Streptomyces spp. are typical soil inhabitants, and produce various carbohydrases, including chitinases. *Streptomyces lividans*, for example, secretes at least four chitinases (Miyashita *et al.*, 1991). The production of these enzymes is induced by chitin, and repressed by glucose. The regulation of chitinase genes in bacteria is still unclear. The mechanism of catabolite repression in *Streptomyces* seems to be quite different from the cAMP-dependent system in *Escherichia coli* (Pastan & Adhya, 1976). Neither the function of the multiple chitinase system in the degradation and utilization of chitin, nor the evolution of these chitinase genes, is well understood.

Our group is interested in the chitinase systems of *Streptomyces*. We have isolated strains of *Streptomyces* with relatively high chitin-degrading activities from soil (Ueno *et al.*, 1990); and to analyse the multiple chitinase system and its genetic control, we have cloned three

* Author for correspondence. Tel. 0298 38 8256; fax 0298 38 8199.

The nucleotide sequence data reported in this paper have been submitted to GenBank, DDBJ and EMBL and have been assigned the accession number D12647.

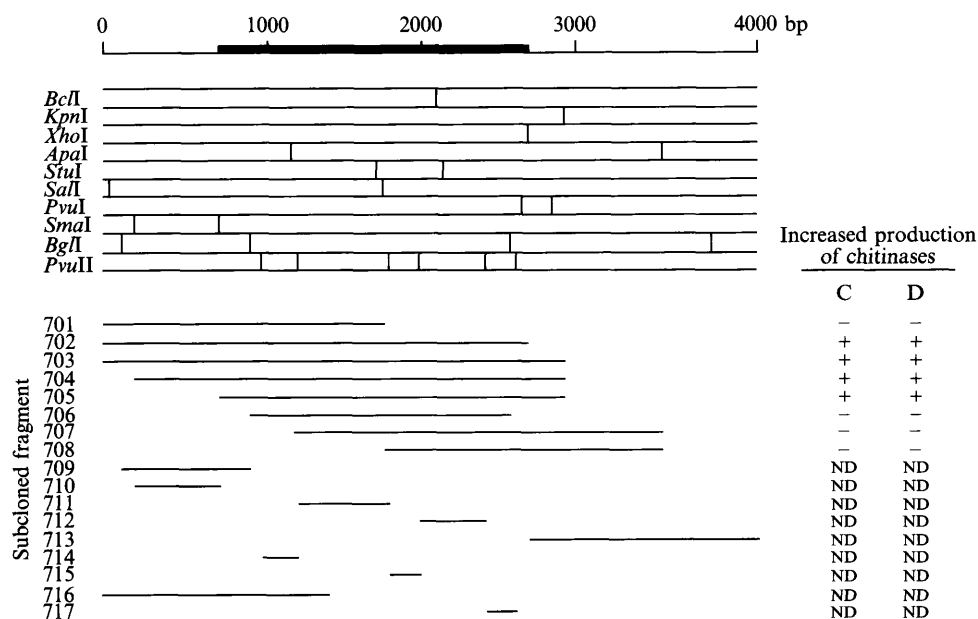


Fig. 1. Restriction map and various subcloned DNA fragments of the insert fragment of pEMJ7. The solid horizontal lines below the restriction map indicate the fragments used in this study to localize the gene and/or for sequencing analysis. The bold line indicates the smallest restriction fragment containing the promoter sequence and the structure gene for chitinase C, as indicated by deletion analysis. Production of chitinases C and D was evaluated by SDS-PAGE of the culture supernatant of each subclone. ND, Not done.

chitinase genes from *S. lividans* by self-cloning (Miyashita *et al.*, 1991). These cloned genes did not hybridize with each other, and were subject to both chitin induction and glucose repression. One of the clones, containing pEMJ7, secreted much chitinase C when grown on a chitin medium. In this report, we describe the structure and expression of the gene which encodes a pre-protein of chitinase C, and discuss the multiple domain structure of the enzyme.

Methods

Bacterial strains, plasmids and media. *Streptomyces lividans* 66 (strains TK24 and TK64) and *Streptomyces* plasmid pIJ486 were kindly supplied by Dr D. A. Hopwood (John Innes Institute, Norwich, UK). Plasmid pEMJ7, a derivative of vector pIJ486, contains a 4.5 kb fragment carrying *chiC* of *S. lividans*. Plasmids pEMJ701, 702, 703, 704, 705, 706, 707 and 708 are derivatives of pEMJ7 containing various restriction fragments of the insert DNA of pEMJ7. Plasmid pUC18 and phages M13mp18 and M13mp19 were used for cloning in *E. coli* JM109 (Yanisch-Perron *et al.*, 1985). Restriction fragments derived from the insert DNA of pEMJ7 (701 to 717 in Fig. 1) were subcloned into the multiple cloning sites of pUC18, and then recloned into M13mp18 (pEMM701, 702, 703, 706, 709, 710, 711, 712, 714, 715, 716 and 717) or M13mp19 (pEMM704R, 705R, 706R, 707R, 708R, 709R, 711R, 712R, 713R, 714R, 715R and 717R) (Fig. 1). The recombinant phage DNA generated was subjected to dideoxynucleotide sequencing. *S. lividans* containing recombinant plasmids was grown at 30 °C with shaking in Luria broth (LB) medium containing thiostreptin (50 µg ml⁻¹) to prepare plasmids. For extraction of RNA, *S. lividans* containing pEMJ7 was grown at 30 °C with shaking in inorganic salts

medium (Miyashita *et al.*, 1991) containing colloidal chitin (0.15 mg ml⁻¹), thiostrepton (50 µg ml⁻¹) and yeast extract (1.5 mg ml⁻¹) with or without glucose (15 mg ml⁻¹). Thiostrepton was a generous gift from S. J. Lucania (E. R. Squibb & Sons). YEME and R2YE medium were used for protoplast formation and transformation of *Streptomyces*, respectively (Hopwood *et al.*, 1985).

Recombinant DNA techniques. Preparation of plasmid, double-stranded and single-stranded M13 DNA, restriction enzyme digestion, recovery of DNA fragments from agarose gels, transformation of *E. coli*, and selection of transformed colonies of *E. coli* and transformed plaques of M13 clones were done according to Sambrook *et al.* (1989). Protoplasting, transformation, selection of transformants of *S. lividans*, and preparation of plasmid DNA from *S. lividans* were done according to Hopwood *et al.* (1985). T4 DNA ligase (ligation kit, Takara Shuzo, Kyoto, Japan), alkaline phosphatase (Toyobo, Osaka, Japan) and T4 DNA polymerase (blunting kit, Takara Shuzo) were used according to the manufacturers' instructions.

Northern hybridization. A 50 ml culture of *S. lividans* containing pEMJ7 was grown at 30 °C in inorganic salts colloidal chitin medium with or without 1.5% (w/v) glucose for 2 d. Total RNA was extracted from the cells as described by Hopwood *et al.* (1985). As a control, RNA was also extracted from cells grown in inorganic salts medium without chitin and glucose. The resulting RNA fractions (30 µg equivalent) were applied to a 14 × 15 cm² 1.2% (w/v) agarose gel containing formaldehyde (2.2 M). After electrophoresis, RNA bands were transferred onto a nylon membrane (Hybond N⁺, Amersham) following the method of Sambrook *et al.* (1989). As a molecular mass marker, a 0.24–9.5 kb RNA ladder (BRL) was used. A single-stranded end-labelled probe of each fragment was synthesized by the method of Ausubel *et al.* (1989) with the following modifications. To prepare templates for primer extension reaction, a *PvuII* restriction fragment (460 bp), which was located into the centre of the *SmaI*–*XhoI* fragment

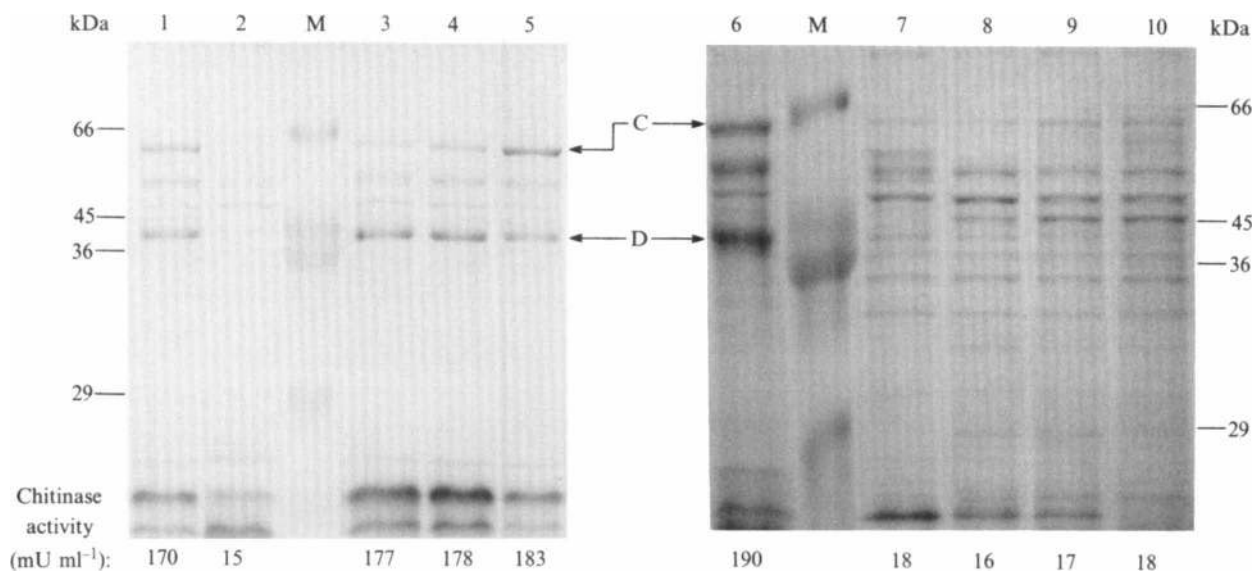


Fig. 2. SDS-PAGE analysis of increased production of chitinases C and D from *S. lividans* containing recombinant plasmids. Increased production of chitinase C and D was monitored in the supernatant of cultures of *S. lividans* containing: pEMJ7 (lane 1); pEMJ486 (lane 2, control); pEMJ702 (lane 3); pEMJ703 (lane 4); pEMJ704 (lane 5); pEMJ705 (lane 6); pEMJ701 (lane 7); pEMJ706 (lane 8); pEMJ707 (lane 9) and pEMJ708 (lane 10). Molecular mass markers (lane M) are indicated. The chitinase C and D bands are indicated by arrows. Chitinase activity in culture supernatant of each clone is indicated at the bottom of each lane.

of the insert DNA of pEMJ7, was cloned into the *Sma*I site of pUC18 to construct pEMU712. The *Bam*HI-*Eco*RI restriction fragment of pEMU712 was subcloned into M13mp18 and M13mp19 to generate pEMM712 and pEMM712R, respectively. Single-stranded template DNAs were prepared from these M13 subclones. M13-M4 primer (0.1 μ g, Takara Shuzo) was end-labelled in the presence of 200 μ Ci (7.4 MBq) [γ - 32 P]dATP. After annealing of labelled primer to the single-stranded DNAs, the strands were extended with Sequenase Version 2 (United States Biochemicals), following the manufacturer's instructions. The resulting double-stranded DNAs were cut by *Hind*III or *Eco*RI. Labelled single-stranded probes were isolated from 1.2% (w/v) low-melting-temperature agarose gel after alkaline agarose gel electrophoresis. The resulting single-stranded probes used to detect transcripts of ORF1 and ORF2 were designated probes 712R and 712, respectively. Northern hybridization and washing of the membrane followed the method supplied by the manufacturer (Amersham).

DNA sequencing. Single-stranded DNAs were extracted from various recombinant M13 clones and sequenced using the 7-deaza DNA sequencing kit (Sequenase Version 2, United States Biochemicals) following the method supplied by the manufacturer. [α - 35 S]dCTP (Amersham) was used for labelling. The sequences were determined at least once for each strand.

N-terminal amino acid sequence analysis. Chitinases were partially purified by Mono-Q ion-exchange column chromatography of crude enzyme obtained from supernatant from a 5-d-old culture (500 ml) as described previously (Miyashita *et al.*, 1991). The peak fractions of chitinase C and D were frozen, lyophilized, and subjected to SDS-PAGE (Laemmli, 1970). The gel was stained with Coomassie Brilliant Blue R250, and the protein bands were electrophoretically blotted onto polyvinylidene difluoride transfer membrane (Immobilon-P, Millipore). The bands corresponding to chitinase C and D were cut from the membrane and subjected to N-terminal amino acid sequence analysis using an automated protein sequencer (model 477A, Applied Biosystems).

S1 mapping. Synthesis of a single-stranded end-labelled probe for S1 mapping was performed as described under *Northern hybridization*, with the following modifications. To prepare templates for the primer extension reaction, the *Bgl*II restriction fragment (900 bp), which contained the translation initiation site and promoter sequence of *chiC*, was cloned into the *Sma*I site of pUC18 to construct pEMU709. The *Bam*HI-*Eco*RI restriction fragment containing the insert DNA of pEMU709 was subcloned into M13mp18 to construct pEMM709, and single-stranded DNA was prepared. The orientation of the *Bgl*II fragment was confirmed by sequencing. A primer (5'TCGCGAGGCCGACCAGA3'), complementary to the sequence located 80 bp downstream from the putative translation initiation site, was synthesized and end-labelled with 200 μ Ci (74 MBq) of [γ - 32 P]dATP. After annealing to the single-stranded DNA, the strand was extended by Sequenase Version 2. The resulting double-stranded DNA was digested by *Nae*I, which cut 200 bp upstream of the putative translation initiation site, and a labelled single-stranded probe was isolated from 1.2% low-melting-temperature agarose gel after alkaline gel electrophoresis. Total RNA was extracted from cultures of *S. lividans* containing plasmid pEMJ7 on inorganic salts colloidal chitin medium with or without 15 mg glucose ml $^{-1}$. The labelled single-stranded probe (equal to 6.25×10^4 Cerenkov counts) and 50 μ g of the resulting total RNA fractions were annealed, and non-hybridized strands were digested by 100 U or 300 U of S1 nuclease (Takara Shuzo) by the method of Ausubel *et al.* (1987). After treatment with S1 nuclease, the radioactive DNA was subjected to denaturing PAGE (using a 4%, w/v, polyacrylamide gel) with 7-deaza DNA sequencing reaction mixtures of pEMU709 prepared using the same synthesized primer, as a control sequence ladder.

Enzyme assay. Chitinase was assayed as described previously (Miyashita *et al.*, 1991), using the fluorogenic substrate 4-methylumbelliferyl *N,N',N''*-triacetyl chitotriose [4-MU-(GlcNAc) $_3$] (Sigma). For evaluation of the chitinase activity in the supernatant of recombinants, supernatant from an 8–12 d culture was used as the

source of enzyme. One unit of chitinase activity was defined as the amount of enzyme that liberated 1 μmol 4-MU from the substrate in 1 min at 37 °C.

SDS-PAGE. SDS-PAGE gels (12%, w/v) were purchased from Funakoshi Chemicals, and electrophoresis was done according to the manufacturer's instructions. After electrophoresis, gels were stained with Coomassie Brilliant Blue G or R. To estimate the protein molecular masses, a marker protein kit (MW-SDS-70; Sigma) was used. A 1.5 ml portion of each supernatant from 8–12 d cultures of *S. lividans* TK24 containing recombinant plasmid was lyophilized. The residue was dissolved in 50 μl of the sample buffer and subjected to SDS-PAGE.

Results and Discussion

Subcloning of chitinase gene in *S. lividans*

The supernatant of *S. lividans* 66 TK24 containing pEMJ7 had 10-fold greater chitinase activity than that of the parent strain *S. lividans* TK24. When total extracellular protein from the transformants with pEMJ7 was subjected to SDS-PAGE, significant bands identical to those of chitinase C and D were observed (Miyashita *et al.*, 1991). To determine the coding region for these enzymes, various restriction fragments of the insert of pEMJ7 were generated and ligated into the *Streptomyces* high-copy-number plasmid pIJ486. The transformants containing recombinants of pIJ486 were assessed for increased chitinase production by measuring the activity in the culture supernatant and by SDS-PAGE analysis of extracellular protein (Fig. 2). In recombinants of *S. lividans* containing pEMJ702, pEMJ703, pEMJ704 and pEMJ705, there was a significant increase in chitinase activity in the culture, and protein bands corresponding to chitinases C and D were observed. No subclones produced either chitinase C or chitinase D alone. These results implied that the coding region for chitinases C and D was on the *Sma*I–*Xho*I fragment (about 2 kb).

Analysis of transcriptional products encoded by the fragment

If chitinases C (65 kDa) and D (41 kDa) were encoded by different genes, the *Sma*I–*Xho*I fragment appeared too small to accommodate both genes. However, chitinase C and chitinase D could have been derived from different, overlapping ORFs on the same strand, or been encoded on opposite strands. To analyse mRNA transcribed from this region by Northern hybridization, we prepared single-stranded probes for each strand. Total RNA fractions were prepared from *S. lividans* containing pEMJ7, grown in the presence of chitin, chitin plus glucose, or without chitin as a control. When the single-stranded probe 712R, prepared from pEMM712R, was used, a single band was detected in RNA isolated from cells grown in the presence of chitin

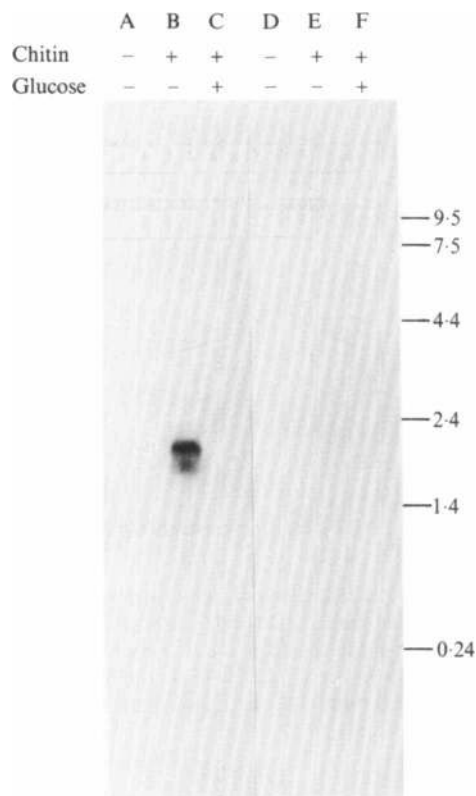


Fig. 3. Northern analysis of the mRNA transcribed from the insert DNA of pEMJ7. Experimental conditions are described in Methods. Inorganic salts medium containing yeast extract (1.5 mg ml⁻¹) and thiostrepton (50 μg ml⁻¹) was used as a basal medium. RNAs were prepared from cells grown in basal medium (lanes A and D), basal medium with chitin (lanes B and E), and basal medium with chitin plus glucose (lanes C and F), and subjected to 1.2% (w/v) agarose gel electrophoresis. Probe 712R (lanes A, B and C) and probe 712 (lanes D, E and F) were used for Northern hybridization. The numbers on the right side indicate the sizes (kb) of the standard RNA ladder.

(Fig. 3, lane B). The probe did not hybridize to RNA isolated from cells grown in the presence of glucose plus chitin, or in the absence of chitin (Fig. 3, lanes A and C). This showed that chitinase production was induced by chitin and repressed by glucose at the stage of transcription. In contrast, probe 712, which was complementary to probe 712R, did not bind to RNA extracted from cells grown under the above conditions (Fig. 3, lanes D, E and F). The RNA molecule (1.8 kb) that hybridized with probe 712R (Fig. 3, lane B) was too small to encode two distinct chitinases of 65 and 41 kDa, suggesting that both chitinases C and D were translated from a single mRNA transcribed from a strand homologous to the 712R probe. It was therefore likely that chitinase D was a proteolytic derivative of chitinase C. This was supported by the similarity between the enzymic properties (K_m ; pH optimum; degradation products of colloidal chitin; substrate specificity) of chitinases C and D (Miyashita *et al.*, 1991).

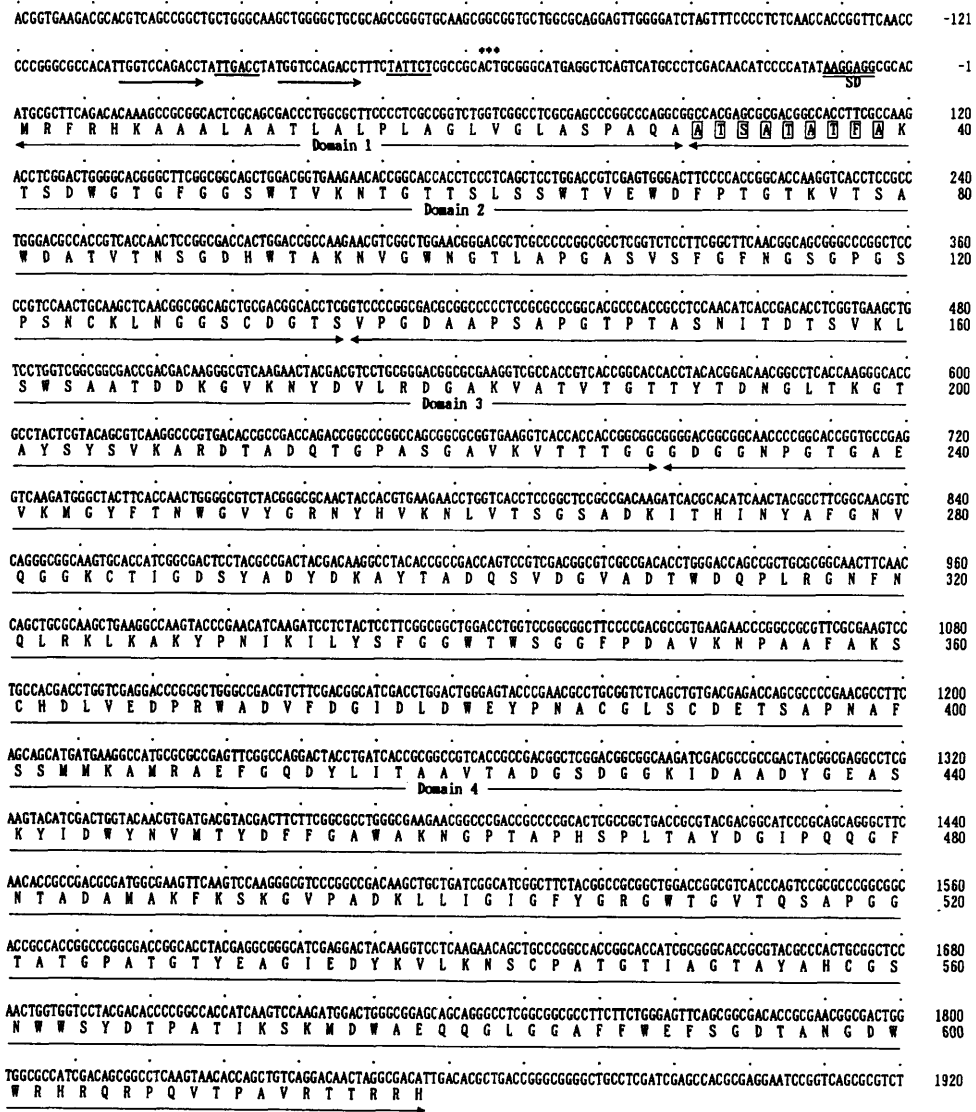


Fig. 4. Nucleotide sequence and deduced amino acid sequence of the *chiC* gene of *S. lividans*. Deduced amino acids are specified in single-letter code under the second nucleotide of each codon. Nucleotides and amino acids are numbered beginning at the start of the coding sequence. The Shine-Dalgarno sequence (SD) is double-underlined. The apparent transcription initiation sites determined by S1 mapping are indicated by asterisks. The -10 and -35 regions of a possible promoter sequence are underlined. The bold arrows indicate the twelve-nucleotide direct repeats in the promoter region. Nine amino acids which corresponded to the N-terminus of mature chitinase C are boxed. The regions of each domain are indicated by thin arrows under the amino acid sequence.

Sequencing

The DNA sequence of the coding region for chitinases C and D was determined by the dideoxy-termination method using single-stranded DNA prepared from various M13 subclones (Fig. 4). Two ORFs (ORF1 and ORF2), which were transcribed in opposite directions, were observed. The following data indicated that ORF1 encoded a precursor of chitinase C: ORF1 was complementary to mRNA detected in Northern hybridization;

the size of ORF1 (1857 nucleotides) was similar to that of the mRNA species (1900 nucleotides) which hybridized to probe 712R (Fig. 3); and in the deduced amino acid sequence, ORF1 encoded a 619 amino acid polypeptide with a molecular mass of 65 198 Da, which agreed with that of chitinase C as estimated by SDS-PAGE (Miyashita *et al.*, 1991). Moreover, ORF1 had a codon distribution typical of *Streptomyces*; the base composition of ORF1 was 68.8 mol% G + C, with more than 90% G or C at the third position of each codon

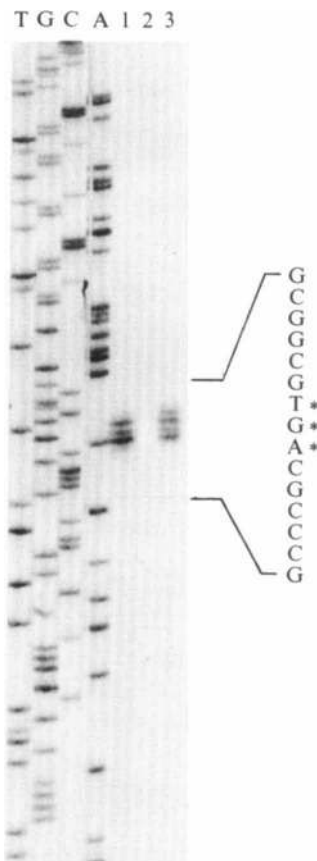


Fig. 5. Analysis of transcription initiation sites of *chiC* from *S. lividans*. A 5' terminus of the chitin-induced transcripts was identified by S1 mapping. RNA was isolated from the cells grown on chitin (lanes 1 and 3) or chitin plus glucose (lane 2). Non-hybridized strands were digested by 100 U (lanes 2 and 3) or 300 U (lane 1) of S1 nuclease. Letters above the lanes indicate the four separate dideoxynucleotide sequence reactions. Asterisks indicate the nucleotides at the transcription start sites.

(Bibb *et al.*, 1984). To confirm that ORF1 actually encoded both mature chitinases C and D, we determined the N-terminal amino acid sequence of chitinases C and D secreted from *S. lividans* TK64 containing pEMJ7 grown in inorganic salts colloidal chitin medium. The N-terminal amino acid sequence of mature chitinase C (Ala-Thr-Ser-Ala-Thr-Ala-Thr-Phe-Ala) was identical to that beginning at Ala 31 of the deduced amino acid sequence of ORF1 (Fig. 4). The N-terminal amino acid sequence of chitinase D could not be determined.

A putative ribosome-binding site complementary to the 3' end of the 16S subunit of rRNA of *S. lividans*, AAGGAGG (Bibb *et al.*, 1982), was located 5 bp upstream of the presumptive start codon of ORF1. The 30-amino-acid sequence from the ATG codon (Met) to Ala-30 exhibited characteristics typical of a signal sequence for peptide secretion (Perlman & Halvorson, 1983), i.e. a positively-charged hydrophilic N-terminus containing Arg-His-Lys residues followed by a long hydrophobic amino acid core sequence, and a serine residue located 5 amino acids upstream of the N-terminus of mature chitinase C, allowing a turn in the polypeptide which presumably exposes the cleavage site Ala-Gln-Ala to the signal peptidase. From these results, we concluded that ORF1 encoded a pre-protein of chitinase C (*chiC*).

Promoter region of the *chiC* gene

The transcription initiation site of *chiC* was determined by S1 mapping using a single-stranded DNA probe. Three major hybrids were located (Fig. 5), indicating that the *chiC* transcription start sites were at positions -56, -57 and -58, with respect to the putative initiation ATG codon (Fig. 4). The probe did not hybridize to RNA extracted from cells grown on chitin plus glucose, indicating that the expression of *chiC* was repressed by glucose. Sequences similar to the most common bacterial promoter consensus sequence were observed around -10 (TATTCT) and -35 (TTGACC) with respect to the apparent transcription initiation site. Two identical 12 bp direct repeat sequences (TGGTCCAGACCT) were present in the promoter region of *chiC* overlapping the putative -35 sequence (TTGACC). Delic *et al.* (1992) found similar direct repeats in the promoter region of chitinase genes (*chi63* and *chi35*) of *S. plucatus*, and suggested that these sequences were involved in the regulation of the expression of these chitinase genes. Similar direct repeat sequences were also observed in the promoter region of *chiA* of *S. lividans* (unpublished data).

The structure of chitinase C

Homology analysis of the deduced amino acid sequence of chitinase C precursor protein revealed that the enzyme

Fig. 6. Sequence similarity of various proteins to each domain of chitinase C of *S. lividans* (SliChiC). Numbers in each line are the positions of the first (left) and last (right) amino acids in the overall sequences of the polypeptides. Residues of each polypeptide identical to those of chitinase C are indicated by dots. Asterisks above the amino acid sequences indicate residues conserved in all polypeptides. Horizontal bars in the region of domain 4 indicate clusters of conserved amino acid residues among the polypeptides. Tandem repeats of fibronectin-type-III-like sequence from chitinase A1 of *B. circulans* and endoglucanase B from *C. fimi* are indicated by BciChi -R1, -R2, and CfiCenB-R1, -R2, -R3. The table below the figure summarizes the proteins compared and the homology observed.

had a multiple domain structure consisting of four discrete domains (Fig. 6).

Domain 1, consisting of the 30 N-terminal amino acids of the chitinase C pre-protein, showed 96% and 40% similarity to the signal sequence of chitinase-63 of *Streptomyces plicatus* (Robbins *et al.*, 1988) and endoglucanase A of *Cellulomonas fimi* (Wong *et al.*, 1986), respectively. The signal sequence processing sites of chitinase C and chitinase-63 were identical, both being between Ala-30 and Ala-31, while that of endoglucanase A was between Ala-31 and Ala-32. The three-amino-acid sequence Ala-Gln-Ala preceding the cleavage sites was conserved in these three proteins, which is consistent with the proposed signal peptidase recognition sequence Ala-X-Ala↓ (Perlman & Halvorson, 1983).

The 105 N-terminal residues of the mature secreted form of chitinase C (domain 2) were similar to cellulose-binding domains located in several cellulases and xylanases. The cellulose-binding domain was first identified in endo- and exoglucanases of *Cellulomonas fimi* (Gilkes *et al.*, 1988), and has since been found in several plant cell wall hydrolases (Gilbert *et al.*, 1990; Meinke *et al.*, 1991a). The ten amino acids which are conserved in the cellulose-binding domains of cellulases and xylanases were also present in domain 2 of chitinase C. A region from amino acids 98 to 120 of chitinase C showed particularly high similarity to cellulase and xylanase cellulose-binding domains. This is the first report of a bacterial chitinase whose sequence exhibits homology with cellulose-binding domains of β -1,4-glycanases. However, the role of domain 2 in chitin degradation activity remains unclear. Deletion experiments should clarify the function of this domain.

The sequence of 94 amino acids from amino acids 136 to 229 (domain 3) showed < 50% similarity to the fibronectin-type-III-like sequences of bacteria. Watanabe *et al.* (1990b) found that two tandem repeats of 95 amino acids in chitinase A1 from *B. circulans* are closely related to the original form of type III repeating units of fibronectin. Fibronectin-type-III-like sequences have also been found in endoglucanase B from *C. fimi* as three tandem repeats (Meinke *et al.*, 1991b). Homology analysis showed that domain 3 of chitinase C was similar not only to these sequences but also to amylase A-180

from an alkaliphilic eubacterium and poly(3-hydroxybutyrate) depolymerase from *Alcaligenes faecalis* (Candussio *et al.*, 1990; Saito *et al.*, 1989). Thirteen out of 93 amino acids (14.0%) were identical in these sequences. It is notable that the fibronectin-type-III-like sequences are widespread in eukaryotic and prokaryotic proteins, especially in bacterial carbohydrases. Although the fibronectin-type-III-like sequence was also found in chitinase A from *S. lividans* (unpublished data), some chitinases lack the sequence, implying that the sequence does not contribute to hydrolytic activity. The function of these sequences and the reasons for their wide distribution are not yet clear.

Domain 4, the longest domain (390 amino acids), at the C-terminus of chitinase C showed similarity to some proteins with chitinolytic activity. In particular, the entire region of domain 4 of chitinase C showed similarity to chitinase A1 from *Bacillus circulans* and chitinase A from *Serratia marcescens*. These similarities suggest that the enzymes are closely related. The identical amino acid sequences among these proteins tended to form clusters (Fig. 6), and clusters 2, 3 and 4 have amino acid sequences which have been identified in a wide range of chitinases, including class III chitinase from higher plants, and are thought to play an important role in the activity of chitinases (Watanabe *et al.*, 1992). Therefore, domain 4 was considered to be a catalytic domain of chitinase C.

Thus, Chitinase C appears to be composed of four discrete domains. The domains do not overlap: the homologous regions between chitinase C and each protein are restricted to the domains shown in Fig. 6. Only chitinase A1 from *B. circulans* and endoglucanases A and B from *C. fimi* showed similarity to chitinase C through two domains. Among the proteins which showed similarities to chitinase C, chitinase A1 of *B. circulans* and endoglucanase B of *C. fimi* are noteworthy. These two proteins also have multiple domain structures (Fig. 7). Besides a catalytic domain, which is related to domain 4 of chitinase C, chitinase A1 has two tandem repeats of fibronectin-type-III-like sequences and a chitin-binding domain. Endoglucanase B of *C. fimi* also consists of four distinct domains: a catalytic domain, *Bacillus subtilis* End1-like sequences, three tandem repeats of fibronectin-

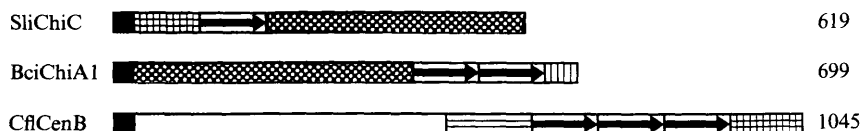


Fig. 7. Sequence organization in proteins containing sequences similar to those in chitinase C. Designations for the enzymes are given in Fig. 6. ■, signal peptide; ▨, cellulose-binding domain; ▩, fibronectin type III-like sequence; ▧, catalytic domain of chitinases; ▤, chitin-binding domain; □, catalytic domain of cellulose; ▥, *B. subtilis* End1-like sequence. Numbers indicate total amino acid residues of the polypeptides.

type-III-like sequences, and the cellulose-binding domain whose sequence is related to domain 2 of chitinase C. Many cellulases have been reported to have one or some of these domains in unique arrangements (Meinke *et al.*, 191*b*). Another chitinase of *B. circulans*, chitinase D, also has one copy of fibronectin-type-III-like sequence and a chitin-binding domain, but they are aligned differently than those in chitinase A1 (Watanabe *et al.*, 1992). Thus, many carbohydrate-degrading enzymes share some common homologous domains, but the arrangements of these domains differ, suggesting that these sequences may have been shuffled during the evolution of these enzymes.

We thank Prof. D. A. Hopwood for his suggestions regarding ORF analysis. This work was supported in part by a Grant-in-Aid (Bio Media Program) from the Ministry of Agriculture, Forestry and Fisheries of Japan (BMP 92-III-2-6).

References

- AUSUBEL, F. M., BRENT, R., KINGSTON, R. E., MOORE, D. D., SEIDMAN, J. G., SMITH, J. A. & STRUHL, K. (1989). *Current Protocols in Molecular Biology*. New York: John Wiley.
- BERGER, E., JONES, W. A., JONES, D. T. & WOODS, D. R. (1989). Cloning and sequencing of an endoglucanase (*end1*) gene from *Butyrivibrio fibrisolvens* H17c. *Molecular and General Genetics* **219**, 193–198.
- BERGER, L. R. & REYNOLDS, D. M. (1958). The chitinase system of a strain of *Streptomyces griseus*. *Biochimica et Biophysica Acta* **29**, 522–534.
- BIBB, M. J. & COHEN, S. N. (1982). Gene expression in *Streptomyces*: construction and application of promoter-probe vectors in *Streptomyces lividans*. *Molecular and General Genetics* **187**, 265–277.
- BIBB, M. J., FINDLAY, P. R. & JOHNSON, M. W. (1984). The relationship between base composition and codon usage in bacterial genes and its use for the simple and reliable identification of protein-coding sequences. *Gene* **30**, 157–166.
- BROGLIE, K. E., GAYNOR, J. J. & BROGLIE, R. M. (1986). Ethylene-regulated gene expression: molecular cloning of the genes encoding an endochitinase from *Phaseolus vulgaris*. *Proceedings of the National Academy of Sciences of the United States of America* **83**, 6820–6824.
- CANDUSSIO, A., SCHMID, G. & BOCK, A. (1990). Biochemical and genetic analysis of a maltopentaose-producing amylase from an alkaliphilic Gram-positive bacterium. *European Journal of Biochemistry* **191**, 177–185.
- DELIC, I., ROBBINS, P. & WESTPHELING, J. (1992). Direct repeat sequences are implicated in the regulation of two *Streptomyces* chitinase promoters that are subject to carbon catabolite control. *Proceedings of the National Academy of Sciences of the United States of America* **89**, 1885–1889.
- GILBERT, H. J., HALL, J., HAZLEWOOD, G. P. & FERREIRA, L. M. A. (1990). The N-terminal region of an endoglucanase from *Pseudomonas fluorescens* subspecies *cellulosa* constitutes a cellulose-binding domain that is distinct from the catalytic centre. *Molecular Microbiology* **4**, 759–767.
- GILKES, N. R., WARREN, R. A. J., MILLER, R. C., JR & KILBURN, D. G. (1988). Precise excision of the cellulose binding domains from two *Cellulomonas fimi* cellulases by a homologous protease and the effect on catalysis. *Journal of Biological Chemistry* **263**, 10401–10407.
- HALL, J. & GILBERT, H. J. (1988). The nucleotide sequence of a carboxymethylcellulase gene from *Pseudomonas fluorescens* subsp. *cellulosa*. *Molecular and General Genetics* **213**, 112–117.
- HALL, J., HAZLEWOOD, G. P., HUSKISSON, N. S., DURRANT, A. J. & GILBERT, H. J. (1989). Conserved serine-rich sequences in xylanase and cellulase from *Pseudomonas fluorescens* subspecies *cellulosa*: internal signal sequence and unusual protein processing. *Molecular Microbiology* **3**, 1211–1219.
- HARPSTER, M. H. & DUNSMUIR, P. (1989). Nucleotide sequence of the chitinase B gene of *Serratia marcescens* QMB1466. *Nucleic Acids Research* **17**, 5395–5395.
- HERGET, T., SCHELL, J. & SCHREIER, P. H. (1990). Elicitor-specific induction of one member of the chitinase gene family in *Arachis hypogaea*. *Molecular and General Genetics* **224**, 469–476.
- HOPWOOD, D. A., BIBB, M. J., CHATER, K. F., KIESER, T., BRUNTON, C. J., KIESER, H. M., LYDIATE, D. L., SMITH, C. P., WARD, J. M. & SCHREMPF, H. S. (1985). *Genetic Manipulation of Streptomyces – a Laboratory Manual*. Norwich, UK: John Innes Foundation.
- JONES, J. D. G., GRADY, K. L., SUSLOW, T. V. & BEDBROOK, J. R. (1986). Isolation and characterization of genes encoding two chitinase enzymes from *Serratia marcescens*. *EMBO Journal* **5**, 467–473.
- KAMEI, K., YAMAMURA, Y., HARA, S. & IKENAKA, T. (1989). Amino acid sequence of chitinase from *Streptomyces erythraeus*. *Journal of Biochemistry* **105**, 979–985.
- KURANDA, M. J. & ROBBINS, P. W. (1991). Chitinase is required for cell separation during growth of *Saccharomyces cerevisiae*. *Journal of Biological Chemistry* **266**, 19758–19767.
- LAEMMLI, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature, London* **227**, 680–685.
- LAFLAMME, D. & ROXBY, R. (1989). Isolation and nucleotide sequence of cDNA clones encoding potato chitinase genes. *Plant Molecular Biology* **13**, 249–250.
- LEGRAND, M., KAUFFMANN, S., GEOFFROY, P. & FRITIG, B. (1987). Biological function of pathogenesis-related proteins: four tobacco pathogenesis-related proteins are chitinases. *Proceedings of the National Academy of Sciences of the United States of America* **84**, 6750–6754.
- MEINKE, A., BRAUN, C., GILKES, N. R., KILBURN, D. G., MILLER, R. C., JR & WARREN, R. A. J. (1991*a*). Unusual sequence organization in *CenB*, an inverting endoglucanase from *Cellulomonas fimi*. *Journal of Bacteriology* **173**, 308–314.
- MEINKE, A., GILKES, N. R., KILBURN, D. G., MILLER, R. C., JR & WARREN, R. A. J. (1991*b*). Multiple domains in endoglucanase B (*CenB*) from *Cellulomonas fimi*: functions and relatedness to domains in other polypeptides. *Journal of Bacteriology* **173**, 7126–7135.
- METRAUX, J. P., BURKHART, W., MOYER, M., DINCHER, S., MIDDLESTEADT, W., WILLIAMS, S., PAYNE, G., CARNES, M. & RYALS, J. (1989). Isolation of a complementary DNA encoding a chitinase with structural homology to a bifunctional lysozyme/chitinase. *Proceedings of the National Academy of Sciences of the United States of America* **86**, 896–900.
- MIYASHITA, K., FUJII, T. & SAWADA, Y. (1991). Molecular cloning and characterization of chitinase genes from *Streptomyces lividans* 66. *Journal of General Microbiology* **137**, 2065–2072.
- MONREAL, J. & REESE, E. T. (1969). The chitinase of *Serratia marcescens*. *Canadian Journal of Microbiology* **15**, 689–696.
- O'NEILL, G., GOH, S. H., WARREN, R. A. J., KILBURN, D. G. & MILLER, R. C., JR. (1986). Structure of the gene encoding the exoglucanase of *Cellulomonas fimi*. *Gene* **44**, 325–330.
- PARSONS, T. J., BRADSHAW, H. D., JR & GORDON, M. P. (1989). Systemic accumulation of specific mRNAs in response to wounding in poplar trees. *Proceedings of the National Academy of Sciences of the United States of America* **86**, 7895–7899.
- PASTAN, I. & ADHYA, S. (1976). Cyclin adenosine 5'-monophosphate in *Escherichia coli*. *Bacteriological Reviews* **40**, 527–551.
- PAYNE, G., AHL, P., MOYER, M., HARPER, A., BECK, J., MEINS, F., JR & RYALS, J. (1990). Isolation of complementary DNA clones encoding pathogenesis-related proteins P and Q, two acidic chitinases from tobacco. *Proceedings of the National Academy of Sciences of the United States of America* **87**, 98–102.
- PERLMAN, D. & HALVORSON, H. O. (1983). A putative signal peptidase

- recognition site and sequence in eukaryotic and prokaryotic signal peptides. *Journal of Molecular Biology* **167**, 391–409.
- ROBBINS, P. W., ALBRIGHT, C. & BENFIELD, B. (1988). Cloning and expression of a *Streptomyces plicatus* chitinase (chitinase-63) in *Escherichia coli*. *Journal of Biological Chemistry* **263**, 443–447.
- SAITO, T., SUZUKI, K., YAMAMOTO, J., FUKUI, T., MIWA, K., TOMITA, K., NAKANISHI, S., ODANI, S., SUZUKI, J. & ISHIKAWA, K. (1989). Cloning, nucleotide sequence, and expression in *Escherichia coli* of the gene for poly(3-hydroxybutyrate) depolymerase from *Alcaligenes faecalis*. *Journal of Bacteriology* **171**, 184–189.
- SAMAC, D. A., HIRONAKA, C. M., YALLALY, P. E. & SHAH, D. M. (1990). Isolation and characterization of the genes encoding basic and acidic chitinase in *Arabidopsis thaliana*. *Plant Physiology* **93**, 907–914.
- SAMBROOK, J., FRITSCH, E. F. & MANIATIS, T. (1989). *Molecular Cloning: a Laboratory Manual*, 2nd edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
- STARK, M. J. R., MILEHAM, A. J., ROMANOS, M. A. & BOYD, A. (1984). Nucleotide sequence and transcription analysis of a linear DNA plasmid associated with the killer character of the yeast *Kluyveromyces lactis*. *Nucleic Acids Research* **12**, 6011–6030.
- UENO, H., MIYASHITA, K., SAWADA, Y. & OBA, Y. (1990). Purification and some properties of extracellular chitinases from *Streptomyces* sp.S-84. *Journal of General and Applied Microbiology* **36**, 377–392.
- WATANABE, T., OYANAGI, W., SUZUKI, K. & TANAKA, H. (1990a). Chitinase system of *Bacillus circulans* WL-12 and importance of chitinase A1 in chitin degradation. *Journal of Bacteriology* **172**, 4017–4022.
- WATANABE, T., SUZUKI, K., OYANAGI, W., OHNISHI, K. & TANAKA, H. (1990b). Gene cloning of chitinase A1 from *Bacillus circulans* WL-12 revealed its evolutionary relationship to *Serratia* chitinase and to the type III homology units of fibronectin. *Journal of Biological Chemistry* **265**, 15659–15665.
- WATANABE, T., OYANAGI, W., SUZUKI, K., OHNISHI, K. & TANAKA, H. (1992). Structure of the gene encoding chitinase D of *Bacillus circulans* WL-12 and possible homology of the enzyme to other prokaryotic chitinases and class III plant chitinase. *Journal of Bacteriology* **174**, 408–414.
- WONG, W. K. R., GERHARD, B., GUO, Z. M., KILBURN, D. G., WARREN, R. A. J. & MILLER, R. C., JR. (1986). Characterization and structure of an endoglucanase gene *cenA* of *Cellulomonas fimi*. *Gene* **44**, 315–324.
- YANISCH-PERRON, C., VIEIRA, J. & MESSING, J. (1985). Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **33**, 103–119.
- ZHU, Q. & LAMB, C. J. (1991). Isolation and characterization of a rice gene encoding a basic chitinase. *Molecular and General Genetics* **226**, 289–296.